

Reaction Engineering for Protein Modification: Tools for Chemistry and Biology



**A thesis submitted in partial fulfillment of the degree of
Doctor of Philosophy at the University of Oxford**

**Justin M. Chalker
New College, Oxford
Trinity Term 2011**

Copyright of a thesis from the University of Oxford is reserved solely for the thesis author. Reproduction of this thesis in any form requires the consent of the author.

© Justin M. Chalker

2011

Short Abstract for Bodleian Library Indexing

Reaction Engineering for Protein Modification: Tools for Chemistry and Biology

Justin M. Chalker
New College

University of Oxford

Submitted for the degree of Doctor of Philosophy in Organic Chemistry
Trinity Term 2011

Chemical modification of proteins is critical for many areas of biochemistry and medicine. Several methods for site-selective protein modification are reported in this Thesis that are useful in accessing both natural and artificial protein architectures. Multiple, complementary methods for the conversion of cysteine to dehydroalanine are described. Dehydroalanine is used as a general precursor to several post-translational modifications and glycosylation, polyprenylation, phosphorylation, and lysine methylation and acetylation are all accessible. These modifications and their mimics were explored on multiple proteins, including histone proteins. Unnatural modifications were also explored. The first examples of olefin metathesis and Suzuki-Miyaura cross-coupling on protein substrates are reported. Allyl sulfides were discovered to be remarkably reactive substrates in olefin metathesis, allowing use of this reaction in water and on proteins. For Suzuki-Miyaura cross-coupling, a new catalyst is described that is fully compatible with proteins. Both olefin metathesis and cross-coupling allow the formation of carbon-carbon bonds on proteins. The prospects of these transformations in chemical biology are discussed. Finally, a novel strategy is reported for the installation of natural, unnatural, and post-translationally modified amino acid residues on proteins. This technology relies on addition of carbon radicals to dehydroalanine. This method of “chemical mutagenesis” is anticipated to complement standard genetic manipulation of protein structure.

Full Abstract

Reaction Engineering for Protein Modification: Tools for Chemistry and Biology

Justin M. Chalker
New College
University of Oxford

Submitted for the degree of Doctor of Philosophy in Organic Chemistry
Trinity Term 2011

The chemical modification of proteins is widely used in biochemistry, medicine, and synthesis. Selective modification at a desired residue, however, is a challenging task. The diverse functionality of polypeptides restricts the variety of reaction types amenable to protein modification. Moreover, the requirement for buffered aqueous media and ambient temperatures adds an additional layer of complexity to such reactions. The limited reaction set used for traditional protein modification—based largely on the reaction of electrophiles with cysteine and lysine—would be complemented by reactions that target other amino acid residues. This Thesis details efforts to bring the precision of chemistry to bear on the problem of selective protein modification. Novel reactions, methods, and materials are described herein that address these challenges. Thesis topics are summarized below:

Methods for Converting Cysteine to Dehydroalanine on Peptides and Proteins

Cysteine is one of the most useful residues for selective protein modification. Typical modifications involve alkylation of the nucleophilic side-chain. Cysteine, however, can be envisioned as a latent *electrophile* after its selective conversion to dehydroalanine. Multiple methods for the conversion of cysteine to dehydroalanine are described. These reactions are complementary, allowing strategic selection of the most appropriate method for a given substrate. These transformations are demonstrated on peptides and proteins, with scope and limitations described. The subsequent reaction of dehydroalanine with thiol nucleophiles allows *specific* access to cysteine derivatives.

Dehydroalanine is a precursor to post-translational modifications

The addition of thiols to dehydroalanine was used to install several natural post-translational modifications, or their mimics, on protein surfaces. Glycosylation, prenylation, phosphorylation, and lysine methylation and acetylation are all accessible through dehydroalanine. These modifications were explored on multiple proteins. One class of proteins, histones, bear several of these modifications. Histones are key components of chromatin and their post-translational modifications are critical regulators of chromatin structure and function. Synthetic access to modified histones is required to study and understand chromatin biology. The synthesis of modified histones using dehydroalanine is described. These modified histones were used in binding studies and as substrates for histone modifying enzymes.

Aqueous Olefin Metathesis and Its Use in Protein Modification

In the course of the dehydroalanine work described above, access to proteins containing *S*-allyl cysteine was made possible by the nucleophilic addition of allyl thiol. It was

discovered that *S*-allyl cysteine is highly active in olefin metathesis, allowing the first instance of this reaction on proteins. A sulfur-relay mechanism is proposed that explains why the sulfur of *S*-allyl cysteine is not only tolerated in olefin metathesis, but *enhances* the reaction rate. The success of this reaction on proteins has also motivated the development of multiple, complementary methods for the chemical incorporation of *S*-allyl cysteine. Among these methods, the first *specific* allylation of cysteine on protein substrates is described. The allylation is based on the rearrangement of the allyl selenenylsulfide of cysteine followed by spontaneous loss of selenium. Access to *S*-allyl cysteine, olefin metathesis at this residue, and prospects of this method are discussed.

Cross-Coupling for Protein Modification

For nearly 10 years, the genetic incorporation of amino acids suitable for cross-coupling has been possible and has been demonstrated in several proteins. Nevertheless, modification of these residues remains an outstanding problem. Described in this Thesis is the discovery of an active, water-soluble Suzuki-Miyaura cross-coupling catalyst and its use on protein substrates. A variety of biaryl domains were installed on a protein surface. Such structures hold potential in fluorogenic labeling of proteins. Moreover, the borate confers water-solubility to an otherwise hydrophobic substrate, allowing site-specific lipidation of proteins without the use of organic solvent. The scope of the catalyst in general synthesis is also described. As little as 1 ppm of this palladium catalyst can catalyze cross-coupling in water without the need for organic solvent. Moreover, hindered *ortho*-substituted couplings are accommodated by this catalyst.

Dehydroalanine: A Handle for the “Chemical Mutagenesis” of Proteins

Versatile access to dehydroalanine on proteins allows exploration of more ambitious prospects in protein modification. The addition of thiol nucleophiles to dehydroalanine provides a useful route to alkylated cysteine residues. The addition of *carbon nucleophiles*, however, dramatically expands the potential of dehydroalanine in protein modification. Specifically, the addition of carbon radicals to dehydroalanine allows its chemical conversion into another residue. This “chemical mutagenesis” is explored as a method to install natural, unnatural, and post-translationally modified amino acids on proteins.

Table of Contents

Preface	xiv
Publication List	xvi
Selection of Abbreviations	xviii
 Chapter 1: An Introduction to the Chemical Modification of Proteins	
 1.1 Introduction	 2
 1.2 Chemical Modification of Proteins at Natural Residues	 2
Lysine modification	3
Glutamate and aspartate modification	5
Cysteine modification	5
Serine modification	8
Tyrosine modification	9
Tryptophan modification	10
 1.3 Chemical Modification of Proteins at Unnatural Residues	 10
Methods for the incorporation of unnatural amino acids into proteins	11
Modification of proteins containing azides and alkynes	15
Modification of proteins containing ketones and aldehydes	16
Modification of proteins containing alkenes	16
Modification of proteins containing aryl halides	18
Miscellaneous methods of protein modification using unnatural amino acids	18
 1.4 Total Chemical Synthesis of Proteins	 19
Native chemical ligation	20
Expressed protein ligation	21
Additional methods of peptide ligation	23
Ligation at residues other than cysteine	23
 1.5 Thesis Objectives	 25

1.6 References	27
-----------------------	----

Chapter 2: Methods for the Chemical Conversion of Cysteine to Dehydroalanine on Protein Substrates

2.1 Introduction	45
-------------------------	----

Dehydroalanine: a precursor to post-translational modifications	45
Previous reports describing the synthesis of dehydroalanine in peptides and proteins	47
Proposed methods for the conversion of cysteine to dehydroalanine	50

2.2 Results and Discussion

Conversion of cysteine to dehydroalanine on amino acid and peptide substrates	52
Conversion of cysteine to dehydroalanine under reducing conditions	52
Direct conversion of cysteine disulfides to dehydroalanine with base	53
Oxidative elimination of cysteine to dehydroalanine	54
Reaction of <i>O</i> -mesitylenesulfonylhydroxylamine (MSH) with functionalized amino acids	56
Bis-alkylation-elimination of cysteine to dehydroalanine	59
Conversion of cysteine to dehydroalanine on proteins	63
Reduction-elimination of cysteine disulfides to dehydroalanine	63
Base-mediated elimination of cysteine disulfides to dehydroalanine	64
Oxidative elimination of cysteine to dehydroalanine using MSH	64
Bis-alkylation-elimination of cysteine to dehydroalanine	68

2.3 Conclusions and Outlook	71
------------------------------------	----

2.4 References	72
-----------------------	----

2.5 Experimental Procedures	78
------------------------------------	----

General considerations	78
Reductive route to Dha using HMPT	80
Base-mediated conversion of cysteine to dehydroalanine	86
MSH in the oxidative elimination of cysteine to dehydroalanine	87
Assessment of reactivity of MSH with functionalized amino acids: substrate synthesis	92

MSH reactivity with functionalized amino acids	102
Bis-alkylation-elimination in the conversion of cysteine to dehydroalanine	113
Protein modification	129
Protein mass spectrometry	129
Reduction-elimination of cysteine to dehydroalanine: attempts on SBL	131
Attempted base-mediated conversion of cysteine disulfides to dehydroalanine on SBL	134
Synthesis of Dha on SBL-S156C using MSH	135
Synthesis of GlcNAc cysteine on SBL at position 156	137
Enzyme kinetics for modified SBL	138
Reaction of Np276 with MSH	139
Protein reactions using dialkylation-elimination of cysteine to dehydroalanine	141
2.6 References for Experimental Procedures	152
2.7 NMR Spectra	153
 Chapter 3: Conversion of Cysteine to Dehydroalanine Allows Access to Proteins Bearing Post-Translational Modifications	
 3.1 Introduction	184
Post-translational modification of proteins	184
Methods for the synthesis of glycoproteins	185
Glycoprotein synthesis by native chemical ligation	185
Glycoprotein synthesis using recombinant expression systems	187
Glycoprotein synthesis using enzymatic glycosylation	188
Glycoprotein synthesis using site-selective chemical glycosylation	188
Methods for the synthesis of lipidated proteins	190
Lipidated proteins by expressed protein ligation	190
Lipidated proteins by site-selective chemical modification	191
Methods for the synthesis of phosphorylated proteins	192
Synthesis of phosphorylated proteins by native chemical ligation	192
Synthesis of phosphorylated proteins by site-selective chemical modification	195
Methods for the synthesis of proteins containing methylated and acetylated lysine residues	195
Synthesis of proteins containing methylated and acetylated lysine by NCL and EPL	196
Genetic incorporation of methylated and acetylated lysine residues into proteins	197

Site-selective chemical installation of methylated and acetylated lysine residues	199
Dehydroalanine as a precursor to post-translational modifications	200
3.2 Results and Discussion	201
Thiol addition to dehydroalanine provides post-translationally modified proteins	201
Dehydroalanine is a precursor to <i>S</i> -linked glycoproteins	202
Dehydroalanine is a precursor to polyprenylated proteins	204
Dehydroalanine is a precursor to phosphocysteine	205
Dehydroalanine is a precursor to methylated lysine mimics	206
Alkyl cysteine modifications are stable, but not necessarily permanent	207
3.3 Conclusions and Outlook	209
3.4 References	210
3.5 Experimental Procedures	222
Synthesis of reagents	222
Protein modification	230
Installation of Dha on SBL using MSH	230
Synthesis of GlcNAc cysteine on SBL at position 156	231
Np276 modification and conversion to a glycoprotein	232
Synthesis of farnesyl cysteine on SBL at position 156	235
Synthesis of monomethyllysine mimic on SBL at position 156	236
Synthesis of dimethyllysine mimic on SBL at position 156	237
Synthesis of trimethyllysine mimic on SBL at position 156	238
Stability of GlcNAc cysteine on SBL in presence of glutathione	238
Reduction of disulfide-linked glycoprotein with glutathione	240
Conversion of an alkyl cysteine to dehydroalanine on SBL at position 156	241
3.6 References for Experimental Procedures	243
3.7 NMR Spectra	243

Chapter 4: The Conversion of Cysteine to Dehydroalanine Enables Access to Histones Bearing Post-Translational Modifications

4.1 Introduction	250
Chromatin structure and the histone code hypothesis	250
Evaluating the histone code hypothesis	253
Synthesis of histones bearing post-translational modifications	254
Histone synthesis by native chemical ligation and expressed protein ligation	254
Site-selective chemical modification of histones	258
Accessing modified histones using translation systems	259
Dehydroalanine as a precursor to post-translationally modified histones	261
4.2 Results and Discussion	262
Histone modification using <i>O</i> -mesitylenesulfonylhydroxylamine (MSH)	262
Histone modification using a bis-alkylation-elimination of cysteine	265
Synthesis of methylated and acetylated lysine mimics	266
Synthesis of a phosphorylated serine mimic on histone H3 <i>via</i> dehydroalanine	268
Synthesis of a glycosylated serine mimic on histone H3 <i>via</i> dehydroalanine	270
The use of histone PTM mimics in the study of histone modifying enzymes	271
4.3 Conclusions and Outlook	273
4.4 References	276
4.5 Experimental Procedures	284
Reagent synthesis	284
Histone sequences	292
Mass spectrometry: correlation of relative peak height to relative amount of protein	293
Histone reactions using MSH	296
Histone modification using a bis-alkylation-elimination of cysteine to dehydroalanine	303
Installation of an acetyllysine mimic at position 9 of histone 3	308
Installation of a monomethyllysine mimic at position 9 of histone 3	309
Installation of a dimethyllysine mimic at position 9 of histone 3	310
Installation of a trimethyllysine mimic at position 9 of histone 3	311
Western blots for methylated and acetylated lysine mimics	312
Synthesis of a phosphoserine mimic at position 10 of histone 3	316

Western blot for phosphoserine mimic at position 10 of histone 3	319
Synthesis of GlcNAc serine mimic at position 10 of histone 3	321
Histone deacetylase assay using acetyllysine mimic	323

4.6 References for Experimental Procedures 329

Chapter 5: Allyl Sulfides are Reactive Substrates in Olefin Metathesis and Enable Cross-Metathesis in Water and on Proteins

5.1 Introduction 331

Ring closing metathesis on amino acids and peptides	332
Cross-metathesis on peptides	334
Olefin metathesis in water	335
Olefin metathesis on proteins	337

5.2 Results and Discussion 338

Amino acid models	338
Discovery of allyl sulfides as reactive substrates in olefin metathesis	338
Olefin metathesis on proteins at <i>S</i> -allyl cysteine	342
Complementary methods for incorporating <i>S</i> -allyl cysteine into proteins	345
Allylation of cysteine with allyl chloride	345
Allylation of cysteine with allyl selenocyanate	346

5.3 Conclusion and Outlook 350

5.4 References 352

5.5 Experimental Procedures 359

Model substrate synthesis	359
Cross-metathesis on amino acids and other model substrates	362
Reagent synthesis for protein modification	365
Protein modification by olefin metathesis	373
Installation of <i>S</i> -allyl cysteine <i>via</i> dehydroalanine	373
Cross-metathesis at <i>S</i> -allyl cysteine made using dehydroalanine	374

Peptidase activity assay of protein modified by olefin metathesis	376
Cross-metathesis on proteins to give glycosylated and PEGylated proteins	377
Installation of <i>S</i> -allyl cysteine using allyl chloride	381
Cross-metathesis at <i>S</i> -allyl cysteine made using allyl chloride	384
Attempted synthesis of <i>S</i> -allyl cysteine from the allyl disulfide of cysteine	387
Installation of <i>S</i> -allyl cysteine using allyl selenocyanate	389
Cross-metathesis at <i>S</i> -allyl cysteine made using allyl selenocyanate	392
Peptidase activity assay of protein modified by olefin metathesis	393
Allylation of GFP using allyl selenocyanate	395
Attempted olefin metathesis on GFP	397

5.6 References for Experimental Procedures 398

5.7 X-Ray Crystallographic Data and NMR Spectra 398

Chapter 6: A convenient Catalyst for Aqueous Suzuki-Miyaura Cross-Coupling and Its Use in Covalent Protein Modification

6.1 Introduction 423

Suzuki-Miyaura coupling in water	423
Palladium-mediated coupling on peptides and proteins	424

6.2 Results and Discussion 429

The discovery of a pyrimidine ligand for Suzuki-Miyaura coupling in water	429
Suzuki-Miyaura coupling on amino acid and peptide substrates	431
Suzuki-Miyaura coupling on protein substrates	432
Scope and limitations of palladium-pyrimidine catalyst in synthesis	435

6.3 Conclusions and Outlook 437

6.4 References 438

6.5 Experimental Procedures 445

Synthesis of cross-coupling substrates	445
--	-----

Model cross-couplings on amino acid substrates	459
Cross-couplings in general biaryl synthesis	469
Cross-coupling on protein substrates	481
Control experiment for cross-coupling on proteins	491
6.6 References for Experimental Procedures	495
6.7 NMR Spectra	496
 Chapter 7: A Safe and Scalable Preparation of Barluenga's Reagent (IPy₂BF₄)	
7.1 Introduction	526
Barluenga's reagent in synthesis	526
Original synthesis of Barluenga's reagent	526
7.2 Results and Discussion	528
Revised synthesis of Barluenga's reagent	528
Iodination of tyrosine with Barluenga's reagent	529
7.3 Conclusions	530
7.4 References	530
7.5 Experimental Procedures	532
Synthesis of Barluenga's reagent	532
Iodination of tyrosine	537
7.6 References for Experimental Procedures	538
7.7 X-Ray Crystallographic Data and NMR Spectra	538

**Chapter 8: Post-Expression Mutagenesis of Proteins:
A Strategy for the Installation of Natural, Unnatural,
and Post-Translationally Modified Amino Acid Residues**

8.1 Introduction	550
Chemical mutagenesis: a milestone in protein science	551
Dual use of genetic and chemical mutations to investigate enzyme catalysis	554
Chemical mutations pertaining to noncanonical amino acids	555
Chemical mutations post-native chemical ligation	557
Chemical mutations to access natural residues bearing post-translational modifications	558
Towards a general chemical mutagenesis	559
8.2 Results and Discussion	561
Chemical mutation of dehydroalanine to a histidine analog	561
Chemical mutation of dehydroalanine to alanine	561
Model studies for carbon-carbon bond formation at dehydroalanine	562
Zn-mediated addition of alkyl iodides to dehydroalanine on a protein	567
8.3 Conclusions and Outlook	572
8.4 References	573
8.5 Experimental Procedures	582
Synthesis of reagents	582
Model studies on the conversion of dehydroalanine to other amino acids	589
Protein modification	607
Chemical mutation of dehydroalanine to histidine analog	609
Chemical mutation of dehydroalanine to alanine	610
Zn-mediated addition of radicals to dehydroalanine on proteins	611
8.6 References for Experimental Procedures	635
8.7 NMR Spectra	635

Preface

This Thesis is the product of nearly five years of research at the University of Oxford Chemistry Research Laboratory. The length of my course is somewhat unusual for an Oxford D.Phil. in Chemistry and I thank my advisor, Prof. Benjamin G. Davis, for supporting my decision to complete such an extended stay. I am extremely grateful for his enthusiasm and unwavering support throughout my time in Oxford. I very much value the freedom I had in my research, especially the opportunity to take on several projects in parallel. During my D.Phil. course I also served as a Lecturer in Organic Chemistry for both Pembroke College and St. Edmund Hall from 2007 to 2010. I thank Prof. Davis for allowing me to pursue my passion for teaching during these three years.

I could not have completed my D.Phil. without the generous financial support of the Rhodes Trust (2006-2009) and the U.S. National Science Foundation (2009-2011). These funding bodies allowed me to pursue and complete an ambitious body of research, unfettered by financial restraints. Their support also afforded me the opportunity to share my research at several international conferences over the last five years. Their investment in science and education should be commended.

The research in this Thesis was greatly facilitated by the excellent service staff at the Chemistry Research Laboratory. I am in their debt. In particular, I thank the NMR, mass spectrometry, and chemical crystallography staff for their valuable assistance. I am especially grateful to Dr. Amber Thompson for X-ray crystallographic analysis. All crystal structures in this Thesis were solved by Dr. Thompson. All crystallographic data included in this Thesis is for reference only and full credit should be attributed to Dr. Thompson.

Next, I would like to comment briefly on the contents and organization of this Thesis. First, several materials are used in multiple Chapters. The same unique number is used to identify the material in all Chapters it appears. At the end of each Chapter, and consistent with the requirements for most chemistry journals, NMR spectra for all new or key compounds are provided. Where compounds are used in more than one Chapter, the NMR is included only in the Chapter of its first occurrence. The preparation of each compound, however, is included in the Experimental Procedures in all Chapters in which it appears. While somewhat redundant, including the same preparation multiple times ensures that each Chapter can be read, understood, and repeated independently. Citations in the body of each Chapter are provided immediately after the concluding remarks and

are listed with the full title for ease of reference. Publications cited in the Experimental Procedures are provided separately, without titles, at the end of each experimental section. Finally, all protein graphics were prepared using MacPyMol (www.pymol.org) and the use of this open source software is acknowledged here.

Portions of this Thesis are collaborative efforts. I have taken great care to attribute the appropriate credit in these cases. Where assays or reactions were assisted or carried out by a colleague, their contribution is stated explicitly in the text and also in the respective experimental section. I thank all of these co-workers; their names are listed in the publications in the next section. I also thank Dr. Nathan Rose and Prof. Christopher Schofield for their support in a recent and promising collaboration in histone modification. I would also like to acknowledge three talented and dedicated Part II undergraduate researchers that I mentored in during my D.Phil.: Yuya Angel Lin, Stefanie Gerstberger, and Charlotte Wood. I strongly encourage readers to consult these three Master's Theses for further elaboration of several discoveries described in this Dissertation. It was a true pleasure to work with Yuya Angel Lin (Keble College 2007-2008) on aqueous olefin-metathesis, a collaboration that has continued into her D.Phil. Stefanie Gerstberger (Pembroke College 2008-2009) was an integral part of the dehydroalanine project and Charlotte Wood (Christ Church 2008-2009) carried out several key experiments in palladium-mediated protein modification; I thank them for their hard work.

A warm thanks goes to Stefanie Gerstberger and Inga Pfeffer for their kind assistance in translating several German articles that are quoted or cited in throughout this Dissertation. I also extend my gratitude to Smita Gunnoo, Jonathan Goodfellow, Lukas Lercher, Yuya Lin, Christopher Spicer, Mitul Patel, and Ben Davis for their critical reading of this Thesis. Their time, consideration, and advice are greatly appreciated. Finally, I would like to thank my family, friends, and loved ones for their patience, understanding, and encouragement throughout my D.Phil.



Justin Chalker

June 10th, 2011

University of Oxford

Chemistry Research Laboratory

Publication List

Selections of this Thesis have been featured in the following publications:

17. Conversion of Cysteine to Dehydroalanine Enables Access to Histones Bearing Diverse Post-Translational Modifications. Chalker, J.M.; Lercher, L.; Rose, N.R.; Schofield, C.J.; Davis, B.G. **2011**, *submitted*.
16. A “Tag-and-Modify” Approach to Site-selective Protein Modification. Chalker, J.M.; Bernardes, G.J.L.; Davis, B.G. *Acc. Chem. Res.* **2011**, *44*, 730-741.
15. Methods for the Conversion of Cysteine to Dehydroalanine on Peptides and Proteins. Chalker, J.M.; Gunnoo, S.B.; Boutureira, O.; Gerstberger, S.C.; Fernández-González, M.; Bernardes, G.J.L.; Griffin, L.; Hailu, H.; Schofield, C.J.; Davis, B.G. *Chem. Sci.* **2011**, *2*, 1666-1676.
14. A Coordinated Synthesis and Conjugation Strategy for the Preparation of Homogeneous Glycoconjugate Vaccine Candidates. Grayson, E.J.; Bernardes, G.J.L.; Chalker, J.M.; Boutureira, O.; Koeppe, J.R.; Davis, B.G. *Angew. Chem. Int. Ed.* **2011**, *50*, 4127-4132.
13. Chemical mutagenesis: Selective post-expression interconversion of protein amino acid residues. Chalker, J.M.; Davis, B.G. *Curr. Opin. Chem. Biol.* **2010**, *14*, 781-789.
12. Olefin Cross-Metathesis on Proteins: Investigation of Allylic Chalcogen Effects and Guiding Principles in Metathesis Partner Selection. Lin, Y.A.; Chalker, J.M.; Davis, B.G. *J. Am. Chem. Soc.* **2010**, *132*, 16805-16811.
11. Site-Selective Chemoenzymatic Synthesis of Glycoproteins using Endoglycosidases. Fernández-González, M.; Bourureira, O.; Bernardes, G.J.L.; Chalker, J.M.; Young, M.A.; Errey, J.C.; Davis, B.G. *Chem. Sci.* **2010**, *1*, 709-715.
10. Safe and Scalable Preparation of Barluenga’s Reagent. Chalker, J.M.; Thompson, A.L.; Davis, B.G. *Org. Synth.* **2010**, *87*, 288-298.
9. Chemical Protein Modification. In *Ideas in Chemistry and Molecular Sciences: Where Chemistry Meets Life*. Wiley, **2010**. Bernardes, G.J.L.; Chalker, J.M.; Davis, B.G.
8. A Convenient Catalyst for Aqueous and Protein Suzuki-Miyaura Cross-Coupling. Chalker, J.M.; Wood, C.S.C.; Davis, B.G. *J. Am. Chem. Soc.*, **2009**, *131*, 16346-16347.
7. Enabling Olefin Metathesis on Proteins: Chemical Incorporation of *S*-Allyl Cysteine. Chalker, J.M.; Lin, Y.A.; Boutureira, O.; Davis, B.G. *Chem. Commun.*, **2009**, 3714-3716.

6. Chemical Modification of Proteins at Cysteine: Opportunities in Chemistry and Biology. Chalker, J.M.; Bernardes, G.J.L.; Lin, Y.A.; Davis, B.G. *Chem. Asian J.*, **2009**, *4*, 630-640.
5. Olefin Metathesis for Site-Selective Protein Modification. Lin, Y.A.; Chalker, J.M.; Davis, B.G. *ChemBioChem*, **2009**, *10*, 959-969.
4. Allyl Sulfides are Privileged Substrates in Aqueous Cross-Metathesis: Application to Site-Selective Protein Modification. Lin, Y.A.; Chalker, J.M.; Floyd, N.; Bernardes, G.J.L.; Davis, B.G. *J. Am. Chem. Soc.*, **2008**, *130*, 9642-9643.
3. Chemical approaches to mapping the function of post-translational modifications. Gamblin, D.P.; van Kasteren, S.I.; Chalker, J.M.; Davis, B.G. *FEBS J.* **2008**, *275*, 1949-1959.
2. Facile Conversion of Cysteine and Alkyl Cysteins to Dehydroalanine on Protein Surfaces: Versatile and Switchable Access to Functionalized Proteins. Bernardes, G.J.L.; Chalker, J.M.; Errey, J.C.; Davis, B.G. *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053.
1. From Disulfide- to Thioether-Linked Glycoproteins. Bernardes, G.J.L.; Grayson, E.J.; Thompson, S.; Chalker, J.M.; Errey, J.C.; El Oualid, F.; Claridge, T.D.W.; Davis, B.G. *Angew. Chem. Int. Ed.* **2008**, *47*, 2244-2247.

Other publications co-authored by Justin M. Chalker:

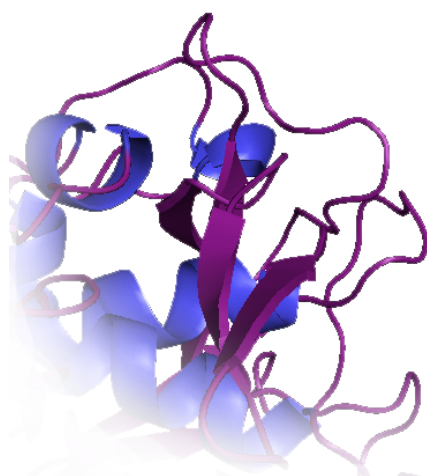
5. Synthesis of (-)- α -Kainic Acid via TMSCl-Promoted Pd-Catalyzed Zinc-ene Cyclization of an Allyl Acetate. Wei, G.; Chalker, J.M.; Cohen, T. *J. Org. Chem.* **2011**, In press. DOI: 10.1021/jo201341q.
4. Analysis of the dispersity in carbohydrate loading of synthetic glycoproteins using MALDI-TOF mass spectrometry. Patel, M.K.; Vijayakrishnan, B.; Koeppe, J.R.; Chalker, J.M.; Doores, K.J.; Davis, B.G. *Chem. Commun.* **2010**, *46*, 9119-9121.
3. Chemical site-selective prenylation of proteins. Gamblin, D.P.; van Kasteren, S.; Bernardes, G.J.L.; Chalker, J.M.; Oldham, N.J.; Fairbanks, A.J.; Davis, B.G. *Mol. Biosyst.*, **2008**, *4*, 558-561.
2. Two Syntheses of (-)-Kainic Acid via Highly Stereoselective Zinc-ene Cyclizations. Chalker, J.M.; Yang, A.; Deng, K.; Cohen, T. *Org. Lett.* **2007**, *9*, 3825-3828.
1. Allyl Sulfones as Precursors to Allylzincs in the Palladium-Catalyzed Zinc-ene Cyclization: Highly Efficient Synthesis of Enantiopure (-)-Erythrodiene. Deng, K.; Chalker J.; Yang, A.; Cohen, T. *Org. Lett.* **2005**, *7*, 3637-3640.

A selection of abbreviations used throughout this Thesis

aq.	aqueous
Boc	<i>tert</i> -butoxycarbonyl
CH ₂ Cl ₂	dichloromethane
Dha	dehydroalanine
DMF	dimethylformamide
Et ₃ N	triethylamine
Et ₂ O	diethyl ether
EtOAc	ethyl acetate
EPL	expressed protein ligation
ESI	electrospray ionization
GFP	green fluorescent protein
GlcNAc	<i>N</i> -acetyl-D-glucosamine
HMPT	hexamethylphosphoroustriamide or tris(dimethylamino)phosphine
ⁱ PrOH	2-propanol
K ₂ CO ₃	potassium carbonate
LC-MS	liquid chromatography—mass spectrometry
LRMS	low resolution mass spectrometry (or spectrum)
HRMS	high resolution mass spectrometry (or spectrum)
MeOH	methanol
MSH	<i>O</i> -mesitylenesulfonylhydroxylamine
NaHCO ₃	sodium hydrogen carbonate
NaOH	sodium hydroxide
NCL	native chemical ligation
NH ₄ Cl	ammonium chloride
NH ₄ OH	ammonium hydroxide
NMR	nuclear magnetic resonance
Np276	protein encoded by gene 276 in <i>Nostoc punctiforme</i>
PBu ₃	tributylphosphine
PTM	post-translational modification
sat.	saturated
SBL	subtilisin from <i>Bacillus lentus</i>

Chapter 1

An Introduction to the Chemical Modification of Proteins



“...the chemical enigma of Life will not be solved until organic chemistry has mastered another, even more difficult subject, the proteins...”

Emil Fischer

Nobel Lecture in Chemistry, 1902

1.1 Introduction

Chemical modification of proteins is central to many scientific endeavors.¹ One common application of protein modification is the analysis of protein location and function. The attachment of a fluorophore, for instance, has been widely used to investigate protein trafficking, binding, and other cellular events.^{2,3} Highly selective chemistry has been developed that even allows this protein labeling in living organisms.⁴⁻⁷ These are critical technologies for the determination of a protein's function in a physiological context.⁸ The selective chemical manipulation of proteins has also played an essential role in understanding the mechanisms of enzyme catalysis.¹ In other cases, chemically modified enzymes can display enhanced or entirely different activity.^{9,10}

Many proteins are *naturally* modified after translation, dramatically increasing their functional capacity.^{11,12} In many cases, the study of these proteins is complicated by their low abundance and tedious purification required for their isolation in homogeneous form. The selective chemical modification of the parent protein can provide ready access to these post-translationally modified proteins, enabling the study of their structure and function.¹³

Modified proteins are also gaining headway in therapeutic arenas with hundreds of protein-based pharmaceuticals under clinical evaluation.¹⁴ Synthetic vaccines,^{15,16} antibody therapies,¹⁷ and other biologic level pharmaceuticals¹⁸⁻²⁰ often feature chemical modification of proteins as a central part of their synthesis.

With so many diverse applications of chemical protein modification, it is not surprising that many methods exist for the chemical alteration of protein structure.^{21,22} In the introduction to this Thesis, several strategies will be reviewed to highlight both the breadth and utility of known protein modifications, but also their inherent limitations. This survey is organized into three general sections: chemical modification of natural residues, incorporation and modification of unnatural residues, and the total chemical synthesis of proteins. For each strategy, its development, utility, and shortcomings will be highlighted. This review will motivate and contextualize the objectives of this Thesis that will be summarized at the end of this Chapter.

1.2 Chemical Modification of Proteins at Natural Residues

The most common strategy for labeling natural residues is to treat a protein with an electrophile for modification of nucleophilic side chains.^{1,21} The modification of natural

residues was a critical development in the study of proteins. With no recombinant technology available at the dawn of protein science, chemical interrogation of protein structure and function was necessary on substrates obtained from natural sources. Many of these techniques were developed over half a century ago, yet the core chemistry is still used today. These key transformations are discussed next.

Lysine modification

The ϵ -amino group of the lysine side chain is a common target for protein modification. Lysine is an abundant residue²³ and protocols for its modification typically result in a statistical distribution of modified residues on the protein surface. The ϵ -amino group can be converted to a variety of derivatives and a selection of transformations is shown in Figure 1.1.

Isocyanates were among the earliest reagents used in protein modification. These reagents react efficiently with lysine, providing the urea derivative shown in Figure 1.1A.^{24,25} This class of reagents was useful in early chemical studies of insulin structure and function.²⁴ Isocyanates were also used to modify lysine residues to assess their role in binding in some of the earliest immunological assays.²⁵ Related modifications using isothiocyanates have also been reported.²⁶ These modifications were inspired in part by Edman's pioneering use of phenyl isothiocyanate in peptide sequencing.^{27,28} The use of isothiocyanates for protein labeling was one of the first methods for attaching antigens to protein carriers for immunological studies,²⁶ a technique still used today in synthetic vaccine campaigns.²⁹

Carbamate and guanidine linkages can also be installed at lysine by reaction with chlorocarbonates or isourea reagents, respectively (Figure 1.1 B-C).^{30,31} Steroid-protein conjugates have been prepared using the carbamate linkage,³⁰ and lysine guanidation has been used to facilitate protein crystallization.³¹

The development of active esters such as the *N*-hydroxysuccinimide (NHS) esters have been particularly useful for the mild and efficient amidation of the lysine ϵ -amino group (Figure 1.1D).^{32,33} NHS and other active esters were used in some of the earliest reports of affinity labeling—the covalent modification of a protein by reaction with a molecule to which the protein binds non-covalently.^{34,35} Homologous amides can also be formed at lysine using Tietze's squarate linker (Figure 1.1E).³⁶ The squarate linker can be

used to cross-link a small molecule containing an amine to the lysine of a protein carrier.³⁷

For the lysine modifications discussed so far, other than guanidation, the charged lysine side chain is neutralized. In some cases it might be advantageous to maintain the overall charge of the protein so the structure and solubility are not adversely affected by the modification. Accordingly, lysine can be converted to an amidine or a secondary amine if the positive charge should be maintained at the side chain (Figure 1.1 F-H). Amidine linkages have been used in the synthesis of glycoproteins³⁸ and reductive alkylation^{39,40} has also been used to attach carbohydrates and other labels to proteins. While reductive alkylation on proteins is typically carried out using an aldehyde and sodium borohydride (Figure 1.1G), milder methods of transfer hydrogenation have recently been developed (Figure 1.1H).⁴¹ In the latter case, the use of transition metal catalysis is a notable and relatively recent addition to protein modification,⁴² a general strategy that will be furthered explored in this Thesis.

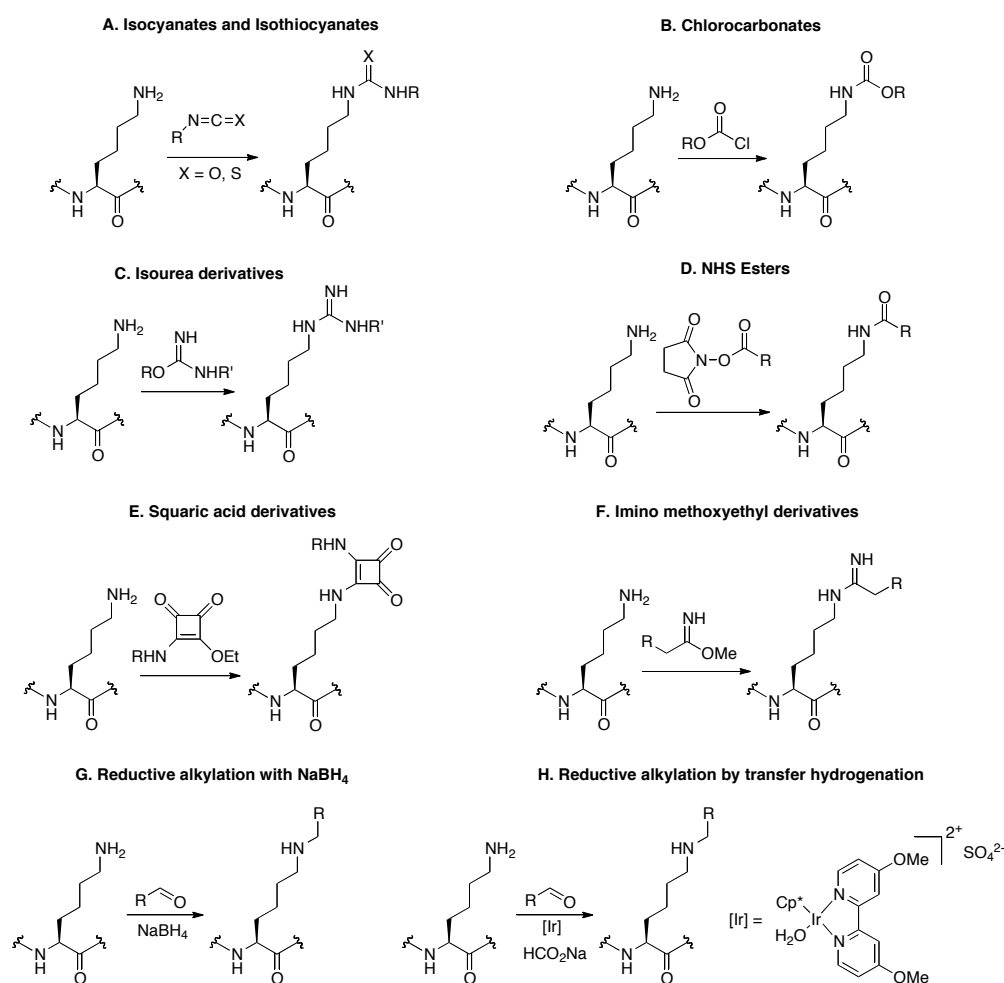


Figure 1.1: Representative reagents and methods for lysine modification

Glutamate and aspartate modification

Lysine is a convenient target for protein modification if high levels of protein labeling are required and site-specificity is not a concern. Glutamate and aspartate are also abundant residues²³ that can be targeted for non-selective modification. The carboxylate side chains of glutamate and aspartate are typically converted to esters or amides. In the early studies of glutamate and aspartate function, acidic methanol or diazomethane was used for their esterification (Figure 1.2A).^{43,44} Milder amide couplings were later made possible by the development of water-soluble carbodiimides (Figure 1.2B).^{45,46} These coupling methods facilitated early efforts to attach antigens to proteins for the production of antibodies.⁴⁷ Amidation of glutamate and aspartate has also been used for glycoprotein synthesis.⁴⁸

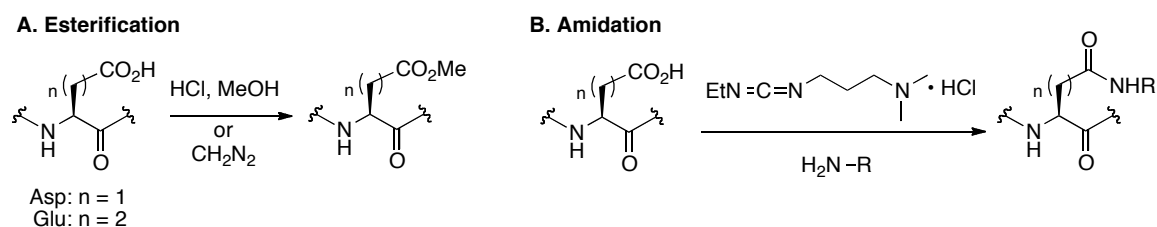


Figure 1.2: Representative methods for glutamate and aspartate modification

Cysteine modification

Cysteine is generally the most nucleophilic residue in proteins. The side chain sulfhydryl reacts with a variety of electrophiles and is central to many efforts in protein modification.^{49,50} Cysteine also has a relatively low natural abundance,²³ often enabling a single reactive cysteine to be genetically positioned at the desired site of modification.⁵⁰ The earliest reagents used to selectively modify cysteine are still used widely today so it is useful to review these methods here.^{21,49,50}

Simple alkylation of cysteine with iodoacetic acid and iodoacetamides is one method to install thioether modifications on proteins (Figure 1.3A). This chemistry was first described on model substrates such as glutathione.⁵¹ While these electrophiles can also react with lysine, cysteine can usually be selectively alkylated.⁵² Iodoacetamide was first used to investigate the number, accessibility, and reactivity of cysteine in various proteins.⁵³ Since these reports, the haloacetamides have been widely used in cysteine alkylation. A notable example is their use in the site-selective synthesis of glycoproteins, as reported by Flitsch.^{54,55}

Another simple, yet powerful, transformation is the aminoethylation of cysteine (Figure 1.3B). This alkylation results in a functional analog of lysine. This cysteine modification was first used to convert cysteine to a lysine analog to synthetically install a site for proteolysis by trypsin.⁵⁶ This protocol was useful in efforts to map protein sequences.^{56,57} These lysine analogs have also been used to investigate catalytic lysines in enzymes.⁵⁸⁻⁶¹ Recent interest in this transformation has been renewed by its utility in accessing methylated lysine analogs.⁶² Lysine methylation is a critical post-translational modification of the histone proteins in chromatin and the aminoethylation of cysteine allows straightforward modification of these proteins.⁶² More on methylated lysine analogs and the chemistry to access these modifications can be found in Chapter 4 where histone modifications are explored.

Cysteine can also be converted to a thioether by conjugate addition to various α,β -unsaturated electrophiles (Figure 1.3 C-D). Maleic acid⁶³ and maleimide⁶⁴ are two such electrophiles used in early efforts to investigate the abundance and function of cysteine residues in proteins. Contemporary protein synthesis has benefited widely from the development of these reagents and other related electrophiles such as vinyl sulfones.⁶⁵ To highlight one example of these cysteine alkylations, Waldmann and co-workers have used maleimide chemistry to attach lipidated peptides to proteins in a site-selective manner.⁶⁶ These proteins were useful in the study of polyprenylated proteins and their role in signal transduction.⁶⁶

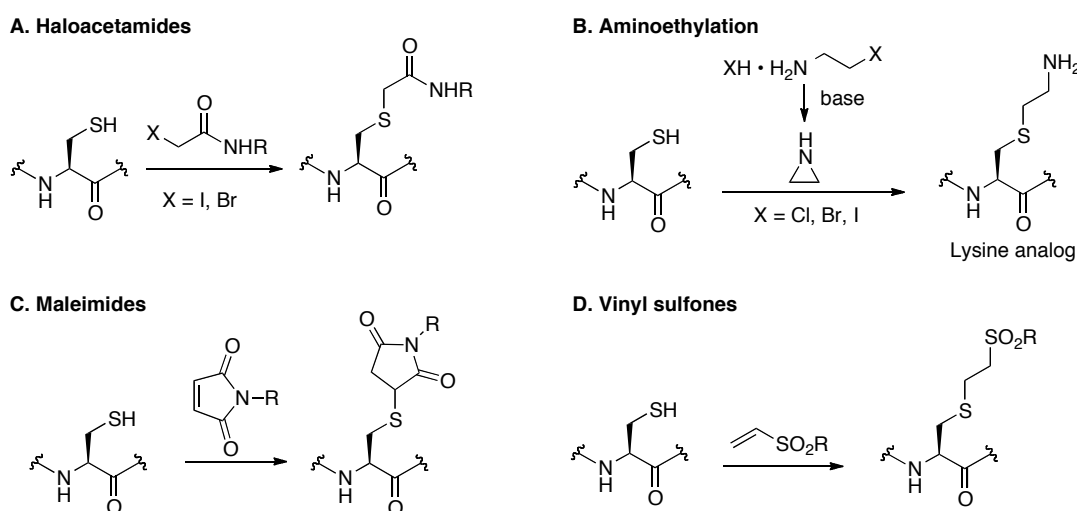


Figure 1.3: Representative reagents and methods for cysteine alkylation

Cysteine is a redox-active residue. The various oxidation states of cysteine are critical in maintaining protein structure, metal-binding, cellular redox regulation, and catalysis.⁶⁷⁻⁶⁹ The chemical manipulation of the oxidation state of cysteine is also useful in labeling proteins. Attaching various modifications through a disulfide-linkage at cysteine is one manner to exploit this unique reactivity of cysteine. Disulfide exchange has long been used to convert cysteine to disulfides. Ellman's method for quantifying thiol content of proteins is based on this simple disulfide exchange.⁷⁰ Several other sulfenylating reagents are also widely used in protein modification and a selection is shown in Figure 1.4.⁷¹⁻⁷⁵ These reagents have been used for the specific modification of cysteine in the synthesis of glycoproteins⁷²⁻⁷⁵ and lipidated proteins.⁷⁶ While disulfide formation is specific for cysteine, the linkage of the modification is susceptible to reduction. In some cases, however, this lability can be used to one's advantage. For instance, drug cargo can be attached to a peptide or protein through a disulfide linkage and then released after entering the reducing environment of the cell.⁷⁷

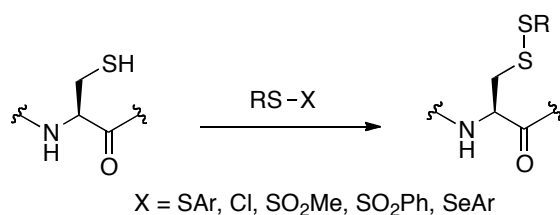
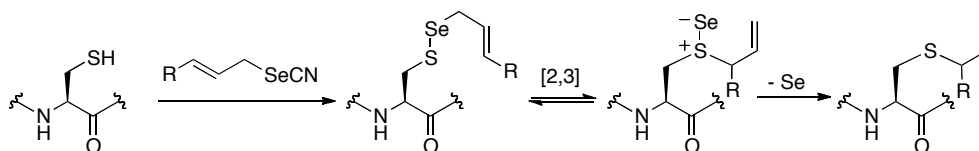
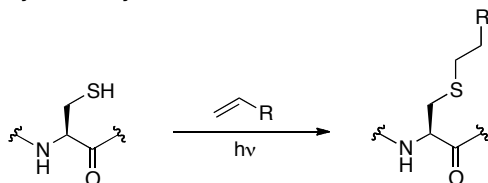
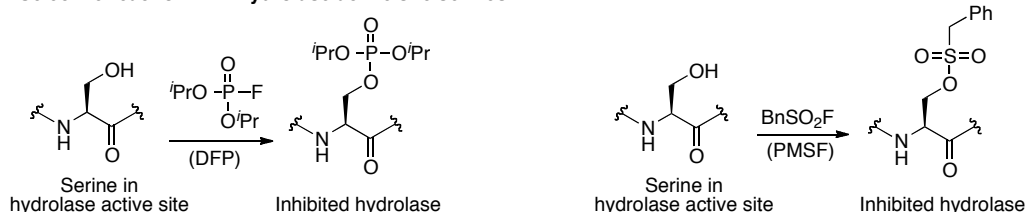
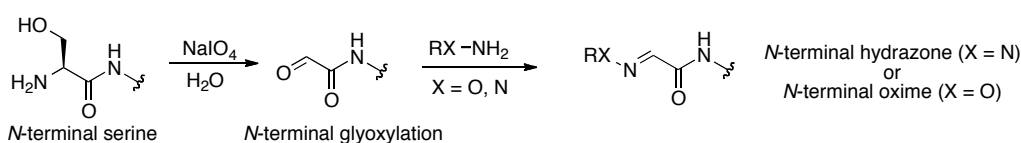


Figure 1.4: Cysteine sulfenylation

Recent developments in cysteine modification have allowed its conversion to a stable thioether in a highly specific manner, avoiding non-selective alkylation that sometime accompanies the use of more traditional electrophiles. The Crich laboratory has described an “electrophile free” dechalcogenative rearrangement of allylic disulfides and allyl selenenylsulfides.⁷⁸⁻⁸¹ A representative allylation is shown in Figure 1.5A. This chemistry benefits from both the high specificity of disulfide and selenenylsulfide formation and the stability of the final thioether product. This chemistry will be discussed in more detail in Chapter 5 where these cysteine-specific allylations were useful in the study of olefin metathesis on proteins. Another highly specific alkylation of cysteine was recently reported by Dondoni which takes advantage of the addition of cysteine thiyl radicals to alkenes (Figure 1.5B).⁸² Dondoni's “thiol-ene” chemistry was used to prepare well-defined glycoproteins.⁸²

A. Dechalcogenative allylation**B. Cysteine thiyl radical addition to alkenes****Figure 1.5:** Alkylations specific for cysteine**Serine modification**

Serine residues are not the most common targets for protein modification, but a few methods are important to highlight. For instance, many hydrolase enzymes contain a highly reactive serine in their active site. Frequently this serine can be selectively modified with reagents such as diisopropyl fluorophosphate (DFP)⁸³ and phenylmethanesulfonyl fluoride (PMSF) (Figure 1.6A).⁸⁴ The use of these reagents as inhibitors of hydrolases was an important development in enzymology and protein modification. These inhibited enzymes can be further manipulated into other residues and enzyme derivatives.⁸⁵⁻⁸⁸ In fact, the first point mutation of an enzyme was accomplished by the chemical conversion of serine to cysteine in the active site of chymotrypsin.^{86,87} The selective sulfonylation of the active site serine was a key step in this transformation. These chemical manipulations of active site residues were essential for the elucidation of their role in catalysis.

A. Selective reaction with hydrolase active site serines**B. Selective oxidation and modification of the N-terminal serine****Figure 1.6:** Selective modification of serine residues

A more general method for serine modification is the oxidation of *N*-terminal serine residues with sodium periodate.⁸⁹ The resulting aldehyde at the *N*-terminus of the protein that can be further modified by reaction with a hydrazide^{89,90} or aminoxy bearing nucleophiles⁹¹ to give hydrazones and oximes, respectively (Figure 1.6B).

Tyrosine modification

Tyrosine can be selectively modified by electrophilic aromatic substitution *ortho* to the tyrosine hydroxyl. These modifications can introduce either heteroatoms or result in the formation of carbon-carbon bonds.

For the introduction of heteroatoms at tyrosine, iodination is the most studied. Iodination is widely used to radiolabel peptides and proteins with the β -emitter ^{131}I for various metabolic or immunological studies. The reaction of Na^{131}I and the oxidant chloramines-T (*N*-chloro tosylamide) generates an electrophilic iodine source (^{131}ICl) that selectively iodinate tyrosine.⁹²

Tyrosine iodination can also aid protein structure determination in X-ray crystallography.^{93,94} Heavy atoms such as iodine display anomalous diffraction patterns that help solve the phase problem in crystallography.⁹⁵ Various iodonium sources have been reported for such protein iodinations (Figure 1.7).^{93,96} A novel synthesis of one such reagent, Barluenga's reagent, is the subject of Chapter 7.

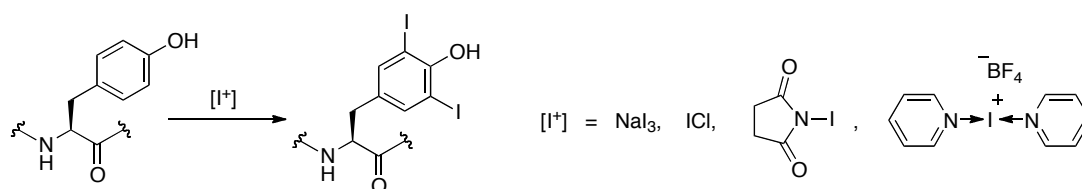
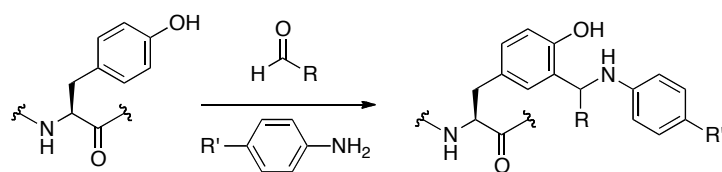


Figure 1.7: Reagents for iodination of tyrosine residues

For the formation of a carbon-carbon bond at tyrosine residues, a three-component Mannich reaction has been described by the Francis laboratory.⁹⁷ In this report, an aldehyde and an amine first condensed to form an imine which then reacted with tyrosine residues of chymotrypsinogen (Figure 1.8A). The same laboratory later reported the selective *O*-alkylation of tyrosine using π -allylpalladium electrophiles.⁹⁸ This transformation was useful for grafting hydrophobic domains onto proteins (Figure 1.8B) and was another recent demonstration of metal-mediated protein modification.⁴²

A. Tyrosine selective Mannich reaction



B. Tyrosine selective alkylation using π -allylpalladium electrophiles

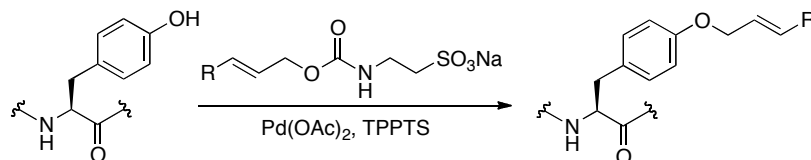


Figure 1.8: Methods for tyrosine alkylation

Tryptophan modification

Tryptophan is a relatively rare amino acid residue²³ and might therefore be a useful target for installing a single, well-defined modification at a desired site. However, very few methods have been reported for the selective functionalization of the indole side-chain.¹ Recently, the paucity of methods for tryptophan modification was addressed by Francis and co-workers who described selective tryptophan labeling using rhodium carbenoids.^{99,100} While the modification results in a regioisomeric mixture of products (Figure 1.9), the transformation illustrates how organometallic chemistry can be applied to unresolved problems in protein modification.⁴²

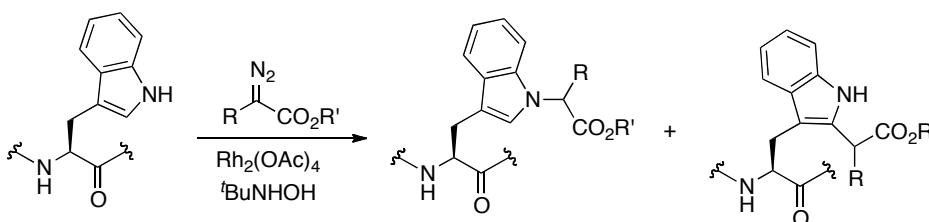


Figure 1.9: Tryptophan alkylation using rhodium carbenoids

1.3 Chemical Modification of Proteins at Unnatural Residues

In the previous section, the diversity of reactions available for modification of natural amino acid residues was discussed. In many cases these reactions can be highly selective, but lack of chemo- and regioselectivity is also common. To maximize selectivity, it is often desirable to use chemical reactions and functional groups that do not interact with the 20 natural residues typically found in proteins. This “bioorthogonal chemistry” can be

realized on proteins through the incorporation of uniquely reactive unnatural amino acids.⁴⁻⁷ There are many methods for the incorporation and modification of unnatural amino acid residues.¹⁰¹ Several of these methods most relevant to this Thesis are highlighted next.

Methods for the incorporation of unnatural amino acids into proteins

Amino acids other than the canonical set of 20 can be incorporated into proteins using several techniques. Non-natural amino acids can be incorporated as analogs of natural amino acids,¹⁰²⁻¹⁰⁸ the genetic code can be reprogrammed¹⁰⁹ or expanded,¹¹⁰ the degeneracy of the genetic code can be broken,¹¹¹ and quadruplet codons can be used in frameshift suppression¹¹² or through use of wholly re-engineered ribosomes.¹¹³ The two most widely used techniques at present are 1) incorporation of unnatural amino acids in place of natural residues using bacteria that cannot biosynthesize the natural residue and 2) genetic code expansion using a mutant tRNA / aminoacyl-tRNA synthetase pair that incorporate an unnatural amino acid in response to the amber stop codon. Both techniques are relevant to the chemistry described in this Thesis and so a brief overview of their technical details are discussed next. For other techniques and a more detailed discussion of unnatural amino acid incorporation into proteins, the reader is referred to the several excellent reviews cited here.^{101,109,110,114,115}

E. coli strains that cannot biosynthesize a particular amino acid are *auxotrophic* for that residue and depend on its uptake from the environment or growth media to survive. *E. coli* that are auxotrophic for methionine or phenylalanine have been used to incorporate unnatural amino acids that are structurally similar to these residues. The unnatural amino acid is simply used as a replacement for its natural counterpart in the growth media. The unnatural amino acid is then incorporated in response to the codons normally reserved for the amino acid it resembles. For example, the amino acids in Figure 1.10 are sufficiently similar to methionine to be efficiently incorporated in its place using methionine auxotrophic bacteria in methionine depleted media. Amino acids with azide, alkene, and alkyne functional groups can all be incorporated by this method.¹⁰²⁻¹⁰⁵ These residues are all useful functional groups for further protein modification and many reactions have been demonstrated or proposed for selective modification of these residues (*vide infra*).

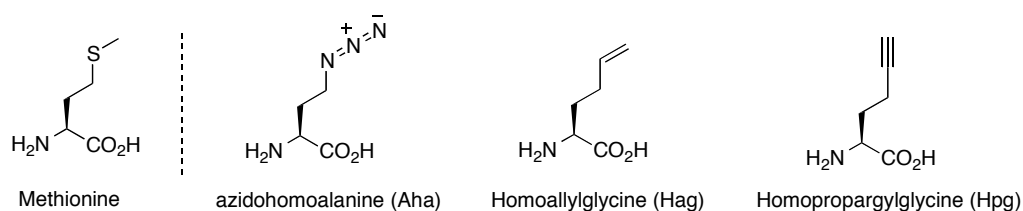


Figure 1.10: A selection of amino acids that can be incorporated into proteins in place of methionine using bacteria that are auxotrophic for methionine

Analogues of phenylalanine have also been incorporated using auxotrophic bacteria. In these cases, incorporation efficiency can be increased by evolving the phenylalanyl-aminoacyl-tRNA synthetase so that its substrate specificity is relaxed. Amino acids containing aryl halides,^{106,107} aryl azides,¹⁰⁷ and aryl ketones¹⁰⁸ can be incorporated into proteins using this strategy (Figure 1.11). Again, these are useful functional groups for further chemical modification of proteins.

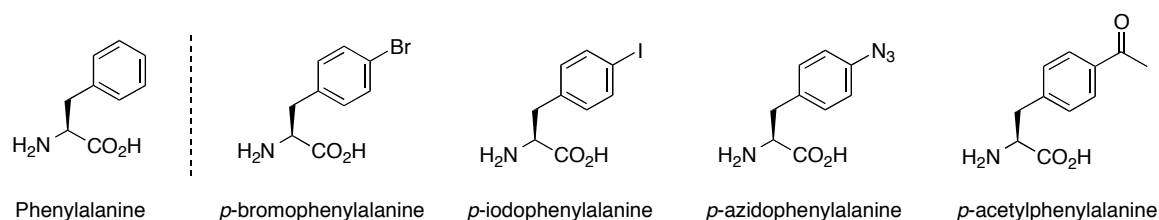


Figure 1.11: A selection of amino acids that can be incorporated into proteins in place of phenylalanine using bacteria that are auxotrophic for phenylalanine

It should be made clear that using auxotrophic bacteria for unnatural amino acid incorporation results in the *replacement* of a natural residue. If the unnatural amino acid needs to be incorporated *in addition to* the standard set of twenty natural residues, a different technique is required. Amber codon (UAG) suppression is one strategy for expanding the genetic code.^{116,117} In its early form, suppressor tRNA was chemically aminoacylated for the incorporation of the unnatural amino acid in response to the amber stop codon.¹¹⁶ In a more general strategy, developed extensively by the Schultz laboratory, the amber stop codon is reassigned to an unnatural amino acid using an evolved tRNA / aminoacyl-tRNA synthetase pair—both are which are synthesized by the host organism using recombinant technology.^{110,117} The mutant tRNA / aminoacyl-tRNA synthetase pairs must be orthogonal to the host translational machinery. In other words, the engineered tRNA should not be aminoacylated by any endogenous aminoacyl-tRNA

synthetase and the mutant aminoacyl-tRNA synthetase should not aminoacylate any host tRNA.

Two sources of tRNA / aminoacyl-tRNA synthetase pairs have been widely used as a starting point for directed evolution for the incorporation of unnatural amino acids into proteins. The first source is the archaea *Methanococcus jannaschii* and its tyrosyl-tRNA / tyrosyl-tRNA synthetase pair.¹¹⁰ The tyrosyl-tRNA from this organism has recognition elements distinct from *E. coli* and the tyrosyl-tRNA synthetase binds to this tRNA in a way that allows mutation of the anticodon loop of the tRNA to recognize the stop codon without losing affinity for the tyrosyl-tRNA synthetase.^{110,117} The second source is the methanogenic bacteria *Methanosarcina barkeri* and its pyrrolysyl-tRNA / pyrrolysyl-tRNA synthetase pair. *Methanosarcina barkeri* naturally incorporates pyrrolysine in response to the amber stop codon present in certain genes of this organism.¹¹⁸

The general strategy for obtaining orthogonal tRNA / aminoacyl-tRNA synthetase pairs by directed evolution is shown in Figure 1.12. The tRNA is first altered to recognize the stop codon, if it does not do so already. Next, key residues in the aminoacyl-tRNA synthetase active site are randomized to create a library of candidate synthetases. For positive selection, the amber codon is inserted into a gene that is essential for the cell's survival. This gene may confer resistance to an antibiotic, for instance. If the cell lives, its aminoacyl-tRNA synthetase has successfully charged an amino acid to the tRNA. This amino acid may be the desired unnatural amino acid or one of the 20 natural amino acids. A negative selection is then carried out in the absence of the unnatural amino acid by incorporating the stop codon into a gene that produces a toxic protein. Only cells harboring an aminoacyl-tRNA synthetase that is specific for the unnatural amino acid will survive. Successive rounds of mutation and positive and negative selection are typically used to evolve high fidelity tRNA / aminoacyl-tRNA synthetase pairs.

A wide range of amino acids have been incorporated into proteins using amber codon suppression.¹¹⁰ The selected examples in Figure 1.13 highlight the remarkable functional diversity that can be genetically encoded into proteins. In the following sections, some reactions are discussed that are useful in the selective modification of these residues.

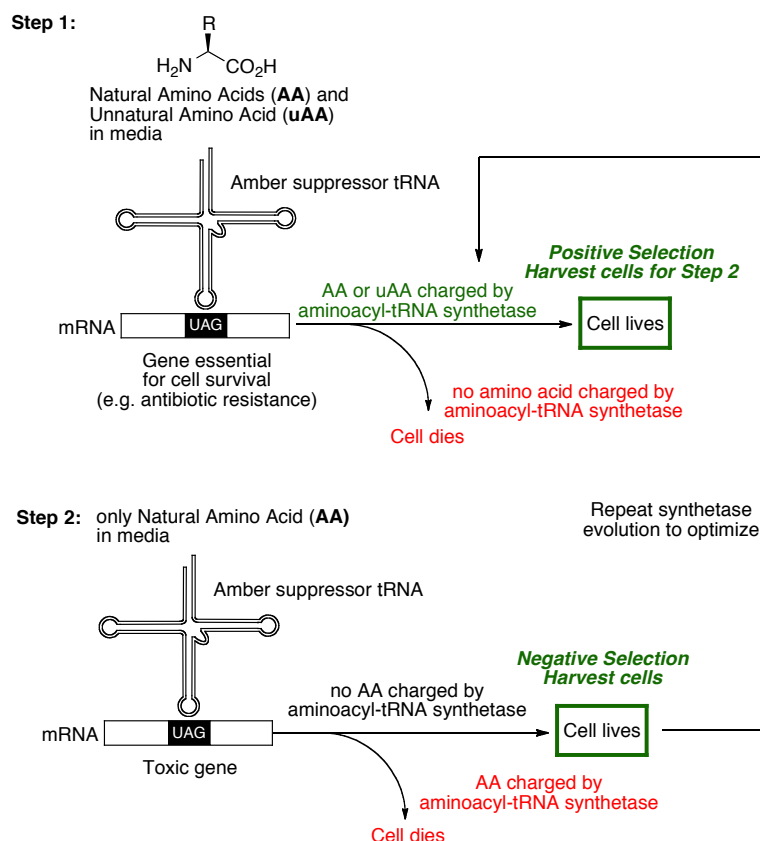


Figure 1.12: Evolution of an orthogonal tRNA / aminoacyl-tRNA synthetase pair for the incorporation of an unnatural amino acid into proteins

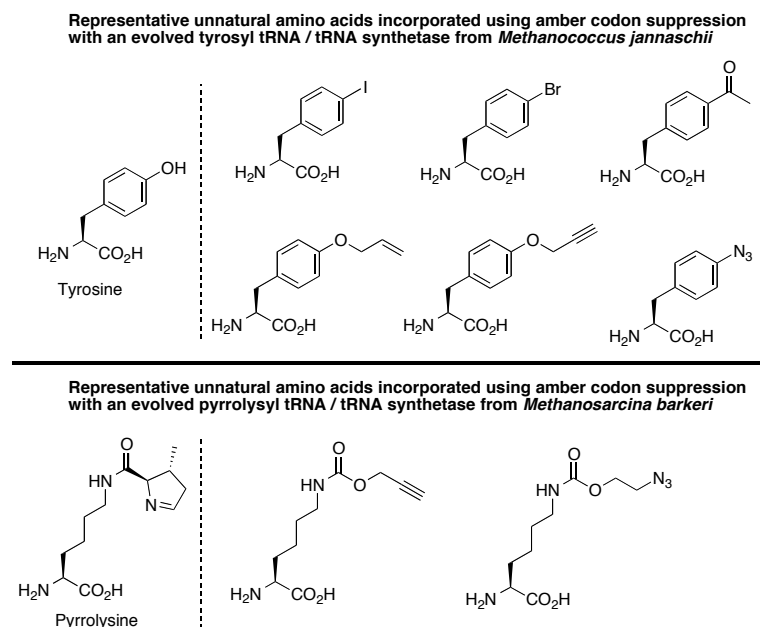


Figure 1.13: A selection of unnatural amino acids that can be incorporated into proteins by amber codon suppression. These amino acids contain functional groups useful in the chemical modification of proteins.

Modification of proteins containing azides and alkynes

As shown in Figures 1.10, 1.11, and 1.13 a variety of residues can be incorporated into proteins that contain an azide^{104,105,119-122} or alkyne.^{103,122,123} These functional groups are largely inert to the biological milieu, but can be selectively modified by a variety of bioorthogonal transformations. The Staudinger ligation and the [3+2]-cycloaddition of azides and alkynes are the two most widely used methods to selectively modify these residues.

The Staudinger ligation (Figure 1.14) is a variation of a classic reduction of azides with phosphines.¹²⁴ The Bertozzi laboratory introduced this reaction for the selective modification and detection of metabolically incorporated azides to monitor the formation of glycoproteins on the surface of cells.¹²⁵ In the classic Staudinger reaction, a phosphine reacts with an azide to give an aza-ylide that is then hydrolyzed in the presence of water to give the amine and phosphine oxide. If an aryl phosphine is used with an ester placed in the *ortho* position, the aza-ylide can be intramolecularly trapped. In this way, the azide can be labeled with a variety of fluorophores or affinity probes.

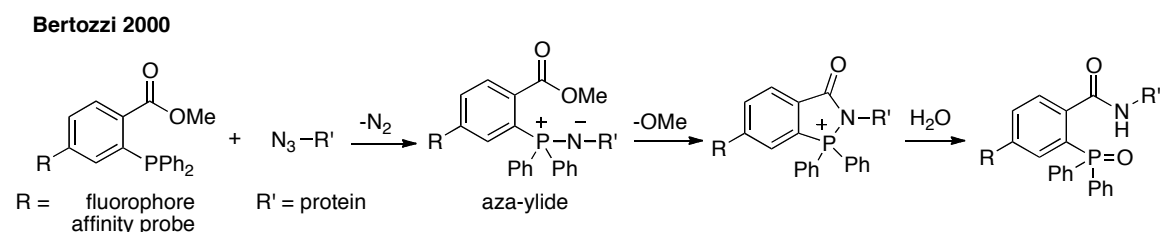
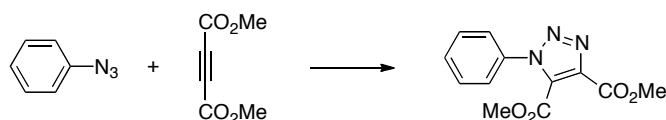


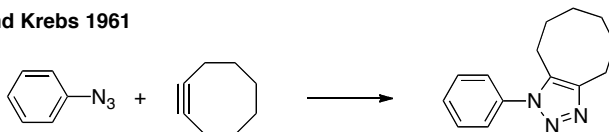
Figure 1.14: The Staudinger ligation

The most common reaction of azides and alkynes, particularly in the context of bioorthogonal chemistry, is their [3+2] cycloaddition. Such cycloadditions have been among the repertoire of organic synthesis for over a century,^{126,127} but their use in protein modification was enabled by the development of the mild copper-catalyzed variant reported independently by Sharpless¹²⁸ and Meldal¹²⁹ in 2002 (Figure 1.15). This reaction is now effectively synonymous with “click chemistry”¹³⁰ because of its high efficiency and selectivity. Copper-free variations of this cycloaddition are also available, and the use of strained alkynes such as cyclooctyne derivatives are particularly efficient (Figure 1.15).^{131,132} The utility of the [3+2] cycloaddition is evident from the large number of reports that use this transformation for *in vitro* protein labeling,¹³³ protein activity profiling,¹³⁴ cell-surface labeling,^{105,132} and selective imaging inside living cells.¹³⁵

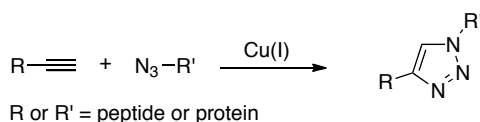
Michael 1893



Wittig and Krebs 1961



Sharpless 2002, 2003 and Meldal 2002



Bertozzi 2004

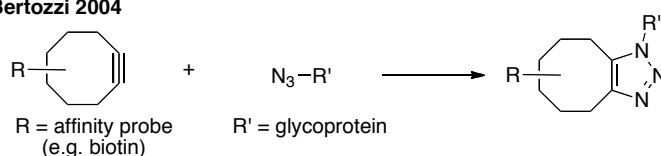


Figure 1.15: The azide-alkyne [3+2] cycloaddition

Modification of proteins containing ketones and aldehydes

Ketones and aldehydes are also useful functional groups for selective labeling of proteins. The most widely used reaction for such modifications is the formation of oximes or hydrazones by the reaction of ketones or aldehydes with hydroxylamines or hydrazines, respectively (Figure 1.16). An aldehyde can be incorporated into proteins by selective oxidation of *N*-terminal serine residues with periodate,⁸⁹ ketones can be genetically encoded,^{108,121,136} and aldehydes such as formylglycine can be introduced into proteins enzymatically.¹³⁷ Dawson has also shown that both hydrazone and oxime ligations are catalyzed by aniline,^{90,91} dramatically increasing the efficiency and generality of this ligation.

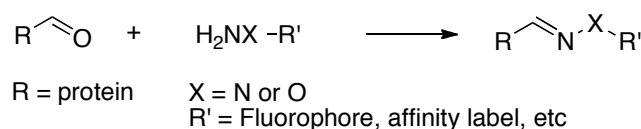


Figure 1.16: Oxime and hydrazone ligation at aldehydes or ketones

Modification of proteins containing alkenes

Alkenes, like alkynes and azides, are often inert to biological systems and are therefore useful handles for bioorthogonal chemistry. One transformation that has gained traction in recent years is the addition of thiyl radicals to alkenes (Figure 1.17). The so-called

“thiol-ene” reaction has been used in the synthesis of modified proteins for vaccine research,¹³⁸ the precise immobilization of proteins on surfaces,¹³⁹ and in glycoprotein synthesis.¹⁴⁰ In these reports, alkenes were introduced to proteins chemically^{138,139} or through auxotrophic expression.^{102,103,140} This transformation complements the reaction of cysteine with alkenes shown in Figure 1.5B.⁸²

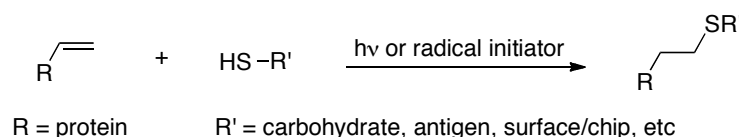


Figure 1.17: The addition of thiol radicals to alkene-containing proteins

Amber codon suppression can also be used to incorporate alkenes into proteins.¹⁴¹ In addition to the other alkene-modifying reactions previously discussed, these genetically encoded alkenes have been shown to react rapidly with diaryltetrazoles under photochemical conditions. In this transformation, the tetrazole extrudes nitrogen to give a nitrile imine that undergoes rapid cycloaddition with the alkene (Figure 1.18).^{142,143} The high specificity of this reaction allowed its use in living cells. This transformation complements a report by the same authors that described the incorporation of the tetrazole into the protein.¹⁴⁴

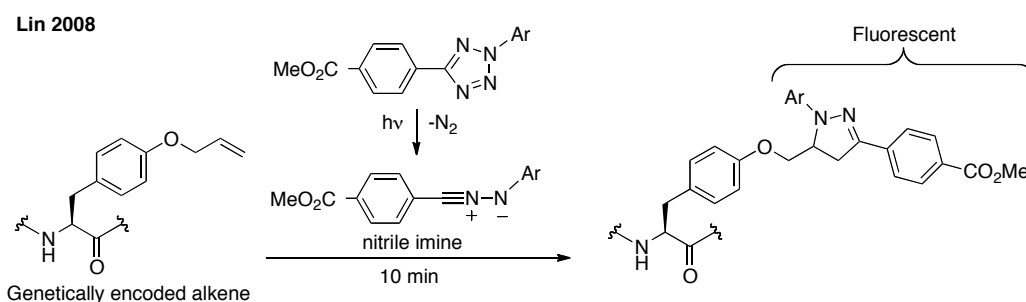


Figure 1.18: Photochemical generation of a nitrile imine from a tetrazole and selective cycloaddition with *O*-allyl tyrosine

A widespread reaction in organic synthesis is olefin metathesis.¹⁴⁵⁻¹⁴⁸ With the ability to incorporate alkenes into proteins, it is not surprising that olefin metathesis has been proposed as a method for protein modification. Indeed, Tirrell proposed the use of olefin metathesis in protein modification more than a decade ago when his lab reported a simple method for the genetic incorporation of alkenes.¹⁰² Olefin metathesis on proteins,

however, was an outstanding problem at the outset of this Thesis. In Chapter 5, a focused effort to adapt olefin metathesis to protein modification is described.

Modification of proteins containing aryl halides

The aryl halide is widely used in organic synthesis in cross-coupling reactions.^{149,150} For protein modification, cross-coupling is an attractive reaction since aryl halides and related cross-coupling partners can be incorporated through a variety of genetic techniques.^{106,107,151-157} Yet, despite the flexibility in the installation of the cross-coupling partner, cross-coupling on proteins has been met with little success. Efforts to carry out Heck and Sonogashira reactions on protein surfaces were met with low conversions (Figure 1.19).^{158,159} An effort to find a more active catalyst that allows efficient cross-coupling on proteins is described in Chapter 6.

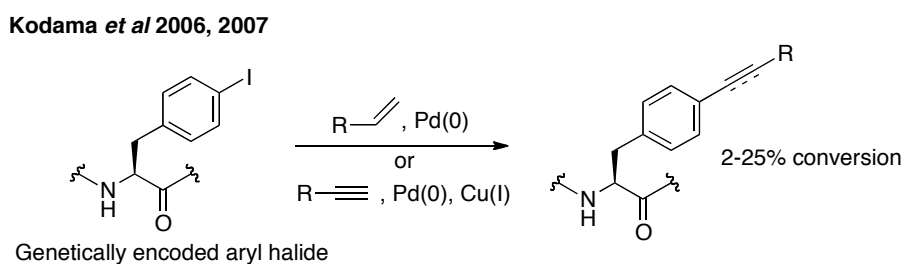


Figure 1.19: Cross-coupling at genetically incorporated aryl halides

Miscellaneous methods of protein modification using unnatural amino acids

A selection of additional methods for protein modification is shown Figure 1.20. The transformations in Figure 1.20 A and B are notable for their high rate and selectivity. The oxidative coupling in Figure 1.20A was used to label virus capsids containing a genetically encoded *p*-amino-phenylalanine.¹⁶⁰ The inverse-electron-demand Diels-Alder reaction in Figure 1.20B is another example of the exquisite selectivity of cycloadditions in protein modification.¹⁶¹ In Figure 1.20C, an example of a genetically encoded photocaged amino acid is depicted.¹⁶² Photocaged residues are particularly useful since their latent functionality can be revealed in a temporally controlled manner. More on photocaged residues can be found in the introductions to Chapters 3 and 4 where the synthesis and study of proteins bearing post-translational modifications is discussed. Many of these residues can be incorporated genetically or through total synthesis. The latter method is discussed next.

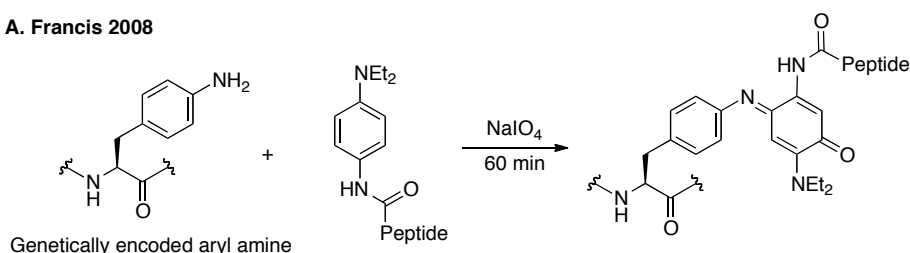
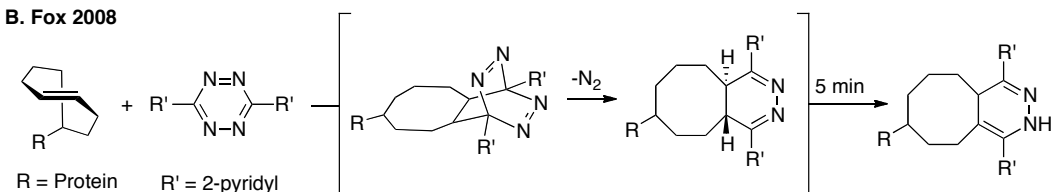
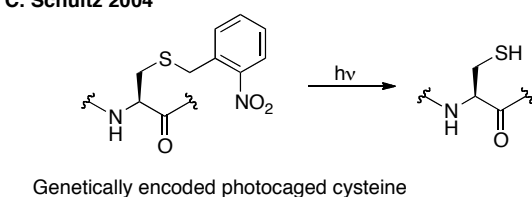
A. Francis 2008**B. Fox 2008****C. Schultz 2004**

Figure 1.20: Additional methods for protein modification using unnatural amino acids

1.4 Total Chemical Synthesis of Proteins

Proteins are not beyond the reach of total chemical synthesis. Since Emil Fischer's first peptide coupling,¹⁶³ the assembly of larger peptides and proteins by purely chemical means has held the attention of organic chemists. A number of technical developments such as the design of peptide protecting group strategies¹⁶⁴ and automated solid phase peptide synthesis¹⁶⁵⁻¹⁶⁷ were key milestones en route to the first total synthesis of a protein. In 1969, Gutte and Merrifield realized this goal in their synthesis of ribonuclease A.¹⁶⁸ They synthesized this 124 residue enzyme by solid phase synthesis. The enzyme was folded and disulfides formed by air oxidation *in vitro*—demonstrating that the sequence of the protein itself is sufficient, in this case, to mediate protein folding. The synthetic enzyme showed the same ribonuclease activity as the natural counterpart.¹⁶⁸ Independently and at the same time as Merrifield's report, scientists at Merck applied a solution phase approach to ribonuclease S.¹⁶⁹⁻¹⁷³ Both syntheses were powerful demonstrations of the application of organic chemistry to the total synthesis of full-length proteins.

The total synthesis of proteins, however, is not without complications. Merrifield's synthesis of ribonuclease A required 369 chemical reactions and nearly 12,000 steps by the peptide synthesizer.¹⁶⁸ Such a synthesis has little room for error and

coupling reactions must be nearly quantitative for acceptable overall yields in a linear synthesis. The synthesis of larger proteins therefore requires convergent assembly of multiple peptide fragments. However, solubility problems of large protected peptides can severely hamper solution-based syntheses. A key concept in the convergent assembly of large unprotected peptides was developed by Kent and co-workers that overcame these difficulties: chemical ligation.^{174,175}

Native chemical ligation

Chemical ligation in protein synthesis relies on the selective or specific condensation of two unprotected peptides. In its original form, this condensation was achieved by the reaction of two synthetic peptides, with one peptide bearing a C-terminal thioacid and the other peptide with N-terminal bromoacetylation.¹⁷⁴ The selective reaction of these peptides results in a thioester backbone at the site of ligation (Figure 1.21). The coupling provides full-length protein in deprotected form. This strategy was used in the synthesis of the HIV-1 protease and the product protein was fully active after refolding.¹⁷⁴

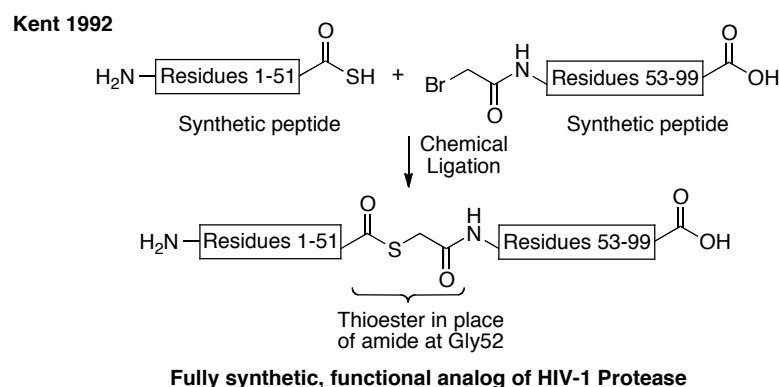


Figure 1.21: “Chemical ligation” in the convergent total synthesis of HIV-1 protease

The ligation of deprotected peptides was an important breakthrough in the total synthesis of proteins. Solid phase synthesis can provide high purity peptides up to 50 residues in length.^{175,176} Methods for the convergent assembly of these fragments were required for protein synthesis and the ligation in Figure 1.21 was one way to accomplish this assembly. The thioester backbone, however, was an unnatural linkage and in some cases access to the native amide bond might be advantageous. Soon after Kent’s synthesis of the HIV-1 protease, the same laboratory reported a method of ligation that generates the natural amide bond. Their strategy was coined “native chemical ligation.”¹⁷⁷

Native chemical ligation involves the merger of two peptides, with one peptide bearing a *C*-terminal thioester and the other peptide containing an *N*-terminal cysteine residue.¹⁷⁷ The *N*-terminal cysteine residue reacts reversibly with the thioester in a trans-thioesterification. The intermediate thioester obtained after this reaction can then undergo an effectively irreversible rearrangement to give the native peptide bond (Figure 1.22).^{177,178} The native chemical ligation (NCL) strategy was first applied to the synthesis of human interleukin 8.¹⁷⁷ In this synthesis, the compatibility of NCL with proteins containing multiple cysteines was also demonstrated. While thiols can add reversibly to the *C*-terminal thioester, only *N*-terminal cysteines rearrange to give the stable amide bond. Since this report from the Kent laboratories, NCL has been used extensively in protein synthesis.^{175,179,180}

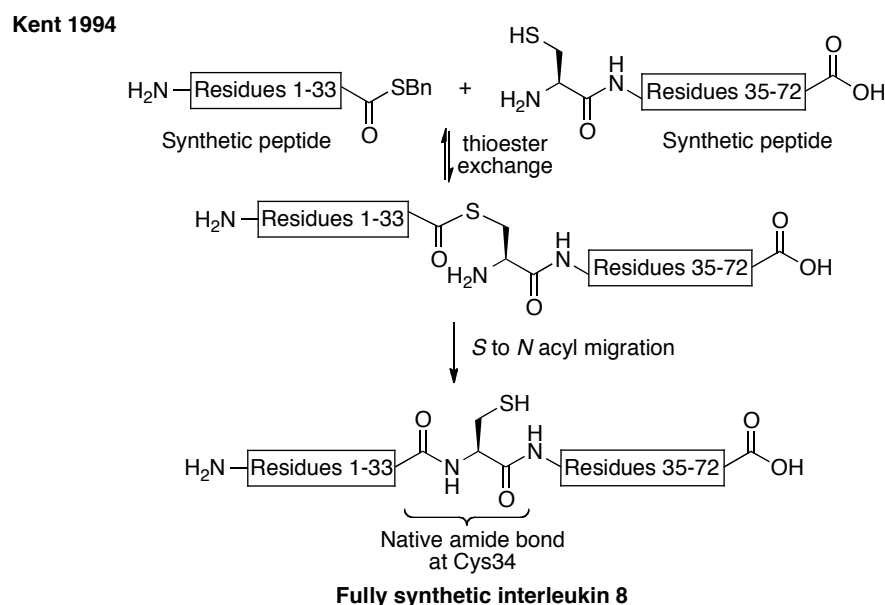


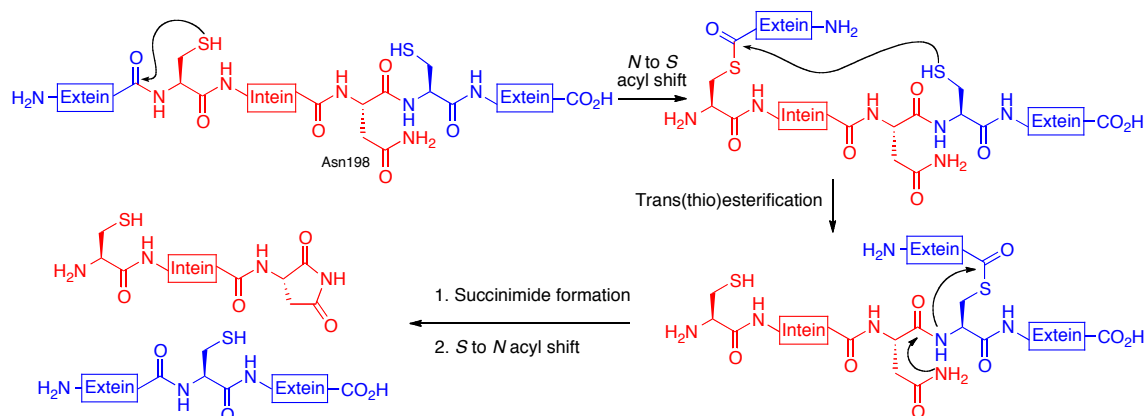
Figure 1.22: Native chemical ligation in the convergent total synthesis of interleukin 8

Expressed protein ligation

One of the challenges of NCL is the efficient synthesis of the peptide that contains the *C*-terminal thioester. A breakthrough in the synthesis of long peptides containing the *C*-terminal thioester was the development of expressed protein ligation (EPL). In this technique, the thioester peptide is produced by recombinant biosynthesis in *E. coli*.^{181,182} A key study that enabled EPL was the mechanistic analysis of natural protein splicing machinery.^{183,184} Protein splicing involves the intramolecular ligation of two external domains of a protein (the exteins) with concomitant excision of an internal domain (the

intein) (Figure 1.23A).^{183,184} Through judicious mutations, a team at New England Biolabs developed a system that prevented splicing but allowed the interception of a thioester intermediate at the C-terminus of the target protein. Their construct also contained an affinity tag at the C-terminus of the intein, allowing straightforward purification of the target protein-intein construct. Interception of the intermediate thioester with a thiol provided the target protein as a C-terminal thioester.¹⁸⁵ The logical intersection of this recombinant technique with native chemical ligation was reported soon after by Muir, Sondhi, and Cole,¹⁸² and also by New England Biolabs (Figure 1.23B).¹⁸¹ In these reports, the recombinant thioester was merged with a synthetic peptide with an N-terminal cysteine, a ligation entirely analogous to NCL. The overall process of recombinant thioester production followed by NCL is now widely referred to as expressed protein ligation.¹⁸²

A. Representative example of natural protein splicing



B. Engineered protein splicing machinery for "expressed protein ligation"

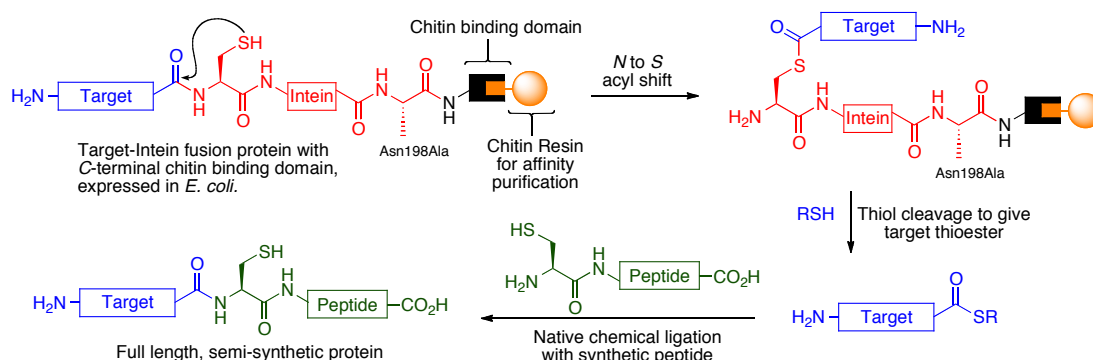


Figure 1.23: Protein splicing (A) and its adaptation to the semi-synthesis of proteins (B)

NCL in its classic form was practically limited to peptide segments of 50 residues or less. The development of EPL allowed the semi-synthesis of proteins without concern

for the size of the target protein. EPL has since been used in the synthesis of a variety of proteins, including those bearing post-translation modifications or unnatural biochemical probes.¹⁸⁶⁻¹⁸⁸ These proteins cannot be accessed purely by biosynthesis and so EPL constituted an important unification of chemical and recombinant protein synthesis.

Additional methods of peptide ligation

NCL and EPL are not without limitations, but many efforts have since been reported for the refinement and generalization of these techniques. For instance, if a protein is to be modified near middle of the sequence, multiple ligations might be required. Complementary methods of ligation have been introduced in an effort to resolve this problem and allow convergent synthesis of the synthetic peptide prior to NCL or EPL. For example, Raines and Kiessling have used the chemoselective ligation of a thioester and an *N*-terminal azide, another variant of the Staudinger ligation, as a method of peptide coupling.^{189,190} This method of ligation was used in conjunction with NCL in the three-fragment assembly of ribonuclease A (Figure 1.24).¹⁹¹

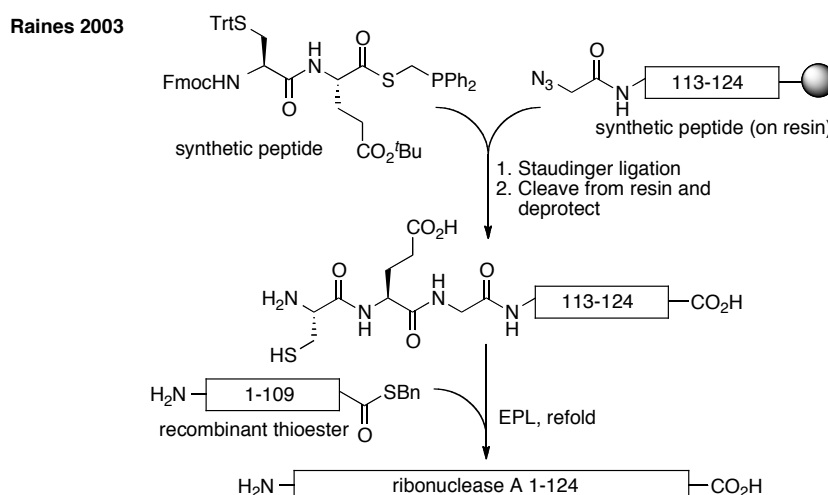


Figure 1.24: Multiple ligations in the total synthesis of ribonuclease A

Ligation at residues other than cysteine

Another apparent limitation of NCL and EPL is the inherent reliance of cysteine at the site of ligation. Two solutions to this problem have emerged. First is the use of auxiliaries that mediate the ligation as a cysteine surrogate, but can then be removed to provide a non-cysteine residue at the site of ligation (Figure 1.25).¹⁹²⁻¹⁹⁴

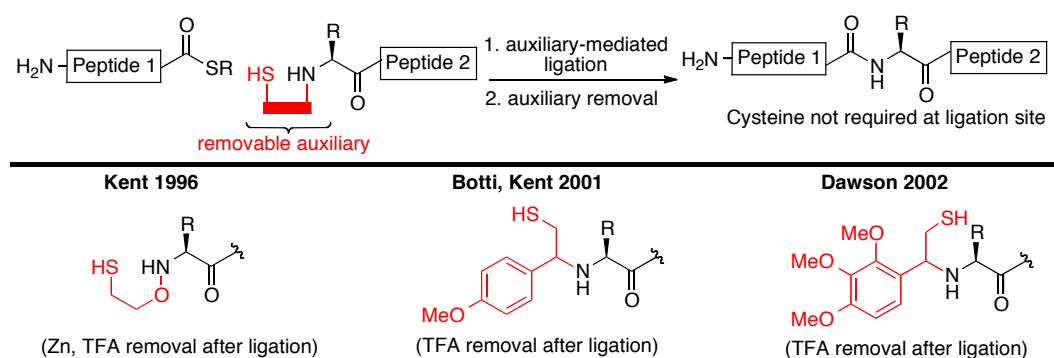


Figure 1.25: Auxiliary mediated ligation at residues other than cysteine

The second strategy to overcome the reliance on cysteine is the development of post-ligation desulfurization. Yan and Dawson first reported the tandem ligation-desulfurization protocol for the synthesis of cysteine free peptides and proteins.¹⁹⁵ After NCL in the usual manner, metal-mediated desulfurization converted cysteine to alanine. Several related strategies have since been reported that constitute, in the formal sense, a ligation at residues other than cysteine (Figure 1.26).¹⁹⁶⁻²⁰³ Notably, a mild radical-based desulfurization was reported by Danishefsky that employs the water-soluble phosphine TCEP, a radical initiator (VA-044), and a thiol source (*t*BuSH). Danishefsky's protocol typically gives higher yields than metal-based desulfurizations.¹⁹⁷

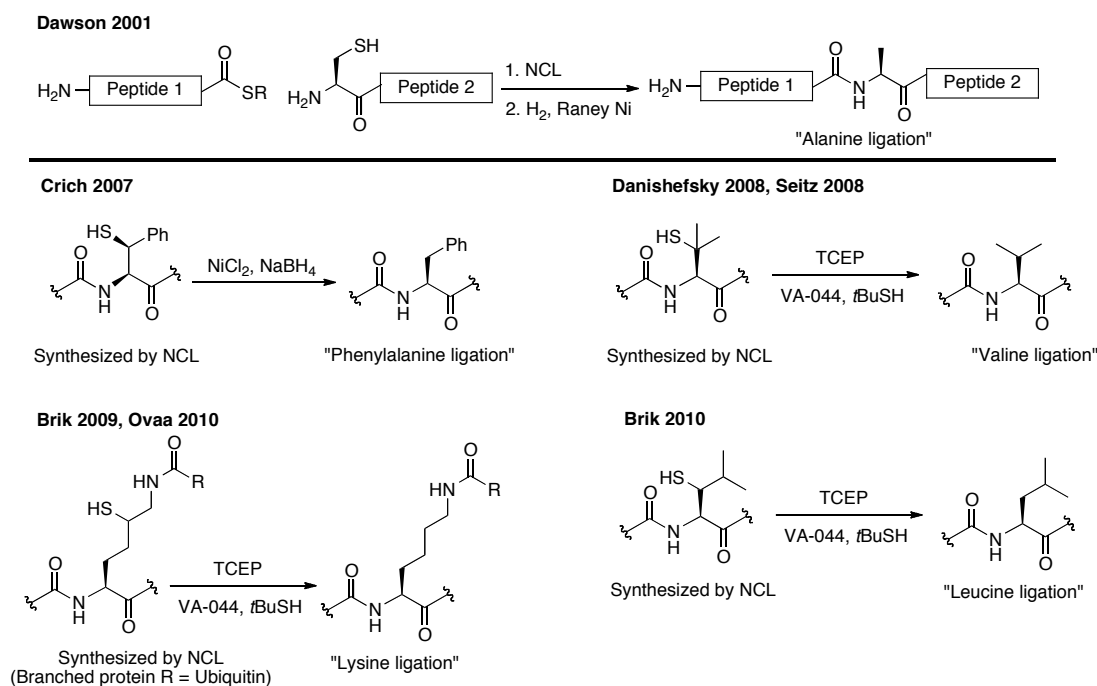


Figure 1.26: Synthesis of peptides and proteins by NCL followed by desulfurization

The total synthesis of proteins is a clear example of how chemistry can impact biology. Several additional examples of NCL and EPL and their utility in chemical biology will be highlighted in the introductions to several Chapters in this Thesis.

Next, the central goals of this Thesis are described. Many unsolved problems in the chemical modification of proteins were mentioned throughout this introduction; some of these problems are addressed explicitly in this Thesis. Additionally, novel methods of protein modification are introduced that reinforce and complement the many accomplishments discussed in this introduction.

1.5 Thesis Objectives

The general goal of the investigations described in this Thesis is to adapt the tools of organic chemistry to protein modification. Chemistry is described that allows the installation of both natural and unnatural modifications on protein substrates. In Chapters 2, 3, and 4, the use of dehydroalanine in protein modification is examined. These studies were pursued with the aim of providing simple chemical access to proteins bearing natural post-translational modifications or their functional analogs. Glycosylation, polyprenylation, lysine methylation and acetylation, and phosphorylation are all modifications that are explored in these Chapters. For several of these natural modifications accessible through dehydroalanine, there is no straightforward biosynthetic alternative. Preliminary validation of these post-translational modification mimics in binding and enzymatic assays is also described.

In Chapters 5 and 6, unnatural modifications on proteins are reported. These studies rely on the use of transition metal chemistry to form carbon-carbon bonds on protein surfaces. In Chapter 5, the first example of olefin metathesis on a protein substrate is described. The key finding in this Chapter was the remarkable reactivity of allyl sulfides in olefin metathesis that allowed the use of this reaction in water and on proteins. This study was motivated in part by the many methods to introduce alkenes to proteins, and the unmet challenge of using such residues in metathesis. With olefin metathesis validated on protein substrates, several prospective uses for this transformation in protein engineering are also discussed.

Chapter 6 describes the discovery of a novel, easy to use catalyst for the modification of proteins by Suzuki-Miyaura cross-coupling. This project was motivated in part by the need for an active cross-coupling catalyst that is compatible with proteins. The successful application of cross-coupling to protein modification reported in this

Chapter vindicates the many efforts to incorporate cross-coupling partners genetically. The Suzuki-Miyaura reaction also confers several synthetic advantages over traditional methods for protein modification such as the potential for fluorogenic labeling and the facile transfer of hydrophobic groups to proteins. These and other prospects for cross-coupling in biological systems are discussed.

Chapter 7 is a brief project that was intimately related to the cross-coupling efforts described in Chapter 6. A simple, safe, and scalable method for the synthesis of Barluenga's reagent is described. This iodinating reagent was needed in large quantities for the selective iodination of tyrosine. A novel synthesis is described in this Chapter that avoids hazardous materials used in the original synthesis.

Chapter 8 is an intersection of an interest in dehydroalanine and organometallic chemistry. The key finding in this Chapter is that carbon-carbon bonds can be formed at dehydroalanine by a Zn-mediated radical addition of alkyl iodides. This chemistry allows the chemical installation of natural, unnatural, and post-translationally modified residues. The concept of "chemical mutagenesis" is explored in this Chapter—the site-selective interconversion of amino acid residues on folded proteins.

Throughout this thesis, the primary interest is the selective modification of protein substrates. Many of the findings, however, are useful for general organic synthesis. Where new reagents, reactions, and catalysts are reported, their use in both protein modification and general synthesis is discussed. Finally, the findings in this Thesis are by no means meant to replace any competing method in protein modification. Rather, the complementary and reinforcing use of chemical modification, genetic engineering, and total synthesis of proteins is advocated throughout this Thesis. Ultimately, the primary goal is to enable access to much needed proteins that are difficult to obtain by other methods. To meet this goal, it is instructive to take a cue from Emil Fischer and draw upon "all the expedients of the present; because only through this hazardous affair can the limitations of the ability of our methods be ascertained."^{167,204}

1.6 References

- (1) Means, G. E.; Feeney, R. E. Chemical modifications of proteins: history and applications. *Bioconjugate Chem.* **1990**, *1*, 2-12.
- (2) Steiner, R. F.; Edelhoch, H. Fluorescent protein conjugates. *Chem. Rev.* **1962**, *62*, 457-483.
- (3) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. The fluorescent toolbox for assessing protein location and function. *Science* **2006**, *312*, 217-224.
- (4) Prescher, J. A.; Bertozzi, C. R. Chemistry in living systems. *Nat. Chem. Biol.* **2005**, *1*, 13-21.
- (5) van Swieten, P. F.; Leeuwenburgh, M. A.; Kessler, B. M.; Overkleeft, H. S. Bioorthogonal organic chemistry in living cells: novel strategies for labeling biomolecules. *Org. Biomol. Chem.* **2005**, *3*, 20-27.
- (6) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974-6998.
- (7) Lim, R. K. V.; Lin, Q. Bioorthogonal chemistry: recent progress and future directions. *Chem. Commun.* **2010**, *46*, 1589-1600.
- (8) Saghatelian, A.; Cravatt, B. F. Assignment of protein function in the postgenomic era. *Nat. Chem. Biol.* **2005**, *1*, 130-142.
- (9) Kaiser, E. T.; Lawrence, D. S. Chemical mutation of enzyme active sites. *Science* **1984**, *226*, 505-511.
- (10) Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M. D. Generation of new enzymes via covalent modification of existing proteins. *Chem. Rev.* **2001**, *101*, 3081-3111.
- (11) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372.
- (12) Walsh, C. T. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*. Roberts and Company Publishers: Englewood, CO, 2006.
- (13) Gamblin, D. P.; van Kasteren, S. I.; Chalker, J. M.; Davis, B. G. Chemical approaches to mapping the function of post-translational modifications. *FEBS J.* **2008**, *275*, 1949-1959.
- (14) Walsh, G. Biopharmaceuticals: recent approvals and likely directions. *Trends Biotechnol.* **2005**, *23*, 553-558.

- (15) *Carbohydrate based vaccines*; Roy, R., Ed., 2008; Vol. 989.
- (16) Zhu, J.; Warren, J. D.; Danishefsky, S. J. Synthetic carbohydrate-based anticancer vaccines: the Memorial Sloan-Kettering experience. *Expert Rev. Vaccines* **2009**, 8, 1399-1413.
- (17) Mimura, Y.; Jefferis, R.; Mimura-Kimura, Y.; Abrahams, J.; Rudd, P. M. In *Therapeutic Monoclonal Antibodies: From Bench to Clinic*; An, Z., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, 2009; pp 67-89.
- (18) Walsh, G.; Jefferis, R. Post-translational modifications in the context of therapeutic proteins. *Nat. Biotech.* **2006**, 10, 1241-1252.
- (19) Doores, K. J.; Gamblin, D. P.; Davis, B. G. Exploring and exploiting the therapeutic potential of glycoconjugates. *Chem. Eur. J.* **2006**, 12, 656-665.
- (20) *Post-translational Modification of Protein Biopharmaceuticals*; Walsh, G., Ed.; Wiley-VCH Verlag GmbH & Co.: Weinheim, Germany, 2009.
- (21) Hermanson, G. T. *Bioconjugation Techniques, 2nd edition*. Academic Press, Inc., 2008.
- (22) Hackenberger, C. P. R.; Schwarzer, D. Chemoselective ligation and modification strategies for peptides and proteins. *Angew. Chem. Int. Ed.* **2008**, 47, 10030-10074.
- (23) Fodje, M. N.; Al-Karadaghi, S. Occurrence, conformational features and amino acid propensities for the π -helix. *Protein Eng.* **2002**, 15, 353-358.
- (24) Hopkins, S. J.; Wormald, A. Action of phenyl isocyanate on insulin. *Nature* **1934**, 134, 290.
- (25) Hopkins, S. J.; Wormald, A. Phenyl isocyanate protein derivatives and their immunological properties. III. The amino-acid derivatives and serological inhibition tests. *Biochem. J.* **1934**, 28, 228-236.
- (26) McBroom, C. R.; Samanen, C. H.; Goldstein, I. J. Carbohydrate antigens: coupling of carbohydrates to proteins by diazonium and phenylisothiocyanate reactions. *Methods Enzymol.* **1972**, 28, 212-219.
- (27) Edman, P. Method for determination of the amino acid sequence in peptides. *Acta Chem. Scand.* **1950**, 4, 283-293.
- (28) Edman, P. Preparation of phenyl thiohydantoins from some natural amino acids. *Acta Chem. Scand.* **1950**, 4, 277-282.
- (29) Astronomo, R. D.; Lee, H.-K.; Scanlan, C. N.; Pantophlet, R.; Huang, C.-Y.; Wilson, I. A.; Blixt, O.; Dwek, R. A.; Wong, C.-H.; Burton, D. R. A

- glycoconjugate antigen based on the recognition motif of a broadly neutralizing human immunodeficiency virus antibody, 2G12, is immunogenic but elicits antibodies unable to bind to the self glycans of gp120. *J. Virol.* **2008**, 82, 6359-6368.
- (30) Erlanger, B. F.; Borek, F.; Beiser, S. M.; Lieberman, S. Steroid-protein conjugates I. Preparation and characterization of conjugates of bovine serum albumin with testosterone and with cortisone. *J. Biol. Chem.* **1957**, 228, 713-727.
 - (31) Hughes, W. L.; Saroff, H. A.; Carney, A. L. Preparation and properties of serum and plasma proteins. XXII. A crystallizable guanidinated derivative of human serum albumin. *J. Am. Chem. Soc.* **1949**, 71, 2476-2480.
 - (32) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. N-Hydroxysuccinimide esters in peptide synthesis. *J. Am. Chem. Soc.* **1963**, 85, 3039.
 - (33) Anderson, G. W.; Zimmerman, J. E.; Callahan, F. M. The use of esters of N-hydroxysuccinimide in peptide synthesis. *J. Am. Chem. Soc.* **1964**, 86, 1839-1842.
 - (34) Becker, J. M.; Wilchek, M.; Katchalski, E. Irreversible inhibition of biotin transport in yeast by biotinyl-*p*-nitrophenyl ester. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, 68, 2604-2607.
 - (35) Becker, J. M.; Wilchek, M. Inactivation by avidin of biotin-modified bacteriophage. *Biochim. Biophys. Acta* **1972**, 264, 165-170.
 - (36) Tietze, L. F.; Arlt, M.; Beller, M.; Glösenkamp, K.-H.; Jähde, E.; Rajewsky, M. F. Squaric acid diethyl ester: a new coupling reagent for the formation of drug biopolymer conjugates. Synthesis of squaric acid ester amides and diamides. *Chem. Ber.* **1991**, 124, 1215-1221.
 - (37) Tietze, L. F.; Schröter, C.; Gabius, S.; Brinck, U.; Goerlach-Graw, A.; Gabius, H.-J. Conjugation of *p*-aminophenyl glycosides with squaric acid diester to a carrier protein and the use of neoglycoprotein in the histochemical detection of lectins. *Bioconjugate Chem.* **1991**, 2, 148-153.
 - (38) Lee, Y. C.; Stowell, C. P.; Krantz, M. J. 2-Imino-2-methoxyethyl 1-thioglycosides: new reagents for attaching sugars to proteins. *Biochemistry* **1976**, 15, 3956-3963.
 - (39) Means, G. E.; Feeney, R. E. Reductive alkylation of amino groups in proteins. *Biochemistry* **1968**, 7, 2192-2201.
 - (40) Gray, G. R. The direct coupling of oligosaccharides to proteins and derivatized gels. *Arch. Biochem. Biophys.* **1974**, 163, 426-428.

- (41) McFarland, J. M.; Francis, M. B. Reductive alkylation of proteins using iridium catalyzed transfer hydrogenation. *J. Am. Chem. Soc.* **2005**, *127*, 13490-13491.
- (42) Antos, J. M.; Francis, M. B. Transition metal catalyzed methods for site-selective protein modification. *Curr. Opin. Chem. Biol.* **2006**, *10*, 253-262.
- (43) Fraenkel-Conrat, H.; Olcott, H. S. Esterification of proteins with alcohols of low molecular weight. *J. Biol. Chem.* **1945**, *161*, 259-268.
- (44) Chibnall, A. C.; Mangan, J. L.; Rees, M. W. Studies on the amide and C-terminal residues in proteins. 3. The esterification of proteins. *Biochem. J.* **1958**, *68*, 114-118.
- (45) Sheehan, J. C.; Hlavka, J. J. The use of water-soluble and basic carbodiimides in peptide synthesis. *J. Org. Chem.* **1956**, *21*, 439-441.
- (46) Sheehan, J. C.; Hlavka, J. J. The cross-linking of gelatin using a water-soluble carbodiimide. *J. Am. Chem. Soc.* **1957**, *79*, 4528-4529.
- (47) Goodfriend, T. L.; Levine, L.; Fasman, G. D. Antibodies to bradykinin and angiotensin: a use of carbodiimides in immunology. *Science* **1964**, *144*, 1344-1346.
- (48) Moczar, E.; Leboul, J. Preparation of *N*-acetylglucosamine derivatives of proteins. *FEBS Lett.* **1975**, *50*, 300-302.
- (49) Crankshaw, M. W.; Grant, G. A. In *Curr. Protoc. Protein Sci.*; John Wiley & Sons, Inc. Unit 15.1, 1996.
- (50) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. Chemical modification of proteins at cysteine: opportunities in chemistry and biology. *Chem. Asian J.* **2009**, *4*, 630-640.
- (51) Dickens, F. Interaction of halogenacetates and SH compounds. The reaction of halogenacetic acids with glutathione and cysteine. The mechanism of iodoacetate poisoning of glyoxalase. *Biochem. J.* **1933**, *27*, 1141-1151.
- (52) Michaelis, L.; Schubert, M. P. The reaction of iodoacetic acid on mercaptans and amines. *J. Biol. Chem.* **1934**, *106*, 331-341.
- (53) Anson, M. L. The reactions of iodine and iodoacetamide with native egg albumin. *J. Gen. Physiol.* **1940**, *23*, 321-331.
- (54) Davis, N. J.; Flitsch, S. L. A novel method for the specific glycosylation of proteins. *Tetrahedron Lett.* **1991**, *32*, 6793-6796.

- (55) Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L. Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin. *Chem. Biol.* **2001**, *8*, 133-145.
- (56) Lindley, H. A new synthetic substrate for trypsin and its application to the determination of the amino-acid aequence of proteins. *Nature* **1956**, *178*, 647-648.
- (57) Raftery, M. A.; Cole, R. D. Tryptic cleavage at cysteinyl peptide bonds. *Biochem. Biophys. Res. Commun.* **1963**, *10*, 467-472.
- (58) Smith, H. B.; Hartman, F. C. Restoration of activity to catalytically deficient mutants of ribulosebisphosphate carboxylase/oxygenase by aminoethylation. *J. Biol. Chem.* **1988**, *263*, 4921-4925.
- (59) Smith, H. B.; Larimer, F. W.; Hartman, F. C. Subtle alteration of the active site of ribulose bisphosphate carboxylase/oxygenase by concerted site-directed mutagenesis and chemical modification. *Biochem. Biophys. Res. Commun.* **1988**, *152*, 579-584.
- (60) Yoshimura, T.; Matsushima, Y.; Tanizawa, K.; Sung, M.-H.; Yamauchi, T.; Wakayama, M.; Esaki, N.; Soda, K. Substitution of *S*-(β -aminoethyl)-cysteine for active-site lysine of thermostable aspartate aminotransferase. *J. Biochem.* **1990**, *108*, 699-700.
- (61) Zheng, R.; Dam, T. K.; Brewer, C. F.; Blanchard, J. S. Active site residues in *Mycobacterium tuberculosis* pantothenate synthetase required in the formation and stabilization of the adenylate intermediate. *Biochemistry* **2004**, *43*, 7171-7178.
- (62) Simon, M. D.; Chu, F.; Racki, L. R.; de la Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **2007**, *128*, 1003-1012.
- (63) Friedmann, E.; Marrian, D. H.; Simon-Reuss, I. Sulphydryl addition compounds of some quinones and related substances in their action on the growth of normal cells. *Brit. J. Pharmacol.* **1948**, *3*, 335-340.
- (64) Tsao, T.-C.; Bailey, K. The extraction, purification and some chemical properties of actin. *Biochim. Biophys. Acta* **1953**, *11*, 102-113.
- (65) Masri, M. S.; Friedman, M. Protein reactions with methyl and ethyl vinyl sulfones. *J. Protein Chem.* **1988**, *7*, 49-54.
- (66) Bader, B.; Kuhn, K.; Owen, D. J.; Waldmann, H.; Wittinghofer, A.; Kuhlmann, J. Bioorganic synthesis of lipid-modified proteins for the study of signal transduction. *Nature* **2000**, *403*, 223-226.

- (67) Giles, N. M.; Watts, A. B.; Giles, G. I.; Fry, F. H.; Littlechild, J. A.; Jacob, C. Metal and redox modulation of cysteine protein function. *Chem. Biol.* **2003**, *10*, 677-693.
- (68) Jacob, C.; Giles, G. I.; Giles, N. M.; Sies, H. Sulfur and selenium: the role of oxidation state in protein structure and function. *Angew. Chem. Int. Ed.* **2003**, *42*, 4742-4758.
- (69) Claiborne, A.; Yeh, J. I.; Mallett, T. C.; Luba, J.; Crane, E. J.; Charrier, V.; Parsonage, D. Protein-sulfenic acids: diverse roles for an unlikely player in enzyme catalysis and redox regulation. *Biochemistry* **1999**, *38*, 15407-15416.
- (70) Ellman, G. L. Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **1959**, *82*, 70-77.
- (71) Fontana, A.; Scoffone, E.; Benassi, C. A. Sulfenyl halides as modifying reagents for polypeptides and proteins. II. Modification of cysteinyl residues. *Biochemistry* **1968**, *7*, 980-986.
- (72) Davis, B. G.; Lloyd, R. C.; Jones, J. B. Controlled site-selective glycosylation of proteins by a combined site-directed mutagenesis and chemical modification approach. *J. Org. Chem.* **1998**, *63*, 9614-9615.
- (73) Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A.; Jones, J. B. Glycomethanethiosulfonates: powerful reagents for protein glycosylation. *Tetrahedron: Asymmetry* **2000**, *11*, 245-262.
- (74) Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glycosyl phenylthiosulfonates (Glyco-PTS): novel reagents for glycoprotein synthesis. *Org. Biomol. Chem.* **2003**, *1*, 3642-3644.
- (75) Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glyco-SeS: selenenylsulfide-mediated protein glycoconjugation--a new strategy in post-translational modification. *Angew. Chem. Int. Ed.* **2004**, *43*, 828-833.
- (76) Gamblin, D. P.; van Kasteren, S.; Bernardes, G. J. L.; Chalker, J. M.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Chemical site-selective prenylation of proteins. *Mol. BioSyst.* **2008**, *4*, 558-561.
- (77) Wender, P. A.; Goun, E. A.; Jones, L. R.; Pillow, T. H.; Rothbard, J. B.; Shinde, R.; Contag, C. H. Real-time analysis of uptake and bioactivatable cleavage of luciferin-transporter conjugates in transgenic reporter mice. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10340-10345.

- (78) Crich, D.; Brebion, F.; Krishnamurthy, V. Allylic disulfide rearrangement and desulfurization: mild, electrophile-free thioether formation from thiols. *Org. Lett.* **2006**, *8*, 3593-3596.
- (79) Crich, D.; Krishnamurthy, V.; Hutton, T. K. Allylic selenosulfide rearrangement: a method for chemical ligation to cysteine and other thiols. *J. Am. Chem. Soc.* **2006**, *128*, 2544-2545.
- (80) Crich, D.; Krishnamurthy, V.; Brebion, F.; Karatholuvhu, M.; Subramanian, V.; Hutton, T. K. Dechalcogenative allylic selenosulfide and disulfide rearrangements: complementary methods for the formation of allylic sulfides in the absence of electrophiles. Scope, limitations, and application to the functionalization of unprotected peptides in aqueous media. *J. Am. Chem. Soc.* **2007**, *129*, 10282-10294.
- (81) Crich, D.; Yang, F. Synthesis of neoglycoconjugates by the desulfurative rearrangement of allylic disulfides. *J. Org. Chem.* **2008**, *73*, 7017-7027.
- (82) Dondoni, A.; Massi, A.; Nanni, P.; Roda, A. A new ligation strategy for peptide and protein glycosylation: photoinduced thiol-ene coupling. *Chem. Eur. J.* **2009**, *15*, 11444-11449.
- (83) Balls, A. K.; Jansen, E. F. Stoichiometric inhibition of chymotrypsin. *Adv. Enzymol.* **1952**, *13*, 321-343.
- (84) Gold, A. M.; Fahrney, D. Sulfonyl fluorides as inhibitors of esterases. II. Formation and reactions of phenylmethanesulfonyl α -chymotrypsin. *Biochemistry* **1964**, *3*, 783-791.
- (85) Strumeyer, D. H.; White, W. N.; Koshland, D. E., Jr. Role of serine in chymotrypsin action. Conversion of the active serine to dehydroalanine. *Proc. Natl. Acad. Sci. U. S. A.* **1963**, *50*, 931-935.
- (86) Neet, K. E.; Koshland, D. E., Jr. The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation". *Proc. Natl. Acad. Sci. U. S. A.* **1966**, *56*, 1606-1611.
- (87) Polgár, L.; Bender, M. L. A new enzyme containing a synthetically formed active site. Thiol-subtilisin. *J. Am. Chem. Soc.* **1966**, *88*, 3153-3154.
- (88) Weiner, H.; White, W. N.; Hoare, D. G.; Koshland, D. E., Jr. The formation of anhydrochymotrypsin by removing the elements of water from the serine at the active site. *J. Am. Chem. Soc.* **1966**, *88*, 3851-3859.

- (89) Geoghegan, K. F.; Stroh, J. G. Site-directed conjugation of nonpeptide groups to peptides and proteins via periodate oxidation of a 2-amino alcohol. Application to modification at N-terminal serine. *Bioconjugate Chem.* **1992**, *3*, 138-146.
- (90) Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E. Nucleophilic catalysis of hydrazone formation and transimination: implications for dynamic covalent chemistry. *J. Am. Chem. Soc.* **2006**, *128*, 15602-15603.
- (91) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. Nucleophilic catalysis of oxime ligation. *Angew. Chem. Int. Ed.* **2006**, *45*, 7581-7584.
- (92) Hunter, W. M.; Greenwood, F. C. Preparation of iodine-131 labelled human growth hormone of high specific activity. *Nature* **1962**, *194*, 495-496.
- (93) Sigler, P. B. Iodination of a single tyrosine in crystals of α -chymotrypsin. *Biochemistry* **1970**, *9*, 3609-3617.
- (94) Ghosh, D.; Erman, M.; Sawicki, M.; Lala, P.; Weeks, D. R.; Li, N.; Pangborn, W.; Thiel, D. J.; Jörnvall, H.; Gutierrez, R.; Eyzaguirre, J. Determination of a protein structure by iodination: the structure of iodinated acetylxylan esterase. *Acta Cryst. D* **1999**, *55*, 779-784.
- (95) Dauter, Z. Phasing in iodine for structure determination. *Nat. Biotech.* **2004**, *22*, 1239-1240.
- (96) Espuña, G.; Andreu, D.; Barluenga, J.; Pérez, X.; Planas, A.; Arsequell, G.; Valencia, G. Iodination of proteins by IPy₂BF₄, a new tool in protein chemistry. *Biochemistry* **2006**, *45*, 5957-5963.
- (97) Joshi, N. S.; Whitaker, L. R.; Francis, M. B. A three-component Mannich-type reaction for selective tyrosine bioconjugation. *J. Am. Chem. Soc.* **2004**, *126*, 15942-15943.
- (98) Tilley, S. D.; Francis, M. B. Tyrosine-selective protein alkylation using π -allylpalladium complexes. *J. Am. Chem. Soc.* **2006**, *128*, 1080-1081.
- (99) Antos, J. M.; Francis, M. B. Selective tryptophan modification with rhodium carbenoids in aqueous solution. *J. Am. Chem. Soc.* **2004**, *126*, 10256-10257.
- (100) Antos, J. M.; McFarland, J. M.; Lavarone, A. T.; Francis, M. B. Chemoselective tryptophan labeling with rhodium carbenoids at mild pH. *J. Am. Chem. Soc.* **2009**, *131*, 6301-6308.
- (101) de Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconjugate Chem.* **2009**, *20*, 1281-1295.

- (102) van Hest, J. C. M.; Tirrell, D. A. Efficient introduction of alkene functionality into proteins in vivo. *FEBS Lett.* **1998**, *428*, 68-70.
- (103) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J. Am. Chem. Soc.* **2000**, *122*, 1282-1288.
- (104) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 19-24.
- (105) Link, A. J.; Tirrell, D. A. Cell surface labeling of *Escherichia coli* via copper(I)-catalyzed [3+2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125*, 11164-11165.
- (106) Sharma, N.; Furter, R.; Kast, P.; Tirrell, D. A. Efficient introduction of aryl bromide functionality into proteins in vivo. *FEBS Lett.* **2000**, *467*, 37-40.
- (107) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues. *ChemBioChem* **2002**, *3*, 235-237.
- (108) Datta, D.; Wang, P.; Carrico, I. S.; Mayo, S. L.; Tirrell, D. A. A designed phenylalanyl-tRNA synthetase variant allows efficient in vivo incorporation of aryl ketone functionality into proteins. *J. Am. Chem. Soc.* **2002**, *124*, 5652-5653.
- (109) Ohta, A.; Yamagishi, Y.; Suga, H. Synthesis of biopolymers using genetic code reprogramming. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159-167.
- (110) Xie, J.; Schultz, P. G. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.
- (111) Kwon, I.; Kirshenbaum, K.; Tirrell, D. A. Breaking the degeneracy of the genetic code. *J. Am. Chem. Soc.* **2003**, *125*, 7512-7513.
- (112) Rodriguez, E. A.; Lester, H. A.; Dougherty, D. A. *In vivo* incorporation of multiple unnatural amino acids through nonsense and frameshift suppression. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 8650-8655.
- (113) Neumann, H.; Wang, K.; Davis, L.; Garcia-Alai, M.; Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **2010**, *464*, 441-444.
- (114) Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **2000**, *4*, 645-652.
- (115) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-canonical amino acids in protein engineering. *Curr. Opin. Biotechnol.* **2003**, *14*, 603-609.

- (116) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **1989**, *244*, 182-188.
- (117) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. Expanding the genetic code of *Escherichia coli*. *Science* **2001**, *292*, 498-500.
- (118) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. Genetically encoding *N*-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **2008**, *4*, 232-234.
- (119) Chin, J. W.; Santoro, S. W.; Martin, A. B.; King, D. S.; Wang, L.; Schultz, P. G. Addition of *p*-azido-L-phenylalanine to the genetic code of *Escherichia coli*. *J. Am. Chem. Soc.* **2002**, *124*, 9026-9027.
- (120) Deiters, A.; Cropp, T. A.; Mukherji, M.; Chin, J. W.; Anderson, J. C.; Schultz, P. G. Adding amino acids with novel reactivity to the genetic code of *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **2003**, *125*, 11782-11783.
- (121) Chin, J. W.; Cropp, T. A.; Anderson, J. C.; Mukherji, M.; Zhang, Z.; Schultz, P. G. An expanded eukaryotic genetic code. *Science* **2003**, *301*, 964-967.
- (122) Nguyen, D. P.; Lusic, H.; Neumann, H.; Kapadnis, P. B.; Deiters, A.; Chin, J. W. Genetic encoding and labeling of aliphatic azides and alkynes in recombinant proteins via a pyrrolysyl-tRNA synthetase/tRNA pair and click chemistry. *J. Am. Chem. Soc.* **2009**, *131*, 8720-8721.
- (123) Dieters, A.; Schultz, P. G. In vivo incorporation of an alkyne into proteins in *Escherichia coli*. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 1521-1524.
- (124) Staudinger, H.; Meyer, J. Über neue organische phosphorverbindungen III. Phosphinmethylderivate und phosphinimine. *Helv. Chim. Acta* **1919**, *2*, 635-646.
- (125) Saxon, E.; Bertozzi, C. R. Cell surface engineering by a modified Staudinger reaction. *Science* **2000**, *287*, 2007-2010.
- (126) Michael, A. Ueber die einwirkung von diazobenzolimid auf acetylendicarbonsäuremethylester. *J. Prakt. Chem.* **1893**, *48*, 94-95.
- (127) Huisgen, R. 1,3-Dipolar cycloadditions past and future. *Angew. Chem. Int. Ed.* **1963**, *2*, 565-598.
- (128) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596-2599.

- (129) Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057-3064.
- (130) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click chemistry: diverse chemical function from a few good reactions. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004-2021.
- (131) Wittig, G.; Krebs, A. Zur existenz niedergliedriger cycloalkine, I. *Chem. Ber.* **1961**, *94*, 3260-3275.
- (132) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. A strain-promoted [3+2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047.
- (133) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. Bioconjugation by copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193.
- (134) Speers, A. E.; Adam, G. C.; Cravatt, B. F. Activity-based profiling in vivo using a copper(I)-catalyzed azide-alkyne [3+2] cycloaddition. *J. Am. Chem. Soc.* **2003**, *125*, 4686-4687.
- (135) Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. Copper-free click chemistry for dynamic *in vivo* imaging. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16793-16797.
- (136) Wang, L.; Zhang, Z.; Brock, A.; Schultz, P. G. Addition of the keto functional group to the genetic code of *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 56-61.
- (137) Carrico, I. S.; Carlson, B. L.; Bertozzi, C. R. Introducing genetically encoded aldehydes into proteins. *Nat. Chem. Biol.* **2007**, *3*, 321-322.
- (138) Wittrock, S.; Becker, T.; Kunz, H. Synthetic vaccines of tumor-associated glycopeptide antigens by immune-compatible thioether linkage to bovine serum albumin. *Angew. Chem. Int. Ed.* **2007**, *46*, 5226-5230.
- (139) Jonkheijm, P.; Weinrich, D.; Köhn, M.; Engelkamp, H.; Christianen, P. C. M.; Kuhlmann, J.; Maan, J. C.; Nüsse, D.; Schroeder, H.; Wacker, R.; Breinbauer, R.; Niemeyer, C. M.; Waldmann, H. Photochemical surface patterning by the thiol-ene reaction. *Angew. Chem. Int. Ed.* **2008**, *47*, 4421-4424.
- (140) Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G. Thiyl glycosylation of olefinic proteins: S-linked glycoconjugate synthesis. *Angew. Chem. Int. Ed.* **2009**, *48*, 7798-7802.

- (141) Zhang, Z.; Wang, L.; Brock, A.; Schultz, P. G. The selective incorporation of alkenes into proteins in *Escherichia coli*. *Angew. Chem. Int. Ed.* **2002**, *41*, 2840-2842.
- (142) Clovis, J. S.; Eckell, A.; Huisgen, R.; Sustmann, R. 1,3-Dipolare cycloadditionen, XXV. Der nachweis des freien diphenylnitrilimins als zwischenstufe bei cycloadditionen. *Chem. Ber.* **1967**, *100*, 60-70.
- (143) Song, W.; Wang, Y.; Qu, J.; Lin, Q. Selective functionalization of a genetically encoded alkene-containing protein via "photoclick chemistry" in bacterial cells. *J. Am. Chem. Soc.* **2008**, *130*, 9654-9655.
- (144) Song, W.; Wang, Y.; Qu, J.; Madden, M. M.; Lin, Q. A photoinducible 1,3-dipolar cycloaddition reaction for rapid, selective modification of tetrazole-containing proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 2832-2835.
- (145) Grubbs, R. H. *Handbook of Metathesis*. Wiley-VCH: Weinheim, Germany, 2003.
- (146) Fürstner, A. Olefin metathesis and beyond. *Angew. Chem. Int. Ed.* **2000**, *39*, 3012-3043.
- (147) Hoveyda, A. H.; Zhugralin, A. R. The remarkable metal-catalysed olefin metathesis reaction. *Nature* **2007**, *450*, 243-251.
- (148) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Metathesis reactions in total synthesis. *Angew. Chem. Int. Ed.* **2005**, *44*, 4490-4527.
- (149) *Handbook of organopalladium chemistry for organic synthesis*. Negishi, E.-i. (Ed.); Wiley Interscience, New York, 2002.
- (150) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Palladium-catalyzed cross-coupling reactions in total synthesis. *Angew. Chem. Int. Ed.* **2005**, *44*, 4442-4489.
- (151) Sakamoto, K.; Hayashi, A.; Sakamoto, A.; Kiga, D.; Nakayama, H.; Soma, A.; Kobayashi, T.; Kitabatake, M.; Takio, K.; Saito, K.; Shirouzu, M.; Hirao, I.; Yokoyama, S. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucl. Acids Res.* **2002**, *30*, 4692-4699.
- (152) Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. The site-specific incorporation of *p*-iodo-L-phenylalanine into proteins for structure determination. *Nat. Biotechnol.* **2004**, *22*, 1297-1301.
- (153) Kodama, K.; Fukuzawa, S.; Sakamoto, K.; Nakayama, H.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. A new protein engineering approach combining chemistry and biology, part I; site-

- specific incorporation of 4-Iodo-L-phenylalanine in vitro by using misacylated suppressor tRNA^{Phe}. *ChemBioChem* **2006**, *7*, 1577-1581.
- (154) Kwon, I.; Wang, P.; Tirrell, D. A. Design of a bacterial host for site-specific incorporation of p-bromophenylalanine into recombinant proteins. *J. Am. Chem. Soc.* **2006**, *128*, 11778-11783.
- (155) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. A genetically encoded boronate-containing amino acid. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220-8223.
- (156) Baldwin, A. J.; Arpino, J. A. J.; Edwards, W. R.; Tippmann, E. M.; Jones, D. D. Expanded chemical diversity sampling through whole protein evolution. *Mol. BioSyst.* **2009**, *5*, 764-766.
- (157) Hancock, S. M.; Uprety, R.; Deiters, A.; Chin, J. W. Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J. Am. Chem. Soc.* **2010**, *132*, 14819-14824.
- (158) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. Regioselective carbon-carbon bond formation in proteins with palladium catalysis; new protein chemistry by organometallic chemistry. *ChemBioChem* **2006**, *7*, 134-139.
- (159) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. Site-specific functionalization of proteins by organopalladium reactions. *ChemBioChem* **2007**, *8*, 232-238.
- (160) Carrico, Z. M.; Romanini, D. W.; Mehl, R. A.; Francis, M. B. Oxidative coupling of peptides to a virus capsid containing unnatural amino acids. *Chem. Commun.* **2008**, 1205-1207.
- (161) Blackman, M. L.; Royzen, M.; Fox, J. M. Tetrazine ligation: fast bioconjugation based on inverse-electron-demand Diels-Alder reactivity. *J. Am. Chem. Soc.* **2008**, *130*, 13518-13519.
- (162) Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G. A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* **2004**, *126*, 14306-14307.
- (163) Fischer, E.; Fourneau, E. Ueber einige derivate des glykocolis. *Ber.* **1901**, *34*, 2868-2877.

- (164) Bergmann, M.; Zervas, L. Über ein allgemeines verfahren der peptid-synthese. *Ber.* **1932**, *65*, 1192-1201.
- (165) Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.
- (166) Merrifield, R. B.; Morrow Stewart, J.; Jernberg, N. Instrument for automated synthesis of peptides. *Anal. Chem.* **1966**, *38*, 1905-1914.
- (167) Merrifield, B. Solid phase synthesis. *Science* **1986**, *232*, 341-347.
- (168) Gutte, B.; Merrifield, R. B. The total synthesis of an enzyme with ribonuclease A activity. *J. Am. Chem. Soc.* **1969**, *91*, 501-502.
- (169) Denkewalter, R. G.; Veber, D. F.; Holly, F. W.; Hirschmann, R. Studies on the total synthesis of an enzyme. I. Objective and strategy. *J. Am. Chem. Soc.* **1969**, *91*, 502-503.
- (170) Strachan, R. G.; Paleveda, W. J.; Nutt, R. F.; Vitali, R. A.; Veber, D. F.; Dickinson, M. J.; Garsky, V.; Deak, J. E.; Walton, E.; Jenkins, S. R.; Holly, F. W.; Hirschmann, R. Studies on the total synthesis of an enzyme. II. Synthesis of a protected tetratetracontapeptide corresponding to the 21-64 sequence of Ribonuclease A. *J. Am. Chem. Soc.* **1969**, *91*, 503-505.
- (171) Jenkins, S. R.; Nutt, R. F.; Dewey, R. S.; Veber, D. F.; Holly, F. W.; Paleveda, W. J.; Lanza, T.; Strachan, R. G.; Schoenewaldt, E. F.; Barkemeyer, H.; Dickinson, M. J.; Sondey, J.; Hirschmann, R.; Walton, E. Studies on the total synthesis of an enzyme. III. Synthesis of a protected hexacontapeptide corresponding to the 65-124 sequence of ribonuclease A. *J. Am. Chem. Soc.* **1969**, *91*, 505-506.
- (172) Veber, D. F.; Varga, S. L.; Milkowski, J. D.; Joshua, H.; Conn, J. B.; Hirschmann, R.; Denkewalter, R. G. Studies on the total synthesis of an enzyme. IV. Some factors affecting the conversion of protected S-protein to ribonuclease S'. *J. Am. Chem. Soc.* **1969**, *91*, 506-507.
- (173) Hirschmann, R.; Nutt, R. F.; Veber, D. F.; Vitali, R. A.; Varga, S. L.; Jacob, T. A.; Holly, F. W.; Denkewalter, R. G. Studies on the total synthesis of an enzyme. V. The preparation of enzymatically active material. *J. Am. Chem. Soc.* **1969**, *91*, 507-508.
- (174) Schnölzer, M.; Kent, S. B. H. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* **1992**, *256*, 221-225.

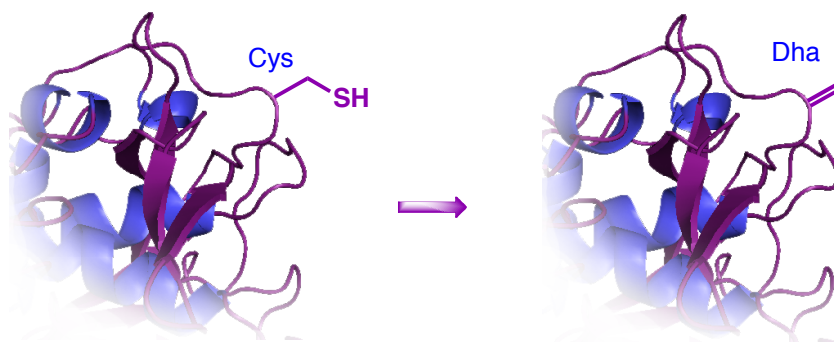
- (175) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, 38, 338-351.
- (176) Kent, S. B. H. Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **1988**, 57, 957-989.
- (177) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **1994**, 266, 776-779.
- (178) Wieland, T.; Bokelmann, E.; Bauer, L.; Lang, H. U.; Lau, H. Über peptidsynthesen. 8. Mitteilung. Bildung von S-haltigen peptiden durch intramolekulare wanderung von amionoacylresten. *Liebigs Ann. Chem.* **1953**, 583, 129-149.
- (179) Dawson, P. E.; Kent, S. B. H. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **2000**, 69, 923-960.
- (180) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, 34, 91-118.
- (181) Evans, T. C.; Benner, J.; Xu, M.-Q. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **1998**, 7, 2256-2264.
- (182) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, 95, 6705-6710.
- (183) Xu, M.-Q.; Perler, F. B. The mechanism of protein splicing and its modulation by mutation. *EMBO J.* **1996**, 15, 5146-5153.
- (184) Paulus, H. Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem.* **2000**, 69, 447-496.
- (185) Chong, S.; Mersha, F. B.; Comb, D. G.; Scott, M. E.; Landry, D.; Vence, L. M.; Perler, F. B.; Benner, J.; Kucera, R. B.; Hirvonen, C. A.; Pelletier, J. J.; Paulus, H.; Xu, M.-Q. Single-column purification of free recombinant proteins using a self-cleavable affinity tag derived from a protein splicing element. *Gene* **1997**, 192, 271-281.
- (186) Severinov, K.; Muir, T. W. Expressed protein ligation, a novel method for studying protein-protein interactions in transcription. *J. Biol. Chem.* **1998**, 273, 16205-16209.
- (187) Schwarzer, D.; Cole, P. A. Protein semisynthesis and expressed protein ligation: chasing a protein's tail. *Curr. Opin. Chem. Biol.* **2005**, 9, 561-569.

- (188) Flavell, R. R.; Muir, T. W. Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology. *Acc. Chem. Res.* **2009**, *42*, 107-116.
- (189) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. Staudinger ligation: a peptide from a thioester and azide. *Org. Lett.* **2000**, *2*, 1939-1941.
- (190) Nilsson, B. L.; Kiessling, L. L.; Raines, R. T. High-yielding Staudinger ligation of a phosphinothioester and azide to form a peptide. *Org. Lett.* **2001**, *3*, 9-12.
- (191) Nilsson, B. L.; Hondal, R. J.; Soellner, M. B.; Raines, R. T. Protein assembly by orthogonal chemical ligation methods. *J. Am. Chem. Soc.* **2003**, *125*, 5268-5269.
- (192) Canne, L. E.; Bark, S. J.; Kent, S. B. H. Extending the applicability of native chemical ligation. *J. Am. Chem. Soc.* **1996**, *118*, 5891-5896.
- (193) Botti, P.; Carrasco, M. R.; Kent, S. B. H. Native chemical ligation using removable *N* α -(1-phenyl-2-mercaptoethyl) auxiliaries. *Tetrahedron Lett.* **2001**, *42*, 1831-1833.
- (194) Offer, J.; Boddy, C. N. C.; Dawson, P. E. Extending synthetic access to proteins with a removable acyl transfer auxiliary. *J. Am. Chem. Soc.* **2002**, *124*, 4642-4646.
- (195) Yan, L. Z.; Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526-533.
- (196) Crich, D.; Banerjee, A. Native chemical ligation at phenylalanine. *J. Am. Chem. Soc.* **2007**, *129*, 10064-10065.
- (197) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-9252.
- (198) Haase, C.; Rohde, H.; Seitz, O. Native chemical ligation at valine. *Angew. Chem. Int. Ed.* **2008**, *47*, 6807-6810.
- (199) Yang, R.; Kumar Pasunooti, K.; Li, F.; Liu, X.-W.; Liu, C.-F. Dual native chemical ligation at lysine. *J. Am. Chem. Soc.* **2009**, *131*, 13592-13593.
- (200) Kumar, K. S. A.; Haj-Yahya, M.; Olschewski, D.; Lashuel, H. A.; Brik, A. Highly efficient and chemoselective peptide ubiquitylation. *Angew. Chem. Int. Ed.* **2009**, *48*, 8090-8094.
- (201) Harpaz, Z.; Siman, P.; Kumar, K. S. A.; Brik, A. Protein synthesis assisted by native chemical ligation at leucine. *ChemBioChem* **2010**, *11*, 1232-1235.

- (202) El Oualid, F.; Merkx, R.; Ekkebus, R.; Hameed, D. S.; Smit, J. J.; de Jong, A.; Hilkmann, H.; Sixma, T. K.; Ova, H. Chemical synthesis of ubiquitin, ubiquitin-based probes, and diubiquitin. *Angew. Chem. Int. Ed.* **2010**, *49*, 10149-10153.
- (203) Macmillan, D. Evolving strategies for protein synthesis converge on native chemical ligation. *Angew. Chem. Int. Ed.* **2006**, *45*, 7668-7672.
- (204) Fischer, E. Untersuchungen über aminosäuren, polypeptide und proteine. *Ber.* **1906**, *39*, 530-610.

Chapter 2

Methods for the Chemical Conversion of Cysteine to Dehydroalanine on Protein Substrates



“...denn nur durch das Wagniss selbst kann die Grenze für die Leistungsfähigkeit unserer Methoden ermittelt werden.”

Emil Fischer

Ber. 1906, 39, 530

2.1 Introduction

Dehydroalanine: a precursor to post-translational modifications

Dehydroalanine (Dha) is an amino acid of both biological and synthetic interest. Dehydroalanine is found naturally in lanthionine containing antibiotic peptides—the so-called lantibiotics—and is formed by the enzymatic dehydration of serine.¹⁻³ Dehydroalanine imparts conformational constraints to these peptides and is a critical determinant in peptide topology.⁴ The electrophilic properties of dehydroalanine are also exploited in nature. In lantibiotics, intramolecular conjugate addition of cysteine to dehydroalanine results in the formation of lanthionine. Dehydroalanine and lanthionine are depicted in a representative lantibiotic (Nisin) in Figure 2.1.¹⁻³

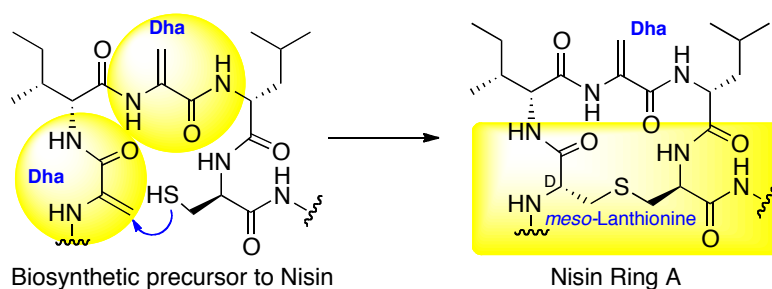


Figure 2.1: Dehydroalanine is an amino acid found in lantibiotics

Dehydroalanine is also found, albeit more rarely, in proteins. In thyroglobulin, for example, iodinated tyrosine is converted to dehydroalanine during thyroid hormone synthesis.^{5,6} The mechanistic course for this transformation is outlined in Figure 2.2.

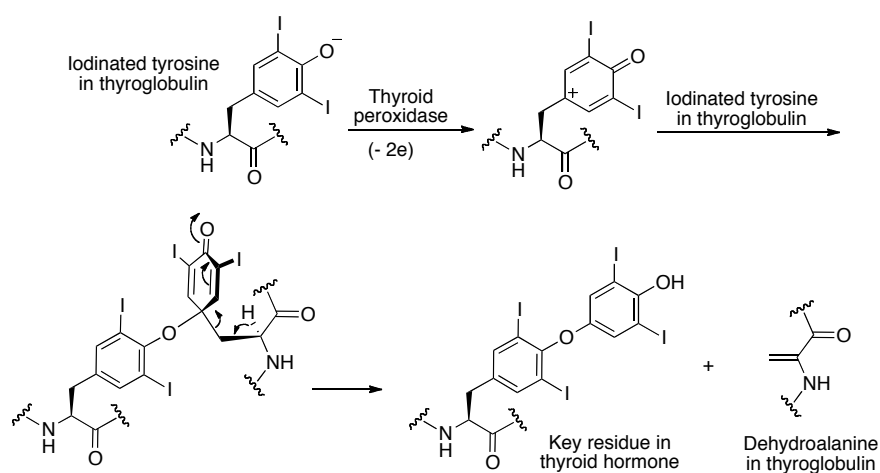


Figure 2.2: Dehydroalanine is formed in thyroglobulin during hormone biosynthesis

In addition to these natural occurrences of dehydroalanine, this residue is useful in the chemical modification of peptides and proteins. The addition of thiol nucleophiles to dehydroalanine provides access to several post-translational modifications and their mimics. Glycosylated and polyprenylated cysteine residues are two such modifications accessible through dehydroalanine (Figure 2.3).⁷⁻¹⁰

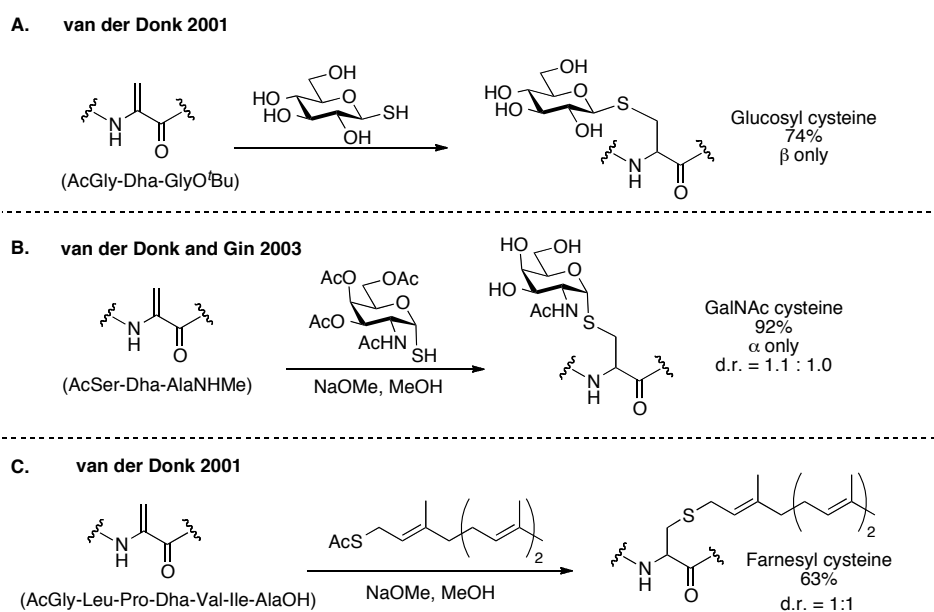


Figure 2.3: Addition of thiol nucleophiles to dehydroalanine enables access to post-translational modifications on peptides and proteins. Representative examples of modified peptides are shown here.^{7,9}

Glycosyl cysteines are naturally occurring¹¹⁻¹³ and are structurally similar to the more common serine-linked glycoproteins.¹⁴ Glycoproteins mediate many aspects of protein folding, cell signaling, inflammation, and immune response.^{14,15} The understanding of glycoproteins is therefore critical in the assessment and treatment of human disease.¹⁶ One central goal in the effort to understand glycoproteins is the synthesis of well-defined glycoproteins.^{17,18} The conjugate addition of glycosyl thiols to dehydroalanine is one method to access such structures (Figure 2.3 A and B). In the examples in Figure 2.3 it is noteworthy that deprotected carbohydrates can be used efficiently. Furthermore, while the diastereoselectivity of the conjugate addition is typically low, high α or β selectivity is observed in certain cases for glycosyl thiol nucleophiles (Figure 2.3 A and B). Finally, thioglycosides are generally more stable to

hydrolysis than their *O*-linked counterparts¹⁹ and the stability of *S*-linked glycoconjugates may therefore confer some metabolic or therapeutic advantage.²⁰

Polyprenylated cysteine residues can also be synthesized from dehydroalanine (Figure 2.3C). In nature, hydrophobic polyprenes²¹ and palmitoyl groups²² attached to cysteine mediate protein-membrane association, as well as various aspects of protein-protein interactions. The synthesis of these cysteine derivatives allows access to these modified proteins that are otherwise difficult to obtain from natural sources in pure form. Dehydroalanine is a convenient precursor to these lipidated cysteine residues.

Access to peptides and proteins bearing post-translational modifications is a current challenge in chemical biology.²³⁻²⁵ As seen in Figure 2.3, dehydroalanine allows access to glycosylated and prenylated peptides and other post-translational modifications are likely possible. A more comprehensive evaluation of dehydroalanine as a precursor to post-translational modifications is undertaken in Chapter 3. In order to take advantage of dehydroalanine as a precursor to post-translational modifications, its efficient incorporation into peptides and proteins is essential. The dehydroalanine residues in Figure 2.3, for instance, were installed by the oxidative elimination of phenylselenocysteine.^{7,9,26-28} There are, however, several other reported methods for the synthesis of dehydroalanine in peptides and proteins. Below is a survey of methods for the synthesis of dehydroalanine and a discussion of the associated scope and limitations. This assessment will help frame a strategy for the design of general methods for the incorporation of dehydroalanine into proteins—a central goal of the efforts described in this Chapter.

Previous reports describing the synthesis of dehydroalanine in peptides and proteins

The chemical installation of dehydroalanine in proteins has been studied for nearly half a century. In 1963, Koshland described the conversion of the active site serine of chymotrypsin to dehydroalanine.²⁹ This transformation was accomplished by the selective sulfonylation of the active site serine with tosyl chloride followed by elimination with sodium hydroxide (Figure 2.4A). In this report, the chemical dehydration of the active site serine was used to investigate this residue's role in catalysis. In a later report, Koshland and co-workers proposed that dehydroalanine could be used as an acceptor for nucleophiles and that these “chemical mutations” could be used to precisely investigate the function of individual residues on a protein.³⁰ Indeed the Koshland labs used the

conjugate addition of thiols to corroborate the formation of dehydroalanine in chymotrypsin.³⁰

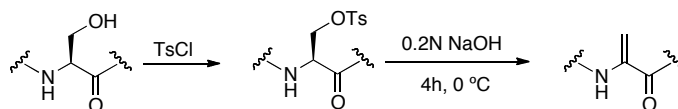
These seminal reports by the Koshland were critical in the elucidation of the catalytic mechanism of serine proteases. However, this method for installing dehydroalanine is unique to hydrolases that possess a highly nucleophilic serine in the active site. For a more general method of dehydroalanine incorporation, other groups have turned to cysteine as a precursor to dehydroalanine. Cysteine is a highly nucleophilic residue and can be selectively modified by a variety of chemical transformations.³¹ Lawton and Holmes, for example, described the conversion of cysteine to dehydroalanine as part of a method to map cysteine residues in proteins. Cysteine was dialkylated and the resulting sulfonium salt was eliminated to dehydroalanine (Figure 2.4B). Dehydroalanine hydrolyzes when heated in acetic acid and the resulting peptide fragments of this chemical digest were used to determine protein sequences.³² A related transformation was recently observed by Kajihara and co-workers as an unwanted reaction in their studies in the synthesis of complex glycopeptides.³³ In this report, the reaction of the diethyl ester of *meso*-2,5-dibromoadipic acid with cysteine yielded dehydroalanine (Figure 2.4C).

The oxidative elimination of alkyl cysteines is another method for the conversion of cysteine derivatives to dehydroalanine.³⁴ However, this reaction typically requires elevated temperatures not compatible with proteins: in Rich's original report the elimination of alkyl cysteine sulfoxides was carried out in refluxing xylene.³⁴ A much milder alternative is the oxidative elimination of alkyl selenocysteines. Phenylselenocysteine and related derivatives undergo oxidative elimination at room temperature in the presence of hydrogen peroxide (Figure 2.4D).²⁸ As mentioned above, this mild method for the incorporation of dehydroalanine has been successfully applied to both peptides^{7-9,26-28,35} and proteins.¹⁰ Notably, Seebeck and Szostak have shown that selenolysine analogs can be incorporated into peptides as a lysine surrogate using *E. coli* translational machinery.³⁵ The Schultz laboratory has reported that phenylselenocysteine can also be incorporated into proteins biosynthetically in response to the amber stop codon using an engineered aminoacyl-tRNA synthetase / tRNA pair.¹⁰ Both methods allow a biosynthetic route to dehydroalanine (Figure 2.4D).

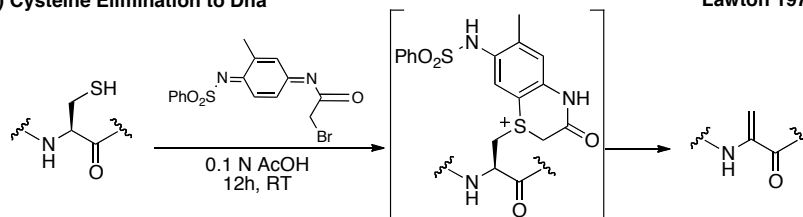
Recently, van der Donk has described an enzymatic route to dehydroalanine that was inspired by the biosynthesis of lantibiotics. Lactacin 481 synthetase (LctM) was used to phosphorylate serine in a variety of peptides attached to a leader sequence. Subsequent

A) Serine Elimination to Dha

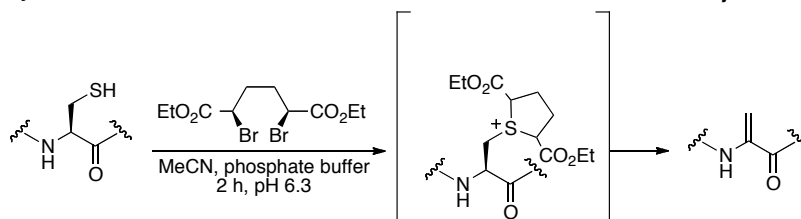
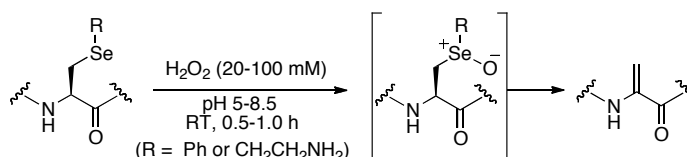
Koshland 1963

**B) Cysteine Elimination to Dha**

Lawton 1977

**C) Cysteine Elimination to Dha**

Kajihara 2009

**D) Alkyl- and Arylselenocysteine elimination to Dha**Shirahama 1996
van der Donk 2000
Szostak 2006
Schultz 2007**E) Enzymatic Elimination to Dha**

van der Donk 2009

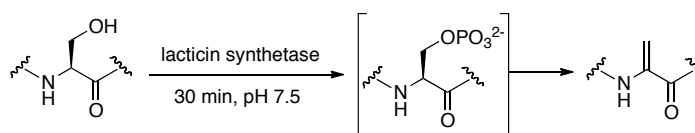


Figure 2.4: Representative methods for the synthesis of dehydroalanine in peptides and proteins

elimination of phosphoserine, also catalyzed by LctM, provided the dehydroalanine-containing peptides (Figure 2.4E).³⁶

Despite an enduring interest in the synthesis of dehydroalanine in peptides and proteins, no single method enables a generally chemo- and site-selective incorporation of dehydroalanine. Koshland's method is specific for hydrolases and requires a high pH that may not be compatible with many proteins (Figure 2.4A). Lawton and Kajihara both use reagents that are not soluble in water (Figure 2.4 B and C)—a clear limitation for most proteins that require buffered aqueous media and may precipitate in the presence of

organic solvent. The oxidative elimination of alkylselenocysteine derivatives (Figure 2.4D) relies on an oxidant (typically hydrogen peroxide) that will oxidize other residues such as methionine.^{10,35} Finally, enzymatic methods for the dehydration of serine (Figure 2.4E) require a leader sequence and may not be directly applicable to any given peptide or protein substrate without that sequence.

In this Chapter, several methods are proposed for the chemical conversion of cysteine to dehydroalanine. These methods were investigated with an aim to overcome the limitations of the techniques in Figure 2.4. Cysteine is a convenient target because of its high nucleophilicity relative to other proteinogenic amino acids. Moreover, cysteine can be incorporated at the site of modification by standard site-directed mutagenesis. Four mechanistically distinct modes of elimination are outlined below. These proposed transformations were then investigated on peptides and proteins with careful attention paid to issues of chemoselectivity and aqueous- and bio-compatibility.

Proposed methods for the conversion of cysteine to dehydroalanine

Four methods for the conversion of cysteine to dehydroalanine are outlined in Figure 2.5. The first method—the tandem reduction-elimination of cysteine disulfides—is based on the reaction of electron rich phosphines with disulfides (Figure 2.5A). The reaction of phosphines and disulfides is an endothermic process³⁷ and, according to the Leffler-Hammond postulate,^{38,39} proceeds through a late transition state. The reaction of phosphines with mixed disulfides should therefore proceed in such a way that the thiol with the lowest pK_a will be the leaving group. It follows that by making a mixed disulfide from an electron-poor thiol and cysteine, phosphine attack on the cysteine sulfur should be faster than attack on the sulfur of the electron-poor thiol. The resulting thiophosphonium salt can then eliminate to dehydroalanine. Harpp and Gleason reported a related elimination from β -keto disulfides using electron rich phosphines,⁴⁰ Xian and co-workers observed the formation of dehydroalanine when cystine was reacted with tris(dimethylamino)phosphine (HMPT),⁴¹ and our own group has observed such an elimination on cysteine-carbohydrate disulfides in the presence of HMPT.⁴²

It was thought that the thiophosphonium salt in Figure 2.5A would eliminate more efficiently than the disulfide alone. However, we did not discount the possibility that a cysteine disulfide could be eliminated directly to dehydroalanine. Thus, the direct elimination of a cysteine disulfide was investigated as a route to dehydroalanine (Figure 2.5B).

A third proposed method for the conversion of cysteine to dehydroalanine is direct oxidative elimination. Rather than use an electrophilic oxygen source that is typically used in the oxidation of thioethers, here it is proposed that an electrophilic nitrogen source may be more efficient. Reports by Claus⁴³ and later by Matsu and co-workers⁴⁴ demonstrated that the *syn*-elimination of sulfilimines (the nitrogen analog of a sulfoxide) proceeds at room temperature while the same elimination with sulfoxides requires much higher temperature. The source of electrophilic nitrogen used in these reports was Tamura's reagent, *O*-mesitylenesulfonylhydroxylamine (MSH, **2.1**).^{45,46} It was not clear at the outset if MSH would react with thiols to give alkenes without the formation of disulfides, but the possibility was intriguing. One possible mechanistic course for the proposed transformation is outlined in Figure 2.5C.

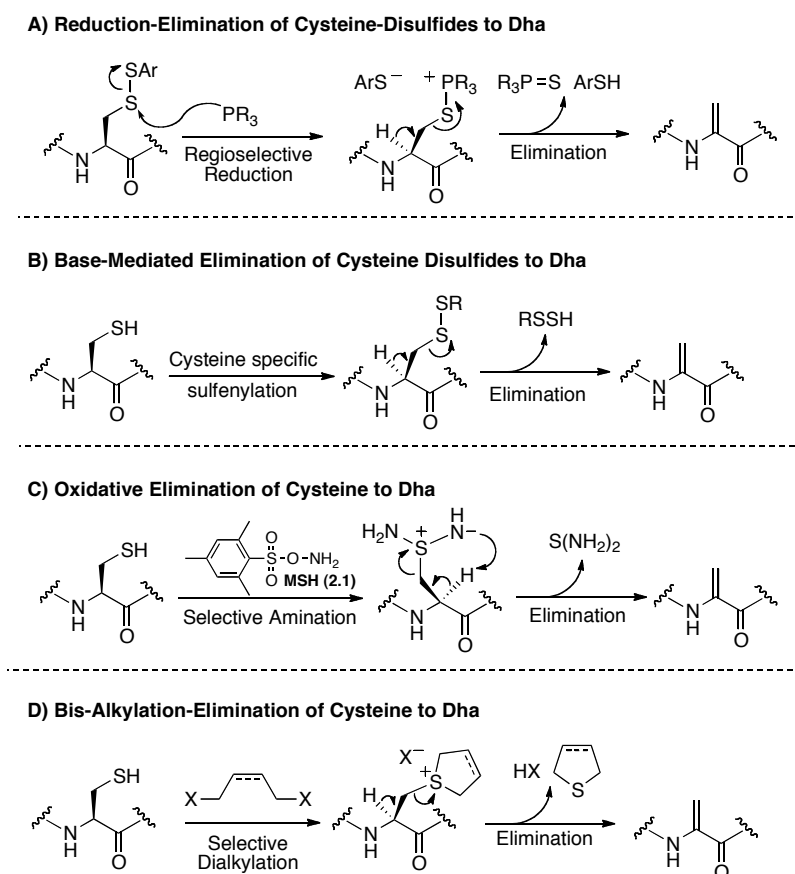


Figure 2.5: Proposed methods for the conversion of cysteine to dehydroalanine

A fourth method for the synthesis of dehydroalanine considered here is the sequential bis-alkylation-elimination of cysteine. This method is inspired in part by the reactions reported by Holmes, Lawton, and Kajihara (Figure 2.4 B and C).^{32,33} It is

noteworthy that related transformations have been observed in both murine and human metabolism of 1,4-dihalobutanes and Busulfan (1,4-bis(methansulfonyl)-butandiol) where the cysteine of glutathione is converted to dehydroalanine.⁴⁷⁻⁵² The occurrence of these reactions under physiological conditions bodes well for the application of these transformations to protein substrates.

The transformations outlined in Figure 2.5 were first investigated on model amino acid and peptide substrates. Careful attention was paid to side reactions, substrate scope, and aqueous compatibility. After an assessment of each method on model systems, these reactions were explored on protein substrates. The most general methods enabled the site-selective synthesis of dehydroalanine in several proteins. With methods for incorporation of dehydroalanine, the use of dehydroalanine as a general precursor to several classes of post-translational modifications could be investigated, as will be seen in Chapter 3.

2.2 Results and Discussion

Conversion of cysteine to dehydroalanine on amino acid and peptide substrates

Conversion of cysteine disulfides to dehydroalanine under reducing conditions

To assess the reduction-elimination strategy outlined in Figure 2.5A, disulfide **2.2** was first synthesized using the methyl ester of Ellman's reagent (**2.3**, Figure 2.6). The mixed disulfide **2.2** possesses the requisite electron-withdrawing leaving group (the thiol derived from Ellman's reagent) that is required for regioselective attack by a phosphine at the sulfur of cysteine. Ellman's reagent itself is advantageous since it is water soluble and has long been used for specific reaction with cysteine in protein samples.⁵³ Tris(dimethylamino)phosphine (HMPT) was used as the phosphine in these test reactions. HMPT is electron rich and renders the intermediate thiophosphonium salt resistant to hydrolysis.

With **2.2** in hand, the reaction was tested. Immediately upon the addition of two equivalents of HMPT, the reaction mixture turned red, consistent with phosphine attack on the disulfide and release of the aryl thiolate. Potassium carbonate was used as the base in this reaction and undried THF as the solvent. The major product of this reaction was AcDhaOMe (**2.4**), which was isolated in 85% yield. Phosphoramidate **2.5** was also isolated from the reaction mixture in 46% yield. The formation of **2.5** is consistent with previous

observations by Harpp and Gleason where thiols attacked the phosphineoxide or (thio)oxide of HMPT or HMPT itself.⁴⁰ Importantly, the product of simple reduction (AcCysOMe) was not observed. It should also be pointed out that carbonate alone is not a strong enough base to eliminate the disulfide to dehydroalanine (carbonate was used to synthesize **2.2** and no dehydroalanine was observed in this synthesis). Therefore the elimination occurs only after attack of the disulfide by HMPT.

With the reduction-elimination reaction validated on substrate **2.2**, a one-pot method was investigated where cysteine is first treated with Ellman's reagent (the free acid, rather than **2.3**) and then HMPT. The dipeptide BocAlaCysOMe (**2.6**) was used in this test reaction. Gratifyingly, the desired peptide BocAlaDhaOMe (**2.7**) was isolated in 83% after only 15 minutes of reaction time (Figure 2.6C).

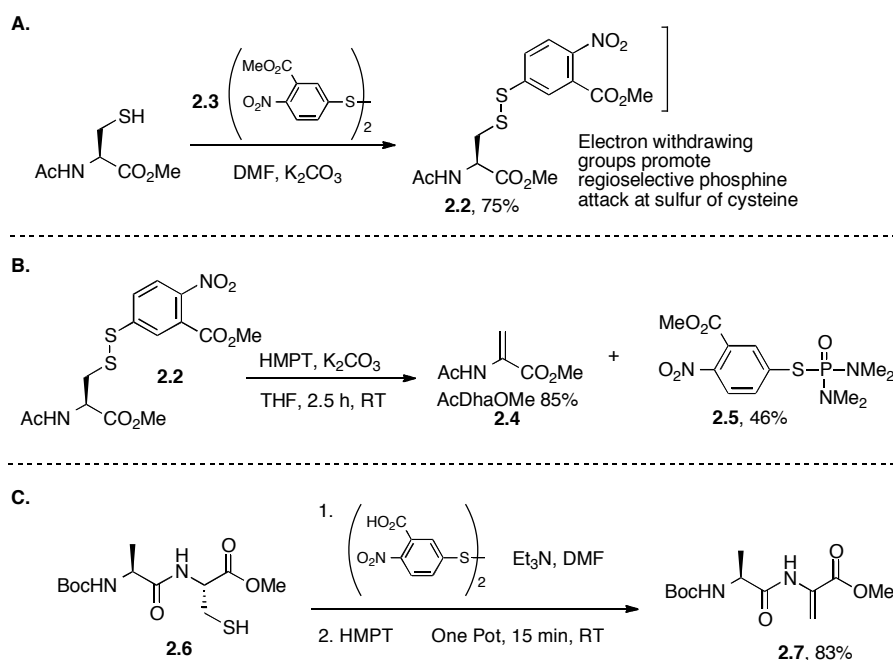


Figure 2.6: Model reduction-elimination of cysteine disulfide

Direct conversion of cysteine disulfides to dehydroalanine with base

The use of disulfides as precursors to dehydroalanine is promising since disulfides are easily formed specifically at cysteine.³¹ The reduction-elimination strategy in Figure 2.6, however, is not the only method for converting cysteine disulfides to dehydroalanine. In the course of investigating the conversion of **2.2** to dehydroalanine, it was found that DBU eliminated this material directly to dehydroalanine (Figure 2.7A). If AcCysOMe is

treated with Ellman's reagent in the presence of DBU, dehydroalanine can be formed directly in a one-pot procedure (Figure 2.7B).

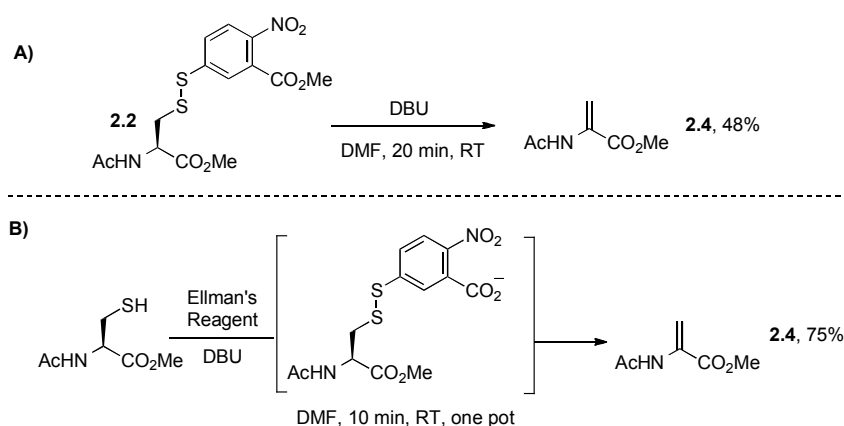


Figure 2.7: Direct elimination of cysteine disulfides with base

The elimination of cysteine disulfides to dehydroalanine using DBU is operationally simple. However, it is known that disulfides react with hydroxide to form sulfinic acids, especially labile disulfides such as those derived from Ellman's reagent.⁵⁴ Therefore this method is likely limited to use in organic solvents. Turning to methods more likely to tolerate water, the oxidative elimination of cysteine to dehydroalanine was investigated next.

Oxidative elimination of cysteine to dehydroalanine using *O*-mesitylenesulfonylhydroxylamine (MSH)

The oxidative elimination of cysteine to dehydroalanine was attempted using *O*-mesitylenesulfonylhydroxylamine (MSH, **2.1**). This reagent was prepared according to the methods of Tamura⁴⁵ and Johnson⁵⁵ (Figure 2.8). CAUTION! Extreme care was taken during the synthesis and subsequent handling of **2.1** since it is known to decompose explosively.⁵⁵ If **2.1** is prepared as small crystals, not dried excessively, stored at -20 °C in unsealed plastic containers, and handled in small quantities (typically less than 1 gram)

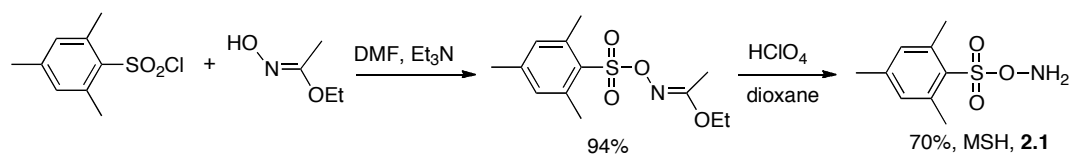


Figure 2.8: Synthesis of *O*-mesitylenesulfonylhydroxylamine (MSH)

the risk of detonation is reduced. Adhering to these precautions, we have encountered no complications to date. For further discussion of MSH stability and kilogram-scale synthesis, the reader is referred to a report by Mendiola and co-workers.⁵⁶

With MSH in hand, the elimination of BocCysOMe was attempted (Figure 2.9). Treatment of BocCysOMe with 2.2 equivalents of MSH resulted in two products after only 2 minutes of reaction time: the desired BocDhaOMe (**2.8**) and the disulfide of BocCysOMe (**2.9**). Fortunately, the formation of **2.9** could be avoided entirely if BocCysOMe was added to excess MSH (Figure 2.9B). Using this procedure, a 98% yield of BocDhaOMe was obtained. The reaction is thought to proceed via *syn*-elimination, as depicted in Figure 2.5C. However, it is not clear at this time whether the cysteine is aminated once or twice before eliminating. When disulfide **2.9** was treated with MSH, no reaction was observed, ruling out **2.9** as an intermediate in the elimination to dehydroalanine. To the best of our knowledge, this is the first direct oxidative elimination of a thiol to an alkene.

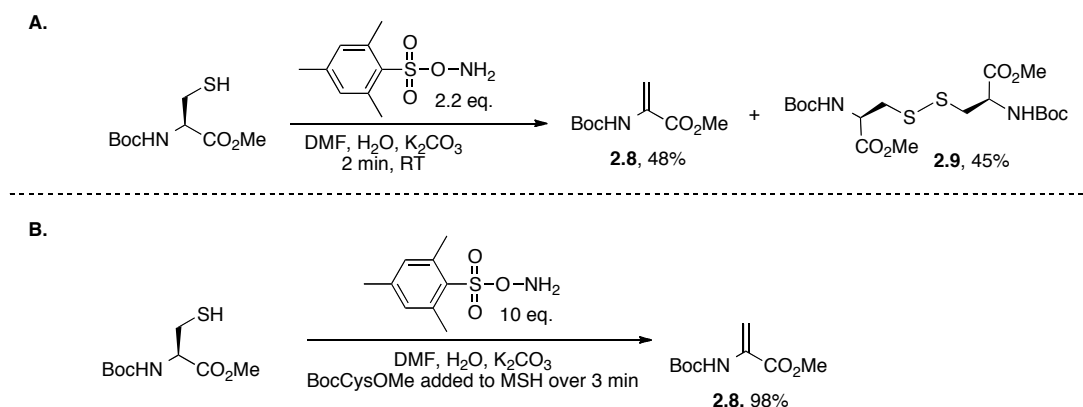


Figure 2.9: Oxidative elimination of cysteine to dehydroalanine

The oxidative elimination of cysteine thioethers was examined next. Again, efficient oxidative elimination was observed: BocDhaOMe **2.8** was isolated in 79% yield after cysteine thioether **2.10** was reacted with MSH (Figure 2.10). The mild conditions and short reaction time of the MSH-mediated elimination compare favorably with the analogous sulfoxide elimination which typically requires prolonged reaction time and elevated temperature.³⁴ Cysteine thioethers such as **2.10** might be useful in peptide synthesis as dehydroalanine surrogates and they do not require side chain protection.

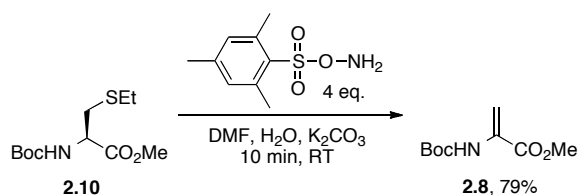


Figure 2.10: Oxidative elimination of a cysteine thioether to dehydroalanine

On the other hand, the reaction of thioether **2.10** with MSH presaged potential compatibility issues with peptides and proteins containing methionine. To investigate whether methionine and other amino acids can react with MSH, partially protected variants of amino acids with functionalized side chains were prepared. These amino acids were then treated with MSH under conditions and times representative for the elimination of cysteine to dehydroalanine. It was thought that by identifying potential side-reactions between MSH and other amino acids, experimental conditions could be adapted to maximize selectivity for cysteine.

Reaction of MSH with functionalized amino acids

MSH did not react with the serine, threonine, asparagine, glutamine, tyrosine and tryptophan side chains under the reaction conditions used for the elimination of cysteine to dehydroalanine. BocArgOH did not react with MSH, likely because its side chain was protected as the hydrochloride salt. The C-terminal carboxy group of BocArgOH also did not react with MSH (Figure 2.11).

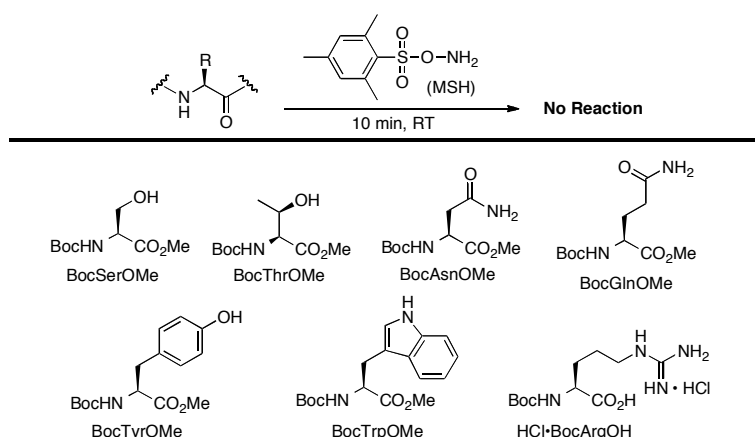


Figure 2.11: Amino acid derivatives that do not react with MSH under the conditions used for the conversion of cysteine to dehydroalanine

Carboxylates can, however, react with MSH under basic conditions. The aspartate and glutamate side chains are both aminated rapidly by MSH (Figure 2.12). Fortunately, the carboxyaminated products can be reduced back to the native side chain. For example, aspartate derivative **2.11** can be reduced to BocAspOBn by treatment with dithiothreitol (DTT). The “reductive recovery” of the native side chain demonstrates that while MSH can react with glutamate and aspartate residues, the product is not permanent.

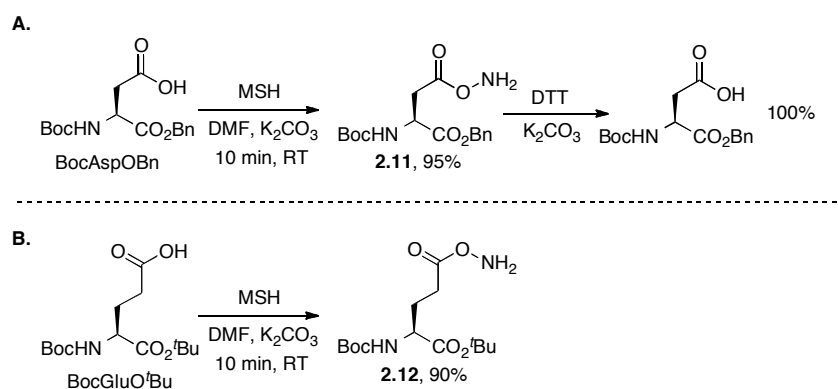


Figure 2.12: Reaction of MSH with glutamate and aspartate and reduction of the carboxyaminated product with DTT

The reaction of MSH and methionine was examined next. Given the reaction of **2.10** and other thioethers with MSH,^{43,44} the amination observed in Figure 2.13 was not surprising. BocMetOMe reacted instantaneously with MSH, and no starting material was observed by TLC. Presumably, a product such as **2.13** is formed. Vinyl glycine and dehydrobutyrine derivatives were not formed under these reaction conditions, suggesting elimination of **2.13** is slow. Moreover, the product of amination can be reduced back to methionine by treatment with DTT (Figure 2.13). Other reducing agents such as phosphines, zinc metal, sodium dithionite, and sodium ascorbate did not regenerate BocMetOMe (see Experimental Procedures for full details). Notably, no sulfoxide or sulfone of BocMetOMe was detected during these reactions.

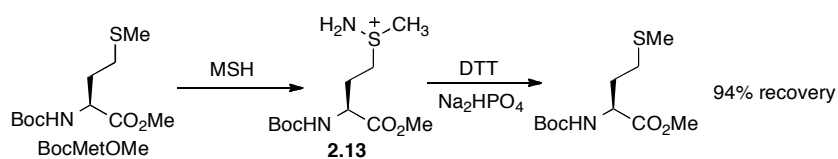


Figure 2.13: Reaction of MSH with methionine and reduction of the product with DTT

The histidine side chain reacted with MSH to give a regioisomeric mixture of aminated products (**2.14** and **2.15**). This observation is consistent with previous reports of the amination of nitrogen heterocycles with MSH and related sources of electrophilic nitrogen.^{46,57} Poor mass balance for this reaction suggests that some material is lost on workup—perhaps as the water soluble imidazolium product of di-*N*-amination of the histidine side chain. This amination of histidine must be kept in mind if MSH is used on a protein containing a polyHis tag, a common tag used for protein purification.

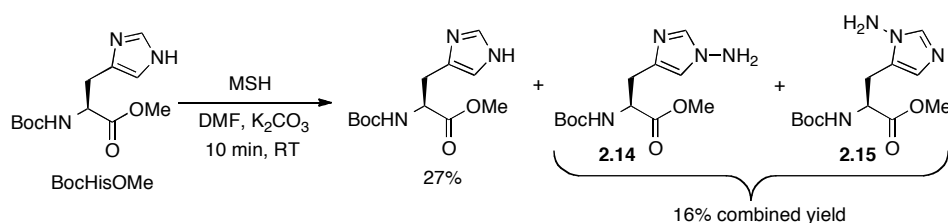


Figure 2.14: Reaction of MSH with histidine

MSH also reacted with the lysine side chain to give a mixture of products. The major product isolated from this complex mixture was the *deaminated* product norleucine (**2.16**, Figure 2.15A). This result was initially counterintuitive: MSH, an aminating and oxidizing agent, effectively deaminated and reduced lysine. This reaction, however, likely proceeds through the hydrazine (the product of *N*-amination of the lysine side chain) before further oxidation to the diimide and extrusion of nitrogen.⁵⁸ This transformation is analogous to a “hydrodeamination” reported by Doldouras and Kollonitsch⁵⁹ after the reaction of primary amines with hydroxylamine-*O*-sulfonic acid. A similar deamination was observed at the *N*-terminal amino group. For instance, after phenylalanine methyl ester was treated with MSH, deaminated product **2.17** was formed in 47% yield (Figure 2.15B).

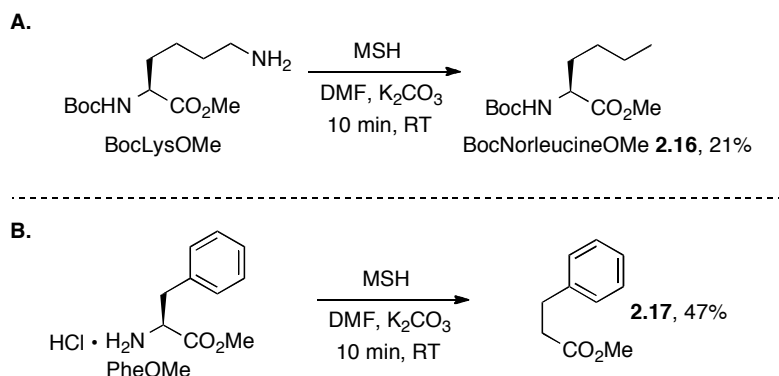


Figure 2.15: Reaction of MSH with lysine and *N*-terminal amino groups

From this assessment of MSH reactivity with functionalized amino acids, it was established that side reactions occur with nucleophilic residues (Asp, Glu, Met, His, and Lys) and *N*-terminal amino groups. For Asp, Glu, and Met, the product of amination can be reduced to the native side chain with DTT. Additionally, reaction with His, Lys, and *N*-terminal amino groups might be avoided by lowering the pH so that these groups are protected in their protonated form. Ultimately, the successful use of MSH in the oxidative elimination of cysteine and alkylcysteines to dehydroalanine will depend on the substrate and conditions employed. Differential accessibility of side chains, pH, and equivalents of MSH will all likely influence the observed selectivity. These considerations were critical in adapting use of MSH to protein modification (*vide infra*). Before considering protein modification, however, a fourth and final method for the conversion of cysteine to dehydroalanine was examined on amino acid models. This bis-alkylation-elimination method is considered next.

Bis-alkylation-elimination of cysteine to dehydroalanine

There were several parameters that were considered in the development of a general bis-alkylation-elimination of cysteine to dehydroalanine: leaving group identity, ring size in the cyclic sulfonium intermediate, saturation level, and water solubility. The general design of this transformation is depicted in Figure 2.5D. The initial goal was to optimize these parameters so that efficient conversion of cysteine to dehydroalanine could be realized at room temperature and at a pH compatible with most proteins. Initially, a model reaction between BocCysOMe and a variety of candidate reagents was examined (Figure 2.16).

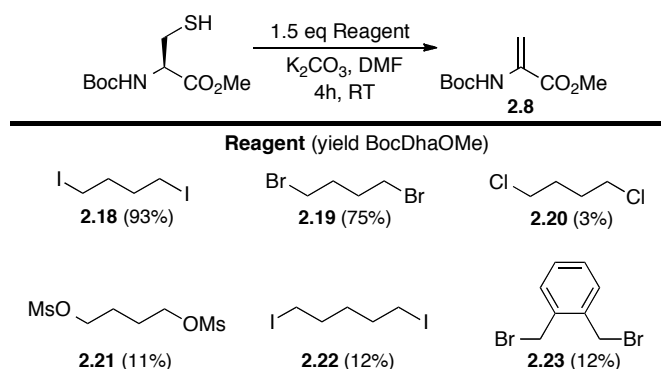


Figure 2.16: Reagents for the bis-alkylation-elimination of cysteine to dehydroalanine

The results in Figure 2.16 are revealing. First, 1,4-diiodobutane (**2.18**) and 1,4-dibromobutane (**2.19**) gave the highest yields in the model elimination. 1,4-Dichlorobutane (**2.20**) was slow to react with cysteine and unreacted BocCysOMe was recovered. While busulfan (**2.21**) is known to convert the cysteinyl residue of glutathione to dehydroalanine *in vivo*,^{47,49-52} the yield for the model reaction in Figure 2.16 was only 11%. 1,5-Diiodopentane (**2.22**) also gave a relatively low yield of dehydroalanine. The low yields of dehydroalanine generated using **2.22** suggest that either cyclization to form the 6-member sulfonium ion is slow or that elimination of this intermediate is slow (or a combination of both). Finally, reagent **2.23** gave similarly low yields of dehydroalanine (Figure 2.17).

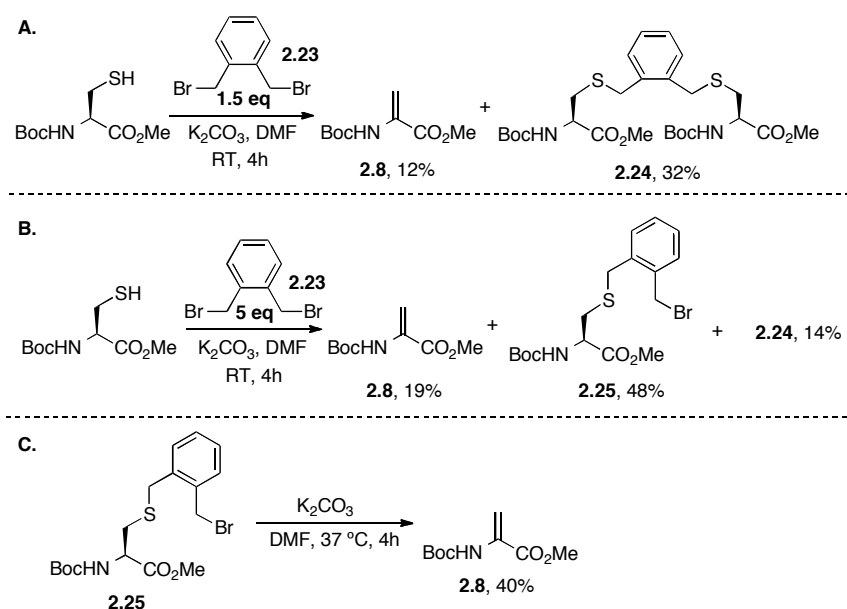


Figure 2.17: Bis-alkylation-elimination of cysteine using **2.23**

When using **2.23**, a significant amount of di-cysteinyated product **2.24** was formed, suggesting that the cyclization is slow (Figure 2.17A). Indeed the uncyclized intermediate **2.25** could be isolated if excess **2.23** was used to minimize the formation of the di-cysteinyated product (Figure 2.17B). To verify that **2.25** is a precursor to dehydroalanine, the uncyclized material was resubjected to the reaction mixture and heated at 37 °C for 4 hours. BocDhaOMe was isolated in 40% yield (Figure 2.17C).

The results in Figures 2.16 and 2.17 guided the synthesis of water-soluble reagents that could be used in the chemical conversion of cysteine to dehydroalanine. These reagents (**2.30** and **2.32**) contain the 1,4-dihalobutane core that gave the best results

in the model transformations. For water solubility, the 1,4-diiodobutane derivative **2.30** contains a hydrophilic tetraethyleneglycol appendage and the 1,4-dibromobutane derivative **2.32** contains two polar amide groups. Neutral water solublizing groups were selected to avoid any charge-charge repulsion with protein side chains near the cysteine of interest. Finally, it should be noted that **2.32** was inspired in part by Kajihara's observation that diethyl *meso*-2,5-dibromoadipate converted cysteine to dehydroalanine on peptide substrates; **2.32** is a water-soluble version of this reagent. The syntheses of **2.30** and **2.32** are outlined in Figure 2.18.

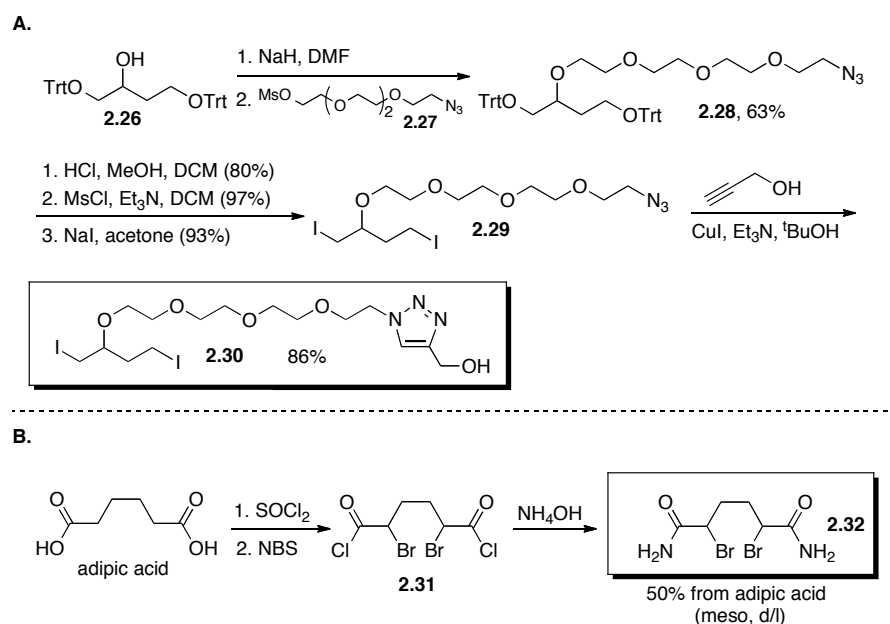


Figure 2.18: Synthesis of water-soluble 1,4-dihalobutane derivatives

Briefly, the synthesis of **2.30** began with commercially available 1,2,4-butanetriol. Selective di-tritylation provided **2.26** which was subsequently alkylated using mesylate **2.27**. The 1,4-di-tritylether **2.28** was then converted to 1,4-diiodide **2.29** by standard transformations. Finally, Cu-catalyzed azide-alkyne cycloaddition^{60,61} provided **2.30**. The final cycloaddition was designed so that any alkyne could be used to impart more hydrophilic character. Propargyl alcohol, however, sufficed and **2.30** is miscible in water.

Reagent **2.32** required fewer steps for its synthesis. Starting from adipic acid, the acid chloride was formed by reaction with thionyl chloride; subsequent di-bromination provided intermediate **2.31**. The crude acid chloride **2.31** was then poured into an aqueous solution of ammonium hydroxide. The bis-amide product **2.32** precipitates directly from the reaction and can be purified by simple filtration and trituration with

water and methanol. It should be noted that **2.32** was isolated as a mixture of *meso* and *d/l* diastereomers. This mixture was not separated and assessed directly in the elimination of cysteine to dehydroalanine.

In the conversion of cysteine to dehydroalanine, BocCysOMe was used and the reaction was first run in DMF to compare directly to the parent compounds in Figure 2.16. Using 1,4-diiodide **2.30**, useful yields of dehydroalanine were obtained at room temperature (Figure 2.19A). Reagent **2.32**, however, gave a disappointingly low yield of 22% (Figure 2.19B). It appeared that the cyclization or elimination step was slow for reagent **2.32** since all cysteine was consumed in the reaction. To overcome these sluggish steps, cysteine was first alkylated at room temperature with **2.32** and then heated at 37 °C to induce cyclization and elimination. Using this protocol, good yields of dehydroalanine were obtained (Figure 2.19C).

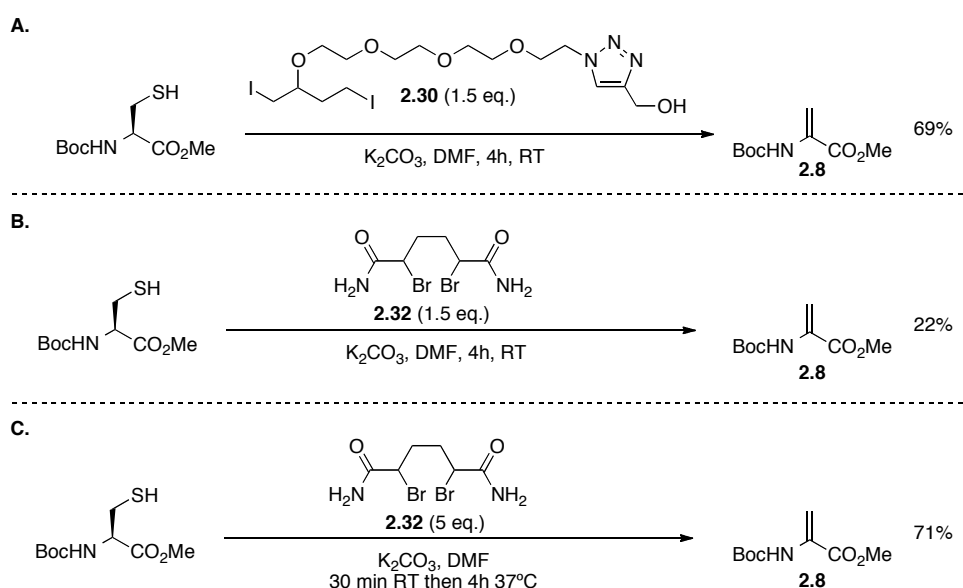


Figure 2.19: Conversion of cysteine to dehydroalanine with **2.30** and **2.32**

The results in Figure 2.19 completed the assessment of the methods for the conversion of cysteine to dehydroalanine proposed in Figure 2.5. Next, these methods were investigated on protein substrates.

Conversion of cysteine to dehydroalanine on proteins.

Reduction-elimination of cysteine disulfides to dehydroalanine

In Figure 2.6, the reduction-elimination of cysteine disulfides to dehydroalanine was demonstrated on simple substrates in organic solvent. To investigate this reaction on a protein substrate, a single cysteine mutant of the protease subtilisin from *Bacillus lentus* (SBL-S156C) was converted to disulfide **2.33** by reaction with Ellman's reagent (Figure 2.20).

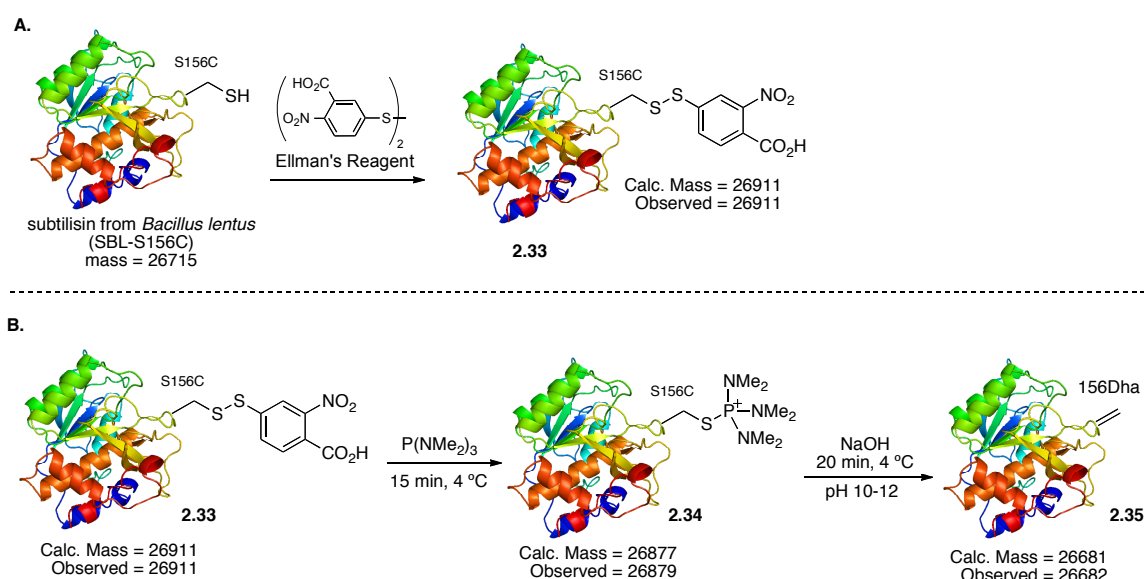


Figure 2.20: Elimination of cysteine disulfides to dehydroalanine on subtilisin from *Bacillus lentus*

Full conversion to **2.33** was verified by LC-MS.^{62,†} Disulfide **2.33** was then purified by dialysis before testing the reduction-elimination. When **2.33** was treated with HMPT (Figure 2.20B), rapid reaction was observed by LC-MS. The mass observed was

[†] Monitoring reaction conversions by ESI-MS have been validated on SBL-S156C. The reader is referred to this publication for further discussion (See the supporting information in reference 62). In short, the concentration of two SBL-S156C derivatives of distinct masses were measured by the bicinchoninic acid (BCA) assay. Mixtures of these two proteins were then prepared in known amounts and analyzed by LC-MS. Good correlations were observed between the relative amount of each protein measured by the BCA assay and the relative amount of each protein according to peak height in the mass spectrum. Note that this validation applies only to modifications of a single residue, where the ionization efficiency of the unmodified and modified protein are likely to be similar.

consistent with the thiophosphonium salt depicted in **2.34**. This thiophosphonium salt, however, proved remarkably stable and no mass change was observed after several hours at pH 8.0. However, after adjusting the pH by the addition of NaOH, the mass of the protein changed immediately and corresponded to the mass of the dehydroalanine containing protein **2.35** (Figure 2.20B). While this result suggested that the elimination might be possible, this reaction was not pursued further on protein substrates because of the high pH required for the elimination.

Base-mediated elimination of cysteine disulfides to dehydroalanine

Subtilisin protein **2.33** was also used as a test substrate to determine if direct elimination to dehydroalanine is possible. However, when **2.33** was treated with NaOH, no elimination was observed by LC-MS. The mass of the product observed was consistent with formation of the cysteine sulfinic acid **2.36** (Figure 2.21).

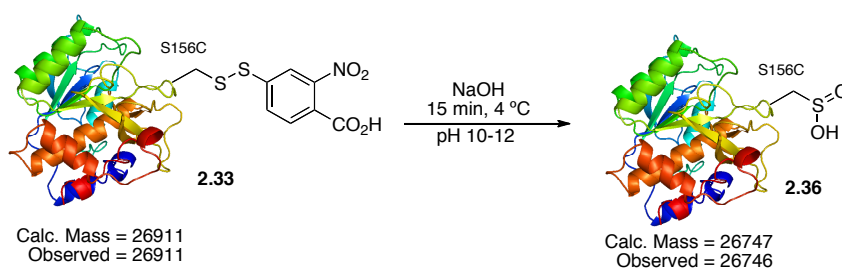


Figure 2.21: Attempted elimination of cysteine disulfides to dehydroalanine on subtilisin from *Bacillus lentus* and possible structure of product.

It is known that disulfides made using Ellman's reagent react with hydroxide to form sulfenic acids that rapidly oxidize to the sulfinic acids.⁵⁴ Thus, while DBU can induce elimination of cysteine disulfides to dehydroalanine in organic solvents (Figure 2.7), this transformation is not readily translated to aqueous systems. Turning to methods that might better tolerate water, the oxidative elimination of cysteine to dehydroalanine and the bis-alkylation-elimination methods were investigated next.

Oxidative elimination of cysteine to dehydroalanine using MSH

In the model studies described above, MSH (**2.1**) reacted with several amino acid residues. However, since cysteine is typically one of the more nucleophilic of these residues, it might react selectively with MSH to provide dehydroalanine. To test this

hypothesis, reaction of MSH with the single cysteine mutant of subtilisin from *Bacillus lentus* was investigated. The reaction was carried out at 4 °C and phosphate buffer at pH 8.0 was used. Organic buffers (TRIS, TAPS, HEPES, etc) were avoided since many contain amine groups that can react with MSH. When MSH was added as a solution in DMF (the total DMF content in the reaction was typically less than 10%), rapid loss of 34 mass units was observed by LC-MS (Figure 2.22). The formation of dehydroalanine was determined by chemical tests. First, the product of the reaction with MSH was treated with Ellman's reagent. No reaction was observed by LC-MS, suggesting that all cysteine was consumed in the reaction. It follows that the reaction at MSH must have occurred at cysteine. When the unmodified SBL-S156C was treated with Ellman's reagent under the same conditions, rapid and complete formation of the disulfide was observed. To chemically corroborate the formation of dehydroalanine, a variety of thiols were added. The protein mass increased by the mass of the thiol, consistent with conjugate addition to dehydroalanine. The addition of the glycosyl thiol of *N*-acetylglucosamine (**2.37**) is one example of this addition and is shown in Figure 2.22. Taken together with the results of the small molecule models, it can be concluded that MSH can react with cysteine to form dehydroalanine and this reaction was successfully applied to a protein substrate. Good protein recovery was obtained after dialysis (typically 80% or higher). Importantly, protein **2.38** remained active over the two-step modification ($k_{\text{cat}} = 5.7 \text{ s}^{-1}$; $K_{\text{M}} = 0.72 \text{ mM}$ using the peptide sucAAPFpNA as the substrate). This result suggests that the conditions used for the modification do not significantly compromise the tertiary structure of the protein.

The addition of other thiols to access post-translational modifications will be discussed further in Chapter 3. For the present discussion, suffice is to note that the addition of glycosyl thiols to dehydroalanine provides straightforward access to glycosyl cysteine residues. As mentioned in the introduction to this chapter, these residues are naturally occurring¹¹⁻¹³ and are also analogs of the more common serine-linked glycoproteins.^{14,63} Thiol addition to dehydroalanine is an operationally simple method for incorporating these residues into proteins, but it was not clear whether the MSH reaction could be generalized to other proteins. It was promising that side reactions were not observed on SBL-S156C. This protein contains several lysine and methionine residues, but under these reaction conditions they did not react with MSH.

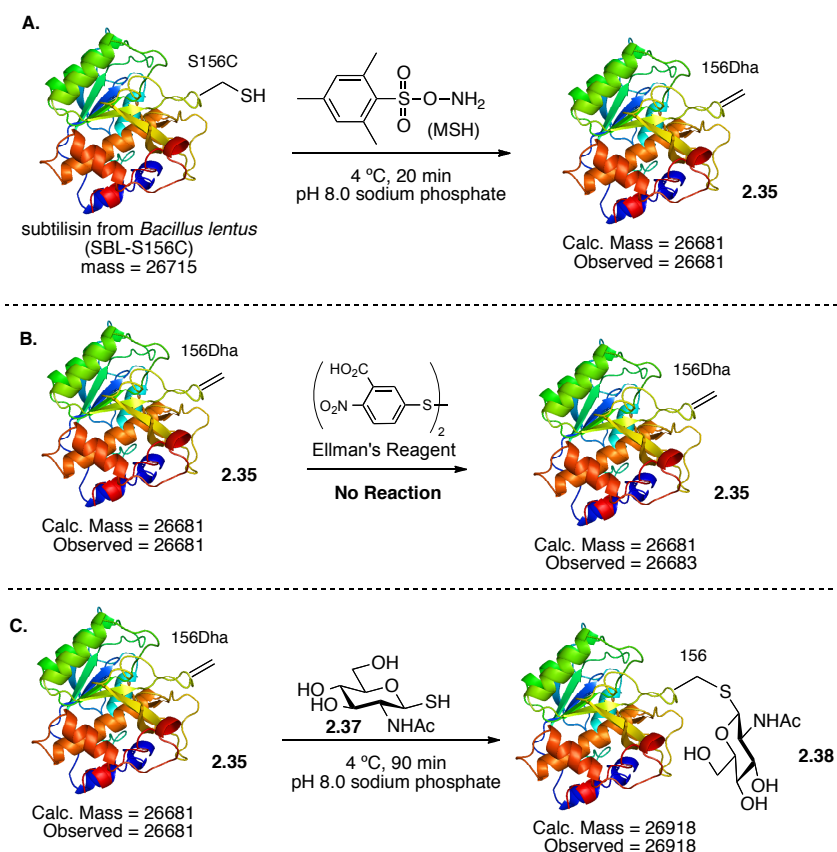


Figure 2.22: Reaction of SBL-156C and MSH and chemical tests to corroborate the formation of dehydroalanine.

To determine the scope of this reaction, MSH was reacted with another single-cysteine protein, Np276 from *Nostoc punctiforme*.[‡] This protein, unlike SBL-S156C, contains an *N*-terminal His tag—a potential site of amination by MSH. Indeed, after treatment with MSH, the Np276 protein reacted to give a mixture of products. LC-MS analysis of the product mixture revealed several products separated by 16 Da, consistent with non-selective amination (e.g. addition of “NH₂”). Lowering the pH to 6.0 resulted in fewer aminations, but it was clear that non-selective amination was still occurring. The MS data is shown in Figure 2.23.

No further effort was made to determine where amination was occurring on the Np276 protein. Rather than optimize the reaction with MSH (perhaps by varying temperature, time, equivalents of MSH, pH, etc), attention was turned to the bis-

[‡] The original construct for the Np276 gene was donated by the Prof. John Blanchard laboratory at the Albert Einstein College of Medicine (New York). The construct for the mutant used in this thesis was prepared by Dr. James Errey and Matthew Young. Protein expression and purification was carried out by Matthew Young.

alkylation-elimination reaction. Should this reaction prove more selective than MSH, the Np276 protein would be a challenging substrate to assess its generality. Accordingly, the bis-alkylation-elimination of cysteine to dehydroalanine was investigated on protein substrates.

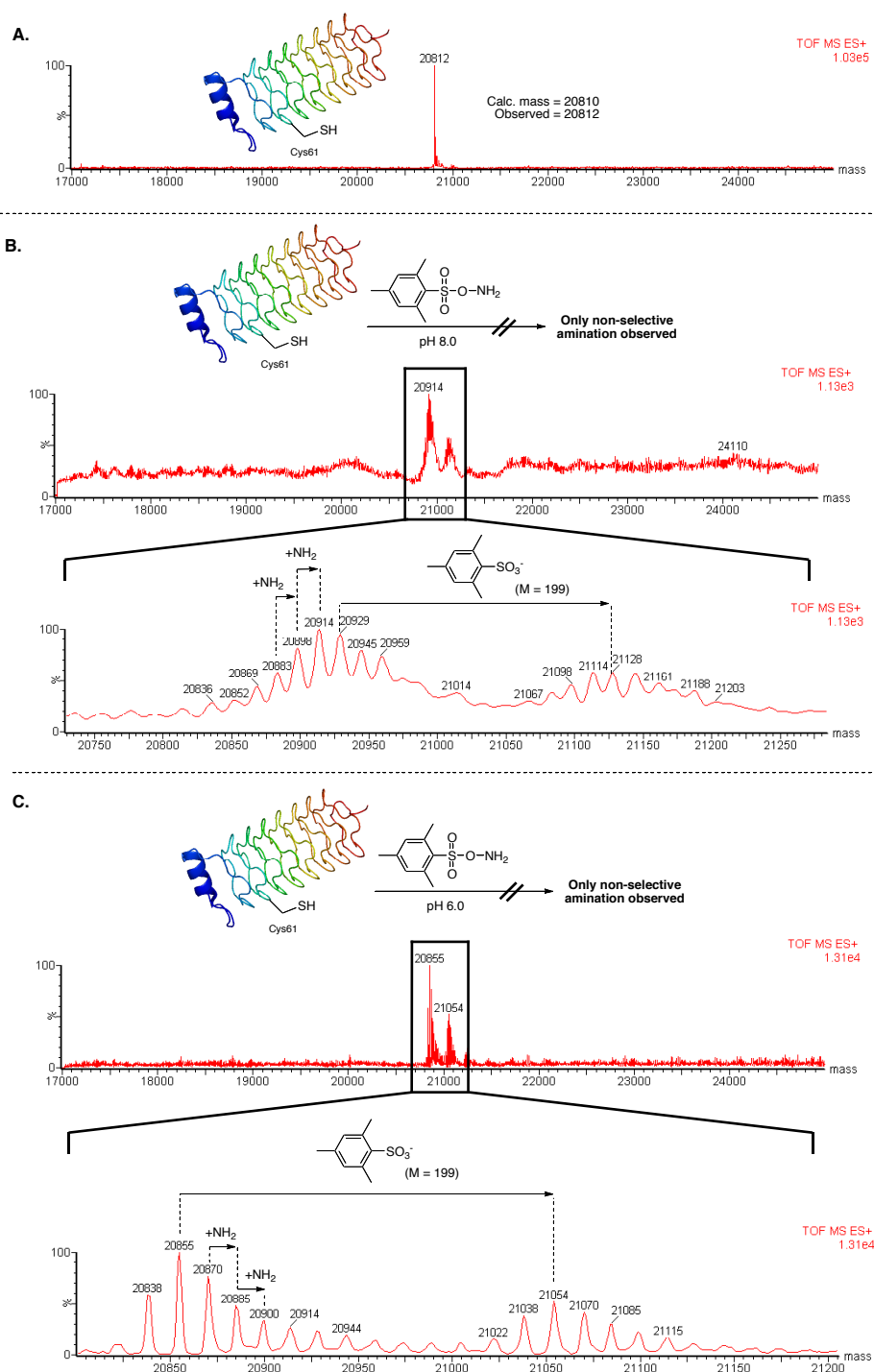


Figure 2.23: Reaction of a single cysteine mutant of Np276 from *Nostoc punctiforme*. Non-selective amination was observed at both pH 8.0 and pH 6.0.

Bis-alkylation-elimination of cysteine to dehydroalanine

Diiodide **2.30** and dibromide **2.32** were first evaluated on SBL-S156C to assess their use on protein substrates. After incubating SBL-S156C with diiodide **2.30** for 1 hour at 37 °C, a loss of 34 mass units was observed by LC-MS (Figure 2.24). The same battery of chemical tests used previously was then carried out to confirm the formation of dehydroalanine: Ellman's reagent did not react with the product, indicating all cysteine was consumed in the reaction. The conjugate addition of a thiol (2-mercaptoethanol) corroborated the formation of dehydroalanine. The overall reaction sequence is shown in Figure 2.24.

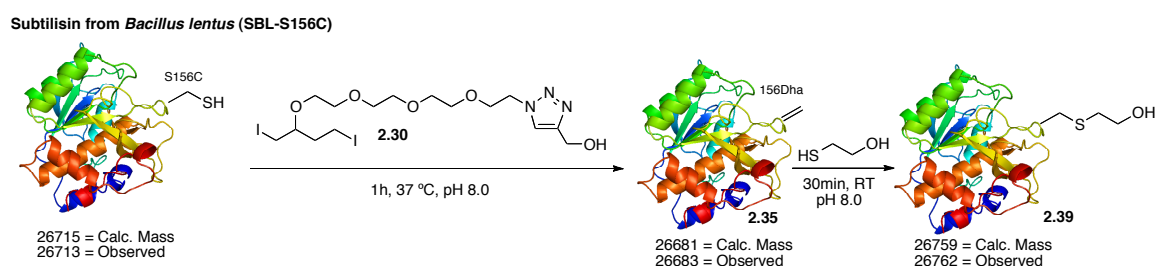


Figure 2.24: Conversion of cysteine to dehydroalanine on SBL-S156C using reagent **2.30**.

The reaction of SBL-S156C with dibromide **2.32** was examined next. The protein was treated with **2.32** as a solution in DMF (the total concentration of DMF was less than 10% of the final volume) and then incubated at room temperature for 30 minutes. After this time, LC-MS revealed the mono-alkylated, uncyclized intermediate **2.40**. This result was consistent with the model reactions in Figure 2.19 that suggested the cyclization step was slow at room temperature. To promote cyclization and elimination, the sample was simply incubated for an additional hour at 37 °C. After this time, clean elimination was observed by LC-MS. Again, the product did not react with Ellman's reagent, confirming reaction occurred at cysteine. Addition of 2-mercaptoethanol corroborated the formation of dehydroalanine. The entire reaction sequence is shown in Figure 2.25.

It should also be pointed out that dibromide **2.32** is synthesized as a mixture of *meso* and *d/l* diastereomers. The separation of these materials was not readily achieved by column chromatography or recrystallization so the mixture of these diastereomers was used in all model reactions on peptides and proteins. The use of this mixture will result in diastereomeric intermediates after both alkylation and cyclization. Though the rate of bis-

alkylation and elimination of these diastereomers may differ, these intermediates converge to the same dehydroalanine product. Since good yields and conversions were obtained in the model reactions on peptides and proteins, any differences in these rates do not appear detrimental to the formation of dehydroalanine.

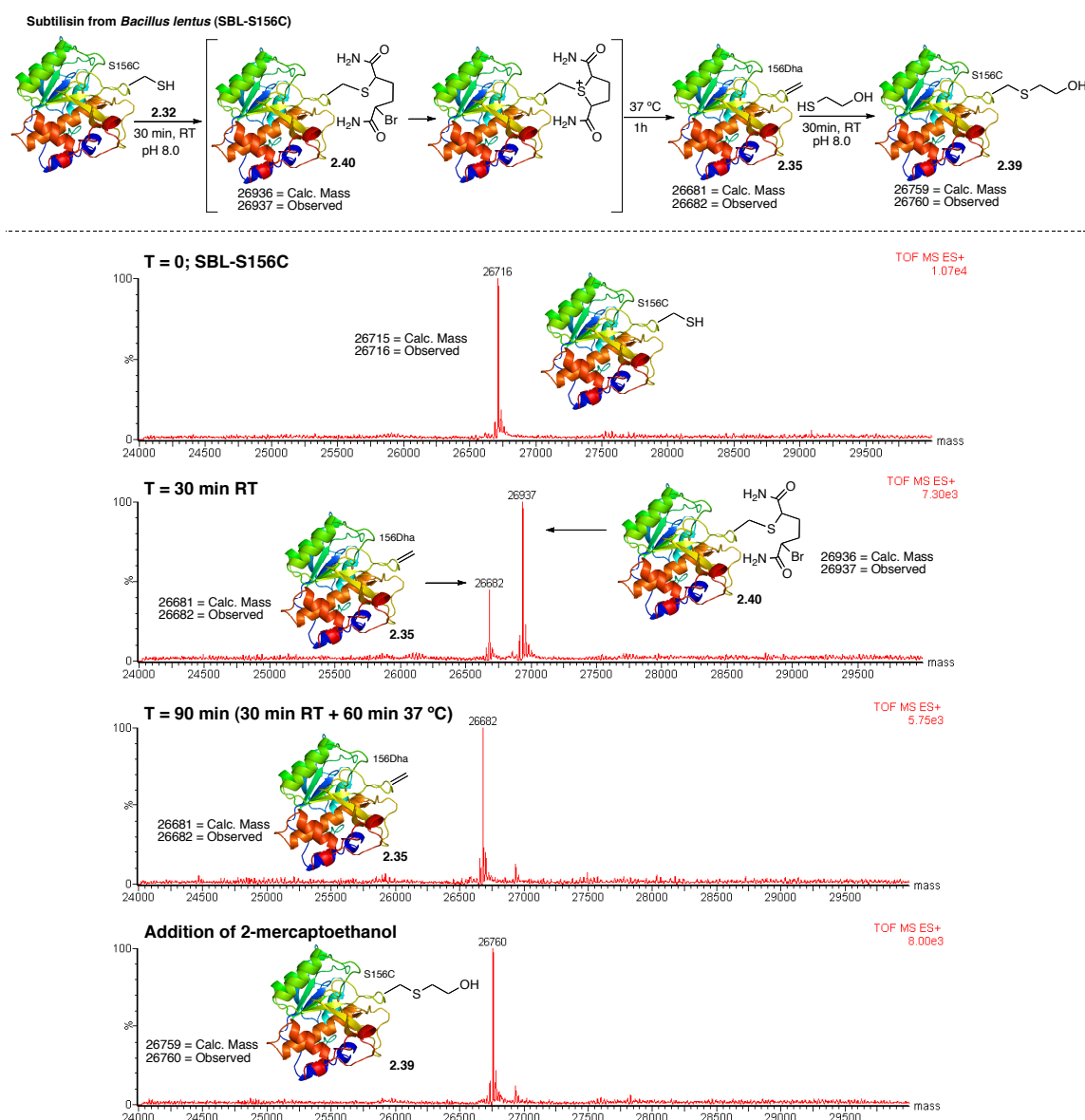


Figure 2.25: Conversion of cysteine to dehydroalanine on SBL-S156C using reagent **2.32**.

With reagents **2.30** and **2.32** validated on SBL-S156C, the stage was set for application of these reagents to the single cysteine Np276 protein that reacted unselectively with MSH. Since the dibromide reagent **2.32** requires a shorter synthesis than **2.30**, the dibromide reagent was used in all subsequent experiments. The reaction of

2.32 with the single cysteine Np276 mutant appeared slower than the reaction with SBL-S156C and unreacted starting material was observed at room temperature. This problem was remedied by simply adding dibromide **2.32** in three portions, all at 37 °C. Clean loss of 34 mass units was observed by LC-MS using this protocol. Chemical and LC-MS analysis confirmed the conversion of cysteine to dehydroalanine: no reaction was observed with Ellman's reagent and 2-mercaptoethanol readily added to the protein product. The reaction sequence and analysis is shown in Figure 2.26.

Np276 from *Nostoc punctiforme* (single Cys at position 61)

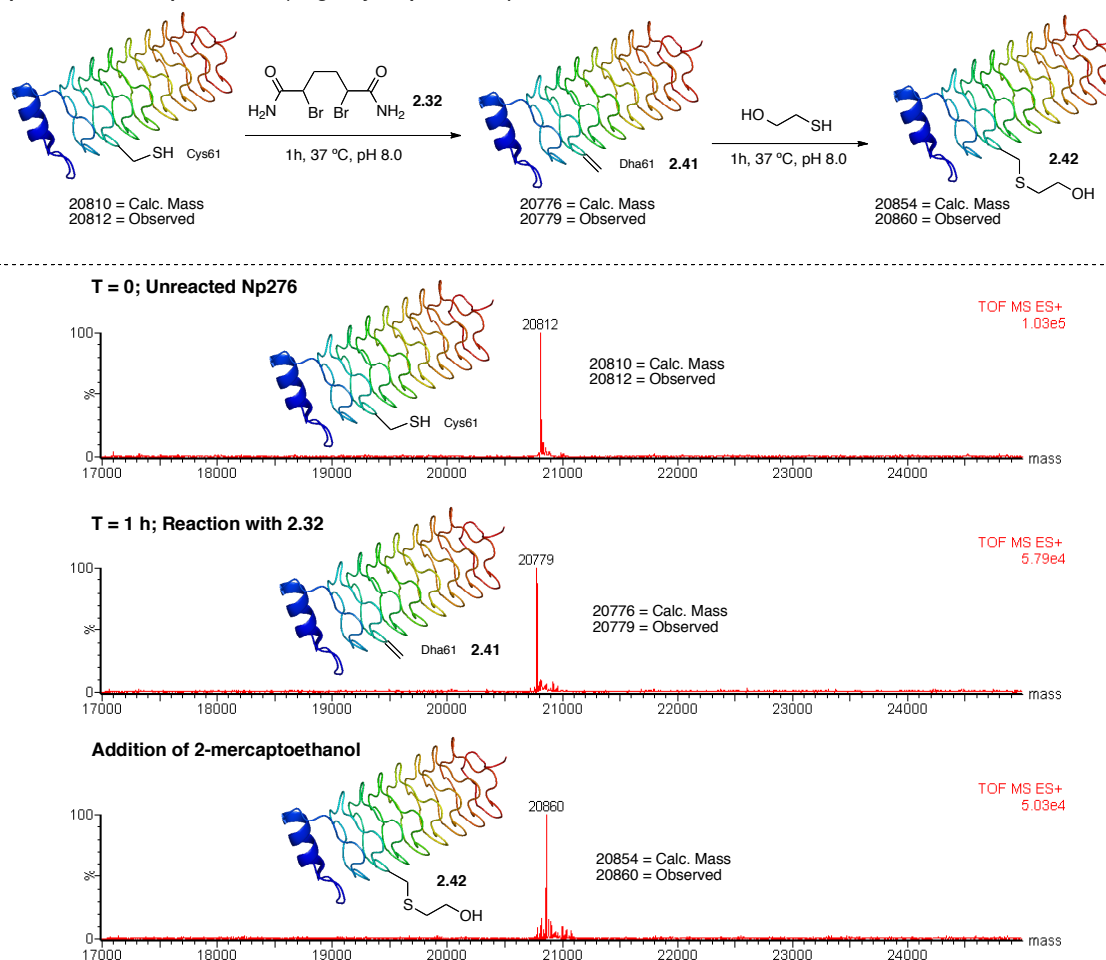


Figure 2.26: Conversion of cysteine to dehydroalanine on Np276 using reagent **2.32**.

Gratifyingly, the use of dibromide **2.32** overcome the limitations of MSH and therefore constitutes a more general reagent for the conversion of cysteine to dehydroalanine on proteins. In the next chapter, the application of dibromide **2.32** in the bis-alkylation-elimination of cysteine to dehydroalanine will be examined in the context of post-translational modifications. More specifically, it will be shown that the chemical

installation of dehydroalanine allows chemical access to several classes of post-translational modifications on protein surfaces. First, a summary of the findings in this chapter is given below.

2.3 Conclusions and Outlook

In this Chapter, several methods for the conversion of cysteine to dehydroalanine were investigated on peptides and proteins. The conversion of cysteine disulfides to dehydroalanine was relatively efficient on peptide substrates under both reducing and basic conditions, but required a pH that is too high for most protein substrates. A novel oxidative elimination of cysteine to dehydroalanine was discovered that employs the electrophilic nitrogen source *O*-mesitylenesulfonylhydroxylamine (MSH). This reaction was successfully applied to a model enzyme that contains a single cysteine, SBL-S156C.⁶⁴ For another single cysteine protein, Np276, MSH reacted unselectively, resulting in unwanted amination. Potential side reactions with MSH were determined using amino acid models. Nucleophilic residues such as lysine and histidine are susceptible to reaction with MSH. Methionine, aspartate, and glutamate also react with MSH, but can be reduced back to the native side chain with DTT. To avoid these side reactions entirely, two reagents were designed that convert cysteine to dehydroalanine through a bis-alkylation-elimination reaction. These reagents, **2.30** and **2.32**, both were successfully applied to protein substrates. Notably, **2.32** was used to convert cysteine to dehydroalanine on Np276—the protein that reacted non-selectively with MSH.

With an efficient conversion of cysteine to dehydroalanine in hand, the use of dehydroalanine in chemical protein modification could be explored. In this chapter, a single example of such a modification was described: the conversion of dehydroalanine to GlcNAc cysteine by the addition of the glycosyl thiol of *N*-acetyl glucosamine. To complement and expand the synthesis of glycoproteins touched upon in this chapter, a more comprehensive investigation of dehydroalanine as a precursor to post-translational modifications was undertaken. This study is the focus of Chapter 3.

2.4 References

- (1) Jung, G. Lantibiotics - ribosomally synthesized biologically active polypeptides containing sulfide bridges and α,β -didehydroamino acids. *Angew. Chem. Int. Ed. Engl.* **1991**, *30*, 1051-1068.
- (2) Jack, R. W.; Jung, G. Lantibiotics and microcins: polypeptides with unusual chemical diversity. *Curr. Opin. Chem. Biol.* **2000**, *4*, 310-317.
- (3) Willey, J. M.; van der Donk, W. A. Lantibiotics: peptides of diverse structure and function. *Annu. Rev. Microbiol.* **2007**, *61*, 477-501.
- (4) Palmer, D. E.; Pattaroni, C.; Nunami, K.; Chadha, R. K.; Goodman, M.; Wakamiya, T.; Fukase, K.; Horimoto, S.; Kitazawa, M.; Fujita, H.; Kubo, A.; Shiba, T. Effects of dehydroalanine on peptide conformations. *J. Am. Chem. Soc.* **1992**, *114*, 5634-5642.
- (5) Gavaret, J.-M.; Nunez, J.; Cahnmann, H. J. Formation of dehydroalanine residues during thyroid hormone synthesis in thyroglobulin. *J. Biol. Chem.* **1980**, *255*, 5281-5285.
- (6) Gavaret, J.-M.; Cahnmann, H. J.; Nunez, J. Thyroid hormone synthesis in thyroglobulin. The mechanism of the coupling reaction. *J. Biol. Chem.* **1981**, *256*, 9167-9173.
- (7) Zhu, Y.; van der Donk, W. A. Convergent synthesis of peptide conjugates using dehydroalanines for chemoselective ligations. *Org. Lett.* **2001**, *3*, 1189-1192.
- (8) Gieselman, M. D.; Zhu, Y.; Zhou, H.; Galonic, D.; van der Donk, W. A. Selenocysteine derivatives for chemoselective ligations. *ChemBioChem* **2002**, *3*, 709-716.
- (9) Galonic, D. P.; van der Donk, W. A.; Gin, D. Y. Oligosaccharide-peptide ligation of glycosyl thiolates with dehydropeptides: synthesis of S-linked mucin-related glycopeptide conjugates. *Chem. Eur. J.* **2003**, *9*, 5997-6006.
- (10) Wang, J.; Schiller, S. M.; Schultz, P. G. A biosynthetic route to dehydroalanine-containing proteins. *Angew. Chem. Int. Ed.* **2007**, *46*, 6849-6851.
- (11) Weiss, J. B.; Lote, C. J.; Bobinski, H. New low molecular weight glycopeptide containing triglucosylcysteine in human erythrocyte membrane. *Nat. New Biol.* **1971**, *234*, 25-26.

- (12) Lote, C. J.; Weiss, J. B. Identification of digalactosylcysteine in a glycopeptide isolated from urine by a new preparative technique. *FEBS Lett.* **1971**, *16*, 81-85.
- (13) Oman, T. J.; Boettcher, J. M.; Wang, H.; Okalibe, X. N.; van der Donk, W. A. Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nat. Chem. Biol.* **2011**, *7*, 78-80.
- (14) *Essentials of Glycobiology*; 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press, 2009.
- (15) Dwek, R. A. Glycobiology: toward understanding the function of sugars. *Chem. Rev.* **1996**, *96*, 683-720.
- (16) Ohtsubo, K.; Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **2006**, *126*, 855-867.
- (17) Davis, B. G. Synthesis of glycoproteins. *Chem. Rev.* **2002**, *102*, 579-601.
- (18) Gamblin, D. P.; Scanlan, E. M.; Davis, B. G. Glycoprotein synthesis: an update. *Chem. Rev.* **2008**, *109*, 131-163.
- (19) Baran, E.; Drabarek, S. Studies on the synthesis of an S-glycosidic bond between cysteine and glucose or galactose. *Pol. J. Chem.* **1978**, *52*, 941-946.
- (20) Doores, K. J.; Gamblin, D. P.; Davis, B. G. Exploring and exploiting the therapeutic potential of glycoconjugates. *Chem. Eur. J.* **2006**, *12*, 656-665.
- (21) Zhang, F. L.; Casey, P. J. Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **1996**, *65*, 241-269.
- (22) Smotrys, J. E.; Linder, M. E. Palmitoylation of intracellular signaling proteins: regulation and function. *Annu. Rev. Biochem.* **2004**, *73*, 559-587.
- (23) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372.
- (24) Walsh, C. T. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*. Roberts and Company Publishers: Englewood, CO, 2006.
- (25) Gamblin, D. P.; van Kasteren, S. I.; Chalker, J. M.; Davis, B. G. Chemical approaches to mapping the function of post-translational modifications. *FEBS J.* **2008**, *275*, 1949-1959.
- (26) Okeley, N. M.; Zhu, Y.; van der Donk, W. A. Facile chemoselective synthesis of dehydroalanine-containing peptides. *Org. Lett.* **2000**, *2*, 3603-3606.

- (27) Levengood, M. R.; van der Donk, W. A. Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies. *Nat. Protoc.* **2007**, *1*, 3001-3010.
- (28) Hashimoto, K.; Sakai, M.; Okuno, T.; Shirahama, H. β -Phenylselenoalanine as a dehydroalanine precursor—efficient synthesis of alternariolide (AM-toxin I). *Chem. Commun.* **1996**, 1139-1140.
- (29) Strumeyer, D. H.; White, W. N.; Koshland, D. E. Role of serine in chymotrypsin acition. Conversion of the active serine to dehydroalanine. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *50*, 931-935.
- (30) Weiner, H.; White, W. N.; Hoare, D. G.; Koshland, D. E. The formation of anhydrochymotrypsin by removing the elements of water from the serine at the active site. *J. Am. Chem. Soc.* **1966**, *88*, 3851-3859.
- (31) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. Chemical modification of proteins at cysteine: opportunities in chemistry and biology. *Chem. Asian J.* **2009**, *4*, 630-640.
- (32) Holmes, T. J.; Lawton, R. G. Cysteine modification and cleavage of proteins with 2-methyl-N¹-benzenesulfonyl-N⁴-bromoacetylquinonediimide. *J. Am. Chem. Soc.* **1977**, *99*, 1984-1986.
- (33) Okamoto, R.; Souma, S.; Kajihara, Y. Efficient substitution reaction from cysteine to the serine residue of glycosylated polypeptide: repetitive peptide segment ligation strategy and the synthesis of glycosylated tetracontapeptide having acid labile sialyl-Tn antigens. *J. Org. Chem.* **2009**, *74*, 2494-2501.
- (34) Rich, D. H.; Tam, J.; Mathiapparanam, P.; Grant, J. A.; Mabuni, C. General synthesis of didehydroamino-acids and peptides. *J. Chem. Soc. Chem. Commun.* **1974**, 897-898.
- (35) Seebeck, F. P.; Szostak, J. W. Ribosomal synthesis of dehydroalanine-containing peptides. *J. Am. Chem. Soc.* **2006**, *128*, 7150-7151.
- (36) You, Y. O.; Levengood, M. R.; Ihnken, L. A. F.; Knowlton, A. K.; van der Donk, W. A. Lactacin 481 synthetase as a general serine/threonine kinase. *ACS Chem. Biol.* **2009**, *4*, 379-385.
- (37) Harpp, D. N.; Gleason, J. G. Organic sulfur chemistry. X. Selective desulfurization of disulfides. Scope and mechanism. *J. Am. Chem. Soc.* **1971**, *93*, 2437-2445.

- (38) Leffler, J. E. Parameters for the description of transition states. *Science* **1953**, *117*, 340-341.
- (39) Hammond, G. S. A correlation of reaction rates. *J. Am. Chem. Soc.* **1955**, *77*, 334-338.
- (40) Harpp, D. N.; Gleason, J. G. The chemistry of small-ring sulfur compounds. Thietanes and 1,2-dithiolanes. *J. Org. Chem.* **1970**, *35*, 3259-3263.
- (41) Wang, H.; Zhang, J.; Xian, M. Facile formation of dehydroalanine from *S*-nitrosocysteines. *J. Am. Chem. Soc.* **2009**, *131*, 13238-13239.
- (42) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. From disulfide- to thioether-linked glycoproteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 2244-2247.
- (43) Franek, W.; Claus, P. K. Alkyl- und arylsulfensäureamide durch cycloeliminierung von alkenen aus sulfimiden, 3: bildung primärer und sekundärer aliphatischer sulfenamide des typs $RSNH_2$ bzw. $RSNHR^1$. *Monatsh. Chem.* **1990**, *121*, 539-547.
- (44) Matsuo, J.; Kozai, T.; Ishibashi, H. Mild preparation of alkenes from phenyl sulfides: one-pot elimination of phenylthio group via sulfilimine at ambient temperature. *Org. Lett.* **2006**, *8*, 6095-6098.
- (45) Tamura, Y.; Minamikawa, J.; Sumoto, K.; Fujii, S.; Ikeda, M. Synthesis and some properties of *O*-acyl- and *O*-nitrophenylhydroxylamines. *J. Org. Chem.* **1973**, *38*, 1239-1241.
- (46) Tamura, Y.; Minamikawa, J.; Ikeda, M. *O*-Mesitylenesulfonylhydroxylamine and related compounds—powerful aminating reagents. *Synthesis* **1977**, 1-17.
- (47) Roberts, J. J.; Warwick, G. P. The mode of action of alkylating agents—II: studies of the metabolism of myleran. The reaction of myleran with some naturally occurring thiols *in vitro*. *Biochem. Pharmacol.* **1961**, *6*, 205-216.
- (48) Onkenhout, W.; van Loon, W. M.; Buijs, W.; van der Gen, A.; Vermeulen, N. P. Biotransformation and quantitative determination of sulfur-containing metabolites of 1,4-dibromobutane in the rat. *Drug Metab. Dispos.* **1986**, *14*, 608-612.
- (49) Marchand, D. H.; Rimmel, R. P.; Abdel-Monem, M. M. Biliary excretion of a glutathione conjugate of busulfan and 1,4-diiodobutane in the rat. *Drug Metab. Dispos.* **1988**, *16*, 85-92.
- (50) Czerwinski, M.; Gibbs, J. P.; Slaterry, J. T. Busulfan conjugation by glutathione *S*-transferases alpha, mu, and pi. *Drug Metab. Dispos.* **1996**, *24*, 1015-1019.

- (51) Cooper, A. J. L.; Younis, I. R.; Niatetskaya, Z. V.; Krasnikov, B. F.; Pinto, J. T.; Petros, W. P.; Callery, P. S. Metabolism of the cysteine *S*-conjugate of busulfan involves a β -lyase reaction. *Drug Metab. Dispos.* **2008**, *36*, 1546-1552.
- (52) Younis, I. R.; Elliott, M.; Peer, C. J.; Cooper, A. J. L.; Pinto, J. T.; Konat, G. W.; Kraszpulski, M.; Petros, W. P.; Callery, P. S. Dehydroalanine analog of glutathione: an electrophilic busulfan metabolite that binds to human glutathione *S*-transferase A1-1. *J. Pharmacol. Exp. Ther.* **2008**, *327*, 770-776.
- (53) Ellman, G. L. Tissue sulfhydryl groups. *Arch. Biochem. Biophys.* **1959**, *82*, 70-77.
- (54) Donovan, J. W.; White, T. M. Alkaline hydrolysis of the disulfide bonds of ovomucoid and of low molecular weight aliphatic and aromatic disulfides. *Biochemistry* **1971**, *10*, 32-38.
- (55) Johnson, C. R.; Kirchhoff, R. A.; Corkins, H. G. Synthesis of optically active sulfoximines from optically active sulfoxides. *J. Org. Chem.* **1974**, *39*, 2458-2459.
- (56) Mendiola, J.; Rincón, J. A.; Mateos, C.; Soriano, J. F.; de Frutos, Ó.; Niemeier, J. K.; Davis, E. M. Preparation, use, and safety of *O*-mesitylenesulfonylhydroxylamine. *Org. Process Res. Dev.* **2009**, *13*, 263-267.
- (57) Wallace, R. G. Hydroxylamine-*O*-sulfonic acid—a versatile synthetic reagent. *Aldrichim. Acta* **1980**, *13*, 3-11.
- (58) Kosower, E. M. Monosubstituted diazenes (diimides). Surprising intermediates. *Acc. Chem. Res.* **1971**, *4*, 193-198.
- (59) Doldouras, G. A.; Kollonitsch, J. A direct, selective, and general method for reductive deamination of primary amines. *J. Am. Chem. Soc.* **1978**, *100*, 341-342.
- (60) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. A stepwise Huisgen cycloaddition process: copper(I)-catalyzed regioselective "ligation" of azides and terminal alkynes. *Angew. Chem. Int. Ed.* **2002**, *41*, 2596-2599.
- (61) Tornøe, C. W.; Christensen, C.; Meldal, M. Peptidotriazoles on solid phase: [1,2,3]-triazoles by regiospecific copper(I)-catalyzed 1,3-dipolar cycloadditions of terminal alkynes to azides. *J. Org. Chem.* **2002**, *67*, 3057-3064.
- (62) Lin, Y. A.; Chalker, J. M.; Davis, B. G. Olefin cross-metathesis on proteins: investigation of allylic chalcogen effects and guiding principles in metathesis partner selection. *J. Am. Chem. Soc.* **2010**, *132*, 16805-16811.
- (63) Zachara, N. E.; Hart, G. W. The emerging significance of *O*-GlcNAc in cellular regulation. *Chem. Rev.* **2002**, *102*, 431-438.

- (64) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins. *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053.

2.5 Experimental Procedures

General Considerations

Melting points (m.p.) were recorded on a Leica Galen III hot stage microscope equipped with a Testo 720 thermocouple probe. Proton nuclear magnetic resonance (^1H NMR) spectra were recorded on a Bruker DPX200 (200 MHz), a Bruker AV400 (400 MHz), or a Bruker AVII500 (500 MHz) spectrometer, as indicated. Carbon nuclear magnetic resonance (^{13}C NMR) spectra were recorded on a Bruker DPX200 (50 MHz), a Bruker AV400 (100 MHz) or on a Bruker AVII500 (125 MHz) spectrometer, as indicated. NMR Spectra were assigned using COSY, HMQC, and DEPT 135. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard (^1H NMR: $\text{CDCl}_3 = 7.26$, $\text{CD}_3\text{OD} = 3.31$; $\text{DMSO}-d_6 = 2.50$; $\text{CD}_3\text{CN} = 1.94$ and ^{13}C NMR: $\text{CDCl}_3 = 77.0$; $\text{CD}_3\text{OD} = 49.0$; $\text{DMSO}-d_6 = 39.5$; $\text{CD}_3\text{CN} = 1.94$). Coupling constants (J) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet, and a = apparent.

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer using thin films on NaCl plates for liquids and oils and KBr discs for solids and crystals. Absorption maxima (ν_{max}) are reported in wavenumbers (cm^{-1}).

Low resolution mass spectra (LRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer using electrospray ionization (ESI) and high resolution mass spectra (HRMS) were recorded on a Bruker MicroTOF ESI mass spectrometer. Nominal and exact m/z values are reported in Daltons. Other methods of ionization (EI, FI, and FAB) are used where indicated and were recorded by the University of Oxford Mass Spectrometry Service in the Department of Chemistry.

Optical rotations were measured on a Perkin-Elmer 241 polarimeter with a path length of 1.0 dm and are reported with implied units of $10^{-1} \text{ deg cm}^2 \text{ g}^{-1}$. Concentrations (c) are given in g/100 ml.

Thin layer chromatography (TLC) was carried out using Merck aluminum backed sheets coated with 60F_{254} silica gel. Visualization of the silica plates was achieved using a UV

lamp ($\lambda_{\text{max}} = 254 \text{ nm}$), and/or ammonium molybdate (5% in 2M H_2SO_4), and/or potassium permanganate (5% KMnO_4 in 1M NaOH with 5% potassium carbonate). Flash column chromatography was carried out using BDH PROLAB[®] 40-63 mm silica gel (VWR). Mobile phases are reported in % volume of more polar solvent in less polar solvent for binary systems (e.g. 20% EtOAc in petrol = 1:4 ethyl acetate:petrol) and in relative composition for ternary systems (e.g. 1:2:4 H_2O : i PrOH:EtOAc)

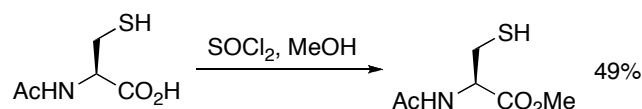
Anhydrous solvents were purchased from Fluka or Acros with the exception of dichloromethane and THF, which were dried in an alumina column under nitrogen. Triethylamine was stored over NaOH pellets. All other solvents were used as supplied (Analytical or HPLC grade), without prior purification. Deionized water was used for chemical reactions and Milli-Q purified water for protein manipulations. Protein concentrations were determined by BCA (bicinchoninic acid) assay (Pierce) and/or OD_{280} . Reagents were purchased from Sigma-Aldrich and used as supplied, unless otherwise indicated. 'Petrol' refers to the fraction of light petroleum ether boiling in the range 40-60 °C. All reactions using anhydrous conditions were performed using flame dried apparatus under an atmosphere of argon or nitrogen. Brine refers to a saturated solution of sodium chloride. Anhydrous magnesium sulfate (MgSO_4) or sodium sulfate (Na_2SO_4) were used as drying agents after reaction workup, as indicated.

DOWEX-50WX8 (H^+ form) was conditioned as follows: 100 g of the commercial resin was placed in a 500 mL sintered filter funnel and allowed to swell with 200 mL of acetone for 5 minutes. The solvent was removed by suction and the resin was washed successively with 800 mL of acetone, 500 mL methanol, 500 mL 5M HCl , and then 1 L of water or until the pH of filtrate was ~ 7 , as indicated by pH paper. The resin was partially dried on the filter and then stored and used as needed.

Further details regarding protein LC-MS analysis are provided in the Protein Modification section of the experimental procedures of this Chapter.

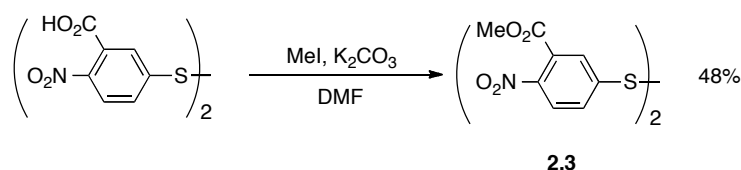
Reductive Route to Dha Using HMPT

AcCysOMe



N-Acetyl cysteine (5.00 g, 30.6 mmol) was added to a flame dried 250 mL round bottom flask under argon and dissolved in MeOH (50 mL). The stirred solution was cooled to 0 °C and thionyl chloride (2.46 mL, 33.7 mmol) was added slowly. The reaction was stirred for 10 minutes at 0 °C and then at room temperature for 1.5 hours. The solvent was then removed under reduced pressure. The resulting residue was redissolved in EtOAc (250 mL) and washed successively with H₂O (200 mL) and brine (200 mL). The organic layer was then dried (MgSO₄), filtered, and concentrated under reduced pressure to give a thick residue that crystallized on standing (2.64 g, 49%). This material was used without further purification. Spectroscopic data was consistent with that previously reported.¹ m.p. = 51-52 °C. $[\alpha]_D^{20} = +96.0$ (*c* = 1.3, CHCl₃). IR (ν_{max} , KBr): 3281, 1743, 1656, 1541, 1437, 1374, 1217. ¹H NMR (400 MHz, CDCl₃): δ = 1.37 (1H, t, *J* = 9.0, SH), 2.03 (3H, s, NHAc), 2.93-2.97 (2H, m, CH₂SH), 3.74 (3H, s, CO₂CH₃), 4.82-4.86 (1H, m, H α), 6.66 (1H, d, *J* = 6.3, NH). ¹³C NMR (100 MHz, CDCl₃): δ = 23.1 (NHAc), 26.8 (CH₂SH), 52.8, 53.6 (CO₂CH₃, C α), 170.1, 170.8 (C=O). LRMS *m/z* (ESI): 176 [M-H]⁻.

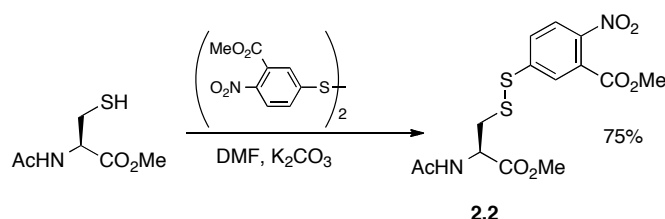
Methyl ester of Ellman's reagent (2.3)



Ellman's reagent (8.00 g, 20.2 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (40 mL). The stirred solution was cooled to 0 °C and K₂CO₃ (14.0 g, 101 mmol) was added. Methyl iodide (3.77 mL, 60.6 mmol) was added dropwise and the reaction stirred (open air) from 0 °C to room temperature over 2 hours. The reaction was diluted with EtOAc (300 mL) and then washed successively with NaHCO₃ (sat. aq., 2 × 100 mL), H₂O (2 × 100 mL), and brine (2 × 100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated to a thick yellow oil. The oil was dissolved in 25 mL of hot toluene (~50 °C) and petrol was added dropwise to the stirred solution until turbid. Additional hot toluene was added until the solution was clear. The solution was left to

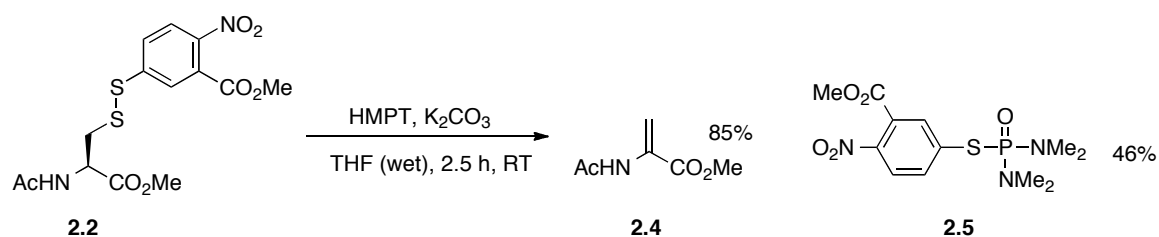
crystallize. The resulting yellow prisms were isolated by filtration and dried under vacuum (4.15 g, 48%). m.p. = 100-102 °C. IR ($\nu_{\max}$, KBr): 2850, 2361, 1738, 1568, 1527, 1435, 1341, 1281, 1134, 1101, 1060. ^1H NMR (400 MHz, CDCl_3): δ = 3.92 (6H, s, CO_2CH_3), 7.68 (2H, dd, J = 8.6, 2.0, ArH), 7.75 (2H, d, J = 2.0, ArH), 7.92 (2H, d, J = 8.6, ArH). ^{13}C NMR (100 MHz, CDCl_3): δ = 53.5 (CO_2CH_3), 125.1, 126.7, 128.3 (CH_{Ar}), 128.9, 142.4, 146.4 (4°_{Ar}), 165.1 ($\text{C}=\text{O}$). HRMS m/z (ESI^+): Found 446.9925 $[\text{M}+\text{Na}]^+$; $\text{C}_{16}\text{H}_{12}\text{N}_2\text{O}_8\text{S}_2\text{Na}$ requires 446.9927. Analysis for $\text{C}_{16}\text{H}_{12}\text{N}_2\text{O}_8\text{S}_2$ requires: C, 45.28%; H, 2.85%; N, 6.60%; Found: C, 45.40%; H, 3.04%; N, 6.60%.

Disulfide 2.2



Ellman's methyl ester **2.3** (2.50 g, 5.89 mmol) was added to a 100 mL round bottom flask along with K_2CO_3 (1.63 g, 11.76 mmol) and DMF (30 mL). To the stirred mixture was added a solution of AcCysOMe (695 mg, 3.92 mmol) in DMF (10 mL) over a 5 minute period. The reaction was stirred for 5 minutes at room temperature (open air) after which time TLC (EtOAc) revealed the desired product (R_f = 0.5) and the thiol from reduced **2.3** (R_f = 0.8). The reaction was diluted with EtOAc (250 mL) and then washed successively with NaHCO_3 (sat. aq. 2×100 mL) and brine (2×100 mL). The organic layer was then dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to give **2.2** as thick yellow oil which crystallized on slow cooling in hot EtOAc. The resulting yellow crystals were isolated by filtration and dried under vacuum (1.14 g, 75%). m.p. = 82-84 °C. $[\alpha]_D^{20}$ = +87.6 (c = 1.2, CHCl_3). IR ($\nu_{\max}$, KBr): 2980, 1740, 1659, 1525, 1341, 1280. ^1H NMR (400 MHz, CDCl_3): δ = 1.97 (3H, s, NHAc), 3.11-3.27 (2H, ABX system, J = 14.5, 6.4, 5.2, CH_2S), 3.68 (3H, s), 3.87 (3H, s) ($2 \times \text{CO}_2\text{CH}_3$), 4.77-4.81 (1H, m, H_α), 6.77 (1H, d, J = 7.6, NH), 7.65 (1H, dd, J = 8.7, 2.2, ArH), 7.71 (1H, d, J = 2.2, ArH), 7.89 (1H, d, J = 8.7, ArH). ^{13}C NMR (100 MHz, CDCl_3): δ = 22.7 (NHAc), 40.4 (CH_2S), 51.4 (C_α), 52.6, 53.3 ($2 \times \text{CO}_2\text{CH}_3$), 124.6, 126.0, 127.9 (CH_{Ar}), 128.6, 144.8, 145.3 (4°_{Ar}), 165.5, 170.1, 170.3 ($3 \times \text{C}=\text{O}$). HRMS m/z (ESI^+): Found 411.0292 $[\text{M}+\text{Na}]^+$; $\text{C}_{14}\text{H}_{16}\text{N}_2\text{O}_7\text{S}_2\text{Na}$ requires 411.0291.

Reaction of 2.2 with HMPT

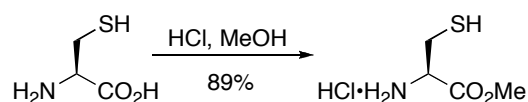


Disulfide **2.2** (32 mg, 0.082 mmol) was added to a 25 mL round bottom flask and dissolved in THF (2 mL, HPLC grade but not anhydrous). K_2CO_3 (57 mg, 0.41 mmol) was added to the stirred solution, followed by HMPT (30 μL , 0.16 mmol). Upon addition of HMPT, the color of the reaction turned red. TLC analysis (EtOAc) after HMPT addition revealed complete consumption of disulfide. The reaction flask was stoppered and the reaction was stirred at room temperature for 2.5 hours. After this time, TLC (EtOAc) indicated the formation of two major products ($R_f = 0.7$ and $R_f = 0.4$). The solvent was removed under reduced pressure and the resulting residue was purified by column chromatography (85% EtOAc in petrol). AcDhaOMe (**2.4**) (10 mg, 85%) and phosphoramidate **2.5** (13 mg, 46% based on disulfide) were isolated.

AcDhaOMe (**2.4**). Spectroscopic data is consistent with that previously reported.² Clear crystals. m.p. = 40–41 °C. IR (ν_{max} , film): 1730, 1677, 1636, 1519, 1440, 1372, 1235, 1203, 1171. ^1H NMR (500 MHz, CDCl_3): δ = 2.14 (3H, s, NHAc), 3.85 (3H, s, CO_2CH_3), 5.89 (1H, d, $J = 1.6$, $\text{C}=\text{CHH}$), 6.60 (1H, app. s, $\text{C}=\text{CHH}$), 7.73 (1H, br. s, NH). ^{13}C NMR (125 MHz, CDCl_3): δ = 24.7 (NHAc), 53.0 (CO_2CH_3), 108.7 ($\text{C}=\text{CH}_2$), 130.8 ($\text{C}=\text{CH}_2$), 164.6, 168.8 ($\text{C}=\text{O}$). LRMS m/z (ES^-): 142.1 $[\text{M}-\text{H}]^-$.

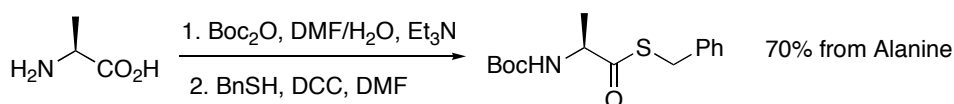
(**2.5**) Clear oil. IR (ν_{max} , film): 3384, 2929, 2490, 1647, 1423, 1072, 880, 813. ^1H NMR (500 MHz, CDCl_3): δ = 2.72 (6H, s, $2 \times \text{NCH}_3$), 2.75 (6H, s, $2 \times \text{NCH}_3$), 3.92 (3H, s, CO_2CH_3), 7.86–7.89 (2H, m, ArH), 7.93–7.96 (1H, m, ArH). ^{13}C NMR (125 MHz, CDCl_3): δ = 36.99, 37.02 (NMe_2), 53.4 (CO_2CH_3), 124.3, 128.1, 133.9 (d, $J_{\text{CP}} = 5.7$), 136.1 (d, $J_{\text{CP}} = 4.8$), 137.8 (d, $J_{\text{CP}} = 4.8$), 146.8 (Ar), 165.4 ($\text{C}=\text{O}$). ^{31}P NMR (202 MHz, CDCl_3): 37.55. HRMS m/z (ESI^+): Found 370.0593 $[\text{M}+\text{Na}]^+$; $\text{C}_{12}\text{H}_{18}\text{N}_3\text{O}_5$ PSNa requires 370.0597.

L-Cysteine methyl ester hydrochloride



MeOH (75 mL) was added to a flame-dried 250 mL round bottom flask and cooled to 0 °C. Acetyl chloride (29.3 mL, 413 mmol) was added slowly to the stirred solution and then stirred an additional 20 minutes at 0 °C to generate methanolic HCl. Cysteine (10.00 g, 82.54 mmol) was added in one portion and the reaction stirred at room temperature for 14 hours under a balloon of nitrogen. The solvent was removed under reduced pressure to give cysteine methyl ester hydrochloride as an off-white solid (12.60 g, 89%) that was used without purification.

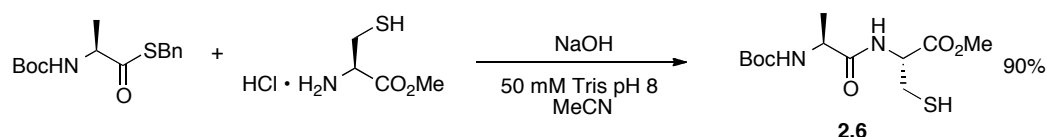
BocAlaSBn



L-Alanine (2.50 g, 28.0 mmol) was added to a 100 mL round bottom flask and suspended in DMF (10 mL). Et₃N (12.7 mL) and Boc₂O (7.96 g, 36.5 mmol) were then added sequentially and the reaction was stirred for 1 hour (open air, RT). H₂O (10 mL) was then added to homogenize the reaction mixture and to the resulting solution was added an additional portion of Boc₂O (1.20 g, 5.50 mmol). The reaction was stirred for an additional 30 minutes at which time the reaction was diluted with NaHCO₃ (sat., 100 mL). The reaction mixture was washed with Et₂O (3 × 100 mL). The organic wash was discarded. The aqueous layer was cooled to 0 °C and acidified to ~ pH 2 with 10 M HCl before extracting with EtOAc (3 × 100 mL). The combined organic layers were washed successively with H₂O (150 mL) and brine (150 mL) before drying over MgSO₄. After filtration, the product was concentrated under reduced pressure to give BocAlaOH which was carried on to the next step without further purification (*R_f* = 0.2, 1% AcOH in EtOAc). BocAlaOH (28 mmol, crude from above) was dissolved in DMF (25 mL) and placed under argon. Benzylthiol (7.23 mL, 61.6 mmol) was added to the stirred solution followed by DCC (7.51 g, 36.4 mmol) and an additional 20 mL of DMF. The reaction was stirred for 2 hours at room temperature before diluting with Et₂O (300 mL) and H₂O (250 mL). After separation, the organic layer was washed successively with 150 mL each of NaHCO₃ (sat.), H₂O, and brine. The organic layer was dried (MgSO₄), filtered, and the solvent removed under reduced pressure. The product was purified by column

chromatography (100% petrol to 15% EtOAc in petrol) to give the titled compound as a white solid ($R_f = 0.6$). The solid was recrystallized from CH_2Cl_2 /petrol to give BocAlaSBn as bright white needles (5.80 g, 70% from Ala). m.p. = 79-80 °C. $[\alpha]_D^{20} = -25.1$ ($c = 1.0$, CHCl_3). IR (ν_{max} , KBr): 3375, 2980, 1689, 1505, 1341, 1383, 1250, 1164; ^1H NMR (CDCl_3 , 400 MHz): $\delta = 1.38$ (3H, d, $J = 7.3$, CH_3), 1.46 (9H, s, Boc), 4.11 (2H, app. d, $J = 4.8$, SCH_2Ph), 4.43 (1H, m, $\text{H}\alpha$), 5.04 (1H, d, $J = 7.6$, NH), 7.32-7.22 (5H, m, Ar). ^{13}C NMR (100 MHz): $\delta = 18.7$ (CH_3), 28.3 (Boc), 33.2 (SCH_2Ph), 56.2 ($\text{C}\alpha$), 80.3 (Boc), 127.3, 128.6, 128.9 (CH_{Ar}), 137.1 (4°Ar), 154.9, 201.5 ($2 \times \text{C}=\text{O}$). HRMS m/z (ESI+): Found 318.1134 ($\text{M}+\text{Na}$) $^+$; $\text{C}_{15}\text{H}_{21}\text{NO}_3\text{SNa}$ requires 318.1140. Analysis for $\text{C}_{15}\text{H}_{21}\text{NO}_3\text{S}$: C: 60.99, H: 7.17, N: 4.74; Found: C: 60.86; H: 7.32; N: 4.66.

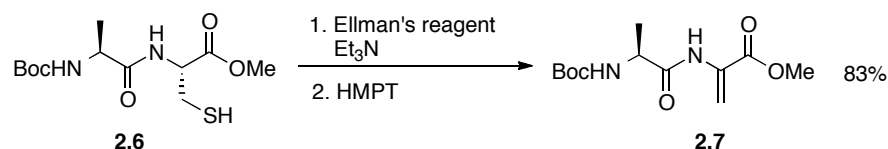
BocAlaCysOMe (2.6)



L-Cysteine methyl ester hydrochloride (5.00 g, 29.1 mmol) was added to a 100 mL, 2-neck round bottom flask and dissolved in 15 mL of pH 8.0 TRIS (50 mM). The solution was cooled to 0 °C and 5.0 mL of 5.82 M NaOH was added dropwise. BocAlaSBn (2.50 g, 8.46 mmol) was added as a solution in MeCN (25 mL). The resulting solution (pH 9, pH paper) was stirred at room temperature for 5 hours after which time TLC indicated complete consumption of thioester ($R_f = 0.4$, 10% EtOAc in petrol) and formation of the ligated product **2.6** ($R_f = 0.5$, 50% EtOAc in petrol). Tributylphosphine (1.05 mL, 4.26 mmol) was added to reduce any disulfide and after 15 min the reaction was diluted with Et_2O (250 mL) and H_2O (150 mL). The layers were separated and the aqueous layer was extracted with Et_2O (100 mL). The combined organics were washed with H_2O (2×150 mL) and brine (150 mL) and then dried (MgSO_4), filtered, and concentrated under reduced pressure. Purification by column chromatography (10% EtOAc in petrol to 50% EtOAc in petrol) provided BocAlaCysOMe (**2.6**) as white crystals (2.59 g, 90%). m.p. = 101-102 °C. $[\alpha]_D^{20} = -1.9$ ($c = 1.0$, CHCl_3). IR (ν_{max} , KBr): 3387, 3298, 2979, 2565, 1746, 1700, 1653, 1503, 1443, 1390, 1362, 1308. ^1H NMR (CDCl_3 , 400 MHz): $\delta = 1.35$ (3H, d, $J = 7.1$, CH_3 Ala), 1.41 (9H, s, Boc), 1.50 (1H, t, $J = 8.7$, SH), 3.03-2.90 (2H, m, CH_2SH), 3.75 (3H, s, CO_2CH_3), 4.21 (1H, br. m, $\text{H}\alpha$ Ala), 4.82 (1H, ddd, $J = 7.8, 4.3, 4.0$, $\text{H}\alpha$ Cys), 5.24 (1H, d, $J = 6.1$, NH Ala), 7.13 (1H, d, $J = 5.8$, NH Cys). ^{13}C NMR (100 MHz):

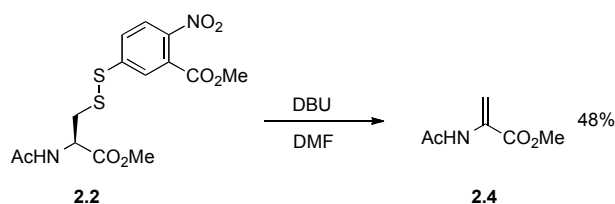
δ = 17.9 (CH_3_{Ala}), 26.6 (CH_2SH), 28.3 (Boc), 50.1 ($\text{C}\alpha_{\text{Ala}}$), 52.8 (CO_2CH_3), 53.7 ($\text{C}\alpha_{\text{Cys}}$), 80.1 (Boc), 155.5, 170.3, 172.7 ($3 \times \text{C}=\text{O}$). HRMS m/z (ESI^+): Found 329.1142 ($\text{M}+\text{Na}$) $^+$; $\text{C}_{12}\text{H}_{22}\text{N}_2\text{O}_5\text{SNa}$ requires 329.1147. Analysis for $\text{C}_{12}\text{H}_{22}\text{N}_2\text{O}_5\text{S}$: C: 47.04, H: 7.24, N: 9.14; Found: C: 47.05, H: 7.25, N: 9.09.

BocAlaDhaOMe (2.7) using HMPT

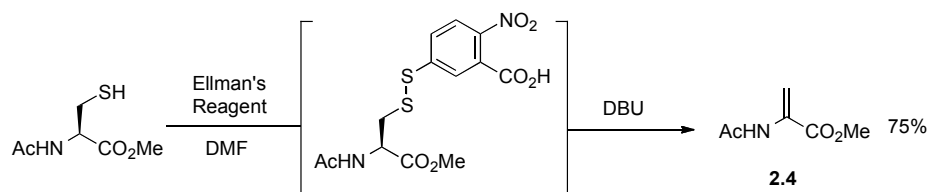


BocAlaCysOMe (**2.6**) (200 mg, 0.065 mmol) was added to a 25 mL pear shaped flask and dissolved in DMF (5 mL). NEt_3 (0.20 mL, 1.4 mmol) was added to the stirred solution followed immediately by Ellman's reagent (0.028 g, 0.071 mmol). The reaction mixture was stirred, open air, for 5 minutes at room temperature at which time TLC (50% EtOAc in petrol) revealed consumption of BocAlaCysOMe ($R_f = 0.5$). HMPT (0.015 mL, 0.083 mmol) was added by microsyringe and the reaction was stirred for an additional 15 minutes at room temperature. TLC (50% EtOAc in petrol) indicated formation of BocAlaDhaOMe ($R_f = 0.7$). The reaction was diluted with Et_2O (150 mL) and washed sequentially with 1 M HCl (75 mL), H_2O (75 mL), and brine (75 mL). The organic layer was dried over MgSO_4 , filtered, and concentrated under reduced pressure. The resulting residual mixture was purified by column chromatography (40% EtOAc in petrol) to afford BocAlaDhaOMe (**2.7**) as a white solid (150 mg, 83%). m.p. = 64-65 °C. $[\alpha]_D^{20} = -53.9$ ($c = 1.1$, CHCl_3). IR (ν_{max} , film): 3332, 2980, 1690, 1523, 1442, 1367, 1249, 1167. ^1H NMR (400 MHz, CDCl_3): δ = 1.40 (3H, d, $J = 7.3$, Me_{Ala}), 1.46 (9H, s, Boc), 3.84 (3H, s, CO_2Me), 4.26 (1H, br. app. s, $\text{H}\alpha_{\text{Ala}}$), 5.01 (1H, br. app. s, NH_{Ala}), 5.91 (1H, d, $J = 1.26$, $\text{C}=\text{CHH}$), 6.60 (1H, app s, $\text{C}=\text{CHH}$), 8.46 (1H, br. s, NH_{Dha}). ^{13}C NMR (100 MHz, CDCl_3): δ = 17.9 (Me_{Ala}), 28.2 (Boc), 50.9 ($\text{C}\alpha_{\text{Ala}}$), 52.9 (CO_2Me), 80.4 (Boc), 109.2 ($\text{C}=\text{CH}_2$), 130.8 ($\text{C}=\text{CH}_2$), 155.5, 164.2, 171.5 ($3 \times \text{C}=\text{O}$). HRMS m/z (ESI^+): Found 295.1262 [$\text{M}+\text{Na}$] $^+$; $\text{C}_{12}\text{H}_{20}\text{N}_2\text{O}_5\text{Na}$ requires 295.1264. Anal. Calcd. for $\text{C}_{12}\text{H}_{20}\text{N}_2\text{O}_5$: C, 52.93; H, 7.40; N, 10.29. Found C, 52.62; H, 7.28; N, 10.16.

Base-mediated conversion of cysteine to dehydroalanine



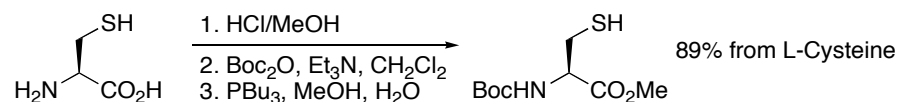
Disulfide **2.2** (40 mg, 0.10 mmol) was added to a 25 mL round bottom flask and dissolved in 3 mL of DMF. DBU (0.03 mL, 0.21 mmol) was added and the reaction turned deep red. The reaction was stirred 20 minutes at room temperature and then diluted with EtOAc (100 mL). The organic layer was washed sequentially with 1 M HCl (100 mL), H₂O (100 mL), and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (65% EtOAc in petrol) to yield AcDhaOMe (**2.4**) (7 mg, 48%). Spectroscopic data was identical to that reported above.



AcCysOMe (50 mg, 0.28 mmol) was added to a 25 mL round bottom flask and dissolved in 5 mL of DMF. Ellman's reagent (123 mg, 0.31 mmol) was added and the reaction stirred 10 minutes at room temperature. DBU (0.42 mL, 2.8 mmol) was then added, upon which the reaction mixture turned dark red. After 3 minutes, TLC revealed formation of AcDhaOMe. After a total of 10 minutes of reaction time the mixture was diluted with EtOAc (150 mL) and then washed sequentially with 1M HCl (2 × 100 mL) and brine (100 mL). The organic layer was then dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (65% EtOAc in petrol) to yield AcDhaOMe (**2.4**) (30 mg, 75%). Spectroscopic data was identical to that reported above.

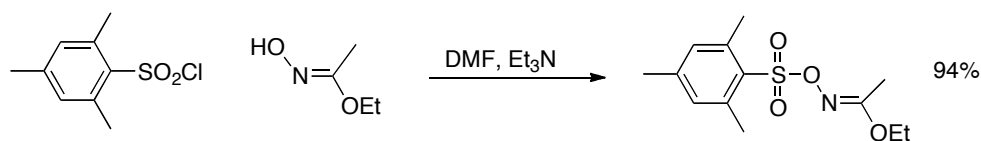
MSH in the Oxidative Elimination of Cysteine to Dehydroalanine

BocCysOMe

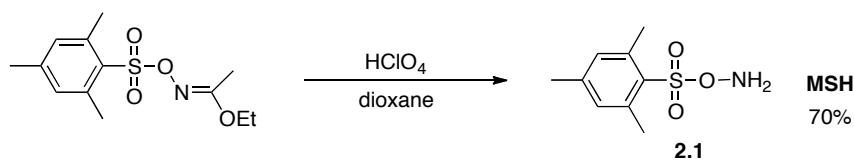


MeOH (100 mL) was added to a flame dried 250 mL round bottom flask under an atmosphere of argon. The solvent was stirred and cooled to 0 °C and acetyl chloride (248 mmol, 17.6 mL) was added dropwise over 5 minutes. The solution was stirred an additional 10 minutes at 0 °C to give a solution of methanolic HCl. L-cysteine (2.00 g, 16.5 mmol) was then added in one portion and the flask flushed with argon. The ice bath was removed and the reaction was stirred at room temperature for 24 hours under a balloon of argon. After this time, the solvent was then removed under reduced pressure to give the crude cysteine methyl ester hydrochloride as a pale yellow solid. This material was used immediately in the next step without purification. The crude ester was suspended in CH₂Cl₂ (100 mL) and cooled to 0 °C. Et₃N (5.06 mL, 36.3 mmol) was added carefully followed by Boc₂O (4.32 g, 19.8 mmol). The reaction was stirred at room temperature for 3.25 hours after which time TLC (30% EtOAc in Petrol) revealed the desired product (R_f = 0.6) and its corresponding disulfide (R_f = 0.3). The solvent was removed under reduced pressure and the resulting residue was redissolved in MeOH (40 mL) and H₂O (8 mL). PBu₃ (2.0 mL, 8.1 mmol) was added dropwise to the stirred solution. TLC revealed reduction of the disulfide. The reaction was diluted with Et₂O (100 mL) and H₂O (50 mL). The organic layer was separated and the aqueous layer was extracted with Et₂O (2 × 50 mL). The combined organics were washed with brine (100 mL), dried over MgSO₄, and filtered. The solvent was removed by rotary evaporation and the residue purified by column chromatography eluting first with 5% EtOAc in petrol and then 20% EtOAc in petrol. BocCysOMe was isolated as a clear oil (3.48 g, 89% from L-cysteine). $[\alpha]_D^{20}$ = +28.3 (c = 7.5, CHCl₃) (Lit.³: $[\alpha]_D^{20}$ = +28.5, c = 7.5, CHCl₃). IR (ν_{max} , film): 2979, 1717, 1506, 1367, 1166. ¹H NMR (CDCl₃, 400 MHz): δ = 1.42 (10H, s, Boc and SH), 2.94 (2H, app. td, J = 8.7, 4.3, CH₂SH), 3.76 (3H, s, CO₂Me), 4.58 (1H, m, H α), 5.44 (1H, d, J = 5.8, NH). ¹³C NMR (100 MHz): δ = 27.3 (CH₂SH), 28.2 (Boc), 52.6 (CO₂Me), 54.8 (C α), 80.2 (Boc), 155.1, 170.8 (2 × C=O). LRMS (ESI⁺): 234 (M-H)⁺.

***O*-mesitylenesulfonylhydroxylamine (MSH) Synthesis**



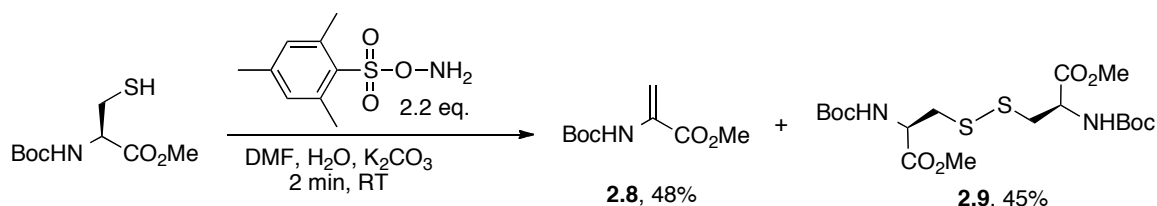
Ethyl *N*-hydroxyacetimidate (2.36 g, 22.9 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (15 mL). Et_3N (6.4 mL, 46 mmol) was added and the solution was cooled to 0 °C. 2-Mesitylenesulfonyl chloride (5.00 g, 22.9 mmol) was added as a solid in several portions. The reaction was stirred vigorously at 0 °C for 30 minutes. A white solid was formed over the course of the reaction. After this time, the reaction was diluted with EtOAc (150 mL) and then washed with H_2O (3×100 mL) and brine (100 mL). The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure to give a pale yellow oil that crystallized on storage at -20 °C (6.11 g, 94%). No purification was required. m.p. = 51-53 °C (Lit⁴ = 57-58 °C). ^1H NMR (400 MHz, CDCl_3): δ = 1.19 (3H, t, J = 7.1, OCH_2CH_3), 2.04 (3H, s, CH_3 imidate), 2.31 (3H, s, Me_{Ar}), 2.64 (6H, s, Me_{Ar}), 3.90 (2H, q, J = 7.1, OCH_2CH_3), 6.97 (2H, s, CH_{Ar}). ^{13}C NMR (100 MHz, CDCl_3): δ = 14.0 (OCH_2CH_3), 14.9 (CH_3 imidate), 21.1 (Me_{Ar}), 22.8 (Me_{Ar}), 63.6 (OCH_2CH_3), 130.3, 131.5, 140.7, 143.3 ($4 \times \text{C}_{\text{Ar}}$), 169.7 ($\text{C}=\text{N}$).



The MSH precursor (6.11 g, 21.4 mmol) was dissolved in dioxane (8 mL) at room temperature and then cooled to 0 °C. Perchloric acid (4 mL, 70% solution) was added dropwise over 2 minutes. The reaction was stirred vigorously at 0 °C. The mixture thickened and solidified over the course of 5 minutes. The reaction was incubated for an additional 5 minutes to ensure complete hydrolysis and then poured onto 50 g of ice. Et_2O (100 mL) and H_2O (100 mL) were both added to the quenched mixture and then transferred to a separatory funnel. The organic layer was separated and then dried / neutralized for 2 minutes with K_2CO_3 . The solid was then removed by filtration and rinsed with Et_2O . The filtrate was concentrated carefully under reduced pressure to about 50 mL total volume. The solution was poured into a beaker with 150 mL ice cold petrol. MSH crystals formed immediately. The beaker was incubated at 0 °C for 30 minutes to complete the crystallization. The crystals were isolated by filtration and then dried under

high vacuum for 15 minutes. The white needles (3.21 g, 70%) were stored in a plastic tube, sealed only with wax film. m.p. = 90-91 °C (Lit⁴ = 95-96 °C). ¹H NMR (400 MHz, CDCl₃): δ = 2.33 (3H, s, Me_{Ar}), 2.65 (6H, s, Me_{Ar}), 5.56 (2H, br. s, NH₂), 7.01 (2H, s, CH_{Ar}). ¹³C NMR (50 MHz, CDCl₃): δ = 21.0 (Me_{Ar}), 22.6 (Me_{Ar}), 129.0, 131.6, 140.9, 143.7 (4 × C_{Ar}). Found: C, 50.18%; H, 5.90%, N, 6.27%. C₉H₁₅NO₄ requires: C, 50.21%; H, 6.09%; N, 6.51%.

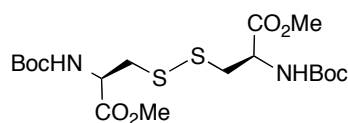
Chemical conversion of BocCysOMe to BocDhaOMe (2.8) using MSH



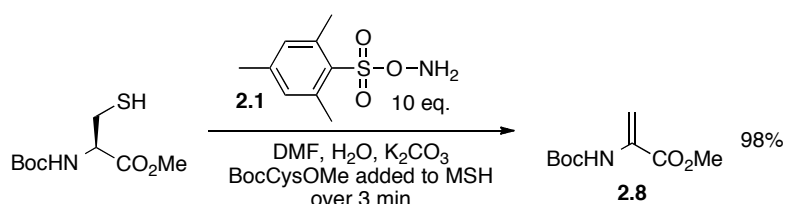
BocCysOMe (100 mg, 0.42 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (8 mL, open air, room temperature). A solution of potassium carbonate (294 mg, 2.12 mmol) in H₂O (2 mL) was added by pipette and stirred 30 seconds. MSH (198 mg, 0.92 mmol) was then added in one portion and stirred for 1 minute at which time the reaction mixture changed from a cloudy suspension to a clear solution. The reaction was checked immediately by TLC (30% EtOAc in petrol), revealing complete consumption of the starting material and the formation of two products: a strongly UV active material (R_f = 0.6) and another product (R_f = 0.2). After 2 minutes of total reaction time, the mixture was diluted with Et₂O (80 mL) and H₂O (80 mL). The organic layer was separated and the aqueous layer was extracted with Et₂O (3 × 40 mL). The combined organics were washed with brine (100 mL), dried (MgSO₄), and filtered. After removal of solvent, the residue was purified by column chromatography by eluting first with 3% ethyl acetate in petrol to give BocDhaOMe (**2.8**) (41 mg, 48%) as a clear oil and then 30% ethyl acetate in petrol to give disulfide **2.9** (45 mg, 45%) as small white needles.

$\text{BocHN-C(=CH}_2\text{)-CO}_2\text{Me}$ Data for BocDhaOMe (**2.8**): IR (ν_{max} , film): 3423, 2980, 1719, 1634, 1513, 1328, 1159, 1068. ¹H NMR (CDCl₃, 400 MHz): δ = 1.46 (9H, s, Boc), 3.80 (3H, s, CO₂CH₃), 5.70 (1H, d, J = 1.5, C=CHH), 6.13 (1H, app s, C=CHH), 7.00 (1H, br s, NH). ¹³C NMR (100 MHz): δ = 28.2 (Boc), 52.8 (CO₂CH₃), 80.6 (Boc), 105.1 (C=CH₂), 131.3

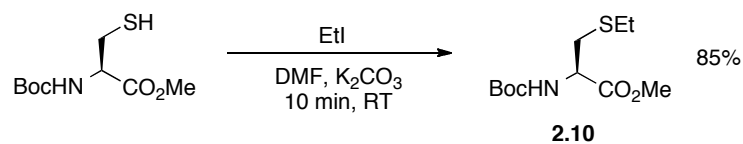
(C=CH₂), 152.5, 164.4 (2 × C=O). Anal. for C₉H₁₅NO₄: C, 53.72; H, 7.51; N, 6.96; found C, 53.95; H, 7.63; N, 6.83.



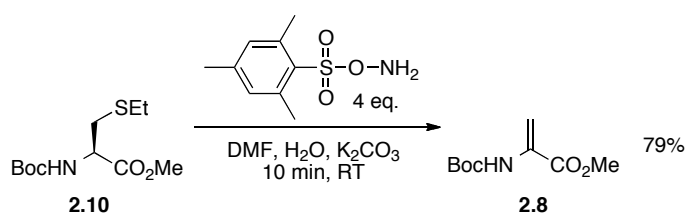
Data for BocCysOMe disulfide (**2.9**): m.p. 89-90 °C [Lit. 96-97 °C]⁵. $[\alpha]_D^{20} +91.4$ ($c = 1$ in CHCl₃). IR ($\nu_{\max}$, KBr): 3377, 2983, 1748, 1686, 1515, 1367, 1167. ¹H NMR (CDCl₃, 400 MHz): $\delta = 1.45$ (18H, s, 2 × Boc), 3.16 (4H, d, $J = 5.1$, 2 × CH₂S), 3.77 (6H, s, 2 × CO₂CH₃), 4.60 (2H, m, 2 × H α), 5.40 (2H, d, $J = 7.6$, NH). ¹³C NMR (100 MHz): $\delta = 28.3$ (Boc), 41.3 (CH₂S), 52.6 (CO₂CH₃), 52.8 (C α), 80.3 (Boc), 155.0, 171.2 (2 × C=O). LRMS m/z (ESI⁺): 469.2 [M+H]⁺.



MSH (**2.1**) (439 mg, 2.0 mmol) was added to a 10 mL round bottom flask and dissolved in DMF (3 mL). In a separate vial, BocCysOMe (48 mg, 0.20 mmol) was added and dissolved in DMF (3 mL). The vial containing BocCysOMe was cooled to 0 °C and a solution of K₂CO₃ (138 mg, 1.00 mmol) in H₂O (3.0 mL) was added. The resulting solution was added dropwise by pipette over a period of 3 min to the stirred MSH solution at room temperature. The vial was rinsed with DMF (2 × 1 mL) to ensure complete transfer. TLC analysis (20% EtOAc in petrol) after completion of the addition revealed a single, UV active product ($R_f = 0.6$). The reaction mixture was transferred to a separatory funnel and diluted with diethyl ether (150 mL) and H₂O (100 mL). After separation, the organic layer was washed successively with water (80 mL) and brine (80 mL) before drying (MgSO₄) and filtering. The solvent was removed under reduced pressure and the resulting residue purified by column chromatography to provide BocDhaOMe **2.8** as a clear oil (40 mg, 98%). Spectroscopic data were identical to that reported above.

BocCys(Et)OMe (2.10)

BocCysOMe (573 mg, 2.44 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (15 mL). K₂CO₃ (674 mg, 4.88 mmol) was added to the solution followed by ethyl iodide (0.30 mL, 3.7 mmol). The reaction was stirred at room temperature for 10 min. TLC (15% EtOAc in petrol) revealed complete consumption of starting material ($R_f = 0.3$) and the formation of new product ($R_f = 0.4$). The reaction mixture was diluted with Et₂O (150 mL) and H₂O (100 mL). The organic layer was separated and the aqueous layer extracted with Et₂O (2 × 100 mL). The combined organics were dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (15% EtOAc in petrol) to give thioether **2.10** as a clear oil (546 mg, 85%). $[\alpha]_D^{20} = +22.0$ ($c = 1.1$, CHCl₃). IR (ν_{max} , film): 2972, 2932, 2361, 2342, 1748, 1717, 1505, 1167. ¹H NMR (CDCl₃, 400MHz): $\delta = 1.23$ (3H, t, $J = 7.4$, SCH₂CH₃), 1.44 (9H, s, Boc), 2.53 (2H, q, $J = 7.4$, SCH₂CH₃), 2.96 (2H, app s, br, CH₂SEt), 3.75 (3H, s, CO₂CH₃), 4.52 (1H, m, H α), 5.36 (1H, d, $J = 7.1$, NH). ¹³C NMR (100 MHz): $\delta = 14.6$ (SCH₂CH₃), 26.5 (SCH₂CH₃), 28.3 (Boc), 33.9 (CH₂SEt), 52.5 (CO₂CH₃), 53.2 (C α), 80.1 (Boc), 155.1, 171.6 (C=O). HRMS m/z (ESI⁺): Found 286.1083 (M+Na)⁺; C₁₁H₂₁NO₄SNa requires 286.1089. Anal. calculated for C₁₁H₂₁NO₄S: C, 50.17; H, 8.04; N, 5.32. Found: C, 49.88; H, 8.09; N, 5.01.



