Olefin Metathesis for Site-Selective Protein Modification

A thesis submitted to the Board of the Faculty of Physical Sciences in partial fulfillment of the degree of

Doctor of Philosophy

at the

University of Oxford

Yuya A. Lin

St Hilda's College, University of Oxford

Hilary 2013

Dedicated to my parents

Thesis Abstract

Olefin Metathesis for Site-Selective Protein Modification

Yuya A. Lin
St. Hilda's College
University of Oxford
Submitted for the degree of Doctor of Philosophy in Organic Chemistry
Hiliary Term 2013

Site-selective protein modification has become an important tool to study protein functions in chemical biology. In the preliminary work, allyl sulfides were found to be reactive substrates in aqueous cross-metathesis (CM) enabling the first examples of protein modification via this approach. In order to access the enhanced CM reactivity of allyl sulfide on proteins, facile chemical methods to install *S*-allyl cysteine on protein surface were developed. In particular, a cysteine-specific allylating reagent – allyl selenocyanate was used on protein substrate for the first time.

The substrate scope of allyl sulfide-tagged proteins and factors that affect the outcome of CM was also investigated. A range of metathesis substrates containing different olefin tether of various lengths were screened; allyl ethers were found to be most suitable as CM partners. By reducing the steric hindrance around the allyl sulfide on protein surface through a chemical spacer, the rate and conversion of metathesis reaction on proteins was greatly enhanced. Moreover, allyl selenides were found to be more reactive than allyl sulfides in CM and enabled reactions with substrates that were previously impossible for the corresponding sulfur-analogue. Through this work, substrate selection guidelines for successful metathesis reaction on proteins were established.

Rapid Se-relayed CM was further investigated through biomimetic chemical access to *Se*-allyl selenocysteine (Seac) via dehydroalanine. On-protein reaction kinetics revealed rate constants of Seac-mediated CM to be comparable or superior to off-protein rates of many current bioconjugations. This CM strategy was applied to histone proteins to install a mimic of acetylated lysine (K9Ac, an epigenetic marker). The resulting synthetic H3 was successfully recognized by antibody that binds natural H3-K9Ac. A Cope-type selenoxide elimination subsequently allowed the removal of such modification to regenerate dehydroalanine. Finally, preliminary research efforts towards metabolic incorporation of allyl sulfide-containing amino acid into proteins, and CM on cell surfaces were discussed

Table of Contents

Acknowledgement iv

Publication List v

Selection of Abbreviations vi

Chapter 1: **An Introduction to Olefin Metathesis for Chemical Modification of Proteins**

1.1 Introduction 1

1.2 Chemical Modification of Proteins via Metal-Mediated Reactions 2

1.3 Olefin-Metathesis on Peptides 7

1.4 Aqueous Olefin-Metathesis 11

1.5 Effect of Allylic Hydroxyl Group in Olefin Metathesis 14

1.6 Allyl Sulfides are Reactive Chemical Handles that Enable Protein Modification by Olefin Metathesis 19

1.6.1 Allyl Sulfides in Aqueous Cross-Metathesis 19

1.6.2 Sulfur-Relayed Cross-Metathesis 20

1.6.3 Strategies to Allow Allyl Sulfide Incorporation into Proteins 24

1.6.4 Applications of Olefin-Metathesis of Allyl Sulfides in Protein Modifications 25

1.7 Thesis Objectives 28

1.8 References 30

Chapter 2: **Enabling Olefin-Metathesis on Proteins: Chemical Routes to *S*-Allyl Cysteine on Protein Surface**

2.1 Introduction 35

2.2 Results and Discussions 36

2.2.1 Direct Allylation 36

2.2.2 Rearrangement and Reduction of Allylic Disulfides and Allyl Selenenylsulfide 38

2.3 Conclusions and Outlook 42

2.4 References 44

2.5 Experimental Procedures 46

2.6 NMR Spectra 66

Chapter 3: **Cross-Metathesis on Proteins: An Investigation of the Substrate Scope and Steric Effects**

3.1 Introduction 69

3.2 Results and Discussions 71

3.2.1 Substrate Scope of Cross-Metathesis on SBL-156Sac (**3.2**) 71

3.2.2 Investigation of the Sterics Around the Alkene Tether on Protein Substrates and Its Effects in Cross-Metathesis 75

3.2.3 Allyl Selenides are Substrates with Cross-Metathesis Reactivity Superior to Allyl Sulfides in Aqueous Media 79

3.2.4 Exploring the Substrate Scope of Cross-Metathesis with Protein Containing an Allyl Selenide Tag 81

3.3 Conclusions and Outlook 83

3.4 References 85

3.5 Experimental Procedures 87

3.6 NMR Spectra 137

Chapter 4: **Rapid Cross Metathesis for Selective Protein Modifications via Chemical Access to *Se*-Allyl Selenocysteine in Proteins**

4.1 Introduction 149

4.2 Results and Discussions 150

4.2.1 Synthesis of *Se*-allyl Selenocysteine (**4.3**) 150

4.2.2 Installation of *Se*-allyl Selenocysteine on SBL 152

4.2.3 Kinetic Studies of Cross-Metathesis on Proteins 154

4.2.4 Cross-Metathesis on SBL-156Seac (**4.5a**) 157

4.2.5 Synthesis of H3-K9Seac (**4.13**) 158

4.2.6 Cross-Metathesis on H3-K9Seac (**4.13**) 159

4.3 Conclusions and Outlook 161

4.4 References 161

4.5 Experimental Procedures 164

4.6 NMR Spectra 199

Chapter 5: **Towards Metabolic Incorporation of Metathesis Handles into Proteins & Olefin Metathesis on Cell Surfaces**

5.1 Introduction 201

5.2 Results and Discussions 202

5.2.1 Metabolic Incorporation of *S*-allyl Homocysteine as Methionine Surrogate 202

5.2.2 Towards Olefin-Metathesis on Cell Surfaces 205

5.3 Conclusions and Outlook 212

5.4 References 213

5.5 Experimental Procedures 214

Acknowledgement

First of all, and most importantly, I would like to thank my parents and brother who provided love and support continuously throughout my time as graduate student in Oxford. Without their encouragement, I would not have been able to accomplish my DPhil study.

Secondly, I would like to express gratitude to my supervisor, Professor Ben Davis, for his guidance and enthusiasm in science. He has also provided numerous opportunities for me to write about and present my work that helped me to develop as a research scientist.

I would also like to give great thanks to Justin Chalker, whom I worked closely with in the olefin metathesis project. I hugely benefited from Justin's day-to-day mentoring in the beginning of my DPhil study, and the interesting scientific discussions we have had. It has been an honor to be able to work with him. It also has been great to be part of the BGD group. During these years, I have met and had the pleasure to work with so many fantastic people from the group. I appreciated their company and assistance in the laboratory. Special thanks go to Dr. Omar Boutureira and Dr. Seung Seo Lee for discussions, technical support and their friendship.

I'm thankful for support from friends at Oxford University Taiwanese Student Society, St Ebbe's Church, and housemates at 11NG. Finally, I am extremely grateful for close friendships I have made in Oxford and Taiwan. The valuable times and conversations I have had with them enriched my student life in Oxford and helped me to get through the tough times in my DPhil study.

Yuya Angel Lin
Hilary 2013
University of Oxford

Publication List

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

(6) Rapid Cross Metathesis for Protein Modifications via Chemical Access to *Se*-Allyl Selenocysteine in Proteins. Lin, Y. A.; Boutureira, O.; Lercher, L.; Bhushan, B.; Davis, B. G. **2013**, submitted.

(5) The Allylic Chalcogen Effect in Olefin Metathesis. Lin, Y. A.; Davis, B. G. *Beilstein J. Org. Chem.* **2010**, *6*, 1219-1228.

(4) 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.

(3) Olefin Metathesis for Site-Selective Protein Modification. Lin, Y. A.; Chalker, J. M.; Davis, B. G. *ChemBioChem* **2009**, *10*, 959-969.

(2) Enabling olefin metathesis on proteins: chemical methods for installation of S-allyl cysteine. Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714-3716.

(1) 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.

Other publications co-authored by Yuya A. Lin:

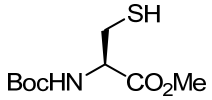
(2) Self-Liganded Suzuki–Miyaura Coupling for Site-Selective Protein PEGylation. Dumas, A.; Spicer, C. D.; Gao, Z.; Takehana, T.; Lin, Y. A.; Yasukohchi, T.; Davis, B. G. *Angew. Chem. Int. Ed.* **2013**, *52*, 3916-3921.

(1) 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.

Selection of Abbreviations

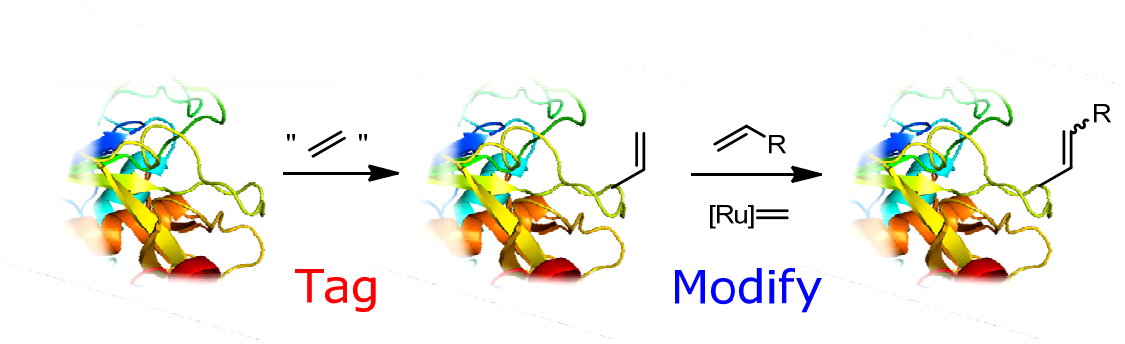
Ac	acetyl	**IR**	infra-red
Ahc	*S*-allyl homocysteine	**Lit.**	literature reference
aq.	aqueous	**m**	meta
Ar	aryl, aromatic	**m. p.**	melting point
Bn	benzyl	**Me**	methyl
Boc	*tert*-butoxycarbonyl	**min**	minute
^{t}Bu	*tert*-butyl	**MSH**	*O*-mesitylenesulfonyl hydroxylamine
C	Celcius	**NMR**	nuclear magnetic resonance
CM	cross-metathesis	**OM**	Olefin metathesis
Cys	cysteine	***p***	*para*
DCM	dichloromethane	**PEG**	polyethylene glycol
Dha	dehydroalanine	**Ph**	phenyl
DMF	dimethylformamide	**^{i}Pr**	isopropyl
DMSO	dimethylsulfoxide	**PTM**	posttranslational modification
DTT	dithiothreitol	**RCM**	ring closing metathesis
ESI	electrospray ionization	**Rf**	retention factor
Et	ethyl	**RT**	room temperature
GFP	green fluorescent protein	**Sac**	*S*-allylcysteine
Glc	glucose	**sat.**	saturated
GlcNAc	*N*-acetyl glucosamine	**SBL**	subtilisin *Bacillus lentus*
h	hour	**Seac**	*Se*-allyl selenocysteine
HRMS	high resolution mass spectrometry	**SM**	starting material
LRMS	low resolution mass spectrometry	**TFA**	trifluoroacetic acid
LC-MS	Liquid chromatography-mass spectrometry	**TLC**	thin layer chromatography

Standard abbreviation for amino acid derivatives are used throughout, for example: -

BocCysOMe = 

Chapter 1

An Introduction to Olefin Metathesis for Chemical Modification of Proteins



1.1 Introduction

Post-translational modification (PTM) of proteins is one of the major factors for complexity of life. The modifications are implicated in important biochemical processes such as cellular communication, protein folding and trafficking, immune responses and many more.[1] Protein modifications are complex to study in nature since proteins of interest are often present in low quantities and in heterogeneous mixtures.[2] Synthetic protein alteration has the potential to widen the scope of studying protein modifications, provided the residue of interest is modified selectively to yield either the natural modification or a suitable mimic.[3] Moreover, unnatural alternations can be made to proteins for desired function via chemical modifications. For example, covalent attachment of fluorescent probes allows tracking of proteins both *in vitro* and *in vivo*;[4] tethering antigens to protein carriers have also shown prominence in the fight against

HIV,[5] cancer,[6] malaria,[7] and other pathogenic bacteria.[8] Chemical modification of proteins has opened up novel and diverse opportunities in chemical biology and medicine. It is not surprising that novel strategies are developing at a rapid pace.

The focus of this thesis is the development and application of olefin metathesis for chemical modifications of proteins. In order to evaluate the viability of this approach, the introductory chapter of this thesis will review a number of transition metal-mediated chemical protein modifications as well as olefin metathesis on peptides and in aqueous systems. The allylic chalcogen effect in metathesis and the preliminary work on olefin metathesis for protein modification will also be discussed in detail. These fundamental studies motivate and build toward the objectives of this thesis which will be summarized at the end of this chapter.

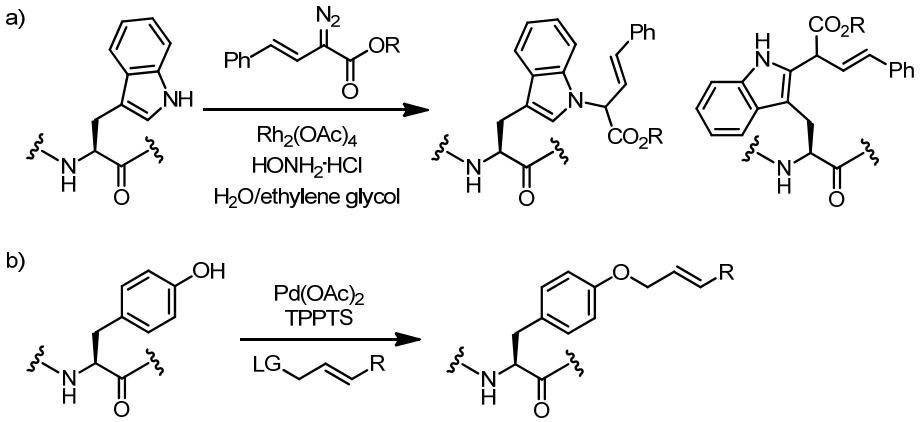
1.2 Chemical Modification of Proteins via Metal-Mediated Reactions

For a reaction to be generally useful for protein modification, it must be able to modify a residue of interest selectively in the presence of other unprotected side chains on the protein surface. The reaction should also be engineered to meet criteria required to prevent protein denaturation. For example, the reaction should be carried out in an aqueous media, low and ambient temperatures, and at or near neutral pH. Furthermore, the reaction must tolerate salts and surfactants that are commonly used to stabilize proteins in solution.[9]

While chemical modifications at some natural residues such as cysteine, lysine and tryptophan can provide some access to modified protein, tuning for desired selectivity among several nucleophilic residues is not a trivial problem and remains highly dependent on the protein in question. An attractive alternative is the use of transition

metal (TM)-catalyzed reactions as these may allow enhanced reactivity and selectivity, for example by targeting non-proteinogenic residues. TM-catalyzed reactions have been widely applied in organic synthesis and enable bond formations that are otherwise difficult or impossible.[10] These reactions are promising candidates for site-selective modification of proteins because their properties and reactivity can be tuned by wide selection of ligands and additives. Moreover, the repertoire of TM-mediated reactions functioning in water has expanded in recent years;[11] while there is still a disparity between the widespread use of transition metals in small molecule synthesis and their modest deployment in protein modification, key advances have been made.[12]

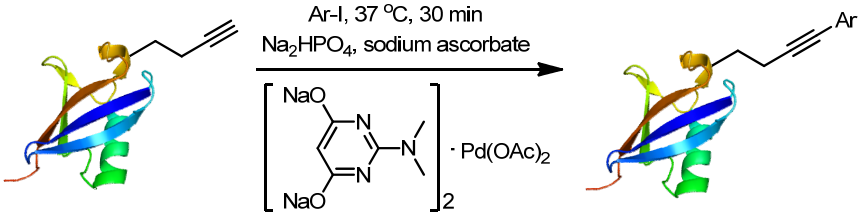
The Francis group has pioneered several modifications at naturally occurring aromatic residues such as tryptophan and tyrosine using transition metals.[12] Alkylation using rhodium carbenoids was achieved on the tryptophan residues of two protein substrates: myoglobin and subtilisin Carlsberg (Scheme 1-1a).[13] From their initial small-molecule studies, both the *N*- and *2*-substituted indole products were observed in the reaction. The widespread successes of palladium π-allyl complexes in synthesis also have found their use in allylation of tyrosine residues to provide proteins modified with fluorescent probes and lipids (Scheme 1-1b).[14]



Scheme 1-1: Metal-mediated protein modification at tryptophan and tyrosine.

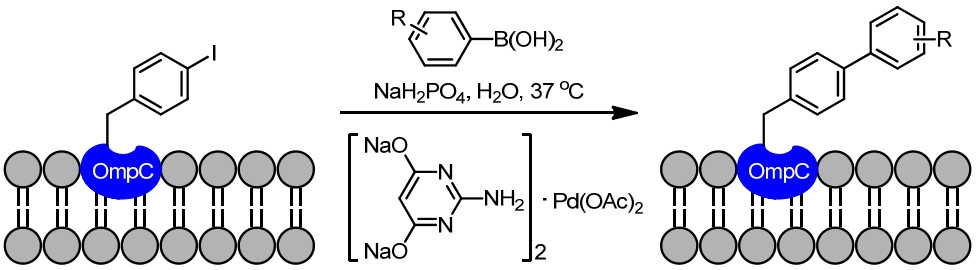
For unnatural residues, site-specific biotinylation via either the Mizoroki-Heck or

Sonogashira reactions have been described at *p*-iodophenylalanine of Ras though only with low yields, 2% and 25% respectively.[15] Significant dehalogenation was also observed in both cases, likely due to slow transmetallation step with the palladium-protein species. Later, Lin and coworkers improved the efficiency and cellular compatibility of the reaction using a water-soluble and copper-free catalytic system. This advancement enabled rapid and efficient attachment of fluorescent probe via Sonogashira coupling on a homopropargylglycine (Hpg)-encoded Ubiquitin and on *Escherichia coli* (*E. coli*) surface (Scheme 1-2).[16]



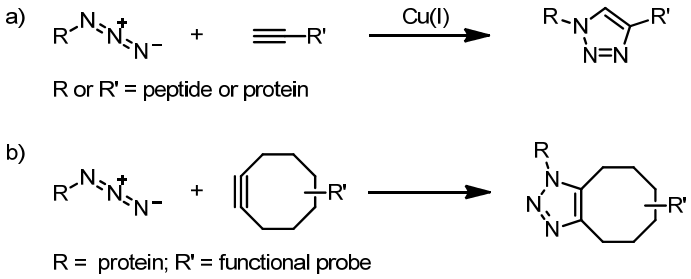
Scheme 1-2: Sonogashira coupling at homopropargylglycine.

Progress of protein modification via Suzuki cross-coupling has been developing rapidly over recent years. One of the first reports on a peptide system was by the Hamachi and coworkers, who introduced *p*-iodophenylalanine at position 34 of the WW domain of Pin1 by solid phase peptide synthesis; subsequent Suzuki coupling was performed using water-soluble Na_2PdCl_4 and fluorescent aryl boronic acids in 50% aqueous glycerol (v/v).[17] The reaction proceeded efficiently with up to 90% conversion by HPLC. Suzuki cross-coupling on T4 lysozyme containing genetically incorporated *p*-borono-phenylalanine with iodinated bodipy molecule using Pd(0) dibenzylidene acetone have also been reported by Brustad *et al.* in moderate yield (30%).[18] In our own group, with recent discovery of a reactive and water-soluble palladium complex, $Pd(OAc)_2(ADHP)_2$, and well-established genetic incorporation of *p*-iodophenyl- alanine in proteins, efficient Suzuki cross-couplings on proteins have been demonstrated *in vitro*[19] and on *E. coli* cell surfaces (Scheme 1-3).[20]



Scheme 1-3: Suzuki cross-coupling on cell surface.

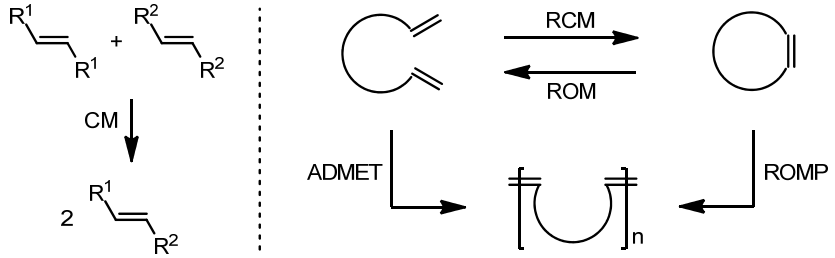
Finally, the azide-alkyne [3+2] cycloadditions have been employed widely in bioconjugation chemistry.[21] Their application on peptides and protein was first reported by Meldal[22] and Sharpless[21a], respectively, with a mild copper-catalyzed variant of the reaction (Scheme 1-4a). Due to potential problems of copper's toxicity for application *in vivo*, copper-free variants have also been developed. The use of strained cyclooctyne derivatives was among the most efficient variants (Scheme 1-4b).[23] These advances have further enabled protein modification via azide-alkyne [3+2] cycloadditions for cell-surface labeling[23a,24] and selective imaging inside living cells.[25]



Scheme 1-4: Azide-alkyne [3+2] cycloaddition.

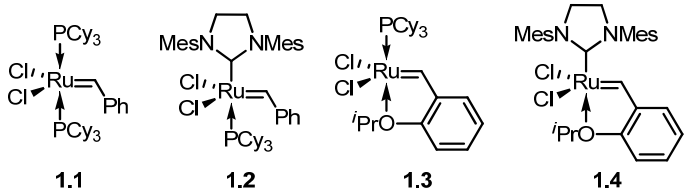
One glaring omission from the list of metal-mediated reactions for protein modification is olefin metathesis, undoubtedly a versatile C-C bond forming reactions (Scheme 1-5). The synthetic strategies that it offers have enabled the construction of biologically active natural products which are otherwise difficult to synthesize.[26] Olefin metathesis is a strong candidate as another bioorthogonal reaction (one that is chemoselective with respect to the functional groups present in proteins) for protein modification owing to the high air-stability, exquisite chemoselectivity and tolerance to various functional groups such as amides, alcohols and carboxylic acids, exhibited by ruthenium-based

metathesis complexes. However, a number of challenges remain for development of efficient olefin metathesis on protein surface.



Scheme 1-5: Selected types of olefin metathesis. ADMET = acyclic diene metathesis, CM = cross-metathesis, RCM = ring-closing metathesis, ROM = ring-opening metathesis, ROMP = ring-opening metathesis polymerization.

First, there must be reliable strategies to introduce olefinic substrates on protein surfaces, either chemically or genetically. Despite the first example of genetic incorporation of homoallylglycine (Hag), as one potential substrate, into proteins in 1998 (where the authors had astutely recognized the potential to modify such residue via olefin metathesis), it had not been possible to realize this potential due to limited development of olefin metathesis in water at the time.[27] Thus a second immediate challenge is the requirement that olefin metathesis must work efficiently under conditions compatible with proteins: for example, in buffered aqueous solutions at or below 37 °C. Since most metathesis reactions with conventional catalysts (Scheme 1-6) involve long reaction time in organic solvent at elevated temperature, engineering a system that allows efficient aqueous metathesis is critical for protein modification. In order to further evaluate the viability of olefin metathesis on protein, research developments of such reaction on amino acids and peptidic system are discussed in the following section.

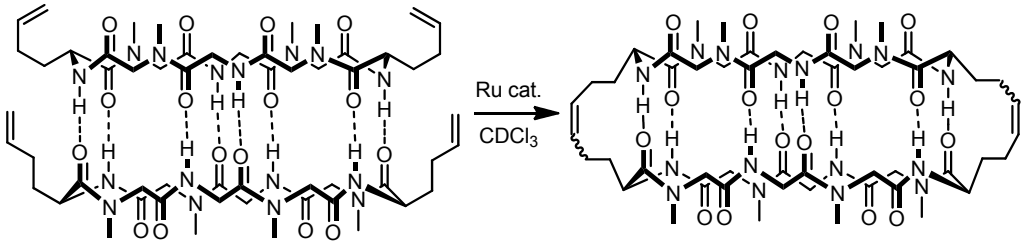


Scheme 1-6: Conventional ruthenium-based metathesis catalysts.

1.3 Olefin-Metathesis on Peptides

The overriding motivation for using olefin metathesis to modify peptides is to install non-labile carbon-carbon bonds. The research aims of these modifications are typically in two directions. First, olefin metathesis is used to cross-link peptides thus stabilizing their secondary structures, in turn to impart better metabolic stability and higher binding affinity towards biological targets. Second, post-synthetic labeling with functional tags such as carbohydrates could be achieved with olefin metathesis to provide non-labile analogues of natural post-translational modifications. Key examples of these modifications will be discussed in turn.

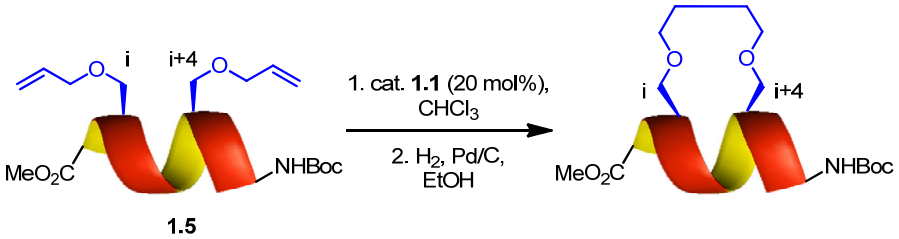
One of the earliest examples of olefin metathesis on peptides was reported by Clark and Ghadiri.[28] The sequential CM and RCM between two cyclic peptides containing Hag residues were demonstrated (Scheme 1-7). The resulting structure is a β-sheet-like cylindrical structure. This example is remarkable in that it is not only one of the earliest examples of olefin metathesis in peptide synthesis, but an example of CM templated by intermolecular hydrogen bonding. This example highlights the influence of peptide pre-organization on the outcome of metathesis reaction and the ability to "covalently capture" self-assembled secondary structure.



Scheme 1-7: CM-RCM of cyclic peptides assisted by intermolecular hydrogen bonding.

Ring-closing metathesis (RCM), probably the most developed metathesis reaction, has proven useful in the synthesis of unnatural amino acids and peptidomimetics.[29] Exploratory work in this area delineated the scope of ring size and side chain useful in

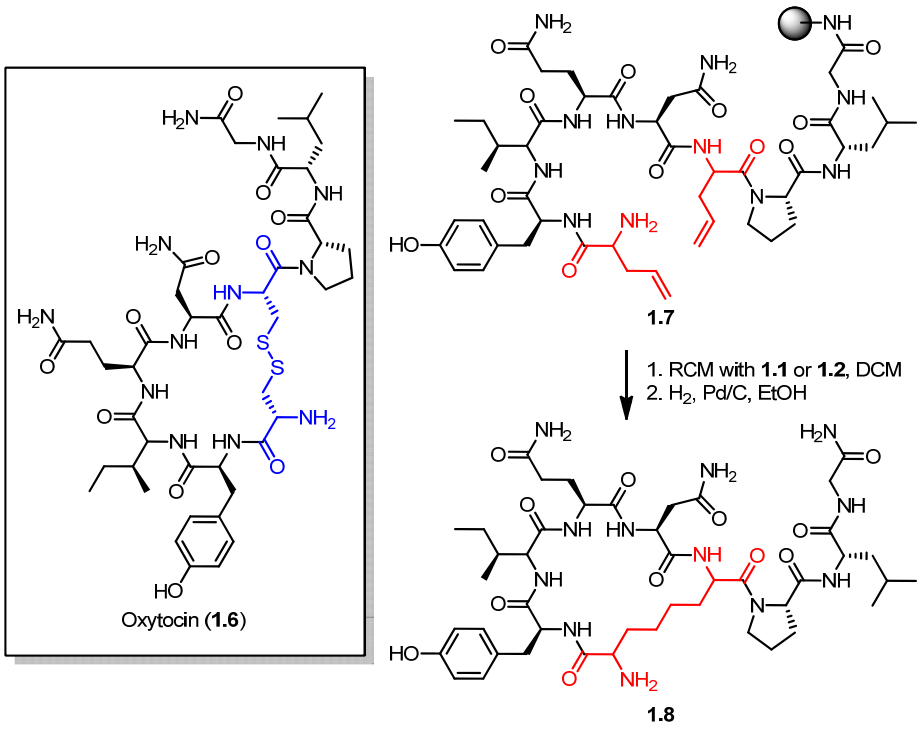
the synthesis of cyclic amino acids.[30] Several β-turn analogues that mimic the natural role of β-turns in stabilizing short peptides could be accessed via this chemistry.[31] β-turns are common secondary structures found in all proteins, and often serve as a recognition elements on the protein surface, making them attractive targets for RCM.[32] RCM has also been used to cross-link α-helices to induce structural rigidity. In one of the first examples, RCM followed by hydrogenation of a heptapeptide α-helix containing *O*-allyl ethers at its *i* and *i*+4 positions, which were aligned on the turn of the helix, provided the cross-linked α-helix (Scheme 1-8).[33] A systematic study of the cross-link position and helix stability was later undertaken by Verdine.[34]



1.5: BocVal-Ase-Leu-Aib-Val-Ase-LeuOMe

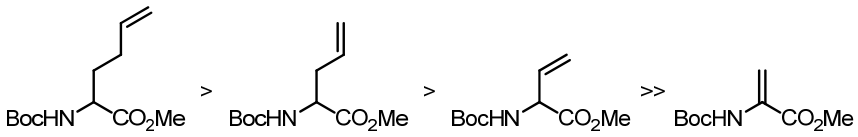
Scheme 1-8: Cross-linked helices by RCM.

Disulfide bridges are important in stabilizing the secondary and tertiary structures of peptides and proteins. However, these are labile under reducing conditions. As a result, many scholars have opted for RCM as a strategy to create the carbon equivalent of disulfide bridges, the so-called "dicarba analogue". Building on Grubbs' early use of RCM to replace disulfides by all-carbon analogues,[30a,b] Vederas and co-workers synthesized the dicarba analogue of the hormone oxytocin (**1.6**).[35] The linear peptide (**1.7**), in which the two cysteine residues were replaced with allylglycine, was constructed by solid phase peptide synthesis. RCM was then carried out on-resin followed by cleavage from the solid support. Hydrogenation afforded the cyclic oxytocin analogue **1.8** (Scheme 1-9) which was shown to have a longer half-life *in vivo* than its natural counterpart.



Scheme 1-9: Synthesis of an oxytocin dicarba analogue by RCM.

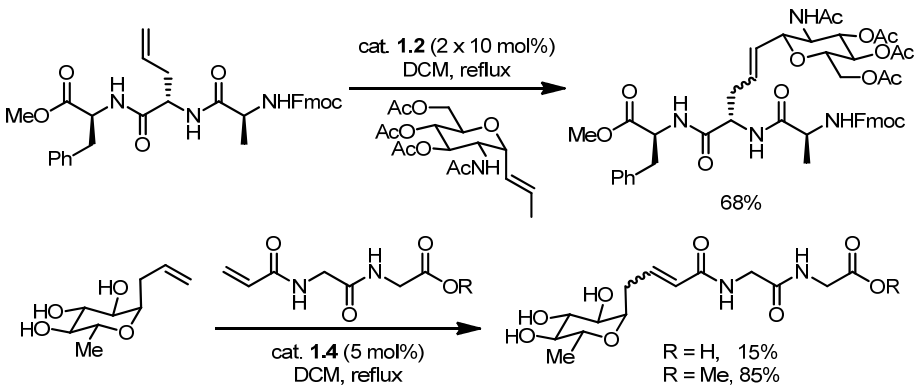
Apart from Ghadiri's cross-metathesis of preorganized peptides, the examples discussed so far are essentially restricted to RCM. To modify proteins post-translationally by cross-metathesis, it is instructive to consider peptide models and the relative reactivity of prospective side chains. Studies by Gibson on CM of unsaturated amino acids revealed that reactivity in metathesis increases with the length of the side chain when catalyst **1.1** was used (Scheme 1-10).[36] It is noteworthy that Schrock type molybdenum catalysts promote cross-metathesis of unreactive partners such as vinylglycine.[37] However, the air- and moisture-sensitivity of this catalyst does not, at present, bode well for its application on metathesis of unprotected peptides in aqueous media.



Scheme 1-10: Relative CM reactivity of a selection of unsaturated amino acids.

Biologically important glycopeptides have also been synthesized by cross-metathesis between sugars and peptides. Early examples by the Roy laboratory demonstrated the

feasibility of glycopeptide synthesis by cross-metathesis of allyl glycosides with glycine and glycylglycine derivatives.[38] McGarvey later refined the choice of protecting groups and reaction conditions to favor cross-metathesis over self-metathesis, with the most promising results being obtained with catalyst **1.2**.[39] Blechert recently revisited glycopeptide synthesis by cross-metathesis and detailed the compatibility of metathesis catalysts with certain amino acid residues.[40] Histidine and tryptophan were problematic, whereas methionine was notably tolerated in metathesis reactions when Hoveyda-Grubbs second generation catalyst (**1.4**) was used. In addition, unprotected carbohydrates were demonstrated as useful substrates in metathesis reactions in the report. Unprotected peptides were also used, but a free carboxy terminus was sometimes found to be detrimental (Scheme 1-11). Collectively, these examples motivate the use of cross-metathesis as a route to non-labile *C*-glycopeptides and glycoprotein analogues of the natural *O*-linked and *N*-linked glycopeptides and glycoproteins.



Scheme 1-11: Glycopeptide syntheses by cross-metathesis.

Although these examples were all carried out in organic solvent, the results showed promising functional group compatibility and substrate scope for olefin metathesis in peptidic systems. The relative reactivity of various unsaturated amino acids in metathesis serves as a useful guide for amino acid selection for incorporation into proteins. Successful results with Hag were particularly promising since it can be genetically incorporated into proteins. Finally, the general strategy of using metathesis to

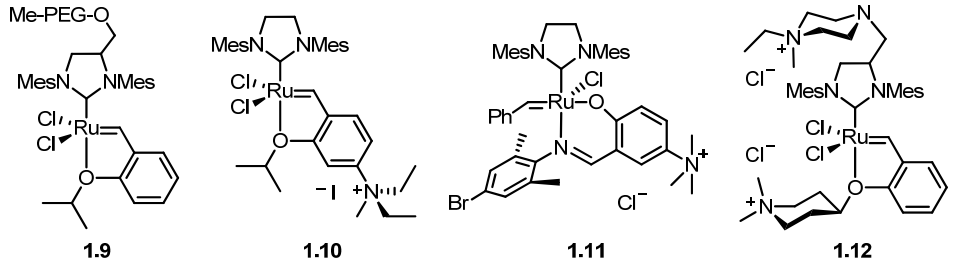
stabilize peptide secondary structure or to covalently tether carbohydrates can, in principle, be extended to proteins. In order for these strategies to be applicable to proteins, developments in aqueous olefin metathesis were required and are briefly reviewed next.

1.4 Aqueous Olefin-Metathesis

Although aqueous olefin metathesis is primarily driven by the ultimate goal of sustainable chemistry, progress in this field is immediately relevant to olefin metathesis on biomolecules that requires aqueous media. Selected catalysts useful in homogenous olefin metathesis are discussed below, along with other notable strategies for aqueous olefin metathesis that might prove useful in bioconjugation.

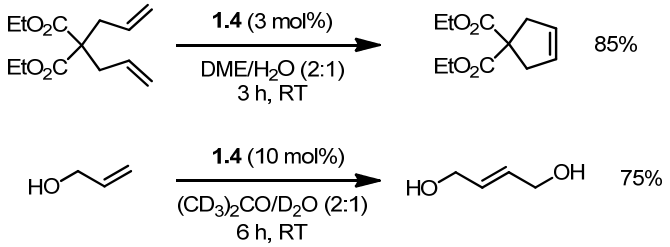
Several water-soluble (pre)catalysts have been developed for olefin metathesis. The design of these water-soluble catalysts are mainly based on ligands containing hydrophilic poly(ethylene glycol) (PEG) or quaternary ammonium groups. A selection of water-soluble ruthenium complex for metathesis is given in Scheme 1-12. The PEGylated catalyst **1.9** was the first reported to promote RCM of water-soluble dienes and homodimerization of allyl alcohol in neat water with good to excellent conversions.[41] Grela and coworkers have reported the installation of quaternary ammonium chloride on the alkylidene portion of the catalyst (**1.10**).[42] The ammonium group imparts increased water solubility to the precatalyst and weakens the O–Ru coordination, resulting in fast initiation. Although this catalyst is only slightly soluble in neat water (< 1 mg/mL) it is efficient in RCM, ene–yne, and cross metathesis of simple substrates in protic solvents, including water. Raines and coworkers have also reported an addition to water-soluble metathesis catalysts containing a salicylaldimine ligand with a pendant ammonium group (**1.11**). This catalyst is active in RCM of a range of

substrates in methanol and methanol-water mixture at a slightly higher temperature (55 °C).[43] The synthetic routes to ammonium tagged ruthenium complexes either involve difficult preparation or purification. In 2012, the Grela laboratory tackled this problem by using an "on-site" quaternization method that enabled facile and efficient preparation of water-soluble metathesis catalysts such as **1.12**.[44] In addition, the presence of a second ammonium group in **1.12**, dramatic improvement in water solubility is observed (35 mg/mL).



Scheme 1-12: Examples of water-soluble metathesis catalysts.

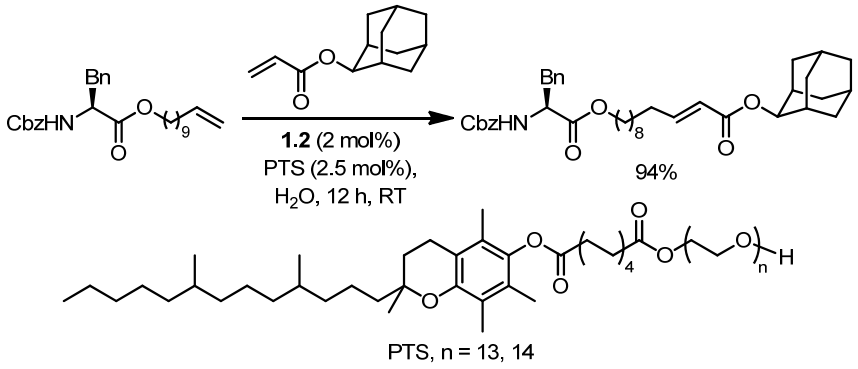
Other studies have focused on the use of conventional catalysts (**1.1**–**1.4**) in aqueous media with the use of a co-solvent. Blechert reported good conversions in RCM and modest reactivity in CM for simple olefins with catalysts **1.2** or **1.4** in aqueous DMF or methanol.[45] The use of dimethoxyethane or acetone in water to aid aqueous RCM of several dienes and homodimerization of allyl alcohol in the presence of catalyst **1.4** was later reported by Raines (Scheme 1-13).[46] Despite these developments, homogeneous CM in water is largely limited to simple, reactive olefins such as allyl alcohol.



Scheme 1-13: Metathesis in homogeneous aqueous media with catalyst **1.4**.

The substrate scope in aqueous CM with commercially available catalyst **1.2** has been expanded to more complex olefins by emulsion chemistry in a recent report by Lipshutz

et al.[47] After screening a range of surfactants such as Triton X-100, Brij 30, PEG-600 and SDS, the nonionic amphiphile PTS was found to be the most effective. A series of water-insoluble olefins and electron-deficient substrates were shown to undergo efficient CM at room temperature with catalyst **1.2** (2 mol %) and PTS (2.5 mol %) (Scheme 1-14). The efficient CM demonstrated by this method is likely due to increased effective concentrations of reactants and catalyst in the micelles formed by the surfactant. It is possible that, for some proteins, nanometer micelles could both solubilize the metathesis catalyst and still allow reaction on a protein surface but this will be highly dependent on the surface hydrophobicity and charge of the protein in question.



Scheme 1-14: CM reactions in the presence of PTS.

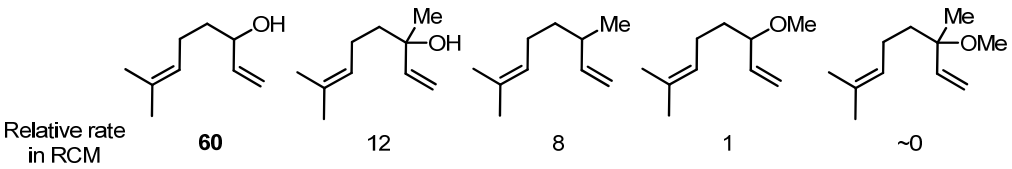
While the development of water-soluble catalysts for olefin metathesis is a milestone in aqueous chemistry, there is not yet a catalyst that is completely resistant to decomposition in water. Substituted olefins have been shown to slow down decomposition because alkylidenes are more stable in water than methylidenes, but the rate of metathesis might also be compromised.[48] Decomposition of metathesis catalysts in water and protic solvents has been investigated, but is still not well understood.[49] In most cases, decomposition results in ruthenium hydrides, which are inactive in metathesis but prone to mediate carbon–carbon double bond isomerization and migration.[50] These side reactions highlight the need for a comprehensive investigation

of the coordination sphere and its ability to attenuate or prevent catalyst decomposition. An immediate compromise to improving catalyst stability in water, especially for the purpose of protein modification, is to use a substrate that can react at a much greater rate than catalyst decomposition.

Among several factors that influence the outcome of metathesis reaction (e.g. the nature of the catalyst and the steric hindrance around the alkene), the effects of nearby heteroatoms are intriguing. The activating effect of allylic alcohols in olefin metathesis has been found to be an important factor that enabled efficient metathesis in many examples. Collectively, these reports suggest that allylic alcohols can indeed modulate the rate of olefin metathesis, and give useful insight in search of a reactive substrate that could potentially enable olefin metathesis on proteins. A selection of pertinent examples is discussed in the following section.

1.5 Effect of Allylic Hydroxyl Group in Olefin Metathesis

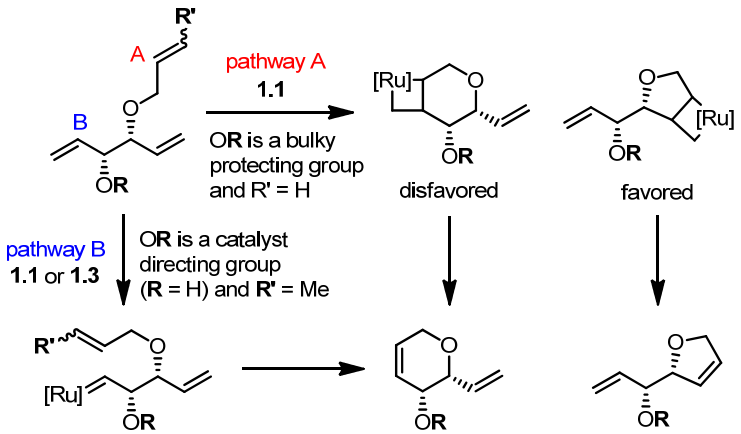
For more than a decade, the directing effect of an allylic hydroxy group in olefin metathesis has been known and taken advantage of by many organic chemists for efficient synthesis of natural products. This particular phenomenon of allylic hydroxy groups in RCM was first identified by Hoye and Zhao in 1999.[51] In this work, the influence of both steric and electronic character of the allylic substituent of linalool and related analogues in RCM was assessed. The free hydroxy group on linalool greatly enhanced RCM relative to the corresponding methyl ether or unsubstituted analogues (Scheme 1-15). This activating effect was marked and initially surprising given that *tert*-butylethylene, containing a fully substituted allylic center, was reported to be almost inert to reaction with catalyst **1.1**.[52]



Scheme 1-15: Allylic hydroxyl activation in RCM.

A number of possible explanations for the rate acceleration due to allylic hydroxy group in olefin metathesis have been proposed. For example, pre-association of the alcohol at ruthenium centre through the rapid and reversible exchange of alkoxy for chloride ligand, or alcohol for phosphine ligand. The hydrogen bonding between the allylic hydroxyl group on the substrate and the chloride ligand on the catalyst had also been postulated. Given the reaction conditions employed by Hoye, hydrogen bonding was most likely the explanation for enhanced reactivity in metathesis albeit no further experimental evidence at the time.

Nevertheless, following Hoye and Zhao's observations, several organic chemists have further studied the effect of allylic alcohols and ethers in metathesis. In 2007, Schmidt demonstrated the preparation of enantiomerically pure dihydropyrans and dihydrofurans bearing an unsaturated olefin tether through a ring size-selective RCM of a triene.[53] In this study, it was found that trienes containing bulky hydroxyl protecting group at the allylic position cyclized selectively to dihydrofurans, whereas the free alcohol yielded 6-membered rings with very high selectivity (Scheme 1-16).

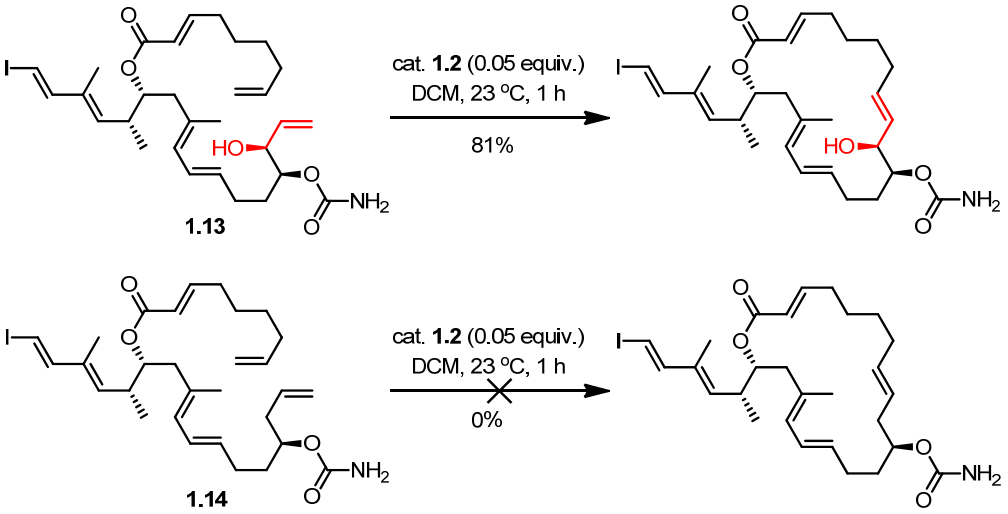


Scheme 1-16: The influence of different OR groups on ring size-selectivity.

This useful selectivity was attributed to the directing effect of the allylic hydroxy group. When the allylic alcohol was protected with a bulky group, the catalyst reacted preferentially at the less hindered allyl ether (pathway A where R′ = H). The formation of 4,6-bicyclic intermediate is apparently disfavored with catalyst **1.1** leading to largely the formation of the dihydrofuran product. It should be noted that when a more reactive second-generation catalyst (**1.2**) the selectivity for the dihydrofuran product is reduced. The selectivity was also found to be highly dependent on the size of protecting group. Significant decrease in ring size selectivity was observed when smaller protecting group such as a methyl group was used. Furthermore, when no protecting group was used (i.e. R = R′ = H) the RCM reaction was unselective and resulted in 1:1 mixture of 5- and 6-membered ring products. The selectivity for dihydropyran formation was therefore tuned by substitution at the terminal position of the allyl ether (e. g. R′ = Me) which directed the catalyst to react via pathway B. The authors were able to obtain significantly improved conversions and yields while maintaining high selectivity for dihydropyran by using, instead of **1.1**, the Hoveyda-Grubbs first generation catalyst (**1.3**) in which the active catalyst species is stabilized by a hemilabile benzylidene ligand. The activating effect of the allylic hydroxy group in RCM was further supported by the dramatic decrease in conversion when the free OH group was protected (i.e. R ≠ H, R′ = Me) in pathway B.

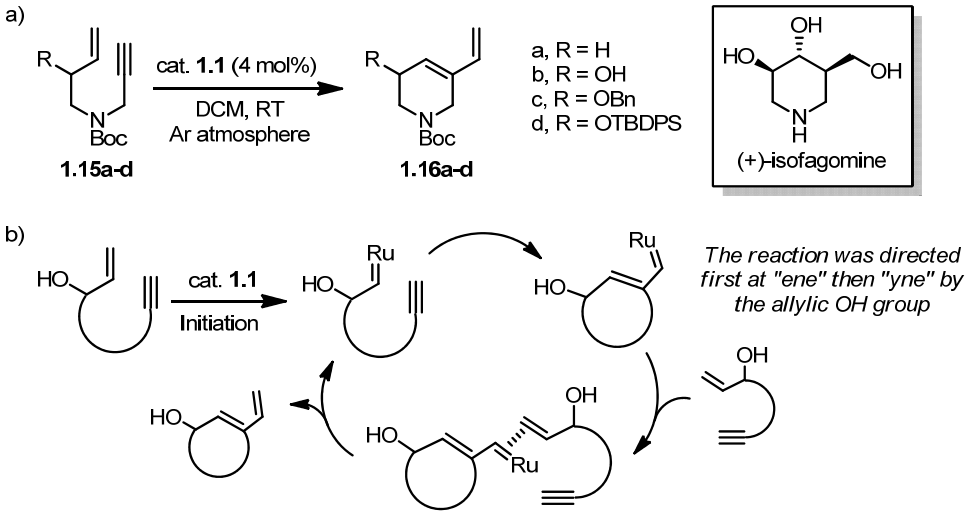
Other examples have also emerged during target syntheses. In the synthesis of Palmerolide A analogues by Nicolaou and coworkers, compound **1.13** was found to undergo smooth macrocyclization via RCM whereas **1.14** (lacking the allylic hydroxyl group) failed to form the desired macrocycle under the same conditions (Scheme 1-17).[54] When **1.14** was treated with more stringent conditions, decomposition and/or polymerization occurred. These observations suggest that the presence of allylic

hydroxyl group in the molecule was crucial for enhancing the reactivity under the mild RCM conditions required by the potentially labile natural product scaffold.



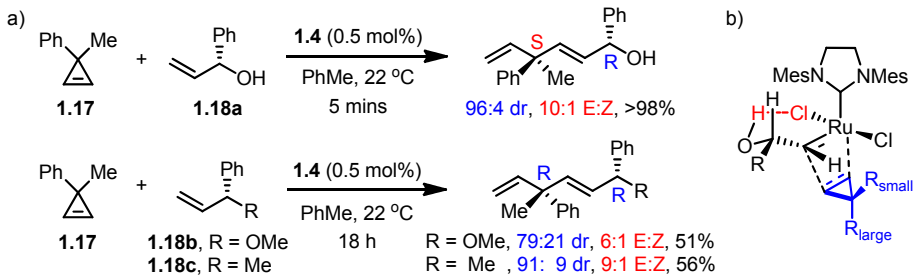
Scheme 1-17: Synthesis of Palmerolide A precursors by RCM.

Enhancements by an allylic hydroxyl group have also been found in ring-closing enyne metathesis. Studies by Imahori *et al.* revealed that the ring-closing enyne metathesis of terminal alkynes containing an allylic hydroxyl group proceeded smoothly without the ethylene atmosphere that is generally required to drive such reactions (Scheme 1-18a).[55] The compound containing the allylic hydroxyl group (**1.15b**) cyclized to the corresponding diene product with quantitative yield in 1.5 hours whereas **1.15a**, without the allylic hydroxyl group, required 41 hours to afford **1.16a** with a diminished yield of 32%. Reduction in yield was also observed with increasing bulk of the protecting group, evident from the metathesis of substituted allyl ethers **1.15c** and **1.15d** (44% and 7% respectively). The authors took advantage of the efficient RCM assisted by allylic alcohol and synthesized (+)-isofagomine from its corresponding acyclic starting material. Associated mechanistic studies suggested that the reaction proceeded via an "ene-then-yne" pathway, further suggesting that rate acceleration is likely due to the directing effect of the allylic hydroxy group on substrates (Scheme 1-18b).



Scheme 1.18: a) Accelerated ring-closing enyne metathesis by allylic hydroxyl group. b) Proposed mode of action by allylic hydroxy group in metathesis.

As discussed earlier, the most likely explanations for observed rate accelerations by allylic hydroxyl groups is hydrogen-bonding. Hoveyda and coworkers recently utilized the hydrogen-bonding of allylic hydroxy groups to ruthenium catalyst for stereoselective ring-opening cross-metathesis (ROCM) (Scheme 1-19).[56] The activating effect from the allylic hydroxy group in metathesis is prominent in this example. Cyclopropene **1.17** with enantiomerically enriched allylic alcohol **1.18a** underwent complete ROCM after 5 minutes with high diastereomeric ratio (dr) (96:4) and *E*:*Z* selectivity (10:1) favoring the *S*,*R*-diasteromer. In contrast, the reaction of methyl ether **1.18b** and the methyl analogue **1.18c** is far less effective (51% and 56% conversion respectively in 18 hours) with lower and opposite stereoselectivity in favor of the *R*,*R*-diastereomers. The observed stereoselectivity can be explained by intramolecular hydrogen-bonding depicted in Scheme 1-19b, where the substituted group of the stereogenic center is situated away from the bulky mesityl groups. On coordination of the cyclopropene to catalyst, formation of the metallacyclobutane species with the larger R group pointing away from the main bulk of the catalyst is preferred thus resulting in the observed stereoselectivity.



Scheme 1-19: a) Effect of allylic hydroxyl group on the rate and steroselectivity of ROCM. b) Proposed H-bonded ruthenium complex for stereoselective ROCM.

It should be noted that all of the illustrated examples regarding allyl hydroxyl activation in olefin metathesis so far are predominantly secondary allylic alcohols. The lack of primary allyl alcohols examples is due to the fact that these can dehydrogenate at elevated temperature in the presence of ruthenium-based metathesis complexes, forming metathesis-unreactive ruthenium hydrides.[49,57]

1.6 Allyl Sulfides are Reactive Chemical Handles that Enable Protein Modification by Olefin Metathesis†

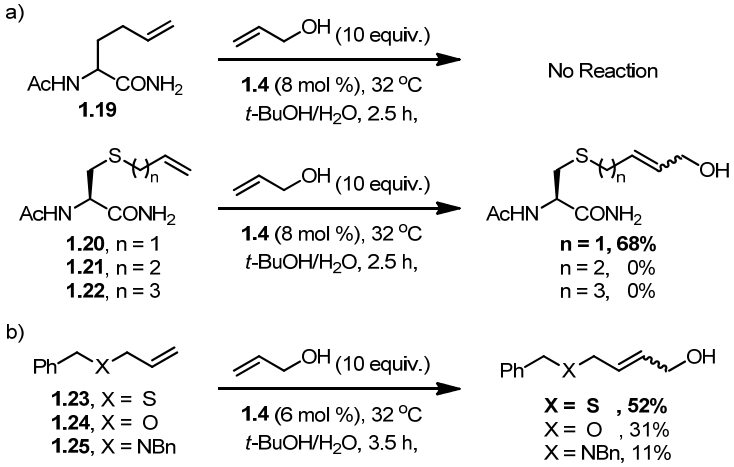
1.6.1. Allyl Sulfides in Aqueous Cross-Metathesis

While there are many examples of allylic alcohols and ethers in metathesis, examples with allyl sulfide substrates were until recently noticeably few. This was to be expected as there have been several cases of olefin metathesis in which sulfides were problematic.[58] In the exploratory work on aqueous metathesis, cross-metathesis of unsaturated amino acids with allyl alcohol in the presence of catalyst **1.4** was

† The exploratory work on protein modification by olefin metathesis described in this section was carried out as the Part II project during the 4th year (2007-2008 academic year) of my undergraduate chemistry degree (MChem) at Keble College, University of Oxford. For additional details and experimental procedures), the reader is referred to the Part II thesis and the publication that resulted from this work (Reference 57).

investigated. Amino acids such as homoallylglycine **1.19** and alkene-tethered cysteine derivatives **1.20**-**1.22** were investigated as substrates of interest due to their potential for incorporation into proteins if they proved to be reactive in metathesis.[59]

Unexpectedly, *S*-allyl cysteine derivative **1.20** was the only substrate that afforded a synthetically useful amount of CM product whereas the reaction of the all carbon analogue homoallylglycine (**1.19**) and sulfide derivatives, *S*-butenyl and *S*-pentenyl cysteine (**1.21** and **1.22**, respectively), failed to work under identical conditions in aqueous media (Scheme 1-20a). In order to compare the relative CM reactivity between other allylic heteroatom derivatives, further studies were carried out on the CM of allyl benzyl ether (**1.24**) and allyl dibenzylamine (**1.25**), but the allyl sulfide analogue **1.23** emerged as the most reactive substrate in aqueous media (Scheme 1-20b).

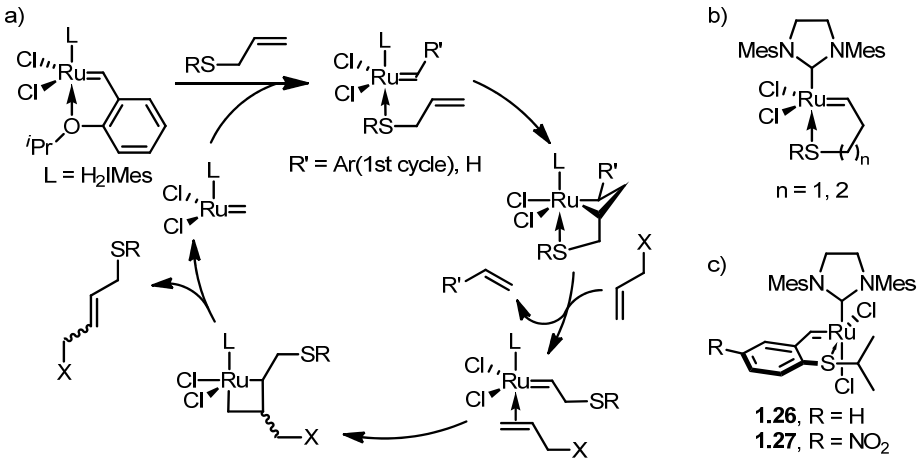


Scheme 1-20: a) Efficient cross-metathesis of *S*-allyl cysteine. b) Comparison of relative reactivity between allylic heteroatom derivatives.

1.6.2 Sulfur-Relayed Cross-Metathesis

These results were interpreted considering early metathesis work that had gained strong evidence for the benefits of suitable Lewis bases as pre-coordinating motifs (or 'relays'). For example, through a "carbonyl-relayed" mechanism, Fürstner has rationalized an observed favoring of macrocyclization over oligomerization, which has a more

favorable change in reaction entropy.[60] It has been suggested that coordination by such groups to ruthenium may bring tethered alkenes closer in proximity to the alkylidene-metal center allowing more ready reaction by exploiting effective molarity. Similarly, the rate enhancement displayed by allyl sulfides could be explained with a sulfur-relayed mechanism where the sulfur atom coordinates to the ruthenium of the second generation precatalyst. A favorable ligand exchange from the benzylidene isopropoxy group of **1.4** to sulfide may be rationalized by the preferred soft acid-base interaction between sulfur and ruthenium. Such coordination brings together the metal alkylidene and alkene substrate leading to rapid formation of the reactive alkylidene species, which initiates subsequent metathesis events (Scheme 1-21a). The coordination depicted in the metallocyclobutane in Scheme 1-21a is based on X-ray analysis of related intermediates.[61] It is currently uncertain whether a similar effect would be observed for phosphine-containing catalysts such as **1.1** and **1.2**. If it would, the ligand exchange might be expected to be slower based on this logic.

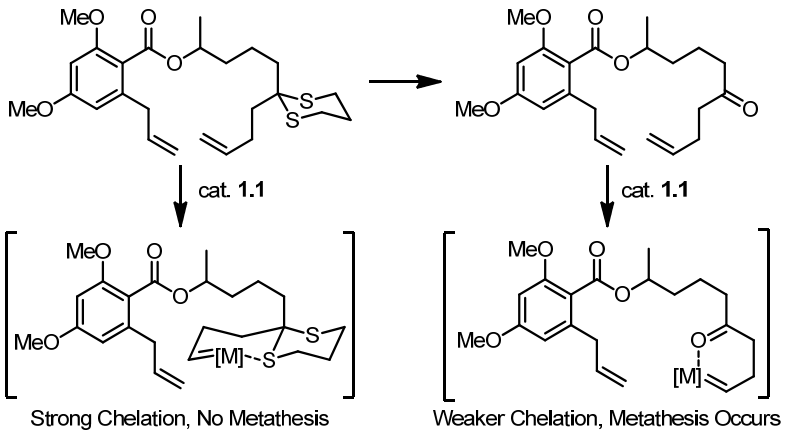


Scheme 1-21: a) Sulfur assisted cross-metathesis. b) Putative non-productive sulfur-chelate with catalyst **1.4**. c) Example of latent sulfur-chelated Ru-complex.

The unproductive CM observed for **1.21** and **1.22** is consistent with Fürstner's observation and can be explained with a chelation model, in which a five- and six-membered chelate, unreactive to metathesis is formed, respectively (Scheme 1-21b).

Notably, Grela and Lemcoff have synthesized the sulfur analogue of the Hoveyda-type catalyst **1.26** and **1.27**,[62] and found that these only initiate at much higher reaction temperatures. The inertness of five-membered sulfur chelates was so compelling that even the *para*-nitro substitution on benzylidene ligand, known to have a significant activation effect in *O*- and *N*-chelated ruthenium catalysts, could not improve the apparent retardation in reaction initiation at lower temperatures.[62c] Their findings are in agreement with the unproductive CM of **1.21** that was observed at room temperature.

The efficient metathesis demonstrated by allyl sulfides in water was initially surprising since sulfur-containing molecules have been well-documented to be detrimental for many transition metal-catalyzed reactions, presumably due to its formation of stable chelates with metal complexes and potential reduction of metals. This phenomenon was so profound, in fact, that Fürstner had suggested "…a quite general incompatibility of the ruthenium based metathesis catalysts with substrates containing sulfur(II) donor sites." in one of his macrocycle synthesis via RCM.[58d] It was proposed in the report that the proximal sulfur atom on the dithiane unit sequestered the ruthenium carbene complex via an unproductive chelation (Scheme 1-22). Nevertheless, upon transformation of the dithiane to the ketone, the weaker coordination from the carbonyl group to the metal center instead provided a beneficial directing effect to allow successful RCM.



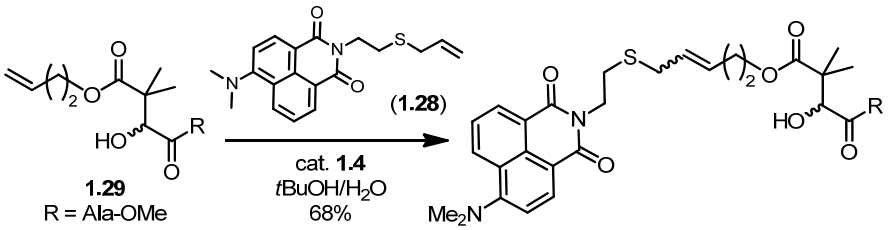
Scheme 1-22: Stable sulfur chelate formation that inhibits RCM.

Detrimental chelation clearly proved not to be the case for allyl sulfides due to having the shorter alkene chain. Any chelate formed with allyl sulfides is apparently too strained to sequester the catalyst effectively and rapid turnover is observed instead. While allyl sulfides had indeed previously been used in olefin metathesis,[63] the highlight of these prior studies had been the tolerance of the catalyst for sulfur-bearing substrates. The results from our laboratory showed that allyl sulfide are not simply tolerated, they can *enhance* the rate of olefin metathesis. The rate enhancement observed for allyl sulfides may be related to other cases in which such remote functionality affects the rate of olefin metathesis.[64] Our findings combined with a reevaluation of several examples (prior and since our initial work) suggest that this might be a general beneficial 'allylic chalcogen effect' that is more keenly felt under certain conditions and is particularly beneficial in aqueous conditions.

The beneficial effects in CM observed from allylic heteroatom coordination have proven to be broadly useful, especially in aqueous olefin metathesis. It should be noted, however, that in such reactions, in necessarily protic solvent, catalyst degradation does occur, especially in aqueous media and in the presence of an allylic alcohol.[49,51,57] It was suspected that one key enabling aspect of allyl sulfide substrates is that reactions proved to be sufficiently high in turnover frequency that catalyst decomposition is effectively outcompeted, likely a key aspect of its success in water.

Ever since the realization of this useful reactivity of allyl sulfides in olefin metathesis, there have been several applications in small molecule synthesis. For example, Hunter *et al.* took advantage of the efficient aqueous CM demonstrated by allyl sulfides to explore the reversibility of such reaction in dynamic combinatorial chemistry in water with the prospect of generating a library of novel nucleic acid-binding compounds.[65] Loh and coworkers have successfully coupled allyl sulfide-derived fluorescent probes (such as

1.28) to olefin-tethered alanine derivatives (such as **1.29**, synthesized using Mukaiyama aldol reaction), via CM adopting the reaction conditions previously optimized in our laboratory (Scheme 1-23).[66] In the latter case, the allyl sulfide based probe was chosen precisely for its enhanced metathesis reactivity in water.



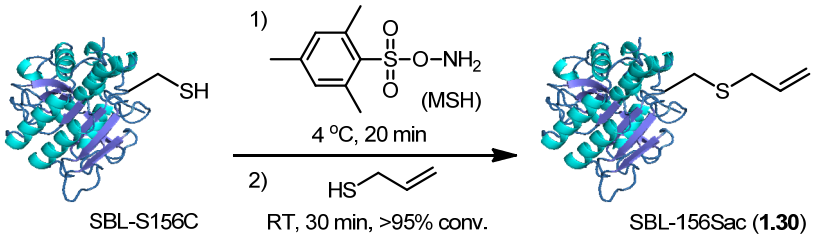
Scheme 1-23: Functionalization of Mukaiyama aldol product by aqueous CM.

1.6.3 Strategies to Allow Allyl Sulfide Incorporation into Proteins

The promising results from the model CM reactions on amino acids immediately motivated the pursuit of *S*-allyl cysteine (Sac) incorporation into proteins. While direct allylation of cysteine might be one obvious chemical route to installing Sac on proteins, non-selective allylation at other nucleophilic residues such as lysine and histidine is one potential major drawback. A cysteine-specific approach was required.

Dehydroalanine (Dha) is an effective Michael-type acceptor for a range of thiol nucleophiles. Multiple routes to Dha-containing proteins have been reported, ranging from early examples of sulfonate elimination (limited in their application due to the required high pH or elevated temperatures),[67] oxidative elimination of phenylseleno-cysteine by treatment with hydrogen peroxide[68] or through various methods from cysteine.[69] A range of thioether-functionalized proteins could then be accessed via the subsequent conjugate addition of the corresponding thiol nucleophile to Dha. In a similar manner, *S*-allyl cysteine was installed on a single-cysteine mutant of subtilisin from *Bacillus lentus* (SBL-S156C) via 1,4-addition of allyl thiol to Dha (obtained, for example, via oxidative elimination of cysteine using *O*-mesitylenesulfonylhydroxyl-

amine (MSH)) (Scheme 1-24).[59]



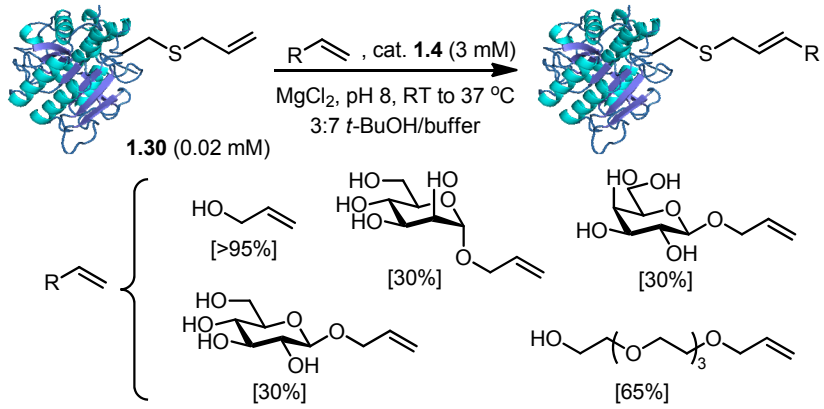
Scheme 1-24: Access to *S*-allyl cysteine on proteins via 1,4-addition to Dha.

Genetic incorporation of some unsaturated amino acids as methionine surrogates in methionine auxotrophic *E. coli* is documented.[27] These surrogates may be successful due to their similarity in size and/or electronic character to methionine. Other than homoallylglycine mentioned earlier, both homopropargylglycine[27b] and azidohomo-alanine[70] have been incorporated as methionine surrogates and have been used extensively in bioconjugation chemistry. Our laboratory considered *S*-allyl cysteine as a potential methionine surrogate. Trial expression of a single methionine mutant of a model glycosidase has revealed unambiguous incorporation but at very low levels.[59] This promising result has served as a starting point in our laboratory to further investigate into other structurally related amino acids as analogues of methionine for incorporation into proteins that will combine the power of olefin metathesis with even more direct control of modification site via genetic manipulation.

1.6.4 Applications of Olefin-Metathesis of Allyl Sulfides in Protein Modification

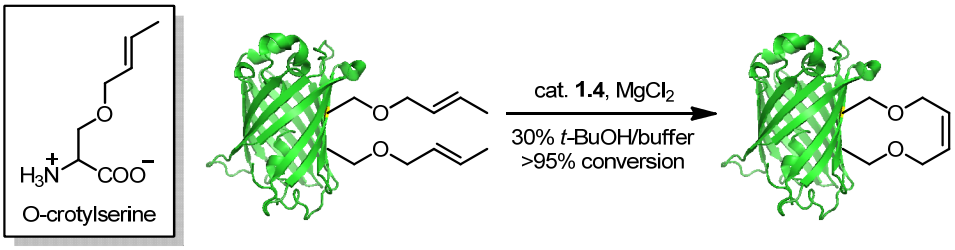
With ready access to *S*-allyl cysteine on protein surface, CM on SBL-156Sac (**1.30**) with allyl alcohol was initially carried in the presence of catalyst **1.4** and *tert*-butanol as a co-solvent.[59] LC-MS analysis of the reaction revealed largely the unmodified protein. Upon addition of simply the catalyst to the protein solution, an adduct speculated to be a metalloprotein derived from metathesis with **1.4** was observed. This species was inactive in CM possibly due to non-productive chelation of ruthenium by nearby Lewis basic

amino acids. Reports from Fürstner *et al.* and others groups have shown that similar unproductive chelation in metathesis could be disrupted using sarcrificial hard Lewis acid such as $Ti(O^iPr)_4$.[71] As a milder and biocompatible alternative, $MgCl_2$ was chosen as an additive in the reaction buffer with the intention to disrupt any such non-productive chelation to ruthenium. Critically, with the inclusion of $MgCl_2$ in the reaction buffer, CM with allyl alcohol proceeded efficiently to >95% conversion at room temperature. This key breakthrough enabled CM with more biologically relevant modifications such as allyl ethers of monosaccharides and oligo(ethylene glycol) (Scheme 1-25); nonetheless, more complex and electronically-demanding metathesis substrates remained challenging.



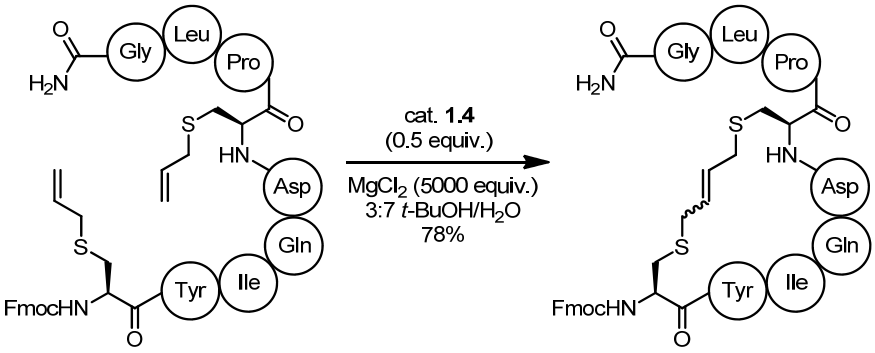
Scheme 1-25: Cross-metathesis on SBL-156Sac (**1.30**)

Following our first report on protein CM, Ai *et al.* investigated Sac and *O*-crotylserine incorporation on GFP.[72] Whilst Sac was found to be oxidized to the corresponding sulfoxide in *Saccharomyces cerevisiae*, *O*-crotylserine displayed both the beneficial effect of allylic heteroatom in metathesis and a more water-stable propagating catalytic species generated from the substituted alkene as compared with the methylidene that results from terminal olefins.[48] Indeed, the resulting RCM on a double *O*-crotylserine mutant with catalyst **1.4** and 30% *t*-BuOH proceeded with near-complete conversion after 5 hours (Scheme 1-26).



Scheme 1-26: RCM on GFP, enabled by allylic chalcogen-containing amino acids

The Vederas group, with long-standing interest in cyclic peptides, has very recently synthesized the dicarba analogues of oxytocin and crotalphine from RCM of *S*-allyl cysteine containing peptides (Scheme 1-27).[73] The authors noted that the presence of the allylic sulfur heteroatom was vital for efficient aqueous RCM; the carbon and oxygen analogues underwent RCM in only low yields. Again, the addition of $MgCl_2$ proved essential. The rate enhancement from allyl sulfide-containing peptides has enabled efficient cyclization using sub-stoichiometric amount of catalyst **1.4** to provide milligram quantities of valuable biologically relevant peptide analogues in a cost-effective manner.



Scheme 1-27: Synthesis of *S*-allyl cysteine analogue of oxytocin via aqueous RCM.

The enhanced reactivity of allyl sulfides (and other allyl chalcogenides) has enabled olefin metathesis on protein surfaces. As well as increasing interests in utilizing metathesis of allyl sulfides for organic synthesis, many opportunities, accordingly, in incorporation of olefin-containing amino acid into proteins have also been created. Other privileged chemical handles that may assist metathesis in the same way as allyl sulfides may also be discovered. However, challenges in olefin metathesis on proteins remain and are worthy of discussion. Some of these include the steric sensitivity of

olefin metathesis that may preclude modification at more hindered protein sites, and the need for organic co-solvent to solubilize conventional metathesis catalysts is limiting its applications in certain intolerant proteins and cellular systems. The goals of this thesis are outlined next.

1.7 Thesis Objective

In order to take advantage of the reactive nature of allyl sulfide handle in olefin metathesis, multiple chemical routes to installing *S*-allyl cysteine on proteins are investigated. These approaches are complementary to the initial route via dehydroalanine to accommodate wider protein targets. The strategy is based on the direct allylation of cysteine residues on a protein surface. In particular, a cysteine-specific allylating reagent – allyl selenocyanate is used on protein substrate for the first time. The results from these research efforts are discussed in Chapter 2.

The necessary accessibility of amino acid metathesis partners has not been fully assessed. The observations on the CM of allyl sulfides alongside reports regarding the activating effect of allylic alcohols and ethers in olefin metathesis suggest that such phenomenon may be general for allylic chalcogens. Therefore, the CM of allyl selenides is studied as a logical extension to this hypothesis. Furthermore, an investigation of various factors that may affect the outcome of cross-metathesis on protein surface is conducted to allow olefin metathesis to be executed more routinely in bioconjugation. A relatively limited set of metathesis partners was screened in our preliminary report. The full scope and functional group tolerance is yet to be determined not only by the protein alone, but also by the metathesis partner and its unique steric and electronic character. The results from these research efforts are discussed in Chapter 3.

Chapter 4 describes the investigation on chemical installation of *Se*-allyl selenocysteine (Seac) into proteins. This work was motivated by the broader substrate scope and faster reaction displayed by allyl selenide-containing proteins shown in Chapter 3. Instead of lengthy genetic manipulations to install selenocysteine into proteins, an alternate and biomimetic chemical route via 1,4-addition of allyl selenolate to dehydroalanine residue was explored. Remarkably, the rapid cross-metathesis shown by Seac-tagged protein is comparable to the rate of the so-called 'click reactions' in bioconjugation chemistry. This prompts for further investigation on the reaction kinetics. The study of CM reaction kinetics on protein substrate was carried out to provide directly relevant kinetic parameters of the protein conjugation reaction.

In Chapter 5, preliminary investigation in metabolic incorporation of olefin-containing amino acid such as *S*-allyl homocysteine is described. *S*-allyl homocysteine was incorporated as a potential methionine surrogate in protein expression using methionine auxotrophic cell line to provide a more general strategy to install metathesis active handles for protein modifications. In addition, an initial research effort to labelling cell surfaces via cross-metathesis, in collaboration with the Hamachi laboratory at Kyoto University, is also described in this chapter.

Ultimately, with advances in water-soluble variants of metathesis complexes, together with increasing numbers of strategies for unsaturated amino acid incorporation, and a more detailed understanding of optimal metathesis partners, the opportunity for *in vivo* protein modification and associated functional investigations by olefin metathesis may become reality. These questions are driving our current efforts to assess fully the scope of olefin metathesis in protein modification and its use in biology.

1.8 References

(1) (a) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto Jr., G. J. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372. (b) Walsh, C. T. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*; Roberts and Co. Press: Englewood, Colorado, 2005.

(2) Gamblin, D. P.; van Kasteren, S. I.; Chalker, J. M.; Davis, B. G. *FEBS J.* **2008**, *275*, 1949-1959.

(3) Davis, B. G. *Science* **2004**, *303*, 480-482.

(4) (a) Steiner, R. F.; Edelhoch, H. *Chem. Rev.* **1962**, *62*, 457-83. (b) Giepmans, B. N. G.; Adams, S. R.; Ellisman, M. H.; Tsien, R. Y. *Science* **2006**, *312*, 217-224.

(5) 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. *J. Virol.* **2008**, *82*, 6359-6368.

(6) Krug, L. M.; Ragupathi, G.; Ng, K. K.; Hood, C.; Jennings, H. J.; Guo, Z.; Kris, M. G.; Miller, V.; Pizzo, B.; Tyson, L.; Baez, V.; Livingston, P. O. *Clin. Cancer Res.* **2004**, *10*, 916-923.

(7) Seeberger, P. H.; Werz, D. B. *Nature* **2007**, *446*, 1046-1051.

(8) Verez-Bencomo, V.; Fernandez-Santana, V.; Hardy, E.; Toledo, M. E.; Rodriguez, M. C.; Heynngnezz, L.; Rodriguez, A.; Baly, A.; Herrera, L.; Izquierdo, M.; Villar, A.; Valdes, Y.; Cosme, K.; Deler, M. L.; Montane, M.; Garcia, E.; Ramos, A.; Aguilar, A.; Medina, E.; Torano, G.; Sosa, I.; Hernandez, I.; Martinez, R.; Muzachio, A.; Carmenates, A.; Costa, L.; Cardoso, F.; Campa, C.; Diaz, M.; Roy, R. *Science* **2004**, *305*, 522-525.

(9) Chi, E. Y.; Krishnan, S.; Randolph, T. W.; Carpenter, J. F. *Pharm. Res.* **2003**, *20*, 1325-1336.

(10) Hegedus, L. S. *Transition Metals in the Synthesis of Complex Organic Molecules*; 2nd Ed., University Science Books, Sausalito, California, 1999.

(11) (a) Herrerías, C. I.; Yao, X.; Li, Z.; Li, C.-J. *Chem. Rev.* **2007**, *107*, 2546-2562. (b) Li, C.-J. *Chem. Rev.* **2005**, *105*, 3095-3166.

(12) Antos, J. M.; Francis, M. B. *Curr. Opin. Chem. Biol.* **2006**, *10*, 253-262.

(13) Antos, J. M.; Francis, M. B. *J. Am. Chem. Soc.* **2004**, *126*, 10256-10257.

(14) Tilley, S. D.; Francis, M. B. *J. Am. Chem. Soc.* **2006**, *128*, 1080-1081.

(15) (a) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Kigawa, T.; Sakamoto, K.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Tachibana, K.; Yokoyama, S. *ChemBioChem* **2006**, *7*, 134-139. (b) Kodama, K.; Fukuzawa, S.; Nakayama, H.; Sakamoto, K.; Kigawa, T.; Yabuki, T.; Matsuda, N.; Shirouzu, M.; Takio, K.; Yokoyama, S.; Tachibana, K. *ChemBioChem* **2007**, *8*, 232-238.

(16) Li, N.; Lim, R. K. V.; Edwardraja, S.; Lin, Q. *J. Am. Chem. Soc.* **2011**, *133*, 15316-15319.

(17) Ojida, A.; Tsutsumi, H.; Kasagi, N.; Hamachi, I. *Tetrahedron Lett.* **2005**, *46*, 3301-3305.

(18) Brustad, E.; Bushey, M. L.; Lee, J. W.; Groff, D.; Liu, W.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2008**, *47*, 8220-8223.

(19) (a) Chalker, J. M.; Wood, C. S. C.; Davis, B. G. *J. Am. Chem. Soc.* **2009**, *131*, 16346-16347. (b) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2011**, *47*, 1698-1700.

(20) (a) Spicer, C. D.; Triemer, T.; Davis, B. G. *J. Am. Chem. Soc.* **2011**, *134*, 800-803. (b) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2013**, *49*, 2747-2749.

(21) (a) Wang, Q.; Chan, T. R.; Hilgraf, R.; Fokin, V. V.; Sharpless, K. B.; Finn, M. G. *J. Am. Chem. Soc.* **2003**, *125*, 3192-3193. (b) Sen Gupta, S.; Kuzelka, J.; Singh, P.; Lewis, W. G.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2005**, *16*, 1572-1579. (c) Lin, P.-C.; Ueng, S.-H.; Tseng, M.-C.; Ko, J.-L.; Huang, K.-T.; Yu, S.-C.; Adak, A. K.; Chen, Y.-J.; Lin, C.-c. *Angew. Chem. Int. Ed.* **2006**, *45*, 4286-4290. (d) van Kasteren, S. I.; Kramer, H. B.; Jensen, H. H.; Campbell, S. J.; Kirkpatrick, J.; Oldham, N. J.; Anthony, D. C.; Davis, B. G. *Nature* **2007**, *446*, 1105-1109.

(22) Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002**, *67*, 3057-3064.

(23) (a) Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2004**, *126*, 15046-15047. (b) Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chemical Biology* **2006**, *1*, 644-648.

(24) Link, A. J.; Tirrell, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 11164-11165.

(25) (a) Baskin, J. M.; Bertozzi, C. R. *QSAR Comb. Sci.* **2007**, *26*, 1211-1219. (b) 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. *Proc. Natl. Acad. Sci.* **2007**, *104*, 16793-16797.

(26) (a) Connon, S. J.; Blechert, S. *Angew. Chem. Int. Ed.* **2003**, *42*, 1900-1923. (b) Grubbs, R. H. *Handbook of metathesis*; Wiley-VCH: Weinheim, Germany, 2003.

(c) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. *Angew. Chem. Int. Ed.* **2005**, *44*, 4490-4527. (d) Fürstner, A. *Chem. Commun.* **2011**, *47*, 6505-6511.

(27) (a) van Hest, J. C. M.; Tirrell, D. A. *FEBS Letters* **1998**, *428*, 68-70. (b) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 1282-1288.

(28) Clark, T. D.; Ghadiri, M. R. *J. Am. Chem. Soc.* **1995**, *117*, 12364-5.

(29) Brik, A. *Adv. Synth. Catal.* **2008**, *350*, 1661-1675.

(30) (a) Miller, S. J.; Grubbs, R. H. *J. Am. Chem. Soc.* **1995**, *117*, 5855-5856. (b) Miller, S. J.; Blackwell, H. E.; Grubbs, R. H. *J. Am. Chem. Soc.* **1996**, *118*, 9606-9614. (c) Rutjes, F. P. J. T.; Schoemaker, H. E. *Tetrahedron Lett.* **1997**, *38*, 677-680.

(31) (a) Pernerstorfer, J.; Schuster, M.; Blechert, S. *Chem. Commun.* **1997**, 1949-1950. (b) Fink, B. E.; Kym, P. R.; Katzenellenbogen, J. A. *J. Am. Chem. Soc.* **1998**, *120*, 4334-4344.

(32) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1-109.

(33) (a) Blackwell, H. E.; Grubbs, R. H. *Angew. Chem. Int. Ed.* **1998**, *37*, 3281-3284. (b) 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. *J. Org. Chem.* **2001**, *66*, 5291-5302.

(34) Schafmeister, C. E.; Po, J.; Verdine, G. L. *J. Am. Chem. Soc.* **2000**, *122*, 5891-5892.

(35) (a) Stymiest, J. L.; Mitchell, B. F.; Wong, S.; Vederas, J. C. *Org. Lett.* **2003**, *5*, 47-49. (b) Stymiest, J. L.; Mitchell, B. F.; Wong, S.; Vederas, J. C. *J. Org. Chem.* **2005**, *70*, 7799-7809.

(36) (a) Gibson, S. E.; Gibson, V. C.; Keen, S. P. *Chem. Commun.* **1997**, 1107-1108. (b) Biagini, S. C. G.; Gibson, S. E.; Keen, S. P. *J. Chem. Soc., Perkin Trans. 1* **1998**, 2485-2499.

(37) Brummer, O.; Ruckert, A.; Blechert, S. *Chem.--Eur. J.* **1997**, *3*, 441-446.

(38) Hu, Y.-J.; Roy, R. *Tetrahedron Lett.* **1999**, *40*, 3305-3308.

(39) McGarvey, G. J.; Benedum, T. E.; Schmidtmann, F. W. *Org. Lett.* **2002**, *4*, 3591-3594.

(40) Bruechner, P.; Koch, D.; Voigtmann, U.; Blechert, S. *Synth. Commun.* **2007**, *37*, 2757-2769.

(41) Hong, S. H.; Grubbs, R. H. *J. Am. Chem. Soc.* **2006**, *128*, 3508-3509.

(42) (a) Michrowska, A.; Gułajski, Ł.; Kaczmarska, Z.; Mennecke, K.; Kirschning, A.; Grela, K. *Green Chem.* **2006**, *8*, 685-688. (b) Gułajski, Ł.; Michrowska, A.;

Narożnik, J.; Kaczmarska, Z.; Rupnicki, L.; Grela, K. *ChemSusChem* **2008**, *1*, 103-109.

(43) Binder, J. B.; Guzei, I. A.; Raines, R. T. *Adv. Synth. Catal.* **2007**, *349*, 395-404.

(44) Skowerski, K.; Szczepaniak, G.; Wierzbicka, C.; Gułajski, Ł.; Bieniek, M.; Grela, K. *Catal. Sci. Technol.* **2012**, *2*, 2424-2427.

(45) Connon, S. J.; Rivard, M.; Zaja, M.; Blechert, S. *Adv. Synth. Catal.* **2003**, *345*, 572-575.

(46) Binder, J. B.; Blank, J. J.; Raines, R. T. *Org. Lett.* **2007**, *9*, 4885-4888.

(47) Lipshutz, B. H.; Aguinaldo, G. T.; Ghorai, S.; Voigtritter, K. *Org. Lett.* **2008**, *10*, 1325-1328.

(48) Lynn, D. M.; Mohr, B.; Grubbs, R. H.; Henling, L. M.; Day, M. W. *J. Am. Chem. Soc.* **2000**, *122*, 6601-6609.

(49) Dinger, M. B.; Mol, J. C. *Organometallics* **2003**, *22*, 1089-1095.

(50) Dinger, M. B.; Mol, J. C. *Eur. J. Inorg. Chem.* **2003**, 2827-2833.

(51) Hoye, T. R.; Zhao, H. *Org. Lett.* **1999**, *1*, 1123-1125.

(52) Ulman, M.; Grubbs, R. H. *Organometallics* **1998**, *17*, 2484-2489.

(53) Schmidt, B.; Nave, S. *Adv. Synth. Catal.* **2007**, *349*, 215-230.

(54) Nicolaou, K. C.; Leung, G. Y. C.; Dethe, D. H.; Guduru, R.; Sun, Y.-P.; Lim, C. S.; Chen, D. Y. K. *J. Am. Chem. Soc.* **2008**, *130*, 10019-10023.

(55) Imahori, T.; Ojima, H.; Yoshimura, Y.; Takahata, H. *Chem. Eur. J.* **2008**, *14*, 10762-10771.

(56) Hoveyda, A. H.; Lombardi, P. J.; O'Brien, R. V.; Zhugralin, A. R. *J. Am. Chem. Soc.* **2009**, *131*, 8378-8379.

(57) Werner, H.; Grünwald, C.; Stüer, W.; Wolf, J. *Organometallics* **2003**, *22*, 1558-1560.

(58) (a) Armstrong, S. K.; Christie, B. A. *Tetrahedron Lett.* **1996**, *37*, 9373-9376. (b) Mascareñas, J. L.; Rumbo, A.; Castedo, L. *J. Org. Chem.* **1997**, *62*, 8620-8621. (c) Shon, Y.-S.; Lee, T. R. *Tetrahedron Lett.* **1997**, *38*, 1283-1286. (d) Fürstner, A.; Seidel, G.; Kindler, N. *Tetrahedron* **1999**, *55*, 8215-8230.

(59) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642-9643.

(60) Fürstner, A.; Langemann, K. *Synthesis* **1997**, 792-803.

(61) Feldman, J.; Murdzek, J. S.; Davis, W. M.; Schrock, R. R. *Organometallics* **1989**, *8*, 2260-2265.

(62) (a) Ben-Asuly, A.; Tzur, E.; Diesendruck, C. E.; Sigalov, M.; Goldberg, I.; Lemcoff, N. G. *Organometallics* **2008**, *27*, 811-813. (b) Kost, T.; Sigalov, M.; Goldberg, I.; Ben-Asuly, A.; Lemcoff, N. G. *J. Organomet. Chem.* **2008**, *693*, 2200-2203. (c) Tzur, E.; Szadkowska, A.; Ben-Asuly, A.; Makal, A.; Goldberg, I.; Wozniak, K.; Grela, K.; Lemcoff, N. G. *Chem. Eur. J.* **2010**, *16*, 8726-8737.

(63) (a) Spagnol, G.; Heck, M.-P.; Nolan, S. P.; Mioskowski, C. *Org. Lett.* **2002**, *4*, 1767-1770. (b) Toste, F. D.; Chatterjee, A. K.; Grubbs, R. H. *Pure Appl. Chem.* **2002**, *74*, 7-10.

(64) Lin, Y. A.; Davis, B. G. *Beilstein J. Org. Chem.* **2010**, *6*, 1219-1228.

(65) Hunter, L.; Condie, G. C.; Harding, M. M. *Tetrahedron Lett.* **2010**, *51*, 5064-5067.

(66) Alam, J.; Keller, T. H.; Loh, T.-P. *J. Am. Chem. Soc.* **2010**, *132*, 9546-9548.

(67) (a) Strumeyer, D. H.; White, W. N.; Koshland, D. E., Jr. *Proc. Natl. Acad. Sci.* **1963**, *50*, 931-935. (b) Weiner, H.; White, W. N.; Hoare, D. G.; Koshland, D. E., Jr. *J. Am. Chem. Soc.* **1966**, *88*, 3851-3859.

(68) Wang, J.; Schiller, S. M.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2007**, *46*, 6849-6851.

(69) (a) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053. (b) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem. Sci.* **2011**, *2*, 1666-1676. (c) Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 1835-1839.

(70) Kiick, K. L.; Saxon, E.; Tirrell, D. A.; Bertozzi, C. R. *Proc. Natl. Acad. Sci.* **2002**, *99*, 19-24.

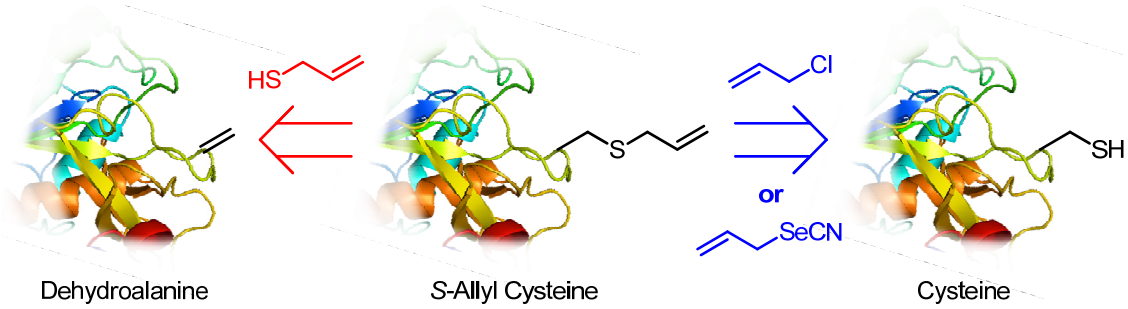
(71) (a) Fürstner, A.; Langemann, K. *J. Am. Chem. Soc.* **1997**, *119*, 9130-9136. (b) Wipf, P.; Weiner, W. S. *J. Org. Chem.* **1999**, *64*, 5321-5324. (c) Yang, Q.; Xiao, W.-J.; Yu, Z. *Org. Lett.* **2005**, *7*, 871-874.

(72) Ai, H.-W.; Shen, W.; Brustad, E.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 935-937.

(73) Cochrane, S. A.; Huang, Z.; Vederas, J. C. *Org. Biomol. Chem.* **2013**.

Chapter 2

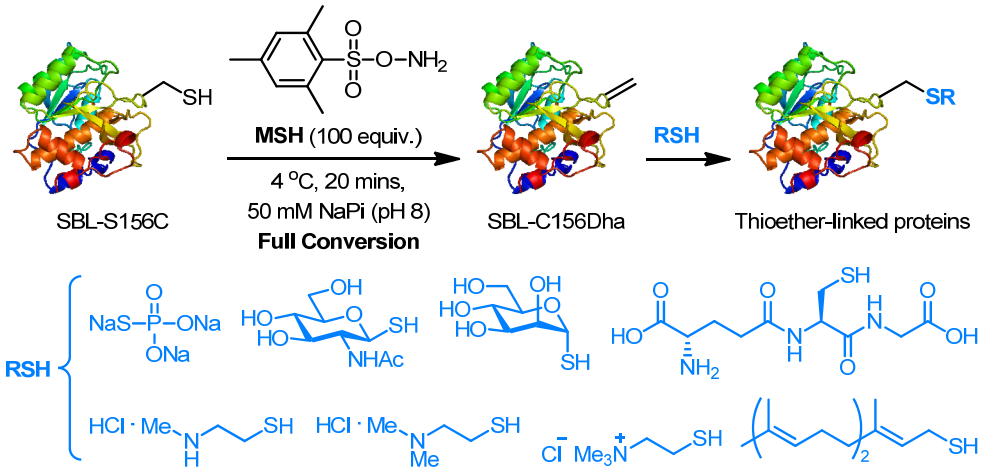
Enabling Olefin-Metathesis on Proteins: Chemical Routes to *S*-Allyl Cysteine on Protein Surface



2.1 Introduction

From the exploratory work on aqueous cross-metathesis, as described in Section 1.6, allyl sulfides were disclosed as privileged substrates in aqueous cross-metathesis. In order to take advantage of this useful reactivity in protein modifications, the non-proteogenic amino acid *S*-allyl cysteine (Sac) was used as the reactive handle for site-specific attachment of simple monosaccharides and oligo(ethylene glycol). Sac on protein was initially obtained from the corresponding dehydroalanine (Dha),[1] a useful chemical handle for the synthesis of a range of thio-ether modified protein conjugates (Scheme 2-1).[2] While this route may be fast and chemoselective, thiol addition to Dha in peptides typically results in epimeric mixtures.[3] Selectivity of the addition will be highly dependent on the local stereochemical environment. Although the diastereoselectivity was not an immediate concern for evaluating the metathesis

reaction on proteins, developing complimentary routes to single diasteromer of Sac would be useful.



Scheme 2-1: Thioether functionalized proteins via 1,4 addition of thiols to Dha.

2.2 Result and Discussion[†]

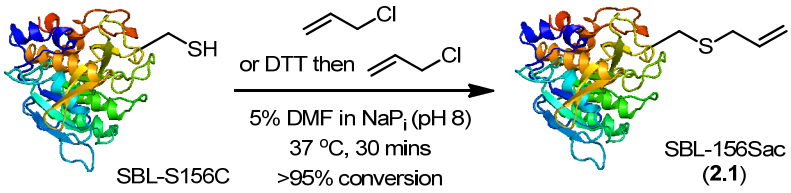
2.2.1 Direct Allylation

To avoid epimerization at cysteine, direct allylation of cysteine on protein surface is one obvious chemical approach as such reaction on cysteine-containing peptides with allyl halide is well-documented.[4] The allylated products have been subsequently coupled to proteins through linkers[4a] or by native chemical ligation.[5] Despite the precedence on peptidic systems, direct allylation of cysteine on protein surfaces has not been reported. This could be due to the limited solubility of allyl halides in water or potential complications in selectivity. Indeed, such alkylation is in risk of non-selective alkylation at other nucleophilic residues such as lysine and histidine but this can

[†] This project was carried out as part of a productive collaboration with Justin M. Chalker, a former DPhil student in the group. Experimental results that were obtained by Mr. Chalker are indicated in the body of this chapter by reference to this footnote and his contributions are stated explicitly in the experimental sections of this chapter.

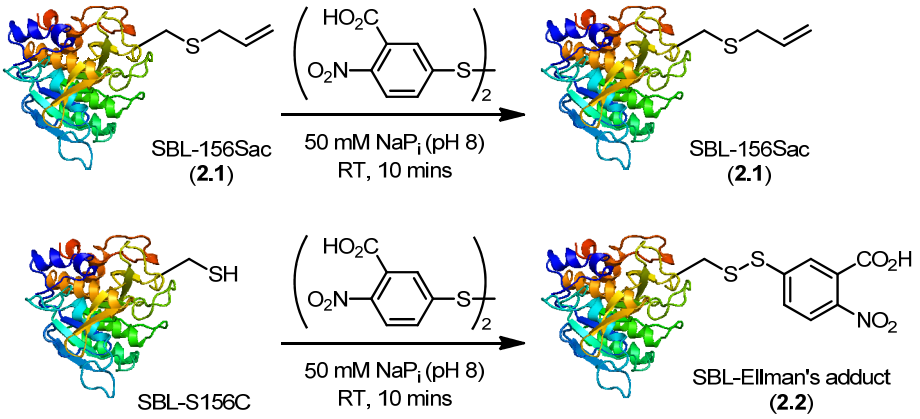
sometimes be avoided by judicious selection of electrophiles, together with discreet control of pH, stoichiometry, reaction time and temperature.[6] Successful cysteine alkylation demonstrating some degree of selectivity has been shown on proteins with electrophiles such as α-halocarbonyls, maleimides, and other alkyl halides.[7]

To investigate this transformation, a single-cysteine mutant of subtilisin from *Bacillus lentus* (SBL-S156C) was chosen as a model protein. This single-cysteine mutant simplifies initial optimization of reaction conditions. In addition, the protease activity of the enzyme can be assayed before and after modification to ensure the protein is not denatured during the reaction. At the outset, allylation was attempted simply by treating a sample of SBL-S156C in pH 8.0 phosphate buffer with 1000 equivalence of allyl chloride solution in DMF. The total amount of DMF needed to achieve complete homogeneity was found to be 5% of the total volume and is sufficiently low to be compatible with most proteins. After incubating the reaction at 37 °C for 30 minutes, full conversion to allylated protein **2.1** (SBL-156Sac) was observed (Scheme 2-2).



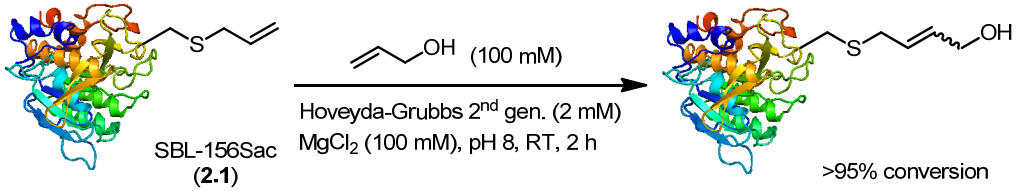
Scheme 2-2: Installing *S*-allyl cysteine on protein via direct allylation with allyl chloride.

Complete consumption of free thiol, and thus reaction at cysteine, was further verified by modified Ellman's assay. In the presence of free sulfhydryl groups, the treatment of Ellman's reagent (5,5′-Dithiobis(2-nitrobenzoic acid))[8] led to formation of the corresponding disulfide **2.2** and releasing 5-mercapto-2-nitrobenzoic acid which is indicated by the yellow appearance of the resulting reaction mixture (Scheme 2-3). No reaction is observed otherwise.



Scheme 2-3: Ellman's assay for free sulfhydryl groups.

The selective allylation of cysteine was observed even when a large excess of allyl chloride was used (>1000 equivalents) – a testament to the unique nucleophilicity of cysteine. Furthermore, a protocol with pre-reduction of cysteine with dithiothreitol (DTT) was also demonstrated[†], a useful option for cysteines prone to oxidation. The allylated protein was also active in cross-metathesis with allyl alcohol; full conversion was observed after 2 hours at RT (Scheme 2-4). Finally, peptidase activity of the allylated protein[†] was retained, as evidenced by liberation of *p*-nitroaniline (*p*-NA) upon treatment with the peptide suc-AAPF-*p*NA (see Experimental procedures).



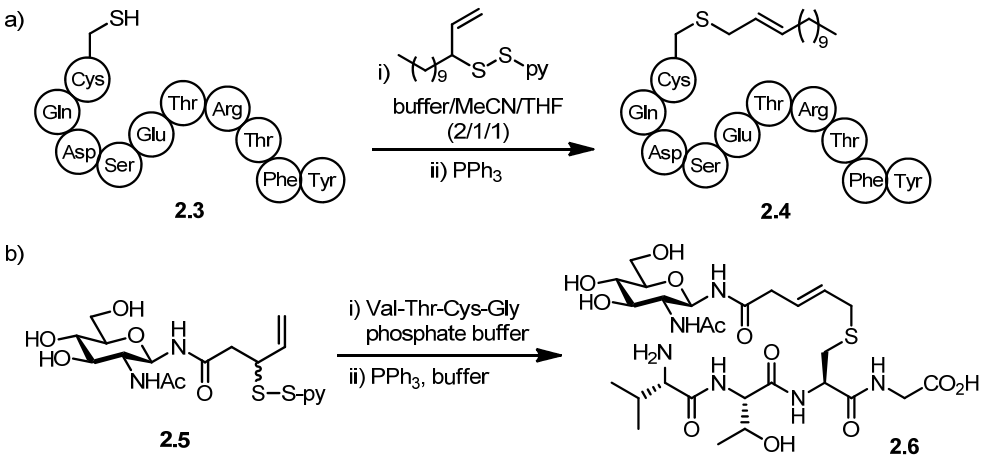
Scheme 2-4: Cross-metathesis with allyl alcohol at *S*-allyl cysteine.

2.2.2 Rearrangement and Reduction of Allylic Disulfides and Allyl Selenenylsulfide

Direct allylation of the model protein with allyl chloride was found to be selective and allows a direct route for the formation of *S*-allyl cysteine. However, this approach may have some drawbacks for proteins containing strongly nucleophilic residues that may be allylated non-selectively, as observed with other alkylating reagents such as maleimides

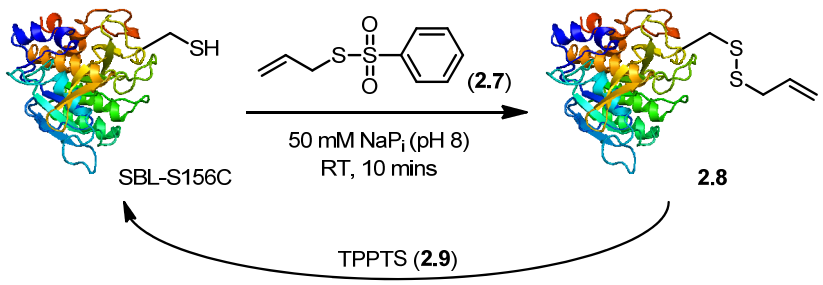
and α-halocarbonyls.[9] This selectivity issue may also be more pronounced when the target cysteine is buried. This allylation is therefore best accomplished with a cysteine-specific method. To fulfill such requirement, Crich's work on the formation and dechalcogenative rearrangement of allyl disulfides and allyl selenenylsulfide are considered next.

The rearrangement and reduction of allylic disulfides and allylic diselenides, was initially reported by Baldwin[10] and Sharpless,[11] respectively. The Crich laboratory later adapted this chemistry as a cysteine-specific ligation strategy.[12] Crich and co-workers have demonstrated the utility of this methodology in specific allylation of cysteine to functionalize Fibronectin fragment (decapeptide **2.3**),[12c] and to synthesize neoglycoconjugates such as **2.6** (Scheme 2-5).[12d] These reactions perform well in aqueous media because the pericyclic transition state of the rate-determining rearrangement step is stabilized by polar solvents.[13] The ease and specificity of preparation of disulfides and selenenylsulfides at cysteine and the abundant precedence for the subsequent reductive rearrangement prompted the investigation of these transformations as potential routes to *S*-allyl cysteine on proteins.



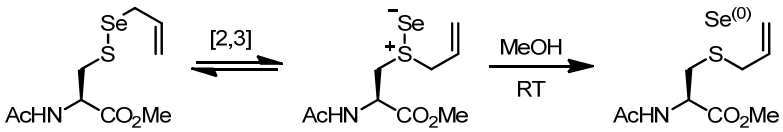
Scheme 2-5: Cysteine-specific functionalization of peptides via dechalcogenative rearrangement of allylic disulfides.

The rearrangement of allyl disulfides was investigated first. When SBL-S156C was treated with allyl phenylthiosulfonate (**2.7**), the corresponding allyl disulfide **2.8** was formed with full conversion. Compound **2.7** belongs to a class of reagents that have previously been shown to efficiently form mixed disulfides on a protein surface.[14] Disulfide **2.8** was stable at room temperature and no spontaneous loss of sulfur could be observed over several hours by LC-MS. In order to promote the desired dechalcogenative rearrangement, the trisodium salt of triphenylphosphine-3,3′,3″-trisulfonic acid (TPPTS, **2.9**) was added to protein **2.8**†. TPPTS was chosen as a water-soluble derivative of triphenylphosphine, the phosphine previously used by Crich to promote desulfurization in peptidic systems in less polar solvents. Unfortunately, only the reduced product was observed (Scheme 2-6).



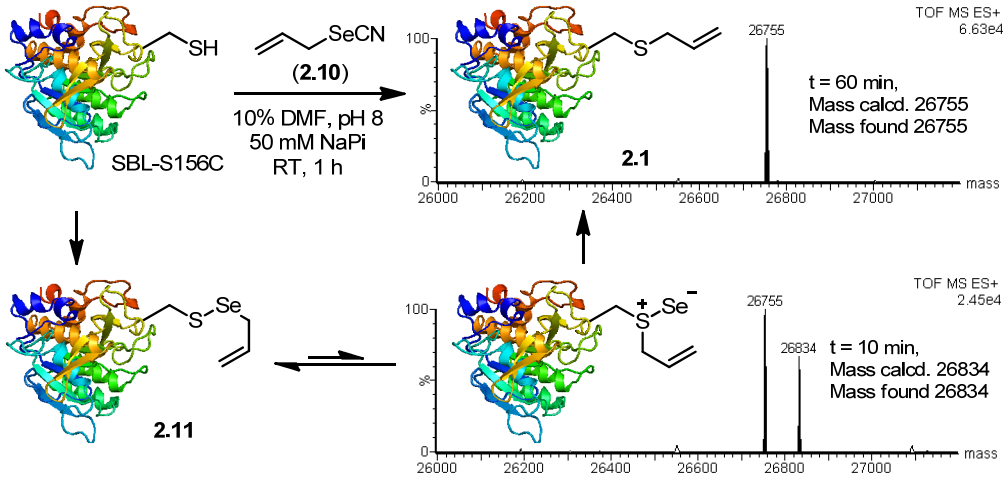
Scheme 2-6: Attempted desulfurization of protein allylic disulfide with TPPTS.

The unsubstituted allylic disulfide was presumably too exposed and prone to direct attack by the nucleophilic TPPTS that was present in large excess. While altering reaction conditions and screening other phosphines may rescue this disulfide route to *S*-allyl cysteine, an alternative route via the dechalcogenative rearrangement of allyl selenenylsulfide is investigated instead. Crich and coworkers have shown that in such system the loss of selenium is far more facile than the desulfurization of allyl disulfides. In many cases, the loss of selenium was found to be spontaneous in the absence of a phosphine,[12b,c] an important feature for proteins with natural disulfides (Scheme 2-7).



Scheme 2-7: Rearrangement in the absence of phosphines.

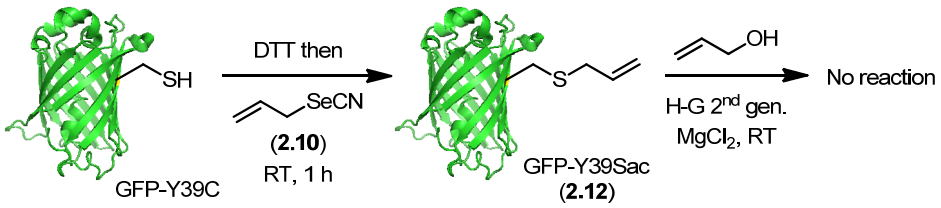
Allyl selenocyanate and *Se*-allyl selenosulfate have both been reported as suitable reagents for the conversion of cysteine to its *Se*-allyl selenenylsulfide.[12b,c] To test the dechalcogenative rearrangement of allyl selenenylsulfide on a protein, allyl selenocyanate (**2.10**) was used to synthesize allyl selenenylsulfide on protein (**2.11**, Scheme 2-8) as previous reports indicated higher overall yields of allylated peptides as compared to when allylic selenosulfates were used. On treating SBL-S156C with **2.10**, the cysteine was reacted completely within 10 minutes at room temperature. LC-MS analysis revealed rapid formation of *Se*-allyl selenenylsulfide/*S*-allyl selenosulfide intermediates that spontaneously lose selenium over the course of 1 hour to smoothly obtain the desired *S*-allyl cysteine with full conversion (Scheme 2-8)[†]. Again, the specificity of the allylation at cysteine and the peptidase activity of modified protein were confirmed by Ellman's assay and treatment of suc-AAPF-*p*NA peptide, respectively (see Experimental procedures). The fact that modified protein remained an active protease indicates the mild nature of such transformation.



Scheme 2-8: Cysteine-specific allylation by reductive rearrangement of *Se*-allyl selenenylsulfide.

The allylated SBL via allyl selenocyanate route also underwent model cross-metathesis with allyl alcohol in the presence of Hoveyda-Grubbs 2nd generation catalyst. Full conversion was observed after two hours at room temperature. As expected, no difference in cross-metathesis reactivity was observed between the sample obtained from allyl chloride and the sample obtained from allyl selenocyanate (see Experimental procedures).

Finally, to show the generality of this cysteine-specific allylation method, a single-cysteine mutant of green fluorescent protein (GFP-Y39C) was allylated with allyl selenocyanate[†]. Efficient allylation was achieved after 1 hour at room temperature to provide the desired protein product (Scheme 2-9). GFP remained fluorescent throughout the chemical manipulations indicating that the tertiary structure of the protein was intact. However, attempts at the subsequent cross-metathesis using previously optimized conditions were unsuccessful. LC-MS analysis of the reaction revealed only the unreacted starting material. This observation was attributed to a more sterically hindered reaction site in GFP, after comparing its crystal structure with that of SBL.



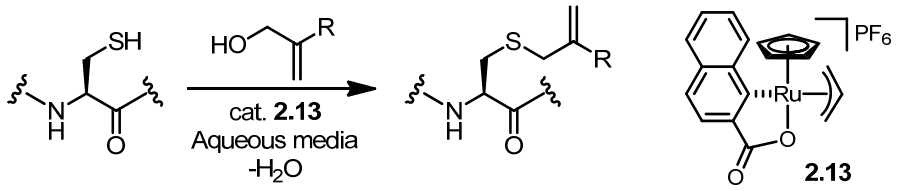
Scheme 2-9: Allylation of GFP-Y39C with allyl selenocyanate and attempted cross-metathesis.

2.3 Conclusions and Outlook

In this chapter, multiple chemical methods to install *S*-allyl cysteine on proteins were investigated. While some success was observed at direct allylation with allyl chloride,

the cysteine-specific allylation via dechalcogenative rearrangement (using allyl selenocyanate) was a more selective and controllable route to *S*-allyl cysteine. This is the first demonstration of *Se*-allyl selenenylsulfide reductive rearrangement on proteins. The main purpose of this chemistry has been to provide a facile access to metathesis reactive handle facilitating the investigation of subsequent metathesis reactions. Nonetheless, as this reductive rearrangement of *Se*-allyl selenenylsulfide has been used by Crich in a variety of allylations on peptide substrates,[12] it would also bode well as a ligation method for the synthesis of other lipidated proteins.

Since the publication of the results in this chapter,[15] other thio-specific allylation methods have emerged. In the reports by Kitamura and coworkers, thiol-containing peptides and cysteine could be successfully allylated with various allylic alcohols in aqueous medium in the presence of catalytic amount of a π-allyl Ru(IV) complex (**2.13**).[16] Allylic alcohols containing long alkyl chain can also be installed via this method to obtain peptides with farnesyl mimics as well as other artificial lipopeptides, suggesting potential application in synthetic biology.



Scheme 2-10: Catalytic dehydrative *S*-allylation of thiols.

The unsuccessful metathesis observed on GFP-Y39Sac (**2.12**) raised several so far unanswered questions in the scope of olefin metathesis for protein modifications: In what way does sterics affect the outcome of metathesis? Can metathesis be executed at the active site or on a hindered helix? What other factors also affect the efficiency of metathesis on proteins? These questions are investigated in the next chapter.

2.4 References

(1) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642-9643.

(2) (a) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053. (b) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem. Sci.* **2011**, *2*, 1666-1676. (c) Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 1835-1839.

(3) (a) Zhu, Y.; van der Donk, W. A. *Org. Lett.* **2001**, *3*, 1189-1192. (b) Galonić, D. P.; van der Donk, W. A.; Gin, D. Y. *Chem. Eur. J.* **2003**, *9*, 5997-6006.

(4) (a) Kuhn, K.; Owen, D. J.; Bader, B.; Wittinghofer, A.; Kuhlmann, J.; Waldmann, H. *J. Am. Chem. Soc.* **2001**, *123*, 1023-1035. (b) Ludolph, B.; Eisele, F.; Waldmann, H. *J. Am. Chem. Soc.* **2002**, *124*, 5954-5955.

(5) Alexandrov, K.; Heinemann, I.; Durek, T.; Sidorovitch, V.; Goody, R. S.; Waldmann, H. *J. Am. Chem. Soc.* **2002**, *124*, 5648-5649.

(6) Crankshaw, M. W.; Grant, G. A. *Modification at Cysteine*; John Wiley & Sons, Inc. Unit 15.1, 1996.

(7) (a) Hermanson, G. T. *Bioconjugation Techniques, 2nd edition*; Academic Press, Inc., 2008. (b) Chalker, J. M.; Bernardes, G. J. L.; Lin, Y. A.; Davis, B. G. *Chem. Asian J.* **2009**, *4*, 630-640.

(8) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70-77.

(9) Nielsen, M. L.; Vermeulen, M.; Bonaldi, T.; Cox, J.; Moroder, L.; Mann, M. *Nat. Methods* **2008**, *5*, 459-460.

(10) Höfle, *G.*; Baldwin, J. E. *J. Am. Chem. Soc.* **1971**, *93*, 6307-6308.

(11) Sharpless, K. B.; Lauer, R. F. *J. Org. Chem.* **1972**, *37*, 3973-3974.

(12) (a) Crich, D.; Brebion, F.; Krishnamurthy, V. *Org. Lett.* **2006**, *8*, 3593-3596. (b) Crich, D.; Krishnamurthy, V.; Hutton, T. K. *J. Am. Chem. Soc.* **2006**, *128*, 2544-2545. (c) Crich, D.; Krishnamurthy, V.; Brebion, F.; Karatholuvhu, M.; Subramanian, V.; Hutton, T. K. *J. Am. Chem. Soc.* **2007**, *129*, 10282-10294. (d) Crich, D.; Yang, F. *J. Org. Chem.* **2008**, *73*, 7017-7027.

(13) Li, Z.; Wang, C.; Fu, Y.; Guo, Q.-X.; Liu, L. *The Journal of Organic Chemistry* **2008**, *73*, 6127-6136.

(14) Gamblin, D. P.; Garnier, P.; Ward, S. J.; Oldham, N. J.; Fairbanks, A. J.; Davis, B. G. *Org. Biomol. Chem.* **2003**, *1*, 3642-3644.

(15) Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714-3716.

(16) (a) Tanaka, S.; Pradhan, P. K.; Maegawa, Y.; Kitamura, M. *Chem. Commun.* **2010**, *46*, 3996-3998. (b) Jaisankar, P.; Tanaka, S.; Kitamura, M. *J. Org. Chem.* **2011**, *76*, 1894-1897.

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 and are uncorrected. Elemental microanalysis was obtained through the London Metropolitan University Elemental Analysis Service.

Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on 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 AV400 (100 MHz) spectrometer or on a Bruker AVII500 (125 MHz) spectrometer, as indicated. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard (^{1}H NMR: $CDCl_3$ = 7.26; DMSO-d_6 = 2.50 and ^{13}C NMR: $CDCl_3$ = 77.0; DMSO-d_6 = 39.5). Coupling constants (*J*) are reported in Hertz (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 (υ_{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.

Thin layer chromatography (TLC) was carried out using Merck aluminium backed sheets coated with $60F_{254}$ silica gel. Visualization of the silica plates was achieved using a UV lamp (λ_{max} = 254 nm), and/or ammonium molybdate (5% in 2 M 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).

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} deg cm^2 g^{-1}. Concentrations (c) are given in g/100 ml.

Anhydrous solvents were purchased from Fluka or Acros with the exception of DCM and THF, which were dried by passing through an activated alumina column. All other solvents were used as supplied (Analytical or HPLC grade), without prior purification. Reagents were purchased from Alfa Aesar or Sigma-Aldrich and used as supplied unless stated otherwise. Distilled water was used for chemical reactions and Milli-Q® purified water for protein manipulations. 'Petrol' refers to the fraction of light petroleum ether boiling in the range 40-60 ºC. 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) was 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 acetone, 500 mL methanol, 500 mL 5 M HCl, and then 1 L 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.

In addition to those specified above, the following abbreviations, designations, and formulas are used throughout the experimental procedures:

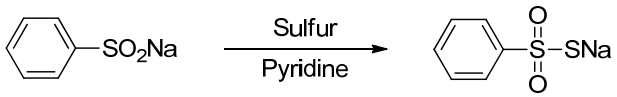
MeOH = methanol	H_2O = water
EtOAc = ethyl acetate	CH_2Cl_2 = dichloromethane
^{i}PrOH = isopropanol	PBu_3 = tributylphosphine
Boc = *tert*-butoxycarbonyl	K_2CO_3 = potassium carbonate
$NaHCO_3$ = sodium bicarbonate	NaOH = sodium hydroxide
sat. = saturated	aq. = aqueous
Cs_2CO_3 = cesium carbonate	NH_4OH = ammonium hydroxide
DMF = dimethylformamide	NH_4Cl = ammonium chloride
Et_2O = diethyl ether	Et_3N = triethylamine
RT = room temperature	NaP_i = sodium phosphate

Protein Mass Spectrometry

Liquid chromatography-mass spectrometry (LC-MS) was performed on a Micromass LCT (ESI-TOF-MS) coupled to a Waters Alliance 2790 HPLC using a Phenomenex Jupiter C4 column (250 × 4.6 mm × 5μm). Water:acetonitrile, 95:5 (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as the mobile phase at a flow rate of 1.0 mL min^{-1}. The gradient was programmed as follows: 95% A (5 min isocratic) to 100% B after 15 min then isocratic for 5 min. The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V for subtilisin *Bacillus lentus* (SBL) and Green Fluorescent Protein (GFP). Nitrogen was used as the nebulizer and desolvation gas at a total flow of 600 L hr^{-1}. Spectra were 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 ion series of protein was generated by combining the chromatogram between 13 and 16 minutes. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v. 4.0 from Waters) according to manufacturer's instructions.

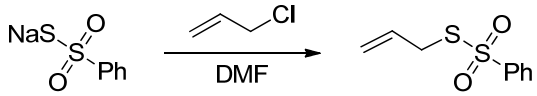
Preparation of Compound 2.7 and 2.10

Sodium phenylthiosulfonate



Sodium benzenesulfinate (10.00 g, 61.0 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 Et_2O. Recrystallization from EtOH afforded sodium phenylthiosulfonate (8.40 g, 70%) as a white crystalline solid. Spectroscopic data was consistent with that previous reported. m.p. = 294-295 °C (Lit.[2.1] = 287 °C). IR (υ_{max}, KBr): 1936, 1900, 1818, 1771, 1684, 1607, 1583, 1475, 1443, 1390, 1333, 1311. ^{1}H NMR (400 MHz, DMSO-d_6): δ_H = 7.19-7.50 (3H, m, *m,p*-ArH), 7.75 (2H, dd, J = 7.7, 1.5, *o*-ArH). ^{13}C NMR (100 MHz, DMSO-d_6): δ_C = 124.0, 127.8, 129.0 (3° Ar), 154.9 (4° Ar). LRMS m/z (ESI^-): Found 173.0 $[M-Na]^-$; $C_6H_5O_2S_2$ requires 173.0.

***S*-Allylphenylthiosulfonate (2.7)**



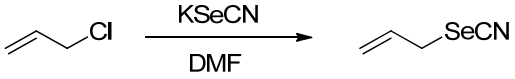
Sodium phenylthiosulfonate (1.5 g, 7.64 mmol) was added to a 25 mL round bottom flask containing allyl chloride (1 mL, 12.3 mmol) in DMF (5 mL). Reaction was stirred at room temperature for 3.5 h. Reaction was diluted with water (100 mL) and transferred to a separation funnel. Aqueous layer was extracted with Et_2O (2 × 100 mL). Combined Et_2O layers were washed with brine (100 mL) and dried with $MgSO_4$, which was filtered. Volatiles were removed in vacuo to give the titled compound as a clear oil (592 mg, 36%). IR (υ_{max}, film): 1447, 1325, 1144, 1078, 988, 929, 755, 716, 685, 598, 537. ^{1}H NMR (400

[2.1] R. Sato, T. Goto, Y. Takikawa, S. Takizawa, Synthesis **1980**, 615.

MHz, $CDCl_3$): δ_H = 3.69 (2H, d, *J* = 7.0, SC<u>H</u>$_2$), 5.10 (1H, dd, *J* = 9.9, 1.0, HC=C<u>H</u>H *cis*), 5.21 (1H, dd, *J* = 17.0, 1.0, HC=CH<u>H</u> *trans*), 5.70 (1H, m, *J* = 17.0, 9.9, 7.0, 7.0, <u>H</u>C=CH_2), 7.53-7.60 (2H, m, *m*-ArH), 7.60-7.69 (1H, m, *p*-ArH), 7.93 (2H, d, *J* = 7.3, *o*-ArH). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 38.9 (S<u>C</u>H_2), 120.1 (HC=<u>C</u>H_2), 127.0, 129.2, 130.5, 133.7 (includes 3º Ar and H<u>C</u>=CH_2), 144.9 (4º Ar). Anal. Calcd. for $C_9H_{10}O_2S_2$: C, 50.44; H, 4.70; S, 29.92. Found C, 50.36; H, 4.76; S, 30.00.

Allyl selenocyanate (2.10)[2.2]

This experiment was carried out with Justin Chalker.



The synthesis of allyl selenocyanate was adapted from the literature.[2.3] KSeCN (3.27 g, 22.72 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, 45.43 mmol) was added slowly to the stirred solution. The reaction was stirred for 20 minutes at room temperature and then diluted with Et_2O (200 mL) and washed sequentially 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 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 (υ_{max}, film): 2151, 1633, 1432, 1403, 1197, 987, 928, 851, 686. ^{1}H NMR (200 MHz, $CDCl_3$): δ_H = 3.65 (2H, dt, *J* = 7.2, 1.0, CH_2SeCN), 5.23 (1H, dd, *J* = 1.0, 9.9, CHH=CH), 5.32 (1H, dq, *J* = 1.0, 16.7, CHH=CH), 6.00 (1H, m, H_2C=CH). ^{13}C NMR (50 MHz, CDCl3): δ_C = 31.3 (CH_2SeCN), 101.4 (C≡N), 120.6 (CH=CH_2), 131.6 (CH=CH_2).

2.2 Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714-3716.
2.3 E. H. Riague, J.-C. Guillemin, *Organometallics* **2002**, *21*, 68-73.

Protein modification on SBL-S156C

Sequence of subtilisin Bacillus lentus (SBL) mutant S156C

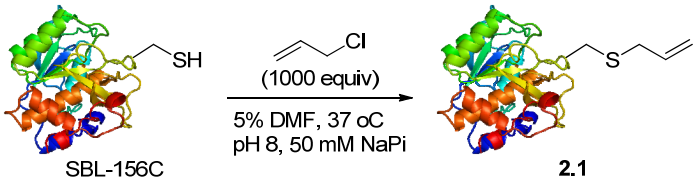
PDB code for wild type = 1GCI

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQD
GNGHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNG
MHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGNCGAGSISYPARYANAMAVGAT
DQNNNRASFSQYGAGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKN
PSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

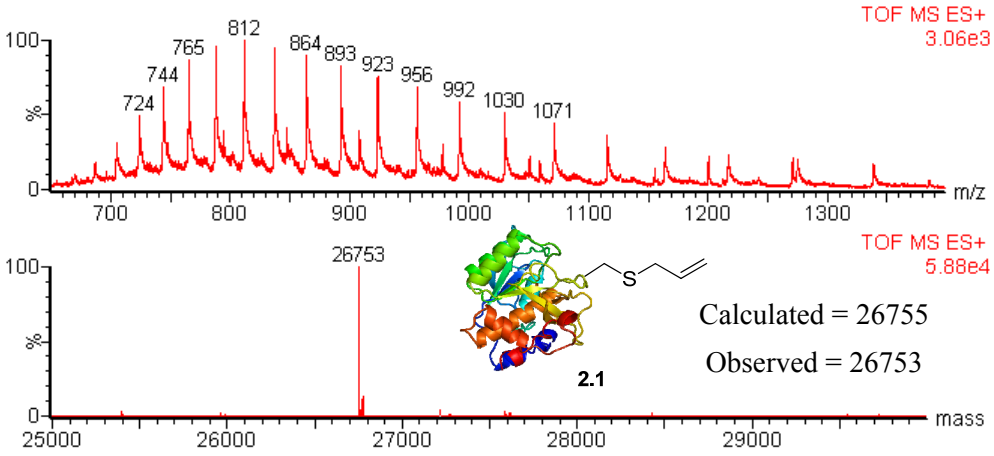
Calculated average isotopic mass = 26714.5

Allylation of SBL-S156C with Allyl Chloride

Method A: Direct Allylation

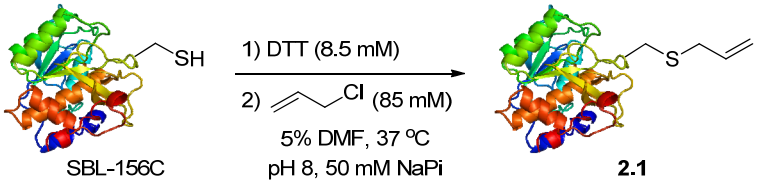


SBL-S156C (2.5 mL, 1 mg/mL, pH 8.0 sodium phosphate, 94 nmol) was added to a 15 mL Falcon tube and stored on ice until needed. Allyl chloride was prepared as a 0.65 M solution in DMF. A 147 µL aliquot of the allyl chloride solution (96 µmol) was added to the protein solution and the reaction was vortexed immediately upon addition. The reaction was incubated at 37 °C for 30 minutes and then analyzed directly by LC-MS, confirming full conversion to the allylated product SBL-156-Sac (**2.1**) (calculated mass = 26755, observed mass = 26753). 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 200 µL aliquots, flash frozen, and stored at −20 °C. ESI-MS are shown below.

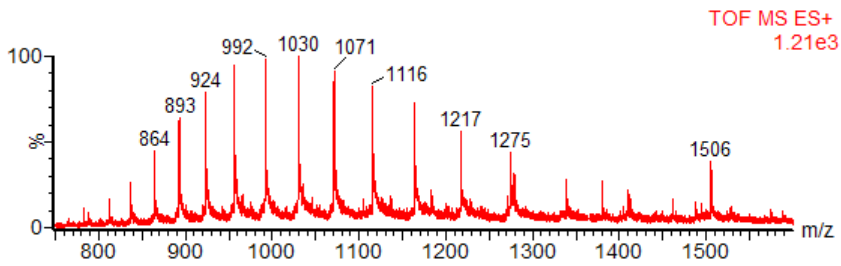


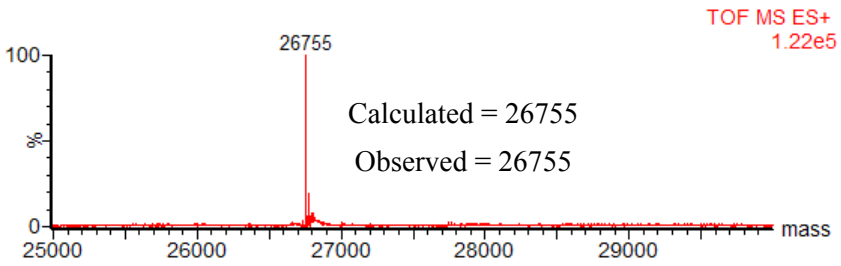
Method B: Pre-reduction with DTT

This experiment was carried out with Justin Chalker.

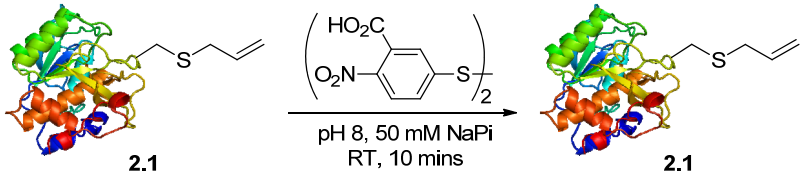


SBL-S156C (2.5 mL, 1.0 mg/mL, pH 8.0 sodium phosphate, 94 nmol) was added to a 15 mL Falcon tube and stored on ice. Dithiothreitol (DTT) (3.6 mg, 23 µmol) was added as a solid to reduce any contaminant disulfide. The solution was vortexed and then shaken for 10 minutes at room temperature. Allyl chloride (19 µL, 230 µmol) was then added as a solution in DMF (200 µL). The mixture was vortexed and then shaken at 37 °C for 30 min. LC-MS analysis revealed full conversion to the allylated product **2.1** (26755 = calculated mass, found 26756). The reaction mixture was passed through a PD10 column previously equilibrated with pH 8.0 sodium phosphate (50 mM). The product was split into 200 µL aliquots and flash-frozen. ESI-MS are shown below.

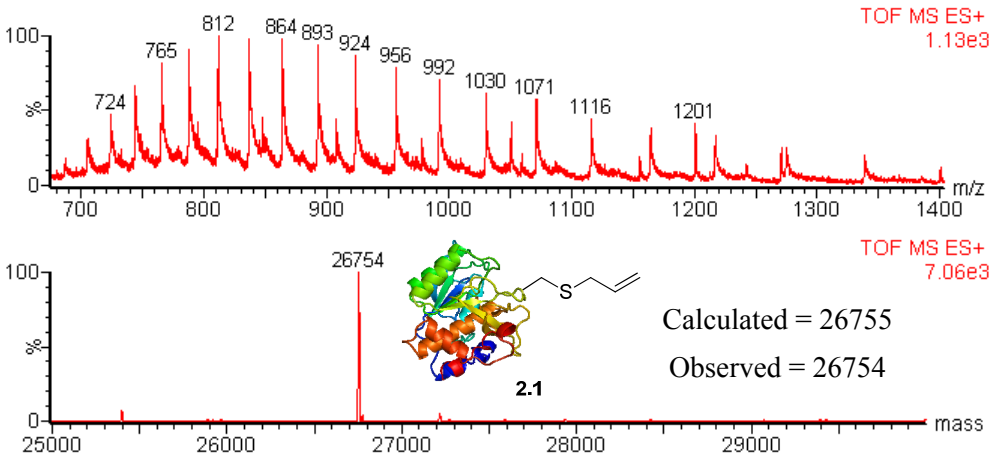


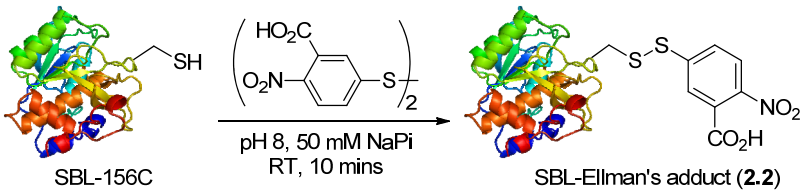


Ellman's Test to Confirm Reaction at Cysteine

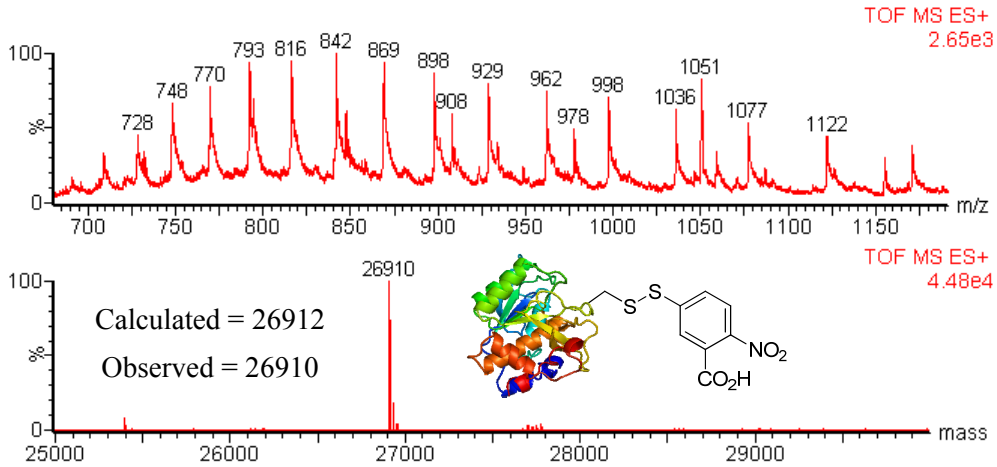


A buffered solution of Ellman's reagent was prepared by dissolving 3.0 mg of Ellman's reagent in 300 µL of 50 mM sodium phosphate buffer (pH 8). A 20 µL of the Ellman's solution (0.5 µmol) was added to SBL-156Sac (**3**) (100 µL, 0.7 mg/mL, 3 nmol, prepared from Method A above) in an eppendorf tube. 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 = 26754). 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).

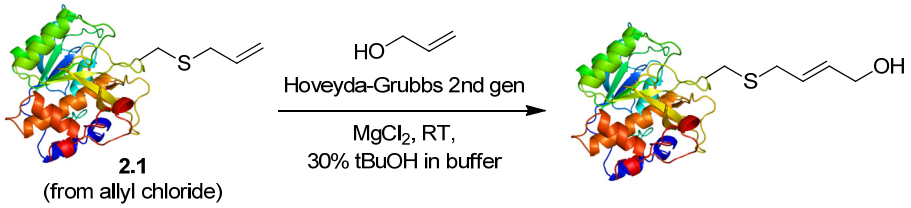




An SBL-S156C solution (100 µL, 0.71 mg/mL) was prepared by diluting 71 µL of SBL-S156C (1 mg/mL) to 100 µL with 50 mM sodium phosphate buffer (pH 8.0). A 20 µL aliquot of the Ellman's solution prepared above was added to protein solution. The solution turned bright yellow immediately upon addition of 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 **2.2** was observed (calculated mass = 26912, observed mass = 26910).



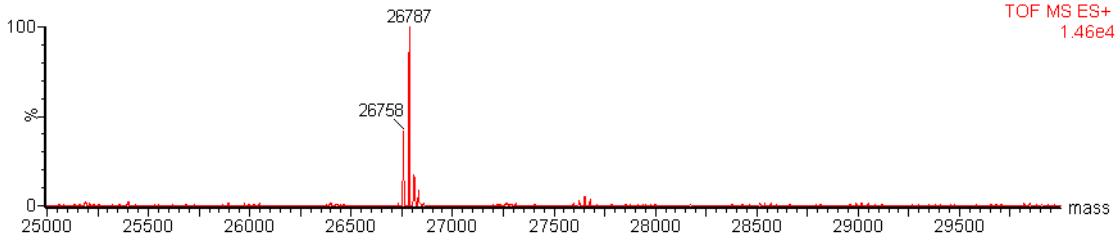
CM between Allyl Alcohol and SBL-Sac (2.1) from Allyl Chloride Allylation



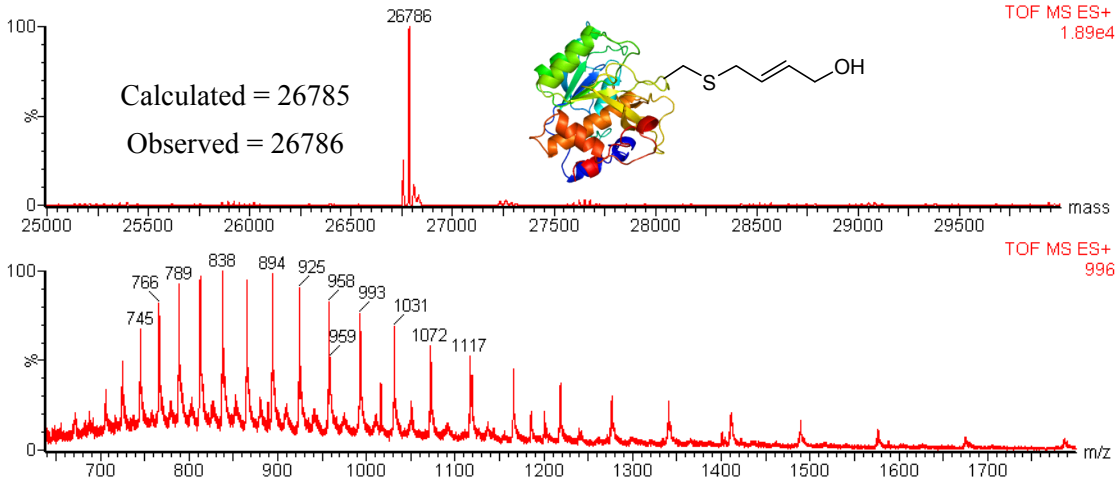
All manipulations were carried out at room temperature. The protein was prepared as described above by allylation with allyl chloride. A solution of Hoveyda-Grubbs 2nd generation catalyst in ^{t}BuOH was prepared by vortexing and gently warming 2.2 mg of Ru catalyst in 355 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (**2.1**) (250 µL, 0.57 mg/mL, 5.3 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.8 mg, 53 µmol) followed

by an aliquot of the catalyst/ ^{t}BuOH solution (107 µL, 1.1 µmol). The reaction was vortexed after each addition to homogenize. Allyl alcohol (3.6 µL, 53 µmol) 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).

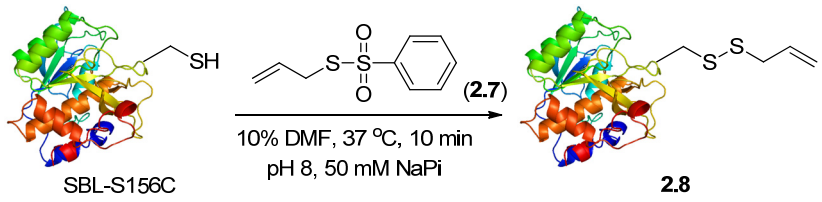
1 h:



2 h:

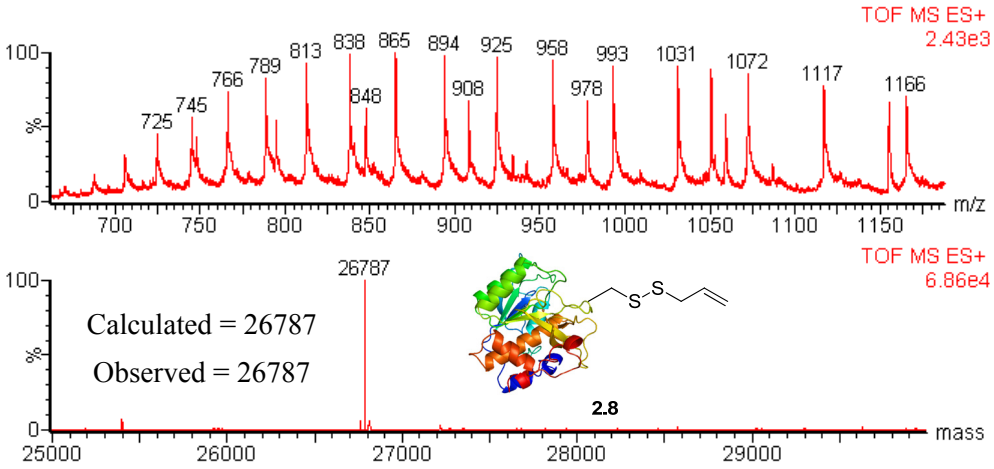


Synthesis of Allyl Disulfide on SBL-S156C



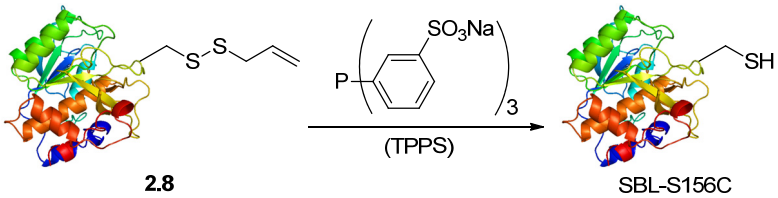
All manipulations were carried out at room temperature. A solution of allyl phenylthiosulfonate (**2.7**) in DMF was prepared by dissolving 3.3 mg of **2.7** in 406 µL DMF. 100 µL of the DMF solution containing **2.7** (0.8 mg, 3.7 µmol) was added to a 1.5 mL eppendorf tube with SBL-S156C (1 mL, 1 mg/mL, 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 **2.8** (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 ºC. The final concentration of protein was 0.3 mg/mL.

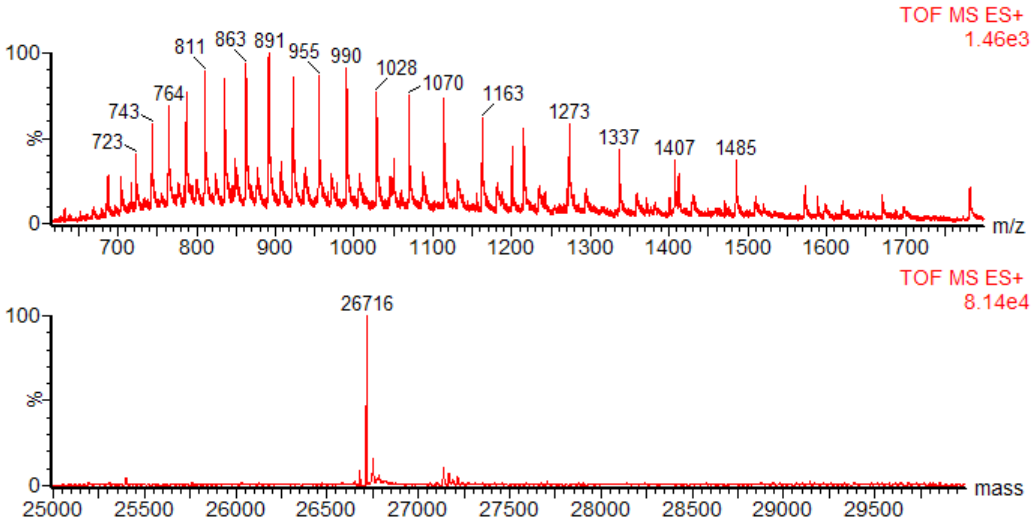


Attempted Desulfurization of Allyl Disulfide 2.8 with TPPTS

This experiment was carried out with Justin Chalker.

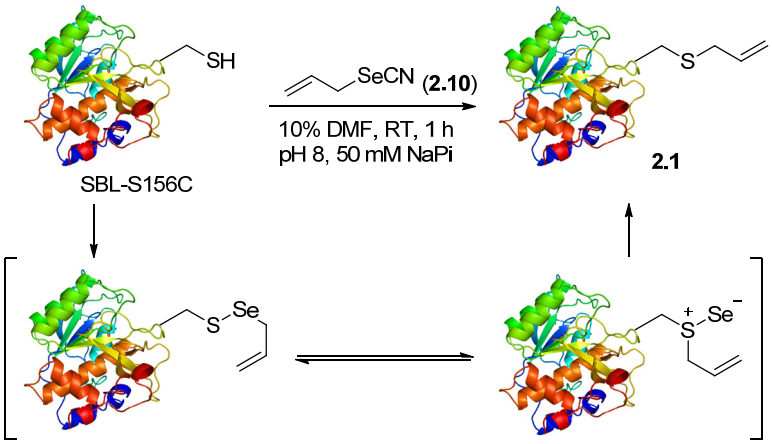


SBL-156-allyl disulfide **2.8** (250 µL, 0.29 mg/mL, pH 8.0 sodium phosphate, 2.7 nmol) was added to a 1.0 mL plastic tube and stored on ice. A stock solution of tris(3-sulfophenyl)phosphine trisodium salt (TPPTS, Aldrich) was prepared by dissolving 0.9 mg in 112.5 µL of H_2O. A 10 µL aliquot of the 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 the free cysteine (product of reduction) was observed. (26715 = calculated mass; 26715 found).



Allylation of SBL-S156C with Allyl Selenocyanate (2.10) and Its Ellman's Test to Confirm Reaction at Cysteine

This experiment was carried out with Justin Chalker.

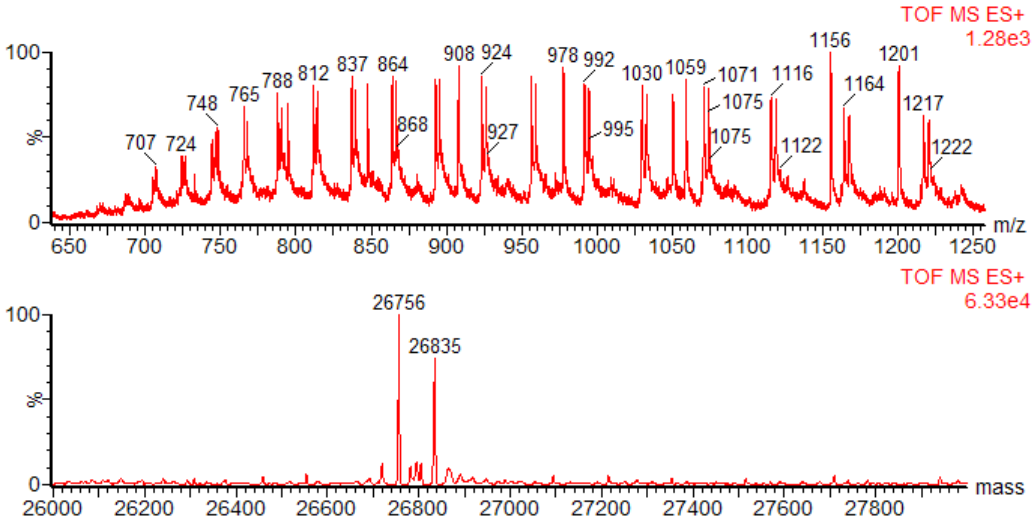


All manipulations were carried out in a well-ventilated hood. SBL-S156C (200 µL, 1.0 mg/mL, pH 8.0 sodium phosphate, 7.5 nmol) was added to a 1.0 mL plastic tube and stored on ice. A stock solution of allyl selenocyanate (**2.10**) (11.9 mg) was prepared in DMF (541 µL). A 5 µL aliquot of the solution containing **2.10** (0.75 µmol) was added to the protein sample. The reaction was shaken for 10 minutes at room temperature and a 30 µL aliquot was analyzed by LC-MS. A mixture of selenenylsulfide and *S*-allyl cysteine was observed (~5:4). (26834 calculated for selenenylsulfide; 26755 calculated for *S*-allyl sulfide). After 1 hour of total reaction time, the reaction was analyzed by

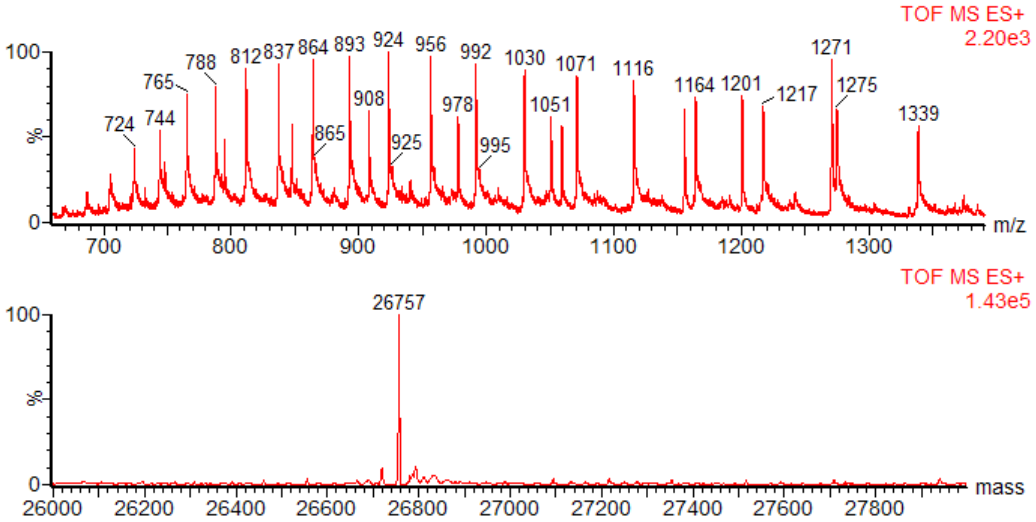
LC-MS. Full conversion to the allyl sulfide **2.1** was observed. (26755 = calculated mass; 26755 found). Spectra from LC-MS analysis are shown below.

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 μmol in 25 μL 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 NaP_i (50 mM). This sample was used for subsequent metathesis reactions.

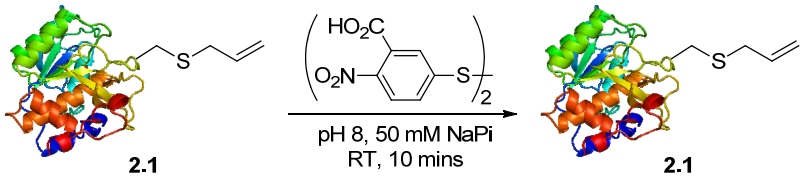
10 minutes:



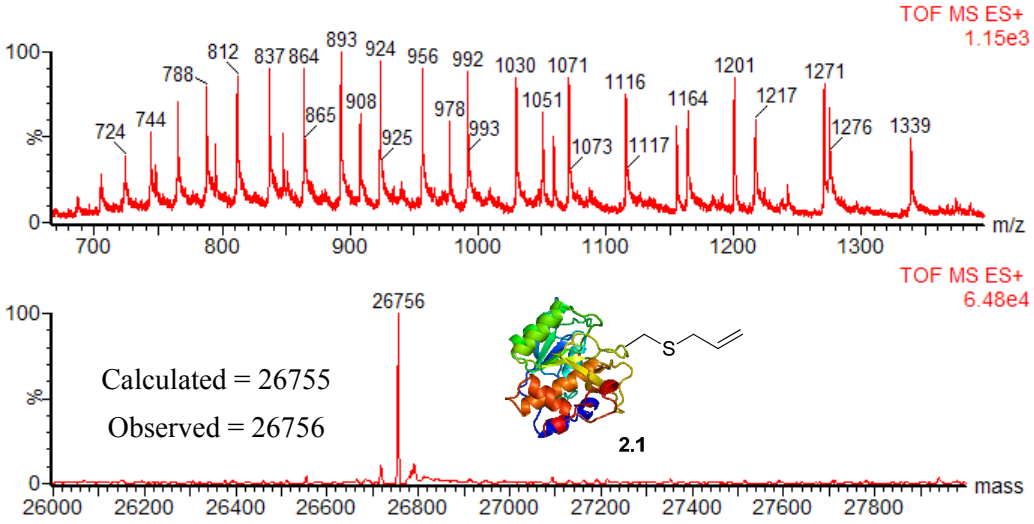
60 minutes:

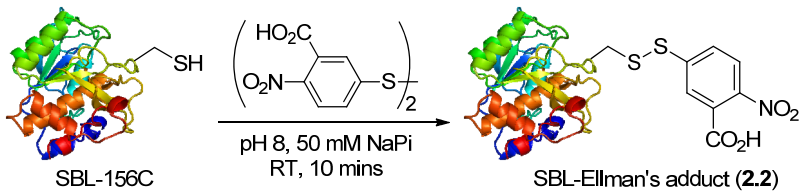


Ellman's Test to Confirm Reaction at Cysteine



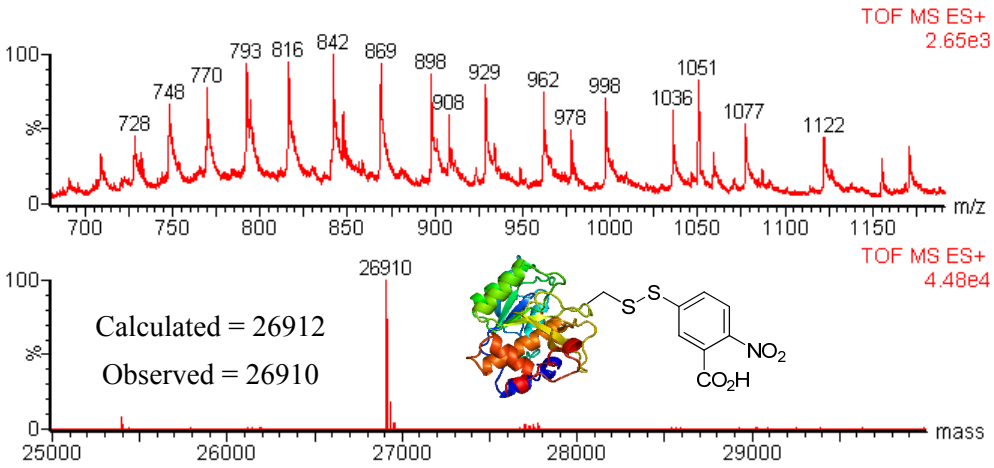
A buffered solution of Ellman's reagent was prepared by dissolving 3.0 mg of Ellman's reagent in 300 µL of 50 mM sodium phosphate buffer (pH 8). A 20 µL of the Ellman's solution (0.5 µmol) was added to SBL-156Sac (**2.1**) (100 µL, 0.7 mg/mL, 3 nmol, prepared from Method A above) in an eppendorf tube. 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: 26756). 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.



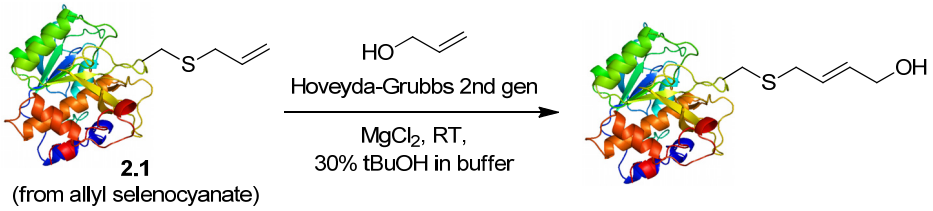


An SBL-S156C solution (100 µL, 0.71 mg/mL) was prepared by diluting 71 µL of SBL-S156C (1 mg/mL) to 100 µL with 50 mM sodium phosphate buffer (pH 8.0). A 20 µL aliquot of the Ellman's solution prepared above was added to protein solution. The

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).



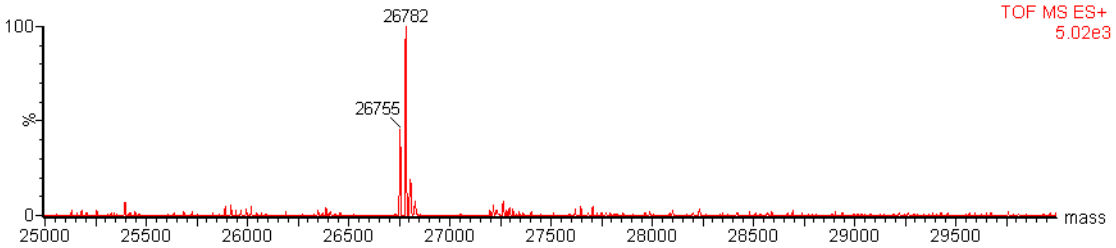
CM between Allyl Alcohol and SBL-156Sac from Allyl Selenocyanate Allylation



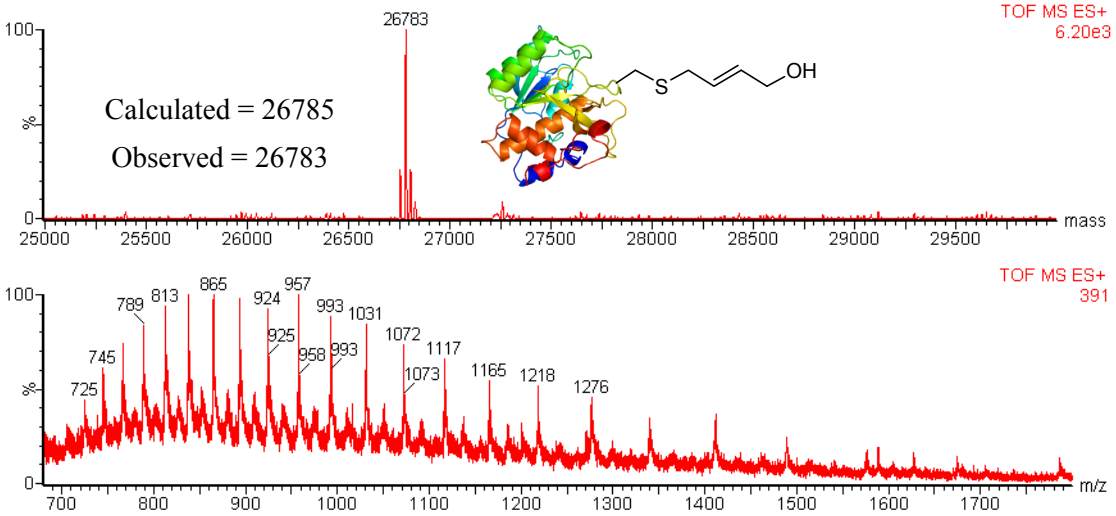
All manipulations were carried out at room temperature. The protein was prepared as described above by allylation with allyl selenocyanate. A solution of Hoveyda-Grubbs 2nd generation catalyst in ^{t}BuOH was prepared by vortexing and gently warming 1.4 mg of Ru catalyst in 444 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (**2.1**) (250 µL, 0.29 mg/mL, 2.7 nmol) was added $MgCl_2 \cdot 6H_2O$ (5.5 mg, 27 µmol) followed by an aliquot of the catalyst/^{t}BuOH solution (107 µL, 0.54 µmol). The reaction was vortexed after each addition to homogenize. Allyl alcohol (1.8 µL, 27 µmol) 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).

1 h:

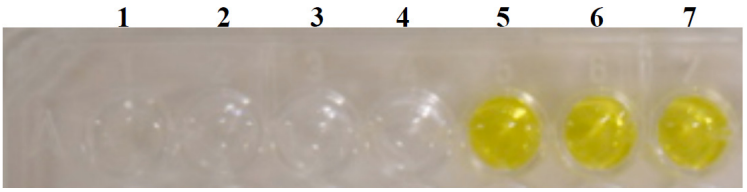


2 h:



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 µL aliquots of each sample were added to a 96-well plate. A 25 µL 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*-nitroanaline (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

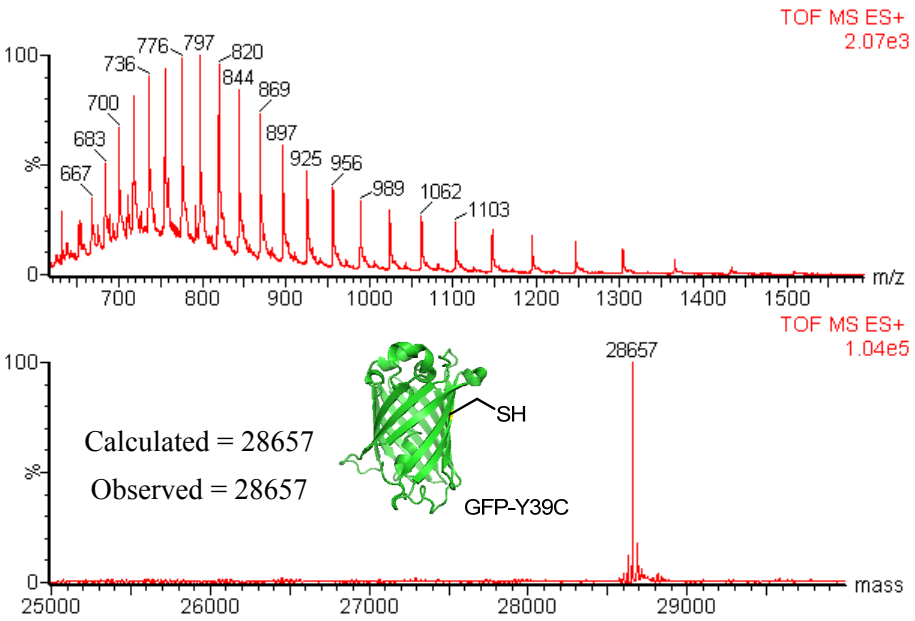
Cross-Metathesis Attempt on Green Fluorescent Protein

The GFP mutant (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:

MDYKDDDDKVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATCGKLTLKFILT
TGKLPVPWPTLVTTLTYGVQVFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYK
TRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKANF
KIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEF
VTAAGITLGMDELYKHHHHHH

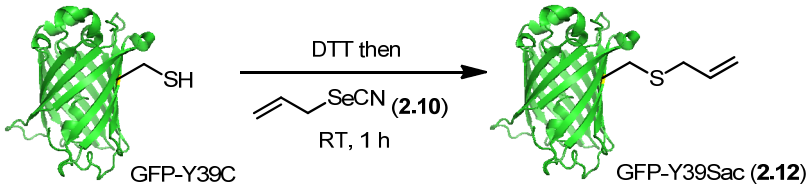
Calculated average isotopic mass = 28677 Daltons. Upon maturation and loss of H_2O and H_2, the calculated average mass of the protein is 28657 Daltons.

The mutant GFP was fluorescent upon receipt and supplied at a concentration of 1.0 mg/mL in pH 8 sodium phosphate buffer (50 mM). A solution of DTT was prepared by dissolving 3.5 mg in 700 µL of pH 8 NaP_i buffer. 4 µL of this DTT solution was added to a 40 µL aliquot of the protein solution. The sample was incubated for 10 minutes at RT and then analyzed by LC-MS. ESI-MS of the protein after DTT treatment is shown below.

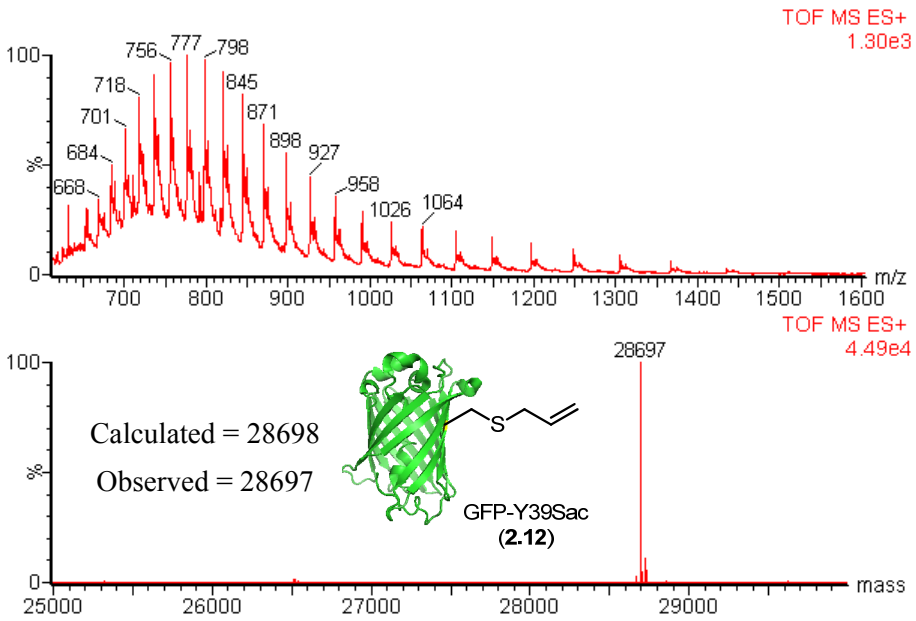


Synthesis of GFP-Y39Sac (2.12) using allyl selenocyanate

This experiment was carried out with Justin Chalker.

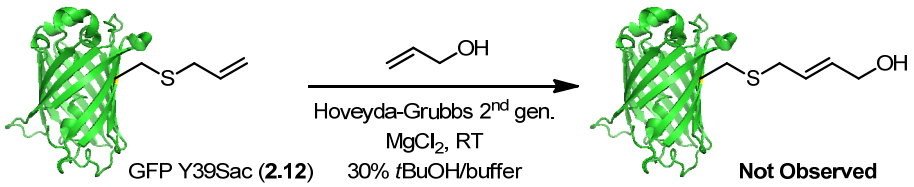


GFP-Y39C (100 µL, 1.0 mg/mL) in 50 mM NaP_i buffer (pH 8) was thawed and stored on ice. A stock solution of DTT was prepared by dissolving 3.5 mg (0.023 mmol) in 700 µL of 50 mM NaP_i buffer (pH 8). A stock solution of allyl selenocyanate was prepared by dissolving 14.3 mg (0.098 mmol) in 140 µL of DMF. A 10 µL of the DTT solution (0.035 µmol) was added to the protein sample and shaken at RT for 10 minutes to reduce any contaminant disulfide. After reduction, 5 µL of the allyl selenocyanate solution (3.5 µmol) was added to the protein solution. The reaction was mixed on a vortex and shaken at RT for 1 hour. The reaction was analyzed directly by LC-MS, revealing complete conversion to the allylated product **2.12** (calculated mass = 28698; observed mass = 28697). MS for the reaction is shown below.



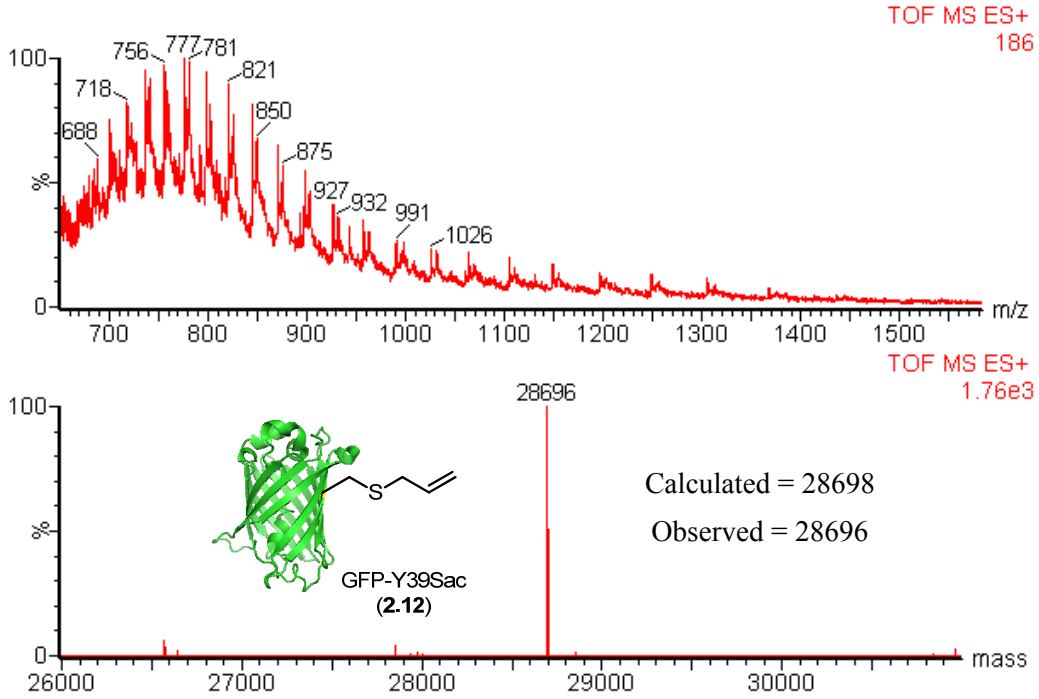
The reaction was repeated on a 400 µL scale in the same manner. LC-MS revealed full conversion to the desired allylated product. The two batches were combined and pass through a PD Minitrap (GE Healthcare) to remove small molecules. The product was eluted with 1 mL of pH 8 NaP_i (50 mM) to obtain the purified product with a final concentration of 0.5 mg/mL. LC-MS analysis revealed no change after purification. The sample was divided into 200 µL aliquots, flash frozen and stored at –20 °C. The protein remained fluorescent throughout all manipulations described above.

Cross-Metathesis of GFP-Y39Sac (2.12) with allyl alcohol

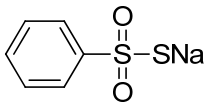


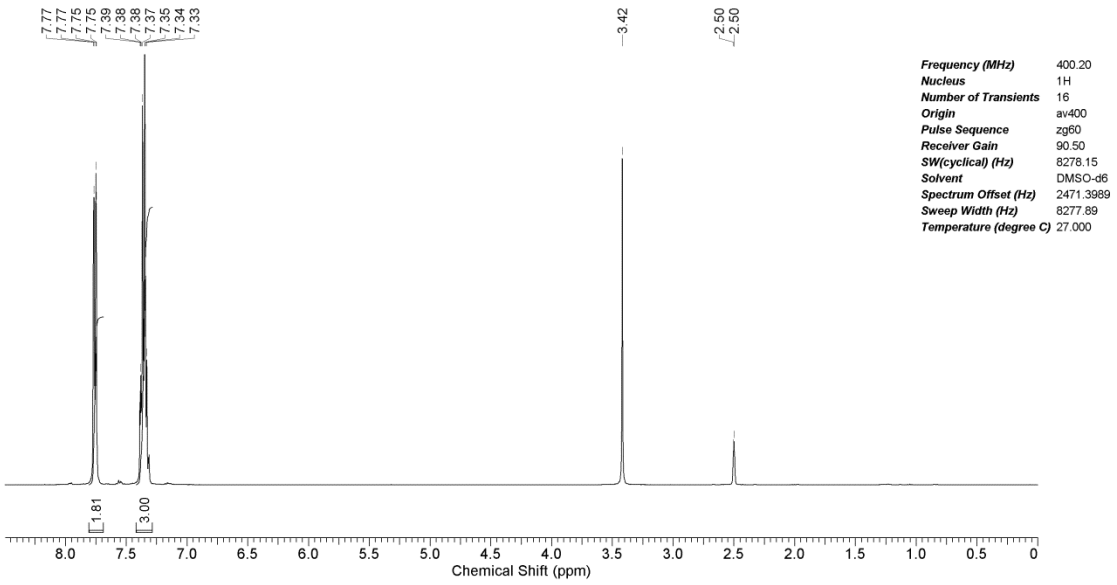
A solution of Hoveyda-Grubbs 2nd generation catalyst in ^{t}BuOH was prepared by vortexing and gently warming 1.3 mg of Ru catalyst in 253 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing GFP-Y39Sac (**2.12**) (200 µL, 0.5 mg/mL, 3.5 nmol) was added $MgCl_2 \cdot 6H_2O$ (6.7 mg, 33 µmol) followed by an aliquot of the catalyst/^{t}BuOH solution (86 µL). The reaction was vortexed immediately to mix after each addition. The

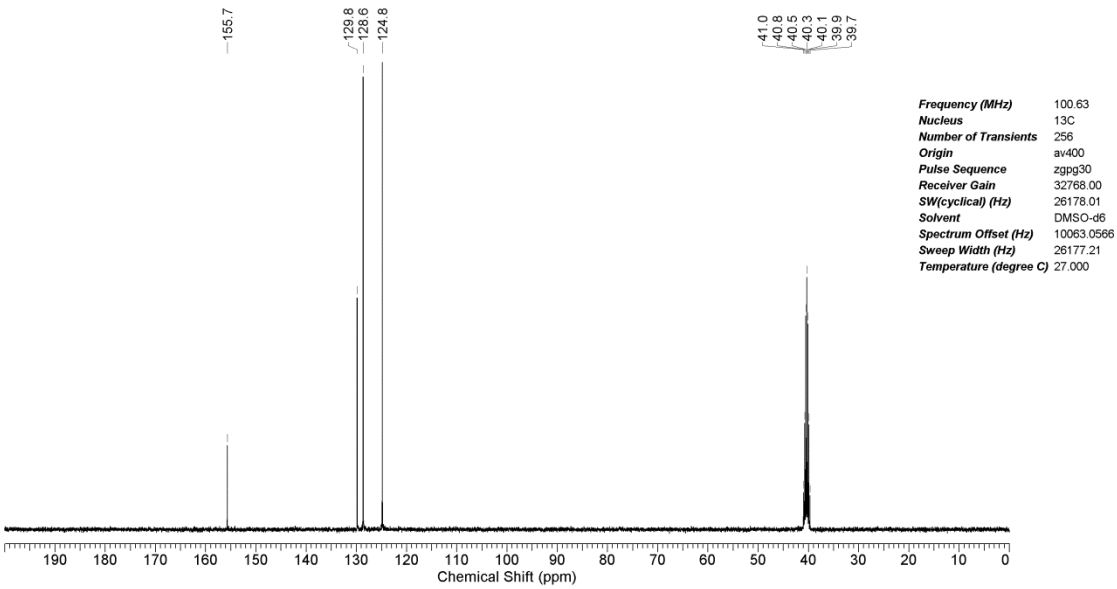
reaction mixture appeared as a pale green emulsion and was shaken at RT for 1 minute before allyl alcohol (2.4 µL, 35 µmol) was added. The reaction was vortexed and then shaken at RT. LC-MS analysis after 1 hour revealed only the starting material **2.12**. After an additional hour, no change was observed by LC-MS. ESI-MS of the reaction is shown below.

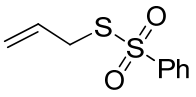


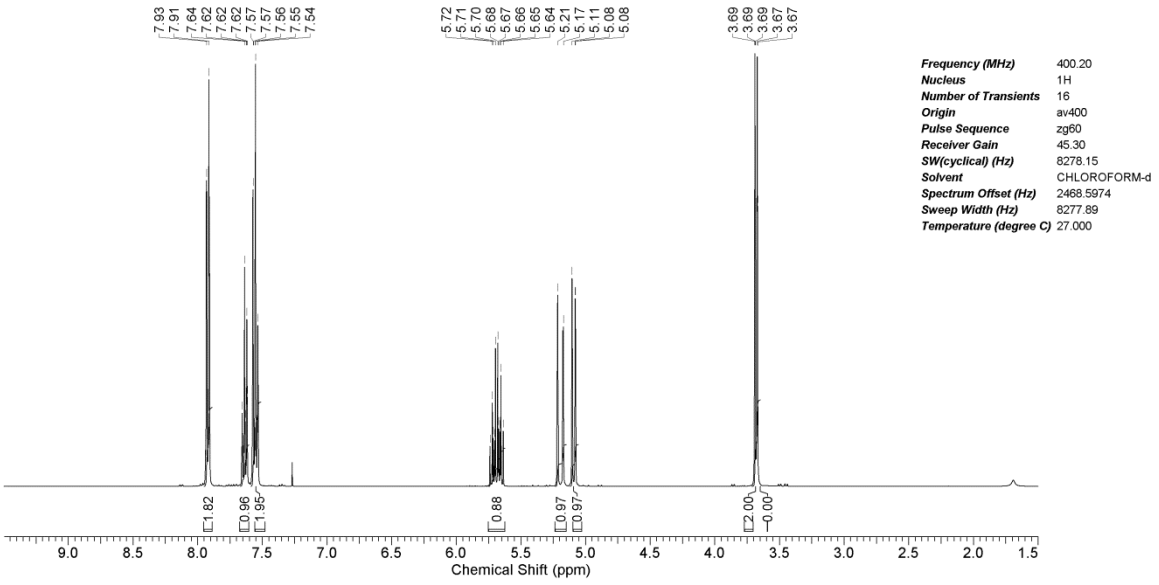
2.6 NMR Spectra

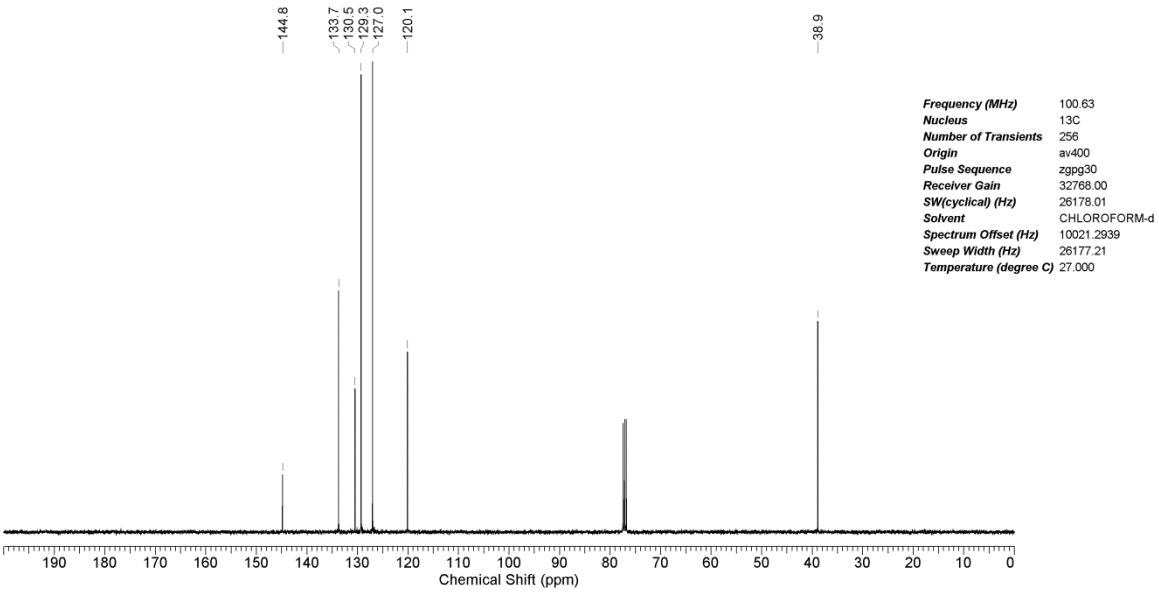


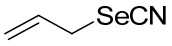


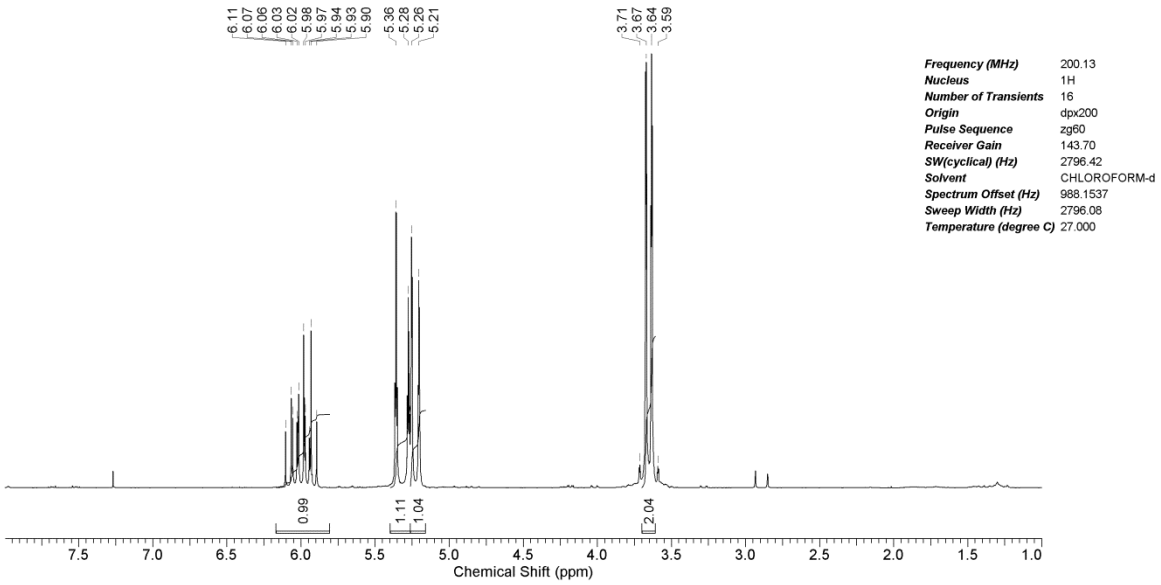


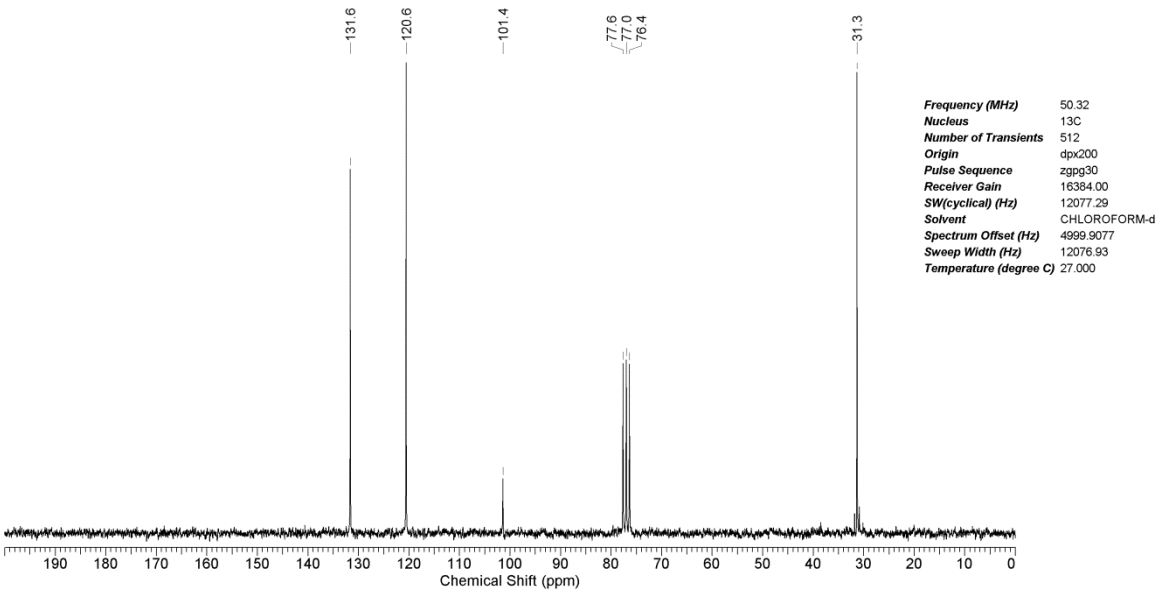

S
O
S
O
Ph


7.93
7.91
7.64
7.62
7.62
7.62
7.57
7.57
7.56
7.55
7.54
5.72
5.71
5.70
5.68
5.67
5.66
5.65
5.64
5.21
5.17
5.11
5.08
5.08
3.69
3.69
3.69
3.67
3.67
1.82
0.96
1.95
0.88
0.97
0.97
2.00
0.00
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
Chemical Shift (ppm)
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 45.30
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000


144.8
133.7
130.5
129.3
127.0
120.1
38.9
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)
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

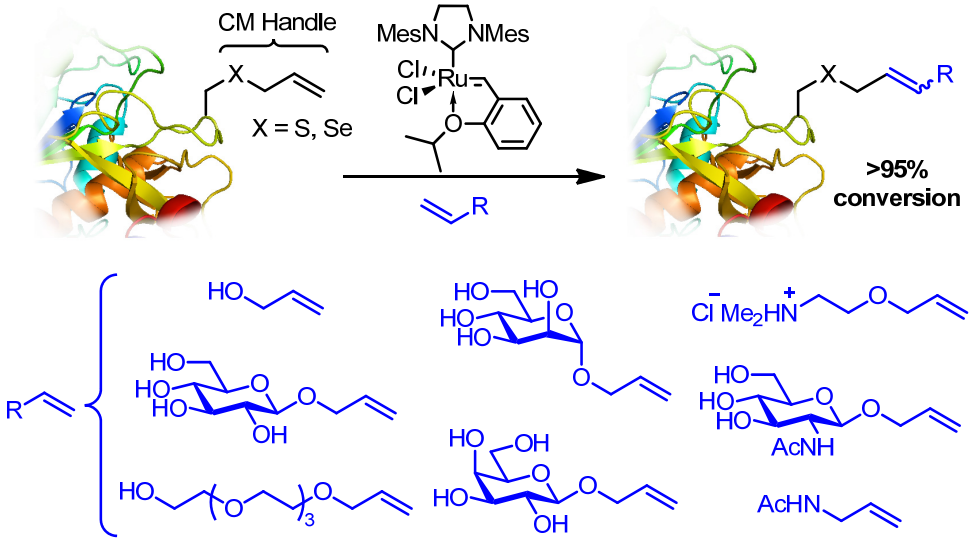

SeCN


6.11
6.07
6.06
6.03
6.02
5.98
5.97
5.94
5.93
5.90
5.36
5.28
5.26
5.21
3.71
3.67
3.64
3.59
0.99
1.11
1.04
2.04
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
Chemical Shift (ppm)
Frequency (MHz) 200.13
Nucleus 1H
Number of Transients 16
Origin dpx200
Pulse Sequence zg60
Receiver Gain 143.70
SW(cyclical) (Hz) 2796.42
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 988.1537
Sweep Width (Hz) 2796.08
Temperature (degree C) 27.000


131.6
120.6
101.4
77.6
77.0
76.4
31.3
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)
Frequency (MHz) 50.32
Nucleus 13C
Number of Transients 512
Origin dpx200
Pulse Sequence zgpg30
Receiver Gain 16384.00
SW(cyclical) (Hz) 12077.29
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 4999.9077
Sweep Width (Hz) 12076.93
Temperature (degree C) 27.000

Chapter 3

Cross-Metathesis on Proteins: An Investigation of the Substrate Scope and Steric Effects

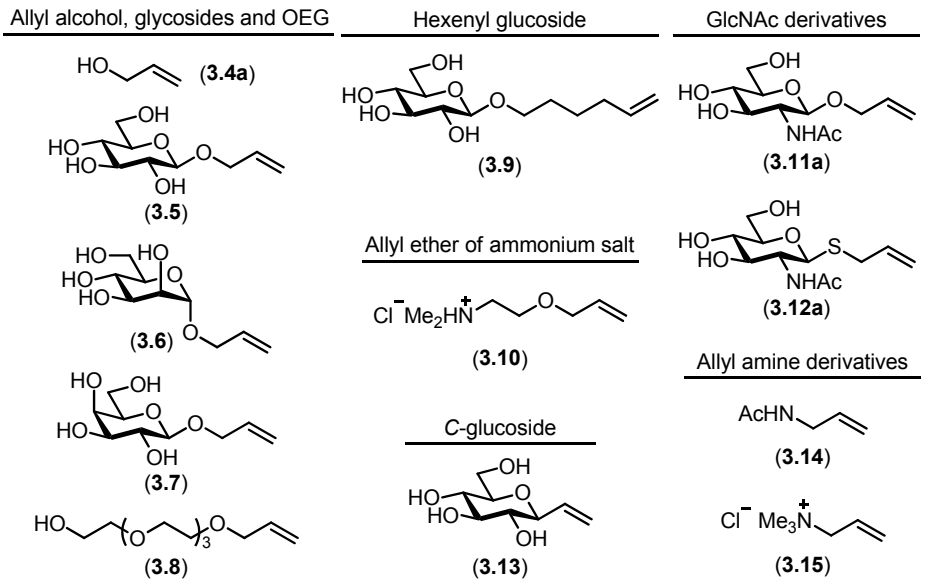


3.1 Introduction

Our preliminary reports demonstrated the enhanced reactivity of *S*-allyl cysteine (Sac) in olefin metathesis and its use in covalent protein modification. However, the cross-metathesis (CM) substrates tested were largely limited to simple allylic alcohols and the full scope of CM as a method for bioconjugation remains unclear. For the general use of olefin metathesis as a bioconjugation technique, understanding the scope and limitations of metathesis substrates is essential. Indeed, Grubbs has established some guiding principles of substrate selection for CM in organic solvents,[1] yet additional factors must be considered for successful CM on protein substrates. Most notably, the

reaction must proceed rapidly in water at, or near, room temperature. The metathesis partners and any intervening linker must therefore be selected with these stringent requirements in mind. Having established various facile chemical routes for installing Sac on protein surfaces, described in Chapter 2,[2] the unsuccessful metathesis observed on GFP-Y39Sac warrants the need for further investigation into the steric sensitivity of olefin metathesis at the reaction site. Moreover, a clearer understanding of the scope and limitations of metathesis partners may also allow olefin metathesis to be deployed more routinely in bioconjugation. Progress in the genetic incorporation of alkene-containing unnatural amino acids further drives the development of olefin metathesis as a strategy for protein modification.[3] Finally, an increased understanding of chemical behavior in aqueous olefin metathesis is useful in general synthetic endeavors.[4] These considerations motivated our investigation of aqueous CM on protein substrates.