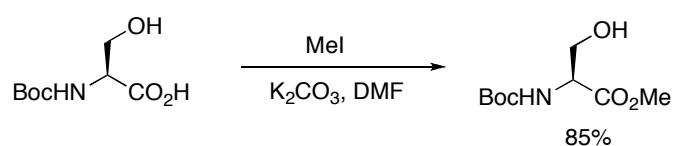
BocCys(Et)OMe (**2.10**) (106 mg, 0.40 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (5 mL). K₂CO₃ (278 mg, 2.01 mmol) was added by pipette as a solution in H₂O (1.0 mL). MSH (172 mg, 0.80 mmol) was added as a solid in one portion (open air, RT). TLC analysis (20% EtOAc in petrol) after 1 min of reaction revealed a strongly UV active product ($R_f = 0.6$) and a trace of starting material ($R_f = 0.5$). A second dose of MSH (172 mg, 0.80 mmol) was added after 5 min of reaction time and TLC analysis revealed only the UV active product. After 10 min of total reaction time, the reaction mixture was diluted with Et₂O (100 mL) and H₂O (50 mL). The organic layer

was separated and the aqueous layer was extracted with Et₂O (2 × 50 mL). The combined organics were dried (MgSO₄), filtered, and concentrated under reduced pressure. Column chromatography (3% EtOAc in petrol) provided BocDhaOMe (**2.8**) (63 mg, 79%). This material was spectroscopically identical to that reported above.

Assessment of Reactivity of MSH with Functionalized Amino Acids:

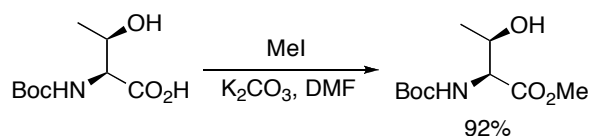
Substrate Synthesis

BocSerOMe



(L)-BocSerOH (1.00 g, 4.87 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (10 mL). K₂CO₃ (1.01 g, 7.31 mmol) was added, followed by MeI (1.52 mL, 24.4 mmol). The reaction was stirred at room temperature for 5 hours. After this time, the reaction was diluted with EtOAc (200 mL) and then washed with H₂O (2 × 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting oil (BocSerOMe) did not require purification (910 mg, 85%). $[\alpha]_D^{20} = +11.3$ (c = 1.6, CHCl₃) (Lit⁶: +9.7 c = 4.3, CHCl₃). IR ($\nu_{\max}$, film): 3396, 2979, 1746, 1714, 1515, 1368, 1351, 1287, 1249, 1214. ¹H NMR (CDCl₃, 200 MHz): δ = 1.40 (9H, s, Boc), 3.30 (1H, s, OH), 3.72 (3H, s, CO₂Me), 3.78-3.94 (2H, m, CH₂OH), 4.31 (1H, m, H α), 5.61 (1H, d, J = 7.9, NH). ¹³C NMR (CDCl₃, 50 MHz): δ = 28.2 (Boc), 52.4 (CO₂Me), 55.6 (C α), 63.0 (CH₂OH), 80.1 (Boc), 155.7, 171.4 (2 × C=O). LRMS m/z (ESI⁺): 242.1 [M+Na]⁺.

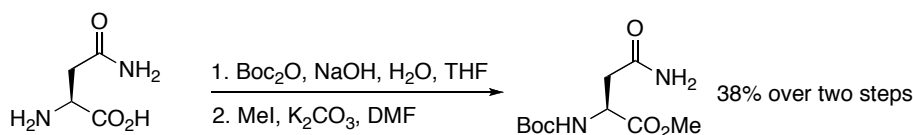
BocThrOMe



(L)-BocThrOH (1.00 g, 4.56 mmol) was added to a 10 mL round bottom flask and suspended in DMF (10 mL). K₂CO₃ (946 mg, 6.84 mmol) was added to the stirred mixture, followed by iodomethane (0.85 mL, 13.7 mmol). The reaction was stirred at

room temperature for 30 minutes. After this time, the reaction was diluted with EtOAc (100 mL) and then washed sequentially with sat. NaHCO₃ (50 mL) and brine (2 × 50 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting oil was sufficiently pure to not require purification (981 mg, 92%). Spectroscopic data was consistent with that previously reported.⁷ $[\alpha]_D^{20} = -10.1$ ($c = 1.0$, CH₂Cl₂) (Lit⁷: -8.77, $c = 0.75$, CH₂Cl₂). IR ($\nu_{\max}$, film): 3436, 2979, 2935, 1748, 1718, 1509, 1367, 1166. ¹H NMR (400 MHz, CDCl₃): $\delta = 1.21$ (3H, d, $J = 6.3$, CH₃), 1.41 (9H, s, Boc), 2.87 (1H, d, $J = 4.3$, OH), 3.73 (3H, s, CO₂Me), 4.21 (1H, m, H α), 4.26 (1H, m, H β), 5.44 (1H, d, $J = 8.6$, NH). ¹³C NMR (100 MHz, CDCl₃): $\delta = 19.8$ (CH₃), 28.2 (Boc), 52.3 (CO₂Me), 60.3 (C α), 67.9 (C β), 79.9 (Boc), 156.1, 172.0 (2 × C=O). LRMS m/z (ESI⁺): 256.1 [M+Na]⁺.

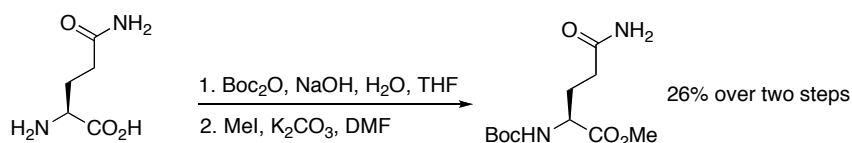
BocAsnOMe



L-Asparagine (2.00 g, 15.1 mmol) was added to a 100 mL round bottom flask and suspended in H₂O (20 mL). NaOH (1.27 g, 31.8 mmol) was added to the flask and the mixture was stirred until all material was dissolved. Boc₂O (3.97 g, 18.2 mmol) was then added as a solution in THF (20 mL). The reaction was stirred at room temperature for 1 hour and then quenched by the addition of 1M HCl (50 mL). The product was extracted with EtOAc (3 × 40 mL). The combined organic layers were then washed with brine (70 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The crude product (2.98 g) was carried on immediately to the next step. Accordingly, the crude (L)-BocAsnOH was dissolved in DMF (10 mL). To the stirred solution was added K₂CO₃ (2.66 g, 19.3 mmol) and methyl iodide (1.60 mL, 25.7 mmol). The reaction was stirred at room temperature for 1 hour and then diluted with EtOAc (100 mL). The mixture was washed sequentially with NaHCO₃ (50 mL sat. aq.), H₂O (50 mL), and brine (50 mL). The organic layer was then dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc) to provide (L)-BocAsnOMe⁸ as a white foam (1.42 g, 38% over two steps). $[\alpha]_D^{20} = -16.2$ ($c = 0.5$, MeOH). IR ($\nu_{\max}$, film): 3354, 2979, 1741, 1680, 1619, 1504, 1438, 1410, 1393, 1367, 1348, 1317, 1286, 1251, 1222, 1167, 1120. ¹H NMR (400 MHz, CDCl₃): $\delta = 1.42$

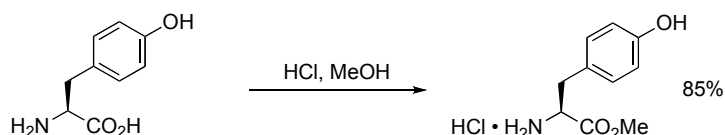
(9H, s, Boc), 2.71 (1H, ABX system, $J = 15.9, 4.3$, H β), 2.91 (1H, ABX system, $J = 15.9, 4.8$, H β'), 3.73 (3H, s, CO₂Me), 4.50 (1H, m, H α), 5.79 (1H, d, $J = 8.3$, NHBoc), 6.06-6.09 (2H, app. d, $J = 9.9$, NH₂). ¹³C NMR (100 MHz, CDCl₃): $\delta = 28.2$ (Boc), 37.3 (C β), 50.2 (C α), 52.6 (CO₂Me), 80.0 (Boc), 155.6, 172.1, 172.5 (3 $\times$ C=O). LRMS m/z (ESI⁺): 269.1 [M+Na]⁺.

BocGlnOMe



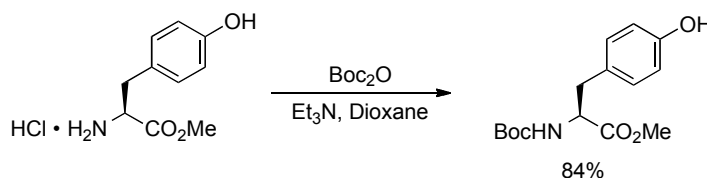
(L)-Glutamine (1.00 g, 6.84 mmol) was added to a 100 mL round bottom flask and suspended in H₂O (10 mL). NaOH (574 mg, 14.4 mmol) was added and the mixture was stirred until all material was dissolved. Boc₂O (1.79 g, 8.20 mmol) was added as a solution in THF (10 mL). The reaction was stirred at room temperature for 1 hour and then quenched by the addition of 1M HCl (50 mL). The product was extracted with EtOAc (3 $\times$ 40 mL). The combined organic layers were then washed with brine (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The crude product (1.68 g) was carried on immediately to the next step. Accordingly, the crude BocGlnOH was dissolved in DMF (14 mL). K₂CO₃ (1.24 g, 8.97 mmol) was added followed by methyl iodide (0.74 mL, 12 mmol). The reaction was stirred for 30 minutes at room temperature and then diluted with EtOAc (100 mL). The organic fraction was washed sequentially with NaHCO₃ (50 mL sat. aq.) and brine (2 $\times$ 50 mL). After drying (MgSO₄), the organic layer was filtered and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc) to afford (L)-BocGlnOMe as a waxy solid (460 mg, 26% over two steps). Spectroscopic data was consistent with that previously reported.⁹ [α]_D²⁰ = -21.2 (c = 0.5, MeOH). IR (ν_{max} , film): 3346, 2978, 2248, 1700, 1682, 1620, 1520, 1454, 1392, 1367, 1218, 1166, 1097, 1054, 1028. ¹H NMR (400 MHz, CDCl₃): $\delta = 1.37$ (9H, s, Boc), 1.87-1.92 (1H, m, H β), 2.05-2.15 (1H, m, H β'), 2.24-2.34 (2H, m, H γ), 3.68 (3H, s, CO₂Me), 4.21-4.26 (1H, m, H α), 5.56 (1H, d, $J = 7.8$, NHBoc), 6.05 (1H, s, CONHH), 6.40 (1H, s, CONHH). ¹³C NMR (100 MHz, CDCl₃): $\delta = 28.1$ (Boc), 28.3 (C β), 31.6 (C γ), 52.4 (CO₂Me), 52.9 (C α), 79.9 (Boc), 155.7, 172.7, 174.7 (3 $\times$ C=O). LRMS m/z (ESI⁺): 283.1 [M+Na]⁺.

(L)-Tyrosine methyl ester hydrochloride



A 250 mL round bottom flask was flame dried and placed under an atmosphere of nitrogen. Anhydrous MeOH (80 mL) was added to the flask and cooled to 0 °C. Acetyl chloride (193 mmol, 13.74 mL) was added slowly with stirring to generate methanolic HCl. After all of the acetyl chloride was added, the solution was stirred an additional 15 minutes at 0 °C. (L)-Tyrosine (7.00 g, 38.6 mmol) was added as a solid in one portion. The ice bath was removed and the reaction stirred at room temperature for 12 hours. After this time, volatiles were removed under reduced pressure to give the crude product as an off-white solid. The solid was triturated by stirring rapidly in 50 mL of boiling acetone for 10 minutes. After cooling to room temperature, the flask was cooled on ice for several minutes. The resulting micro-crystals were isolated by filtration and washed with 50 mL of cold acetone. The crystals were then dried under vacuum (7.65 g, 85% yield). m.p. = 172-174 °C, decomposes at 190 °C (Lit¹⁰: 190-192, decomp.). $[\alpha]_D^{20} = +69.8$ (c = 3.0, pyridine) (Lit¹¹: +72.0, c = 3.0, pyridine). IR ($\nu_{\max}$, KBr): 3343, 2879, 1958, 1744, 1614, 1592, 1515, 1448, 1397, 1349, 1226, 1141, 1107, 1060. ¹H NMR (400 MHz, MeOD): δ = 3.10-3.22 (2H, ABX system, J = 14.3, 7.2, 6.1, $\underline{\text{CH}_2\text{Ar}}$), 3.82 (3H, s, $\text{CO}_2\underline{\text{CH}_3}$), 4.28 (1H, app t. J = 6.7, H α), 6.81 (2H, d, J = 8.5, ArH), 7.10 (2H, d, J = 8.5, ArH). ¹³C NMR (50 MHz, MeOD): δ = 36.7 ($\underline{\text{CH}_2\text{Ar}}$), 53.7, 55.6 (C α and $\text{CO}_2\underline{\text{CH}_3}$), 117.0, 125.8, 131.7, 158.4 (Ar), 170.6 (C=O). LRMS m/z (ESI): 194.1 $[\text{M-H}]^-$.

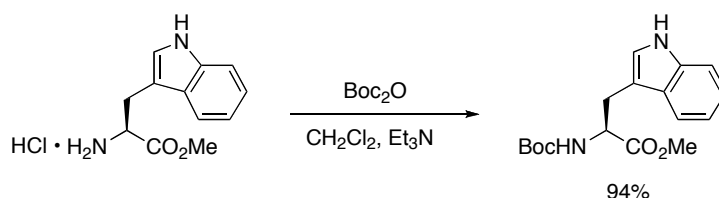
BocTyrOMe



L-Tyrosine methyl ester hydrochloride (3.23 g, 16.6 mmol) was added to a 250 mL round bottom flask and suspended in dioxane (50 mL). To the stirred suspension was added Et_3N (4.65 mL, 33.1 mmol). Boc_2O (4.70 g, 21.5 mmol) was added in one portion and the reaction was stirred at room temperature. After 5 hours of reaction time, the reaction was diluted with EtOAc (150 mL) and washed with H_2O (150 mL). The aqueous layer was extracted with EtOAc (2 $\times$ 150 mL) and the combined organic layers were washed with

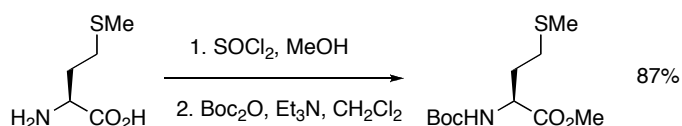
brine (200 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography by eluting first with 15% EtOAc in petrol and then 50% EtOAc in petrol. The product was isolated as a white solid (4.13 g, 84% yield). m.p. = 96-99 °C (Lit.¹² = 102 °C). $[\alpha]_D^{20} = +41.2$ (c = 1.0, CHCl₃) (Lit¹³: +53.1, c = 1.0, CHCl₃), IR ($\nu_{\max}$, KBr): 3370, 2979, 1687, 1516, 1445, 1368, 1224, 1166. ¹H NMR (200 MHz, CDCl₃): δ = 1.42 (9H, s, Boc), 3.00 (2H, ABX system, J = 5.8, 6.0, 13.6, CH₂Ar), 3.71 (3H, s, CO₂CH₃), 4.54 (1H, m, H α), 5.06 (1H, d, J = 4.0, NH), 6.48 (1H, br s, OH), 6.73 (2H, d, J = 4.0, CH_{Ar}), 6.96 (2H, d, J = 4.0, CH_{Ar}). ¹³C NMR (50 MHz, CDCl₃): δ = 28.3 (Boc), 37.6 (CH₂Ar), 52.4 (CO₂CH₃), 54.6 (C α), 80.3 (Boc), 115.5, 127.3, 130.3, 155.3, 155.4, 172.7 (C_{Ar} and C=O). LRMS m/z (ESI⁻): 294.1 [M-H]⁻.

BocTrpOMe



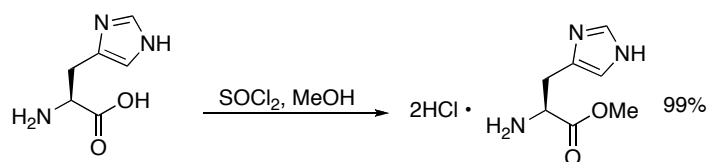
L-Tryptophan methyl ester hydrochloride (1.00 g, 3.93 mmol) was added to a 100 mL round bottom flask and suspended in CH₂Cl₂ (20 mL). Et₃N (1.64 mL, 11.8 mmol) was added and the mixture was stirred until all material was dissolved. Boc₂O (943 mg, 4.32 mmol) was added and the reaction was stirred at room temperature for 4 hours. After this time, the reaction was diluted with CH₂Cl₂ (100 mL) and washed with 1M HCl (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting solid was recrystallized (CH₂Cl₂/petrol) and isolated as small white needles (1.179 g, 94%). m.p. = 143-144 °C (Lit¹⁴: 151-152 °C). $[\alpha]_D^{20} = +55.7$ (c = 1.0, CHCl₃) (Lit⁶: +47.5, c = 0.05, CHCl₃). IR ($\nu_{\max}$, film): 3357, 2978, 1739, 1697, 1503, 1457, 1437, 1367, 1249, 1218, 1167. 1097, 1063. ¹H NMR (CDCl₃, 200MHz): δ = 1.45 (9H, s, Boc), 3.30 (2H, d, J = 5.4, CH₂Ar), 3.68 (3H, s, CO₂Me), 4.62-4.72 (1H, m, H α), 5.13 (1H, d, J = 7.9, NH), 6.97 (1H, s, CH_{Ar}), 7.09-7.24 (2H, m, CH_{Ar}), 7.32-7.37 (1H, m, CH_{Ar}), 7.55-7.59 (1H, m, CH_{Ar}), 8.40 (1H, s, NH). ¹³C NMR (CDCl₃, 50 MHz): δ = 27.9 (CH₂Ar), 28.3 (Boc), 52.2 (CO₂Me), 54.2 (H α), 79.9 (Boc), 109.9 (4^oAr), 111.2, 118.6, 119.5, 122.0, 122.8 (5 × CH_{Ar}), 127.6, 136.1 (2 × 4^oAr), 155.3, 172.8 (2 × C=O). LRMS m/z (ESI⁺): 341.2 [M+Na]⁺.

BocMetOMe



To a flame dried flask under nitrogen was added MeOH (30 mL). The stirred solution was cooled to 0 °C before SOCl₂ (3.70 mL, 50.7 mmol) was added dropwise. The solution was stirred 10 min at 0 °C before methionine (5.04 g, 33.8 mmol) was added in one portion. The reaction was stirred at room temperature overnight after which time the volatiles were removed under reduced pressure to give the crude methionine methyl ester hydrochloride as a yellow-white solid. This solid was dissolved in CH₂Cl₂ (150 mL) and cooled to 0 °C. Et₃N (17.5 mL, 126 mmol) was added carefully followed by Boc₂O (13.6 g, 62.5 mmol). After stirring for 7 hours at room temperature, the reaction was diluted with CH₂Cl₂ (100mL) and washed with H₂O (2 × 100 mL). The organic layer was dried (MgSO₄), filtered, and the solvent removed under reduced pressure. The product (R_f = 0.6 in 20% EtOAc in petrol) was purified by column chromatography (20% EtOAc in petrol) to afford the titled compound as a clear oil (7.75 g, 87% yield). [α]_D²⁰ = -29.5, c = 1.0, MeOH (Lit¹⁵ = -34.0, c = 1.0, MeOH). IR (ν_{max}, film): 3359, 2977, 2920, 1744, 1716, 1515, 1438, 1391, 1366, 1251, 1221, 1167. ¹H NMR (400 MHz CDCl₃): δ = 1.39 (9H, s, Boc), 1.88 (1H, m, CHHCH₂SMe), 2.12-2.04 (4H, includes 3H, s, SMe and CHHCH₂SMe), 2.49 (2H, t, J = 8.0, CH₂SMe), 3.70 (3H, s, CO₂Me), 4.37 (1H, q, J = 7.1, Hα), 5.20 (1H, d, J = 7.1, NH). ¹³C NMR (100 MHz): δ = 15.3 (SCH₃), 28.1 (Boc), 29.8 (CH₂CH₂SMe), 31.9 (CH₂SMe), 52.2 (CO₂CH₃), 52.6 (Cα), 79.8 (Boc), 155.2, 172.7 (2 × C=O). LRMS m/z (ESI⁺): 286.1 [M+Na]⁺.

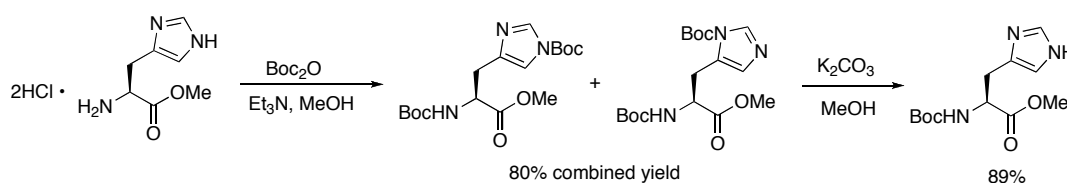
(L)-Histidine methyl ester di-hydrochloride



L-Histidine (20.00 g, 128.9 mmol) was added to a 500 mL round bottom flask and dissolved in MeOH (100 mL, anhydrous). SOCl₂ (11.2 mL, 154.7 mmol) was added carefully by syringe under a balloon of nitrogen. After the addition of thionyl chloride, the flask was equipped with a condenser and heated at reflux for 10 hours. After this time the reaction was cooled to room temperature and the solvent removed under reduced

pressure. The resulting solid was triturated by stirring vigorously in Et₂O (200 mL) for 20 minutes. The product was isolated by filtration as hygroscopic crystals (31.05 g, 99%). Spectroscopic data was consistent with that previously reported.¹⁶ m.p. = 200-203 °C. $[\alpha]_D^{20} = +10.3$ (c = 2.0, H₂O). IR ($\nu_{\max}$, KBr): 2969, 2027, 1996, 1760, 1625, 1336, 1291, 1065. ¹H NMR (400 MHz, D₂O): δ = 3.30-3.43 (2H, ABX system, *J* = 15.9, 7.1, 6.8, H β), 3.73 (3H, s, CO₂Me), 4.41 (1H, t, *J* = 6.8, H α), 7.35 (1H, s, CH_{Ar}), 8.61 (1H, s, CH_{Ar}). ¹³C NMR (100 MHz, D₂O): δ = 25.3 (C β), 52.2 (C α), 54.4 (CO₂Me), 118.7 (CH_{Ar}), 126.5 (4_{Ar}), 134.7 (CH_{Ar}), 169.2 (C=O). LRMS *m/z* (ESI⁺): 170.1 [M+H]⁺.

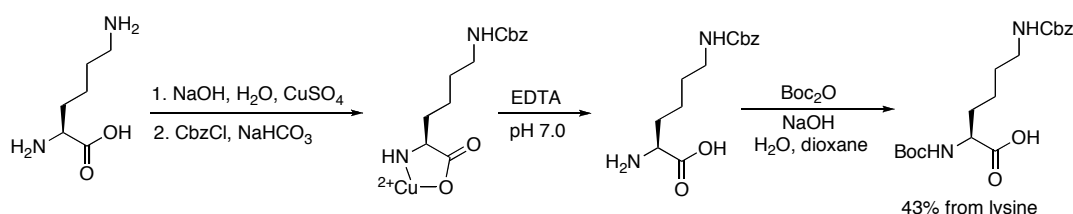
BocHisOMe



The following procedure was adapted from the literature.¹⁶ L-Histidine methyl ester dihydrochloride (1.00 g, 4.13 mmol) was added to a 100 mL round bottom flask and suspended in MeOH (25 mL). Et₃N (2.9 mL, 20.7 mmol) was added slowly to the stirred suspension at room temperature. Boc₂O (2.07 g, 9.48 mmol) was added as a solid in one portion to the reaction flask. The reaction was stirred at room temperature for 2 hours and then diluted with EtOAc (200 mL). The mixture was washed sequentially with H₂O (150 mL) and brine (150 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (60% EtOAc in petrol) to give a mixture of regioisomers (1.24 g, 80% combined yield). Note: these isomers are separable by chromatography (*R_f* = 0.25, *R_f* = 0.20, 60% EtOAc in petrol), but were still combined since they converged to BocHisOMe on deprotection. Accordingly, a portion of the two regioisomers (906 mg, 2.45 mmol) was dissolved in MeOH (25 mL). K₂CO₃ (34 mg, 0.25 mmol) was added and the reaction mixture heated at reflux for 3 hours. TLC (EtOAc) revealed complete consumption of both regioisomers. The reaction was filtered and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% MeOH in CH₂Cl₂) to provide BocHisOMe as a thick clear oil that crystallized upon standing (586 mg, 89%). m.p. = 135-136 °C (Lit¹⁷ = 128-129 °C). $[\alpha]_D^{20} = -4.2$ (c = 2.0, MeOH). IR ($\nu_{\max}$, film): 3300, 2978, 1743, 1710, 1512, 1437, 1392, 1367, 1288, 1252, 1215, 1167, 1055, 1024. ¹H NMR (400 MHz, MeOD): δ = 1.41 (9H, s, Boc), 2.92-3.09 (2H, ABX system, *J* = 14.9,

8.6, 5.3, H β), 3.70 (3H, s, CO₂Me), 4.40 (1H, dd, J = 8.6, 5.3, H α), 6.87 (1H, s, CH_{Ar}), 7.60 (1H, s, CH_{Ar}). ¹³C NMR (100 MHz, MeOD): δ = 28.8 (Boc), 30.4 (C β), 52.8 (CO₂Me), 55.4 (C α), 80.8 (Boc), 118.3 (CH_{Ar}), 134.7 (4 $^{\circ}$ _{Ar}), 136.4 (CH_{Ar}), 157.9, 174.3 (2 $\times$ C=O). LRMS m/z (ESI⁺): 270.1 [M+H]⁺.

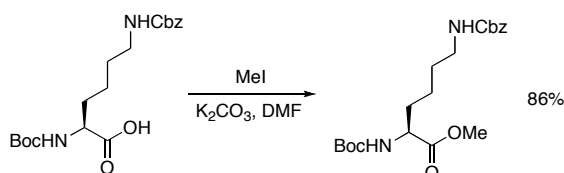
BocLys(Cbz)OH



This preparation was adapted from the literature.^{18,19} L-Lysine (40.00 g, 219 mmol) was added to a 1 L flask and dissolved in H₂O (160 mL). The stirred solution was cooled to 0 °C and NaOH (17.52 g, 438.0 mmol) was added as a solid. The reaction was stirred until all material was dissolved. CuSO₄ (17.48 g, 109.5 mmol) was prepared as a solution in H₂O (80 mL) in a separate flask and then poured into the solution of lysine. The resulting blue solution was cooled to 0 °C and then NaHCO₃ (22.08 g, 262.8 mmol) was added. When all material was dissolved, benzylchloroformate (40.5 mL, 284.7 mmol) was poured into the reaction mixture at 0 °C. The reaction was stoppered and stirred vigorously at 0 °C for 1 hour and then at room temperature for 5 hours. After this time, the blue precipitate was isolated by filtration and washed with H₂O (200 mL). The solid was partially dried on the filter before washing with acetone (200 mL) to remove any excess CbzCl. The Lys(Cbz) copper complex was dried to a blue powder on the filter. In a 2 L beaker, a suspension of EDTA (70.4 g, 241 mmol) in H₂O (800 mL) was prepared and heated to 80 °C. The Lys(Cbz) copper complex was added as a solid to the stirred solution of EDTA. A color change from lavender to bright aqua blue was observed as the EDTA copper complex formed. After all of the lysine derivative was added, the pH was adjusted to 7 by adding ~ 100 mL of 5M NaOH (the pH was checked using pH paper). At pH 7, a white precipitate was formed (Lys(Cbz)OH). The hot suspension was filtered and the solid isolated. The solid was then washed with H₂O (2 L) and then MeOH (500 mL). The white solid was dried on the filter and then under high vacuum. The crude material was used without purification in the next step. Accordingly, Lys(Cbz)OH (15.00 g, 53.51 mmol crude from above) was suspended in H₂O (50 mL). NaOH (4.49 g, 112 mmol) was added to the stirred suspension and the mixture was stirred at room temperature to

dissolve. Once all material had dissolved, Boc₂O (11.68 g, 53.51 mmol) was added as a solution in 1,4-dioxane (50 mL). The reaction mixture was stirred vigorously at room temperature. Solid was formed during the reaction. After 4 hours, the reaction mixture was diluted with EtOAc (100 mL) and 1M HCl (100 mL). This mixture was stirred vigorously at room temperature to dissolve all material. The mixture was further diluted with EtOAc (100 mL) and 1M HCl (100 mL). The organic layer was separated, dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% MeOH in CH₂Cl₂), to give BocLys(Cbz)OH as a thick, clear oil (8.73 g, 43% from lysine). $[\alpha]_D^{20} = -5.6$ (*c* = 3.3, MeOH). IR ($\nu_{\max}$, film): 3332, 2977, 2936, 1707, 1529, 1455, 1409, 1393, 1367, 1253, 1165, 1048, 1025. ¹H NMR (400 MHz, DMSO): δ = 1.26-1.37 (13H, m, contains Boc, H γ , H δ), 1.53-1.65 (2H, m, H β), 2.98 (2H, q, *J* = 5.5, CH₂NHCbz), 3.80-3.85 (1H, m, H α), 5.00 (2H, s, CH₂Ph), 6.90 (1H, d, *J* = 5.6, NH_{Boc}), 7.23 (1H, t, *J* = 5.5, NHCbz), 7.28-7.38 (5H, m, CH_{Ar}). ¹³C NMR (100 MHz, DMSO): δ = 22.9, 28.3, 29.1 (C γ , Boc, C δ), 30.7 (C β), 40.1 (CH₂NHCbz), 53.7 (C α), 65.2 (CH₂Ph), 77.9 (Boc), 127.8 (2 × C), 128.4 (3 × CH_{Ar}), 137.3 (4 $^{\circ}$ _{Ar}), 155.6, 156.1, 174.6 (3 × C=O). LRMS *m/z* (ESI⁺): 403.2 [M+Na]⁺.

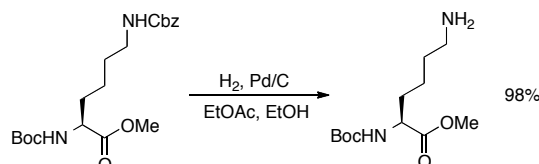
BocLys(Cbz)OMe



BocLys(Cbz)OH (7.23 g, 19.0 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (40 mL). K₂CO₃ (5.25 g, 38.0 mmol) was added and the stirred solution was cooled to 0 °C. Iodomethane (1.05 mL, 24.7 mmol) was added and the reaction was stirred at 0 °C for 5 minutes and then room temperature for 2 hours. After this time, the reaction was diluted with EtOAc (250 mL) and washed with H₂O (300 mL) and brine (2 × 300 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (45% EtOAc in petrol) to give BocLys(Cbz)OMe as a thick, clear oil (6.41 g, 86%). Spectroscopic data was consistent with that previously reported.¹⁹ $[\alpha]_D^{20} = -22.5$ (*c* = 1.0, MeOH). IR ($\nu_{\max}$, film): 3343, 2951, 2866, 1714, 1526, 1455, 1366, 1251, 1167, 1047, 1024. ¹H NMR (400 MHz, CDCl₃): δ = 1.23-1.48 (13H, m, includes Boc, H γ , H δ), 1.53-1.78 (2H, m, H β), 3.07-3.11 (2H, q, *J* = 6.3, CH₂NHCbz), 3.63 (3H, s, CO₂Me), 4.16-

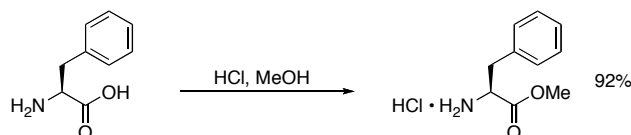
4.23 (1H, m, H α). 5.02 (2H, s, CH₂Ph), 5.26-5.35 (2H, m, NH₂Boc and NH₂Cbz), 7.20-7.30 (5H, m, CH_{Ar}). ¹³C NMR (100 MHz, CDCl₃): δ = 22.1, 28.0, 29.0 (C γ , Boc, C δ), 31.7 (C β), 40.2 (CH₂NHCbz), 51.9 (CO₂Me), 53.0 (C α), 66.1 (CH₂Ph), 79.4 (Boc), 127.68, 127.74, 128.1 (3 $\times$ CH_{Ar}), 136.4 (4 $^{\circ}$ _{Ar}), 155.3, 156.3, 173.3 (3 $\times$ C=O). LRMS m/z (ESI⁺): 417.2 [M+Na]⁺.

BocLysOMe



BocLys(Cbz)OMe (6.13 g, 15.5 mmol) was dissolved in a mixture of EtOAc (80 mL) and EtOH (20 mL). Palladium (5% wt. on carbon, 826 mg, 0.39 mmol) was added to the stirred solution. The flask was sealed and hydrogen was bubbled through the solution for 5 minutes. After 1 hour of stirring under an H₂ atmosphere (balloon), TLC (50% EtOAc in petrol) revealed complete consumption of starting material. The palladium and carbon was removed by filtration through a pad of celite. The solids were washed with an additional 100 mL of MeOH. The filtrate was concentrated under reduced pressure and the resulting residue was purified by column chromatography (10% MeOH in CH₂Cl₂ with 2% Et₃N). BocLysOMe was isolated as a thick, clear oil¹⁹ (3.95 g, 98%). [α]_D²⁰ = -29.0 (c = 1.0, MeOH). IR ($\nu_{\max}$, film): 3365, 2933, 1744, 1713, 1521, 1366, 1167, 1050, 1021. ¹H NMR (400 MHz, CDCl₃): δ = 1.27-1.38 (15H, m, contains Boc, H γ , H δ , and NH₂), 1.46-1.76 (2H, m, H β), 2.59 (2H, t, J = 6.7, CH₂NH₂), 3.63 (3H, s, CO₂Me), 4.14-4.22 (1H, m, H α), 5.29 (1H, d, J = 7.8, NH₂Boc). ¹³C NMR (100 MHz, CDCl₃): δ = 22.4, 28.1, 32.2, 32.8 (C γ , Boc, C δ , C β), 41.5 (CH₂NH₂), 51.9 (CO₂Me), 53.2 (C α), 79.5 (Boc), 155.3, 173.2 (2 $\times$ C=O). LRMS m/z (ESI⁺): 261.2 [M+H]⁺.

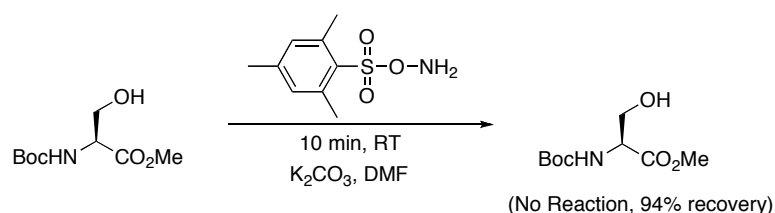
(L)-Phenylalanine methyl ester hydrochloride



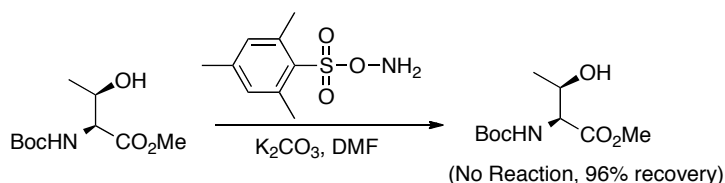
MeOH (20 mL) was added to a flame dried 100 mL round bottom flask under an atmosphere of argon. The flask was cooled to 0 °C and acetyl chloride (4.30 mL, 60.5 mmol) was added carefully. The reaction was stirred at 0 °C for 10 minutes to generate

methanolic HCl. L-Phenylalanine (2.00 g, 12.1 mmol) was then added as a solid in one portion. The ice bath was removed and the reaction was stirred at room temperature for 12 hours. After this time, the solvent was removed under reduced pressure to give HCl•PheOMe as a white solid that did not require purification (2.41 g, 92%). m.p. = 150-151 °C (Lit²⁰ = 158-159 °C). $[\alpha]^{20}_D = +32.5$ (c = 2.0, EtOH) (Lit²⁰: +36.2, c = 2, EtOH). IR ($\nu_{\max}$, KBr): 2843, 2011, 1748, 1584, 1496, 1447, 1292, 1242, 1084, 1060, 1032. ¹H NMR (400 MHz, D₂O): δ = 3.15 (2H, ABX system, J = 14.4, 7.6, 5.8, CH₂Ph), 3.70 (3H, s, CO₂Me), 4.29 (1H, dd, J = 7.3, 6.1, H α), 7.14-7.31 (5H, m, ArH). ¹³C NMR (100 MHz, D₂O): δ = 35.9 (CH₂Ph), 53.9, 54.4 (CO₂Me, C α), 128.5, 129.6, 129.7 (CH_{Ar}), 134.0 (4 $^{\circ}$ Ar), 170.4 (C=O). LRMS m/z (ESI⁺): 180.1 [M+H]⁺.

MSH Reactivity with Functionalized Amino Acids:

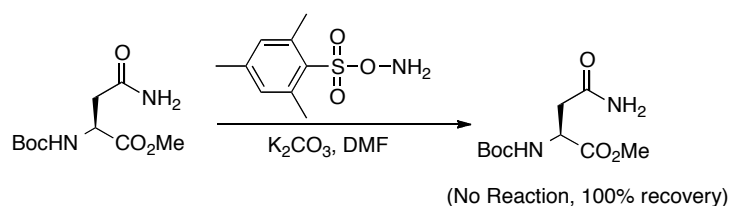


BocSerOMe (50 mg, 0.23 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (5 mL). K₂CO₃ (158 mg, 1.14 mmol) was added to the stirred solution, followed by MSH (124 mg, 0.58 mmol). The reaction was stirred at room temperature for 10 minutes and then diluted with EtOAc (100 mL). The solution was washed with H₂O (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc) to provide unreacted BocSerOMe (47 mg, 94% recovery). Spectroscopic and analytical data was identical to that reported above for BocSerOMe.

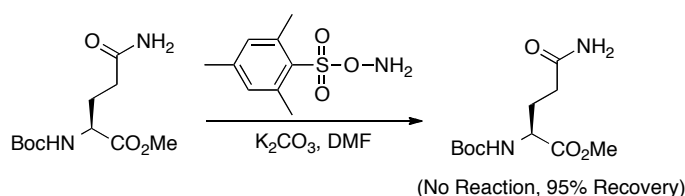


(L)-BocThrOMe (50 mg, 0.21 mmol) was added to a 10 mL round bottom flask and dissolved in DMF (4 mL). K₂CO₃ (148 mg, 1.07 mmol) was added, followed by MSH

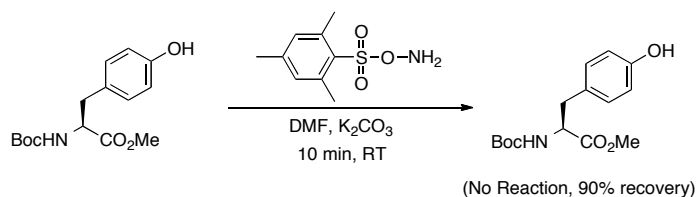
(113 mg, 0.53 mmol). The reaction was stirred at room temperature for 10 minutes. After this time, the mixture was diluted with EtOAc (60 mL) and washed sequentially with H₂O (60 mL) and brine (60 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography to afford 48 mg of unreacted BocThrOMe (96% recovery). Spectroscopic data was identical to that reported above for BocThrOMe.



(L)-BocAsnOMe (20 mg, 0.081 mmol) was added to a 5 mL round bottom flask and dissolved in DMF (0.5 mL). K₂CO₃ (56 mg, 0.41 mmol) was added, followed by MSH (44 mg, 0.20 mmol). The reaction was stirred 10 minutes at room temperature and then purified directly by column chromatography (EtOAc). Unreacted (L)-BocAsnOMe was recovered in quantitative yield (20 mg, 100% recovery). Spectroscopic data was identical to that reported above.

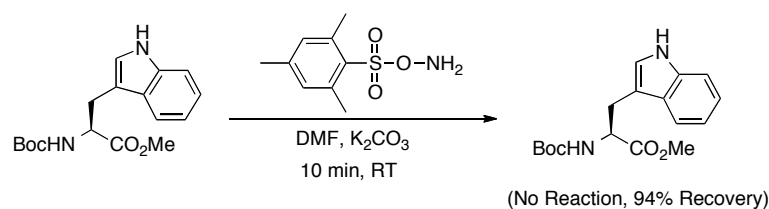


(L)-BocGlnOMe (20 mg, 0.077 mmol), was added to a 5 mL round bottom flask and dissolved in DMF (1 mL). K₂CO₃ (53 mg, 0.38 mmol) was added followed by MSH (42 mg, 0.19 mmol). The reaction was stirred for 10 minutes at room temperature and then purified directly by column chromatography (10% MeOH in CH₂Cl₂). Unreacted (L)-BocGlnOMe was recovered (19 mg, 95% recovery). Spectroscopic data was identical to that reported above.

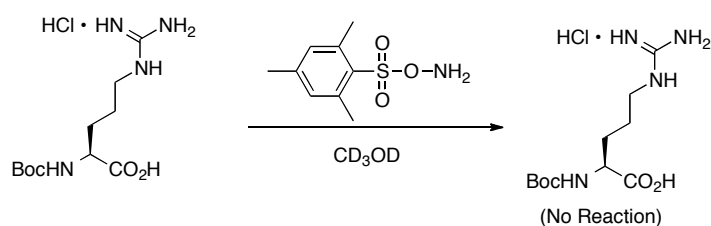


BocTyrOMe (100 mg, 0.34 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (5 mL). K₂CO₃ (234 mg, 1.69 mmol) was added, followed by MSH

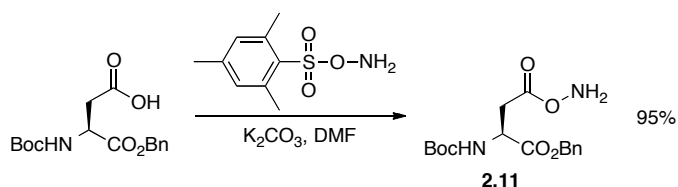
(183 mg, 0.85 mmol). The reaction was stirred at room temperature for 10 minutes and then diluted with EtOAc (100 mL). The solution was washed with H₂O (100 mL), brine (100 mL), and then dried over MgSO₄. After filtration, the solvent was concentrated under reduced pressure. The resulting residue was purified by column chromatography (50% EtOAc in petrol) to afford 90 mg of unreacted BocTyrOMe (90% recovery). Spectroscopic and analytical data were identical to that reported above for the starting material.



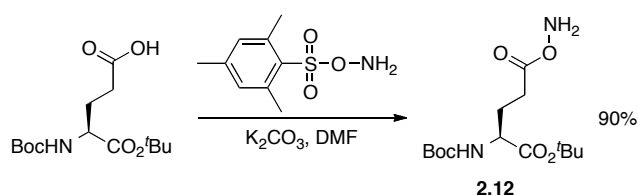
BocTrpOMe (50 mg, 0.16 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (5 mL). K₂CO₃ (109 mg, 0.79 mmol) was added to the stirred solution, followed by MSH (84 mg, 0.39 mmol). The reaction was stirred at room temperature for 10 minutes and then diluted with EtOAc (100 mL). The solution was washed with H₂O (100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (60% EtOAc in petrol), providing unreacted BocTrpOMe (47 mg, 94% recovery). Spectroscopic and analytical data was identical to that reported above for BocTrpOMe.



BocArgOH hydrochloride (5.0 mg, 0.016 mmol) was added to a glass vial and dissolved in 300 μ L CD₃OD. MSH (9.0 mg, 0.042 mmol) was dissolved in 300 μ L CD₃OD and then added to the solution of BocArgOH. The reaction was incubated at room temperature for 10 minutes and then analyzed directly by ¹H NMR and ESI-MS. No reaction was observed and only unreacted BocArgOH was detected in both NMR and ESI-MS.



(L)-BocAspOBn (Fluka, 20 mg, 0.062 mmol) was added to a 5 mL round bottom flask and dissolved in DMF (1 mL). K₂CO₃ (51 mg, 0.37 mmol) and MSH (33 mg, 0.16 mmol) were added sequentially to the stirred solution. The reaction was stirred for 10 minutes at room temperature and then purified directly by column chromatography (10% MeOH in CH₂Cl₂). A single product was isolated (*R*_f = 0.7, 10% MeOH in CH₂Cl₂) and determined to be the product of carboxyamination (**2.11**) (clear oil, 20 mg, 95%). [α]²⁰_D = -20.9 (*c* = 1.0, MeOH). IR (ν_{max} , film): 3314, 2978, 2934, 2361, 1744, 1713, 1500, 1367, 1342, 1289, 1247, 1215, 1163. ¹H NMR (500 MHz, CDCl₃): δ = 1.43 (9H, s, Boc), 2.95 (2H, ABX system, *J* = 16.7, 4.7, H β), 4.65 (1H, m, H α), 5.20 (2H, AB quart., *J* = 12.3, CH₂Ph), 5.49 (1H, d, *J* = 7.9, NHBoc), 6.28 (2H, br. s, NH₂), 7.34-7.39 (5H, m, ArH). ¹³C NMR (125 MHz, CDCl₃): δ = 28.2 (Boc), 35.2 (C β), 50.0 (C α), 67.6 (CH₂Ph), 80.4 (Boc), 128.4, 128.5, 128.6 (CH_{Ar}), 135.1 (4^o_{Ar}), 155.2, 170.5, 171.5 (3 $\times$ C=O). HRMS *m/z* (ESI⁺): Found 361.1369 [M+Na]⁺; C₁₆H₂₂N₂O₆Na requires 361.1370.



(L)-BocGluO'Bu (Fluka, 20 mg, 0.066 mmol) was added to a 5 mL round bottom flask and dissolved in DMF (0.5 mL). K₂CO₃ (55 mg, 0.40 mmol) was added to the stirred solution, followed by MSH (36 mg, 0.17 mmol). The reaction was stirred 10 minutes at room temperature and then purified directly by column chromatography (EtOAc). A single product was isolated (*R*_f = 0.6, EtOAc) and identified as the product of carboxyamination **2.12** (clear oil, 19 mg, 90%). [α]²⁰_D = -22.5 (*c* = 1.0, MeOH). IR (ν_{max} , film): 3313, 2978, 2934, 2360, 2342, 1734, 1714, 1558, 1508, 1455, 1392, 1367, 1250, 1154. ¹H NMR (500 MHz, CDCl₃): δ = 1.44 (9H, s), 1.47 (9H, s) (2 $\times$ ^{*t*}Bu), 1.91-1.97 (1H, m H β), 2.15-2.24 (1H, m, H β '), 2.37-2.49 (2H, m, H γ), 4.20-4.25 (1H, m, H α), 5.10 (1H, d, *J* = 6.9, NHBoc), 6.36 (2H, br. s, NH₂). ¹³C NMR (125 MHz, CDCl₃): δ = 28.0 (^{*t*}Bu), 28.1 (C β), 28.3 (^{*t*}Bu), 28.5 (C γ), 53.1 (C α), 79.9, 82.4 (2 $\times$ ^{*t*}Bu), 155.4, 171.1,

173.6 ($3 \times \text{C}=\text{O}$). HRMS m/z (ESI⁺): Found 341.1680 [M+Na]⁺; C₁₄H₂₆N₂O₆Na requires 341.1683.

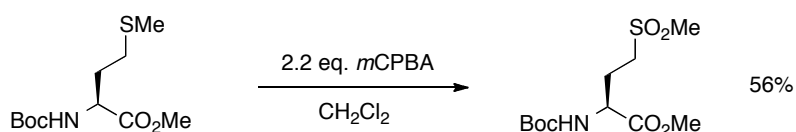


Carboxyaminated BocAspOBn (**2.11**) (20 mg, 0.059 mmol) was added to a 5 mL round bottom flask and dissolved in DMF (1 mL). K₂CO₃ (33 mg, 0.24 mmol) and dithiothreitol (DTT, 18 mg, 0.12 mmol) were added to the stirred solution. After 5 minutes, TLC and ESI-MS indicated that all aminated product had been reduced to BocAspOBn. After 30 minutes of total reaction time, the mixture was purified directly by column chromatography (EtOAc). BocAspOBn (deaminated) was isolated in quantitative yield (19 mg, 100%). Spectroscopic data was identical to the commercial material (Fluka).

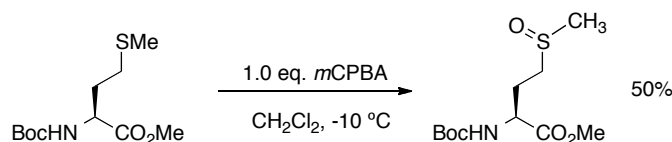
Summary of Reaction of MSH with Methionine and Attempted Reduction Back to Thioether:

Entry	Reductant	Conditions	Yield BocMetOMe
1	 (DTT)	Na ₂ HPO ₄ , DMF, H ₂ O, 2h, RT	94%
2	none (control)	Na ₂ HPO ₄ , DMF, H ₂ O, 2h, RT	20%
3	P(NMe ₂) ₃	DMF, H ₂ O, 3h, RT	0%
4	PBu ₃	DMF, H ₂ O, 3h, RT	0%
5	DTT (5 mol%), PBu ₃	Na ₂ HPO ₄ , DMF, H ₂ O, 3h, RT	17%
6	Zn powder	DMF, H ₂ O, AcOH, 8h, RT	0%
7	Na ₂ S ₂ O ₆	Na ₂ CO ₃ , DMF, H ₂ O, 2h, RT	11%
8	sodium ascorbate	DMF, H ₂ O, 3h, RT	0%

BocMetOMe sulfoxide and sulfone (synthesized for TLC reference during the reaction of BocMetOMe with MSH):

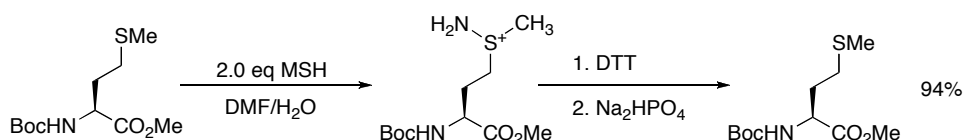


BocMetOMe (876 mg, 3.33 mmol) was added to a 100 mL round bottom flask and dissolved in CH_2Cl_2 (15 mL). The stirred solution was cooled to 0 °C and 1.80 g *m*CPBA (70%, 7.32 mmol) was added as a suspension in 10 mL of CH_2Cl_2 by pipette over a period of 5 minutes (open air). An additional 3 mL of CH_2Cl_2 was used to complete the transfer of oxidant. The reaction was stirred for 1.5 hours at room temperature at which time it was diluted with Et_2O (200 mL) and NaHCO_3 (sat., 100 mL). The layers were separated and the organic fraction was washed successively with NaHCO_3 (2×100 mL) and brine (150 mL). The organics were dried (MgSO_4), filtered, and concentrated under reduced pressure. The product was purified by column chromatography (60% EtOAc in petrol) to afford the sulfone of BocMetOMe ($R_f = 0.33$) as a thick, clear oil (547 mg, 56%). $[\alpha]_D^{20} = +17.8$ ($c = 1.8$, CHCl_3). IR (ν_{max} , film): 3357, 3000, 2380, 1712, 1520, 1368, 1299, 1165. ^1H NMR (400 MHz, CDCl_3): $\delta = 1.41$ (9H, s, Boc), 2.13 (1H, m, $\text{CHHCH}_2\text{SO}_2\text{Me}$), 2.38 (1H, m, $\text{CHHCH}_2\text{SO}_2\text{Me}$), 2.91 (3H, s, SO_2Me), 3.18-3.03 (2H, m, $\text{CH}_2\text{SO}_2\text{Me}$), 3.74 (3H, s, CO_2CH_3), 4.38 (1H, br. s, $\text{H}\alpha$), 5.38 (1H, d, $J = 7.6$, NH). ^{13}C NMR (100 MHz): $\delta = 25.6$ ($\text{CH}_2\text{CH}_2\text{SO}_2\text{Me}$), 28.2 (Boc), 40.7 (SO_2Me), 51.0 ($\text{CH}_2\text{SO}_2\text{Me}$), 52.0 ($\text{C}\alpha$), 52.8 (CO_2CH_3), 80.5 (Boc), 155.4, 171.7 ($2 \times \text{C=O}$). HRMS m/z (ESI $^+$): Found 318.0982 ($\text{M}+\text{Na}$) $^+$; $\text{C}_{11}\text{H}_{21}\text{NO}_6\text{SNa}$ requires 318.0987.

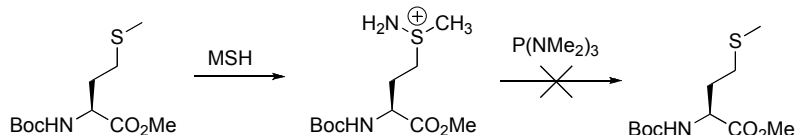


BocMetOMe (1.52 g, 5.76 mmol) was added to a 100 mL round bottom flask and dissolved in CH_2Cl_2 (30 mL). The stirred solution was cooled to -10 °C (NaCl/ice bath) and *m*CPBA (1.36 g, 5.76 mmol) was added by pipette as a suspension in CH_2Cl_2 (10 mL) over a period of 10 min. The temperature was maintained between -5 °C and -10 °C over 1 hr. TLC (EtOAc) revealed consumption of the starting material and formation of the sulfoxide ($R_f = 0.2$). The reaction was then diluted with EtOAc (250 mL) and NaHCO_3 (sat., 200 mL). After separation, the organics were washed successively with

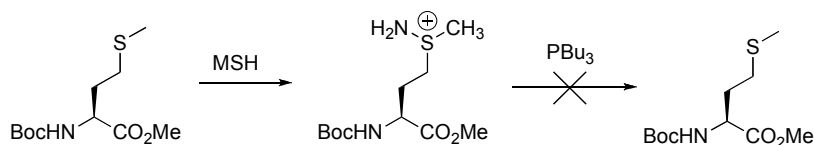
NaHCO₃ (sat., 200 mL) and brine (200 mL) and then dried over MgSO₄. The solution was filtered and the solvent removed under reduced pressure. The residue was purified by column chromatography (EtOAc to 10% MeOH in EtOAc) to give the sulfoxide of BocMetOMe as a thick, clear oil and as a mixture of diastereomers (806 mg, 50%). $[\alpha]_D^{20} = +17.1$ ($c = 0.60$, CHCl₃). IR ($\nu_{\max}$, film): 1699, 1165, 1017. ¹H NMR (CDCl₃, 400 MHz): $\delta = 1.25$ (9H, s, Boc), 1.94 (1H, m, CHHCH₂S(O)Me), 2.13 (1H, m, CHHCH₂S(O)Me), 2.42 (3H, s, S(O)Me), 2.70-2.55 (2H, m, CH₂S(O)Me), 3.57 (3H, s, CO₂CH₃), 4.24-4.16 (1H, m, H α), 5.78 (1H, app. t, $J = 9.4$, NH). ¹³C NMR (100 MHz, data reported for both diastereomers where peaks resolved): $\delta = 24.9$, 25.2 (C β), 27.8 (Boc), 38.0 (S(O)CH₃), 49.8, 50.0 (CH₂S(O)Me), 52.0, 52.1 (CO₂CH₃), 52.3 (C α), 79.5 (Boc), 155.2, 155.3 (C=O) 171.7, 171.8 (C=O). HRMS m/z (EI+): Found 302.1033 (M+Na)⁺; C₁₁H₂₁NO₅SNa requires 302.1038.



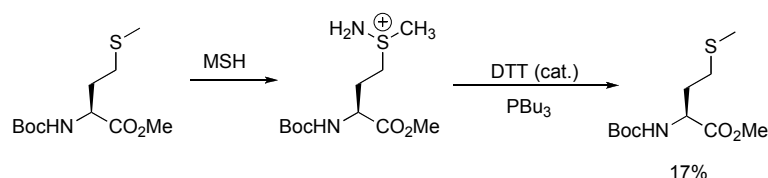
BocMetOMe (245 mg, 0.93 mmol) was added to a 50 mL round bottom flask and then dissolved in DMF (5 mL). The solution was stirred vigorously while H₂O (5 mL) was added by pipette. MSH (400 mg, 1.86 mmol) was added to the solution in one portion and the cloudy suspension homogenized after 30 seconds of stirring. After 5 minutes, TLC analysis (30% EtOAc in petrol) revealed complete consumption of BocMetOMe. All material was located on the baseline. After 20 minutes of stirring, DTT (1.43 g, 1.86 mmol) was added as a solid. TLC analysis revealed no change after 1 hour of stirring. After 1 hour of total reaction time, NaHPO₄•12H₂O (3.33 g, 9.30 mmol) was added to give a saturated solution of phosphate salts. After 2 hours of total reaction time (1 hour with base), TLC (30% EtOAc in petrol) revealed the regeneration of BocMetOMe. A final hour of reaction time revealed no further change. The reaction was then diluted with Et₂O (150 mL) and H₂O (150 mL) and separated. The organic layer was washed sequentially with H₂O (150 mL) and brine (150 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The product was purified by column chromatography (30% EtOAc in petrol) to give recovered BocMetOMe (230 mg, 94%). A similar experiment without DTT returned only 20% of BocMetOMe.



BocMetOMe (100 mg, 0.38 mmol) was added to a 25 mL round bottom flask and dissolved in a mixture of DMF (5 mL) and H_2O (2 mL). MSH (123 mg, 0.57 mmol) was added at room temperature and the reaction stirred for 5 minutes. TLC revealed complete consumption of starting material. HMPT (0.14 mL, 0.76 mmol) was added and the reaction monitored periodically over the course of three hours. No regeneration of starting material was observed.

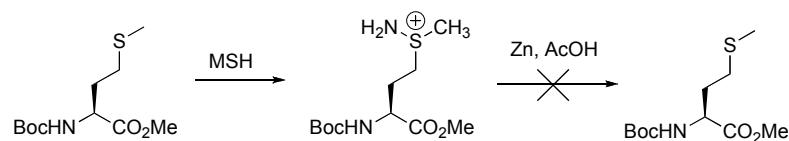


BocMetOMe (100 mg, 0.38 mmol) was added to a 25 mL round bottom flask and dissolved in a mixture of DMF (5 mL) and H_2O (2 mL). MSH (123 mg, 0.57 mmol) was added at room temperature and the reaction stirred for 5 minutes. TLC revealed complete consumption of starting material. PBu_3 (0.23 mL, 0.95 mmol) was added and the reaction stirred 10 minutes at room temperature. An additional 0.5 mL of PBu_3 was added and the reaction stirred for 3 hours. No BocMetOMe was regenerated.

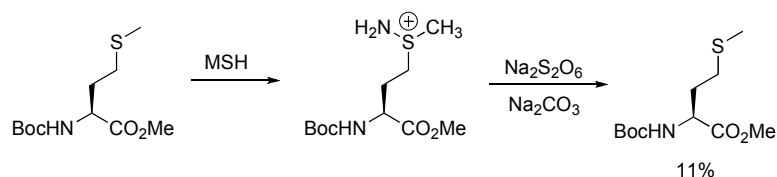


BocMetOMe (100 mg, 0.38 mmol) was added to a 25 mL round bottom flask and dissolved in a mixture of DMF (5 mL) and H_2O (2 mL). MSH (123 mg, 0.57 mmol) was added at room temperature and the reaction stirred for 5 minutes. TLC revealed complete consumption of starting material. Na_2CO_3 (201 mg, 1.9 mmol), PBu_3 (0.19 mL, 0.76 mmol), and DTT (3 mg, 0.02 mmol) were added sequentially. After 10 minutes, an additional 3 mg of DTT (0.02 mmol) was added. After 3 hours at room temperature, the reaction was diluted with Et_2O (150 mL) and then washed with H_2O (100 mL) and brine (100 mL). The organic layer was then dried, filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (35% EtOAc in

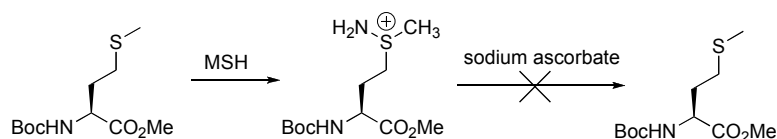
petrol) to give 17 mg (17%) of recovered BocMetOMe. Spectral data was identical to the starting material.



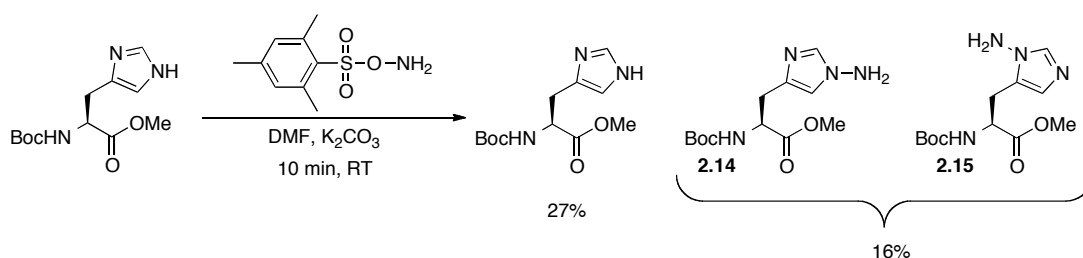
BocMetOMe (100 mg, 0.38 mmol) was added to a 25 mL round bottom flask and dissolved in a mixture of DMF (5 mL) and H₂O (2 mL). MSH (123 mg, 0.57 mmol) was added at room temperature and the reaction stirred for 5 minutes. TLC revealed complete consumption of starting material. AcOH (2.0 mL) and zinc powder (100 mg) were then added sequentially and the reaction stirred vigorously at room temperature. No BocMetOMe was detected after an hour so an additional 500 mg of zinc powder was added. No BocMetOMe was regenerated after 8 hours of reaction.



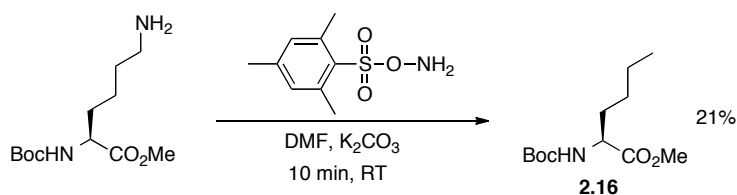
BocMetOMe (100 mg, 0.38 mmol) was added to a 25 mL round bottom flask and dissolved in a mixture of DMF (5 mL) and H₂O (2 mL). MSH (123 mg, 0.57 mmol) was added at room temperature and the reaction stirred for 5 minutes. TLC revealed complete consumption of starting material. Na₂CO₃ (121 mg, 1.14 mmol) was added followed by sodium dithionite (330 mg, 1.9 mmol). After 2 hours at room temperature, some starting material was regenerated. The reaction was diluted with Et₂O (150 mL) and H₂O (100 mL). The organic layer was separated and then washed with H₂O (2 × 100 mL) and brine (100 mL). The organic layer was dried, filtered, and then concentrated under reduced pressure. The resulting residue was purified by column chromatography (35% EtOAc in petrol) to give 11 mg of BocMetOMe (11%). Spectral data was identical to the starting material.



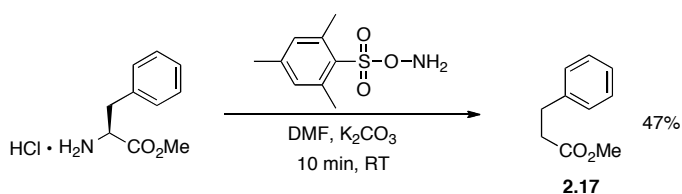
BocMetOMe (100 mg, 0.38 mmol) was added to a 25 mL round bottom flask and dissolved in a mixture of DMF (5 mL) and H₂O (2 mL). MSH (123 mg, 0.57 mmol) was added at room temperature and the reaction stirred for 5 minutes. TLC revealed complete consumption of starting material. Sodium ascorbate (376 mg, 1.9 mmol) was added and the reaction followed by periodically by TLC. After three hours no BocMetOMe was regenerated.



BocHisOMe (500 mg, 1.86 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (8 mL). K₂CO₃ (1.29 g, 9.30 mmol) was added, followed by MSH (450 mg, 2.09 mmol). The reaction was stirred at room temperature for 10 minutes and then diluted with EtOAc (200 mL). The solution was washed with H₂O (200 mL) and brine (200 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. TLC (10% MeOH in EtOAc) revealed 3 products. Careful separation by column chromatography (10% MeOH in EtOAc) provided unreacted BocHisOMe (134 mg, 27%) and a mixture of regioisomers of the amination products **2.14** and **2.15** (87 mg). One of the regioisomers was isolated by a second purification by column chromatography (10% MeOH in EtOAc). Data for the isolated regioisomer: $[\alpha]_D^{20} = -1.0$ (*c* = 0.5, MeOH). IR (ν_{max} , film): 3345, 3221, 2978, 1741, 1699, 1519, 1438, 1392, 1367, 1289, 1253, 1219, 1166. ¹H NMR (400 MHz, CDCl₃): δ = 1.39 (9H, s, Boc), 3.02 (1H, dd, *J* = 15.0, 6.7, CHHAr), 3.19 (1H, dd, *J* = 15.0, 5.3, CHHAr), 3.75 (3H, s, CO₂Me), 4.62 (1H, m, H α), 4.85 (2H, s, NH₂), 5.26 (1H, app. s, NHBoc), 6.72 (1H, s, CHAr), 7.49 (1H, s, CHAr). ¹³C NMR (125 MHz, CDCl₃): δ = 26.5 (CH₂Ar), 28.2 (Boc), 52.6 (CO₂Me), 57.7 (C α), 80.4 (Boc), 126.6 (CHAr), 127.0 (4^oAr), 138.1 (CHAr), 135.2, 172.0 (2 $\times$ C=O). HRMS *m/z* (ESI⁺): Found 285.1556 [M+H]⁺; C₁₂H₂₁N₄O₄ requires 285.1557.



BocLysOMe (229 mg, 0.88 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (10 mL). K_2CO_3 (608 mg, 4.40 mmol) was added, followed by MSH (473 mg, 2.2 mmol). The reaction was stirred 10 minutes at room temperature and then diluted with EtOAc (150 mL) and $NaHCO_3$ (sat. aq., 100 mL). The organic layer was separated, dried ($MgSO_4$), filtered, and concentrated under reduced pressure. TLC of the product mixture revealed that all BocLysOMe was consumed in the reaction and that a complicated product mixture had formed. The major product by TLC was isolated after careful purification by column chromatography (30% EtOAc in petrol). This material was assigned as Boc-norleucine methyl ester (**2.16**). (45 mg, 21%). When the reaction was repeated and stirred for 1 hour, a similar yield (20%) was obtained. $[\alpha]_D^{20} = -22.7$ ($c = 0.93$, MeOH). IR (ν_{max} , film): 2959, 2944, 2887, 2438, 2348, 1735, 1716, 1516, 1475, 1366, 1164. 1H NMR (500 MHz, $CDCl_3$): $\delta = 0.89$ (3H, t, $J = 7.1$, CH_2CH_3), 1.27-1.36 (4H, m, $CH_2CH_2CH_3$), 1.44 (9H, s, Boc), 1.56-1.64 (1H, m, H_β), 1.74-1.82 (1H, m, $H_{\beta'}$), 3.73 (3H, s, CO_2Me), 4.28 (1H, m, H_α), 4.99 (1H, d, $J = 6.6$, NH). ^{13}C NMR (125 MHz, $CDCl_3$): $\delta = 13.8$ (CH_2CH_3), 22.3, 27.4 ($CH_2CH_2CH_3$), 28.3 (Boc), 32.5 (C_β), 52.2 (CO_2Me), 53.4 (C_α), 79.8 (Boc), 155.4, 173.5 (2 $\times$ C=O). HRMS m/z (ESI $^+$): Found 268.1519 $[M+Na]^+$; $C_{12}H_{23}NO_4Na$ requires 268.1519.

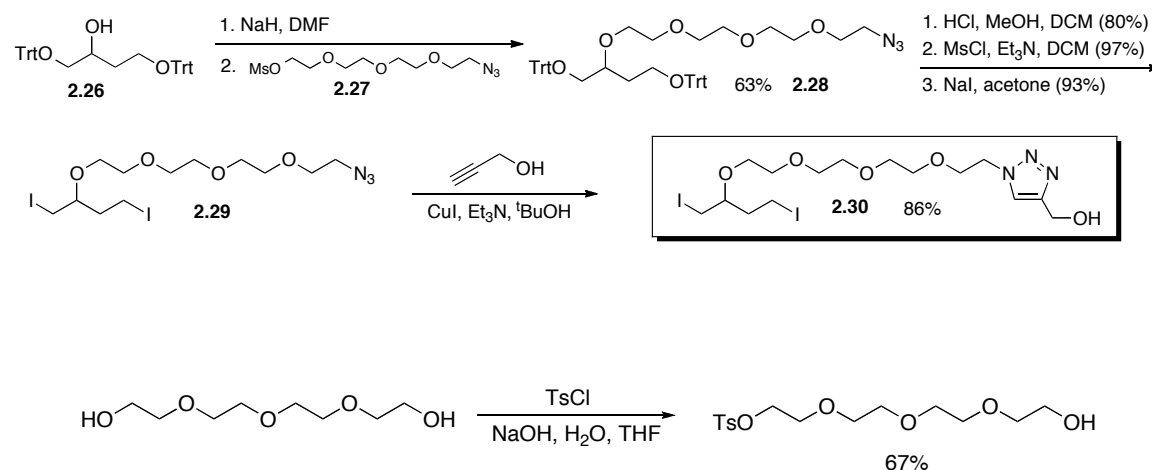


Phenylalanine methyl ester hydrochloride (100 mg, 0.46 mmol) was added to a 25 mL round bottom flask and suspended in DMF (5 mL). K_2CO_3 (384 mg, 2.78 mmol) was added and the mixture was stirred 5 minutes at room temperature before MSH (247 mg, 1.15 mmol) was added as a solid. The reaction was stirred for 10 minutes at room temperature and then diluted with Et_2O (150 mL) and washed with $NaHCO_3$ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure. TLC (20% EtOAc in petrol) revealed a major product spot ($R_f = 0.4$). This product was purified by column chromatography (20% EtOAc in petrol) and

identified as the deamination product **2.17** (36 mg, 47%). Analytical data of this material was consistent with that previously reported.²¹ IR (ν_{max} , film): 3064, 3003, 2360, 2341, 1738, 1604, 1496, 1454, 1436, 1364, 1294, 1256, 1196, 1164. ^1H NMR (500 MHz, CDCl_3): δ = 2.66 (2H, t, J = 7.9), 2.98 (2H, t, J = 7.9) (CH_2CH_2), 3.69 (3H, s, CO_2Me), 7.20-7.25 (3H, m, ArH), 7.28-7.33 (2H, m, ArH). ^{13}C NMR (125 MHz, CDCl_3): δ = 30.9, 35.7 (CH_2CH_2), 51.6 (CO_2Me), 126.2, 128.2, 128.5 (CH_{Ar}), 140.5 (4°_{Ar}), 173.3 (C=O). LRMS m/z (ESI^+): 187.1 $[\text{M}+\text{Na}]^+$.

Bis-Alkylation-Elimination in the Conversion of Cysteine to Dehydroalanine

Synthesis of diiodide 2.30

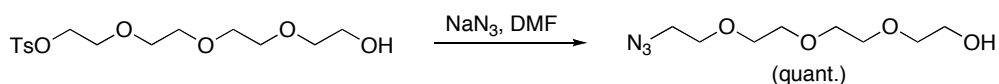


Tetraethylene glycol (20.0 g, 103 mmol) was added to a 250 mL round bottom flask and dissolved in THF (20 mL). The stirred solution was cooled to 0 °C and then NaOH (1.32 g, 33.0 mmol) was added as a solution in H_2O (20 mL). Tosyl chloride (3.93 g, 20.6 mmol) was prepared as a solution in THF (20 mL) and then poured into the reaction flask. The reaction was stirred at 0 °C, open to air, for 2 hours. After this time, the reaction mixture was diluted with EtOAc (200 mL) and H_2O (200 mL). The organic layer was separated, washed with brine (200 mL), dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (5% MeOH in EtOAc) to provide the tosylate as a clear oil (4.82 g, 67%). Spectroscopic data was consistent with that previously reported.²² IR (ν_{max} , film): 3447, 2874, 1597, 1454,

1355, 1177, 1098, 1017, 924. ^1H NMR (400 MHz, CDCl_3): δ = 2.41 (3H, s, CH_3), 3.55-3.68 (14H, m) 4.13 (2H, t, J = 4.8) ($-\text{OCH}_2\text{CH}_2\text{O}-$), 7.31 (2H, d, J = 8.1, ArH), 7.76 (2H, d, J = 8.1, ArH). ^{13}C NMR (100 MHz, CDCl_3): δ = 21.5 (CH_3), 61.5, 68.5, 69.1, 70.1, 70.3, 70.47, 70.54, 72.3 ($-\text{OCH}_2\text{CH}_2\text{O}-$), 127.8, 129.7 (CH_{Ar}), 132.7, 144.7 (4°_{Ar}). LRMS (ESI+): 371.1 ($\text{M}+\text{Na}$) $^+$.

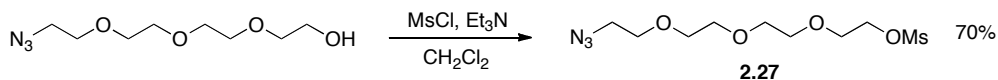
This reaction can also be done on a larger scale and no purification is required if excess tetraethylene glycol is used:

Tetraethylene glycol (200.0 g, 1.03 mol) was added to a 1 L round bottom flask and dissolved in THF (200 mL). The stirred solution was cooled to 0 °C and NaOH (6.18 g, 154 mmol) was added as a solution in H_2O (100 mL). The solution was stirred at 0 °C for 10 minutes before tosyl chloride (24.54 g, 128.7 mmol) was added as a solution in THF (50 mL). The reaction was stirred for 2 hours at 0 °C, open to air. After this time, the reaction was diluted with EtOAc (500 mL) and H_2O (500 mL). The organic layer was separated, washed with brine (500 mL), dried (MgSO_4), and filtered. Concentrating under reduced pressure provided the tosylate as a clear oil spectroscopically identical to the purified sample prepared above (32.92 g, 73%). This material was used in the next step without further purification.



The tosylate prepared above (32.91 g, 94.46 mmol) was dissolved in DMF (50 mL). Sodium azide (18.42 g, 283.4 mmol) was added to the stirred solution and the reaction was heated to 90 °C. The reaction was stirred, open to air, for 15 minutes. After this time, an additional 20 mL DMF was added to help with solubility. After a total of 1 hour at 90 °C, TLC (5% MeOH in EtOAc) revealed complete consumption of starting material. The reaction was then diluted with H_2O (500 mL) and extracted with CH_2Cl_2 (3 $\times$ 200 mL). The combined organic layers were dried (MgSO_4), filtered, and concentrated under reduced pressure. The product azide was isolated as a clear oil and did not require purification (20.71 g, quant.). Spectroscopic data was consistent with that previously reported.²² IR (ν_{max} , film): 3444, 2871, 2107, 1645, 1454, 1349, 1287, 1125, 937. ^1H NMR (400 MHz, CDCl_3): δ = 3.36 (2H, t, J = 4.8), 3.57 (2H, t, J = 4.3), 3.61-3.67 (10H,

m), 3.69 (2H, t, $J = 4.3$). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 50.5, 61.5, 69.9, 70.2, 70.43, 70.49, 70.53, 72.4$. LRMS (ESI $^{+}$): 242.1 ($\text{M}+\text{Na}$) $^{+}$.

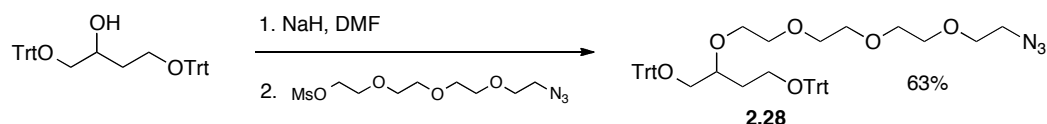


The azide derivative (20.71 g, 94.46 mmol) was added to a 500 mL round bottom flask and dissolved in CH_2Cl_2 (150 mL). Et_3N (40 mL, 283 mmol) was added and the stirred solution was cooled to 0 °C. Methane sulfonylchloride (11.0 mL, 142 mmol) was added slowly and the reaction was stirred at 0 °C for 10 minutes. The ice bath was then removed and the reaction stirred at room temperature for 1 hour. After this time, TLC (2% MeOH in EtOAc) revealed complete consumption of starting material. The reaction was diluted with CH_2Cl_2 (200 mL) and then washed with 1M HCl (200 mL). The aqueous layer was extracted with CH_2Cl_2 (200 mL). The combined organic layers were dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was then purified by column chromatography (2% MeOH in EtOAc) to afford the product mesylate as a yellow oil (19.73 g, 70%). Spectroscopic data was consistent with that previously reported.²³ IR (ν_{max} , film): 2872, 2109, 1454, 1414, 1174, 1128, 1017, 974, 922. ^1H NMR (400 MHz, CDCl_3): $\delta = 3.05$ (3H, s, SO_2Me), 3.36 (2H, t, $J = 5.1$, CH_2N_3), 3.63-3.65 (10H, m, $-\text{OCH}_2\text{CH}_2\text{O}-$), 3.74 (2H, t, $J = 4.6$, $\text{CH}_2\text{CH}_2\text{OMs}$), 4.35 (2H, t, $J = 4.6$, $\text{CH}_2\text{CH}_2\text{OMs}$). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 37.5$ (SO_2Me), 50.5 (CH_2N_3), 68.8, 69.2, 69.9, 70.43 ($2 \times \text{C}$), 70.47, 70.50 ($-\text{OCH}_2\text{CH}_2\text{O}-$). LRMS (ESI $^{+}$): 320.1 ($\text{M}+\text{Na}$) $^{+}$.



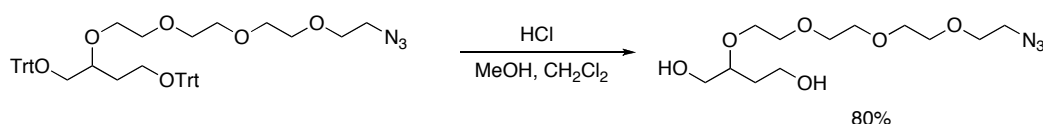
This procedure was adapted from the literature.²⁴ 1,2,4-Butanetriol (1.00 g, 9.42 mmol) was added to a 100 mL round bottom flask and dried under high vacuum for 15 minutes. The flask was then placed under an atmosphere of nitrogen and the triol was dissolved in pyridine (25 mL). Triphenylmethyl chloride (6.30 g, 22.6 mmol) was added in one portion and the reaction stirred at room temperature for 20 hours. The reaction was then diluted with EtOAc (300 mL) and then washed with 1M HCl (2×150 mL). The organic

layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was crystallized from CH_2Cl_2 and petrol. The product was isolated as small white crystals and dried under high vacuum (3.84 g, 69%). m.p. = 152-154 °C. IR (ν_{max} , film): 3566, 3085, 2926, 2876, 2360, 1596, 1490, 1447, 1032, 762, 647, 633. ^1H NMR (400 MHz, CDCl_3): δ = 1.83 (2H, q, J = 5.8, $\text{CH}_2\text{CH}_2\text{OTrt}$), 2.86 (1H, app s, OH), 3.12-3.16 (1H, m), 3.18-3.23 (2H, m), 3.29-3.35 (1H, m) ($2 \times \text{CH}_2\text{OTrt}$), 4.07 (1H, br app s, CHOH), 7.27-7.34 (18H, m), 7.44-7.48 (12H, m) ($2 \times \text{Trt}$). ^{13}C NMR (100 MHz, CDCl_3): δ = 33.7 ($\text{CH}_2\text{CH}_2\text{OTrt}$), 61.2 ($\text{CH}_2\text{CH}_2\text{OTrt}$), 67.5 (CH_2OTrt), 69.6 (CHOH), 86.5, 86.9 ($2 \times \text{Trt}$), 126.9, 127.0, 127.8 ($2 \times \text{C}$), 128.5, 128.6, 143.9, 144.0 ($2 \times \text{Trt}$). LRMS (ESI $^+$): 613.2 ($\text{M}+\text{Na}$) $^+$.

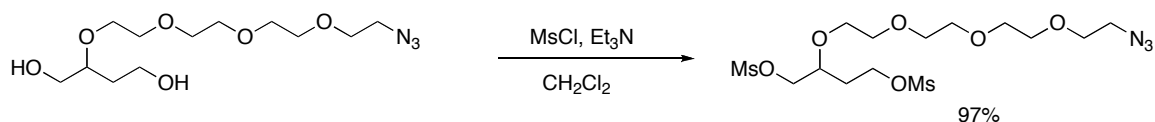


The mesylate (12.04 g, 40.49 mmol) was added to a 25 mL round bottom flask and dried under high vacuum for 30 minutes prior to use. The bis-trityl ether (15.95 g, 26.99 mmol) was added to a flame dried 2-neck, 250 mL round bottom flask and placed under an atmosphere of nitrogen before dissolving in anhydrous DMF (50 mL). Sodium hydride (1.62 g, 60% in mineral oil, 40.5 mmol) was added carefully at room temperature and an efficient exit for hydrogen gas was ensured. The reaction mixture was stirred for 10 minutes at room temperature before the dried mesylate was added as a solution in DMF (10 mL). After 3 hours at room temperature, the reaction was cooled to 0 °C and quenched by the careful addition of saturated aqueous NH_4Cl (50 mL). The mixture was then diluted with EtOAc (300 mL) and washed with H_2O (300 mL) and brine (300 mL). The organic layer was then dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (25% EtOAc in petrol) to afford alkylated product as a thick, clear oil (13.49 g, 63%). IR (ν_{max} , film): 3085, 3058, 3031, 2872, 2104, 1596, 1490, 1448, 1285, 1182, 1089, 763, 747, 706, 632. ^1H NMR (400 MHz, CDCl_3): δ = 1.75-1.87 (2H, m, $\text{CH}_2\text{CH}_2\text{OTrt}$), 3.09-3.13 (4H, m, $2 \times \text{CH}_2\text{OTrt}$), 3.34 (2H, t, J = 5.0, CH_2N_3), 3.45-3.50 (3H, m), 3.54-3.57 (4H, m), 3.60-3.70 (7H, m), 3.73-3.78 (1H, m) ($-\text{OCH}_2\text{CH}_2\text{O}-$ and CHCH_2OTrt), 7.19-7.29 (18H, m), 7.38-7.45 (12H, m) ($2 \times \text{Trt}$). ^{13}C NMR (100 MHz, CDCl_3): δ = 32.7 ($\text{CH}_2\text{CH}_2\text{OTrt}$), 50.5

(CH₂N₃), 59.9, 66.0, 69.62, 69.89, 70.44, 70.52 (2 × C), 70.55, 70.66 (-OCH₂CH₂O- and 2 × CH₂OTrt), 76.8 (CHCH₂OTrt), 86.3, 86.4 (2 × Trt), 126.74, 126.79, 127.62, 127.67, 128.55, 128.65, 144.06, 144.24 (2 × Trt). HRMS m/z (ESI⁺): Found 814.3832 [M+Na]⁺; C₅₀H₅₃N₃O₆Na requires 814.3827.

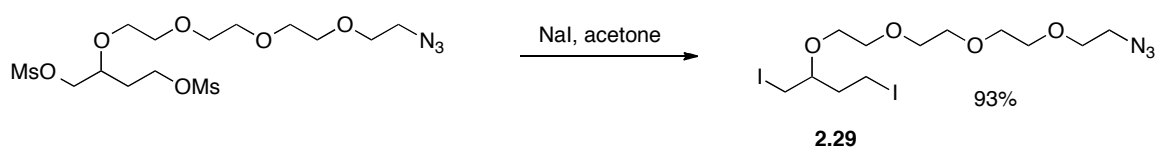


Anhydrous MeOH (50 mL) was added to a 250 mL round bottom flask and placed under an atmosphere of nitrogen. The flask was cooled to 0 °C and acetyl chloride (12.2 mL, 170.3 mmol) was added slowly. The solution was stirred for 10 minutes to generate methanolic HCl. The bis-trityl ether (13.49 g, 17.03 mmol) was added to the reaction flask as a solution in CH₂Cl₂ (30 mL). The reaction was then warmed to room temperature and stirred for 2 hours after which time TLC (5% MeOH in CH₂Cl₂) revealed complete consumption of starting material. Volatile components were removed under reduced pressure and the resulting residue was purified directly by column chromatography (5% MeOH in CH₂Cl₂ then 7% MeOH in CH₂Cl₂). The product was isolated as a thick, clear oil (4.20 g, 80%). IR (ν_{max}, film): 3418, 2873, 2108, 1646, 1454, 1347, 1300, 1120, 945, 850. ¹H NMR (400 MHz, MeOD): δ = 1.72 (2H, q, *J* = 6.1, CH₂CH₂OH), 3.39 (2H, t, *J* = 4.9, CH₂N₃), 3.50-3.58 (2H, m), 3.61-3.72 (16H, m), 3.76-3.79 (1H, m) (-OCH₂CH₂O-, 2 × CH₂OH, and CHCH₂OH). ¹³C NMR (100 MHz, MeOD): δ = 35.5 (CH₂CH₂OH), 51.9 (CH₂N₃), 59.8, 65.1, 70.36, 71.26, 71.65 (2 × C), 71.67, 71.74, 72.05 (-OCH₂CH₂O- and 2 × CH₂OH), 79.7 (CHCH₂OH). HRMS m/z (ESI⁺): Found 330.1637 [M+Na]⁺; C₁₂H₂₅N₃O₆Na requires 330.1636.



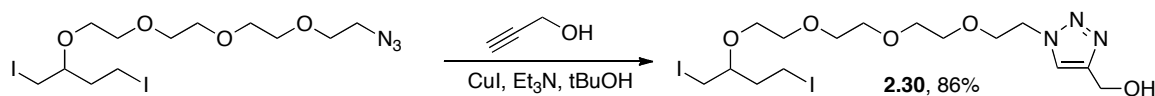
The diol (4.20 g, 13.7 mmol) was added to a 250 mL round bottom flask and dissolved in CH₂Cl₂ (60 mL). Et₃N (11.4 mL, 82.0 mmol) was added and the resulting solution was cooled to 0 °C. Methane sulfonylchloride (3.17 mL, 41.0 mmol) was added slowly to the reaction. After the addition was complete, the reaction was stirred for 5 minutes before

the ice bath was removed. The reaction was then stirred at room temperature for 1 hour. After this time the reaction was quenched by diluting with CH₂Cl₂ (100 mL) and 1M HCl (100 mL). The organic layer was separated and the aqueous layer was extracted with CH₂Cl₂ (2 × 70 mL). The combined organic layers were dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (2% MeOH in CH₂Cl₂) to afford the busulfan derivative as a clear oil (6.13 g, 97%). IR (ν_{max} , film): 3022, 2937, 2874, 2109, 1463, 1352, 1174, 1123, 950, 836. ¹H NMR (400 MHz, CDCl₃): δ = 1.92-1.99 (2H, m, CH₂CH₂OMs), 3.01 (3H, s), 3.05 (3H, s) (2 × SO₂Me), 3.36 (2H, t, *J* = 5.1, CH₂N₃), 3.58-3.66 (13H, m), 3.75-3.82 (2H, m) (-OCH₂CH₂O- and CHCH₂OMs), 4.16 (1H, dd, *J* = 11.1, 4.8), 4.29-4.43 (3H, m) (2 × CH₂OMs). ¹³C NMR (100 MHz, CDCl₃): δ = 31.1 (CH₂CH₂OMs), 37.1, 37.4 (2 × SO₂Me), 50.5 (CH₂N₃), 66.0, 69.8, 69.9, 70.27, 70.32, 70.37, 70.43, 70.44, 70.47 (-OCH₂CH₂O- and 2 × CH₂OMs), 73.7 (CHCH₂OMs). HRMS *m/z* (ESI⁺): Found 486.1189 [M+Na]⁺; C₁₄H₂₉N₃O₁₀SNa requires 486.1187.



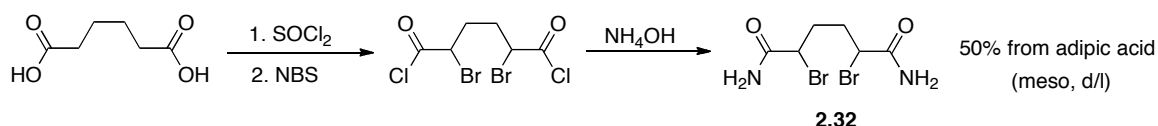
The Busulfan derivative (2.00 g, 4.31 mmol) was added to a 100 mL round bottom flask and dissolved in acetone (25 mL). Sodium iodide (6.47 g, 43.2 mmol) was added and the flask was then equipped with a condenser and placed under an atmosphere of nitrogen. The reaction was heated to reflux and stirred for 5 hours. After this time, the reaction was cooled to room temperature and diluted with EtOAc (200 mL) and H₂O (200 mL). The organic layer was separated and then washed with brine (200 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (60% EtOAc in petrol) to afford the di-iodide as a clear oil (2.12 g, 93%). IR (ν_{max} , film): 2868, 2103, 1439, 1347, 1297, 1246, 1114. ¹H NMR (400 MHz, CDCl₃): δ = 2.04-2.17 (2H, m, CH₂CH₂I), 3.24-3.31 (4H, m, 2 × CH₂I), 3.37-3.42 (3H, m, CH₂N₃ and CHCH₂I), 3.59-3.68 (13H, m), 3.76-3.82 (1H, m) (-OCH₂CH₂O-). ¹³C NMR (100 MHz, CDCl₃): δ = 2.19, 8.59 (2 × CH₂I), 38.7 (CH₂CH₂I), 50.6 (CH₂N₃), 69.25, 69.95, 70.59 (br, 3 × C), 70.63, 70.64 (-OCH₂CH₂O-), 78.3

(CHCH₂I). HRMS *m/z* (ESI⁺): Found 549.9647 [M+Na]⁺; C₁₂H₂₃N₃O₄I₂Na requires 549.9670.



The azide (149 mg, 0.28 mmol) was added to a 5 mL round bottom flask and dissolved in *t*BuOH (0.40 mL). Propargyl alcohol (33 μ L, 0.56 mmol), Et₃N (0.20 mL, 1.4 mmol), and CuI (5 mg, 0.028 mmol) were added sequentially to the stirred solution. The flask was sealed with a septum and placed under an atmosphere of nitrogen. The reaction was stirred at room temperature for 3 hours after which time all starting material had been consumed (TLC, 5% MeOH in CH₂Cl₂). The reaction mixture was loaded directly to a silica column and purified by elution with 5% MeOH in CH₂Cl₂. Triazole **2.30** was isolated as a clear liquid (142 mg, 86%). IR (ν_{max} , film): 3385, 2869, 1600, 1456, 1351, 1181, 1096. ¹H NMR (400 MHz, DMSO): δ = 1.91-2.06 (2H, m, CH₂CH₂I), 3.23-3.54 (3H, m, contains CHCH₂I and CH₂CH₂I), 3.37-3.46 (2H, m, CHCH₂I), 3.48-3.57 (11H, m), 3.63-3.69 (1H, m) (-OCH₂CH₂O-), 3.81 (2H, t, *J* = 5.3, OCH₂CH₂N), 4.48-4.51 (4H, m, includes CH₂OH and OCH₂CH₂N), 5.16 (1H, app s, OH), 7.93 (1H, s, CH_{Ar}). ¹³C NMR (100 MHz, DMSO): δ = 3.5, 11.3 (2 $\times$ CH₂I), 38.7 (CH₂CH₂I), 49.2, 55.1 (CH₂OH and OCH₂CH₂N), 68.14, 68.80, 69.56, 69.67, 69.76 (br 2 $\times$ C), 69.84 (-OCH₂CH₂O-), 77.4 (CHCH₂I), 123.1 (CH_{Ar}), 147.9 (4°_{Ar}). HRMS *m/z* (ESI⁺): Found 605.9938 [M+Na]⁺; C₁₅H₂₇N₃O₅I₂Na requires 605.9932.

Synthesis of dibromide 2.32

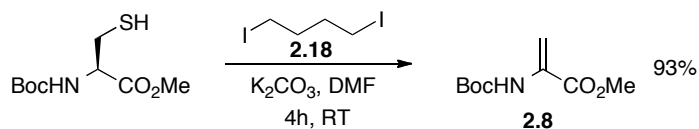


α,α' -Di-bromo-adipyl(bis)amide was prepared using a procedure adapted from the literature.^{25, 26} Adipic acid (25.00 g, 171.1 mmol) was added to a 500 mL round bottom flask and suspended in thionylchloride (75.0 mL, 1034 mmol). The flask was equipped with a condenser and the reaction was heated to reflux (open to air, bath temp 80 °C). After 30 minutes at reflux, all adipic acid had dissolved. The reaction was stirred for an

additional 60 minutes at reflux and then cooled to room temperature. CCl₄ (100 mL) was added to the reaction followed by NBS (73.1 g, 411 mmol). The reaction was stirred vigorously and 10 drops of HBr (48% aq.) was added by pipette. The reaction was heated to reflux, again open to air. The reaction gradually turns from red to black over the course of an hour. After 2 hours at reflux, the reaction was cooled to room temperature and then to 0 °C. The mixture was stirred at 0 °C to ensure all succinimide had precipitated. The solid was removed by filtration. Et₂O (50 mL) was used to rinse and complete the filtration. The filtrate was concentrated under reduced pressure to give a thick, dark red liquid. In a 500 mL round bottom flask, 200 mL of NH₄OH (25% aq.) was cooled to 0 °C. The crude acid chloride was added dropwise over 20 minutes to the ammonia solution with rapid stirring. After the addition was complete, the reaction was stirred vigorously at 0 °C for 1 hour. The bis-amide product precipitated from the reaction mixture. The dark solid was isolated by filtration and partially dried. The product was purified by triturating in MeOH/H₂O: The solid was suspended in H₂O (100 mL) and MeOH (100 mL) and heated to 60 °C. The mixture was stirred rapidly at 60 °C for 30 minutes. After this time, the mixture was cooled to room temperature. The resulting white solid (a mixture of *meso* and *d/l* diastereomers) was isolated by filtration and washed with MeOH (200 mL). The product (**2.32**) was dried under high vacuum (25.75 g, 50%). IR (ν_{max} , KBr): 3302, 2946, 2801, 1684, 1418, 1319, 1277, 1250, 1220, 1194, 978, 876. ¹H NMR (400 MHz, DMSO): δ = 1.75-2.08 (4H, m, CH₂CH₂), 4.28-4.36 (2H, m, 2 × CHBr), 7.30 (2H, s), 7.69 (2H, s) (2 × NH₂). ¹³C NMR (100 MHz, DMSO): (both diastereomers reported), δ = 32.5, 32.6 (CH₂CHBr), 48.2, 48.5 (CH₂CHBr), 169.87, 169.92 (C=O). LRMS *m/z* (ESI⁺): 324.9 [M+Na]⁺.

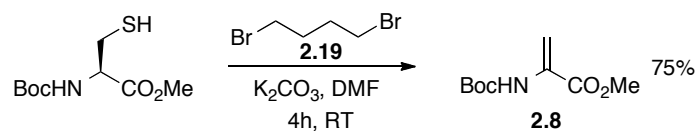
Di-alkylative conversion of cysteine to dehydroalanine:

BocDhaOMe using 1,4-diiodobutane



BocCysOMe (50 mg, 0.21 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (5 mL). K_2CO_3 (147 mg, 1.06 mmol) was added to the stirred solution followed immediately by 1,4-diiodobutane (42 μ L, 0.32 mmol). The flask was stoppered and the reaction stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et_2O (100 mL) and washed sequentially with H_2O (100 mL) and brine (100 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to provide BocDhaOMe as a clear oil (40 mg, 93%). IR (ν_{max} , film): 3423, 2980, 1719, 1634, 1513, 1328, 1159, 1068. 1H NMR ($CDCl_3$, 400 MHz): δ = 1.46 (9H, s, Boc), 3.80 (3H, s, CO_2CH_3), 5.70 (1H, d, J = 1.5, $C=CHH$), 6.13 (1H, app s, $C=CHH$), 7.00 (1H, br s, NH). ^{13}C NMR (100 MHz): δ = 28.2 (Boc), 52.8 (CO_2CH_3), 80.6 (Boc), 105.1 ($C=CH_2$), 131.3 ($C=CH_2$), 152.5, 164.4 (2 $\times$ $C=O$). Anal. for $C_9H_{15}NO_4$: C, 53.72; H, 7.51; N, 6.96; found C, 53.95; H, 7.63; N, 6.83.

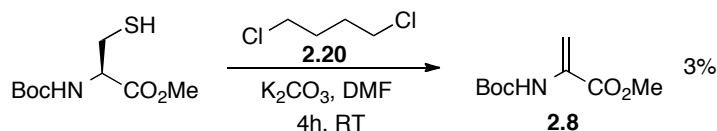
BocDhaOMe using 1,4-dibromobutane



BocCysOMe (50 mg, 0.21 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (5 mL). K_2CO_3 (147 mg, 1.06 mmol) was added to the stirred solution followed immediately by 1,4-dibromobutane (38 μ L, 0.32 mmol). The flask was stoppered and the reaction stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et_2O (100 mL) and washed sequentially with H_2O (100 mL) and brine (100 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography

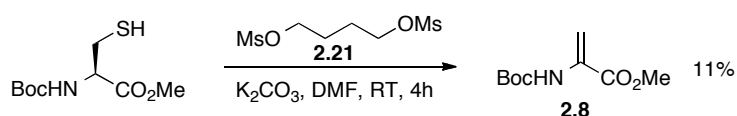
(10% EtOAc in petrol) to provide BocDhaOMe as a clear oil (32 mg, 75%). Spectroscopic data was identical to that reported above.

BocDhaOMe using 1,4-dichlorobutane



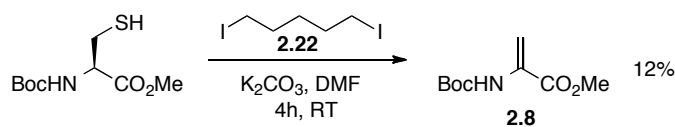
BocCysOMe (50 mg, 0.21 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (5 mL). K₂CO₃ (147 mg, 1.06 mmol) was added to the stirred solution followed immediately by 1,4-dichlorobutane (36 μ L, 0.32 mmol). The flask was stoppered and the reaction stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et₂O (100 mL) and washed sequentially with H₂O (100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to provide BocDhaOMe as a clear oil (1.4 mg, 3%). Spectroscopic data was identical to that reported above.

BocDhaOMe using Busulfan



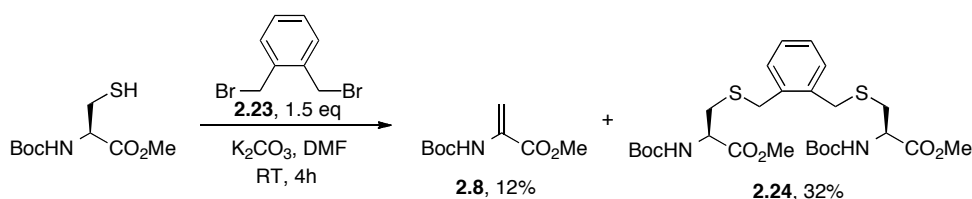
BocCysOMe (55 mg, 0.23 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (5 mL). K₂CO₃ (162 mg, 1.17 mmol) was added to the stirred solution, followed immediately by Busulfan (**2.21**) (85 mg, 0.35 mmol). The flask was stoppered and the reaction was stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et₂O (100 mL) and washed sequentially with H₂O (100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to provide BocDhaOMe (5 mg, 11%). Spectroscopic data was identical to that reported above.

BocDhaOMe using 1,5-diiodopentane

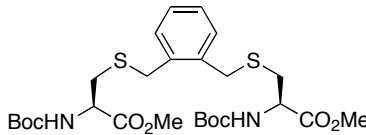


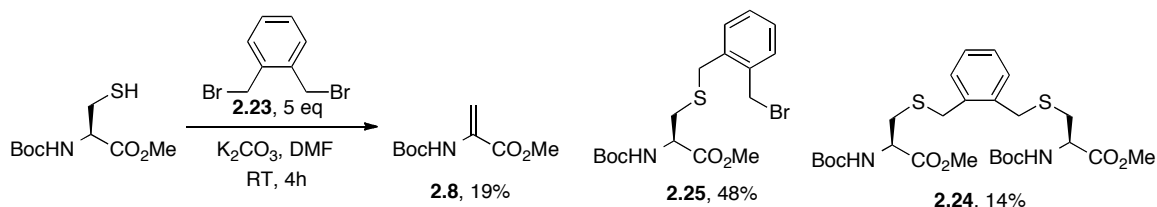
BocCysOMe (50 mg, 0.21 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (5 mL). K_2CO_3 (147 mg, 1.06 mmol) was added followed immediately by 1,5-diiodopentane (47 μ L, 0.32 mmol). The reaction was stoppered and stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et_2O (100 mL) and washed sequentially with H_2O (100 mL) and brine (100 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to provide BocDhaOMe (5 mg, 12%). Spectroscopic data was identical to that reported above.

BocDhaOMe using α,α' -dibromo-*o*-xylene

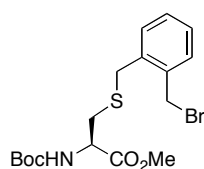


BocCysOMe (50 mg, 0.21 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (5 mL). K_2CO_3 (147 mg, 1.06 mmol) was added, followed immediately by α,α' -dibromo-*o*-xylene (84 mg, 0.32 mmol). The reaction was stoppered and stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et_2O (100 mL) and washed sequentially with H_2O (100 mL) and brine (100 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography by first eluting with 10% EtOAc in petrol to provide BocDhaOMe (5 mg, 12%) and then with 40% EtOAc in petrol to provide **2.24** (19 mg, 32%). Spectroscopic data for BocDhaOMe was identical to that described above.


 Data for **2.24**: m.p. = 130-131 °C (MeOH). $[\alpha]_D^{20} = -3.7$ ($c = 0.8$, CHCl_3). IR (ν_{max} , film): 3366, 2977, 1746, 1715, 1505, 1392, 1249, 1214, 1166, 1053, 1016. ^1H NMR (CDCl_3 , 500MHz): $\delta = 1.46$ (18H, s, Boc), 2.86-2.98 (4H, ABX system, $J = 13.9, 5.4, 4.7$, $\text{H}\beta_{\text{Cys}}$), 3.76 (6H, s, CO_2Me), 3.84-3.94 (4H, AB System, $J = 12.9$, SCH_2Ar), 4.57 (2H, m, $\text{H}\alpha_{\text{Cys}}$), 5.39 (2H, d, $J = 7.9$, NH), 7.21-7.29 (4H, m, CH_{Ar}). ^{13}C NMR (125 MHz, CDCl_3): $\delta = 28.3$ (Boc), 33.9, 34.2 ($\text{C}\beta_{\text{Cys}}$ and SCH_2Ar), 52.6 (CO_2Me), 53.2 ($\text{C}\alpha_{\text{Cys}}$), 80.2 (Boc), 127.7, 130.9, 135.7 (Ar), 155.1, 171.6 ($2 \times \text{C}=\text{O}$). HRMS m/z (ESI^+): Found 595.2125 $[\text{M}+\text{Na}]^+$; $\text{C}_{26}\text{H}_{40}\text{N}_2\text{O}_8\text{S}_2\text{Na}$ requires 595.2118.

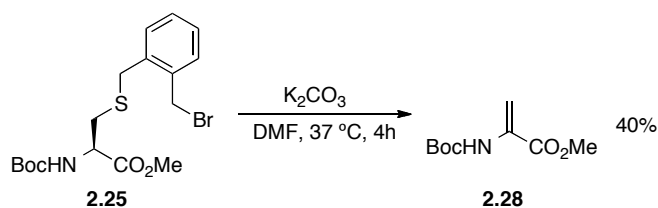


BocCysOMe (50 mg, 0.21 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (3 mL). K_2CO_3 (147 mg, 1.06 mmol) was added followed immediately by a solution of α, α' -dibromo-*o*-xylene (280 mg, 1.06 mmol) in DMF (1 mL). An additional 1 mL of DMF was used to rinse and ensure complete transfer. The reaction was stoppered and stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et_2O (100 mL) and washed sequentially with H_2O (100 mL) and brine (100 mL). The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography by first eluting with 10% EtOAc in petrol to provide BocDhaOMe (**2.8**, 8 mg, 19%) and then with 40% EtOAc in petrol to provide uncyclized intermediate **2.25** (41 mg, 48%) and dicysteinylated product **2.24** (8 mg, 14%). Spectroscopic data for BocDhaOMe and **2.24** was identical to that described above.



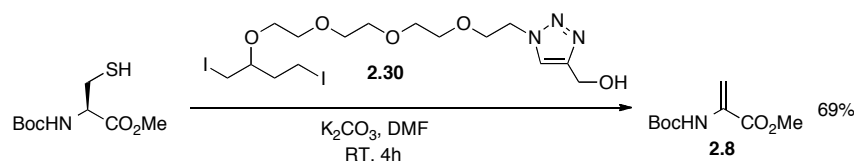
Data for uncyclized product **2.25**: clear oil. $[\alpha]_D^{20} = -5.3$ ($c = 1.7$, CHCl₃). IR ($\nu_{\max}$, film): 3376, 2977, 2360, 2341, 1746, 1714, 1498, 1366, 1165. ¹H NMR (CDCl₃, 400MHz): $\delta = 1.46$ (9H, s, Boc), 2.87-3.02 (2H, ABX system, $J = 13.9, 5.6, 4.6$, H β_{Cys}), 3.77 (CO₂Me), 3.92 (2H, AB system, $J = 12.9$, SCH₂Ar), 4.58 (1H, m, H α_{Cys}), 4.67 (2H, AB system, $J = 10.4$, CH₂Br), 5.35 (1H, d, $J = 7.1$, NH), 7.25-7.36 (4H, m, CH_{Ar}). ¹³C NMR (100 MHz, CDCl₃): $\delta = 28.3$ (Boc), 30.9 (CH₂Br), 33.8, 34.5 (C β_{Cys} and SCH₂Ar), 52.6 (CO₂Me), 53.2 (C α_{Cys}), 80.2 (Boc), 128.1, 129.0, 130.9, 131.1 (CH_{Ar}), 136.16, 136.22 (4° Ar), 155.1, 171.5 (2 × C=O). HRMS m/z (ESI⁺): Found 440.0502 [M+Na]⁺; C₁₇H₂₄NO₄SBrNa requires 440.0502.

Conversion of **2.25** to BocDhaOMe



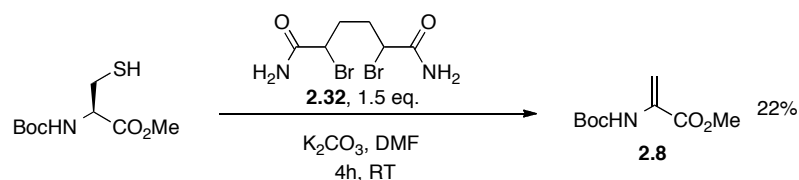
Uncyclized cysteine derivative **2.25** (236 mg, 0.560 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (6 mL). K₂CO₃ (390 mg, 2.82 mmol) was added to the stirred solution and the reaction was heated to 37 °C. After 4 hours, the reaction was diluted with Et₂O (150 mL) and then washed with H₂O (100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to provide BocDhaOMe as a clear oil (45 mg, 40%). Spectroscopic data was identical to that reported above.

BocDhaOMe using diiodide **2.30**

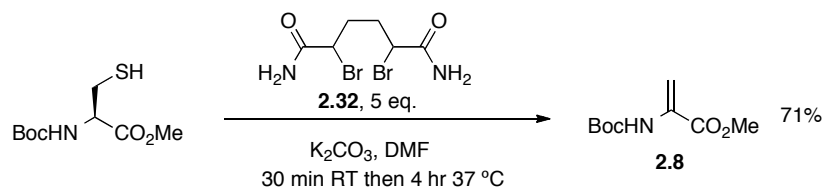


BocCysOMe (34 mg, 0.14 mmol) was added to a 25 mL round bottom flask and dissolved in DMF (2 mL). K₂CO₃ (97 mg, 0.70 mmol) was added followed immediately by a solution of diiodide **2.30** (126 mg, 0.220 mmol) in DMF (1.5 mL). An additional 1.5 mL DMF was used to rinse and complete the transfer. The flask was stoppered and the reaction stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et₂O (100 mL) and washed sequentially with H₂O (100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to provide BocDhaOMe (20 mg, 69%). Spectroscopic data was identical to that reported above.

BocDhaOMe using di-bromide **2.32**



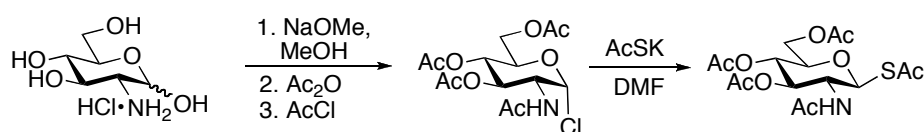
BocCysOMe (100 mg, 0.42 mmol) was added to a 10 mL round bottom flask and dissolved in DMF (3 mL). K₂CO₃ (290 mg, 2.10 mmol) was added to the stirred solution followed immediately by a solution of di-bromide **2.32** (192 mg, 0.640 mmol) in DMF (2 mL). An additional 1 mL of DMF was used to wash and complete the transfer. The flask was stoppered and the reaction stirred at room temperature for 4 hours. After this time, the reaction was diluted with Et₂O (80 mL) and washed sequentially with H₂O (80 mL) and brine (80 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to provide BocDhaOMe as a clear oil (19 mg, 22%). Spectroscopic data was identical to that reported above.



BocCysOMe (459 mg, 1.95 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (20 mL). K₂CO₃ (1.35 g, 9.77 mmol) and di-bromide **2.32** (2.95 g, 9.77 mmol) were both added as solids. The reaction was stirred at room temperature for 30 minutes and then placed in a water bath pre-heated to 37 °C. After stirring for 4 hours at

37 °C, the reaction was diluted with EtOAc (250 mL) and then washed with H₂O (150 mL) and brine (150 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% EtOAc in petrol) to afford BocDhaOMe as a clear oil (279 mg, 71%). Spectroscopic data was identical to that reported above.

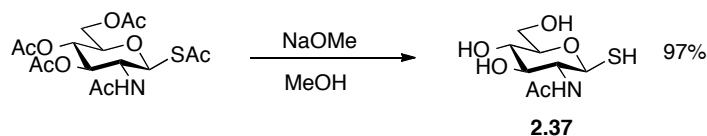
3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio-β-*D*-glucosamine



(D)-Glucosamine hydrochloride (20.0 g, 92.7 mmol) was added to a 100 mL round bottom flask and suspended in MeOH (20 mL). Sodium methoxide (6.00 g, 111 mmol) was added in several portions and the mixture stirred vigorously at room temperature for 1 hour before NaCl was removed by filtration. The solids were washed with MeOH (20 mL). Acetic anhydride (12 mL, 127 mmol) was added to the stirred filtrate at 0 °C and warmed slowly to room temperature. After 12 hours, the reaction was concentrated to ~20 mL total volume and the product precipitated by the addition of Et₂O (100 mL). The solid was isolated by filtration and dried under vacuum before returning to an oven dried 2-neck 250 mL round bottom flask equipped with a condenser. The flask was placed under and argon atmosphere and cooled to 0 °C. Acetyl chloride (40.0 mL, 563 mmol) was added slowly and the mixture stirred vigorously. After the addition, the reaction was heated to reflux. Gradually the reaction turned black. After 3 hours at reflux, the reaction was cooled to room temperature, diluted with CH₂Cl₂ (200 mL), and poured into a beaker of ice (~200 g). The organics were washed with NaHCO₃ (sat. aq., 2 × 200 mL), dried (MgSO₄), filtered, and concentrated to a black residue. The residue was dissolved in EtOAc and filtered through a short column of silica, eluting with EtOAc. The product was precipitated by the addition of cold Et₂O and petrol. The resulting powder was isolated by filtration and dried under vacuum (10.19 g, 50% from glucosamine hydrochloride). The crude product (~4:1 mixture of α:β) was used in the final step without purification. Accordingly, a portion of the anomeric chloride (1.00 g, 2.73 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (15 mL). Potassium thioacetate (1.56 g, 13.7 mmol) was then added in one portion. The reaction was flushed with nitrogen and stirred at room temperature for 3 hours. After this time, no starting

material was detected by TLC (70% EtOAc in petrol). The reaction mixture was diluted with EtOAc (250 mL) and washed sequentially with NaHCO₃ (sat. aq., 200 mL), H₂O (200 mL), and brine (200 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure to give a brown solid. This solid was recrystallized from CH₂Cl₂ and petrol to afford the product as white crystals (661 mg, 60%). m.p. = 186-188 °C (Lit²⁷: 196-197, *i*PrOH). $[\alpha]_D^{20} = +12.1$ (*c* = 2.2, CHCl₃) (Lit²⁷: +14.1 *c* = 2.16, CHCl₃). IR ($\nu_{\max}$, KBr): 3733, 3064, 2961, 2360, 2341, 1747, 1707, 1665, 1538, 1433, 1370, 1230, 1048. ¹H NMR (400 MHz, CDCl₃): δ = 1.90 (3H, s), 2.02 (6H, s), 2.05 (3H, s), 2.35 (3H, s) (5 × Ac), 3.81 (1H, ddd, *J* = 9.4, 4.6, 1.3, H5), 4.08 (1H, dd, *J* = 12.4, 1.3, H6), 4.22 (1H, dd, *J* = 12.4, 4.6, H6'), 4.34 (1H, q, *J* = 10.3, H2), 5.10-5.16 (2H, m, H3, H4), 5.17 (1H, d, *J* = 10.3, H1). ¹³C NMR (100 MHz, CDCl₃): δ = 20.49, 20.59, 20.64, 23.0, 30.8 (5 × Ac), 51.9 (C2), 61.8 (C6), 67.8, 74.0 (C3, C4), 76.3 (C5), 81.5 (C1), 169.2, 170.0, 170.6, 171.3, 193.5 (5 × C=O). LRMS (ESI⁺): 428.1 (M+Na)⁺.

***N*-acetyl-1-thio- β -D-glucosamine (GlcNAc thiol **2.37**)**



3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio- β -D-glucosamine (250 mg, 0.62 mmol) was added to a 50 mL round bottom flask and placed under an atmosphere of nitrogen before dissolving in anhydrous MeOH (8 mL). Sodium methoxide (67 mg, 1.2 mmol) was added under a stream of nitrogen at room temperature. The reaction was stirred for 1 hour at room temperature before quenching with DOWEX-50WX8 (H⁺ form) until the pH of the reaction mixture was neutral (pH paper). The resin was removed by filtration and washed with MeOH (~ 50 mL). The filtrate was evaporated to provide GlcNAc thiol **2.37** as a thick syrup that crystallized upon standing (142 mg, 97%). m.p. = 170-171 °C (Lit²⁸: 174-176 °C). $[\alpha]_D^{20} = -13.5$ (*c* = 1.0, MeOH) (Lit²⁸: -10.4, *c* = 1.0, MeOH). IR ($\nu_{\max}$, KBr): 3287, 2360, 2341, 1647, 1558, 1374, 1313, 1056. ¹H NMR (400 MHz, D₂O): δ = 1.90 (3H, s, Ac), 3.30-3.39 (3H, m, H3, H4, H5), 3.56 (1H, dd, *J* = 11.1, 5.3, H6), 3.59 (1H, t, *J* = 10.2, H2), 3.74 (1H, dd, *J* = 11.1, 1.2, H6'), 4.53 (1H, d, *J* = 10.2, H1). ¹³C NMR (50 MHz, D₂O): δ = 22.6 (Ac), 58.2 (C2), 61.1 (C6), 70.0, 75.3, 79.3, 80.52 (C3, C4, C5, C1). LRMS (ESI⁻): 236.1 (M-H)⁻. Note: in D₂O this material anomerizes to a β : α over a period of 24 hours.

PROTEIN MODIFICATION

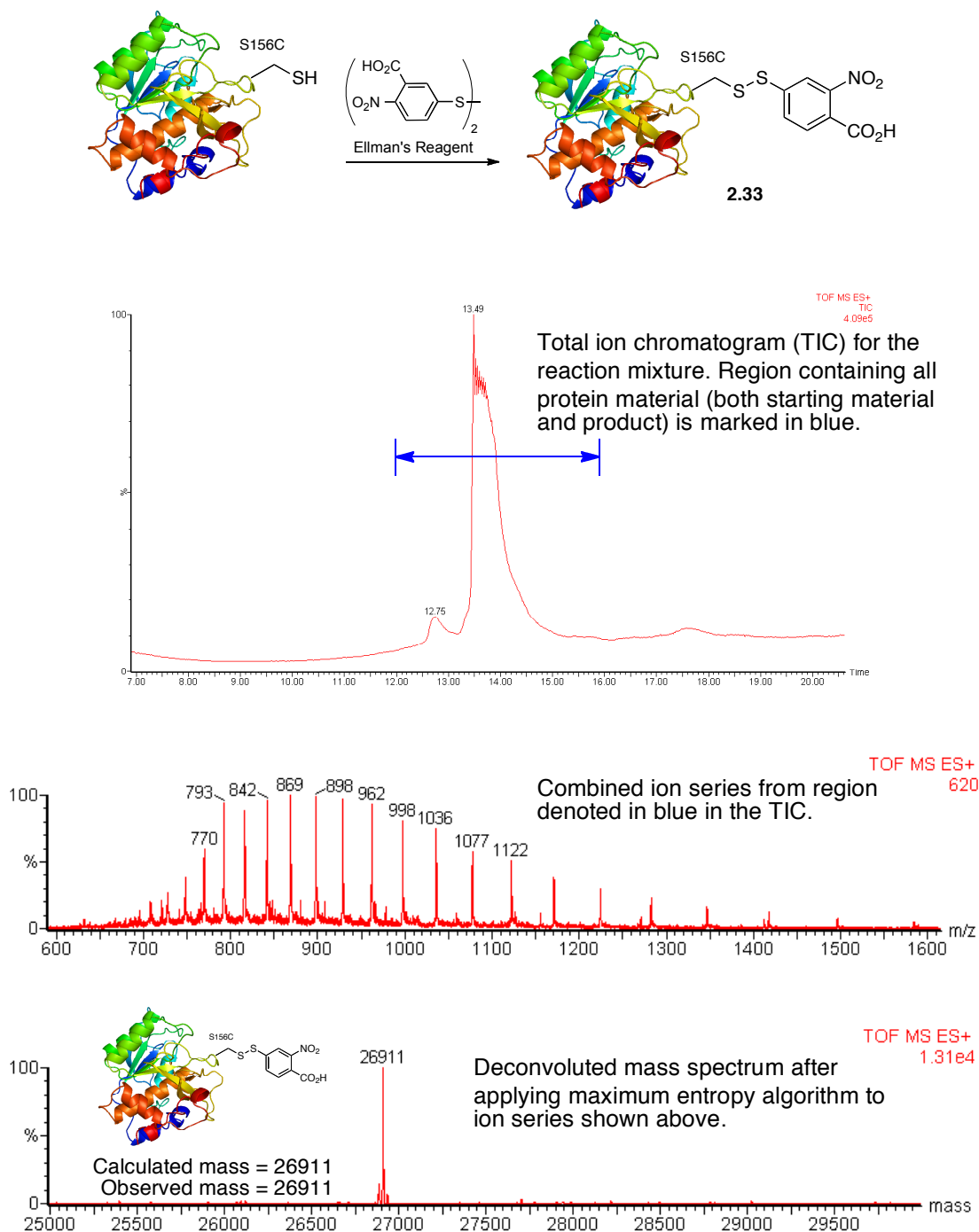
Protein Mass Spectrometry

Liquid chromatography-mass spectrometry (LC-MS) was performed on a Micromass LCT (ESI-TOF-MS) coupled to an Agilent 1100 Series HPLC using a Phenomenex Jupiter 5 μ m C4 column (250 $\times$ 4.6 mm). Water:acetonitrile, 95:5 (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid by volume, were used as the mobile phase at a flow rate of 1.0 mL/min. The gradient was programmed as follows: 95% A for 5 min to desalt and then a linear gradient to 100% B over 15 min followed by 100% B for an additional 5 min. A linear gradient over 10 minutes back to 95% A was used to re-equilibrate the column. The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 l/hr. Under these conditions, all protein material typically co-elutes in a single peak between 13 and 18 minutes. For reaction analysis, the mass spectra for all protein material in this peak were combined using MassLynx software (v. 4.0 from Waters). Mass spectra were then calibrated using a calibration curve constructed from a minimum of 17 matched peaks from the multiply charged ion series of equine myoglobin obtained at a cone voltage of 25V. The calibrated, combined ion series was deconvoluted using a maximum entropy algorithm that is preinstalled on the MassLynx software. If multiple proteins are present, the relative peak height is used to calculate the relative amount of each protein and reaction conversions.

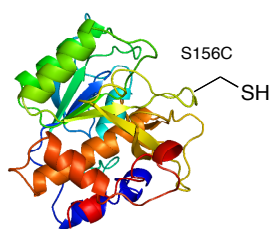
For a discussion of using ESI-MS to following reaction conversions on protein substrates, the reader is referred to a published validation of this technique.²⁹ In short, two derivatives of subtilisin from *Bacillus lentus* (SBL) were mixed in known amounts, as measured by the bicinchoninic acid (BCA) assay. The relative amounts of each protein were then correlated to relative peak heights on ESI-MS. Excellent correlation was observed. Therefore, for reactions on SBL that result in a change at a single residue, the unmodified and modified proteins ionize with the same efficiency using ESI. Thus, to a first approximation, the relative peak height observed by ESI-MS can be used to calculate reaction conversion. For all protein reactions in this Thesis, reaction conversions are reported based on relative peak height measure in the ESI-MS. For those proteins that were purified by dialysis or size-exclusion chromatography after modification, protein

recovery is reported based on concentrations measurements by bicinchoninic acid (BCA) assay.

A representative LC-MS analysis for a protein reaction is shown below:



Reduction-Elimination of Cysteine to Dehydroalanine: Attempts on SBL

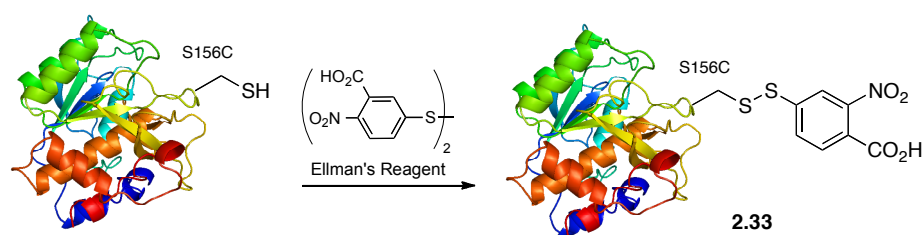


SBL-S156C (PDB for wild type = 1GCI)

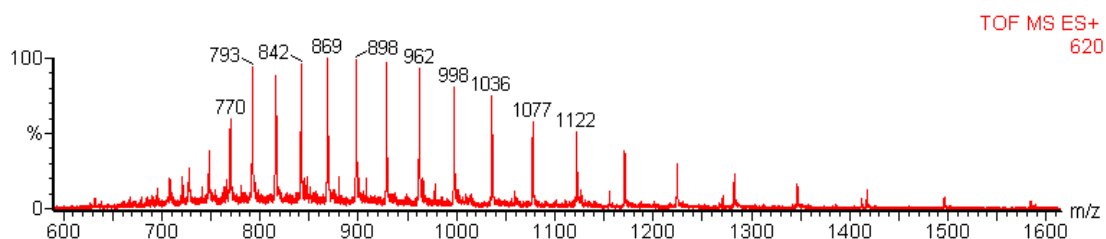
Sequence:

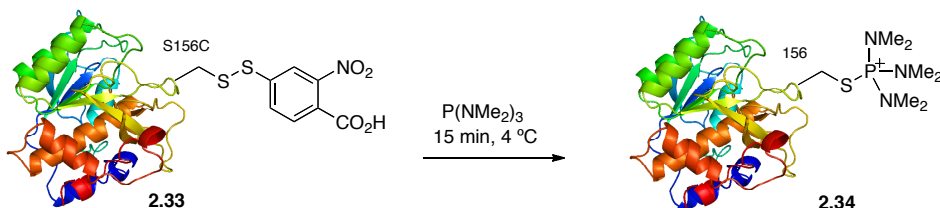
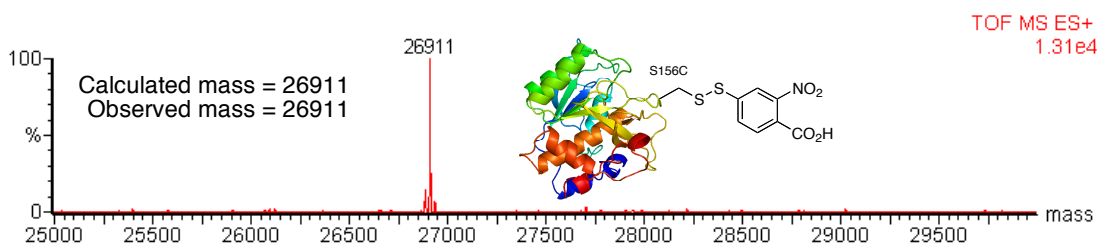
AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPS
 TQDGNGHGHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGL
 EWAGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGN**C**GAGSISYPA
 RYANAMAVGATDQNNNRASFQYAGGLDIVAPGVNVQSTYPGSTYASLNGTSM
 ATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Average Mass = 26715

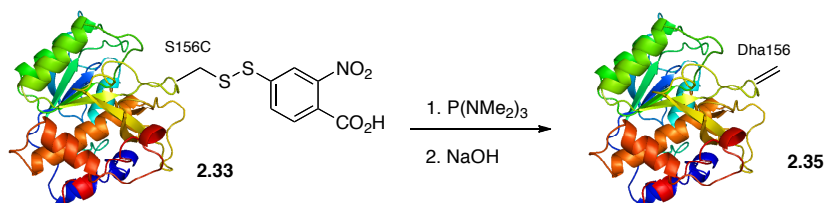
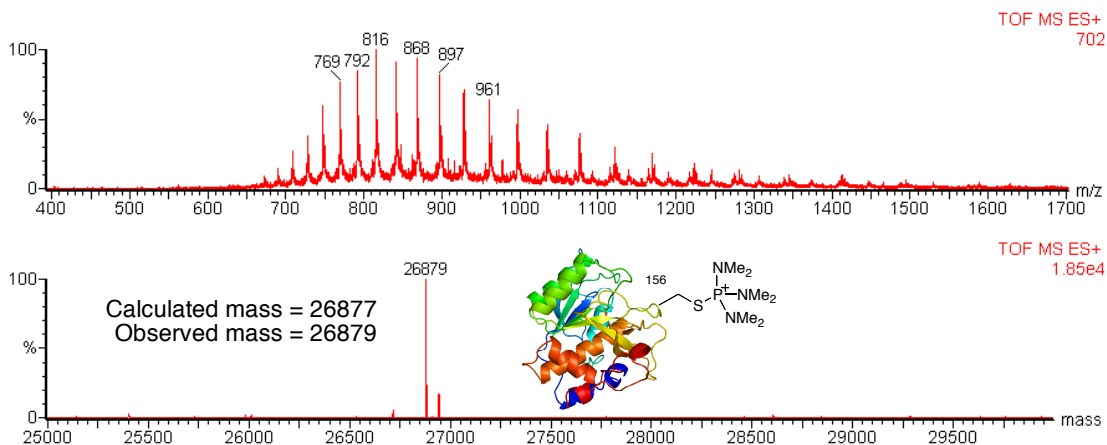


SBL-S156C was prepared as a 1 mg/mL solution in 50 mM potassium phosphate buffer (pH 8.0) and 200 μ L were added to a 1.5 mL plastic tube. In a separate tube, Ellman's reagent (1.2 mg, 3.0 μ mol) was dissolved in 200 μ L of the same buffer. A 50 μ L aliquot of the Ellman's solution was added to the protein and the reaction vortexed to homogenize. The resulting yellow solution was then shaken at 4 $^{\circ}$ C for 20 minutes. LC-MS analysis revealed full conversion to the Ellman adduct **2.33** (calculated mass = 26911; observed mass = 26911).



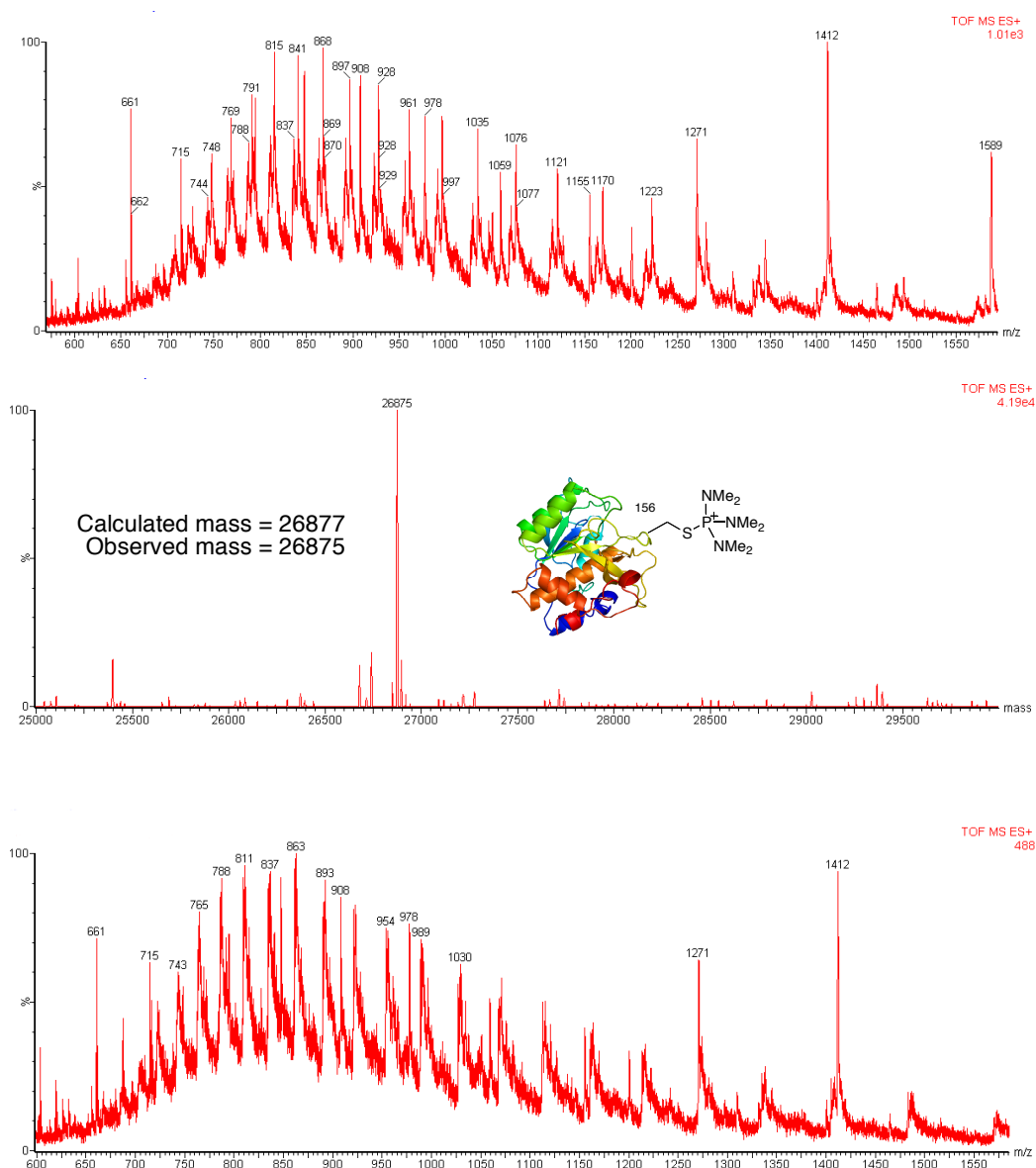


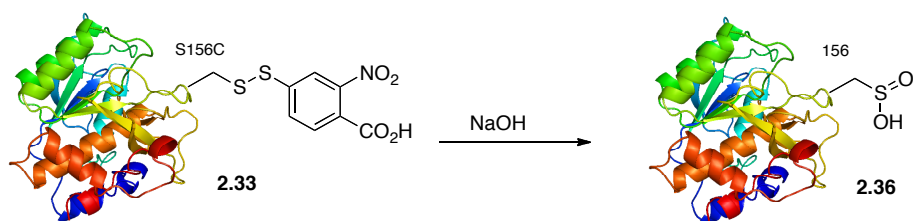
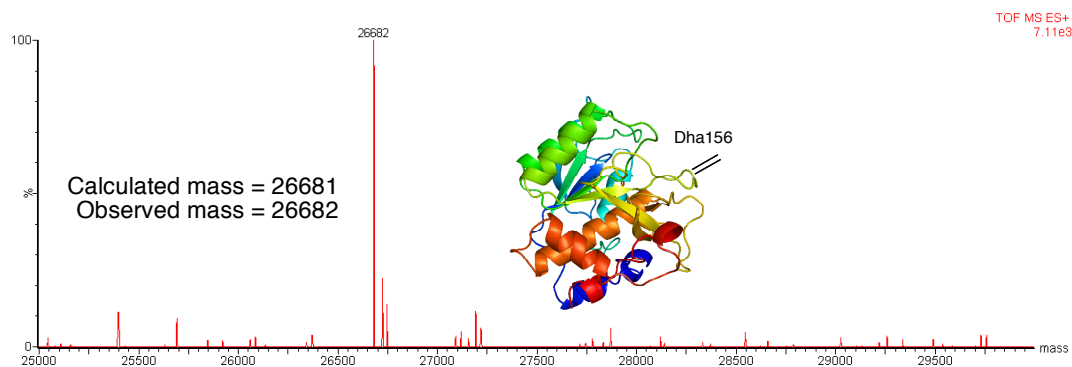
The SBL-S156C adduct of Ellman's reagent (**2.33**) was purified by dialysis twice against 4 L of 50 mM potassium phosphate (pH 8.0). HMPT (5 μL of a 0.06 mM solution in DMF, 0.3 μmol) was added to a 150 μL aliquot of the protein solution (0.3 mg/mL, 0.002 μmol). The reaction was shaken for 15 min at 4 $^\circ\text{C}$ and then analyzed by LC-MS. The phosphonium salt **2.34** was observed as the major product (calculated mass = 26877; observed mass = 26879).



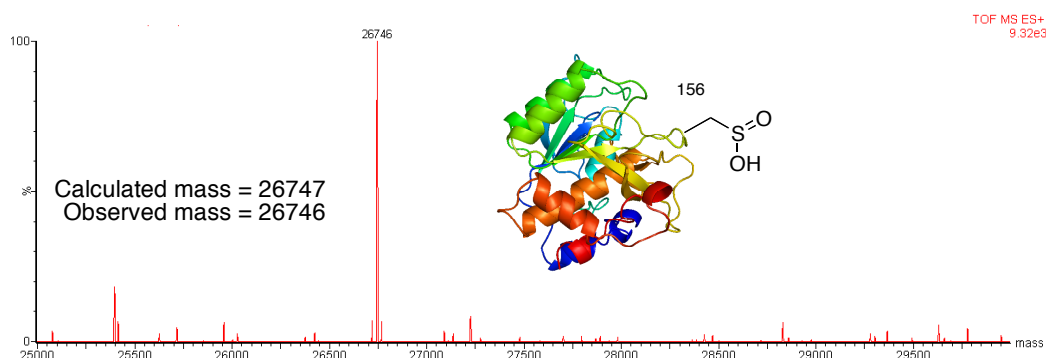
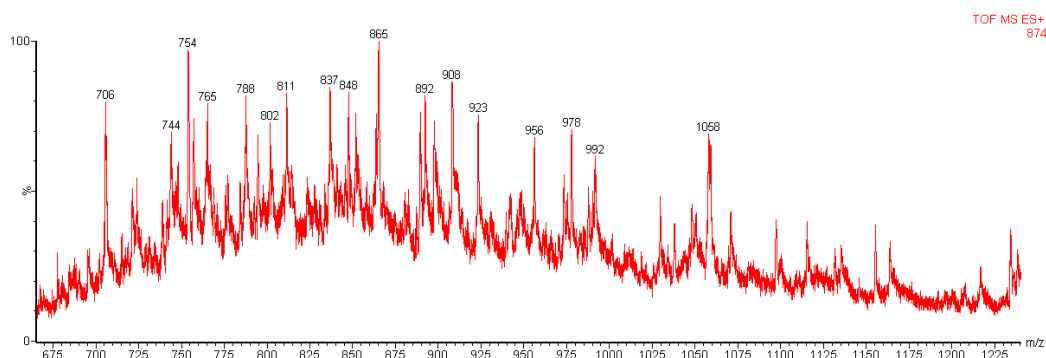
A 200 μL aliquot of SBL-S156C Ellman disulfide (**2.33**) (0.29 mg/mL in pH 8.0 sodium phosphate, 50 mM, 0.002 μmol) was added to a 1.0 mL plastic tube. Na_2CO_3 (5 mg) was added to the protein solution and the mixture vortexed to dissolve. A stock solution of

HMPT was prepared by dissolving 2 μL in 250 μL of DMF. 25 μL of the HMPT solution was added to the protein solution and the reaction shaken at 4 $^{\circ}\text{C}$ for 15 minutes. LC-MS analysis of the reaction mixture revealed full conversion to the HMPT phosphonium adduct **2.34** (calculated mass = 26877; observed mass = 26875). NaOH (50 μL of a 1.0M solution) was then added and then reaction shaken for 20 minutes at 4 $^{\circ}\text{C}$. LC-MS analysis revealed elimination to dehydroalanine (**2.35**) (calculated mass = 26681, observed mass = 26682). ESI-MS are shown below. Above: HMPT adduct. Below: SBL-156Dha

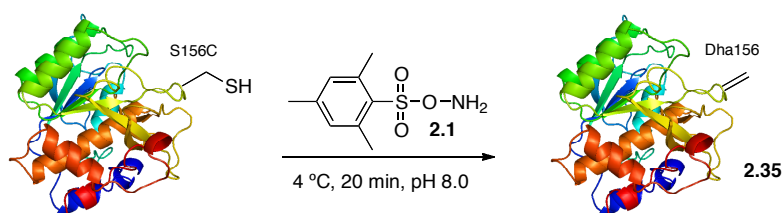




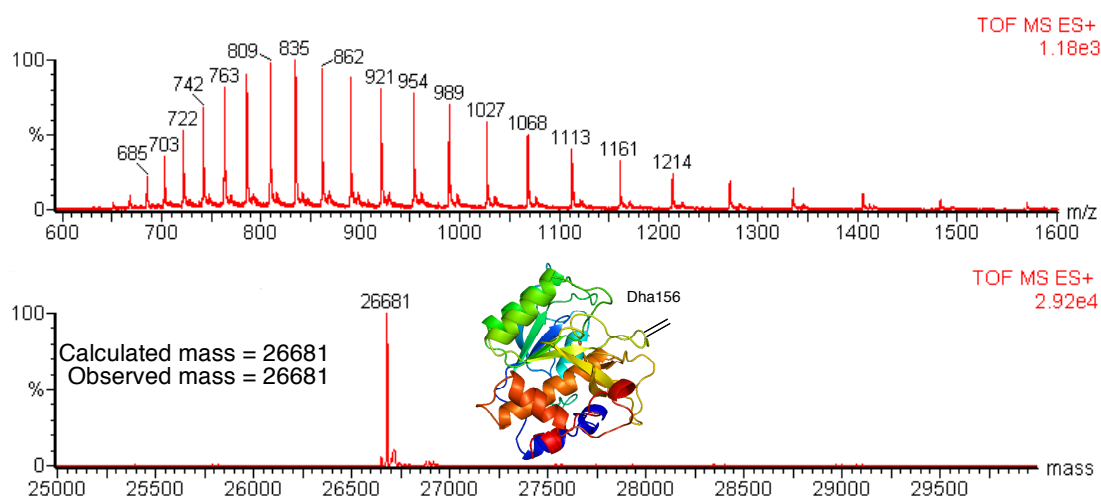
A 50 μ L aliquot of SBL-S156C Ellman disulfide (0.29 mg/mL in pH 8.0 sodium phosphate, 50 mM, 0.5 nmol) was added to a 1.0 mL plastic tube. To this solution was added 50 μ L of 1.0 M NaOH. The mixture was shaken for 15 minutes at 4 $^{\circ}$ C and then analyzed by LC-MS. No dehydroalanine was generated, only conversion to a protein with a mass corresponding to the sulfenic acid was observed (calculated mass = 26747, observed mass = 26746). ESI-MS are shown below.



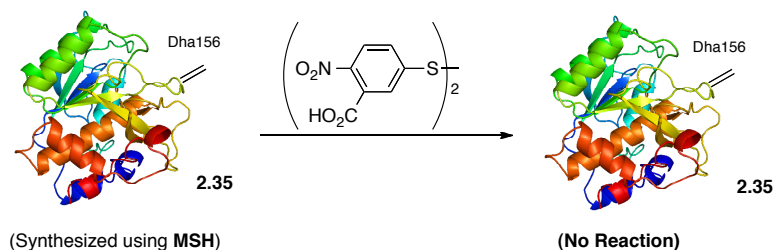
Synthesis of Dha on SBL-S156C using MSH



SBL-S156C was prepared as a 1.4 mg/mL solution in 50 mM sodium phosphate buffer (pH 8.0) and 1.00 mL (0.052 μ mol) was added to a 1.5 mL plastic tube. MSH (1.2 mg, 5.6 μ mol) was added as a solution in DMF (50 μ L) and the reaction vortexed immediately upon addition. The homogenized sample was shaken at 4 °C for 20 minutes before a 40 μ L aliquot was analyzed by LC-MS. A mass corresponding to SBL-Dha156 was observed (calculated mass = 26681; observed mass = 26681). This protein does not react with Ellman's reagent and a variety of thiols can be added by conjugate addition (see next). ESI-MS are shown below.

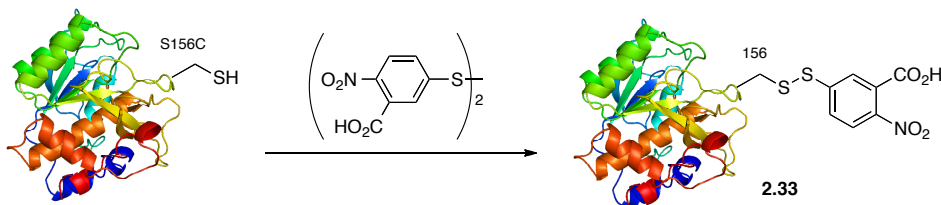
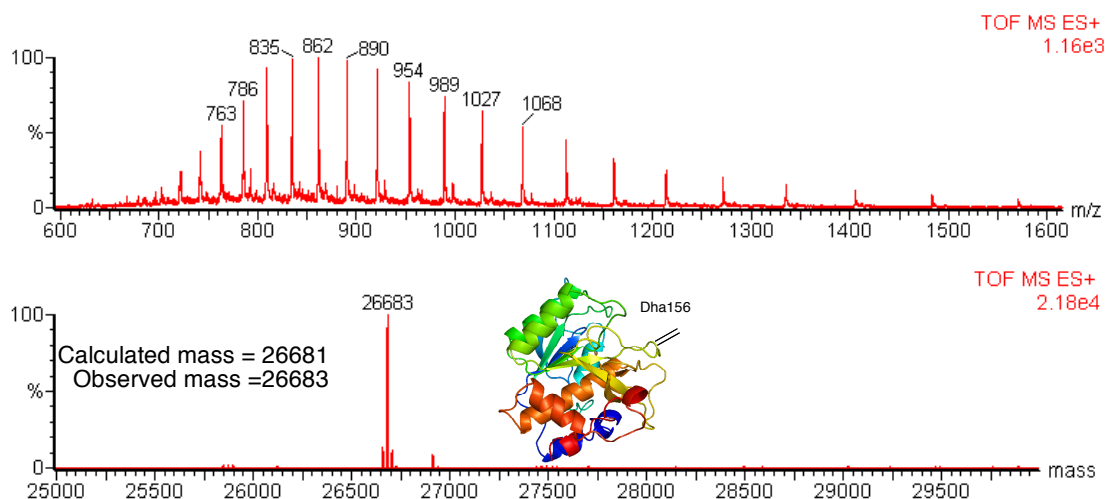


Ellman's Test to Confirm Reaction at Cysteine:

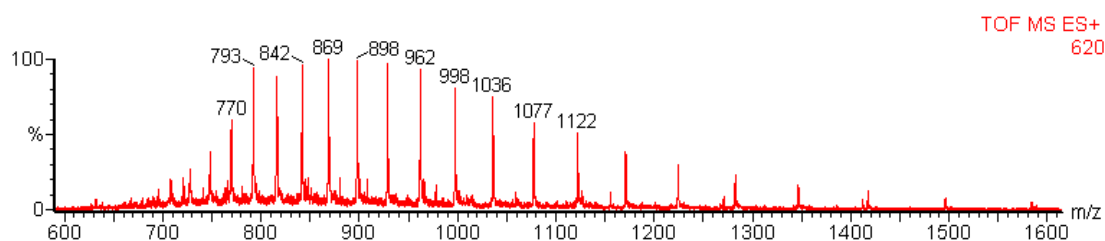


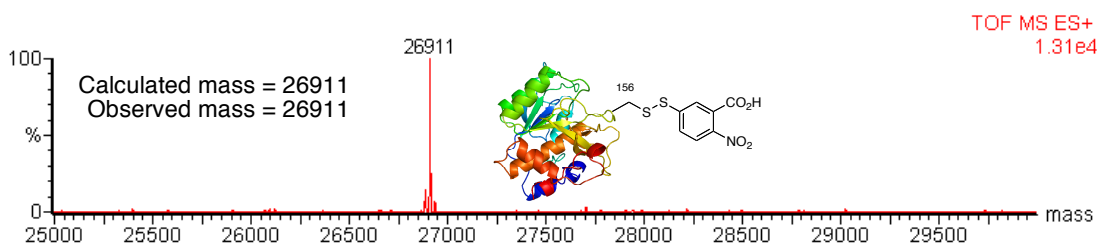
The SBL-Dha156 prepared above was diluted to 0.7 mg/mL with sodium phosphate buffer (50 mM, pH 8.0). Ellman's reagent (0.52 mg, 1.3 μ mol, 250 eq.) was added as a solid to a 200 μ L aliquot of the protein solution. The solution was vortexed and placed on

a lab rotisserie for 20 min at room temperature. LC-MS analysis of a 50 μ L aliquot revealed no reaction, indicating absence of free thiol and that all cysteine was consumed in the reaction (calculated mass = 26681; observed mass = 26683). When SBL-S156C is treated with Ellmna's reagent, the protein is converted to the corresponding disulfide (see next). ESI-MS are shown below.



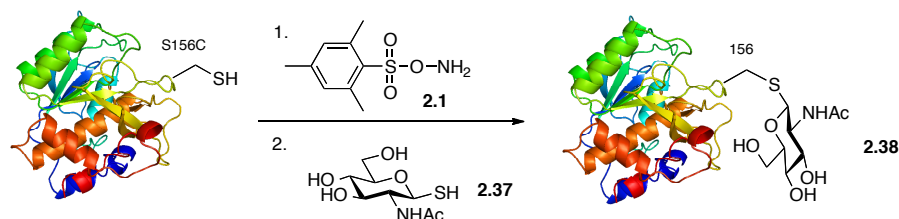
SBL-S156C was prepared as a 1 mg/mL solution in 50 mM potassium phosphate buffer (pH 8.0) and 200 μ L were added to a 1.5 mL plastic tube. In a separate tube, Ellman's reagent (1.2 mg, 3 μ mol) was dissolved in 200 μ L of the same buffer. A 50 μ L aliquot of the Ellman's solution was added to the protein and the reaction vortexed to homogenize. The resulting yellow solution was then shaken at 4 $^{\circ}$ C for 20 minutes. LC-MS analysis revealed full conversion to the Ellman adduct **2.33** (calculated mass = 26911; observed mass = 26911).



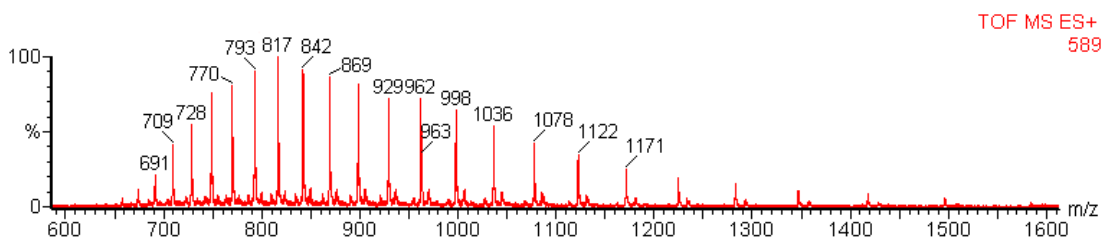


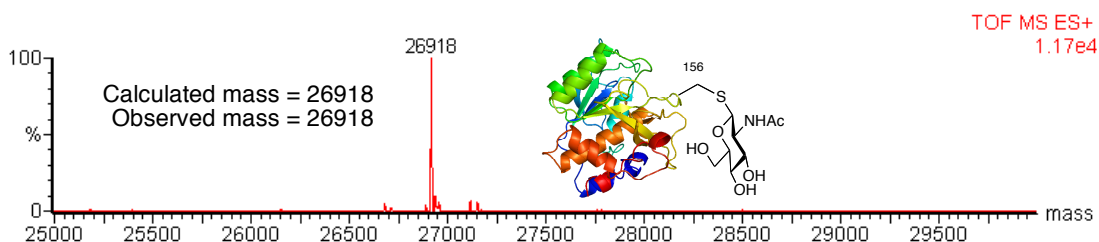
Chemical Test for Dehydroalanine: Conjugate Addition of Nucleophiles

Synthesis of GlcNAc cysteine on SBL at position 156 (**2.38**)



All manipulations were carried out in a cold room maintained at 4 °C. A 1 mg sample of lyophilized SBL-S156C (0.037 μ mol) was dissolved in 1.00 mL in pH 8.0 sodium phosphate buffer (50 mM). A solution of MSH was prepared by dissolving 1.8 mg (8.4 μ mol) in 100 μ L DMF. A 50 μ L portion of the MSH solution was added to the protein by micropipette. The reaction was vortexed periodically over 1 minute and then rotated on a lab rotisserie for an additional 19 minutes at 4 °C. A 50 μ L aliquot was analyzed by LC-MS to confirm the conversion of Cys156 to Dha156 (26681 calculated, 26681 found). To the reaction mixture was added GlcNAc thiol **2.37** as a solid (8.8 mg, 1000 eq.) to give a 39 mM solution in thiol. After 90 minutes of shaking at 4 °C, the reaction was analyzed directly by LC-MS. Complete conversion to SBL-156SGlcNAc **2.38** was observed (calculated mass = 26918, observed mass = 26918). The protein was purified by dialysis (2 $\times$ against 4L pH 7.0 sodium phosphate, 50 mM). Protein recovery after dialysis was 0.8 mg (80%, BCA assay). ESI-MS are shown below.



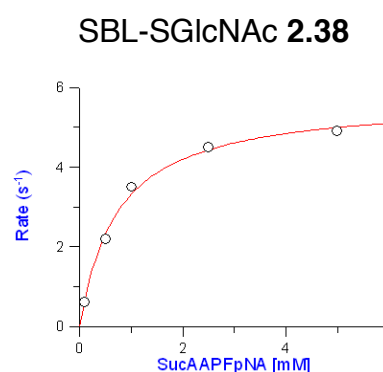
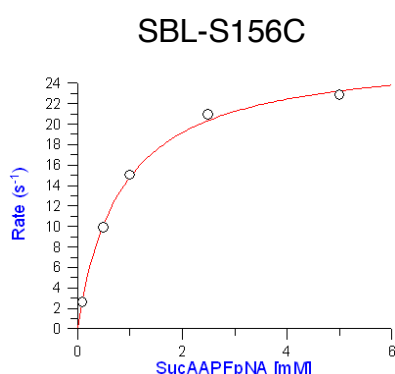


Enzyme kinetics for modified protein 2.38

This experiment was carried out with the assistance of Dr. James Errey.

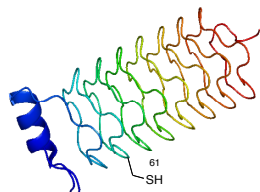
The enzyme concentration was determined using the bicinchoninic acid protein assay (Pierce) with bovine serum albumin as a standard. Turnover numbers are based on an enzyme monomer. Initial velocities for SBL-S156C and SBL-SGlcNAc **2.38** were determined using sucAAPFpNA (Bachem Biosciences Inc) with continuous detection of the formation of the product pNA at 410 nm (pNA: $\epsilon = 8,800 \text{ M}^{-1} \text{ cm}^{-1}$) at 25°C. A typical reaction mixture contained 100 mM sodium phosphate, pH 7.5, 500 mM NaCl, 1 mM suc-AAPF-pNA in a final volume of 1 ml. Reactions were initiated by the addition of enzyme, typically 15 nM final concentration. Initial velocity kinetic data were fitted using GraFit 5. Plots and kinetic parameters are shown below.

Protein	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{M}^{-1} \text{s}^{-1}$)
SBL-S156C	0.83 ± 0.07	27.1 ± 0.7	$(3.3 \pm 0.1) \times 10^4$
SBL-SGlcNAc 2.38	0.72 ± 0.09	5.7 ± 0.2	$(7.9 \pm 0.3) \times 10^3$



Reaction of Np276 with MSH

The following Np276 mutant was expressed and purified by Matthew Young. For the expression and purification of Np276 and further use of this model protein, the reader is referred to the published procedures.³⁰



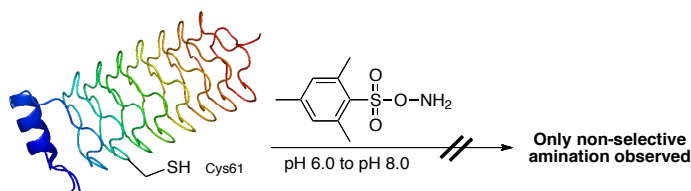
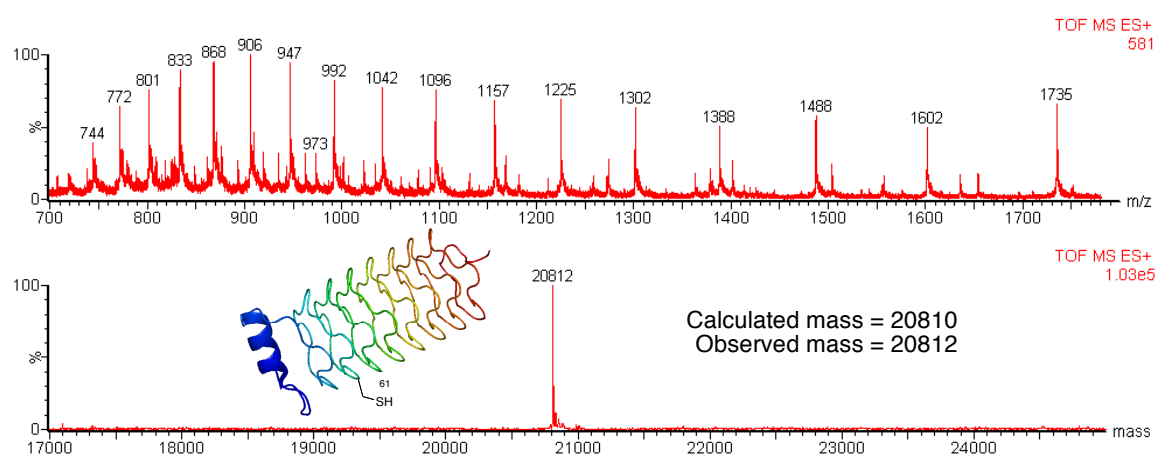
Np276 (PDB = 2J8K)

Sequence:

GSSHHHHHHSSGLVPRGSHIDVGKLRQLYAAGERDFSIVDLRGAVLENINLSGAI
 LHGA⁶¹LDLEANLQQANLSRADLSGATLNGADLRGANLSKADLSDAILDNAILEGA
 ILDEAVLNQANLKAANLEQAILSHANIREADLSEANLEAADLSGADLAIADLHQA
 NLHQAALERANLTGANLEDANLEGTILEGGNNNLAT

Average Mass = 20810

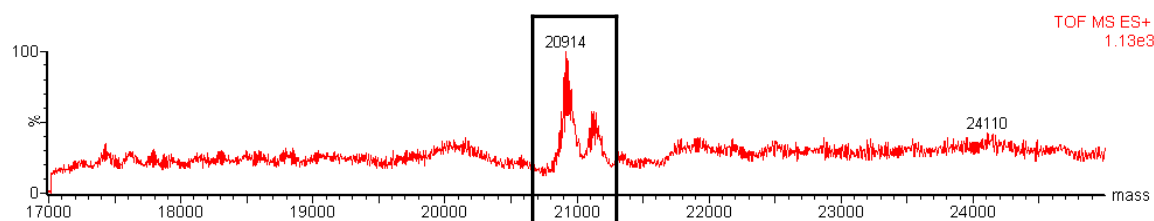
ESI-MS for unmodified Np276 mutant is shown below



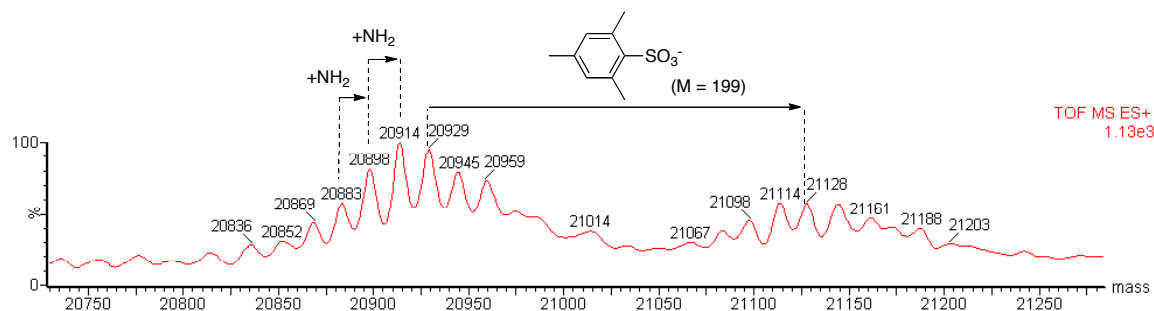
Np276-Cys61 (1.00 mL, c = 0.79 mg/mL, pH 8.0 sodium phosphate) was reduced with DTT (5 mg) and shaken at room temperature for 15 minutes. 500 μ L of the reduced protein was passed through a PDmini column into pH 8.0 sodium phosphate buffer and the remaining 500 μ L was passed through a PDmini column into pH 6.0 sodium phosphate buffer. The proteins solutions were kept on ice. A stock solution of MSH was

prepared by dissolving 3.4 mg in DMF (567 μ L). Two reactions were set up: one at pH 8.0 and one at pH 6.0. Accordingly, 100 μ L of the protein solution was mixed with 5 μ L of the MSH solution at 4 $^{\circ}$ C. The reaction was shaken at 4 $^{\circ}$ C and then analyzed by LC-MS. A mixture of NH_2 adducts was observed at both pH values. Fewer NH_2 groups were observed at pH 6.0, but non-selective amination was still observed. ESI-MS are shown below.

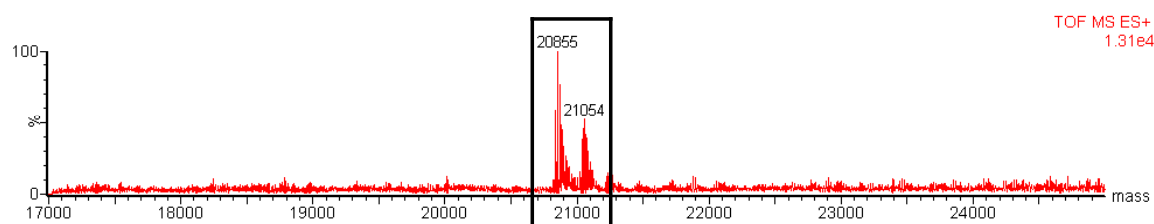
pH 8.0:



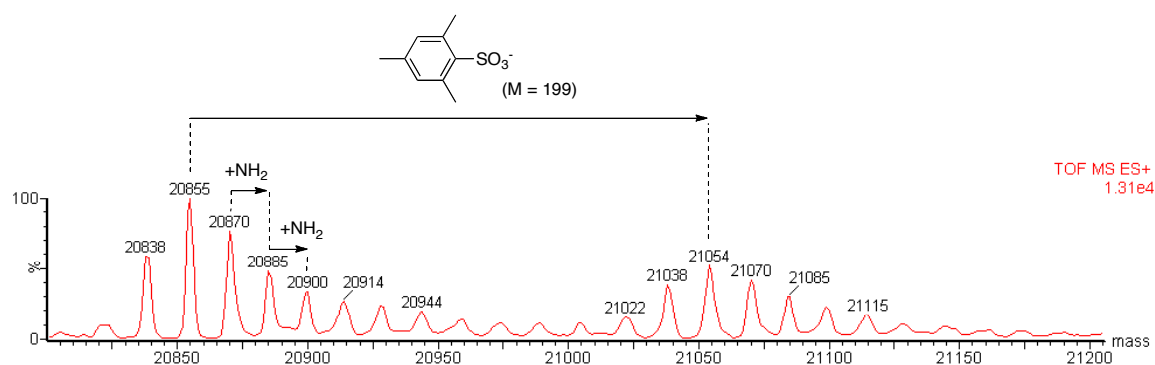
(Magnified region, denoted in box): Representative peaks separated by units of “ NH_2 ” are shown below. Two clusters of peaks are also observed that are separated by the mass of the mesitylene sulfonate anion.



pH 6.0:

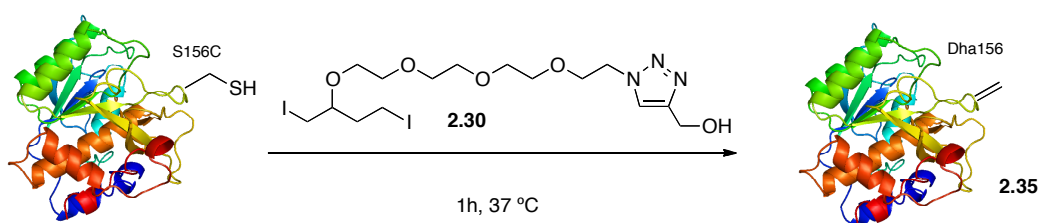
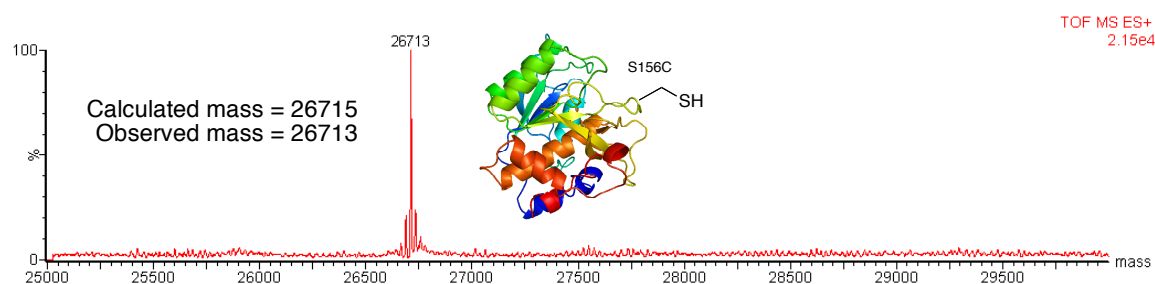
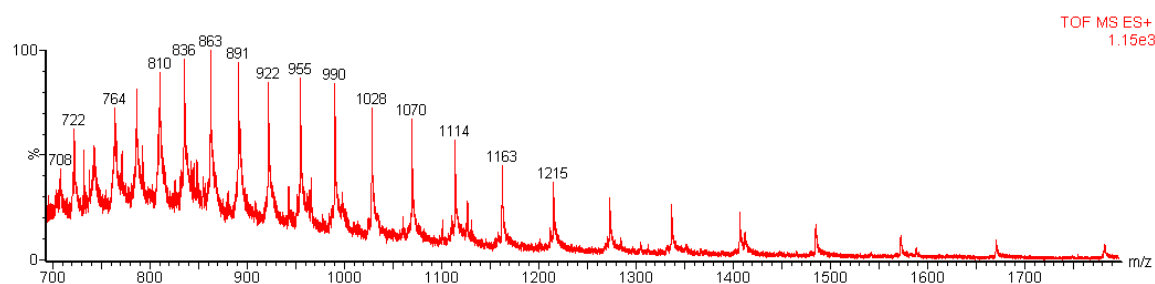


(Magnified region denoted in box): Representative peaks separated by units of “ NH_2 ” are shown below. Two clusters of peaks are also observed that are separated by the mass of the mesitylene sulfonate anion.



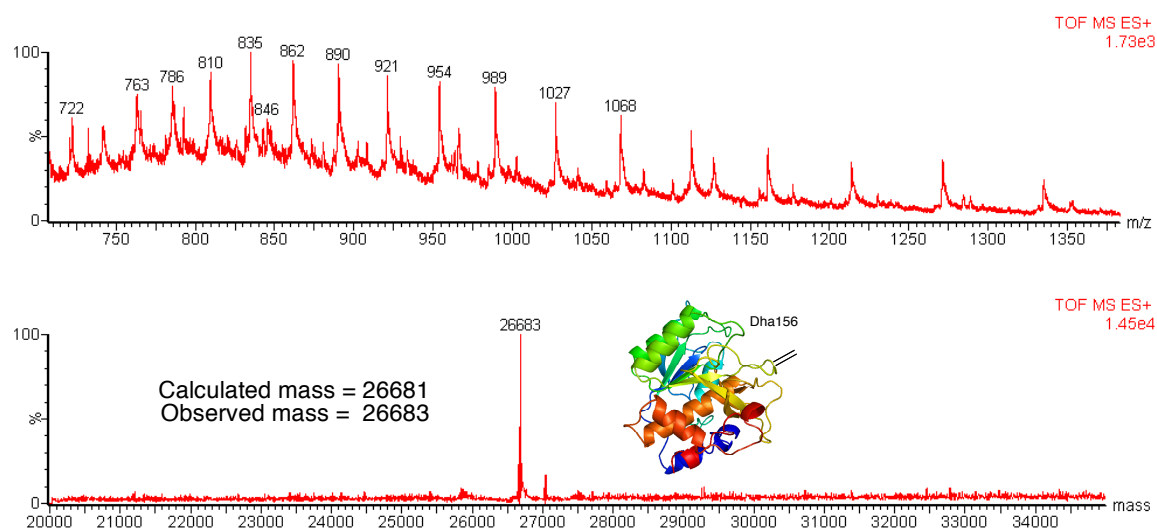
Protein Reactions using Dialkylation-Elimination of Cysteine to Dehydroalanine

SBL-S156C was prepared at 2.0 mg/mL in pH 8.0 sodium phosphate buffer (50 mM) and reduced by the addition of DTT (6 mg). After shaking at room temperature for 20 minutes, the solution was passed through a PD minitrap to remove excess DTT. The protein was eluted with sodium phosphate buffer (pH 8.0, 50 mM). The final concentration was 1.0 mg/mL. The protein was stored on ice until needed. ESI-MS are shown below.

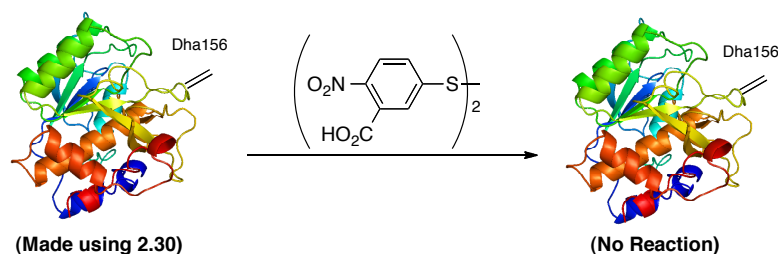


A solution of the di-iodide **2.30** was prepared by dissolving 11 mg (0.019 mmol) in 52.6 μ L of DMF. A 500 μ L solution of SBL-S156C (1 mg / mL in pH 8.0 sodium phosphate buffer, 19 nmol protein) was added to the triazole solution and the mixture was vortexed

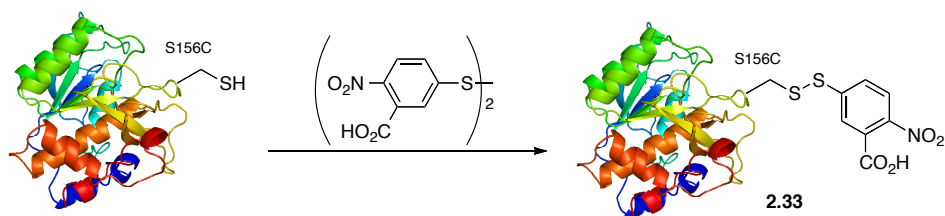
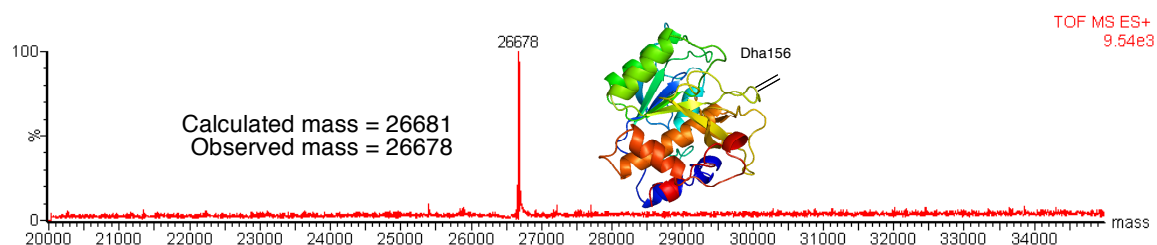
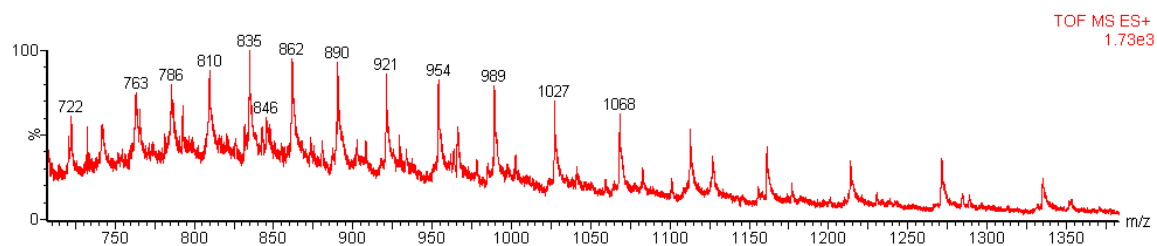
immediately. The reaction was shaken at 37 °C for 1 hour. After this time, any precipitate was removed by centrifugation (2 min, 14K g). The supernatant was analyzed by LC-MS, revealing complete conversion to dehydroalanine (calculated mass = 26681; observed mass = 26683 found). LC-MS is show below.



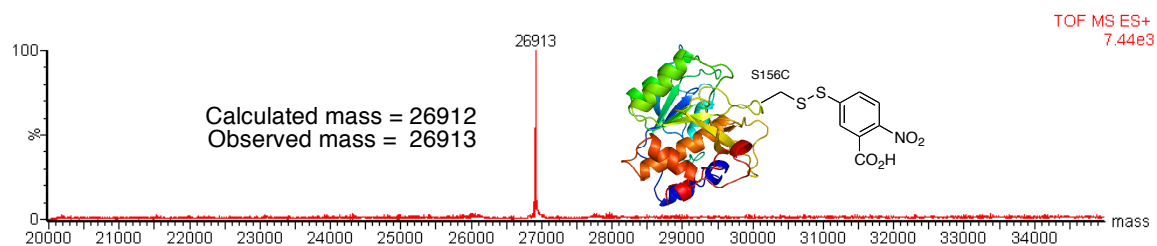
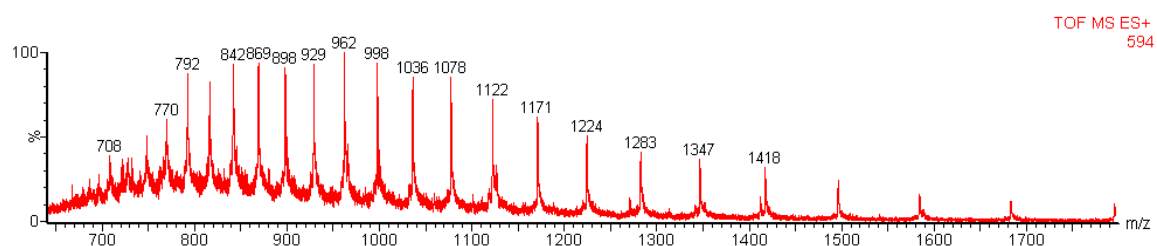
Ellman's Test to Confirm Reaction at Cysteine:



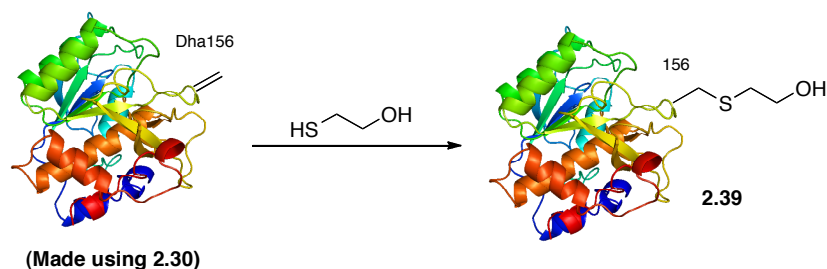
Ellman's solution was prepared by dissolving 0.4 mg in 108 μ L of sodium phosphate buffer (pH 8.0, 50 mM). A 10 μ L aliquot of Ellman's solution was added to a 50 μ L sample of SBL-Dha prepared above using diiodide **2.30** (0.5 mg/mL, pH 8.0 sodium phosphate). The reaction was shaken at room temperature for 15 minutes and then analyzed by LC-MS. No reaction was observed, indicating all cysteine was consumed in the elimination. (calculated mass = 26681; observed mass = 26678). Under identical conditions, SBL-S156C is converted to the Ellman disulfide (see next). ESI-MS are shown below.



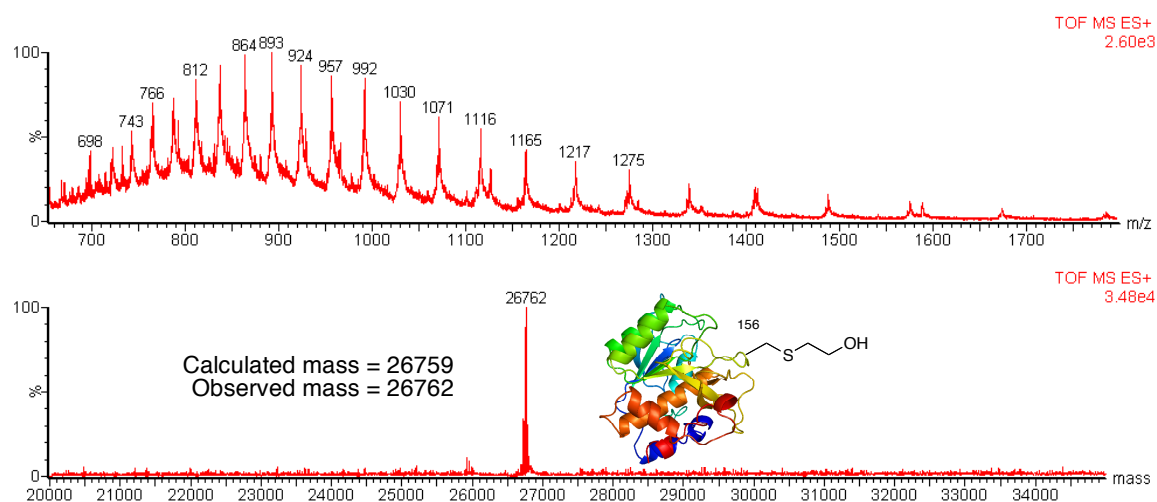
A solution of SBL-S156C was prepared at 0.5 mg/mL in pH 8.0 sodium phosphate buffer (50 mM). A 10 μ L aliquot of the Ellman's solution prepared in the previous reaction was added to 50 μ L of SBL-(S156C). The reaction was shaken at room temperature for 15 minutes and then analyzed by LC-MS. Full conversion to the disulfide was observed (calculated mass = 26912; observed mass = 26913). ESI-MS are shown below.



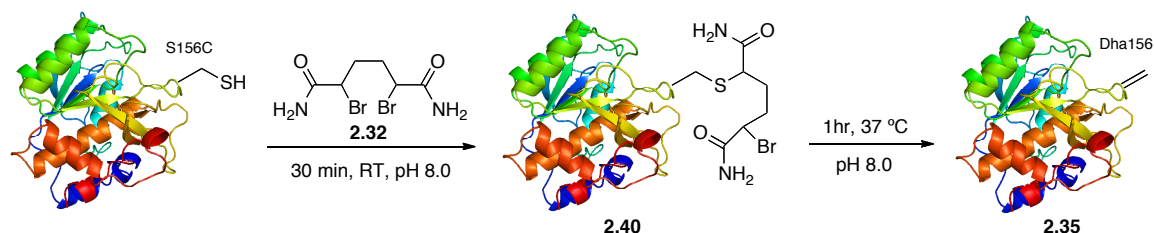
Chemical Test for Dehydroalanine: Addition of 2-Mercaptoethanol



SBL-156Dha was prepared using diiodide **2.30** as described above. A 50 μ L aliquot of the SBL-156Dha was mixed with 3.3 μ L of 2-mercaptoethanol. The reaction was incubated at room temperature for 30 minutes and then analyzed by LC-MS. Full conversion to the thioether **2.39** was observed (calculated mass = 26759; observed mass = 26762). ESI-MS are shown below.



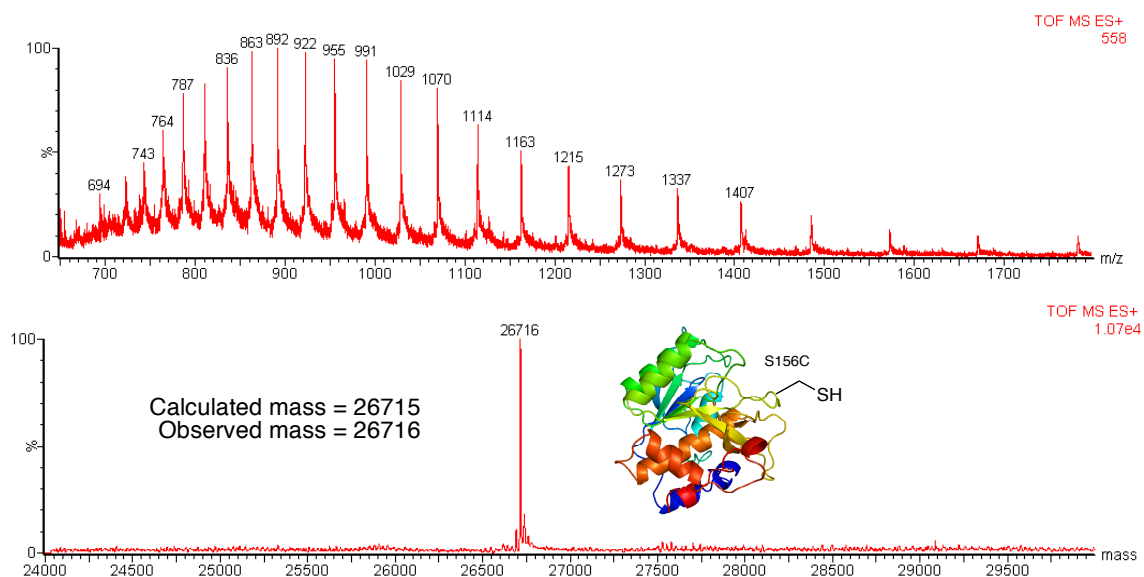
SBL-156Dha synthesis using dibromide 2.32



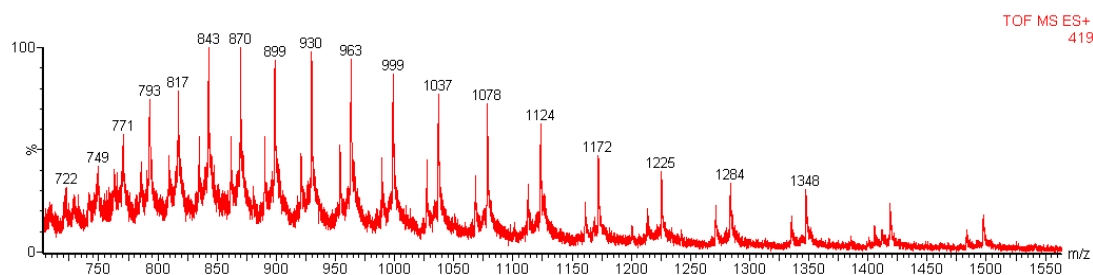
SBL-S156C was prepared at 1.0 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). DTT (1.4 mg) was added as a solid to a 500 μ L aliquot of the protein solution to reduce any contaminant disulfide. After 15 minutes of incubation at room temperature, the DTT was removed by passing the protein solution through a PD minitrapp (GE Healthcare), eluting with 1 mL of sodium phosphate buffer (pH 8.0, 50 mM). A 100 μ L aliquot of the

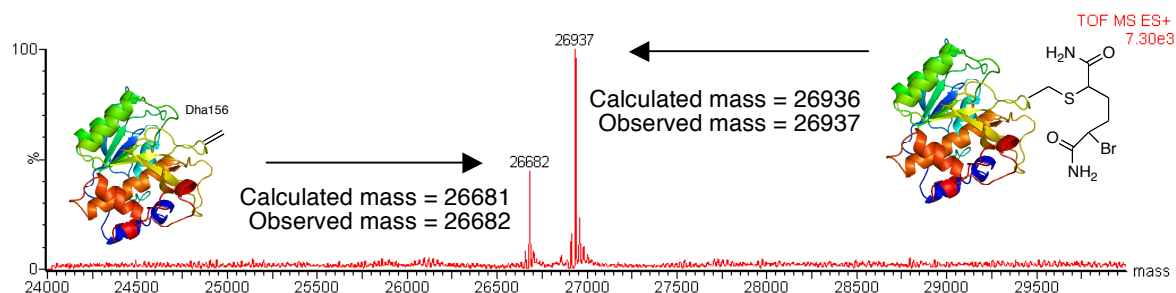
reduced protein (0.5 mg/mL, 0.002 μ mol) was used for the elimination reaction. Accordingly, a stock solution of dibromide **2.32** was prepared by dissolving 35.5 mg in 418 μ L DMF. A 10 μ L aliquot of the di-bromide solution (3 μ mol) was added to the 100 μ L protein solution at room temperature. The reaction was vortexed to homogenize and shaken at room temperature for 30 minutes. After this time, solids were removed by centrifugation (1 min, 14K g). (Note: some di-bromide crystallizes / precipitates from the reaction mixture). The supernatant was analyzed by LC-MS, revealing two products: alkylated, uncyclized intermediate (calculated mass = 26936; observed mass = 26937) and dehydroalanine (calculated mass = 26681; observed mass = 26682). The reaction was then shaken at 37 $^{\circ}$ C for an hour to complete the cyclization and elimination. After this time, solids were removed by centrifugation and the supernatant analyzed by LC-MS. A single protein was detected with a mass corresponding to the dehydroalanine containing protein (calculated mass = 26681; observed mass = 26682). ESI-MS for the entire reaction sequence are shown below.

t = 0 (Starting material SBL-S156C)

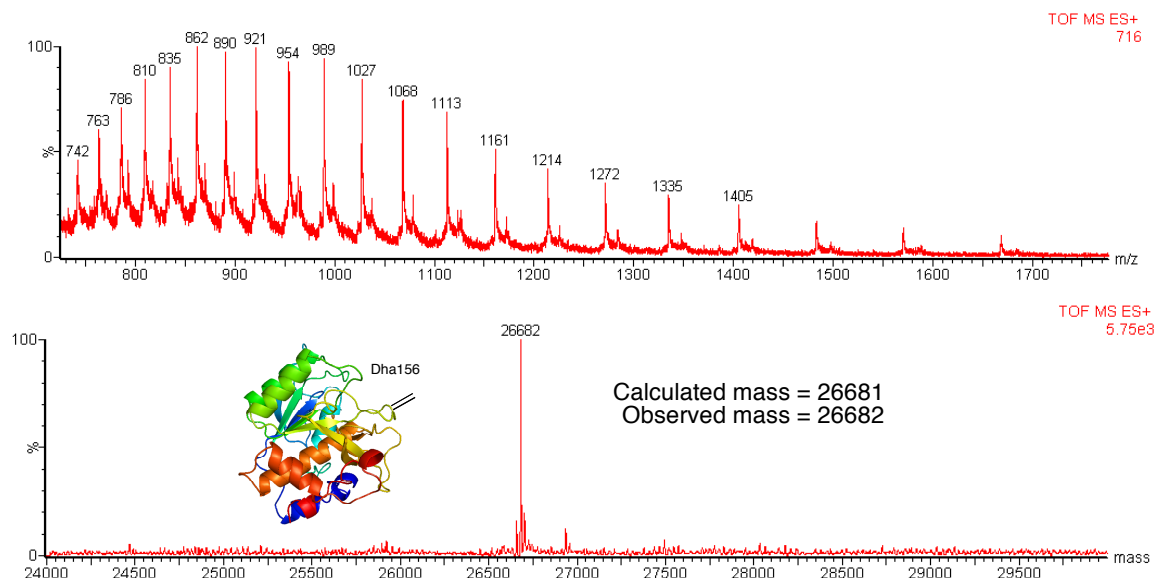


t = 30 min, RT



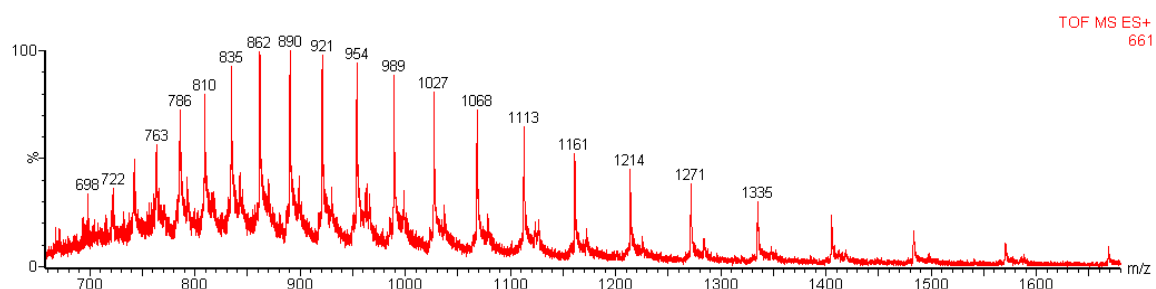


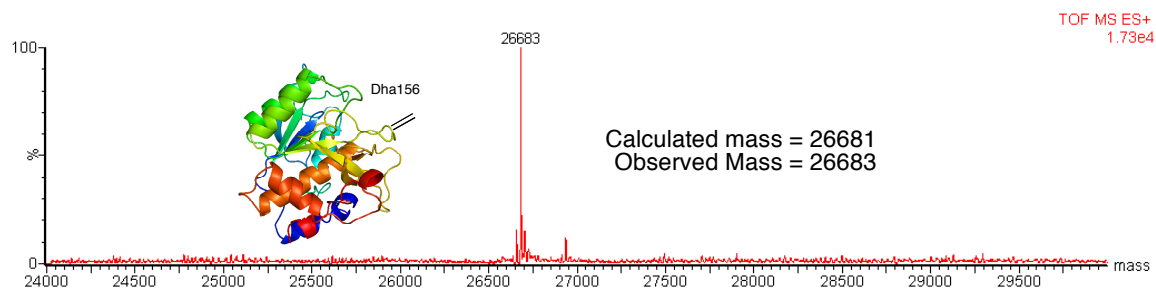
t = 90 min (after 60 min at 37 °C)



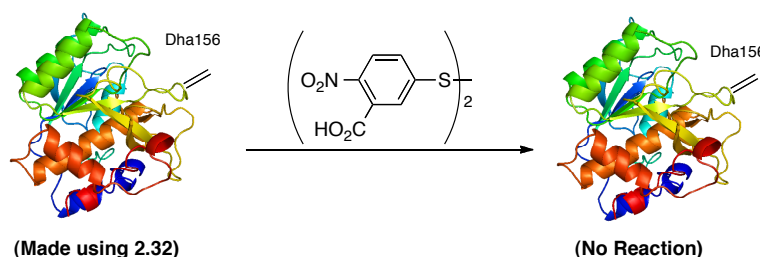
The procedure was repeated to obtain more SBL-Dha:

A 450 μ L aliquot of SBL-S156C (c = 0.5 mg/mL, pH 8.0 sodium phosphate buffer) was warmed to room temperature. A 45 μ L aliquot of the di-bromide solution prepared above was added to the protein. The reaction was shaken at room temperature for 30 minutes and then at 37 °C for 1 hour. After this time, solids were removed by centrifugation (1 min, 14K g) and the solid was passed through a PD minitrap, eluting with 1 mL of the reaction buffer (pH 8.0 sodium phosphate, 50 mM). LC-MS analysis of the product revealed a mass corresponding to the dehydroalanine containing protein (calculated mass = 26681; observed mass = 26682). ESI-MS are shown below.

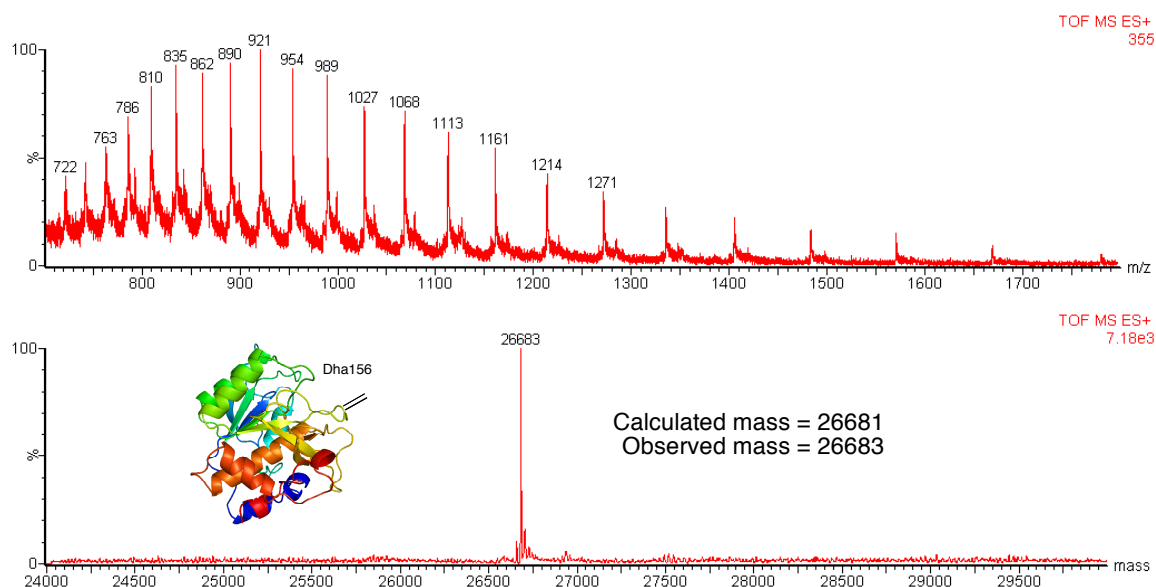


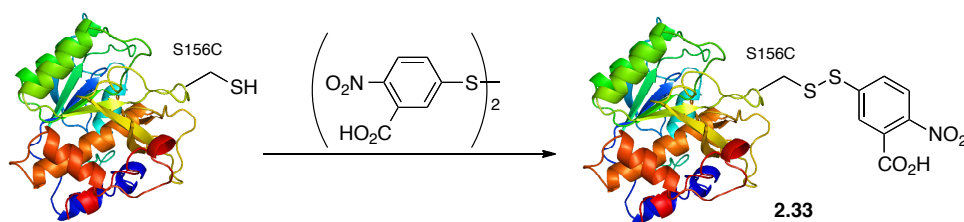


Ellman's Test to Confirm Reaction at Cysteine:

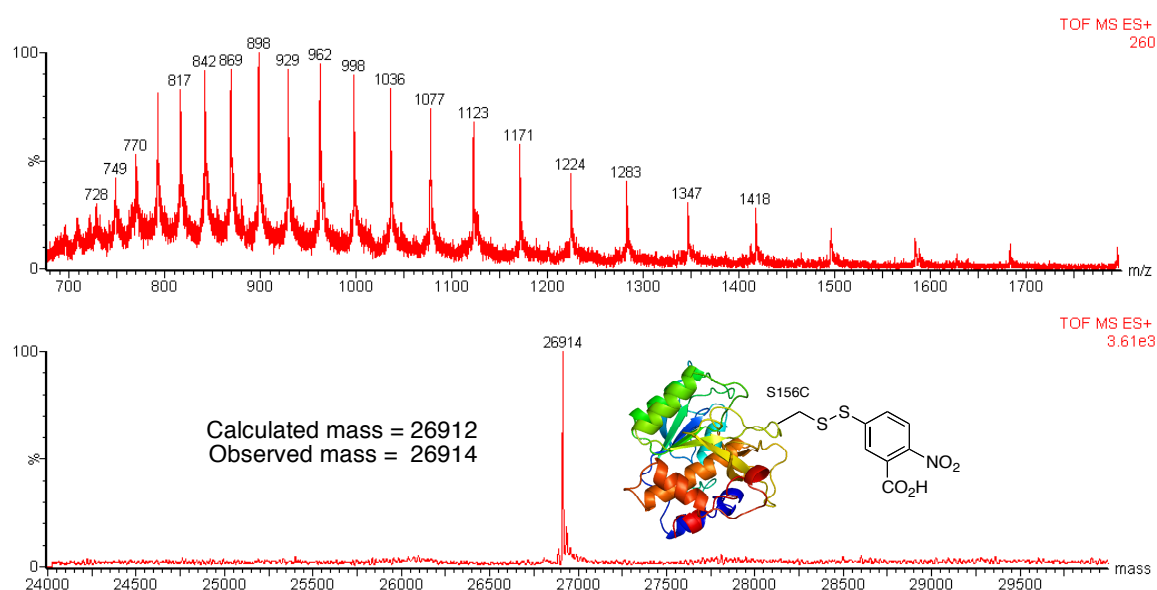


Ellman's assay was carried out to verify consumption of free thiol, and hence reaction at cysteine. Accordingly, a stock solution of Ellman's reagent was prepared by dissolving 0.5 mg in 10.2 mL of sodium phosphate buffer (pH 8.0, 50 mM). 2 μ L of the Ellman's solution was added to a 30 μ L aliquot of the SBL-Dha solution (0.22 mg/mL) prepared above. The solution was shaken for 15 minutes at room temperature and then analyzed by LC-MS. No reaction was observed (calculated mass = 26681; observed mass = 26683). Under the same reaction conditions, SBL-S156C is converted to the Ellman disulfide (see next). ESI-MS are shown below.

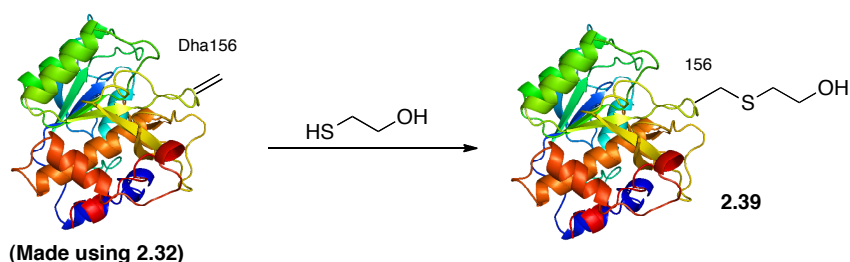




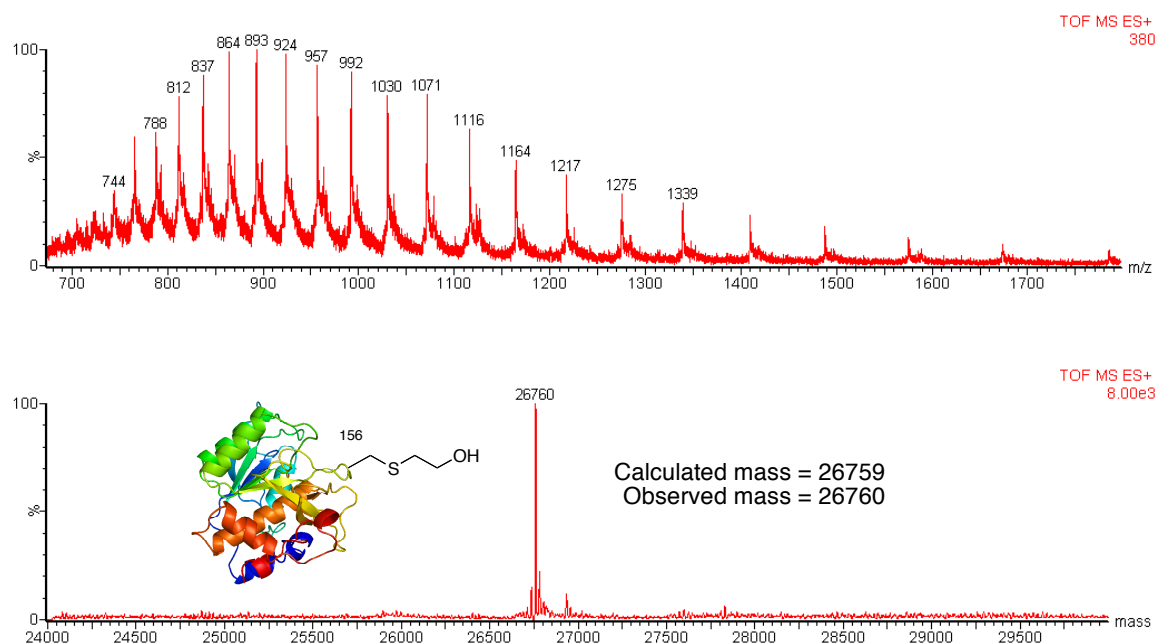
SBL-S156C was prepared at 0.22 mg/mL in pH 8.0 sodium phosphate buffer. A 2 μ L aliquot of the Ellman's reagent solution prepared above was added to 30 μ L of SBL-Dha solution. The reaction was shaken 15 minutes at room temperature, after which time full conversion to the corresponding disulfide was observed (calculated mass = 26912; observed mass = 26914). ESI-MS are shown below.



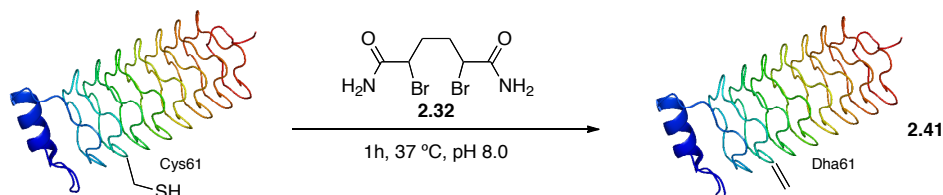
Chemical Test for Dehydroalanine: Addition of 2-Mercaptoethanol



The presence of dehydroalanine was corroborated by the addition of β -mercaptoethanol. Accordingly, a 25 μ L aliquot of the SBL-Dha protein (0.22 mg/mL) was mixed with 0.7 μ L of β -mercaptoethanol. The reaction was shaken at room temperature for 30 minutes and then analyzed by LC-MS. Full conversion to the corresponding thioether was observed (calculated mass = 26759; observed mass = 26760).

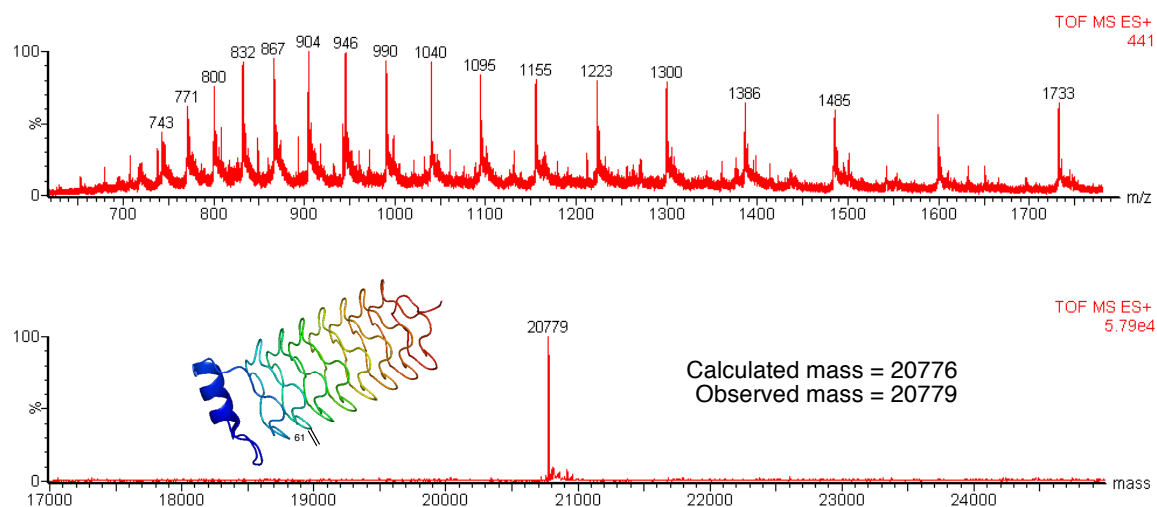


Dha formation on Np276 using dibromide **2.32**

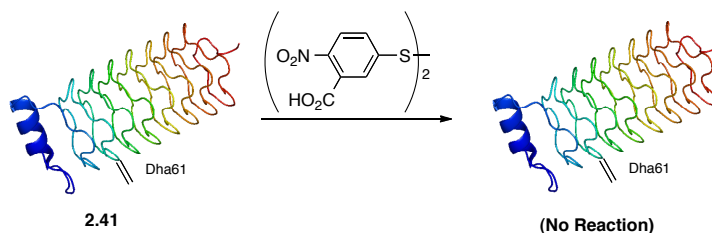


Np276-Cys61 (2.5 mL, 1.1 mg / mL in pH 8.0 sodium phosphate, 50 mM) was reduced by the addition of 10 mg DTT. The solution was shaken at room temperature for 15 minutes to reduce any contaminant disulfide. After this time, the protein solution was passed through a PD10 column (GE Healthcare), eluting with the same buffer. The resulting protein solution was used immediately. Accordingly, 2.5 mL of the NP276 protein ($c = 0.79$ mg/mL) was added to a 15 mL plastic tube and kept at room temperature. A stock solution of dibromide **2.32** was prepared by dissolving 218 mg in 1.91 mL DMF. A 125 μ L aliquot of the dibromide solution was added to the protein solution and the reaction was vortexed to homogenize and immediately placed in a 37 °C incubator with shaking. After 15 minutes, an additional 125 μ L aliquot of the di-bromide solution was added to the reaction. After another 15 minutes, a third and final 125 μ L aliquot of the di-bromide solution was added. The reaction was shaken at 37 °C for a final 30 minutes and then analyzed by LC-MS. A protein mass corresponding to the Dha protein **2.41** was observed (calculated mass = 20776; observed mass = 20779).

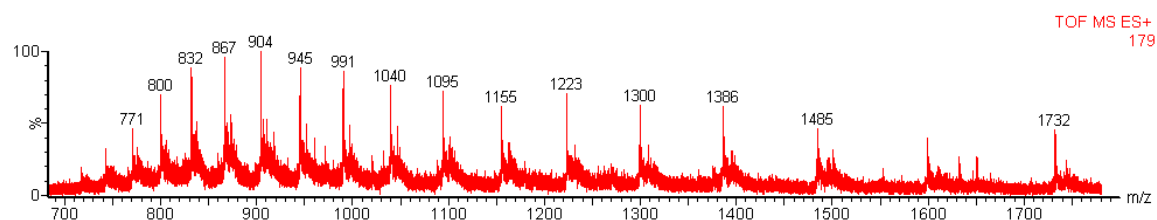
Precipitated dibromide was removed by centrifugation and the protein solution was stored on ice. ESI-MS are shown below.

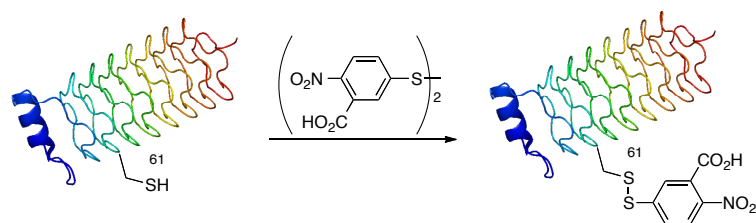
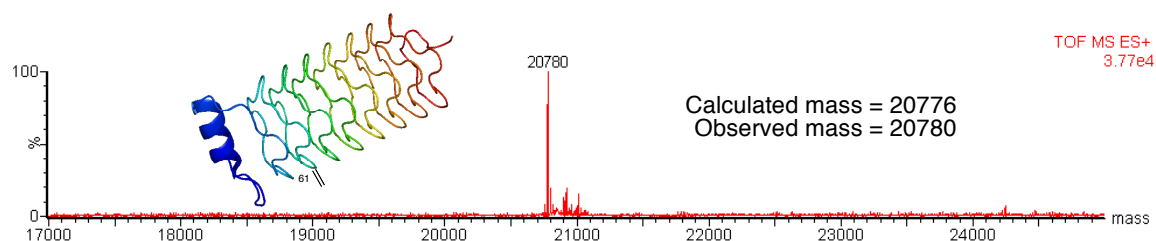


Ellman's Test to Confirm Reaction at Cysteine:

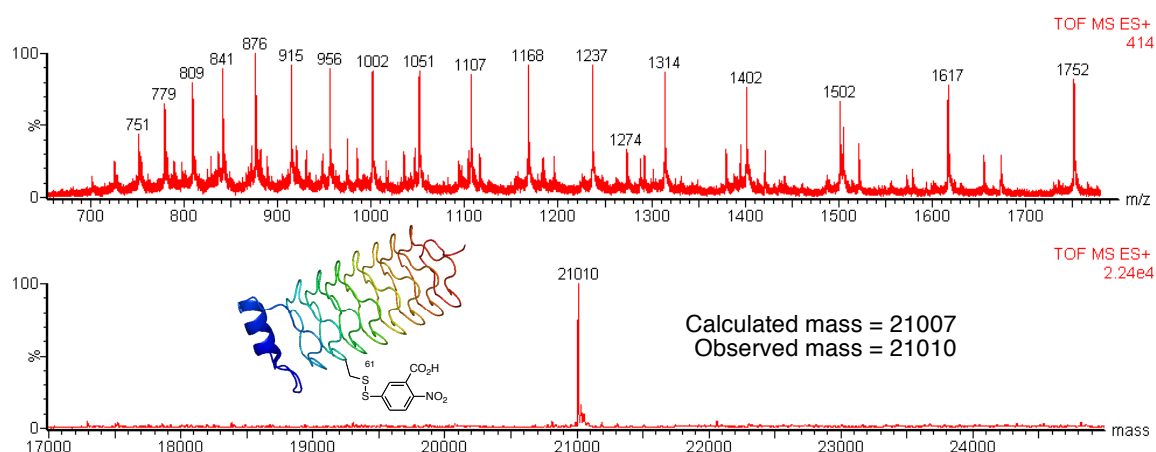


A stock solution of Ellman's reagent was prepared by dissolving 0.5 mg in 250 μ L of sodium phosphate buffer (pH 8.0, 50 mM). A 140 μ L aliquot of the dehydroalanine containing protein **2.41** (0.79 mg/mL in pH 8.0 sodium phosphate buffer) was mixed with 10 μ L of the Ellman's solution. The solution was incubated at room temperature for 15 minutes and then analyzed by LC-MS. No reaction was observed, indicating complete consumption of cysteine during the elimination reaction (calculated mass = 20776 calculated; observed mass = 20780). ESI-MS are shown below. Under identical conditions, Np276-Cys61 is converted to the Ellman disulfide (see next).

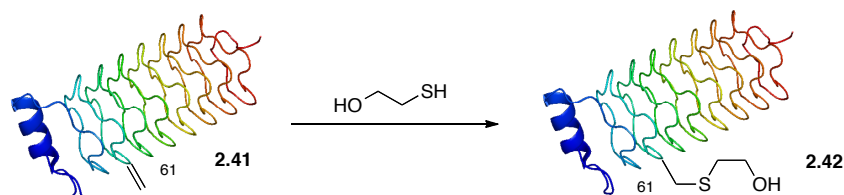




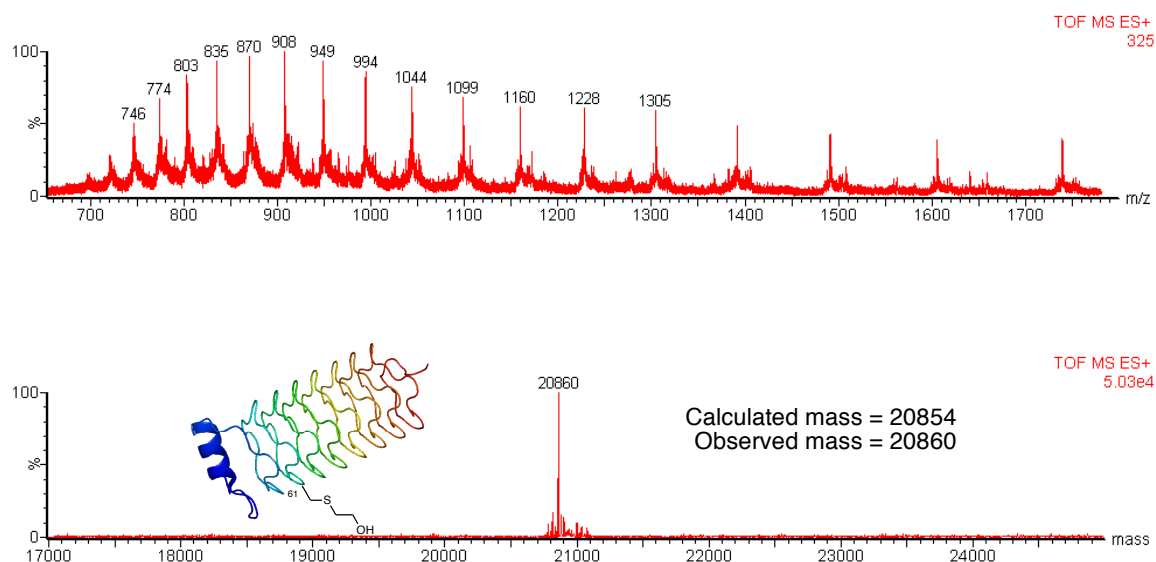
Np276-Cys61 (140 μ L of a 0.79 mg/mL solution in pH 8.0 sodium phosphate buffer, 50 mM) was mixed with 10 μ L of the Ellman's reagent solution prepared above. The reaction was shaken at room temperature for 15 minutes and then analyzed by LC-MS. Full conversion to the Ellman disulfide was observed (calculated mass = 21007; observed mass = 21010). ESI-MS are shown below.



Chemical Test for Dehydroalanine: Addition of 2-Mercaptoethanol



A 100 μ L aliquot of the Dha containing Np276 was mixed with 5 μ L of 2-mercaptoethanol. The reaction was shaken at 37 $^{\circ}$ C for 1 hour. After this time, full conversion to the thioether **2.42** was observed by LC-MS (calculated mass = 20854; observed mass = 20860). ESI-MS are shown below.

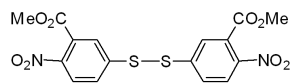


2.6 References for Experimental Procedures

- (1) Aroyan, C. E.; Miller, S. J. *J. Am. Chem. Soc.* **2007**, *129*, 256-257.
- (2) Adams, L. A.; Aggarwal, V. K.; Bonnert, R. V.; Bressel, B.; Cox, R. J.; Shepherd, J.; de Vicente, J.; Walter, M.; Whittingham, W. G.; Winn, C. L. *J. Org. Chem.* **2003**, *68*, 9433-9440.
- (3) Threadgill, M. D.; Gledhill, A. P. *J. Org. Chem.* **1989**, *54*, 2940-2949.
- (4) Johnson, C. R.; Kirchhoff, R. A.; Corkings, H. G. *J. Org. Chem.* **1974**, *39*, 2458-2459.
- (5) Kuligowski, C.; Bezzenine-Lafollee, S.; Chaume, G.; Mahuteau, J.; Barriere, J. C.; Bacque, E.; Pancrazi, A.; Ardisson, J. *J. Org. Chem.* **2002**, *67*, 4565-4568.
- (6) Vaswani, R. G.; Chamberlin, A. R. *J. Org. Chem.* **2008**, *73*, 1661-1681.
- (7) Lassen, K. M.; Lee, J.; Joullié, M. M. *Tetrahedron Lett.* **2010**, *51*, 1635-1638.
- (8) Capasso, S.; Mazzarella, L.; Sica, F.; Zagari, A. *Int. J. Pept. Prot. Res.* **1989**, *33*, 124-132.
- (9) Crosignani, S.; White, P. D.; Linciau, B. *J. Org. Chem.* **2004**, *69*, 5897-5905.
- (10) Vogler, K.; Studer, R. O.; Lergier, W.; Lanz, P. *Helv. Chim. Acta.* **1965**, *48*, 1407-1414.
- (11) Groselj, U.; Bevk, D.; Jakse, R.; Meden, A.; Pirc, S.; Recnik, S.; Stanovnik, B.; Svete, J. *Tetrahedron: Asymmetry* **2004**, *15*, 2367-2383.
- (12) Chankeshwara, S. V.; Chakraborti, A. K. *Org. Lett.* **2006**, *8*, 3259-3262.

-
- (13) Oswald, C. L.; Carrillo-Márquez, T.; Caggiano, L.; Jackson, R. F. W. *Tetrahedron* **2008**, *64*, 681-687.
- (14) Qing, G.; He, Y.; Wang, F.; Qin, H.; Hu, C.; Yang, X. *Eur. J. Org. Chem.* **2007**, 1768-1778.
- (15) Dhaon, M. K.; Olsen, R. K.; Ramasamy, K. *J. Org. Chem.* **1982**, *47*, 1962-1965.
- (16) Abdo, M.-R.; Joseph, P.; Boigegrain, R.-A.; Liautard, J.-P.; Montero, J.-L.; Köhler, S.; Winum, J.-Y. *Bioorg. Med. Chem.* **2007**, *15*, 4427-4433.
- (17) Ishizaki, T.; Hashimoto, Y.; Shudo, K.; Okamoto, T. *Heterocycles* **1983**, *20*, 1481-1485.
- (18) Balajthy, Z. *Org. Prep. Proced. Int.* **1991**, *23*, 375-376.
- (19) Lüning, U.; Kühl, C.; Uphoff, A. *Eur. J. Org. Chem.* **2002**, 4063-4070.
- (20) White, B. D.; Mallen, J.; Arnold, K.A.; Fronczek, F. R.; Gandour, R. D.; Gehrig, M. B.; Gokel, G. W. *J. Org. Chem.* **1989**, *54*, 937-947.
- (21) Black, P. J.; Edwards, M. G.; Williams, J. M. J. *Eur. J. Org. Chem.* **2006**, 4367-4378.
- (22) Shirude, P. S.; Kumar, V. A.; Ganesh, K. N. *Eur. J. Org. Chem.* **2005**, 5207-5215.
- (23) Bertozzi, C. R.; M. D. Bednarski, M. D. *J. Org. Chem.* **1991**, *56*, 4326-4329.
- (24) Yamashita, H.; Mukaiyama, T. *Chem. Lett.* **1985**, 363-366.
- (25) Harpp, D. N.; Smith, R. A.; Steliou, K. *J. Org. Chem.* **1981**, *46*, 2072-2079.
- (26) Treibs, A.; Neumayr, F. *Chem. Ber.* **1957**, *90*, 76-79.
- (27) Meyer, W.; Reckendorf, Z.; Bonner, W. A. *J. Org. Chem.* **1961**, *26*, 4596-4599.
- (28) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 2244-2247.
- (29) Lin, Y. A.; Chalker, J. M.; Davis, B. G. *J. Am. Chem. Soc.* **2010**, *132*, 16805-16811.
- (30) Fernández-González, M.; Bourureira, O.; Bernardes, G.J.L.; Chalker, J.M.; Young, M.A.; Errey, J.C.; Davis, B.G. *Chem. Sci.* **2010**, 709-715.

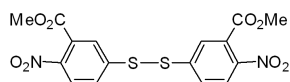
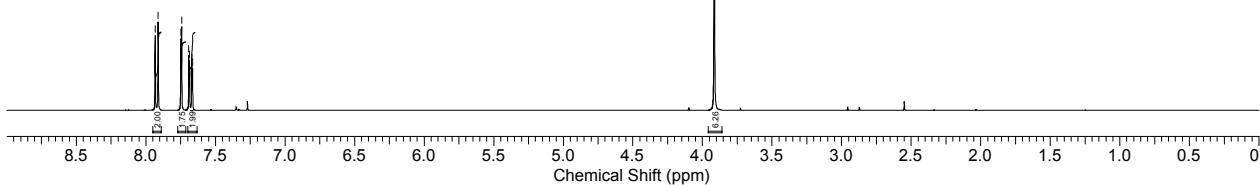
2.7 NMR Spectra



7.93
7.75
7.74
7.69
7.67
7.67

3.92

Frequency (MHz) 400.20
Nucleus ¹H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 90.50
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.6377
Sweep Width (Hz) 8277.89
Temperature (degree C) 21.300



165.1

146.4

142.4

128.9

128.3

126.7

125.1

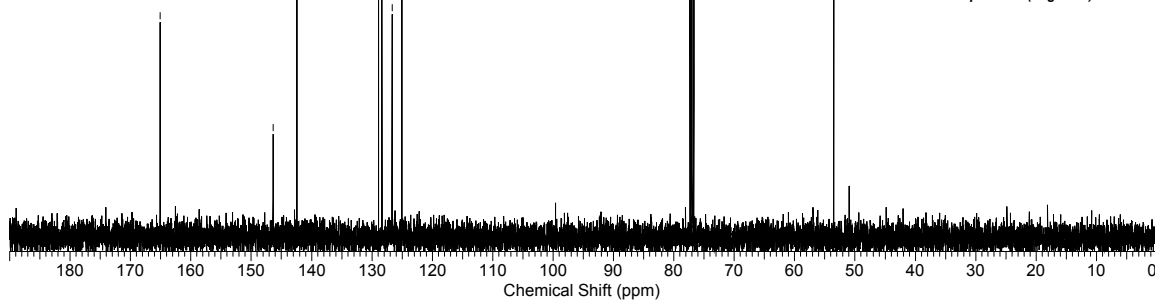
77.3

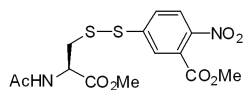
77.0

76.7

53.5

Frequency (MHz) 100.64
Nucleus ¹³C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10012.5645
Sweep Width (Hz) 26176.41
Temperature (degree C) 21.600





7.90
7.88
7.71
7.66
7.65
7.64
7.63
7.27

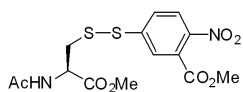
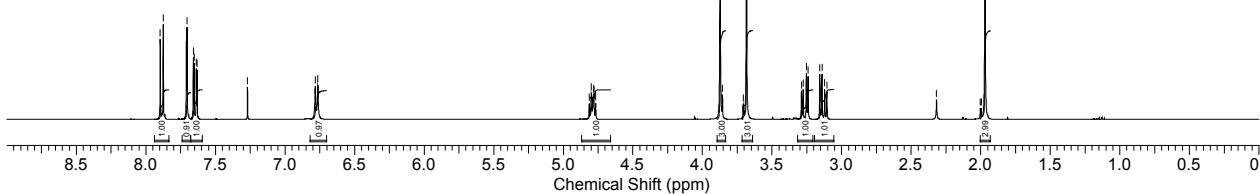
6.78
6.76

4.81
4.80
4.78
4.76
4.77

3.87
3.86
3.71
3.68
3.29
3.28
3.25
3.16
3.14
3.12
3.11

2.32
2.00
1.97

Frequency (MHz) 400.20
Nucleus ¹H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 25.40
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.6377
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000



170.3
170.0
165.5

145.3
144.8

128.6
127.9
126.0
124.6

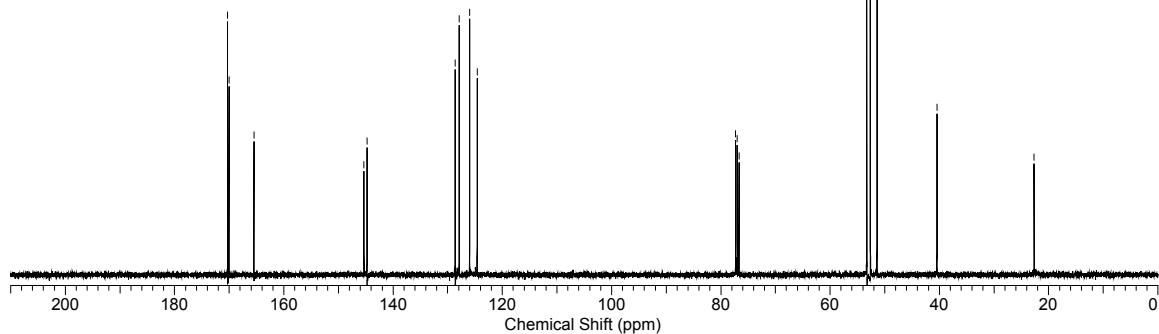
77.3
77.0
76.7

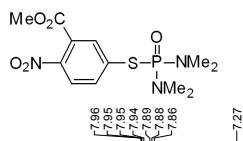
53.3
52.6
51.4

40.4

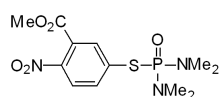
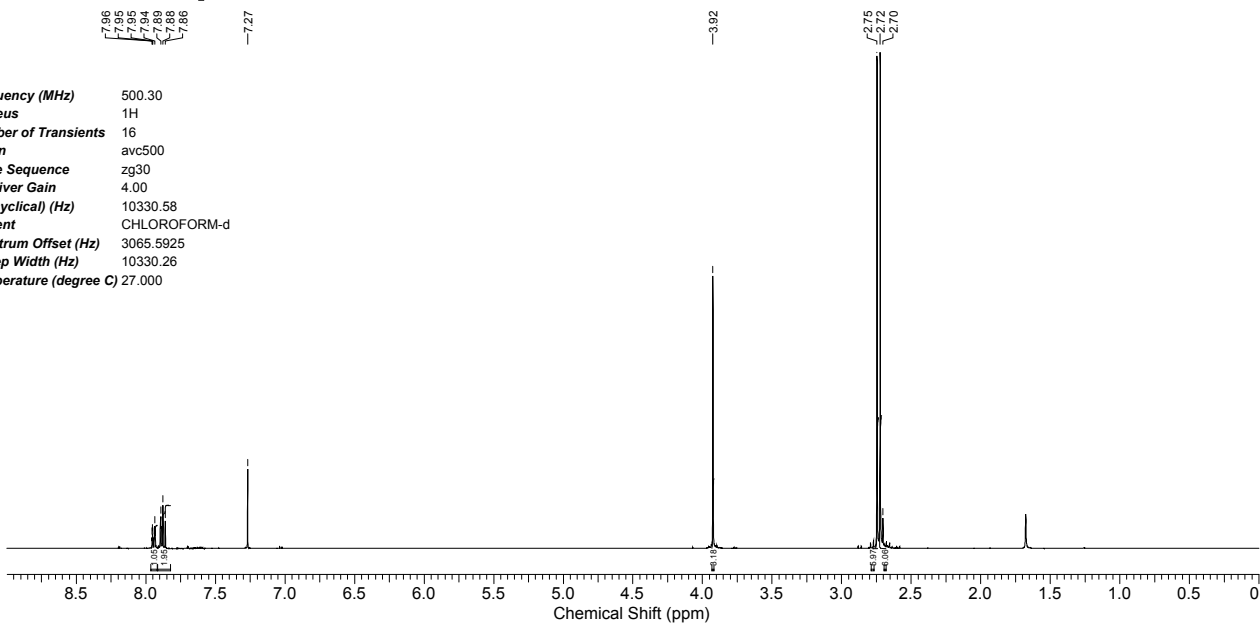
22.7

Frequency (MHz) 100.64
Nucleus ¹³C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 9999.7822
Sweep Width (Hz) 26176.41
Temperature (degree C) 27.000

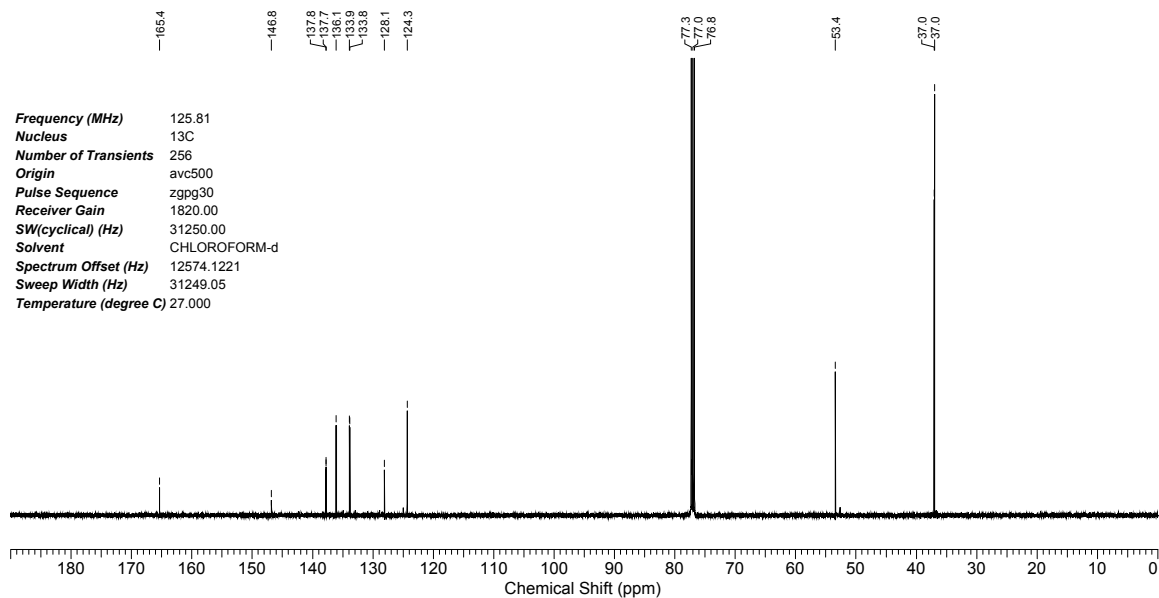


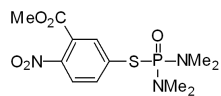


Frequency (MHz) 500.30
 Nucleus ^1H
 Number of Transients 16
 Origin avc500
 Pulse Sequence zg30
 Receiver Gain 4.00
 SW(cyclical) (Hz) 10330.58
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 3065.5925
 Sweep Width (Hz) 10330.26
 Temperature (degree C) 27.000



Frequency (MHz) 125.81
 Nucleus ^{13}C
 Number of Transients 256
 Origin avc500
 Pulse Sequence zgpg30
 Receiver Gain 1820.00
 SW(cyclical) (Hz) 31250.00
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 12574.1221
 Sweep Width (Hz) 31249.05
 Temperature (degree C) 27.000



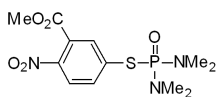
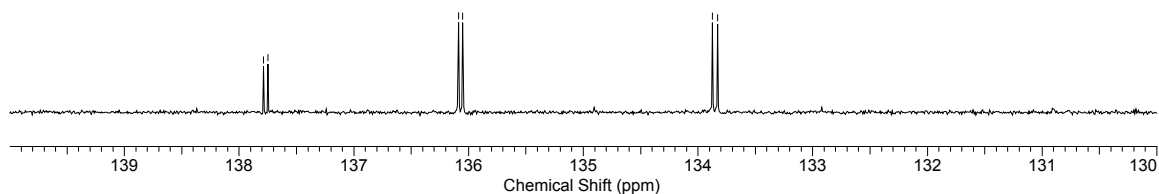


137.8
137.7

136.1
136.1

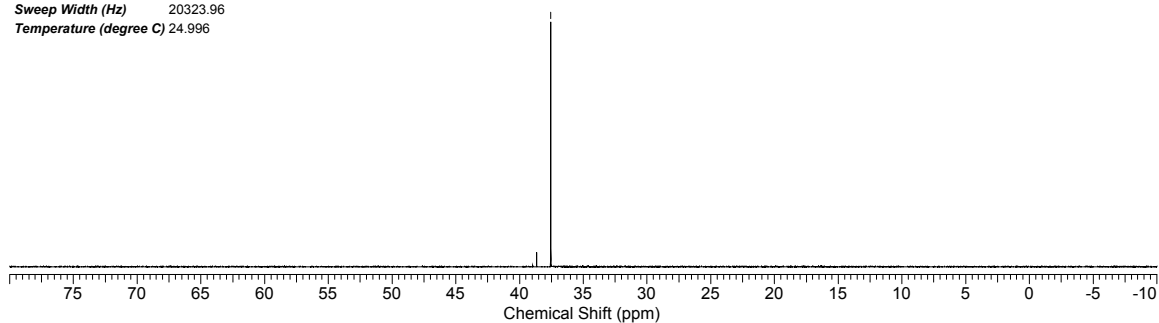
133.9
133.8

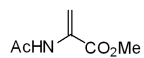
Frequency (MHz) 125.81
Nucleus ¹³C
Number of Transients 256
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12574.1221
Sweep Width (Hz) 31249.05
Temperature (degree C) 27.000



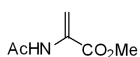
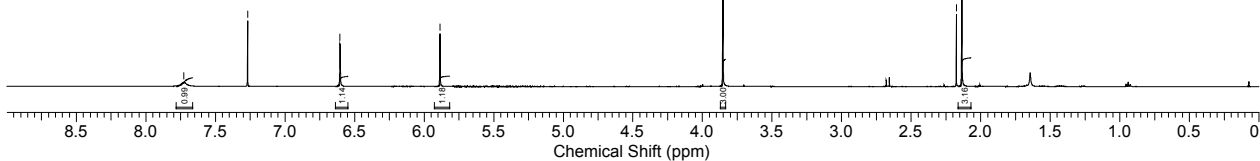
-37.6

Frequency (MHz) 202.46
Nucleus ³¹P
Number of Transients 719
Origin drx500
Pulse Sequence zgpg30
Receiver Gain 9195.20
SW(cyclical) (Hz) 20325.20
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 7490.8750
Sweep Width (Hz) 20323.96
Temperature (degree C) 24.996

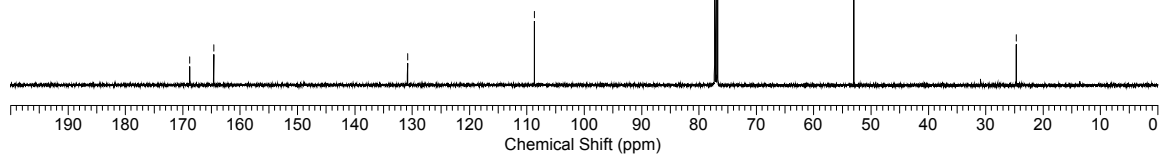


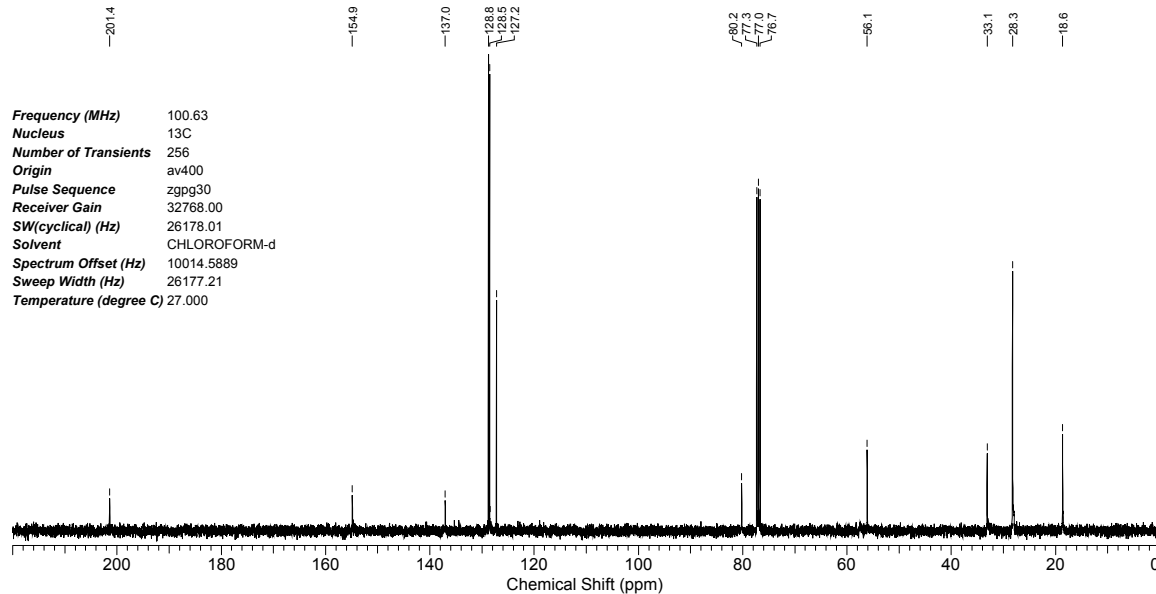
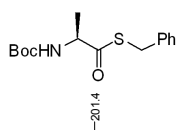
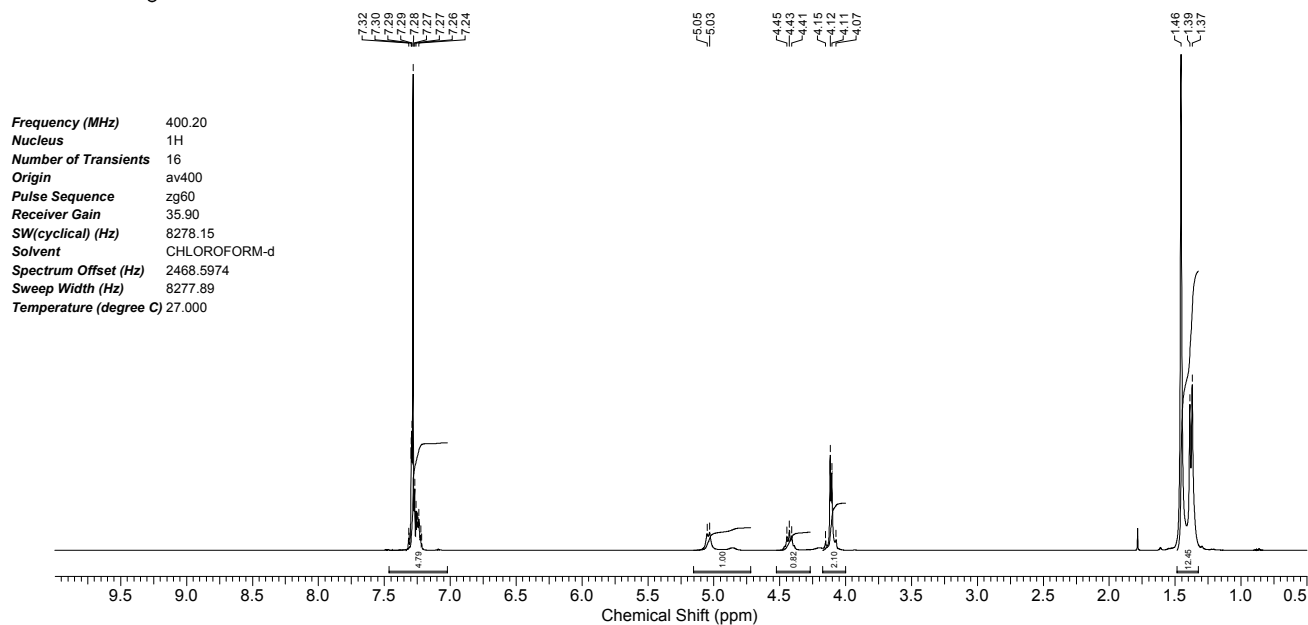
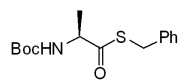


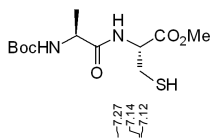
Frequency (MHz) 500.30
 Nucleus ^1H
 Number of Transients 16
 Origin avc500
 Pulse Sequence zg30
 Receiver Gain 4.00
 SW(cyclical) (Hz) 10330.58
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 3065.5925
 Sweep Width (Hz) 10330.26
 Temperature (degree C) 27.000



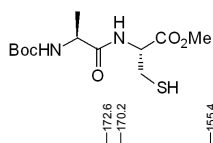
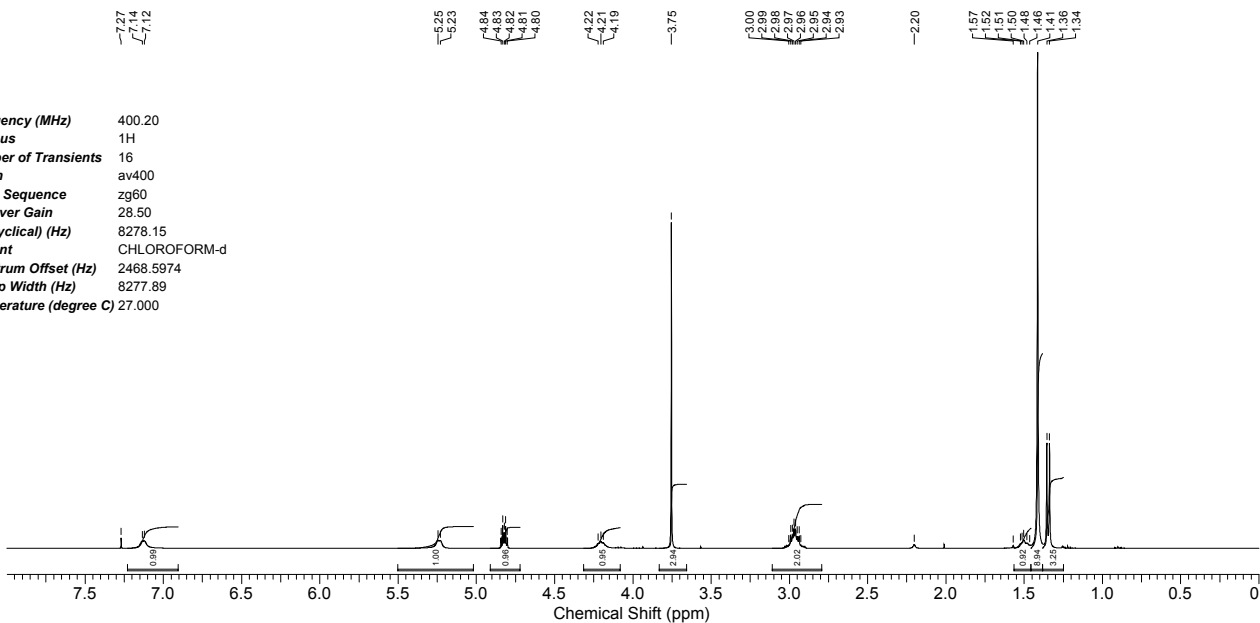
Frequency (MHz) 125.81
 Nucleus ^{13}C
 Number of Transients 512
 Origin avc500
 Pulse Sequence zgpg30
 Receiver Gain 1820.00
 SW(cyclical) (Hz) 31250.00
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 12574.1221
 Sweep Width (Hz) 31249.05
 Temperature (degree C) 27.000



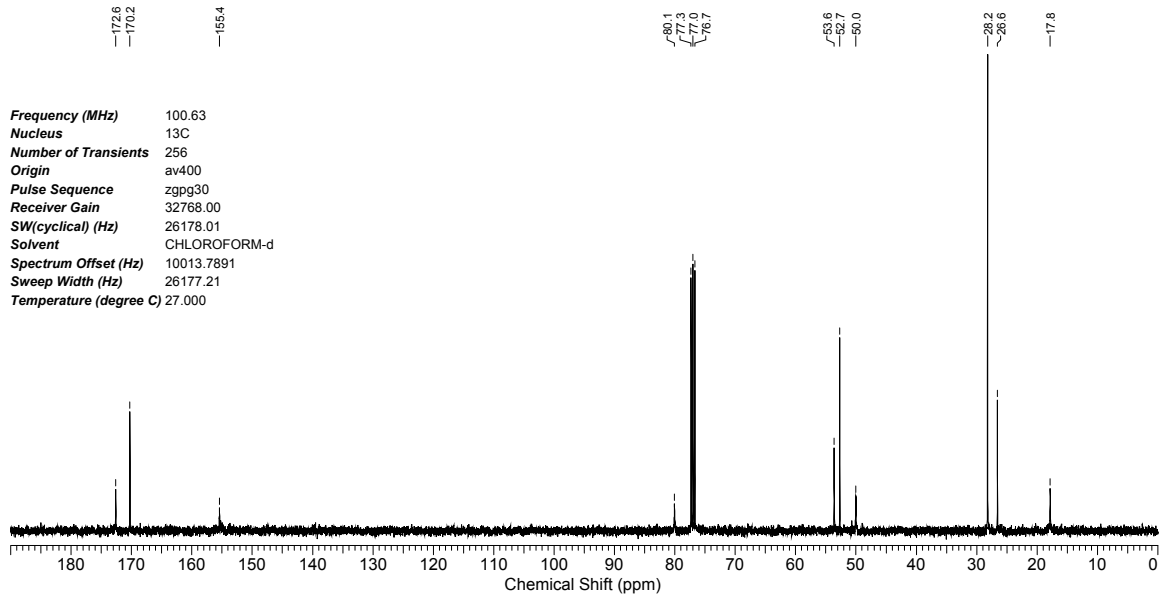


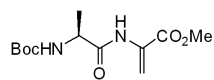


Frequency (MHz) 400.20
 Nucleus 1H
 Number of Transients 16
 Origin av400
 Pulse Sequence zg60
 Receiver Gain 28.50
 SW(cyclical) (Hz) 8278.15
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 2468.5974
 Sweep Width (Hz) 8277.89
 Temperature (degree C) 27.000

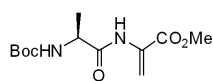
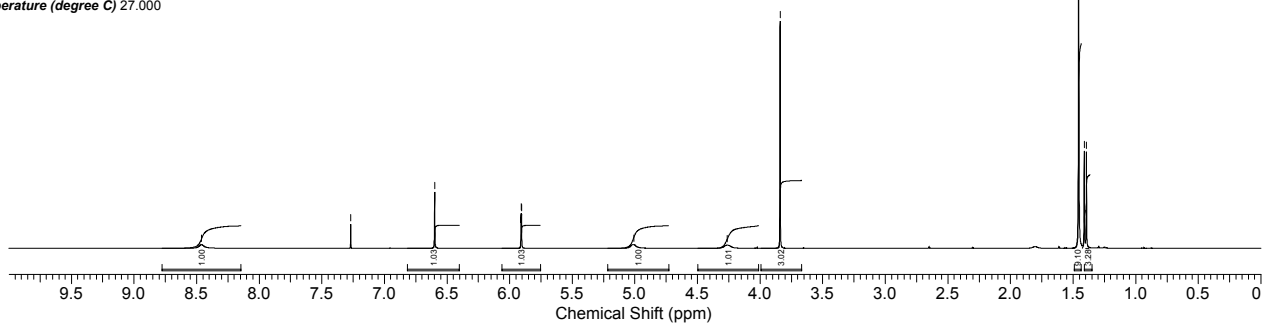


Frequency (MHz) 100.63
 Nucleus 13C
 Number of Transients 256
 Origin av400
 Pulse Sequence zgpg30
 Receiver Gain 32768.00
 SW(cyclical) (Hz) 26178.01
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 10013.7891
 Sweep Width (Hz) 26177.21
 Temperature (degree C) 27.000

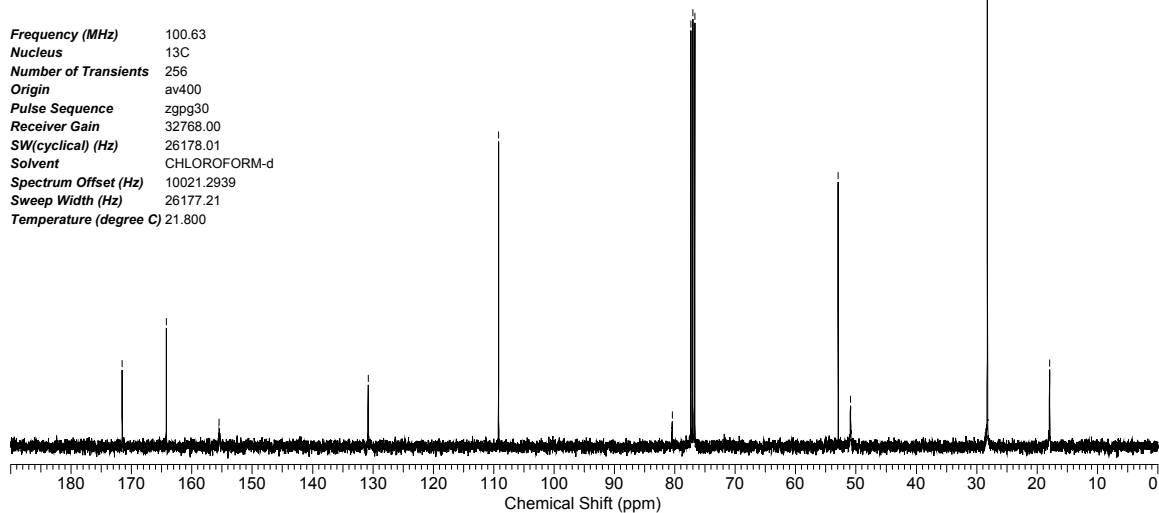


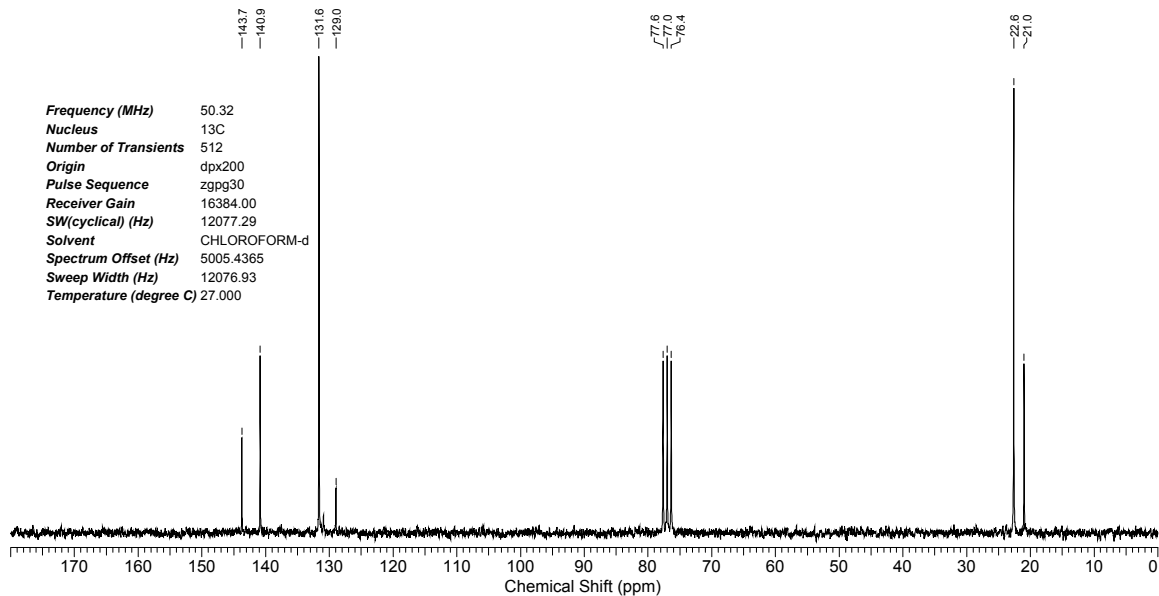
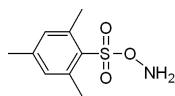
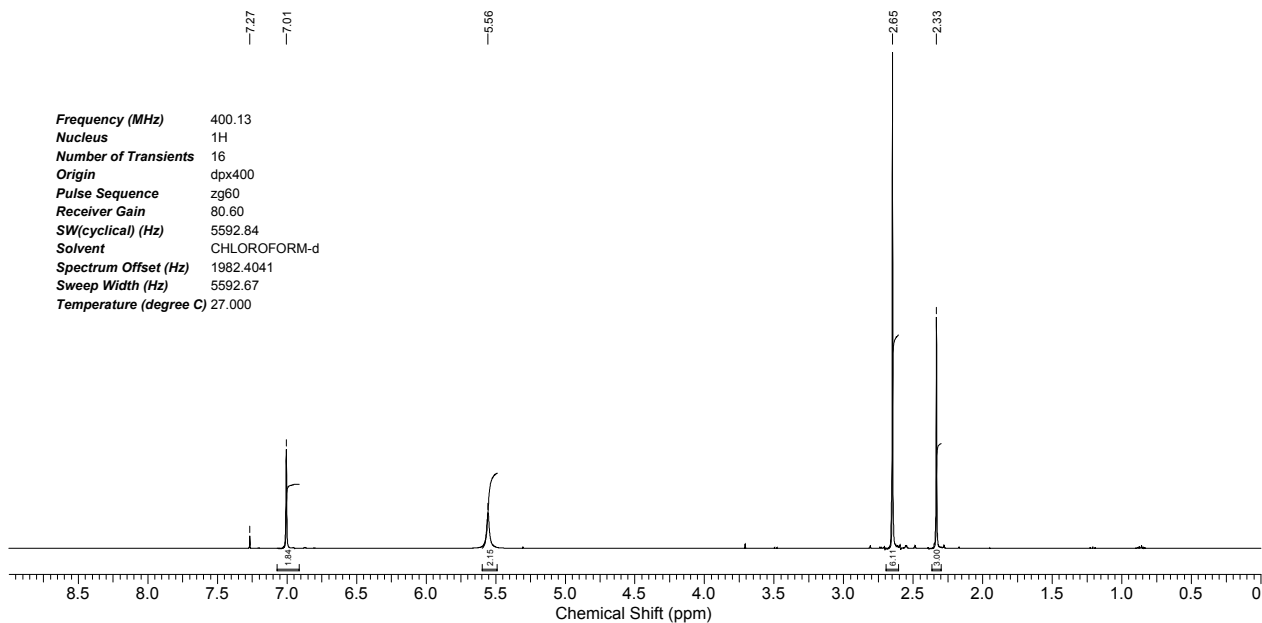
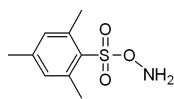


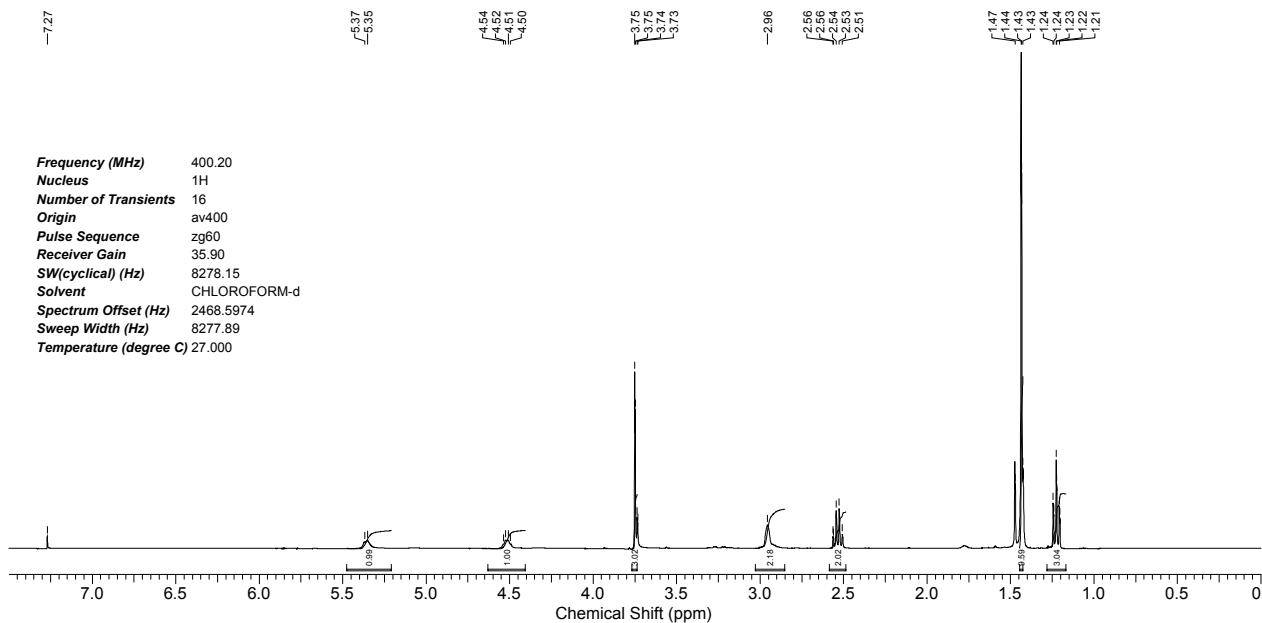
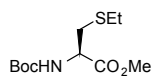
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 90.50
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000



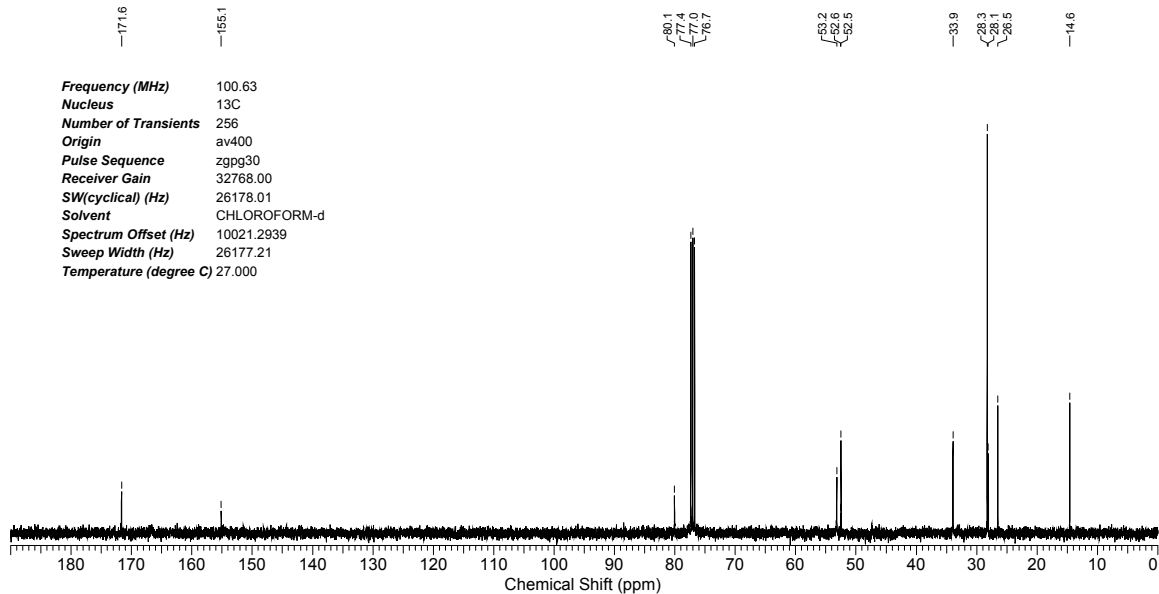
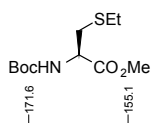
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 21.800



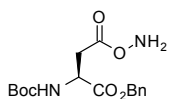




Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 35.90
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000



Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000



7.38
7.37
7.37
7.36
7.35
7.34
7.34
7.27

6.28

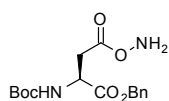
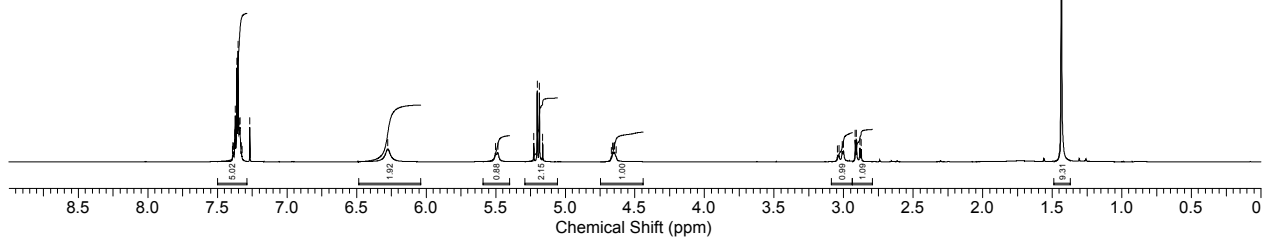
5.50
5.49

5.23
5.20
5.19
5.16
4.87
4.86
4.85
4.64

3.04
3.03
3.01
2.90
2.91
2.88
2.87

1.43

Frequency (MHz) 500.30
Nucleus 1H
Number of Transients 16
Origin avc500
Pulse Sequence zg30
Receiver Gain 4.00
SW(cyclical) (Hz) 10330.58
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 3065.5417
Sweep Width (Hz) 10330.26
Temperature (degree C) 24.970



171.5
170.5

155.2

135.1

128.6
128.5
128.4

80.3
77.3
77.2
76.8

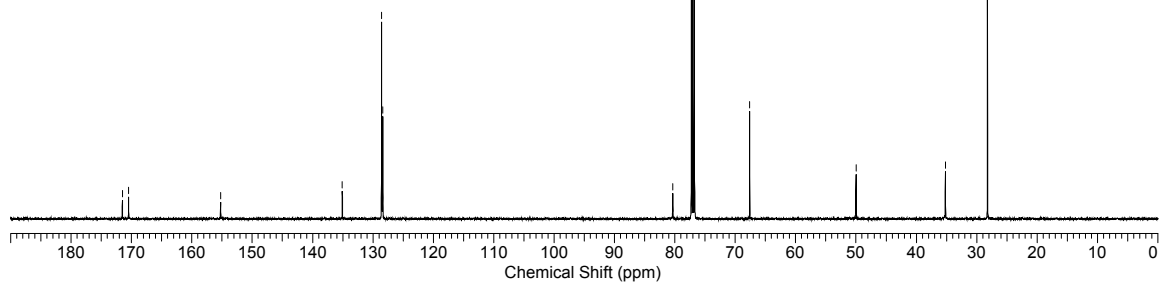
67.6

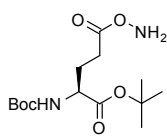
50.0

35.2

28.2

Frequency (MHz) 125.80
Nucleus 13C
Number of Transients 256
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12571.3047
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970





7.27

6.36

5.10

5.09

4.22

4.21

2.48

2.48

2.48

2.44

2.43

2.42

2.40

2.38

2.21

2.19

1.93

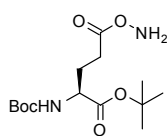
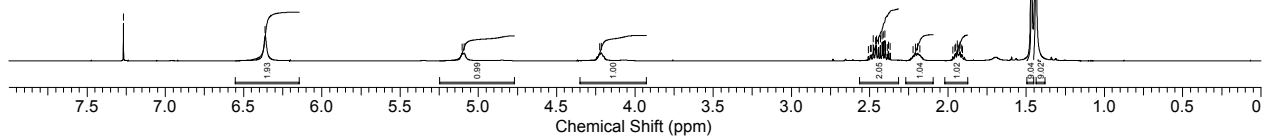
1.92

1.91

1.47

1.44

Frequency (MHz) 500.30
Nucleus 1H
Number of Transients 16
Origin avc500
Pulse Sequence zg30
Receiver Gain 4.00
SW(cyclical) (Hz) 10330.58
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 3065.5417
Sweep Width (Hz) 10330.26
Temperature (degree C) 24.970



173.6

171.1

155.3

82.4

79.9

77.3

77.0

76.7

53.1

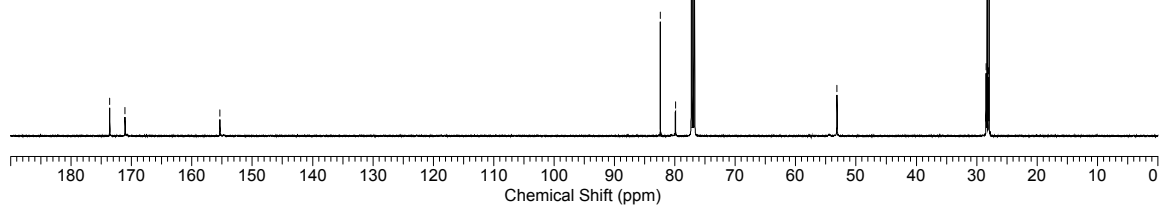
28.5

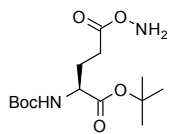
28.3

28.1

28.0

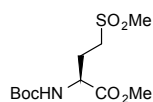
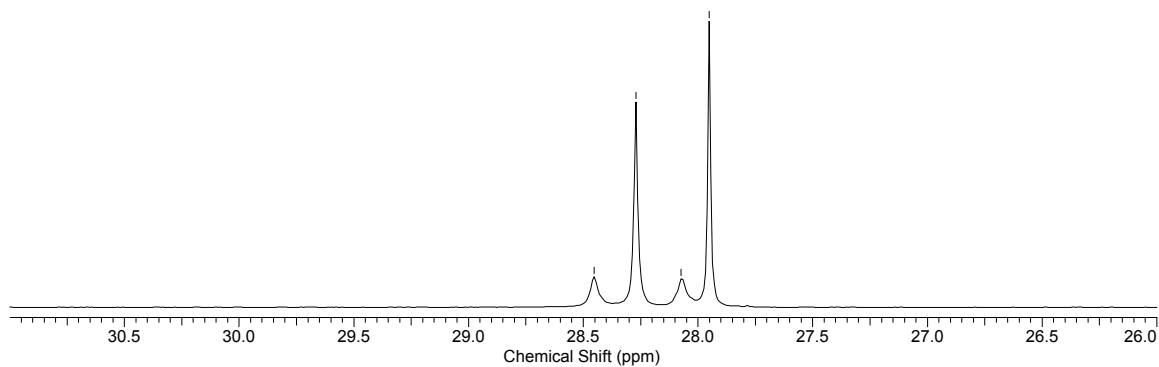
Frequency (MHz) 125.80
Nucleus 13C
Number of Transients 1024
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12571.9844
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970





Frequency (MHz) 125.80
 Nucleus ^{13}C
 Number of Transients 1024
 Origin avc500
 Pulse Sequence zgpg30
 Receiver Gain 1820.00
 SW(cyclical) (Hz) 31250.00
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 12571.9844
 Sweep Width (Hz) 31249.05
 Temperature (degree C) 24.970

-28.5
 -28.3
 -28.1
 -28.0



-7.27

5.39
5.37

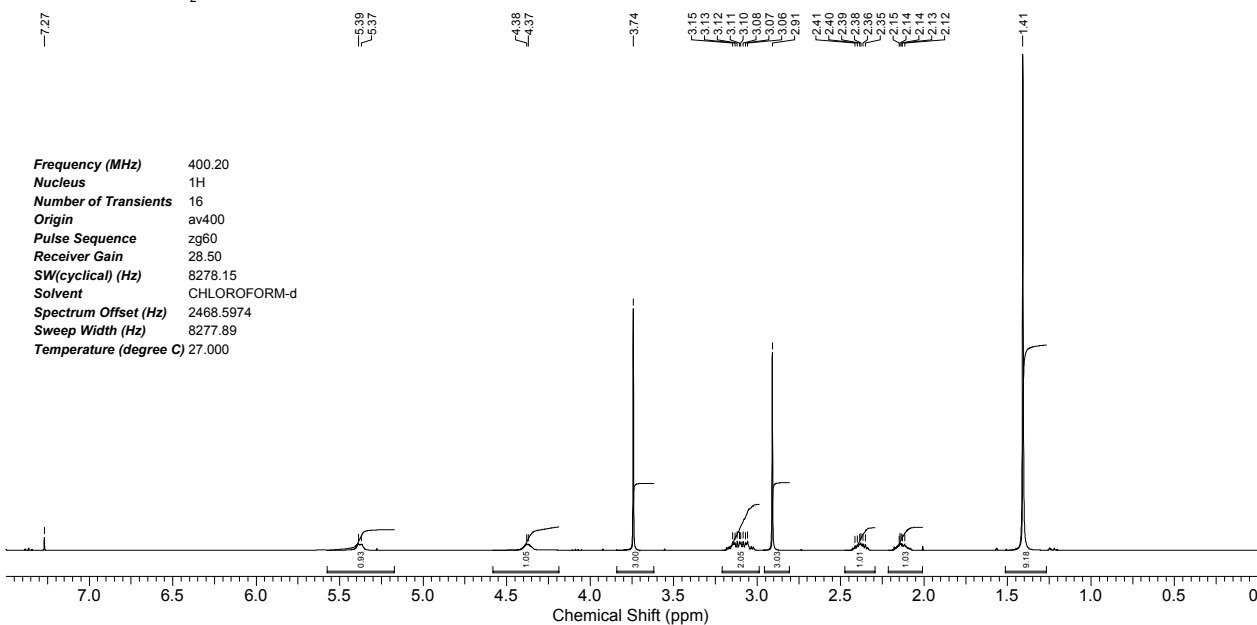
4.38
4.37

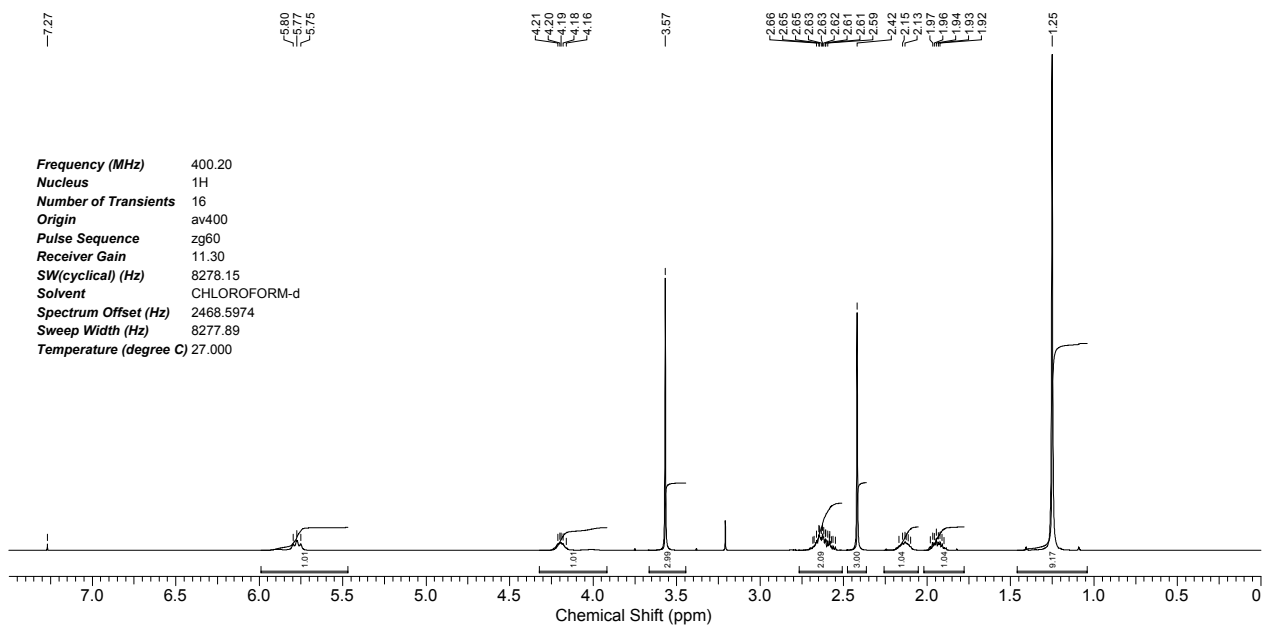
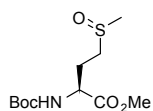
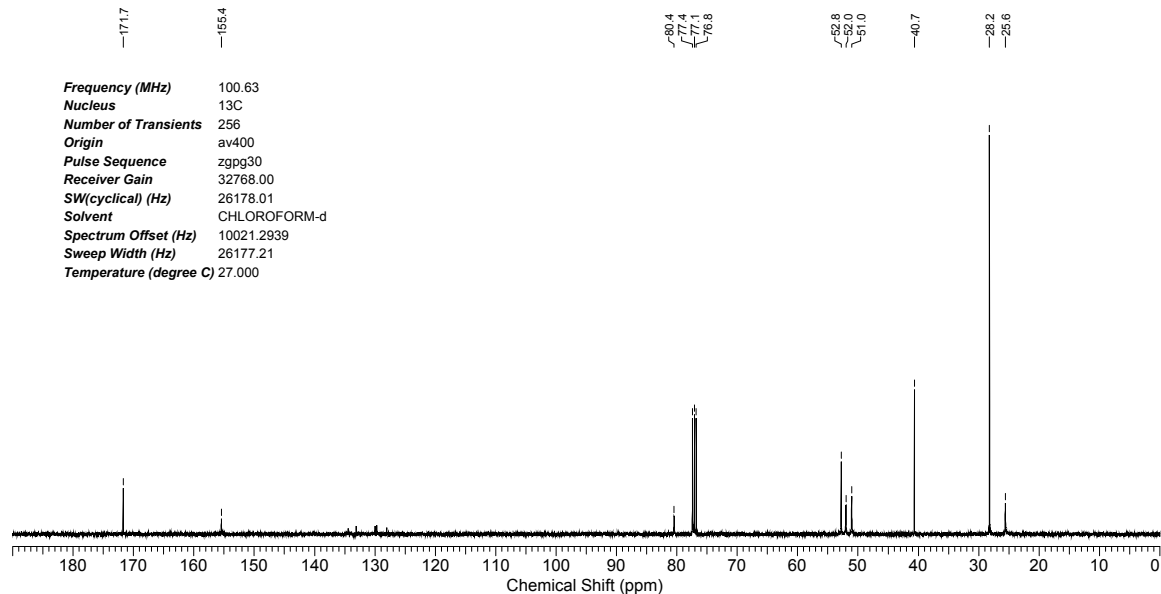
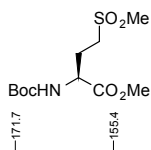
-3.74

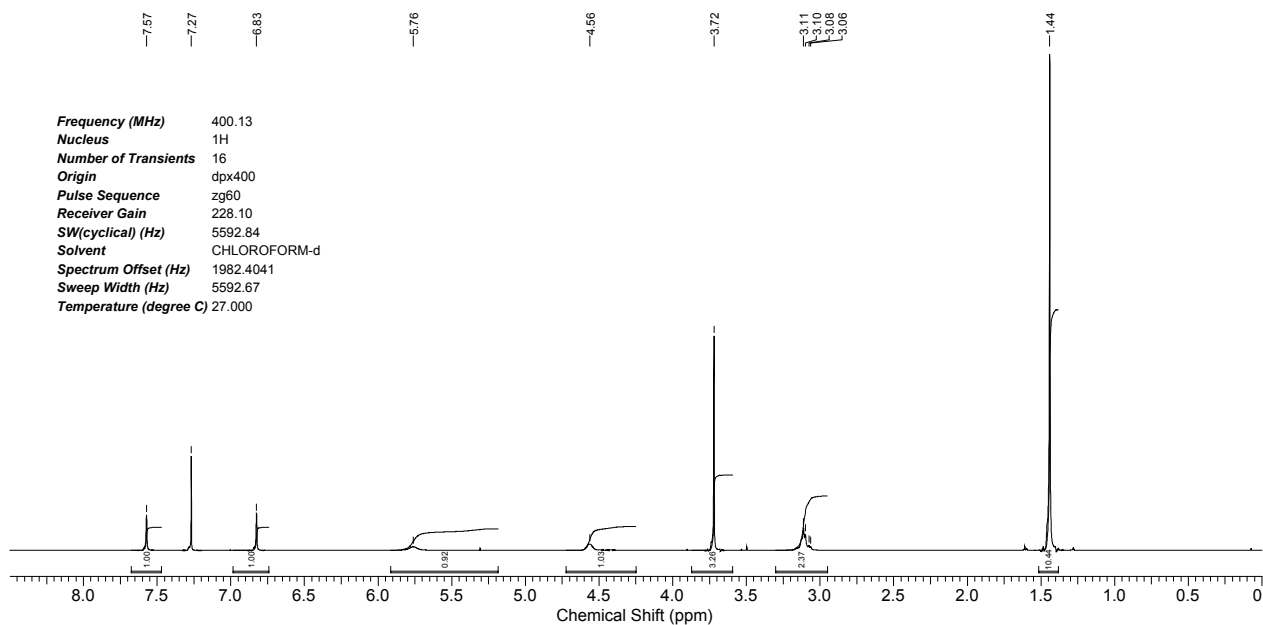
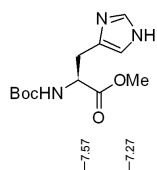
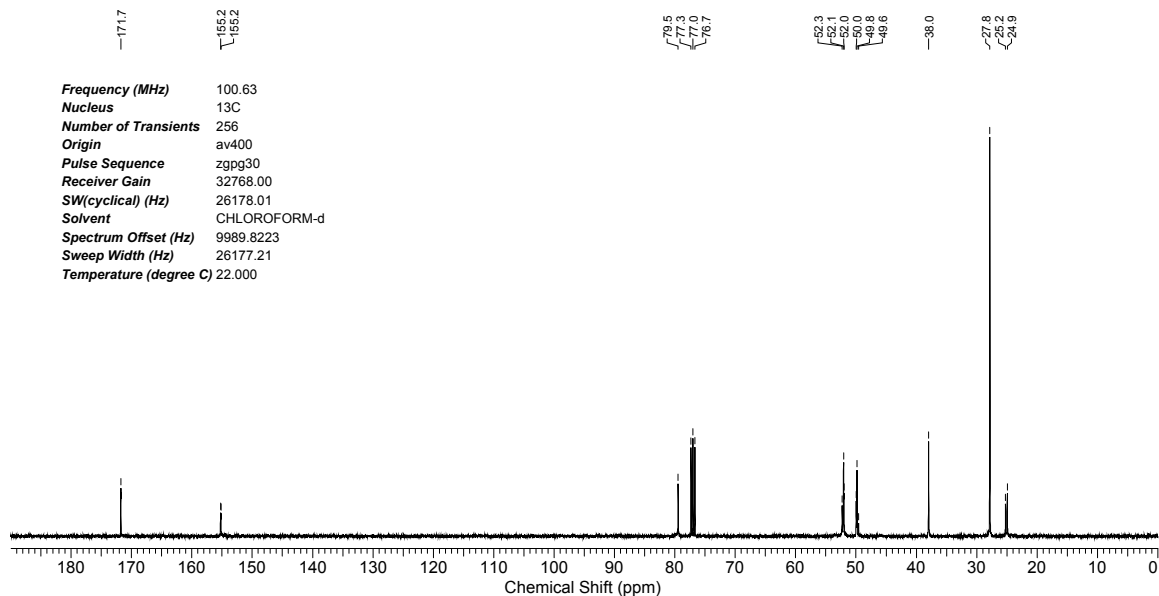
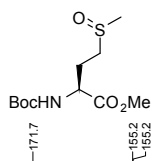
3.15
 3.13
 3.12
 3.11
 3.10
 3.08
 3.07
 3.06
 2.91
 2.41
 2.40
 2.39
 2.38
 2.36
 2.35
 2.15
 2.14
 2.13
 2.12

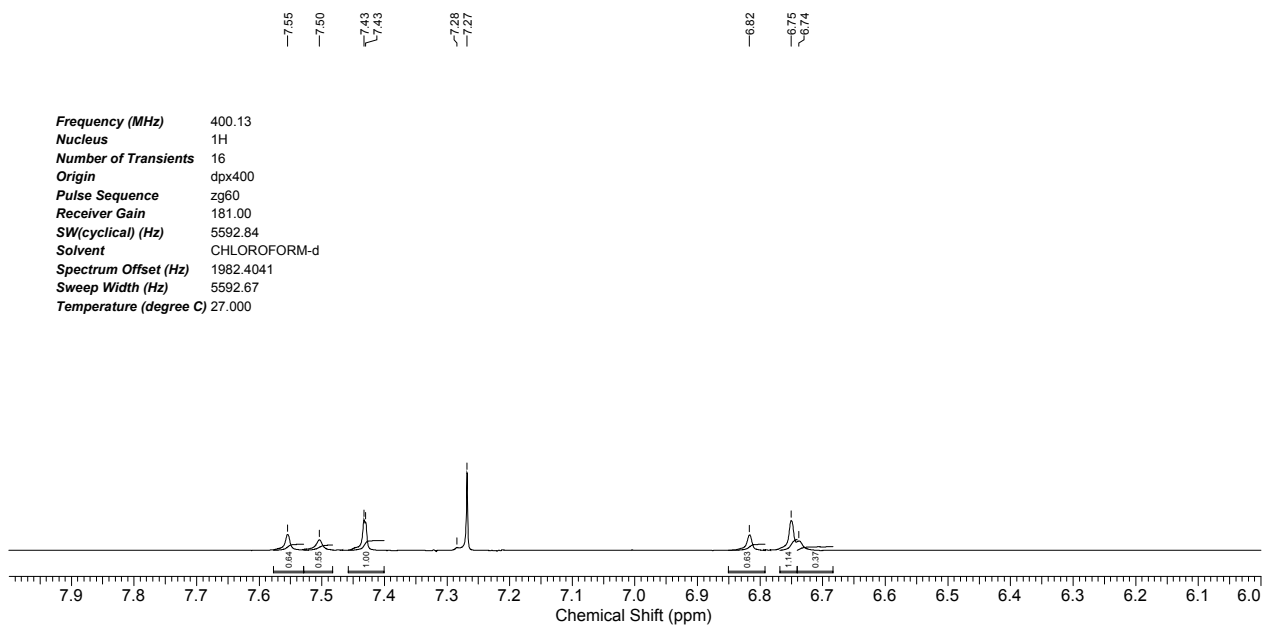
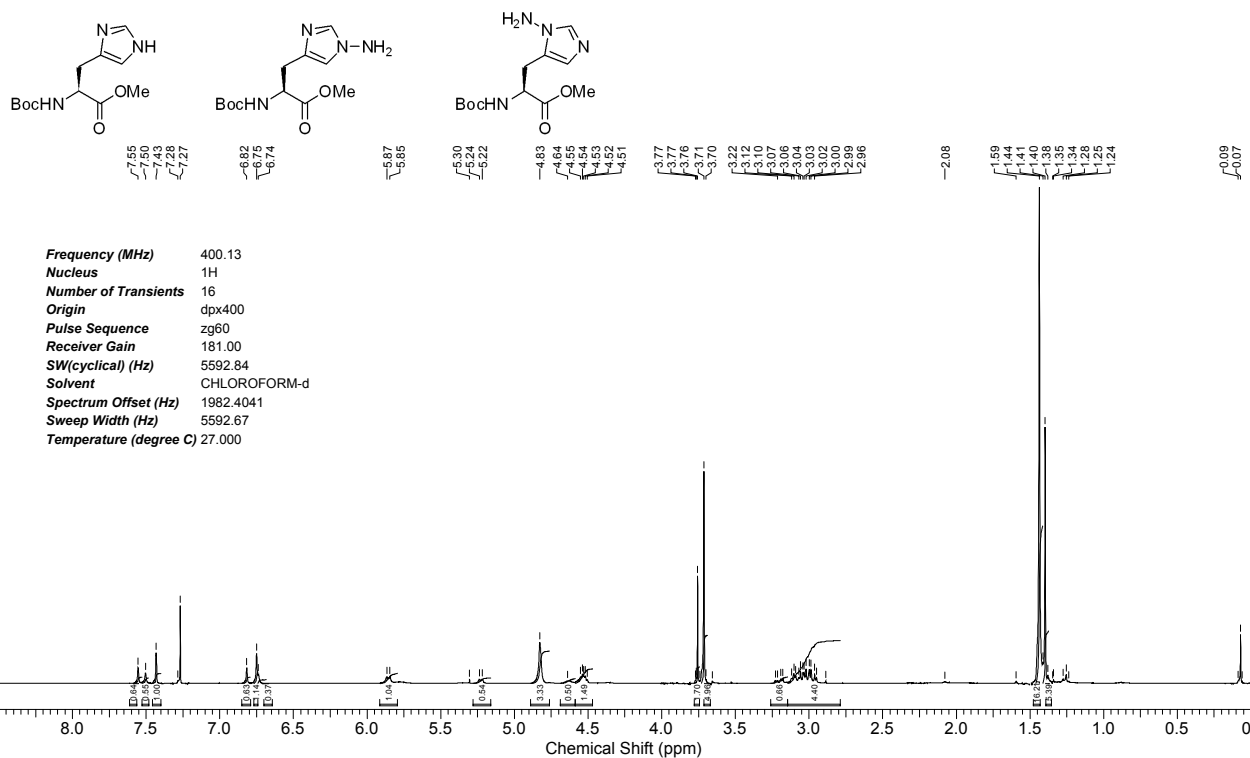
-1.41

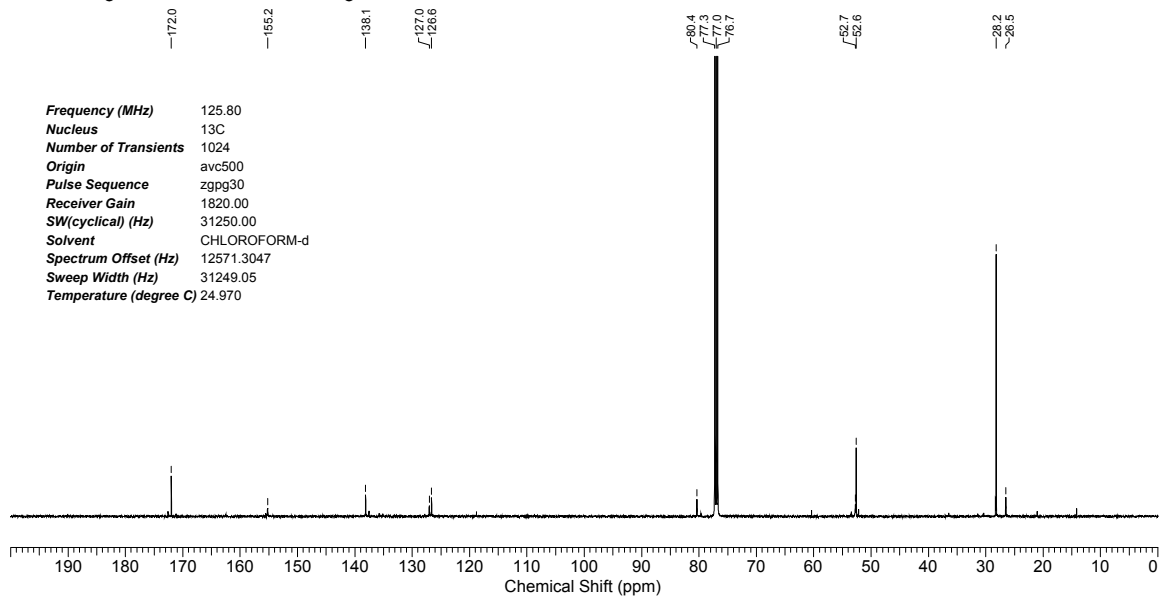
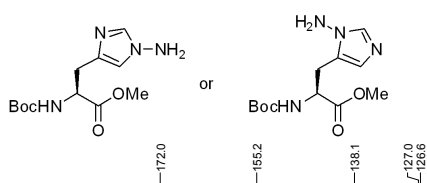
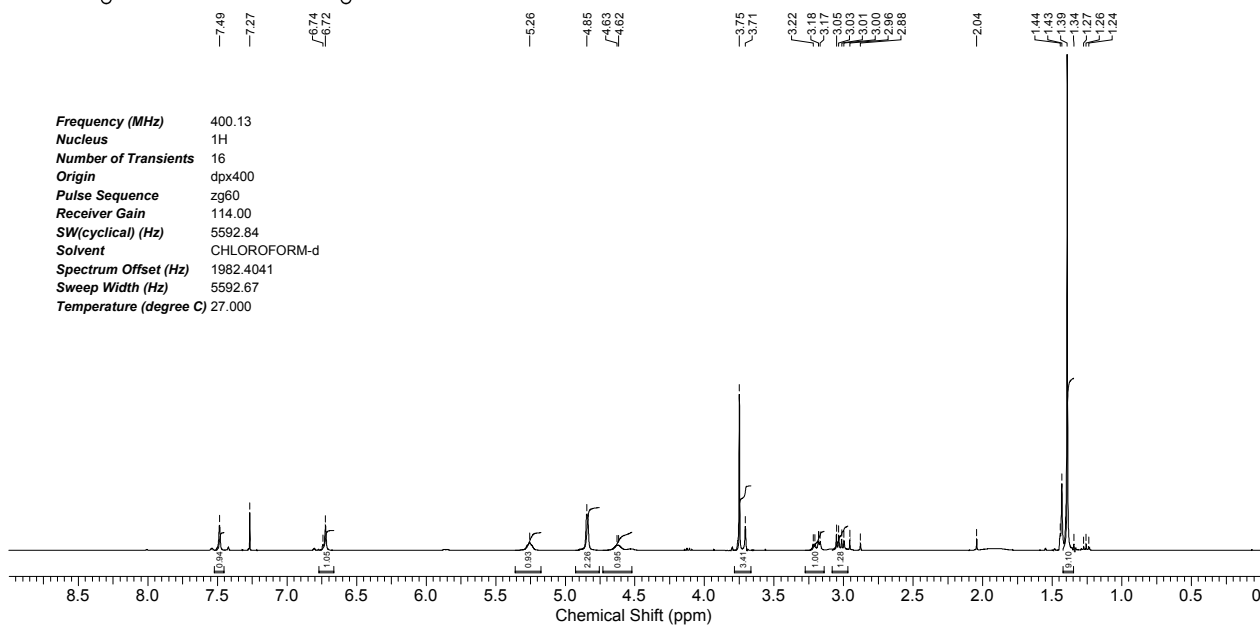
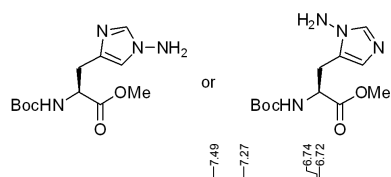
Frequency (MHz) 400.20
 Nucleus ^1H
 Number of Transients 16
 Origin av400
 Pulse Sequence zg60
 Receiver Gain 28.50
 SW(cyclical) (Hz) 8278.15
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 2468.5974
 Sweep Width (Hz) 8277.89
 Temperature (degree C) 27.000

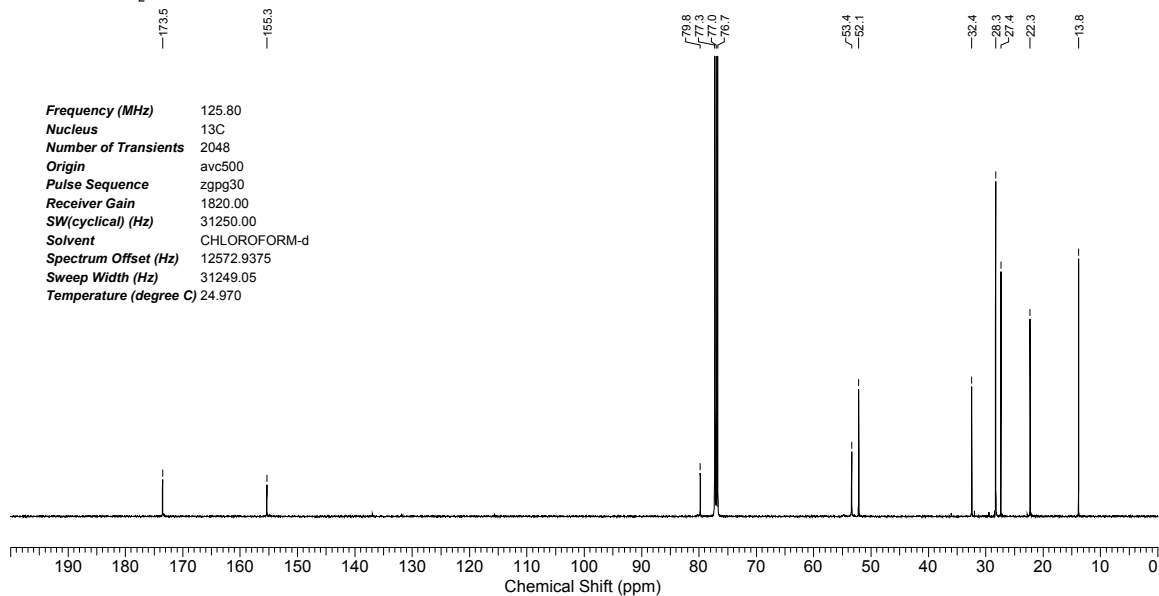
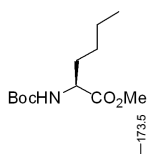
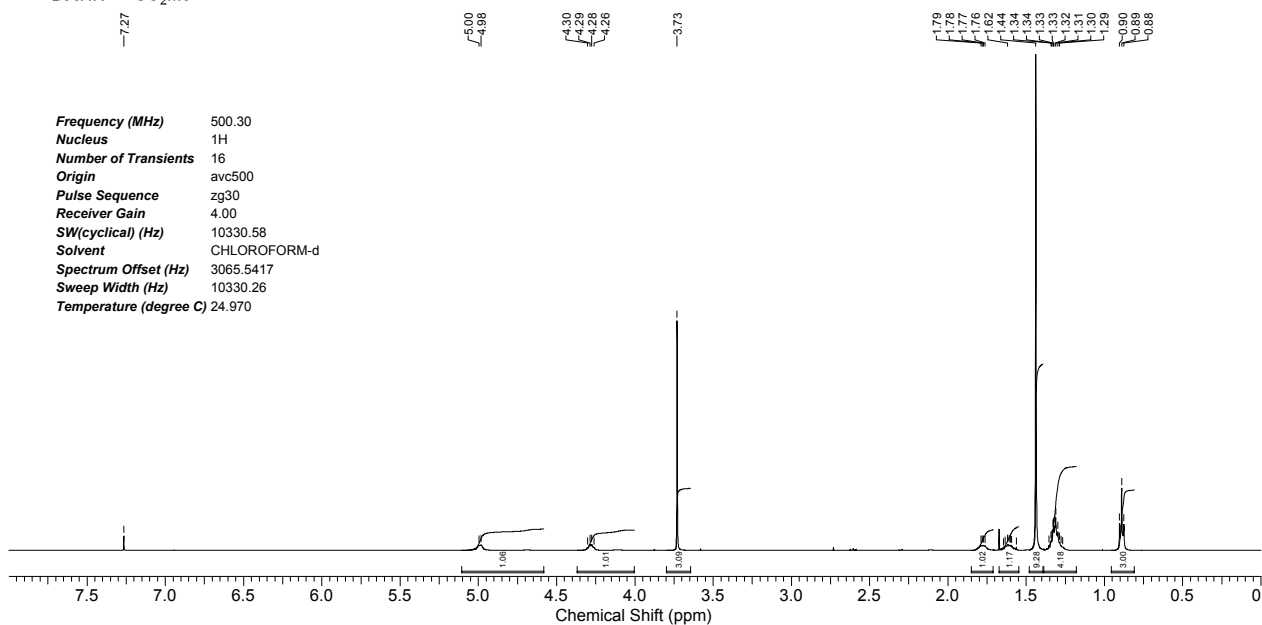
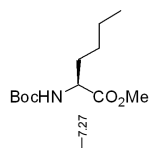


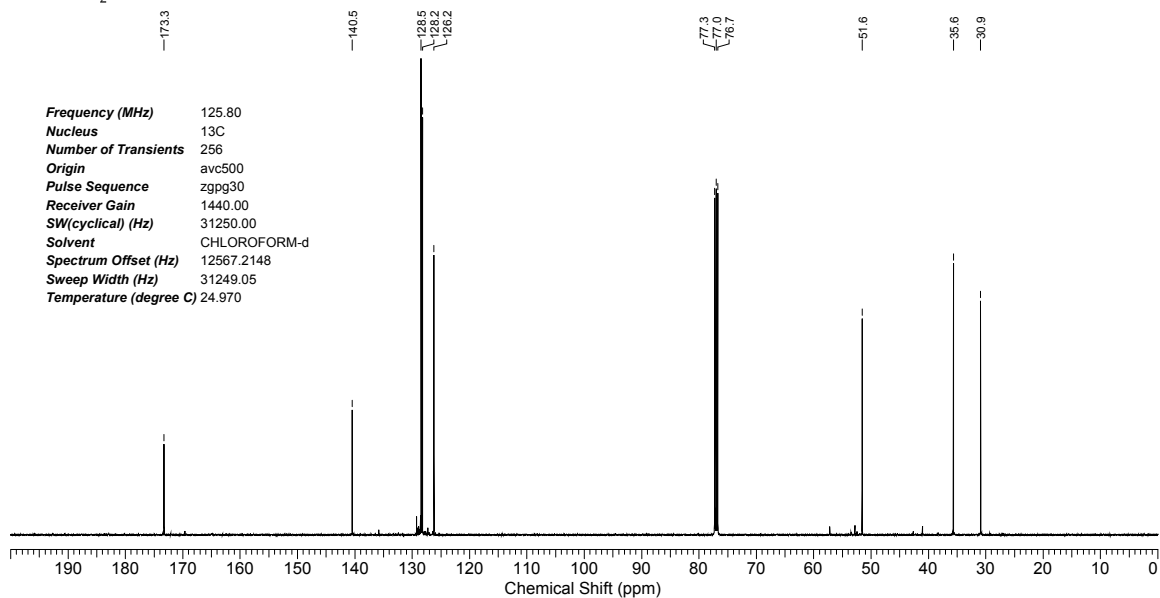
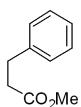
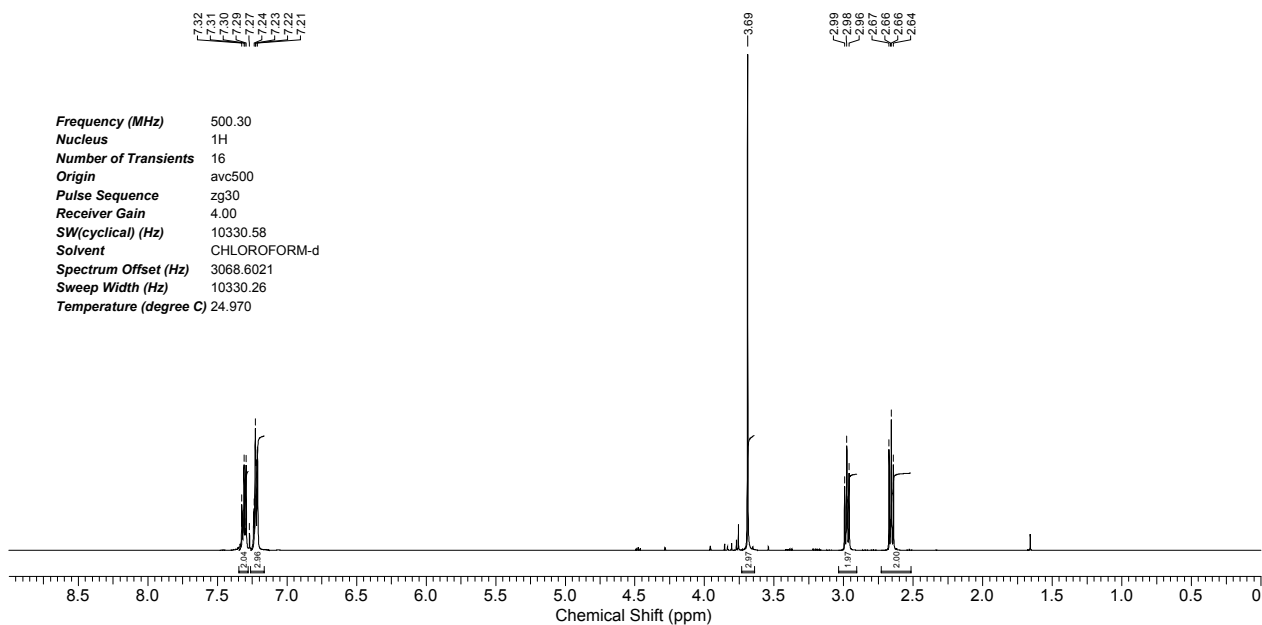
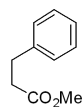


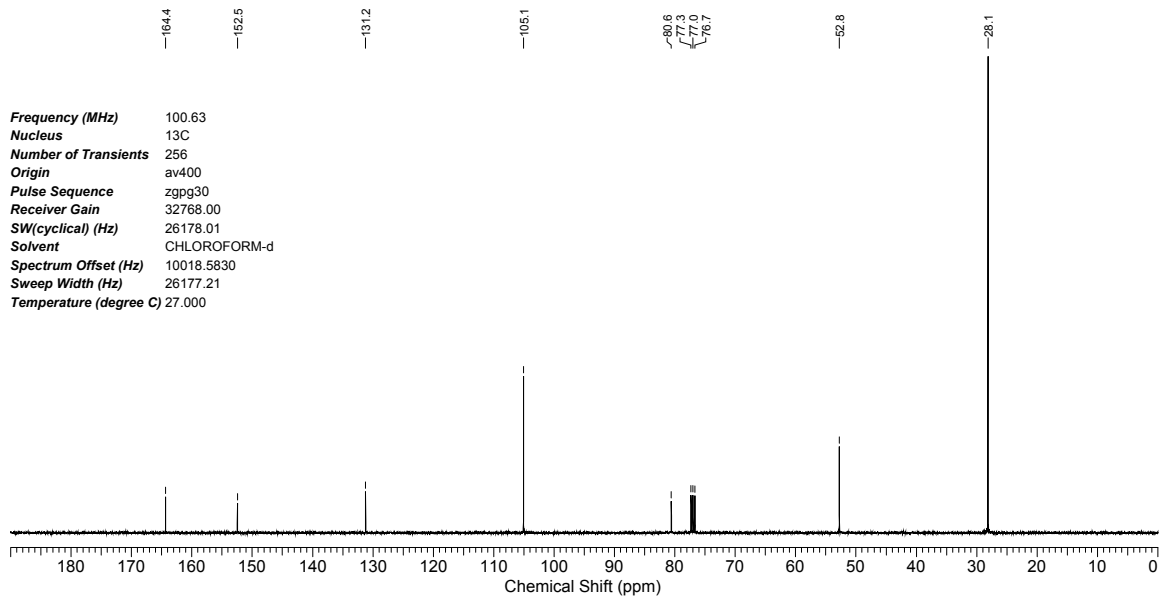
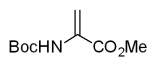
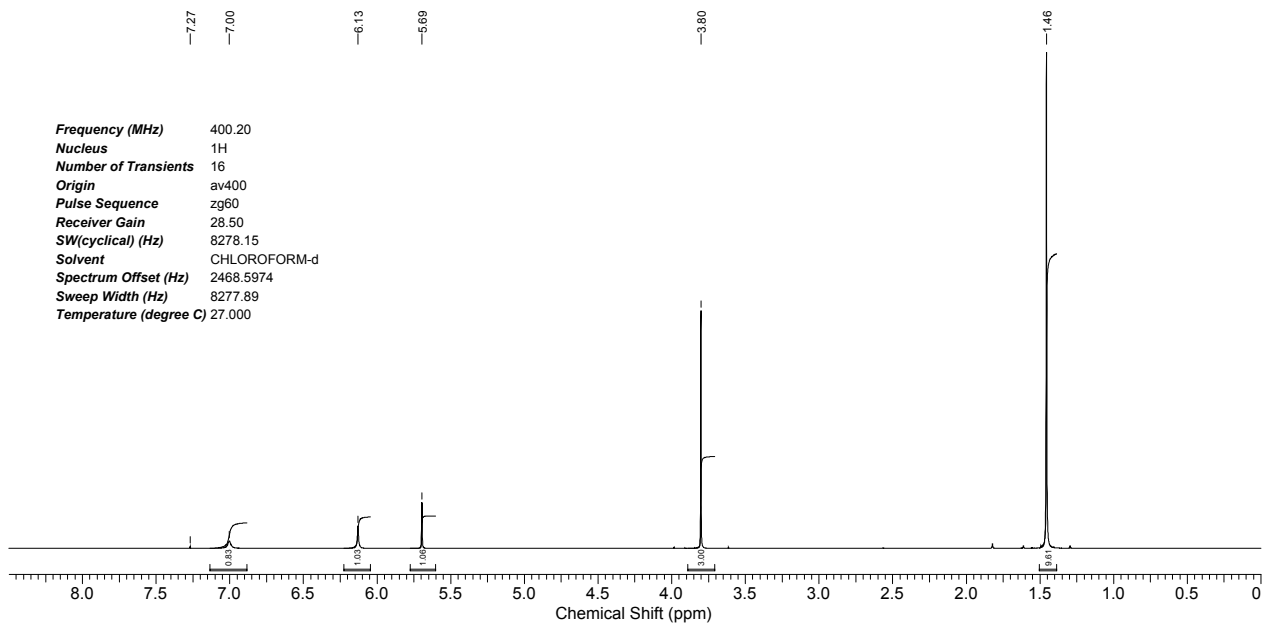
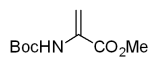


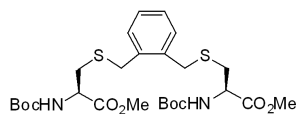












7.29
7.27
7.26
7.26
7.23
7.23
7.22
7.22
7.21

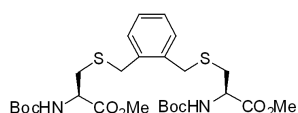
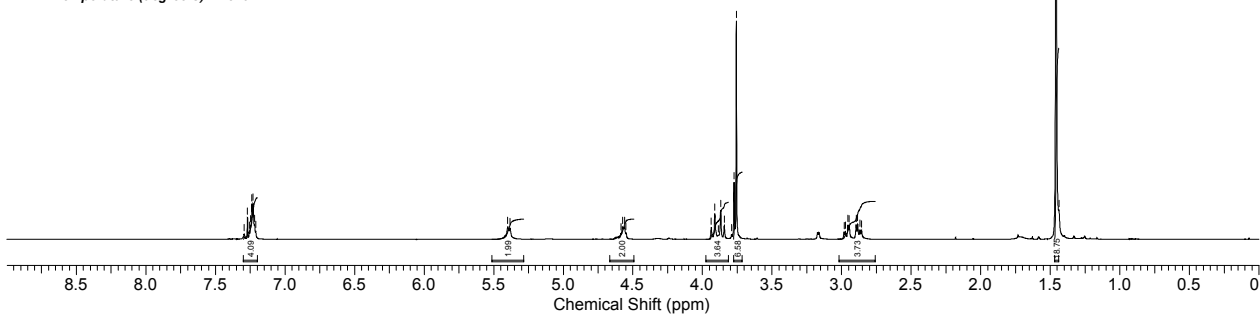
5.40
5.38

4.59
4.58
4.56
4.55
3.94
3.91
3.89
3.87
3.84
3.79
3.76

2.98
2.97
2.95
2.94
2.90
2.89
2.87
2.86

1.46
1.44

Frequency (MHz) 500.30
Nucleus 1H
Number of Transients 16
Origin avc500
Pulse Sequence zg30
Receiver Gain 4.00
SW(cyclical) (Hz) 10330.58
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 3065.5417
Sweep Width (Hz) 10330.26
Temperature (degree C) 14.370



171.6
155.1
135.7
130.9
127.7

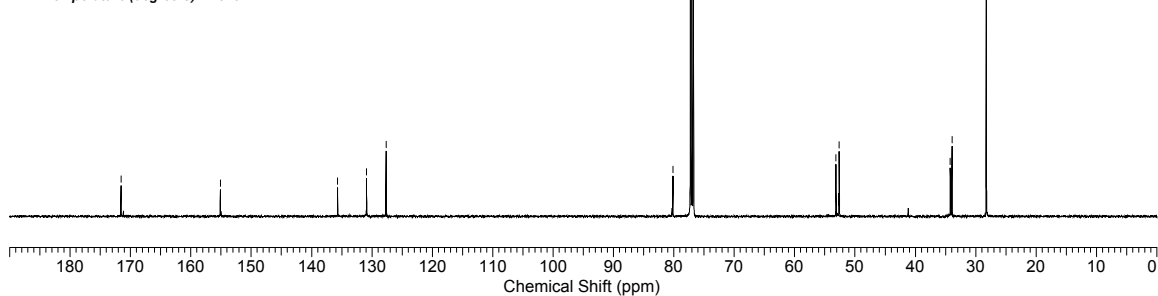
80.2
77.3
77.0
76.7

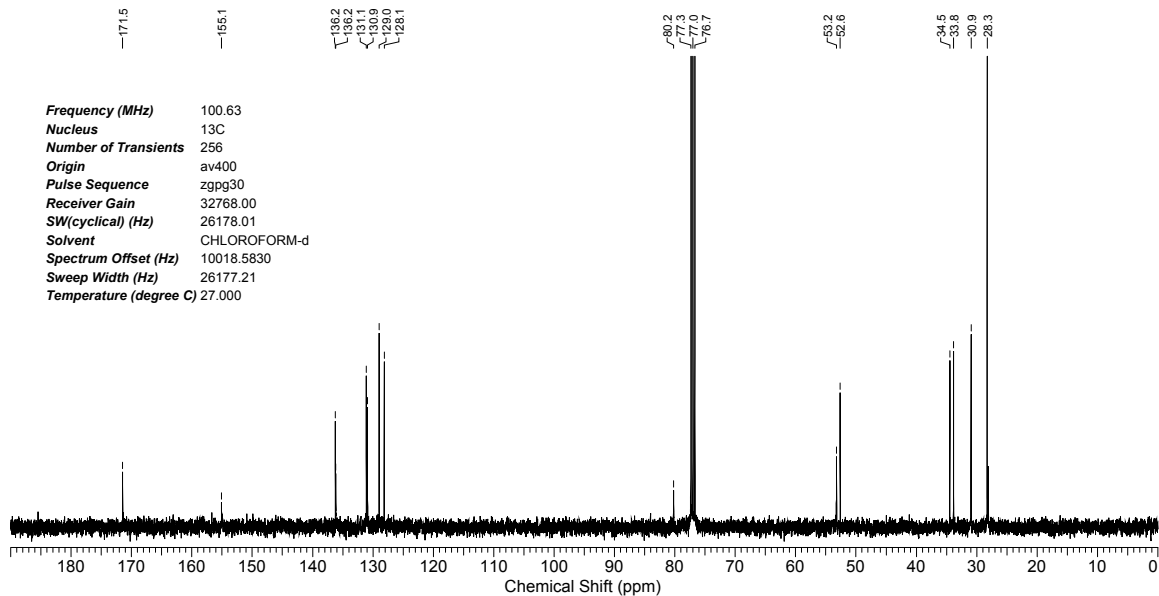
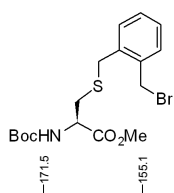
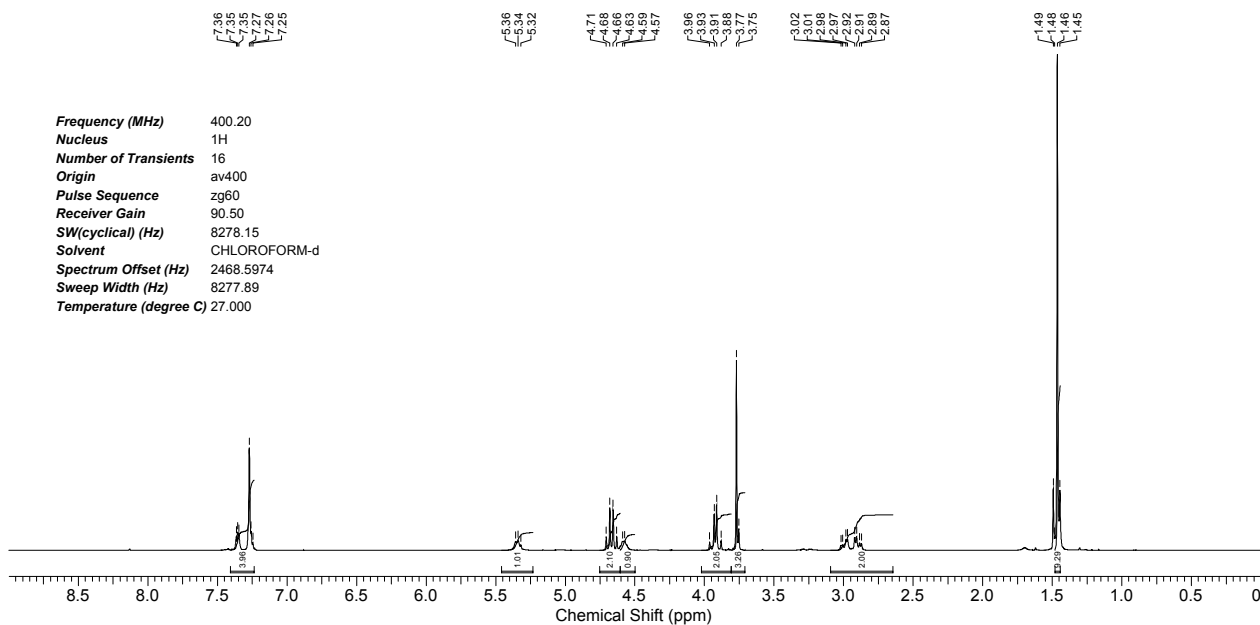
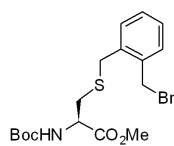
53.2
52.6

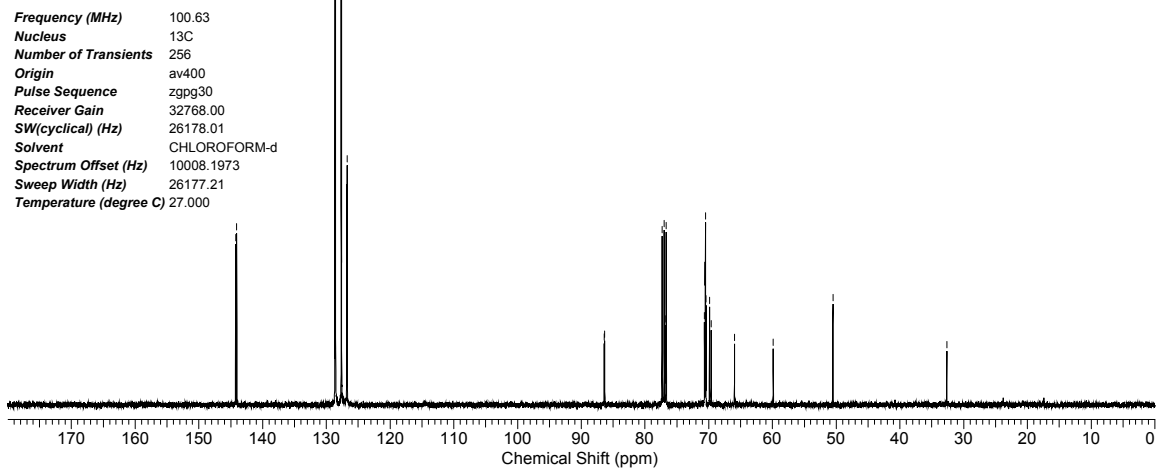
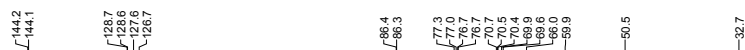
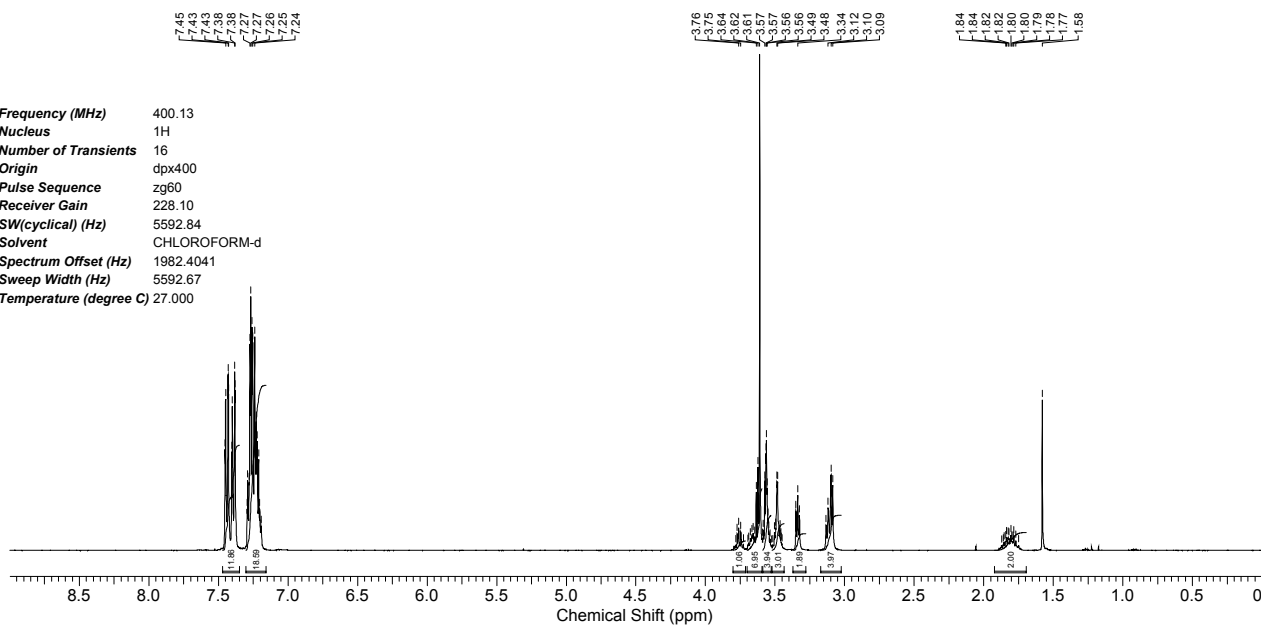
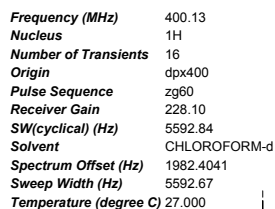
34.2
33.9

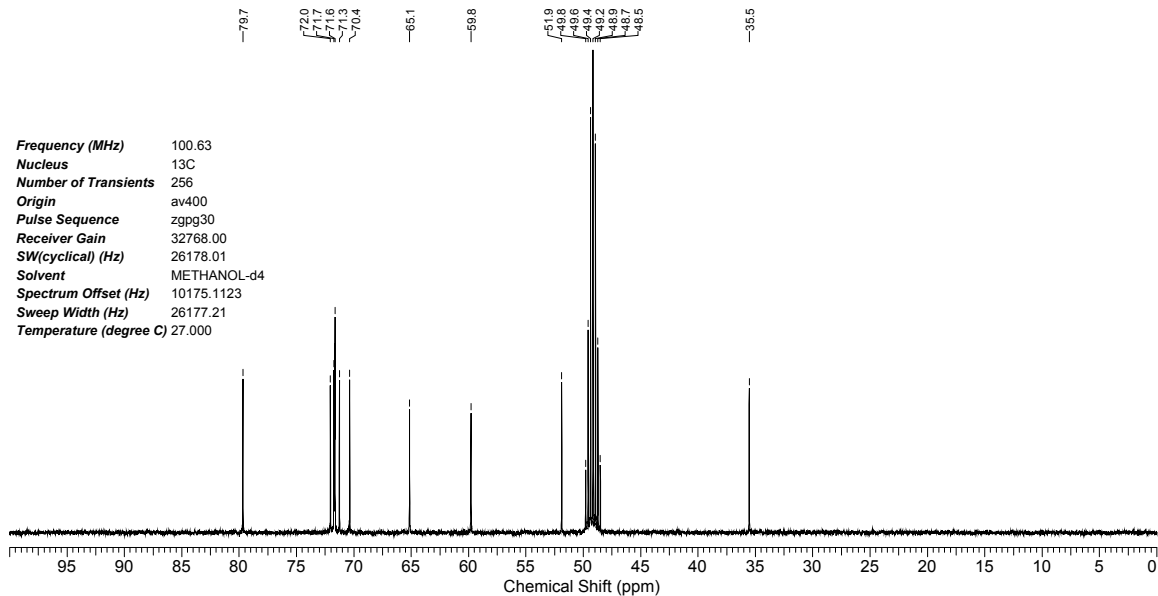
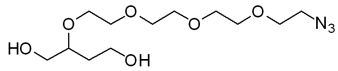
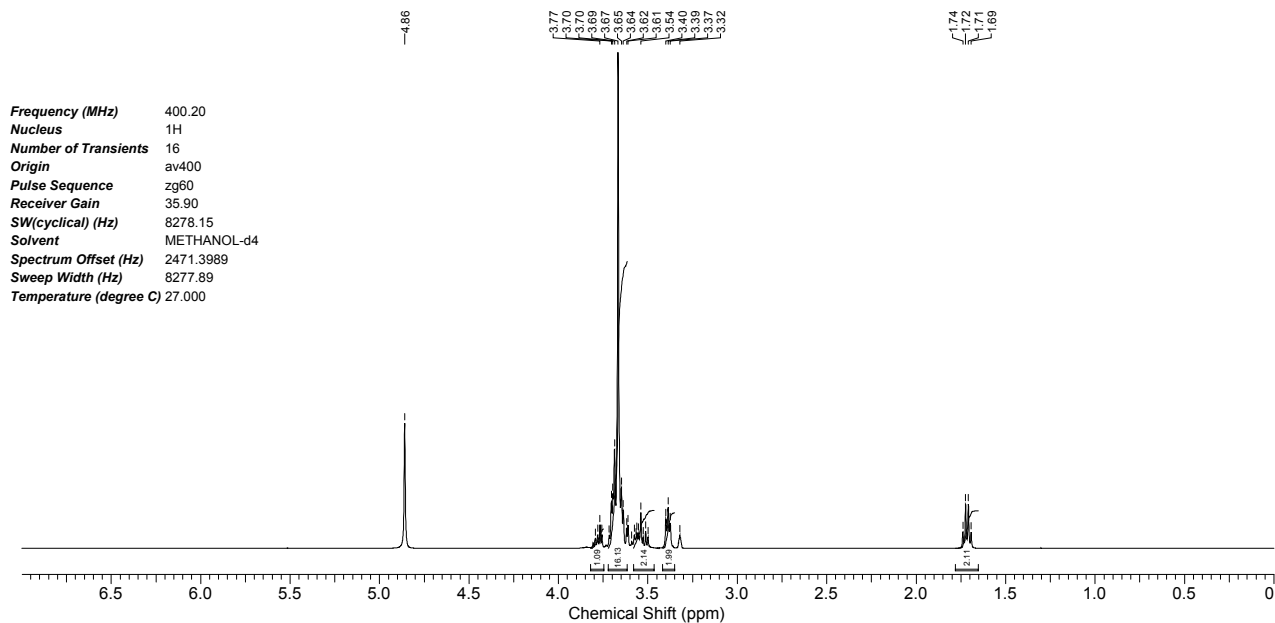
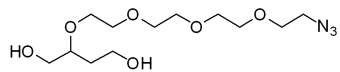
28.3

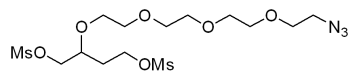
Frequency (MHz) 125.80
Nucleus 13C
Number of Transients 512
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12567.2148
Sweep Width (Hz) 31249.05
Temperature (degree C) 14.370



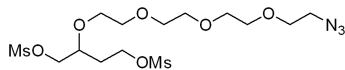
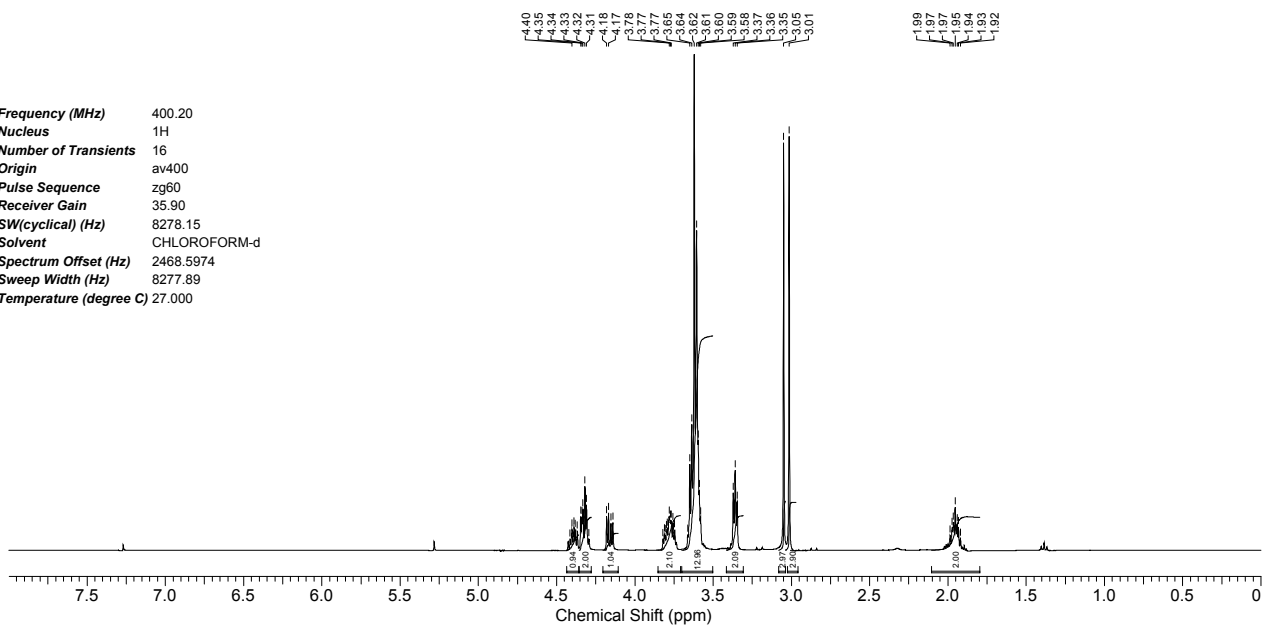




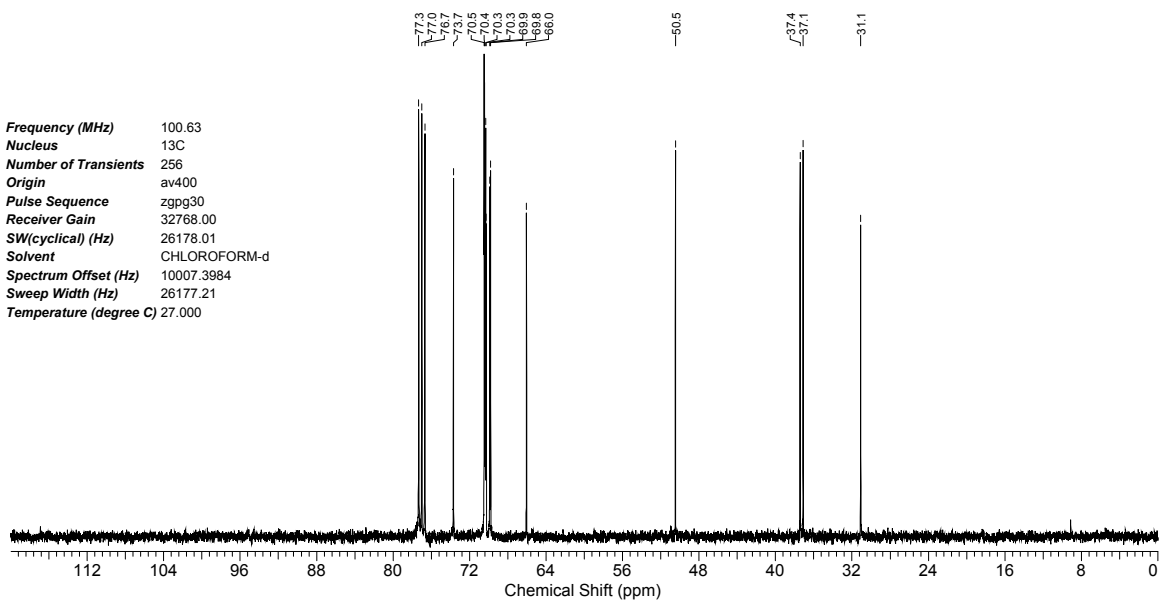


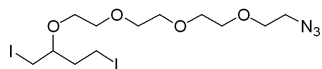


Frequency (MHz) 400.20
 Nucleus ¹H
 Number of Transients 16
 Origin av400
 Pulse Sequence zg60
 Receiver Gain 35.90
 SW(cyclical) (Hz) 8278.15
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 2468.5974
 Sweep Width (Hz) 8277.89
 Temperature (degree C) 27.000

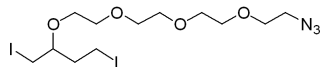
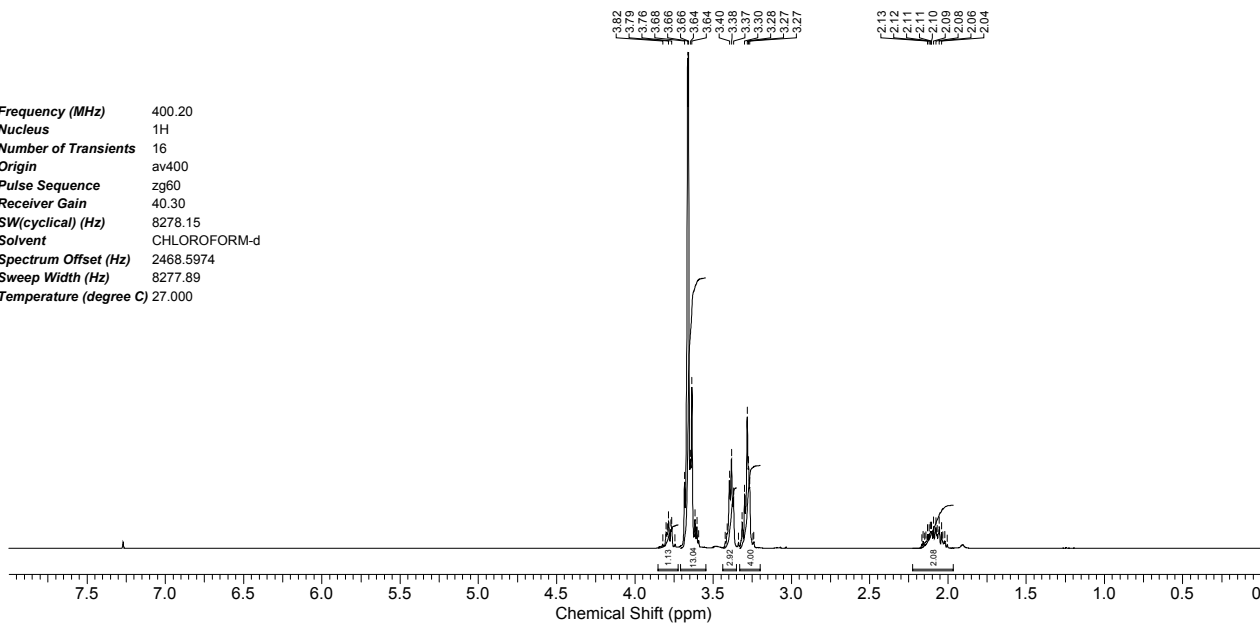


Frequency (MHz) 100.63
 Nucleus ¹³C
 Number of Transients 256
 Origin av400
 Pulse Sequence zgpg30
 Receiver Gain 32768.00
 SW(cyclical) (Hz) 26178.01
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 10007.3984
 Sweep Width (Hz) 26177.21
 Temperature (degree C) 27.000

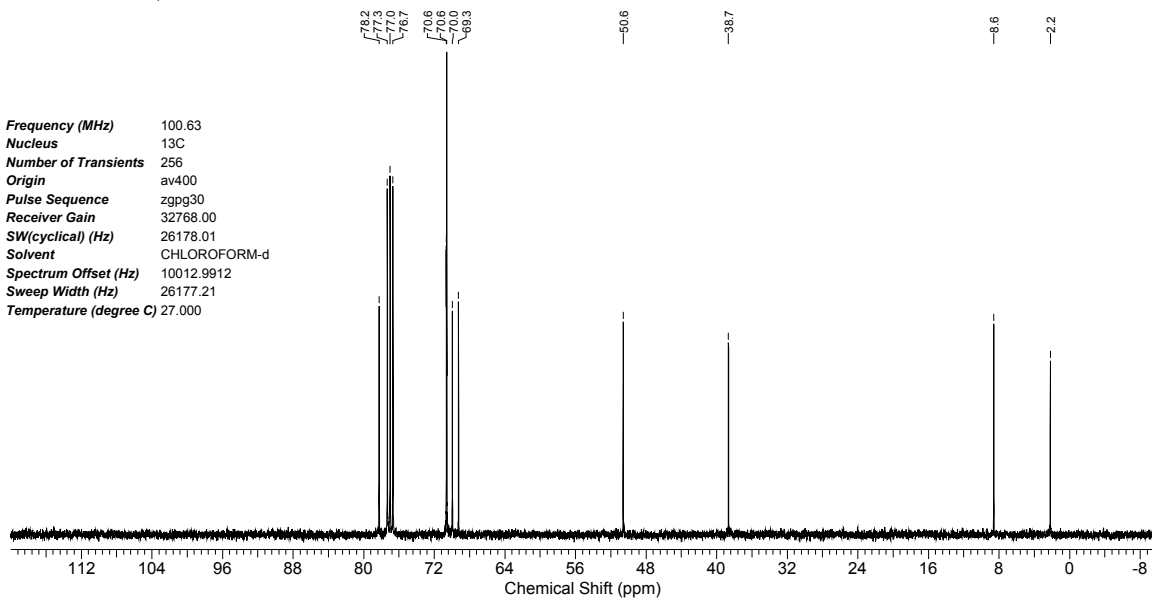


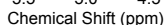


Frequency (MHz) 400.20
 Nucleus ^1H
 Number of Transients 16
 Origin av400
 Pulse Sequence zg60
 Receiver Gain 40.30
 SW(cyclical) (Hz) 8278.15
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 2468.5974
 Sweep Width (Hz) 8277.89
 Temperature (degree C) 27.000

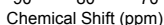


Frequency (MHz) 100.63
 Nucleus ^{13}C
 Number of Transients 256
 Origin av400
 Pulse Sequence zgpg30
 Receiver Gain 32768.00
 SW(cyclical) (Hz) 26178.01
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 10012.9912
 Sweep Width (Hz) 26177.21
 Temperature (degree C) 27.000

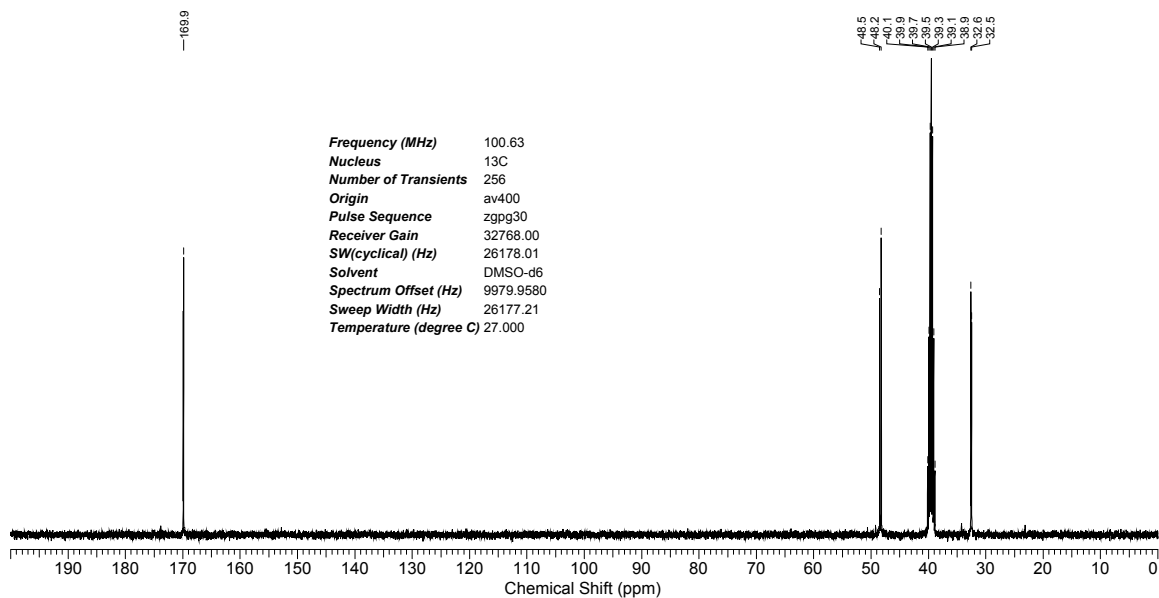
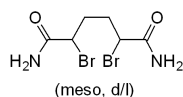
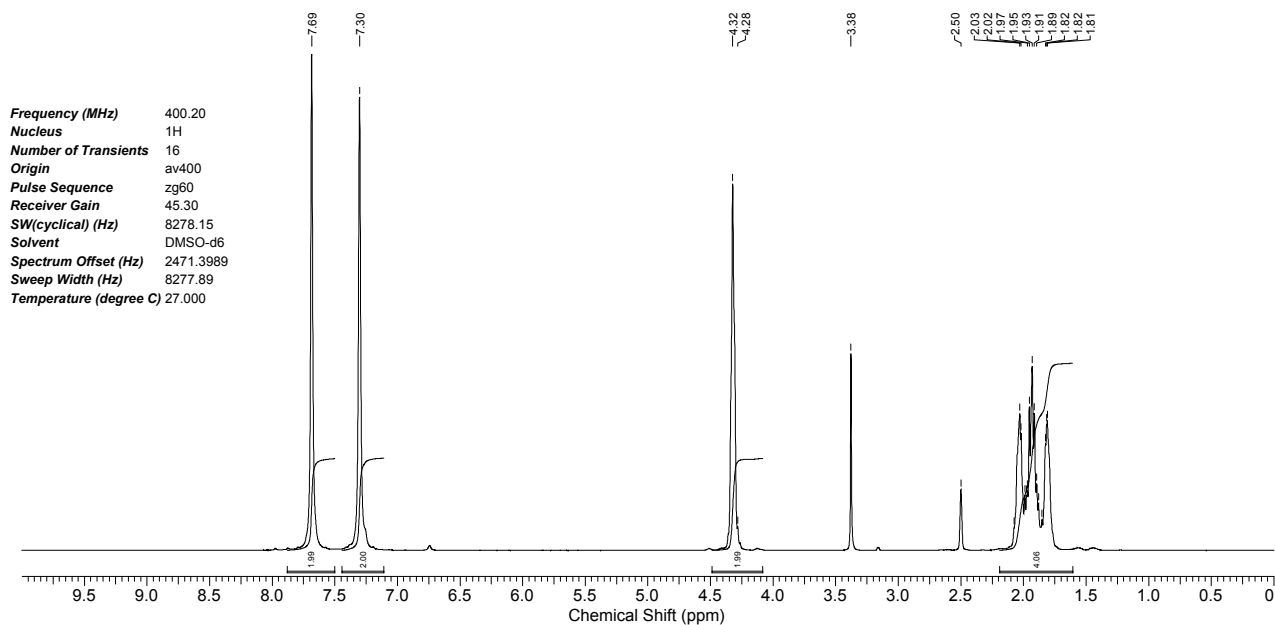
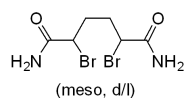


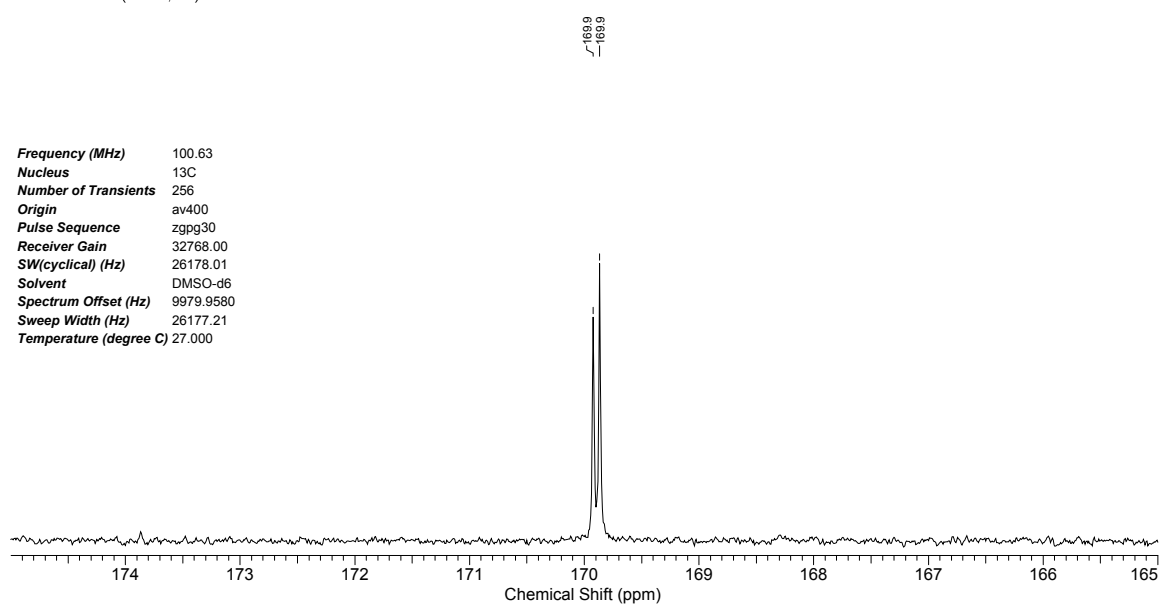
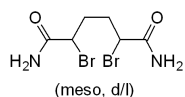
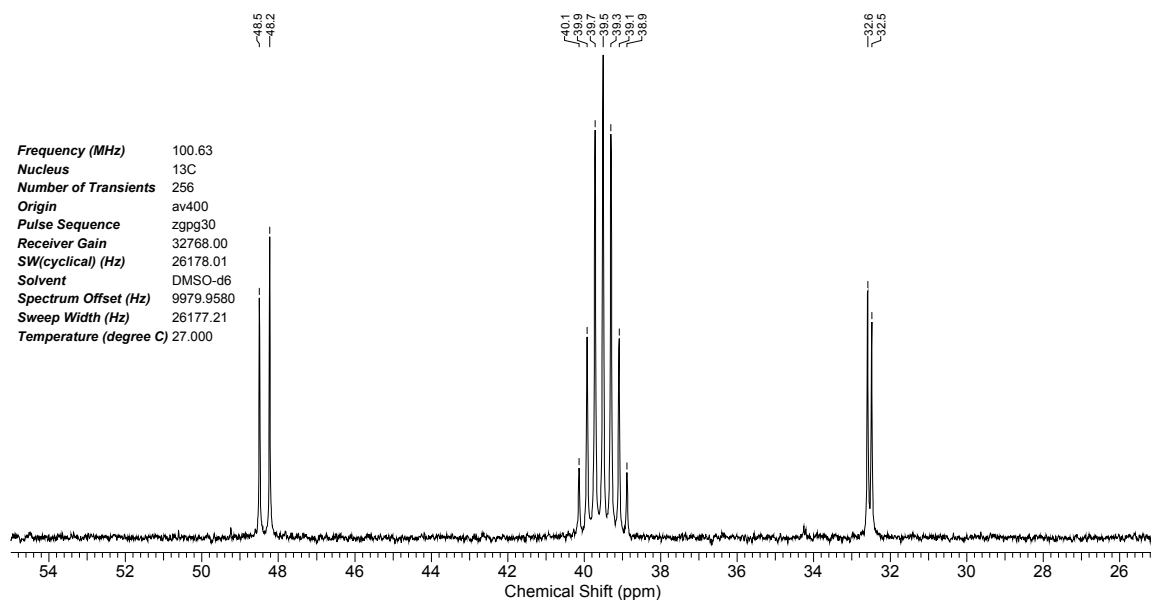
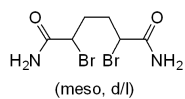


Frequency (MHz)	400.20
Nucleus	¹ H
Number of Transients	16
Origin	av400
Pulse Sequence	zg60
Receiver Gain	90.50
SW(cyclical) (Hz)	8278.15
Solvent	DMSO-d6
Spectrum Offset (Hz)	2471.3989
Sweep Width (Hz)	8277.89
Temperature (degree C)	27.00



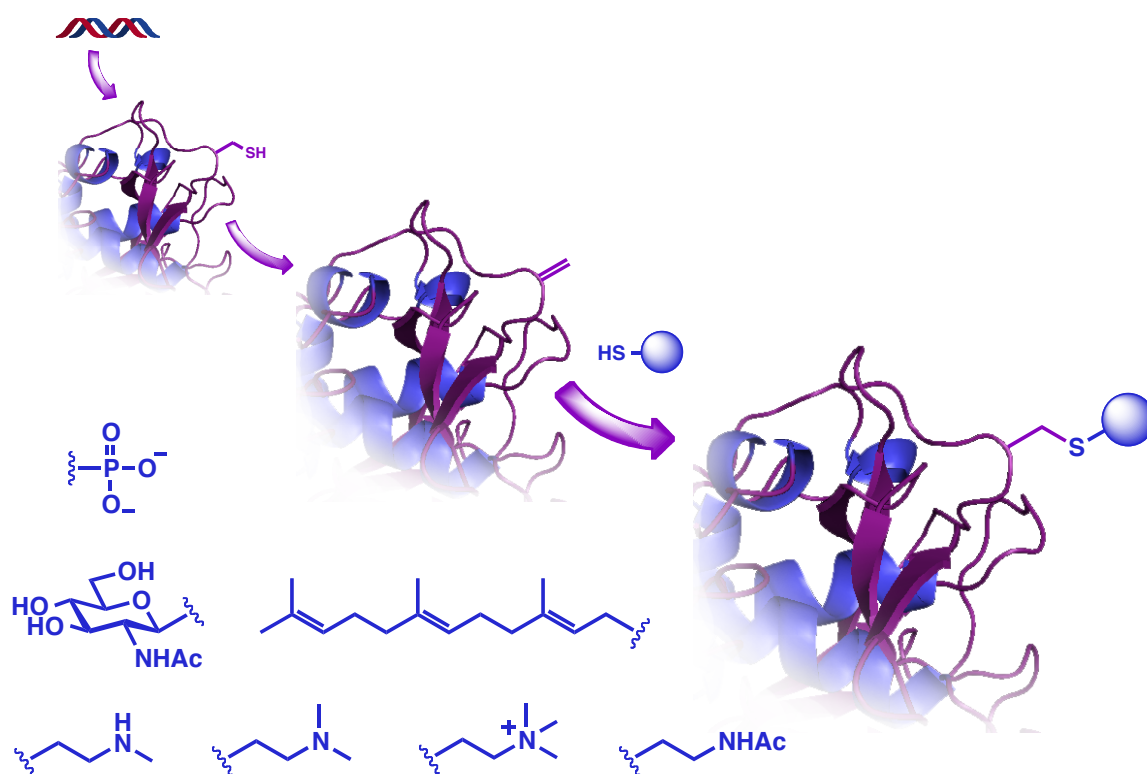
Frequency (MHz)	100.63
Nucleus	¹³ C
Number of Transients	256
Origin	av400
Pulse Sequence	zgpg30
Receiver Gain	32768.00
SW(cyclical) (Hz)	26178.01
Solvent	DMSO-d6
Spectrum Offset (Hz)	9975.16500
Sweep Width (Hz)	26177.21
Temperature (degree C)	27.000





Chapter 3

Conversion of Cysteine to Dehydroalanine Allows Access to Proteins Bearing Post-Translational Modifications



“What I cannot create, I do not understand”

Richard P. Feynman

California Institute of Technology, 1988

3.1 Introduction

Post-translational modification of proteins

After translation, proteins are often naturally tailored with various covalent modifications that dramatically expand their functional capacity.^{1,2} A selection of these post-translational modifications (PTMs) is shown in Figure 3.1.

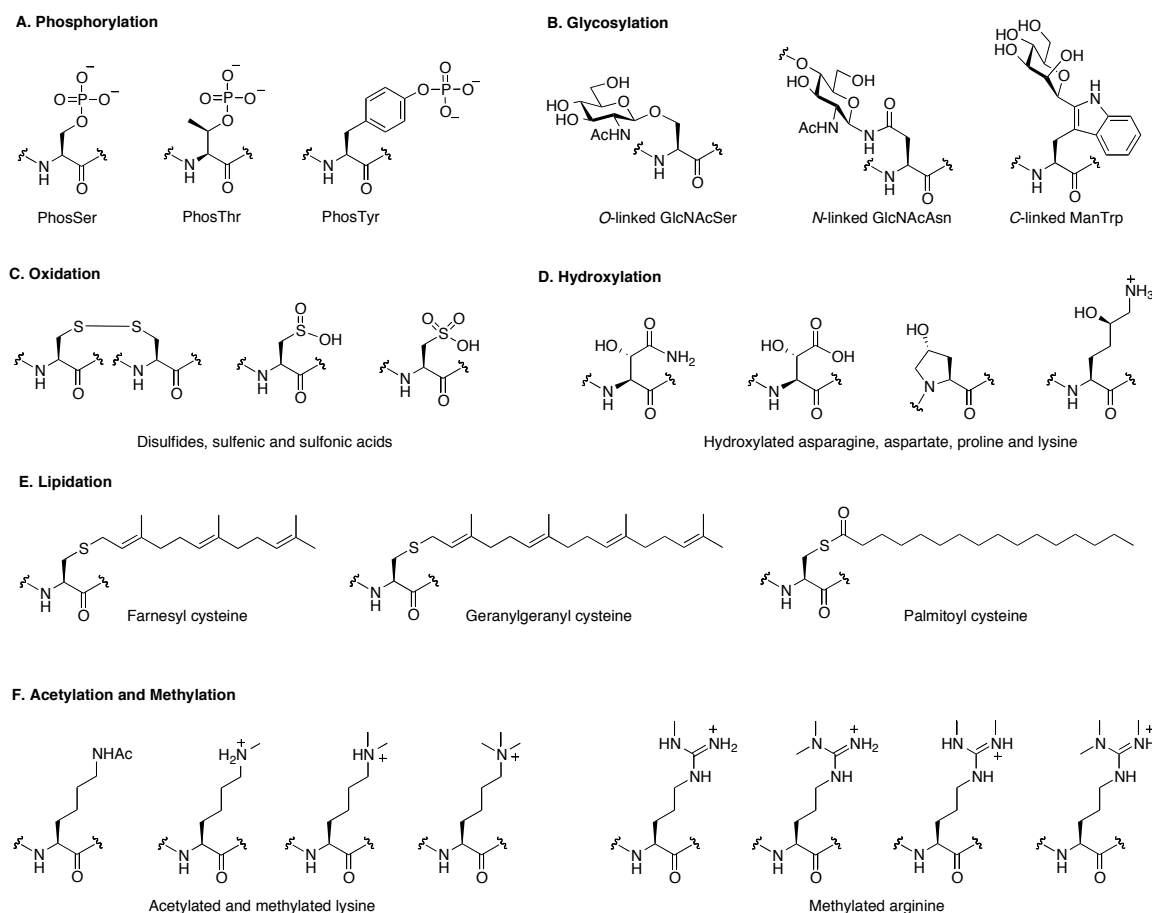


Figure 3.1: Selected examples of post-translational modifications

The functions of these modifications are critical in a host of cellular processes.^{1,2} Phosphorylation of serine, threonine, and tyrosine (Figure 3.1A) is a fundamental means of cellular signaling and influences many essential processes such as transcription.³⁻⁶ Understanding the complex network of kinases and phosphatases that mediate cellular signaling is a current frontier in chemical and systems biology.^{6,7} Glycosylation (Figure 3.1B) is also involved in diverse processes such as signaling, protein folding, and immune response.⁸⁻¹⁰ The oxidation (and reduction) of cysteine (Figure 3.1C) is a key process in metal binding and biocatalysis; disulfides are also essential structural elements in many

proteins.¹¹ Hydroxylation also contributes to protein structural stability and is widespread in collagen.¹² This PTM is also critical in oxygen sensing¹³ (Figure 3.1D). Lipidation of cysteine (Figure 3.1E) is a modification involved in protein-protein interactions, protein trafficking, and membrane localization.^{14,15} Lysine acetylation is a widespread modification¹⁶ and is an important regulator of cellular metabolism.¹⁷ Finally, lysine acetylation¹⁸ and methylation¹⁹ and arginine methylation²⁰ (Figure 3.1F) play important roles in chromatin biology.²¹

While the general functions of the modifications listed above have been established in various *in vitro* and *in vivo* assays, a full understanding of PTMs is often lacking. Many of the PTMs in Figure 3.1 are dynamic and are difficult to isolate in homogeneous form from natural sources.^{1,2} Therefore, a critical goal in contemporary protein science is the production of homogeneous, post-translationally modified proteins. Supply of these proteins enables the study of structure-function relationships and the consequences of post-translational modification on binding and activity.

In Chapter 2, dehydroalanine was used as a precursor to glycosylated proteins. The use of dehydroalanine as a precursor to polyprenylated peptides was also discussed. The addition of other nucleophiles to dehydroalanine may also provide access to other PTMs or their mimics. For instance, phosphorylated cysteine or analogs of acetylated or methylated lysine should, in principle, be accessible through dehydroalanine. To motivate the synthesis of these PTMs, previous reports of their installation on proteins will be discussed. This discussion will help highlight any strategic advantage or disadvantage of using dehydroalanine to access post-translational modifications.

Methods for the synthesis of glycoproteins

Glycoprotein synthesis by native chemical ligation

As discussed in Chapter 1, the synthesis of proteins by native chemical ligation (NCL)^{22,23} or expressed protein ligation (EPL)²⁴⁻²⁶ is a powerful method for accessing post-translationally modified proteins.^{27,28} Glycoproteins are no exception. In 1999, Bertozzi and Kent reported the application of NCL to the total chemical synthesis of the antimicrobial *O*-linked glycoprotein diptericin.²⁹ Two peptide fragments were synthesized; one as a *C*-terminal thioester and the other with an *N*-terminal cysteine. Each fragment contained an α -*N*-acetylgalactosamine (GalNAc) at threonine residues that are naturally glycosylated (Figure 3.2A). After coupling and purification, the synthetic

glycoprotein displayed antimicrobial activity comparable to the native protein. This report validated total synthesis as a route to homogeneous, functional glycoproteins.

Since this report, NCL and EPL have been applied in a number of glycoprotein syntheses. These syntheses include glycoproteins bearing multiple GalNAc residues^{30,31} and also proteins bearing complex *N*-linked glycans.³²⁻³⁶ These syntheses all feature the merger of *C*-terminal thioester and *N*-terminal cysteine peptide fragments characteristic to NCL and EPL. In a notable exception, Brik and co-workers used a serine-linked *N*-acetylglucosamine derivative bearing a thiol (rather than an *N*-terminal cysteine) to mediate the peptide coupling.³⁷ Subsequent desulfurization provided the *O*-GlcNAc glycopeptide (Figure 3.2B). In this way, the desired PTM is used to promote the desired coupling. Together, these ligation methods have provided homogeneous glycoproteins that have enabled the study of various biological activities,²⁹ protein folding,³⁸ and cancer therapies³⁹ in which glycoproteins play a central role.

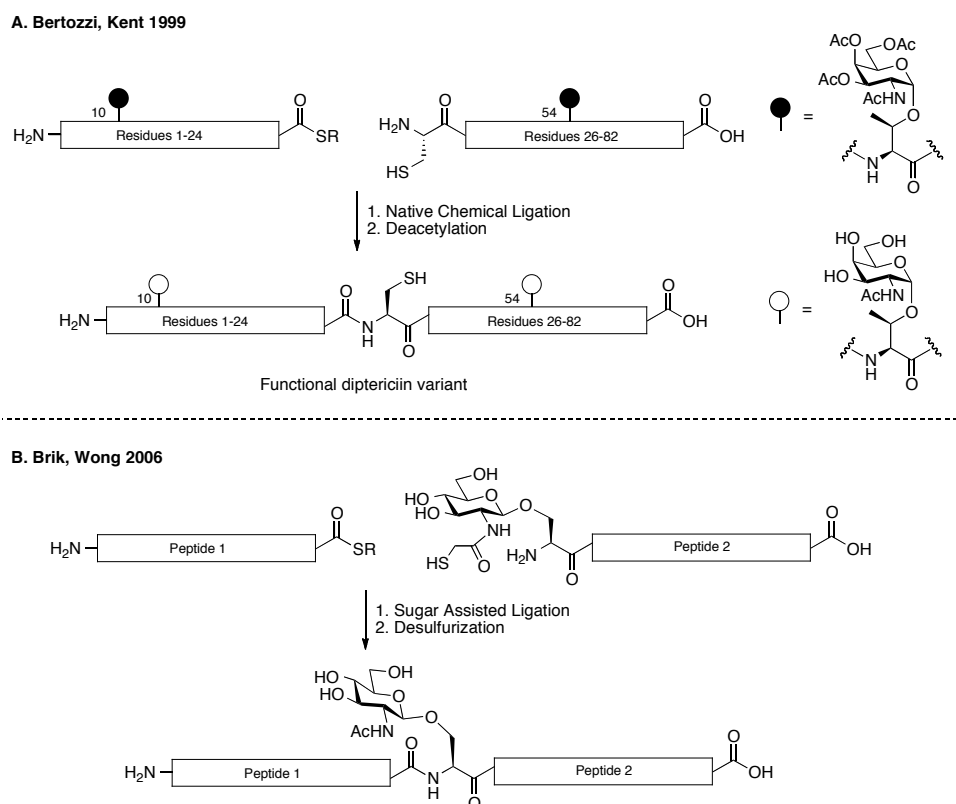


Figure 3.2: Selected examples of glycoprotein synthesis by native chemical ligation

While NCL and EPL offer the atomic level precision of chemical synthesis, these methods necessitate the synthesis and purification of the required peptide fragments that may prove technically challenging. Proteins that bear multiple modifications that are

separated by more than 50 residues, and so require multiple ligations, further compound this technical demand. For these reasons, other approaches to glycoprotein synthesis have been reported that may complement total chemical synthesis. Recombinant translation, enzymatic glycosylation, and site-selective chemical modification are three such strategies.

Glycoprotein synthesis using recombinant expression systems

While total chemical synthesis of proteins by NCL and EPL is a viable route to glycoproteins, purely biological methods for glycoprotein synthesis have recently emerged. For example, Gerngross and co-workers have genetically engineered the yeast strain *Pichia pastoris* to carry out protein glycosylation characteristic of mammalian species.⁴⁰⁻⁴² This technology has allowed large scale fermentation for the production of proteins with sialylated complex *N*-glycans. Erythropoietin⁴⁰ and glycosylated antibodies⁴¹ have both been produced by these “humanized” yeast strains (Figure 3.3A).

Very recently, the glycosylation machinery of *Campylobacter jejuni* was successfully transferred to *E. coli* hosts. Schwarz and co-workers demonstrated that with judicious engineering, this system could produce proteins bearing *N*-glycans—including the ubiquitous mammalian Asn-GlcNAc linkage (Figure 3.3B).⁴³ To produce homogeneous glycoproteins, the product after expression was trimmed and remodeled *in vitro* with various carbohydrate processing enzymes.⁴³ The combination of high fidelity expression systems and efficient enzymatic remodeling is a promising new development in the synthesis of glycoproteins.

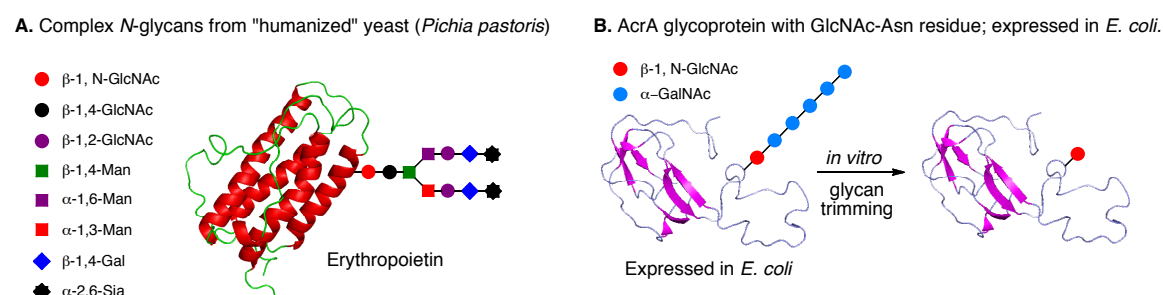


Figure 3.3: Representative glycoproteins accessible through expression in engineered yeast^{40,41} (A) and *E. coli*⁴³ (B).

Glycoprotein synthesis using enzymatic glycosylation

The use of enzymes to remodel and install carbohydrate domains is a viable strategy for glycoprotein synthesis.⁴³ Takegawa and co-workers have shown that endo- β -*N*-acetylglucosaminidase from *Arthrobacter protophormiae* (Endo A) can catalyze the transglycosylation of *N*-linked carbohydrates from one substrate to another.⁴⁴ The Wong laboratory has also shown that several enzymes can be used to remodel *N*-linked glycans, as demonstrated in their enzymatic installation of sialyl Lewis X on RNase A.⁴⁵ A significant advance in enzymatic glycoprotein synthesis was the discovery that Endo A catalyzes the transfer of synthetic sugar oxazoline donors to proteins bearing GlcNAc asparagine (Figure 3.4).⁴⁶ The Wang laboratory has expanded on this discovery and synthesized many glycoproteins,⁴⁷ including homogenously glycosylated antibodies.⁴⁸ Further discoveries regarding the transglycosylation activity of the endo- β -*N*-acetylglucosaminidase from *Mucor hiemalis* (Endo M)⁴⁹ and the engineering of highly active mutants of both Endo A and Endo M have broadened the scope of this method. In particular, these mutant glycosidases favor transferase activity over their innate glycosidase activity (the reverse reaction).⁵⁰⁻⁵²

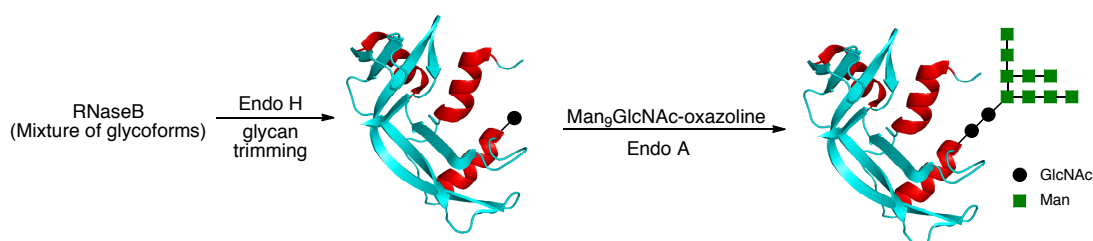


Figure 3.4: Representative example of the enzymatic synthesis of glycoproteins⁵²

The main limitation of the strategy depicted in Figure 3.4 is that the site of glycosylation (i.e. the site of GlcNAc-Asn residues) cannot be easily manipulated. In fact, it is determined entirely by source organism and its endogenous glycosyl transferases. Should a different or non-natural site be desired, the requisite GlcNAc-Asn residue (or a suitable mimic) must be installed chemically. These chemical methods are discussed next.

Glycoprotein synthesis using site-selective chemical glycosylation

The direct chemical glycosylation of proteins requires a uniquely reactive residue to ensure selectivity. While many unnatural residues have been used to install carbohydrates, these transformations often result in linkages that do not resemble the

native structure.^{53,54} In this survey, the examples discussed are those that give access to a natural glycosylated residue or structurally similar analog.

Cysteine has been targeted to form disulfide-linked glycoproteins. Cysteine has a relatively low natural abundance in proteins and can be genetically incorporated at the desired site of modification. The Davis group has introduced several procedures for the chemoselective sulfenylation of cysteine, allowing access to these glycoproteins.⁵⁵⁻⁵⁸ These methods constituted some of the earliest examples of site-specific chemical glycosylation of proteins (Figure 3.5A).

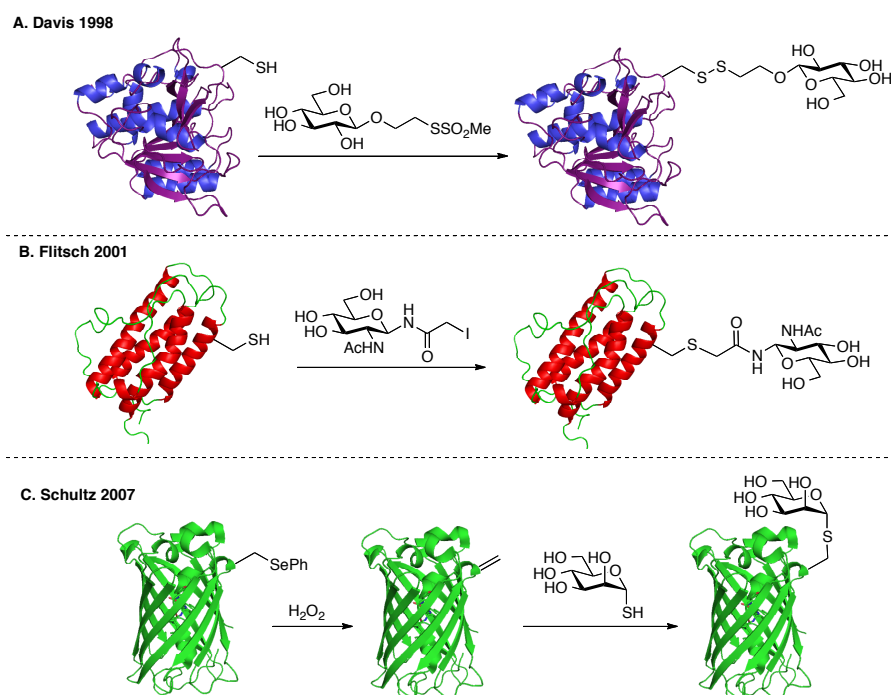


Figure 3.5: Selected examples of chemical glycosylation of proteins

For *N*-linked glycoproteins, site-selective coupling to an asparagine or aspartic acid residue is complicated by the high abundance of these residues on the surface of a protein. To avoid this complication, Flitsch has introduced a cysteine alkylation procedure that results in a structural analog of *N*-linked glycosylation (Figure 3.5B). Flitsch and co-workers took advantage of this strategy to chemically access homogeneously glycosylated erythropoietin.⁵⁹

As discussed in Chapter 2, dehydroalanine is a precursor to glycosylated cysteine residues. These residues are naturally occurring in various peptides⁶⁰⁻⁶² and their synthesis from dehydroalanine has been shown on peptides⁶³⁻⁶⁶ and proteins.⁶⁷ For proteins, the lone example of using dehydroalanine as a precursor to glycosyl cysteines

Lipidated proteins by site-selective chemical modification

Site-selective chemical modification has complemented EPL and NCL in the synthesis of lipidated proteins. One of the earliest reports on the synthesis and study of lipid-modified proteins relied on selective alkylation of a *C*-terminal cysteine (Figure 3.7A).⁷⁴ In this report, Bader and co-workers showed that the synthetic farnesylated and palmitoylated H-Ras proteins behaved as the natural counterparts in their interaction with membranes and activity in rat pheochromocytoma cell lines.⁷⁴

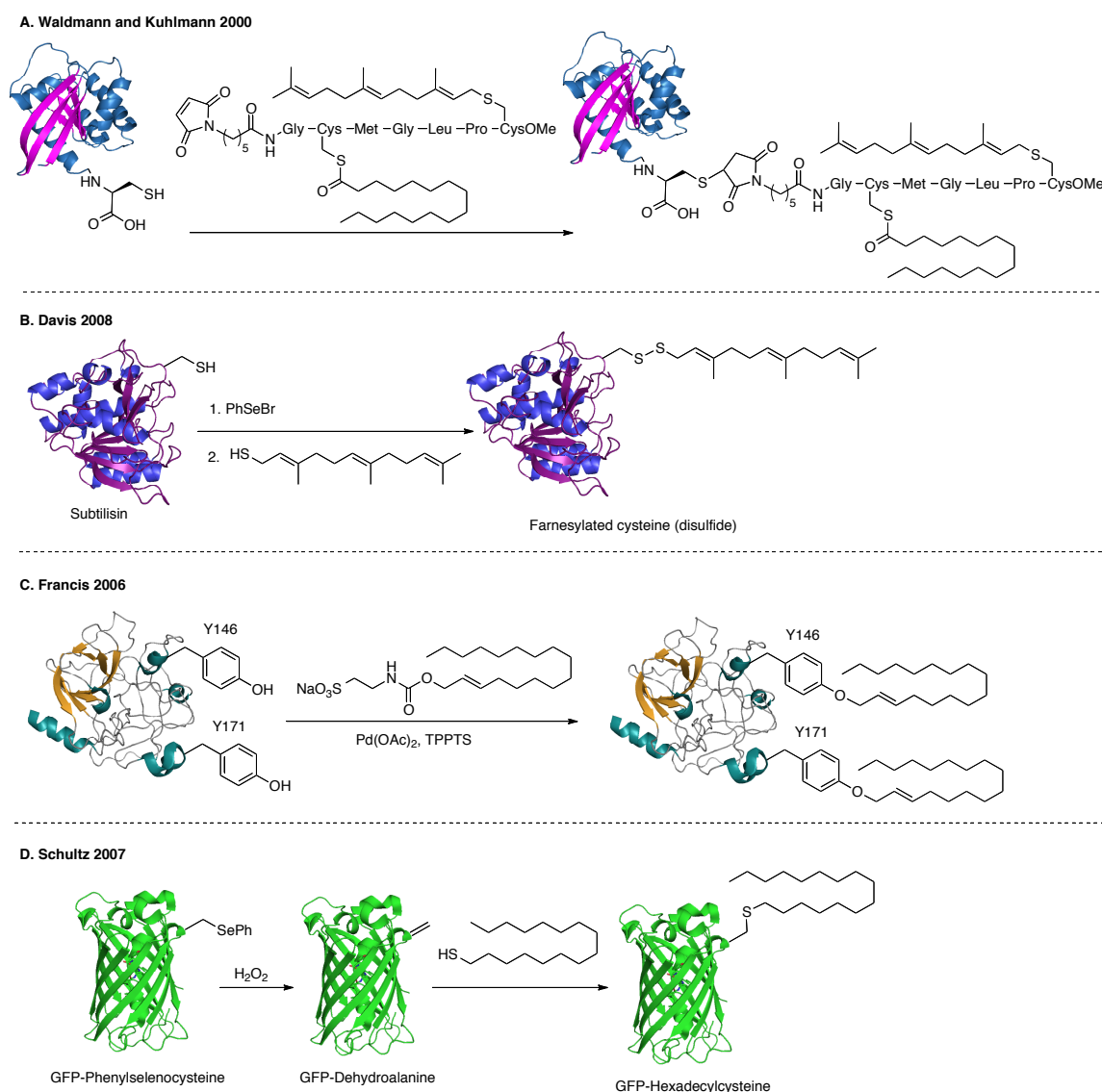


Figure 3.7: Representative examples of site-selective chemical lipidation of proteins

In another example of cysteine modification, Davis and co-workers reported the selective sulfenylation of cysteine to give the disulfide analog of farnesyl cysteine.⁷⁵ This

modification was carried out on a single-cysteine protease which was shown to retain enzymatic activity after modification (Figure 3.7B)⁷⁵

Targeting a different residue for selective lipidation, Francis and co-workers reported the selective *O*-alkylation of tyrosine side chains using π -allylpalladium electrophiles (Figure 3.7C).⁷⁶ Notably, no organic solvent was required in this modification because of the use of a water soluble leaving group on the lipidated allyl carbamate (the precursor to the π -allylpalladium intermediate) and a water soluble phosphine ligand (TPPTS). Chymotrypsinogen A was modified using this chemistry and the lipidated proteins were shown to associate with the lipid bilayer of unilamellar vesicles.⁷⁶

The Schultz laboratory has used dehydroalanine as a precursor to lipidated proteins.⁶⁷ Dehydroalanine was installed by the oxidative elimination of phenylselenenylcysteine. Phenylselenenylcysteine was genetically incorporated into green fluorescent protein in response to the amber stop codon using a mutant tyrosyl tRNA / tRNA synthetase pair.⁶⁷ By the conjugate addition of hexadecylthiol to dehydroalanine, a non-hydrolyzable analog of palmitoylated cysteine was obtained (Figure 3.7D).

Methods for the synthesis of phosphorylated proteins

Phosphorylated proteins are difficult to access in homogeneous form by recombinant expression or through the use of kinases *in vitro*.⁶ The site and level of phosphorylation is difficult to control by these methods and for *in vitro* methods the kinase itself often must be phosphorylated itself, adding another level of complexity to the synthesis.⁶ Furthermore, phosphorylation is dynamic and susceptible to hydrolysis (both chemically and enzymatically), prompting the development of methods for the chemical synthesis of phosphorylated proteins and analogs.⁶ These methods rely almost exclusively on native chemical ligation and expressed protein ligation, though a few methods for the chemical phosphorylation of proteins have been reported. Both approaches are described below.

Synthesis of phosphorylated proteins by native chemical ligation

In the first report of expressed protein ligation (EPL) by Muir, Sondhi, and Cole; the first chemical synthesis of a phosphorylated protein was reported.²⁴ A peptide containing an *N*-terminal cysteine and a phosphorylated tyrosine was ligated to a *C*-terminal thioester of the protein tyrosine kinase Csk (Figure 3.8A). A fluorescein label was also appended to

the peptide to facilitate analysis. The semi-synthetic protein was used to determine the effect of tyrosine phosphorylation on conformation and catalytic activity in Csk and related kinases.²⁴ Since this report, EPL has been used extensively to access many proteins,^{25,26} including phosphorylated proteins. For instance, the synthesis of proteins containing multiple phosphoserine and phosphothreonine residues by EPL⁷⁷⁻⁷⁹ and NCL⁸⁰ was critical in the study of transforming growth factor- β signaling. Additionally, NCL has provided access to phosphorylated histone proteins.^{81,82} These modified proteins are critical in the study of chromatin biology.⁸³⁻⁸⁶

A distinct advantage of using NCL and EPL over traditional expression systems or other enzymatic modification of proteins is the ability to incorporate unnatural analogs of amino acids for functional assays. For instance, Cole and co-workers have used non-hydrolyzable mimics of phosphotyrosine⁸⁷⁻⁸⁹ and phosphoserine/threonine,^{90,91} each installed by EPL, to assess the effects of phosphorylation *in vitro* and *in vivo* (Figure 3.8B). In a related strategy, the Muir lab has recently reported stable analogs of phosphohistidine, a particularly fragile and poorly understood PTM.⁹² These phosphohistidine analogs are compatible with solid phase peptide synthesis and were incorporated into peptide thioesters ultimately used in the synthesis of histones bearing phosphorylated histidines (Figure 3.8C). Together, these studies eliminate the complication of phosphate hydrolysis and provide more direct evidence of phosphorylation function than classic point mutation approaches.

Another advantage of NCL and EPL is the capacity to install latent or masked post-translational modifications in protein structures. Muir and co-workers have used such a strategy to incorporate photocaged phosphoserine into the cellular signaling protein Smad2 (Figure 3.8D).⁹³ The temporally controlled uncaging by irradiation with 325 nm light allowed direct observation of Smad2 trimerization as well as nuclear localization—two processes dependent on serine phosphorylation at the caged sites.

Other developments in the use of NCL to access phosphorylated proteins have been reported recently. For example, *N*-terminal phosphoserine residues have been shown to promote the ligation with thioesters.⁹⁴ This “phosphate-assisted ligation” means that the post-translational modification itself can mediate the peptide coupling (Figure 3.8E). While this strategy has yet to be demonstrated in the synthesis of phosphorylated proteins, it is noteworthy that it does not require cysteine at the site of ligation. Another development is the synthesis of thioester peptides that contain phosphorylated residues. Unlike the use of EPL in which the thioester contains only unmodified residues, it might

be necessary in some cases to synthesize the thioester with phosphorylated side chains. Macmillan and co-workers have explored conditions that accommodate both phosphoserine and thioester synthesis in the same peptide.⁹⁵

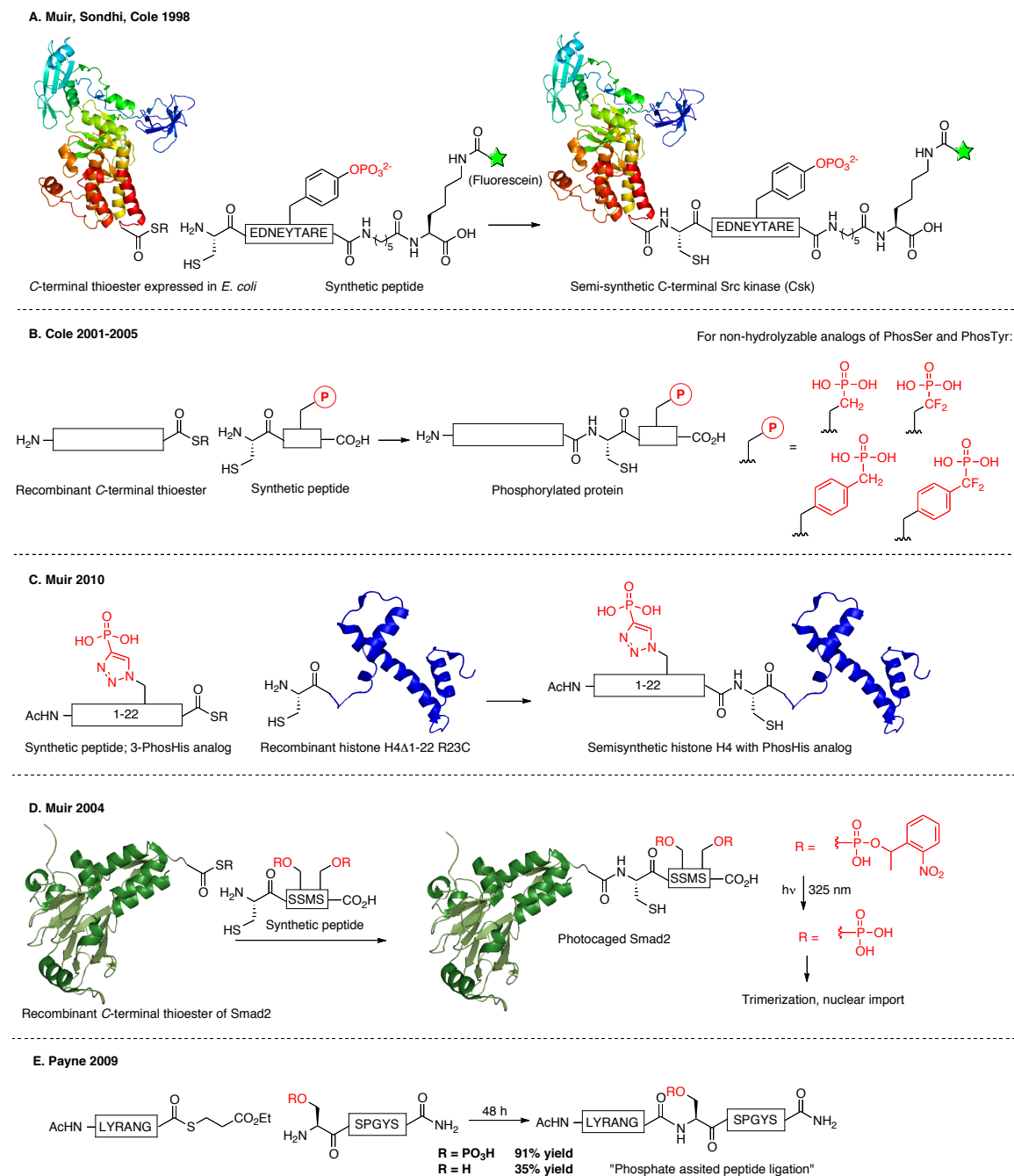


Figure 3.8: Phosphorylated proteins by EPL and NCL

Synthesis of phosphorylated proteins by site-selective chemical modification

Comparatively few methods for the chemical phosphorylation of proteins exist. The formation of a disulfide between cysteine and thiophosphate has been used as an analog of phosphorylated aspartate (Figure 3.9A).^{96,97} This modification was used in the study of bacterial chemotaxis proteins CheB⁹⁷ and CheY.⁹⁶ For analogs of phosphotyrosine, the Hackenberger labs have used the Staudinger-Phosphite reaction at *p*-azidophenylalanine.⁹⁸ The combined use of non-natural amino acid incorporation and chemical modification was used to install the phosphotyrosine mimics as well as photocaged analogs (Figure 3.9B).⁹⁸ While these methods are promising tools for the study of their respective modifications, the relative paucity of chemical modifications useful for the study of such a common modification is striking. In particular, reliable chemical modifications to mimic phosphoserine/threonine have not been reported—a shortcoming that will be addressed in this Thesis.

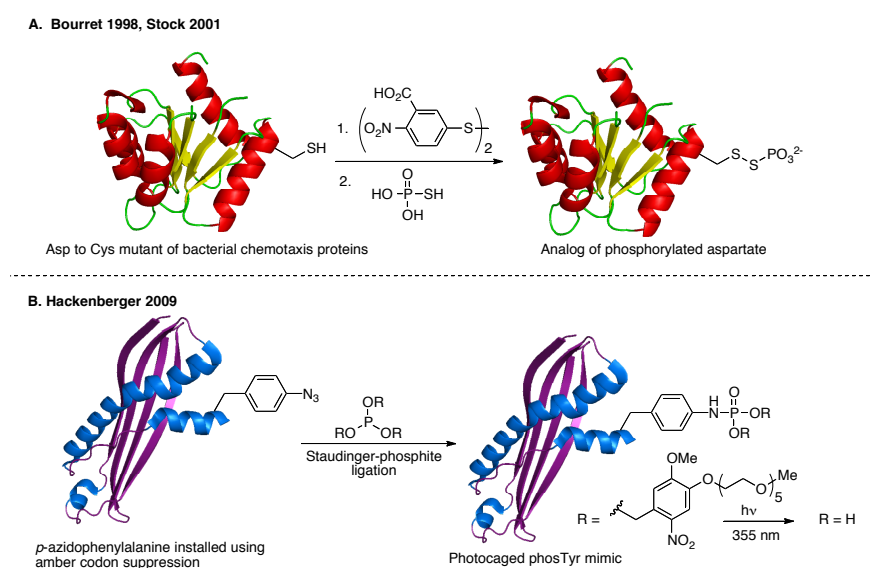


Figure 3.9: Site-selective chemical protein phosphorylation

Methods for the synthesis of proteins containing methylated and acetylated lysine residues

As shown in Figure 3.1, lysine side chains are subject to several post-translational modifications.^{1,2} Among these modifications, lysine acetylation¹⁸ and methylation¹⁹ play essential roles in chromatin biology. These modifications occur widely on histone proteins²¹ and their role in epigenetics will be discussed in more detail in Chapter 4. For the present discussion, it suffices to say that accessing proteins bearing this modification

will help elucidate their structural and functional roles.⁸³⁻⁸⁶ In this Chapter and in Chapter 4, it will be shown that dehydroalanine is a precursor to analogs of methylated and acetylated lysine residues. In order to evaluate the advantages and limitations of this chemistry, alternative methods for accessing proteins with methylated and acetylated lysine residues are reviewed next.

Synthesis of proteins containing methylated and acetylated lysine by NCL and EPL

Like other proteins bearing post-translational modifications, those with methylated and acetylated lysine residues can be prepared by native chemical ligation and expressed protein ligation.⁸³ Because these modifications and their role in chromatin biology are a current frontier in epigenetics,⁸⁴ they have been synthesized and investigated most commonly on histone proteins. For histones, the *N*-terminal tails are densely modified with several modifications,²¹ including lysine methylation and acetylation.^{18,19} Native chemical ligation is well poised to access these modified proteins by the merger of a synthetic peptide bearing the modifications (and a *C*-terminal thioester) with a truncated histone possessing an *N*-terminal cysteine. McCafferty and co-workers have employed such a strategy to synthesize several histone 3 and histone 4 proteins that contain acetylated or methylated lysine residues.⁹⁹ Two examples are shown in Figure 3.10A. The ligation site was selected so that post-ligation desulfurization¹⁰⁰ would provide alanine and the native protein sequence. These synthetic proteins were used for *in vitro* chromatin reconstitution experiments and also as substrates for histone modifying enzymes.⁹⁹ Related strategies for the synthesis of acetylated histones have since been reported for the study of chromatin assembly¹⁰¹ and the interaction of acetylated histones with chromatin remodeling enzymes.¹⁰²

NCL can also supply histones bearing multiple post-translational modifications. The Muir lab, for instance, has synthesized a histone H2B variant that contains both acetylated lysine and phosphorylated serine on the *N*-terminal tail (Figure 3.10B).⁸² Post-ligation desulfurization^{100,103} was also used in this example for a formal ligation at alanine. In this report it was shown that when lysine 5, 11, 12, and 15 were acetylated, antibodies for phosphorylated serine 14 did not bind. This observation highlights the caution required in interpreting antibody-based assays that may be influenced not only by the modification to which they were raised, but also by the modification of neighboring residues.

For accessing histones that are modified closer to the C-terminus, EPL has proven particularly useful. Poirier, Ottesen and co-workers have employed such an approach to synthesize histone H3 with acetylation at lysine 115 and 122.¹⁰⁴ Histone 3 contains a single cysteine at position 110 and served as a convenient site for ligation (Figure 3.10C). The histones obtained were used in DNA binding studies. Interestingly, the results obtained with the synthetic acetylated histones differed from histones where these lysine residues were mutated to glutamine (a mutation sometimes used to mimic lysine acetylation). This observation highlights the utility and precision of NCL and EPL in accessing modified proteins.

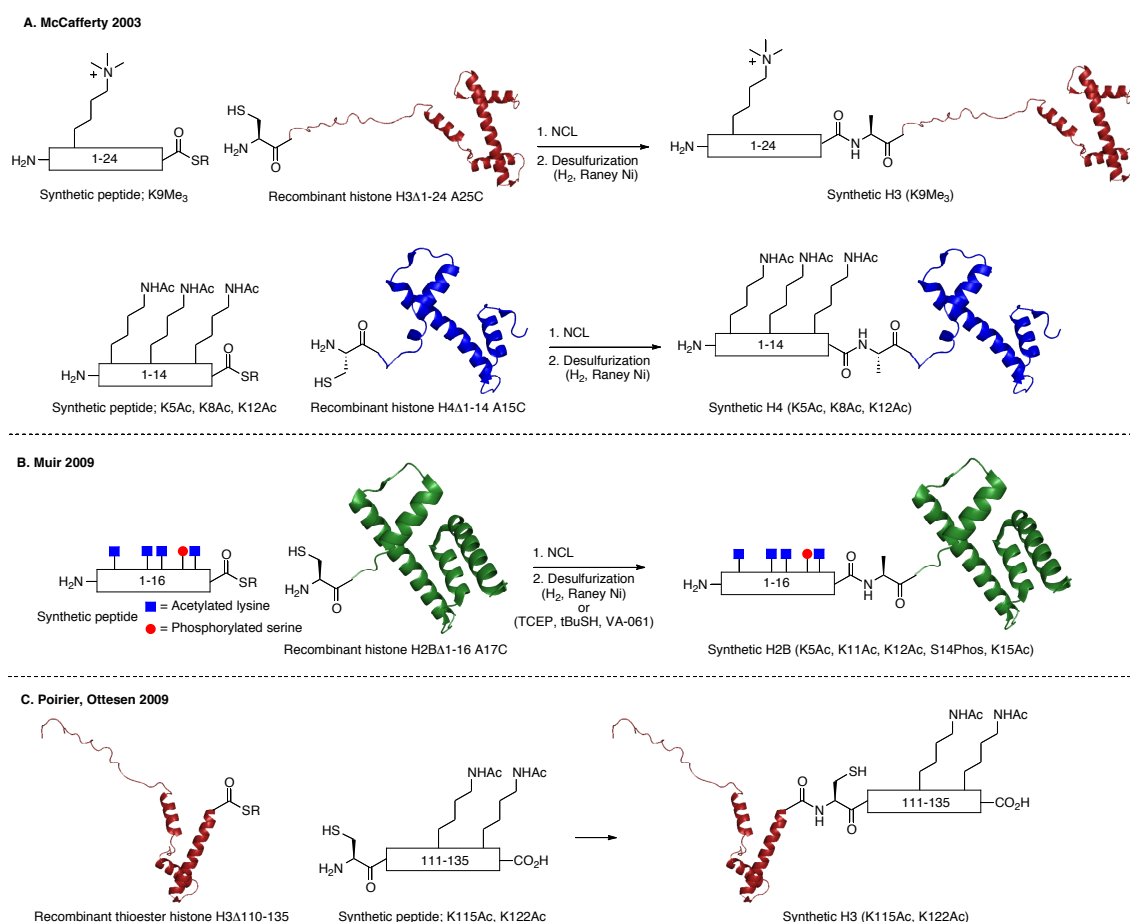


Figure 3.10: Methylated and acetylated proteins by NCL and EPL

Genetic incorporation of methylated and acetylated lysine residues into proteins

The genetic incorporation of methylated and acetylated lysine residues was recently made possible by the discovery that the pyrrolysyl tRNA / tRNA synthetase pair from *Methanosarcina barkeri* can be evolved to incorporate related amino acid derivatives into proteins in response to the amber stop codon.¹⁰⁵ This technology, developed largely by

the Chin,^{105,106} Schultz,^{107,108} and Liu groups,¹⁰⁹ relies on the fortuitous use of the UAG codon by methanogenic bacteria for the incorporation of pyrrolysine (Figure 3.11A) into proteins. Noting that acetyl lysine is a substructure of pyrrolysine (Figure 3.11A), Chin and co-workers evolved the synthetase to recognize acetyl lysine specifically, and incorporated this residue into sperm whale myoglobin (Figure 3.11A) and manganese superoxide dismutase.¹⁰⁵ It was also found during these studies that the *E. coli* deacetylase enzyme CobB must be inhibited to prevent post-translational deacetylation of the lysine of interest.¹⁰⁵ The Chin lab has since used this technology to access acetylated histones¹¹⁰ and study the effect of lysine acetylation on cyclophilin A catalysis.¹¹¹

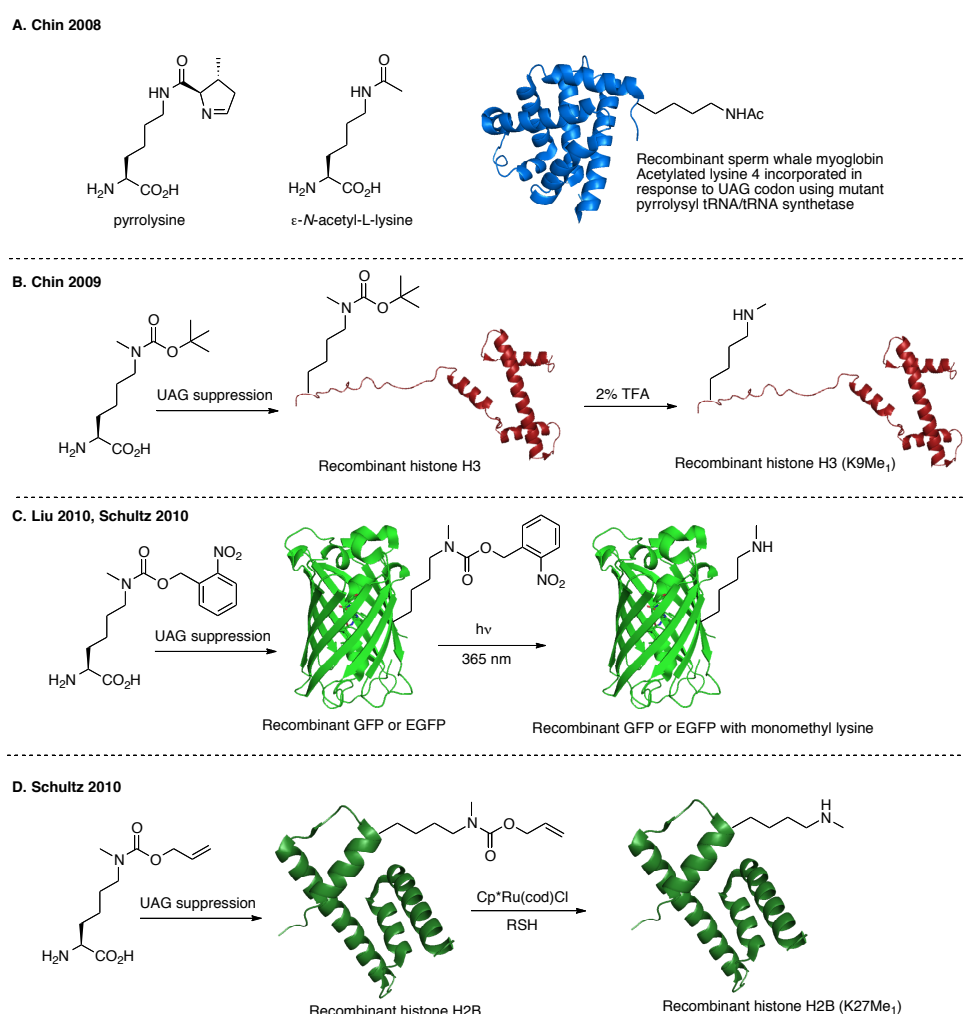


Figure 3.11: Mutant pyrrolysyl tRNA / tRNA synthetase pairs can be used to genetically incorporate acetylated and monomethylated lysine in response to the amber stop codon

Following a similar strategy, Chin and co-workers have also reported the incorporation of monomethylated lysine using a mutant pyrrolysyl tRNA / tRNA synthetase pair.¹⁰⁶ The incorporation, however, is indirect—only the Boc protected derivative can be used. For acid-stable histone proteins, deprotection with TFA provides the monomethyllysine residue (Figure 3.11B).¹⁰⁶ Other latent monomethyllysine derivatives have been installed in a similar fashion by the Schultz^{107,108} and Liu groups¹⁰⁹ (Figure 3.11 C-D). The use of a photocaged monomethyllysine derivative (Figure 3.11C) is notable in that the monomethyllysine residue can be revealed *in vivo*.^{108,109} Unfortunately, the pyrrolysyl aminoacyl-tRNA synthetase has not yet been evolved to accommodate the charged side chain of trimethyllysine and dimethyllysine can only be incorporated indirectly and necessitates several chemical steps.¹¹² Therefore, to access these residues, one must resort to NCL, EPL, or one of the site-selective chemical methods discussed next.

Site-selective chemical installation of methylated and acetylated lysine residues

A long-standing method for the chemical incorporation of lysine into proteins is the aminoethylation of cysteine. More than half a century ago, cysteine was alkylated with aziridine and related electrophiles to access these lysine mimics. The chemical conversion of cysteine to these lysine analogs imparted additional proteolytic sites for trypsin, greatly facilitating protein sequencing.¹¹³ Since this time, aminoethylation of cysteine has been used to investigate catalytic roles of lysine¹¹⁴ and most recently as a method to access methylated lysine analogs.¹¹⁵ The Shokat lab has demonstrated the ease and utility of this reaction in accessing histone proteins containing methylated lysines (Figure 3.12A).¹¹⁵ Fortuitously, histone proteins contain few (if any) cysteine residues, allowing a single cysteine to be genetically incorporated at the site of interest. Facile alkylation with the appropriate electrophile provides a mimic of mono-, di-, or trimethyllysine. While this method is typically used to access proteins with a single modified lysine, these proteins have nonetheless proven useful in several studies of chromatin structure and associated biology.¹¹⁶⁻¹²¹

The analogous aminoethylation to provide the acetyllysine mimic fails (Figure 3.12B).¹²² To access mimics of acetyllysine by alkylation of cysteine, Cole and co-workers have developed a reagent to alkylate cysteine that results in a thiocarbamate analog of acetyllysine (Figure 3.12B).¹²² This acetyllysine analog is resistant to hydrolysis by histone deacetylase enzymes, and so may be useful in cell-based assays

where the resilience of the modification may be crucial. The analog was incorporated in histones and shown to bind to antibodies and binding partners that normally recognize the natural acetyllysine residue.¹²²

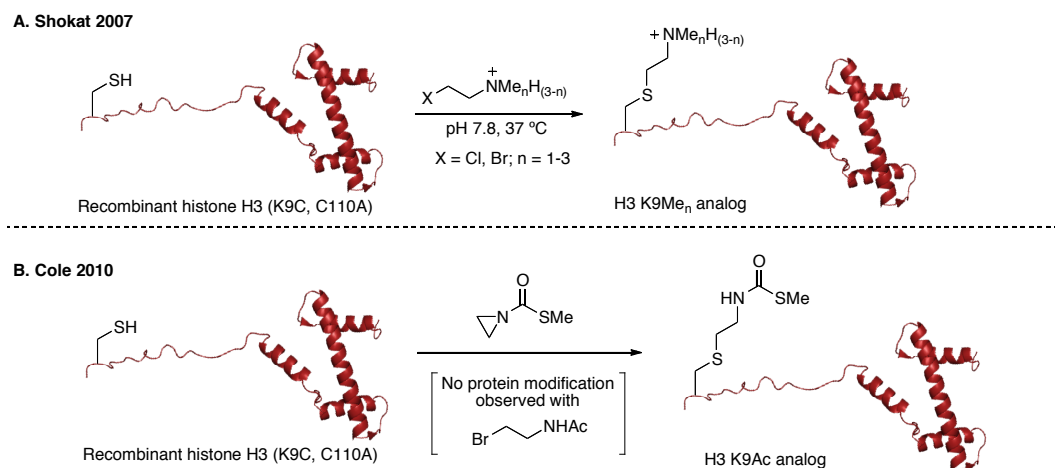


Figure 3.12: Chemical access to methylated and acetylated lysine derivatives in histone proteins

Dehydroalanine as a precursor to post-translational modifications

From the survey above, it is clear that there are several strategies for accessing post-translational modifications. Total chemical synthesis, recombinant technologies, and various enzymatic and chemical modifications can all provide proteins bearing PTMs. These techniques, of course, are not without limitations. For instance, native chemical ligation and expressed protein ligation are unrivaled in accessing precisely modified proteins—especially those that contain several modifications. However, the technical demands associated with multiple ligations typically limit the modifications to the 50 residues of either terminus. In contrast, the incorporation of modified amino acids by translation, for instance by amber codon suppression, is not restricted to any particular site. Still, each modification using nonsense suppression necessitates the evolution of a unique synthetase for its incorporation and many important modifications (e.g. lysine trimethylation, serine glycosylation, and serine phosphorylation, among others) are not yet amenable to this technique. Chemical techniques are also limited in several respects. These methods are typically restricted to a single modification and are predicated on incorporation of a uniquely reactive side chain at the site of interest. Even then, many important modifications (serine phosphorylation, for instance) have not been reported by chemical methods.

In Chapters 3 and 4, the use of dehydroalanine as a precursor to post-translational modifications will be explored in an effort to resolve some of the limitations described above. A motivating factor for this study was the potential for a single residue—dehydroalanine—to serve as a precursor for several classes of PTMs. In Figure 3.13, several PTMs or analogs to natural PTMs are shown that are, in principle, accessible from Dha. Preliminary validation of this strategy was already reported in Chapter 2 where dehydroalanine was chemically converted to a cysteine residue bearing an *N*-acetylglucosamine modification. The proposed synthesis of these PTMs using reagents **2.1** and **2.32** is the focus of Chapters 3 and 4.

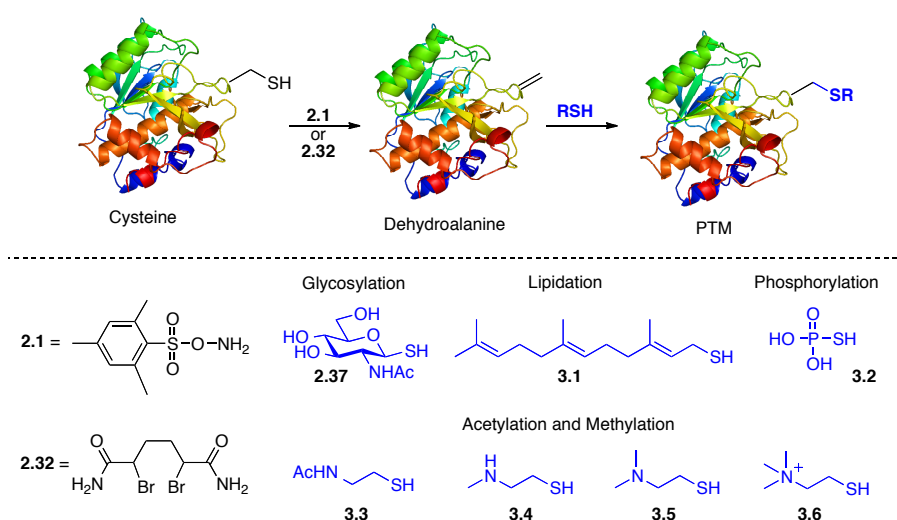


Figure 3.13: Post-translational modifications (and analogs) accessible, in principle, through dehydroalanine

3.2 Results and Discussion

Thiol addition to dehydroalanine provides post-translationally modified proteins.

The first experimental goal was to demonstrate that dehydroalanine, synthesized using MSH (**2.1**) or dibromide **2.32**, is an acceptor for thiols that provide the natural post-translational modifications or mimics shown in Figure 3.13. These transformations were first investigated on subtilisin from *Bacillus lentus* (SBL-S156C) and Np276 from *Nostoc punctiforme* since dehydroalanine was readily installed on these proteins at a single site using **2.1** or **2.32**.

Dehydroalanine is a precursor to *S*-linked glycoproteins.

Glycosyl thiol **2.37** was shown in Chapter 2 to add to dehydroalanine to provide GlcNAc cysteine. As discussed previously, GlcNAc cysteine is a natural residue^{60,61} and an analog of the more common GlcNAc serine residue.^{10,123} The addition to dehydroalanine is operationally simple and can be carried out in a one-pot procedure whereby cysteine is first converted to dehydroalanine using MSH and then reacted with a thiol in the same reaction vessel. This transformation is shown in Figure 3.14 on SBL-S156C.

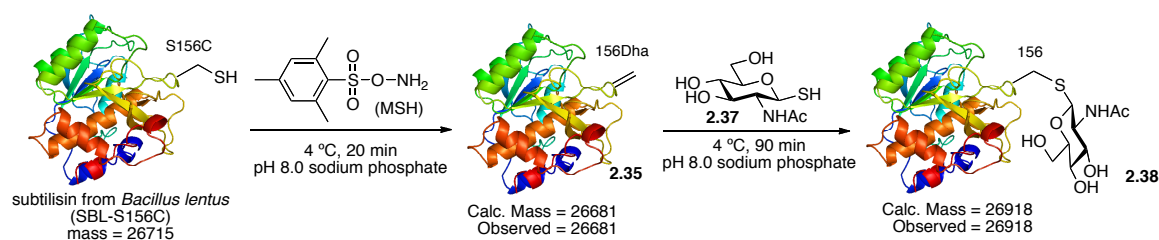


Figure 3.14 Synthesis of glycoproteins using dehydroalanine

This simple procedure for accessing GlcNAc cysteine should find use in the emerging areas of chemical biology that involve proteins bearing GlcNAc (*vide infra*).^{10,123-125} Moreover, outside of total or semi-synthesis of glycoproteins by native chemical ligation, no simple conjugation reaction for accessing GlcNAc cysteine has been reported for protein substrates. In contrast, glycosyl thiols are easily prepared by a number of methods^{126,127} and their addition to dehydroalanine is operationally simple and highly chemoselective.⁶³⁻⁶⁷ Finally, the resulting protein products can be purified by simple dialysis or size exclusion chromatography to remove small molecule reagents. After this purification, protein recovery is typically 80% or higher.

An inherent limitation of using dehydroalanine is the epimeric mixture that results from the conjugate addition of a thiol to Dha. Diastereoselectivity is typically low for such additions⁶³⁻⁶⁵ since the stereochemically determining step is protonation of an intermediate enolate, likely a fast and unselective step in water. On the protein substrates in this and subsequent Chapters, no effort was made to determine the diastereoselectivity. It is assumed that a mixture of epimers results, analogous to peptide models.⁶³⁻⁶⁵ However, it is essential to see if these post-translationally modified proteins are still useful in a biochemical or structural assay. If so, the ease of synthesis might mitigate the stereochemical disadvantage of using dehydroalanine as a precursor to post-translationally modified proteins. To this end, glycoprotein **2.38**, synthesized using

dehydroalanine and the conjugate addition of glycosyl thiol **2.37**, was used as a substrate in an enzymatic assay with endoglycosidase A (Endo A).[†] It was discovered that Endo A could catalyze the transfer of oxazoline carbohydrate donors to the 4-OH of the GlcNAc acceptor on glycoprotein **2.38**. Full conversion was observed for model reactions, suggesting both epimers are acceptors in this reaction (Scheme 3.15).¹²⁸

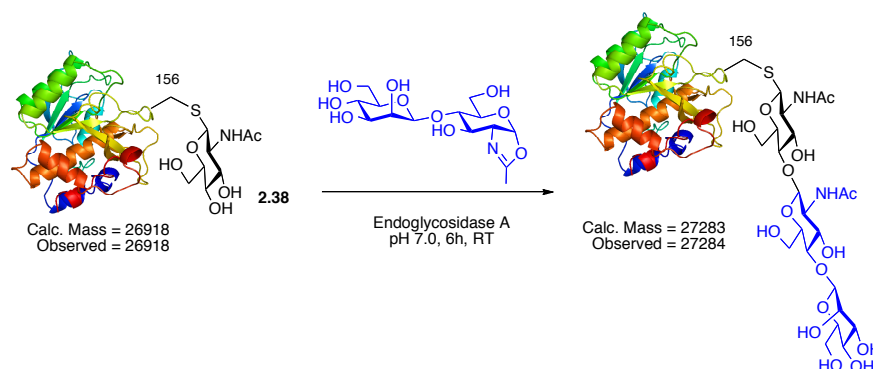


Figure 3.15: Glycoprotein **2.38**, synthesized *via* dehydroalanine, is an acceptor for Endoglycosidase A catalyzed oxazoline transfer. Full conversion was observed in this transformation.^{128,†}

Glycosyl thiols also added to dehydroalanine on the model protein Np276 (Figure 3.16). Dehydroalanine was generated at a single site on this protein using reagent **2.32**, as reported in Chapter 2. Glycosyl thiol **2.37** added to dehydroalanine to provide glycoprotein **3.7**. The protein recovery after dialysis was 88%.

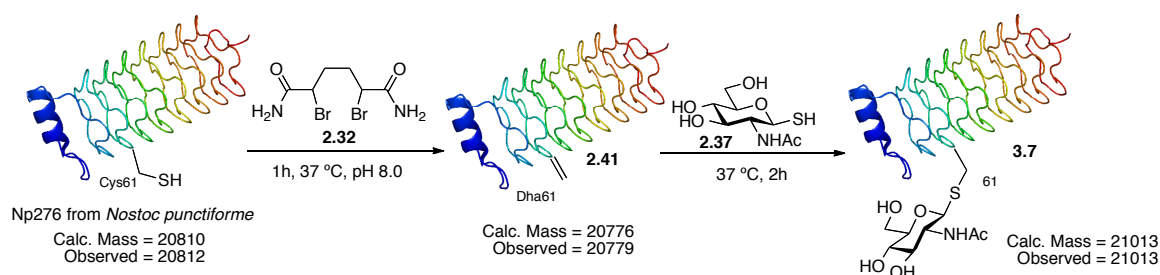


Figure 3.16: Installation of GlcNAc cysteine on model protein Np276

[†] This experiment with endoglycosidase A was carried out by Dr. Marta Fernández-González as part of a larger study on the promiscuity of endoglycosidases and the chemoenzymatic synthesis of glycoproteins. The result is included here only for discussion of the product **2.38** and its use in a biochemical assay that might be affected by the epimeric mixture that results from the conjugate addition to dehydroalanine. The reader is referred to reference 128 for full details of the endoglycosidase study.

The conversion of cysteine to dehydroalanine was accomplished on two different proteins, both of which could be further transformed to *S*-linked glycoproteins by the addition of a glycosyl thiol. While the addition of glycosyl thiols to dehydroalanine has been reported previously on peptides,^{63-66,129} there is only a single example of using dehydroalanine as a precursor to glycoproteins.⁶⁷ In this report by Schultz, the addition of the glycosyl thiol of mannose was added to dehydroalanine that was generated by the oxidative elimination of a phenylselenocysteine residue in green fluorescent protein. All methionines were oxidized in this procedure. In contrast, the transformations in Figure 3.15 and Figure 3.16 are highly chemoselective and demonstrated the first conversion of dehydroalanine to GlcNAc cysteine on a protein substrate. Next, other thiols were examined that would provide access to other classes of post-translational modifications after reaction with dehydroalanine.

Dehydroalanine is a precursor to polyprenylated proteins.

The addition of prenylated thiols to dehydroalanine yields prenylated cysteine,^{63,64} a natural post-translational modification.¹⁴ To validate this route to prenylated proteins, SBL-156Dha (**2.35**) was reacted with farnesyl thiol (**3.1**). Since farnesyl thiol is not soluble in water, it was first dissolved in DMSO. The addition of this solution to the protein solution resulted in an emulsion. The total DMSO content did not exceed 10% of the total reaction volume. The water soluble reducing agent tris-(2-carboxyethyl)phosphine (TCEP) was included in the reaction mixture to reduce any farnesyl disulfide. Without TCEP, farnesyl thiol formed disulfide rapidly under these conditions, lowering the amount of free thiol and slowing the conjugate addition. Gratifyingly, the addition of farnesyl thiol to dehydroalanine was complete at room temperature after 90 minutes of reaction time (Figure 3.17).

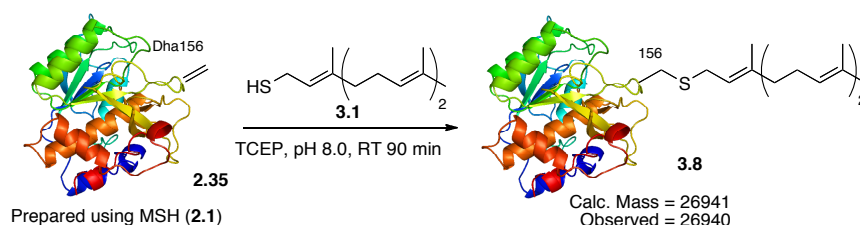


Figure 3.17: Addition of farnesyl thiol to dehydroalanine

Farnesylated proteins have been prepared by various ligation methods such as NCL and EPL, as discussed in the introduction. The reaction sequence in Figure 3.17 complements these efforts in accessing the natural farnesylated cysteine residue. While the epimeric problem cannot be avoided using the dehydroalanine method, this protocol remains one of the most technically simple methods for accessing prenylated proteins. Moreover, the addition to dehydroalanine results in a thioether that is not susceptible to reduction like disulfide-linked lipidated proteins.⁷⁵

While glycosylation and lipidation were important demonstrations of the utility of dehydroalanine (and the methods to generate dehydroalanine), other post-translational modifications might be possible that had not been previously synthesized through dehydroalanine. We turn to these modifications next.

Dehydroalanine is a precursor to phosphocysteine.

Phosphorylation is one of the most prevalent post-translational modifications.⁶ Should thiophosphate add to dehydroalanine, a simple method would be established for the synthesis of phosphocysteine. Phosphocysteine is a natural residue, but might serve as a mimic for the ubiquitous phosphoserine.¹³⁰ It was not clear at the outset whether phosphocysteine would be stable under basic conditions and hydrolysis to cysteine or elimination back to dehydroalanine might complicate the proposed transformation. However, these concerns were disqualified when SBL-156Dha (**2.35**) was reacted with thiophosphate: rapid and complete addition was observed within an hour at 4 °C (Figure 3.18).

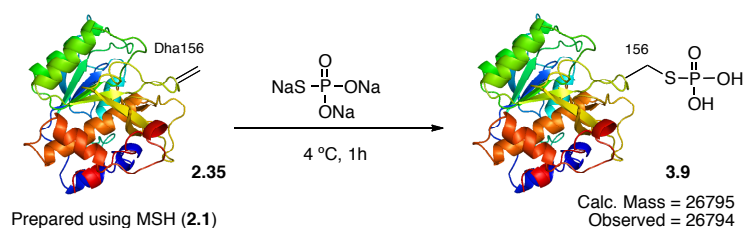


Figure 3.18: Synthesis of a phosphorylated protein by reaction of thiophosphate with dehydroalanine

The addition of thiophosphate to dehydroalanine is a simple and direct route to phosphorylated proteins. As discussed in the introduction to this Chapter, total chemical synthesis by native chemical ligation can supply phosphorylated proteins. Chemical methods to selectively access phosphoserine or phosphocysteine, however, have not been

reported. It remained to be determined if phosphocysteine is a functional mimic of phosphoserine—either in enzymatic or various binding assays. This validation will be discussed in Chapter 4. First, the synthesis of other post-translational modifications from dehydroalanine was explored.

Dehydroalanine is a precursor to methylated lysine mimics.

The methylation of lysine is a dynamic post-translational modification. Perhaps the most intense interest in these PTMs is in the context of histone modifications and epigenetics.⁸⁴ Histone proteins and DNA are the fundamental constituents of chromatin and post-translational modifications of histones are intimately involved in the regulation of DNA transcription,¹³¹ replication, and repair.¹³² Access to modified histones will help elucidate the structural basis for these processes and aid in deconvoluting the “histone code”.^{83-86,133} The use of dehydroalanine as a precursor to modified lysine residues was inspired by Shokat’s report of cysteine alkylation to access mimics of methylated lysine.¹¹⁵ Using dehydroalanine as a precursor to these methylated lysine analogs would complement direct alkylation of cysteine. For instance, if a cysteine and dehydroalanine residue were incorporated into a protein at two different sites, alkylation of cysteine and addition of a thiol to dehydroalanine would provide a route to two distinct modifications on the same protein. For histone proteins, multiple modifications are especially common.

The first task was again to establish conditions for the synthesis of the methylated lysine analogs. SBL-156Dha was used for this initial testing. Since the thiols necessary for this addition were synthesized as their hydrochloride salts, sodium or potassium hydrogenphosphate was used to first neutralize the reagents and give a final pH of the reaction mixture between 8 and 9. Gratifyingly, the addition of thiols **3.4-3.6** proceeded smoothly under these conditions to provide the methylated lysine analogs (Figure 3.19). The establishment of these reaction conditions was important for the application of this chemistry to histone proteins, the focus of Chapter 4.

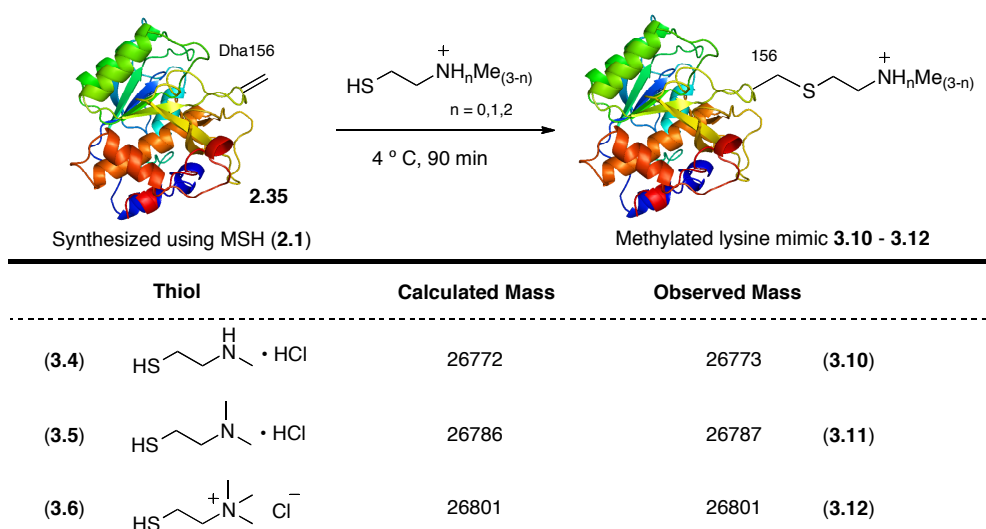


Figure 3.19: Synthesis of methylated lysine analogs from dehydroalanine

Alkyl cysteine modifications are stable, but not necessarily permanent.

One feature of using dehydroalanine (or cysteine itself) as a precursor to alkyl cysteines is that the product thioether is a stable modification compared to other cysteine modification techniques such as disulfide formation. This stability was demonstrated experimentally by incubating the thioether-linked glycoprotein **2.38** with 10 mM glutathione, the natural cellular redox buffer. After 3 hours at 37 °C, the protein was recovered unchanged. When the same experiment was run with the corresponding disulfide-linked glycoprotein (**3.13**) complete reduction was observed (Figure 3.20).

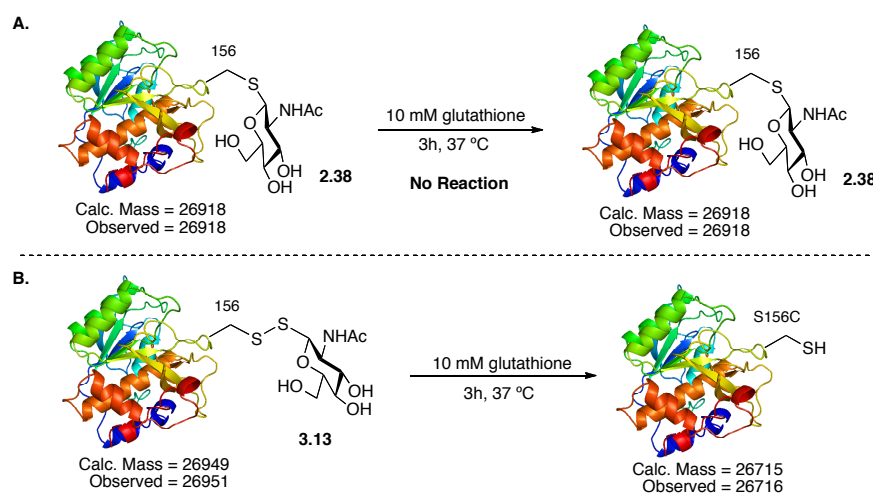


Figure 3.20: Relative stability of thioethers and disulfides in the presence of glutathione

The stability of thioethers, however, does not necessarily imply permanence. In Chapter 2, it was shown that alkyl cysteines can be eliminated to dehydroalanine by

reaction with MSH. The newly regenerated dehydroalanine can then be converted to another modification. This reaction sequence was demonstrated on SBL-S156C (Figure 3.21). Ethanethiol was added to dehydroalanine to give *S*-ethyl cysteine (**3.14**). Oxidative elimination of this thioether by reaction with MSH regenerated dehydroalanine. A final addition of GlcNAc thiol **2.37** was used to demonstrate the chemical transformation to a different modification. The ability to install and remove PTMs by selective chemical modification might be useful in the study of these proteins in binding or catalysis. The oxidative elimination of thioethers might also find use in mapping polyprenylation: after eliminating polyprenylated cysteine residues to dehydroalanine, various affinity tags could be added to map this PTM.

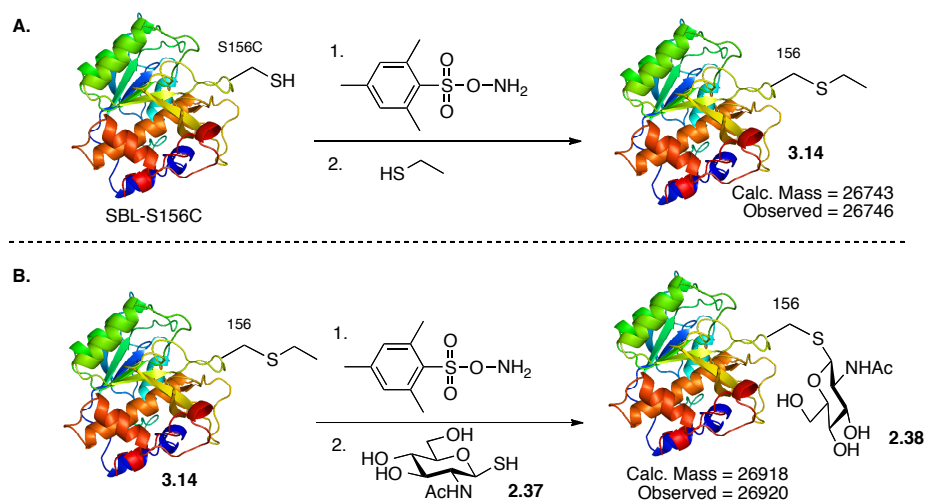


Figure 3.21: Oxidative elimination of *S*-ethyl cysteine to dehydroalanine and addition of GlcNAc thiol **2.37**.

The results above establish the use of dehydroalanine as a precursor to a variety of post-translational modifications. Glycosylation, lipidation, phosphorylation and methylation were all demonstrated on a model protein substrate. The next goal of this project was to apply this chemistry to a protein that is naturally modified with one or more of the post-translational modifications described above. By accessing these proteins, the synthetic modifications could be validated in some biological context. Furthermore, application of this chemistry to another protein substrate would allow a further assessment of scope for the methods of dehydroalanine synthesis as well as its subsequent conversion to a PTM or PTM mimic. Histone proteins are a class of proteins that bear many of the PTMs accessible through dehydroalanine.²¹ Moreover, the structural and

functional consequences of many of these modifications are not entirely elucidated and so chemical access to these modified proteins is of current interest to chemical biologists.⁸³⁻

⁸⁶ The incorporation of dehydroalanine into histones and its conversion to PTMs is the focus of Chapter 4.

3.3 Conclusions and Outlook

In this Chapter, dehydroalanine was used as a precursor to several classes of post-translational modifications. This study was enabled by efficient and reliable conversion of cysteine to dehydroalanine described in the previous Chapter. Proteins containing glycans, lipids, phosphates, and methylated lysines were all synthesized. For many of these modified proteins, there is no straightforward alternative for their synthesis outside native chemical ligation or expressed protein ligation. In particular, the synthesis of proteins containing glycosyl cysteines and phosphocysteines was proposed here as a method to mimic the natural glycosyl serine and phosphoserine residues, respectively. No other chemical methods have been reported to access these residues.

Methylated lysine analogs are also accessible from dehydroalanine. While these residues can easily be accessed by simple alkylation of cysteine,¹¹⁵ it is conceivable that dehydroalanine and cysteine could be installed at two different sites and allow the chemical installation of multiple modifications. Soon after the preliminary results of this Chapter were published,¹³⁴ the Schultz labs reported a biosynthetic route to dehydroalanine and its subsequent conversion to the methylated lysine analogs by the methods depicted in Figure 3.19.¹³⁵ Similarly, van der Donk and co-workers have since reported an enzymatic route to dehydroalanine and its conversion to methylated lysine mimics by the addition of an appropriate thiol.¹²⁹ With multiple methods to access dehydroalanine, the chemical installation of multiple PTMs is feasible.

Acetylated lysine analogs are also accessible from dehydroalanine, as will be demonstrated in Chapter 4 in the synthesis of modified histones. Histone proteins carry many of the PTMs accessible from dehydroalanine and are therefore an excellent testing ground for the chemistry developed in Chapters 2 and 3. The synthesis of modified histones using dehydroalanine also allows the opportunity to validate the PTM mimics as functional surrogates for the natural modifications, a topic explored in the following Chapter.

3.4 References

- (1) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372.
- (2) Walsh, C. T. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*. Roberts and Company Publishers: Englewood, CO, 2006.
- (3) Hunter, T.; Karin, M. The regulation of transcription by phosphorylation. *Cell* **1992**, *70*, 375-387.
- (4) Whitmarsh, A. J.; Davis, R. J. Regulation of transcription factor function by phosphorylation. *Cell. Mol. Life Sci.* **2000**, *57*, 1172-1183.
- (5) Cohen, P. The origins of protein phosphorylation. *Nat. Cell Biol.* **2002**, *4*, E127-E130.
- (6) Tarrant, M. K.; Cole, P. A. The chemical biology of protein phosphorylation. *Annu. Rev. Biochem.* **2009**, *78*, 797-825.
- (7) Manning, G.; Whyte, D. B.; Martinez, R.; Hunter, T.; Sudarsanam, S. The protein kinase complement of the human genome. *Science* **2002**, *298*, 1912-1934.
- (8) Dwek, R. A. Glycobiology: toward understanding the function of sugars. *Chem. Rev.* **1996**, *96*, 683-720.
- (9) Ohtsubo, K.; Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **2006**, *126*, 855-867.
- (10) *Essentials of Glycobiology*; 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press, 2009.
- (11) Jacob, C.; Giles, G. I.; Giles, N. M.; Sies, H. Sulfur and selenium: the role of oxidation state in protein structure and function. *Angew. Chem. Int. Ed.* **2003**, *42*, 4742-4758.
- (12) Shoulders, M. D.; Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **2009**, *78*, 929-958.
- (13) Schofield, C. J.; Ratcliffe, P. J. Oxygen sensing by HIF hydroxylases. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 343-354.
- (14) Zhang, F. L.; Casey, P. J. Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **1996**, *65*, 241-269.

- (15) Gelb, M. H.; Brunsfeld, L.; Hrycyna, C. A.; Michaelis, S.; Tamanoi, F.; Van Voorhis, W. C.; Waldmann, H. Therapeutic intervention based on protein prenylation and associated modifications. *Nat. Chem. Biol.* **2006**, *2*, 518-528.
- (16) Kim, S. C.; Sprung, R.; Chen, Y.; Xu, Y.; Ball, H.; Pei, J.; Cheng, T.; Kho, Y.; Xiao, H.; Xiao, L.; Grishin, N. V.; White, M.; Yang, X.-J.; Zhao, Y. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol. Cell* **2006**, *23*, 607-618.
- (17) Zhao, S.; Xu, W.; Jiang, W.; Yu, W.; Lin, Y.; Zhang, T.; Yao, J.; Zhou, L.; Zeng, Y.; Li, H.; Li, Y.; Shi, J.; An, W.; Hancock, S. M.; He, F.; Qin, L.; Chin, J.; Yang, P.; Chen, X.; Lei, Q.; Xiong, Y.; Guan, K.-L. Regulation of cellular metabolism by protein lysine acetylation. *Science* **2010**, *327*, 1000-1004.
- (18) Kimura, A.; Matsubara, K.; Horikoshi, M. A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. *J. Biochem.* **2005**, *138*, 647-662.
- (19) Martin, C.; Zhang, Y. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 838-849.
- (20) Bedford, M. T.; Richard, S. Arginine methylation: an emerging regulator of protein function. *Mol. Cell* **2005**, *18*, 263-272.
- (21) Kouzarides, T. Chromatin modifications and their function. *Cell* **2007**, *128*, 693-705.
- (22) Dawson, P. E.; Kent, S. B. H. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923-960.
- (23) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338-351.
- (24) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705-6710.
- (25) Flavell, R. R.; Muir, T. W. Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology. *Acc. Chem. Res.* **2009**, *42*, 107-116.
- (26) Schwarzer, D.; Cole, P. A. Protein semisynthesis and expressed protein ligation: chasing a protein's tail. *Curr. Opin. Chem. Biol.* **2005**, *9*, 561-569.
- (27) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91-118.

- (28) Kumar, K. S. A.; Brik, A. Accessing posttranslationally modified proteins through thiol positioning. *J. Pept. Sci.* **2010**, *16*, 524-529.
- (29) Shin, Y.; Winans, K. A.; Backes, B. J.; Kent, S. B. H.; Ellman, J. A.; Bertozzi, C. R. Fmoc-based synthesis of peptide-thioesters: application to the total chemical synthesis of a glycoprotein by native chemical ligation. *J. Am. Chem. Soc.* **1999**, *121*, 11684-11689.
- (30) Macmillan, D.; Bertozzi, C. R. Modular assembly of glycoproteins: towards the synthesis of GlyCAM-1 by using expressed protein ligation. *Angew. Chem. Int. Ed.* **2004**, *43*, 1355-1359.
- (31) Hojo, H.; Matsumoto, Y.; Nakahara, Y.; Ito, E.; Suzuki, Y.; Suzuki, M.; Suzuki, A.; Nakahara, Y. Chemical synthesis of 23 kDa glycoprotein by repetitive segment condensation: a synthesis of MUC2 basal motif carrying multiple O-GlcNAc moieties. *J. Am. Chem. Soc.* **2005**, *127*, 13720-13725.
- (32) Yamamoto, N.; Tanabe, Y.; Okamoto, R.; Dawson, P. E.; Kajihara, Y. Chemical synthesis of a glycoprotein having an intact human complex-type sialyloligosaccharide under the Boc and Fmoc synthetic strategies. *J. Am. Chem. Soc.* **2008**, *130*, 501-510.
- (33) Kajihara, Y.; Yamamoto, N.; Okamoto, R.; Hirano, K.; Murase, T. Chemical synthesis of homogeneous glycopeptides and glycoproteins. *Chem. Rec.* **2010**, *10*, 80-100.
- (34) Yuan, Y.; Chen, J.; Wan, Q.; Wilson, R. M.; Danishefsky, S. J. Toward fully synthetic, homogeneous glycoproteins: advances in chemical ligation. *Peptide Science* **2010**, *94*, 373-384.
- (35) Piontek, C.; Ring, P.; Harjes, O.; Heinlein, C.; Mezzato, S.; Lombana, N.; Pöhner, C.; Püttner, M.; Varón Silva, D.; Martin, A.; Xaver Schmid, F.; Unverzagt, C. Semisynthesis of a homogeneous glycoprotein enzyme: ribonuclease C: part 1. *Angew. Chem. Int. Ed.* **2009**, *48*, 1936-1940.
- (36) Piontek, C.; Varón Silva, D.; Heinlein, C.; Pöhner, C.; Mezzato, S.; Ring, P.; Martin, A.; Xaver Schmid, F.; Unverzagt, C. Semisynthesis of a homogeneous glycoprotein enzyme: ribonuclease C: part 2. *Angew. Chem. Int. Ed.* **2009**, *48*, 1941-1945.
- (37) Brik, A.; Yang, Y.-Y.; Ficht, S.; Wong, C.-H. Sugar-assisted glycopeptide ligation. *J. Am. Chem. Soc.* **2006**, *128*, 5626-5627.

- (38) Hackenberger, C. P. R.; Friel, C. T.; Radford, S. E.; Imperiali, B. Semisynthesis of a glycosylated Im7 analogue for protein folding studies. *J. Am. Chem. Soc.* **2005**, *127*, 12882-12889.
- (39) Zhu, J.; Warren, J. D.; Danishefsky, S. J. Synthetic carbohydrate-based anticancer vaccines: the Memorial Sloan-Kettering experience. *Expert Rev. Vaccines* **2009**, *8*, 1399-1413.
- (40) Hamilton, S. R.; Davidson, R. C.; Sethuraman, N.; Nett, J. H.; Jiang, Y.; Rios, S.; Bobrowicz, P.; Stadheim, T. A.; Li, H.; Choi, B.-K.; Hopkins, D.; Wischniewski, H.; Roser, J.; Mitchell, T.; Strawbridge, R.; Hoopes, J.; Wildt, S.; Gerngross, T. U. Humanization of yeast to produce complex terminally sialylated glycoproteins. *Science* **2006**, *313*, 1441-1443.
- (41) Li, H.; Sethuraman, N.; Stadheim, T. A.; Zha, D.; Prinz, B.; Ballew, N.; Bobrowicz, P.; Choi, B.-K.; Cook, W. J.; Cukan, M.; Houston-Cummings, N. R.; Davidson, R.; Gong, B.; Hamilton, S. R.; Hoopes, J. P.; Jiang, Y.; Kim, N.; Mansfield, R.; Nett, J. H.; Rios, S.; Strawbridge, R.; Wildt, S.; Gerngross, T. U. Optimization of humanized IgGs in glycoengineered *Pichia pastoris*. *Nat. Biotechnol.* **2006**, *24*, 210-215.
- (42) Hamilton, S. R.; Gerngross, T. U. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr. Opin. Biotechnol.* **2007**, *18*, 387-392.
- (43) Schwarz, F.; Huang, W.; Li, C.; Schulz, B. L.; Lizak, C.; Palumbo, A.; Numao, S.; Neri, D.; Aebi, M.; Wang, L.-X. A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation. *Nat. Chem. Biol.* **2010**, *6*, 264-266.
- (44) Takegawa, K.; Tabuchi, M.; Yamaguchi, S.; Kondo, A.; Kato, I.; Iwahara, S. Synthesis of neoglycoproteins using oligosaccharide-transfer activity with endo- β -N-acetylglucosaminidase. *J. Biol. Chem.* **1995**, *270*, 3094-3099.
- (45) Witte, K.; Sears, P.; Martin, R.; Wong, C.-H. Enzymatic glycoprotein synthesis: preparation of ribonuclease glycoforms via enzymatic glycopeptide condensation and glycosylation. *J. Am. Chem. Soc.* **1997**, *119*, 2114-2118.
- (46) Li, B.; Song, H.; Hauser, S.; Wang, L.-X. A highly efficient chemoenzymatic approach toward glycoprotein synthesis. *Org. Lett.* **2006**, *8*, 3081-3084.
- (47) Wang, L.-X. Chemoenzymatic synthesis of glycopeptides and glycoproteins through endoglycosidase-catalyzed transglycosylation. *Carb. Res.* **2008**, *343*, 1509-1522.

- (48) Wei, Y.; Li, C.; Huang, W.; Li, B.; Strome, S.; Wang, L.-X. Glycoengineering of human IgG1-Fc through combined yeast expression and *in vitro* chemoenzymatic glycosylation. *Biochemistry* **2008**, *47*, 10294-10304.
- (49) Fujita, K.; Yamamoto, K. A remodeling system for the oligosaccharide chains on glycoproteins with microbial endo- β -*N*-acetylglucosaminidases. *Biochim. Biophys. Acta* **2006**, *1760*, 1631-1635.
- (50) Rising, T. W. D. F.; Heidecke, C. D.; Moir, J. W. B.; Ling, Z.; Fairbanks, A. J. Endohexosaminidase-catalysed glycosylation with oxazoline donors: fine tuning of catalytic efficiency and reversibility. *Chem. Eur. J.* **2008**, *14*, 6444-6464.
- (51) Heidecke, C. D.; Ling, Z.; Bruce, N. C.; Moir, J. W. B.; Parsons, T. B.; Fairbanks, A. J. Enhanced glycosylation with mutants of endohexosaminidase A (Endo A). *ChemBioChem* **2008**, *9*, 2045-2051.
- (52) Huang, W.; Li, C.; Li, B.; Umekawa, M.; Yamamoto, K.; Zhang, X.; Wang, L.-X. Glycosynthases enable a highly efficient chemoenzymatic synthesis of *N*-glycoproteins carrying intact natural *N*-glycans. *J. Am. Chem. Soc.* **2009**, *131*, 2214-2223.
- (53) Davis, B. G. Synthesis of glycoproteins. *Chem. Rev.* **2002**, *102*, 579-601.
- (54) Gamblin, D. P.; Scanlan, E. M.; Davis, B. G. Glycoprotein synthesis: an update. *Chem. Rev.* **2009**, *109*, 131-163.
- (55) Davis, B. G.; Lloyd, R. C.; Jones, J. B. Controlled site-selective glycosylation of proteins by a combined site-directed mutagenesis and chemical modification approach. *J. Org. Chem.* **1998**, *63*, 9614-9615.
- (56) Davis, B. G.; Maughan, M. A. T.; Green, M. P.; Ullman, A.; Jones, J. B. Glycomethanethiosulfonates: powerful reagents for protein glycosylation. *Tetrahedron: Asymmetry* **2000**, *11*, 245-262.
- (57) Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glycosyl phenylthiosulfonates (glyco-PTS): novel reagents for glycoprotein synthesis. *Org. Biomol. Chem.* **2003**, *1*, 3642-3644.
- (58) Gamblin, D. P.; Garnier, P.; van Kasteren, S.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glyco-SeS: selenenylsulfide-mediated protein glycoconjugation-a new strategy in post-translational modification. *Angew. Chem. Int. Ed.* **2004**, *43*, 828-833.

- (59) Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L. Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin. *Chem. Biol.* **2001**, *8*, 133-145.
- (60) Lote, C. J.; Weiss, J. B. Identification of digalactosylcysteine in a glycopeptide isolated from urine by a new preparative technique. *FEBS Lett.* **1971**, *16*, 81-85.
- (61) Weiss, J. B.; Lote, C. J.; Bobinski, H. New low molecular weight glycopeptide containing triglucosylcysteine in human erythrocyte membrane. *Nat. New Biol.* **1971**, *234*, 25-26.
- (62) Oman, T. J.; Boettcher, J. M.; Wang, H.; Okalibe, X. N.; van der Donk, W. A. Sublancin is not a lantibiotic but an S-linked glycopeptide. *Nat. Chem. Biol.* **2011**, *7*, 78-80.
- (63) Zhu, Y.; van der Donk, W. A. Convergent synthesis of peptide conjugates using dehydroalanines for chemoselective ligations. *Org. Lett.* **2001**, *3*, 1189-1192.
- (64) Gieselman, M. D.; Zhu, Y.; Zhou, H.; Galonic, D.; van der Donk, W. A. Selenocysteine derivatives for chemoselective ligations. *ChemBioChem* **2002**, *3*, 709-716.
- (65) Galonic, D. P.; van der Donk, W. A.; Gin, D. Y. Oligosaccharide-peptide ligation of glycosyl thiolates with dehydropeptides: synthesis of S-linked mucin-related glycopeptide conjugates. *Chem. Eur. J.* **2003**, *9*, 5997-6006.
- (66) Levengood, M. R.; van der Donk, W. A. Dehydroalanine-containing peptides: preparation from phenylselenocysteine and utility in convergent ligation strategies. *Nat. Protoc.* **2007**, *1*, 3001-3010.
- (67) Wang, J.; Schiller, S. M.; Schultz, P. G. A biosynthetic route to dehydroalanine-containing proteins. *Angew. Chem. Int. Ed.* **2007**, *46*, 6849-6851.
- (68) Jenkins, C. M.; Han, X.; Yang, J.; Mancuso, D. J.; Sims, H. F.; Muslin, A. J.; Gross, R. W. Purification of recombinant human cPLA₂ γ and identification of C-terminal farnesylation, proteolytic processing, and carboxymethylation by MALDI-TOF-TOF analysis. *Biochemistry* **2003**, *42*, 11798-11807.
- (69) Kalinin, A.; Thomä, N. H.; Iakovenko, A.; Heinemann, I.; Rostkova, E.; Constantinescu, A. T.; Alexandrov, K. Expression of mammalian geranylgeranyltransferase type-II in *Escherichia coli* and its application for *in vitro* prenylation of Rab proteins. *Protein Expr. Purif.* **2001**, *22*, 84-91.

- (70) Rak, A.; Pylypenko, O.; Durek, T.; Watzke, A.; Kushnir, S.; Brunsveld, L.; Waldmann, H.; Goody, R. S.; Alexandrov, K. Structure of Rab GDP-dissociation inhibitor in complex with prenylated YPT1 GTPase. *Science* **2003**, *302*, 646-650.
- (71) Durek, T.; Alexandrov, K.; Goody, R. S.; Hildebrand, A.; Heinemann, I.; Waldmann, H. Synthesis of fluorescently labeled mono- and diprenylated Rab7 GTPase. *J. Am. Chem. Soc.* **2004**, *126*, 16368-16378.
- (72) Brunsveld, L.; Watzke, A.; Durek, T.; Alexandrov, K.; Goody, R. S.; Waldmann, H. Synthesis of functionalized Rab GTPases by a combination of solution- or solid-phase lipopeptide synthesis with expressed protein ligation. *Chem. Eur. J.* **2005**, *11*, 2756-2772.
- (73) Wu, Y.-W.; Tan, K.-T.; Waldmann, H.; Goody, R. S.; Alexandrov, K. Interaction analysis of prenylated Rab GTPase with Rab escort protein and GDP dissociation inhibitor explains the need for both regulators. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12294-12299.
- (74) Bader, B.; Kuhn, K.; Owen, D. J.; Waldmann, H.; Wittinghofer, A.; Kuhlmann, J. Bioorganic synthesis of lipid-modified proteins for the study of signal transduction. *Nature* **2000**, *403*, 223-226.
- (75) Gamblin, D. P.; van Kasteren, S.; Bernardes, G. J. L.; Chalker, J. M.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Chemical site-selective prenylation of proteins. *Mol. BioSyst.* **2008**, *4*, 558-561.
- (76) Tilley, S. D.; Francis, M. B. Tyrosine-selective protein alkylation using π -allylpalladium complexes. *J. Am. Chem. Soc.* **2006**, *128*, 1080-1081.
- (77) Huse, M.; Holford, M. N.; Kuriyan, J.; Muir, T. W. Semisynthesis of hyperphosphorylated type I TGF β receptor: addressing the mechanism of kinase activation. *J. Am. Chem. Soc.* **2000**, *122*, 8337-8338.
- (78) Huse, M.; Muir, T. W.; Xu, L.; Chen, Y.-G.; Kuriyan, J.; Massagué, J. The TGF β receptor activation process: an inhibitor- to substrate-binding switch. *Mol. Cell* **2001**, *8*, 671-682.
- (79) Ottesen, J. J.; Huse, M.; Sekedat, M. D.; Muir, T. W. Semisynthesis of phosphovariants of Smad2 reveals a substrate preference of the activated T β RI kinase. *Biochemistry* **2004**, *43*, 5698-5706.
- (80) Flavell, R. R.; Huse, M.; Goger, M.; Trester-Zedlitz, M.; Kuriyan, J.; Muir, T. W. Efficient semisynthesis of a tetraphosphorylated analogue of the type I TGF β receptor. *Org. Lett.* **2002**, *4*, 165-168.

- (81) Shogren-Knaak, M. A.; Fry, C. J.; Peterson, C. L. A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* **2003**, *278*, 15744-15748.
- (82) Chiang, K. P.; Jensen, M. S.; McGinty, R. K.; Muir, T. W. A semisynthetic strategy to generate phosphorylated and acetylated histone H2B. *ChemBioChem* **2009**, *10*, 2182-2187.
- (83) Chatterjee, C.; Muir, T. W. Chemical approaches for studying histone modifications. *J. Biol. Chem.* **2010**, *285*, 11045-11050.
- (84) Allis, C. D.; Muir, T. W. Spreading chromatin into chemical biology. *ChemBioChem* **2011**, *12*, 264-279.
- (85) Oliver, S. S.; Denu, J. M. Dynamic interplay between histone H3 modifications and protein interpreters: emerging evidence for a "histone language". *ChemBioChem* **2011**, *12*, 299-307.
- (86) Voigt, P.; Reinberg, D. Histone tails: ideal motifs for probing epigenetics through chemical biology approaches. *ChemBioChem* **2011**, *12*, 236-252.
- (87) Lu, W.; Gong, D.; Bar-Sagi, D.; Cole, P. A. Site-specific incorporation of a phosphotyrosine mimetic reveals a role for tyrosine phosphorylation of SHP-2 in cell signaling. *Mol. Cell* **2001**, *8*, 759-769.
- (88) Lu, W.; Shen, K.; Cole, P. A. Chemical dissection of the effects of tyrosine phosphorylation of SHP-2. *Biochemistry* **2003**, *42*, 5461-5468.
- (89) Zhang, Z.; Shen, K.; Lu, W.; Cole, P. A. The role of C-terminal tyrosine phosphorylation in the regulation of SHP-1 explored via expressed protein ligation. *J. Biol. Chem.* **2003**, *278*, 4668-4674.
- (90) Zheng, W.; Zhang, Z.; Ganguly, S.; Weller, J. L.; Klein, D. C.; Cole, P. A. Cellular stabilization of the melatonin rhythm enzyme induced by nonhydrolyzable phosphonate incorporation. *Nat. Struct. Biol.* **2003**, *10*, 1054-1057.
- (91) Zheng, W.; Schwarzer, D.; LeBeau, A.; Weller, J. L.; Klein, D. C.; Cole, P. A. Cellular stability of serotonin *N*-acetyltransferase conferred by phosphonodifluoromethylene alanine (Pfa) substitution for Ser-205. *J. Biol. Chem.* **2005**, *280*, 10462-10467.
- (92) Kee, J.-M.; Villani, B.; Carpenter, L. R.; Muir, T. W. Development of stable phosphohistidine analogues. *J. Am. Chem. Soc.* **2010**, *132*, 14327-14329.

- (93) Hahn, M. E.; Muir, T. W. Photocontrol of Smad2, a multiphosphorylated cell-signaling protein, through caging of activating phosphoserines. *Angew. Chem. Int. Ed.* **2004**, *43*, 5800-5803.
- (94) Thomas, G. L.; Payne, R. J. Phosphate-assisted peptide ligation. *Chem. Commun.* **2009**, 4260-4262.
- (95) Masania, J.; Li, J.; Smerdon, S. J.; Macmillan, D. Access to phosphoproteins and glycoproteins through semi-synthesis, native chemical ligation, and N→S acyl transfer. *Org. Biomol. Chem.* **2010**, *8*, 5113-5119.
- (96) Silversmith, R. E.; Bourret, R. B. Synthesis and characterization of a stable analog of the phosphorylated form of the chemotaxis protein CheY. *Protein Eng.* **1998**, *11*, 205-212.
- (97) Saxl, R. L.; Anand, G. S.; Stock, A. M. Synthesis and biochemical characterization of a phosphorylated analogue of the response regulator CheB. *Biochemistry* **2001**, *40*, 12896-12903.
- (98) Serwa, R.; Wilkening, I.; Del Signore, G.; Mühlberg, M.; Claußnitzer, I.; Weise, C.; Gerrits, M.; Hackenberger, C. P. R. Chemoselective Staudinger-phosphite reaction of azides for the phosphorylation of proteins. *Angew. Chem. Int. Ed.* **2009**, *48*, 8234-8239.
- (99) He, S.; Bauman, D.; Davis, J. S.; Loyola, A.; Nishioka, K.; Gronlund, J. L.; Reinberg, D.; Meng, F.; Kelleher, N.; McCafferty, D. G. Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12033-12038.
- (100) Yan, L. Z.; Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526-533.
- (101) Shogren-Knaak, M.; Ishii, H.; Sun, J.-M.; Pazin, M. J.; Davie, J. R.; Peterson, C. L. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **2006**, *311*, 844-847.
- (102) Ferreira, H.; Flaus, A.; Owen-Hughes, T. Histone modifications influence the action of Snf2 family remodelling enzymes by different mechanisms. *J. Mol. Biol.* **2007**, *374*, 563-579.

- (103) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-9252.
- (104) Manohar, M.; Mooney, A. M.; North, J. A.; Nakkula, R. J.; Picking, J. W.; Edon, A.; Fishel, R.; Poirier, M. G.; Ottesen, J. J. Acetylation of histone H3 at the nucleosome dyad alters DNA-histone binding. *J. Biol. Chem.* **2009**, *284*, 23312-23321.
- (105) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. Genetically encoding *N*-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **2008**, *4*, 232-234.
- (106) Nguyen, D. P.; Garcia Alai, M. M.; Kapadnis, P. B.; Neumann, H.; Chin, J. W. Genetically encoding *N*-methyl-L-lysine in recombinant histones. *J. Am. Chem. Soc.* **2009**, *131*, 14194-14195.
- (107) Ai, H.-w.; Lee, J. W.; Schultz, P. G. A method to site-specifically introduce methyllysine into proteins in *E. coli*. *Chem. Commun.* **2010**, *46*, 5506-5508.
- (108) Groff, D.; Chen, P. R.; Peters, F. B.; Schultz, P. G. A genetically encoded ϵ -*N*-methyl lysine in mammalian cells. *ChemBioChem* **2010**, *11*, 1066-1068.
- (109) Wang, Y.-S.; Wu, B.; Wang, Z.; Huang, Y.; Wan, W.; Russell, W. K.; Pai, P.-J.; Moe, Y. N.; Russell, D. H.; Liu, W. R. A genetically encoded photocaged *N*-methyl-L-lysine. *Mol. BioSyst.* **2010**, *6*, 1557-1560.
- (110) Neumann, H.; Hancock, S. M.; Buning, R.; Routh, A.; Chapman, L.; Somers, J.; Owen-Hughes, T.; van Noort, J.; Rhodes, D.; Chin, J. W. A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol. Cell* **2009**, *36*, 153-163.
- (111) Lammers, M.; Neumann, H.; Chin, J. W.; James, L. C. Acetylation regulates cyclophilin A catalysis, immunosuppression and HIV isomerization. *Nat. Chem. Biol.* **2010**, *6*, 331-337.
- (112) Nguyen, D. P.; Garcia Alai, M. M.; Virdee, S.; Chin, J. W. Genetically directing ϵ -*N*, *N*-dimethyl-L-lysine in recombinant histones. *Chem. Biol.* **2010**, *17*, 1072-1076.
- (113) Lindley, H. A new synthetic substrate for trypsin and its application to the determination of the amino-acid sequence of proteins. *Nature* **1956**, *178*, 647-648.
- (114) Smith, H. B.; Hartman, F. C. Restoration of activity to catalytically deficient mutants of ribulose biphosphate carboxylase/oxygenase by aminoethylation. *J. Biol. Chem.* **1988**, *263*, 4921-4925.

- (115) Simon, M. D.; Chu, F.; Racki, L. R.; de la Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **2007**, *128*, 1003-1012.
- (116) Xu, C.; Cui, G.; Botuyan, M. V.; Mer, G. Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. *Structure* **2008**, *16*, 1740-1750.
- (117) Li, B.; Jackson, J.; Simon, M. D.; Fleharty, B.; Gogol, M.; Seidel, C.; Workman, J. L.; Shilatifard, A. Histone H3 lysine 36 dimethylation (H3K36me₂) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. *J. Biol. Chem.* **2009**, *284*, 7970-7976.
- (118) Lu, X.; Simon, M. D.; Chodaparambil, J. V.; Hansen, J. C.; Shokat, K. M.; Luger, K. The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1122-1124.
- (119) Francis, N. J.; Follmer, N. E.; Simon, M. D.; Aghia, G.; Butler, J. D. Polycomb proteins remain bound to chromatin and DNA during DNA replication in vitro. *Cell* **2009**, *137*, 110-122.
- (120) Hung, T.; Binda, O.; Champagne, K. S.; Kuo, A. J.; Johnson, K.; Chang, H. Y.; Simon, M. D.; Kutateladze, T. G.; Gozani, O. ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Mol. Cell* **2009**, *33*, 248-256.
- (121) Margueron, R.; Justin, N.; Ohno, K.; Sharpe, M. L.; Son, J.; Drury, W. J.; Voigt, P.; Martin, S. R.; Taylor, W. R.; De Marco, V.; Pirrotta, V.; Reinberg, D.; Gambin, S. J. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **2009**, *461*, 762-767.
- (122) Huang, R.; Holbert, M. A.; Tarrant, M. K.; Curtet, S.; Colquhoun, D. R.; Dancy, B. M.; Dancy, B. C.; Hwang, Y.; Tang, Y.; Meeth, K.; Marmorstein, R.; Cole, R. N.; Khochbin, S.; Cole, P. A. Site-specific introduction of an acetyl-lysine mimic into peptides and proteins by cysteine alkylation. *J. Am. Chem. Soc.* **2010**, *132*, 9986-9987.
- (123) Zachara, N. E.; Hart, G. W. The emerging significance of O-GlcNAc in cellular regulation. *Chem. Rev.* **2002**, *102*, 431-438.
- (124) Rexach, J. E.; Clark, P. M.; Hsieh-Wilson, L. C. Chemical approaches to understanding O-GlcNAc glycosylation in the brain. *Nat. Chem. Biol.* **2008**, *4*, 97-106.

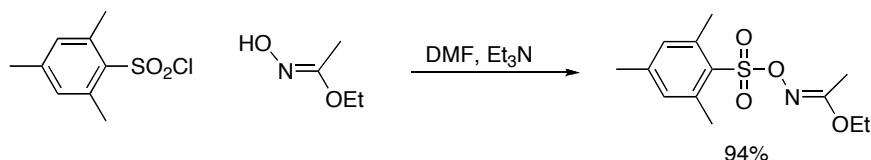
- (125) Sakabe, K.; Wang, Z.; Hart, G. W. β -N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19915-19920.
- (126) Pachamuthu, K.; Schmidt, R. R. Synthetic routes to thiooligosaccharides and thioglycopeptides. *Chem. Rev.* **2006**, *106*, 160-187.
- (127) Bernardes, G. J. L.; Gamblin, D. P.; Davis, B. G. The direct formation of glycosyl thiols from reducing sugars allows one-pot protein glycoconjugation. *Angew. Chem. Int. Ed.* **2006**, *45*, 4007-4011.
- (128) Fernández-González, M.; Boutureira, O.; Bernardes, G. J. L.; Chalker, J. M.; Young, M. A.; Errey, J. C.; Davis, B. G. Site-selective chemoenzymatic construction of synthetic glycoproteins using endoglycosidases. *Chem. Sci.* **2010**, *1*, 709-715.
- (129) You, Y. O.; Levengood, M. R.; Ihnken, L. A. F.; Knowlton, A. K.; van der Donk, W. A. Lactacin 481 synthetase as a general serine/threonine kinase. *ACS Chem. Biol.* **2009**, *4*, 379-385.
- (130) Stadtman, T. C. Emerging awareness of the critical roles of S-phospho-cysteine and selenophosphate in biological systems. *BioFactors* **1994**, *4*, 181-185.
- (131) Li, B.; Carey, M.; Workman, J. L. The role of chromatin during transcription. *Cell* **2007**, *128*, 707-719.
- (132) Groth, A.; Rocha, W.; Verreault, A.; Almouzni, G. Chromatin challenges during DNA replication and repair. *Cell* **2007**, *128*, 721-733.
- (133) Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41-45.
- (134) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins. *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053.
- (135) Guo, J.; Wang, J.; Lee, J. S.; Schultz, P. G. Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 6399-6401.

3.5 Experimental Procedures

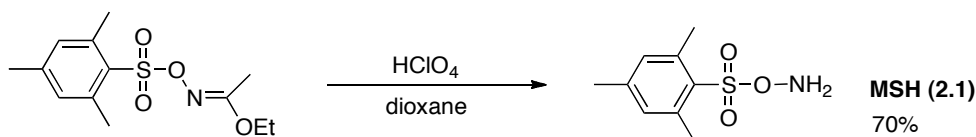
General Considerations

See Chapter 2 Experimental Procedures for general considerations in synthesis and protein analysis.

O-mesitylenesulfonylhydroxylamine (MSH, 2.1) Synthesis



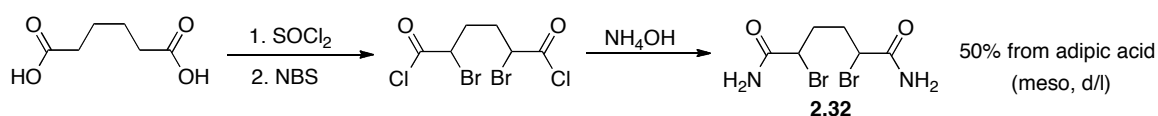
Ethyl *N*-hydroxyacetimidate (2.36 g, 22.9 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (15 mL). Et₃N (6.4 mL, 46 mmol) was added and the solution was cooled to 0 °C. 2-Mesitylenesulfonyl chloride (5.00 g, 22.9 mmol) was added as a solid in several portions. The reaction was stirred vigorously at 0 °C for 30 minutes. A white solid was formed over the course of the reaction. After this time, the reaction was diluted with EtOAc (150 mL) and then washed with H₂O (3 × 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure to give a pale yellow oil that crystallized on storage at -20 °C (6.11 g, 94%). No purification was required. m.p. = 51-53 °C (Lit¹ = 57-58 °C). ¹H NMR (400 MHz, CDCl₃): δ = 1.19 (3H, t, *J* = 7.1, OCH₂CH₃), 2.04 (3H, s, CH₃ imidate), 2.31 (3H, s, Me_{Ar}), 2.64 (6H, s, Me_{Ar}), 3.90 (2H, q, *J* = 7.1, OCH₂CH₃), 6.97 (2H, s, CH_{Ar}). ¹³C NMR (100 MHz, CDCl₃): δ = 14.0 (OCH₂CH₃), 14.9 (CH₃ imidate), 21.1 (Me_{Ar}), 22.8 (Me_{Ar}), 63.6 (OCH₂CH₃), 130.3, 131.5, 140.7, 143.3 (4 × C_{Ar}), 169.7 (C=N).



The MSH precursor (6.11 g, 21.4 mmol) was dissolved in dioxane (8 mL) at room temperature and then cooled to 0 °C. Perchloric acid (4 mL, 70% solution) was added dropwise over 2 minutes. The reaction was stirred vigorously at 0 °C and thickened and solidified over the course of 5 minutes. The reaction was incubated for an additional 5 minutes to ensure complete hydrolysis and then poured onto 50 g of ice. Et₂O (100 mL) and H₂O (100 mL) were both added to the quenched mixture and then transferred to a

separatory funnel. The organic layer was separated and then dried / neutralized for 2 minutes with K_2CO_3 . The solid was then removed by filtration and rinsed with Et_2O . The filtrate was concentrated carefully under reduced pressure to about 50 mL total volume. The solution was poured into a beaker with 150 mL ice cold petrol. MSH crystals formed immediately. The beaker was incubated at 0 °C for 30 minutes to complete the crystallization. The crystals were isolated by filtration and then dried under high vacuum for 15 minutes. The white needles (3.21 g, 70%) were stored in a plastic tube, sealed only with wax film. m.p. = 90-91 °C (Lit¹ = 95-96 °C). 1H NMR (400 MHz, $CDCl_3$): δ = 2.33 (3H, s, Me_{Ar}), 2.65 (6H, s, Me_{Ar}), 5.56 (2H, br. s, NH_2), 7.01 (2H, s, CH_{Ar}). ^{13}C NMR (50 MHz, $CDCl_3$): δ = 21.0 (Me_{Ar}), 22.6 (Me_{Ar}), 129.0, 131.6, 140.9, 143.7 ($4 \times C_{Ar}$). Found: C, 50.18%; H, 5.90%, N, 6.27%. $C_9H_{15}NO_4$ requires: C, 50.21%; H, 6.09%; N, 6.51%.

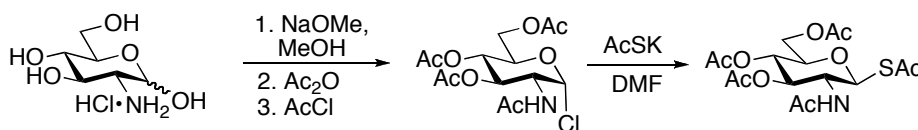
Synthesis of dibromide 2.32



α,α' -Di-bromo-adipyl(bis)amide was prepared using a procedure adapted from the literature.^{2, 3} Adipic acid (25.00 g, 171.1 mmol) was added to a 500 mL round bottom flask and suspended in thionylchloride (75.0 mL, 1.03 mol). The flask was equipped with a condenser and the reaction was heated to reflux (open to air, bath temp 80 °C). After 30 minutes at reflux, all adipic acid had dissolved. The reaction was stirred for an additional 60 minutes at reflux and then cooled to room temperature. CCl_4 (100 mL) was added to the reaction followed by NBS (73.1 g, 411 mmol). The reaction was stirred vigorously and 10 drops of HBr (48% aq.) was added by pipette. The reaction was heated to reflux, again open to air. The reaction gradually turns from red to black over the course of an hour. After 2 hours at reflux, the reaction was cooled to room temperature and then to 0 °C. The mixture was stirred at 0 °C to ensure all succinimide had precipitated. The solid was removed by filtration. Et_2O (50 mL) was used to rinse and complete the filtration. The filtrate was concentrated under reduced pressure to give a thick, dark red liquid. In a 500 mL round bottom flask, 200 mL of NH_4OH (25% aq.) was cooled to 0 °C. The crude acid chloride was added dropwise over 20 minutes to the ammonia solution with rapid

stirring. After the addition was complete, the reaction was stirred vigorously at 0 °C for 1 hour. The bis-amide product precipitated from the reaction mixture. The dark solid was isolated by filtration and partially dried. The product was purified by triturating in MeOH/H₂O. The solid was suspended in H₂O (100 mL) and MeOH (100 mL) and heated to 60 °C. The mixture was stirred rapidly at 60 °C for 30 minutes. After this time, the mixture was cooled to room temperature. The resulting white solid (a mixture of *meso* and *d/l* diastereomers) was isolated by filtration and washed with MeOH (200 mL). The product (**2.32**) was dried under high vacuum (25.75 g, 50%). IR (ν_{max} , KBr): 3302, 2946, 2801, 1684, 1418, 1319, 1277, 1250, 1220, 1194, 978, 876. ¹H NMR (400 MHz, DMSO): δ = 1.75-2.08 (4H, m, CH₂CH₂), 4.28-4.36 (2H, m, 2 × CHBr), 7.30 (2H, s), 7.69 (2H, s) (2 × NH₂). ¹³C NMR (100 MHz, DMSO): (both diastereomers reported), δ = 32.5, 32.6 (CH₂CHBr), 48.2, 48.5 (CH₂CHBr), 169.87, 169.92 (C=O). LRMS *m/z* (ESI⁺): 324.9 [M+Na]⁺.

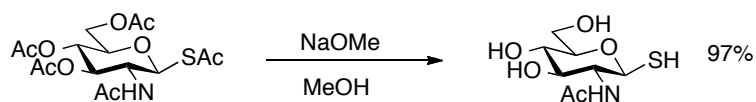
3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio- β -D-glucosamine



(D)-Glucosamine hydrochloride (20.0 g, 92.7 mmol) was added to a 100 mL round bottom flask and suspended in MeOH (20 mL). Sodium methoxide (6.00 g, 111 mmol) was added in several portions and the mixture stirred vigorously at room temperature for 1 hour before NaCl was removed by filtration. The solids were washed with MeOH (20 mL). Acetic anhydride (12 mL, 127 mmol) was added to the stirred filtrate at 0 °C and warmed slowly to room temperature. After 12 hours, the reaction was concentrated to ~20 mL total volume and the product precipitated by the addition of Et₂O (100 mL). The solid was isolated by filtration and dried under vacuum before returning to an oven dried 2-neck 250 mL round bottom flask equipped with a condenser. The flask was placed under argon atmosphere and cooled to 0 °C. Acetyl chloride (40.0 mL, 563 mmol) was added slowly and the mixture stirred vigorously. After the addition, the reaction was heated to reflux. Gradually the reaction turned black. After 3 hours at reflux, the reaction was cooled to room temperature, diluted with CH₂Cl₂ (200 mL), and poured into a beaker of ice (~200 g). The organics were washed with NaHCO₃ (sat. aq., 2 × 200 mL), dried

(MgSO₄), filtered, and concentrated to a black residue. The residue was dissolved in EtOAc and filtered through a short column of silica, eluting with EtOAc. The product was precipitated by the addition of cold Et₂O and petrol. The resulting powder was isolated by filtration and dried under vacuum (10.19 g, 50% from glucosamine hydrochloride). The crude product (~4:1 mixture of α : β) was used in the final step without purification. Accordingly, a portion of the anomeric chloride (1.00 g, 2.73 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (15 mL). Potassium thioacetate (1.56 g, 13.7 mmol) was then added in one portion. The reaction was flushed with nitrogen and stirred at room temperature for 3 hours. After this time, no starting material was detected by TLC (70% EtOAc in petrol). The reaction mixture was diluted with EtOAc (250 mL) and washed sequentially with NaHCO₃ (sat. aq., 200 mL), H₂O (200 mL), and brine (200 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure to give a brown solid. This solid was recrystallized from CH₂Cl₂ and petrol to afford the product as white crystals (661 mg, 60%). m.p. = 186-188 °C (Lit⁴: 196-197, *i*PrOH). [α]_D²⁰ = +12.1 (c = 2.2, CHCl₃) (Lit⁴: +14.1 c = 2.16, CHCl₃). IR (ν_{max} , KBr): 3733, 3064, 2961, 2360, 2341, 1747, 1707, 1665, 1538, 1433, 1370, 1230, 1048. ¹H NMR (400 MHz, CDCl₃): δ = 1.90 (3H, s), 2.02 (6H, s), 2.05 (3H, s), 2.35 (3H, s) (5 $\times$ OAc), 3.81 (1H, ddd, *J* = 9.4, 4.6, 1.3, H5), 4.08 (1H, dd, *J* = 12.4, 1.3, H6), 4.22 (1H, dd, *J* = 12.4, 4.6, H6'), 4.34 (1H, q, *J* = 10.3, H2), 5.10-5.16 (2H, m, H3, H4), 5.17 (1H, d, *J* = 10.3, H1). ¹³C NMR (100 MHz, CDCl₃): δ = 20.49, 20.59, 20.64, 23.0, 30.8 (5 $\times$ OAc), 51.9 (C2), 61.8 (C6), 67.8, 74.0 (C3, C4), 76.3 (C5), 81.5 (C1), 169.2, 170.0, 170.6, 171.3, 193.5 (5 $\times$ C=O). LRMS (ESI⁺): 428.1 (M+Na)⁺.

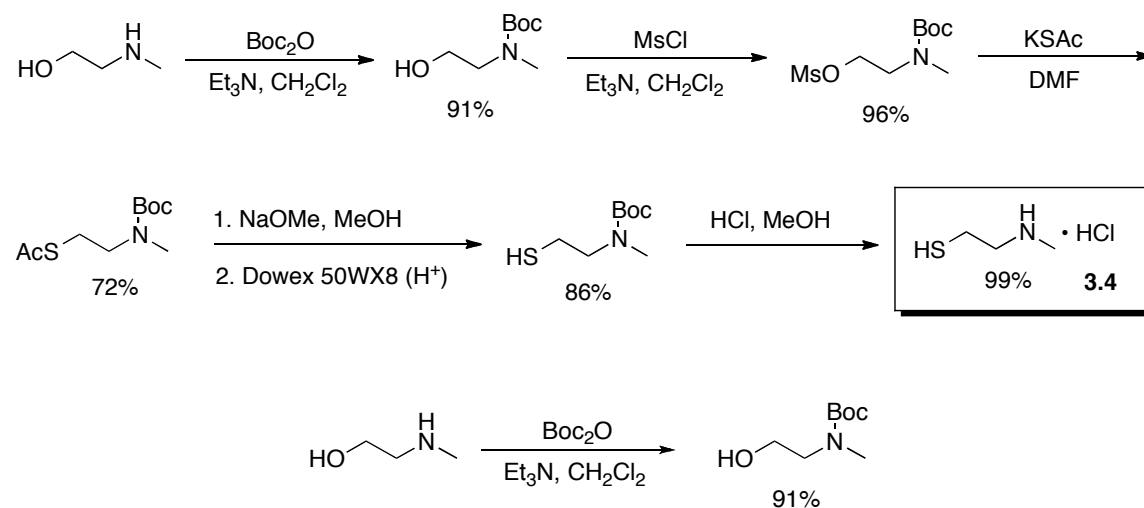
***N*-acetyl-1-thio- β -D-glucosamine (GlcNAc thiol **2.37**)**



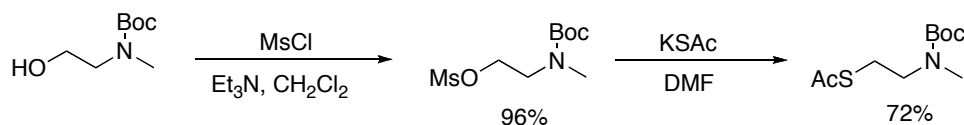
3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio- β -D-glucosamine (250 mg, 0.62 mmol) was added to a 50 mL round bottom flask and placed under an atmosphere of nitrogen before dissolving in anhydrous MeOH (8 mL). Sodium methoxide (67 mg, 1.2 mmol) was added under a stream of nitrogen at room temperature. The reaction was stirred for 1 hour at room temperature before quenching with DOWEX-50WX8 (H⁺ form) until the pH of the reaction mixture was neutral (pH paper). The resin was removed by filtration and washed with MeOH (~ 50 mL). The filtrate was evaporated to provide GlcNAc thiol **2.37** as a

thick syrup that crystallized upon standing (142 mg, 97%). m.p. = 170-171 °C (Lit⁵: 174-176 °C). $[\alpha]_D^{20} = -13.5$ (c = 1.0, MeOH) (Lit⁵: -10.4, c = 1.0, MeOH). IR ($\nu_{\max}$, KBr): 3287, 2360, 2341, 1647, 1558, 1374, 1313, 1056. ¹H NMR (400 MHz, D₂O): δ = 1.90 (3H, s, Ac), 3.30-3.39 (3H, m, H3, H4, H5), 3.56 (1H, dd, J = 11.1, 5.3, H6), 3.59 (1H, t, J = 10.2, H2), 3.74 (1H, dd, J = 11.1, 1.2, H6'), 4.53 (1H, d, J = 10.2, H1). ¹³C NMR (50 MHz, D₂O): δ = 22.6 (Ac), 58.2 (C2), 61.1 (C6), 70.0, 75.3, 79.3, 80.52 (C3, C4, C5, C1). LRMS (ESI⁻): 236.1 (M-H)⁻. Note: in D₂O this material anomerizes to a 3.6:1.0 mixture of β : α over a period of 24 hours.

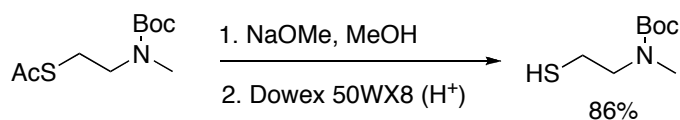
Synthetic Scheme for Thiol Precursor to Monomethyllysine Analog (3.4)



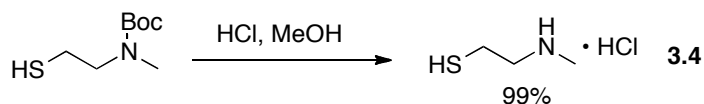
N-methylethanolamine (4.13 g, 55.0 mmol) was added to a 250 mL round bottom flask and dissolved in CH_2Cl_2 (100 mL). Et_3N (9.6 mL, 69 mmol) was added and the stirred solution was cooled to 0 °C. Boc_2O (10.00 g, 45.82 mmol) was added and the reaction was stirred for 5 min at 0 °C and then 4 hours at room temperature. After this time, the reaction was cooled to 0 °C and quenched with 1M HCl (200 mL). The mixture was diluted with CH_2Cl_2 (100 mL) and the organic layer was separated. After washing with 1M HCl (2 $\times$ 100 mL), the organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting oil (7.30 g, 91%) did not require further purification. Spectroscopic data was consistent with that previously reported.⁶ IR ($\nu_{\max}$, film): 3441, 2977, 1681, 1484, 1396, 1305, 1226, 1156. ¹H NMR (CDCl_3 , 400 MHz): δ = 1.47 (9H, s, Boc), 2.93 (3H, s, NMe), 3.40 (2H, t, J = 4.8, $\text{CH}_2\text{CH}_2\text{OH}$), 3.74-3.78 (2H, m, CH_2OH). ¹³C NMR (CDCl_3 , 50 MHz): δ = 28.3 (Boc), 35.3 (NMe), 51.2 ($\text{CH}_2\text{CH}_2\text{OH}$), 60.9 (CH_2OH), 79.7 (Boc), 157.0 (C=O). LRMS m/z (ESI⁺): 198.1 $[\text{M}+\text{Na}]^+$.



2-[*N*-(*tert*-Butyloxycarbonyl)-*N*-methylamino]ethanol (7.21 g, 41.2 mmol) was added to a 250 mL round bottom flask and dissolved in CH₂Cl₂ (85 mL). Et₃N (7.5 mL, 53.5 mmol) was added and the stirred solution was cooled to 0 °C. Methanesulfonyl chloride (3.34 mL, 43.2 mmol) was added dropwise and 10 mL CH₂Cl₂ was used to rinse the flask on completion of the addition. The flask was stoppered and stirred at 0 °C for 1 hour. After this time, the reaction was quenched at 0 °C with 1M HCl (100 mL). After diluting with 50 mL CH₂Cl₂, the organic layer was separated and then washed with 1M HCl (2 × 100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The yellow oil obtained was sufficiently pure to not require further purification (9.96 g, 96%). However, this material is not stable and was used immediately in the next transformation. Accordingly, the product mesylate (9.56 g, 37.8 mmol) was added to a 250 mL round bottom flask and dissolved in DMF (100 mL). Potassium thioacetate (6.47 g, 56.6 mmol) was added and the reaction was stirred vigorously as the reaction became viscous. The flask was stoppered and the reaction heated to 60 °C. After 1 hour at 60 °C, the reaction was cooled to room temperature. TLC (50% EtOAc in petrol) revealed complete consumption of starting material (*R*_f = 0.3) and the formation of thioester product (*R*_f = 0.7). The reaction mixture was diluted with Et₂O (200 mL) and washed with H₂O (3 × 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (20% EtOAc in petrol) to provide the product thioester as a yellow oil (6.32 g, 72%). IR (*ν*_{max}, film): 2976, 2932, 1698, 1480, 1457, 1425, 1392, 1365, 1301, 1236, 1214. ¹H NMR (CDCl₃, 200 MHz): δ = 1.40 (9H, s, Boc), 2.28 (3H, s), 2.84 (3H, s) (NMe and SAc), 2.95 (2H, t, *J* = 7.1), 3.30 (2H, t, *J* = 7.1) (CH₂CH₂). ¹³C NMR (CDCl₃, 50 MHz): δ = 27.0 (CH₂SAc), 28.2 (Boc), 30.5 (SAc), 34.6 (NMe), 48.1 (CH₂NMe), 79.5 (Boc), 155.4, 195.1 (2 × C=O). HRMS *m/z* (ESI⁺): Found 256.0977 [M+Na]⁺; C₁₀H₁₉NO₃SNa requires 256.0978.

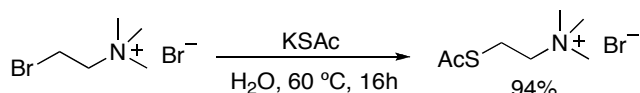
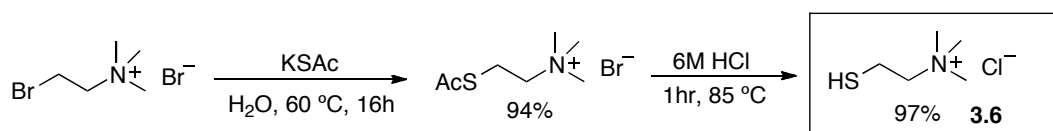


2-[*N*-(*tert*-Butyloxycarbonyl)-*N*-methyamino]acetylethanthiol (500 mg, 2.14 mmol) was added to a 25 mL round bottom flask and dissolved in MeOH (5 mL). The flask was sealed and placed under a nitrogen atmosphere. NaOMe (173 mg, 3.21 mmol) was added and the reaction was stirred at room temperature for 30 minutes. After this time the reaction was quenched with Dowex-50WX8 (H⁺) until the pH was neutral (pH paper). The resin was removed by filtration and washed with MeOH (25 mL). The filtrate was concentrated under reduced pressure and the resulting residue was purified by column chromatography (20% EtOAc in petrol). The product thiol was isolated as a yellow oil (353 mg, 86%). IR (ν_{max} , film): 3004, 2976, 2554, 2360, 2341, 1697, 1480, 1457, 1422, 1392, 1365, 1301, 1241, 1214, 1169. ¹H NMR (CDCl₃, 200 MHz): δ = 1.25 (1H, t, *J* = 8.3, SH), 1.38 (9H, s, Boc), 2.53-2.64 (2H, m, CH₂SH), 2.82 (3H, s, NMe), 3.30 (2H, t, *J* = 7.3, CH₂CH₂SH). ¹³C NMR (CDCl₃, 50 MHz): δ = 22.4 (CH₂SH), 28.2 (Boc), 34.6 (NMe), 52.0 (CH₂CH₂SH), 79.4 (Boc), 155.4 (C=O). HRMS *m/z* (ESI⁺): Found 214.0870 [M+Na]⁺; C₈H₁₇NO₂SNa requires 214.0872.

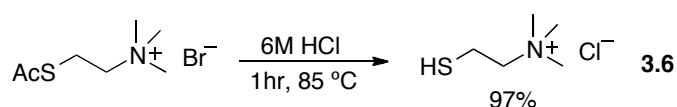


A 50 mL round bottom flask was flame dried and placed under an atmosphere of nitrogen. MeOH (4 mL) was added and the flask was cooled to 0 °C. Acetyl chloride (1.7 mL, 24 mmol) was added and the reaction was stirred at 0 °C for 10 min to generate a solution of HCl in MeOH. 2-[*N*-(*tert*-Butyloxycarbonyl)-*N*-methyamino]ethanthiol (302 mg, 1.58 mmol) was added to the reaction flask as a solution in 1 mL MeOH. An additional 1 mL MeOH was used to complete the transfer. The ice bath was removed and the reaction was stirred for 30 minutes, after which time TLC (20% EtOAc in petrol) revealed complete consumption of starting material. The solvent was removed under reduced pressure to give the product as an extremely hygroscopic residue (200 mg, 99%). ¹H NMR (CDCl₃, 200 MHz): δ = 2.77 (3H, s, NMe), 2.85 (2H, t, *J* = 6.7), 3.24 (2H, t, *J* = 6.7) (CH₂CH₂). ¹³C NMR (CDCl₃, 50 MHz): δ = 21.1 (CH₂SH), 33.7 (NMe), 52.8 (CH₂CH₂SH).

Synthetic Scheme for Thiol Precursor to Trimethyllysine Analog (3.6)

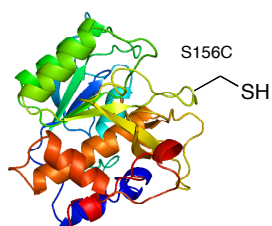


2-(Bromoethyl)-trimethylammonium bromide (5.00 g, 20.2 mmol) was added to a 250 mL round bottom flask and dissolved in H₂O (25 mL). Potassium thioacetate (3.01 g, 26.3 mmol) was added to the stirred solution and the resulting mixture was heated to 60 °C. The flask was stoppered and stirred at 60 °C for 16 hours. After this time, the reaction mixture was concentrated under reduced pressure to give a brown solid. The product was extracted by stirring the solid in 100 mL of MeOH:CH₂Cl₂ (1:1) at room temperature for 15 minutes. KBr was removed by filtration and the filtrate was concentrated under reduced pressure. The extraction and filtration was repeated twice more to ensure removal of KBr. After concentrating under reduced pressure, the resulting reddish solid was sufficiently pure to carry on to the next step (4.59 g, 94%). IR (ν_{max} , KBr): 1694, 1563, 1484, 1420, 1355, 1267, 1200, 1132, 1107, 1059, 1024, 1001, 971. ¹H NMR (D₂O, 400 MHz): δ = 2.31 (3H, s, SAc), 3.09 (9H, s, NMe₃), 3.17-3.21 (2H, m), 3.35-3.39 (2H, m) (CH₂CH₂). ¹³C NMR (MeOD, 50 MHz): δ = 23.1, 30.6, 53.9, 66.3, 196.2. LRMS m/z (ESI⁺): 162.1 [M]⁺.



2-(acetothioethyl)-trimethylammonium bromide (1.18 g, 4.88 mmol) was added to a 100 mL round bottom flask and dissolved in 6M HCl (10 mL). The stirred solution was stoppered and heated to 85 °C. After 1 hour, the reaction was concentrated under reduced pressure and dried under high vacuum to provide the product thiol as a pale yellow, very hygroscopic solid (740 mg, 97%). This material was used without further purification. ¹H NMR (D₂O, 400 MHz): δ = 2.82-2.86 (2H, m, CH₂), 3.04 (9H, s, NMe₃), 3.42-3.46 (2H, m, CH₂). ¹³C NMR (MeOD, 50 MHz): δ = 18.1, 53.9, 69.5. LRMS m/z (ESI⁺): 120.1 [M]⁺.

Protein Modification



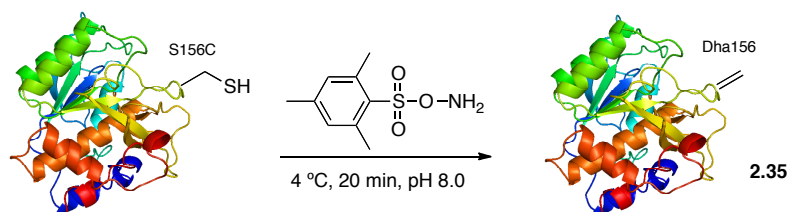
SBL-S156C (PDB for wild type = 1GCI)

Sequence:

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPS
TQDGNGHGHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGL
EWAGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGN**C**GAGSISYPA
RYANAMAVGATDQNNNRASFQYGAGLDIVAPGVNVQSTYPGSTYASLNGTSM
ATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAAEAATR

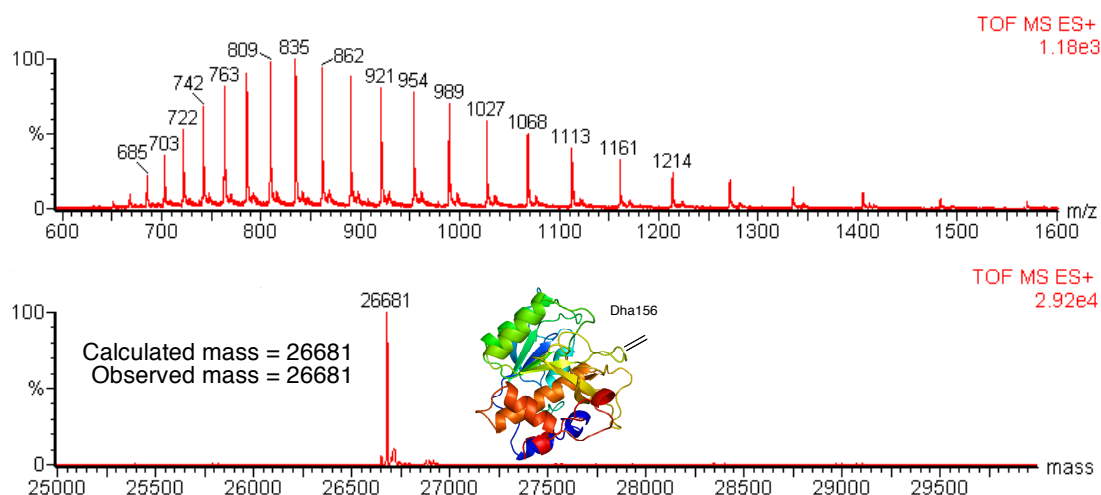
Average Mass = 26715

SBL-156Dha (2.35) using MSH

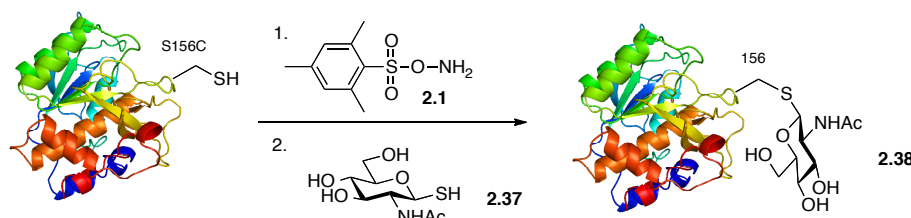


SBL-156Dha was prepared according to the procedure established in Chapter 2. This protocol is repeated here for reference. For full details and control experiments, the reader is referred to Chapter 2.

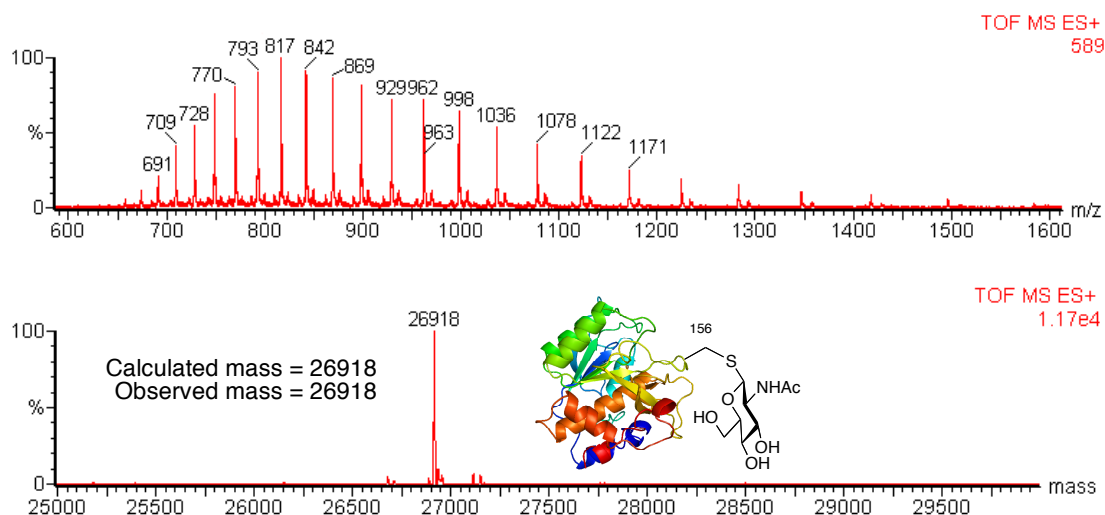
SBL-S156C was prepared as a 1.4 mg/mL solution in 50 mM sodium phosphate buffer (pH 8.0) and 1.00 mL (0.052 μ mol) was added to a 1.50 mL plastic tube. MSH (1.2 mg, 5.6 μ mol) was added as a solution in DMF (50 μ L) and the reaction vortexed immediately upon addition. The homogenized sample was shaken at 4 °C for 20 minutes before a 40 μ L aliquot was analyzed by LC-MS. A mass corresponding to SBL-Dha156 was observed (calculated mass = 26681; observed mass = 26681). This protein does not react with Ellman's reagent and a variety of thiols can be added by conjugate addition (see next). ESI-MS are shown below.



Synthesis of GlcNAc cysteine on SBL at position 156 (**2.38**)



All manipulations were carried out in a cold room maintained at 4 °C. A 1 mg sample of lyophilized SBL-S156C (0.037 μmol) was dissolved in 1.00 mL in pH 8.0 sodium phosphate buffer (50 mM). A solution of MSH was prepared by dissolving 1.8 mg (8.4 μmol) in 100 μL DMF. A 50 μL portion of the MSH solution was added to the protein by micropipette. The reaction was vortexed periodically over 1 minute and then rotated on a lab rotisserie for an additional 19 minutes at 4 °C. A 50 μL aliquot was analyzed by LC-MS to confirm the conversion of Cys156 to Dha156 (calculated mass = 26681; observed mass = 26681). To the reaction mixture was added GlcNAc thiol **2.37** as a solid (8.8 mg, 1000 eq.) to give a 39 mM solution in thiol. After 90 minutes of shaking at 4 °C, the reaction was analyzed directly by LC-MS. Complete conversion to SBL-156SGlcNAc (**2.38**) was observed. (calculated mass = 26918, observed mass = 26918). The protein was purified by dialysis (2 $\times$ against 4L pH 7.0 sodium phosphate, 50 mM). Protein recovery after dialysis was 0.8 mg (80%, BCA assay) ESI-MS are shown below.

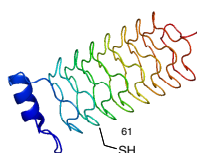


Np276 modification

The following Np276 mutant was expressed and purified by Matthew Young. For the expression and purification of Np276 and further use of this model protein, the reader is referred to the published procedures.⁷

Np276 (PDB = 2J8K)

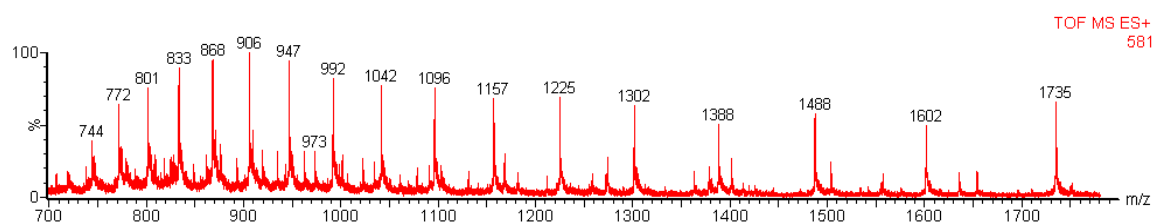
Sequence:

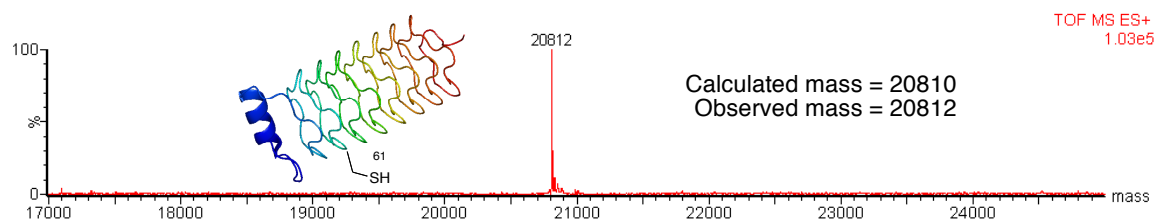


GSSHHHHHHSSGLVPRGSHIDVGKLRQLYAAGERDFSIVDLRGAVLENINLSGAI
 LHGA⁶¹LDEANLQQANLSRADLSGATLNGADLRGANLSKADLSDAILDNAILEGA
 ILDEAVLNQANLKAANLEQAILSHANIREADLSEANLEAADLSGADLAIADLHQA
 NLHQAALERANLTGANLEDANLEGTTILEGGNNNLAT

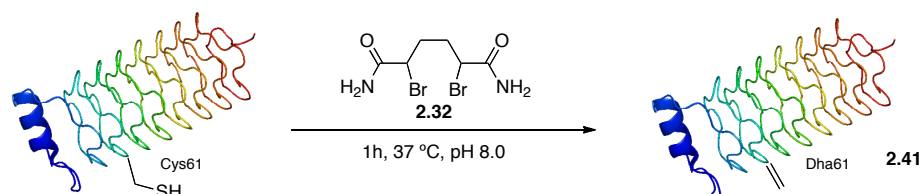
Average Mass = 20810

ESI-MS for unmodified Np276 mutant is shown below



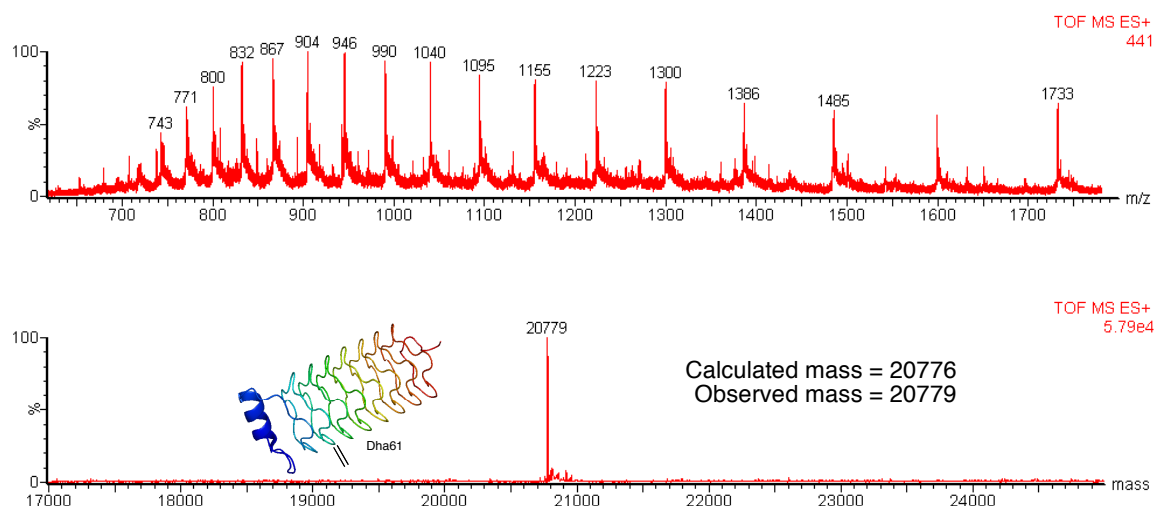


Dha formation on Np276 (**2.41**) using dibromide **2.32**

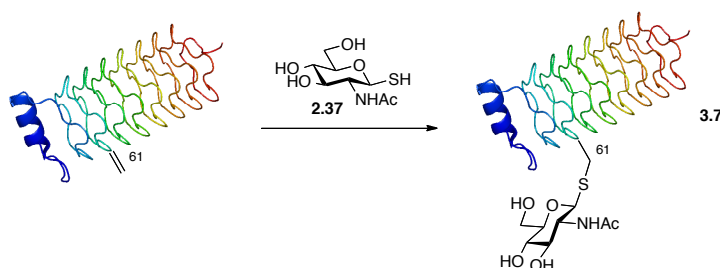


Dha was installed at position 61 on Np276 as described in Chapter 2. The protocol is reproduced here for reference. The reader is referred to Chapter 2 for further details and control experiments.

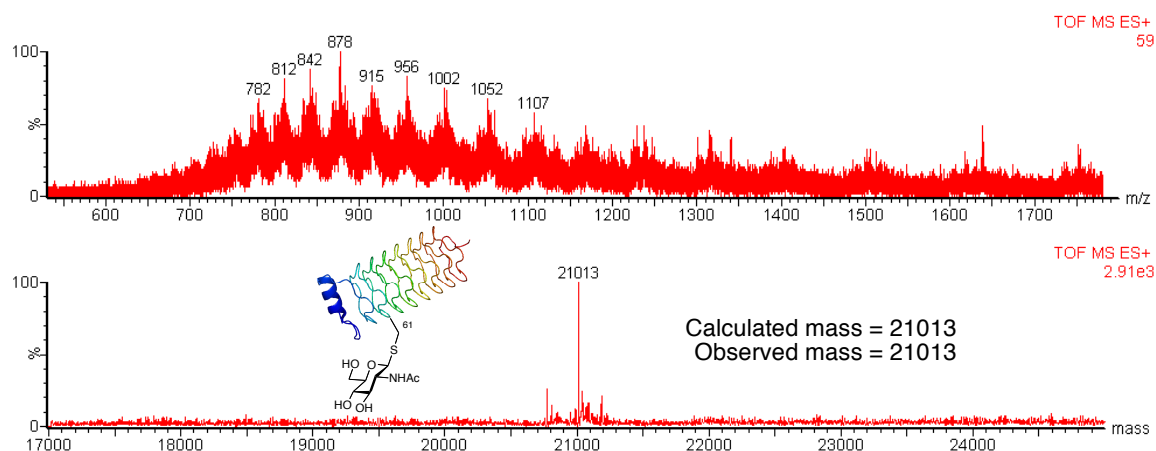
Np276-Cys61 (2.5 mL, 1.1 mg / mL in pH 8.0 sodium phosphate, 50 mM) was reduced by the addition of 10 mg DTT. The solution was shaken at room temperature for 15 minutes to reduce any contaminant disulfide. After this time, the protein solution was passed through a PD10 column (GE Healthcare), eluting with the same buffer. The resulting protein solution was used immediately. Accordingly, 2.5 mL of the NP276 protein ($c = 0.79$ mg/mL) was added to a 15 mL plastic tube and kept at room temperature. A stock solution of dibromide **2.32** was prepared by dissolving 218 mg in 1.91 mL DMF. A 125 μ L aliquot of the dibromide solution was added to the protein solution and the reaction was vortexed to homogenize and immediately placed in a 37 °C incubator with shaking. After 15 minutes, an additional 125 μ L aliquot of the dibromide solution was added to the reaction. After another 15 minutes, a third and final 125 μ L aliquot of the dibromide solution was added. The reaction was shaken at 37 °C for a final 30 minutes and then analyzed by LC-MS. A protein mass corresponding to the Dha protein **2.41** was observed (calculated mass = 20776; observed mass = 20779). Precipitated dibromide was removed by centrifugation and the protein solution was stored on ice. ESI-MS are shown below.



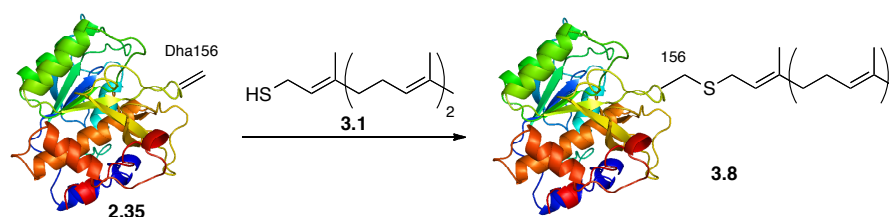
Synthesis of Np276 glycoprotein 3.7



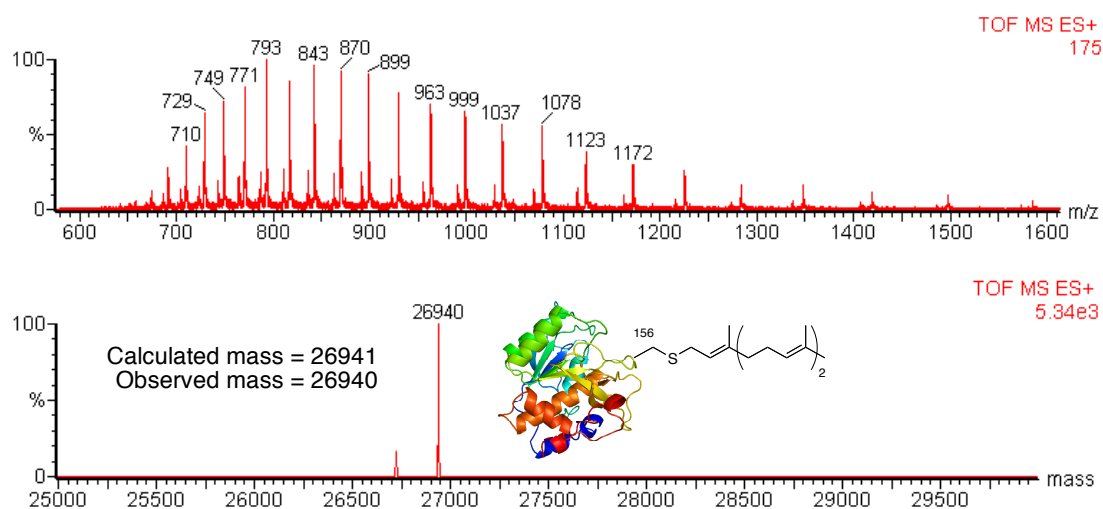
A 1.00 mL aliquot of the Dha containing Np276 ($c = 0.79$ mg/mL) was mixed with 135 mg of GlcNAc thiol **2.37**. The reaction was shaken at 37 °C for 2 hours. After this time, small molecules were removed by dialysis (2 × against 4L pH 7.0 sodium phosphate, 50 mM). Full conversion to the thioether **2.37** was observed by LC-MS (calculated mass = 21013; observed mass = 21013). Protein recovery after dialysis was 0.7 mg (88%, BCA). ESI-MS are shown below.



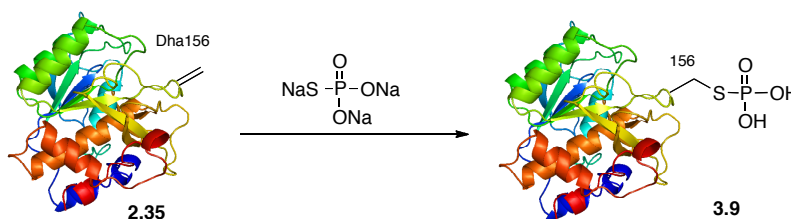
Synthesis of farnesyl cysteine on SBL at position 156 (3.8)



A 200 μL aliquot of 0.7 mg/mL SBL-Dha156 (**2.35**) previously prepared using MSH was thawed and kept on ice. A (0.35 M) farnesyl thiol solution in DMSO was prepared alongside an aqueous solution of TCEP $\cdot\text{HCl}$. (Farnesyl thiol **3.1** was prepared by Gonalo Bernardes according to the published procedure)⁸ The TCEP was neutralized to pH 7.0 with NaOH to give a final concentration of 0.20 M TCEP. The farnesyl thiol (15 μL) and TCEP solution (52 μL) were added in succession to the protein to give a cloudy emulsion. The reaction was rotated on a lab rotisserie for 90 minutes at room temperature and then analyzed directly by LC-MS. A protein species with a mass of 26940 was found which corresponds to the farnesyl thioether sodium adduct (calculated mass = 26941). ESI-MS are shown below.

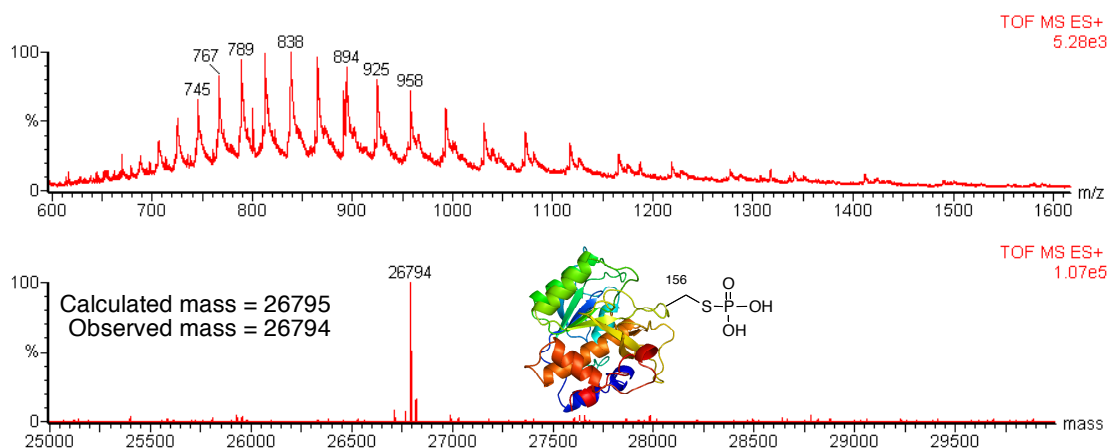


Synthesis of phosphocysteine on SBL at position 156 (3.9)

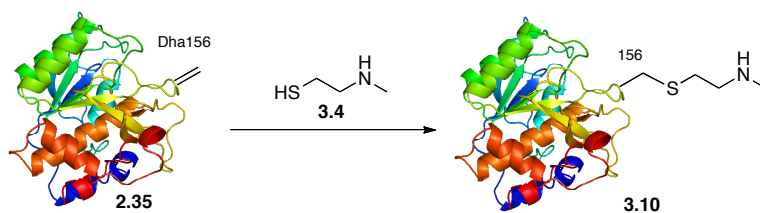


All manipulations were carried out in a cold room maintained at 4 $^{\circ}\text{C}$. A 250 μL aliquot of 1.0 mg/mL SBL-Dha156 (**2.35**) (pH 8.0 sodium phosphate, 50 mM) previously

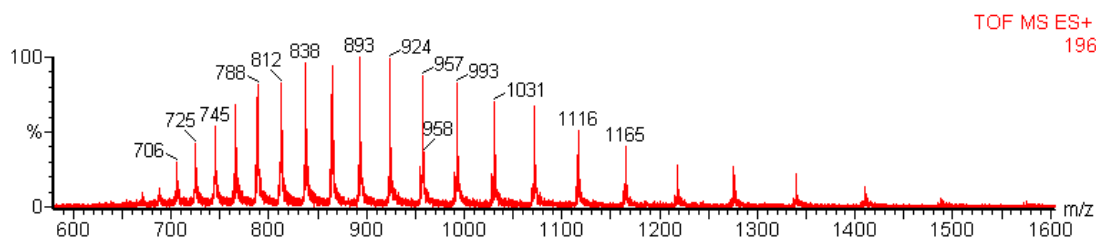
prepared was thawed and kept on ice until needed. Sodium thiophosphate (8.4 mg, 0.050 mmol) was added as a solid along with 100 μ L of pH 8.0 sodium phosphate buffer (50 mM). The reaction was vortexed to homogenize and then shaken at 4 $^{\circ}$ C for 1 h. LC-MS analysis of the reaction mixture revealed complete conversion to SBL-PhosCys156 (**3.9**) (calculated mass = 26795, observed mass = 26794). ESI-MS are shown below.

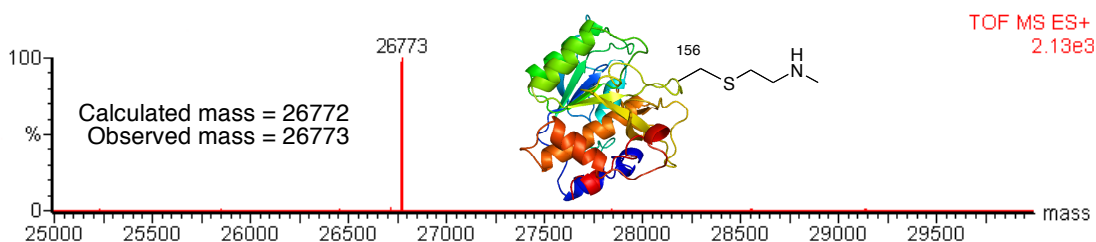


Synthesis of monomethyllysine mimic on SBL at position 156 (**3.10**)

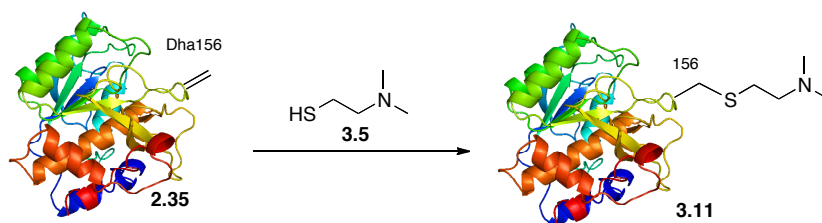


All manipulations were carried out in a cold room at 4 $^{\circ}$ C. A 200 μ L aliquot of 0.7 mg/mL SBL-Dha156 (**2.35**) previously prepared was thawed and kept on ice until needed. A solution of thiol **3.4** was prepared by dissolving 6.6 mg (0.052 mmol) in 200 μ L H₂O. K₂HPO₄ (45 mg, 0.26 mmol) was added to the thiol solution as a solid and the solution vortexed. All of the thiol solution was transferred to the protein by micropipette to give a reaction mixture of pH 9.0. The reaction was vortexed and then shaken at 4 $^{\circ}$ C for 90 minutes. LC-MS analysis revealed full conversion to the monomethyllysine mimic **3.10** (calculated mass = 26772; observed mass = 26773). ESI-MS are shown below.

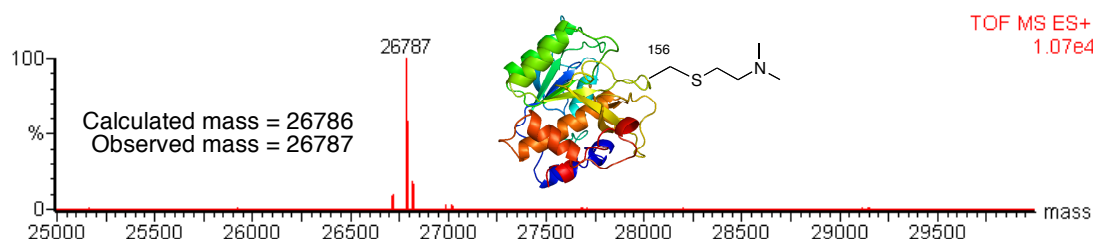
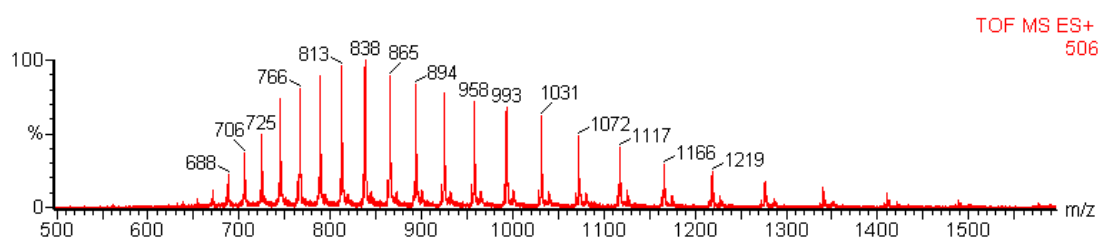




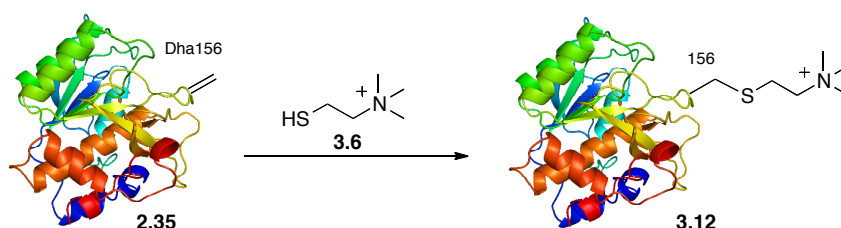
Synthesis of dimethyllysine mimic on SBL at position 156 (**3.11**)



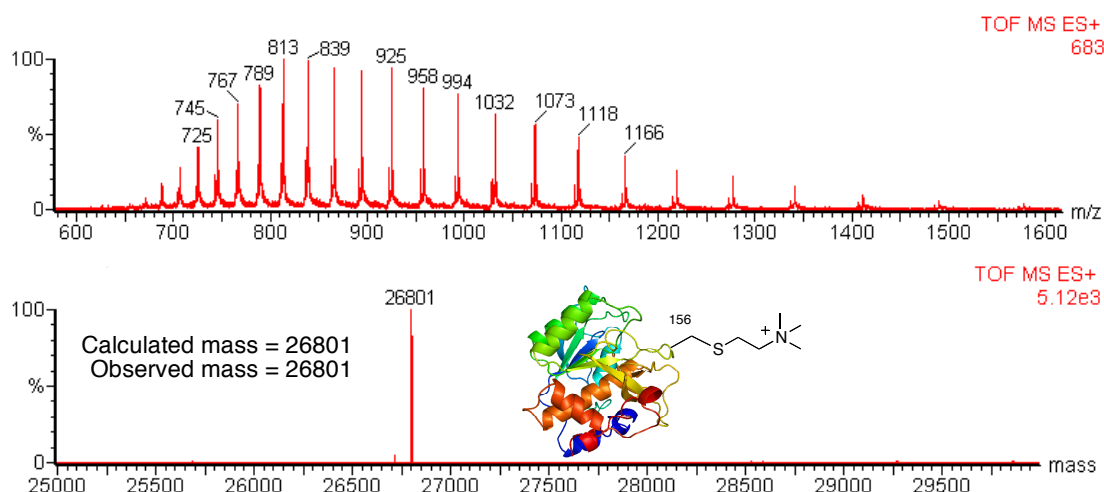
All manipulations were carried out in a cold room maintained at 4 °C. A 200 μ L aliquot of 0.7 mg/mL SBL-Dha156 (**2.35**) previously prepared was thawed and kept on ice until needed. A solution of dimethylaminothiol **3.5** was prepared by dissolving 3.2 mg (0.022 mmol) of the hydrochloride salt (Aldrich) in 150 μ L of pH 8.0 phosphate buffer (50 mM). A 50 μ L aliquot of the thiol solution was added to the protein and the reaction was shaken at 4 °C for 90 minutes at which time a 40 μ L aliquot was taken for LC-MS analysis. Full conversion to the desired dimethyllysine mimic **3.11** was observed. (calculated mass = 26786; observed mass = 26787). ESI-MS are shown below.



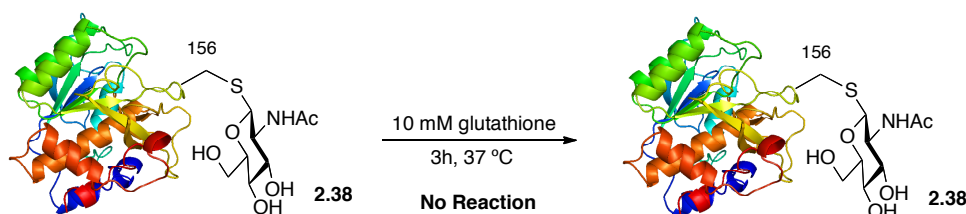
Synthesis of trimethyllysine mimic on SBL at position 156 (**3.12**)



All manipulations were carried out in a cold room maintained at 4 °C. A 200 μL aliquot of 0.7 mg/mL SBL-Dha156 (**2.35**) previously prepared was thawed and kept on ice until needed. A solution of thiol **3.6** was prepared by dissolving 8.2 mg of **3.6** (0.052 mmol) and 27 mg K_2HPO_4 (0.20 mmol) in 200 μL of H_2O . All of the thiol solution was added to protein to give a reaction mixture at pH 9.0. The reaction was shaken at 4 °C for 90 min before a 50 μL aliquot was analyzed by LC-MS. Full conversion to the trimethyllysine mimic **3.12** was observed. (calculated mass = 26801; observed mass = 26801). ESI-MS are shown below.

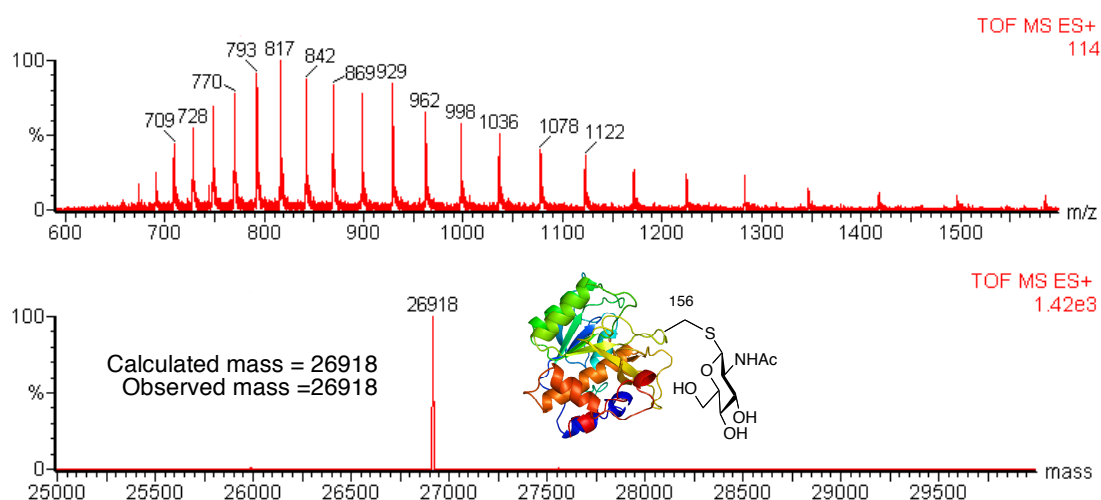


Stability of glycoprotein **2.38** in presence of glutathione

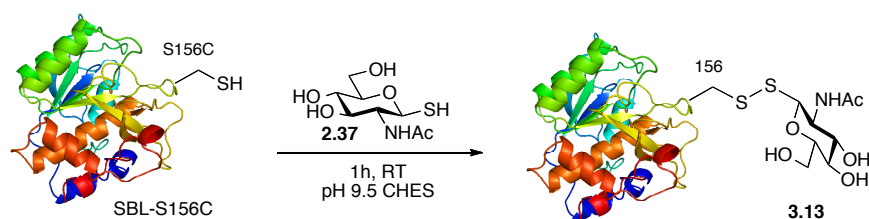


Thioether-linked GlcNAc-SBL **2.38** was prepared as described above. The product protein was purified by loading the sample in 2.50 mL of pH 8.0 sodium phosphate buffer (50 mM) onto a PD10 desalting column (GE Healthcare) previously equilibrated with 10 column volumes of pH 8.0 sodium phosphate buffer (50 mM) and eluting with 3.50 mL

of the same buffer. The collected sample was concentrated to 1 mg/L using 10K molecular weight cutoff Viva Spin column. A 250 μ L aliquot of glycoprotein **2.38** was added to a 1.5 mL plastic tube and diluted to 500 μ L with pH 8.0 sodium phosphate buffer (50 mM). Glutathione (1.5 mg, 5.0 μ mol) was added as a solid to the tube (final glutathione concentration = 10 mM) and the mixture incubated in a 37 $^{\circ}$ C shaker for 3 hours. A 40 μ L aliquot was analyzed by LC-MS to show that glycoprotein **2.38** was still intact and the only protein species detected (calculated mass = 26918; observed mass = 26918). ESI-MS are shown below.



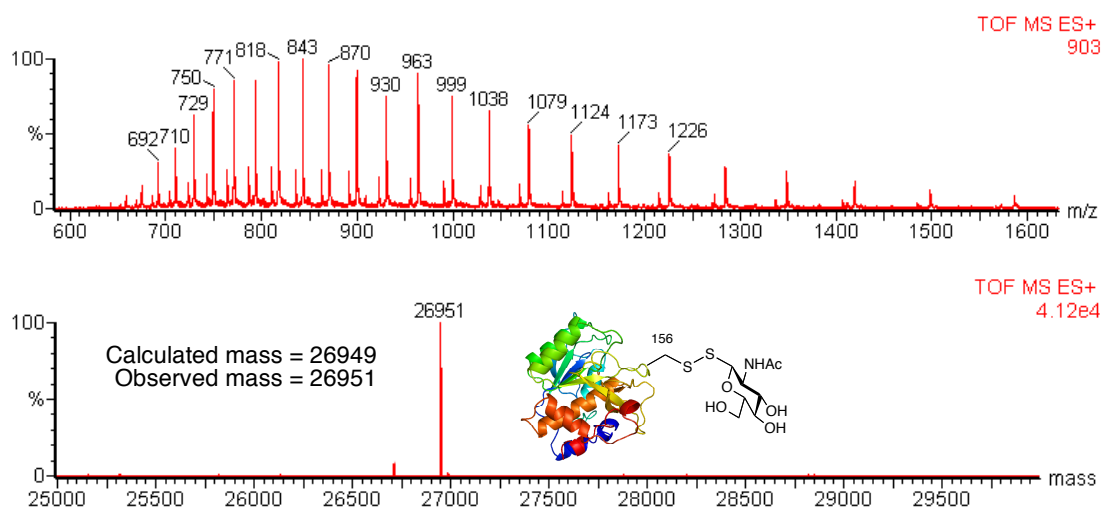
Synthesis of disulfide-linked glycoprotein **3.13**



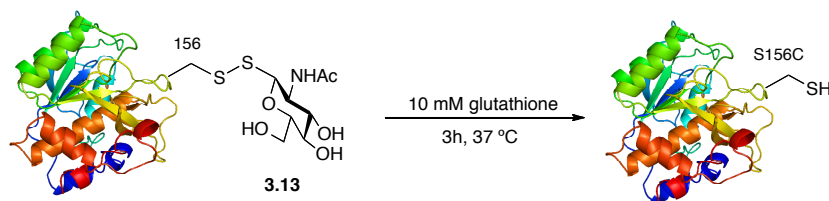
This experiment was carried out by Dr. Gonalo Bernardes. It is included here for reference.

SBL-S156C (3.0 mg, 0.11 μ mol) was added to a 1.5 mL plastic tube as a lyophilized solid and dissolved in 1.0 mL buffer comprised of 70 mM CHES, 5.0 mM MES, and 2.0 mM CaCl_2 (pH 9.5). GlcNAc thiol **2.37** (2.7 mg, 11 μ mol) was added as a solid to the protein solution along with 200 μ L acetonitrile. The reaction was vortexed to homogenize and rotated on a lab rotisserie for 1 hour at room temperature. LC-MS analysis of a 50 μ L aliquot of the reaction mixture revealed complete conversion to the corresponding

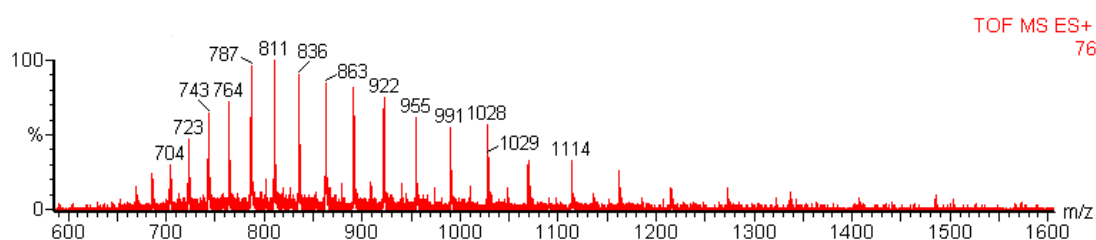
disulfide, **3.13** (calculated mass = 26949; observed mass = 26951). ESI-MS are shown below.

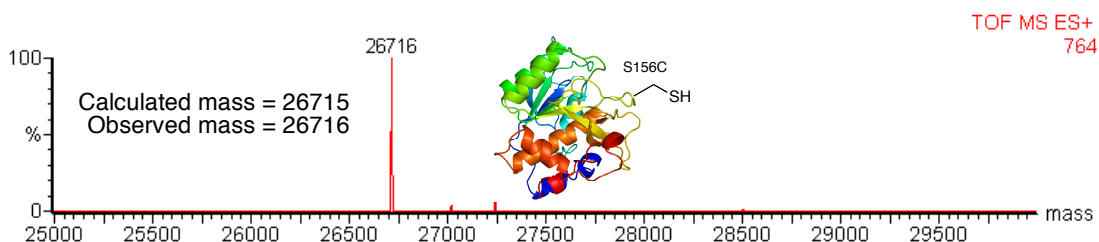


Reduction of disulfide-linked glycoprotein **3.13** with glutathione

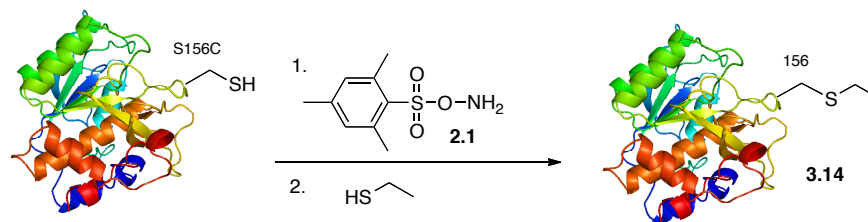


A 500 μ L solution of the disulfide-linked SBL-GlcNAc **3.13** was prepared at a concentration of 0.5 mg/mL in pH 8.0 sodium phosphate (50 mM) and added to a 1.5 mL plastic tube along with glutathione (1.5 mg, 5.0 μ mol). The final glutathione concentration was 10 mM. The solution was incubated for 3 hours at 37 °C. Complete reduction was observed by LC-MS (calculated mass = 26715, observed mass = 26716). ESI-MS are shown below.

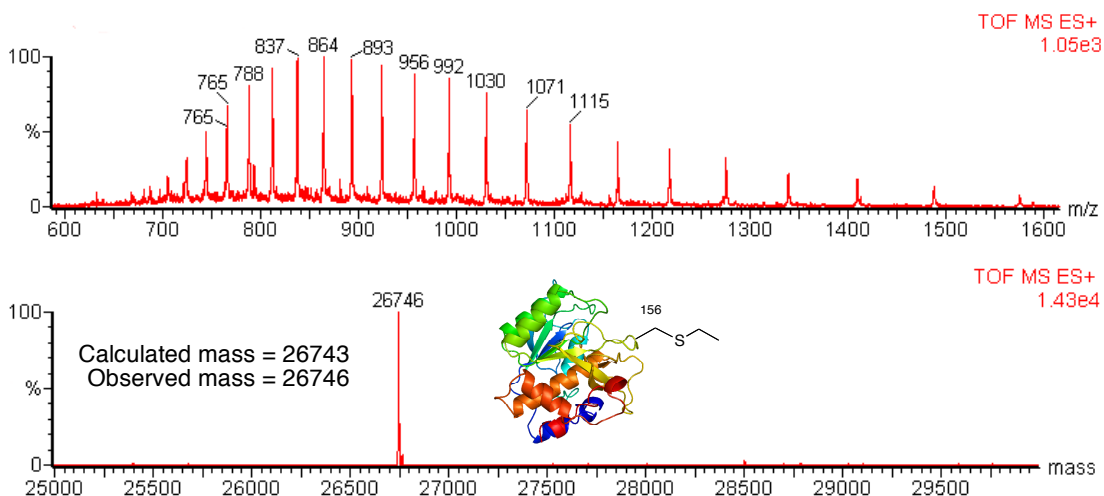


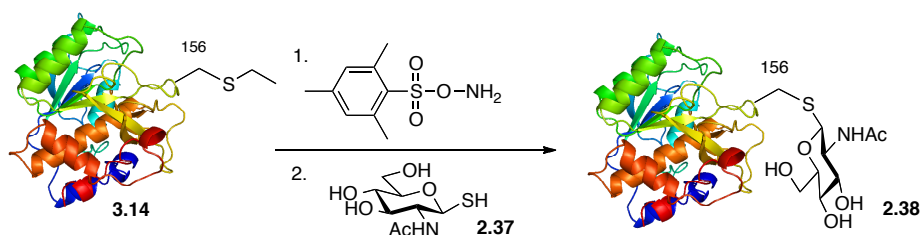


Synthesis of *S*-ethylcysteine on SBL at position 156 (**3.14**)



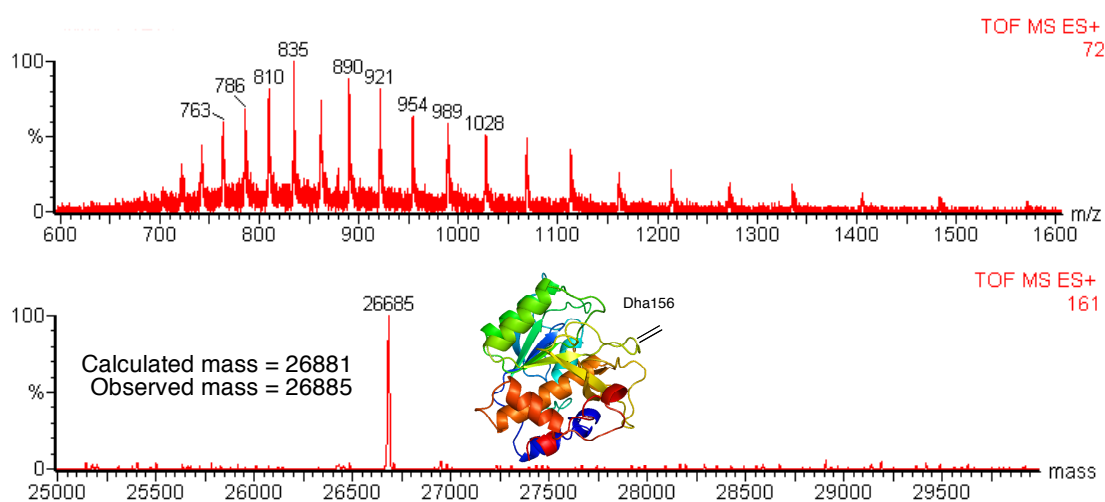
A fresh sample of SBL-156Dha was prepared as described above by reaction with MSH (**2.1**). SBL-156Dha (**2.35**) was used directly. Accordingly, 35 μL of ethane thiol was added to a 625 μL of a 1 mg/mL solution of SBL-156Dha (0.05 μmol) in 50 mM sodium phosphate (pH 8.0). The sample was vortexed to homogenize and then rotated for 30 minutes at room temperature. LC-MS analysis of the reaction mixture showed full conversion to the ethylthioether **3.14** (calculated mass = 26743; observed mass = 26746). The reaction mixture was passed through a PD10 column to remove the bulk of small molecules, eluting with pH 8.0 sodium phosphate (50 mM) and then further purified by dialysis ($2 \times 4\text{L}$ of the same buffer). Protein recovery of the purified protein was 0.5 mg (80%, BCA). The protein was concentrated to 0.36 mg/mL (vivaspin, 10K MWCO) and then split into aliquots and flash frozen. ESI-MS for the thioether are shown below.

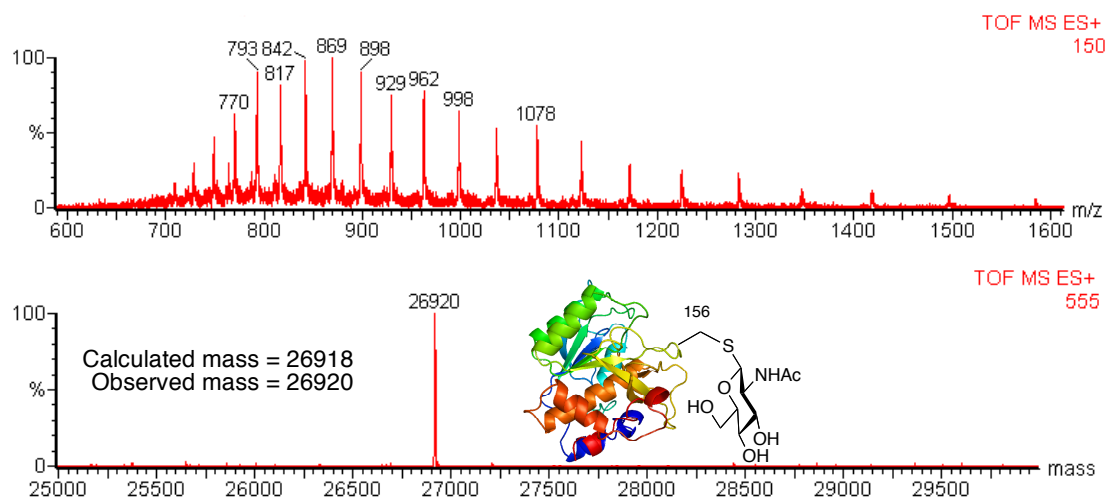




Regeneration of Dha from *S*-ethylcysteine (3.14) on SBL at position 156

A 500 μL sample of thioether **3.14** was thawed and kept on ice until needed. This sample was previously purified by dialysis to remove excess thiol (see above). An MSH solution (1.8 mg, 8 μmol) was prepared in DMF (720 μL) and 14 μL (0.16 μmol) was added to a 250 μL sample of SBL-156SEt. The reaction was vortexed to homogenize and then shaken for 20 minutes at 4 $^{\circ}\text{C}$. A 40 μL aliquot was taken for LC-MS analysis which showed full conversion to SBL-Dha **2.35** (calculated mass = 26681; observed mass = 26685). To verify that this material corresponds to the dehydroalanine containing protein, GlcNAc thiol **2.37** (4.0 mg, 17 μmol) was added as a solid to the reaction mixture. The solution was rotated at room temperature for 30 minutes. Full conversion to SBL-156GlcNAc cysteine **2.38** was observed (calculated mass = 26918; observed mass = 26920). The ESI-MS spectra for both the regenerated SBL-Dha and the subsequent addition of GlcNAcSH are shown below.

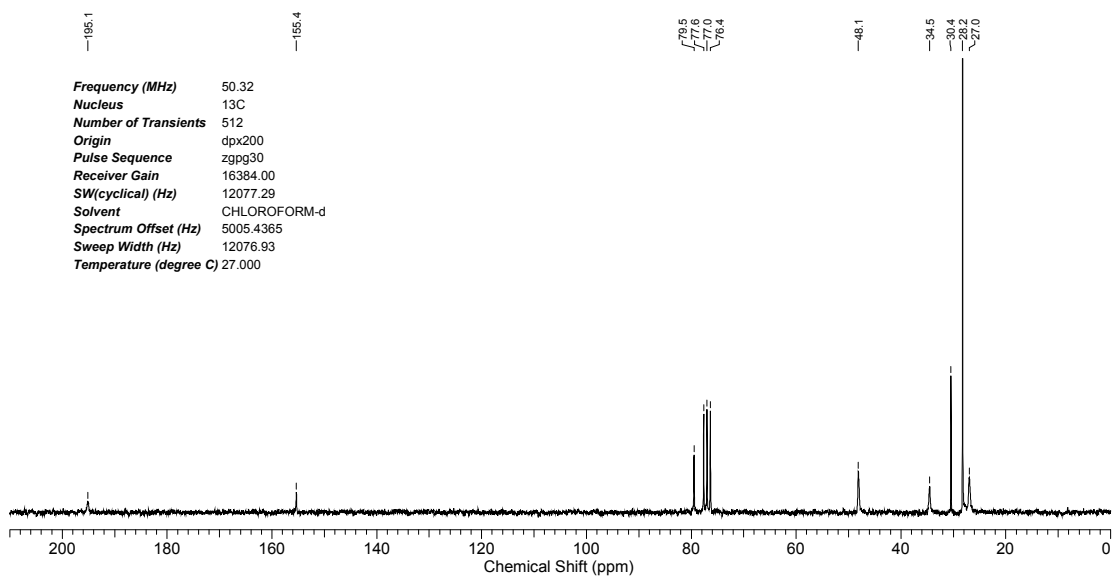
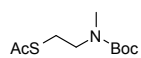
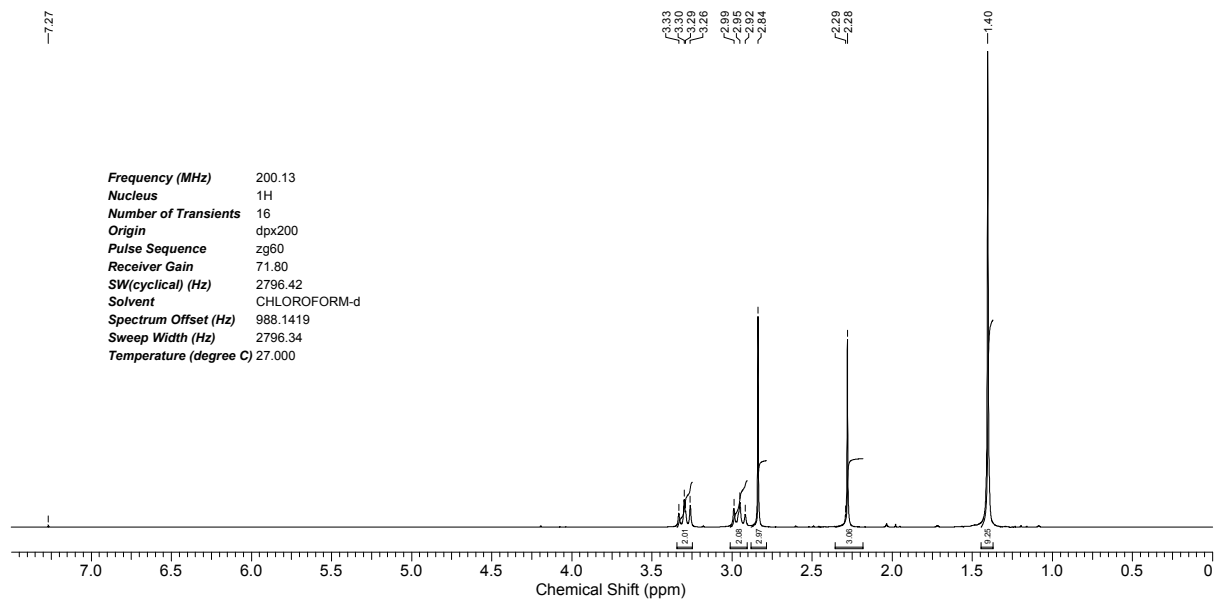
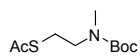


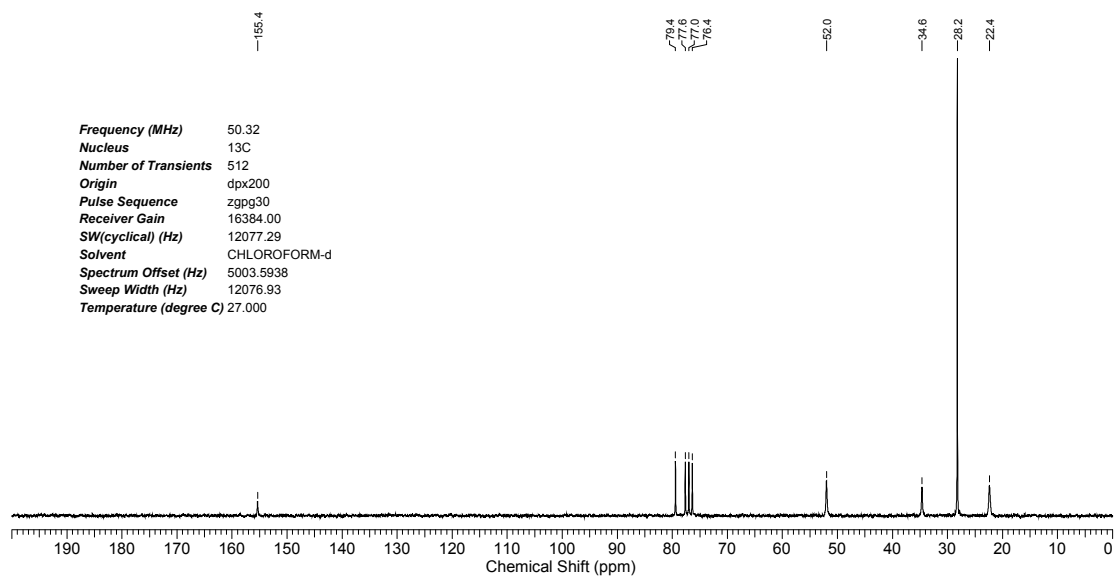
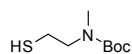
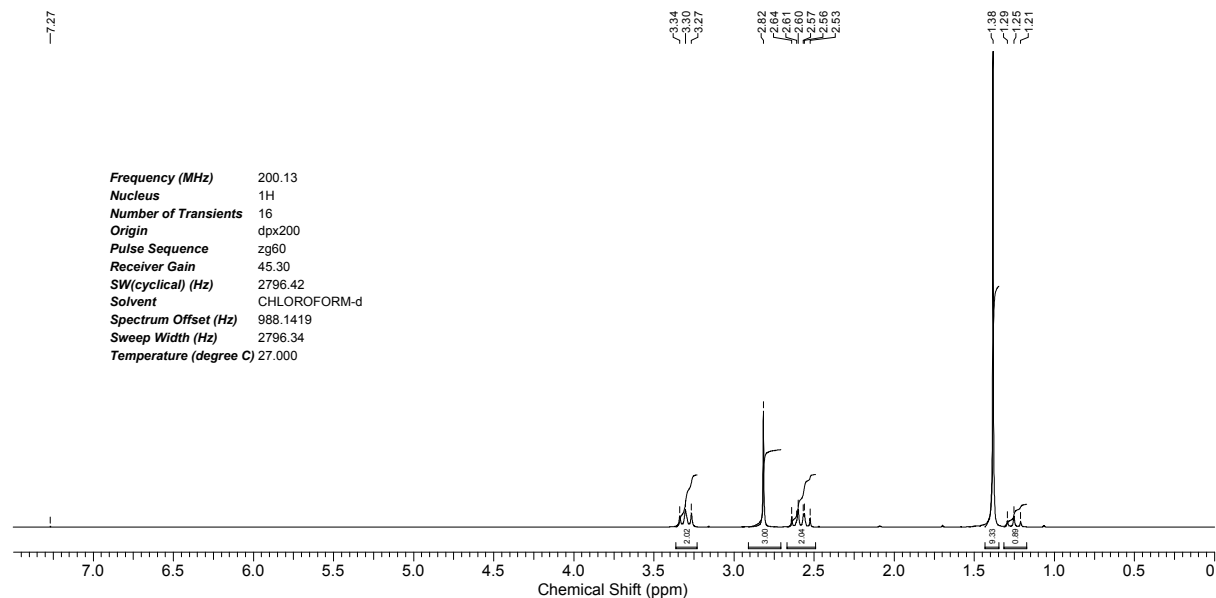
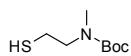


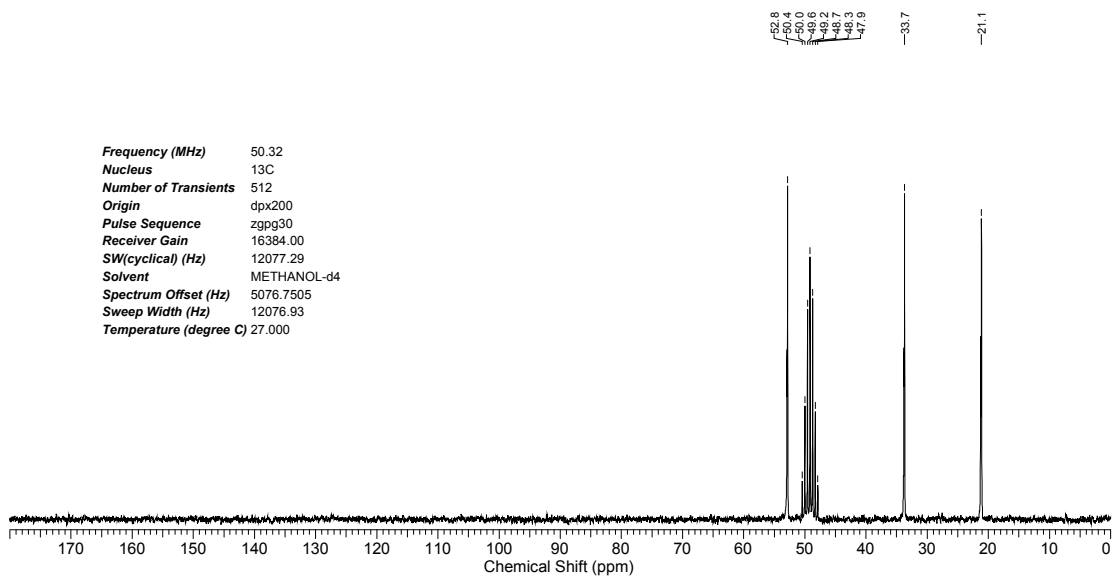
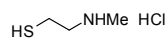
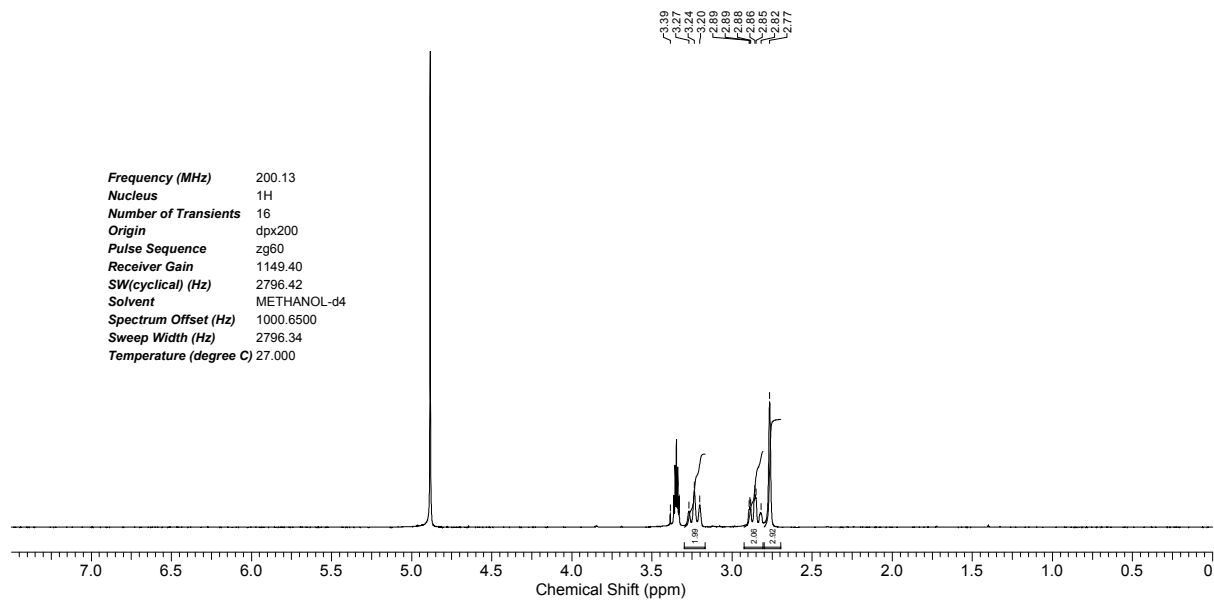
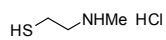
3.6 References for Experimental Procedures

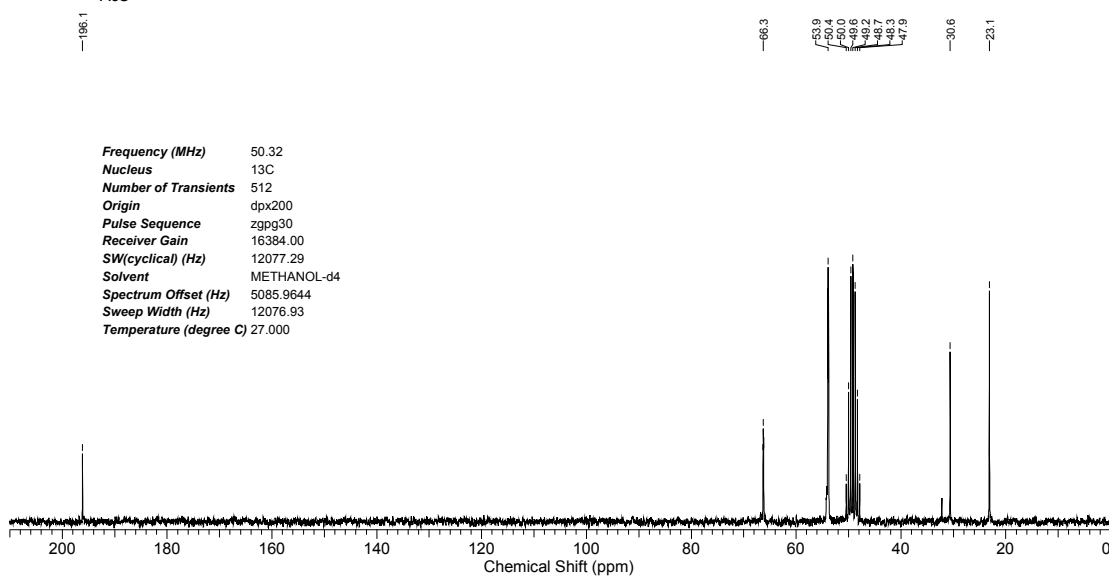
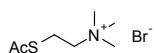
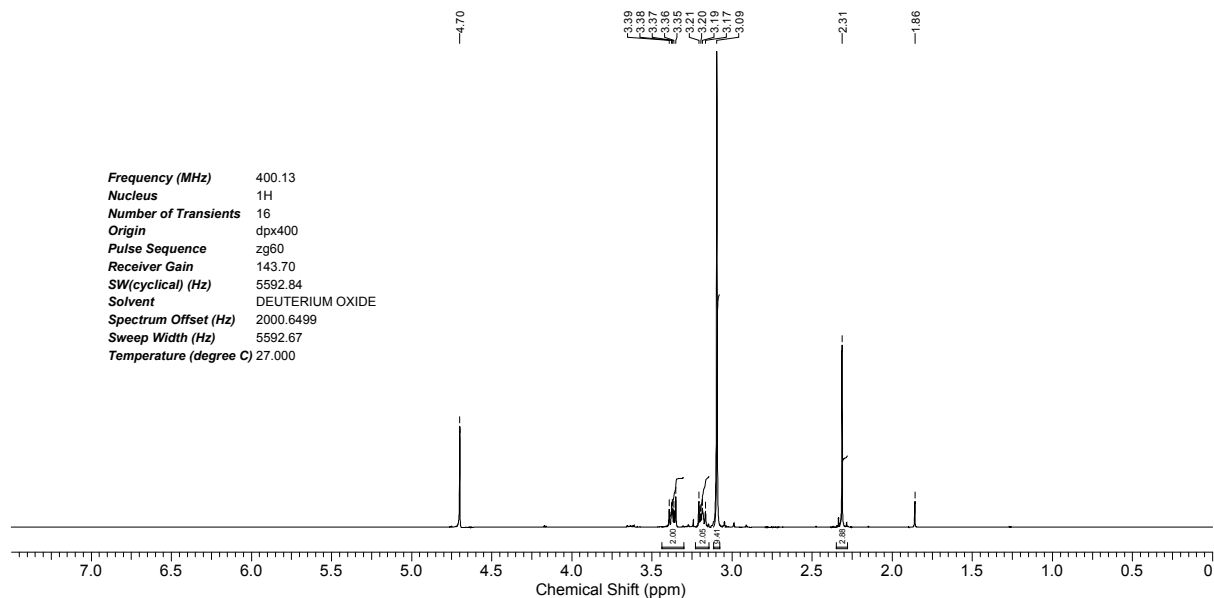
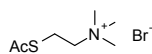
- (1) Johnson, C. R.; Kirchhoff, R. A.; Corkings, H. G. *J. Org. Chem.* **1974**, *39*, 2458-2459.
- (2) Harpp, D. N.; Smith, R. A.; Steliou, K. *J. Org. Chem.* **1981**, *46*, 2072-2079.
- (3) Treibs, A.; Neumayr, F. *Chem. Ber.* **1957**, *90*, 76-79.
- (4) Meyer, W.; Reckendorf, Z.; Bonner, W. A. *J. Org. Chem.* **1961**, *26*, 4596-4599.
- (5) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 2244-2247.
- (6) El Alaoui, A.; Schmidt, F.; Amessou, M.; Sarr, M.; Decaudin, D.; Florent, J.-C.; Johannes, L. *Angew. Chem. Int. Ed.* **2007**, *46*, 6469-6472.
- (7) Fernández-González, M.; Bourureira, O.; Bernardes, G.J.L.; Chalker, J.M.; Young, M.A.; Errey, J.C.; Davis, B.G. *Chem. Sci.* **2010**, *1*, 709-715.
- (8) Gamblin, D. P.; van Kasteren, S.; Bernardes, G. J. L.; Chalker, J. M.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. *Mol. Biosyst.* **2008**, *4*, 558-561.

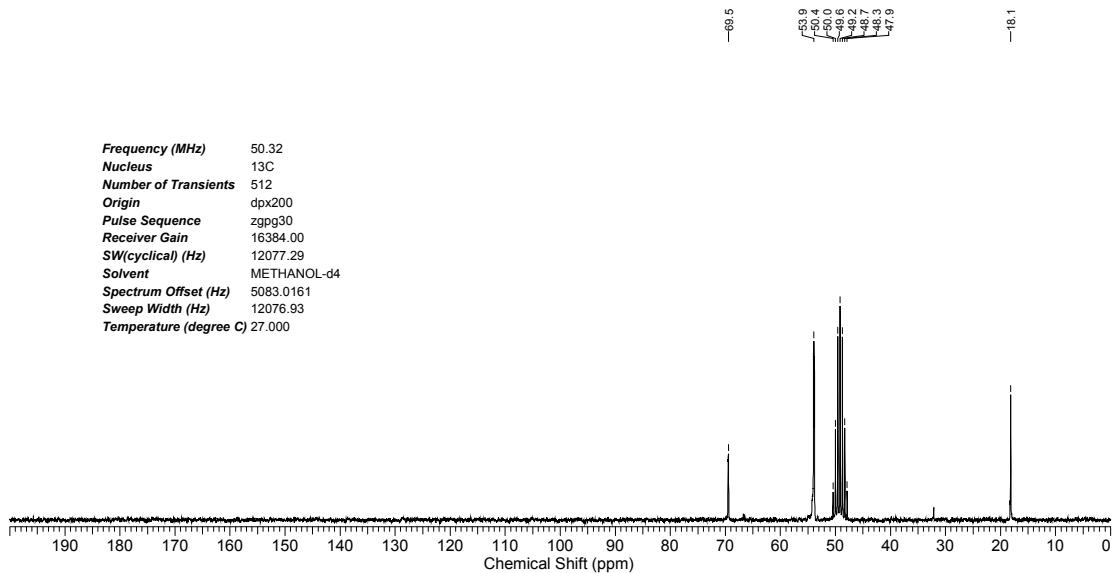
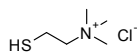
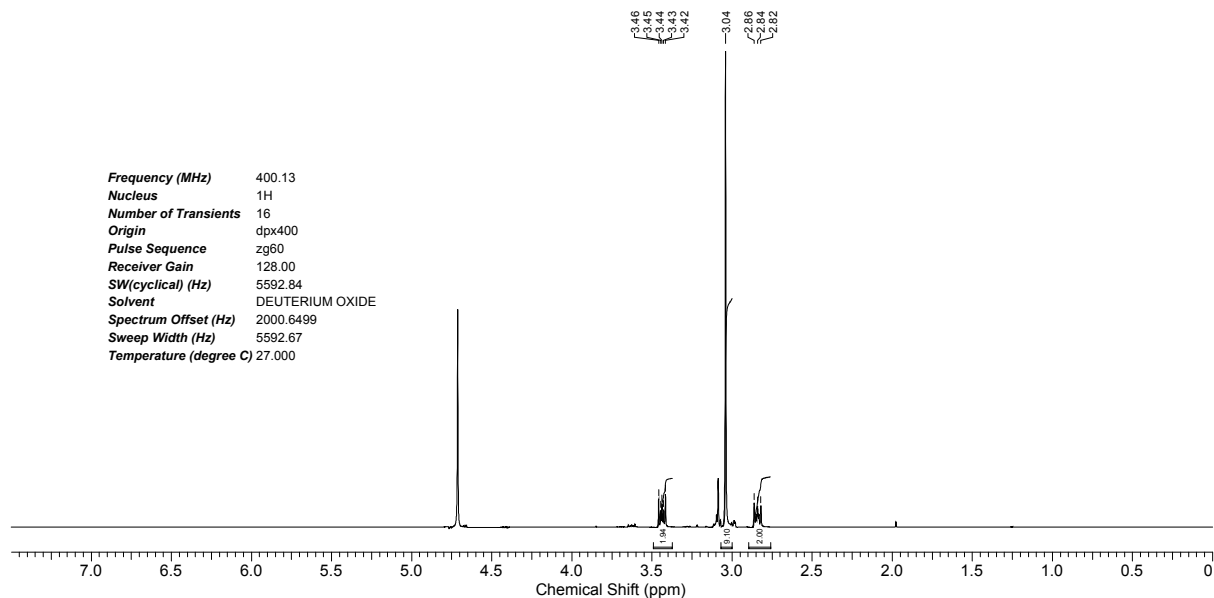
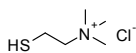
3.7 NMR Spectra





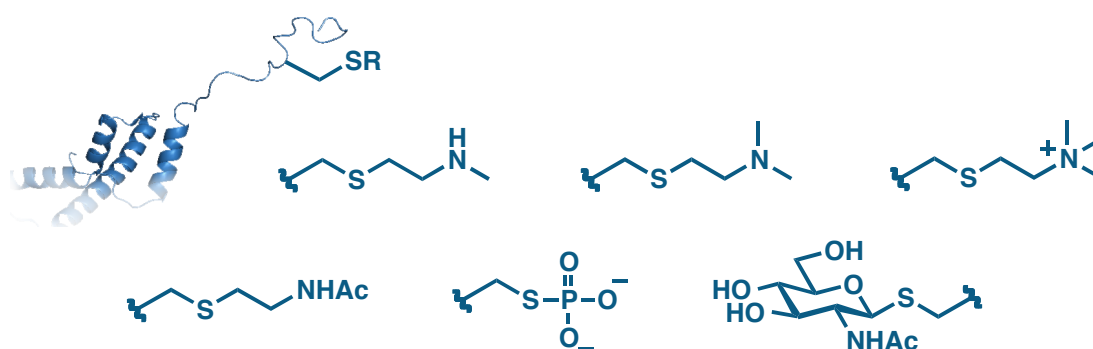






Chapter 4

The Conversion of Cysteine to Dehydroalanine Enables Access to Histones Bearing Post-Translational Modifications



“The histone code described here is by no means deciphered...”

Brian D. Strahl & C. Davis Allis

Nature, **2000**, 403, 41-45

4.1 Introduction

Chromatin structure and the histone code hypothesis

Eukaryotic cells face the staggering task of compacting the DNA of their entire genome in each cell nucleus. This task is further complicated by the necessity to package DNA in a way that allows tight regulation of transcription, replication and repair.^{1,2} The complex of DNA and proteins that make up this genetic package is called chromatin. Chromatin is made up of fundamental repeating units called nucleosomes.³⁻⁶ Each nucleosome is composed of an octamer of proteins (histones) and about 147 base pairs of DNA that wrap around this octamer.^{7,8} Two copies each of these highly conserved histones (H2A, H2B, H3, and H4) make up each octamer. Linker DNA and linker histone proteins string together each nucleosome to complete the primary structure of chromatin. A schematic representation of chromatin structure—from the chromosome to DNA level—is shown in Figure 4.1.

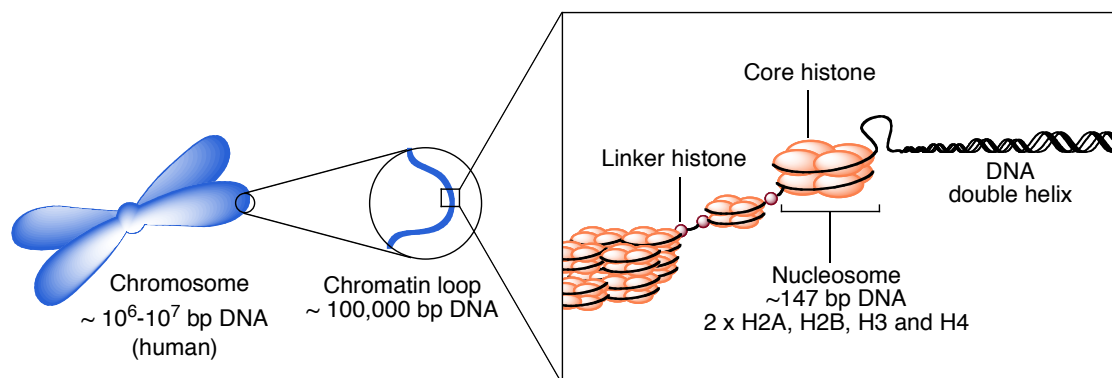


Figure 4.1: Schematic representation of chromatin

The structure and function of chromatin has been the subject of intense research for well over half a century. A fundamental problem in cellular biology is to derive a complete account of how the constituents of chromatin—DNA and histone proteins—interact to ensure high fidelity transcription and replication required for each cell. In 1950, Stedman and Stedman pointed out that such an account should explain how cells can differ so widely in physiology, yet contain the same genomic DNA.^{9,10} They hypothesized that the proteins associated with DNA, the histone proteins, might act as gene suppressors. This hypothesis was advanced further by Huang and Bonner who

demonstrated that histones bind to DNA *in vitro* and, in doing so, prevent RNA synthesis.¹¹ The physiological relevance of histone suppression of RNA synthesis was validated soon after by Allfrey, Littau, and Mirsky.¹² Despite clear evidence that histones bound to DNA and inhibited or somehow regulated RNA synthesis, it was not clear how so few and similar histone proteins could maintain the precise control required in the cell.¹³

Some clues for how histones might control gene expression came with the discovery that these proteins contained lysine side chains that were both *N*-acetylated^{14,15} and *N*-methylated.¹⁶ Upon investigation of these histone modifications, Allfrey, Faulkner, and Mirsky demonstrated that methylation and acetylation occur after translation of the histone proteins on a time scale comparable to RNA synthesis.¹⁷ Furthermore, it was shown that acetylated histones bind to DNA, but in a way that allowed RNA synthesis. This observation led the authors to propose that genetic control might be related to these post-translational modifications and that such modifications might be dynamic.¹⁷

Since these remarkable discoveries, the complex roles of histone acetylation¹⁸ and methylation¹⁹ have been intensely studied. Moreover, the ensuing decades revealed that histones are subjected to several other post-translational modifications, further complicating the structural and functional account of chromatin.^{5,20,21} For example, histone serine and threonine residues are subject to phosphorylation;²²⁻²⁶ lysine can be modified by biotinylation,²⁷⁻²⁹ ubiquitination,³⁰⁻³⁵ and sumoylation;³⁶ arginine can be methylated^{37,38} or deiminated;³⁹⁻⁴² histones can bear ADP-ribose;⁴³ and even just recently, serine and threonine glycosylation was found on histones.^{44,45} The structures and known sites for a selection of these modifications on the core histones are shown in Figure 4.2 and Figure 4.3, respectively. It should be noted that the modifications and their locations in Figures 4.2 and 4.3 are not exhaustive and many modifications remain to be fully mapped.^{36,45} Furthermore, Figure 4.3 only depicts the core histones; the linker histones (H1) are also modified, though less densely.⁴³

With such a large variety of post-translational modifications on histone proteins, Stedman and Stedman's original hypothesis that histones are gene suppressors^{9,10} becomes even more complicated to account for mechanistically. Moreover, a growing body of evidence suggests that histones not only influence transcription by contact with DNA, but also by interacting with several chromatin modifying enzymes.⁴⁶ In particular,

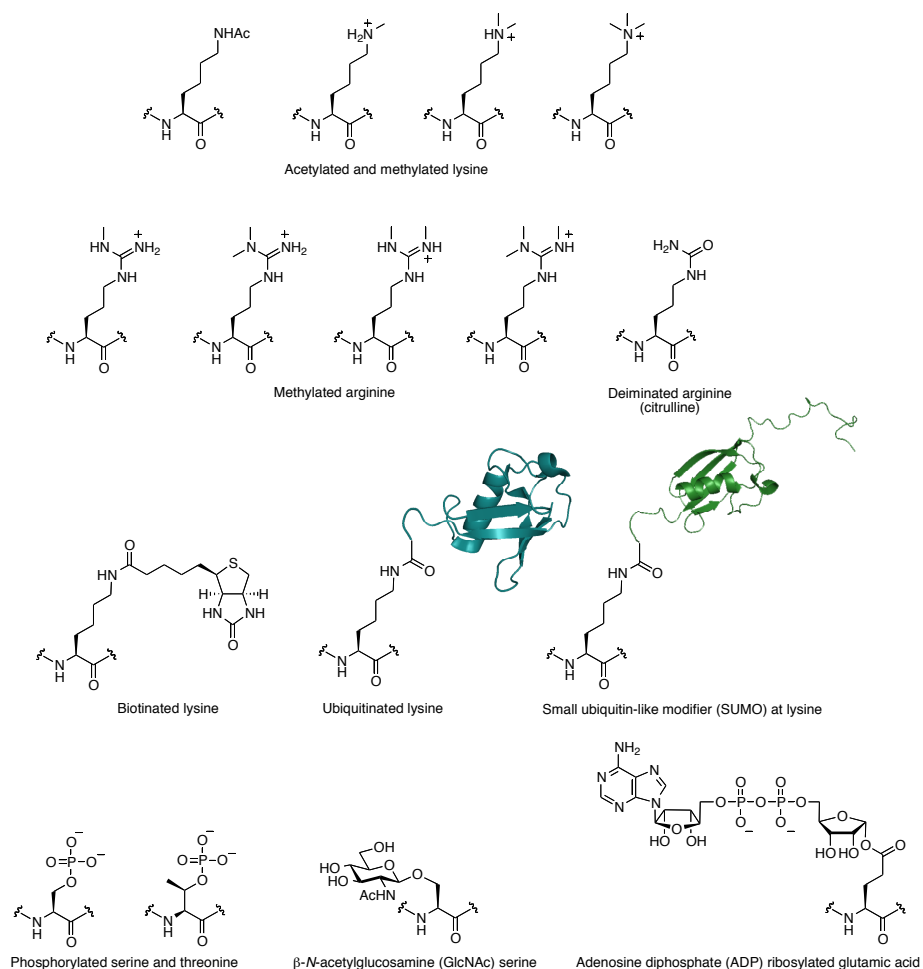


Figure 4.2: A selection of histone post-translational modifications.