The model protein used in the following studies was a single cysteine mutant of subtilisin from *Bacillus lentus* (SBL-S156C). While a few examples of cross-metathesis on the single Sac mutant of SBL (**3.2**, SBL-156Sac) using Hoveyda-Grubbs second generation catalyst (**3.1**)[5] were demonstrated in the preliminary work (Chapter 1, Section 1.6), the substrate scope of the reaction was not fully assessed. The substrates screened previously included simple allylic alcohols and ethers.[2,6] In fact, only the cross-metathesis between the protein substrate and allyl alcohol proceeded to completion. Since the project goal is aimed at biochemically relevant protein modifications, allyl ethers containing carbohydrates, oligo-ethylene glycols, and charged groups are among the metathesis substrates synthesized for this study (see experimental procedures). Compounds containing an allyl sulfide, *N*-allyl amines, or longer alkene tethers were also metathesis partners of interest for the assessment of substrate scope (Scheme 3-1).



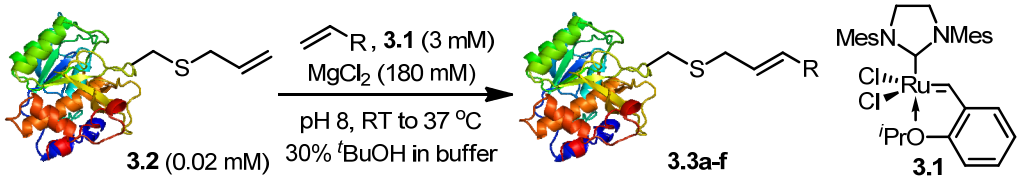
Scheme 3-1: Alkene Substrates Screened in Cross-Metathesis on Proteins.

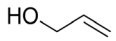
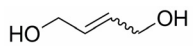
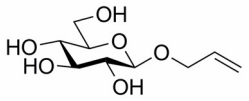
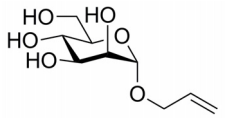
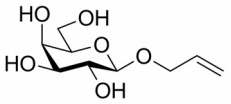
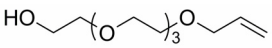
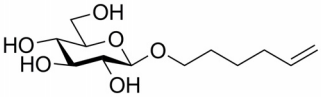
3.2 Results and Discussions

3.2.1 Substrate Scope of Cross-Metathesis on SBL-156Sac (3.2)

S-allylcysteine (Sac) was chemically installed on SBL-S156C by direct allylation with allyl chloride, as described in Chapter 2 to generate **3.2**.[7] The progress of CM reaction was monitored by ESI-MS for up to 2 hours, the reaction time necessary for complete conversion with allyl alcohol as the metathesis partner (Table 3-1, Entry 1). When protein **3.2** was tested with each of the CM substrates, best results were obtained with allylic alcohols, ethers, and hexenyl glucoside **3.9**. The reaction worked moderately well with allyl glycosides **3.5** to **3.7**, and oligo-ethylene glycol derivative **3.8**, with conversions ranging from 30% to 65% (Table 3-1, Entry 3–6). CM with hexenyl glucoside **3.9** importantly revealed the sensitivity of the reaction to linker length, with full conversion to the modified protein after only one hour at room temperature being observed (Table 3-1, Entry 7). This result compares favorably to the allyl glycosides in Entries 3–5 and was the first carbohydrate bearing substrate to proceed with full conversion.

Table 3-1 Substrate Scope of Cross-Metathesis with SBL-156Sac (**3.2**)



Entry	Alkene (mM)	Conditions	Prod.	Conv. (%)[a]
1	 **3.4a** (180)	RT, 2 h	**3.3a**	>95
2	 **3.4b** (180)	RT, 2 h then 37 °C, 30 mins	**3.3a** **3.3a**	28 >95
3	 **3.5** (92)	37 °C, 1 h	**3.3b**	30
4	 **3.6** (89)	37 °C, 1 h	**3.3c**	30
5	 **3.7** (92)	37 °C, 1 h	**3.3d**	30
6[b]	 **3.8** (92)	37 °C, 30 mins	**3.3e**	65
7	 **3.9** (92)	RT, 1 h	**3.3f**	>95

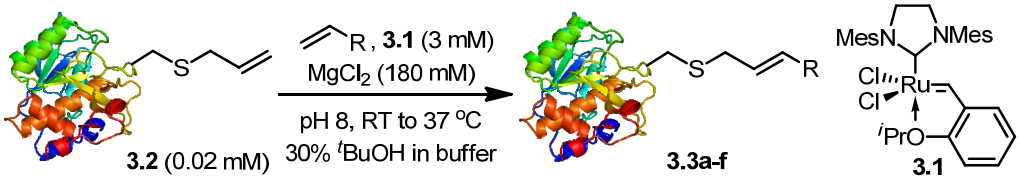
[a] Determined by LC-MS. [b] First 2 h at RT

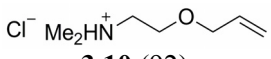
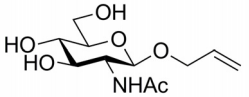
As an additional guide to mechanism, CM of protein **3.2** with the self-metathesis product of allyl alcohol (**3.4b**) was carried out to test whether, and under what conditions it is a reactive substrate. The self-metathesis product of allyl alcohol used in Table 3-1 was isolated from a model cross-metathesis of small molecules in water as described in previously reported work.[6] Interestingly, CM of **3.4b** with protein **3.2** only reached 28% conversion under the same reaction conditions used for allyl alcohol (Table 3-1, Entry 2) thus suggesting that the CM observed in Entry 1 was mainly with allyl alcohol and not the corresponding homodimer. This difference in conversion is likely due to a higher rate

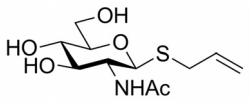
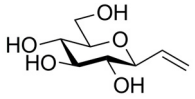
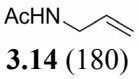
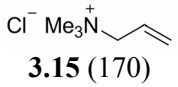
of metathesis of **3.4a** than **3.4b** since the latter is a more substituted alkene and therefore generally slower in olefin metathesis. However, when heated to 37 °C the reaction with **3.4b** proceeded with over 95% conversion after 30 minutes. This result indicates that while slower than allyl alcohol, the homodimer can successfully re-enter the metathesis cycle (Table 3-1, Entry 2). This observation is consistent with other reports on the reversible nature of olefin CM.[8]

A number of CM partners failed to react with the Sac containing proteins. No product formation was observed in the CM between protein **3.2** and substrates **3.10**–**3.15** (Table 3-2, Entries 8–13). In all of these attempts only starting material was detected by LC-MS. These results indicated the structural features of metathesis partners that may adversely affect the rate of reaction. For instance, alkenes **3.10** and **3.15** contain electron withdrawing groups known, to retard olefin metathesis.[9] *N*-Acetyl glucosamine derived substrates **3.11a** and **3.12a** also gave no detectable CM product. A comparison of the chemical structure of GlcNAc **3.11a** with glucoside **3.5** suggests that the acetamide at C-2 in **3.11a** must contribute some adverse steric, electronic, or chelating influence that impedes metathesis.

Table 3-2 Substrate Scope of Cross-Metathesis with SBL-156Sac (**3.2**) – Continued

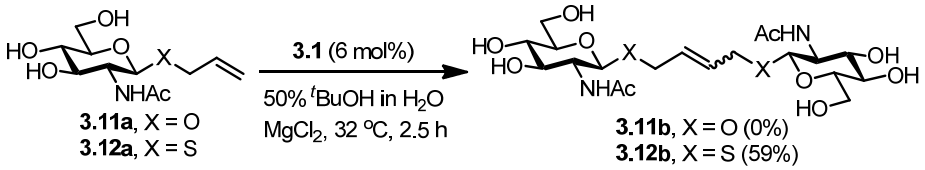


Entry	Alkene (mM)	Conditions	Prod.	Conv. (%)[a]
8	 **3.10** (92)	37 °C, 1 h	-	0
9	 **3.11a** (85)	37 °C, 2 h	-	0

10	 **3.12a** (85)	37 °C, 1 h	-	0
11	 **3.13** (85)	37 °C, 1 h	-	0
12	 **3.14** (180)	37 °C, 1 h	-	0
13	 **3.15** (170)	37 °C, 1 h	-	0

[a] Determined by LC-MS.

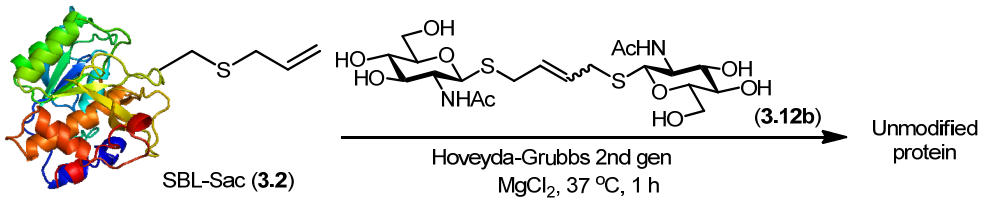
The GlcNAc derivative **3.12a** was synthesized with the intention of potentially overcoming these obstacles, since the allyl sulfide linker might be expected to be more reactive. However, only self-metathesis was observed (See Experimental Section). Apparently, the self-metathesis product of **3.12a** (**3.12b**), unlike the homodimer of allyl alcohol, cannot re-enter the metathesis cycle. We attributed this low reactivity of **3.12b** to its hindered structure. The propensity for self-metathesis and the resulting reactivity of the self-metathesis product are therefore important considerations in substrate selection. To further investigate this observation, a model reaction comparing the rate of self-metathesis between **3.11a** and **3.12a** was carried out (Scheme 3-2).



Scheme 3-2: Comparison in rate of self-metathesis between **3.11a** and **3.12a**.

The formation of the self-metathesis product of **3.12a** was observed within the first 30 minutes of reaction. After 2.5 hours, the self-metathesis product **3.12b** was isolated in a yield of 59%. In contrast, the self-metathesis of **3.11a** only resulted in >95% recovery of starting material under the same reaction conditions. When **3.12b** was tested in CM with

protein **3.2**, no reaction was observed after 1 hour at 37 °C (Scheme 3-3).



Scheme 3-3: CM between protein **3.2** and metathesis partner **3.12b**.

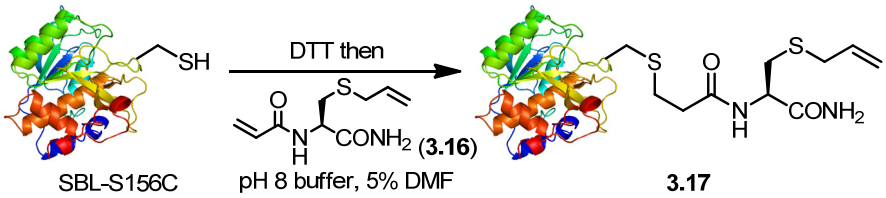
Reaction with vinyl *C*-glucoside **3.13** was unfruitful likely because of the steric congestion at the alkene. Allyl acetamide **3.14** performed poorly in CM with protein **3.2** possibly due to the formation of a stable six-member ring chelate via coordination of the carbonyl oxygen to ruthenium, poisoning the catalyst.

From these initial results, it seems that in order for CM on the Sac-containing protein to work efficiently, the metathesis partner needs to be less reactive than the allyl sulfide, but not so unreactive that no metathesis occurs. If the metathesis partner is highly reactive (e.g. **3.4a** or **3.12a**), the product of self-metathesis must be able to re-enter the cycle or no protein modification is observed. Of the substrates tested, allylic alcohols and ethers, and hexenyl glucoside **3.9** stood out as the most productive metathesis partners for Sac-containing protein **3.2**. The remaining substrates (Table 3-2) provide a benchmark of challenging transformations that can perhaps be achieved by altering the protein metathesis partner or linker. Accordingly, the accessibility of the Sac residue on the protein surface and its effect on cross-metathesis is assessed next.

3.2.2 Investigation of the Sterics Around the Alkene Tether on Protein Substrates and Its Effects in Cross-Metathesis

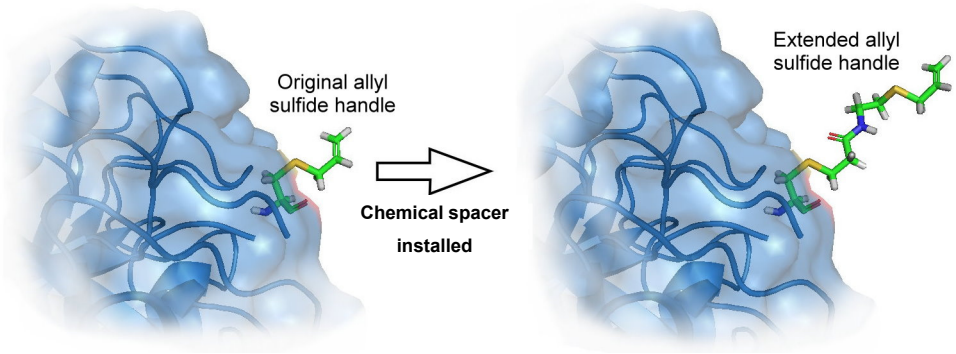
The CM reactions failed to reach full conversion with metathesis partners such as allyl glycosides and oligo-ethylene glycols. In other cases no productive CM was observed,

particularly for sterically demanding and electron-poor substrates. Referring back to the proposed sulfur-relayed metathesis mechanism (Chapter 1, Scheme 1-21a), coordination of the Sac residue to the ruthenium is critical for a rapid reaction. In addition, olefin metathesis on small molecules has been shown to be sensitive to steric hindrance. Thus the influence of sterics on CM at the protein surface was investigated. Accordingly, compound **3.16** was synthesized and used to install a Sac residue at the same protein site, but extended from the protein surface to create a less hindered olefin. Conjugate addition of the cysteinyl residue on SBL-S156C to acrylamide **3.16** led to full conversion of the alkylated protein **3.17** after incubation at 37 °C for 1 hour (Scheme 3-4). The conjugation at cysteine was verified by Ellman's assay (See Experimental Section).



Scheme 3-4: Installing extended *S*-allylcysteine on protein surface.

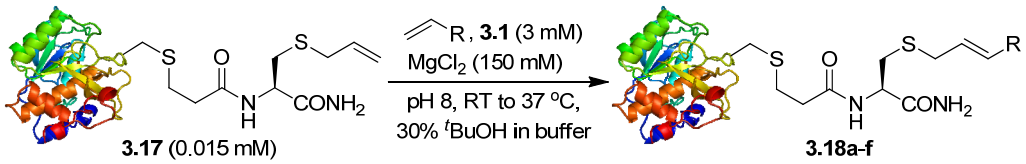
A computer protein model was also generated with PyMol® showing before and after the chemical spacer installation (Scheme 3-5). The model showed that the C–C double bond of the allyl sulfide moiety on SBL is approximately 3.0 Å from the protein surface. The bulkiness of the NHC ligand on Hoveyda-Grubbs 2nd generation catalyst together with the PyMol® model, further suggest that carrying out CM at the active sites or α-helices of proteins would be challenging. Nevertheless, the chemical spacer extended the reactive allyl sulfide metathesis handle to 10.7 Å away from the protein surface suggesting that a more rapid pre-coordination of the protein to the catalyst could be expected and hence results in a faster rate of reaction.

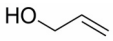
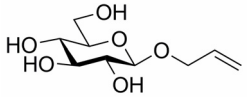
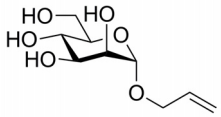


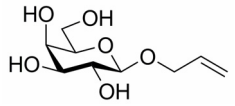
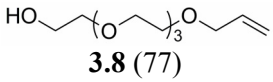
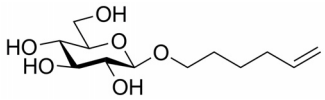
Scheme 3-5: PyMol model of SBL-156Sac (**3.2**) and SBL-Sac extended (**3.17**).

With protein **3.17** in hand, CM with allyl alcohol (**3.4a**) and ether substrates **3.5** to **3.9** was carried out. Notably, all reactions proceeded to full conversions. The results summarized in Table 3-3 demonstrated that indeed the local protein environment affect CM on the SBL protein surface. CM with allyl alcohol reached full conversion after just 30 minutes of reaction time at room temperature, while the same reaction with protein **3.2** required 2 hours to reach completion. Protein CM with hexenyl glucoside **3.9** also proceeded with full conversion. Again, a shorter reaction time was required with the protein containing the extended linker.

Table 3-3 Cross-Metathesis on SBL **3.17** – Extended Sac

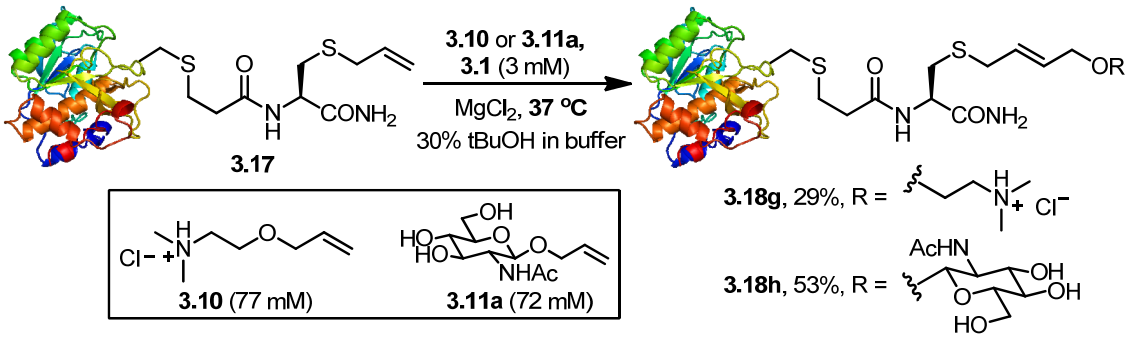


Entry	Alkene (mM)	Conditions	Prod.	Conv. (%)[a]
1	 **3.4a** (76)	RT, 30 mins	**3.18a**	>95
2	 **3.5** (77)	37 °C, 1 h	**3.18b**	>95
3	 **3.6** (72)	37 °C, 1 h	**3.18c**	>95

4	 **3.7** (77)	37 °C, 1 h	**3.18d**	>95
5	 **3.8** (77)	RT, 2 h	**3.18e**	>95
6	 **3.9** (80)	RT, 30 mins	**3.18f**	>95

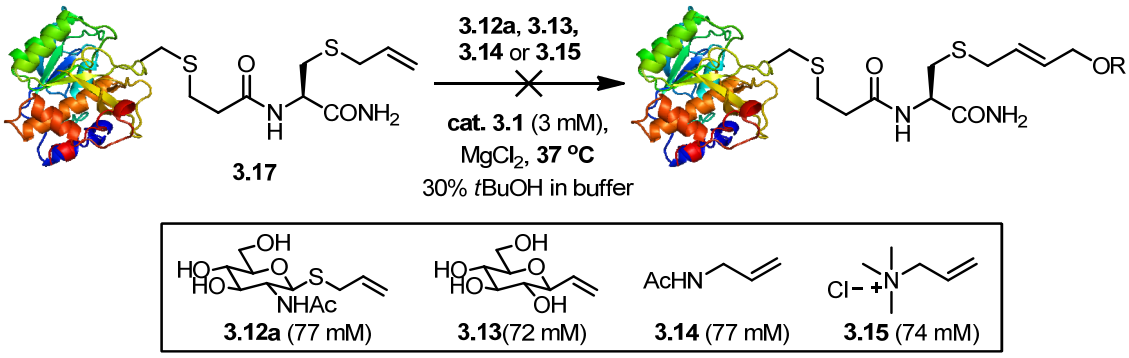
[a] Determined by LC-MS.

More challenging substrates such as ethanolamine **3.10** and GlcNAc **3.11a** only gave 29% and 53% conversion, respectively with the Sac extended protein **3.17** (Scheme 3-6). Despite the low conversions, the use of a linker to increase the efficiency of protein modification via CM was promising since these were the only two allyl ether substrates that had failed to work with protein **3.2**. The influence of steric effects should thus be considered, especially if the modification is to be at a hindered active site or α-helix.



Scheme 3-6: Cross-Metathesis of Protein **3.17** with Ethanolamine **3.10**/GlcNAc **3.11a**.

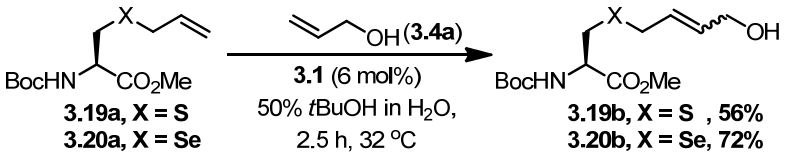
Substrates **3.12a** to **3.15** remained challenging, even on the Sac-extended protein **3.17** (Scheme 3-7). Substrate **3.12a** again only resulted in unproductive self-metathesis (See Experimental Section) whereas compounds **3.13** to **3.15** were either too electron-deficient or sterically demanding to participate in CM.



Scheme 3-7: Substrates that failed to react with protein **3.17** by cross-metathesis.

3.2.3 Allyl Selenides are Substrates with Cross-Metathesis Reactivity Superior to Allyl Sulfides in Aqueous Media

From literature reports regarding the positive influence of allylic alcohols[10] and our initial report on allyl sulfides[6] in olefin metathesis, it was tempting to consider if such rate enhancing observation were general for allylic chalcogenides.[11] Pursuing this hypothesis, allyl selenides were naturally the next candidate of interest in the chalcogen group to be examined in model CM reactions. Thus *Se*-allylselenocysteine (Seac) derivative **3.20a** was synthesized and tested along with Sac derivative **3.19a** in a model aqueous CM reaction with allyl alcohol under identical conditions. The reaction with Seac **3.20a** was found to be higher yielding than the equivalent sulfur case, with respective yields of 72% and 56% (Scheme 3-8).

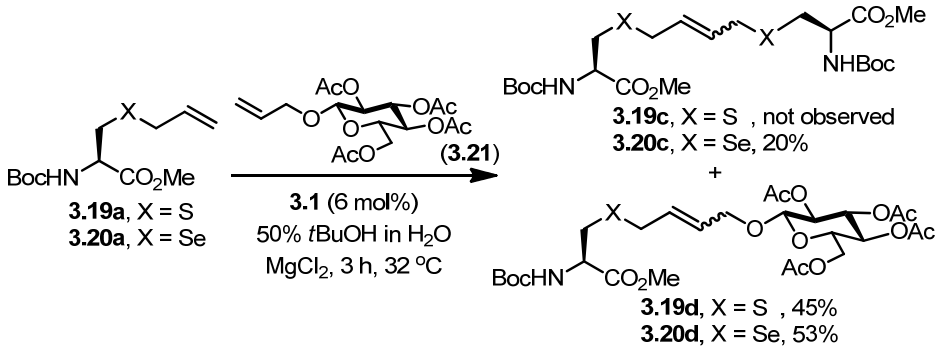


Scheme 3-8: Comparison of reactivity between Sac and Seac in aqueous CM with allyl alcohol.

This difference in reactivity may be attributed to the softness of selenium which makes the coordination to ruthenium even more favorable than for sulfur in Sac. Remarkably, there are only few examples in the literature describing olefin metathesis with selenium-

containing compounds. In one instance, Koketsu and co-workers used RCM of an allyl selenide derivative as a key step for the synthesis of selenium-containing bicyclic-lactams.[12] However, the scope of olefinic selenoethers in olefin metathesis was not manifested in these reports, and their enhanced metathesis reactivity was not noted.

Next, a more complex and biochemically important carbohydrate metathesis partner **3.21** was used in place of allyl alcohol. The CM reaction of glucoside **3.21** with Sac **3.19a** and Seac **3.20a** gave moderate yields of 45% and 53%, respectively. Initial inspection of the yields of CM products **3.19d** and **3.20d** suggests no significant difference in reactivity between Sac and Seac. However, self-metathesis product **3.20c** was also isolated from the reaction of Seac, whereas no self-metathesis product of Sac (**3.19c**) was observed (Scheme 3-9).

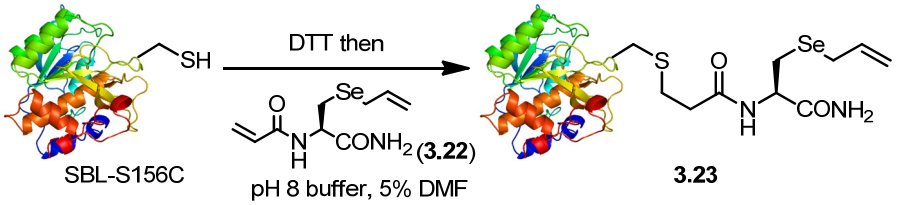


Scheme 3-9: Comparison of reactivity between Sac and Seac in aqueous CM with **3.21**.

The total CM yields (CM and self-metathesis) are therefore 73% for Seac and 45% for Sac, a clear indication that not only are allyl selenides reactive in cross-metathesis but are also more reactive than the corresponding allyl sulfide. Taking advantage of this enhanced reactivity of allyl selenides in protein conjugation, especially its ability to promote challenging CM with substrates that were sluggish or unreactive with Sac, the substrate scope of proteins containing an allyl selenide tag was examined next.

3.2.4 Exploring the Substrate Scope of Cross-Metathesis with Proteins Containing an Allyl Selenide Tag

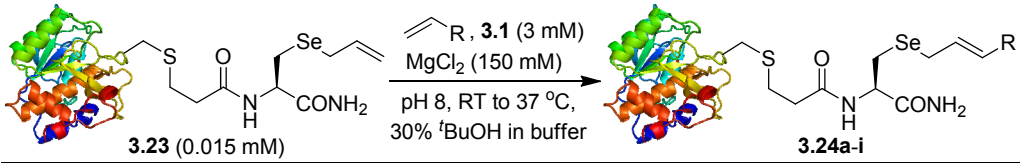
After establishing the increased metathesis reactivity of Seac in small molecules, an allyl selenide tag was introduced into the protein SBL-S156C to see if the rate enhancement is translated in protein systems. In order to directly compare the CM reactivity with protein **3.17**, Seac-containing protein **3.23** was synthesized in a similar manner using Seac acrylamide **3.22** (Scheme 3-10). Again, reaction at cysteine was verified via Ellman's assay (See Experimental Section).

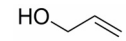
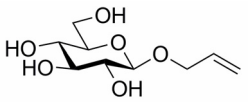
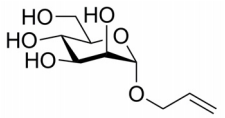
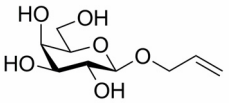
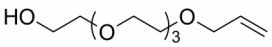
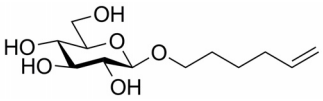
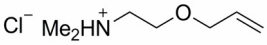
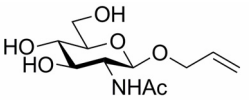
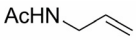


Scheme 3-10: Installing extended *Se*-allylselenocysteine on protein surface.

The remarkable reactivity of allyl selenide containing protein **3.23** was clear after the first test in CM. The CM between allyl alcohol and protein **3.23** required only *15 minutes at room temperature to reach completion* (Table 3-4, Entry 1). Ether-based metathesis substrates also reached full conversion under mild reaction conditions (Table 3-4, Entries 2-8). Notably, these substrates included the more challenging ethanolamine **3.10** and GlcNAc **3.11a**, which proceeded with poor conversion in previous attempts on protein **3.17**. CM with the reactive hexenyl glucoside **3.9** also reached full conversion (Table 3-4, Entry 6). Moreover, allyl acetamide **3.14**, a substrate that was previously unreactive in all previous CM reactions, also gave productive CM with protein **3.23** (Table 3-4, Entry 9). Either the unhindered Seac on protein **3.23** was able to initiate rapid CM with **3.14** before the catalyst was sequestered by the acetamide, or the allyl selenide is simply a better coordinator of ruthenium than the chelating acetamide.

Table 3-4 Cross-Metathesis on SBL **3.23** – Extended Seac

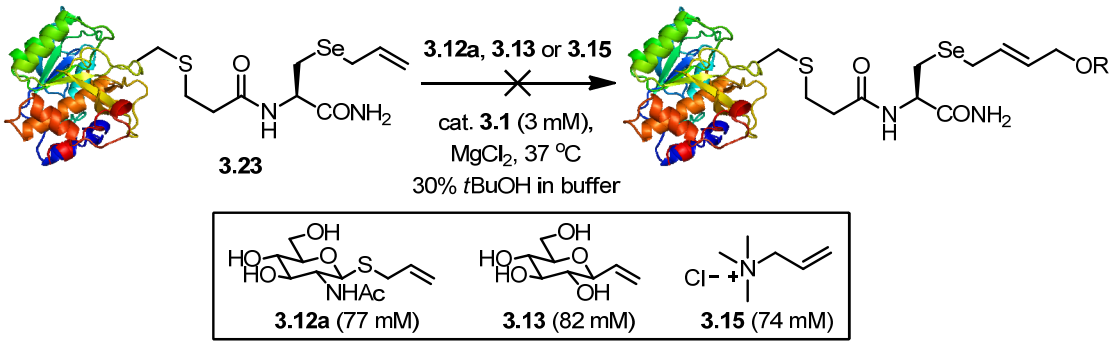


Entry	Alkene (mM)	Conditions	Prod.	Conv. (%)[a]
1	 **3.4a** (76)	RT, 15 mins	**3.24a**	>95
2	 **3.5** (77)	37 °C, 1 h	**3.24b**	>95
3	 **3.6** (74)	37 °C, 1 h	**3.24c**	>95
4	 **3.7** (77)	37 °C, 1 h	**3.24d**	>95
5	 **3.8** (77)	RT, 1 h	**3.24e**	>95
6	 **3.9** (85)	RT, 30 mins	**3.24f**	>95
7	 **3.10** (76)	37 °C, 30 mins	**3.24g**	>95
8	 **3.11** (73)	37 °C, 1 h	**3.24h**	>95
9	 **3.14** (77)	37 °C, 30 mins	**3.24i**	90

[a] Determined by LC-MS.

Among the CM reactions carried out on protein **3.23**, the modifications with GlcNAc **3.11a** and acetamide **3.14** are particularly biologically relevant modifications. The GlcNAc moiety on the glycosylated protein **3.24h** is an anchor for many modifications and bio-processes such as celluar signalling.[13] Moreover, acetylation of amines such as that found in **3.24i** is an important protein post-translational modification (PTM).[14] In

nature, N-acetyllysines has vital roles in epigenetics by regulating the binding of histone proteins to DNA.[15] The attachment of an acetyl group on proteins by CM might suffice as a mimic for the natural PTM of lysine residues in proteins.[16] The results in Table 3-4 are highly promising and it is clear that unhindered allyl selenides allow unprecedented reactivity in bioconjugation by olefin metathesis. Nonetheless, some limitations for CM using **3.23** remain. Substrates **3.12a**, **3.13**, and **3.15** did not participate in productive CM (Scheme 3-11). Likely *C*-vinyl glucoside **3.13** is too hindered and **3.14** both too hindered and electron deficient. Again, **3.12a** underwent preferential self-metathesis and no CM product was detected.

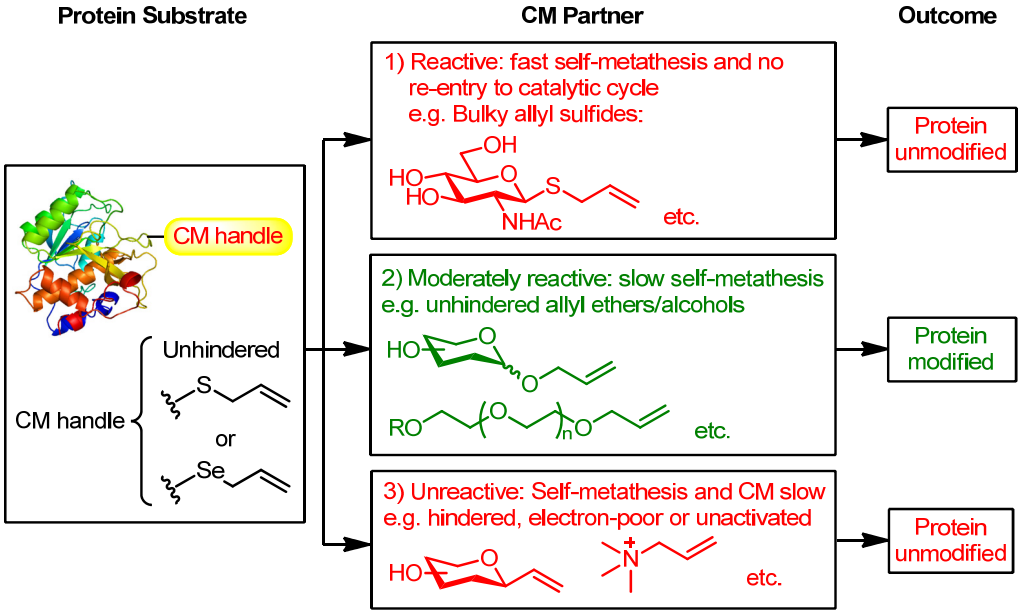


Scheme 3-11: Substrates that failed to react with protein **3.23** by cross-metathesis.

3.3 Conclusions and Outlook

In summary, by examining a range of sterically and electronically diverse olefin substrates in protein CM, we have gleaned some guiding principles for successful CM on protein substrates (Scheme 3-12). In general, an unhindered allyl sulfide or allyl selenide chemical tag on the protein surface at the site of modification is desirable. The role of this olefin partner is distinct since there is minimal risk of protein self-metathesis.[17] Extending the site of the reaction from the surface of the protein was suspected to enhance reactivity simply through steric relief and increased solvent accessibility. However, other subtle changes in the complex chemical environment of the protein that may account for this difference in reactivity cannot be ruled out. For Sac and Seac

containing proteins, allyl ethers make good metathesis partners as they undergo slow self-metathesis compared to allyl sulfides, allowing sufficient amount of unsubstituted alkene for productive CM. However, allyl ethers are not the only effective metathesis partners as other olefins such as hexenyl glucoside **3.9** can also be reactive in CM. Importantly, the metathesis partner must not form a stable chelate. If this occurs, the metathesis rate drops and little or no protein modification is observed.



Scheme 3-12: Guide to Substrate and Linker Selection for Cross-Metathesis on Proteins.

Throughout the course of our investigation, we have demonstrated that by relieving steric hindrance around the alkene and protein surface, the rate of cross-metathesis is increased. However, the steric-sensitive nature of olefin metathesis also means that modification at more hindered protein sites is a current, unmet challenge with conventional metathesis catalysts. This limitation prompts the need for a new class of metathesis catalysts for bioconjugation, where the ligand binding to the metal should be both small and water-soluble—a significant challenge given that the sterically encumbered NHC ligands impart the stability necessary for use in air and water. Additionally, allyl selenides were discovered to be superior to allyl sulfides in aqueous CM. For unhindered allyl

selenide-containing proteins, efficient CM was achieved with several substrates including carbohydrates, oligo-ethylene glycols, allyl acetamides, and even alkenes with electron withdrawing ammonium salts. It is also worth noting that at the time of first report of olefin metathesis on proteins,[6] examples of homogenous cross-metathesis in water were largely limited to simple alkenols.[9,18] In this work, a new benchmark in substrate complexity is set for olefin cross-metathesis. Complex macromolecules and metathesis partners were joined efficiently by virtue of the innate affinity of allyl sulfides and allyl selenides for ruthenium, an affinity that orchestrates rapid, productive metathesis of the alkene and alkylidene.

The promising results from CM of allyl selenides is the driving force for developing chemical and genetic strategies for incorporation of Seac and other allyl selenide derivatives on protein surfaces. The outlook of this project is that the application of these concepts and techniques in bioconjugation and synthetic chemistry will allow a largely untapped potential for allyl sulfides and allyl selenides to be realized.

3.4 References

(1) Chatterjee, A. K.; Choi, T.-L.; Sanders, D. P.; Grubbs, R. H. *J. Am. Chem. Soc.* **2003**, *125*, 11360-11370.

(2) Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714-3716.

(3) (a) van Hest, J. C. M.; Tirrell, D. A. *FEBS Letters* **1998**, *428*, 68-70. (b) Zhang, Z.; Wang, L.; Brock, A.; Schultz, P. G. *Angew. Chem., Int. Ed.* **2002**, *41*, 2840-2842. (c) Ai, H.-W.; Shen, W.; Brustad, E.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 935-937.

(4) (a) Fürstner, A. *Angew. Chem. Int. Ed.* **2000**, *39*, 3012-3043. (b) Nicolaou, K. C.; Bulger, P. G.; Sarlah, D. *Angew. Chem. Int. Ed.* **2005**, *44*, 4490-4527. (c) Hoveyda, A. H.; Zhugralin, A. R. *Nature* **2007**, *450*, 243-251.

(5) Garber, S. B.; Kingsbury, J. S.; Gray, B. L.; Hoveyda, A. H. *J. Am. Chem. Soc.* **2000**, *122*, 8168-8179.

(6) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642-9643.

(7) Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714-3716.

(8) (a) Fürstner, A.; Thiel, O. R.; Ackermann, L. *Org. Lett.* **2001**, *3*, 449-451. (b) Smith, A. B., III; Adams, C. M.; Kozmin, S. A. *J. Am. Chem. Soc.* **2001**, *123*, 990-991.

(9) (a) Hong, S. H.; Grubbs, R. H. *J. Am. Chem. Soc.* **2006**, *128*, 3508-3509. (b) Jordan, J. P.; Grubbs, R. H. *Angew. Chem. Int. Ed.* **2007**, *46*, 5152-5155.

(10) (a) Hoye, T. R.; Zhao, H. *Org. Lett.* **1999**, *1*, 1123-1125. (b) Hoveyda, A. H.; Lombardi, P. J.; O'Brien, R. V.; Zhugralin, A. R. *J. Am. Chem. Soc.* **2009**, *131*, 8378-8379.

(11) Lin, Y. A.; Davis, B. G. *Beilstein J. Org. Chem.* **2010**, *6*, 1219-1228.

(12) Garud, D. R.; Garud, D. D.; Koketsu, M. *Org. Biomol. Chem.* **2009**, *7*, 2591-2598.

(13) Hart, G. W.; Akimoto, Y. In *Essentials of Glycobiology*; 2nd ed.; Cold Spring Harbor Laboratory Press, New York, 2nd Ed.: 2009, p 263-279.

(14) Walsh, C. T.; Garneau-Tsodikova, S.; Gatto, G. J., Jr. *Angew. Chem. Int. Ed.* **2005**, *44*, 7342-7372.

(15) (a) Kouzarides, T. *Cell* **2007**, *128*, 693-705. (b) Bannister, A. J.; Kouzarides, T. *Cell Res.* **2011**, *21*, 381-395.

(16) Davis, B. G. *Science* **2004**, *303*, 480-482.

(17) No protein self-metathesis was observed by SDS-PAGE gel analysis (See ESI).

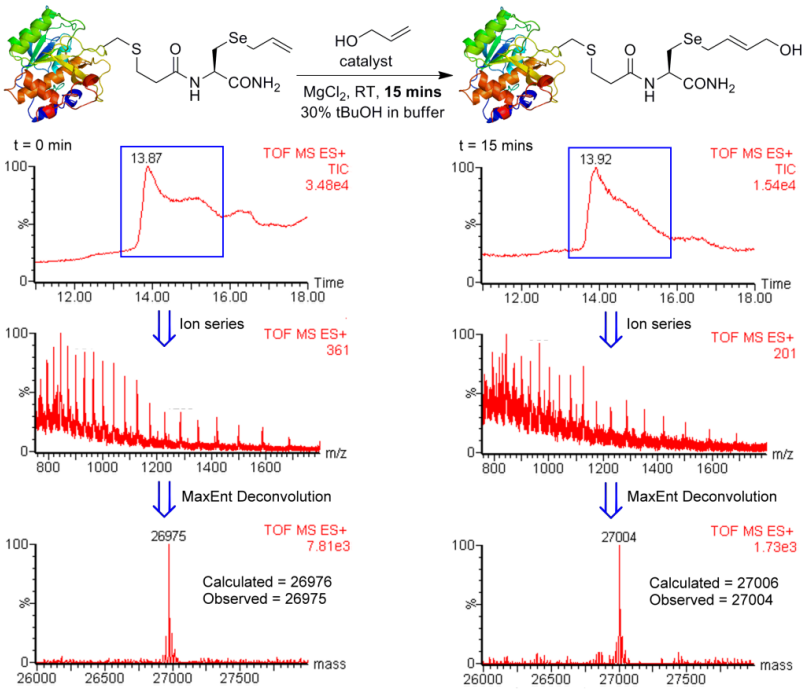
(18) (a) Binder, J. B.; Blank, J. J.; Raines, R. T. *Org. Lett.* **2007**, *9*, 4885-4888. (b) Gulajski, L.; Michrowska, A.; Naroznik, J.; Kaczmarska, Z.; Rupnicki, L.; Grela, K. *ChemSusChem* **2008**, *1*, 103-109.

3.5 Experimental Procedures

General Considerations

See Chapter 2 Experimental Procedures for general considerations in synthesis and protein analysis. In addition to the general consideration in protein analysis, a validation for using ESI-MS to follow reaction conversions on protein substrates is outlined below. Compounds **3.6–3.8**, **3.10**, **3.12a**, **3.13–3.15** were kindly provided by a formal DPhil student, Justin M. Chalker, whom I collaborated in this project. For detailed preparation of these compounds, reader is to refer to the publication that resulted from this work.[3.1]

Validation of Following Reaction Conversions on Protein Substrates by Mass Spectrometry

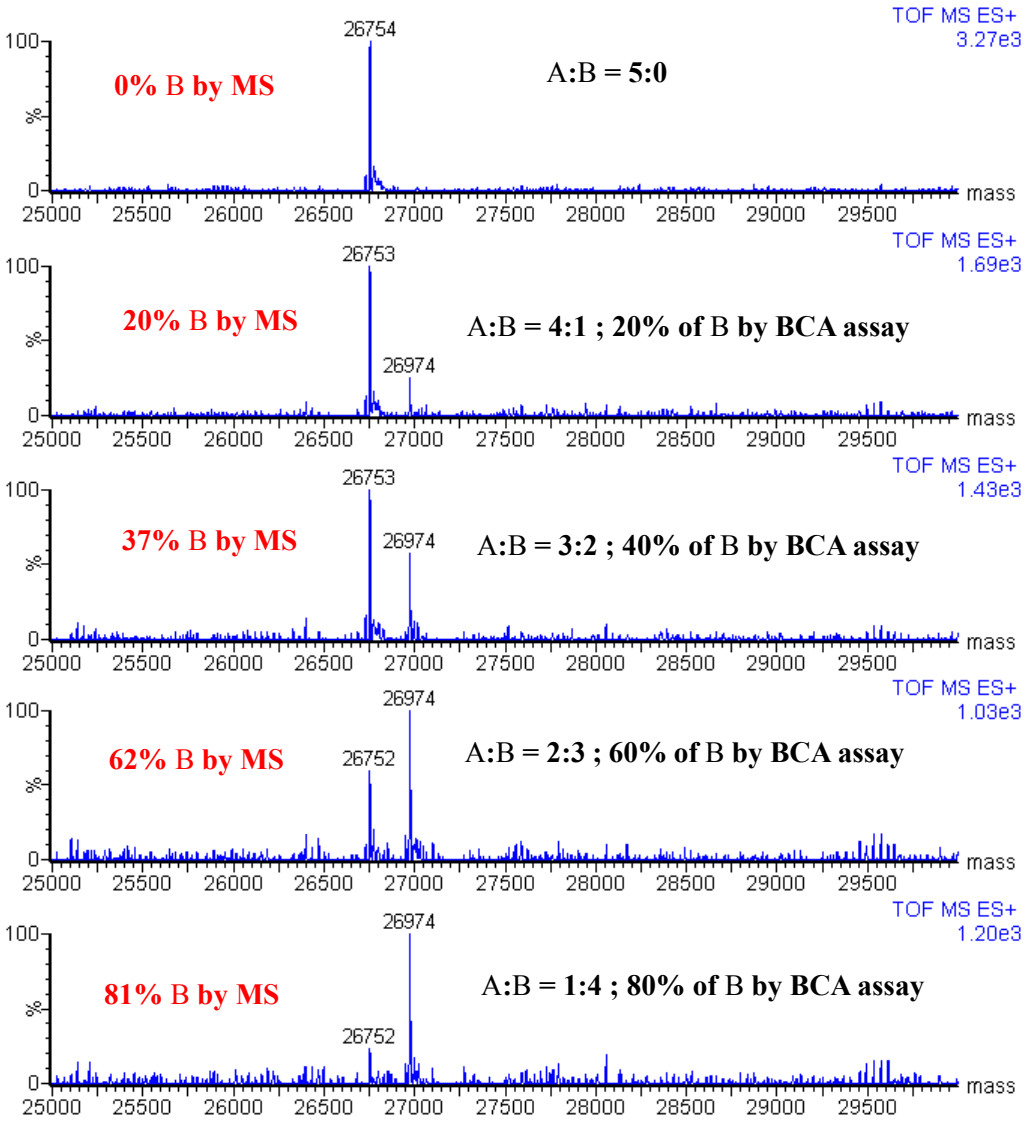


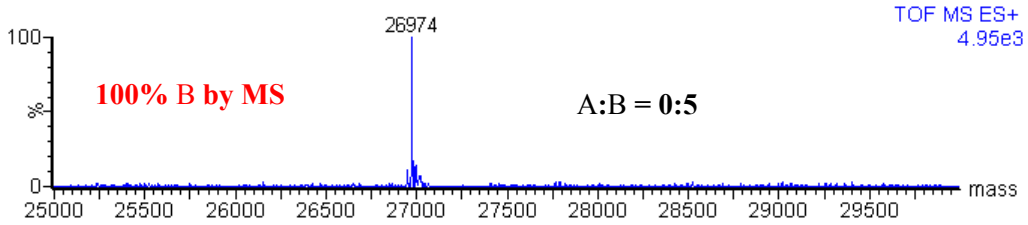
Modified and un-modified proteins were not resolved during chromatography and typically eluted between 13 and 16 minutes. MS data therefore reflects the total protein content of the reaction mixture. Conversions were calculated from the relative intensities of mass peaks after MaxEnt processing. Both the modified and unmodified proteins are

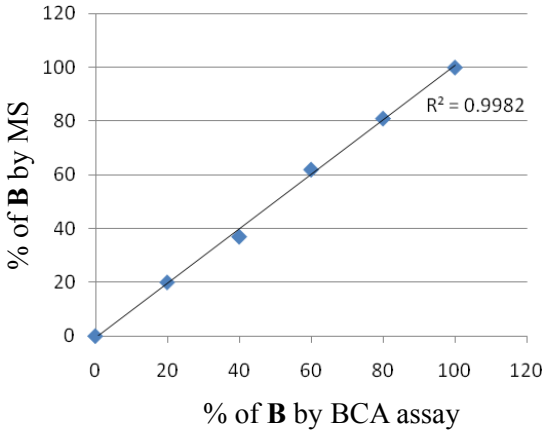
[3.1] Lin, Y. A.; Chalker, J. M.; Davis, B. G. *J. Am. Chem. Soc.* **2010**, *132*, 16805-16811.

assumed to be ionized with similar efficiency. This assumption and method for calculating reaction conversions was validated by correlating relative MS peak intensities to mixtures of proteins with known relative concentrations. This method is described below:

Six samples containing a mixture of protein **3.2** (**A**) and **3.23** (**B**) (c = 0.14 mg/mL, measured by standard bicinchoninic acid (BCA) assay from Thermo Scientific) were prepared according to the following **A**:**B** ratios (5:0, 4:1, 3:2, 2:3, 1:4 and 0:5). 45 µL of each sample was analyzed by LC-MS. The LC-MS spectra of these mixtures are shown below. The calculated masses for **A** and **B** are 26755 and 26976, respectively. A graph of the percentage of **B** by MS against percentage of **B** measured was then plotted.

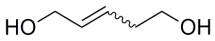






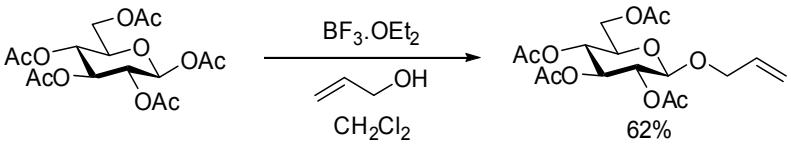
Preparation of Metathesis Substrates

2-Buten-1,4-diol (3.4b)



This is a by-product isolated from CM between BocSacOMe (**3.19a**) and allyl alcohol (**3.4a**) (See page 119 for characterization).

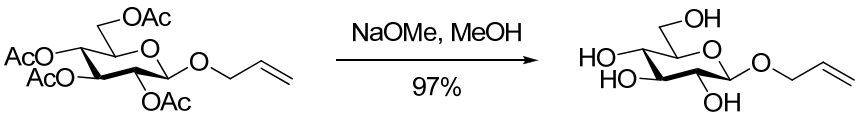
Allyl 2,3,4,6-tetra-*O*-acetyl-*β*-(D)-glucopyranoside (3.21)



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, anhydrous). The stirred solution was cooled to 0 °C and $BF_3{\cdot}OEt_2$ (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 oganic 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.[3.2] m.p. = 73-75 °C. $[\alpha]^{20}_D = -13.4$ (c = 2.0, $CHCl_3$). IR (υ_{max}, KBr): 1755, 1373, 1232, 1042. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.91, 1.94, 1.96, 1.99 (12H, 4 × s, 4 × OAc), 3.62 (1H, ddd, *J* = 9.9, 4.8, 2.5, H5), 4.02 (1H, dd, *J* = 13.1, 6.1, CH<u>H</u>CH=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, C<u>H</u>HCH=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 HC=C<u>H</u>H *cis* and H3), 5.19 (1H, dd, *J* = 17.2, 1.5, HC=CH<u>H</u> *trans*), 5.76 (1H, m, <u>H</u>C=CH_2). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 20.5 (br, contains 2 × OAc), 20.57, 20.63 (2 × OAc), 61.9 (<u>C</u>H_2CH=CH_2), 68.3 (C4), 69.9 (C6), 71.2, 71.7, 72.8 (C2, C3, C5), 99.4 (C1), 117.5 (HC=<u>C</u>H_2), 133.3 (H<u>C</u>=CH_2), 169.2, 169.3, 170.1, 170.5 (4 × C=O). LRMS m/z (ESI^+): Found 406 $[M+NH_4]^+$, 799 $[2M+Na]^+$.

Allyl-*β*-(D)-glucoside (3.5)

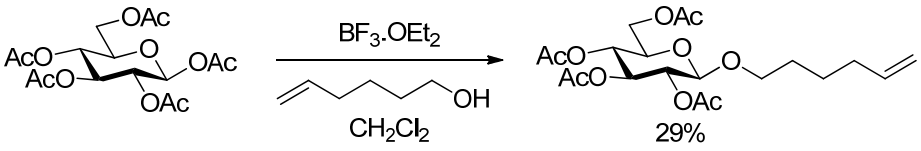


To a stirred solution of allyl 2,3,4,6-tetra-*O*-acetyl-*β*-(D)-glucopyranoside (**3.21**) (2.00 g, 5.15 mmol) in MeOH (20 ml) was added sodium methoxide (139 mg, 0.26 mmol). After stirring for 20 minutes, the reaction was neutralized with Dowex 50WX8 (H^+ form) and

3.2 Rodriguez, E. B.; Stick, R. V. *Aust. J. Chem*. **1990**, *43*, 665-679.

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.[3.3] m.p. = 96-98 °C [Lit.[3.3] = 102-103 °C, EtOH (aq)]. $[\alpha]^{20}_{D} = -39.0$ (c = 1, MeOH) [Lit.[3.3] = –41.4, c = 0.43, MeOH]. IR (υ_{max}, KBr): 3417, 2873, 1645, 1425, 1368, 1258, 1082, 919, 617. ^{1}H NMR (400 MHz, CD_3OD): δ_H = 3.18-3.41 (3H, m, contains H2, H3 and H4), 3.68 (1H, dd, *J* = 12.0, 5.5 Hz, H6), 3.89 (1H, dd, *J* = 12.0, 1.7 Hz, H6′), 3.94 (1H, dt, *J* = 12.5, 6.2 Hz, H5), 4.17 (1H, dd, *J* = 12.9, 6.0 Hz, OC<u>H</u>HCH=CH_2), 4.32 (1H, d, *J* = 7.9 Hz, H1), 4.40 (1H, dd, *J* = 12.9, 5.1 Hz, OCH<u>H</u>CH=CH_2), 5.18 (1H, dd, *J* = 10.6, 1.6 Hz, HC=C<u>H</u>H *cis*), 5.35 (1H, dd, *J* = 17.4, 1.7 Hz, HC=CH<u>H</u> *trans*), 5.99 (1H, m, <u>H</u>C=CH_2). ^{13}C NMR (100 MHz, CD_3OD): δ_C = 61.7 (C6), 70.1 (O<u>C</u>H_2CH=), 70.6 (C5), 74.0 (C2), 76.9, 77.0 (C3, C4), 102.3 (C1), 116.6 (CH=<u>C</u>H_2), 134.7 (<u>C</u>H=CH2). LRMS m/z (ESI^+): 238 $[M+NH_4]^+$, 243 $[M+Na]^+$.

Hexenyl 2,3,4,6-tetra-*O*-acetyl-*β*-(D)-glucopyranoside

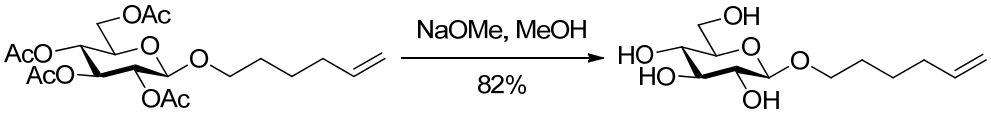


β-D-glucose pentaacetate (5.00 g, 12.8 mmol) was added to a 100 mL flame-dried round bottom flask under argon and dissolved in CH_2Cl_2 (25 mL, anhydrous). The stirred solution was cooled to 0 ºC and $BF_3{\cdot}OEt_2$ (3.2 mL, 25.6 mmol) was added by syringe. After stirring 10 minutes at 0 ºC, 5-hexen-1-ol (3.1 mL, 25.6 mmol) was added. The ice bath was removed after completion of the addition and the reaction stirred at room temperature for 18 hours. The reaction was then cooled to 0 ºC and quenched with

[3.3] Kishida, M.; Akita, H. *Tetrahedron* **2005**, *61*, 10559-10568.

$NaHCO_3$ (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 (50 mL). The combined oganic layers were washed with brine (100 mL), dried over $MgSO_4$, filtered, and concentrated *in vacuo*. The product was purified by column chromatography (20% EtOAc in petrol) to give the titled compound as a thick clear oil which solidified to a waxy solid on storage at –20 °C (1.62 g, 29%). Spectroscopic data was consistent with that previously reported.[3.4] $[\alpha]^{20}_D = -19.7$ (c = 1.0, MeOH). IR (υ_{max}, KBr): 1741, 1641, 1379, 1257, 1037, 904, 744, 697, 677, 576, 561. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 1.33-1.50 (2H, m, C<u>H</u>$_2$CH$_2$CH=CH$_2$), 1.50-1.67 (2H, m, OCH$_2$C<u>H</u>$_2$), 1.99-2.10 (14H, 4 × s and 1 × m, 4 × COC<u>H</u>$_3$ and C<u>H</u>$_2$CH=CH$_2$), 3.48 (1H, dt, *J* = 9.6, 6.7, OC<u>H</u>H′), 3.69 (1H, ddd, *J* = 10.1, 4.6, 2.5, H5), 3.88 (1H, dt, *J* = 9.6, 6.2, OCH<u>H</u>′), 4.13 (1H, dd, *J* = 12.4, 2.0, H6), 4.26 (1H, dd, *J* = 12.4, 5.1, H6′), 4.49 (1H, d, *J* = 7.8, H1), 4.92-5.02 (3H, contain CH=C<u>H</u>$_2$ and H2), 5.08 (1H, t, *J* = 9.6, H4), 5.20 (1H, t, *J* = 9.6, H3), 5.78 (1H, ddt, *J* = 17.0, 10.3, 6.7, 6.7, C<u>H</u>=CH$_2$). ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 20.58, 20.61, 20.64, 20.72 (4 × CO<u>C</u>H$_3$), 25.0 (<u>C</u>H$_2$CH$_2$CH=CH$_2$), 28.8 (OCH$_2$<u>C</u>H$_2$), 33.3 (<u>C</u>H$_2$CH=CH$_2$), 62.0 (C6), 68.4 (C4), 69.9 (OCH$_2$), 71.3 (C2), 71.7 (C5), 72.8 (C3), 100.8 (C1), 114.7 (CH=<u>C</u>H$_2$), 138.5 (<u>C</u>H=CH$_2$), 169.3, 169.4, 170.3, 170.7 (4 × C=O). HRMS m/z (ESI$^+$): Found 453.1729 [M+Na]$^+$; $C_{20}H_{30}O_{10}Na$ requires 453.1731.

Hexenyl-*β*-D-glucoside (3.9)

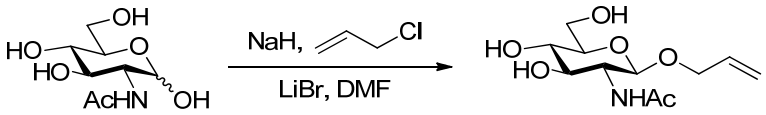


To a stirred solution of hexenyl 2,3,4,6-tetra-*O*-acetyl-*β*-(D)-glucopyranoside (1.11 g, 2.58 mmol) in anhydrous MeOH (10 ml) was added sodium methoxide (70 mg, 1.29 mmol). After stirring for 1 hour, the reaction was neutralized with Dowex 50WX8 (H^+

[3.4] Rodebaugh, R.; Fraser-Reid, B. *Tetrahedron* **1996**, *52*, 7663-7678.

form) and then filtered. The resulting solution was concentrated *in vacuo* to give a thick oil. Purification by column chromatography (20% MeOH in EtOAc) afforded the titled compound as a thick clear oil (545 mg, 82%). $[\alpha]^{20}_{D} = -31.7$ (c = 1.2, MeOH). IR (υ_{max}, film): 3384, 2932, 2502, 1641, 1378, 1164, 1078, 910. ^{1}H NMR (400 MHz, CD_3OD) δ_H = 1.45-1.55 (2H, m, $C\underline{H}_2CH_2CH=CH_2$), 1.59-1.70 (2H, m, $OCH_2C\underline{H}_2$), 2.09 (2H, q, *J* = 6.9, $C\underline{H}_2CH=CH_2$), 3.18 (1H, t, *J* = 8.3, H2), 3.23-3.40 (3H, m, contains H3, H4 and H5), 3.56 (1H, dt, *J* = 9.6, 6.6, OC<u>H</u>H′), 3.68 (1H, dd, *J* = 11.9, 5.1, H6), 3.87 (1H, dd, *J* = 11.9, 1.8, H6′), 3.92 (1H, dt, *J* = 9.9, 6.8, OCH<u>H</u>′), 4.26 (1H, d, *J* = 7.8, H1), 4.94 (1H, dd, *J* = 10.3, 0.9, CH=C<u>H</u>H *cis*), 5.01 (1H, dd, *J* = 17.1, 2.0, CH=CH<u>H</u> *trans*), 5.83 (1H, ddt, *J* = 17.1, 10.3, 6.9, 6.9, C<u>H</u>=CH_2). ^{13}C NMR (100 MHz, CD_3OD) δ_C = 25.5 ($\underline{C}H_2CH_2CH=CH_2$), 29.2 ($OCH_2\underline{C}H_2$), 33.6 ($\underline{C}H_2CH=CH_2$), 61.8 (C6), 69.7 ($OCH_2$), 70.6 (C5), 74.1 (C2), 76.9, 77.1 (C3, C4), 103.3 (C1), 114.0 (CH=$\underline{C}H_2$), 138.9 ($\underline{C}H=CH_2$). HRMS m/z (ESI$^+$): Found 285.1308 $[M+Na]^+$; $C_{12}H_{22}O_6Na$ requires 285.1309.

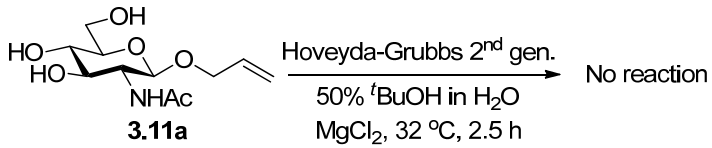
***O*-Allyl-*N*-Acetyl-D-glucosamine (3.11)**



A 25 mL two-necked round-bottomed flask was flamed dried under argon. GlcNAc (500 mg, 2.26 mmol), LiBr (392 mg, 4.52 mmol) and DMF (4 mL, anhydrous) were added sequentially under a stream of argon. The suspension was stirred at RT until homogeneous (~1 h). NaH (108 mg, 2.71 mmol) was added, followed by allyl chloride (0.56 mL, 6.78 mmol) (Note: mixture turned cloudy on addition of NaH). The reaction was monitored by TLC (0.5:3:6.5, H_2O:^{i}PrOH:EtOAc). After 2 h, an additional 3 equivalents of allyl chloride was added (0.56 mL, 6.78 mmol). After 2 h, the solvent was removed *in vacuo*. The resulting yellow residue was purified directly by column

chromatography (0.5:3:6.5, H_2O:^{i}PrOH:EtOAc). A clear oil was obtained that contained residual DMF. A second purification by column chromatography (20% MeOH in EtOAc) provided white needle-like crystals which are very hydroscopic (321 mg, 54%). Spectroscopic data was consistent with that previously reported.[3.5] $[\alpha]^{20}_D$ = −5.0 (c = 1.0, MeOH). IR (υ_{max}, KBr): 3451, 1647, 1556, 1378, 1074, 946, 626. ^{1}H NMR (400 MHz, CD_3OD) δ_H = 2.02 (3H, s, $NHCO_2C\underline{H}_3$), 3.25-3.59 (2 H, m, contain H4 and H5), 3.65-3.96 (4H, m, contain H2, H3 and OCH_2), 4.10 (1H, dd, *J* = 13.4, 5.8, H6), 4.35 (1H, dd, *J* = 13.3, 4.9 Hz, H′6), 4.50 (1H, d, *J* = 8.6, H1), 5.09-5.19 (1H, m, CH=C<u>H</u>H *cis*), 5.28 (1H, dd, J = 17.2, 1.5, CH=CH<u>H</u> *trans*), 5.82-5.99 (1H, m, C<u>H</u>=CH_2). ^{13}C NMR (100 MHz, CD_3OD) δ_C = 22.2 ($NHCO_2\underline{C}H_3$), 56.2, 61.7 (contain C2, C3), 69.9 (C6), 71.0, 75.0, 76.8 (contain C4, C5 and OCH_2), 100.9 (C1), 116.2 (CH=$\underline{C}H_2$), 134.6 ($\underline{C}$H=CH_2), 173.0 (C=O). HRMS m/z (ESI^+): Found 284.1103 $[M+Na]^+$; $C_{11}H_{19}NO_6Na$ requires 284.1105.

Self-metathesis of *O*-Allyl GlcNAc 3.11a



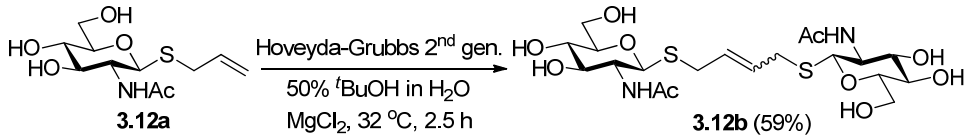
3.11a

O-Allyl GlcNAc **3.11a** (83 mg, 0.31 mmol) and $MgCl_2{\cdot}6H_2O$ (64.2 mg, 0.31 mmol) was added to a 25 mL round bottomed flask and dissolved with ^{t}BuOH/H_2O (1:2, 3 mL). The mixture was stirred and warmed to 32 °C. Hoveyda-Grubbs 2nd generation catalyst (6 mg, 10 μmol) was dissolved in ^{t}BuOH (1 mL) with gentle heating. The catalyst solution was then added to the reaction and the reaction was stirred for 1 hour at 32 °C. A second dose of catalyst (6 mg in 1 mL ^{t}BuOH) followed by H_2O (1 mL) was added. After further 1.5 hours of reaction time, TLC analysis (EtOAc:^{i}PrOH:H_2O, 2:2:1) of the reaction revealed

[3.5] Huang, G.-L.; Zhang, D.-W.; Zhao, H.-J.; Zhang, H.-C.; Wang, P.-G. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2042 - 2043.

only the catalyst (R_f = 0.9) and starting material (R_f = 0.5). The formation of the self-metathesis product was not observed. The solvent was removed under reduced pressure to give a dark residue. Purification by flash column chromatography (20% MeOH in EtOAc then EtOAc:^{i}PrOH:H_2O, 5:4:1) provided the recovered starting material (80 mg, 96% recov.).

Self-metathesis of GlcNAc derived allyl sulfide 3.12a



GlcNAc derived allyl sulfide **3.12a** (121 mg, 0.43 mmol) and $MgCl_2 \cdot 6H_2O$ (88.3 mg, 0.43 mmol) was added to a 25 mL round bottomed flask and dissolved with ^{t}BuOH/H_2O (1:2, 3 mL). The mixture was stirred and warmed to 32 °C. Hoveyda-Grubbs 2nd generation catalyst (8 mg, 13 μmol) was dissolved in ^{t}BuOH (1 mL) with gentle heating. The catalyst solution was then added to the reaction and the reaction was stirred for 1 hour at 32 °C (within the first 30 minutes, the formation of the self-metathesis product was observed by TLC (R_f = 0.2, EtOAc:^{i}PrOH:H_2O, 5:4:1). A second dose of catalyst (8 mg in 1 mL ^{t}BuOH) followed by H_2O (1 mL) was added. After further 1.5 hours of reaction time, the solvent was removed under reduced pressure to give a dark residue. Purification by flash column chromatography (EtOAc:^{i}PrOH:H_2O, 5:4:1) provided the self-metathesis product **3.12b** as a white foam (68 mg, 59%). IR (υ_{max}, KBr): 1640, 1377, 1058, 944, 613. ^{1}H NMR (500 MHz, D_2O): δ_H = 1.98 (3H, s, NHAc), 3.25 (1H, dd, *J* = 14.5, 4.4, SC<u>H</u>H′), 3.37-3.41 (2H, m, contain H4, H5, SCH<u>H</u>′), 3.43 (1H, dd, J = 10.0, 2.2, SCH<u>H</u>′), 3.48 (1H, t, *J* = 9.0, H3), 3.67 (1H, dd, *J* = 12.5, 5.4, H6), 3.71 (1H, t, *J* = 10.0, H2), 3.85 (1H, dd, *J* = 12.5, 1.7, H6′), 4.49 (1H, d, *J* = 10.4, H1), 5.59-5.65 (1H, m, CH=). ^{13}C NMR (125 MHz, D_2O): δ_C = 22.2 (NHAc), 31.5 (SCH_2), 54.6 (C2), 60.9 (C6),

69.8 (C4), 75.1 (C3), 79.8 (C5), 83.1 (C1), 128.9 (CH=), 174.4 (C=O). HRMS m/z (ESI^+): Found 549.1544 $[M+Na]^+$; $C_{20}H_{34}N_2O_{10}S_2Na$ requires 549.1547.

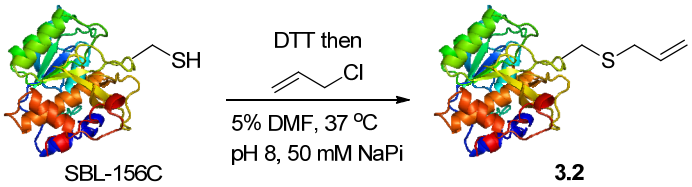
Sequence of subtilisin Bacillus lentus (SBL) mutant S156C

PDB code for wild type = 1GCI

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQD
GNGHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNG
MHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGNCGAGSISYPARYANAMAVGAT
DQNNNRASFSQYGAGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKN
PSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Calculated average isotopic mass = 26714.5

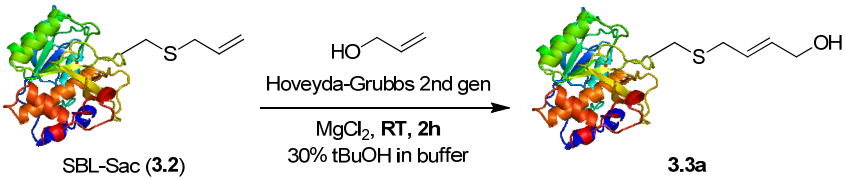
Synthesis of SBL-156Sac (3.2)



SBL-Sac (**3.2**) was synthesized by direct allylation with allyl chloride from SBL-S156C as reported previously in Chapter 2.[3.6]

Cross Metathesis Reactions on SBL-156Sac (3.2)

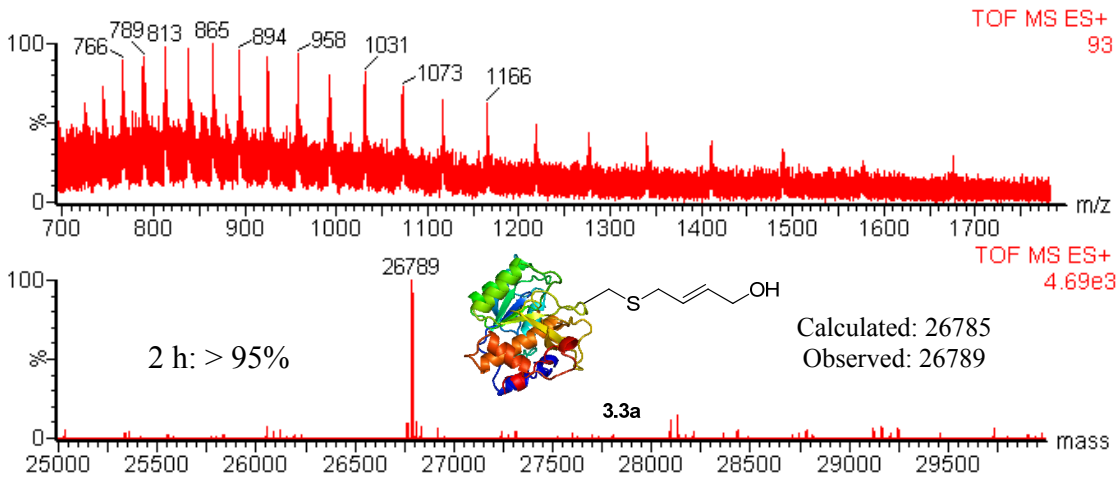
CM between SBL-156Sac (3.2) and Allyl alcohol (3.4a)



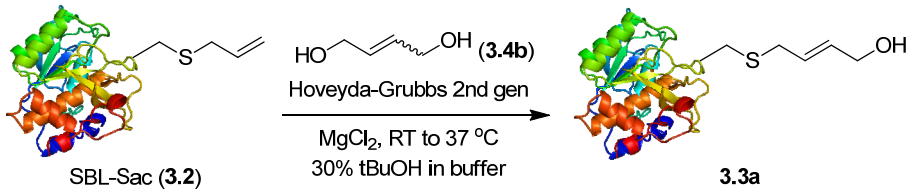
All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 2.6 mg of **3.1** in 337.6 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200

[3.6] Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714-3716.

µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) followed by catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each addition to homogenize. Reaction was placed on shaker for 2 minute before adding allyl alcohol (3.56 µL, 52.3 µmol) to the mixture. The reaction was vortexed and placed on shaker for further 2 hours. Analysis with LC-MS showed >95% conversion to the CM product **3.3a** (26785 calculated mass, found 26789). ESI-MS spectra are shown below.

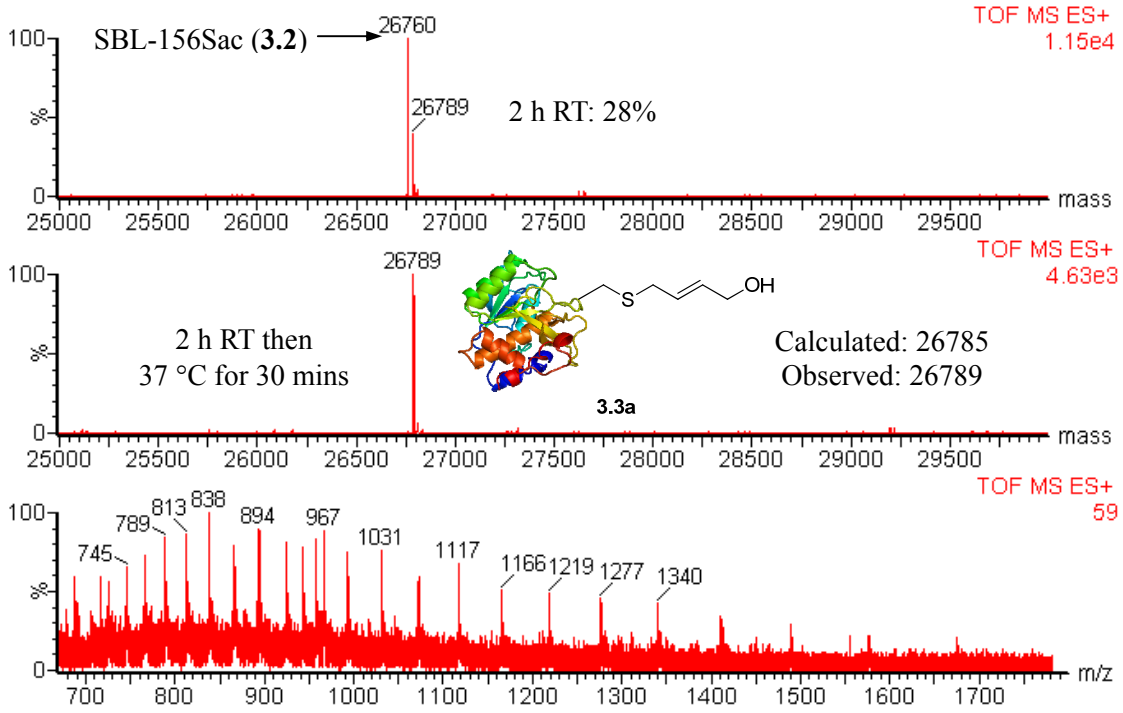


Cross-Metathesis on SBL-156Sac (3.2) with Allyl alcohol homodimer (3.4b)

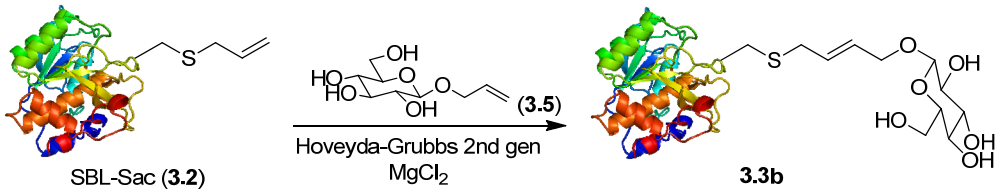


All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 2.1 mg of **3.1** in 272.7 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) followed by catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each addition to homogenize. The reaction was placed on shaker for 2 minute before adding allyl alcohol homodimer (**3.4b**) (4.6 mg, 52.3 µmol) to the mixture. The reaction was

vortexed and placed on shaker for 2 hours. Analysis with LC-MS showed 28% conversion to the CM product **3.3a**. Reaction was then incubated at 37 °C for further 30 minutes, after which time LC-MS showed >95% conversion to product **3.3a** (26785 calculated mass, found 26786). ESI-MS spectra are shown below.

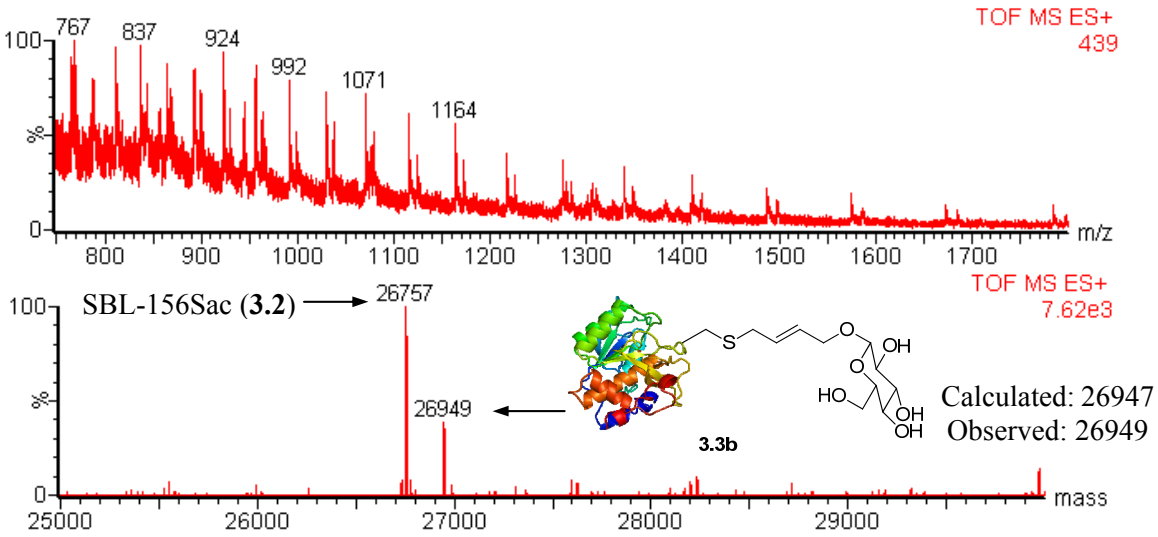


Cross Metathesis on SBL-Sac (3.2) with Allyl glucoside 3.5

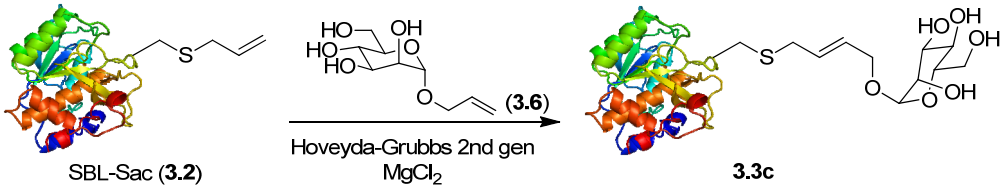


All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 1.2 mg of **3.1** in 156 µL ^{t}BuOH. To an 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) followed by catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each

addition to homogenize. The reaction was placed on shaker for 2 minutes before adding glucoside **3.5** (5.8 mg, 26 µmol) to the mixture. Reaction was vortexed and incubated at 37 °C for 1 hour. Analysis with LC-MS showed 30% conversion to the CM product **3.3b** (26947 calculated mass, found 26949). ESI-MS spectra are shown below.

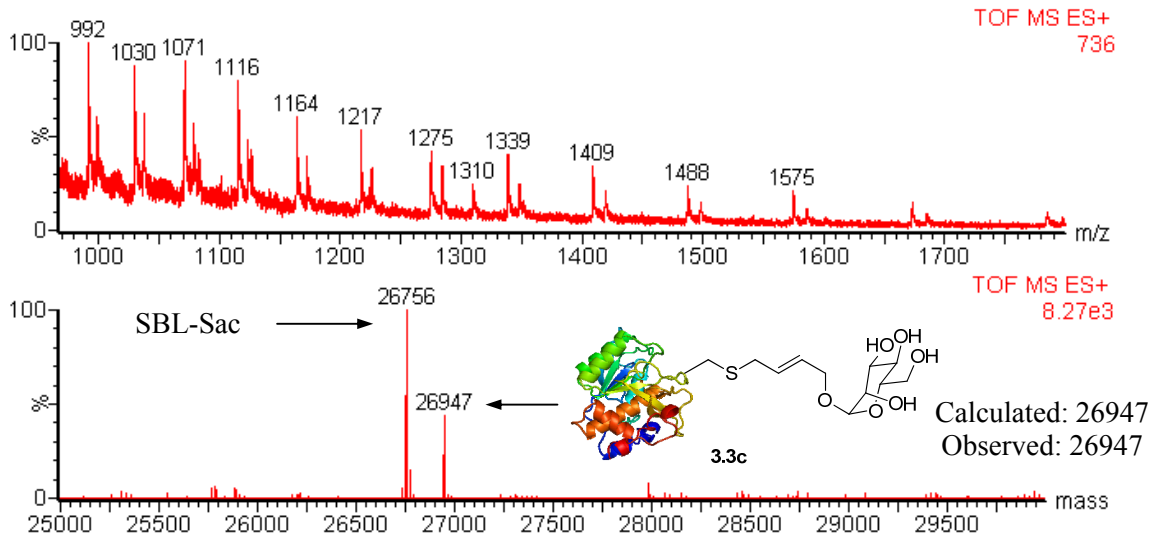


Cross-Metathesis on SBL-Sac (3.2) with Allyl mannoside 3.6

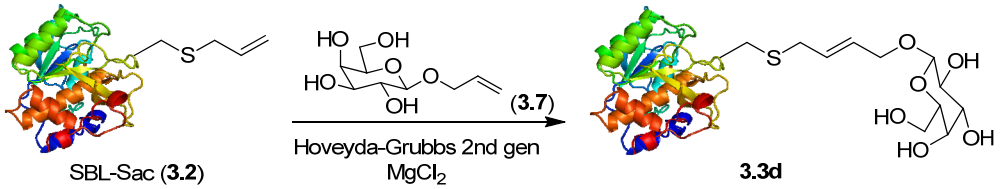


All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 1.9 mg of **3.1** in 247 µL ^{t}BuOH. To an 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) followed by catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each addition to homogenize. The reaction was placed on shaker for 2 minutes before adding mannoside **3.6** (5.8 mg, 26.2 µmol) in 20 µL buffer to the mixture. The reaction was vortexed and incubated at 37 °C for 1 hour. Analysis with LC-MS showed 30%

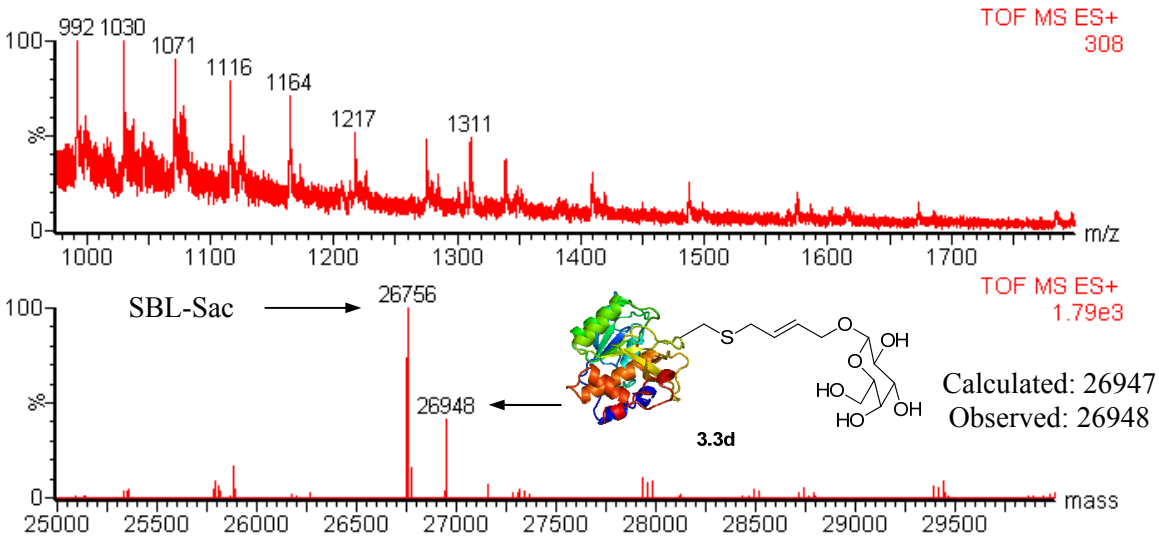
conversion to the CM product **3.3c** (26947 calculated mass, found 26947). ESI-MS spectra are shown below.



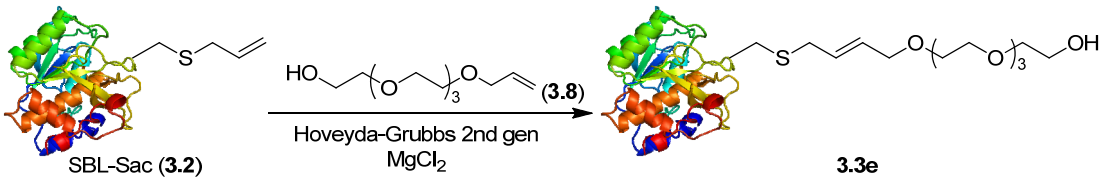
Cross-Metathesis on SBL-Sac (3.2) with Allyl galactoside 3.7



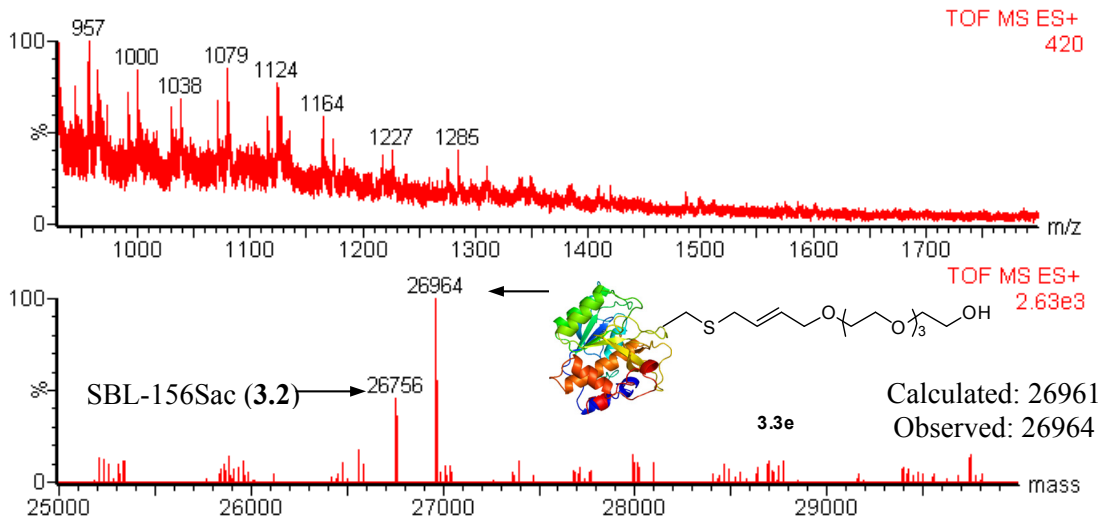
All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 1.9 mg of **3.1** in 247 µL ^{t}BuOH. To an 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) followed by catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each addition to homogenize. The reaction was placed on shaker for 2 minutes before adding mannoside **3.7** (5.8 mg, 26.2 µmol) to the mixture. The reaction was vortexed and incubated at 37 °C for 1 hour. Analysis with LC-MS showed 30% conversion to the CM product **3.3d** (26947 calculated mass, found 26948). ESI-MS spectra are shown below.



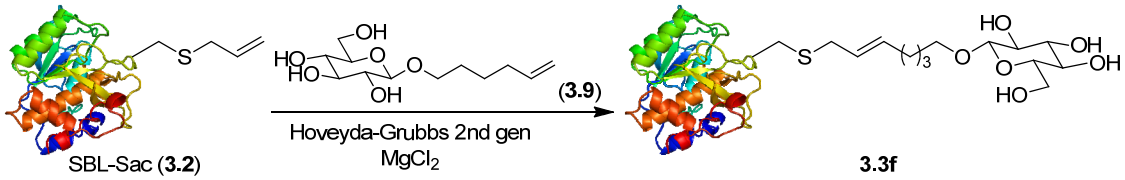
Cross Metathesis on SBL-156Sac (3.2) with Allyl tetraethylene glycol (3.8)



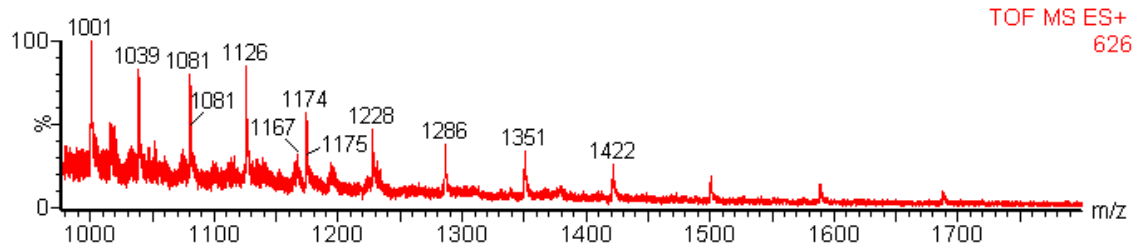
All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 2.4 mg of **3.1** in 312 µL ^{t}BuOH. To an 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) then catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each addition to homogenize. Reaction was placed on shaker for 2 minutes before adding Allyl tetraethylene glycol (**3.8**) (6.1 mg (6.1 µL, neat), 26.2 µmol) to the mixture. The reaction was vortexed and shaken at RT for 2 hour. The reaction was then incubated at 37 °C for further 30 minutes. Analysis with LC-MS showed 65% conversion to the CM product **3.3e** (26961 calculated mass, found 26964). ESI-MS spectra are shown below.

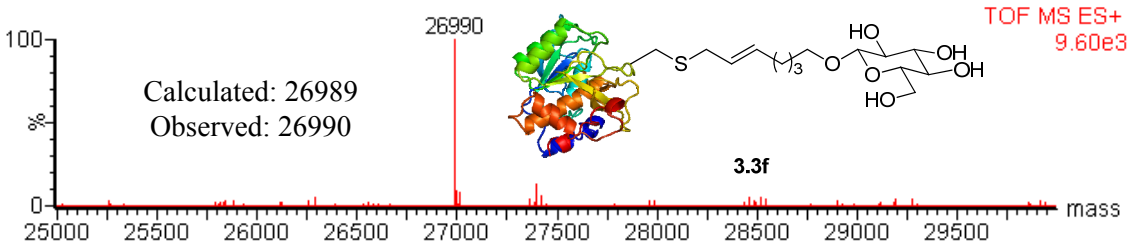


Cross Metathesis on SBL-156Sac (3.2) with Hexenyl glucoside 3.9

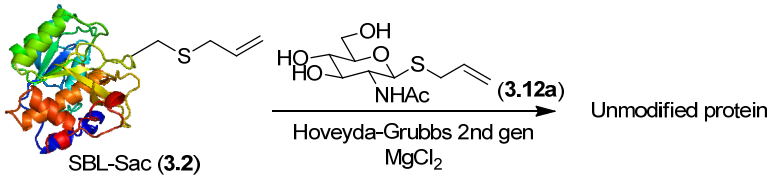


All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 1.9 mg of **3.1** in 246.8 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) then catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each addition to homogenize. The reaction was placed on shaker for 2 minutes before adding hexenyl glucoside **3.9** (6.9 mg, 26.2 µmol) to the mixture. The reaction was vortexed and shaken at RT for 1 hour. Analysis with LC-MS showed >95% conversion to the CM product **3.3f** (26989 calculated mass, found 26990). ESI-MS spectra are shown below.

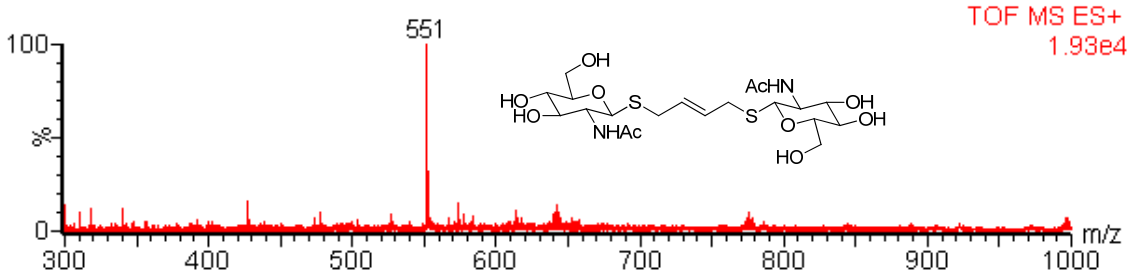


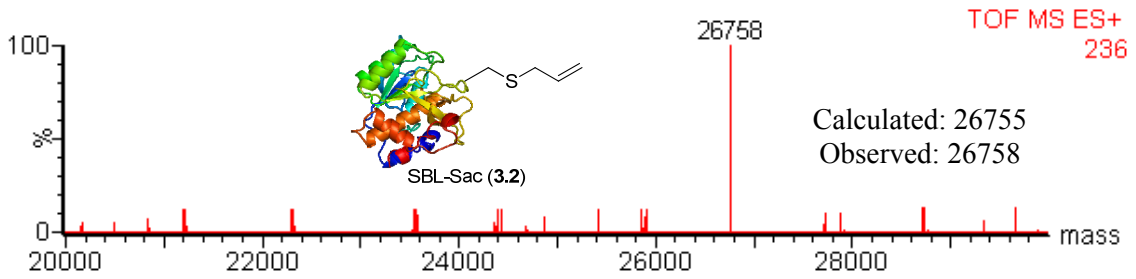


Cross Metathesis on SBL-156Sac (3.2) with GlcNAc derived allyl sulfide 3.12a

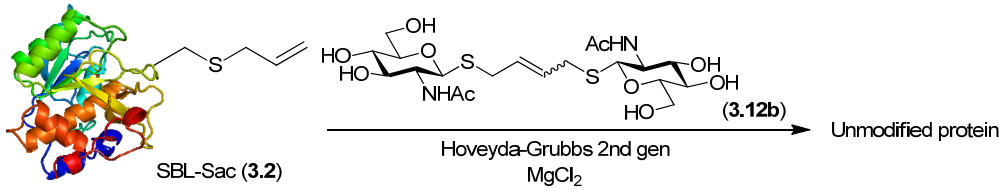


All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 1.9 mg of **3.1** in 246.8 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.7 mg/mL, 5.23 nmol) was added $MgCl_2 \cdot 6H_2O$ (10.6 mg, 52.3 µmol) then catalyst/ ^{t}BuOH solution (85.7 µL, ~1.05 µmol). The reaction was vortexed after each addition to homogenize. The reaction was placed on shaker for 2 minutes before adding GlcNAc allyl sulfide **3.12a** (7.2 mg, 26.2 µmol, in 20 µL buffer) to the mixture. The reaction was vortexed and incubated at 37 °C for 1 hour. Analysis with LC-MS only detected the GlcNAc allyl sulfide homodimer (549 calculated mass (+Na ion), found 551) together with the unmodified protein (26755 calculated mass, found 26758). ESI-MS spectra are shown below.

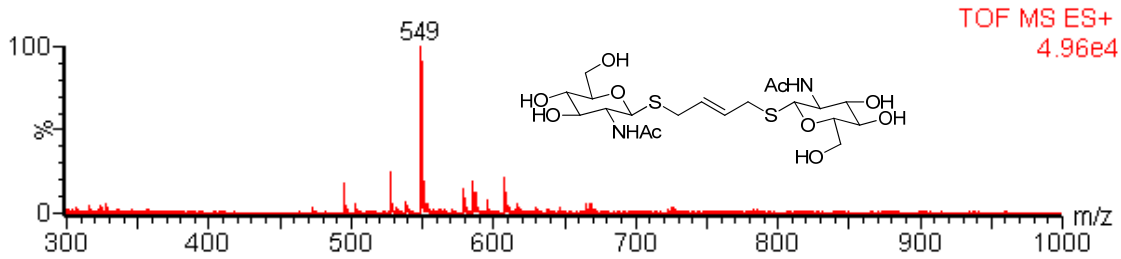


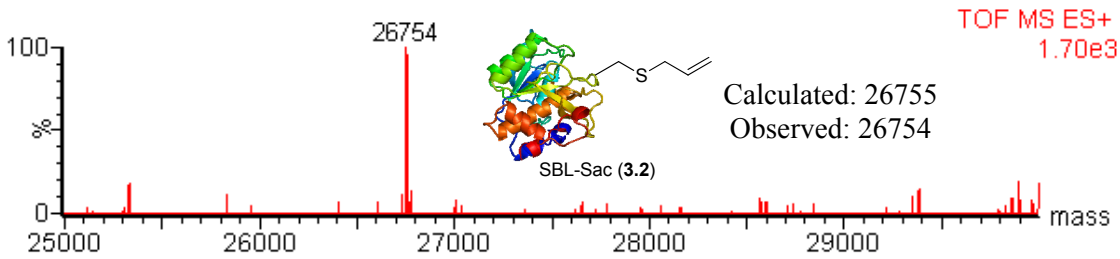


Cross-metathesis between 3.12b and SBL-Sac (3.2)

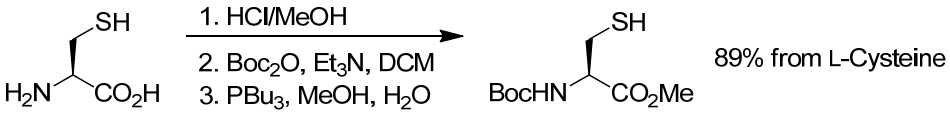


All manipulations were carried out at room temperature. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH was prepared by vortexing and gentle warming 1.0 mg of **3.1** in 153 µL ^{t}BuOH. To a 1.5 mL eppendorf tube containing SBL-156Sac (**3.2**) (200 µL, 0.6 mg/mL, 4.49 nmol) was added $MgCl_2 \cdot 6H_2O$ (8.5 mg, 41.8 µmol) then catalyst/ ^{t}BuOH solution (85.7 µL, ~0.9 µmol). The reaction was vortexed after each addition to homogenize. The reaction was placed on shaker for 5 minutes before adding **3.12b** (11.8 mg, 22.4 µmol, in 20 µL buffer) to the mixture. The reaction was vortexed and incubated at 37 °C for 1 hour. Analysis by LC-MS revealed only **3.12b** (549 calculated mass (+Na ion), found 549) together with the unmodified protein (26755 calculated mass, found 26758). ESI-MS spectra are shown below.





***N*-Boc-L-Cysteine methyl ester (BocCysOMe)**

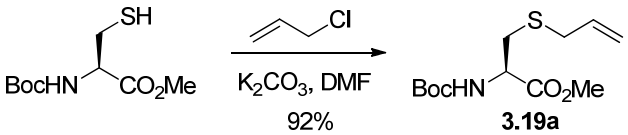


Anhydrous MeOH (100 mL) was added to a flame dried 250 mL round bottom flask equipped with a Teflon coated stir bar. The solvent was stirred and cooled to 0 °C and acetyl chloride (17.6 mL, 248 mmol) was added dropwise over 5 minutes. The solution was stirred an additional 10 minutes at 0 °C to give a concentrated solution of HCl. L-Cysteine (2.00 g, 16.5 mmol) was then added in one portion and the flask flushed 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 CH_2Cl_2 (100 mL) and cooled to 0 °C. Et_3N (5.0 mL, 36.3 mmol) was added carefully followed by di-*tert*-butyl dicarbonate (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). 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 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 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). $[\alpha]^{20}_D$ = +28.3 (c = 7.5, $CHCl_3$) (Lit. [3.7]: $[\alpha]^{20}_D$ = +28.5, c = 7.5, $CHCl_3$). IR (υ_{max}, film): 2979, 1717, 1506, 1367, 1166. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.42 (10H, s, includes Boc and SH), 2.94 (2H, app. td, *J* = 4.3, 8.7, C<u>H</u>$_2$SH), 3.76 (3H, s, $CO_2C\underline{H}_3$), 4.58 (1H, m, H_α), 5.44 (1H, d, *J* = 5.8, NH). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 27.3 (<u>C</u>H_2SH), 28.2 (Boc), 52.6 ($CO_2\underline{C}H_3$), 54.8 (C_α), 80.2 (Boc), 155.1, 170.8 (2 × C=O). LRMS m/z (ESI$^-$): 234 $[M-H]^-$.

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. [3.8] = 96-97 °C). $[\alpha]^{20}_D$ = +91.4 (c = 1.0, $CHCl_3$). IR (υ_{max}, KBr): 3377, 2983, 1748, 1686, 1515, 1367, 1167. ^{1}H NMR (400MHz, $CDCl_3$): δ_H = 1.45 (18H, s, 2 × Boc), 3.16 (4H, d, *J* = 5.1, 2 × C<u>H</u>$_2$S), 3.77 (6H, s, 2 × CO_2CH_3), 4.60 (2H, m, 2 × H_α), 5.40 (2H, d, *J* = 7.6, 2 × NH). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 28.3 (Boc), 41.3 (CH_2S), 52.6 ($CO_2\underline{C}H_3$), 52.8 (C_α), 80.3 (Boc), 155.0, 171.2 (C=O). LRMS m/z (ESI$^+$): 469 $[M+H]^+$, 486 $[M+NH_4]^+$, 491 $[M+NH_4+MeCN]^+$.

_N_-Boc-_S_-allyl-cysteine methyl ester (BocSacOMe, 3.19a)

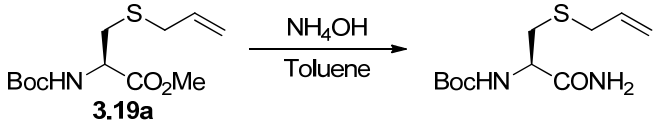


BocCysOMe (16.8 g, 71.2 mmol) was added to a 250 mL round bottom flask and placed under an atmosphere of argon before dissolving in DMF (70 mL). K_2CO_3 (24.6 g, 178

[3.7] Threadgill, M. D.; Gledhill, A. P. *J. Org. Chem.* **1989**, *54*, 2940-2949.
[3.8] Kuligowski, C.; Bezzenine-Lafollé, S.; Chaume, G.; Mahuteau, J.; Barrière, J. -C.; Bacqué, E.; Pancrazi, A.; Ardisson, J. *J. Org. Chem.* **2002**, *67*, 4565-4568.

mmol) was added to the stirred solution and the mixture was cooled to 0 °C. Allyl chloride (27.5 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 **3.19a** (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 18.0 g of **3.19a** as a clear oil (92%) which solidified to waxy prisms upon storage at –20 °C. m.p. = 37-38 °C. $[\alpha]^{20}_D$ = +13.5 (c = 1.0, $CHCl_3$). IR (υ_{max}, film): 3375, 2978, 1747, 1715, 1503, 1367, 1166. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.46 (9H, s, Boc), 2.83-2.96 (2H, ABX System, *J* = 14.0, 4.8, 5.5, C<u>H</u>$_2$SAllyl), 3.09-3.19 (2H, m, SC<u>H</u>$_2$CH=CH$_2$), 3.77 (3H, s, CO$_2$C<u>H</u>$_3$), 4.53 (1H, m, H$_\alpha$), 5.11-5.15 (2H, m, HC=C<u>H</u>$_2$), 5.33 (1H, d, *J* = 6.8, NH), 5.76 (1H, m, <u>H</u>C=CH$_2$). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 28.3 (Boc), 32.8 (<u>C</u>H$_2$SAllyl), 35.1 (S<u>C</u>H$_2$CH=CH$_2$), 52.5 (CO$_2$<u>C</u>H$_3$), 53.1 (C$_\alpha$), 80.0 (Boc), 117.8 (HC=<u>C</u>H$_2$), 133.6 (H<u>C</u>=CH$_2$), 155.1, 171.6 (2 × C=O). HRMS m/z (ESI$^+$): Found 298.1083 [M+Na]$^+$; $C_{12}H_{21}NO_4SNa$ requires 298.1089. Analysis for $C_{12}H_{21}NO_4S$ requires C, 52.34%; H, 7.69%; N, 5.09%; Found: C, 51.94%; H, 7.62%; N, 4.73%.

***N*-Boc-*S*-allyl-L-cysteine amide (BocSacNH$_2$)**

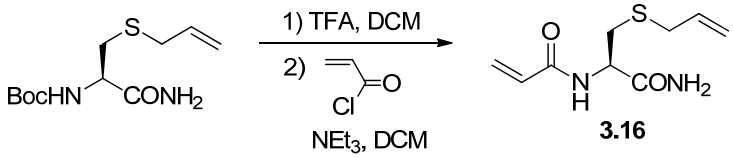


3.19a

BocSacOMe (**3.19a**) (1.00 g, 3.63 mmol) was dissolved in toluene (10 mL) and transferred to an argon-filled 50 mL round bottomed flask. NH_4OH (10 mL, 25%, aq.) was then added. The resulting emulsion was stirred vigorously at room temperature for 18 hours. The reaction was diluted with H_2O (50 mL) and transferred to a separatory

funnel. The aqueous layer was extracted with EtOAc (2 × 50 mL). The combined organic layers were washed with brine (100 mL), dried ($MgSO_4$) and filtered. The solvent was removed under reduced pressure. After purification by flash column chromatography (60% EtOAc in petrol) unreacted starting material (482 mg, 48% recovery) and the titled product as a white solid (429 mg, 45%, 88% brsm) were both isolated. $[\alpha]^{20}_D = -18.7$ (c = 1.0, MeOH). m.p. = 106-107 °C. IR (υ_{max}, KBr): 3389, 3190, 2979, 2935, 1662, 1519, 1425, 1392, 1315, 1271, 1167, 1048, 916, 856, 780, 723, 627. ^{1}H NMR (400 MHz, $CDCl_3$) δ_H = 1.41 (9H, s, Boc), 2.79 (2H, d, *J* = 6.1, H_β), 3.14 (2H, d, *J* = 7.2, $C\underline{H}_2CH{=}CH_2$), 4.28 (1H, m, H_α), 5.09 (1H, d, *J* = 10.0, HC=C<u>H</u>H *cis*), 5.13 (1H, dd, *J* = 17.1, 1.19, HC=CH<u>H</u> *trans*), 5.60 (1H, br. s., NHBoc), 5.73 (1H, m, *J* = 17.1, 10.0, 7.2, <u>H</u>C=CH_2), 6.46 (1H, br. s., CON<u>H</u>H), 6.69 (1H, br. s., CONH<u>H</u>). ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 28.1, 28.3 (Boc, rotamers), 32.9 (C_β), 34.9 ($\underline{C}H_2CH{=}CH_2$), 53.4 ($C_\alpha$), 80.3 (4° C Boc), 117.9 (CH=$\underline{C}H_2$), 133.7 ($\underline{C}H{=}CH_2$), 155.5 (C=O, Boc), 173.7 (C=O, amide). HRMS m/z (ESI$^+$): Found 283.1087 $[M+Na]^+$; $C_{11}H_{20}N_2O_3SNa$ requires 283.1087.

***N*-acryl-*S*-allyl cysteine amide (3.16)**

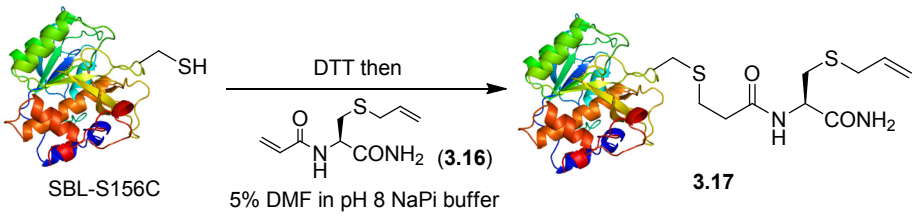


3.16

$BocSacNH_2$ (287 mg, 1.10 mmol) was added to a 50 mL round bottomed flask and dissolved in DCM (15 mL). The solution was stirred and placed under an argon atmosphere. Trifluoroacetic acid (TFA) (1.5 mL, 20.2 mmol) was added as one portion at room temperature. The reaction was stirred for 1.5 hours, at which time TLC (70% EtOAc in petrol) revealed complete consumption of starting material. Solvent and excess TFA was removed under reduced pressure. The resulting residue was dried briefly under high vacuum before re-dissolving in DCM (13 mL). The solution was cooled to 0 °C and NEt_3 (0.62 mL, 4.42 mmol) was added dropwise over 5 minutes. Acryloyl chloride (98.6

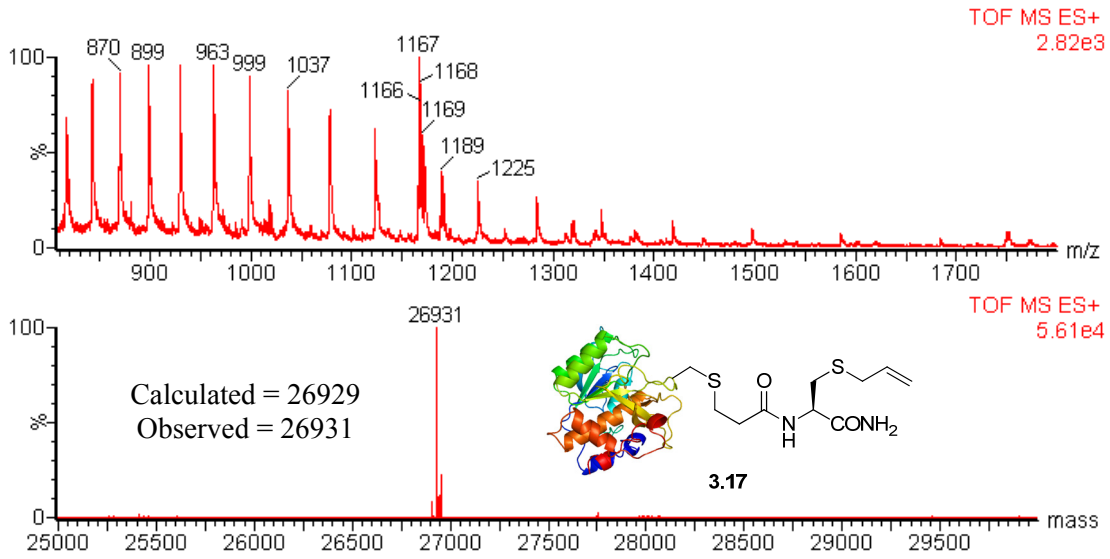
μL, 1.21 mmol) was then added. The ice bath was removed and the reaction was stirred at room temperature for 17 hours. The solvent was removed under reduced pressure. The resulting crude solid was dried briefly under high vacuum before re-dissolving with EtOAc (50 mL) and H_2O (50 mL). Organic layer was isolated and aqueous layer was further extracted with EtOAc (2 × 60 mL). The combined EtOAc layers were washed with $NaHCO_3$ solution (saturated, 75 mL), brine (75 mL), dried with $MgSO_4$ and filtered. Solvent was removed under reduced pressure to reveal product **3.16** as a white crystalline solid. NMR showed that the product is pure and was used without further purification (122 mg, 52%). $[\alpha]^{20}_D = -39.8$ (c = 1.0, MeOH). m.p. = 118-119 °C. IR (υ_{max}, KBr): 3382, 2926, 1654, 1529, 1437, 1311, 1234, 1177, 1130, 1078, 965, 915, 804, 659. ^{1}H NMR (400 MHz, CD_3OD) $\delta_H = 2.75$ (1H, dd, J = 13.8, 8.0, H_β), 2.94 (1H, dd, J = 13.8, 5.9, H_β'), 3.20 (2H, d, J = 7.2, $C\underline{H}_2CH=CH_2$), 4.61 (1H, dd, J = 8.0, 5.9, H_α), 5.12 (1H, d, J = 9.9, $CH_2CH=C\underline{H}H$ cis), 5.18 (1H, dd, J = 17.0, 1.4, $CH_2CH=CH\underline{H}$ trans), 5.73 (1H, dd, J = 9.8, 2.1, $COCH=C\underline{H}H$ cis), 5.80 (1H, m, J = 17.0, 9.9, 7.2, $CH_2C\underline{H}=CH_2$), 6.28 (1H, m, J = 17.2, 2.1, $COCH=CH\underline{H}$ trans), 6.36 (1H, dd, J = 17.2, 9.8, $COC\underline{H}=CH_2$). ^{13}C NMR (100 MHz, CD_3OD) $\delta_C = 32.3$ (C_β), 34.6 ($\underline{C}H_2CH=CH_2$), 52.8 (C_α), 117.0 ($CH_2CH=\underline{C}H_2$), 126.6 ($COCH=\underline{C}H_2$), 130.6 ($CO\underline{C}H=CH_2$), 134.4 ($CH_2\underline{C}H=CH_2$), 166.9, 174.2 (2 × C=O). HRMS m/z (ESI$^+$): Found 237.0672 $[M+Na]^+$; $C_9H_{14}N_2O_2SNa$ requires 237.0668.

SBL-S156C and Acrylamide 3.16 conjugate addition reaction

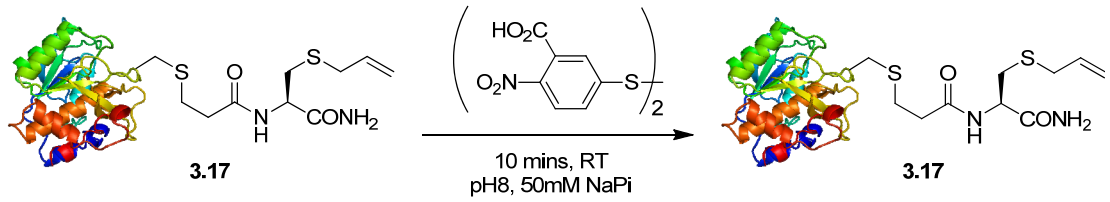


SBL-S156C (2.5 mL, 1.0 mg/mL, 94 nmol, pH 8.0 sodium phosphate) was thawed and put on ice. The protein solution was transferred to a 15 mL falcon tube and DTT (2.2 mg, 14.3 μmol) was added as a solid at RT. The solution was vortexed and placed on a shaker

for 10 minutes. *N*-acryl-*S*-allyl cysteine amide (**3.16**) (16.0 mg, 74.9 µmol) was weighed into a 1.5 mL eppendorf tube and dissolved with DMF (131.6 µL). This DMF solution was then transferred to the protein solution. The reaction was vortexed and shaken at 37 °C for one hour. LC-MS analysis showed full conversion to the Michael addition product **3.17** (calculated = 26929, found = 26931). Small molecules were removed by PD10 (GE Healthcare), eluting with 50 mM sodium phosphate buffer (pH 8.0). The sample was then split into 500 µL aliquots, flash frozen, and stored at −80 ºC. The final concentration of protein **3.17** was 0.71 mg/mL.

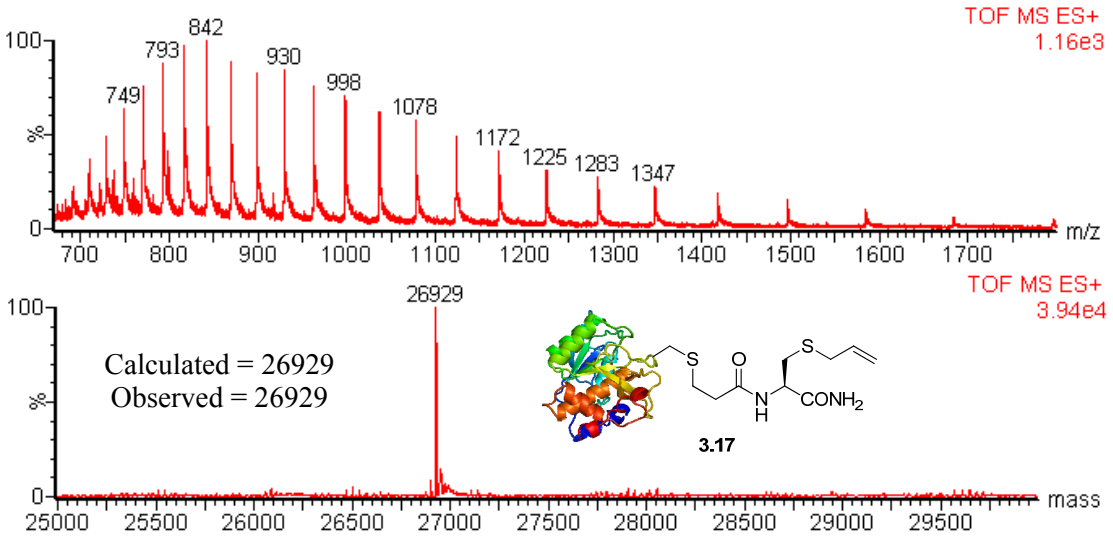


Ellman's Test on Protein 3.17 to Confirm Reaction at Cysteine

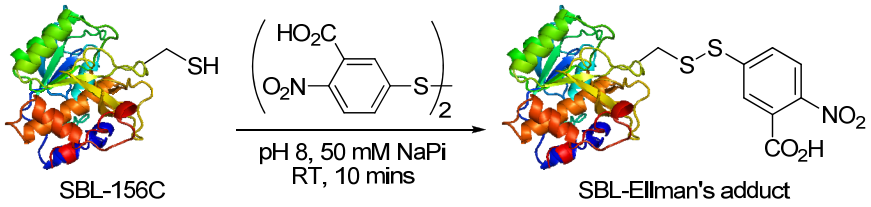


A buffered solution of Ellman's reagent was prepared by dissolving 0.9 mg of Ellman's reagent in 216 µL of 50 mM sodium phosphate buffer (pH 8). A 20 µL of the Ellman's solution (0.2 µmol) was added to protein **3.17** (100 µL, 0.59 mg/mL, 2.2 nmol) in an eppendorf tube. The reaction was vortexed and shaken at room temperature for 10 minutes. The reaction was then analyzed by LC-MS, revealing only the starting material

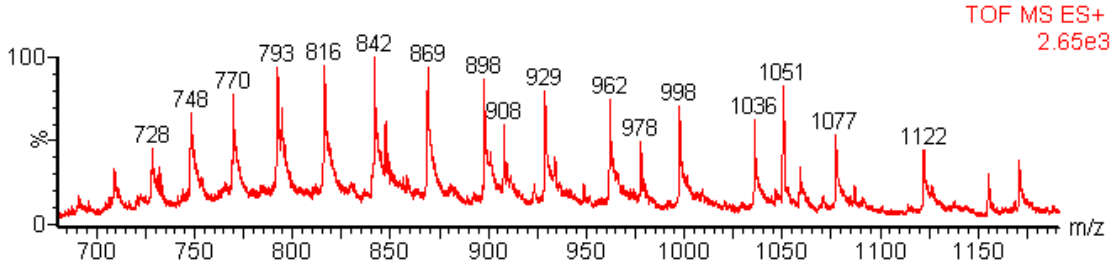
(calculated = 26929, found = 26929). This assay indicates that all free thiol is consumed, thus the alkylation occurred at cysteine.

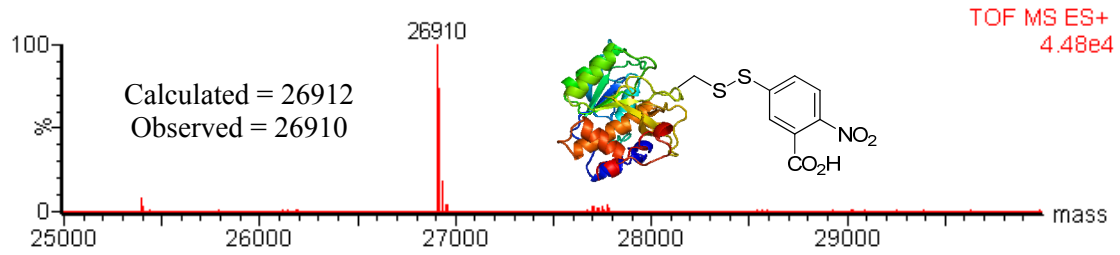


If free cysteine was present the Ellman's adduct would form, as demonstrated below:



An SBL-S156C solution (100 µL, 0.71 mg/mL) was prepared by diluting 71 µL of SBL-S156C (1 mg/mL) to 100 µL with 50 mM sodium phosphate buffer (pH 8.0). A 20 µL 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 = 26912, found = 26910).



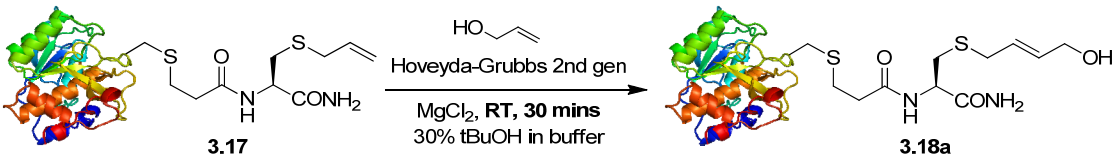


Cross-Metathesis of SBL-extended Sac (3.17)

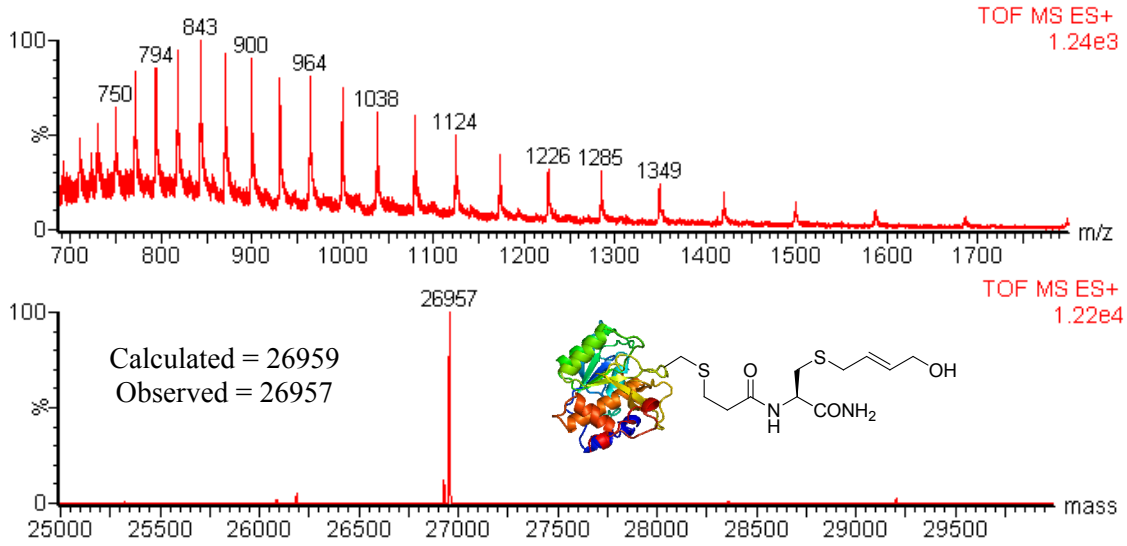
General Procedure for Cross-Metathesis of SBL-extended Sac (3.17)

All manipulations were carried out at room temperature unless stated otherwise. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH (6.4 mg/mL) was prepared by vortexing and gentle warming **3.1** in *t*BuOH. SBL-Sac amide **3.17** (200 µL, 0.59 mg/mL, 4.38 nmol) was transferred to a 1.5 mL eppendorf tube and placed on ice. $MgCl_2.6H_2O$ (8.7 mg, 42.8 µmol) was added as a solid to the protein solution and vortexed to mix. Catalyst/*t*BuOH solution (85.7 µL, ~0.88 µmol) was added to the reaction and vortexed immediately to mix. The reaction mixture appeared as a green emulsion. The reaction was shaken at RT for 1-2 minute before adding the alkene substrate (5000 equiv., 21.9 µmol) as a solid or neat liquid unless stated otherwise. The reaction was mixed on a vortex immediately after addition and was shaken at RT or 37 °C, as indicated. All reactions were monitored by LC-MS.

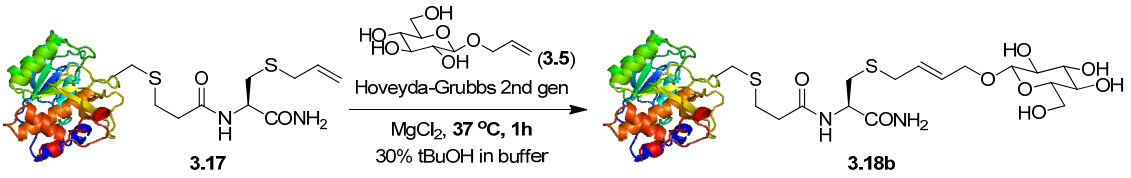
CM between protein 3.17 and allyl alcohol (3.4a)



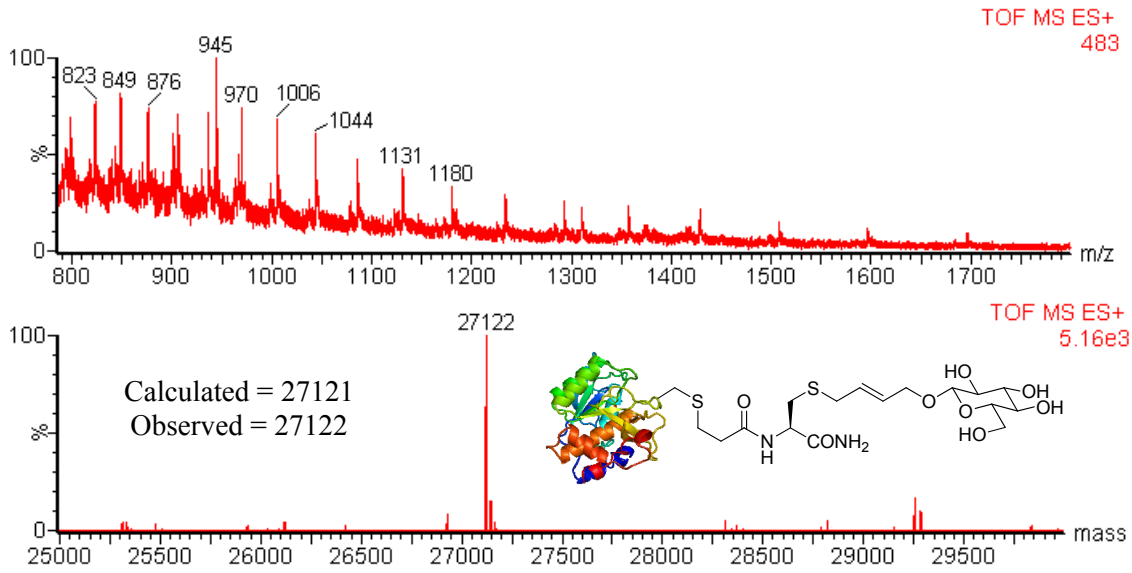
T = RT. After 30 minutes, LC-MS analysis of the product showed full conversion to the CM product **3.18a** (calculated: 26959, found: 26957). ESI-MS of the reaction is shown below:



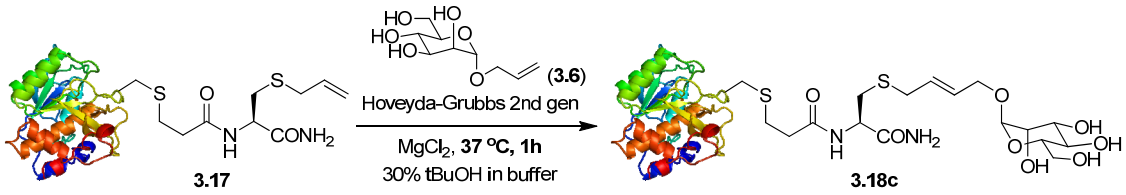
CM between protein 3.17 and Allyl glucoside 3.5



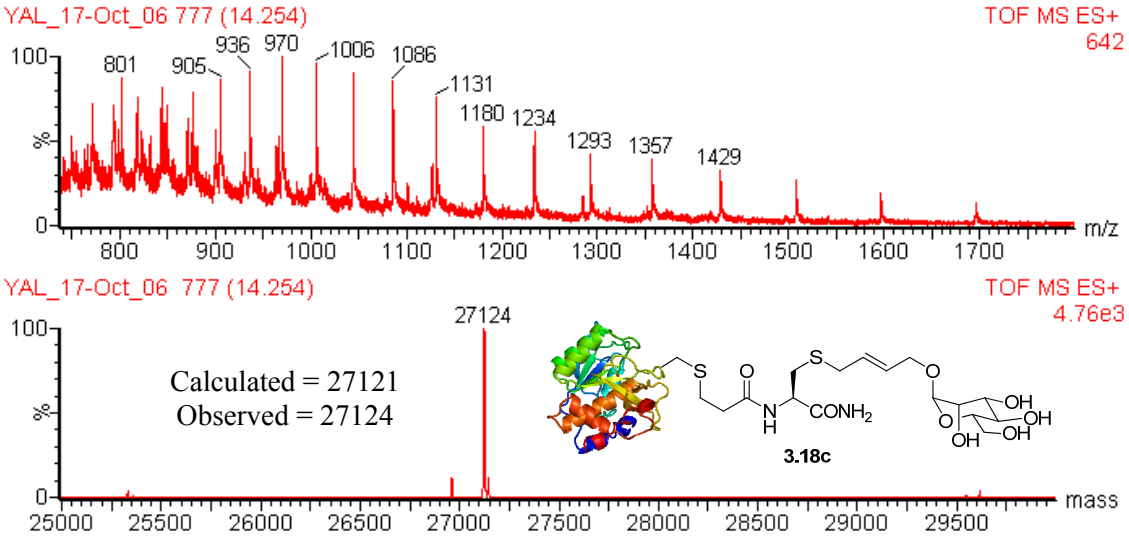
T = 37 °C. After 1 hour LC-MS analysis of the product showed full conversion to the glycosylated protein **3.18b** (calculated: 27121, found: 27122). ESI-MS for the reaction is shown below:



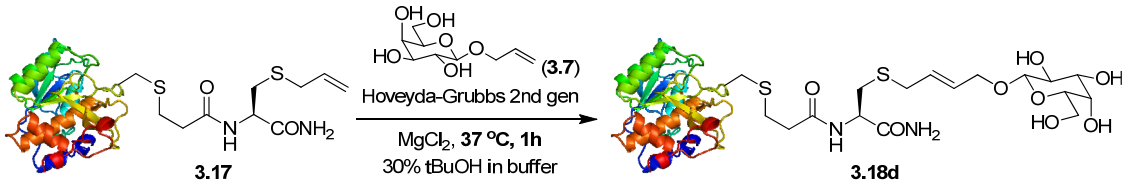
CM between protein 3.17 and Allyl mannoside 3.6



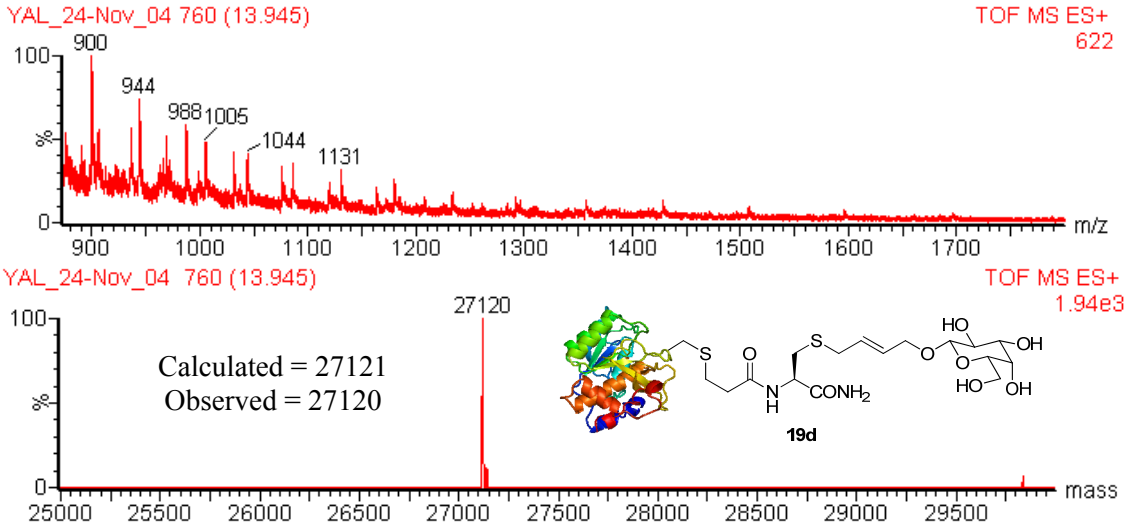
α-O-Allyl-D-mannose (**3.6**) was added as a solution in pH 8 NaPi (50 mM) buffer (20 µL of 1.1 M solution, 4.8 mg, 21.9 µmol). T = 37 °C. After 1 hour LC-MS analysis of the product showed full conversion to the glycosylated protein **3.18c** (calculated: 27121, found: 27124). ESI-MS for the reaction is shown below:



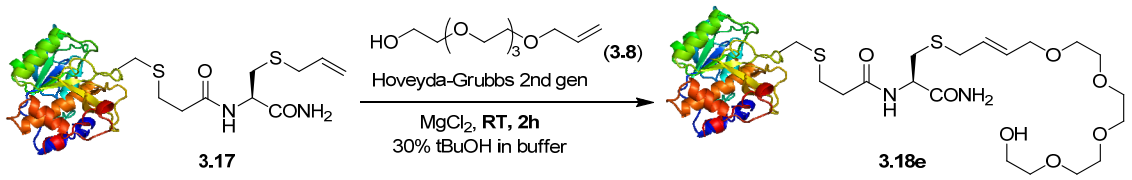
CM between protein 3.17 and Allyl galactoside 3.7



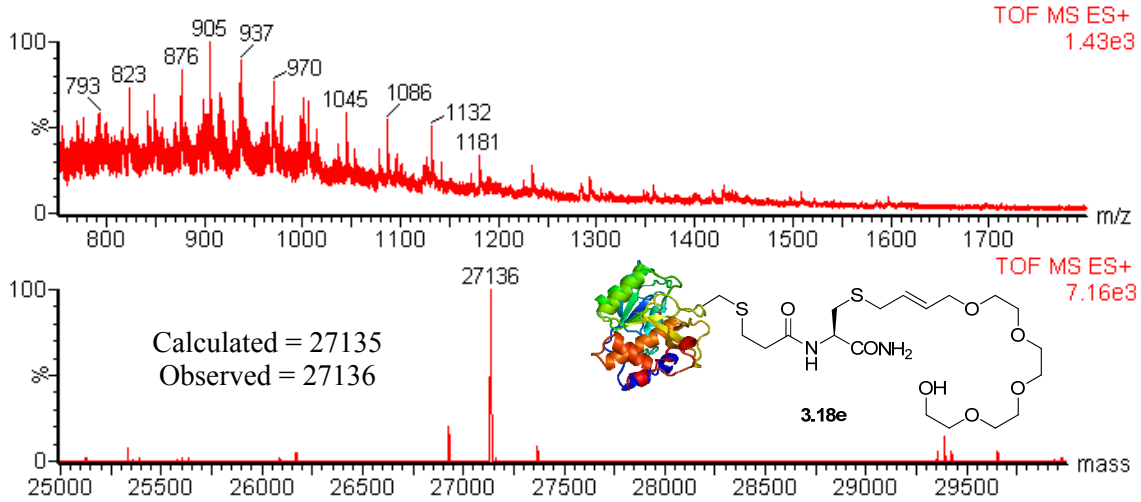
T = 37 °C. After 1 hour LC-MS analysis of the product showed full conversion to the glycosylated protein **3.18d** (calculated: 27121, found: 27120). ESI-MS for the reactions is shown below:



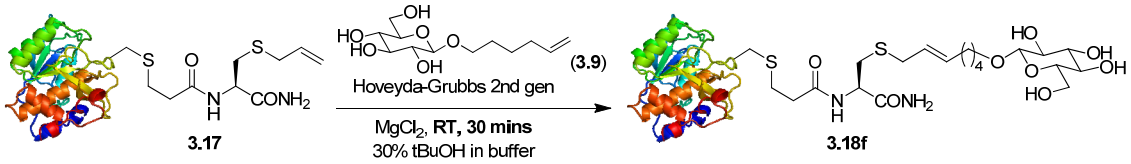
CM between protein 3.17 and Allyl tetraethylene glycol (3.8)



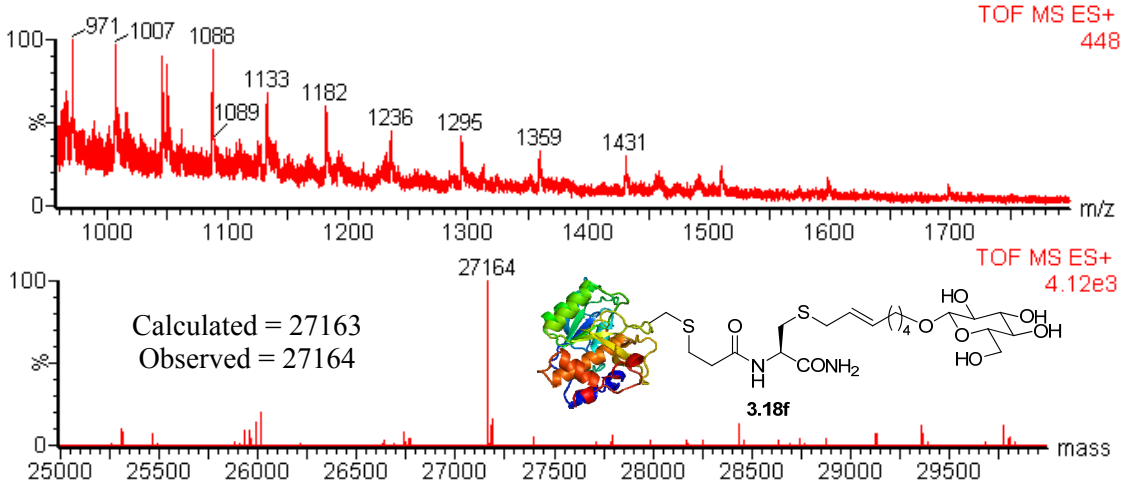
T = RT. After 2 hour LC-MS analysis of the product showed full conversion to the modified protein **3.18e** (calculated: 27135, found: 27136). ESI-MS for the reaction is shown below:



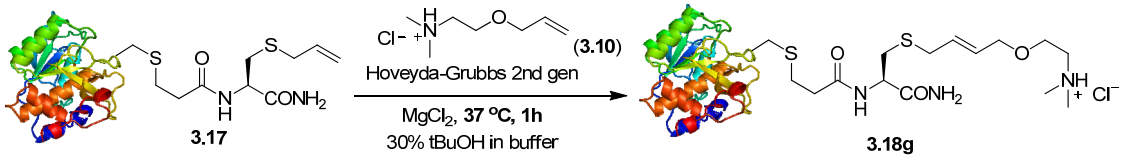
CM between protein 3.17 and hexenyl glucoside 3.9



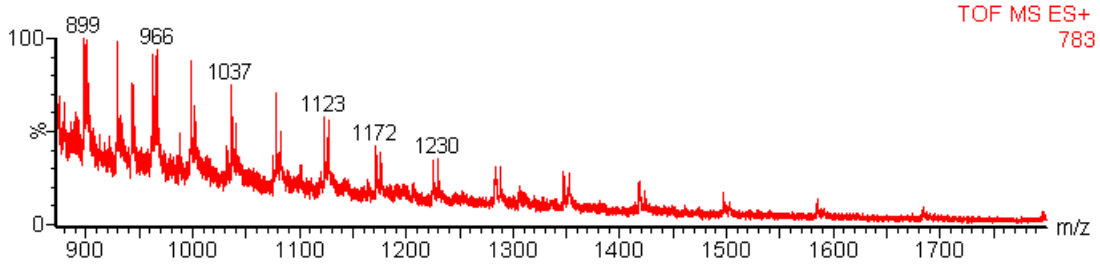
T = RT. After 30 minutes LC-MS analysis of the product showed full conversion to the glycosylated protein **3.18f** (calculated: 27163, found: 27164). ESI-MS for these reactions are shown below:

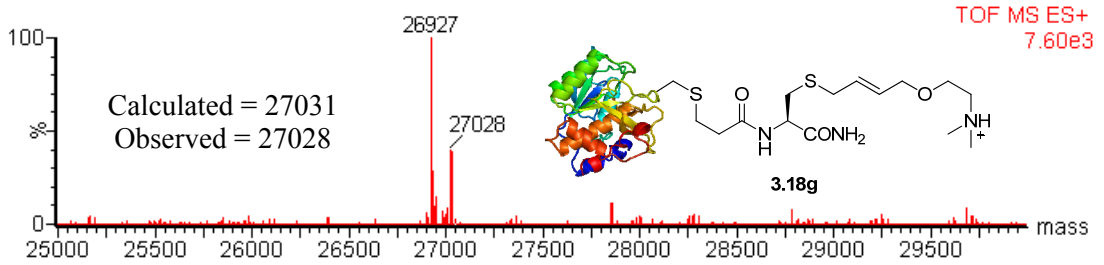


CM of protein 3.17 with *N*,*N*-dimethyl-*O*-Allyl-ethanolamine hydrochloride (3.10)

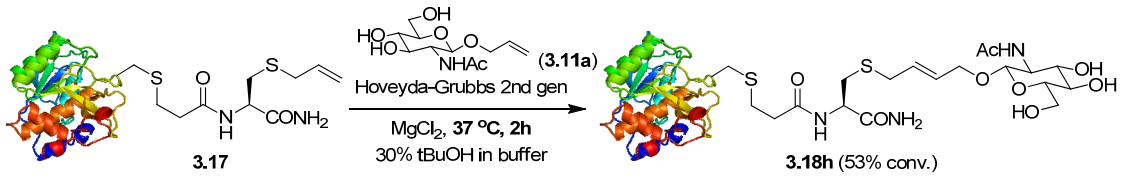


T = 37 °C. After 30 minutes at 37 °C, LC-MS analysis of the product showed 29% conversion to the modified protein **3.18g** (without Cl^-) (calculated: 27031, found: 27028). ESI-MS for the reaction is shown below:

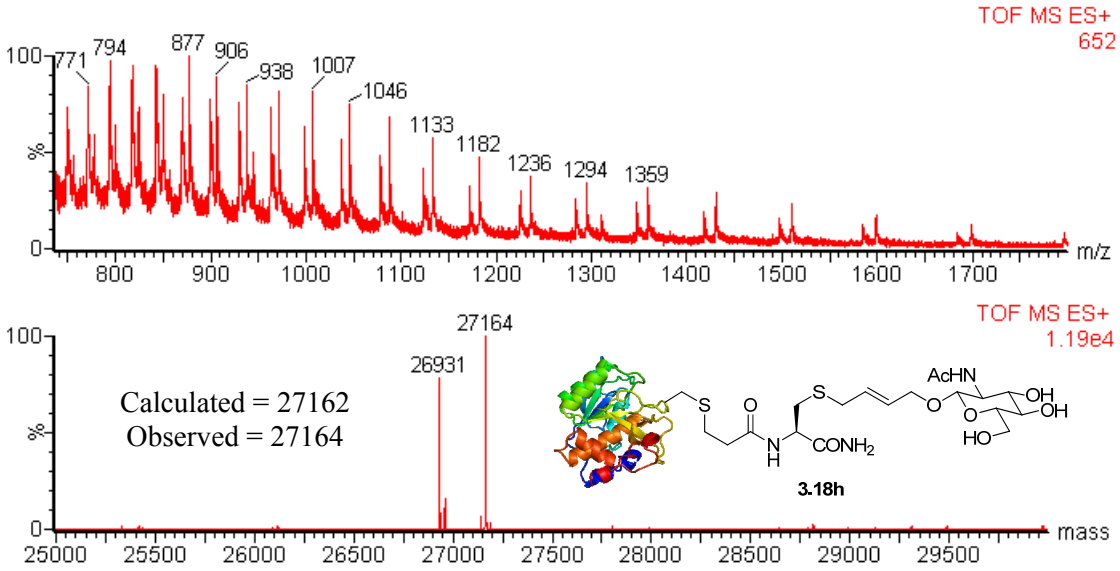




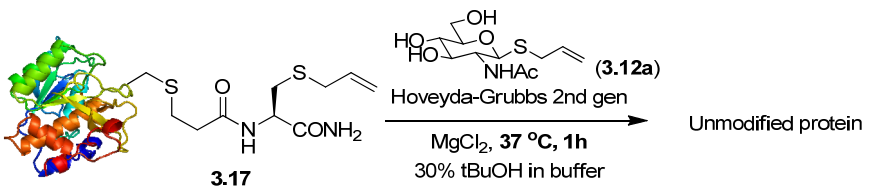
CM between protein 3.17 and *β-O*-allyl-D-GlcNAc (3.11a)



β-O-allyl-D-GlcNAc (**3.11a**) (5.7 mg, 21.9 µmol) was added as a solution in 20 µL of pH 8 NaP_i (50 mM) buffer. T = 37 °C. After 2 hour LC-MS analysis of the product showed 53% conversion to the glycosylated protein **3.18h** (calculated: 27162, found: 27164). ESI-MS for the reaction is shown below:

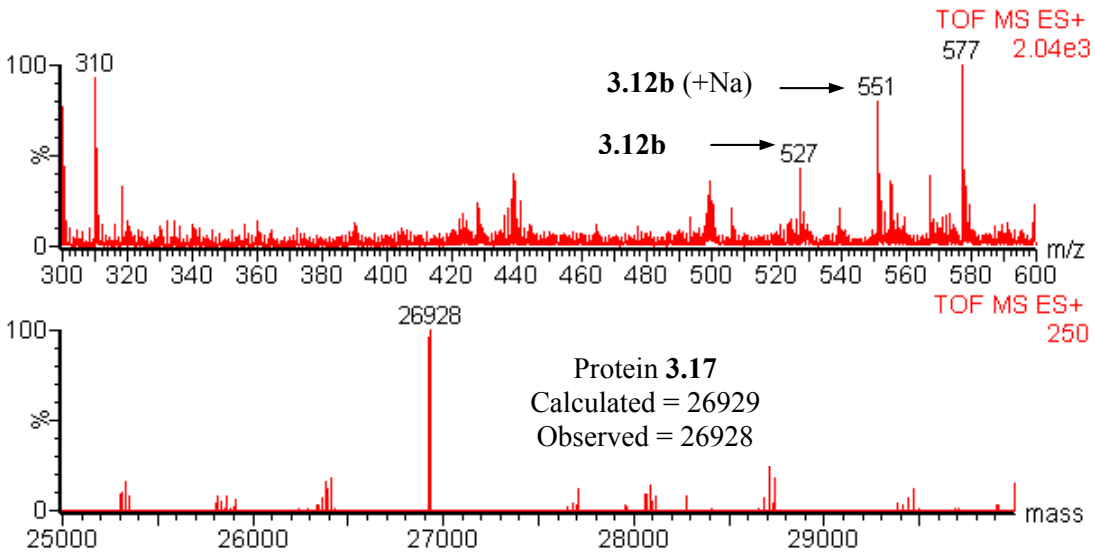


CM between protein 3.17 and GlcNAc allyl sulfide (3.12a)

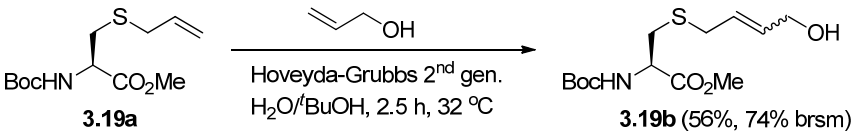


GlcNAc allyl sulfide **3.12a** (6.1 mg, 21.9 µmol) was added as a solution in 20 µL of pH 8

NaP_i (50 mM) buffer. T = 37 °C. After 1 hour LC-MS analysis of the product showed only self-metathesis product of **3.12a** (**3.12b**) and unmodified protein **3.17** (calculated = 26929, found = 26928). ESI-MS for the reaction is shown below:



CM of BocSacOMe (3.19a) and Allyl alcohol (3.4a)

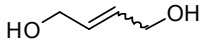


3.19a **3.19b** (56%, 74% brsm)

A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) (12 mg, 0.02 mmol) in ^{t}BuOH (2.0 ml) was added to a stirred mixture containing BocSacOMe (**3.19a**) (126 mg, 0.46 mmol) and allyl alcohol (0.16 ml, 2.29 mmol) in ^{t}BuOH:H_2O (1:2, 6 ml) at 32 ºC. A second dose of the catalyst solution (12 mg of **3.1** in 2.0 ml ^{t}BuOH) and allyl alcohol (**3.4a**) (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 to yield a dark brown residue. Purification by column chromatography (50% EtOAc in petrol) afforded the starting material (30 mg, 24%) and cross-metathesis product **3.19b** (78 mg, 56%, 74% brsm). IR (υ_{max}, film): 3376, 2978, 1713, 1513, 1367, 1166. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.43 (9H, s, Boc), 2.49 (1H, br. s, OH), 2.76-2.94 (2 H, ABX System, H_β), 3.14 (2H, d, *J* = 7.1, SC$\underline{H_2}$CH=), 3.75 (3H, s,

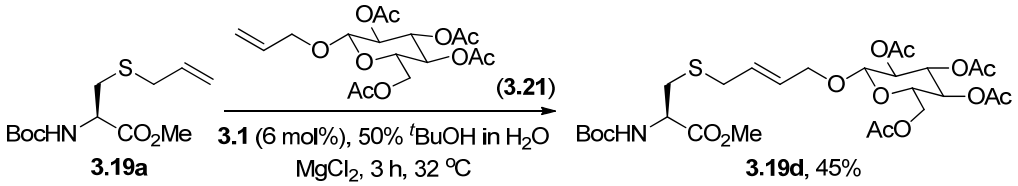
$CO_2C\underline{H}_3$), 4.11 (2H, d, J = 3.8, $C\underline{H}_2OH$), 4.41-4.56 (1H, m, H_α), 5.35 (1H, d, J = 7.8, NH), 5.58-5.70 (1H, m, $SCH_2C\underline{H}$=), 5.70-5.83 (1H, m, =$C\underline{H}CH_2OH$). ^{13}C NMR (100 MHz, $CDCl_3$): δ_c = 28.3 (Boc), 33.0 (C_β), 34.0 ($S\underline{C}H_2CH$=), 52.6 ($CO_2\underline{C}H_3$), 53.3 (C_α), 62.8 ($\underline{C}H_2OH$), 80.4 (Boc), 127.3 (=$\underline{C}HCH_2OH$), 133.0 ($SCH_2\underline{C}H$=), 155.2 ($\underline{C}O_2CH_3$), 171.7 ($NH\underline{C}OCH_3$). HRMS m/z (ESI^+): found 328.1189 $[M+Na]^+$; $C_{11}H_{23}NO_5Na$ requires 328.1195.

Note: Allyl alcohol self-metathesis product (**3.4b**) is also isolated (206 mg, 99%).



2-Buten-1,4-diol (3.4b): (95% *E*-isomer.) IR (υ_{max}, film): 3331, 1421, 1085, 992. ^{1}H NMR (400 MHz, DMSO-d_6): δ_H = 3.92 (4H, d, J = 2.6, 2 × $C\underline{H}_2OH$), 4.68 (2H, br. s, 2 × $CH_2O\underline{H}$), 5.68 (2H, m, $C\underline{H}$=$C\underline{H}$). ^{13}C NMR (100 MHz, $CHCl_3$): δ_C = 62.8 ($\underline{C}H_2OH$), 130.5 ($\underline{C}H$=CH), identical to reported ^{13}C NMR for **(*E*)-2-Buten-1,4-diol**.[3.9]

CM between 3.19a and Allyl glucoside tetraacetate 3.21

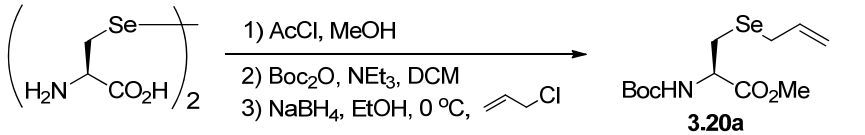


BocSacOMe (**3.19a**) (90.5 mg, 0.33 mmol) and $MgCl_2.6H_2O$ (100 mg, 0.49 mmol) were added to a 25 mL round bottom flask was dissolved with ^{t}BuOH/H_2O (1:2, 4 mL). The mixture was stirred and warmed to 32 °C. Allyl glucoside tetraacetate **3.21** (255.3 mg, 0.66 mmol) was added as a solid to the stirred mixture. Hoveyda-Grubbs 2nd generation catalyst (6.2 mg, 9.89 μmol) was dissolved in ^{t}BuOH (1 mL) with gentle heating. The catalyst solution was then added to the reaction and the reaction was stirred for 1 hour at 32 °C. Second dose of catalyst (6.2 mg in 1 mL ^{t}BuOH) followed by H_2O (1 mL) was added. After further 2 hours of reaction time, TLC (30% EtOAc in petrol) revealed the formation of the desired CM product (R_f = 0.08). Solvent was removed under reduced

[3.9] Crimmins, M. T.; DeBaillie, A. C. *J. Am. Chem. Soc.* **2006**, *128*, 4936-4937.

pressure to give a dark residue. Purification by flash column chromatography (20% EtOAc in petrol then 40% EtOAc in petrol) provided the starting material **3.19a** (19.6 mg, 22% recovery) and the CM product **3.19d** (94.2 mg, 45%). IR (υ_{max}, film): 3377, 2977, 1748, 1506, 1437, 1368, 1223, 1042, 907, 779, 735, 701, 600. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.41 (9H, s, Boc), 1.96, 1.98, 2.02, 2.05 (12H, 4 × s, 4 × OAc), 2.80 (1H, dd, *J* = 13.9, 5.5, H_β), 2.89 (1H, dd, *J* = 13.9, 5.3, H_β'), 3.11 (2H, dd, *J* = 6.1, 3.6, SC<u>H</u>$_2$), 3.67 (1H, ddd, *J* = 10.1, 4.7, 2.4, H5), 3.73 (3H, s, $CO_2C\underline{H}_3$), 4.07 (1H, dd, *J* = 12.7, 5.0, OC<u>H</u>H'), 4.11 (1H, dd, *J* = 12.5, 2.4, H6), 4.23 (1H, dd, *J* = 12.5, 4.7, H6'), 4.30 (1H, dd, *J* = 12.7, 4.4, OCH<u>H</u>'), 4.43-4.50 (1H, m, H_α), 4.53 (1H, d, *J* = 8.0, H1), 4.97 (1H, dd, *J* = 9.5, 8.0, H2), 5.06 (1H, t, *J* = 9.6, H4), 5.18 (1H, t, *J* = 9.4, H3), 5.33 (1H, d, *J* = 7.9, NH), 5.57 (1H, dt, *J* = 15.4, 4.8, $OCH_2C\underline{H}$=), 5.63 (1H, dt, *J* = 15.4, 6.1, $SCH_2C\underline{H}$=). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 20.56, 20.57, 20.66, 20.72 (4 × OAc), 28.3 (Boc), 33.1 (C_β), 33.8 (S<u>C</u>H$_2$), 52.5 ($CO_2\underline{C}H_3$), 53.2 (C_α), 61.9 (C6), 68.3 (C4), 68.8 (O<u>C</u>H$_2$), 71.2 (C2), 71.7 (C5), 72.8 (C3), 80.1 (4° C, Boc), 99.4 (C1), 128.5, 129.3 (H<u>C</u>=<u>C</u>H), 155.0 (C=O Boc), 169.3, 169.4, 170.2, 170.6 (4 × C=O, Ac), 171.5 (C=O, ester). LRMS m/z (ESI^+): Found 636.26 $[M+H]^+$, 658.22 $[M+Na]^+$. HRMS m/z (ESI^+): Found 658.2133 $[M+Na]^+$; $C_{27}H_{41}NO_{14}SNa$ requires 658.2140.

_N_-Boc-_Se_-allyl-L-selenocysteine methyl ester (BocSeacOMe, 3.20a)

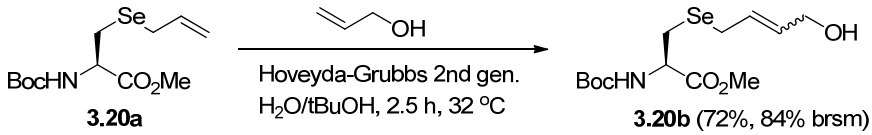


Anhydrous MeOH (30 mL) was added to a flame dried 100 mL round bottom flask equipped with a Teflon coated stir bar. The solvent was stirred and cooled to 0 ºC and acetyl chloride (44.9 mmol, 3.2 mL) was added dropwise over 5 minutes. The solution was stirred an additional 10 minutes at 0 ºC to give a concentrated solution of HCl. Seleno-L-cystine (1.00 g, 2.99 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 selenocystine methyl ester hydrochloride as a yellow solid. This material was used immediately in the next step without purification. The crude ester was suspended in DCM (30 mL) and cooled to 0 °C. Et_3N (1.7 mL, 12.0 mmol) was added carefully followed by di-tert-butyl dicarbonate (Boc_2O, 1.63 g, 7.48 mmol). The reaction was stirred at room temperature for 3.5 hours after which time TLC (30% EtOAc in Petrol) revealed the Boc protected diselenide ($R_f = 0.2$). The solvent was removed under reduced pressure and the resulting residue was redissolved in with Et_2O (50 mL) and H_2O (50 mL). The organic layer was separated and the aqueous layer was extracted with Et_2O (2 × 25 mL). The combined organics were washed with brine (50 mL), dried over $MgSO_4$, and filtered. The solvent was removed under reduced pressure to obtain the protected selenocystine, which was used in the next step without further purification. The yellow solid was dissolved in EtOH (40 mL) and cooled to 0 ºC. $NaBH_4$ (225.4 mg, 5.96 mmol) was added carefully. After stirring for 15 minutes at 0 °C, allyl chloride (1.2 mL, 15.0 mmol) was added all at once. After stirring for an additional 30 minutes, 1M HCl (8 mL) followed by H_2O (80 mL) was added to quench excess $NaBH_4$. The mixture was extracted with Et_2O (2 × 75 mL). The combined organic layers were washed with brine (100 mL), dried with $MgSO_4$, and filtered. The solvent was evaporated in vacuo to give a yellow oil. Purification by column chromatography (30% EtOAc in petrol) provided the titled compound as a pale yellow oil which solidified on storing at −20 °C (1.68 g, 87% from Seleno-L-cystine). **Data for BocSeacOMe (3.20a):** $[\alpha]^{20}_D = -37.1$ (c = 1.0, MeOH). m.p. = 37-38 °C. IR (υ_{max}, KBr): 3372, 2980, 1678, 1519, 1437, 1411, 1166, 1051, 987, 946, 912, 871, 851, 778. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.45 (9H, s, Boc), 2.89 (1H, dd, *J* = 13.0, 5.3, H_β), 2.94 (1H, dd, *J* = 13.0, 5.1, H_β'), 3.18 (2H, dd, *J* = 7.6, 4.0,

C<u>H</u>$_2$CH=), 3.76 (3H, s, CO_2CH_3), 4.54-4.66 (1H, m, H_α), 5.04 (1H, dd, *J* = 17.0, 1.3, CH=C<u>H</u>H *trans*), 5.07 (1H, d, *J* = 9.9 Hz, CH=CH<u>H</u> *cis*), 5.34 (1H, d, *J* = 7.3, NH), 5.84 (1H, m, *J* = 17.0, 9.9, 7.7, C<u>H</u>=CH_2). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 25.2 (C_β), 26.7 (<u>C</u>H_2CH=), 28.3 (Boc), 52.5 (CO_2CH_3), 53.4 (C_α), 80.1 (Boc), 117.1 (CH=<u>C</u>H_2), 134.4 (<u>C</u>H=CH_2), 155.1, 171.7 (2 × C=O). HRMS m/z (ESI$^+$): Found 346.0527 $[M+Na]^+$; $C_{12}H_{21}NO_4SeNa$ requires 346.0528. **Data for *N*-Boc-L-selenocystine methyl ester:** IR (υ_{max}, film): 3367, 2978, 2254, 1713, 1506, 1437, 1393, 1367, 1166, 1050, 1023, 914, 861, 777. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.46 (18H, s, 2 × Boc), 3.29-3.48 (4H, m, H_β), 3.77 (6H, s, 2 × CO_2C<u>H</u>$_3$), 4.54-4.71 (2H, m, 2 × H_α), 5.40 (2H, d, *J* = 7.33, 2 × NH). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 28.2 (Boc), 32.2 (C_β), 52.5 (CO_2<u>C</u>H_3), 53.6 (C_α), 80.1 (Boc), 155.0, 171.3 (2 × C=O). HRMS m/z (ESI$^+$): Found 587.0381 $[M+Na]^+$; $C_{18}H_{32}N_2O_8Se_2Na$ requires 587.0385.

CM between 3.20a and Allyl Alcohol (3.4a)



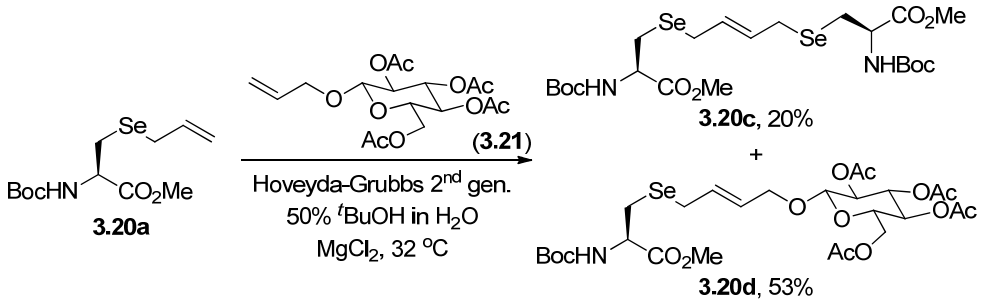
3.20a **3.20b** (72%, 84% brsm)

A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1)** (3.6 mg, 0.0058 mmol) in ^{t}BuOH (0.5 ml) was added to a stirred mixture containing BocSeacOMe (**3.20a**) (62.0 mg, 0.19 mmol) and allyl alcohol (0.07 mL, 0.96 mmol) in ^{t}BuOH:H_2O (1:1.5, 2.5 mL) at 32 ºC. After 1 hour, a second dose of the catalyst solution (3.6 mg catalyst in 0.5 mL ^{t}BuOH) and allyl alcohol (0.07 mL, 0.96 mmol) followed by 0.5 mL of water was added. The reaction was stirred for an additional 1.5 hours before removing the solvents in vacuo to give a dark residue. Purification by column chromatography (50% EtOAc in petrol then 15% MeOH in EtOAc) afforded the starting material (8.4 mg, 13.5%) and cross-metathesis product **3.20b** (49.0 mg, 72%, 84% brsm). IR (υ_{max}, film): 3379, 2978, 2250, 1699, 1506, 1438, 1166, 1051, 1008, 915, 859, 733. ^{1}H NMR (400 MHz, $CDCl_3$)

δ_H = 1.44 (9H, s, Boc), 2.36 (1H, br. s., OH), 2.84 (1H, dd, *J* = 12.9, 5.6, H_β), 2.91 (1H, dd, *J* = 12.9, 6.1, H_β'), 3.20 (2H, d, *J* = 7.3, SeC<u>H</u>$_2$CH=), 3.75 (3H, s, $CO_2C\underline{H}_3$), 4.13 (2H, d, *J* = 4.3, C<u>H</u>$_2$OH), 4.46-4.62 (1H, m, H_α), 5.34 (1H, d, *J* = 7.8, NH), 5.60-5.85 (2H, m, <u>H</u>C=C<u>H</u>). ^{13}C NMR (100 MHz, $CDCl_3$) δ_C = 24.9 (C_β), 25.5 (Se<u>C</u>H$_2$CH=), 28.3 (1 C, s, Boc), 52.6 ($CO_2\underline{C}H_3$), 53.6 (C_α), 62.8 (<u>C</u>H$_2$OH), 80.5 (Boc), 128.1, 132.4 (H<u>C</u>=<u>C</u>H), 155.1, 171.7 (2 × C=O). HRMS m/z (ESI$^+$): Found 354.0810 $[M+H]^+$; $C_{13}H_{24}NO_5Se$ requires 354.0815.

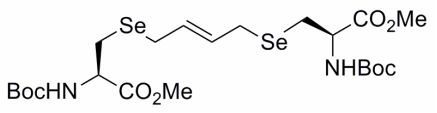
Note: Allyl alcohol self-metathesis product (**3.4b**) is also isolated (76.6 mg, 90%).

CM between 3.20a and Allyl glucoside tetraacetate 3.21

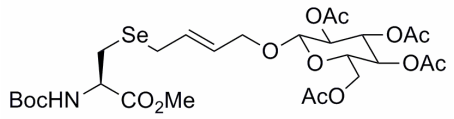


BocSeacOMe (**3.20a**) (64.3 mg, 0.20 mmol) and $MgCl_2.6H_2O$ (60.8 mg, 0.30 mmol) was added to a 25 mL round bottom flask was dissolved in ^{t}BuOH/H_2O (1:2, 4 mL). The mixture was stirred and warmed to 32 °C. Allyl glucoside tetraacetate **3.21** (155 mg, 0.40 mmol) was added as a solid to the stirred mixture. Hoveyda-Grubbs 2nd generation catalyst (3.8 mg, 6.1 μmol) was dissolved in ^{t}BuOH (1 mL) with gentle heating. The catalyst solution was then added to the reaction and the reaction was stirred for 1 hour at 32 °C. Second dose of catalyst (3.8 mg in 1 mL ^{t}BuOH) followed by H_2O (1 mL) was added. After a further 2 hours of reaction time, TLC (40% EtOAc in petrol) revealed the formation of the amino acid self-metathesis product (**3.20c**) (R_f = 0.44) and the desired CM product (**3.20d**) (R_f = 0.15). The reaction was diluted with EtOAc (50 mL) and H_2O (50 mL) then transferred to a separatory funnel. The organic layer was isolated and aqueous layer was further extracted with Et_2O (50 mL). The combined organic layers

were washed with brine (50 mL), dried with $MgSO_4$ and filtered. The resulting solution was concentrated under reduced pressure, providing a dark residue. Purification by flash column chromatography (20% EtOAc in petrol then 40% EtOAc in petrol) provided the amino acid self-metathesis product **3.20c** (12.2 mg, 20%) and the CM product **3.20d** (71.9 mg, 53%).



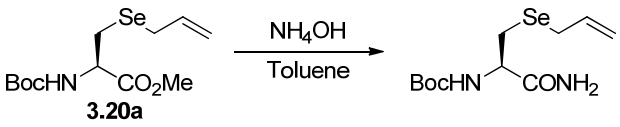
Amino acid homodimer 3.20c: IR (υ_{max}, film): 3368, 2978, 1714, 1505, 1437, 1367, 1214, 1166, 1050, 914, 733. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.45 (18H, s, Boc), 2.82-3.02 (4H, m, H_β), 3.09-3.24 (4H, m, C<u>H</u>$_2$CH=), 3.76 (6H, s, $CO_2C\underline{H}_3$), 4.51-4.69 (2H, m, H_α), 5.36 (2H, d, *J* = 7.8, NH), 5.54-5.61 (2H, m, C<u>H</u>=C<u>H</u>). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 25.3 (C_β, <u>C</u>H_2CH=), 28.3 (Boc), 52.6 ($CO_2\underline{C}H_3$), 53.4 (C_α), 80.2 (Boc 4°), 129.2 (C=C), 155.1, 171.6 (2 × C=O). HRMS m/z (ESI$^+$): Found 641.0855 $[M+Na]^+$; $C_{22}H_{38}N_2O_8Se_2Na$ requires 641.0855.



CM product 3.20d: IR (υ_{max}, film): 3375, 2977, 1755, 1509, 1437, 1368, 1226, 1166, 1042, 908, 735. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.44 (9H, s, Boc), 1.99, 2.01, 2.05, 2.08 (12H, 4 × s, 4 × OAc), 2.86 (1H, dd, *J* = 12.9, 5.5, H_β), 2.93 (1H, dd, *J* = 12.9, 5.3, H_β'), 3.17 (2H, dd, *J* = 7.6, 3.0, SeC<u>H</u>$_2$), 3.69 (1H, ddd, *J* = 9.9, 4.7, 2.3, H5), 3.75 (3H, s, $CO_2C\underline{H}_3$), 4.08 (1H, dd, *J* = 12.7, 6.3, OC<u>H</u>H'), 4.13 (1H, dd, *J* = 12.4, 2.3, H6), 4.25 (1H, dd, *J* = 12.4, 4.7, H6'), 4.32 (1H, dd, *J* = 12.7, 5.4, OCH<u>H</u>'), 4.54 (1H, d, *J* = 7.9, H1), 4.54-4.61 (1H, m, H_α), 4.99 (1H, dd, *J* = 9.6, 7.9, H2), 5.08 (1H, t, *J* = 9.6, H4), 5.20 (1H, t, *J* = 9.5, H3), 5.35 (1H, d, *J* = 7.7, NH), 5.53 (1H, dt, *J* = 15.3, 6.0, $OCH_2C\underline{H}$=), 5.74 (1H, dt, *J* = 15.3, 7.6, $SeCH_2C\underline{H}$=). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 20.57, 20.59, 20.68, 20.73 (4 × OAc), 25.2 (C_β), 25.4 (Se<u>C</u>H_2), 28.3 (Boc), 52.6 ($CO_2\underline{C}H_3$), 53.4 (C_α), 61.9 (C6), 68.3

(C4), 68.9 (O<u>C</u>H$_2$), 71.2 (C2), 71.7 (C5), 72.8 (C3), 80.1 (4° C, Boc), 99.4 (C1), 127.6 (OCH$_2$<u>C</u>H=), 130.4 (SeCH$_2$<u>C</u>H=), 155.0 (C=O, Boc), 169.3, 169.4, 170.2, 170.7 (4 × C=O, Ac), 171.5 (C=O, ester). HRMS m/z (ESI$^+$): Found 706.1588 $[M+Na]^+$; $C_{27}H_{41}NO_{14}SeNa$ requires 706.1587.

_N_-Boc-_Se_-allyl-L-selenocysteine amide (BocSeAcNH$_2$)

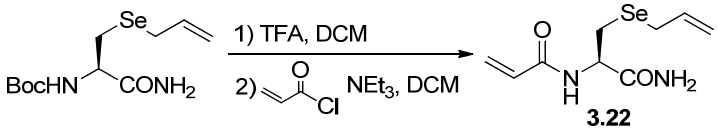


3.20a

BocSeacOMe (**3.20a**) (1.5 g, 1.60 mmol) was dissolved in toluene (15 mL) and transferred to an argon-filled 25 mL round bottomed flask. 25% NH_4OH (aq., 15 mL) was then added. The resulting emulsion was stirred vigorously at room temperature for 30 hours. The reaction was diluted with H_2O (100 mL) and transferred to a separatory funnel. The aqueous layer was extracted with EtOAc (2 × 100 mL). The combined organic layers were washed with brine (100 mL), dried ($MgSO_4$) and filtered. The solvent was removed under reduced pressure providing a yellow oil. After purification by flash column chromatography (60% EtOAc in petrol), unreacted starting material was isolated as a yellow oil (459 mg, 31% recovery) and the titled product as a white solid (571 mg, 40%, 58% brsm). $[\alpha]^{20}_D = -26.0$ (c = 1.0, MeOH). m.p. = 98-99 °C. IR (υ_{max}, KBr): 3390, 3189, 2981, 2935, 1663, 1517, 1425, 1368, 1312, 1252, 1165, 1115, 1047, 1018, 911, 821, 778, 616. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.42 (9H, s, Boc), 2.79 (1H, dd, *J* = 12.9, 6.1, H_β), 2.975 (1H, dd, *J* = 12.9, 6.3, H_β'), 3.19 (2H, d, *J* = 7.6, C<u>H</u>$_2$CH=CH$_2$), 4.25-4.45 (1H, m, H_α), 5.00 (1H, dd, *J* = 9.8, 1.3, CH$_2$CH=C<u>H</u>H *cis*), 5.06 (1H, dd, *J* = 16.9, 1.3, CH$_2$CH=CH<u>H</u> *trans*), 5.54 (1H, d, *J* = 6.6, NHBoc), 5.84 (1H, m, J = 16.9, 9.8, 7.6, CH$_2$C<u>H</u>=CH$_2$), 6.34 (1H, br. s., CON<u>H</u>H), 6.60 (1H, br. s., CONH<u>H</u>). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 25.1 (C_β), 26.8 (Se<u>C</u>H$_2$), 28.3 (Boc), 53.8 (C_α), 80.3 (4° C Boc), 117.1 (CH=<u>C</u>H$_2$), 134.5 (<u>C</u>H=CH$_2$), 155.4 (C=O, Boc), 173.7 (C=O, amide). HRMS m/z

(ESI$^+$): Found 331.0523 $[M+Na]^+$; $C_{11}H_{20}N_2O_3SeNa$ requires 331.0532.

_N_-acryl-_Se_-allyl selenocysteine amide (3.22)

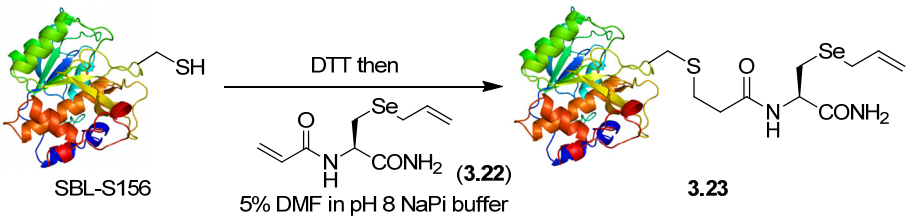


3.22

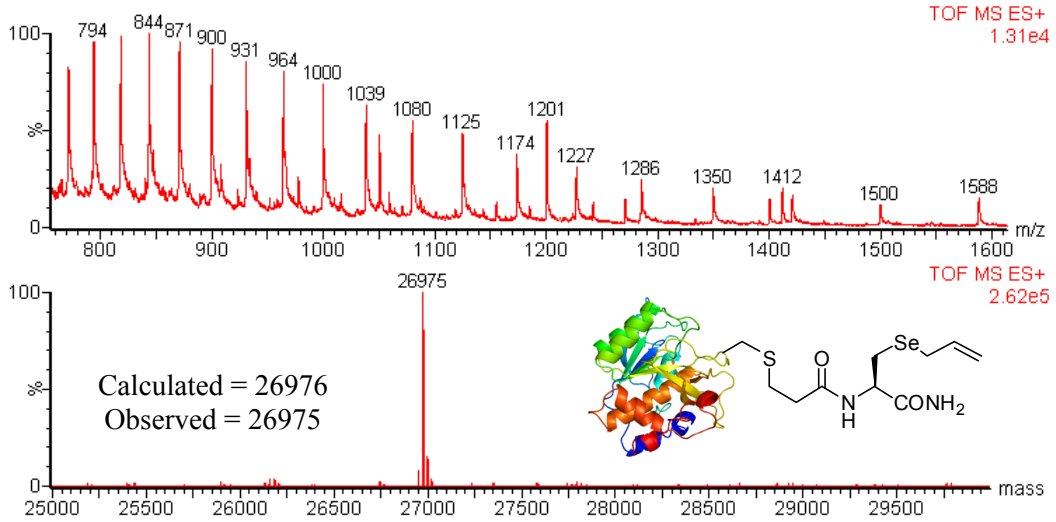
$BocSeacNH_2$ (421 mg, 1.37 mmol) was added to a 25 mL round bottomed flask was dissolved with DCM (8 mL). The solution was stirred and placed under an argon atmosphere. Trifluoroacetic acid (TFA) (0.8 mL, 10.77 mmol) was added as one portion at room temperature. One hour later, a second dose of TFA (0.8 mL, 10.77 mmol) was added. The reaction was stirred for an additional 3 hours, after which time TLC (70% EtOAc in petrol) revealed complete consumption of starting material. Solvent and excess TFA was removed under reduced pressure. The resulting residue was dried briefly under high vacuum before re-dissolving in DCM (8 mL). The solution was cooled to 0 °C and NEt_3 (0.8 mL, 5.48 mmol) was added dropwise over 1 minute. The solution was stirred and placed under an argon atmosphere. Acryloyl chloride (0.13 mL, 1.51 mmol) was then added. The ice bath was removed and reaction was stirred at room temperature for 16 hours. The solvent was removed under reduced pressure. The resulting crude solid was dried briefly under high vacuum before re-dissolving with EtOAc (100 mL) and H_2O (100 mL). The organic layer was isolated and aqueous layer was further extracted with EtOAc (2 × 50 mL). The combined EtOAc layers were washed with brine (150 mL), dried with $MgSO_4$ and filtered. Solvent was removed under reduced pressure to reveal product **3.22** as a white crystalline solid. NMR showed that the product is spectroscopically pure and required no further purification (311 mg, 87%). $[\alpha]^{20}_D$ = −42.0 (c = 1.0, MeOH). m.p. = 123-124 °C. IR (υ_{max}, KBr): 3296, 1926, 1653, 1529, 1404, 1310, 1236, 1120, 1072, 963, 912, 804, 692. ^{1}H NMR (400 MHz, CD_3OD): δ_H = 2.78 (1H, dd, J = 12.9, 7.8, H_β), 2.95 (1H, dd, J = 12.9, 6.1, H_β'), 3.26 (2H, d, J = 7.7,

C<u>H</u>$_2$CH=CH$_2$), 4.63 (1H, dd, *J* = 7.8, 6.1, H$_\alpha$), 5.01 (1H, dd, *J* = 9.8, 1.5, CH$_2$CH=C<u>H</u>H *cis*), 5.10 (1H, dq, *J* = 17.0, 1.5, CH$_2$CH=CH<u>H</u> *trans*), 5.72 (1H, dd, *J* = 9.8, 2.3, COCH=C<u>H</u>H cis), 5.90 (1H, m, *J* = 17.0, 9.8, 7.7, CH$_2$C<u>H</u>=CH$_2$), 6.27 (1H, dd, *J* = 17.1, 2.3, COCH=CH<u>H</u> trans), 6.355 (1H, dd, *J* = 17.1, 9.8, COC<u>H</u>=CH$_2$). ^{13}C NMR (100 MHz, CD$_3$OD): δ_C = 24.0 (C$_\beta$), 26.1 (<u>C</u>H$_2$CH=CH$_2$), 53.5 (C$_\alpha$), 116.0 (CH$_2$CH=<u>C</u>H$_2$), 126.5 (COCH=<u>C</u>H$_2$), 130.6 (CO<u>C</u>H=CH$_2$), 135.1 (CH$_2$<u>C</u>H=CH$_2$), 166.8, 174.3 (2 × C=O). HRMS m/z (ESI$^+$): Found 285.0108 [M+Na]$^+$; $C_9H_{14}N_2O_2SeNa$ requires 285.0113.