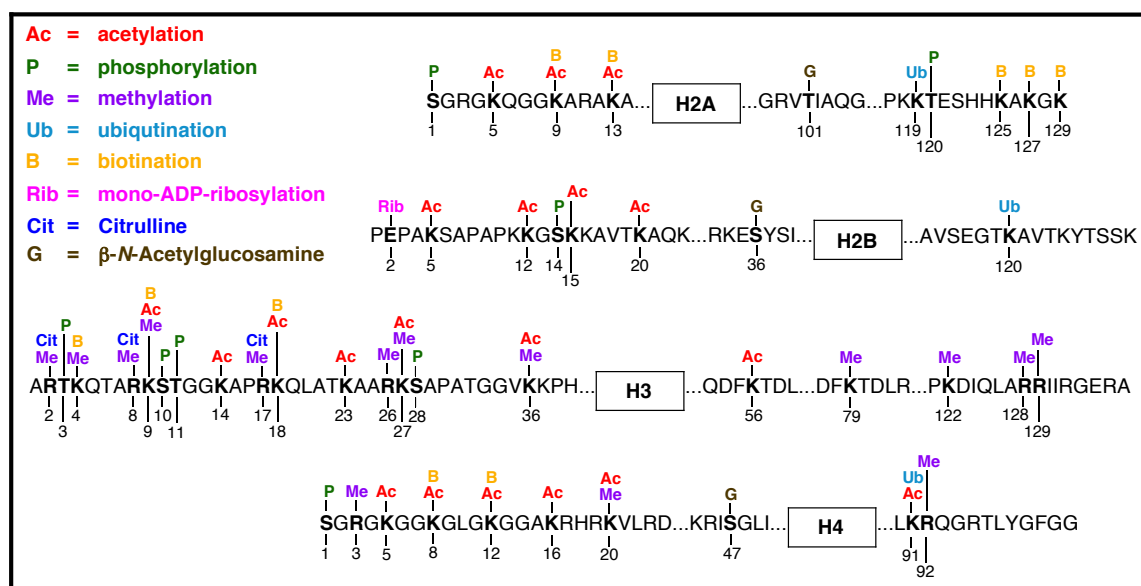


Figure 4.3: Location of a selection of post-translational modifications for the core histones^{43,45,47}

the *N*-terminal tail of the core histones (residues 1-30 of H2, H3, and H4) protrude from the nucleosome and interact with other nuclear proteins. These residues are densely modified (cf. Figure 4.3), often in distinct patterns that correlate with a specific function or phenotype. These relatively recent discoveries led Strahl and Allis to propose that the post-translational modification of histone tails might code for the processes required in chromatin maintenance, regulation, and general function. Their proposal is referred to as the “histone code hypothesis.”^{46,48}

Evaluating the histone code hypothesis

The histone code hypothesis put forth by Strahl and Allis is an intriguing proposal for the function of histone post-translational modifications. Deciphering this code, however, is no easy task.^{46,48} One necessary step is to map all histone PTMs—the vocabulary of the histone language. Figure 4.3 shows the progress to date, but a complete PTM map is essential. The evaluation of the histone code itself is riddled with complications. For example, there is not necessarily a one-to-one correspondence for a given PTM and its functional outcome; the outcome may be influenced by other histone PTMs. For example, lysine acetylation at a given site may be associated with transcription, but only when a serine at another site is phosphorylated.⁴⁶ The combinatorial nature of the histone code and the many types of histone modifications complicate efforts to translate the histone language.

The histone code hypothesis prompts many questions in chromatin biology.^{46,48,49} A few of these unresolved problems are highlighted here. First, it is not clear how the histone code is maintained. The nature of binding and catalysis of the enzymes that install and remove histone PTMs will help resolve this issue.²¹ Second, the histone code is read by “effector” proteins that bind to PTMs.⁵⁰ These proteins, also called readers, may facilitate further downstream events, stabilize relevant chromatin structure, and recruit various other binding modules and chromatin modifying enzymes. A full account of code readers is far from complete. Third, the histone PTMs themselves may influence higher order chromatin structure. In doing so, the histone PTMs may not only code for interaction with other proteins, but code for the folding of chromatin itself. Studying this relationship between histone PTMs and chromatin folding and higher order structure is yet another line of investigation prompted by the histone code hypothesis.⁴⁹ Finally, it is not clear how the histone code is established or maintained during cell replication. A mechanistic account of histone PTM inheritance is far from complete.⁴⁹

An understanding of the histone code has far reaching implications. Chromatin regulation is central to the fate of cells and is therefore a fundamental problem in cellular biology. At the organism level, an understanding of the histone code may help elucidate the molecular basis of cancer, aging, and other aspects of human health.^{46,48} The far reaching impact of the histone code has understandably drawn the attention of biologists and chemists alike and there is no doubt that evaluation of the histone code hypothesis will require the full resources of both chemistry and biology.⁴⁹ One distinctively chemical contribution is the synthesis and supply of histones bearing post-translational modifications. In order to study the influence of histone PTMs on chromatin structure and chromatin modifying enzymes, these modified histones must be readily accessible in homogenous form. In this Chapter, dehydroalanine is used to install PTMs on histones. Before discussing this strategy, it is necessary to review other approaches used to access histones bearing post-translational modifications. This survey will help highlight the scope and limitations of each method so that all can be used in a reinforcing, complementary manner in ongoing efforts to decipher the histone code.

Synthesis of histones bearing post-translational modifications

The synthesis of full-length histones bearing post-translational modifications will help determine their structural and function consequences in chromatin biology. While synthetic peptides can be used to assay chromatin-modifying enzymes and binding modules,^{47,51} only the full-length histones can be used to investigate the structure and function of chromatin in its physiologically relevant form.⁴⁷ Several approaches exist for the synthesis of modified histones and three general strategies will be outlined below: histone synthesis using native chemical ligation or expressed protein ligation, site-selective chemical modification of histone mutants, and recombinant translation using an expanded genetic code. While these strategies differ conceptually, the reinforcing use of chemical and biological techniques is a common feature in all approaches.

Histone synthesis by native chemical ligation and expressed protein ligation

As seen in the introduction to Chapter 3, native chemical ligation (NCL)⁵² and expressed protein ligation (EPL)^{53,54} have been used extensively in the synthesis of post-translationally modified proteins.⁵⁵⁻⁵⁷ The synthesis of modified histones is no exception.^{47,49,58} In fact, NCL was one of the first strategies used to evaluate the histone code hypothesis.⁵⁹ McCafferty, for example, used native chemical ligation to prepare

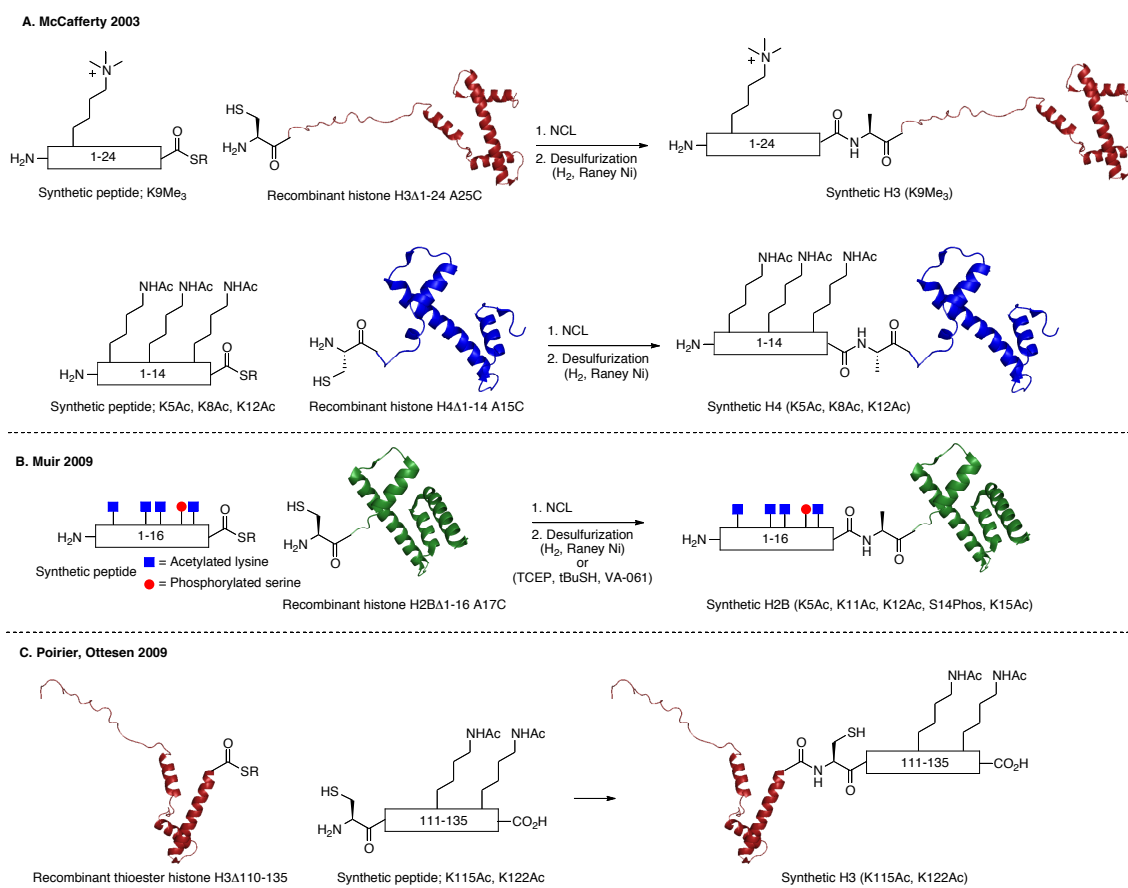


Figure 4.4: Representative syntheses of modified histones by NCL and EPL

acetylated and methylated histones (Figure 4.4A). Since there are few cysteine residues in histones, post-ligation desulfurization^{60,61} was used to provide a formal ligation at alanine. These modified histones were used to prepare chromatin assemblies and study the action of chromatin modifying enzymes. A similar NCL strategy has been adopted by other researchers to investigate acetylated histones and their effect on chromatin folding and remodeling.^{62,63}

Native chemical ligation has also been used in the synthesis of phosphorylated histones. Peterson and co-workers used NCL to prepare histone H3 with phosphorylation at serine 10.⁶⁴ This synthetic protein was used in nucleosomal arrays to study enzymatic modifications of the nucleosome. Among these assays, it was found that in certain cases histone acetyltransferase activity differed from a similar assay on a truncated peptide substrate. This observation highlights the importance of studying chromatin-modifying enzymes on full-length histones, a more physiologically relevant substrate than peptide fragments.

Evaluation of the histone code hypothesis necessitates access to histones bearing not only multiple modifications, but different types of modifications on the same protein. The Muir lab has demonstrated that NCL is well-suited for such synthesis by the preparation of an H2B variant carrying both acetyllysine and phosphorylated serine (Figure 4.4B).⁶⁵ Since H2B does not have cysteine in its native sequence, post-ligation desulfurization was required. Danishefsky's radical induced desulfurization⁶¹ was particularly effective and compatible with both acetyllysine and phosphorylated serine.⁶⁵ Further investigation of this modified protein highlighted the caution required in antibody-based assays of post-translational modifications: an antibody specific to the phosphorylated serine was blocked by the nearby acetyllysine residues. The Muir lab has since prepared other phosphorylated histones by NCL, including histones bearing stable analogs of phosphorylated histidine—a fragile and poorly understood PTM.⁶⁶

In the histone syntheses discussed so far, the PTMs were located on the *N*-terminal histone tails. Accordingly, a synthetic peptide bearing PTMs and the requisite thioester at the *C*-terminus of that peptide was ligated to a truncated histone with an *N*-terminal cysteine residue (Figure 4.4 A-B). For the synthesis of histones with PTMs near the *C*-terminus of the protein, expressed protein ligation is particularly useful. For example, Poirier, Ottesen, and co-workers used EPL to prepare H3 with acetylation at lysines 115 and 122 (Figure 4.4C).⁶⁷ It was found that this acetylated histone reduced the free energy of histone octamer binding and increased the rate on nucleosome repositioning on DNA. In this report, the authors also compared acetyllysine to glutamine at the same residues (a lysine to glutamine mutation is often used to mimic lysine acetylation) and found that while glutamine mimics the neutralization of charge that accompanies lysine acetylation, glutamine is a poor steric mimic for lysine acetylation.⁶⁷ This discovery once again highlights the need to assay histones and higher order chromatin with the natural modification.

Histone ubiquitination is also accessible using EPL. In 2008, the Muir laboratory reported a three component approach in the synthesis of ubiquitinated H2B.⁶⁸ The synthetic strategy is summarized in Figure 4.5A with retrosynthetic disconnections denoted in blue. A synthetic peptide containing residues 117-125 of H2B (A117C) with an isopeptide bond between the side chain of lysine 120 of H2B and glycine 76 of ubiquitin was first prepared. This synthetic peptide was then ligated to a recombinant thioester of ubiquitin Δ76. Glycine 76 of the synthetic peptide was derivitized with an

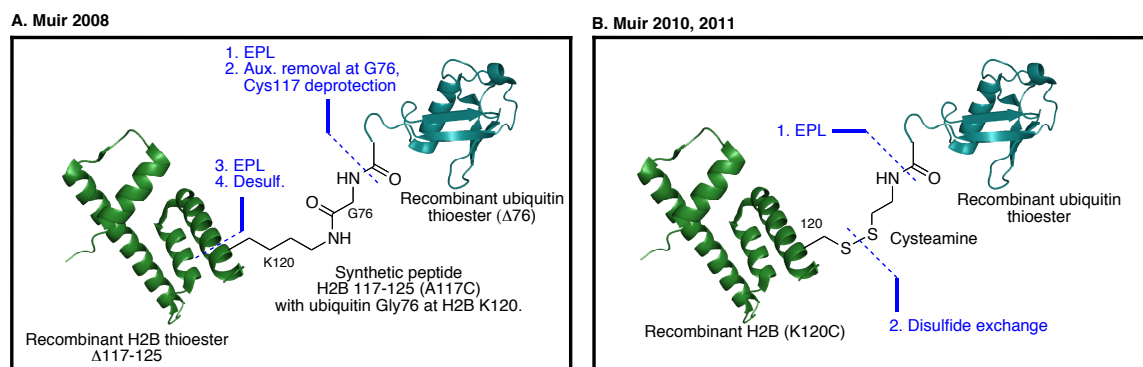


Figure 4.5: Syntheses of ubiquitinated histone H2B

auxiliary to mediate the ligation. Removal of this auxiliary after ligation and concomitant deprotection of cysteine 117 set the stage for the final ligation with the H2B thioester ($\Delta 117-125$). A final desulfurization converted cysteine 117 to the native alanine residue of H2B. The final construct (Figure 4.5A) is the natural constitution of H2B that has been ubiquitinated at lysine 120. This ubiquitinated H2B construct was incorporated into a nucleosome to investigate its influence on methylation by the methyltransferase hDot1L. This “cross-talk” study (the investigation of how one histone modification may modulate the rate and level of other histone modifications), was carried out in an intranucleosomal and internucleosomal manner. In the latter case, the synthesis of a dinucleosome was demonstrated, a significant accomplishment in the context of designer chromatin.^{49,68}

A related analog of ubiquitinated H2B was prepared soon after by the same laboratory.⁶⁹ The synthesis of this disulfide-linked ubiquitin (Figure 4.5B) overcame many of the technical challenges associated with multiple ligations required in the synthesis of the native structure, allowing access to larger quantities of ubiquitinated H2B. In nucleosomal assays, the H2-ubiquitin analog behaved similarly to the native construct. Furthermore, the ease of synthesis also enabled the preparation of several H2B constructs that differed in the site of ubiquitination. Comparison of these constructs revealed that stimulation of the methyltransferase activity of hDot1L was not restricted to a single site of ubiquitination.⁶⁹ Further studies using the disulfide-linked H2B-ubiquitin revealed its inhibitory effect on higher-order chromatin structure.⁷⁰

The reinforcing use of EPL and site-selective chemical modification is evident in the synthesis of ubiquitinated H2B. Other post-translational modifications are also accessible through chemical modification of histones, as discussed in the next section.

Site-selective chemical modification of histones

Chemical modification of histones was one of the earliest approaches to understand the effects of post-translational modifications. In Mirsky's landmark publication in 1964 on histone methylation and acetylation, histones were acetylated by careful treatment with acetic anhydride.¹⁷ The chemically acetylated histones, in contrast to the native histones, did not inhibit RNA synthesis and provided one line of evidence suggesting this PTM was involved in genetic regulation. To evaluate the histone code hypothesis, more precise modifications are required. Cysteine modification is the primary means to chemically install PTMs. Indeed, disulfide formation was described in the previous section for the preparation of ubiquitinated H2B (Figure 4.5B).⁶⁹ Other modifications—in particular lysine methylation and acetylation—can be mimicked by cysteine aminoethylation.^{71,72}

As discussed in Chapter 1, cysteine aminoethylation has been used for more than half a century to convert cysteine into a functional lysine analog.⁷³ It was a natural extension of this chemistry to access methylated lysine analogs. Shokat and co-workers demonstrated the ease and utility of this transformation in the synthesis of methylated histones (Figure 4.6A).⁷¹ These methylated lysine analogs are recognized by antibodies raised to the natural residue,⁷¹ validating their use as functional mimics. Fortuitously, histones contain very few cysteines so site-directed mutagenesis can be used to install a single cysteine at the desired site of modification. This strategy has been widely adopted for the study of lysine methylation in the context of chromatin remodeling,⁷¹ nucleosome structure,⁷⁴ crystallographic analysis of chromatin binding proteins,⁷⁵ epigenetic inheritance,^{76,77} and cross-talk with other histone modifications.^{78,79}

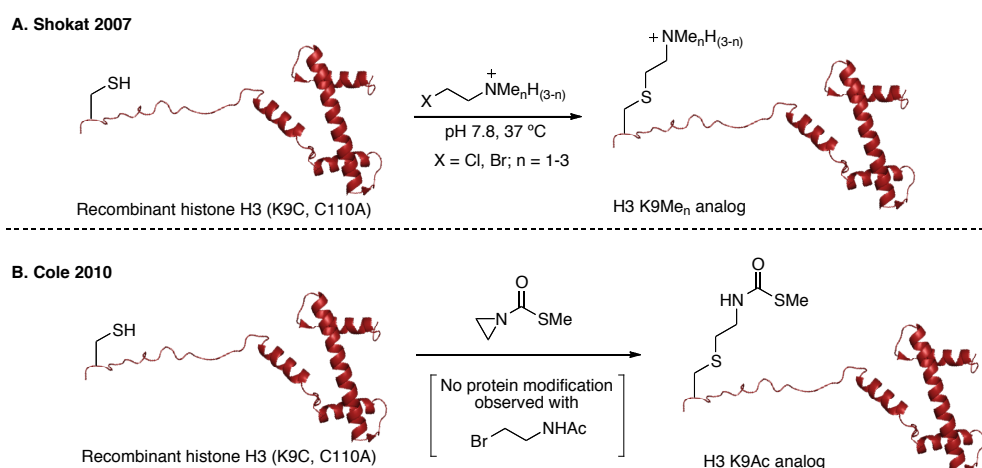


Figure 4.6: Representative examples of site-selective chemical modification of histones

The analogous transformation of cysteine to an acetyllysine mimic requires a slight modification (Figure 4.6B).⁷² The Cole laboratory has introduced a thiocarbamate analog of acetyllysine, accessible again by alkylation of cysteine. This acetyllysine mimic is recognized by antibodies raised to the natural modification, but it is not hydrolyzed by histone deacetylase enzymes. This resistance to enzymatic hydrolysis may prove useful in cellular assays that benefit from the analog's stability.⁷²

The hallmark of cysteine modification to access methylated and acetylated lysine mimics is the operational simplicity of the procedure. A single chemical step can be used on a recombinant histone and purification can be as simple as dialysis. However, cysteine modification has only been used to install a single type of modification for any given protein since regioselective cysteine modification is rare⁸⁰⁻⁸² and has not been reported for histones. There is also a need to devise transformations that allow the conversion of cysteine to mimics of other modifications such as phosphorylation and glycosylation. In this Chapter, efforts to resolve these shortcomings will be addressed. Before these experimental pursuits, methods for the synthesis of modified histones using translation systems will be discussed. This assessment will further illustrate the complementary nature of chemical and biological methods for the synthesis of modified histones.

Accessing modified histones using translation systems

As discussed in Chapter 1, various translation systems have been developed for the incorporation of non-canonical amino acids into proteins.⁸³⁻⁸⁷ It is perhaps not surprising that such strategies have been applied to the biosynthesis of modified histones. For instance, as discussed in Chapter 3, the Schultz group have pioneered a technique for amber codon suppression where the stop codon is reassigned to a non-canonical amino acid using evolved tyrosyl tRNA / tRNA synthetase pairs.⁸⁵ Using this technique, phenylselenocysteine can be incorporated at a genetically defined site.⁸⁸ Oxidative elimination to dehydroalanine and subsequent addition of various thiols provide methylated and acetylated lysine mimics (Figure 4.7A). This technology was applied to histone H3, but the oxidative elimination step also led to unwanted methionine oxidation. A methionine free H3 mutant was therefore used in this investigation.⁸⁹ The acetylated lysine mimic obtained by this method was used to assay histone deacetylase activity.⁸⁹

Related amber codon suppression strategies have since been reported for the incorporation of the *natural* acetylated or methylated lysine. These systems rely on

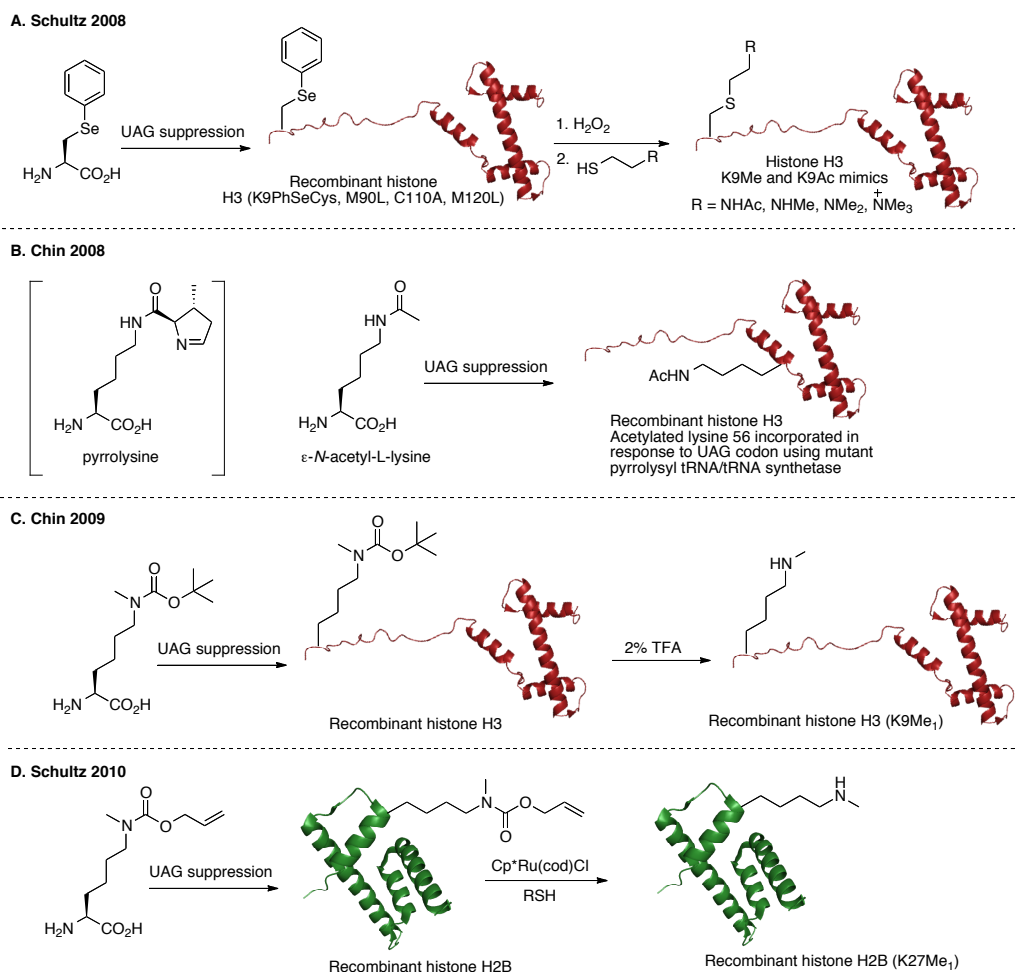


Figure 4.7: Representative examples of amber codon suppression in the synthesis of histones bearing post-translational modifications.

mutant pyrrolysyl tRNA / tRNA synthetase pairs evolved to charge various lysine derivatives in response to the amber stop codon (Figure 4.7 B-D). Acetyllysine can be incorporated directly by this method,⁹⁰ allowing the biosynthesis of several histones containing acetylated lysine (Figure 4.7B).⁹¹ This technique was used to investigate the effects of acetylation at lysine 56 in histone H3 on DNA binding and chromatin packing.⁹¹ The authors noted that incorporation of histone modifications during translation is particularly useful for modifications near the middle of the protein sequence. Synthesis of such proteins by native chemical ligation, while certainly possible, will nonetheless be more technically demanding since multiple ligations will likely be required.

Unlike acetylated lysine, methylated lysine residues cannot yet be incorporated directly by amber codon suppression. However, monomethyllysine has been incorporated indirectly in protected form (Figure 4.7 C-D).⁹²⁻⁹⁴ These latent methylated lysine residues

can then be deprotected chemically to provide the final histone product.^{92,93} Related efforts to incorporate dimethyllysine have been reported, but require several chemical steps to reach the final product.⁹⁵ There is no method yet reported for incorporation of trimethyllysine into full-length histones by translation.

Dehydroalanine as a precursor to post-translationally modified histones

The objective in this Chapter is to explore the use of dehydroalanine in the synthesis of histones bearing post-translational modifications. In particular, the chemistry described in Chapters 2 and 3 was investigated on histone H3. Histone H3 is densely modified, particularly in the *N*-terminal tail region (cf. Figure 4.3),⁹⁶ and the synthesis and evaluation of modified H3 is required to determine its role in the histone code.⁴⁶⁻⁴⁹ In these studies, cysteine was used as a precursor to dehydroalanine. H3 contains a single cysteine at position 110 which can be mutated to alanine without any discernable effects in chromatin structure,⁷¹ allowing genetic placement of the cysteine at the desired site of modification.

In the introduction to this Chapter, it is clear that there are many approaches to the synthesis of modified histones. The chemistry pursued in this Chapter was designed to complement these techniques, especially in situations where the results reported here might confer some advantage, technically or otherwise. For example, the Schultz group has used dehydroalanine to access modified histones.⁸⁹ However, their method for installing dehydroalanine oxidizes native methionine residues; perhaps the chemistry developed in Chapters 2 and 3 can overcome this limitation. Moreover, dehydroalanine might be a precursor to modifications not yet accessible by other methods of site-selective modification. Phosphorylation and glycosylation are two such modifications explored in this Chapter. NCL and EPL can undoubtedly provide these modified histones, but the chemistry described in this Chapter can be used in conjunction with NCL post-ligation or as an operationally simple alternative when only a single modification is required. The relative advantages and disadvantages of alternative methods of histone modification will be discussed throughout this chapter.

Finally, for the post-translational modification mimics accessible through dehydroalanine, it is necessary to validate their use as functional mimics. Binding studies and enzymatic assays are two approaches described in this Chapter to demonstrate how these modified histones might be used to study histone PTMs.

4.2 Results and Discussion

In Chapters 2 and 3, the conversion of cysteine to dehydroalanine was accomplished using oxidative elimination with *O*-mesitylsulfonylhydroxylamine (MSH, **2.1**) or through a bis-alkylation-elimination pathway using di-bromide **2.32**. Both methods were explored on the single cysteine histone mutant H3 (K9C, C110A) (**4.1**) for initial evaluations.[†] Position 9 of H3 is a lysine residue that is subject to methylation and acetylation—two modifications that can be mimicked by the addition of a suitable thiol to dehydroalanine. The conversion of cysteine to dehydroalanine at this residue is therefore an opportunity to test the applicability of reagents **2.1** and **2.32** in histone modification (Figure 4.8)

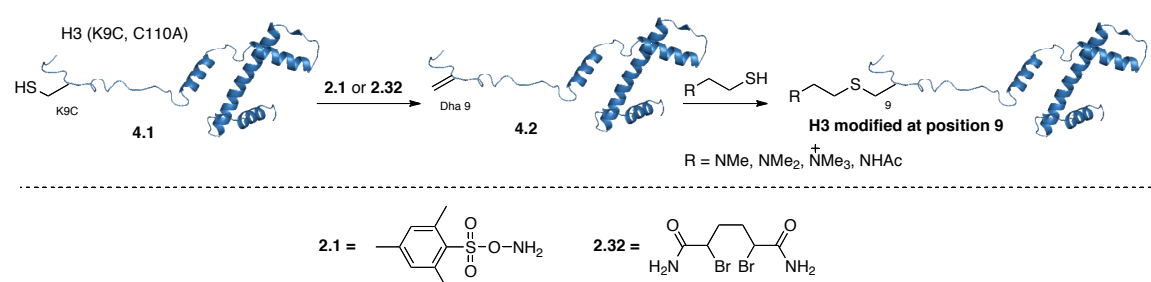


Figure 4.8: Proposed modifications of H3 at position 9.

Histone modification using *O*-mesitylsulfonylhydroxylamine (MSH, **2.1**)

In Chapter 2, several selectivity issues were discovered using MSH in protein modification. MSH was observed to react with nucleophilic residues and histones might be challenging substrates for this reagent. Indeed, a first attempt with MSH at pH 9 demonstrated that non-selective amination was again a problem (Figure 4.9A). Adducts of 16 Da were observed in the ESI-MS, consistent with the non-selective amination observed for other proteins in Chapter 2.

Reasoning that non-selective amination could be minimized by simply lowering the pH, the reaction was repeated at pH 7.0. Amination was still observed (there was a mass increase, rather than the desired elimination which would result in a loss of 34 mass units), but the reaction appeared much more selective. One of several possible products is proposed in Figure 4.9B. Amination of methionine is proposed since the rate of amination of the thioether should be relatively insensitive to pH, while other residues such as lysine

[†] All histone mutants used in this Chapter were expressed and purified by Dr. Nathan Rose of the Prof. Christopher Schofield research group in the University of Oxford Department of Chemistry.

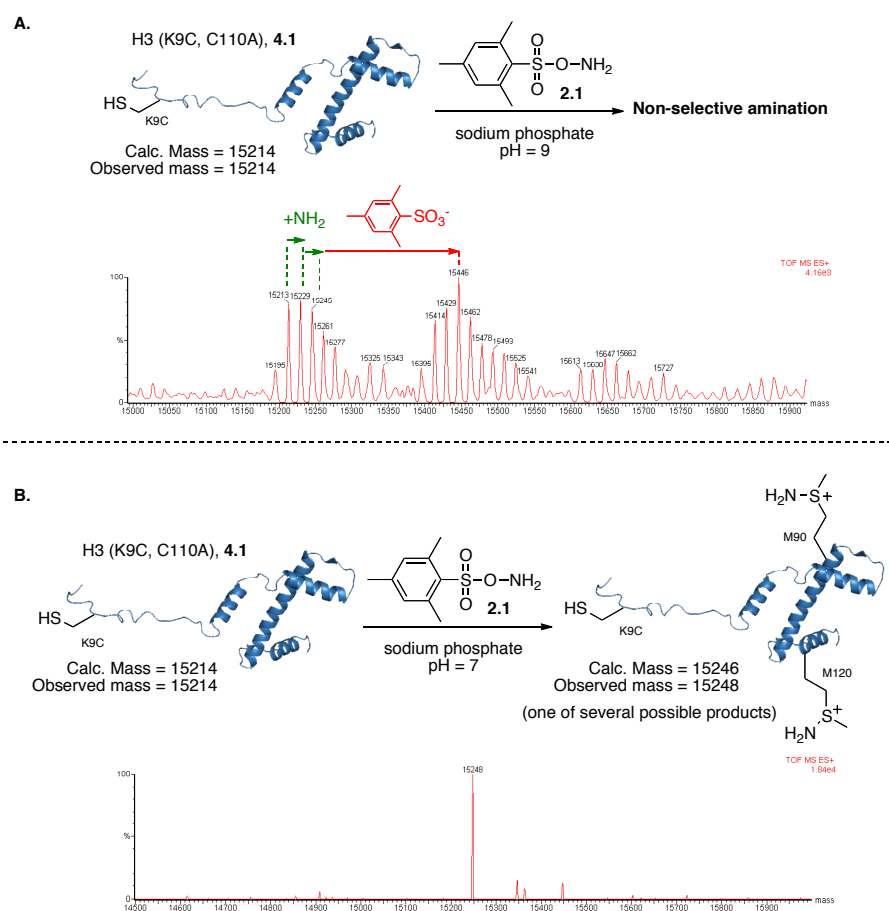


Figure 4.9: Reaction of H3 (K9C, C110A) with MSH (**2.1**) at pH 9.0 (A) and at pH 7.0 (B)

and histidine should be much less reactive in their protonated form at the lower pH.

To determine if methionine was interfering with the desired reaction, a methionine-free mutant was prepared, H3 (K9C, M90L, C110A, M120L).[†] Non-selective reaction with MSH was still observed at pH 7.0 with this mutant (see Experimental Procedures), but lowering the pH to 6.0 resulted in the desired loss of 34 mass units (Figure 4.10A). The product of this reaction did not react with Ellman's reagent, consistent with reaction at cysteine (4.10B). The addition of thiol **3.4** also corroborates the formation of dehydroalanine and provides a mimic of H3 with monomethylation at lysine 9 (Figure 4.10C).

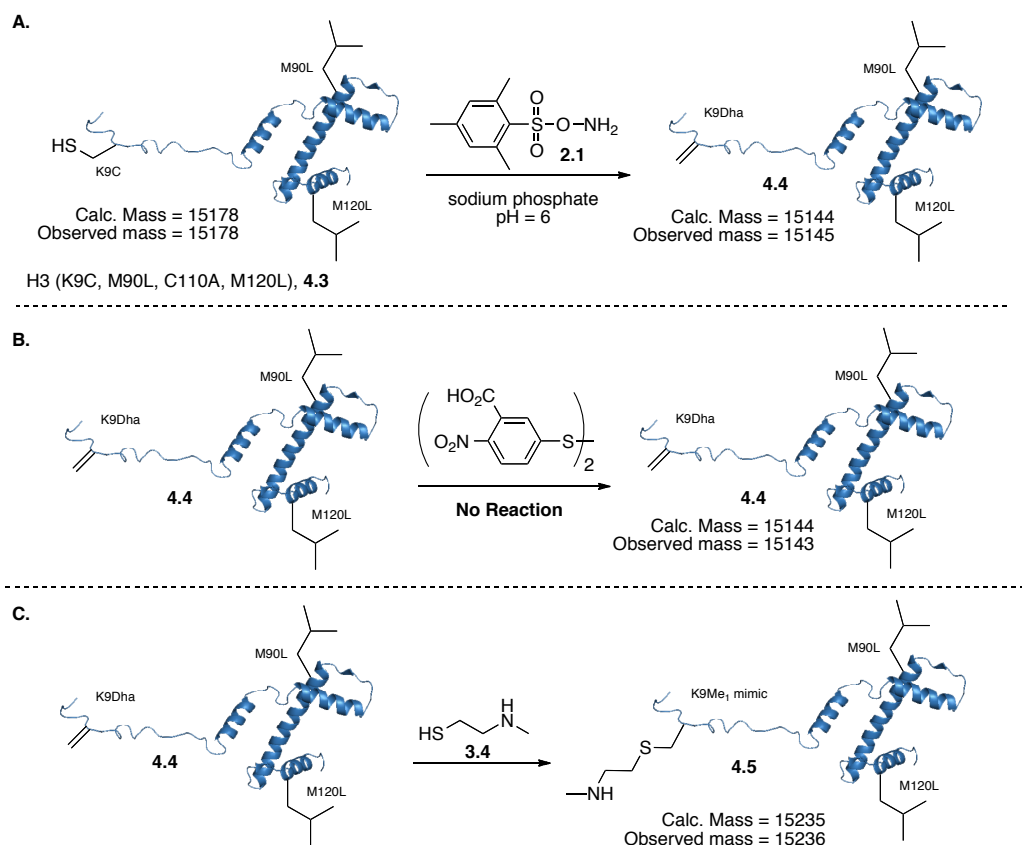


Figure 4.10: Reaction of H3 (K9C, M90L, C110A, M120L) at pH 6.0 and chemical tests for the conversion of cysteine to dehydroalanine.

The results in Figures 4.9 and 4.10 suggest that the selectivity of MSH can be controlled to some extent by pH adjustments. Using these experimental precautions, it was possible to synthesize a monomethyllysine mimic at position 9 of H3. For general utility in histone modification, however, it is desirable to keep the histone sequence as similar as possible to the native sequence and mutating methionines 90 and 120 to leucine is sidestepping this issue. While it was shown in Chapter 2 that the product from the reaction of MSH and methionine can be reduced to the native side chain with DTT, this is not an ideal protocol for histone modification since the MS analysis becomes more complicated and DTT may react with dehydroalanine. Moreover, it was a stated goal at the beginning of this Chapter to overcome side reactions at Met90 and M120 that were also observed by the Schultz group in their oxidative eliminations on H3 using hydrogen peroxide.⁸⁹ To install dehydroalanine at position 9 on H3 and avoid side reactions with methionine, di-bromide **2.32** was investigated next.

Histone modification using a bis-alkylation-elimination protocol with reagent **2.32**: synthesis of methylated and acetylated lysine mimics

Histone H3 (K9C, C110A) (**4.1**) was reacted with di-bromide **2.32** to determine if cysteine could be selectively converted to dehydroalanine in the presence of the native methionine residues of H3. The reaction course was followed by LC-MS and the results were entirely consistent with those observed on other proteins in Chapters 2 and 3. After 30 minutes at room temperature, all starting material was consumed and the major product was the alkylated product **4.6** (Figure 4.11A). A small amount of dehydroalanine was also detected by ESI-MS after this time. The reaction was driven to completion by simply incubating at 37 °C for an additional hour. The product did not react with Ellman's reagent, consistent with elimination at cysteine (Figure 4.11B) and the presence of dehydroalanine was corroborated by the addition of 2-mercaptoethanol to give thioether **4.7** (Figure 4.11C).

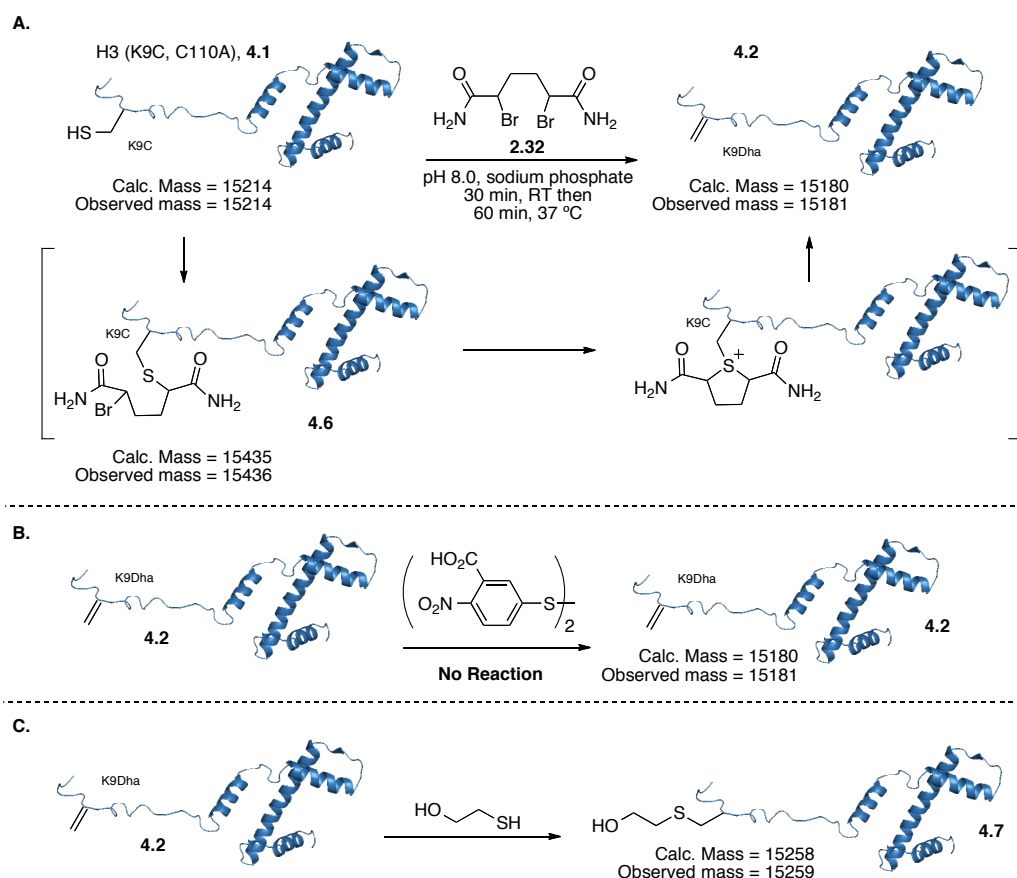


Figure 4.11: Reaction of H3 (K9C, C110A) with di-bromide **2.32** and chemical tests for the conversion of cysteine to dehydroalanine.

With a reliable route to dehydroalanine now established on histone H3, its use in the installation of post-translational modifications was investigated next. After conversion to dehydroalanine, the appropriate thiol was simply added for the conjugate addition. In this way, a one-pot procedure was established for the conversion of cysteine to mimics of methylated and acetylated lysine (Figure 4.12).

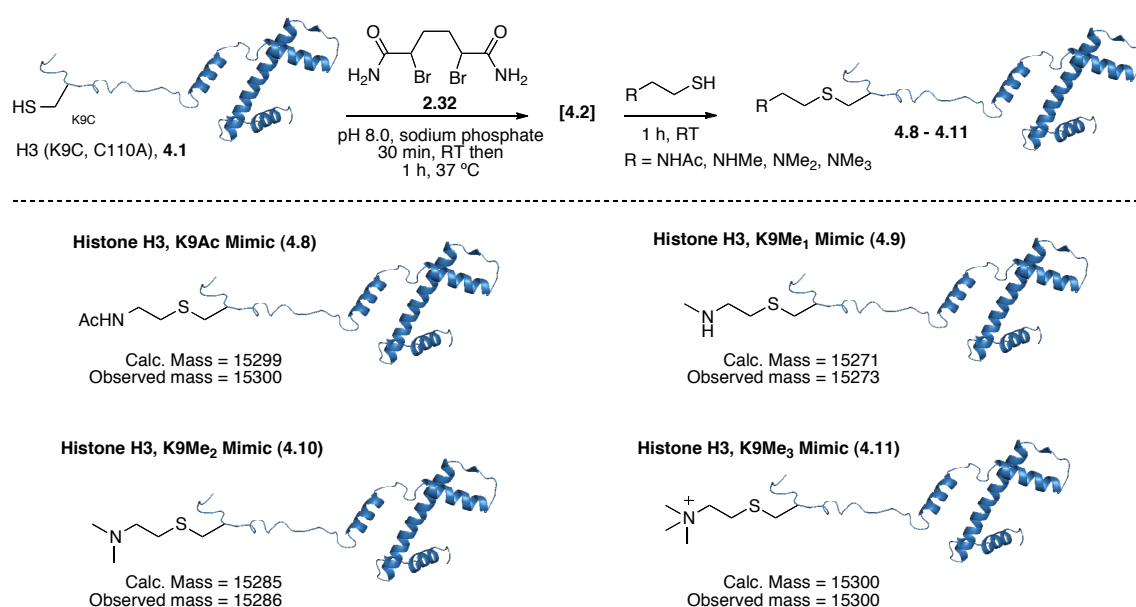


Figure 4.12: One-pot conversion of cysteine to acetylated and methylated lysine mimics at position 9 of histone H3.

Full conversions were observed in all cases by LC-MS.* After separation of the protein from small molecules using a size-exclusion column or dialysis, good protein recovery was obtained (typically 80-95% from the single cysteine mutant). Both the conversion to dehydroalanine and subsequent addition of the thiol are operationally simple and provide, from a single precursor, access to methylated and acetylated lysine mimics. Of course the methylated lysine mimics can be accessed by simple alkylation of cysteine,⁷¹ but the acetylated lysine mimic **4.8** cannot be accessed by simple alkylation

* Reaction conversions reported correspond to the relative peak height of starting material and product in the ESI-MS. For histone H3, good correlation between relative peak height and independently measured amounts of protein was observed. The full details are given in the Experimental Procedures. In short, two histone mutants of different mass were mixed in known amounts, as measured by bicinchoninic acid (BCA) assay. These mixtures were analyzed by LC-MS and the relative peak height by ESI-MS correlated well ($R^2 > 0.99$) to the relative amount of histone measured by BCA assay.

(cf. Figure 4.6B).⁷² Dehydroalanine therefore provides a convenient route to acetylated histones.

Next, the validity of these modifications as mimics for the natural acetylated and methylated lysine was investigated using western blot analysis (Figure 4.13). In these blots, commercially available antibodies were used that are raised to the natural acetylated or methylated lysine at position 9 of histone H3. In all cases, antibody recognition was observed. While it is not clear if the antibody recognizes one or both epimers presumably formed from thiol addition to dehydroalanine, the binding observed in Figure 4.13 corroborates the site-specificity of the modification reaction.

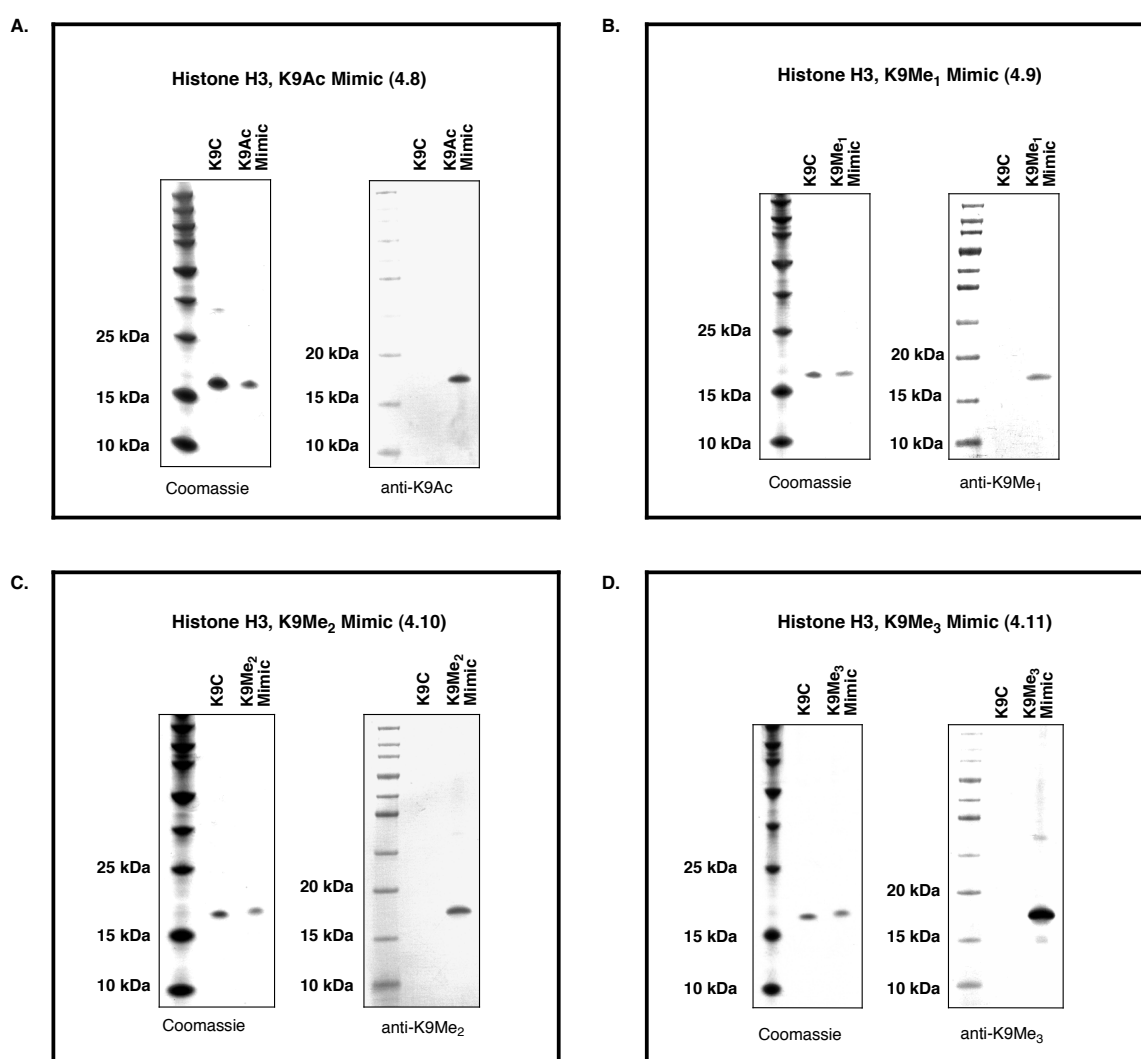


Figure 4.13: Western blot of acetylated and methylated lysine mimics 4.8 – 4.11. The antibodies used in this assay were raised to the natural modification.

Consistent with the results in Figure 4.13, Shokat and co-workers have shown that antibodies raised to the natural mono-, di-, and trimethyllysine recognize the corresponding cysteine derived analogs,⁷¹ but it was necessary to show that methylated histones **4.9** – **4.11** are also recognized since they are derived from dehydroalanine and presumably exist as a mixture of epimers. Moreover, while the Schultz laboratory has installed the same acetyllysine mimic used in protein **4.8** on a methionine-free H3 mutant, no immunoblot of this mimic was performed and only indirect evidence was presented for its validity as an acetyllysine mimic.⁸⁹ The western blot of histone **4.8** (Figure 4.13A) is the first direct evidence that this modification can mimic acetyllysine.

Two major classes of lysine modification, methylation and acetylation, are accessible from dehydroalanine which in turn is installed simply and selectively using reagent **2.32**. Other classes of modifications are also accessible from dehydroalanine, as demonstrated in Chapter 3. Two of these PTMs—phosphorylation and glycosylation—were explored next on histone H3.

Synthesis of a phosphorylated serine mimic on histone H3 *via* dehydroalanine

In Chapter 3, it was demonstrated that sodium thiophosphate can add to dehydroalanine, generating phosphocysteine. It was suggested in that Chapter that phosphorylated cysteine might be a suitable mimic for phosphorylated serine. This is a particularly attractive hypothesis since no chemical method has been reported for the site-selective phosphorylation of serine or cysteine on proteins. Histone H3 is also an interesting target for phosphorylation since it bears this modification at several residues (Figure 4.3). Residue 10 was selected as a target since phosphorylation at serine 10 of H3 is a critical modification in transcriptional activation and cell cycle progression.²⁴⁻²⁶

To determine if phosphocysteine can mimic phosphoserine on H3, another histone mutant was prepared with a single cysteine at position 10.[†] This mutant, H3 (S10C, C110A), was first reacted with di-bromide **2.32** to test the conversion of cysteine to dehydroalanine. Using the same protocol described for the other histone mutant, dehydroalanine was efficiently installed at position 10 of H3 (Figure 4.14A). The same battery of chemical tests was performed to verify the conversion of cysteine to dehydroalanine: the product did not react with Ellman's reagent, indicating consumption of cysteine (Figure 4.14B), and the efficient addition of 2-mercaptoethanol corroborated the presence of dehydroalanine (Figure 4.14C).

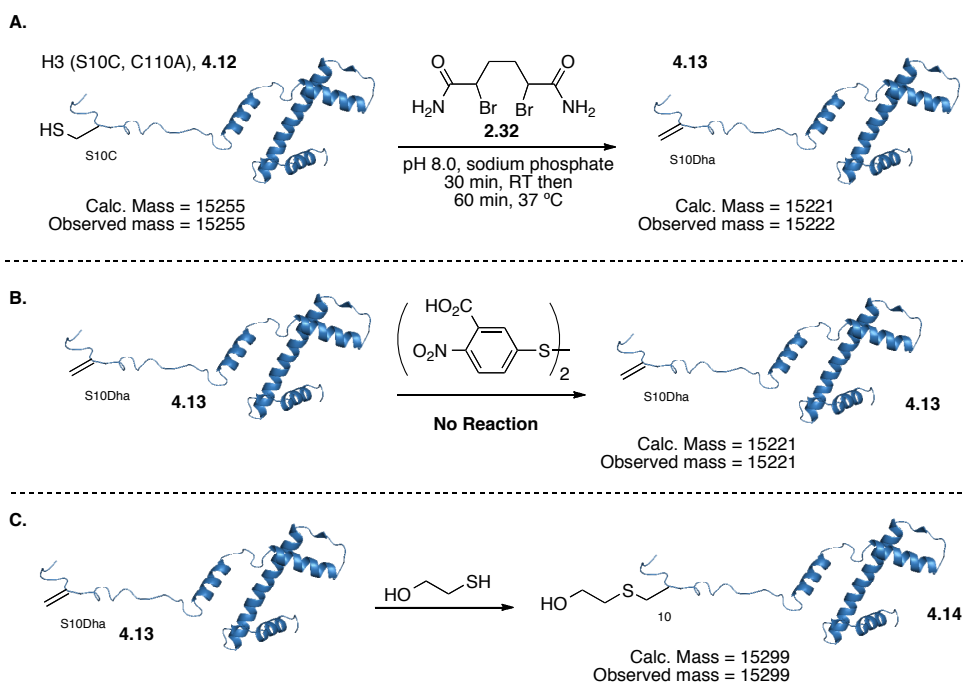


Figure 4.14: Reaction of H3 (S10C, C110A) with di-bromide **2.32** and chemical tests for the conversion of cysteine to dehydroalanine.

Next, the conversion of dehydroalanine to phosphocysteine was tested. Simply incubating thiophosphate with the protein **4.13** for an hour at 37 °C resulted in full conversion to the phosphorylated histone **4.15** (Figure 4.15). Validation of phosphorylated histone **4.15** as a mimic to the natural variant was carried out by western blot analysis. The antibody used in this assay was raised to the natural phosphorylated serine residue at position 10 of H3. Clear binding was observed, indicating the structural similarity of **4.15** to the natural protein (Figure 4.16). This is the first demonstration that phosphocysteine can function as a surrogate for phosphorylated serine. Moreover, while phosphorylated histones have been synthesized using NCL and EPL,^{64-66,97} this is the first example of a site-selective chemical installation of a phosphoserine mimic.

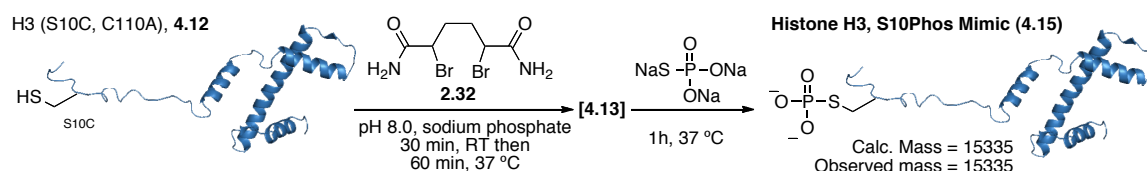


Figure 4.15: One-pot conversion of cysteine to a phosphocysteine, a mimic of phosphoserine, at position 10 of histone H3.

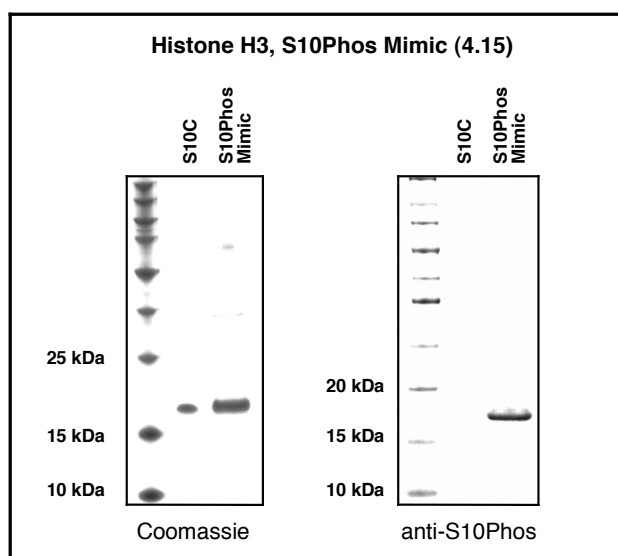


Figure 4.16: Western blot of phosphorylated serine mimic **4.15**. The antibody used in this assay was raised to the natural modification.

Synthesis of a glycosylated serine mimic on histone H3 *via* dehydroalanine

Many proteins, including histones, are modified at serine and threonine with β -*N*-acetyl-D-glucosamine (*O*-GlcNAcylation).^{44,45,98} This modification is intimately linked to phosphorylation: both modifications are dynamically cycled on and off the same residues and there is strong evidence for cross-talk between these modifications on various proteins.⁹⁹ The relationship between *O*-GlcNAcylation and phosphorylation in the context of the histone code, however, has not been studied. In fact, only recently was *O*-GlcNAc established unambiguously as a histone PTM.⁴⁵ Like other histone post-translational modifications, studies of the role of *O*-GlcNAc in chromatin regulation would benefit from access to full-length histones bearing *O*-GlcNAc or a suitable mimic. In Chapters 2 and 3, it was shown that dehydroalanine is a convenient precursor to cysteine-linked GlcNAc (*S*-GlcNAc). This chemistry was next applied to H3 at position 10 to establish a simple method for the synthesis of glycosylated histones.

GlcNAcylated H3 was synthesized by a one-pot procedure where the single cysteine at position 10 was first converted to dehydroalanine and then to GlcNAc cysteine by the addition of glycosyl thiol **2.37** (Figure 4.17). Full conversion to the GlcNAcylated histone **4.16** was observed. Hart has shown that all core histones (H2A, H2B, H3, and H4) carry the *O*-GlcNAc PTM.⁴⁵ The site of *O*-GlcNAc on H3, however, has not been mapped. The synthesis in Figure 4.17 is therefore ‘bio-predictive’ and may help in assays to determine the natural site and function of this modification. For instance, proteins such

as **4.16** could be immobilized and used in pull-down assays in the search for binding partners. Furthermore, the synthesis in Figure 4.17 should be applicable to any serine to cysteine mutation in H3 or other histones.

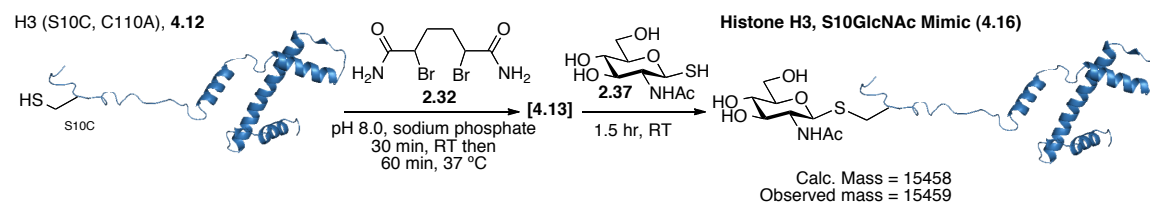


Figure 4.17: One-pot conversion of cysteine to a glycosyl cysteine at position 10 of histone H3.

The synthesis of phosphorylated and glycosylated histones was accomplished using site-selective chemical modification of dehydroalanine. For these modifications, no alternative chemical modification has been reported for protein substrates outside total synthesis by NCL or EPL. The technical simplicity of the dehydroalanine strategy bodes well for the study of these modifications and their place in the histone code. However, the mixture of epimers that likely results from these modifications must always be taken into consideration for any assay. The next section illustrates how even a mixture of these epimers might be useful in an enzymatic assay.

The use of histone PTM mimics in the study of histone modifying enzymes

In the western blots of the methylated and acetylated lysine mimics, as well as the phosphorylated serine mimic, it is not clear whether the antibody binds to one or both epimers. Nevertheless, these proteins can be used in enzymatic assays. For instance, if an enzyme is expected to have some function but evidence from peptide substrates is inconclusive, the mimics reported in this Chapter can be assayed. A simple experiment with histone deacetylase 1 (HDAC1) and histone deacetylase 2 (HDAC2) illustrates how such an experiment might be revealing (Figure 4.18).

A batch of H3 with the acetyllysine mimic at position 9 (histone **4.8**) was prepared as described above and purified by dialysis. This protein was then incubated with HDAC1 or HDAC2 for 12 hours at 37 °C. Control experiments were run in parallel where histone **4.8** was incubated for the same time and at the same concentration, but without any HDAC. After the incubation period, the reaction mixtures and controls were analyzed by SDS-PAGE, western blot, and LC-MS. The results are shown in Figure 4.18 and 4.19.

It is clear that when the acetyllysine mimic **4.8** is incubated with either HDAC enzyme, very little binding is observed in the western blot for the H3 K9Ac modification. No protein is lost in the assay—coomassie staining shows equal amounts of histone in the reaction and the control experiment without HDAC. Additionally, the western blot shows strong binding in the control experiments so the acetyl group is not lost by adventitious hydrolysis. Together, these results indicate that the lysine mimic is a substrate for HDAC1 and HDAC2.

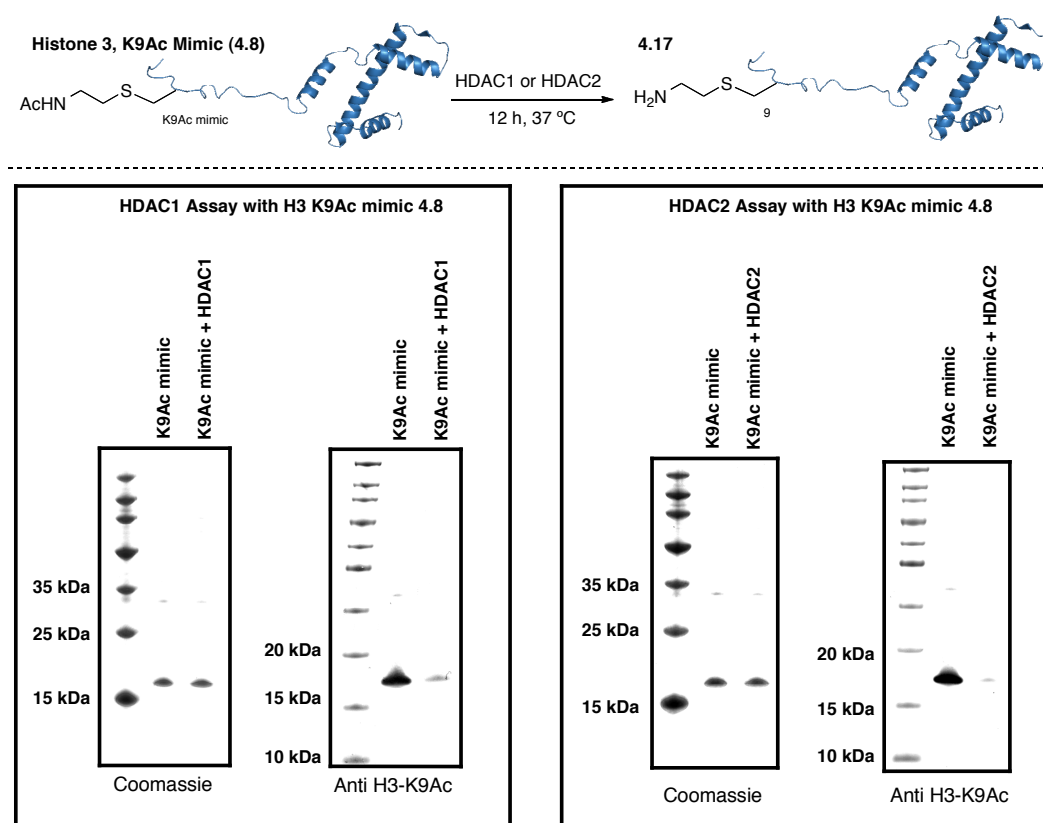


Figure 4.18: Histone deacetylase (HDAC) assay using H3 K9Ac mimic **4.8** as a substrate for HDAC1 and HDAC2

While western blot analysis clearly indicates histone deacetylation, LC-MS analysis of the reactions revealed only 42% ($\pm 3\%$) and 55% ($\pm 3\%$) deacetylation by HDAC1 and HDAC2, respectively (Figure 4.19). Given the western blot analysis in Figure 4.18 and LC-MS analysis in Figure 4.19, these results are consistent with selective deacetylation of the natural L-epimer of **8** as well as selective binding to the L-epimer by the primary antibody used in the western blot. While kinetic data might be difficult to

extract from such an assay, nominal HDAC activity can be determined unambiguously. In fact, the western blot and LC-MS analysis in Figures 4.18 and 4.19 are the first direct evidence for HDAC activity on the acetyllysine mimic in **4.8**. Additionally, the only other acetyllysine mimic available, the thiocarbamate analog introduced by Cole and co-workers (Scheme 1A), is not a substrate for HDAC enzymes. Therefore, the use of Dha and its conversion to acetyllysine mimics such as **8** is currently the most convenient strategy for investigating HDAC activity on full-length histones.

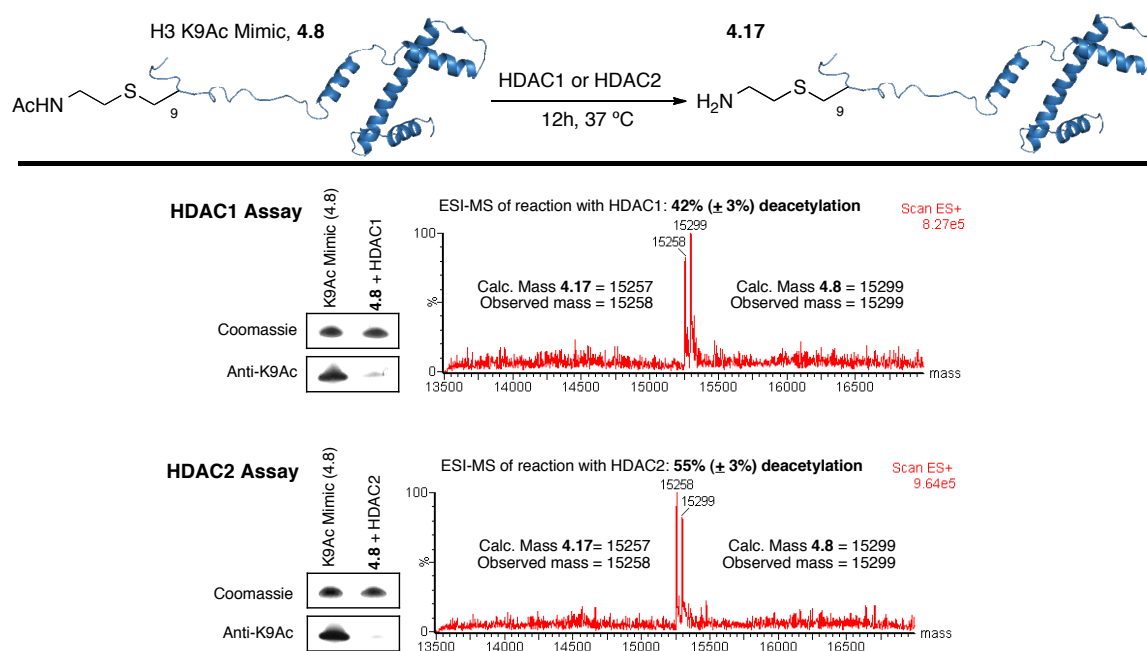


Figure 4.19: LC-MS analysis of histone deacetylase (HDAC) assay using H3 K9Ac mimic **4.8** as a substrate for HDAC1 and HDAC2

4.3 Conclusions and Outlook

In this Chapter, the conversion of cysteine to dehydroalanine was studied on histone proteins. Two reagents—MSH (**2.1**) and di-bromide **2.32**—were used in this study. MSH was discovered to be more selective at lower pH, but it is too reactive for general use in histone modification. Di-bromide **2.32**, however, is mild and selective and allowed the installation of dehydroalanine on multiple mutants of histone H3. The reaction is operationally simple and full conversion and good protein recovery was observed in all cases examined. Dehydroalanine has previously been installed on histone H3 by the oxidative elimination of a genetically encoded phenylselenocysteine residue.⁸⁹ This

method, however, led to unwanted oxidation of methionine side chains. The use of reagent **2.32** resolves this problem.

With a reliable route to dehydroalanine on histone proteins, several post-translational modifications were installed. Mimics of methylated and acetylated lysine, phosphorylated serine, and glycosylated serine were all synthesized from dehydroalanine. All transformations can be completed in only a few hours from recombinant protein and all reagents are easy to prepare and many are commercially available. The methylated lysine mimics can be accessed by simple alkylation of cysteine⁷¹ so the use of dehydroalanine is not meant to replace this method. However, complementary use can be envisioned. For instance, a photocaged cysteine could be incorporated genetically¹⁰⁰ at one site while free cysteine could be incorporated at a second site. Modification of the free cysteine and deprotection and modification of the second cysteine might allow two different modifications to be installed chemically. Dehydroalanine would be useful in this scenario to access modifications not easily accessible from cysteine itself. Acetyllysine, phosphoserine, and GlcNAc serine are such modifications. All three of these modifications were demonstrated on H3 in this Chapter.

Acetyllysine can be mimicked by the thiocarbamate derivative reported by Cole and co-workers (Figure 4.6B). This analog is recognized by antibodies raised to the natural modification and it is stable to HDAC enzymes. This stability can be used to one's advantage in cellular assays where the resilience of the acetyllysine mimic is essential. To assay HDAC activity on full-length histones, however, a different mimic or the natural residue is needed. Such a mimic was synthesized from dehydroalanine on H3. This protein (**4.8**) was recognized by antibodies raised to the natural modification and was also used as a substrate for HDAC1 and HDAC2. Of course, the natural acetyllysine residue can be installed by NCL or EPL.^{59,62,63,65} When multiple, different modifications are required, these are the methods of choice. Recombinant biosynthesis of acetylated histones is also possible, though not yet widely available.^{90,91} For the installation of a single acetyllysine mimic, however, the synthesis from dehydroalanine is comparatively simple. Using dehydroalanine requires no more specialized equipment or techniques than normally required in routine protein expression and analysis. This chemistry should therefore be accessible to chemists and non-chemists alike.

The use of dehydroalanine to access phosphorylated and glycosylated serine mimics on histones was also reported. Other than NCL or EPL, no other chemical methods have been described to access these mimics on histones. Again, the operational

simplicity of installing these modifications through dehydroalanine is advantageous. The demonstration that antibodies raised to phosphoserine bind to phosphocysteine verify that these residues are indeed structurally similar. The use of phosphocysteine in kinase and phosphatase pathways is currently under investigation. Access to GlcNAcylated histones also provide the opportunity to dissect their role in chromatin structure and function.

The chemistry reported in this Chapter has prompted several other studies that are currently underway. One goal is to determine the influence of a given modification on the action of other chromatin modifying enzymes. For instance, “cross-talk” between phosphorylation or glycosylation at position 10 of H3 and other modifications at neighboring residues can be studied using **4.15** and **4.16**. Structural investigations are also underway. It might be the case that one epimer of a histone modified through dehydroalanine will bind selectively to an effector protein. This “epimeric resolution” might be revealed in ongoing crystallographic studies. If such selective crystallization is observed, another scenario will be demonstrated where the epimeric mixture of modified histones is not necessarily a problem.

Dehydroalanine is a versatile precursor to mimics of natural post-translational modifications. Nevertheless, as a target for synthesis it is often more satisfying—both aesthetically and experimentally—to access the product in its natural form. Is it possible to use dehydroalanine to synthesize the *natural* constitution of these residues, rather than mere analogs? This intriguing question will be explored in more detail in Chapter 8. At this point, suffice is to say that such a transformation would require entirely different chemistry than the thiol additions so pervasive in this and previous Chapters. Indeed, a carbon-carbon bond would be required for the synthesis of lysine residues from dehydroalanine—no small feat on protein substrates in aqueous media. Before addressing this question head-on, various methods for the formation of carbon-carbon bonds on proteins were studied. Chapter 5 and Chapter 6 describe the use of olefin metathesis and Suzuki-Miyaura cross-coupling on protein substrates, respectively. Unlike Chapters 2, 3, and 4—the next two chapters focus on *unnatural* protein modification and the opportunities in protein chemistry enabled by such transformations.

4.4 References

- (1) Li, B.; Carey, M.; Workman, J. L. The role of chromatin during transcription. *Cell* **2007**, *128*, 707-719.
- (2) Groth, A.; Rocha, W.; Verreault, A.; Almouzni, G. Chromatin challenges during DNA replication and repair. *Cell* **2007**, *128*, 721-733.
- (3) Kornberg, R. D.; Thomas, J. O. Chromatin structure: oligomers of the histones. *Science* **1974**, *184*, 865-868.
- (4) Kornberg, R. D. Chromatin structure: a repeating unit of histones and DNA. *Science* **1974**, *184*, 868-871.
- (5) Kornberg, R. D.; Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **1999**, *98*, 285-294.
- (6) Oudet, P.; Gross-Bellard, M.; Chambon, P. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* **1975**, *4*, 281-300.
- (7) Arents, G.; Burlingame, R. W.; Wang, B.-C.; Love, W. E.; Moudrianakis, E. N. The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10148-10152.
- (8) Luger, K.; Mäder, A. W.; Richmond, R. K.; Sargent, D. F.; Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **1997**, *389*, 251-260.
- (9) Stedman, E.; Stedman, E. Cell specificity of histones. *Nature* **1950**, *166*, 780-781.
- (10) Stedman, E.; Stedman, E. The basic proteins of cell nuclei. *Phil. Trans. Roy. Soc. London, Ser. B* **1951**, *235*, 565-595.
- (11) Huang, R.-C. C.; Bonner, J. Histone, a suppressor of chromosomal RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1962**, *48*, 1216-1222.
- (12) Allfrey, V. G.; Littau, V. C.; Mirsky, A. E. On the role of histones in regulating ribonucleic acid synthesis in the cell nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **1963**, *49*, 414-421.
- (13) Barr, G. C.; Butler, J. A. V. Histones and gene function. *Nature* **1963**, *199*, 1170-1172.
- (14) Phillips, D. M. P. Acetyl groups as *N*-terminal substituents in calf-thymus histones. *Biochem. J.* **1961**, *80*, 40P-41P.

- (15) Phillips, D. M. P. The presence of acetyl groups in histones. *Biochem. J.* **1963**, *87*, 258-263.
- (16) Murray, K. The occurrence of ϵ -N-methyl lysine in histones. *Biochemistry* **1964**, *3*, 10-15.
- (17) Allfrey, V. G.; Faulkner, R.; Mirsky, A. E. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **1964**, *51*, 786-794.
- (18) Kimura, A.; Matsubara, K.; Horikoshi, M. A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. *J. Biochem.* **2005**, *138*, 647-662.
- (19) Martin, C.; Zhang, Y. The diverse functions of histone lysine methylation. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 838-849.
- (20) Isenberg, I. Histones. *Ann. Rev. Biochem.* **1979**, *48*, 159-191.
- (21) Kouzarides, T. Chromatin modifications and their function. *Cell* **2007**, *128*, 693-705.
- (22) Gurley, L. R.; D'Anna, J. A.; Barham, S. S.; Deaven, L. L.; Tobey, R. A. Histone phosphorylation and chromatin structure during mitosis in chinese hamster cells. *Eur. J. Biochem.* **1978**, *84*, 1-15.
- (23) Green, G. R. Phosphorylation of histone variant regions in chromatin: unlocking the linker? *Biochem. Cell Biol.* **2001**, *79*, 275-287.
- (24) Hans, F.; Dimitrov, S. Histone H3 phosphorylation and cell division. *Oncogene* **2001**, *20*, 3021-3027.
- (25) Nowak, S. J.; Corces, V. G. Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. *Trends Genet.* **2004**, *20*, 214-220.
- (26) Dyson, M. H.; Thomson, S.; Mahadevan, L. C. Heat shock, histone H3 phosphorylation and the cell cycle. *Cell Cycle* **2005**, *4*, 13-17.
- (27) Kobza, K.; Camporeale, G.; Rueckert, B.; Keuh, A.; Griffin, J. B.; Sarath, G.; Zemleni, J. K4, K9 and K18 in human histone H3 are targets for biotinylation by biotinidase. *FEBS J.* **2005**, *272*, 4249-4259.
- (28) Ching Chew, Y.; Camporeale, G.; Kothapalli, N.; Sarath, G.; Zemleni, J. Lysine residues in N-terminal and C-terminal regions of human histone H2A are targets for biotinylation by biotinidase. *J. Nutr. Biochem.* **2006**, *17*, 225-233.

- (29) Kothapalli, N.; Camporeale, G.; Keuh, A.; Chew, Y. C.; Oommen, A. M.; Griffin, J. B.; Zemleni, J. Biological functions of biotinylated histones. *J. Nutr. Biochem.* **2005**, *16*, 446-448.
- (30) Goldknopf, I. L.; Taylor, C. W.; Baum, R. M.; Yeoman, L. C.; Olson, M. O. J.; Prestayko, A. W.; Busch, H. Isolation and characterization of protein A24, a "histone-like" non-histone chromosomal protein. *J. Biol. Chem.* **1975**, *250*, 7182-7187.
- (31) Levinger, L.; Varshavsky, A. Selective arrangement of ubiquitinated and D1 protein-containing nucleosomes within the drosophila genome. *Cell* **1982**, *28*, 375-385.
- (32) Jason, L. J. M.; Moore, S. C.; Lewis, J. D.; Lindsey, G.; Ausió, J. Histone ubiquitination: a tagging tail unfolds? *BioEssays* **2002**, *24*, 166-174.
- (33) Muratani, M.; Tansey, W. P. How the ubiquitin-proteasome system controls transcription. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 1-10.
- (34) Osley, M. A. H2B ubiquitylation: the end is in sight. *Biochim. Biophys. Acta* **2004**, *1677*, 74-78.
- (35) Higashi, M.; Inoue, S.; Ito, T. Core histone H2A ubiquitylation and transcriptional regulation. *Exp. Cell Res.* **2010**, *316*, 2707-2712.
- (36) Shiio, Y.; Eisenman, R. N. Histone sumoylation is associated with transcriptional repression. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13225-13230.
- (37) Gary, J. D.; Clarke, S. RNA and protein interactions modulated by protein arginine methylation. *Prog. Nucleic Acid Res. Mol. Biol.* **1998**, *61*, 65-131.
- (38) Bedford, M. T.; Richard, S. Arginine methylation: an emerging regulator of protein function. *Mol. Cell* **2005**, *18*, 263-272.
- (39) Hagiwara, T.; Nakashima, K.; Hirano, H.; Senshu, T.; Yamada, M. Deimination of arginine residues in nucleophosmin/B23 and histones in HL-60 granulocytes. *Biochem. Biophys. Res. Commun.* **2002**, *290*, 979-983.
- (40) Cuthbert, G. L.; Daujat, S.; Snowden, A. W.; Erdjument-Bromage, H.; Hagiwara, T.; Yamada, M.; Schneider, R.; Gregory, P. D.; Tempst, P.; Bannister, A. J.; Kouzarides, T. Histone deimination antagonizes arginine methylation. *Cell* **2004**, *118*, 545-553.
- (41) Wang, Y.; Wysocka, J.; Sayegh, J.; Lee, Y.-H.; Perlin, J. R.; Leonelli, L.; Sonbuchner, L. S.; McDonald, C. H.; Cook, R. G.; Dou, Y.; Roeder, R. G.; Clarke, S.; Stallcup, M. R.; Allis, C. D.; Coonrod, S. A. Human PAD4 regulates

- histone arginine methylation levels via demethyliminination. *Science* **2004**, *306*, 279-283.
- (42) Thompson, P. R.; Fast, W. Histone citrullination by protein arginine deiminase: is arginine methylation a green light or a roadblock? *ACS Chem. Biol.* **2006**, *1*, 433-441.
 - (43) Hassa, P. O.; Haenni, S. S.; Elser, M.; Hottiger, M. O. Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 789-829.
 - (44) Kelly, W. G.; Hart, G. W. Glycosylation of chromosomal proteins: localization of O-linked N-acetylglucosamine in drosophila chromatin. *Cell* **1989**, *57*, 243-251.
 - (45) Sakabe, K.; Wang, Z.; Hart, G. W. β -N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 19915-19920.
 - (46) Strahl, B. D.; Allis, C. D. The language of covalent histone modifications. *Nature* **2000**, *403*, 41-45.
 - (47) Chatterjee, C.; Muir, T. W. Chemical approaches for studying histone modifications. *J. Biol. Chem.* **2010**, *285*, 11045-11050.
 - (48) Jenuwein, T.; Allis, C. D. Translating the histone code. *Science* **2001**, *293*, 1074-1080.
 - (49) Allis, C. D.; Muir, T. W. Spreading chromatin into chemical biology. *ChemBioChem* **2011**, *12*, 264-279.
 - (50) Taverna, S. D.; Li, H.; Ruthenburg, A. J.; Allis, C. D.; Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat. Struct. Mol. Biol.* **2007**, *14*, 1025-1040.
 - (51) Voigt, P.; Reinberg, D. Histone tails: ideal motifs for probing epigenetics through chemical biology approaches. *ChemBioChem* **2011**, *12*, 236-252.
 - (52) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776-779.
 - (53) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705-6710.
 - (54) Evans, T. C.; Benner, J.; Xu, M.-Q. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **1998**, *7*, 2256-2264.
 - (55) Dawson, P. E.; Kent, S. B. H. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923-960.

- (56) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91-118.
- (57) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338-351.
- (58) Flavell, R. R.; Muir, T. W. Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology. *Acc. Chem. Res.* **2009**, *42*, 107-116.
- (59) He, S.; Bauman, D.; Davis, J. S.; Loyola, A.; Nishioka, K.; Gronlund, J. L.; Reinberg, D.; Meng, F.; Kelleher, N.; McCafferty, D. G. Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12033-12038.
- (60) Yan, L. Z.; Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526-533.
- (61) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-9252.
- (62) Shogren-Knaak, M.; Ishii, H.; Sun, J.-M.; Pazin, M. J.; Davie, J. R.; Peterson, C. L. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **2006**, *311*, 844-847.
- (63) Ferreira, H.; Flaus, A.; Owen-Hughes, T. Histone modifications influence the action of Snf2 family remodelling enzymes by different mechanisms. *J. Mol. Biol.* **2007**, *374*, 563-579.
- (64) Shogren-Knaak, M. A.; Fry, C. J.; Peterson, C. L. A native peptide ligation strategy for deciphering nucleosomal histone modifications. *J. Biol. Chem.* **2003**, *278*, 15744-15748.
- (65) Chiang, K. P.; Jensen, M. S.; McGinty, R. K.; Muir, T. W. A semisynthetic strategy to generate phosphorylated and acetylated histone H2B. *ChemBioChem* **2009**, *10*, 2182-2187.
- (66) Kee, J.-M.; Villani, B.; Carpenter, L. R.; Muir, T. W. Development of stable phosphohistidine analogues. *J. Am. Chem. Soc.* **2010**, *132*, 14327-14329.
- (67) Manohar, M.; Mooney, A. M.; North, J. A.; Nakkula, R. J.; Picking, J. W.; Edon, A.; Fishel, R.; Poirier, M. G.; Ottesen, J. J. Acetylation of histone H3 at the

- nucleosome dyad alters DNA-histone binding. *J. Biol. Chem.* **2009**, *284*, 23312-23321.
- (68) McGinty, R. K.; Kim, J.; Chatterjee, C.; Roeder, R. G.; Muir, T. W. Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. *Nature* **2008**, *453*, 812-816.
 - (69) Chatterjee, C.; McGinty, R. K.; Muir, T. W. Disulfide-directed histone ubiquitylation reveals plasticity in hDot1L activation. *Nat. Chem. Biol.* **2010**, *6*, 267-269.
 - (70) Fierz, B.; Chatterjee, C.; McGinty, R. K.; Bar-Dagan, M.; Raleigh, D. P.; Muir, T. W. Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. *Nat. Chem. Biol.* **2011**, *7*, 113-119.
 - (71) Simon, M. D.; Chu, F.; Racki, L. R.; de la Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **2007**, *128*, 1003-1012.
 - (72) Huang, R.; Holbert, M. A.; Tarrant, M. K.; Curtet, S.; Colquhoun, D. R.; Dancy, B. M.; Dancy, B. C.; Hwang, Y.; Tang, Y.; Meeth, K.; Marmorstein, R.; Cole, R. N.; Khochbin, S.; Cole, P. A. Site-specific introduction of an acetyl-lysine mimic into peptides and proteins by cysteine alkylation. *J. Am. Chem. Soc.* **2010**, *132*, 9986-9987.
 - (73) Lindley, H. A new synthetic substrate for trypsin and its application to the determination of the amino-acid sequence of proteins. *Nature* **1956**, *178*, 647-648.
 - (74) Lu, X.; Simon, M. D.; Chodaparambil, J. V.; Hansen, J. C.; Shokat, K. M.; Luger, K. The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat. Struct. Mol. Biol.* **2008**, *15*, 1122-1124.
 - (75) Xu, C.; Cui, G.; Botuyan, M. V.; Mer, G. Structural basis for the recognition of methylated histone H3K36 by the Eaf3 subunit of histone deacetylase complex Rpd3S. *Structure* **2008**, *16*, 1740-1750.
 - (76) Francis, N. J.; Follmer, N. E.; Simon, M. D.; Aghia, G.; Butler, J. D. Polycomb proteins remain bound to chromatin and DNA during DNA replication in vitro. *Cell* **2009**, *137*, 110-122.
 - (77) Margueron, R.; Justin, N.; Ohno, K.; Sharpe, M. L.; Son, J.; Drury, W. J.; Voigt, P.; Martin, S. R.; Taylor, W. R.; De Marco, V.; Pirrotta, V.; Reinberg, D.; Gamblin, S. J. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **2009**, *461*, 762-767.

- (78) Hung, T.; Binda, O.; Champagne, K. S.; Kuo, A. J.; Johnson, K.; Chang, H. Y.; Simon, M. D.; Kutateladze, T. G.; Gozani, O. ING4 mediates crosstalk between histone H3 K4 trimethylation and H3 acetylation to attenuate cellular transformation. *Mol. Cell* **2009**, *33*, 248-256.
- (79) Li, B.; Jackson, J.; Simon, M. D.; Fleharty, B.; Gogol, M.; Seidel, C.; Workman, J. L.; Shilatifard, A. Histone H3 lysine 36 dimethylation (H3K36me2) is sufficient to recruit the Rpd3s histone deacetylase complex and to repress spurious transcription. *J. Biol. Chem.* **2009**, *284*, 7970-7976.
- (80) Bader, B.; Kuhn, K.; Owen, D. J.; Waldmann, H.; Wittinghofer, A.; Kuhlmann, J. Bioorganic synthesis of lipid-modified proteins for the study of signal transduction. *Nature* **2000**, *403*, 223-226.
- (81) Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L. Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin. *Chem. Biol.* **2001**, *8*, 133-145.
- (82) Smith, J. J.; Conrad, D. W.; Cuneo, M. J.; Hellinga, H. W. Orthogonal site-specific protein modification by engineering reversible thiol protection mechanisms. *Protein Sci.* **2005**, *14*, 64-73.
- (83) Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **2000**, *4*, 645-652.
- (84) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-canonical amino acids in protein engineering. *Curr. Opin. Biotechnol.* **2003**, *14*, 603-609.
- (85) Xie, J.; Schultz, P. G. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.
- (86) Ohta, A.; Yamagishi, Y.; Suga, H. Synthesis of biopolymers using genetic code reprogramming. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159-167.
- (87) de Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconjugate Chem.* **2009**, *20*, 1281-1295.
- (88) Wang, J.; Schiller, S. M.; Schultz, P. G. A biosynthetic route to dehydroalanine-containing proteins. *Angew. Chem. Int. Ed.* **2007**, *46*, 6849-6851.
- (89) Guo, J.; Wang, J.; Lee, J. S.; Schultz, P. G. Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 6399-6401.

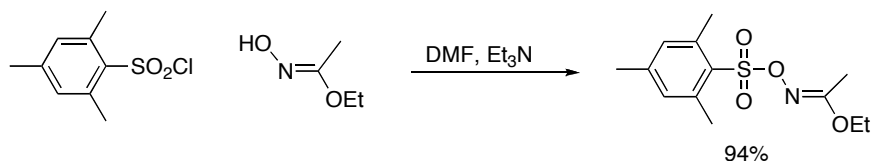
- (90) Neumann, H.; Peak-Chew, S. Y.; Chin, J. W. Genetically encoding *N*-acetyllysine in recombinant proteins. *Nat. Chem. Biol.* **2008**, *4*, 232-234.
- (91) Neumann, H.; Hancock, S. M.; Buning, R.; Routh, A.; Chapman, L.; Somers, J.; Owen-Hughes, T.; van Noort, J.; Rhodes, D.; Chin, J. W. A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol. Cell* **2009**, *36*, 153-163.
- (92) Nguyen, D. P.; Garcia Alai, M. M.; Kapadnis, P. B.; Neumann, H.; Chin, J. W. Genetically encoding *N*-methyl-L-lysine in recombinant histones. *J. Am. Chem. Soc.* **2009**, *131*, 14194-14195.
- (93) Ai, H.-w.; Lee, J. W.; Schultz, P. G. A method to site-specifically introduce methyllysine into proteins in *E. coli*. *Chem. Commun.* **2010**, *46*, 5506-5508.
- (94) Wang, Y.-S.; Wu, B.; Wang, Z.; Huang, Y.; Wan, W.; Russell, W. K.; Pai, P.-J.; Moe, Y. N.; Russell, D. H.; Liu, W. R. A genetically encoded photocaged *N*-methyl-L-lysine. *Mol. BioSyst.* **2010**, *6*, 1557-1560.
- (95) Nguyen, D. P.; Garcia Alai, M. M.; Virdee, S.; Chin, J. W. Genetically directing ϵ -*N,N*-dimethyl-L-lysine in recombinant histones. *Chem. Biol.* **2010**, *17*, 1072-1076.
- (96) Oliver, S. S.; Denu, J. M. Dynamic interplay between histone H3 modifications and protein interpreters: emerging evidence for a "histone language". *ChemBioChem* **2011**, *12*, 299-307.
- (97) Tarrant, M. K.; Cole, P. A. The chemical biology of protein phosphorylation. *Annu. Rev. Biochem.* **2009**, *78*, 797-825.
- (98) Zachara, N. E.; Hart, G. W. The emerging significance of *O*-GlcNAc in cellular regulation. *Chem. Rev.* **2002**, *102*, 431-438.
- (99) Wang, Z.; Gucek, M.; Hart, G. W. Cross-talk between GlcNAcylation and phosphorylation: site-specific phosphorylation dynamics in response to globally elevated *O*-GlcNAc. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 13793-13798.
- (100) Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G. A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* **2004**, *126*, 14306-14307.

4.5 Experimental Procedures

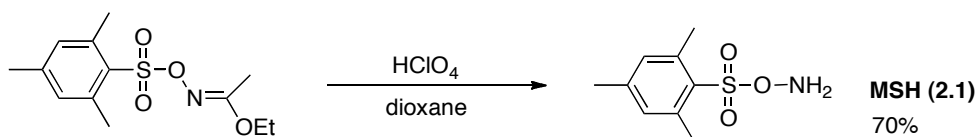
General Considerations

See Chapter 2 Experimental Procedures for general considerations in synthesis and protein analysis.

O-mesitylenesulfonylhydroxylamine (MSH, 2.1) Synthesis



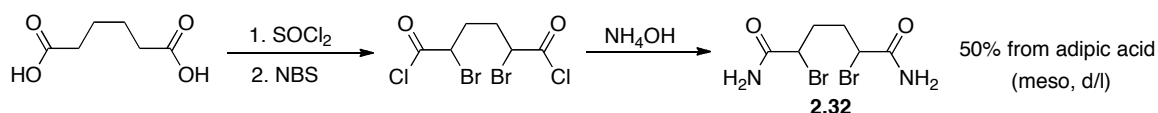
Ethyl *N*-hydroxyacetimidate (2.36 g, 22.9 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (15 mL). Et_3N (6.4 mL, 46 mmol) was added and the solution was cooled to 0 °C. 2-Mesitylenesulfonyl chloride (5.00 g, 22.9 mmol) was added as a solid in several portions. The reaction was stirred vigorously at 0 °C for 30 minutes. A white solid was formed over the course of the reaction. After this time, the reaction was diluted with EtOAc (150 mL) and then washed with H_2O (3×100 mL) and brine (100 mL). The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure to give a pale yellow oil that crystallized on storage at -20 °C (6.11 g, 94%). No purification was required. m.p. = 51-53 °C (Lit¹ = 57-58 °C). ^1H NMR (400 MHz, CDCl_3): δ = 1.19 (3H, t, J = 7.1, OCH_2CH_3), 2.04 (3H, s, CH_3 imidate), 2.31 (3H, s, Me_{Ar}), 2.64 (6H, s, Me_{Ar}), 3.90 (2H, q, J = 7.1, OCH_2CH_3), 6.97 (2H, s, CH_{Ar}). ^{13}C NMR (100 MHz, CDCl_3): δ = 14.0 (OCH_2CH_3), 14.9 (CH_3 imidate), 21.1 (Me_{Ar}), 22.8 (Me_{Ar}), 63.6 (OCH_2CH_3), 130.3, 131.5, 140.7, 143.3 ($4 \times \text{C}_{\text{Ar}}$), 169.7 ($\text{C}=\text{N}$).



The MSH precursor (6.11 g, 21.4 mmol) was dissolved in dioxane (8 mL) at room temperature and then cooled to 0 °C. Perchloric acid (4 mL, 70% solution) was added dropwise over 2 minutes. The reaction was stirred vigorously at 0 °C and thickens and solidifies over the course of 5 minutes. The reaction was incubated for an additional 5 minutes to ensure complete hydrolysis and then poured onto 50 g of ice. Et_2O (100 mL) and H_2O (100 mL) were both added to the quenched mixture and then transferred to a

separatory funnel. The organic layer was separated and then dried / neutralized for 2 minutes with K_2CO_3 . The solid was then removed by filtration and rinsed with Et_2O . The filtrate was concentrated carefully under reduced pressure to about 50 mL total volume. The solution was poured into a beaker with 150 mL ice cold petrol. MSH crystals formed immediately. The beaker was incubated at 0 °C for 30 minutes to complete the crystallization. The crystals were isolated by filtration and then dried under high vacuum for 15 minutes. The white needles (3.21 g, 70%) were stored in a plastic tube, sealed only with wax film. m.p. = 90-91 °C (Lit¹ = 95-96 °C). 1H NMR (400 MHz, $CDCl_3$): δ = 2.33 (3H, s, Me_{Ar}), 2.65 (6H, s, Me_{Ar}), 5.56 (2H, br. s, NH_2), 7.01 (2H, s, CH_{Ar}). ^{13}C NMR (50 MHz, $CDCl_3$): δ = 21.0 (Me_{Ar}), 22.6 (Me_{Ar}), 129.0, 131.6, 140.9, 143.7 ($4 \times C_{Ar}$). Found: C, 50.18%; H, 5.90%, N, 6.27%. $C_9H_{15}NO_4$ requires: C, 50.21%; H, 6.09%; N, 6.51%.

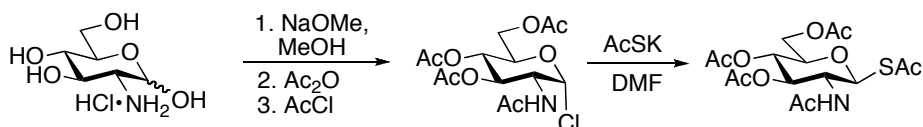
Synthesis of dibromide 2.32



α, α' -Di-bromo-adipyl(bis)amide was prepared using a procedure adapted from the literature.^{2, 3} Adipic acid (25.00 g, 171.1 mmol) was added to a 500 mL round bottom flask and suspended in thionylchloride (75.0 mL, 1034 mmol). The flask was equipped with a condenser and the reaction was heated to reflux (open to air, bath temp 80 °C). After 30 minutes at reflux, all adipic acid had dissolved. The reaction was stirred for an additional 60 minutes at reflux and then cooled to room temperature. CCl_4 (100 mL) was added to the reaction followed by NBS (73.1 g, 411 mmol). The reaction was stirred vigorously and 10 drops of HBr (48% aq.) was added by pipette. The reaction was heated to reflux, again open to air. The reaction gradually turns from red to black over the course of an hour. After 2 hours at reflux, the reaction was cooled to room temperature and then to 0 °C. The mixture was stirred at 0 °C to ensure all succinimide had precipitated. The solid was removed by filtration. Et_2O (50 mL) was used to rinse and complete the filtration. The filtrate was concentrated under reduced pressure to give a thick, dark red liquid. In a 500 mL round bottom flask, 200 mL of NH_4OH (25% aq.) was cooled to 0 °C. The crude acid chloride was added dropwise over 20 minutes to the ammonia solution with rapid stirring. After the addition was complete, the reaction was stirred vigorously at

0 °C for 1 hour. The bis-amide product precipitated from the reaction mixture. The dark solid was isolated by filtration and partially dried. The product was purified by triturating in MeOH/H₂O: The solid was suspended in H₂O (100 mL) and MeOH (100 mL) and heated to 60 °C. The mixture was stirred rapidly at 60 °C for 30 minutes. After this time, the mixture was cooled to room temperature. The resulting white solid (a mixture of *meso* and *d/l* diastereomers) was isolated by filtration and washed with MeOH (200 mL). The product (**2.32**) was dried under high vacuum (25.75 g, 50%). IR ($\nu_{\max}$, KBr): 3302, 2946, 2801, 1684, 1418, 1319, 1277, 1250, 1220, 1194, 978, 876. ¹H NMR (400 MHz, DMSO): δ = 1.75-2.08 (4H, m, $\underline{\text{CH}_2\text{CH}_2}$), 4.28-4.36 (2H, m, 2 $\times$ CHBr), 7.30 (2H, s), 7.69 (2H, s) (2 $\times$ NH₂). ¹³C NMR (100 MHz, DMSO): (both diastereomers reported), δ = 32.5, 32.6 ($\underline{\text{CH}_2\text{CHBr}}$), 48.2, 48.5 ($\text{CH}_2\underline{\text{CHBr}}$), 169.87, 169.92 (C=O). LRMS m/z (ESI⁺): 324.9 [M+Na]⁺.

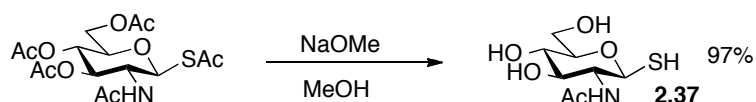
3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio- β -D-glucosamine



(D)-Glucosamine hydrochloride (20.0 g, 92.7 mmol) was added to a 100 mL round bottom flask and suspended in MeOH (20 mL). Sodium methoxide (6.00 g, 111 mmol) was added in several portions and the mixture stirred vigorously at room temperature for 1 hour before NaCl was removed by filtration. The solids were washed with MeOH (20 mL). Acetic anhydride (12 mL, 127 mmol) was added to the stirred filtrate at 0 °C and warmed slowly to room temperature. After 12 hours, the reaction was concentrated to ~20 mL total volume and the product precipitated by the addition of Et₂O (100 mL). The solid was isolated by filtration and dried under vacuum before returning to an oven dried 2-neck 250 mL round bottom flask equipped with a condenser. The flask was placed under and argon atmosphere and cooled to 0 °C. Acetyl chloride (40.0 mL, 563 mmol) was added slowly and the mixture stirred vigorously. After the addition, the reaction was heated to reflux. Gradually the reaction turned black. After 3 hours at reflux, the reaction was cooled to room temperature, diluted with CH₂Cl₂ (200 mL), and poured into a beaker of ice (~200 g). The organic layer was washed with NaHCO₃ (sat. aq., 2 $\times$ 200 mL), dried (MgSO₄), filtered, and concentrated to a black residue. The residue was dissolved in EtOAc and filtered through a short column of silica, eluting with EtOAc. The product was precipitated by the addition of cold Et₂O and petrol. The resulting powder was

isolated by filtration and dried under vacuum (10.19 g, 50% from glucosamine hydrochloride). The crude product (~4:1 mixture of α : β) was used in the final step without purification. Accordingly, a portion of the anomeric chloride (1.00 g, 2.73 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (15 mL). Potassium thioacetate (1.56 g, 13.7 mmol) was then added in one portion. The reaction was flushed with nitrogen and stirred at room temperature for 3 hours. After this time, no starting material was detected by TLC (70% EtOAc in petrol). The reaction mixture was diluted with EtOAc (250 mL) and washed sequentially with NaHCO₃ (sat. aq., 200 mL), H₂O (200 mL), and brine (200 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure to give a brown solid. This solid was recrystallized from CH₂Cl₂ and petrol to afford the product as white crystals (661 mg, 60%). m.p. = 186-188 °C (Lit⁴: 196-197, *i*PrOH). $[\alpha]_D^{20} = +12.1$ ($c = 2.2$, CHCl₃) (Lit⁴: +14.1 $c = 2.16$, CHCl₃). IR ($\nu_{\max}$, KBr): 3733, 3064, 2961, 2360, 2341, 1747, 1707, 1665, 1538, 1433, 1370, 1230, 1048. ¹H NMR (400 MHz, CDCl₃): $\delta = 1.90$ (3H, s), 2.02 (6H, s), 2.05 (3H, s), 2.35 (3H, s) (5 $\times$ Ac), 3.81 (1H, ddd, $J = 9.4, 4.6, 1.3$, H5), 4.08 (1H, dd, $J = 12.4, 1.3$, H6), 4.22 (1H, dd, $J = 12.4, 4.6$, H6'), 4.34 (1H, q, $J = 10.3$, H2), 5.10-5.16 (2H, m, H3, H4), 5.17 (1H, d, $J = 10.3$, H1). ¹³C NMR (100 MHz, CDCl₃): $\delta = 20.49, 20.59, 20.64, 23.0, 30.8$ (5 $\times$ Ac), 51.9 (C2), 61.8 (C6), 67.8, 74.0 (C3, C4), 76.3 (C5), 81.5 (C1), 169.2, 170.0, 170.6, 171.3, 193.5 (5 $\times$ C=O). LRMS (ESI⁺): 428.1 (M+Na)⁺.

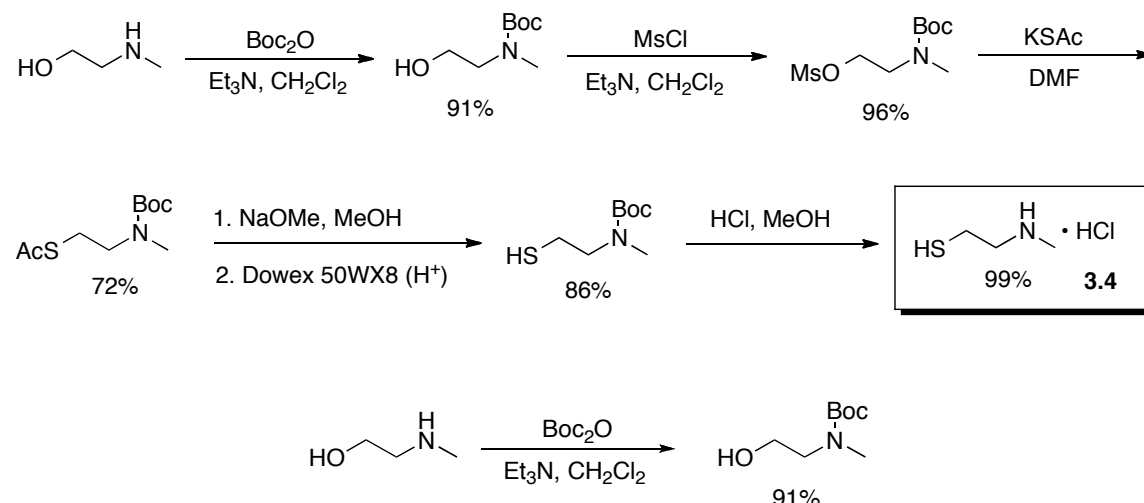
***N*-acetyl-1-thio- β -D-glucosamine (GlcNAc thiol **2.37**)**



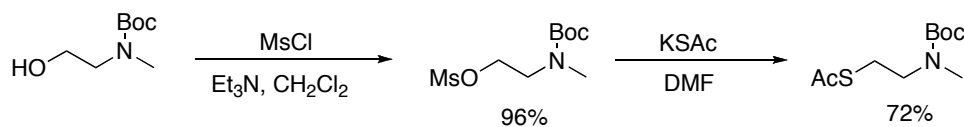
3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio- β -D-glucosamine (250 mg, 0.62 mmol) was added to a 50 mL round bottom flask and placed under an atmosphere of nitrogen before dissolving in anhydrous MeOH (8 mL). Sodium methoxide (67 mg, 1.2 mmol) was added under a stream of nitrogen at room temperature. The reaction was stirred for 1 hour at room temperature before quenching with DOWEX-50WX8 (H⁺ form) until the pH of the reaction mixture was neutral (pH paper). The resin was removed by filtration and washed with MeOH (~ 50 mL). The filtrate was evaporated to provide GlcNAc thiol **2.37** as a thick syrup that crystallized upon standing (142 mg, 97%). m.p. = 170-171 °C (Lit⁵: 174-176 °C). $[\alpha]_D^{20} = -13.5$ ($c = 1.0$, MeOH) (Lit⁵: -10.4, $c = 1.0$, MeOH). IR ($\nu_{\max}$, KBr): 3287, 2360, 2341, 1647, 1558, 1374, 1313, 1056. ¹H NMR (400 MHz, D₂O): $\delta = 1.90$

(3H, s, Ac), 3.30-3.39 (3H, m, H3, H4, H5), 3.56 (1H, dd, $J = 11.1, 5.3$, H6), 3.59 (1H, t, $J = 10.2$, H2), 3.74 (1H, dd, $J = 11.1, 1.2$, H6'), 4.53 (1H, d, $J = 10.2$, H1). ^{13}C NMR (50 MHz, D_2O): $\delta = 22.6$ (Ac), 58.2 (C2), 61.1 (C6), 70.0, 75.3, 79.3, 80.52 (C3, C4, C5, C1). LRMS (ESI $^-$): 236.1 (M-H) $^-$. Note: in D_2O this material anomerizes to a 3.6:1.0 mixture of β : α over a period of 24 hours.

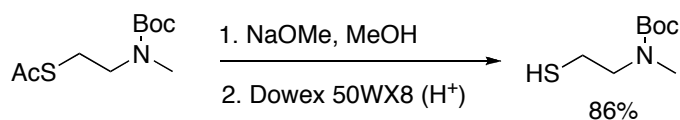
Synthetic Scheme for Thiol Precursor to Monomethyl Lysine Analog (3.4)



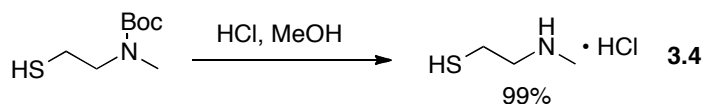
N-methylethanolamine (4.13 g, 55.0 mmol) was added to a 250 mL round bottom flask and dissolved in CH_2Cl_2 (100 mL). Et_3N (9.6 mL, 69 mmol) was added and the stirred solution was cooled to 0 °C. Boc_2O (10.00 g, 45.82 mmol) was added and the reaction was stirred for 5 min at 0 °C and then 4 hours at room temperature. After this time, the reaction was cooled to 0 °C and quenched with 1M HCl (200 mL). The mixture was diluted with CH_2Cl_2 (100 mL) and the organic layer was separated. After washing with 1M HCl (2×100 mL), the organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting oil (7.30 g, 91%) did not require further purification. Spectroscopic data was consistent with that previously reported.⁶ IR (ν_{max} , film): 3441, 2977, 1681, 1484, 1396, 1305, 1226, 1156. ^1H NMR (CDCl_3 , 400 MHz): $\delta = 1.47$ (9H, s, Boc), 2.93 (3H, s, NMe), 3.40 (2H, t, $J = 4.8$, $\text{CH}_2\text{CH}_2\text{OH}$), 3.74-3.78 (2H, m, CH_2OH). ^{13}C NMR (CDCl_3 , 50 MHz): $\delta = 28.3$ (Boc), 35.3 (NMe), 51.2 ($\text{CH}_2\text{CH}_2\text{OH}$), 60.9 (CH_2OH), 79.7 (Boc), 157.0 (C=O). LRMS m/z (ESI $^+$): 198.1 $[\text{M}+\text{Na}]^+$.



2-[*N*-(*tert*-Butyloxycarbonyl)-*N*-methylamino]ethanol (7.21 g, 41.2 mmol) was added to a 250 mL round bottom flask and dissolved in 85 mL CH₂Cl₂. Et₃N (7.5 mL, 53.5 mmol) was added and the stirred solution was cooled to 0 °C. Methanesulfonyl chloride (3.34 mL, 43.2 mmol) was added dropwise and 10 mL CH₂Cl₂ was used to rinse the flask on completion of the addition. The flask was stoppered and stirred at 0 °C for 1 hour. After this time, the reaction was quenched at 0 °C with 1M HCl (100 mL). After diluting with 50 mL CH₂Cl₂, the organic layer was separated and then washed with 1M HCl (2 × 100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The yellow oil obtained was sufficiently pure to not require further purification (9.96 g, 96%). However, this material is not stable and was used immediately in the next transformation. Accordingly, the product mesylate (9.56 g, 37.8 mmol) was added to a 250 mL round bottom flask and dissolved in DMF (100 mL). Potassium thioacetate (6.47 g, 56.6 mmol) was added and the reaction was stirred vigorously as the reaction became viscous. The flask was stoppered and the reaction heated to 60 °C. After 1 hour at 60 °C, the reaction was cooled to room temperature. TLC (50% EtOAc in petrol) revealed complete consumption of starting material (*R*_f = 0.3) and the formation of thioester product (*R*_f = 0.7). The reaction mixture was diluted with Et₂O (200 mL) and washed with H₂O (3 × 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (20% EtOAc in petrol) to provide the product thioester as a yellow oil (6.32 g, 72%). IR (ν_{max}, film): 2976, 2932, 1698, 1480, 1457, 1425, 1392, 1365, 1301, 1236, 1214. ¹H NMR (CDCl₃, 200 MHz): δ = 1.40 (9H, s, Boc), 2.28 (3H, s) 2.84 (3H, s) (NMe and SAc), 2.95 (2H, t, *J* = 7.1), 3.30 (2H, t, *J* = 7.1) (CH₂CH₂). ¹³C NMR (CDCl₃, 50 MHz): δ = 27.0 (CH₂SAc), 28.2 (Boc), 30.5 (SAc), 34.6 (NMe), 48.1 (CH₂NMe), 79.5 (Boc), 155.4, 195.1 (2 × C=O). HRMS *m/z* (ESI⁺): Found 256.0977 [M+Na]⁺; C₁₀H₁₉NO₃SNa requires 256.0978.

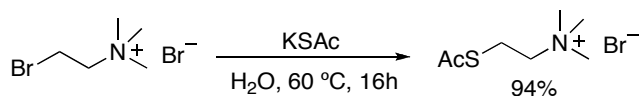
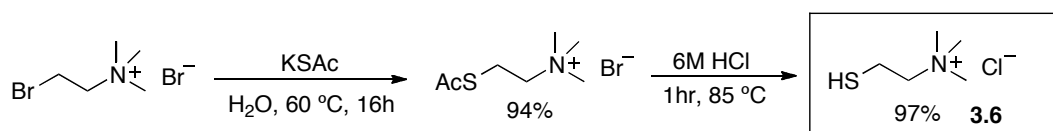


2-[*N*-(*tert*-Butyloxycarbonyl)-*N*-methyamino]acetylethanthiol (500 mg, 2.14 mmol) was added to a 25 mL round bottom flask and dissolved in MeOH (5 mL). The flask was sealed and placed under a nitrogen atmosphere. NaOMe (173 mg, 3.21 mmol) was added and the reaction was stirred at room temperature for 30 minutes. After this time the reaction was quenched with Dowex-50WX8 (H⁺) until the pH was neutral (pH paper). The resin was removed by filtration and washed with MeOH (25 mL). The filtrate was concentrated under reduced pressure and the resulting residue was purified by column chromatography (20% EtOAc in petrol). The product thiol was isolated as a yellow oil (353 mg, 86%). IR (ν_{max} , film): 3004, 2976, 2554, 2360, 2341, 1697, 1480, 1457, 1422, 1392, 1365, 1301, 1241, 1214, 1169. ¹H NMR (CDCl₃, 200 MHz): δ = 1.25 (1H, t, *J* = 8.3, SH), 1.38 (9H, s, Boc), 2.53-2.64 (2H, m, CH₂SH), 2.82 (3H, s, NMe), 3.30 (2H, t, *J* = 7.3, CH₂CH₂SH). ¹³C NMR (CDCl₃, 50 MHz): δ = 22.4 (CH₂SH), 28.2 (Boc), 34.6 (NMe), 52.0 (CH₂CH₂SH), 79.4 (Boc), 155.4 (C=O). HRMS *m/z* (ESI⁺): Found 214.0870 [M+Na]⁺; C₈H₁₇NO₂SNa requires 214.0872.

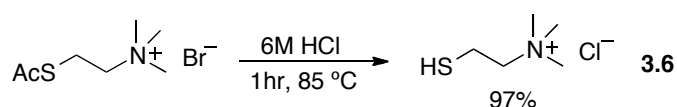


A 50 mL round bottom flask was flame dried and placed under an atmosphere of nitrogen. MeOH (4 mL) was added and the flask was cooled to 0 °C. Acetyl chloride (1.7 mL, 23.7 mmol) was added and the reaction was stirred at 0 °C for 10 min to generate a solution of HCl in MeOH. 2-[*N*-(*tert*-Butyloxycarbonyl)-*N*-methyamino]ethanthiol (302 mg, 1.58 mmol) was added to the reaction flask as a solution in 1 mL MeOH. An additional 1 mL MeOH was used to complete the transfer. The ice bath was removed and the reaction was stirred for 30 minutes, after which time TLC (20% EtOAc in petrol) revealed complete consumption of starting material. The solvent was removed under reduced pressure to give the product as an extremely hygroscopic residue (200 mg, 99%). ¹H NMR (CDCl₃, 200 MHz): δ = 2.77 (3H, s, NMe), 2.85 (2H, t, *J* = 6.7), 3.24 (2H, t, *J* = 6.7) (CH₂CH₂). ¹³C NMR (CDCl₃, 50 MHz): δ = 21.1 (CH₂SH), 33.7 (NMe), 52.8 (CH₂CH₂SH).

Synthetic Scheme for Thiol Precursor to Trimethyl Lysine Analog (3.6)



2-(Bromoethyl)-trimethylammonium bromide (5.00 g, 20.2 mmol) was added to a 250 mL round bottom flask and dissolved in H₂O (25 mL). Potassium thioacetate (3.01 g, 26.3 mmol) was added to the stirred solution and the resulting mixture was heated to 60 °C. The flask was stoppered and stirred at 60 °C for 16 hours. After this time, the reaction mixture was concentrated under reduced pressure to give a brown solid. The product was extracted by stirring the solid in 100 mL of MeOH:CH₂Cl₂ (1:1) at room temperature for 15 minutes. KBr was removed by filtration and the filtrate was concentrated under reduced pressure. The extraction and filtration was repeated twice more to ensure removal of KBr. After concentrating under reduced pressure, the resulting reddish solid was sufficiently pure to carry on to the next step (4.59 g, 94%). IR (ν_{max} , KBr): 1694, 1563, 1484, 1420, 1355, 1267, 1200, 1132, 1107, 1059, 1024, 1001, 971. ¹H NMR (D₂O, 400 MHz): δ = 2.31 (3H, s, SAc), 3.09 (9H, s, NMe₃), 3.17-3.21 (2H, m), 3.35-3.39 (2H, m) (CH₂CH₂). ¹³C NMR (MeOD, 50 MHz): δ = 23.1, 30.6, 53.9, 66.3, 196.2. LRMS m/z (ESI⁺): 162.1 [M]⁺.



2-(acetothioethyl)-trimethylammonium bromide (1.18 g, 4.88 mmol) was added to a 100 mL round bottom flask and dissolved in 6M HCl (10 mL). The stirred solution was stoppered and heated to 85 °C. After 1 hour, the reaction was concentrated under reduced pressure and dried under high vacuum to provide the product thiol as a pale yellow, very hygroscopic solid (740 mg, 97%). This material was used without further purification. ¹H NMR (D₂O, 400 MHz): δ = 2.82-2.86 (2H, m, CH₂), 3.04 (9H, s, NMe₃), 3.42-3.46 (2H, m, CH₂). ¹³C NMR (MeOD, 50 MHz): δ = 18.1, 53.9, 69.5. LRMS m/z (ESI⁺): 120.1 [M]⁺.

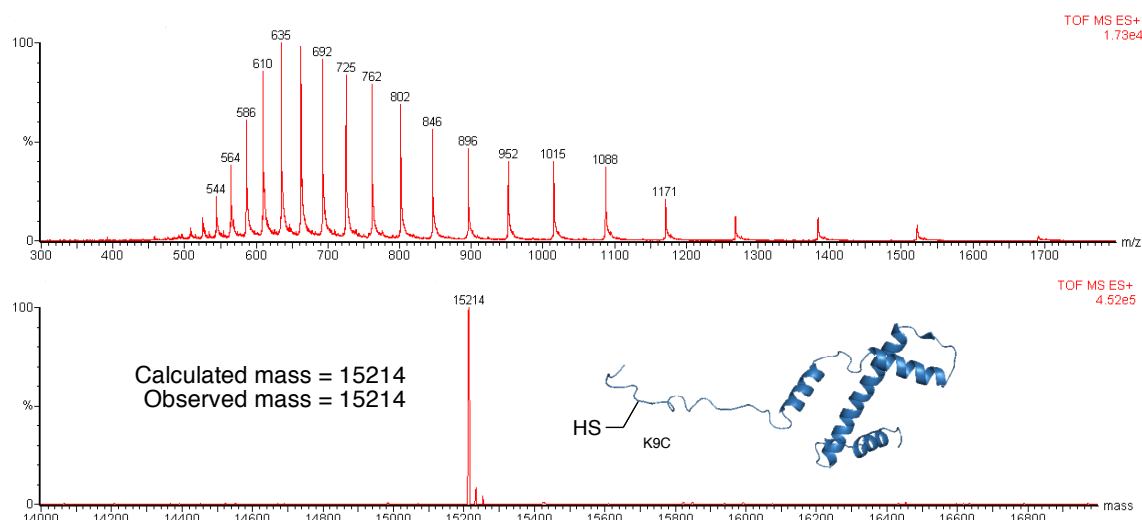
Histones used in this study

All histones were expressed and purified by Dr. Nathan Rose of the Prof. Christopher J. Schofield research group at the University of Oxford Department of Chemistry. Dr. Rob Klose (Department of Biochemistry, University of Oxford) is gratefully acknowledged for donation of the plasmid encoding the H3 (K9C, C110A) mutant. Amino acid sequences are shown below with point mutations highlighted. Calculated average molecular mass and ESI-MS are also shown.

Histone H3 (K9C, C110A) (4.1)

ARTKQTAR**C**STGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKST
ELLIRKLFPQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVALFEDTNLA**A**AIHAKRVTI
MPKDIQLARRIRGERA

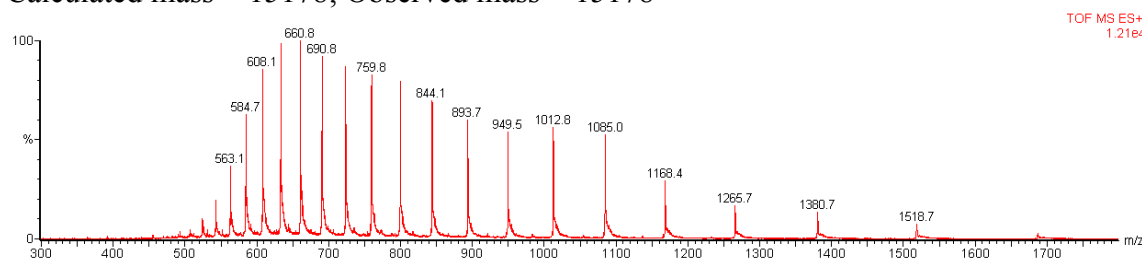
Calculated mass = 15214; Observed mass = 15214

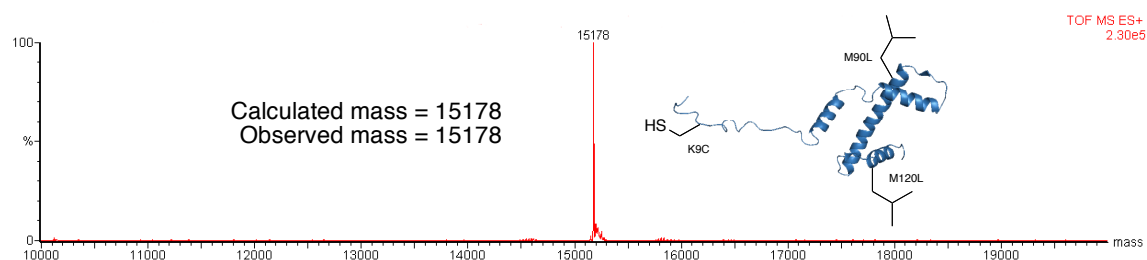


Histone H3 (K9C, M90L, C110A, M120L) (4.3)

ARTKQTAR**C**STGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKST
ELLIRKLFPQRL**V**REIAQDFKTDLRFQSSAVLALQEASEAYLVALFEDTNLA**A**AIHAKRVTI**L**
PKDIQLARRIRGERA

Calculated mass = 15178; Observed mass = 15178



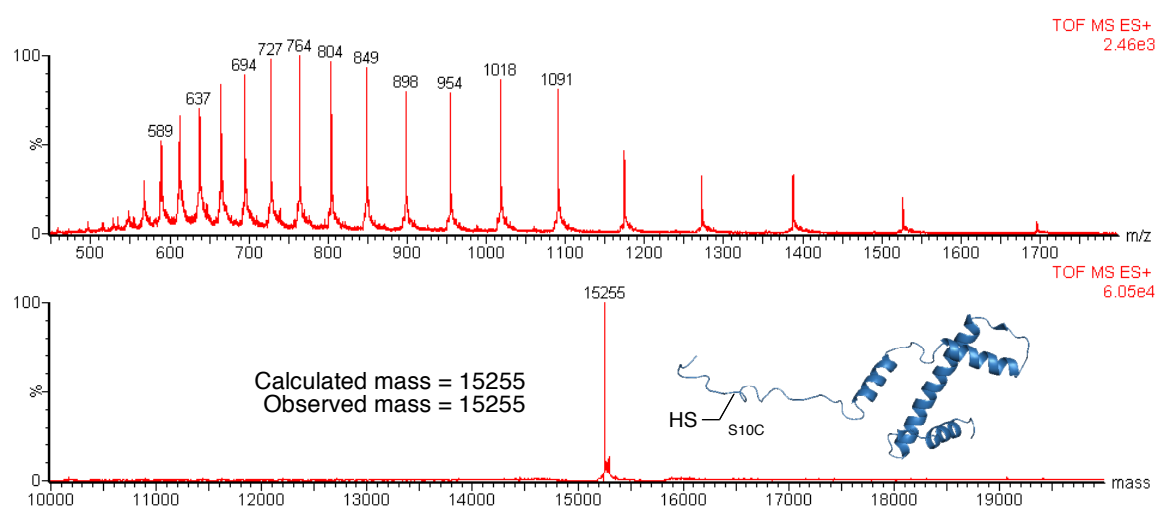


Histone H3 (S10C, C110A) (4.12)

ARTKQTARKCTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKST
ELLIRKLPFQRLVREIAQDFKTDLRFSAVMALQEASEAYLVALFEDTNLAIIHAKRVTI
MPKDIQLARRIRGERA

Calculated mass = 15255

Observed mass = 15255



Mass Spectrometry:

Correlation of Relative Peak Height to Relative Amount of Protein

HPLC conditions and mass spectrometer settings were identical to those in Chapter 2.

The correlation of peak height of mass spectra to relative amount of protein was validated using the two histone mutants shown below. Briefly, the protein concentrations were measured using the bicinchoninic acid assay (Pierce) and then the proteins were combined in known amounts and the mixture analyzed by LC-MS. The amount of each protein was then correlated to the relative peak height in the corresponding mass spectrum of the mixture.

Histone mutants used in this experiment:

Protein A: H3 (K9C, C110A): Calc. Mass = 15214

Protein B: H3 (S10C, C110A): Calc. Mass = 15255

The two histone mutants were prepared at 1.0 mg/mL in H₂O. Concentrations were measured using the bicinchoninic acid assay (BSA assay, Pierce) according to the manufacturer's instructions. A 250 µL aliquot of each protein was reduced by the addition of DTT (2 mg in 250 µL pH 8.0 sodium phosphate buffer). Both samples were incubated at room temperature for 10 minutes to ensure reduction. After this incubation the proteins were mixed in the following relative amounts:

Protein A : Protein B

5 : 0

4 : 1

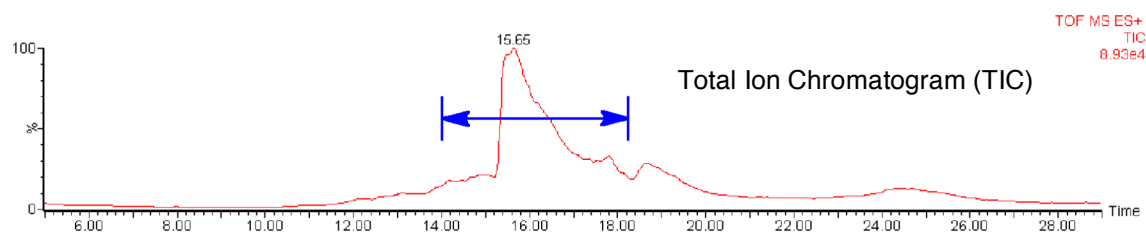
3 : 2

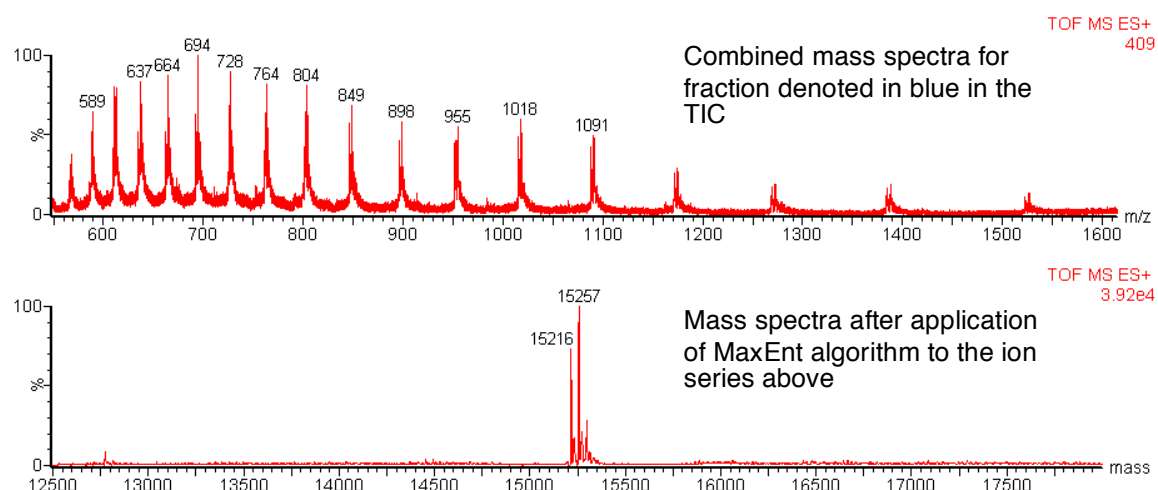
2 : 3

1 : 4

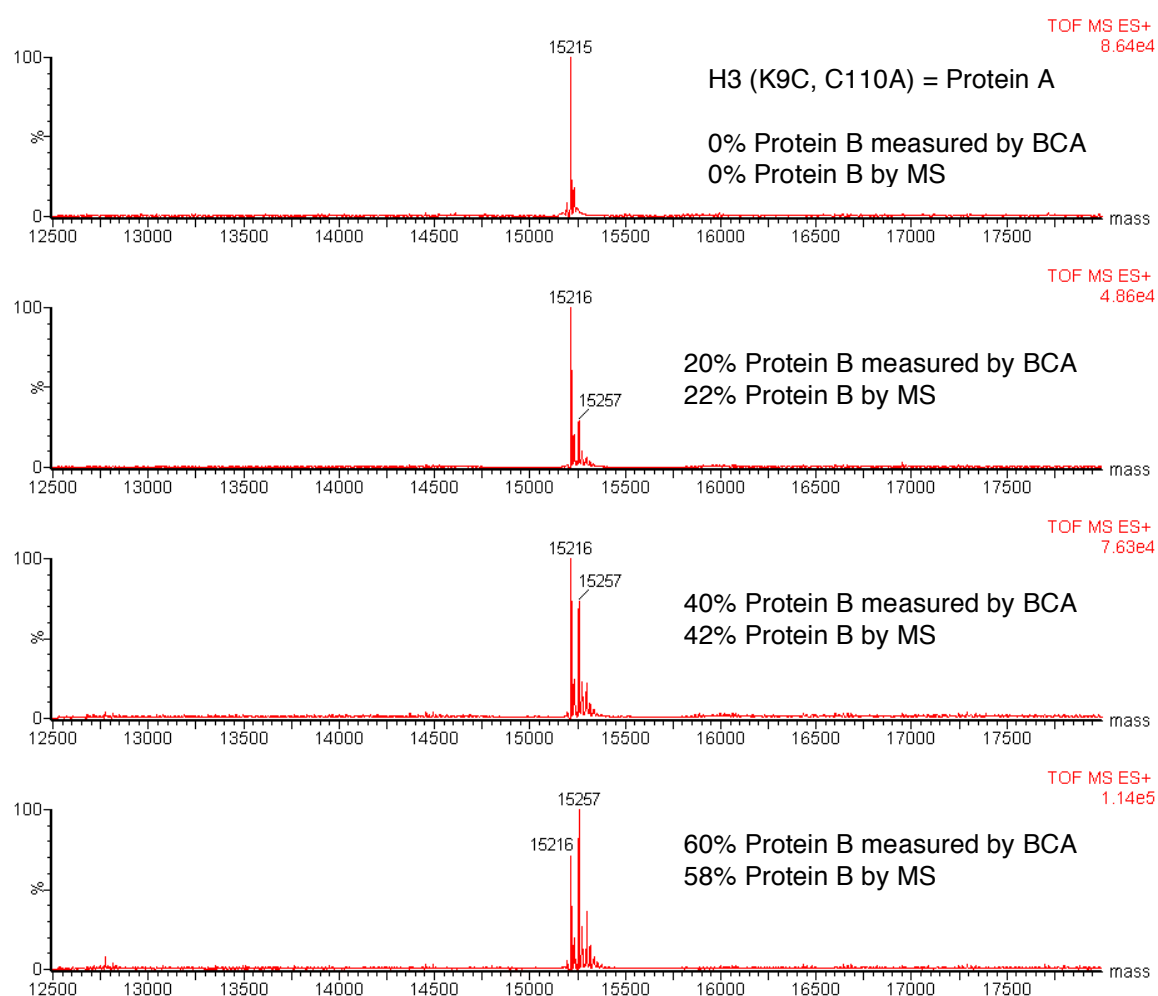
0 : 5

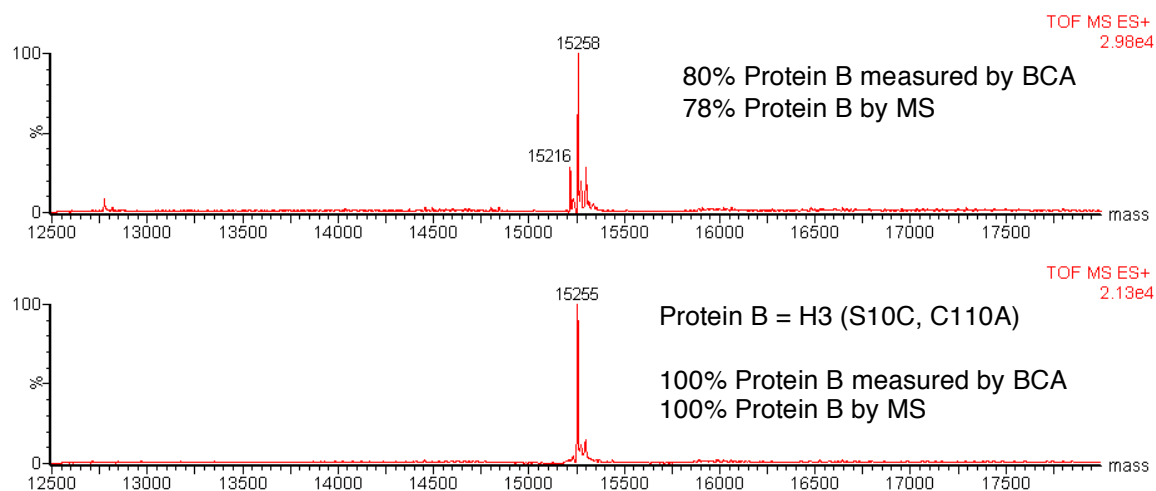
A 25 µL injection was used for each LC-MS run. The HPLC and mass spectrometer settings were identical to those used for reaction analysis (see General Considerations in Chapter 2). Both proteins co-elute and all protein material typically elutes between 14 and 18.5 minutes for the chromatographic conditions outlined in the General Considerations. The mass spectra collected over this time period were combined and the resulting ion series obtained was processed to convergence using the Maximum Entropy algorithm. The relative peak height obtained after MaxEnt processing was then correlated to the relative measured amount that was measured by the BCA assay. An outline of the MS processing for a typical total ion chromatogram is shown below.



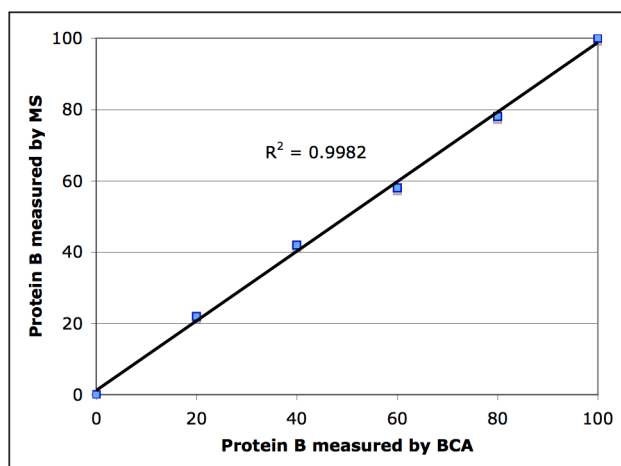


The mass spectra after MaxEnt analysis for the mixtures of protein A and protein B:



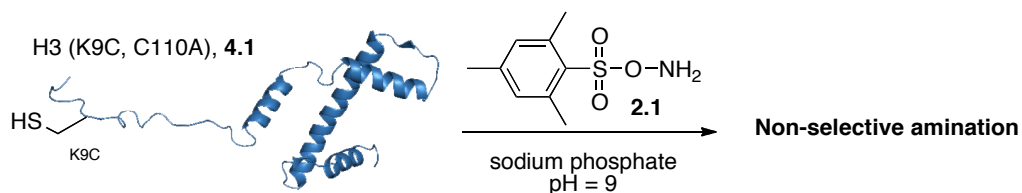


Correlation of protein amount by BCA to relative peak height in ESI-MS



Histone Reactions using MSH (2.1)

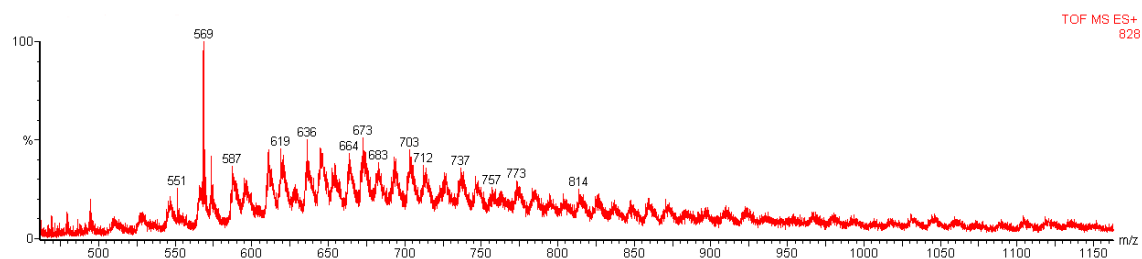
Reaction of H3 (K9C, C110A) and MSH at pH 9.0



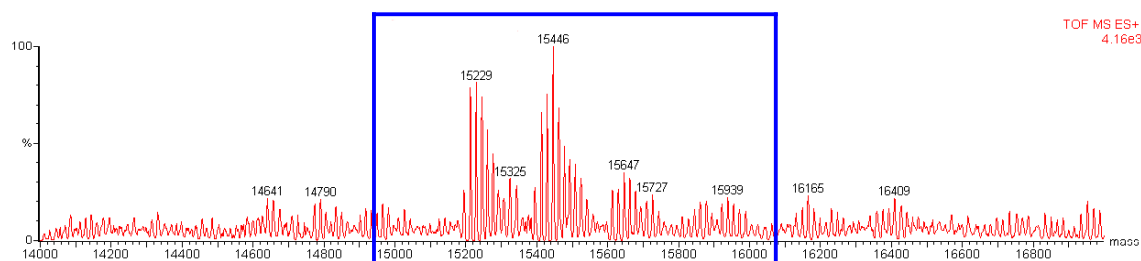
All manipulations were carried out at 4 °C. A 250 μ L aliquot of H3 (K9C, C110A; 0.4 mg/mL in H₂O; $\sim$ 0.007 μ mol) was added to a 1.0 mL plastic tube and stored on ice until needed. Sodium phosphate dibasic dihydrate (17 mg) was added to bring the pH to $\sim$ 9. A stock solution of MSH (2.1) was prepared by dissolving 0.7 mg in 17.5 μ L of DMF. A 10 μ L aliquot of the MSH solution (1.9 μ mol) was added to the protein solution and the

reaction vortexed to mix and then shaken at 4 °C for 20 minutes. The reaction mixture was analyzed by LC-MS, revealing the formation of several adducts from reaction with MSH. This outcome was consistent with non-selective amination described in Chapter 2 (ESI-MS shown below).

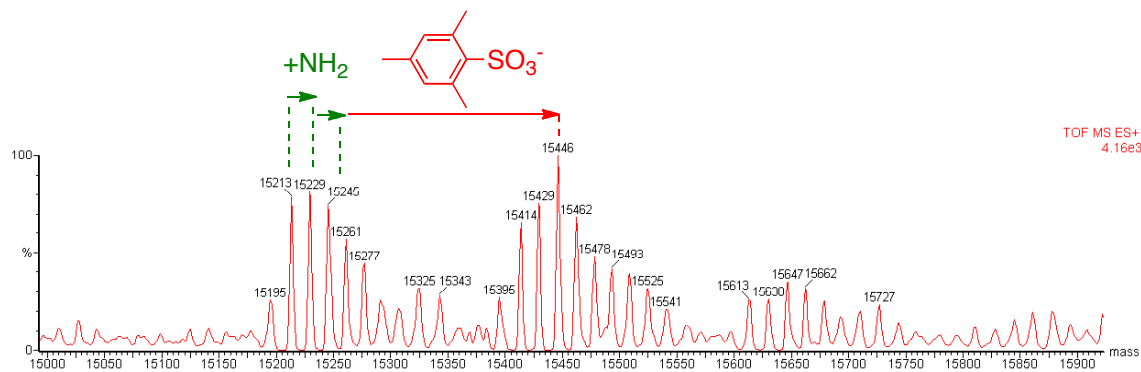
Ion series:



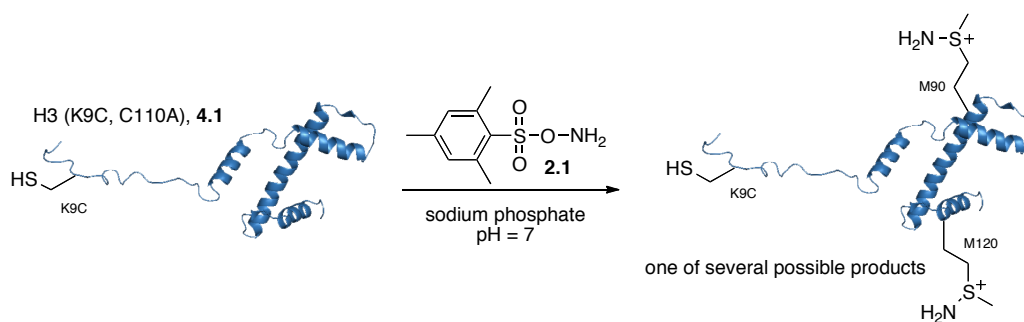
Deconvoluted MS:



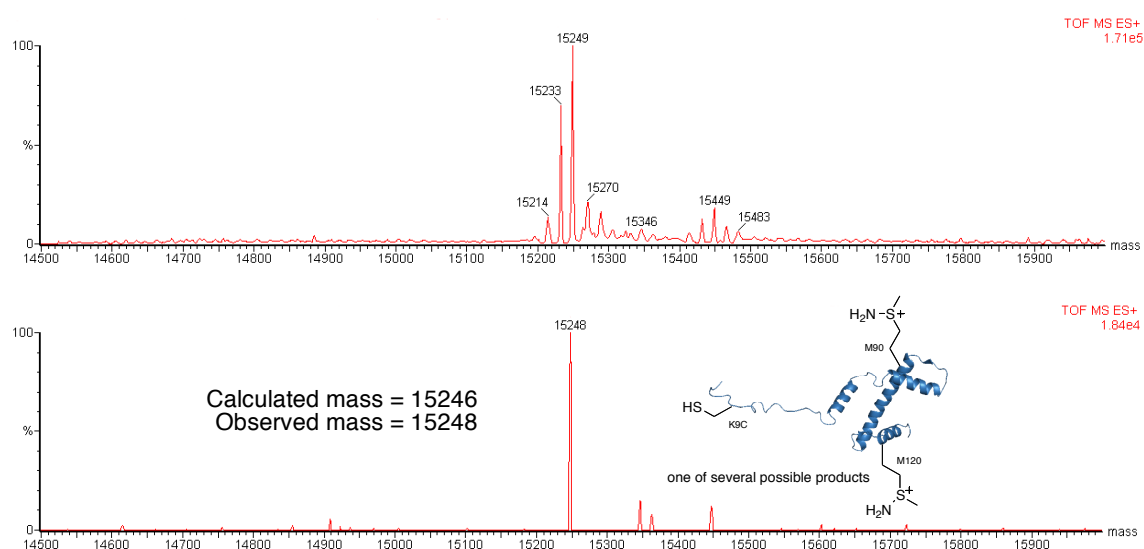
Blue inset from MS shown above:



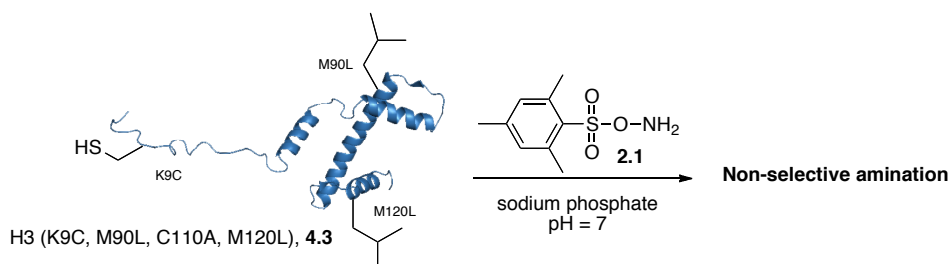
Reaction of H3 (K9C, C110A) and MSH at pH 7.0



All manipulations were carried out at 4 °C. A 200 μ L aliquot of H3 (K9C, C110A; 0.4 mg/mL; $\sim$ 0.005 μ mol) in pH 7.0 sodium phosphate (50 mM) was added to a 1.0 mL plastic tube and stored on ice. A stock solution of MSH (**2.1**) was prepared by dissolving 1.8 mg in 600 μ L of DMF. A 6 μ L aliquot of the MSH solution (0.08 μ mol) was added to the protein solution and the reaction vortexed and then shaken for 25 minutes. LC-MS analysis of the reaction mixture revealed a mixture of adducts (Mass = 15214, 15233, and 15249). An additional 4 μ L of the MSH solution was added and the reaction shaken for 15 minutes. LC-MS analysis revealed full conversion to a species of a mass of 15248. One possible product with this mass is the double sulfilimine adduct shown above. ESI-MS are shown below.

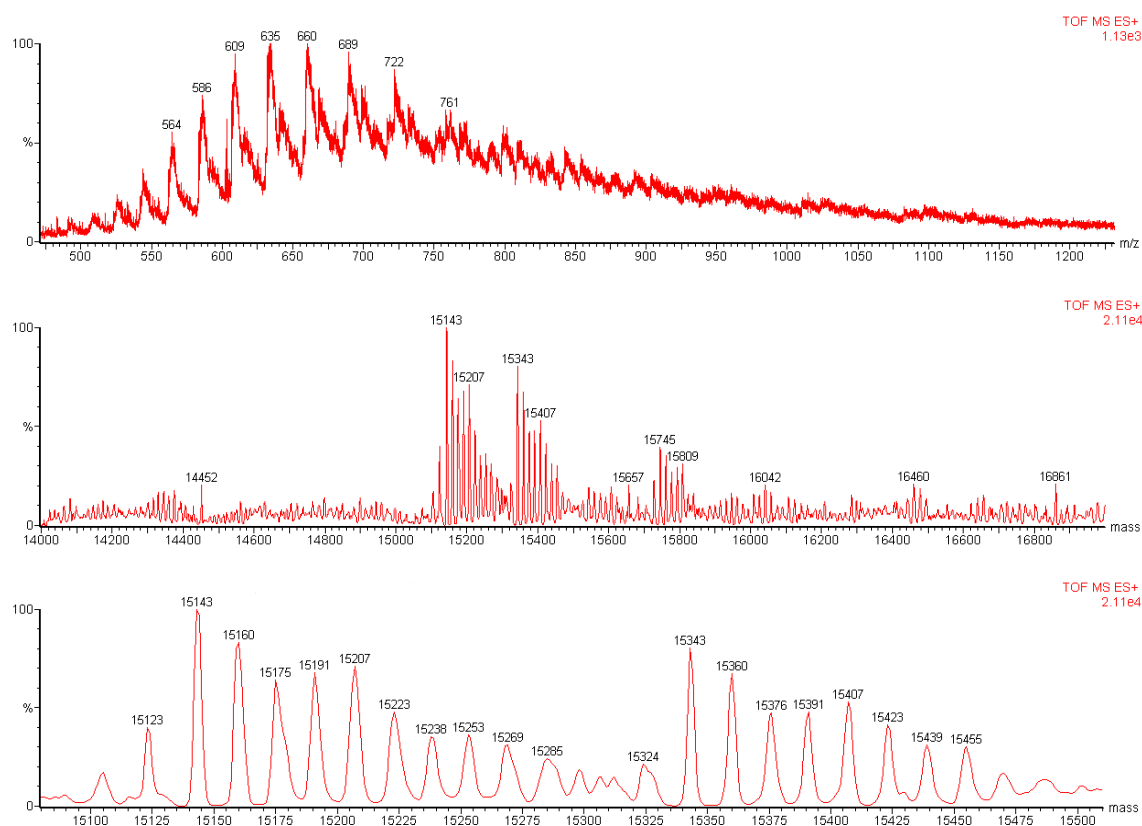


Reaction of methionine-free H3 (K9C, M90L, C110A, M120L) and MSH at pH 7.0

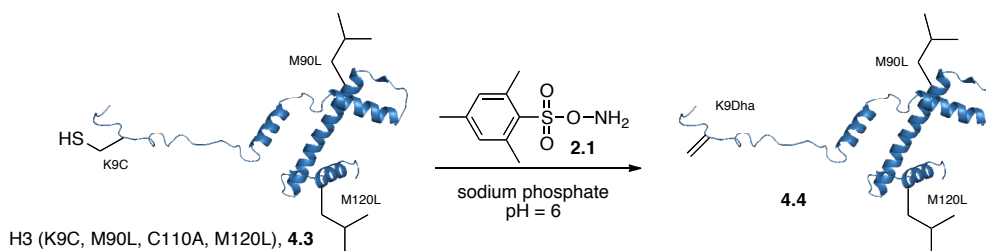


All manipulations were carried out at 4 °C. A 250 μ L aliquot of H3 (K9C, M90L, C110A, M120L; 0.3 mg/mL in H₂O; $\sim$ 0.005 μ mol) was adjusted to $\sim$ pH 7 by the addition of 8 mg of Na₂HPO₄•H₂O and 15 mg of NaH₂PO₄•H₂O. A stock solution of MSH was prepared by dissolving 1.3 mg in 118 μ L DMF. 5 μ L of the MSH solution (2.6 μ mol) was added to the protein sample and shaken for 10 minutes. LC-MS analysis revealed a

mixture of MSH adducts along with peak that is the mass of the Dha product (15144 calculated, 15143 found). Fewer equivalents of MSH (down to 0.04 μmol) led to similar MSH adducts. ESI-MS are shown below.

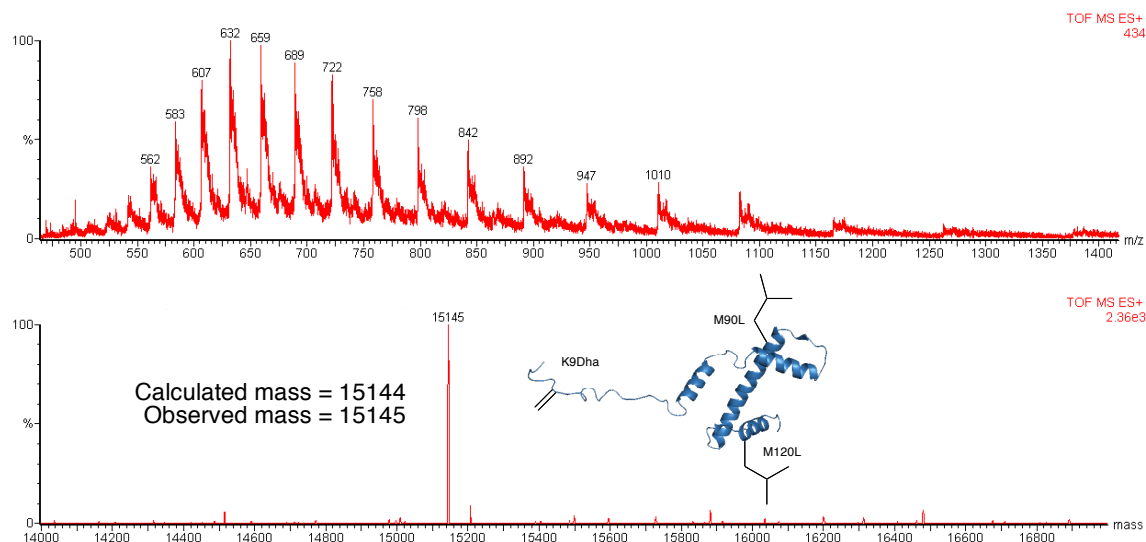


Reaction of methionine-free H3 (K9C, M90L, C110A, M120L) and MSH at pH 6.0

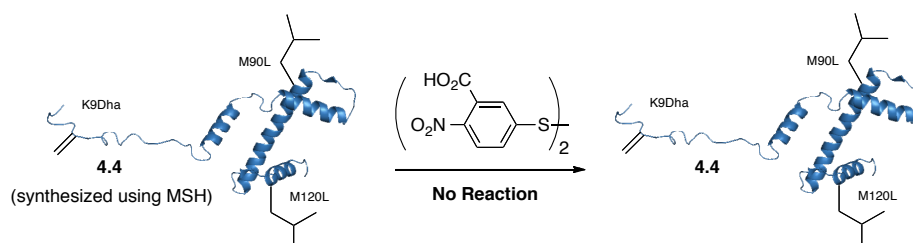


All manipulations were carried out at 4 °C. To a 2.5 mL aliquot of H3 (K9C, M90L, C110A, M120L; 0.3 mg/mL in H_2O ; $\sim 0.05 \mu\text{mol}$) was added 150 mg of $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ and 50 mg of $\text{Na}_2\text{HPO}_4 \cdot \text{H}_2\text{O}$. The mixture was vortexed to dissolve. The final pH was 6.0 (pH meter). A stock solution of MSH (**2.1**) was prepared by dissolving 1.3 mg in 491 μL DMF. A 200 μL aliquot of the MSH solution (2.5 μmol) was added to the protein solution and vortexed immediately upon addition to ensure efficient mixing. The reaction was shaken for 20 minutes and then analyzed by LC-MS. A protein with a mass consistent with the formation of dehydroalanine was observed (calculated mass = 15144;

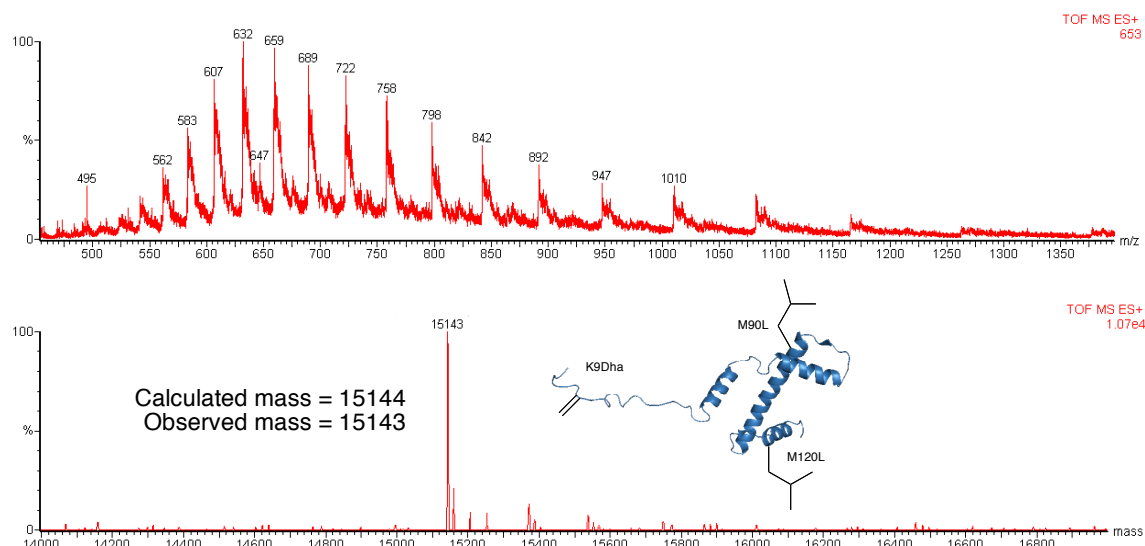
observed mass = 15145). The protein was purified by passing the reaction mixture through a PD10 column (GE Healthcare) previously equilibrated with H₂O. The final concentration of the protein was 0.21 mg/mL. The protein solution was split into 250 µL aliquots and flash frozen. ESI-MS are shown below.



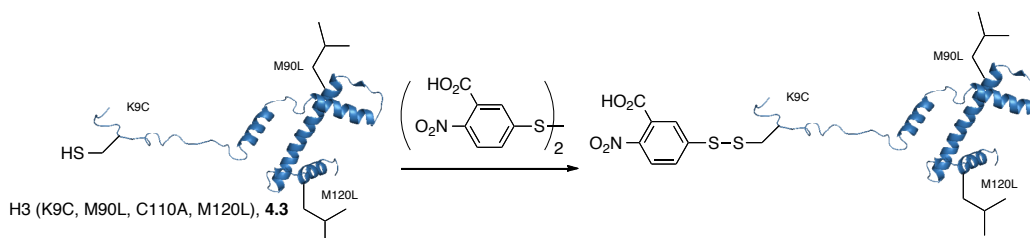
Ellman's test to confirm reaction at cysteine



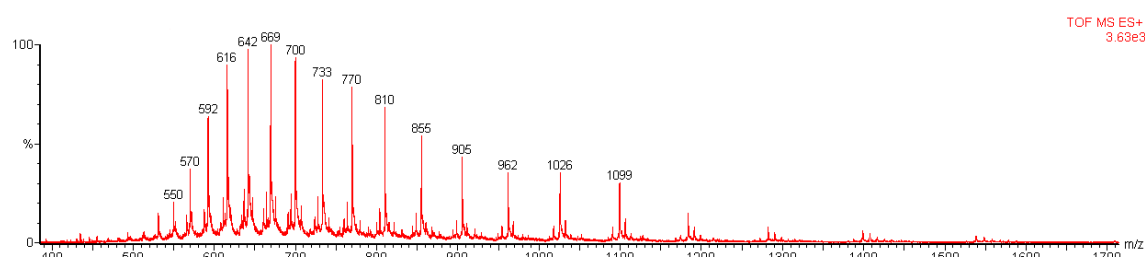
All manipulations were carried out at 4 °C. To a 200 µL aliquot of H3 (K9Dha, M90L, C110A, M120L; 0.21 mg/mL in H₂O; 0.003 µmol) was added 5 mg of Na₂HPO₄. A stock solution of Ellman's reagent was prepared by dissolving 3 mg in 273 µL of pH 8.0 sodium phosphate (50 mM). A 2.5 µL aliquot of the Ellman's solution (0.07 µmol) was added to the protein sample. The reaction was vortexed and then shaken for 15 minutes. LC-MS analysis revealed no reaction, indicating that all free cysteine was consumed during its conversion to Dha (calculated mass = 15144; observed mass = 15143). ESI-MS are shown below. Under the same reaction conditions, H3 (K9C, M90L, C110A, M120L) is converted to the disulfide (see next).

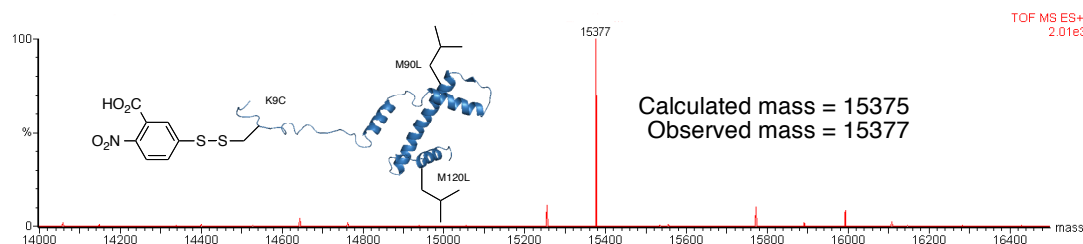


Control: H3 (K9C, M90L, C110A, M120L) and Ellman's reagent.

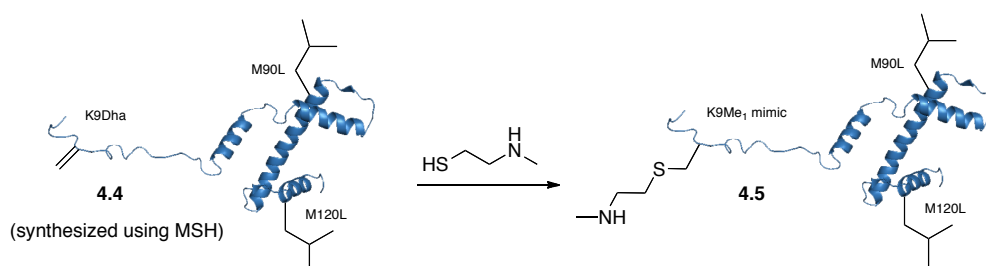


All manipulations were carried out at 4 °C in the same manner as for the K9Dha sample described above. Accordingly, to a 200 μL aliquot of H3 (K9C, M90L, C110A, M120L; 0.21 mg/mL in H_2O ; 0.003 μmol) was added 5 mg of Na_2HPO_4 . A stock solution of Ellman's reagent was prepared by dissolving 3 mg in 273 μL of pH 8.0 sodium phosphate (50 mM). A 2.5 μL aliquot of the Ellman's solution (0.07 μmol) was added to the protein sample. The reaction was vortexed and then shaken for 15 minutes. LC-MS analysis revealed full conversion to the Ellman disulfide (calculated mass = 15375; observed mass = 15377). ESI-MS are shown below.

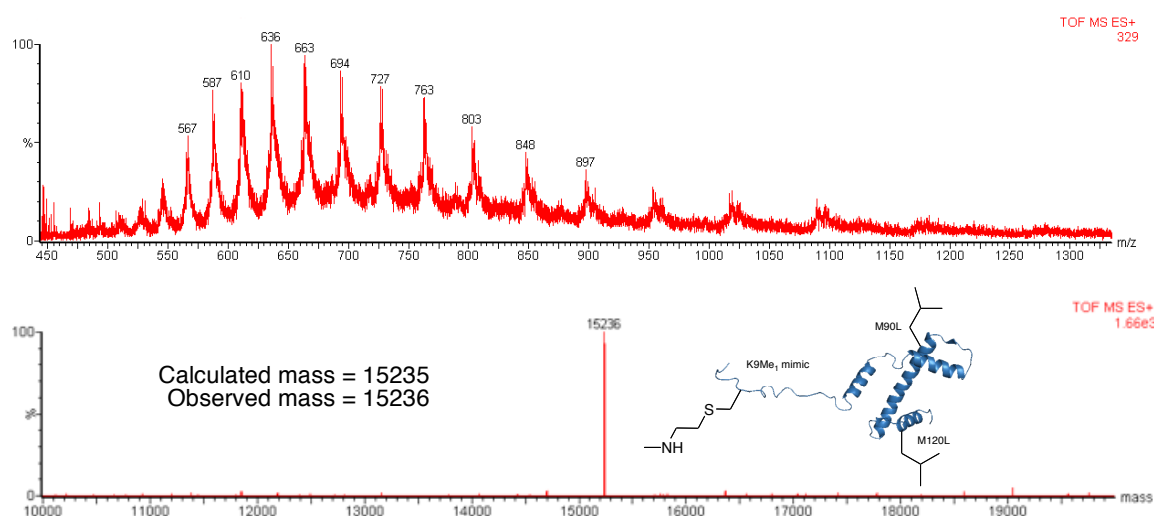




Chemical test for dehydroalanine: thiol addition to dehydroalanine and synthesis on a monomethyl lysine analog at position 9 of H3.



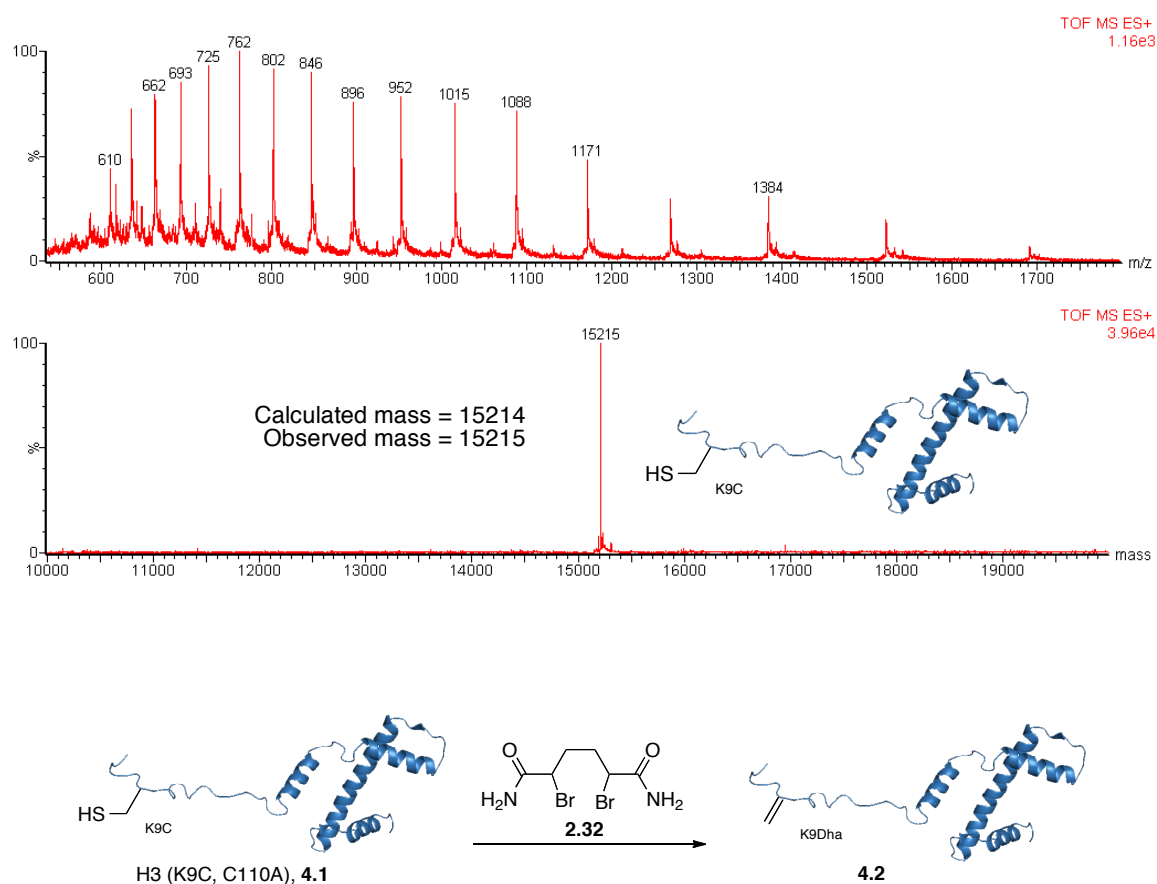
To a 200 μ L aliquot of H3 (K9Dha, M90L, C110A, M120L; 0.21 mg/mL in H₂O; 0.003 μ mol) was added 10 mg of Na₂HPO₄. The mixture was vortexed to dissolve and then stored on ice. A stock solution of 2-(methylamino)ethanthiol hydrochloride (**3.4**) was prepared by dissolving 10 mg in 520 μ L of H₂O. A 50 μ L aliquot of the thiol solution (7.5 μ mol) was added to the protein. The reaction was vortexed and then rotated at room temperature for 25 minutes. LC-MS of the reaction mixture revealed full conversion to the monomethyl lysine mimic (calculated mass = 15235; observed mass = 15236). ESI-MS are shown below.



Histone modification using di-bromide **2.32**

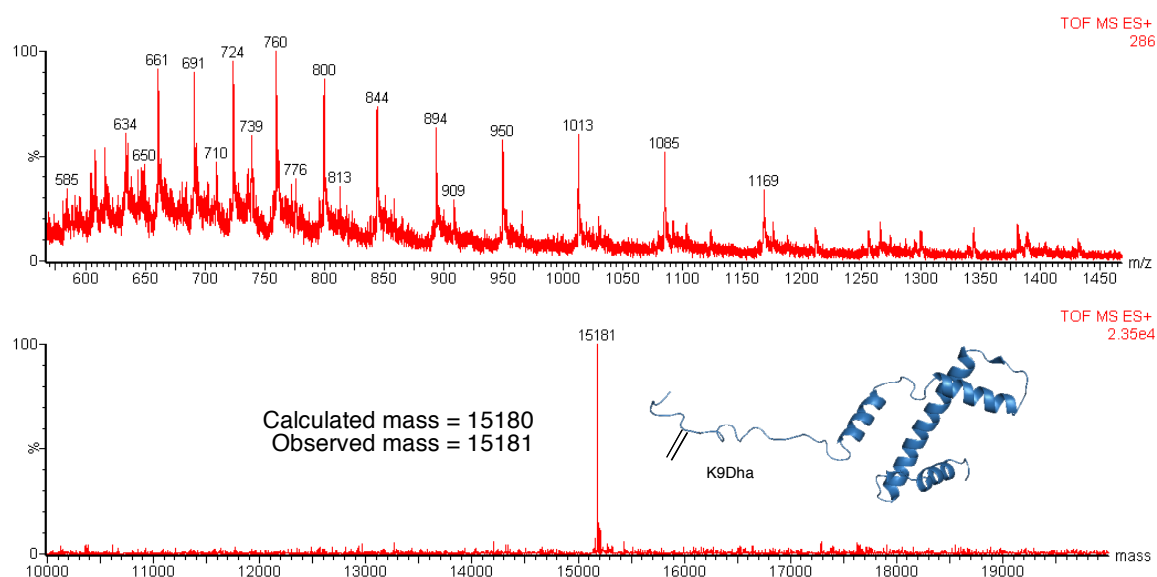
H3 (K9C, C110A) (**4.1**) was used for the following reactions. This mutant contains two natural methionine residues.

H3 (K9C, C110A) (3.3 mg, lyophilized powder) was dissolved in 250 μ L water (MilliQ). Sodium phosphate buffer (250 μ L, pH 8.0, 50 mM) was added to the protein solution followed by dithiothreitol (DTT, 17 mg). The reaction was shaken at room temperature for 15 minutes to reduce any contaminant disulfide. To remove the DTT, the protein solution was passed through a PD minitrapp (GE Healthcare), eluting with 1.0 mL H₂O. The reduced protein was collected kept on ice ($c = 3.3$ mg/mL). ESI-MS are shown below (calculated mass = 15214; observed mass = 15215).



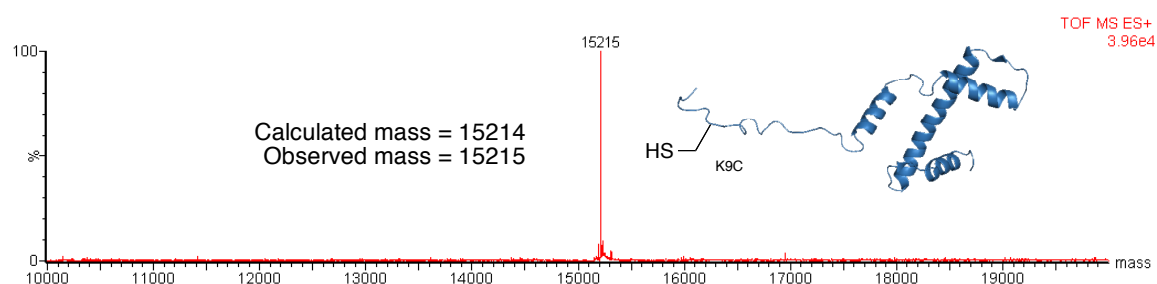
H3 (K9C, C110A) (reduced with DTT and passed through a PD minitrapp, see above) was diluted with H₂O to a concentration of 1.0 mg/mL. A 250 μ L aliquot of this protein solution (~ 16 nmol) was further diluted with 250 μ L of sodium phosphate buffer (pH 8.0, 50 mM). The protein solution was kept on ice until needed. A stock solution of α,α' -di-bromoadipyl(bis)amide (**2.32**) was prepared by dissolving 16.9 mg in 128 μ L of DMF. The protein solution was warmed to room temperature and then a 56 μ L aliquot of the di-

bromide solution (25 μmol) was added. The reaction was vortexed immediately to give a clear solution. The reaction was shaken at room temperature for 30 minutes and then 1 hour at 37 $^{\circ}\text{C}$. After this time, the reaction was cooled to room temperature and analyzed by LC-MS. Complete conversion to dehydroalanine was observed (calculated mass = 15180; observed mass = 15181). (Note: some of di-bromide **2.32** crystallizes over the course of the reaction and can be removed by centrifugation). Small molecules were removed by passing the reaction mixture through a PD mini-trap, previously equilibrated with H_2O . The protein was eluted with H_2O and kept on ice. ESI-MS are shown below.

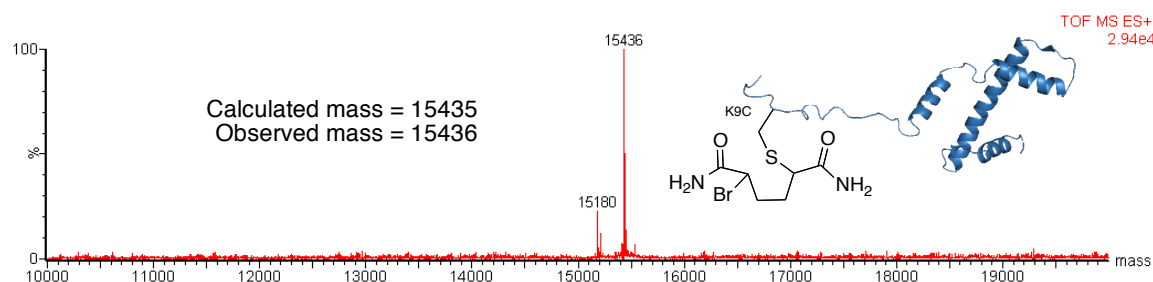


Note that after the 30 minute incubation at room temperature, the main product is alkylated, uncyclized cysteine—consistent with observations on other proteins in Chapters 2 and 3. Incubation at 37 $^{\circ}\text{C}$ promotes cyclization and elimination to dehydroalanine. ESI-MS for this intermediate (and the entire reaction sequence) is shown below.

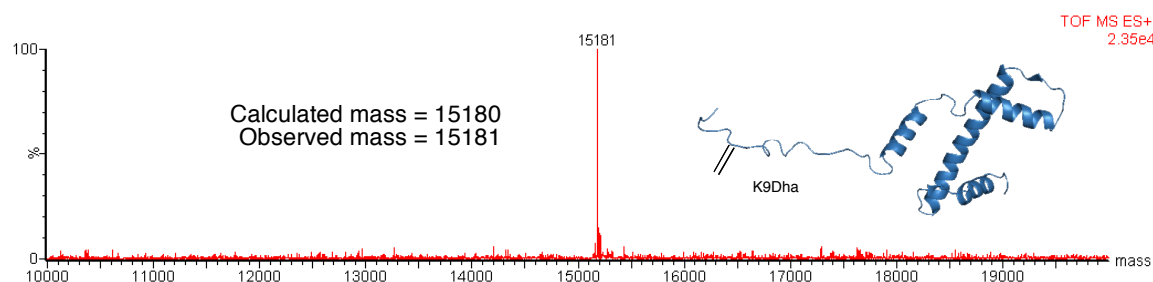
t = 0 min



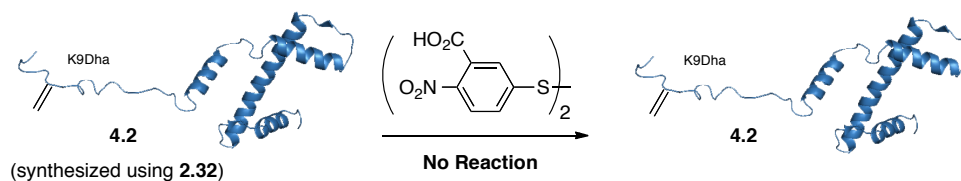
t = 30 min RT



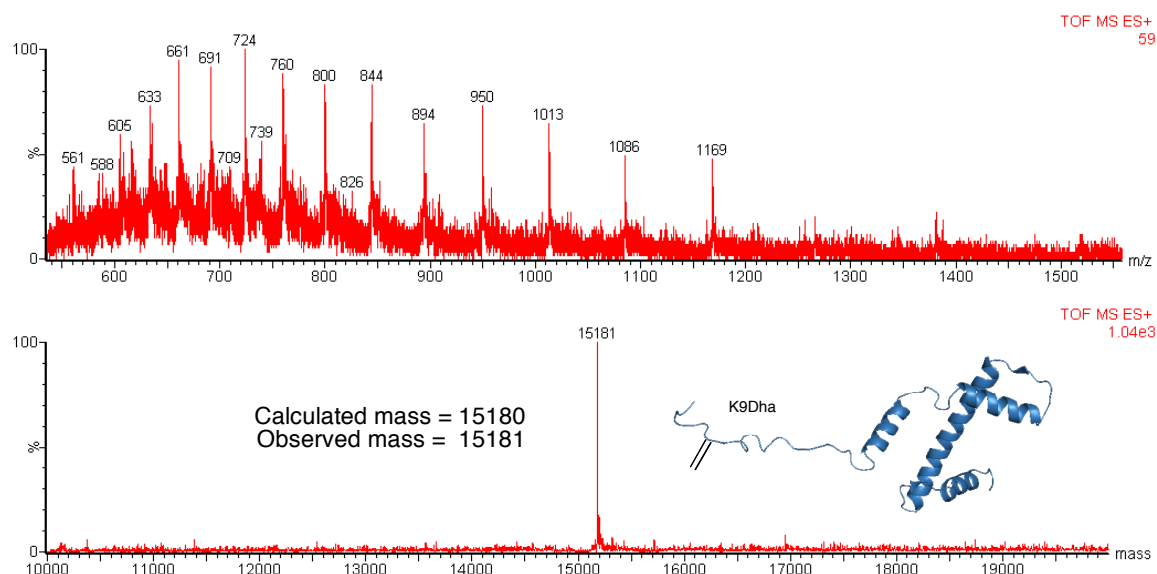
t = 90 min (60 min at 37 °C)



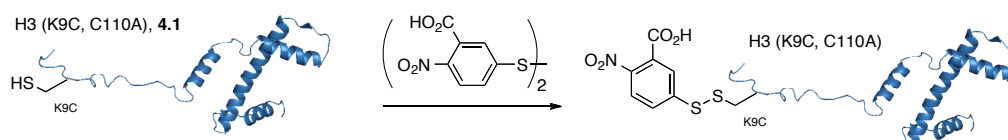
Ellman's test to show all cysteine consumed in reaction with **2.32**



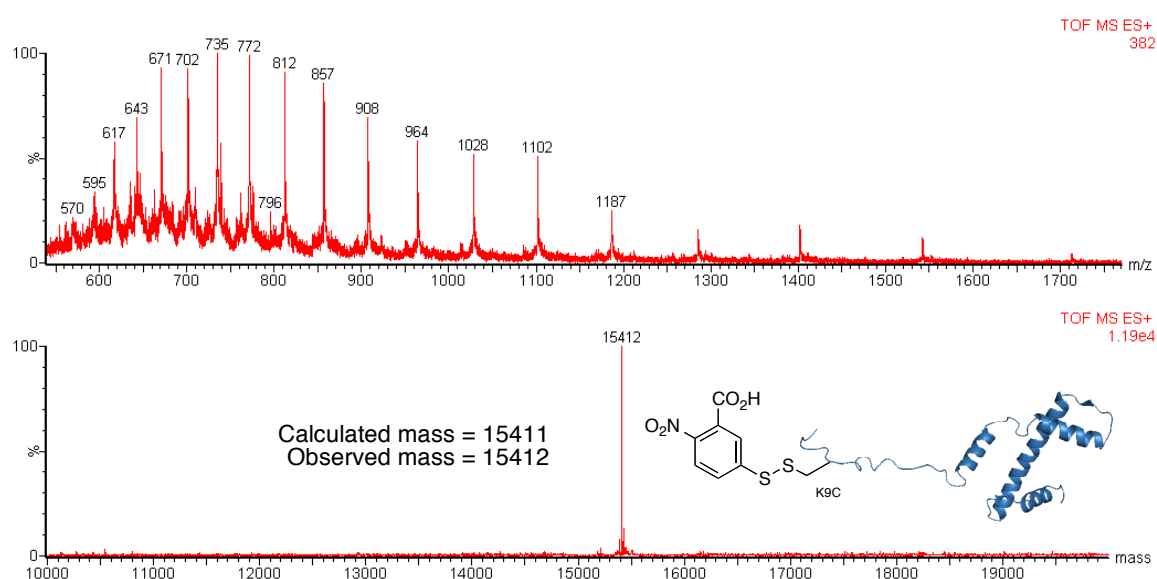
A stock solution of Ellman's reagent was prepared by dissolving 0.3 mg in 3.75 mL pH 8.0 sodium phosphate buffer (50 mM). A 50 μ L aliquot of H3 (K9Dha, C110A) prepared above using **2.32**, 0.25 mg/mL in H₂O after PD minitrapp, was warmed to room temperature. A 20 μ L portion of the Ellman's stock solution was added to the protein. The reaction was shaken for 20 minutes and then analyzed directly by LC-MS. No reaction was observed, indication that all free cysteine was consumed in the previous elimination (calculated mass = 15180; observed mass = 15181). Under the same conditions, the free cysteine of H3 (K9C, C110A) is quantitatively converted to the Ellman disulfide (see next).



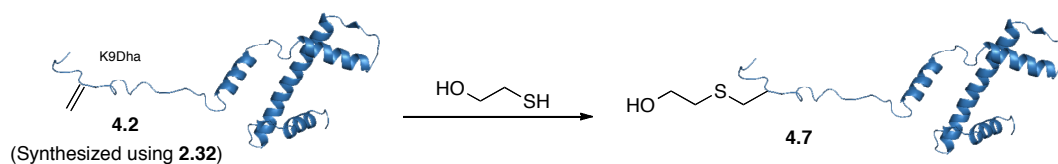
Control: Reaction of H3 (K9C, C110A) with Ellman's reagent



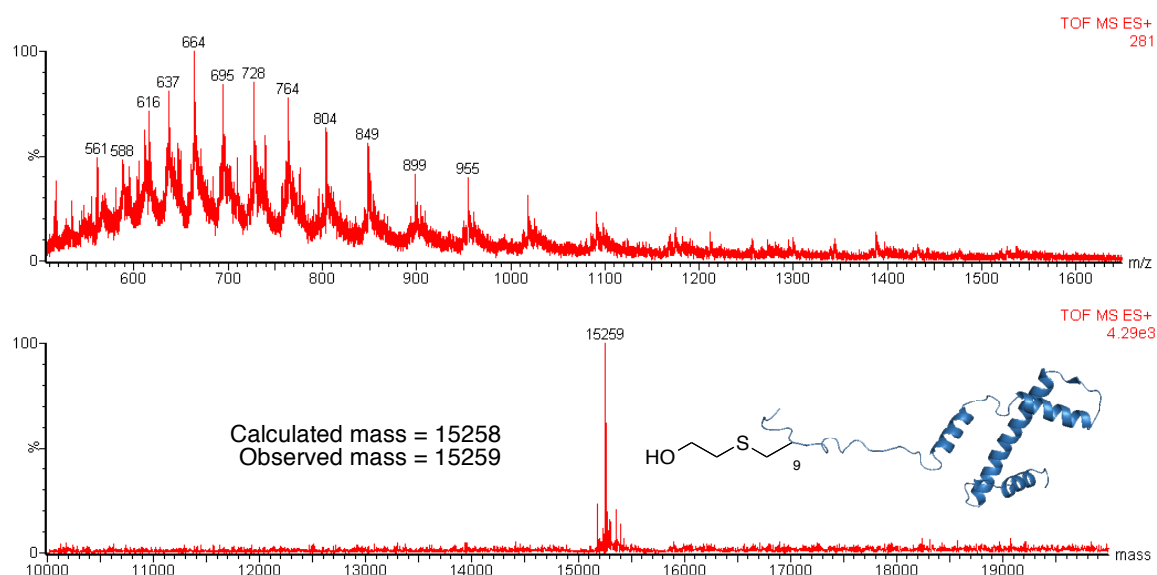
H3 (K9C, C110A) was prepared at 0.25 mg/mL in H₂O. A 20 μ L aliquot of the Ellman's stock solution prepared above was added to 50 μ L of the protein solution. The reaction was shaken at room temperature for 20 minutes and then analyzed directly by LC-MS. Full conversion to the Ellman disulfide was observed (calculated mass = 15411; observed mass = 15412).



Chemical Test for Dehydroalanine

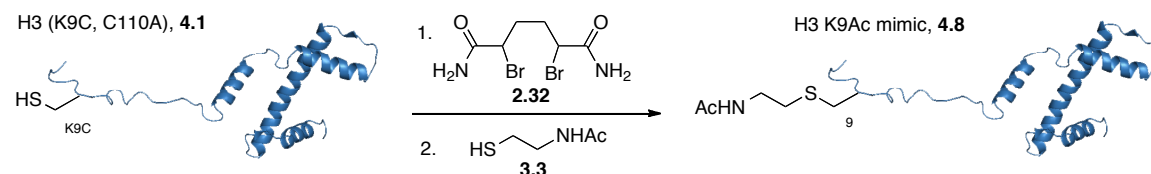


H3 (K9C, C110A) (reduced with DTT and passed through a PD minitrap, see above) was diluted with H₂O to a concentration of 1.0 mg/mL. A 250 μL aliquot of this protein solution (~ 16 nmol) was further diluted with 250 μL of sodium phosphate buffer (pH 8.0, 50 mM). The protein solution was kept on ice until needed. A stock solution of α,α' -di-bromoadipyl(bis)amide (**2.32**) was prepared by dissolving 16.9 mg in 128 μL of DMF. The protein solution was warmed to room temperature and then a 56 μL aliquot of the di-bromide solution (25 μmol) was added. The reaction was vortexed immediately to give a clear solution. The reaction was shaken at room temperature for 30 minutes and then 1 hour at 37 $^{\circ}\text{C}$. After this time, the reaction was cooled to room temperature and analyzed by LC-MS. Complete conversion to dehydroalanine was observed (calculated mass = 15180; observed mass = 15181). A 70 μL aliquot of the reaction mixture (~ 2.3 nmol) was removed and mixed with 0.84 μL of 2-mercaptoethanol (0.012 mmol). After 30 minutes of shaking at room temperature, the reaction was analyzed directly by LC-MS. Full conversion to the thioether was observed (calculated mass = 15258; observed mass = 15259). ESI-MS are shown below.

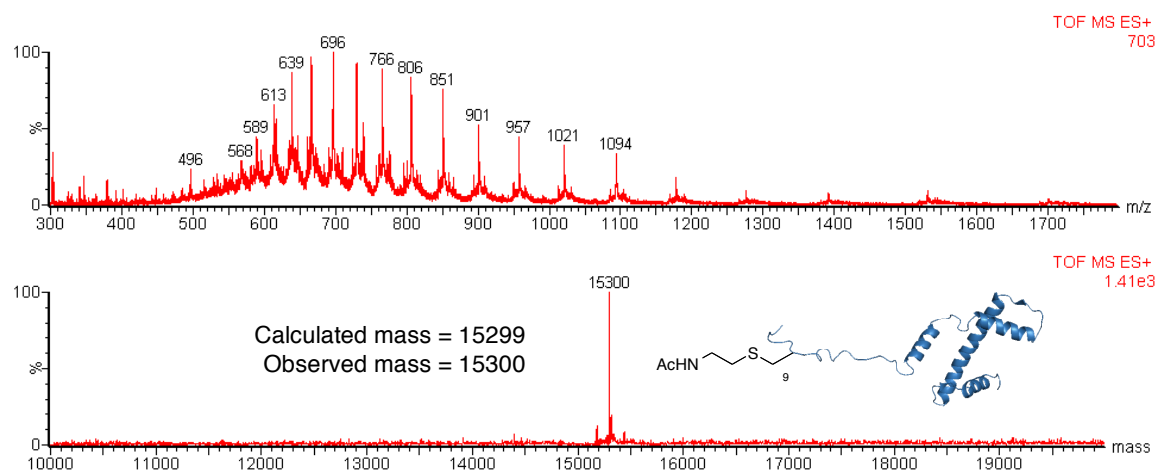


Chemical Access to Histones Bearing Post-Translational Modifications

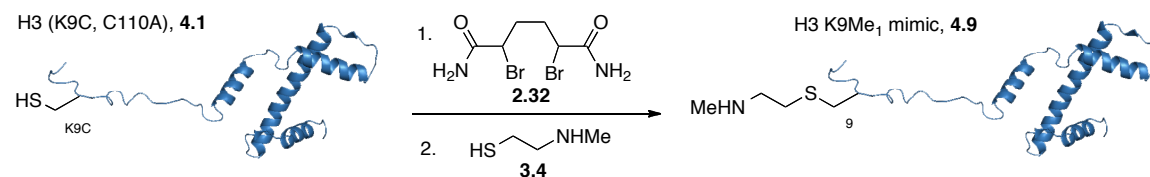
Installation of an acetyllysine mimic at position 9 of histone 3



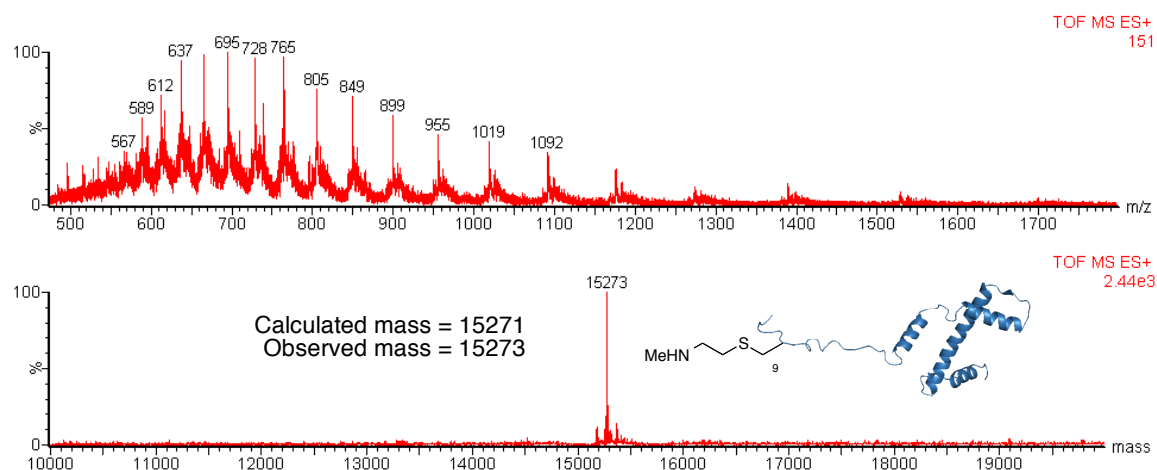
H3 (K9C, C110A) (reduced with DTT and passed through a PD minitrapp, see above) was diluted with H₂O to a concentration of 1.0 mg/mL. A 250 μ L aliquot of this protein solution (~16 nmol) was further diluted with 250 μ L of sodium phosphate buffer (pH 8.0, 50 mM). The protein solution was kept on ice until needed. A stock solution of α,α' -di-bromoadipyl(bis)amide (**2.32**) was prepared by dissolving 16.9 mg in 128 μ L of DMF. The protein solution was warmed to room temperature and then a 56 μ L aliquot of the di-bromide solution (25 μ mol) was added. The reaction was vortexed immediately to give a clear solution. The reaction was shaken at room temperature for 30 minutes and then 1 hour at 37 $^{\circ}$ C. After this time, precipitated di-bromide **2.32** was removed by centrifugation (2 min, 14K g) and the supernatant was carefully transferred to a new reaction tube. LC-MS revealed full conversion to dehydroalanine (calculated mass = 15180; observed mass = 15181). The acetyl-lysine analog was obtained by treating the protein solution with 14 μ L of 2-mercapto-1-*N*-acetyylethylamine (**3.3**, neat). The reaction was shaken at room temperature for 1 hour. After this time, any precipitate was removed by centrifugation and the supernatant analyzed by LC-MS. Full conversion to the acetyllysine mimic was observed (15299 calculated mass, 15300 found). ESI-MS for the final product is shown below.



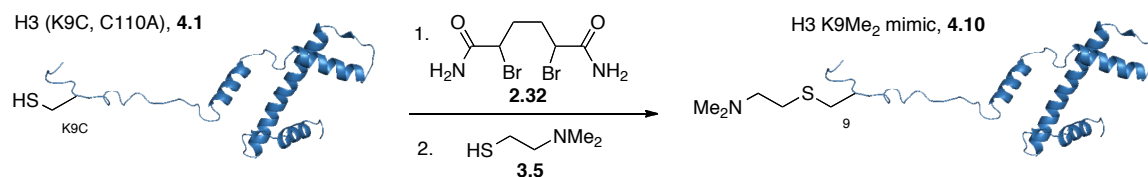
Installation of monomethyllysine mimic at position 9 of histone 3



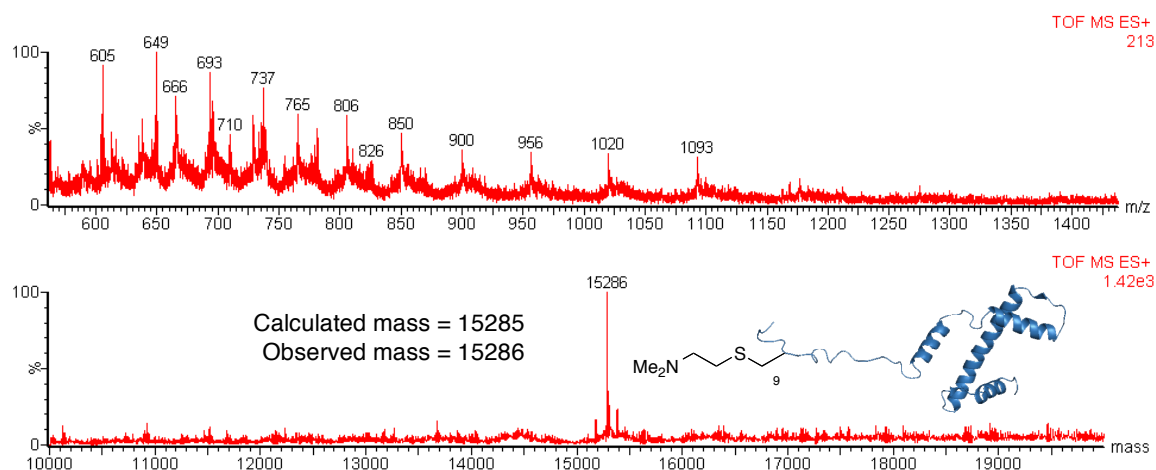
H3 (K9C, C110A) (reduced with DTT and passed through a PD minitrapp, see above) was diluted with H₂O to a concentration of 1.0 mg/mL. A 250 μ L aliquot of this protein solution ($\sim$ 16 nmol) was further diluted with 250 μ L of sodium phosphate buffer (pH 8.0, 50 mM). The protein solution was kept on ice until needed. A stock solution of α,α' -di-bromoadipyl(bis)amide (**2.32**) was prepared by dissolving 16.9 mg in 128 μ L of DMF. The protein solution was warmed to room temperature and then a 56 μ L aliquot of the di-bromide solution (25 μ mol) was added. The reaction was vortexed immediately to give a clear solution. The reaction was shaken at room temperature for 30 minutes and then 1 hour at 37 $^{\circ}$ C. After this time, precipitated di-bromide **2.32** was removed by centrifugation (2 min, 14K g) and the supernatant was carefully transferred to a new reaction tube. LC-MS revealed full conversion to dehydroalanine (calculated mass = 15180; observed mass = 15182). A stock solution of 2-mercapto-*N*-methyl-aminoethane (**3.4**) was prepared by dissolving 21 mg of the hydrochloride salt in 46 μ L of H₂O. The pH was adjusted to $\sim$ 8 by the addition of 16.5 μ L of 5M NaOH. 50 μ L of the resulting thiol solution was added to the dehydroalanine containing protein. After 1 hour at room temperature, small molecules were removed using a PD minitrapp, eluting with 1 mL H₂O. LC-MS of the resulting protein revealed full conversion to the monomethyllysine mimic (calculated mass = 15271; observed mass = 15273). Protein recovery was measured by BCA assay (0.2 mg, 80% recovery). ESI-MS for the modified histone is shown below.



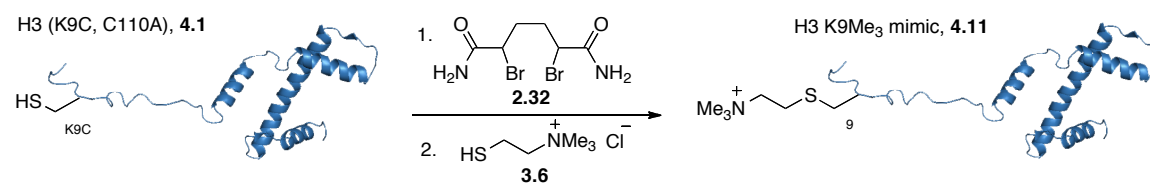
Installation of dimethyllysine mimic at position 9 of histone 3



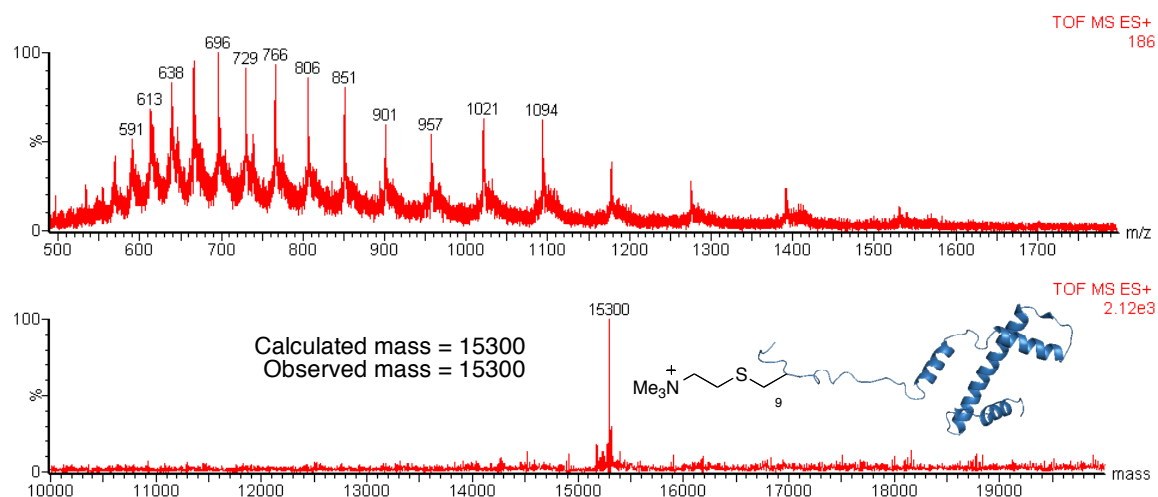
H3 (K9C, C110A) (reduced with DTT and passed through a PD minitrapp, see above) was diluted with H₂O to a concentration of 1.0 mg/mL. A 250 μ L aliquot of this protein solution (~16 nmol) was further diluted with 250 μ L of sodium phosphate buffer (pH 8.0, 50 mM). The protein solution was kept on ice until needed. A stock solution of α, α' -dibromoadipyl(bis)amide (**2.32**) was prepared by dissolving 16.9 mg in 128 μ L of DMF. The protein solution was warmed to room temperature and then a 56 μ L aliquot of the di-bromide solution (25 μ mol) was added. The reaction was vortexed immediately to give a clear solution. The reaction was shaken at room temperature for 30 minutes and then 1 hour at 37 °C. After this time, precipitated di-bromide **2.32** was removed by centrifugation (2 min, 14K g) and the supernatant was carefully transferred to a new reaction tube. A stock solution of 2-mercapto-*N,N*-dimethyl-ethylamine (**3.5**) was prepared by dissolving 15 mg of the hydrochloride salt (Aldrich) in 20 μ L H₂O. The pH of the thiol solution was adjusted to ~8 by the addition of 10.7 μ L 5M NaOH. A 23 μ L aliquot of the resulting thiol solution was added to the dehydroalanine containing histone. After 1 hour at room temperature, the reaction solution was passed through a PD minitrapp to remove small molecules, eluting with 1 mL H₂O. LC-MS analysis of the product solution revealed full conversion to the dimethyllysine mimic (calculated mass = 15285; observed mass = 15286). Protein recovery was measured by BCA assay (0.2 mg, 80% recovery). The ESI-MS for the modified histone is shown below.



Installation of trimethyllysine mimic at position 9 of histone 3

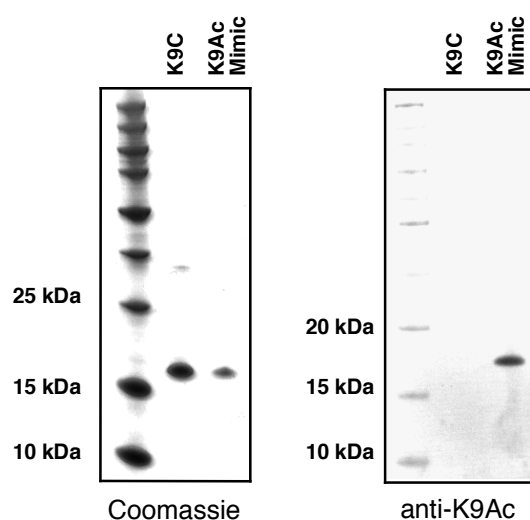


H3 (K9C, C110A) (reduced with DTT and passed through a PD minitrapp, see above) was diluted with H₂O to a concentration of 1.0 mg/mL. A 250 μL aliquot of this protein solution (~16 nmol) was further diluted with 250 μL of sodium phosphate buffer (pH 8.0, 50 mM). The protein solution was kept on ice until needed. A stock solution of α,α'-di-bromoadipyl(bis)amide (**2.32**) was prepared by dissolving 16.9 mg in 128 μL of DMF. The protein solution was warmed to room temperature and then a 56 μL aliquot of the di-bromide solution (25 μmol) was added. The reaction was vortexed immediately to give a clear solution. The reaction was shaken at room temperature for 30 minutes and then 1 hour at 37 °C. After this time, precipitated di-bromide **2.32** was removed by centrifugation (2 min, 14K g) and the supernatant was carefully transferred to a new reaction tube. A stock solution of 2-mercaptoethyl-*N,N,N*-trimethylammonium chloride (**3.6**) was prepared by dissolving 35 mg of the thiol in 49 μL H₂O. The pH of this solution was adjusted to ~8 by the addition of 36 μL of 5M NaOH. 50 μL of the resulting thiol solution was added to the dehydroalanine containing histone. The reaction was shaken at room temperature for 1 hour and then passed through a PD minitrapp, eluting with 1 mL H₂O. Protein recovery was measured by BCA assay (0.2 mg, 80% recovery). The modified protein was analyzed by LC-MS, revealing full conversion to the trimethyllysine mimic (calculated mass = 15300; observed mass = 15300). ESI-MS for the modified protein is shown below.



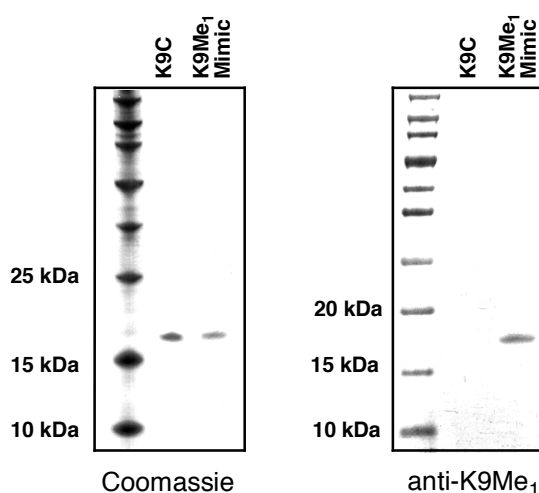
Western Blot for H3 K9Ac Mimic (4.8)

Two aliquots of the H3 K9Ac mimic (4.8) prepared above were subjected to SDS-PAGE using a 12% NuPAGE pre-cast gel (Invitrogen) and MES running buffer (Invitrogen). One aliquot was visualized using Instant Blue (coomassie) staining. The other sample was used for western blot analysis. A lane with unmodified H3 (K9C, C110A) was used alongside both samples as a control along with standard molecular weight markers. For the western blot, the protein was transferred to a PVDF membrane using an iBlot transfer apparatus (Invitrogen). The membrane was blocked with 30 mg/mL BSA in TBST buffer for 30 minutes at room temperature with rocking (TBST = 10 mM Tris, pH 7.8 with 200 mM NaCl and 0.02% (w/w) Tween 20). After blocking, the membrane was washed with TBST buffer (5 × 15 mL). The primary antibody (anti-H3-K9Ac, rabbit polyclonal, Abcam ab10812) was diluted from 80 µL to 10 mL in TBST buffer. The membrane was incubated with the primary antibody solution for 1 hour at room temperature with rocking. The membrane was then washed with TBST buffer (5 × 15 mL). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Thermo Scientific Pierce 31343) was diluted from 10 µL to 10 mL in TBST buffer. The membrane was incubated with the secondary antibody solution for 1 hour at room temperature with rocking. After incubating with the secondary antibody, the membrane was washed with TBST buffer (5 × 15 mL) and then water (2 × 15 mL). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). A clear band was observed at the modified protein and no band was observed at the unmodified H3 (K9C, C110A). The SDS-PAGE gel with coomassie staining and western blot are shown below.



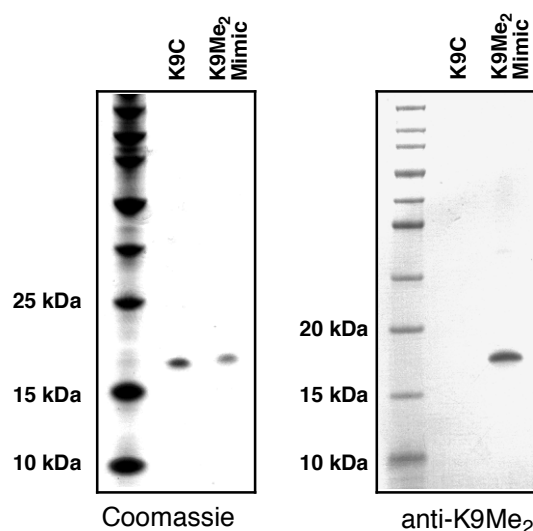
Western Blot for H3 K9Me₁ Mimic (4.9)

Two aliquots of the H3 K9Me₁ mimic (4.9) prepared above were subjected to SDS-PAGE using a 12% NuPAGE pre-cast gel (Invitrogen) and MES running buffer (Invitrogen). One aliquot was visualized using Instant Blue (coomassie) staining. The other sample was used for western blot analysis. A lane with unmodified H3 (K9C, C110A) was used alongside both samples as a control along with standard molecular weight markers. For the western blot, the protein was transferred to a PVDF membrane using an iBlot transfer apparatus (Invitrogen). The membrane was blocked with 30 mg/mL BSA in TBST buffer for 30 minutes at room temperature with rocking (TBST = 10 mM Tris, pH 7.8 with 200 mM NaCl and 0.02% (w/w) Tween 20). After blocking, the membrane was washed with TBST buffer (5 × 15 mL). The primary antibody (anti-H3-K9Me₁, rabbit polyclonal, Millipore 07-450) was diluted from 10 µL to 10 mL in TBST buffer. The membrane was incubated with the primary antibody solution for 1 hour at room temperature with rocking. The membrane was then washed with TBST buffer (5 × 15 mL). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687) was diluted from 5 µL to 10 mL in TBST buffer. The membrane was incubated with the secondary antibody solution for 30 minutes at room temperature with rocking. After incubating with the secondary antibody, the membrane was washed with TBST buffer (5 × 15 mL) and then water (2 × 15 mL). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). A clear band was observed at the modified protein and no band was observed at the unmodified H3 (K9C, C110A). The SDS-PAGE gel with coomassie staining and western blot are shown below.



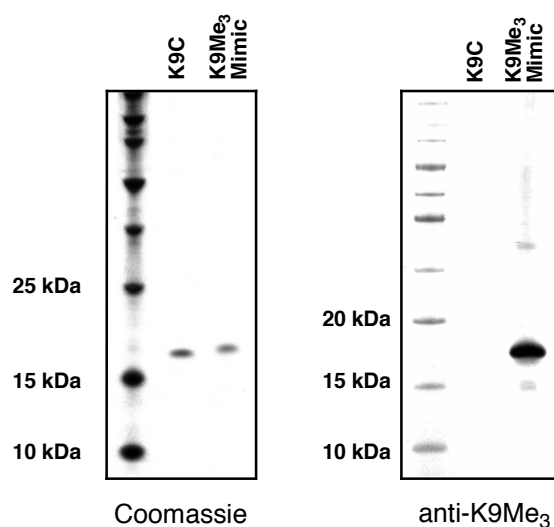
Western Blot for H3 K9Me₂ Mimic (4.10)

Two aliquots of the H3 K9Me₂ mimic (4.10) prepared above were subjected to SDS-PAGE using a 12% NuPAGE pre-cast gel (Invitrogen) and MES running buffer (Invitrogen). One aliquot was visualized using Instant Blue (coomassie) staining. The other sample was used for western blot analysis. A lane with unmodified H3 (K9C, C110A) was used alongside both samples as a control along with standard molecular weight markers. For the western blot, the protein was transferred to a PVDF membrane using an iBlot transfer apparatus (Invitrogen). The membrane was blocked with 30 mg/mL BSA in TBST buffer for 30 minutes at room temperature with rocking (TBST = 10 mM Tris, pH 7.8 with 200 mM NaCl and 0.02% (w/w) Tween 20). After blocking, the membrane was washed with TBST buffer (5 × 15 mL). The primary antibody (anti-H3-K9Me₂, rabbit polyclonal, Millipore 07-441) was diluted from 20 µL to 10 mL in TBST buffer. The membrane was incubated with the primary antibody solution for 1 hour at room temperature with rocking. The membrane was then washed with TBST buffer (5 × 15 mL). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687) was diluted from 5 µL to 10 mL in TBST buffer. The membrane was incubated with the secondary antibody solution for 30 minutes at room temperature with rocking. After incubating with the secondary antibody, the membrane was washed with TBST buffer (5 × 15 mL) and then water (2 × 15 mL). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). A clear band was observed at the modified protein and no band was observed at the unmodified H3 (K9C, C110A). The SDS-PAGE gel with coomassie staining and western blot are shown below.

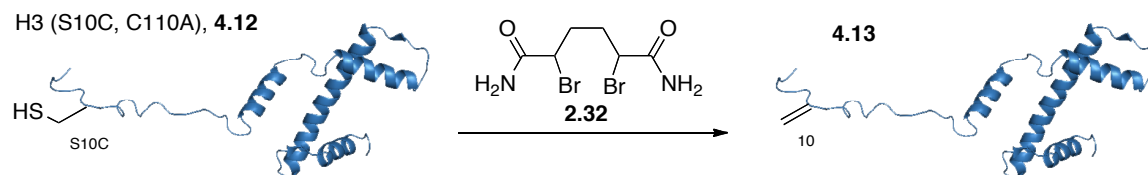


Western Blot for H3 K9Me₃ Mimic (4.11)

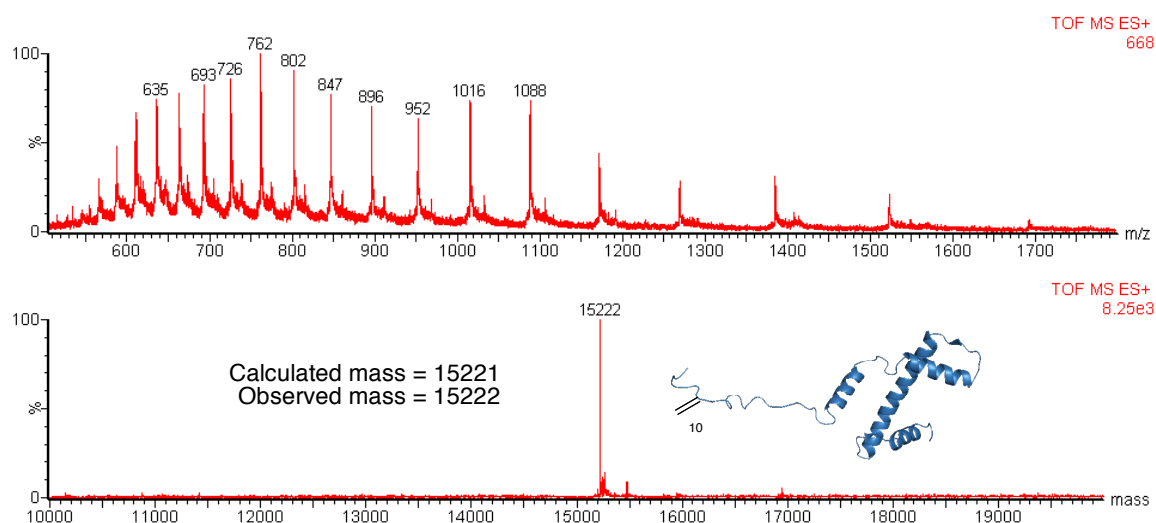
Two aliquots of the H3 K9Me₃ mimic (4.11) prepared above were subjected to SDS-PAGE using a 12% NuPAGE pre-cast gel (Invitrogen) and MES running buffer (Invitrogen). One aliquot was visualized using Instant Blue (coomassie) staining. The other sample was used for western blot analysis. A lane with unmodified H3 (K9C, C110A) was used alongside both samples as a control along with standard molecular weight markers. For the western blot, the protein was transferred to a PVDF membrane using an iBlot transfer apparatus (Invitrogen). The membrane was blocked with 30 mg/mL BSA in TBST buffer for 30 minutes at room temperature with rocking (TBST = 10 mM Tris, pH 7.8 with 200 mM NaCl and 0.02% (w/w) Tween 20). After blocking, the membrane was washed with TBST buffer (5 × 15 mL). The primary antibody (anti-H3-K9Me₃, rabbit polyclonal, Abcam ab8898) was diluted from 10 µL to 10 mL in TBST buffer. The membrane was incubated with the primary antibody solution for 1 hour at room temperature with rocking. The membrane was then washed with TBST buffer (5 × 15 mL). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687) was diluted from 5 µL to 10 mL in TBST buffer. The membrane was incubated with the secondary antibody solution for 30 minutes at room temperature with rocking. After incubating with the secondary antibody, the membrane was washed with TBST buffer (5 × 15 mL) and then water (2 × 15 mL). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). A clear band was observed at the modified protein and no band was observed at the unmodified H3 (K9C, C110A). The SDS-PAGE gel with coomassie staining and western blot are shown below.



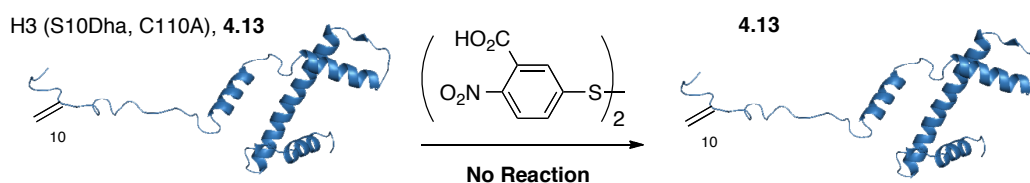
H3 (S10Dha, C110A) synthesis using di-bromide **2.32**



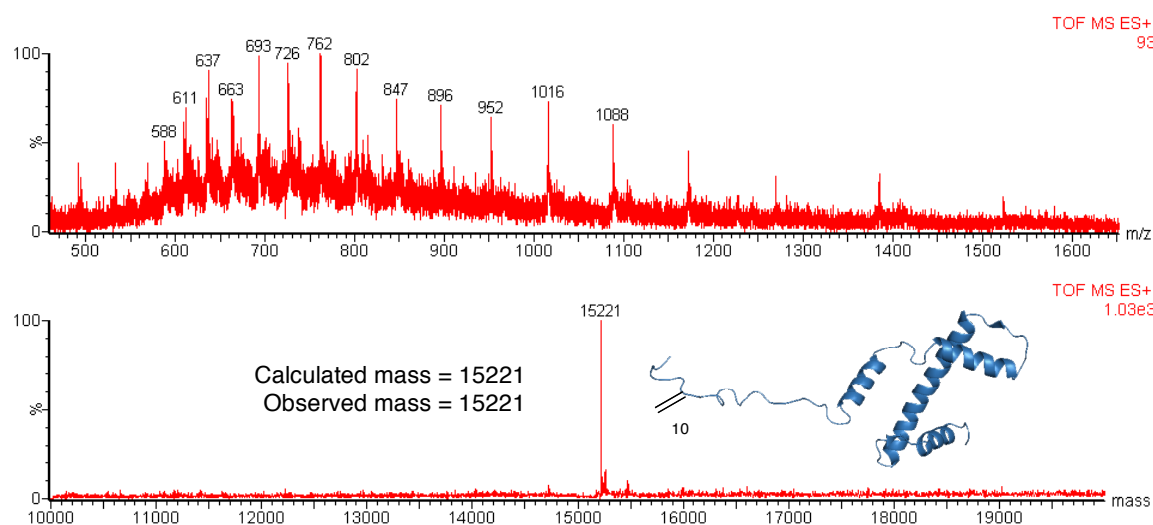
A solution of H3 (S10C, C110A) was prepared at a concentration of 2.0 mg/mL in water. A 250 μ L aliquot of the protein solution was diluted with 250 μ L of pH 8.0 sodium phosphate buffer (50 mM). To this diluted protein solution was then added DTT (10 mg). The solution was shaken at room temperature for 15 minutes to reduce any contaminant disulfide. After this time, the protein solution was passed through a PD minitrapp, previously equilibrated with H₂O, and eluted with 1.0 mL of H₂O. The histone solution was stored on ice until needed. A stock solution of di-bromide **2.32** was prepared by dissolving 84 mg in 564 μ L of DMF. A 250 μ L aliquot of the reduced protein was diluted with 250 μ L of pH 8.0 sodium phosphate (50 mM) and warmed to room temperature. Once the protein was at room temperature, a 50 μ L aliquot of the di-bromide solution was added. The reaction was vortexed and then shaken at room temperature for 30 minutes and then 37 °C for 60 minutes. The reaction mixture was then cooled to room temperature and precipitated di-bromide **2.32** was removed by centrifugation (2 min, 14K g). LC-MS analysis of the supernatant revealed full conversion to a protein with a mass corresponding to H3 (S10Dha) (calculated mass = 15221; observed mass = 15222). ESI-MS are shown below.



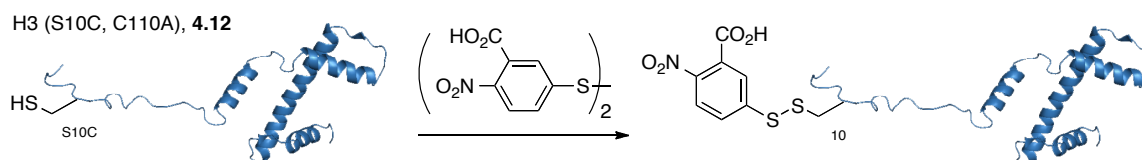
Ellman's test to confirm reaction at cysteine



A stock solution of Ellman's reagent was prepared by dissolving 3.0 mg in 577 μL of 50 mM sodium phosphate buffer (pH 8.0). A 100 μL aliquot of H3 (S10Dha, C110A) prepared above ($c = 0.5 \text{ mg/mL}$ in 25 mM sodium phosphate buffer, pH 8.0) was mixed with 5 μL of the Ellman's reagent stock solution. The solution was shaken at room temperature for 15 minutes. After this time, the reaction was analyzed directly by LC-MS. No reaction was observed and only unreacted starting material was obtained, consistent with complete consumption of cysteine in the elimination reaction (calculated mass = 15221; observed mass = 15221). ESI-MS are shown below. Under identical conditions, H3 (S10C, C110A) reacts with Ellman's reagent to give the corresponding disulfide (see next).

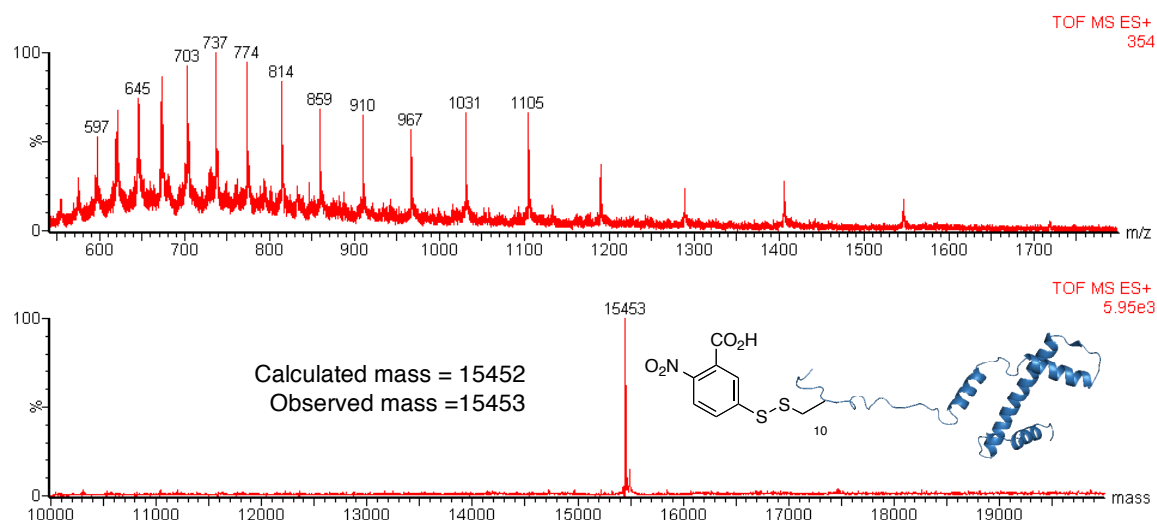


Control: H3 (S10C, C110A) Ellman's Test



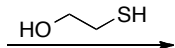
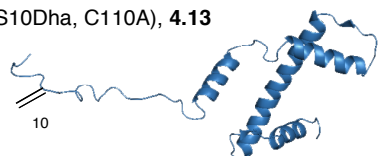
A solution of H3 (S10C, C110A) was prepared at 0.5 mg/mL in 25 mM sodium phosphate (pH 8.0). To 100 μL of the protein solution was added 5 μL of the Ellman's reagent stock solution prepared above. The reaction was shaken at room temperature for

15 minutes and then analyzed directly by LC-MS. Full conversion to the corresponding disulfide was observed (calculated mass = 15452; observed mass = 15453). ESI-MS are shown below.

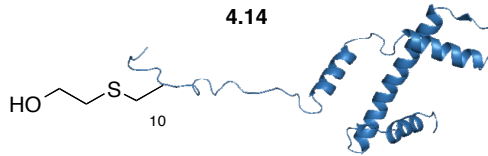


Chemical Test for Dha: Reaction of H3 (S10Dha, C110A) and 2-mercaptoethanol

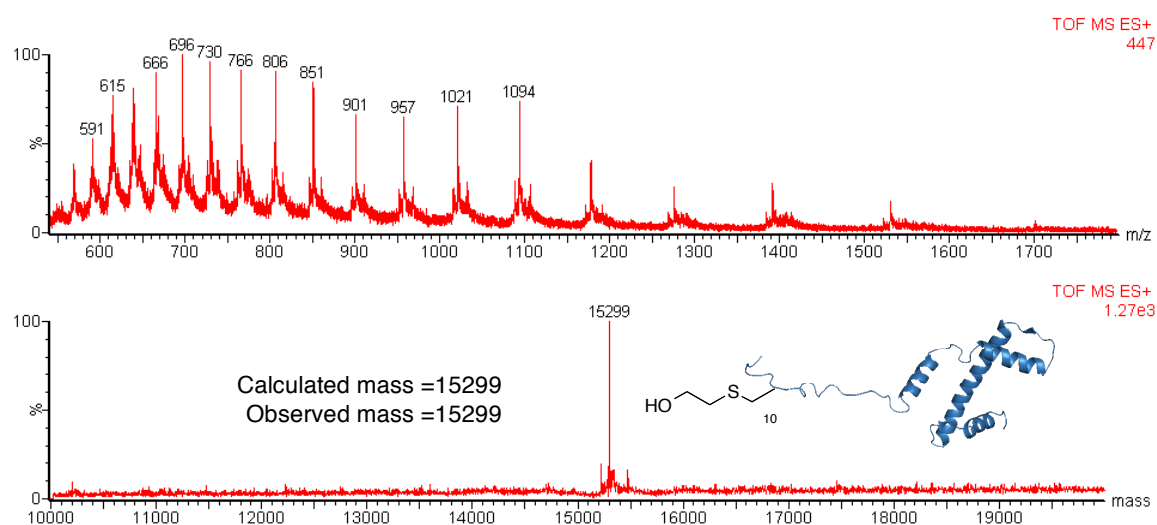
H3 (S10Dha, C110A), 4.13



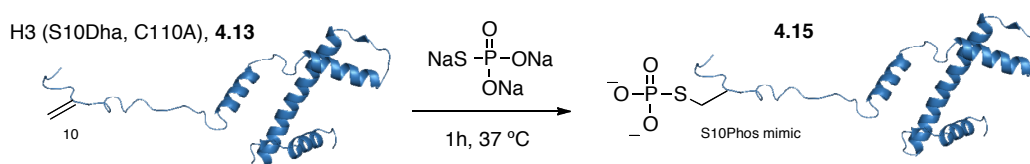
4.14



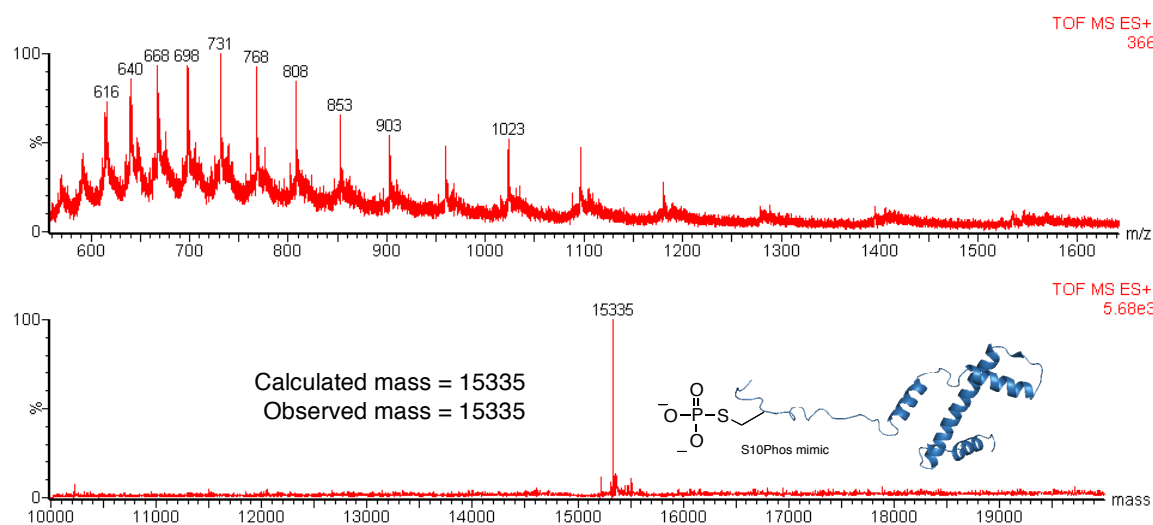
A 100 μ L solution of H3 (S10Dha, C110A) prepared above was mixed with 5 μ L of 2-mercaptoethanol (neat). The reaction was shaken at room temperature for 40 minutes and then analyzed directly by LC-MS. Full conversion to the corresponding thioether was observed. (calculated mass = 15299; observed mass = 15299). ESI-MS are shown below.



Synthesis of PhosSer mimic on H3 at position 10



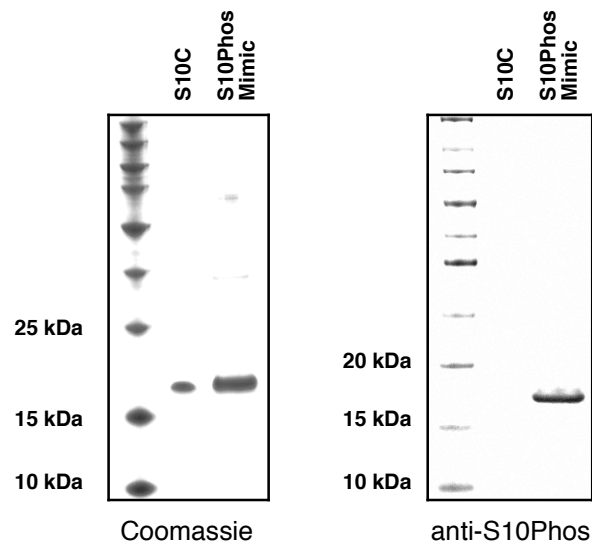
A stock solution of sodium thiophosphate was prepared by dissolving 46 mg in 18.6 μL H_2O and 20 μL of 5M HCl. The final pH of this stock solution was approximately 8. A 200 μL aliquot of H3 (S10Dha, C110A) prepared above ($c = 0.5$ mg/mL in 25 mM sodium phosphate buffer, pH 8.0) was mixed with 25 μL of the stock solution of thiophosphate. The reaction was incubated at 37 °C for 1 hour and then analyzed directly by LC-MS. Full conversion to phosphocysteine was observed (Calculated mass = 15335; observed mass = 15335). ESI-MS are shown below. The protein solution was split into aliquots and flash frozen. These aliquots were used directly for western blot analysis (see next).



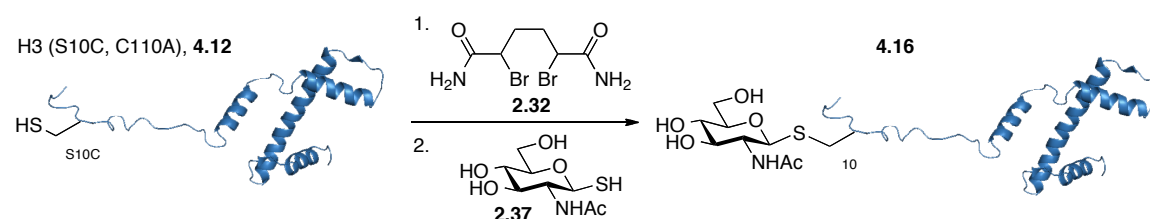
Western Blot for H3 S10Phos Mimic

Two aliquots of the H3 S10Phos mimic prepared above were subjected to SDS-PAGE using a 12% NuPAGE pre-cast gel (Invitrogen) and MES running buffer (Invitrogen). One aliquot was visualized using Instant Blue (coomassie) staining. The other sample was used for western blot analysis. A lane with unmodified H3 (S10C, C110A) was used alongside both samples as a control along with standard molecular weight markers. For the western blot, the protein was transferred to a PVDF membrane using an iBlot transfer

apparatus (Invitrogen). The membrane was blocked with 30 mg/mL BSA in TBST buffer for 30 minutes at room temperature with rocking (TBST = 10 mM Tris, pH 7.8 with 200 mM NaCl and 0.02% (w/w) Tween 20). After blocking, the membrane was washed with TBST buffer (5×15 mL). The primary antibody (anti-H3-S10Phos, rabbit polyclonal, Abcam ab5176) was diluted from 20 μ L to 10 mL in TBST buffer. The membrane was incubated with the primary antibody solution for 1 hour at room temperature with rocking. The membrane was then washed with TBST buffer (5×15 mL). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687) was diluted from 2 μ L to 10 mL in TBST buffer. The membrane was incubated with the secondary antibody solution for 15 minutes at room temperature with rocking (a short incubation time was used for the secondary antibody to minimize any hydrolysis of the H3 phosphocysteine by the alkaline phosphatase). After incubating with the secondary antibody, the membrane was washed with TBST buffer (5×15 mL) and then water (2×15 mL). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). A clear band was observed at the modified protein and no band was observed at the unmodified H3 (S10C, C110A). The SDS-PAGE gel with coomassie staining and western blot are shown below.

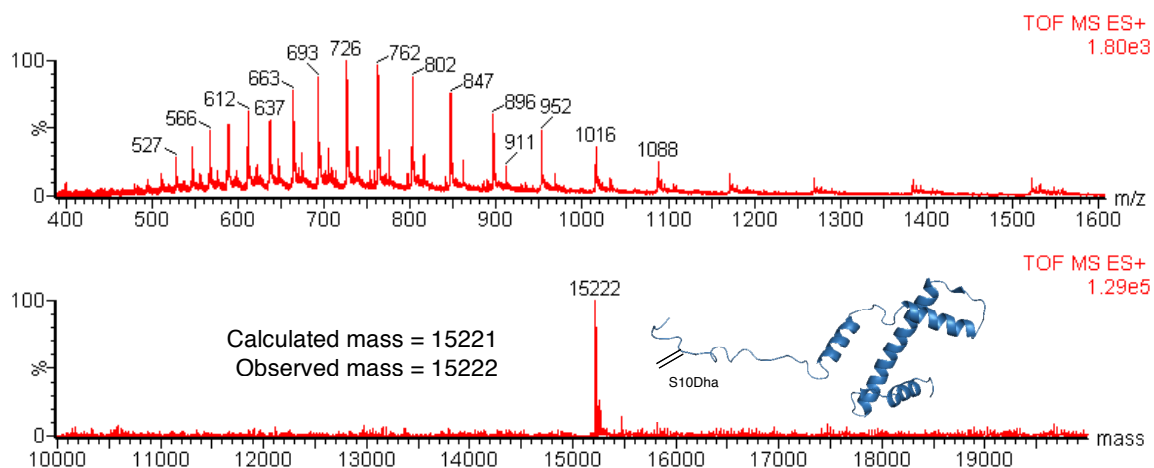


Synthesis of GlcNAc serine mimic on H3 at position 10:

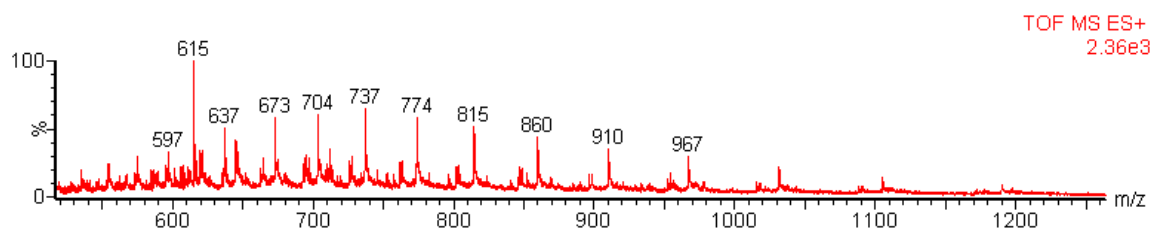


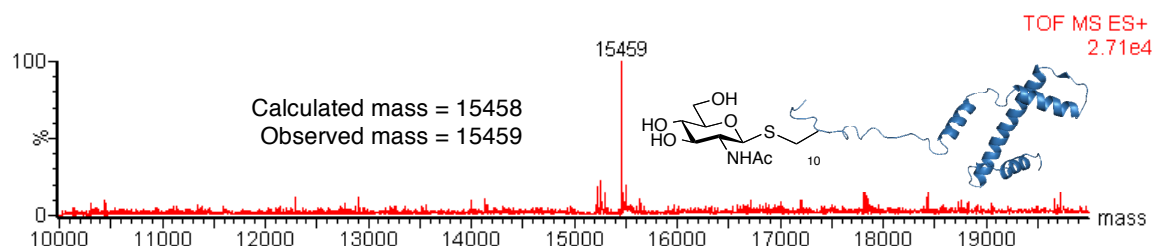
H3 (S10C, C110A) was prepared at 1.0 mg/mL in H₂O. A 200 μ L aliquot of this H3 solution was diluted with 100 μ L of sodium phosphate buffer (50 mM, pH 8.0). A stock solution of di-bromide **2.32** was prepared by dissolving 34 mg in 229 μ L DMF. A 20 μ L aliquot of the di-bromide solution was added to the protein solution. The reaction was shaken at room temperature for 30 minutes and then at 37 $^{\circ}$ C for 1 hour. LC-MS analysis showed full conversion to dehydroalanine after this time (calculated mass = 15221; observed mass = 15222). This dehydroalanine solution was used directly for the addition of GlcNAc derivative **2.37**. Accordingly, 6.2 mg of the GlcNAc derivative **2.37** (26 μ mol) was added as a solid to the protein solution. The reaction was shaken at room temperature for 90 minutes after which time LC-MS revealed full conversion to the GlcNAc-histone (calculated mass = 15458; observed mass = 15459). ESI-MS are shown below.

ESI-MS for S10Dha:

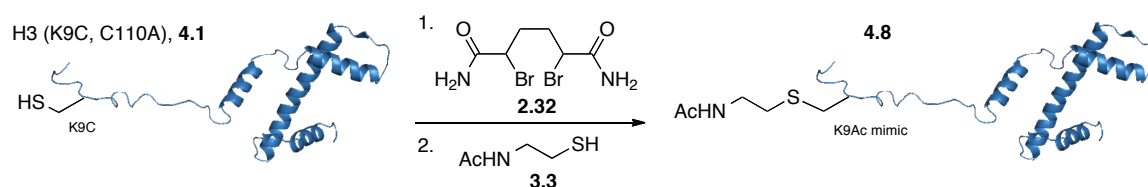


ESI-MS for GlcNAc serine mimic:

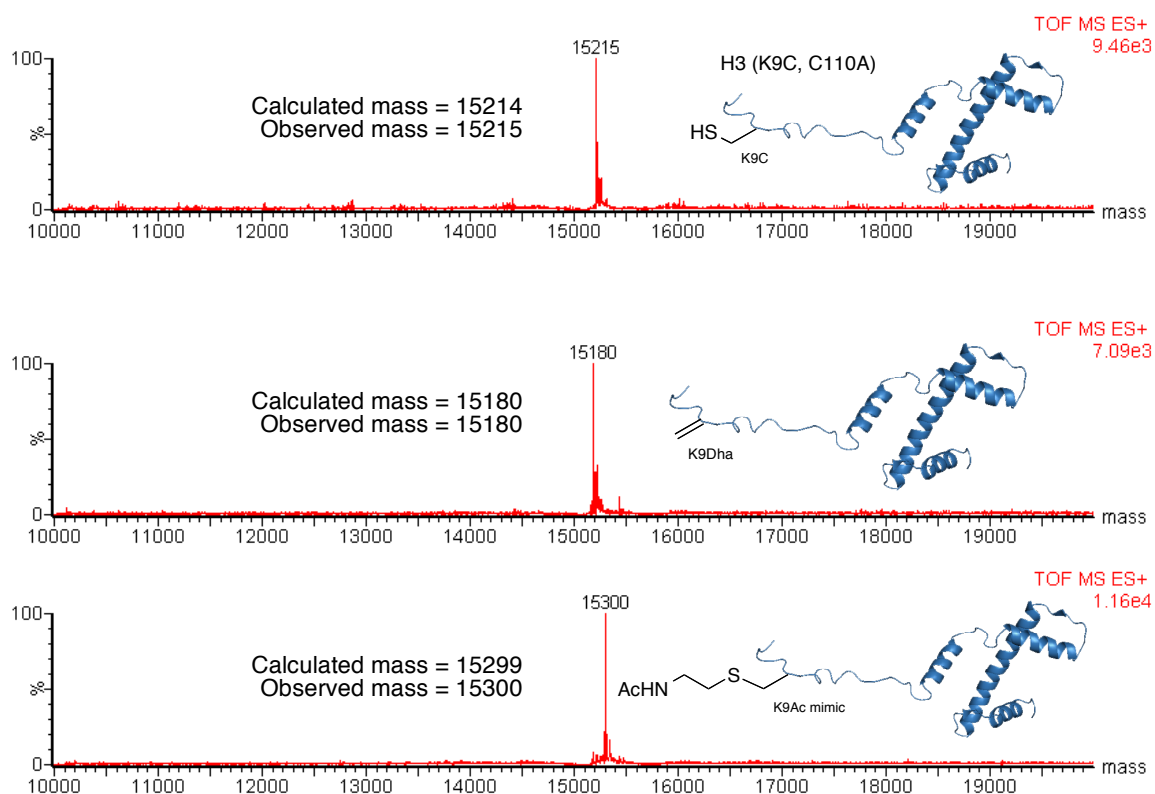




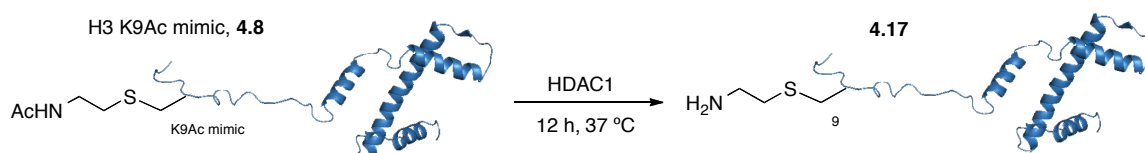
Preparation of H3 K9Ac mimic for HDAC assay



H3 (K9C, C110A) was prepared as a 2.9 mg/mL solution in water. A 250 μ L aliquot of this protein solution was diluted with 250 μ L of sodium phosphate buffer (pH 8.0, 50 mM). DTT (11 mg) was added as a solid to reduce any contaminant disulfide. The solution was shaken at room temperature for 15 minutes to ensure complete reduction. After this time, the protein solution was passed through a PD minitrapp (GE Healthcare), eluting with 50 mM sodium phosphate (1.0 mL, pH 8.0). A stock solution of dibromide **2.32** was prepared by dissolving 30 mg in 250 μ L of DMF. An 89 μ L aliquot of the dibromide solution was added to the protein solution obtained after the PD minitrapp. The reaction was shaken at room temperature for 30 minutes and then 37 $^{\circ}$ C for 1 hour. After this time, solids were removed by centrifugation (1 min, 14K g) and the reaction was analyzed by LC-MS. Full conversion to dehydroalanine was observed (calculated mass = 15180; observed mass = 15180). *N*-acetyl cysteamine (**3.3**, 46 μ L neat) was added to the reaction mixture. The reaction was shaken at room temperature for 1 hour. After this time, the reaction was analyzed directly by LC-MS, revealing full conversion to the H3 K9Ac mimic (15299 calc. mass; 15300 observed). ESI-MS for H3 (K9C, C110A); H3 containing dehydroalanine at position 9; and the H3 K9Ac mimic are shown below. After the synthesis of the K9Ac mimic was complete, small molecules were removed by dialysis in 4 L of water (Slide-A-Lyzer, 10K MWCO, Pierce). After dialysis, protein recovery was measured by BCA assay (0.7 mg, 96% recovery). The purified protein solution was then split into aliquots and flash frozen.



Deacetylation of H3 K9Ac mimic (4.8) with HDAC1



Two samples were prepared for the HDAC1 assay; each sample was prepared in duplicate. Of the duplicate samples, one was used for SDS-PAGE and western blot and the other was used for LC-MS analysis.

Sample 1 (Negative control)

5 μ L of the H3-K9Ac mimic **4.8** prepared above and purified by dialysis

20 μ L of the reaction buffer (25 mM Tris, pH 8.0, 138 mM NaCl)

(The H3 K9Ac concentration was measured to 0.6 mg / mL by BCA assay after dilution in the reaction buffer)

Sample 2 (Reaction)

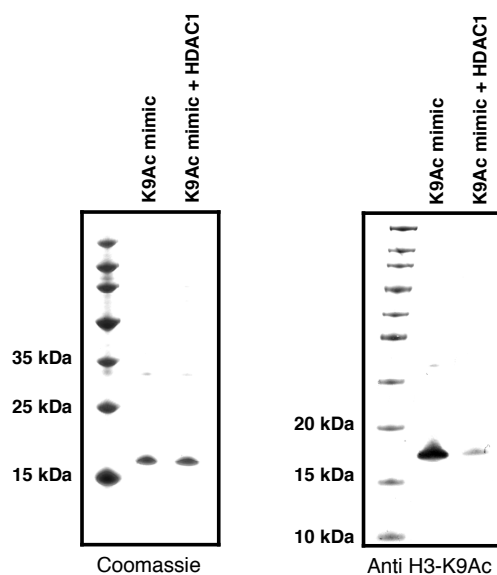
5 μ L of the H3-K9Ac mimic **4.8** prepared above and purified by dialysis

10 μ L of the reaction buffer (25 mM Tris, pH 8.0, 138 mM NaCl)

10 μ L of recombinant HDAC1 (Active Motif 31342, 0.27 μ g/ μ L in 40 mM Tris, pH 8.0, 110 mM NaCl, 2.2 mM KCl, 80 μ g/mL FLAG peptide, 20% glycerol)

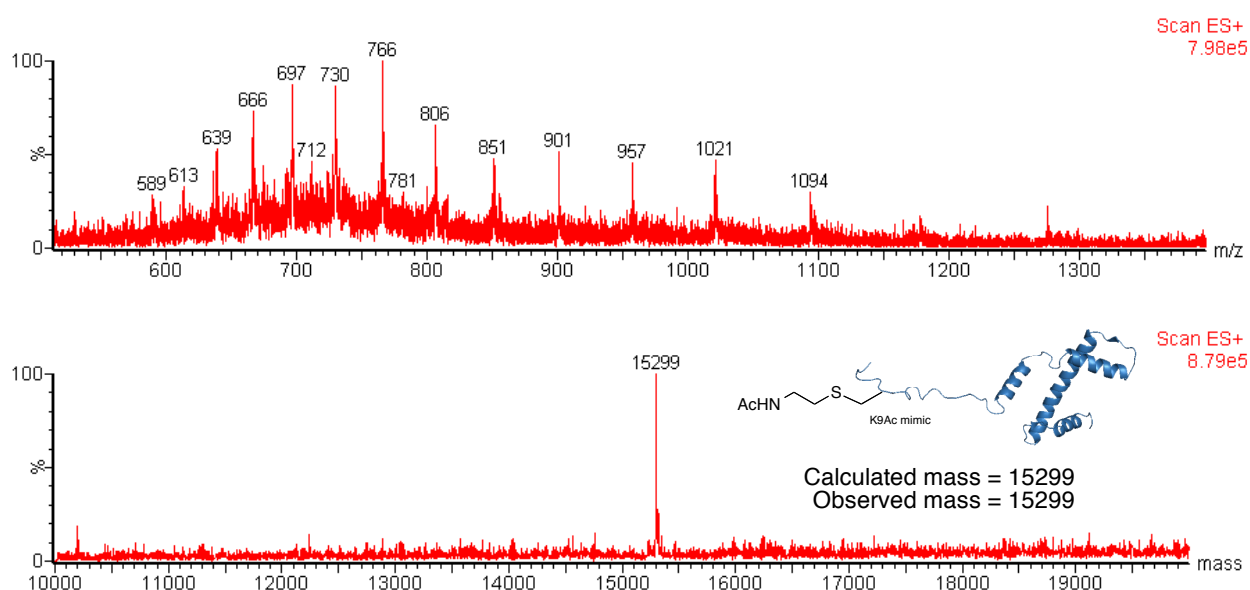
Both samples were vortexed and then shaken at 37 °C for 12 hours. After this time, SDS-PAGE loading buffer was added directly to both Sample 1 and Sample 2, the samples were heated at 95 °C for 5 minutes, and then analyzed by SDS-PAGE and western blot. The duplicate samples were analyzed by LC-MS.

Western Blot: After SDS-PAGE (12% NuPAGE pre-cast gel, MES running buffer, Invitrogen), proteins were transferred to a PVDF membrane using an iBlot transfer apparatus (Invitrogen). The membrane was blocked with 30 mg/mL BSA in TBST buffer for 30 minutes at room temperature with rocking (TBST = 10 mM Tris, pH 7.8 with 200 mM NaCl and 0.02% (w/w) Tween 20). After blocking, the membrane was washed with TBST buffer (5 $\times$ 15 mL). The primary antibody (anti-H3-K9Ac, rabbit polyclonal, Abcam ab10812) was diluted from 40 μ L to 10 mL in TBST buffer. The membrane was incubated with the primary antibody solution for 1 hour at room temperature with rocking. The membrane was then washed with TBST buffer (5 $\times$ 15 mL). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687) was diluted from 2 μ L to 10 mL in TBST buffer. The membrane was incubated with the secondary antibody solution for 1 hour at room temperature with rocking. After incubating with the secondary antibody, the membrane was washed with TBST buffer (5 $\times$ 15 mL) and then water (2 $\times$ 15 mL). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). The lane with H3-K9Ac mimic and no HDAC1 (Sample 1) showed a strong band. The lane containing both the H3-K9Ac mimic and HDAC1 showed a very weak band, consistent with deacetylation of the lysine mimic. The SDS-PAGE gel with coomassie staining and the WB membrane are shown on the following page.



LC-MS: 15 μ L of each sample were mixed with 10 μ L of 3M guanidinium hydrochloride (pH 7.6). The mixture was homogenized by vortex and then analyzed directly by LC-MS. No reaction was observed for Sample 1 (no HDAC1) and only unreacted **4.8** was observed (calculated mass = 15299, observed mass = 15299). For sample 2 (reaction with HDAC1), 42% deacetylation was observed, according to the relative peak height of **4.8** (calculated mass = 15299, observed mass = 15299) and the deacetylated protein **4.17** (calculated mass = 15257, observed mass = 15258). ESI-MS are shown below.

Sample 1 (control, no HDAC):



Reaction scheme showing the conversion of H3 K9Ac Mimic, **8** to compound **15** using HDAC2. The reaction conditions are 12 h, 37 °C. The structure of **8** is shown as a peptide chain with a K9Ac mimic and an AcHN group. The structure of **15** is shown as a peptide chain with a K9Ac mimic and an H₂N group.

Sample 1 (Negative control)

20 μ L of the reaction buffer (25 mM Tris, pH 8.0, 138 mM NaCl)

(The concentration of **4.8** was measured to 0.6 mg / mL by BCA assay after dilution in the reaction buffer)

Sample 2 (Reaction)

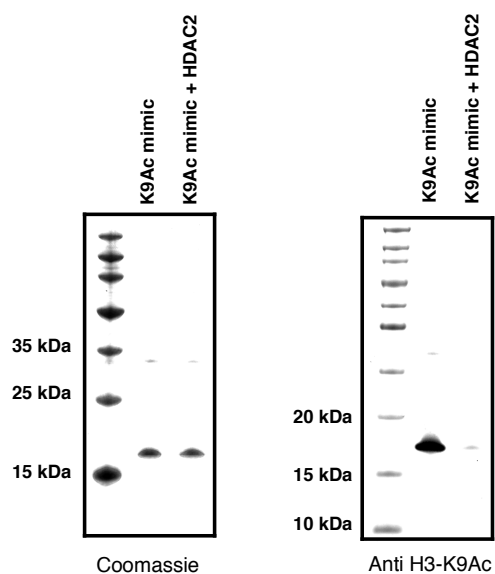
5 μ L of the H3 K9Ac mimic **4.8** prepared above and purified by dialysis

10 μ L of the reaction buffer (25 mM Tris, pH 8.0, 138 mM NaCl)

10 μ L of recombinant HDAC2 (Active Motif 31343, 0.24 μ g/ μ L in 25 mM Tris, pH 8.0, 139 mM NaCl, 0.05% Tween-20, 10% glycerol)

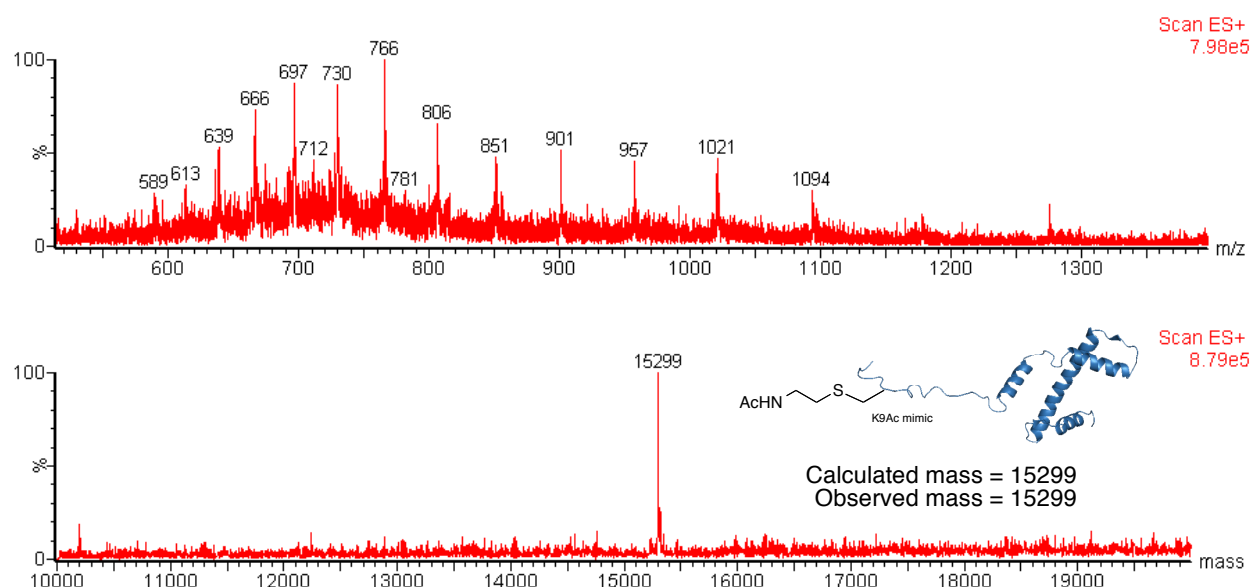
Both samples were vortexed to homogenize and then shaken at 37 °C for 12 hours. After this time, SDS-PAGE loading buffer was added directly to both samples, the samples were heated at 95 °C for 5 minutes, and then analyzed by SDS-PAGE and western blot. The duplicate samples were analyzed by LC-MS.

Western Blot: After SDS-PAGE (12% NuPAGE pre-cast gel, MES running buffer, Invitrogen), proteins were transferred to a PVDF membrane using an iBlot transfer apparatus (Invitrogen). The membrane was blocked with 30 mg/mL BSA in TBST buffer for 30 minutes at room temperature with rocking (TBST = 10 mM Tris, pH 7.8 with 200 mM NaCl and 0.02% (w/w) Tween 20). After blocking, the membrane was washed with TBST buffer (5 $\times$ 15 mL). The primary antibody (anti-H3-K9Ac, rabbit polyclonal, Abcam ab10812) was diluted from 40 μ L to 10 mL in TBST buffer. The membrane was incubated with the primary antibody solution for 1 hour at room temperature with rocking. The membrane was then washed with TBST buffer (5 $\times$ 15 mL). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687) was diluted from 2 μ L to 10 mL in TBST buffer. The membrane was incubated with the secondary antibody solution for 1 hour at room temperature with rocking. After incubating with the secondary antibody, the membrane was washed with TBST buffer (5 $\times$ 15 mL) and then water (2 $\times$ 15 mL). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). The lane with H3 K9Ac mimic and no HDAC2 (Sample 1) showed a strong band. The lane containing both the H3 K9Ac mimic and HDAC2 showed only a very weak band, consistent with deacetylation of the lysine mimic. The SDS-PAGE gel with coomassie staining and the WB membrane are shown on the following page.



LC-MS: 15 μ L of each sample were mixed with 10 μ L of 3M guanidinium hydrochloride (pH 7.6). The mixture was homogenized by vortex and then analyzed directly by LC-MS. No reaction was observed for Sample 1 (no HDAC1) and only unreacted **4.8** was observed (calculated mass = 15299, observed mass = 15299). For sample 2 (reaction with HDAC2), 55% deacetylation was observed, according to the relative peak height of **4.8** (calculated mass = 15299, observed mass = 15299) and the deacetylated protein **4.17** (calculated mass = 15257, observed mass = 15258). ESI-MS are shown below.

Sample 1 (control, no HDAC):



Scan ES+ 1.30e5

100

%

637 667 695 728 764 804 849 880 902 922 1021 1094

m/z

600 650 700 750 800 850 900 950 1000 1050 1100 1150 1200 1250

0

Calculated mass = 15257
Observed mass = 15258

15258 15299

100

%

10000 11000 12000 13000 14000 15000 16000 17000 18000 19000

mass

0

Calculated mass = 15299
Observed mass = 15299

Scan ES+ 4.92e5

AcHN

K9Ac mimic

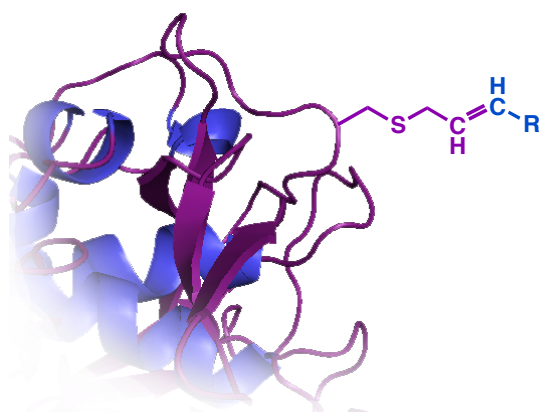
H₂N

9

- (1) Johnson, C. R.; Kirchhoff, R. A.; Corkings, H. G. *J. Org. Chem.* **1974**, *39*, 2458-2459.
- (2) Harpp, D. N.; Smith, R. A.; Steliou, K. *J. Org. Chem.* **1981**, *46*, 2072-2079.
- (3) Treibs, A.; Neumayr, F. *Chem. Ber.* **1957**, *90*, 76-79.
- (4) Meyer, W.; Reckendorf, Z.; Bonner, W. A. *J. Org. Chem.* **1961**, *26*, 4596-4599.
- (5) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 2244-2247.
- (6) El Alaoui, A.; Schmidt, F.; Amessou, M.; Sarr, M.; Decaudin, D.; Florent, J.-C.; Johannes, L. *Angew. Chem. Int. Ed.* **2007**, *46*, 6469-6472.

Chapter 5

Allyl Sulfides are Reactive Substrates in Olefin Metathesis and Enable Cross-Metathesis in Water and on Proteins



*“Dans les champs de l’observation
le hasard ne favorise que les esprits préparé”*

Louis Pasteur
University of Lille, 1854

5.1 Introduction

Olefin metathesis is a central transformation in organic synthesis.¹⁻⁴ The wide scope and functional group tolerance of ruthenium based olefin metathesis is enabled by the availability of several catalysts. A schematic of these reactions and a selection of commercially available catalysts⁵⁻⁷ is depicted in Figure 5.1.

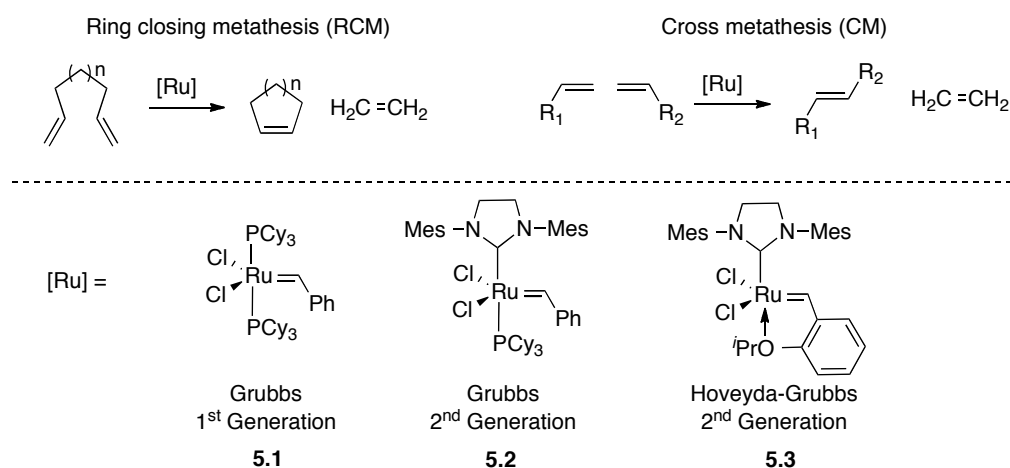


Figure 5.1: Olefin metathesis and common ruthenium-based catalysts

The capacity to forge a carbon-carbon bond in a mild and selective manner has inspired the application of both ring-closing metathesis (RCM) and cross-metathesis (CM) in the modification of biomolecules.^{8, 9} Our group has a long-term interest in adapting the tools of organic synthesis to protein modification and olefin metathesis offers many opportunities in engineering protein function. In fact, the potential of olefin metathesis for protein modification was recognized more than a decade ago when the first alkenes were incorporated into proteins by the Tirrell laboratory.¹⁰ Yet, despite the prospect for selective carbon-carbon bond formation on proteins, the successful application of olefin metathesis to protein modification was an outstanding problem at the start of this project. This Chapter is an account of our efforts to carry out the first olefin metathesis reaction on a protein substrate.[†] To help frame our motivation and strategy in this endeavor, the use of olefin metathesis on peptides is reviewed next.

[†] This project was part of a productive collaboration with Yuya A. Lin, a talented and dedicated Part II undergraduate researcher I mentored in the 2007-2008 academic year. Yuya continued her D.Phil. studies in the Davis group in 2008. Experimental results that were obtained by Ms. Lin are indicated in the body of this Chapter by reference to this footnote and her contributions are stated explicitly in the experimental sections of this

Ring closing metathesis on amino acids and peptides

Ring closing metathesis allows access to at least three general structural motifs in peptide chemistry: β -turn mimics, “stapled” α -helical peptides, and carbon analogs of disulfide bridges. These structures were a proving ground for the use of olefin metathesis on peptides. They are each discussed in turn below.

Early reports by Grubbs and co-workers described the various ring sizes accessible through RCM on amino acid and peptide substrates.^{11, 12} These ring closures were typically carried out in dichloromethane using first generation catalysts such as **5.1**.⁵ These reports, along with related studies by Rutjes,¹³ established that ruthenium based olefin metathesis is compatible with peptide substrates and can provide β -turn mimics (Figure 5.2). The β -turn is a common secondary structure in proteins that is critical as a folding and recognition element.¹⁴ It was postulated that turns supported by a rigid carbon-carbon bond, rather than by hydrogen bonding or disulfide bridges,¹⁴ may impart increased conformational and metabolic stability to the peptide (Figure 5.2). This concept was later investigated by the Katzenellenbogen labs¹⁵ and has been reviewed in detail by Brik.¹⁶

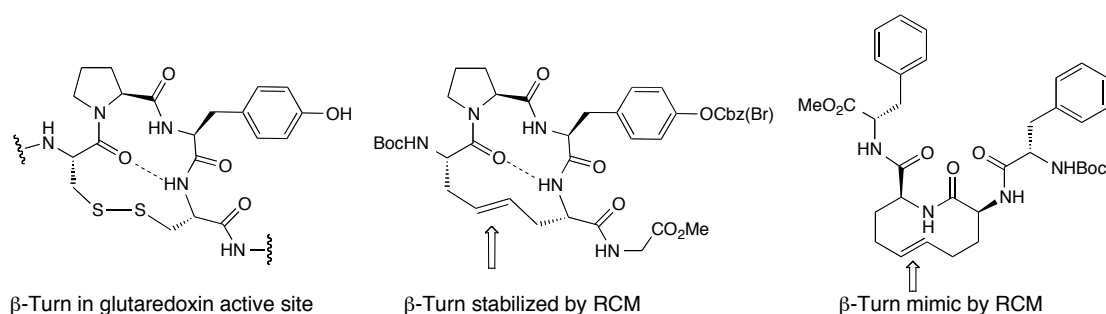


Figure 5.2: The β -turn can be stabilized^{11, 12} or mimicked¹⁵ with ring closing metathesis. The peptides shown were synthesized using Grubbs first generation catalyst (**5.1**).

β -Turn mimics installed by metathesis capture a desired peptide conformer. Ring closing metathesis can also stabilize other structural elements like the α -helix. If two alkenes are appropriately positioned on the same face of the helix, RCM can crosslink the helix through a carbon-carbon bond. It was proposed that such a tether might confer stability to the α -helix. This hypothesis was first pursued by Blackwell and Grubbs who

Chapter. For additional details of olefin metathesis on proteins that are not included here, the reader is referred to Yuya A. Lin's Part II Thesis (Keble College, University of Oxford) and the publications that resulted from these efforts (References 54, 67, and 77)

demonstrated that *O*-allyl serine (and *O*-allyl homoserine) residues at positions *i* and *i* + 4 on the peptide backbone do indeed undergo RCM and that the peptide retained its helical structure (Figure 5.3).^{17, 18} The Verdine laboratory expanded on this work and systematically investigated the position of the unmetathesized side chain and the efficiency of RCM on peptide substrates.¹⁹

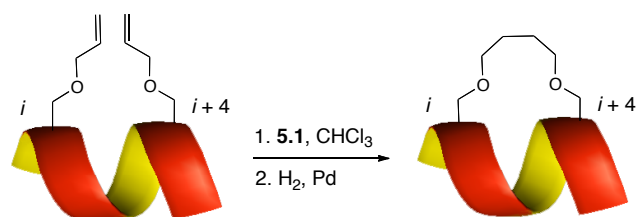


Figure 5.3: Ring closing metathesis can “staple” α -helical peptides.^{17, 18}

Disulfide bridges are a natural cross-link in peptides and proteins. Disulfides are, however, susceptible to reduction and it may be advantageous to synthesize stable analogs. Ring closing metathesis provides the opportunity to mimic disulfide bridges through a stable carbon-carbon bond. Grubbs and co-workers first proposed RCM as a route to disulfide surrogates in their studies on β -turn mimics.^{11, 12} These principles have since been applied in a number of contexts. One example is highlighted here: Vederas' synthesis of “dicarba analogues” of the peptidic hormone oxytocin (Figure 5.4).^{20, 21} Oxytocin is a cyclic peptide by virtue of an intramolecular disulfide bond. The hormone has a variety of physiological functions and is used clinically to induce labor. The half-life of the natural peptide, however, is short. The synthetic peptides prepared by Vederas, in contrast, displayed a longer *in vivo* half-life than the disulfide counterpart and were potent antagonists of the oxytocin receptors.

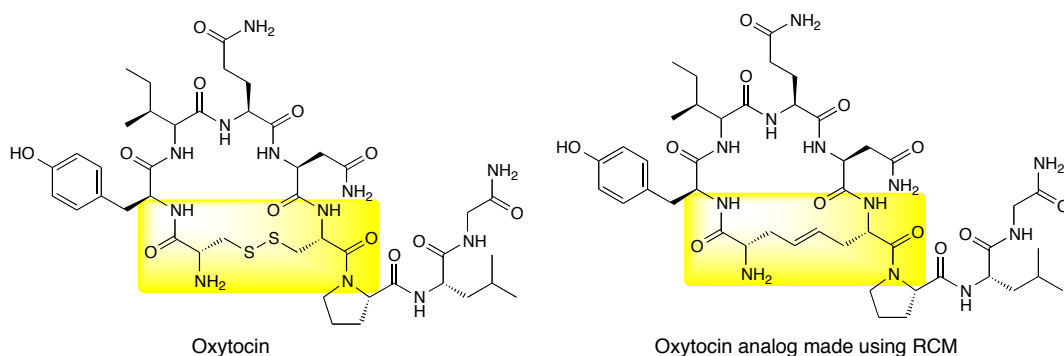


Figure 5.4: Carbon analog of oxytocin prepared by ring closing metathesis using Grubbs first generation catalyst **5.1**.^{20, 21}

From the examples cited above, it is clear that ring closing metathesis can stabilize peptide secondary structure. Olefin cross-metathesis, on the other hand, allows the appendage of other molecules to the peptide. Nature naturally modifies peptides and proteins after translation,²² and cross-metathesis offers a method to install stable analogs of such covalent modifications. Reports of cross-metathesis on peptides are discussed next.

Cross-metathesis on peptides

Cross-metathesis is developing at a rapid pace²³ and empirical guidelines have been proposed to rationalize selectivity.²⁴ This selectivity is particularly important when self-metathesis is a competing reaction. Cross-metathesis does not always benefit from proximity effects that are often inherent to the success of ring closing metathesis. For selective cross-metathesis on peptides, it is therefore essential to consider the relative reactivity of unsaturated amino acids. Studies by Gibson established that the rate of cross-metathesis generally increases with the length of the amino acid side chain (Figure 5.5). These model reactions were carried out in dichloroethane using catalyst **5.1** with 1-hexene as the cross-metathesis partner.

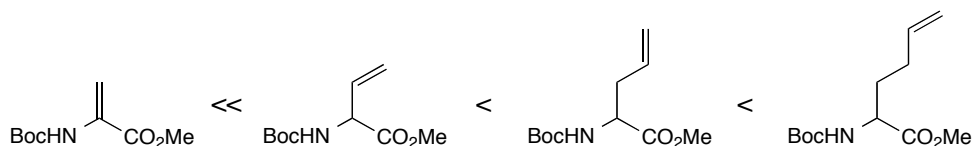


Figure 5.5: Relative reactivity of unsaturated amino acids in cross-metathesis with 1-hexene using Grubbs first generation catalyst **5.1**^{25, 26}

While these simple reactions reveal much about the reactivity of the unsaturated amino acid, biologically relevant cross-metathesis partners add another level of complexity. For instance, glycopeptides and glycoproteins mediate a variety of cellular processes²⁷⁻²⁹ and their synthesis is critical for their study.^{30, 31} Cross-metathesis allows, in principle, the synthesis of stable glycopeptide analogs. To this end, Roy has reported the cross metathesis between a variety of *C*- and *O*-linked glycosides and peptoids.³² While self-metathesis of both carbohydrate and alkene-containing peptoid compete with the desired cross-metathesis, the tolerance of **5.1** to a variety of common protecting groups was established. McGarvey later described the utility of Grubbs second generation catalyst **5.2** and Hoveyda-Grubbs catalyst **5.3** in *C*-glycopeptide synthesis.^{33, 34} In these

reports, selective cross-metathesis (over self-metathesis) was achieved by using a terminal substituted olefin on the carbohydrate metathesis substrate. Finally, Blechert has exploited the remarkable activity of the Hoveyda-Grubbs second generation catalyst **5.3**.³⁵ This catalyst allowed mild and efficient cross-metathesis between *C*-glycosides and peptides bearing an *N*-terminal acrylamide. Notably, fully deprotected carbohydrates gave useful yields of cross-metathesis product. Blechert also reported that histidine and free carboxy groups can inhibit catalysis, presumably by non-productive chelation. A selection of these glycopeptide syntheses is shown in Figure 5.6.

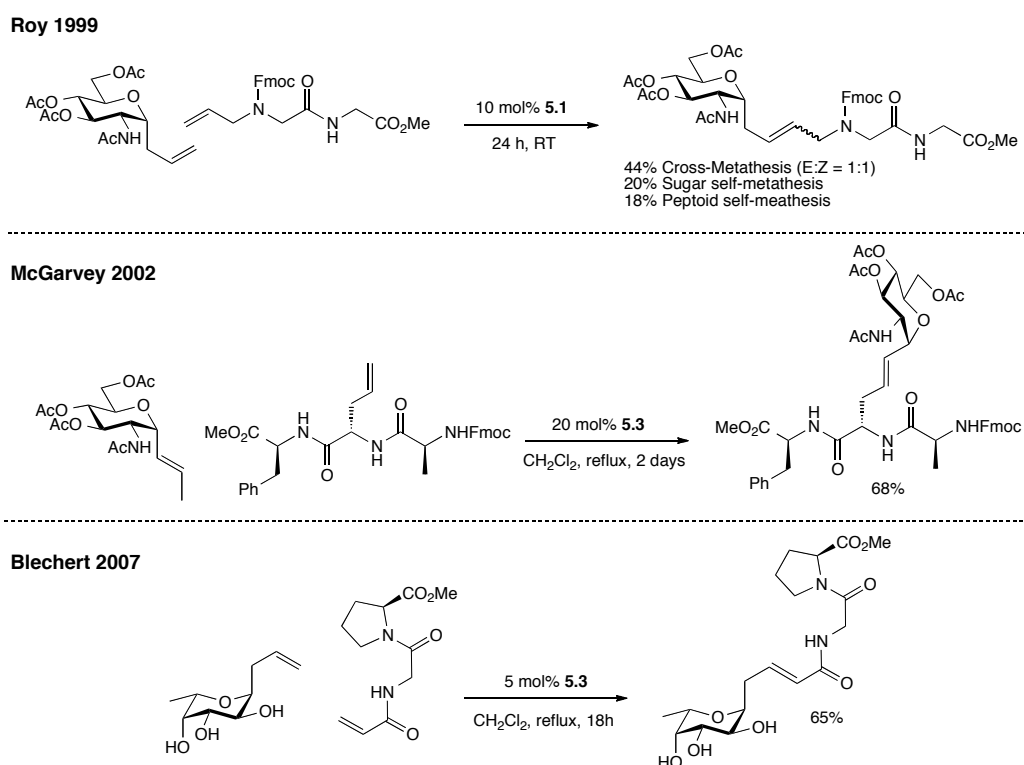


Figure 5.6: Select glycopeptide syntheses by olefin cross-metathesis^{32, 33, 35}

Thus far, the examples of olefin metathesis on peptides were invariably carried out in organic solvent at elevated temperature. For proteins, aqueous media is essential. It is therefore prudent to discuss the precedence for olefin metathesis in water.

Olefin metathesis in water

Before water-soluble olefin metathesis catalysts were available, there were several clues in the literature that ruthenium-based metathesis catalysts are active in the presence of water. In 1996, for instance, Grubbs reported ring-opening metathesis polymerization in

aqueous media using catalysts such as **5.1**.³⁶ Blechert later described the tolerance of catalysts such as **5.2** and **5.3** to air and protic solvents such as methanol and water. In Blechert's report, several examples of ring-closing metathesis were described along with cross-metathesis of simple allylic alcohols in aqueous solvents.³⁷ Raines has since described the use of dimethoxyethane or acetone as a co-solvent for olefin metathesis in homogenous aqueous media. Hoveyda-Grubbs second generation catalyst **5.3** was the most active and promoted a range of RCM transformations and the self-metathesis of allyl alcohol.³⁸ With an eye toward biocompatibility, Raines also described RCM of small molecules in the presence of Ribonuclease A. The protein was completely soluble in the dimethoxyethane/water mixture, as was catalyst **5.3**.³⁸ This observation suggests that conventional commercially available ruthenium-based olefin metathesis catalysts could be used for metathesis on protein substrates.

For more general application to protein modification, metathesis in pure water is a fundamental goal. The use of non-ionic surfactants enables use of conventional catalysts in pure water, but since these reactions occur in micelles they might not be suitable for protein modification.³⁹ On the other hand, several water-soluble catalysts have been reported that could be considered for protein modification. Representative examples are shown in Figure 5.7. All of these examples are variants of the Hoveyda-Grubbs second generation catalyst. The water-solubilizing group can be attached to the *N*-heterocyclic carbene ligand (**5.4**, **5.5**)^{40, 41} or the alkylidene (**5.5-5.8**).⁴¹⁻⁴⁴ Catalyst **5.5** bears these polar groups on both the NHC ligand and the alkylidene.⁴¹ The catalysts mediate a range of standard ring-closing metathesis reactions but their use in cross-metathesis is comparatively limited. In fact, only simple alkenols (such as allyl alcohol) were reported as cross-metathesis substrates.

The catalysts in Figure 5.7 are representative of an ever-growing interest in aqueous olefin metathesis.⁴⁵ Still, these water-soluble catalysts typically suffer from the same limited stability to water and air as conventional catalysts. It has been shown that the methyldiene intermediate in the catalytic cycle (the propagating species in metathesis of terminal olefins) is particularly prone to degradation in water and forms ruthenium hydrides.⁴⁶ These hydrides can lead to alkene isomerization and migration—complicating purification and analysis. On the other hand, these double bond isomers can be converged, in some cases, to a single alkane product by post-metathesis hydrogenation.^{47, 48} Oxygen can also result in the formation of ruthenium monocarbonyl derivatives.^{47, 48} For our goal of carrying out olefin metathesis on protein substrates, air and water

tolerance is highly desirable and it is therefore necessary to be mindful of these potential decompositions pathways.

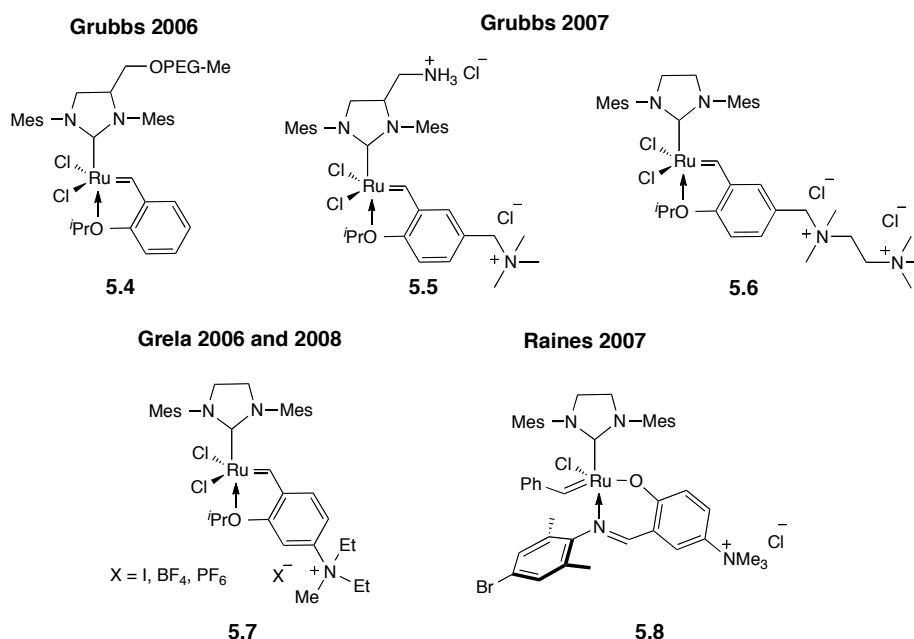


Figure 5.7: Representative water-soluble olefin metathesis catalysts⁴⁰⁻⁴⁴

Olefin metathesis on proteins

The survey above has highlighted the progress of olefin metathesis on peptides and in protic media. Our goal was to extend these preliminary studies to protein substrates. In peptides, stabilizing β -turns and helices or replacing label disulfides with carbon-carbon cross-linking is enabled by ring-closing olefin metathesis. Attaching carbohydrates and other appendages to peptides is also possible through cross-metathesis. These modifications, in principle, should be applicable to protein substrates. One can imagine, for instance, that stapling helices or replacing disulfides with carbon-carbon cross linking might lead to higher stability and thermophilicity—an common goal in industrial biosynthesis. Stable *C*-analogs of glycoproteins might also confer metabolic or therapeutic advantage in biological systems. These prospects motivate the work described in this Chapter.

The direct transition to proteins, however, is complicated by several factors. For example, the olefin metathesis reactions on peptides described above all employed organic solvent, high reaction temperatures, and prolonged reaction time. These conditions are not generally suitable for proteins which usually require mild temperature (37 °C or lower) and buffered aqueous media. For examples of metathesis in water, it is

striking to note the substrate scope: at the outset of this project, cross-metathesis in homogeneous aqueous media was limited to the self-metathesis simple alkenes such as allyl alcohol.^{37, 38, 40-45} Finally, there is the challenge to incorporate the alkene into the protein—either chemically or genetically—and the reactivity of these residues in metathesis may dictate the scope of olefin metathesis as a tool in protein modification. To adapt aqueous metathesis to proteins, expanded substrate scope and complexity in olefin metathesis was essential.

5.2 Results and Discussion

Amino acid models and the discovery of allyl sulfides as reactive substrates in olefin metathesis

The first goal in this project was to identify an amino acid residue that was active in olefin metathesis under conditions compatible with protein substrates. During these initial tests, all reactions were run open to air and below 37 °C. Mixtures of *t*BuOH and water were used as solvent and Hoveyda-Grubbs second generation catalyst (**5.3**) was employed in all cases. *t*BuOH was selected as the co-solvent because it is fully miscible in water and, unlike methanol, cannot generate ruthenium hydrides through β -hydride elimination. Hoveyda-Grubbs second generation catalyst was the catalyst of choice because of its established chemoselectivity, its relatively high stability to air and water, and its commercial availability.⁷

As a first test, cross-metathesis between homoallylglycine (Hag) derivative **5.9** and allyl alcohol was attempted. Hag was selected as a model substrate since it can be incorporated into proteins as a methionine analog in host organisms that cannot biosynthesize methionine.^{10, 49} In fact, when Hag was first incorporated to proteins genetically in 1998, Tirrell proposed this residue as a target for olefin metathesis. Nevertheless, olefin metathesis on proteins containing Hag has not been reported. Our disappointing results suggest that Hag is simply not reactive under the conditions employed: no metathesis at Hag was observed (Figure 5.8)

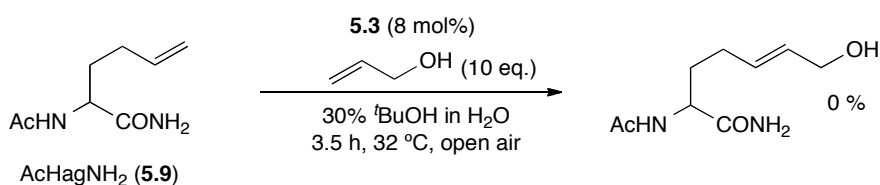


Figure 5.8: Attempted cross-metathesis using AcHagNH₂[†]

The result in Figure 5.8 does not imply that metathesis at Hag cannot be productive—after all, Gibson has shown that cross-metathesis at Hag is possible.^{25, 26} Rather, the result demonstrates that under these conditions (open air, 32 °C, aqueous solvent) the rate of metathesis at Hag is prohibitively slow.

Disappointed, but not deterred, we turned to cysteine derivatives that, if reactive in olefin metathesis, could be incorporated into proteins chemically for the sake of validating olefin metathesis in protein modification. First, we examined the reactivity of alkene-containing cysteine derivatives in olefin metathesis. Accordingly, *S*-allyl cysteine (Sac) derivative **5.10** was screened as a metathesis substrate under our restrictive conditions. The result was a surprise. The cross-metathesis product **5.11** was isolated in 56% yield after only 2.5 hours at 32 °C in a mixture of *t*BuOH and H₂O. No effort was made to exclude oxygen in this reaction.

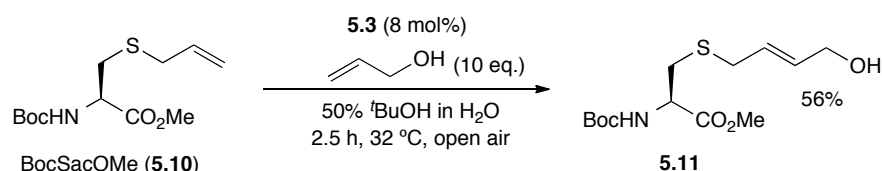


Figure 5.9: Cross-metathesis of *S*-allyl cysteine derivative **5.10** with allyl alcohol[†]

This result was probed more systematically by comparison of a series of cysteine derivatives (Figure 5.10). Remarkably, substrates derived from *S*-allyl cysteine were the only substrates that displayed any reactivity in metathesis. Extending the alkenyl chain from allyl to butenyl- and pentenyl sulfides resulted in no detectable cross-metathesis.

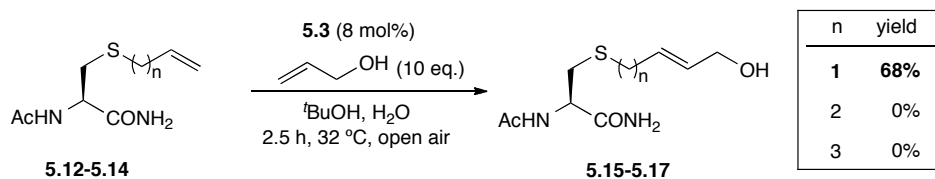


Figure 5.10: Alkenyl sulfides in cross-metathesis with allyl alcohol[†]

The results in Figures 5.9 and 5.10 were unexpected. It is well-documented that sulfides can inhibit olefin metathesis, presumably by non-productive chelation.⁵⁰⁻⁵³ This inhibition is so pronounced that Fürstner has concluded that there is a “general incompatibility of the ruthenium based metathesis catalysts with substrates containing sulfur(II) donor sites.”⁵³ Consistent with Fürstner’s observations, the butenyl- and

pentenyl cysteine derivatives **5.13** and **5.14** did not show any reaction (Figure 5.10, $n = 2$ or 3). Allyl sulfides, however, were an exception to this rule. But, in light of the inactivity of Hag derivatives in metathesis reactions under similar conditions, (Figure 5.8) it seemed that allyl sulfides are not simply tolerated by catalyst **5.3**; *allyl sulfides enhance the rate of olefin metathesis*. This rate enhancement is sufficiently high to out-compete catalyst decomposition in aqueous media and provide useful yields of cross-metathesis products. The following “sulfur-relay” mechanism⁵⁴ was proposed to account for these observations:

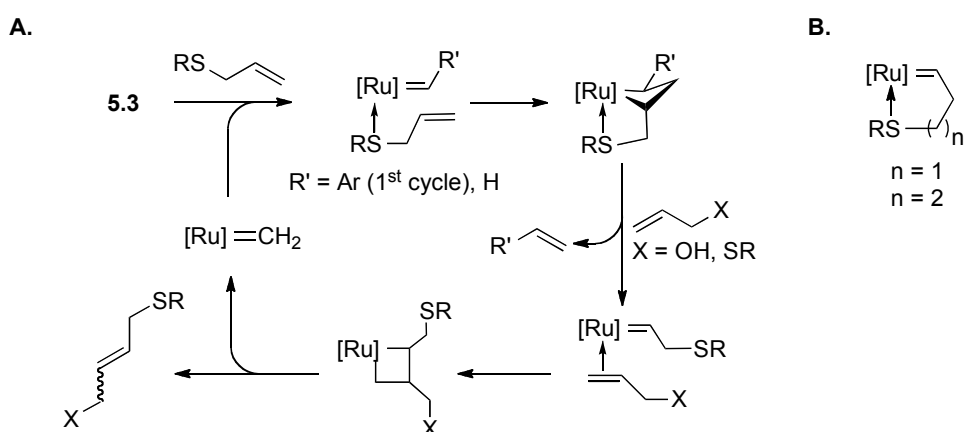


Figure 5.11: A. Allyl sulfides are reactive substrates in olefin cross-metathesis B. Butenyl and pentenyl sulfides may sequester the catalyst in non-productive chelates.

In this mechanism, the sulfide is thought to coordinate rapidly and favorably with the ruthenium catalyst since ruthenium and sulfur are soft acids and bases, respectively.⁵⁵ For Hoveyda-Grubbs second generation catalyst **5.3**, this step may involve a favorable ligand exchange from the isopropyl ether of the alkylidene to the sulfide. With the sulfide coordinated to the ruthenium center, the alkene is now proximal to the alkylidene and provokes rapid metathesis (Figure 5.11A). When the sulfide is a butenyl or pentenyl sulfide, the product of the first metathesis step results in a 5 and 6 member chelate that may be too stable for catalytic turnover (Figure 5.11B). In the case of allyl sulfides, such a chelate would be strained, facilitating ligand exchange and coordination of the cross-metathesis partner.

A number of observations by other researchers are consistent with the mechanism proposed in Figure 5.11. Hoyer, for instance, has noted that allylic alcohols enhance the rate of ring-closing metathesis, perhaps through a coordination similar to that of allyl

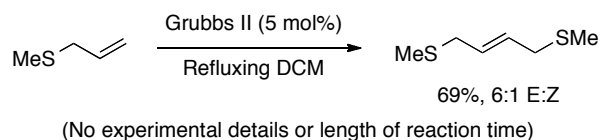
sulfides described here.⁵⁶ Fürstner has also described templated macrocycle synthesis by ring-closing metathesis where a carbonyl of the RCM substrate coordinates to the ruthenium and positions the alkene and alkylidene into close proximity.^{57, 58} Regarding the proposed chelates in Figure 5.11B., Grela has prepared metathesis catalysts with sulfide-chelated alkylidenes.⁵⁹ These catalysts generally require high temperatures (90 °C) to activate and are stable and unreactive at room temperature, consistent with the proposed inactive chelates in Figure 5.11B.⁵⁹

It should also be pointed out that allyl sulfides have been used previously in metathesis reactions. However, in these reports the remarkable reactivity of allyl sulfides in metathesis was not apparent. Grubbs, for instance, reported the self-metathesis of allyl methyl sulfide in refluxing dichloromethane. Using 5 mol% of Grubbs second generation catalyst **5.2**, the self-metathesis product was isolated in 69% yield.⁶⁰ Similarly, Heck reported that the self-metathesis product of allyl phenyl sulfide can be isolated in 75% yield after heating at 80 °C for 2 hours in toluene.⁶¹ There is no question that these results establish the compatibility of allyl sulfides with Grubbs second generation catalysts. The enhanced reactivity of allyl sulfides proposed in this Thesis, however, is not apparent from these reactions. Our mechanistic framework in Figure 5.11 should apply in both aqueous and purely organic solvents. Indeed, when the self-metathesis of allyl phenyl sulfide (**5.18**) is run in dichloromethane using 5 mol% Hoveyda-Grubbs second generation catalyst, *the self-metathesis product 5.19 was formed in 87% yield after only 5 minutes of reaction time*. Clearly such rate and yield suggest that allyl sulfides possess a remarkable reactivity in olefin metathesis that has gone overlooked in synthesis (Figure 5.12).

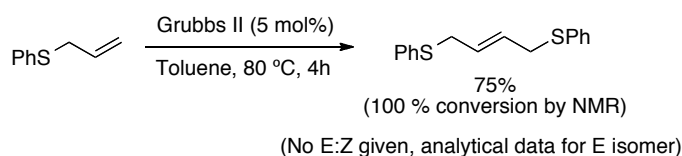
It should be made clear that the result in Figure 5.12B. was obtained with Hoveyda-Grubbs second generation catalyst **5.3** while the reactions in Figure 5.12A. were carried out using Grubbs second generation catalyst **5.2**. The result in Figure 5.12B. is consistent with rapid initiation of catalyst **5.3** which is facilitated by displacement of the ether ligand of the alkylidene by sulfur. In catalyst **5.2**, which bears a phosphine ligand, initiation is likely slower. It should also be noted that while the extremely rapid metathesis of allyl sulfides bodes well for application to biomolecules where rapid ligation is necessary to outpace catalyst decomposition, E:Z ratios are improved with longer reaction time. Thus, when the same reaction was run for two hours, the E:Z ratio of **5.19** improved to 11.1 : 1.0. This result suggests that the product of self-metathesis can re-enter the catalytic cycle and the double bond geometry can be equilibrated by extended reaction

time. The product could be further purified by crystallization and single crystals were prepared for X-ray analysis to confirm the *trans* double bond geometry (Figure 5.12).

A. Grubbs 2001



Heck 2002



B. This Thesis

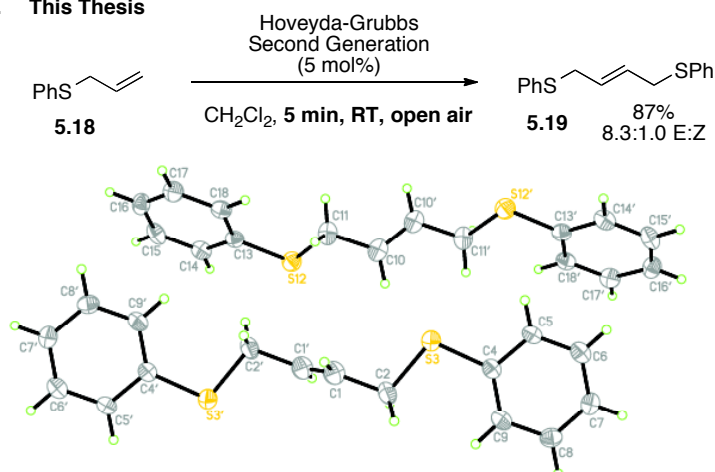


Figure 5.12: **A.** Demonstration of allyl sulfide tolerance in metathesis reactions using Grubbs second generation catalyst. **B.** Demonstration of the high reactivity of allyl sulfides in metathesis reactions using Hoveyda-Grubbs second generation catalyst. Single crystal X-ray analysis was carried out on the major *trans* isomer (all diffraction data is tabulated in the experimental section).

Olefin metathesis on proteins

By identifying *S*-allyl cysteine as a reactive substrate for olefin metathesis, we arrived at a fortuitous intersection of the dehydroalanine campaign discussed in Chapter 2 and our interest in olefin metathesis on proteins. With ready access to dehydroalanine, we could install *S*-allyl cysteine by the conjugate addition of allyl thiol. Direct allylation with allyl

chloride was later shown to be a viable route to *S*-allyl cysteine (*vide infra*) but initially we opted for the dehydroalanine route to avoid unselective allylation. Using *O*-mesitylenesulfonylhydroxylamine (MSH), dehydroalanine was installed in the model protease subtilisin from *Bacillus lentus*. This protein has a single cysteine and its conversion to dehydroalanine was discussed in detail in Chapter 2. Addition of allyl thiol provided *S*-allyl cysteine on the protein surface (**5.20**, Figure 5.13).

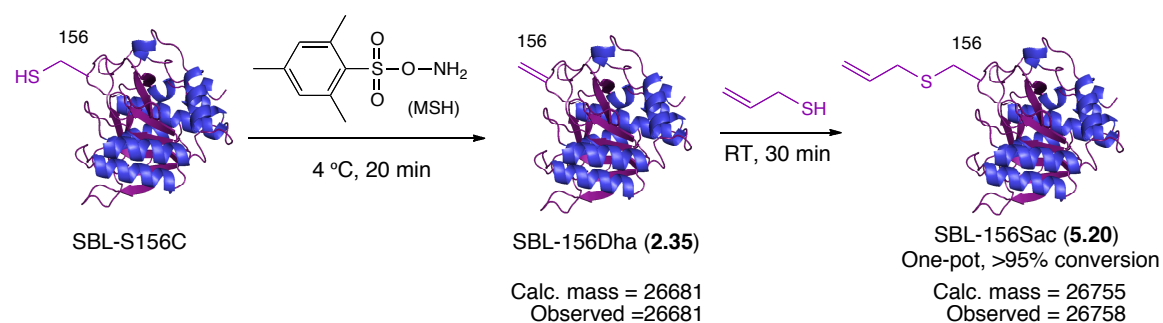


Figure 5.13: Installation of *S*-allyl cysteine into a model protein

With an allyl sulfide installed on a protein substrate, olefin metathesis could be attempted. First attempts were made with allyl alcohol using Hoveyda-Grubbs second generation catalyst **5.3**. An excess of **5.3** (200 equiv.) and allyl alcohol (10,000 equiv.) were used in initial attempts. The large excess of reagents used in protein modification can be removed by dialysis or size exclusion chromatography so for exploratory work this was not necessarily a problem. The reaction was run open air at room temperature at pH 8.0. No cross-metathesis was observed with allyl alcohol, but a peak of *M*+406 was observed that did not change after 5 hours of reaction time. Suspecting that this adduct was a metalloprotein that was inactive because of chelation, MgCl₂ was added as a Lewis acid. It has been shown that when olefin metathesis is hindered by substrate chelation, the addition of a Lewis acid can disrupt this coordination and allow metathesis to proceed.^{58, 62, 63} It has also been shown that MgCl₂ was an essential additive for other metal mediated protein modifications that could be hindered by non-productive chelation.^{64, 65} Gratifyingly, when MgCl₂ was included in the reaction mixture (10,000 equiv. relative to protein), cross-metathesis was observed with over 90% conversion in 5 hours at room temperature (Figure 5.14). Importantly, the protein product was still active as a protease. The model protein cleaved the protease substrate SucAAPFpNA and released *p*-

nitroanilide, indicating that the protein was stable to the reaction conditions (See experimental section for protease activity assay).

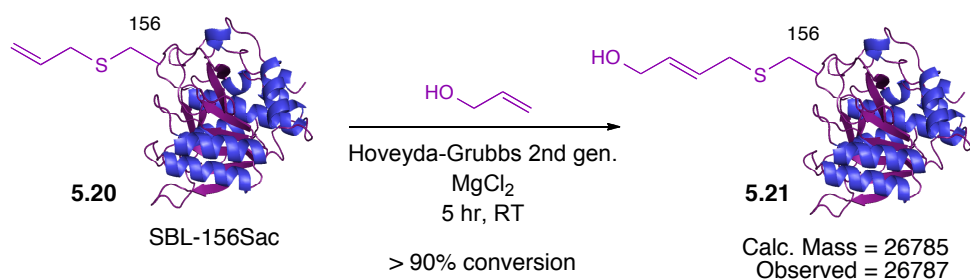


Figure 5.14: First cross-metathesis on a protein substrate⁵⁴

With conditions established for successful cross-metathesis on a protein surface, other cross-metathesis partners were screened. Carbohydrates and oligoethylene glycols were synthesized and subjected to the metathesis reaction with the model protein (Figure 5.15). Carbohydrates are involved in a variety of cell-signaling events and protein folding²⁷⁻²⁹ and the synthesis of glycoproteins is a current and necessary challenge to understand their biological role.^{30, 31} Attaching polyethyleneglycol (PEG) to a protein increases its effective molecular mass and confers many advantages in the circulatory half-life of protein therapeutics.⁶⁶ The results of these conjugation experiments are summarized in Figure 5.15.

Allyl glycosides (**5.26** and **5.27**) and allyl PEG derivatives (**5.28** and **5.29**) underwent productive cross-metathesis with protein **5.20**. Longer reaction times did not improve conversions. While the conversions in Figure 5.15 are modest, these reactions were the first demonstrations of olefin metathesis on protein substrates. The high reactivity of allyl sulfides in olefin metathesis enables these reactions. While in these examples *S*-allyl cysteine was installed using dehydroalanine, other methods can be envisioned that may be more general and allow a more thorough investigation of the scope and limitations of olefin metathesis on proteins. In the final section of this Chapter, multiple complementary methods for the incorporation of *S*-allyl cysteine into proteins are described.

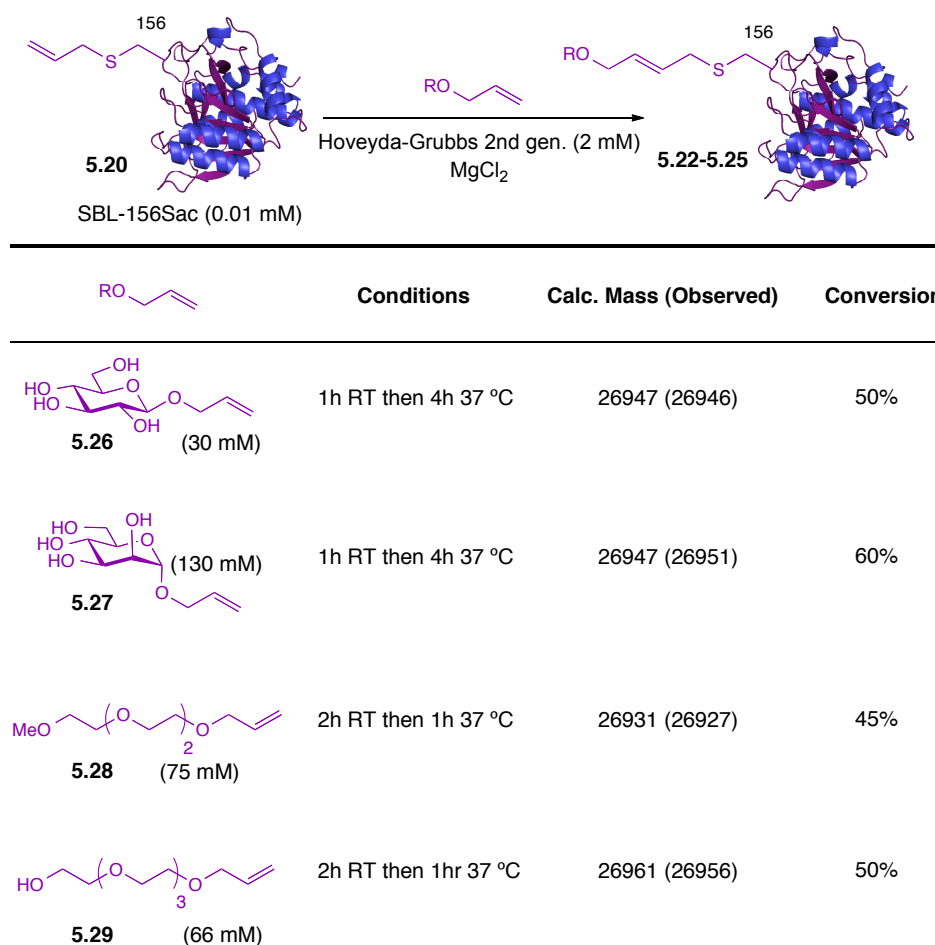


Figure 5.15: Cross-metathesis between a model protein and carbohydrate and PEG metathesis partners⁵⁴

Complementary methods for incorporating *S*-allyl cysteine into proteins

The conversion of dehydroalanine to *S*-allyl cysteine enabled the first examples of olefin metathesis on a protein substrate.⁵⁴ The conjugate addition of thiols to dehydroalanine, however, likely results in a mixture of diastereomers. The direct allylation of cysteine provides the same reactive amino acid side chain as a single diastereomer. Two methods were explored for the direct allylation of cysteine: reaction with allyl chloride and reaction with allyl selenocyanate.

Allylation of cysteine with allyl chloride

It was initially thought allylation with allyl chloride would suffer from non-selective reaction with other nucleophilic side chains. However, it was found that the highly nucleophilic cysteine sulfhydryl can be allylated selectively at pH 8.0. The reaction is

shown in Figure 5.16. First, any contaminant disulfide was reduced with DTT and then allyl chloride was added as a solution in DMF. The transfer of an allyl group was observed by LC-MS after only 30 minutes at 37 °C. The allylated product did not react with Ellman's reagent, indicating that the reaction was at the single cysteine of the SBL-S156C mutant. The product of allylation was also active in olefin metathesis and full conversion was observed in the cross-metathesis with allyl alcohol (Figure 5.16).

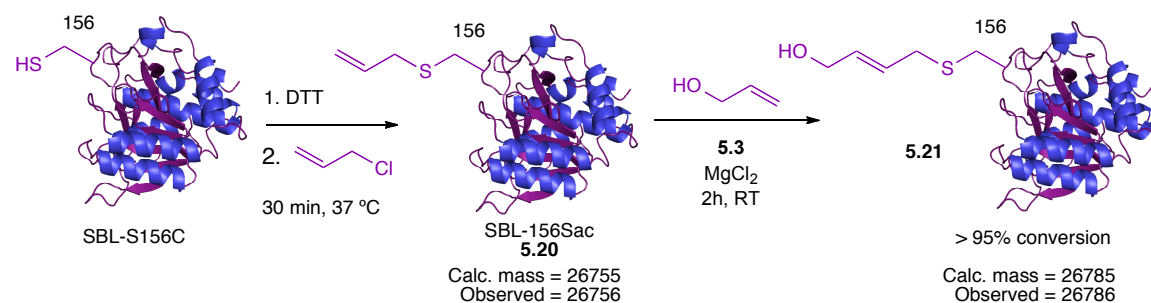


Figure 5.16: Cysteine allylation with allyl chloride and use in olefin cross-metathesis⁶⁷

Allylation of cysteine with allyl selenocyanate

The allylation of cysteine with allyl chloride is a simple, direct, and selective method to access *S*-allyl cysteine in proteins. While non-selective allylation was not observed on the model protein in Figure 5.16, the allylation of more hindered cysteine residues would be best accomplished by a *specific* allylation. That is, a transformation that can only react with cysteine and no other residue. Crich's dechalcogenative rearrangement of allylic disulfides and allyl selenenylsulfides are two reactions that could satisfy this criterion.⁶⁸⁻⁷¹ These transformations draw upon reports by Baldwin⁷² and Sharpless⁷³ that described the dechalcogenative contraction of allylic disulfides and diselenides, to the thio- and selenoethers, respectively (Figure 5.17).

For the rearrangements shown in Figure 5.17, the addition of a thio- or selenophilic nucleophile (PPh_3) is required for the dechalcogenation in some cases (Figure 5.17 A, C, D). In other cases, the extrusion of sulfur or selenium is spontaneous (Figure 5.17 B, E). It was not clear at the outset which method would most reliably allylate cysteine. Therefore both the allylic disulfide and allyl selenenylsulfide were investigated on the model protein SBL-S156C.

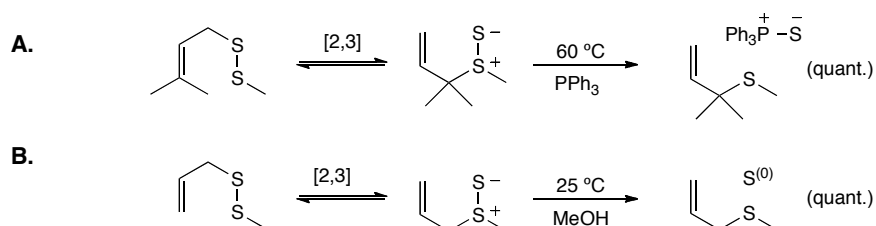
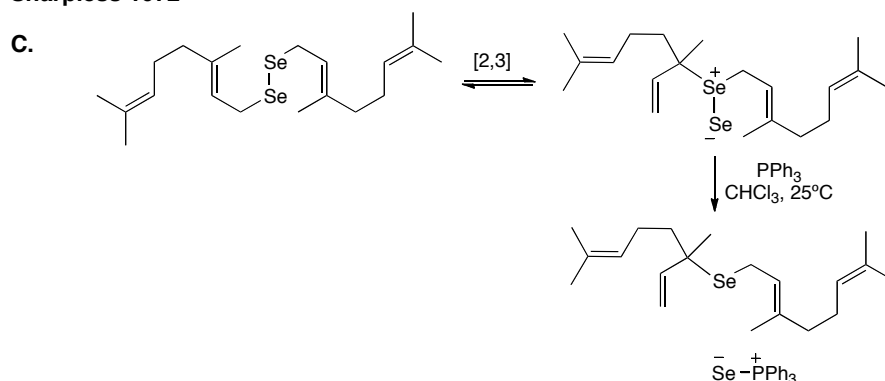
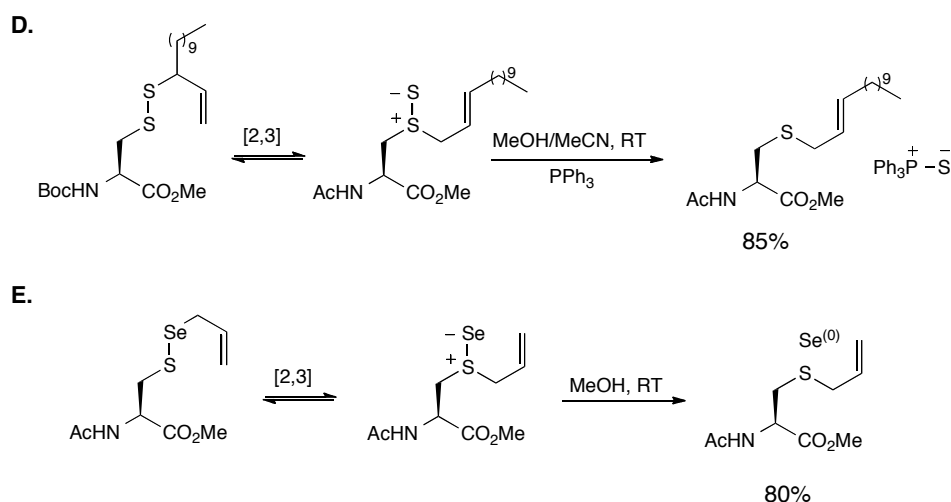
Baldwin 1971**Sharpless 1972****Crich 2006, 2007, 2008**

Figure 5.17: Dechalcogenative rearrangement of allylic disulfides, allylic diselenides, and allyl selenenyldisulfides

For the synthesis of allyl cysteine disulfide, SBL-S156C was treated with allyl phenylthiosulfonate (**5.30**). Reagent **5.30** is a member of a class of reagents that have been used previously to make mixed disulfides on protein surfaces.⁷⁴ The disulfide **5.31** was formed efficiently and was stable at room temperature; no spontaneous desulfurization was observed over several hours by LC-MS. In an attempt to provoke the desired desulfurization, protein **5.31** was treated with a water-soluble analog of triphenyl

phosphine (TPPTS, Figure 5.18). However, only rapid reduction to the free cysteine was observed. Apparently the rearrangement of the disulfide is not favored or TPPTS is too nucleophilic and only attacks the disulfide.

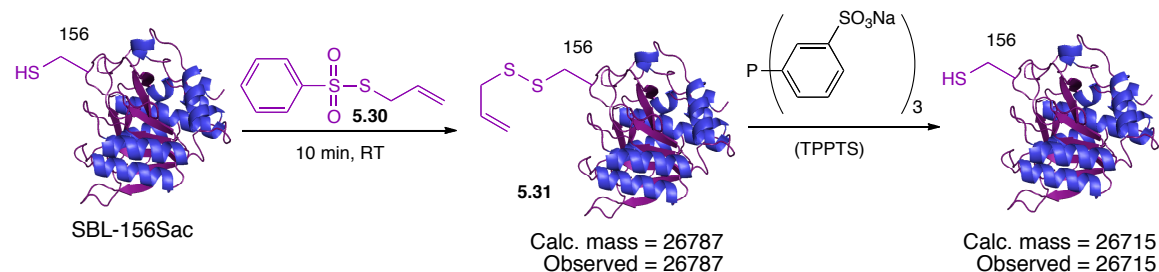


Figure 5.18: Synthesis and attempted desulfurization of allyl cysteine disulfide⁶⁷

While the rearrangement of disulfide **5.31** proved recalcitrant, Crich has shown that the analogous rearrangement of allyl selenenylsulfides is comparatively facile and can proceed at room temperature. Moreover, the extrusion of selenium is spontaneous and the addition of triphenylphosphine is not required (Figure 5.17E).^{68, 70} To test the dechalcogenative rearrangement of the allyl selenenylsulfide on a protein, SBL-S156C was treated with allyl selenocyanate (**5.32**). Within 10 minutes at room temperature, all starting material was consumed. A mixture of the allyl selenenylsulfide **5.33** and allyl sulfide **5.20** were observed by LC-MS (Figure 5.19A). After an hour at room temperature, the rearrangement and loss of selenium was complete, providing *S*-allyl cysteine. The product of the reaction did not react with Ellman's reagent, indicating that all cysteine was consumed in the reaction. The product **5.20** was also used in a model cross-metathesis with allyl alcohol (Figure 5.19B).

The efficient and specific allylation of cysteine using allyl selenocyanate should enable further exploration of olefin metathesis on protein substrates. This allylation also complements direct allylation with allyl chloride and the addition of allyl thiol to dehydroalanine. Since all three routes to *S*-allyl cysteine are mechanistically distinct, they can be deployed in a manner that is most suitable for the protein target.

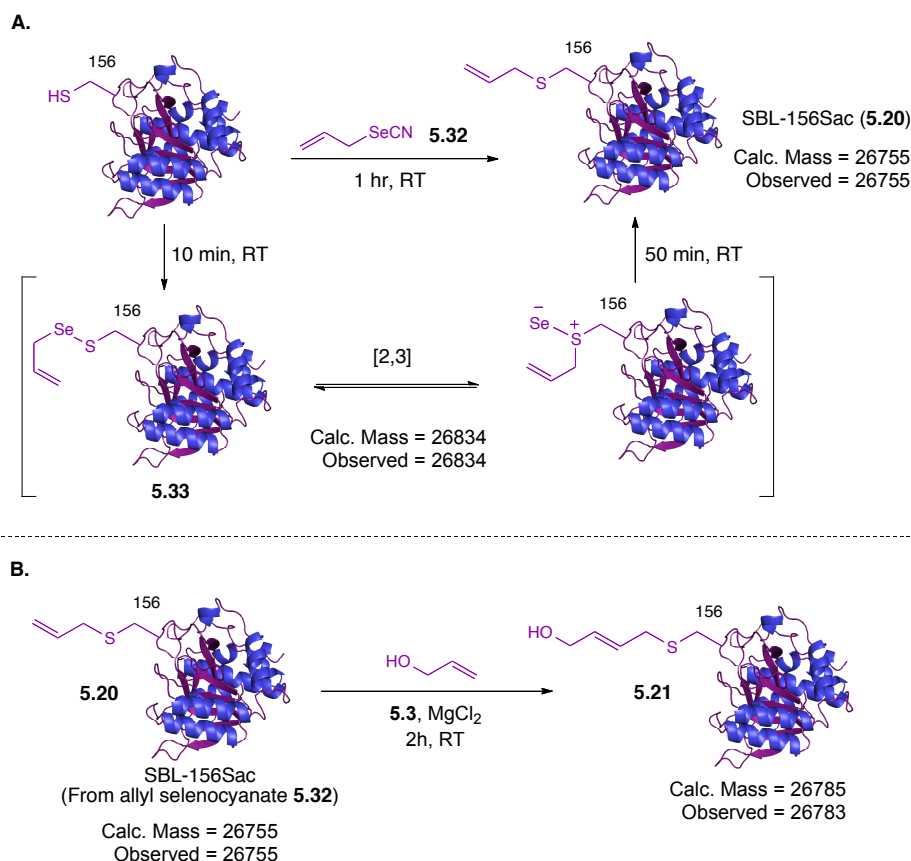


Figure 5.19: Cysteine specific allylation with allyl selenocyanate (**5.32**) and use in a model cross-metathesis reaction.⁶⁷

As a final example of the incorporation of *S*-allyl cysteine into protein substrates, a single cysteine mutant of green fluorescent protein (Y39C) was allylated using allyl selenocyanate **5.32**. Again, efficient allylation was observed within 1 hour at room temperature to provide protein **5.34** (Figure 5.20). Reaction at cysteine was verified by Ellman's assay. Also, protein fluorescence was maintained throughout the allylation, indicating the protein retained tertiary structure under the reaction conditions. Attempts at cross-metathesis on this allylated protein, however, were unsuccessful. We tentatively attribute this result to steric hindrance after inspection of the crystal structure of SBL and GFP. On one hand, if the cysteine on GFP is hindered, the efficient allylation with allyl selenocyanate is noteworthy. On the other hand, this result prompts several questions regarding the scope of olefin metathesis in protein modification: What role does sterics play in successful modification? Can metathesis be executed at a hindered helix or active site? What is the optimal distance from the surface of the protein for efficient metathesis? Studies are ongoing to answer these questions.

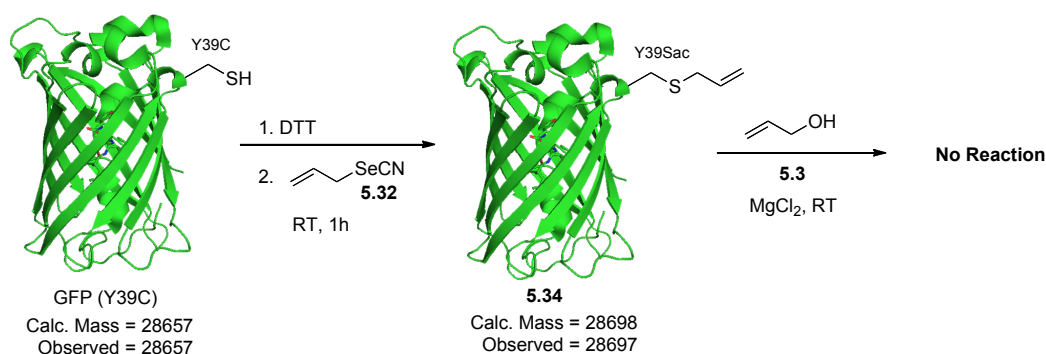


Figure 5.20: Allylation of GFP (Y39C) with allyl selenocyanate and attempted cross-metathesis

As a final note, demonstrating the dechalcogenative allylation of cysteine on multiple protein substrates should encourage use of the Crich ligation in other venues of protein modification. The allylation is extremely fast and specific for cysteine. While we have restricted the application of this reaction to the synthesis of *S*-allyl cysteine to investigate olefin metathesis, Crich has shown this method can be used in a variety of allylations on peptide substrates.⁶⁸⁻⁷¹

5.3 Conclusions and Outlook

In this chapter, the remarkable reactivity of allyl sulfides in olefin metathesis was described. This reactivity was used to enable the first examples of olefin metathesis on protein substrates. Multiple methods for the incorporation of *S*-allyl cysteine into proteins were reported, along with their subsequent reaction with alkene-containing carbohydrates and oligoethyleneglycols using Hoveyda-Grubbs second generation catalyst **5.3**.

Since the publication of the results in this chapter,^{9, 54, 67} interest in applying olefin metathesis to protein substrates has grown.^{75, 76} The Schultz group, for instance, has disclosed a method for the genetic incorporation of crotyl serine into proteins for the purpose of post-translational ring-closing metathesis.⁷⁶ In this report, the authors attempted to evolve a synthetase to incorporate *S*-allyl cysteine, but apparently this amino acid is not metabolically stable in yeast, and instead the synthetase was evolved to incorporate the *S*-allyl cysteine sulfoxide (the apparent product of metabolism). In their report, our conditions for olefin metathesis were adopted for the successful reaction at crotyl serine. In related efforts, and concurrent with studies on the chemical incorporation of *S*-allyl cysteine, our group attempted incorporation of *S*-allyl cysteine as a methionine surrogate in Met-auxotrophic *E. coli*.^{54†} (The synthesis of deprotected *S*-allyl cysteine is

reported in the experimental section of this Chapter). Unambiguous incorporation was verified by tryptic digest and MS-MS of expressed protein, but the incorporation levels were prohibitively low for studies in olefin metathesis.^{54†} These results highlight the unsolved problem of high fidelity genetic incorporation of the reactive allyl sulfide.

At the outset of this project, cross-metathesis in homogenous aqueous media was limited to simple alkenes like allyl alcohol. The cross-metathesis on proteins reported in this chapter sets a new benchmark for complexity in olefin metathesis. Nonetheless, the scope and limitations of olefin metathesis on proteins remains to be determined. For instance, the examples of olefin metathesis described in this Chapter were carried out on a single protein on a highly exposed *S*-allyl cysteine residue. Moreover, the cross-metathesis partners were largely restricted to simple allyl ethers. Work is ongoing to generalize olefin metathesis as a method for protein modification and determine the factors that govern successful cross-metathesis and what role the secondary structure of the protein plays in successful metathesis. As part of these efforts, it has since been shown that not just allyl sulfides, but allyl chalcogenides in general, are reactive substrates in olefin metathesis.⁷⁷ The reactivity in metathesis was found to increase from allyl ether to allyl sulfide to allyl selenide.⁷⁷ Indeed the chemical incorporation of allyl selenides into proteins greatly expands the metathesis partners useful in protein modification.⁷⁷ These results further motivate the need to incorporate such reactive residues into proteins genetically.

While our primary objective was to adapt the olefin metathesis reaction to protein modification, the reactivity of allyl sulfides in olefin metathesis is of general synthetic interest. Allyl sulfides are widely used in synthesis and now olefin metathesis should be considered a valuable transformation for these synthons. Soon after our preliminary reports, several groups have incorporated allyl sulfides into their synthetic designs in order to take advantage of their high reactivity in olefin metathesis.⁷⁸⁻⁸⁰ These applications of allyl sulfides are diverse: asymmetric synthesis,⁷⁸ dynamic combinatorial library synthesis,⁷⁹ and peptide modification⁸⁰ all benefited from the high reactivity of allyl sulfides in olefin metathesis.

There are many opportunities for olefin metathesis in protein modification. Stapled helices or induced β -turns may allow unprecedented design of protein secondary and tertiary structure. Cross-metathesis may allow the synthesis of biological probes or tracking of proteins *in vivo*. Many of these applications require the genetic incorporation of allyl chalcogenides and so constitute an opportunity for protein engineering. These

opportunities and challenges are driving our current research and the results of these efforts will be published in due course.

5.4 References

- (1) Grubbs, R. H. *Handbook of Metathesis*. Wiley-VCH: Weinheim, Germany, 2003.
- (2) Fürstner, A. Olefin metathesis and beyond. *Angew. Chem. Int. Ed.* **2000**, *39*, 3012-3043.
- (3) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Metathesis reactions in total synthesis. *Angew. Chem. Int. Ed.* **2005**, *44*, 4490-4527.
- (4) Hoveyda, A. H.; Zhugralin, A. R. The remarkable metal-catalysed olefin metathesis reaction. *Nature* **2007**, *450*, 243-251.
- (5) Schwab, P.; Grubbs, R. H.; Ziller, J. W. Synthesis and applications of $\text{RuCl}_2(=\text{CHR}')(\text{PR}_3)_2$: the influence of the alkylidene moiety on metathesis activity. *J. Am. Chem. Soc.* **1996**, *118*, 100-110.
- (6) Scholl, M.; Ding, S.; Woo Lee, C.; Grubbs, R. H. Synthesis and activity of a new generation of ruthenium-based olefin metathesis catalysts coordinated with 1,3-dimesityl-4,5-dihydroimidazol-2-ylidene ligands. *Org. Lett.* **1999**, *1*, 953-956.
- (7) Garber, S. B.; Kingsbury, J. S.; Gray, B. L.; Hoveyda, A. H. Efficient and recyclable monomeric and dendritic Ru-based metathesis catalysts. *J. Am. Chem. Soc.* **2000**, *122*, 8168-8179.
- (8) Binder, J. B.; Raines, R. T. Olefin metathesis for chemical biology. *Curr. Opin. Chem. Biol.* **2008**, *12*, 767-773.
- (9) Lin, Y. A.; Chalker, J. M.; Davis, B. G. Olefin metathesis for site-selective protein modification. *ChemBioChem* **2009**, *10*, 959-969.
- (10) van Hest, J. C. M.; Tirrell, D. A. Efficient introduction of alkene functionality into proteins in vivo. *FEBS Lett.* **1998**, *428*, 68-70.
- (11) Miller, S. J.; Grubbs, R. H. Synthesis of conformationally restricted amino acids and peptides employing olefin metathesis. *J. Am. Chem. Soc.* **1995**, *117*, 5855-5856.
- (12) Miller, S. J.; Blackwell, H. E.; Grubbs, R. H. Application of ring-closing metathesis to the synthesis of rigidified amino acids and peptides. *J. Am. Chem. Soc.* **1996**, *118*, 9606-9614.

- (13) Rutjes, F. P. J. T.; Schoemaker, H. E. Ruthenium-catalyzed ring closing olefin metathesis of non-natural α -amino acids. *Tetrahedron Lett.* **1997**, *38*, 677-680.
- (14) Rose, G. D.; Glerasch, L. M.; Smith, J. A. Turns in peptides and proteins. *Adv. Protein Chem.* **1985**, *37*, 1-109.
- (15) Fink, B. E.; Kym, P. R.; Katzenellenbogen, J. A. Design, synthesis, and conformational analysis of a proposed type I β -turn mimic. *J. Am. Chem. Soc.* **1998**, *120*, 4334-4344.
- (16) Brik, A. Metathesis in peptides and peptidomimetics. *Adv. Synth. Catal.* **2008**, *350*, 1661-1675.
- (17) Blackwell, H. E.; Grubbs, R. H. Highly efficient synthesis of covalently cross-linked peptide helices by ring-closing metathesis. *Angew. Chem. Int. Ed.* **1998**, *37*, 3281-3284.
- (18) Blackwell, H. E.; Sadowsky, J. D.; Howard, R. J.; Sampson, J. N.; Chao, J. A.; Steinmetz, W. E.; O'Leary, D. J.; Grubbs, R. H. Ring-closing metathesis of olefinic peptides: design, synthesis, and structural characterization of macrocyclic helical peptides. *J. Org. Chem.* **2001**, *66*, 5291-5302.
- (19) Schafmeister, C. E.; Po, J.; Verdine, G. L. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J. Am. Chem. Soc.* **2000**, *122*, 5891-5892.
- (20) Stymiest, J. L.; Mitchell, B. F.; Wong, S.; Vederas, J. C. Synthesis of biologically active dicarba analogues of the peptide hormone oxytocin using ring-closing metathesis. *Org. Lett.* **2003**, *5*, 47-49.
- (21) Stymiest, J. L.; Mitchell, B. F.; Wong, S.; Vederas, J. C. Synthesis of oxytocin analogues with replacement of sulfur by carbon gives potent antagonists with increased stability. *J. Org. Chem.* **2005**, *70*, 7799-7809.
- (22) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372.
- (23) Connon, S. J.; Blechert, S. Recent developments in olefin cross-metathesis. *Angew. Chem. Int. Ed.* **2003**, *42*, 1900-1923.
- (24) Chatterjee, A. K.; Choi, T.-L.; Sanders, D. P.; Grubbs, R. H. A general model for selectivity in olefin cross metathesis. *J. Am. Chem. Soc.* **2003**, *125*, 11360-11370.
- (25) Gibson, S. E.; Gibson, V. C.; Keen, S. P. Cross metathesis of the amino acid homoallylglycine. *Chem. Commun.* **1997**, 1107-1108.

- (26) Biagini, S. C. G.; Gibson, S. E.; Keen, S. P. Cross-metathesis of unsaturated α -amino acid derivatives. *J. Chem. Soc. Perkin Trans. I* **1998**, 2485-2499.
- (27) *Essentials of Glycobiology*; 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press, 2009.
- (28) Dwek, R. A. Glycobiology: toward understanding the function of sugars. *Chem. Rev.* **1996**, *96*, 683-720.
- (29) Ohtsubo, K.; Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **2006**, *126*, 855-867.
- (30) Davis, B. G. Synthesis of glycoproteins. *Chem. Rev.* **2002**, *102*, 579-601.
- (31) Gamblin, D. P.; Scanlan, E. M.; Davis, B. G. Glycoprotein synthesis: an update. *Chem. Rev.* **2009**, *109*, 131-163.
- (32) Hu, Y.-J.; Roy, R. Cross-metathesis of N-alkenyl peptoids with O- or C-allyl glycosides. *Tetrahedron Lett.* **1999**, *40*, 3305-3308.
- (33) McGarvey, G. J.; Benedum, T. E.; Schmidtman, F. W. Development of co- and post-translational synthetic strategies to C-neoglycopeptides. *Org. Lett.* **2002**, *4*, 3591-3594.
- (34) Schmidtman, F. W.; Benedum, T. E.; McGarvey, G. J. The preparation of C-glycosyl amino acids—an examination of olefin cross-metathesis. *Tetrahedron Lett.* **2005**, *46*, 4677-4681.
- (35) Brüchner, P.; Koch, D.; Voigtman, U.; Blechert, S. Cross-metathesis of C-glycosides and peptides. *Synth. Commun.* **2007**, *37*, 2757-2769.
- (36) Lynn, D. M.; Kanaoka, S.; Grubbs, R. H. Living ring-opening metathesis polymerization in aqueous media catalyzed by well-defined ruthenium carbene complexes. *J. Am. Chem. Soc.* **1996**, *118*, 784-790.
- (37) Connon, S. J.; Rivard, M.; Zaja, M.; Blechert, S. Practical olefin metathesis in protic media under an air atmosphere. *Adv. Synth. Catal.* **2003**, *345*, 572-575.
- (38) Binder, J. B.; Blank, J. J.; Raines, R. T. Olefin metathesis in homogeneous aqueous media catalyzed by conventional ruthenium catalysts. *Org. Lett.* **2007**, *9*, 4885-4888.
- (39) Lipshutz, B. H.; Aguinaldo, G. T.; Ghorai, S.; Voigtritter, K. Olefin cross-metathesis reactions at room temperature using the nonionic amphiphile "PTS": just add water. *Org. Lett.* **2008**, *10*, 1325-1328.

- (40) Hong, S. H.; Grubbs, R. H. Highly active water-soluble olefin metathesis catalyst. *J. Am. Chem. Soc.* **2006**, *128*, 3508-3509.
- (41) Jordan, J. P.; Grubbs, R. H. Small-molecule N-heterocyclic-carbene-containing olefin-metathesis catalysts for use in water. *Angew. Chem. Int. Ed.* **2007**, *46*, 5152-5155.
- (42) Michrowska, A.; Gułajski, Ł.; Kaczmarska, Z.; Mennecke, K.; Kirschning, A.; Grela, K. A green catalyst for green chemistry: synthesis and application of an olefin metathesis catalyst bearing a quaternary ammonium group. *Green Chem.* **2006**, *8*, 685-688.
- (43) Gułajski, Ł.; Michrowska, A.; Narożnik, J.; Kaczmarska, Z.; Rupnicki, L.; Grela, K. A highly active aqueous olefin metathesis catalyst bearing a quaternary ammonium group. *ChemSusChem* **2008**, *1*, 103-109.
- (44) Binder, J. B.; Guzei, I. A.; Raines, R. T. Salicylaldimine ruthenium alkylidene complexes: metathesis catalysts tuned for protic solvents. *Adv. Synth. Catal.* **2007**, *349*, 395-404.
- (45) Burtscher, D.; Grela, K. Aqueous olefin metathesis. *Angew. Chem. Int. Ed.* **2009**, *48*, 442-454.
- (46) Lynn, D. M.; Mohr, B.; Grubbs, R. H.; Henling, L. M.; Day, M. W. Water-soluble ruthenium alkylidenes: synthesis, characterization, and application to olefin metathesis in protic solvents. *J. Am. Chem. Soc.* **2000**, *122*, 6601-6609.
- (47) Dinger, M. B.; Mol, J. C. Degradation of the first-generation Grubbs metathesis catalyst with primary alcohols, water, and oxygen. Formation and catalytic activity of ruthenium(II) monocarbonyl species. *Organometallics* **2003**, *22*, 1089-1095.
- (48) Dinger, M. B.; Mol, J. C. Degradation of the second-generation Grubbs metathesis catalyst with primary alcohols and oxygen— isomerization and hydrogenation activities of monocarbonyl complexes. *Eur. J. Inorg. Chem.* **2003**, 2827-2833.
- (49) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J. Am. Chem. Soc.* **2000**, *122*, 1282-1288.
- (50) Armstrong, S. K.; Christie, B. A. Synthesis of sulphur-containing heterocycles by ring closing diene metathesis. *Tetrahedron Lett.* **1996**, *37*, 9373-9376.

- (51) Shon, Y.-S.; Lee, T. R. Catalytic ring-closing olefin metathesis of sulfur-containing species: heteroatom and other effects. *Tetrahedron Lett.* **1997**, *38*, 1283-1286.
- (52) Mascareñas, J. L.; Rumbo, A.; Castedo, L. A practical approach to *cis*-2,5-disubstituted tetrahydrofurans and O-bridged medium-sized carbocycles from [5+2] pyrone-alkene cycloadducts. *J. Org. Chem.* **1997**, *62*, 8620-8621.
- (53) Fürstner, A.; Seidel, G.; Kindler, N. Macrocycles by ring-closing-metathesis, XI: syntheses of (*R*)-(+)-lasiodiopodin, zeranol and truncated salicylihalamides. *Tetrahedron* **1999**, *55*, 8215-8230.
- (54) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. Allyl sulfides are privileged substrates in aqueous cross-metathesis: application to site-selective protein modification. *J. Am. Chem. Soc.* **2008**, *130*, 9642-9643.
- (55) Pearson, R. G. Hard and soft acids and bases. *J. Am. Chem. Soc.* **1963**, *85*, 3533-3539.
- (56) Hoye, T. R.; Zhao, H. Some allylic substituent effects in ring-closing metathesis reactions: allylic alcohol activation. *Org. Lett.* **1999**, *1*, 1123-1125.
- (57) Fürstner, A.; Langemann, K. Macrocycles by ring-closing metathesis. *Synthesis* **1997**, 792-803.
- (58) Fürstner, A.; Langemann, K. Total syntheses of (+)-ricinelaidic acid lactone and of (–)-gloeosporone based on transition-metal-catalyzed C-C bond formations. *J. Am. Chem. Soc.* **1997**, *119*, 9130-9136.
- (59) Tzur, E.; Szadkowska, A.; Ben-Asuly, A.; Makal, A.; Goldberg, I.; Woźniak, K.; Grela, K.; Lemcoff, N. G. Studies on electronic effects in O-, N-, and S-chelated ruthenium olefin-metathesis catalysts. *Chem. Eur. J.* **2010**, *16*, 8726-8737.
- (60) Toste, F. D.; Chatterjee, A. K.; Grubbs, R. H. Functional group diversity by ruthenium-catalyzed olefin cross-metathesis. *Pure Appl. Chem.* **2002**, *74*, 7-10.
- (61) Spagnol, G.; Heck, M.-P.; Nolan, S. P.; Mioskowski, C. Efficiency of a ruthenium catalyst in metathesis reactions of sulfur-containing compounds. *Org. Lett.* **2002**, *4*, 1767-1770.
- (62) Wipf, P.; Weiner, W. S. Enantioselective synthesis and photoracemization studies of (+)-2-cyclopropyl-7,8-dimethoxy-2*H*-chromene-5-carboxylic acid methyl ester, an advanced intermediate of a dihydrofolate reductase inhibitor. *J. Org. Chem.* **1999**, *64*, 5321-5324.

- (63) Yang, Q.; Xiao, W.-J.; Yu, Z. Lewis acid assisted ring-closing metathesis of chiral diallylamines: an efficient approach to enantiopure pyrrolidine derivatives. *Org. Lett.* **2005**, *7*, 871-874.
- (64) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. Regioselective carbon-carbon bond formation in proteins with palladium catalysis; new protein chemistry by organometallic chemistry. *ChemBioChem* **2006**, *7*, 134-139.
- (65) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. Site-specific functionalization of proteins by organopalladium reactions. *ChemBioChem* **2007**, *8*, 232-238.
- (66) Harris, J. M.; Chess, R. B. Effect of pegylation on pharmaceuticals. *Nat. Rev. Drug Discovery* **2003**, *2*, 214-221.
- (67) Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. Enabling olefin metathesis on proteins: chemical methods for installation of *S*-allyl cysteine. *Chem. Commun.* **2009**, 3714-3716.
- (68) Crich, D.; Krishnamurthy, V.; Hutton, T. K. Allylic selenosulfide rearrangement: a method for chemical ligation to cysteine and other thiols. *J. Am. Chem. Soc.* **2006**, *128*, 2544-2545.
- (69) Crich, D.; Brebion, F.; Krishnamurthy, V. Allylic disulfide rearrangement and desulfurization: mild, electrophile-free thioether formation from thiols. *Org. Lett.* **2006**, *8*, 3593-3596.
- (70) Crich, D.; Krishnamurthy, V.; Brebion, F.; Karatholuvhu, M.; Subramanian, V.; Hutton, T. K. Dechalcogenative allylic selenosulfide and disulfide rearrangements: complementary methods for the formation of allylic sulfides in the absence of electrophiles. Scope, limitations, and application to the functionalization of unprotected peptides in aqueous media. *J. Am. Chem. Soc.* **2007**, *129*, 10282-10294.
- (71) Crich, D.; Yang, F. Synthesis of neoglycoconjugates by the desulfurative rearrangement of allylic disulfides. *J. Org. Chem.* **2008**, *73*, 7017-7027.
- (72) Höfle, G.; Baldwin, J. E. Thiosulfoxides. The intermediates in rearrangement and reduction of allylic disulfides. *J. Am. Chem. Soc.* **1971**, *93*, 6307-6308.

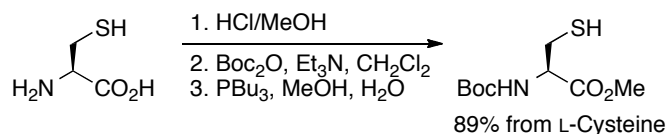
- (73) Sharpless, K. B.; Lauer, R. F. Facile thermal rearrangements of allyl selenides and diselenides. [1,3] and [2,3] shifts. *J. Org. Chem.* **1972**, *37*, 3973-3974.
- (74) Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. Glycosyl phenylthiosulfonates (Glyco-PTS): novel reagents for glycoprotein synthesis. *Org. Biomol. Chem.* **2003**, *1*, 3642-3644.
- (75) Kirshenbaum, K.; Arora, P. S. Cross-dressing proteins by olefin metathesis. *Nat. Chem. Biol.* **2008**, *4*, 527-528.
- (76) Ai, H.-w.; Shen, W.; Brustad, E.; Schultz, P. G. Genetically encoded alkenes in yeast. *Angew. Chem. Int. Ed.* **2010**, *49*, 935-937.
- (77) Lin, Y. A.; Chalker, J. M.; Davis, B. G. Olefin cross-metathesis on proteins: investigation of allylic chalcogen effects and guiding principles in metathesis partner selection. *J. Am. Chem. Soc.* **2010**, *132*, 16805-16811.
- (78) Zheng, S.; Gao, N.; Liu, W.; Liu, D.; Zhao, X.; Cohen, T. Regio- and enantioselective iridium-catalyzed allylation of thiophenol: synthesis of enantiopure allyl phenyl sulfides. *Org. Lett.* **2010**, *12*, 4454-4457.
- (79) Hunter, L.; Condie, G. C.; Harding, M. M. Reversible aqueous metathesis reactions for potential application in dynamic combinatorial chemistry. *Tetrahedron Lett.* **2010**, *51*, 5064-5067.
- (80) Alam, J.; Keller, T. H.; Loh, T.-P. Functionalization of peptides and proteins by Mukaiyama aldol reaction. *J. Am. Chem. Soc.* **2010**, *132*, 9546-9548.

5.5 Experimental Procedures

General Considerations

See Chapter 2 Experimental Procedures for general considerations in synthesis and protein analysis.

N-Boc-L-Cysteine methyl ester (BocCysOMe)

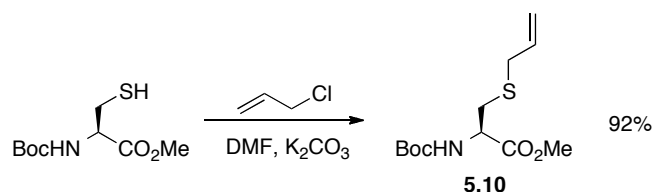


Anhydrous MeOH (100 mL) was added to a flame dried 250 mL round bottom flask. The solvent was stirred and cooled to 0 °C and acetyl chloride (248 mmol, 17.6 mL) was added dropwise over 5 minutes. The solution was stirred an additional 10 minutes at 0 °C to give a concentrated solution of HCl. L-Cysteine (2.00 g, 16.51 mmol) was then added in one portion and the flask flushed briefly with argon. The ice bath was removed and the reaction was stirred at room temperature for 24 hours. The solvent was then removed under reduced pressure to give the crude cysteine methyl ester hydrochloride as a pale yellow solid. This material was used immediately in the next step without purification. The crude ester was suspended in CH₂Cl₂ (100 mL) and cooled to 0 °C. Et₃N (5.06 mL, 36.3 mmol) was added carefully followed by di-*tert*-butyl dicarbonate (Boc₂O, 4.32 g, 19.81 mmol). The reaction was stirred at room temperature for 3.25 hours after which time TLC (30% EtOAc in Petrol) revealed the desired product (*R*_f = 0.6) and its corresponding disulfide (*R*_f = 0.3). The solvent was removed under reduced pressure and the resulting residue was redissolved in MeOH (40 mL) and H₂O (8 mL). Tributylphosphine (2.0 mL, 8.1 mmol) was added dropwise to the stirred solution. TLC revealed reduction of the disulfide. The reaction was diluted with Et₂O (100 mL) and H₂O (50 mL). The organic layer was separated and the aqueous layer was extracted with Et₂O (2 × 50 mL). The combined organics were washed with brine (100 mL), dried over MgSO₄, and filtered. The solvent was removed under reduced pressure and the residue purified by column chromatography eluting first with 5% EtOAc in petrol and then 20% EtOAc in petrol. The titled compound was isolated as a clear oil (3.48 g, 89% from L-cysteine). $[\alpha]_D^{20} = +28.3$ (*c* = 7.5, CHCl₃) (Lit.¹: $[\alpha]_D^{20} = +28.5$, *c* = 7.5, CHCl₃). IR (*ν*_{max}, film): 2979, 1717, 1506, 1367, 1166. ¹H NMR (400 MHz, CDCl₃): δ = 1.42 (10H,

s, includes Boc and SH), 2.94 (2H, app. td, $J = 4.3, 8.7$, CH_2SH), 3.76 (3H, s, CO_2CH_3), 4.58 (1H, m, $\text{H}\alpha$), 5.44 (1H, d, $J = 5.8$, NH). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 27.3$ (CH_2SH), 28.2 (Boc), 52.6 (CO_2CH_3), 54.8 ($\text{C}\alpha$), 80.2 (Boc), 155.1, 170.8 ($2 \times \text{C}=\text{O}$). LRMS m/z (ESI $^-$): 234.1 $[\text{M}-\text{H}]^-$.

If PBU_3 was not used to reduce disulfide, the reaction mixture could be purified by column chromatography to give BocCysOMe and the corresponding disulfide. Data for disulfide: Small white needles: m.p. = 89-90 °C (Lit.² = 96-97 °C). $[\alpha]_D^{20} = +91.4$ ($c = 1.0$, CHCl_3). IR (ν_{max} , KBr): 3377, 2983, 1748, 1686, 1515, 1367, 1167. ^1H NMR (400MHz, CDCl_3): $\delta = 1.45$ (18H, s, $2 \times \text{Boc}$), 3.16 (4H, d, $J = 5.1$ Hz, $2 \times \text{CH}_2\text{S}$), 3.77 (6H, s, $2 \times \text{CO}_2\text{CH}_3$), 4.60 (2H, m, $2 \times \text{H}\alpha$), 5.40 (2H, d, $J = 7.6$, $2 \times \text{NH}$). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 28.3$ (Boc), 41.3 (CH_2S), 52.6 (CO_2CH_3), 52.8 ($\text{C}\alpha$), 80.3 (Boc), 155.0, 171.2 ($\text{C}=\text{O}$). LRMS m/z (ESI $^+$): 469.2 $[\text{M}+\text{H}]^+$.

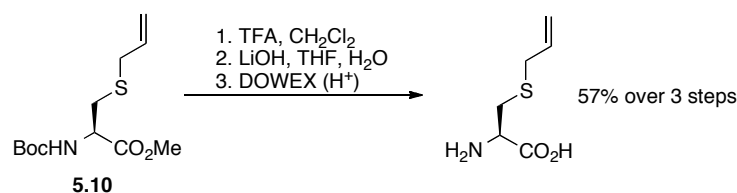
***N*-Boc-*S*-allyl-cysteine methyl ester (**5.10**)**



BocCysOMe (16.75 g, 71.17 mmol) was added to a 250 mL round bottom flask and placed under an atmosphere of argon before dissolving in DMF (71 mL). K_2CO_3 (24.59 g, 178 mmol) was added to the stirred solution and the mixture was cooled to 0 °C. Allyl chloride (27.54 mL, 356 mmol) was added by syringe, the ice bath was removed, and the reaction was stirred vigorously at room temperature for 15 hours. TLC (25% EtOAc in petrol) indicated complete consumption of BocCysOMe ($R_f = 0.5$) and formation of the allylated product **5.10** ($R_f = 0.6$). The reaction was diluted with 400 mL each of Et_2O and H_2O and separated. The organic layer was washed with H_2O (2×200 mL) and then brine (2×200 mL). The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The product was purified by column chromatography (15% EtOAc in petrol) yielding 18.05 g of **5.10** as a clear oil (92%) which solidified to waxy prisms upon storage at -20 °C. m.p. = 37-38 °C. $[\alpha]_D^{20} = +13.5$ ($c = 1.0$, CHCl_3). IR (ν_{max} , film): 3375, 2978, 1747, 1715, 1503, 1367, 1166. ^1H NMR (400 MHz, CDCl_3): $\delta = 1.46$ (9H, s, Boc), 2.83-2.96 (2H, ABX System, $J = 14.0, 5.5, 4.8$, CH_2Sallyl), 3.09-3.19 (2H, m,

SCH₂CH=CH₂), 3.77 (3H, s, CO₂CH₃), 4.53 (1H, m, H α), 5.11-5.15 (2H, m, HC=CH₂), 5.33 (1H, d, J = 6.8, NH), 5.76 (1H, m, HC=CH₂). ¹³C NMR (100 MHz, CDCl₃): δ = 28.3 (Boc), 32.8 (CH₂Sallyl), 35.1 (SCH₂CH=CH₂), 52.5 (CO₂CH₃), 53.1 (C α), 80.0 (Boc), 117.8 (HC=CH₂), 133.6 (HC=CH₂), 155.1, 171.6 (2 $\times$ C=O). HRMS m/z (ESI⁺): Found 298.1083 [M+Na]⁺; C₁₂H₂₁NO₄SNa requires 298.1089. Analysis for C₁₂H₂₁NO₄S requires C, 52.34%; H, 7.69%; N, 5.09%; Found: C, 51.94%; H, 7.62%; N, 4.73%.

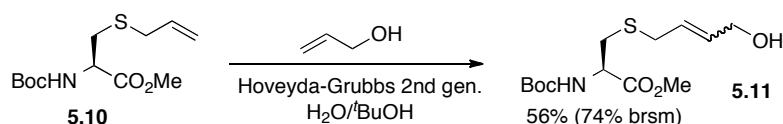
S-Allylcysteine (Sac)



BocSacOMe (9.92 g, 36.0 mmol) was added to a 250 mL round bottom flask and dissolved in CH₂Cl₂ (100 mL). The solution was placed under argon and TFA (10 mL) was added at room temperature. The reaction was stirred at room temperature for 3 hours after which time TLC (15% EtOAc in petrol) indicated consumption of **5.10** (R_f = 0.4). The solvent was then removed by rotary evaporation and the resulting yellow residue was dried briefly under vacuum. The viscous residue was then dissolved in THF (30 mL) and the stirred solution was cooled to 0 °C. LiOH (72 mL of a 5M solution) was then poured into the reaction mixture. The ice bath was removed and the reaction stirred at room temperature for 1 hour after which time the reaction was diluted with 100 mL H₂O and then neutralized (pH < 7, pH paper) with DOWEX® 50WX8 (H⁺, ~120 g). All of the resin was poured into an empty column and washed with 500 mL of H₂O (gravity flow). The flow-through was discarded. The product was eluted with 5% NH₄OH (aq) and the fractions containing the titled compound were collected [R_f = 0.6; 7:2:1 *i*PrOH:MeOH:NH₄OH (25% aq)]. The product was concentrated by rotary evaporation to give a yellow solid which was then purified by column chromatography [7:2:1 *i*PrOH:MeOH:NH₄OH (25% aq.)] to give Sac as white crystals (3.29 g, 57%). m.p. = 212-213 °C. [α]_D²⁰ = -13.7 (c = 1.0, H₂O). IR ($\nu_{\max}$, KBr): 1505, 1396, 1300, 993, 921, 842, 545, 524. ¹H NMR (400 MHz, D₂O): δ = 2.57-2.70 (2H, ABX system, J = 13.3, 5.3, 6.7, CH₂Sallyl), 3.06 (2H, d, J = 7.2, CH₂CH=CH₂), 3.28 (1H, dd, J = 6.7, 5.3, H α), 5.02-5.09 (2H, m, HC=CH₂), 5.71 (1H, m, HC=CH₂). ¹³C NMR (100 MHz, D₂O): δ = 31.2 (CH₂Sallyl), 34.2 (CH₂CH=CH₂), 53.9 (C α), 118.6 (HC=CH₂), 133.8 (HC=CH₂), 173.7

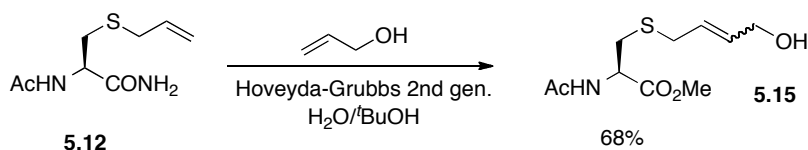
(C=O). HRMS m/z (ESI⁺): Found 184.0403 [M+Na]⁺; C₆H₁₁NO₂SNa requires 184.0408. Analysis for C₆H₁₁NO₂S requires: C, 44.70%; H, 6.88%; N, 8.69%; Found: C, 44.74%; H, 6.66%; N, 8.38%. Note: As with other deprotected amino acids, ¹H and ¹³C shifts in D₂O are dependent on sample concentration.

Cross-metathesis between BocSacOMe and allyl alcohol



This experiment was performed by Angel Lin (Part II, 2007) and is included here as a representative example of olefin metathesis of allyl sulfides in aqueous media.

A solution of Hoveyda-Grubbs 2nd generation catalyst (12 mg, 0.02 mmol) in ^tBuOH (2.0 mL) was added to a stirred mixture containing BocSacOMe (126 mg, 0.46 mmol) and allyl alcohol (0.16 mL, 2.3 mmol) in ^tBuOH:H₂O (1:2, 6 mL) at 32 °C. A second dose of the catalyst solution (12 mg in 2.0 mL ^tBuOH) and allyl alcohol (0.16 mL, 2.29 mmol) followed by 2.0 mL of water was added one hour later. After an additional 1.5 hours of reaction time, the brown mixture was concentrated under reduced pressure. The resulting residue was purified by column chromatography (50% EtOAc in petrol) to afford recovered BocSacOMe (30 mg, 24%) and cross-metathesis product (78 mg, 56%, 74% brsm). IR ($\nu_{\max}$, film): 3376, 2978, 1713, 1513, 1367, 1166. ¹H NMR (400 MHz, CDCl₃): δ = 1.43 (9H, s, Boc), 2.49 (1H, br s, OH), 2.76-2.94 (2H, ABX System, J = 13.8, 5.8, 5.0, H β), 3.14 (2H, d, J = 7.1 Hz, SCH₂CH=), 3.75 (3H, s, CO₂CH₃), 4.11 (2H, d, J = 3.8 Hz, CH₂OH), 4.41-4.56 (1H, m, H α), 5.35 (1H, d, J = 7.8 Hz, NH), 5.58-5.70 (1H, m, SCH₂CH=), 5.70-5.83 (1H, m, =CHCH₂OH). ¹³C NMR (100 MHz, CDCl₃): δ = 28.3 (Boc), 33.0 (C β), 34.03 (SCH₂CH=), 52.6 (CO₂CH₃), 53.3 (C α), 62.8 (CH₂OH), 80.4 (Boc), 127.3 (=CHCH₂OH), 133.0 (SCH₂CH=), 155.2, 171.7 (2 $\times$ C=O). HRMS m/z (ESI⁺): found 328.1189 [M+Na]⁺; C₁₁H₂₃NO₅Na requires 328.1195.

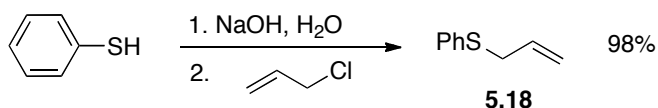


This experiment was performed by Angel Lin (Part II, 2007) and is included here as a representative example of olefin metathesis of allyl sulfides in aqueous media. For the synthesis of 5.12, the reader is referred to the published procedure.³

A solution of Hoveyda-Grubbs 2nd generation catalyst **5.3** (6.5 mg, 0.01 mmol) in ^tBuOH (1.0 mL) was added to a stirred mixture containing amide **5.12** (52 mg, 0.26 mmol) and allyl alcohol (0.10 mL, 1.47 mmol) in water (4 mL). The light green mixture was heated to 32 °C and was allowed to stir for 1 hour at which time TLC indicated the formation of a product ($R_f = 0.3$, EtOAc:ⁱPrOH:H₂O = 8:2:0.5). An additional portion of catalyst solution (6.5 mg **5.3** in 1.0 mL ^tBuOH) and allyl alcohol (0.10 mL, 1.47 mmol) was added at this time. After an additional 1.5 hours of reaction time, no **5.12** was detected by TLC and the reaction was concentrated under reduced pressure to yield a dark brown residue. Purification by column chromatography (EtOAc:ⁱPrOH:H₂O = 8:2:0.5) afforded cross-metathesis product **5.15** (42 mg, 68%). This reaction was run under identical conditions for substrates **5.13** and **5.14**. In these cases, no cross-metathesis product was observed and only starting material was recovered.

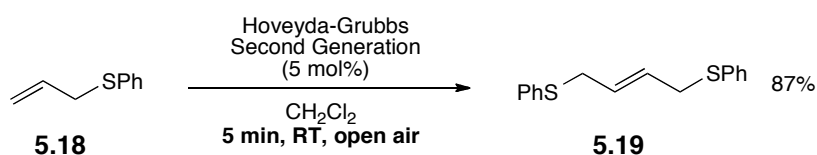
Data for **5.15**: IR ($\nu_{\max}$, film): 3301, 1659, 1543, 1416, 1297, 1089, 1002. ¹H NMR (400 MHz, CD₃OD): δ = 2.02 (3H, s, NHAc), 2.70 (1H, dd, J = 14.0, 8.1, H β), 2.92 (1H, dd, J = 13.9, 5.8, H β'), 3.20 (2H, d, J = 7.1, SCH₂CH=), 4.07 (2H, d, J = 5.1, CH₂OH), 4.51 (1H, dd, J = 8.1, 5.8, H α), 5.63-5.73 (1H, m, SCH₂CH=), 5.73-5.85 (1H, m, =CHCH₂OH). ¹³C NMR (100 MHz, CD₃OD): δ = 21.6 (NHAc), 32.4 (C β), 33.3 (SCH₂CH=), 52.9 (C α), 62.0 (CH₂OH), 127.2 (SCH₂CH=), 132.9 (=CHCH₂OH), 172.3, 174.5 (2 $\times$ C=O). HRMS m/z (ESI⁺): Found 255.0774 [M+Na]⁺; C₉H₁₆N₂NaO₃S requires 255.0779.

Allyl phenyl sulfide (**5.18**)



NaOH (1.56 g, 39.0 mmol) was added to a 100 mL round bottom flask and dissolved in 50 mL of deionized H₂O. The stirred solution was cooled to 0 °C and thiophenol (4.0 mL, 39.0 mmol) was added. After 5 minutes of stirring, allyl chloride (3.3 mL, 42.9 mmol) was added. The ice bath was removed and the resulting cloudy reaction mixture was stirred vigorously for 6 hours at room temperature. TLC (10% EtOAc in petrol) indicated the formation of a single product ($R_f = 0.7$). The reaction was diluted with 250 mL Et₂O and washed successively with 0.1 M NaOH (3 $\times$ 100 mL) and brine (200 mL). The

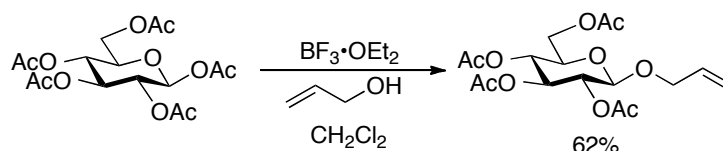
organic layer was dried over MgSO_4 , filtered, and concentrated under reduced pressure to give allyl phenyl sulfide as a clear oil that required no purification (5.73 g, 98%). Spectroscopic data was consistent with that previously reported.⁴ IR (ν_{max} , film): 3078, 2916, 1944, 1852, 1636, 1584, 1480, 1438, 1229, 1089, 919, 737, 690. ^1H NMR (400 MHz, CDCl_3): δ = 3.59 (2H, d, J = 6.8, PhSCH_2), 5.12 (1H, dd, J = 10.0, 1.3, $\text{HC}=\text{CHH}$), 5.18 (1H, dd, J = 16.9, 1.3, $\text{HC}=\text{CHH}$), 5.93 (1H, m, $\text{HC}=\text{CH}_2$), 7.12-7.41 (5H, m, Ar). ^{13}C NMR (100 MHz, CDCl_3): δ = 37.2 (PhSCH_2), 117.7 ($\text{HC}=\text{CH}_2$), 126.3, 128.9, 129.8, 133.6 (includes 3° Ar and $\text{HC}=\text{CH}_2$), 136.0 (4° Ar). LRMS (ESI): 148.9 $[\text{M}-\text{H}]^-$.



Allyl phenyl sulfide (195 mg, 1.30 mmol) was added to a 50 mL round bottom flask and dissolved in 13 mL dichloromethane (HPLC grade). Hoveyda-Grubbs second generation catalyst (41 mg, 0.065 mmol) was added to the stirred solution. The reaction was stirred, open to air, at room temperature for 5 minutes. After this time, the solvent was removed by rotary evaporation (bath temperature = 23 °C) and the resulting residue was purified directly by column chromatography (5% EtOAc in petrol). The cross-metathesis product (R_f = 0.2, 5% EtOAc in petrol) was isolated as a clear oil which crystallized on standing (154 mg, 87% yield). The ratio of *trans*:*cis* was 8.3 : 1.0, as determined by the relative integration of the allylic protons. Spectroscopic data was also consistent with that previously reported for the *trans* isomer.⁵ Note: when the reaction was run under identical conditions for 2 hours, the ratio of olefin isomers was 11.1 : 1.0 and the yield was 92%. This result suggests that while most of the allyl phenyl sulfide is consumed in the first few minutes of the reaction, the self-metathesis product can re-enter the catalytic cycle and the double bond isomers can equilibrate. Data for *trans* isomer. m.p. = 65-67 °C (analytical sample crystallized from EtOAc and petrol). IR (ν_{max} , KBr): 2360, 2341, 1749, 1437, 1223, 1093, 963, 893. ^1H NMR (400 MHz, CDCl_3): (Data for *trans* isomer) δ = 3.50 (4H, dd, J = 3.8, 1.8, $2 \times \text{PhSCH}_2$), 5.64 (2H, m, J = 3.8, 1.8, $\text{HC}=\text{CH}$), 7.20-7.35 (10H, m, ArH). ^{13}C NMR (100 MHz, CDCl_3): (Data for *trans* isomer) δ = 35.9 (PhSCH_2), 126.2, 128.6, 128.8, 129.9 (includes CH_{Ar} and $\text{HC}=\text{CH}$), 135.7 (4° Ar). HRMS m/z (EI^+): Found 272.0697 $[\text{M}]^+$; $\text{C}_{16}\text{H}_{16}\text{NS}_2$ requires 272.0693.

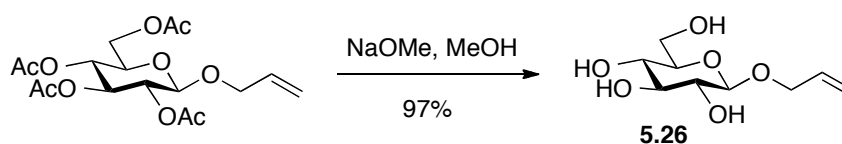
The *trans* product could be purified by recrystallization from dichloromethane and petrol. Single crystals were then prepared by a second recrystallization from dichloromethane and petrol. X-ray diffraction data was collected and analyzed by Dr. Amber Thompson of the University of Oxford crystallography service. Atomic coordinates and diffraction data is tabulated at the end of this chapter.

Allyl 2,3,4,6-tetra-*O*-acetyl- β -(D)-glucopyranoside



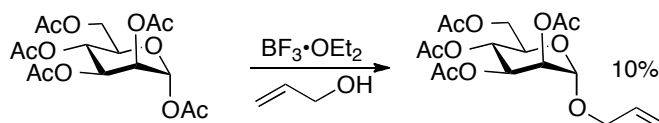
β -(D)-glucose pentaacetate (10.00 g, 25.60 mmol) was added to a 250 mL flame dried round bottom flask under argon and dissolved in CH_2Cl_2 (60 mL). The stirred solution was cooled to 0 °C and $\text{BF}_3\cdot\text{OEt}_2$ (4.87 mL, 38.43 mmol) was added by syringe. After stirring 10 minutes at 0 °C, allyl alcohol (2.61 mL, 38.43 mmol) was added. The ice bath was removed after completion of the addition and the reaction stirred at room temperature for 8.5 hours. The reaction was then cooled to 0 °C and quenched with NaHCO_3 (sat. aq., 50 mL). After dilution with H_2O (50 mL), the organic layer was separated and the aqueous layer was extracted with CH_2Cl_2 (2×100 mL). The combined organic layers were washed with brine (100 mL), dried over MgSO_4 , filtered, and concentrated under reduced pressure. The product was purified by column chromatography (40% EtOAc in petrol) to give the titled compound as a bright white solid (6.14 g, 62%). This material was spectroscopically identical to that previously reported.⁶ m.p. = 73-75 °C. $[\alpha]_D^{20} = -13.4$ ($c = 2.0$, CHCl_3). IR (ν_{max} , KBr): 1755, 1373, 1232, 1042. ^1H NMR (400 MHz, CDCl_3): $\delta = 1.91, 1.94, 1.96, 1.99$ (12H, 4 $\times$ s, 4 $\times$ OAc), 3.62 (1H, ddd, $J = 9.9, 4.8, 2.5$, H5), 4.02 (1H, dd, $J = 13.1, 6.1$, $\text{CHHCH}=\text{CH}_2$), 4.05 (1H, dd, $J = 12.4, 2.5$, H6), 4.18 (1H, dd, $J = 12.4, 4.8$, H6'), 4.25 (1H, dd, $J = 13.1, 4.8$, $\text{CHHCH}=\text{CH}_2$), 4.49 (1H, d, $J = 7.9$, H1), 4.93 (1H, dd, $J = 9.5, 7.9$, H2), 5.00 (1H, t, $J = 9.6$, H4), 5.10-5.15 (2H, m, contains $\text{HC}=\text{CHH}$ and H3), 5.19 (1H, dd, $J = 17.2, 1.5$, $\text{HC}=\text{CHH}$), 5.76 (1H, m, $\text{HC}=\text{CH}_2$). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 20.5$ (br, contains 2 $\times$ OAc), 20.57, 20.63 (2 $\times$ OAc), 61.9 ($\text{CH}_2\text{CH}=\text{CH}_2$), 68.3 (C4), 69.9 (C6), 71.2, 71.7, 72.8 (C2, C3, C5), 99.4 (C1), 117.5 ($\text{HC}=\text{CH}_2$), 133.3 ($\text{HC}=\text{CH}_2$), 169.2, 169.3, 170.1, 170.5 (4 $\times$ C=O). LRMS m/z (ESI⁺): Found 406.2 $[\text{M}+\text{NH}_4]^+$, 799.2 $[2\text{M}+\text{Na}]^+$.

Allyl- β -(D)-glucopyranoside (5.26)



To a stirred solution of allyl 2,3,4,6-tetra-*O*-acetyl- β -(D)-glucopyranoside (2.00 g, 5.15 mmol) in MeOH (20 ml) was added sodium methoxide (139 mg, 0.26 mmol). After stirring for 20 minutes, the reaction was neutralized with Dowex 50WX8 (H^+ form) and then filtered. The resulting solution was concentrated under reduced pressure to give a thick oil. Purification by column chromatography (30% i PrOH in EtOAc) afforded the titled compound as a thick clear oil which crystallized on standing (1.10 g, 97%). Spectroscopic data was consistent with that previously reported.⁷ m.p. = 96-98 °C [Lit.⁷ = 102-103 °C, EtOH (aq)]. $[\alpha]_D^{20} = -39.0$ ($c = 1$, MeOH) [Lit.⁷ = -41.4, $c = 0.43$, MeOH]. IR (ν_{max} , KBr): 3417, 2873, 1645, 1425, 1368, 1258, 1082, 919, 617. 1H NMR (400 MHz, CD_3OD): $\delta = 3.18$ -3.41 (3H, m, contains H2, H3 and H4), 3.68 (1H, dd, $J = 12.0$, 5.5, H6), 3.89 (1H, dd, $J = 12.0$, 1.7, H6'), 3.94 (1H, dt, $J = 12.5$, 6.2, H5), 4.17 (1H, dd, $J = 12.9$, 6.0, OCHHCH=CH₂), 4.32 (1H, d, $J = 7.9$, H1), 4.40 (1H, dd, $J = 12.9$, 5.1, OCHHCH=CH₂), 5.18 (1H, dd, $J = 10.6$, 1.6, HC=CHH), 5.35 (1H, dd, $J = 17.4$, 1.7, HC=CHH), 5.99 (1H, m, HC=CH₂). ^{13}C NMR (100 MHz, MeOD): $\delta = 61.7$ (C6), 70.1 (OCH₂CH=), 70.6 (C5), 74.0 (C2), 76.9, 77.0 (C3, C4), 102.3 (C1), 116.6 (CH=CH₂), 134.7 (CH=CH₂). LRMS m/z (ESI⁺): 238.3 $[M+NH_4]^+$, 243.3 $[M+Na]^+$.

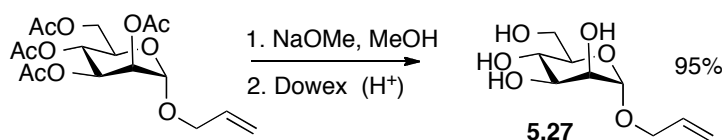
Allyl 2,3,4,6-tetra-*O*-acetyl- α -(D)-mannopyranoside



α -(D)-Mannose pentaacetate (3.50 g, 8.97 mmol) was added to a 100 mL 2-neck round-bottom flask and dissolved in CH_2Cl_2 (50 mL) and placed under nitrogen. The stirred solution was cooled to 0 °C and $BF_3 \cdot OEt_2$ (1.36 mL, 10.80 mmol) was added followed by allyl alcohol (0.92 mL, 13.46 mmol). The reaction was warmed to room temperature over several hours. After 12 hours, the reaction was quenched at 0 °C with $NaHCO_3$ (10 mL, sat., aq.) and diluted with H_2O (100 mL). The organic layer was separated and the aqueous layer was extracted with CH_2Cl_2 (2×100 mL). The combined organics were dried ($MgSO_4$), filtered, and concentrated under reduced pressure. Column

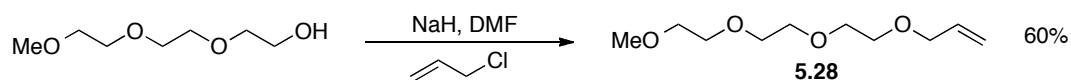
chromatography (50% EtOAc in petrol) afforded titled compound as a clear oil (348 mg, 10%). Spectroscopic data was consistent with that previously reported.⁸ $[\alpha]_D^{20} = +50.9$ ($c = 1.2$, MeOH). IR ($\nu_{\max}$, film): 1751, 1432, 1371, 1227, 1136, 1049. ^1H NMR (400 MHz, CDCl_3): $\delta = 1.90, 1.96, 2.02, 2.07$ (3H, s, $4 \times \text{OAc}$), 3.90-3.97 (2H, m, contains H5 and $\text{CHHCH}=\text{CH}_2$), 4.01 (1H, dd, $J = 12.3, 2.4$, H6), 4.08-4.14 (1H, m, $\text{CHHCH}=\text{CH}_2$), 4.20 (1H, dd, $J = 12.3, 5.3$, H6'), 4.78 (1H, dd, $J = 1.8$, H1), 5.13-5.22 (3H, m, contains H2 and $\text{CH}=\text{CH}_2$), 5.24-5.29 (2H, m, H3 and H4), 5.76-5.86 (1H, m, $\text{CH}=\text{CH}_2$). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 20.61, 20.62, 20.7, 20.8$ ($4 \times \text{OAc}$), 62.4 (C6), 66.0 (C2), 68.5 ($\text{CH}_2\text{CH}=\text{CH}_2$), 68.5, 69.0, 69.5 (C3, C4, C5), 96.5 (C1), 118.3 ($\text{CH}=\text{CH}_2$), 132.9 ($\text{CH}=\text{CH}_2$) 169.6, 169.8, 169.9, 170.5 ($4 \times \text{OAc}$). LRMS m/z (ESI^+): 411.1 $[\text{M}+\text{Na}]^+$.

Allyl- α -(D)-mannopyranoside (5.27)



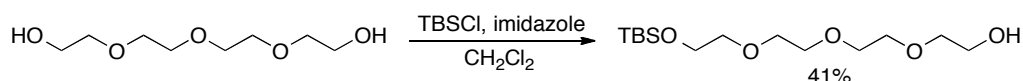
Allyl 2,3,4,6-tetra-*O*-acetyl- α -(D)-mannopyranoside (260 mg, 0.67 mmol) was added to a 25 mL round bottom flask and dissolved in 10 mL MeOH. To the stirred solution at room temperature was added NaOMe (36 mg, 0.67 mmol). The solution was stirred at room temperature for 25 minutes after which time TLC (20% MeOH in EtOAc) indicated consumption of the starting material and formation of a single product ($R_f = 0.3$). The reaction mixture was quenched by the addition of Dowex-50WX8 (H^+ form) until the pH was ~ 7 (pH paper). The solution was filtered and rinsed with MeOH. The solvent was evaporated to give the titled compound as a thick clear oil (141 mg, 95%). Spectroscopic data was consistent with that previously reported.⁸ $[\alpha]_D^{20} = +78.2$ ($c = 0.4$, MeOH). ^1H NMR (400 MHz, CDCl_3): $\delta = 3.54$ (1H, ddd, $J = 9.4, 5.6, 2.3$, H5), 3.64 (1H, t, $J = 9.4$, H4), 3.71-3.75 (2H, m, H3 and H6), 3.83-3.86 (2H, m, H2 and H6'), 3.99-4.04 (1H, m, $\text{CHHCH}=\text{CH}_2$), 4.20-4.25 (1H, m, $\text{CHHCH}=\text{CH}_2$), 4.81 (1H, d, $J = 1.5$, H1), 5.18 (1H, dd, $J = 10.4, 1.8$, $\text{CH}=\text{CHH}$), 5.31 (1H, dd $J = 17.2, 1.8$, $\text{CH}=\text{CHH}$), 5.90-5.99 (1H, m, $\text{CH}=\text{CH}_2$). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 61.9$ (C6), 67.6 (C4), 67.9 ($\text{CH}_2\text{CH}=\text{CH}_2$), 71.2 (C2), 71.7 (C3), 73.7 (C5), 99.7 (C1), 116.4 ($\text{CH}=\text{CH}_2$), 134.5 ($\text{CH}=\text{CH}_2$). LRMS m/z (ESI^+): 243.1 $[\text{M}+\text{Na}]^+$.

Allyl methyl triethylene glycol (5.28)



To a flame dried 100 mL three-neck round bottom flask under nitrogen was added DMF (40 mL, anhydrous). Sodium hydride (306 mg, 60% wt. in mineral oil, 7.66 mmol) was added under a stream of nitrogen. The stirred suspension was cooled to 0 °C and triethylene glycol monomethyl ether (1.00 mL, 6.38 mmol) was added dropwise and then stirred for 5 minutes upon completion of addition. Allyl chloride (1.31 mL, 16.0 mmol) was then added and the reaction was stirred at 0 °C for 40 minutes before quenching carefully with 20 mL H₂O. The mixture was diluted with 300 mL EtOAc and washed successively with H₂O (2 × 150 mL) and brine (200 mL). The organic layer was dried (Na₂SO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc) to give allyl ether **5.28** as a clear oil (778 mg, 60%). IR (ν_{max} , film): 2870, 1625, 1446, 1312, 1218, 1190, 1109. ¹H NMR (400 MHz, CD₃OD): δ = 3.38 (3H, s, OCH₃), 3.54-3.57 (2H, m, CH₂OMe), 3.60-3.67 (10H, m, -OCH₂CH₂O-), 4.03 (2H, dt, J = 5.5, 1.4, OCH₂CH=CH₂), 5.16-5.20 (1H, m, CH=CHH), 5.27-5.33 (1H, m, CH=CHH), 5.89-5.97 (1H, m, CH=CH₂). ¹³C NMR (100 MHz, CD₃OD): δ = 58.2 (CH₃), 69.7, 70.4, 70.7 (br), 72.0, 72.1 (peak overlap; includes -OCH₂CH₂O- and CH₂CH=CH₂), 116.2 (CH=CH₂), 135.3 (CH=CH₂). HRMS m/z (ESI⁺): Found 227.1256 [M+Na]⁺; C₁₀H₂₀O₄Na requires 227.1254.

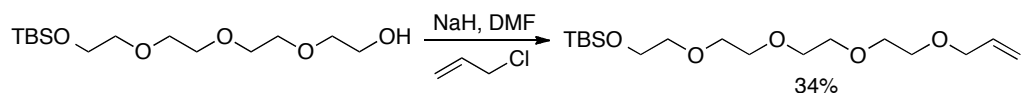
tetraethylene glycol mono(*tert*-butyl dimethylsilyl) ether



CH₂Cl₂ (100 mL) was added to a 250 mL round bottom flask and flushed with nitrogen. Tetraethylene glycol (4.00 mL, 23.17 mmol) and imidazole (4.73 g, 69.51 mmol) were added and stirred to dissolve. The stirred solution was cooled to 0 °C and *tert*-butyl dimethylsilyl chloride (TBSCl, 4.32 g, 27.8 mmol) was added in a single portion. After stirring for 40 min at 0 °C, TLC (EtOAc) revealed the di-silyl- (R_f = 0.6) and mono-silyl ether (R_f = 0.3). The reaction was quenched with 250 mL of H₂O and the organic layer separated. The aqueous layer was then extracted with CH₂Cl₂ (2 × 100 mL). The combined organic layers were washed with brine (200 mL), dried (Na₂SO₄), filtered, and concentrated under reduced pressure. Purification by column chromatography (EtOAc) afforded the titled compound as a clear oil (2.92 g, 41%). IR (ν_{max} , film): 3451, 2926,

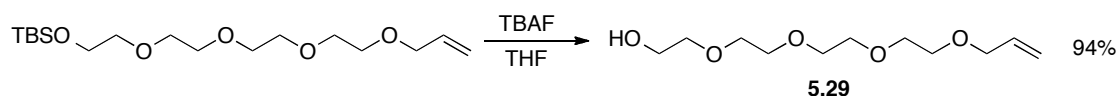
2245, 1753, 1647, 1472, 1361, 1254, 1094. ^1H NMR (400 MHz, CDCl_3): δ = 0.01 (6H, s, TBS), 0.83 (9H, s, TBS), 2.97 (1H, br. s, OH), 3.48-3.72 (16H, m, $-\text{OCH}_2\text{CH}_2\text{O}-$). ^{13}C NMR (100 MHz, CDCl_3): δ = -5.3, 18.3, 25.9 (TBS), 61.6, 62.7, 70.3, 70.6 (br), 70.7, 72.5, 72.6 (peak overlap; $-\text{OCH}_2\text{CH}_2\text{O}-$). HRMS m/z (ESI^+): Found 331.1913 $[\text{M}+\text{Na}]^+$; $\text{C}_{14}\text{H}_{32}\text{O}_5\text{SiNa}$ requires 331.1911.

Allyl (dimethyl-*tert*-butyl silyl) tetraethyleneglycol



To a flame dried 250 mL 2-neck round bottom flask under nitrogen was added DMF (40 mL, anhydrous). Sodium hydride (311 mg, 60% wt. in mineral oil, 7.78 mmol) was added under a stream of nitrogen and the stirred suspension was cooled to 0 °C. Tetraethylene glycol mono(*tert*-butyl dimethylsilyl) ether (2.00 g, 6.48 mmol) was added dropwise to the reaction and then stirred at 0 °C for 3 min. Allyl chloride (1.33 mL, 16.20 mmol) was then added and the reaction stirred at 0 °C for 40 minutes. The reaction was then quenched at 0 °C by slow addition of H_2O (10 mL). The mixture was diluted with EtOAc (250 mL) and H_2O (250 mL). The organic layer was separated, washed successively with H_2O (2 $\times$ 250 mL) and brine (150 mL), dried (Na_2SO_4), and filtered. After concentrating under reduced pressure, the resulting residue was purified by column chromatography (50% EtOAc in petrol) to afford the titled compound as a clear oil (777 mg, 34%). IR (ν_{max} , film): 2858, 2360, 1463, 1340, 1253, 1106. ^1H NMR (400 MHz, CDCl_3): δ = 0.04 (6H, s, TBS), 0.87 (9H, s, TBS), 3.53 (2H, t, J = 5.4), 3.57-3.59 (2H, m), 3.61-3.65 (10H, m), 3.74 (2H, t, J = 5.4) ($-\text{OCH}_2\text{CH}_2\text{O}-$), 4.00 (2H, dt, J = 1.4, 5.6, $\text{CH}_2\text{CH}=\text{CH}_2$), 5.13-5.17 (1H, m, $\text{CH}=\text{CHH}$), 5.21-5.28 (1H, m, $\text{CH}=\text{CHH}$), 5.84-5.94 (1H, m, $\text{CH}=\text{CH}_2$). ^{13}C NMR (100 MHz, CDCl_3): δ = -5.3, 18.4, 25.9 (TBS), 62.7, 69.4, 70.62 (br), 70.65, 70.73, 72.2, 72.6 (peak overlap; includes $-\text{OCH}_2\text{CH}_2\text{O}-$ and $\text{CH}_2\text{CH}=\text{CH}_2$), 117.1 ($\text{CH}=\text{CH}_2$), 134.7 ($\text{CH}=\text{CH}_2$). HRMS m/z (ESI^+): Found 371.2224 $[\text{M}+\text{Na}]^+$; $\text{C}_{17}\text{H}_{36}\text{O}_5\text{SiNa}$ requires 371.2224.

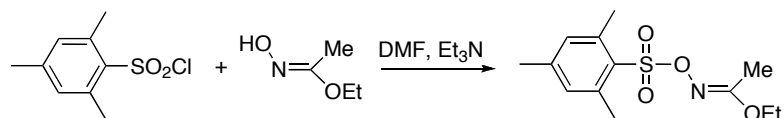
Tetraethylene glycol monoallyl ether (5.29)



Allyl (dimethyl-*tert*-butyl silyl) tetraethyleneglycol (275 mg, 0.79 mmol) was added to a 50 mL round bottom flask and flushed with argon before dissolving in THF (10 mL, anhydrous). The stirred solution was cooled to 0 °C and TBAF (1.18 mL of a 1.0 M solution in THF) was added dropwise. The reaction was stirred for 30 min and another portion of TBAF (1.0 mL of a 1.0 M solution in THF) was added. After 10 minutes, TLC (10% MeOH in EtOAc) revealed complete consumption of the starting material (R_f = 0.7) and formation of the product (R_f = 0.3). The reaction was quenched by the addition of Dowex 50WX8 (H^+). The resin was filtered off and rinsed with MeOH. The filtrate was concentrated and then purified by column chromatography (10% MeOH in EtOAc) to afford the titled compound as a clear oil (174 mg, 94%). IR (ν_{max} , film): 3454, 3080, 2868, 1964, 1646, 1455, 1350, 1249, 1104, 931. 1H NMR (400 MHz, $CDCl_3$): δ = 3.56-3.68 (16H, m, $-OCH_2CH_2O-$), 4.03 (2H, dt, J = 5.6, 1.5, $CH_2CH=CH_2$), 5.15-5.20 (1H, m, $CH=CHH$), 5.27-5.33 (1H, m, $CH=CHH$), 5.88-5.98 (1H, m, $CH=CH_2$). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 61.3, 69.6, 70.4, 70.59 (br), 70.63, 72.1, 72.7 (peak overlap; includes $-OCH_2CH_2O-$ and $CH_2CH=CH_2$), 116.3 ($CH=CH_2$), 135.1 ($CH=CH_2$). HRMS m/z (ESI^+): Found 257.1363 $[M+Na]^+$; $C_{11}H_{22}O_5Na$ requires 257.1359.

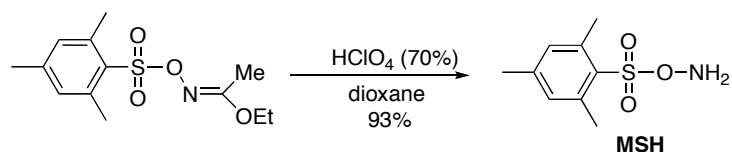
Preparation of MSH

Ethyl-*O*-(mesitylenesulfonyl)acetohydroxamate (MSH precursor)



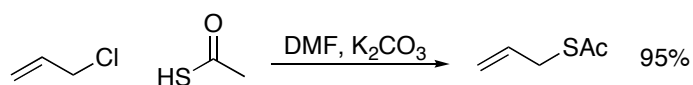
Ethyl *N*-hydroxyacetimidate (2.36 g, 22.9 mmol) was dissolved in DMF (6 mL). Triethylamine (3 mL) was added and the stirred solution was cooled to 0 °C. 2-Mesitylenesulfonyl chloride (5.00 g, 22.9 mmol) was added in two portions and the resulting white slurry was stirred vigorously for 15 minutes. The mixture was then diluted with CH_2Cl_2 (100 mL) and washed repeatedly with H_2O . The organic layer was dried ($MgSO_4$), filtered, and the solvent removed under reduced pressure to give the titled compound as a white solid that was used without further purification. Spectroscopic data was identical to that previously reported. m.p. = 51-53 °C [Lit.⁹ = 57-58 °C]. 1H NMR (400 MHz, $CDCl_3$): δ = 1.19 (3H, t, J = 7.1 Hz, CH_2CH_3), 2.04 (3H, s, CH_3), 2.31 (3H, s, CH_3Ar), 2.64 (6H, s, $2 \times CH_3Ar$), 3.90 (2H, q, J = 7.1 Hz, CH_2CH_3), 6.97 (2H, s, ArH). ^{13}C NMR (100 MHz, $CDCl_3$): δ = 14.0 (OCH_2CH_3), 14.9 (CH_3), 21.1 (CH_3Ar), 22.8 ($2 \times CH_3Ar$), 63.6 (OCH_2CH_3), 130.3, 131.5, 140.7, 143.3 (Ar), 169.2 ($C=N$).

***O*-Mesitylsulfonylhydroxylamine (MSH)**



A solution of ethyl-*O*-(mesitylsulfonyl)acetohydroxamate (4.42 g, 15.5 mmol) in dioxane (4 mL) was cooled to 0 °C. Perchloric acid (70%, 1.80 mL) was added dropwise by pipette over 2 minutes. After stirring for 5 minutes the mixture solidified. The contents of the reaction were transferred to 200 mL of ice water and the flask rinsed with H₂O (50 mL) and Et₂O (50 mL). The contents were transferred to a separatory funnel and extracted with Et₂O (40 mL). The organic layer was neutralized and partially dried with anhydrous potassium carbonate and then filtered. The filtrate was concentrated to less than 100 mL total volume and then poured into 150 mL of ice cold petrol and left to crystallize for 30 min. The white crystals (small needles) were isolated by filtration, transferred to a plastic Falcon tube, and dried under vacuum. The dried product (3.11 g, 93% yield) was stored at -20 °C and sealed with no more than wax film. m.p. = 90-91 °C [Lit.⁹ = 95-96 °C]. ¹H NMR (400 MHz, CDCl₃): δ = 2.32 (3H, s, CH₃Ar), 2.63 (6H, s, 2 × CH₃Ar), 6.58 (2H, br s, NH₂), 6.98 (2H, s, ArH). ¹³C NMR (100 MHz, CDCl₃): δ = 21.1 (CH₃Ar), 22.7 (2 × CH₃Ar), 128.9, 131.7, 141.0, 143.9 (Ar). Analysis for C₉H₁₅NO₄ requires C, 50.21%; H, 6.09%; N, 6.51%; Found: C, 50.18%; H, 5.90%; N, 6.27%.

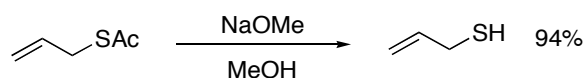
Allyl thioacetate



To a 500 mL round bottom flask equipped with a Teflon coated stir bar was added DMF (70 mL), K₂CO₃ (24.87 g, 180 mmol), and allyl chloride (20.0 mL, 245 mmol). The flask was equipped with a rubber septum and placed under argon. The stirred mixture was cooled to 0 °C and thioacetic acid (11.64 mL, 164 mmol) was added slowly. An efficient exit through a bubbler was ensured to compensate for CO₂ evolution. After gas evolution had subsided, the ice bath was removed and the reaction stirred vigorously at room temperature for 1.5 hours. TLC (5% EtOAc in petrol) indicated consumption of thioacetic acid (*R*_f = 0.0 to 0.1) and formation of allyl thioacetate (*R*_f = 0.7). The reaction mixture was transferred to a separatory funnel and diluted with Et₂O (600 mL) and H₂O (250

mL). The layers were separated and the organic layer was washed successively with NaHCO₃ (sat. aq., 250 mL), H₂O (250 mL), and brine (250 mL). The organic layer was dried (MgSO₄) and filtered. The solution was concentrated to ~100 mL by rotary evaporation at a pressure of 300 mm Hg with care taken not to exceed 25 °C. The resulting yellow solution was purified by distillation (b.p. = 54 °C, 38 mm Hg) to give the titled compound as a clear liquid (17.98 g, 95%). IR (ν_{max} , film): 1695, 1423, 1355, 1137, 959, 921. ¹H NMR (400 MHz, CDCl₃): δ = 2.33 (3H, s, SAc), 3.53 (2H, dd, J = 7.1, 1.3, CH₂SAc), 5.09 (1H, dt, J = 9.9, 1.3, HC=CHH), 5.23 (1H, dt, J = 16.9, 1.3, HC=CHH), 5.79 (1H, m, HC=CH₂). ¹³C NMR (100 MHz, CDCl₃): δ = 30.5 (SAc), 32.0 (CH₂SAc), 117.9 (HC=CH₂), 133.0 (HC=CH₂), 195.1 (C=O). Attempts at microanalysis were complicated by volatility.

Allylthiol (~1M in MeOH/MeOAc)



Allyl thiol has an unpleasant stench. Keep in fumehood at all times.

Allyl thioacetate, (16.62 g, 143.0 mmol) was added to a 250 round bottom flask equipped with a Teflon coated stir bar. Anhydrous methanol (80 mL) was added and the stirred solution was cooled to 0 °C. Sodium methoxide (8.11 g, 150 mmol) was added in four portions over 10 minutes. After 10 minutes of stirring, TLC (30% EtOAc in petrol) revealed that all thioester had been consumed, with all product at the baseline. The reaction was quenched with DOWEX 50WX8 (H⁺ form) until the pH was ~7, as indicated by pH paper. The reaction was filtered and then distilled directly (b.p. = 67 °C, 760 mm Hg) to give the titled compound as a solution in MeOH and MeOAc (MeOH and MeOAc co-distilled with allyl thiol). Approximating the density of the solution as the density of MeOH (~ 0.8 g/mL), the concentration of allylthiol is ~1.08 M, as judged by ¹H NMR. Yield = 125 mL = 135 mmol = 94%. ¹H NMR (400 MHz, CDCl₃): δ = 1.41 (1H, t, J = 7.8, SH), 3.13 (2H, t, J = 7.8, CH₂SH), 4.97 (1H, ad, J = 10.1, CH=CHH), 5.12 (1H, dd, J = 16.8, 1.4, CH=CHH), 5.90 (1H, m, CH=CH₂). This solution was used directly in subsequent reactions.

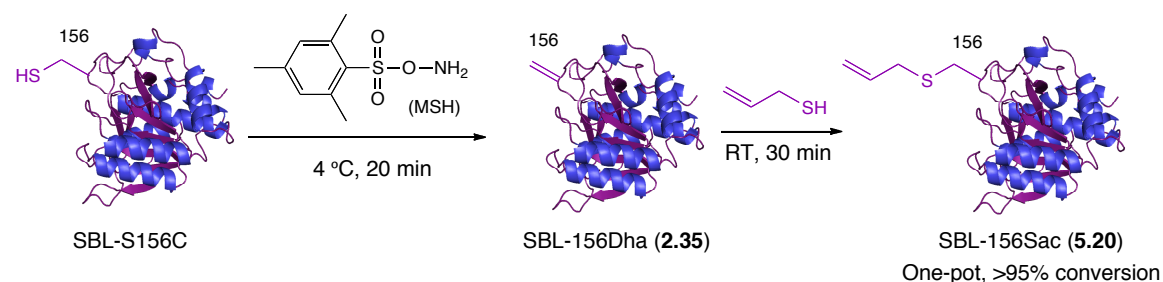
Protein Modification by Olefin Metathesis

Sequence of SBL-S156C (BPN' numbering, PDB code for wild type SBL = 1GCI)

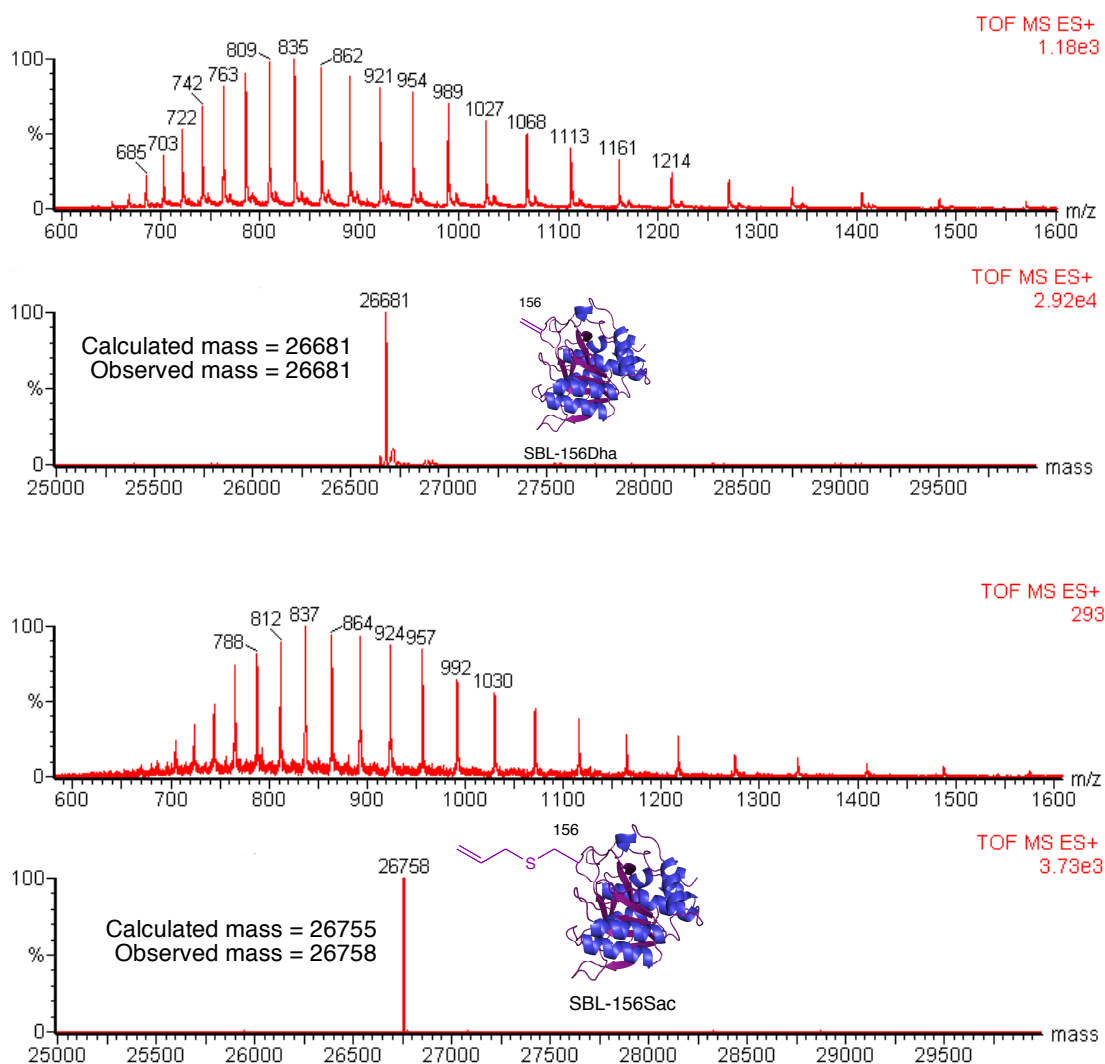
AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPS
TQDGNGHGHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGL
EWAGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGN**C**GAGSISYPA
RYANAMAVGATDQNNNRASFQYAGGLDIVAPGVNVQSTYPGSTYASLNGTSM
ATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Calculated average isotopic mass = 26715

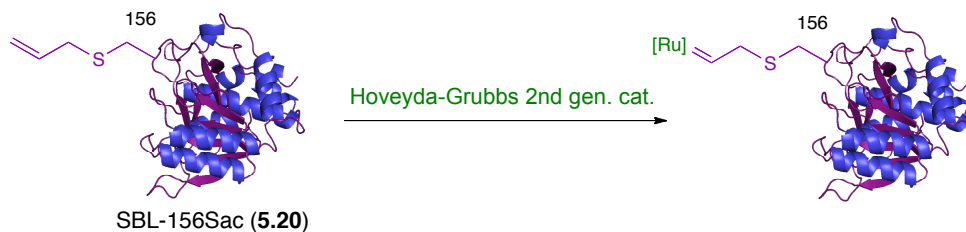
SBL-156-*S*-Allyl-Cysteine (SBL-156Sac, **5.20**) made via dehydroalanine



SBL-S156C was prepared as a 1.4 mg/mL solution in 50 mM sodium phosphate buffer (pH 8.0) and 1.00 mL (0.052 μmol) was added to a 1.50 mL plastic tube. MSH (1.2 mg, 5.6 μmol) was added as a solution in DMF (50 μL) and the reaction vortexed immediately upon addition. The homogenized sample was shaken at 4 $^{\circ}\text{C}$ for 20 minutes before a 40 μL aliquot was analyzed by LC-MS, confirming full conversion of Cys156 to Dha156 (calculated mass = 26681; observed mass = 26681). Allylthiol (190 μL of the 1.08 M solution in MeOH prepared above) was added directly to the protein mixture and vortexed to homogenize. The reaction was rotated at room temperature for 30 minutes and then analyzed directly by LC-MS, confirming full conversion to *S*-allyl cysteine (Sac) (calculated mass = 26755; observed mass = 26758). Small molecules were removed with a PD10 column (GE Healthcare), eluting with 50 mM sodium phosphate buffer (pH 8.0). The sample was further purified by dialysis (2 $\times$ against 4L of 50 mM sodium phosphate, pH 8.0). The sample was then split into 350 μL aliquots, flash frozen, and stored at -80 $^{\circ}\text{C}$. The final concentration of **5.20** was 0.4 mg/mL, protein recovery was >95%. ESI-MS for both SBL-156Dha and SBL-156Sac are shown below:



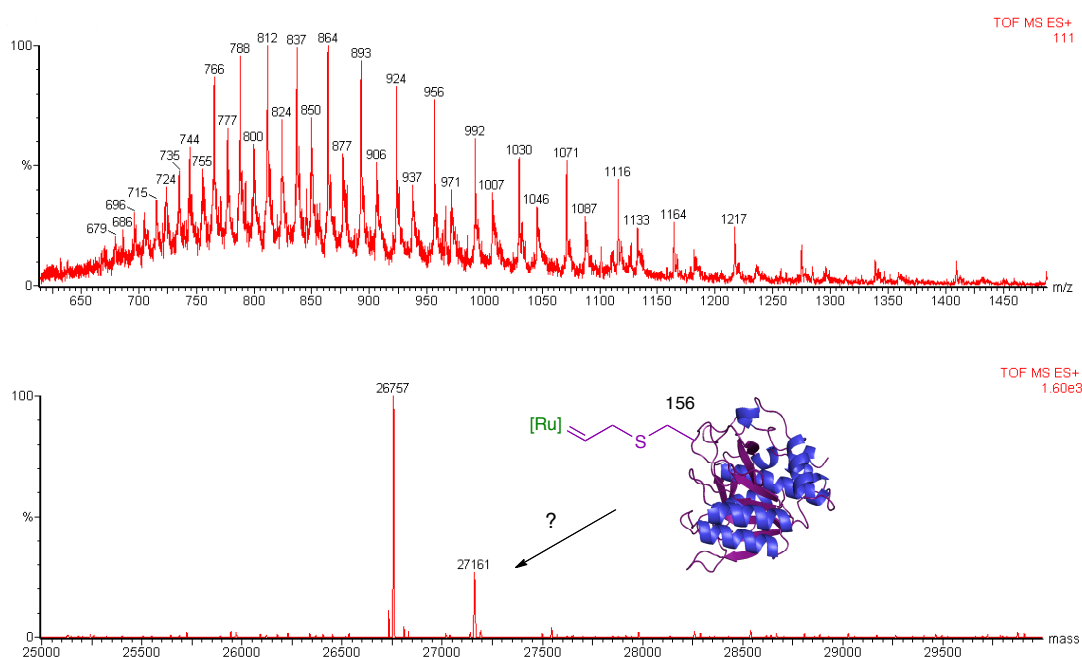
Putative metalloprotein intermediate



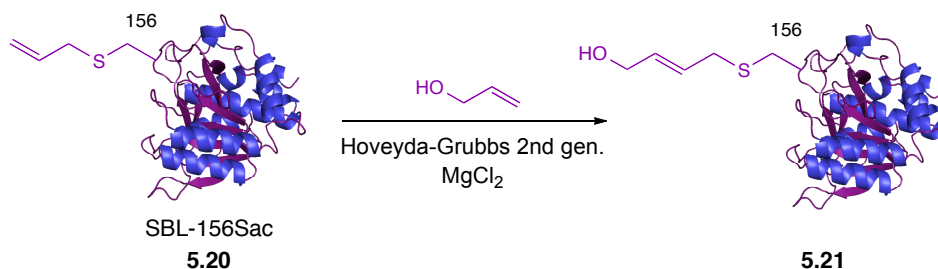
This experiment was carried out with Angel Lin.

All manipulations were carried out at room temperature. SBL-156Sac **5.20** was thawed and stored on ice until needed (previously prepared as described above, 0.4 mg/mL in 50 mM sodium phosphate, pH 8.0). A saturated solution of Grubbs-Hoveyda II was prepared by repeatedly vortexing and sonicating 0.8 mg of the catalyst in 115 μ L t -BuOH. A 250 μ L aliquot of SBL-C156Sac **5.20** (0.004 μ mol) was transferred to a 1.5 mL plastic tube.

An aliquot of the Grubbs-Hoveyda catalyst solution was then added (71.5 μL , 0.75 μmol) and the sample vortexed. Finally, allyl alcohol (2.5 μL , 37 μmol) was added. The reaction tube was rotated on a lab rotisserie at room temperature. The progress of the reaction was monitored directly by LC-MS. A mass of 27161 [M+406] was observed after 1 hour. The reaction was monitored up to 5 hours with little change observed. While only speculative, the [M+406] adduct could be a metalloprotein. This result provided the motivation to use MgCl_2 as an additive in future cross-metathesis attempts. ESI-MS are shown below.



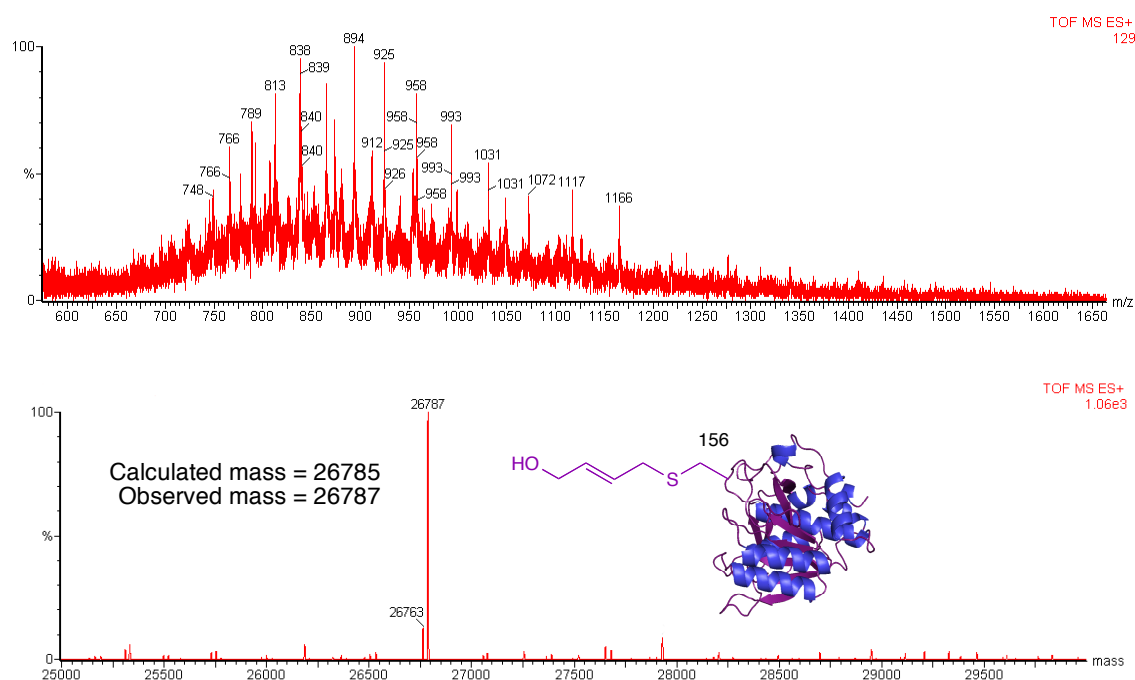
Cross metathesis on SBL-156Sac with allyl alcohol using MgCl_2 additive



This experiment was carried out with Angel Lin.

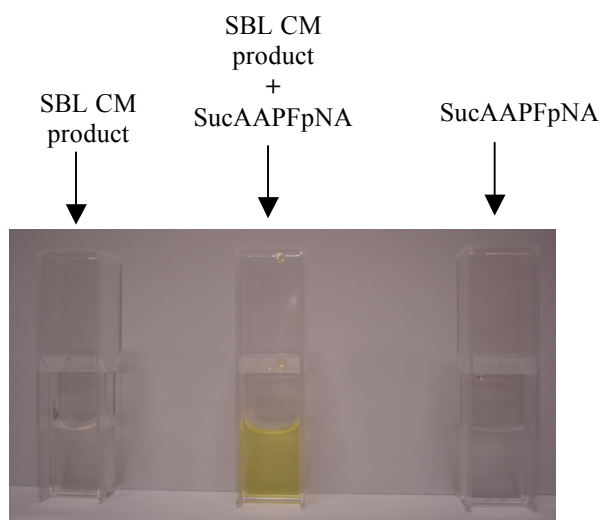
All manipulations were carried out at room temperature. SBL-156Sac **5.20** was thawed and stored on ice until needed (previously prepared as described above, 0.4 mg/mL in 50 mM sodium phosphate, pH 8.0). A saturated solution of Grubbs-Hoveyda II was prepared by repeatedly vortexing and sonicating 0.8 mg of the catalyst in 115 μL $t\text{-BuOH}$. A 250

μL aliquot of SBL-C156Sac **5.20** (0.004 μmol) was transferred to a 1.5 mL plastic tube. $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (7.1 mg, 35 μmol) and $t\text{BuOH}$ (36 μL) were both added to the protein solution and the sample vortexed to homogenize. An aliquot of the Grubbs-Hoveyda catalyst solution was then added (71.5 μL , ~ 0.75 μmol) and the sample vortexed. Finally, allyl alcohol (2.5 μL , 37 μmol) was added. The reaction tube was rotated on a lab rotisserie at room temperature. The progress of the reaction was monitored directly by LC-MS. After 5 hours of reaction time, $>90\%$ conversion to cross metathesis product **5.21** was observed (calculated mass = 26785; observed mass = 26787). ESI-MS are shown below.



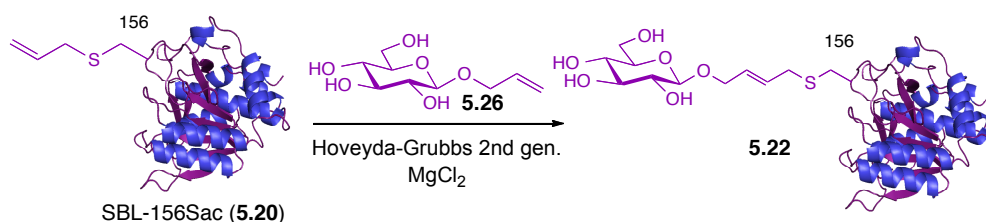
Peptidase activity assay of cross-metathesis modified SBL **5.21**

An aliquot of the reaction mixture of the cross-metathesis product **5.21** was prepared at a concentration of 0.1 mg/mL in pH 8.0 sodium phosphate buffer (50 mM) and 250 μL was added to 500 μL plastic tube. A 25 μL aliquot of the peptide SucAAPFpNA (0.20 M in DMSO, Bachem) was added and the reaction mixture turned bright yellow immediately upon addition. The resulting yellow solution indicates amide hydrolysis and liberation of *para*-nitroaniline (pNA), confirming peptidase activity of the modified SBL. The modified protein and sucAAPFpNA alone are clear solutions at the same concentration (see figure below).



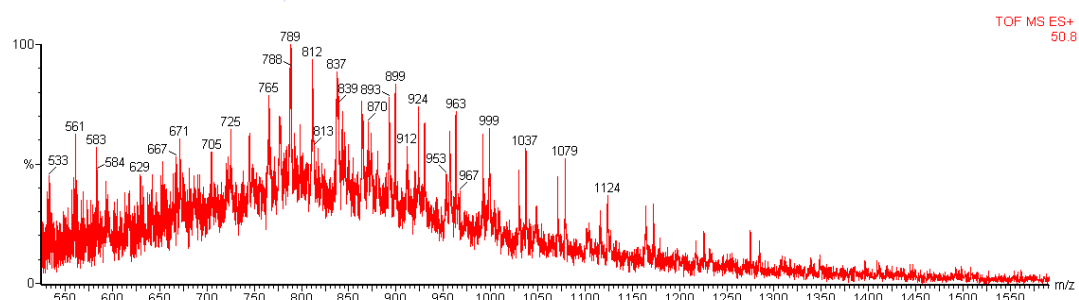
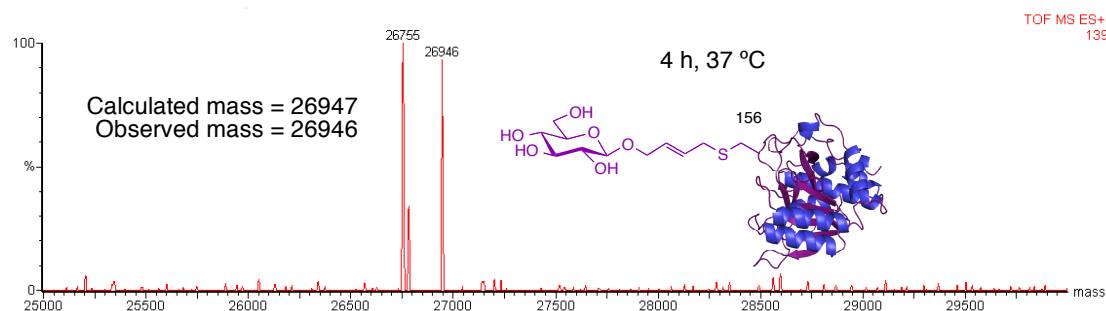
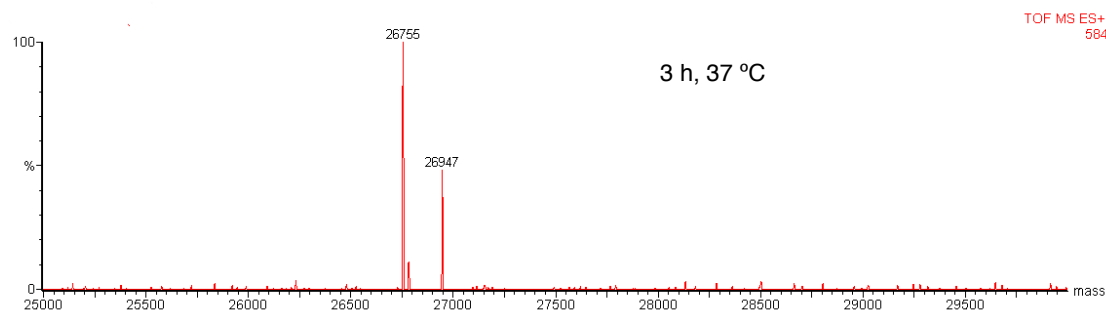
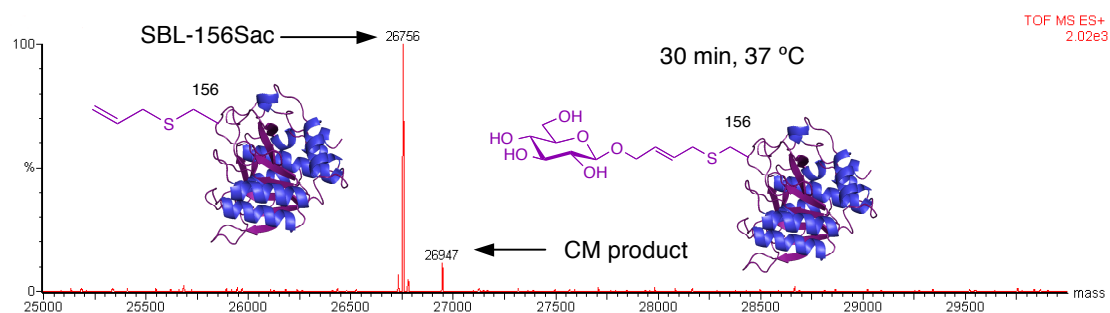
Peptidase Activity of CM Modified SBL 5.21

Cross metathesis on SBL-156Sac 5.20 with allyl glucoside 5.26

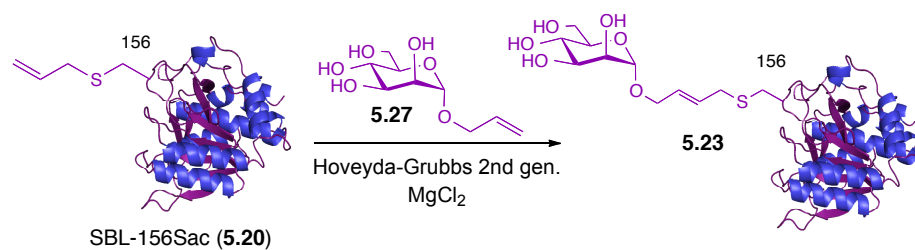


This experiment was carried out with Angel Lin.

All manipulations were carried out at room temperature. SBL-156Sac **5.20** was thawed and stored on ice until needed (previously prepared as described above, 0.4 mg/mL in 50 mM sodium phosphate, pH 8.0). A saturated solution of Grubbs-Hoveyda II was prepared by repeatedly vortexing and sonicating 0.9 mg of the catalyst in 141 μ L t BuOH. A 250 μ L aliquot of SBL-C156Sac **5.20** (0.004 μ mol) was transferred to a 1.5 mL plastic tube. $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (6.5 mg, 30 μ mol) and t BuOH (36 μ L) were both added to the protein solution and the sample vortexed to homogenize. An aliquot of the Grubbs-Hoveyda catalyst solution was then added (71.5 μ L, $\sim$ 0.73 μ mol) and the sample vortexed. Finally, allyl glucoside **5.26** (2.5 mg, 11 μ mol) was added and the reaction vortexed and then rotated on a lab rotisserie at room temperature. The progress of the reaction was monitored by LC-MS. After 1 hour at room temperature, no product formation was observed so the reaction tube was incubated at 37 $^\circ\text{C}$ with shaking. After 4 hours at 37 $^\circ\text{C}$, 50% conversion of to cross metathesis product was observed. ESI-MS are shown below.

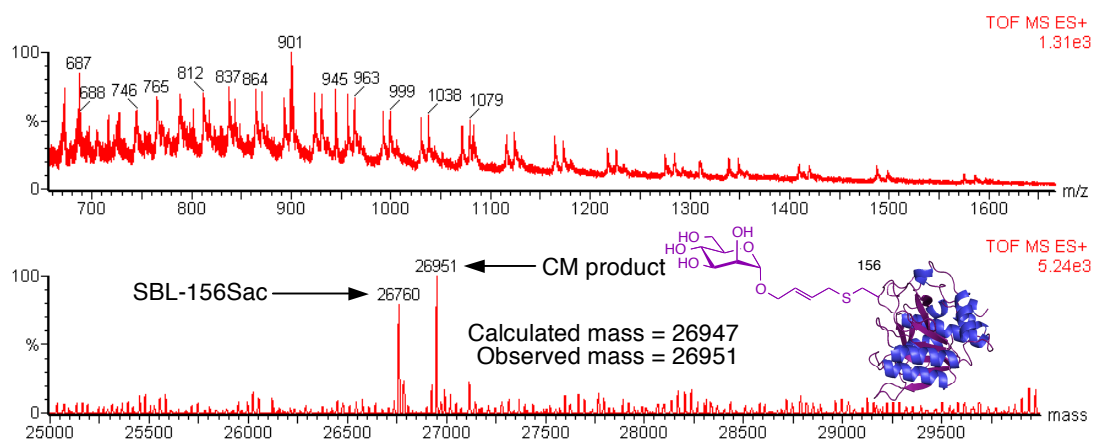


Cross metathesis on SBL-S156Sac 5.20 with allyl mannoside 5.27

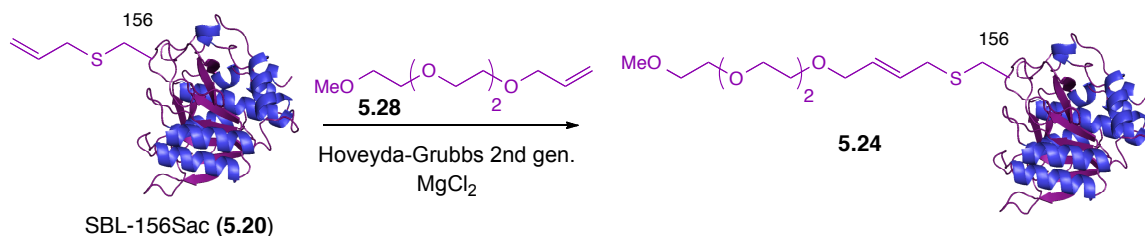


This experiment was carried out with Angel Lin.

All manipulations were carried out at room temperature. SBL-156Sac **5.20** was thawed and stored on ice until needed (previously prepared as described above, but at a concentration of 0.57 mg/mL in 50 mM sodium phosphate, pH 8.0). A saturated solution of Grubbs-Hoveyda II was prepared by repeatedly vortexing and sonicating 0.83 mg of the catalyst in 100 μ L t BuOH. A 250 μ L aliquot of SBL-156Sac **5.20** (0.005 μ mol) was transferred to a 1.5 mL plastic tube. $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (11 mg, 53 μ mol) and t BuOH (36 μ L) were both added to the protein solution and the sample vortexed to homogenize. An aliquot of the Grubbs-Hoveyda catalyst solution was then added (71.5 μ L, $\sim$ 1.1 μ mol) and the sample vortexed. Finally, allyl mannoside **5.27** (12 mg, 53 μ mol) was added and the reaction vortexed and then on a lab rotisserie at room temperature. The progress of the reaction was monitored by LC-MS. After 1 hour at room temperature, no product formation was observed so the reaction tube was incubated at 37 $^\circ\text{C}$. After 4 hours at 37 $^\circ\text{C}$, 60% conversion to the cross metathesis product was observed (calculated mass = 26947; observed mass = 26951). ESI-MS are shown below.



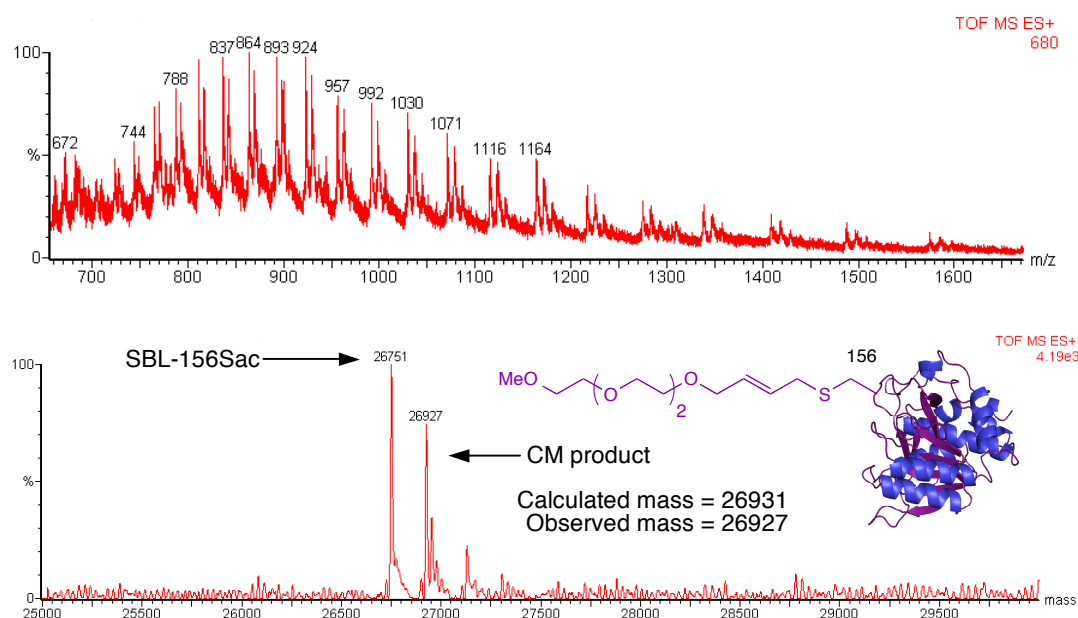
Cross metathesis on SBL-156Sac **5.20** with allyl methyl triethylene glycol **5.28**



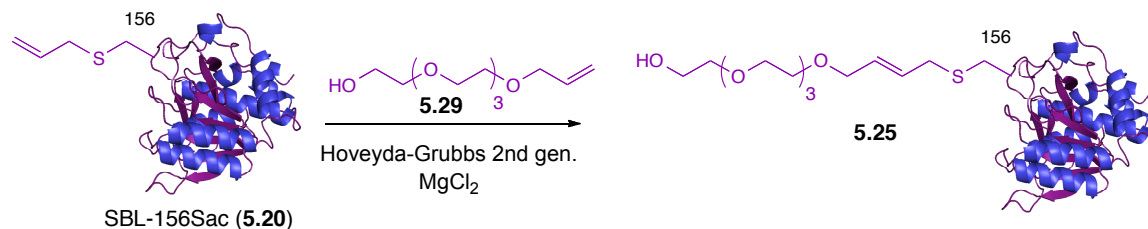
This experiment was carried out with Angel Lin.

All manipulations were carried out at room temperature. SBL-156Sac **5.20** was thawed and stored on ice until needed (0.57 mg/mL in 50 mM sodium phosphate, pH 8.0). A saturated solution of Grubbs-Hoveyda II was prepared by repeatedly vortexing and

sonicating 2.0 mg of the catalyst in 330 μL $t\text{BuOH}$. A 250 μL aliquot of SBL-C156Sac **5.20** (0.005 μmol) was transferred to a 1.5 mL plastic tube. $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (11.7 mg, 57 μmol) and $t\text{BuOH}$ (36 μL) were both added to the protein solution and the sample vortexed to homogenize. An aliquot of the Grubbs-Hoveyda catalyst solution was then added (71.5 μL , ~ 1.1 μmol) and the sample vortexed. Finally, allyl methyl triethylene glycol **5.28** (5.4 mg, 27 μmol) was added and the reaction, vortexed, and rotated on a lab rotisserie at room temperature. The progress of the reaction was monitored by LC-MS. After 2 hour at room temperature, 20% product formation was observed. The reaction tube was then incubated at 37 $^\circ\text{C}$. After 1 hour at 37 $^\circ\text{C}$, 45% conversion of to cross metathesis product was observed (calculated mass = 26931; observed mass = 26927). ESI-MS are shown below.



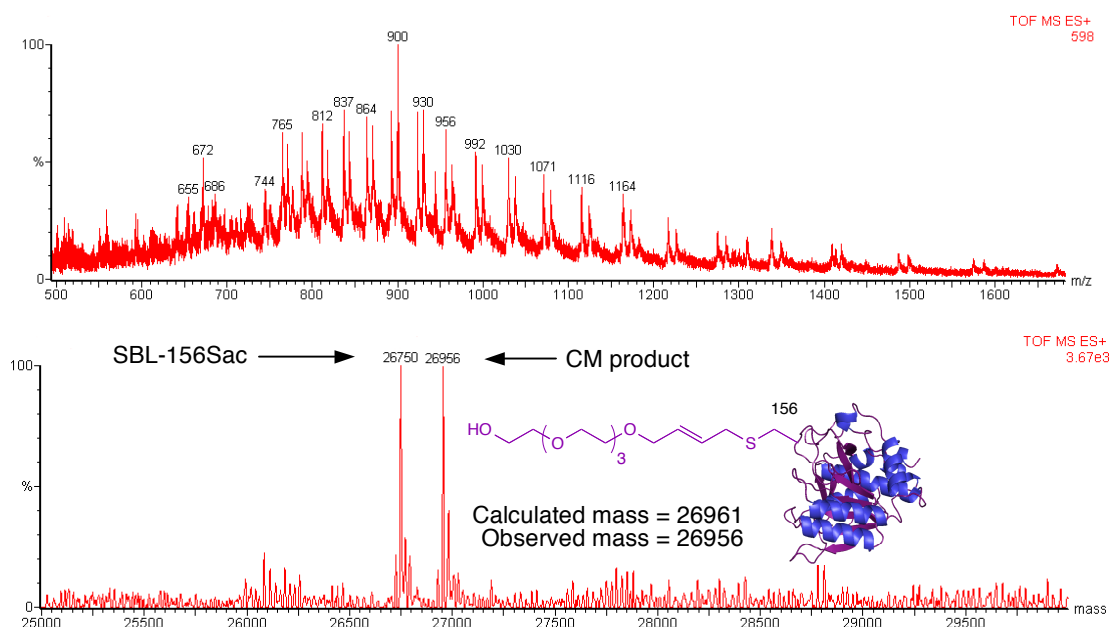
Cross metathesis on SBL-S156Sac **5.20** with allyl tetraethylene glycol **5.29**



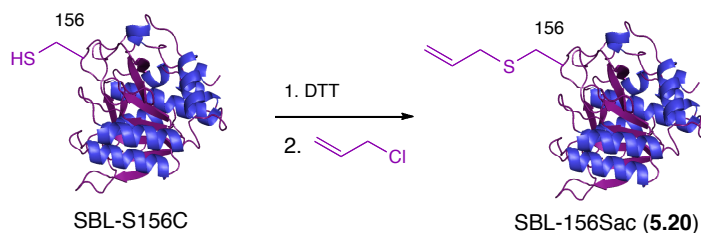
This experiment was carried out with Angel Lin.

All manipulations were carried out at room temperature. SBL-C156Sac **5.20** was thawed and stored on ice until needed (previously prepared as described above, 0.57 mg/mL in 50

mM sodium phosphate, pH 8.0). A saturated solution of Grubbs-Hoveyda II was prepared by repeatedly vortexing and sonicating 2.0 mg of the catalyst in 330 μL $t\text{BuOH}$. A 250 μL aliquot of SBL-C156Sac **5.20** (0.005 μmol) was transferred to a 1.5 mL plastic tube. $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (11.0 mg, 54 μmol) and $t\text{BuOH}$ (36 μL) were both added to the protein solution and the sample vortexed to homogenize. An aliquot of the Grubbs-Hoveyda catalyst solution was then added (71.5 μL , ~ 1.1 μmol) and the sample vortexed. A stock solution of tetraethylene glycol monoallyl ether **5.29** was prepared by dissolving 18.6 mg in 50 μL of 30% $t\text{BuOH}$ in 50 mM sodium phosphate (pH 8.0). A 16.8 μL aliquot of this solution (6.2 mg, 27 μmol) was added to the reaction and the tube was vortexed and then rotated on a lab rotisserie at room temperature for 2 hours and then 1 hour at 37 $^\circ\text{C}$. LC-MS analysis revealed 50% conversion of to cross metathesis product (calculated mass = 26961; observed mass = 26956). ESI-MS are shown below.

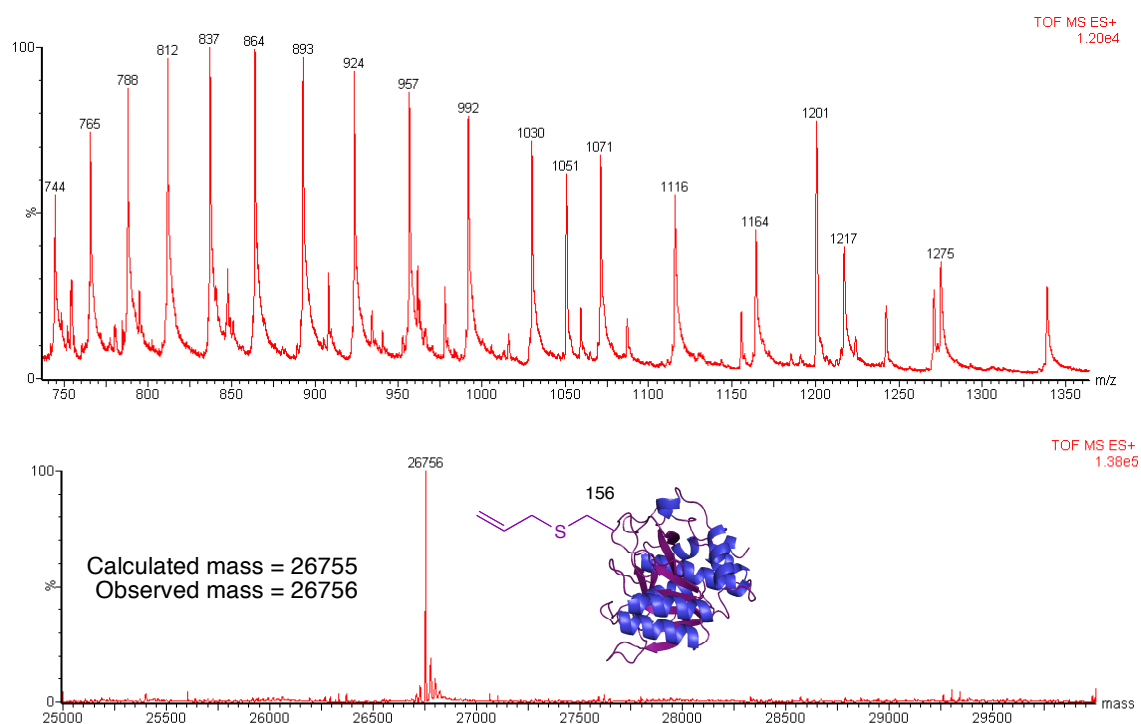


Allylation of SBL-S156C with allyl chloride

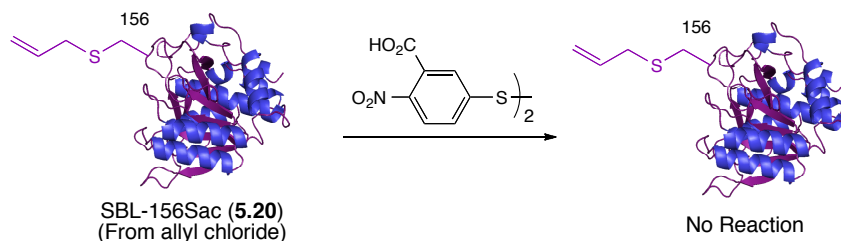


SBL-S156C (2.5 mL, 1.0 mg/mL, pH 8.0 sodium phosphate, 94 nmol) was added to a 15 mL Falcon tube and stored on ice. Dithiothreitol (DTT) (3.6 mg, 23 μmol) was added as a

solid to reduce any contaminant disulfide. The solution was vortexed and then shaken for 10 minutes at room temperature. Allyl chloride (19 μL , 230 μmol) was then added as a solution in DMF (200 μL). The mixture was vortexed and then shaken at 37 $^{\circ}\text{C}$ for 30 min. LC-MS analysis revealed full conversion to the allylated product **5.20** (calculated mass = 26755; observed mass = 26756). The reaction mixture was passed through a PD10 column previously equilibrated with pH 8.0 sodium phosphate (50 mM). The protein concentration was 0.7 mg/mL with >95% protein recovery. The product was split into 200 μL aliquots and flash frozen. ESI-MS are shown below.

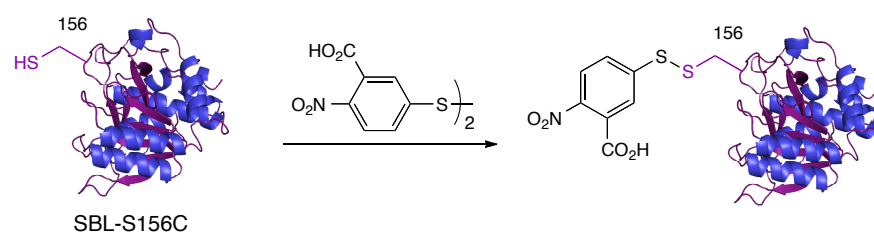
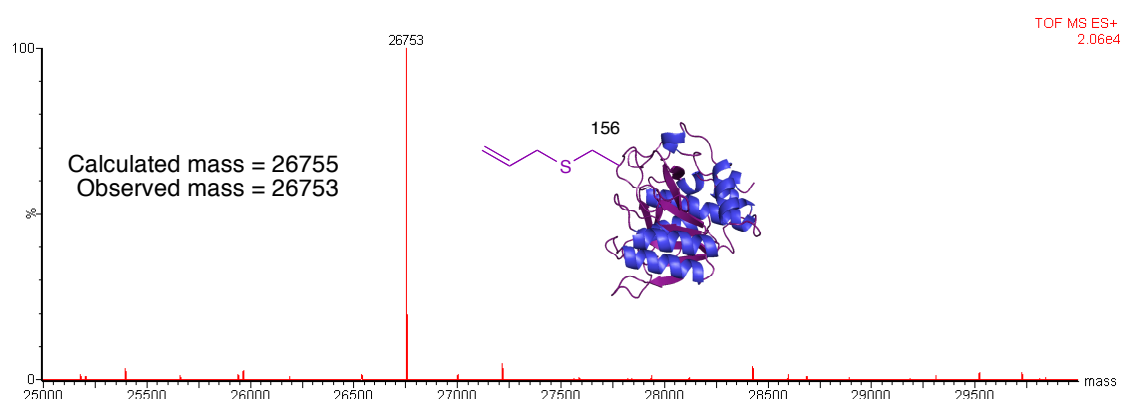
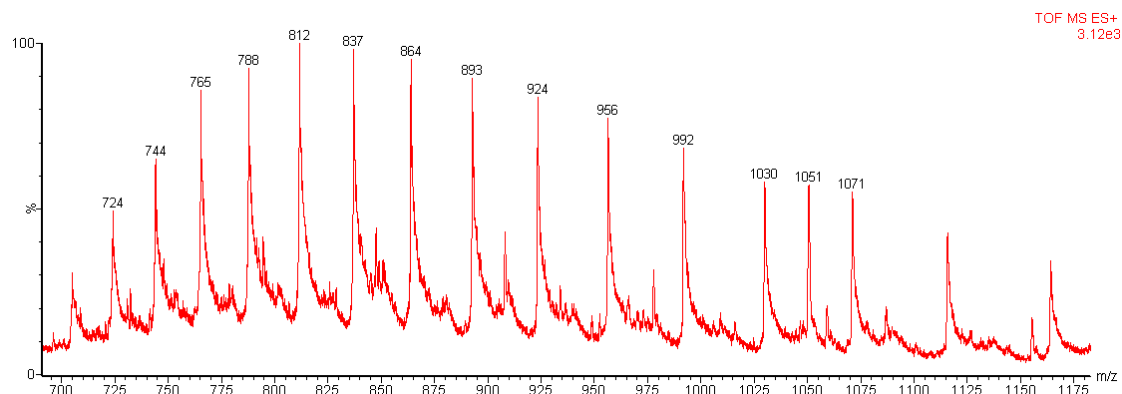


Ellman's test to confirm reaction at cysteine

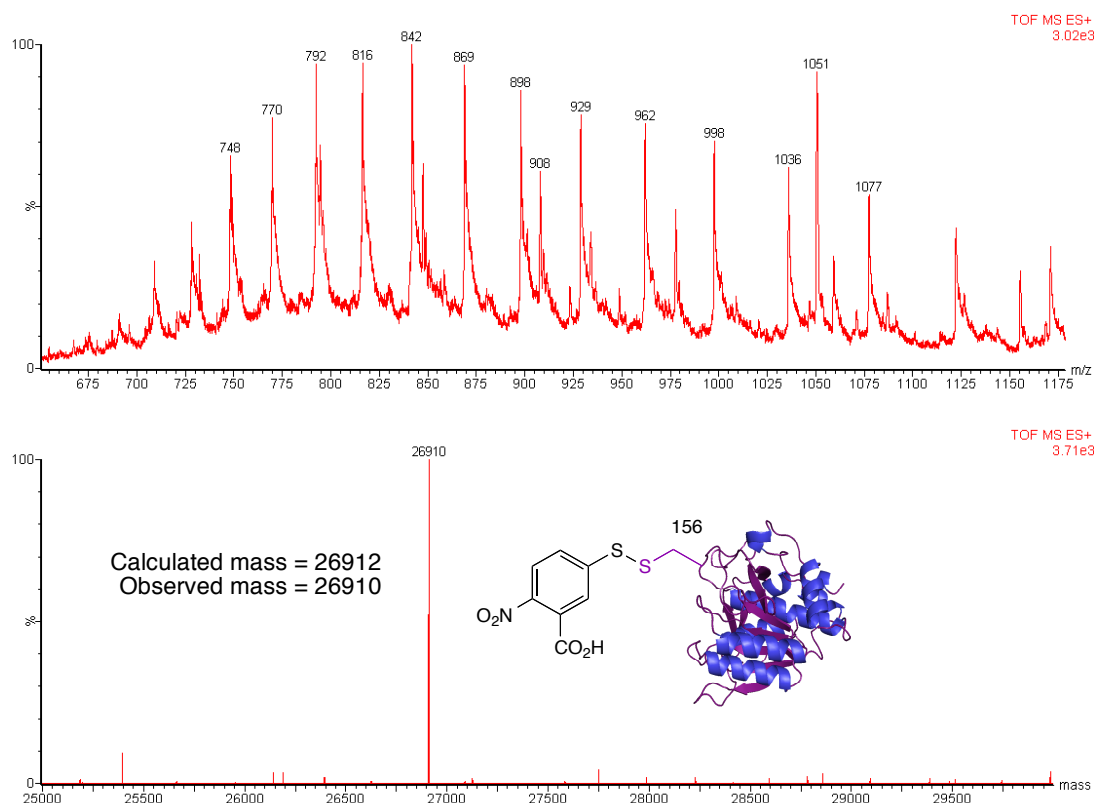


A buffered solution of Ellman's reagent was prepared by dissolving 3.0 mg of Ellman's reagent in 300 μL of 50 mM sodium phosphate buffer (pH 8). A 20 μL of the Ellman's solution (0.5 μmol) was added to SBL-156Sac **5.20** (100 μL , 0.7 mg/mL, 3 nmol, prepared using allyl chloride). The reaction was vortexed and shaken at room temperature

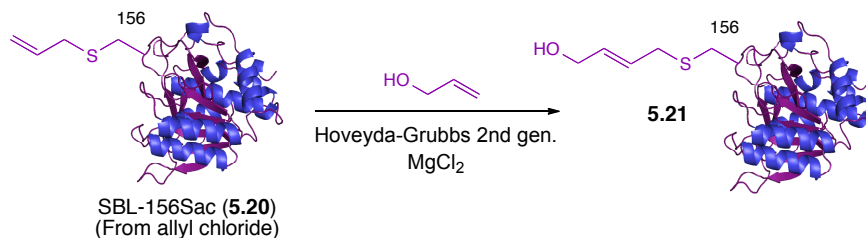
for 10 minutes. The reaction was then analyzed by LC-MS, revealing only the starting material (calculated mass = 26755; observed mass = 26753). This assay indicates that all free thiol is consumed, thus the allylation occurred at cysteine. Under identical conditions, the unmodified SBL-S156C forms the Ellman disulfide (see next).



An SBL-S156C solution (100 μ L, 0.7 mg/mL) was prepared by diluting 71 μ L of SBL-S156C (1 mg/mL) to 100 μ L with 50 mM sodium phosphate buffer (pH 8.0). A 20 μ L aliquot of the Ellman's solution prepared above was added to protein solution. The solution turned bright yellow immediately upon addition of the Ellman's reagent. The reaction was shaken for 10 minutes at room temperature and was analyzed directly with LC-MS. Full conversion to the SBL-S156C Ellman's adduct was observed (calculated mass = 26912; observed mass = 26910).



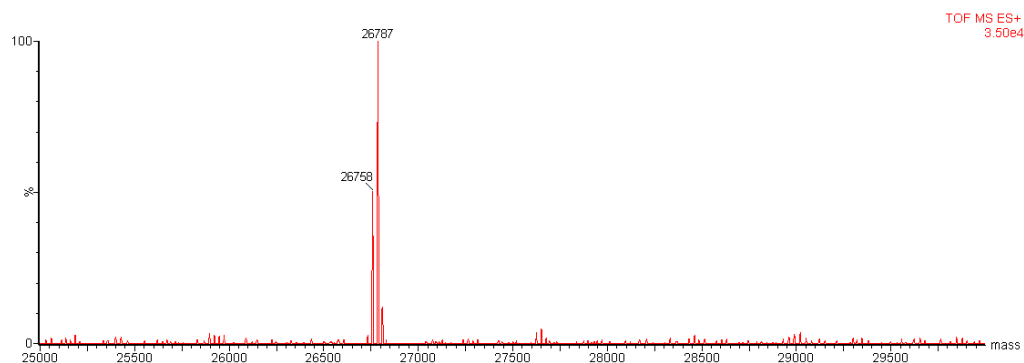
Olefin metathesis on SBL-156Sac after allylation with allyl chloride



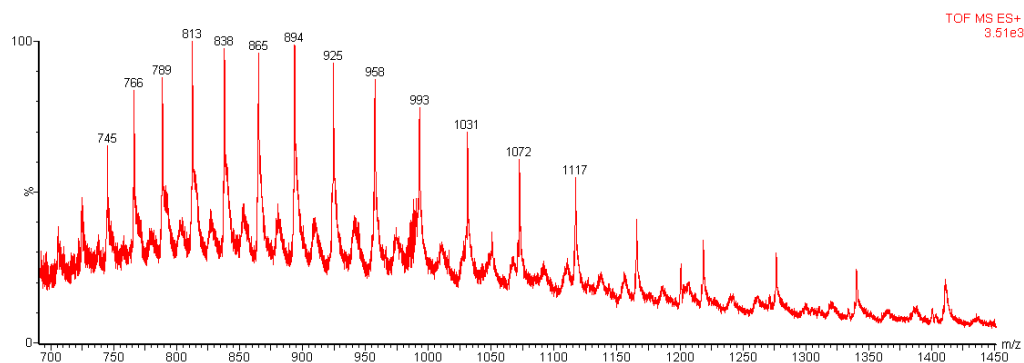
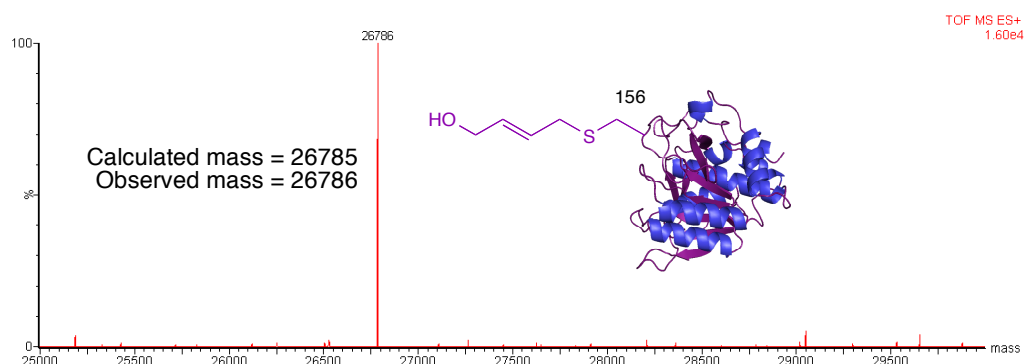
This experiment was carried out with Angel Lin.

All manipulations were carried out at room temperature. The protein was prepared as described above by allylation with allyl chloride. A solution of Hoveyda-Grubbs 2nd generation catalyst in *t*BuOH was prepared by vortexing and gently warming 2.2 mg of catalyst in 355 μ L *t*BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (250 μ L, 0.57 mg/mL, 5.3 nmol) was added $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (10.8 mg, 53 μ mol) followed by an aliquot of the catalyst solution (107 μ L, 1.1 μ mol). The reaction was vortexed after each addition to homogenize. Allyl alcohol (3.6 μ L, 53 μ mol) was then added to the mixture. The reaction was vortexed and then shaken at room temperature. LC-MS after 1 hour of reaction time revealed formation of product. Full conversion to the CM product was observed after 2 hours (calculated mass = 26785; observed mass = 26786).

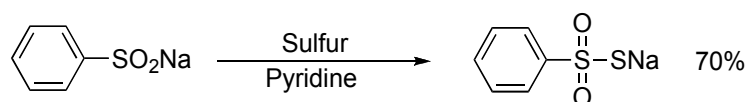
1h:



2h:



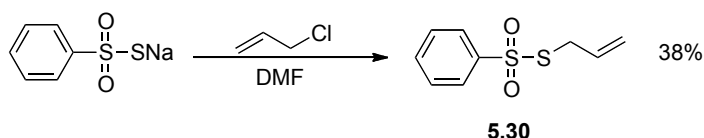
Sodium phenylthiosulfonate



This experiment was performed by Angel Lin (Part II, 2007) and is included here for reference. Sodium benzenesulfinate (10.00 g, 61.00 mmol) and sulfur (1.95 g, 61.0 mmol) were dissolved in anhydrous pyridine (60 mL) to give a yellow solution. The reaction was stirred for 1 h at room temperature under argon, after which time a white suspension had formed. The solid formed was isolated by filtration and washed with Et₂O. Recrystallization from EtOH afforded sodium phenylthiosulfonate (8.40 g, 70%) as a white crystalline solid. Spectroscopic data was consistent with that previous reported.¹⁰

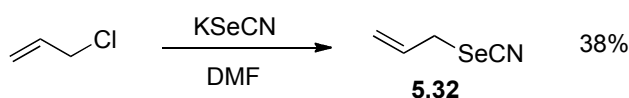
m.p. = 294–295 °C (Lit¹⁰ = 287 °C). IR ($\nu_{\max}$, KBr): 1936, 1900, 1818, 1771, 1684, 1607, 1583, 1475, 1443, 1390, 1333, 1311. ¹H NMR (400 MHz, DMSO) δ = 7.19–7.50 (3H, m, *m,p*-ArH), 7.75 (2H, dd, J = 7.7, 1.5, *o*-ArH). ¹³C NMR (100 MHz, DMSO): δ = 124.0, 127.8, 129.0 (3° Ar), 154.9 (4° Ar). LRMS m/z (ESI⁺): Found 173.0 [M–Na][–]; C₆H₅O₂S₂ requires 173.0.

Allyl phenylthiosulfonate (5.30)



This experiment was performed by Angel Lin (Part II, 2007) and is included here for reference. Sodium phenylthiosulfonate (2.00 g, 10.2 mmol) was added to a 50 mL round bottom flask containing allyl chloride (2.50 mL, 30.6 mmol) in DMF (10 mL). The reaction was stirred at room temperature for 3 hours. After this time, the reaction was diluted with water (100 mL) and transferred to a separatory funnel. The aqueous layer was extracted with Et₂O (2 × 100 mL). The combined Et₂O layers were washed sequentially with H₂O (100 mL) and brine (100 mL) and then dried (MgSO₄) and filtered. The solvent was removed under reduced pressure and the resulting residue was purified by column chromatography (5% EtOAc in petrol). The titled compound was isolated as a clear oil (830 mg, 38%). IR ($\nu_{\max}$, film): 1447, 1325, 1144, 1078, 988, 929, 755, 716, 685, 598, 537. ¹H NMR (400 MHz, CDCl₃): δ = 3.68 (2H, dt, J = 7.1, 1.1, SCH₂), 5.09 (1H, dd, J = 10.0, 1.1, HC=CHH), 5.21 (1H, dd, J = 17.0, 1.1, HC=CHH), 5.70 (1H, m, J = 17.0, 10.0, 7.1, HC=CH₂), 7.51–7.59 (2H, m, ArH), 7.60–7.68 (1H, m, ArH), 7.93 (2H, d, J = 7.3, ArH). ¹³C NMR (100 MHz, CDCl₃): δ = 38.8 (SCH₂), 120.0 (HC=CH₂), 126.9, 129.2, 130.4, 133.7 (includes 3° Ar and HC=CH₂), 144.8 (4° Ar). Anal. Calcd. for C₉H₁₀O₂S₂: C, 50.44; H, 4.70; S, 29.92. Found C, 50.36; H, 4.76; S, 30.00.

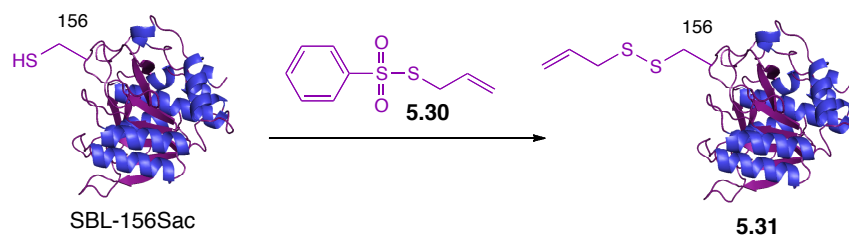
Allyl selenocyanate (5.32)



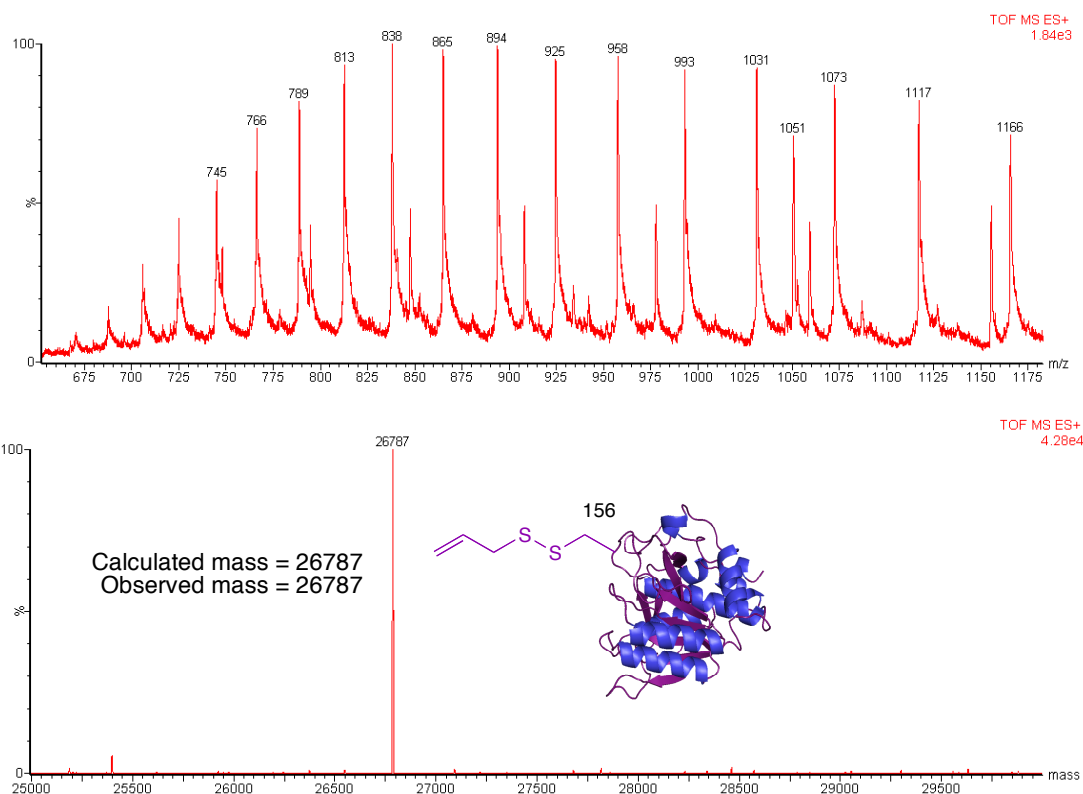
The synthesis of allyl selenocyanate was adapted from the literature.¹¹ KSeCN (3.27 g, 22.7 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (25 mL).

The solution was placed under an atmosphere of nitrogen and allyl chloride (3.72 mL, 45.4 mmol) was added slowly to the stirred solution. The reaction was stirred for 20 minutes at room temperature and then diluted with Et₂O (200 mL) and washed sequentially with H₂O (2 × 200 mL) and brine (200 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The product was isolated as a pale yellow liquid and was sufficiently pure to use in subsequent manipulations (1.25 g, 38%). This material has a sharp, lingering odor and should be used only in a well-ventilated fume hood. IR (ν_{max} , film): 2151, 1633, 1432, 1403, 1197, 987, 928, 851, 686. ¹H NMR (200 MHz, CDCl₃): δ = 3.65 (2H, dt, J = 7.2, 1.0, CH_2SeCN), 5.23 (1H, dd, J = 9.9, 1.0, $\text{CHH}=\text{CH}$), 5.32 (1H, dq, J = 16.7, 1.0, $\text{CHH}=\text{CH}$), 6.00 (1H, m, $\text{H}_2\text{C}=\text{CH}$). ¹³C NMR (50 MHz, CDCl₃): δ = 31.3 (CH_2SeCN), 101.4 ($\text{C}\equiv\text{N}$), 120.6 ($\text{CH}=\text{CH}_2$), 131.6 ($\text{CH}=\text{CH}_2$).

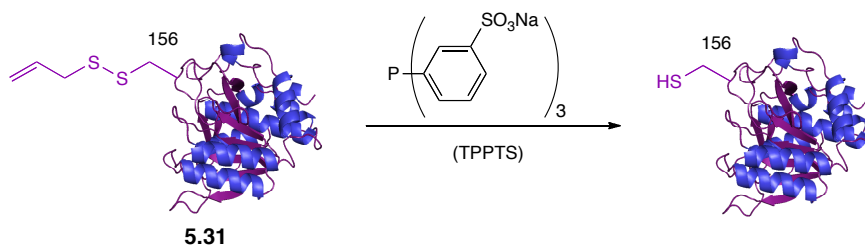
Synthesis of allyl disulfide on SBL-S156C



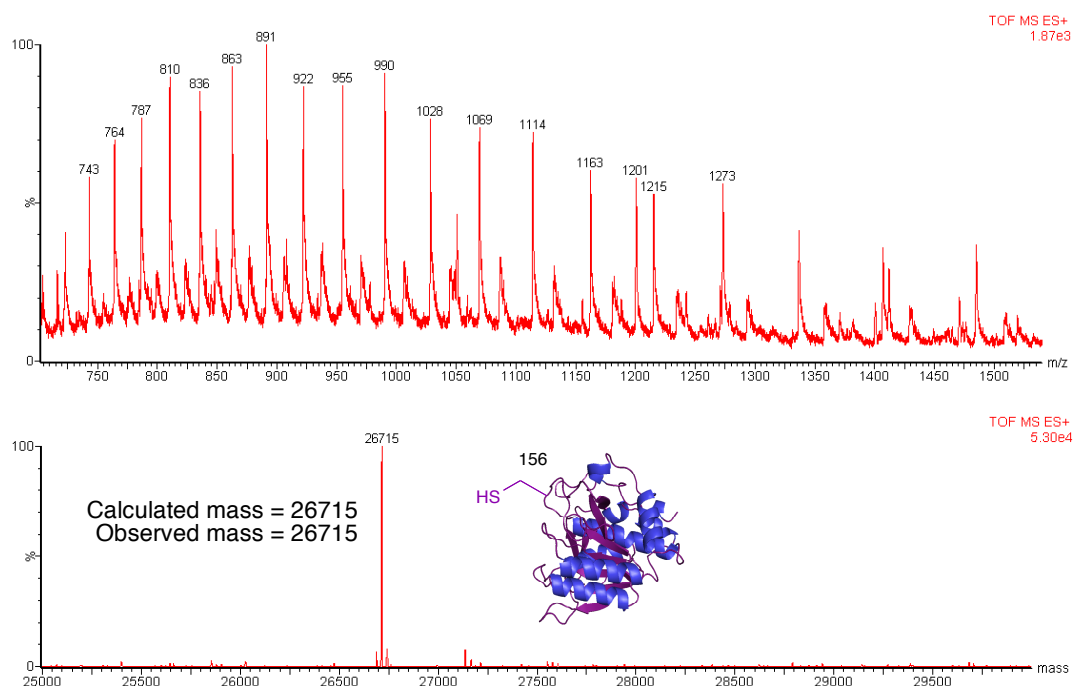
All manipulations were carried out at room temperature. A solution of allyl phenylthiosulfonate **5.30** in DMF was prepared by dissolving 3.3 mg in 406 μL DMF. 100 μL of allyl-PTS solution (0.8 mg, 3.7 μmol) was added to a 1.5 mL eppendorf tube containing SBL-S156C (1 mL, 1 mg/mL, 0.037 μmol , pH 8.0 sodium phosphate). The reaction was vortexed immediately and placed on shaker for 10 minutes. LC-MS analysis revealed full conversion to the allyl disulfide **5.31** (calculated mass = 26787; observed mass = 26787). Small molecules were removed with a PD10 column (GE Healthcare), eluting with 3.5 mL 50 mM sodium phosphate buffer (pH 8.0). The sample was then split into 250 μL aliquots, flash frozen, and stored at -20°C . The final concentration of protein was 0.3 mg/mL and protein recovery was >95%. ESI-MS are shown below.



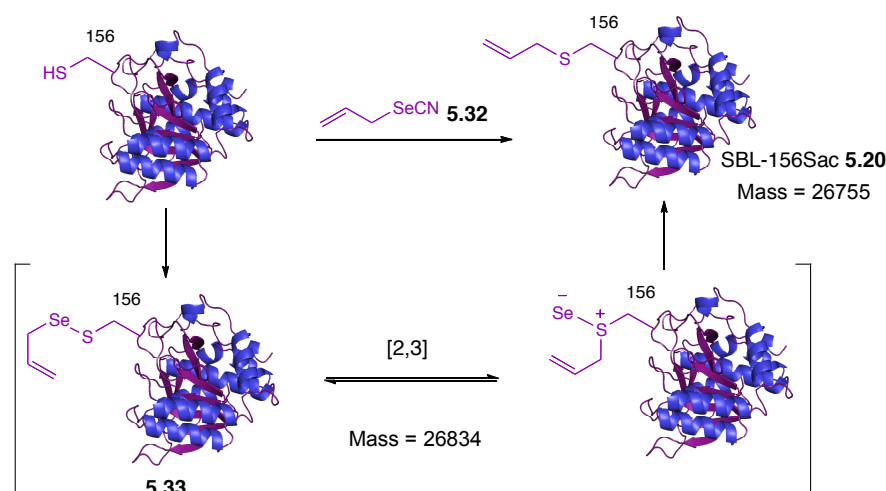
Attempted desulfurization of allyl disulfide with TPPTS



SBL-156-Allyl disulfide **5.31** (250 μL , 0.3 mg/mL, pH 8.0 sodium phosphate, 2.7 nmol) was added to a 1.0 mL plastic tube and stored on ice. A stock solution of tris(3-sulfophenyl)phosphine trisodium salt (TPPTS, Aldrich) was prepared by dissolving 0.9 mg in 112.5 μL of H_2O . A 10 μL aliquot of the TPPTS solution (135 nmol) was added to the protein solution at room temperature. The reaction was shaken at room temperature for 15 minutes and then analyzed by LC-MS. Only free cysteine (product of reduction) was observed. (calculated mass = 26715; observed mass = 26715 found). ESI-MS are shown below.



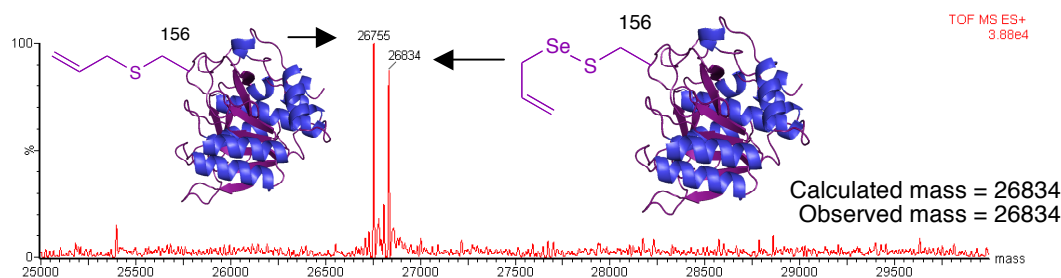
Allylation of SBL-S156C with Allyl Selenocyanate (**5.32**)



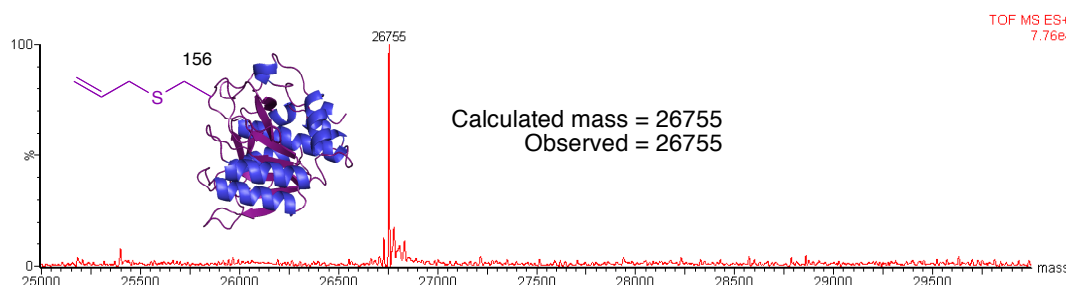
All manipulations were carried out in a well-ventilated hood. SBL-S156C (200 μ L, 1.0 mg/mL, pH 8.0 sodium phosphate, 7.5 nmol) was added to a 1.0 mL plastic tube and stored on ice. A stock solution of allyl selenocyanate **5.32** (11.9 mg) was prepared in DMF (541 μ L). A 5 μ L aliquot of the selenide solution (0.75 μ mol) was added to the protein sample. The reaction was shaken for 10 minutes at room temperature and a 30 μ L aliquot was analyzed by LC-MS. A mixture of selenenyl sulfide **5.33** and allylcysteine **5.20** was observed ($\sim$ 5:4). After 1 hour of total reaction time, the reaction was analyzed again by LC-MS. Full conversion to the allyl sulfide was observed (calculated mass = 26755; observed mass = 26755). ESI-MS shown below.

This transformation was scaled up using 1.0 mL protein (1.0 mg/mL, pH 8.0 sodium phosphate, 37 nmol) and 0.55 mg allyl selenocyanate (3.7 μ mol in 25 μ L DMF). Complete conversion was observed after 1 hour at room temperature. Small molecules were removed by passing the sample through a PD10 column previously equilibrated with pH 8.0 sodium phosphate (50 mM). The protein concentration was 0.3 mg/mL, with >95% protein recovery. This sample was used for subsequent metathesis reactions.

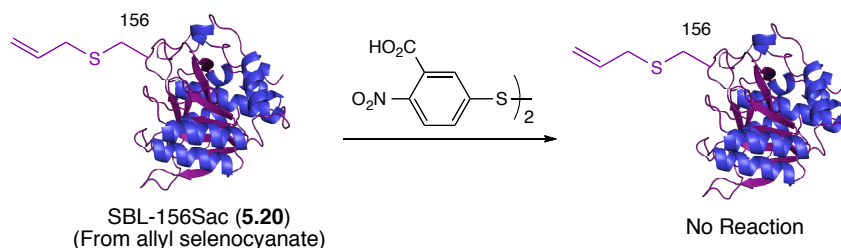
10 min RT:



60 min RT:

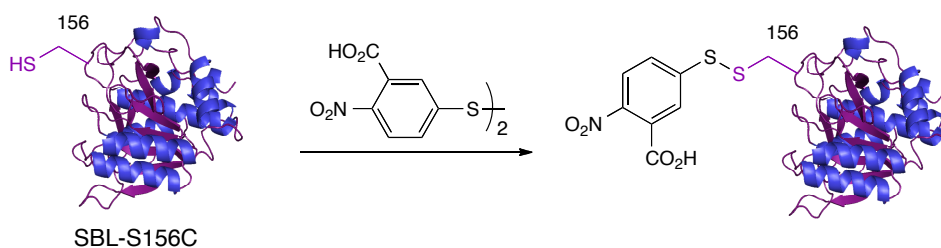
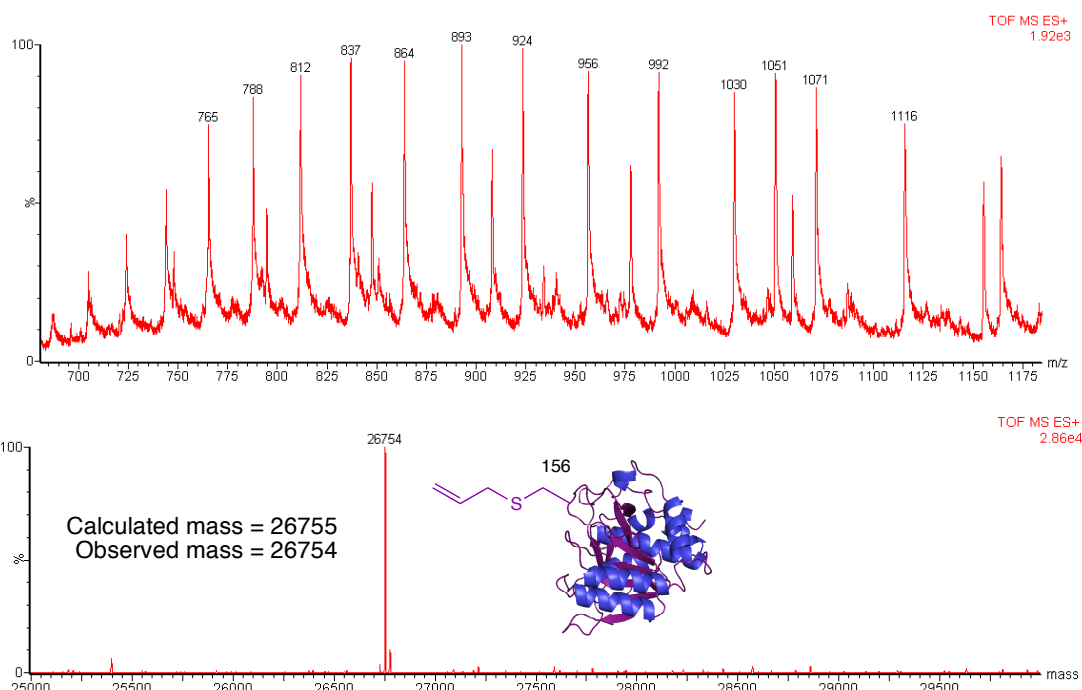


Ellman's Test to Confirm Reaction at Cysteine

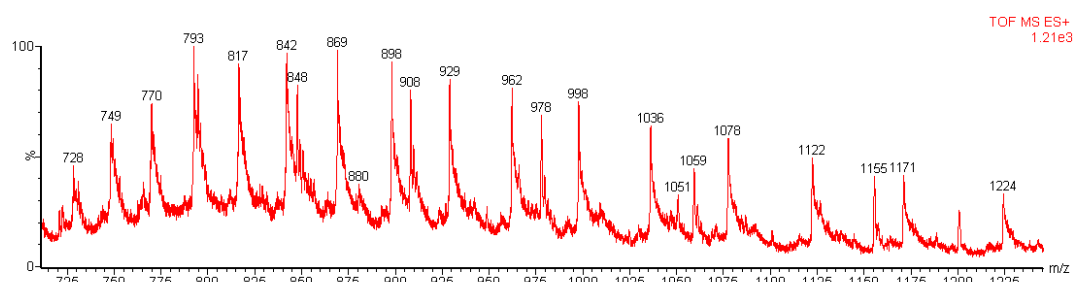


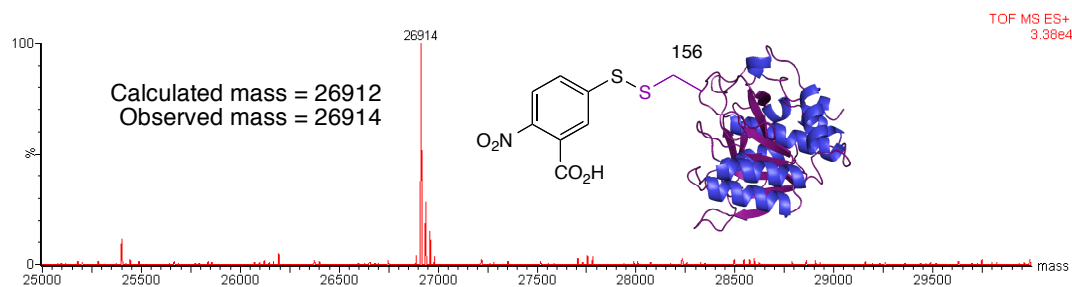
SBL-156Sac (prepared above from allyl selenocyanate and purified by PD10) (100 μ L, 0.3 mg/mL, 1.1 nmol) was added to 250 μ L plastic tube. A stock solution of Ellman's reagent was prepared by dissolving 1.1 mg in 244 μ L of pH 8.0 sodium phosphate (50 mM). 20 μ L of the Ellman's solution (217 nmol) was added to the protein sample and shaken 10 minutes at room temperature. The reaction was analyzed by LC-MS and only allyl cysteine was observed, indicating no free cysteine was present in the protein and the

allylation occurred at cysteine (calculated mass = 26755; observed mass = 26754). Under identical conditions SBL-S156C gave full conversion to the Ellman's disulfide (see below). ESI-MS are shown below.

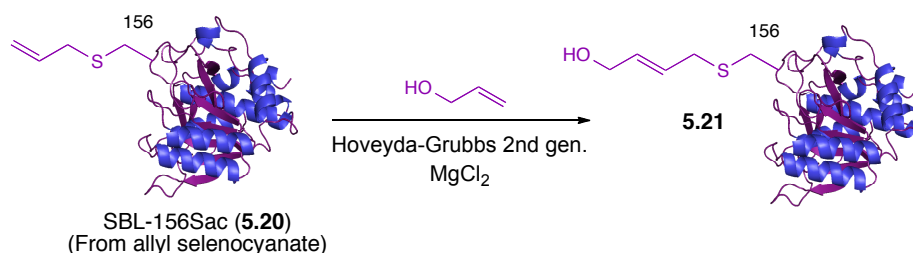


SBL-S156C (100 μ L, 0.3 mg/mL, 1.1 nmol) was added to 250 μ L plastic tube. A 20 μ L aliquot of the Ellman's solution prepared above (217 nmol) was added to the protein sample. The reaction was shaken 10 minutes at room temperature. The reaction was analyzed by LC-MS revealing full conversion to the Ellman disulfide adduct (calculated mass = 26912; observed mass = 26914). ESI-MS are shown below.





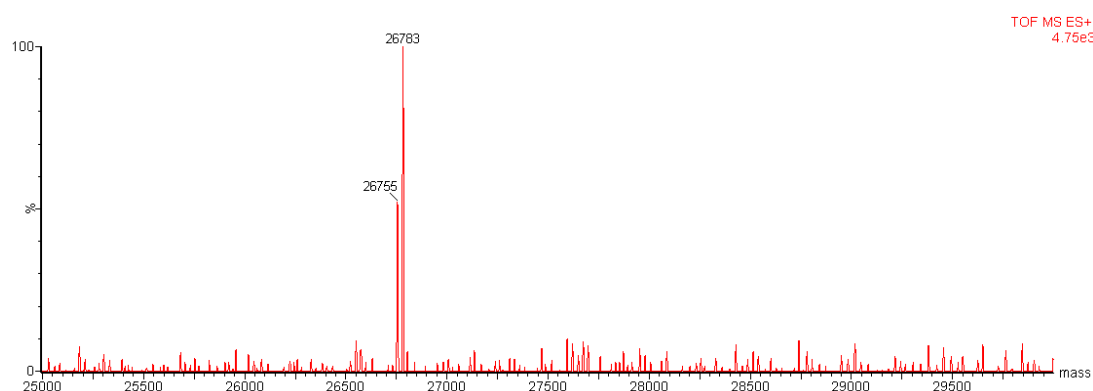
Olefin metathesis on SBL-156Sac after allylation using allyl selenocyanate



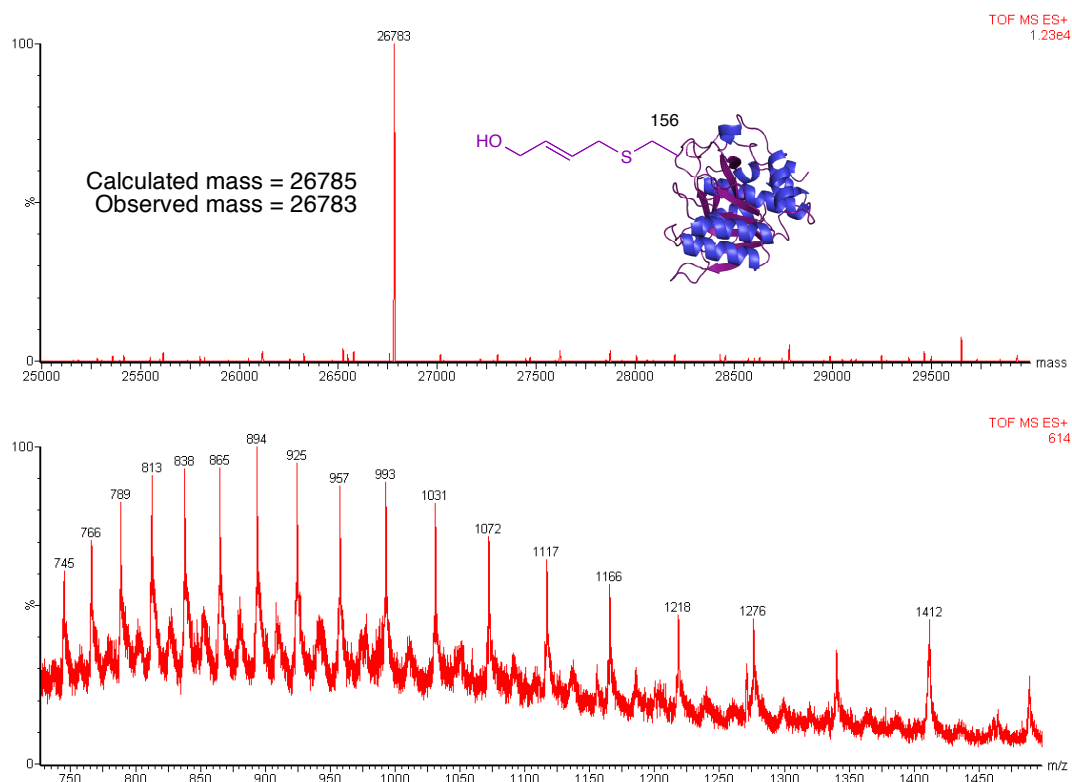
This experiment was carried out with Angel Lin.

All manipulations were carried out at room temperature. The protein was prepared as described above by allylation with allyl selenocyanate. A solution of Hoveyda-Grubbs 2nd generation catalyst (HG-II) in *t*BuOH was prepared by vortexing and gently warming 1.4 mg of HG-II in 444 μ L *t*BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (250 μ L, 0.29 mg/mL, 2.7 nmol) was added MgCl₂·6H₂O (5.5 mg, 27 μ mol) followed by an aliquot of the HG-II/ *t*BuOH solution (107 μ L, 0.54 μ mol). The reaction was vortexed after each addition to homogenize. Allyl alcohol (1.8 μ L, 27 μ mol) was then added to the mixture. The reaction was vortexed and then shaken at room temperature. LC-MS after 1 hour of reaction time revealed formation of product. Full conversion to the CM product was observed after 2 hours (calculated mass = 26785; observed mass = 26783). ESI-MS are shown below.

1h RT:



2h RT:



Peptidase Activity Assay of Allylated SBL-S156C

SBL-S156C (unmodified), SBL-156Sac (from allyl chloride), and SBL-156Sac (from allyl selenocyanate) were prepared at a concentration of 0.1 mg/mL in pH 8.0 sodium phosphate (50 mM). 225 μ L aliquots of each sample were added to a 96-well plate. A 25 μ L aliquot of SucAAPFpNA (0.20 M in DMSO, Bachem) was added to each of the protein samples. All reactions turned yellow immediately upon addition of the peptide substrate. The yellow solution indicates liberation of *p*-nitroaniline (pNA), confirming peptidase activity of all samples. All protein solutions and the peptide solution alone at the same concentration are colorless (See below).



- Well 1: SucAAPFpNA
 Well 2: SBL-S156C
 Well 3: SBL-156Sac (from allyl chloride)
 Well 4: SBL-156Sac (from allyl selenocyanate)
 Well 5: SBL-S156C + SucAAPFpNA
 Well 6: SBL-156Sac (from allyl chloride) + SucAAPFpNA
 Well 7: SBL-156Sac (from allyl selenocyanate) + SucAAPFpNA

Olefin Metathesis Studies on Green Fluorescent Protein

The triple mutant GFP (Y39C, C48L, C72V) was expressed and purified by Dr. Huiwang Ai of The Scripps Research Institute (Prof. Peter Schultz group). The sequence is shown below with mutations highlighted:

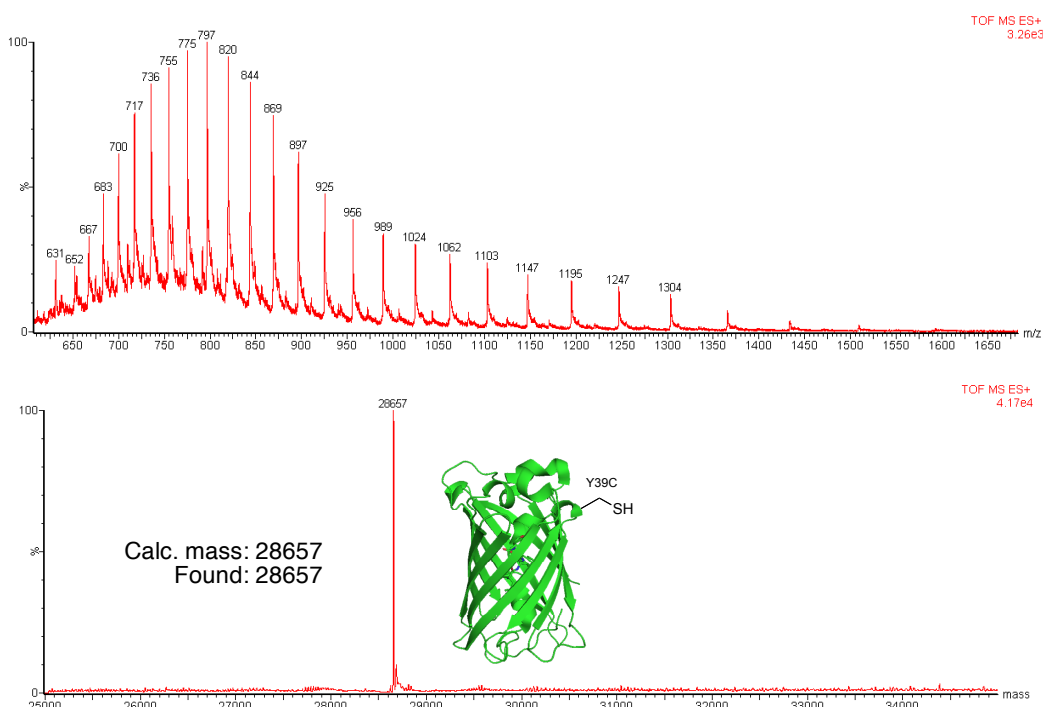
```

M D Y K D D D D K V S K G E E L F T G
V V P I L V E L D G D V N G H K F S V S
G E G E G D A T C G K L T L K F I L T T
G K L P V P W P T L V T T L T Y G V Q V
F S R Y P D H M K Q H D F F K S A M P E
G Y V Q E R T I F F K D D G N Y K T R A
E V K F E G D T L V N R I E L K G I D F
K E D G N I L G H K L E Y N Y N S H N V
Y I M A D K Q K N G I K A N F K I R H N
I E D G S V Q L A D H Y Q Q N T P I G D
G P V L L P D N H Y L S T Q S A L S K D
P N E K R D H M V L L E F V T A A G I T
L G M D E L Y K H H H H H H
  
```

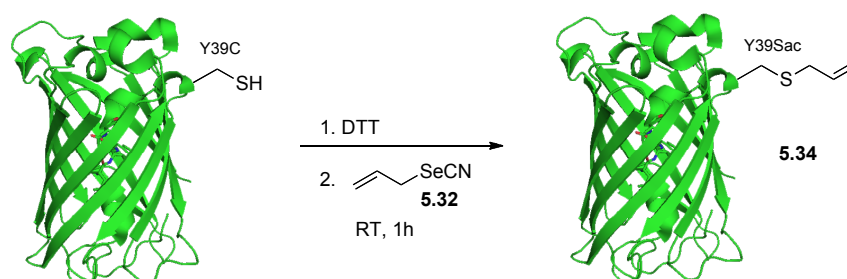
The calculated average mass of this sequence is 28677 Daltons. Upon maturation and loss of H₂O and H₂, the calculated average mass of the protein is **28657** Daltons.

GFP (Y39C, C48L, C72V)

The triple mutant of GFP was supplied by Huiwang Ai of the Scripps Research Institute (Prof. Peter Schultz group). The sample was fluorescent upon receipt and supplied at a concentration of 1.0 mg/mL in pH 8.0 sodium phosphate buffer (50 mM). A solution of DTT was prepared by dissolving 3.5 mg in 700 μ L of pH 8.0 sodium phosphate buffer (50 mM). 4 μ L of this DTT solution was added to a 40 μ L aliquot of the GFP protein solution. The sample was incubated for 10 minutes at room temperature and then analyzed by LC-MS. A single peak was observed (calculated mass = 28657; observed mass = 28657). ESI-MS are shown below.

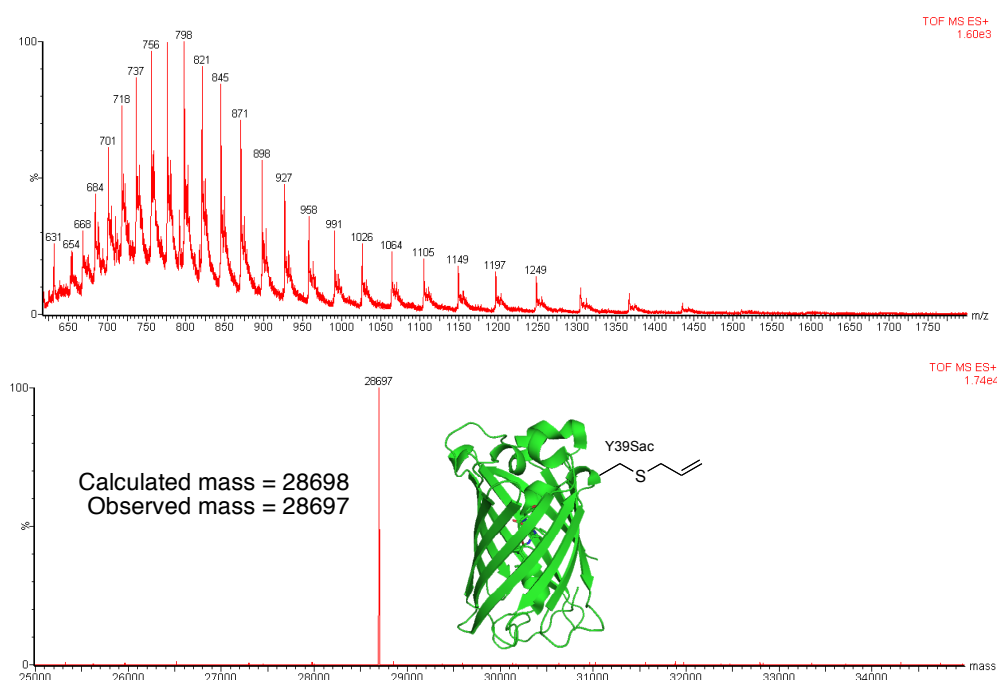


Allylation of GFP using allyl selenocyanate

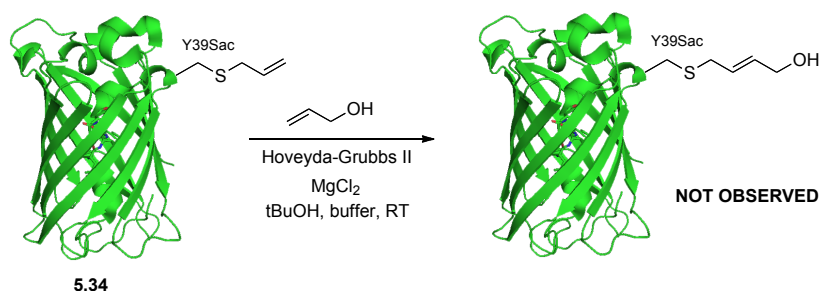


A 100 μ L aliquot of GFP (Y39C, C48L, C72V) was thawed and stored on ice. The protein was supplied at a concentration of 1.0 mg/mL in sodium phosphate buffer (50 mM). A stock solution of DTT was prepared by dissolving 3.5 mg (0.023 mmol) in 700

μL of pH 8.0 sodium phosphate buffer (50 mM). A stock solution of allyl selenocyanate was prepared by dissolving 14.3 mg (0.098 mmol) in 140 μL of DMF. A 10 μL aliquot of the DTT solution (0.035 μmol) was added to the GFP sample and shaken at room temperature for 10 minutes to reduce any contaminant disulfide. After reduction, a 5 μL aliquot of the allyl selenocyanate (3.5 μmol) was added to the protein solution. The reaction was vortexed and shaken at room temperature for 1 hour. The reaction was analyzed directly by LC-MS, revealing complete conversion to the allylated product (calculated mass = 28698; observed mass = 28697). ESI-MS are shown below.

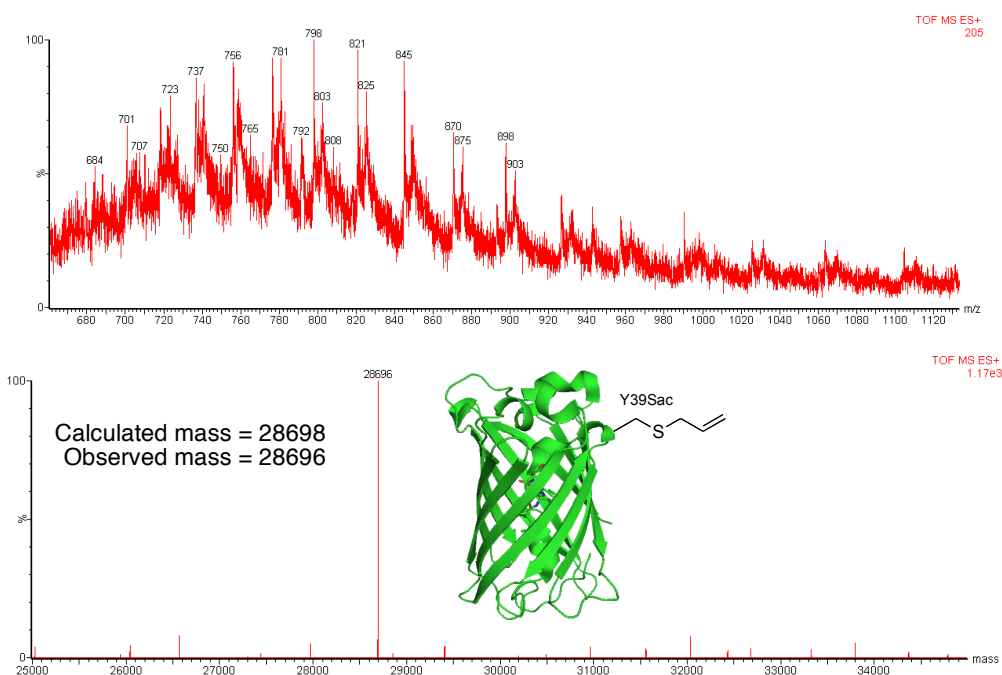


A second batch of GFP-Y39Sac was prepared in the same manner on a 400 μL scale. LC-MS revealed full conversion to the allylated product. Both batches were combined and passed through a PD Minitrapp to remove small molecules. The product was eluted with pH 8.0 sodium phosphate (50 mM). The final concentration was 0.5 mg/mL and ESI-MS revealed no change after purification. The sample was split into 200 μL aliquots. One sample was used immediately for a metathesis reaction and the others were flash frozen and stored at $-20\text{ }^{\circ}\text{C}$. The protein remained fluorescent throughout all manipulations described above.



This experiment was carried out with Angel Lin.

GFP-39Sac **5.34** (mw = 28698, 0.5 mg/mL, 200 μ L, 3.48 nmol) was transferred to a 1.5 mL eppendorf tube and placed on ice. A t BuOH solution of Hoveyda-Grubbs 2nd generation catalyst was prepared by dissolving 1.3 mg of Ru catalyst with 253.2 μ L t BuOH in a 1.5 mL eppendorf tube. $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$ (6.7 mg, 33 μ mol) was added as a solid to the protein solution and vortexed immediately to mix. The protein is still fluorescent at this stage. 85.7 μ L of catalyst/ t BuOH solution was added to the reaction and vortexed immediately to mix. The reaction mixture appeared as a pale green emulsion. The reaction was shaken at RT for 1 minute before adding allyl alcohol (2.4 μ L, 34.8 μ mol). The reaction was vortexed immediately after addition and was shaken at RT. After 1 hour LC-MS analysis of the product showed only the starting material. The reaction was analyzed by LC-MS after an additional hour and only the starting material was found. ESI-MS are shown below.



5.6 References for Experimental Procedures

-
- (1) Threadgill, M. D.; Gledhill, A. P. *J. Org. Chem.* **1989**, *54*, 2940-2949.
 - (2) Kuligowski, C.; Bezenine-Lafollé, S.; Chaume, G.; Mahuteau, J.; Barrière, J. -C.; Bacqué, E.; Pancrazi, A.; Ardisson, J. *J. Org. Chem.* **2002**, *67*, 4565-4568.
 - (3) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642-9643.
 - (4) Streiff, S.; Ribeiro, N.; Désaubry, L. *J. Org. Chem.* **2004**, *69*, 7592-7598.
 - (5) Spagnol, G.; Heck, M. -P.; Nolan, S. P.; Mioskowski, C. *Org. Lett.* **2002**, *4*, 1767-1770.
 - (6) Rodriguez, E. B.; Stick, R. V. *Aust. J. Chem.* **1990**, *43*, 665-679.
 - (7) Kishida, M.; Akita, H. *Tetrahedron* **2005**, *61*, 10559-10568.
 - (8) Utile, J. -P.; Priem, B. *Carbohydr. Res.* **2000**, *329*, 431-439.
 - (9) Tamura, Y.; Minamikawa, J.; Sumoto, K.; Fujii, S.; Ikeda, M. *J. Org. Chem.* **1973**, *38*, 1239-1241.
 - (10) R. Sato, T. Goto, Y. Takikawa, S. Takizawa, *Synthesis* 1980, 615.
 - (11) E. H. Riague, J.-C. Guillemin, *Organometallics* 2002, **21**, 68-73.

5.7 X-Ray Crystallographic Data and NMR Spectra

X-Ray crystallographic analysis was carried out by Dr. Amber Thompson.

Table 1. Crystal data and structure refinement for 6318.

Identification code	6318	
Empirical formula	C ₁₆ H ₁₆ S ₂	
Formula weight	272.43	
Temperature	150 K	
Wavelength	0.71073 Å	
Crystal system	Triclinic	
Space group	P -1	
Unit cell dimensions	a = 5.5297(3) Å	α = 91.576(2)°.
	b = 7.7052(4) Å	β = 95.898(2)°.
	c = 16.3123(10) Å	γ = 90.972(2)°.
Volume	690.95(7) Å ³	
Z	2	
Density (calculated)	1.309 Mg/m ³	
Absorption coefficient	0.364 mm ⁻¹	
F(000)	288	
Crystal size	0.360 x 0.300 x 0.060 mm ³	
Theta range for data collection	5.296 to 26.026°.	
Index ranges	-6 ≤ h ≤ 6, -9 ≤ k ≤ 8, -20 ≤ l ≤ 20	
Reflections collected	9658	
Independent reflections	2693 [R(int) = 0.050]	
Completeness to theta = 25.505°	98.9 %	
Absorption correction	Semi-empirical from equivalents	
Max. and min. transmission	0.98 and 0.89	
Refinement method	Full-matrix least-squares on F ²	
Data / restraints / parameters	2692 / 0 / 163	
Goodness-of-fit on F ²	0.9305	
Final R indices [I > 2σ(I)]	R ₁ = 0.0612, wR ₂ = 0.1250	
R indices (all data)	R ₁ = 0.0895, wR ₂ = 0.1386	
Largest diff. peak and hole	0.59 and -0.61 e.Å ⁻³	

Table 2. Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for 6318. $U(\text{eq})$ is defined as one third of the trace of the orthogonalized U^{ij} tensor.

	x	y	z	U(eq)
C(1)	693	5425	5280	39
C(2)	934	5003	6184	37
S(3)	3968	4206	6437	31
C(4)	4146	3880	7509	24
C(5)	6161	2940	7838	26
C(6)	6472	2620	8672	31
C(7)	4813	3214	9191	29
C(8)	2834	4144	8870	29
C(9)	2486	4472	8031	29
C(10)	4128	516	4844	35
C(11)	3804	1003	3966	36
S(12)	820	244	3516	31
C(13)	718	941	2493	26
C(14)	-1288	402	1948	30
C(15)	-1459	868	1134	33
C(16)	325	1869	839	32
C(17)	2298	2438	1375	31
C(18)	2516	1981	2200	28

Table 3. Bond lengths [Å] and angles [°] for 6318.

C(1)-C(1)#1	1.287	C(17)-H(171)	0.944
C(1)-C(2)	1.513	C(18)-H(181)	0.943
C(1)-H(11)	0.951		
C(2)-S(3)	1.808	C(1)#1-C(1)-C(2)	124.066
C(2)-H(21)	0.966	C(1)#1-C(1)-H(11)	118.7
C(2)-H(22)	0.969	C(2)-C(1)-H(11)	117.2
S(3)-C(4)	1.766	C(1)-C(2)-S(3)	107.254
C(4)-C(5)	1.407	C(1)-C(2)-H(21)	109.7
C(4)-C(9)	1.388	S(3)-C(2)-H(21)	111.6
C(5)-C(6)	1.383	C(1)-C(2)-H(22)	108.7
C(5)-H(51)	0.938	S(3)-C(2)-H(22)	109.5
C(6)-C(7)	1.386	H(21)-C(2)-H(22)	110.0
C(6)-H(61)	0.924	C(2)-S(3)-C(4)	104.059
C(7)-C(8)	1.383	S(3)-C(4)-C(5)	115.455
C(7)-H(71)	0.919	S(3)-C(4)-C(9)	125.337
C(8)-C(9)	1.394	C(5)-C(4)-C(9)	119.207
C(8)-H(81)	0.952	C(4)-C(5)-C(6)	119.969
C(9)-H(91)	0.928	C(4)-C(5)-H(51)	118.4
C(10)-C(10)#2	1.330	C(6)-C(5)-H(51)	121.6
C(10)-C(11)	1.485	C(5)-C(6)-C(7)	120.691
C(10)-H(101)	0.953	C(5)-C(6)-H(61)	118.7
C(11)-S(12)	1.816	C(7)-C(6)-H(61)	120.6
C(11)-H(111)	0.969	C(6)-C(7)-C(8)	119.493
C(11)-H(112)	0.986	C(6)-C(7)-H(71)	120.6
S(12)-C(13)	1.763	C(8)-C(7)-H(71)	119.8
C(13)-C(14)	1.398	C(7)-C(8)-C(9)	120.586
C(13)-C(18)	1.397	C(7)-C(8)-H(81)	118.9
C(14)-C(15)	1.378	C(9)-C(8)-H(81)	120.5
C(14)-H(141)	0.940	C(8)-C(9)-C(4)	120.050
C(15)-C(16)	1.375	C(8)-C(9)-H(91)	119.0
C(15)-H(151)	0.922	C(4)-C(9)-H(91)	120.9
C(16)-C(17)	1.383	C(10)#2-C(10)-C(11)	123.653
C(16)-H(161)	0.922	C(10)#2-C(10)-H(101)	117.3
C(17)-C(18)	1.393	C(11)-C(10)-H(101)	119.1

C(10)-C(11)-S(12)	108.722	C(14)-C(15)-C(16)	121.137
C(10)-C(11)-H(111)	111.6	C(14)-C(15)-H(151)	119.1
S(12)-C(11)-H(111)	108.6	C(16)-C(15)-H(151)	119.8
C(10)-C(11)-H(112)	110.6	C(15)-C(16)-C(17)	119.077
S(12)-C(11)-H(112)	108.9	C(15)-C(16)-H(161)	119.8
H(111)-C(11)-H(112)	108.5	C(17)-C(16)-H(161)	121.1
C(11)-S(12)-C(13)	102.908	C(16)-C(17)-C(18)	120.878
S(12)-C(13)-C(14)	117.341	C(16)-C(17)-H(171)	118.8
S(12)-C(13)-C(18)	123.962	C(18)-C(17)-H(171)	120.3
C(14)-C(13)-C(18)	118.696	C(13)-C(18)-C(17)	119.779
C(13)-C(14)-C(15)	120.415	C(13)-C(18)-H(181)	119.0
C(13)-C(14)-H(141)	119.7	C(17)-C(18)-H(181)	121.2
C(15)-C(14)-H(141)	119.8		

Symmetry transformations used to generate equivalent atoms:

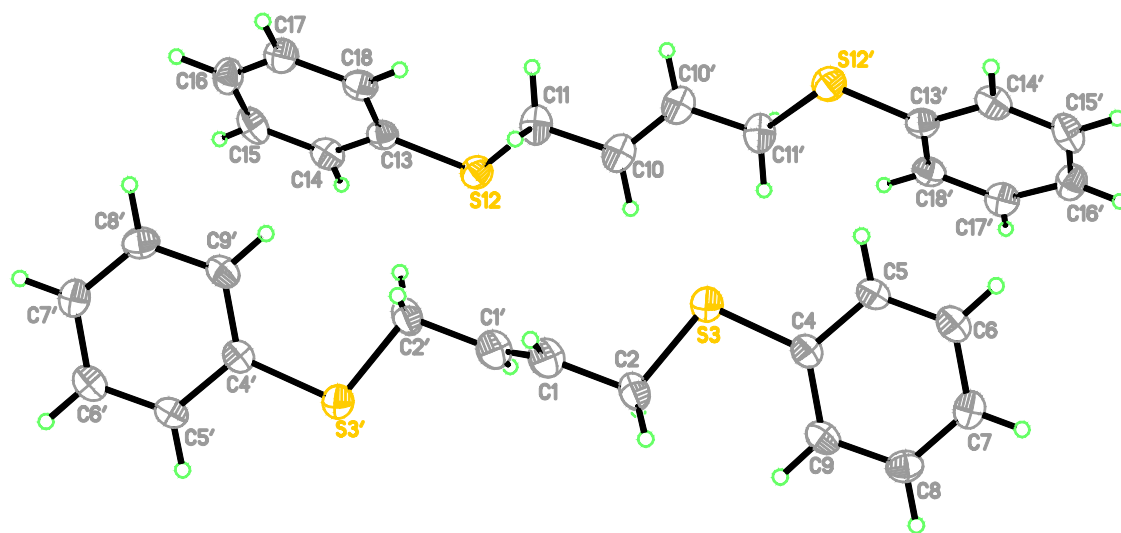
#1 -x,-y+1,-z+1 #2 -x+1,-y,-z+1

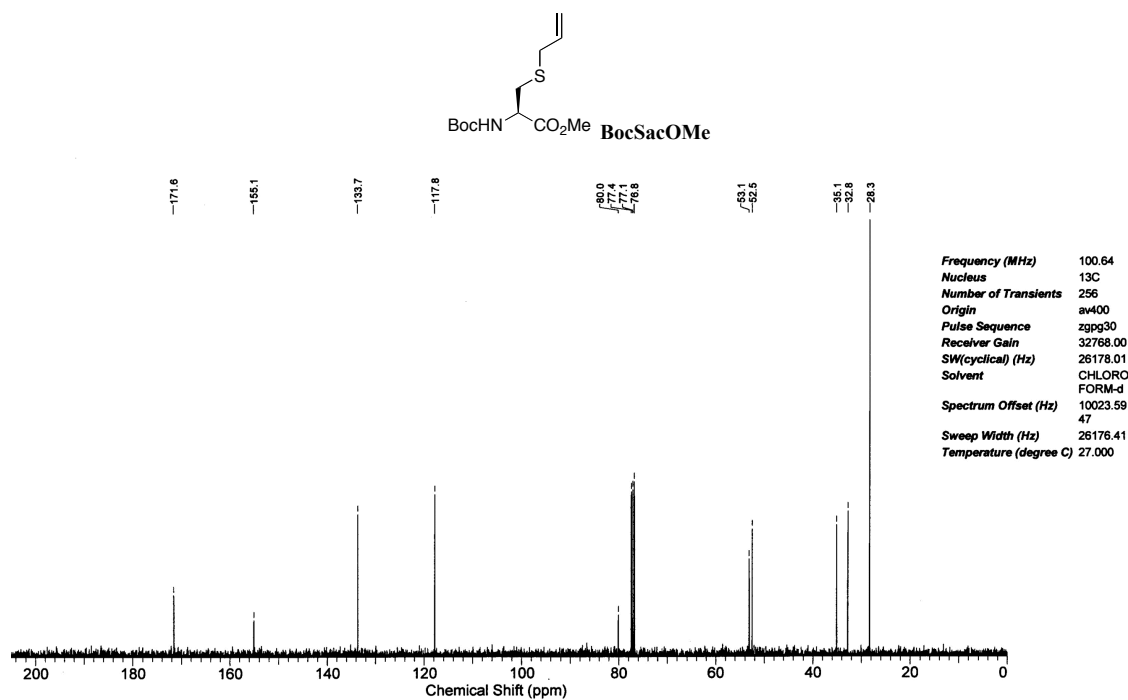
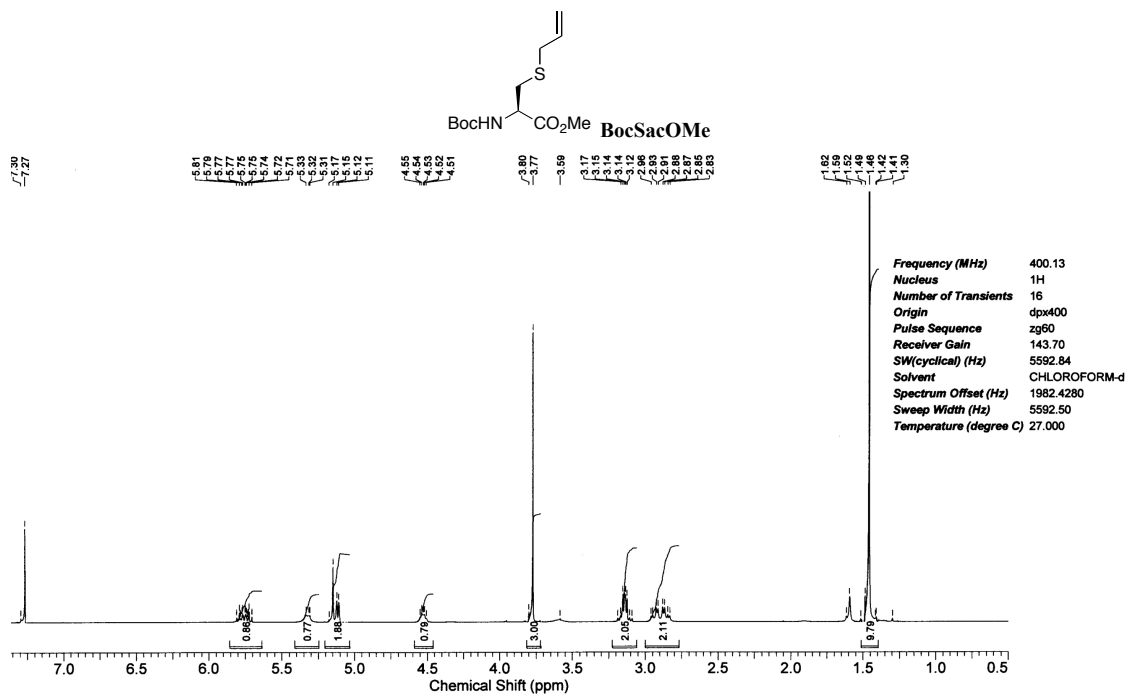
Table 4. Anisotropic displacement parameters ($\text{\AA}^2 \times 10^3$) for 6318. The anisotropic displacement factor exponent takes the form: $-2\pi^2 [h^2 a^{*2} U^{11} + \dots + 2 h k a^* b^* U^{12}]$

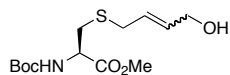
	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
C(1)	42	38	38	7	4	8
C(2)	36	43	31	2	-6	12
S(3)	30	38	26	3	2	3
C(4)	23	21	28	0	-1	-2
C(5)	22	25	32	0	2	3
C(6)	27	28	37	2	-1	3
C(7)	35	26	26	2	0	-3
C(8)	28	28	33	-2	8	-2
C(9)	25	25	36	4	2	4
C(10)	40	36	29	2	3	-4
C(11)	39	37	32	4	-2	-3
S(12)	32	30	31	4	5	-1
C(13)	28	20	30	-1	5	3
C(14)	26	24	39	3	4	-2
C(15)	30	27	40	0	-9	2
C(16)	42	26	28	3	2	4
C(17)	33	26	35	4	5	-1
C(18)	23	29	34	-1	4	0

Table 5. Hydrogen coordinates ($\times 10^4$) and isotropic displacement parameters ($\text{\AA}^2 \times 10^{-3}$) for 6318.

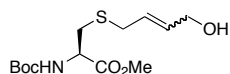
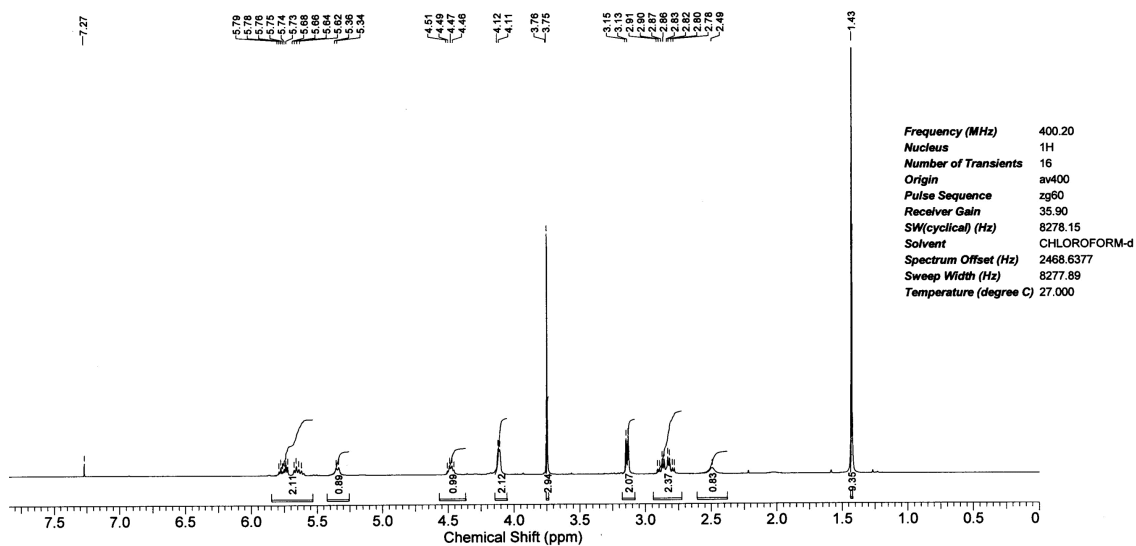
	x	y	z	U(eq)
H(11)	1676	6355	5118	48
H(21)	647	6031	6510	44
H(22)	-246	4101	6270	45
H(51)	7255	2532	7478	30
H(61)	7817	2013	8877	37
H(71)	4998	2975	9742	34
H(81)	1707	4550	9232	36
H(91)	1154	5108	7831	34
H(101)	3056	976	5214	41
H(111)	5023	485	3655	44
H(112)	3912	2275	3920	44
H(141)	-2516	-295	2136	36
H(151)	-2806	512	786	40
H(161)	184	2159	292	37
H(171)	3479	3170	1176	36
H(181)	3889	2327	2563	33



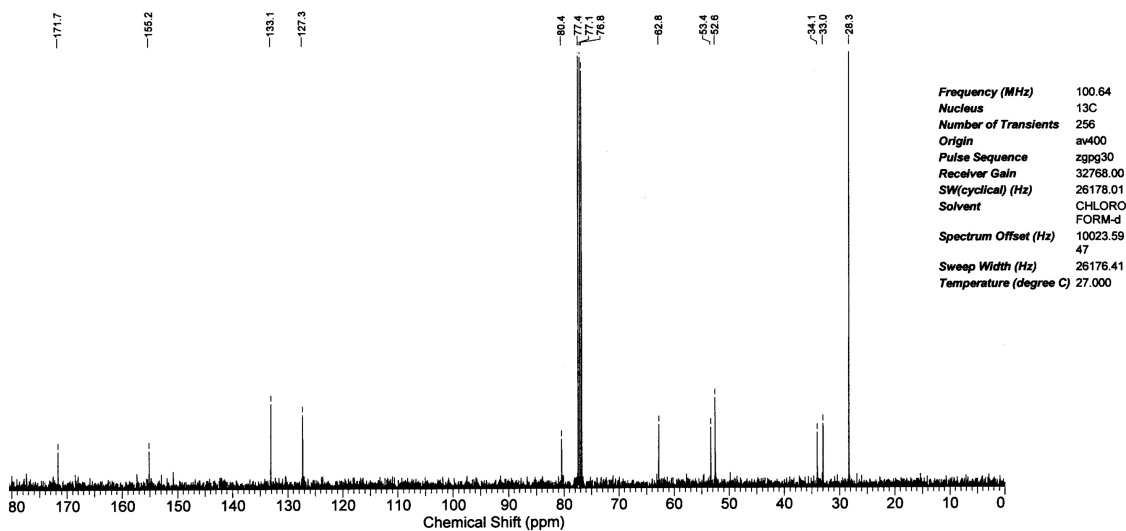


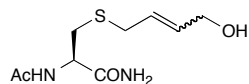


Cross-Metathesis Product of BocSacOMe and allyl alcohol

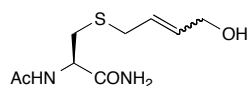
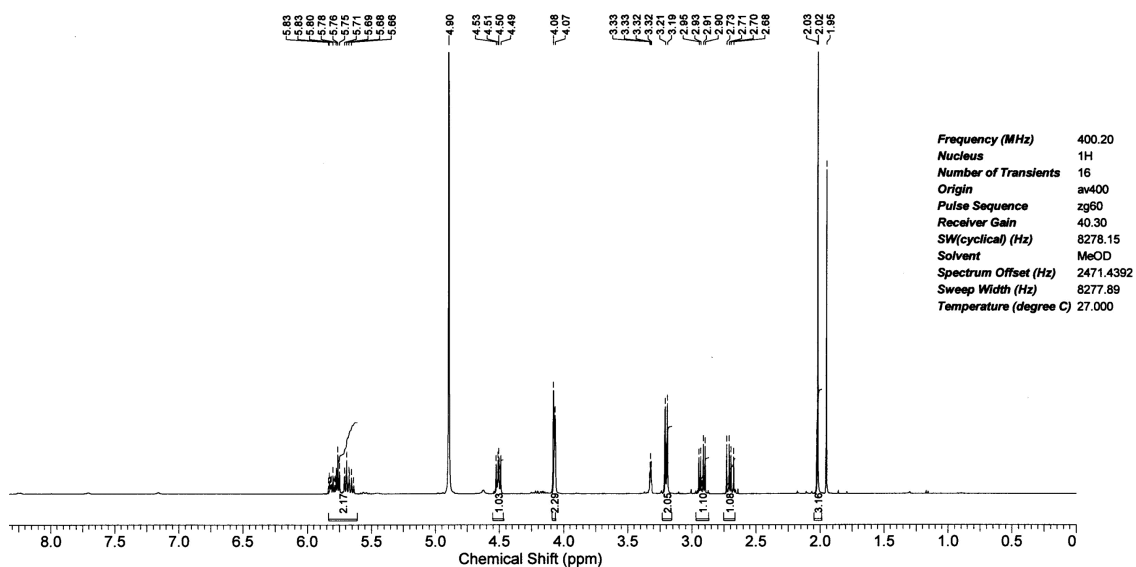


Cross-Metathesis Product of BocSacOMe and allyl alcohol

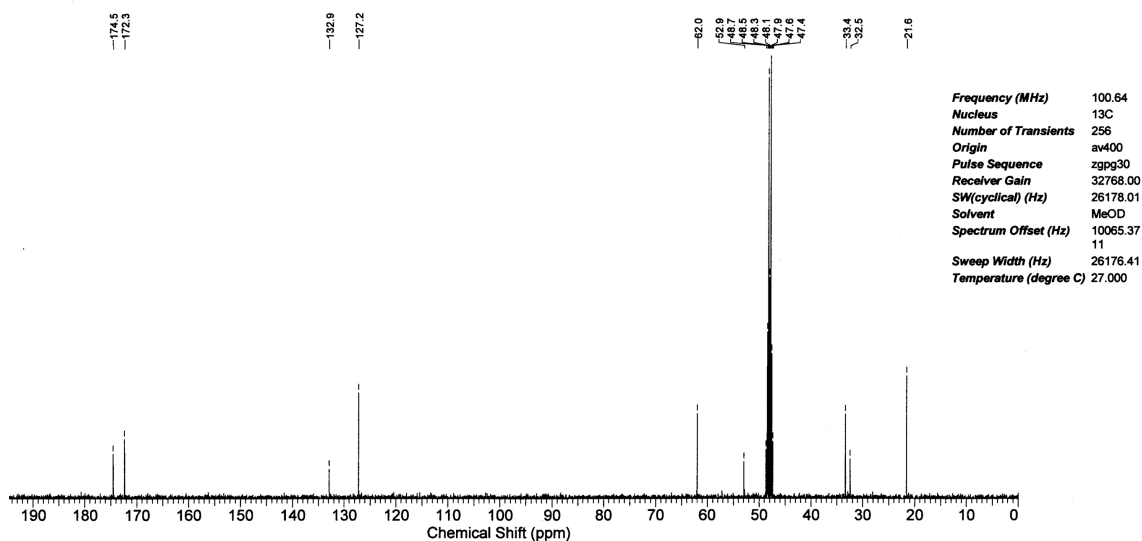


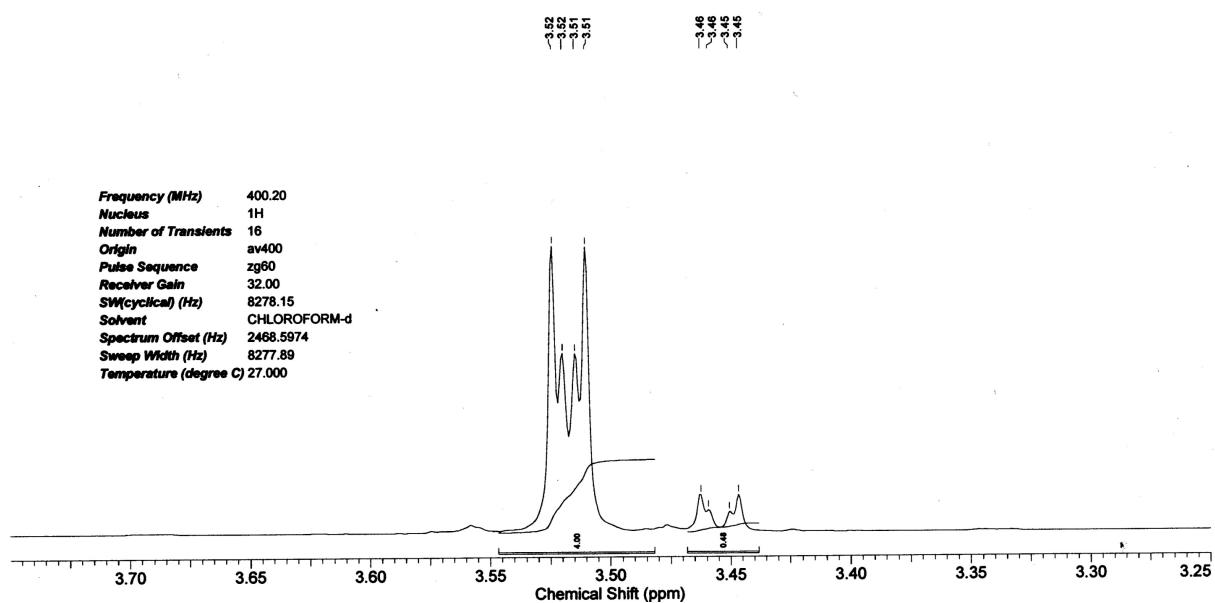
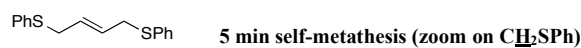
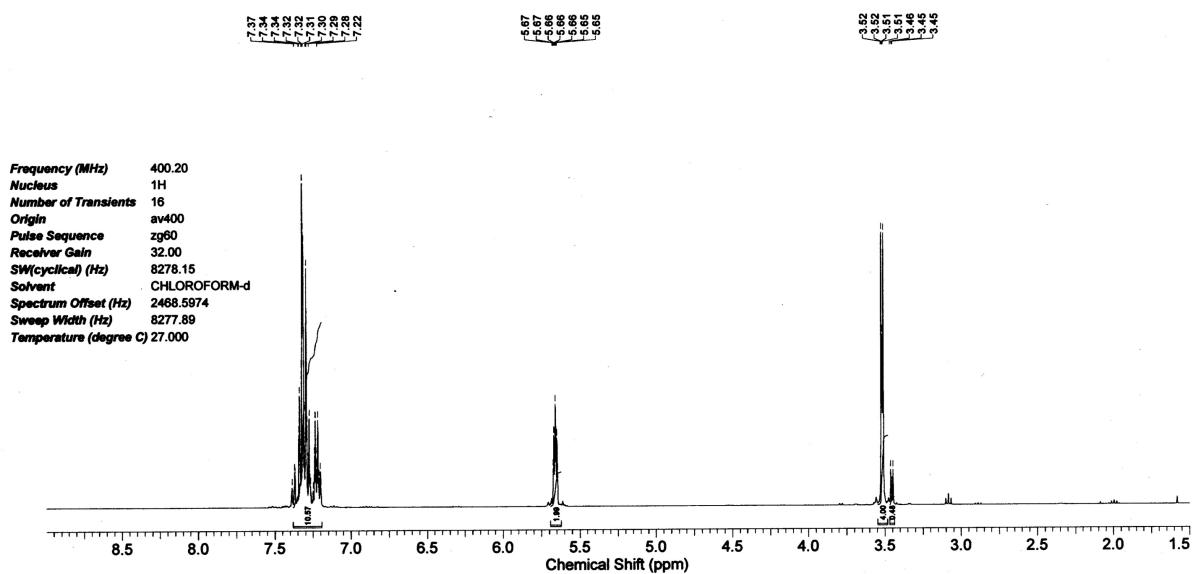
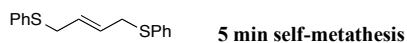


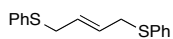
Cross-Metathesis Product of AcSacNH₂ and allyl alcohol



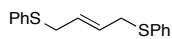
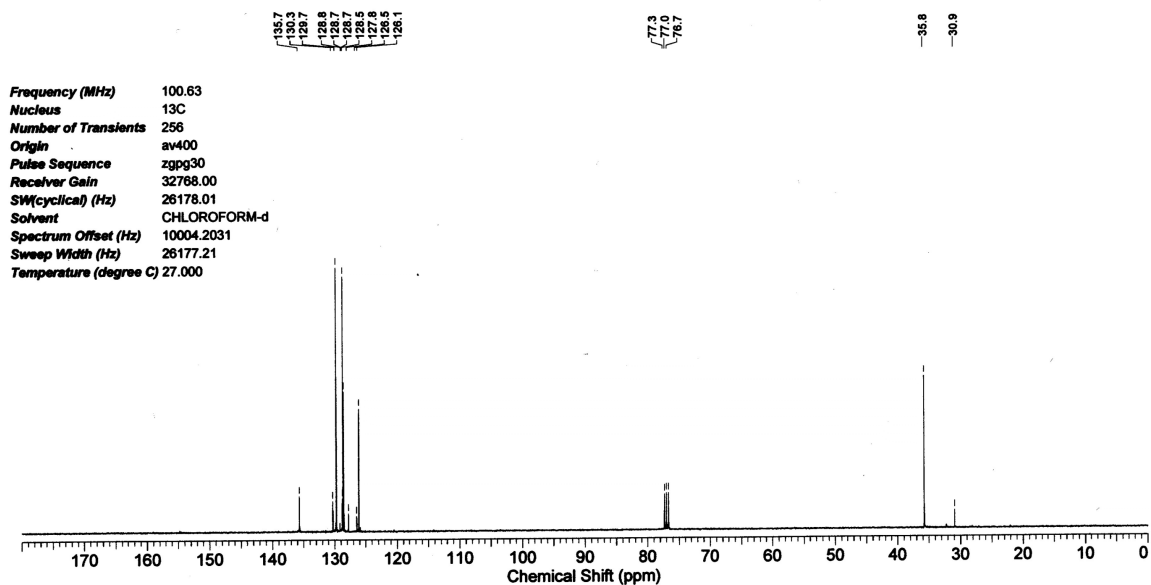
Cross-Metathesis Product of AcSacNH₂ and allyl alcohol



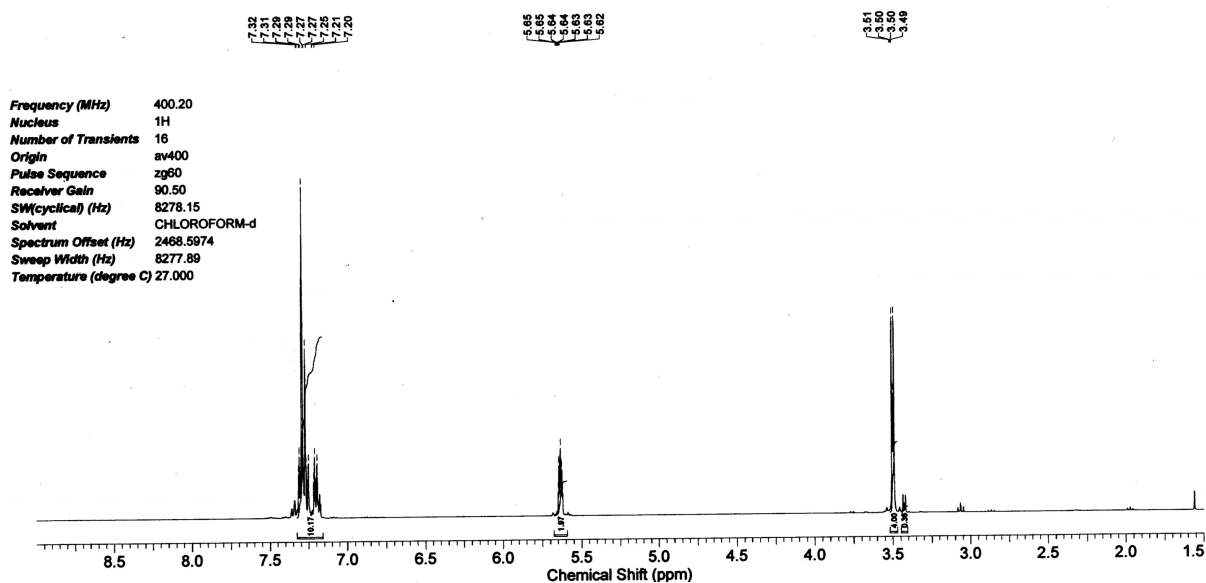


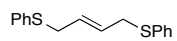


5 min self-metathesis

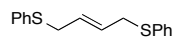
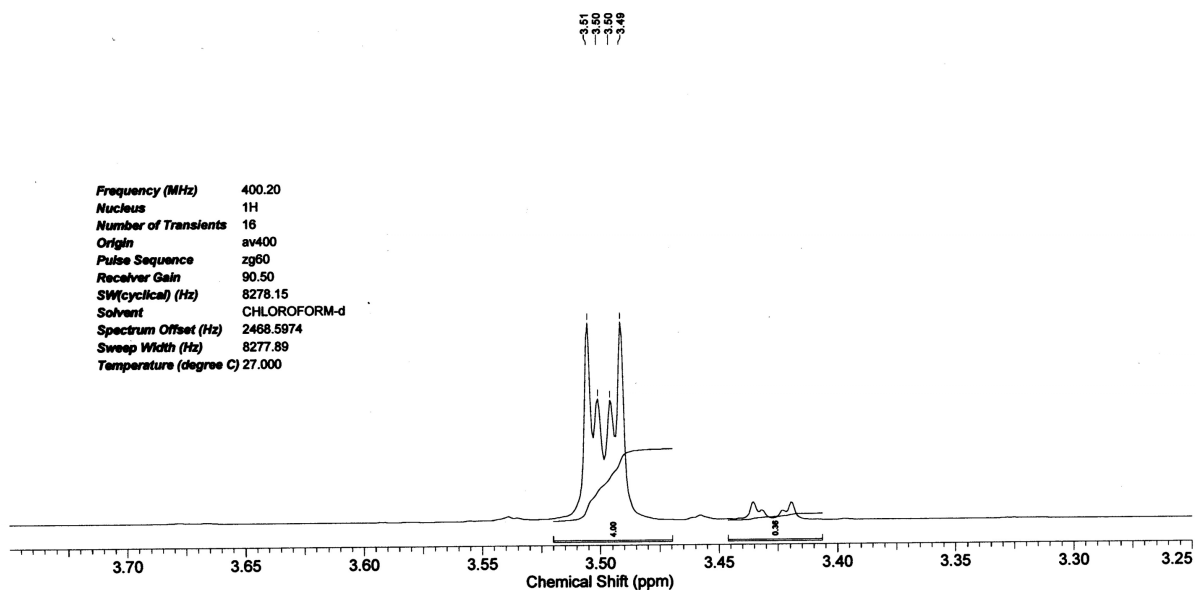


2 hr self-metathesis

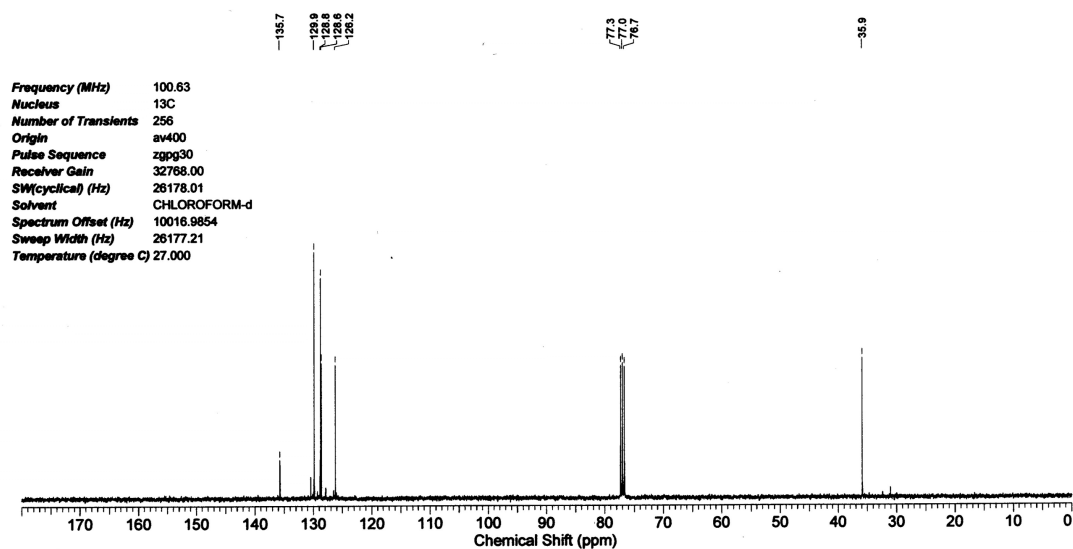


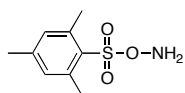


2 hr self-metathesis (zoom on CH₂SPh)

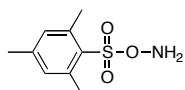
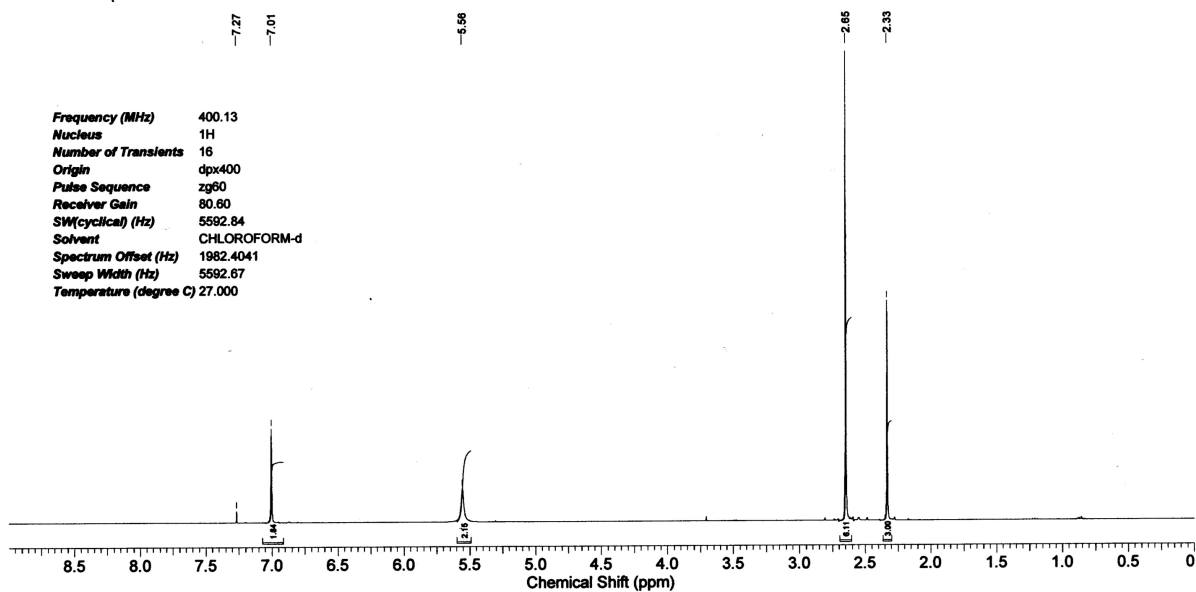


2 hr self-metathesis

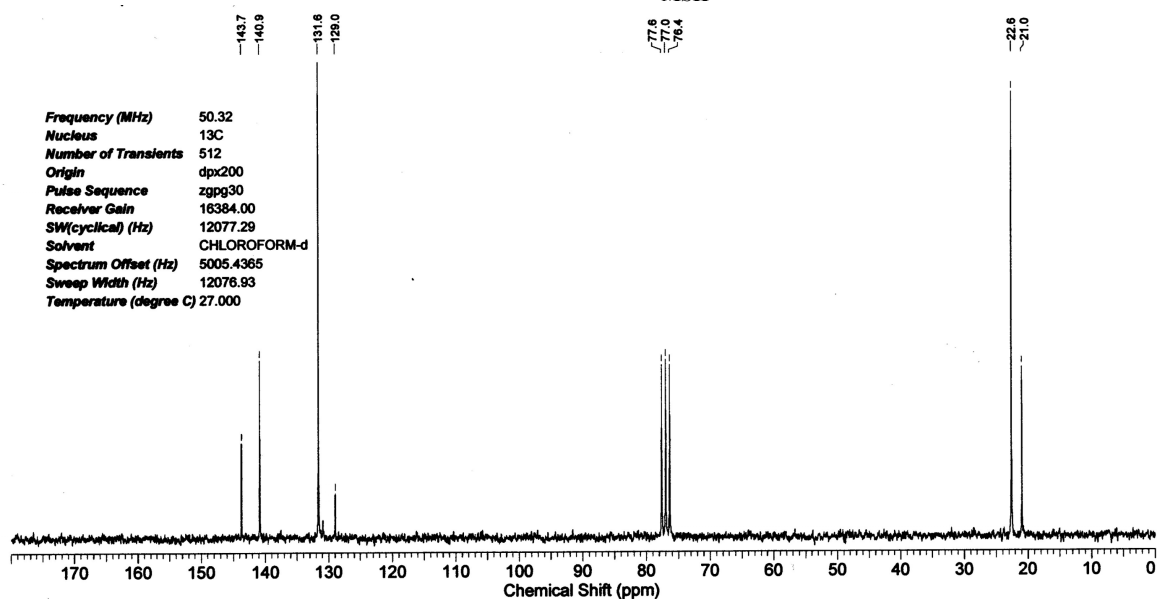


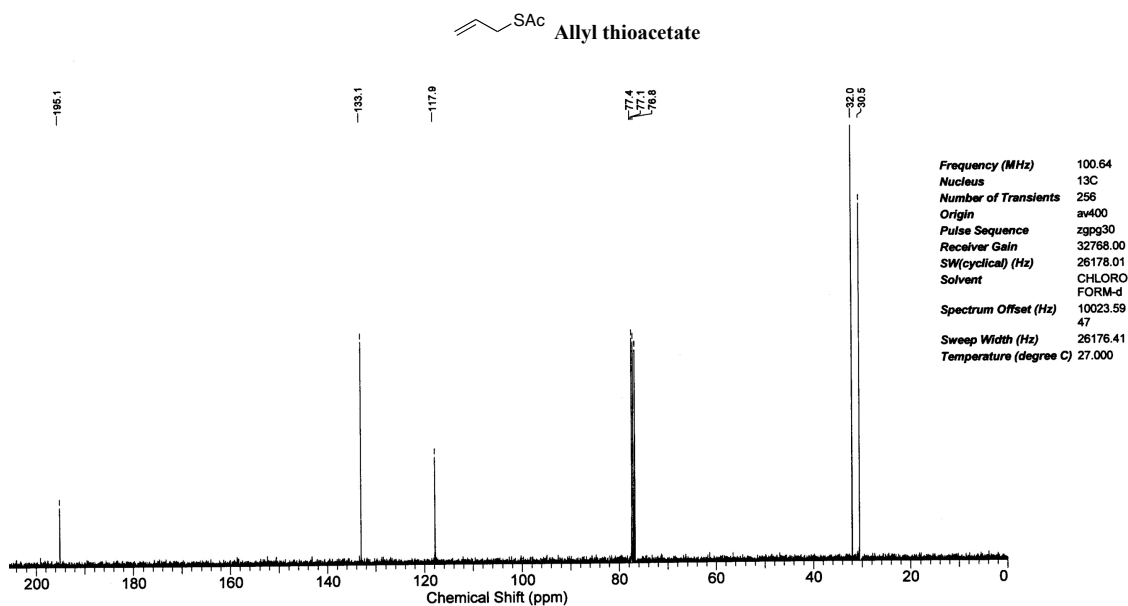
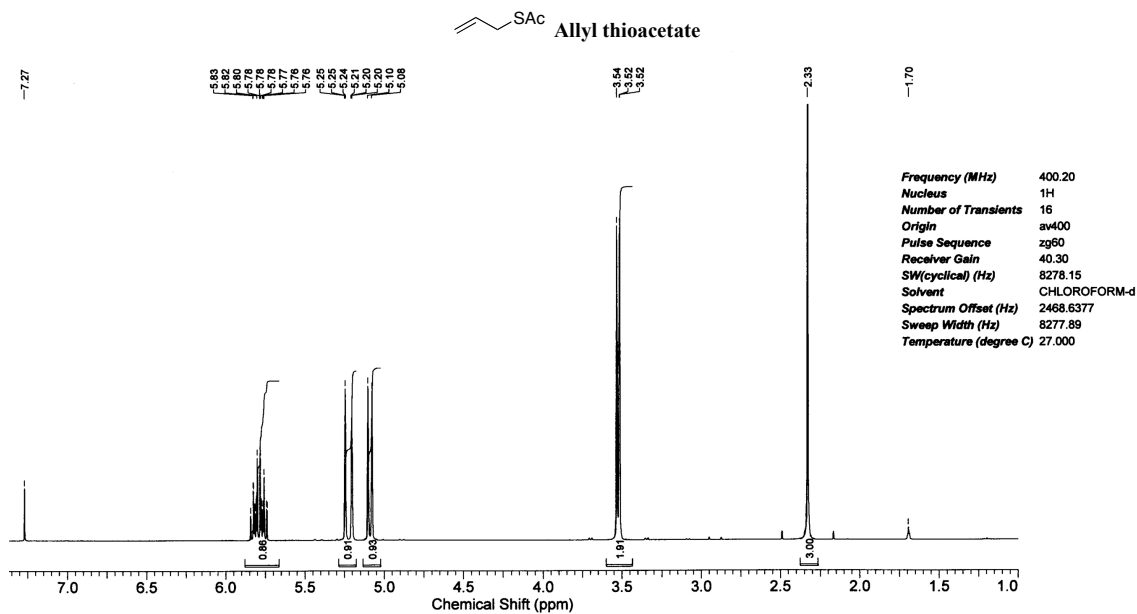


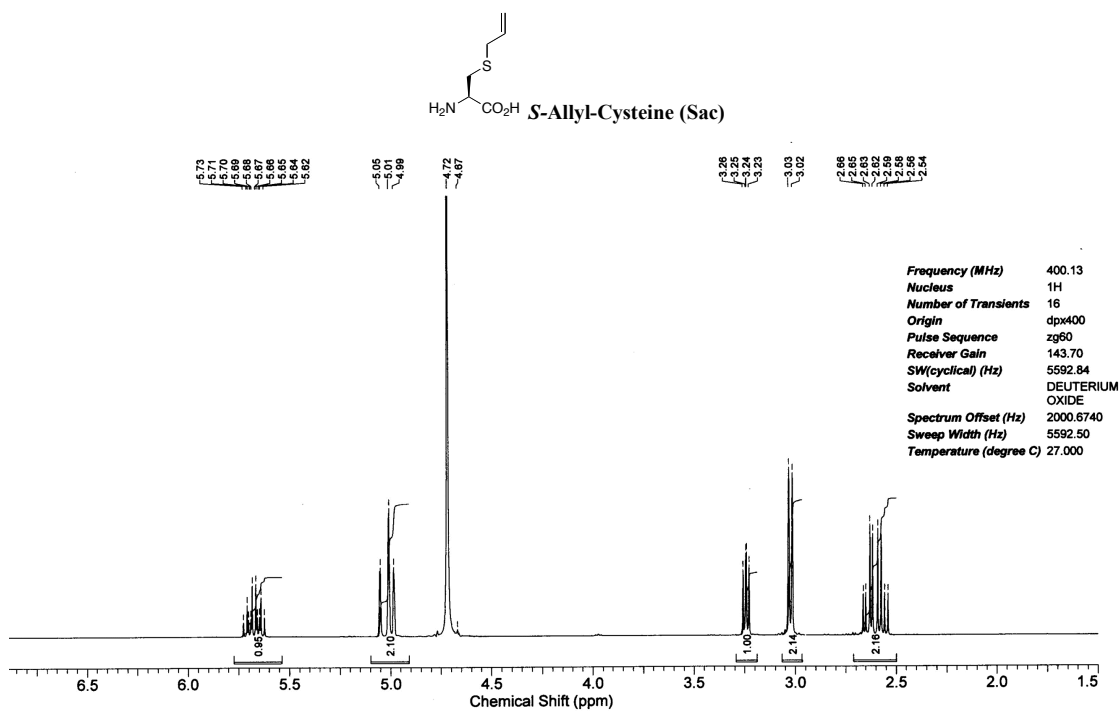
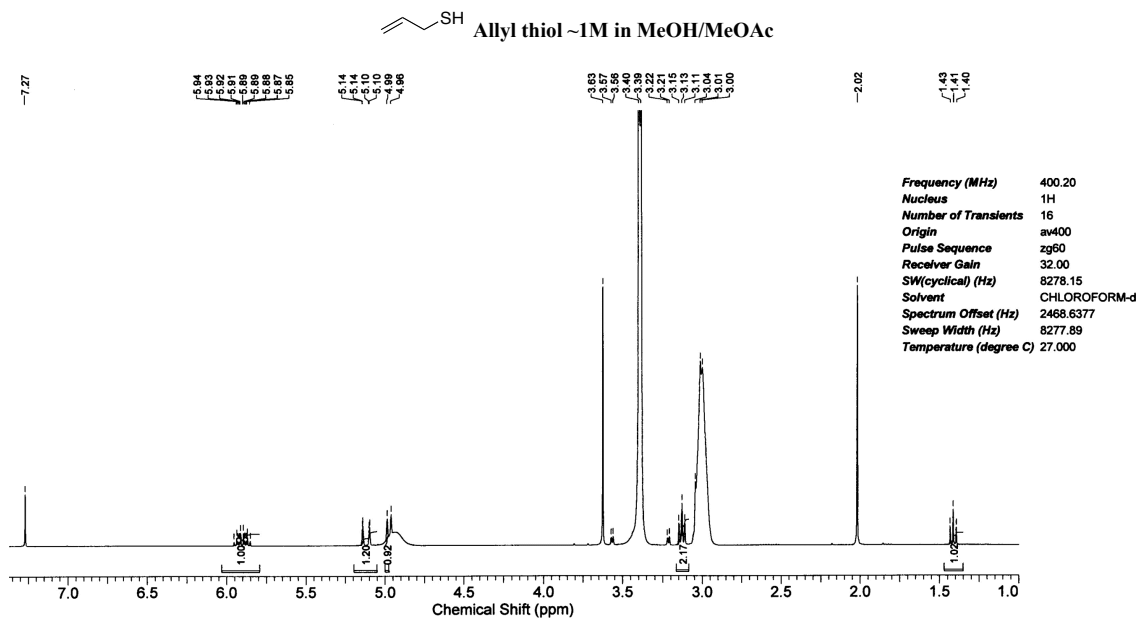
MSH

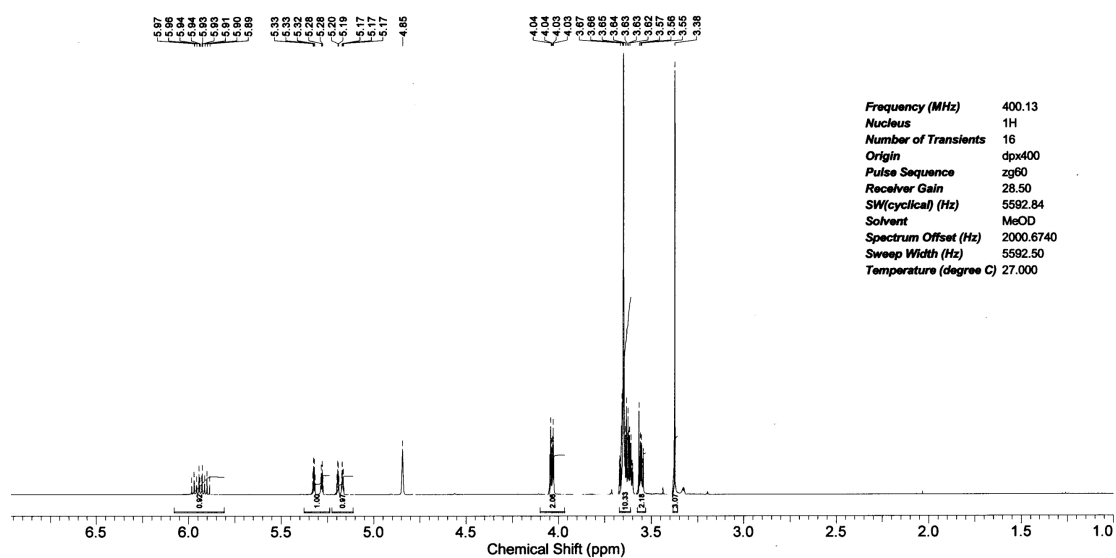
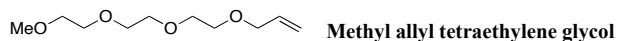
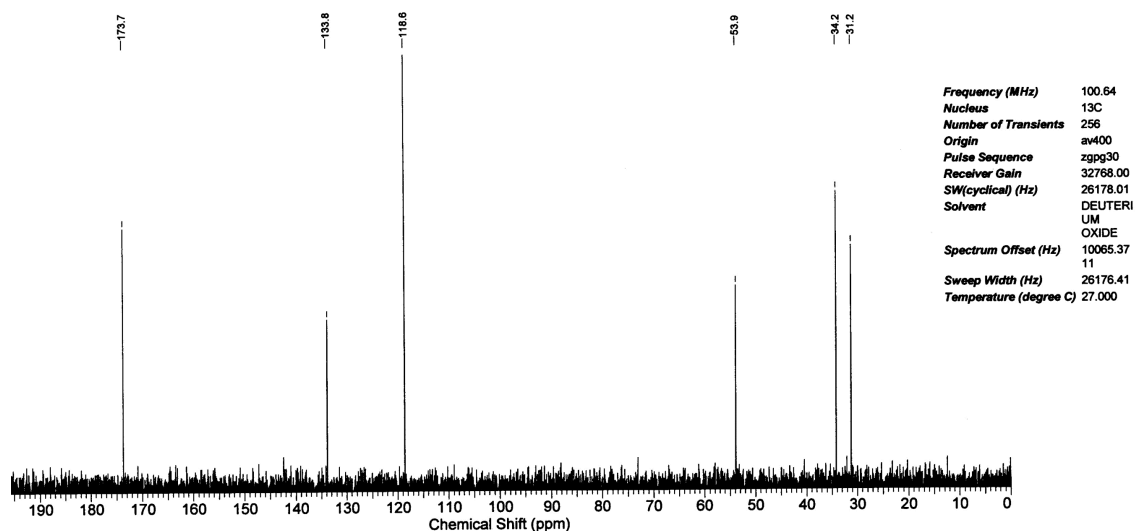
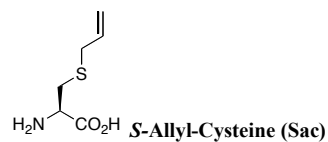


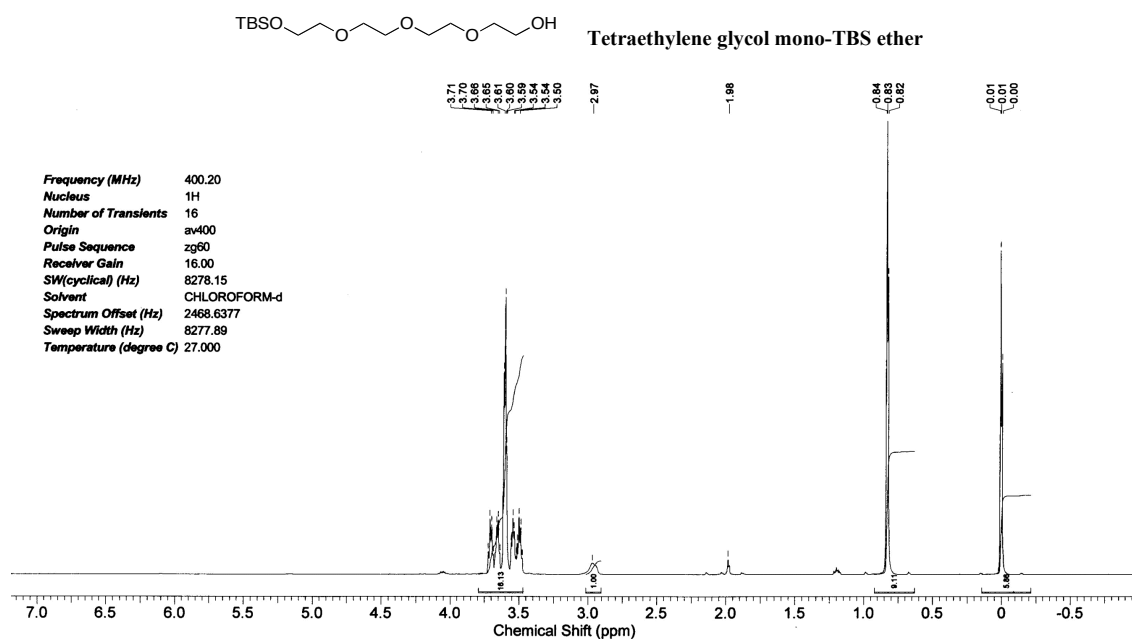
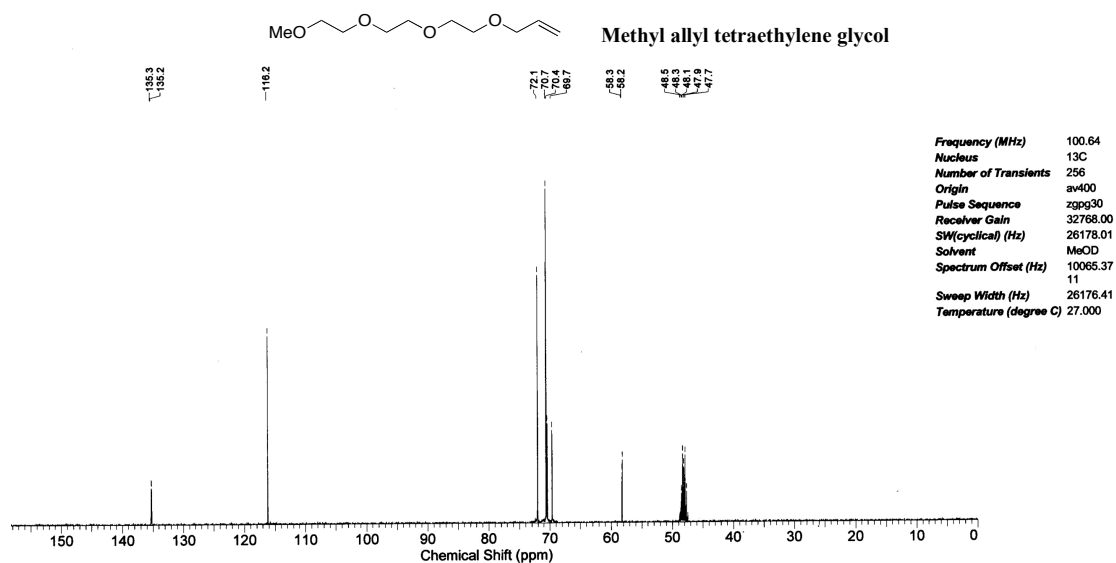
MSH

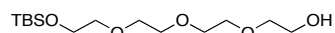




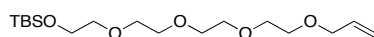
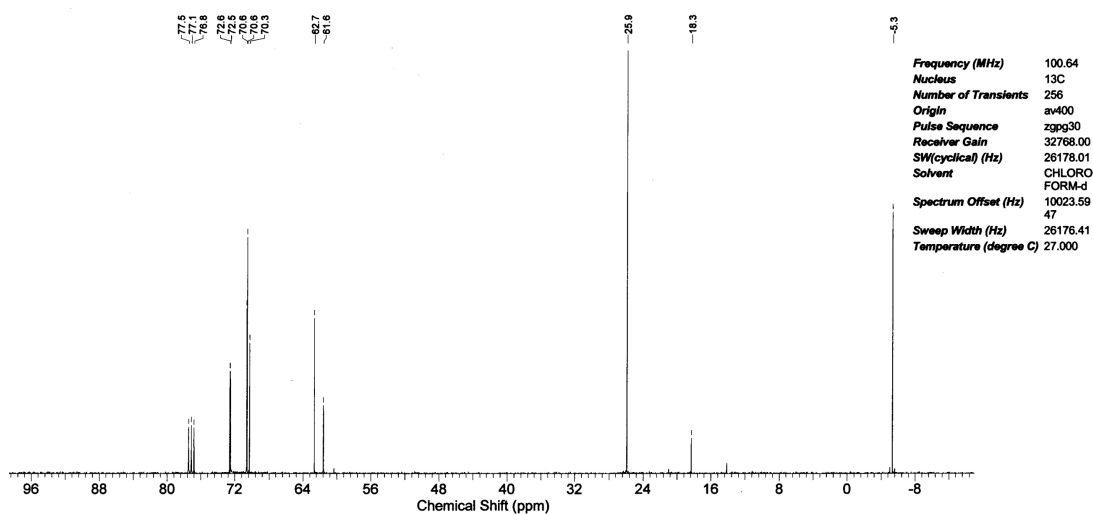




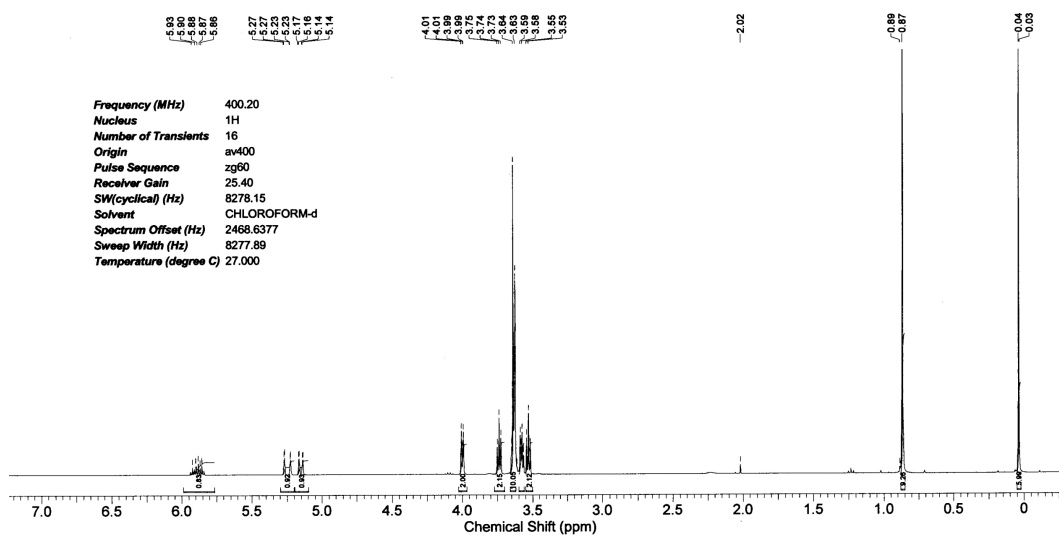


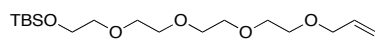


Tetraethylene glycol mono-TBS ether

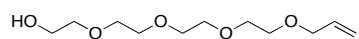
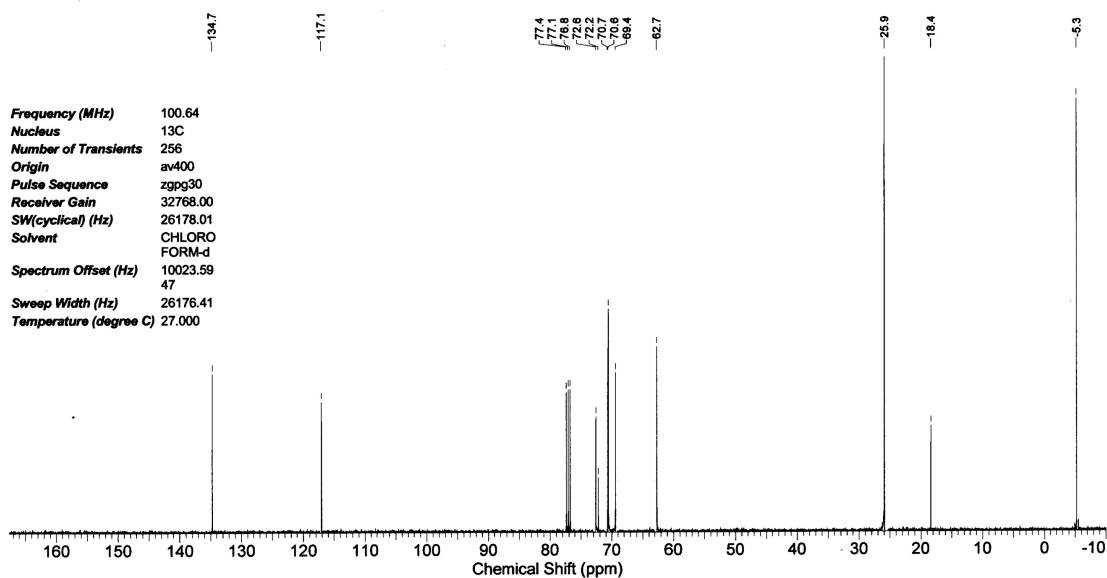


TBS allyl tetraethylene glycol

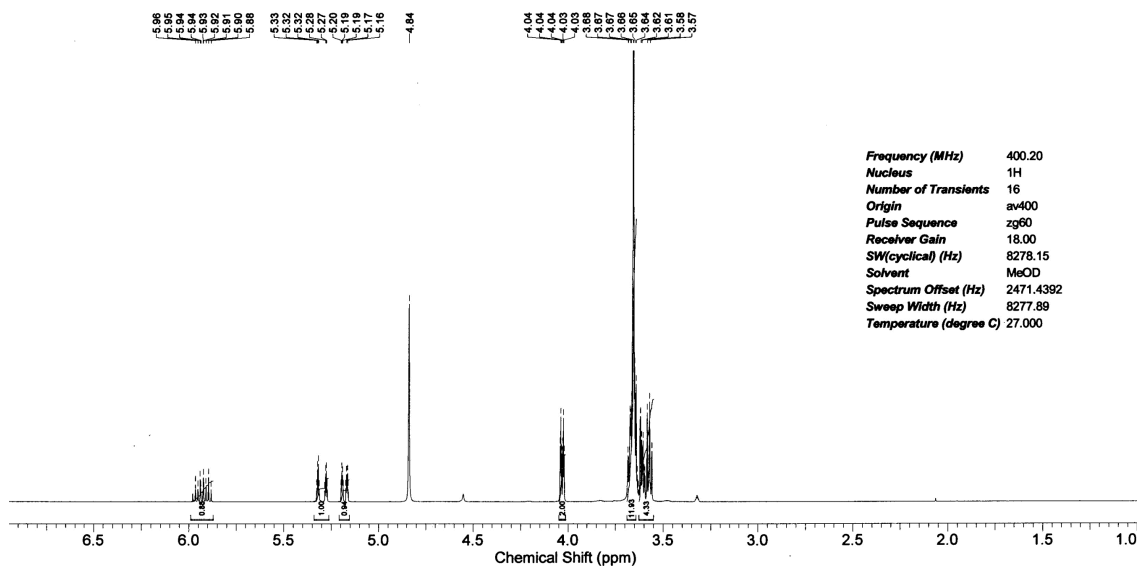


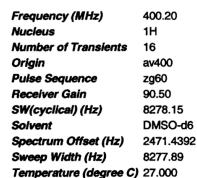
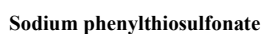
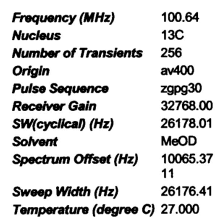
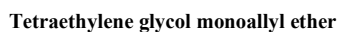


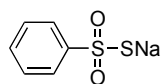
TBS allyl tetraethylene glycol



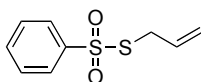
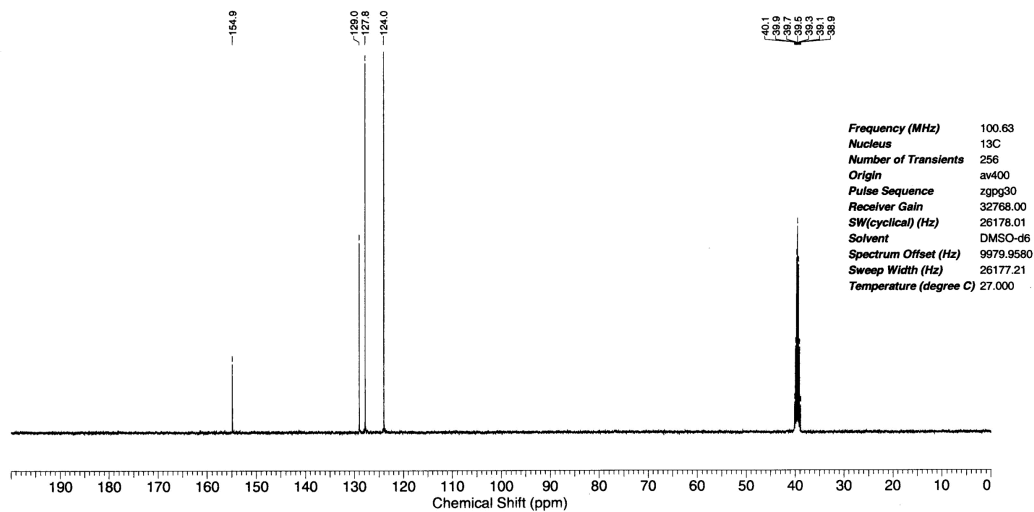
Tetraethylene glycol monoallyl ether



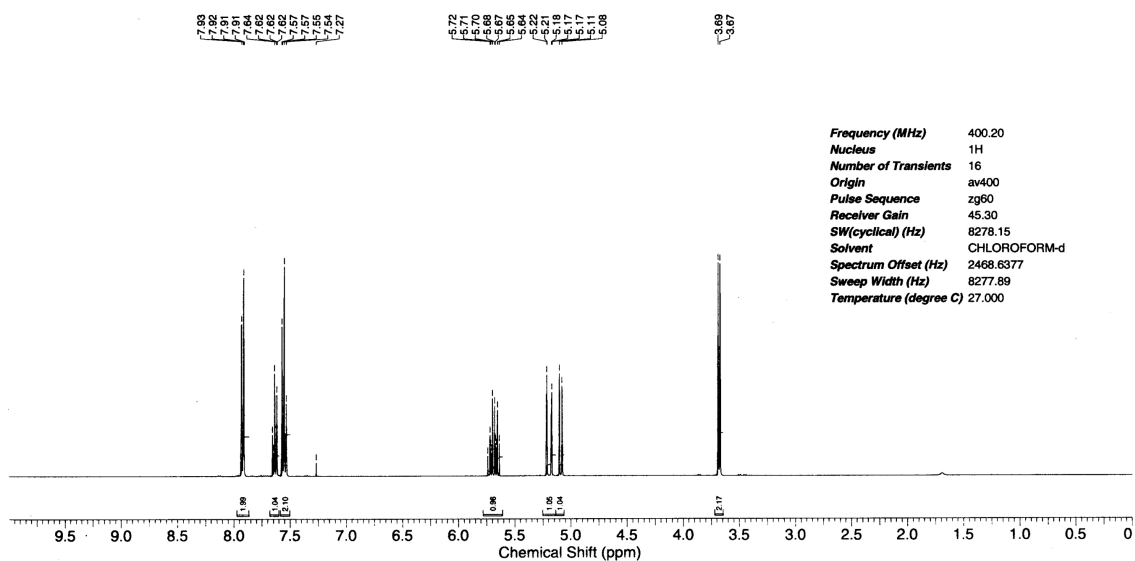


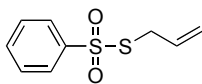


Sodium phenylthiosulfonate

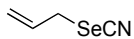
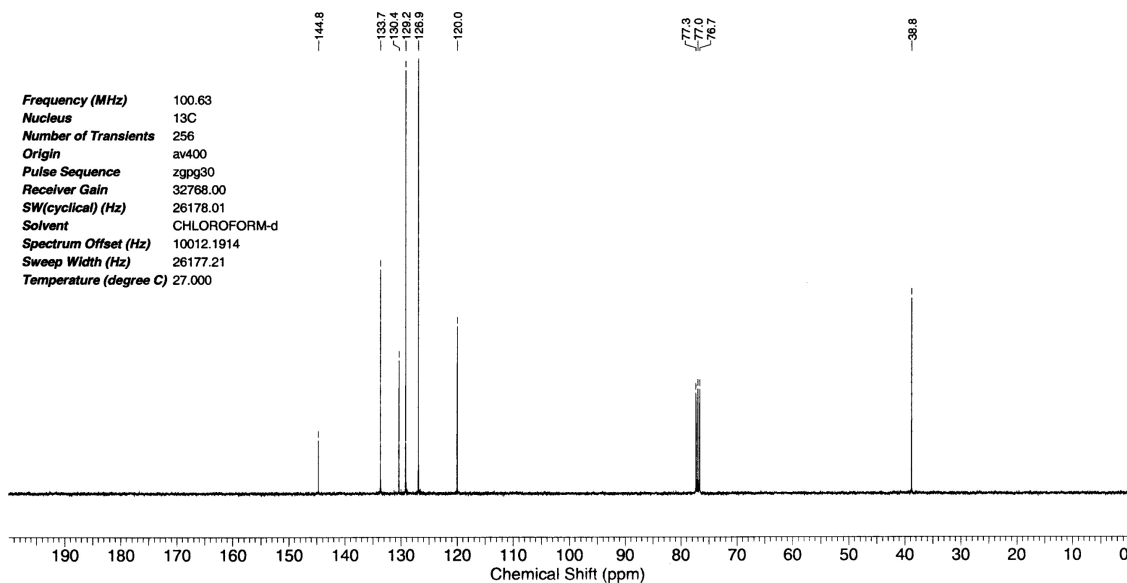


Allyl phenylthiosulfonate

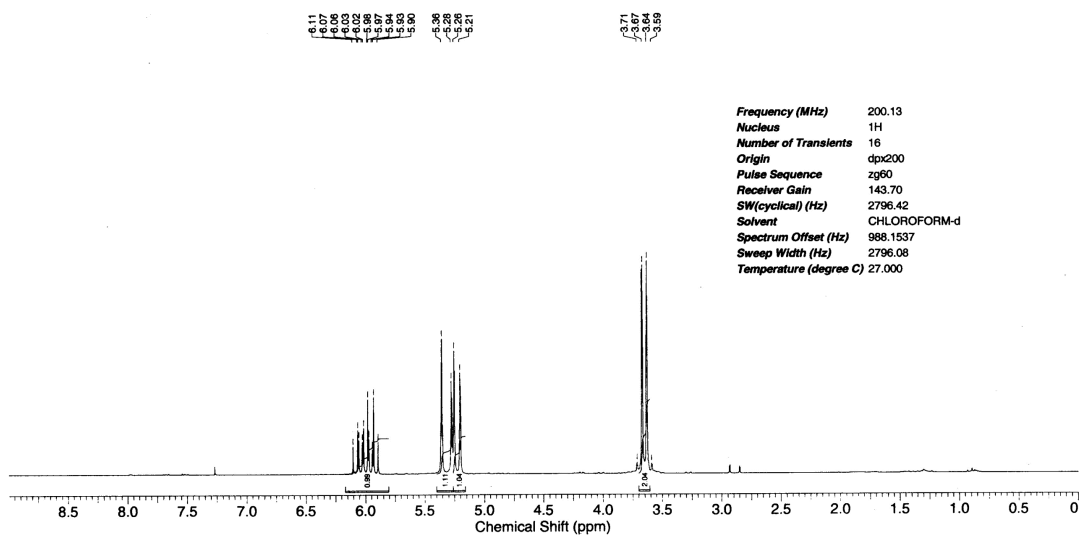


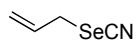


Allyl phenylthiosulfonate

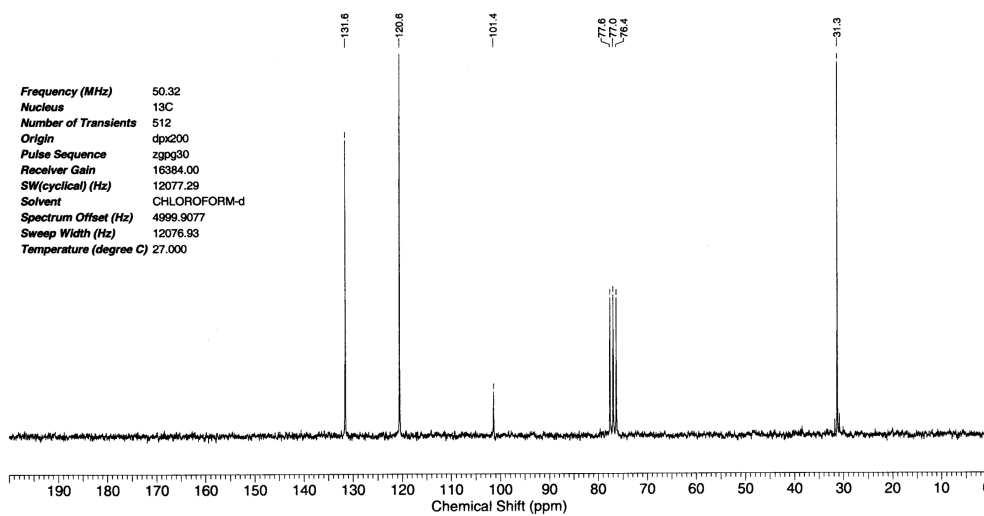


Allyl selenocyanate



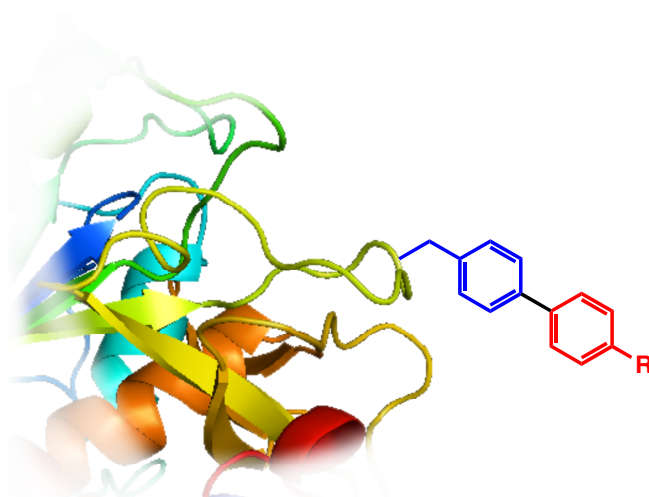


Allyl selenocyanate



Chapter 6

A Convenient Catalyst for Aqueous Suzuki-Miyaura Cross-Coupling and Its Use in Covalent Protein Modification



“Tall oaks from little acorns grow”

From *Lines Written for a School Declamation*

David Everett (1769-1813)

Quoted by Herbert C. Brown, Nobel Lecture in Chemistry, 1979

6.1 Introduction

Cross-coupling reactions are a mainstay in organic synthesis and are among the most widely used methods for carbon-carbon bond construction.¹⁻⁴ Among these reactions, the palladium catalyzed coupling between organoboron derivatives and aryl- or vinyl halides is of particular importance.^{5,6} A typical example and the catalytic cycle of this transformation are depicted in Figure 6.1. Now widely referred to as the Suzuki-Miyaura reaction in recognition of its discoverers, this cross-coupling is ubiquitous in general synthesis due to its high chemo- and stereoselectivity in addition to its broad functional group tolerance.^{7,8} The Suzuki-Miyaura reaction's pervasive use in the synthesis of biaryls for pharmaceuticals and agrochemicals,^{9,10} wide application in polymer and material science,^{11,12} and employment in the most advanced stages of complex total synthesis is a testament to its utility.¹³⁻¹⁶

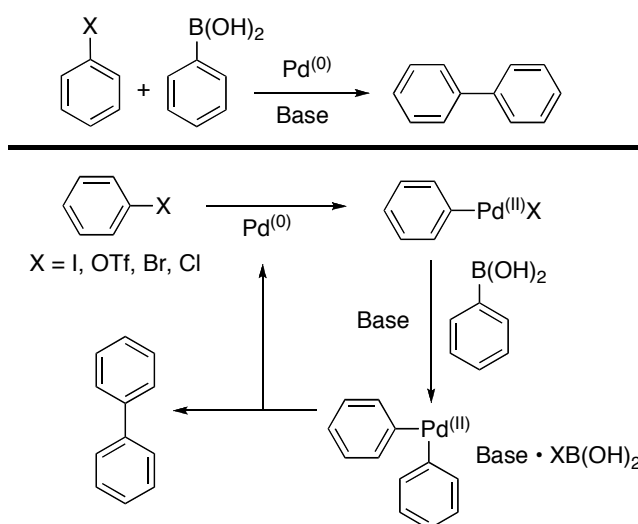


Figure 6.1: Suzuki-Miyaura Cross Coupling⁷

Suzuki-Miyaura coupling in water

The widespread use and reliability of this cross-coupling is made possible by the continual development of high-performance catalysts.¹⁷ Among these developments, the use of water-soluble ligands and other additives allows use of the Suzuki-Miyaura reaction in water (Figure 6.2).¹⁸⁻²² There are several reasons why water is a desirable solvent in synthesis. First, water is a non-flammable, non-toxic solvent and its use may therefore lessen the environmental impact of synthesis in certain instances.^{19,23-25} Second, water can positively influence the rate and selectivity of reactions due to its ability to stabilize polar transition states.²⁶⁻²⁸ Third, the availability of water-soluble catalysts for

Suzuki-Miyaura coupling enables use of this reaction on molecules that are only soluble in water. This final point is of particular relevance to bio-molecules such as carbohydrates, nucleic acid, peptides, and proteins that typically require aqueous reaction media. With our interest in expanding the repertoire of reactions useful in protein modification, we set out to explore the Suzuki-Miyaura reaction for use in peptide and protein modification.

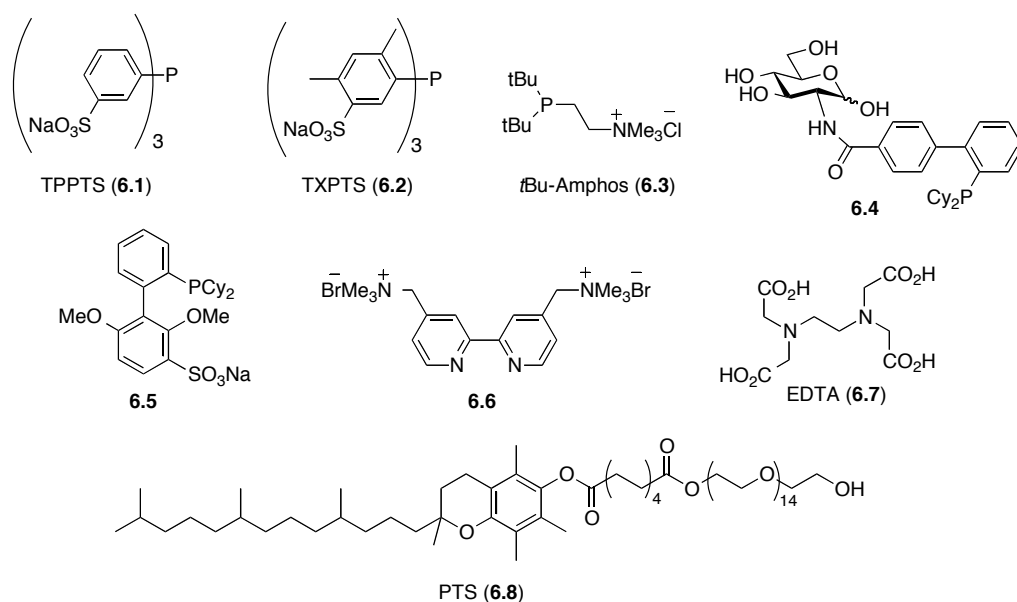


Figure 6.2: Selection of water-soluble ligands and a surfactant (PTS) useful for aqueous cross-coupling²²

Palladium-mediated coupling on peptides and proteins

Given the broad functional group tolerance of the Suzuki-Miyaura reaction and ample precedence for its use in water, it is not surprising that this reaction has previously been investigated on peptides and proteins. The advantages are clear: the modification is installed through a non-labile carbon-carbon bond and the cross-coupling partners are largely inert to biological systems. Additionally, we anticipated that this reaction would be well poised for transferring hydrophobic groups to peptides and proteins without the need for organic solvent since the borate salt can confer water-solubility to the otherwise hydrophobic aryl or alkenyl cross-coupling partner. Yet, despite these positive aspects of the Suzuki-Miyaura reaction, its use in bioconjugation was limited to short peptides at the start of this research.^{29,30} What follows is a brief discussion on the development of the

Suzuki-Miyaura reaction for peptide and protein modification to set the stage for our own efforts in this area.

In 2005, Hamachi reported the first Suzuki-Miyaura reaction on a synthetic peptide.²⁹ In this report, the WW domain of the Pin1 protein was synthesized by solid phase peptide synthesis with *p*-iodophenylalanine installed at various positions in the 34-residue peptide. The WW domain of the Pin1 protein is involved in a protein-protein interaction during a phosphorylation cascade.^{31,32} The Suzuki-Miyaura reaction was used to install a coumarin-derived fluorophore for the purpose of sensing phosphorylated peptides that bind to the Pin1 WW domain (Figure 6.3). Upon optimization, efficient reaction (up to 90% conversion by HPLC) was found when the coupling was run in 50% aqueous glycerol (v/v) using the water-soluble catalyst Na₂PdCl₄. The fluorescently labeled peptides were then shown to decrease in fluorescence upon binding to proline-rich phosphorylated peptides, validating the use of the modified peptide as a fluorescent sensor. This report was a valuable addition to the palladium-mediated modifications of peptides previously reported by Schmidtchen³³ and Ghadiri³⁴ in 1998 and 2001, respectively, both of whom described Sonogashira couplings³⁵ on short peptides.

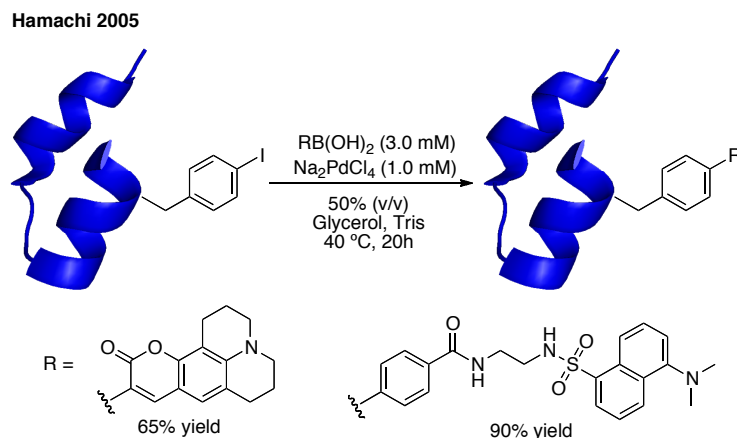


Figure 6.3: Suzuki-Miyaura coupling on the synthetic WW domain of Pin1²⁹

A second report of Suzuki-Miyaura coupling on peptide substrates was recently reported by Barluenga and co-workers.³⁰ In this report, the cross-coupling partners were *p*-iodophenylalanine and 3,5-diiodotyrosine. In the latter case, the residue was obtained by selective di-iodination of tyrosine using Barluenga's reagent (IPy₂BF₄ **6.9**) (Figure 6.4)^{36,37} It was shown that the iodination-cross-coupling could be carried out in a one-pot operation. This strategy was used for the selective iodination and di-arylation of Leu-

enkephalin peptide, demonstrating the capacity for selective iodination and cross-coupling on bio-active peptides. Importantly, both steps were fully compatible with water and good yields were obtained. Furthermore, this method can be applied to *native* peptides and proteins since the iodination can be carried out after peptide synthesis or protein translation. Indeed the selective iodination of tyrosine on proteins using Barluenga's Reagent has been described³⁸ and while the incorporation of iodine serves other purposes, such as assisting x-ray structure determination,³⁹ its selective introduction bodes well for use in modification by cross-coupling.

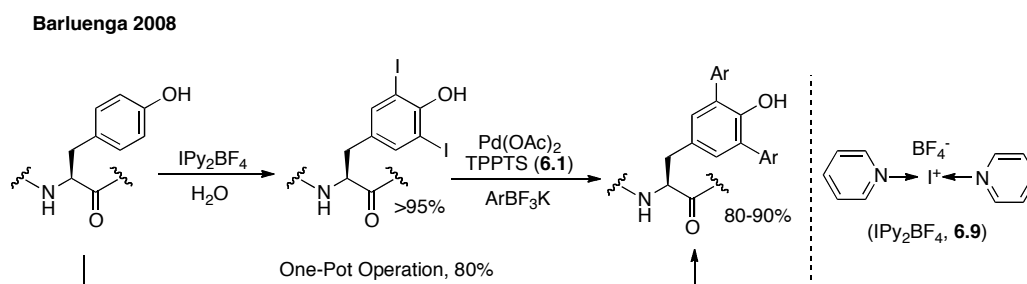


Figure 6.4: Selective diiodination of tyrosine with IPy₂BF₄ and subsequent Suzuki-Miyaura coupling.³⁰

Despite the precedence set by Hamachi and Barluenga for Suzuki-Miyaura coupling on peptide substrates, cross-coupling on protein substrates has been met with little success. There are several instructive points that can be gleaned from published attempts of cross-coupling on protein substrates. These reports are discussed next.

It should be pointed out that even before Hamachi's report of the first Suzuki-Miyaura coupling on a synthetic peptide, the idea of using this reaction on protein substrates had been proposed in the literature by Tirrell.^{40,41} Such proposals were based on the emerging technologies that allowed the incorporation—genetically or chemically—of unnatural amino acids into protein backbones that were suitable for cross-coupling reactions.³⁸⁻⁵² A selection of these amino acids is shown in Figure 6.5.

In general, four strategies have been used for incorporating cross-coupling partners into proteins. The first is to incorporate the unnatural amino acid as a surrogate for a natural amino acid by expressing the protein in auxotrophic cells using media depleted of the natural amino acid.⁴⁰⁻⁴⁴ The second strategy is to incorporate the unnatural amino acid in response to the amber stop codon using a mutant suppressor tRNA and aminoacyl-tRNA synthetase pair.^{39,45-47,51,52} The third option is to chemically or

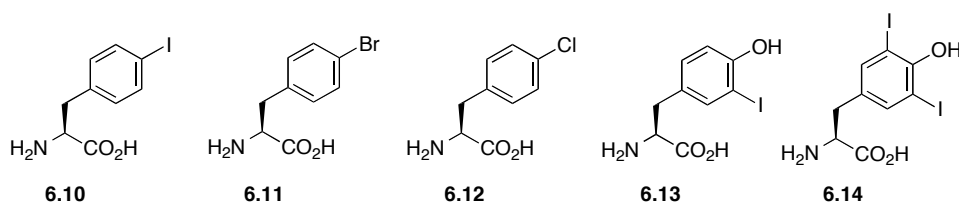


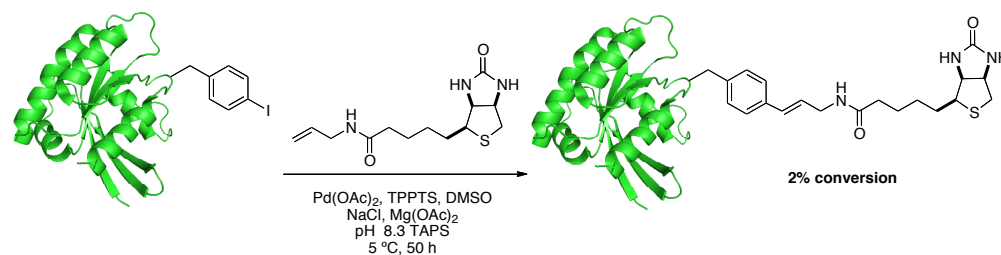
Figure 6.5: Examples of cross-coupling partners that can be incorporated into proteins

enzymatically misacylate tRNA, effectively rewriting the genetic code.^{50,53} The fourth strategy is to introduce the cross-coupling partner through some selective chemical transformation, for instance by selective iodination of tyrosine using IPy_2BF_4 to give **6.14**.³⁸ Figure 6.5 depicts examples of the cross-coupling partners that can be incorporated by these methods and the reader is referred to several authoritative reviews devoted to these techniques and their applicability to various host species and cell-free translation systems.⁵⁴⁻⁵⁸

Despite the ability to incorporate cross-coupling partners into proteins and ample precedence for coupling on short peptides, there are no reports of high-yielding Suzuki-Miyaura coupling on full-length proteins. In fact, all reported cross-coupling reactions on proteins were met with low conversions. The first example of palladium mediated carbon-carbon bond formation on protein substrates was described by Kodama and co-workers who reported first a Heck modification of *p*-iodophenylalanine⁴⁸ and later a Sonogashira reaction at the same residue.⁴⁹ In both cases, the Ras GTP binding protein was used as a model substrate. *p*-Iodophenylalanine was incorporated at position 32 using a cell-free expression system.⁵⁹ A variety of conditions and additives were explored for the conjugation of biotin but conversions did not exceed 2% for the Heck route and 25% for the Sonogashira method (Figure 6.6). It was found that ascorbic acid, DMSO, and MgCl_2 were critical for successful cross-coupling. It was suggested that ascorbic acid maintained a reducing environment that kept palladium and copper at the $\text{Pd}^{(0)}$ and $\text{Cu}^{(I)}$ oxidation states, respectively. The roles of DMSO and MgCl_2 were not clear but they appeared to minimize protein precipitation under the reaction conditions. While the overall conversions for these reactions are low, it should be noted that these are two of the earliest examples of carbon-carbon bond forming reactions on proteins using bio-orthogonal chemistry.

In a recent report, the Schultz group described the biosynthetic incorporation of *p*-boronophenylalanine in response to the amber stop codon.⁵² The primary use of this amino acid was to allow traceless affinity purification: the boronic acid binds to

Kodama *et al* 2006



Kodama *et al* 2007

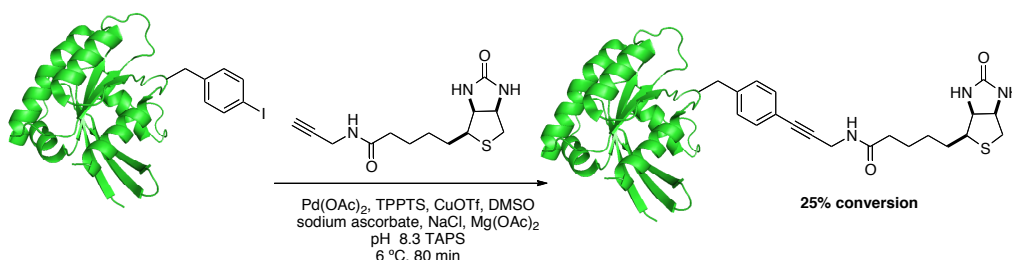


Figure 6.6: Modification of *p*-iodophenylalanine in the Ras protein by the Heck reaction (top) and Sonogashira coupling (bottom)^{48,49}

polyhydroxylated resins and then can be eluted with an oxidant (e.g. hydrogen peroxide) or reductant (silver diammonium nitrate) to yield tyrosine or phenylalanine, respectively. The authors also recognized the opportunity for the use of this residue in Suzuki-Miyaura cross-coupling. However, even after multiple attempts, the only sign of reaction was observed after prolonged incubation at high temperature (12 h at 70 °C) (Figure 6.7). Even under these harsh conditions, the conversion was estimated to be only 30%. The difficulty of this coupling prompted the authors to call for the development of new catalysts so that Suzuki-Miyaura cross-coupling may be used as a general strategy for protein modification.

Schultz 2008

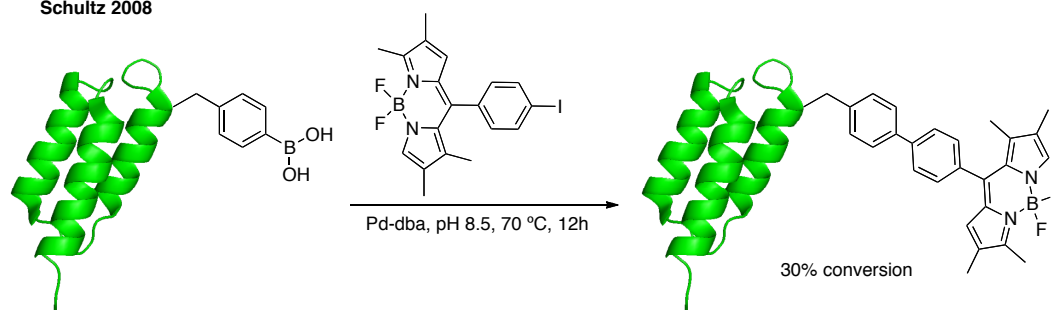


Figure 6.7: Fluorescent labeling of Z-domain of staphylococcal protein A by Suzuki-Miyaura cross-coupling⁵²

Given the potential benefits of the Suzuki-Miyaura reaction to protein modification—carbon-carbon bond formation and high selectivity—and the ready availability of cross-coupling partners in protein substrates, there was a need for an easy to use catalyst that was highly active and could mediate the cross-coupling on protein substrates. As with other methods for modifying proteins, this catalyst would have to work efficiently at room temperature in buffered water and tolerate the various functionalities of natural amino acid side chains. Next, an account of the discovery of such a catalyst and its use in both small-molecule and protein chemistry is presented.

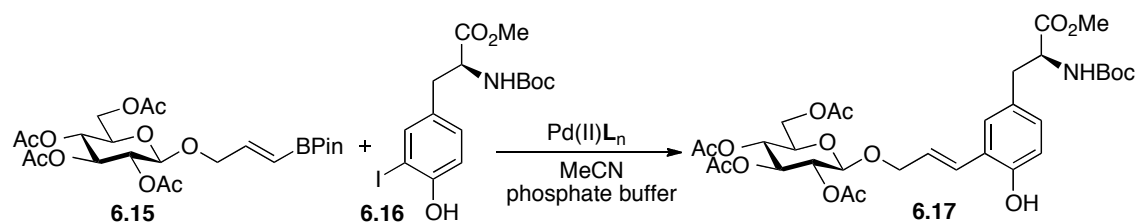
6.2 Results and Discussion

The discovery of a pyrimidine ligand for Suzuki-Miyaura coupling in water

The immediate goal en route to enabling Suzuki-Miyaura cross-coupling on protein substrates was to identify a suitable catalyst. Candidate catalysts were restricted to commercially available ligands and palladium salts. Additionally, a Pd^(II) pre-catalyst was desirable for optimization since stock solutions could be conveniently prepared and stored without the concern of oxidation that would accompany a Pd⁽⁰⁾ source. This same concern for oxidation also prompted a search for phosphine-free ligands. It was anticipated that this restrictive search would focus attention on catalysts that could mediate the desired transformation without the need for tedious degassing and an inert atmosphere during the reaction.

Several phosphine-free palladium catalysts were screened for the model coupling of carbohydrate-derived boronic ester **6.15** and 3-iodotyrosine derivative **6.16** (Table 6.1). The former substrate was selected because of the importance of glycoproteins and their synthesis,⁶⁰⁻⁶⁴ and the latter since 3-iodotyrosine can be incorporated into proteins using genetic techniques.^{45,46} For these exploratory experiments, protected substrates were used for ease of handling and the reactions were run in aqueous acetonitrile. At the outset, the temperature was restricted to 37 °C since that is generally a suitable temperature for protein substrates.

Attempts with EDTA^{65,66} and NaCl²⁹ as ligands gave low yields of cross-coupling product **6.17** (Entries 1-4), even with stoichiometric palladium (Entry 4). Palladium black was observed invariably during these attempts. These results were disappointing given



Entry	Pd Salt	Pd Loading	Ligand (L)	L : Pd	Conditions	Yield (%)
1 ^a	PdCl_2	10 to 40%	EDTA	1:1	RT, 12 h	0
2 ^a	PdCl_2	40%	EDTA	1:1	37 °C, 12 h	25
3 ^a	PdCl_2	10%	NaCl	2:1	RT, 12 h	0
4 ^a	PdCl_2	110%	NaCl	2:1	RT, 12 h	19
5 ^b	Pd(OAc)_2 or PdCl_2	N/A		1:1	N/A	N/A
6	Pd(OAc)_2	5%		2:1	37 °C, 8 h	0
7	Pd(OAc)_2	20%		2:1	37 °C, 8 h	74
8	Pd(OAc)_2	2%		2:1	37 °C, 33 h	82

^a Palladium black observed

^b Palladium/ligand not soluble in water

Table 6.1: Catalyst screening for model Suzuki-Miyaura cross-coupling

the success of these catalytic systems on other substrates, including peptides.²⁹ Still intent on finding a phosphine-free, readily available ligand, other candidates were screened. The bipy-dicarboxylate in Entry 5 was not sufficiently soluble in water and while palladium acetate readily dissolved in water in the presence of the thiourea in Entry 6, this system did not promote cross-coupling. Attention then turned to 2-amino-4,6-dihydroxypyrimidine as a ligand (Entries 7-8). This material had been reported as a ligand in Sonogashira cross-couplings³⁵ in organic solvents.⁶⁷ It occurred to us that the hydroxyl groups should be sufficiently acidic to render this ligand soluble in alkaline water. Indeed, 2 equivalents of sodium hydroxide generated a salt that was completely

water-soluble. Moreover, palladium acetate was freely soluble in water in the presence of the sodium salt of this ligand. Gratifyingly, this catalytic system promoted the model Suzuki-Miyaura cross-coupling. A high loading of palladium was used at first to push the reaction to completion (Entry 7), but comparable yields were obtained with 2% palladium and longer reaction time (Entry 8).

Suzuki-Miyaura coupling on amino acid and peptide substrates

Following the lead from Table 6.1, the Pd-pyrimidine catalyst was explored in more detail on other amino acid and peptide substrates. Indeed, efficient cross-coupling was observed at 37 °C with 1 mol% palladium in buffered water on several amino acids (Table 6.2). No organic solvent was used in these reactions and no effort was made to exclude oxygen. Formate was added to facilitate reduction of Pd^(II), but control reactions later showed this additive was not necessary for these substrates and the boronic acid can apparently serve as the reducing agent. The amino acid substrates were selected because they can be incorporated into proteins genetically^{39-47,50,51,53} (Entries 1-4) or chemically^{38,68} (Entries 5-7), as described in the introduction to this Chapter. 4-Chlorophenylalanine was inert to these coupling conditions (Entry 3), suggesting this catalytic system is best suited to more reactive aryl iodides and aryl bromides. Entry 5 demonstrated that double modification is possible with this catalyst and that tyrosine is a potential target for sequential iodination with IPy₂BF₄ and cross-coupling.^{30,38} Entries 6 and 7 indicate that a C-terminal cysteine or related derivatives may inhibit this catalytic system, perhaps by detrimental chelation of the palladium catalyst, but internal alkyl-cysteines are tolerated and high yields were observed with the glutathione-derived peptide in Entry 7.

$\text{ArX} + \text{PhB(OH)}_2 \xrightarrow[\text{H}_2\text{O, K}_2\text{HPO}_4, \text{HCO}_2\text{H}]{[\text{Pd}], 37^\circ\text{C, open air}} \text{Ar-Ph}$ $[\text{Pd}] = \left[\begin{array}{c} \text{NaO} \\ \\ \text{C}_5\text{H}_3\text{N}_2 \\ \\ \text{NaO} \end{array} \right]_2 \cdot \text{Pd(OAc)}_2$					
Entry	ArX	Pd Loading	Conditions	Coupling Product	Yield (%)
1		1%	37 °C, 4 h		95
2		1%	37 °C, 4 h		98
3		1%	37 °C, 4 h		0
4		1%	37 °C, 4 h		94
5		2%	37 °C, 4 h		95
6		1%–10%	37 °C, 6 h		0
7		4%	37 °C, 6 h		92

Table 6.2: Model cross-couplings on amino acid and peptide substrates using the palladium-pyrimidine catalyst.

Suzuki-Miyaura coupling on protein substrates

The promising results in Table 6.2 prompted validation of this catalytic system on a protein surface. A model protein **6.18** suitable for cross-coupling was obtained by alkylating the single cysteine mutant of subtilisin from *Bacillus lentus* (SBL S156C).⁶⁹ While the ultimate goal was to adapt the cross-coupling to genetically incorporated cross-coupling partners, **6.18** could be used to simply validate Suzuki-Miyaura cross-coupling as a viable method for protein modification. SBL is a serine protease and boronic acids are inhibitors of this protein, forming borates at the serine of the active site.⁷⁰ The protein

was therefore irreversibly inhibited by sulfonylation at S221 with phenylmethylsulfonylfluoride (PMSF) to obviate any potential complications in the coupling and to prevent any influence of the protease activity on the coupling mechanism. Dithiothreitol (DTT) was used to pre-reduce and contaminant disulfide before alkylation. All transformations were monitored by LC-MS. **6.18** and **6.19** did not react with Ellman's reagent, indicating that all cysteine had been consumed in the alkylation. Benzylated protein **6.19** was synthesized as a negative control (*vide infra*) and monitored and characterized in a similar fashion (Figure 6.8).

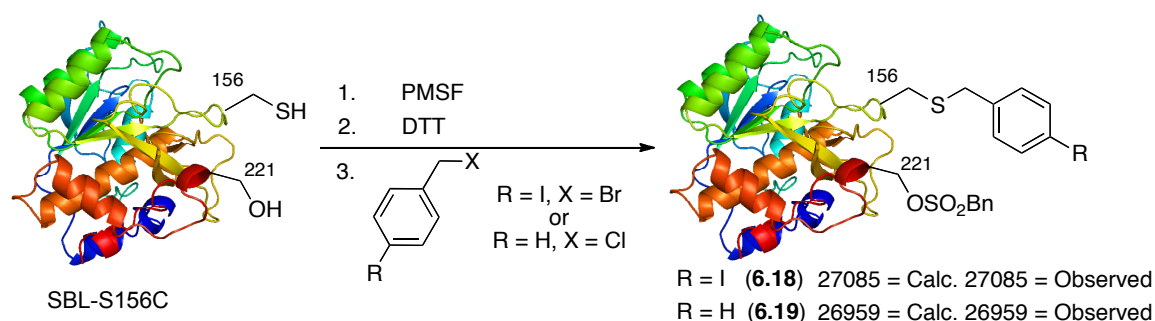


Figure 6.8: Synthesis of protein for cross-coupling (**6.18**) and an un-iodinated substrate (**6.19**) as a control.

As a first attempt, phenylboronic acid was used as the cross-coupling partner, and an excess of palladium (50 equiv.) and boronic acid (500 equiv.) were employed. The reaction buffer was pH 8.0 sodium phosphate (Table 6.3, Entry 1). Gratifyingly, the coupling proceeded to full conversion within 30 minutes at 37 °C, as monitored by LC-MS. While the palladium and boronic acid were in excess, these loadings are still lower than the thousands of equivalents often employed in more traditional protein modification.⁷¹ Several controls were carried out to corroborate the cross-coupling. The palladium, phenylboronic acid, and aryl iodide are all necessary for reaction to occur. Only unreacted protein was observed when **6.18** was treated with palladium or phenylboronic acid alone under otherwise identical conditions. If un-iodinated protein **6.19** was subjected to the coupling conditions, no reaction as observed and only unmodified **6.19** was observed by LC-MS (see experimental procedures for full details).

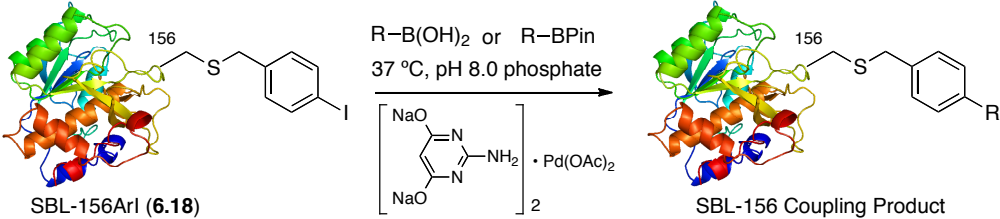
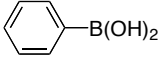
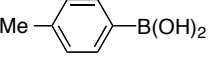
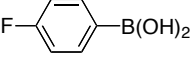
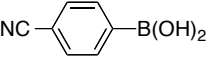
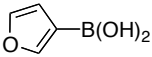
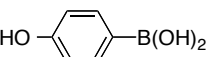
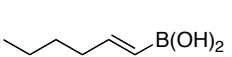
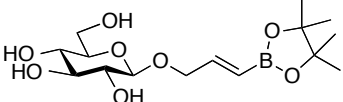
					
Entry	R-B(OH) ₂ / R-BPin	Time	Mass Calculated	Mass Observed	Conversion
1		30 min	27035	27035	> 95%
2		30 min	27049	27049	> 95%
3		30 min	27053	27053	> 95%
4		30 min	27060	27060	> 95%
5		30 min	27113	27113	> 95%
6		30 min	27025	27025	> 95%
7		30 min	27051	27052	> 95%
8		30 min	27041	27041	> 95%
9		60 min	27177	27176	> 95%

Table 6.3: Suzuki-Miyaura cross-coupling on subtilisin derivative **6.18**

A variety of aryl- and vinylboronic acids could be coupled to protein **6.18** (Table 6.3). It is worth noting that no organic solvent was used in these reactions. The borate confers water solubility to otherwise hydrophobic substrates and allows efficient conjugation of a variety of non-polar aryl groups and lipids (Entries 1-8). The synthesis of biaryls on protein surfaces opens the potential for fluorogenic labeling—a method that would complement other fluorogenic labeling methods.⁷² The importance of lipidated proteins in protein-protein interaction and membrane localization has been documented⁷³ and the Suzuki-Miyaura cross-coupling allows the site specific installation of these hydrophobic domains. Finally, a glycoprotein was synthesized by cross-coupling (Entry 9). Synthetic access to well-defined glycoproteins is an ongoing effort due to their ubiquity in nature and the many roles they serve in cell-cell communication, signaling, and disease.⁶⁰⁻⁶⁴

Scope and limitations of palladium-pyrimidine catalyst in synthesis

While the primary goal of validating Suzuki-Miyaura cross-coupling on protein surfaces was met, it was also of interest to determine the scope and limitations of the palladium-pyrimidine catalyst in general synthesis. A variety of aryl halides were screened to this end and the results are compiled in Table 6.4. Efficient coupling was observed with a variety of aryl iodides and aryl bromides. Activity with aryl chlorides was limited (Entry 6). Microwave irradiation was used to induce rapid coupling, though conventional heating sufficed in most cases. Palladium acetate alone gave inferior yields, indicating the pyrimidine ligand is necessary for a competent catalytic system. For instance, in a control reaction with 1 mol% Pd(OAc)₂ and no ligand, the reaction in Entry 1 proceeded in only 20% yield under otherwise identical conditions (See Experimental Procedures for full details).

The catalyst was tolerant of a variety of functional groups, as demonstrated in the synthesis of phenols, anilines, thiophenes, and pyridines (Entries 8-11). Thiols inhibited catalysis, however, and when glutathione is present in equal amounts to palladium, no cross-coupling was observed (Figure 6.9).

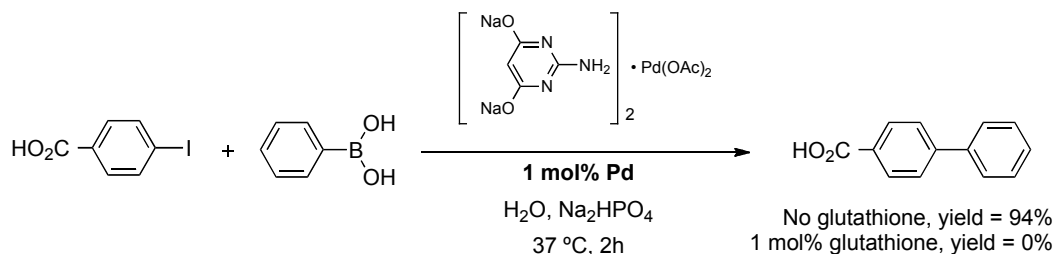
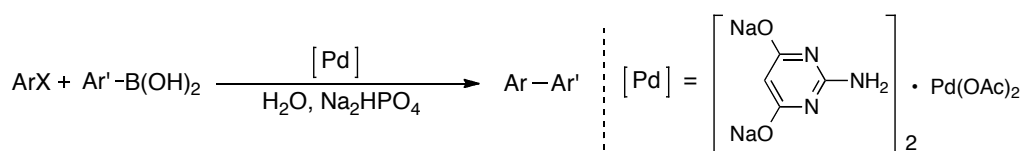


Figure 6.9: Glutathione inhibits catalysis

Hindered *ortho*-substituted biaryls were synthesized in good yield (Table 6.4, Entries 12-15). These congested couplings are notoriously difficult and a common testing ground for catalyst activity.⁷⁴⁻⁷⁶ Catalyst loading was also investigated and as little as 1 ppm of palladium was needed for simple biaryl synthesis with turnover frequencies approaching 10⁶ mol product • mol Pd⁻¹ h⁻¹ (Entries 1-4). The cross-coupling is also easily applied to large scale reactions: A 50-gram coupling of 4-bromobenzoic acid and phenylboronic acid was carried out in a beaker of water, open to air, using only 0.01 mol% palladium (Entry 5). The product biaryl crystallized directly from the reaction mixture and was isolated in excellent yield. No organic solvent was required in the reaction or workup—an important feature in the context of green chemistry.²³⁻²⁵



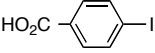
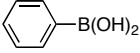
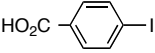
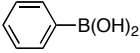
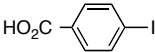
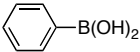
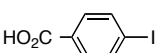
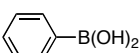
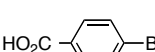
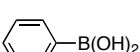
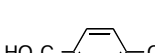
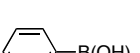
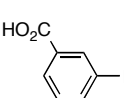
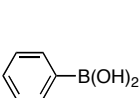
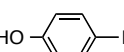
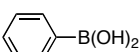
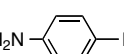
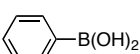
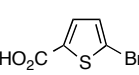
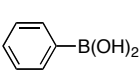
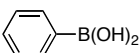
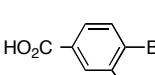
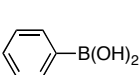
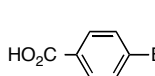
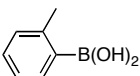
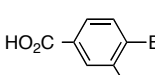
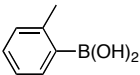
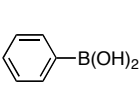
Entry	ArX	Ar'-B(OH) ₂	Pd (mol%)	Conditions	Yield (%)
1			0.01	120 °C, 10 min, μ wave	96
2			0.001	120 °C, 10 min, μ wave	96
3			0.0001	120 °C, 10 min, μ wave	19
4			0.0001	120 °C, 60 min, μ wave	90
5			0.01	90 °C, 2h, open air, 50 g	97
6			1%	120 °C, 10 min, μ wave	52
7			1%	120 °C, 10 min, μ wave	98
8			1%	120 °C, 10 min, μ wave	94
9			1%	120 °C, 10 min, μ wave	91
10			1%	120 °C, 10 min, μ wave	96
11			1%	120 °C, 10 min, μ wave	94
12			1%	120 °C, 10 min, μ wave	99
13			1%	120 °C, 10 min, μ wave	99
14			1%	120 °C, 10 min, μ wave	99
15			1%	120 °C, 10 min, μ wave (1:1 H ₂ O:DMF)	82

Table 6.4: Substrate scope and performance of palladium-pyrimidine catalyst

6.3 Conclusions and Outlook

In conclusion, a convenient catalytic system for aqueous Suzuki-Miyaura Cross-Coupling was discovered. The catalyst system is mild enough to promote cross-coupling on a protein surface and sufficiently active to mediate hindered biaryl coupling. The cross-couplings reported on protein substrates are the first to proceed to completion. Moreover, the use of boronic acids as cross-coupling partners allows bio-conjugation of otherwise water-insoluble substrates without the need for organic solvents which may denature the protein. In the case of biaryls, the coupling product is also of interest since these structures often possess an absorption and emission profile different to that of the coupling partners. Efforts to use Suzuki-Miyaura cross-coupling as a method for fluorogenic protein labeling are underway.

It is also noteworthy that this catalyst system has since been adopted in other reports of protein modification. For instance, Spicer and Davis recently described the use of the palladium-pyrimidine catalyst system for modification of maltose binding protein by cross-coupling at a genetically encoded *p*-iodophenylalanine.⁷⁷ This was the first cross-coupling at a genetically incorporated aryl halide to proceed with full conversion, vindicating the many efforts to genetically incorporate cross-coupling partners into proteins. Moreover, cross-coupling at genetically incorporated aryl halides may enable the use of this chemistry for *in vivo* applications such as cell-surface labeling. The Liu laboratory has also used this catalyst in protein modification.⁷⁸ In this report, the authors used the Pd-pyrimidine catalyst to mediate Suzuki-Miyaura coupling at the Z domain of protein A. The coupling was carried out between a fluorophore and *p*-iodophenylalanine; the latter was incorporated into the protein genetically using an evolved pyrrolysyl-tRNA / tRNA synthetase pair.

While the primary motivation for the development of this catalyst was for applications in bioconjugation, this catalytic system should be useful in general synthetic endeavors. The 2-amino-4,6-dihydroxypyrimidine ligand is commercially available at low cost—typically \$0.13 (USD) per mmol from multiple suppliers. This low cost compares favorably to other water-soluble ligands used in aqueous cross-coupling. For instance, 2-amino-4,6-dihydroxypyrimidine is nearly 1000 times cheaper than TPPTS (Figure 6.2), with the sulfonated phosphine typically selling at \$113 (USD) per mmol. While for small scale applications the cost of TPPTS is not prohibitive, gram to kilogram scale syntheses are expensive. Moreover, TPPTS is also highly sensitive to oxidation and

increases the procedural burden for high-throughput synthesis and analysis where catalyst solutions must be continually protected from oxygen. The Pd-pyrimidine system, on the other hand, is a bench stable Pd(II) source. Pfizer has since taken advantage of this catalyst in their medicinal chemistry program for the synthesis of heterocycles and a collaboration has been established to assess the utility of this catalyst in the synthesis of pharmaceuticals.⁷⁹

Currently, the use of this Pd-pyrimidine catalyst in other transformations in water is under investigation. NMR and X-ray analysis is also underway to gain some insight into the catalyst's structure. Finally, the scope of the Suzuki-Miyaura cross-coupling in bioconjugation, particularly with cross-coupling partners that are incorporated biosynthetically, is an active area of inquiry.

6.4 References

- (1) Hegedus, L. S. *Transition metals in the synthesis of complex organic molecules*, 2nd ed. University Science Books, Sausalito, 1999.
- (2) *Handbook of organopalladium chemistry for organic synthesis*. Negishi, E.-i. (Ed.); Wiley Interscience, New York, 2002.
- (3) *Cross-coupling reactions: a practical guide*. Miyaura, N. (Ed.); Springer, Berlin, 2002.
- (4) *Metal-catalyzed cross-coupling reactions*, 2nd ed. de Meijere, A., Diederich, F. (Eds.); Wiley-VCH, Weinheim, 2004.
- (5) Miyaura, N.; Yamada, K.; Suzuki, A. A new stereospecific cross-coupling by the palladium-catalyzed reaction of 1-alkenylboranes with 1-alkenyl or 1-alkynyl halides. *Tetrahedron Lett.* **1979**, 20, 3437-3440.
- (6) Miyaura, N.; Suzuki, A. Stereoselective synthesis of arylated (*E*)-alkenes by the reaction of alk-1-enylboranes with aryl halides in the presence of palladium catalyst. *J. Chem. Soc., Chem. Commun.* **1979**, 866-867.
- (7) Miyaura, N.; Suzuki, A. Palladium-catalyzed cross-coupling reactions of organoboron compounds. *Chem. Rev.* **1995**, 95, 2457-2483.
- (8) Suzuki, A. Synthetic studies via the cross-coupling reaction of organoboron derivatives with organic halides. *Pure Appl. Chem.* **1991**, 63, 419-422.

- (9) King, A. O.; Yasuda, N. Palladium-catalyzed cross-coupling reactions in the synthesis of pharmaceuticals. *Top. Organomet. Chem.* **2004**, *6*, 205-245.
- (10) Torborg, C.; Beller, M. Recent applications of palladium-catalyzed coupling reactions in the pharmaceutical, agrochemical, and fine chemical industries. *Adv. Synth. Catal.* **2009**, *351*, 3027-3043.
- (11) Schlüter, A. D. The tenth anniversary of Suzuki polycondensation (SPC). *J. Polym. Sci., Part A: Polym. Chem.* **2001**, *39*, 1533-1556.
- (12) Schlüter, A. D.; Bo, Z. In *Handbook of Organopalladium Chemistry for Organic Synthesis*; Negishi, E.-i., Ed., 2002; pp 825-861.
- (13) Kotha, S.; Lahiri, K.; Kashinath, D. Recent applications of the Suzuki-Miyaura cross-coupling reaction in organic synthesis. *Tetrahedron* **2002**, *58*, 9633-9695.
- (14) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. Palladium-catalyzed cross-coupling reactions in total synthesis. *Angew. Chem. Int. Ed.* **2005**, *44*, 4442-4489.
- (15) Armstrong, R. W.; Beau, J.-M.; Cheon, S. H.; Christ, W. J.; Fujioka, H.; Ham, W.-H.; Hawkins, L. D.; Jin, H.; Kang, S. H.; Kishi, Y.; Martinelli, M. J.; McWhorter, W. W.; Mizuno, M.; Nakata, M.; Stutz, A. E.; Talamas, F. X.; Taniguchi, M.; Tino, J. A.; Ueda, K.; Uenishi, J.-i.; White, J. B.; Yonaga, M. Total synthesis of a fully protected palytoxin carboxylic acid. *J. Am. Chem. Soc.* **1989**, *111*, 7525-7530.
- (16) Armstrong, R. W.; Beau, J.-M.; Cheon, S. H.; Christ, W. J.; Fujioka, H.; Ham, W.-H.; Hawkins, L. D.; Jin, H.; Kang, S. H.; Kishi, Y.; Martinelli, M. J.; McWhorter, W. W.; Mizuno, M.; Nakata, M.; Stutz, A. E.; Talamas, F. X.; Taniguchi, M.; Tino, J. A.; Ueda, K.; Uenishi, J.-i.; White, J. B.; Yonaga, M. Total synthesis of palytoxin carboxylic acid and palytoxin amide. *J. Am. Chem. Soc.* **1989**, *111*, 7530-7533.
- (17) Bellina, F.; Carpita, A.; Rossi, R. Palladium catalysts for the Suzuki cross-coupling reaction: an overview of recent advances. *Synthesis* **2004**, 2419-2440.
- (18) Genet, J. P.; Savignac, M. Recent developments of palladium(0) catalyzed reactions in aqueous medium. *J. Organomet. Chem.* **1999**, *576*, 305-317.
- (19) Franzen, R.; Xu, Y. Review on green chemistry - Suzuki cross coupling in aqueous media. *Can. J. Chem.* **2005**, *83*, 266-272.
- (20) Shaughnessy, K. H.; DeVasher, R. B. Palladium-catalyzed cross-coupling in aqueous media: recent progress and current applications. *Curr. Org. Chem.* **2005**, *9*, 585-604.

- (21) Shaughnessy, K. H. Beyond TPPTS: new approaches to the development of efficient palladium-catalyzed aqueous-phase cross-coupling reactions. *Eur. J. Org. Chem.* **2006**, 1827-1835.
- (22) Polshettiwar, V.; Decottignies, A.; Len, C.; Fihri, A. Suzuki-Miyaura cross-coupling reactions in aqueous media: green and sustainable syntheses of biaryls. *ChemSusChem* **2010**, 3, 502-522.
- (23) Sheldon, R. A. Organic synthesis; past, present and future. *Chem. Ind. (London)* **1992**, 903-906.
- (24) Anastas, P. T.; Warner, J. C. *Green chemistry: theory and practice*. Oxford University Press: New York, 1998.
- (25) Sheldon, R. A. The E Factor: fifteen years on. *Green Chem.* **2007**, 9, 1273-1283.
- (26) Li, C.-J.; Chan, T.-H. *Organic reactions in aqueous media*. Wiley: New York, 1997.
- (27) *Organic synthesis in water*. Grieco, P.A. (Ed.); Blackie: London, 1998.
- (28) Chanda, A.; Fokin, V. V. Organic synthesis "on water". *Chem. Rev.* **2009**, 109, 725-748.
- (29) Ojida, A.; Tsutsumi, H.; Kasagi, N.; Hamachi, I. Suzuki coupling for protein modification. *Tetrahedron Lett.* **2005**, 46, 3301-3305.
- (30) Vilaró, M.; Arsequell, G.; Valencia, G.; Ballesteros, A.; Barluenga, J. Arylation of Phe and Tyr side chains of unprotected peptides by a Suzuki-Miyaura reaction in water. *Org. Lett.* **2008**, 10, 3243-3245.
- (31) Sudol, M. Structure and function of the WW domain. *Prog. Biophys. Molec. Biol.* **1996**, 65, 113-132.
- (32) Sudol, M.; Sliwa, K.; Russo, T. Functions of WW domains in the nucleus. *FEBS Lett.* **2001**, 490, 190-195.
- (33) Dibowski, H.; Schmidtchen, F. P. Bioconjugation of peptides by palladium-catalyzed C-C cross-coupling in water. *Angew. Chem. Int. Ed.* **1998**, 37, 476-478.
- (34) Bong, D. T.; Ghadiri, M. R. Chemoselective Pd(0)-catalyzed peptide coupling in water. *Org. Lett.* **2001**, 3, 2509-2511.
- (35) Sonogashira, K.; Tohda, Y.; Hagihara, N. A convenient synthesis of acetylenes: catalytic substitutions of acetylenic hydrogen with bromoalkenes, iodoarenes and bromopyridines. *Tetrahedron Lett.* **1975**, 16, 4467-4470.

- (36) Barluenga, J.; González, J. M.; Campos, P. J.; Asensio, G. I(py)₂BF₄, a new reagent in organic synthesis: general method for the 1,2-iodofunctionalization of olefins. *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 319-320.
- (37) Barluenga, J.; García-Martín, M. A.; González, J. M.; Clapés, P.; Valencia, G. Iodination of aromatic residues in peptides by reaction with IPy₂BF₄. *Chem. Commun.* **1996**, 1505-1506.
- (38) Espuña, G.; Andreu, D.; Barluenga, J.; Pérez, X.; Planas, A.; Arsequell, G.; Valencia, G. Iodination of proteins by IPy₂BF₄, a new tool in protein chemistry. *Biochemistry* **2006**, *45*, 5957-5963.
- (39) Xie, J.; Wang, L.; Wu, N.; Brock, A.; Spraggon, G.; Schultz, P. G. The site-specific incorporation of *p*-iodo-L-phenylalanine into proteins for structure determination. *Nat. Biotech.* **2004**, *22*, 1297-1301.
- (40) Sharma, N.; Furter, R.; Kast, P.; Tirrell, D. A. Efficient introduction of aryl bromide functionality into proteins in vivo. *FEBS Lett.* **2000**, *467*, 37-40.
- (41) Kirshenbaum, K.; Carrico, I. S.; Tirrell, D. A. Biosynthesis of proteins incorporating a versatile set of phenylalanine analogues. *ChemBioChem* **2002**, *3*, 235-237.
- (42) Kast, P.; Hennecke, H. Amino acid substrate specificity of *Escherichia coli* phenylalanyl-tRNA synthetase altered by distinct mutations. *J. Mol. Biol.* **1991**, *222*, 99-124.
- (43) Ibba, M.; Kast, P.; Hennecke, H. Substrate specificity is determined by amino acid binding pocket size in *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry* **1994**, *33*, 7107-7112.
- (44) Ibba, M.; Hennecke, H. Relaxing the substrate specificity of an aminoacyl-tRNA synthetase allows in vitro and in vivo synthesis of proteins containing unnatural amino acids. *FEBS Lett.* **1995**, *364*, 272-275.
- (45) Kiga, D.; Sakamoto, K.; Kodama, K.; Kigawa, T.; Matsuda, T.; Yabuki, T.; Shirouzu, M.; Harada, Y.; Nakayama, H.; Takio, K.; Hasegawa, Y.; Endo, Y.; Hirao, I.; Yokoyama, S. An engineered *Escherichia coli* tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9715-9723.
- (46) Sakamoto, K.; Hayashi, A.; Sakamoto, A.; Kiga, D.; Nakayama, H.; Soma, A.; Kobayashi, T.; Kitabatake, M.; Takio, K.; Saito, K.; Shirouzu, M.; Hirao, I.;

- Yokoyama, S. Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucl. Acids Res.* **2002**, *30*, 4692-4699.
- (47) Kwon, I.; Wang, P.; Tirrell, D. A. Design of a bacterial host for site-specific incorporation of *p*-bromophenylalanine into recombinant proteins. *J. Am. Chem. Soc.* **2006**, *128*, 11778-11783.
- (48) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. Regioselective carbon-carbon bond formation in proteins with palladium catalysis; new protein chemistry by organometallic chemistry. *ChemBioChem* **2006**, *7*, 134-139.
- (49) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. Site-specific functionalization of proteins by organopalladium reactions. *ChemBioChem* **2007**, *8*, 232-238.
- (50) Hartman, M. C. T.; Josephson, K.; Lin, C.-W.; Szostak, J. W. An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PLoS ONE* **2007**, *2*, e972.
- (51) Baldwin, A. J.; Arpino, J. A. J.; Edwards, W. R.; Tippmann, E. M.; Jones, D. D. Expanded chemical diversity sampling through whole protein evolution. *Mol. BioSyst.* **2009**, *5*, 764-766.
- (52) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. A genetically encoded boronate-containing amino acid. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220-8223.
- (53) Kodama, K.; Fukuzawa, S.; Sakamoto, K.; Nakayama, H.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. A new protein engineering approach combining chemistry and biology, part I; site-specific incorporation of 4-iodo-L-phenylalanine in vitro by using misacylated suppressor tRNA^{Phe}. *ChemBioChem* **2006**, *7*, 1577-1581.
- (54) Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **2000**, *4*, 645-652.
- (55) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-canonical amino acids in protein engineering. *Curr. Opin. Biotechnol.* **2003**, *14*, 603-609.
- (56) Xie, J.; Schultz, P. G. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.

- (57) Ohta, A.; Yamagishi, Y.; Suga, H. Synthesis of biopolymers using genetic code reprogramming. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159-167.
- (58) de Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconjugate Chem.* **2009**, *20*, 1281-1295.
- (59) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishiwaka, K.; Yokoyama, S. An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotech.* **2002**, *20*, 177-182.
- (60) *Essentials of Glycobiology*; 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press, 2009.
- (61) Dwek, R. A. Glycobiology: toward understanding the function of sugars. *Chem. Rev.* **1996**, *96*, 683-720.
- (62) Ohtsubo, K.; Marth, J. D. Glycosylation in cellular mechanisms of health and disease. *Cell* **2006**, *126*, 855-867.
- (63) Davis, B. G. Synthesis of glycoproteins. *Chem. Rev.* **2002**, *102*, 579-601.
- (64) Gamblin, D. P.; Scanlan, E. M.; Davis, B. G. Glycoprotein synthesis: an update. *Chem. Rev.* **2009**, *109*, 131-163.
- (65) Korolev, D. N.; Bumagin, N. A. Pd-EDTA as an efficient catalyst for Suzuki-Miyaura reactions in water. *Tetrahedron Lett.* **2005**, *46*, 5751-5754.
- (66) Wei, J.-F.; Jiao, J.; Feng, J.-J.; Lv, J.; Zhang, X.-R.; Shi, X.-Y.; Chen, Z.-G. PdEDTA held in an ionic liquid brush as a highly efficient and reusable catalyst for Suzuki reactions in water. *J. Org. Chem.* **2009**, *74*, 6283-6286.
- (67) Li, J.-H.; Zhang, X.-D.; Xie, Y.-X. Efficient and copper-free Sonogashira cross-coupling reaction catalyzed by Pd(OAc)₂/pyrimidines catalytic system. *Eur. J. Org. Chem.* **2005**, 4256-4259.
- (68) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. Chemical modification of proteins at cysteine: opportunities in chemistry and biology. *Chem. Asian J.* **2009**, *4*, 630-640.
- (69) DeSantis, G.; Peach, C.; Jones, J. B. Benzophenone boronic acid photoaffinity labeling of subtilisin CMMS to probe altered specificity. *Bioorg. Med. Chem.* **2000**, *8*, 563-570.

- (70) Stabile, M. R.; Lai, W. G.; DeSantis, G.; Gold, M.; Jones, J. B.; Mitchinson, C.; Bott, R. R.; Graycar, T. P.; Liu, C.-C. Probing the specificity of the S₁ binding site of M222 mutants of subtilisin *B. lentus* with boronic acid inhibitors. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2501-2506.
- (71) Hermanson, G. T. *Bioconjugate techniques*, 2nd Ed. Academic Press, San Diego, 2008.
- (72) Le Droumaguet, C.; Wang, C.; Wang, Q. Fluorogenic click reaction. *Chem. Soc. Rev.* **2010**, *39*, 1233-1239.
- (73) Zhang, F. L.; Casey, P. J. Protein prenylation: molecular mechanisms and functional consequences. *Annu. Rev. Biochem.* **1996**, *65*, 241-269.
- (74) Moore, L. R.; Shaughnessy, K. H. Efficient aqueous-phase Heck and Suzuki couplings of aryl bromides using tri(4,6-dimethyl-3-sulfonatophenyl)phosphine trisodium salt (TXPTS). *Org. Lett.* **2004**, *6*, 225-228.
- (75) Anderson, K. W.; Buchwald, S. L. General catalysts for the Suzuki-Miyaura and Sonogashira coupling reactions of aryl chlorides and for the coupling of challenging substrate combinations in water. *Angew. Chem. Int. Ed.* **2005**, *44*, 6173-6177.
- (76) Lipshutz, B. H.; Petersen, T. B.; Abela, A. R. Room-temperature Suzuki-Miyaura couplings in water facilitated by nonionic amphiphiles. *Org. Lett.* **2008**, *10*, 1333-1336.
- (77) Spicer, C. D.; Davis, B. G. Palladium-mediated site-selective Suzuki-Miyaura protein modification at genetically encoded aryl halides. *Chem. Commun.* **2011**, *47*, 1698-1700.
- (78) Wang, Y.-S.; Russell, W. K.; Wang, Z.; Wan, W.; Dodd, L. E.; Pai, P.-J.; Russell, D. H.; Liu, W. R. The *de novo* engineering of pyrrolysyl-tRNA synthetase for genetic incorporation of L-phenylalanine and its derivatives. *Mol. BioSyst.* **2011**, *7*, 714-717.
- (79) Beal, D. M.; Burslem, G. M. Private communication. A cheap, stable, phosphine free and water soluble palladium catalyst for Suzuki-Miyaura coupling. *Synthon* **2009**, (Internal publication of Pfizer).

6.5 Experimental Procedures

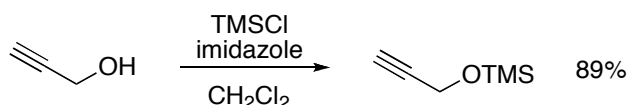
General Considerations

See Chapter 2 Experimental Procedures for general considerations in synthesis and protein analysis.

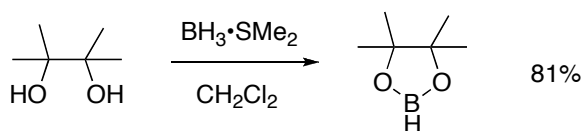
Microwave Reactions

Microwave reactions were carried out in a CEM Discover microwave reactor. All reactions were performed at variable power so that a constant internal temperature was maintained. The maximum power was set at 300 W. The maximum pressure was set to 300 psi and cooling was facilitated by a stream of nitrogen. Stirring was used in all reactions.

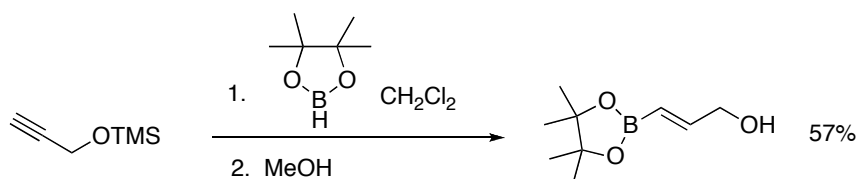
Synthesis of Cross-Coupling Substrates



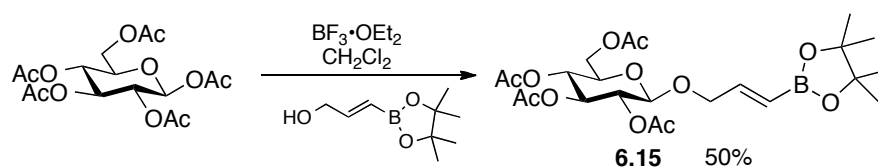
This synthesis was adapted from literature.¹ To a flame dried flask under an atmosphere of nitrogen was added CH_2Cl_2 (100 mL), propargyl alcohol (9.61 mL, 165 mmol), and imidazole (14.61 g, 214.6 mmol). The stirred solution was cooled to 0 °C and chlorotrimethylsilane (23.2 mL, 182 mmol) was added at a steady rate. The reaction was allowed to warm to room temperature over 1.25 hours at which time the reaction was judged to be complete by TLC. The reaction mixture was cooled to 0 °C and quenched with NaHCO_3 (50 mL, saturated aq.) and then diluted with 100 mL each of H_2O and CH_2Cl_2 . The organic layer was separated and washed with H_2O (2 × 100mL), dried (Na_2SO_4), and filtered. The product was concentrated in an ice bath under reduced pressure (240 mm Hg). The resulting residue was distilled (b.p. = 36-38 °C, 30 mmHg) to give the titled compound as a clear liquid (18.84 g, 89% yield). ^1H NMR (200 MHz, CDCl_3): δ = 0.15 (9H, s, 3 × CH_3), 2.38 (1H, t, J = 2.4, $\text{C}\equiv\text{CH}$), 4.26 (2H, d, J = 2.4, CH_2OTMS). ^{13}C NMR (50 MHz, CDCl_3): δ = 0.5 (TMS), 50.7 (CH_2OTMS), 73.0 ($\text{C}\equiv\text{CH}$), 82.1 ($\text{C}\equiv\text{CH}$).



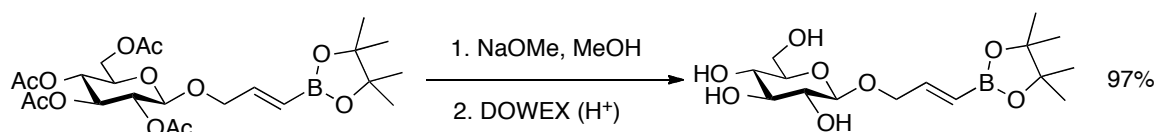
To a flame dried flask under nitrogen was added pinacol (15.0 g, 127 mmol) and CH_2Cl_2 (15 mL). The stirred solution was cooled to 0 °C and borane-dimethylsulfide complex (13.5 mL, 132 mmol) was added. Adequate headspace and nitrogen flow was ensured to compensate for the evolution of hydrogen gas. The solution was stirred at 0 °C for 20 min and then at room temperature 1.25 hours at which time no more hydrogen evolution was observed. The mixture was distilled directly with CH_2Cl_2 removal first followed by the pinacol borane (b.p. = 38-41 °C, 30 mmHg) which was isolated as a clear liquid (13.23 g, 81% yield). Spectroscopic data was consistent with that previously reported.² IR (ν_{max} , film): 3203, 1459, 1196. ^1H NMR (200 MHz, CDCl_3): δ = 1.24 (12H, s, 4 × CH_3), 2.23 (1H, s, BH). ^{13}C NMR (50 MHz): δ = 24.8 (CH_3), 83.2 (C_{OBH}).



This synthesis was adapted from the literature.¹ To a flame dried flask under an atmosphere of nitrogen was added propargylic alcohol TMS silyl ether (5.17 g, 40.3 mmol) and CH_2Cl_2 (20 mL). The stirred solution was cooled to 0 °C and pinacol borane (9.20 mL, 63.4 mmol) was added neat. The reaction mixture was allowed to warm to room temperature over two hours and then stirred for 20 hours at room temperature. The reaction was then cooled to 0 °C and MeOH was added *dropwise* by pipette with rapid stirring. After hydrogen evolution had ceased, an additional 50 mL of MeOH was added and the reaction mixture was stirred for 30 minutes at room temperature. The solvent was removed under reduced pressure and residual volatiles were removed by distillation (bath temp = 100 °C, 30 mmHg). The remaining yellow oil was purified by column chromatography (80% EtOAc in petrol) to give the titled compound as a clear oil (4.52 g, 57% yield). IR (ν_{max} , film): 3424, 2979, 1645, 1478, 1371, 1145. ^1H NMR (400 MHz, CDCl_3): δ = 1.21 (12H, s, 4 × CH_3), 4.12 (2H, app s, CH_2OH), 5.61 (1H, d, J = 18.4, $\text{CH}=\text{CHBPin}$), 6.63 (1H, dt, J = 18.4, 4.0). ^{13}C NMR (100 MHz, CDCl_3): δ = 24.6 (CH_3), 64.0 (CH_2OH), 83.1 (C_{Me_2}), 116.8 (br, $\text{C}=\text{CHBPin}$), 152.1 ($\text{CH}=\text{CHBPin}$).

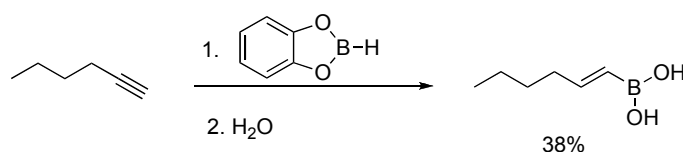


To a flame dried flask under an atmosphere of nitrogen was added β -D-glucose pentaacetate (15.82 g, 40.53 mmol) and CH_2Cl_2 (60 mL). The stirred solution was cooled to 0 °C and $\text{BF}_3 \cdot \text{OEt}_2$ (6.68 mL, 52.7 mmol) was added. The solution was stirred for 10 minutes before the pinacolboronic ester prepared above (8.95 g, 48.64 mmol) was added as a solution in CH_2Cl_2 (10 mL) followed by a CH_2Cl_2 rinse (10 mL) to ensure complete transfer. The reaction was stirred 18 hours at room temperature before it was cooled to 0 °C, quenched with H_2O (100 mL), extracted with CH_2Cl_2 ($3 \times 100\text{ mL}$), dried (MgSO_4), and filtered. The solvent was removed under reduced pressure and the residue purified by column chromatography (gradient of 20% to 50% EtOAc in petrol). The product ($R_f = 0.34$, 30% EtOAc in petrol) was collected, the solvent removed, and the resulting solid recrystallized ($\text{CH}_2\text{Cl}_2/\text{petrol}$) to give the glucose boronic ester derivative as white needles (10.33 g, 50% yield). m.p. = 140-141 °C. $[\alpha]_D^{20} = -1.90$ ($c = 0.51$, CHCl_3). IR (ν_{max} , KBr): 2980, 1752, 1642. ^1H NMR (400 MHz, CDCl_3): 1.25 (12H, s, $4 \times \text{CH}_3$, Pin), 1.99, 2.00, 2.05, 2.07 (3H each, s, $4 \times \text{OAc}$), 3.66 (1H, ddd, $J = 9.7, 4.8, 2.4$, H5), 4.11 (1H, dd, $J = 12.3, 2.4$, H6), 4.16 (1H, ddd, $J = 14.8, 4.7, 1.6$, CHHCH=), 4.23 (1H, dd, $J = 12.3, 4.8$, H6'), 4.39 (1H, ddd, $J = 14.8, 4.7, 1.6$, CHHCH=), 4.53 (1H, d, $J = 8.0$, H1), 5.01 (1H, dd, $J = 9.5, 8.0$, H2), 5.07 (1H, t, $J = 9.7$, H4), 5.18 (1H, t, $J = 9.5$, H3), 5.65 (1H, td, $J = 18.2, 1.6$, CH=CHBPin), 6.53 (1H, td, $J = 18.2, 4.7$, CH=CHBPin). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 20.5, 20.55, 20.63, 20.67$ ($4 \times \text{OAc}$), 24.69, 24.72 (contains $4 \times \text{CH}_3$ Pin), 61.8 (C6), 68.3 (C4), 70.4 ($\text{CH}_2\text{CH=}$), 71.1, 71.7 (C2 and C5), 72.8 (C3), 83.3 (CMe_2 Pin), 99.6 (C1), 120.5 (br, CH=CHBPin), 147.1 (CH=CHBPin), 169.28, 169.32, 170.2, 170.6 ($4 \times \text{OAc}$). HRMS m/z (ESI+): Found 537.2114 ($\text{M} + \text{Na}^+$); $\text{C}_{23}\text{H}_{35}\text{BO}_{12}\text{Na}$ requires 537.2119. Anal. calcd. for $\text{C}_{23}\text{H}_{35}\text{BO}_{12}\text{Na}$: C, 53.71; H, 6.86; N, -; found C, 53.62; H, 6.85; N, -.



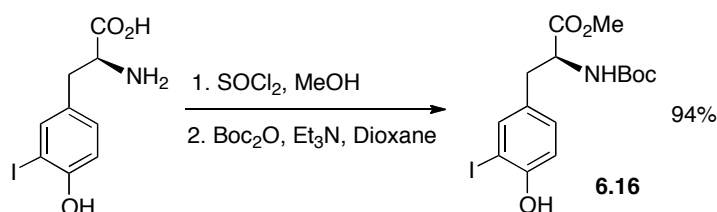
The glucose boronic ester derivative prepared above (150 mg, 0.29 mmol) was added to a 10 mL round bottom flask and dissolved in MeOH (5 mL). NaOMe (24 mg, 0.44 mmol)

was added in one portion to the stirred solution at room temperature. The reaction flask was sealed with a rubber septum and the mixture stirred at room temperature for 45 minutes. After this time, TLC analysis (5% MeOH in EtOAc) revealed complete consumption of starting material. Dowex-50WX8 was pre-washed with MeOH, acetone, 2M HCl, then H₂O until the filtrate was neutral. This conditioned resin was then added to the reaction mixture in small portions until the pH of the reaction mixture was approximately neutral (pH paper). The quenched reaction was stirred for an additional 30 minutes at room temperature to neutralize any remaining borate. The resin was filtered off and washed with 50 mL of MeOH. The solvent was removed under reduced pressure to give the deprotected sugar as a clear, hygroscopic liquid (98 mg, 97% yield). $[\alpha]_D^{20} = -3.4$ ($c = 1.1$, MeOH). IR ($\nu_{\max}$, film): 3300, 1644. ¹H NMR (400 MHz, MeOD): 1.27 (12H, s, 4 × CH₃), 3.23-3.41 (4H, m, includes H2, H3, H4, and H5), 3.69 (1H, dd, $J = 12.0, 5.2$, H6), 3.87 (1H, dd, $J = 12.0, 1.5$, H6'), 4.25 (1H, ddd, $J = 14.8, 4.7, 1.6$, CHHCH=), 4.32 (1H, d, $J = 7.6$, H1), 4.46 (1H, ddd, $J = 14.8, 4.7, 1.6$, CHHCH=), 5.72 (1H, app d, $J = 18.2$, CH=CHBPIn), 6.65 (1H, dt, $J = 18.2, 4.7$, HC=CHBPIn). ¹³C (100 MHz, MeOD): $\delta = 25.2$ (CH₃), 62.7 (C6), 71.6 (CH₂CH=), 71.5, 75.0, 77.9, 78.0 (C2, C3, C4, C5), 84.6 (CMe₂), 103.6 (C1), 120.5 (br, CH=CHBPIn), 150.3 (CH=CHBPIn). HRMS m/z (ESI⁺): found 369.1691 (M+Na)⁺; C₁₅H₂₇BO₈Na requires 369.1697.

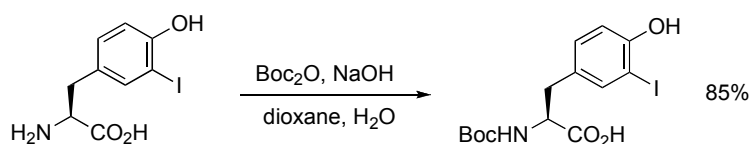


1-Hexyne (0.96 mL, 8.3 mmol) was added to a flame dried, 50 mL pear-shaped flask under a nitrogen atmosphere. Catechol borane (0.89 mL, 8.3 mmol) was added to the reaction vessel at room temperature (no solvent was used in this reaction). The reaction was heated to 50 °C and stirred for 2 hours. After this time, the reaction was cooled to room temperature and then placed on ice. H₂O (10 mL) was added slowly to the reaction mixture, resulting in the formation of a white precipitate. The mixture was stirred vigorously for 1 hour to ensure complete hydrolysis of the catechol boronic ester. The solid was isolated by filtration and washed with H₂O (20 mL). The solid was recrystallized (MeOH/H₂O) and the white crystals collected by filtration and dried under high vacuum (406 mg, 38%). m.p. = 68-70 °C. IR ($\nu_{\max}$, KBr): 2994, 1639, 1465, 1357, 1264, 1164, 996. ¹H NMR (400 MHz, DMSO): $\delta = 0.87$ (3H, t, $J = 7.2$, CH₃), 1.24-1.38

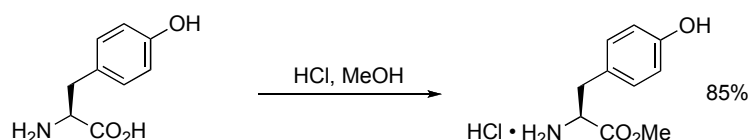
(4H, m, CH₃CH₂CH₂), 2.07 (2H, q, $J = 6.4$, CH₂CH=CH), 5.30 (1H, d, $J = 17.9$, CH=CHB(OH)₂), 6.42 (1H, dt, $J = 17.9, 6.4$, CH₂CH=CH), 7.48 (2H, s, B(OH)₂). ¹³C NMR (50 MHz, DMSO): $\delta = 13.9$ (CH₃), 21.8, 30.4, 34.8 (3 $\times$ CH₂), 124.8 (br, CH=CHB(OH)₂), 150.1 (CH=CHB(OH)₂). HRMS m/z (FI⁺): found 128.1008 (M)⁺; C₆H₁₃BO₂ requires 128.1009.



To a flame dried flask under nitrogen was added anhydrous MeOH (10 mL). The flask was cooled to 0 °C and SOCl₂ (0.26 mL, 3.6 mmol) was added dropwise with stirring. After stirring for 5 minutes at 0 °C, 3-iodo-L-tyrosine (1.00 g, 3.26 mmol) was added in one portion. The reaction was stirred for 20 hours at room temperature after which time the solvent was removed to give an off white solid (3-iodotyrosine methyl ester hydrochloride) which was used without purification. The solid was suspended in dioxane (50 mL). Boc₂O (1.28 g, 4.74 mmol) and triethylamine (1.0 mL, 7.1 mmol) were added sequentially and the reaction was stirred at room temperature for 1.5 hours. The reaction was judged complete by TLC and the dioxane was evaporated under reduced pressure. The residue was partitioned between EtOAc and H₂O (100 mL each) and separated. The aqueous layer was further extracted with EtOAc (2 $\times$ 50mL). The combined organics were washed with brine, dried (MgSO₄), filtered, and concentrated under reduced pressure. Purification by column chromatography (35% EtOAc in petrol) provided *N*-Boc-3-iodo- L-tyrosine methyl ester (1.29 g, 94% yield over two steps) as a thick, clear oil which solidified upon storage at -20 °C. m.p. = 110-112 °C (Lit³: 114-115 °C). $[\alpha]^{20}_D = +48.8$ ($c = .57$, CHCl₃). IR ($\nu_{\max}$, KBr): 3390, 1715, 1690, 1606, 1578, 1525, 1444, 1420, 1376, 1307, 1158, 1067, 1028, 994. ¹H NMR (200 MHz, CDCl₃): $\delta = 1.43$ (9H, s, Boc), 2.97 (2H, m, CH₂Ar), 3.73 (3H, s, CO₂CH₃), 4.51 (1H, q, $J = 8.0$, C α), 5.06 (1H, d, $J = 8.0$, NH), 5.99 (1H, s, OH), 6.84 (1H, d, $J = 8.4$, CH_{Ar}), 6.98 (1H, dd, $J = 8.4, 1.6$, CH_{Ar}), 7.42 (1H, d, $J = 1.6$, CH_{Ar}). ¹³C NMR (50 MHz): $\delta = 28.3$ (Boc), 36.9 (CH₂Ar), 52.3 (CO₂CH₃), 54.4 (C α), 80.2 (Boc), 85.2 (C-I), 115.0, 129.9, 130.8, 140.0, 154.2, 155.1, 172.2 (5 $\times$ C_{Ar} and 2 $\times$ C=O). LRMS (ESI⁺): 480.1 (M+MeCN+NH₄)⁺.

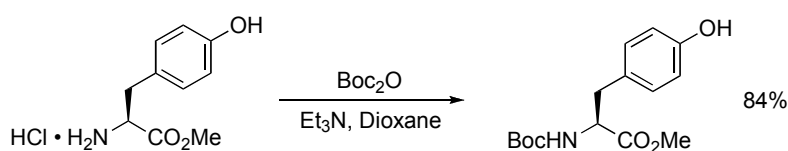


3-iodo-L-tyrosine (400 mg, 1.30 mmol) was added to a 100 mL round bottom flask and suspended in H₂O (5 mL). The stirred suspension was cooled to 0 °C and 0.78 mL of 5M NaOH was added. In a separate vial, Boc₂O (299 mg, 1.37 mmol) was dissolved in 5 mL dioxane. The solution of Boc₂O was added by pipette to the reaction flask at 0 °C. The reaction was stoppered and warmed from 0 °C to room temperature over a period of 3 hours. After this time, the reaction was quenched with 1M HCl (100 mL) and extracted with EtOAc (3 × 50 mL). The combined organics were washed with brine (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (15% MeOH in CH₂Cl₂) to afford 449 mg of *N*-Boc-3-iodo-L-tyrosine (85% yield) as a white solid. m.p. = 88-90 °C (Lit⁴: 83-85 °C). $[\alpha]_D^{20} = +28.0$ (c = 0.25, MeOH). IR ($\nu_{\max}$, film): 3330, 2978, 2360, 1683, 1575, 1412, 1368, 1289, 1252, 1161, 1054, 1036. ¹H NMR (400 MHz, MeOD): δ = 1.38 (Boc), 2.92 (2H, ABX system J = 13.6, 8.3, 4.6, CH₂Ar), 4.23 (1H, dd, J = 8.3, 4.6, H α), 6.75 (1H, d, J = 8.3, CH_{Ar}), 7.04 (1H, d, J = 8.3, CH_{Ar}), 7.54 (1H, s, CH_{Ar}). ¹³C NMR (100 MHz, MeOD): δ = 28.8 (Boc), 37.5 (CH₂Ar), 56.9 (C α), 80.6, 84.5 (C-I and Boc), 115.6, 131.5, 131.7, 141.0, 156.6, 157.6, 176.9 (5 × C_{Ar} and 2 × C=O). LRMS m/z (ESI): 405.9 [M-H]⁻.



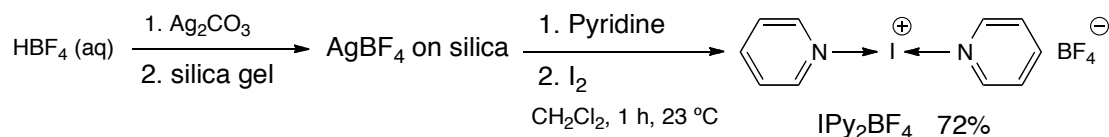
A 250 mL round bottom flask was flame dried and placed under an atmosphere of nitrogen. Anhydrous MeOH (80 mL) was added to the flask and cooled to 0 °C. Acetyl chloride (193 mmol, 13.74 mL) was added slowly with stirring to generate methanolic HCl. After all of the acetyl chloride was added, the solution was stirred an additional 15 minutes at 0 °C. L-Tyrosine (7.00 g, 38.6 mmol) was added as a solid in one portion. The ice bath was removed and the reaction stirred at room temperature for 12 hours. After this time, volatiles were removed under reduced pressure to give the crude product as an off-white solid. The solid was triturated by stirring rapidly in 50 mL of boiling acetone for 10 minutes. After cooling to room temperature, the flask was cooled on ice for several

minutes. The resulting micro-crystals were isolated by filtration and washed with 50 mL of cold acetone. The crystals were then dried under vacuum (7.65 g, 85% yield). m.p. = 172-174 °C, decomposes at 190 °C (Lit.⁵: 190-192, decomp.). $[\alpha]^{20}_D = +69.8$ (c = 3.0, pyridine) (Lit.⁶: +72.0, c = 3.0, pyridine). IR ($\nu_{\max}$, KBr): 3343, 2879, 1958, 1744, 1614, 1592, 1515, 1448, 1397, 1349, 1226, 1141, 1107, 1060. ¹H NMR (400 MHz, MeOD): δ = 3.10-3.22 (2H, ABX system, J = 14.3, 7.2, 6.1, $\underline{\text{CH}_2\text{Ar}}$), 3.82 (3H, s, $\text{CO}_2\underline{\text{CH}_3}$), 4.28 (1H, app t. J = 6.7, $\text{H}\alpha$), 6.81 (2H, d, J = 8.5, ArH), 7.10 (2H, d, J = 8.5, ArH). ¹³C NMR (50 MHz, MeOD): δ = 36.7 ($\underline{\text{CH}_2\text{Ar}}$), 53.7, 55.6 ($\text{C}\alpha$ and $\text{CO}_2\underline{\text{CH}_3}$), 117.0, 125.8, 131.7, 158.4 (Ar), 170.6 (C=O). LRMS m/z (ESI): 194.1 $[\text{M}-\text{H}]^-$.



L-Tyrosine methyl ester hydrochloride (3.23 g, 16.6 mmol) was added to a 250 mL round bottom flask and suspended in dioxane (50 mL). To the stirred suspension was added Et₃N (4.65 mL, 33.1 mmol). Boc₂O (4.70 g, 21.5 mmol) was then added in one portion and the reaction was stirred at room temperature. After 5 hours of reaction time, the reaction was diluted with EtOAc (150 mL) and washed with H₂O (150 mL). The aqueous layer was extracted with EtOAc (2 × 150 mL) and the combined organic layers were washed with brine (200 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography by eluting first with 15% EtOAc in petrol and then 50% EtOAc in petrol. The product was isolated as a white solid (4.13 g, 84% yield). m.p. = 96-99 °C (Lit.⁷ = 102 °C). $[\alpha]^{20}_D = +41.2$ (c = 1.0, CHCl₃) (Lit.⁸: +53.1, c = 1.0, CHCl₃), IR ($\nu_{\max}$, KBr): 3370, 2979, 1687, 1516, 1445, 1368, 1224, 1166. ¹H (200 MHz, CDCl₃): δ = 1.42 (9H, s, Boc), 3.00 (2H, ABX system, J = 5.8, 6.0, 13.6, $\underline{\text{CH}_2\text{Ar}}$), 3.71 (3H, s, $\text{CO}_2\underline{\text{CH}_3}$), 4.54 (1H, m, $\text{H}\alpha$), 5.06 (1H, d, J = 4.0, NH), 6.48 (1H, br s, OH), 6.73 (2H, d, J = 4.0, CH_{Ar}), 6.96 (2H, d, J = 4.0, CH_{Ar}). ¹³C (50 MHz, CDCl₃): δ = 28.3 (Boc), 37.6 ($\underline{\text{CH}_2\text{Ar}}$), 52.4 ($\text{CO}_2\underline{\text{CH}_3}$), 54.6 ($\text{C}\alpha$), 80.3 (Boc), 115.5, 127.3, 130.3, 155.3, 155.4, 172.7 (C_{Ar} and C=O). LRMS m/z (ESI): 294.1 $[\text{M}-\text{H}]^-$.

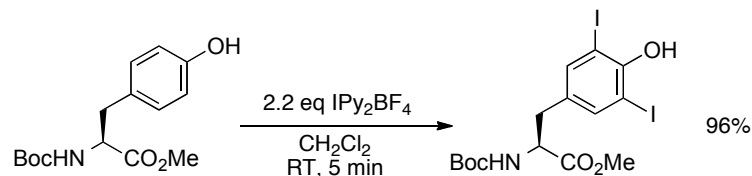
Barluenga's Reagent (IPy₂BF₄)



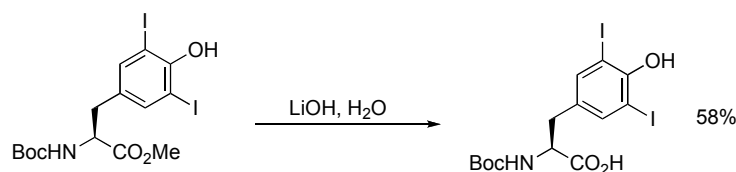
Barluenga's reagent⁹ (IPy₂BF₄) was prepared according to a modified synthesis recently reported by our group.¹⁰ The reader is referred to this latter report and to Chapter 7 of this thesis for a detailed procedure for the preparation of this useful reagent. The preparation is summarized below:

To a 500 mL round bottom flask was added deionized H₂O (50 mL) and HBF₄ (6.3 mL of an 8.0 M solution, 50 mmol). Ag₂CO₃ (6.89 g, 25.0 mmol) was added to the stirred solution in several portions as a solid. The mixture was stirred vigorously for 20 minutes to dissolve all material. To the resulting solution was added silica gel (10.0 g). The slurry was stirred at room temperature for 5 minutes. After this time, the stirrer bar was removed and the water was evaporated under reduced pressure (55 °C, 30 mm Hg) to give AgBF₄ on silica gel. The solid was dried by rotary evaporation until the AgBF₄ on silica was free flowing. The flask was cooled to room temperature and then carried on to the next step. Accordingly, CH₂Cl₂ (300 mL) was added and the mixture stirred. Care was taken to ensure all silica gel was suspended in the solvent. Pyridine (8.1 mL, 100 mmol) and iodine (12.69 g, 50.0 mmol) were added sequentially to the reaction flask. AgI precipitated immediately. The flask was stoppered and the reaction stirred vigorously at room temperature for 1 hour. After this time, solids (AgI and silica) were removed by filtration. The filter cake was washed with CH₂Cl₂ (100 mL). The filtrate was concentrated under reduced pressure to give a red solid. The solid was dissolved in CH₂Cl₂ (100 mL) and stirred at 0 °C. Et₂O (200 mL) was poured into the stirred solution to precipitate the IPy₂BF₄. The product was isolated by filtration and dried under vacuum to give Barluenga's reagent as a white powder (13.30 g, 72% yield). This material is sufficiently pure for most applications, but can be recrystallized from CH₂Cl₂ to give the titled compound as colorless prisms (12.03 g combined from 3 crops of crystals, 65% yield). The following analytical data was obtained from the first crop of crystals: m.p = 148-151 °C (Lit⁹: 149-151 °C). IR (KBr) cm⁻¹: 1603, 1447, 1209, 1056, 760, 689, 636, 521, 440. ¹H NMR (400 MHz, CD₃CN) δ = 7.63 (4H, dd, *J* = 7.7, 6.4 Hz, H-3), 8.24 (2H,

tt, $J = 7.7, 1.4$ Hz, H-4), 8.78 (4H, dd, $J = 6.4, 1.4$, H-2). ^{13}C NMR (100 MHz, CD_3CN) δ : 129.2 (C-3), 143.5 (C-4), 150.9 (C-2). ^{19}F NMR (377 MHz, CD_3CN) δ : -151.5 (BF_4^-). Anal. Calcd. for $\text{C}_{10}\text{H}_{10}\text{BF}_4\text{IN}_2$: C, 32.29; H, 2.71; N, 7.53. Found C, 32.33; H, 2.70; N, 7.53.

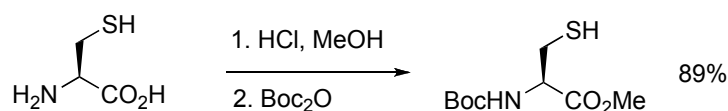


To a solution of *N*-Boc-L-tyrosine-methyl ester (1.76 g, 5.97 mmol) in 50 mL CH_2Cl_2 was added IPy_2BF_4 (4.88 g, 13.1 mmol) in one portion. The TLC was checked immediately and showed complete consumption of starting material (less than 5 min at room temperature). The reaction was quenched with $\text{Na}_2\text{S}_2\text{O}_3$ (50 mL, 20% wt.) and extracted with CH_2Cl_2 (2×50 mL). The combined organic layers were dried over MgSO_4 , filtered, and concentrated under reduced pressure. The product was purified by column chromatography (40% EtOAc in petrol). The di-iodinated product was isolated as a white solid (3.08 g, 96%) that was spectroscopically identical to that reported in the literature.¹¹ m.p. = 117-119 °C. $[\alpha]_D^{20} = +42.2$ ($c = 1.1$, CHCl_3). IR (ν_{max} , KBr): 3455, 3363, 2979, 1759, 1729, 1687, 1523, 1458, 1406, 1254, 1164. ^1H (200 MHz, CDCl_3): $\delta = 1.44$ (9H, s, Boc), 2.95 (2H, ABX system, $J = 5.8, 6.2, 13.8$, CH_2Ar), 3.73 (3H, s, CO_2CH_3), 4.48 (1H, m, H α), 5.03 (1H, d, $J = 7.6$, NH), 7.43 (2H, s, $2 \times \text{CH}_{\text{Ar}}$). ^{13}C (50 MHz, CDCl_3): $\delta = 28.3$ (Boc), 36.3 (CH_2Ar), 52.4 (CO_2CH_3), 54.3 (C α), 80.2, 82.2 (C-I and Boc), 132.2, 139.9, 152.6, 154.9, 171.8 ($3 \times \text{C}_{\text{Ar}}$ and $2 \times \text{C}=\text{O}$). LRMS m/z (ESI): 545.9 $[\text{M}-\text{H}]^-$.



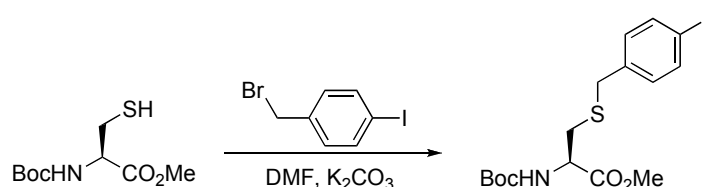
N-Boc-3,5-di-iodo-tyrosine methyl ester (1.45 g, 2.64 mmol) was added to a 100 mL round bottom flask and dissolved in THF (10 mL). A solution of $\text{LiOH} \cdot \text{H}_2\text{O}$ (1.11 g, 26.5 mmol) in H_2O (10 mL) was added by pipette to the stirred solution at room temperature. The reaction was stirred at room temperature for 1 hour, after which time no starting material was detected by TLC (EtOAc). The reaction was cooled to 0 °C and quenched by the addition of 1M HCl (100 mL). The product was extracted with EtOAc (3×75 mL). The combined organics were washed with brine (100 mL), dried (MgSO_4), filtered, and

concentrated under reduced pressure to give the crude product as a solid. The solid was recrystallized (EtOAc/petrol) to give the product as tiny white needles. The needles were collected by filtration and dried under vacuum (815 mg, 58% yield). m.p. = 170-171 °C. $[\alpha]_D^{20} = +16.3$ ($c = 1.0$, MeOH). IR ($\nu_{\max}$, KBr): 3489, 3339, 2979, 1691, 1256, 1458, 1403, 1275, 1166, 1056. ^1H (400 MHz, MeOD): $\delta = 1.41$ (9H, s, Boc), 2.77 (1H, dd, $J = 13.9, 9.4$, CHHAr), 3.06 (1H, dd, $J = 13.9, 4.8$, CHHAr), 4.28 (1H, dd, $J = 9.8, 4.8$, $\text{H}\alpha$), 7.61 (2H, s, CHAr). ^{13}C (100 MHz, MeOD): $\delta = 28.9$ (Boc), 36.8 (CH_2Ar), 56.1 ($\text{C}\alpha$), 80.7, 85.3 (C-I and Boc), 134.7, 141.5, 155.6, 157.8, 174.9 ($3 \times \text{CAr}$ and $2 \times \text{C=O}$). HRMS m/z (ESI+): Found 555.9094 ($\text{M} + \text{Na}^+$); $\text{C}_{14}\text{H}_{17}\text{I}_2\text{O}_5\text{Na}$ requires 555.9094.

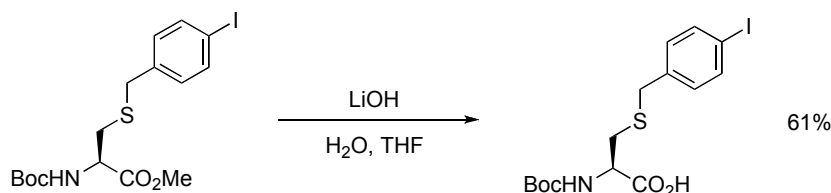


MeOH (100 mL) was added to a flame dried 250 mL round bottom flask. The solvent was stirred and cooled to 0 °C and acetyl chloride (248 mmol, 17.6 mL) was added dropwise over 5 minutes. The solution was stirred an additional 10 minutes at 0 °C to give a concentrated solution of HCl. L-Cysteine (2.00 g, 16.5 mmol) was then added in one portion and the flask flushed with argon. The ice bath was removed and the reaction was stirred at room temperature for 24 hours. The solvent was then removed under reduced pressure to give the crude cysteine methyl ester hydrochloride as a pale yellow solid. This material was used immediately in the next step without purification. The crude ester was suspended in CH_2Cl_2 (100 mL) and cooled to 0 °C. Et_3N (5.06 mL, 36.3 mmol) was added carefully followed by Boc_2O (4.32 g, 19.8 mmol). The reaction was stirred at room temperature for 3.25 hours after which time TLC (30% EtOAc in Petrol) revealed the desired product ($R_f = 0.6$) and its corresponding disulfide ($R_f = 0.3$). The solvent was removed under reduced pressure and the resulting residue was redissolved in MeOH (40 mL) and H_2O (8 mL). PBU_3 (2.0 mL, 8.1 mmol) was added dropwise to the stirred solution. TLC revealed reduction of the disulfide. The reaction was diluted with Et_2O (100 mL) and H_2O (50 mL). The organic layer was separated and the aqueous layer was extracted with Et_2O (2×50 mL). The combined organics were washed with brine (100 mL), dried over MgSO_4 , and filtered. The solvent was removed by rotary evaporation and the residue purified by column chromatography eluting first with 5% EtOAc in petrol and then 20% EtOAc in petrol. BocCysOMe was isolated as a clear oil (3.48 g, 89% yield).

from L-cysteine). $[\alpha]_D^{20} = +28.3$ ($c = 7.5$, CHCl_3) (Lit.¹²: $[\alpha]_D^{20} = +28.5$, $c = 7.5$, CHCl_3). IR (ν_{max} , film): 2979, 1717, 1506, 1367, 1166. ^1H NMR (CDCl_3 , 400MHz): $\delta = 1.42$ (10H, s, includes Boc and SH), 2.94 (2H, app. td, $J = 8.7, 4.3$, CH_2SH), 3.76 (3H, s, CO_2CH_3), 4.58 (1H, m, $\text{H}\alpha$), 5.44 (1H, d, $J = 5.8$, NH). ^{13}C NMR (100 MHz, CDCl_3): $\delta = 27.3$ (CH_2SH), 28.2 (Boc), 52.6 (CO_2CH_3), 54.8 ($\text{C}\alpha$), 80.2 (Boc), 155.1, 170.8 ($2 \times \text{C=O}$). LRMS (ESI): 234.1 (M-H)⁻.

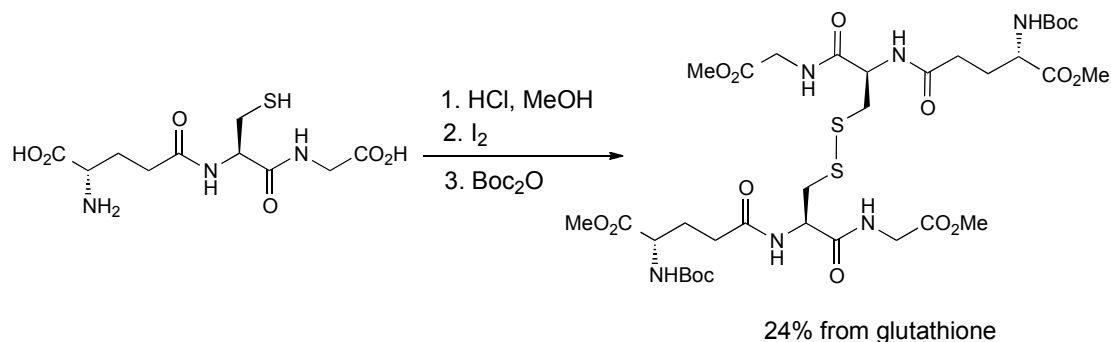


BocCysOMe (1.40 g, 4.17 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (20 mL). K_2CO_3 (1.73 g, 12.5 mmol) and 4-iodo-benzylbromide (742 mg, 2.50 mmol) were added and the reaction flushed with nitrogen. The reaction was stirred at room temperature for 2 hours and then diluted with EtOAc (200 mL). The solution was washed with 200 mL each of H_2O and brine. The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The residue was then passed through a silica plug, eluting with 20% EtOAc in petrol. The resulting product was carried to the next step without further purification (see next).



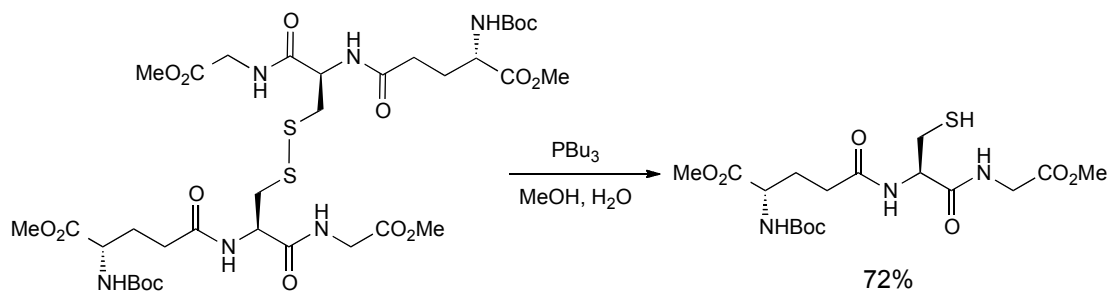
The aryl iodide from above (1.13 g, 2.50 mmol) was dissolved in THF (10 mL) and stirred at room temperature. To the stirred solution was added an aqueous solution of $\text{LiOH}\cdot\text{H}_2\text{O}$ (1.05 g, 25.0 mmol in 10 mL H_2O). The reaction mixture was stirred vigorously (open air) for 1 hour. TLC analysis (40% EtOAc in petrol) revealed complete consumption of starting material. The reaction was quenched with 1M HCl (100 mL) and then extracted with EtOAc (3×75 mL). The combined organic layers were washed with brine (100 mL), dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (20% MeOH in CH_2Cl_2 with 2% NH_4OH). The product was isolated as a thick oil which crystallized upon drying under high vacuum (670 mg, 61% yield). m.p. = 92-94 °C. $[\alpha]_D^{20} = -4.8$ ($c = 0.25$, MeOH). IR (ν_{max} , KBr): 3159, 2977, 2360, 1901, 1694, 1585, 1483, 1454, 1393, 1367,

1248, 1164, 1058, 1024. ^1H NMR (MeOD, 400MHz): δ = 1.45 (9H, s, Boc), 2.74-2.92 (2H, ABX system, J = 13.6, 6.8, 4.6, $\text{CH}_2\text{SCH}_2\text{Ar}$), 3.71 (2H, s, SCH_2Ar), 4.23 (1H, t, J = 5.4, $\text{H}\alpha$), 7.12 (2H, d, J = 8.1, ArH), 7.62 (2H, d, J = 8.1, ArH). ^{13}C NMR (100 MHz, MeOD): δ = 28.9 (Boc), 35.2 ($\text{C}\beta$), 36.7 (SCH_2Ar), 55.9 ($\text{C}\alpha$), 79.6, 80.5 (Boc, rotamers), 93.0 (C-I), 132.3, 138.6 (CH_{Ar}), 139.9 (C_{Ar}), 157.6, 176.6 ($2 \times \text{C}=\text{O}$). HRMS m/z (ESI $^+$): Found 460.0055 ($\text{M}+\text{Na}$) $^+$; $\text{C}_{15}\text{H}_{20}\text{NO}_4\text{INa}$ requires 460.0055.

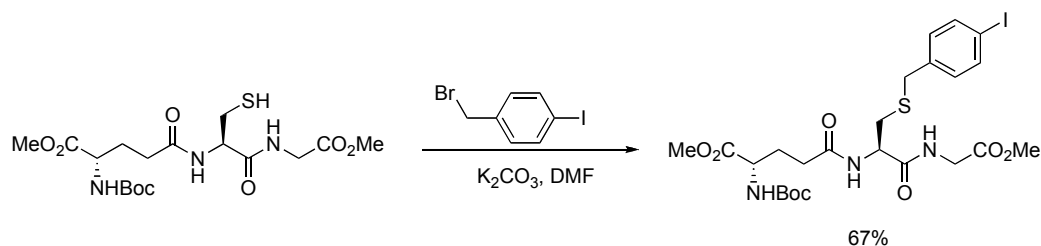


To a flame dried 250 mL round bottom flask was added MeOH (30 mL). The flask was cooled to 0 °C and acetyl chloride (6.97 mL, 97.6 mmol) was added slowly. The solution was stirred 15 minutes to generate methanolic HCl. L-Glutathione (3.00 g, 9.76 mmol) was added in a single portion and the ice bath removed. The reaction was stirred at room temperature for 6 hours and then the volatiles removed by rotary evaporation. The crude residue was dissolved in DMF (50 mL) and cooled to 0 °C. Et_3N (6.80 mL, 48.8 mmol) was slowly added to the stirred solution (open air). Iodine beads (1.24 g, 4.88 mmol) were then added slowly with vigorous stirring. After 15 minutes, Boc_2O (2.13 g, 9.76 mmol) was added as a solid and the reaction flask was rinsed with 10 mL DMF. The reaction was stirred 1 hours and then diluted with EtOAc (250 mL). The solution was successively washed with 1M HCl (250 mL), H_2O (100 mL), and brine (100 mL). The organic layer was then dried (MgSO_4), filtered, and concentrated under reduced pressure. The residue was purified by column chromatography (5% MeOH in CH_2Cl_2) and the isolated product recrystallized from CH_2Cl_2 /petrol. The white crystals were isolated by filtration and dried under high vacuum (1.05 g, 24% yield from glutathione). m.p. = 121-123 °C. $[\alpha]_{\text{D}}^{20}$ = -85.6 (c = 0.5, MeOH). IR (ν_{max} , KBr): 3300, 2360, 2341, 1746, 1699, 1652, 1522, 1437, 1367, 1213, 1167. ^1H NMR (CDCl_3 , 400MHz): δ = 1.41 (18H, s, Boc), 1.92-2.01 (2H, m, $\text{H}\beta_{\text{Glu}}$), 2.13-2.21 (2H, m, $\text{H}\beta'_{\text{Glu}}$), 2.36-2.42 (4H, m, $\text{H}\gamma_{\text{Glu}}$), 2.92 (2H, dd, J = 15.4, 4.6, CHHS), 3.02 (2H, dd, J = 15.4, 3.5, CHHS), 3.71 (6H, s, CO_2CH_3), 3.74 (6H, s, CO_2CH_3), 3.94-4.17 (4H, m, $\text{H}\alpha_{\text{Gly}}$), 4.32-4.38 (2H, m, $\text{H}\alpha_{\text{Glu}}$), 5.40 (2H, d, J = 8.3,

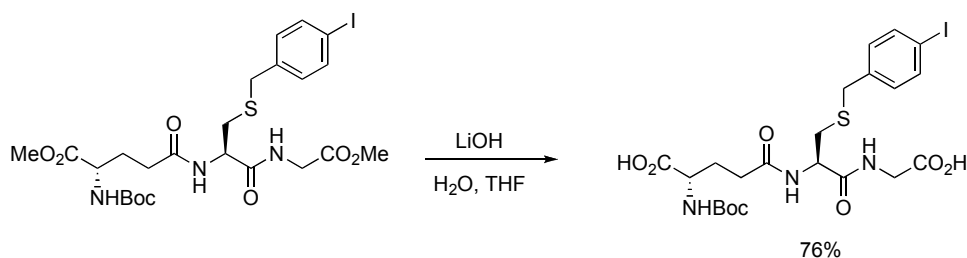
NH_{Glu}), 5.50 (2H, m, H α _{Cys}), 7.05 (2H, d, J = 9.1, NH_{Cys}), 8.56 (2H, app. s, NH_{Gly}). ¹³C NMR (100 MHz, CDCl₃): δ = 28.2 (Boc), 28.4 (C β _{Glu}), 32.2 (C γ _{Glu}), 41.2 (C α _{Gly}), 45.9 (CH₂S), 52.25, 52.33, 52.8, 52.9 (contains 2 $\times$ CO₂CH₃, C α _{Cys}, C α _{Glu}), 80.0 (Boc), 155.6, 169.9, 170.8, 172.5, 172.9 (5 $\times$ C=O). HRMS m/z (ESI⁺): Found 891.3094 (M+Na)⁺; C₃₄H₅₆N₆O₁₆S₂Na requires 891.3092.



BocGSHOME disulfide (802 mg, 0.92 mmol) was added to a 100 mL round bottom flask and dissolved in MeOH (10 mL). H₂O (2 mL) and PBu₃ (0.30 mL, 1.20 mmol) were added sequentially to the reaction mixture and the solution stirred vigorously for 30 minutes. The solvent was then removed under reduced pressure and the resulting residue was purified directly by column chromatography (5% MeOH in CH₂Cl₂). The isolated product was then recrystallized (CH₂Cl₂/petrol) to give the reduced glutathione derivative as white crystals (581 mg, 72% yield). m.p. = 112-113 °C. [α]_D²⁰ = -36.4 (c = 0.5, MeOH). IR (ν_{max} , KBr): 3311, 2956, 1746, 1712, 1653, 1523, 1438, 1367, 1212, 1168, 1052, 1028. ¹H NMR (CDCl₃, 400MHz): δ = 1.41 (9H, s, Boc), 1.80 (1H, t, J = 8.8, SH), 1.91-1.98 (1H, m, H β _{Glu}), 2.13-2.20 (1H, m, H β' _{Glu}), 2.37 (2H, t, J = 7.1, H γ _{Glu}), 2.78-2.85 (1H, m, CHHSH), 2.98-3.05 (1H, m, CHHSH), 3.72 (3H, s, CO₂CH₃), 3.73 (3H, s, CO₂CH₃), 4.03 (2H, d, J = 5.6, H α _{Gly}), 4.30 (1H, m, H α _{Glu}), 4.74 (1H, m, H α _{Cys}), 5.45 (1H, d, J = 7.8, NH_{Glu}), 7.02 (1H, d, J = 7.6, NH_{Cys}), 7.37 (1H, br. s, NH_{Gly}). ¹³C NMR (100 MHz, CDCl₃): δ = 26.5 (CH₂SH), 28.2 (Boc), 28.4 (C β _{Glu}), 32.0 (C γ _{Glu}), 41.1 (C α _{Gly}), 52.3, 52.5, 52.7 (C α _{Glu} and 2 $\times$ CO₂CH₃), 54.3 (C α _{Cys}), 80.1 (Boc), 155.7, 170.0, 170.3, 172.3, 172.8 (5 $\times$ C=O). HRMS m/z (ESI⁺): Found 458.1566 (M+Na)⁺; C₁₇H₂₉N₃O₈SNa requires 458.1568.



BocGSHOMe (500 mg, 1.15 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (10 mL). K_2CO_3 (317 mg, 2.30 mmol) and 4-iodo-benzylbromide (444 mg, 1.50 mmol) were added sequentially. The reaction was flushed with nitrogen and stirred at room temperature for 3 hours. TLC analysis (5% MeOH in CH_2Cl_2) revealed complete consumption of thiol. The reaction was diluted with EtOAc (250 mL), and then washed with H_2O (2×200 mL) and brine (200 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure. The product was purified by column chromatography and the isolated material recrystallized (CH_2Cl_2 /petrol) to afford the product as a white solid (748 mg, 67% yield). m.p. = 79-80 °C. $[\alpha]_D^{20} = -41.8$ ($c = 0.5$, MeOH). IR (ν_{max} , KBr): 2978, 2952, 2252, 1747, 1700, 1645, 1530, 1484, 1437, 1392, 1367, 1213, 1168, 1058, 1007. 1H NMR ($CDCl_3$, 400MHz): $\delta = 1.38$ (9H, s, Boc), 1.87-1.96 (1H, m, $H\beta_{Glu}$), 2.07-2.18 (1H, m $H\beta'_{Glu}$), 2.31 (2H, t, $J = 7.3$, $H\gamma_{Glu}$), 2.76 (2H, d, $J = 6.1$, $H\beta_{Cys}$), 3.64 (2H, s, SCH_2Ar), 3.68 (3H, s, CO_2CH_3), 3.69 (3H, s, CO_2CH_3), 3.98 (2H, m, $H\alpha_{Gly}$), 4.30 (1H, m, $H\alpha_{Glu}$), 4.70 (1H, m, $H\alpha_{Cys}$), 5.51 (1H, d, $J = 7.8$, NH_{Glu}), 7.04 (3H, d, $J = 8.3$, contains CH_{Ar} and NH_{Cys}), 7.41 (1H, br. s, NH_{Gly}), 7.57 (2H, d, $J = 8.3$, CH_{Ar}). ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 28.1$ (Boc), 28.2 ($C\beta_{Glu}$), 31.9 ($C\gamma_{Glu}$), 33.1 ($C\beta_{Cys}$), 35.6 (SCH_2Ar), 41.1 ($C\alpha_{Gly}$), 51.9, 52.2, 52.3, 52.7 ($C\alpha_{Glu}$, $C\alpha_{Cys}$, $2 \times CO_2CH_3$), 79.9 (Boc), 92.4 (C-I), 130.9 (CH_{Ar}), 137.5 (C_{Ar}), 155.6, 169.8, 170.7, 172.2, 172.8 ($5 \times C=O$). HRMS m/z (ESI $^+$): Found 674.1019 ($M+Na$) $^+$; $C_{24}H_{34}N_3O_8ISNa$ requires 674.1009.

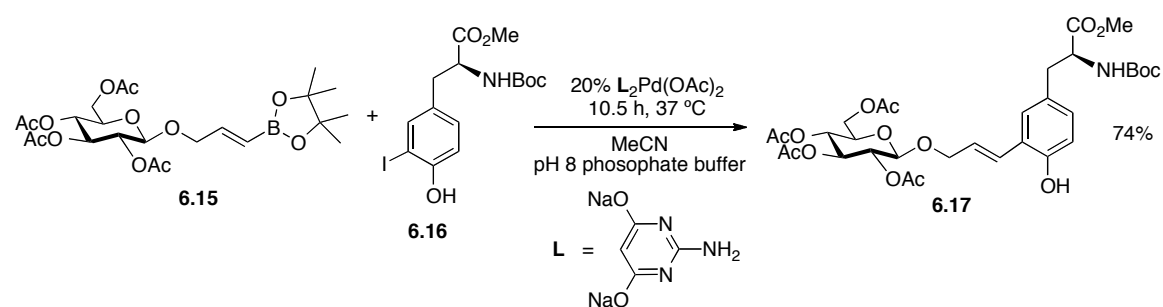


The glutathione methyl ester derivative (460 mg, 0.71 mmol) was added to a 50 mL round bottom flask and dissolved in THF (7 mL). An aqueous solution of $LiOH \cdot H_2O$

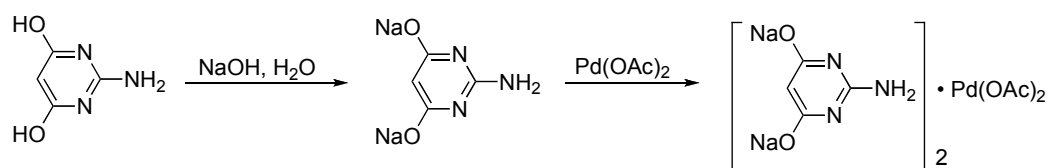
(296 mg, 7.06 mmol in 7 mL H₂O) was added dropwise by pipette to the stirred peptide solution. The reaction was stirred vigorously at room temperature for 1.5 hours and then quenched with 1M HCl (50 mL). The product was extracted with EtOAc (4 × 50 mL). The combined organic layers were then washed with brine (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (7:2:1 ⁱPrOH:MeOH:NH₄OH) to afford the product as a glassy foam (335 mg, 76% yield). $[\alpha]_D^{20} = -22.4$ (*c* = 0.25, MeOH). IR ($\nu_{\max}$, film): 3253, 3056, 2978, 1651, 1585, 1483, 1395, 1251, 1166, 1057, 1023, 1008. ¹H NMR (MeOD, 400MHz): δ = 1.44 (9H, s, Boc), 1.91-2.02 (1H, m, H β _{Glu}), 2.07-2.18 (1H, m, H β' _{Glu}), 2.37 (2H, t, *J* = 7.6, H γ _{Glu}), 2.70 (1H, dd, *J* = 13.8, 8.7, CHH₂SCH₂Ar), 2.94 (1H, dd, *J* = 13.8, 5.2, CHH₂SCH₂Ar), 3.72 (2H, s, SCH₂Ar), 3.77 (2H, AB system, *J* = 17.2, H α _{Gly}), 4.02 (1H, t, *J* = 6.2, H α _{Glu}), 4.60 (1H, dd, *J* = 8.7, 5.2, H α _{Cys}), 7.14 (2H, d, *J* = 8.1, CH_{Ar}), 7.64 (2H, d, *J* = 8.3, CH_{Ar}). ¹³C NMR (100 MHz, MeOD): δ = 29.0 (Boc), 30.2 (C β _{Glu}), 33.4 (C γ _{Glu}), 34.2 (C β _{Cys}), 36.5 (SCH₂Ar), 44.4 (C α _{Gly}), 54.1 (C α _{Cys}), 56.4 (C α _{Glu}), 80.3 (Boc), 93.0 (C-I), 132.4 (CH_{Ar}), 138.8 (CH_{Ar}), 139.7, 157.8, 172.6, 175.8, 175.9, 178.6 (C_{Ar} and 5 × C=O). HRMS *m/z* (ESI⁺): Found 646.0694 (M+Na)⁺; C₂₂H₃₀N₃O₈ISNa requires 646.0696.

Cross-Couplings from Table 6.1:

Model Suzuki-Miyaura coupling with 2-amino-4,6-dihydroxypyrimidine ligand



Catalyst Preparation



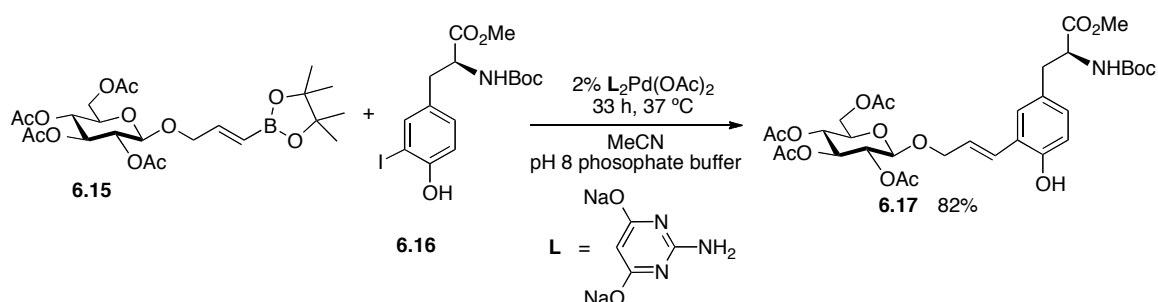
To a 5.00 mL volumetric flask was added 2-amino-4,6-dihydroxypyrimidine (13 mg, 0.10 mmol) and 2 mL of a 0.10 M NaOH stock solution. The pyrimidine ligand was dissolved

completely by stirring for 2 minutes in a water bath preheated to 65 °C. To the resulting solution was added Pd(OAc)₂ (11.0 mg, 0.05 mmol). The mixture was stirred vigorously at 65 °C for 30 minutes (open air) to give a homogenous yellow-orange solution. After cooling to room temperature, the stir bar was removed and the solution diluted to 5.00 mL with distilled water to give a catalyst solution 0.01 M in Pd(II). This catalyst solution is active after storage in a capped vial for over 1 year.

Suzuki-Miyaura Coupling

The glucose boronic ester derivative **6.15** (147 mg, 0.28 mmol) and *N*-Boc-3-iodo-L-tyrosine methyl ester **6.16** (40 mg, 0.10 mmol) were added to a two-neck flask under a stream of nitrogen and dissolved in MeCN (4 mL). To the stirred solution was added pH 8.0 sodium phosphate buffer (3 mL, 68 mM). The palladium solution prepared above (0.47 mL, 4.7 μmol, 5 mol%) was added to give a bright yellow solution. The reaction mixture was placed in an oil bath preheated to 37 °C and stirred under nitrogen for 5 hours. Three additional doses of Pd (each 0.47 mL, 4.7 μmol, 5 mol%) were added to the reaction mixture after 5, 7, and 8 hours of reaction time in an effort to push the coupling to completion (total Pd loading = 20 mol% after final dose). After the final dose, 1.5 mL of MeCN was also added to maintain solubility of the coupling partners. After 10.5 hours of total reaction time at 37 °C, TLC analysis (60% EtOAc in petrol) indicated total consumption of the aryl iodide (*R_f* = 0.7) and formation of a highly UV active product spot (*R_f* = 0.4). The reaction was cooled to room temperature, diluted with Et₂O (50 mL) and brine (50 mL, saturated). The organic layer was separated and the aqueous layer extracted with Et₂O (2 × 50 mL). The combined organics were washed with brine (50 mL), dried (MgSO₄), and filtered. The solvent was removed under reduced pressure and the residue purified by column chromatography (45% to 60% EtOAc in petrol) to give the desired coupling product **6.17** as a clear oil (48 mg, 74% yield). $[\alpha]_D^{20} = +8.6$ (*c* = 2.0, CHCl₃). IR (*ν*_{max}, film): 3396, 1753, 1507, 1376, 1226, 1040. ¹H NMR (500 MHz, CDCl₃): δ = 1.41 (9H, s, Boc), 2.00, 2.02, 2.06, 2.08 (3H each, s, 4 × OAc), 2.98 (2H, ABX system, *J* = 13.6, 6.0, 5.2, CH₂Ar), 3.71-3.73 (4H, m, CO₂CH₃ and H₅), 4.18 (1H, dd, *J* = 12.3, 2.2, H₆), 4.25-4.31 (2H, m, H_{6'} and CHHCH=), 4.47-4.54 (2H, m, H_α and CHHCH=), 4.64 (1H, d, *J* = 8.0, H₁), 5.01-5.06 (1H, m, NH), 5.05-5.06 (1H, dd, *J* = 9.6, 8.0, H₂), 5.11 (1H, t, *J* = 9.6, H₄), 5.21 (1H, t, *J* = 9.6, H₃), 6.20 (1H, dt, *J* = 16.0, 6.0, CH₂CH=), 6.20-6.30 (1H, br s, OH), 6.69 (1H, d, *J* = 7.9, CH_{Ar}), 6.81 (1H, d, *J* = 16.0,

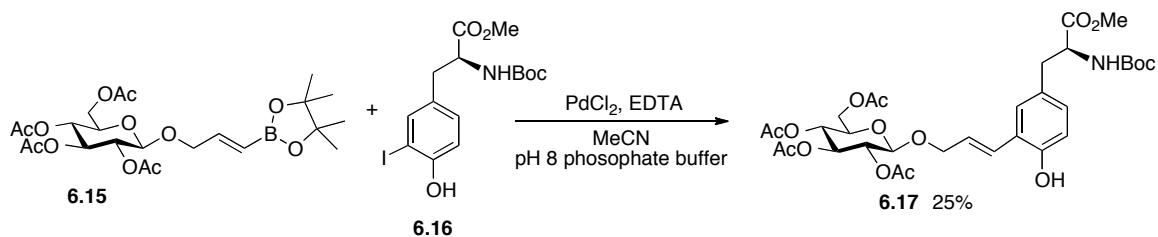
CH=CHAr), 6.87 (1H, d, $J = 7.9$, CH_{Ar}), 7.10 (1H, d, $J = 1.3$, CH_{Ar}). ¹³C NMR (125 MHz, CDCl₃): $\delta = 20.57, 20.59, 20.70$ (2 peaks) (4 $\times$ OAc), 28.3 (Boc), 37.6 (CH₂Ar), 52.2 (CO₂CH₃), 54.6 (C α), 61.9 (C6), 68.4 (C4), 70.1 (CH₂CH=), 71.4, 71.7 (C2 and C5), 72.9 (C3), 80.1 (Boc), 99.5 (C1), 116.1 (CH_{Ar}), 123.5 (C_{Ar}), 125.5 (CH₂CH=), 127.7 (CH=CHAr), 127.8 (C_{Ar}), 128.2 (CH_{Ar}), 129.6 (CH_{Ar}), 152.6, 155.2, 169.5, 169.6, 170.3, 170.9, 172.5 (C_{Ar} and 6 $\times$ C=O). HRMS m/z (ESI⁺): Found 704.2525 (M+Na⁺); C₃₂H₄₃NO₁₅Na requires 704.2530.



The glucose boronic ester derivative **6.15** (92 mg, 0.18 mmol) and *N*-Boc-3-iodotyrosine methyl ester **6.16** (50 mg, 0.12 mmol) were added to a two-neck flask under a stream of nitrogen and dissolved in MeCN (3 mL). To the stirred solution was added pH 8.0 sodium phosphate buffer (3 mL, 68 mM). The palladium solution prepared above (0.12 mL, 1 μ mol, 1 mol%) was added to give a faintly yellow solution. The reaction mixture was heated at 37 °C for 24 hours at which time another 1% Pd dose (0.12 mL) was added. After 9 additional hours at 37 °C, all of the aryl iodide had been consumed (TLC). The reaction was worked up and purified as described above to give cross-coupled product **6.17** (68 mg, 82% yield).

Model Suzuki-Miyaura coupling with EDTA ligand

For use of EDTA as a ligand in aqueous Suzuki-Miyaura cross-coupling, see references 13 and 14.



Preparation of catalyst solution

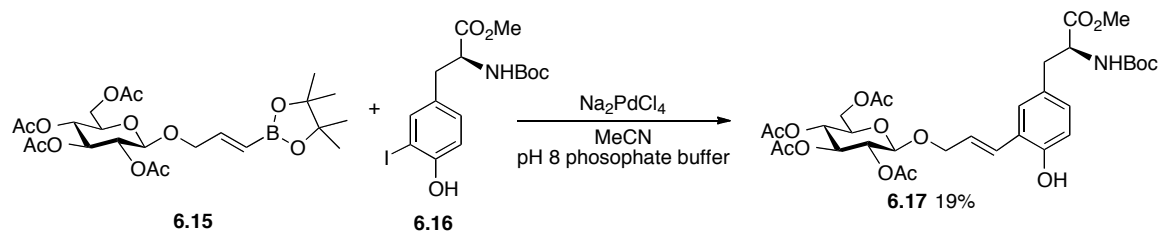
EDTA (15 mg, 0.05 mmol) was added to a 5.0 mL volumetric flask along with deionized H₂O (2 mL) and 0.1M NaOH (1.0 mL, 0.10 mmol). The solution was stirred at 50 °C to dissolve and then PdCl₂ (9 mg, 0.05 mmol) was added to the flask. The mixture was stirred at 70 °C for 45 min at which time a homogenous yellow solution resulted. The solution was diluted to 5.00 mL with distilled water. The final concentration was 0.01M in Pd.

Suzuki-Miyaura Coupling

The glucose boronic ester derivative **6.15** (92 mg, 0.18 mmol) and *N*-Boc-3-iodotyrosine methyl ester **6.16** (50 mg, 0.12 mmol) were dissolved in a 2:1 mixture of MeCN and pH 8.0 phosphate buffer (4 mL) previously degassed by simultaneous sonication and N₂ sparging. The Pd-EDTA solution prepared above was added (0.59 mL, 5.9 μmol, 5%) and the reaction warmed to 37 °C. After one hour an additional 5% of Pd solution was added (0.59 mL) and after a second hour an additional 5% of Pd (0.59 mL) was added along with additional boronic ester (92 mg, 0.18 mmol). Each hour thereafter an additional 5% Pd solution was added until the total palladium loading was 40 mol%. The reaction was stirred a final hour at 37 °C, then cooled to room temperature, and stirred overnight (10 additional hours). The reaction was then diluted with brine (50 mL) and Et₂O (50 mL). The organic layer was separated and the aqueous extracted with Et₂O (2 × 50mL). The combined organics were dried (MgSO₄), filtered, and the solvent removed under reduced pressure. Column chromatography afforded the cross-coupled product **6.17** (20 mg, 25% yield) with spectroscopic properties identical to that shown previously.

Model Suzuki-Miyaura coupling with NaCl ligand

For use of NaCl as a ligand for palladium in aqueous Suzuki-Miyaura cross-coupling, see reference 15.



Preparation of catalyst solution

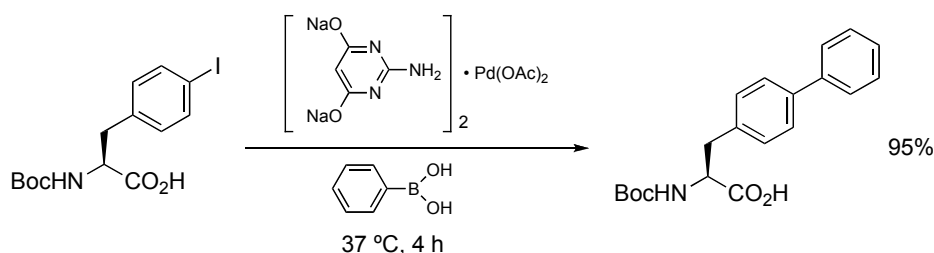
To a glass vial was added PdCl₂ (14 mg, 0.08 mmol), NaCl (18 mg, 0.31 mmol), and deionized H₂O (2.0 mL). The mixture was stirred at 50 °C until a homogenous yellow solution resulted.

Suzuki-Miyaura Coupling (Stoichiometric Pd)

To a flask containing the glucose boronic ester derivative **6.15** (147 mg, 0.29 mmol) and *N*-Boc-3-iodotyrosine methyl ester **6.16** (30 mg, 0.07 mmol) was added MeCN (4 mL) and 68 mM, pH 8.0 phosphate buffer (3 mL) sequentially. All of the catalyst solution prepared above was transferred under a nitrogen atmosphere (1.1 eq Pd). The reaction mixture immediately darkened and a black solid precipitated. The coupling product was observed immediately by TLC. The reaction was stirred at room temperature overnight, with no change observed by TLC. The reaction was diluted with brine (50 mL) and Et₂O (50 mL). The organic layer was separated and the aqueous extracted with Et₂O (2 × 50 mL). The combined organics were dried (MgSO₄), filtered, and the solvent removed under reduced pressure. Column chromatography afforded the cross coupled product **6.17** (9 mg, 19% yield) as a clear residue with spectral properties identical to that shown previously.

Cross-Couplings from Table 6.2

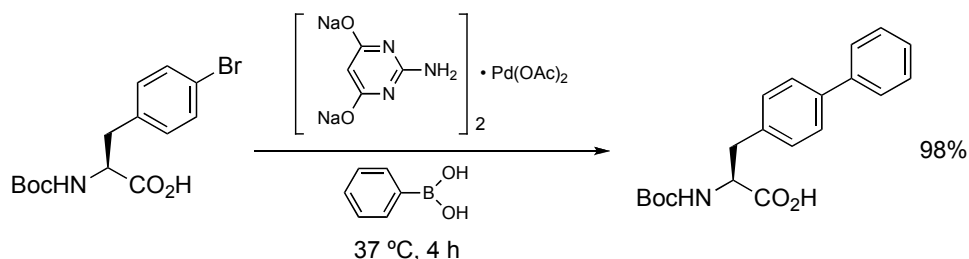
Table 6.2, Entry 1



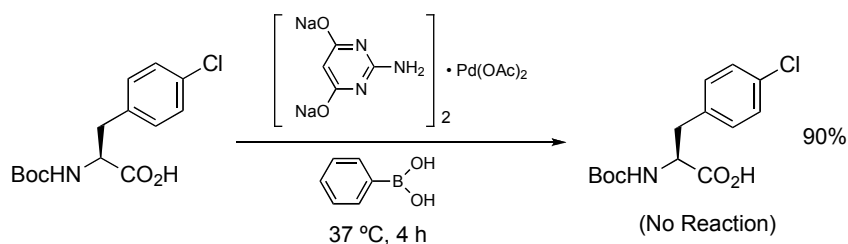
Boc-4-iodo-L-phenylalanine (50 mg, 0.13 mmol), phenyl boronic acid (23 mg, 0.19 mmol), and K₂HPO₄ (113 mg, 0.65 mmol) were added to a 10 mL glass vial. The solids were dissolved in deionized H₂O (3.0 mL) by stirring for 2 minutes at 37 °C. Formic acid (0.25 μL in 10 μL of H₂O, 6.5 μmol) was added by micropipette, followed by 130 μL of the 0.01M palladium-pyrimidine solution prepared above (1.3 μmol). The vial was capped and the reaction was stirred vigorously at 37 °C. No effort was made to exclude

oxygen. After 4 hours, the reaction was cooled to room temperature. The reaction was diluted with 1M HCl (50 mL) and then extracted with EtOAc (3 × 50 mL). The combined organics were washed with brine (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (15% MeOH in CH₂Cl₂). The product was isolated as a thick, clear oil (42 mg, 95% yield). $[\alpha]_D^{20} = +23.8$ (c = 3.1, MeOH). IR ($\nu_{\max}$, film): 3332, 2978, 1909, 1697, 1601, 1487, 1408, 1368, 1252, 1166, 1054. ¹H NMR (400 MHz, DMSO): δ = 1.32 (9H, s, Boc), 2.96 (1H, dd, J = 13.1, 8.4, CHHAr), 3.17 (1H, app d, J = 10.4, CHHAr), 4.17 (1H, app s, H α), 6.67 (1H, br s, NH), 7.30-7.34 (3H, m, CHAr), 7.42 (2H, t, J = 7.6, CHAr), 7.53 (2H, d, J = 8.1, CHAr), 7.60 (2H, d, J = 7.6, CHAr). ¹³C NMR (100 MHz, DMSO): δ = 28.2 (Boc), 36.8 (CH₂Ar), 55.7 (C α), 77.8 (Boc), 126.3, 126.5, 127.2, 128.9, 130.0, 134.1, 138.0, 140.2, 155.2, 175.0 (includes 8 × C_{Ar} and 2 × C=O). HRMS m/z (ESI⁺): Found 364.1515 (M+Na⁺); C₂₀H₂₃NO₄Na requires 364.1519.

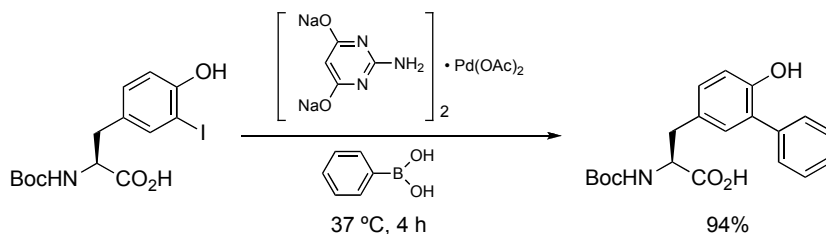
Table 6.2, Entry 2



Boc-4-bromo-L-phenylalanine (250 mg, 0.73 mmol), phenyl boronic acid (133 mg, 1.09 mmol), and K₂HPO₄ (636 mg, 3.65 mmol) were added to a 25 mL round bottom flask. The solids were dissolved in H₂O (15 mL) by stirring at 37 °C for 5 minutes. Formic acid (1.4 μ L in 50 μ L of H₂O, 0.04 mmol) and a 730 μ L aliquot of the 0.01M palladium catalyst solution prepared above (7.3 μ mol) were added sequentially to the reaction. The flask was stoppered and stirred vigorously at 37 °C for 4 hours. After this time, the reaction was diluted with 50 mL each of 1M HCl and EtOAc. The organic layer was separated and the aqueous layer extracted with EtOAc (2 × 50 mL). The combined organic layers were washed with brine (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (15% MeOH in CH₂Cl₂), affording 243 mg of product (98% yield). Spectroscopic and analytical data was identical to that reported above.

Table 6.2, Entry 3

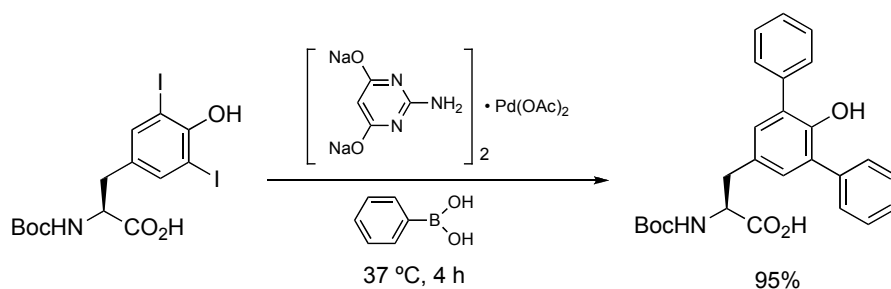
Boc-4-chloro-L-phenylalanine (50 mg, 0.17 mmol), phenyl boronic acid (31 mg, 0.25 mmol), and K_2HPO_4 (148 mg, 0.85 mmol) were added to a 10 mL glass vial. The solids were dissolved in H_2O (3 mL) by stirring at 37 °C for 2 minutes. Formic acid (0.32 μL in 10 μL of H_2O) was added by micropipette, followed by 170 μL of the 0.01M palladium solution prepared above (1.7 μmol). The reaction was capped and stirred at 37 °C for 4 hours. After this time, the reaction was acidified with 50 mL of 1M HCl and then extracted with EtOAc (3 $\times$ 50 mL). The combined organics were washed with brine (100 mL), dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography to afford 45 mg (90% yield) of recovered Boc-4-chloro-L-phenylalanine. Spectroscopic data was consistent with the commercially available starting material.

Table 6.2, Entry 4

Boc-3-iodotyrosine (56 mg, 0.14 mmol), phenyl boronic acid (25 mg, 0.21 mmol), and K_2HPO_4 (120 mg, 0.69 mmol) were added to a 10 mL glass vial and dissolved in H_2O by stirring for 2 minutes at 37 °C. Formic acid (0.26 μL in 10 μL of H_2O) was added by micropipette, followed by 140 μL of the 0.01 M palladium solution prepared above (1.4 μmol). The resulting pale yellow solution was capped and stirred at 37 °C for 4 hours. After this time, the reaction was quenched with 50 mL of 1M HCl and then diluted with 50 mL of EtOAc. The organic layer was separated and the aqueous fraction was extracted with EtOAc (2 $\times$ 50 mL). The combined organic layers were dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (15% MeOH in CH_2Cl_2) to afford the cross-coupled product as a thick,

clear oil that dried to a glassy foam under high vacuum (46 mg, 94% yield). $[\alpha]_D^{20} = +25.0$ ($c = 0.1$, MeOH). IR ($\nu_{\max}$, film): 3356, 2978, 2930, 1684, 1611, 1509, 1490, 1417, 1386, 1271, 1164, 1054. ^1H NMR (500 MHz, DMSO, 343K): $\delta = 1.33$ (9H, s, Boc), 2.85 (1H, dd, $J = 13.9, 7.9$, CHHAr), 3.04 (1H, dd, $J = 13.9, 4.7$, CHHAr), 4.07 (1H, m, H α), 6.26 (1H, br s, NH), 6.87 (1H, d, $J = 8.2$, CHAr), 7.00 (1H, dd, $J = 8.2, 1.9$, CHAr), 7.11 (1H, d, $J = 1.9$, CHAr), 7.27 (1H, t, $J = 7.3$, CHAr), 7.37 (2H, t, $J = 7.3$, CHAr), 7.56 (2H, d, $J = 7.3$, CHAr). ^{13}C NMR (125 MHz, DMSO, 343K): $\delta = 27.9$ (Boc), 36.3 (CH₂Ar), 55.5 (C α), 77.5 (Boc), 115.6, 125.9, 126.9, 127.4, 128.7, 128.8, 130.9, 133.7, 138.7, 152.4 (10 $\times$ CAr), 154.6, 174.3 (2 $\times$ C=O). HRMS m/z (ESI+): Found 380.1471 (M+Na⁺); C₂₀H₂₃NO₅Na requires 380.1468.

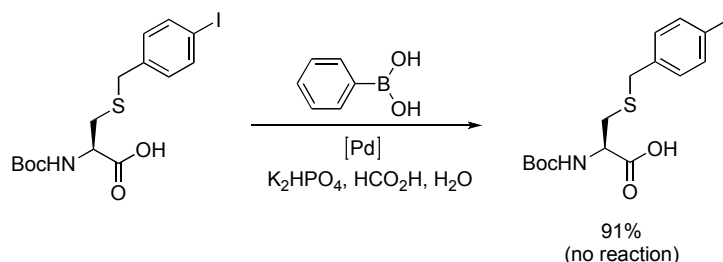
Table 6.2, Entry 5



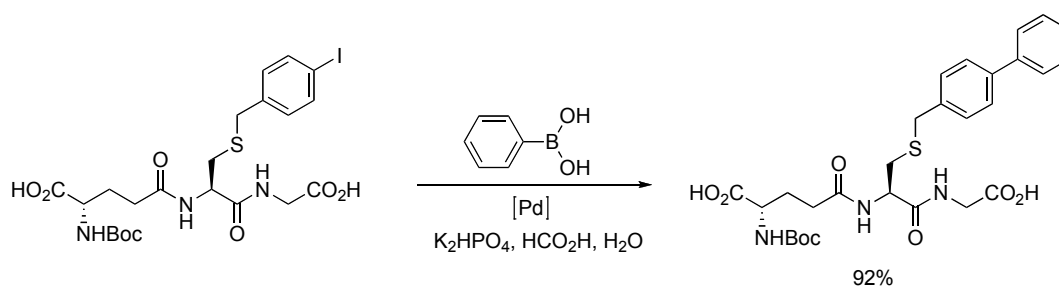
N-Boc-3,5-di-iodo tyrosine (50 mg, 0.094 mmol), phenyl boronic acid (34 mg, 0.28 mmol), and K₂HPO₄ (164 mg, 0.94 mmol) were added to a 6 mL glass vial and dissolved in H₂O (2 mL) at 37 °C. Formic acid (0.35 μL in 10 μL of H₂O) was added by micropipette, followed by 188 μL of the 0.01 M palladium solution prepared above (1.8 μmol). The pale yellow solution was stirred at 37 °C for 2 hours after which time another dose of formic acid (0.35 μL in 10 μL of H₂O) and the palladium catalyst was added (188 μL , 1.8 μmol). The reaction was stirred a final 2 hours and then quenched with 1M HCl (50 mL) and extracted with EtOAc (3 $\times$ 50 mL). The combined organic layers were washed with brine (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (15% MeOH in CH₂Cl₂) to afford the product as a white solid (39 mg, 96%). m.p. = 186 °C (decomposes). $[\alpha]_D^{20} = +36.0$ ($c = 0.25$, MeOH). IR ($\nu_{\max}$, KBr): 3583, 3539, 2976, 2927, 2360, 1684, 1600, 1497, 1422, 1228, 1165, 1122, 1055, 1027. ^1H NMR (400 MHz, MeOD, 343K): $\delta = 1.37$ (9H, s, Boc), 2.95 (1H, dd, $J = 13.7, 7.6$, CHHAr), 3.21 (1H, dd, $J = 13.7, 4.9$, CHHAr), 4.30 (1H, dd, $J = 7.6, 4.9$, H α), 7.11 (2H, s, CHAr), 7.33 (2H, tt, J

= 7.5, 1.3, CH_{Ar}), 7.42 (4H, t, *J* = 7.5, CH_{Ar}), 7.57 (4H, dd, *J* = 7.5, 1.3, CH_{Ar}). ¹³C NMR (50 MHz, MeOD, data for major rotamer): δ = 28.8 (Boc), 38.0 (CH₂Ar), 56.5 (Cα), 80.6 (Boc), 128.2, 129.4 (CH_{Ar}), 130.5 (4°_{Ar}), 130.8 (CH_{Ar}), 131.7 (4°_{Ar}), 132.2 (CH_{Ar}), 140.2, 150.3, 157.7, 176.0 (2 × 4°_{Ar} and 2 × C=O). HRMS *m/z* (ESI⁺): Found 456.1780 (M+Na⁺); C₂₆H₂₇NO₅Na requires 456.1781.

Table 6.2, Entry 6



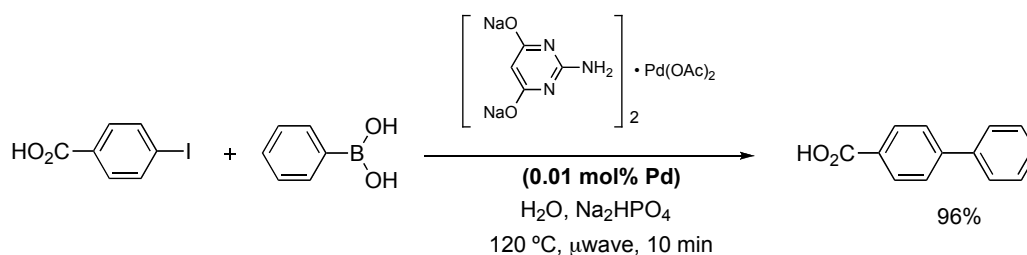
The aryl iodide (87 mg, 0.20 mmol), phenylboronic acid (36 mg, 0.30 mmol), and K₂HPO₄ (174 mg, 1.0 mmol) were added to a 10 mL glass vial and dissolved in 3 mL H₂O at 37 °C. A solution of formic acid was added (10 μL of 1 μM solution), followed by 200 μL of the 0.01 M palladium catalyst solution. The vial was capped and stirred at 37 °C for 4 hours. No reaction was observed by TLC or ESI-MS. An additional dose of the palladium catalyst (200 μL of the 0.01M solution) and phenylboronic acid (36 mg) were added and the reaction stirred 30 minutes at 37 °C. Additional doses of the palladium catalyst solution (400 μL each) were added every 20 minutes thereafter until the total palladium loading was 10 mol%. No reaction was observed throughout this sequence. After 6 hours of reaction time, the solution was poured into 100 mL 1M HCl. The organics were extracted with EtOAc (3 × 50 mL), washed with brine (100 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (20% MeOH in CH₂Cl₂), providing 79 mg of recovered aryl iodide (91% recovery). Spectroscopic data was identical to that reported for the starting material.

Table 6.2, Entry 7

The glutathione derived aryl iodide (40 mg, 0.064 mmol), phenyl boronic acid (12 mg, 0.096 mmol), and K_2HPO_4 (56 mg, 0.32 mmol) were added to a 5 mL glass vial and dissolved in H_2O (2.0 mL). An aqueous solution of formic acid was prepared by dissolving 12 μL in 1.00 mL H_2O ; 10 μL of this solution was added to the reaction mixture (0.003 mmol). An aliquot of the palladium solution prepared above (64 μL of 0.01 M, 1 mol%) was added and the vial capped and stirred at 37 $^{\circ}C$. Additional doses of phenyl boronic acid (12 mg, 0.096 mmol), K_2HPO_4 (56 mg, 0.32 mmol), formic acid (10 μL of the solution prepared above), and palladium (64 μL of the solution prepared above) were added each hour for three hours. The final palladium loading was 4 mol%. After 6 hours of reaction time, the reaction was quenched by pouring the solution into 75 mL of 1M HCl. The product was extracted with EtOAc (4 $\times$ 50 mL). The combined organic layer was washed with brine (100 mL), dried ($MgSO_4$), filtered, and evaporated under reduced pressure. The product was purified by column chromatography (3:2:1 EtOAc: i PrOH: H_2O) to afford the cross-coupled product as a white, glassy solid (34 mg, 92% yield). $[\alpha]_D^{20} = -36.7$ ($c = 0.65$, DMSO). IR (ν_{max} , KBr): 3299, 2977, 2360, 2388, 1536, 1591, 1367, 1250, 1166, 1027. 1H NMR (400 MHz, D_2O with 4eq $NaHPO_4$): $\delta = 1.29$ (9H, s, Boc), 1.67-1.80 (1H, m, $H\beta_{Glu}$), 1.91-2.01 (1H, m, $H\beta'_{Glu}$), 2.23 (2H, br app s, $H\gamma_{Glu}$), 2.70-2.76 (1H, m, $H\beta_{Cys}$), 2.90-2.93 (1H, m, $H\beta'_{Cys}$), 3.65 (2H, AB system, $J = 17.3$, $H\alpha_{Gly}$), 3.74 (2H, s, SCH_2Ar), 3.76-3.83 (1H, m, $H\alpha_{Glu}$), 4.39 (1H, m, $H\alpha_{Cys}$), 7.33 (1H, t, $J = 7.6$, CH_{Ar}), 7.37 (2H, d, $J = 8.1$, CH_{Ar}), 7.43 (2H, t, $J = 7.6$, CH_{Ar}), 7.56 (2H, d, $J = 8.1$, CH_{Ar}), 7.60 (2H, d, $J = 7.6$, CH_{Ar}). ^{13}C NMR (100 MHz, D_2O with 4eq $NaHPO_4$): $\delta = 27.8$ (Contains Boc and $C\beta_{Glu}$), 31.7 ($C\gamma_{Glu}$), 32.9 ($C\beta_{Cys}$), 35.5 (SCH_2Ar), 43.3 ($C\alpha_{Gly}$), 52.8 ($C\alpha_{Cys}$), 55.2 ($C\alpha_{Glu}$), 80.5 (Boc), 126.7, 127.0, 127.4, 128.8, 129.4, 137.1, 139.3, 140.0 (8 $\times$ C_{Ar}), 157.0, 171.8, 175.4, 176.2, 178.9 (5 $\times$ $C=O$). HRMS m/z (ESI $^{+}$): Found 572.2076 (M-H) $^{-}$; $C_{28}H_{34}N_3O_8S$ requires 572.2072.

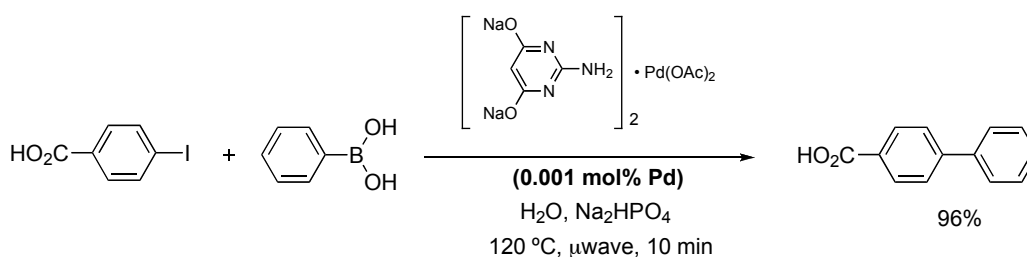
Cross-Couplings from Table 6.4:

Table 6.4, Entry 1



4-Iodobenzoic acid (100 mg, 0.40 mmol), phenylboronic acid (59 mg, 0.48 mmol), and Na_2HPO_4 (284 mg, 2.00 mmol) were all added as solids to a 10 mL microwave reaction vessel. The solids were dissolved in 4 mL H_2O by stirring at 50°C for 5 minutes. The palladium catalyst solution was diluted to 0.001M with H_2O . A 40 μL aliquot of this palladium solution (0.01 mol%) was added to the reaction solution and heated immediately to 120°C under microwave radiation for 10 minutes with stirring. The product crystallized from the reaction mixture. The crystals were isolated by filtration and washed with 20 mL H_2O . The resulting product was transferred to a separatory funnel using 50 mL 1M HCl and 50 mL EtOAc . The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure. The isolated material was pure and did not require further purification (77 mg, 96% yield). Spectroscopic data was consistent with that previously reported.¹⁶ m.p. = $228\text{--}229^\circ\text{C}$. IR (ν_{max} , KBr): 3000, 1675, 1606, 1421, 1317, 1287, 1192, 1121, 1006, 940, 862, 813, 751, 695. ^1H NMR (400 MHz, DMSO): δ = 7.39 (1H, t, J = 7.5), 7.47 (2H, t, J = 7.5), 7.70 (2H, d, J = 7.5), 7.76 (2H, d, J = 8.3), 8.05 (2H, d, J = 8.3). ^{13}C NMR (100 MHz, DMSO): δ = 126.8, 127.0, 128.3, 129.1, 130.1 ($5 \times \text{CH}_{\text{Ar}}$), 130.4, 139.2, 144.1 ($3 \times 4^\circ \text{Ar}$), 167.5 ($\text{C}=\text{O}$). LRMS m/z (ESI): 197.0 $[\text{M}-\text{H}]^-$.

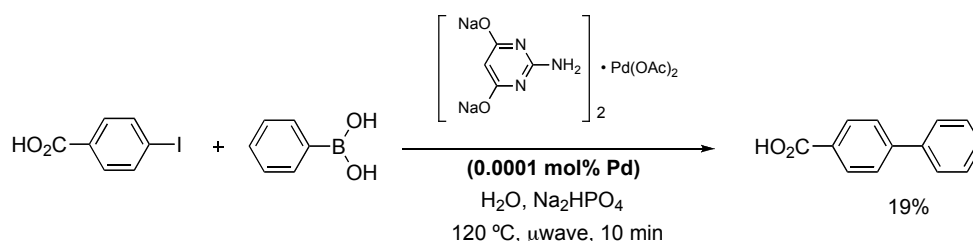
Table 6.4, Entry 2



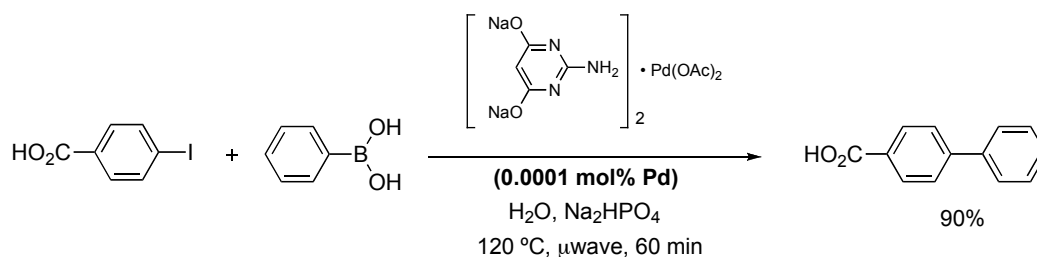
4-Iodobenzoic acid (100 mg, 0.40 mmol), phenylboronic acid (59 mg, 0.48 mmol), and Na_2HPO_4 (284 mg, 2.00 mmol) were all added as solids to a 10 mL microwave reaction

vessel. The solids were dissolved in 4 mL H₂O by stirring at 50 °C for 5 minutes. The palladium catalyst solution was diluted to 0.001M with H₂O. A 4 µL aliquot of this palladium solution (0.001 mol%) was added to the reaction solution by micropipette and heated immediately to 120 °C under microwave radiation for 10 minutes with stirring. The product crystallized from the reaction mixture. The crystals were isolated by filtration and washed with 20 mL H₂O. The resulting product was transferred to a separatory funnel using 50 mL 1M HCl and 50 mL EtOAc. The organic layer was separated, dried (MgSO₄), filtered, and concentrated under reduced pressure. The isolated material was pure and did not require further purification (77 mg, 96% yield). Spectroscopic data was identical to that reported above.

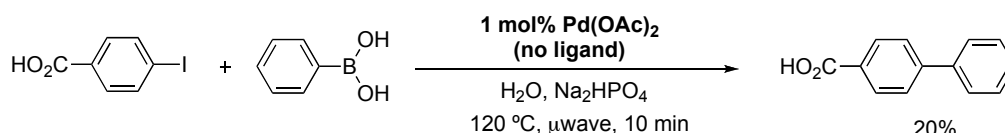
Table 6.4, Entry 3



4-Iodobenzoic acid (100 mg, 0.40 mmol), phenylboronic acid (59 mg, 0.48 mmol), and Na₂HPO₄ (284 mg, 2.00 mmol) were all added as solids to a 10 mL microwave reaction vessel. The solids were dissolved in 4 mL H₂O by stirring at 50 °C for 5 minutes. The palladium catalyst solution was diluted to 0.001M with H₂O. A 0.4 µL aliquot of this palladium solution (0.0001 mol%) was added to the reaction solution by micropipette and heated immediately to 120 °C under microwave radiation for 10 minutes with stirring. The product crystallized from the reaction mixture. The crystals were isolated by filtration and washed with 20 mL H₂O. The resulting product was transferred to a separatory funnel using 50 mL 1M HCl and 50 mL EtOAc. The organic layer was separated, dried (MgSO₄), filtered, and concentrated under reduced pressure. The isolated material was pure and did not require further purification (15 mg, 19% yield). Spectroscopic data was identical to that reported above. For this loading and yield, (0.4 µL of 0.001M Pd = 0.0004 µmol Pd = 0.0000004 mmol; 15.0 mg product = 0.0757 mmol) 189,184 turnovers are required. At 10 minutes of reaction time, this is an average of 315 turnovers / second. To determine whether the catalyst can complete the reaction at this loading (1 ppm Pd) the reaction was run again for 1 hour under otherwise identical conditions (see next).

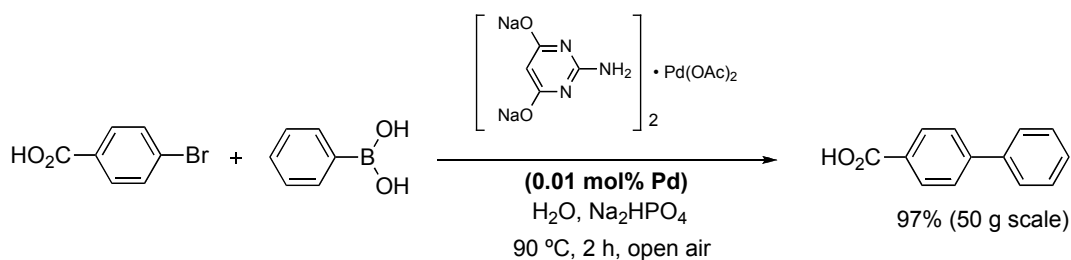
Table 6.4, Entry 4

4-Iodobenzoic acid (100 mg, 0.40 mmol), phenylboronic acid (59 mg, 0.48 mmol), and Na_2HPO_4 (284 mg, 2.00 mmol) were all added as solids to a 10 mL microwave reaction vessel. The solids were dissolved in 4 mL H_2O by stirring at $50\text{ }^\circ\text{C}$ for 5 minutes. The palladium catalyst solution was diluted to 0.001M with H_2O . A 0.4 μL aliquot of this palladium solution (0.0001 mol%) was added to the reaction solution by micropipette and heated immediately to $120\text{ }^\circ\text{C}$ under microwave radiation for 60 minutes with stirring. The product crystallized from the reaction mixture. The crystals were isolated by filtration and washed with 20 mL H_2O . The resulting product was transferred to a separatory funnel using 50 mL 1M HCl and 50 mL EtOAc . The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure. The isolated material was pure and did not require further purification (72 mg, 90% yield). Spectroscopic data was identical to that reported above. This experiment indicates that the catalyst is active for nearly 10^6 turnovers for this cross-coupling.

Control: Cross-Coupling without pyrimidine ligand

4-Iodobenzoic acid (76 mg, 0.31 mmol), phenylboronic acid (45 mg, 0.37 mmol), and Na_2HPO_4 (220 mg, 1.55 mmol) were added to a 10 mL microwave reaction tube and dissolved in 2 mL H_2O at $60\text{ }^\circ\text{C}$. $\text{Pd}(\text{OAc})_2$ (0.7 mg, 3.1 μmol) was weighed on a microbalance and then added to the reaction tube as a solid. The reaction was then heated to $120\text{ }^\circ\text{C}$ under microwave irradiation for 10 minutes with stirring. Upon cooling the product crystallized. The white solid was isolated by filtration and washed with 15 mL H_2O . The solid was then transferred to a separatory funnel with 50 mL EtOAc and 50 mL 1M HCl . The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure to give the cross-coupled product as white crystals (12 mg, 20% yield). Spectroscopic data was identical to that reported above.

Table 6.4, Entry 5

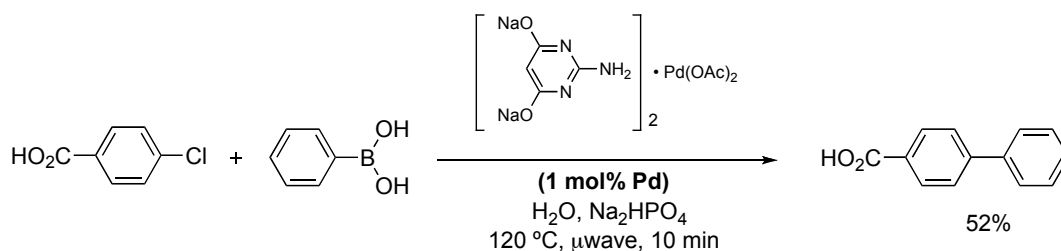


Na₂HPO₄ (176.6 g, 1.244 mol) was added to a 2L beaker along with a large stir bar. H₂O (1 L, deionized) was poured into the beaker and the mixture stirred to give a clear solution. When all phosphate salts were dissolved, the solution was heated to 50 °C. The internal temperature was monitored throughout the course of the reaction. 4-bromobenzoic acid (50.00 g, 248.7 mmol) was added in several portions to the stirred solution. When all of the 4-bromobenzoic acid was dissolved, phenylboronic acid (30.93 g, 253.7 mmol) was added in several portions. The mixture was then heated to 90 °C with rapid stirring. Once all material was dissolved, 2.49 mL of the 0.01 M palladium catalyst solution was added (0.025 mmol, 0.01 mol%). The reaction was stirred rapidly at 90 °C, open to air. Within one minute of reaction time, product began to crystallize from the reaction mixture. Efficient stirring was ensured to prevent the reaction from solidifying. The reaction was stirred for 2 hours to ensure completion. After this time the reaction was cooled to room temperature with stirring. Once the reaction was cooled to room temperature, the solid was isolated by filtration and dried on the filter with suction. A nitrogen stream was also used to assist drying. The material isolated (the sodium carboxylate of the cross-coupled product) was pure by ¹H NMR. To exchange the sodium salt for the free acid, the solid was returned to a 2 L beaker and suspended in 500 mL deionized H₂O. 10 M HCl (100 mL) was poured into the stirred suspension. The mixture was stirred at room temperature for 2 hours. The crystals were then isolated by filtration and dried under suction and then collected and dried under high vacuum with periodic heating (47.82 g, 97% yield).

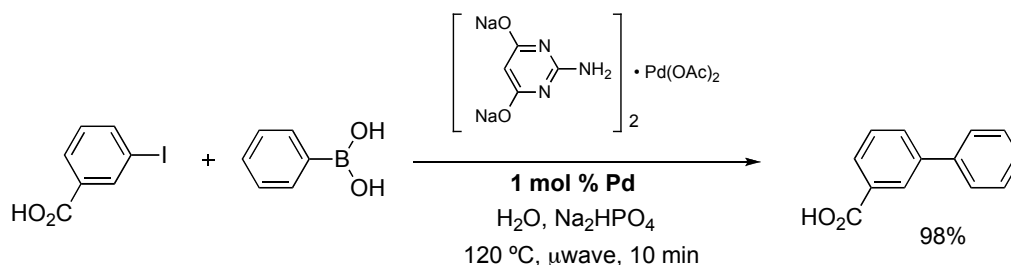
¹H NMR for crude carboxylate:

¹H NMR (400 MHz, DMSO): δ = 7.36 (1H, t, *J* = 7.5), 7.46 (2H, t, *J* = 7.5), 7.59 (2H, d, *J* = 7.5), 7.68 (2H, d, *J* = 8.2), 7.95 (2H, d, *J* = 8.2).

Spectroscopic data for free acid was identical to that reported above.

Table 6.4, Entry 6

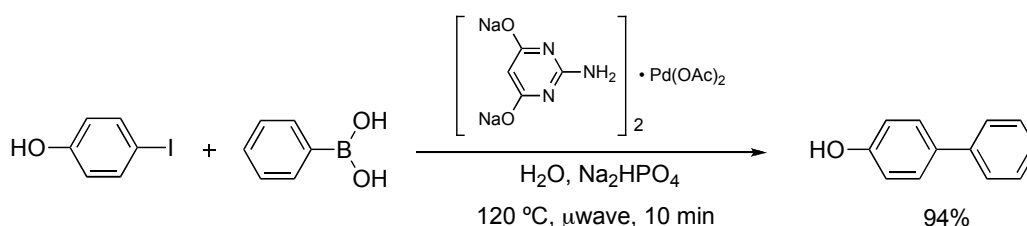
4-chlorobenzoic acid (100 mg, 0.64 mmol), phenylboronic acid (93 mg, 0.77 mmol), and Na_2HPO_4 (455 mg, 3.20 mmol) were all added as solids to a 10 mL microwave reaction vessel. H_2O (4 mL) was added and the mixture was stirred at $50\text{ }^\circ\text{C}$ to dissolve. A 640 μL aliquot of 0.01M palladium catalyst was added and the reaction was immediately heated to $120\text{ }^\circ\text{C}$ under microwave radiation for 10 minutes. Upon cooling, crystals were observed. The crystals were isolated by filtration and then stirred in 5M HCl for 30 minutes. The product crystals were again isolated by filtration and dried under vacuum (66 mg, 52% yield). Spectroscopic data was identical to that reported above from the aryl iodide and aryl bromide.

Table 6.4, Entry 7

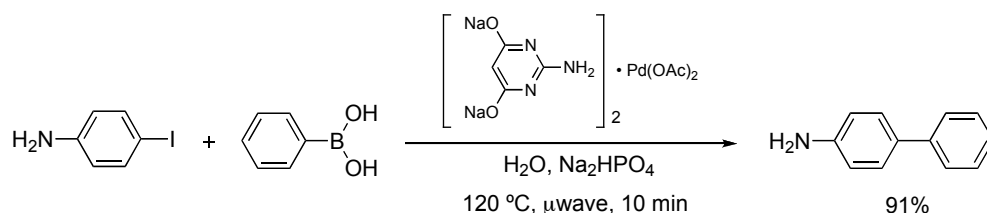
3-Iodobenzoic acid (100 mg, 0.40 mmol), phenylboronic acid (59 mg, 0.48 mmol), and Na_2HPO_4 (284 mg, 2.00 mmol) were all added as solids to a 10 mL microwave reaction vessel. H_2O (2 mL) was added and the mixture stirred at $50\text{ }^\circ\text{C}$ for 5 minutes to dissolve the reactants. An aliquot of the palladium catalyst solution (400 μL of the 0.01M Pd solution, 1 mol%) was added to the mixture and the reaction was immediately heated to $120\text{ }^\circ\text{C}$ under microwave irradiation with stirring for 10 minutes. After cooling to room temperature, the reaction was partitioned between 50 mL 1M HCl and 50 mL EtOAc. The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting solid was purified by column chromatography (10% MeOH in CH_2Cl_2) to afford the product as white crystals (78 mg, 98% yield). Spectroscopic data

was consistent with that previously reported.¹⁷ m.p. = 160-161 °C. IR (ν_{max} , KBr): 3051, 1681, 1584, 1482, 1456, 1319, 1258, 930, 743, 698. ¹H NMR (400 MHz, DMSO): δ = 7.42 (1H, tt, J = 7.3, 1.1), 7.50 (2H, t, J = 7.3), 7.58 (1H, t, J = 7.7), 7.66 (2H, m), 7.87 (1H, dq, J = 7.7, 1.4), 8.14 (1H, dt, J = 7.7, 1.4), 8.40 (1H, t, J = 1.4). ¹³C NMR (100 MHz, DMSO): δ = 127.2, 127.8, 128.85, 128.93, 128.96, 128.99 (6 $\times$ CH_{Ar}), 129.8 (4° Ar), 132.4 (CH_{Ar}), 139.9, 141.6 (2 $\times$ CH_{Ar}), 172.3 (C=O). LRMS m/z (ESI⁺): 197.0 [M-H]⁺.

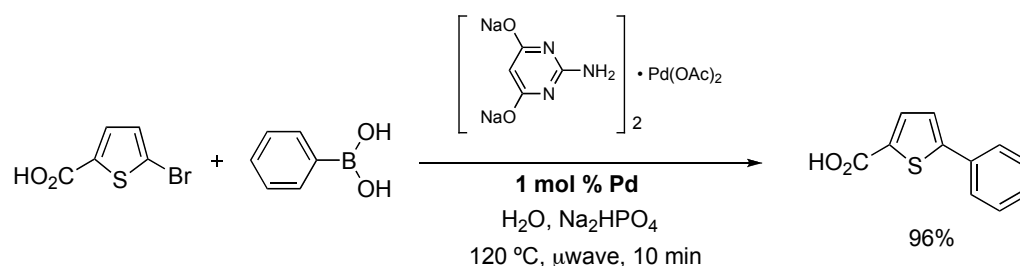
Table 6.4, Entry 8



4-Iodophenol (100 mg, 0.45 mmol), phenylboronic acid (67 mg, 0.55 mmol), and Na_2HPO_4 (319 mg, 2.25 mmol) were all added as solids to a 10 mL microwave reaction tube. H_2O (3.0 mL) was added and the mixture stirred for 5 minutes at room temperature to dissolve the bulk of the material. (Note: 4-Iodophenol has limited solubility in this mixture at room temperature). An aliquot of the 0.01M palladium catalyst solution (0.45 mL, 0.0045 mmol, 1 mol%) was added to the reaction mixture and immediately sealed and placed in the microwave reactor. The reaction was heated at 120 °C with stirring for 10 minutes. Upon cooling, the product crystallized from the reaction mixture. The reaction tube was cooled on ice to complete the crystallization and the product was isolated by filtration and dried under vacuum. (72 mg, 94% yield). Spectroscopic data was consistent with that previously reported.¹⁸ m.p. = 161-162 °C. IR (ν_{max} , KBr): 3375, 2361, 1598, 1488, 1424, 1262, 1113, 832, 757, 686. ¹H NMR (400 MHz, DMSO): δ = 6.85 (2H, d, J = 8.5), 7.27 (1H, t, J = 7.4), 7.40 (2H, t, J = 7.4), 7.48 (2H, d, J = 8.5), 7.57 (2H, d, J = 8.5), 9.58 (1H, br s, OH). ¹³C NMR (100 MHz, DMSO): δ = 115.8, 126.0, 126.4, 127.8, 128.8 (5 $\times$ CH_{Ar}), 131.0, 140.3, 157.2 (3 $\times$ 4° Ar). LRMS m/z (ESI⁺): 169.1 [M-H]⁺.

Table 6.4, Entry 9

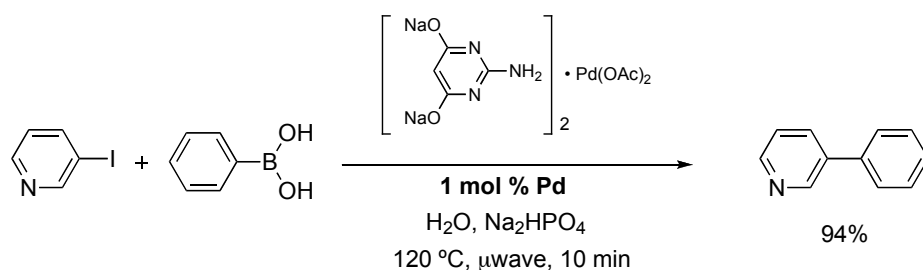
4-Iodoaniline (100 mg, 0.46 mmol), phenylboronic acid (67 mg, 0.55 mmol), and Na_2HPO_4 (327 mg, 2.30 mmol) were all added as solids to a 10 mL microwave reaction tube. H_2O (3.0 mL) was added and the mixture stirred at room temperature to partially dissolve the reactants. (Note: 4-iodoaniline is not soluble under these conditions at room temperature). An aliquot of the 0.01 M palladium catalyst solution (0.460 mL, 0.0046 mmol, 1 mol%) was added to the reaction and the tube was immediately sealed and placed in the microwave reactor. The reaction was heated to 120 °C and stirred for 10 minutes. After the reaction, the product was as a dark liquid layer that crystallized upon cooling on ice. The product was isolated by filtration and dried under vacuum (70 mg, 91% yield). Spectroscopic data was consistent with that previously reported.¹⁸ m.p. = 39–40 °C. IR (ν_{max} , KBr): 3423, 3392, 3299, 3200, 1899, 1618, 1520, 1487, 1419, 1266, 832, 762, 694. ^1H NMR (400 MHz, DMSO): δ = 5.23 (2H, br s, NH_2), 6.67 (2H, d, J = 8.3), 7.20 (1H, t, J = 7.2), 7.34–7.38 (4H, m), 7.53 (2H, d, J = 7.6). ^{13}C NMR (100 MHz, DMSO): δ = 114.3, 125.4, 125.7, 127.2 ($4 \times \text{CH}_{\text{Ar}}$), 127.5 (4°Ar), 128.8 (CH_{Ar}), 140.7, 148.4 (4°Ar). LRMS m/z (ESI^+): 170.1 $[\text{M}+\text{H}]^+$.

Table 6.4, Entry 10

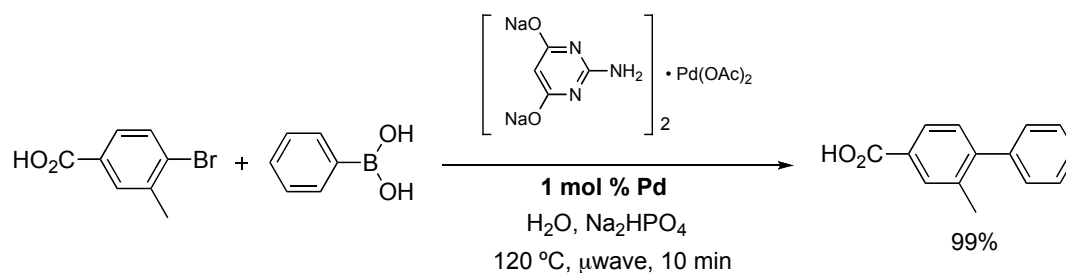
5-bromothiophene-2-carboxylic acid (100 mg, 0.48 mmol), phenylboronic acid (71 mg, 0.58 mmol), and Na_2HPO_4 (341 mg, 2.40 mmol) were all added as solids to a 10 mL microwave reaction vessel. H_2O (2 mL) was added and the mixture was stirred at 50 °C to dissolve all material. A aliquot of the palladium catalyst was added to the reaction mixture (0.460 mL of a 0.01M solution; 1 mol% Pd) and the reaction was immediately heated to 120 °C under microwave irradiation for 10 minutes with stirring. After cooling

to room temperature, the reaction was partitioned between 50 mL 1M HCl and 50 mL EtOAc. The organic layer was separated, dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting solid was purified by column chromatography (10% MeOH in CH₂Cl₂), affording the product as white crystals (95 mg, 96% yield).¹⁹ m.p. = 179-180 °C. IR (ν_{max} , KBr): 3000, 1656, 1539, 1461, 1348, 1289, 1116, 1056, 928, 909, 819, 744, 677. ¹H NMR (500 MHz, DMSO): δ = 7.36 (1H, t, J = 7.5), 7.44 (2H, t, J = 7.5), 7.51 (1H, d, J = 3.8), 7.64 (1H, d, J = 3.8), 7.71 (2H, d, J = 7.5). ¹³C NMR (125 MHz, DMSO): δ = 124.3, 125.7, 128.6, 129.3 (4 $\times$ CH_{Ar}), 133.1 (4° Ar), 133.2 (CH_{Ar}), 136.6, 148.4 (2 $\times$ 4° Ar), 163.8 (C=O). LRMS m/z (ESI⁻): 203.0 [M-H]⁻.

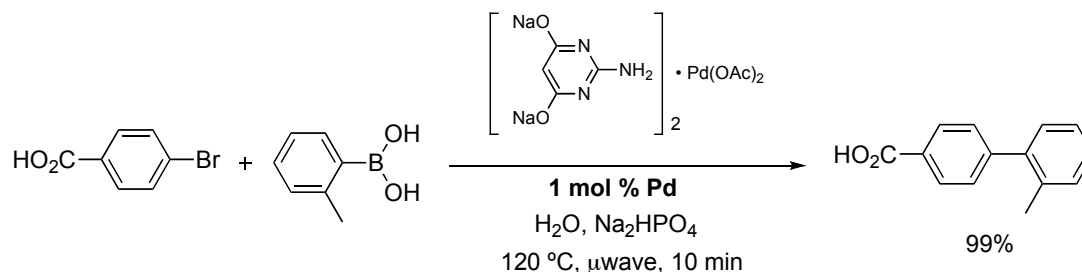
Table 6.4, Entry 11



3-Iodopyridine (100 mg, 0.49 mmol), phenylboronic acid (71 mg, 0.59 mmol), and Na₂HPO₄ (348 mg, 2.45 mmol) were added to a 10 mL microwave reaction vessel. The solids were suspended in 3 mL H₂O and stirred for 5 minutes at 50 °C to partially dissolve the reactants. An aliquot of the 0.01M palladium solution (490 μ L, 1 mol%) was added to the reaction mixture and the tube was immediately heated to 120 °C and stirred for 10 minutes. After cooling to room temperature, the reaction was diluted with 50 mL EtOAc and 50 mL H₂O. The organic layer was separated, dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc), affording the cross-coupled product as a clear oil (71 mg, 94% yield). Spectroscopic data was consistent with that previously reported.²⁰ IR (ν_{max} , film): 3055, 3030, 2360, 2341, 1581, 1499, 1450, 1407, 1024, 1005, 812, 754, 697. ¹H NMR (400 MHz, CDCl₃): δ = 7.34 (1H, dt, J = 7.8, 3.8), 7.38-7.41 (1H, m), 7.45-7.49 (2H, m), 7.56-7.58 (2H, m), 7.83-7.87 (1H, m), 8.59 (1H, app. s), 8.86 (1H, s). ¹³C NMR (100 MHz, CDCl₃): δ = 123.4, 127.0, 128.0, 128.9, 134.2 (5 $\times$ CH_{Ar}), 136.5, 137.7 (2 $\times$ 4° Ar), 148.2, 148.3 (2 $\times$ CH_{Ar}). LRMS m/z (ESI⁺): 156.1 [M+H]⁺.

Table 6.4, Entry 12

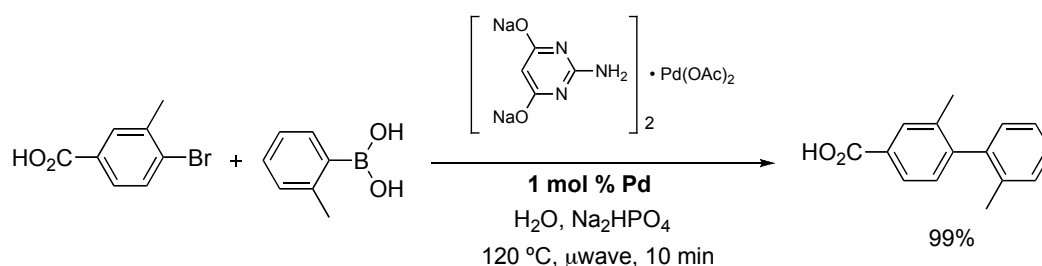
4-Bromo-3-methylbenzoic acid (100 mg, 0.47 mmol), phenylboronic acid (69 mg, 0.56 mmol), and Na_2HPO_4 (334 mg, 2.35 mmol) were all added as solids to a 10 mL microwave reaction vessel. The solids were dissolved in 3 mL H_2O by stirring at 50 $^\circ\text{C}$ for 5 minutes. An aliquot of the palladium catalyst (470 μL of the 0.01M Pd solution, 1 mol% Pd) was added to the reaction mixture and the reaction immediately heated to 120 $^\circ\text{C}$ under microwave irradiation and stirred for 10 minutes. After cooling to room temperature, the reaction was partitioned between 50 mL 1M HCl and 50 mL EtOAc. The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography to afford the cross-coupled product as white crystals (98 mg, 99% yield). m.p. = 139-140 $^\circ\text{C}$. IR (ν_{max} , KBr): 2980, 1685, 1483, 1285, 1008, 780, 700. ^1H NMR (400 MHz, DMSO): δ = 2.26 (3H, s, CH_3), 7.29 (1H, d, J = 7.9), 7.34 (2H, d, J = 7.7), 7.39 (1H, d, J = 6.7), 7.45 (2H, t, J = 7.7), 7.83 (1H, d, J = 7.9), 7.90 (1H, s). ^{13}C NMR (100 MHz, DMSO): δ = 20.2 (CH_3), 127.0, 127.5, 128.4, 128.8, 129.9, 131.3 (6 $\times$ CH_{Ar}), 134.2, 135.2, 140.4, 145.6 (4 $\times$ 4 $^\circ$ Ar), 167.4 (C=O). HRMS m/z (ESI-): Found 211.0767 (M-H $^-$); $\text{C}_{14}\text{H}_{11}\text{O}_2$ requires 211.0765.

Table 6.4, Entry 13

4-bromobenzoic acid (100 mg, 0.50 mmol), *o*-tolylboronic acid (81 mg, 0.60 mmol), and Na_2HPO_4 (355 mg, 2.50 mmol) were all added as solids to a 10 mL microwave reaction vessel. The solids were dissolved in 3 mL of H_2O by stirring at 50 $^\circ\text{C}$ for 5 minutes. An aliquot of the palladium solution (500 μL of the 0.01M Pd solution, 1 mol% Pd) was

added to the mixture and the reaction was immediately heated to 120 °C under microwave irradiation for 10 minutes with stirring. After cooling to room temperature, the product was partitioned between 50 mL 1M HCl and 50 mL H₂O. The organic layer was separated, dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10 % MeOH in CH₂Cl₂) to afford the cross-coupled product as white crystals (105 mg, 99% yield). m.p. = 185-186 °C. IR (ν_{max} , KBr): 3019, 2360, 1672, 1609, 1585, 1558, 1481, 1320, 1289. ¹H NMR (400 MHz, DMSO): δ = 2.22 (3H, s, CH₃), 7.20-7.28 (4H, m, Tol), 7.43 (2H, d, J = 7.9), 8.02 (2H, d, J = 7.9). ¹³C NMR (50 MHz, DMSO): δ = 20.1 (CH₃), 126.1, 127.8, 129.2, 129.3, 129.4 (5 $\times$ CH_{Ar}), 129.9 (4° Ar), 130.5 (CH_{Ar}), 134.7, 140.4, 145.6 (3 $\times$ 4° Ar), 167.6 (C=O). HRMS m/z (ESI-): Found 211.0766 (M-H⁻); C₁₄H₁₁O₂ requires 211.0765.

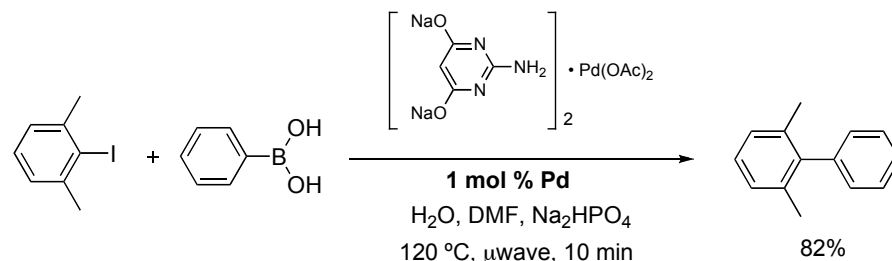
Table 6.4, Entry 14



3-Methyl-4-bromobenzoic acid (100 mg, 0.47 mmol), *o*-tolylboronic acid (76 mg, 0.56 mmol), and Na₂HPO₄ (334 mg, 2.35 mmol) were all added as solids to a 10 mL microwave reaction vessel. H₂O (3 mL) was added and the mixture stirred at 50 °C to dissolve most of the material. An aliquot of the 0.01 M palladium catalyst solution (470 μ L, 1 mol%) was added to the reaction and the tube was immediately heated to 120 °C under microwave irradiation for 10 minutes with stirring. After cooling to room temperature, the reaction was partitioned between 50 mL 1M HCl and 50 mL EtOAc. The organic layer was separated, dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (10% MeOH in CH₂Cl₂). The product was isolated as a white, waxy solid (104 mg, 99% yield). m.p. = 136-137 °C. IR (ν_{max} , film): 2922, 1690, 1611, 1564, 1425, 1304, 1199, 1125, 1007, 761, 730, 705. ¹H NMR (400 MHz, DMSO): δ = 1.96 (3H, d, J = 3.3, CH₃), 2.02 (3H, d, J = 2.8, CH₃), 7.03 (1H, d, J = 7.1), 7.15 (1H, dd, J = 7.8, 3.3), 7.20-7.31 (3H, m), 7.81 (1H, d, J = 7.6), 7.89 (1H, s). ¹³C NMR (100 MHz, DMSO): δ = 19.42, 19.45 (2 $\times$ CH₃), 125.8, 126.8, 127.7, 128.6, 129.4, 130.0 (6 $\times$ CH_{Ar}), 130.1 (4° Ar), 130.8 (CH_{Ar}), 134.8,

135.7, 140.3, 145.6, 145.6 ($4 \times 4^\circ \text{Ar}$), 167.5 ($\text{C}=\text{O}$). HRMS m/z (ESI⁻): Found 225.0924 ($\text{M}-\text{H}^-$); $\text{C}_{15}\text{H}_{13}\text{O}_2$ requires 225.0921.

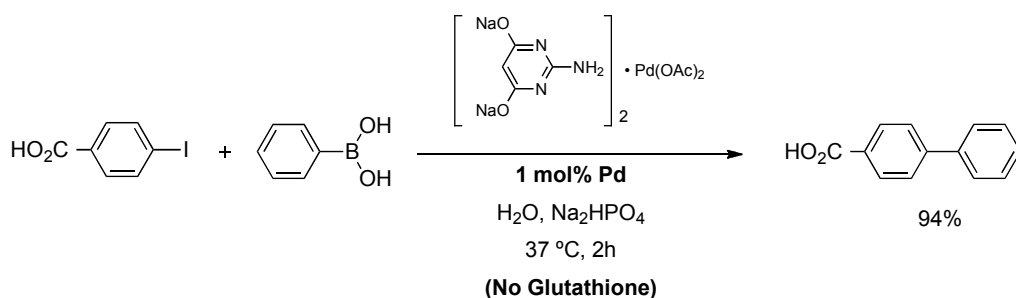
Table 6.4, Entry 15



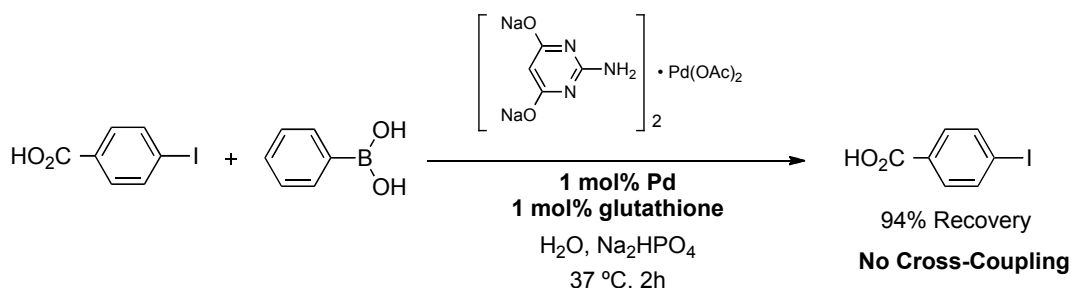
2,5-Dimethyl-iodobenzene (100 mg, 0.43 mmol) was added to a 10 mL microwave reaction vessel and dissolved in DMF (1.43 mL). Phenylboronic acid (105 mg, 0.86 mmol) was added to the stirred solution at room temperature. In a separate vial, Na_2HPO_4 (244 mg, 1.72 mmol) was dissolved in 1.00 mL of H_2O . The solution of base was then added to the reaction vessel. Two layers formed on addition of the base. An aliquot of the 0.01 M Pd catalyst solution (430 μL , 1 mol%) was added and the mixture was immediately heated to 120°C under microwave irradiation with stirring for 10 minutes. After cooling to room temperature, the reaction was partitioned between EtOAc (50 mL) and H_2O (50 mL). The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (petrol), affording the cross-coupled product as a clear oil (64 mg, 82% yield). IR (ν_{max} , film): 3058, 3020, 2952, 2921, 1463, 1442, 1099, 1031, 847, 742. ^1H NMR (400 MHz, CDCl_3): δ = 2.10 (6H, s, $2 \times \text{CH}_3$), 7.16-7.24 (5H, m), 7.39 (1H, dt, J = 7.0, 1.2), 7.48 (2H, dt, J = 7.0, 1.1). ^{13}C NMR (100 MHz, CDCl_3): δ = 20.8 (CH_3), 126.6, 127.0, 127.2, 128.4, 129.0 ($5 \times \text{CH}_{\text{Ar}}$), 136.0, 141.0, 141.8 ($3 \times 4^\circ \text{Ar}$). HRMS m/z (FI⁺): Found 182.1089 (M^+); $\text{C}_{14}\text{H}_{14}$ requires 182.1096.

Cysteine Compatibility Experiments

To assess the cysteine tolerance of the palladium catalyst, glutathione was added to a model cross-coupling. In the absence of glutathione, the cross-coupling proceeded in 94% yield. With a 1:1 molar ratio of palladium and glutathione, no cross-coupling was observed under the same conditions and only unreacted aryl iodide was recovered:



4-iodobenzoic acid (100 mg, 0.40 mmol), phenylboronic acid (74 mg, 0.60 mmol), and Na_2HPO_4 (284 mg, 2.00 mmol) were added to a 10 mL glass vial and dissolved in 2 mL H_2O . The mixture was stirred at 50 °C to dissolve and then cooled to 37 °C. The palladium catalyst (400 μL of a 0.01 M solution, 1 mol%) was added to the reaction and the vial was capped and stirred vigorously at 37 °C for two hours. After this time, the reaction was complete by TLC (EtOAc). The reaction was then partitioned between 1M HCl (50 mL) and EtOAc (50 mL). The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting solid was purified by column chromatography (EtOAc) to afford the cross-coupled product as white crystals (75 mg, 94%). Spectroscopic data was identical to that reported above.



4-iodobenzoic acid (100 mg, 0.40 mmol), phenylboronic acid (74 mg, 0.60 mmol), Na_2HPO_4 (284 mg, 2.00 mmol), and glutathione (1.2 mg, 0.0040 mmol) were added to a 10 mL glass vial and dissolved in 2 mL H_2O . The mixture was stirred at 50 °C to dissolve and then cooled to 37 °C. The palladium catalyst (400 μL of a 0.01 M solution, 1 mol%) was added to the reaction and the vial was capped and stirred vigorously at 37 °C for two hours. After this time, the reaction was partitioned between 1M HCl (50 mL) and EtOAc (50 mL). The organic layer was separated, dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting solid was purified by column chromatography (EtOAc) to afford unreacted 4-iodobenzoic acid as white crystals (94 mg, 94% recovery). Spectroscopic data was identical to the commercially available starting material.

Protein Experiments

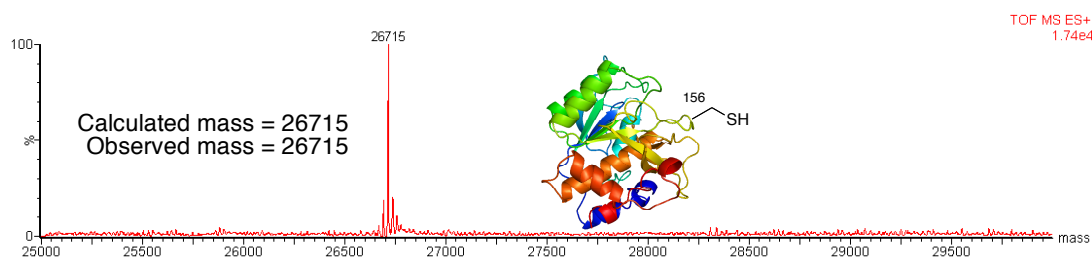
Sequence of subtilisin *Bacillus lentus* (SBL) mutant S156C

PDB code for wild type = 1GCI

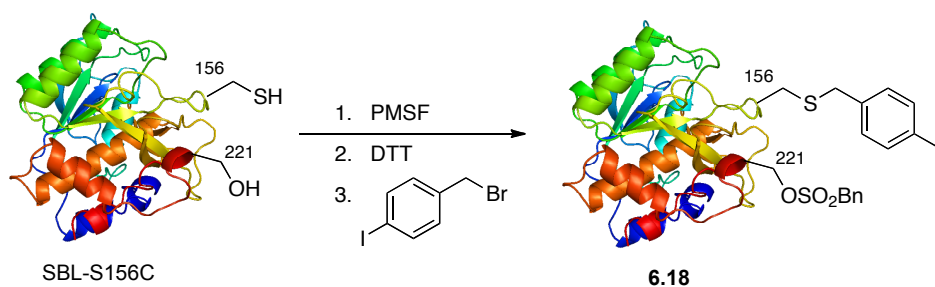
AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPS
TQDGNGHGHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGL
EWAGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGNCGAGSISYPA
RYANAMAVGATDQNNNRASFQYAGGLDIVAPGVNVQSTYPGSTYASLNGTSM
ATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Calculated average isotopic mass = 26714.5

ESI-MS is shown below.

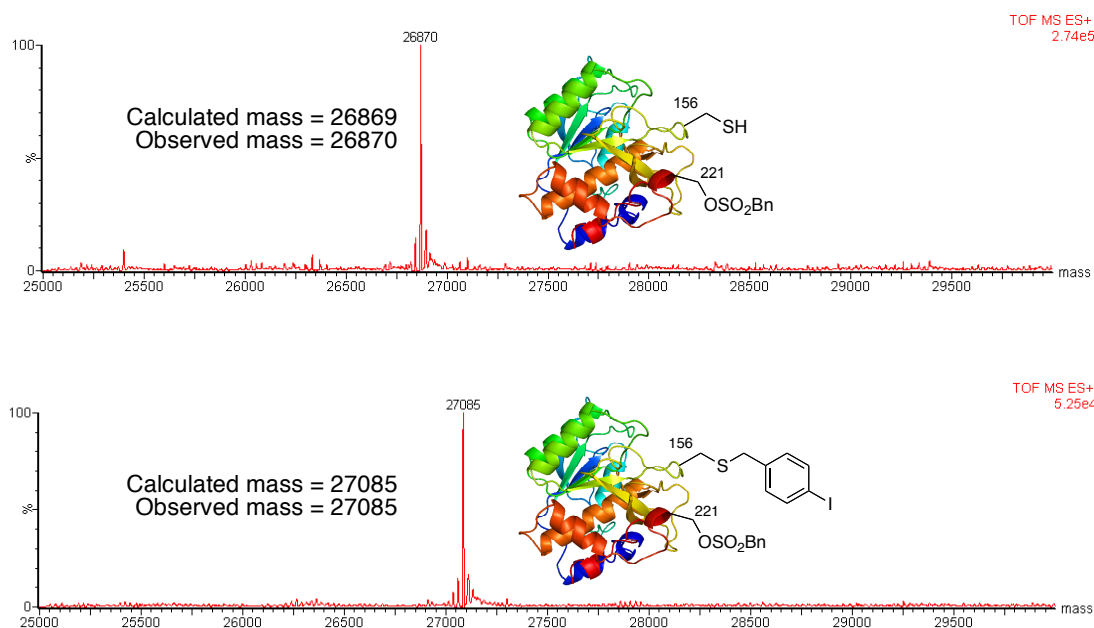


Aryl iodide incorporation on SBL-S156C: Synthesis of SBL-ArI (6.18)

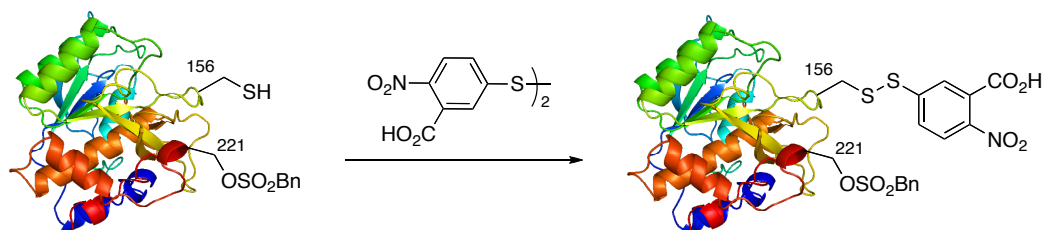


SBL-S156C was prepared as a 6.0 mg/mL solution in sodium phosphate buffer (pH 8.0, 50 mM). A 350 μ L aliquot of the protein solution (0.11 μ mol) was added to a 1.5 mL plastic tube and stored on ice until needed. Phenylmethylsulfonyl fluoride (PMSF, 1.0 mg, 5.6 μ mol) was prepared as a solution in MeCN (50 μ L). All of the PMSF solution was added to the protein and the reaction was shaken 5 minutes at room temperature. After this time, precipitated PMSF was removed by centrifugation (2 min, 4 $^{\circ}$ C, 14K g).

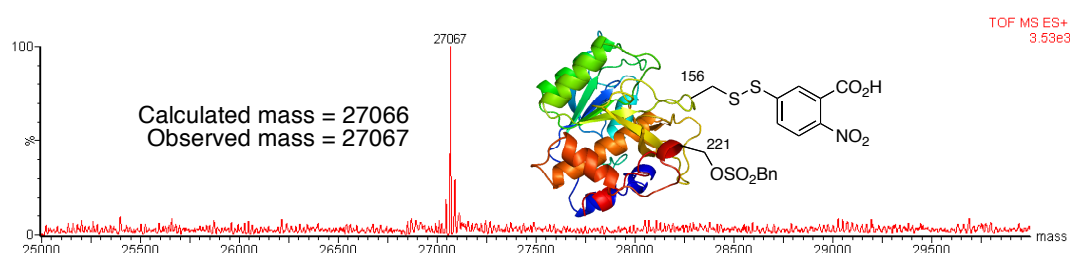
The protein solution was transferred to a new vial and contaminant disulfide was reduced by the addition of DTT (1.7 mg in 20 μ L pH 8.0 sodium phosphate buffer). The reaction was shaken for 10 minutes at room temperature and then a 5 μ L aliquot was diluted to 30 μ L with the same buffer and analyzed by LC-MS. The SBL-PMSF adduct was the only species observed (calculated mass = 26869; observed mass = 26870). Immediately after 10 minutes of incubation with DTT, a solution of 4-iodobenzylbromide (33 mg, 0.11 mmol) in DMF (150 μ L) was added to the protein solution at room temperature. The reaction was vortexed and then shaken vigorously at 37 $^{\circ}$ C for 1 hour. Solids were then removed from the reaction mixture by centrifugation (2 min, 4 $^{\circ}$ C, 14K G). The supernatant was then loaded onto a PD Minitrapp (GE Healthcare) previously equilibrated with 10 mL of pH 8.0 sodium phosphate buffer (50 mM). The protein was eluted with 1.0 mL of the same buffer. LC-MS analysis of the product solution revealed a single protein with a mass corresponding the SBL-ArI **6.18** (calculated mass = 27085; observed mass = 27085). ESI-MS are shown below.



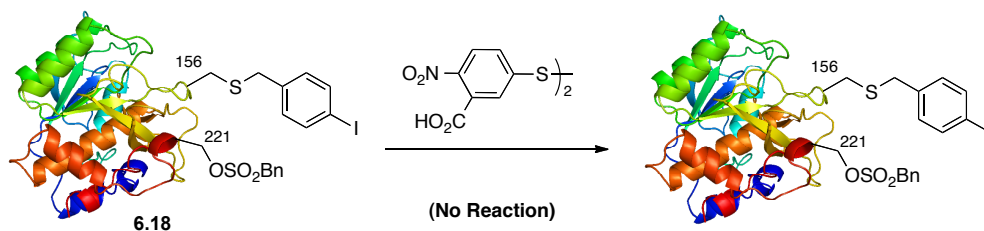
Control: SBL-PMSF and Ellman's Reagent



A sample of PMSF inhibited SBL-S156C was prepared as described above. A 100 μ L aliquot was prepared at 0.14 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A stock solution of Ellman's reagent was prepared by dissolving 2 mg in 1.00 mL pH 8.0 sodium phosphate buffer. To the 100 μ L sample of protein (5.2 nmol) was added 50 μ L of the Ellman's solution (260 nmol). The reaction was shaken for 5 minutes at 4 $^{\circ}$ C and then analyzed directly by LC-MS. A single protein species was detected with a mass corresponding to the Ellman disulfide (calculated mass = 27066; observed mass = 27067). This experiment shows that PMSF does not react with Cys156. This experiment is also the positive control for the following reaction, demonstrating that the Ellman disulfide readily forms at Cys156. ESI-MS is shown below.



Control: SBL-ArI and Ellman's Reagent



A sample of SBL-ArI **6.18** was diluted to 0.14 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). To a 100 μ L aliquot of the protein sample (5.2 nmol) was added 50 μ L of the Ellman's solution prepared above (260 nmol). The mixture was shaken at 4 $^{\circ}$ C for 5 minutes and then analyzed directly by LC-MS. A single protein species was detected with a mass corresponding to the unreacted aryl iodide (calculated mass = 27085; observed mass = 27086). This experiment indicates that all cysteine is consumed during the benzylation to form SBL-ArI **6.18**, and is consistent with alkylation at cysteine. Under the same reaction conditions, unalkylated Cys156 is converted to the corresponding disulfide (see previous experiment). ESI-MS is shown below.

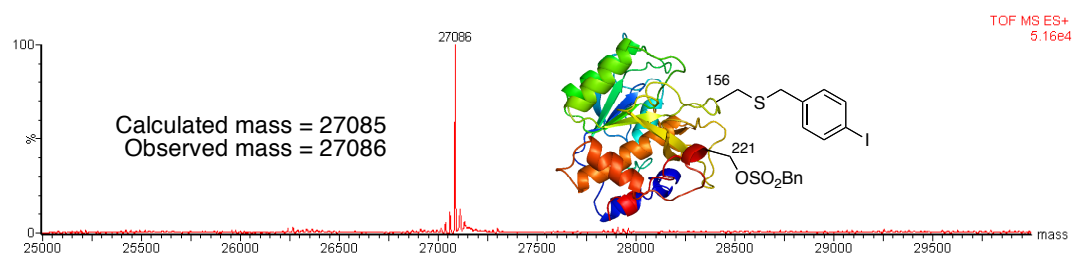
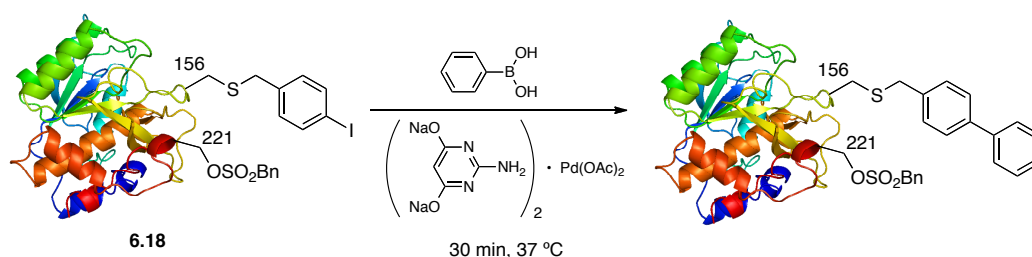


Table 6.3: Cross-Coupling on SBL-156ArI (6.18)

Table 6.3, Entry 1



SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of phenylboronic acid was prepared by dissolving 9.7 mg of phenylboronic acid and 11.9 mg of Na_2HPO_4 in 808 μ L of H_2O at room temperature. The protein sample was warmed to room temperature and an aliquot of the phenylborate solution (7.5 μ L, 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μ L, 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27035; observed mass = 27035). ESI-MS is shown below.

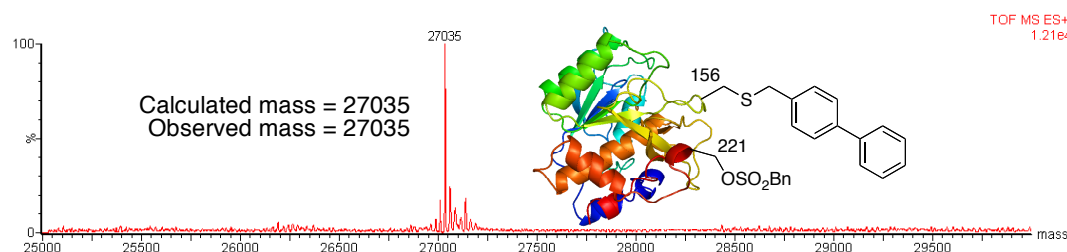
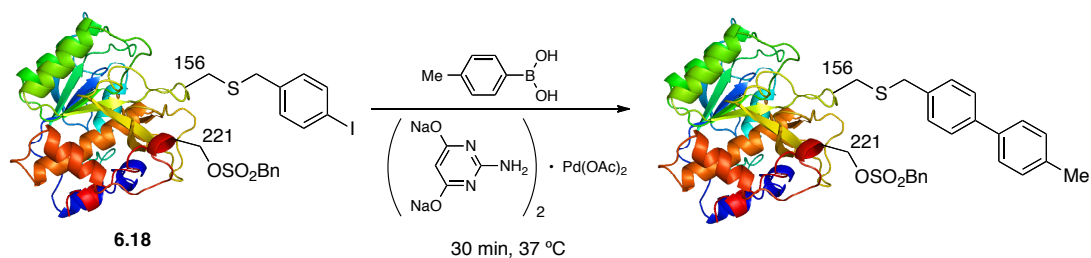
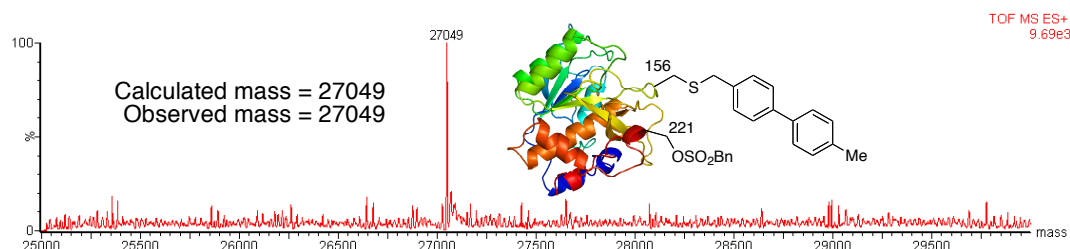
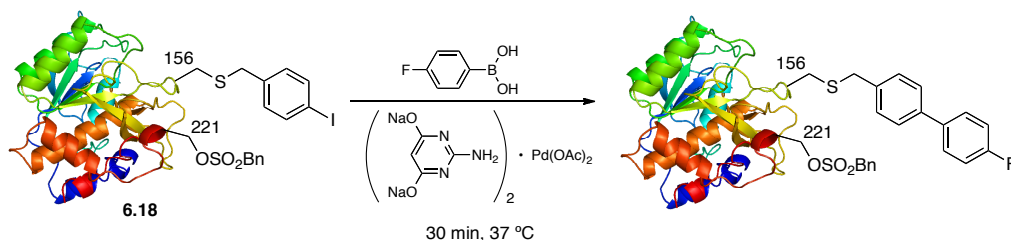


Table 6.3, Entry 2

SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of 4-tolylboronic acid was prepared by dissolving 5.8 mg of 4-tolylboronic acid and 6.4 mg Na_2HPO_4 in 870 μ L H_2O at 60 °C. Once the solution was clear, it was kept at 37 °C to prevent crystallization. The protein sample was warmed to room temperature and an aliquot of the 4-tolylborate solution (15 μ L, 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μ L, 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27049; observed mass = 27049). ESI-MS is shown below.

**Table 6.3, Entry 3**

SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of 4-fluorophenylboronic acid was prepared by dissolving 6.1 mg of 4-fluorophenylboronic acid and 6.1 mg Na_2HPO_4 in 457

$\mu\text{L H}_2\text{O}$ at 50 °C. Once the solution was dissolved it was stored at 37 °C. The protein sample was warmed to room temperature and an aliquot of the 4-fluorophenylborate solution (7.5 μL , 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μL , 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27053; observed mass = 27053). ESI-MS is shown below.

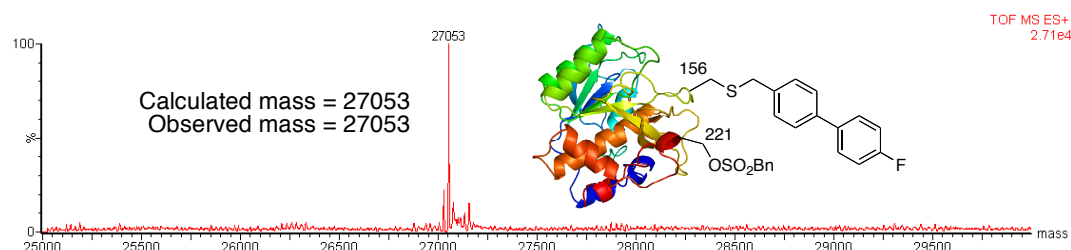
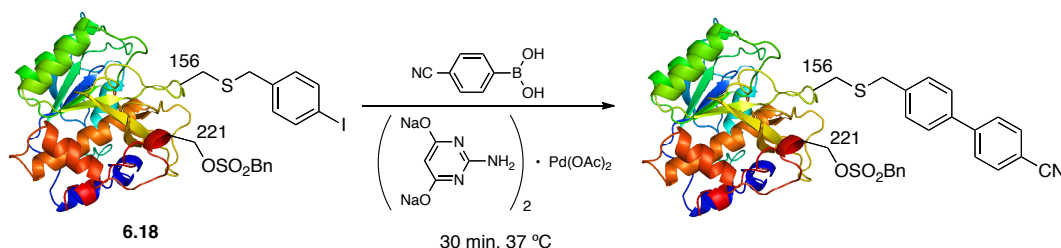


Table 6.3, Entry 4



SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μL aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of 4-cyanophenylboronic acid was prepared by dissolving 3.8 mg of 4-cyanophenylboronic acid and 3.8 mg Na₂HPO₄ in 518 $\mu\text{L H}_2\text{O}$ at 50 °C. Once the solution was clear, it was kept at 37 °C to avoid crystallization. The protein sample was warmed to room temperature and an aliquot of the 4-cyanophenylborate solution (15 μL , 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μL , 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27060; observed mass = 27060). ESI-MS is shown below.

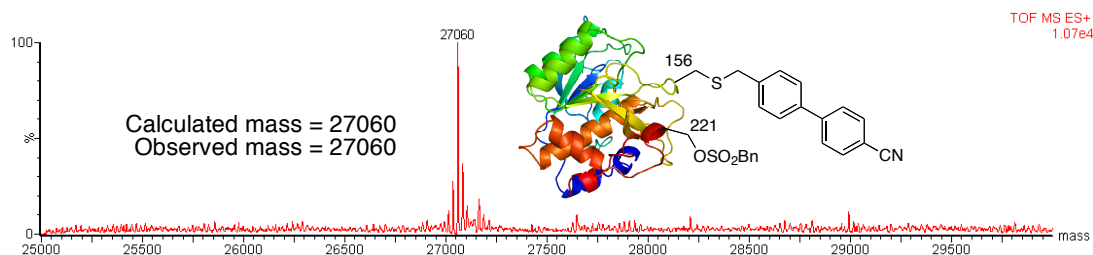
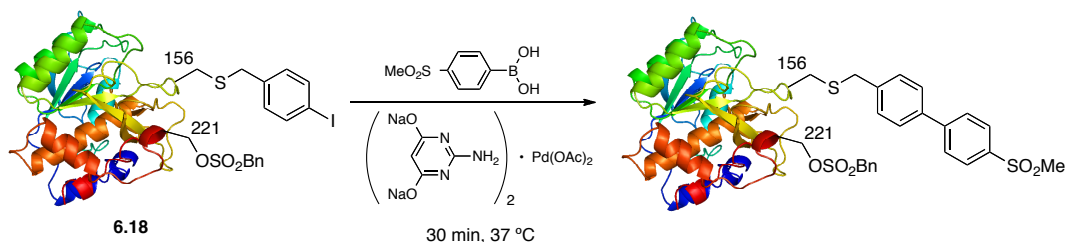


Table 6.3, Entry 5



SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of 4-(methanesulfonyl)phenylboronic acid was prepared by dissolving 5.3 mg of 4-(methanesulfonyl)phenylboronic acid and 4.9 mg Na_2HPO_4 in 663 μ L H_2O at 50 °C. Once the solution was clear, it was kept at 37 °C to prevent crystallization. The protein sample was warmed to room temperature and an aliquot of the 4-(methanesulfonyl)phenylborate solution (15 μ L, 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μ L, 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27113; observed mass = 27113). ESI-MS is shown below.

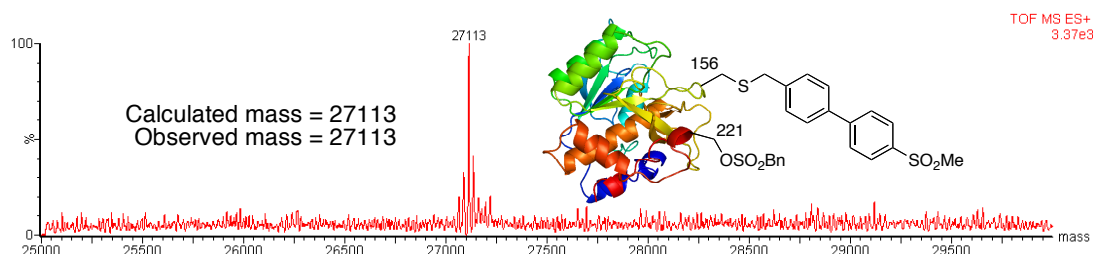
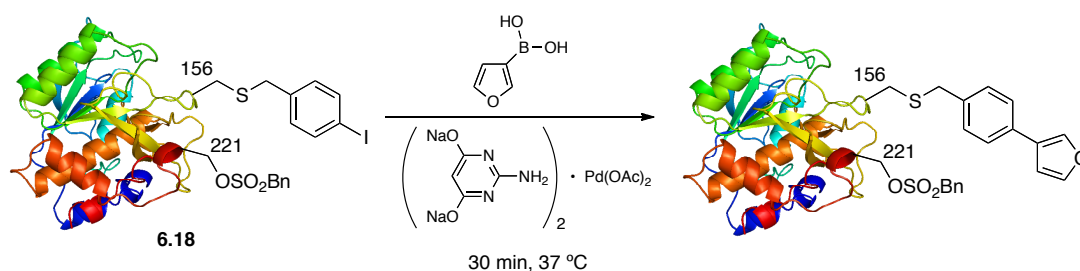
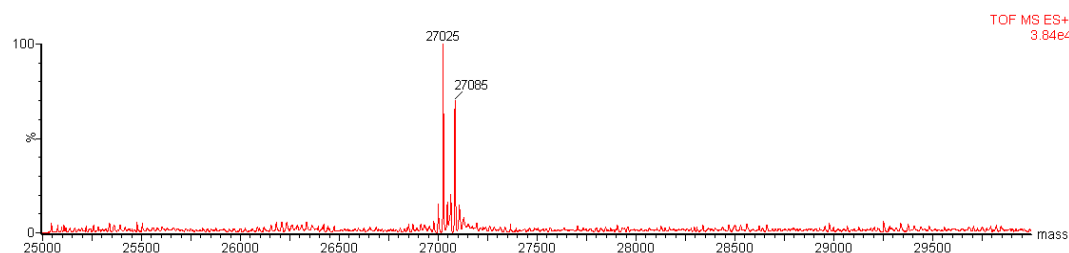


Table 6.3, Entry 6



SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of furan-3-boronic acid was prepared by dissolving 5.3 mg of furan-3-boronic acid and 7.3 mg Na_2HPO_4 in 497 μ L H_2O at room temperature. The protein sample was warmed to room temperature and an aliquot of the furan-3-borate solution (7.5 μ L, 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μ L, 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27025; observed mass = 27025). If the reaction was run for only 15 minutes at 37 °C, only 60% conversion was observed. ESI-MS are shown below.

15 Minutes:



30 Minutes:

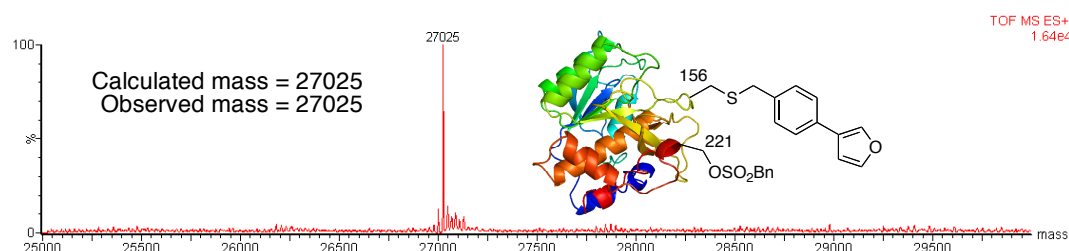
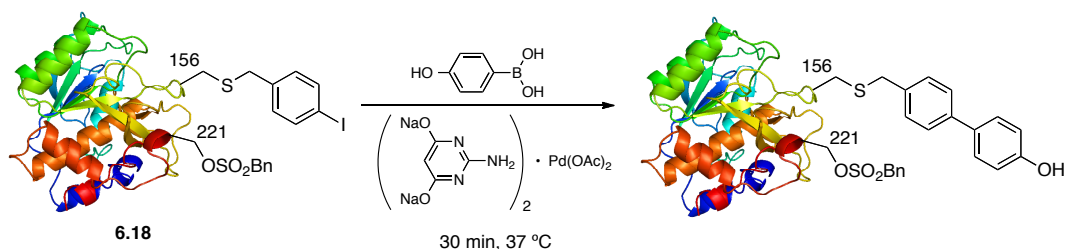
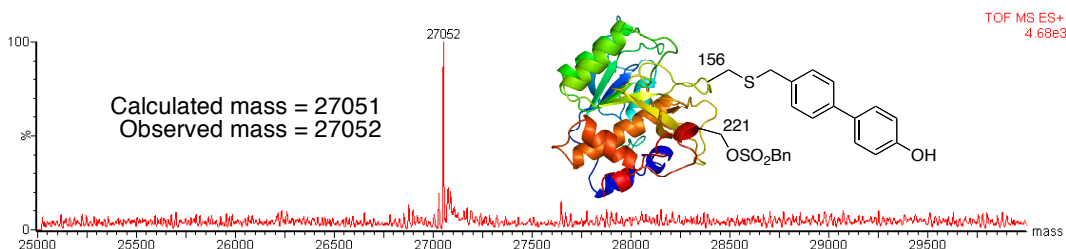
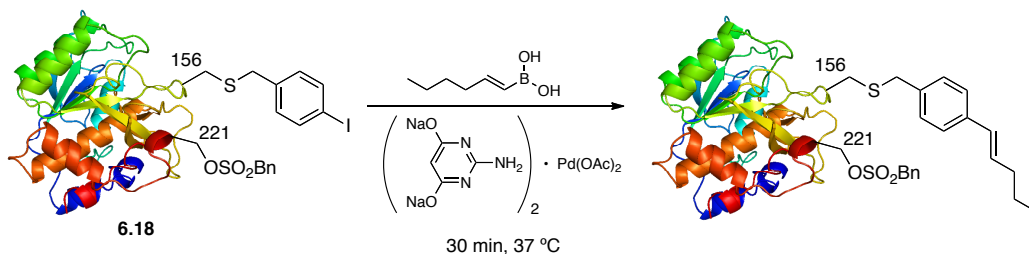


Table 6.3, Entry 7

SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of 4-hydroxyphenylboronic acid was prepared by dissolving 3.2 mg of 4-hydroxyphenylboronic acid and 3.5 mg Na_2HPO_4 in 240 μ L H_2O at room temperature. The protein sample was warmed to room temperature and an aliquot of the 4-hydroxyphenylborate solution (7.5 μ L, 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μ L, 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27051; observed mass = 27052). ESI-MS is shown below.

**Table 6.3, Entry 8**

SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of 1-hexenylboronic acid was prepared by vigorously stirring 4.2 mg of 1-hexenylboronic acid and 4.6 mg Na_2HPO_4 in 630 μ L H_2O at 60 °C. Once the solution was clear, it was kept at 37 °C to prevent

crystallization. The protein sample was warmed to room temperature and an aliquot of the 1-hexenylborate solution (15 μ L, 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μ L, 75 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 $^{\circ}$ C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27041; observed mass = 27041). ESI-MS is shown below.

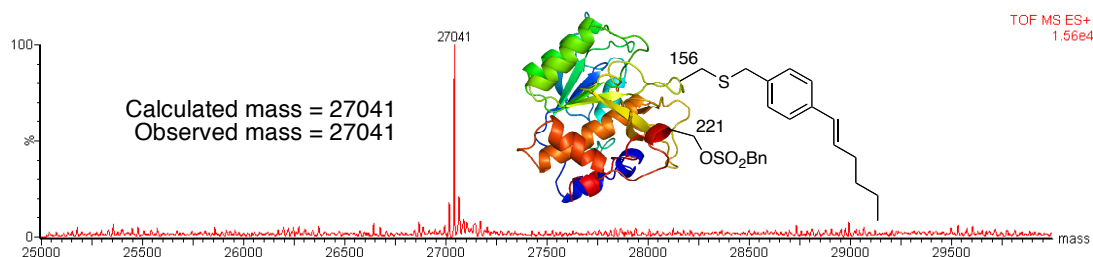
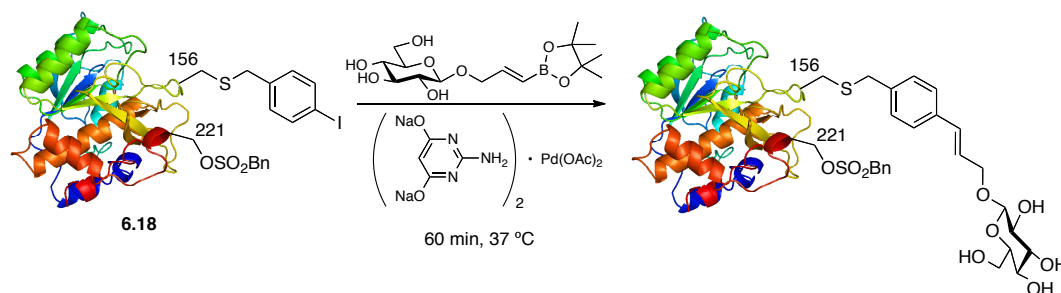
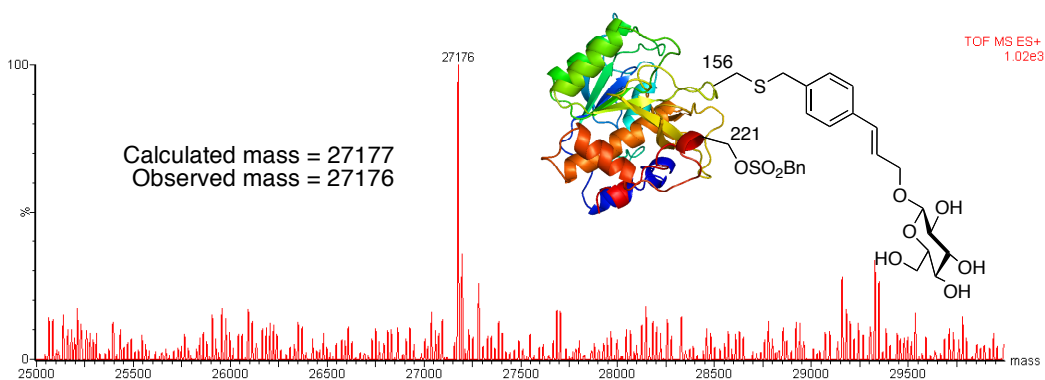


Table 6.3, Entry 9

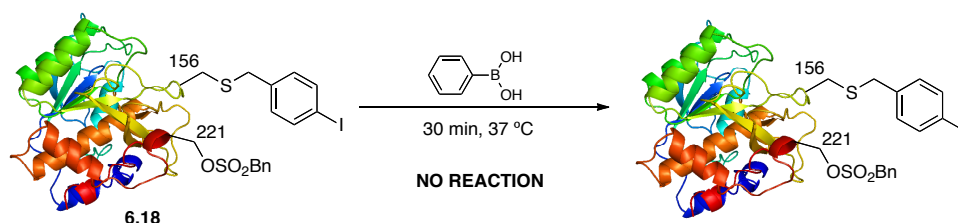


SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of the glucose pinacolboronic ester was prepared by dissolving 3.3 mg of the sugar and 6.3 mg Na_2HPO_4 in 185 μ L H_2O . The protein sample was warmed to room temperature and an aliquot of the boronic ester solution (50 μ L, 2.6 μ mol) and an aliquot of the 0.01M Pd catalyst solution (15 μ L, 150 nmol) were added sequentially. The reaction was vortexed to homogenize and then incubated at 37 $^{\circ}$ C for 60 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the cross-coupled product (calculated mass = 27177; observed mass = 27176). ESI-MS is shown below.

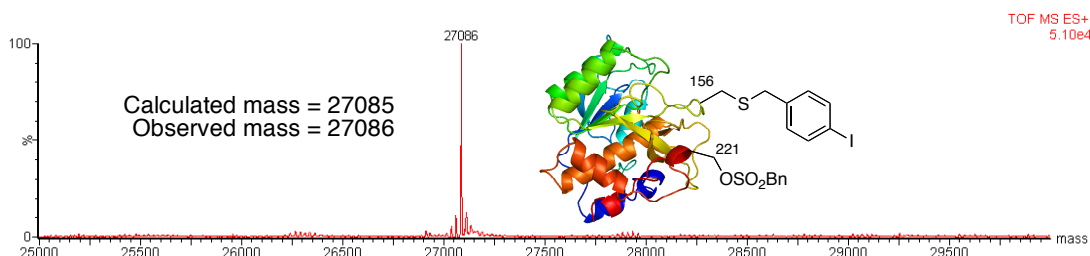


Cross-Coupling Controls for Protein Modification

Boronic Acid Only:



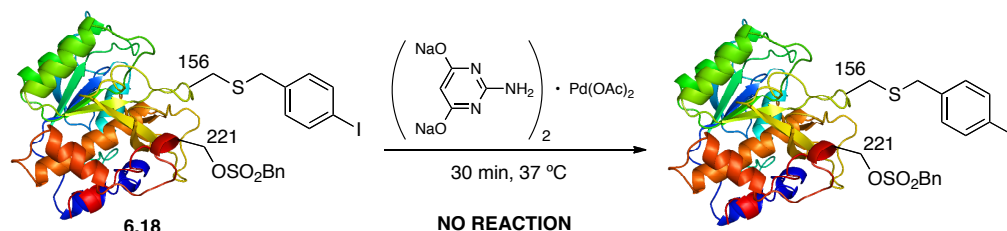
SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of phenylboronic acid was prepared by dissolving 9.7 mg of phenylboronic acid and 11.9 mg of Na₂HPO₄ in 808 μ L of H₂O. The protein sample was warmed to room temperature and an aliquot of the phenylborate solution (7.5 μ L, 750 nmol) was added. The mixture was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the unreacted aryl iodide (calculated mass = 27085; observed mass = 27086). ESI-MS is shown below.



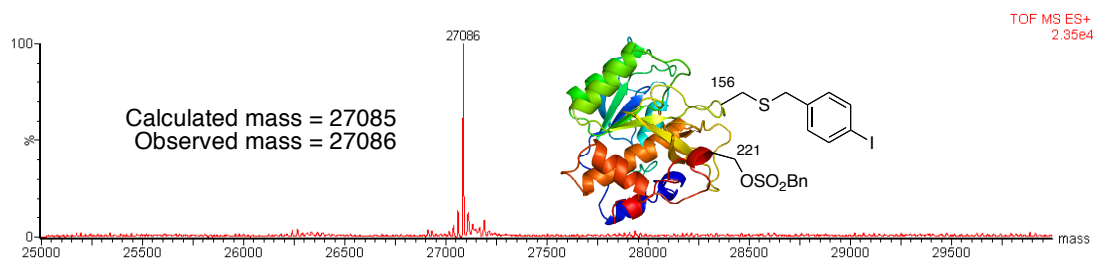
Note regarding boronic acid compatibility with protein substrates: Since boronic acids can condense with alcohols or amines it is possible that for some protein substrates boronic acids may be problematic. Under the conditions used for the cross-coupling,

however, we have seen no such non-selective labelling. Moreover, even if boronic acids coordinated to residues such as serine or lysine, such bonding is reversible and the addition of diols or diamines to the reaction mixture after cross-coupling may be a strategy to resolve such complications, should they arise. For the case of the model protein used in the present studies, Ser221 is unique in that it is a catalytic residue of the serine protease and is highly nucleophilic, hence the use of PMSF to block this residue that is known to attack boronic acids. The remaining 32 serine residues and 5 lysine residues are unprotected and did not affect the coupling. It should also be noted that inhibition of the protease by a boronic acid would not prohibit the cross-coupling. The purpose of PMSF inhibition is to decouple the formation of borate by the protease from the formation of borate by the buffer. This distinction is important to illustrate that the cross-coupling is not dependent on the protein's ability to convert boronic acids to borates.

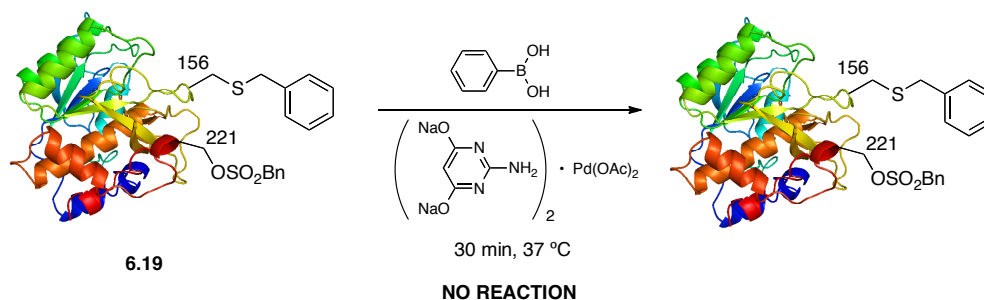
Palladium Only:



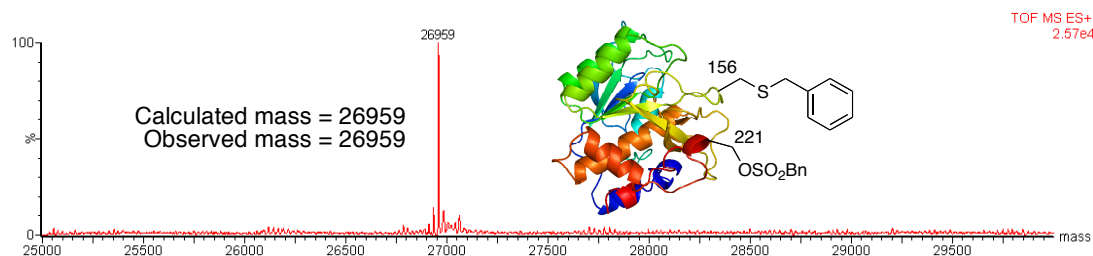
SBL-ArI **6.18** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). A 50 μL aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. The protein sample was warmed to room temperature and an aliquot of the 0.01M Pd catalyst solution (7.5 μL , 75 nmol) was added. The mixture was vortexed to homogenize and then incubated at 37 °C for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the unreacted aryl iodide (calculated mass = 27085; observed mass = 27086). ESI-MS is shown below.



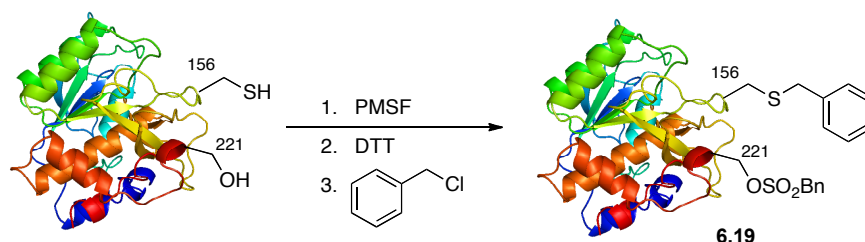
Boronic Acid and Palladium, No Aryl Iodide:



SBL-156CysBn **6.19** was prepared at 0.8 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM) (see next experiment for the synthesis of **6.19**). A 50 μ L aliquot of the protein solution (1.5 nmol) was transferred to a 1.5 mL plastic tube and kept on ice until needed. A stock solution of phenylboronic acid was prepared by dissolving 9.7 mg of phenylboronic acid and 11.9 mg of Na_2HPO_4 in 808 μ L of H_2O . The protein sample was warmed to room temperature and an aliquot of the phenylborate solution (7.5 μ L, 750 nmol) and an aliquot of the 0.01M Pd catalyst solution (7.5 μ L, 75 nmol) were added sequentially. The mixture was vortexed to homogenize and then incubated at 37 $^\circ\text{C}$ for 30 minutes. After this time, the reaction was analyzed directly by LC-MS. The only protein detected displayed a mass corresponding to the unmodified protein (calculated mass = 26959; observed mass = 26959). ESI-MS is shown below.

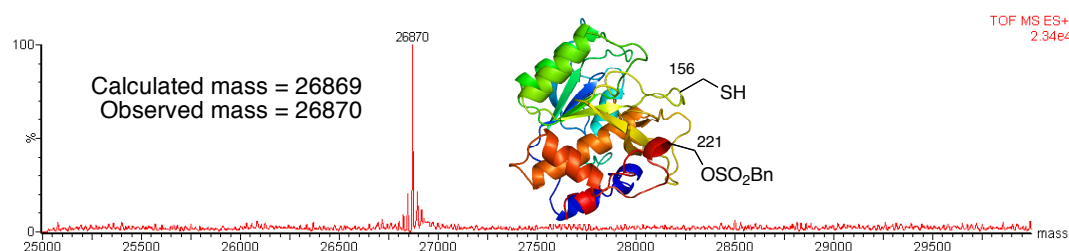


Synthesis of benzylated, uniodinated protein substrate 6.19:

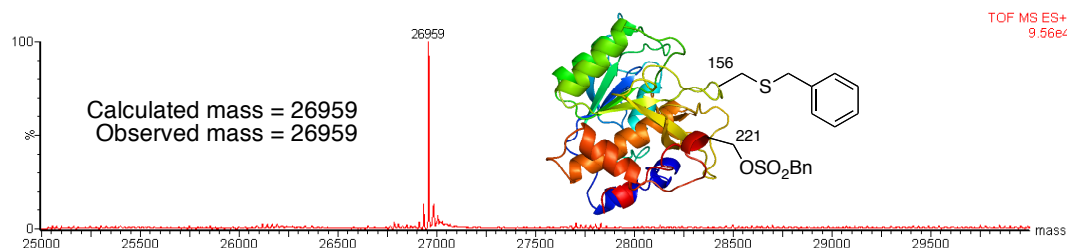


SBL-Cys156 was prepared at 2.0 mg/mL in pH 8.0 sodium phosphate buffer (50 mM) and kept on ice until needed. A solution of PMSF was prepared by dissolving 3.9 mg in 443 μ L of MeCN. A 25 μ L aliquot of the PMSF solution (1.3 μ mol) was added to 350 μ L of the protein (26 nmol) and the reaction shaken for 5 minutes at room temperature. A stock solution of DTT was prepared by dissolving 4.7 mg in 235 mL pH 8 sodium phosphate buffer (50 mM). 10 μ L of the DTT solution (1.3 μ mol) was added to the protein solution to reduce contaminant disulfide. The reaction was shaken 10 minutes at room temperature and then a 10 μ L aliquot was taken for LC-MS analysis (calculated mass = 26869; observed mass = 26870). Meanwhile, a stock solution of benzyl chloride was prepared by dissolving 5 μ L in 500 μ L of DMF. A 150 μ L aliquot of the benzyl chloride solution (13 μ mol) was then added to the protein solution. The reaction was vortexed to homogenize and then shaken vigorously at room temperature. After 30 minutes, small molecules were removed by passing the reaction solution through a PD MiniTrap (GE Healthcare), eluting with pH 8.0 sodium phosphate buffer (50 mM). The protein solution was analyzed by LC-MS, revealing a single protein species with a mass corresponding to the PMSF inhibited and benzylated product (calculated mass = 26959; observed mass = 26959). The product protein did not react with Ellman's reagent, indicating consumption of cysteine—consistent with benzylation at Cys156 (see next). ESI-MS are shown below.

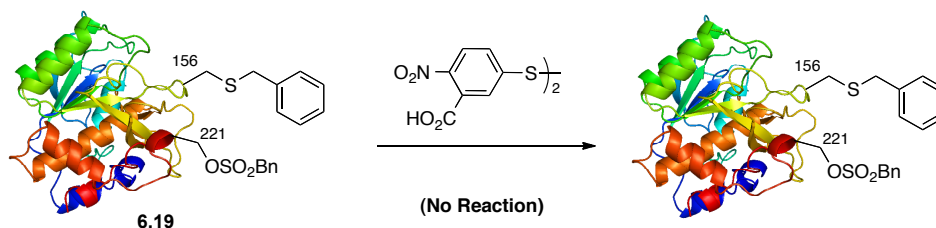
PMSF Inhibition:



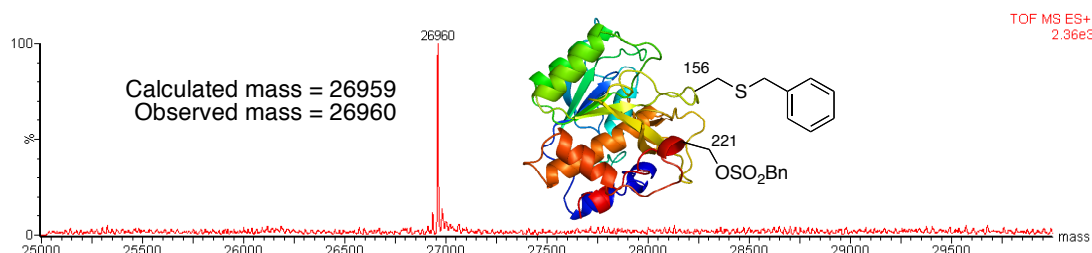
Benylation:



Control: Ellman's Test on benzylated, uniodinated protein



A sample of SBL-156SBn **6.19** was diluted to 0.14 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). To a 100 μ L aliquot of the protein sample (5.2 nmol) was added 50 μ L of the Ellman's solution prepared above (260 nmol). The mixture was shaken at 4 $^{\circ}$ C for 5 minutes and then analyzed directly by LC-MS. A single protein species was detected with a mass corresponding to the unreacted protein (calculated mass = 26959; observed mass = 26960). This experiment indicates that all cysteine is consumed during the benzylation to form SBL-156SBn, and is consistent with alkylation at cysteine. Under the same reaction conditions, unalkylated Cys156 is converted to the corresponding disulfide. ESI-MS is shown below.

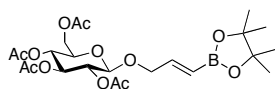


6.6 References for Experimental Procedures

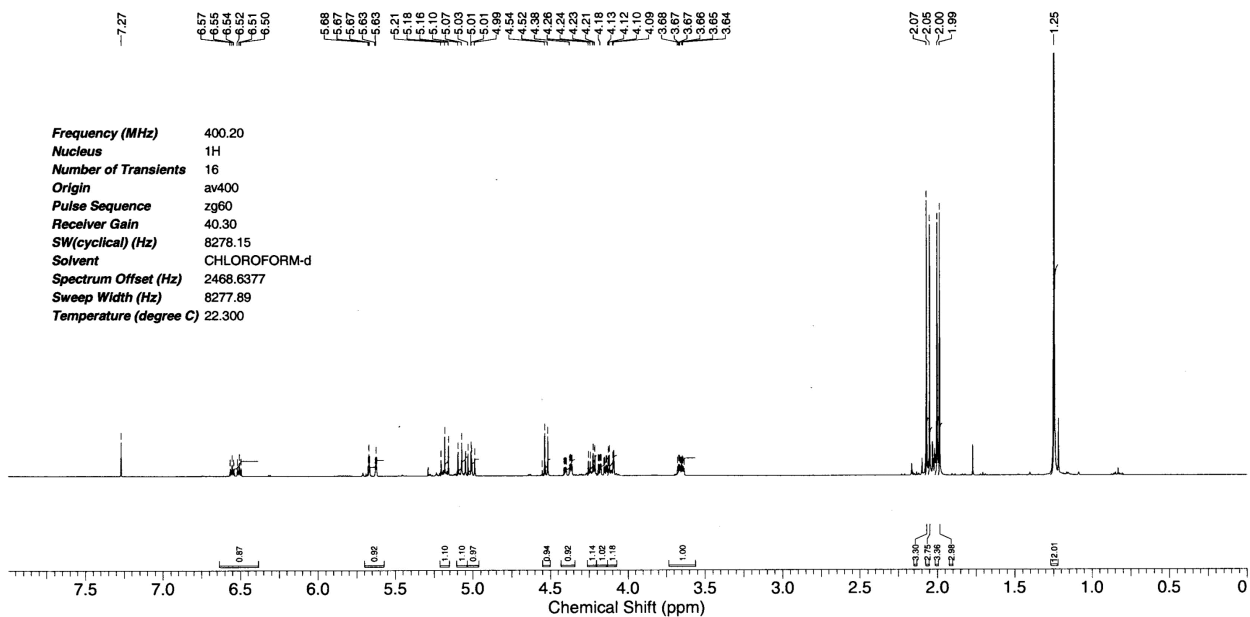
- (1) Jehanno, E.; Vaultier, M. *Tetrahedron Lett.* **1995**, 36, 4439-4442.
- (2) Tucker, C. E.; Davidson, J.; Knochel, P. J. *Org. Chem.* **1992**, 57, 3482-3485.

-
- (3) Walker, W. H.; Rokita, S. E. *J. Org. Chem.* **2003**, *68*, 1563-1566.
- (4) Knör, S.; Laufer, B.; Kessler, H. *J. Org. Chem.* **2006**, *71*, 5625-5630.
- (5) Vogler, K.; Studer, R. O.; Lergier, W.; Lanz, P. *Helv. Chim. Acta.* **1965**, *48*, 1407-1414.
- (6) Groselj, U.; Bevk, D.; Jakse, R.; Meden, A.; Pirc, S.; Recnik, S.; Stanovnik, B.; Svete, J. *Tetrahedron: Asymmetry* **2004**, *15*, 2367-2383.
- (7) Chankeshwara, S. V.; Chakraborti, A. K. *Org. Lett.* **2006**, *8*, 3259.
- (8) Oswald, C. L.; Carrillo-Márquez, T.; Caggiano, L.; Jackson, R. F. W. *Tetrahedron* **2008**, *64*, 681-687.
- (9) Barluenga, J.; González, J. M.; Campos, P. J.; Asensio, G. *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 319-320.
- (10) Chalker, J. M.; Thompson, A. L.; Davis, B. G. *Org. Synth.* **2010**, *87*, 288-298.
- (11) a) Brink, H.; Rijkers, D.; Liskamp, R. M. J. *J. Org. Chem.* **2006**, *71*, 1817-1824.
b) Nishiyama, S.; Kim, M. H.; Yamamura, S.; *Tetrahedron Lett.* **1994**, *35*, 8397-8400. c) Ocasio, C. A.; Scanlan, T. S. *ACS Chem. Biol.* **2006**, *1*, 585-593.
- (12) Threadgill, M. D.; Gledhill, A. P. *J. Org. Chem.* **1989**, *54*, 2940-2949.
- (13) Korolev, D. N.; Bumagin, N. A. *Tetrahedron Lett.* **2005**, *46*, 5751-5754.
- (14) Wei, J.-F.; Jiao, J.; Feng, J.-J.; Lv, J.; Zhang, X.-R.; Shi, X.-Y.; Chen, Z.-G. *J. Org. Chem.* **2009**, *74*, 6283-6286.
- (15) Ojida, A.; Tsutsumi, H.; Kasagi, N.; Hamachi, I. *Tetrahedron Lett.* **2005**, *46*, 3301-3305.
- (16) Ohishi, T.; Nishiura, M.; Hou, Z. *Angew. Chem. Intl. Ed.* **2008**, *47*, 5792-5795.
- (17) Tao, B.; Boykin, D. W. *J. Org. Chem.* **2004**, *69*, 4330-4335.
- (18) Nájera, C.; Gil-Moltó, J.; Karlström, S.; Falvello, L. R. *Org. Lett.* **2003**, *5*, 1451-1454.
- (19) Campaigne, E. E.; Cline, R. E. *J. Org. Chem.* **1956**, *21*, 39-44.
- (20) Alacid, E.; Nájera, C. *Org. Lett.* **2008**, *10*, 5011-5014.

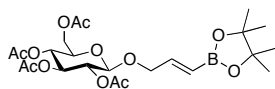
6.7 NMR Spectra



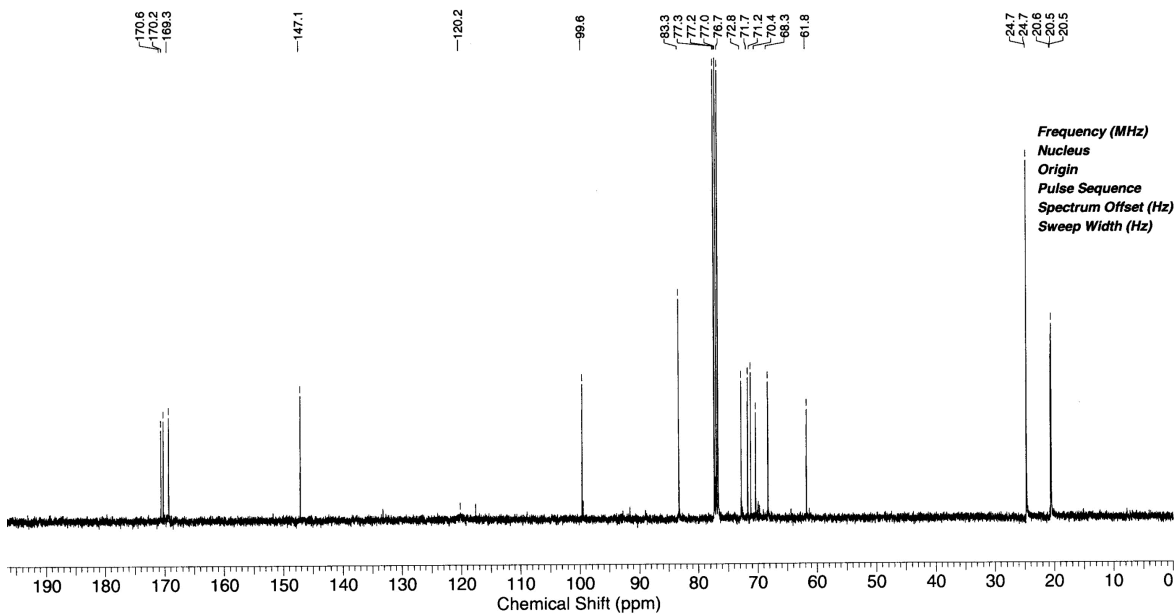
^1H NMR (400 MHz, CDCl_3)



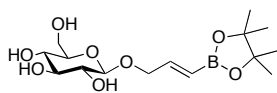
Frequency (MHz) 400.20
Nucleus ^1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 40.30
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.6377
Sweep Width (Hz) 8277.89
Temperature (degree C) 22.300



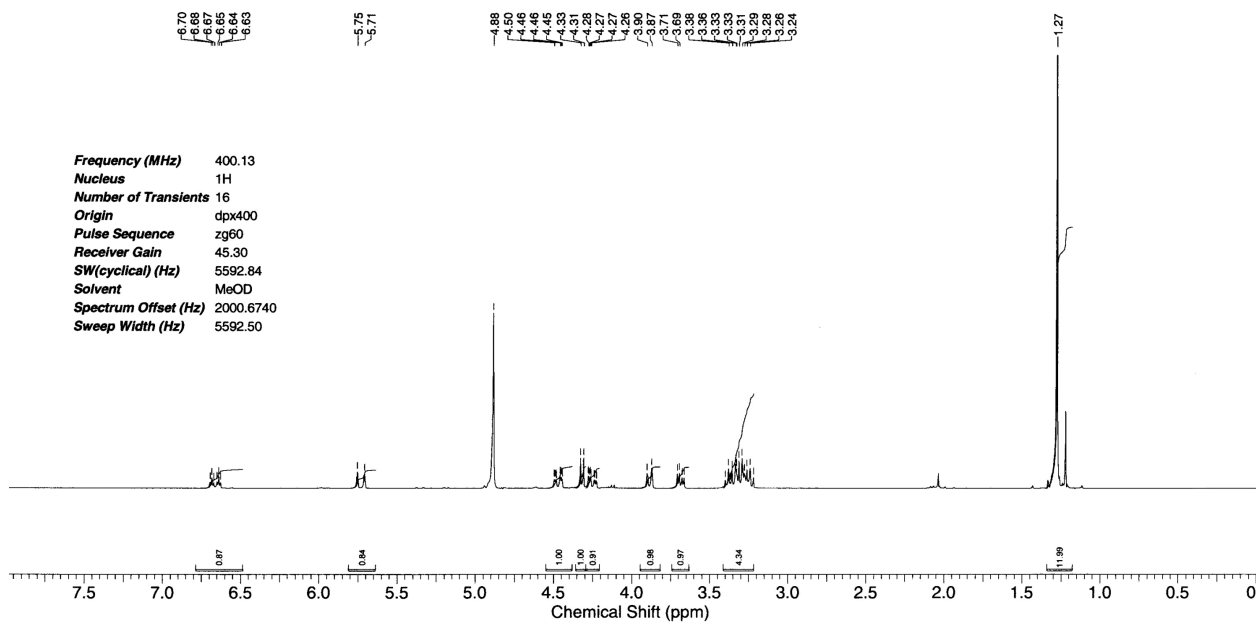
^{13}C NMR (100 MHz, CDCl_3)



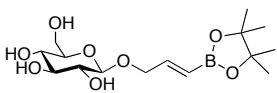
Frequency (MHz) 100.64
Nucleus ^{13}C
Origin Bruker
Pulse Sequence ZGPG30
Spectrum Offset (Hz) 10017.4893
Sweep Width (Hz) 26178.01



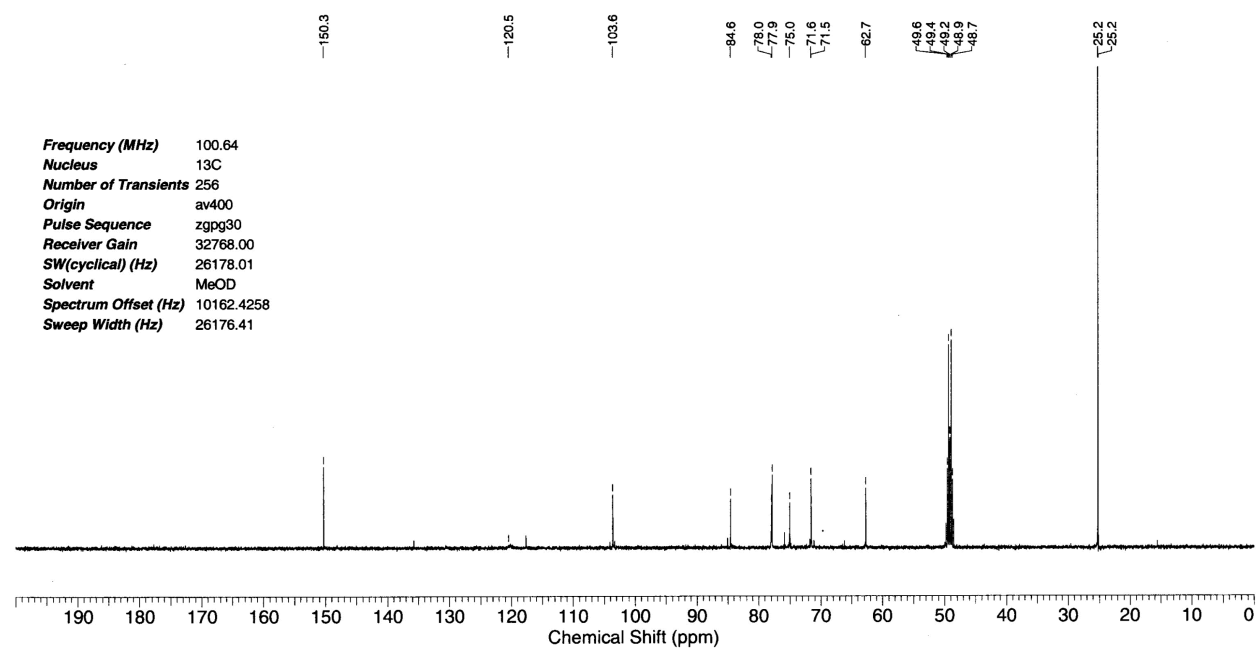
¹H NMR (400 MHz, MeOD)



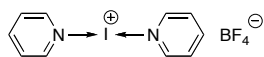
Frequency (MHz) 400.13
Nucleus 1H
Number of Transients 16
Origin dpx400
Pulse Sequence zg60
Receiver Gain 45.30
SW(cyclical) (Hz) 5592.84
Solvent MeOD
Spectrum Offset (Hz) 2000.6740
Sweep Width (Hz) 5592.50



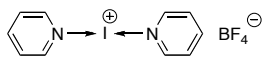
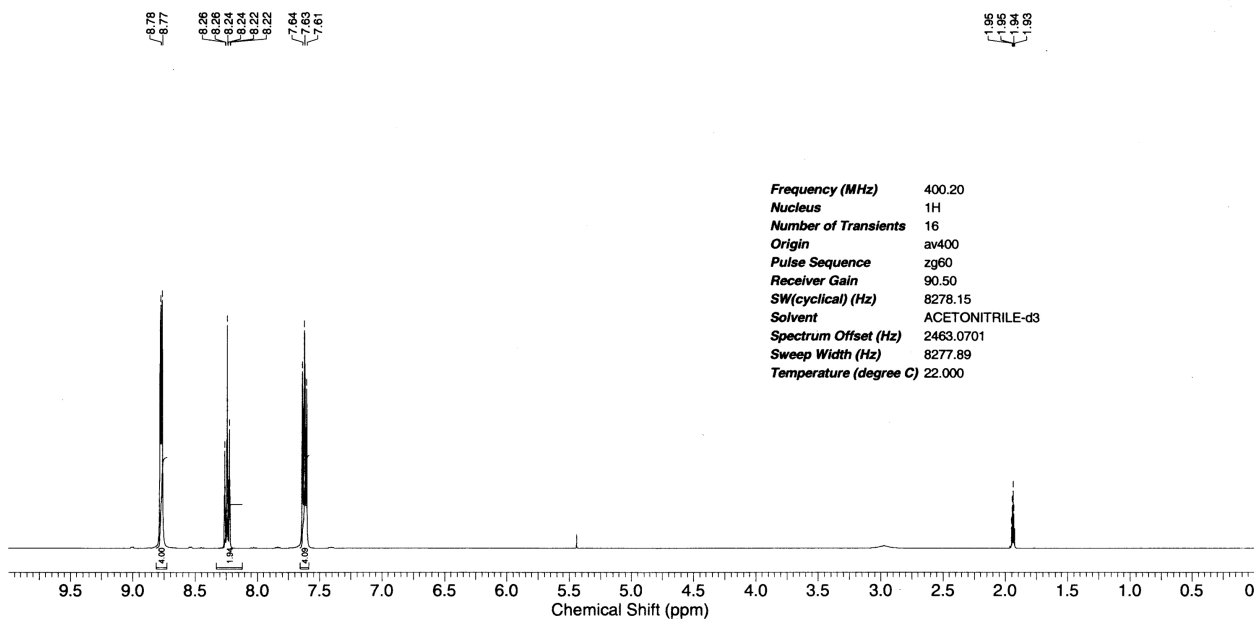
¹³C NMR (100 MHz, MeOD)



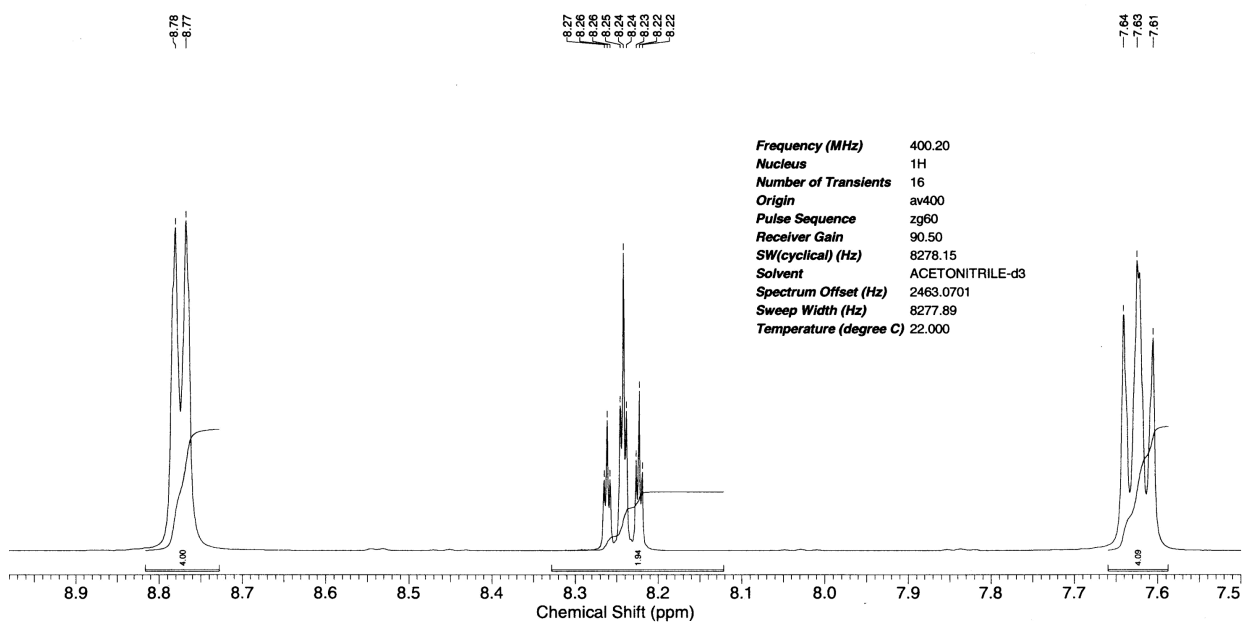
Frequency (MHz) 100.64
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent MeOD
Spectrum Offset (Hz) 10162.4258
Sweep Width (Hz) 26176.41

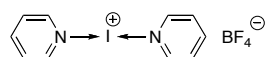


Barluenga's Reagent ^1H NMR (400 MHz, CD_3CN)

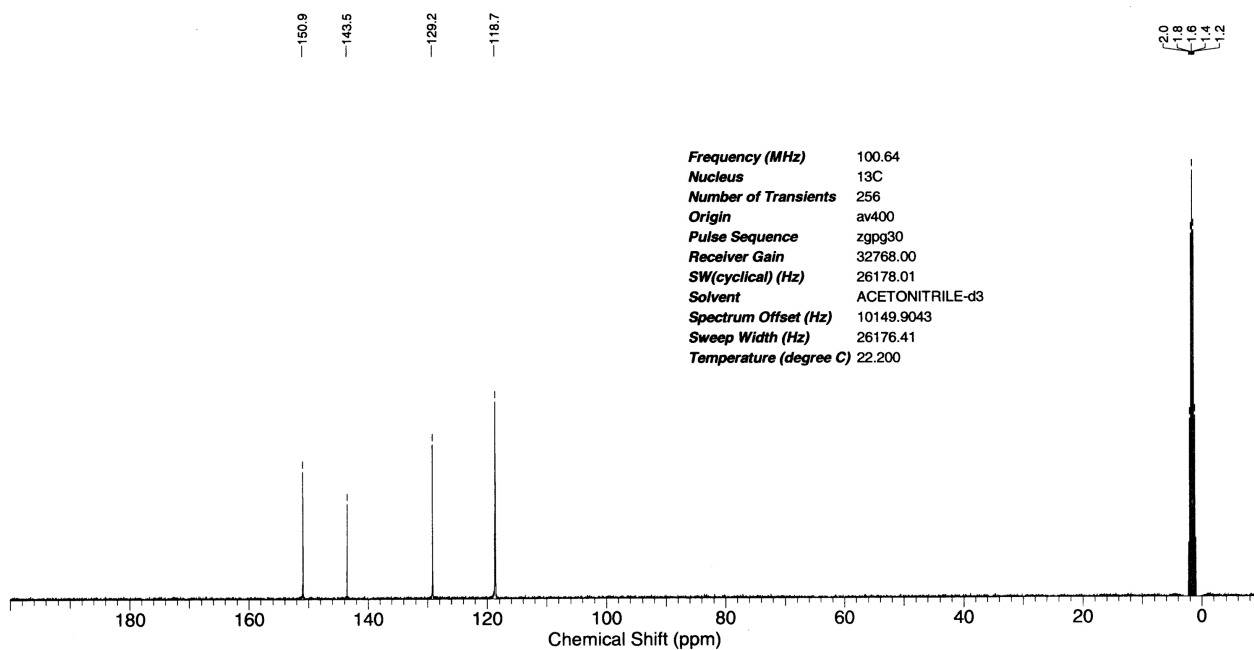


Barluenga's Reagent ^1H NMR (400 MHz, CD_3CN) (zoom)

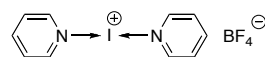




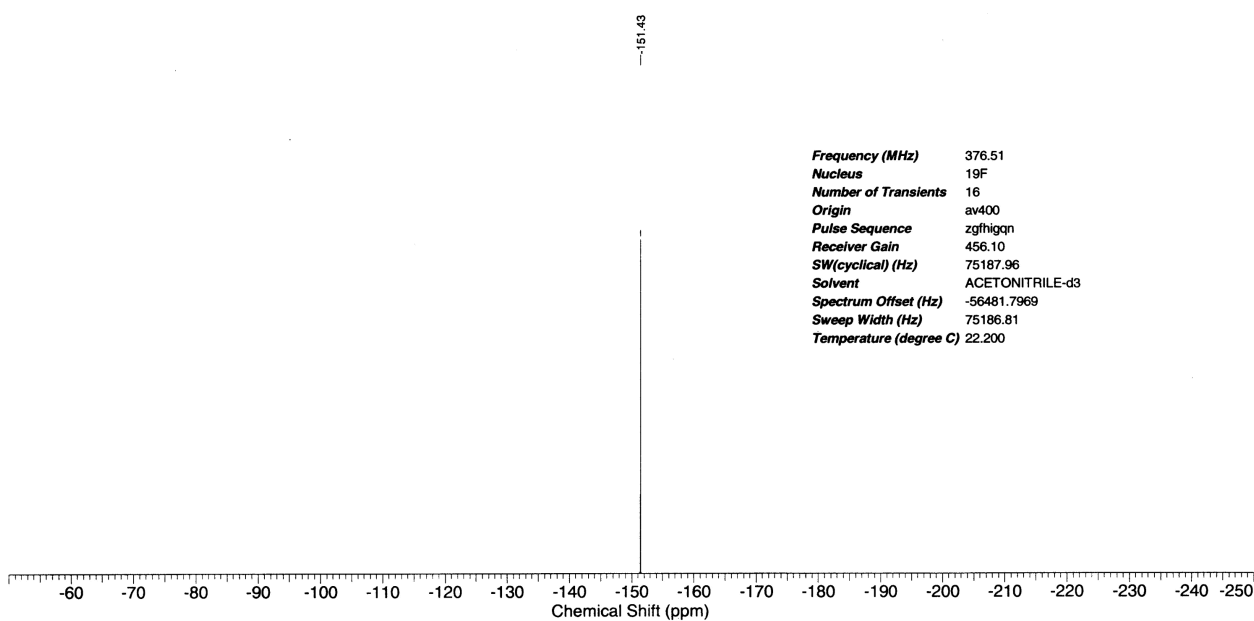
Barluenga's Reagent ^{13}C NMR (100 MHz, CD_3CN)



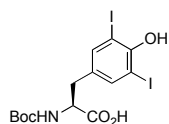
Frequency (MHz) 100.64
 Nucleus ^{13}C
 Number of Transients 256
 Origin av400
 Pulse Sequence zgpg30
 Receiver Gain 32768.00
 SW(cyclical) (Hz) 26178.01
 Solvent ACETONITRILE-d3
 Spectrum Offset (Hz) 10149.9043
 Sweep Width (Hz) 26176.41
 Temperature (degree C) 22.200



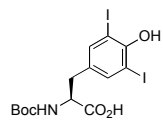
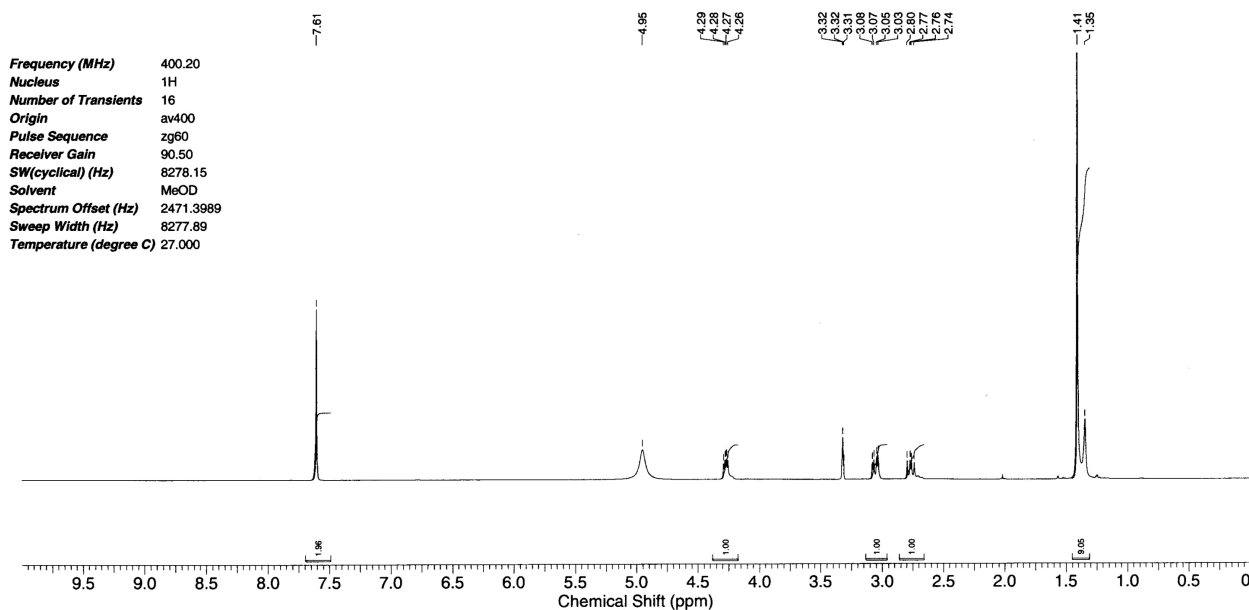
Barluenga's Reagent ^{19}F NMR (377 MHz, CD_3CN)



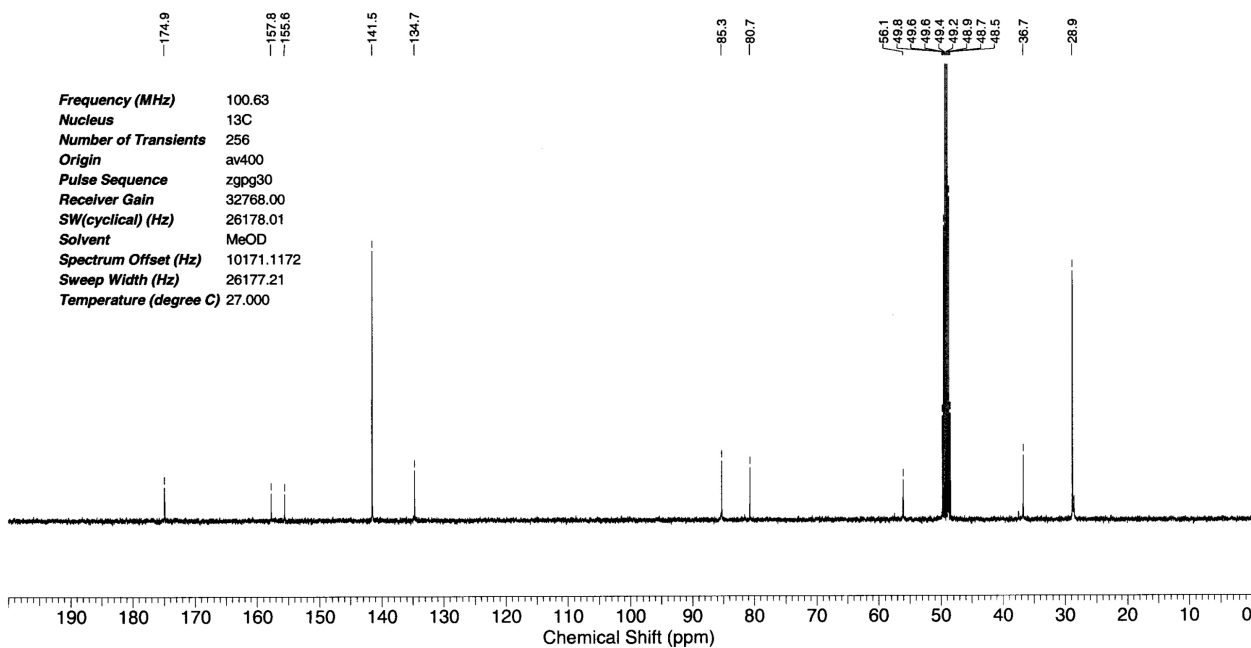
Frequency (MHz) 376.51
 Nucleus ^{19}F
 Number of Transients 16
 Origin av400
 Pulse Sequence zgfhgqn
 Receiver Gain 456.10
 SW(cyclical) (Hz) 75187.96
 Solvent ACETONITRILE-d3
 Spectrum Offset (Hz) -56481.7969
 Sweep Width (Hz) 75186.81
 Temperature (degree C) 22.200

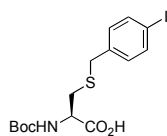


¹H NMR (400 MHz, MeOD)

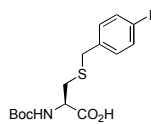
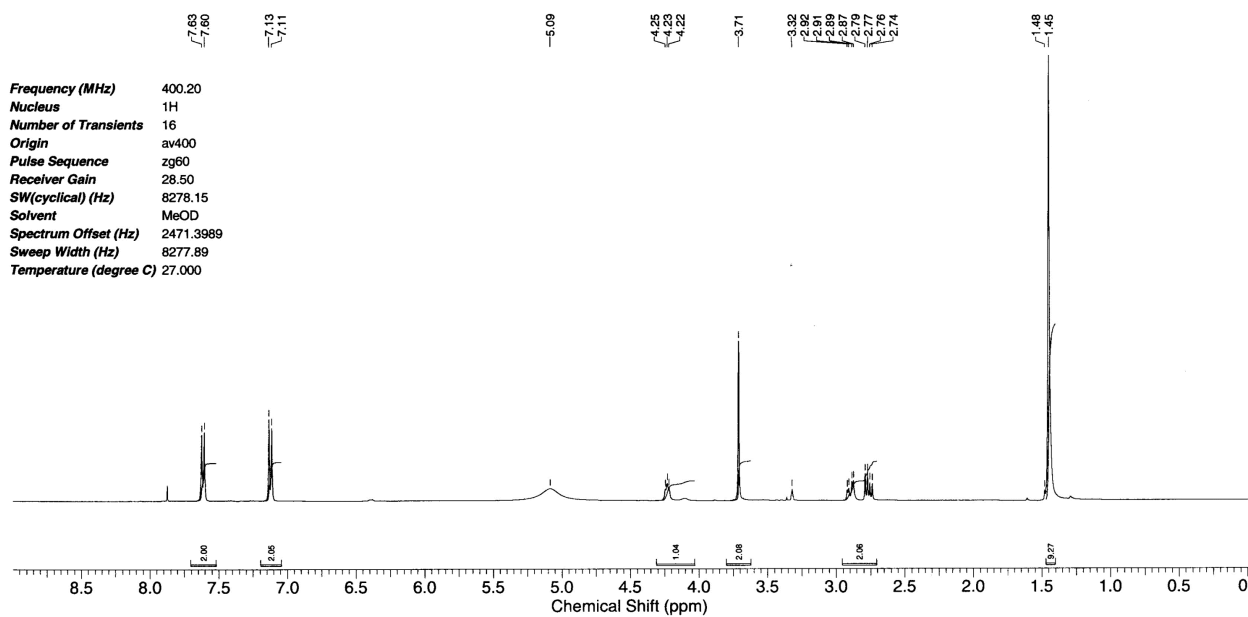


¹³C NMR (100 MHz, MeOD)

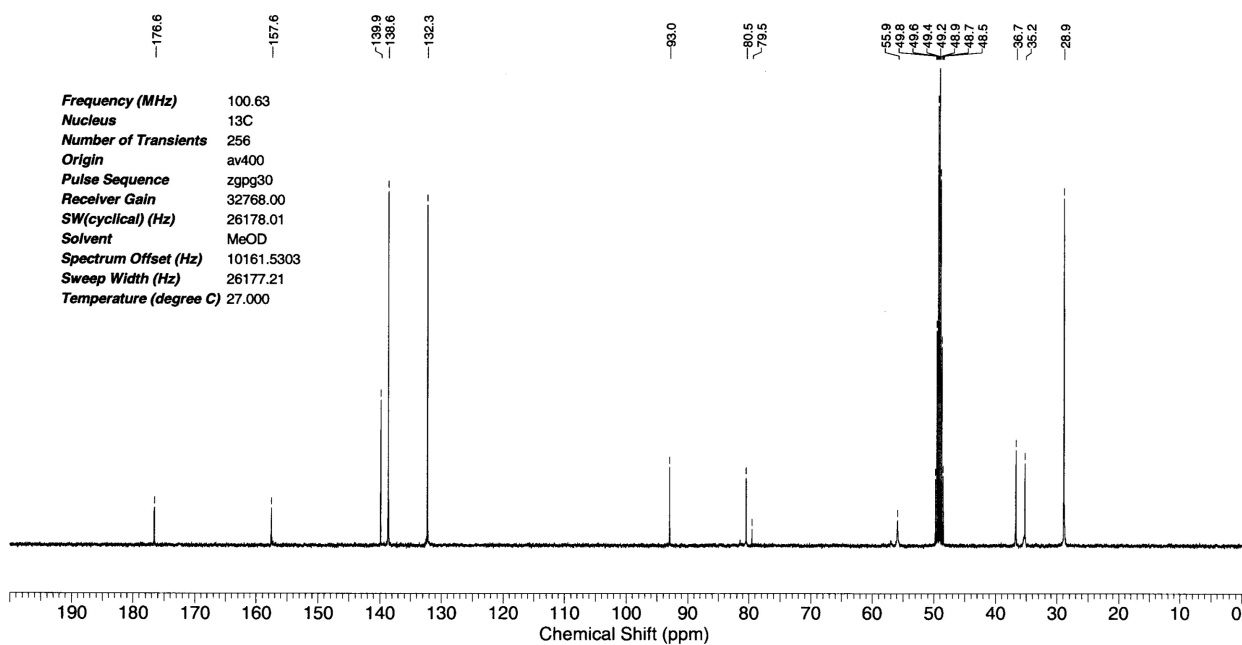


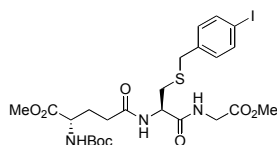


¹H NMR (400 MHz, MeOD)

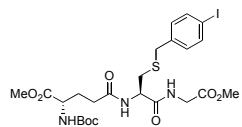
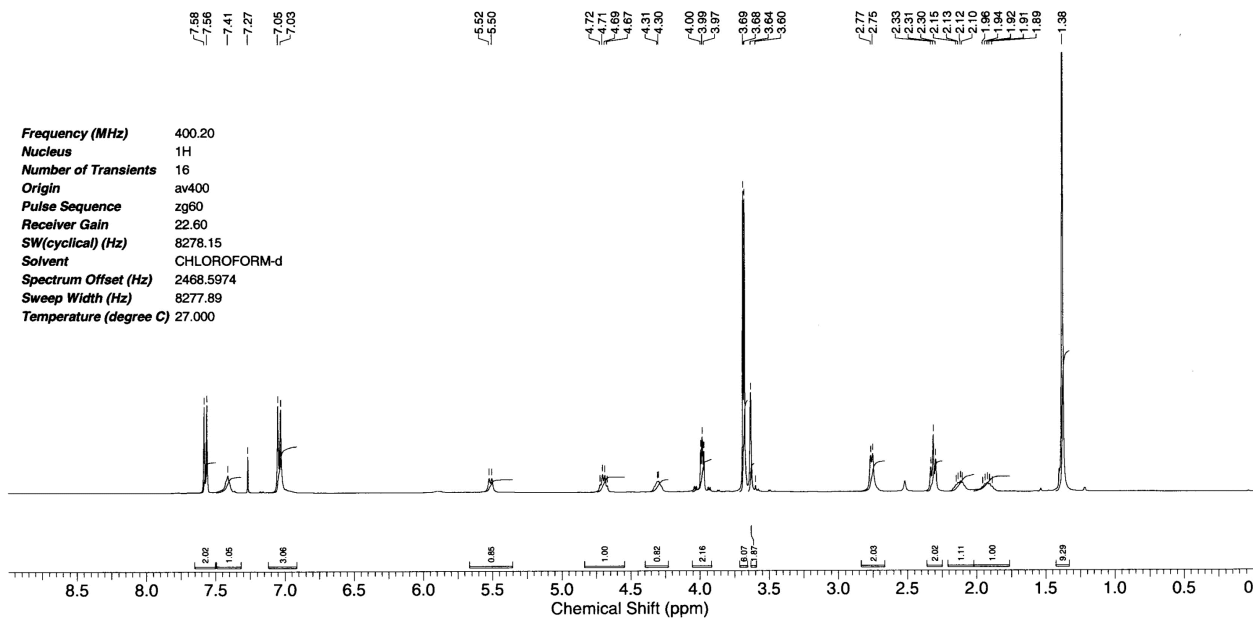


¹³C NMR (100 MHz, MeOD)

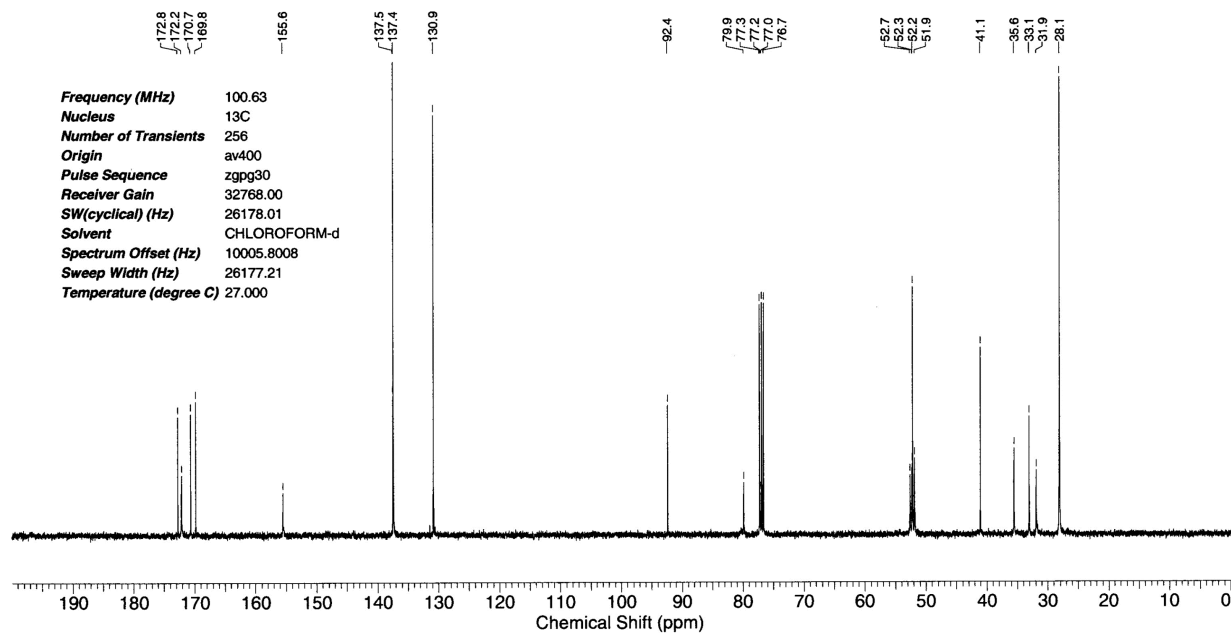


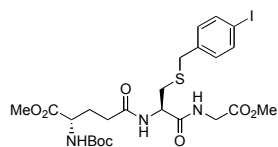


¹H NMR (400 MHz, CDCl₃)

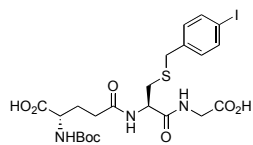
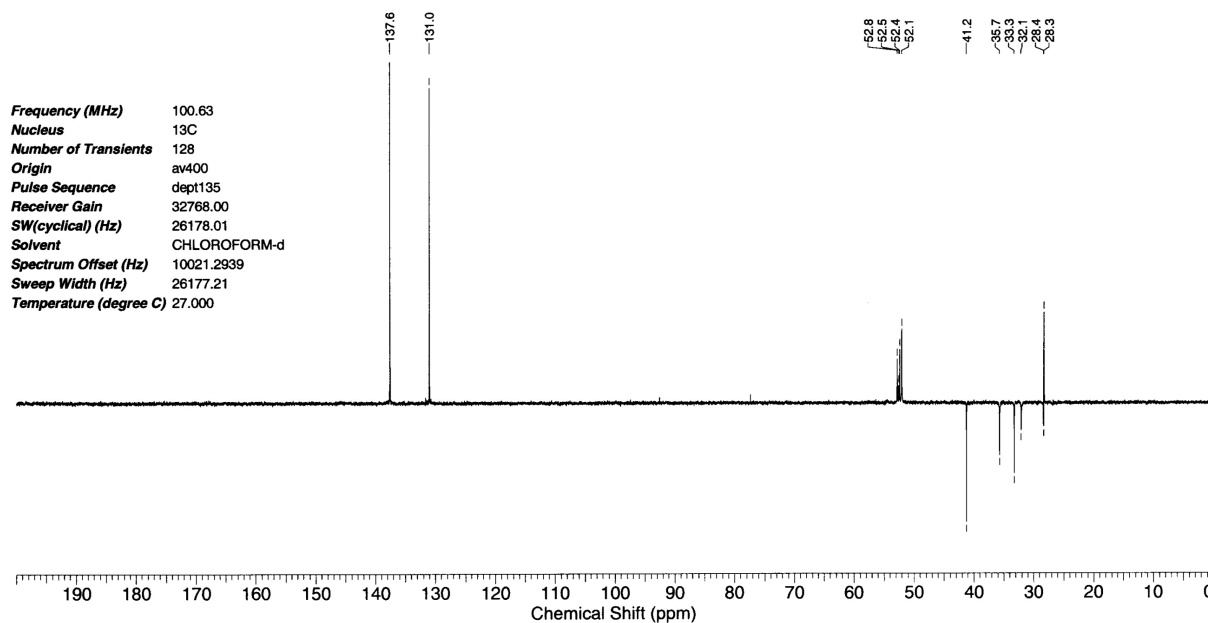


¹³C NMR (100 MHz, CDCl₃)

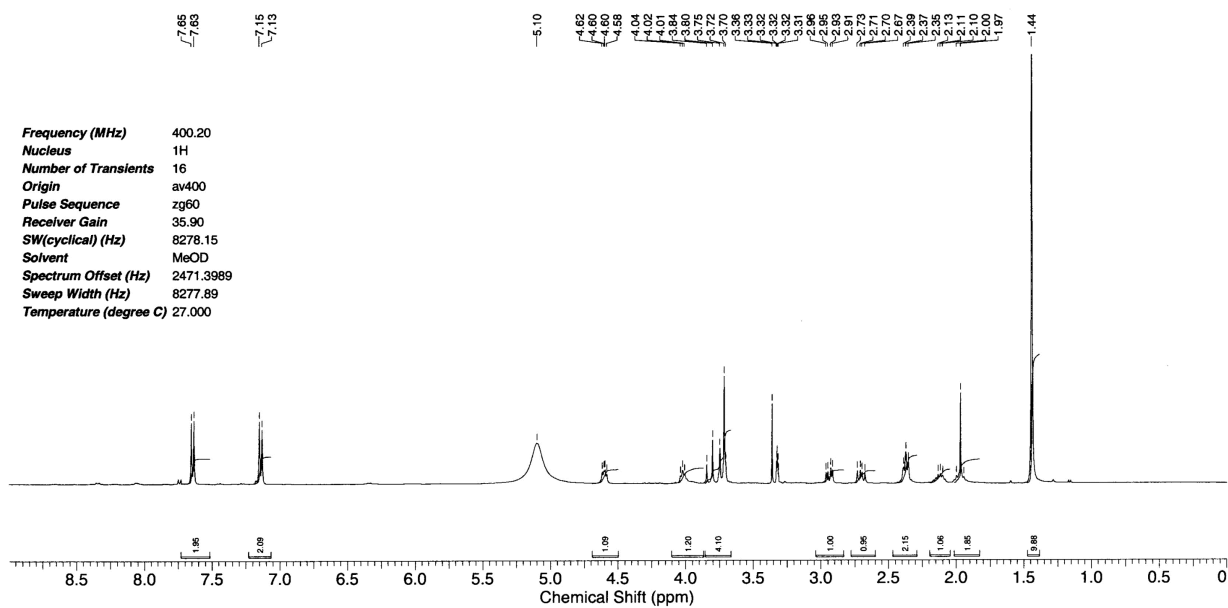


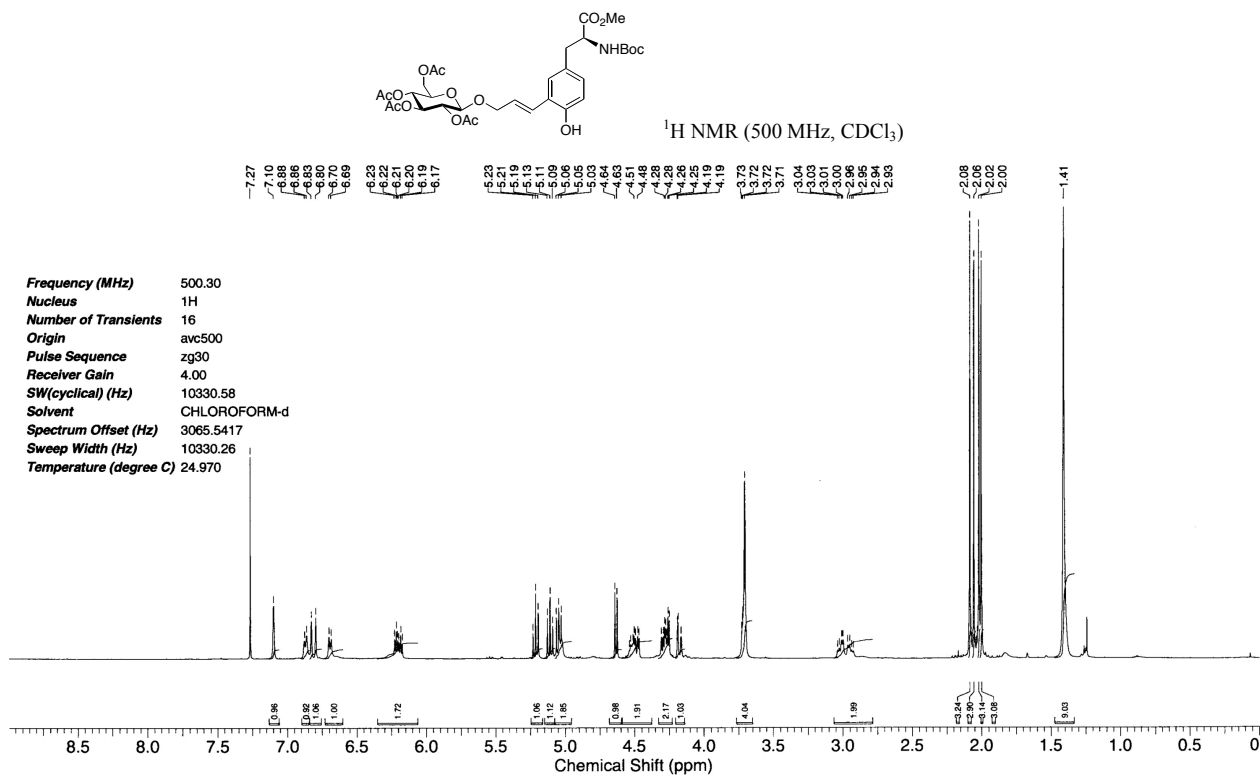
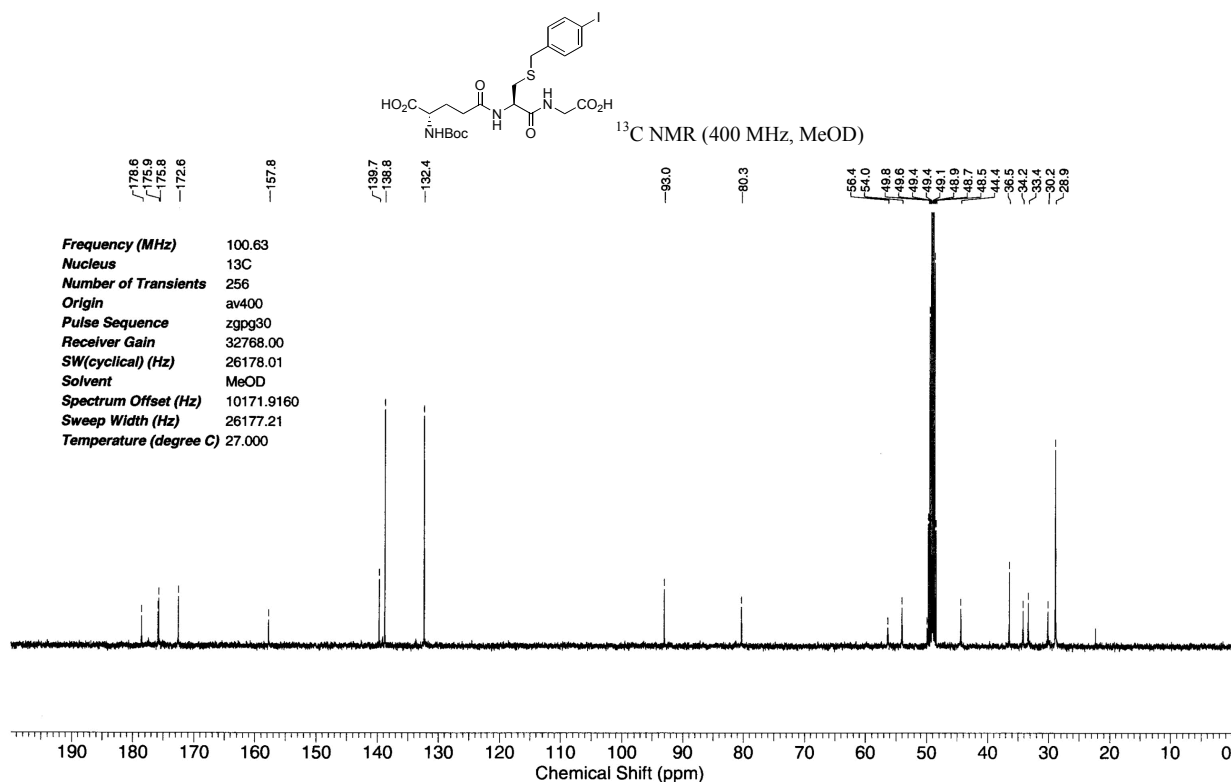


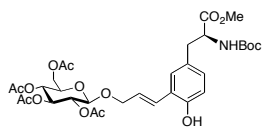
DEPT 135 (100 MHz, CDCl₃)



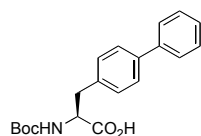
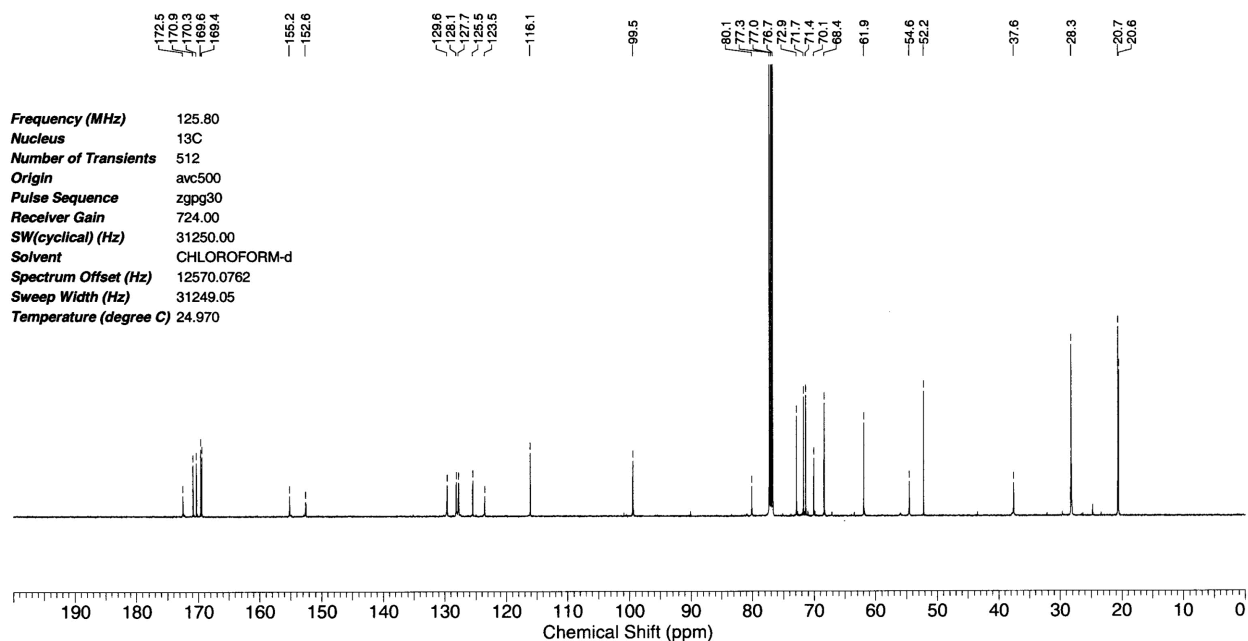
¹H NMR (400 MHz, MeOD)



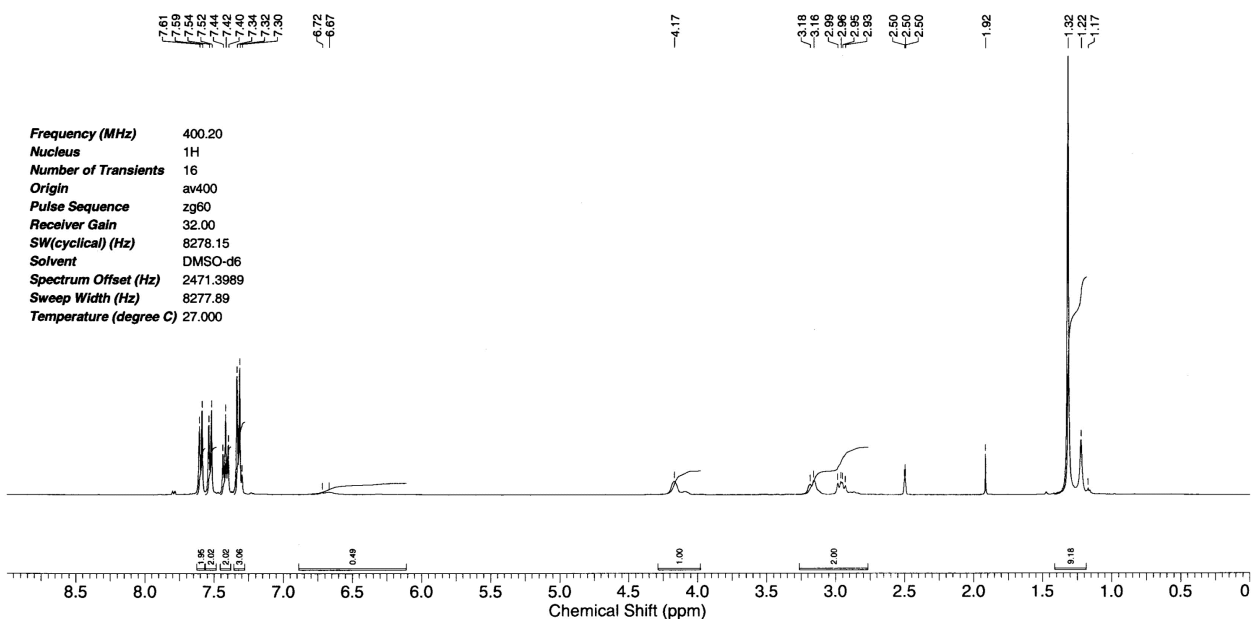


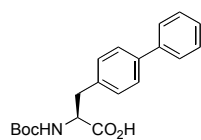


^{13}C NMR (125 MHz, CDCl_3)

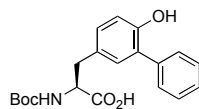
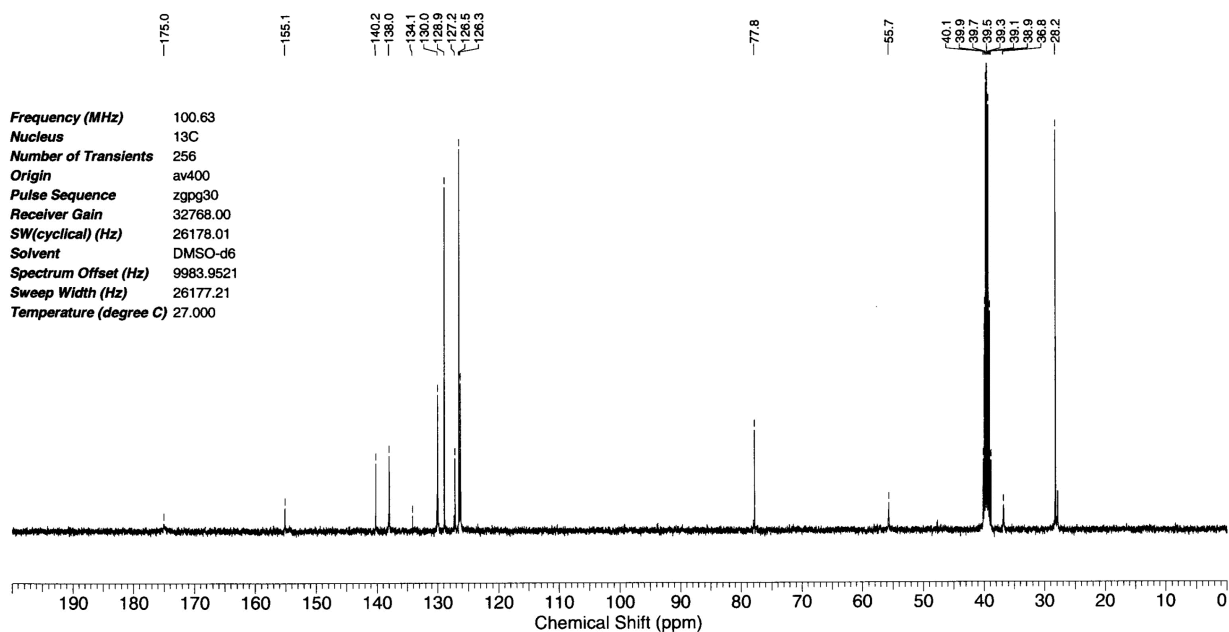


(Table 6.2, Entry 1 and 2) ^1H NMR (400 MHz, DMSO)

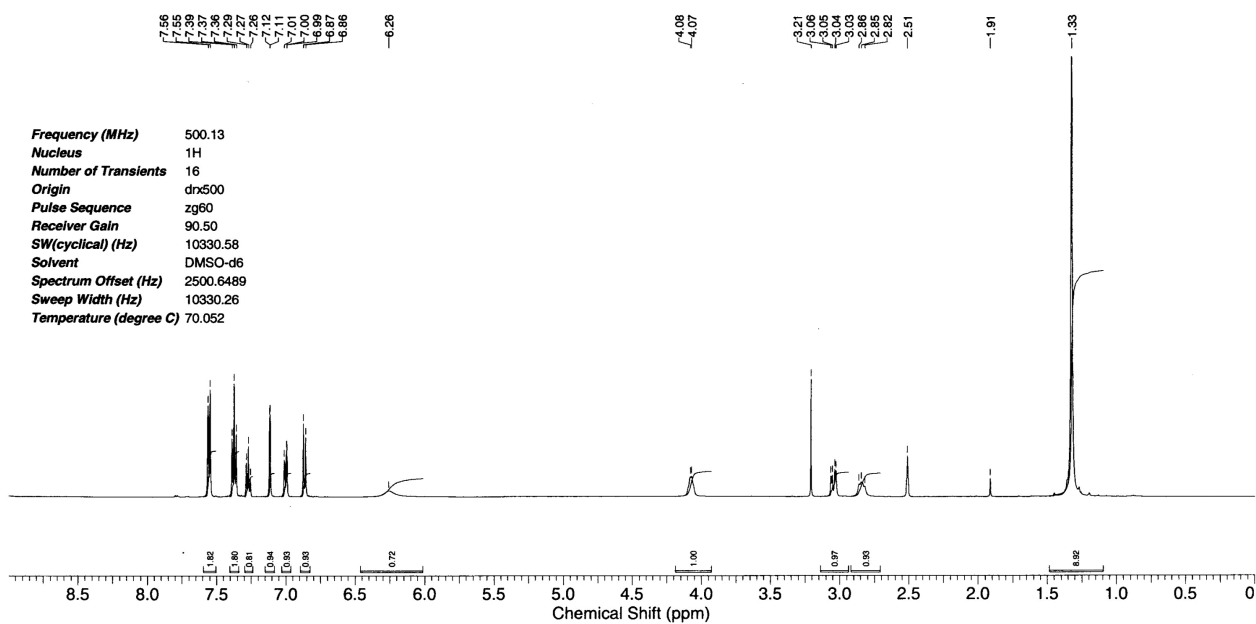


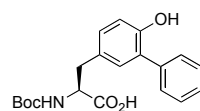


(Table 6.2, Entry 1 and 2) ^{13}C NMR (100 MHz, DMSO)

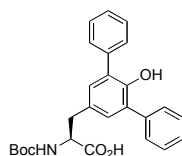
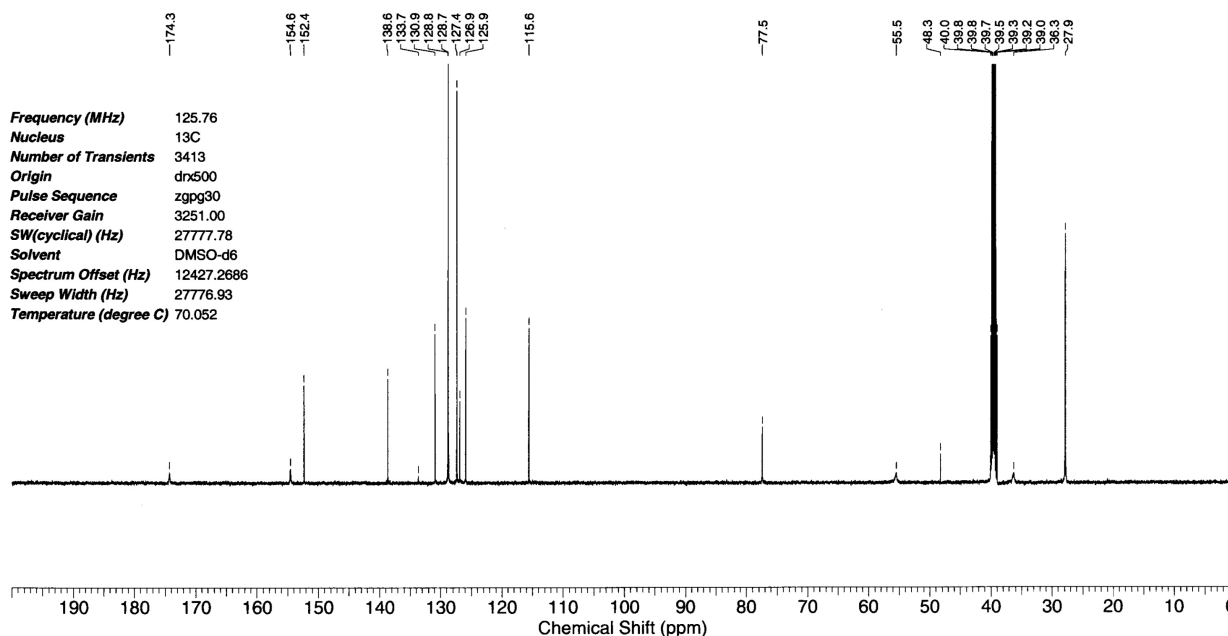


(Table 6.2, Entry 4) ^1H NMR (500 MHz, DMSO, 70 °C)

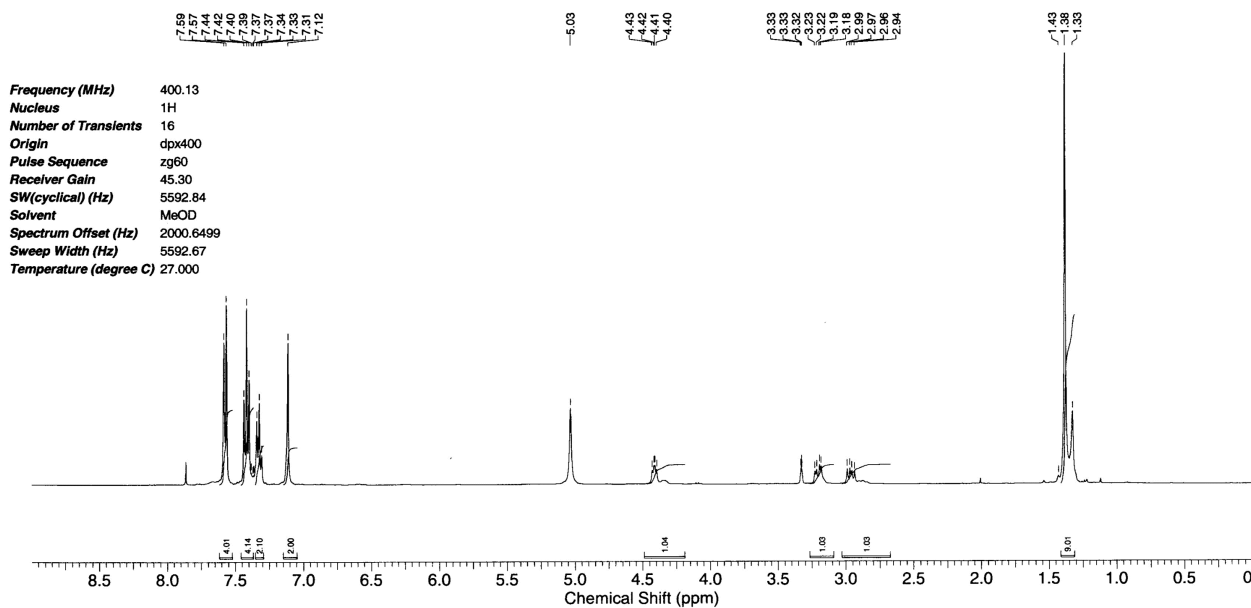


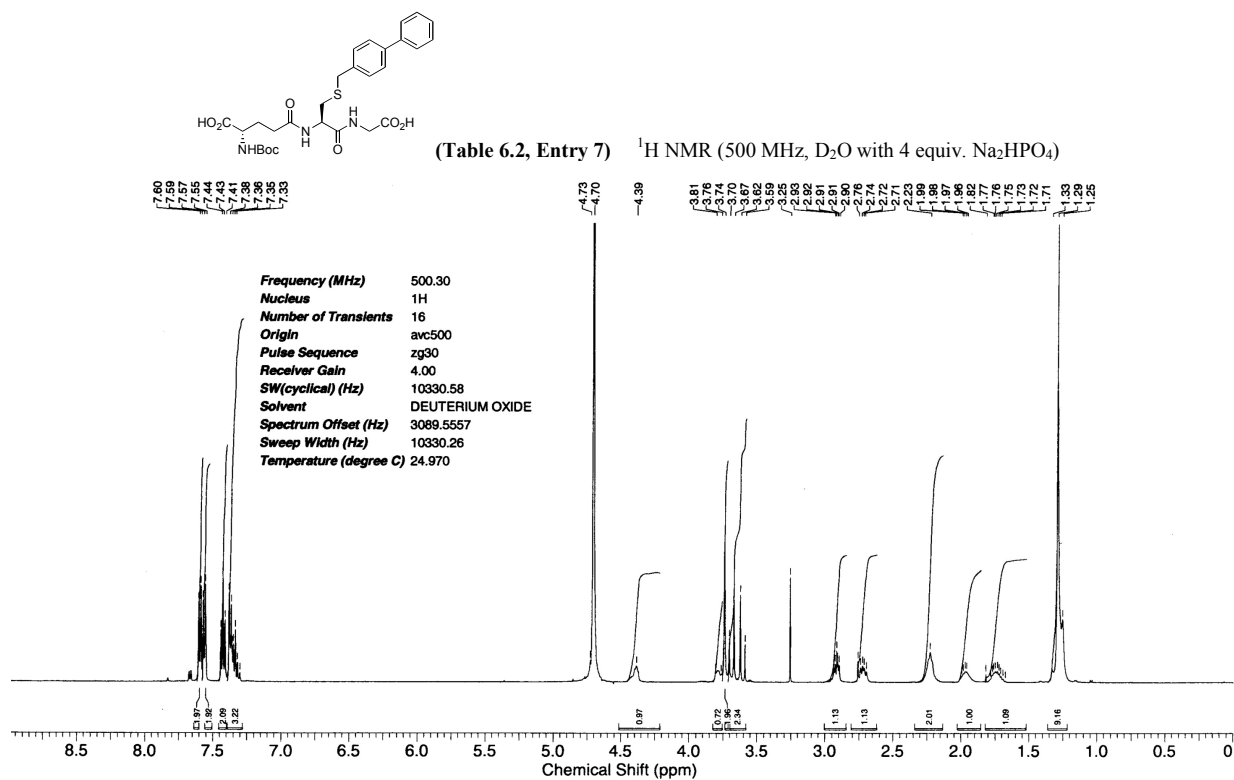
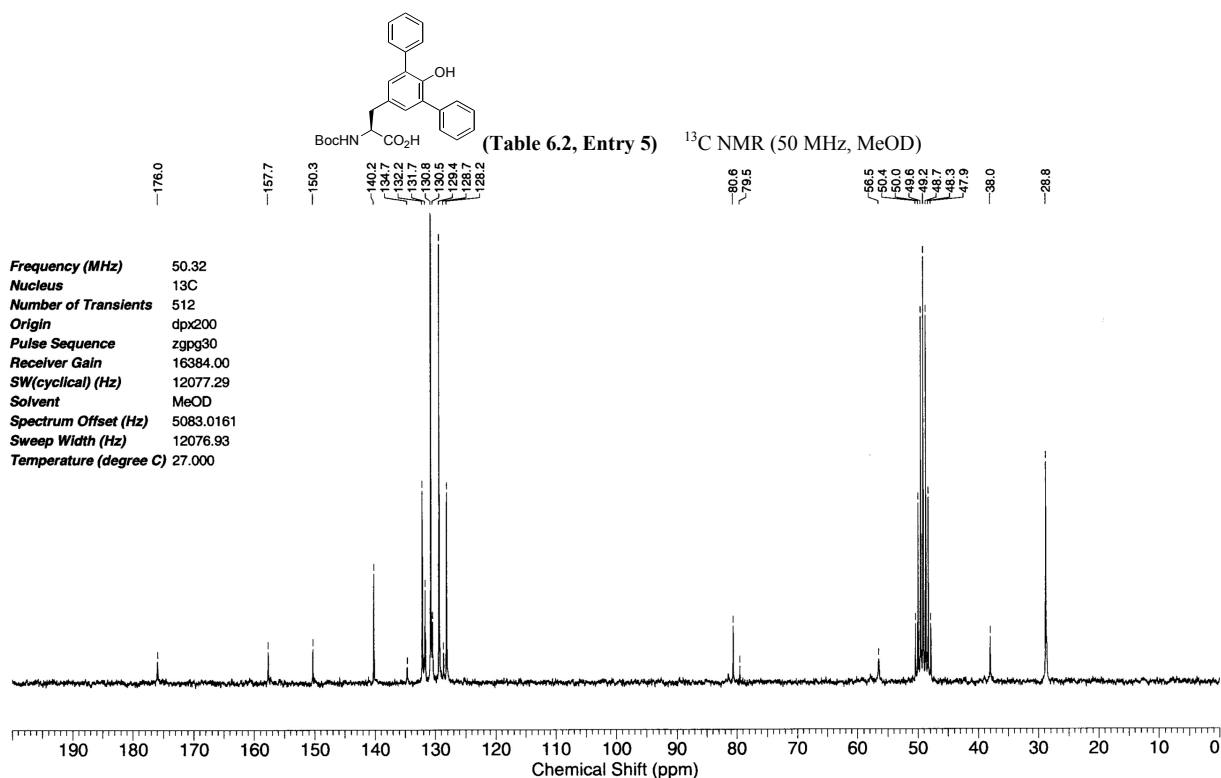


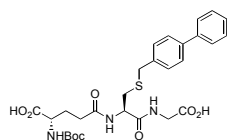
(Table 6.2, Entry 4) ^{13}C NMR (125 MHz, DMSO, 70 °C)



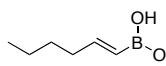
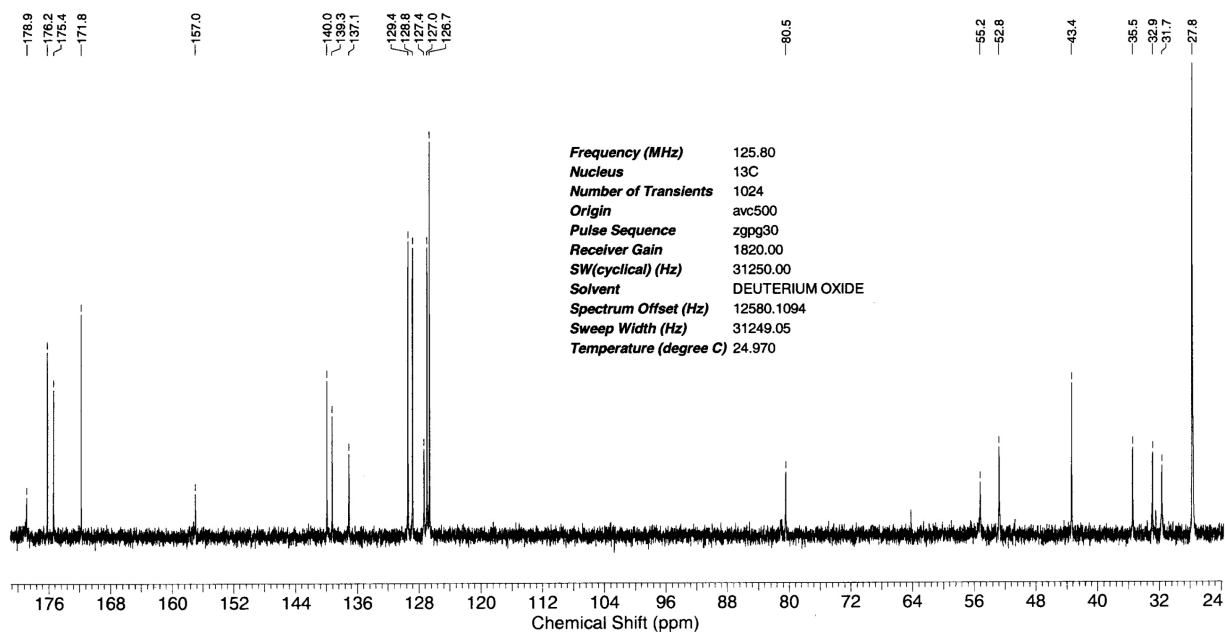
(Table 6.2, Entry 5) ^1H NMR (400 MHz, MeOD)



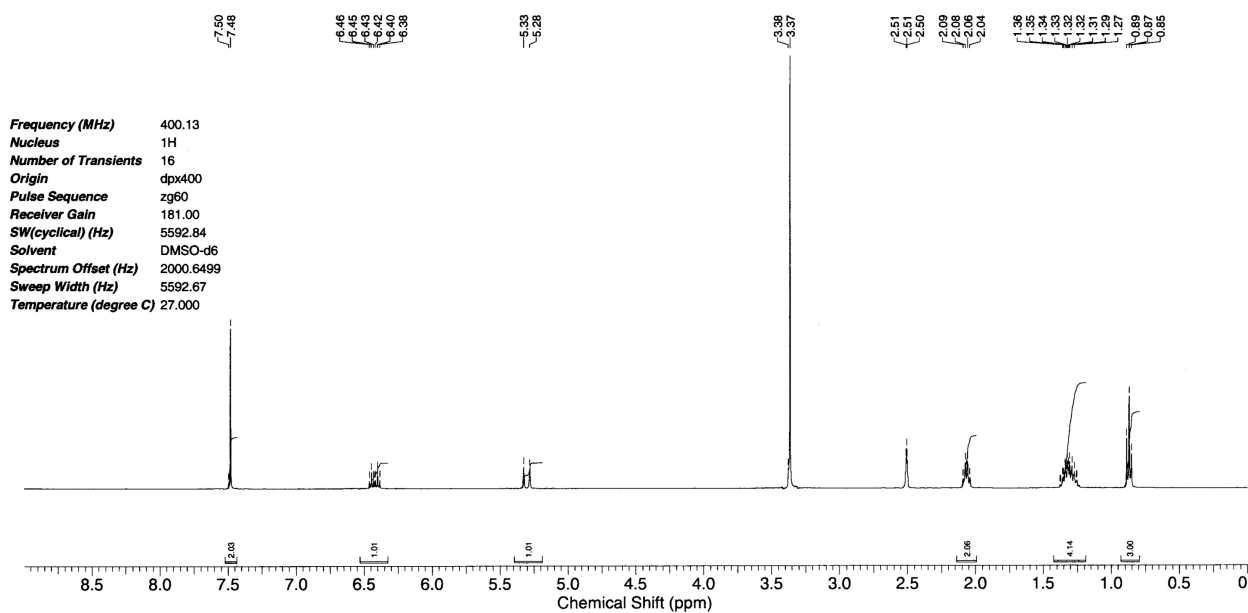


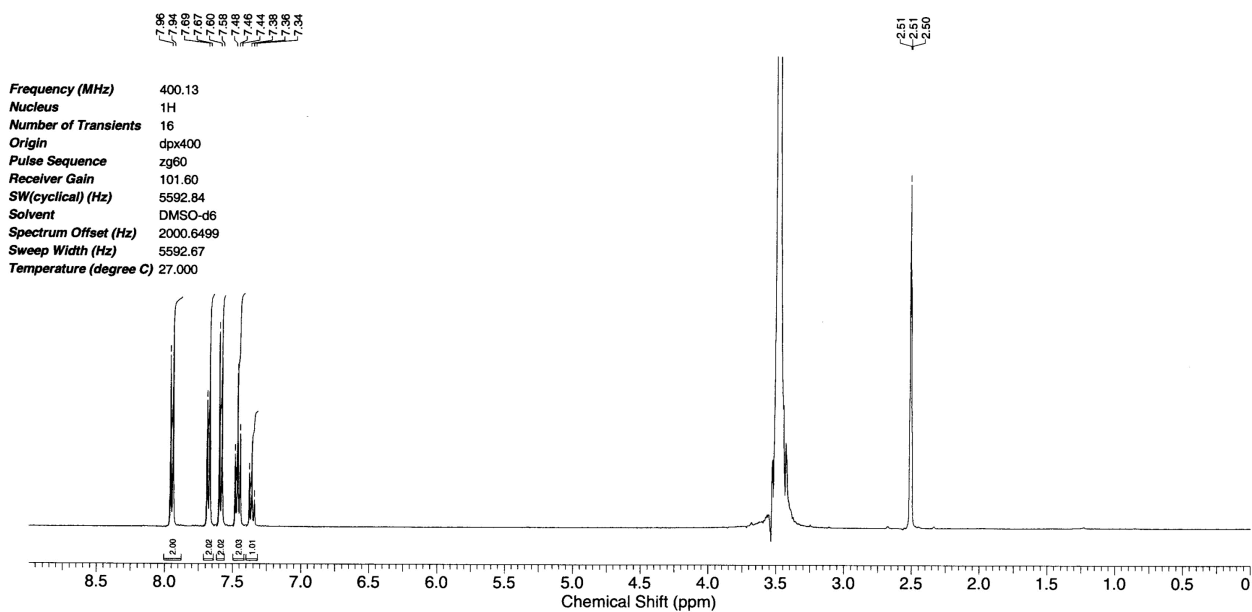
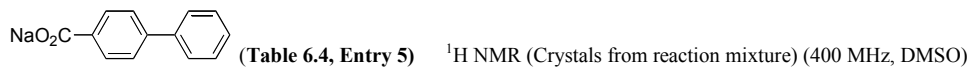
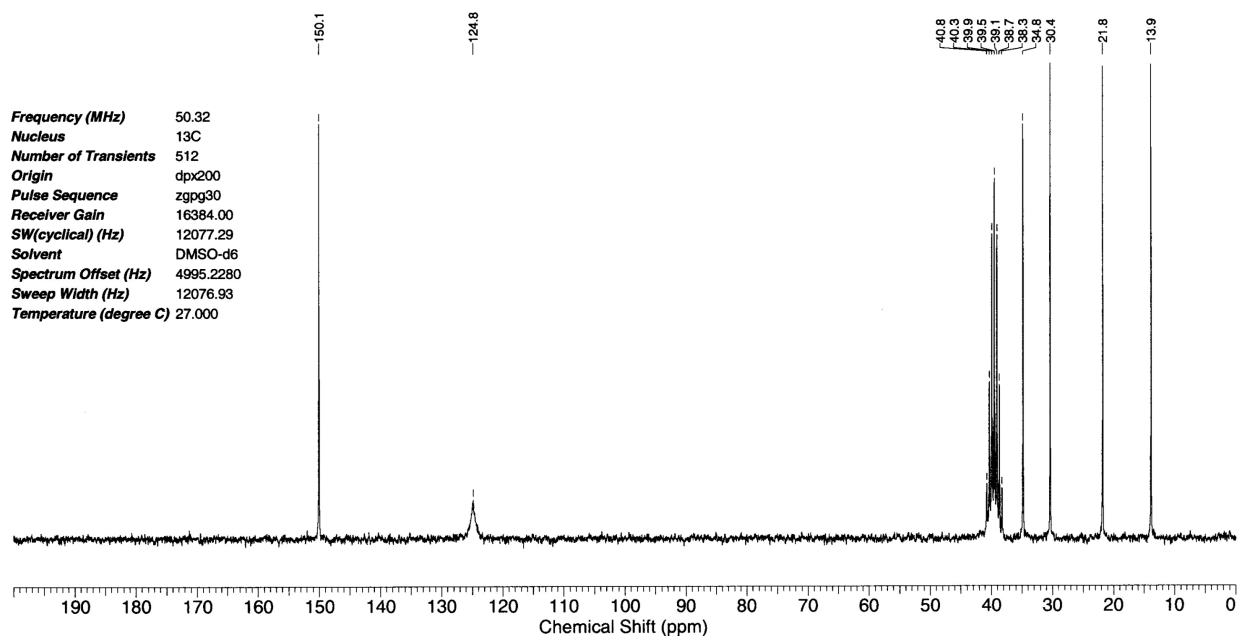
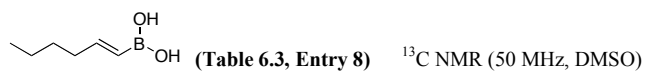


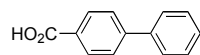
(Table 6.2, Entry 7) ^{13}C NMR (125 MHz, D_2O with 4 equiv. Na_2HPO_4)



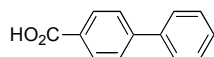
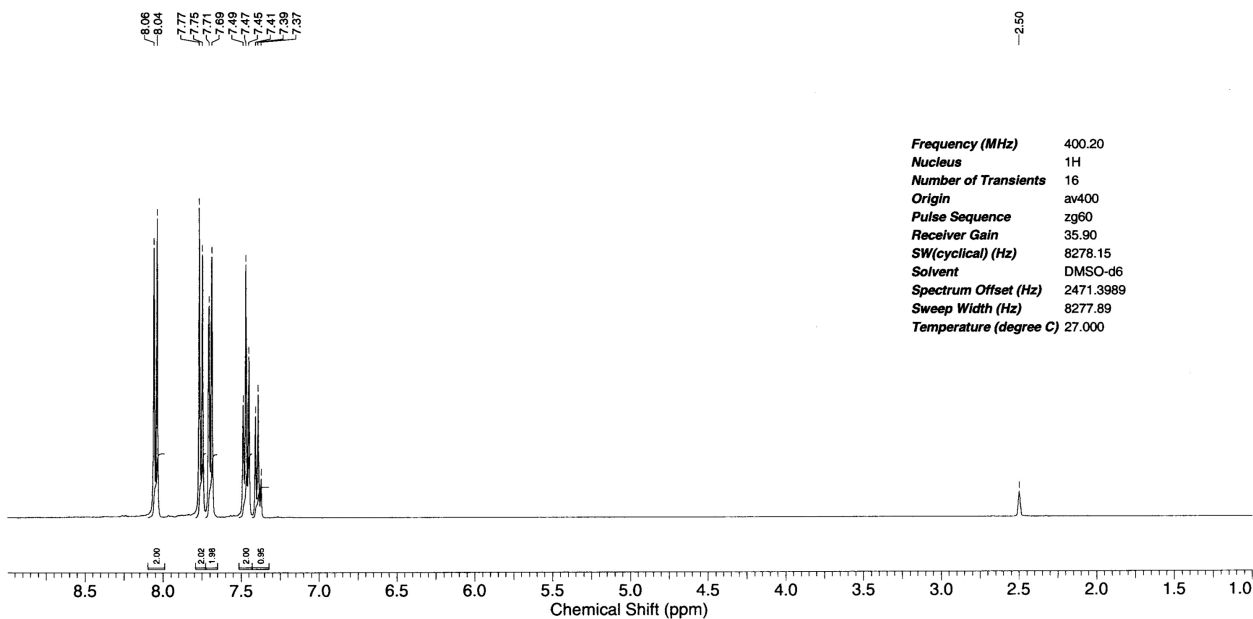
(Table 6.3, Entry 8) ^1H NMR (400 MHz, DMSO)



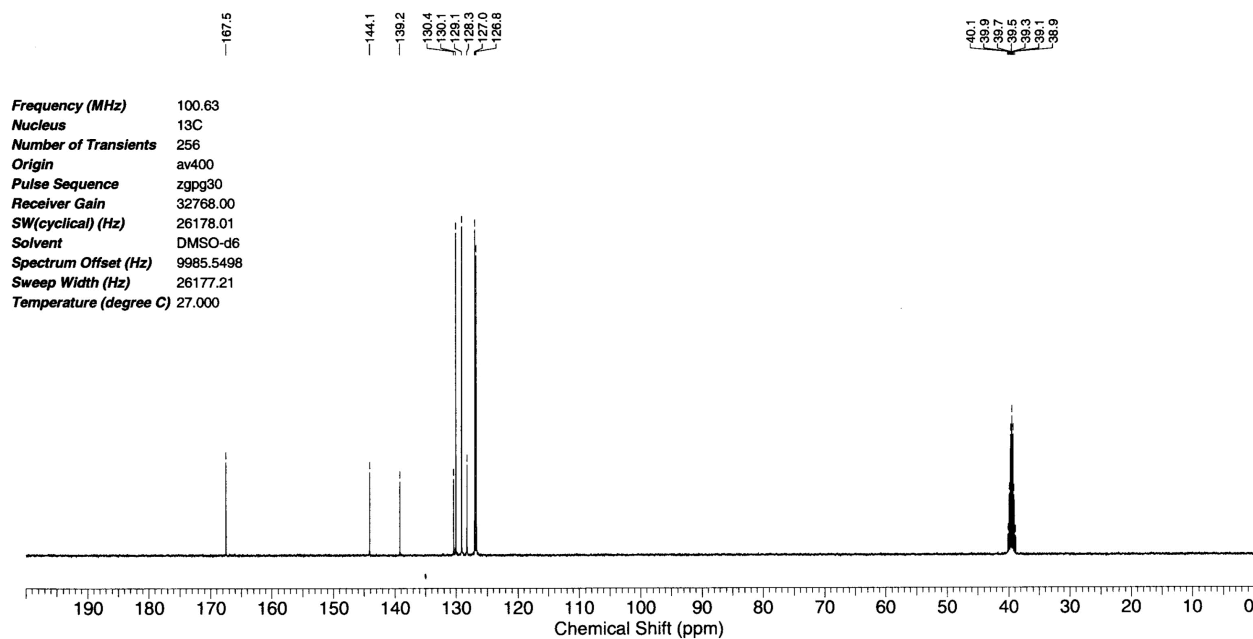


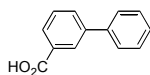


(Table 6.4, Entry 5) ^1H NMR (from 50 g scale) (400 MHz, DMSO)

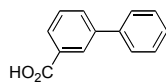
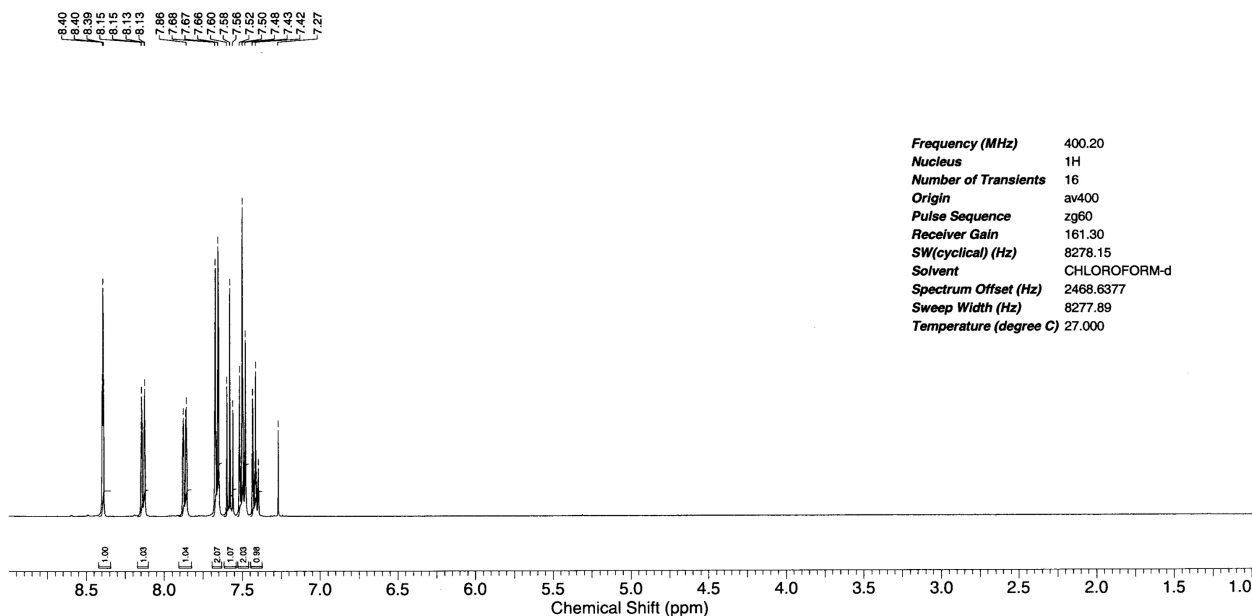


(Table 6.4, Entry 5) ^{13}C NMR (from 50 g scale) (100 MHz, DMSO)

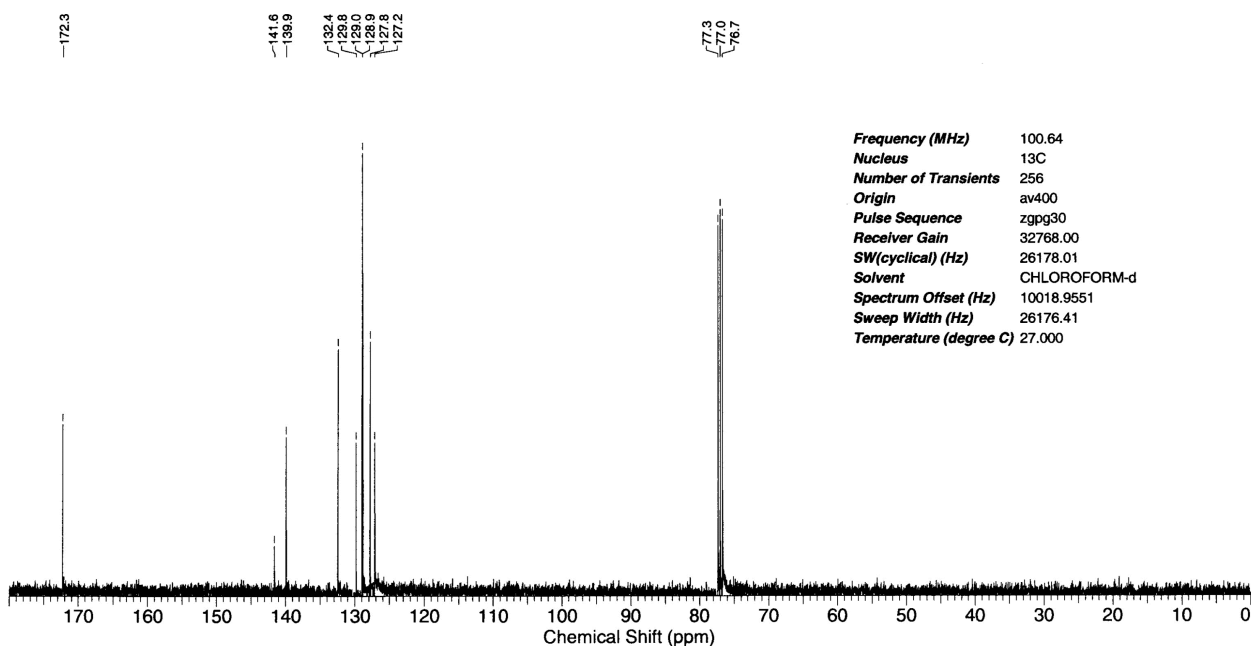


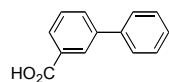


(Table 6.4, Entry 7) ^1H NMR (400 MHz, CDCl_3)

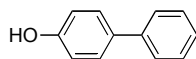
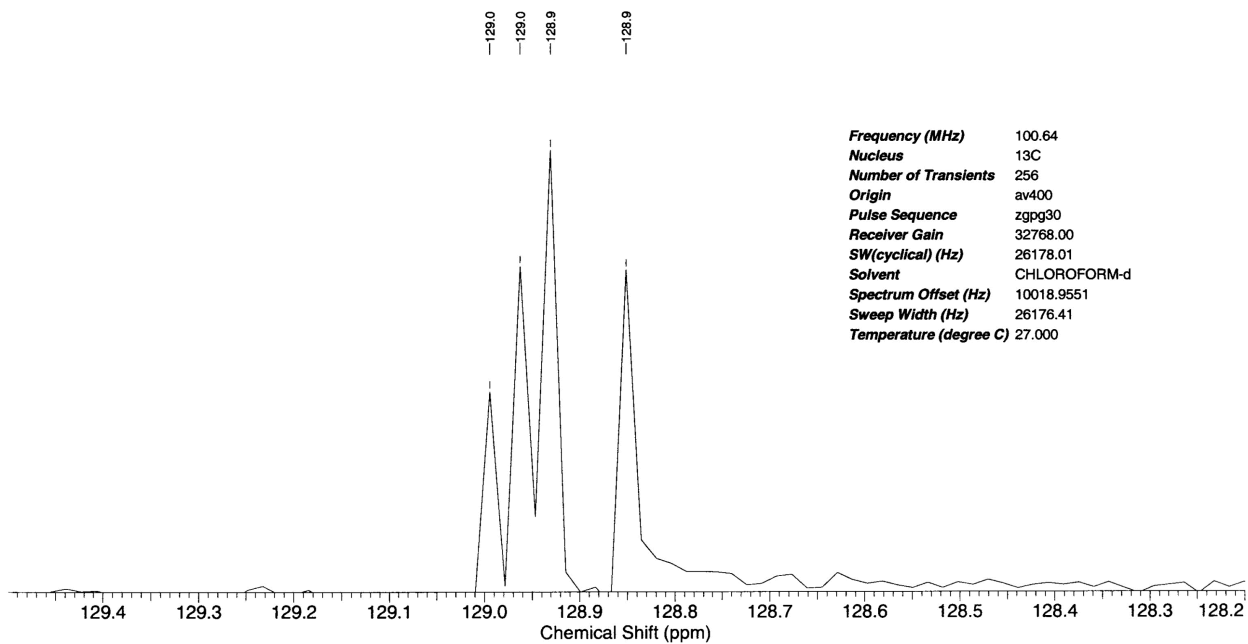


(Table 6.4, Entry 7) ^{13}C NMR (100 MHz, CDCl_3)

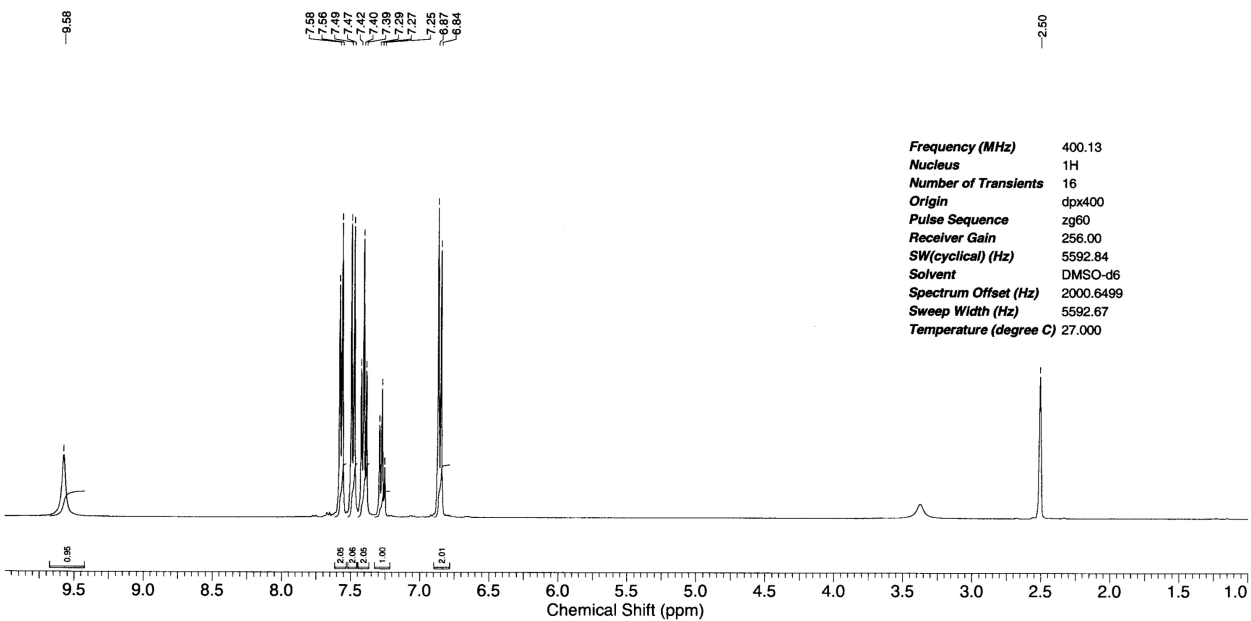


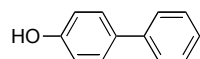


(Table 6.4, Entry 7) ^{13}C NMR (100 MHz, CDCl_3 , zoom)



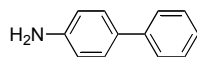
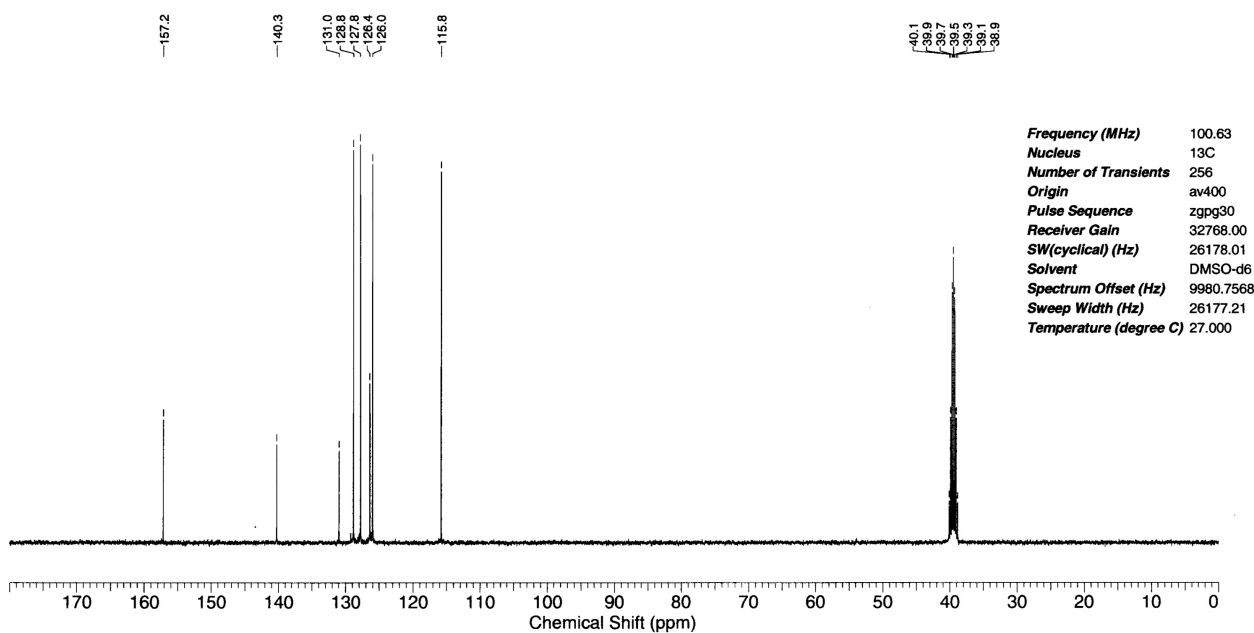
(Table 6.4, Entry 8) ^1H NMR (400 MHz, DMSO)





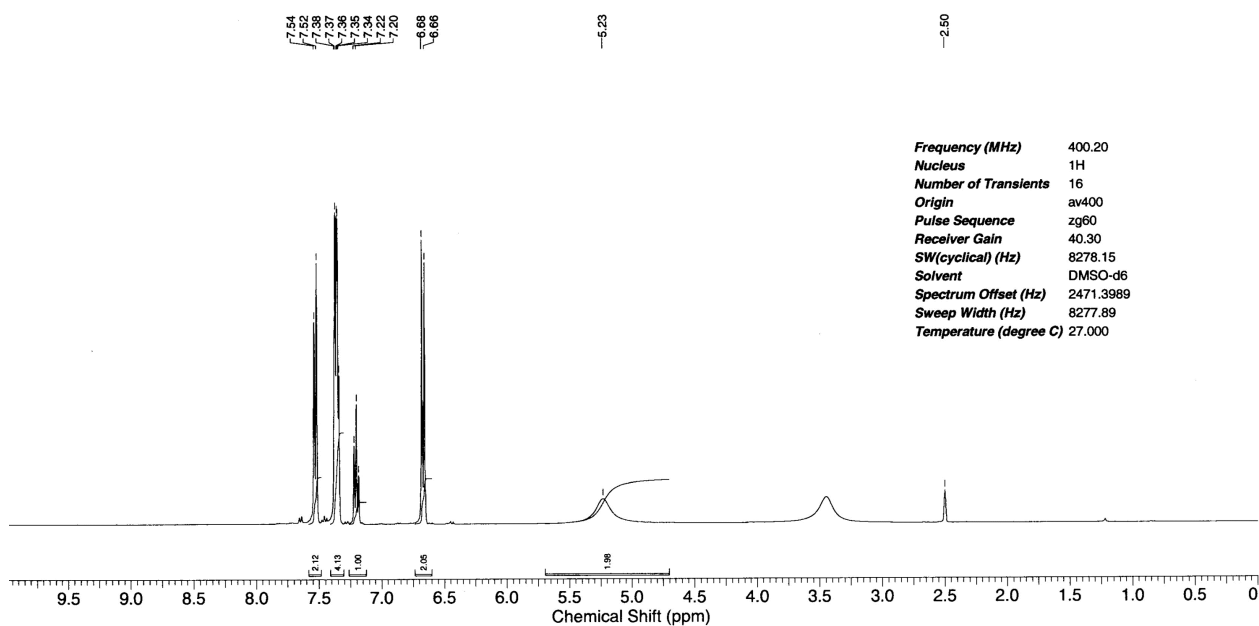
(Table 6.4, Entry 8)

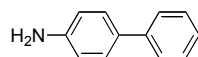
^{13}C NMR (100 MHz, DMSO)



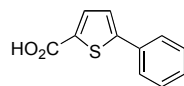
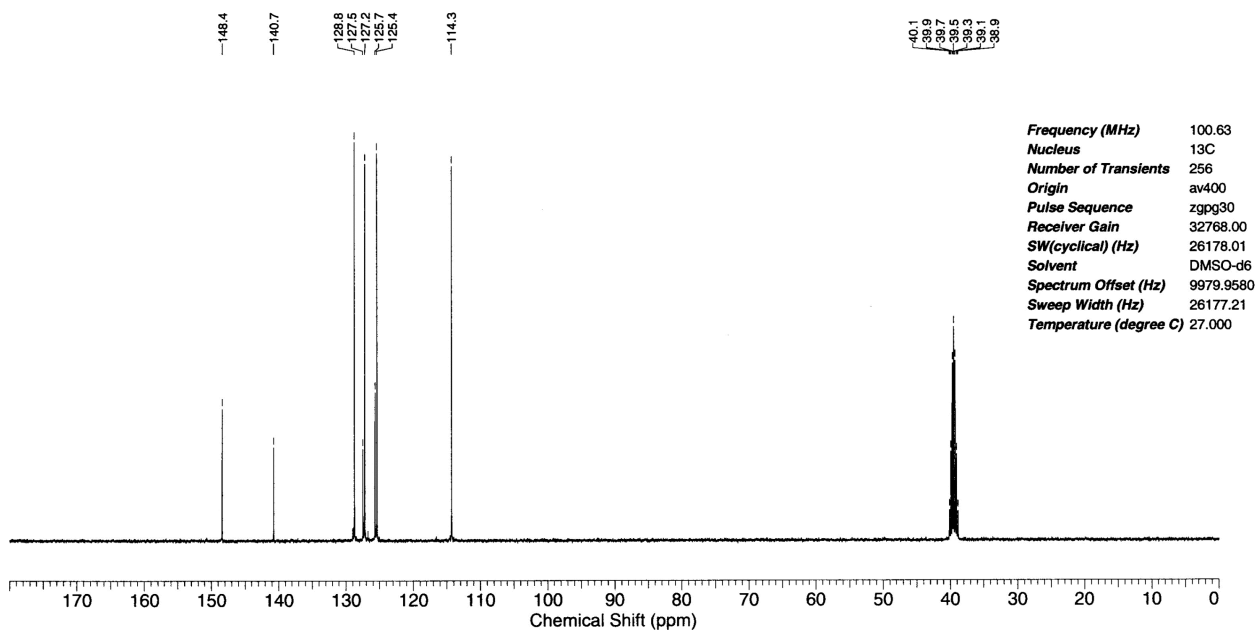
(Table 6.4, Entry 9)

^1H NMR (400 MHz, DMSO)

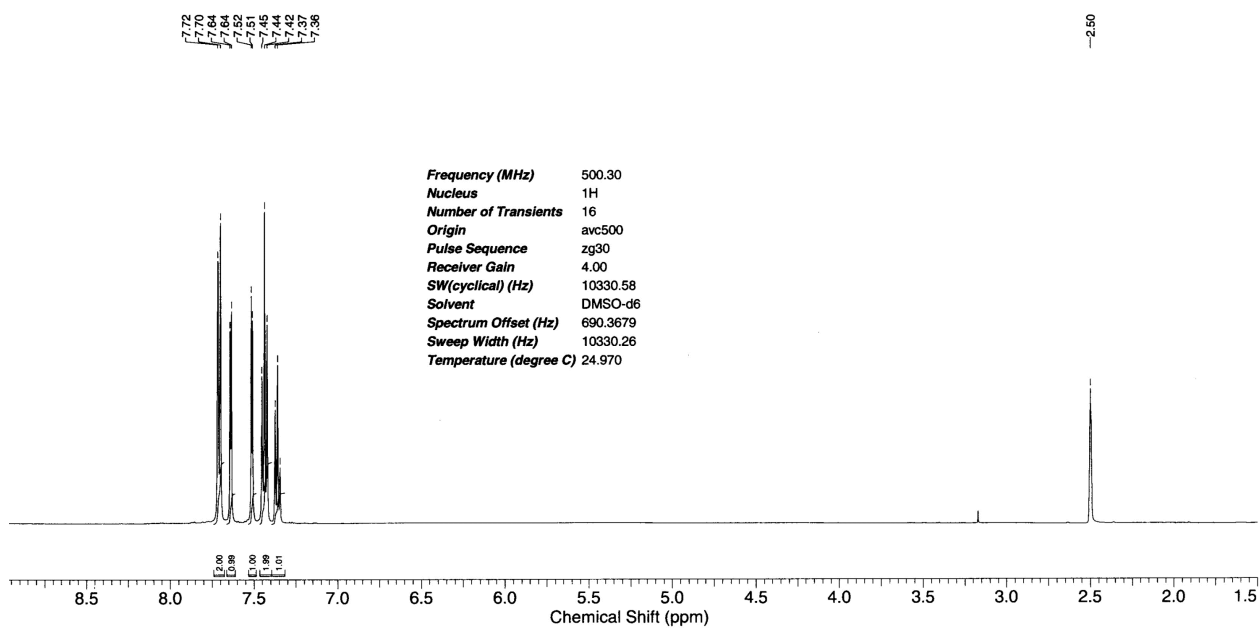


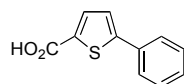


(Table 6.4, Entry 9) ^{13}C NMR (100 MHz, DMSO)

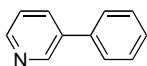
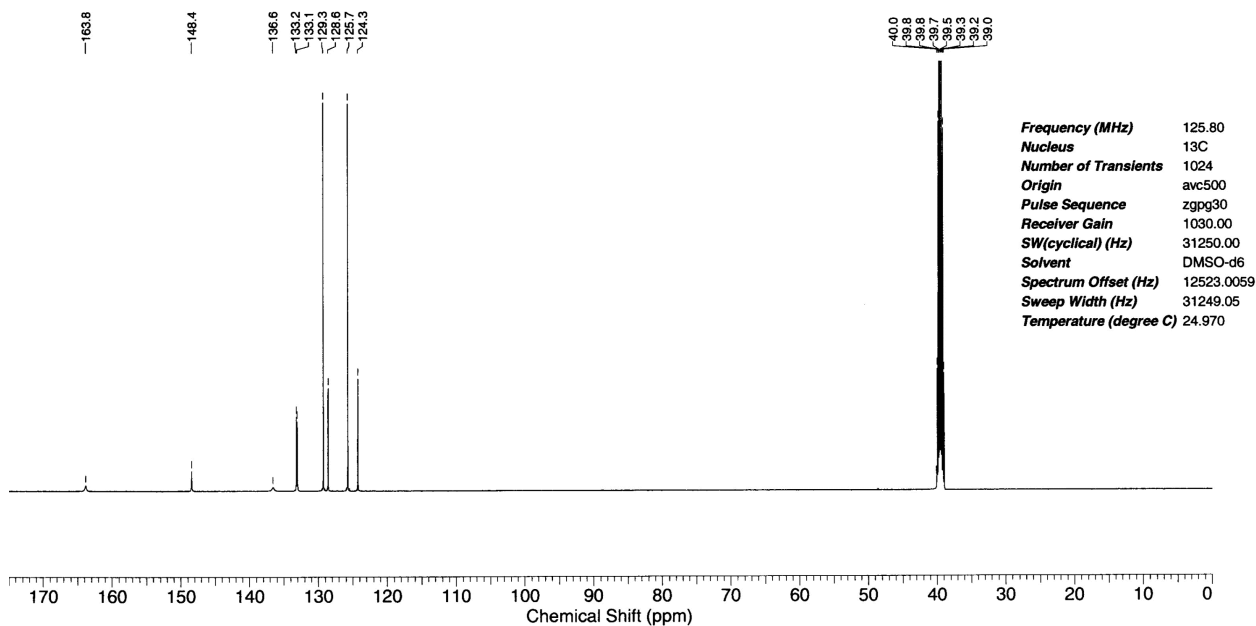


(Table 6.4, Entry 10) ^1H NMR (500 MHz, DMSO)

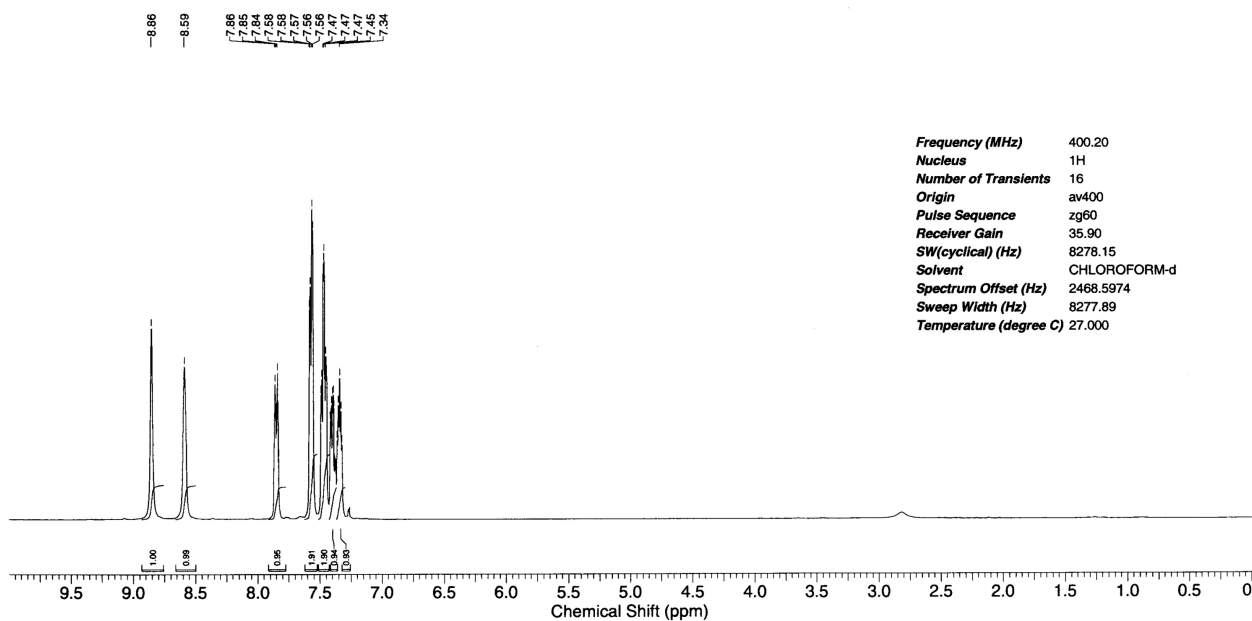


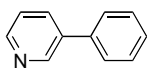


(Table 6.4, Entry 10) ^{13}C NMR (125 MHz, DMSO)

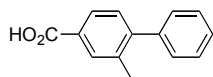
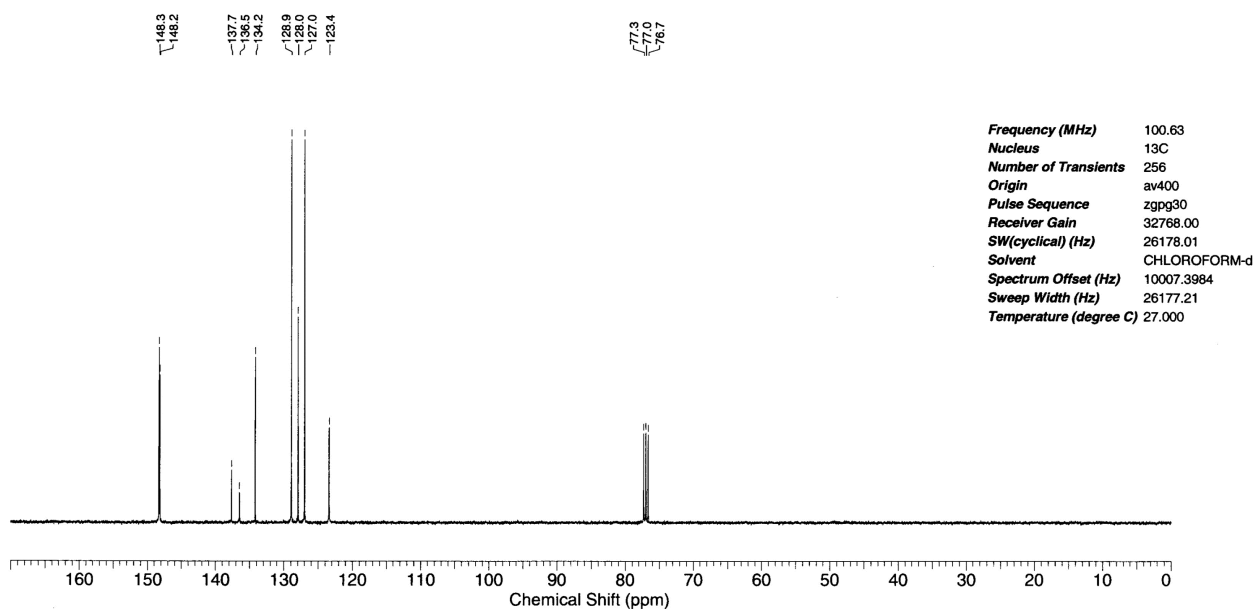


(Table 6.4, Entry 11) ^1H NMR (400 MHz, CDCl_3)

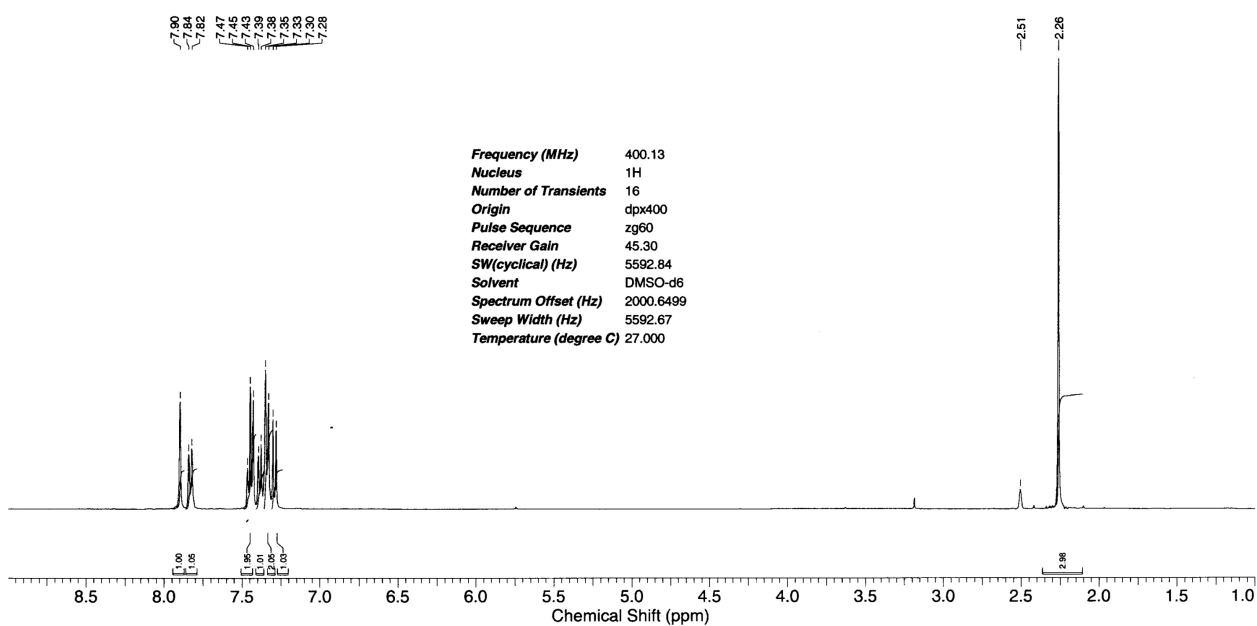


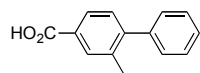


(Table 6.4, Entry 11) ^{13}C NMR (100 MHz, CDCl_3)

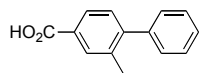
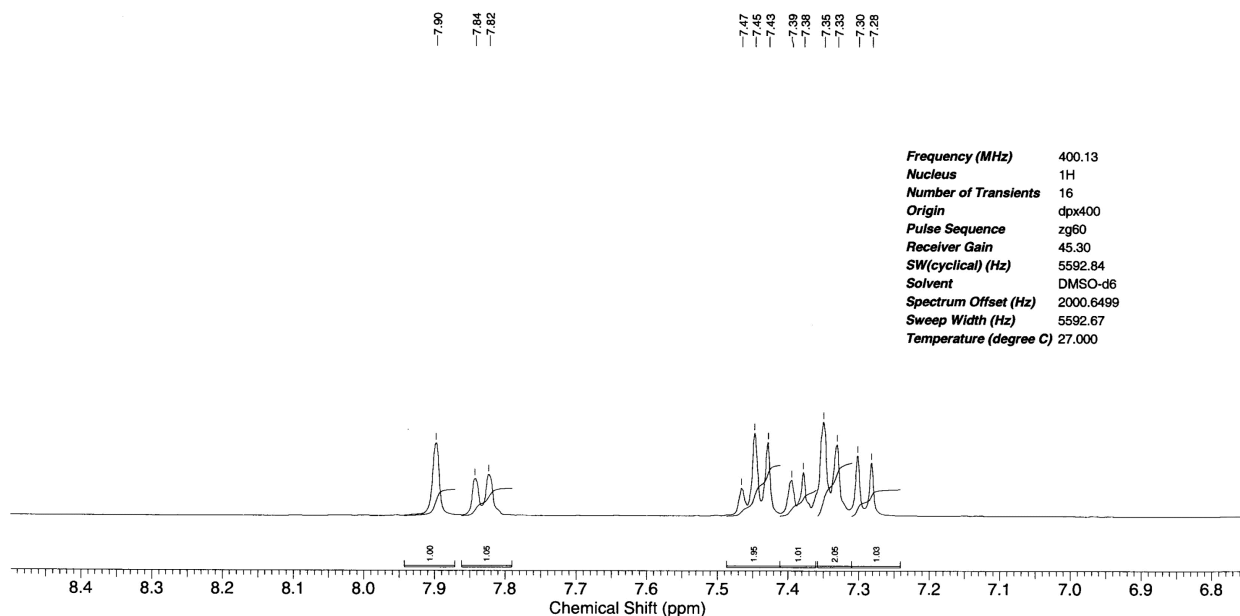


(Table 6.4, Entry 12) ^1H NMR (400 MHz, DMSO)

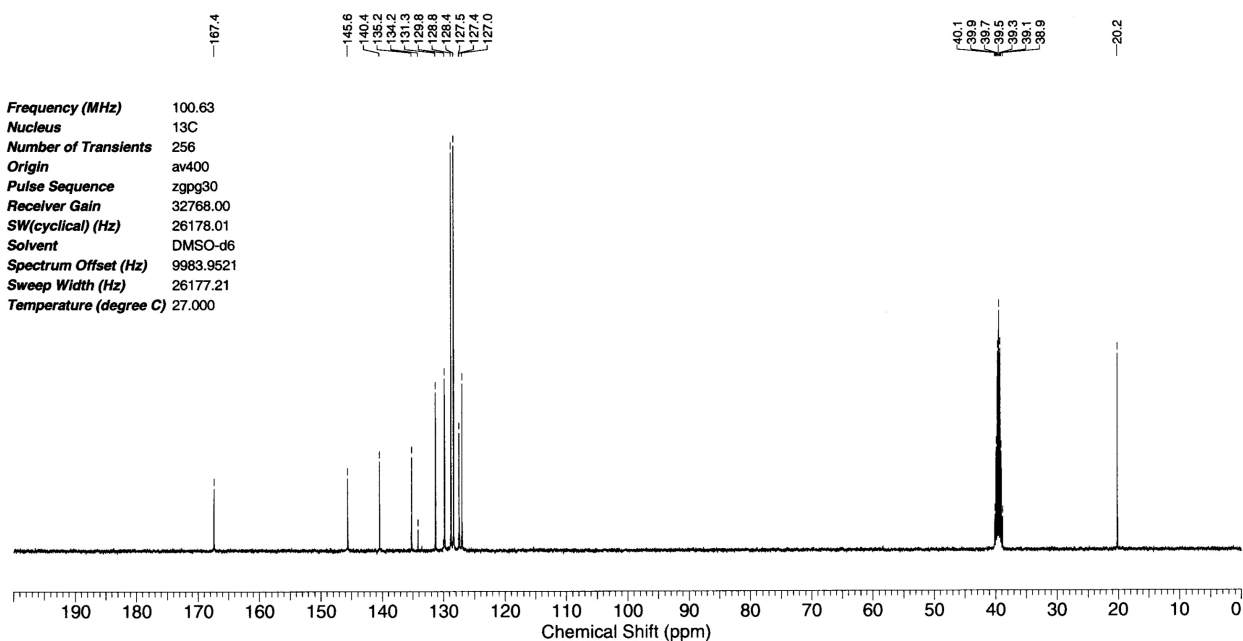


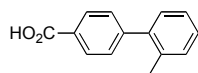


(Table 6.4, Entry 12) ^1H NMR (400 MHz, DMSO) (zoom)

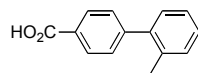
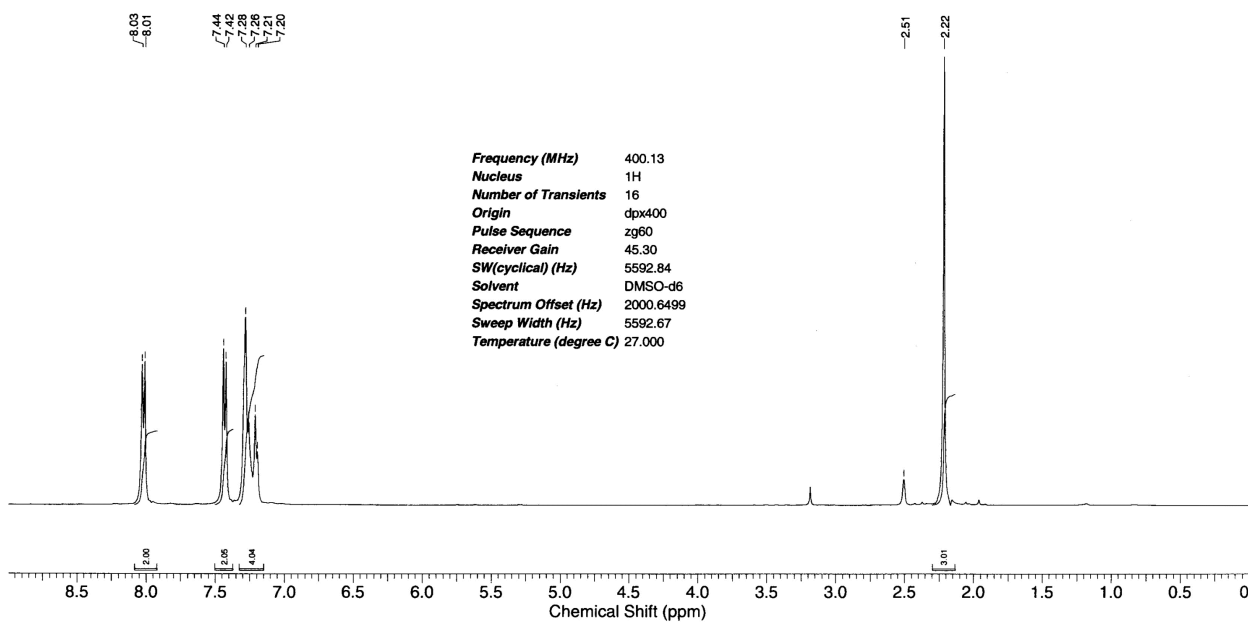


(Table 6.4, Entry 12) ^{13}C NMR (100 MHz, DMSO)

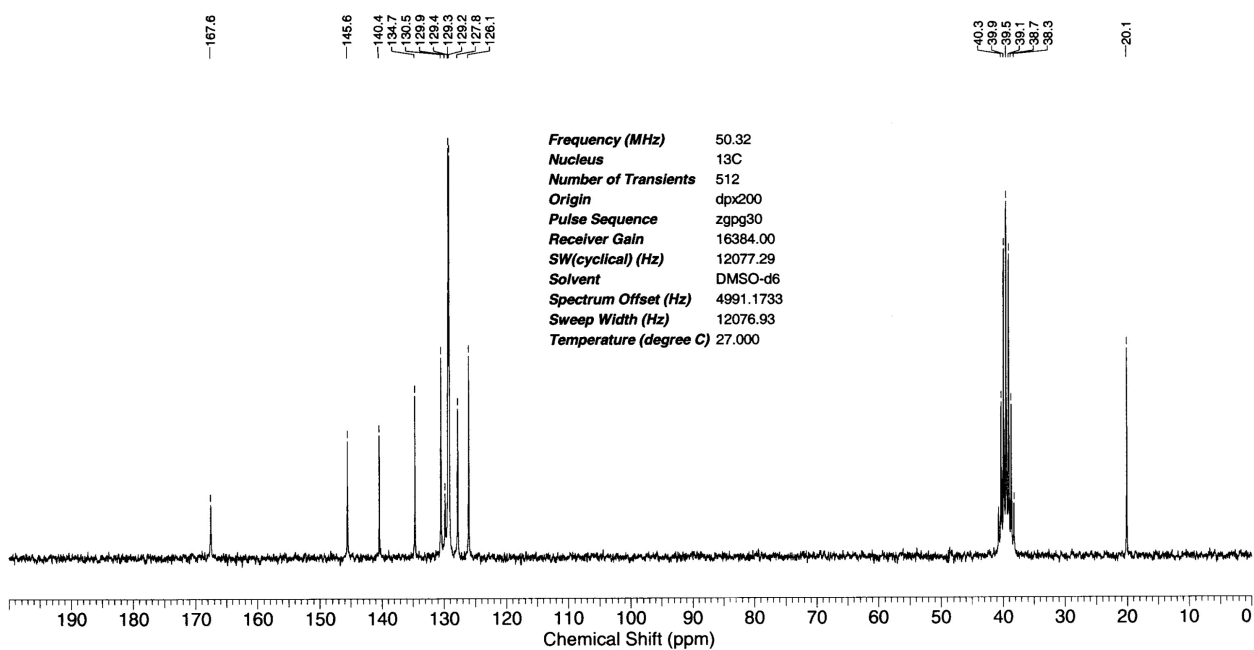


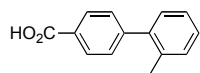


(Table 6.4, Entry 13) ^1H NMR (400 MHz, DMSO)

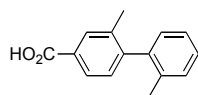
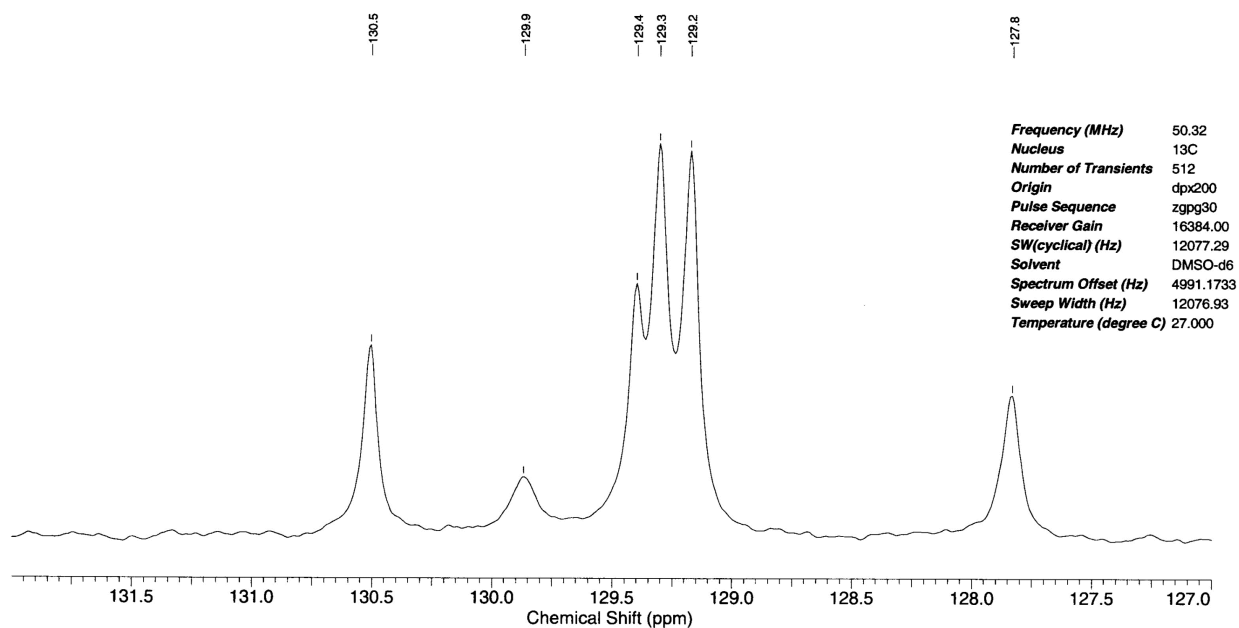


(Table 6.4, Entry 13) ^{13}C NMR (50 MHz, DMSO)

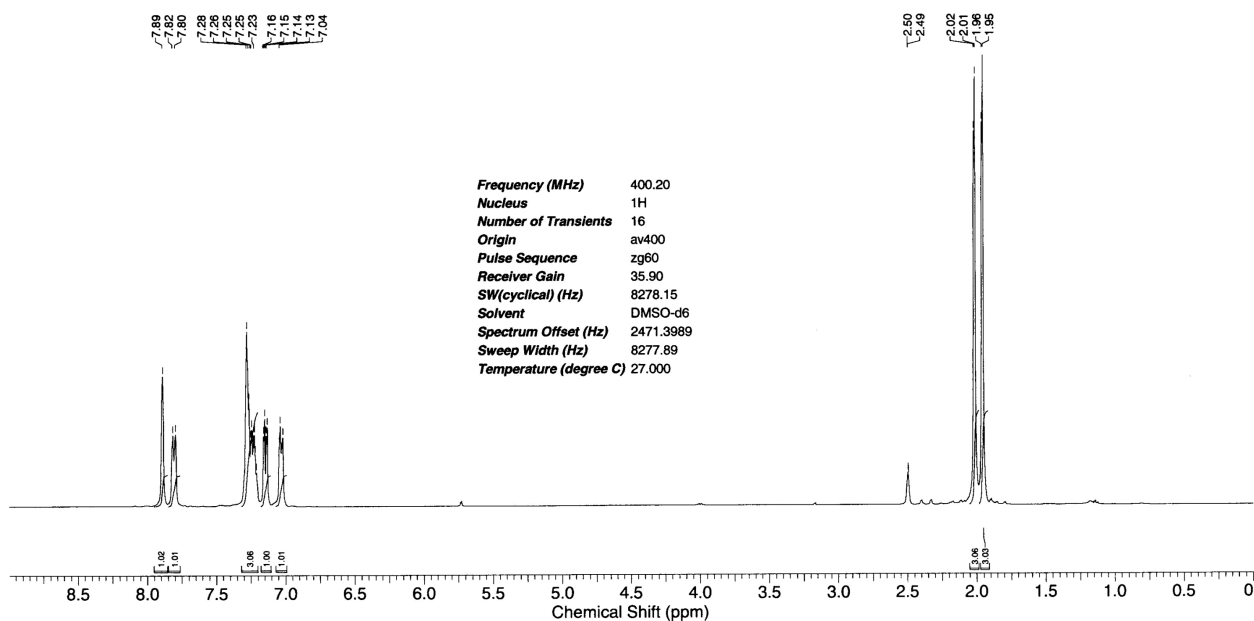


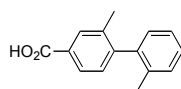


(Table 6.4, Entry 13) ^{13}C NMR (50 MHz, DMSO) (zoom)

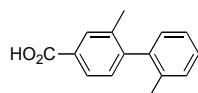
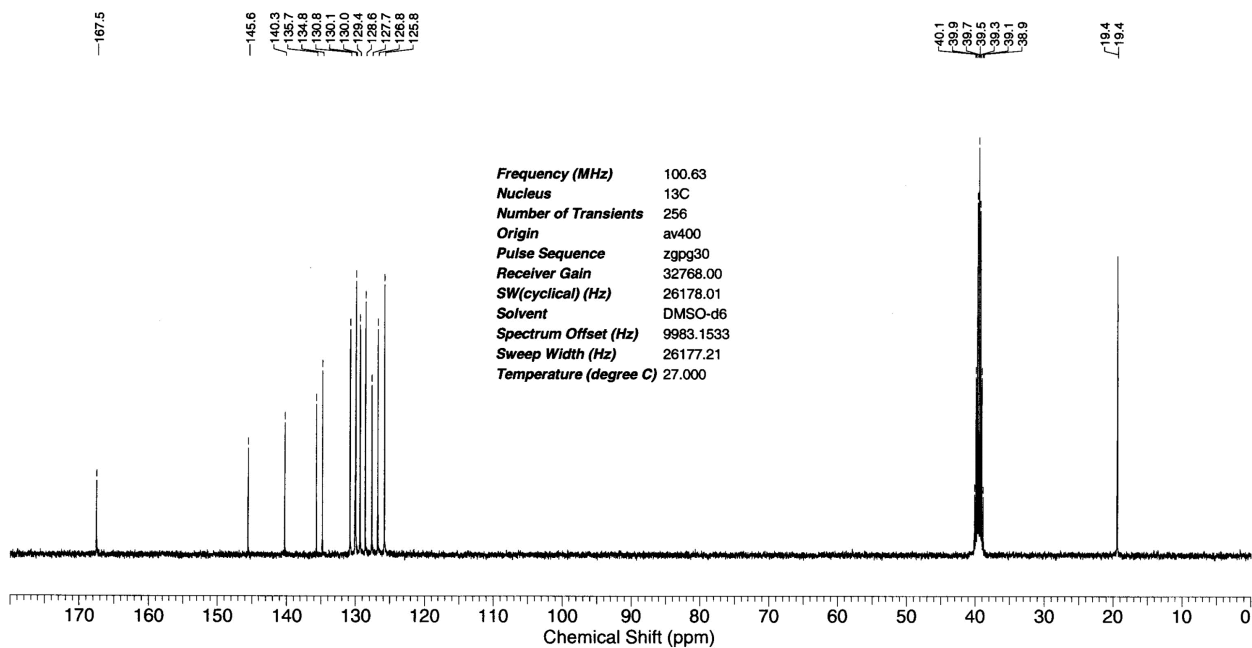


(Table 6.4, Entry 14) ^1H NMR (400 MHz, DMSO)

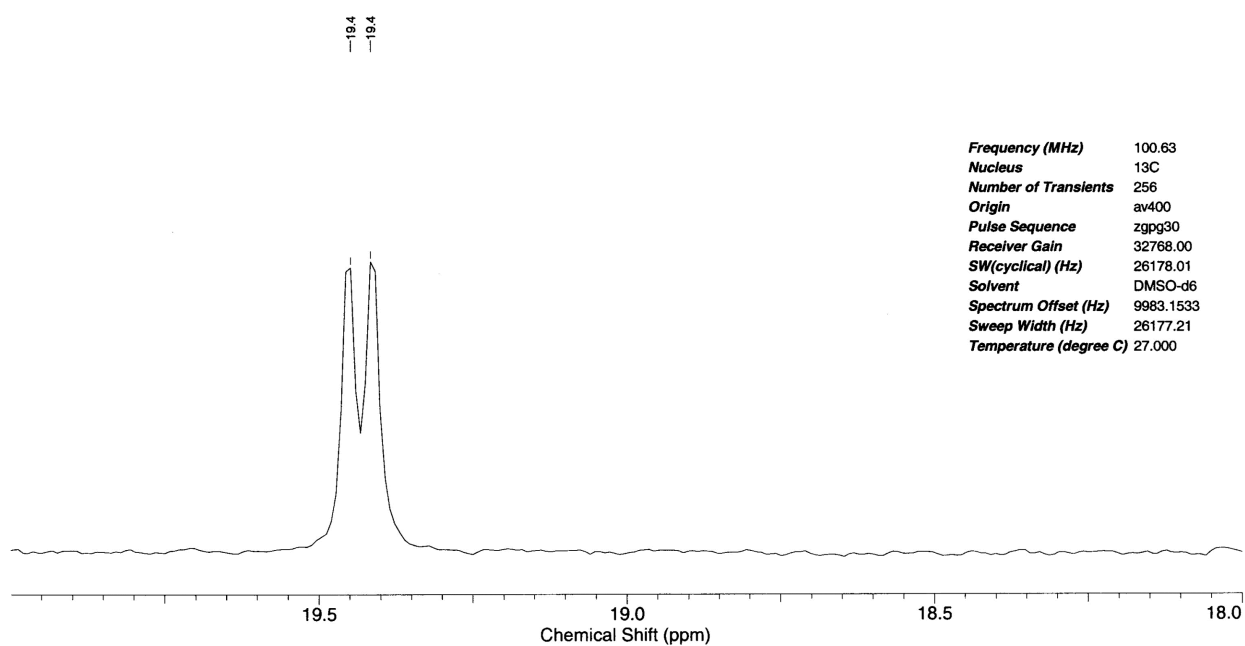


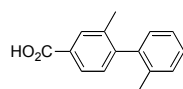


(Table 6.4, Entry 14) ^{13}C NMR (100 MHz, DMSO)

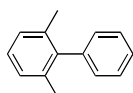
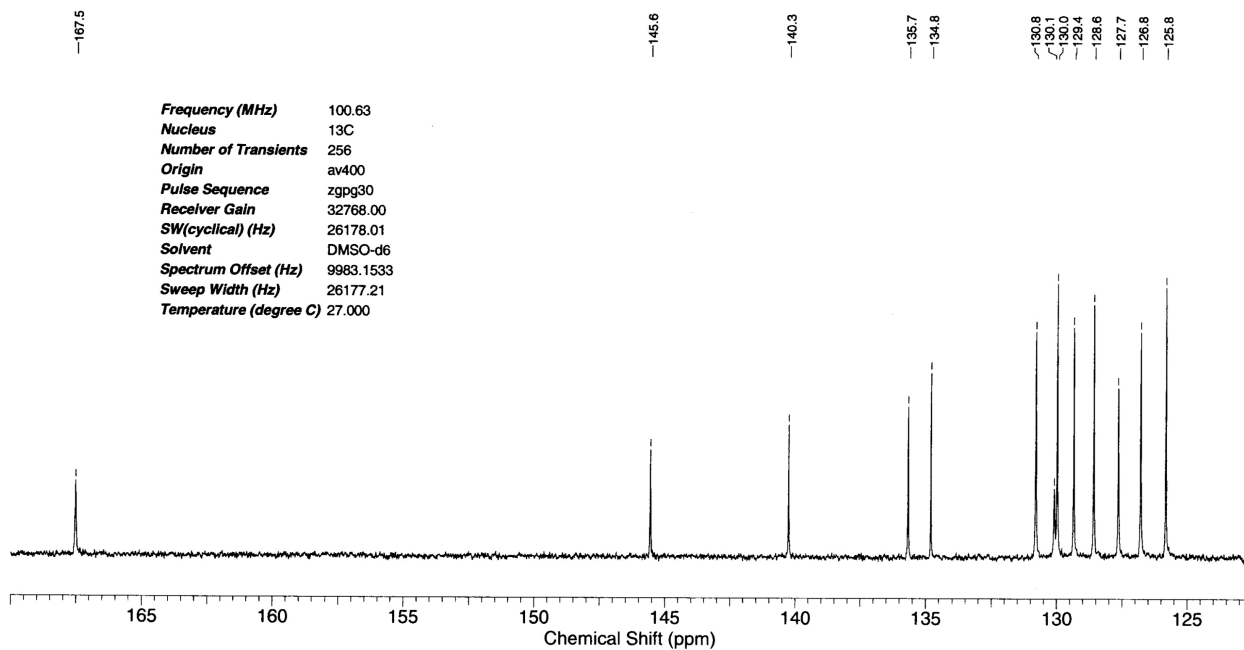


(Table 6.4, Entry 14) ^{13}C NMR (100 MHz, DMSO) (zoom)

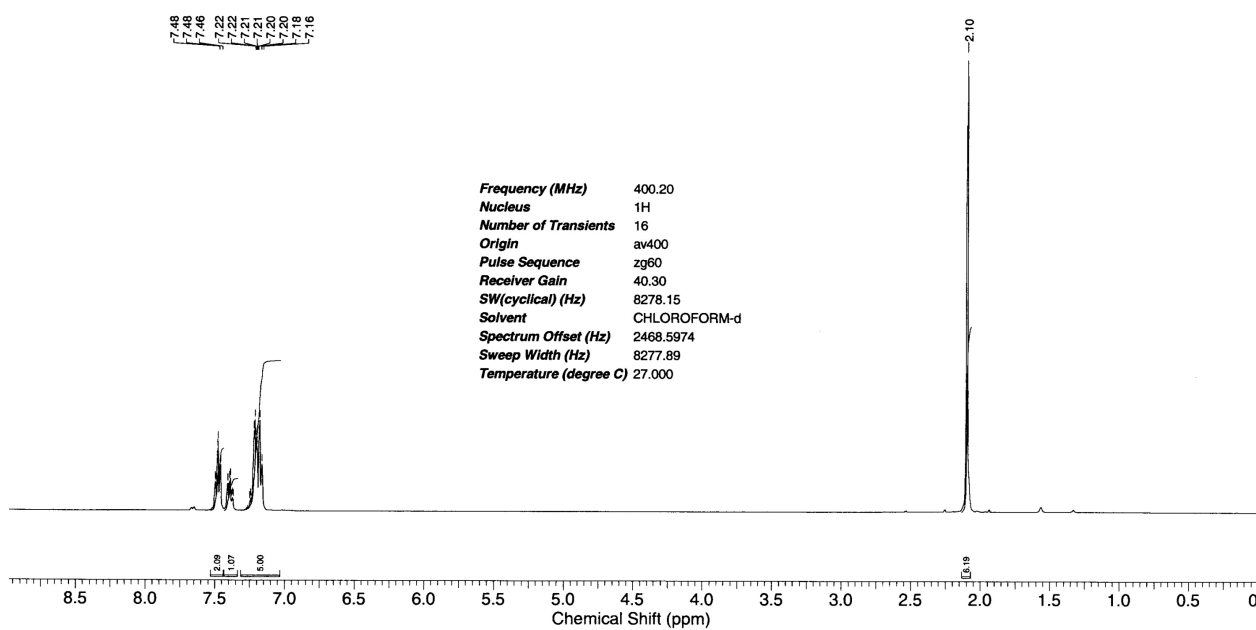


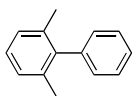


(Table 6.4, Entry 14) ^{13}C NMR (100 MHz, DMSO) (zoom)

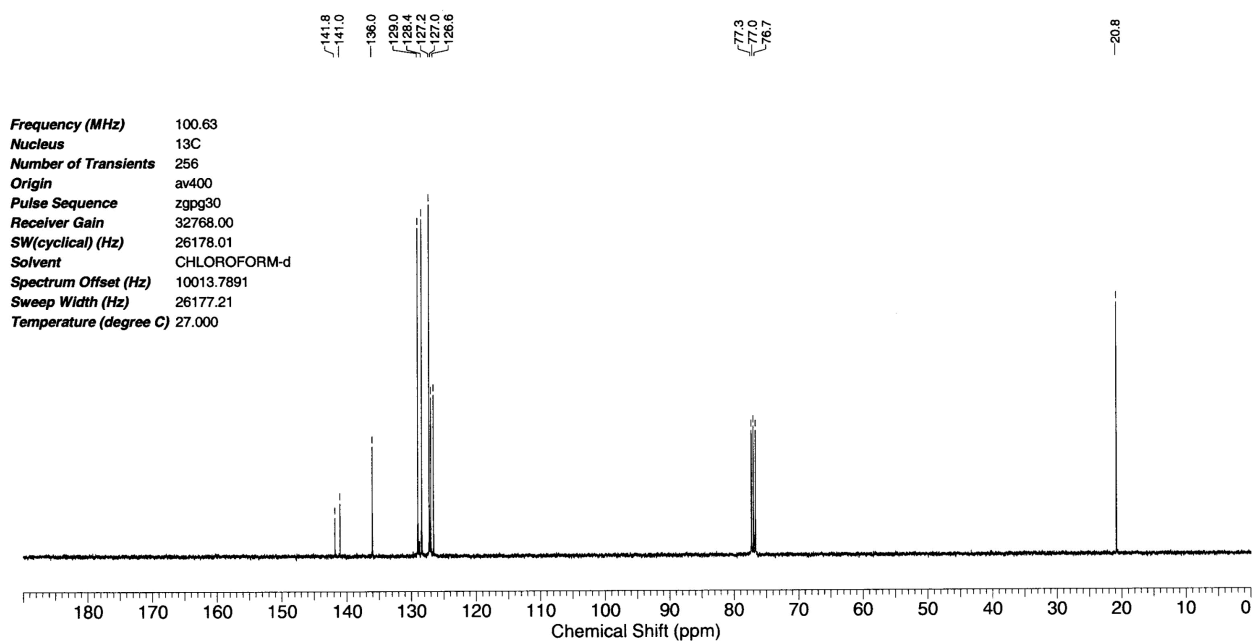


(Table 6.4, Entry 15) ^1H NMR (400 MHz, CDCl_3)



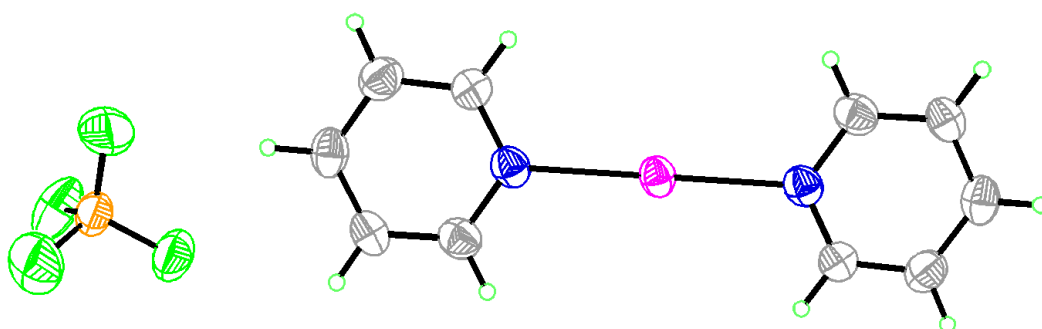


(Table 6.4, Entry 15) ^{13}C NMR (100 MHz, CDCl_3)



Chapter 7

A Safe and Scalable Preparation of Barluenga's Reagent (IPy₂BF₄)



“...at the borderline of physics and biology there are still enormous possibilities to advance science without missing our identity as chemists.”

José Barluenga, University of Oviedo
Angew. Chem. Int. Ed. **2010**, 49, 6250-6252

7.1 Introduction

Bis(pyridine)iodonium(I) tetrafluoroborate (IPy₂BF₄, **7.1**)—referred to as Barluenga's Reagent in recognition of its discoverer¹—is a mild iodonium source and oxidant. IPy₂BF₄ is a relatively stable solid and it is soluble in both organic and aqueous solvent systems.

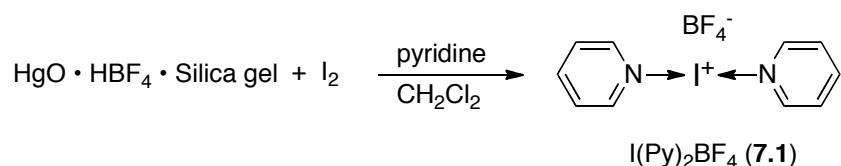


Figure 7.1: Original preparation of IPy₂BF₄, Barluenga's Reagent

The utility of **7.1** in organic synthesis has been widely demonstrated and a selection of transformations promoted by IPy₂BF₄ is shown in Figure 7.2. Alkenes and alkynes are iodinated by **7.1** under mild conditions, and lead to 1,2-functionalized products after attack by a nucleophile (Figure 7.2, A-E).¹⁻³ In cases where the nucleophile attacks intramolecularly, carbo- and heterocycles are formed (Figure 7.2, C-E).⁴⁻⁶ Often **7.1** alone can iodinate these saturated systems. In other cases, HBF₄ is used to protonate the pyridine and activate **7.1**. Arenes are iodinated by **7.1** through electrophilic aromatic substitution and the reaction is mild enough that selective iodination of tyrosine in peptides^{7,8} and proteins⁹ has been realized (Figure 7.2, F-G).

Barluenga's reagent is useful in oxidation chemistry. **7.1** can oxidize secondary alcohols to ketones under thermal conditions (Figure 7.2, H). Under photolytic conditions, cycloalkanols are oxidatively cleaved to the corresponding ω-iodocarbonyls when treated with **7.1**, and 1,2-diols are converted to the respective dicarbonyls (Figure 7.2, I-J).^{10,11} C-H functionalization¹² is also enabled by **7.1** (Figure 7.2, K).

IPy₂BF₄ has been used as an activator in carbohydrate synthesis (Figure 7.2, L-M). Thio- and *n*-pentenyl glycosides can be activated for glycosylation or converted to the corresponding glycosyl fluoride with **7.1**.^{13,14} Finally, it is noteworthy that many of the reactions in Figure 7.2 involve carbon-carbon bond formations (A, B, D, E, K).

Given the synthetic versatility of IPy₂BF₄, it is useful to have ready access to multi-gram quantities of this reagent. Even though Barluenga's reagent is available from several commercial suppliers, the cost may preclude its use in large-scale synthetic endeavors. For example, 5 grams of **7.1** is typically \$150-200 (USD). Moreover, the preparation of **7.1** described in the literature^{1,15} uses extremely toxic Hg(II) salts¹⁶ that necessitate a

risky and tedious workup. The original synthesis of **7.1** is depicted in Figure 7.1. Mercury(II) oxide and tetrafluoroboric acid were adsorbed onto silica gel. Iodine and pyridine were added to this system and **7.1** could be isolated after filtration, concentration, and precipitation from diethyl ether. Since mercury(II) is extremely toxic, the product obtained from this preparation had to be re-dissolved in dichloromethane and precipitated by the addition of diethyl ether several times until the filtrate tested negative for mercury(II). A safe and scalable procedure would necessitate an alternative protocol that avoids the use of mercury salts.

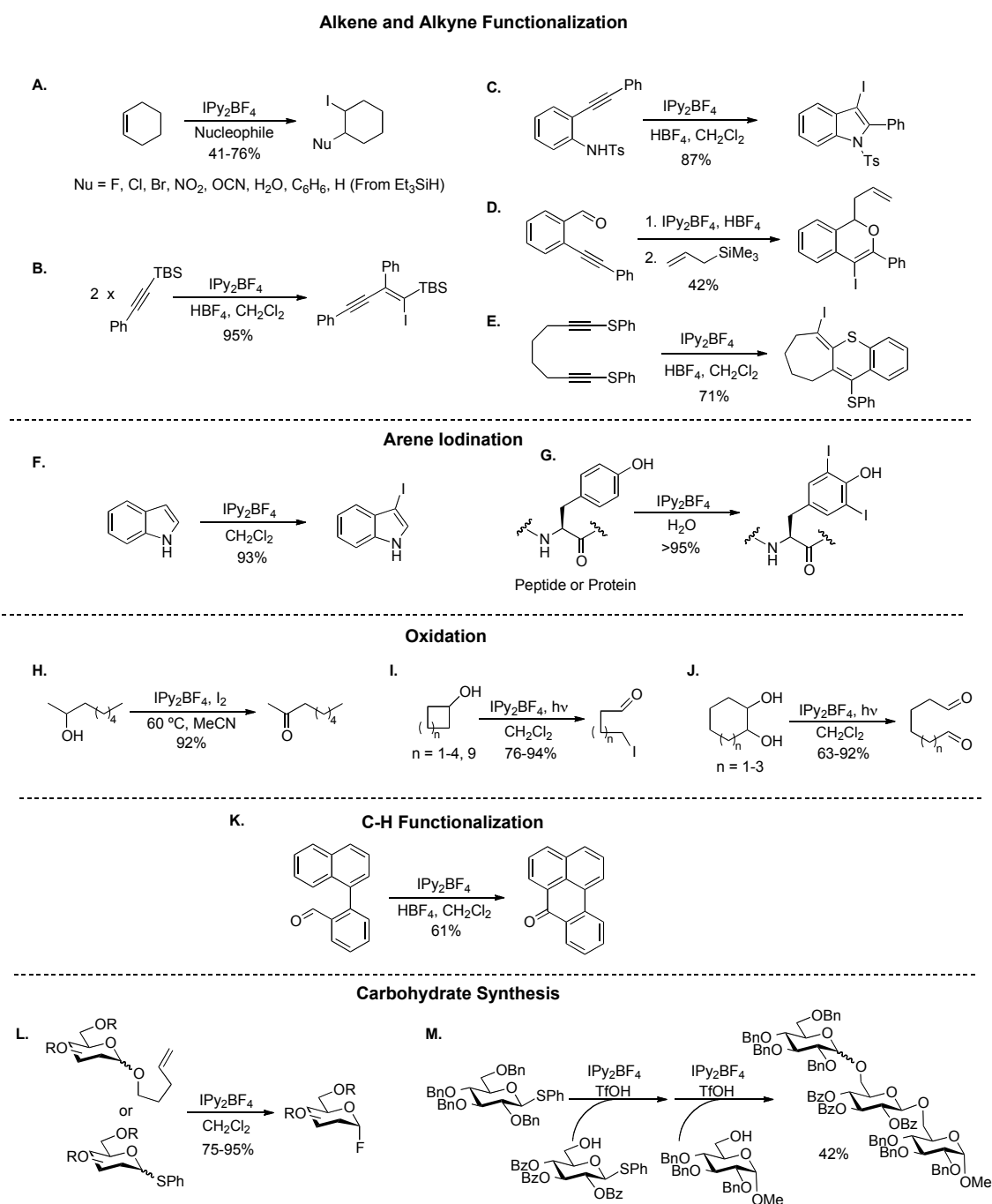


Figure 7.2: Barluenga's Reagent (IPy₂BF₄) in Organic Synthesis

In the cross-coupling studies described in Chapter 6, **7.1** was used to access diiodinated tyrosine as a substrate for the Suzuki-Miyaura reaction. For model reactions and exploratory work, large quantities of Barluenga's reagent were required. For instance, it would require at least 12.6 g of IPy₂BF₄ to prepare 5 grams of diiodinated BocTyrOMe (Figure 7.3). This quantity of Barluenga's reagent is commercially available at approximately \$350-500 (USD), depending on the supplier. Since **7.1** is used in stoichiometric quantities for all reactions in Figure 7.2, similarly prohibitive costs would be met for these transformations in large-scale endeavors.

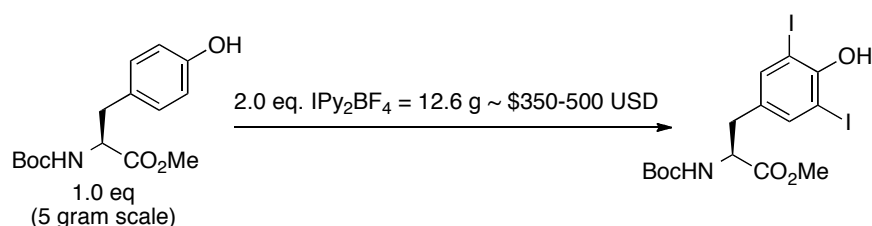


Figure 7.3: Cost of gram-scale iodinations with IPy₂BF₄

Given the safety risks of Barluenga's original preparation of **7.1** and the relatively high commercial cost for **7.1**, it was necessary to devise an alternative synthesis that was simple, safe, and scalable. Since the role of mercury(II) was ostensibly to abstract and trap iodide ion (I⁻), a halophilic species was needed as a mercury surrogate. Due to the high affinity of silver(I) salts for halides, and in particular iodide, it was thought that AgBF₄ would be a suitable replacement for HgO in the synthesis of Barluenga's Reagent.

7.2 Results and Discussion

While AgBF₄ is commercially available as a solution in water, Ag₂CO₃ is less costly—less than \$1 (USD) / mmol—and can be treated with HBF₄ to generate AgBF₄ and CO₂. Since AgBF₄ is not very soluble in CH₂Cl₂, the silver salt was adsorbed onto silica gel to aid in efficient mixing. This strategy is analogous to the preparation reported by Barluenga using HgO•HBF₄ on silica gel.^{1,15,17} Accordingly, after generating an aqueous solution of AgBF₄ from Ag₂CO₃ and HBF₄, silica gel was added to the reaction mixture and the water was removed by rotary evaporation to give the AgBF₄ on silica gel. The silica gel supported AgBF₄ was then suspended in CH₂Cl₂ and pyridine and iodine were added sequentially. AgI precipitated as a yellow solid immediately upon the addition of iodine. After 1 hour of reaction time, the solids (AgI and silica) were removed by

filtration. The filtrate was concentrated and the product could be precipitated simply by the addition of diethyl ether. The white solid isolated was spectroscopically identical to **7.1**. The yield by this method (72%) is comparable to that reported by Barluenga (70-80%).¹⁵

IPy₂BF₄ obtained by this method is sufficiently pure to be used in most synthetic applications, including the iodination depicted in Figure 7.4. Indeed nearly a quantitative yield of BocTyrOMe (**7.3**) was obtained using IPy₂BF₄ prepared by this method.

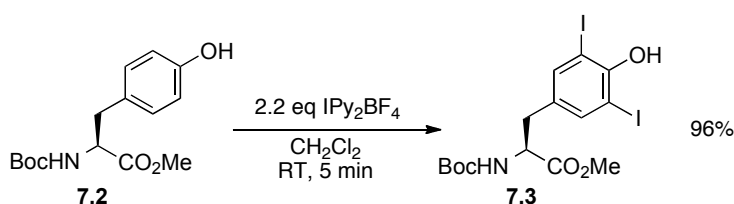


Figure 7.4: Diiodination of BocTyrOMe

IPy₂BF₄ can be crystallized from CH₂Cl₂ if required. To unambiguously confirm that the Ag(I) based synthesis leads to **7.1**, single crystals were prepared for X-ray analysis. The X-ray crystal structure is shown in Figure 7.5.

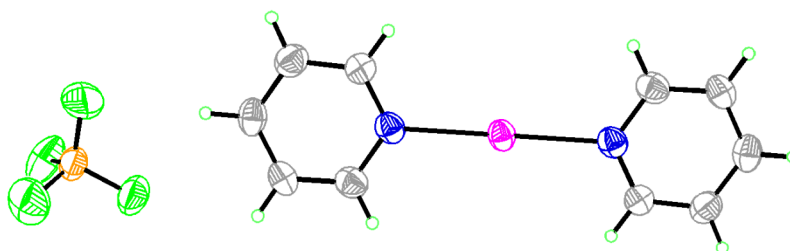


Figure 7.5: The molecular structure of bis(pyridine)iodonium(I) tetrafluoroborate, from the crystal structure at 150 K. Only one molecule is shown for clarity and the atomic displacement parameters are drawn at 50% probability.

In addition to confirming that the Ag(I) protocol does indeed provide IPy₂BF₄, the X-ray analysis also revealed an interesting phase change.¹⁸ When the diffraction data was recorded at 150 K, a C-centered monoclinic cell was determined. When the temperature was warmed to 250 K, a related, but smaller cell of about half the original size was observed. Returning to 150 K, the original cell was observed, suggesting that the phase change is reversible. This observation helped clarify certain aspects of disorder in the original report of the crystal structure of **7.1**.¹⁹

7.3 Conclusions

The preparation described herein is adapted procedurally from Barluenga's initial reports, but replaces Hg(II) with Ag(I). The resulting procedure is a much safer and more convenient synthesis, with yields comparable to that reported by Barluenga. The protocol reported herein can conveniently supply more than 10 g of IPy₂BF₄ in less than a day. The most expensive reagent, Ag₂CO₃, is less than one tenth the price per mole of commercial IPy₂BF₄. Moreover, this procedure is robust and reliable and has been repeated in other laboratories.²⁰ Additionally, X-ray analysis revealed a phase change in the crystals of IPy₂BF₄ not previously reported. With a safe and scalable synthesis of IPy₂BF₄ available, increased deployment of Barluenga's reagent in synthetic endeavors is anticipated.

7.4 References

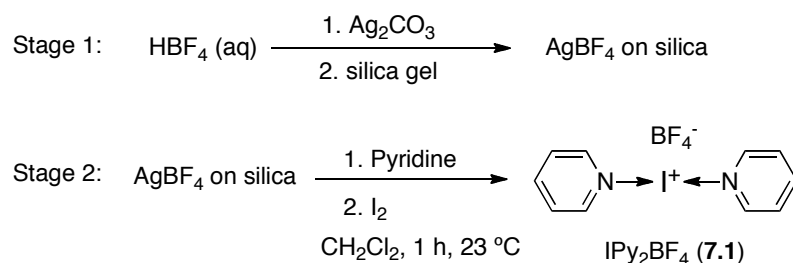
- (1) Barluenga, J.; González, J. M.; Campos, P. J.; Asensio, G. I(py)₂BF₄, a new reagent in organic synthesis: general method for the 1,2-iodofunctionalization of olefins. *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 319-320.
- (2) Barluenga, J.; Llorente, I.; Alvarez-García, L. J.; González, J. M.; Campos, P. J.; Díaz, M. R.; García-Granda, S. Reactions of (tert-butyldimethylsilyl)alkynes with IPy₂BF₄: selective synthesis of novel head-to-head dimers. *J. Am. Chem. Soc.* **1997**, *119*, 6933-6934.
- (3) Barluenga, J. Transferring iodine: more than a simple functional group exchange in organic synthesis. *Pure Appl. Chem.* **1999**, *71*, 431-436.
- (4) Barluenga, J.; Trincado, M.; Rubio, E.; González, J. M. IPy₂BF₄-promoted intramolecular addition of masked and unmasked anilines to alkynes: direct assembly of 3-iodoindole cores. *Angew. Chem. Int. Ed.* **2003**, *42*, 2406-2409.
- (5) Barluenga, J.; Vázquez-Villa, H.; Ballesteros, A.; González, J. M. Cyclization of carbonyl groups onto alkynes upon reaction with IPy₂BF₄ and their trapping with nucleophiles: a versatile trigger for assembling oxygen heterocycles. *J. Am. Chem. Soc.* **2003**, *125*, 9028-9029.
- (6) Barluenga, J.; Romanelli, G. P.; Alvarez-García, L. J.; Llorente, I.; González, J. M.; García-Rodríguez, E.; García-Granda, S. A novel domino reaction: cyclization of alkynyl sulfides by reaction with IPy₂BF₄. *Angew. Chem. Int. Ed.* **1998**, *37*, 3136-3139.

- (7) Vilaró, M.; Arsequell, G.; Valencia, G.; Ballesteros, A.; Barluenga, J. Arylation of Phe and Tyr side chains of unprotected peptides by a Suzuki-Miyaura reaction in water. *Org. Lett.* **2008**, *10*, 3243-3245.
- (8) Barluenga, J.; García-Martín, M. A.; González, J. M.; Clapés, P.; Valencia, G. Iodination of aromatic residues in peptides by reaction with IPy₂BF₄. *Chem. Commun.* **1996**, 1505-1506.
- (9) Espuña, G.; Andreu, D.; Barluenga, J.; Pérez, X.; Planas, A.; Arsequell, G.; Valencia, G. Iodination of proteins by IPy₂BF₄, a new tool in protein chemistry. *Biochemistry* **2006**, *45*, 5957-5963.
- (10) Barluenga, J.; González-Bobes, F.; Ananthoju, S. R.; García-Martín, M. A.; González, J. M. Oxidative opening of cycloalkanols: an efficient entry to ω-iodocarbonyl compounds. *Angew. Chem. Int. Ed.* **2001**, *40*, 3389-3392.
- (11) Barluenga, J.; González-Bobes, F.; Murguía, M. C.; Ananthoju, S. R.; González, J. M. Bis(pyridine)iodonium tetrafluoroborate (IPy₂BF₄): a versatile oxidizing reagent. *Chem. Eur. J.* **2004**, *10*, 4206-4213.
- (12) Barluenga, J.; Trincado, M.; Rubio, E.; González, J. M. Direct intramolecular arylation of aldehydes promoted by reaction with IPy₂BF₄/HBF₄: synthesis of benzocyclic ketones. *Angew. Chem. Int. Ed.* **2006**, *45*, 3140-4143.
- (13) Huang, K. T.; Winssinger, N. IPy₂BF₄-mediated glycosylation and glycosyl fluoride formation *Eur. J. Org. Chem.* **2007**, 1887-1890.
- (14) López, J. C.; Uriel, C.; Guillamón-Martín, A.; Valverde, S.; Gómez, A. M. IPy₂BF₄-mediated transformation of n-pentenyl glycosides to glycosyl fluorides: a new pair of semiorthogonal glycosyl donors. *Org. Lett.* **2007**, *9*, 2759-2762.
- (15) Barluenga, J.; Rodríguez, M. A.; Campos, P. J. Electrophilic additions of positive iodine to alkynes through an iodonium mechanism. *J. Org. Chem.* **1990**, *55*, 3104-3106.
- (16) <http://msds.chem.ox.ac.uk/>.
- (17) Barluenga, J.; Campos, P. J.; González, J. M.; Asensio, G. A simple and general route to aryl iodides from arenes. *J. Chem. Soc., Perkin Trans. 1* **1984**, 2623.
- (18) X-Ray and phase analysis were carried out by Dr. Amber Thompson of the University of Oxford Crystallography Service. For an extended discussion the crystal analysis, also contributed by Dr. Amber Thompson, see the experimental section of this Chapter.

- (19) Álvarez-Rúa, C.; García-Granda, S.; Ballesteros, A.; González-Bobes, F.; González, J. M. Bis(pyridine)iodonium(I) tetrafluoroborate. *Acta Cryst.* **2002**, E58, o1381-o1383.
- (20) Chalker, J. M.; Thompson, A. L.; Davis, B. G. Safe and scalable preparation of Barluenga's reagent (bis(pyridine)iodonium(I) tetrafluoroborate). *Org. Synth.* **2010**, 87, 288-298.

7.5 Experimental Procedures

The following preparation of IPy₂BF₄ is adapted from that recently published in *Organic Syntheses*.¹ The reader is referred to this publication for additional details and observations by the lab that repeated and checked the procedure.



Stage 1. *Silver tetrafluoroborate on silica*. A 500 mL, 1-neck round-bottom flask is equipped with a Teflon-coated magnetic stir bar. Deionized water (Millipore, 50 mL) is poured into the flask and tetrafluoroboric acid (6.3 mL of an 8.0M aqueous solution, 50 mmol, 1.0 equiv, Fluka) is added by syringe. The solution is stirred, open air, at room temperature (23 °C) while silver carbonate (6.89 g, 25.0 mmol, 0.50 equiv, Alfa Aesar) is added in several portions as a solid over 2 minutes. Carbon dioxide gas evolves. The resulting mixture is stirred rapidly (1500 rpm) until all solids are dissolved to give a grey solution (typically 20 minutes). Silica gel (10.0 g, SiliaFlash® P60, VWR) is poured into the solution and the mixture stirred for 5 minutes. The stir bar is removed using a magnetic wand and the flask transferred to a rotary evaporator with a water bath preheated to 55 °C. The water is evaporated (55 °C, 30 mm Hg, dry ice condenser) to give the AgBF₄ on silica as an off-white/grey solid on the walls of the flask. A spatula is used to scrape all silica to the bottom of the flask. The flask is then returned to the rotary evaporator and dried with rapid spinning (55 °C, 20 mm Hg). After 10 minutes of drying

in this manner, a free flowing powder is obtained. The flask is allowed to stand and cool at room temperature for 5 minutes before carrying on to the next step.

Stage 2. *Bis(pyridine)iodonium(I) tetrafluoroborate* (IPy₂BF₄, 7.1). A Teflon-coated magnetic stir bar is added to the 500 mL one-neck flask from above containing the AgBF₄ on silica. Dichloromethane (300 mL, HPLC grade, Fisher Scientific) is poured into the flask and the mixture is stirred at room temperature, open to the atmosphere. Pyridine (8.1 mL, 100 mmol, 2.0 equiv, glass distilled grade, Rathburn Chemicals Ltd.) is poured into the reaction flask from a 10.0 mL measuring cylinder and the mixture stirred at room temperature. One minute after the addition of pyridine, iodine beads (12.69 g, 50.0 mmol, 1.0 equiv, Fisher Scientific) are poured into the reaction flask from a weighing boat. A yellow precipitate (AgI) is formed immediately upon addition of the iodine. A glass stopper is used to loosely cap the reaction flask and the mixture is stirred vigorously (1500 rpm) for 1 hour. The formation of the product, IPy₂BF₄, is difficult to monitor by conventional methods such as TLC. Qualitatively, the formation of AgI (yellow solid) and the dissolution of the iodine beads are indicators of reaction progression. Typically 1 hour of stirring was sufficient for a preparation on the scale described and prolonged reaction time did not improve the final yield (*vide infra*). The solution gradually turns red as the iodine is dissolved. After 1 hour of stirring, the solids are removed by filtration through a sintered glass funnel (porosity 3 glass frit, 250 mL capacity) under gentle suction and the dark red filtrate is collected in a 1 L, 1 neck round-bottom flask. The reaction flask is rinsed with 100 mL dichloromethane to ensure complete transfer and the silica/AgI filter cake is then washed with an additional 100 mL dichloromethane. The solvent is removed by rotary evaporation (35 °C, 285 mm Hg) to give a reddish-yellow solid. This solid is then dissolved in dichloromethane (100 mL) and a Teflon-coated stir bar is added. The stirred solution is cooled to 0 °C using an ice bath and then diethyl ether (200 mL, glass distilled grade, Rathburn Chemicals Ltd.) is poured into the stirred solution to precipitate the product. The mixture is stirred at 0 °C for 10 minutes to ensure all IPy₂BF₄ is precipitated. The product is isolated by filtration under gentle suction, using a filter with a porosity 3 glass frit that was lined with filter paper (QL 100, Fisher Scientific). The flask is rinsed with diethyl ether (50 mL) to complete the transfer. The white power was dried on the filter for 5 minutes and then transferred to a jar or flask and dried under vacuum in a desiccator for 1 hour (0.1 mm Hg). The powdered IPy₂BF₄ (13.30 g, 72% yield) is sufficiently pure for most synthetic applications and can be stored

for months if protected from light and kept at -20 °C. The preparation was run in triplicate on the described scale with yields of 13.0 to 13.6 g (70-73%) after precipitation with diethyl ether.

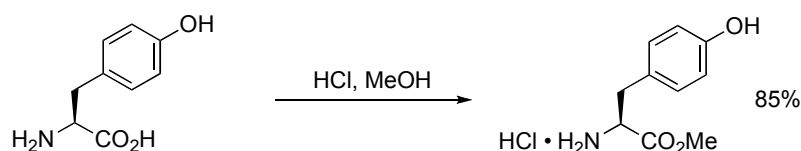
Spectroscopic data for precipitated **7.1**: ^1H NMR (400 MHz, CD_3CN) δ : 7.63 (4H, dd, $J = 7.7, 6.4$ Hz, H-3), 8.24 (2H, tt, $J = 7.7, 1.4$ Hz, H-4), 8.78 (4H, dd, $J = 6.4, 1.4$, H-2); ^{13}C NMR (100 MHz, CD_3CN) δ : 129.2 (C-3), 143.5 (C-4), 150.9 (C-2); ^{19}F NMR (377 MHz, CD_3CN) δ : -151.5 (BF_4^-).

IPy_2BF_4 can also be crystallized from dichloromethane. Accordingly, the powdered IPy_2BF_4 (13.30 g) is added to a 100 mL, 1 neck round-bottom flask with a Teflon-coated stir bar. A water bath is preheated to 50 °C and 25 mL of dichloromethane is added to the solid. The suspension is stirred and lowered to the water bath. Once the dichloromethane is boiling (ca. 2 minutes), additional dichloromethane is added by pipette, with stirring, until all solids are dissolved (30 mL additional dichloromethane, ca. 55 mL total solvent). Once a clear solution is obtained, the stirring is stopped and the flask removed from the water bath. Prisms are observed within 10 minutes. After 30 minutes, the flask is cooled to 0 °C (ice bath) to facilitate the crystallization. After 2 hours on ice, the crystals are isolated by filtration with gentle suction using a filter with a porosity 3 glass frit that was lined with filter paper (QL 100, Fisher Scientific). The suction is removed and the crystals are triturated with diethyl ether (30 mL) on the filter. The crystals are then dried on the filter for 5 minutes under suction and then transferred to a jar and dried under vacuum in a desiccator for 2 hours (0.1 mm Hg). The first crop of crystals yields 8.11 g of white prisms. All remaining material in the crystallization flask and filter is collected by rinsing with dichloromethane (ca. 50 mL). The combined filtrates and mother liquor are concentrated by rotary evaporation (35 °C, 285 mm Hg) to give a red solid. The recrystallization is repeated twice more as above to recover the remaining material, using 25 mL of dichloromethane in a 50 mL round-bottom flask in the second recrystallization (2.79 g of crystals isolated) and 8 mL of dichloromethane in a 25 mL round-bottom flask in the third and final recrystallization (1.13 g crystals isolated). The total recovery from recrystallization is 12.03 g (90% yield from powdered IPy_2BF_4 and 65% overall yield).

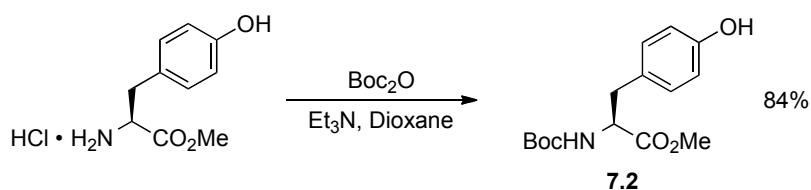
Analytical data for recrystallized **7.1**: ^1H , ^{13}C , and ^{19}F NMR spectra after recrystallization were identical to the product precipitated from diethyl ether (*vide supra*). The following analytical data was obtained from the first crop of crystals: mp 148-151 °C (Lit²: 149-151 °C); IR (KBr) cm^{-1} : 1603, 1447, 1209, 1056, 760, 689, 636, 521, 440; Anal. Calcd. for $\text{C}_{10}\text{H}_{10}\text{BF}_4\text{IN}_2$: C, 32.29; H, 2.71; N, 7.53. Found C, 32.33; H, 2.70 ; N, 7.53.

Single crystals suitable for X-ray analysis were prepared to unambiguously confirm the identity of the product. Accordingly, 500 mg of the first crop of crystals was dissolved in 3.0 mL of boiling dichloromethane in a straight-sided glass tube (2.0 × 5.0 cm) and left to cool, open air. After 30 minutes, the vial was capped, covered in foil, and left to stand overnight at room temperature (23 °C). The crystals obtained were submitted to the University of Oxford Crystallography Service for analysis by Dr. Amber Thompson. The results of her analysis are summarized here. Colorless plate-like crystals were isolated from the mother liquor and coated in inert perfluoro-polyether oil, and mounted on a hair and frozen in place at 150 K using an Oxford Cryosystems Cryostream N2 open flow cooling device.³ Diffraction data were collected using a Nonius Kappa-CCD diffractometer and results compared with those reported by Alvarez-Rua et al.⁴ Three crystals were tested in all by the collection of a 10° ω -scan which indexed in each case to give the same monoclinic cell (by transformation, similar to that reported previously⁵; see extended discussion at the end of this Chapter for further details). A complete dataset to a maximum resolution of 0.77 Å was recorded for the third crystal. On completion, the crystal was warmed to 250 K at 120 K/h, where the cell was redetermined and found to be a smaller, related C-centered monoclinic cell. A second dataset was recorded for comparison before the crystal was returned to 150 K (at 120 K/h) where the cell was found to have returned to that seen previously at that temperature, demonstrating the occurrence of a reversible structural phase transition. Cell parameters and intensity data for both data sets were processed using the DENZO-SMN package⁶ and the structures solved by direct methods using SIR92.⁷ Both structures were refined by full-matrix least squares on F^2 using the CRYSTALS suite (Figure 1).⁸ Where possible, non-hydrogen atoms were refined with anisotropic displacement parameters and the hydrogen atoms were initially refined with soft restraints, then constrained using a riding model. In the case of the 250 K data, the BF_4 counter-ion was found to be disordered and a full description of the way this was modeled and a brief comparison of the structures is given in the extended discussion at the end of this Chapter, together with the structural data in

CIF format. These data have also been deposited with the Cambridge Crystallographic Data Centre as supplementary publications no. CCDC 772585 and CCDC 772586, respectively, and copies can be obtained, free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif. The X-ray structure is shown in Figure 7.5.

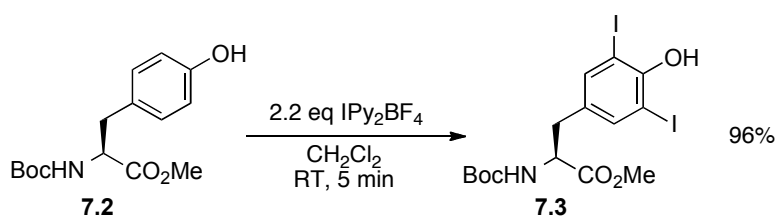


A 250 mL round bottom flask was flame dried and placed under an atmosphere of nitrogen. Anhydrous MeOH (80 mL) was added to the flask and cooled to 0 °C. Acetyl chloride (193 mmol, 13.74 mL) was added slowly with stirring to generate methanolic HCl. After all of the acetyl chloride was added, the solution was stirred an additional 15 minutes at 0 °C. L-Tyrosine (7.00 g, 38.6 mmol) was added as a solid in one portion. The ice bath was removed and the reaction stirred at room temperature for 12 hours. After this time, volatiles were removed under reduced pressure to give the crude product as an off-white solid. The solid was triturated by stirring rapidly in 50 mL of boiling acetone for 10 minutes. After cooling to room temperature, the flask was cooled on ice for several minutes. The resulting micro-crystals were isolated by filtration and washed with 50 mL of cold acetone. The crystals were then dried under vacuum (7.65 g, 85% yield). m.p. = 172-174 °C, decomposes at 190 °C (Lit⁹: 190-192, decomp.). $[\alpha]_D^{20} = +69.8$ (c = 3.0, pyridine) (Lit¹⁰: +72.0, c = 3.0, pyridine). IR ($\nu_{\max}$, KBr): 3343, 2879, 1958, 1744, 1614, 1592, 1515, 1448, 1397, 1349, 1226, 1141, 1107, 1060. ¹H NMR (400 MHz, MeOD): δ = 3.10-3.22 (2H, ABX system, J = 14.3, 7.2, 6.1, $\underline{\text{CH}_2\text{Ar}}$), 3.82 (3H, s, $\text{CO}_2\underline{\text{CH}_3}$), 4.28 (1H, app t, J = 6.7, H α), 6.81 (2H, d, J = 8.5, ArH), 7.10 (2H, d, J = 8.5, ArH). ¹³C NMR (50 MHz, MeOD): δ = 36.7 ($\underline{\text{CH}_2\text{Ar}}$), 53.7, 55.6 (C α and $\text{CO}_2\underline{\text{CH}_3}$), 117.0, 125.8, 131.7, 158.4 (Ar), 170.6 (C=O). LRMS m/z (ESI): 194.1 [M-H]⁻.



L-Tyrosine methyl ester hydrochloride (3.23 g, 16.6 mmol) was added to a 250 mL round bottom flask and suspended in dioxane (50 mL). To the stirred suspension was added

Et₃N (4.65 mL, 33.1 mmol). Boc₂O (4.70 g, 21.5 mmol) was then added in one portion and the reaction was stirred at room temperature. After 5 hours of reaction time, the reaction was diluted with EtOAc (150 mL) and washed with H₂O (150 mL). The aqueous layer was extracted with EtOAc (2 × 150 mL) and the combined organic layers were washed with brine (200 mL), dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography by eluting first with 15% EtOAc in petrol and then 50% EtOAc in petrol. The product was isolated as a white solid (4.13 g, 84% yield). m.p. = 96-99 °C (Lit.¹¹ = 102 °C). [α]_D²⁰ = +41.2 (c = 1.0, CHCl₃) (Lit.¹²: +53.1, c = 1.0, CHCl₃), IR ($\nu_{\max}$, KBr): 3370, 2979, 1687, 1516, 1445, 1368, 1224, 1166. ¹H NMR (200 MHz, CDCl₃): δ = 1.42 (9H, s, Boc), 3.00 (2H, ABX system, *J* = 5.8, 6.0, 13.6, CH₂Ar), 3.71 (3H, s, CO₂CH₃), 4.54 (1H, m, H α), 5.06 (1H, d, *J* = 4.0, NH), 6.48 (1H, br s, OH), 6.73 (2H, d, *J* = 4.0, CH_{Ar}), 6.96 (2H, d, *J* = 4.0, CH_{Ar}). ¹³C NMR (50 MHz, CDCl₃): δ = 28.3 (Boc), 37.6 (CH₂Ar), 52.4 (CO₂CH₃), 54.6 (C α), 80.3 (Boc), 115.5, 127.3, 130.3, 155.3, 155.4, 172.7 (4 × C_{Ar} and 2 × C=O). LRMS *m/z* (ESI): 294.1 [M-H].



To a solution of *N*-Boc-L-tyrosine-methyl ester (1.76 g, 5.97 mmol) in 50 mL CH₂Cl₂ was added IPy₂BF₄ (4.88 g, 13.1 mmol) in one portion. The TLC was checked immediately and showed complete consumption of starting material (less than 5 min at room temperature). The reaction was quenched with Na₂S₂O₃ (50 mL, 20% wt.) and extracted with CH₂Cl₂ (2 × 50 mL). The combined organic layers were dried over MgSO₄, filtered, and concentrated under reduced pressure. The product was purified by column chromatography (40% EtOAc in petrol). The di-iodinated product was isolated as a white solid (3.08 g, 96%) that was spectroscopically identical to that reported in the literature.¹³ m.p. = 117-119 °C. [α]_D²⁰ = +42.2 (c = 1.1, CHCl₃). IR ($\nu_{\max}$, KBr): 3455, 3363, 2979, 1759, 1729, 1687, 1523, 1458, 1406, 1254, 1164. ¹H NMR (200 MHz, CDCl₃): δ = 1.44 (9H, s, Boc), 2.95 (2H, ABX system, *J* = 5.8, 6.2, 13.8, CH₂Ar), 3.73 (3H, s, CO₂CH₃), 4.48 (1H, m, H α), 5.03 (1H, d, *J* = 7.6, NH), 7.43 (2H, s, 2 × CH_{Ar}). ¹³C NMR (50 MHz, CDCl₃): δ = 28.3 (Boc), 36.3 (CH₂Ar), 52.4 (CO₂CH₃), 54.3 (C α), 80.2, 82.2 (C-I and

Boc), 132.2, 139.9, 152.6, 154.9, 171.8 ($3 \times C_{Ar}$ and $2 \times C=O$). LRMS m/z (ESI⁻): 545.9 [M-H]⁻.

7.6 References for Experimental Procedures

-
- (1) Chalker, J. M.; Thompson, A. L.; Davis, B. G. *Org. Synth.* **2010**, *87*, 288-298.
 - (2) Barluenga, J.; González, J. M.; Campos, P. J.; Asensio, G. *Angew. Chem. Int. Ed. Engl.* **1985**, *24*, 319-320.
 - (3) Cosier, J.; Glazer, A. M. *J. Appl. Cryst.* **1986**, *19*, 105-107.
 - (4) Álvarez-Rúa, C.; García-Granda, S.; Ballesteros, A.; González-Bobes, F.; González, J. M. *Acta Cryst.* **2002**, *E58*, o1381-o1383.
 - (5) Feast, G. C.; Haestier, J.; Page, L. W.; Robertson, J.; Thompson, A. L.; Watkin, D. J. *Acta Cryst.* **2009**, *C65*, o635-o638.
 - (6) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307-326.
 - (7) Altomare, A.; Cascarano, G.; Giacovazzo, C.; Guagliardi, A.; Burla, M. C.; Polidori, G.; Camalli, M. *J. Appl. Cryst.* **1994**, *27*, 435.
 - (8) Betteridge, P. W.; Carruthers, J. R.; Cooper, R. I.; Prout, K.; Watkin, D. J. *J. Appl. Cryst.* **2003**, *36*, 1487.
 - (9) Vogler, K.; Studer, R. O.; Lergier, W.; Lanz, P. *Helv. Chim. Acta.* **1965**, *48*, 1407-1414.
 - (10) Groselj, U.; Bevk, D.; Jakse, R.; Meden, A.; Pirc, S.; Recnik, S.; Stanovnik, B.; Svete, J. *Tetrahedron: Asymmetry* **2004**, *15*, 2367-2383.
 - (11) Chankeshwara, S. V.; Chakraborti, A. K. *Org. Lett.* **2006**, *8*, 3259.
 - (12) Oswald, C. L.; Carrillo-Márquez, T.; Caggiano, L.; Jackson, R. F. W. *Tetrahedron* **2008**, *64*, 681-687.
 - (13) a) Brink, H.; Rijkers, D.; Liskamp, R. M. J. *J. Org. Chem.* **2006**, *71*, 1817-1824.
b) Nishiyama, S.; Kim, M. H.; Yamamura, S.; *Tetrahedron Lett.* **1994**, *35*, 8397-8400. c) Ocasio, C. A.; Scanlan, T. S. *ACS Chem. Biol.* **2006**, *1*, 585-593.

7.7 X-Ray Crystallographic Data and NMR and other Spectra

X-Ray crystallographic analysis was carried out by Dr. Amber Thompson.

Crystal Structure Determination

by X-ray Diffraction

Colourless plate-like crystals were isolated from the mother liquor and coated in inert perfluoro-polyether oil, and mounted on a hair and frozen in place at 150 K using an Oxford Cryosystems Cryostream N2 open flow cooling device.¹ Diffraction data were collected using a Nonius Kappa-CCD diffractometer and results compared with those reported by Alvarez-Rua *et al.*² Three crystals were tested in all by the collection of a 10° ω -scan which indexed in each case to give the same primitive monoclinic cell after transformation, similar to that reported previously (Table 1). A complete dataset to a maximum resolution of 0.77 Å was recorded for the third crystal. On completion, the crystal was warmed to 250 K at 120 K/hr, where the cell was re-determined and found to be a smaller, related C-centred monoclinic cell. A second dataset was recorded for comparison before the crystal was returned to 150 K (at 120 K/hr) which gave a similar cell to that seen previously at that temperature, suggesting that the phase transition is reversible. Cell parameters and intensity data for both data sets were processed using the DENZO-SMN package³ and the structures solved by direct methods using SIR92.⁴ Both structures were refined by full-matrix least squares on F^2 using the CRYSTALS suite.⁵ Where possible, non-hydrogen atoms were refined with anisotropic displacement parameters and the hydrogen atoms were initially refined with soft restraints, then constrained using a riding model. In the case of the 250 K data, the BF_4 counter-ion was found to be disordered around a two-fold axis, so the structure was modelled with isotropic boron and fluorine atoms.

	Temp (K)	<i>a</i> (Å)	<i>b</i> (Å)	<i>c</i> (Å)	α (°)	β (°)	γ (°)	<i>V</i> (Å ³)
Crystal 1	150	12.2544	15.8555	14.6954	89.932	113.409	90.036	2620.31
Crystal 2	150	12.2361	15.8295	14.6585	89.968	113.351	90.000	2606.69
Crystal 3	150*	12.4888	15.8540	14.6695	89.986	113.295	89.916	2616.46
Crystal 3	250†	12.3206	7.9297	14.7047	90.004	113.290	90.162	1319.55
Crystal 3	150‡	12.1662	15.7653	14.6030	89.960	113.489	89.860	2568.80
Known Cell	120	12.1844	15.7830	14.8550	90	115.176	90	2585.25
Transformed§	120	12.1844	15.7830	14.6674	90	113.569	90	2585.25

Table 1. Cell parameters of the three crystals and the structure reported by Alvarez-Rua *et al.*, including the cell for the published structure transformed to the same cell setting (for comparison).

* Cell recorded after quench-cooling to 150 K.

† Cell recorded after quench-cooling to 150 K, and then warming to 250 K at 120 K/hr.

‡ Cell recorded after quench-cooling to 150 K, warming to 250 K, followed by cooling to 150 K at 120 K/hr.

§ The matrix used to transform the known cell to the cell used here was:
$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

The cell found at 250 K had a volume that was approximately half that reported previously, and in addition, examination of the systematic absences suggested that it was C-centred monoclinic.** The lattice exceptions for the data at 150 and 250 K are shown below where the data were indexed using the smaller monoclinic cell in order to aid identification of the transition.

150 K

Lattice exceptions	P	A	B	C	I	F	Obv	Rev	All
Total no of reflections	0	2679	2679	2666	2667	4012	3571	3562	5357
Total with $I > 3\sigma(I)$	0	650	507	299	1013	728	1395	1378	2072
Mean intensity for all	0	0	0	0	6	0	6	6	6
Mean intensity $I > 3\sigma(I)$	0	1	1	0	17	1	16	17	17

250 K

Lattice exceptions	P	A	B	C	I	F	Obv	Rev	All
Total no of reflections	0	2847	2840	2833	2845	4260	3803	3791	5683
Total with $I > 3\sigma(I)$	0	346	347	3	845	348	1131	1134	1699
Mean intensity for all	0	0	0	0	5	0	6	6	6
Mean intensity $I > 3\sigma(I)$	0	2	2	0	20	2	20	20	20

This showed that at 250 K, the lattice points with $h+k$ odd are consistently weak with only three reflections for which $I > 3\sigma(I)$, whereas at 150 K there are no systematically weak conditions.

Where both data sets were processed using the larger monoclinic cell, further examination of the systematic absences showed that there were subtle changes in the systematic absences. These are most marked for the conditions $0kl$, $k=2n$ and $hk0$, $k=2n$ due to the doubling of the b-axis, confirming the presence of the phase transition:

150 K

Reflection class	For this condition	For the reverse
	$\langle F^2 \rangle$ No of reflns	$\langle F^2 \rangle$ No of reflns
hkl: $k+l=2n$	25.4703 4701	40.0190 4708
hkl: $h+l=2n$	62.2845 4722	2.9951 4687
hkl: $h+k=2n$	25.4534 4694	40.0141 4715
hkl: $h+k+l=2n$	59.3293 4691	6.3228 4718
hkl: $h+k, k+l=2n$	47.8284 2354	27.7189 7055
Ok1: $k=2n$	74.4436 235	8.4635 237
Ok1: $l=2n$	79.1775 228	5.9329 244
Ok1: $k+l=2n$	73.7489 239	8.0434 233
h0l: $h=2n$	144.9250 186	0.7745 204
h0l: $l=2n$	142.0093 190	0.6615 200
h0l: $h+l=2n$	139.7139 194	0.0488 196
hk0: $h=2n$	83.0346 178	16.0525 197
hk0: $k=2n$	93.1741 183	4.6439 192
hk0: $h+k=2n$	81.8626 184	15.0774 191
h00: $h=2n$	319.7149 6	0.3495 6
Ok0: $k=2n$	269.9869 15	-0.0638 14
00l: $l=2n$	598.9078 7	-0.2594 7

250 K

Reflection class	For this condition	For the reverse
	$\langle F^2 \rangle$ No of reflns	$\langle F^2 \rangle$ No of reflns
hkl: $k+l=2n$	20.1871 5007	33.1896 5011
hkl: $h+l=2n$	52.0580 5019	1.2225 4999
hkl: $h+k=2n$	20.2265 5020	33.1838 4998
hkl: $h+k+l=2n$	52.1519 5003	1.2909 5015
hkl: $h+k, k+l=2n$	39.0817 2514	22.5397 7504
Ok1: $k=2n$	74.4217 237	0.5342 245
Ok1: $l=2n$	72.7630 231	3.8270 251
Ok1: $k+l=2n$	69.2754 242	4.1841 240
h0l: $h=2n$	155.7308 195	-0.1574 209
h0l: $l=2n$	154.8313 196	-0.0593 208
h0l: $h+l=2n$	150.9611 201	-0.0421 203
hk0: $h=2n$	79.8464 192	14.9406 205
hk0: $k=2n$	94.6271 194	0.1757 203
hk0: $h+k=2n$	77.2262 199	15.2794 198
h00: $h=2n$	311.4204 6	0.1695 6
Ok0: $k=2n$	491.3418 14	-0.2619 13
00l: $l=2n$	657.6426 7	-0.8972 7

** In order to make comparison between the low and high temperature structures easier, the structure at 150 K was determined in the space group $P2_1/n$ rather than $P2_1/c$ (as reported before by Alvarez-Rua *et al.*). For the same reason, the high temperature structure was reported in the non-standard setting of $C2/n$, which is identical to $C2/c$, but the cell origin is located on the second inversion centre.

Examination of the systematic absences when the data were processed in the appropriate settings suggests the space group at 150 K was $P2_1/n$ (as reported previously) and that at 250 K it was $C2/n$.^{**} The 150 K structure was found to have a single whole molecule and two half molecules with two BF_4^- counter-ions in the asymmetric unit (in agreement with that reported previously). In contrast, the 250 K structure was found to have only one half molecule in the asymmetric unit, with the BF_4^- disordered around a two-fold rotation axis. In addition to the BF_4^- disorder, there are also indications from the ADPs that the pyridine ring is also moving considerably. Order-disorder transitions leading to an increase in symmetry coupled with disorder on heating are well known for example cyclohexane.⁶

Given the disorder at 150 K and the pseudo-symmetry at 150 K (discussed in depth by Alvarez-Rua *et al.*), attempts were made to solve the structure in $C2$, Cn and $C2/m$. None however, were satisfactory and it was concluded that the best space group was $C2/n$. Nevertheless, the presence of the extensive disorder in the BF_4^- , which cannot be modelled very well, leads to elevated R-indices ($R1 = 9.71\%$ ($I > -3\sigma$); 7.30% ($I > 2\sigma$), whereas at 150 K $R1 = 9.23\%$ ($I > -3\sigma$); 4.04% ($I > 2\sigma$)). To demonstrate that the disorder was the origin of the poor statistics, the BF_4^- was removed and SQUEEZE⁷ within PLATON⁸ was used to model the disorder by carrying out a discrete Fourier transform of the void region, which was then treated as a contribution to the A and B parts of the calculated structure factors. This led to a reduction in $R1$ to 7.27% ($I > -3\sigma$) and 5.07% for the strong data ($I > 2\sigma$). The large number of weak, systematic, pseudo-absences, mean it is difficult to be certain of the correct space group. This is especially true since the phase diagram has not been studied and there is uncertainty over how many phases there are and what temperature the transitions occur at. However, the results from the SQUEEZE model support the conclusion that deficiencies in the model were principally due to difficulties modelling the tumbling BF_4^- , rather than errors in the symmetry. Nevertheless, it would seem likely, that the disorder in the BF_4^- counter-ion is key to the phase transition, with the ordering caused for example, by internal pressure caused by short contacts between the Ipy_2^+ and the BF_4^- counter-ion.

A brief comparison of the two structures demonstrates that the principal difference between them is the loss of the two-fold rotor in the b direction, the c -glide perpendicular to b and the associated inversion centres (Figure 1). This leaves the 2_1 screw axis, the n -glide and (the consequent inversion centres). The loss of the blue inversion centres forces the cell to double (resulting in the iodine being now on a general position). It seems likely that the driving force for the transition is probably the BF_4^- ion, which on cooling becomes more ordered, breaking the symmetry; an effect that is propagated through the crystal.

Work is underway to further characterise the temperature dependence of this material.

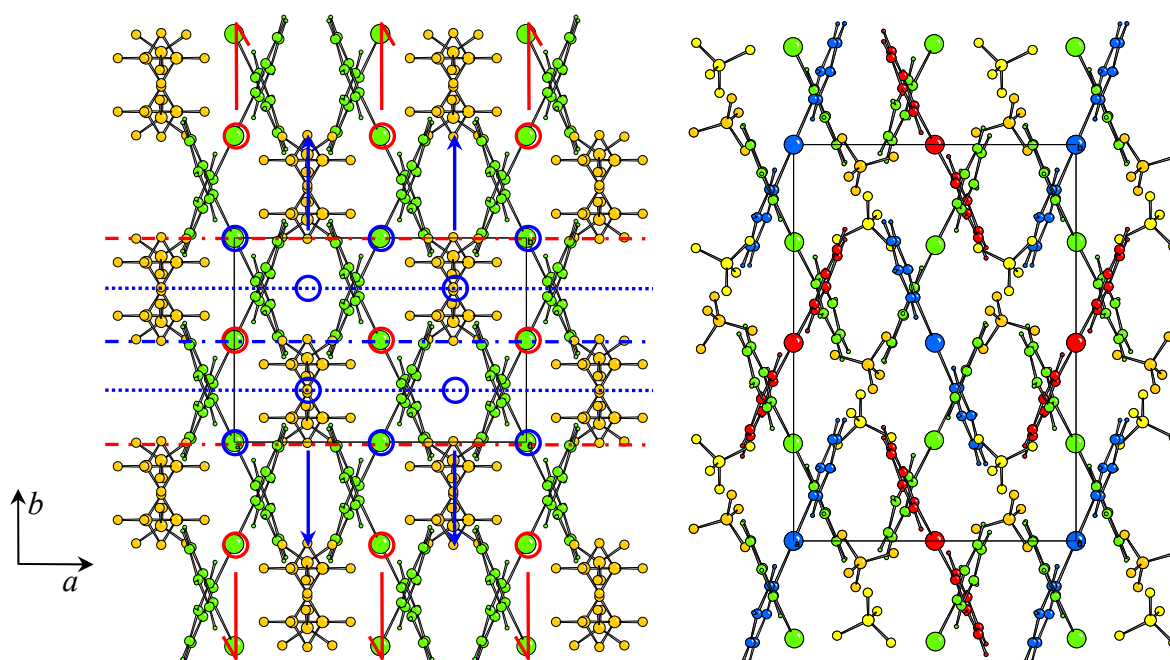
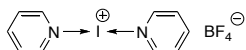
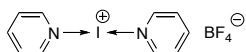
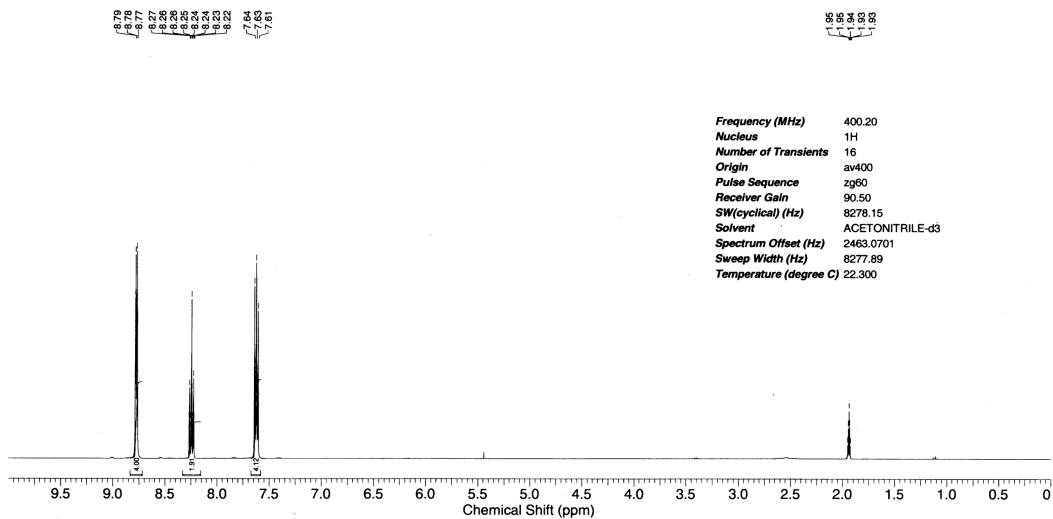


Figure 1. The structure at 250 K (left) and at 150 K (right), viewed down the *c*-axis with the *b* axis vertical. The Ipy_2^+ are shown in green, red and blue; the BF_4^- are shown in orange and yellow. The high temperature structure is overlaid with the symmetry operators for C2/n . The operators that are lost as a function of the transition are shown in blue and those retained are shown in red (arrows = two-fold axis; fish-hooks = 2_1 screw axis; dotted line = *c*-glide; dot-dash line = *n*-glide open circles = inversion centre). Thus, at low temperature the blue and red iodine atoms remain on inversion centres (I(1) and I(11)), while the green iodine is on a general position (I(21)).

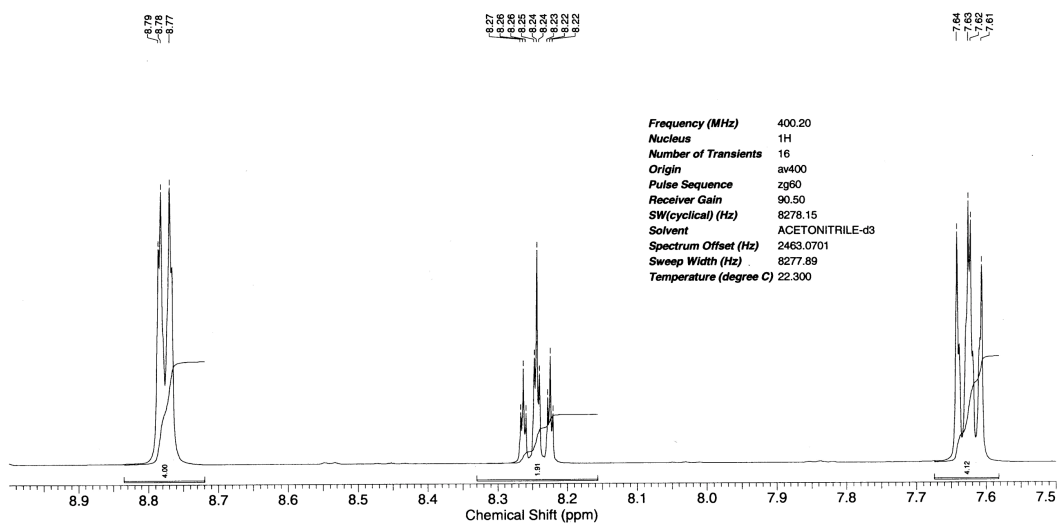
1. J. Cosier and A. M. Glazer, *J. Appl. Cryst.* (1986), 19, 105-107.
2. C. Álvarez-Rúa, S. García-Granda, A. Ballesteros, F. González-Bobes & J. M. González, *Acta Cryst.* (2002), E58, o1381-o1383.
3. Z. Otwinowski, W. Minor, Processing of X-ray Diffraction Data Collected in Oscillation Mode, *Methods Enzymol.* 1997, 276, Eds C. W. Carter, R. M. Sweet, Academic Press.
4. A. Altomare, G. Cascarano, C. Giacovazzo, A. Guagliardi, M. C. Burla, G. Polidori & M. Camalli, *J. Appl. Cryst.* (1994), 27, 435.
5. P. W. Betteridge, J. R. Carruthers, R. I. Cooper, K. Prout, & D. J. Watkin, *J. Appl. Cryst.* (2003), 36, 1487.
6. R. Kahn, R. Fourme, D. André & M. Renaud, *Acta Cryst.*, (1973), **B29**, 131-138.
7. A. Spek, *J. Appl. Cryst.* (2003), 36, 7-13; P. van der Sluis & A. L. Spek, *Acta Cryst.* (1990). A46, 194-201.
8. PLATON, A Multipurpose Crystallographic Tool, A. L. Spek. Utrecht, the Netherlands, 1998.

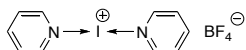


IPy₂BF₄ Precipitated from Et₂O: 1H NMR (400 MHz, CD₃CN)

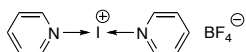
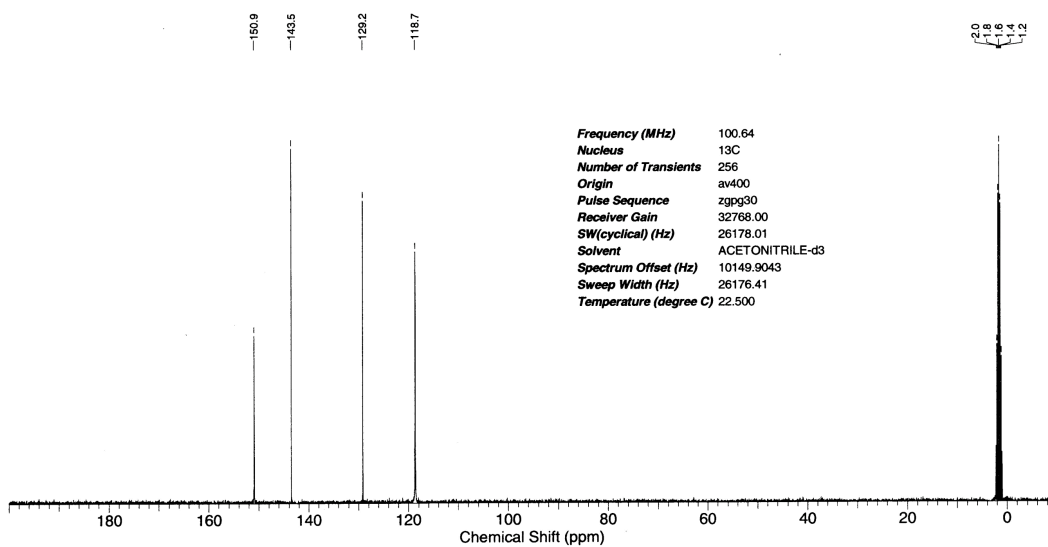


IPy₂BF₄ Precipitated from Et₂O: 1H NMR (400 MHz, CD₃CN); Expansion

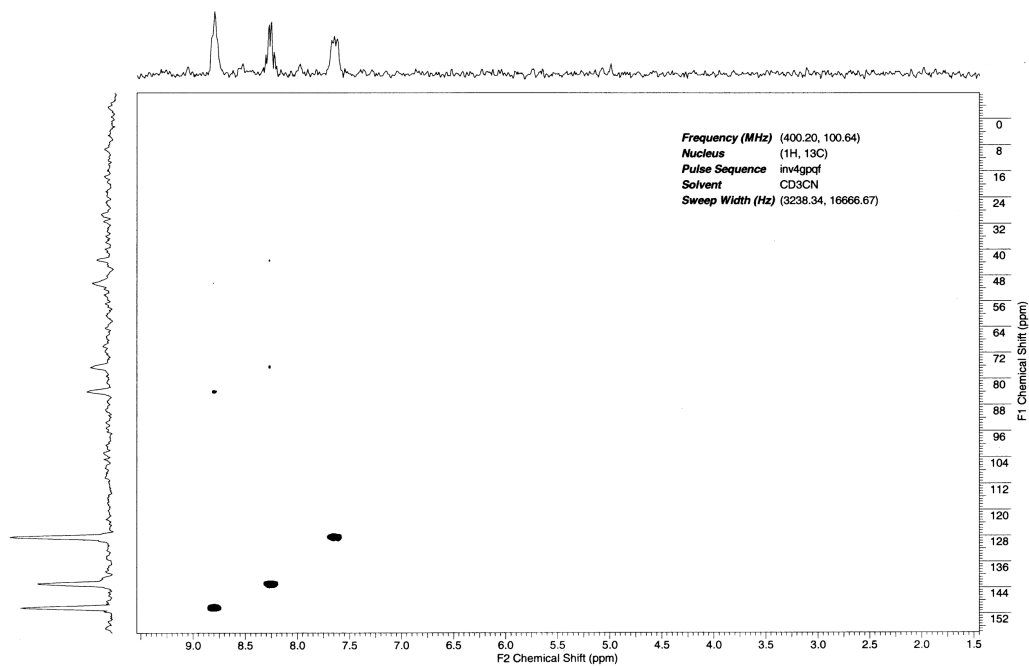


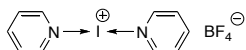


IPy₂BF₄ Precipitated from Et₂O: 13C NMR (100 MHz, CD₃CN)

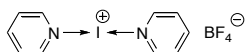
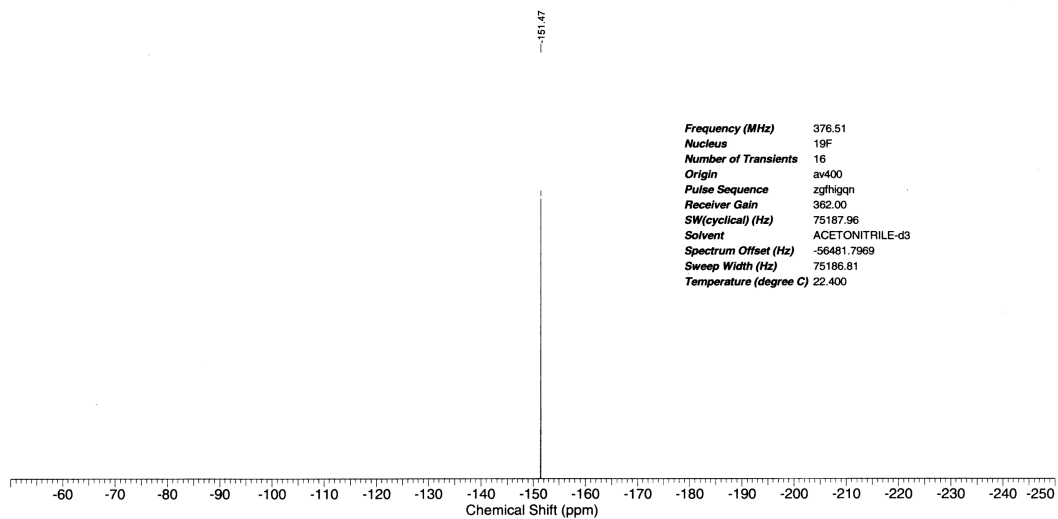


IPy₂BF₄ Precipitated from Et₂O: HMQC (CD₃CN)

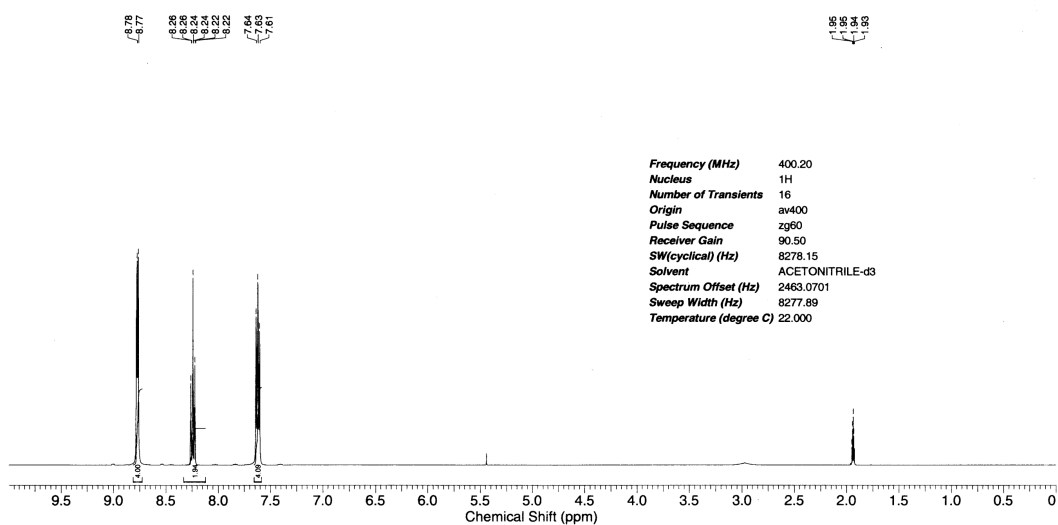


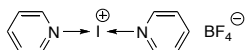


IPy₂BF₄ Precipitated from Et₂O: 19F NMR (377 MHz, CD₃CN)

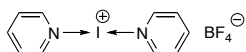
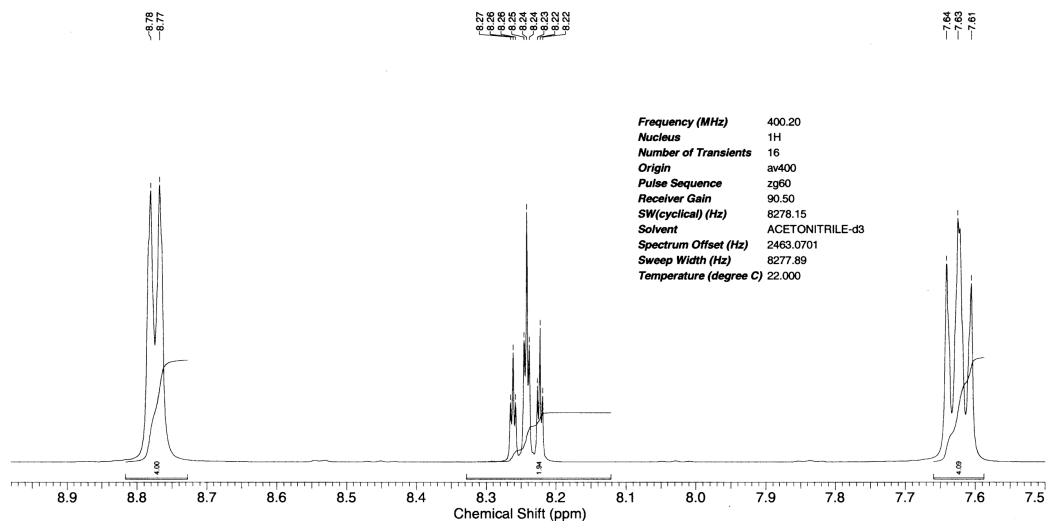


IPy₂BF₄ Recrystallized from CH₂Cl₂: 1H NMR (400 MHz, CD₃CN)

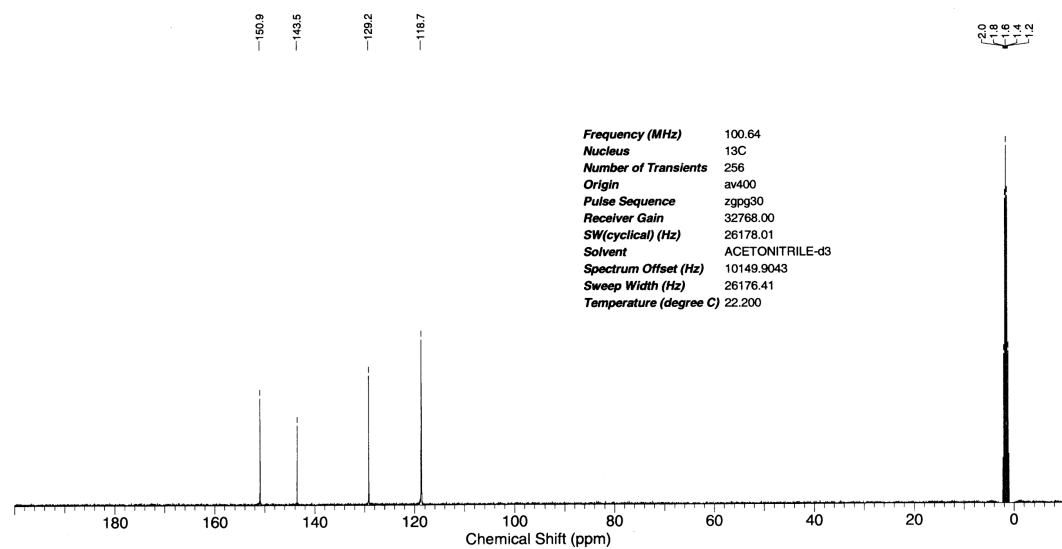


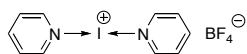


IPy₂BF₄ Recrystallized from CH₂Cl₂: ¹H NMR (400 MHz, CD₃CN); Expansion

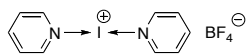
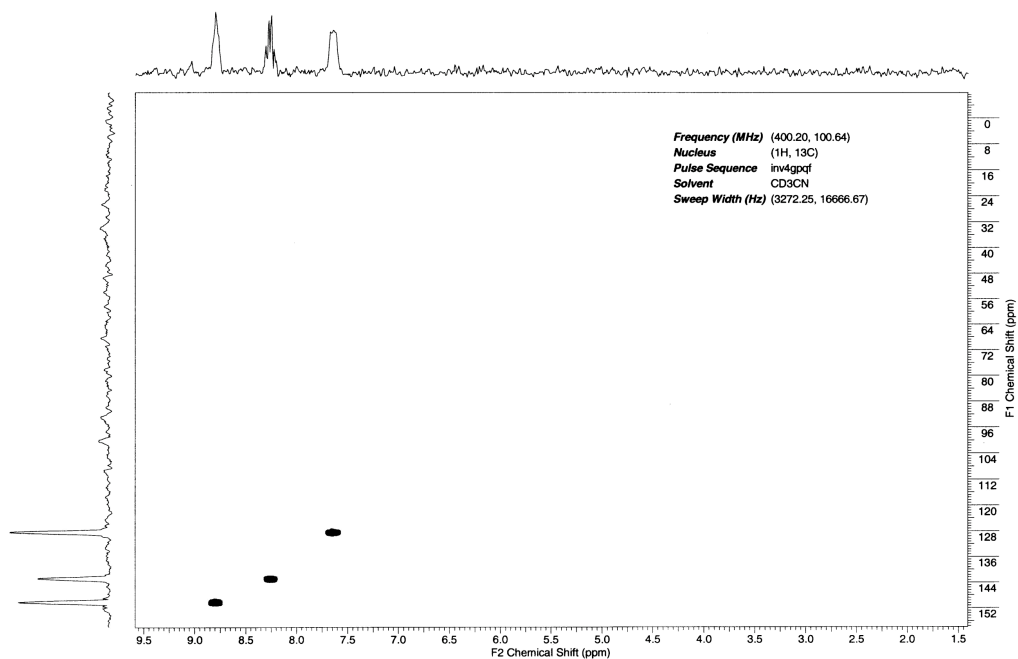


IPy₂BF₄ Recrystallized from CH₂Cl₂: ¹³C NMR (100 MHz, CD₃CN)

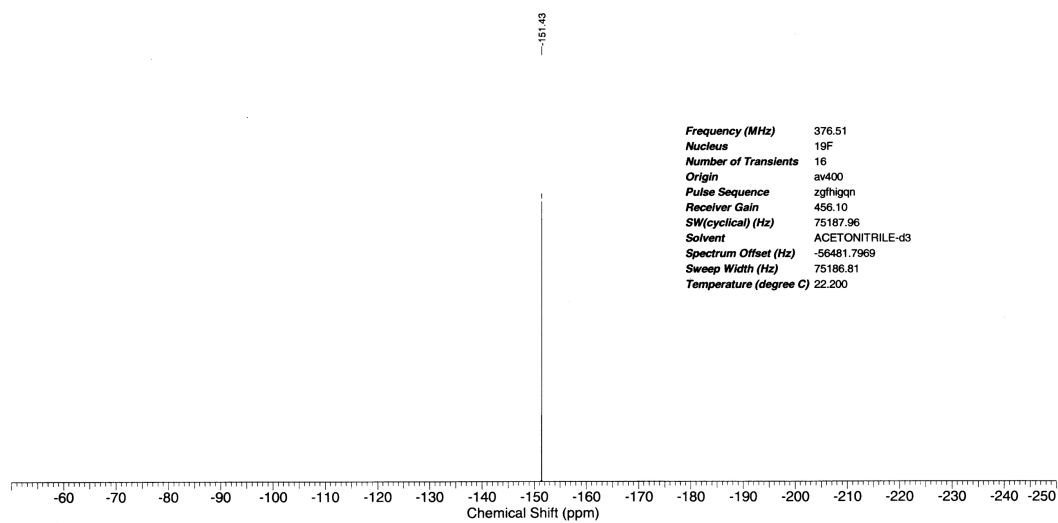


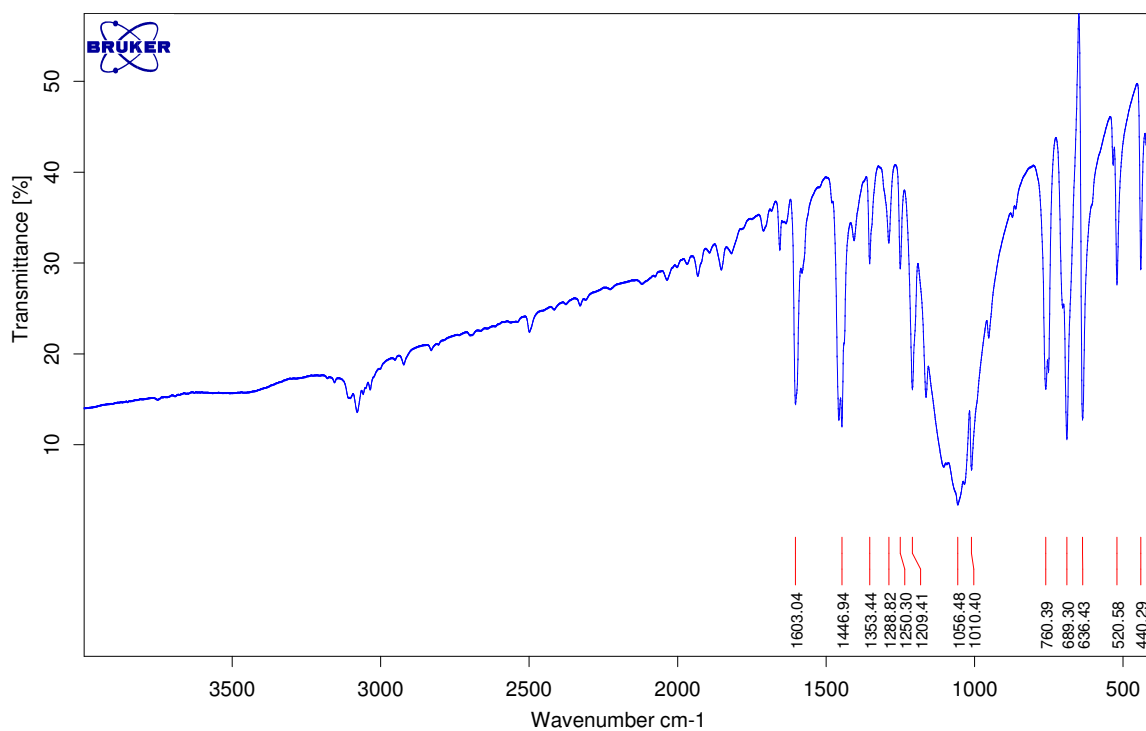


IPy₂BF₄ Recrystallized from CH₂Cl₂: HMQC (CD₃CN)



IPy₂BF₄ Recrystallized from CH₂Cl₂: 19F NMR (376 MHz, CD₃CN)



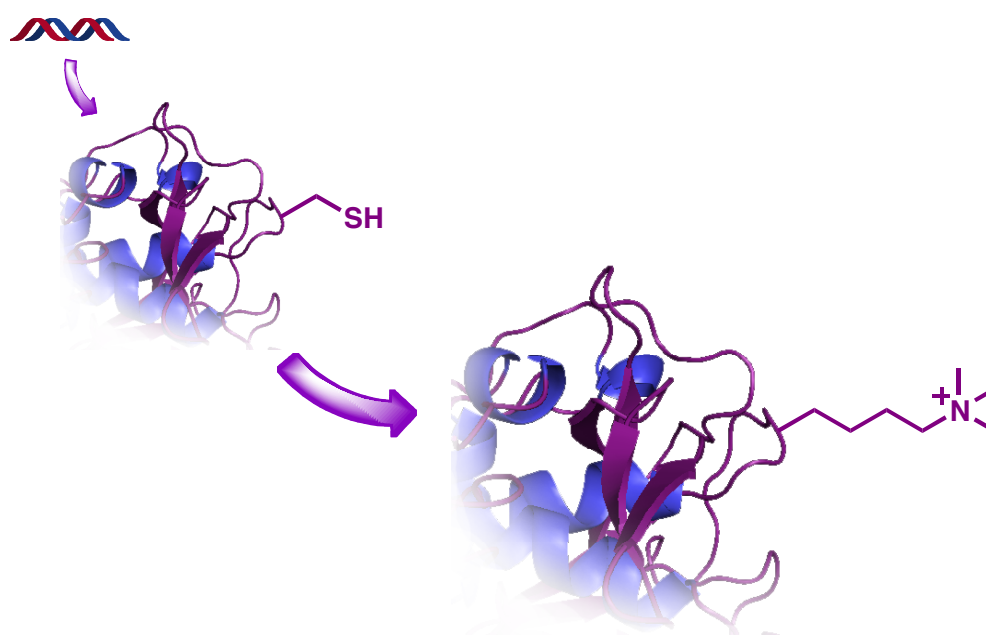


C:\Test\Test.23231 IPy2BF4 TENSOR 27, transmission

01/02/2009

Chapter 8

Post-Expression Mutagenesis of Proteins: A Strategy for the Installation of Natural, Unnatural, and Post- Translationally Modified Amino Acid Residues



“...by chemical or enzymatic means it is possible to substitute one amino acid of the protein molecule by another, simulating the effect of mutation.”

Lázló Polgár & Myron L. Bender

Advances in Enzymology and Related Areas of Molecular Biology

1970, 33, 381-400

8.1 Introduction

The development of site-directed mutagenesis revolutionized biochemistry.¹⁻⁴ Its value cannot be overstated: the precise manipulation of gene sequences and the proteins that they encode have enabled the study of enzymatic catalysis, protein-protein interactions, cellular signaling networks, and even organism level function—all at atomic resolution. Nonetheless, the genetic code and standard site-directed mutagenesis is generally limited to twenty proteinogenic amino acids. The ability to manipulate protein function by the incorporation of amino acids beyond the set of twenty canonical residues is active area of contemporary protein science.⁵⁻¹⁰ The opportunities created by such efforts are manifold: incorporation of reactive chemical groups allow further modification, activity profiling, and biophysical analysis *in vitro* and *in vivo*.¹¹⁻¹⁴ Installation of amino acids bearing post-translational modifications resolves supply issues that often hamper the study of such proteins.¹⁵⁻¹⁹ Additionally, installation of unnatural functionality to a protein can impart novel catalytic activity.^{20,21}

With such a wide array of opportunities in protein science, it is perhaps not surprising that many techniques have been developed for the precise manipulation or installation of natural, unnatural, and post-translationally modified amino acid residues. The introduction to this Thesis highlights the breadth of these efforts and so only three general strategies are listed here for reference: Biosynthetically, modified proteins can be accessed through translational machinery using amber codon suppression and other methods of genetic code reprogramming.⁵⁻⁸ Chemically, these proteins are accessible through total synthesis, due in large part to the development of automated peptide synthesis and efficient methods for peptide ligation.^{9,10,22-27} A third strategy is the selective chemical modification of the folded protein.^{11-13,28}

It is this final chemical approach that is the focus of this Chapter. Specifically, the objective of this Chapter is to explore the concept of a “chemical mutation” or a “simulated mutation” put forth by Koshland and Bender, respectively, during their pioneering work that culminated in the first point mutation of an enzyme in 1966.²⁹⁻³¹ After revisiting Koshland’s and Bender’s seminal work and several efforts in protein modification it inspired, a hypothesis will be formed that draws upon and indeed expands the original notion of Koshland’s and Bender’s mutations. Namely, it is proposed in this Chapter that “post-expression mutagenesis”—the selective chemical conversion an amino acid residue on a folded protein to a different natural, unnatural, or post-translationally

modified residue—should complement and reinforce other methods for accessing modified proteins. Before further elaboration of this proposal, it is necessary to explore its original conception by Koshland and Bender and how their ideas have been integrated with chemistry and biochemistry since their prescient studies.

Chemical mutagenesis: a milestone in protein science

The first point mutation of an enzyme was reported independently by the labs of Daniel E. Koshland, Jr. and Myron L. Bender in 1966.^{29,30} This mutation was accomplished chemically. In both cases the active site serine of subtilisin (a serine protease) was converted to cysteine by sulfonylation with phenylmethanesulfonyl fluoride (PMSF) and subsequent displacement with thioacetate. The thioacyl-enzyme obtained was hydrolyzed through the innate activity of the protease, providing the thiol-subtilisin product. The overall transformation is shown in Figure 8.1.

Koshland and Bender 1966

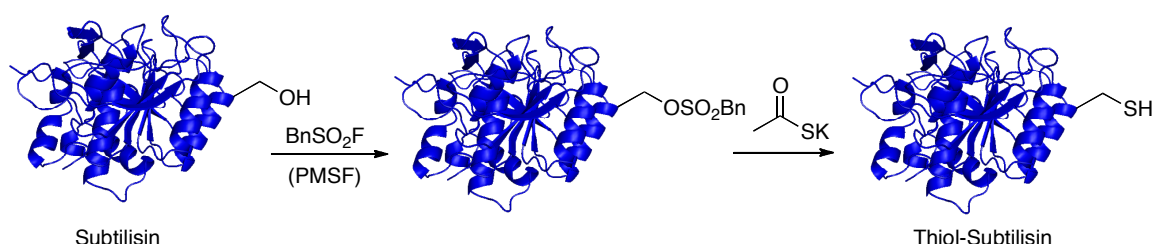


Figure 8.1: Chemical mutation of serine to cysteine in the subtilisin active site, reported independently by Koshland and Bender^{29,30}

Koshland and Bender both used the transformation in Figure 8.1 to investigate the nature of the active site of proteases. Koshland referred to the conversion of serine to cysteine as a “chemical mutation”²⁹ while Bender opted for the term “simulated mutation”³¹ so as not to confuse the process with mutations that arise from the chemical modification of DNA.^{32†} Semantics aside, Koshland and Bender accomplished the first point mutation of an enzyme before site-specific mutations at the DNA level were possible. Remarkably, both Koshland and Bender anticipated precise genetic manipulation to alter protein structure,^{29,31} but until such technology was available they

[†] Throughout this Chapter the term “chemical mutation” will refer exclusively to the chemical conversion of one amino acid to another on full-length protein after expression or total synthesis.

proposed the use of chemistry on the expressed and folded protein to access mutant enzymes.

Soon after Koshland and Bender's landmark papers, Laskowski reported an enzyme-mediated mutation on soybean trypsin inhibitor. Laskowski's "enzymatic mutations" were carried out using trypsin and carboxypeptidase B to both excise an amino acid and re-ligate a different amino acid at the site of mutation. The equilibrium could be tilted toward the mutant soybean trypsin inhibitor by careful control of amino acid and trypsin/carboxypeptidase concentrations. The overall process is shown in Figure 8.2.

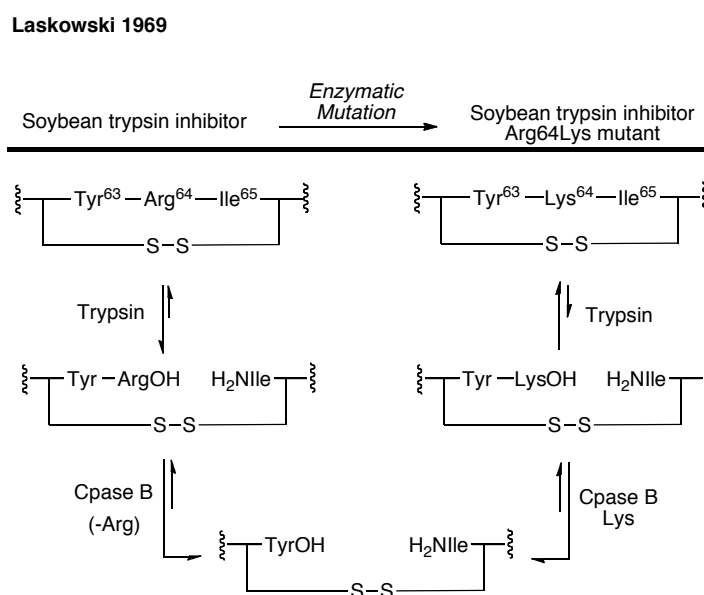


Figure 8.2: Laskowski's enzymatic mutation of arginine to lysine at position 64 in soybean trypsin inhibitor³³

Laskowski's strategy was further developed by Tschesche who took advantage of various proteases to expand the type of residues that could be re-ligated into the soybean trypsin inhibitor.³⁴ Tschesche later used sequential amino acid excision, chemical coupling, and elastase ligation to obtain additional mutants.³⁵ Importantly, the reports from Laskowski and Tschesche demonstrated that chemical mutations are not limited to selective chemical reactions, but can employ judicious use of enzymes.

Nearly a decade after Koshland and Bender's chemical mutations on subtilisin, Peter Clark and Gordon Lowe published their "chemical mutations of papain."^{36,37} In these reports, a cysteine residue was used as a chemical precursor to formylglycine, serine, and glycine. The transformations are summarized in Figure 8.3. Cysteine 35 of

papain was first alkylated with a phenacyl bromide. Photolysis of the resulting intermediate provided the thioaldehyde, dehydrocysteine. Hydration of the thiocarbonyl and loss of hydrogen sulfide led to the formation of formylglycine (Figure 8.3A). A final reduction with NaBH_4 completed the chemical mutation of cysteine to serine (Figure 8.3B). Alternatively, formylglycine could be converted to glycine by treatment with sodium hydroxide (Figure 8.3C).

Clark and Lowe 1977, 1978

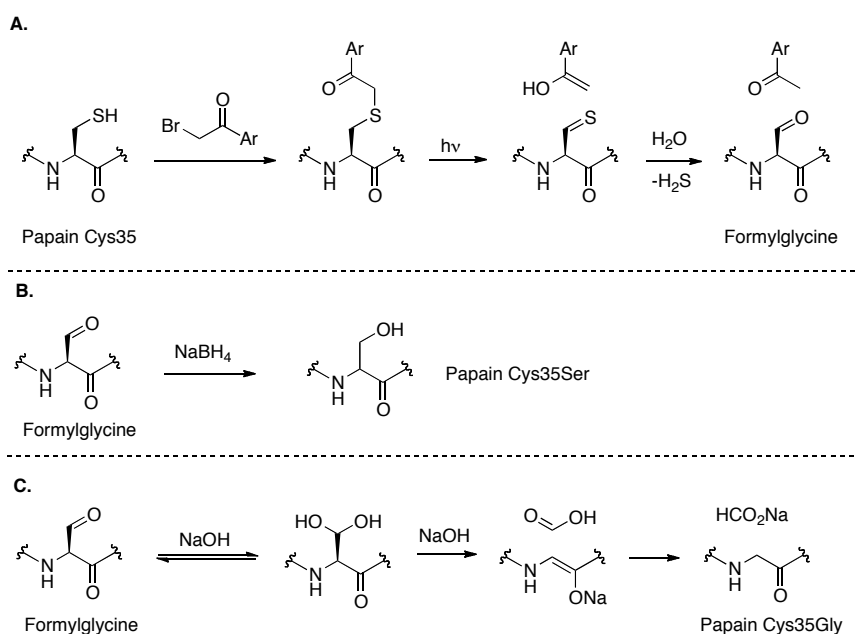


Figure 8.3: Clark and Lowe's "chemical mutations of papain"^{36,37}

Unlike the chemical mutations described thus far, Clark and Lowe's mutations of cysteine are not substrate dependent and, in principle, could be applied to any free cysteine residue. Koshland and Bender's mutations relied on the unique reactivity of the active serine in subtilisin while Laskowski and Tschesche's mutations were restricted to soybean trypsin inhibitor and its innate affinity for the proteases and ligases employed in their method. Notwithstanding the relatively limited substrate scope, the reports from Koshland, Bender, Laskowski, Tschesche, and Lowe described the most precise manipulation of protein sequence that preceded site-directed mutagenesis of DNA.

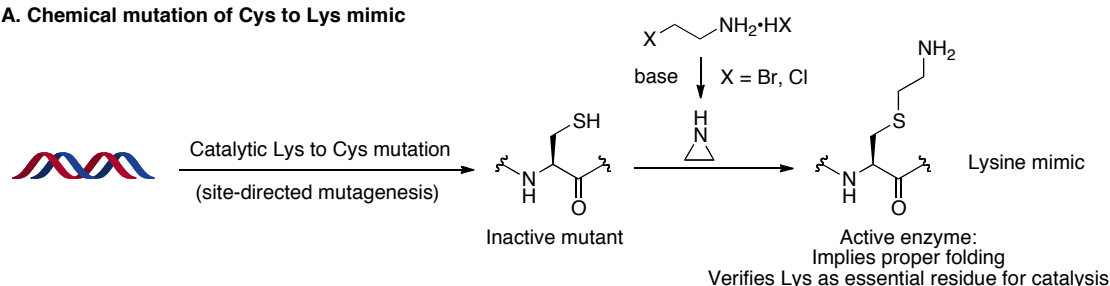
In the decade following Clark and Lowe's publications, genetic technology developed at a rapid pace¹⁻⁴ and the concept of chemical mutagenesis may have seemed obsolete. Other than reports on the chemical interconversion of asparagine and glutamine to aspartic and glutamic acid, respectively, few investigations of chemical mutations were

disclosed.^{38,39} It was not until the late 1980's and through the turn of the century that several scenarios were encountered where researchers benefited from chemical mutations or a dual and reinforcing use of genetic and chemical mutation. These situations are discussed next to demonstrate how the conceptual advances of Koshland and Bender serve the protein chemist even nearly half a century after their inception.

Dual use of genetic and chemical mutations to investigate enzyme catalysis

Site-directed mutations at the genetic level are now routinely used to investigate catalytically relevant residues in enzymes. However, if a catalytically inactive mutant is obtained it is not always straightforward to attribute inactivity to the altered active site or to misfolding that results from such a mutation. If the residue of interest is lysine, a genetic mutation to cysteine will often result in diminished catalytic activity. If the activity is recovered by chemical conversion of cysteine to a lysine analog by aminoethylation, the altered activity can be attributed to the mutation rather than misfolding. The strategy, outlined in Figure 8.4A, has been used in several investigations of catalytic lysine residues.⁴⁰⁻⁴⁴ Related strategies have been employed in the studies of arginine residues as well (Figure 8.4B).⁴⁵

A. Chemical mutation of Cys to Lys mimic



B. Chemical mutation of Cys to Arg mimic

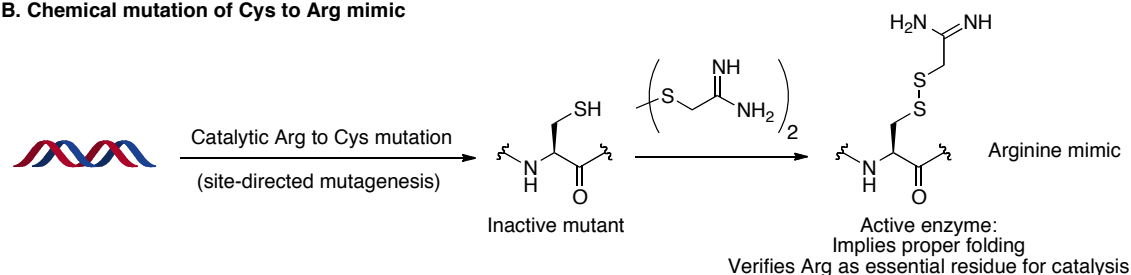


Figure 8.4: Combined use of genetic and chemical mutation to investigate catalytic lysine and arginine residues

As a brief aside, aminoethylation of cysteine as a method to chemically install a lysine analog (Figure 8.4A) is one of the earliest methods for chemoselective protein modification. As early as 1956 this method was used to install additional sites for protease action since enzymes such as trypsin and carboxypeptidase-B cleave peptides and proteins at lysine residues. Thus, while this method provides only an analog of lysine, it is a functional analog that aided early efforts in protein sequencing.⁴⁶⁻⁴⁸ As a functional mimic of lysine, the aminoethylation of cysteine can be considered a chemical mutation from cysteine to lysine—one of the earliest experimental demonstrations of the chemical mutation strategy later formalized by Koshland and Bender.

Chemical mutations pertaining to noncanonical amino acids

The installation or modification of unnatural amino acid residues is another instance where chemical mutations are useful. Taking a cue from Bender and Koshland, Hilvert converted the active serine of subtilisin to selenocysteine by sequential reaction with PMSF and hydrogen selenide (Figure 8.5).⁴⁹ The “selenosubtilisin” obtained was remarkably active as an acyltransferase in contrast to the native protease activity of subtilisin itself. While selenocysteine is sometimes incorporated naturally into proteins,⁵⁰ Hilvert’s chemical route provides convenient access to this residue. Related studies have since been reported where a chemically installed selenocysteine^{51,52} or tellurocysteine⁵³ was used to impart glutathione peroxidase^{51,53} or deiodinase⁵² activity to the modified protein.

Hilvert 1989

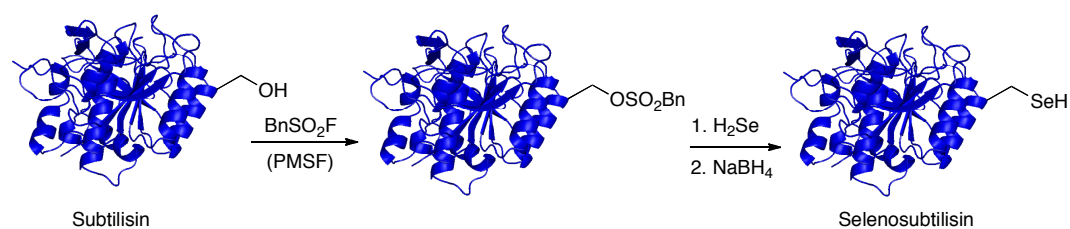


Figure 8.5: Chemical mutation of serine to selenocysteine at the active site of subtilisin

In the examples just discussed, serine was chemically mutated to a noncanonical residue. In the reverse sense, an unnatural residue can be converted to a natural residue by chemical mutation. Photocaged or otherwise protected amino acids have been incorporated into proteins on many occasions either through amber codon suppression or

by total chemical synthesis.⁵⁴⁻⁶¹ Chemical mutation (in this case deprotection) reveals the natural residue. A selection of these residues are shown in Figure 8.6.

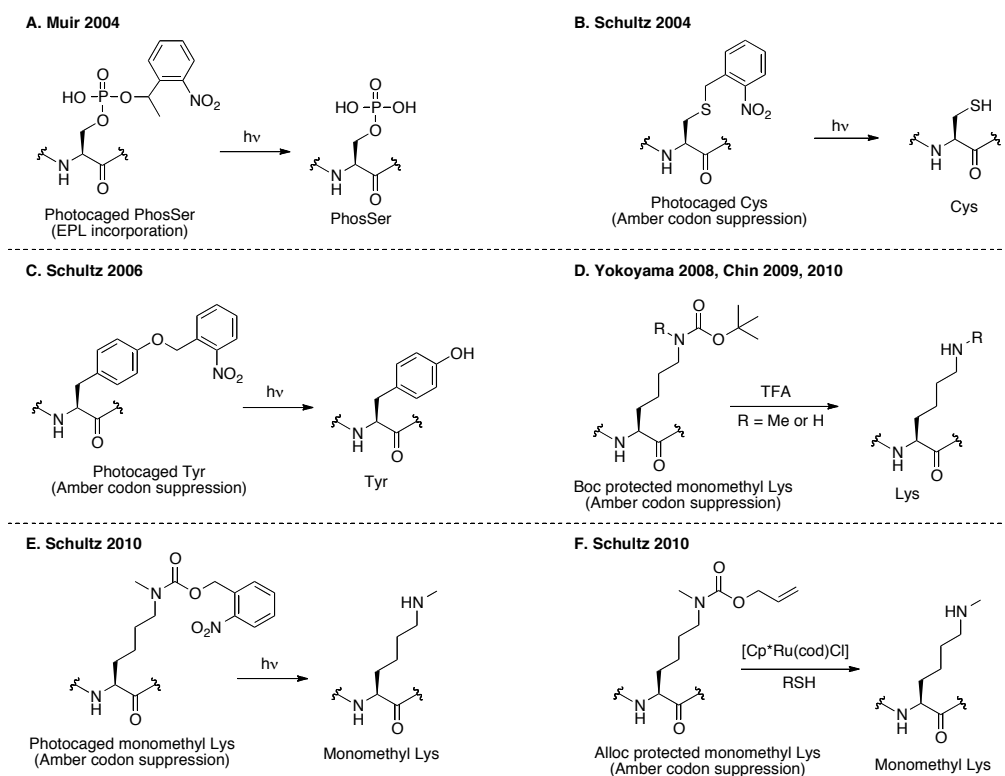


Figure 8.6: Protected amino acids can be deprotected by chemical mutation. These residues can be incorporated into proteins by synthesis or translation.

The chemical mutation of an unnatural residue to a native sidechain has also been exploited by the Schultz group in traceless affinity chromatography.⁶² Genetically encoded *p*-boronophenylalanine can be used to bind an expressed protein to a polyhydroxylated resin. Elution with hydrogen peroxide oxidizes the unnatural residue to tyrosine; elution with silver diammonium nitrate reduces the affinity tag to phenylalanine (Figure 8.7).⁶²

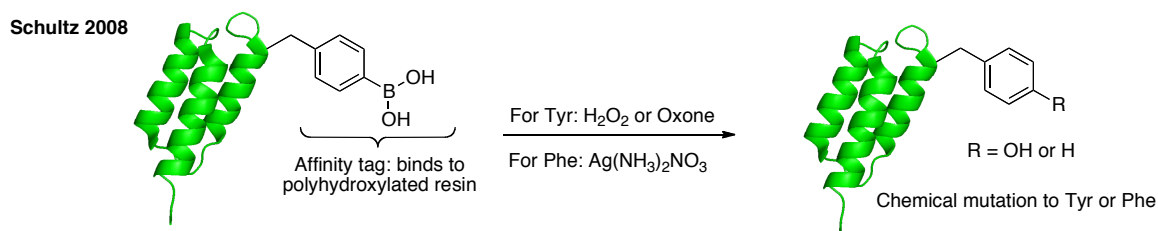


Figure 8.7: Chemical mutation of *p*-boronophenylalanine to tyrosine or phenylalanine allows traceless affinity chromatography.

Chemical mutations post-native chemical ligation

The general utility of native chemical ligation²⁴ and expressed protein ligation^{25,26} in the synthesis of proteins is evident in the large number of targets prepared by these methods.^{9,10,27} One apparent limitation of these methods is the need for cysteine at the site of ligation. If, however, cysteine or a related derivative can be chemically mutated to a different residue after ligation, the ligation is effectively carried out at this final residue. Using cysteine as the penultimate residue in native chemical ligation was first advanced by Yan and Dawson who showed that desulfurization of cysteine after ligation provided alanine.⁶³ The authors could therefore use native chemical ligation to synthesize peptides and proteins that do not contain cysteine residues. Related strategies for the chemical mutation of cysteine derivatives to natural residues following native chemical ligation have since emerged. This concept allows ligation, in the formal sense, at several other residues.⁶⁴⁻⁷¹ A selection of these chemical mutations is shown in Figure 8.8.

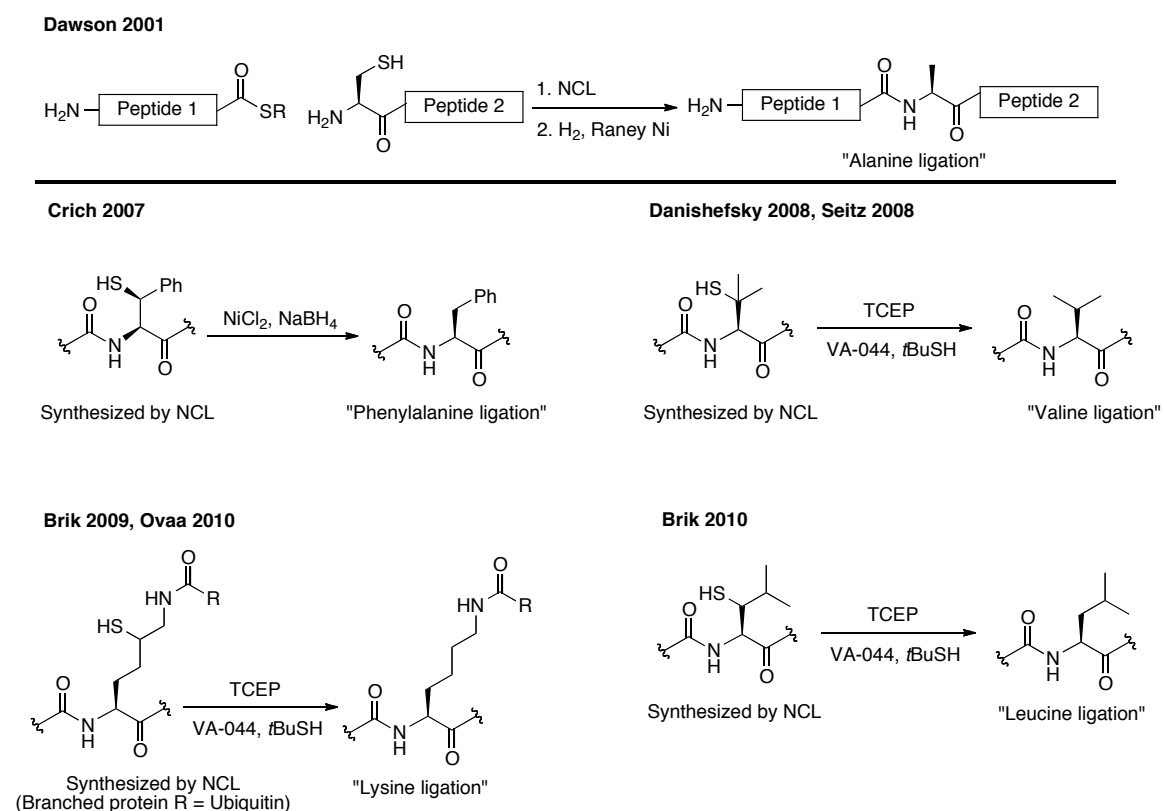


Figure 8.8: Chemical mutation of cysteine and related derivatives after native chemical ligation allows the synthesis of cysteine free peptides and proteins.

Chemical mutations to access natural residues bearing post-translational modifications

As discussed in Chapters 3 and 4, chemical access to proteins bearing post-translational modifications is one attractive way to supply such proteins for biochemical and biophysical analysis. The concept of chemical mutagenesis was implicit in many of the protein syntheses discussed in these Chapters. For instance, incorporation of amino acids bearing photocaged PTMs or other forms of protection rely on a selective chemical mutation to reveal the natural residue.^{54,58-60} Several examples in Figure 8.6 depict such transformations. Many PTMs, however, are not accessible by a simple chemical reaction on a translated, folded protein. In these cases, conversion to a suitable mimic may suffice. While one of the primary interests in this Chapter is chemical access to post-translational residues in their natural form, it is worth pointing out several PTM mimics that are accessible by chemical mutation. As shown in Figure 8.9, the chemical mutation of a suitable amino acid precursor can provide access to mimics of glycosyl serine⁷²⁻⁷⁴ and asparagine,^{75,76} phosphorylated⁷⁷ or sulfated⁷⁸ tyrosine, phosphorylated serine,⁷² and methylated^{72,74,79} and acetylated lysine.^{74,80}

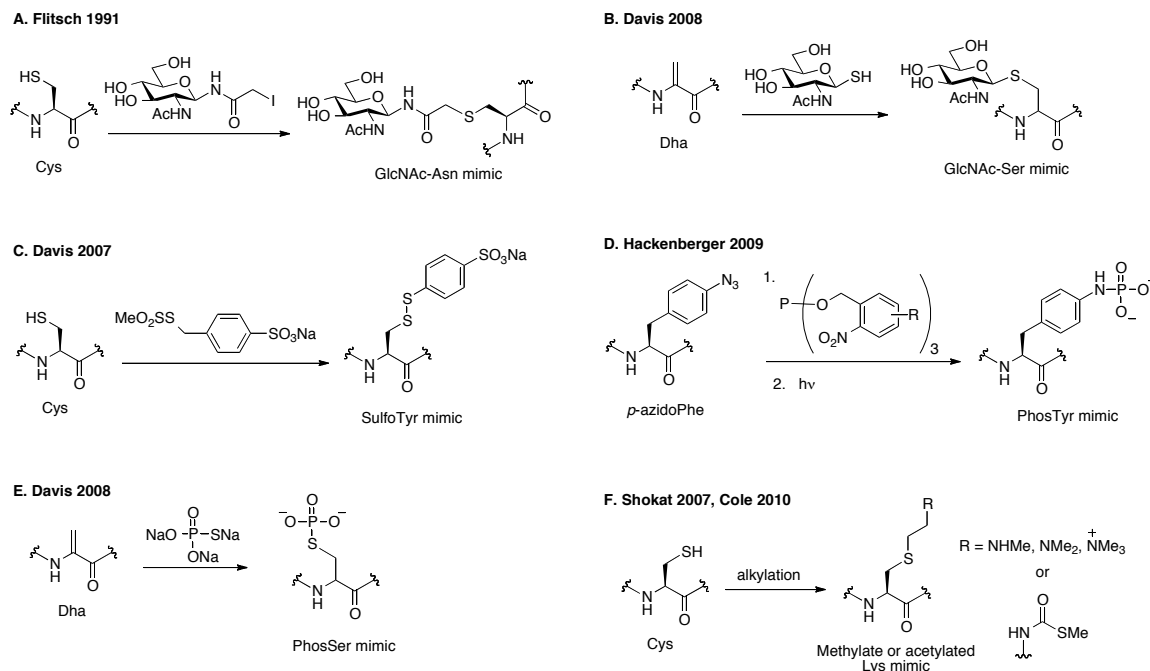


Figure 8.9: Chemical mutations that provide access to proteins bearing PTM mimics

Since the PTM mimics in Figure 8.9 were discussed and in some cases examined experimentally in Chapters 2, 3, and 4, their use in biochemistry will not be discussed

further. For the purposes of this Chapter, the examples in Figure 8.9 highlight an opportunity in chemical mutagenesis: the selective chemical conversion of an amino acid residue (natural or unnatural) to an amino acid residue bearing a post-translational modification in its natural constitution, rather than an analogous structure. Next, the prospects of consolidating and generalizing these efforts are discussed, setting the stage for the experimental pursuits of this Chapter.

Towards a general chemical mutagenesis

Many examples were described in the previous sections that embody the concept of chemical mutagenesis. However, many of these demonstrations are highly substrate specific or allow conversion to only a single type of residue. For instance, the mutation of serine to cysteine reported by Koshland and Bender relies on the unique reactivity of serine proteases and the chemistry employed does not accommodate a mutation from that same serine to other residues (such as lysine, methionine, phenylalanine, etc).^{29,30} Clark and Lowe advanced a more general concept of chemical mutagenesis wherein one residue, cysteine, was shown to be a chemical precursor to serine or glycine.^{36,37} To further develop the concept of chemical mutagenesis into a method that generally complements genetic and recombinant methods for protein engineering, reactions must be developed that enable the conversion of a uniquely reactive residue to a diverse set of amino acid sidechains. In the search for such a residue, a cue can once again be taken from Koshland. In his studies of chymotrypsin, the active site serine was converted to dehydroalanine.^{81,82} Koshland proposed that dehydroalanine could be used as a precursor to other amino acid residues by the addition of a nucleophile.⁸² Indeed Koshland demonstrated such additions experimentally to chemically corroborate the formation of dehydroalanine.^{81,82}

An exercise in retrosynthesis reveals that several proteinogenic residues should be accessible from dehydroalanine (Figure 8.10). For instance, dehydroalanine can be converted to alanine by hydrogenation. For other sidechains, a carbon-carbon bond must be formed at the β -position of dehydroalanine. A review of the literature reveals many options for such transformations that are compatible with buffered water—a requirement for protein substrates. Two candidate transformations are particularly attractive. First, aliphatic side chains should be accessible by the addition of alkyl radicals to dehydroalanine. Pioneering studies by Luche demonstrated that alkyl halides and zinc

metal react to form radicals that undergo conjugate addition to a wide variety of α,β -unsaturated electrophiles in aqueous media.⁸³⁻⁸⁸ Moreover, Li has used this reaction in the synthesis of amino acids.⁸⁹ Second, aromatic side chains should be accessible by the Rh(I)-catalyzed conjugate addition of arylboronic acids.⁹⁰ While this reaction is typically carried out at higher temperature, its use on peptides and in aqueous solvent bodes well for application to protein substrates.^{91,92} Finally, prospects for diastereoselective variants of these transformations—particularly in the conversion of dehydroalanine to alanine and the aromatic side chains⁹⁰⁻⁹⁹—may overcome the limitations that complicate simple thiol addition to Dha.

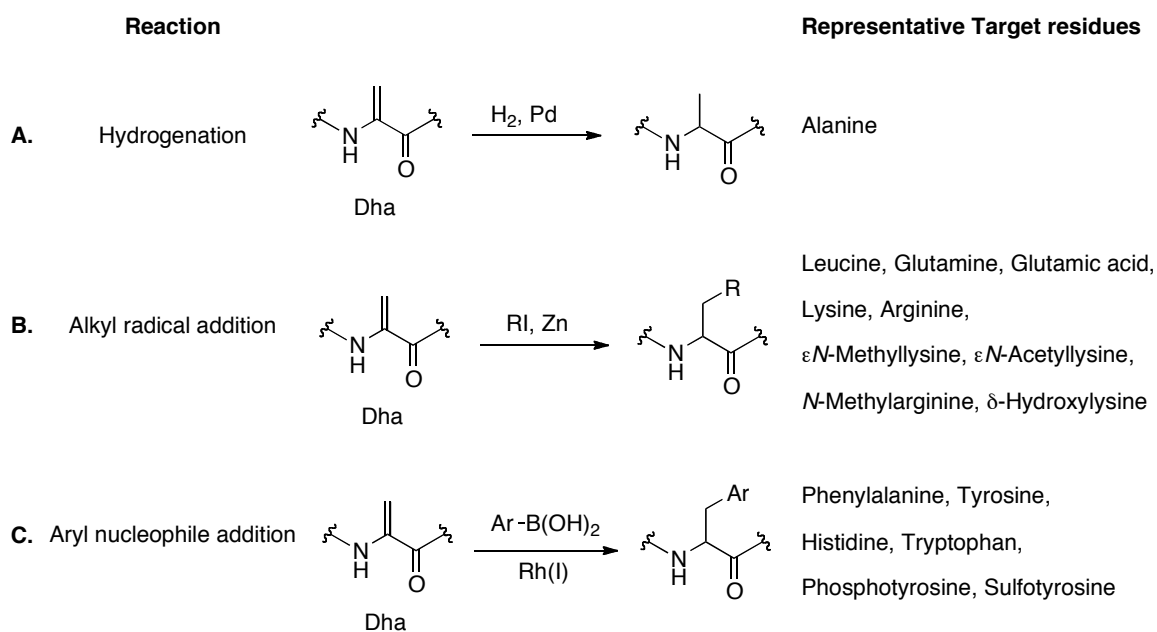


Figure 8.10: Dehydroalanine as a precursor to other amino acids

The remainder of this Chapter is concerned with a preliminary validation of dehydroalanine as a general precursor to other amino acid residues on proteins. Reactions of dehydroalanine and non-thiol nucleophiles are explored, with a focused effort on the transformation in Figure 8.10B. In general, all of the transformations in Figure 8.10 serve as a guide in this experimental pursuit of Koshland's conjecture that dehydroalanine is generally useful in chemical mutagenesis.

8.2 Results and Discussion

Chemical mutation of dehydroalanine to a histidine analog

The addition of thiol nucleophiles to dehydroalanine was explored in Chapters 2, 3, and 4. These additions are operationally simple and provide access to several mimics of post-translational modifications. Before launching an investigation into the transformations outlined in Figure 8.10, a first objective was to determine whether other heteroatom nucleophiles add to dehydroalanine on protein substrates. The addition of *N*-nucleophiles to dehydroalanine-containing peptides has been reported and provides histidine and tryptophan analogs.¹⁰⁰⁻¹⁰² For instance imidazole can add to dehydroalanine to provide an analog of histidine. Since *N*-nucleophiles typically react more slowly with dehydroalanine than thiols,^{101,102} it was not clear whether such a reaction was suitable for protein substrates. However, when imidazole was added to dehydroalanine-containing subtilisin mutant **2.35**—the model protein prepared in Chapter 2—complete conversion to the product of histidine addition (protein **8.1**) was observed over the course of 5 hours at 37 °C (Figure 8.11). The full course of the reaction, followed by LC-MS, can be found in the experimental section.

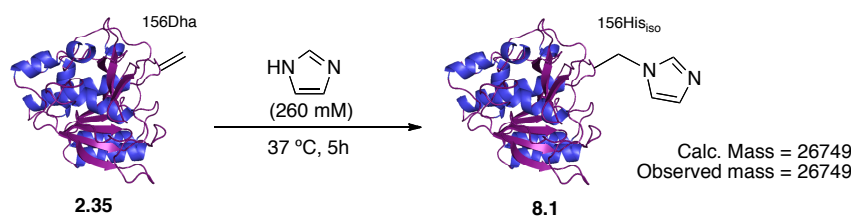


Figure 8.11: Addition of imidazole to dehydroalanine

The addition of imidazole to dehydroalanine provides an analog of histidine that might prove useful in investigating these residues in catalysis. The incorporation of this residue into enzyme active sites is an ongoing effort of which the results are too preliminary to report here. Nevertheless, the result in Figure 8.11 suggests that nitrogen nucleophiles can add to proteins that contain dehydroalanine and analogs of histidine are a potential target.

Chemical mutation of dehydroalanine to alanine

Hydrogenation provides a route from dehydroalanine to alanine. This reaction was tested on protein **2.35**. A water-soluble palladium source was prepared using Pd(OAc)₂ and the

water-soluble phosphine TPPTS.¹⁰³ Hydrogenation was carried out under an atmosphere of hydrogen for 3 hours at 37 °C. Analysis by LC-MS revealed a protein with the appropriate mass, but the resolution of the spectrometer is not sufficient to distinguish the starting material and product. To show that dehydroalanine was consumed in the reaction, thiol **2.38** was added to product protein. No addition was observed, consistent with conversion of dehydroalanine to alanine. The presence of palladium does not prevent the addition of **2.38** to dehydroalanine: in a control reaction, thiol **2.38** smoothly added to the dehydroalanine of protein **2.35** even in the presence of palladium.

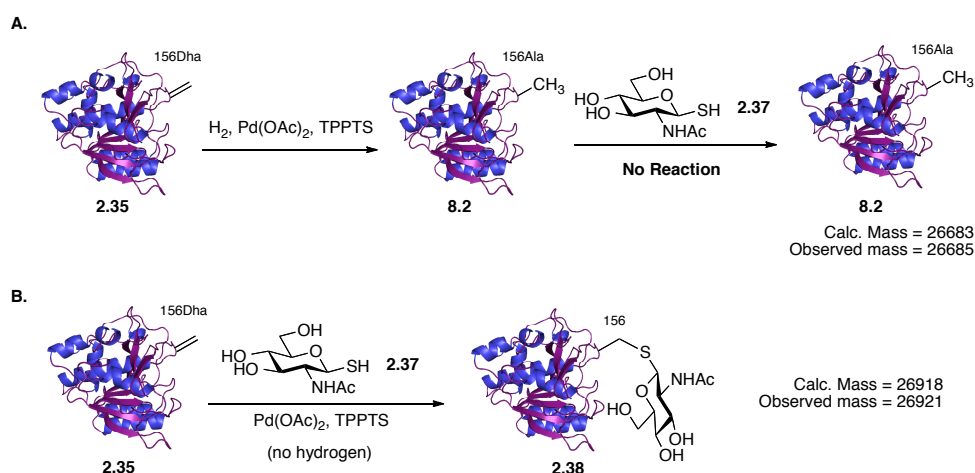


Figure 8.12: Hydrogenation of dehydroalanine to alanine and chemical test to corroborate the reduction

The conversion of dehydroalanine to alanine could potentially be carried out in an asymmetric fashion, though such hydrogenations in water are rare since the majority of chiral ligands are not soluble in water.¹⁰⁴ Moreover, as seen in Figure 8.8, a variety of methods are already established for the direct conversion of cysteine to alanine by desulfurization.^{63,64} A more important contribution to the concept of chemical mutagenesis would be the addition of carbon nucleophiles to dehydroalanine. The exploratory work for such transformations is discussed next.

Model studies for carbon-carbon bond formation at dehydroalanine

The zinc-mediated generation of alkyl radicals from alkyl iodides and their subsequent addition to α,β -unsaturated electrophiles is a reaction well-poised for peptide and protein modification: it is typically carried out at room temperature in aqueous solvent⁸³⁻⁸⁸ and has even been used in the synthesis of amino acids.⁸⁹ Particularly intriguing is the

prospect of installing amino acids into proteins that are not currently accessible by translation (e.g. trimethyl lysine, among others shown in Figure 8.10B). To ascertain the scope, limitations, and general chemical behavior of this transformation, several model reactions were examined. The first of which was the chemical conversion of AcDhaOMe (**2.4**) to AcLeuOMe (**8.3**) (Table 8.1).

At the outset of these studies, no effort was made to exclude oxygen and all reactions were run at room temperature or lower. Ammonium chloride has been shown to activate zinc metal¹⁰⁵ and was used in Luche's original studies,^{83-85,88} so a similar additive, ammonium acetate, was used in the model studies reported here. Ammonium acetate can be buffered (important for protein samples) and the pH was adjusted to 6.0. A variety of conditions were examined and the results are summarized in Table 8.1.

<chem>CC(=O)NC(=C)C(=O)OC</chem> $\xrightarrow[\text{NH}_4\text{OAc, solvent, additives}]{\text{I-CH(CH}_3)_2, \text{Zn powder}}$ <chem>CC(C)C(C(=O)OC)NC(=O)C</chem> AcDhaOMe (2.4) AcLeuOMe (8.3)								
Entry	Temp.	Zn (eq.)	<i>i</i> PrI (eq.)	NH ₄ OAc (eq.) 5M, pH 6.0	Solvent	Additive (eq)	Time	Yield
1	RT	15	2	10	MeOH	None	30 min	28%
2	RT	15	10	10	MeOH	None	30 min	92%
3	RT	15	10	10	H ₂ O	None	30 min	64%
4	RT	15	10	10	H ₂ O	None	4 hr	64%
5	RT	15	10	10	H ₂ O	Nal (10 eq)	30 min	64%
6	RT	3 x 15	3 x 10	10	H ₂ O	None	30 min	87%

Table 8.1: Model chemical mutation of dehydroalanine to leucine

As a starting point, methanol was used as the solvent to help dissolve both AcDhaOMe (**2.4**) and 2-iodopropane. Zinc was used in excess in all experiments and ammonium acetate was added as a 5M solution in water. Using only 2 equivalents of 2-iodopropane, 28% yield of AcLeuOMe (**8.3**) was obtained after only 30 minutes of reaction time (Table 8.1, Entry 1). This result was promising since the desired product

was obtained and no side reaction was observed—unreacted **2.4** was the only other material observed. Suspecting the reduction of 2-iodopropane to propane was a competing process, the reaction was repeated with 10 equivalents of 2-iodopropane (Entry 2). Gratifyingly, using an excess of the alkyl iodide resulted in excellent isolated yield (92%) of AcLeuOMe (**8.3**).

Solvent was examined next. While the excellent yield in MeOH was promising, purely aqueous conditions are more suitable for protein samples. In Entry 3, the same reaction was tested, again with excess 2-iodopropane, except MeOH was replaced by H₂O. While AcDhaOMe (**2.4**) was soluble in the reaction mixture, 2-iodopropane and zinc are not soluble in water. Nevertheless, a moderate yield (64%) of the desired product was obtained after only 30 minutes of reaction. Extending the reaction time to 4 hours did not improve the yield (Entry 4). Luche has shown that the addition of NaI in similar reactions has improved yields,⁸⁸ possibly by serving as the final reductant,¹⁰⁶ but no such improvement was observed in this case (Entry 5). The low conversions were finally overcome by simply adding additional 2-iodopropane and zinc metal (Entry 6). In this final entry, only water was used as the solvent and both 2-iodopropane and zinc metal were added in excess, in three portions. The isolated yield was 87% after only 30 minutes of reaction time at room temperature.

A few observations of the reactions in Table 8.1 are worth noting. First, the reaction between zinc and 2-iodopropane under these conditions is fast and exothermic. In fact, most zinc metal is dissolved by the end of the reaction. For larger scale reactions or for protein samples, the exotherm might need to be controlled by slow addition of the iodide or lower temperature. Since excess 2-iodopropane was required, but longer reaction time did not improve yields, it appears that reduction to propane is the predominant side reaction. As long as the reduced product does not interfere with the reaction, simply adding additional alkyl iodide should resolve this problem. Finally, it is remarkable that the reaction is so fast despite the presence of three phases: zinc as a solid, 2-iodopropane as an organic liquid, and the aqueous solution of dehydroalanine. While the yields obtained suggest miscibility of reagents is not essential, vigorous stirring was used to ensure efficient mixing.

Several other alkyl iodides were tested next to assess the generality of the addition to dehydroalanine and determine any caveats or clues that might aid the deployment of this reaction on proteins. The results of this assessment are tabulated in Table 8.2.

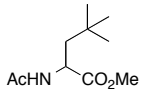
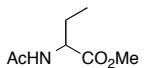
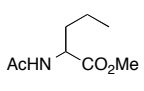
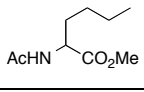
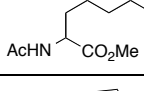
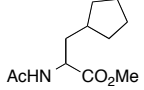
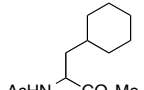
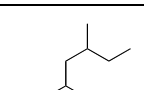
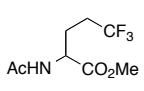
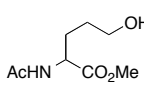
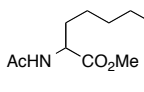
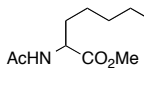
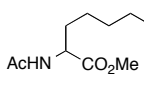
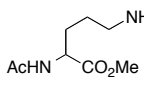
$\text{AcHN}-\text{C}(\text{CO}_2\text{Me})=\text{CH}_2 \xrightarrow[\text{NH}_4\text{OAc, H}_2\text{O}]{\text{RI, Zn powder}} \text{AcHN}-\text{CH}(\text{R})-\text{CO}_2\text{Me}$ AcDhaOMe (2.4)								
Entry	Temp.	Zn	NH ₄ OAc (5M, pH 6.0)	RI	Time	Product		Yield
1	RT	3 x 15 eq.	10 eq.	 3 x 10 eq.	30 min.		8.4	93%
2	0 °C	6 x 15 eq.	10 eq.	I-CH ₃ 6 x 10 eq.	2 h		8.5	52%
3	RT	3 x 15 eq.	10 eq.	 3 x 10 eq.	30 min.		8.6	75%
4	RT	3 x 15 eq.	10 eq.	 3 x 10 eq.	30 min.		8.7	74%
5	RT	3 x 15 eq.	10 eq.	 3 x 10 eq.	30 min.		8.8	76%
6	RT	3 x 15 eq.	10 eq.	 3 x 10 eq.	30 min.		8.9	85%
7	RT	3 x 15 eq.	10 eq.	 3 x 10 eq.	30 min.		8.10	92%
8	RT	3 x 15 eq.	10 eq.	 3 x 10 eq.	30 min.		8.11	90%
9	RT	3 x 15 eq.	10 eq.	I-CH ₂ CF ₃ 3 x 10 eq.	30 min.		8.12	74%
10	0 °C	45 eq.	10 eq.	 12 eq. over 20 min	1h		8.13	83%
11	0 °C	45 eq.	10 eq.	 10 eq. over 20 min	1h		8.14	77%
12	0 °C	45 eq.	10 eq.	 30 eq. over 30 min	1h		8.15	78% (After Boc ₂ O)
13	0 °C	45 eq.	10 eq.	 30 eq. over 30 min	1h		8.16	91% (After Boc ₂ O)
14	0 °C	45 eq.	10 eq.	 30 eq. over 30 min	1h		8.17	77% (After Boc ₂ O)

Table 8.2: Zinc-mediated addition of alkyl iodides to dehydroalanine

As shown in Table 8.2, the zinc-mediated addition of alkyl iodides to dehydroalanine enables the synthesis of a variety of unnatural and natural amino acid residues. No organic solvent was used in these reactions other than the acetic acid present in the ammonium acetate buffer. Primary, secondary, and tertiary iodides are all accommodated. Iodomethane (Entry 2) was an exception and required a larger excess to achieve even moderate yields. This result likely reflects the difficult formation of an unstable methyl radical intermediate and is consistent with the reactivity trends observed by Luche.⁸⁸ The use of iodomethane is further complicated by its low boiling point and the reaction was run at 0 °C to prevent loss of the reagent by evaporation. For alkyl iodides that are insoluble in water (Entries 1-9), the conditions optimized for 2-iodopropane were used: both zinc and the alkyl iodide were added in 3 portions over a 30 minute period at room temperature.

For alkyl iodides that are soluble in water (Entries 10-14), the reaction was run at 0 °C with slow addition of the alkyl iodide by syringe pump. If the alkyl iodide is added in one portion without cooling, even on the milligram scale employed, the large exotherm can lead to uncontrolled bumping. No overheating was observed with a controlled addition of the iodide to an ice-cold solution of dehydroalanine. The yields for Entries 10-14 paralleled that of the other primary iodides (Entries 3,4,5, and 9). Notably, this reaction was suitable for the conversion of dehydroalanine to lysine, homolysine, and ornithine (Entries 12, 13, and 14, respectively). The crude product in these cases, after workup, was protected with Boc₂O to facilitate purification. It is also fortuitous that this reaction works at low pH since the iodides used in Entries 12-14 will likely cyclize under basic conditions. Under the reaction conditions used in Table 8.2, these amines are protected as their hydroiodide salts.

The results in Tables 8.1 and 8.2 were encouraging. Natural and unnatural amino acids were synthesized from dehydroalanine in buffered aqueous media at or below room temperature. Other than reduction of the alkyl iodide, no side reactions were observed in this model system. For protein systems, the excess alkyl iodide and any reduced product should be easily removed by dialysis or by passing through a size-exclusion column. It was also suspected that the protein would be protected from non-selective alkylation because of the pH of the reaction. To test this hypothesis, the reaction was examined on a protein substrate.

Zn-mediated addition of alkyl iodides to dehydroalanine on a protein

The single cysteine mutant of subtilisin from *Bacillus lentus* (SBL-S56C) was once again used in these model studies. The reader is referred to Chapter 2 for the conversion of this single cysteine to dehydroalanine. To test the zinc-mediated addition of alkyl iodides to dehydroalanine, protein **2.35** was first prepared in 500 mM ammonium acetate buffer at pH 6.0. The buffer strength and pH were selected to approximate the conditions of the model reactions in Table 8.1 and 8.2. As a first test, a chemical mutation to leucine was examined using 2-iodopropane as the alkyl iodide. A large excess of both zinc (~10,000 eq.) and 2-iodopropane (~7000 eq.) was used in the test reaction. On the reaction scale examined (nM in protein), this corresponds to 2 mg zinc metal and 2 μ L 2-iodopropane for a 250 μ L reaction. The reaction mixture was shaken vigorously at room temperature to ensure efficient mixing. After 15 minutes, the remaining zinc and 2-iodopropane was conveniently removed by centrifugation. The supernatant was analyzed by LC-MS.

The result was encouraging. The only product observed was a protein with a mass corresponding to the leucine mutant (calculated mass = 26725; observed mass = 26728) (Figure 8.13A, Entry 6). Consumption of dehydroalanine, and thus reaction at position 156, was confirmed chemically: after dialysis to pH 8.0, no reaction was observed with thiol **2.38**. Under otherwise identical conditions, thiol **2.37** adds to dehydroalanine (Figure 8.13B). Additional control experiments revealed no reaction between SBL-156Dha (**2.35**) and Zn or 2-iodopropane alone (Figure 8.13A, Entries 2-3). Ammonium acetate is also required; no reaction was observed when this buffer was replaced with sodium phosphate at pH 8.0 or pH 6.0 (Figure 8.13A, Entries 4-5). Together these reactions support a chemical mutation of dehydroalanine to leucine that parallels the model studies in Table 8.1. This series of experiments is summarized in Figure 8.13.

It is interesting that ammonium acetate buffer is necessary. It seems that activation of the zinc metal is not accomplished by acid alone since phosphate buffer at the same pH led to no reaction. This buffer dependence is consistent with observations of Cava,¹⁰⁵ Luche^{83-85,88} and Li⁸⁹ who showed that ammonium salts activate zinc metal for reductions. Several other observations are also worth noting. First, it should be reiterated that no reaction was observed between the protein and 2-iodopropane alone. The low solubility of the alkyl iodide in water and the low pH likely slow non-selective alkylation. The conversion of dehydroalanine to Leu was repeated several times with the same result. The product protein can be easily separated from the zinc and excess iodide by centrifugation

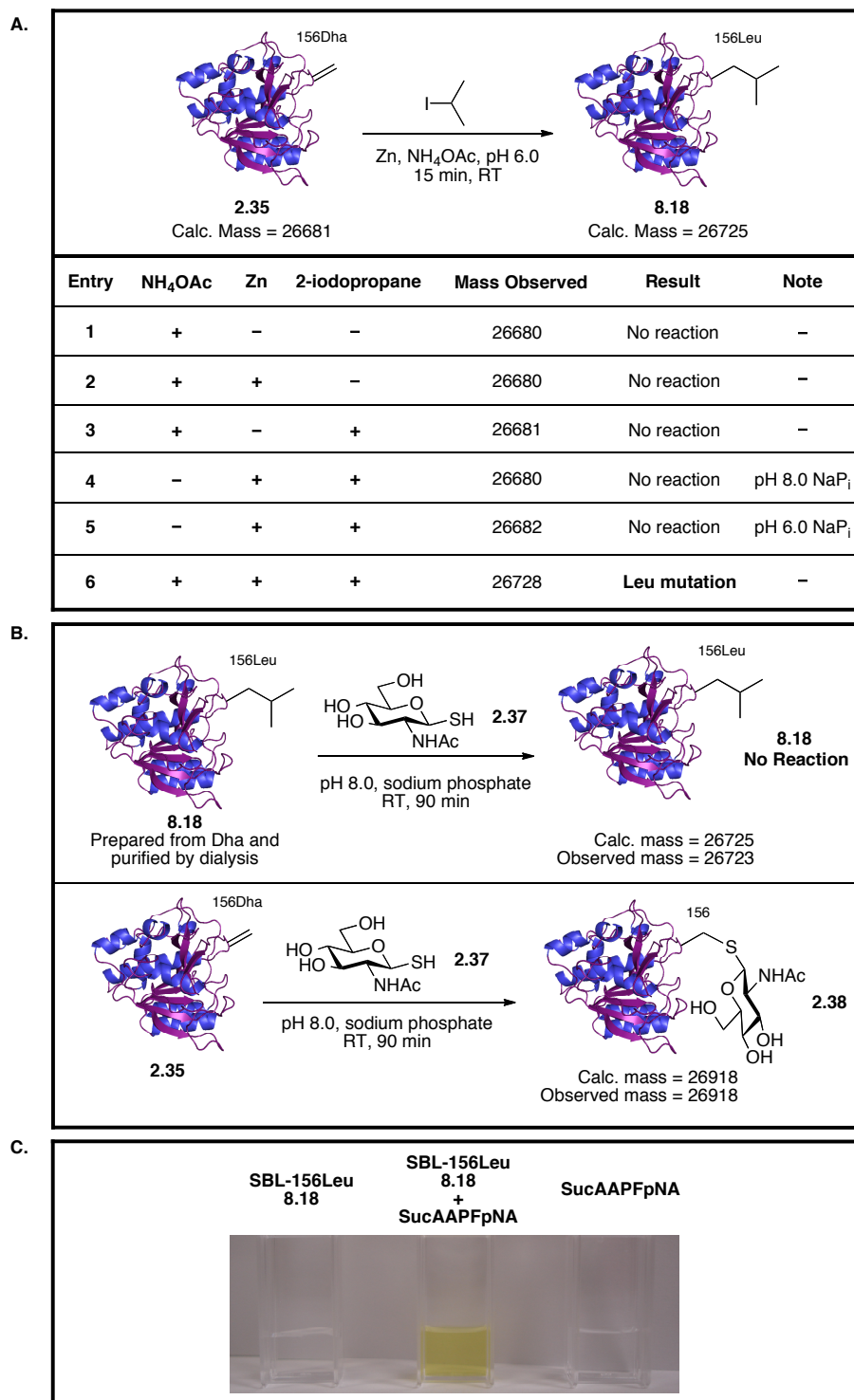


Figure 8.13: A chemical mutation from dehydroalanine to leucine and related control reactions. A) pH 6.0 Ammonium acetate, 2-iodopropane, and Zn are all required for the reaction. B) Chemical test to demonstrate reaction occurred at dehydroalanine. C) Activity assay of Leu mutant: adding the protease substrate sucAAPFpNA results in instantaneous color change, indicating proteolytic release of *p*-nitroanilide (pNA).

and then further purified by dialysis or by passing through a size-exclusion column. Protein recovery was typically between 70% and 90% over the chemical mutation and purification steps, as determined by concentration measurements of the purified protein. Finally the Leu mutant **8.18** was still an active protease, indicating it was not denatured during the chemical modification (Figure 8.13C).

With the promising results shown in Figure 8.13, several other iodides were examined in the zinc-mediated addition to dehydroalanine.[‡] The results are compiled in Figure 8.14.

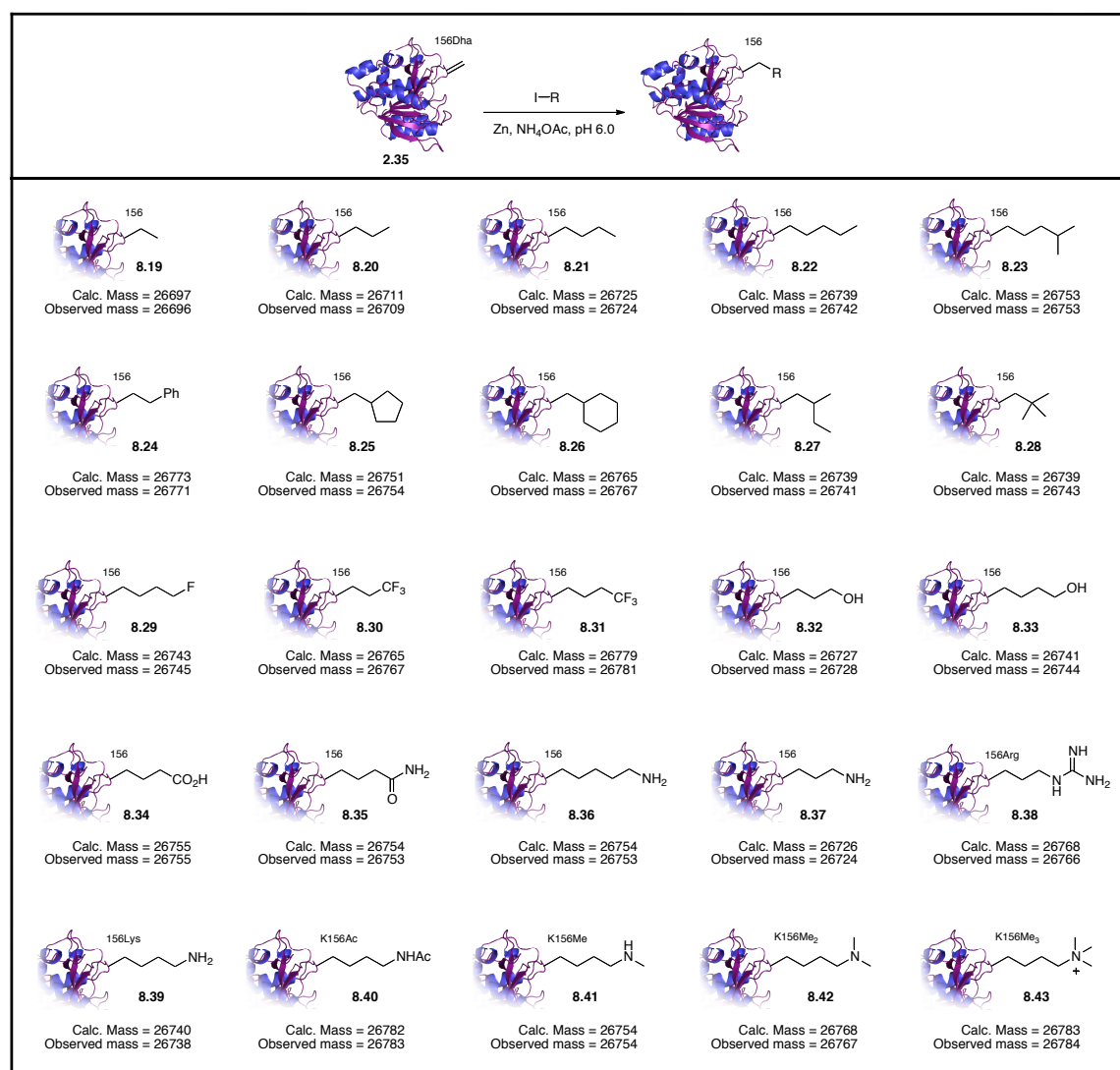


Figure 8.14: Chemical mutations of dehydroalanine to unnatural, natural, and post-translationally modified residues

[‡] The results in Figure 8.14 were obtained with the technical assistance of Dr. Gonalo Bernardes.

Reaction conditions were similar to those used above for the Dha to Leu mutation. Reaction progress was monitored by LC-MS; zinc and the corresponding iodide were added portionwise until no dehydroalanine remained. A variety of natural, unnatural, and post-translationally modified amino acids are accessible. A series of hydrocarbon side chains were installed (**8.19-8.28**) that varied in chain length and steric bulk. Fluorinated side chains were synthesized (**8.29-8.31**) and homologs of serine, glutamic acid, glutamine, and lysine (**8.32-8.37**). The natural amino acids arginine and lysine, along with acetylated and methylated lysine, were also obtained (**8.38-8.43**).

The uses envisioned for these mutations are many and varied. For instance, the non-natural alkyl side chains might be useful in the chemical remodeling of enzymes for catalysis. In the same way that directed evolution has relied on various methods of genetic mutations, it may also benefit from the chemical mutations shown in Figure 8.14. While an epimeric mixture likely results from chemical mutations of dehydroalanine, this does not preclude their use in catalysis—even asymmetric catalysis. For instance, several mutants (e.g. those in Figure 8.14) can be prepared in short order from a single protein and assayed. Any leads could be further optimized by standard genetic mutations or by pursuing a genetic route to install the unnatural amino acid identified in the initial screen. In this way, the “post-expression mutagenesis” described in this Chapter can reinforce standard genetic techniques and biological methods for unnatural amino acid incorporation.

Another application of the chemistry describe in Figure 8.14 is the supply of proteins bearing post-translational modifications. In Chapters 3 and 4, dehydroalanine was converted to sulfur-based mimics of several PTMs, including methylated and acetylated lysine residues. The chemistry described in this Chapter enables the synthesis of the *natural* methylated and acetylated lysine residues. The installation of these residues on histone proteins is the next logical step and an active pursuit in our laboratories. Furthermore, many other post-translational modifications should be accessible by the chemistry in this chapter: δ -hydroxylysine,¹⁰⁷ methylated arginine,¹⁰⁸ and citrulline¹⁰⁹ are all residues that are current targets for this chemistry due to their many biological functions, of which many are not clarified.

The chemistry depicted in Figure 8.14 is, of course, not without limitations. For instance, the natural aromatic residues are not accessible. An attempt with 4-iodophenol resulted in only unreacted starting material (Figure 8.15). Luche had also noted that aryl iodides were not reactive in this transformation, presumably because of the high energy

required to generate an aryl radical.⁸⁶ As mentioned in the introduction, other chemistry might be required to chemically access aryl side chains (cf. Figure 8.10C).^{90-99, §} Another glaring omission from Figure 8.14 is glutamic acid and glutamine. When iodoacetic acid or iodoacetamide were used in the reaction, only unreacted protein was recovered. Apparently, iodoacetic acid or iodoacetamide are reduced to the corresponding enolate and quenched by protonation faster than any radical intermediate (if formed) can add to dehydroalanine.

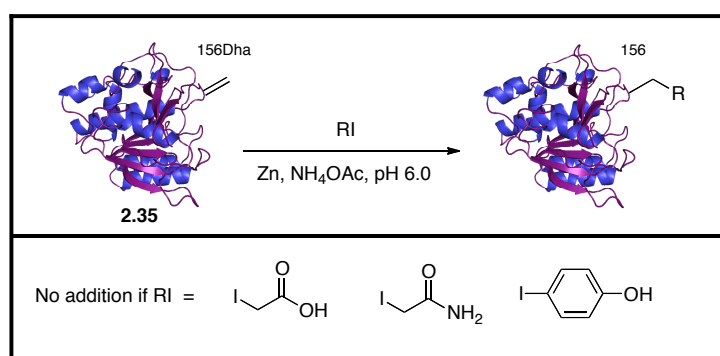


Figure 8.15: Unsuccessful chemical mutations

Finally, the results reported in this Chapter are preliminary. All reactions were carried out on a single protein, at a single site. Further exploration on other proteins is ongoing and the scope and limitations of these transformations are under investigation.[¶] Key experiments include installation of residues bearing post-translational modifications at sites that naturally bear these modifications. Histone proteins are one such target. Another unanswered question is the application of the zinc-mediated addition of alkyl iodides to a dehydroalanine residue in an enzyme active site. The predominant side-

[§] The Rh(I)-catalyzed addition of arylboronic acids to dehydroalanine is an ongoing effort to access aromatic sidechains. This reaction was briefly examined on both peptides and proteins, but the results obtained are too preliminary to present in this Thesis. The key problem encountered was the high temperature required for catalysis. No reaction was observed on protein **2.35** between room temperature and 37 °C using [Rh(I)Cl(cod)]₂ and phenylboronic acid. A comprehensive examination of Rh or Pd catalysts should resolve this problem. Alternatively, the preformed aryl-Rh species can be synthesized separately and then used as a stoichiometric reagent in protein modification.

[¶] Preliminary results suggest that hydrolysis at dehydroalanine is a competing reaction during exothermic modifications. This hydrolysis is especially pronounced in the acidic medium of the Zn mediated radical addition of alkyl iodides. Slow addition of reagents at lower temperature should be considered to minimize this side reaction.

reaction for this reaction is reduction of the radical intermediate before addition to dehydroalanine. It is not clear at present whether the radical intermediate can access a hindered active site before non-productive reduction to the alkane. If hindered sites cannot be modified, this chemistry might be limited to the surface of proteins. Additionally, the addition of imidazole to dehydroalanine, demonstrated at the beginning of this Chapter, provides access to a histidine analog. The validation of this residue in catalysis is an additional line of investigation that is currently underway.

8.3 Conclusions and Outlook

In 1966, Koshland and Bender independently reported the first point mutation of an enzyme.^{29,30} This mutation was accomplished by selective chemical reactions on the expressed and folded protein. In this Chapter, a concerted effort was made to generalize their original conception of “chemical mutagenesis”. In these efforts, reactions were developed for the conversion of dehydroalanine to natural, unnatural, and post-translationally modified residues. The use of dehydroalanine and organometallic chemistry is a unifying theme in this Thesis and their intersection in this Chapter has enabled a step forward in protein modification.

Many challenges and opportunities lie ahead in the pursuit of a general “chemical mutagenesis”. If dehydroalanine is used as a chemical precursor to other amino acids, a key breakthrough will be the analysis and control of the stereochemical outcome of any reaction—a non-trivial feat on protein substrates. Alternatively, other amino acids could be developed and incorporated into proteins as precursors to other residues—perhaps through transformations that bypass the problem of stereochemistry entirely. Another challenge is the installation of multiple modifications. Orthogonal chemistry is required for accessing such proteins. Finally, while total chemical synthesis of proteins by NCL or EPL can supply these modified proteins, the concept of chemical mutagenesis can be coordinated with these efforts. For instance, the conversion of cysteine to another residue post-NCL constitutes a formal ligation at the final residue. Desulfurization of cysteine to alanine is one widely used transformation in this context^{63,64} and perhaps some of the reactions reported in this Chapter will have some bearing on the total synthesis of proteins. Regardless of the specific method or reaction, the fundamental goal is to access modified proteins for their use in biology. No one method will resolve all supply issues,

but a combination of complementary and reinforcing strategies should facilitate selection and deployment of the most appropriate or convenient method for the problem at hand.

8.4 References

- (1) Hutchison, C. A., III; Phillips, S.; Edgell, M. H.; Gillam, S.; Jahnke, P.; Smith, M. Mutagenesis at a specific position in a DNA sequence. *J. Biol. Chem.* **1978**, *253*, 6551-6560.
- (2) Shortle, D.; DiMaio, D.; Nathans, D. Directed mutagenesis. *Ann. Rev. Genet.* **1981**, *15*, 265-294.
- (3) Smith, M. In vitro mutagenesis. *Ann. Rev. Genet.* **1985**, *19*, 423-462.
- (4) Smith, M. Site-directed mutagenesis. *Phil. Trans. R. Soc. Lond. A* **1986**, *317*, 295-304.
- (5) Xie, J.; Schultz, P. G. A chemical toolkit for proteins—an expanded genetic code. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 775-782.
- (6) Dougherty, D. A. Unnatural amino acids as probes of protein structure and function. *Curr. Opin. Chem. Biol.* **2000**, *4*, 645-652.
- (7) Link, A. J.; Mock, M. L.; Tirrell, D. A. Non-canonical amino acids in protein engineering. *Curr. Opin. Biotechnol.* **2003**, *14*, 603-609.
- (8) Ohta, A.; Yamagishi, Y.; Suga, H. Synthesis of biopolymers using genetic code reprogramming. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159-167.
- (9) Dawson, P. E.; Kent, S. B. H. Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **2000**, *69*, 923-960.
- (10) Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **2009**, *38*, 338-351.
- (11) Hackenberger, C. P. R.; Schwarzer, D. Chemoselective ligation and modification strategies for peptides and proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 10030-10074.
- (12) Sletten, E. M.; Bertozzi, C. R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974-6998.

- (13) de Graaf, A. J.; Kooijman, M.; Hennink, W. E.; Mastrobattista, E. Nonnatural amino acids for site-specific protein conjugation. *Bioconjugate Chem.* **2009**, *20*, 1281-1295.
- (14) Saghatelian, A.; Cravatt, B. F. Assignment of protein function in the postgenomic era. *Nat. Chem. Biol.* **2005**, *1*, 130-142.
- (15) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372.
- (16) Walsh, C. T. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*. Roberts and Company Publishers: Englewood, CO, 2006.
- (17) Rauh, D.; Waldmann, H. Linking chemistry and biology for the study of protein function. *Angew. Chem. Int. Ed.* **2007**, *46*, 826-829.
- (18) Gamblin, D. P.; van Kasteren, S. I.; Chalker, J. M.; Davis, B. G. Chemical approaches to mapping the function of post-translational modifications. *FEBS J.* **2008**, *275*, 1949-1959.
- (19) Flavell, R. R.; Muir, T. W. Expressed protein ligation (EPL) in the study of signal transduction, ion conduction, and chromatin biology. *Acc. Chem. Res.* **2009**, *42*, 107-116.
- (20) Kaiser, E. T.; Lawrence, D. S. Chemical mutation of enzyme active sites. *Science* **1984**, *226*, 505-511.
- (21) Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M. D. Generation of new enzymes via covalent modification of existing proteins. *Chem. Rev.* **2001**, *101*, 3081-3111.
- (22) Merrifield, B. Solid phase synthesis. *Science* **1986**, *232*, 341-347.
- (23) Kent, S. B. H. Chemical synthesis of peptides and proteins. *Annu. Rev. Biochem.* **1988**, *57*, 957-989.
- (24) Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. H. Synthesis of proteins by native chemical ligation. *Science* **1994**, *266*, 776-779.
- (25) Muir, T. W.; Sondhi, D.; Cole, P. A. Expressed protein ligation: a general method for protein engineering. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 6705-6710.
- (26) Evans, T. C.; Benner, J.; Xu, M.-Q. Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **1998**, *7*, 2256-2264.
- (27) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91-118.

- (28) Means, G. E.; Feeney, R. E. Chemical modifications of proteins: history and applications. *Bioconjugate Chem.* **1990**, *1*, 2-12.
- (29) Neet, K. E.; Koshland, D. E., Jr. The conversion of serine at the active site of subtilisin to cysteine: a "chemical mutation". *Proc. Natl. Acad. Sci. USA* **1966**, *56*, 1606-1611.
- (30) Polgár, L.; Bender, M. L. A new enzyme containing a synthetically formed active site. Thiol-subtilisin. *J. Am. Chem. Soc.* **1966**, *88*, 3153-3154.
- (31) Polgár, L.; Bender, M. L. In *Advances in Enzymology and Related Areas of Molecular Biology*; Nord, F. F., Ed.; Interscience Publishers: New York, 1970; pp 381-400.
- (32) Singer, B.; Kusmierek, J. T. Chemical mutagenesis. *Ann. Rev. Biochem.* **1982**, *52*, 655-693.
- (33) Sealock, R. W.; Laskowski, M., Jr. Enzymic replacement of the arginyl by a lysyl residue in the reactive site of soybean trypsin inhibitor. *Biochemistry* **1969**, *8*, 3703-3710.
- (34) Jering, H.; Tschesche, H. Replacement of lysine by arginine, phenylalanine and tryptophan in the reactive site of the bovine trypsin-kallikrein inhibitor (Kunitz) and change of the inhibitory properties. *Eur. J. Biochem.* **1976**, *61*, 453-463.
- (35) Wenzel, H. R.; Tschesche, H. "Chemical mutation" by amino acid exchange in the reactive site of a proteinase inhibitor and alteration of its inhibitor specificity. *Angew. Chem. Int. Ed. Engl.* **1981**, *20*, 295-296.
- (36) Clark, P. I.; Lowe, G. Chemical mutations of papain. The preparation of Ser 25- and Gly 25-papain. *J. Chem. Soc., Chem. Commun.* **1977**, 923-924.
- (37) Clark, P. I.; Lowe, G. Conversion of the active-site cysteine residue of papain into a dehydro-serine, a serine and a glycine residue. *Eur. J. Biochem.* **1978**, *84*, 293-299.
- (38) Venkatesh, Y. P.; Vithayathil, P. J. Influence of deamidation(s) in the 67-74 region of ribonuclease on its refolding. *Int. J. Pept. Protein Res.* **1985**, *25*, 27-32.
- (39) Kuroki, R.; Yamada, H.; Moriyama, T.; Imoto, T. Chemical mutations of the catalytic carboxyl groups in lysozyme to the corresponding amides. *J. Biol. Chem.* **1986**, *261*, 13571-13574.
- (40) Smith, H. B.; Hartman, F. C. Restoration of activity to catalytically deficient mutants of ribulosebisphosphate carboxylase/oxygenase by aminoethylation. *J. Biol. Chem.* **1988**, *263*, 4921-4925.

- (41) Smith, H. B.; Larimer, F. W.; Hartman, F. C. Subtle alteration of the active site of ribulose biphosphate carboxylase/oxygenase by concerted site-directed mutagenesis and chemical modification. *Biochem. Biophys. Res. Commun.* **1988**, *152*, 579-584.
- (42) Yoshimura, T.; Matsushima, Y.; Tanizawa, K.; Sung, M.-H.; Yamauchi, T.; Wakayama, M.; Esaki, N.; Soda, K. Substitution of *S*-(β -aminoethyl)-cysteine for active-site lysine of thermostable aspartate aminotransferase. *J. Biochem.* **1990**, *108*, 699-700.
- (43) Zheng, R.; Dam, T. K.; Brewer, C. F.; Blanchard, J. S. Active site residues in *Mycobacterium tuberculosis* pantothenate synthetase required in the formation and stabilization of the adenylate intermediate. *Biochemistry* **2004**, *43*, 7171-7178.
- (44) Planas, A.; Kirsch, J. F. Reengineering the catalytic lysine of aspartate aminotransferase by chemical elaboration of a genetically introduced cysteine. *Biochemistry* **1991**, *30*, 8268-8276.
- (45) Dhalla, A. M.; Li, B.; Alibhai, M. F.; Yost, K. J.; Hemmingsen, J. M.; Atkins, W. M.; Schineller, J.; Villafranca, J. J. Regeneration of catalytic activity of glutamine synthetase mutants by chemical activation: exploration of the role of arginines 339 and 359 in activity. *Protein Sci.* **1994**, *3*, 476-481.
- (46) Lindley, H. A new synthetic substrate for trypsin and its application to the determination of the amino-acid aequence of proteins. *Nature* **1956**, *178*, 647-648.
- (47) Tietze, F.; Gladner, J. A.; Folk, J. E. Release of C-terminal *S*-(β -aminoethyl)-cysteine residues by carboxypeptidase-B. *Biochim. Biophys. Acta* **1957**, *26*, 659.
- (48) Raftery, M. A.; Cole, R. D. Tryptic cleavage at cysteinyl peptide bonds. *Biochem. Biophys. Res. Commun.* **1963**, *10*, 467-472.
- (49) Wu, Z.-P.; Hilvert, D. Conversion of a protease into an acyl transferase: selenolsubtilisin. *J. Am. Chem. Soc.* **1989**, *111*, 4513-4514.
- (50) Stadtman, T. C. Selenocysteine. *Annu. Rev. Biochem.* **1996**, *65*, 83-100.
- (51) Liu, J.-Q.; Jiang, M.-S.; Luo, G.-M.; Yan, G.-L.; Shen, J.-C. Conversion of trypsin into a selenium-containing enzyme by using chemical mutation. *Biotechnol. Lett.* **1998**, *20*, 693-696.
- (52) Lian, G.; Ding, L.; Chen, M.; Liu, L.; Zhao, D.; Ni, J. A selenium-containing catalytic antibody with type I deiodinase activity. *Biochem. Biophys. Res. Commun.* **2001**, *283*, 1007-1012.

- (53) Mao, S.; Dong, Z.; Liu, J.; Li, X.; Liu, X.; Luo, G.; Shen, J. Semisynthetic tellurosubtilisin with glutathione peroxidase activity. *J. Am. Chem. Soc.* **2005**, *127*, 11588-11589.
- (54) Hahn, M. E.; Muir, T. W. Photocontrol of Smad2, a multiphosphorylated cell-signaling protein, through caging of activating phosphoserines. *Angew. Chem. Int. Ed.* **2004**, *43*, 5800-5803.
- (55) Wu, N.; Deiters, A.; Cropp, T. A.; King, D.; Schultz, P. G. A genetically encoded photocaged amino acid. *J. Am. Chem. Soc.* **2004**, *126*, 14306-14307.
- (56) Deiters, A.; Groff, D.; Ryu, Y.; Xie, J.; Schultz, P. G. A genetically encoded photocaged tyrosine. *Angew. Chem. Int. Ed.* **2006**, *45*, 2728-2731.
- (57) Mukai, T.; Kobayashi, T.; Hino, N.; Yanagisawa, T.; Sakamoto, K.; Yokoyama, S. Adding L-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem. Biophys. Res. Commun.* **2008**, *371*, 818-822.
- (58) Nguyen, D. P.; Garcia Alai, M. M.; Kapadnis, P. B.; Neumann, H.; Chin, J. W. Genetically encoding *N*-methyl-L-lysine in recombinant histones. *J. Am. Chem. Soc.* **2009**, *131*, 14194-14195.
- (59) Groff, D.; Chen, P. R.; Peters, F. B.; Schultz, P. G. A genetically encoded ϵ -*N*-methyl lysine in mammalian cells. *ChemBioChem* **2010**, *11*, 1066-1068.
- (60) Ai, H.-w.; Lee, J. W.; Schultz, P. G. A method to site-specifically introduce methyllysine into proteins in *E. coli*. *Chem. Commun.* **2010**, *46*, 5506-5508.
- (61) Nguyen, D. P.; Garcia Alai, M. M.; Virdee, S.; Chin, J. W. Genetically directing ϵ -*N,N*-dimethyl-L-lysine in recombinant histones. *Chem. Biol.* **2010**, *17*, 1072-1076.
- (62) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. A genetically encoded boronate-containing amino acid. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220-8223.
- (63) Yan, L. Z.; Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **2001**, *123*, 526-533.
- (64) Wan, Q.; Danishefsky, S. J. Free-radical-based, specific desulfurization of cysteine: a powerful advance in the synthesis of polypeptides and glycopolypeptides. *Angew. Chem. Int. Ed.* **2007**, *46*, 9248-9252.

- (65) Crich, D.; Banerjee, A. Native chemical ligation at phenylalanine. *J. Am. Chem. Soc.* **2007**, *129*, 10064-10065.
- (66) Chen, J.; Wan, Q.; Yuan, Y.; Zhu, J.; Danishefsky, S. J. Native chemical ligation at valine: a contribution to peptide and glycopeptide synthesis. *Angew. Chem. Int. Ed.* **2008**, *47*, 8521-8524.
- (67) Haase, C.; Rohde, H.; Seitz, O. Native chemical ligation at valine. *Angew. Chem. Int. Ed.* **2008**, *47*, 6807-6810.
- (68) Kumar, K. S. A.; Haj-Yahya, M.; Olschewski, D.; Lashuel, H. A.; Brik, A. Highly efficient and chemoselective peptide ubiquitylation. *Angew. Chem. Int. Ed.* **2009**, *48*, 8090-8094.
- (69) Yang, R.; Kumar Pasunooti, K.; Li, F.; Liu, X.-W.; Liu, C.-F. Dual native chemical ligation at lysine. *J. Am. Chem. Soc.* **2009**, *131*, 13592-13593.
- (70) Harpaz, Z.; Siman, P.; Kumar, K. S. A.; Brik, A. Protein synthesis assisted by native chemical ligation at leucine. *ChemBioChem* **2010**, *11*, 1232-1235.
- (71) El Oualid, F.; Merckx, R.; Ekkebus, R.; Hameed, D. S.; Smit, J. J.; de Jong, A.; Hilkmann, H.; Sixma, T. K.; Ovaa, H. Chemical synthesis of ubiquitin, ubiquitin-based probes, and diubiquitin. *Angew. Chem. Int. Ed.* **2010**, *49*, 10149-10153.
- (72) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: versatile and switchable access to functionalized proteins. *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053.
- (73) Fernández-González, M.; Boutureira, O.; Bernardes, G. J. L.; Chalker, J. M.; Young, M. A.; Errey, J. C.; Davis, B. G. Site-selective chemoenzymatic construction of synthetic glycoproteins using endoglycosidases. *Chem. Sci.* **2010**, *1*, 709-715.
- (74) Guo, J.; Wang, J.; Lee, J. S.; Schultz, P. G. Site-specific incorporation of methyl- and acetyl-lysine analogues into recombinant proteins. *Angew. Chem. Int. Ed.* **2008**, *47*, 6399-6401.
- (75) Davis, N. J.; Flitsch, S. L. A novel method for the specific glycosylation of proteins. *Tetrahedron Lett.* **1991**, *32*, 6793-6796.
- (76) Macmillan, D.; Bill, R. M.; Sage, K. A.; Fern, D.; Flitsch, S. L. Selective in vitro glycosylation of recombinant proteins: semi-synthesis of novel homogeneous glycoforms of human erythropoietin. *Chem. Biol.* **2001**, *8*, 133-145.

- (77) Serwa, R.; Wilkening, I.; Del Signore, G.; Mühlberg, M.; Claußnitzer, I.; Weise, C.; Gerrits, M.; Hackenberger, C. P. R. Chemoselective Staudinger-phosphite reaction of azides for the phosphorylation of proteins. *Angew. Chem. Int. Ed.* **2009**, *48*, 8234-8239.
- (78) van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G. Expanding the diversity of chemical protein modification allows post-translational mimicry. *Nature* **2007**, *446*, 1105-1109.
- (79) Simon, M. D.; Chu, F.; Racki, L. R.; de la Cruz, C. C.; Burlingame, A. L.; Panning, B.; Narlikar, G. J.; Shokat, K. M. The site-specific installation of methyl-lysine analogs into recombinant histones. *Cell* **2007**, *128*, 1003-1012.
- (80) Huang, R.; Holbert, M. A.; Tarrant, M. K.; Curtet, S.; Colquhoun, D. R.; Dancy, B. M.; Dancy, B. C.; Hwang, Y.; Tang, Y.; Meeth, K.; Marmorstein, R.; Cole, R. N.; Khochbin, S.; Cole, P. A. Site-specific introduction of an acetyl-lysine mimic into peptides and proteins by cysteine alkylation. *J. Am. Chem. Soc.* **2010**, *132*, 9986-9987.
- (81) Strumeyer, D. H.; White, W. N.; Koshland, D. E., Jr. Role of serine in chymotrypsin action. Conversion of the active serine to dehydroalanine. *Proc. Natl. Acad. Sci. USA* **1963**, *50*, 931-935.
- (82) Weiner, H.; White, W. N.; Hoare, D. G.; Koshland, D. E., Jr. The formation of anhydrochymotrypsin by removing the elements of water from the serine at the active site. *J. Am. Chem. Soc.* **1966**, *88*, 3851-3859.
- (83) Pétrier, C.; Luche, J.-L. Allylzinc reagent additions in aqueous media. *J. Org. Chem.* **1985**, *50*, 910-912.
- (84) Petrier, C.; Einhorn, J.; Luche, J. L. Selective tin and zinc mediated allylations. *Tetrahedron Lett.* **1985**, *26*, 1449-1452.
- (85) Petrier, C.; Dupuy, C.; Luche, J. L. Conjugate additions to α,β -unsaturated carbonyl compounds in aqueous media. *Tetrahedron Lett.* **1986**, *27*, 3149-3152.
- (86) Luche, J. L.; Allavena, C. Ultrasound in organic synthesis 16. Optimisation of the conjugate additions to α,β -unsaturated carbonyl compounds in aqueous media. *Tetrahedron Lett.* **1988**, *29*, 5369-5372.
- (87) Luche, J. L.; Allavena, C.; Petrier, C.; Dupuy, C. Ultrasound in organic synthesis 17. Mechanistic aspects of the conjugate additions to α -enones in aqueous media. *Tetrahedron Lett.* **1988**, *29*, 5373-5374.

- (88) Dupuy, C.; Petrier, C.; Sarandeses, L. A.; Luche, J. L. Ultrasound in organic synthesis. 19. Further studies on the conjugate additions to electron deficient olefins in aqueous media. *Synth. Commun.* **1991**, *21*, 643-651.
- (89) Huang, T.; Keh, C. C. K.; Li, C.-J. Synthesis of α -amino acid derivatives and amines *via* activation of simple alkyl halides by zinc in water. *Chem. Commun.* **2002**, 2440-2441.
- (90) Hayashi, T.; Yamasaki, K. Rhodium-catalyzed asymmetric 1,4-addition and its related asymmetric reactions. *Chem. Rev.* **2003**, *103*, 2829-2844.
- (91) Chapman, C. J.; Matsuno, A.; Frost, C. G.; Willis, M. C. Site-selective modification of peptides using rhodium and palladium catalysis: complementary electrophilic and nucleophilic arylation. *Chem. Commun.* **2007**, 3903-3905.
- (92) Chapman, C. J.; Hargrave, J. D.; Bish, G.; Frost, C. G. Peptide modification through site-selective residue interconversion: application of the rhodium-catalysed 1,4-addition of aryl siloxanes and boronates. *Tetrahedron* **2008**, *64*, 9528-9539.
- (93) Reetz, M. T.; Moulin, D.; Gosberg, A. BINOL-based diphosphonites as ligands in the asymmetric Rh-catalyzed conjugate addition of arylboronic acids. *Org. Lett.* **2001**, *3*, 4083-4085.
- (94) Chapman, C. J.; Frost, C. G. Synthesis of functionalised phenylalanines using rhodium catalysis in water. *Adv. Synth. Catal.* **2003**, *345*, 353-355.
- (95) Navarre, L.; Darses, S.; Genet, J.-P. Tandem 1,4-addition/enantioselective protonation catalyzed by rhodium complexes: efficient acces to α -amino acids. *Angew. Chem. Int. Ed.* **2004**, *43*, 719-723.
- (96) Navarre, L.; Darses, S.; Genet, J.-P. Efficient access to alanine derivatives by 1,4-additions of potassium trifluoro(organo)borates. *Eur. J. Org. Chem.* **2004**, 69-73.
- (97) Navarre, L.; Martinez, R.; Genet, J.-P.; Darses, S. Access to enantioenriched α -amino esters via rhodium-catalyzed 1,4-addition/enantioselective protonation. *J. Am. Chem. Soc.* **2008**, *130*, 6159-6169.
- (98) Ray, D.; Nyong, A. M.; Natarajan, A. Synthesis of unnatural amino acid derivatives via palladium-catalyzed 1,4-addition of boronic acids. *Tetrahedron Lett.* **2010**, *51*, 2655-2656.
- (99) Hargrave, J. D.; Bish, G.; Kociok Köhn, G.; Frost, C. G. Rhodium-catalysed conjugate addition of arylboronic acids to enantiopure dehydroamino acid derivatives. *Org. Biomol. Chem.* **2010**, *8*, 5120-5125.

- (100) Pérez, M.; Pleixats, R. FeCl₃-catalyzed conjugate addition of secondary amines, imidazole and pyrazole to methyl 2-acetamidoacrylate. Preparation of β -dialkylamino- α -alanine and β -(*N*-heteroaryl)- α -alanine derivatives. *Tetrahedron* **1995**, *51*, 8355-8362.
- (101) Ferreira, P. M. T.; Maia, H. L. S.; Monteiro, L. S.; Sacramento, J.; Sebastião, J. Synthesis of β -substituted alanines *via* Michael addition of nucleophiles to dehydroalanine derivatives. *J. Chem. Soc., Perkin Trans. I* **2000**, 3317-3324.
- (102) Ferreira, P. M. T.; Maia, H. L. S.; Monteiro, L. S.; Sacramento, J. Michael addition of thiols, carbon nucleophiles and amines to dehydroamino acid and dehydropolypeptide derivatives. *J. Chem. Soc., Perkin Trans. I* **2001**, 3167-3173.
- (103) Genet, J.-P.; Blart, E.; Savignac, M. Palladium-catalyzed cross-coupling reactions in a homogeneous aqueous medium. *Synlett* **1992**, 715-717.
- (104) Ager, D. J.; Lefort, L.; de Vries, J. G. In *Asymmetric synthesis and application of α -amino acids*; Soloshonok, V. A., Izawa, K., Eds., 2009; Vol. 1009; pp 239-250.
- (105) Kerdesky, F. A. J.; Ardecky, R. J.; Lakshmikantham, M. V.; Cava, M. P. Simple *o*-quinodimethane route to ($\pm$)-4-demethoxydaunomycinone. *J. Am. Chem. Soc.* **1981**, *103*, 1992-1996.
- (106) Russell, G. A. Free radical chain reactions involving alkyl- and alkenylmercurials. *Acc. Chem. Res.* **1989**, *22*, 1-8.
- (107) Shoulders, M. D.; Raines, R. T. Collagen structure and stability. *Annu. Rev. Biochem.* **2009**, *78*, 929-958.
- (108) Bedford, M. T.; Richard, S. Arginine methylation: an emerging regulator of protein function. *Mol. Cell* **2005**, *18*, 263-272.
- (109) Thompson, P. R.; Fast, W. Histone citrullination by protein arginine deiminase: is arginine methylation a green light or a roadblock? *ACS Chem. Biol.* **2006**, *1*, 433-441.

8.5 Experimental Procedures

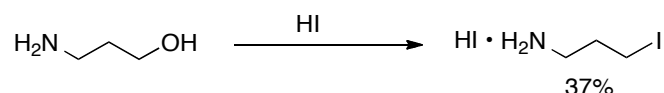
General Considerations

See Chapter 2 Experimental Procedures for general considerations in synthesis and protein analysis.

Synthesis of alkyl iodides used in Table 8.2 and Figure 8.14

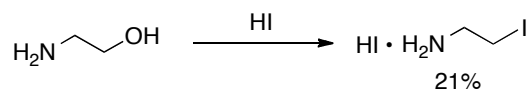
(Alkyl iodides not shown here are commercially available)

(3-iodopropyl)amine hydroiodide



Hydriodic acid (55%, 13.0 mL) was added to a 100 mL round bottom flask and cooled to 0 °C. To the stirred solution was added 3-amino-1-propanol (2.0 mL, 26.6 mmol). The ice bath was removed after completion of the addition and the reaction heated to 125 °C (bath temp) and then stirred (open air) for 2 hours. The reaction mixture, now approximately half the original volume was cooled to room temperature. The reaction mixture solidified. The dark red solid was recrystallized twice (EtOH/CHCl₃) to give the titled compound as clear prisms (3.06 g, 37%). m.p. = 163-164 °C. IR (ν_{max} , KBr): 2991, 2288, 1860, 1582, 1588, 1309, 1194, 1120, 1059. ¹H NMR (400 MHz, MeOD): δ = 2.20 (2H, quin, J = 7.3, CH₂CH₂I), 3.09 (2H, t, J = 7.3, CH₂NH₂), 3.33 (2H, t, J = 7.3, CH₂I). ¹³C NMR (100 MHz, MeOD): δ = 0.1 (CH₂I), 31.2 (CH₂CH₂I), 40.6 (CH₂NH₂). HRMS m/z (ESI⁺): Found 185.9774 [M+H]⁺. C₃H₉NI requires 185.9775.

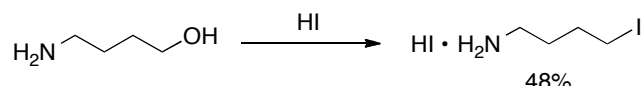
(2-iodoethyl)amine hydroiodide



Ethanolamine (1.00 g, 16.37 mmol) was added to a 50 mL round bottom flask. Hydriodic acid (55%, 8 mL) was added slowly to the flask with stirring. The mixture was then placed in an oil bath preheated to 125 °C. The reaction was heated (open air) for 2 hours and then cooled to room temperature and then on ice. The product crystallized from the reaction mixture. The crystals were isolated by filtration and then recrystallized (CHCl₃/EtOH) to give white flakes (1.02 g, 21%). m.p. = 189-190 °C. IR (ν_{max} , KBr):

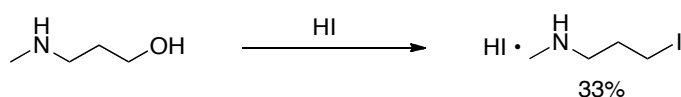
2995, 1552, 1488, 1435, 1378, 1314, 1206, 1057, 1014, 910, 860. ^1H NMR (400 MHz, MeOD): δ = 3.37-3.46 (4H, m, CH_2CH_2). ^{13}C NMR (100 MHz, MeOD): δ = -3.21 (CH_2I), 42.3 (CH_2NH_2). HRMS m/z (ESI^+): Found 171.9615 $[\text{M}+\text{H}]^+$; $\text{C}_2\text{H}_7\text{NI}$ requires 171.9618.

(4-iodobutyl)amine hydroiodide



4-amino-butanol (2.00 g, 22.44 mmol) was added to a 50 mL round bottom flask. Hydriodic acid (55%, 11 mL) was added slowly to the flask with stirring. The mixture was then placed in an oil bath preheated to 125 °C. The reaction was heated (open air) for 2 hours and then cooled to room temperature. The product crystallized from the reaction mixture. The product was isolated by filtration and then recrystallized ($\text{CHCl}_3/\text{EtOH}$). The crystals (pale yellow flakes) were isolated by filtration and triturated with CHCl_3 to remove remaining iodine. The product was dried under vacuum. (3.55 g, 48%). m.p. = 97-99 °C. IR (ν_{max} , KBr): 2969, 1950, 1472, 1387, 1281, 1175, 1034, 952, 849. 761, 719. ^1H NMR (400 MHz, MeOD): δ = 1.78-1.85 (2H, m, $\text{CH}_2\text{CH}_2\text{NH}_2$), 1.89-1.97 (2H, m, $\text{CH}_2\text{CH}_2\text{I}$), 3.01 (2H, t, J = 7.6, CH_2NH_2), 3.31 (2H, t, J = 6.7, CH_2I). ^{13}C NMR (100 MHz, MeOD): δ = 4.8 (CH_2I), 28.6 ($\text{CH}_2\text{CH}_2\text{NH}_2$), 30.4 ($\text{CH}_2\text{CH}_2\text{I}$), 38.9 (CH_2NH_2). HRMS m/z (ESI^+): Found 199.9928 $[\text{M}+\text{H}]^+$; $\text{C}_4\text{H}_{11}\text{NI}$ requires 199.9931.

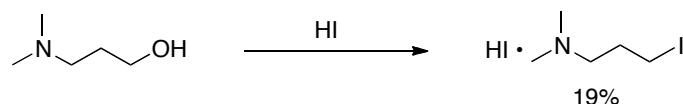
(3-iodopropyl)methylamine hydroiodide



3-methylamino-1-propanol (2.00 g, 22.44 mmol) was added to a 50 mL round bottom flask. Hydriodic acid (55%, 12 mL) was added slowly to the stirred alcohol. The reaction mixture was heated at 125 °C (open air) for 2 hours and then cooled to room temperature. The solution was then stored at -20 °C for several hours during which time the product crystallized from the reaction mixture. The crystals were isolated by filtration and washed with Et_2O (100 mL). The product was then recrystallized ($\text{CHCl}_3/\text{EtOH}$) to give yellow needles that were isolated by filtration, washed with Et_2O (100 mL), and dried under vacuum. (2.41 g, 33%). m.p. = 53-55 °C. IR (ν_{max} , KBr): 3480, 2935, 2760, 2398, 1457, 1416, 1268, 1203, 1038. ^1H NMR (400 MHz, MeOD): δ = 2.24 (2H, m, $\text{CH}_2\text{CH}_2\text{I}$), 2.76

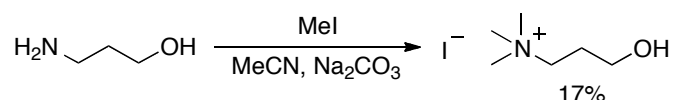
(3H, s, NHCH_3), 3.16 (2H, t, $J = 7.8$, CH_2NHMe), 3.33 (2H, t, $J = 6.8$, CH_2I). ^{13}C NMR (100 MHz, MeOD): $\delta = -0.40$ (CH_2I), 30.0 ($\text{CH}_2\text{CH}_2\text{I}$), 33.1 (NHCH_3), 50.1 (CH_2NHCH_3). HRMS m/z (ESI^+): Found 199.9928 $[\text{M}+\text{H}]^+$; $\text{C}_4\text{H}_{11}\text{NI}$ requires 199.9930.

(3-iodopropyl)dimethylamine hydroiodide



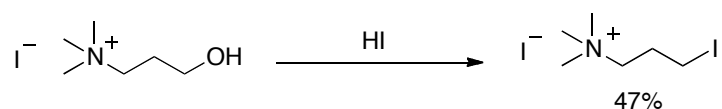
3-dimethylamino-1-propanol (2.00 g, 19.39 mmol) was added to a 25 mL round bottom flask. Hydriodic acid (55%, 12 mL) was added slowly to the stirred alcohol. The reaction mixture was heated at 125 °C (open air) for 2 hours and then cooled to room temperature. The product crystallized from the reaction mixture. The crystals were isolated by filtration and then recrystallized from CHCl_3 . The white crystals were isolated by filtration and triturated with Et_2O to remove remaining iodine before drying under vacuum. (1.24 g, 19%). m.p. = 110-111 °C. IR (ν_{max} , KBr): 3010, 2921, 2690, 1470, 1314, 1189, 1101, 1036, 1004, 960. ^1H NMR (400 MHz, MeOD): $\delta = 2.29$ (2H, m, $\text{CH}_2\text{CH}_2\text{I}$), 2.95 (6H, s, NMe_2), 3.28-3.33 (4H, m, contains CH_2NMe_2 and CH_2I). ^{13}C NMR (100 MHz, MeOD): $\delta = -1.03$ (CH_2I), 28.4 ($\text{CH}_2\text{CH}_2\text{I}$), 42.9 (NMe_2), 58.6 (CH_2NMe_2). HRMS m/z (ESI^+): Found 214.0084 $[\text{M}+\text{H}]^+$; $\text{C}_5\text{H}_{13}\text{NI}$ requires 214.0087.

(3-hydroxypropyl)-trimethylammonium iodide



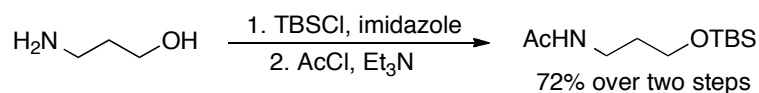
3-amino-1-propanol (1.00 mL, 13.07 mmol) was added to a 100 mL round bottom flask and dissolved in MeCN (50 mL). Na_2CO_3 (6.93 g, 65.4 mmol) was added and the mixture stirred vigorously. MeI (4.90 mL, 78.42 mmol) was added dropwise at room temperature. After stirring (open air) at room temperature for 1 hour, the solids were filtered and rinsed with 50 mL MeCN. The filtrate was evaporated to give a solid residue. The solid was recrystallized ($\text{EtOH}/\text{CHCl}_3$) to give the titled compound as clear flakes (530 mg, 17%). m.p. = 183 °C (decomposes). IR (ν_{max} , KBr): 3386, 3008, 1477, 1402, 1313, 1249, 1132, 1067, 959, 909. ^1H NMR (400 MHz, D_2O): $\delta = 1.95$ (2H, m, $\text{CH}_2\text{CH}_2\text{OH}$), 3.05 (9H, s, NMe_3), 3.34 (2H, m, CH_2NMe_3), 3.60 (2H, t, $J = 5.9$, CH_2OH). ^{13}C NMR (100 MHz, D_2O): $\delta = 25.7$ ($\text{CH}_2\text{CH}_2\text{OH}$), 53.5 (NMe_3), 58.5 (CH_2OH), 64.5 (CH_2NMe_3). HRMS m/z (ESI^+): Found 118.1226 $[\text{M}]^+$; $\text{C}_6\text{H}_{16}\text{NO}$ requires 118.1226.

(3-iodopropyl)-trimethylammonium iodide



(3-Hydroxypropyl)-trimethylammonium iodide (320 mg, 1.31 mmol) was added to a 25 mL round bottom flask and dissolved in hydriodic acid (55%, 7 mL). The solution was heated (open air) at 125 °C for 2 hours and then cooled to room temperature. Upon cooling the product crystallized from the reaction mixture. The crystals were isolated by filtration and rinsed with CHCl_3 (~20 mL). The product was then recrystallized ($\text{EtOH}/\text{CHCl}_3$) to give the product as reddish-brown needles (218 mg, 47%). m.p. = 128-129 °C. IR (ν_{max} , KBr): 3061, 1469, 1240, 1182, 1013, 961, 947. ^1H NMR (400 MHz, DMSO): δ = 2.23 (2H, m, $\text{CH}_2\text{CH}_2\text{I}$), 3.06 (9H, s, NMe_3), 3.23 (2H, t, J = 6.7, CH_2I), 3.32 (2H, m, CH_2NMe_3). ^{13}C NMR (100 MHz, DMSO): δ = 2.8 (CH_2I), 30.0 ($\text{CH}_2\text{CH}_2\text{I}$), 53.4 (NMe_3), 66.7 (CH_2NMe_3). HRMS m/z (ESI^+): Found 228.0244 $[\text{M}]^+$; $\text{C}_6\text{H}_{15}\text{IN}$ requires 228.0242.

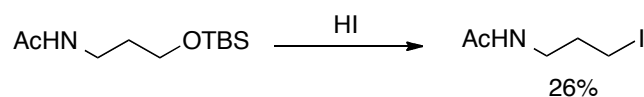
N-acetyl-3-aminopropanol-*tert*-butyldimethylsilyl ether



3-aminopropanol (3.00 g, 39.9 mmol) was added to a 500 mL round bottom flask and dissolved in 100 mL of CH_2Cl_2 . Imidazole (4.5 g, 66.5 mmol) was added and the mixture stirred at room temperature until a clear solution was obtained. TBSCl (2.34 g, 15.1 mmol) was added in one portion and the reaction flushed with nitrogen. After 1.5 hours, the reaction was quenched with NaHCO_3 (sat. aq., 200 mL) and diluted with an additional 100 mL of CH_2Cl_2 . The organic layer was separated and then washed with H_2O (3×150 mL). The organic layer was dried (Na_2SO_4), filtered, and concentrated under reduced pressure. The crude amine was sufficiently pure by NMR to carry on to the next step. Accordingly the amine was redissolved in CH_2Cl_2 (100 mL). Et_3N (7.4 mL, 53.2 mmol) was added and the stirred solution cooled to 0 °C. AcCl (1.13 mL, 16.0 mmol) was added slowly and the reaction placed under nitrogen. After 1.5 hours at 0 °C, the reaction was quenched with NaHCO_3 (sat. aq., 100 mL) and diluted with 100 mL of CH_2Cl_2 . The organic layer was separated and then washed with NaHCO_3 (sat. aq. 100 mL), dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc) to afford the desired compound as a clear

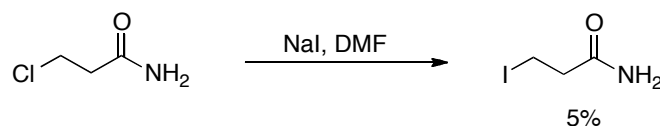
oil (2.45 g, 72% over two steps). IR ($\nu_{\max}$, film): 3291, 3088, 2931, 1654, 1558, 1472, 1370, 1295, 1256, 1219, 1100. ^1H NMR (400 MHz, CDCl_3): δ = 0.05 (6H, s, TBS), 0.88 (9H, s, TBS), 1.70 (2H, m, $\text{CH}_2\text{CH}_2\text{OTBS}$), 1.92 (3H, s, NHAc), 3.35 (2H, m, CH_2NHAc), 3.72 (2H, m, CH_2OTBS), 6.36 (1H, br. s, NH). ^{13}C NMR (100 MHz, CDCl_3): δ = -5.5 (TBS), 18.14 (TBS), 23.3 (NHAc), 25.9 (TBS), 31.3 ($\text{CH}_2\text{CH}_2\text{OTBS}$), 38.6 (CH_2NHAc), 62.7 (CH_2OTBS), 169.8 (NHAc). HRMS m/z (ESI^+): Found 254.1547 $[\text{M}+\text{Na}]^+$; $\text{C}_{11}\text{H}_{25}\text{NO}_2\text{SiNa}$ requires 254.1547.

1-iodo-3-acetamidopropane



N-acetyl-3-aminopropanol-*tert*-butyldimethylsilyl ether (200 mg, 0.86 mmol) was added to a 25 mL round bottom flask along with 5 mL of hydriodic acid (55%). The mixture was heated to 100 °C (open air) and stirred for 2 hours. After cooling to room temperature, the reaction was poured onto ~50 g ice. The product was extracted with CH_2Cl_2 (3 $\times$ 50 mL) and the combined organic layers were washed with NaHCO_3 (sat. aq., 100 mL). The organics were then dried (MgSO_4), filtered, and concentrated under reduced pressure. The product was purified by column chromatography (EtOAc) to give the titled compound as a yellow oil (50 mg, 26%). The product is unstable and is best used immediately after purification. IR ($\nu_{\max}$, film): 2966, 2934, 2362, 2344, 1780, 1700, 1386, 1239, 1202. ^1H NMR (400 MHz, CDCl_3): δ = 2.04 (3H, s, NHAc), 2.06 (2H, t, J = 6.7, $\text{CH}_2\text{CH}_2\text{I}$), 3.21 (2H, t, J = 6.7, CH_2I), 3.35 (2H, q, J = 6.7, CH_2NHAc), 6.04 (1H, br. s, NHAc). ^{13}C NMR (100 MHz, CDCl_3): δ = 3.5 (CH_2I), 23.7 (NHAc), 32.8 ($\text{CH}_2\text{CH}_2\text{I}$), 40.3 (CH_2NHAc), 171.1 (NHAc). HRMS m/z (ESI^+): Found 227.9882 $[\text{M}+\text{H}]^+$; $\text{C}_5\text{H}_{11}\text{NOI}$ requires 227.9880.

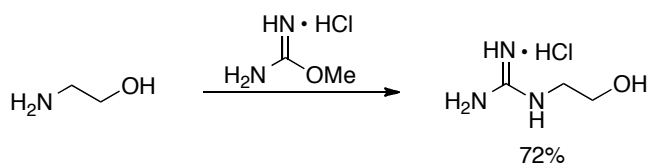
3-iodopropionamide



3-chloropropionamide (2.00 g, 18.60 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (15 mL). Sodium iodide (13.94 g, 92.99 mmol) was added to give a saturated solution. The reaction mixture was heated to 100 °C and stirred for 3 hours.

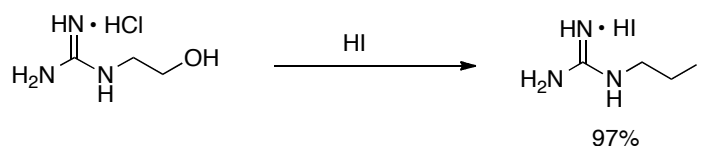
After this time no starting material was observed by ESI-MS. The reaction was cooled to room temperature and then transferred to a separatory funnel and diluted with EtOAc (150 mL) and H₂O (100 mL). The organic layer was separated and then washed successively with H₂O (100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting red solid was recrystallized by first dissolving in CHCl₃/EtOH at 70 °C and then adding petrol after cooling to room temperature. The product was isolated as white crystals (143 mg, 5% yield). Note this yield is unoptimized and the product is likely lost during extraction. Spectroscopic data was consistent with that previously reported.¹ m.p. = 130-132 °C (Lit.¹ 135-136 °C). IR (ν_{max} , KBr): 1662, 1630, 1415, 1290, 1169, 1136. ¹H NMR (500 MHz, MeOD): δ = 2.84 (2H, t, J = 6.9, CH₂CH₂I), 3.37 (2H, t, J = 6.9, CH₂CH₂I). ¹³C NMR (125 MHz, MeOD): δ = -1.8 (CH₂CH₂I), 40.6 (CH₂CH₂I), 176.1 (C=O). HRMS m/z (ESI⁺): Found 221.9390 [M+Na]⁺; C₃H₆NOINa requires 221.9386.

***N*-(2-hydroxyethyl)-guanidine hydrochloride**



O-methylisourea hydrochloride (1.19 g, 10.8 mmol) was added to a 10 mL vial and dissolved in ethanolamine (1.0 mL, 16.6 mmol). The mixture was stirred for 5 minutes at room temperature and then heated to 100 °C (open air) and stirred for 3 hours. The reaction was cooled to room temperature, dissolved in methanol, transferred to a 50 mL round bottom flask, and then concentrated under reduced pressure. Excess ethanolamine was then removed by distillation (170 °C bath temp at ~3-8 mm Hg). The thick residue was then stored at -20 °C overnight and the product solidified. The white solid was recrystallized from MeOH and EtOAc to give the titled compound as white needles (1.09 g, 72%). Spectroscopic data was consistent with that previously reported.² m.p. = 90-91 °C. IR (ν_{max} , KBr): 3229, 1650, 1457, 1410, 1353, 1278, 1222, 1055, 898, 857. ¹H NMR (400 MHz, D₂O): δ = 3.22 (2H, t, J = 5.3), 3.60 (2H, t, J = 5.3). ¹³C NMR (100 MHz, D₂O): δ = 43.8, 60.2, 157.6. LRMS m/z (ESI⁺): 104.1 (M+H)⁺.

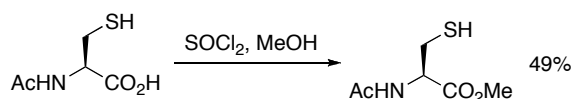
***N*-(2-iodoethyl)-guanidine hydroiodide**



N-(2-hydroxyethyl)-guanidine hydrochloride (419 mg, 3.00 mmol) was added to a 25 mL round bottom flask and dissolved in hydriodic acid (55%, 6 mL). The reaction was heated to 125 °C and stirred open air for 2 hours. The reaction was then cooled to room temperature and water was removed by evaporation. The resulting residue was dried under high vacuum and was sufficiently pure to use in additions to dehydroalanine (996 mg, 97%, very hygroscopic salt). ¹H NMR (400 MHz, D₂O): δ = 3.24 (2H, t, *J* = 6.5, CH₂I), 3.48 (2H, t, *J* = 6.5, CH₂CH₂I). ¹³C NMR (125 MHz, D₂O): δ = 2.4 (CH₂I), 43.4 (CH₂CH₂I), 156.7 (C=NH). HRMS *m/z* (ESI⁺): Found 213.9836 [M+H]⁺; C₃H₉N₃I requires 213.9833.

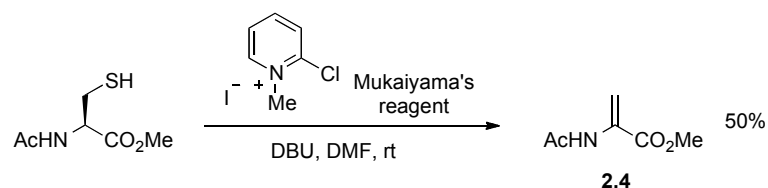
Synthesis of model dehydroalanine substrate

AcCysOMe



N-Acetyl-L-cysteine (5.00 g, 30.6 mmol) was added to a flame dried 250 mL round bottom flask under argon and dissolved in MeOH (50 mL). The stirred solution was cooled to 0 °C and thionyl chloride (2.46 mL, 33.7 mmol) was added slowly. The reaction was stirred for 10 minutes at 0 °C and then at room temperature for 1.5 hours. The solvent was then removed under reduced pressure. The resulting residue was redissolved in EtOAc (250 mL) and washed successively with H₂O (200 mL) and brine (200 mL). The organic layer was then dried (MgSO₄), filtered, and concentrated under reduced pressure to give a thick residue that crystallized on standing (2.64 g, 49%). This material was used without further purification. Spectroscopic data was consistent with that previously reported.³ *m.p.* = 51-52 °C. [α]_D²⁰ = +96.0 (*c* = 1.3, CHCl₃). IR (ν_{max}, KBr): 3281, 1743, 1656, 1541, 1437, 1374, 1217. ¹H NMR (400 MHz, CDCl₃): δ = 1.37 (1H, t, *J* = 9.0, SH), 2.03 (3H, s, NHAc), 2.93-2.97 (2H, m, CH₂SH), 3.74 (3H, s, CO₂CH₃), 4.82-4.86 (1H, m, H_α), 6.66 (1H, d, *J* = 6.3, NH). ¹³C NMR (100 MHz, CDCl₃): δ = 23.1 (NHAc), 26.8 (CH₂SH), 52.8, 53.6 (CO₂CH₃, C_α), 170.1, 170.8 (C=O). LRMS *m/z* (ESI⁻): 176 [M-H]⁻.

AcDhaOMe (2.4)

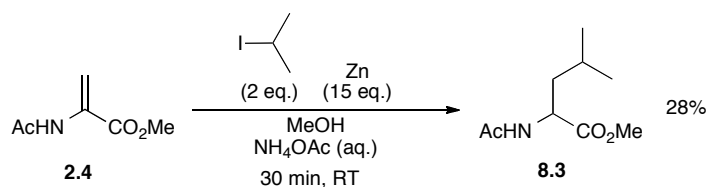


The following procedure was developed by Dr. Omar Boutureira as part of the comprehensive study of converting cysteine to dehydroalanine. While not discussed in Chapter 2, this protocol is useful for preparing relatively large quantities of AcDhaOMe.

DBU (2.50 mL, 17.0 mmol) was added dropwise to a dispersion of *N*-acetyl-L-cysteine methyl ester (1.00 g, 5.64 mmol) and 2-chloro-1-methylpyridinium iodide (1.63 g, 6.4 mmol) in dry and degassed DMF (5.6 mL) at room temperature under an argon atmosphere. After 5 minutes at the same temperature, the reaction mixture was diluted with EtOAc and washed with 1 M aqueous HCl and water. The combined organic layers were dried over MgSO₄, filtered, and concentrated under reduced pressure. The residue was purified by column chromatography (50% EtOAc in petrol) to afford AcDhaOMe (404 mg, 50%) as a white crystalline solid. Spectroscopic data was consistent with that previously reported.⁴ m.p. = 40–41 °C. IR (ν_{max} , film): 1730, 1677, 1636, 1519, 1440, 1372, 1235, 1203, 1171. ¹H NMR (500 MHz, CDCl₃): δ = 2.14 (3H, s, NHAc), 3.85 (3H, s, CO₂CH₃), 5.89 (1H, d, J = 1.6, C=CHH), 6.60 (1H, app. s, C=CHH), 7.73 (1H, br. s, NH). ¹³C NMR (125 MHz, CDCl₃): δ = 24.7 (NHAc), 53.0 (CO₂CH₃), 108.7 (C=CH₂), 130.8 (C=CH₂), 164.6, 168.8 (C=O). LRMS m/z (ESI): 142.1 [M-H][−].

Model studies on the chemical conversion of dehydroalanine to leucine (Table 8.1)

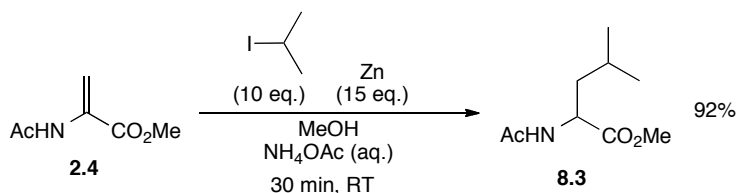
Table 8.1, Entry 1



AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and dissolved in MeOH (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added to the stirred solution followed by zinc powder (206 mg, 3.15 mmol). The mixture was cooled to 0 °C and 2-iodopropane (42 μ L, 0.42 mmol) was added. After stirring 10 minutes at 0 °C, the reaction was removed from the ice bath and stirred vigorously at room temperature for 30

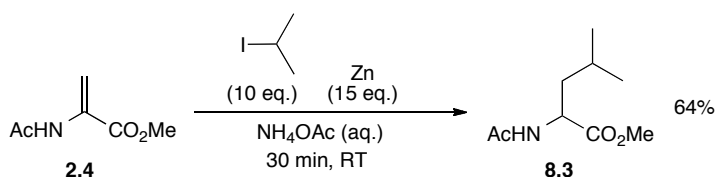
minutes. After this time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-AcLeuOMe (11 mg, 28% yield). Spectroscopic data was consistent with that previously reported.⁵ IR (ν_{max} , film): 3283, 2957, 1747, 1655, 1543, 1437, 1372, 1204, 1158. ¹H NMR (CDCl₃, 400 MHz): δ = 0.92 (3H, d, J = 6.3, CHMe), 0.93 (3H, d, J = 6.1, CHMe'), 1.49-1.55 (1H, m, H β), 1.59-1.69 (2H, m, H β ' and CHMe₂), 2.01 (3H, s, NHAc), 3.72 (3H, s, CO₂Me), 4.63 (1H, m, H α), 6.01 (1H, d, J = 7.3, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 22.0, 22.8, 23.1, 24.8 (NHAc, CHMe₂, and CHMe₂), 41.7 (C β), 50.7 (C α), 52.3 (CO₂Me), 169.8, 173.7 (2 $\times$ C=O). LRMS m/z (ESI⁺): 210.1 [M+Na]⁺.

Table 8.1, Entry 2



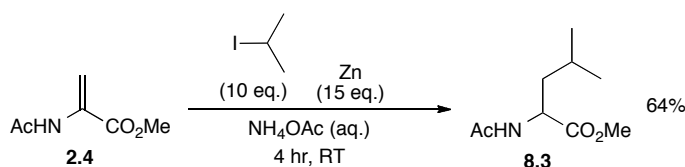
AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and dissolved in MeOH (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added to the stirred solution followed by zinc powder (206 mg, 3.15 mmol). The mixture was cooled to 0 °C and 2-iodopropane (0.21 mL) was added dropwise over 1 minute. After stirring 10 minutes at 0 °C, the reaction was removed from the ice bath and stirred vigorously at room temperature for 30 minutes. After this time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-AcLeuOMe (36 mg, 92% yield). Spectroscopic data was identical to that reported above.

Table 8.1, Entry 3



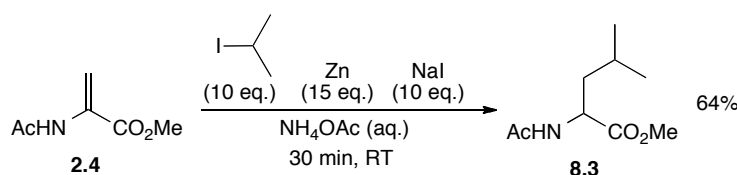
AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added and the reaction was cooled to 0 °C. 2-iodopropane (0.21 mL) was added dropwise over 1 minute. After 10 minutes at 0 °C, the reaction was removed from the ice bath and stirred at room temperature for 30 minutes. After this time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-AcLeuOMe (25 mg, 64% yield). Spectroscopic data was identical to that reported above.

Table 8.1, Entry 4



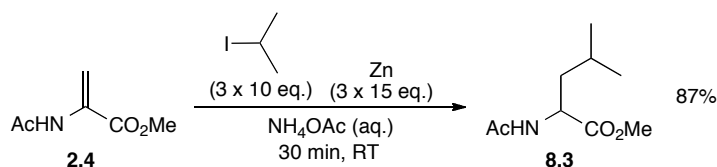
AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added and the reaction was cooled to 0 °C. 2-iodopropane (0.21 mL) was added dropwise over 1 minute. After 10 minutes at 0 °C, the reaction was removed from the ice bath and stirred at room temperature for 4 hours. After this time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-AcLeuOMe (25 mg, 64% yield). Spectroscopic data was identical to that reported above.

Table 8.1, Entry 5



AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Sodium iodide (315 mg, 2.1 mmol) and zinc powder (206 mg, 3.15 mmol) were added to the stirred solution. 2-iodopropane (0.21 mL) was then added dropwise over 1 minute at room temperature. After 30 minutes, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-AcLeuOMe (25 mg, 64% yield). Spectroscopic data was identical to that reported above.

Table 8.1, Entry 6

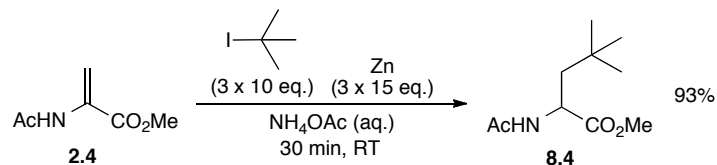


AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by 2-iodopropane (0.21 mL). The reaction was stirred at room temperature for 10 minutes and then a second dose of zinc powder (206 mg, 3.15 mmol) and 2-iodopropane (0.21 mL) was added. After 10 additional minutes, a third and final dose of zinc powder (206 mg, 3.15 mmol) and 2-iodopropane (0.21 mL) was added. After 30 minutes of total reaction time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-AcLeuOMe (34 mg, 87% yield). Spectroscopic data was identical to that reported above.

Model chemical mutations from Table 8.2

Table 8.2, Entry 1

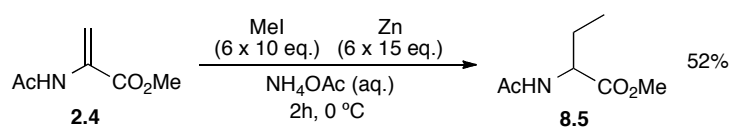
Chemical conversion of dehydroalanine to *tert*-butylalanine



AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by 2-methyl-2-iodopropane (0.25 mL, 2.1 mmol). The reaction was stirred at room temperature for 10 minutes and then a second dose of zinc powder (206 mg, 3.15 mmol) and 2-methyl-2-iodopropane (0.25 mL, 2.1 mmol) was added. After 10 additional minutes, a third and final dose of zinc powder (206 mg, 3.15 mmol) and 2-methyl-2-iodopropane (0.25 mL, 2.1 mmol) was added. After 30 minutes of total reaction time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-Ac(*t*BuAla)OMe (39 mg, 93% yield). IR (ν_{max} , film): 3285, 3072, 2955, 1748, 1657, 1548, 1470, 1436, 1370, 1243, 1175, 1124. ¹H NMR (CDCl₃, 500 MHz): δ = 0.94 (9H, s, CMe₃), 1.46 (1H, ABX system, J = 14.2, 8.5, H β), 1.74 (1H, ABX system, J = 14.2, 4.1, H β'), 1.98 (3H, s, NHAc), 3.70 (3H, s, CO₂Me), 4.64 (1H, td, J = 8.5, 4.1, H α), 5.98 (1H, d, J = 8.5, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 23.1 (NHAc), 29.5 (CMe₃), 30.7 (CMe₃), 46.1 (C β), 49.8 (C α), 52.2 (CO₂Me), 169.6, 174.1 (2 × C=O). HRMS m/z (ESI⁺): Found 224.1257 [M+Na]⁺; C₁₀H₁₉NO₃Na requires 224.1257.

Table 8.2, Entry 2

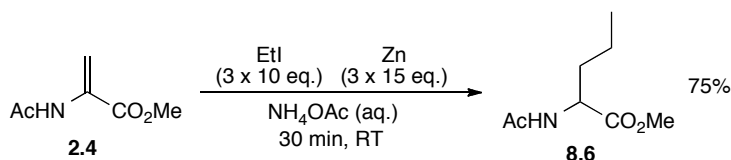
Chemical conversion of dehydroalanine to methylalanine



AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. The reaction mixture was cooled to 0 °C. Every twenty minutes, a dose of zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by methyl iodide (0.13 mL, 2.1 mmol). This portion-wise addition of zinc and methyl iodide was repeated for a total of 6 additions, all at 0 °C. After the final addition, the reaction was stirred for twenty minutes at 0 °C. After this time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-*N*-acetyl-methylalanine methyl ester (17 mg, 52% yield). IR (ν_{max} , film): 2970, 2360, 2341, 1744, 1654, 1541, 1507, 1457, 1436, 1374, 1299, 1258, 1208, 1155. ¹H NMR (CDCl₃, 400 MHz): δ = 0.90 (3H, t, *J* = 7.4, CH₂CH₃), 1.66-1.77 (1H, m, H β), 1.83-1.92 (1H, m, H β'), 2.03 (3H, s, NHAc), 3.75 (3H, s, CO₂Me), 4.56-4.61 (1H, m, H α), 6.07 (1H, br. s, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 9.4 (CH₂CH₃), 23.2 (NHAc), 25.6 (C β), 52.3 (CO₂Me), 53.2 (C α), 169.7, 173.0 (2 $\times$ C=O). HRMS *m/z* (ESI⁺): Found 182.0782 [M+Na]⁺; C₇H₁₃NO₃Na requires 182.0788.

Table 8.2, Entry 3

Chemical conversion of dehydroalanine to norvaline

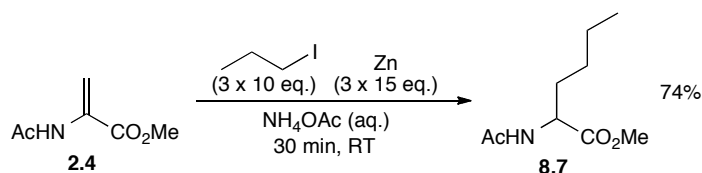


AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by ethyl iodide (0.17 mL, 2.1 mmol). The reaction was stirred at room temperature for 10 minutes and then a second dose of zinc powder (206 mg, 3.15 mmol) and ethyl iodide (0.17 mL, 2.1 mmol) was added. After 10 additional minutes, a third and final dose of zinc powder (206 mg, 3.15 mmol) and ethyl iodide (0.17 mL, 2.1 mmol) was added. After 30 minutes of total reaction time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced

pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-*N*-acetyl-norvaline methyl ester. (27 mg, 75% yield). IR ($\nu_{\max}$, film): 3284, 3070, 2961, 2876, 1746, 1656, 1546, 1438, 1375, 1229, 1204, 1156. ^1H NMR (CDCl_3 , 500 MHz): δ = 0.91 (3H, t, J = 7.3, CH_2CH_3), 1.25-1.41 (2H, m, CH_2CH_3), 1.59-1.66 (1H, m, $\text{H}\beta$), 1.75-1.82 (1H, m, $\text{H}\beta'$), 2.01 (3H, s, NHAc), 3.73 (3H, s, CO_2Me), 4.60 (1H, td, J = 7.8, 5.5, $\text{H}\alpha$), 6.10 (1H, d, J = 7.8, NHAc). ^{13}C NMR (CDCl_3 , 125 MHz): δ = 13.6 (CH_2CH_3), 18.5 (CH_2CH_3), 23.1 (NHAc), 34.5 ($\text{C}\beta$), 51.9, 52.2 ($\text{C}\alpha$, CO_2Me), 169.8, 173.3 ($2 \times \text{C=O}$). HRMS m/z (ESI^+): Found 196.0940 $[\text{M}+\text{Na}]^+$; $\text{C}_8\text{H}_{15}\text{NO}_3\text{Na}$ requires 196.0944.

Table 8.2, Entry 4

Chemical conversion of dehydroalanine to norleucine

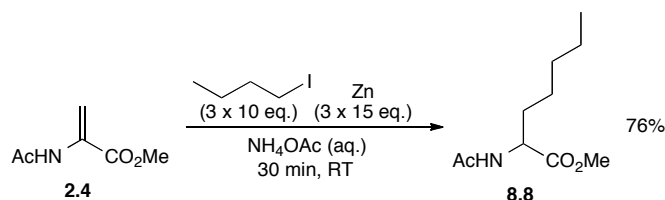


AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H_2O (2.0 mL). Aqueous NH_4OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by 1-iodopropane (0.20 mL, 2.1 mmol). The reaction was stirred at room temperature for 10 minutes and then a second dose of zinc powder (206 mg, 3.15 mmol) and 1-iodopropane (0.20 mL, 2.1 mmol) was added. After 10 additional minutes, a third and final dose of zinc powder (206 mg, 3.15 mmol) and 1-iodopropane (0.20 mL, 2.1 mmol) was added. After 30 minutes of total reaction time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO_3 (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-*N*-acetyl-norleucine methyl ester. (29 mg, 74% yield). IR ($\nu_{\max}$, film): 3284, 3062, 2957, 2934, 2863, 2360, 1746, 1655, 1542, 1436, 1374, 1218, 1200, 1154. ^1H NMR (CDCl_3 , 500 MHz): δ = 0.88 (3H, t, J = 7.3, CH_2CH_3), 1.22-1.33 (4H, m, $\text{CH}_2\text{CH}_2\text{CH}_3$), 1.61-1.69 (1H, m, $\text{H}\beta$), 1.78-1.85 (1H, m, $\text{H}\beta'$), 2.01 (3H, s, NHAc), 3.73 (3H, s, CO_2Me), 4.59 (1H, td, J = 7.6, 5.5, $\text{H}\alpha$), 6.07 (1H, d, J = 7.6, NHAc). ^{13}C NMR (CDCl_3 , 125 MHz): δ = 13.8 (CH_2CH_3), 22.3, 23.1, 27.3 ($\text{CH}_2\text{CH}_2\text{CH}_3$).

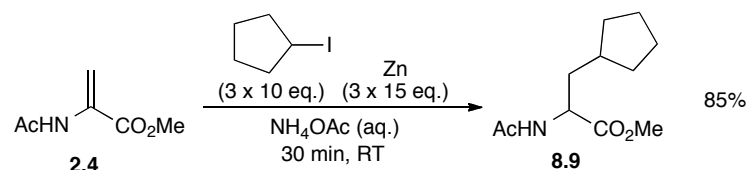
and NHAc), 32.2 (C β), 52.1, 52.3 (C α and CO₂Me), 169.7, 173.3 (2 $\times$ C=O). HRMS *m/z* (ESI⁺): Found 210.1100 [M+Na]⁺; C₉H₁₇NO₃Na requires 210.1101.

Table 8.2, Entry 5

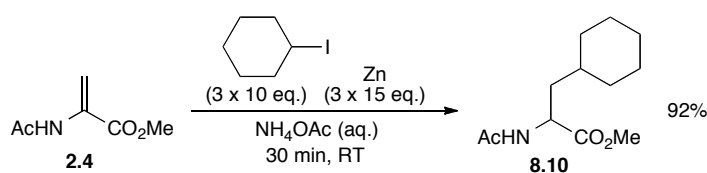
Chemical conversion of dehydroalanine to *n*-butylalanine



AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by 1-iodobutane (0.24 mL, 2.1 mmol). The reaction was stirred at room temperature for 10 minutes and then a second dose of zinc powder (206 mg, 3.15 mmol) and 1-iodobutane (0.24 mL, 2.1 mmol) was added. After 10 additional minutes, a third and final dose of zinc powder (206 mg, 3.15 mmol) and 1-iodobutane (0.24 mL, 2.1 mmol) was added. After 30 minutes of total reaction time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-*N*-acetyl-*n*-butylalanine methyl ester. (32 mg, 74% yield). IR (ν_{max} , film): 3284, 3063, 2956, 2931, 2861, 2360, 2341, 1746, 1655, 1543, 1437, 1374, 1292, 1259, 1210, 1153. ¹H NMR (CDCl₃, 500 MHz): δ = 0.88 (3H, t, *J* = 6.9, CH₂CH₃), 1.26-1.35 (6H, m, CH₂CH₂CH₂CH₃), 1.62-1.70 (1H, m, H β), 1.80-1.87 (1H, m, H β'), 2.03 (3H, s, NHAc), 3.75 (3H, s, CO₂Me), 4.62 (1H, td, *J* = 7.6, 5.5, H α), 5.94 (1H, d, *J* = 7.6, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 13.9 (CH₂CH₃), 22.4, 23.2, 24.8, 31.3 (CH₂CH₂CH₂CH₃ and NHAc), 32.5 (C β), 52.1 (C α), 52.3 (CO₂Me), 169.7, 173.3 (2 $\times$ C=O). HRMS *m/z* (ESI⁺): Found 224.1257 [M+Na]⁺; C₁₀H₁₉NO₃Na requires 224.1257.

Table 8.2, Entry 6**Conversion of dehydroalanine to cyclopentylalanine**

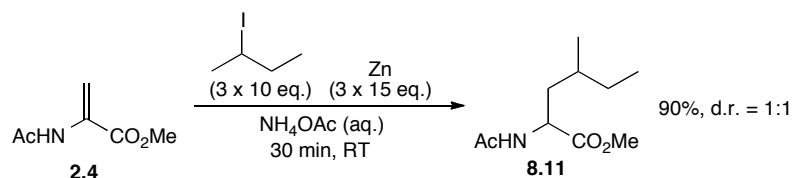
AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, and then cyclopentyl iodide was added dropwise over 5 minutes (0.24 mL, 2.1 mmol). After completion of the addition, the reaction was stirred 5 minutes at room temperature and then a second dose of zinc powder (206 mg, 3.15 mmol) was added followed by cyclopentyl iodide (0.24 mL, 2.1 mmol, added over 5 minutes). After completion of the addition, the reaction was stirred 5 minutes at room temperature and then a third and final dose of zinc powder (206 mg, 3.15 mmol) was added followed by cyclopentyl iodide (0.27 mL, 2.1 mmol, added over 5 minutes). After a final 5 minutes of stirring, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-*N*-acetyl-cyclopentylalanine methyl ester. (38 mg, 85% yield). IR (ν_{max} , film): 3282, 3063, 2951, 2868, 2360, 2342, 1747, 1656, 1542, 1437, 1373, 1206, 1168. ¹H NMR (CDCl₃, 500 MHz): δ = 1.08-1.13 (2H, m, Cp), 1.49-1.62 (4H, m), 1.66-1.82 (5H, m) (Cp and H β), 2.01 (3H, s, NHAc), 3.72 (3H, s, CO₂Me), 4.60 (1H, td, J = 8.2, 5.0, H α), 6.03 (1H, d, J = 8.2, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 23.1 (NHAc), 24.9, 25.1, 32.6, 32.7, 36.5 (Cp), 38.8 (C β), 51.8 (C α), 52.2 (CO₂Me), 169.7, 173.5 (2 $\times$ C=O). HRMS m/z (ESI⁺): Found 236.1255 [M+Na]⁺; C₁₁H₁₉NO₃Na requires 236.1257.

Table 8.2, Entry 7**Chemical conversion of dehydroalanine to cyclohexylalanine**

AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, and then cyclohexyl iodide was added dropwise over 5 minutes (0.27 mL, 2.1 mmol). After completion of the addition, the reaction was stirred 5 minutes at room temperature and then a second dose of zinc powder (206 mg, 3.15 mmol) was added followed by cyclohexyl iodide (0.27 mL, 2.1 mmol, added over 5 minutes). After completion of the addition, the reaction was stirred 5 minutes at room temperature and then a third and final dose of zinc powder (206 mg, 3.15 mmol) was added followed by cyclohexyl iodide (0.27 mL, 2.1 mmol, added over 5 minutes). After a final 5 minutes of stirring, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-*N*-acetyl-cyclohexylalanine methyl ester (44 mg, 92% yield). IR (ν_{max} , film): 3283, 3067, 2996, 2924, 2851, 1747, 1659, 1540, 1448, 1373, 1301, 1278, 1250, 1221, 1175. ¹H NMR (CDCl₃, 500 MHz): δ = 0.86-0.96 (2H, m, Cy), 1.10-1.34 (4H, m, Cy), 1.48-1.54 (1H, m, H β), 1.64-1.72 (5H, m, H β' and Cy), 1.78-1.81 (1H, m, Cy), 2.03 (3H, s, NHAc), 3.74 (3H, s, CO₂Me), 4.64 (1H, td, J = 8.7, 5.4, H α), 5.83 (1H, d, J = 8.7, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 23.2 (NHAc), 26.0, 26.1, 26.3, 32.6, 33.4, 34.1 (Cy), 40.2 (C β), 50.1 (C α), 52.3 (CO₂Me), 169.8, 173.8 (2 $\times$ C=O). HRMS m/z (ESI⁺): Found 250.1410 [M+Na]⁺; C₁₂H₂₁NO₃Na requires 250.1414.

Table 8.2, Entry 8

Chemical conversion of dehydroalanine to *s*-butylalanine

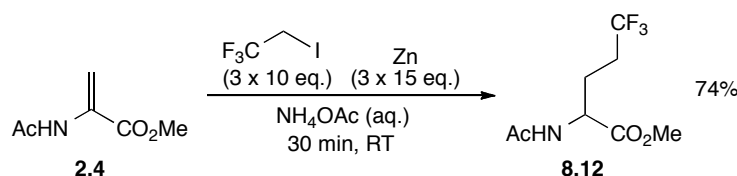


AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by 2-iodobutane (0.24 mL, 2.1 mmol). The reaction was stirred at room temperature for 10 minutes and then a second dose of zinc powder (206 mg, 3.15

mmol) and 2-iodobutane (0.24 mL, 2.1 mmol) was added. After 10 additional minutes, a third and final dose of zinc powder (206 mg, 3.15 mmol) and 2-iodobutane (0.24 mL, 2.1 mmol) was added. After 30 minutes of total reaction time, the reaction was diluted with EtOAc (100 mL) and washed sequentially with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide (±)-*N*-acetyl-*sec*-butylalanine methyl ester as a mixture of diastereomers (38 mg, 90% yield). The d.r. was 1:1, as determined by ¹H NMR. IR (ν_{max}, film): 3282, 3070, 2960, 2933, 2876, 2360, 2341, 1747, 1655, 1543, 1460, 1437, 1375, 1265, 1202, 1161. ¹H NMR (CDCl₃, 500 MHz, data for both diastereomers shown): δ = 0.84-0.88 (6H, 2 × t, *J* = 6.0, CH₂CH₃), 0.92 (6H, 2 × d, *J* = 6.6, CHCH₃), 1.15-1.25 (2H, m), 1.30-1.48 (4H, m), 1.53-1.63 (2H, m), 1.73-1.80 (2H, m) (CHCH₂, Hβ), 2.01, 2.02 (6H, 2 × s, NHAc), 3.73 (6H, s, CO₂Me), 4.61-4.68 (2H, m, Hα), 5.88-5.95 (2H, 2 × d, *J* = 8.2, 7.3, NHAc). ¹³C NMR (CDCl₃, 125 MHz, data for both diastereomers shown): δ = 10.8, 11.2 (CH₂CH₃), 18.6, 19.1 (CHCH₃), 23.1 (br, NHAc), 28.7, 29.7, 30.9, 31.1 (CHCH₂), 39.6, 39.7 (Cβ), 50.4, 50.6 (Cα), 52.2, 52.3 (CO₂Me), 169.7, 169.9, 173.3, 173.8 (C=O). HRMS *m/z* (ESI⁺): Found 224.1258 [M+Na]⁺; C₁₀H₁₉NO₃Na requires 224.1257.

Table 8.2, Entry 9

Chemical conversion of dehydroalanine to δ,δ,δ-trifluoronorvaline

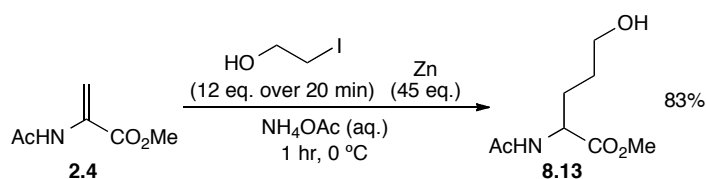


AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (206 mg, 3.15 mmol) was added to the stirred solution, followed by 1,1,1-trifluoro-2-iodoethane (0.21 mL, 2.1 mmol, dropwise over 2 minutes). The reaction was stirred at room temperature for 8 minutes and then a second dose of zinc powder (206 mg, 3.15 mmol) and 1,1,1-trifluoro-2-iodoethane (0.21 mL, 2.1 mmol, dropwise over 2 minutes) was added. After 8 additional minutes, a third and final dose of zinc powder (206 mg, 3.15 mmol) and 1,1,1-trifluoro-2-iodoethane (0.21 mL, 2.1 mmol, dropwise over 2 minutes) was added. After 30 minutes of total reaction time, the reaction was diluted with EtOAc (100 mL) and washed sequentially

with NaHCO₃ (sat. aq., 100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (80% EtOAc in petrol) to provide desired product as a white crystals (35 mg, 74%). m.p. = 86-87 °C. IR ($\nu_{\max}$, film): 3075, 2959, 2360, 2341, 1747, 1658, 1551, 1456, 1439, 1396, 1376, 1340, 1279, 1255, 1210. ¹H NMR (CDCl₃, 500 MHz): δ = 1.85-1.91 (1H, m, H β), 2.04 (3H, s, NHAc), 2.07-2.24 (3H, m, H β' and CH₂CF₃), 3.78 (3H, s, CO₂Me), 4.67 (1H, td, J = 7.7, 4.7, H α), 6.26 (1H, d, J = 7.7, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 23.0 (NHAc), 25.4 (q, J_{CF} = 2.9, C β), 30.2 (q, J_{CF} = 29.6, CH₂CF₃), 50.9 (C α), 52.7 (CO₂Me), 126.5 (q, J_{CF} = 275.6, CF₃), 170.1, 172.0 (2 $\times$ C=O). ¹⁹F NMR (CDCl₃, 377 MHz): δ = -66.5. HRMS m/z (ESI⁺): Found 250.0656 [M+Na]⁺; C₈H₁₂F₃NO₃Na requires 250.0662.

Table 8.2, Entry 10

Chemical conversion of dehydroalanine to *bis*-homoserine

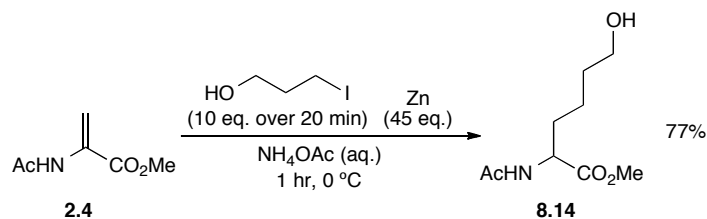


AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H₂O (2.0 mL). Aqueous NH₄OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (618 mg, 9.45 mmol) was added to the stirred solution and the mixture was cooled to 0 °C. 2-Iodoethanol (0.20 mL, 2.6 mmol) was added by syringe pump over 20 minutes. The reaction was stirred at 0 °C for an additional 40 minutes upon completion of the addition of iodoethanol. After this time, the reaction mixture was filtered through cotton, rinsing with 20 mL MeOH. The filtrate was concentrated under reduced pressure and partially purified by column chromatography (10% MeOH in EtOAc). The fractions containing the product were combined, concentrated, and purified again by column chromatography (5% MeOH in EtOAc). The product was isolated as a pale yellow oil (33 mg, 83% yield). IR ($\nu_{\max}$, film): 3283, 3072, 2954, 2869, 2360, 2341, 1742, 1655, 1542, 1437, 1375, 1269, 1212, 1176, 1060. ¹H NMR (CDCl₃, 500 MHz): δ = 1.55-1.66 (2H, m, CH₂CH₂OH), 1.75-1.82 (1H, m, H β), 1.89-1.96 (1H, m, H β'), 2.03 (3H, s, NHAc), 2.45 (1H, br. s, OH), 3.66 (2H, t, J = 6.1, CH₂OH), 3.75 (3H, s, CO₂Me), 4.63 (1H, td, J = 7.6, 5.5, H α), 6.45 (1H, d, J = 7.6,

NHAc). ^{13}C NMR (CDCl_3 , 125 MHz): δ = 23.1 (NHAc), 28.1 ($\text{CH}_2\text{CH}_2\text{OH}$), 29.3 ($\text{C}\beta$), 51.9 ($\text{C}\alpha$), 52.4 (CO_2Me), 61.9 (CH_2OH), 170.2, 173.0 ($2 \times \text{C}=\text{O}$). HRMS m/z (ESI^+): Found 212.0892 $[\text{M}+\text{Na}]^+$; $\text{C}_8\text{H}_{15}\text{NO}_4\text{Na}$ requires 212.0893.

Table 8.2, Entry 11

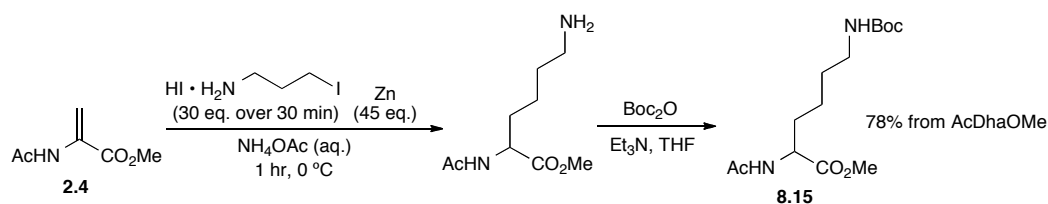
Chemical conversion of dehydroalanine to ϵ -hydroxynorleucine



AcDhaOMe (30 mg, 0.21 mmol) was added to a 10 mL vial and suspended in H_2O (2.0 mL). Aqueous NH_4OAc (0.42 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (618 mg, 9.45 mmol) was added to the stirred solution and the mixture was cooled to 0 °C. 3-Iodo-1-propanol (0.20 mL, 2.1 mmol) was added by syringe pump over 20 minutes. The reaction was stirred at 0 °C for an additional 40 minutes upon completion of the addition of 3-iodo-1-propanol. After this time, the reaction mixture was filtered through cotton, rinsing with 20 mL MeOH. The filtrate was concentrated under reduced pressure and partially purified by column chromatography (5% MeOH in EtOAc). The fractions containing the product were combined, concentrated, and purified again by column chromatography (5% MeOH in EtOAc). The product was isolated as a pale yellow oil (33 mg, 77% yield). IR (ν_{max} , film): 3285, 3073, 2949, 2866, 2361, 2341, 1744, 1657, 1548, 1437, 1375, 1295, 1261, 1213, 1173. ^1H NMR (CDCl_3 , 500 MHz): δ = 1.34-1.47 (2H, m, $\text{H}\gamma$), 1.51-1.63 (2H, m, $\text{CH}_2\text{CH}_2\text{OH}$), 1.65-1.73 (1H, m, $\text{H}\beta$), 1.81-1.89 (1H, m, $\text{H}\beta'$), 2.02 (3H, s, NHAc), 2.20 (1H, br. s, OH), 3.62 (1H, td, J = 6.3, 2.5, CH_2OH), 3.73 (3H, s, CO_2Me), 4.59 (1H, td, J = 7.9, 5.0, $\text{H}\alpha$), 6.31 (1H, d, J = 7.9, NHAc). ^{13}C NMR (CDCl_3 , 125 MHz): δ = 21.6 ($\text{C}\gamma$), 23.1 (NHAc), 31.9, 32.2 ($\text{C}\beta$ and $\text{CH}_2\text{CH}_2\text{OH}$), 52.1 ($\text{C}\alpha$), 52.4 (CO_2Me), 62.2 (CH_2OH), 170.2, 173.1 ($2 \times \text{C}=\text{O}$). HRMS m/z (ESI^+): Found 226.1050 $[\text{M}+\text{Na}]^+$; $\text{C}_9\text{H}_{17}\text{NO}_4\text{Na}$ requires 226.1050.

Table 8.2, Entry 12

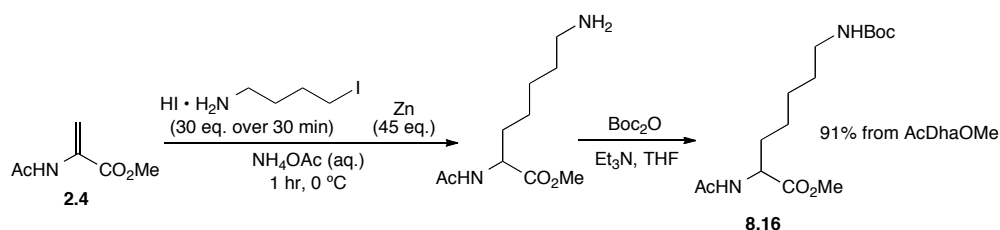
Chemical conversion of dehydroalanine to lysine



AcDhaOMe (15 mg, 0.10 mmol) was added to a 10 mL vial and suspended in H₂O (0.5 mL). Aqueous NH₄OAc (0.21 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (294 mg, 4.5 mmol) was added to the stirred solution and the mixture was cooled to 0 °C. 3-Iodo-1-aminopropane hydroiodide (984 mg, 3.0 mmol) was dissolved in H₂O (1 mL) and then added by syringe pump to the reaction mixture over a period of 30 minutes at 0 °C. The reaction was stirred at 0 °C for an additional 30 minutes upon completion of the addition. After this time, the reaction mixture was filtered through cotton, rinsing with 20 mL MeOH. The filtrate was concentrated under reduced pressure. The residue was then dissolved in THF (5 mL) and Boc₂O (1.1 g, 5.0 mmol) was added to the stirred solution. After cooling to 0 °C, Et₃N (1.4 mL, 10 mmol) was added. After the addition, the reaction was stirred at room temperature for one hour. After this time, the reaction was diluted with EtOAc (150 mL) and washed sequentially with H₂O (100 mL) and brine (100 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc), providing (±)-AcLys(Boc)OMe as a thick clear oil (25 mg, 78% yield from AcDhaOMe). IR (ν_{max} , film): 3313, 3064, 2934, 1746, 1688, 1661, 1533, 1438, 1366, 1272, 1251, 1210, 1172. ¹H NMR (CDCl₃, 500 MHz): δ = 1.27-1.41 (2H, m, CH₂CH₂CH₂NHBoc), 1.44 (9H, s, Boc), 1.44-1.52 (2H, m, CH₂CH₂NHBoc), 1.66-1.73 (1H, m, H β), 1.80-1.88 (1H, m, H β'), 2.03 (3H, s, NHAc), 3.10 (2H, q, J = 6.1, CH₂NHBoc), 3.74 (3H, s, CO₂Me), 4.56-4.60 (2H, m, H α and NHBoc), 6.16 (1H, d, J = 6.0, NHAc). ¹³C NMR (CDCl₃, 125 MHz): δ = 22.3 (CH₂CH₂CH₂NHBoc), 23.2 (NHAc), 28.4 (Boc), 29.7 (CH₂CH₂NHBoc), 31.9 (C β), 39.9 (CH₂NHBoc), 52.0 (C α), 52.4 (CO₂Me), 79.2 (Boc), 156.1, 169.9, 173.0 (3 × C=O). HRMS m/z (ESI⁺): Found 325.1727 [M+Na]⁺; C₁₄H₂₆N₂O₅Na requires 325.1734.

Table 8.2, Entry 13

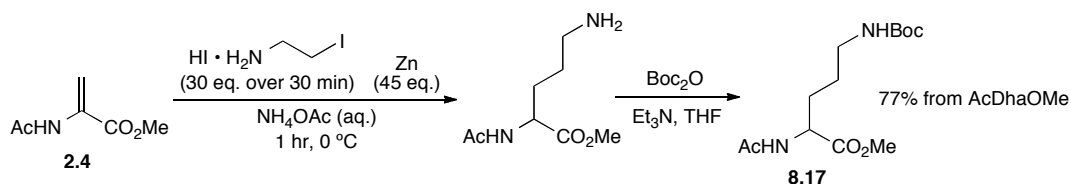
Chemical conversion of dehydroalanine to homolysine



AcDhaOMe (15 mg, 0.10 mmol) was added to a 10 mL vial and suspended in H_2O (0.5 mL). Aqueous NH_4OAc (0.21 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (294 mg, 4.5 mmol) was added to the stirred solution and the mixture was cooled to $0\text{ }^\circ\text{C}$. 4-Iodo-1-aminobutane hydroiodide (981 mg, 3.0 mmol) was dissolved in H_2O (1 mL) and then added by syringe pump to the reaction mixture over a period of 30 minutes at $0\text{ }^\circ\text{C}$. The reaction was stirred at $0\text{ }^\circ\text{C}$ for an additional 30 minutes upon completion of the addition. After this time, the reaction mixture was filtered through cotton, rinsing with 20 mL MeOH . The filtrate was concentrated under reduced pressure. The residue was then dissolved in THF (5 mL) and Boc_2O (1.1 g, 5.0 mmol) was added to the stirred solution. After cooling to $0\text{ }^\circ\text{C}$, Et_3N (1.4 mL, 10 mmol) was added. After the addition, the reaction was stirred at room temperature for one hour. After this time, the reaction was diluted with EtOAc (150 mL) and washed sequentially with H_2O (100 mL) and brine (100 mL). The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc), providing the racemic homolysine derivative as a pale yellow oil (30 mg, 91% yield from AcDhaOMe). IR (ν_{max} , film): 3314, 2975, 2934, 2863, 3260, 2341, 1746, 1692, 1662, 1537, 1438, 1391, 1366, 1274, 1251, 1209, 1173. ^1H NMR (CDCl_3 , 500 MHz): δ = 1.24-1.36 (4H, m), 1.40-1.47 (11H, m) (contains Boc and $\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{NHBoc}$), 1.60-1.69 (1H, m, $\text{H}\beta$), 1.77-1.84 (1H, m, $\text{H}\beta'$), 2.01 (3H, s, NHAc), 3.07 (2H, m, CH_2NHBoc), 3.73 (3H, s, CO_2Me), 4.56-4.60 (2H, m, $\text{H}\alpha$ and NHBoc), 6.16 (1H, d, J = 6.3, NHAc). ^{13}C NMR (CDCl_3 , 125 MHz): δ = 23.1 (NHAc), 24.7, 26.2, 28.4, 29.7 (contains Boc and $\text{CH}_2\text{CH}_2\text{CH}_2\text{CH}_2\text{NHBoc}$), 32.3 ($\text{C}\beta$), 40.2 (CH_2NHBoc), 52.0 ($\text{C}\alpha$), 52.3 (CO_2Me), 79.0 (Boc), 156.0, 169.8, 173.1 ($3 \times \text{C}=\text{O}$). HRMS m/z (ESI^+): Found 339.1881 $[\text{M}+\text{Na}]^+$; $\text{C}_{15}\text{H}_{28}\text{N}_2\text{O}_5\text{Na}$ requires 339.1890.

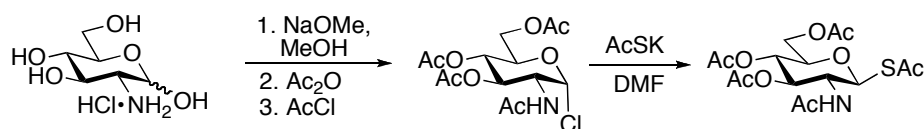
Table 8.2, Entry 14

Chemical conversion of dehydroalanine to ornithine



AcDhaOMe (15 mg, 0.10 mmol) was added to a 10 mL vial and suspended in H_2O (0.5 mL). Aqueous NH_4OAc (0.21 mL, 5M, pH 6.0) was added and the mixture was stirred until all material had dissolved. Zinc powder (294 mg, 4.5 mmol) was added to the stirred solution and the mixture was cooled to $0\text{ }^\circ\text{C}$. 2-Iodo-1-aminoethane hydroiodide (897 mg, 3.0 mmol) was dissolved in H_2O (1 mL) and then added by syringe pump to the reaction mixture over a period of 30 minutes at $0\text{ }^\circ\text{C}$. The reaction was stirred at $0\text{ }^\circ\text{C}$ for an additional 30 minutes upon completion of the addition. After this time, the reaction mixture was filtered through cotton, rinsing with 20 mL MeOH. The filtrate was concentrated under reduced pressure. The residue was then dissolved in THF (5 mL) and Boc_2O (1.1 g, 5.0 mmol) was added to the stirred solution. After cooling to $0\text{ }^\circ\text{C}$, Et_3N (1.4 mL, 10 mmol) was added. After the addition, the reaction was stirred at room temperature for one hour. After this time, the reaction was diluted with EtOAc (150 mL) and washed sequentially with H_2O (100 mL) and brine (100 mL). The organic layer was dried (MgSO_4), filtered, and concentrated under reduced pressure. The resulting residue was purified by column chromatography (EtOAc), providing the racemic ornithine derivative as a pale yellow oil (23 mg, 77% yield from AcDhaOMe). IR (ν_{max} , film): 3313, 3067, 2976, 2934, 2869, 2361, 1744, 1689, 1662, 1533, 1451, 1438, 1391, 1366, 1275, 1252, 1215, 1172. ^1H NMR (CDCl_3 , 500 MHz): δ = 1.43 (9H, s, Boc), 1.48-1.57 (2H, m, $\text{CH}_2\text{CH}_2\text{NHBoc}$), 1.64-1.71 (1H, m, $\text{H}\beta$), 1.83-1.88 (1H, m, $\text{H}\beta'$), 2.03 (3H, s, NHAc), 3.13 (2H, q, J = 6.3, CH_2NHBoc), 3.74 (3H, s, CO_2Me), 4.60 (1H, td, J = 7.7, 5.4, $\text{H}\alpha$), 4.67 (1H, br. s, NHBoc), 6.32 (1H, d, J = 7.7, NHAc). ^{13}C NMR (CDCl_3 , 125 MHz): δ = 23.1 (NHAc), 26.2 ($\text{CH}_2\text{CH}_2\text{NHBoc}$), 28.4 (Boc), 29.5 ($\text{C}\beta$), 39.9 (CH_2NHBoc), 51.9 ($\text{C}\alpha$), 52.4 (CO_2Me), 79.3 (Boc), 156.1, 170.0, 172.9 (3 $\times$ $\text{C}=\text{O}$). HRMS m/z (ESI^+): Found 311.1569 [$\text{M}+\text{Na}$] $^+$; $\text{C}_{13}\text{H}_{24}\text{N}_2\text{O}_5\text{Na}$ requires 311.1577.

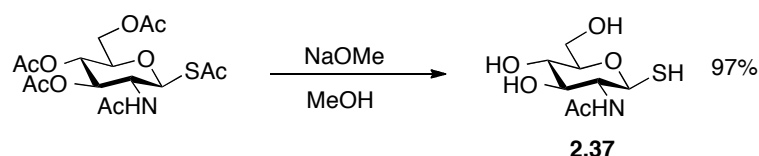
3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio- β -D-glucosamine



(D)-Glucosamine hydrochloride (20.0 g, 92.7 mmol) was added to a 100 mL round bottom flask and suspended in MeOH (20 mL). Sodium methoxide (6.00 g, 111 mmol) was added in several portions and the mixture was stirred vigorously at room temperature for 1 hour before NaCl was removed by filtration. The solids were washed with MeOH (20 mL). Acetic anhydride (12 mL, 127 mmol) was added to the stirred filtrate at 0 °C and warmed slowly to room temperature. After 12 hours, the reaction was concentrated to ~20 mL total volume and the product precipitated by the addition of Et₂O (100 mL). The solid was isolated by filtration and dried under vacuum before returning to an oven dried 2-neck 250 mL round bottom flask equipped with a condenser. The flask was placed under an argon atmosphere and cooled to 0 °C. Acetyl chloride (40.0 mL, 563 mmol) was added slowly and the mixture stirred vigorously. After the addition, the reaction was heated to reflux. Gradually the reaction turned black. After 3 hours at reflux, the reaction was cooled to room temperature, diluted with CH₂Cl₂ (200 mL), and poured into a beaker of ice (~200 g). The organics were washed with NaHCO₃ (sat. aq., 2 × 200 mL), dried (MgSO₄), filtered, and concentrated to a black residue. The residue was dissolved in EtOAc and filtered through a short column of silica, eluting with EtOAc. The product was precipitated by the addition of cold Et₂O and petrol. The resulting powder was isolated by filtration and dried under vacuum (10.19 g, 50% from glucosamine hydrochloride). The crude product (~4:1 mixture of α : β) was used in the final step without purification. Accordingly, a portion of the anomeric chloride (1.00 g, 2.73 mmol) was added to a 50 mL round bottom flask and dissolved in DMF (15 mL). Potassium thioacetate (1.56 g, 13.7 mmol) was then added in one portion. The reaction was flushed with nitrogen and stirred at room temperature for 3 hours. After this time, no starting material was detected by TLC (70% EtOAc in petrol). The reaction mixture was diluted with EtOAc (250 mL) and washed sequentially with NaHCO₃ (sat. aq., 200 mL), H₂O (200 mL), and brine (200 mL). The organic layer was dried (MgSO₄), filtered, and concentrated under reduced pressure to give a brown solid. This solid was recrystallized from CH₂Cl₂ and petrol to afford the product as white crystals (661 mg, 60%). m.p. = 186-188 °C (Lit⁶: 196-197, *i*PrOH). $[\alpha]_D^{20} = +12.1$ (*c* = 2.2, CHCl₃) (Lit⁶: +14.1 *c* = 2.16, CHCl₃). IR ($\nu_{\max}$, KBr): 3733, 3064, 2961, 2360, 2341, 1747, 1707, 1665, 1538, 1433,

1370, 1230, 1048. ^1H NMR (400 MHz, CDCl_3): δ = 1.90 (3H, s), 2.02 (6H, s), 2.05 (3H, s), 2.35 (3H, s) ($5 \times \text{Ac}$), 3.81 (1H, ddd, J = 9.4, 4.6, 1.3, H5), 4.08 (1H, dd, J = 12.4, 1.3, H6), 4.22 (1H, dd, J = 12.4, 4.6, H6'), 4.34 (1H, q, J = 10.3, H2), 5.10-5.16 (2H, m, H3, H4), 5.17 (1H, d, J = 10.3, H1). ^{13}C NMR (100 MHz, CDCl_3): δ = 20.49, 20.59, 20.64, 23.0, 30.8 ($5 \times \text{Ac}$), 51.9 (C2), 61.8 (C6), 67.8, 74.0 (C3, C4), 76.3 (C5), 81.5 (C1), 169.2, 170.0, 170.6, 171.3, 193.5 ($5 \times \text{C}=\text{O}$). LRMS (ESI $^+$): 428.1 ($\text{M}+\text{Na}$) $^+$.

***N*-acetyl-1-thio- β -D-glucosamine (GlcNAc thiol **2.37**)**



3,4,6-Tri-*O*-acetyl-*N*-acetyl-*S*-acetyl-1-thio- β -D-glucosamine (250 mg, 0.62 mmol) was added to a 50 mL round bottom flask and placed under an atmosphere of nitrogen before dissolving in anhydrous MeOH (8 mL). Sodium methoxide (67 mg, 1.2 mmol) was added under a stream of nitrogen at room temperature. The reaction was stirred for 1 hour at room temperature before quenching with DOWEX-50WX8 (H^+ form) until the pH of the reaction mixture was neutral (pH paper). The resin was removed by filtration and washed with MeOH (~ 50 mL). The filtrate was evaporated to provide GlcNAc thiol **2.37** as a thick syrup that crystallized upon standing (142 mg, 97%). m.p. = 170-171 $^\circ\text{C}$ (Lit 7 : 174-176 $^\circ\text{C}$). $[\alpha]_{\text{D}}^{20}$ = -13.5 (c = 1.0, MeOH) (Lit 7 : -10.4, c = 1.0, MeOH). IR (ν_{max} , KBr): 3287, 2360, 2341, 1647, 1558, 1374, 1313, 1056. ^1H NMR (400 MHz, D_2O): δ = 1.90 (3H, s, Ac), 3.30-3.39 (3H, m, H3, H4, H5), 3.56 (1H, dd, J = 11.1, 5.3, H6), 3.59 (1H, t, J = 10.2, H2), 3.74 (1H, dd, J = 11.1, 1.2, H6'), 4.53 (1H, d, J = 10.2, H1). ^{13}C NMR (50 MHz, D_2O): δ = 22.6 (Ac), 58.2 (C2), 61.1 (C6), 70.0, 75.3, 79.3, 80.52 (C3, C4, C5, C1). LRMS (ESI $^-$): 236.1 ($\text{M}-\text{H}$) $^-$. Note: in D_2O this material anomerizes to a 3.6:1.0 mixture of β : α over a period of 24 hours.

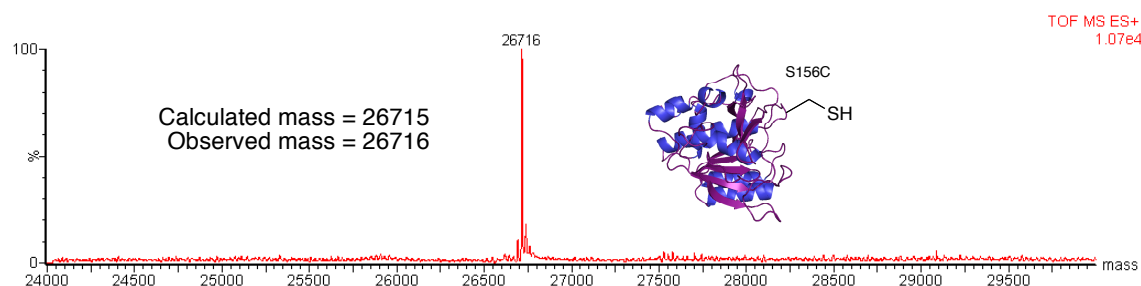
Protein Modification

The single cysteine mutant of subtilisin from *Bacillus lentus* (SBL-S156C) was used in this study (PDB for wild type = 1GCI). The amino acid sequence, average mass, and ESI-MS are shown below.

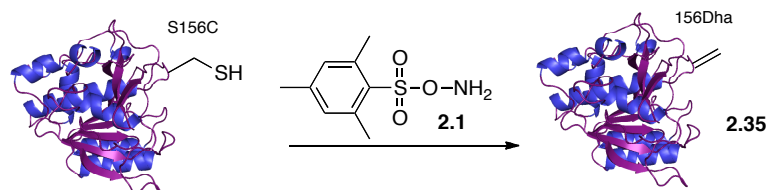
Sequence:

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPS
TQDGNGHGHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGL
EWAGNNGMHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGN**C**GAGSISYPA
RYANAMAVGATDQNNNRASFSQYGAGLDIVAPGVNVQSTYPGSTYASLNGTSM
ATPHVAGAAALVKQKNPSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR
Average Mass = 26715

ESI-MS for SBL-S156C



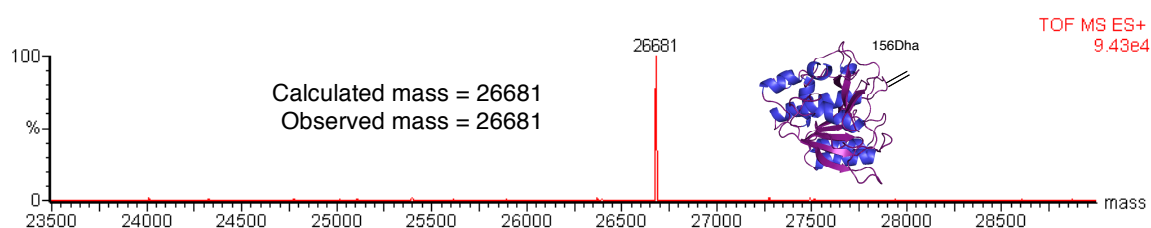
Conversion of cysteine to dehydroalanine using MSH (2.1) on SBL-S156C



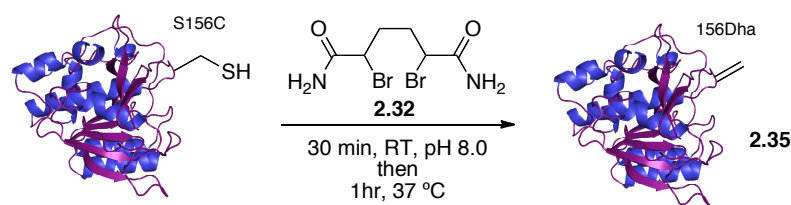
For full details of the use of MSH in the conversion of cysteine to dehydroalanine on proteins, including all control reactions, the reader is referred to Chapter 2.

All manipulations were carried out at 4 °C. SBL-S156C was prepared as a 1 mg/mL solution in pH 8.0 sodium phosphate buffer (50 mM) and 2.0 mL was added to 15 mL Falcon tube. A solution of MSH was prepared by dissolving 5.0 mg (0.023 mmol) in 200 μ L of DMF. All of the MSH solution was added to the protein solution by micropipette and the reaction vortexed immediately to ensure rapid mixing. The reaction was shaken

for 20 minutes and then diluted to 2.50 mL total volume. The reaction mixture was then loaded onto a PD10 column (GE Healthcare) pre-equilibrated with the reaction buffer used in subsequent steps (e.g. 50 mM sodium phosphate pH 8.0 or 500 mM ammonium acetate pH 6.0). The protein was eluted with 3.50 mL of the same buffer. A sample of the PD10 purified protein was analyzed by LC-MS (calculated mass = 26681; observed mass = 26681). ESI-MS is shown below.

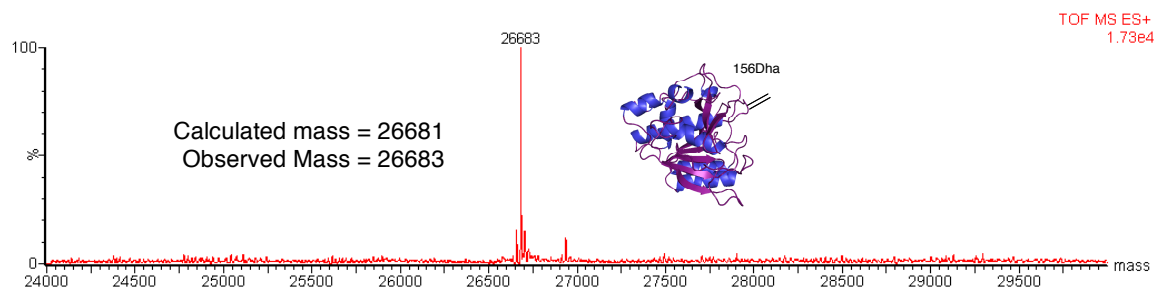


Conversion of cysteine to dehydroalanine using di-bromide **2.32** on SBL-156C

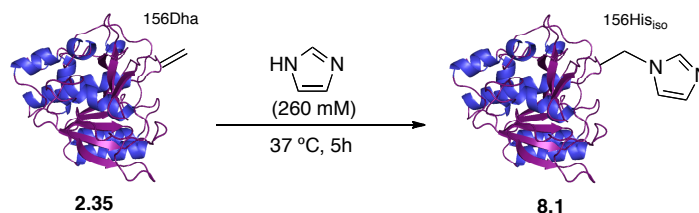


*For full details of the use of **2.32** in the conversion of cysteine to dehydroalanine on proteins, including all control reactions, the reader is referred to Chapter 2.*

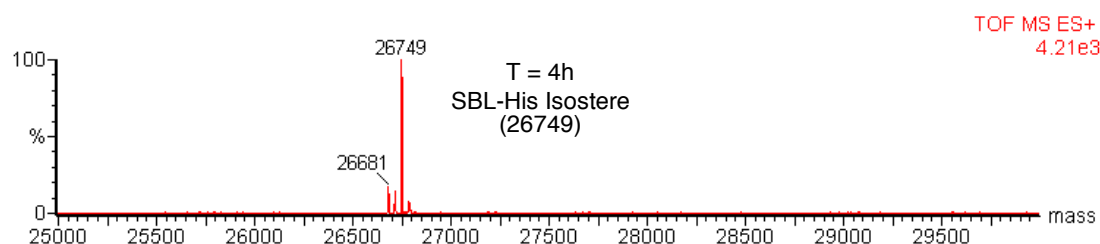
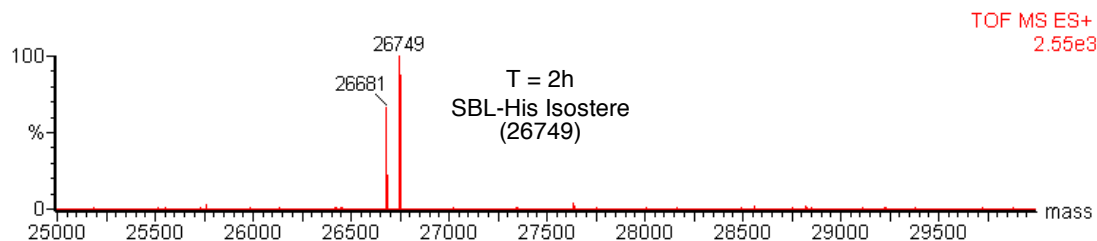
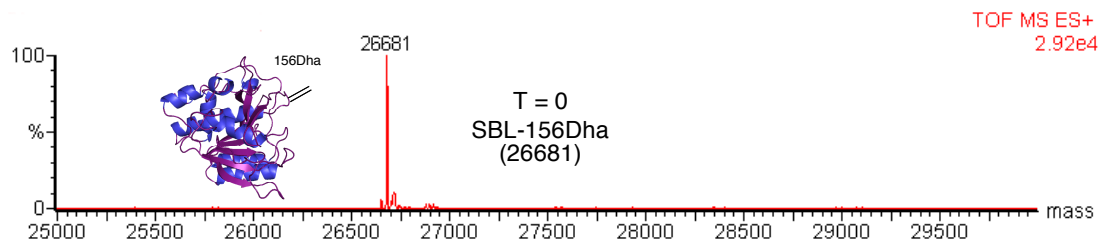
SBL-S156C was prepared at 1.0 mg/mL in sodium phosphate buffer (pH 8.0, 50 mM). DTT (1.4 mg) was added as a solid to a 500 μ L aliquot of the protein solution to reduce any contaminant disulfide. After 15 minutes of incubation at room temperature, the DTT was removed by passing the protein solution through a PD minitrapp (GE Healthcare), eluting with 1 mL of sodium phosphate buffer (pH 8.0, 50 mM). The protein solution was stored on ice until needed. A stock solution of dibromide **2.32** was prepared by dissolving 35.5 mg in 418 μ L DMF. A 450 μ L aliquot of SBL-S156C ($c = 0.5$ mg/mL, pH 8.0 sodium phosphate buffer) was warmed to room temperature. A 45 μ L aliquot of the dibromide solution was added to the protein. The reaction was shaken at room temperature for 30 minutes and then at 37 $^{\circ}$ C for 1 hour. After this time, solids were removed by centrifugation (1 min, 14K g) and the solid was passed through a PD minitrapp, eluting with 1 mL of the reaction buffer used in the subsequent step. LC-MS analysis of the product revealed a mass corresponding to the dehydroalanine containing protein (calculated mass = 26681; observed mass = 26682). ESI-MS is shown below.

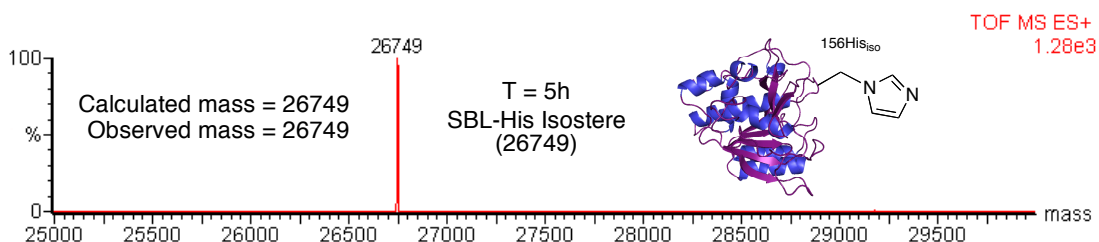


Chemical mutation of dehydroalanine to histidine analog (8.1)

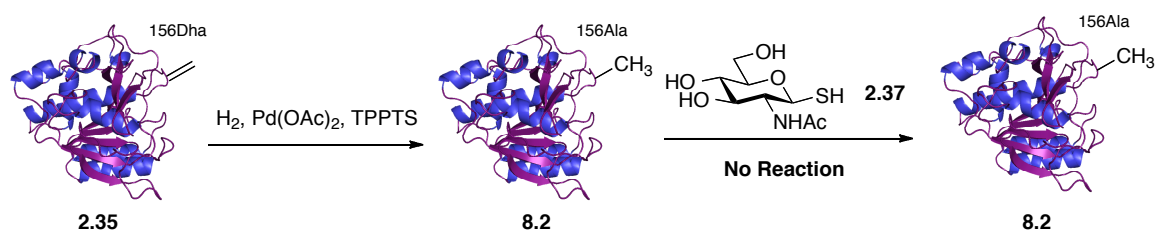


A 200 μ L aliquot of SBL-156Dha (**2.35**) was prepared at 0.7 mg/mL in sodium phosphate buffer (pH 8.0). Imidazole (3.6 mg, 0.052 mmol) was added to the protein solution as a solid. The reaction was incubated at 37 °C and analyzed by LCMS at 2, 4, and 5 hours after which time complete conversion to the histidine isostere was observed (calculated mass = 26749, observed mass = 26749). ESI-MS are shown below.

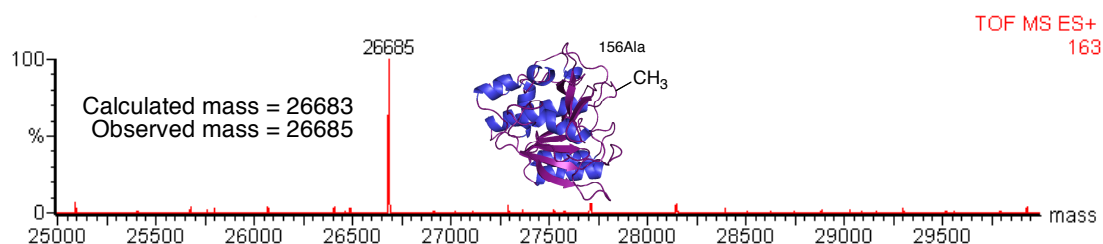




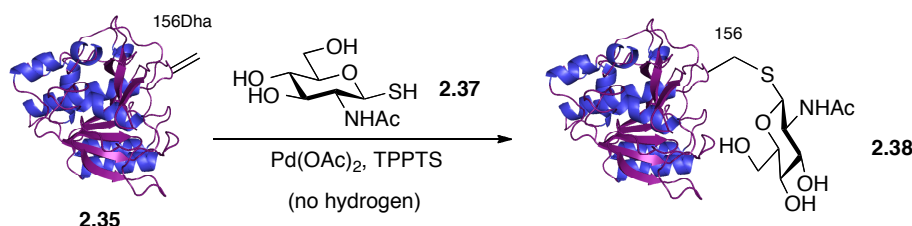
Chemical mutation of dehydroalanine to alanine (8.2)



A 200 μL sample of SBL-156Dha (**2.35**) was prepared at 0.30 mg/mL in 50 mM sodium phosphate buffer (pH 8.0, ~ 0.003 μmol protein) and was added to 1.50 mL plastic tube. A stock solution of catalyst was prepared by adding 0.8 mg $\text{Pd}(\text{OAc})_2$ (3.6 μmol) and 6.0 mg of tris(3-sulfophenyl)phosphine trisodium salt (TPPTS, 10.6 μmol) to a 1.50 mL plastic tube and dissolving in 200 μL of 50 mM sodium phosphate (pH 8.0) with the aid of sonication. This solution is approximately 18 mM in Pd. A 20 μL aliquot of the catalyst solution (~ 0.3 μmol) was added to the protein solution which was then vortexed to homogenize and sealed with a rubber septa. Hydrogen (1 atm, balloon) was bubbled through the solution for 5 minutes and the reaction incubated in a 37 $^\circ\text{C}$ water bath under an H_2 atmosphere for 3 hours. Since the mass change expected in the hydrogenation (+ 2 Da) is within the error of the mass spectrometer, conversion of dehydroalanine to alanine cannot be inferred from the mass spectrum. Therefore, hydrogenation was inferred indirectly: a 60 μL aliquot of the reaction mixture was added directly to 1 mg of GlcNAc thiol **2.37** (4.2 μmol ; ~ 5000 eq). The mixture was vortexed to dissolve the thiol and then rotated at room temperature for 30 minutes. LC-MS analysis of the sample showed no addition of **2.37** indicating that dehydroalanine had been consumed in the hydrogenation. ESI-MS is shown below. Under the same conditions (Pd, TPPTS) but without hydrogen, **2.37** adds to dehydroalanine (see next).



Control: addition of thiol **2.37 to dehydroalanine in presence of Pd(OAc)₂ and TPPTS**



To verify that GlcNAc thiol **2.37** can add to SBL-Dha in the presence of palladium, a control without hydrogen was run. Accordingly, 56 μ L of the protein at 0.30 mg/mL was mixed with 4 μ L of the 18 mM palladium solution prepared above from Pd(OAc)₂ and TPPTS. The resulting solution was added to 1 mg of GlcNAcSH, vortexed, and rotated at room temperature for 30 minutes. LC-MS analysis of the reaction mixture showed full conversion to glycoprotein **2.38** (calculated mass = 26918; observed mass = 26921). ESI-MS is shown below.

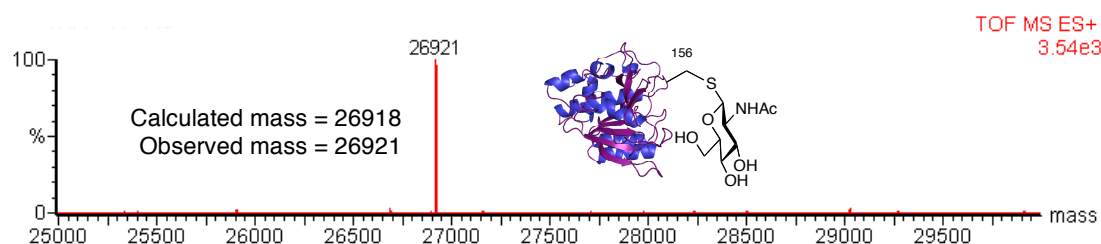
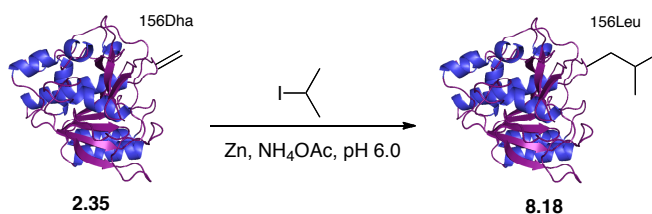
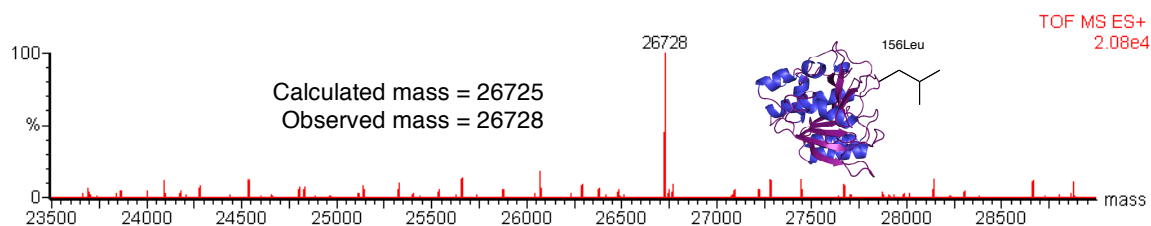


Figure 8.13A, Entry 6

Chemical mutation of dehydroalanine to leucine (8.18**)**



A 250 μL sample of SBL-Dha156 (**2.35**, 0.29 mg/mL) was prepared in 500 mM NH_4OAc (pH 6.0). Zn powder (2 mg) was added to the protein solution followed by 2-iodopropane (2 μL). The reaction mixture was vortexed and then shaken vigorously at room temperature for 15 minutes. Solids were removed by centrifugation and a 35 μL aliquot was analyzed by LC-MS. Full conversion to a protein with a mass consistent with the formation of SBL-156Leu was observed (calculated mass = 26725, observed mass = 26728). ESI-MS is shown below.

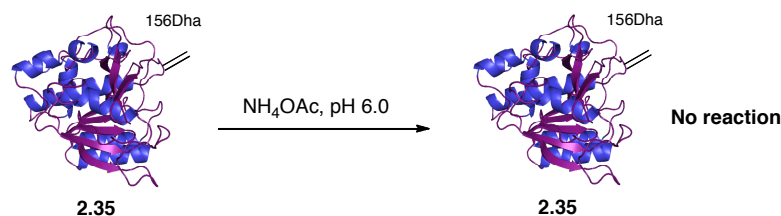


Control experiments

The following control experiments were carried out to show that the alkyl iodide, zinc, and ammonium acetate buffer are all required for the reaction to occur:

Figure 8.13A, Entry 1

SBL-Dha (**2.35**) and NH_4OAc : no reaction



A 20 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL, pH 6.0 NH_4OAc , 500 mM) was shaken at room temperature for 15 minutes. LC-MS analysis showed only unchanged starting material (calculated mass = 26681; observed mass = 26680). ESI-MS is shown below.

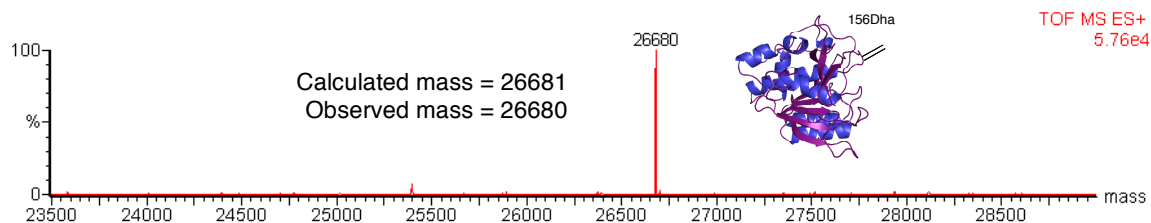
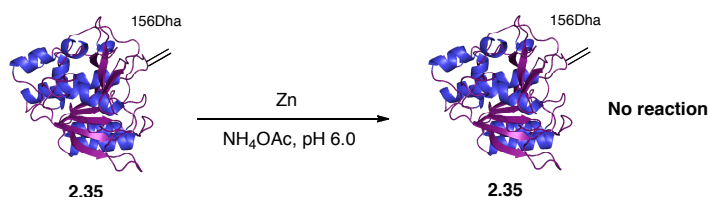


Figure 8.13A, Entry 2

SBL-156Dha (2.35) and Zn: no reaction



A 115 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL, pH 6.0 NH₄OAc, 500 mM) prepared above was mixed with 1 mg of Zn powder and shaken at room temperature for 15 minutes. The zinc was then separated by centrifugation and the supernatant analyzed by LC-MS. Only unmodified starting material was observed (calculated mass = 26681; observed mass = 26680). ESI-MS is shown below.

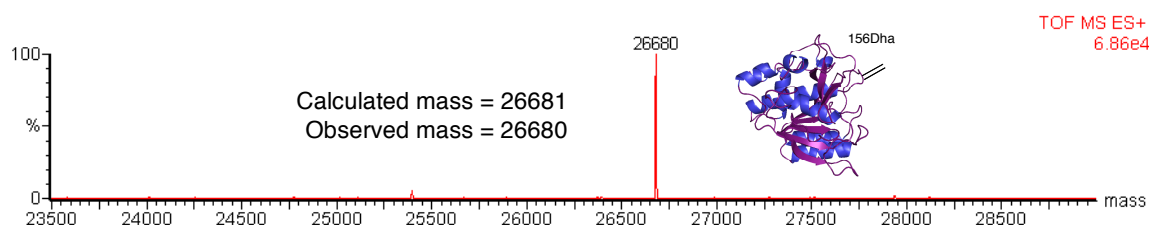
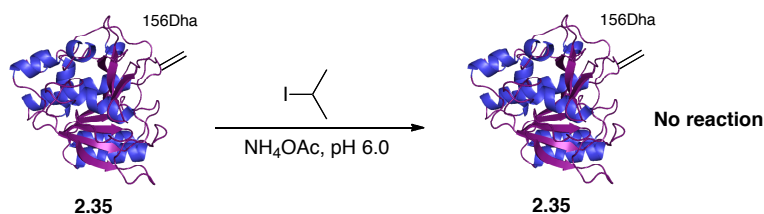


Figure 8.13A, Entry 3

SBL-S156C (2.35) and 2-iodopropane: no reaction



A 115 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL, pH 6.0 NH₄OAc, 500 mM) and 2 μ L of 2-iodopropane were mixed and shaken at room temperature for 15 minutes. LC-MS analysis of the mixture showed only unmodified starting material (calculated mass = 26681; observed mass = 26681). ESI-MS is shown below.

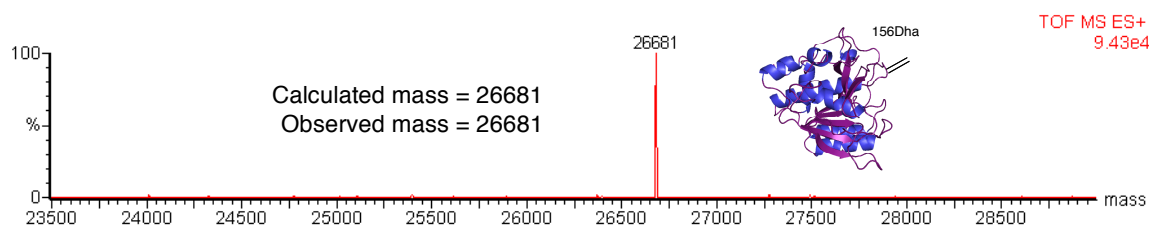
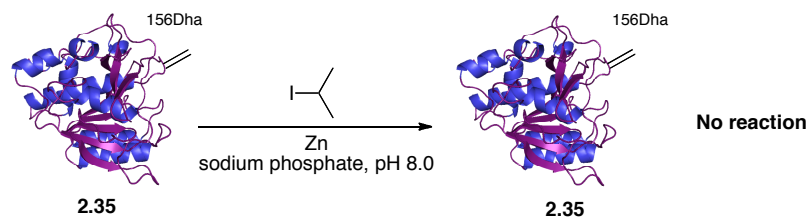


Figure 8.13A, Entry 4

SBL-156Dha (**2.35**), 2-iodopropane, and Zn: no reaction in pH 8.0 sodium phosphate



A 250 μL sample of SBL-Dha156 (**2.35**, 0.29 mg/mL) was prepared in 50 mM sodium phosphate buffer (pH 8.0). Zn powder (2 mg) was added to the protein solution followed by 2-iodopropane (2 μL). The reaction mixture was vortexed and then shaken vigorously at room temperature for 15 minutes. Solids were removed by centrifugation and a 35 μL aliquot was analyzed by LC-MS. Only unreacted **2.35** was observed (calculated mass = 26681, observed mass = 26680). ESI-MS is shown below.

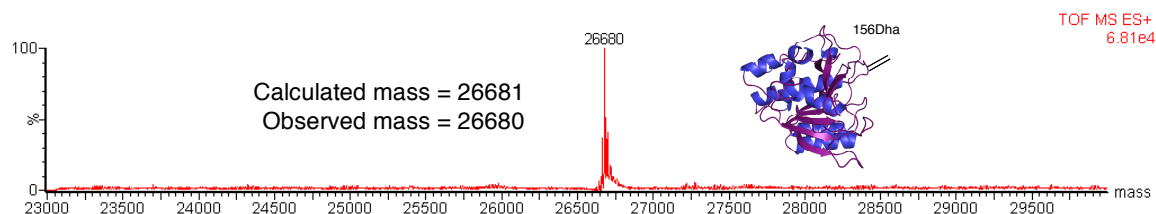
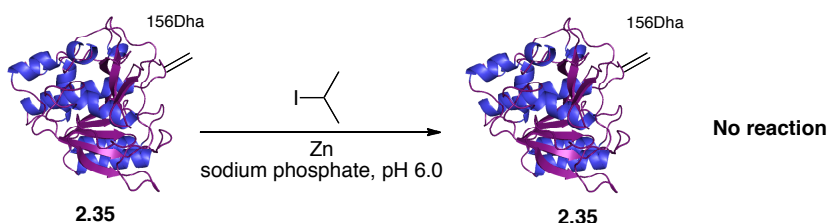


Figure 8.13A, Entry 5

SBL-156Dha (**2.35**), 2-iodopropane, and Zn: no reaction in pH 6.0 sodium phosphate



A 250 μL sample of SBL-Dha156 (**2.35**, 0.29 mg/mL) was prepared in 50 mM sodium phosphate buffer (pH 6.0). Zn powder (2 mg) was added to the protein solution followed by 2-iodopropane (2 μL). The reaction mixture was vortexed and then shaken vigorously at room temperature for 15 minutes. Solids were removed by centrifugation and a 35 μL aliquot was analyzed by LC-MS. Only unreacted **2.35** was observed (calculated mass = 26681, observed mass = 26682). ESI-MS is shown below.

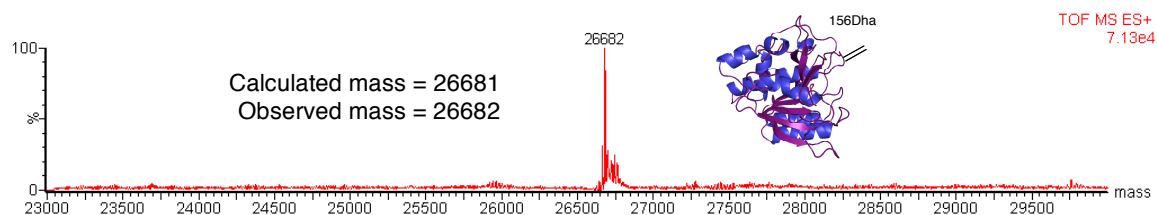
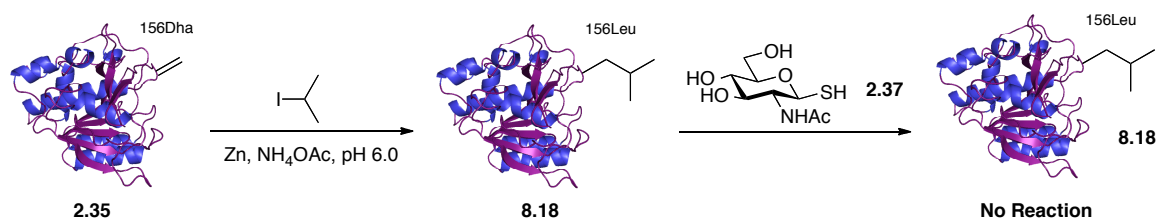
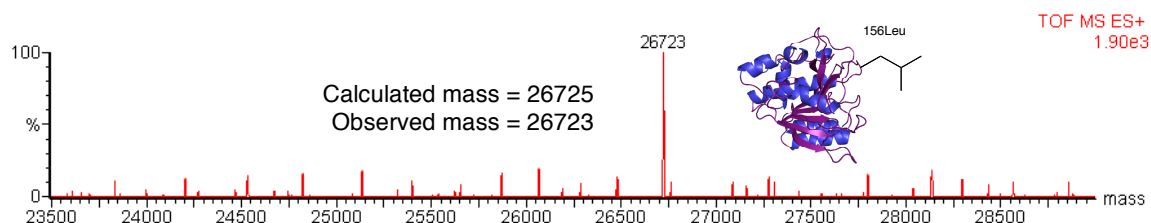


Figure 8.13B

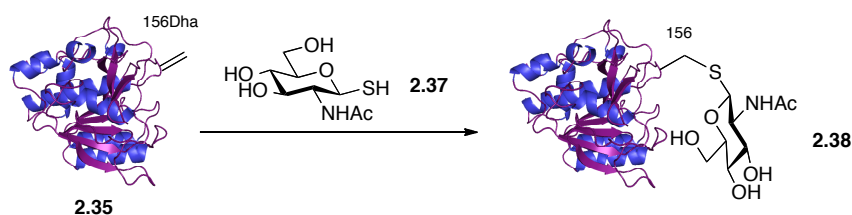
Chemical test for consumption of dehydroalanine



The reaction of SBL-156Dha **2.35** with Zn and 2-iodopropane was repeated on a 500 μ L scale under otherwise identical reaction conditions. After the product was obtained (calculated mass = 26725; observed mass = 26727), the buffer was exchanged by dialysis against 4L of 50 mM sodium phosphate (pH 8.0). After dialysis, the concentration of the protein was measured to be 0.15 mg/mL; protein recovery was 0.12 mg (83%) (BCA assay, Pierce). Next, GlcNAc thiol **2.37** (0.7 mg, 3 μ mol) was added to a 100 μ L aliquot of this protein solution (0.6 nmol protein). The reaction was shaken at room temperature for 90 minutes and then analyzed directly by LC-MS. No thiol addition was observed, consistent with reaction at dehydroalanine during the chemical mutation (calculated mass = 26725; observed mass = 26723). ESI-MS is shown below. Under otherwise identical conditions, thiol **2.37** adds to SBL-156Dha (**2.35**) (see next).



Control: Addition of thiol **2.37** to SBL-156Dha **2.35**



SBL-156Dha (**2.35**) was prepared at a concentration of 0.15 mg/mL in sodium phosphate buffer (pH 8.0). GlcNAc thiol **2.37** (0.7 mg, 3 μ mol) was added to a 100 μ L aliquot of this protein solution (0.6 nmol protein). The reaction was shaken at room temperature for 90 minutes and then analyzed directly by LC-MS. Complete conversion to the glycoprotein was observed (calculated mass = 26918; observed mass = 26918). ESI-MS is shown below.

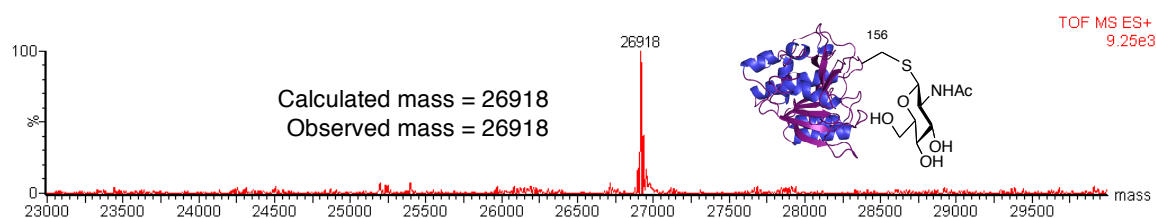
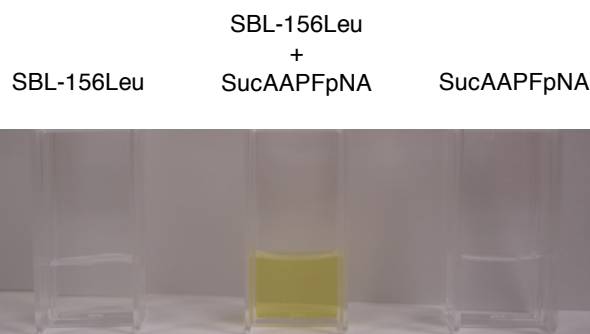


Figure 8.13C

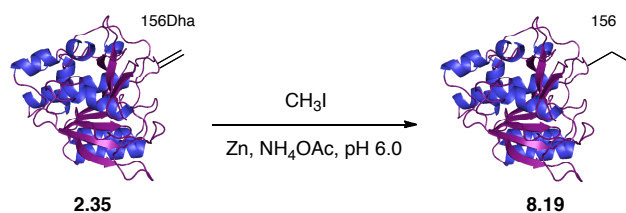
Peptidase activity after chemical mutation of Dha to Leu

The SBL-156Leu mutant prepared above and purified by dialysis against 4L of 50 mM sodium phosphate (pH 8.0) was assayed for peptidase activity. A 167 μ L of this protein solution was diluted to 250 μ L with sodium phosphate buffer (pH 8.0). To this solution was added 25 μ L of a 0.2M solution of the peptide substrate SucAAPFpNA in DMSO. The reaction solution turned bright yellow immediately upon addition of the peptide substrate, indicating liberation of the *p*-nitroanilide after reaction with SBL. A solution of the SBL-156Leu mutant or the peptide substrate alone are clear solutions (see figure below).

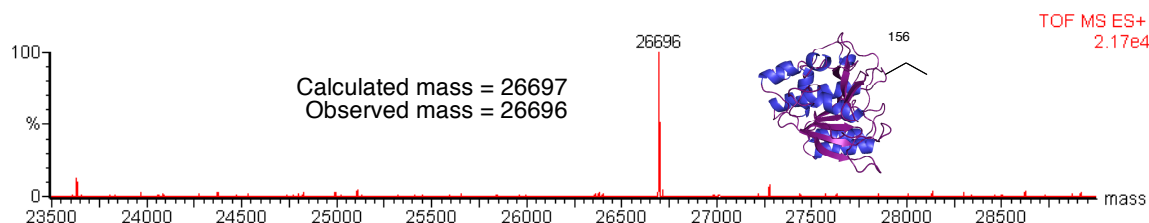


Chemical mutations from Figure 8.14:

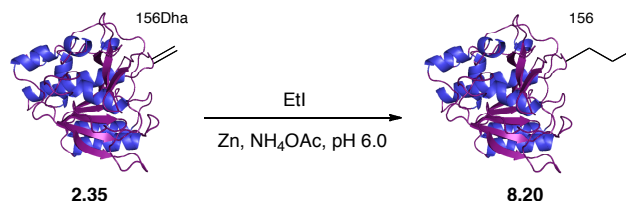
Chemical mutation of dehydroalanine to methylalanine (8.19)



A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Two doses each of zinc powder (4 mg) and methyl iodide (2 μL) were added every five minutes at room temperature. The reaction was shaken vigorously after each addition. The insoluble materials were spun down by centrifugation and the supernatant was analyzed by LC-MS. Approximately 30% conversion was observed so an additional 3 doses of 4 mg of fresh zinc powder and 2 μL of methyl iodide were added to the reaction every 5 minutes. The insoluble materials were removed by centrifugation and the supernatant was analyzed by LC-MS. Full conversion to SBL-156-methylalanine was observed (calculated mass = 26697; observed mass = 26696). ESI-MS is shown below.

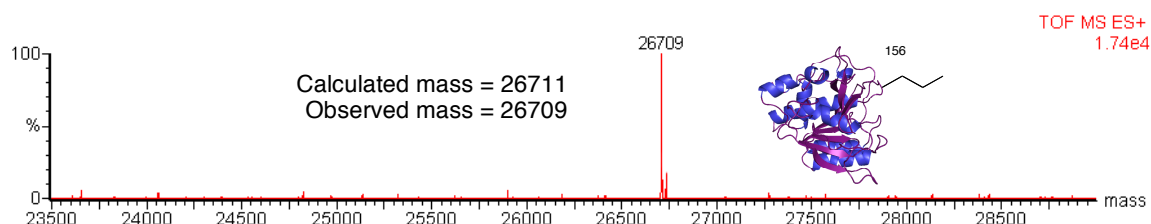


Chemical mutation of dehydroalanine to norvaline (8.20)

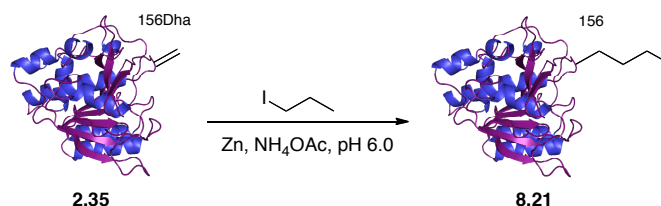


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Zinc powder (2 mg) was added to the protein solution followed by 2 μL of ethyl iodide. The reaction was shaken vigorously at room temperature for 15 minutes. The insoluble materials were spun down by centrifugation and the supernatant was analyzed by LC-MS. Approximately 30% conversion was

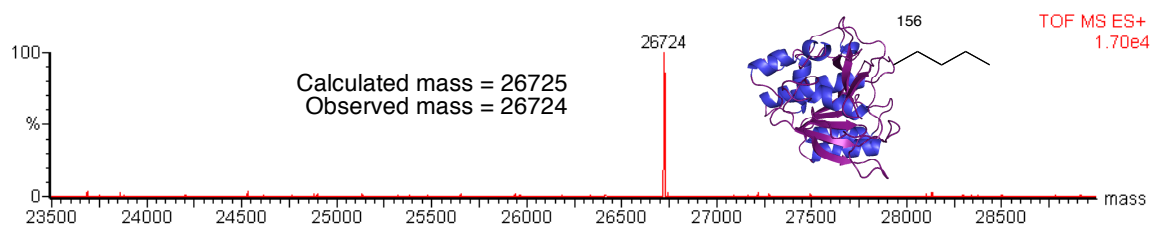
observed so an additional 2 mg of fresh zinc powder and 4 μL of ethyl iodide were added to the reaction. After shaking vigorously for 30 minutes, the reaction was analyzed by LC-MS in the same manner as before, indicating approximately 55% conversion. The reaction was then diluted with 80 μL of pH 6.0 NH_4OAc (500 mM) and 7 mg of zinc powder was added. Ethyl iodide was added in three portions of 2 μL every 10 minutes. Analysis by LC-MS revealed full conversion to SBL-156-norvaline (calculated mass = 26711; observed mass = 26709) ESI-MS is shown below.



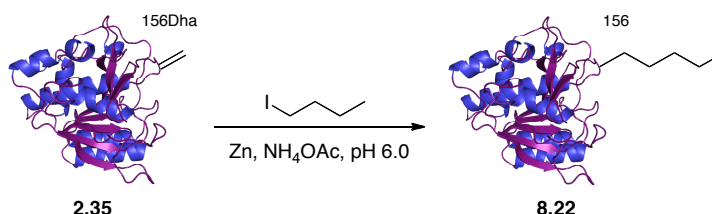
Chemical mutation of dehydroalanine to norleucine (8.21)



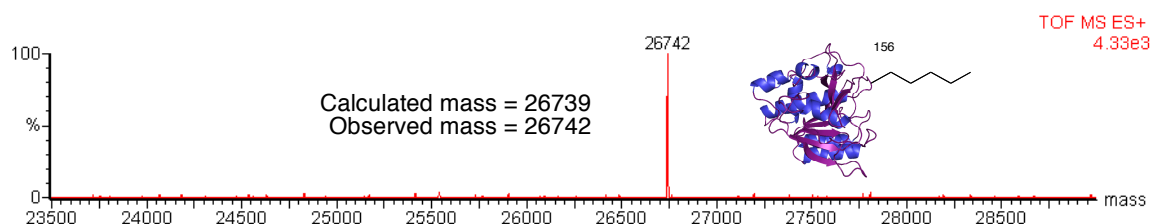
A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (2 mg) and 1-iodopropane (2.0 μL) were added every 10 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. ~50% conversion was observed. The reaction mixture was transferred to a new tube and two additional doses of zinc powder (2 mg) and 1-iodopropane (2.0 μL) were added every 5 minutes. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. ~85% conversion was observed. The reaction mixture was transferred to a new tube along with 100 μL of fresh buffer and three additional doses of zinc powder (2 mg) and 1-iodopropane (2.0 μL) were added every 5 minutes. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to norleucine was observed (calculated mass = 26725; observed mass = 26724). ESI-MS is shown below.



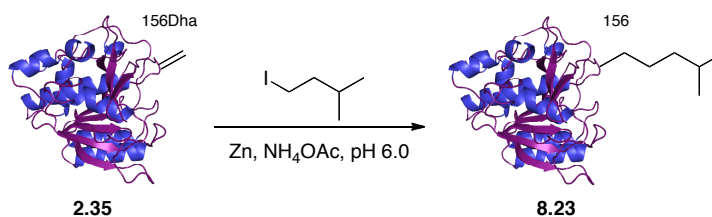
Chemical mutation of dehydroalanine to *n*-butylalanine (8.22)



A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (2 mg) and 1-iodobutane (3.0 μL) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. ~30% conversion was observed. The reaction mixture was transferred to a new tube along with 100 μL of fresh buffer and seven additional doses of zinc powder (2 mg) and 1-iodobutane (2.0 μL) were added every 5 minutes. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to *n*-butylalanine was observed (calculated mass = 26739; observed mass = 26742). ESI-MS is shown below.

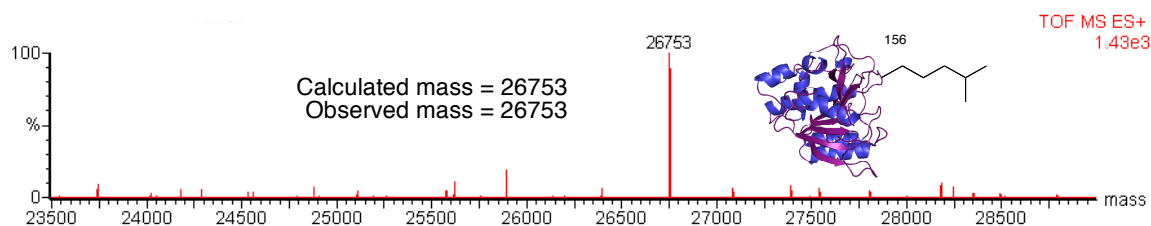


Chemical mutation of dehydroalanine to bishomoleucine (8.23)

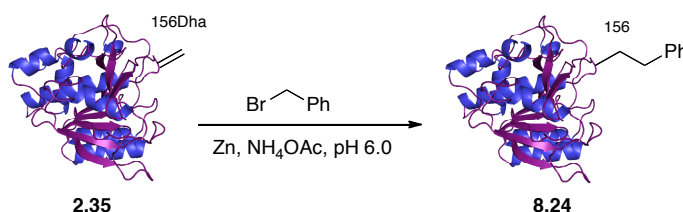


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and

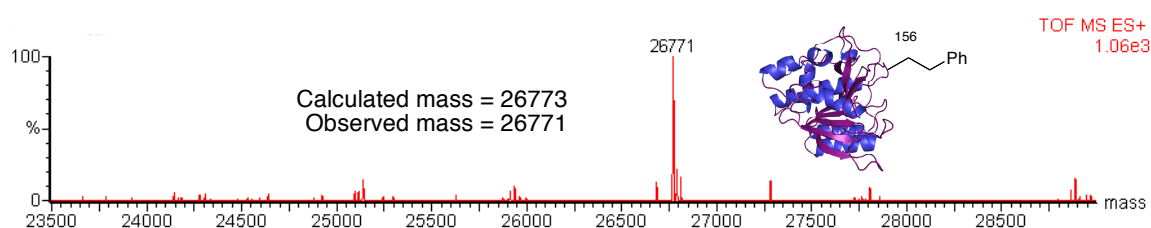
3-methyl-1-iodobutane (2.0 μL) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and the supernatant transferred to a new tube along with fresh buffer (40 μL) and AcOH (2 μL). These round of three doses were repeated 5 additional times. A final centrifugation and analysis of the supernatant revealed full conversion to bishomoleucine (calculated mass = 26753; observed mass = 26753). ESI-MS is shown below.



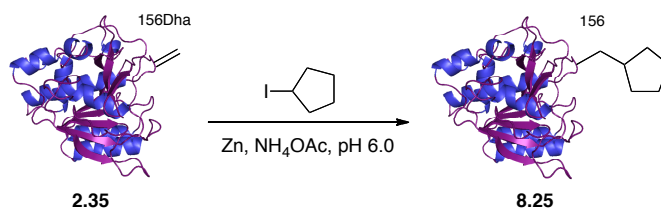
Chemical mutation of dehydroalanine to homophenylalanine (8.24)



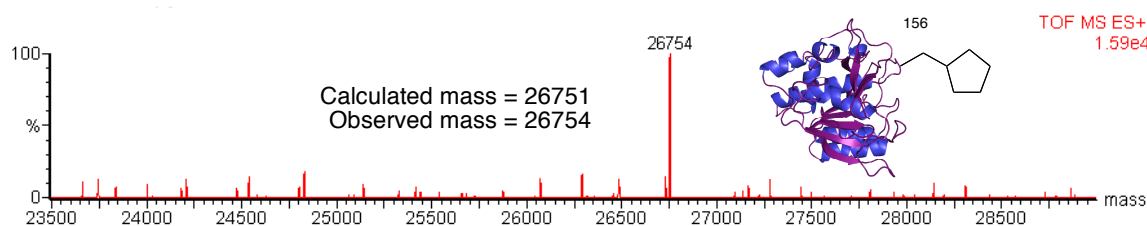
A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (4 mg) and benzylbromide (2 μL) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Five minutes after the third dose, 100 μL of additional buffer was added followed by two doses each of zinc powder (4 mg) and benzylbromide (2 μL), every 5 minutes as before. Solids were again removed by centrifugation and the supernatant transferred to another tube. A final five doses of zinc (4 mg) and benzylbromide (2 μL) were added every 5 minutes to complete the reaction. Solids were again removed by centrifugation and the supernatant analyzed by LC-MS, revealing complete conversion to homophenylalanine (calculated Mass = 26773; observed mass = 26771). ESI-MS is shown below.



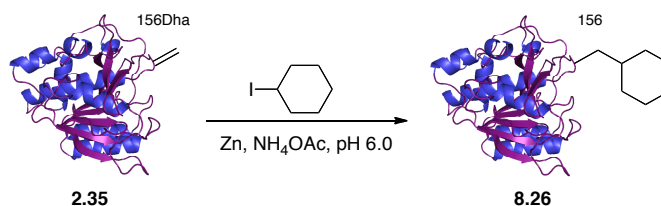
Chemical mutation of dehydroalanine to cyclopentylalanine (8.25)



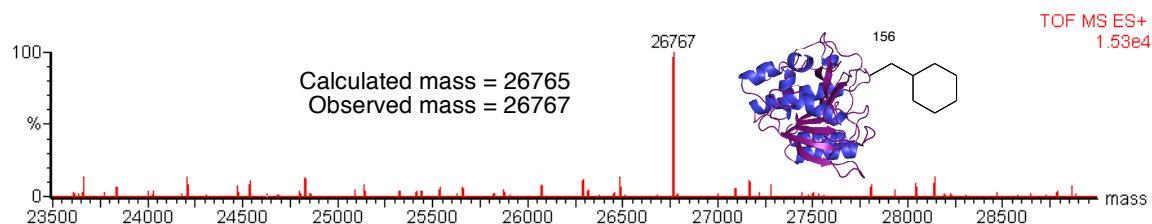
A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Zinc powder (2 mg) was added followed by three doses of iodocyclopentane (2.5 μL every 5 minutes at room temperature). The reaction was vortexed between each addition. An additional two doses of both 5 mg of Zn powder and 2.5 μL of iodocyclopentane was then added in 5 minute intervals. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to cyclopentylalanine was observed (calculated mass = 26751, observed mass = 26754). ESI-MS is shown below.



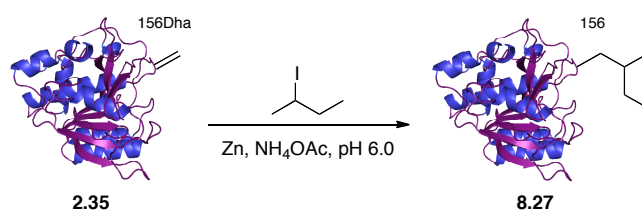
Chemical mutation of dehydroalanine to cyclohexylalanine (8.26)



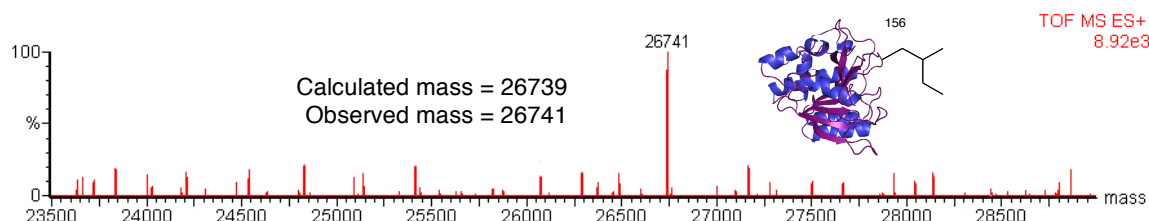
A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Zinc powder (2 mg) was added followed by iodocyclohexane (2.8 μL). The reaction was shaken vigorously at room temperature for 15 minutes and then every 5 minutes an additional portion of Zn (2 mg) and iodocyclohexane (2.8 μL) was added. (5 total doses of Zn and iodocyclohexane were used). Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to cyclohexylalanine was observed (calculated mass = 26765; observed mass = 26767). ESI-MS is shown below.



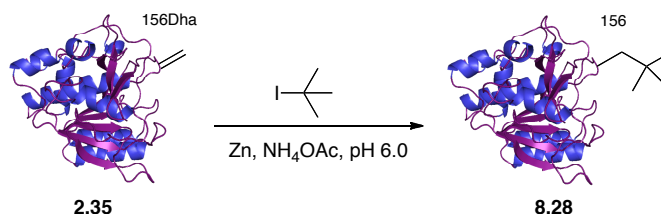
Chemical mutation of dehydroalanine to *sec*-butylalanine (8.27)



A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (2 mg) and 2-iodobutane (3.0 μL) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to *sec*-butylalanine was observed (calculated mass = 26739; observed mass = 26741). ESI-MS is shown below.

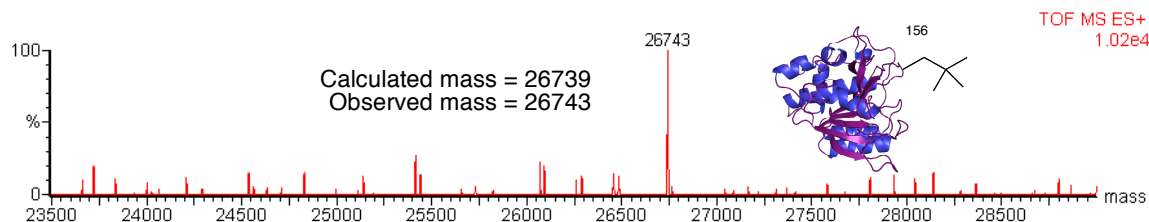


Chemical mutation of dehydroalanine to *tert*-butylalanine (8.28)

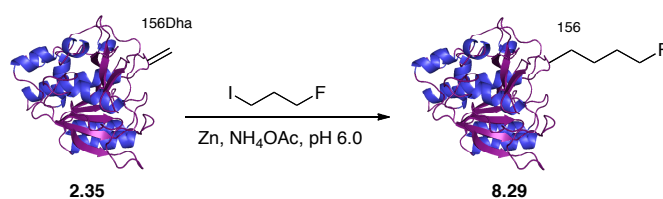


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (2 mg) and *tert*-butyl iodide (3.0 μL) were added every 5 minutes at room temperature. The reaction

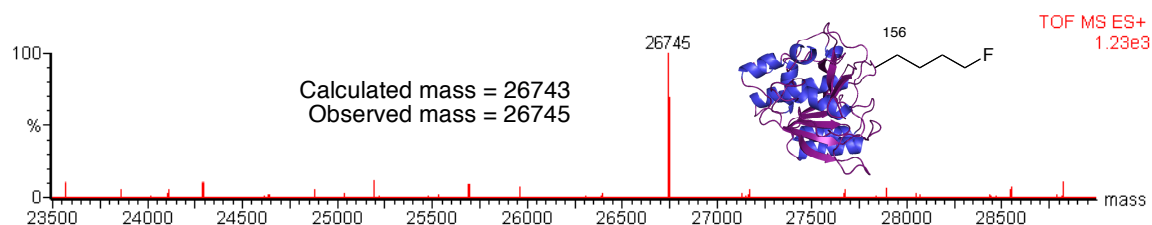
was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to *tert*-butylalanine was observed (calculated mass = 26739; observed mass = 26743). ESI-MS is shown below.



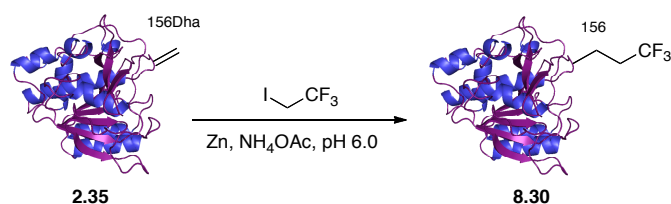
Chemical mutation of dehydroalanine to ϵ -fluoronorleucine (8.29)



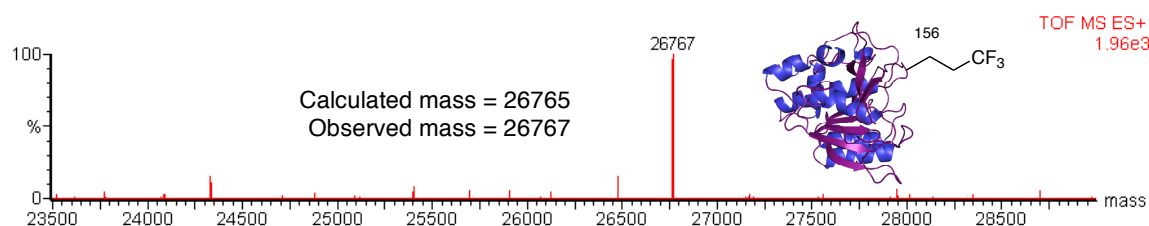
A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and 1-iodo-3-fluoropropane (2 μL) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Three minutes after the third dose, solids were removed by centrifugation and the supernatant transferred to a new tube along with 40 μL of additional buffer and 2 μL AcOH. Three additional doses of zinc powder (3 mg) and 1-iodo-3-fluoropropane (2 μL) were added every 3 minutes at room temperature as before. Three minutes after the final dose, solids were again removed by centrifugation and the supernatant analyzed by LC-MS. Full conversion to ϵ -fluoronorleucine was observed (calculated mass = 26743; observed mass = 26745). ESI-MS is shown below.



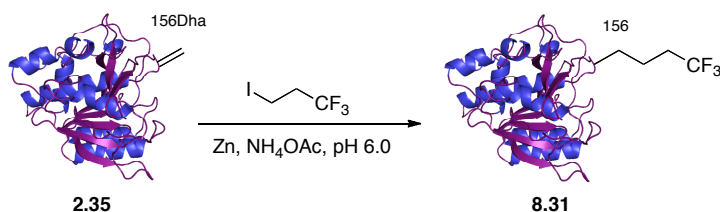
Chemical mutation of dehydroalanine to δ,δ,δ -trifluoronorvaline (8.30)



A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and 1-iodo-2,2,2-trifluoroethane (2 μL) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Three minutes after the third dose, solids were removed by centrifugation and the supernatant transferred to a new tube along with 40 μL of additional buffer and 2 μL AcOH. Three additional doses of zinc powder (3 mg) and 1-iodo-2,2,2-trifluoroethane (2 μL) were added every 3 minutes at room temperature as before. Three minutes after the final dose, solids were again removed by centrifugation and the supernatant analyzed by LC-MS. Full conversion to trifluoronorvaline was observed (calculated mass = 26765; observed mass = 26767). ESI-MS is shown below.

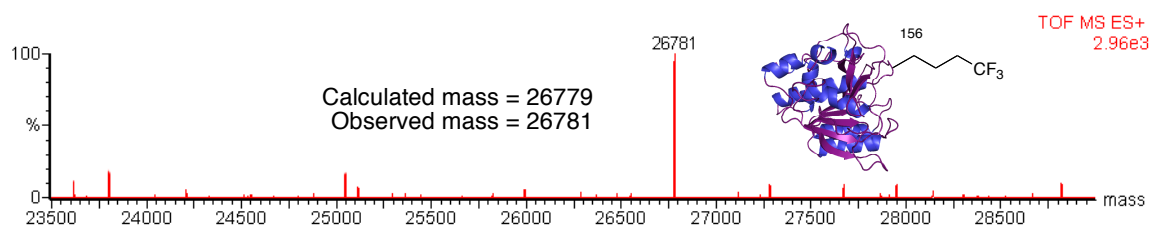


Chemical mutation of dehydroalanine to $\epsilon,\epsilon,\epsilon$ -trifluoronorleucine (8.31)

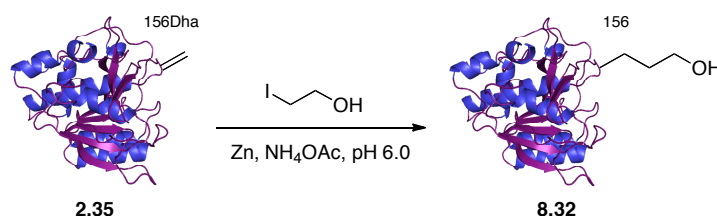


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and 1-iodo-3,3,3-trifluoropropane (2 μL) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Three minutes after the third dose, solids were removed by centrifugation and the supernatant transferred to a new tube along with 40 μL of additional buffer and 2 μL AcOH. Three additional doses

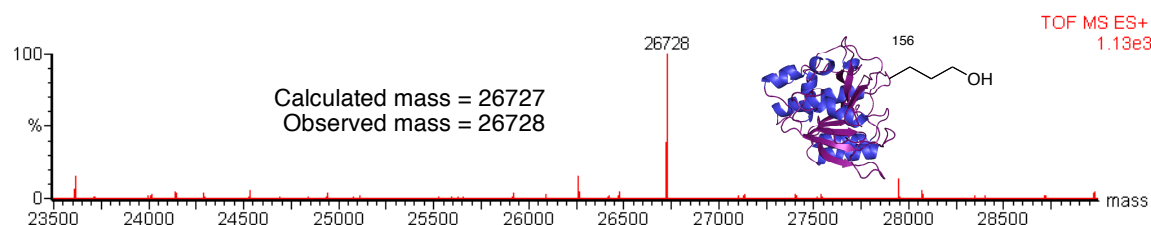
of zinc powder (3 mg) and 1-iodo-3,3,3-trifluoropropane (2 μ L) were added every 3 minutes at room temperature as before. Three minutes after the final dose, solids were again removed by centrifugation and the supernatant analyzed by LC-MS. Full conversion to trifluoronorleucine was observed (calculated mass = 26779; observed mass = 26781). ESI-MS is shown below.



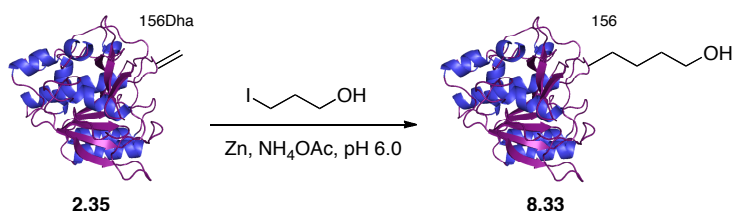
Chemical mutation of dehydroalanine to *bis*-homoserine (8.32)



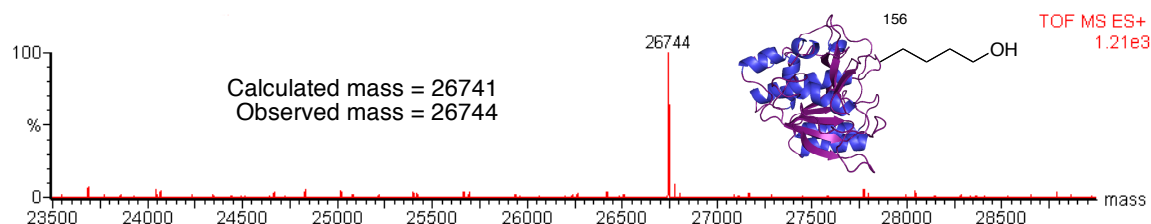
A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and 2-iodoethanol (2.0 μ L) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and the supernatant transferred to a new tube along with fresh buffer (40 μ L) and AcOH (2 μ L) to maintain a pH of 6.0. These round of three doses were repeated 2 additional times. A final centrifugation and analysis of the supernatant revealed full conversion to *bis*-homoserine (calculated mass = 26727; observed mass = 26728). ESI-MS is shown below.



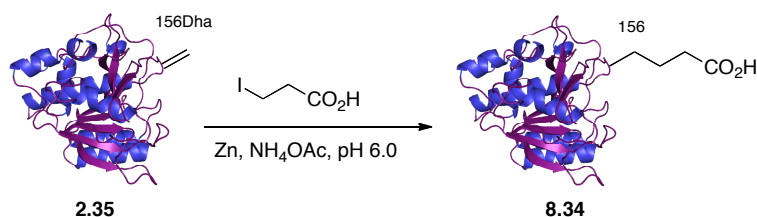
Chemical mutation of dehydroalanine to ϵ -hydroxynorleucine (8.33)



A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and 3-iodopropanol (1.0 μL) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and the supernatant transferred to a new tube along with fresh buffer (20 μL) and AcOH (1 μL) to maintain a pH of 6.0. This round of three doses was then repeated once more. A final centrifugation and analysis of the supernatant revealed full conversion to ϵ -hydroxynorleucine (calculated mass = 26741; observed mass = 26744). ESI-MS is shown below.

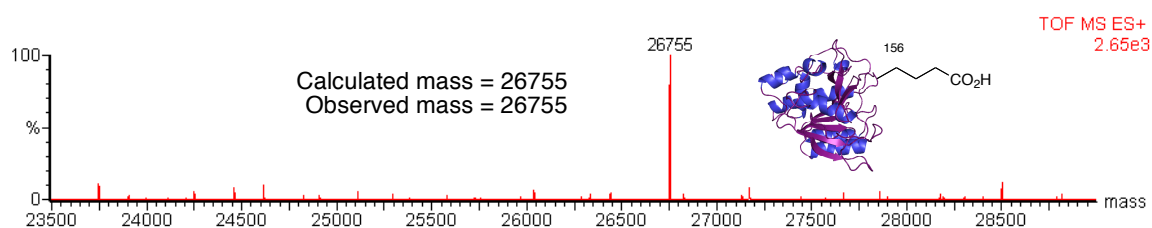


Chemical mutation of dehydroalanine to homoglutamic acid (8.34)

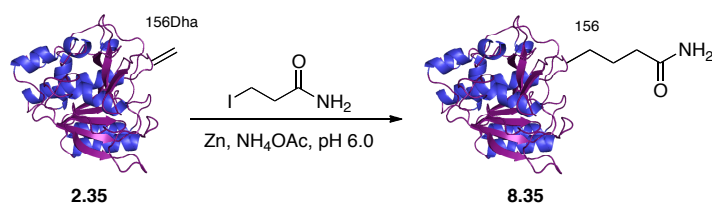


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Four doses each of zinc powder (3 mg) and 3-iodopropionic acid (2 mg; recrystallized from CHCl_3 /Petrol before use) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and the supernatant transferred to a new tube along with 80 μL of additional buffer. Four additional doses of zinc powder (3 mg) and 3-iodopropionic acid (2 mg; recrystallized from CHCl_3 /Petrol before use) were added every 3 minutes at room temperature to complete the reaction.

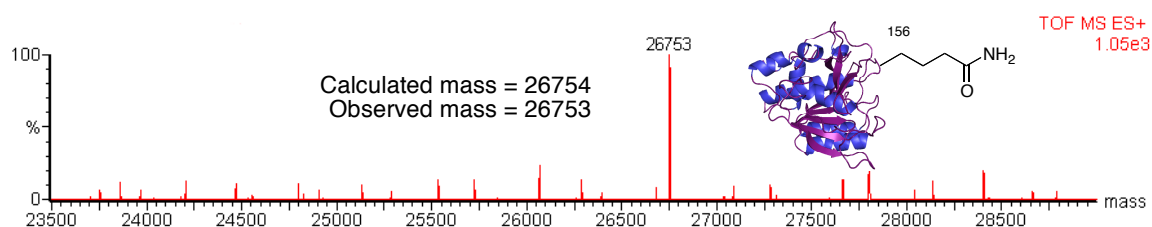
Solids were again removed by centrifugation and the supernatant analyzed by LC-MS. Full conversion to homoglutamic acid was observed (calculated mass = 26755; observed mass = 26755). ESI-MS is shown below.



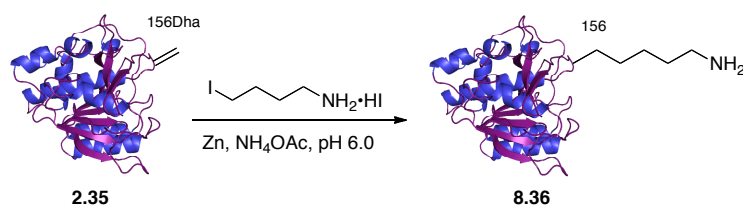
Chemical mutation of dehydroalanine to homoglutamine (8.35)



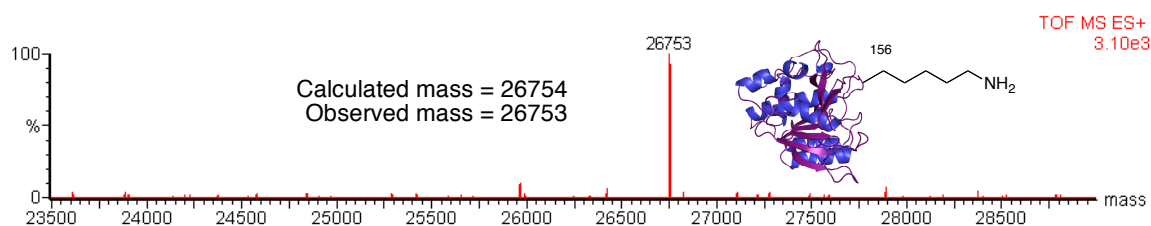
A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (4 mg) and 3-iodopropionamide (2 mg) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and the supernatant transferred to a new tube along with 80 μ L of additional buffer. Three additional doses of zinc powder (4 mg) and 3-iodopropionamide (2 mg) were added every 3 minutes at room temperature. Solids were again removed by centrifugation and the supernatant transferred to another tube along with 80 μ L fresh buffer. A final three doses of zinc (4 mg) and 3-iodopropionamide (2 mg) were added every 3 minutes. Solids were again removed by centrifugation and the supernatant analyzed by LC-MS, revealing complete conversion to homoglutamine (calculated mass = 26754; observed mass = 26753). ESI-MS is shown below.



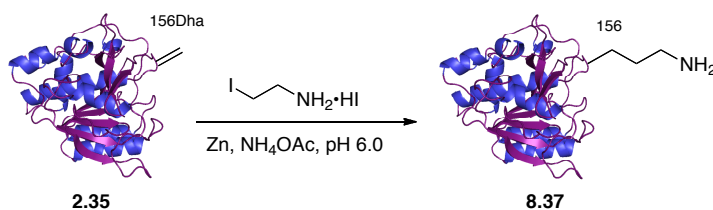
Chemical mutation of dehydroalanine to homolysine (8.36)



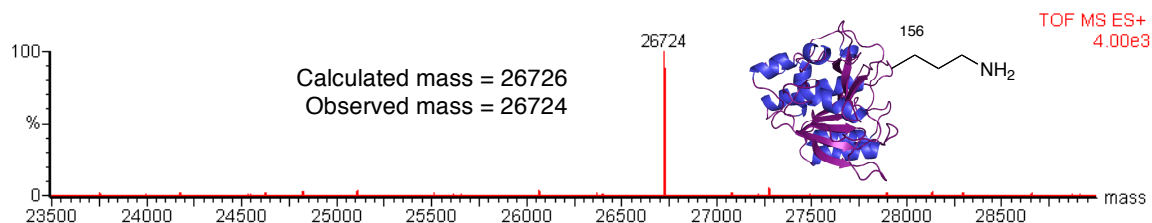
A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH₄OAc, 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and (4-iodobutyl)amine hydroiodide (3 mg) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. An additional 100 μ L of fresh buffer was then added to the reaction mixture followed by two additional doses of zinc powder (3 mg) and (4-iodobutyl)amine hydroiodide (3 mg) every five minutes. Solids were removed by centrifugation and 40 μ L were analyzed by LC-MS. Full conversion to homolysine was observed (calculated mass = 26754; observed mass = 26753).



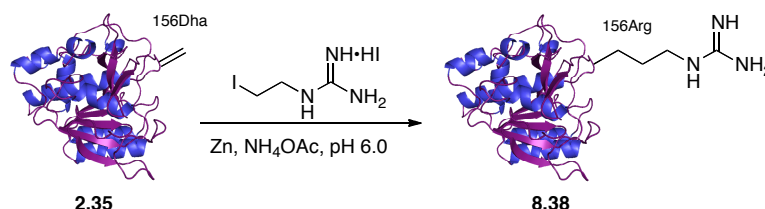
Chemical mutation of dehydroalanine to ornithine (8.37)



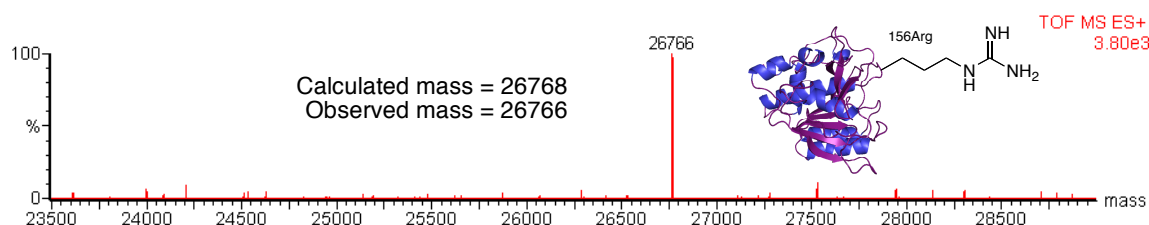
A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH₄OAc, 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and (2-iodoethyl)amine hydroiodide (3 mg) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. An additional 100 μ L of fresh buffer was then added to the reaction mixture followed by two additional doses of zinc powder (3 mg) and (2-iodoethyl)amine hydroiodide (3 mg) every five minutes. Solids were removed by centrifugation and 40 μ L were analyzed by LC-MS. Full conversion to ornithine was observed (calculated mass = 26726; observed mass = 26724). ESI-MS is shown below.



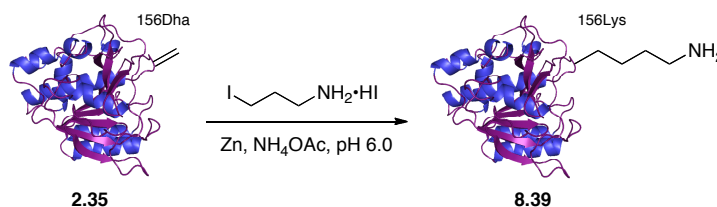
Chemical mutation of dehydroalanine to arginine (8.38)



A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and (2-iodoethyl)guanidine hydroiodide (2 μL) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. An additional 100 μL of fresh buffer was then added to the reaction mixture followed by two additional doses of zinc powder (3 mg) and (2-iodoethyl)guanidine hydroiodide (2 μL) every five minutes. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to arginine was observed (calculated mass = 26768; observed mass = 26766). ESI-MS is shown below.

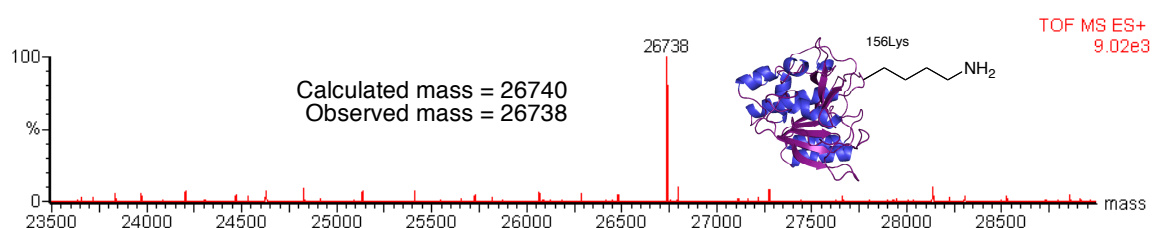


Chemical mutation of dehydroalanine to lysine (8.39)

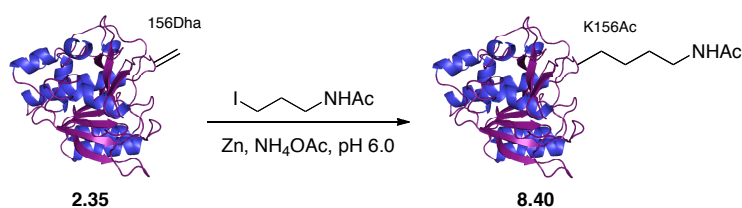


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. A solution of 3-iodopropylamine hydroiodide was prepared in a separate tube by dissolving 20 mg in 100 μL of the same

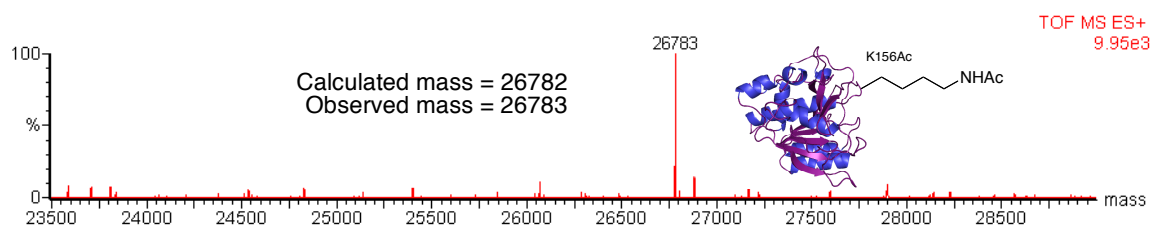
buffer. Five doses each of zinc powder (2 mg) and the 3-iodopropylamine hydroiodide solution (20 μ L) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation five minutes after the final addition. LC-MS analysis of the supernatant revealed full conversion to lysine (calculated mass = 26740; observed mass = 26738). ESI-MS is shown below.



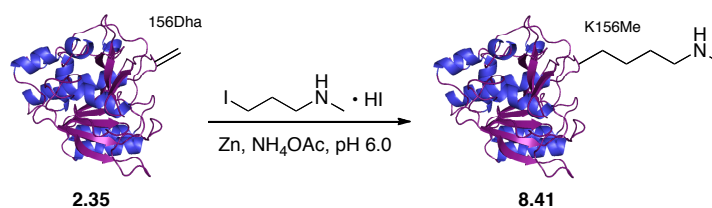
Chemical mutation of dehydroalanine to *N*-acetyllysine (8.40)



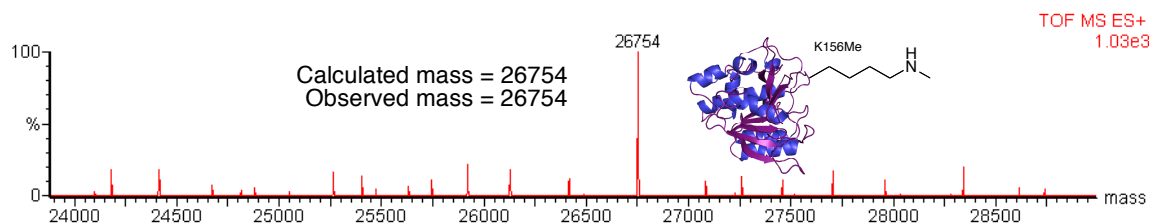
A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH₄OAc, 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Four doses each of zinc powder (3 mg) and 1-iodo-3-acetamidopropane (2.0 μ L) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. An additional 100 μ L of fresh buffer was then added to the reaction mixture followed by a final dose of zinc powder (3 mg) and 1-iodo-3-acetamidopropane (2.0 μ L). Solids were removed by centrifugation and 40 μ L were analyzed by LC-MS. Full conversion to *N*-acetyllysine was observed (calculated mass = 26782; observed mass = 26783). ESI-MS is shown below.



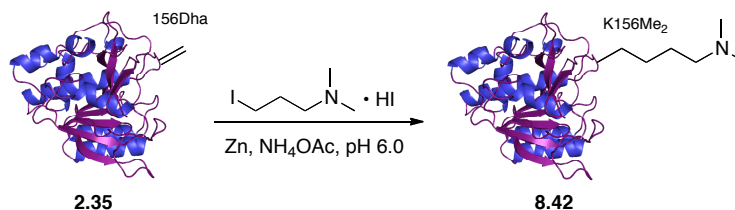
Chemical mutation of dehydroalanine to monomethyllysine (8.41)



A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH₄OAc, 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (4 mg) and 1-iodo-3-methylaminopropane hydroiodide (2 mg) were added to the protein solution every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Five minutes after the third addition, solids were removed by centrifugation and transferred to a new vial along with 50 μ L fresh buffer and 2 μ L AcOH. An additional three doses each of zinc powder (4 mg) and 1-iodo-3-methylaminopropane hydroiodide (2 mg) were added to the protein solution every 5 minutes at room temperature as before. Five minutes after the third addition, solids were removed by centrifugation and the supernatant analyzed by LC-MS. Full conversion to monomethyllysine was observed (calculated mass = 26754; observed mass = 26754). ESI-MS is shown below.

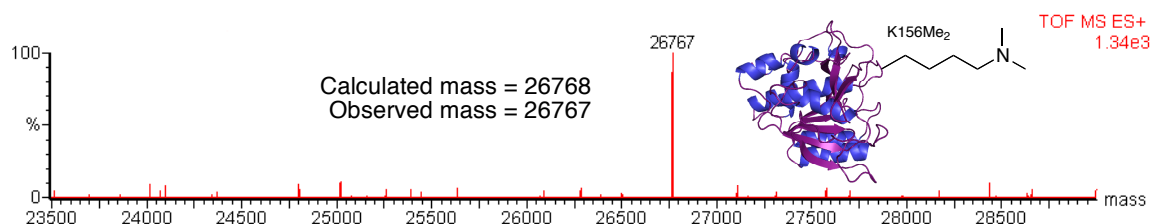


Chemical mutation of dehydroalanine to dimethyllysine (8.42)

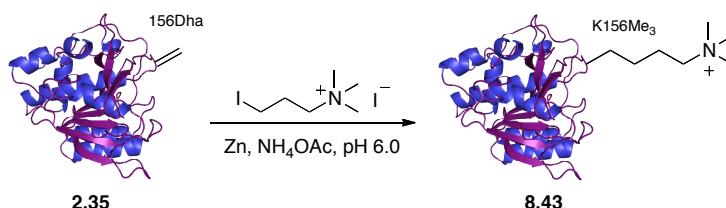


A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH₄OAc, 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Five doses each of zinc powder (2 mg) and (3-iodopropyl)-dimethylamine hydroiodide (2 mg) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. After the fifth addition, AcOH (2 μ L) was added and then solids were removed by

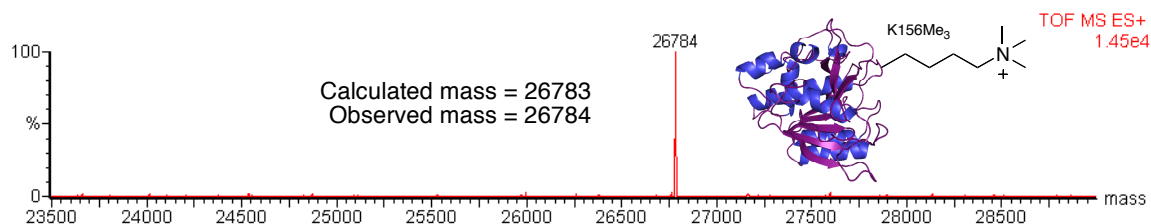
centrifugation. The supernatant was transferred to a new tube and an additional 100 μL of fresh buffer was added along with 1 μL of AcOH. Three additional doses each of zinc powder (2 mg) and (3-iodopropyl)-dimethylamine hydroiodide (2 mg) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Solids were removed by centrifugation and the supernatant analyzed by LC-MS. Full conversion to dimethyllysine was observed (calculated mass = 26768; observed mass = 26767). ESI-MS is shown below.



Chemical mutation of dehydroalanine to trimethyllysine (8.43)

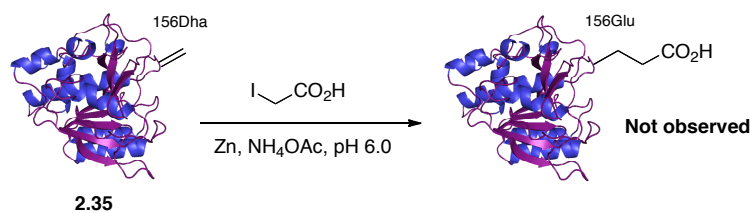


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (3 mg) and (3-iodopropyl)-trimethylammonium iodide (3 mg) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. An additional 100 μL of fresh buffer was then added to the reaction mixture followed by two additional doses of zinc powder (3 mg) and (3-iodopropyl)-trimethylammonium iodide (3 mg) every five minutes. Solids were removed by centrifugation and 40 μL were analyzed by LC-MS. Full conversion to trimethyllysine was observed (calculated mass = 26783; observed mass = 26784). ESI-MS is shown below.

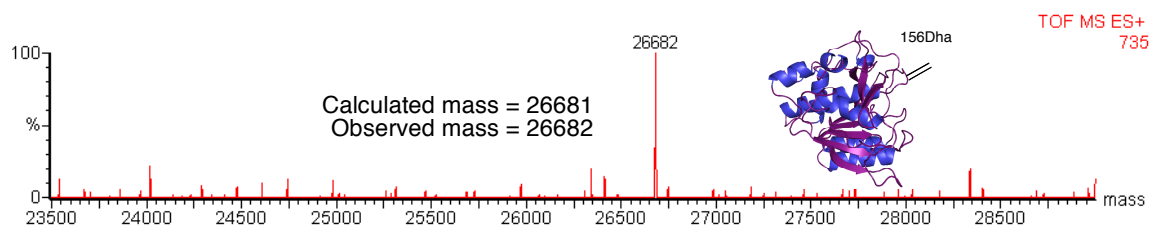


Unsuccessful chemical mutations:

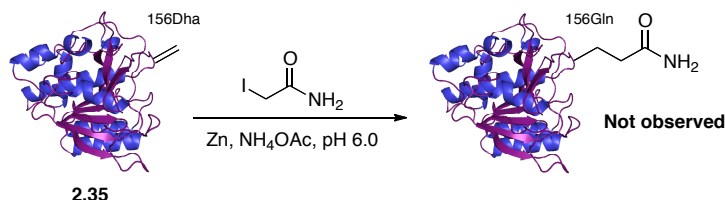
Attempted chemical mutation of dehydroalanine to glutamic acid



A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (4 mg) and iodoacetic acid (2 mg) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Five minutes after the third dose, solids were removed by centrifugation and the supernatant transferred to a new tube along with 80 μL fresh buffer. Three additional doses each of zinc powder (4 mg) and iodoacetic acid (2 mg) were added every 5 minutes at room temperature as before. Five minutes after the final dose, solids were removed by centrifugation and the supernatant analyzed by LC-MS. Only unreacted SBL-156Dha (**2.35**) was observed (calculated mass = 26681; observed mass = 26682). ESI-MS is shown below.

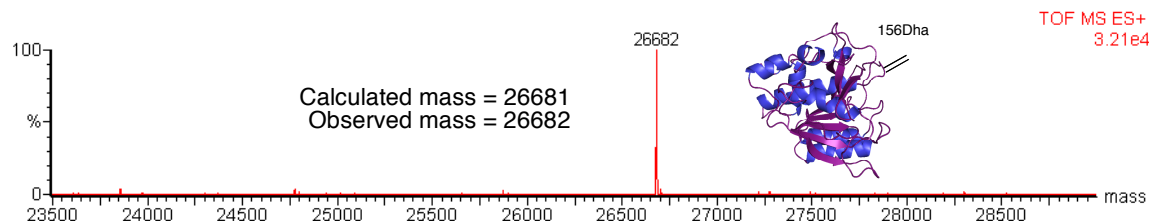


Attempted chemical mutation of dehydroalanine to glutamine

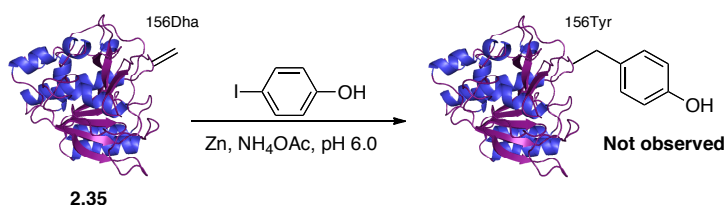


A 250 μL aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH_4OAc , 500 mM buffer) was added to a 1.5 mL plastic reaction tube. Three doses each of zinc powder (4 mg) and iodoacetamide (3 mg) were added every 5 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Five minutes after the third dose, an additional 100 μL of fresh buffer was added before two more doses each of zinc powder

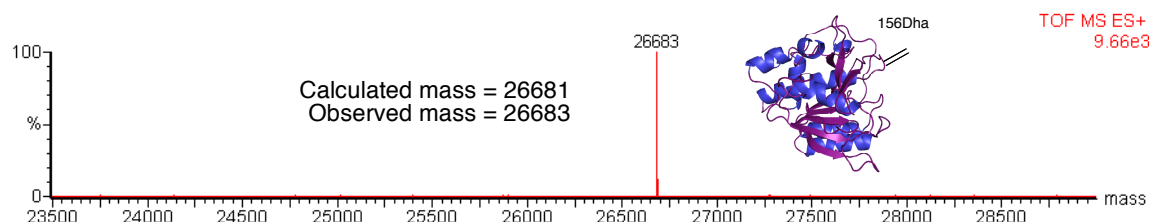
(4 mg) and iodoacetamide (3 mg) were added every 5 minutes at room temperature as before. Five minutes after the final dose, solids were removed by centrifugation and the supernatant analyzed by LC-MS. Only unreacted SBL-156Dha was observed (calculated mass = 26681; observed mass = 26682). ESI-MS is shown below.



Attempted chemical mutation of dehydroalanine to tyrosine



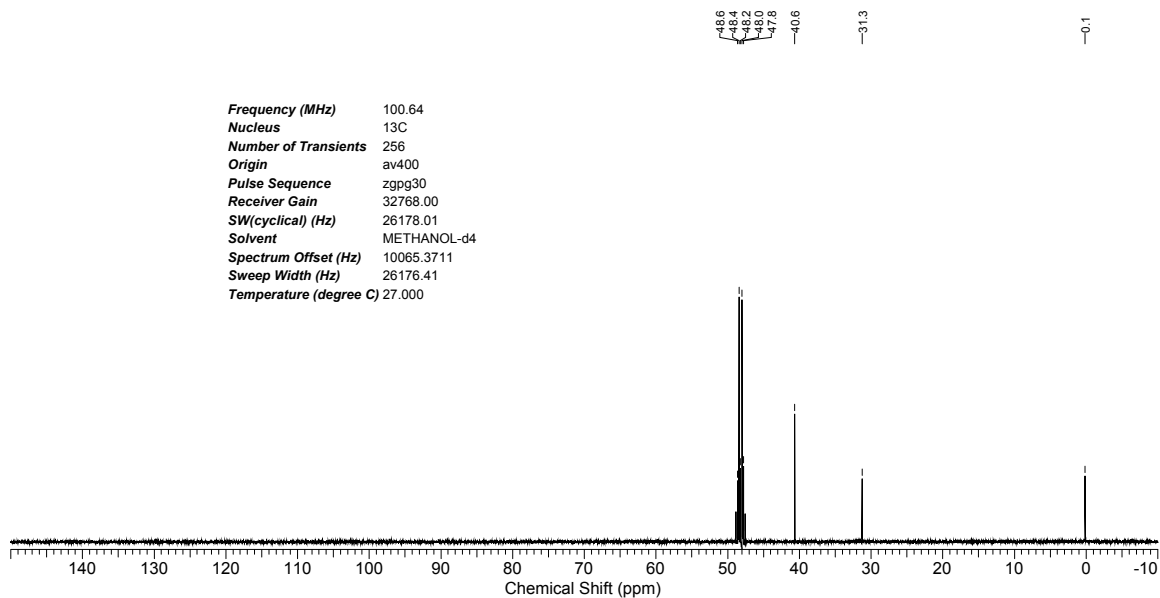
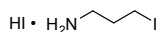
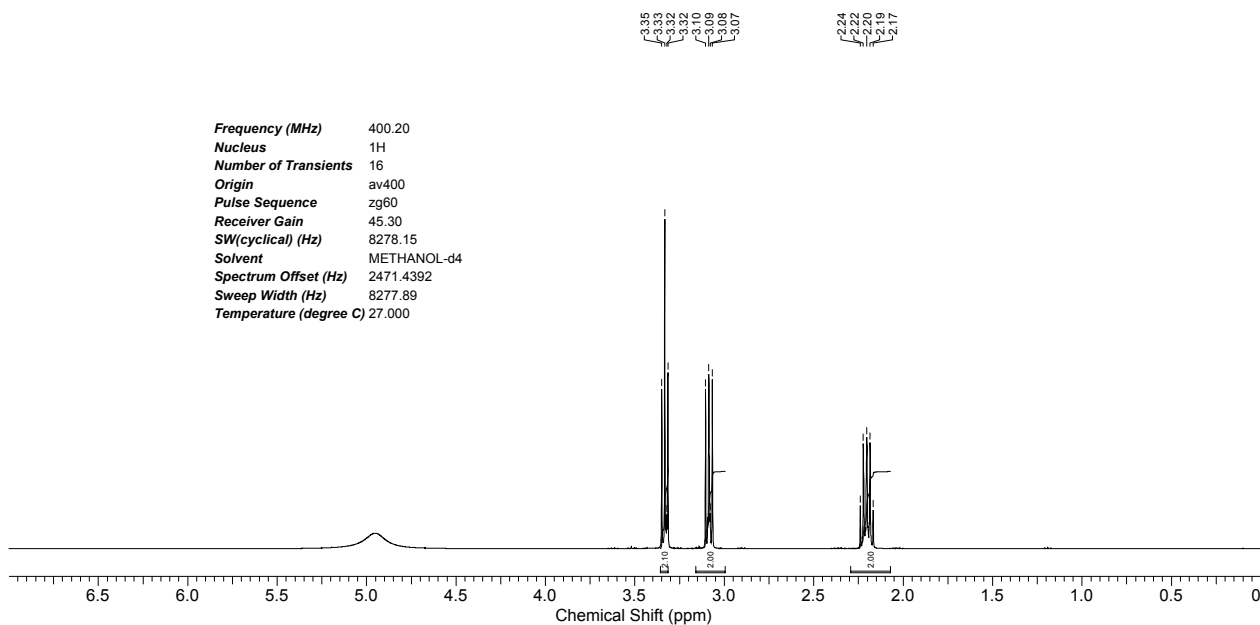
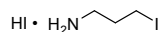
A 250 μ L aliquot of SBL-156Dha **2.35** (0.29 mg/mL in pH 6.0 NH₄OAc, 500 mM buffer) was added to a 1.5 mL plastic reaction tube. A solution of 4-iodophenol was prepared by dissolving 10 mg in 20 μ L of THF. Five doses each of zinc powder (3 mg) and the 4-iodophenol solution (4 μ L) were added every 3 minutes at room temperature. The reaction was vortexed and shaken vigorously after each addition. Three minutes after the fifth dose, solids were removed by centrifugation and the supernatant analyzed by LC-MS. Only unreacted SBL-156Dha was observed (calculated mass = 26681; observed mass = 26683). ESI-MS is shown below.

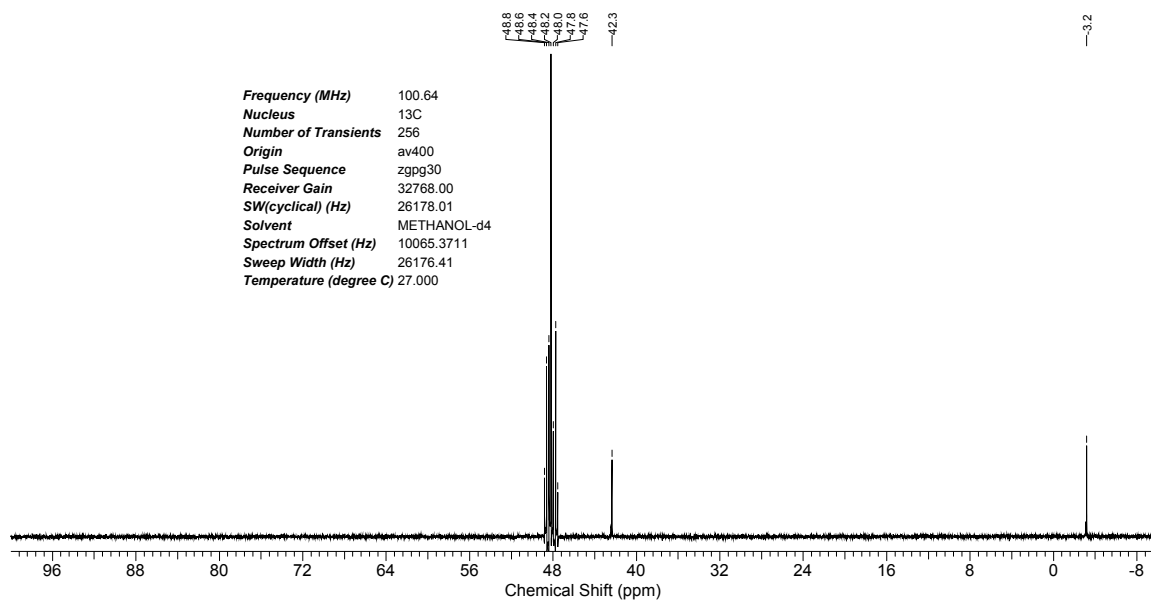
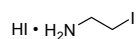
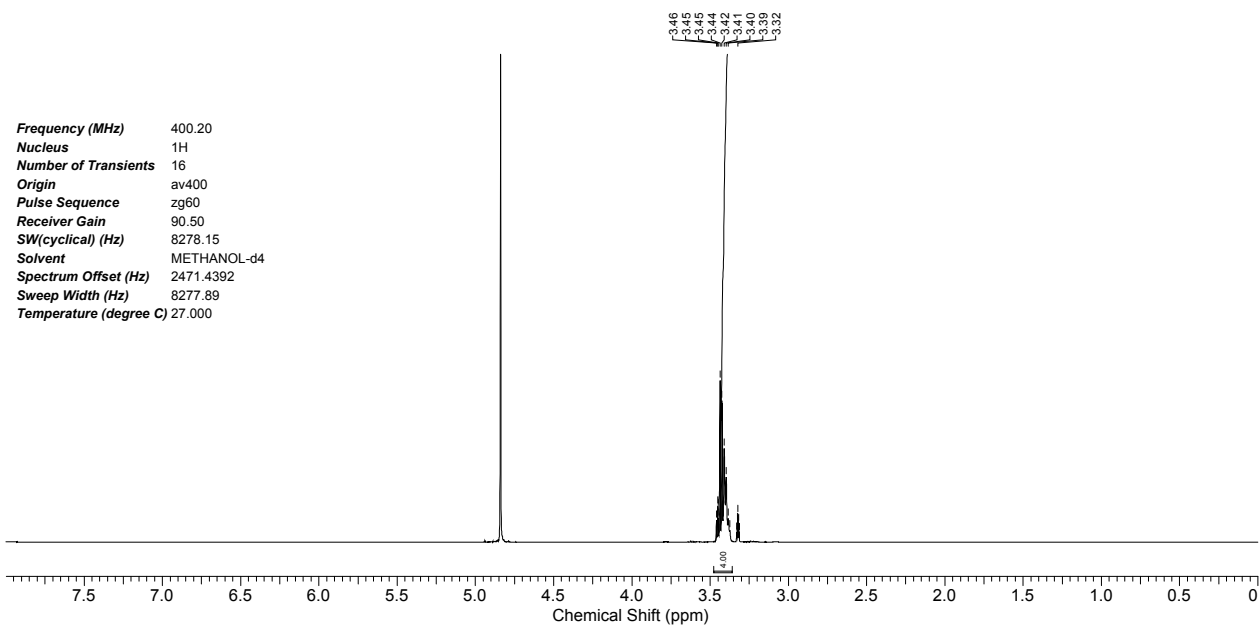
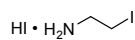


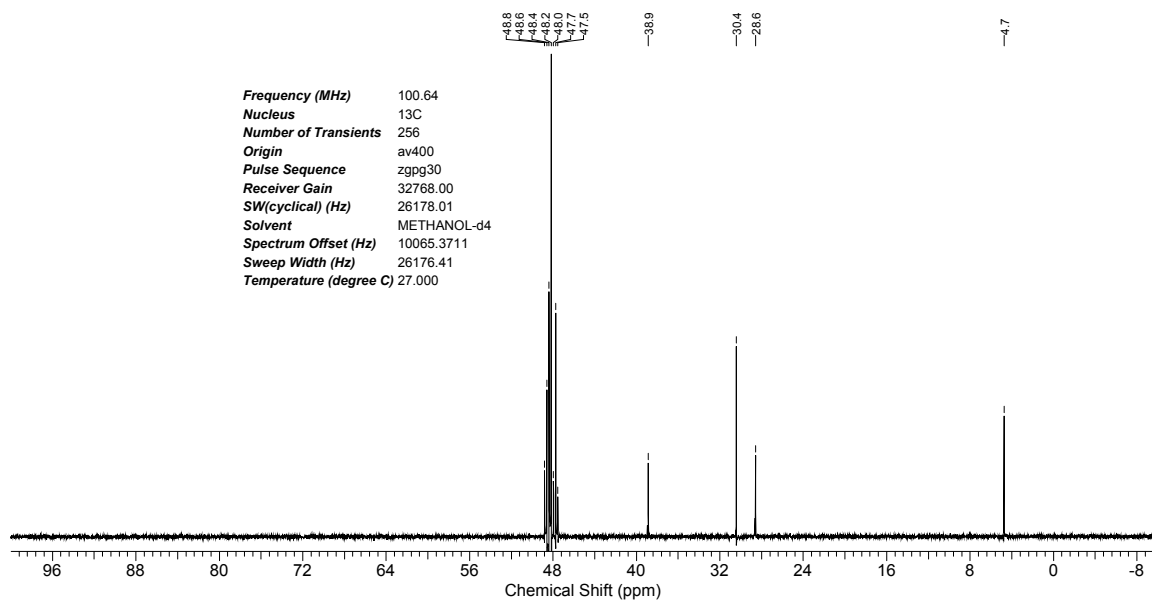
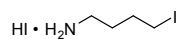
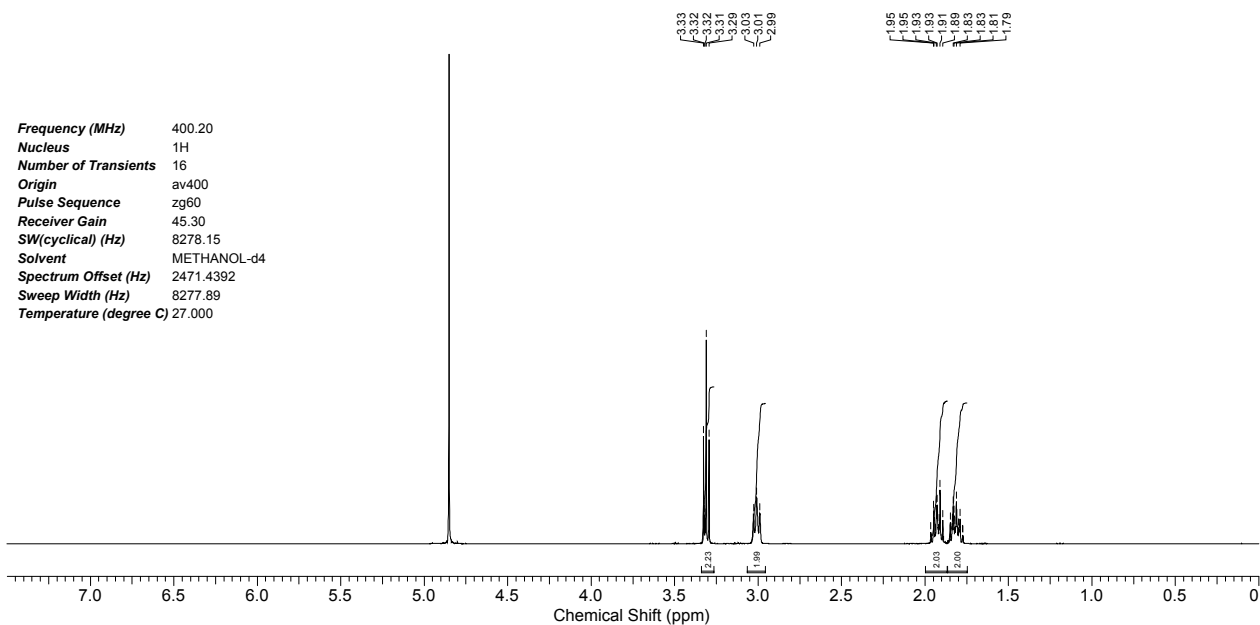
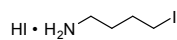
8.6 References for Experimental Procedures

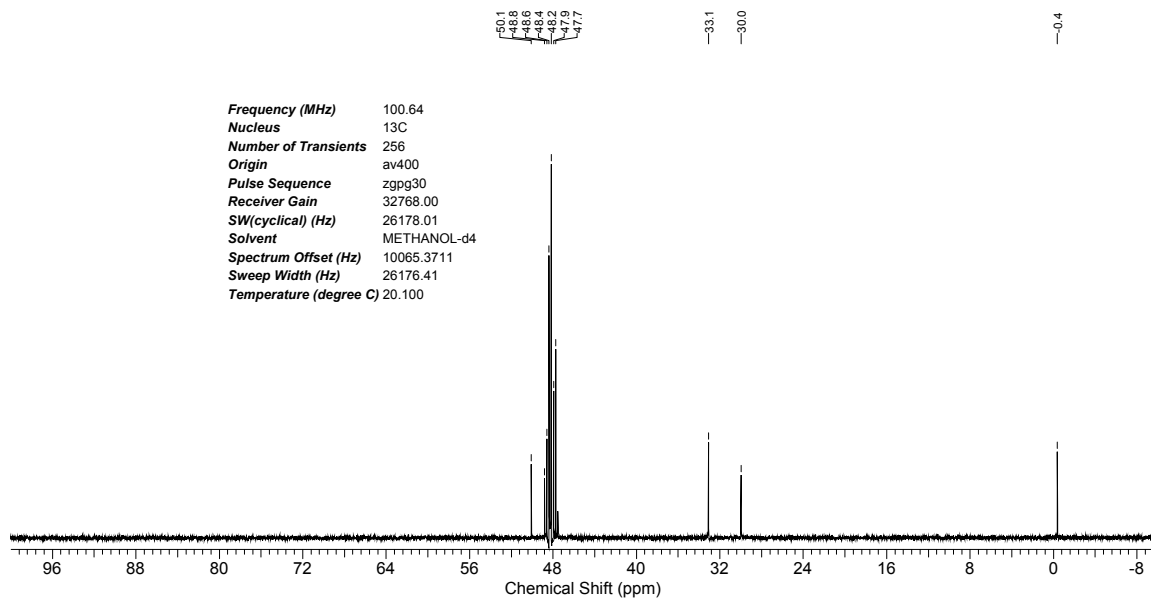
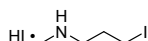
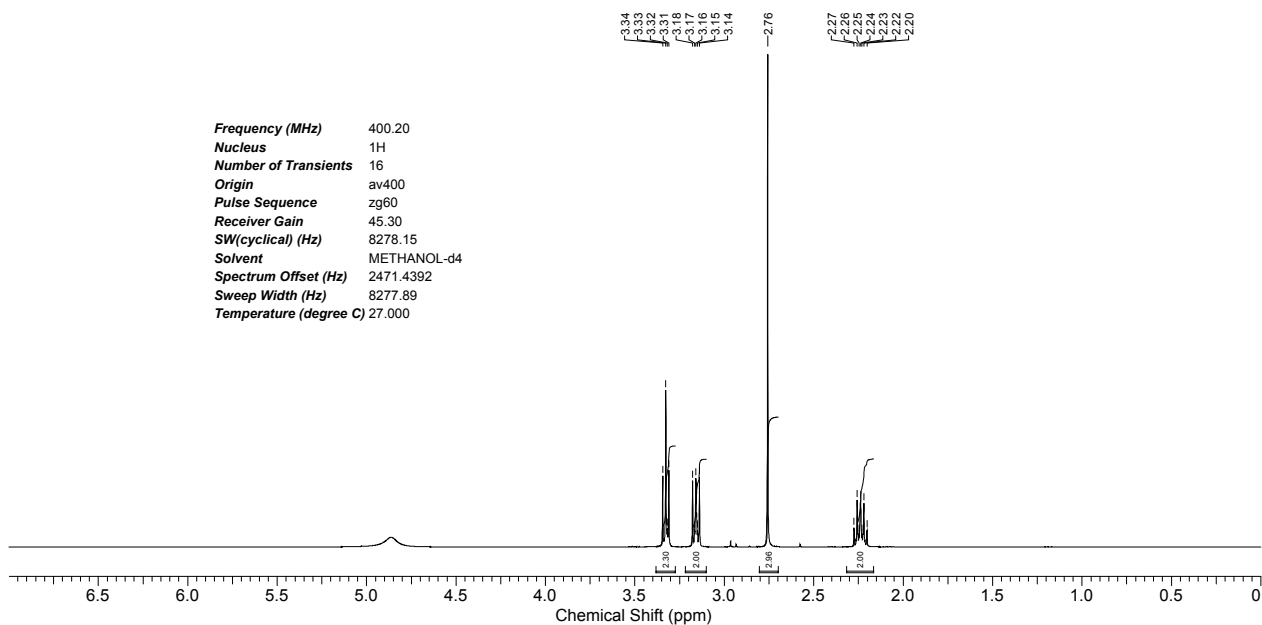
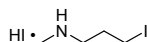
- (1) Tanner, D. D.; Meintzer, C. P. *J. Am. Chem. Soc.* **1985**, *107*, 6584-6589.
- (2) Lim, H. D.; Smits, R. A.; Bakker, R. A.; van Dam, C. M. E.; de Esch, I. J. P.; Leurs, R. *J. Med. Chem.* **2006**, *49*, 6650-6651.
- (3) Aroyan, C. E; Miller, S. J. *J. Am. Chem. Soc.*, **2007**, *129*, 256-257.
- (4) Adams, L. A.; Aggarwal, V. K.; Bonnert, R. V.; Bressel, B.; Cox, R. J.; Shepherd, J.; de Vicente, J.; Walter, M.; Whittingham, W. G.; Winn, C. L. *J. Org. Chem.*, **2003**, *68*, 9433-9440.
- (5) Rosaria Rella, M.; Williard, P. G. *J. Org. Chem.* **2007**, *72*, 525-531.
- (6) Meyer, W.; Reckendorf, Z.; Bonner, W. A. *J. Org. Chem.* **1961**, *26*, 4596-4599.
- (7) Bernardes, G. J. L.; Grayson, E. J.; Thompson, S.; Chalker, J. M.; Errey, J. C.; El Oualid, F.; Claridge, T. D. W.; Davis, B. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 2244-2247.

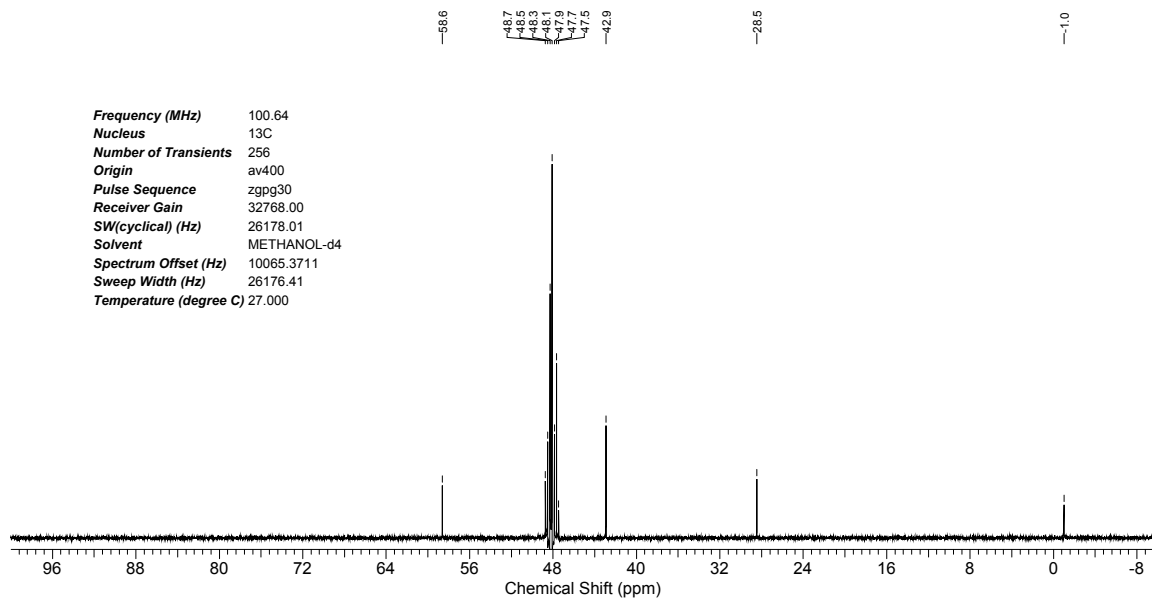
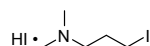
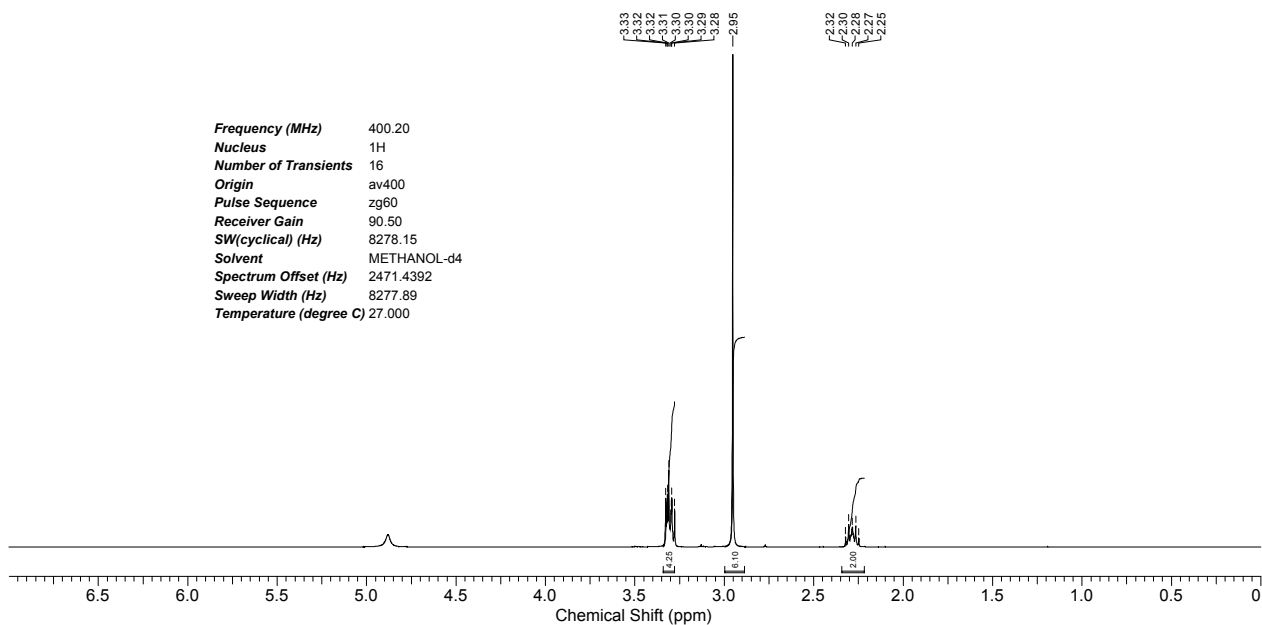
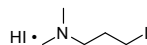
8.7 NMR Spectra

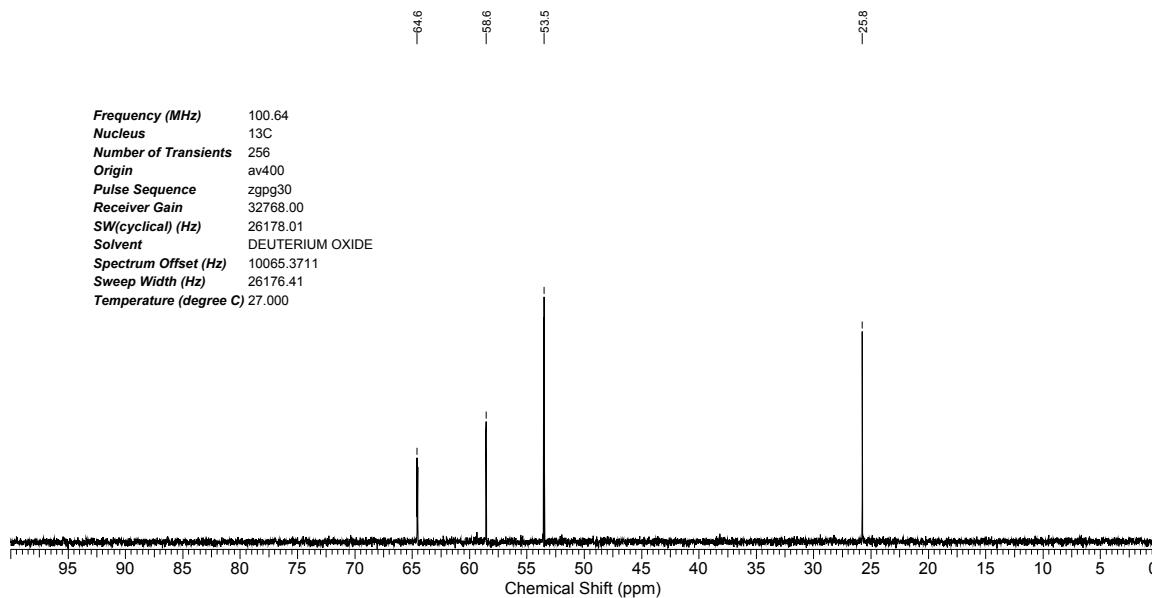
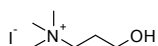
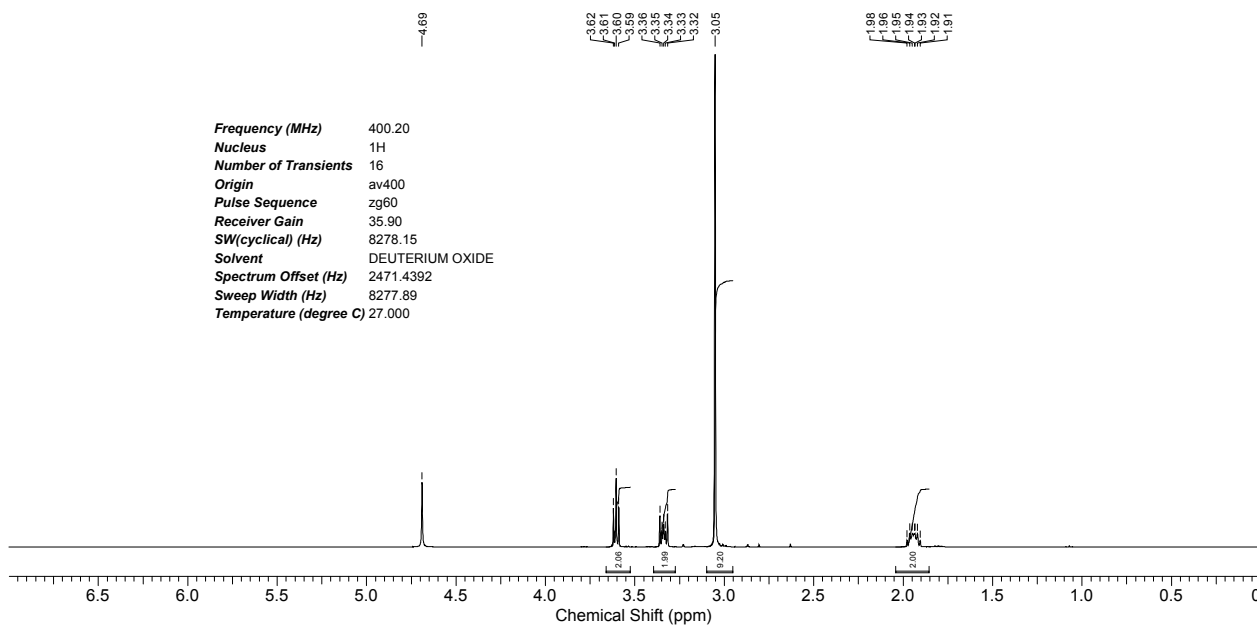
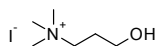


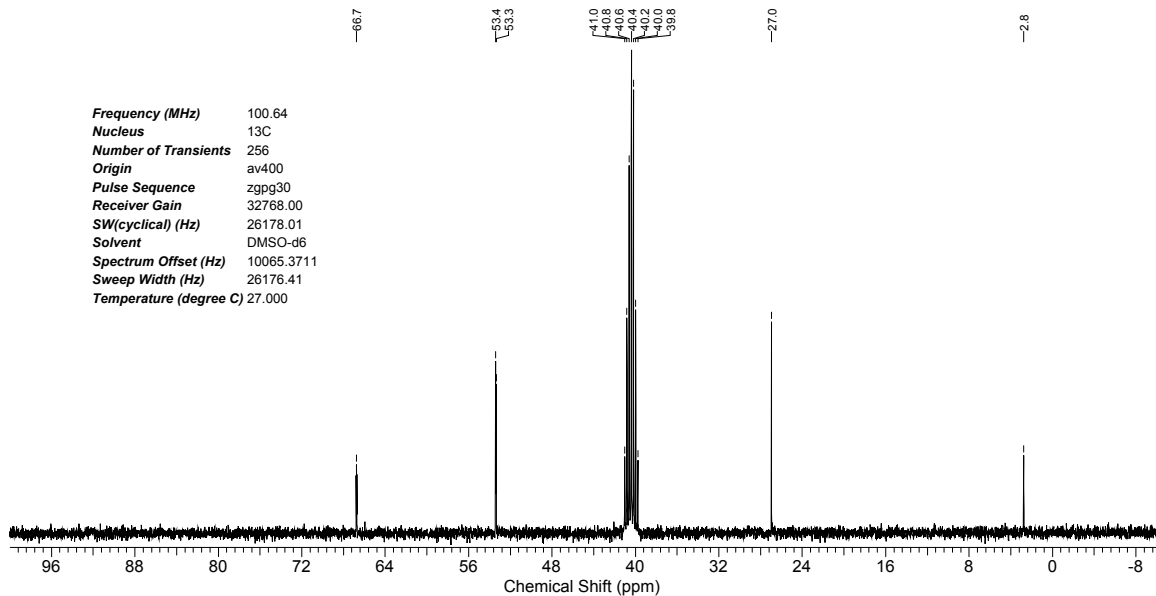
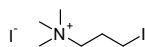
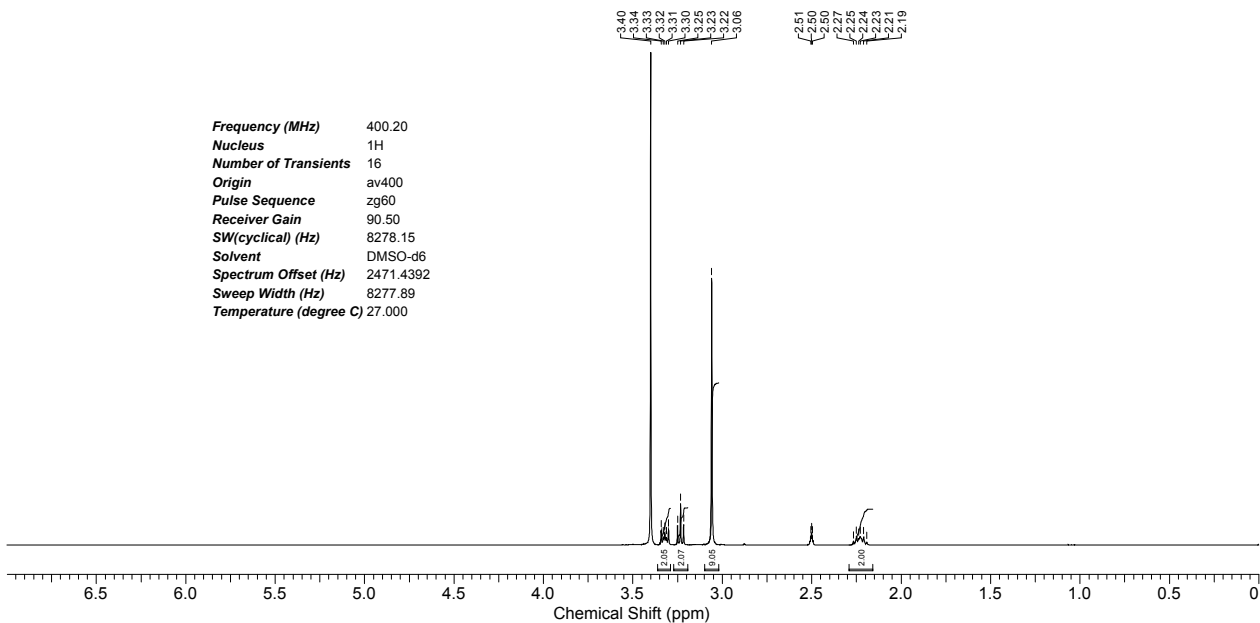
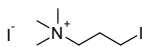


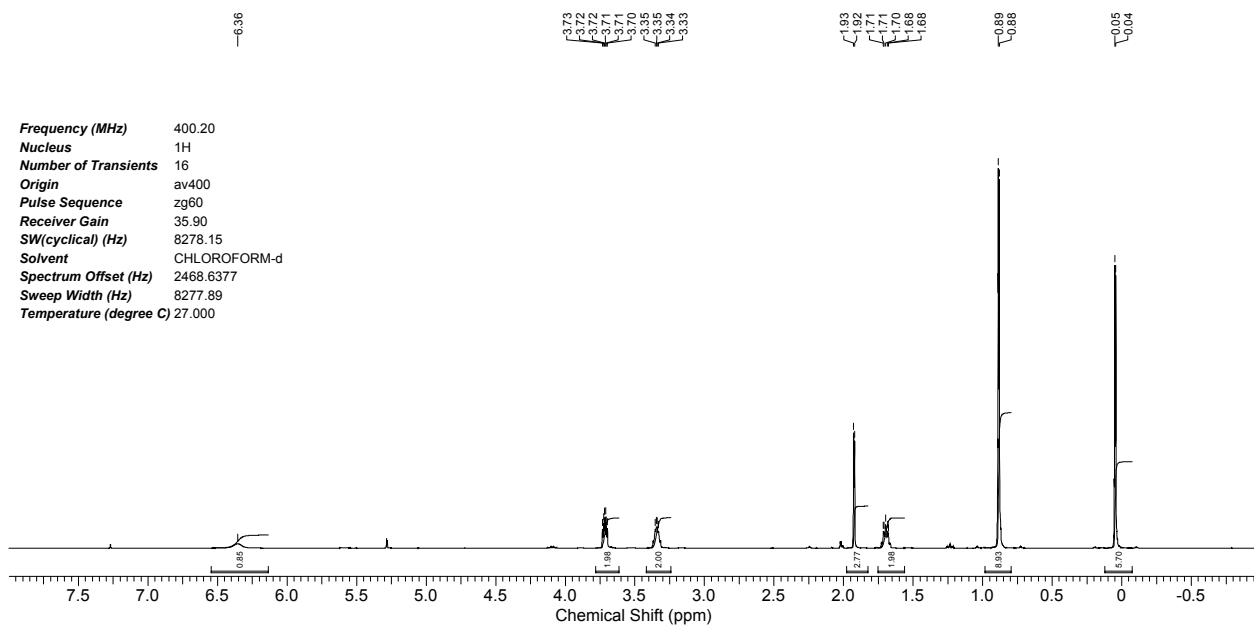
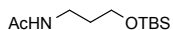




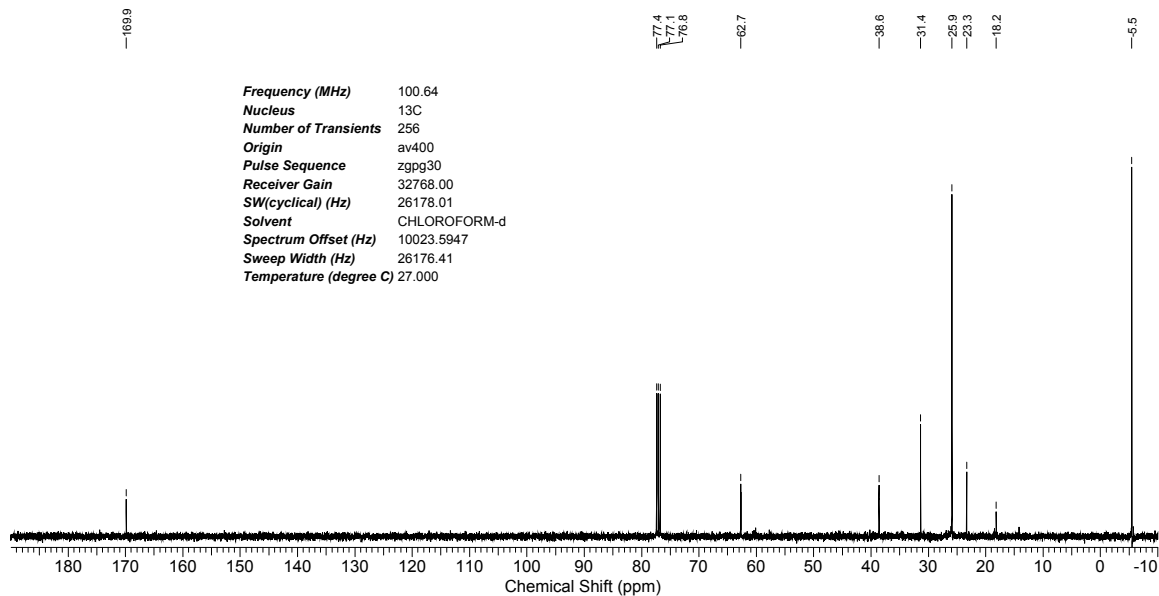
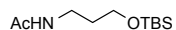




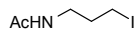




Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 35.90
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.6377
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000



Frequency (MHz) 100.64
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10023.5947
Sweep Width (Hz) 26176.41
Temperature (degree C) 27.000

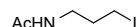
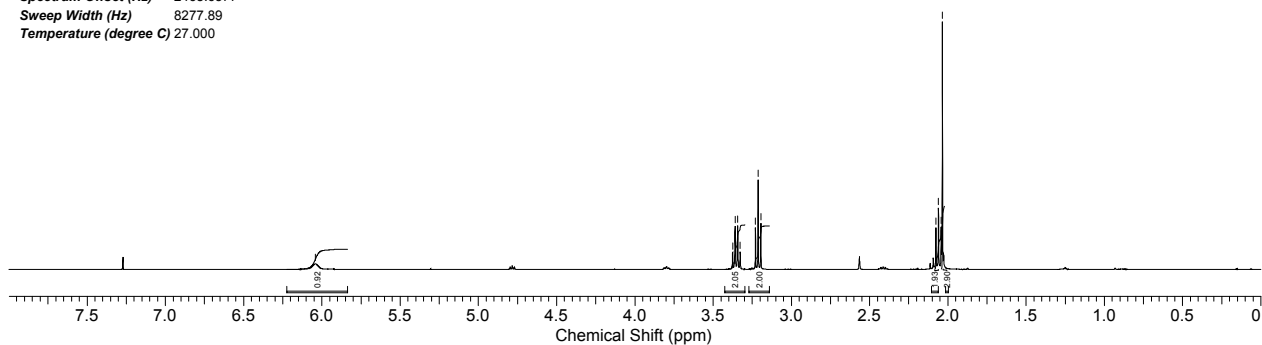


0.04

3.38
3.36
3.34
3.33
3.23
3.21
3.20

2.08
2.06
2.04

Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 161.30
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.6377
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000



171.1

77.4
77.2
76.8

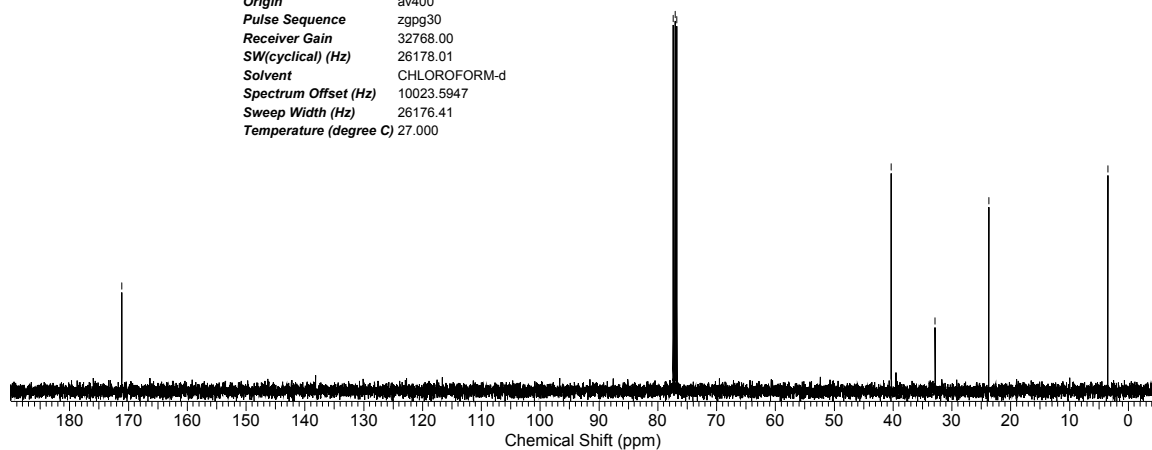
40.3

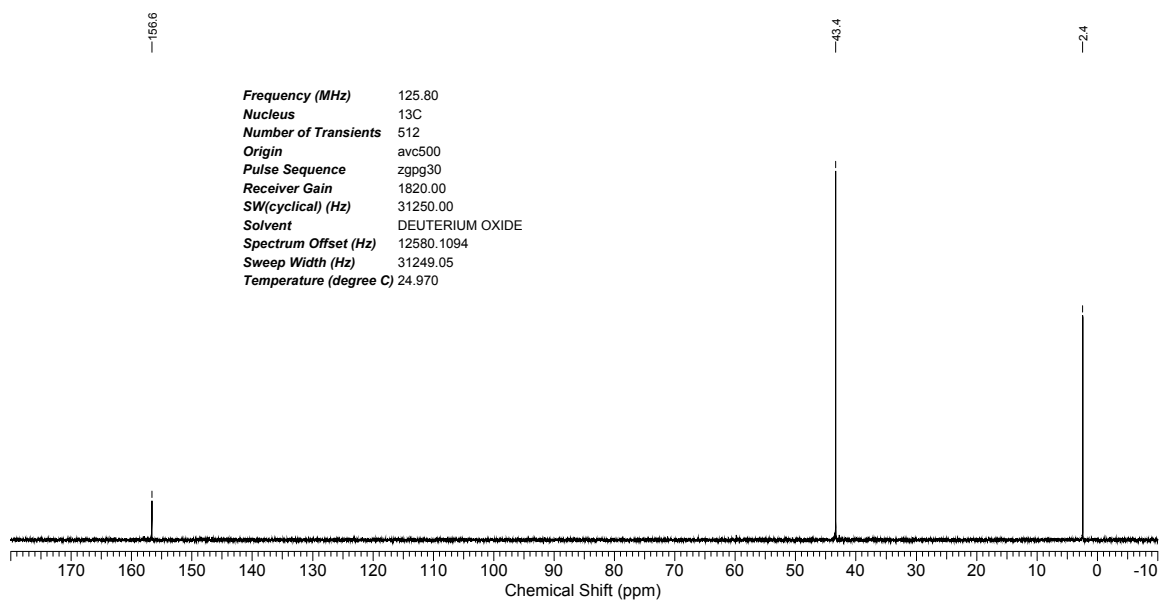
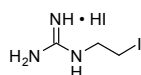
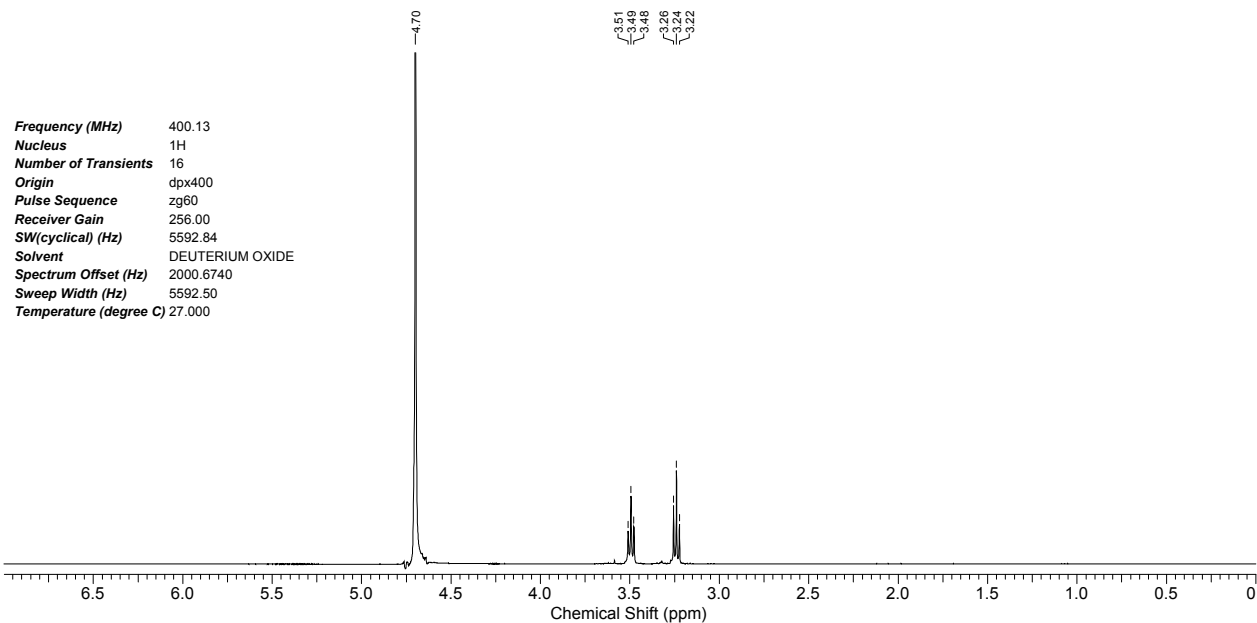
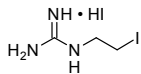
32.8

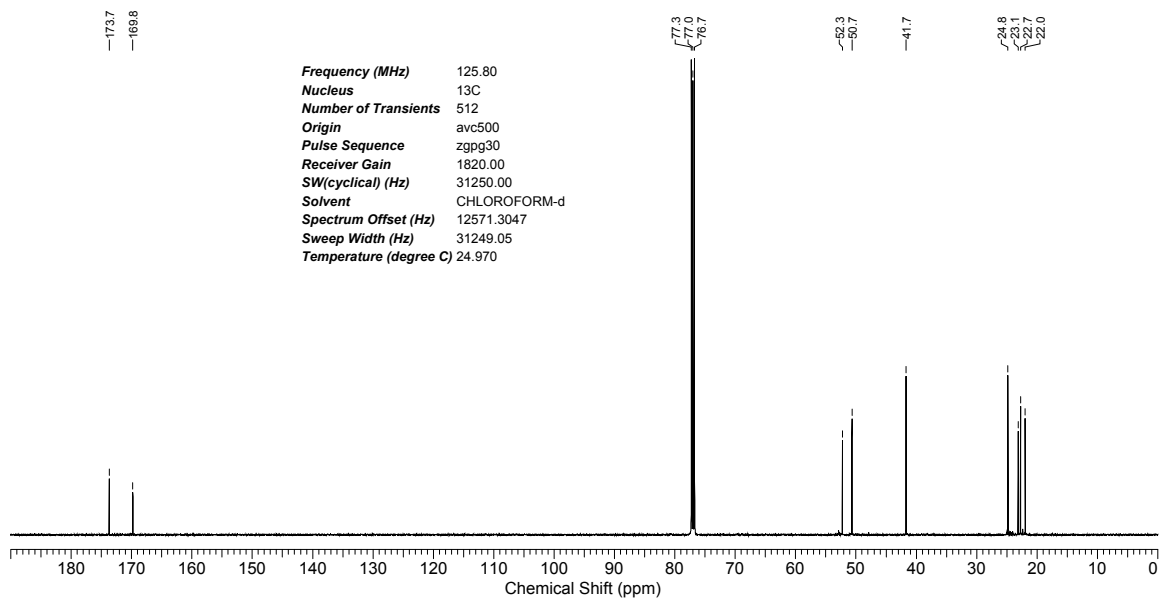
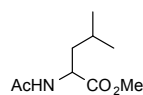
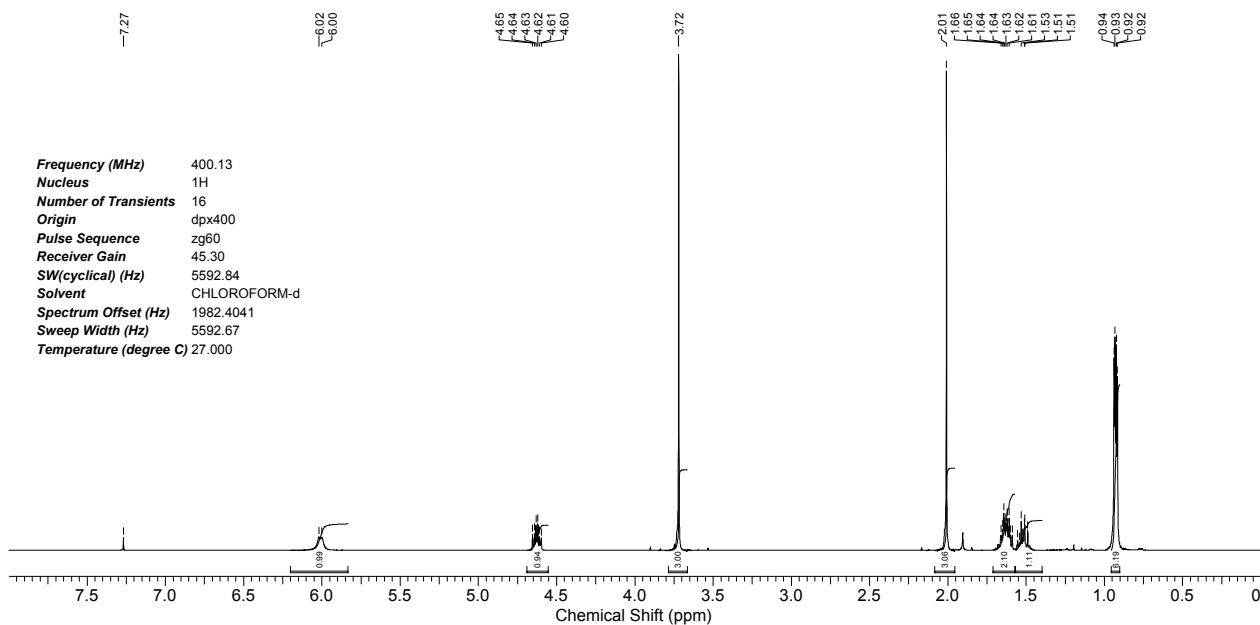
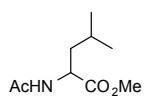
23.7

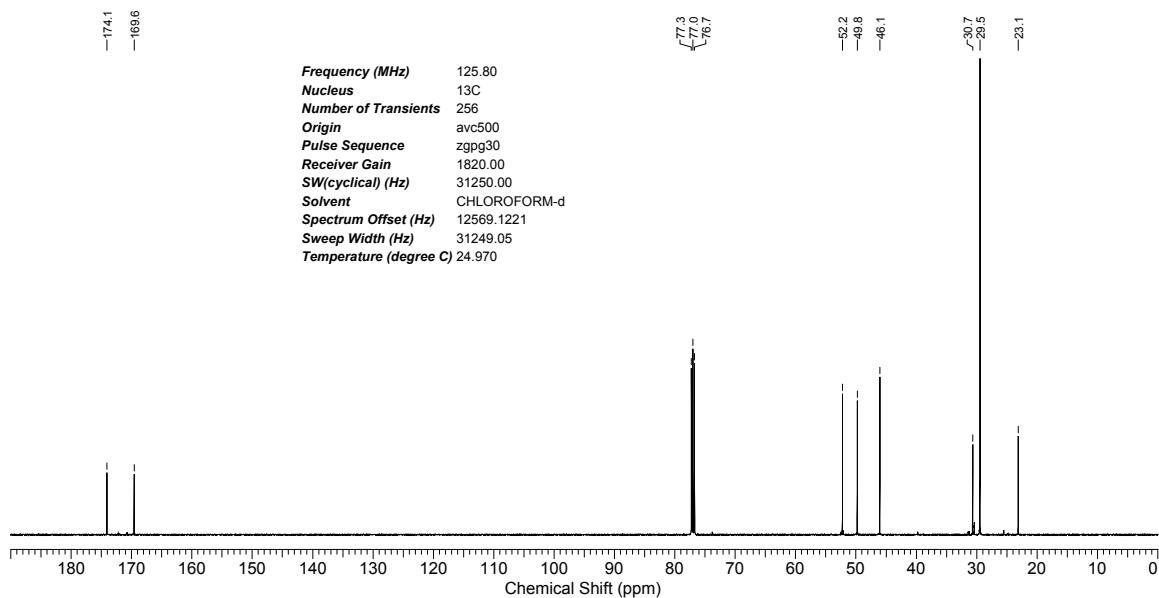
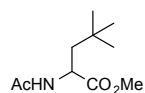
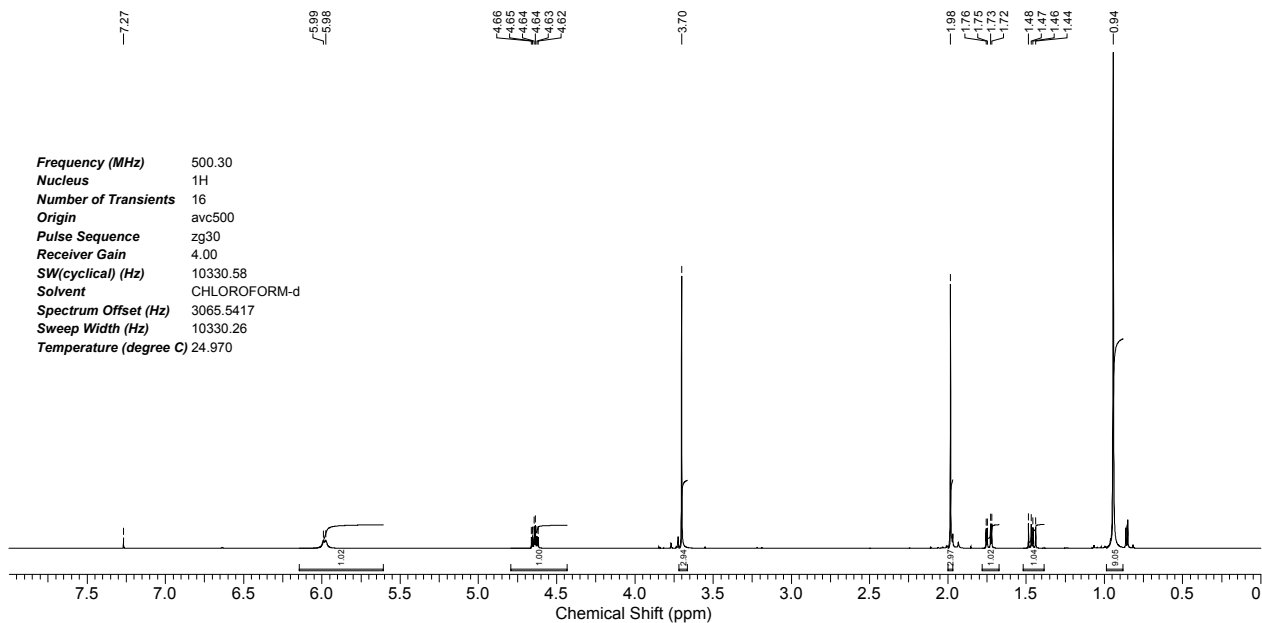
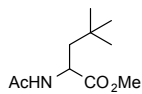
3.5

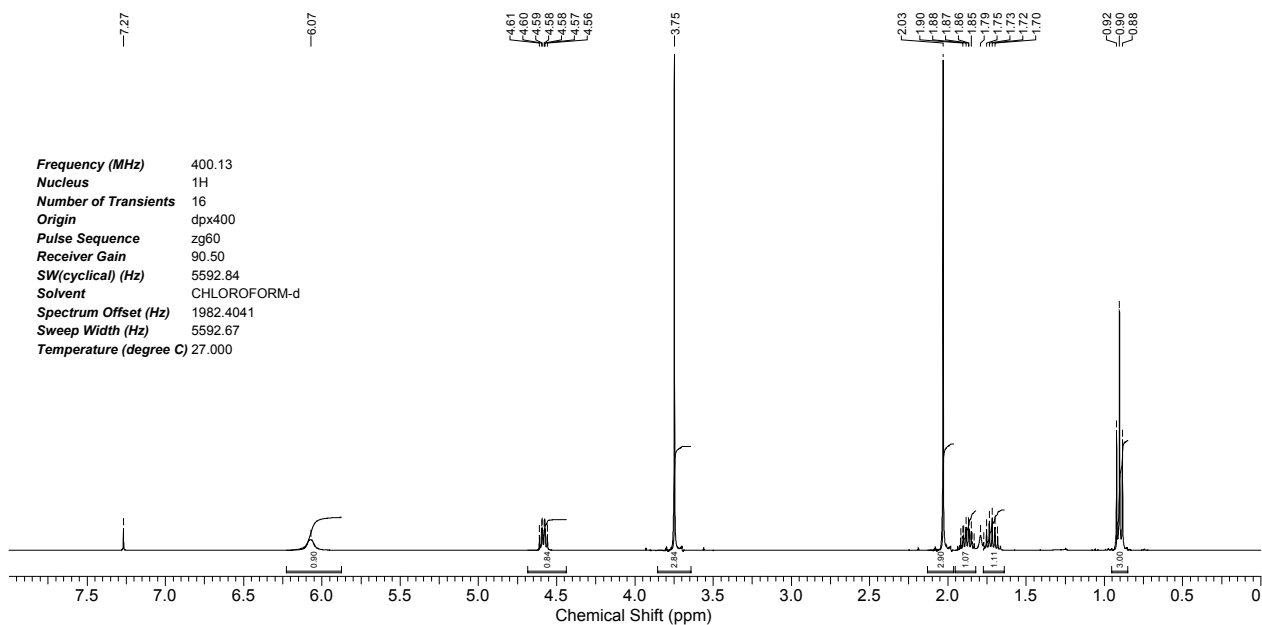
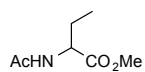
Frequency (MHz) 100.64
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10023.5947
Sweep Width (Hz) 26176.41
Temperature (degree C) 27.000



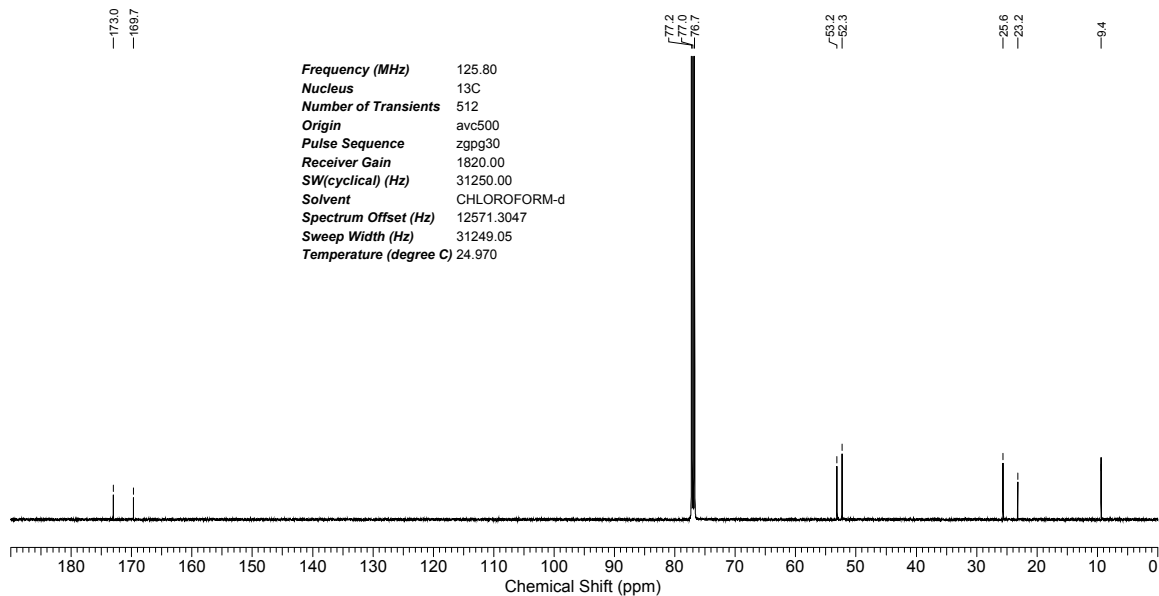
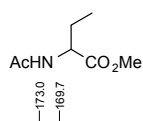




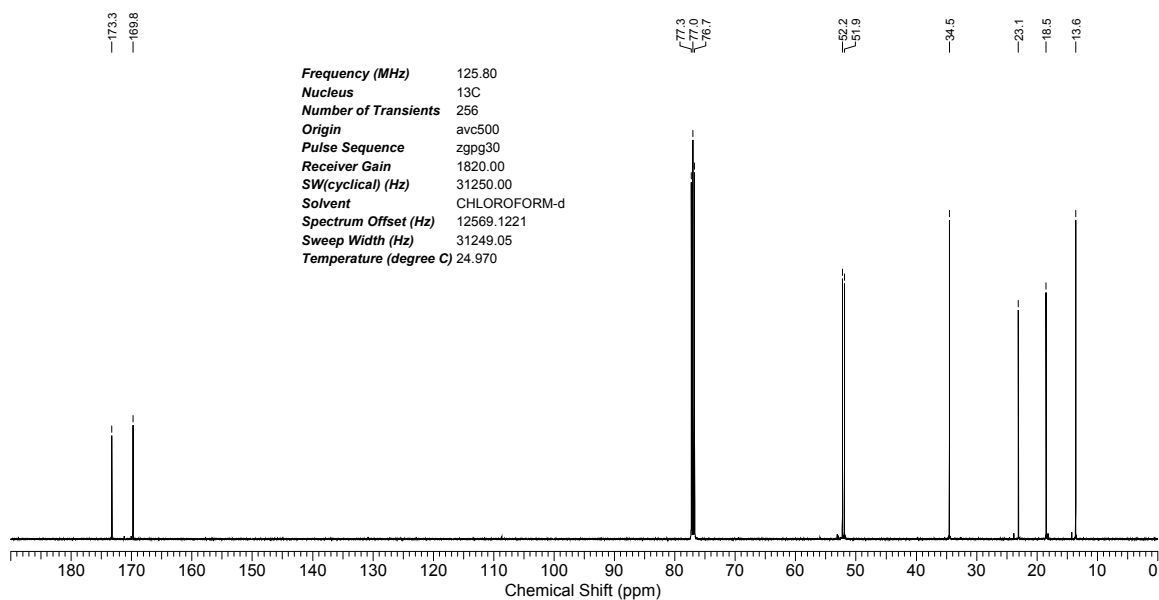
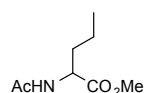
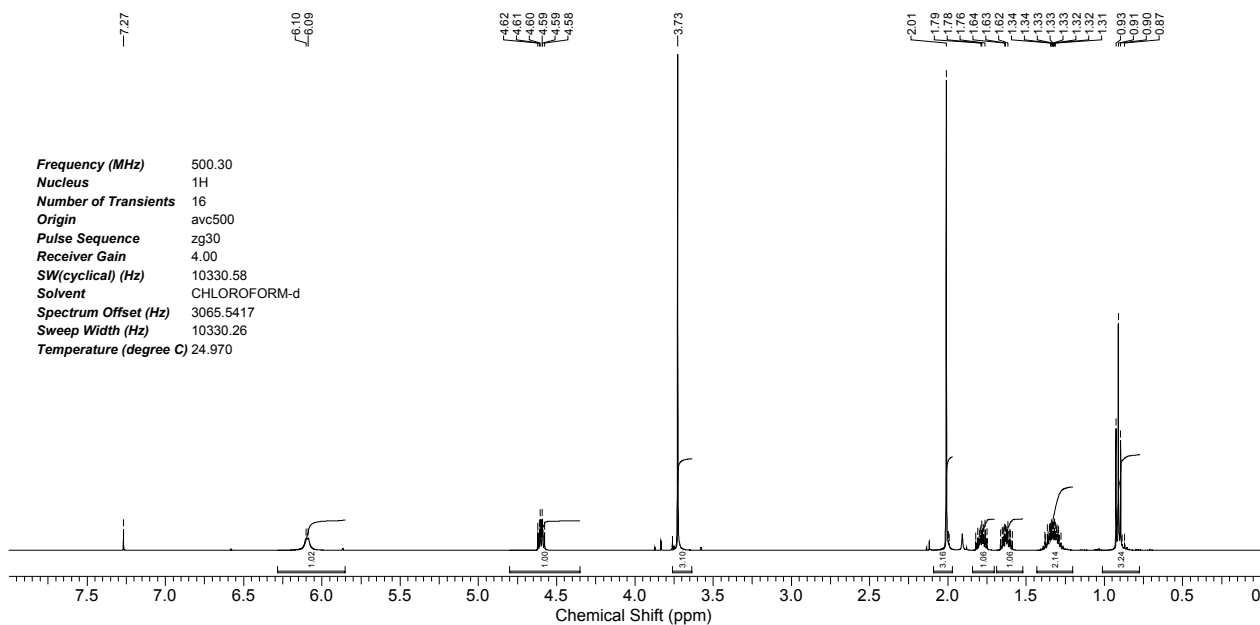
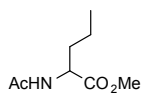


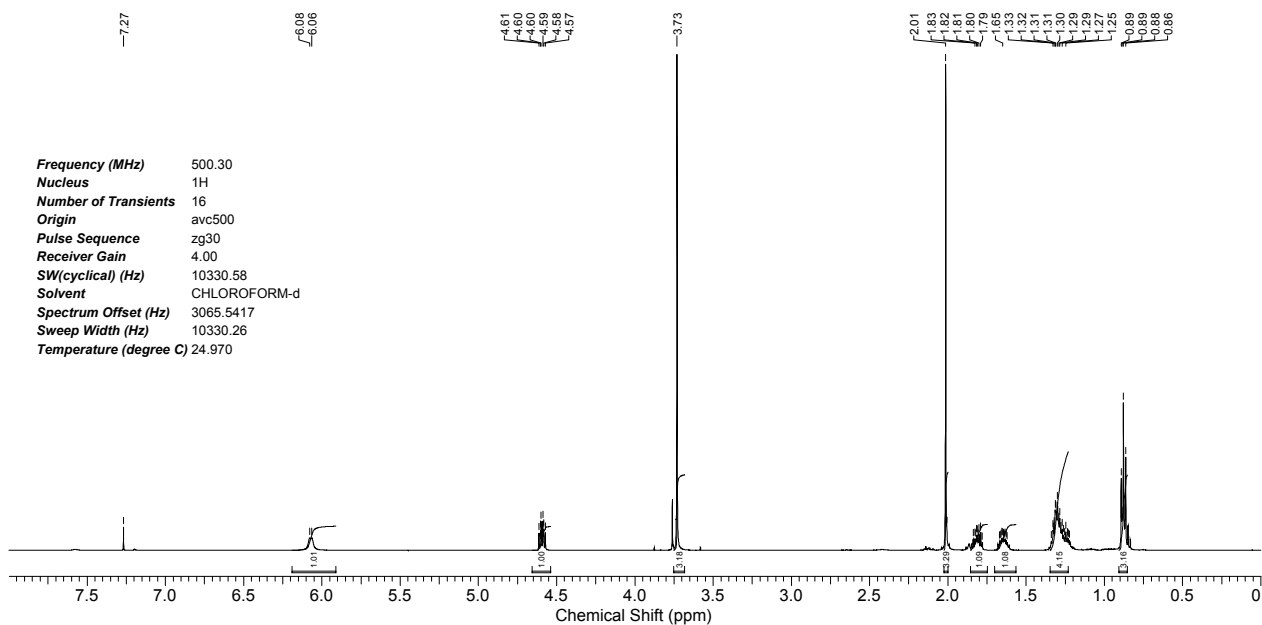
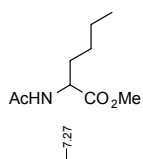


Frequency (MHz) 400.13
 Nucleus 1H
 Number of Transients 16
 Origin dpx400
 Pulse Sequence zg60
 Receiver Gain 90.50
 SW(cyclical) (Hz) 5592.84
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 1982.4041
 Sweep Width (Hz) 5592.67
 Temperature (degree C) 27.000

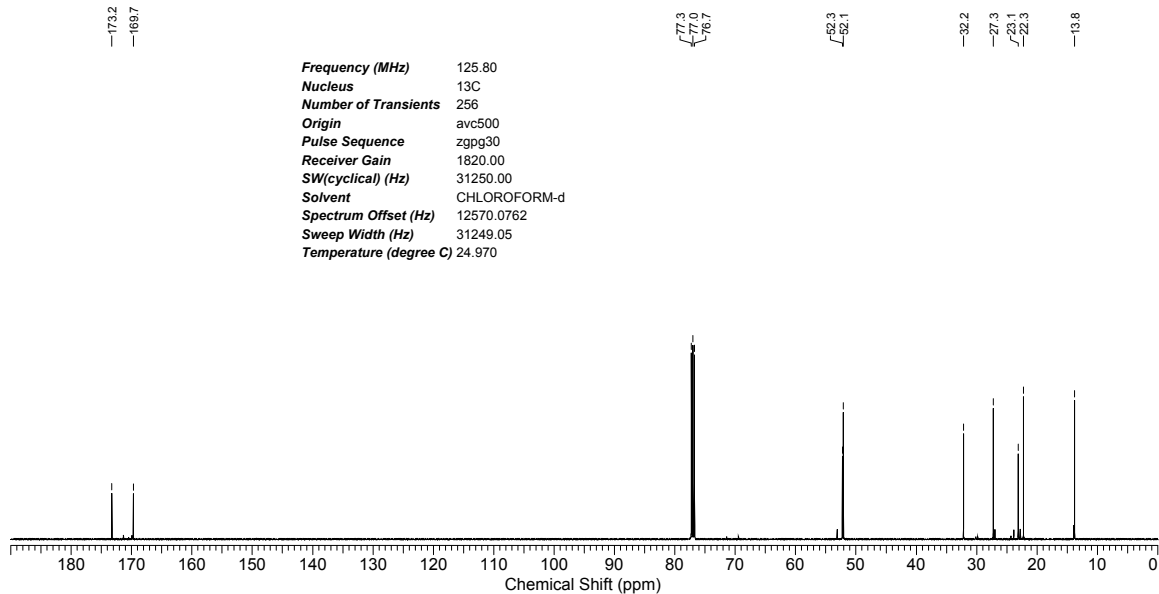
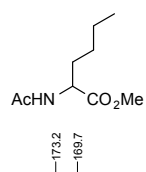


Frequency (MHz) 125.80
 Nucleus 13C
 Number of Transients 512
 Origin avc500
 Pulse Sequence zgpg30
 Receiver Gain 1820.00
 SW(cyclical) (Hz) 31250.00
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 12571.3047
 Sweep Width (Hz) 31249.05
 Temperature (degree C) 24.970

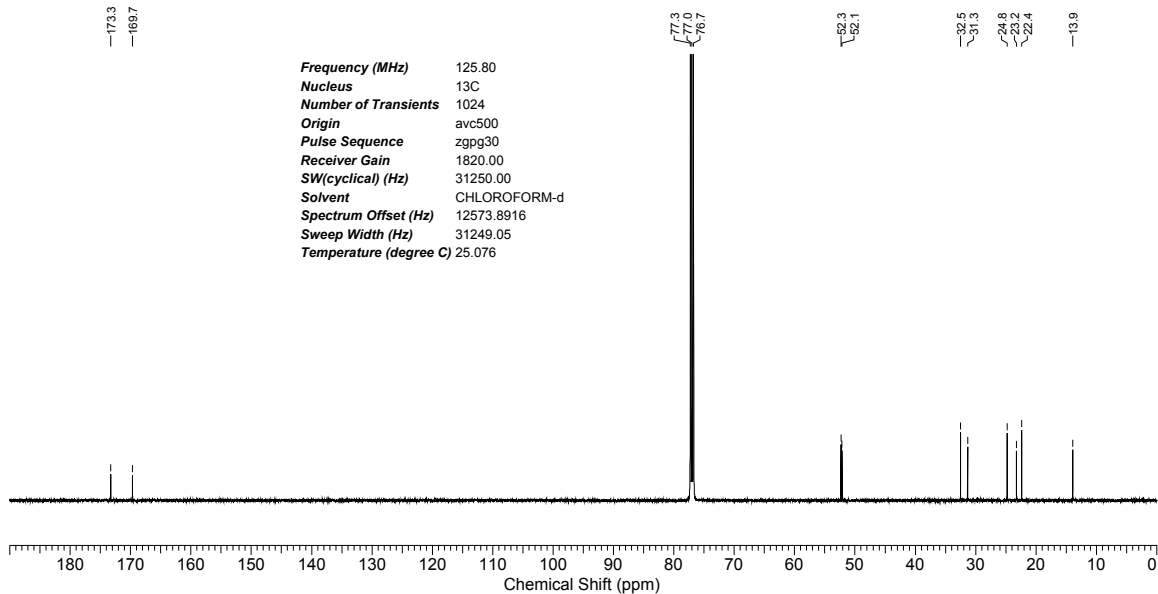
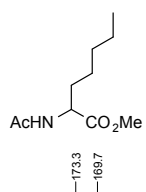
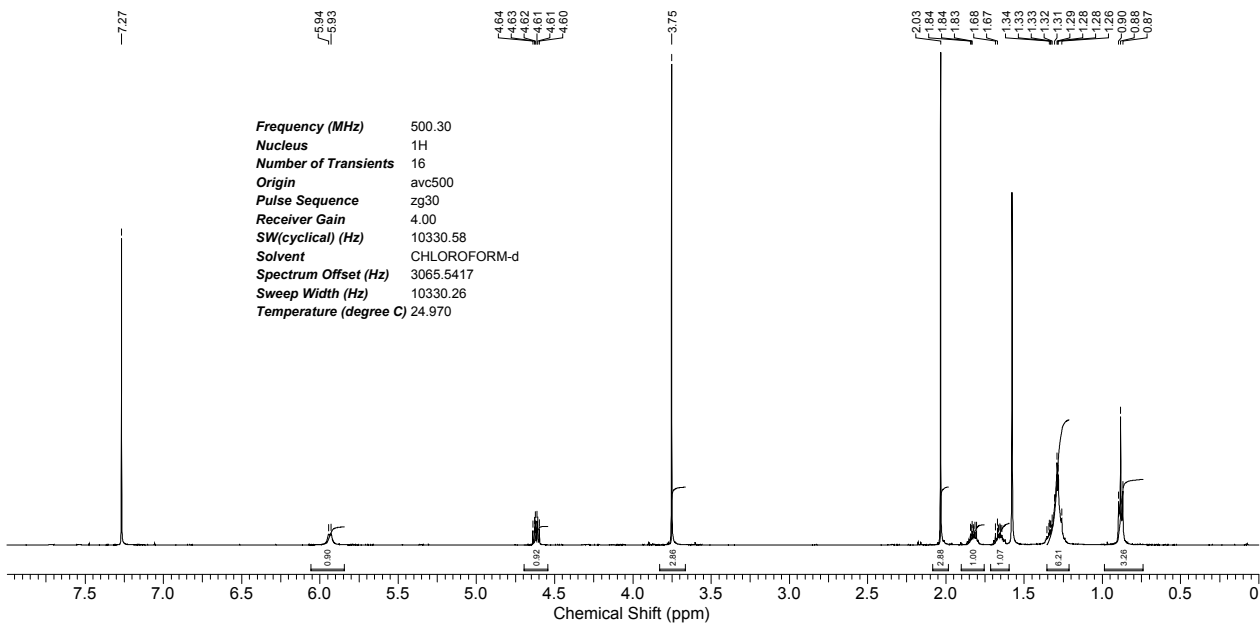
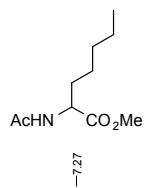


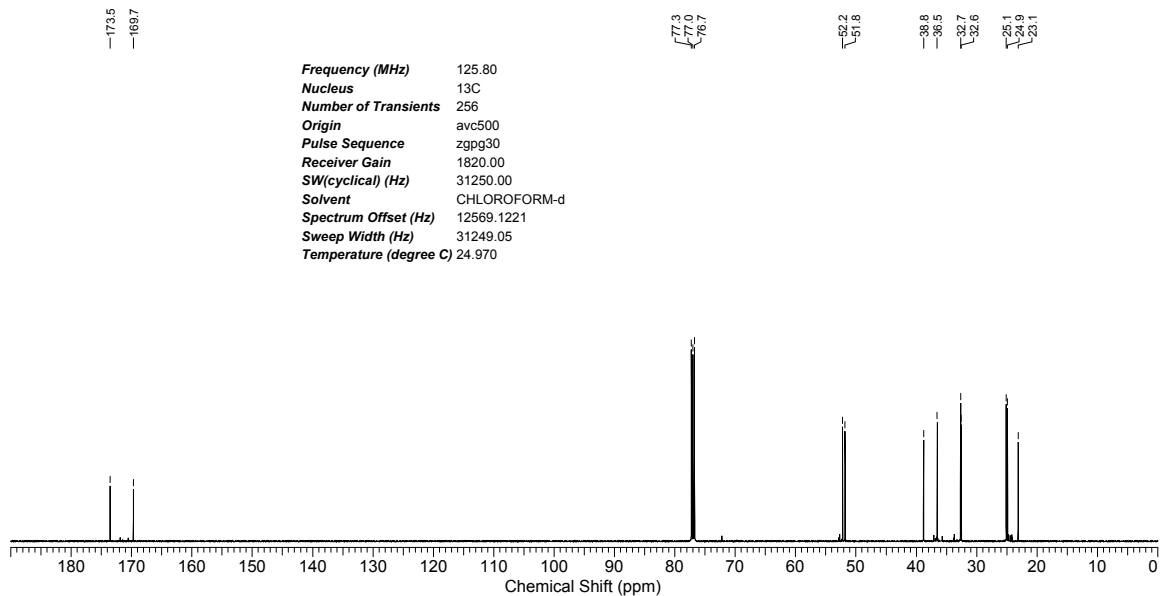
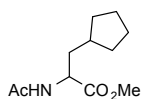
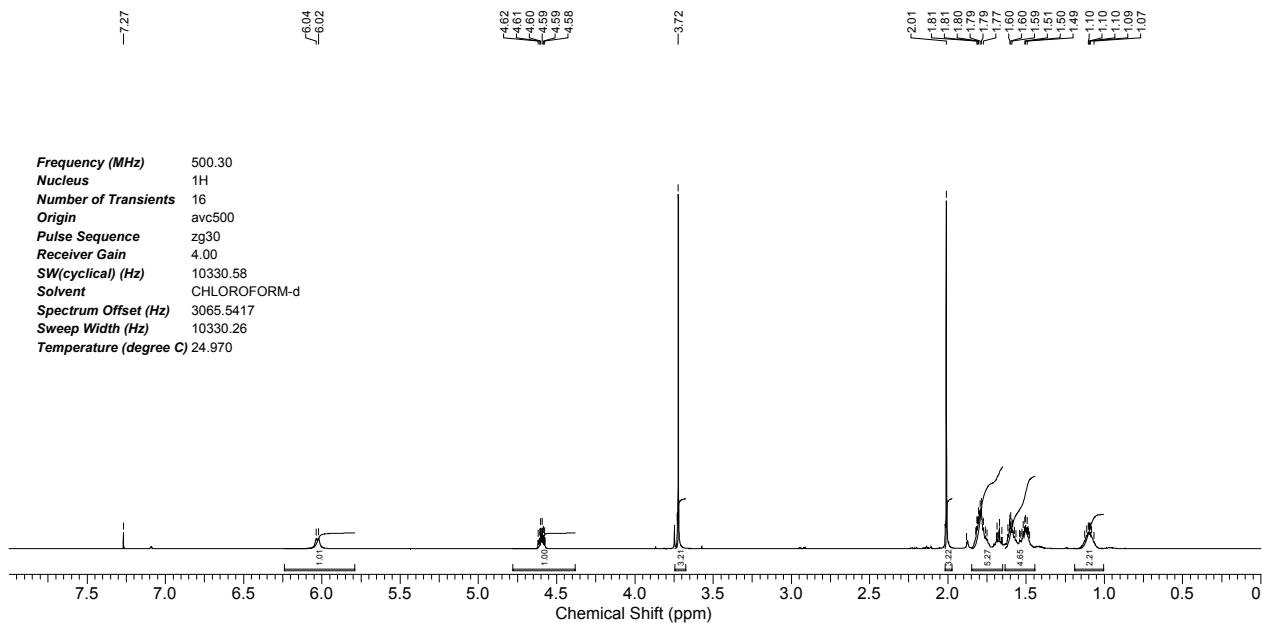
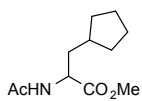


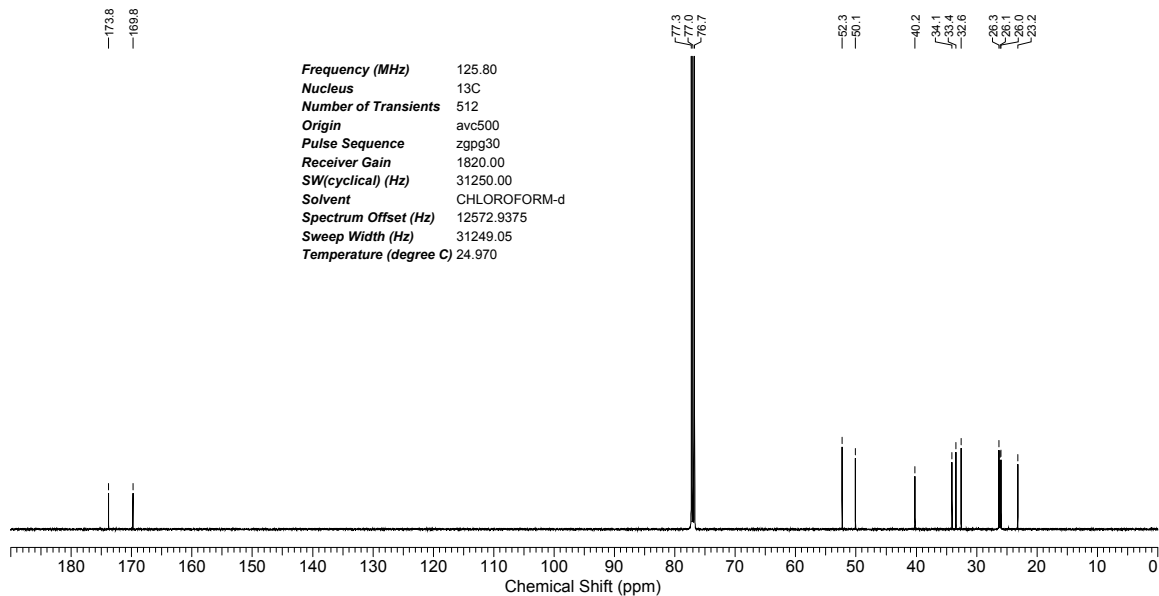
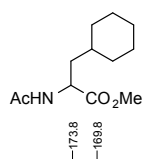
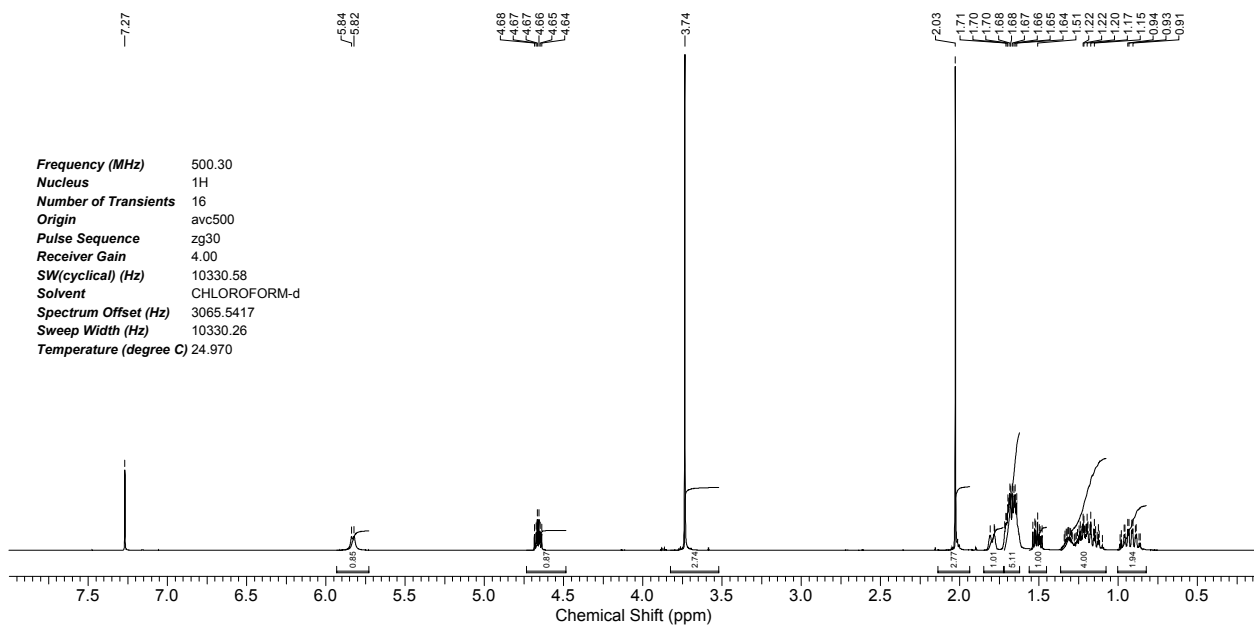
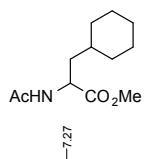
Frequency (MHz) 500.30
 Nucleus 1H
 Number of Transients 16
 Origin avc500
 Pulse Sequence zg30
 Receiver Gain 4.00
 SW(cyclical) (Hz) 10330.58
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 3065.5417
 Sweep Width (Hz) 10330.26
 Temperature (degree C) 24.970

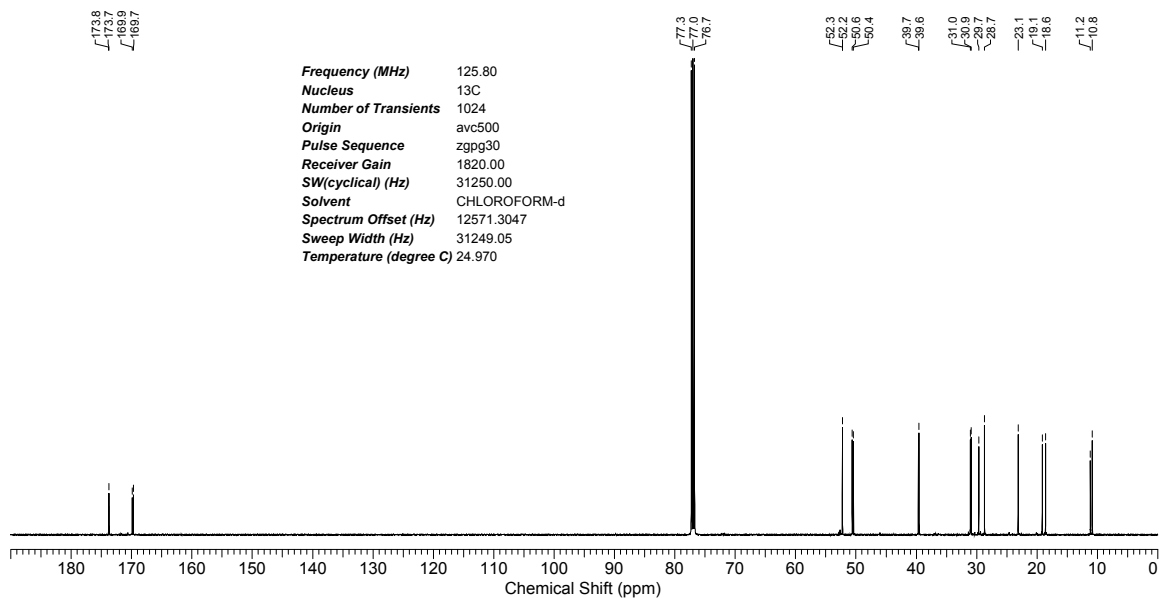
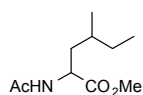
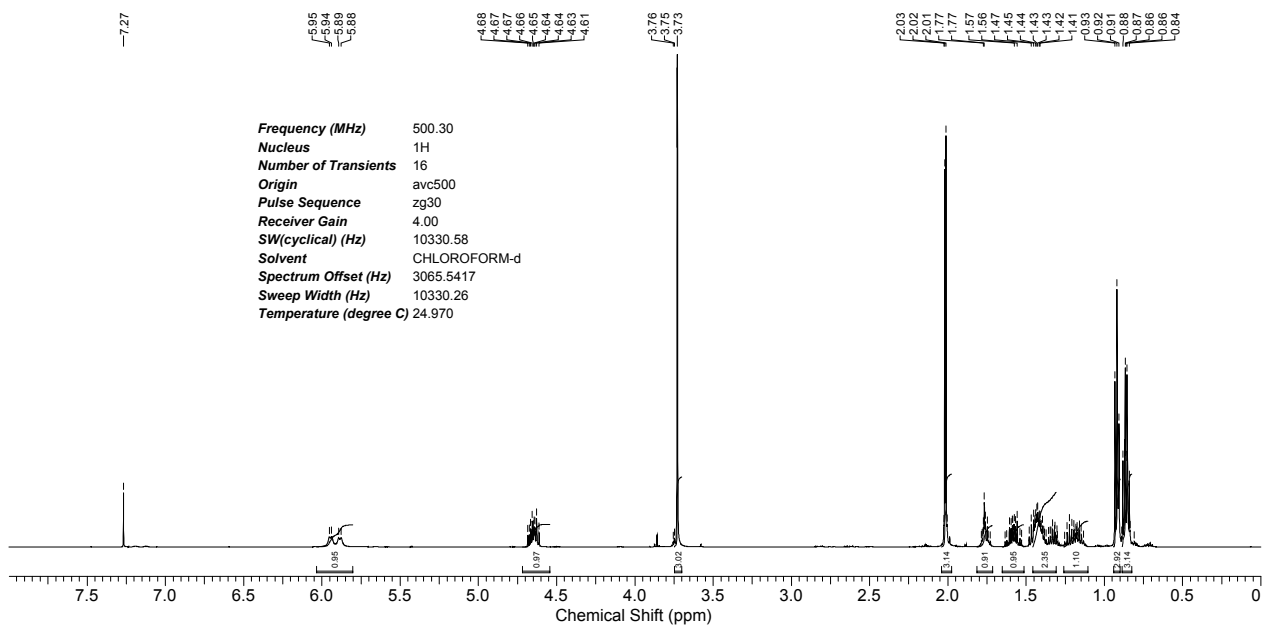
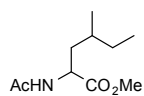


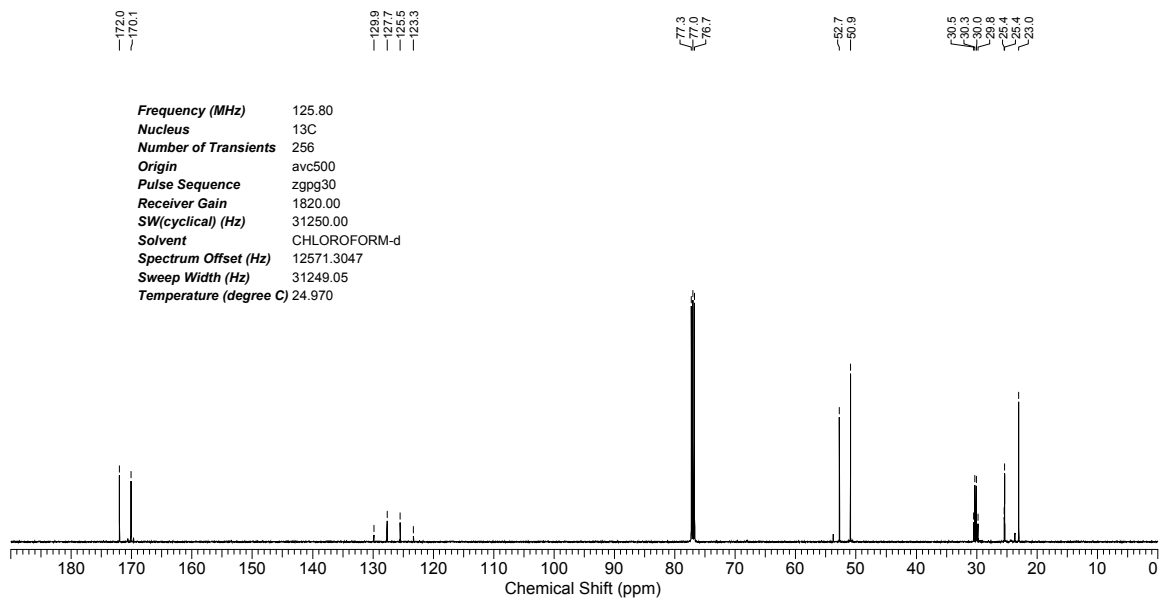
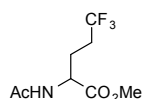
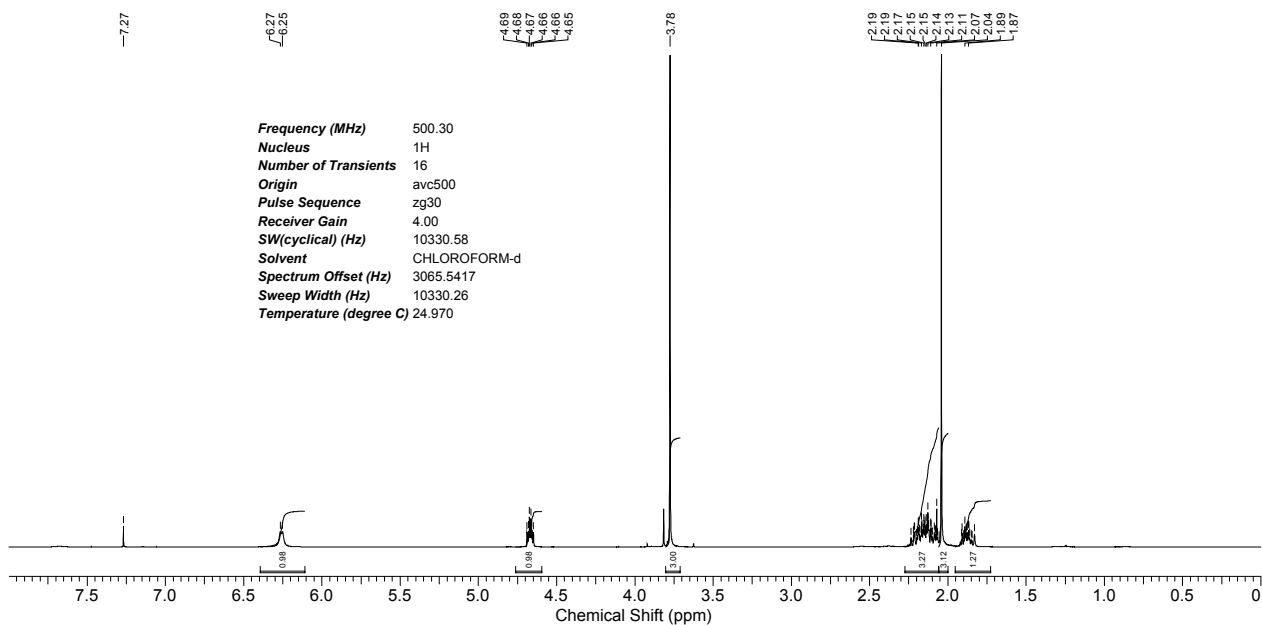
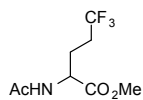
Frequency (MHz) 125.80
 Nucleus 13C
 Number of Transients 256
 Origin avc500
 Pulse Sequence zgpg30
 Receiver Gain 1820.00
 SW(cyclical) (Hz) 31250.00
 Solvent CHLOROFORM-d
 Spectrum Offset (Hz) 12570.0762
 Sweep Width (Hz) 31249.05
 Temperature (degree C) 24.970

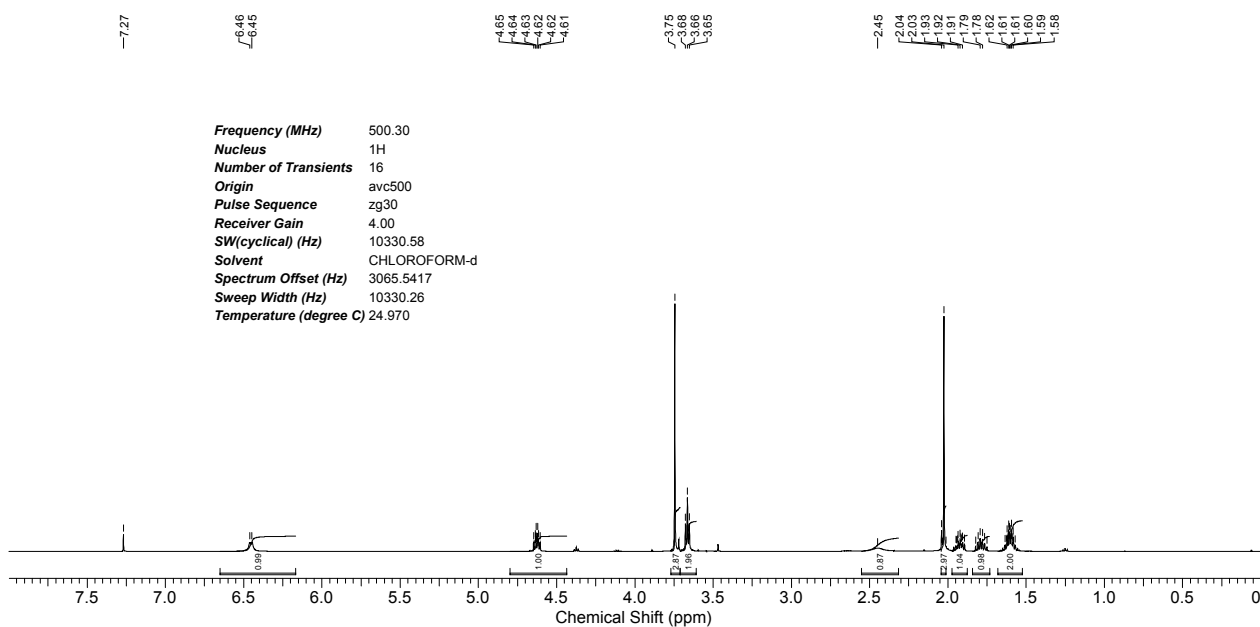
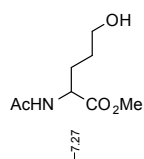
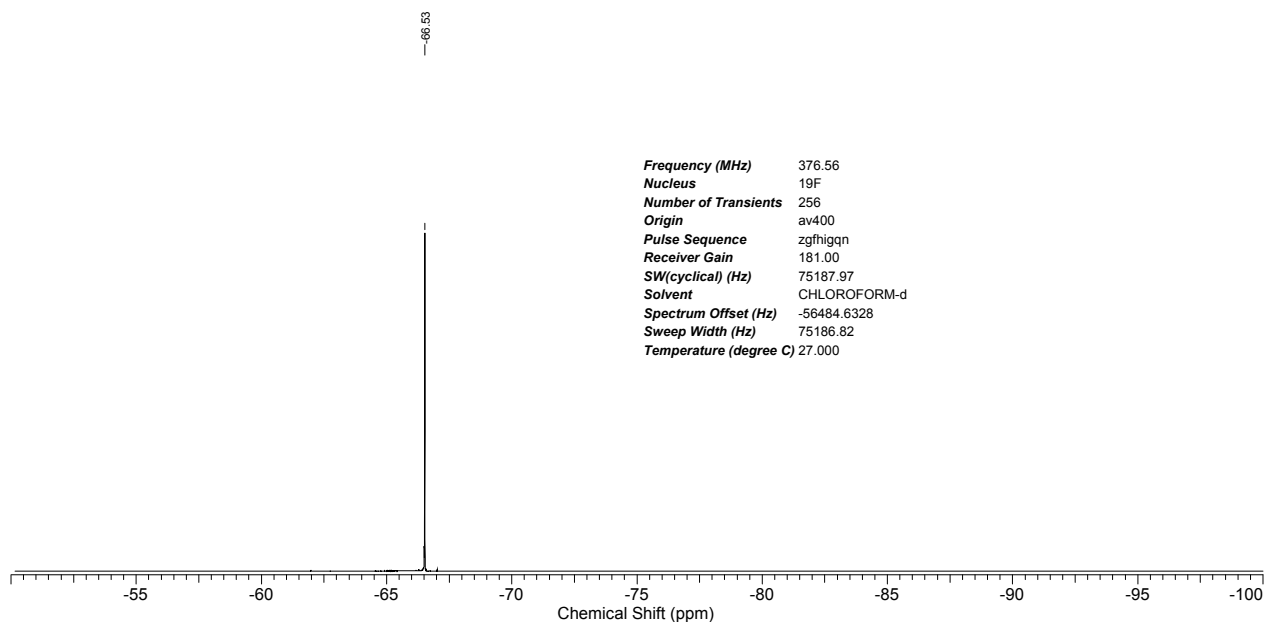
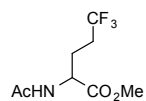


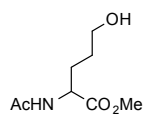












173.0
170.2

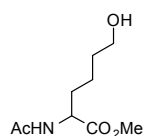
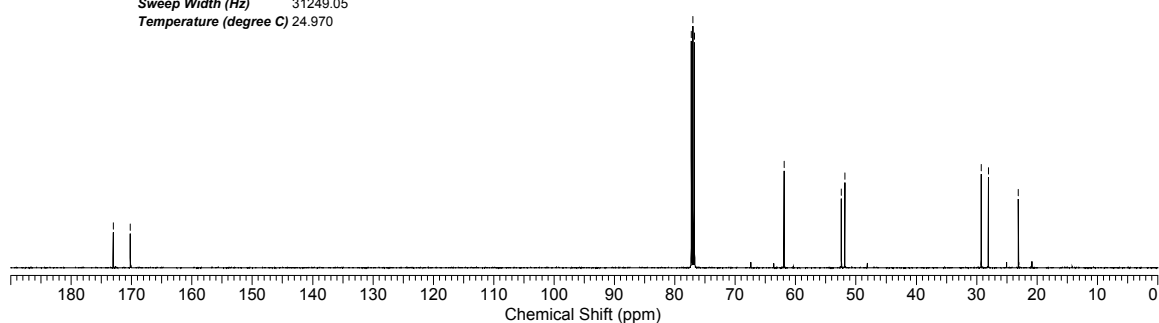
77.3
77.0
76.8

61.9

52.4
51.9

29.3
28.1
23.1

Frequency (MHz) 125.80
Nucleus ¹³C
Number of Transients 512
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12571.3047
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970



7.27

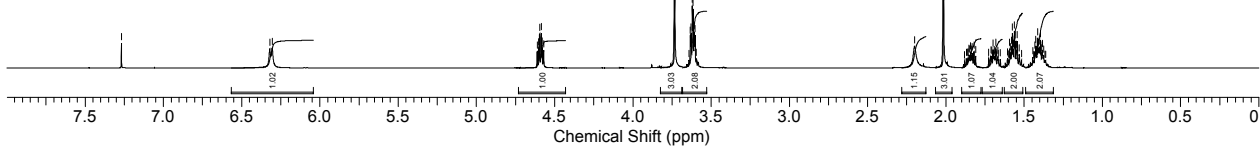
6.32
6.30

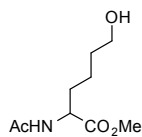
4.61
4.60
4.60
4.59
4.59
4.57

3.73
3.63
3.63
3.63
3.62
3.61
3.60

2.20
2.02
1.87
1.83
1.82
1.68
1.59
1.59
1.56
1.56
1.54
1.43
1.40
1.40

Frequency (MHz) 500.30
Nucleus ¹H
Number of Transients 16
Origin avc500
Pulse Sequence zg30
Receiver Gain 4.00
SW(cyclical) (Hz) 10330.58
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 3065.5417
Sweep Width (Hz) 10330.26
Temperature (degree C) 24.970





173.1
170.1

77.3
77.0
76.8

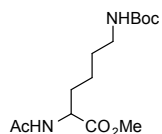
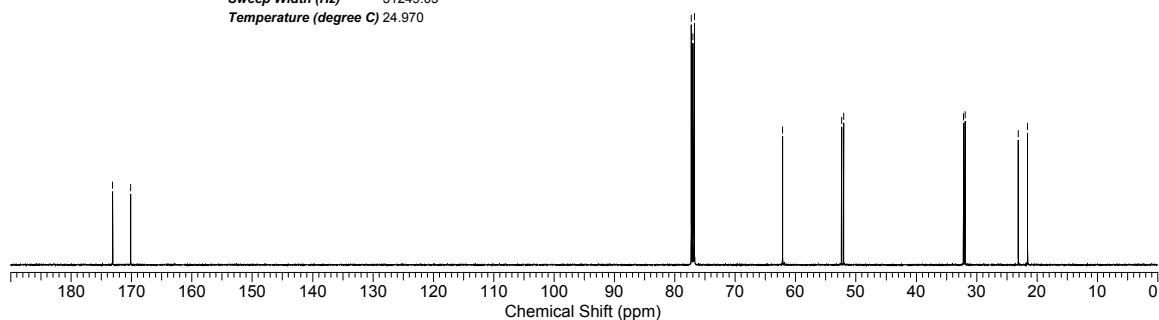
62.2

52.4
52.1

32.2
31.9

23.1
21.6

Frequency (MHz) 125.80
Nucleus ¹³C
Number of Transients 256
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12571.3047
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970



7.27

6.17
6.16

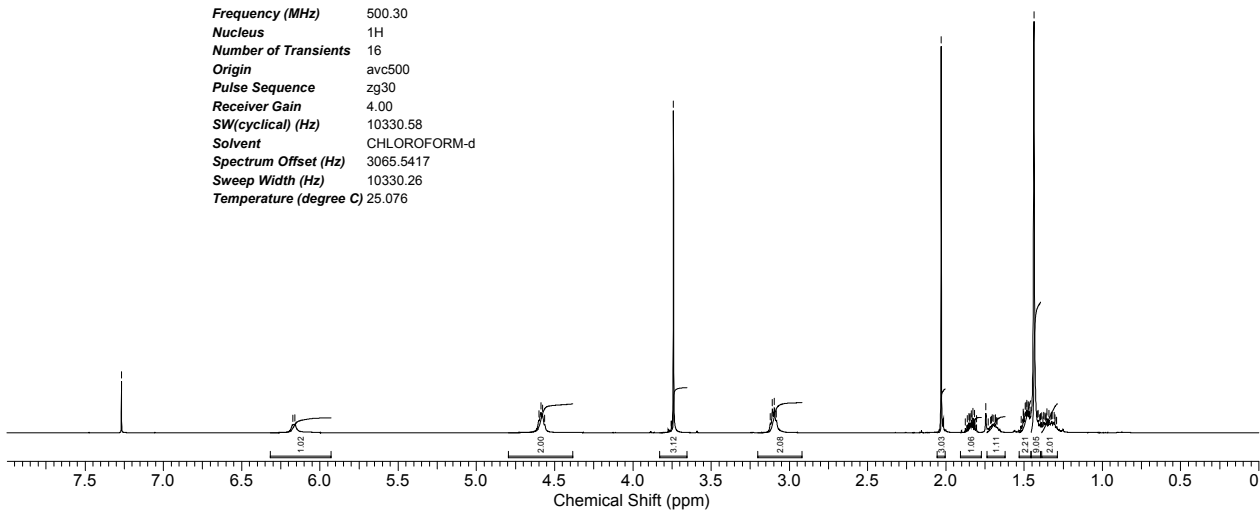
4.60
4.59
4.58
4.56

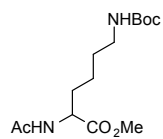
3.74

3.12
3.11
3.10
3.08

2.03
1.86
1.86
1.83
1.74
1.51
1.49
1.48
1.46
1.45
1.44
1.41
1.35

Frequency (MHz) 500.30
Nucleus ¹H
Number of Transients 16
Origin avc500
Pulse Sequence zg30
Receiver Gain 4.00
SW(cyclical) (Hz) 10330.58
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 3065.5417
Sweep Width (Hz) 10330.26
Temperature (degree C) 25.076





—173.0
—169.9
—156.1

79.1
77.3
77.0
76.7

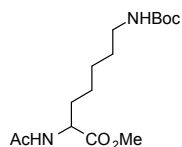
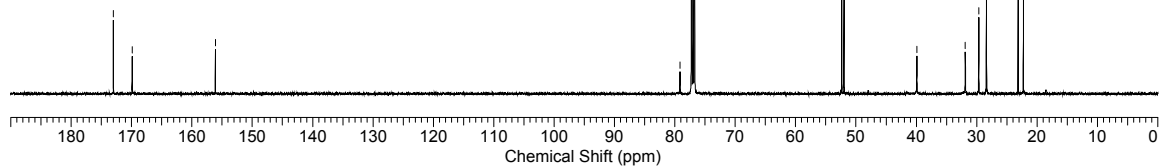
52.4
52.0

—39.9

31.9
30.7
28.4

23.1
22.3

Frequency (MHz) 125.80
Nucleus 13C
Number of Transients 512
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12571.9844
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970



—7.27

6.17
6.15

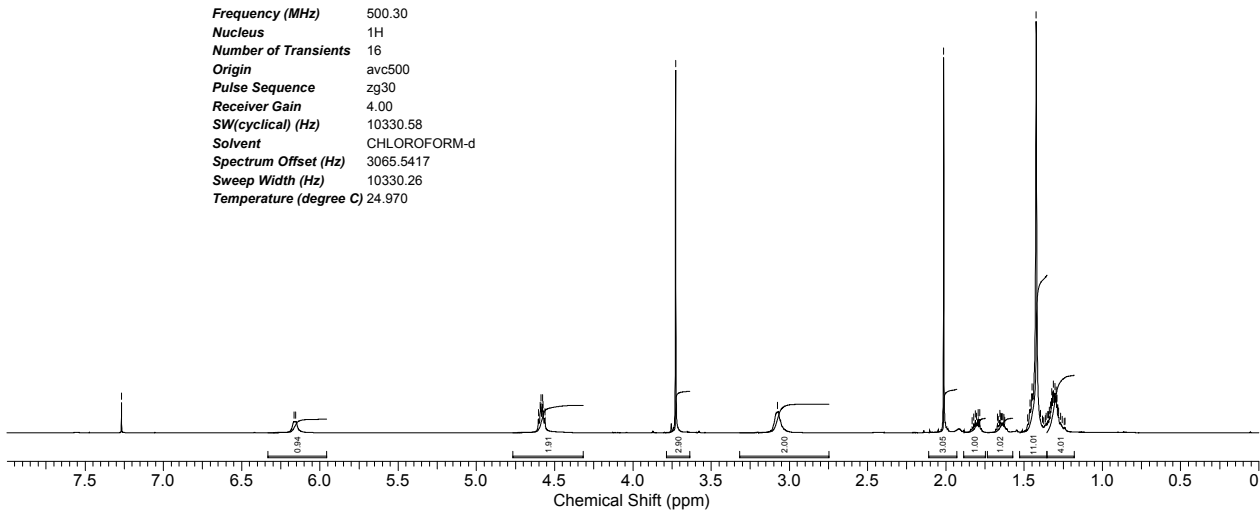
4.60
4.59
4.58
4.57
4.56

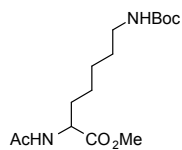
—3.73

—3.08

2.01
1.81
1.80
1.79
1.78
1.46
1.45
1.42
1.34
1.32
1.32
1.30
1.30
1.29

Frequency (MHz) 500.30
Nucleus 1H
Number of Transients 16
Origin avc500
Pulse Sequence zg30
Receiver Gain 4.00
SW(cyclical) (Hz) 10330.58
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 3065.5417
Sweep Width (Hz) 10330.26
Temperature (degree C) 24.970





173.1
169.8
156.0

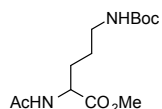
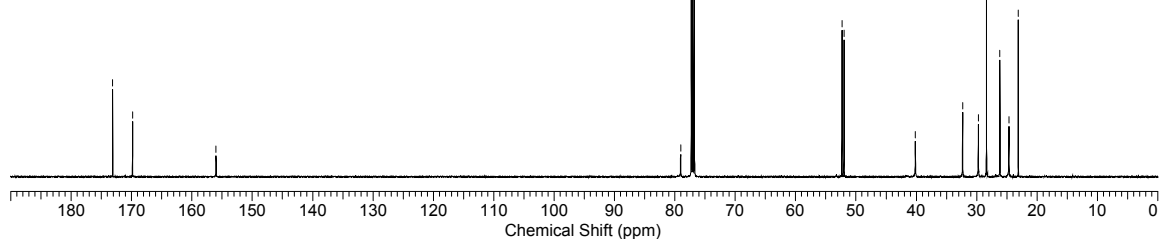
79.0
77.3
77.0
76.8

52.3
52.0

40.2

32.3
29.7
28.4
26.1
24.7
23.1

Frequency (MHz) 125.80
Nucleus 13C
Number of Transients 512
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12571.3047
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970



7.27

6.33
6.32

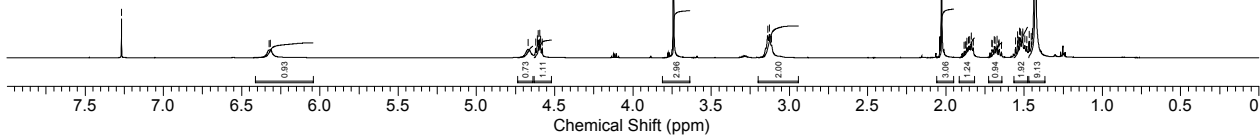
4.67
4.62
4.61
4.60
4.59
4.58

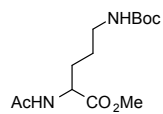
3.74

3.14
3.13
3.11

2.09
1.85
1.84
1.66
1.54
1.53
1.52
1.51
1.47
1.45
1.43

Frequency (MHz) 500.30
Nucleus 1H
Number of Transients 16
Origin avc500
Pulse Sequence zg30
Receiver Gain 4.00
SW(cyclical) (Hz) 10330.58
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 3065.5417
Sweep Width (Hz) 10330.26
Temperature (degree C) 24.970





Chemical Shifts (ppm):
 172.8, 170.0, 158.1

Chemical Shifts (ppm):
 79.3, 77.3, 77.0, 76.8

Chemical Shifts (ppm):
 52.4, 51.9

Chemical Shift (ppm):
 39.9

Chemical Shifts (ppm):
 29.5, 28.4, 26.1, 23.1

Frequency (MHz) 125.80
Nucleus ¹³C
Number of Transients 512
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 12571.3047
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970