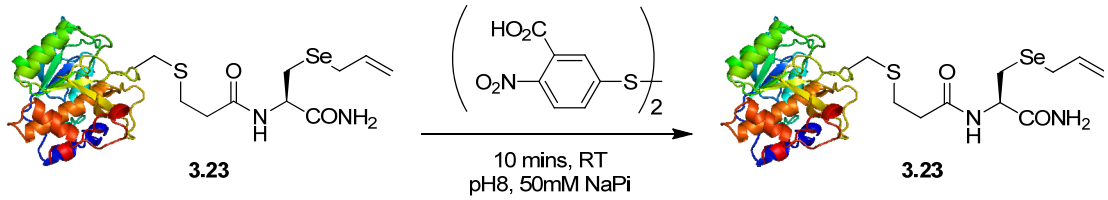
SBL-S156C conjugate addition to *N*-acryl-*Se*-allyl selenocysteine amide (3.22)



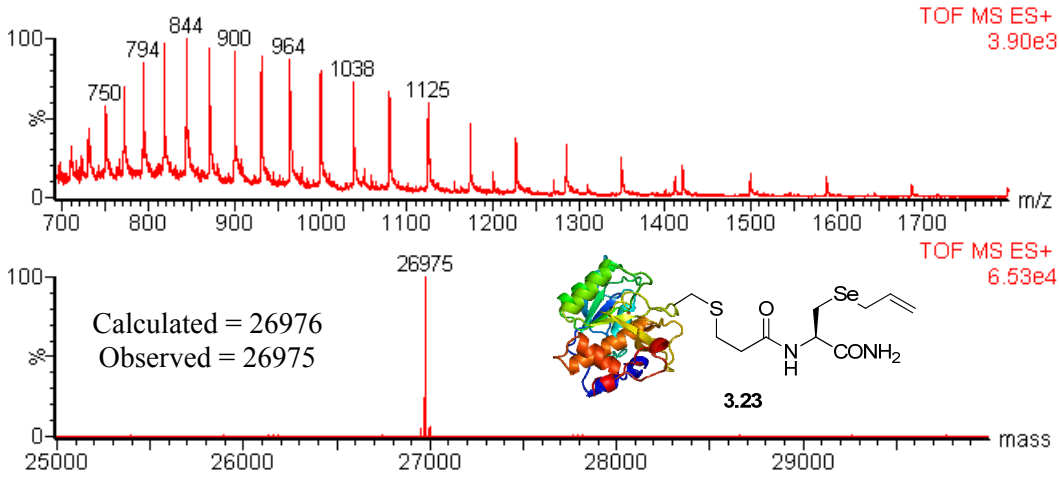
SBL-S156C (2.5 mL, 1 mg/mL, 93.6 nmol, pH 8.0 sodium phosphate) was thawed and put on ice. The protein solution was transferred to a 15 mL falcon tube and DTT (2.0 mg, 13.0 µmol) was added as a solid at RT. Solution was vortexed to mix and shaken for 10 minutes. *N*-acryl-*Se*-allyl selenocysteine amide (**3.22**) (19.6 mg, 74.9 µmol) was weighed into a 1.5 mL eppendorf tube and dissolved with DMF (131.6 µL). This DMF solution was then transferred to the protein solution. The reaction was vortexed and shaken at 37 °C for one hour. LC-MS analysis showed full conversion to the Michael addition product (calculated = 26976, found = 26975). Small molecules were removed by PD10 (GE Healthcare), eluting with 50 mM sodium phosphate buffer (pH 8.0). The sample was then split into 250 µL aliquots, flash frozen, and stored at −80 ºC. The final concentration of protein **3.23** was 0.71 mg/mL.



Ellman's Test on Protein 3.23 to Confirm Reaction at Cysteine



A buffered solution of Ellman's reagent was prepared by dissolving 3.6 mg of Ellman's reagent in 842 µL of 50 mM sodium phosphate buffer (pH 8). A 20 µL of the Ellman's solution (0.2 µmol) was added to protein **3.23** (100 µL, 0.59 mg/mL, 2.2 nmol) in an eppendorf tube. 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 = 26976, found = 26975). This assay indicates that all free thiol is consumed, thus the alkylation occurred at cysteine.



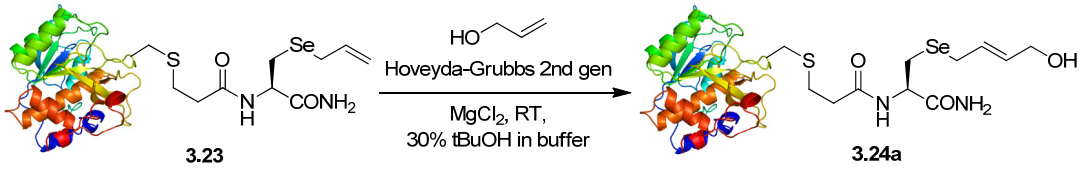
If free cysteine was present the Ellman's adduct would form, as shown on page 110).

Cross-Metathesis of SBL-extended Seac (3.23)

General Procedure for Cross-Metathesis of SBL-extended Seac (3.23)

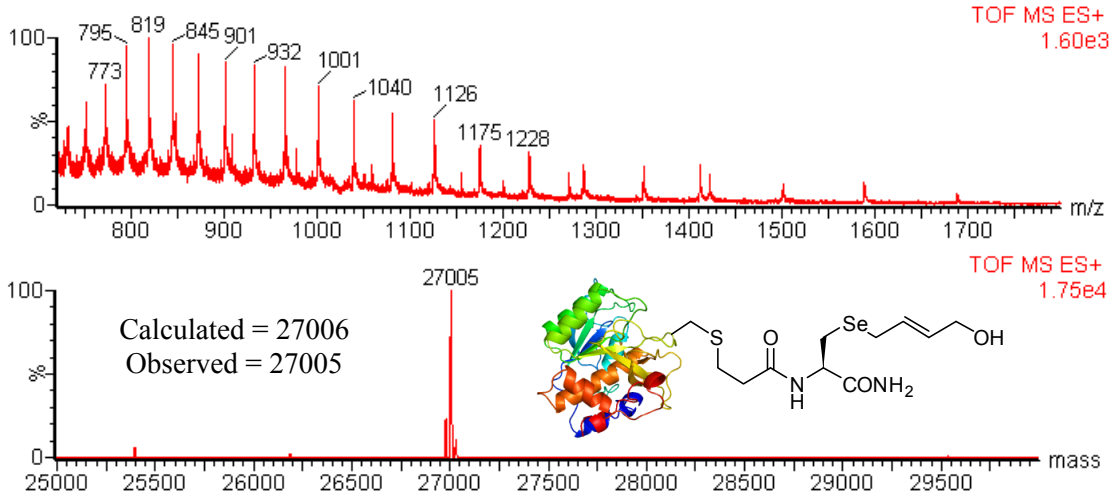
All manipulations were carried out at room temperature unless stated otherwise. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH (6.4 mg/mL) was prepared by vortexing and gentle warming **3.1** in *t*BuOH. Protein **3.23** (200 µL, 0.59 mg/mL, 4.37 nmol) was transferred to a 1.5 mL eppendorf tube and placed on ice. $MgCl_2.6H_2O$ (8.7 mg, 42.8 µmol) was added as a solid to the protein solution and vortexed to mix. The catalyst/*t*BuOH solution (85.7 µL, ~0.875 µmol) was added to the reaction and vortexed immediately to mix. The reaction mixture appeared as a green emulsion. The reaction was shaken at RT for 1-2 minute before adding the alkene substrate (5000 equiv., 21.9 µmol) as a solid or neat liquid unless stated otherwise. The reaction was vortexed immediately after addition and was shaken at RT or 37 °C, as indicated. All reaction was monitored by LC-MS.

CM between protein 3.23 and allyl alcohol (3.4a)

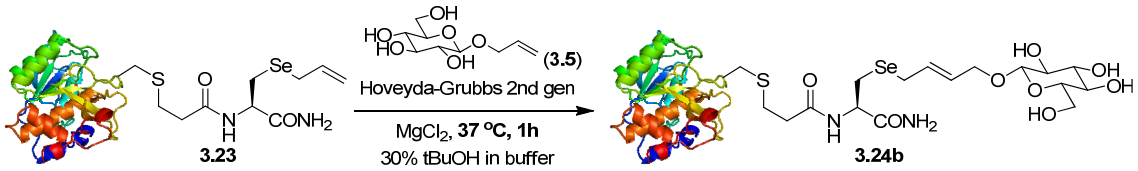


3.23 3.24a

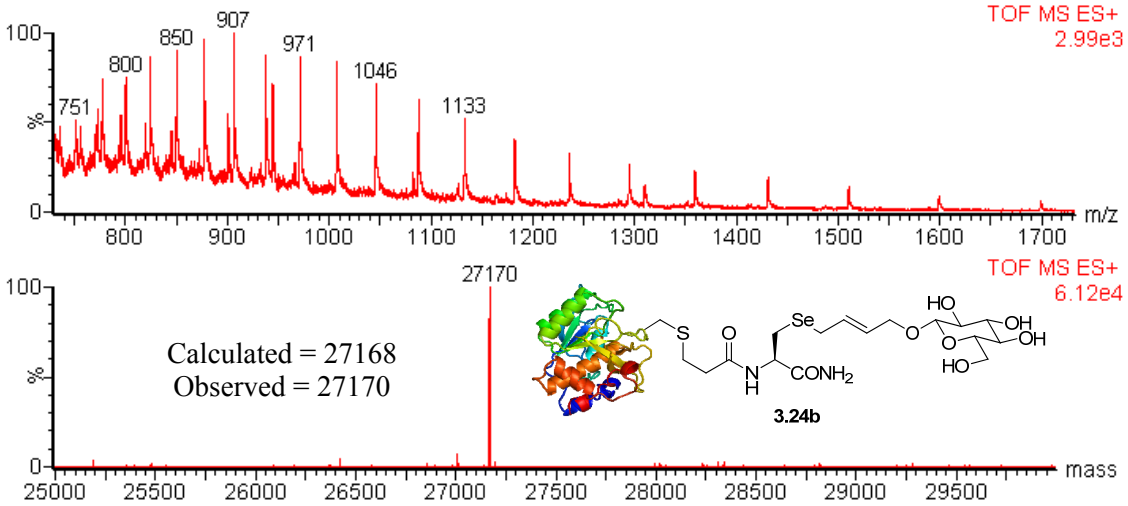
T = RT. After 15 minutes, LC-MS analysis of the product showed full conversion to the CM product **3.24a** (calculated: 27006, found: 27005). ESI-MS for the reaction is shown below:



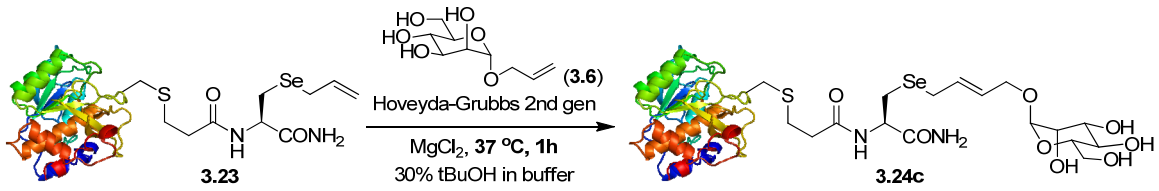
CM between protein 3.23 and *β-O*-Allyl-D-glucose (3.5)



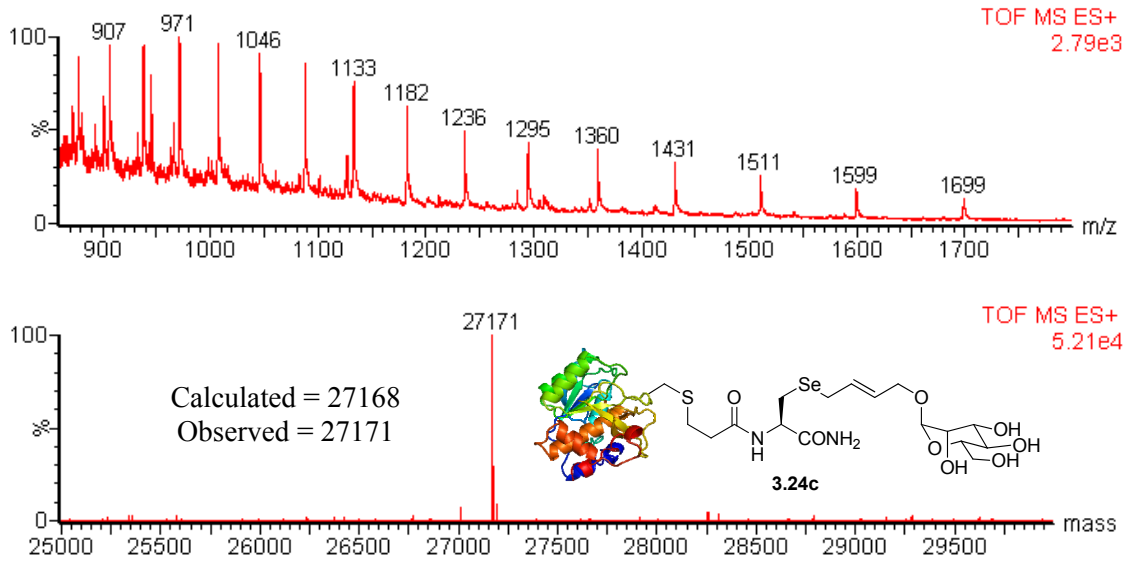
T = 37 °C. After 1 hour LC-MS analysis of the product showed full conversion to the glycosylated protein **3.24b** (calculated: 27168, found: 27170). ESI-MS for the reaction is shown below:



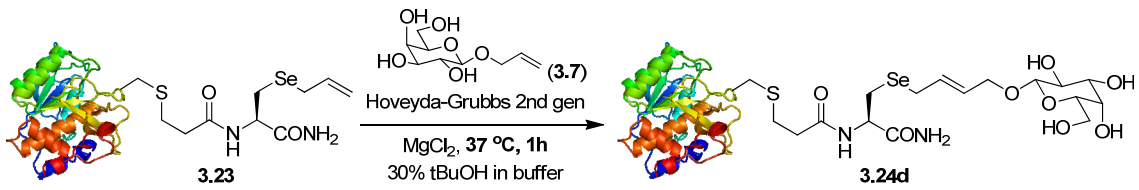
CM between protein 3.24 and *α*-*O*-Allyl-D-mannose (3.6)



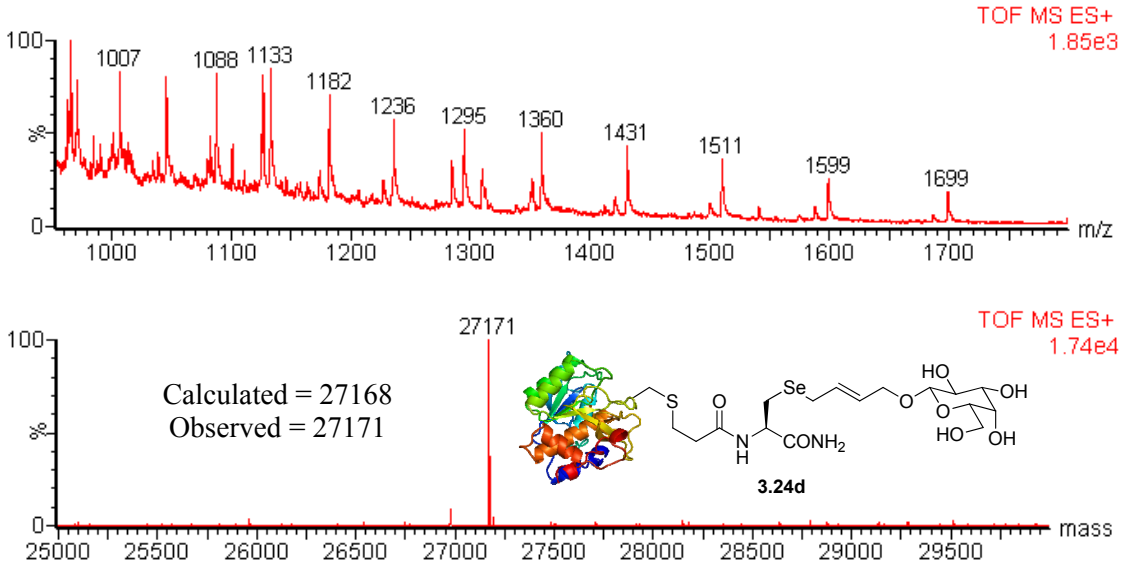
α-*O*-Allyl-D-mannose (**3.6**) was added as a solution in pH 8 NaPi (50 mM) buffer (15 µL of 2 M solution, 6.6 mg, 30.0 µmol). T = 37 °C. After 1 hour LC-MS analysis of the product showed full conversion to the glycosylated protein **3.24c** (calculated: 27168, found: 27171). ESI-MS for the reaction is shown below:



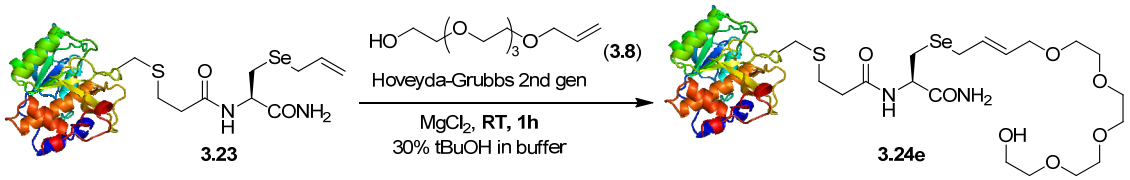
CM between protein 3.23 and *β*-*O*-Allyl-D-galactose (3.7)



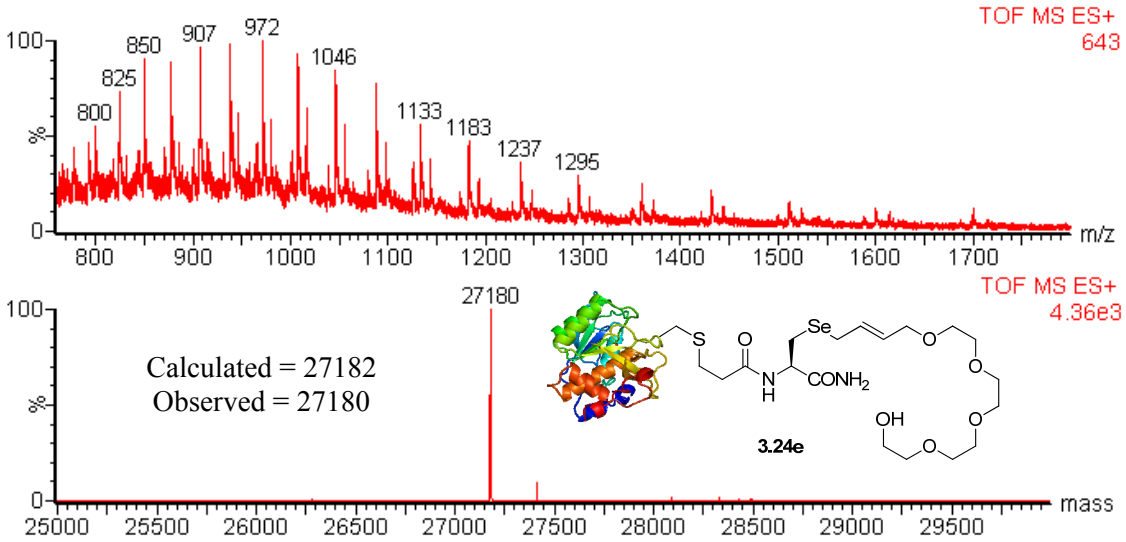
T = 37 °C. After 1 hour LC-MS analysis of the product showed full conversion to the glycosylated protein **3.24d** (calculated: 27168, found: 27171). ESI-MS for the reaction is shown below:



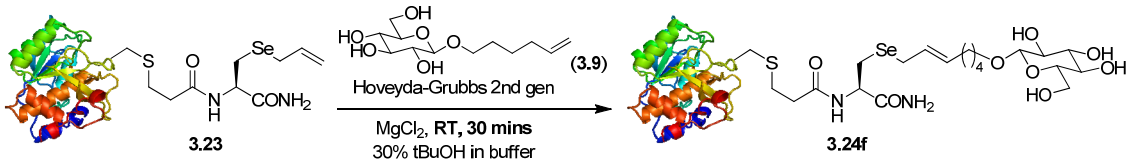
CM between protein 3.23 and Allyl tetraethylene glycol (3.8)



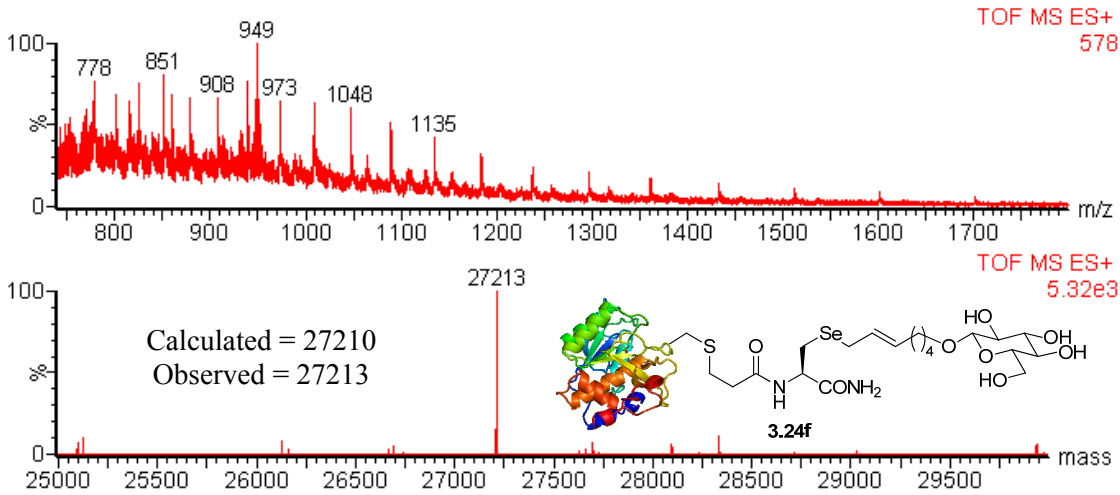
T = RT. After 1 hour LC-MS analysis of the product showed full conversion to the modified protein **3.24e** (calculated: 27182, found: 27180). ESI-MS for the reaction is shown below:



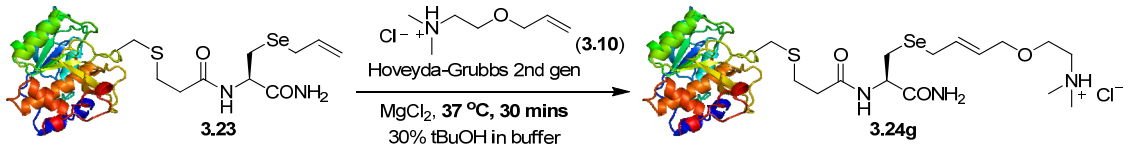
CM between protein 3.23 and *β-O*-hexenyl-D-glucose (3.9)



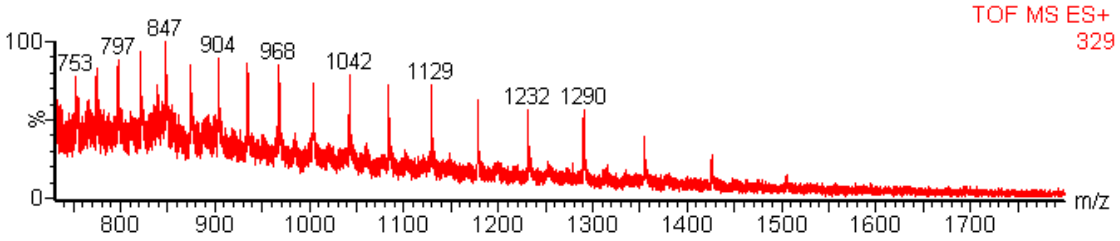
T = RT. After 30 minutes LC-MS analysis of the product showed full conversion to the modified protein **3.24f** (calculated: 27210, found: 27213). ESI-MS for the reaction is shown below:

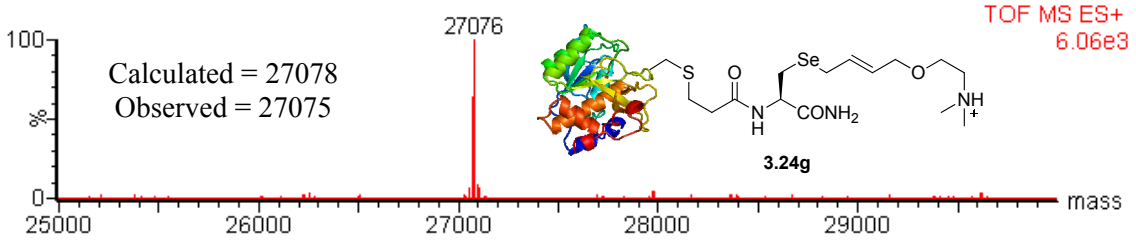


CM of protein 3.23 with *N,N*-dimethyl-*O*-Allyl-ethanolamine hydrochloride (3.10)

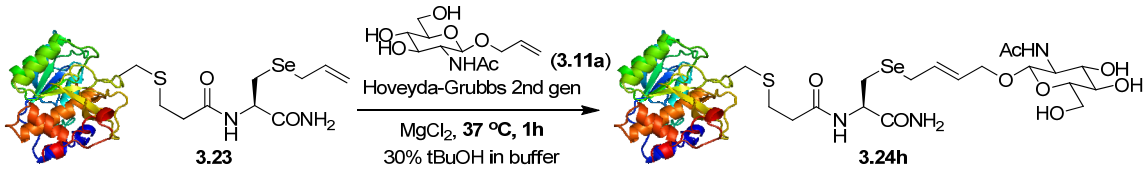


T = 37 °C. After 30 minutes, LC-MS analysis of the product showed full conversion to the modified protein **3.24g** (in ES+ mode, mass without Cl^- was detected. Calculated: 27078, found: 27076). ESI-MS for the reaction is shown below:

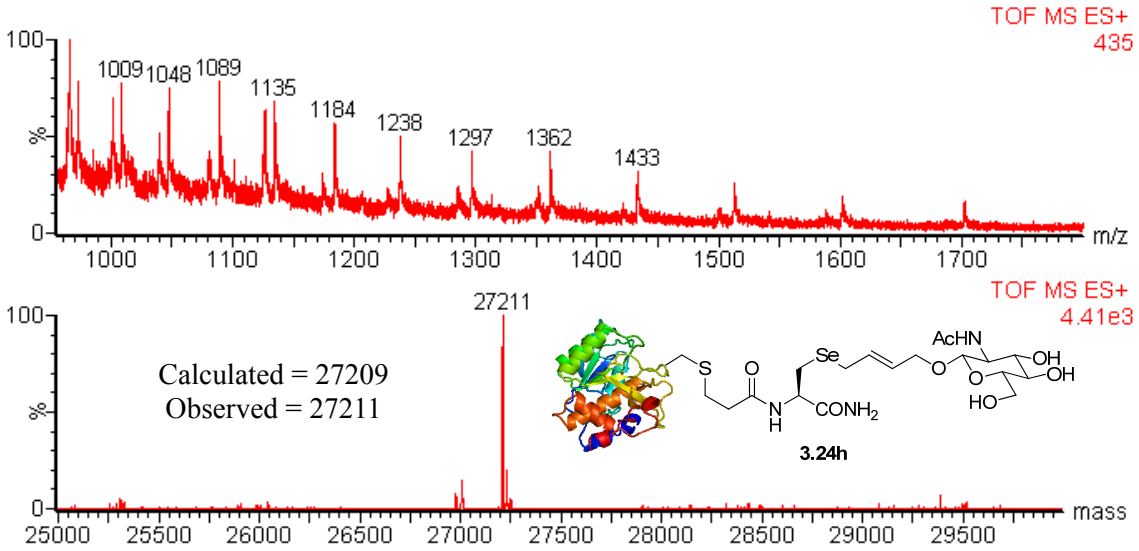




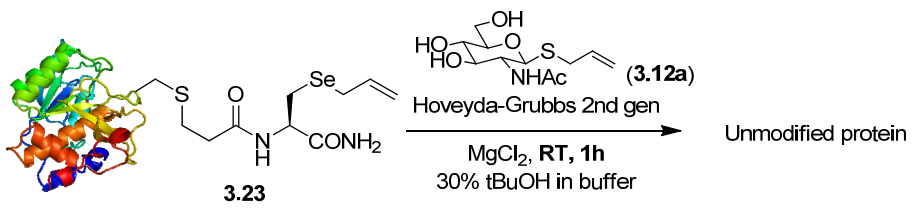
CM between protein 3.23 and *β-O*-allyl-D-GlcNAc (3.11a)



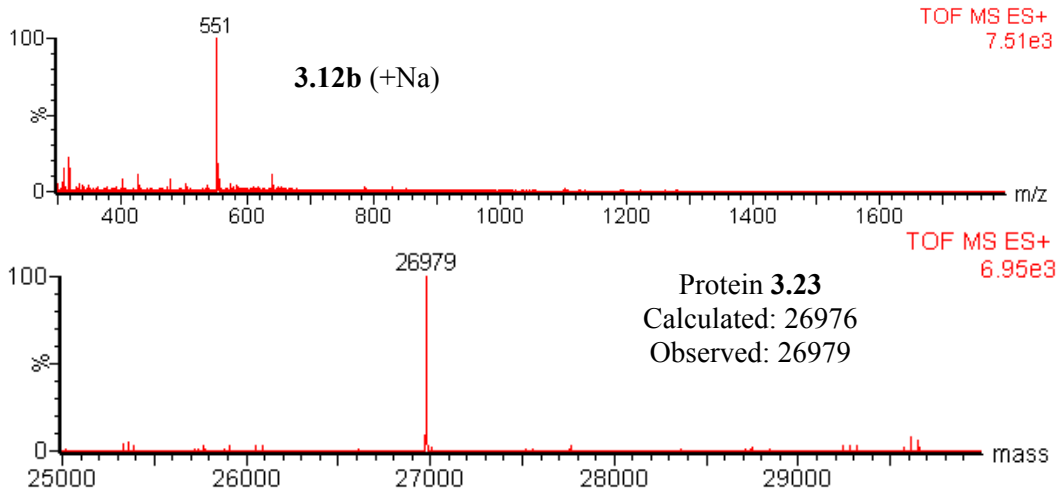
β-O-allyl-D-GlcNAc (**3.11a**) (5.7 mg, 21.9 µmol) was added as a solution in 20 µL of pH 8 NaP_i (50 mM) buffer. T = 37 °C. After 1 hour LC-MS analysis of the product showed full conversion to the glycosylated protein **3.24h** (calculated: 27209, found: 27211). ESI-MS for the reaction is shown below:



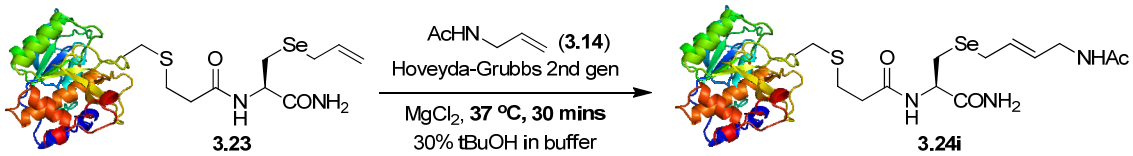
CM between protein 3.23 and GlcNAc allyl sulfide (3.12a)



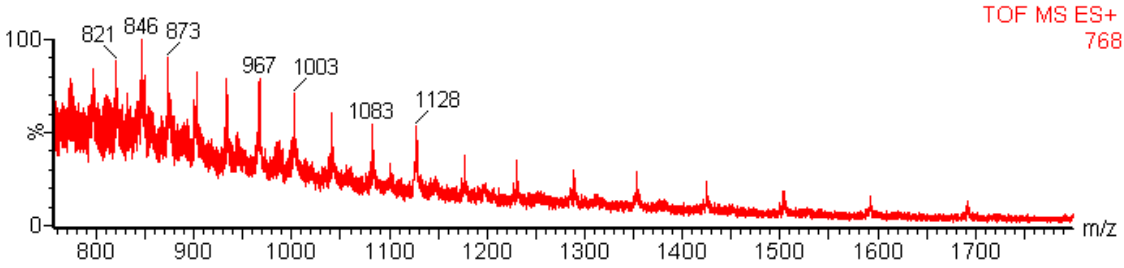
GlcNAc allyl sulfide **3.12a** (6.1 mg, 21.9 µmol) was added as a solution in 20 µL of pH 8 NaP_i (50 mM) buffer. T = RT. After 1 hour LC-MS analysis of the product showed only **3.12b** and unmodified protein **3.23** (calculated: 26976, found: 26979). ESI-MS for the reaction is shown below:

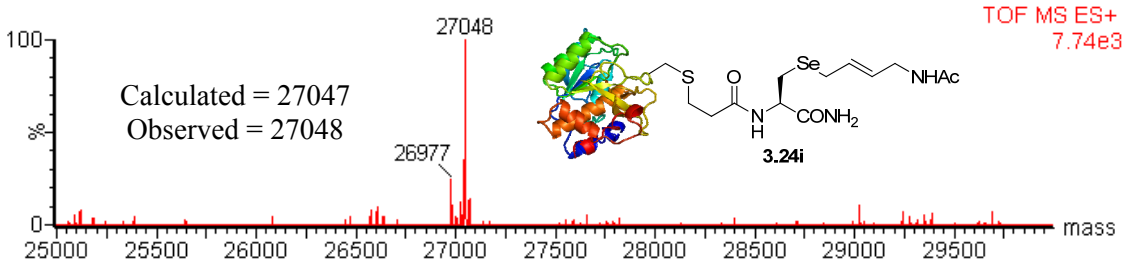


CM between protein 3.23 and *N*-acetylallylamine (3.14)



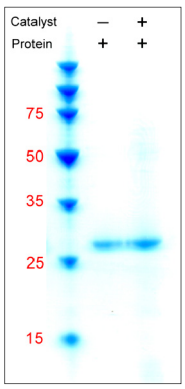
T = 37 °C. After 30 minutes, LC-MS analysis of the product showed 90% conversion to the modified protein **3.24i** (calculated: 27047, found: 27048). ESI-MS for these reactions are shown below:





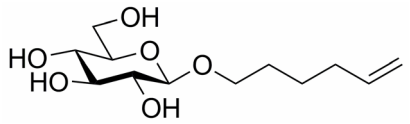
Protein 3.23 self-metathesis attempt

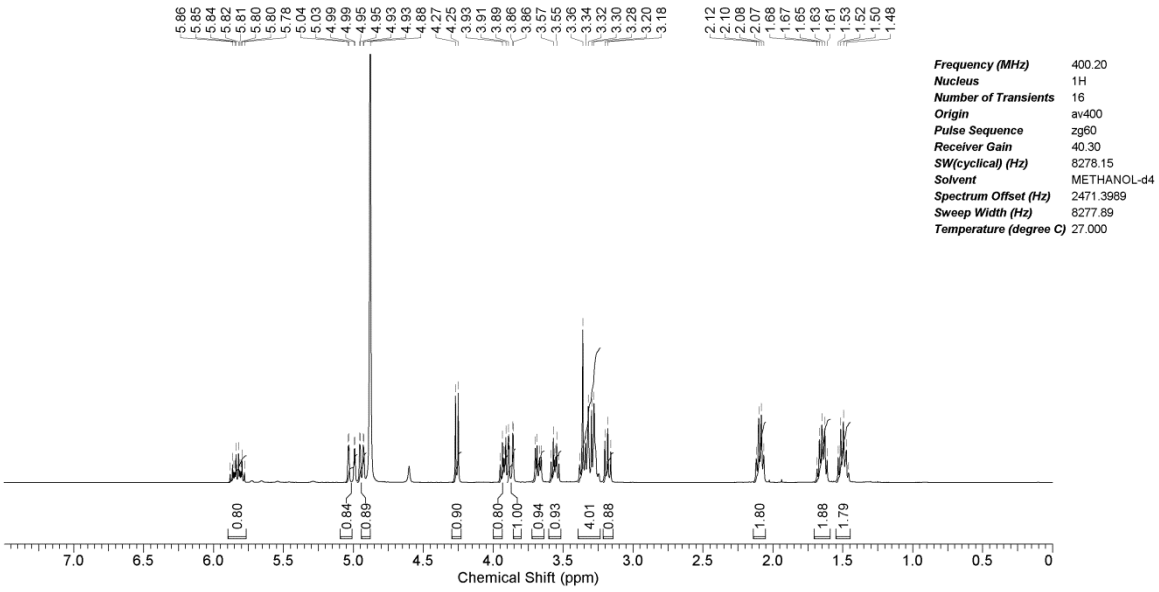
All manipulations were carried out at room temperature unless stated otherwise. The protein solution was prepared in 50 mM sodium phosphate buffer at pH 8. A solution of Hoveyda-Grubbs 2nd generation catalyst (**3.1**) in ^{t}BuOH (3.5 mg/mL) was prepared by vortexing and gentle warming 1.9 mg of **3.1** in 543 µL *t*BuOH. Protein **3.23** (100 µL, 0.63 mg/mL, 2.34 nmol) was transferred to a 1.5 mL eppendorf tube and placed on ice. $MgCl_2.6H_2O$ (4.7 mg, 23.4 µmol) was added as a solid to the protein solution and vortexed to mix. The catalyst/*t*BuOH solution (42.9 µL, ~0.234 µmol) was added to the reaction, vortexed immediately to mix. The reaction mixture appeared as a pale green emulsion. Reaction was shaken at RT for 5 minute before adding second equivalent of protein **3.23** (100 µL, 0.63 mg/mL, 2.34 nmol). Reaction was vortexed and shaken at RT for 2 hours. No protein self-metathesis was observed by SDS-PAGE gel analysis.

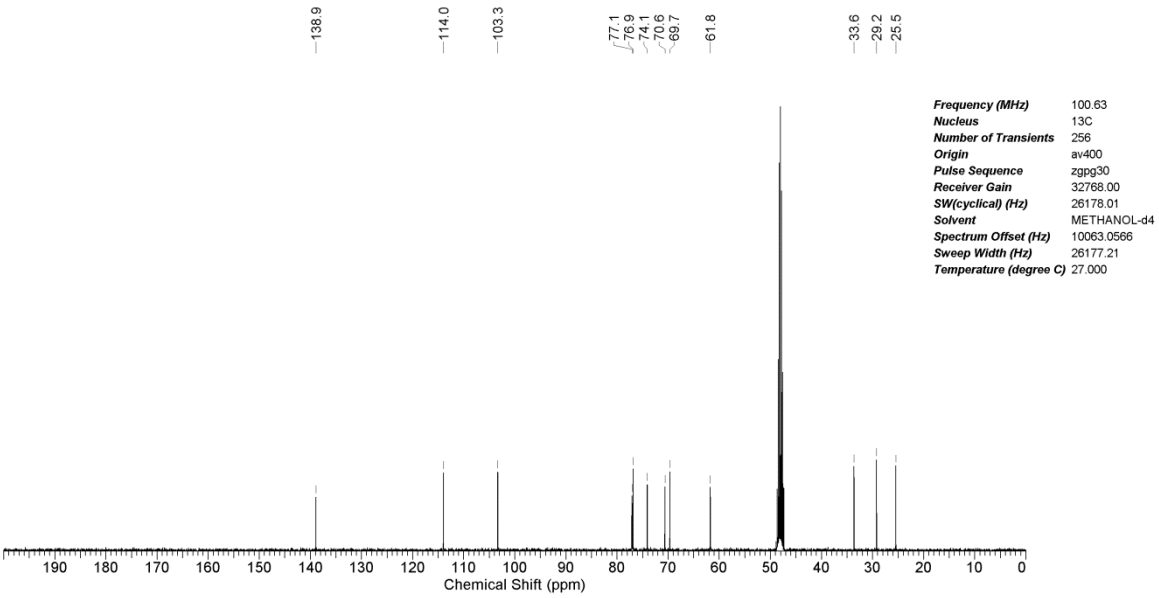


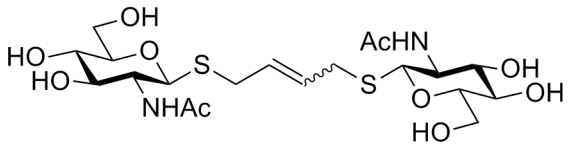
Protein **3.23**
Calculated mass: 26796

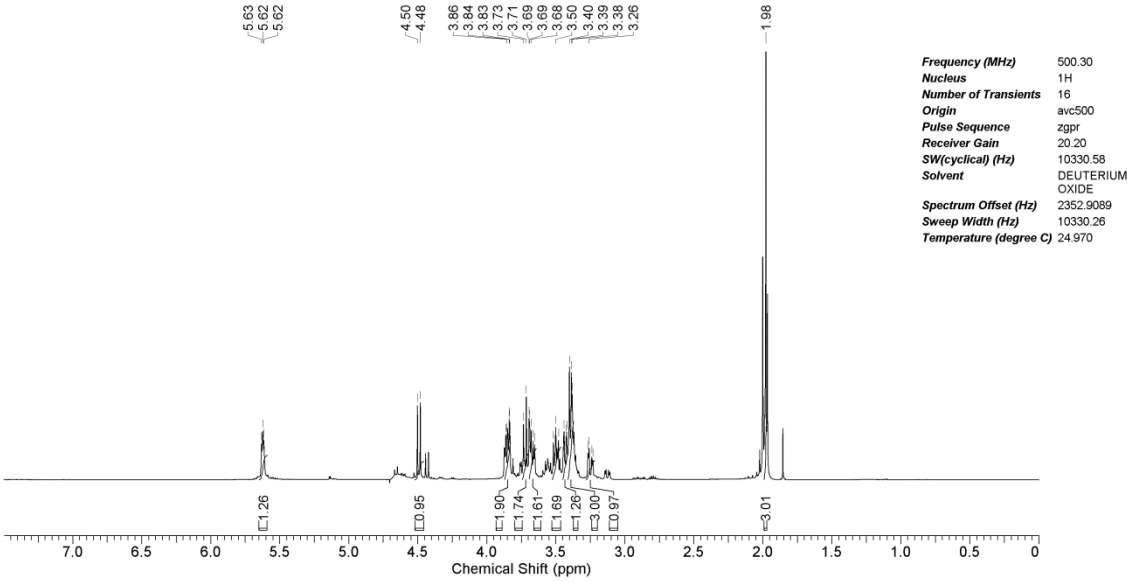
3.6 NMR Spectra

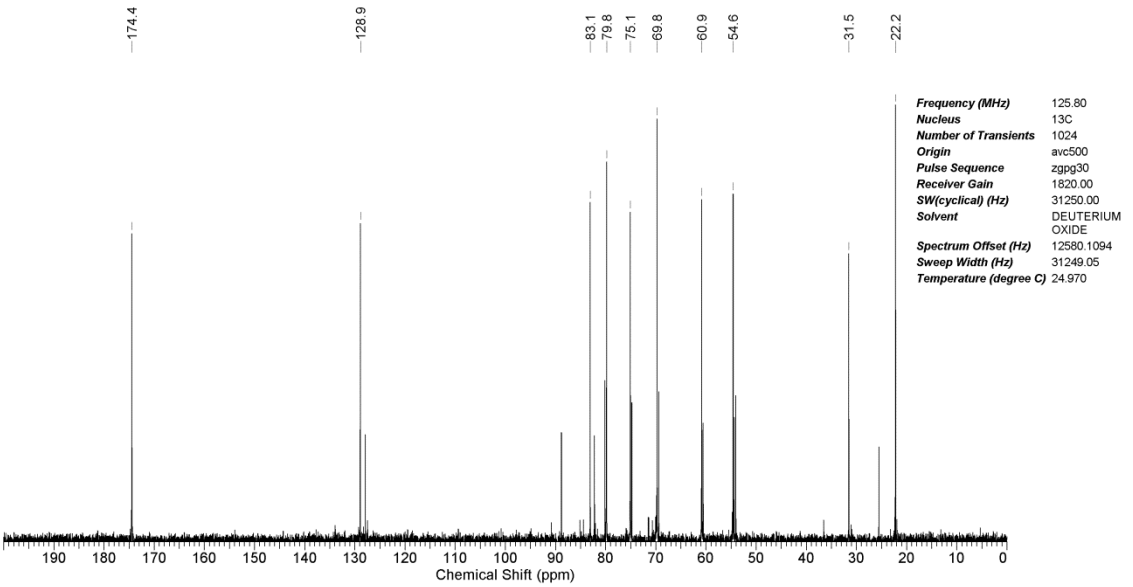


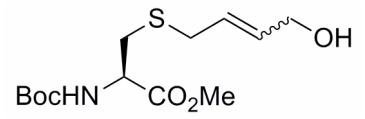


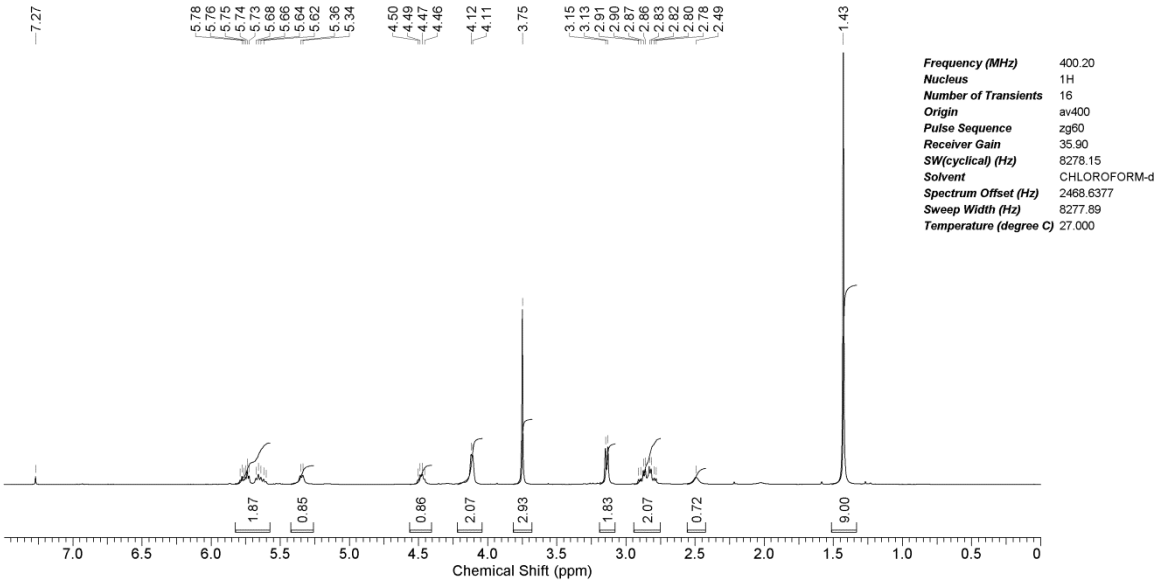


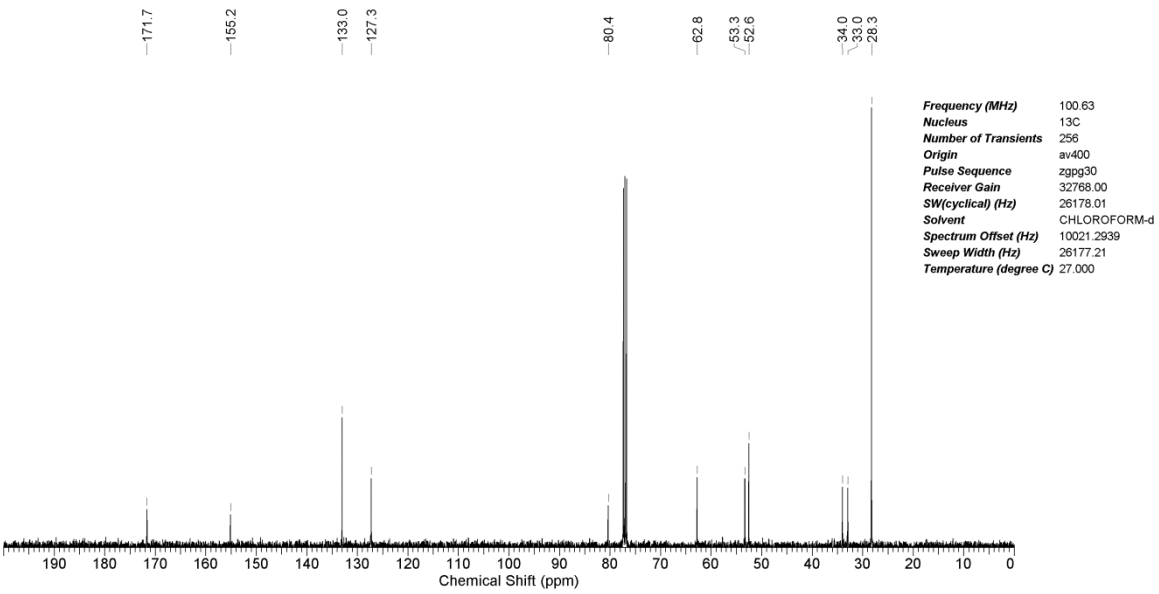

OH
HO
O
HO
S
NHAc
AcHN
S
OH
O
OH
HO

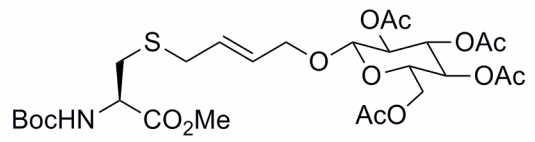

5.63
5.62
5.62
4.50
4.48
3.86
3.84
3.83
3.73
3.71
3.69
3.69
3.68
3.50
3.40
3.39
3.38
3.26
1.98
Frequency (MHz) 500.30
Nucleus 1H
Number of Transients 16
Origin avc500
Pulse Sequence zgpr
Receiver Gain 20.20
SW(cyclical) (Hz) 10330.58
Solvent DEUTERIUM OXIDE
Spectrum Offset (Hz) 2352.9089
Sweep Width (Hz) 10330.26
Temperature (degree C) 24.970
1.26
0.95
1.90
1.74
1.61
1.69
1.26
3.00
0.97
3.01
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)

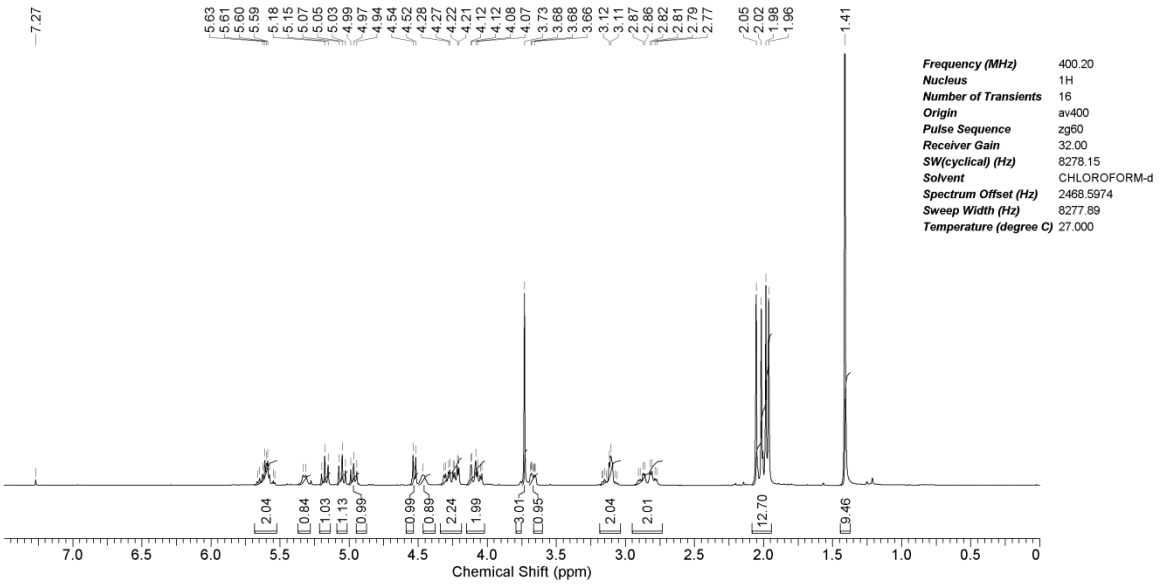

174.4
128.9
83.1
79.8
75.1
69.8
60.9
54.6
31.5
22.2
Frequency (MHz) 125.80
Nucleus 13C
Number of Transients 1024
Origin avc500
Pulse Sequence zgpg30
Receiver Gain 1820.00
SW(cyclical) (Hz) 31250.00
Solvent DEUTERIUM OXIDE
Spectrum Offset (Hz) 12580.1094
Sweep Width (Hz) 31249.05
Temperature (degree C) 24.970
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

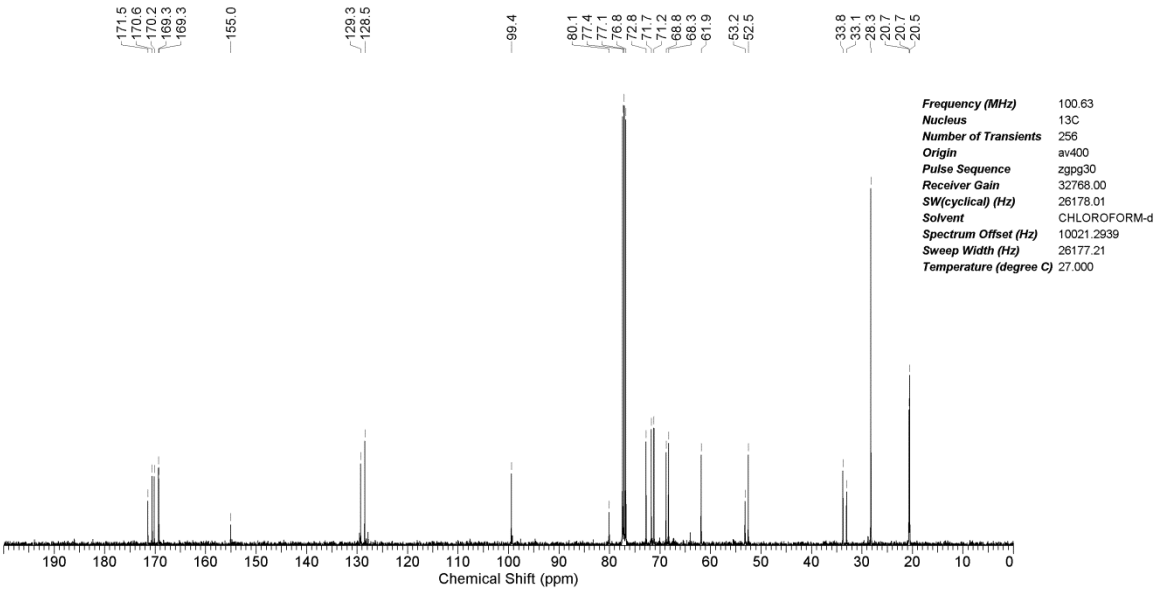

S
OH
BocHN
CO2Me

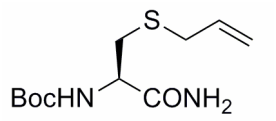

7.27
5.78
5.76
5.75
5.74
5.73
5.68
5.66
5.64
5.62
5.36
5.34
4.50
4.49
4.47
4.46
4.12
4.11
3.75
3.15
3.13
2.91
2.90
2.87
2.86
2.83
2.82
2.80
2.78
2.49
1.43
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 35.90
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.6377
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000
1.87
0.85
0.86
2.07
2.93
1.83
2.07
0.72
9.00
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)

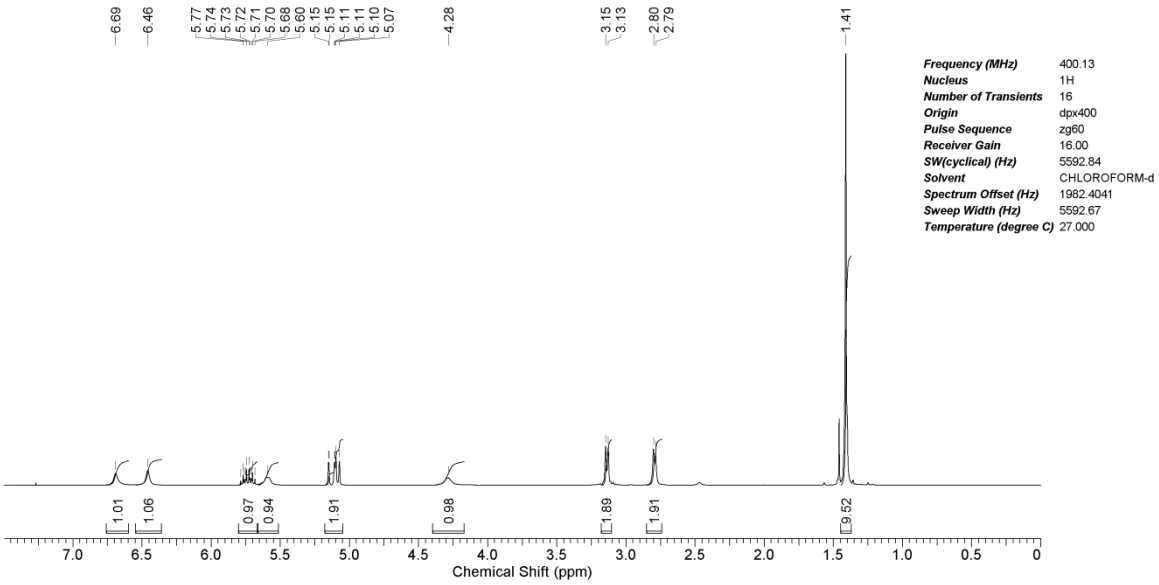

171.7
155.2
133.0
127.3
80.4
62.8
53.3
52.6
34.0
33.0
28.3
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

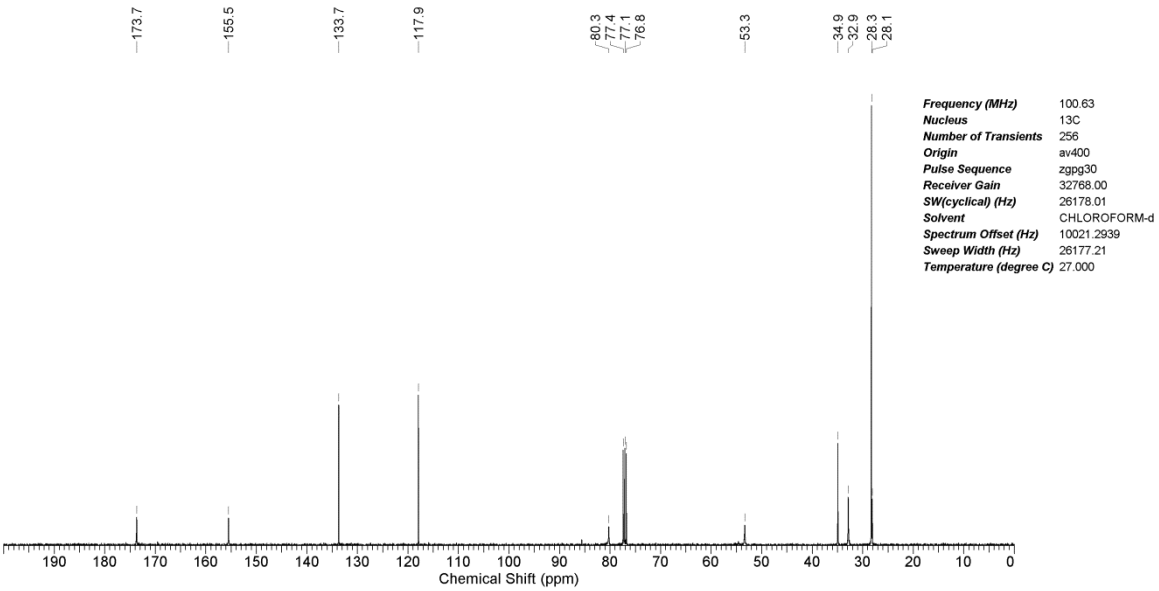

OAc
OAc
OAc
S
O
O
BocHN
CO2Me
AcO

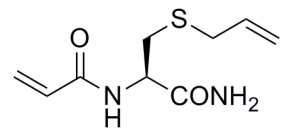

7.27
5.63
5.61
5.60
5.59
5.18
5.15
5.07
5.05
5.03
4.99
4.97
4.94
4.54
4.52
4.28
4.27
4.22
4.21
4.12
4.12
4.08
4.07
3.73
3.68
3.68
3.66
3.12
3.11
2.87
2.86
2.82
2.81
2.79
2.77
2.05
2.02
1.98
1.96
1.41
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 32.00
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000
2.04
0.84
1.03
1.13
0.99
0.99
0.89
2.24
1.99
3.01
0.95
2.04
2.01
12.70
9.46
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)

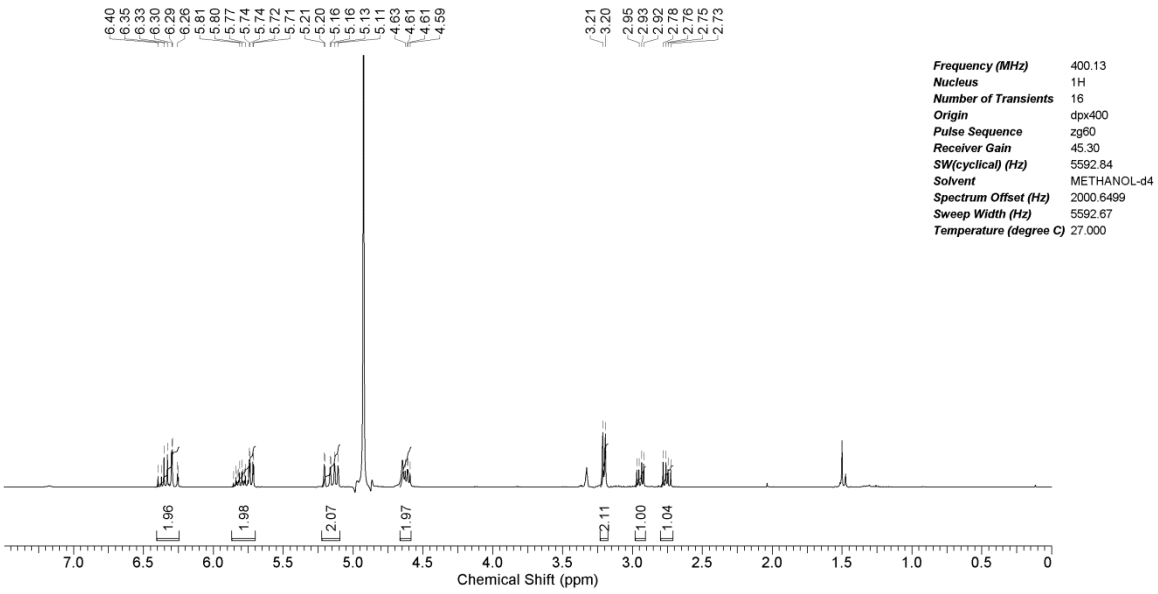

171.5
170.6
170.2
169.3
169.3
155.0
129.3
128.5
99.4
80.1
77.4
77.1
76.8
72.8
71.7
71.2
68.8
68.3
61.9
53.2
52.5
33.8
33.1
28.3
20.7
20.7
20.5
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

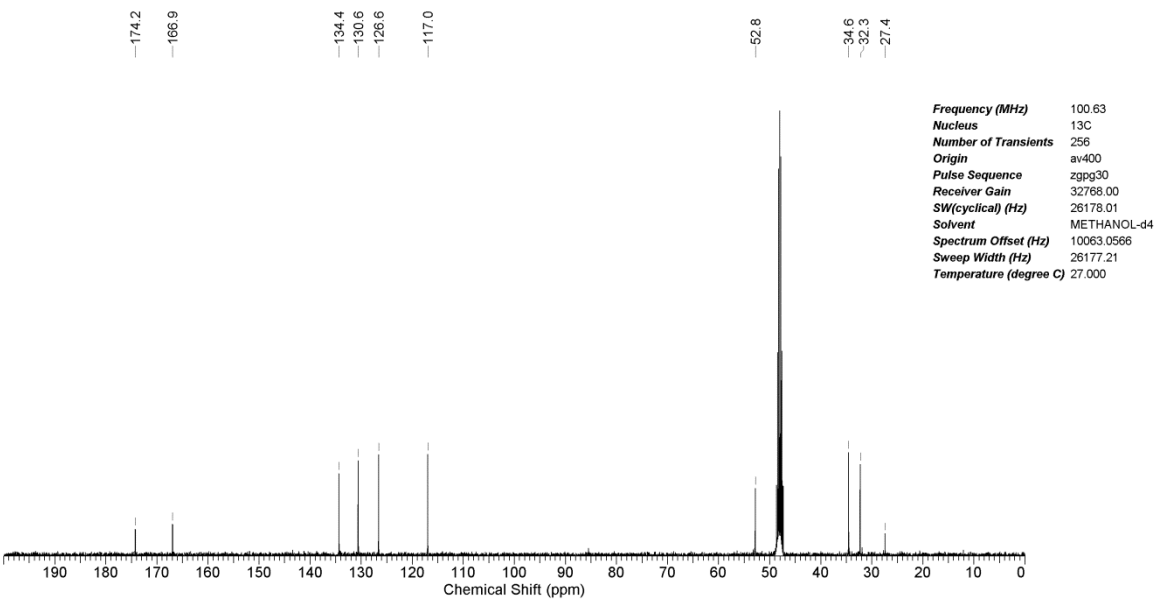

S
BocHN
CONH2

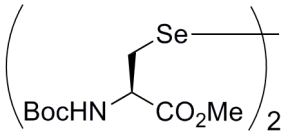

6.69
6.46
5.77
5.74
5.73
5.72
5.71
5.70
5.68
5.60
5.15
5.15
5.11
5.11
5.10
5.07
4.28
3.15
3.13
2.80
2.79
1.41
1.01
1.06
0.97
0.94
1.91
0.98
1.89
1.91
9.52
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)
Frequency (MHz) 400.13
Nucleus 1H
Number of Transients 16
Origin dpx400
Pulse Sequence zg60
Receiver Gain 16.00
SW(cyclical) (Hz) 5592.84
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 1982.4041
Sweep Width (Hz) 5592.67
Temperature (degree C) 27.000

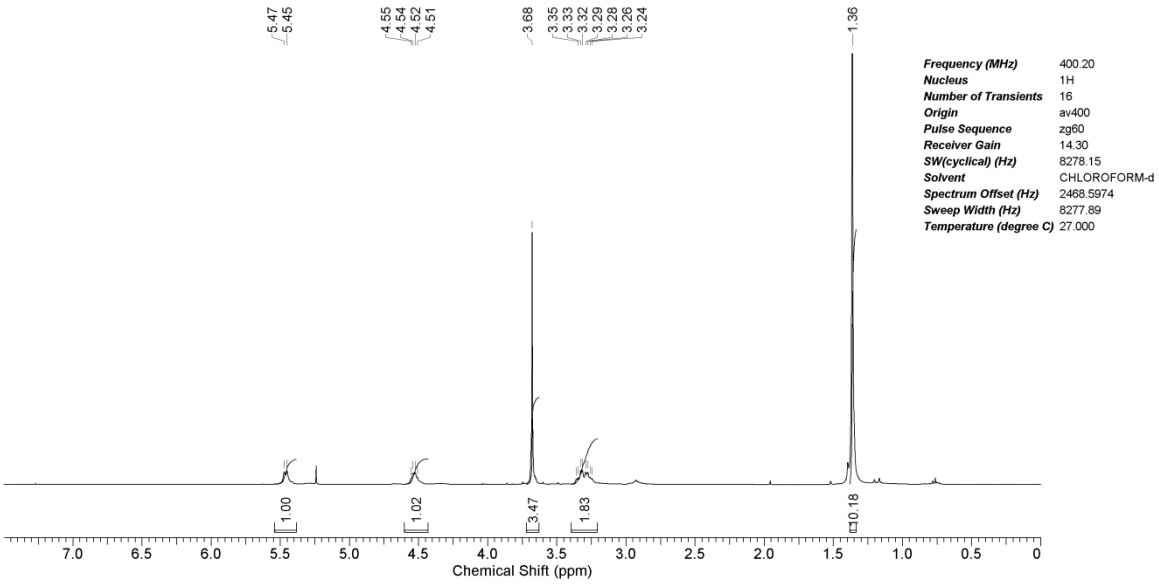

173.7
155.5
133.7
117.9
80.3
77.4
77.1
76.8
53.3
34.9
32.9
28.3
28.1
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000

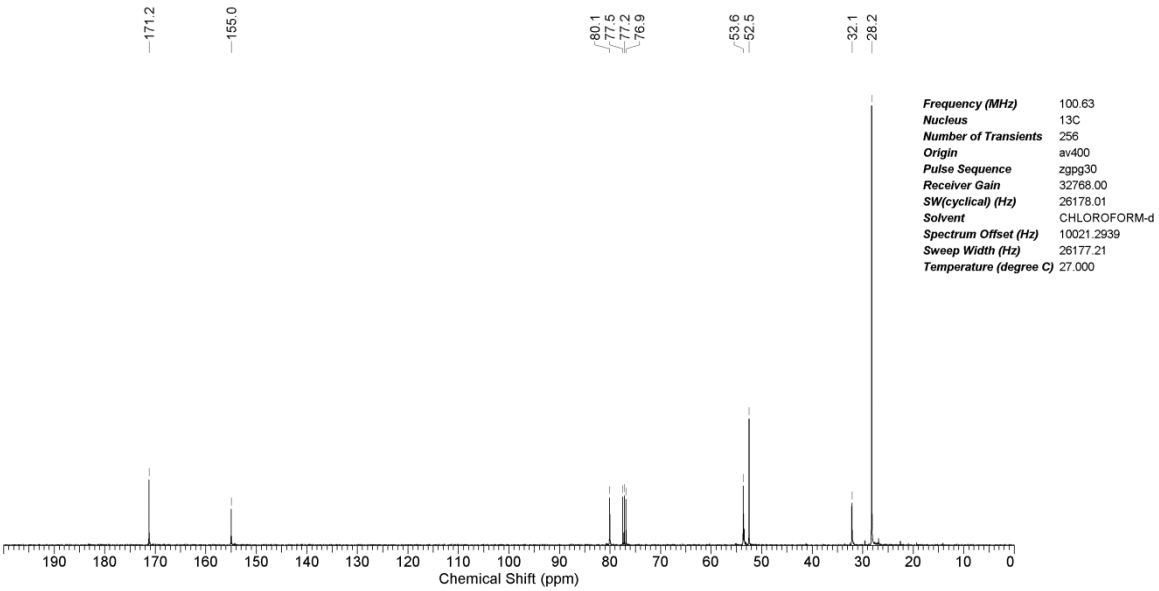

O
S
N
H
CONH2

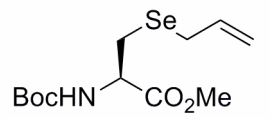

6.40
6.35
6.33
6.30
6.29
6.26
5.81
5.80
5.77
5.74
5.74
5.72
5.71
5.21
5.20
5.16
5.13
5.11
4.63
4.61
4.61
4.59
3.21
3.20
2.95
2.93
2.92
2.78
2.76
2.75
2.73
Frequency (MHz) 400.13
Nucleus 1H
Number of Transients 16
Origin dpx400
Pulse Sequence zg60
Receiver Gain 45.30
SW(cyclical) (Hz) 5592.84
Solvent METHANOL-d4
Spectrum Offset (Hz) 2000.6499
Sweep Width (Hz) 5592.67
Temperature (degree C) 27.000
1.96
1.98
2.07
1.97
2.11
1.00
1.04
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)

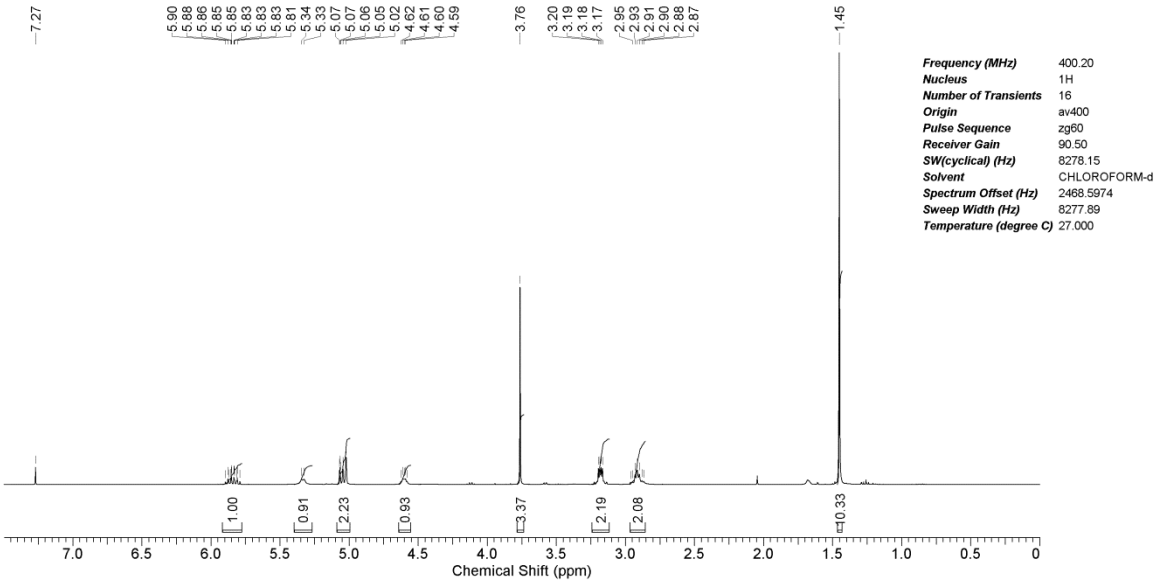

174.2
166.9
134.4
130.6
126.6
117.0
52.8
34.6
32.3
27.4
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent METHANOL-d4
Spectrum Offset (Hz) 10063.0566
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

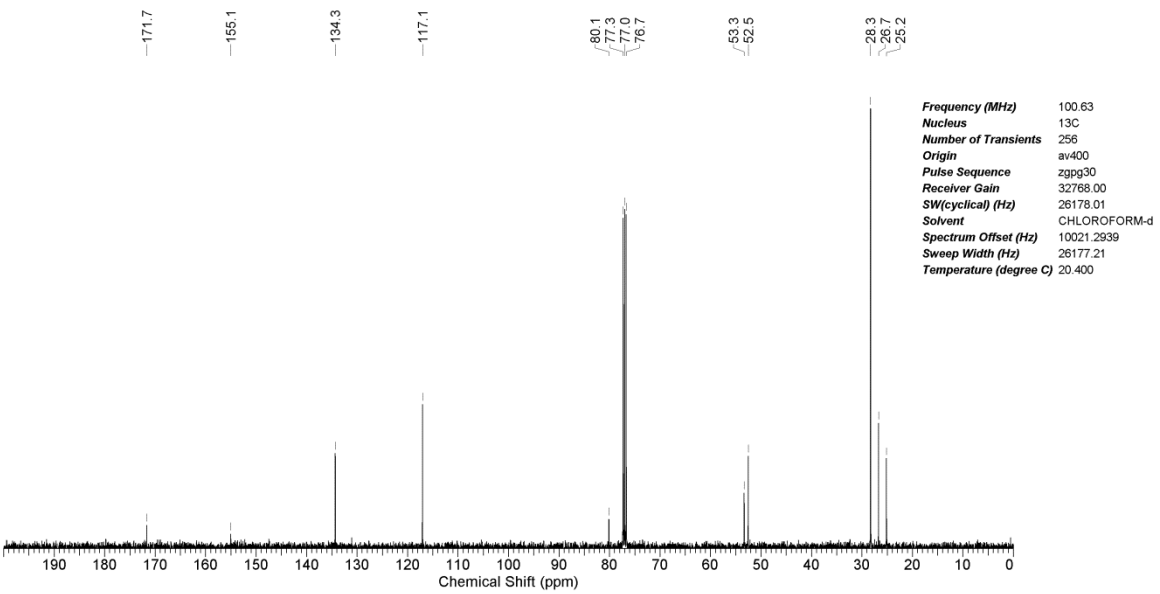

Se
BocHN
CO2Me
2

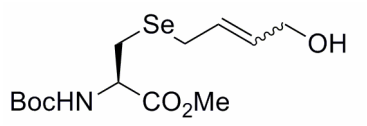

5.47
5.45
4.55
4.54
4.52
4.51
3.68
3.35
3.33
3.32
3.29
3.28
3.26
3.24
1.36
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 14.30
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000
1.00
1.02
3.47
1.83
10.18
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)

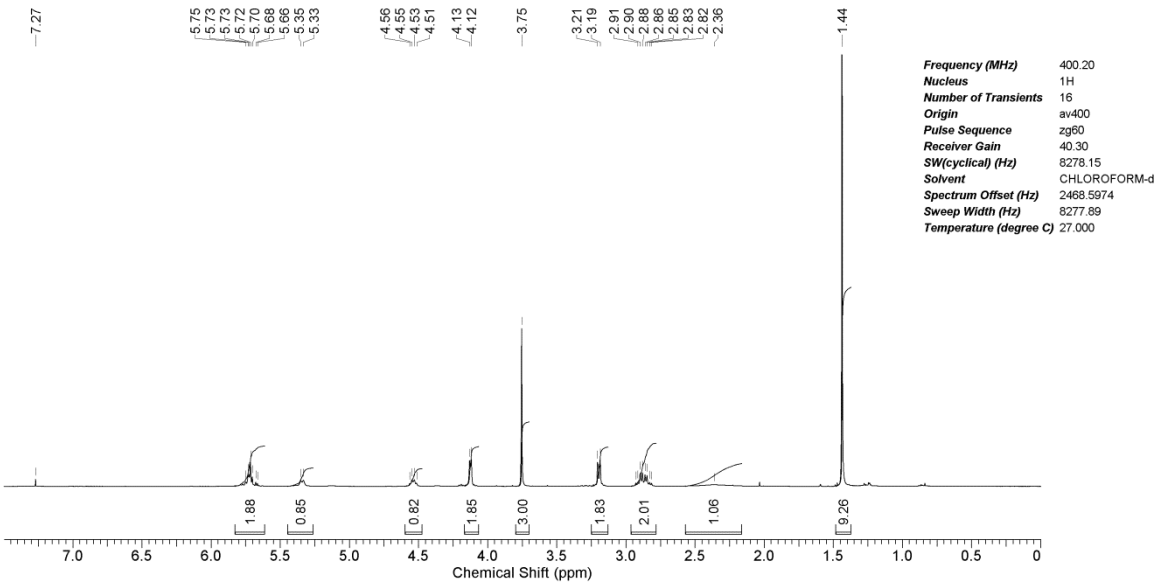

171.2
155.0
80.1
77.5
77.2
76.9
53.6
52.5
32.1
28.2
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

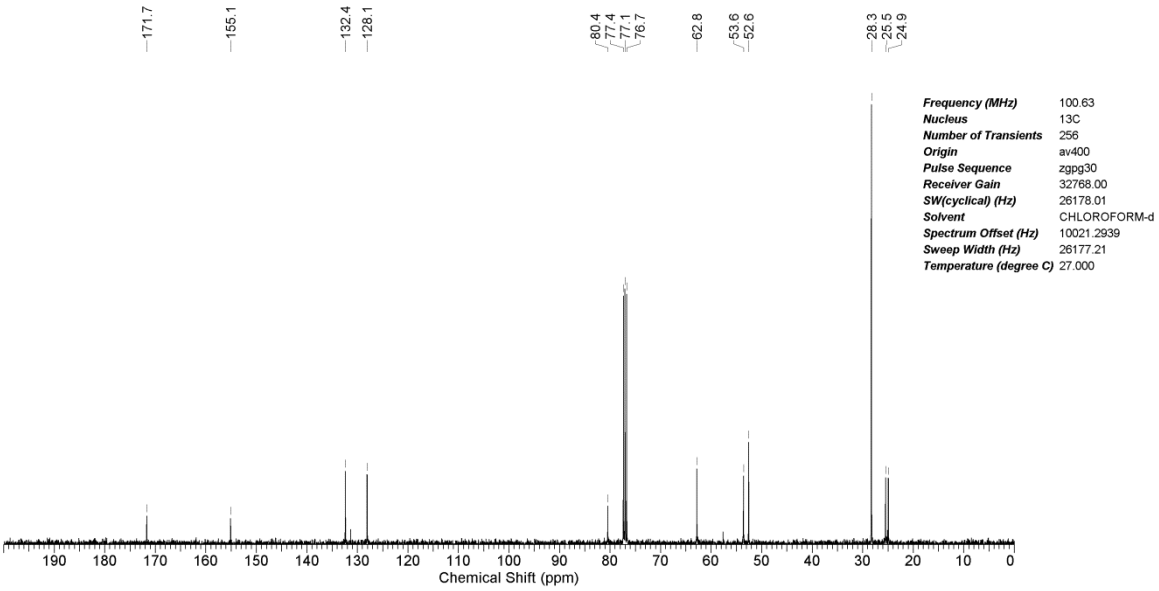

Se
BocHN
CO2Me

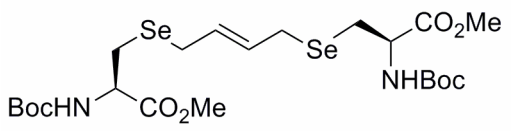

7.27
5.90
5.88
5.86
5.85
5.85
5.83
5.83
5.83
5.81
5.34
5.33
5.07
5.07
5.06
5.05
5.02
4.62
4.61
4.60
4.59
3.76
3.20
3.19
3.18
3.17
2.95
2.93
2.91
2.90
2.88
2.87
1.45
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 90.50
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000
1.00
0.91
2.23
0.93
3.37
2.19
2.08
10.33
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)

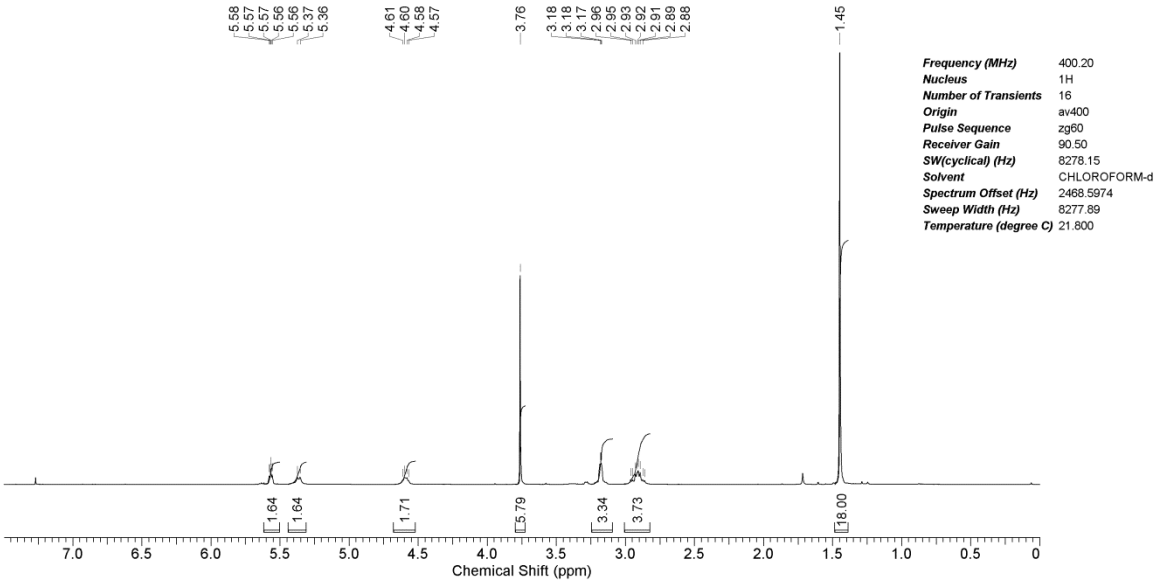

171.7
155.1
134.3
117.1
80.1
77.3
77.0
76.7
53.3
52.5
28.3
26.7
25.2
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 20.400
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

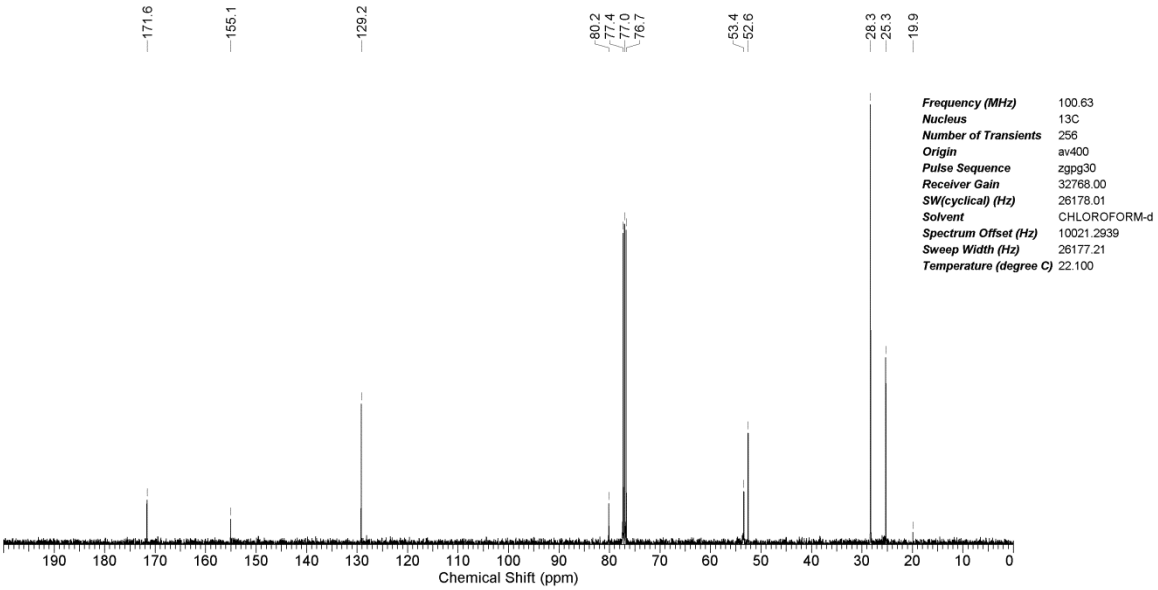

Se
OH
BocHN
CO2Me

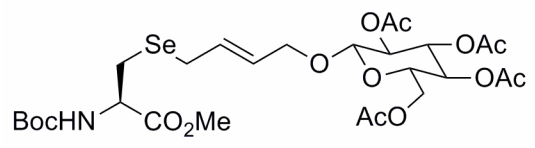

7.27
5.75
5.73
5.73
5.72
5.70
5.68
5.66
5.35
5.33
4.56
4.55
4.53
4.51
4.13
4.12
3.75
3.21
3.19
2.91
2.90
2.88
2.86
2.85
2.83
2.82
2.36
1.44
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 40.30
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 27.000
1.88
0.85
0.82
1.85
3.00
1.83
2.01
1.06
9.26
7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0
Chemical Shift (ppm)

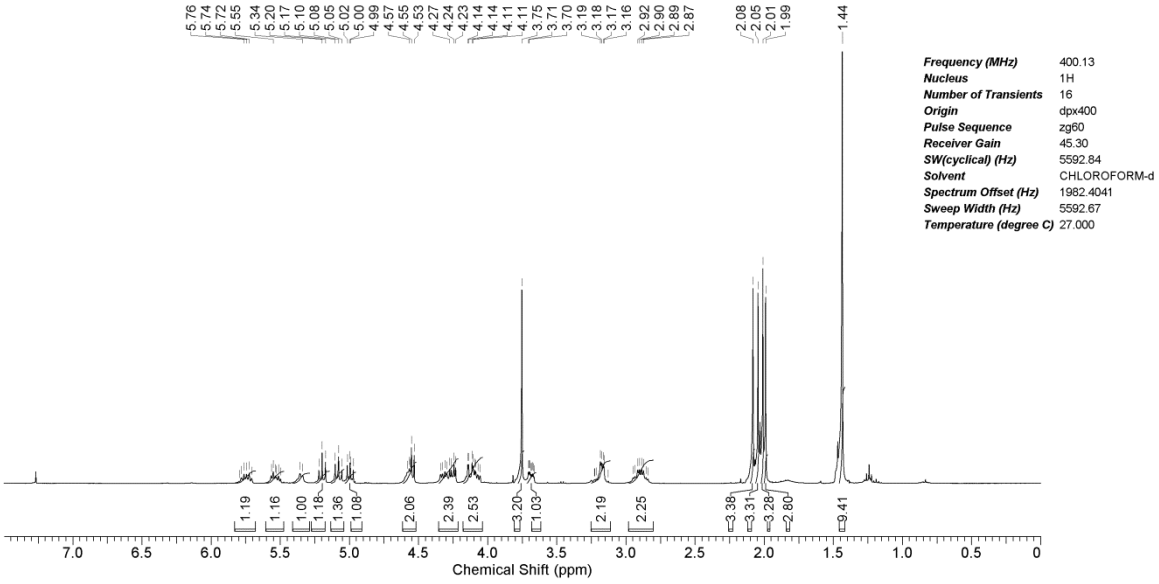

171.7
155.1
132.4
128.1
80.4
77.4
77.1
76.7
62.8
53.6
52.6
28.3
25.5
24.9
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000
190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
Chemical Shift (ppm)

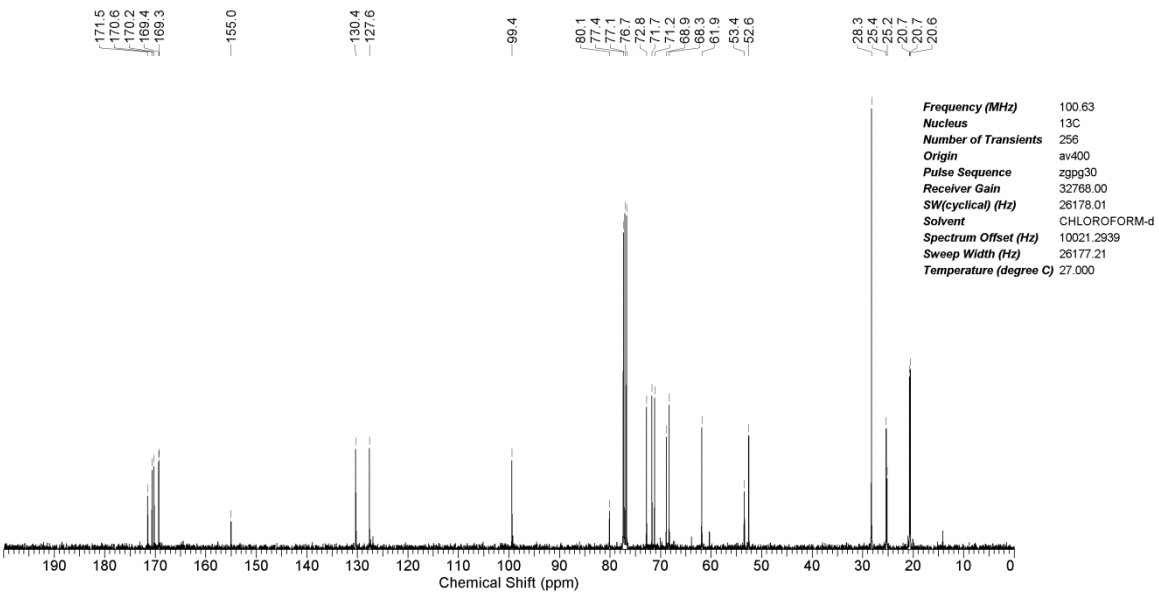

Se
Se
CO2Me
NHBoc
BocHN
CO2Me

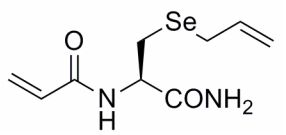

5.58
5.57
5.57
5.56
5.56
5.56
5.37
5.36
4.61
4.60
4.58
4.57
3.76
3.18
3.18
3.17
2.96
2.95
2.93
2.92
2.91
2.89
2.88
1.45
Frequency (MHz) 400.20
Nucleus 1H
Number of Transients 16
Origin av400
Pulse Sequence zg60
Receiver Gain 90.50
SW(cyclical) (Hz) 8278.15
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 2468.5974
Sweep Width (Hz) 8277.89
Temperature (degree C) 21.800
1.64
1.64
1.71
5.79
3.34
3.73
18.00
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)

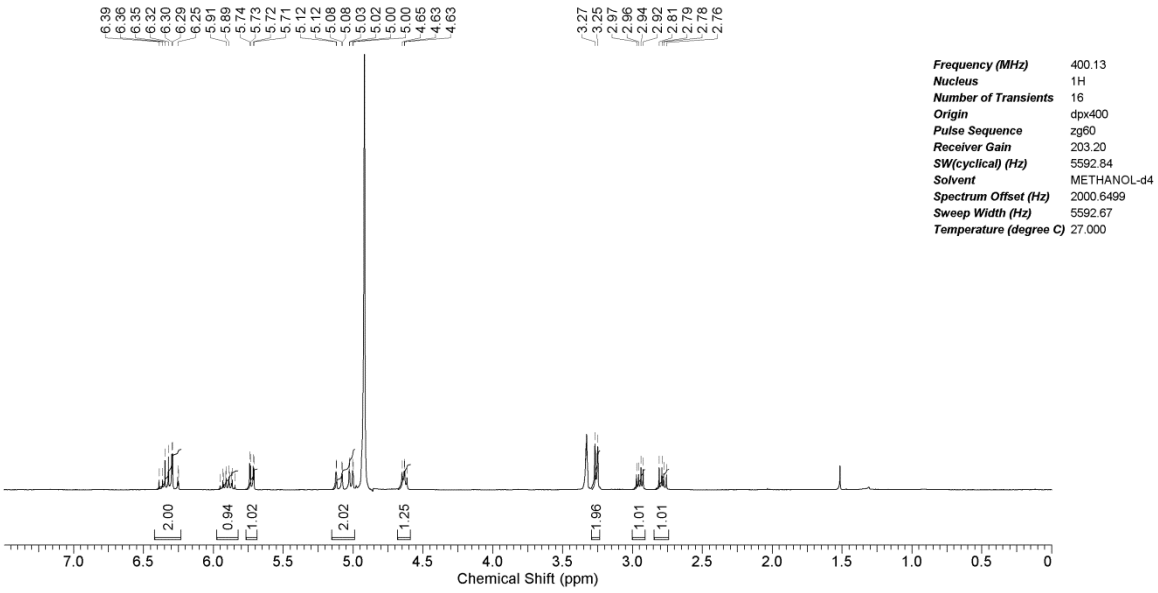

171.6
155.1
129.2
80.2
77.4
77.0
76.7
53.4
52.6
28.3
25.3
19.9
Frequency (MHz) 100.63
Nucleus 13C
Number of Transients 256
Origin av400
Pulse Sequence zgpg30
Receiver Gain 32768.00
SW(cyclical) (Hz) 26178.01
Solvent CHLOROFORM-d
Spectrum Offset (Hz) 10021.2939
Sweep Width (Hz) 26177.21
Temperature (degree C) 22.100
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

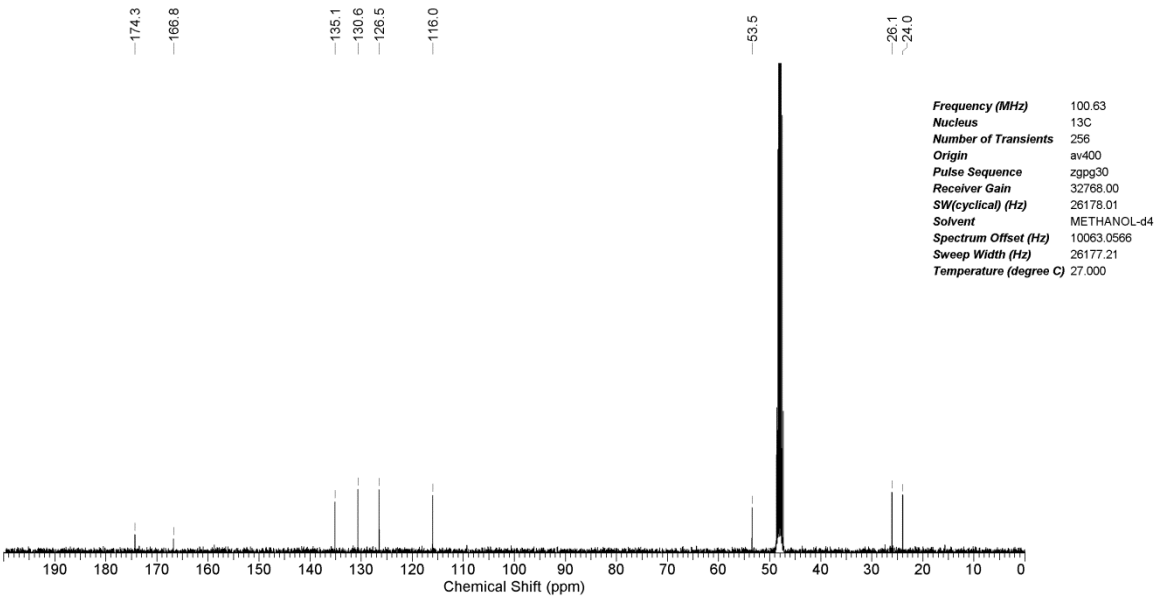

OAc
Se
OAc
O
O
OAc
BocHN
CO2Me
AcO


5.76
5.74
5.72
5.55
5.34
5.20
5.17
5.10
5.08
5.05
5.02
5.00
4.99
4.57
4.55
4.53
4.27
4.24
4.23
4.14
4.14
4.11
4.11
3.75
3.71
3.70
3.19
3.18
3.17
3.16
2.92
2.90
2.89
2.87
2.08
2.05
2.01
1.99
1.44
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 CHLOROFORM-d
Spectrum Offset (Hz) 1982.4041
Sweep Width (Hz) 5592.67
Temperature (degree C) 27.000
1.19
1.16
1.00
1.18
1.36
1.08
2.06
2.39
2.53
3.20
1.03
2.19
2.25
3.38
3.31
3.28
2.80
9.41
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)


171.5
170.6
170.2
169.4
169.3
155.0
130.4
127.6
99.4
80.1
77.4
77.1
76.7
72.8
71.7
71.2
68.9
68.3
61.9
53.4
52.6
28.3
25.4
25.2
20.7
20.7
20.6
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
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

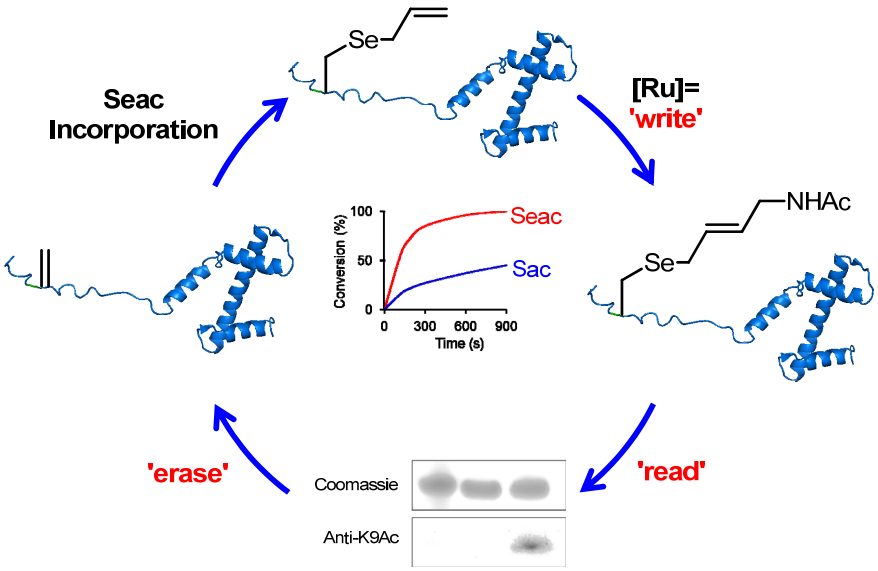

O
Se
N
H
CONH2


6.39
6.36
6.35
6.32
6.30
6.29
6.25
5.91
5.89
5.74
5.73
5.72
5.71
5.12
5.12
5.08
5.08
5.03
5.02
5.00
5.00
4.65
4.63
4.63
3.27
3.25
2.97
2.96
2.94
2.92
2.81
2.79
2.78
2.76
Frequency (MHz) 400.13
Nucleus 1H
Number of Transients 16
Origin dpx400
Pulse Sequence zg60
Receiver Gain 203.20
SW(cyclical) (Hz) 5592.84
Solvent METHANOL-d4
Spectrum Offset (Hz) 2000.6499
Sweep Width (Hz) 5592.67
Temperature (degree C) 27.000
2.00
0.94
1.02
2.02
1.25
1.96
1.01
1.01
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
Chemical Shift (ppm)


174.3
166.8
135.1
130.6
126.5
116.0
53.5
26.1
24.0
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 METHANOL-d4
Spectrum Offset (Hz) 10063.0566
Sweep Width (Hz) 26177.21
Temperature (degree C) 27.000
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)

Chapter 4

Rapid Cross Metathesis for Selective Protein Modifications via Chemical Access to *Se*-Allyl Selenocysteine in Proteins



4.1 Introduction

Developing novel strategies for site-specific protein modifications is an ongoing challenge in chemical biology.[1] As discussed in Chapter 1, allylic alcohols have been identified as reactive substrates in olefin metathesis for over a decade.[2] Through the preliminary work described in the introductory chapter (Section 1.6) and the related extended studies in Chapter 3, a generalized allylic heteroatom effect has begun to emerge in cross-metathesis (CM) enabling selective protein modifications.[3] Such putative beneficial effects of allylic chalcogens suggest a likely increase in the reactivity of higher chalcogens as Lewis bases that are more preferred, softer ligands for

ruthenium complexes. It was demonstrated in Section 3.2.3 that simple allyl selenides are indeed reactive substrates and enabled efficient cross metathesis with broader substrate scope and shorter reaction time.[3b] In this chapter, the installation into proteins the simplest amino acid residue containing an allyl selenide moiety, *Se*-Allyl-selenocysteine (Seac), was investigated.

4.2 Result and Discussion[†]

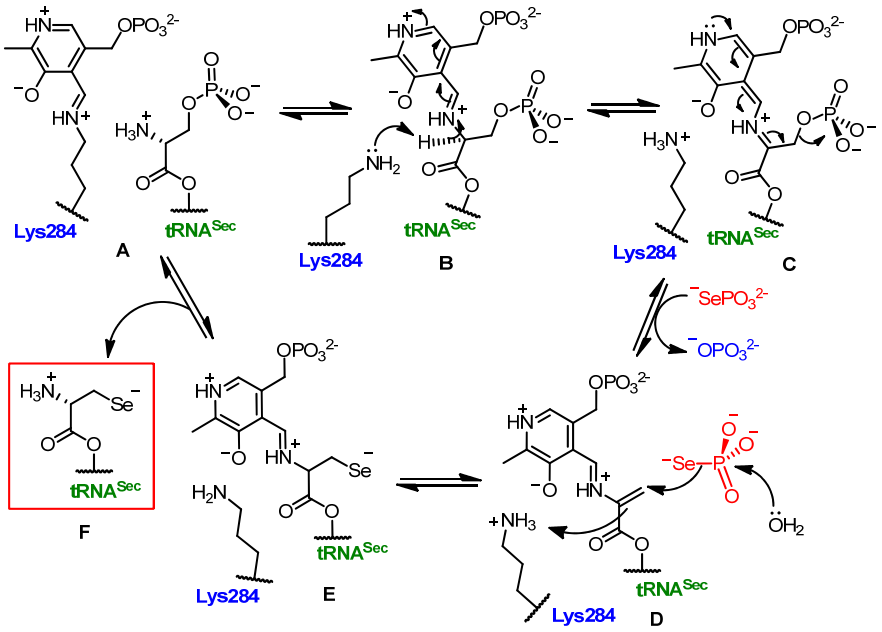
4.2.1 Synthesis of *Se*-allyl Selenocysteine (4.3)

Possible approaches to incorporating Seac into proteins were considered. These relied on installation of the Se heteroatom either co-(using Seac directly or from a precursor such as selenocysteine (Sec)) or post-translationally. Direct co-translational incorporation of Seac has not been demonstrated likely due to metabolism that occurs readily in cell, similar to what has been observed for selenomethionine and *S*-methyl-selenocysteine.[4] Despite the precedence of metabolic incorporation of selenocysteine into recombinant proteins in *E. coli*, this approach requires complex genetic manipulation of both DNA and RNA[5] that limits widespread utility. Notably, a recent study by Palioura *et al.* revealed a mechanism for selenocysteine formation via C–Se bond formation from a dehydroalanyl-tRNA^{Sec} intermediate (Scheme 4-1).[6]

Their study shows that the dehyroalanyl-tRNA^{Sec} is generated from elimination of phosphate in *O*-phosphoseryl- tRNA^{Sec} (Sep-tRNA^{Sec}) (**C**-**D**), and the subsequent attack by selenium from selenophosphate leads to the formation of selenocysteinyl moiety (**D**-**E**). Following the cleavage of pyridoxal phosphate (PLP) group by nearby lysine

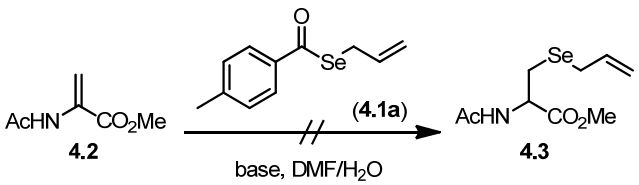
[†] This project was carried out as part of collaboration with Dr. Omar Boutureira, a former post-doctoral research worker in the group. Experiments that were obtained by Dr. Boutureira are indicated in the body of this chapter by reference to this footnote and his contributions are stated explicitly in the experimental sections of this chapter.

residue, the oxidized form of selenocysteinyl-tRNA^{Sec} is released (**E-F**). Inspired by nature, we therefore considered an alternative biomimetic route to incorporate Seac into proteins via 1-4, addition of allyl selenolate to dehydroalanine, which has recently emerged as a versatile chemical handle for proteins.[7]



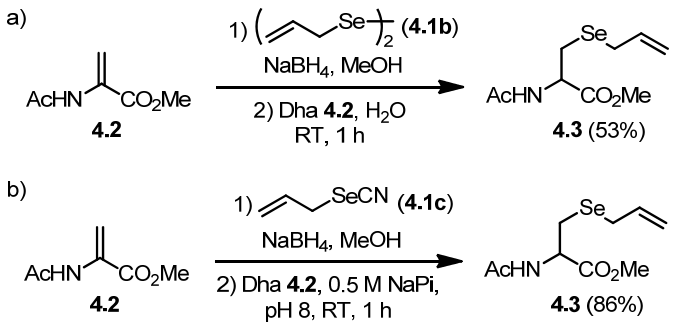
Scheme 4-1: The mechanism of selenocysteine formation in nature.

At the outset, the generation of allyl selenolate nucleophile was explored on model Dha **4.2**. *In situ* generation of allyl selenolate via cleavage of *Se*-allyl-4-methyl-benzoselenoate (**4.1a**) yielded no desired Seac product and only starting material was recovered (Scheme 4-2 and Experimental procedures).[†] Next, we considered access to allyl selenolate nucleophile via hydride reduction, which has been utilized in the synthesis of allyl selenol by Riague *et al.*[8] Allyl selenol was synthesized either via Bu_3SnH reduction of diallyl diselenide (**4.1b**) or $LiAlH_4$ reduction of allyl selenocyanate (**4.1c**) followed by acidification with succinic acid.[8]



Scheme 4-2: Attempted synthesis of Seac **4.3** via hydrolysis of **4.1a**.

The generation of the selenium nucleophile via reduction of diallyl diselenide was investigated first. Diallyl diselenide is unstable and exists in mixture with diallyl selenide.[8] Nevertheless, treatment of **4.1b** with stoichiometric amount of sodium borohydride in degassed methanol generated the corresponding allyl selenolate in solution. The subsequent addition of this solution to Dha **4.2** in water yielded 53% of the desired Seac product **4.3** (Scheme 4-3a).[†] Allyl selenocyanate (**4.1c**), a potentially more stable, alternative allyl selenolate source was investigated next. Pleasingly, due to the cleaner pre-reduction step, the desired addition product **4.3** was obtained with an improved yield of 86% after 1 hour at room temperature (Scheme 4-3b). Notably, although allyl selenolate is sensitive to acidic conditions (allyl selenol in solution quickly decomposes to **4.1b** in the absence of a radical inhibitor[8]), under buffered basic conditions the lifetime of allyl selenolate is apparently long enough to successfully react with Dha in aqueous media.

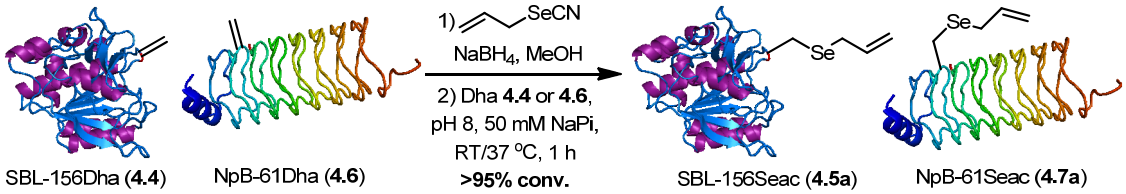


Scheme 4-3: Synthesis of Seac derivative **4.3** from Dha derivative **4.2**.

4.2.2 Installation of *Se*-allyl Selenocysteine (Seac) into proteins

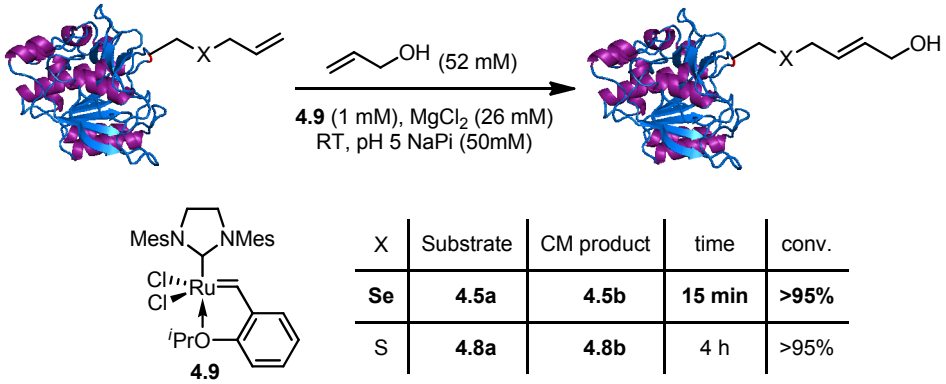
The promising results observed on the amino acid model motivated the installation of Seac via this biomimetic strategy in two differing protein structural motifs/folds: 3-layer-α/β-Rossman- fold protein subtilisin from *Bacillus lentus* (SBL) and all-β-helix protein 275-276 from *Nostoc punctiforme* (Npβ),[9] respectively. Accordingly, single Dha-containing protein substrates SBL-156Dha (**4.4**) and Npβ-61Dha (**4.6**) were

synthesized from corresponding cysteine mutants following reported bis-alkylative elimination procedures (see Experimental section).[7e] Use of **4.1a** or **4.1b** failed on proteins and led to only a mixture of products. However, allyl selenolate solution generated from **4.1c** allowed successful biomimetic conjugate addition (Scheme 4-4); LC-MS analysis of the reaction mixture revealed >95% conversion to the desired product SBL-156Seac (**4.5a**) after 1 hour at room temperature. Even addition to the more sterically demanding[10] 61 site in Npβ-61Dha under slightly elevated temperature (37 °C) afforded expected Seac-containing protein **4.7a** with >95% conversion after 2 hours.[†] Importantly, the modified SBL was shown to retain its peptidase activity (see Experimental section).[†]



Scheme 4-4: Biomimetic installation of Seac into proteins.

In order to assess the reactivity of Seac-tagged protein **4.5a**, standard model CM with allyl alcohol was carried out in the presence of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) under previously optimized reaction conditions. Consistent with previous results on the Seac-tagged protein substrate **3.23** (see Section 3.2.4),[3b] protein **4.5** undergoes rapid CM with allyl alcohol. The reaction was complete *after only 15 minutes* at room temperature (Scheme 4-5). In comparison, the sulfur-equivalent protein substrate (SBL-156Sac, **4.8a**) required 4 hours to reach completion under identical reaction conditions. Although Npβ-61Seac (**4.7a**) demonstrated the efficiency of the Se-nucleophile addition at hindered sites, the corresponding CM is rather sensitive to such a crowded environment and attempts resulted unfruitfully. Importantly, this result reinforces the use of an appropriate linker-extended Seac handle for hindered protein sites with conventional metathesis catalysts.[3b]



X	Substrate	CM product	time	conv.
Se	**4.5a**	**4.5b**	**15 min**	**>95%**
S	**4.8a**	**4.8b**	4 h	>95%

Scheme 4-5: Cross-Metathesis of SBL-156Seac (**4.5a**)/SBL-156Sac (**4.8a**) with allyl alcohol.

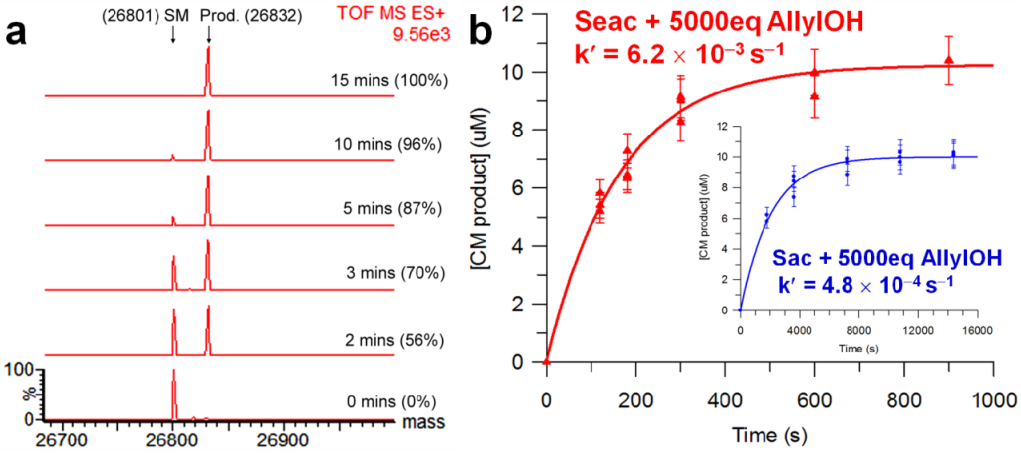
4.2.3 Kinetic Studies of Cross-Metathesis on Proteins

The remarkable CM rate displayed by SBL-156Seac inspired us to conduct a quantitative comparison with popular bioconjugation techniques such as Staudinger ligation,[11] azide-alkyne cycloaddition[12] and inverse-electron-demand Diels-Alder (IEDDA) reaction,[13] most of which have been studied only in small molecules ('off-protein'). Fox and Mehl have conducted *in vitro* and *in vivo* kinetic study of IEDDA reaction utilizing the fluorescent properties of tetrazine-encoded GFP (k = 880 and 330 $M^{-1}s^{-1}$, respectively).[14] The fluorescence of GFP is quenched by the tetrazine prior to reaction and "turned on" upon coupling with transcyclooctene. The reaction kinetics were then determined by fluorescence measurements.

To investigate the rate of CM on protein substrate, the ratio of starting material and product were monitored by LC-MS at several time points over the course of reaction under pseudo-first order conditions with respect to the protein (Scheme 4-6a). A small aliquot was taken from reaction at each time point and immediately desalted (quenched) with small amount of Sephadex® G-25 resin prior to LC-MS analysis. As no significant protein precipitation is observed during reaction, the percentage conversions can be converted directly to product concentrations.‡ The data was fitted to a non-linear, single

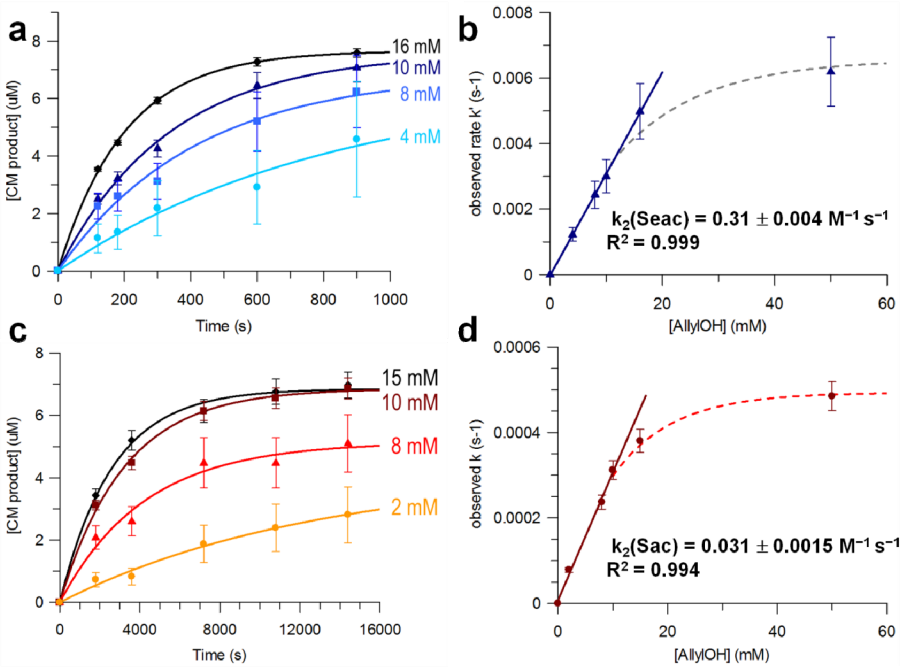
‡ For validation of ESI-MS in protein reaction monitoring in reaction see Section 3.5.

exponential, regression to obtain pseudo-first order rate constants (Scheme 4-6b).



Scheme 4-6: (a) LC-MS of the CM between SBL-156Seac and allyl alcohol over 15 minutes. (b) Changes in [**4.5b**] as a function of time. The inset shows the kinetics of the CM with SBL-156Sac; note the difference in time scale. Reaction conditions: [protein] = 0.01 mM, [**4.9**] = 1 mM, [allyl alcohol] = 52 mM, [$MgCl_2 \cdot 6H_2O$] = 26 mM, in 3:7 *t*BuOH/NaPi (50 mM, pH 5).

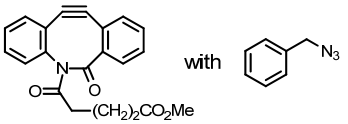
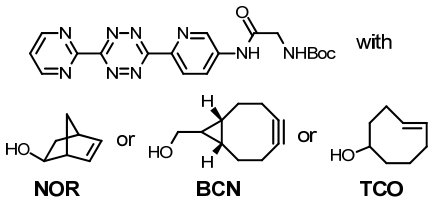
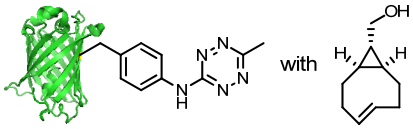
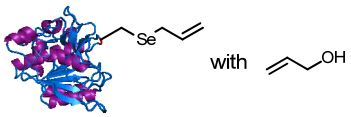
Determination of pseudo-first order rate constants at various allyl alcohol concentrations and subjecting the data to a linear fit revealed second order rate constant for Seac-mediated CM (Scheme 4-7). The on-protein second-order rate constant for the CM of allyl alcohol with Seac-tagged protein **4.5a** was determined to be 0.31 ± 0.004 M^{-1} s^{-1} (Scheme 4-7 a,b) whereas the value for on-protein Sac-mediated-CM in an identical protein context (site 156) was some 10-fold lower (0.031 ± 0.0015 M^{-1} s^{-1}) under the same conditions (Scheme 4-7 c,d). Interestingly, Scheme 4-7b indicated an initially linear increase in observed rate with respect to allyl alcohol concentration and gradually plateaued at concentrations >20 mM. This suggested that the catalyst plays a critical role under these conditions, presumably becoming 'saturated' at high concentration of allyl alcohol leading to a shift in rate-limited step (change in mechanism). A similar 'leveling off' was observed for SBL-156Sac but at lower concentration of allyl alcohol at 10 mM (Scheme 4-7d).



Scheme 4-7: Rate of CM varying allyl alcohol concentration with (a) SBL-156Seac or (c) SBL-156Sac. Determination of the second-order rate constant for CM between allyl alcohol and (b) SBL-156Seac or (d) SBL-156Sac.

When comparing the second-order rate constants with widely-used bioconjugation techniques in chemical biology, the on-protein rate of Seac-assisted CM was found to be about 150-folds faster than off-protein Staudinger ligation (k = 0.002 $M^{-1}s^{-1}$),[15] and comparable to faster variants of off-protein strain-promoted azide-alkyne cycloaddition (k = 0.002–0.9 $M^{-1}s^{-1}$)[16] and certain tetrazine-based cycloadditions (k > 0.15 $M^{-1}s^{-1}$) (Scheme 4-8).[17] However, selected examples of off-protein tetrazine-based reactions are still more rapid, highlighting the specific potential of that particular class of reaction.[17-18] It seems highly likely that local protein environment will have a strong effect upon on-protein reactions as compared with off-protein. The results presented in this chapter suggest that this retards rate by up to four orders of magnitude in some reaction systems. Consistent with the steric sensitivity of CM,[3b] an estimated off-protein rates for Seac-mediated CM from initial kinetic results of protein substrate suggests k > 1.8 $M^{-1}s^{-1}$ (obtained by comparing the on-protein and off-protein IEDDA rates from studies done by Mehl and Chin, respectively). Since the extent of rate retardation could

vary from protein to protein, these off-protein values are of less biological relevance unless accurate estimates can be made.

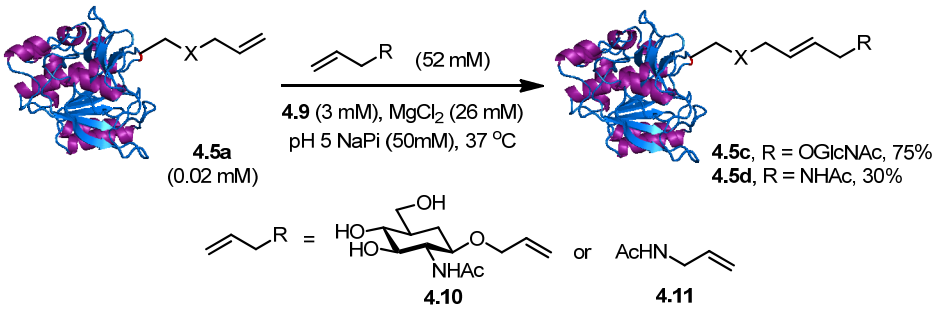
Reaction Type	Reaction Components	2nd order rate constant (M^{-1} s^{-1})
Staudinger ligation		0.002 (in CD_3CN)
Strain-promoted azide-alkyne cycloaddition		0.3 (in CD_3CN)
Inverse-electron demand Diels-Alder reaction (off-protein)		0.47[a]/437[b]/5235[c] (in MeOH/H_2O, 55:45)
Inverse-electron demand Diels-Alder reaction (on-protein)		880 (in PBS)
Seac-assisted cross-metathesis		0.3 (in BuOH/NaP_i buffer, 3:7)

Scheme 4-8: Comparison of rate constants of selected examples of bioconjugation reactions. Rate constants were measured with [a] norbornenol (NOR), [b] bicyclo[6.1.0]- non-4-yn-9-ylmethanol (BCN), [c] trans-cyclooctenol (TCO).

4.2.4 Cross-Metathesis on SBL-156Seac (4.5a)

Next, having established usefully enhanced on-protein rates, the substrate breadth of this accelerated reaction was tested. Model Seac-tagged protein, SBL-156Seac (**4.5a**), was subjected to CM with challenging metathesis substrate GlcNAc **4.10**, which is essentially unreactive in sulfur-relayed CM with SBL-156Sac (**4.8a**) (See Section 3.2.1). Pleasingly, the reaction reached 75% conversion after 1.5 hours at 37 °C (Scheme 4-9). Another biologically relevant substrate, *N*-allyl acetamide (**4.11**), was used in CM with protein **4.5a**. Acetamide **4.11** has been suggested to poison catalyst **4.7**, likely through formation of a stable six-member chelate with the ruthenium complex via coordination of

the carbonyl oxygen atom.[3b,19] Despite this poisoning, LC-MS analysis revealed 30% conversion to the desired CM product **4.13** after 20 minutes at 37 °C (Scheme 4-9). The reaction mixture appeared brown in color consistent with catalyst degradation and no further reaction could be observed (see Experimental procedures). The previously reported CM for GlcNAc **4.10** and acetamide **4.11** on SBL containing extended Seac both reached >90% conversions.[3b] The further enhanced CM conversions presumably was due to the reduced steric hindrance around the reactive alkene, provided by the chemical spacer, and could outcompete catalyst degradation in the case for *N*-allyl acetamide.

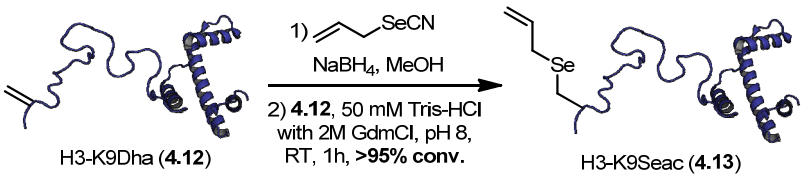


Scheme 4-9: Cross-metathesis of SBL-156Seac (**4.5a**) with substrate **4.10** and **4.11**.

4.2.5 Synthesis of H3-K9Seac (4.13)

After demonstrating the efficient CM on the model protein SBL-Seac, it was important to test whether Se-relayed CM could be translated to more functionally relevant protein systems. Histone proteins are key components of chromatin and their posttranslational modification (e.g. lysine methylation and acetylation) are critical regulators of chromatin structure and function.[20] First of all, Seac must be installed on histone H3. Following the synthesis of Dha on H3 using procedures established by Chalker *et al.*,[7f] the subsequent allyl selenolate addition using the protocols for protein **4.4** only led to protein precipitation. The presence of 10% methanol in reaction mixture is apparently not compatible with histone proteins. A range of denaturing conditions was screened for the conjugate addition. Finally, by carrying out the reaction in 2 M guanidinium chloride

(GdmCl) (with 50 mM Tris, pH 8), allyl selenolate addition to Dha **4.12** proceeded successfully to yield the desired Seac product **4.13** after 1 hour at room temperature (Scheme 4-10).

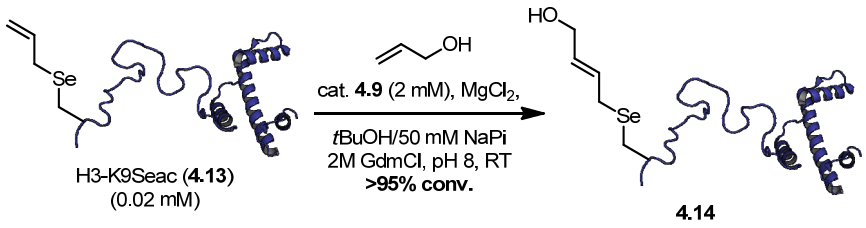


Scheme 4-10: Synthesis of H3-K9Seac (**4.13**).

4.2.6 Cross-Metathesis on H3-K9Seac (4.13)

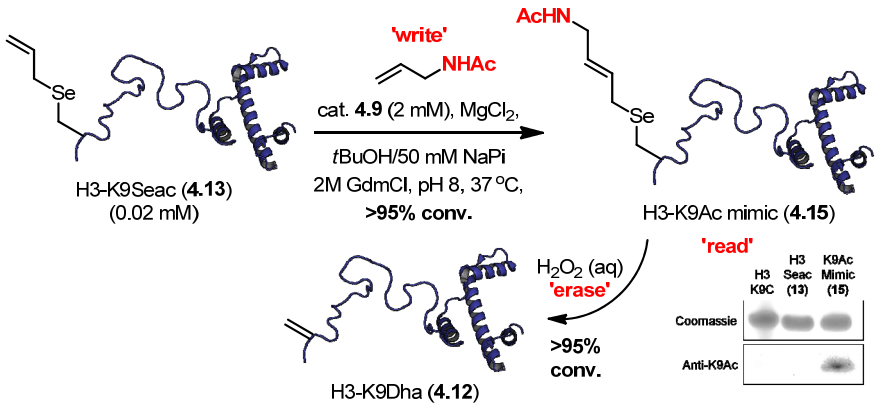
With H3-K9Seac (**4.13**) readily in hand, model CM with allyl alcohol was carried out. The intolerance of organic solvent of H3 was an immediate concern since the CM reaction requires 30% of *t*BuOH to aid solubility of catalyst **4.9** in water. However, in the presence of 2 M GdmCl, no protein precipitation was observed but protein could not be detected by MS despite a clear band on SDS-PAGE gel. In addition, the fact that the brown color (indicating presence of metal catalyst) from the reaction mixture retained in the Vivaspin® during protein purification implied metal coordination to the protein could be interfering with MS detection. Metal coordination to proteins is not uncommon,[21] and such phenomenon is likely to be even more pronounced when protein is in an unfolded state exposing higher number of binding sites.

These observations suggest that ruthenium scavenging is required prior to detection by MS. 3-mercaptopropionic acid (3-MP) was considered as a ruthenium scavenger. 3-MP has been previously shown to remove residual ruthenium catalyst in RCM reactions[22] and palladium bounded to protein surface in Suzuki couplings.[21b] The used of 3-MP as Ru-scavenger allowed monitoring of CM reactions on H3 by MS. Pleasingly, this revealed CM of H3-K9Seac (**4.13**) with allyl alcohol with near-full conversion of the CM-modified H3 (**4.14**) (Scheme 4-11a).



Scheme 4-11: Cross-Metathesis of H3-K9Seac (**4.13**) with allyl alcohol.

The ability to modify position 9 of H3 suggested a strategy for chemically-enabled 'write-read-erase' sequence that is observed in (enzyme-mediated) histone epigenetics.[20] As a 'write' step, H3-K9Seac was subjected to CM with the more challenging *N*-allyl acetamide CM substrate and complete conversion was observed (Scheme 4-11a). The higher conversion observed in H3-K9Seac compared with SBL-156Seac presumably reflect the more exposed reaction site under denaturing conditions, highlighting once again the sensitivity of CM to sterics. As a 'read' step, the CM- modified H3 (**4.15**) was analyzed with antibody raised to *N*-acetyl lysine at site 9 of H3 in a western blot assay. The antibody was found to successfully recognize the CM-installed modification at site 9 as a K9Ac PTM mimic (Scheme 4-12). Finally, the K9Ac mimic was 'erased' using Cope-type elimination via the generation of labile selenoxide using mild peroxide oxidation to re-generate H3-K9Dha (**4.12**). Notably, over-oxidation at residues such as methionine and tyrosine was observed under the reaction conditions employed (see Experimental procedure).



Scheme 4-12: Chemically-enabled 'write-read-erase' sequence: cross-metathesis of H3-K9Seac (**4.13**) with acetamide **4.11**, western blot analysis and oxidative elimination of **4.15**.

4.3 Conclusions and Outlook

In conclusion, the efficient chemical incorporation of Seac into proteins was demonstrated. This has facilitated rapid cross-metathesis for protein modifications; determination of on-protein rate constants shows that these outstrip or are comparable to many of the so-called 'click' reactions in chemical biology. Using allyl selenide- (e.g. Seac) containing proteins, broader substrate scope was accessed in Se-relayed CM than was possible with allyl sulfides. Direct access to Seac in proteins opens up opportunities to create potential PTM mimics via CM as demonstrated in this chapter: a K9Ac mimic of H3 was successfully installed and was recognized and removed (write-read-erase) in histone H3. It would be interesting to see whether re-writing the "erased" histone is chemically possible since the enzyme-mediated acetylation in nature is reversible.

This synthetic manipulation of biology also highlights opportunities for new metathesis catalysts in which the use of CM on proteins or cellular system may be enabled by tuned solubilities, compatibilities and permeabilities. Finally, this work also provided an alternative route to a protected (allylated) Sec in proteins (that could be revealed using, e.g., Pd(0)),[23] and also further motivates developments in genetic incorporation of allyl selenide containing amino acids into proteins in order to access more generally the wider applications in chemical biology that may be offered by rapid allyl selenide-mediated cross-metathesis.

4.4 References

(1) (a) Sletten, E. M.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 6974-6998. (b) Chalker, J. M.; Bernardes, G. J. L.; Davis, B. G. *Acc. Chem. Res.* **2011**, *44*, 730-741. (c) Takaoka, Y.; Ojida, A.; Hamachi, I. *Angew. Chem. Int. Ed.* **2013**, 4088-4106.

(2) (a) Hoye, T. R.; Zhao, H. *Org. Lett.* **1999**, *1*, 1123-1125. (b) Lin, Y. A.; Davis, B. G. *Beilstein J. Org. Chem.* **2010**, *6*, 1219-1228.

(3) (a) Lin, Y. A.; Chalker, J. M.; Davis, B. G. *ChemBioChem* **2009**, *10*, 959-969. (b) Lin, Y. A.; Chalker, J. M.; Davis, B. G. *J. Am. Chem. Soc.* **2010**, *132*, 16805-16811. (c) Kirshenbaum, K.; Arora, P. S. *Nat. Chem. Biol.* **2008**, *4*, 527-528. (d) Ai, H.-W.; Shen, W.; Brustad, E.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 935-937.

(4) (a) Stolz, J. F.; Basu, P.; Santini, J. M.; Oremland, R. S. *Annu. Rev. Microbiol.* **2006**, *60*, 107-130. (b) Foster, S. J.; Kraus, R. J.; Ganther, H. E. *Arch. Biochem. Biophys.* **1986**, *251*, 77-86. (c) Ganther, H. E. *Carcinogenesis* **1999**, *20*, 1657-1666.

(5) (a) Arnér, E. S. J.; Sarioglu, H.; Lottspeich, F.; Holmgren, A.; Böck, A. *J. Mol. Biol.* **1999**, *292*, 1003-1016. (b) Aldag, C.; Bröcker, M. J.; Hohn, M. J.; Prat, L.; Hammond, G.; Plummer, A.; Söll, D. *Angew. Chem. Int. Ed.* **2013**, *52*, 1441-1445.

(6) Palioura, S.; Sherrer, R. L.; Steitz, T. A.; Soell, D.; Simonovic, M. *Science* **2009**, *325*, 321-325.

(7) (a) Levengood, M. R.; van der Donk, W. A. *Nat. Protocols* **2007**, *1*, 3001-3010. (b) Wang, J.; Schiller, S. M.; Schultz, P. G. *Angew. Chem. Int. Ed.* **2007**, *46*, 6849-6851. (c) Bernardes, G. J. L.; Chalker, J. M.; Errey, J. C.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 5052-5053. (d) Chalker, J. M.; Davis, B. G. *Curr. Opin. Chem. Biol.* **2010**, *14*, 781-789. (e) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. *Chem. Sci.* **2011**, *2*, 1666-1676. (f) Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 1835-1839.

(8) Riague, E. H.; Guillemin, J.-C. *Organometallics* **2001**, *21*, 68-73.

(9) Vetting, M. W.; Hegde, S. S.; Hazleton, K. Z.; Blanchard, J. S. *Protein Sci.* **2007**, *16*, 755-760.

(10) Fernandez-Gonzalez, M.; Boutureira, O.; Bernardes, G. J. L.; Chalker, J. M.; Young, M. A.; Errey, J. C.; Davis, B. G. *Chem. Sci.* **2010**, *1*, 709-715.

(11) (a) Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007-2010. (b) Köhn, M.; Breinbauer, R. *Angew. Chem. Int. Ed.* **2004**, *43*, 3106-3116.

(12) (a) Hong, V.; Steinmetz, N. F.; Manchester, M.; Finn, M. G. *Bioconjugate Chem.* **2010**, *21*, 1912-1916. (b) van Hest, J. C. M.; van Delft, F. L. *ChemBioChem* **2011**, *12*, 1309-1312.

(13) (a) Blackman, M. L.; Royzen, M.; Fox, J. M. *J. Am. Chem. Soc.* **2008**, *130*, 13518-13519. (b) Devaraj, N. K.; Weissleder, R.; Hilderbrand, S. A. *Bioconjugate Chem.* **2008**, *19*, 2297-2299. (c) Schoch, J.; Wiessler, M.; Jäschke, A. *J. Am. Chem. Soc.* **2010**, *132*, 8846-8847. (d) Devaraj, N. K.; Weissleder, R. *Acc. Chem. Res.* **2011**, *44*, 816-827.

(14) Seitchik, J. L.; Peeler, J. C.; Taylor, M. T.; Blackman, M. L.; Rhoads, T. W.; Cooley, R. B.; Refakis, C.; Fox, J. M.; Mehl, R. A. *J. Am. Chem. Soc.* **2012**, *134*, 2898-2901.

(15) Lin, F. L.; Hoyt, H. M.; van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 2686-2695.

(16) (a) Debets, M. F.; van Berkel, S. S.; Schoffelen, S.; Rutjes, F. P. J. T.; van Hest, J. C. M.; van Delft, F. L. *Chem. Commun.* **2010**, *46*, 97-99. (b) Jewett, J. C.; Sletten, E. M.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 3688-3690. (c) Debets, M. F.; van Berkel, S. S.; Dommerholt, J.; Dirks, A. J.; Rutjes, F. P. J. T.; van Delft, F. L. *Acc. Chem. Res.* **2011**, *44*, 805-815.

(17) Lang, K.; Davis, L.; Torres-Kolbus, J.; Chou, C.; Deiters, A.; Chin, J. W. *Nat.Chem.* **2012**, *4*, 298-304.

(18) (a) Lang, K.; Davis, L.; Wallace, S.; Mahesh, M.; Cox, D. J.; Blackman, M. L.; Fox, J. M.; Chin, J. W. *J. Am. Chem. Soc.* **2012**, *134*, 10317-10320. (b) Plass, T.; Milles, S.; Koehler, C.; Szymański, J.; Mueller, R.; Wießler, M.; Schultz, C.; Lemke, E. A. *Angew. Chem. Int. Ed.* **2012**, *51*, 4166-4170.

(19) (a) Fürstner, A.; Langemann, K. *J. Am. Chem. Soc.* **1997**, *119*, 9130-9136. (b) Fürstner, A.; Seidel, G.; Kindler, N. *Tetrahedron* **1999**, *55*, 8215-8230.

(20) (a) Bannister, A. J.; Kouzarides, T. *Cell Res.* **2011**, *21*, 381-395. (b) Kouzarides, T. *Cell* **2007**, *128*, 693-705.

(21) (a) Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642-9643. (b) Spicer, C. D.; Davis, B. G. *Chem. Commun.* **2011**, *47*, 1698-1700.

(22) Macdonald, J. E.; Veinot, J. G. C. *Langmuir* **2008**, *24*, 7169-7177.

(23) Shimizu, I. In *Handbook of Organopalladium Chemistry for Organic Synthesis*; John Wiley & Sons, Inc.: 2003, p 1981-1994.

4.5. Experimental Procedures

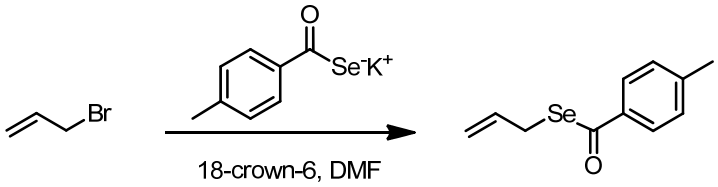
General Considerations

See Chapter 2 Experimental Procedures for general considerations in synthesis. See Chapter 3 Experimental for a validation of using ESI-MS to follow reaction conversions on protein substrates. Compound **4.1a**, **4.1b** and **4.7a** were synthesized by Dr. Omar Boutureira. The preparation of these compounds is included here for reference.

Protein Mass Spectrometry. Liquid chromatography-mass spectrometry (LC-MS) was performed on a Waters Micromass LCT Premier XE (ESI-TOF-MS) coupled to a Waters 1525 Binary HPLC pump using Chromolith FastGradient RP-18 2×50 mm column from Merck. Water (solvent A) and acetonitrile (solvent B), each containing 0.1% formic acid, were used as the mobile phase. The gradient was programmed as in shown in the table below. The electrospray source was operated with a capillary voltage of 3.2 kV and a cone voltage of 25 V for all the protein used in this report. Nitrogen was used as the nebulizer and desolvation gas at a total flow of 700 L/hr. Total mass spectra were reconstructed from the ion series using the MaxEnt algorithm preinstalled on MassLynx software (v. 4.1 from Waters) according to manufacturer's instructions.

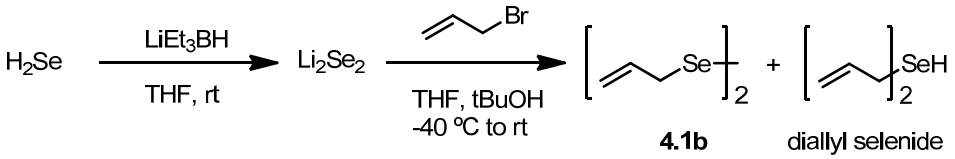
Time(min)	Flow (mL)	%A	%B
0	0.4	95	5
1	0.4	95	5
5	0.4	0	100
8	1.0	0	100
8.1	1.0	95	5
9.5	1.0	95	5

***Se*-Allyl 4-methylbenzoselenoate (4.1a)**



Potassium 4-methylselenobenzoate[4.1] (1.98 g, 8.36 mmol) was added to a mixture of allyl bromide (362 μL, 4.18 mmol) and 18-crown-6 (2.21 g, 8.36 mmol) in dry and degassed DMF (15 mL) at room temperature. The reaction mixture was stirred at the same temperature for 15 minutes. The crude was then diluted with EtOAc and washed with water. The combined organic layers were dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The residue was purified by column chromatography (from petrol to 1:5 EtOAc:petrol) to afford **4.1a** (686.4 mg, 70%) as a yellow liquid. R_f (1:5 EtOAc:petrol): 0.60. ^{1}H NMR ($CDCl_3$, 400 MHz) δ_H = 7.80 (2H, d, *J* = 8.3 Hz, Ar), 7.25 (2H, d, *J* = 8.3 Hz Ar), 5.99 (1H, m, CH=), 5.28 (1H, dd, *J* = 16.9, 1.3 Hz, =C<u>H</u>H), 5.06 (1H, dd, *J* = 9.9, 0.8 Hz, =CH<u>H</u>), 3.75 (2H, d, *J* = 7.6 Hz, CH_2), 2.40 (3H, s, CH_3, Ar). ^{13}C NMR ($CDCl_3$, 100 MHz) δ_C = 193.7 (C=O), 144.6, 136.3 (C, Ar), 134.6 (CH=), 129.4, 127.3 (CH, Ar), 117.3 (=CH_2), 27.7 (CH_2), 21.7 (CH_3, Ar). HRMS (FI+) for (M^+) $C_{11}H_{12}OSe$ (*m/z*): calc. 240.0054; found 240.0050. Spectroscopic data was identical to that previously reported.[4.2]

Diallyl diselenide (4.1b) [4.3]

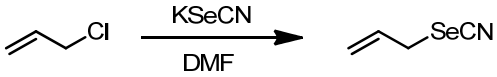


Selenium powder (2 g, 25.3 mmol) was added portionwise to 1M $LiEt_3BH$ in THF (25 mL, 25.3 mmol) at room temperature. Gas evolution occurred, and the suspension turned dark brown-red. The Li_2Se_2 thus formed was stirred at the same temperature for 30 minutes. Allyl bromide (2.2 mL, 25.3 mmol) in THF (33 mL) was added dropwise to this solution suspended in *t*BuOH (3.3 mL) and THF (63 mL) at –40 ºC. The solution was then warmed to room temperature and stirred for 8 h. The reaction mixture was

[4.1] Kawai, Y.; Ando, H.; Ozeki, H.; Koketsu, M.; Ishihara, H. *Org. Lett.* **2005**, *7*, 4653–4656.
[4.2] Garud, D. R.; Garud, D. D.; Koketsu, M. *Org. Biomol. Chem.* **2009**, *7*, 2591–2598.
[4.3] Riague, E. H.; Guillemin, J.-C. *Organometallics* **2002**, *21*, 68–73.

diluted with Et_2O and washed with water. The combined organic layers were dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford 3 g of an inseparable 1:4 mixture (by ^1H NMR) of diallyl diselenide **4.1b** (27%) and diallyl selenide (53%) as an orange syrup. Used in the next step without further purification. Data for **4.1b**: R_f (1:4 EtOAc:petrol): 0.73. ^1H NMR ($CDCl_3$, 400 MHz) δ_H = 5.86 (2H, m, CH=), 5.12–5.08 (4H, m, $=CH_2$), 3.54 (4H, d, J = 7.9 Hz, CH_2). ^{13}C NMR ($CDCl_3$, 100 MHz) δ_C = 134.7 (CH=), 117.1 ($=CH_2$), 31.9 (CH_2). Data for diallyl selenide: R_f (1:4 EtOAc:petrol): 0.73. ^1H NMR ($CDCl_3$, 400 MHz) δ_H = 5.86 (2H, m, CH=), 5.03–4.99 (4H, m, $=CH_2$), 3.14 (2H, d, J = 7.7 Hz, CH_2). ^{13}C NMR ($CDCl_3$, 100 MHz) δ_C = 134.7 (CH=), 117.3 ($=CH_2$), 25.3 (CH_2). Spectroscopic data was identical to that previously reported.[4.4]

Allyl selenocyanate (4.1c)[4.5]



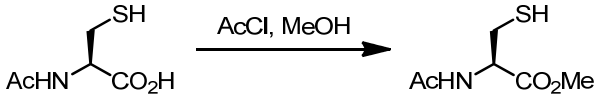
KSeCN (2.17 g, 15.06 mmol) was added to a 100 mL round bottom flask and dissolved in DMF (15 mL). The solution was placed under an atmosphere of nitrogen and allyl chloride (4.0 mL, 49.08 mmol) was added slowly to the stirred solution. The reaction was stirred for 1 hour at room temperature and then diluted with Et_2O (150 mL) and washed sequentially with H_2O (2 × 150 mL) and brine (150 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure. The product was isolated as a pale yellow liquid and was sufficiently pure to use in subsequent manipulations (1.90 g, 86%). This material has a sharp, lingering odor and should be used only in a well-ventilated fume hood. IR (υ_{max}, film): 2151, 1633, 1432, 1403, 1197, 987, 928, 851, 686. ^1H NMR (400 MHz, $CDCl_3$): δ_H = 3.65 (2H, dt, J = 7.2, 1.0,

4.4 Ouchi, A.; Liu, S.; Li, Z.; Kumar, A.; Suzuki, T.; Hyugano, T.; Kitahara. *J. Org. Chem.* **2007**, *72*, 8700–8706.

4.5 Chalker, J. M.; Lin, Y. A.; Boutureira, O.; Davis, B. G. *Chem. Commun.* **2009**, 3714-3716.

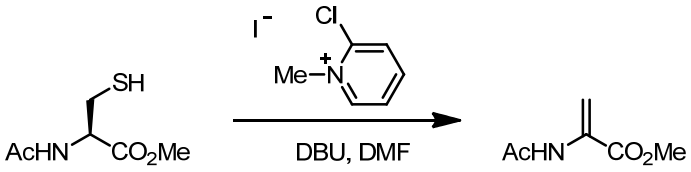
CH_2SeCN), 5.23 (1H, dd, J = 1.0, 9.9, CHH=CH), 5.32 (1H, dq, J = 1.0, 16.7, CHH=CH), 6.00 (1H, m, H_2C=CH). ^{13}C NMR (50 MHz, CDCl3): δ_C = 31.3 (CH_2SeCN), 101.4 (C≡N), 120.6 ($CH=CH_2$), 131.6 ($CH=CH_2$).

_N_-Acetyl-L-cysteine Methyl Ester (AcCysOMe)[4.6]



Acetyl chloride (43 mL, 0.6 mol) was added dropwise to anhydrous MeOH (183 mL) at 0 ºC under argon atmosphere. After 10 minutes at the same temperature, *N*-acetyl cysteine (20 g, 0.12 mol) was added as a solid in one portion. The reaction was gradually warmed up to room temperature and stirred for 5 hours. The solvent was removed under reduced pressure and the crude diluted with EtOAc (400 mL) and washed with saturated aqueous $NaHCO_3$ (400 mL), water (400 mL) and brine (400 mL). The combined organic layers were dried over $MgSO_4$, filtered, and concentrated under reduced pressure to afford the desired methyl ester AcCysOMe (16 g, 75%) as a white solid. Used in the next step without further purification. R_f (9:1 EtOAc/MeOH) = 0.52. ^1H NMR ($CDCl_3$, 200 MHz): δ_H =: 1.35 (1H, t, J = 9.0, SH), 2.07 (3H, s, $COCH_3$), 3.01 (2H, dd, J = 9.0, 3.9, CH_2S), 3.79 (3H, s, 3H, CO_2CH_3), 4.89 (m, 1H, H_α), 6.44 (1H, br.s, NH). LRMS (ES^-) for (M–H) $C_6H_{10}NO_3S$ (*m/z*): calc. 176.0; found 176.1.

Methyl 2-Acetoamidoacrylate (AcDhaOMe, 4.2)[4.7]



1,8-Diazabicycloundec-7-ene (DBU) (2.6 mL, 16.9 mmol) was added dropwise to a dispersion of AcCysOMe (1 g, 5.6 mmol) and 2-chloro-1-methylpyridinium iodide

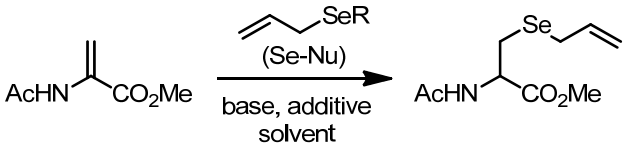
4.6 Lin, Y. A.; Chalker, J. M.; Floyd, N.; Bernardes, G. J. L.; Davis, B. G. *J. Am. Chem. Soc.* **2008**, *130*, 9642–9643.

4.7 Leroy, C.; Dupas, G.; Bourguignon, J.; Quéguiner, G. *Tetrahedron* **1994**, *50*, 13135–13144.

(1.63 g, 6.2 mmol) in dry and degassed DMF (5.6 mL) at room temperature under argon atmosphere. After 5 minutes at the same temperature, the reaction mixture was diluted with EtOAc (200 mL) and washed with 1 M aqueous HCl (2 × 200 mL), water (200 mL) and brine (200 mL). The combined organic layers were dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The residue was purified by column chromatography (from 1:1 EtOAc/petrol to EtOAc) to afford **4.2** (404 mg, 50%) as a colourless crystalline solid. R_f (1:1 EtOAc/petrol) = 0.43. ^{1}H NMR ($CDCl_3$, 400 MHz): δ_H = 2.13 (3H, s, $COCH_3$), 3.84 (3H, s, CO_2CH_3), 5.87 (1H, d, *J* = 1.5, =CH<u>H</u>), 6.59 (1H, aps, =C<u>H</u>H), 7.74 (1H, br. s, NH). ^{13}C NMR ($CDCl_3$, 100 MHz): δ_C = 24.7 (CH_3, Ac), 53.0 (OCH_3), 108.7 (C=<u>C</u>H_2), 130.8 (<u>C</u>=CH_2), 164.6 (C=O, Ac), 168.8 (C=O, <u>C</u>O_2Me) LRMS (ES^-): (M–H) $C_6H_9NO_3$ (m/z): calc. 142.1; found 142.1.

Optimization of the *Se*-allyl selenocysteine synthesis *via* conjugate addition.[a]

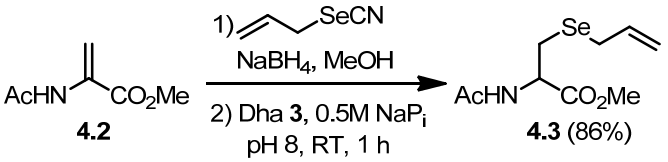
**Entry 1-10 were carried out by Dr. Omar Boutureira.*



Entry	*Se*-Nu (equiv.)	Base (equiv.)	Solvent (v/v)	Additive (equiv.)	*t* (h)	Product (%)
1	**4.1a** (1)	K_2CO_3 (2)	DMF/H_2O (4:1)	-	4	NR[b]
2	**4.1a** (1)	Cs_2CO_3 (2)	DMF	Piperazine (1)	7	mixture
3	**4.1a** (1)	K_2CO_3 (2)	DMF/H_2O (4:1)	Et_2NH (4)	4	NR[b]
4	**4.1a** (1)	Cs_2CO_3 (2)	DMF	$NH_2NH{\cdot}AcOH$ (1)	0.5	NR[c]
5[d]	**4.1a** (1)	K_2CO_3 (1)	MeOH/H_2O (4:1)	$NaBH_4$ (2)	3.5	<10
6	**4.1b** (2)	K_2CO_3 (4)	DMF/H_2O (4:1)	-	0.5	NR[b]
7[e]	**4.1b** (2)	K_2CO_3 (4)	MeOH/H_2O (4:1)	$NaBH_4$ (2)	1	53
8	**4.1b** (2)	K_2CO_3 (4)	DMF/H_2O (4:1)	DTT (2)	0.5	40
9[e]	**4.1b** (2)	K_2CO_3 (4)	DMF/H_2O (4:1)	DTT (2)	1	45
10[e]	**4.1c** (2)	K_2CO_3 (4)	MeOH/H_2O (4:1)	$NaBH_4$ (2)	2	22
11[e]	**4.1c (4)**	**-**	**0.5 M NaP_i (pH 8)**	**$NaBH_4$ (4)**	**1**	**86**

(a) General conditions: AcDhaOMe (1 equiv.), *Se*-allyl reagent (1–4 equiv.), additive (1–4 equiv.) and base (1–4 equiv.) in solvent (13.2 mL/mmol) at room temperature unless otherwise indicated. (b) NR= no reaction (>95% of the starting material was recovered). (c) NR= no reaction (61% of the starting material was recovered). (d) Pre-reduction of 4.1a in methanol at room temperature for 1 h prior to the addition of AcDhaOMe. (e) Pre-reduction of *Se*-allyl reagent in methanol at room temperature for 5 min prior to the addition of AcDhaOMe.

N-Acetyl-*Se*-allyl-DL-cysteine methyl ester (AcSeacOMe, 4.3)

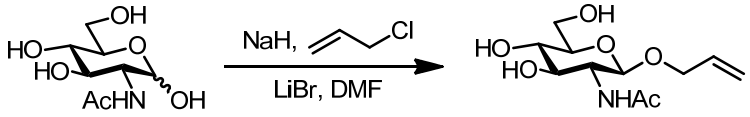


$NaBH_4$ (53 mg, 1.40 mmol) was added portionwise to a solution of allyl selenocyanate (**4.1c**) (200 mg, 1.40 mmol) in degassed MeOH (7 mL) at room temperature. Hydrogen gas evolution occurred, and the solution turned from cloudy to clear. After 5 minutes, this solution was added dropwise to AcDhaOMe (**4.2**) (50 mg, 0.35 mmol) in 0.5 M sodium phosphate (pH 8, 2 mL) and stirred at room temperature for 1 h. The reaction diluted with EtOAc (100 mL) and washed subsequently with 1 M HCl (aq) (50 mL)*, saturated $NaHCO_3$ (aq) (50 mL), water (50 mL) and brine (50 mL). The organic layers were dried over $MgSO_4$, filtered, and concentrated under reduced pressure. The residue was purified by column chromatography (3:2 EtOAc/petrol) to afford **4.3** (79 mg, 86%) as a colourless syrup. R_f (1:1 EtOAc/petrol) = 0.2. ^{1}H NMR ($CDCl_3$, 400 MHz): δ_H = 2.05 (3H, s, $COCH_3$), 2.92 (1H, dd, J_{AB} = 13.1, $J_{B,\alpha}$ = 5.3, CH_2Se), 2.98 (1H, dd, J_{AB} = 13.1, $J_{A,\alpha}$ = 5.1, CH_2Se), 3.16 (2H, dd, J = 7.6, 0.8, $SeCH_2$), 3.77 (3H, s, OCH_3), 4.88 (1H, dt, J = 7.6, 5.2, H_α), 5.00–5.09 (2H, m, $=CH_2$), 5.84 (1H, dtd, J = 15.3, 9.6, 7.7, CH=), 6.31 (1H, br. s, NH),. ^{13}C NMR ($CDCl_3$, 100 MHz): δ_C = 23.2 (CH_3, Ac), 25.0 (CH_2Se), 26.8 ($SeCH_2$), 51.9 (C_α), 52.7 (OCH_3), 117.1 ($=CH_2$), 134.3 (CH=), 169.7 (C=O, Ac), 171.4 (C=O, $\underline{C}O_2Me$). IR (υ_{max}, film): 3280, 3077, 2953, 2927, 2852, 1746,

1656, 1541, 1213. HRMS (ES^+) for (M+Na) $C_9H_{15}NaO_3Se$ (*m/z*): calc. 288.0109; found 288.0111.

*Excess allyl selenolate is quenched on addition of acid. Allyl selenol in solution is unstable and forms diallyl diselenide, which readily decomposes to diallyl selenide and selenium (brick red solid) via allylic rearrangement.

_O_-Allyl-_N_-Acetyl-D-glucosamine (4.10)

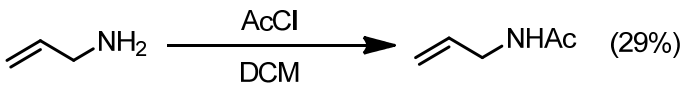


A 25 mL two-necked round-bottomed flask was flamed dried under argon. GlcNAc (500 mg, 2.26 mmol), LiBr (392 mg, 4.52 mmol) and DMF (4 mL, anhydrous) were added sequentially under a stream of argon. The suspension was stirred at RT until homogeneous (~1 h). NaH (108 mg, 2.71 mmol) was added, followed by allyl chloride (0.56 mL, 6.78 mmol) (Note: mixture turned cloudy on addition of NaH). The reaction was monitored by TLC (0.5:3:6.5, H_2O:^{i}PrOH:EtOAc). After 2 h, an additional 3 equivalents of allyl chloride was added (0.56 mL, 6.78 mmol). After 2 h, the solvent was removed under reduced pressure. The resulting yellow residue was purified directly by column chromatography (0.5:3:6.5, H_2O:^{i}PrOH:EtOAc). A clear oil was obtained that contained residual DMF. A second purification by column chromatography (20% MeOH in EtOAc) provided white needle-like crystals which are very hydroscopic (321 mg, 54%). Spectroscopic data was consistent with that previously reported.[4.8] $[\alpha]^{20}_D$ = −5.0 (c = 1.0, MeOH). IR (υ_{max}, KBr): 3451, 1647, 1556, 1378, 1074, 946, 626. ^{1}H NMR (400 MHz, CD_3OD) δ_H = 2.02 (3H, s, $NHCO_2C\underline{H}_3$), 3.25-3.59 (2 H, m, contain H4 and H5), 3.65-3.96 (4H, m, contain H2, H3 and OCH_2), 4.10 (1H, dd, *J* = 13.4, 5.8, H6), 4.35 (1H, dd, *J* = 13.3, 4.9 Hz, H′6), 4.50 (1H, d, *J* = 8.6, H1), 5.09-5.19 (1H, m,

[4.8] Huang, G.-L.; Zhang, D.-W.; Zhao, H.-J.; Zhang, H.-C.; Wang, P.-G. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2042 - 2043.

CH=C<u>H</u>H *cis*), 5.28 (1H, dd, J = 17.2, 1.5, CH=CH<u>H</u> *trans*), 5.82-5.99 (1H, m, C<u>H</u>=CH_2). ^{13}C NMR (100 MHz, CD_3OD) δ_C = 22.2 ($NHCO_2\underline{C}H_3$), 56.2, 61.7 (contain C2, C3), 69.9 (C6), 71.0, 75.0, 76.8 (contain C4, C5 and OCH_2), 100.9 (C1), 116.2 (CH=$\underline{C}H_2$), 134.6 ($\underline{C}$H=CH_2), 173.0 (C=O). HRMS m/z (ESI$^+$): Found 284.1103 $[M+Na]^+$; $C_{11}H_{19}NO_6Na$ requires 284.1105.

_N_-acetyl-allylamine (4.11)



Allylamine (3.0 mL, 37.6 mmol) was added to a 25 mL round bottom flask and dissolved in CH_2Cl_2 (10 mL). The stirred solution was cooled to 0 ºC and acetyl chloride (0.89 mL, 12.5 mmol) was added slowly. The reaction was stirred, open to air, for 5 minutes at 0 ºC. After this time, the reaction was diluted with EtOAc (100 mL) and washed with 1M HCl (2 × 100 mL). The organic layer was dried ($MgSO_4$), filtered, and concentrated under reduced pressure (115 mbar, 30 ºC bath temp) to give the product as a clear oil (360 mg, 29%). Spectroscopic data was consistent with that previously reported.[4.9] IR (υ_{max}, film): 3287, 3083, 2922, 1652, 1555, 1429, 1374, 1284, 1040, 920. ^{1}H NMR (200 MHz, $CDCl_3$): δ_H = 1.99 (3H, s, NH<u>Ac</u>), 3.84 (2H, tt, *J* = 1.4, 5.6, C<u>H</u>$_2$NHAc), 5.07-5.91 (2H, m, C<u>H</u>$_2$=CH), 5.71-5.91 (1H, m, CH_2=C<u>H</u>), 6.13 (br. s, 1H, N<u>H</u>Ac). ^{13}C NMR (50 MHz, $CDCl_3$): δ_C = 23.0 (NH<u>Ac</u>), 42.0 (<u>C</u>H_2NHAc), 116.3 (<u>C</u>H_2=CH), 134.0 (CH_2=<u>C</u>H), 170.3 (C=O). LRMS (ESI$^+$): 122.1 $(M+Na)^+$.

Synthesis of Seac-tagged proteins

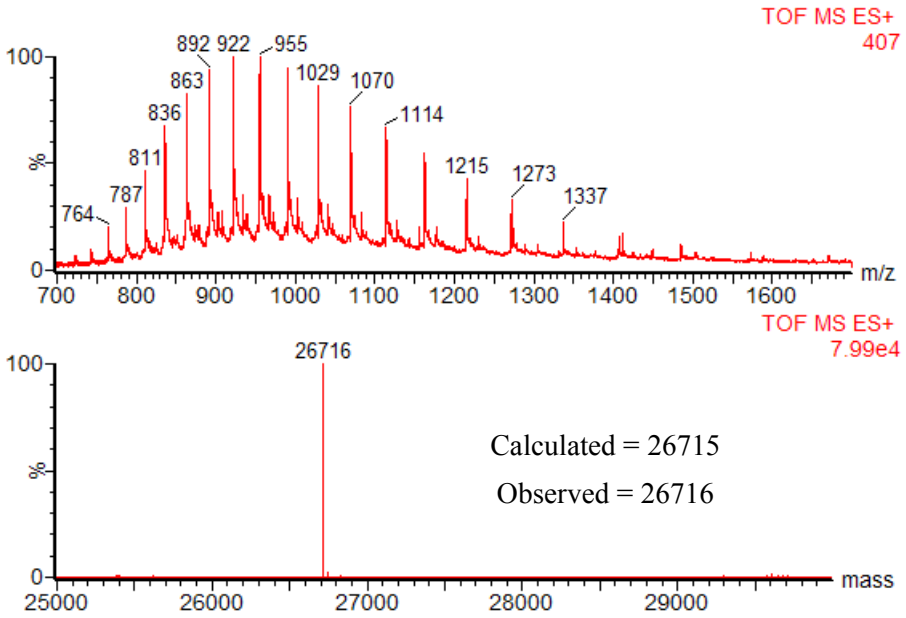
Sequence of subtilisin Bacillus lentus (SBL) mutant S156C

PDB code for wild type = 1GCI

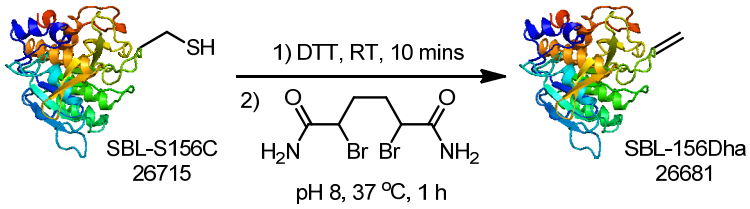
[4.9] Schmidt, A. M.; Eilbracht, P. *Org. Biomol. Chem.* **2005**, *3*, 2333-2343.

AQSVPWGISRVQAPAAHNRGLTGSGVKVAVLDTGISTHPDLNIRGGASFVPGEPSTQD
GNGHGTHVAGTIAALNNSIGVLGVAPSAELYAVKVLGASGSGSVSSIAQGLEWAGNNG
MHVANLSLGSPSPSATLEQAVNSATSRGVLVVAASGNCGAGSISYPARYANAMAVGAT
DQNNNRASFSQYGAGLDIVAPGVNVQSTYPGSTYASLNGTSMATPHVAGAAALVKQKN
PSWSNVQIRNHLKNTATSLGSTNLYGSGLVNAEAATR

Calculated average isotopic mass = 26715

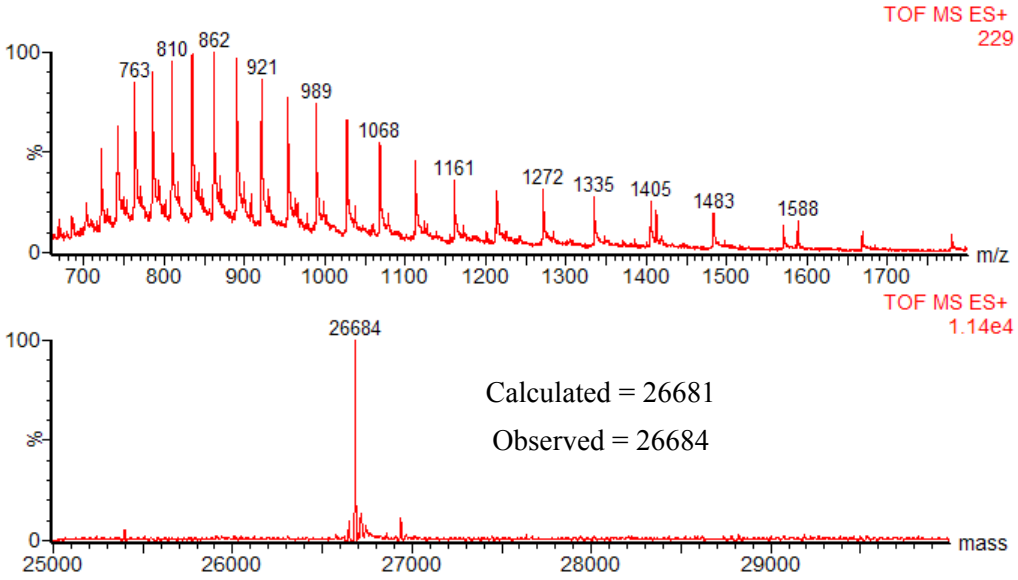


SBL-156Dha (4.4)

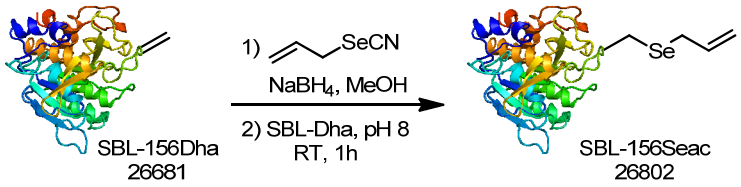


To a solution of SBL-S156C (2.5 mL, 1.2 mg/mL, 0.11 µmol) in 50 mM sodium phosphate buffer (pH 8) was added DTT (1.3 mg, 8.4 µmol). After mixing on a vortex the solution was shaken for further 10 minutes at RT. The reaction was then passed through a PD-10 desalting column (GE Healthcare) pre-equilibrated with 50 mM sodium phosphate buffer (pH 8) followed by elution with 3.5 mL of the buffer. The resulting solution was concentrated to 500 µL (6.0 mg/mL) with To the DTT-treated SBL-S156C solution (500 µL, 6.0 mg/mL, 0.11 µmol) was added 2,5-dibromohexane-

diamide (17 mg, 56 µmol, in 112 µL of DMF) was added to the protein solution. The reaction was mixed on a vortex and shaken at RT for 30 minutes then incubated at 37 °C for further 1 hour. To purify the protein, the reaction was loaded on a PD MiniTrap (GE Healthcare) pre-equilibrated with 50 mM sodium phosphate buffer (pH 8) and 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-Dha (**4.4**) (26681 calculated mass, 26684 found). ESI-MS are shown below. The protein solution was flash frozen with liquid nitrogen and stored at –20 °C.

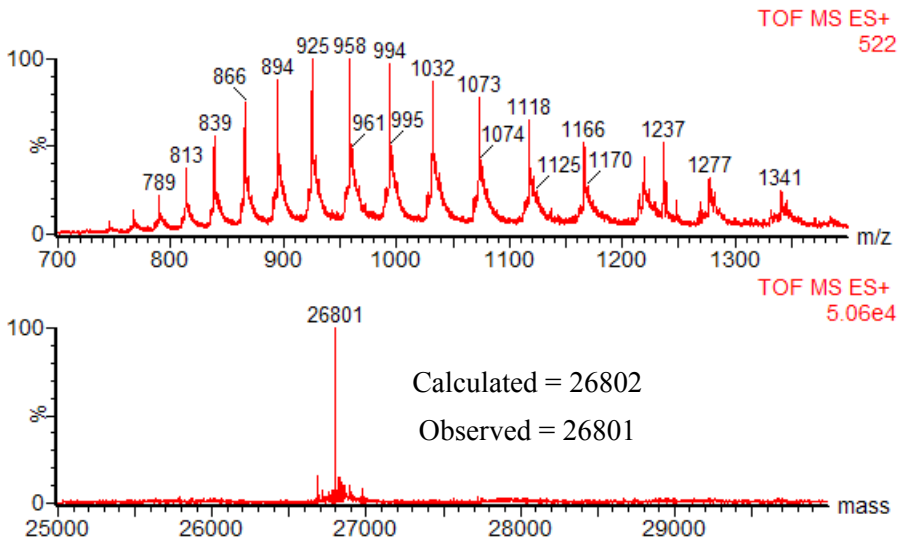


SBL-156Seac (4.5a)



$NaBH_4$ (2.52 mg, 65.5 µmol) was added portionwise to a solution of allyl selenocyanate (**4.1c**) (9.6 mg, 65.5 µmol) in degassed MeOH (375 µL) at room temperature and the resulting mixture vortexed for 5 minutes. Gas evolution occurred, and the solution turned colourless. This solution was added dropwise to a solution of SBL-156Dha (**4.4**) (2.5 mL of 0.7 mg/mL, 65.5 nmol) in 50 mM sodium phosphate buffer (pH 8.0) at room

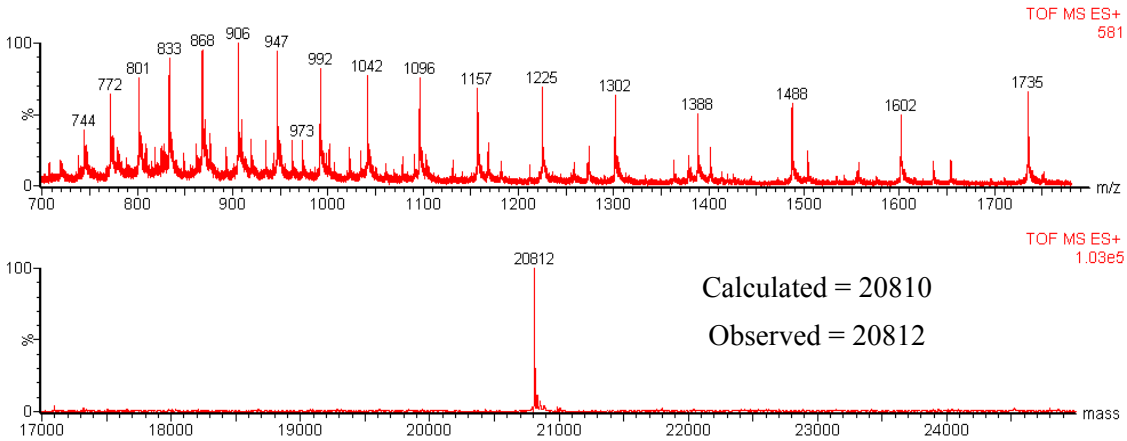
temperature and the resulting mixture vortexed for 30 seconds. After 1 h of additional shaking, a 50 µL aliquot was analyzed directly by LC–MS and complete conversion to SBL-156Seac (**4.5a**) (26802 calculated; 26803 found) was observed. Small molecules were removed from the reaction mixture by loading the sample onto a PD10 desalting column (GE Healthcare) previously equilibrated with 10 column volumes of pH 5.0 sodium phosphate buffer (50 mM) and eluting with 3.5 mL of the same buffer. The solution was concentrated on a vivaspin membrane concentrator (10 kDa molecular weight cut off) and washed with sodium phosphate buffer (3 x 500 µL, 50 mM, pH 8.0). Finally, the solution was concentrated to 1 mL (1.17 mg/mL by BCA assay) and was flash frozen with liquid nitrogen and stored at –80 ºC. ESI-MS of the reaction is shown below.



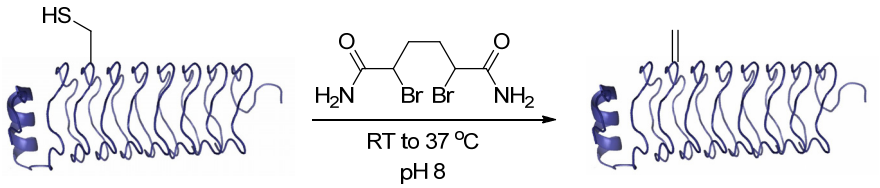
Sequence for Npβ M21I M61C

GSSHHHHHHSSGLVPRGSHIDVGKLRQLYAAGERDFSIVDLRGAVLENINLSGAILHG
ACLDEANLQQANLSRADLSGATLNGADLRGANLSKADLSDAILDNAILEGAILDEAVL
NQANLKAANLEQAILSHANIREADLSEANLEAADLSGADLAIADLHQANLHQAALERA
NLTGANLEDANLEGTILEGGNNNLAT

Calculated average isotopic mass = 20810 (*N*-terminal Met cleaved)

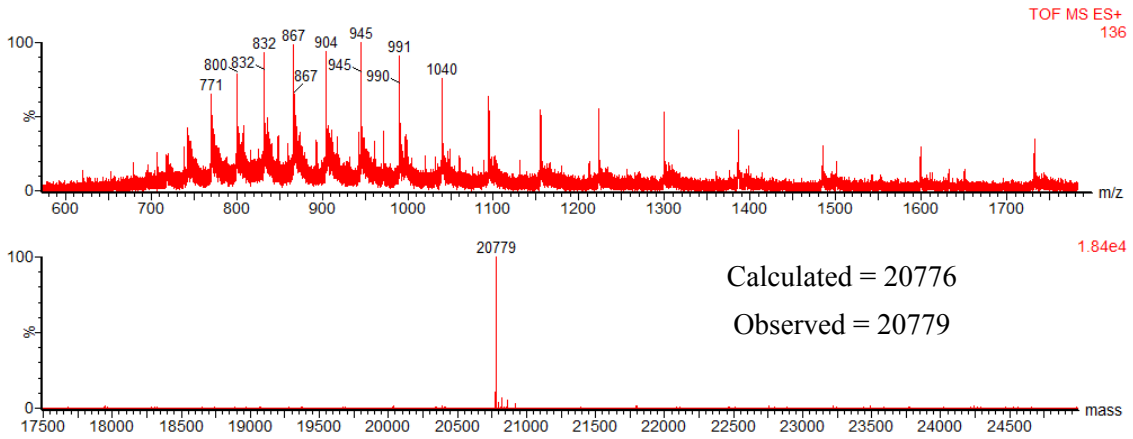


Npβ-61Dha (4.6)

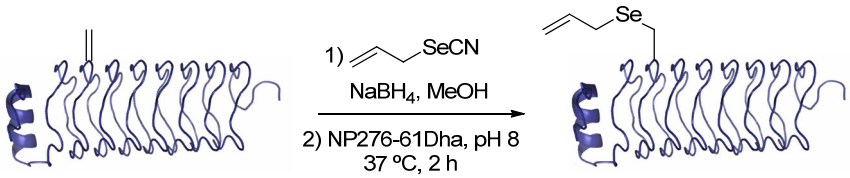


D/L-Dithiothreitol (DTT) (10 mg, 64.9 µmol) was added as a solid to a solution of Npβ M21I M61C (2.5 mL, 1.1 mg/mL, 132.1 nmol) in 50 mM sodium phosphate buffer at pH 8.0 and the resulting mixture shaken for 15 minutes. Small molecules were removed from the reaction mixture by loading the sample onto a PD10 desalting column (GE Healthcare) previously equilibrated with 10 column volumes of pH 8.0 sodium phosphate buffer (50 mM) and eluting with 3.5 mL of the same buffer. To a solution of Npβ M21I M61C (2.5 mL, 0.79 mg/mL, 95.1 nmol) in 50 mM sodium phosphate buffer at pH 8.0, a freshly prepared solution of α,α′-di-bromo-adipyl(bis)amide (125 µL of 114 mg/mL, 48.1 µmol) in DMF was added and the resulting mixture vortexed for 30 seconds and placed in a 37 ºC incubator with shaking. After 15 minutes, an additional 125 µL aliquot of the di-bromide solution was added to the reaction. After another 15 minutes, a third and final 125 µL aliquot of the di-bromide solution was added. The reaction was shaken at 37 ºC for a final 30 minutes. A 20 µL aliquot was analyzed directly by LC–MS and complete conversion to Npβ-61Dha (**4.6**) was observed (calculated mass, 20776; observed mass,

20779). Excess di-bromide was removed by centrifugation (10000 rcf, 1 min) and the protein solution was flash frozen with liquid nitrogen and stored at –20 ºC.

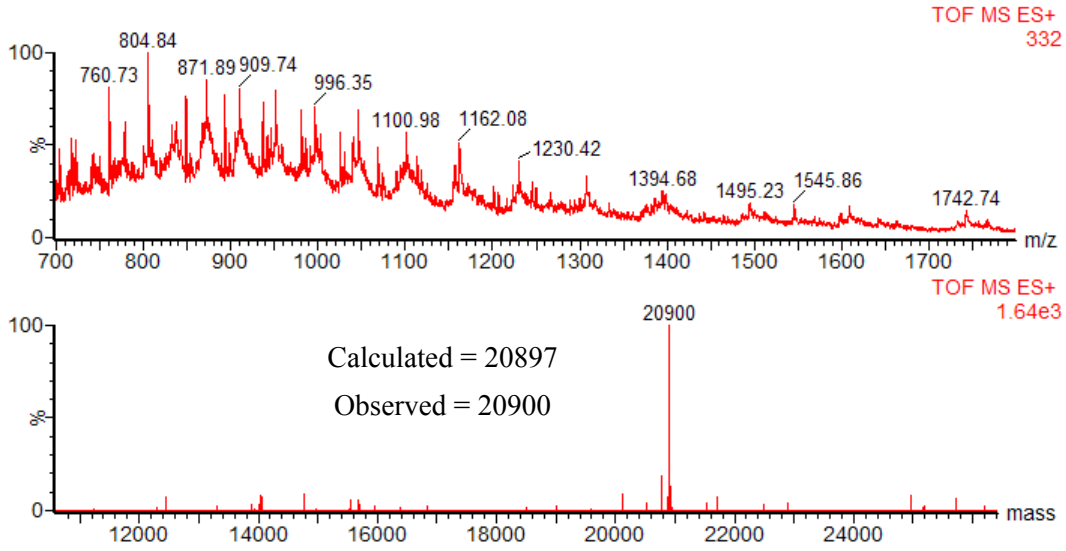


Npβ-61Seac (4.7a)



$NaBH_4$ (0.5 mg, 13.2 µmol) was added portionwise to a solution of allyl selenocyanate **4.1c** (1.9 mg, 13.2 µmol) in degassed MeOH (75 µL) at room temperature and the resulting mixture vortexed for 5 minutes. Gas evolution occurred, and the solution turned colourless. This solution was added dropwise to a solution of Npβ-61Dha (**4.6**) (0.5 mL of 0.55 mg/mL, 13.2 nmol) in 50 mM sodium phosphate buffer (pH 8.0) at room temperature and the resulting mixture vortexed for 30 seconds. After 2 h of additional shaking, a 20 µL aliquot was analyzed directly by LC–MS and >95% conversion to Npβ-61Seac (**4.7a**) (calculated mass, 20897; observed mass, 20900) was observed. Small molecules were removed from the reaction mixture by loading the sample onto a PD MiniTrap desalting column (GE Healthcare) previously equilibrated with 10 column volumes of pH 8.0 sodium phosphate buffer (50 mM) and eluting with 1 mL of the same buffer. The solution was concentrated on a Vivaspin® membrane concentrator (10 kDa

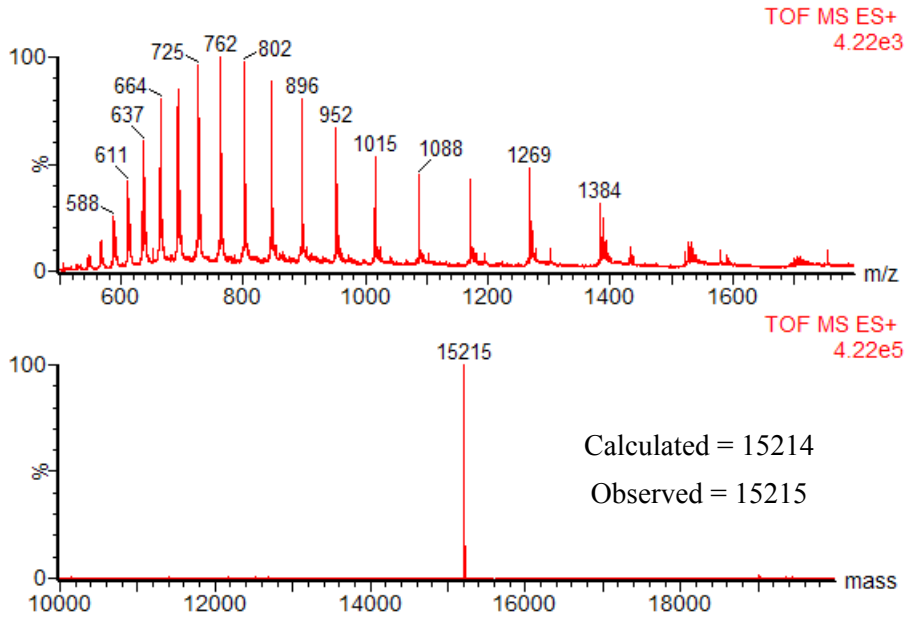
molecular weight cut off) and washed with sodium phosphate buffer (3 x 100 µL, 50 mM, pH 8.0). Finally, the solution was concentrated to 250 µL (0.51 mg/mL by BCA assay) and was flash frozen with liquid nitrogen and stored at –20 ºC.



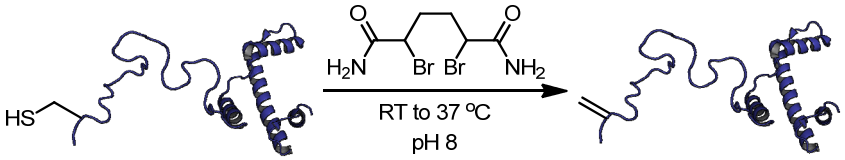
Sequence for H3 K9C C110A

ARTKQTARCSTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKST
ELLIRKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVALFEDTNLAAIHAKR
VTIMPKDIQLARRIRGERA

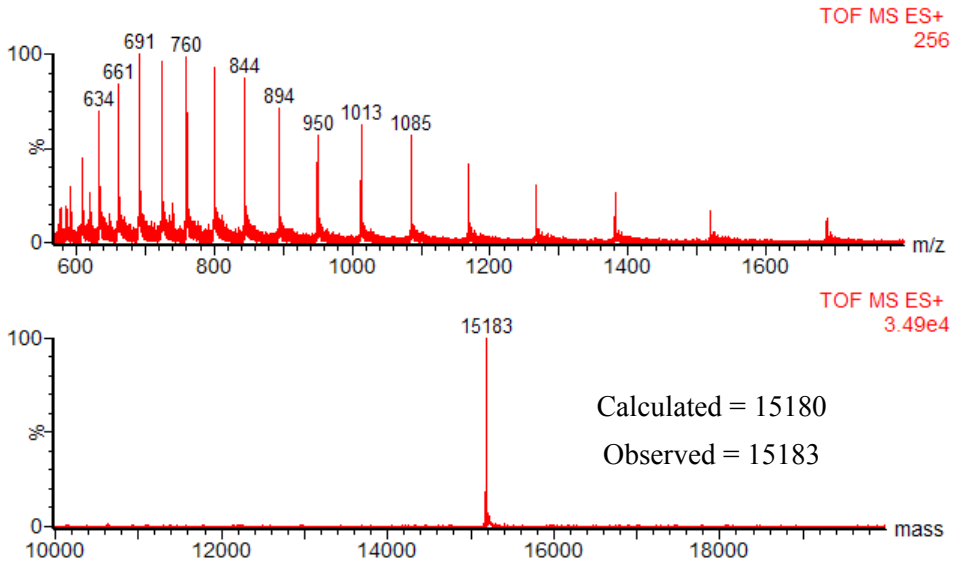
Calculated average isotopic mass = 15214



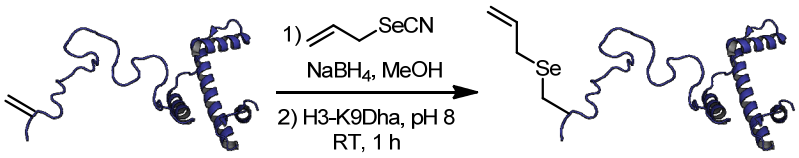
H3-K9Dha (4.12)



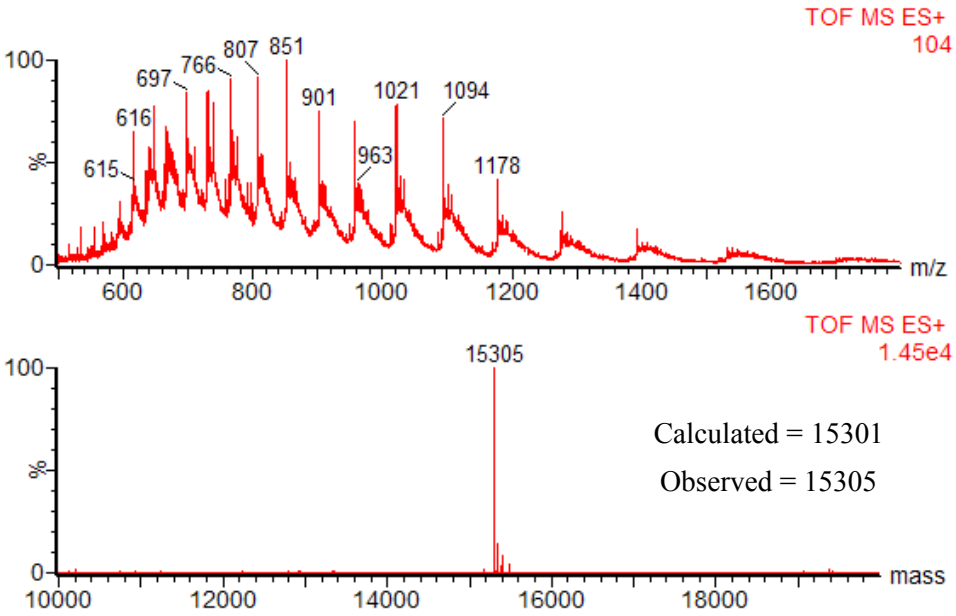
To a solution of H3-K9C (0.5 mL, 4 mg/mL, 0.13 µmol) in 50 mM Tris (pH 8) with 5 M guanidium chloride (GdmCl) was added DTT (2.0 mg, 13 µmol). After mixing on a vortex the solution was shaken for further 30 minutes at RT. The reaction was then passed through a PD MiniTrap (GE Healthcare) pre-equilibrated with 50 mM Tris (pH 8) with 5 M GdmCl followed by elution with 1.0 mL of the same buffer. This solution was concentrated to 0.5 mL with Vivaspin® (10 kDa MWCO). 2,5-dibromohexane-diamide (7.6 mg, 25 µmol, in 50 µL of DMF) was then added to the protein solution. The reaction was mixed on a vortex and shaken at RT for 45 minutes then incubated at 37 °C for further 1.5 hour. To purify the protein, the reaction was loaded on a PD MiniTrap (GE Healthcare) pre-equilibrated with 50 mM Tris (pH 8) with 5 M GdmCl and eluted with 1.0 mL of the same buffer (final concentration = 2 mg/mL). LC–MS analysis of the product solution revealed a single protein with a mass corresponding to H3-K9Dha **4.12** (15180 calculated mass, 15181 found). ESI–MS are shown below. The protein solution was stored at –20 °C.



H3-K9Seac (4.13)

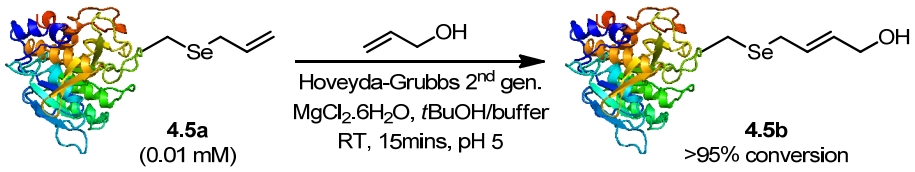


A solution of H3-K9Dha (**4.12**) (500 µL, 0.8 mg/mL, 0.026 µmol) was prepared by diluting H3-K9Dha (200 µL, 2 mg/mL in 5M GdmCl with 50 mM Tris, pH 8) with 300 µL of 50 mM Tris (pH 8). $NaBH_4$ (1.0 mg, 26 µmol) was added to a solution of allyl selenocyanate (**4.1c**) (3.8 mg, 26 µmol) in degassed MeOH (80 µL) at room temperature. After H_2 gas evolution, the resulting mixture was vortexed for 5 minutes before adding to the protein solution. The reaction was mixed on a vortex and then shaken at room temperature for 1 hour. LC–MS revealed complete conversion to H3-K9Seac (**4.13**) (15301 calculated mass; 15305 found) was observed. Small molecules were removed by loading the reaction mixture onto a PD10 column (pre-equilibrated with 2 M GdmCl with 50 mM sodium phosphate buffer, pH 8) and eluting with 1 mL of the same buffer (final concentration is 0.4 mg/mL). The resulting protein solution was then stored at –20 °C.

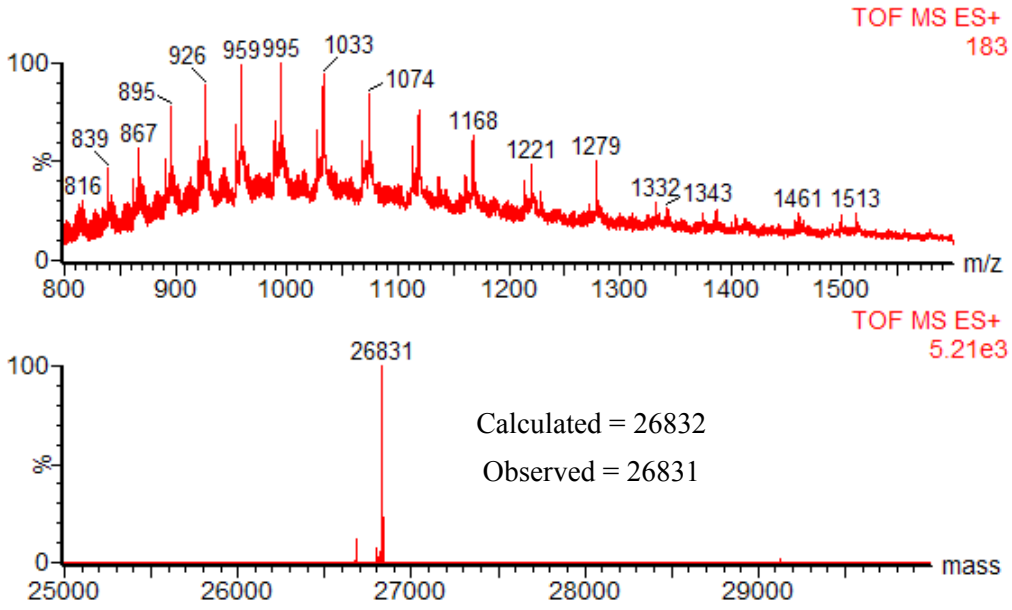


Cross-metathesis reactions on proteins

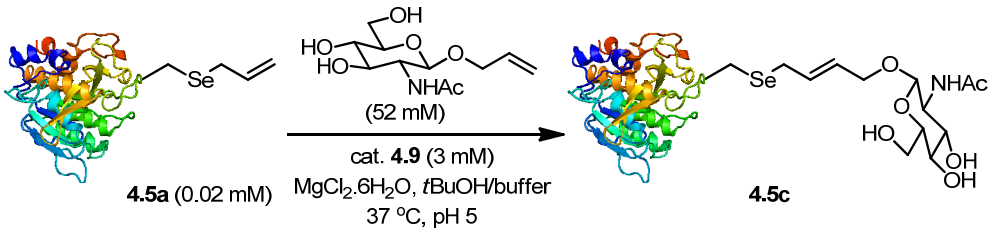
CM between SBL-156Seac (4.5a) and allyl alcohol



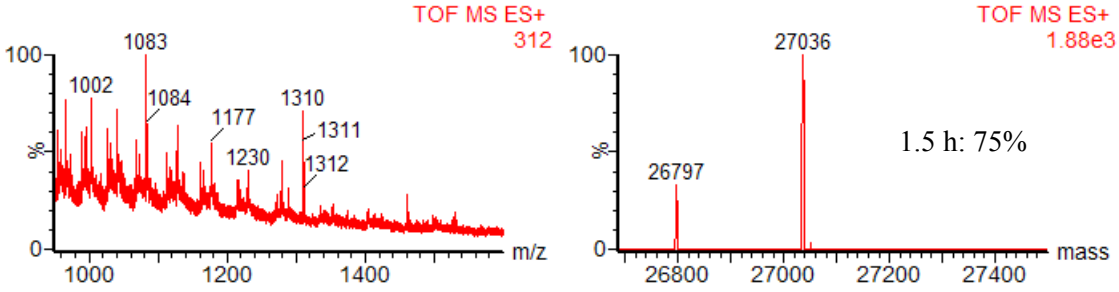
A solution of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) in ^{t}BuOH was prepared by vortexing and gentle warming 0.6 mg of **4.9** in 287 µL *t*BuOH. To a solution of SBL-156Seac (**4.5a**) (10.4 µM, 100 µL of 0.4 mg/mL solution) in 50 mM sodium phosphate buffer (pH 5) was added $MgCl_2.6H_2O$ (0.7 µL of a 1 mg/µL solution in water, 3.7 µmol) followed by the catalyst/*t*BuOH solution (43 µL, 0.15 µmol). The reaction was mixed on a vortex after each addition. The reaction mixture appeared as a green emulsion. Reaction was shaken at RT for 1 minute before adding allyl alcohol (0.5 µL, 7.5 µmol). Reaction was mixed on a vortex immediately after addition and was shaken at RT. After 15 minutes, the reaction was passed through a PD SpinTrap G-25 (GE Healthcare) to remove small molecules. The resulting solution was analyzed by LC-MS which showed >90% conversion to the CM product (26832 calculated mass, 26831 found).



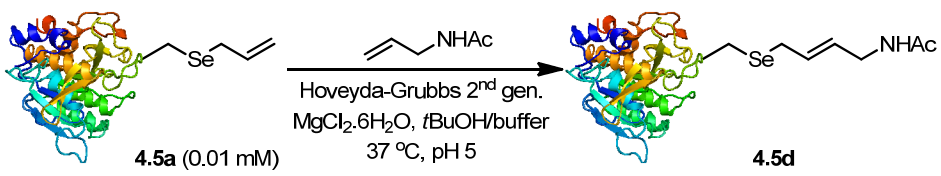
CM between SBL-156Seac (4.5a) and *O*-allyl GlcNAc (4.10)



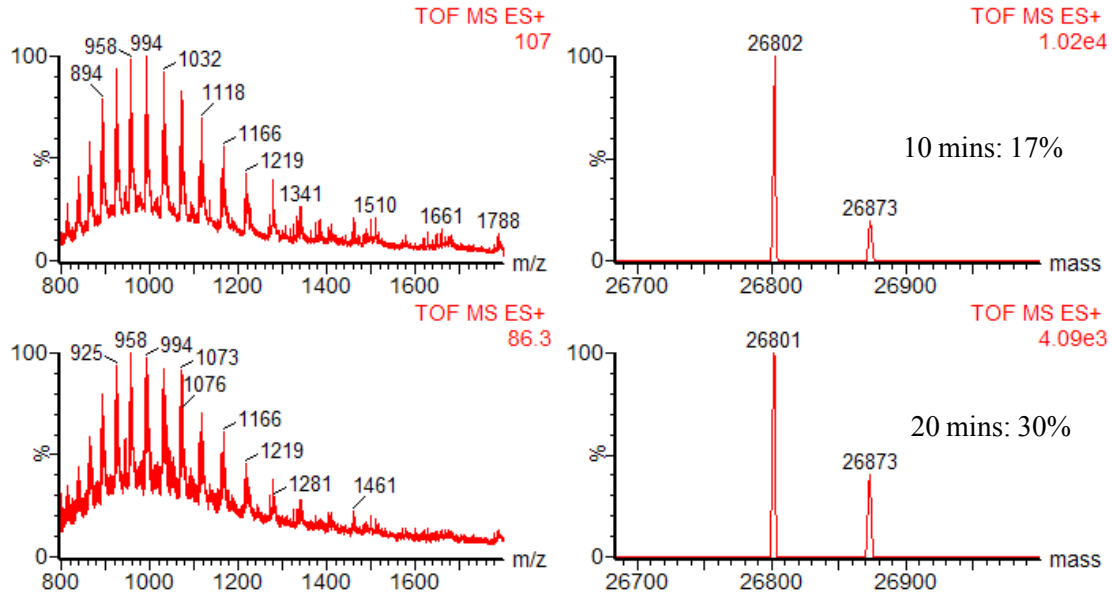
A solution of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) in ^{t}BuOH was prepared by vortexing and gentle warming 0.8 mg of **4.9** in 120 µL *t*BuOH. To a solution of SBL-156Seac (**4.5a**) (20 µM, 50 µL) in 50 mM sodium phosphate buffer (pH 5) was added $MgCl_2.6H_2O$ (0.4 µL of a 1 mg/µL solution in water, 2 µmol) followed by the catalyst/*t*BuOH solution (20 µL, 0.15 µmol). The reaction was mixed on a vortex each time after an addition. The reaction mixture appeared as a green emulsion. Reaction was shaken at RT for 1 minute before adding *O*-allyl GlcNAc (**4.10**) (1 µL of a 1 mg/µL solution, 3.7 µmol). Reaction was mixed on a vortex immediately after addition and was incubated at 37 °C. After 1.5 hours, the protein was purified by passing through a PD SpinTrap G-25 (GE Healthcare) to remove small molecules prior to analysis by LC-MS. LC-MS showed a 75% conversion to the CM product **4.5c** (calculated mass = 27035, observed mass = 27036).

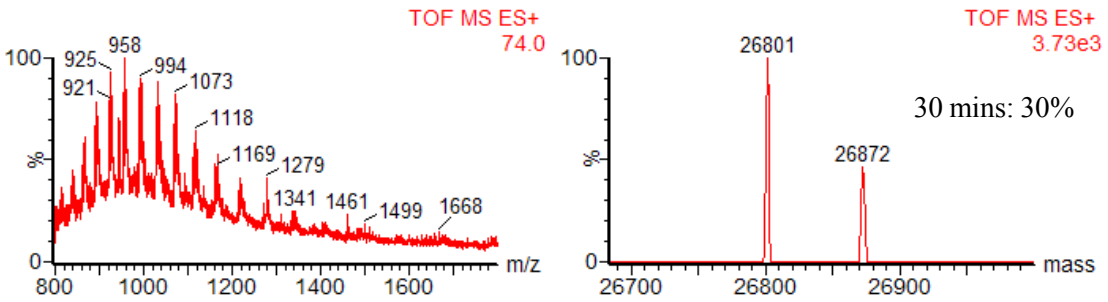


CM between SBL-156Seac (4.5a) and *N*-allyl acetamide (4.11)

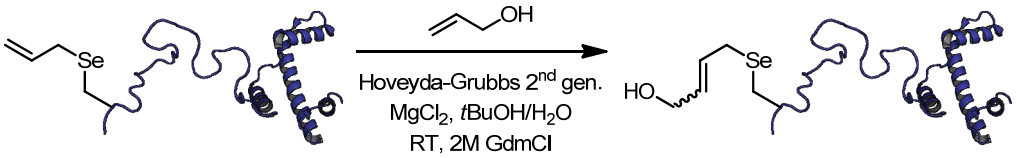


A solution of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) in ^{t}BuOH was prepared by vortexing and gentle warming 0.7 mg of **4.9** in 156 µL *t*BuOH. To a solution of SBL-156Seac (**4.5a**) (10.7 µM, 50 µL) in 50 mM sodium phosphate buffer (pH 5) was added $MgCl_2.6H_2O$ (0.4 µL of a 1 mg/µL solution in water, 2 µmol) followed by the catalyst/*t*BuOH solution (20 µL, 0.15 µmol). The reaction was mixed on a vortex each time after an addition. The reaction mixture appeared as a green emulsion. Reaction was shaken at RT for 1 minute before adding *N*-allyl acetamide (**4.11**) (0.4 µL, 0.37 mg, 3.7 µmol). Reaction was mixed on a vortex immediately after addition and was incubated at 37 °C. 15 µL aliquots were taken from reaction mixture at 10, 20 and 30 minutes, each time the aliquot was passed through a PD SpinTrap G-25 (GE Healthcare) to remove small molecules prior to analysis by LC-MS. LC-MS showed 30% conversion to the CM product **4.5d** (calculated mass = 26873, observed mass = 26873). At 20 minutes the reaction mixture turned brown in colour, a sign of catalyst degradation, hence no further reaction was observed at 30 minutes.

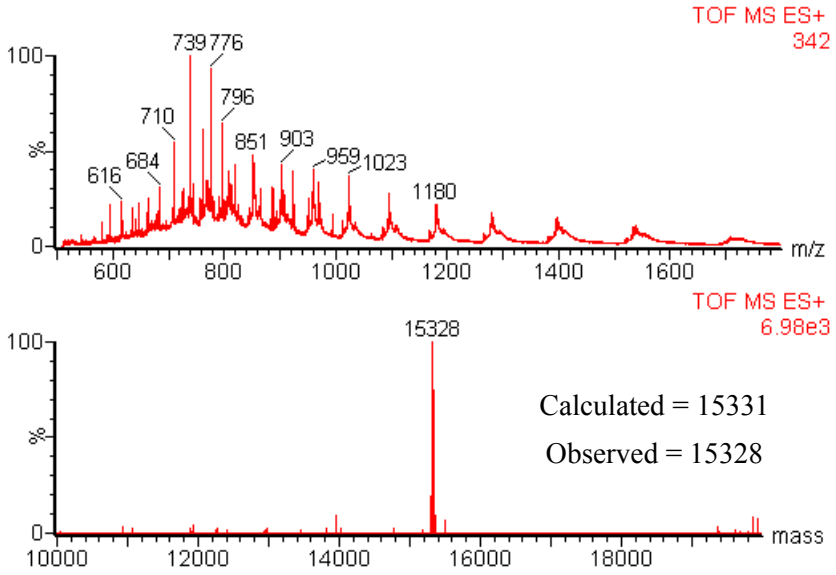




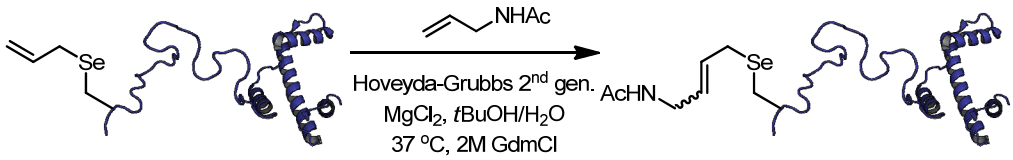
CM between H3-K9Seac (4.13) and allyl alcohol



A solution of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) in ^{t}BuOH was prepared by vortexing and gentle warming 0.62 mg of **4.9** in 132 µL *t*BuOH. To a solution of H3-K9Seac (**4.13**) (150 µL, c = 0.4 mg/mL, 3.9 nmol) in 50 mM sodium phosphate buffer (pH 8) with 2 M GdmCl was added $MgCl_2.6H_2O$ (4 µL of a 1 mg/µL solution in water, 20 µmol) and the catalyst/*t*BuOH solution (64 µL, 0.48 µmol). The reaction was mixed on a vortex each time after an addition. Reaction was shaken at RT for 1 minute before adding allyl alcohol (1.3 µL, 20 µmol). Reaction was mixed on a vortex immediately after addition and was shaken at RT for 15 minutes. 3-mercaptopropionic acid (1.5 µL, 17 µmol) was added as a metal scavenger for ruthenium and the mixture was shaken for further 30 minutes at RT. The protein sample was then purified from small molecules by dilution with 50 mM sodium phosphate (pH 8) with 2 M GdmCl (300 µL) and then concentration by Vivaspin® (5 kDa MWCO) to 150 µL. This process was repeated for four times. LC-MS analysis of the protein sample revealed >95% conversion to the desired CM product **4.14** (calculated = 15331, found = 15328).

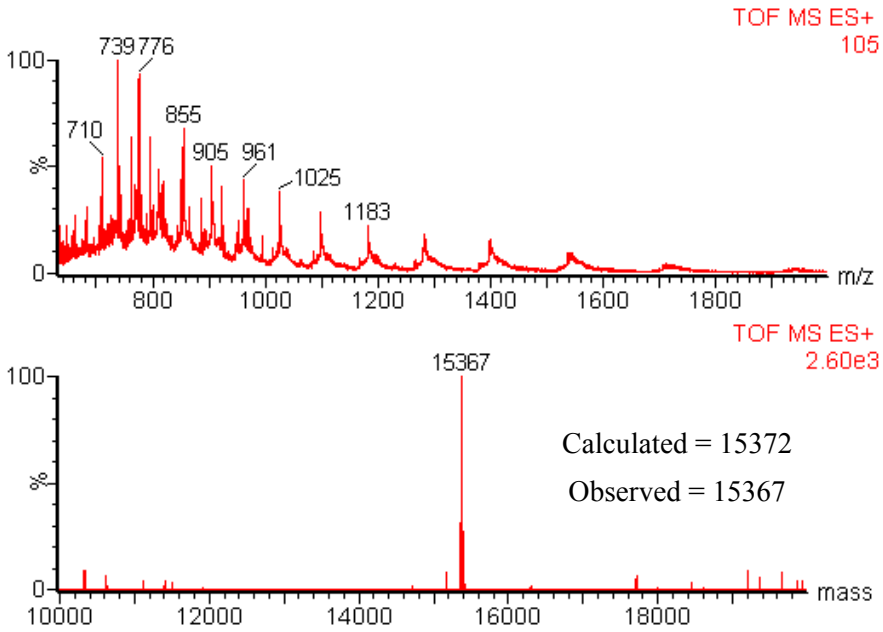


CM between H3-K9Seac (4.13) and *N*-acetyl-allylamine (4.11)



A solution of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) in ^{t}BuOH was prepared by vortexing and gentle warming 0.43 mg of **4.9** in 92 µL *t*BuOH. To a solution of H3-K9Seac (**4.13**) (100 µL, c = 0.4 mg/mL, 2.6 nmol) in 50 mM sodium phosphate buffer (pH 8) with 2 M GdmCl was added $MgCl_2.6H_2O$ (2.6 µL of a 1 mg/µL solution in water, 13 µmol) and the catalyst/*t*BuOH solution (43 µL, 0.3 µmol). The reaction was mixed on a vortex each time after an addition. Reaction was shaken at RT for 1 minute before adding *N*-allyl acetamide (**4.11**) (1 mg, 10 µmol). Reaction was mixed on a vortex immediately after addition and was incubated at 37 °C for 20 minutes. 3-mercapto-propionic acid (1 µL, 11 µmol) was added as a metal scavenger for ruthenium and the mixture was shaken for further 30 minutes at RT. The protein sample was then purified from small molecules by dilution with 50 mM sodium phosphate (pH 8) with 2 M GdmCl (300 µL) and then concentration by Vivaspin® (5kDa MWCO) to 100 µL. This process was repeated for four times. LC-MS analysis of the protein

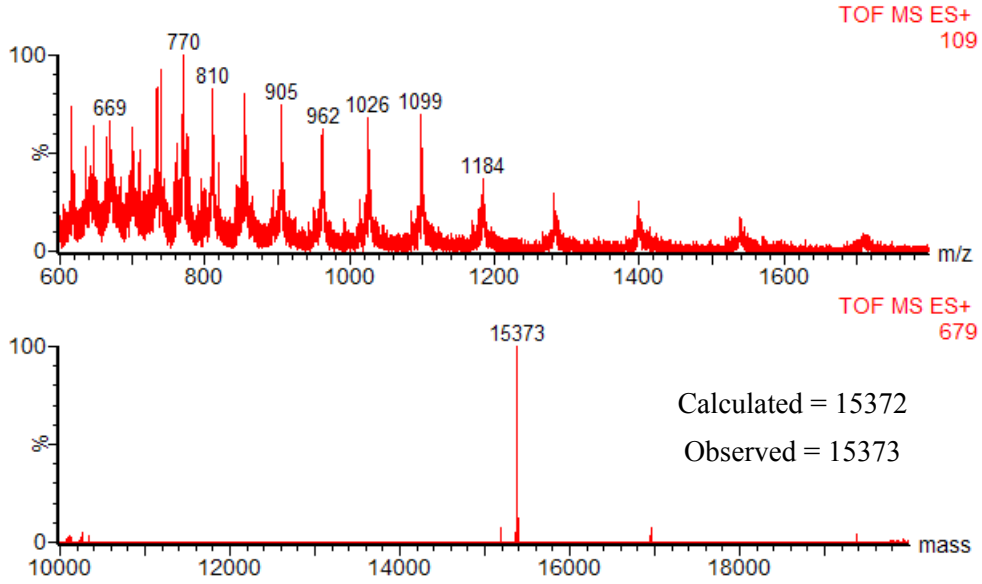
sample revealed >95% conversion to the desired CM product **4.15** (calculated = 15372, found = 15367).



This cross-metathesis was repeated at 350 µL scale following a similar procedure.

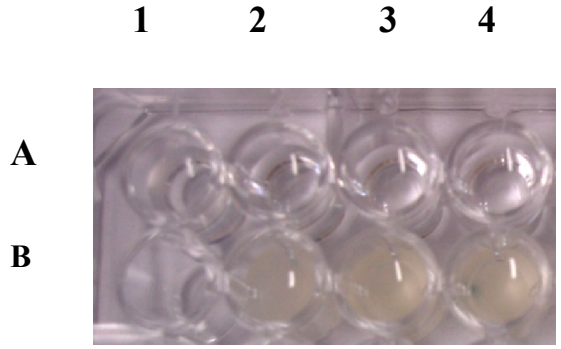
A solution of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) in *t*-BuOH was prepared by vortexing and gentle warming 0.67 mg of **4.9** in 160 µL *t*-BuOH. To a solution of H3-K9Seac (**4.13**) (350 µL, c = 0.4 mg/mL, 9.2 nmol) in 50 mM sodium phosphate buffer (pH 8) with 2 M GdmCl was added $MgCl_2.6H_2O$ (10 mg, 49 µmol) and the catalyst/*t*-BuOH solution (75 µL, 0.5 µmol). The reaction was mixed on a vortex each time after an addition. Reaction was shaken at RT for 1 minute before adding *N*-allyl acetamide (**4.11**) (2.5 mg, 26 µmol). Reaction was mixed on a vortex immediately after addition and was incubated at 37 °C for 20 minutes. A second dose of catalyst/*t*-BuOH solution (75 µL, 0.5 µmol) and acetamide **4.11** (2.5 mg, 26 µmol) was added to the reaction then further incubated at 37 °C for 20 minutes. After this time, 3-mercaptopropionic acid (5 µL, µmol) was added as a metal scavenger for ruthenium and the mixture was shaken for 1 hour minutes at RT. The protein sample was then purified from small molecules by dilution to 1 mL with 50 mM sodium phosphate (pH

8) with 2 M GdmCl and then concentration by Vivaspin® (5kDa MWCO) to 500 µL. The concentrate was then loaded onto a PDmini desalting column (GE Healthcare) pre-equilibrated with 50 mM sodium phosphate (pH 8) with 2 M GdmCl and eluted using 1 mL of the same buffer. LC-MS analysis of the protein sample revealed >95% conversion to the desired CM product.



Peptidase Activity of SBL *Se*-Proteins

SBL-S156C (unmodified), SBL-156Dha (**4.4**) (modified), SBL-156Seac (**4.5a**) (modified) were prepared at a concentration of 0.01 mg/mL in pH 8.0 sodium phosphate (50 mM). 200 µL aliquots of each protein sample were added to a 96-well plate. A 2 µL aliquot of SucAAPFpNA (0.20 M in DMSO, Bachem) was added to each of the protein samples. All SBL samples turned yellow immediately upon addition of the peptide substrate. The yellow solution indicates liberation of *p*-nitroaniline (pNA), confirming peptidase activity of all SBL samples. All protein solutions and the peptide solution alone at the same concentration are colorless (See below).



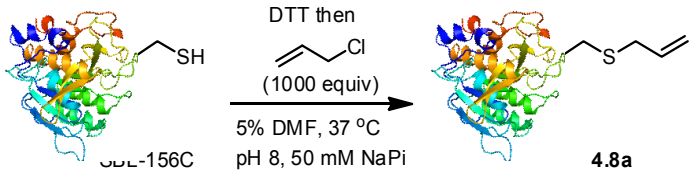
Well A1: SucAAPFpNA Well B1: (empty)

Well A2: SBL-S156C Well B2: SBL-S156C + SucAAPFpNA

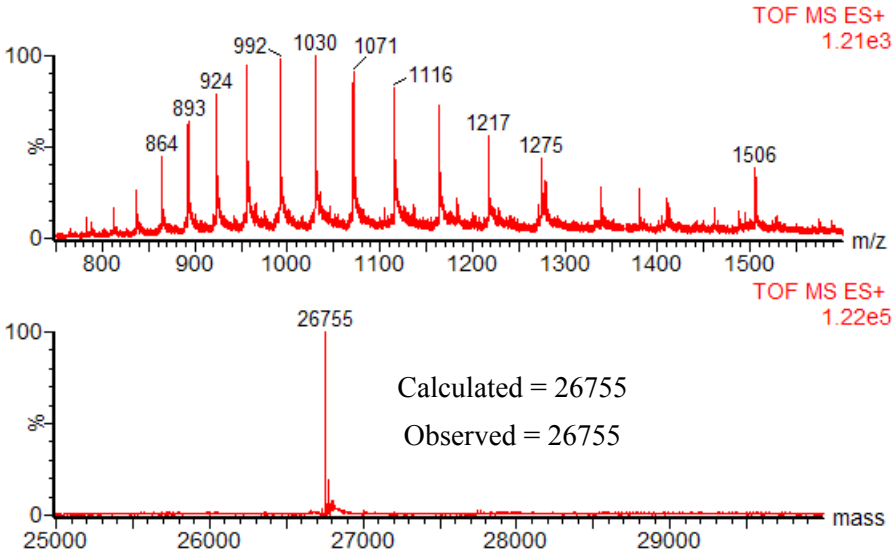
Well A3: SBL-156Dha **4.4** Well B3: SBL-156Dha **4.4** + SucAAPFpNA

Well A4: SBL-156Seac **4.5a** Well B4: SBL-156Seac **4.5a** + SucAAPFpNA

SBL-156Sac (4.8a)

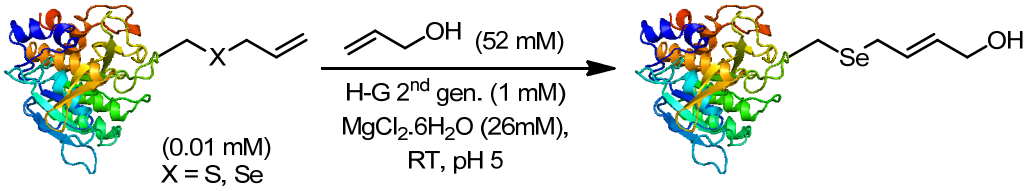


SBL-156Sac (**4.8a**) was synthesized following previously reported procedures.[4.5] SBL-S156C (2.5 mL, 1.0 mg/mL, pH 8.0 sodium phosphate, 94 nmol) was added to a 15 mL Falcon tube and stored on ice. Dithiothreitol (DTT) (3.6 mg, 23 µmol) was added as a solid to reduce any contaminant disulfide. The 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 °C for 30 min. LC-MS analysis revealed full conversion to the allylated product. (calculated mass = 26755, observed mass = 26755). The reaction mixture was passed through a PD10 column previously equilibrated with pH 5.0 sodium phosphate (50 mM). The product was split into 100 µL aliquots and flash-frozen. ESI-MS are shown below.



Kinetic Studies for rate of CM between SBL-156Seac/SBL-156Sac and allyl alcohol

General procedure for determining pseudo-1st-order rate constant for protein cross-metathesis

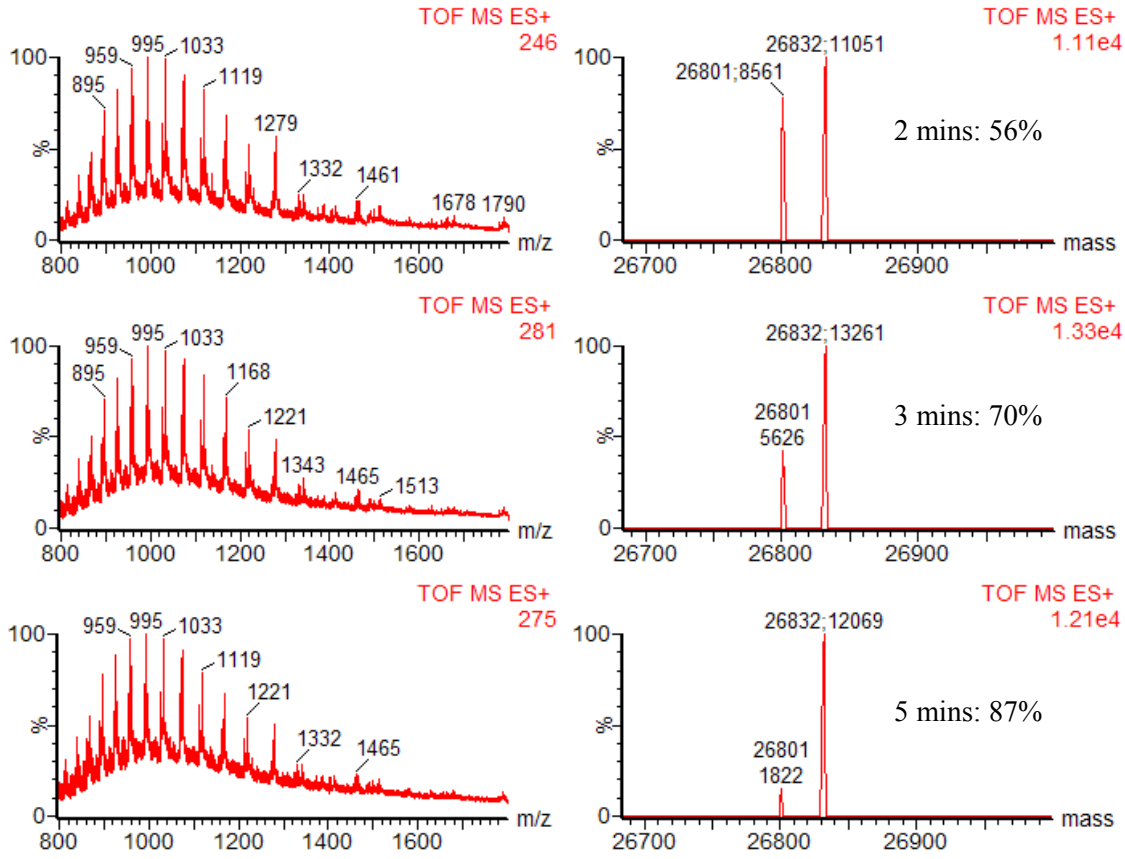


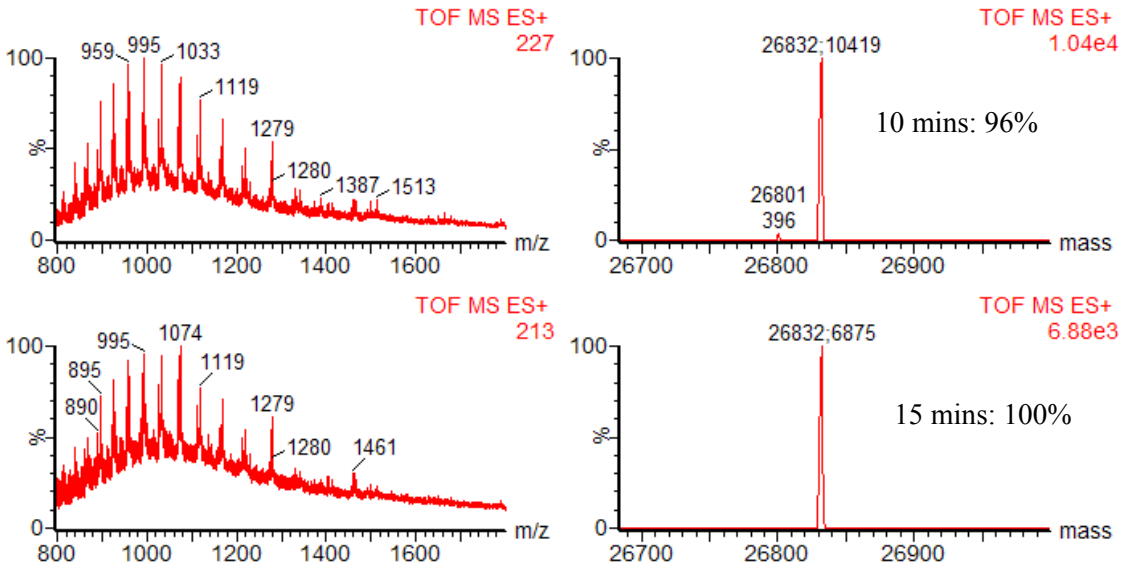
A solution of Hoveyda-Grubbs 2nd generation catalyst (**4.9**) in ^{t}BuOH was prepared by vortexing and gentle warming 0.76 mg of **4.9** in 363 µL *t*BuOH. To a solution of SBL-156Seac (**4.5a**) or SBL-156Sac (**4.8a**) (10.4 µM, 100 µL of a 0.4 mg/mL solution) in 50 mM sodium phosphate buffer (pH 5) was added $MgCl_2.6H_2O$ (0.7 µL of a 1 mg/µL solution in water, 3.7 µmol) followed by the catalyst/*t*BuOH solution (43 µL, 0.15 µmol). The reaction was mixed on a vortex after each addition. The reaction mixture appeared as a green emulsion. Reaction was shaken at RT for 1 minute before adding allyl alcohol (0.5 µL, 7.5 µmol). Reaction was mixed on a vortex immediately after addition and was shaken at RT. 20 µL from reaction was taken at 2, 3, 5, 10 and 15 minutes in the case of SBL-Seac, and 30 minutes, 1, 2, 3 and 4 hours in the case of SBL-Sac. The aliquots were immediately passed through a PD SpinTrap (G-25

Sephadex® superfine) (GE Healthcare) to remove small molecules thus quenching the reaction. The resulting solution was analyzed by LC-MS. The reaction was carried out in triplicate to obtain three data points for each time point. The LC-MS data for one set of time points is shown below. The number under the mass observed indicates the intensity of the peak and is used to calculate the percentage of product conversion. As no significant protein precipitation was observed during the reaction, the percentage of product conversion is directly converted to the product concentration in the reaction mixture. The reaction was assumed to be pseudo-first order with respect to the protein and the data was fitted to the rate equation $\mathbf{A_t = A_0\ (1-e^{-kt})}$ using the data analysis software, GraFit 7.0 (Erithacus Software Limited), to obtain pseudo-first-order rate constants.

LC-MS Data for CM between SBL-156Seac (4.5a) and Allyl alcohol

Calculated mass for Seac CM product **4.5b**: 26832; Observed mass: 26831.





Time (s)	[CM product **4.5b**] (μM)
0	0.00
120	5.41; 5.21; 5.82
180	6.45; 6.33; 7.28
300	9.15; 8.27; 9.05
600	9.98; 9.15; 9.98
900	10.4; 10.4; 10.4

Table S1. SBL-156Seac (**4.5a**) and allyl alcohol CM product concentrations collected in triplicate at various time points

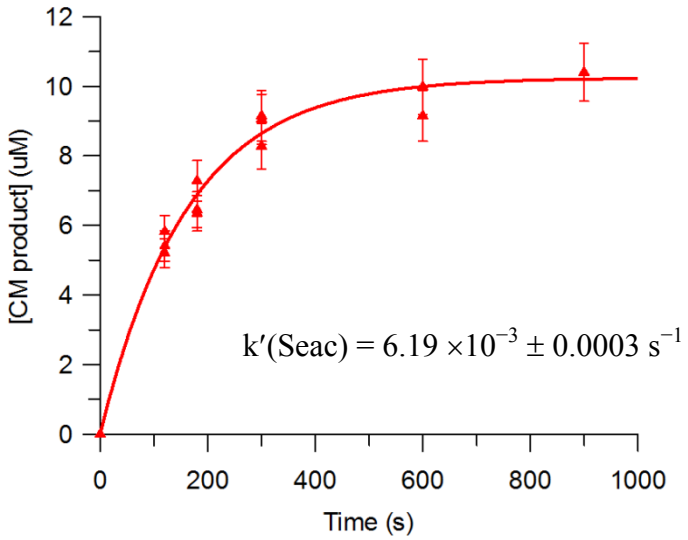
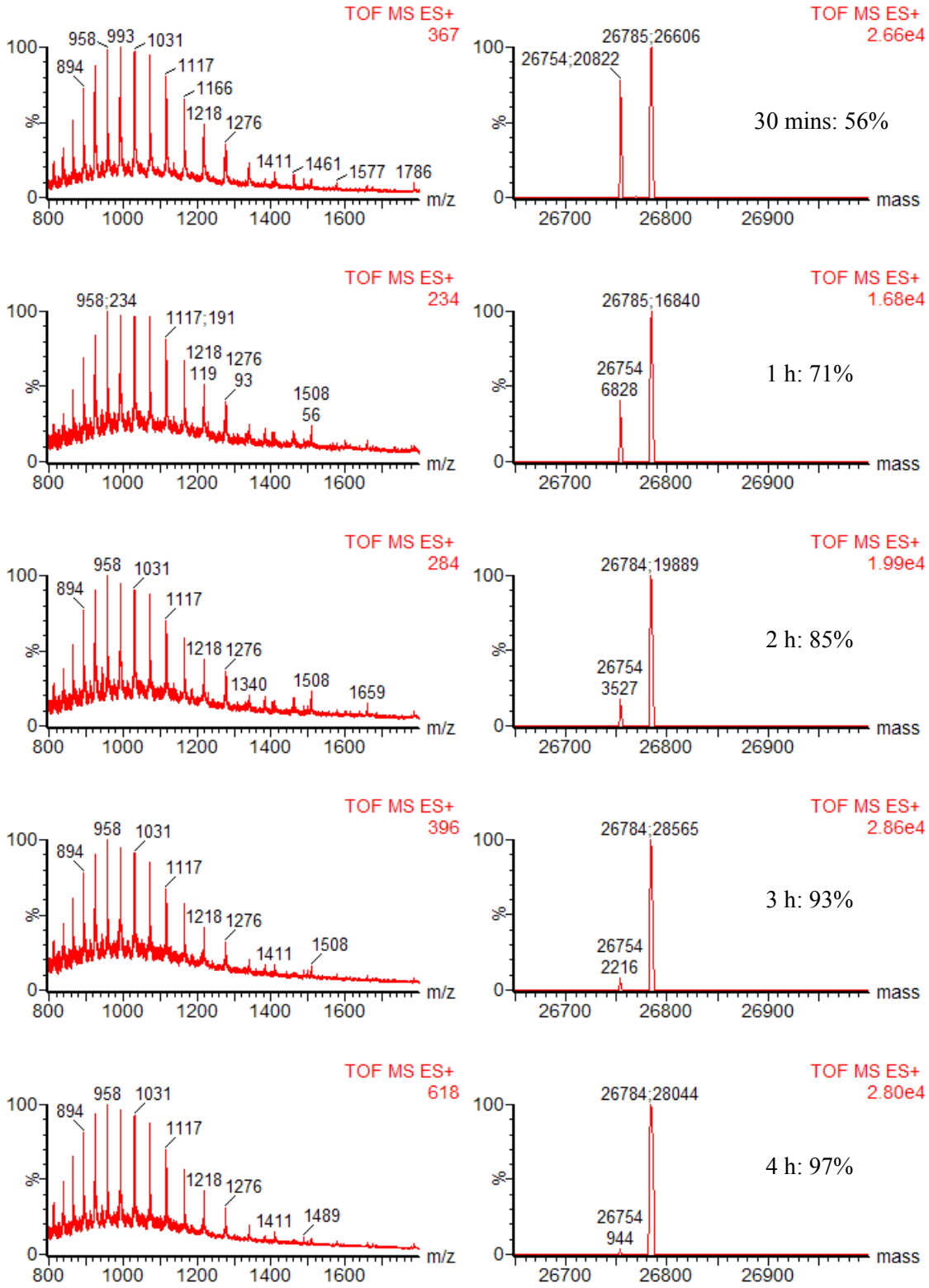


Figure S1. Graph of change in concentration of SBL-Seac CM product **4.5b** as a function of time

LC-MS Data for CM between SBL-156Sac (4.8a) and Allyl alcohol

Calculated mass for Sac CM product **4.8b**: 26785; Observed mass: 26785.



Time (s)	[CM product **4.8b**] (μM)
0	0.00
1800	6.24; 5.82; 5.82
3600	8.74; 7.38; 8.42
7200	9.67; 8.84; 9.88
10800	9.98; 9.67; 10.3
14400	10.19; 10.09; 10.3

Table S2. SBL-156Sac (**4.8a**) and allyl alcohol CM product concentrations collected in triplicate at various time points

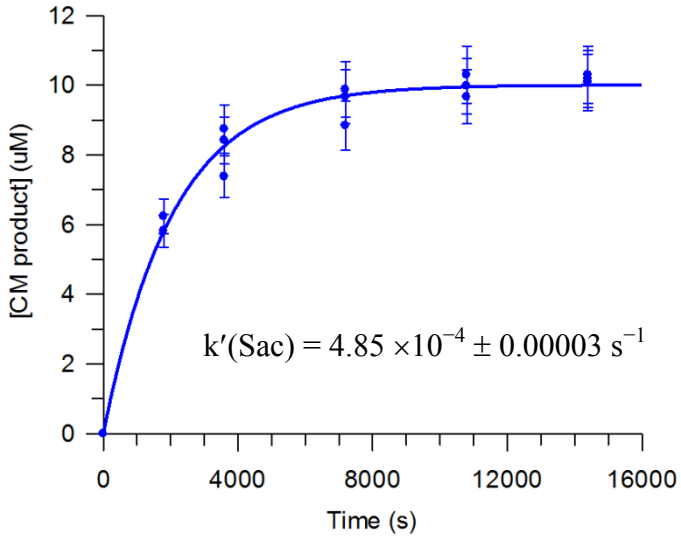


Figure S2. Graph of change in concentration of SBL-Sac CM product **4.8b** as a function of time

Determination of second-order rate constant for protein cross-metathesis

Cross metathesis between SBL-156Seac/SBL-156Sac and different concentrations of allyl alcohol was carried out according to the general procedure for determining the pseudo-first-order rate constant. The data collected were shown in the table below. The first order rate constants were then plotted against allyl alcohol concentration and subjected to a linear fit to obtain the second order rate constant.

Kinetic data for CM of SBL-156Seac (4.5a):

Time (s)	4 mM AllylOH	8 mM AllylOH	10 mM AllylOH	16 mM AllylOH
0	0.00	0.00	0.00	0.00
120	1.14	2.25	2.50	3.54
180	1.35	2.60	3.22	4.47
300	2.18	3.12	4.26	5.93
600	2.91	5.20	6.45	7.28
900	4.58	6.24	7.07	7.59

Table S3. Concentrations of CM product **4.5b** at various time points for CM carried out at the stated concentration of Allyl alcohol

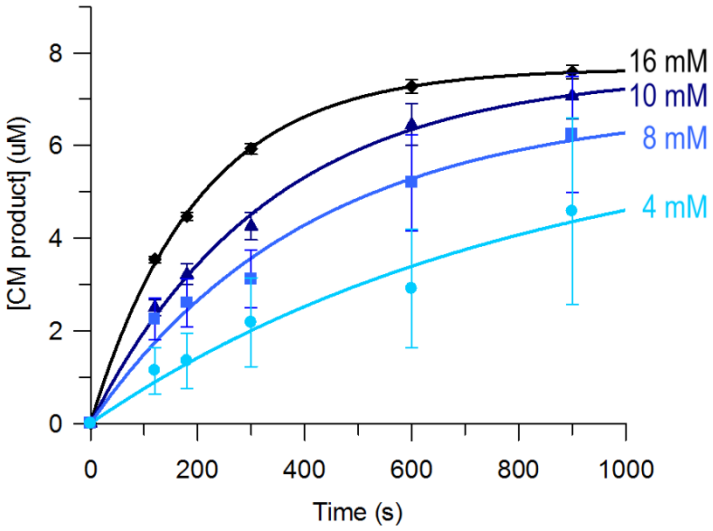


Figure S3. Graph of the change in the concentration of SBL-Seac CM product **4.5b** concentration as a function of time at various concentration of allyl alcohol

[AllylOH](mM)	observed rate k' (s^{-1})
0.0	0
4.0	0.00122 ± 0.00054
8.0	0.00244 ± 0.00050
10.0	0.00301 ± 0.00022
16.0	0.00498 ± 0.00009

Table S4. The corresponding observed rate k' for SBL-156Seac (**4.5a**) CM at different concentrations of Allyl alcohol

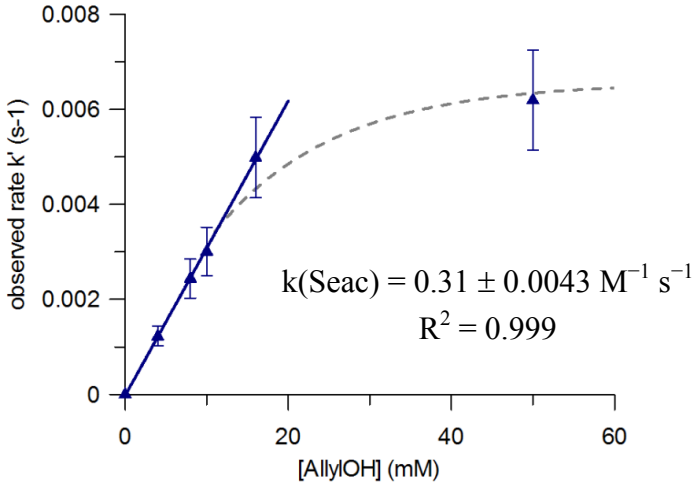


Figure S4. Determination of second-order rate constant for CM between SBL-156Seac (**4.5a**) and allyl alcohol

Kinetic data for CM of SBL-156Sac (4.8a):

Time (s)	2 mM AllylOH	8 mM AllylOH	10 mM AllylOH	15 mM AllylOH
0	0.00	0.00	0.00	0.00
1800	0.73	2.08	3.12	3.43
3600	0.83	2.6	4.47	5.20
7200	1.87	4.47	6.14	6.14
10800	2.39	4.47	6.55	6.76
14400	2.81	5.10	6.86	6.97

Table S5. Concentrations of CM product **4.8b** at various time points for CM carried out at the stated concentration of Allyl alcohol

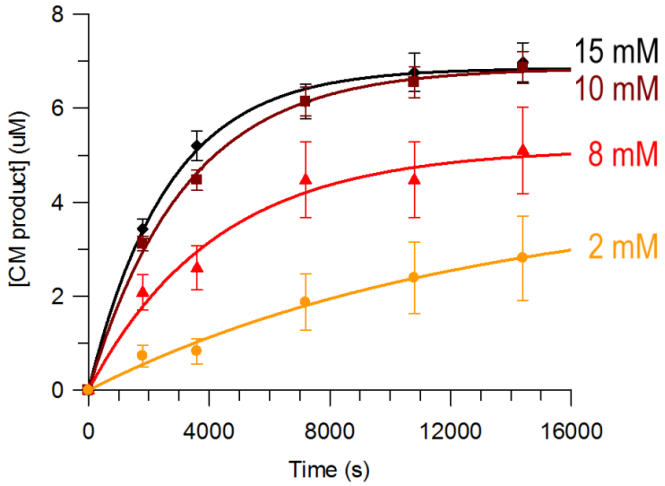


Figure S5. Graph of the change in the concentration of SBL-Sac CM product **4.8b** concentration as a function of time at various concentration of allyl alcohol

[AllylOH] (mM)	observed rate k' (s^{-1})
0	0
2.0	0.0000786 ± 0.000025
8.0	0.000237 ± 0.000043
10.0	0.000312 ± 0.000016
15.0	0.000380± 0.000023

Table S6. The corresponding observed rate k' for SBL-156Sac (**4.8a**) CM at different concentrations of Allyl alcohol

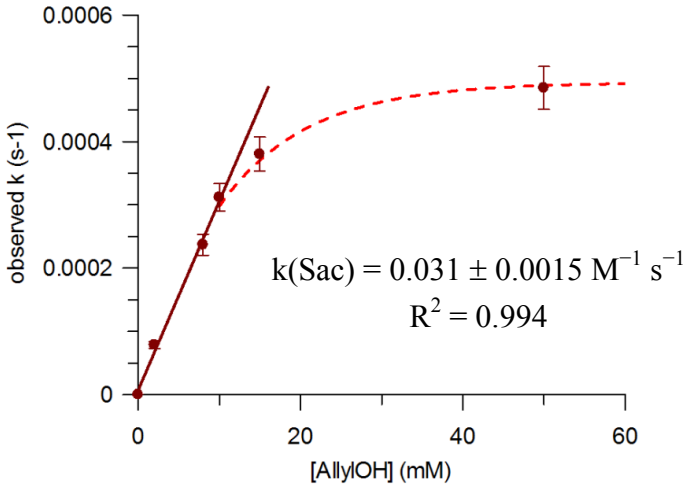
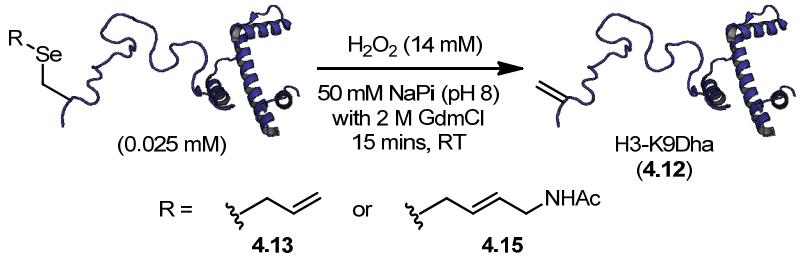


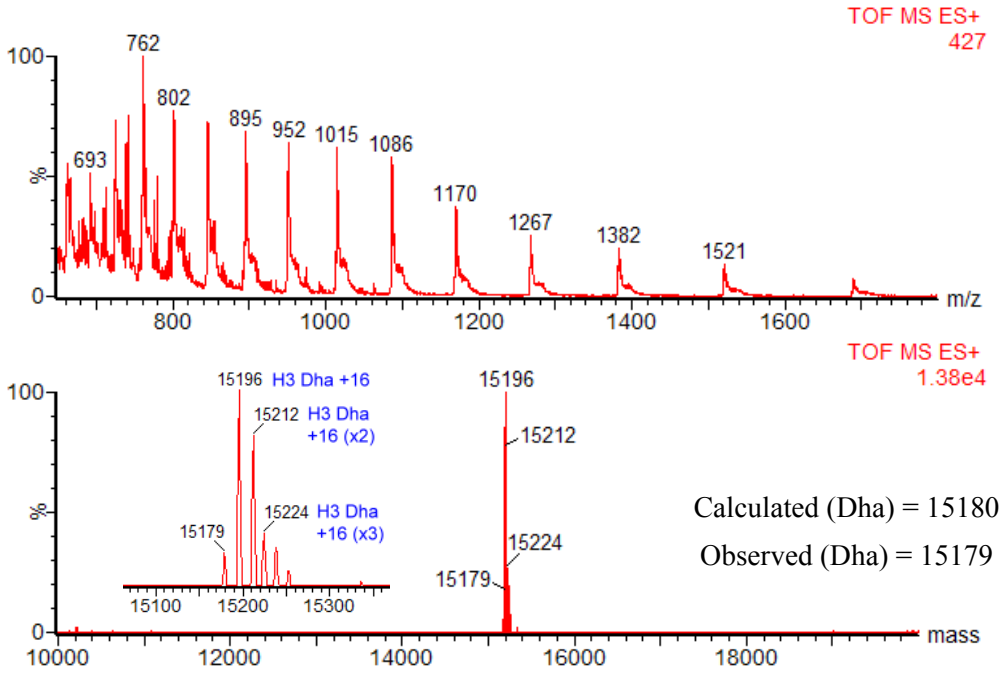
Figure S6. Determination of second-order rate constant for CM between SBL-156Sac (**4.8a**) and allyl alcohol

Oxidative Elimination of H3-K9Seac (4.13) and H3-K9Ac mimic 4.15



To a solution of H3-K9Seac (**4.13**) or K3-K9Ac mimic **4.15** (20 µL, c = 0.4 mg/mL, 0.52 nmol) in 50 mM sodium phosphate (pH 8, with 2 M GdmCl) was added hydrogen peroxide (1 µL of 1% solution in water, 0.29 µmol). The reaction was mixed on a vortex and then incubated at room temperature for 15 minutes with shaking. After this

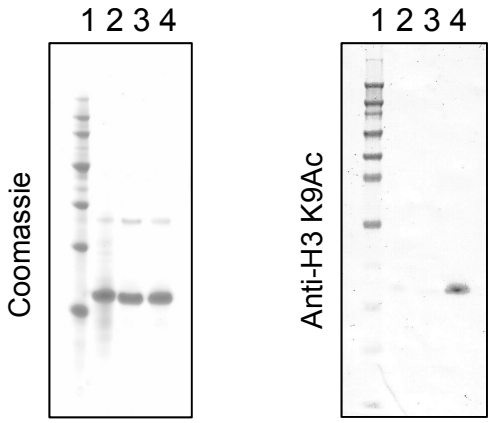
time, an aliquot was analyzed by LC-MS which revealed the oxidative elimination product H3-K9Dha (**4.12**) (calculated mass = 15180, observed = 15180) with some unavoidable over oxidation (on Met, Tyr etc.). Both protein **4.13** and **4.15** eliminate oxidatively to the corresponding Dha. The ESI-MS of the reaction is shown below.



Western Blot for H3 K9Ac Mimic (4.15)

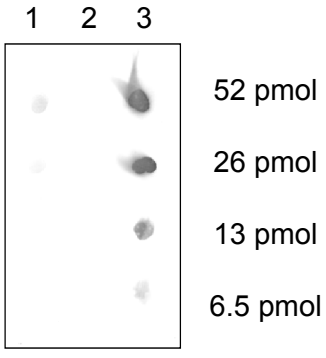
Two aliquots of the H3 K9Ac mimic (**4.15**) were subjected to SDS-PAGE using a 10% NuPAGE pre-cast gel (Invitrogen) and MES running buffer (Invitrogen). One aliquot was visualized using Instant Blue (coomassie) staining. The other sample was used for western blot analysis. Two lanes, one with unmodified H3 (K9C, C110A) and the other H3-K9Seac (**4.13**), were 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 × 2 minutes). The primary antibody (anti-H3-K9Ac, rabbit polyclonal, Abcam ab10812) was diluted from 20 μL to 20 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 × 2 minutes). The secondary antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687) was diluted from 1 μL to 30 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 × 2 minutes) and then water (2 × 2 minutes). 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 (left), and western blot (right) are shown below (Lane 1: protein marker; Lane 2: H3-K9C; Lane 3: H3-K9Seac (**4.13**); Lane 4: H3-K9Ac mimic (**4.15**). Lane 2, 3 and 4 each contain 2.5 μg of protein).



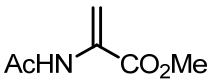
Dot blot for H3 K9Ac Mimic 4.15

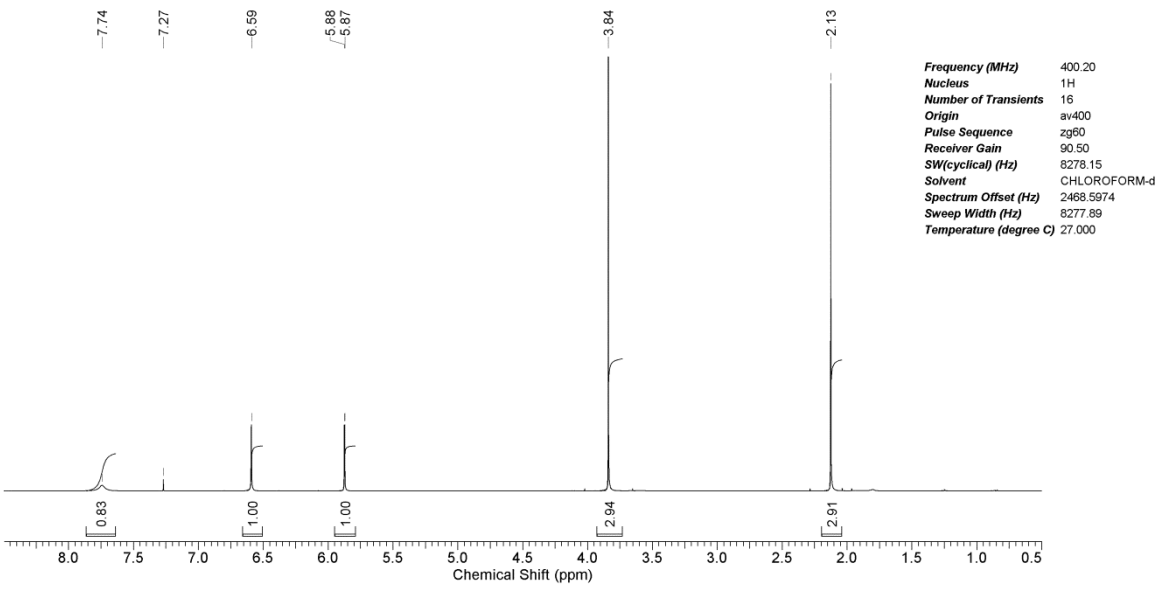
A PVDF membrane (0.2 μm pore size; Invitrogen) was wetted with Methanol, washed with water (2 × 10 mL) and soaked in PBS. 2 μL of serial dilutions of **4.15**, thialysine K9Ac mimic[4.10] and H3 K9C were spotted on the membrane and allowed to dry overnight. The dry membrane was again wetted with Methanol, washed with water (2 × 10 mL) and blocked with BSA (3% w/v; 10 mL) in TBST for an hour at room temperature while rocking. The membrane was washed with TBST (5 × 2 minutes). The primary antibody (anti-H3-K9Ac, rabbit polyclonal, Abcam ab10812) was diluted from 20 μL (8 μg) to 20 mL in TBST buffer (with 3% w/v BSA) and then incubated with the membrane for 1 hour at room temperature while rocking. After washing with TBST (5 × 2 minutes), the membrane was incubated with the second antibody (anti-rabbit IgG produced in goat, conjugated with alkaline phosphatase, Sigma-Aldrich A3687; 1 μL in 30 mL TBST with 3% w/v BSA) for additional hour at room temperature while rocking. After incubation the membrane was washed with TBST (5 × 2 minutes) and water (3 × 2 minutes). The membrane was then developed using BCIP/NBT liquid substrate system (Sigma-Aldrich B1911). See top right of the page for the blot (Lane 1: K9Ac mimic **4.15**; Lane 2: H3 K9C; Lane 3: K9Ac thialysine mimic).

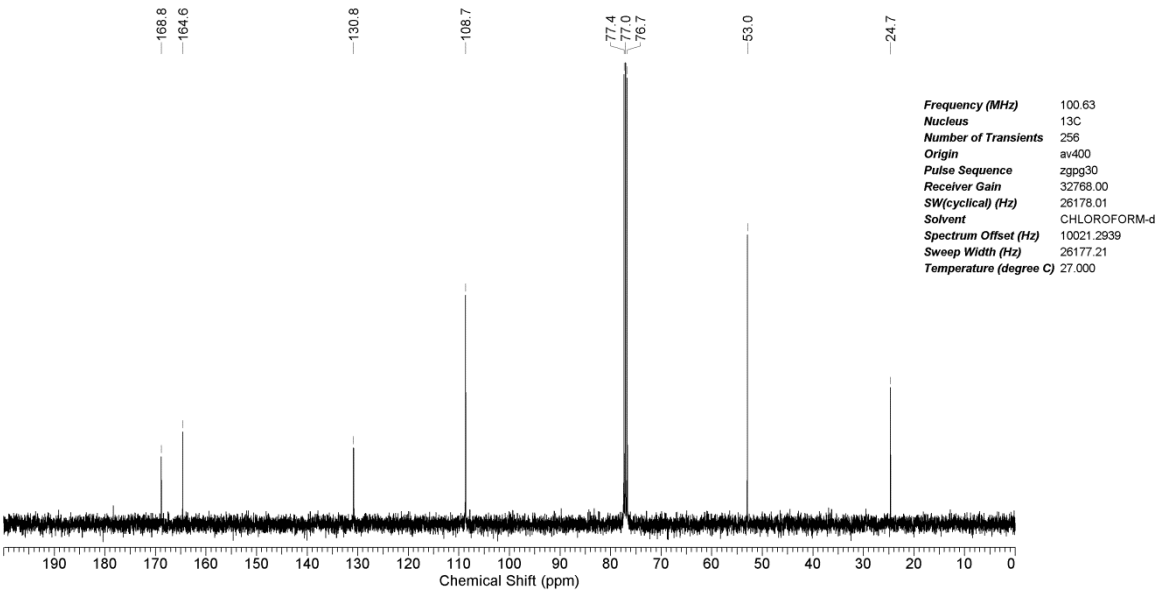


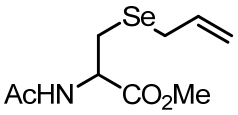
[4.10] The thialysine K9Ac mimic was synthesized according to previously reported literature: Chalker, J. M.; Lercher, L.; Rose, N. R.; Schofield, C. J.; Davis, B. G. *Angew. Chem. Int. Ed.* **2012**, *51*, 1835-1839.

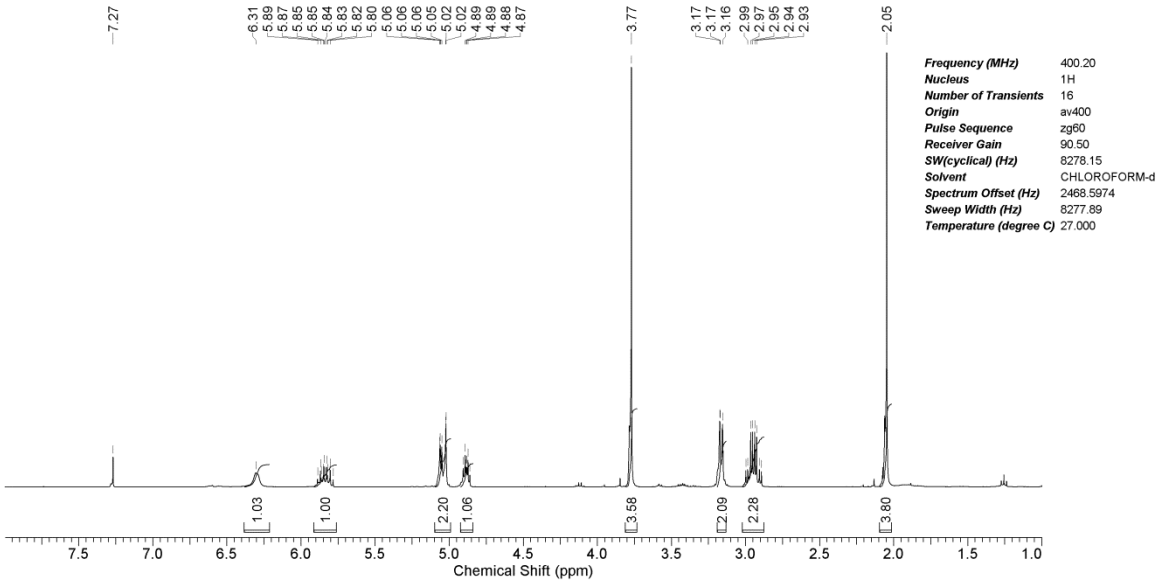
4.6. NMR Spectra

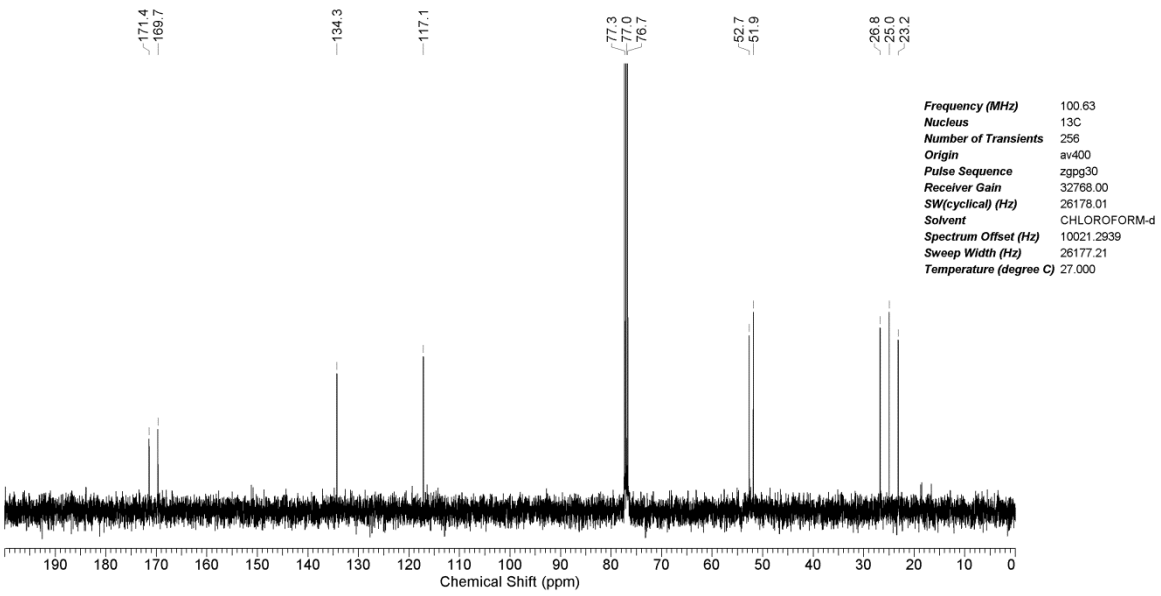






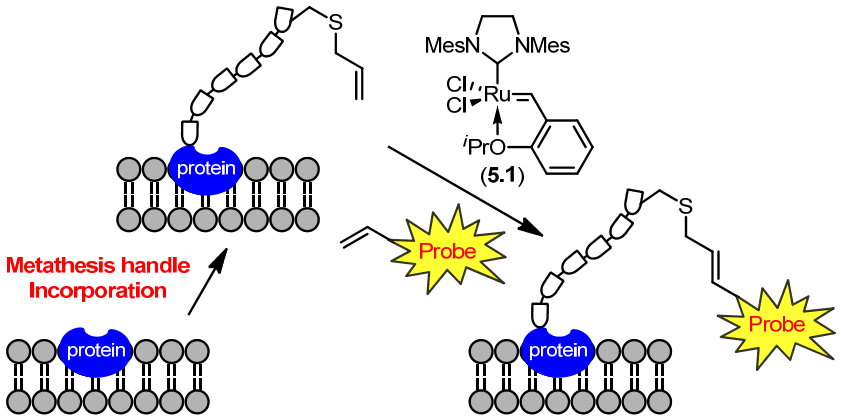

Se
AcHN
CO2Me


7.27
6.31
5.89
5.87
5.85
5.85
5.84
5.83
5.82
5.80
5.06
5.06
5.06
5.05
5.02
5.02
4.89
4.89
4.88
4.87
3.77
3.17
3.17
3.16
2.99
2.97
2.95
2.94
2.93
2.05
1.03
1.00
2.20
1.06
3.58
2.09
2.28
3.80
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
Chemical Shift (ppm)
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


171.4
169.7
134.3
117.1
77.3
77.0
76.7
52.7
51.9
26.8
25.0
23.2
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
Chemical Shift (ppm)
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

Chapter 5

Towards Metabolic Incorporation of Metathesis Handles into Proteins & Olefin Metathesis on Cell Surfaces



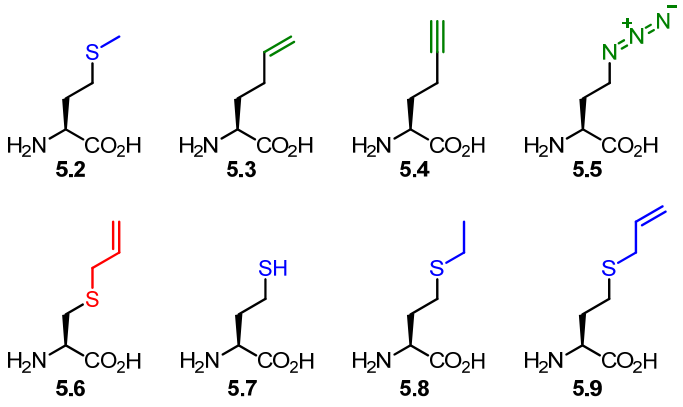
5.1 Introduction

So far, successful cross-metathesis was demonstrated on proteins to attach functional modifications. Various chemical methods of incorporating metathesis reactive tags (allyl sulfides and allyl selenides) have also been developed. The studies on various aspects of olefin metathesis on proteins described in this thesis have expanded its potential applications in chemical biology. This final chapter describes some of the research efforts towards wider application of olefin metathesis in chemical biology. Preliminary results from two research projects will be discussed. The metabolic incorporation of metathesis handles into proteins will be discussed first in section **5.2.1**, and then the work on olefin metathesis on cell surfaces follows on from section **5.2.2**.

5.2 Results and Discussions

5.2.1 Metabolic Incorporation of *S*-Allyl Homocysteine as a Methionine Surrogate[†]

In the preliminary work on genetic incorporation of metathesis handles into proteins, described in the introductory chapter (**1.6.3**), only partial incorporation was observed when *S*-allyl cysteine (Sac) (**5.6**) was used as a methionine surrogate in the expression of a model glycosidase. Nevertheless, Sac incorporation was confirmed by MS-MS analysis of the expressed protein after tryptic digestion. Studies have shown that by hi-jacking the metabolic and translational machinery of *E. coli* unnatural amino acids such as homoallylglycine (**5.3**), homopropargylglycine (**5.4**) and azidohomoalanine (**5.5**), which are useful tags for biocojugation, can be incorporated into proteins as methionine analogues.[1] Encouraged by our initial results and the promiscuity of methionyl-tRNA synthase (Met-RS), the genetic incorporation of amino acids containing metathesis- reactive handles using methionine auxotrophic *E. coli* was further investigated.

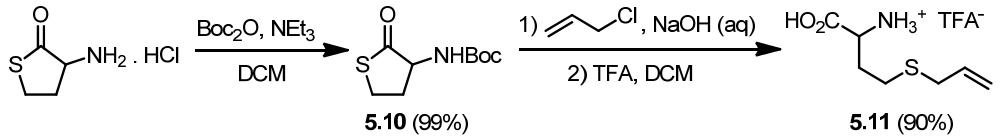


Scheme 5-1: Structures of selected methionine analogues.

[†] This project was carried out in collaboration with Mr. Bhaskar Bhushan, a DPhil student in the group. Experiments carried out by Mr. Bhushan are indicated with reference to this footnote. His contributions in this project are stated explicitly in the experimental section of this chapter.

Fersht and Dingwall have demonstrated that L-ethionine (**5.8**) may be mischarged by the Met-RS without being metabolically edited.[2] In contrast, L-homocysteine (**5.7**), a natural competitor of methionine, is transformed to the corresponding thiolactone to avoid being misread in translation.[2-3] Together with other modeling studies on amino acid binding in Met-RS,[4] we reasoned that the observed low-level incorporation of Sac as Met analogue was due to the bulky sulfur atom located at an unfavorable position for binding to Met-RS. *S*-allyl homocysteine (Ahc) (**5.9**) was envisioned to be a better Met analogue than Sac because the sulfur atom of Ahc is appropriately placed at the same position as that of Met, and only one carbon unit more than L-ethionine.

A short synthesis of DL-Ahc was designed from the readily available homocysteine thiolactone. The protected homocysteine thiolactone **5.10** can be allylated via a one-pot thiolactone opening/allylation strategy using a mixture of allyl chloride and aqueous sodium hydroxide. Following standard Boc deprotection, the TFA salt of DL-Ahc (**5.11**) can be obtained in excellent yields (Scheme 5-2).

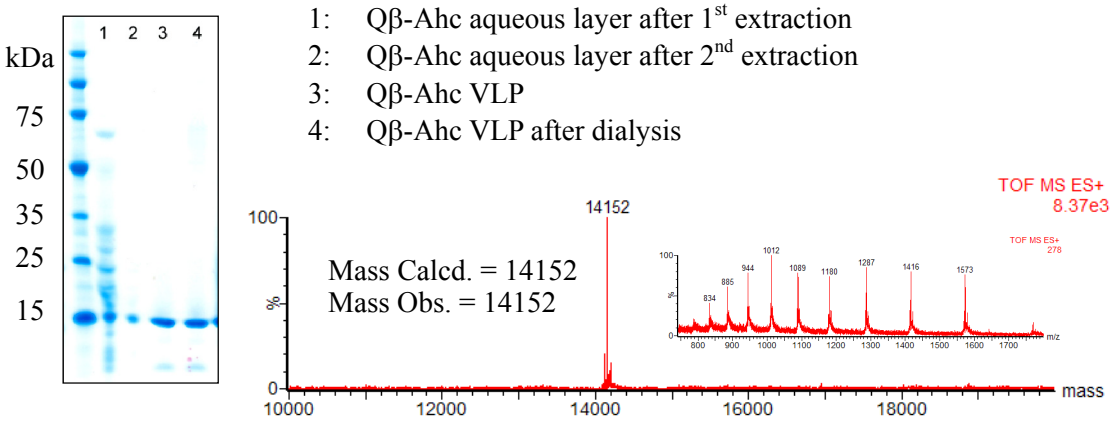


5.10 (99%) **5.11** (90%)

Scheme 5-2: Synthesis of DL-*S*-allyl homocysteine.

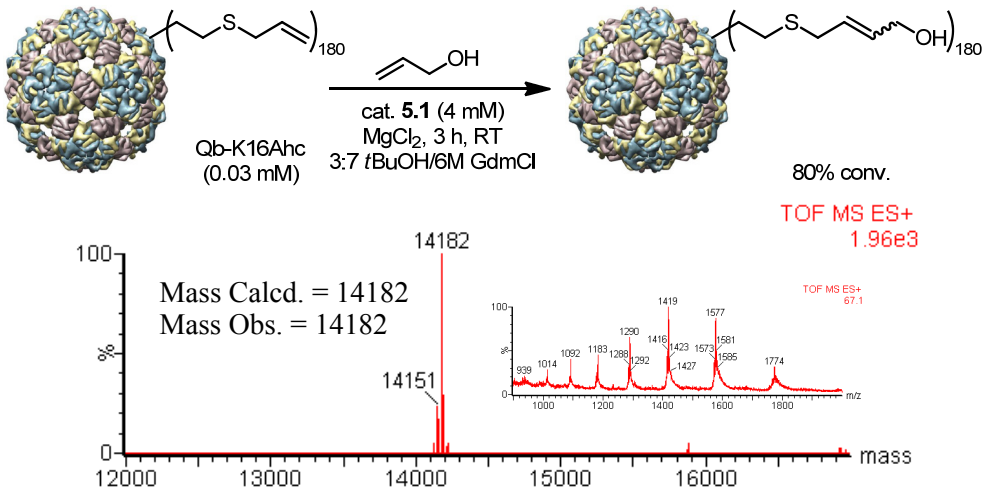
With DL-Ahc in hand, the protein Qβ was chosen as a model protein to investigate the metabolic incorporation of Ahc. Monomeric Qβ assembles into a 180-copy, virus-like particles (VLPs) which has been used in antigen display for therapeutic applications. Incorporation of homoallylglycine,[5] homopropargylglycine[6] and azidohomoalanine[7] into Qβ with methionine auxotrophic *E. coli* is well-established. Accordingly, Ahc incorporation was investigated adapting the reported procedure: methionine auxotrophic strain of *E. coli* (B834(DE3)) containing the gene of single-Met mutant of Qβ-K16M were first grown to an OD_{600} between 0.6 and 0.8 in media supplemented with

methionine. The cells were then pelleted and washed thoroughly with media without methionine and then resuspended in media supplemented with DL-Ahc at a concentration of 1.7 mM. The protein expression was induced by addition of IPTG (1 mM) (see Experimental procedures for details). After purification, Qβ-VLPs were treated with reducing agent to obtain the monomer prior to analysis by SDS-PAGE and LC-MS. The SDS-PAGE showed a clean band at the correct mass range and LC-MS confirmed that Ahc has been success incorporated into Qβ (Scheme 5-3). 1.8 mg of Qβ-K16Ahc was obtained from 625 mL of cell culture. The incorporation of Ahc as Met surrogate at position 16 was further verified by MS-MS analysis after trypsin digestion (see Experimental procedure).



Scheme 5-3: SDS-PAGE and LC-MS analysis of monomeric Qβ-K16Ahc.

Having obtained useful amount of Qβ-K16Ahc, model cross-metathesis reaction was carried out with allyl alcohol in the presence of Hoveyda-Grubbs 2nd generation catalyst (**5.1**). In the initial attempt, protein precipitation was observed when the CM was carried out in sodium phosphate buffer with 30% *t*-BuOH. Conducting CM reactions in guanidinium chloride, in which successful reaction was observed for organic-solvent-intolerant H3, was considered next. Pleasingly, LC-MS analysis after Ru-scavenging with 3-mercaptopropionic acid (3-MP) revealed the CM-modified Qβ with 80% conversion after 3 hours at room temperature (Scheme 5-4).†



Scheme 5-4: Cross-metathesis of Qβ-K16Ahc.

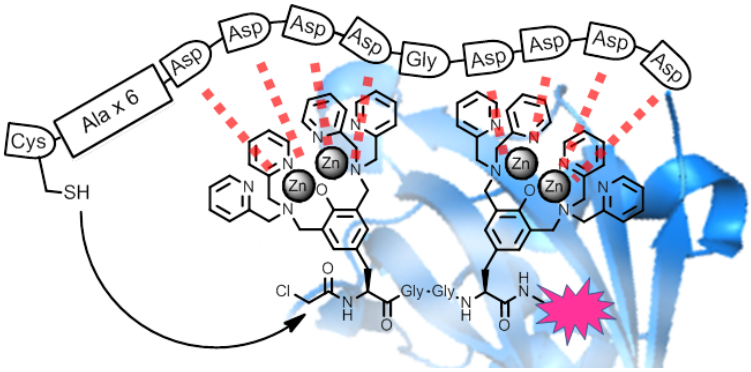
In this section, successful Ahc incorporation into protein was demonstrated. The subsequent CM reaction also proceeded readily under previously optimized conditions. The results presented here are preliminary and are being further investigated in our laboratory (see later in Conclusions and Outlook). Next, our initial research efforts towards olefin metathesis on cell surfaces are discussed.

5.2.2 Towards Olefin-Metathesis on Cell Surfaces[‡]

The Hamachi Laboratory has developed a method to selectively label G-protein coupled receptor (GPCR) proteins on surfaces of HEK293 cells.[8] This approach involves rapid reaction between a short peptide tag (CAAAAAADDDDGDDDD, CA6D4x2) and a tetranuclear Zn(II)-dipicolylamine tyrosine- (DpaTyr) based probe containing a reactive α-chloroacetyl group. In their report, the CA6D4x2 tag is fused on to the outer membrane N-terminus of GPCR such as bradykinin receptor (B2R) and forms a tight tag-probe pair with the probe-appended Zn(II)-DpaTyr complex. This interaction hence

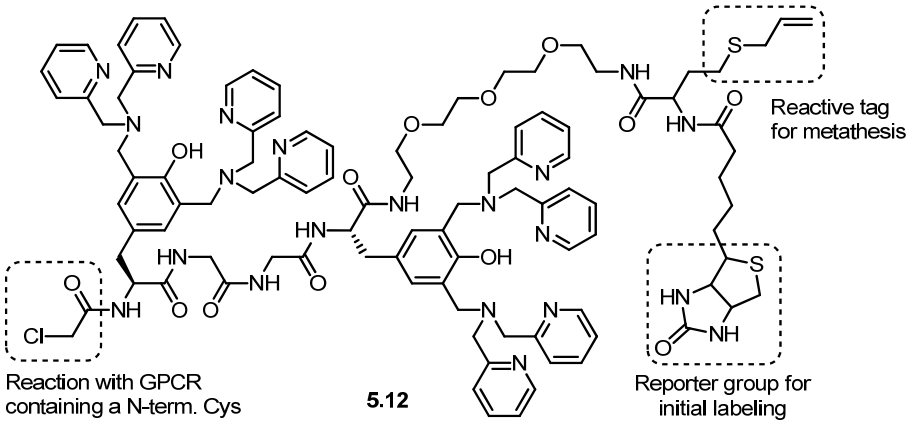
[‡] This project was a three-month-collaborative project supported by JSPS Postdoctoral Fellowship for Foreign Researchers (Short-Term) (October 2011 to January 2012) and was carried out under the supervision of Professor Itaru Hamachi at Kyoto University in Japan.

brings the α-chloroacetyl group and the N-terminal cysteine into close proximity allowing the covalent labeling to take place (Scheme 5-5).



Scheme 5-5: Covalent labeling of tag-fused GPCR with Zn(II)-DpaTyr probe.

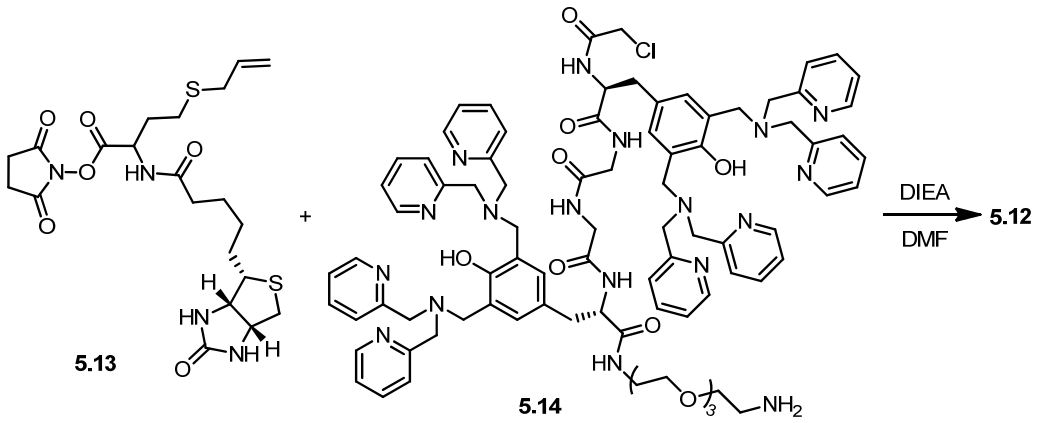
In order to assess CM on cell surfaces, it was envisioned that the covalent labeling described above may be used to introduce a metathesis handle on the cell surface. A secondary probe could then be attached subsequently via CM. The monitoring of this CM reaction could be conducted either by fluorescent microscopy or western blotting. Accordingly, DpaTyr based probe containing a metathesis-reactive handle was designed (Scheme 5-6). In addition to the DpaTyr dimer and α-chloro-acetyl group, required for initial covalent labeling, probe **5.12** also contains an *S*-allyl homocysteine unit for the subsequent CM and a biotin reporter group to verify the first-step labeling.



Scheme 5-6: The design of DpaTyr-based probe **5.12**.

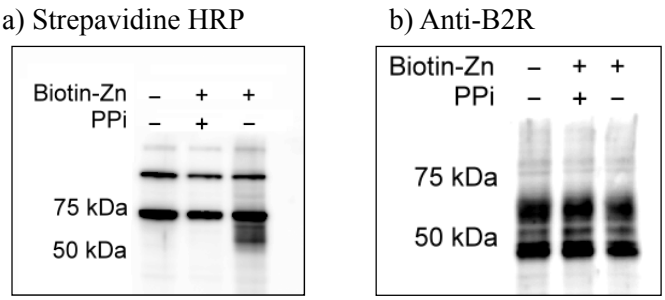
Probe **5.12** was obtained after reaction between fragment **5.13** and **5.14** followed by purification with reverse-phase HPLC (Scheme 5-7). Fragment **5.13** and **5.14** were

synthesized independently from homocysteine thiolactone and BocDpaTyrOH, respectively (see Experimental procedure for detailed synthetic procedures). In order to synthesize the Zn(II) complex for protein labeling, equal volume of **5.12** (2 mM) and $ZnCl_2$ (8 mM) in 50 mM HEPES buffer (pH 7.5) were mixed. The desired zinc complex (**5.12-Zn**) was obtained in quantitative yield as 1 mM solutions.



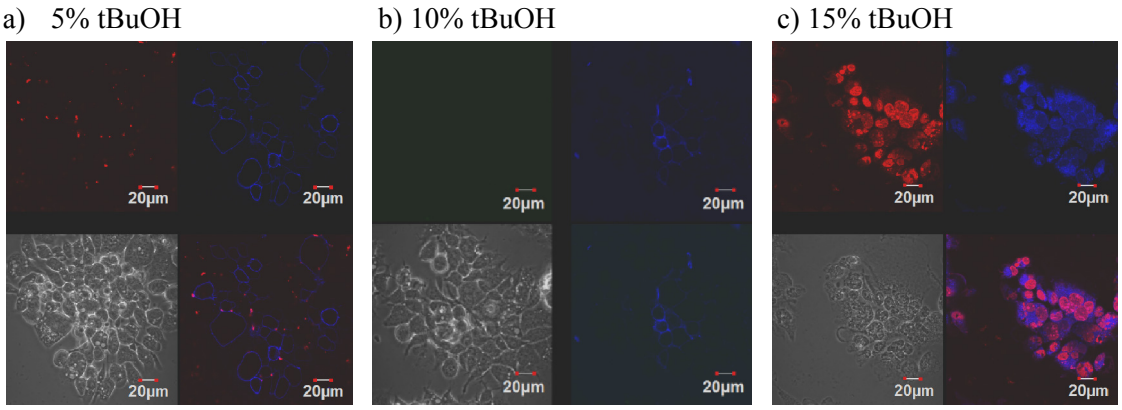
Scheme 5-7: Synthesis of DpaTyr probe **5.12**.

To prepare the initial covalent labeling on cell, HEK239 cells were transfected with CA6D4x2 Tag-fused B2R DNA using Lipofectamine LTX (Invitrogen) according to a general protocol. After 48 hours, HEK293 cells expressing the tag-fused B2R was incubated with a 1 μM solution of **5.12-Zn** complex in 50 mM HEPES (pH 7.5) for 10 minutes at room temperature (see experimental section). The labeled cells were harvested, lysed and then analyzed by western blotting. The western blot showed successful labeling of B2R with **5.12-Zn** (Scheme 5-8a, Lane 3). The smear between 50 and 75 kDa indicated the over expressed B2R on the cell membrane. In the presence of pyrophosphate, the coordination of zinc complex is perturbed and indeed no labeling was observed (Scheme 5-8a, Lane 2). The anti-B2R blot shown in Scheme 5-8b indicated that all three cell lysate samples indeed expressed B2R and at similar concentrations.



Scheme 5-8: Western blotting analysis of B2R labeling by **5.12-Zn** (Biotin-Zn).

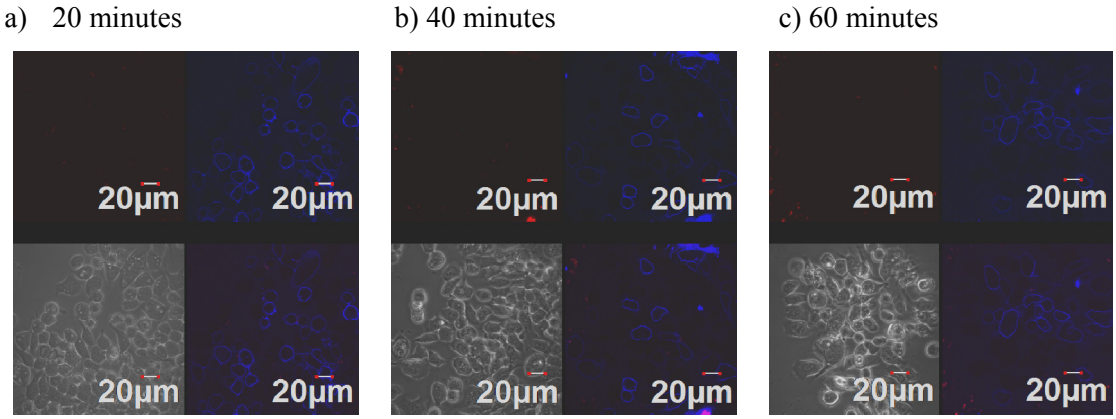
HEK239 cells expressing tag-fused B2R were then subjected to cell viability test with various amount of catalyst **5.1** in *tert*-butanol (1.75 mg/mL). HEK cells were incubated with 5, 10 and 15% of catalyst/*t*-BuOH solution in PBS buffer for 10 minutes. The cells were stained with the excess amount of Cy5-conjugated B2R antagonist peptide (1 μM in final concentration) to check for B2R expression, and propidium iodide (PI) for cell death. The stained cells were then visualized using a confocal laser scanning microscope (CLSM). Images were shown Scheme 5-9.



Scheme 5-9: Cell viability screen at various percentages of *t*BuOH.

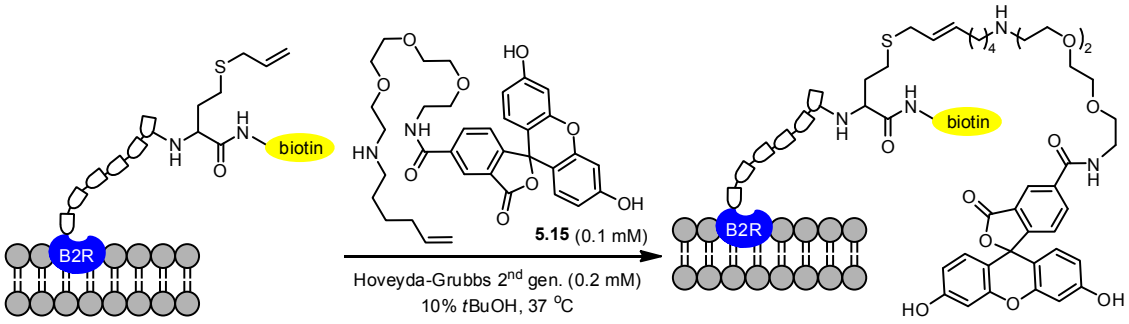
The results showed that the transfection of B2R containing plasmid was successful and that HEK cells can tolerate up to 10% *t*BuOH (Scheme 5-9a, b). At higher percentage of tBuOH, the cells burst allowing PI to stain the nucleus as indicated by the red fluorescence shown in Scheme 5-9c. Hoveyda-Grubbs catalyst is not soluble in water and partially precipitated out of solution at 10% *t*-BuOH in PBS. However, it is hoped that some catalyst is still present in solution and able to perform CM reaction.

Next, the incubation time for metathesis reaction was optimized. The incubation time of HEK cells with 5% catalyst/*t*-BuOH in PBS solution (prepared from 3 mg/mL catalyst solution in *t*-BuOH) was investigated. The cells were stained with Cy5-conjugated B2R antagonist peptide and PI at 20, 40 and 60 minutes and then visualized by CLSM. In all cases, healthy cells present even at 60 minutes as shown by the CLSM images (Scheme 5-10).



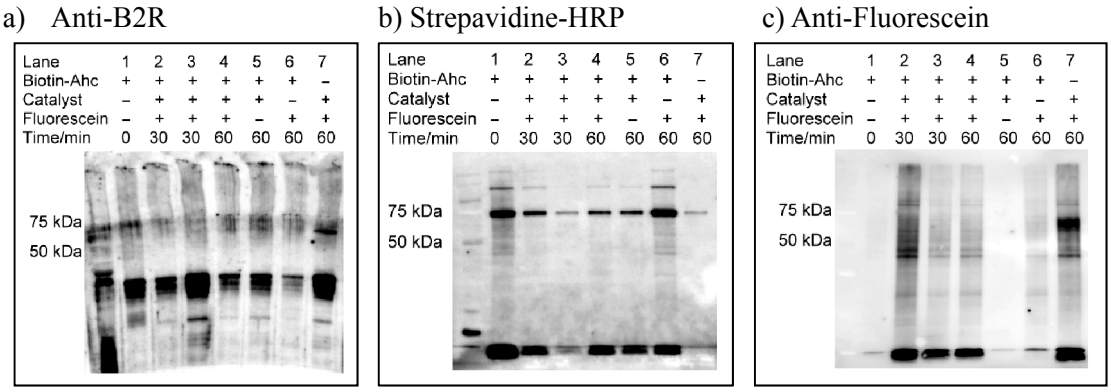
Scheme 5-10: Cell viability screen at various incubation times with 5% **5.1**/*t*BuOH in PBS.

After successful labeling of probe **5.12-Zn**, and optimizing the reaction time and percentage of *t*BuOH, secondary labeling reaction via CM was investigated with catalyst **5.1** and alkene-tethered fluorescein probe **5.15** (Scheme 5.11). A fluorescein-based probe was chosen because it allows the CM reaction to be analyzed by CLSM as well as western blotting (with anti-fluorescein antibody), which is a more sensitive analysis method. The synthesis of **5.15** is detailed in the experimental section.



Scheme 5-11: Secondary labeling of B2R with fluorecein **5.15** via CM.

After the initial labeling of the probe **5.12-Zn** (following procedures described earlier), the cells were washed and 10% catalyst/*t*BuOH solution (2.5 mg/mL) in PBS buffer was added. After 1 minute incubation at room temperature, a solution of fluorescein **5.15** was added and incubated at 37 °C for 30 minutes and 60 minutes. Reaction mixture was then removed and the cell was washed with PBS and harvested. After cell lysis, the samples were then analyzed western blotting (Scheme 5-12).

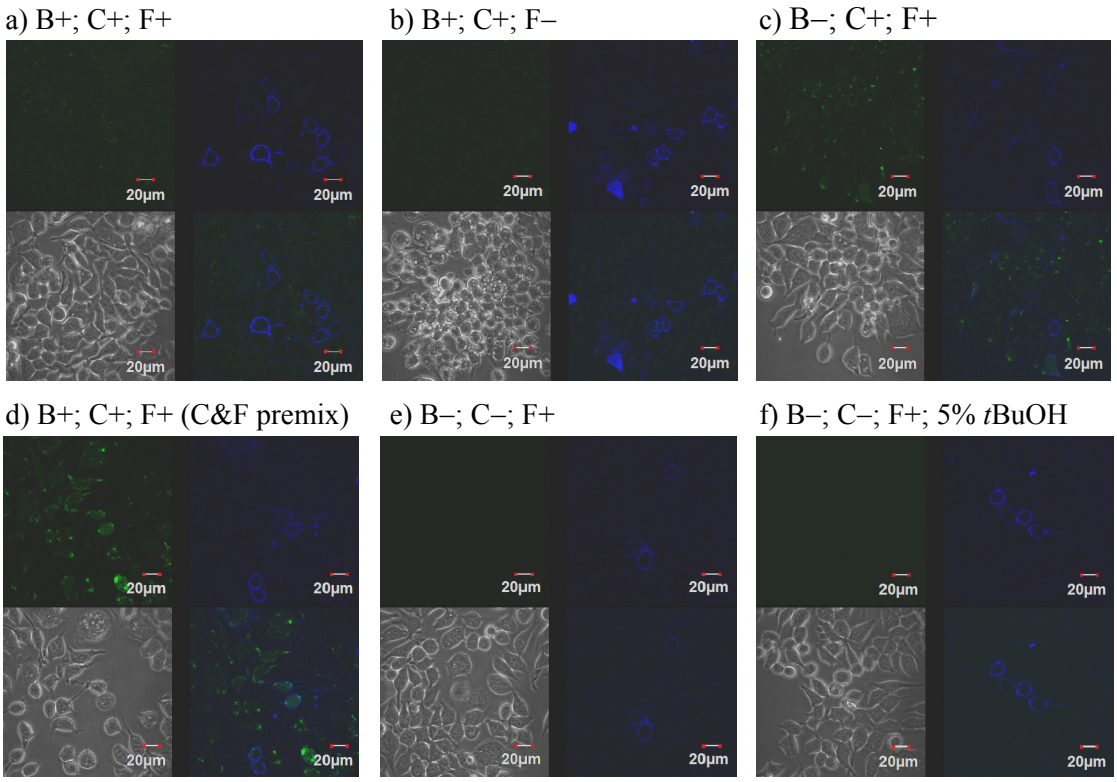


Scheme 5-12: Western Blotting analysis of CM on **5.12-Zn** labeled B2R.

The anti-B2R and strepavidine western blot showed that all samples expressed B2R and the initial covalent labeling by **5.12-Zn** was successful (Scheme 5-12a, b). The intensities of the B2R bands in both blots (between 50 and 75 kDa) were not consistent; this is likely due to the lost of cells from several washing steps. The anti-fluorescein western blot indicated some extent of labeling via CM occurred and 30 minutes of incubation time is sufficient for labeling (no significant difference between Lane 3 and 4 in Scheme 5-12c). However, the result from lane 7 (no initial labeling by **5.12-Zn**) suggested that the fluorescein labeling might not be site-specific and were not due to CM, or that anti-fluorescein antibody binds non-specifically.

We decided to monitor the metathesis reaction by CLSM as an alternative. CM was carried out on cells as described previously. After removing the reaction mixture, the cells were stained with Cy5-conjugated B2R antagonist peptide and then visualized by

CLSM through the appropriate emission filter (Cy5 and fluorescein). The fluorescent images were shown in Scheme 5-13. From the results of the confocal imaging, it was revealed that no labeling via CM had occurred. It could be that at 10% *t*BuOH in PBS there were simply just not enough active catalyst present in solution to for CM to take place. In one case, fluorescein was observed to penetrate into the cell (Scheme 5-13d); the solution from premixing the catalyst with fluorescein probe apparently weakens cell membrane and making it more permeable for organic substances to enter.



Scheme 5-13: CLSM analysis of cross metathesis reaction on cell surfaces. **B** = **5.12-Zn**; **C** = 10% cat. **5.1**/*t*BuOH solution in PBS; **F**: fluorescein **5.15** at 0.1 mM.

While western blotting was found to be a better method to analyze the CM reaction on cell surface in terms of its detection sensitivity, false positive results may be observed due to non-specific binding of antibodies. Therefore necessary control reactions must be carefully conducted in the future to eliminate potential errors. From these preliminary results, it is clear that CM on cell surface still require further optimizations. There might also be additional factors in cellular system that can potentially interfere with olefin

metathesis, for instance, non-desired coordination from the cell surface to catalyst. These questions await further investigations.

5.3 Conclusions and Outlook

The successful incorporation of *S*-allyl homocysteine as a methionine surrogate into proteins is a step towards a more general application of olefin metathesis in chemical biology. Given the promiscuity of the Met tRNA synthase, it is likely that the selenium analogues of Ahc (*Se*-allyl homoselenocysteine, Ahs) could also be a suitable surrogate for methionine providing that the Se-containing amino acid is tolerated by metabolism. This would provide alternate access to allyl selenide moiety in proteins allowing remarkable reactivity of Se-relayed CM to be explored in chemical modification of proteins.

From the initial efforts of CM on cell surfaces, a metathesis-reactive handle (allyl sulfide) was successfully installed on the surfaces of HEK293 cells but the subsequent CM proved to be difficult. These preliminary results highlighted the current limitation of CM for bioconjugation using conventional metathesis catalyst. For efficient bioconjugations on cellular system, the homogeneity of reaction mixture (in aqueous media) and rate of the conjugation is important. With the use of novel water-soluble catalyst, together with increased understanding on various aspects of olefin-metathesis for protein modification (the use of an appropriate chemical spacer and/or more reactive allyl selenide), developed and discussed in this thesis, we anticipate that CM will become a viable strategy for labeling proteins in cellular systems.

5.4 References

(1) (a) van Hest, J. C. M.; Tirrell, D. A. *FEBS Letters* **1998**, *428*, 68-70. (b) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **2000**, *122*, 1282-1288.

(2) Fersht, A. R.; Dingwall, C. *Biochemistry* **1979**, *18*, 1250-1256.

(3) (a) Jakubowski, H.; R.Fersht, A. *Nucleic Acids Research* **1981**, *9*, 3105-3117. (b) Jakubowski, H. *Proc. Natl. Acad. Sci.* **1990**, *87*, 4504-4508. (c) Jakubowski, H. *Journal of Biological Chemistry* **1995**, *270*, 17672-17673.

(4) (a) Serre, L.; Verdon, G.; Choinowski, T.; Hervouet, N.; Risler, J.-L.; Zelwer, C. *J. Mol. Biol.* **2001**, *306*, 863-876. (b) Datta, D.; Vaidehi, N.; Zhang, D.; Goddard, W. A., III *Protein Sci.* **2004**, *13*, 2693-2705.

(5) (a) Floyd, N.; Vijayakrishnan, B.; Koeppe, J. R.; Davis, B. G. *Angew. Chem. Int. Ed.* **2009**, *48*, 7798-7802. (b) 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.

(6) (a) Boutureira, O.; D'Hooge, F.; Fernandez-Gonzalez, M.; Bernardes, G. J. L.; Sanchez-Navarro, M.; Koeppe, J. R.; Davis, B. G. *Chem. Commun.* **2010**, *46*, 8142-8144. (b) Ribeiro-Viana, R.; Sánchez-Navarro, M.; Luczkowiak, J.; Koeppe, J. R.; Delgado, R.; Rojo, J.; Davis, B. G. *Nat. Commun.* **2012**, *3*, 1303.

(7) Strable, E.; Prasuhn, D. E.; Udit, A. K.; Brown, S.; Link, A. J.; Ngo, J. T.; Lander, G.; Quispe, J.; Potter, C. S.; Carragher, B.; Tirrell, D. A.; Finn, M. G. *Bioconjugate Chem.* **2008**, *19*, 866-875.

(8) Nonaka, H.; Fujishima, S.-h.; Uchinomiya, S.-h.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc.* **2010**, *132*, 9301-9309.

5.5. Experimental Procedures

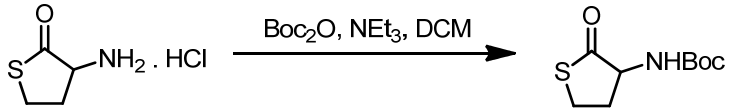
Experimental Procedures for Ahc Incorporation Section

General Considerations & Protein Mass Spectrometry

See Chapter 2 and Chapter 4 Experimental Procedures for general considerations in synthesis and protein analysis methods used in this study, respectively.

Synthesis of DL-*S*-allyl homocysteine

N-Boc-DL-Homocysteine thiolactone (5.10)

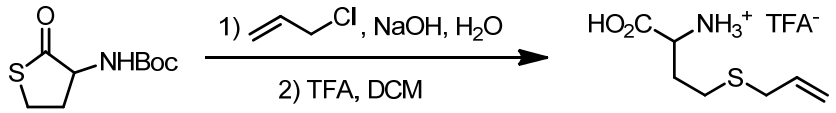


DL-Homocysteine thiolactone hydrochloride (3.0 g, 19.5 mmol) was added to a 100 mL round-bottomed flask with a Teflon™-coated stirrer bar and suspended in DCM (50 mL). The flask was put under argon atmosphere, stirred and cooled to 0 °C with an ice bath. Et_3N (5.0 mL, 39.0 mmol) was added portionwise. Boc_2O (5.0 g, 23.4 mmol) was added and the reaction was stirred at room temperature for 4 hours. After this time the reaction was evaporated to dryness with rotary evaporation. The resulting white solid was re-dissolved with EtOAc (150 mL), diluted with H_2O (300 mL) and transferred to a separatory funnel. The organic layer was isolated and the aqueous layer was further extracted with EtOAc (150 mL). The combined EtOAc was washed with brine (200 mL), dried ($MgSO_4$) and then filtered. The solvent was removed under vacuum to give the desired compound as a white solid (4.2 g, 99%). NMR shows the desire product was pure enough without further purification. Spectroscopic data was identical to that previously reported.[5.1] m.p. = 130-133 °C. IR (υ_{max}, solid): 3350, 2978, 1677, 1519, 1368, 1300, 1239, 1157, 1053, 972, 863, 779, 612. ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.45 (9H, s, Boc), 1.98 (1H, qd, J = 12.2, 7.5, H_β), 2.72-2.94 (1H, m, H_β'), 3.23 (1H,

[5.1] Llewellyn, D. B.; Wahhab, A. *Tetrahedron Lett.* **2009**, *50*, 3939-3941.

ddd, J = 11.6, 7.1, 1.0, H_γ), 3.31 (1H, td, J = 12.1, 5.1, H_γ'), 4.12-4.45 (1H, m, H_α), 5.01 (1H, br. s., NH). ^{13}C NMR (100 MHz, $CDCl_3$): δ_C = 27.2 (C_γ), 28.3 (Boc), 32.1 (H_β), 60.5 (H_α), 80.4 (Boc, 4°), 155.5, 174.1 (2 × C=O). LRMS m/z (ESI+): Found 240.04 $[M+Na]^+$; $C_9H_{15}NO_3SNa$ requires 240.07.

_S_-Allyl-DL-Homocysteine TFA salt (5.11)



N-Boc-DL-Homocysteine thiolactone (2.0 g, 9.20 mmol) was added to a 100 mL round-bottomed flask and suspended in H_2O (40 mL). The suspension was stirred at room temperature and NaOH (1.5 g, 36.8 mmol) was added as a solid. After NaOH was completely dissolved, allyl chloride (4.0 mL, 46.0 mmol) was added. The emulsion was stirred vigorously at room temperature for 3 hours. After this time, TLC (50% EOAc in petrol with 0.5% AcOH) revealed disappearance of starting material (R_f = 0.6) and formation of the allylated product (R_f = 0.3). Reaction was acidified by adding 1M HCl until pH < 5 and then transferred to a separatory funnel. The product was extracted from the aqueous layer with EtOAc (2 × 100 mL). The organic layer was washed with brine (100 mL), dried with $MgSO_4$, filtered and then concentrated under reduced pressure to yield a light yellow oil. This was re-dissolved in DCM (30 mL) and TFA (6 mL) was then added. The reaction was stirred at room temperature for 2 hours after which TLC (50% EOAc in petrol with 0.5% AcOH) revealed complete consumption of starting material. The reaction was evaporated to dryness under reduced pressure to give a dark oil. The dark residue was diluted with a little bit of H_2O and cooled to 0 °C before adding NH_4OH (25% aq) dropwise until basic (pH > 8, pH paper). During this process a white, very insoluble solid crashed out. Purification by column chromatography (1:2:7 25% NH_4OH (25% aq):MeOH:^{i}PrOH) gave the desired amino acid TFA salt as a white

powder (2.4 g, 90%). m.p. = 225-226. IR (υ_{max}, solid): 2920, 1657, 1578, 1412, 1339, 1192, 1127, 988, 916, 840, 801, 724. ^{1}H NMR (400 MHz, D_2O): δ_H = 1.93-2.14 (2H, m, H_β), 2.49-2.57 (2H, m, H_γ), 3.13 (2H, d, J = 7.1, SCH_2), 3.74 (1H, dd, J = 6.9, 5.7, H_α), 5.06-5.13 (2H, m, $CH=CH_2$), 5.75 (1H, ddt, J = 17.1, 9.9, 7.2, 7.2, $CH=CH_2$). ^{13}C NMR (100 MHz, D_2O): δ_C = 25.7 (C_γ), 30.4 (SCH_2), 33.8 (C_β), 54.3 (C_α), 118.2 ($CH=CH_2$), 134.2 ($CH=CH_2$), 174.5 (C=O). LRMS m/z (ESI+): Found 176.07 $[M+H]^+$; $C_7H_{14}NO_2S$ requires 176.07.

General procedure for expression of Qβ-K16M

Plasmid carrying the M16K Qβ gene and p75M vector was transformed into *E. Coli* B834(DE3) competent cells and was plated on LB-agar supplemented with ampicillin (100 μg/mL) and incubated at 37 °C for 18 hours. A single colony from the plate was used to inoculate SelenoMet media (10 mL) containing ampicillin (100 μg/mL) and supplemented with L-methionine (40 μg/mL). The overnight culture were used to inoculate SelenoMet media (625 mL) containing ampicillin (100 μg/mL) and supplemented with L-methionine (40 μg/mL). The cultures were incubated (37 °C, 180 rpm) and grown to OD_{600} between 0.6 and 0.8. The cells were then pelleted by centrifugation (7,000 rpm, 8 minutes). The pellets were washed with SelenoMet media (3 × 300 mL) containing ampicillin (100 μg/mL) and then transferred to prewarmed (37 °C) SelenoMet media (625 mL) with containing ampicillin (100 μg/mL) and supplemented with amino acid **5.11** (500 mg/L, 1.7 mM) in place of methionine. The culture was incubated for 30 minutes at 37 °C, followed by 1 hour at 30 °C. Expression was induced by addition of IPTG (1 mM) and was continued for 6 hours (37 °C, 180 rpm). The cells were harvested by centrifugation (7,000 rpm, 8 minutes) and were collected and suspended in TBS buffer (20 mM Tris-HCl, 150 mM NaCl, 2 mM EDTA, pH 7.5, 30 mL) and stored at −80 °C.

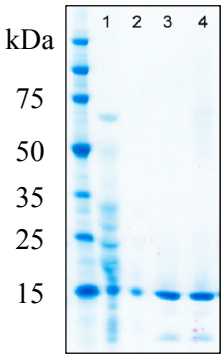
Purification procedure for Qβ-K16M

The cells were thawed and sonicated (Soniprep at 10 micron amplitude, 30 seconds with 1 minute rest, 5 times, on ice) and the cell debris removed by centrifugation (20,000 rpm, 40 minutes). The clear lysate was added to an equal volume of 1:1 n-BuOH:$CHCl_3$ and mixed by inversion. The layers were separated by centrifugation (3750 rpm, swing-out, 30 minutes). The aqueous layer was isolated and the VLPs were precipitated by addition of PEG8000 (10% w/v) and NaCl (100 mM) to the aqueous layer, dissolved by vortexing and the suspension rocked at 4 °C for 30 minutes. The precipitated Qβ particles were collected by centrifugation (3750 rpm, swing-out, 30 minutes) and resuspended in 10 mL TBS buffer with the supernatant discarded. The extraction step was repeated a second time on the resuspended Qβ. The resulting aqueous layer was isolated and dialyzed into 50 mM sodium phosphate buffer (pH 8).

DNA Sequence Data for Qβ-K16Ahc

AKLETVTLGN IGKDGXQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK NYKVQVKIQN PTACTANGSC DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL LIDAIDQLNP AY

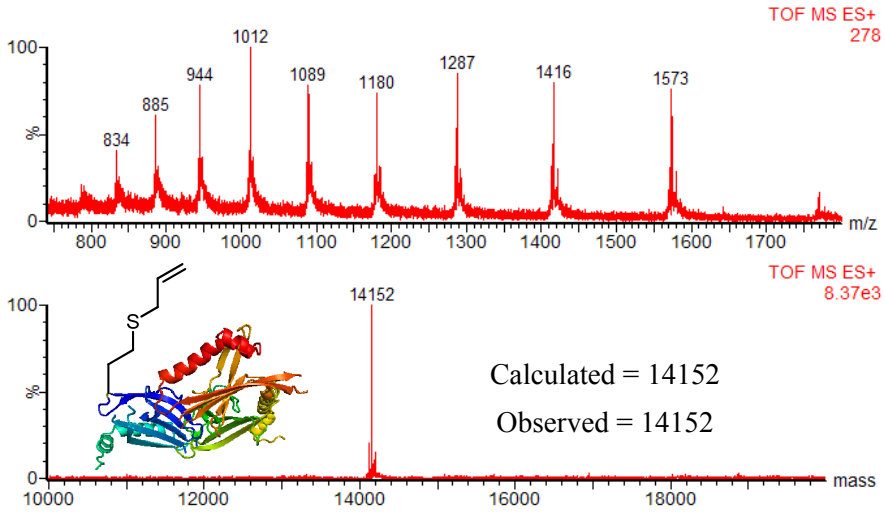
If X = Met, Calculated mass = 14126
If X = Ahc, Calculated mass = 14126 + 26 = 14152



1: Qβ-Ahc aqueous layer after 1st extraction
2: Qβ-Ahc aqueous layer after 2nd extraction
3: Qβ-Ahc VLP
4: Qβ-Ahc VLP after dialysis

Clean band for Qβ K16Ahc was observed on SDS-PAGE, 1.8 mg of the protein was obtained from 625 mL cell culture. The Qβ-VLPs were reduced by adding 2 μL of

0.5M TCEP solution to 50 µL of protein solution (c = 0.2) then small scale dialysis (30 minutes, 500 mL sodium phosphate buffer (50 mM, pH 8)) prior to analysis by LCMS. The LCMS spectra of the protein monomer are shown below.



"In-gel" Tryptic Digestion-MSMS Analysis of Expressed Proteins

"In-gel" tryptic digestions were performed following an MS-compatible protocol. Following SDS-PAGE of the protein, the gel band of the desired protein were excised from the gel, diced into 1 mm cubes and placed in Eppendorf tubes. The bands were washed with 100 µL of Solution A (40 mg ammonium bicarbonate in 10 mL Milli-Q water and 10 mL HPLC grade acetonitrile) for 30 minutes and the supernatant was discarded. The process was repeated 2 times. The gel was then washed with 100 µL acetonitrile until the gel pieces dehydrated and turned white (10 minutes). The supernatant was discarded and the samples were left open until all the gel pieces were dry. 100 µL of 10 mM DTT solution in 25 mM ammonium bicarbonate solution was added and incubated at 37 °C for 30 minutes. Supernatant was discarded. The gel pieces were then washed twice with 25 mM ammonium bicarbonate solution, followed by acetonitrile. 100 µL of 55 mM iodoacetamide solution in 25 mM ammonium bicarbonate solution were then added and the mixture was incubated for 60 minutes in

the dark. The supernatant was then discarded and the gel pieces were washed twice with 100 µL Solution A for 10 minutes. The pieces were then washed with 100 µL acetonitrile until the gel pieces dehydrated and turned white (10 minutes). The supernatant was discarded and the samples were left open until all the gel pieces were dry. 20 µL of reconstituted Promega Sequencing Grade Modified Trypsin solution (1 µg) was then added and the digest was incubated at 37 °C overnight. 1 µL of formic was then added in order to stop the digest, and the supernatant containing the peptides was transferred to a new tube. 50 µL of extraction buffer (20 µL formic acid in 10 mL Milli-Q water and 10 mL HPLC grade acetonitrile) was then added and the mixture was incubated for 30 minutes and the supernatant was pooled with the existing supernatant. The combined supernatant were then concentrated to dryness by SpeedVac and then transferred to MS vials with 10 µL water. The samples were analyzed by MS/MS-MS.

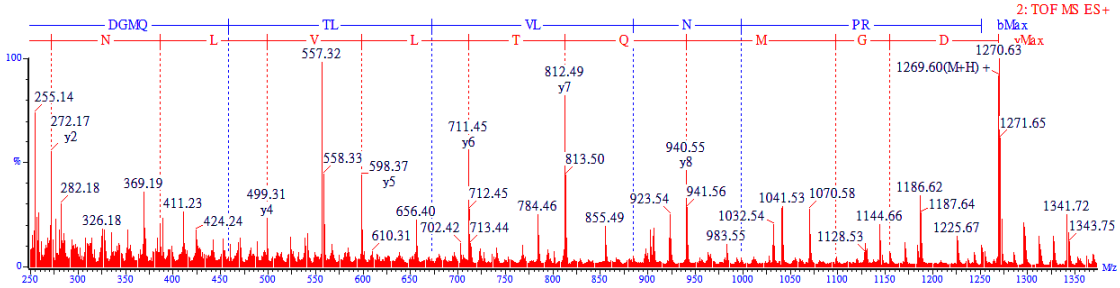
MSMS Data of Tryptic Digested Qβ-K16Ahc

AKLETVTLGN IGKDGMQTLV LNPRGVNPTN GVASLSQAGA VPALEKRVTV SVSQPSRNRK NYKVQVKIQN PTACTANGSC DPSVTRQAYA DVTFSFTQYS TDEERAFVRT ELAALLASPL LIDAIDQLNP AY

The peptide containing the unnatural amino acid Ahc, denoted as "M" in this analysis, after tryptic digestion of Qβ-K16Ahc is highlighted in the sequence above. The MSMS data showed good coverage of the fragmented peptide and received a score of 123 on BioLynx analysis (MassLynx v4.1, from Waters).

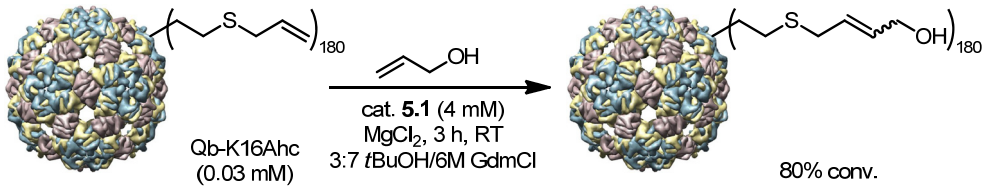
Observed MW: 1268.6522 Precursor ion charge state: 1
M/z tolerance: 0.30 Intensity threshold: 3970 (0.750%)
Modifications: Ahc (+)

a	88.04	145.06	302.12	430.18	531.23	644.31	743.38	856.46	970.51	1067.56	1223.66
	---	---	-0.05	---	---	---	---	---	---	---	---
b	116.03	173.06	330.12	458.18	559.22	672.31	771.38	884.46	998.50	1095.56	1251.66
	---	---	---	-0.04	---	-0.03	---	-0.03	-0.02	---	---
	Asp	Gly	Ahc	Gln	Thr	Leu	Val	Leu	Asn	Pro	Arg
y	1269.67	1154.64	1097.62	940.56	812.50	711.45	598.37	499.30	386.22	272.17	175.12
	---	-0.00	0.01	0.00	0.01	-0.00	-0.00	-0.01	-0.01	0.00	---
z	1252.64	1137.61	1080.59	923.53	795.47	694.42	581.34	482.27	369.18	255.14	158.09
	---	0.02	0.01	-0.01	---	---	---	-0.01	-0.01	-0.00	---

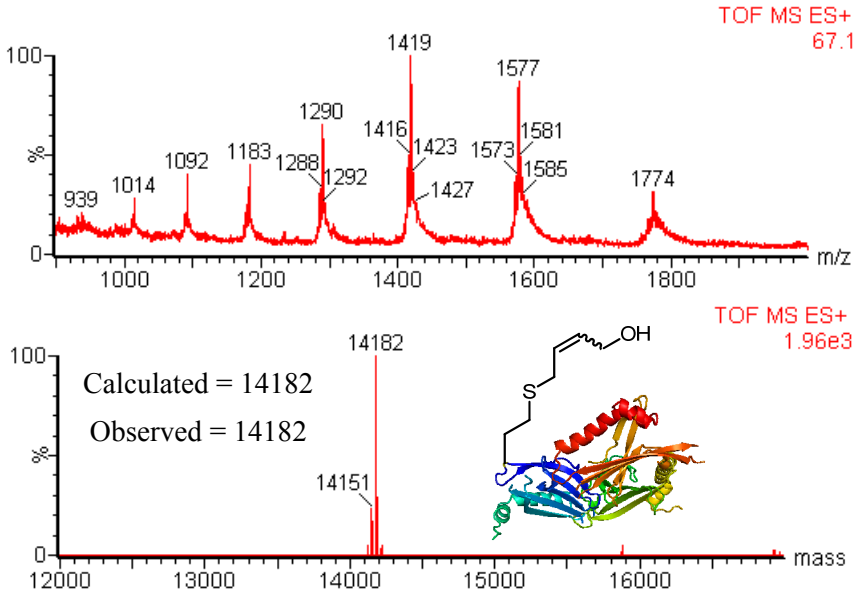


Cross Metathesis on Qb-K16Ahc with allyl alcohol

This experiment was carried out with Bhaskar Bhushan.



A solution of Qβ-K16Ahc (c = 2.0 mg/mL, 35 µL) in 50 mM sodium phosphate buffer (pH 8) was diluted to 700 µL with 6M guanidinium chloride (GdmCl) and concentrated back down with a Vivaspin® (10 kDa MWCO) to a final volume of 100 µL (c = 0.7 mg/mL, 0.03 mM). This process was repeated twice. To this protein solution was added $MgCl_2 \cdot 6H_2O$ (10 µL of a 1 mg/µL solution in water, 49 µmol) and Hoveyda-Grubbs 2nd generation catalyst in *t*BuOH solution (0.4 mg in 43 µL *t*BuOH, 0.64 µmol). The reaction was mixed on a vortex each time after an addition. Reaction was shaken at RT for 1 minute before adding allyl alcohol (1.1 µL, 17 µmol). Reaction was mixed on a vortex immediately after addition and was shaken at RT for 3 hours. 3-mercaptopropionic acid (1.7 µL, 20 µmol) was added as a metal scavenger for ruthenium and the mixture was shaken for further 15 minutes at RT. The protein sample was then purified from small molecules by dilution with 6 M GdmCl (400 µL) and then concentration by Vivaspin® (10 kDa MWCO) to 100 µL. This process was repeated for four times. The protein sample was treated with TCEP before analysis by LC-MS. LC-MS analysis of the protein sample revealed 80% conversion to the desired CM product (calculated = 14182, found = 14182).



Experimental Procedures from Cell Labeling Section[5.2]

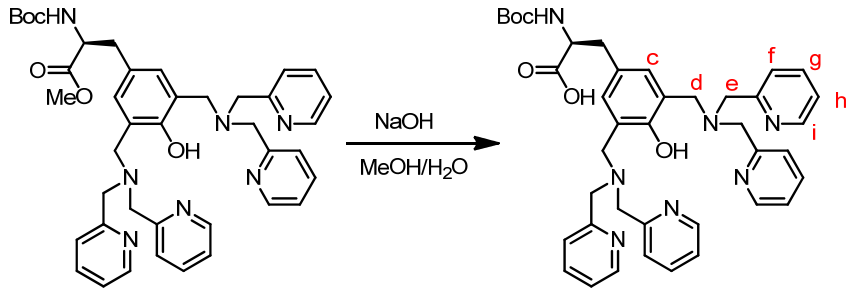
General Considerations

^{1}H NMR spectra were recorded using a Varian Mercury 400 (400 MHz) spectrometer, and the chemical shifts (δ, ppm) are referenced to the respective solvent. The coupling constant, *J*, is reported in the unit of Hertz (Hz). Reverse-phase HPLC was conducted with a Lachrom (Hitachi) instrument with C18 columns. Unless noted otherwise, all chemical reagents and solvents were obtained from commercial suppliers (Sigma-Aldrich, Tokyo Chemical Industry (TCI), Wako Pure Chemical Industries, Acros Organics, Sasaki Chemical, or Watanabe Chemical Industries) and used without further purification. The starting material *N*-Boc-2,2′-dipicolylamine-L-tyrosine methyl ester (BocDpaTyrOMe) was kindly supplied by Mr. Sho-hei Uchinomiya in the Hamachi Laboratory.

[5.2] The experiments in this section were all carried out in the Department of Synthetic Chemistry and Biological Chemistry at Kyoto University, Japan. Due to limited availability of analysis instruments, only ^{1}H NMR was obtained for the compounds synthesized in this section.

Synthesis of Probe 5.12

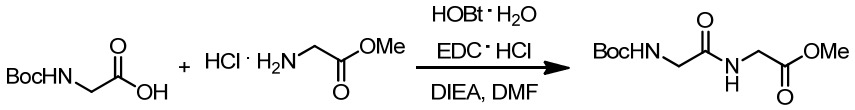
BocDpaTyrOH



BocDpaTyrOMe (1.24 g, 1.73 mmol) was added to a 100 mL round-bottomed flask equipped with a magnetic stirrer and then dissolved in MeOH (20 mL). The solution was stirred and cooled to 0 °C. 1N NaOH solution (5 mL) was then added. The reaction was stirred at 0 °C for 17 hours. After which time TLC analysis ($CHCl_3$:MeOH = 10:1, with few drops of NH_4OH (aq)) revealed complete consumption of starting material (R_f = 0.6) and formation of a product spot (R_f = 0.2). The reaction was diluted with H_2O (100 mL) and transferred to a separatory funnel. The aqueous phase was extracted with Et_2O (50 mL × 2) then acidified to pH 4 with 1N HCl.* The aqueous layer was further extracted with $CHCl_3$ (100 mL × 2). Combined chloroform layers were washed with brine (100 mL), dried (Na_2SO_4) and filtered. The resulting solution was concentrated under reduced pressure to obtain the desired product as a white foam (0.94 g, 77%). Spectroscopic data was identical to that previously reported.[5.3] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.47 (9H, s, Boc), 2.98 (1H, dd, *J* = 13.2, 5.2, Hβ), 3.24 (1H, dd, *J* = 13.2, 4.0, Hβ′), 3.72-3.96 (12H, m, contain H_d and H_e), 4.51-4.53 (1H, m, Hα), 5.22 (1H, d, *J* = 7.6, NH), 7.12 (2H, s, H_c), 7.18 (4H, t, *J* = 6.2, H_h), 7.45 (4H, d, *J* = 8.0, H_f), 7.65 (4H, dt, *J* = 7.6, 1.6, H_g), 8.54 (4H, dd, *J* = 5.2, 0.8, H_i). *pH is very important. It must be 4, if any lower then pyridine groups can protonate and product cannot be extracted.

[5.3] Ojida, A.; Honda, K.; Shinmi, D.; Kiyonaka, S.; Mori, Y.; Hamachi, I. *J. Am. Chem. Soc.* **2006**, *128*, 10452-10459.

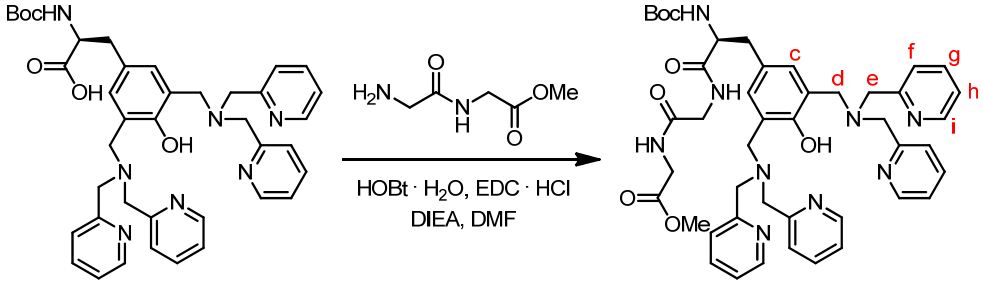
BocGlyGlyOMe



A two-necked, 100 mL round-bottomed-flask was purged with N_2. BocGlyOH (1.67 g, 9.56 mmol) and HGlyOMe hydrochloride (1.0 g, 7.96 mmol) was added and dissolved with anhydrous DMF (30 mL). *N*,*N*′-diisopropylethylamine (DIEA, 6 mL, 43,4 mmol), 1-hydroxybenzotriazole monohydrate (HOBt·H_2O, 1.83 g, 11.9 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl, 2.29 g, 11.9 mmol) was added to the reaction subsequently. Reaction was stirred at room temperature for 3 hours. Reaction was diluted with H_2O (200 mL)* and extracted with EtOAc (2 × 150 mL). Combined EtOAc was washed with saturated bicarbonate solution (2 × 150 mL), brine (150 mL), dried with Na_2SO_4 and then filtered. The resulting solution was concentrated under reduced pressure to yield the product as a thick clear oil (1.7 g, 87%). Spectroscopic data was identical to that previously reported.[5.4] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.46 (9H, s, Boc), 3.76 (3H, s, CO_2Me), 3.86 (2H, d, *J* = 5.6, Gly-CH_2), 4.07 (2H, d, *J* = 5.6, Gly-CH_2), 5.24-5.26 (1H, m, NH), 6.74-6.76 (1H, m, NH).

*heat was generated. If large scale, the solution should be cooled first.

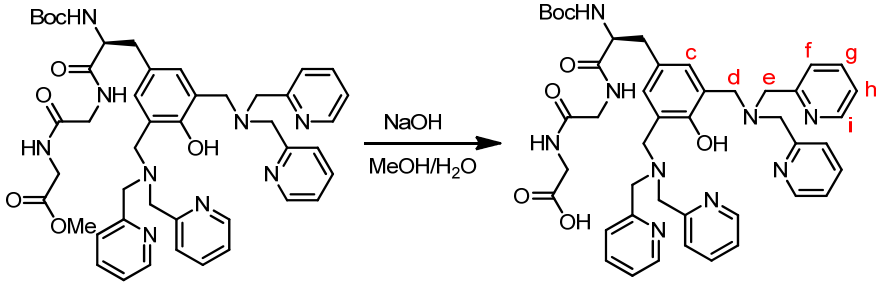
BocDpaTyrGlyGlyOMe



[5.4] Jacobsen, O.; Klaveness, J.; Petter Ottersen, O.; Reza Amiry-Moghaddam, M.; Rongved, P. *Org. Biomol. Chem.* **2009**, *7*, 1599-1611.

A two-necked, 25 mL round-bottomed-flask was purged with N_2. BocGlyGlyOMe (315 mg, 1.28 mmol) was added and dissolved with anhydrous DCM (4 mL). TFA (4 mL) was added dropwise and the reaction was stirred at RT for 1 hour. The reaction was then evaporated to dryness under reduced pressure. The resulting residue was re-dissolved with anhydrous DMF (6 mL) and cooled to 0 °C. *N*,*N*′-diisopropylethylamine (DIEA) was added until pH = 10 and then BocDpaTyrOH (625 mg, 0.89 mmol) 1-hydroxy-benzotriazole monohydrate (HOBt·H_2O, 196 mg, 1.28 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl, 245 mg, 1.28 mmol) were added subsequently. Reaction was stirred at room temperature for 18 hours. Reaction was diluted with H_2O (100 mL) and extracted with EtOAc (2 × 80 mL). Combined EtOAc was washed with saturated bicarbonate solution (2 × 80 mL), brine (100 mL), dried with Na_2SO_4 and then filtered. The resulting solution was concentrated under reduced pressure to yield the product as a light yellow foam (662 mg, 90%). Spectroscopic data was identical to that previously reported.[5.3] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.39 (9H, s, Boc), 2.83-2.90 (1H, m, Hβ), 3.00 (1H, dd, *J* = 13.2, 5.6, Hβ′), 3.53-4.10 (19H, m, contain H_d, H_e, Gly-CH_2, CO_2Me), 4.42-4.48 (1H, m, Hα), 5.35 (1H, d, *J* = 5.6, NH), 7.07 (2H, s, H_c), 7.16 (4H, t, *J* = 6.2, H_h), 7.44 (4H, d, *J* = 8.0, H_f), 7.63 (4H, dt, *J* = 7.6, 1.6, H_g), 8.59 (4H, d, *J* = 4.4, H_i), 11.17 (1H, br. s, OH).

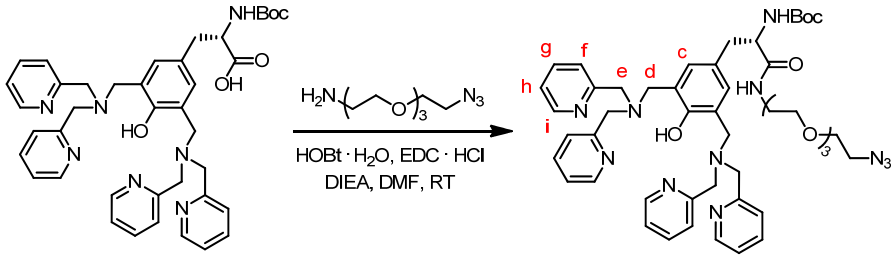
BocDpaTyrGlyGlyOH



BocDpaTyrGlyGlyOMe (520 mg, 0.63 mmol) was added to a 25 mL round-bottomed flask equipped with a magnetic stirrer and then dissolved in MeOH (8 mL). The solution was stirred and cooled to 0 °C. 1N NaOH solution (2 mL) was then added. The reaction was stirred at 0 °C for 18 hours. The reaction was diluted with H_2O (40 mL) and cooled to 0 °C. The solution was adjusted to pH 4* with 1N HCl and then extracted with chloroform (40 mL × 2). Combined chloroform layers were washed with brine (40 mL), dried (Na_2SO_4) and filtered. The resulting solution was concentrated under reduced pressure to obtain the desired product as a white foam. (445 mg, 87%). Spectroscopic data was identical to that previously reported.[5.3] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.36 (9H, s, Boc), 2.89 (1H, dd, *J* = 13.6, 8.4, Hβ), 2.99 (1H, dd, *J* = 13.6, 6.0, Hβ′), 3.80-4.00 (16H, m, contain H_d, H_e and Gly-CH_2), 4.46-4.47 (1H, m, Hα), 5.50 (1H, d, *J* = 7.6, NH), 7.01 (2H, s, H_c), 7.05 (1H, br. s, NH), 7.16-7.19 (4H, m, H_h), 7.47 (4H, d, *J* = 7.6, H_f), 7.64 (4H, dt, *J* = 7.6, 1.6, H_g), 8.55 (4H, dd, *J* = 5.2, 0.8, H_i).

*If extraction failed at pH 4, the extraction is re-attempted at pH 3 or 5.

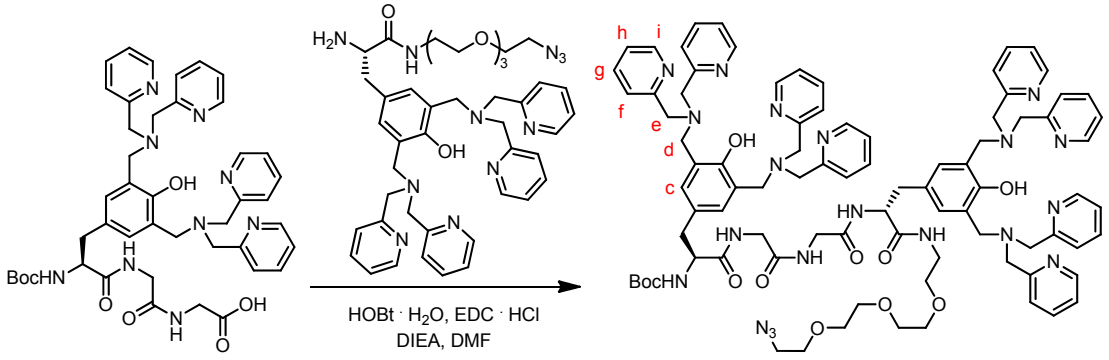
BocDpaTyr-triEG-N_3



A two-necked, 50 mL round-bottomed-flask was purged with N_2. BocDpaTyrOH (250 mg, 0.36 mmol) was added and dissolved with anhydrous DMF (10 mL). 11-azido-3,6,9-trioxaundecan-1-amine (85 µL, 0.43 mmol), *N*,*N*′-diisopropylethylamine (DIEA, 310 µL, 1.78 mmol), 1-hydroxybenzotriazole monohydrate (HOBt·H_2O, 82 mg, 0.53 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC·HCl, 102 mg, 0.53 mmol) was added subsequently. Reaction was stirred at room temperature for 4 hours. Reaction was diluted with H_2O (100 mL) and extracted with

EtOAc (2 × 100 mL). Combined EtOAc was washed with saturated bicarbonate solution (2 × 100 mL), brine (100 mL), dried with Na_2SO_4 and then filtered. The resulting solution was concentrated under reduced pressure to yield the product as a thick clear oil (280 mg, 96%). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.34 (9H, s, Boc), 2.88-2.96 (2H, m, Hβ), 3.23-3.64 (16H, m, contain tri-EG CH_2), 3.68-3.85 (12H, m, contain H_d and H_e), 4.39-4.42 (1H, m, Hα), 5.37 (1H, br. s, NH), 6.95 (1H, br. s, NH), 7.06 (2H, s, H_c), 7.12-7.16 (4H, m, H_h), 7.47 (4H, d, J = 7.6, H_f), 7.63 (4H, dt, J = 7.6, 1.6, H_g), 8.54 (4H, dd, J = 4.8, 0.8, H_i).

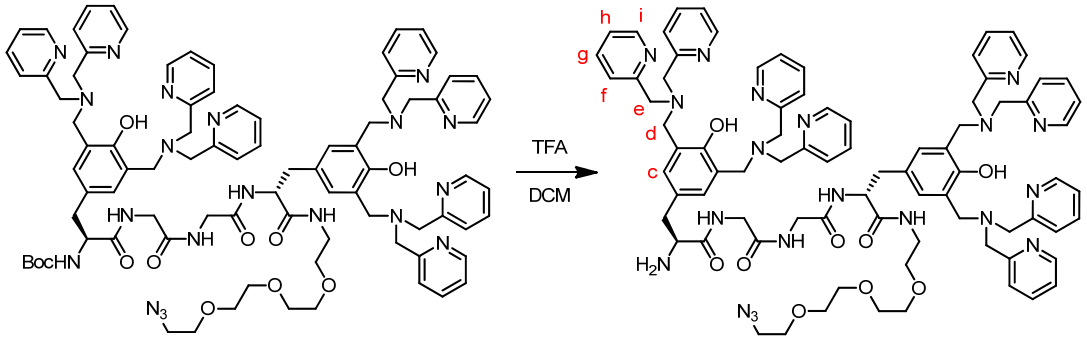
BocDpaTyrGlyGlyDpaTyr-triEG-N_3



A two-necked, 25 mL round-bottomed-flask was purged with N_2. BocDpaTyr-triEG-N_3 (283 mg, 0.35 mmol) was added and dissolved with anhydrous DCM (4 mL). TFA (4 mL) was added dropwise and the reaction was stirred at RT for 6 hour. The reaction was then evaporated to dryness under reduced pressure. The resulting residue was re-dissolved with anhydrous DMF (6 mL) and cooled to 0 °C. *N*,*N*′-diisopropyl-ethylamine (DIEA) was added until pH = 10 and then BocDpaTyrGlyGlyOH (308 mg, 0.38 mmol) 1-hydroxy-benzotriazole monohydrate (HOBt·H_2O, 80 mg, 0.52 mmol) and 1-ethyl-3-(3-dimethyl-aminopropyl)carbodiimide hydrochloride (EDC·HCl, 100 mg, 0.52 mmol) were added subsequently. Reaction was stirred at room temperature for 17 hours. Reaction was diluted with H_2O (100 mL) and extracted with EtOAc (2 × 80 mL). Combined EtOAc was washed with saturated bicarbonate solution (2 × 80 mL), brine

(100 mL), dried with Na_2SO_4 and then filtered. The resulting solution was concentrated under reduced pressure to yield the product as a pale red foam (444 mg, 80%). Spectroscopic data was identical to that previously reported.[5.5] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.33 (9H, s, Boc), 2.92-3.13 (4H, m, Hβ), 3.23-3.88 (44H, m), 4.39-4.42 (1H, m, Hα), 4.67 (1H, m, Hα), 5.80 (1H, br. s, NH), 7.00 (2H, s, H_c), 7.09-7.22 (11H, m, contain H_h and NH), 7.35 (1H, br. s, NH), 7.43-7.48 (8H, m, H_f), 7.58-7.63 (8H, m, H_g), 7.85 (1H, br. s, NH), 8.51 (4H, dd, J = 4.8, 0.8, H_i), 8.55 (4H, d, J = 4.4, H_i), 8.74 (1H, br. s, NH), 10.96 (1H, br. s, OH), 11.13 (1H, br. s, OH).

H-DpaTyrGlyGlyDpaTyr-triEG-N_3

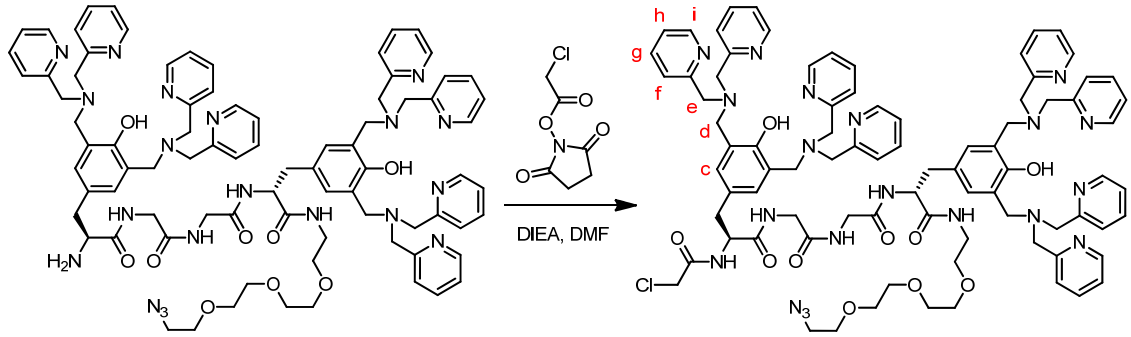


A two-necked, 25 mL round-bottomed-flask was purged with nitrogen. BocDpaTyr-GlyGly-DpaTyr-triEG-N_3 (422 mg, 0.26 mmol) was added and dissolved with anhydrous DCM (5 mL) then cooled to 0 °C. TFA (5 mL) was added dropwise and the reaction was stirred at RT for 6 hour. The reaction was then evaporated to dryness under reduced pressure. The residue was dissolved with H_2O (50 mL) and extracted with $CHCl_3$ (50 mL × 2). Combined organic layer was washed with brine (50 mL), dried (Na_2SO_4) and filtered. The resulting solution was concentrated in vacuo to obtain a pale brown foam. Crude NMR shows some impurity between 1-3 ppm. The product was purified by back extraction. First, the product was dissolved in $CHCl_3$ (50 mL) and washed with acidic water (50 mL, pH = 1-2). This organic fraction was discarded. The

[5.5] Nonaka, H.; Fujishima, S.-h.; Uchinomiya, S.-h.; Ojida, A.; Hamachi, I. *J. Am. Chem. Soc.* **2010**, *132*, 9301-9309.

aqueous layer was then adjusted to pH 9-10 and extracted with $CHCl_3$ (50 mL × 2). Combined $CHCl_3$ was washed with brine (50 mL), dried (Na_2SO_4) and filtered. The resulting solution was concentrated in vacuo to obtain a white foam (260 mg, 66%). Spectroscopic data was identical to that previously reported.[5.5] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 2.58 (1H, dd, *J* = 13.6, 9.2, Hβ_1), 3.00 (1H, dd, *J* = 14.0, 9.2, Hβ_2), 3.08 (1H, dd, *J* = 14.0, 4.8, Hβ_2'), 3.15 (1H, dd, *J* = 13.6, 6.4, Hβ_1'), 3.27-3.94 (44H, m), 4.62-4.65 (1H, m, Hα), 7.01 (2H, s, H_c), 7.09-7.14 (10H, m, contain H_h and H_c'), 7.35 (1H, d, *J* = 8.0, NH), 7.47 (8H, t, *J* = 7.2, H_f), 7.60 (8H, dq, *J* = 8.0, 1.6, H_g), 8.19-8.21 (2H, m, NH), 8.29 (2H, t, *J* = 5.6, NH), 8.51 (8H, dd, *J* = 4.4, 0.8, H_i), 8.53-8.55 (1H, m, NH), 11.00-11.05 (2H, br. s, OH).

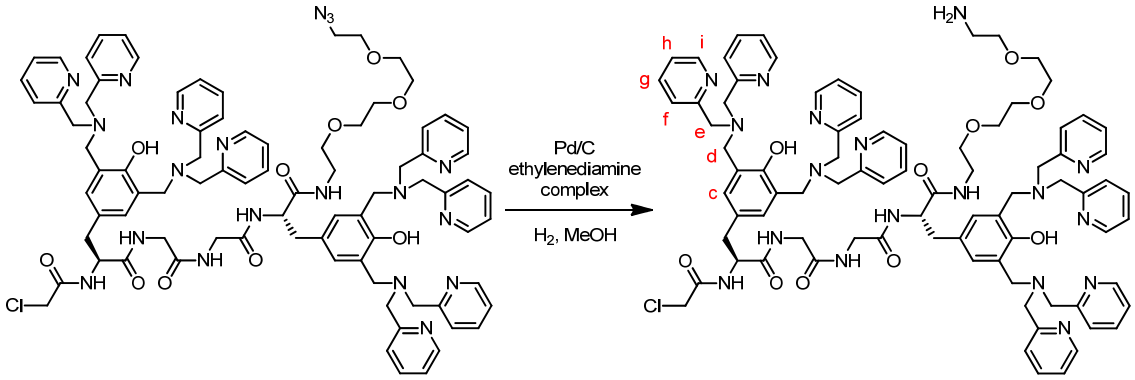
_N_-αClAc-DpaTyrGlyGlyDpaTyr-triEG-N_3



A 25 mL round-bottomed flask was purged with N_2. H-DpaTyrGlyGlyDpaTyr-triEG-N_3 (100 mg, 0.07 mmol) was added and dissolved with anhydrous DMF (5 mL). *N*,*N*′-diisopropylethylamine (DIEA, 23 μL, 0.13 mmol) was added followed by chloroacetic acid *N*-hydroxysuccinimide ester (20 mg, 0.1 mmol). The reaction was stirred at RT for 3 hour. The reaction was diluted with H_2O (50 mL) and extracted with EtOAc (40 mL × 2). Combined organic layer was washed with $NaHCO_3$ (sat. solution, 40 mL × 2), brine (40 mL), dried (Na_2SO_4) and filtered. The resulting solution was concentrated in vacuo to obtain a thick yellow oil (102 mg, 97%). Spectroscopic data was identical to that previously reported.[5.5] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 2.97-3.22 (4H, m, Hβ), 3.32-3.91 (44H, m), 4.03 (1H, d, *J* = 14.2, αCl-C*H*H′), 4.25 (1H, d, *J* =

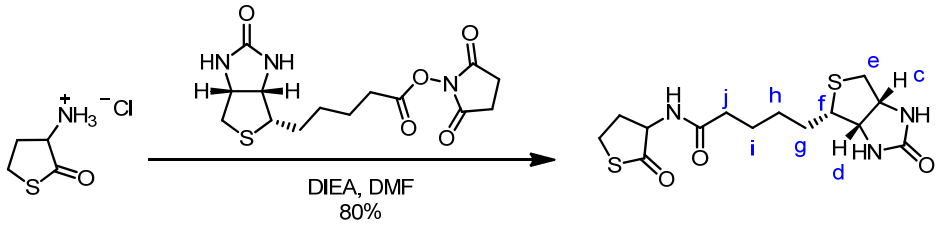
14.2, αCl-CH*H′*), 4.65 (1H, dd, *J* = 16.0, 7.6, Hα), 4.73-4.78 (1H, m, Hα′), 7.00 (2H, s, H_c), 7.05-7.07 (1H, m, NH), 7.11-7.17 (10H, m, contain H_h and H_c′), 7.30 (1H, m, NH), 7.46 (8H, t, *J* = 6.4, H_f), 7.61 (8H, dq, *J* = 7.6, 1.6, H_g), 7.70-7.76 (2H, m, NH), 8.54-8.58 (8H, m, H_i), 9.26-9.27 (1H, m, NH). HRMS m/z (ESI+): Found 1579.7295 $[M+H]^+$; $C_{84}H_{96}ClN_{20}O_{10}$ requires 1579.7301.

_N_-αClAc-DpaTyrGlyGlyDpaTyr-triEG-NH$_2$ (5.14)



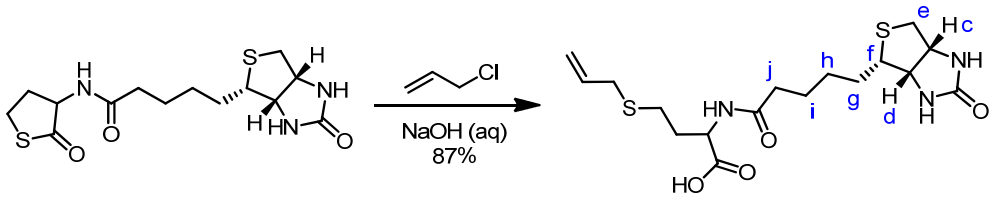
N-αClAc-DpaTyrGlyGlyDpaTyr-triEG-N_3 (146 mg, 0.09 mmol) was added to a 2-necked 50 mL round-bottomed flask and purged with N_2. Anhydrous DCM and MeOH (1:1, 10 mL) were added and the solution was stirred at RT. Pd/C-ethylenediamine complex (50 mg, 30% wt of starting material) was then added to the stirring solution. Finally, the gas in the flask was exchanged with H_2. The reaction was stirred vigorously for 5 hours under H_2 atmosphere (1 atm) at RT. Reaction was then filtered through a pad of celite and concentrated by rotary evaporation under reduced pressure to give a pale yellow solid (126 mg, 88%). Spectroscopic data was identical to that previously reported.[5.5] ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 2.99-3.21 (4H, m, Hβ), 3.33-4.08 (46H, m), 4.56-4.58 (1H, m, Hα), 4.72-4.74 (1H, m, Hα′), 7.00-7.17 (12H, m, contain H_h and H_c′), 7.22-7.24 (1H, m, NH), 7.42-7.47, (8H, m, H_f), 7.59-7.64 (8H, m, H_g), 7.79-7.87 (2H, m, NH), 8.18-8.19 (1H, m, NH), 8.55 (8H, d, *J* = 4.4, H_i), 8.60-8.62 (1H, m, NH). HRMS m/z (ESI+): Found 1553.7345 $[M+H]^+$; $C_{84}H_{98}ClN_{18}O_{10}$ requires 1553.7396.

Biotinylated homocysteine thiolactone



DL-homocysteine thiolactone hydrochloride (50 mg, 0.33 mmol) was added to a 10 mL round-bottomed flask and dissolved in anhydrous DMF (3 mL). *N*,*N*′-diisopropyl-ethylamine (DIEA, 140 μL, 0.80 mmol) was added followed by biotin *N*-hydroxysuccinimide ester (120 mg, 0.36 mmol). The reaction was stirred under N_2 atmosphere at RT for 24 hours. DMF was removed under high vacuum. The resulting residue was dissolved with $CHCl_3$ (50 mL) and washed with water (10 mL × 3) then dried with Na_2SO_4. After filtration, the solvent was removed to give a white solid. Purification by silica gel column chromatography (20% MeOH in $CHCl_3$) obtained the desired product as a white solid (90 mg, 80%). ^{1}H NMR (400 MHz, CD_3OD): δ_H = 1.48 (2H, quin, *J* = 7.6, H_h), 1.56-1.79 (4H, m, H_g, H_i), 2.09-2.21 (1H, m, Hβ), 2.22-2.31 (2H, m, H_j), 2.57 (1H, m, Hβ′), 2.70 (1H, d, *J* = 12.8, H_e), 2.93 (1H, dd, *J* = 12.8, 4.8, H_e′), 3.19-3.24 (1H, m, Hγ), 3.26-3.30 (1H, m, H_f), 3.41 (1H, dt, *J* = 11.6, 5.2, Hγ′), 4.31 (1H, dd, *J* = 8.0, 4.8, H_c), 4.49 (1H, dd, *J* = 8.0, 4.0, H_d), 4.65 (1H, dd, *J* = 12.8, 6.8, Hα).

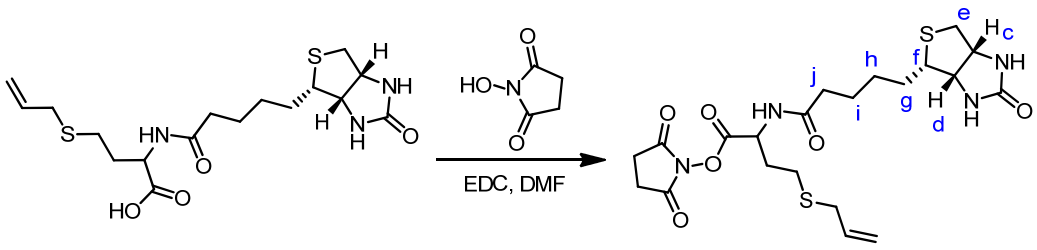
Biotinylated *S*-allyl-homocysteine



Biotinylated homocysteine thiolactone (56 mg, 0.16 mmol), NaOH (1 mL of 1N solution) and H_2O (4 mL) was added to a 25 mL round-bottomed flask. The mixture was stirred until all solid has dissolved before adding allyl chloride (75 μL, 0.92 mmol)

and stirred for further 3 hours at RT. The reaction was diluted with H_2O (5 mL) and acidified to pH 1-2 with 1N HCl before extracted with $CHCl_3$ (50 mL x 2). The resulting organic layer was slightly cloudy and MeOH was added until solution become clear. The organic layer was dried with Na_2SO_4, filtered and then concentrated under reduced pressure to yield a white solid (56 mg, 86%). ^{1}H NMR (400 MHz, CD_3OD): δ_H = 1.32-1.41 (2H, m, H_h), 1.45-1.67 (4H, m, H_g, H_i), 1.78-1.88 (1H, m, Hβ), 1.96-2.03 (1H, m Hβ′), 2.17 (2H, t, J = 7.0, H_j), 2.35-2.49 (2H, m, Hγ), 2.60 (1H, d, J = 12.8, H_e), 2.83 (1H, dd, J = 12.8, 4.8, H_e′), 3.05 (2H, d, J = 7.2, SCH_2), 3.08-3.13 (1H, m, H_f), 4.21 (1H, dd, J = 7.6, 4.4, H_c), 4.40 (1H, dd, J = 8.0, 4.4, H_d), 4.44 (1H, dd, J = 9.2, 4.8, Hα), 4.91-5.04 (2H, m, $=CH_2$), 5.69 (1H, dt, J = 17.2, 7.2, CH=).

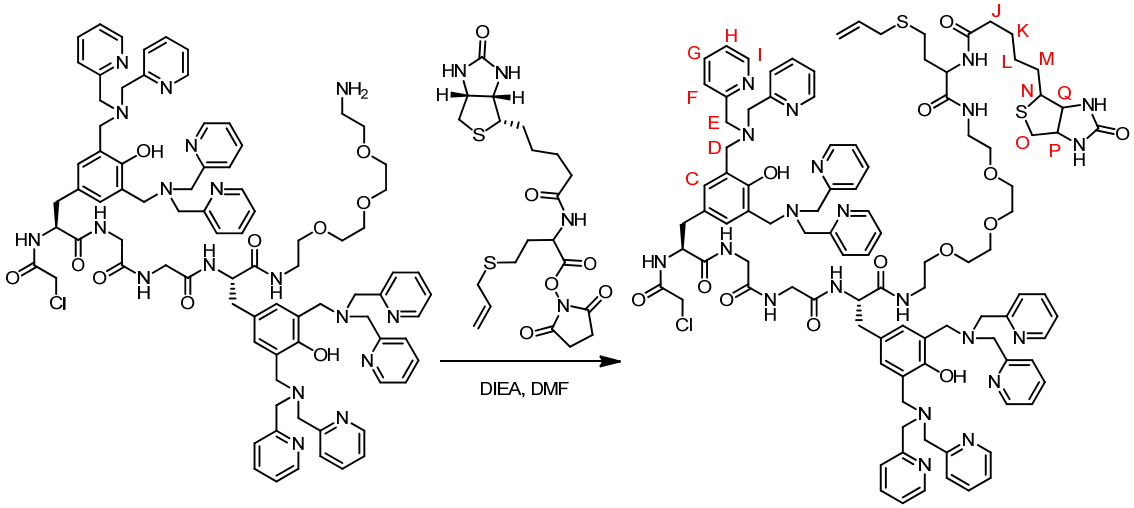
Biotinylated *S*-allyl-homocysteine *N*-hydroxysuccinimde ester (5.13)



Biotinylated *S*-allyl-homocysteine (92 mg, 0.23 mmol) was added to a 25 mL round-bottomed flask and dissolved in anhydrous DMF (5 mL). *N*-hydroxysuccinimide (32 mg, 0.28 mmol) followed by 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC · HCl, 53 mg, 0.28 mmol) was added to the reaction and stirred for 24 hours at RT. DMF was removed under high vacuum. The resulting residue was dissolved in $CHCl_3$ (100 mL) and washed with H_2O (10 mL × 3), dried with Na_2SO_4 and then filtered. Solvent was removed by rotary evaporation under reduced pressure to yield a white amorphous solid (103 mg, 90%). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.42-1.48 (2H, m, H_h), 1.60-1.80 (4H, m, H_g and H_i), 2.07-2.32 (4H, m, Hβ and H_j), 2.57-2.66 (2H, m, Hγ), 2.72 (1H, d, J = 12.8, H_e), 2.85 (4H, s, NHS-CH_2), 2.91 (1H, dd, J = 12.8, 4.8, H_e′), 3.11-3.14 (1H, m, H_f), 3.16 (2H, d, J = 7.2, SCH_2), 4.28-4.33 (1H, m,

H_c), 4.49-4.52 (1H, m, H_d), 5.03-5.15 (3H, m, Hα and $=CH_2$), 5.78 (1H, dt, J = 17.2, 7.6, CH=). HRMS m/z (ESI+): Found 521.1497 $[M+Na]^+$; $C_{21}H_{30}N_4O_6S_2Na$ requires 521.1449.

_N_-αClAc-DpaTyrGlyGlyDpaTyr-triEG-_S_-allyl homocysteine biotin (5.12)



A mixture of *N*-αClAc-DpaTyrGlyGlyDpaTyr-triEG-NH_2 (32 mg, 0.020 mmol), *N*,*N*′-diisopropyl-ethylamine (DIEA, 35 μL, 0.20 mmol) and biotinylated *S*-allyl-homocysteine *N*-hydroxysuccinimde ester (12 mg, 0.024 mmol) in anhydrous DMF (2 mL) was stirred at room temperature for 3 hours. TLC revealed complete consumption of the amine starting material. The product was purified by reverse-phase HPLC (product peak at 37.8 minutes) and subsequent lyophilization afforded the resired compound as a thick oil (13 mg). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.28-1.46 (2H, m, *L*), 1.50-1.75 (4H, m, *K* and *M*), 1.80-2.03 (2H, m, Ahc-β), 2.26 (2H, t, J = 7.2, *J*), 2.68 (1H, d, *J* = 12.4, *O*), 2.82-2.93 (3H, m, *O*′and Tyr-β), 3.03-3.07 (1H, m, *N*), 3.11 (2H, d, *J* = 7.6, SCH_2), 3.15-3.21 (2H, m, NCH_2CH_2O), 3.26-3.36 (4H, m, Gly-CH_2), 3.48-3.62 (14H, triEG), 3.76-4.24 (26H, m, $ClCH_2$, *D* and *E*), 4.28-4.44 (2H, m, *P* and *Q*), 4.48 (1H, dd, *J* = 7.6, 4.4, Tyr-α), 4.59 (1H, dd, *J* = 9.6, 4.8, Tyr-α), 4.67 (1H, dd, *J* = 9.2, 5.2, Ahc-α), 5.00-5.10 (2H, m, $=CH_2$), 5.74 (1H, ddt, *J* = 17.2, 10.0, 7.2, CH=), 7.16, 7.17 (4H, s, *C*), 7.49-7.55 (16H, m, *F* and *H*), 7.89-7.98 (8H, m, G), 8.62-8.65 (8H, m,

I). HRMS m/z (ESI+): Found 969.4429 $[M+2H]^{2+}$; $C_{101}H_{124}ClN_{21}O_{13}S_2$ requires 969.4420.

HPLC conditions are as follows. Column: YMC-pack ODS-A, C18, 250 × 20 mm; flow rate: 10 mL/min; detection: UV (220 nm); mobile phase: MeCN (with 0.1% TFA)/H_2O (with 0.1% TFA), step gradient as follows:

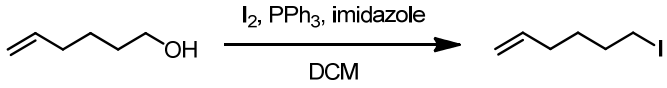
t (min)	0	10	20	45	60	70	75
H_2O (%)	100	100	75	65	50	0	100
MeCN (%)	0	0	25	35	50	100	0

***N*-αClAc-DpaTyrGlyGlyDpaTyr-triEG-*S*-allyl homocysteine biotin zinc complex**

Equal volume of **5.12** (2 mM) and $ZnCl_2$ (8 mM) in HEPES buffer (50 mM HEPES, pH 7.5) were mixed to obtain 1 mM solution of the desired zinc complex (**5.12-Zn**) in quantitative yield. The complex was diluted to 1 μM with HEPES buffer (1 μL in 1 mL buffer) for cell labeling reaction.

Synthesis of fluorescein probe 5.15

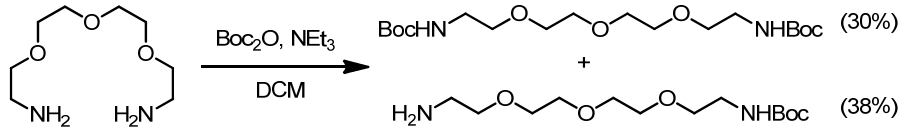
6-Iodohexene



A stirred solution of 5-hexen-1-ol (1.3 mL, 10.8 mmol), triphenylphosphine (3.40 g, 13.0 mmol) and imidazole (1.10 g, 16.2 mmol) in anhydrous DCM (40 mL) was cooled to 0 °C. Iodine (3.30 g, 13.0 mmol) was then added to the stirring mixture. The reaction was stirred vigorously at 0 °C for 1.5 hour when TLC revealed complete consumption of starting material. The reaction was quenched with saturated NH_4Cl (40 mL) (cloudy yellow to colorless) and then transferred to a separatory funnel. The organic layer was separated and the aqueous layer was extracted with DCM (2 × 40 mL). The combines organic layers were washed with brine (80 mL), dried over $MgSO_4$, filtered, and concentrated under reduced pressure. Purification by flash column chromatography

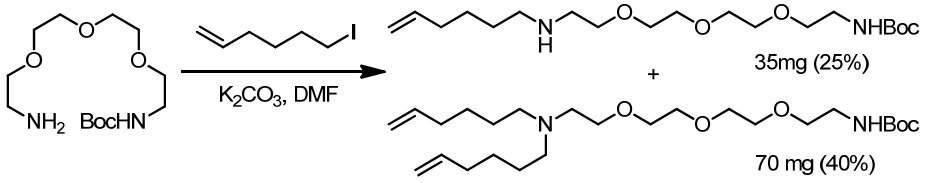
(hexane) afforded 6-iodohexene as a thin volatile oil (2.07 g, 90%). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.51 (2H, quin, *J* = 7.45, C<u>H</u>$_2$CH$_2$CH$_2$I), 1.85 (2H, quin, *J* = 7.26, C<u>H</u>$_2$CH$_2$I), 2.09 (2H, q, *J* = 7.07, C<u>H</u>$_2$CH=CH$_2$), 3.20 (2H, t, *J* = 6.95, C<u>H</u>$_2$I), 4.98 (1H, d, *J* = 10.36, CH=C<u>H</u>H *cis*), 5.02 (1H, d, *J* = 17.18, CH=CH<u>H</u> *trans*), 5.80 (1H, ddt, *J* = 16.96, 10.20, 6.63, 6.63, C<u>H</u>=CH$_2$).

_N_-Boc-tetraethylene glycol



1,11-diamino-3,6,9-trioxaundecane (220 mg, 1.14 mmol) was added to a 250 mL round-bottomed flask with a Teflon™-coated stirrer bar and dissolved in anhydrous DCM (4 mL). The flask was purged with N_2 atmosphere and stirred. Et_3N (240 µL, 1.72 mmol) followed by Boc_2O (250 mg, 1.14 mmol) was added and the reaction was stirred at room temperature for 3 hours. After this time the reaction was evaporated to dryness with rotary evaporation. The resulting residue was purified directly by silica gel column chromatography (MeOH:$CHCl_3$, 1:9, with 0.1% NH_4OH) to obtained the unwanted di-Boc protected compound (138 mg, 30%) and the desired product as a colourless oil (128 mg, 38%). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.45 (9H, s, Boc), 2.87 (2H, t, *J* = 5.2, BocNHCH$_2$C<u>H</u>$_2$O), 3.31 (2H, q, *J* = 4.8, BocNHC<u>H</u>$_2$), 3.50-3.56 (4H, m, NCH$_2$CH$_2$O), 3.60-3.68 (8H, m, diEG-CH$_2$).

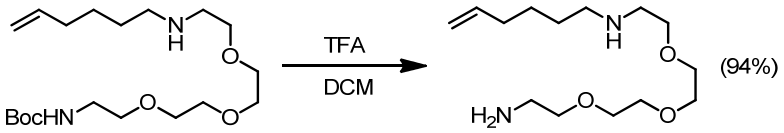
_N_-Boc-_N'_-hexenyl tetraethylene glycol



N-Boc-tetraethylene glycol (110 mg, 0.38 mmol) was added to a 25 mL round bottom flask and dissolved in 2 mL of anhydrous DMF. K_2CO_3 (78 mg, 0.56 mmol) was added

to the stirred solution. 6-Iodohexene (100 mg, 0.45 mmol, in 1 mL DMF) was then added by pipette to the reaction. The reaction was stoppered and stirred at room temperature for 24 hours. DMF was removed under high vacuum. The resulting residue was purified by column chromatography to afford the unwanted di-alkylated product (70 mg, 40%) and the desired mono-alkylated product (35 mg, 25%). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.39-1.46 (11H, m, Boc and C$\underline{H}_2$CH_2CH_2N), 1.49-1.60 (2H, m, C$\underline{H}_2$CH_2N), 2.07 (2H, m, *J* = 7.2, 6.8, 1.6, 1.2, C$\underline{H}_2$CH=CH_2), 2.65 (2H, t, *J* = 6.8, C$\underline{H}_2$N), 2.81-2.84 (2H, m, NC$\underline{H}_2$CH_2O), 3.29-3.32 (2H, m, BocNHCH_2C$\underline{H}_2$), 3.54 (2H, t, J = 4.8, BocNHC$\underline{H}_2$), 3.60-3.72 (10H, m, diEG-CH_2 and NCH_2C$\underline{H}_2$O), 4.94 (1H, ddt, *J* = 10.2, 2.0, 1.2, CH=C$\underline{H}$H *cis*), 5.00 (1H, ddt, *J* = 17.2, 2.0, 1.6, CH=CH$\underline{H}$ *trans*), 5.80 (1H, ddt, *J* = 17.2, 10.2, 6.8, C$\underline{H}$=CH_2).

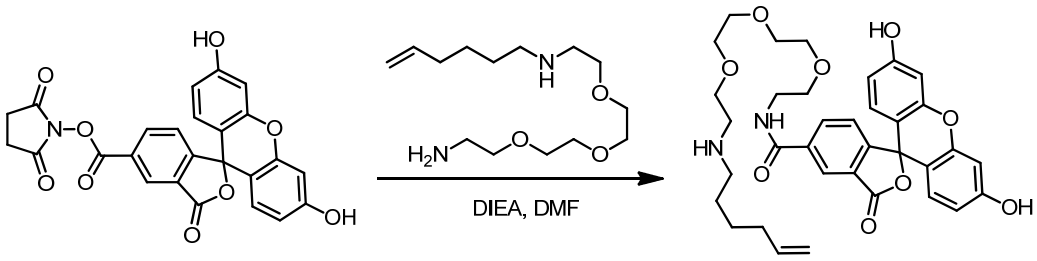
_N_-hexenyl tetraethylene glycol



N-Boc-*N*′-hexenyl tetraethylene glycol (35 mg, 0.09 mmol) in a 25 mL round-bottomed-flask was purged with N_2 and dissolved with anhydrous DCM (1.5 mL). TFA (1.5 mL) was added dropwise and the reaction was stirred at RT for 3 hour. The reaction was then evaporated to dryness under reduced pressure. The residue was dissolved with H_2O (25 mL) and the pH was adjusted to 11 with NH_4OH. Aqueous phase was extracted with $CHCl_3$ (2 × 25 mL). Combined organic layer was washed with brine (25 mL), dried (Na_2SO_4) and filtered. The resulting solution was concentrated in vacuo to obtain a colourless oil (24 mg, 94%). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 1.44 (2H, quin, *J* = 6.8, C$\underline{H}_2$CH_2CH_2N), 1.51 (2H, quin, *J* = 6.8, C$\underline{H}_2$CH_2N), 2.07 (2H, m, *J* = 7.2, 6.8, 1.6, 1.2, C$\underline{H}_2$CH=CH_2), 2.61 (2H, t, *J* = 7.2, C$\underline{H}_2$N), 2.79 (2H, t, J = 5.2, hex-NC$\underline{H}_2$CH_2O), 2.87 (2H, t, *J* = 5.2, hex-NCH_2C$\underline{H}_2$O), 3.52 (2H, t, *J* = 4.8, NH_2C$\underline{H}_2$),

3.59 (2H, t, J = 5.2, $NH_2CH_2C\underline{H}_2$), 3.61-3.68 (8H, m, diEG-CH_2), 4.94 (1H, ddt, J = 10.2, 2.0, 1.2, CH=C$\underline{H}$H *cis*), 5.00 (1H, ddt, J = 17.2, 2.0, 1.6, CH=CH$\underline{H}$ *trans*), 5.80 (1H, ddt, J = 17.2, 10.2, 6.8, C$\underline{H}$=CH_2).

5′-triEG-hexenyl fluorescein



5′-carboxy-fluorescein NHS ester (30 mg, 0.063 mmol) and *N*-hexenyl tetraethylene glycol (24 mg, 0.087 mmol) was added to a 25 mL round-bottomed flask and dissolved with DMF. *N*,*N*′-diisopropyl-ethylamine (DIEA, 30 μL, 0.17 mmol) was added and the reaction was stirred at room temperature for 14 hours. DMF was removed under high vacuum and the resulting residue was purified by silica gel column chromatography (30% MeOH in $CHCl_3$ with 0.1% AcOH). The desired product was obtained as a bright orange solid (35 mg, 87%). ^{1}H NMR (400 MHz, CD_3OD): δ_H = 1.34 (2H, quin, J = 7.6, $C\underline{H}_2CH_2CH=CH_2$), 1.55 (2H, quin, J = 7.6, $NCH_2C\underline{H}_2$), 1.97 (2H, q, J = 6.8, $C\underline{H}_2CH=CH_2$), 2.87 (2H, t, J = 8.0, NCH_2), 3.06 (2H, t, J = 5.2, CH_2N-hexenyl), 3.53-3.63 (14H, m, triEG-CH_2), 4.86 (1H, dd, J = 10.4, 1.0, CH=C$\underline{H}$H *cis*), 4.91 (1H, dd, J = 17.2, 1.8, CH=CH$\underline{H}$ *trans*), 5.68 (1H, ddt, J = 17.2, 10.4, 6.8, C$\underline{H}$=CH_2), 6.50 (2H, dd, J = 9.2, 2.0), 6.56 (2H, d, J = 2.0), 6.86 (2H, d, J = 9.2) (6 × CH xanthene), 7.26 (1H, d, J = 8.0), 7.99 (1H, dd, J = 8.0, 1.2), 8.43 (1H, s) (3 × CH benzoate). HRMS m/z (ESI$^+$): Found 633.2796 $[M+H]^+$; $C_{35}H_{41}N_2O_9$ requires 633.2807.

Cell Culture and Recombinant Protein Expression in HEK293 Cells

HEK293 cells were cultured in high glucose Dulbecco's Modified Eagle Medium (DMEM, 4.5 g of glucose/L) supplemented with 10% fetal bovine serum (FBS),

penicillin (100 units/mL), streptomycin (100 μg/mL), and amphotericin B (250 ng/mL). Cells were maintained at 37 °C in a humidified atmosphere of 5% CO_2 in air. A subculture was performed every 3-4 days from subconfluent (<80%) cultures using a trypsin-EDTA solution. Transfection of cDNA plasmids was carried out in a 35 mm glass-bottomed dish (Iwaki) using Lipofectamine LTX (Invitrogen) according to a general protocol. The cells were subjected to labeling experiments after 44-52 h of the transfection.

General Procedure for Covalent Labeling of tag-fused B2R by 5.12-Zn

The HEK293 cells expressing the tag fused B2R (~1×10^6) were incubated with 1 mL of HEPES-buffered saline (HBS, containing 107 mM NaCl, 6 mM KCl, 1.2 mM $MgSO_4$, 2 mM $CaCl_2$, 11.5 mM glucose, 20 mM HEPES, adjusted to pH 7.4 with NaOH) containing 0.5 mM TCEP for 10 min at room temperature. After removing TCEP solution, the cells were treated with 1 mL of HBS (pH 7.4) containing the labeling probe **5.12-Zn** (1 μM) for 10 min at room temperature. The probe solution was removed and the cells were washed with 1.0 mL of Phosphate Buffered Saline (PBS) containing 1 mM pyrophosphate (PP_i) to remove the unreacted probe. The sample is then ready for analysis by western blotting (see below for general procedures) or the subsequent CM reaction.

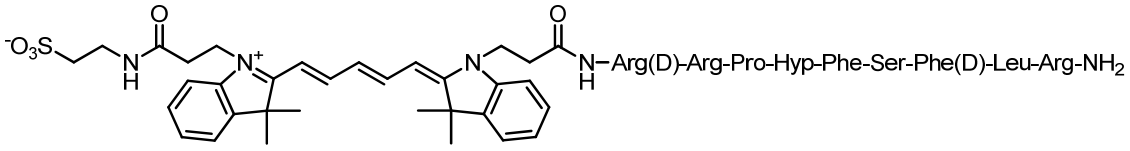
General Procedure for Cross-Metathesis on HEK239 cells

After the initial labeling with probe **5.12-Zn**, catalyst/*t*-BuOH solution (50 μL of 2.5 mg/mL solution in *t*-BuOH) was mixed with PBS buffer (450 μL) and then added to cell dish. After 1 minute, fluorescein solution was added (500 μL of 200 μM solution in PBS) and incubated at 37 °C for 30 to 60 minutes. The reaction mixture was then removed and the cell was washed with PBS (1 mL × 3). The sample is then ready for analysis either by CLMS or western blotting. In the experiment where either catalyst or

fluorescein is absent the reaction mixture, blank PBS buffer is used instead (Scheme 5-13b,e,f). In the case of C&F premix (Scheme 5-13d), the catalyst and fluorescein solutions are premixed prior to addition to cell culture dish.

Sample Preparation for Fluorescence Imaging

After being filled with PBS containing Cy5-appended B2R antagonist peptide (0.5 μM, 1 mL) with/without propidium iodide (PI), the cells were analyzed using a confocal laser scanning microscope (CLSM; Olympus, FLUOVIEW FV1000) equipped with a ×100 lens and the appropriate emission filters for fluorescein, Oregongreen488, EGFP, TAMRA, PI and Cy5. Cy5-Appended B2R Antagonist Peptide was kindly supplied by the Hamachi Laboratory (see below for the chemical structure).



Cy5-Appended B2R Antagonist Peptide

Sample Preparation for Western Blotting

The labeled HEK293 cells expressing the tag fused. The cells were collected by cell scraper and lysed with 100 μL of RIPA lysis buffer (pH 7.4, 25 mM Tris·HCl, 150 mM NaCl, 0.1% SDS, 1% Nonidet P-40, 0.25% Deoxycholic acid) containing 1% protease inhibitor cocktail set III (Novagen) at 4 °C for 60 min. The insoluble material was removed by centrifugation at 12 000 rpm for 10 min. The supernatant was mixed with 2× sample loading buffer for membrane protein (pH 6.8, 125 mM Tris·HCl, 20% glycerol, 4% SDS, 0.01% bromophenol blue, 100 mM DTT) and incubated at RT for 1-2 hours (instead of heat denaturation). The sample was then subjected to SDS-PAGE followed by electrotransfer onto an Immun-Blot PVDF membrane (Bio-Rad). The membrane was blocked with 5% skimmed milk in TBST buffer for 1 hour at RT (TBST

= 10 mM Tris, 200 mM NaCl, 0.02% (w/w) Tween 20, pH 7.8). After washing with TBST (5 x 15 mL), the membrane was incubated with the desired antibody (see below for details).

The immune-detection of the B2R was performed by treating the membrane with rabbit anti-B2R polyclonal antibody for 1 hour at RT (60 μL in 15 mL TBST, 1:250 dilution, Santa Cruz Biotechnology), washed with TBST (5 x 15 mL) and then incubated for further 1 hour with the secondary antibody: horseradish peroxidase (HRP)-conjugated anti-rabbit IgG antibody (3 μL in 15 mL TBST, 1:5000 dilution, Santa Cruz Biotechnology). After washing with TBST (5 x 15 mL), the chemiluminescence signal using Chemi-Lumi One (Nacalai Tesque) was detected with ChemiDoc XRS (BIO-RAD).

The strepavidine blot (to check for attachment of biotin-appended probe) was prepared treating the membrane with HRP-conjugated streptavidin (3 μL in 15 mL TBST, 1: 5000 dilution, Invitrogen) for 1 hour at RT. After washing with TBST (5 x 15 mL), the membrane was subjected to the chemiluminescence detection as described above.

The anti-fluorescein blot was prepared by treating the membrane with rabbit anti-fluorescein polyclonal antibody for 1 hour at RT (7.5 μL in 15 mL TBST, 1:2000 dilution, Abcam), washed with TBST (5 x 15 mL) and then incubated for further 1 hour with the secondary antibody: horseradish peroxidase (HRP)-conjugated anti-rabbit IgG antibody (3 μL in 15 mL TBST, 1:5000 dilution, Santa Cruz Biotechnology). After washing with TBST (5 x 15 mL), the membrane was subjected to the chemiluminescence detection using Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare).