

Novel Protein Modification Strategies

Doctor of Philosophy (DPhil) in Organic Chemistry

The University of Oxford, *Pembroke College*

Department of Chemistry

Submitted: **Trinity 2018**

By: **Ben J. Bower** *under the supervision of* **Prof. Benjamin G. Davis**

0.1 Abstract

Site-directed mutagenesis allows the ready variation of proteinogenic amino acids; genetic code expansion permit the incorporation of a few more. The widespread use of these genetic techniques has caused a revolution in molecular biology. A crucial limitation of genetic techniques is the ‘ribosomal filter’, where an amino acid has to be processed by the ribosome or other translational enzymes.

I have developed a chemical technique for the direct modification of proteins allowing the incorporation of a large scope of amino acid side-chains. Notably, this technique is independent of the ribosome. Radical addition to protein backbone dehydroalanine (Dha) forms $C(sp^3)-C(sp^3)$ bonds directly onto a protein to facilitate “chemical mutagenesis”.

Protein glycosylation is difficult to study due to mixtures associated with biological glycosylation machinery. In Chapter 1: I have employed the above chemical mutagenesis method to create homogenous glycoproteins which retain a native bond. During this study, I have demonstrated the biological relevance of these transformations by showing reactivity with glycosidases and glycosyl transferases, found interesting reactivity of an extremely polarised amide substrate, and have used the method on an aggregation prone periplasmic protein.

In Chapter 2: I used techniques developed in Chapter 1 to create intermediate mimics of the Sirt2 catalysed deacetylation of Histones. I used the mimics of Sirt2 intermediate 2 to determine the geometry around an unstable reaction intermediate. Having atomic control over the amino acid allowed us to create mimics of both the **R** and **S** diastereomers. Using analytical size exclusion chromatography I have obtained qualitative results which show the **R** isomer is bound by Sirt2 much stronger than the **S**, hinting that the **R** isomer is that which is present in biology.

In Chapter 3: I demonstrated a green synthesis of 2,5-dibromohexadiamide (DBHDA); which has been used extensively in the prior two chapters to modify protein Cys residues into Dha. The developed method of producing DBHDA replaces CCl_4 with cyclohexane and gives a yield of 31 % over 3 steps and 6 functional group interconversions. The DBHDA produced was found to be of extremely high purity ($99.7 \% \pm 1.21$), which was determined by the use of quantitative ^{13}C NMR spectroscopy.

0.2 Acknowledgments

Firstly: I would like to thank my supervisors Ben Davis and Alistair Henry for providing exciting, ground-breaking and challenging projects to work on. They both have piloted my projects to success through the wavy waters of chemical biology. Providing ample suggestions for experiments when I was out of ideas has not only kept me busy but also taken projects in new and exciting directions, while keeping the focus of my work on hypothesis driven science.

The Department of Chemistry at The University of Oxford have provided a lively atmosphere to work in with a focus on innovation which has helped me to focus on experimental work; even if Jack Cox glory was never found. Pembroke College and my College Supervisor: Andy Baldwin have provided a friendly and supportive habitat during the course of my DPhil. I would like to thank UCB not only for funding and the use of their facilities but for: contributing ideas to my project; organising events which has allowed my work to be showcased while allowing me to hear work from diverse array of fields and; for giving me a taste of how chemical biology is conducted in industry. I am very grateful for the studentship provided by UCB and the BBSRC which has allowed me to conduct this DPhil degree. The Bodleian has never once disappointed, no matter what obscure journal I have requested; even if the periodical provided is in a foreign language and held together only by bootlace, what else should a scholar expect?

Throughout the course of DPhil degree, I have collaborated with numerous people without whom success would have been much delayed, if attained at all: I worked extensively with Tom Wright on the development of radical additions to dehydroalanine. The synthesis of the *O*-GlcNAc reagent was conducted in collaboration with Ollie Coleman and Tom Fleming, they both worked on the expression and modification of cAb-Lys which was used in one chapter 1 experiment. I worked with Ritu Raj and Jens Willwacher on the synthesis of ribose derivatives used in chapter 2. Tom Parson kindly provided me with tetrasaccharide oxazoline used in chapter 1, as well as the suggestion of using the ENGases in the first place! The tetrasaccharide was also synthesised later in collaboration with Aziz Khan. Cristina Marculescu kindly provided me with a sample of EndoS which saved a few weeks of work in getting a vital result. ESR experiments were performed with Will Meyers which provided a crucial insight into radical formation. While working at UCB Paul Creeke not only provided excellent technical assistance with surface plasmon resonance but also numerous (albeit unsuccessful) ideas on how to reduce unwanted binding to CM resin. Numerous conversations with Andy Baldwin have allowed me to become an aggregate connoisseur much like himself, without which I would not have been able to complete the AcrA work. Tim Claridge has provided numerous suggestions for NMR spectroscopy experiments, which I otherwise would never have thought of, as well as checking my work on some more challenging spectra. Charlie Fehl has managed to turn an old sample of 1-(IAc)- β -GlcNAc into a new project for which I am very

grateful. Although this is at the very start of the collaboration, work with Brian Josephson and Manfred Jung offers exciting prospects for the characterisation of Sirt2 deacetylation intermediates. Last but by no means least: James Wickens has provided reliable technical support, new experimental ideas and untold patience while aiding me with mass spectrometry. James unfortunately has a few times had to miss his bus home in order to change a needle or polish a cone, so I could watch the mass spectrometer well into the night.

Technical support has been critical throughout the course of my DPhil degree I would like to thank the technicians: Amy Webb, Syvain Royer, Abdel Boumertit and Vincent Mason for untold help surrounding numerous issues within the Ben Davis group. I would like to thank James Wickens and Elisabete Pires for technical assistance with mass spectrometry. Barbara Odell and Tina Jackson for technical support with NMR spectroscopy. Alison Turner for technical assistance with SEC-MALS and James Heads for technical assistance with thermal florescence. The X-Ray crystallography service at the University of Oxford's chemistry department has allowed unambiguous characterization of numerous samples, this was particular useful in chapter 2, I would like to thank all the team and Amber Thompson in particular.

Numerous people have proofread this thesis including: Ben Davis, Alistair Henry, Tom Wright, Kim Wals, Brian Joesephson, Patrick Issenegger, Gustaf Hember and Aziz Kahn. I am very thankful to them for performing a challenging and somewhat tedious task.

Success in my DPhil degree has been aided by the experience I had prior to starting at Oxford. Work in the R Alan Aitken Laboratory gave me experience in chemical synthesis which transformed my skills from those of a student to those of a researcher, I appreciate help from both Alan and Youcef Balouta in this. Work in the laboratory of Doug Philp not only honed experimental skill but stated an interest in carbohydrate chemistry which has been a theme for my doctoral work: I would like to thank Doug and Tamara Kosikova for this. Time in the group of Terry K Smith was my first experience of research and taught me to persevere when projects don't seem to be moving and also how to work independently. Time spent working with Iain Smellie was 'enlightening', I am also thankful for what he taught me about radicals.

I have made too many friends to mention during my time in the Ben Davis group. Their support and conversations have been invaluable for the development of this work and myself as a scientist. Cycling with: Leanne Minall, Charlie Fehl, Costi Giurgiu, Gustaf Hemberg and Brian Josephson has led to many great ideas while rolling through the Cotswolds. Tom Parsons has suggested many ideas while stood next to me that turned into major parts of this thesis and encouraged me to read a set of novels I enjoyed greatly. The beer enthusiasts in the group (who shall remain anonymous) have led to many unforgettable memories and enjoyable occasions. Mateusz Imiolek has been a great flatmate, even if he does keep on me up every other Monday discussing chemistry problems. Kim Wals and Tom Wright have taught me most of what I know about biological work, as well as being very good friends. I've

had a lot of fun with Aziz Kahn while working on chemistry. To those specifically mentioned and many more: it's been a fun and exciting 4 years.

My parents Patty and John have been a continued source of support throughout my DPhil study. Abigail has not once complained about how much cycling I have watched while writing this thesis and has provided some excellent illustrations for title pages.

Finally: I wish to thank my examiners Akane Kawamura and John Denu. I hope they enjoy reading this thesis as much as I have enjoyed creating it.

0.3 Contents

Novel Protein Modification Strategies	i
0.1 Abstract.....	i
0.2 Acknowledgments.....	iii
0.3 Contents	viii
0.4 Experiments performed in collaboration	xxii
0.5 Abbreviations	xxiii
0.6 Preface.....	xxviii
References:	xxxi
Chapter 1: Synthetic Glycoproteins	0
1.1.0.0 Introduction.....	1
1.1.1.0 Classes of Glycoproteins.....	1
1.1.1.1 <i>O</i> -Linked Glycoproteins	3
1.1.1.2 <i>N</i> -Linked Glycoproteins	7
1.1.2.0 Synthetic Production of Glycoproteins	9
1.1.2.1 Introduction.....	9
1.1.2.2 Nonspecific Methods.....	10
1.1.2.3 Glycosylation using Genetically Incorporated Tags.....	12

1.1.2.4 Glyco-Conjugations Using Cysteine	17
1.1.2.5 Glyco-conjugations Using Dehydroalanine	21
1.1.3.0 Introduction to Radical Addition to Dha.....	23
1.1.4.0 Aims	27
1.1.4.1 Choice of Protein Sites	28
1.2.0.0 Results and Discussion.....	29
1.2.1.0 <i>O</i> -GlcNAcylation.....	29
1.2.2.0 <i>N</i> -GlcNAcylation.....	33
1.2.3.0 Glycoengineering of Synthetically Incorporated Glycans	40
1.2.3.1 Extending Enzymes	40
1.2.3.2 Trimming Enzymes	48
1.2.4.0 Expansion.....	58
1.2.4.1 General <i>O</i> -Glycosylation	58
1.2.4.2 <i>O</i> -Glc.....	59
1.2.4.3 <i>O</i> -Man	62
1.2.4.4 F ₂ - <i>N</i> -GlcNAc	65
1.2.5.0 AcrA.....	68
1.2.5.1 Introduction.....	68

1.2.5.2 AcrA problems.....	70
1.2.5.3 Optimising AcrA Dha Formation.....	70
1.2.5.6 Optimising Radical Addition to AcrA Dha	82
1.2.5.5 AcrA Enzyme Reactions and Conclusion	91
1.3.0.0 Conclusion	94
1.3.1.0 Summary	94
1.3.2.0 Perspective	94
1.4.0.0 Experimental	98
1.4.1.0 General Chemical Experimental Considerations	98
1.4.1.1 NMR Spectroscopy General.....	98
1.4.1.2 ¹ H NMR Spectroscopy.....	99
1.4.1.3 ¹³ C NMR Spectroscopy	99
1.4.1.4 ¹⁹ F NMR Spectroscopy.....	100
1.4.1.5 IR Spectroscopy.....	100
1.4.1.6 Small Molecule Mass Spectrometry	100
1.4.1.7 Optical Rotation.....	101
1.4.1.8 UV-Vis Spectroscopy	101
1.4.1.9 Glovebox.....	101

1.4.1.10 TLC Dips.....	102
1.4.1.11 Ion Exchange Resin Activation.....	103
1.4.1.12 Molecular Sieves Activation.....	103
1.4.2.0 Small Molecule Compound Characterisation	104
1.4.2.1 Peracetyl- β -D-GlcNAc 1	104
1.4.2.2 Peracetyl 1-iodoethyl- β -D-GlcNAc 2.....	106
1.4.2.3 1-iodoethyl- β -D-GlcNAc 3	108
1.4.2.4 Peracetyl 2-chloro- α -D-GlcNAc 5	110
1.4.2.5 Peracetyl 1-azido- β -D-GlcNAc 6	112
1.4.2.6 Peracetyl 1-amino- β -D-GlcNAc 7	114
1,3-Propane Dithiol method	114
Adam's Catalyst Method:.....	115
1.4.2.7 Peracetyl 1-(4-iodobutyramido)- β -D-GlcNAc 10	117
1.4.2.8 1-(4-iodobutyramido)- β -D-GlcNAc 11	120
1.4.2.9 Peracetyl 1-(iodoactamide)- β -D-GlcNAc 8	122
EEDQ method:.....	122
Acid chloride method:	123
1.4.2.10 1-(iodoacetamide)- β -D-GlcNAc 9	125

1.4.2.11 2,2-Difluoro-2-iodoacetic Acid 18	127
1.4.2.12 Peracetyl 1-(2,2-difluoro-2-iodoacetamido)- β -D-GlcNAc 19	128
1.4.2.13 1-(2,2-difluoro-2-iodoacetamido)- β -D-GlcNAc 20	130
1.4.2.14 General method for sugar peracetylation	132
1.4.2.15 General method for sugar iodoethanol glycosylation	133
1.4.2.16 General Method for 1-(2-Iodoethoxy)aldohexose deacetylation	133
1.4.2.17 Peracetyl α,β -D-Mannose 15	134
1.4.2.18 Peracetyl 1-(2-iodoethoxy)- α -D-Mannose 16	135
1.4.2.19 1-(2-iodoethoxy)- α -D-Mannose 17	136
1.4.2.20 Peracetyl α,β -D-Glucose 12	137
1.4.2.21 Peracetyl 1-(2-iodoethoxy)- β -D-Glucose 13	138
1.4.2.22 1-(2-iodoethoxy)- β -D-Glucose 14	139
1.4.3.0 General Protein Experimental Considerations	140
1.4.3.1 General	140
1.4.3.2 Common Protein Solutions	141
Expression Stocks:	141
Broths Employed:	142
Common Buffers:	143

1.4.3.3 Protein Mass Spectrometry	145
1.4.3.4 Circular Dichroism	146
1.4.3.5 Fast Liquid Protein Chromatography	146
1.4.3.6 Thermal Florescence.....	146
1.4.3.7 Transformation Protocol	147
1.4.3.8 Protein Concentration Measurement.....	147
1.4.4.0 Mutagenesis.....	147
1.4.4.1 AcrA wt to AcrA N123C	147
1.4.5.0 Images	148
1.4.6.0 Protein Expression	149
1.4.6.1 Multidrug efflux pump subunit AcrA Expression (wt and 123 Cys). 149	
1.4.6.2 Family 84 human <i>O</i> -GlcNAcase (hOGA) Expression	151
1.4.6.3 Endo- β - <i>N</i> -acetylglucosaminidase from <i>Arthrobacter protophormiae</i> Y299F (EndoA Y299F) Expression.....	153
1.4.6.4 Histone H3 K27C.....	155
1.4.7.0 Chemical Reactions of Proteins.....	158
1.4.7.1 Histone H3 k27 Dha	158
1.4.7.2 Histone H3 S10 Dha	160

1.4.7.3 H3 K27OGlcNAc	162
1.4.7.4 H3 K27NGlcNAc	163
1.4.7.5 H3 S10 NF ₂ GlcNAc	164
1.4.7.6 H3 S10OGlc	168
1.4.7.7 H3 S10OMAn	170
1.4.7.8 H3 S10NGlcNAc	172
1.4.7.9 AcrA N123Dha	174
1.4.7.10 AcrA N123OGlcNAc.....	176
1.4.7.11 AcrA N123OGlcNAc Degas Time Screen.....	178
1.4.7.12 AcrA N123Dha with DBHDA – Failed	180
1.4.7.13 AcrA N123OGlcNAc Zn Degas Time Screen	181
1.4.7.14 AcrA N123Dha - Failed	183
1.4.8.0 Glyco-engineering of histone H3.....	185
1.4.8.1 EndoA.....	185
1.4.8.1.1 N-Linked	186
1.4.8.1.2 O-Linked	187
1.4.8.2 GalT	188
1.4.8.2.1 N-Linked	189

1.4.8.2.2 O-Linked	190
1.4.8.3 hOGA (human <i>O</i> -GlcNAcase).....	191
1.4.8.4 Jackbean <i>O</i> -GlcNAcase	194
1.4.8.5 PNGase F	196
1.4.8.5.1 Native.....	196
1.4.8.5.2 Denatured.....	197
1.4.8.5.3 PNGase F activity control	198
1.4.8.6.4 Her EndoS Cleavage.....	200
1.4.8.5.5 Trimmed Her PNGase F cleavage.....	203
1.4.9.0 Glyco-engineering of cAb-Lys.....	204
1.4.9.1 EndoA extension with tetrasaccharide oxazoline.....	204
1.4.10.0 Glyco-engineering of AcrA	206
1.4.10.1 <i>Streptococcus pneumoniae</i> <i>O</i> -GlcNAcase cleavage test	206
1.4.10.2 <i>Human</i> <i>O</i> -GlcNAcase cleavage test	208
1.4.11.0 Protein Sequences of Expressed Proteins:	209
1.4.11.1 His ₆ -AcrA wt:.....	209
1.4.11.2 His ₆ -AcrA N123C:	209
1.4.11.3 Histone H3 K27C, C110A:	209

1.4.11.4 Histone H3 S10C, C110A:.....	209
1.4.11.5 His ₆ -Family 84 <i>human O</i> -GlcNAcase (hOGA):.....	210
1.4.11.6 GST-Endoglycosidase from <i>Streptococcus pyogenes</i> D233Q	210
1.4.11.7 GST-Endo- β - <i>N</i> -acetylglucosaminidase from <i>Arthrobacter</i> <i>protophormiae</i> Y299F:.....	211
1.5.0.0 References	212
Chapter 2: Determining the Stereochemistry of a Sirt2 Intermediate	222
2.1.1.0 Introduction.....	223
2.1.2.0 Sirt2 Structure and Function	224
2.1.3.0 Sirt2 medical implications	225
2.1.4.0 Sirt2 Mechanism.....	226
2.1.5.0 Aims	230
2.2.0.0 Results and Discussion.....	232
2.2.1.0 Synthesis of Aminoacetal Mimic	232
2.2.2.0 Conjugation to Histones.....	234
2.2.3.0 – Histone Refolding	235
2.2.4.0 Expression of Sirt2	238
2.2.5.0 Binding by Surface Plasmon Resonance.....	244

2.2.6.0 Binding by Size Exclusion Chromatography	248
2.3.0.0 Conclusion	254
2.4.0.0 Experimental	258
2.4.1.0 Compound Characterisation.....	258
2.4.1.1 1-Iodohexan-5-one 21.....	258
2.4.1.2 5,3-Dibenzoylriboacetal.....	259
2.4.1.2.1 R 23.....	261
2.4.1.2.2 S 22	262
2.4.1.3 S-Riboacetal 24.....	263
2.4.1.4 R-Riboacetal 25	265
2.4.3.0 Protein Expression	267
2.4.3.1 <i>Human</i> Sirtuin 2 (56-365 and 50-365)	267
2.4.3.2 Histone H4 K16C.....	269
2.4.3.3 Histone H3 K18C.....	272
2.4.5.0 Chemical Reactions of Proteins.....	275
2.4.5.1 Histone H4 K16Dha.....	275
2.4.5.2 Histone H4 K16R-Riboacetal.....	277
2.4.5.3 Histone H4 K16S-Riboacetal.....	278

2.4.5.4 Histone H3 K18Dha.....	279
2.4.5.5 Histone H3 K18R-Riboacetal.....	281
2.4.5.6 Histone H3 K18S-Riboacetal.....	282
2.4.6.0 Biacore Experiments	283
2.4.6.1 CM5 Coupling	283
2.4.6.2 CM5 Calibration.....	285
2.4.6.3 CM5 Histone Test	286
2.4.6.4 CM5 Binding Experiment.....	288
2.4.6.5 CM5 Histone Test with CM-Dextran	292
2.4.6.6 CM4 Coupling	293
2.4.6.7 CM4 Calibration.....	295
2.4.6.8 CM4 Histone Test	296
2.4.6.9 CM4 Binding Experiment.....	300
2.4.6.10 Amine Surface Antibody coupling	303
2.4.6.11 Amine Surface Sirt2 Calibration.....	306
2.4.6.12 Carbohydrazide Oxidised Antibody Histone Test	307
2.4.6.13 Noncovalent Histone Test.....	308
2.4.7.0 SEC-MALS Experiments.....	309

2.4.7.1 Sirt2 Control.....	310
2.4.7.2 S Control	311
2.4.7.3 R Control.....	312
2.4.7.4 S 0.5 eq	313
2.4.7.6 S 1.0 eq	315
2.4.7.9 R 1.5 eq	318
2.4.8.0 Histone Refolding Protocol.....	319
2.4.9.0 Sequences	320
2.4.9.1 His ₆ TEV <i>Human</i> Sirtuin 2 56-365:.....	320
2.4.9.2 His ₆ TEV <i>Human</i> Sirtuin 2 50-365:.....	320
2.4.9.3 Histone H3 K18C, C110A:.....	320
2.4.9.4 Histone H4 K16C:.....	320
2.5.0.0 References	321
Chapter 3: A Greener Route to DBHDA.....	325
3.1.0.0 Introduction	326
3.1.1.0 Past Syntheses	326
3.1.2.0 CCl ₄ Impacts	328
3.1.3.0 Aims	329

3.2.0.0 Results and Discussion.....	329
3.2.1.0 Planning	329
3.2.2.0 Developed Procedure.....	331
3.2.2.1 2,5-dibromohexanediamide 27	331
3.2.3.0 Discussion.....	335
3.2.3.1 Reaction Monitoring.....	335
3.2.3.2 NBS Quality	338
3.2.3.3 NBS Quantity.....	339
3.2.3.4 Reaction Temperatures.....	340
3.2.3.5 Ammonia Vapour	341
3.2.3.6 Product Purity	343
3.3.0.0 Conclusion	345
3.4.0.0 Experimental	346
3.4.1.0 TLC Procedures	346
3.4.1.1 Chlorination.....	346
3.4.1.2 Bromination.....	346
3.4.2.0 Quantitative ¹³ C NMR Spectra.....	347
3.4.3.1. Chlorination.....	349

3.4.3.2 Bromination.....	349
3.4.4.0 DBHDA Characterisation data	349
3.5.0.0 References	351

0.4 Experiments performed in collaboration

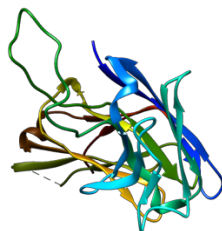
cAb-Lys Glycoengineering	Tom Wright, Oliver Coleman
Synthesis of <i>O</i> -GlcNAc Reagent	Oliver Coleman, Tom Fleming
Synthesis of Chapter 2 Riboacetals	Ritu Raj, Jens Willwacher
Crystallography	Amber Thompson

0.5 Abbreviations

2xYT	Twice Yeast Tryptone
Ac	Acetyl
	Multidrug Efflux Pump A cridine Resistance Protein, followed by a single letter which denotes the part
AcrX	
ADPR	A denosine 5'- D iphosphoribose
Ala	A lanine
Arg	A rginine
aSEC	A nalytical S ize E xclusion C hromatography
Asn	A sparagine
Asp	A spartic Acid
bpy	2,2'- B ipyridine
brine	Saturated Aqueous Sodium Chloride
BSA	B ovine S erum A lbumin
Bz	B enzoyl
<i>C. Jenuni</i>	<i>Campylobacter jejuni</i>
Cassava Blight Bacteria	<i>Xanthomonas manihotis</i>
CD	C ircular D ichroism
CDI	C arbonyl d iimidazole
CI	C hemical Ionisation
CML	C arboxy m ethyl L ysine
COSY	Homonuclear c orrelation s pectroscopy
CuAAC	C opper(I)-Catalyzed A zide- A lkyne C ycloaddition
Cys	C ysteine
DBADC	2,5- D ibromoadipyl D ichloride
DBHDA	2,5- D ibromohex s diamide
DCM	D ichlorom e thane
	Distorsionless enhancement by polarization transfer (Q denotes detection of quaternary nuclei) otherwise a tilt angle is given
DEPTx	
Dha	D ehydro a lanine
diNAcBac	<i>N,N'</i> - D iacetyl b acillosamine
DIPEA	<i>N,N</i> - D iisopropylethylamine
DMAP	4-(d imethyl) a minopyridine
DMSO	D imethyl S ulfoxide
DNA	D eoxyribon n ucleic A cid
DTT	D ithio t hreitol
<i>E. coli</i>	<i>Escherichia coli</i>
EDC	1- E thyl-3-(3- d imethylaminopropyl) c arbodiimide
EDG	E lectron D onating G roup

EDTA	E thyl e di a min e t e t r a a cetic acid
EEDQ	<i>N</i> - E thoxycarbonyl-2- E thoxy-1,2- D ihydro q uino l ine
EndoX	Endo - β - <i>N</i> -acetylglucosaminidase followed by a letter which denotes....
ENGase	Endo - β - <i>N</i> -acetylglucosaminid a s e followed by a letter which denotes....
ESI	E lectro s pray Ionisation
Et	E thyl
Et Al.	<i>Et Alia</i>
EWG	E lectron W ithdrawing G roup
fPBS	F luorinated P hosphate B uffered S aline
FPLC	F ast P rotein L iquid C hromatography
FTIR	F ourier transform I nfrared
Fuc	F ucose
Gal	G alactose
GalNAc	<i>N</i> - A cetyl g alactosamine
GalT	β -1,4- G alactosyl t ransferase
Gdm.Cl	G uanidinium chloride
Glc	G lucose
GlcNAc	<i>N</i> - A cetyl g lucosamine
Gln	G lutamine
Glu	G lutamic acid
Gly	G lycine
GPI	G lycosyl p hosphatidylinositol
Hag	H omoallylglycine
HAT	H istone A cetyltransferase
HBS-P	H epes B uffered S aline with P 20 surfactant
HDAC	H istone D eacetylase
Her	H erceptin
hhGln	<i>homohomo</i> G lutamine
hhSer	<i>homohomo</i> S erine
His	H istamine
HMBC	H eteronuclear M ultiple B ond C orrelation
hOGA	H uman <i>O</i> - G lc N Acase
HPLC	H igh P erformance L iquid C hromatography
HSQC	H eteronuclear single quantum coherence spectroscopy
Human	<i>Homo Sapiens</i>
Ile	I soleucine
IR	I nfrared
Jackbean	<i>Canavalia Ensiformis</i>
LB	L ysogeny B roth

LCMS	Liquid Chromatography Mass Spectrometry
Leu	Leucine
Lys	Lysine
MALS	Multangle Light Scattering
Man	Mannose
MDBP	Methyl 2,5-Dibromopentanoate
Me	Methyl
MES	2-(<i>N</i> -Morpholino)ethanesulfonic Acid
Met	Methionine
MHB	Muller Hinton Broth
MS	Mass Spectrometry
NAD ⁺	Nicotinamide Adenine Dinucleotide
NBS	<i>N</i> -Bromosuccinimide
NHS	<i>N</i> -Hydroxysuccinimide
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
OGA	<i>O</i> -GlcNAcase
OGT	<i>O</i> -GlcNActransferase
PBS	Phosphate Buffered Saline
PEST	Proline Glutamic Acid Serine Threonine
PG	Protecting Group
pH	<i>Pondus Hydrogenii</i> , $-\log_{10}[H_3O^+]$
Phe	Phenylalanine
PNGase F	Peptide: <i>N</i> -Glycosidase F
<i>p</i> NP	<i>para</i> -Nitrophenol
ppm	Parts Per Million
Pro	Proline
PTM	Post-Translation Modification
Py	Pyridine
qNMR	Quantitative Nuclear Magnetic Resonance Spectroscopy
RM	Reaction Mixture typically used on a TLC
RNA	Ribonucleic acid
RPM	Revolutions Per Minute
RT	Room temperature (and pressure)
S200	Superdex 200
SAU-X	Sodium Acetate Urea, followed by a number which denotes the concentration of sodium chloride
SB	Super Broth
SDS-PAGE	Sodium Dodecyl Sulfate Poly(acrylamide) Gel Electrophoresis

SEC	Size E xclusion C hromatography
Ser	S erine
SirtX	S irtuin followed by a number which denotes the type
SM	S tarting M aterial
SOC	S uper O ptimal B roth with C atabolite R epression
SPR	S urface P lasmon R esonance
SPPS	S olid P hase P eptide S ynthesis
TB	T errific B roth
TBAHS	T etrabutylammonium h ydrogensulfate
TBS	T ris B uffered S aline
THF	T etra h ydrofuran
Thr	T hreonine
TIC	T otal I on C ount
TLC	T hin L ayer C hromatography
TOCSY	T otal C orrelated S pectroscopy
Tris	T ris(hydroxymethyl)aminomethane
Trp	T ryptophan
Tyr	T yrosine
UDP	U ridine D iphosphate
UV	U ltra- V iolet
Val	V aline
wt	W ild T ype
X-Y-OAADPR	O - A cetyl A denosine D iphosphate R ibose, where X represents the position of acetylation and Y the anomer
α	Denotes an antibody
$^{\circ}\text{C}$	Degrees C elsius
	Denotes an unspecific protein
aq.	A queous
sat.	S aturated
calc.	C alculated
Da	D alton
eq.	E quivalents
ES	E lectrospray
FA	F ormic A cid
HRMS	H igh R esolution M ass S pectrometry
h	H our
min	M inute

s	Second
br	Broad
s	Singlet
d	Doublet
dd	Doublet of doublet
ddd	Doublet of doublet of doublet
dt	Doublet of triplet
t	Triplet
q	Quartet
m	Multiplet
OD	Optical Density
R	General Organic Chemical Group

0.6 Preface

When mammalian sperm and egg meet, numerous interactions between proteins, carbohydrates and glycoproteins commence: Firstly, the sperm cell breaks through a protective layer of cells surrounding the egg known as the *cumulus oophorus*, this is accomplished by the hydrolysis of the glycocalyx using a GPI anchored hyaluronidase.¹ The next obstacle for the sperm is an extracellular matrix known as the zona pellucida, an extra cellular matrix composed of three glycoproteins: ZP1, ZP2 and ZP3.² A molecular recognition event occurs between a galactosyl transferase on the surface of the sperm cell and ZP3 silyl-Lewis^{x3} which triggers^{4,5} the acrosome to react and release enzymes (proteases⁶ and glycosidases⁷) supplementary to the anchored proteases⁸ which digest the zona pellucida. Passing through the zona pellucida allows the fusion of the sperm cell to the egg cell plasma membrane through the interaction with the egg's integrin 61 and the sperms fertilin- β and cyritestin.⁹ In many ways it would be fair to say that human life begins with glycoproteins.

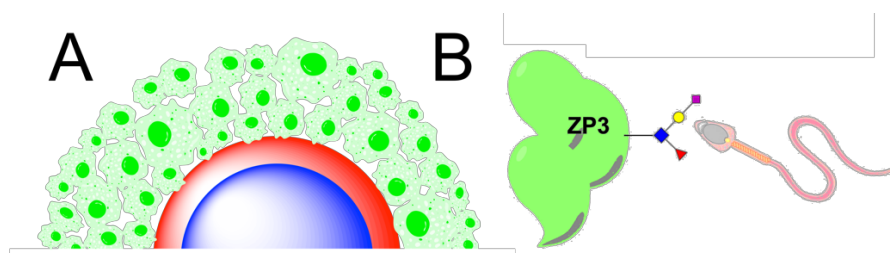


Figure 1 - A Diagram of the mammalian egg, showing the cumulus oophorus in green, zona pellucida in red and the egg in blue B Diagram of a mammalian sperm cell binding to the carbohydrate part (silyl-Lewisx) of glycoprotein ZP3

While the importance of glycoproteins needs little introduction to biologists, chemistry has failed to provide an economic route to preparing homogenous glycoproteins. For instance, blockbuster drug Epogen® (epoetin alfa) is sold as a mixture of at least 52 different glycoforms,¹⁰ obtained from expression in Chinese hamster ovary cells. The microheterogeneity of glycoforms makes studying the subtlety of protein glycosylation a difficult challenge. Danishefsky *et al.*¹¹ spent numerous years chemically synthesising a single glycoform of epoetin alfa; this yielded only 110 µg (4.17 nmol) of protein. A chemical method with the precision to access a single glycoform has great potential; possibly leading to longer lasting or more active biologics, as it has been demonstrated that certain glycoforms are more active than others.¹²

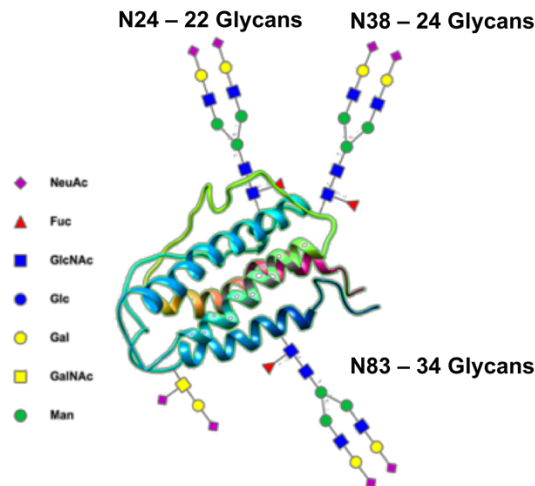


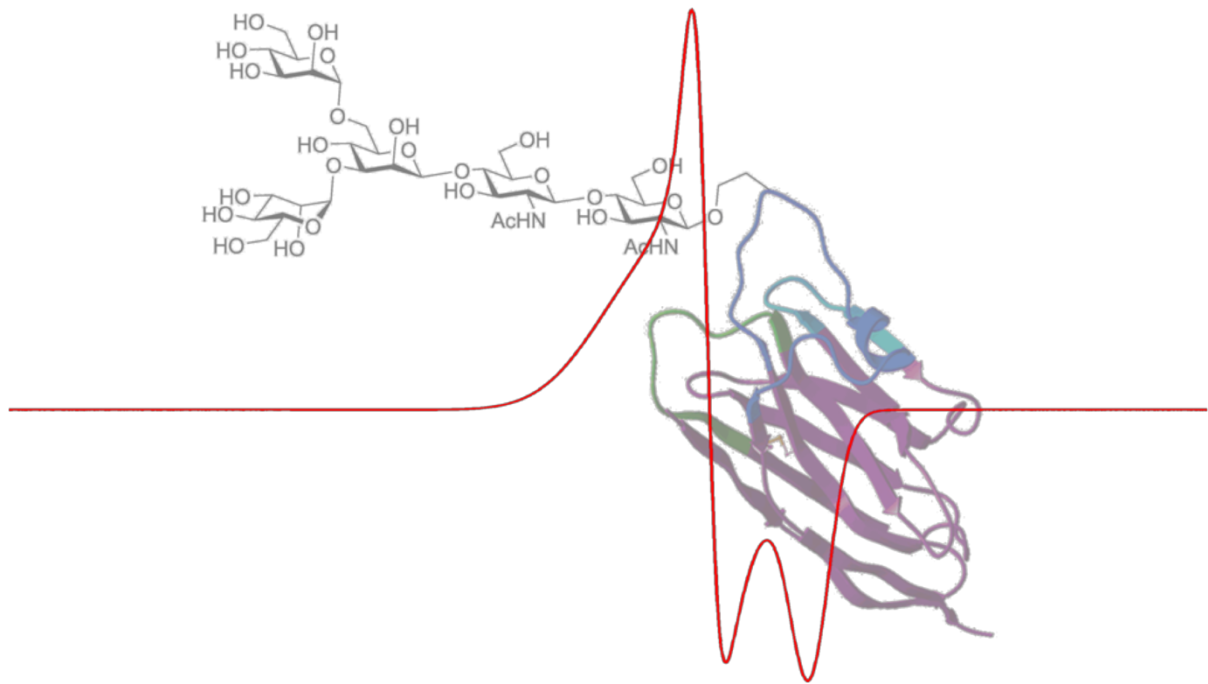
Figure 2 - The glycoform of epoetin alfa synthesised chemically by Danishefsky et al.¹¹

With the above in mind, and the additional fact that, in 2006, 70 % of proteins in clinical trials were glycosylated,^{13, 14} I set out to create a broadly applicable method of incorporating carbohydrates and their derivatives homogenously into proteins through a late stage recreation of their native linkage to the peptide backbone. I speculated that this 'late stage diversification'¹⁵ of expressed proteins by chemical mutagenesis would be the best way of achieving our goal of providing a broad tool for the glycobiologist to explore the glycoproteome. Therefore, I present to you 'Novel Protein Modification Strategies':

References:

1. Y. Lin, K. Mahan, W. F. Lathrop, D. G. Myles and P. Primakoff, *J. Cell Biol.*, 1994, **125**, 1157-1163.
2. J. M. Greve and P. M. Wassarman, *J. Mol. Biol.*, 1985, **181**, 253-264.
3. P.-C. Pang, P. C. N. Chiu, C.-L. Lee, L.-Y. Chang, M. Panico, H. R. Morris, S. M. Haslam, K.-H. Khoo, G. F. Clark, W. S. B. Yeung and A. Dell, *Science*, 2011, **333**, 1761-1764.
4. C. M. B. O'Toole, C. Arnoult, A. Darszon, R. A. Steinhardt, H. M. Florman and T. Hunter, *Mol. Biol. Cell*, 2000, **11**, 1571-1584.
5. M. K. Jungnickel, H. Marrero, L. Birnbaumer, J. R. Lemos and H. M. Florman, *Nat. Cell Biol.*, 2001, **3**, 499.
6. K. Ohmura, N. Kohno, Y. Kobayashi, K. Yamagata, S. Sato, S.-i. Kashiwabara and T. Baba, *J. Biol. Chem.*, 1999, **274**, 29426-29432.
7. D. J. Miller, X. Gong and B. D. Shur, *Development*, 1993, **118**, 1279-1289.
8. G. Z. Zhu, D. G. Myles and P. Primakoff, *J. Cell Sci.*, 2001, **114**, 1787-1794.
9. P. Primakoff and D. G. Myles, *Trends in Genet.*, 2000, **16**, 83-87.
10. R. S. Rush, P. L. Derby, D. M. Smith, C. Merry, G. Rogers, M. F. Rohde and V. Katta, *Anal. Chem.*, 1995, **67**, 1442-1452.
11. P. Wang, S. Dong, J.-H. Shieh, E. Peguero, R. Hendrickson, M. A. S. Moore and S. J. Danishefsky, *Science*, 2013, **342**, 1357-1360.
12. T. Suzuki, K. Kitajima, S. Inoue and Y. Inoue, *Glycoconj. J.*, 1995, **12**, 183-193.
13. N. Sethuraman and T. A. Stadheim, *Curr. Opin. Biotechnol.*, 2006, **17**, 341-346.
14. P. M. Rudd, H. C. Joao, E. Coghill, P. Fiten, M. R. Saunders, G. Opdenakker and R. A. Dwek, *Biochem.*, 1994, **33**, 17-22.
15. T. Cernak, K. D. Dykstra, S. Tyagarajan, P. Vachal and S. W. Krska, *Chem Soc Rev*, 2016, **45**, 546-576.

Chapter 1: Synthetic Glycoproteins



1.1.0.0 Introduction

It is predicted that more than two thirds of eukaryotic proteins are glycoproteins,^{16, 17} making glycosylation the most abundant post translational modification in Nature. Therefore, it should come as no surprise that the idea of carbohydrates as a component of proteins can be traced back to the earliest days of biochemistry. In 1885 Berzelius noticed that protein hydrolysed with acid was similar to a carbohydrate which had been decomposed in the same manner. Schutzenberger demonstrated that acid decomposed protein gives a positive result on testing with Fehling's solution, a chemical test for aldehydes. Many more unequivocal results^{18, 19} were obtained before Neuberger's²⁰ 1938 work flow for the isolation of protein bound carbohydrates was published. While it is feasible that there could be up to 10^{12} different dimers through to heptamers²¹ recent analysis considering known enzymes estimates ~7000 glycans in mammals.²²

1.1.1.0 Classes of Glycoproteins

Glycoproteins are typically classified²³ by the type of atom through which the glycan is bound to the protein. There are two large classes of glycoprotein and 4 smaller classes. The two large classes of protein glycosylation are *O*-Linked²⁴ and *N*-Linked,²⁵ which this chapter will focus on. Two of the classes are rare cases with different linkages: *C*-Linked, where mannose is connected to the C2 of a tryptophan indole;²⁶

and *S*-Linked where a GlcNAc is connected to a cysteine sulfur atom.²⁷ These smaller cases will not be discussed again in this thesis, but can be seen in *figure 1*.

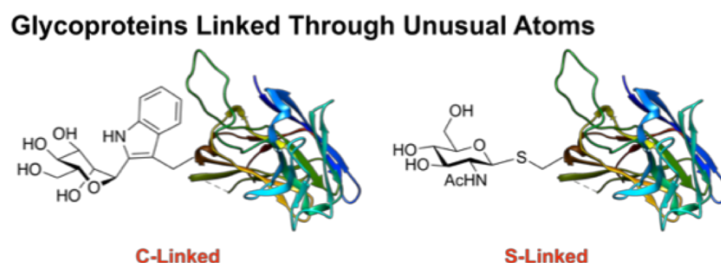
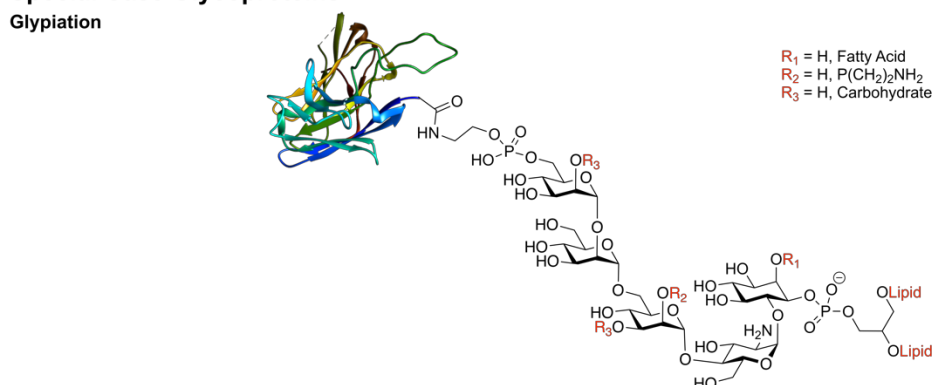


Figure 1: Structures of a generic protein linked to glycans, through atypical atoms: Carbon (C-Linked) and Sulfur (S-Linked).

The two remaining types of glycoprotein are both special cases: glypiation²⁸ refers to the attachment of glycosylphosphatidylinositol (GPI), a glycolipid, to the C terminus of the protein which is used to attach a protein to membranes; and glycation²⁹ which is the non-enzymatic and nonspecific attachment of a carbohydrate to a protein through Amadori,³⁰ Maillard^{31, 32} or Schiff base³³ reactivity, all are depicted in *figure 2*.

Glypiation



Glycation

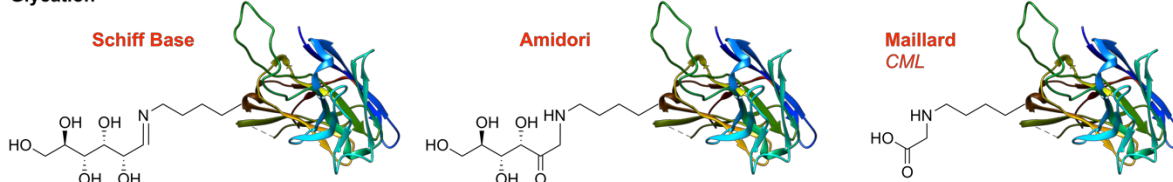


Figure 2 - Structures of atypical glycoproteins showing the C-terminal attachment of a GPI anchor, glypiation and the unspecific attachment of glycans characteristic of Schiff base, Amadori and Maillard Reactivity. CML Refers to carboxymethyl lysine.

1.1.1.1 O-Linked Glycoproteins

O-linked glycosylation is predominantly found on the side chain oxygen atoms of serine and threonine²⁴ residues but is also seen on the phenolic oxygen of tyrosine^{34, 35} side chains. Tyrosine glycosylation occurs most notably in glycogen.^{36, 37} *O*-linked glycosylation can also be seen on side chains which have been modified post-translationally such as hydroxylysine,³⁸ hydroxytryptophan,³⁹ hydroxyproline^{40, 41} and through the phosphate oxygen atom of a phosphoserine.⁴² Depiction of these *O*-Linked Glycans can be seen in *figure 3*.

O-Glycoproteins

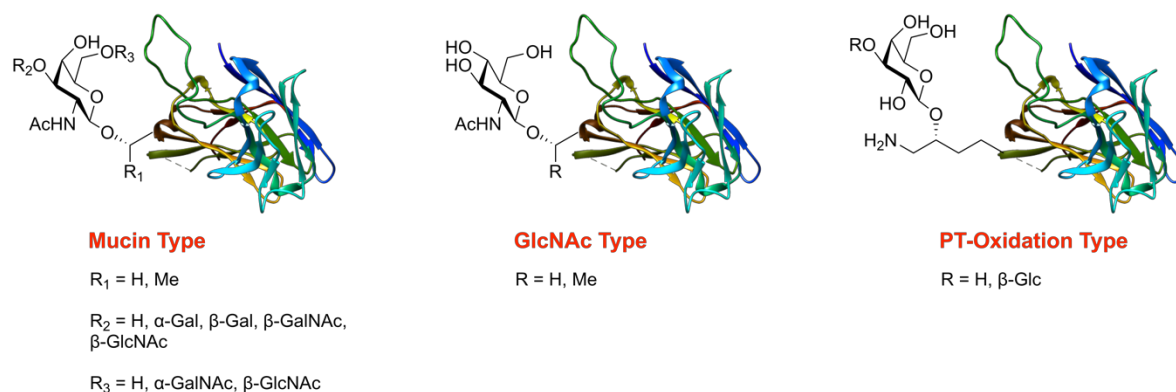


Figure 3 - Structures of some O-Linked Glycoproteins, showing GalNAc Ser/Thr and the substitutions which form mucin cores, GlcNAc Ser/Thr and Gal-Hydroxylysine a typical glycosylation of a posttranslational modification.

A significant portion of O-Linked glycoproteins have a GalNAc residue linked to the protein.²⁴ In eukaryotic cells GalNAc O-linked glycosylation typically occurs in the Golgi apparatus⁴³ by the addition of monosaccharide GalNAc to a Ser or Thr side chain; glycoside-transferases extend this monosaccharide to yield one of 8 core structures²⁴ which are then diversified further. Enzymes then modify these cores by the addition of further carbohydrate residues,⁴⁴ sulfate,^{45, 46} and acetyl^{47, 48} groups, this is known as mucin type glycosylation. Besides mucin type O-glycans several others can be found: O-Mannosylation can be found extensively in yeast cells⁴⁹⁻⁵² in contrast to mammalian cells where this modification rarely⁵³⁻⁵⁵ occurs. O-Fuc⁵⁶ and O-Glc⁵⁷ have also been observed in mammals but are rare. Finally, there is O-GlcNAc, which I shall discuss below.

Establishment that *O*-GlcNAcylation was localised in the nucleus⁵⁸ challenged the historic dogma⁵⁹ that glycosylation is restricted to compartments of the protein secretion machinery, cell surface and extracellular matrix. *O*-GlcNAcylation is distinct from other *O*-glycosylation for numerous other reasons: *O*-GlcNAc has exclusively been detected as a monosaccharide with no other modification, such as sulfation, acetylation or methylation and *O*-GlcNAc cycles on and off proteins (at a similar rate to phosphorylation).⁶⁰ In contrast to *O*-GalNAc glycosylation *O*-GlcNAcylation⁶¹ occurs primarily in the nucleus and cytoplasm,⁶² and is a much more dynamic modification.^{62, 63}

The cycling of *O*-GlcNAc is controlled by two enzymes, which are highly conserved^{64, 65} across species. *O*-GlcNAc transferase (OGT)⁶⁶ catalyses the addition of the monosaccharide while *O*-GlcNAcase (OGA)^{67, 68} removes it by hydrolysis. The *O*-GlcNAc modification like *O*-phosphorylation cycles on and off however, hundreds of kinases and phosphatases have been identified, in contrast to the two which control *O*-GlcNAcylation.⁶⁹

Approximately 50 % of known *O*-GlcNAcylation sites fit the Pro-Val-Ser sequence, which is also recognised by proline directed kinases. The fact that OGT will complex the catalytic subunit of protein phosphatase I⁷⁰ would hint that *in vivo* a single complex can both remove the *O*-phosphate PTM and add an *O*-GlcNAc unit. It has been demonstrated that OGT-deletion in neuron models leads to rapid

hyperphosphorylation of tau proteins,⁷¹ causing cell death. It has also been shown that *O*-GlcNAcylation can prevent the formation of toxic aggregates.⁷²⁻⁷⁴

Many *O*-GlcNAcylation sites have a high 'PEST' score⁷⁵ indicating that *O*-GlcNAcylation may slow down⁷⁶ the rapid degradation associated with these protein sites. PEST scoring,⁷⁷ originating from proline (P) glutamic acid (E) serine (S) threonine (T) is an algorithm which searches for clusters of the PEST amino acids and hydrophobic regions then returns a real number value which ranges from -45 for poly(isoleucine) to 50 for poly(aspartate) ($-45 < \text{PEST score} \in \mathbb{R} < 50$) with a single proline, a PEST score > 5 indicates that *in vivo* it is likely the protein will be rapidly degraded. It has been reported that *O*-GlcNAcylation can slow the proteasomal degradation of proteins.^{78, 79}

A ubiquitous response to cellular stress is the swift and wide increase in *O*-GlcNAcylation.⁸⁰ The increased *O*-GlcNAcylation level serves to protect cells by increasing the concentration of heatshock proteins,^{81, 82} by turning on their expression. Also, it is speculated that many more protective mechanisms⁸³ may be at play here; such as *O*-GlcNAcylation preventing protein aggregation as discussed in refs.⁷²⁻⁷⁴

As mentioned before, *O*-GlcNAcylation can mainly be found in the nucleus. RNA polymerase II as well as various transcription factors are *O*-GlcNAcylated, which can either activate or suppress transcription.^{84, 85} The role of *O*-GlcNAcylation in transcription and translation was discovered⁸⁶ when chromatinic proteins, histones

were found to be *O*-GlcNAcylated. During diabetes, transcription factor SP1 is hyper-*O*-GlcNAcylated. The *O*-GlcNAcylation of SP1 activates transcription of several genes associated with diabetes but, suppresses the transcription of several other genes.⁷¹ It has been proposed that hypo-*O*-GlcNAcylation contributes to Alzheimer's disease, while hyper-*O*-GlcNAcylation is associated with diabetes, this continuum has been reviewed by Dias and Hart.⁷¹

1.1.1.2 *N*-Linked Glycoproteins

N-Linked glycosylation almost exclusively occurs on asparagine residues. In eukaryotes asparagine glycosylation occurs co-translationally.⁸⁷ As the newly expressed protein emerges through the SEC61 translocon⁸⁸ an oligosaccharyl transferases add a core glycan (GlcNAc₂Man₉Glc₃) *en bloc* from a dolichol linked precursor⁸⁹ to sites fitting the consensus sequence Asn-X-Ser/Thr (Where X represents a non-proline amino acid).⁹⁰ The terminal glucose residues are subsequently removed by the calnexin/calreticulin protein folding quality control machinery⁹¹ to yield a protein containing a core polysaccharide of GlcNAc₂Man₉, which can then be further processed.⁹² In contrast to eukaryotes bacterial *N*-glycosylation occurs on folded proteins, separate from the translocation machinery.⁹³

The core glycan of an *N*-linked glycoprotein is constructed from nucleotide-activated monosaccharides prior to attachment to the protein fragment in the endoplasmic reticulum by monosaccharyl transferases.⁹⁴ After attachment to the protein, the glycan

can be processed by downstream enzymes to yield a diverse selection of glycan structures.⁹⁵

N-glycans have the role of increasing protein solubility, maintaining structure and providing a ‘tag’ for molecular recognition events. Specific examples of the *N*-Linked glycans being used as a molecular recognition tag include: the quality control mechanism for protein folding,⁹¹ cell signalling⁹⁶ and cell adhesion.⁹⁷

The other observed types of *N*-glycosylation are: Man-Trp (although the evidence for Man-Trp is not unambiguous),⁹⁸ where a mannose monosaccharide is covalently attached to the indole *N* on the Trp^{98, 99} sidechain. Various carbohydrates are found linked to the guanidine N atom of Arg including Glc,¹⁰⁰ GlcNAc^{101, 102} and rhamnose.¹⁰³ The structures of these *N*-linked glycoproteins can be seen in *figure 4*.

***N*-Glycoproteins**

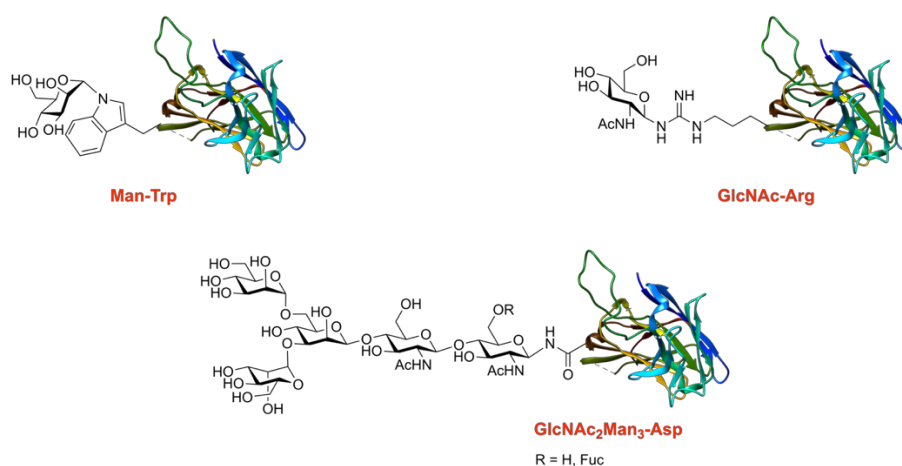


Figure 4 - Examples of N-Linked Glycoproteins showing: N- α -Man-Trp, β -GlcNAc-Arg and GlcNAc₂Man₃-Asp which can be fucosylated, GlcNAc₂Man₃ forms the basis of many glycoproteins.

1.1.2.0 Synthetic Production of Glycoproteins

1.1.2.1 Introduction

Recombinant protein expression of a glycoprotein using laboratory work horse *E. coli* yields only the poly(peptide) aglycone, while expression in eukaryotic cells will give a mixture of glycoforms. Numerous chemical and 'glycoengineering' strategies have been developed to work around the problem of glycoprotein microheterogeneity.¹⁰⁴

It is likely that both the carbohydrate and the protein segment of a glycoprotein will be in short supply therefore it is essential that their conjugation is efficient. As the chromatographic resolution of glycoforms is extremely troublesome, the method should ideally yield a single product in near quantitative conversion. Further to this efficiency requirement it is essential that the reaction method must be mild enough to avoid disrupting protein structure. In biology the glycan of a glycoprotein is placed on a specific site, therefore it should come as no surprise that the method of glycosylation must also be specific, related to this is the fact that random reaction with a charged amino acid such as lysine will cause significant disruption to the charge state of the protein and thus will likely affect its structure and function.

An obvious strategy to obtain a single modification is total synthesis, first synthesising the glycan containing amino acids, then constructing the protein using solid phase peptide synthesis^{11, 105} (SPPS) and native chemical ligations.¹⁰⁶ However, as I discussed in the introduction to this thesis the approach is lengthy, laborious and not scalable.

1.1.2.2 Nonspecific Methods

Chemical glycosylation of a protein has been accomplished by numerous methods.¹⁰⁴

A plethora of nonspecific techniques based on the modification of amines are available.¹⁰⁴ Most notably of these nonspecific techniques is the reductive amination of the reducing end of a carbohydrate to either the *N*-terminus of the protein or a lysine side chain.¹⁰⁷

Due to the unspecific nature of reductive amination in the presence of a large number of primary amines this technique leads to difficulty controlling both the number of carbohydrates added and the position on the protein to which they go. The conjugation reaction proceeds through the monosaccharide on the reducing end being in the open chain form, if this is not available then the carbohydrate can be oxidised with IO_4^- to provide aldehydes¹⁰⁸⁻¹¹⁰ however, this only leads to more microheterogeneity. The reductive amination method yields a bond which more closely resembles a glycation product rather than that of a glycosylation. Schemes for reductive amination methods can be seen in *figure 5*.

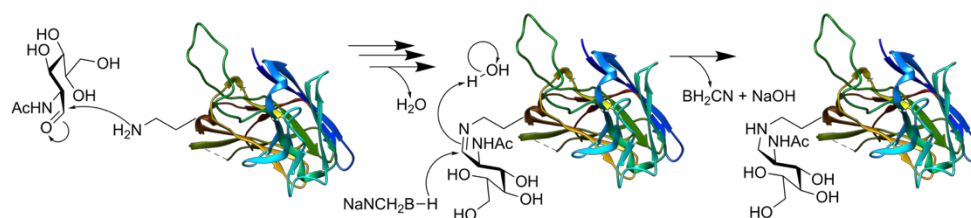
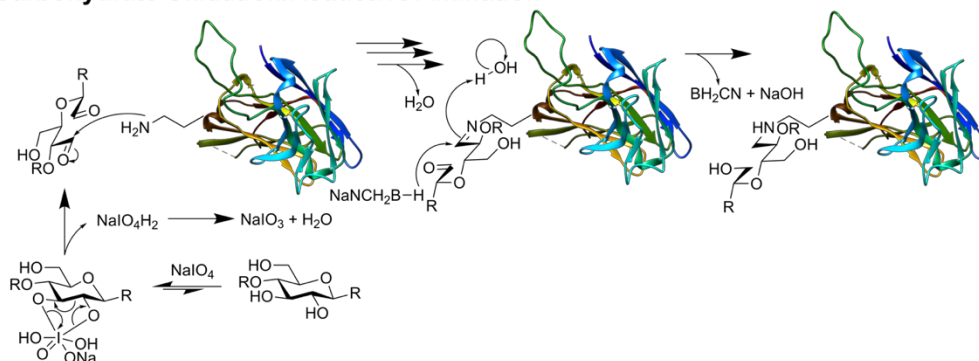
Reductive Amination**Carbohydrate Oxidation/Reductive Amination**

Figure 5 - Reductive Amination Methods for Protein Carbohydrate Conjugation: a lysine side chain amine group attacks the aldehyde of an open chain carbohydrate, eliminating water to form an imine which is then reduced with sodium cyanoborohydride. If the reducing end of the carbohydrate is not free then oxidation of an available 1,2,-diol with sodium metaperiodate can be used to oxidize to a dialdehyde which can then undergo the same process as above.

The introduction of bifunctional linkers to the carbohydrates can be an efficient method of crosslinking, many of these have been developed including: Introducing an activated ester (such as a *para*-nitrophenol or *N*-hydroxysuccinimide) on a uronic acid¹¹¹ which allows reaction with protein amines. Activation of carbohydrate hydroxyl groups with cyanogen bromide¹¹² or carboxydiimidazole (CDI),¹¹³ which after reaction with the protein amine will yield an isourea or carbamate respectively; as all are non-specific I will not dwell on them. Again, it is difficult to control both the number of

carbohydrates added and the position on the protein to which they go furthermore the bond does not resemble what biology produces.

The use of azide-alkyne dipolar cycloadditions (copper catalysed^{114, 115} is typically used) has been extensively¹¹⁶ employed as a protein carbohydrate conjugation method. CuAAC has been used non-specifically by performing azide transfer onto protein amines followed by reaction with carbohydrate bound alkynes.¹¹⁷ The reaction mechanism for CuAAC can be seen in *figure 6*.

1.1.2.3 Glycosylation using Genetically Incorporated Tags

The introduction of a non-canonical amino acids to the protein backbone using techniques such as amber codon suppression¹¹⁸ allows amino acids containing either an azide or an alkyne to be introduced at a known location. Despite getting around the specificity problem, the use of azide-alkyne dipolar cycloaddition introduces a triazole linkage, which is bulky, rigid and non-biological.

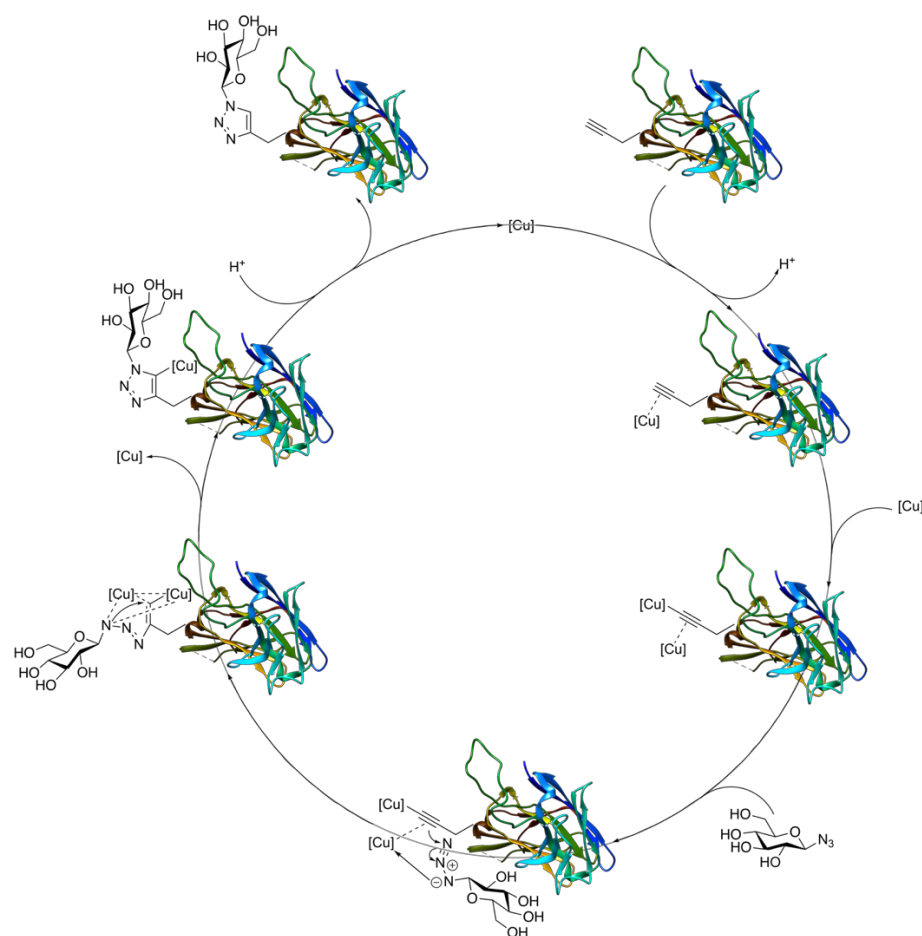
Copper Catalysed Azide-Alkyne Dipolar Cycloaddition

Figure 6 - Mechanism of Copper Catalysed Azide Alkyne Dipolar Cycloaddition¹¹⁹
the reaction proceeds by coordination of an alkyne substrate to the copper ion in η^2 coordination, a second ion of copper coordinates to the deprotonated terminus of the alkyne. The activated alkyne can now attack the terminal nitrogen atom of the azide which allows the negative nitrogen atom to coordinate to the copper atom. Attack of the first nitrogen atom to the alkene then displaces a copper atom and forms the triazole. The organic portion of the organocuprate then loses copper and gains a proton to complete the catalytic cycle.

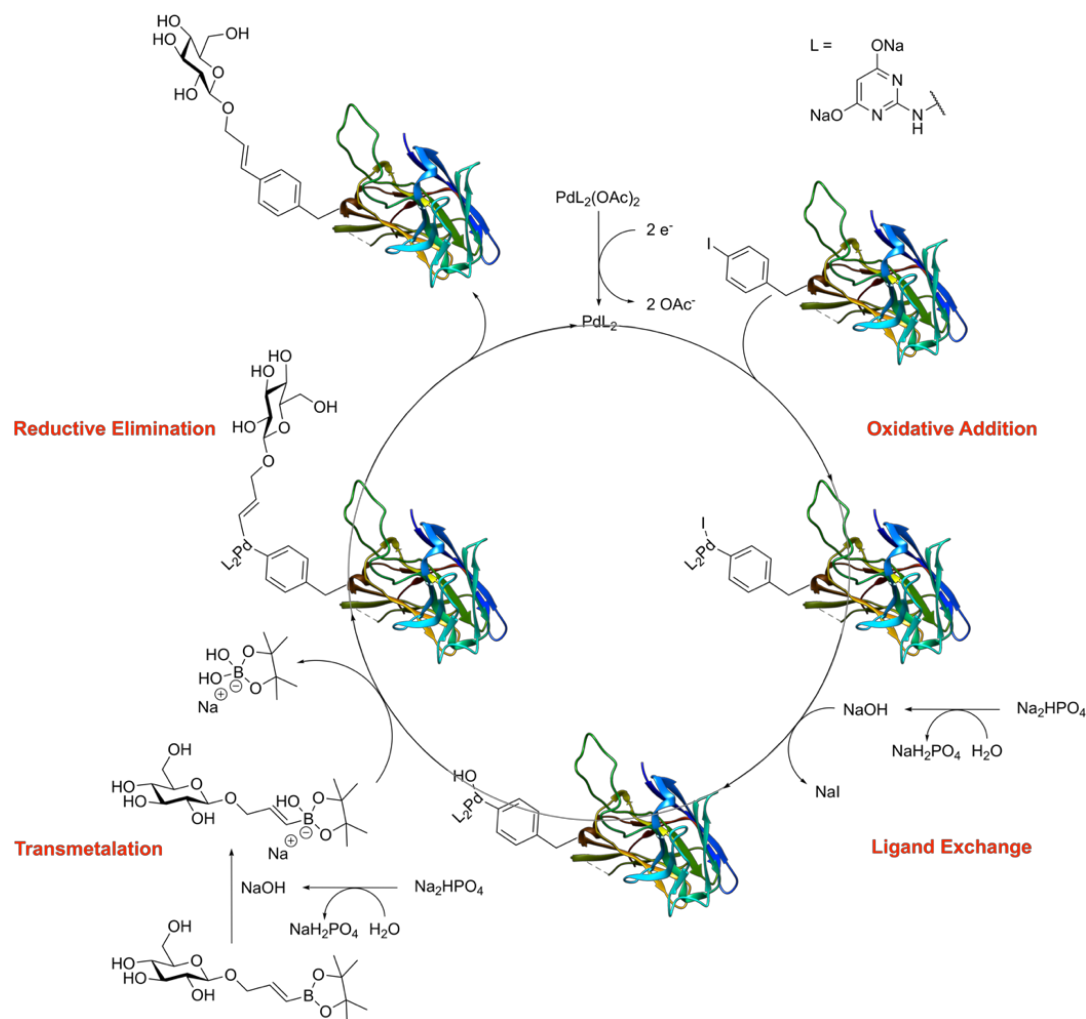
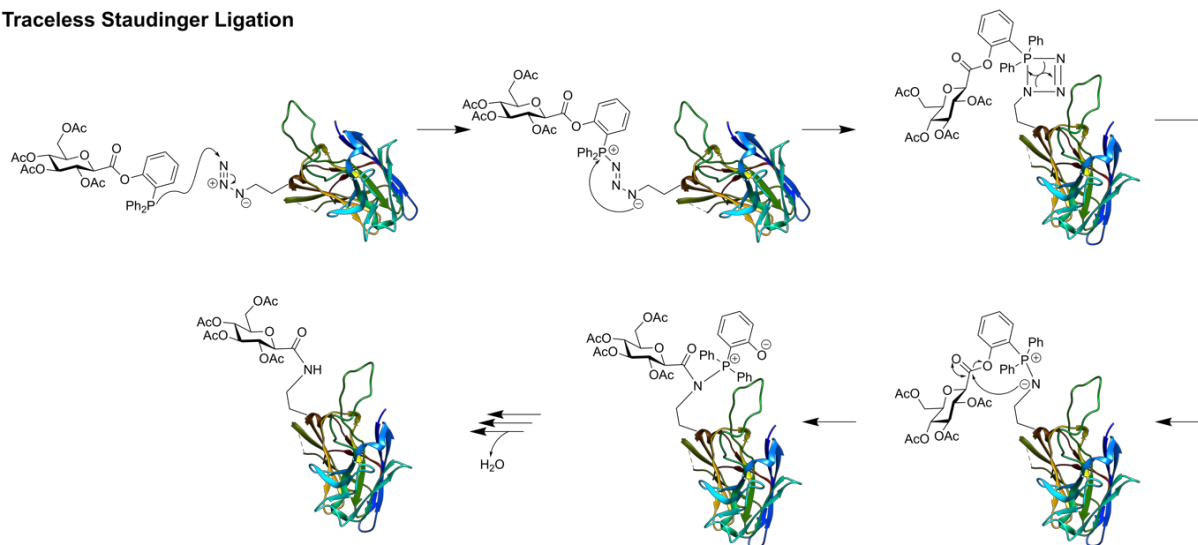
Suzuki-Miyaura Coupling

Figure 7 - Mechanism of Suzuki-Miyaura Coupling for Carbohydrate Conjugation,¹²⁰ the palladium II species is first reduced to Pd 0, which then can participate in the oxidative addition of the side chain of *para*-iodophenylalanine, the resultant iodide is exchanged with hydroxide which is then transmetalated to boron, which leaves the alkyne substituent around the palladium centre. Reductive elimination of the arene and alkene then completes the catalytic cycle.

Use of the expanded genetic code to incorporate a reactive residue for carbohydrate conjugations has been employed by more methods than just CuAAC. The genetic incorporation of *p*-iodo-L-phenylalanine serves as a handle for Suzuki-Miyaura

coupling.¹²¹ Davis previous employed a carbohydrate boronic acid to couple with a chemically incorporated aryl-iodide to generate a mimetic of Glc-tyrosine.¹²² *Figure 7* shows the reaction mechanism for Suzuki-Miyaura coupling conjugating a carbohydrate to a protein.

Traceless Staudinger Ligation



*Figure 8 - Mechanism of Traceless Staudinger Ligation:*¹²³ Attack of the phosphine to the terminal azide atom followed by cyclisation from the negatively charged nitrogen atom forms the four membered triazaphosphete ring which rapidly decomposed to release nitrogen. The negatively charged nitrogen atom which remains then attacks the ester to form an amide and hydrolysis of the P-N bond leads to the desired product.

Use of the traceless-Staudinger ligation¹²⁴ between genetically incorporated azido-*homo*-alanine and a protected Glc C-glycoside containing a *o*-diphenylphosphino-phenolic ester, yields a linkage similar to that seen in Asp glycosylation with the amide reversed.¹²⁵ The amide carbonyl is present on the carbohydrate and the amine on the protein, this can be seen in *figure 8*, along with the reaction mechanism.

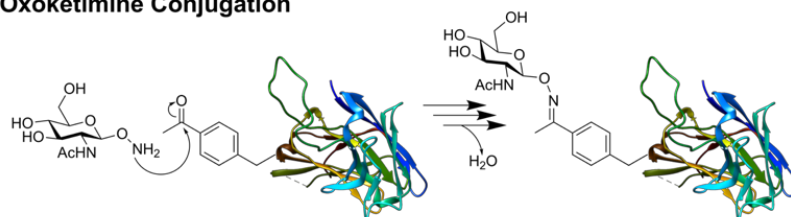
Oxoketimine Conjugation

Figure 9 – Scheme for hydroxylamine/ketone glycoconjugation. The hydroxylamine nitrogen atom forms an O-substituted ketoxamine bond.

Schultz demonstrated¹²⁶ a glycoconjugation method for reacting a carbohydrate bound hydroxylamine with the benzylic ketone of genetically incorporated *p*-acetyl-L-phenylalanine. The use of a small molecule nucleophile against a protein electrophilic site was revolutionary; this significantly lowers the potential of undesired reactivity as canonical amino acid side chains are either neutral or nucleophilic, allowing orthogonal reaction, *figure 9* shows a scheme for this glycoconjugation.

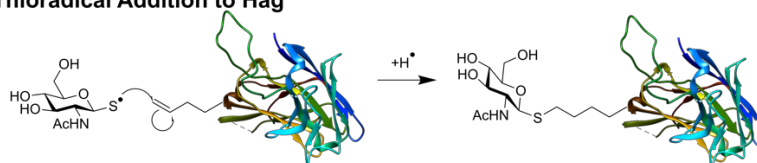
Thioradical Addition to Hag

Figure 10 - Scheme for the addition of a Thioradical to Genetically Encoded Homoallylglycine (Hag).¹²⁷ A negatively charged sulfur atom on a thioglycoside attacks the least substituted end of the alkene which hangs from the Hag sidechain.

Davis¹²⁷ employed a similar approach by genetically incorporating *homoallylglycine* (Hag) into the protein then using addition of a thio-radical to the alkene this reaction is orthogonal to other protein functionality, as shown in *figure 10*. The bond formed

by the addition of a carbohydrate thio-radical to Hag is similar to what is found in biological Cys-GlcNAc (*Figure 1*).

1.1.2.4 Glyco-Conjugations Using Cysteine

Despite showing promise the genetic incorporation of uncanonical amino acids often results in very low protein yields (single digit milligrams per litre is typical). The selective chemical modification of low abundance proteinogenic amino acids offers much more potential for the production of large amounts of protein, as the efficiency of the translational machinery is not affected.

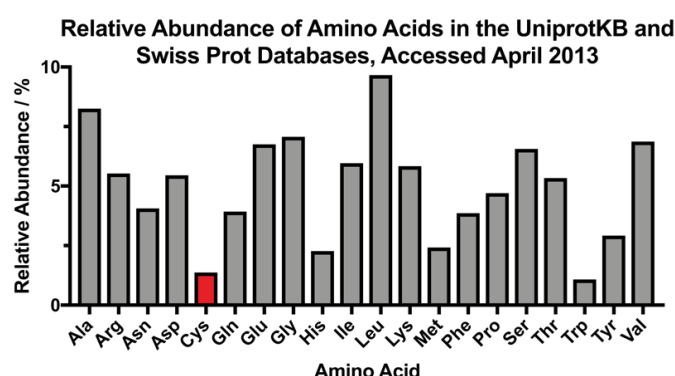


Figure 11 - Relative Abundance of Cysteine compared to other proteinogenic amino acids.¹²⁸ Cysteine is the second lowest abundance at 1.37 %. As many of cysteine residues are involved in (somewhat inert) disulfide bonds the abundance of reactive cysteine residues are lower still, making it an attractive residue for chemical modification.

Cysteine is an amino acid of particularly low abundance (1.37 %, of proteins listed in the UniProtKB and Swiss Prot databases as of April 2013, depicted graphically in *figure 11*)¹²⁸ and if present many of the cysteine sidechains are locked up in disulfide

bonds. Cysteine offers potential for site specific modification. While cysteine does not have orthogonal reactivity to proteinogenic amino acid side chains which are either nucleophilic or unreactive, it is the most¹²⁹ nucleophilic of the amino acids. With careful control of reaction conditions, the enhanced nucleophilicity of cysteine allows its selective modification by treatment with electrophiles.

Alkylation

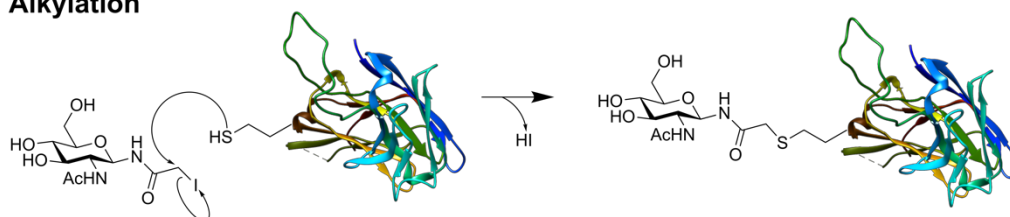


Figure 12 - Scheme of Alkylation of Cysteine with a Haloacyl linked Carbohydrate. A nucleophilic substitution reaction takes place between a cysteine side chain S atom and an α -halogeniated carbonyl, substituting the halogen atom.

Site directed mutagenesis allows placement of a cysteine residue where ever a natural glycosylation site is found, which can then be further reacted to introduce a glycan. The alkylation of cysteine residues has been employed since the 1930s,¹³⁰ but only gained popularity as a glycosylation method in the 90s. For example, to introduce an asparagine-carbohydrate like residue 1-haloacetamido- β -saccharides are incubated with a protein solution containing a single cysteine at slightly basic pH, this has been utilised by Flitsch for GlcNAcylation¹³¹ and more generally by Arsequell,¹³² figure 12 shows a scheme for these alkylations. The alkylation of cysteine provides a much more 'biological' bond than prior methods, although these are significantly further from the protein backbone than bonds found in biology and contain an abiological thioether

group. Despite cysteine being significantly more reactive than other nucleophilic amino acid side chains, there is the potential for off target reactions; especially on lysine, histidine and methionine residues.

Maleimide Conjugation



Figure 13 - Scheme of Maleimide Conjugation with a Carbohydrate Linked Maleimide. First a maleimido sugar is prepared from a 1-amino carbohydrate (protected with generic protecting groups (Pg)) and maleic anhydride by a condensation reaction. After deprotection this will react with a cysteine side chain S atom, the S atom attacks the conjugated alkene at the β position to perform a 1,4-addition.

Cysteine has also been used for nucleophilic attack to carbohydrate bound maleimides^{133, 134} (as shown in figure 13) however much like the haloacyl reagents these also show reactivity to other protein nucleophiles such as lysine¹³⁵ and also contain an abiological linkage.

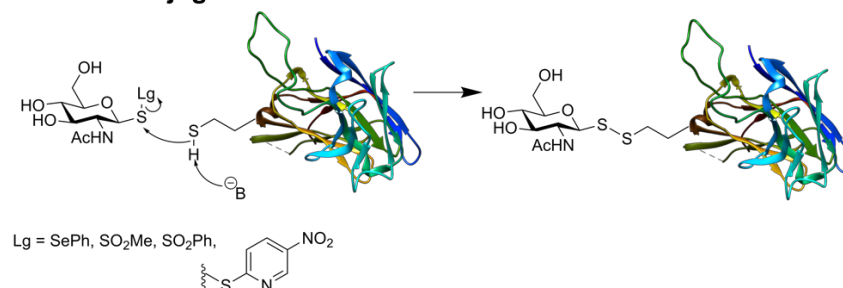
Disulfide Conjugation

Figure 14 - Scheme of Disulfide Carbohydrate Conjugation, a variety of leaving groups (Lg) can be used to form this disulfide bond. A nucleophilic substitution reaction takes place between a cysteine side chain S atom (activated by a general base (B⁻)) and the activated S atom, it is also possible that an addition-elimination reaction mechanism could occur here.¹³⁶

By activating a small molecule or protein thiol with a leaving group then treating with the corresponding partner under basic conditions a disulfide bond can be formed, as shown in *figure 14*. Disulfides are stable enough to study biological systems. Davis has used this approach to demonstrate¹³⁷ how glycans can protect a protein from proteolysis and that the strength of the effect is proportional to the size of the glycan.

1.1.2.5 Glyco-conjugations Using Dehydroalanine

Dha Formation Using DBHDA

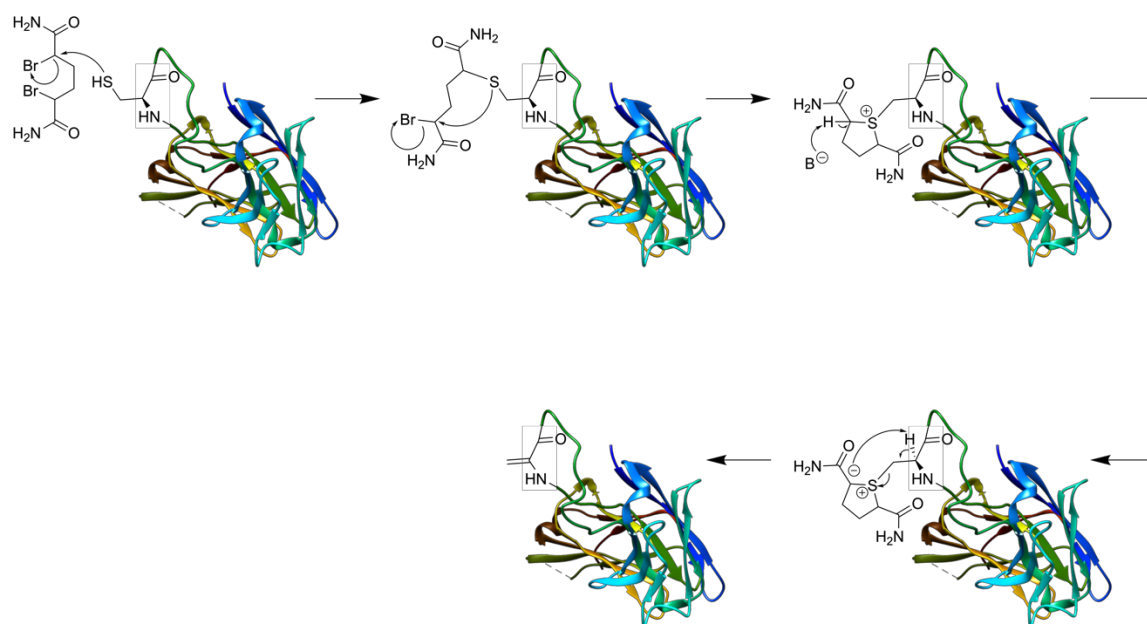


Figure 15 - Mechanism of Dehydroalanine (Dha) Formation using 2,5-Dibromohexadiamide (DBHDA): The sulfur atom of a protein bound cysteine residue attacks the bromine group of DBHDA, this is repeated to form a tetrahydrothiophenium species which is deprotonated (β -with respect to the sulfur atom) by a general base (B^-) in the buffer. In a manner similar to Swern oxidation,¹³⁸ the deprotonated carbon then deprotonates the protein backbone, which causes its elimination to give Dha.¹³⁹

One of the earliest chemical mutations performed by protein modification pioneers Koshland along with Bender¹⁴⁰ and Patchornik¹⁴¹ circumvents the promiscuous reactivity problem, by first chemically converting an amino acid side chain to dehydroalanine before reacting it further.¹⁴² Davis has employed

2,5-dibromohexadiamide (DBHDA) as a reagent for the mutation of Cys to Dha,¹⁴³ the scheme for this bisalkylation elimination mechanism¹³⁹ can be seen in *figure 15*.

Thiol 1,4-Addition to Dehydroalanine

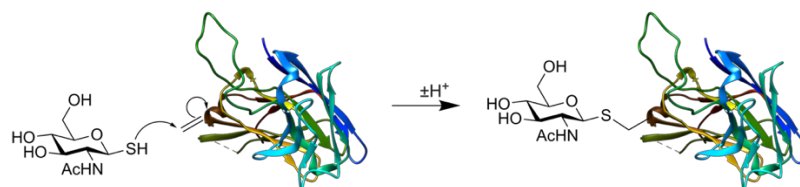


Figure 16 - Scheme of a carbohydrate-1-thiol adding to Dehydroalanine (Dha). A 1-thiolglycoside adds to α,β -unsaturated amide Dha by attack of the nucleophilic thiol group in a 1,4-manner to the α,β -unsaturated amide.

Dehydroalanine (and related dehydrobutyrine) offer a unique opportunity for protein modification due to the electrophilic nature of α,β -unsaturated amides. The strategy of first forming Dha before reacting with a thiol nucleophile was employed by Davis for protein glycosylation (*Figure 16*).¹⁴⁴ While the *S*-GlcNAc thioether resembles the biological thioether (although it is racemic) *O*-GlcNAc cases cannot cleave the *S*-GlcNAc bond.¹⁴⁵ While the stability of *S*-GlcNAc could be considered useful for the study of *O*-GlcNAcylation through such methods as interactions proteomics¹⁴⁶ and thermal-florescence¹⁴⁷ it is not effective for other applications such as the study of deglycosylation kinetics, an area which is unexplored on proteins. It is also likely that many interactions require the ether bond and will not recognise the thioether.

1.1.3.0 Introduction to Radical Addition to Dha

Organic chemistry has often been focussed on the formation of carbon-carbon bonds, dehydroalanine gives an opportunity for the conjugate addition of a species to the carbonyl group. In traditional organic chemistry this would typically be accomplished with a Gilman (organocuprate) reagent.¹⁴⁸ Obviously organometallic reagents will not be compatible with proteins due to the presence of water in which proteins are dissolved and the potential for promiscuous reaction.

Luche demonstrated¹⁴⁹ the conjugate addition of organo-radicals to α,β -unsaturated ketones and α,β -unsaturated aldehydes in pyridine enriched aqueous media, generating radicals from treating the corresponding iodo- compound with a zinc-copper couple. While the utilisation of radical reactions on proteins (and DNA)¹⁵⁰ is established,¹⁵¹ most notably in Danishefsky desulfurisation,¹⁵² there are challenges in adapting Luche's method for the use on protein bound Dha.

Conjugate radical addition to the α,β -unsaturated amide will cause a captodatively¹⁵³ stabilised radical to form on the protein backbone. The stability of this captodative radical presents potential for further reaction. Indeed these concerns for further reactivity were realised, both with atmospheric oxygen and a second organo-radical.¹⁵⁴

When atmospheric $^3\Sigma_g^-$ oxygen combines with the backbone radical, the protein fragments forming a ketone at the C-terminal fragment and an amine forms at the *N*-terminal fragment.¹⁵⁴ Sequentially the C-terminal fragment can be trapped as a biotin

hydrazone.¹⁵⁴ Use of a nitrogen filled glovebox prevented this reactivity, showing no significant amount of cleaved protein when the atmosphere has $[O_2] < 10$ ppm provided the protein had been pre-equilibrated in this atmosphere before reaction.¹⁵⁴

Recombination of the backbone radical with a second organo-radical leads to a 'double addition' product with substitution directly on the protein backbone. The first attempt to prevent double addition was to replace the Zn initiator with In which has a much lower 1st ionisation potential (906 kJ mol⁻¹ and 558 kJ mol⁻¹ respectively, cf. Na 496 kJ mol⁻¹).¹⁵⁵ The hope that a more potent initiator would lead to a high concentration of hydrogen atom type species to combine with was partially realised, showing reduced amounts of the double addition product. The small molecule and protein 1,4 radical additions are compared in *Figure 17*. Despite some success the prevention of the double addition product has to be complete due to a very limited scope for purification, especially for small modifications.

The next attempt to prevent double addition utilised dichloroindium hydride as the initiator, as popularised by Singram¹⁵⁶ with the hope that an easily available source of hydrogen would aid radical quenching. The reagent was generated *in situ* by dissolving indium (III) chloride and the iodo-compound in the protein solution then transferring by pipette directly on top of a single crystal of sodium borohydride. Although very little double addition product was observed in most cases, the reaction is accompanied by a large exotherm caused the initiation event, which could easily be felt through the

glove box glove. A control experiment lead to the discovery that sodium borohydride alone could initiate the reaction, with the *caveat* that if a high enough concentration of radicals was not formed, some or all of the Dha would be reduced to inert alanine.¹⁵⁴

While the conversion of Dha to alanine presents little change in mass and is therefore difficult to monitor by mass spectrometry reduction can be checked for by a chemical test such as reacting sample with a thiol such as β -mercaptan ethanol.^{144, 154} An overview of this reaction can be seen in *Figure 17*.

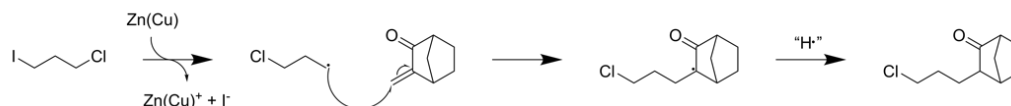
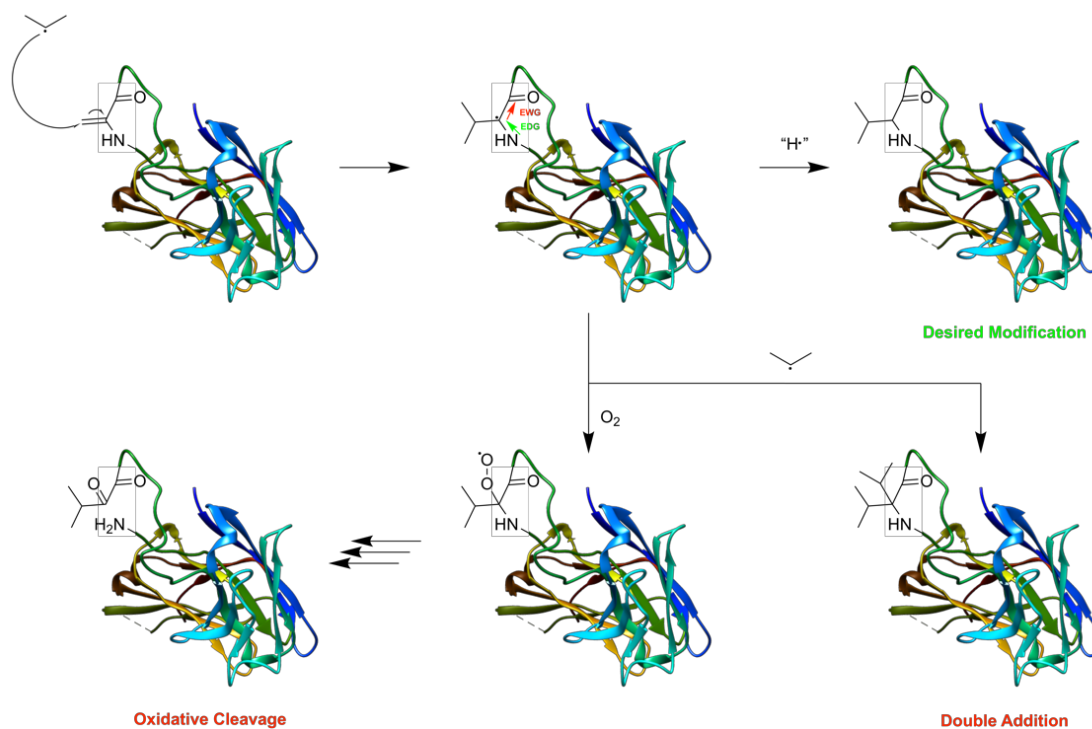
Small Molecule Radical Addition to α,β -unsaturated Carbonyl**Macromolecule Radical Addition to α,β -unsaturated Carbonyl**

Figure 17 - Schemes Comparing Organo-radical Addition to an α,β -unsaturated carbonyl as part of a small molecule and as part of a protein: For a small molecule the α -radical generated is rapidly quenched, however for a protein the amide bonds either side of the radical centre provide captodative stability due to electron withdrawing group (EWG) carbonyl and electron donating group (EDG) ketone, which can allow this centre to react with molecular oxygen of a second organo-radical instead of the desired hydrogen atom.

1.1.4.0 Aims

While numerous site specific and nonspecific techniques have been developed to introduce glycans to proteins, the bonds that links the carbohydrate to the protein are not found in biology, which can have large effects on their function. I thought that glycosylation would be an exciting application of our fledgling chemical mutagenesis method.

I set out initially to produce two distinct protein GlcNAcylation methods, to form mimics of: *O*-linked GlcNAc-Ser and *N*-linked GlcNAc-Asn. Should the additions prove successful I wanted to demonstrate that the glycans would be recognised by enzymes, showing their relevance to biology. Demonstrating the action of enzymes on my synthetic glycoprotein will allow potential for the study of enzymatic action on a homogenous glycoprotein *in vitro* which would be extremely difficult to do by any other method. I aimed to demonstrate that glycosylation with chemical mutagenesis methodology could be applied to a challenging protein, not just model substrates.

1.1.4.1 Choice of Protein Sites

While the sites at which glycosylation occurs are strictly defined in biology it has been previously demonstrated¹⁵⁴ that chemical mutagenetic methods allow ready variation of reaction site. As I did not have a specific application in mind, I have employed both Histone H3 S10C and H3 K27C as reaction substrates, primarily as I had both of substrates available in large quantities. Modification sites for each chemical mutation in Chapter 1 of this thesis are tabulated in *Table 1*.

Modification	Histone H3 Site
hhSGlcNAc	K27
hhQGlcNAc	K27
QGlcNAc	S10
F ₂ QGlcNAc	S10
hhSMan	K27
hhSGlc	K27

Table 1 - Table showing the sites on Histone H3 which the various modifications were performed. Modifications are homohomo-Serine-GlcNAc (hhSGlcNAc), homohomo-Glutamine-GlcNAc(hhQGlcNAc), Glutamine-GlcNAc (QGlcNAc), 4,4-difluoro-glutamine-GlcNAc (F₂QGlcNAc), homohomo-Serine-Mannose (hhSMan), homohomo-Serine-Glucose (hhSGlc).

1.2.0.0 Results and Discussion

1.2.1.0 O-GlcNAcylation

It became immediately apparent that I could not introduce a GlcNAc serine whilst maintaining a sidechain the same length as occurs in biology. Serine has only one carbon atom in the side chain, in order to introduce a similar O-GlcNAc sidechain using the developed chemical mutagenesis method¹⁵⁴ I needed to employ a carbon centred radical. Adding an iodomethyl group to the O-1 of the GlcNAc residue would be unstable with respect to oxygen mediated elimination, therefore being the next best option a 2-iodoethyl group was chosen. These two methylene groups, combined with the single methylene group of Dha would give a chain size of three methylene groups, or a *homohomoserine* (*hhSer*, *hhS*).

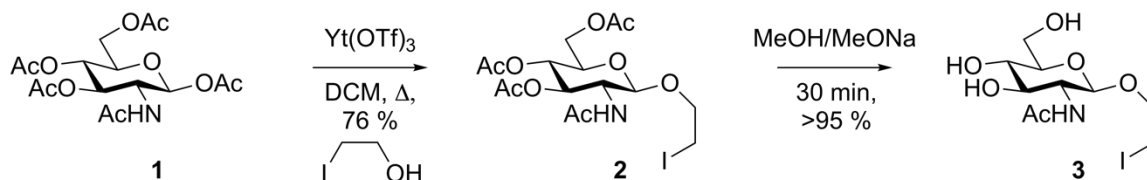
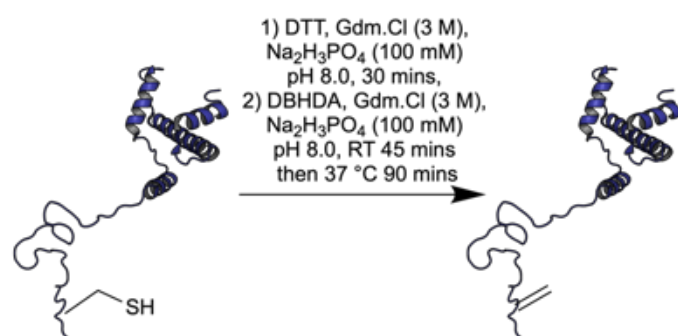


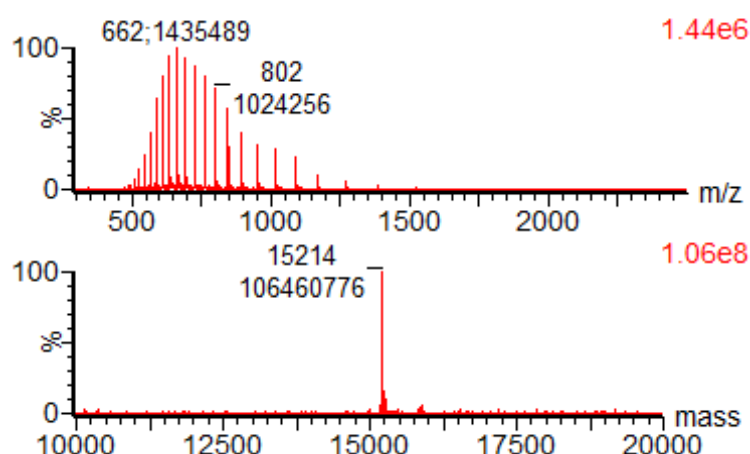
Figure 18 - Scheme for the synthesis of the O-GlcNAc Reagent: Peracetyl-β-GlcNAc **1** was heated to reflux with 2-iodoethanol and yttrium (III) triflate in dichloromethane (DCM) to obtain peracetyl-1-(2-iodoethyl)-β-GlcNAc **2** which was deprotected with sodium methoxide in methanol and quenched with Ion exchange Resin DOWEX H⁺ to give the desired product **3**.

The synthesis of the radical precursor (*Figure 18*) relied on the glycosylation of peracetylated-β-D-GlcNAc **1** with 2-iodoethanol using rare-earth Lewis acid catalysis¹⁵⁷

which gave the desired product **2** in high yield after purification by column chromatography. The β -anomer is formed due to neighbouring group participation from the adjacent amide bond. Deprotection with methanolic sodium methoxide¹⁵⁸ gave us the desired radical precursor **3**, which was crystallised from methanol. The synthesis is described in more detail in section 1.4.2. The first attempts at the addition of the *O*-GlcNAc reagent to histone H3 K27Dha (scheme of preparation from histone H3 K7C and mass spectrum provided in *Figure 19*, experimental details in section 1.4.7.1) proved successful, using 3 mg of carbohydrate and 1 mg of NaBH₄. A more detailed description of the experiments can be found in section 1.4.7.3 and the mass spectrum can be seen in *Figure 20*.



Starting Material – Histone H3 K27C



Product – Histone H3 K27Dha

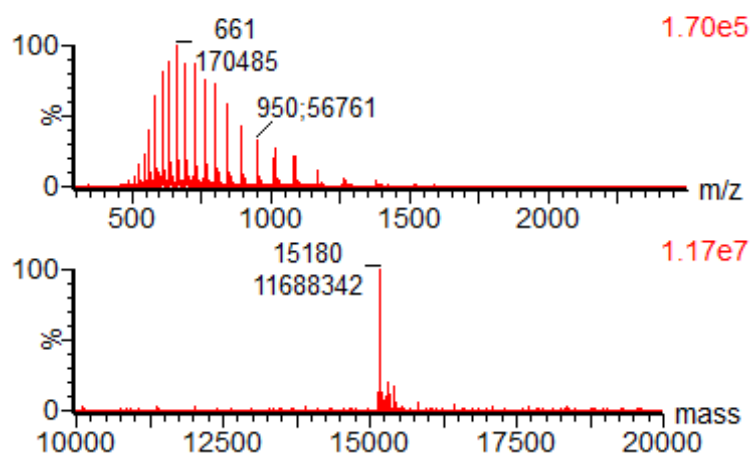
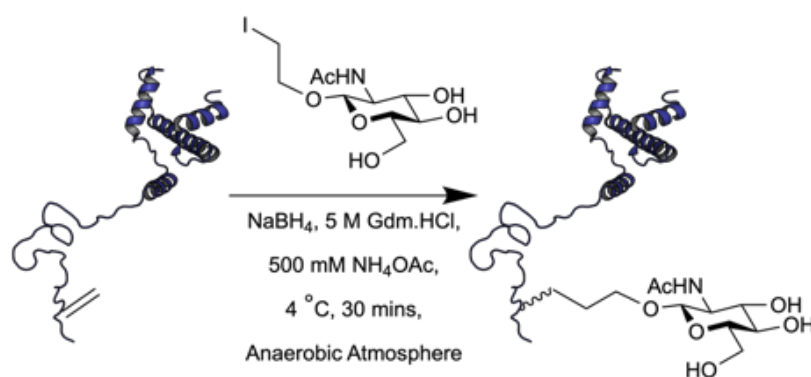
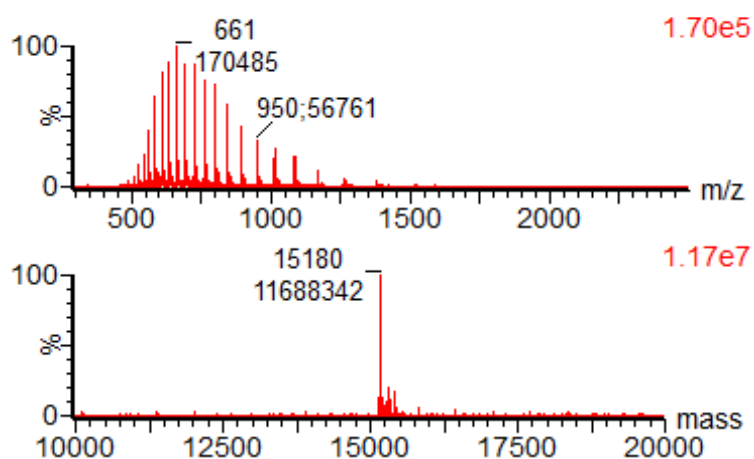


Figure 19 – Scheme for the chemical conversion of Histone H3 K27C to Histone H3 K27Dha. Histone H3 K27C was first treated with dithiothreitol (DTT) in a guanidinium chloride (Gdm.Cl), phosphate buffer. Next the protein was treated with 2,5-dibromohexanediamide (DBHDA). Note the mass difference between the starting material and product mass spectra corresponding to the loss of SH₂. Calculated Mass: 15181, Observed mass: 15180



Starting Material – Histone H3 K27Dha



Product – Histone H3 K27hhSGlcNAc

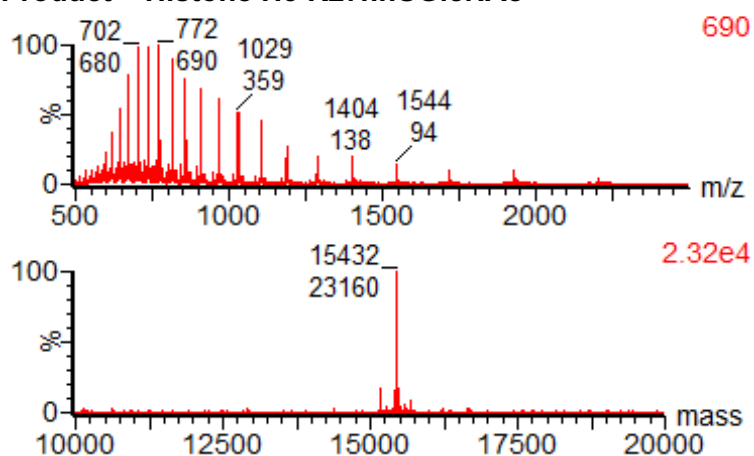


Figure 20 - Scheme for the chemical conversion of Histone H3 K27Dha to Histone H3 K27homohomo-Serine-GlcNAc. Histone H3 K27Dha in guanidinium chloride (Gdm.Cl) containing acetate buffer was sequentially treated with 1-iodoethyl-β-D-GlcNAc (**3**) and NaBH_4 . Note the mass difference between the starting material and product mass spectra corresponding to the addition of $\text{C}_{10}\text{H}_{19}\text{NO}_6$. Calculated Mass: 15430 Observed mass: 15432

1.2.2.0 *N*-GlcNAcylation

Using the same logic as applied to the *O*-linked GlcNAc I ruled out the possibility of introducing GlcNAc-Asn, extending the aglycone by an extra methyl group would make the amino acid glutamine, which was set as my target. The desired iodoacetyl-1- β -D-GlcNAc **9** had previously been synthesised by Flitsch *et al.*¹³¹

I mimicked the strategy of Flitsch,¹³¹ however replacing the stannic chloride/TMSN₃ azide coupling with NaN₃ under phase transfer conditions; EtOAc was employed as a solvent here rather than the more conventional¹³³ DCM due to safety concern. Despite concerns about the efficiency of this reaction due to the decomposition of the solvent,¹⁵⁹ they were not realised and the reaction proceeded in quantitative yield. It is likely that the side reaction problems did not occur due to how rapid the reaction is between electrophilic 1- α -chloro centre and the nucleophilic azide anion. The synthesis is described in more detail in section 1.4.2 and can be seen in *figure 21*.

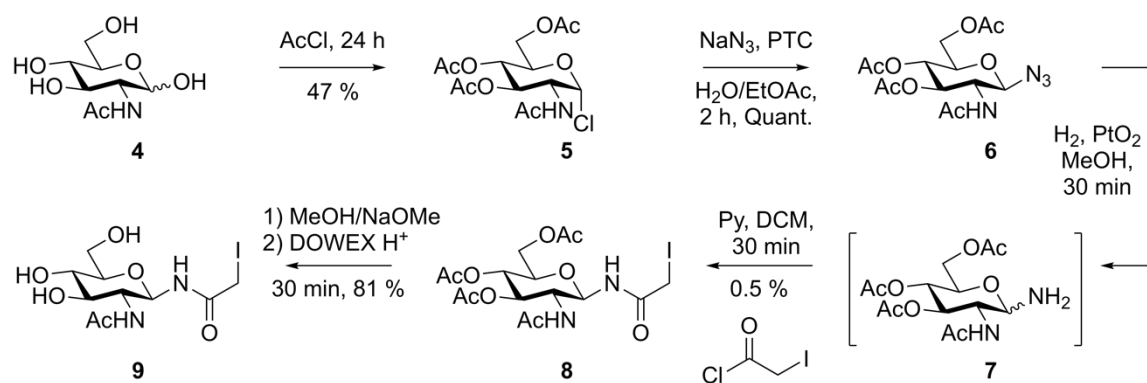
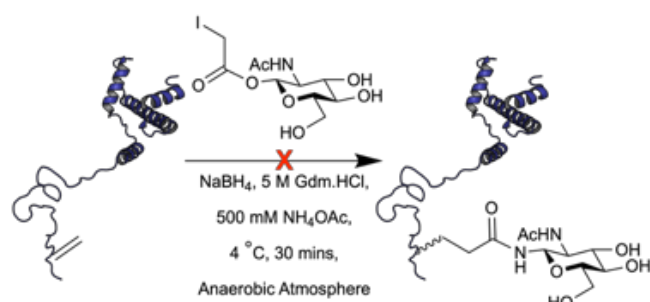


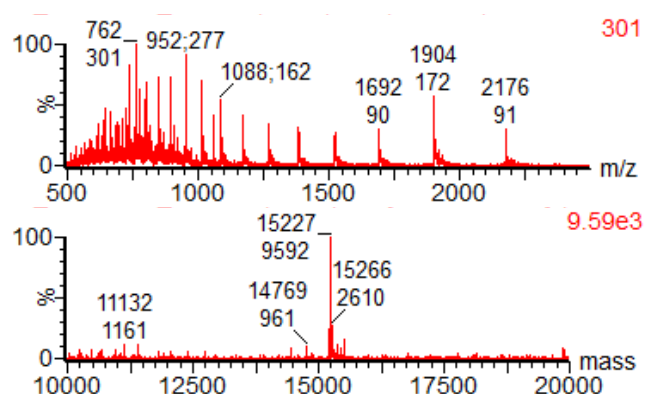
Figure 21 - Scheme for the synthesis of the *N*-GlcNAc Reagent **9**: *D*-GlcNAc is treated with acetyl chloride (AcCl) to protect and chlorinate yielding α -1-chloroperacetylGlcNAc **5**, which is treated with sodium azide under phase transfer catalysis (NBu_4HSO_4 was employed) conditions to yield the 1- β -azidoperactylGlcNAc **6**, that is reduced by catalytic hydrogenation using Adam's catalyst (PtO_2), the corresponding amine (**7**) is treated with iodoacetyl chloride in dichloromethane (DCM) with pyridine (Py) which gives us our desired product (**9**) after deprotection with methanolic sodium methoxide followed by acid work up with ion exchange resin DOWEX.

In contrast to the *O*-GlcNAc reagent, I could not successfully add the *N*-GlcNAc reagent to histone H3 S10Dha (Figure 22, experimental detail is in section 1.4.7.8) despite numerous attempts. I attributed this to the low nucleophilicity of the radical generated, which can be explained by both the electron withdrawing effect of the adjacent carbonyl group¹⁶⁰ and that the resonance structures¹⁶¹ could be drawn which would delocalise the radical into the amide bond.¹⁶² I next tried adding a methylene group to create the propyl analogue, despite trying numerous methods I could not remove the acetyl protecting groups without eliminating the iodine atom to create a stable α,β -unsaturated amide, therefore butyl was the next choice. A butyl linker would

yield a sidechain three methylene groups longer than asparagine and two methylene groups longer than glutamine, *homohomoglutamine* (*hhGln*, *hhQ*).



Starting Material – Histone H3 S10Dha



Product – Histone H3 S10A not the expected Histone H3 S10QGlcNAc

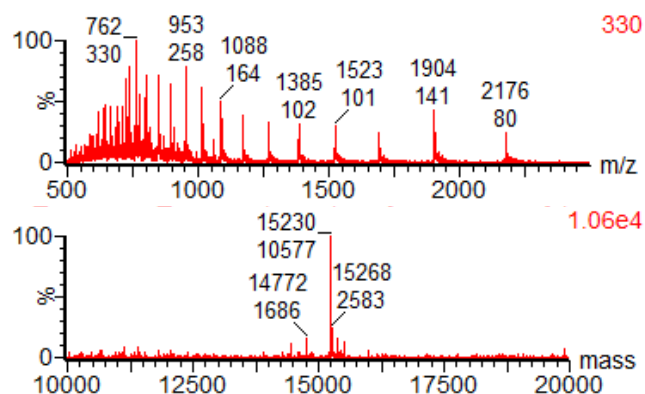


Figure 22 - Scheme for the attempted chemical conversion of Histone H3 K27Dha to Histone H3 K27GlcNAc. Histone H3 K27Dha in guanidinium chloride (Gdm.Cl) containing acetate buffer was sequentially treated with 1-(iodoacetamide)-β-D-GlcNAc and NaBH_4 . Note little mass difference between the starting material and product mass spectra. Calculated Mass: 15483 Observed mass: 15230 (either unreacted starting material or Histone H3 S10A)

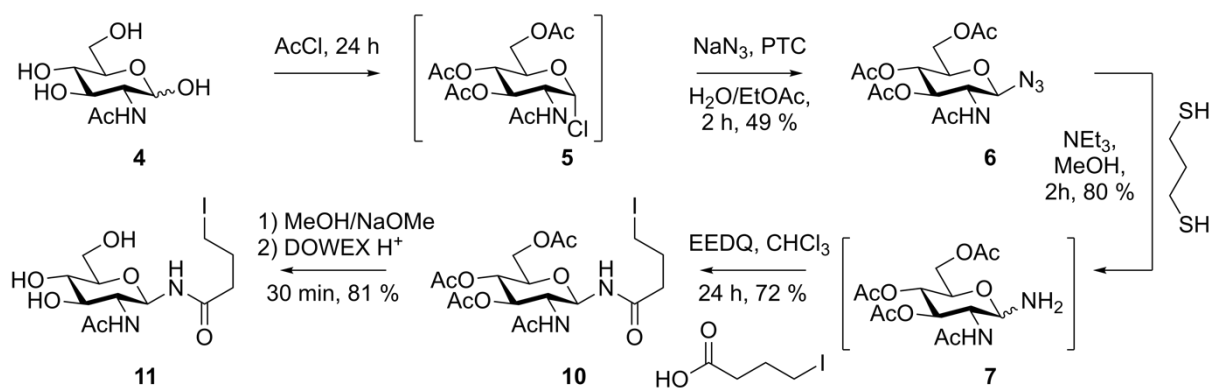
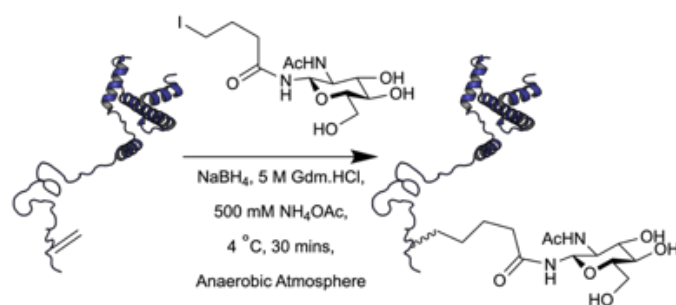


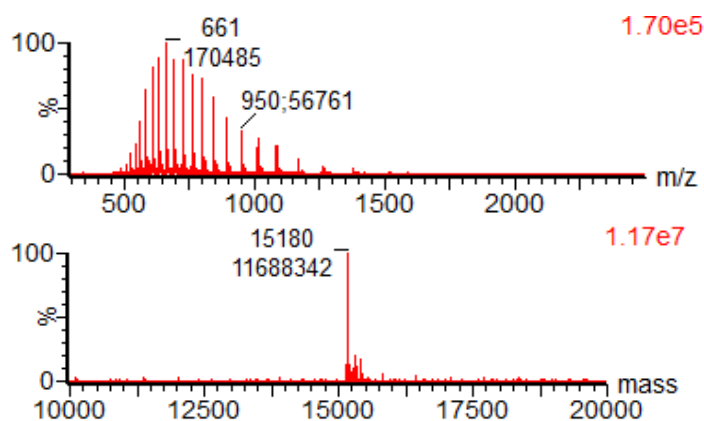
Figure 23 - Scheme for the synthesis of the N-GlcNAc Reagent: GlcNAc is treated with AcCl to protect and chlorinate yielding α -1-chloroperacetylGlcNAc **5**, which is treated with sodium azide under phase transfer catalysis conditions to yield the 1- β -azidoperacetylGlcNAc **6**, reduction with 1,3-propanedithiol with triethylamine in methanol yields the 1- β -amidoperacetylGlcNAc **7** which is treated with 4-iodobutyric acid and N-ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ) in chloroform which gives us our desired product **11** after deprotection with methanolic sodium methoxide followed by acid work up with ion exchange resin DOWEX H^+ .

Radical precursor synthesis was again conducted in a similar manner as GlcNAc-AcI, however reduction of the azide was carried out with 1,3-propanedithiol which eased isolation and amide coupling utilised an EEDQ coupling as treatment with the acidic and basic conditions required for first synthesising the acid chloride then coupling to the amine resulted in elimination of the iodine atom even at -80°C . The synthesis is described in more detail in section 1.4.2 and shown in Figure 23. Like with the O-GlcNAc reagent, I found that 3 mg (~ 500 eq) of carbohydrate and 1 mg of NaBH_4 ($\sim 2\text{k}$ eq) was required to convert Dha to GlcNAc-*hh*Gln; I speculated I was hitting the golden concentration of organoradicals at this point. A more detailed description

of the experiments can be found in section 1.4.7.4 and mass spectrometry data can be seen in *Figure 24*.



Starting Material – Histone H3 K27Dha



Product – Histone H3 K27 hh QGlcNAc

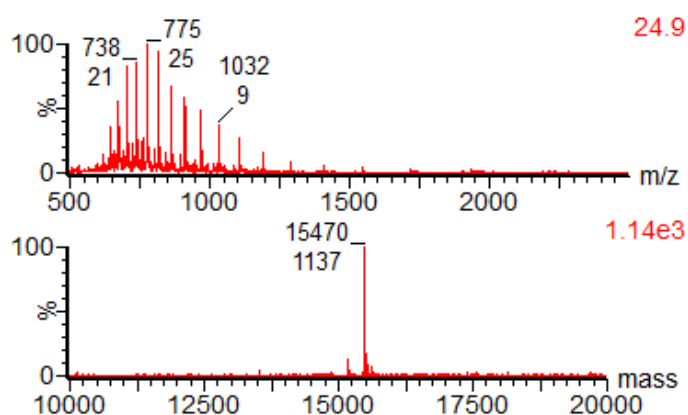


Figure 24 - Scheme for the attempted chemical conversion of Histone H3 K27Dha to Histone H3 K27homohomoGlutamine-GlcNAc. Histone H3 K27Dha in guanidinium chloride (Gdm.Cl) containing acetate buffer was sequentially treated with 1-(4-iodobutylamido)- β -D-GlcNAc and NaBH_4 . Note the difference between the starting material and product mass spectra corresponding to the addition of $\text{C}_{13}\text{H}_{24}\text{N}_2\text{O}_6$. Calculated Mass: 15471 Observed mass: 15470

1.2.3.0 Glycoengineering of Synthetically Incorporated Glycans

While the addition of monosaccharides containing native linkages to proteins could be considered a success, by enzymatically modifying the glycan I demonstrated not only that glycans conjugated by this radical method are relevant to biology but that the scope for this method is much broader if I can diversify the glycans created. Broadly glycan modifying enzymes come in three classes, enzymes which extend the glycan, enzymes which trim the glycan and enzymes which perform functional group interconversions on the glycan. I aimed to find examples of both extending and trimming enzymes which were functional on my glycoprotein substrates. For all of this section more detailed description of the experiments can be found in section 1.4.8.

1.2.3.1 Extending Enzymes

Endoglycosidases (ENGases, EndoX) are used by biology to hydrolyse the β -1,4 bond between the *N,N'*-diacetylchitobiosyl residue of the GlcNAc₂Man₃ core, and will process a wide variety of carbohydrates including those with core fucosylation.¹⁶³

Biochemists have used ENGases as a powerful tool for glycan remodelling by first removing all but the leading GlcNAc ($\pm$ Fuc) then extending the remaining glycan enzymatically to give a homogenous glycoprotein.¹⁶⁴⁻¹⁶⁸

A remarkable feature of ENGases is transglycosylation;¹⁶⁸ the ability to transfer glycans to proteins, as well as their conventional cleavage. A carbohydrate activated as an oxazoline (at least Man- β -1,4-GlcNAc)¹⁴³ and a protein containing an Asp- β -

GlcNAc (or similar)¹⁴³ is required. The reaction scheme of ENGases is shown in *Figure 25*.

ENGase Reactivity

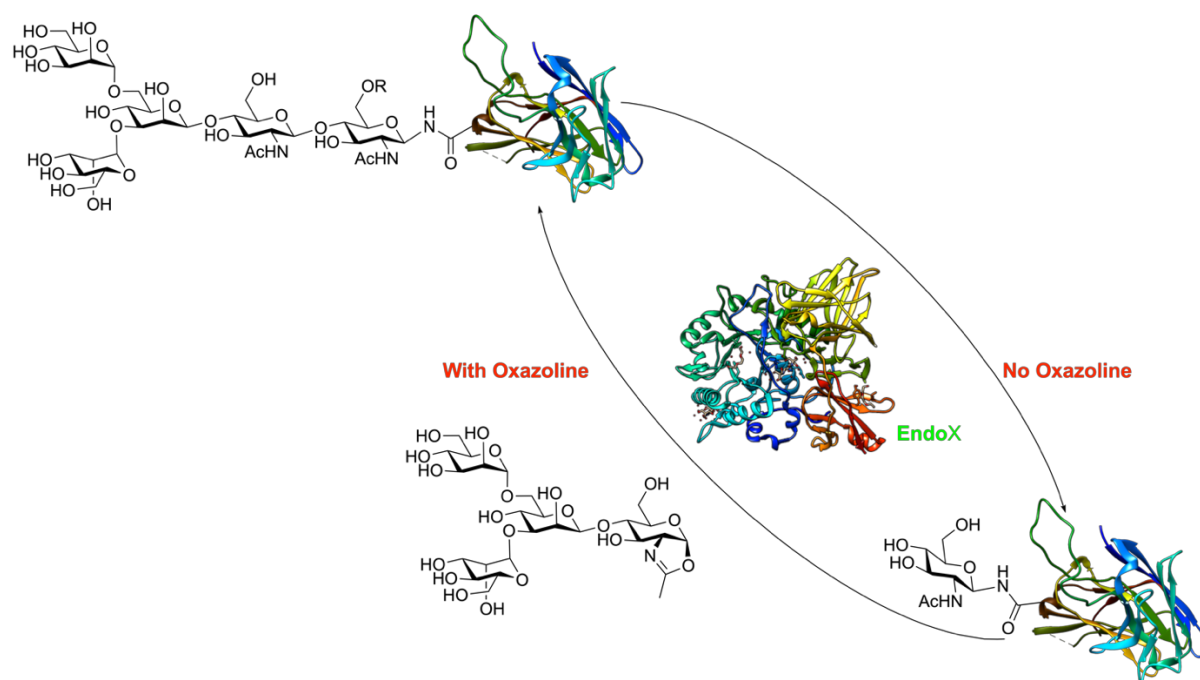
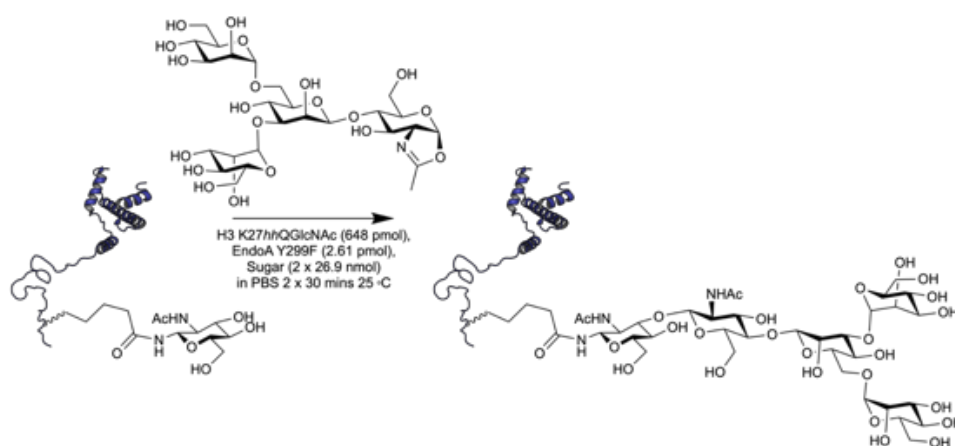


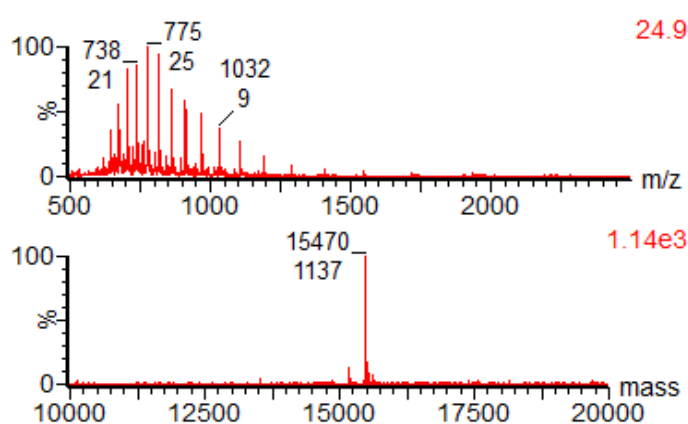
Figure 25 - Scheme of ENGase reactivity,¹⁶⁸ when a $\text{Man}_3\text{GlcNAc}_2\text{-Asp}$ Glycoprotein encounters an ENGase (the figure shows the crystal structure for EndoA) then the bond between the two GlcNAc residues is hydrolysed however, the monosaccharyl glycoprotein which results can be converted back into the original if treated with the appropriate carbohydrate oxazoline and the same enzyme

While the function of ENGases shows remarkable potential, it was not fully realised until mutants were developed¹⁶⁹⁻¹⁷¹ which restricted the hydrolysis pathway; to produce a stable glycosylation rather than a dynamic one. I set out to test if I could extend the glycan using EndoA Y299F,¹⁷¹ a particularly broad and efficient Endo variant. The extension proved successful on both the *hhGln-GlcNAc* (*Figure 26*,

experimental details in section 1.4.8.1.1) and *hh*Ser-GlcNAc (*Figure 27*, experimental details in section 1.4.8.1.2) which I had chemically introduced.¹⁵⁴ The success of these carbohydrates with native bonds is contrasted to those with a thioether bond, which are not recognised by the enzyme,¹⁶³ meaning that no trans glycosylation reaction occurs. Neither the *N*- nor *O*- GlcNAc modified proteins achieved complete conversion for transglycosylation, likely because EndoA Y299F still maintains some hydrolytic activity.¹⁷¹



Starting Material – Histone H3 K27hhQGlcNAc



Product – Histone H3 K27hhQGlcNAc₂Man₃

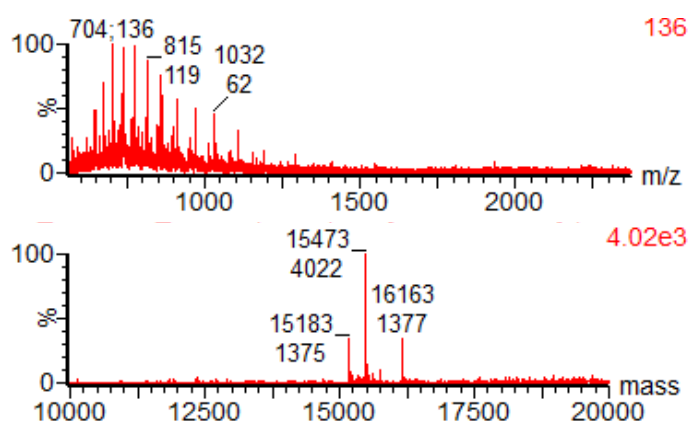
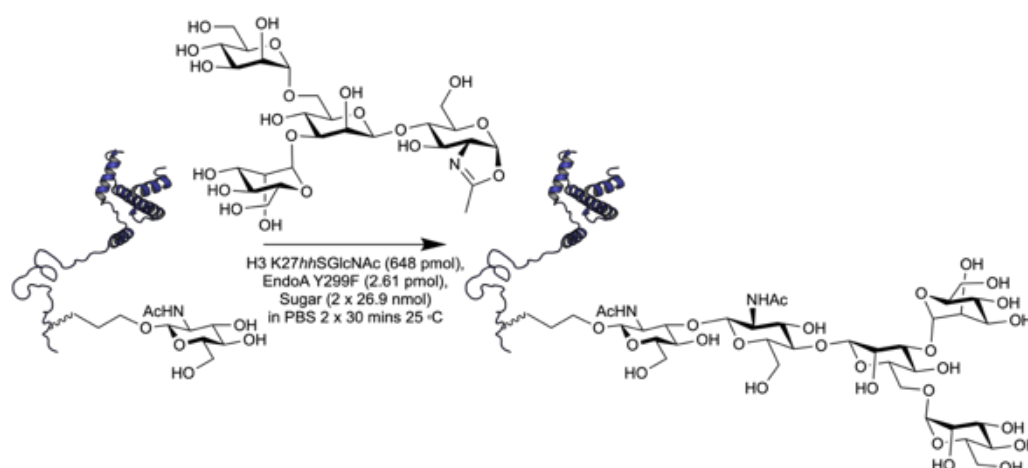
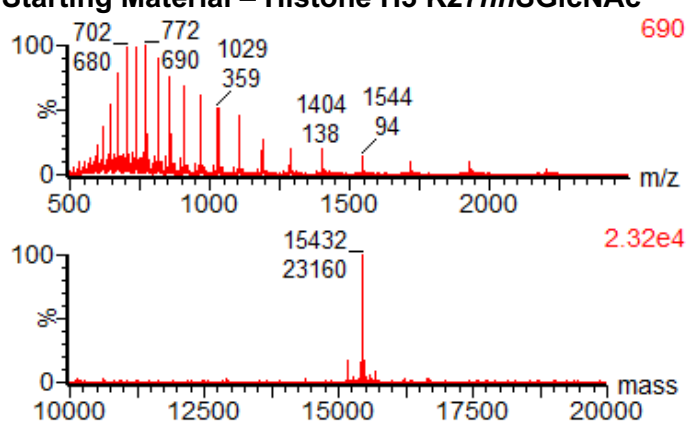


Figure 26 – Scheme for the enzymatic conversion of Histone H3 K27hhQ-GlcNAc to Histone H3 K27hhQ-GlcNAc₂Man₃. Histone H3 K27hhQ-GlcNAc in phosphate buffered saline (PBS) was treated with EndoA Y299F and the GlcNAc₂Man₃ Oxazoline at 25 °C for 30 minutes. After this period a second addition of EndoA Y299F and the GlcNAc₂Man₃ Oxazoline was carried and the reaction was allowed 30 further minutes at the same temperature. Note the difference between the starting material and product mass spectra corresponding to the addition of C₂₆H₁₄NO₂₀. Calculated Mass: 16160, Observed Mass: 16163 and 15437 (unreacted starting material).



Starting Material – Histone H3 K27hhSGlcNAc



Product – Histone H3 K27hhSGlcNAc₂Man₃

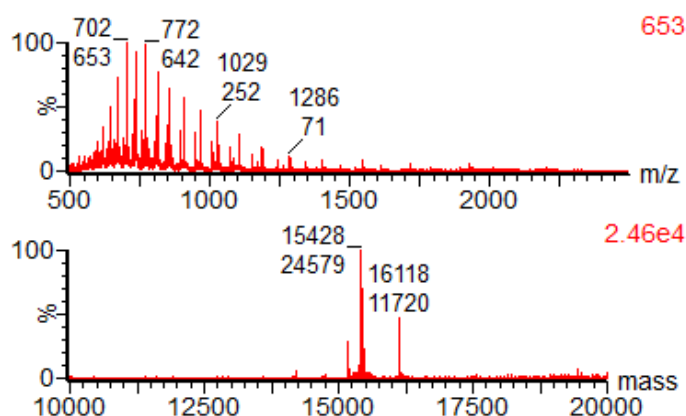


Figure 27 - Scheme for the enzymatic conversion of Histone H3 K27hhS-GlcNAc to Histone H3 K27hhS-GlcNAc₂Man₃. Histone H3 K27hhS-GlcNAc in phosphate buffered saline (PBS) was treated with EndoA Y299F and the GlcNAc₂Man₃ Oxazoline at 25 °C for 30 minutes. After this period a second addition of EndoA Y299F and the GlcNAc₂Man₃ Oxazoline was carried and the reaction was allowed 30 further minutes at the same temperature. Note the difference between the starting material and product mass spectra corresponding to the addition of C₂₆H₁₄NO₂₀. Calculated Mass: 16119, Observed Mass: 16118 and 15428 (unreacted starting material).

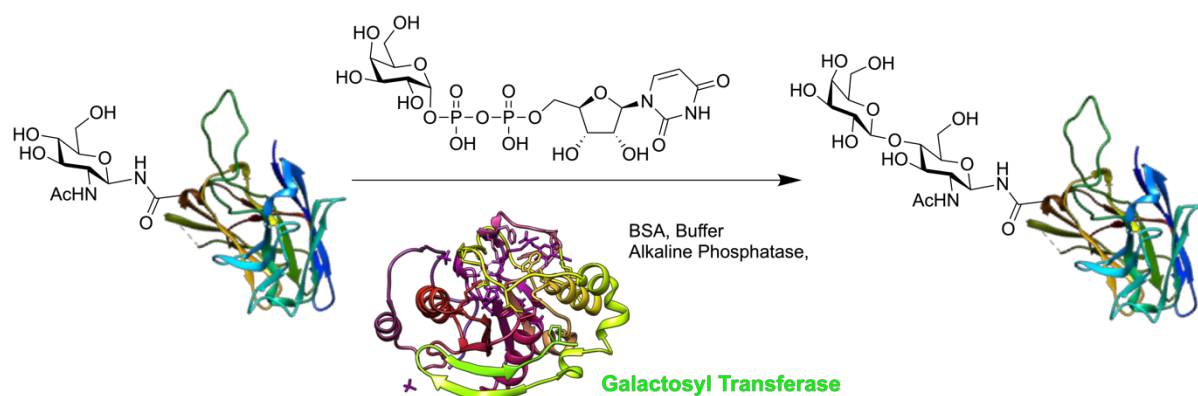
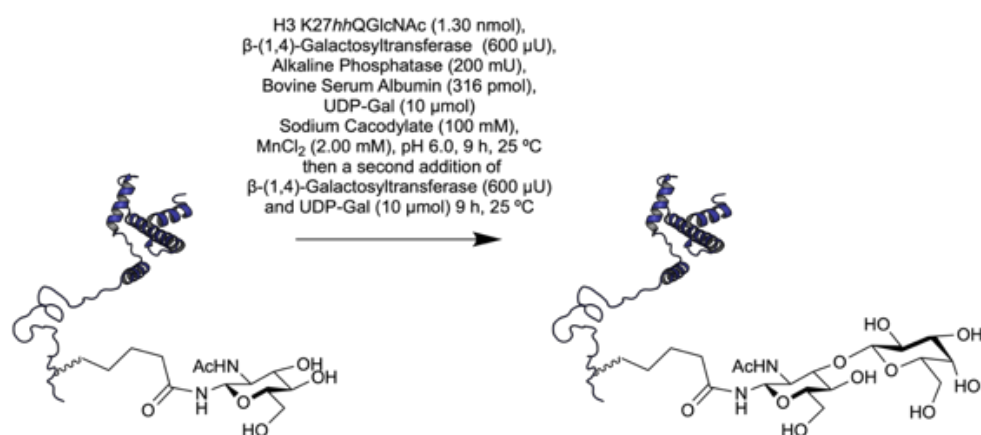
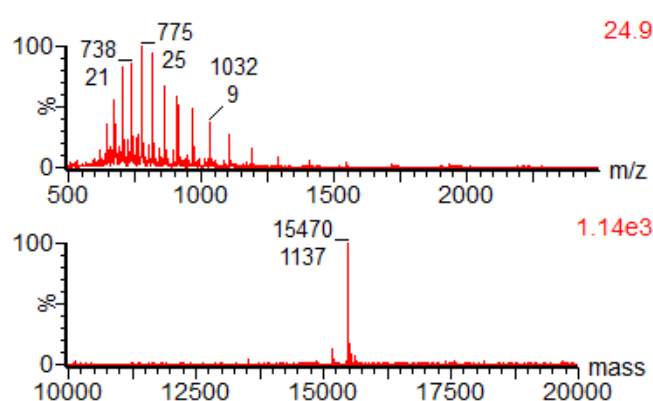


Figure 28 - Scheme of Galactose Transfer by Galactosyl Transferase (GalT), bovine serum albumin (BSA) is required to stabilise the protein while alkaline phosphatase is required to prevent inhibition by uridine diphosphate the structure for GalT represented is from a bovine (*Bos taurus*) source which is what I employed.

Galactosyl transferases¹⁷² (GalT) broadly add a galactose residue to a GlcNAc, forming a β ,1-4 bond. Activated galactose in the form of an UDP donor is used as the galactose source. Additionally, to the GalT, two further proteins are required for the reaction to proceed: alkaline phosphatase to prevent product inhibition of the enzyme and a second protein such as bovine serum albumin (BSA) or lactalbumin to stabilise the GalT, the reaction scheme is shown in *figure 28*. In order to demonstrate that our method could create glycoproteins which are not found in biology I reacted this with both radically prepared *hhGln-GlcNAc* (*Figure 29*, experimental details in section 1.4.8.2.1) and *hhSer-GlcNAc* (*Figure 30*, experimental details in section 1.4.8.2.2), both reactions proved successful, however complete conversion was not achieved with the *hhGln-GlcNA* mutant (~75 %) a longer reaction time could allow a greater conversion.



Starting Material – Histone H3 K27hhQGlcNAc



Product – Histone H3 K27hhQGlcNAc-Gal

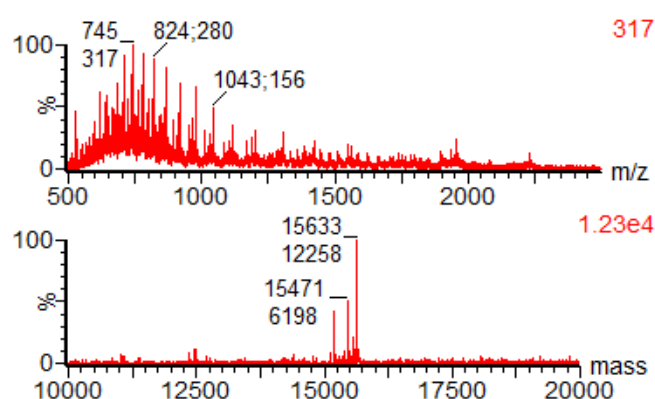


Figure 29 - Scheme for the enzymatic conversion of Histone H3 K27hhQ-GlcNAc to Histone H3 K27hhQ-GlcNAcGal. Histone H3 K27hhQ-GlcNAc in cacodylate buffer (100 mM, pH 6.0) with manganese (II) chloride (2.0 mM) was treated with Bovine β 1,4-Galactosyltransferase (GalT), uridine diphosphate galactose (UDP-Gal), bovine serum albumin (BSA) and alkaline phosphatase, then stood at 25 °C for 9 hours. After this period a second addition of GalT and the UDP-Gal was undertaken, and the reaction was allowed 9 further hours at the same temperature. Note the difference between the starting material and product mass spectra corresponding to the addition of C₆H₁₁O₅. Calculated Mass: 15633, Observed Mass: 15633 and 15471 (unreacted starting material).

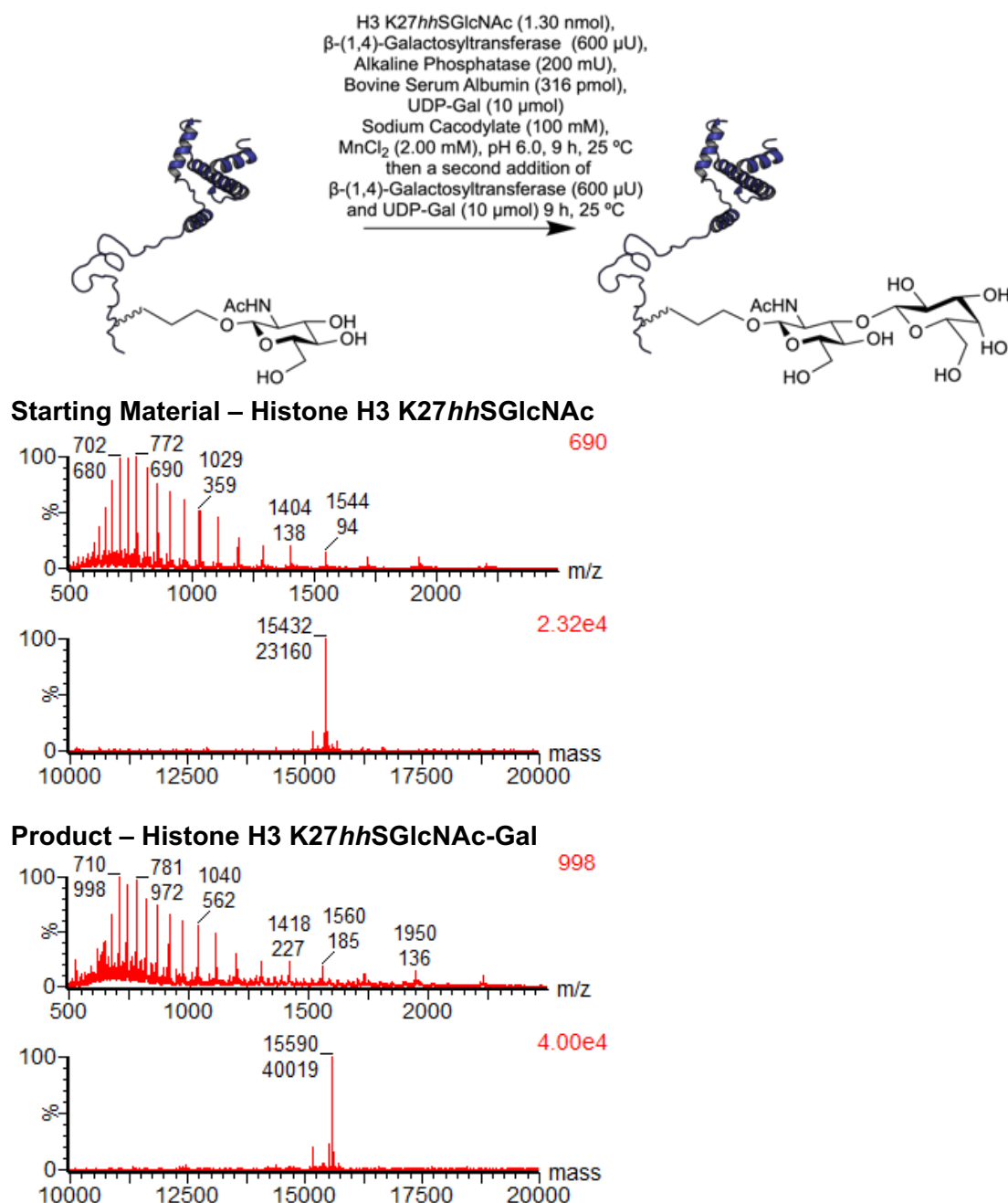


Figure 30 - Scheme for the enzymatic conversion of Histone H3 K27 hh S-GlcNAc to Histone H3 K27 hh S-GlcNAcGal. Histone H3 K27 hh S-GlcNAc in cacodylate buffer (100 mM, pH 6.0) with manganese (II) chloride (2.0 mM) was treated with Bovine β 1,4-Galactosyltransferase (GalT), uridine diphosphate galactose (UDP-Gal), bovine serum albumin (BSA) and alkaline phosphatase, then stood at 25 °C for 9 hours. After this period a second addition of GalT and the UDP-Gal was undertaken and the reaction was allowed 9 further hours at the same temperature. Note the difference between the starting material and product mass spectra corresponding to the addition of C₆H₁₁O₅. Calculated Mass: 15592, Observed Mass: 15590.

1.2.3.2 Trimming Enzymes

In contrast to extending enzymes, *hh*Gln-GlcNAc and *hh*Ser-GlcNAc require a specialised enzyme in order to remove the carbohydrate. *O*-GlcNAcases (OGAs) hydrolyse the Ser/Thr- β -GlcNAc bond, as discussed in this chapter's introduction. The mechanism of OGAs proceeds by an aspartate residue activating the amide which encourages oxazoline formation and the ejection of the aglycone fragment of the glycoprotein, this then deprotonates a second aspartic acid residue. The second aspartic acid, now in basic aspartate form, in the enzyme active site, then deprotonates a molecule of water, which opens the oxazoline ring and takes back the proton removed in the prior step, to yield GlcNAc,¹⁷³ the mechanism is shown in *figure 23*.

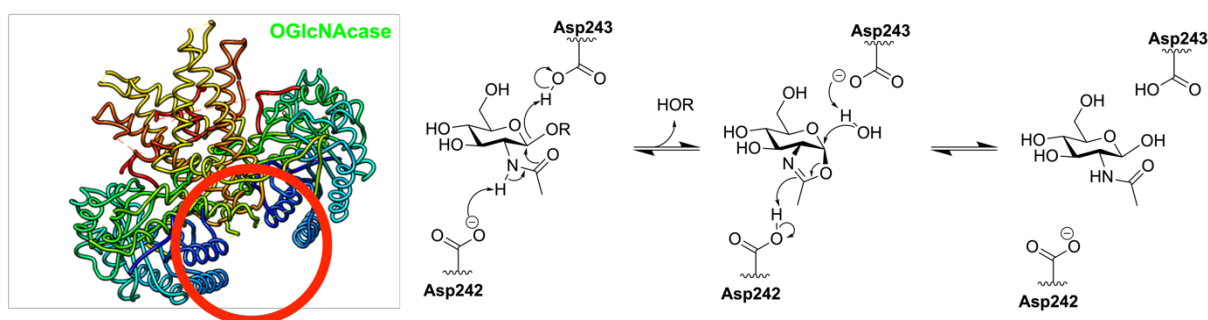


Figure 31 - Mechanism of *O*-GlcNAc Hydrolysis by an *O*-GlcNAcase (OGA):¹⁷³ Neighbouring group participation is essential for the function of OGAs in Human OGA (*h*OGA, depicted) Asp242 deprotonates the amide which starts a cascade of reactions first causing cyclisation to an oxaline which ejects the aglycone. The alkoxide aglycone deprotonates Asp243. The resulting alcohol is cleared from the enzyme active site and a molecule of water enters, much like in the previous step the deprotonation of water by Asp243 starts a cascade of reaction, first attacking and opening the oxazoline ring which deprotonates Asp242 to yield GlcNAc and returns the enzyme to its initial state.

I obtained OGAs from 2 sources Human (*Homo sapiens*)^{67, 174} (Figure 32, experimental details in section 1.4.8.3) and Jackbean (*Canavalia ensiformis*)¹⁷⁵ (Figure 33, experimental details in section 1.4.8.4) both of which could successfully cleave both *hh*Ser-GlcNAc and the control *para*-nitrophenol- β -GlcNAc (*p*NP- β -GlcAc) monitored by observing 405 nm absorbance, which corresponds to the release of *para*-nitrophenoxide from *p*NP- β -GlcNAc (experimental details in section 1.4.8.3 and 1.4.8.4).

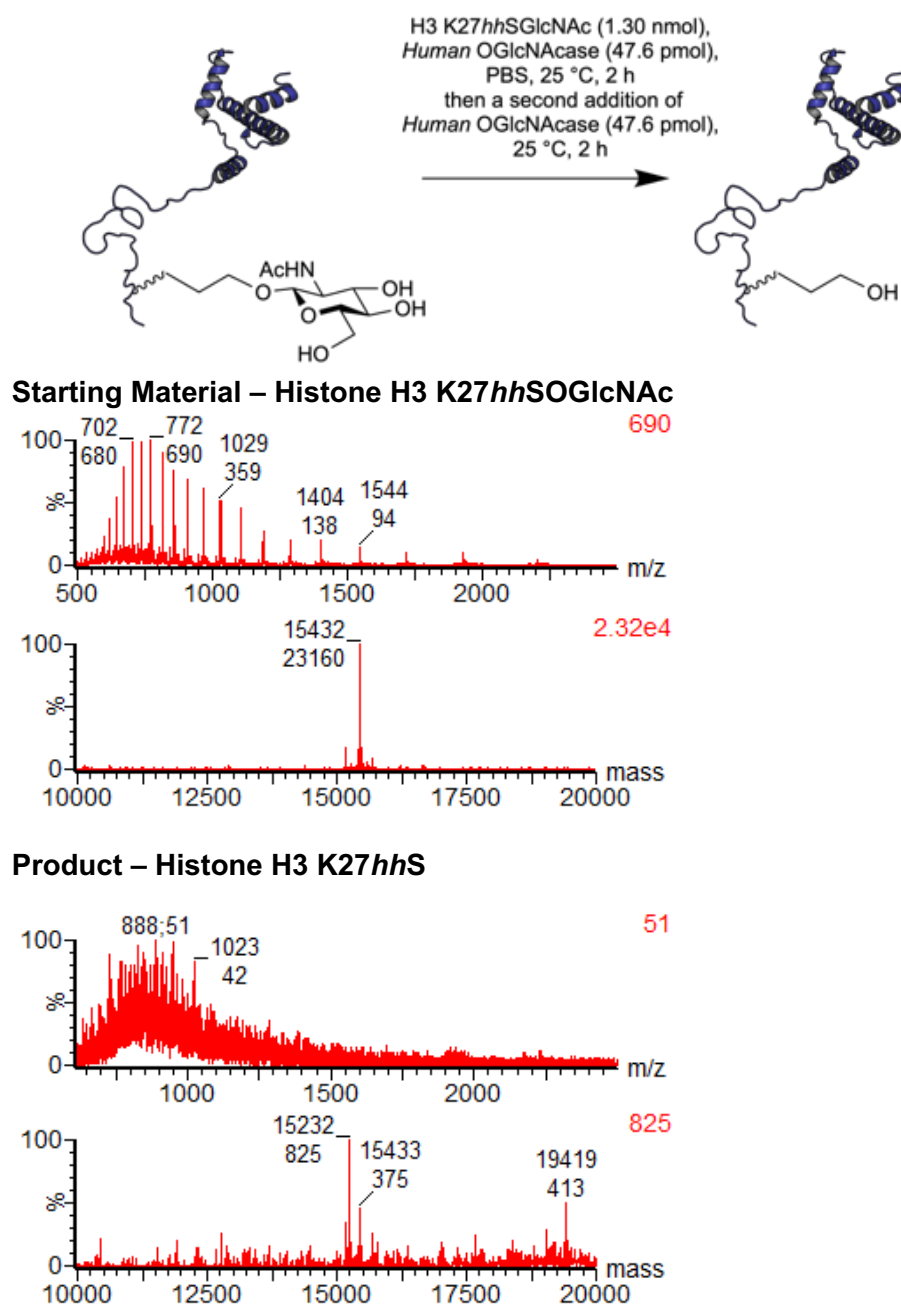


Figure 32 - Scheme for the enzymatic conversion of Histone H3 K27hhS-GlcNAc to Histone H3 K27hhS. Histone H3 K27hhS-GlcNAc in phosphate buffered saline (PBS) was treated with Human O-GlcNAcase (hOGA) then stood at 25 °C for 2 hours. After this period a second addition of hOGA was undertaken, and the reaction was allowed 2 further hours at the same temperature. Note the difference between the starting material and product mass spectra corresponding to the loss of $C_8H_{14}NO_5$. Calculated Mass: 15227, Observed Mass: 15232, 15433 (unreacted starting material) and 19419 (from the enzyme)

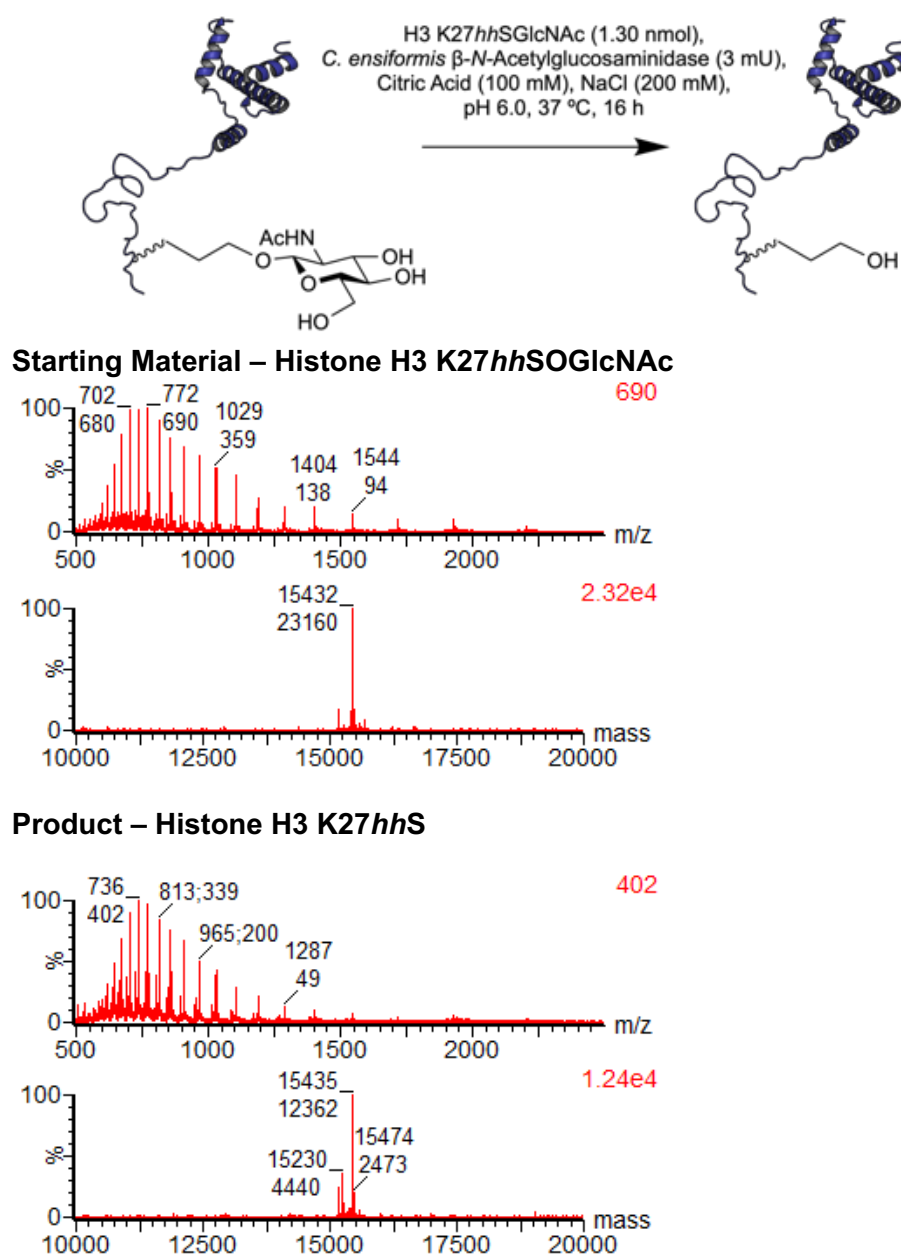
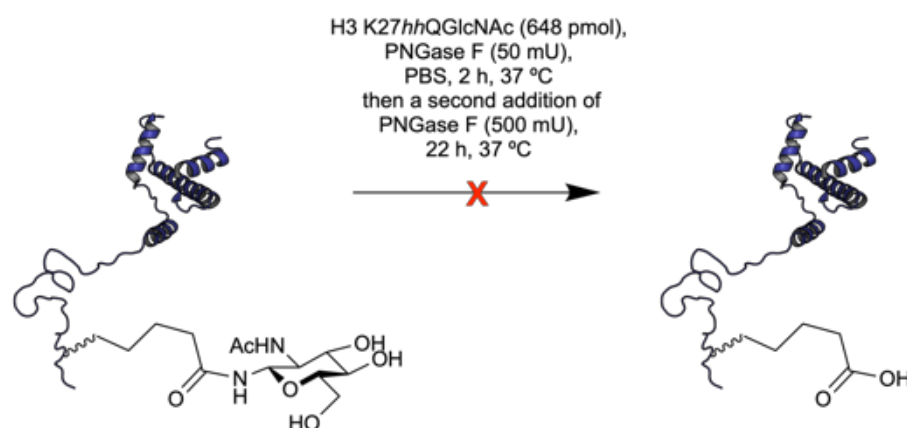
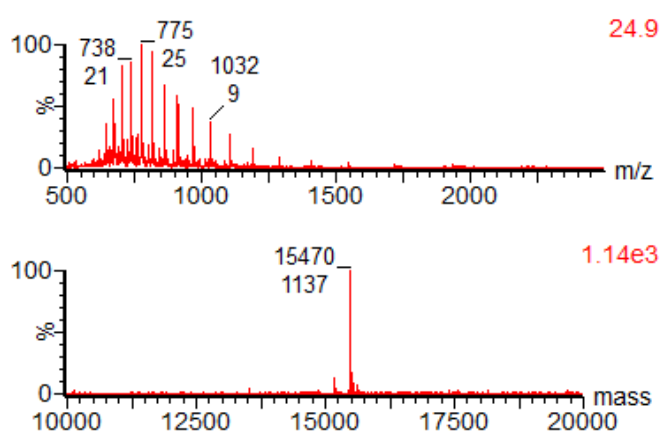


Figure 33 - Scheme for the enzymatic conversion of Histone H3 K27hhS-GlcNAc to Histone H3 K27hhS. Histone H3 K27hhS-GlcNAc in citrate buffered (100 mM, pH 6.0) saline (200 mM) was treated *Canavalia ensiformis* β -N-Acetylglucosaminidase then stood at 37 °C for 16 hours. Note the difference between the starting material and product mass spectra corresponding to the loss of $C_8H_{14}NO_5$. Calculated Mass: 15227, Observed Mass: 15230 and 15435 (unreacted starting material)

Peptide-*N*-glycosidase F (PNGase F)¹⁷⁶ is a biologically occurring enzyme in *Flavobacterium meningosepticum*¹⁷⁷ which hydrolyses a wide variety¹⁷⁸ of asparagine linked glycans. The glycosyl-asparagine hydrolysis results in a protein containing aspartic acid and an unbound glycan. Despite my sample of PNGase F being able to rapidly remove all glycans from a sample of Herceptin¹⁷⁹ (experimental detail in section 1.4.8.5.3) it failed to cleave *hh*Gln-GlcNAc (*Figure 34*, experimental detail in section 1.4.8.5.1). It has previously been reported¹⁸⁰ that PNGase F cannot cleave monosaccharides and required at least *N,N'*-diacetyl chitobiosyl glycan core in order to function.



Starting Material – Histone H3 K27 hh QGlcNAc



Product – Histone H3 K27 hh QGlcNAc not the expected Histone H3 K27 hh E

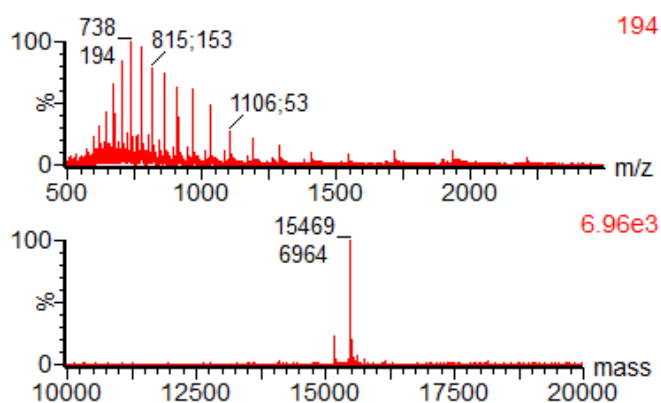


Figure 34 - Scheme for the attempted enzymatic conversion of Histone H3 K27 hh Q-GlcNAc to Histone H3 K27 hh S. Histone H3 K27 hh E-GlcNAc in phosphate buffered saline (PBS) was treated Peptide N-Glycosidase F (PNGase F) then stood at 37 °C for 2 hours. After this period a second addition of PNGase F was undertaken, and the reaction was allowed 22 further hours at the same temperature. Note little difference between the starting material and product mass spectra. Calculated Mass: 15268, Observed Mass: 15469 (starting material)

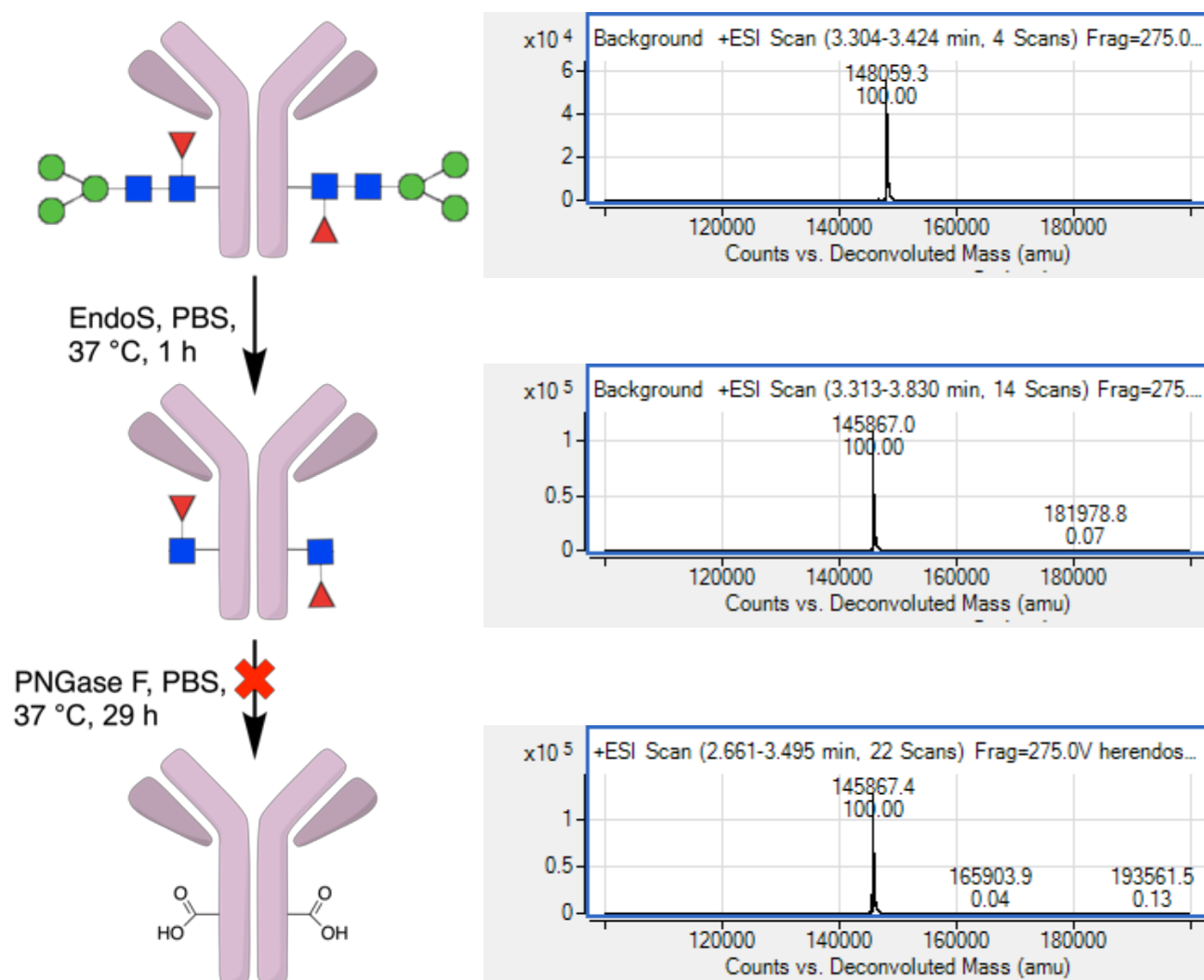


Figure 35 - Scheme of PNGase F treatment of Endo S trimmed Herceptin. A sample of Herceptin is trimmed using Endo S in PBS to give only Asp-GlcNAc-(α -1,6)-Fuc. Calculated Mass: 145868, Observed Mass: 145867, when treated with PNGase F no reaction was observed Calculated Mass: 145166, Observed Mass: 145867 (starting material), this is in contrast to a control which allows full removal when the glycan is intact. This is in agreement with literature¹⁸⁰ which states that PNGase F requires at least N,N'-diacetyl chitobiosyl glycan core in order to function

I demonstrated that PNGase F cannot cleave the disaccharide that remains when a sample of Herceptin is first treated with EndoS to leave only the GlcNAc-Fuc (α -1,6) (as seen in *figure 35*, experimental detail in sections 1.4.8.6.4 and 1.4.8.6.5). In order to prove that it is the lack of the *N,N'*-diacetyl chitobiose glycan which is responsible for the absence of activity a good experiment would be: to first prepare a protein containing GlcNAc₂Man₃-*hh*Gln with EndoA and the GlcNAc₂Man₃ oxazoline (as described in **section 1.2.3.1** and *Figure 26*) then test the activity of PNGase F, this potential scheme is depicted in *figure 36*. However, as I had exhausted my stock of tetrasaccharide, and my result agreed with literature precedent I did not investigate the problem further. An overview of all enzymatic reactions performed on my synthetic glycoproteins can be seen in *Figure 37*.

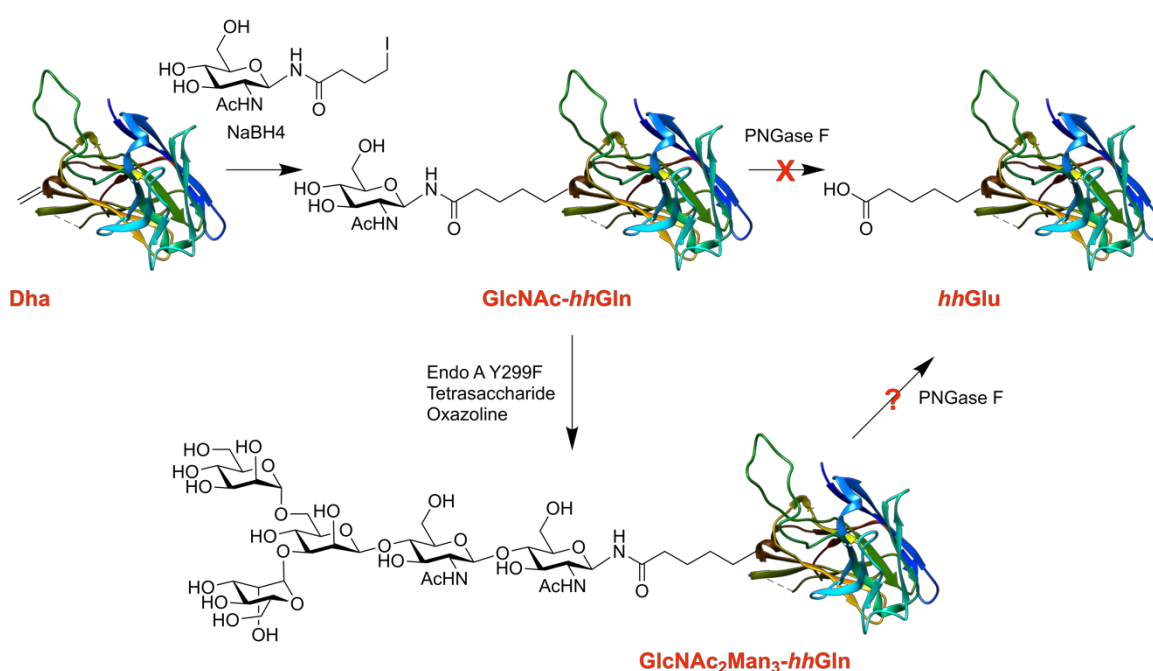


Figure 36 – Method for determining the ability of PNGase F to process synthetically incorporated N-GlcNAc. With current data it is impossible to determine whether peptide N-glycosidase F (PNGase F) is not cleaving hhQ-GlcNAc to hhE because the amino acid part of the conjugate has three methylene groups more than PNGase F's biological substrate, asparagyl glycans, or because the glycan I have incorporated is not large enough. In order to test the specificity the hhQ-GlcNAc must first be extended to hhQ-GlcNAc₂Man₃ using Endo A Y299F (as demonstrated previously in this chapter **section 1.2.3.1** and Figure 26), then treated with PNGase F. If the cleavage can occur it will be obvious that PNGase F can tolerate these 3 extra methylene groups.

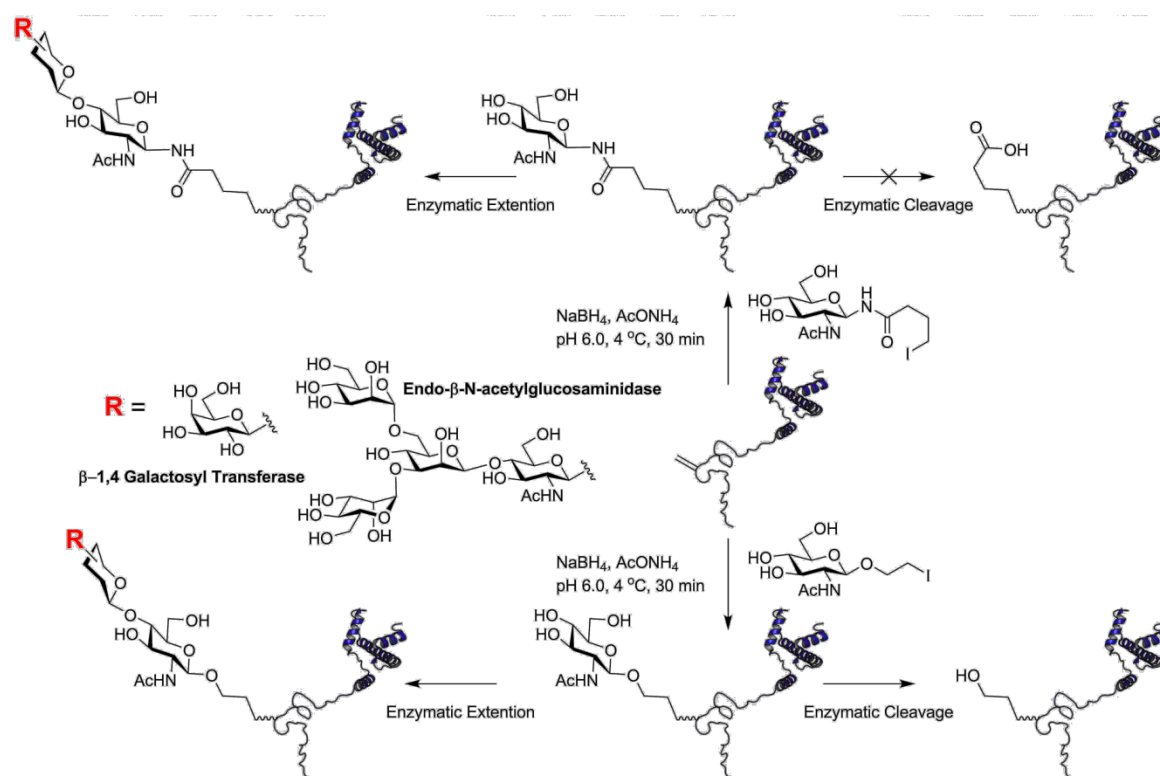


Figure 37 - Overview of Enzymatic Reactions Performed on Histone H3 glycosylated using 1,4 radical addition to dehydroalanine. Both O-Linked GlcNAc and N-linked GlcNAc can be extended with Galactosyl Transferase (to add a single galactose monosaccharide β -1,4 to the leading GlcNAc residue) and Endo- β -N-acetylglucosaminidase (to add the depicted GlcNAcMan₃ polysaccharide β -1,4 to the leading GlcNAc residue). While the O-Linked GlcNAc can be cleaved by the actions of O-GlcNAcases. The N-Linked GlcNAc monosaccharide is resistant to the action Peptide N-Glycosidase F (PNGase F).

1.2.4.0 Expansion

1.2.4.1 General O-Glycosylation

I wished to expand my method of protein glycosylation by providing evidence that this method will be successful with more carbohydrates and proteins. Initially I thought I could create a general method for O-monosaccharylation, this is accomplished by the preparation of 1-iodoethyl glycosides.

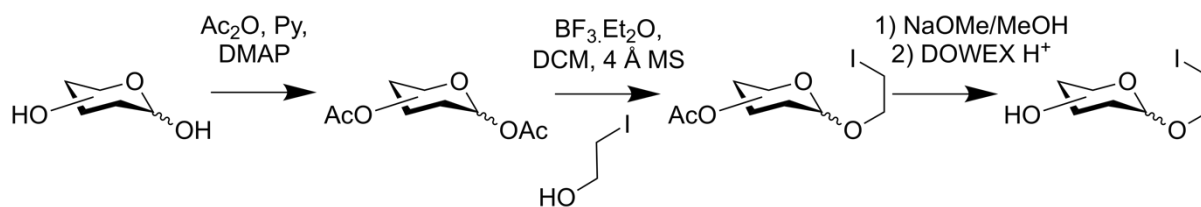
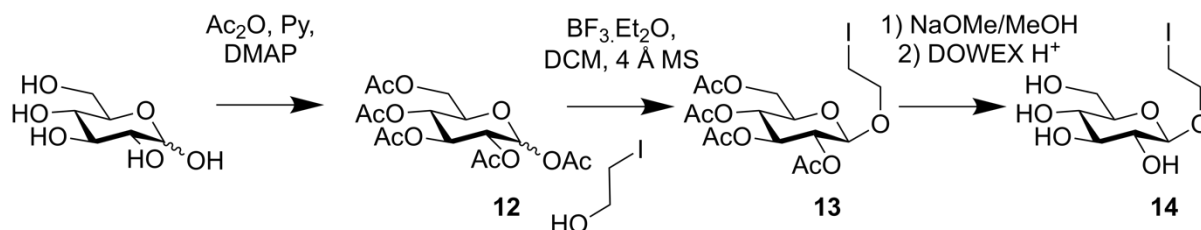


Figure 38 - Scheme for the generation of a general O-Glycosyl reagent. A monosaccharide can be protected with acetic anhydride, using pyridine (Py) as a solvent and N,N -dimethylaminopyridine (DMAP), then glycosylated with 2-iodoethanol using boron trifluoride diethyletherate as an activating agent, dichloromethane (DCM) as a solvent and 4 Å molecular sieves (MS) as a drying agent, deprotection with methanolic sodium methoxide and acidic work up with ion exchange resin DOWEX H^+ leads to get my desired product.

A general method of preparation would be to first protect the free monosaccharide with acetate groups, by suspending in ice-cold pyridine and Ac_2O , then cautiously adding crystals of N,N -dimethylpyridine (DMAP). The protected carbohydrate is glycosylated with 2-iodoethanol and borontrifluoride diethyletherate as an activator, after the glycosylation is quenched any anomeric mixture can be separated by column

chromatography. Deprotection by Zemplén deacetylation will yield the desired product. This scheme is depicted in *Figure 38*.

1.2.4.2 O-Glc



*Figure 39 - Scheme for the synthesis of the O-Glc reagent **14** D-glucose is protected with acetate groups using acetic anhydride and pyridine and a N,N-dimethylaminopyridine catalyst, followed by glycosylation with 2-iodoethanol using boron trifluoride diethyletherate as an activator in dichloromethane (DCM) with the drying agent 4 Å molecular sieves (MS) to give **13**, which after deprotection with methanolic sodium methoxide and acid work up using ion exchange resin DOWEX H^+ yields desired product **14** (1-(2-iodoethoxy)- β -D-Glucose)*

A reagent for O-Glucosylation was prepared as described in **section 1.2.4.1** and shown in *figure 31*, more detailed experimental details are shown in **section 1.4.2**. The carbohydrate was successfully added to H3 K27Dha, experimental details can be found in **section 1.4.7.6** As in previous examples 3 mg of sugar **14** (~500 eq.) and 1 mg of NaBH_4 (~2 keq.) resulted in near complete conversion to the desired product, this is depicted graphically in *figure 40*. A scheme of the reaction with mass spectrometry data can be seen in *figure 41*.

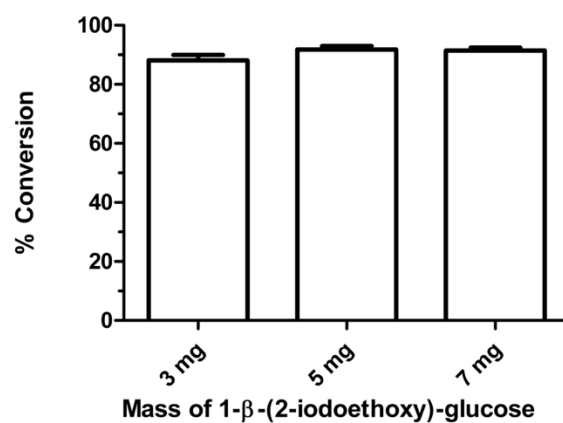
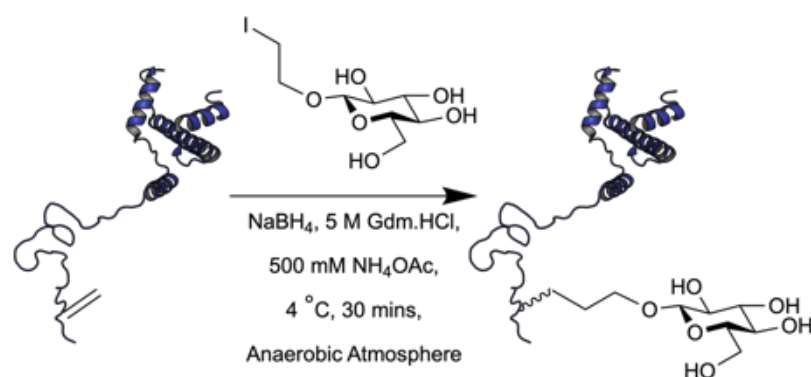
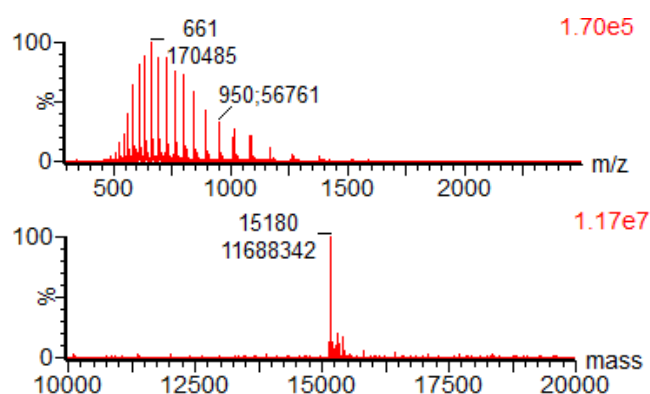


Figure 40 - Bar chart showing the conversion of H3 K27Dha (1 mg mL^{-1}) to H3 K27OGlc using varying amount of reagent **14** (1-(2-iodoethyl)-β-D-Glc) and NaBH_4 (1.0 mg). Error bars are based on SEM of triplicate experiments. Conversions were determined using mass spectrometry.



Starting Material – Histone H3 K27Dha



Product – Histone H3 K27hhS-Glc

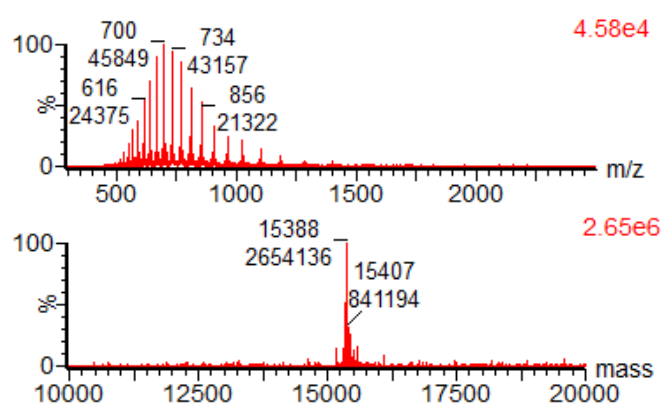


Figure 41 - Scheme for the chemical conversion of Histone H3 K27Dha to Histone H3 K27homohomo-Serine-Glc. Histone H3 K27Dha in guanidinium chloride (Gdm.Cl) containing acetate buffer was sequentially treated with 1-(2-iodoethyl)- β -D-Glc (**14**) and NaBH_4 . Note the mass difference between the starting material and product mass spectra corresponding to the addition of $\text{C}_8\text{H}_{16}\text{O}_6$. Calculated Mass: 15387 Observed Mass: 15388

1.2.4.3 O-Man

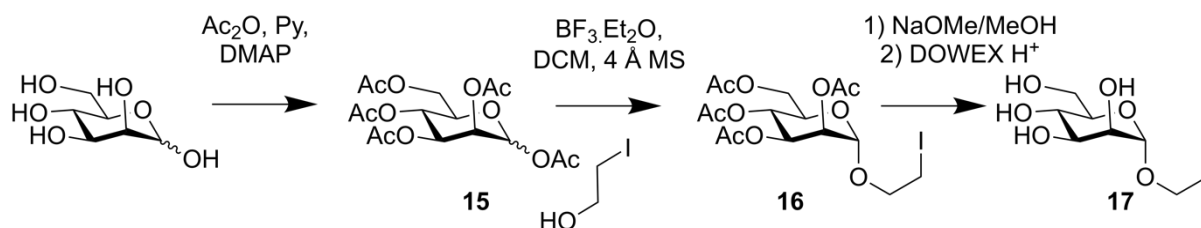
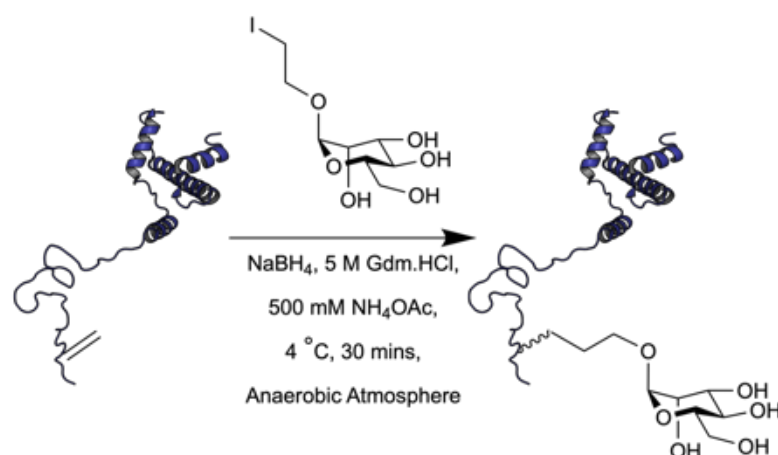


Figure 42 - Scheme for the synthesis of the O-Man reagent **17** D-Mannose is protected with acetate groups using acetic anhydride and pyridine and a *N,N*-dimethylaminopyridine catalyst, followed by glycosylation with 2-iodoethanol using boron trifluoride diethyletherate as an activator in dichloromethane (DCM) with the drying agent 4 Å molecular sieves (MS) to give **16**, which after deprotection with methanolic sodium methoxide and acid work up using ion exchange resin DOWEX H^+ yields desired product **17** (1-(2-iodoethoxy)- α -D-Mannose)

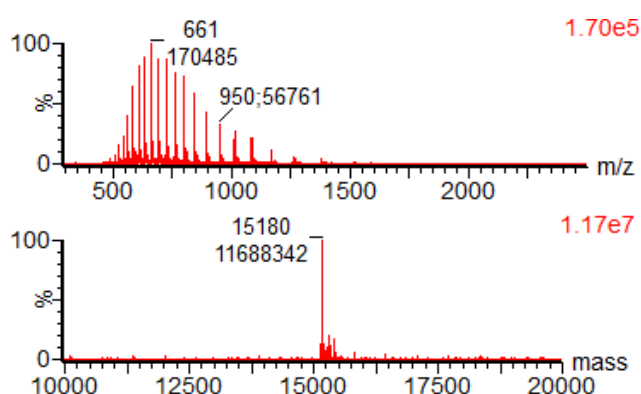
A reagent for O-Mannosylation was prepared as described in section 1.2.4.1 and depicted in figure 42, experimental details are shown in section 1.4.2. Despite numerous attempts I could not get the addition to give more than 80 %, as shown in figure 43 and with experimental details being given in section 1.4.7.7 1-(2-Iodoethoxy)-1- α -mannose (**17**) was isolated an extremely viscous oil, I had great difficulty in dissolving **17** in the protein solution. By first dissolving the mannose compound in a large quantity of water followed by lyophilization I was able to isolate the sugar as an amorphous solid. Unfortunately, after being exposed to air it rapidly turned back into a gum, but quick work allowed the weighing of around 15 samples for protein reaction. A scheme of the reaction with mass spectrometry data can be seen in figure 44.



Figure 43 - Bar chart showing the conversion of H3 K27Dha ($200 \mu\text{L } 1 \text{ mg mL}^{-1}$) to H3 K27OMan using varying amount of reagent **17** (1-(2-iodoethoxy)-1- α -mannose) and NaBH_4 (1.0 mg). Error bars are based on SEM of triplicate experiments. Conversions were determined using mass spectrometry.



Starting Material – Histone H3 K27Dha



Product – Histone H3 K27hhS-Man

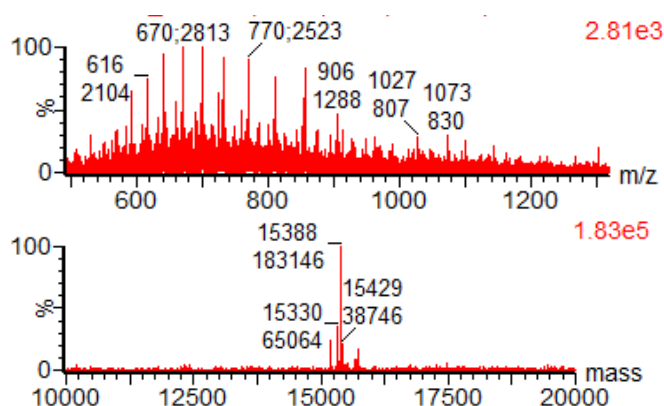


Figure 44 - Scheme for the chemical conversion of Histone H3 K27Dha to Histone H3 K27homohomo-Serine-Glc. Histone H3 K27Dha in guanidinium chloride (Gdm.Cl) containing acetate buffer was sequentially treated with 1-(2-iodoethyl)- α -D-Man (**17**) and NaBH_4 . Note the mass difference between the starting material and product mass spectra corresponding to the addition of $\text{C}_8\text{H}_{16}\text{O}_6$. Calculated Mass: 15389 Observed Mass: 15388

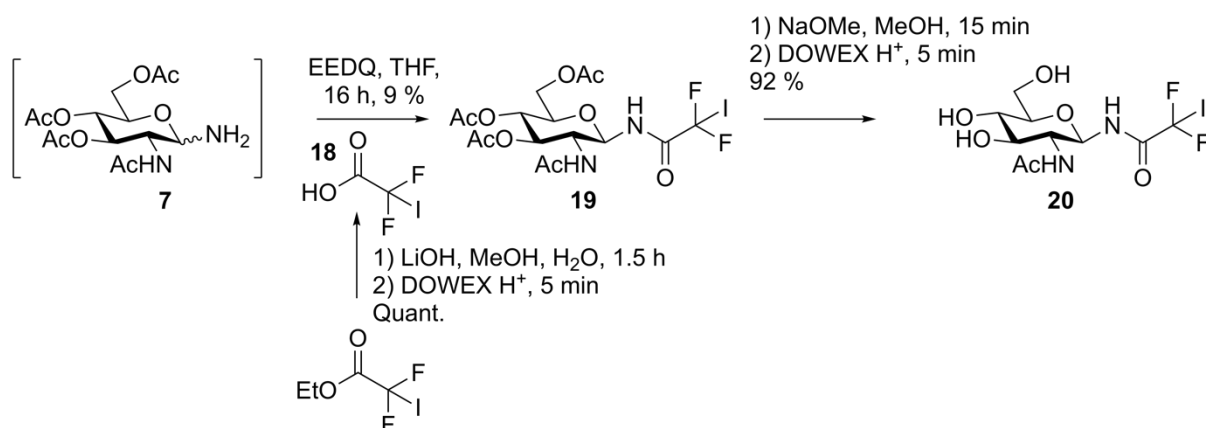
1.2.4.4 F₂-N-GlcNAc

Figure 45 - Scheme for the synthesis of the F₂ N-GlcNAc Reagent **20**: 1-Aminoperacetyl GlcNAc (**7**) is treated with N-Ethoxycarbonyl-2-ethoxy-1,2-dihydroquinoline (EEDQ) and 2,2-difluoro-2-iodoacetic acid (**18**) in THF, followed by deprotection with sodium methoxide gives our desired product **20**.

I wished to incorporate an AcF₂I linked GlcNAc; as I assumed that the difluoro radical centre would be more nucleophilic than I had observed with AcI (section 1.2.2.0) thus I could potentially see greater reactivity. On top of this increased reactivity ¹⁹F is an NMR active nucleus (spin = + $\frac{1}{2}$), previously employed for tagging proteins.¹⁸¹ As ¹⁹F is extremely sensitive to its microenvironment I could see the potential to use NMR spectroscopy to monitor the progress of enzymatic reactions. The carbohydrate was synthesised as described by figure 45 and with experimental details being given in section 1.4.2.

Surprisingly when NaBH₄ was employed as a radical initiator I observed mass difference of +82, corresponding to the addition of 2,2-difluoro-2-iodoethanol. The strong electron withdrawing capacity of the difluoroiodomethyl group could have

sufficiently polarised the amide carbonyl bond to allow reaction with NaBH_4 . Amides are typically inert with respect to NaBH_4 . I do not know if the reduction is happening before or after radical initiation or before. Likewise, I also do not know if the reduction is happening on the protein or small molecule. The experimental details are given in section 1.4.7.5.

The employment of metals as radical initiators removes the mechanism by which amide reduction can occur, therefore using Zn or In metals to initiate radical precursor 20 does not lead to the undesired side product. Despite no reduced product being observed the metal-initiated reactions did not have complete conversion in any case despite numerous conditions attempted, heatmaps of conversions can be seen in *Figure 47*, with experimental details in 1.4.7.5. Mass spectrometry data of the addition reduced products can be seen in *Figure 46*.

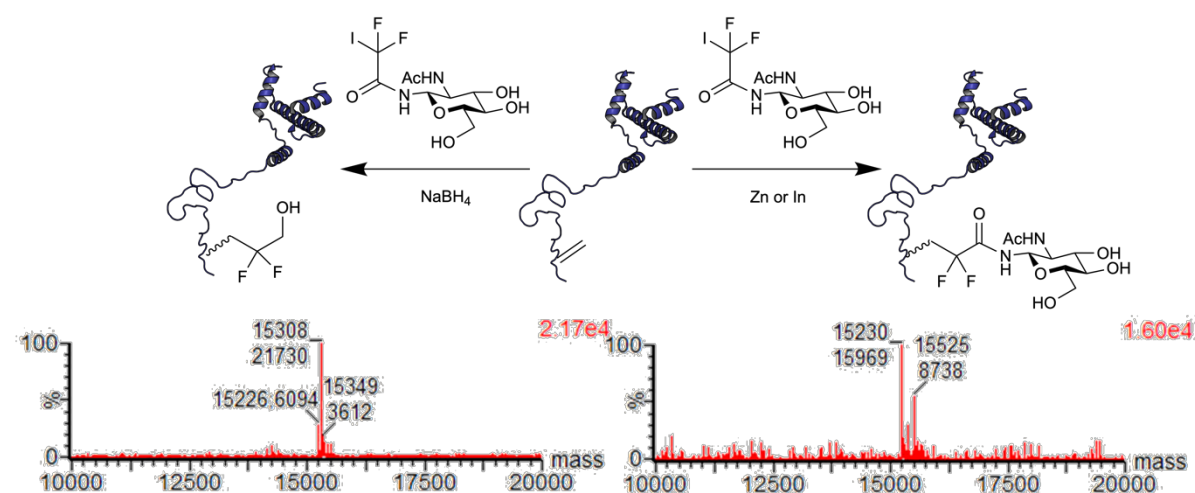


Figure 46 - Scheme of Histone H3 S10Dha treated with F_2N -GlcNAc reagent and an initiator: different outcomes are observed depending on what initiator is used. Sodium borohydride results in the addition of 2,2-difluoroethanol, while other initiators such as Zn or $InCl_3/NaBH_4$ results in the addition of the carbohydrate . Left (NaBH₄ initiated) Calculated Mass:15521 Observed Mass: 15308 (Addition of 2,2-difluoroethanol)

Right (Zn initiated) Calculated Mass:15521 Observed Mass: 15525 + 15230 (Unreacted starting material)

Equivs. Zn / Equivs. Sugar	897	1794	2333
5952	14	28	31
11904	17	24	33
17857	14	24	41

Equivs. In / Equivs. Sugar	179	538	897
1.66E+03	0.0	9.7	14.6
3.31E+03	0.0	0.0	0.0
4.97E+03	0.0	0.0	0.0

Figure 47 – Heat maps for the conversion Histone H3S10Dha to Histone H3 S10QF₂GlcNAc treated with the specified amount of radical precursor carbohydrate **20** (1-(2,2-difluoro-2-iodoacetamido)-β-D-GlcNAc) and the indicated amount of radical initiator (either Zn (top) or In (bottom)). Conversions are based on mass spectrometry and all numbers are the result of a single experiment.

1.2.5.0 AcrA

1.2.5.1 Introduction

Multidrug Efflux Pump Acridine resistance protein A (AcrA)¹⁸² forms part of the AcrA-AcrB-AcrZ-TolC drug efflux protein complex^{183, 184} which serves to move a broad variety of substrates outside of bacterial cells using allostercs driven by proton motive force as a source of energy.^{183, 184} AcrA is the middle protein of the channel and is located in the periplasm, forming both part of the pore and also functions as a linking protein holding AcrB and TolC together.¹⁸⁴ While Six individual units of AcrA are required to form the AcrA-AcrB-TolC complex, it is monomeric in solution.¹⁸⁵ Several views of AcrA's structure are shown in *Figure 48*.

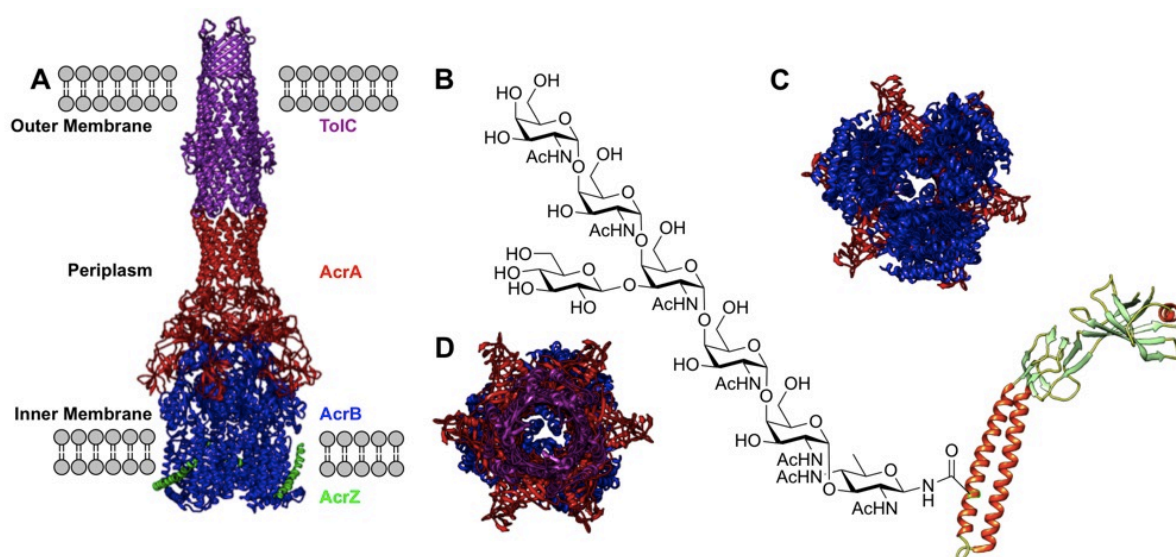


Figure 48 - The structure of the AcrA-AcrB-TolC multidrug efflux pump A) the complex spans the periplasm with AcrB (supported by AcrZ) sits spanning the inner membrane then is connected to 6 units of AcrA which sits in the periplasm and interacts with TolC which spans the outer membrane. B) AcrA found in *Campylobacter jejuni*¹⁸⁶ is glycosylated the structure of the glycan is shown C) AcrB up view of the complex D) TolC up view of the complex

Numerous bacteria express AcrA but *Campylobacter jejuni*¹⁸⁶ is unique in that its variant of AcrA is glycosylated, *C. jejuni* AcrA is glycosylated at sites N123 and N273.⁹³ The *N*-glycan found on AcrA is substantially different to those typically found on the glycoproteins of eukaryotes. The AcrA glycan consists of β -diNAcBac linkage to the protein followed by a pentameric chain of α -1,4-galactose linked the diNAcBac with an α -1,3 linkage. On the third galactose unit of the AcrA glycan there is glucose monomer linked to the galactose ring with by β -1,3 linkage. This is a typical glycan found in *C. jejuni*.^{187, 188}

AcrA forms an interesting target for glycosylation for two reasons, first it is a difficult protein to work with as it is prone to aggregation and precipitation therefore its successful glycosylation will demonstrate the ability of my method to be utilised on more challenging proteins. Secondly, by introducing unnatural glycans, I offer the potential to explore why this protein is glycosylated in only *C. jejuni*.

1.2.5.2 AcrA problems

I worked previously with AcrA and successfully glycosylated it using the same methodology as applied to H3 and other proteins.¹⁵⁴ Despite repeating the exact experimental conditions numerous times I found replicating the previously obtained results very challenging; the protein precipitated from solution very easily. Even if the protein did not precipitate, analysis of soluble aggregate by mass spectrometry is challenging. The following section of my thesis documents how I addressed the problems encountered.

1.2.5.3 Optimising AcrA Dha Formation

As the protein was precipitating, our first assumption was that the protein was aggregating, this suspicion was confirmed by inspection of circular dichroism (CD) spectrum, Ultra-violet (UV) spectrum and size exclusion chromatography (SEC) chromatogram.

	5	5.5	6	6.5	7	7.5	8	pH
0	49.33	48.90	48.48	48.08	47.70	47.15	44.80	
50	50.85	50.50	50.95	50.63	50.53	50.28	49.23	
100	50.95	51.17	51.37	52.10	51.60	51.50	51.03	
150	51.90	52.20	51.90	52.60	52.80	52.08	51.60	
200	51.28	52.00	52.70	53.00	53.05		51.95	
250		51.30	52.13		53.20	53.00	51.93	
[NaCl] / mM								

Figure 49 - Heatmap of melting temperatures (in °C) of AcrA N123C in a variety of buffers varying pH using citrate buffer and concentrations of sodium chloride, in blank squares the data obtained was unreliable. The numbers given are the mean of 4 experiments.

I suspected AcrA was incompatible with the buffer employed. To quantify buffer compatibility, I performed melting temperature analysis of the protein in a variety of buffers. I used thermal florescence to measure the melting temperature as this allows the analysis of large number of samples simultaneously in quadruplet. I varied the [NaCl] in 50 mM increments between 0 mM and 250 mM and the pH (using citrate buffer which maintains its pH reasonably well as temperature is varied) between 5 and 8 in 0.5 unit increment, a heat map of melting temperatures obtained can be seen in *figure 49*. I found that AcrA is more stable in high salt neutral pH buffers, with a ~10 °C of stability between the best and worst performing buffers. I initially decided to use PBS as a buffer for reaction: its success as a buffer for the reactions I wished to accomplish had previously been established¹⁵⁴ and the loss of 1 °C of stability seemed a small price to pay for the convenience of a buffer which can be bought as pellets.

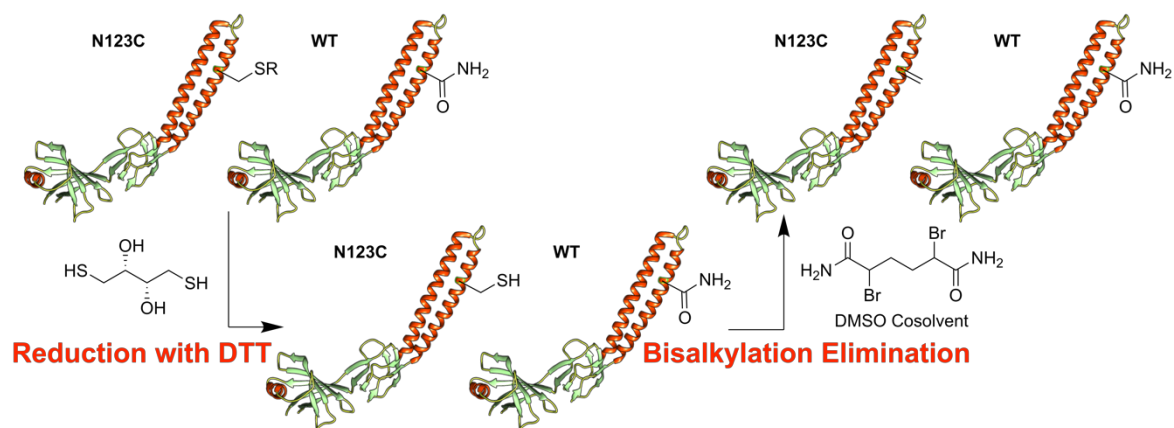


Figure 50 - Scheme for the first attempt at making AcrA N123Dha: simultaneously AcrA N123C and AcrA wildtype (WT) (as a control)) were reduced with DTT to cleave any disulfides which had formed followed by treatment with 2,5-dibromohexadiamide (DBHDA) in dimethylsulfoxide (DMSO) to give AcrA N123Dha. AcrA WT should remain unreacted.

With the melting temperature in mind I set out to produce AcrA Dha in a reliable and reproducible manner. First, I exposed AcrA to the ‘traditional’ conditions (shown in *Figure 50*, experimental details in section 1.4.7.12), that is DBHDA added as a DMSO solution then turned gently at room temperature. I applied these conditions on both the cysteine mutant and the wild type (wt) in order to differentiate between problems with the method I use and inherent instability of the Dha product. I also simultaneously exposed samples of the Cys mutant and the wt protein to the same conditions (temperature and agitation) without the reagents to differentiate between instability from the reaction conditions as opposed to the reagents.

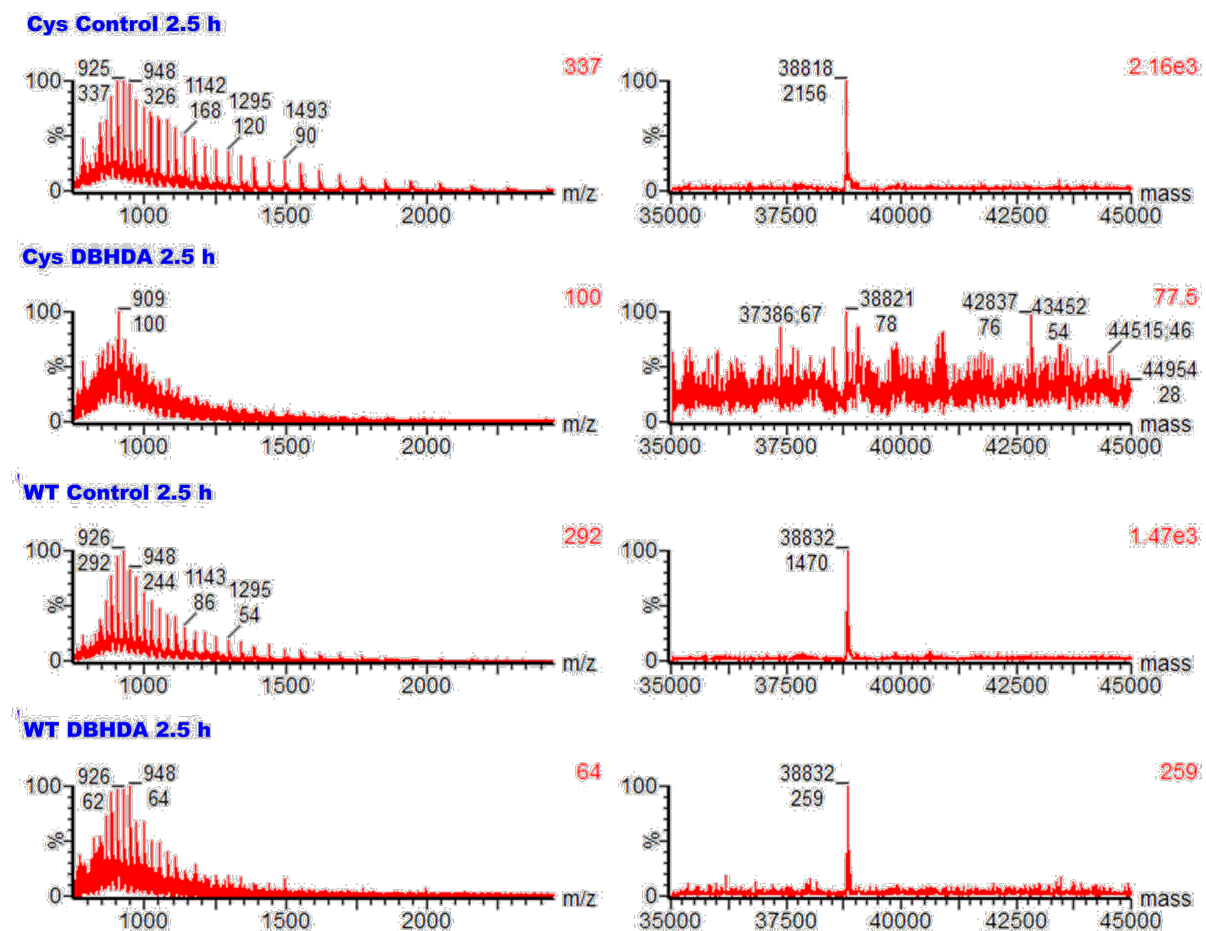


Figure 51 - Mass Spectra of the first attempts to make AcrA N123Dha using 2,5-dibromohexadiamide (DBHDA)/ dimethylsulfoxide (DMSO) after 2.5 h - Note how the ion series and spectra for samples treated with DBHDA are significantly 'messier' than the control samples. Calculated WT Mass: 38827, Calculated Cys Mass: 38816

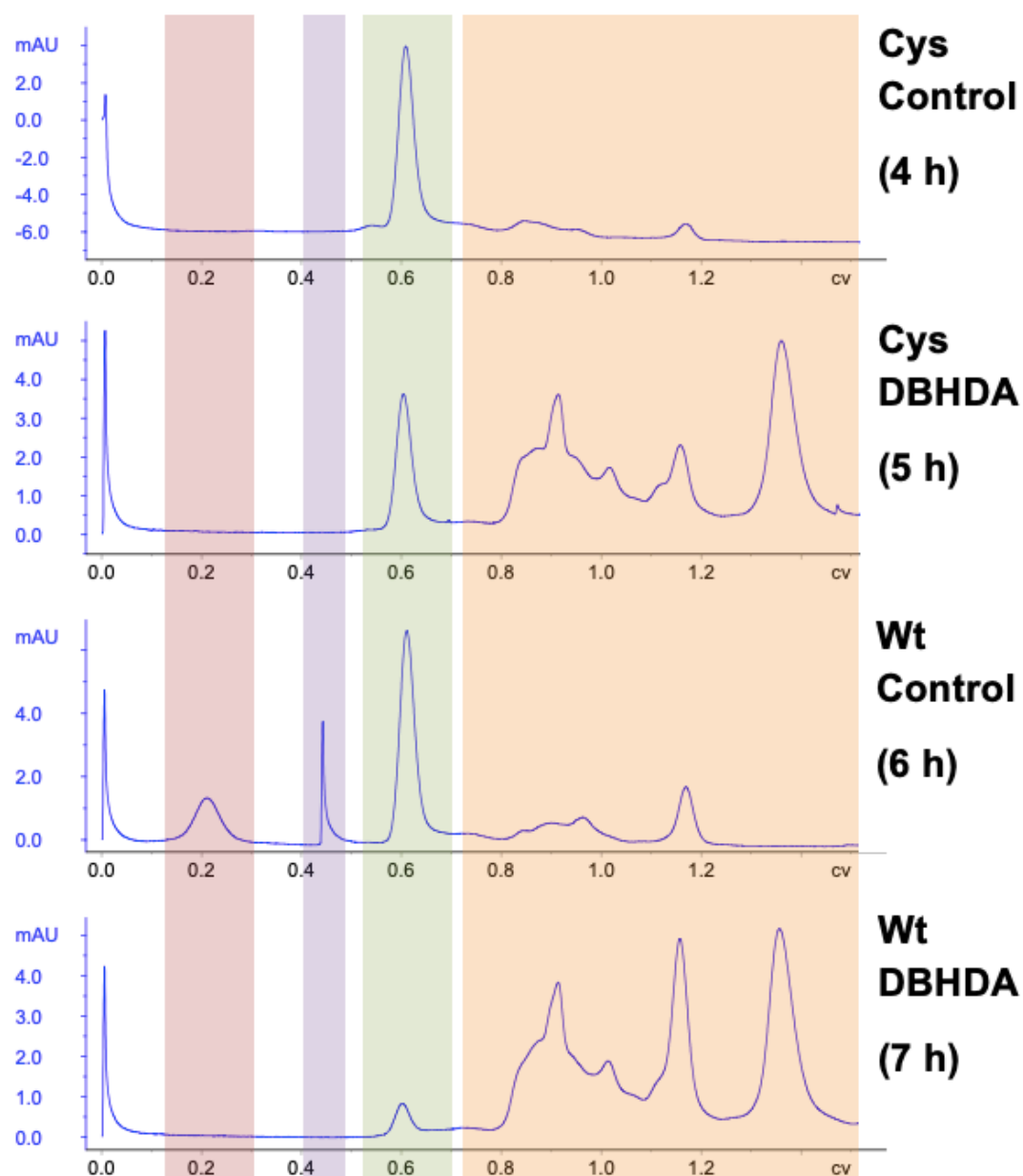


Figure 52 – Size exclusion chromatograms (280 nm, S200 10/300, in phosphate buffered saline (PBS)) of the first attempts to make AcrA N123Dha using DBHDA/DMSO at 4 to 7 h with interesting features highlighted: **Red** shows an aggregate peak, **Blue** shows the presence of an artifact caused by the erroneous removal of the buffer line, **Green** shows the peak from AcrA and **Orange** shows the presence of lower molecular weight species. Note how the intensity of the desired peak (**Green**) is lower in samples with DBHDA.

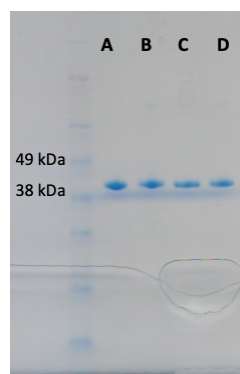


Figure 53 - Sodium dodecylsulfate poly(acrylamide) gel (SDS-PAGE) (4-12% Bis-Tris, MES running buffer, constant potential (200 V), InstantBlue™ development) of the first attempts to make AcrA N123Dha (~39 kDa) using DBHDA/DMSO after 8 h showing **A:** Cys control, **B:** wt control, **C:** Cys + DBHDA, **D:** Wt + DBHDA note the lack of lower molecular mass species. This demonstrates that while protein cannot be seen by mass spectrometry the protein is still present and not cleaved.

After 2.5 h the proteins treated with DBHDA were very difficult to observe by mass spectrometry, *figure 51*, with a significant loss of intensity of the ion series. Analytical size exclusion chromatograms (aSEC) were ran simultaneously however as each run takes 70 minutes there is a gap between each sample. In aSEC chromatograms (*figure 52*) I observe an obvious aggregate peak in the wt control and loss of intensity in the wt with reagents indicative of the formation of aggregates so large that they could no longer pass through the size exclusion resin employed (superdex s200). It was obvious an alternative method was required. Analysis by sodium dodecyl sulfate poly(acrylamide) gel electrophoresis (SDS-PAGE) however, showed no significant loss

of intensity when compared the analysis by MS and aSEC which is expected as the detergent and heating used in sample preparation disrupt any aggregates, *figure 53*.

Comparison of AcrA WT in buffer with 5 % MeCN

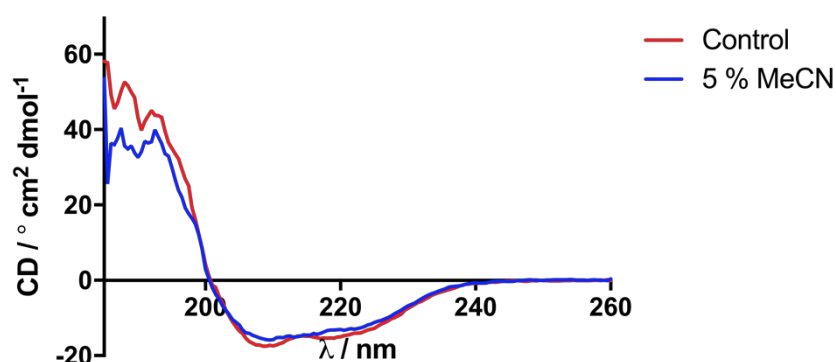


Figure 54 - Circular Dichroism Spectrum of AcrA N123C in fPBS and fPBS with 5 % MeCN, showing no significant differences. fPBS refers to buffer (NaF 150 mM Na₂HPO₄ 10 mM KH₂PO₄ 1.8 mM pH 7.4), it is essential to avoid Cl⁻ ions for CD measurements if you wish to acquire data below 200 nm. The curves are not normalised for concentration. Control sample concentration: 0.20 mg mL⁻¹, 5 % MeCN samples concentration: 0.19 mg mL⁻¹.

As spectra and chromatograms for DBHDA/DMSO treated proteins were significantly lower quality than those without DBHDA/DMSO I concluded that the co-solvent or the reagent were causing the problems. I explored changing both the co-solvent and the reagent. As the circular dichroism spectrum (CD) of AcrA showed little difference between the wt in fPBS (NaF 150 mM Na₂HPO₄ 10 mM KH₂PO₄ 1.8 mM pH 7.4) or in fPBS with 5 % MeCN (*figure 54*) I decided that this would be a suitable co-solvent. For dialkylation reagents I did not take such a methodical approach and by serendipity

methyl 2,5diboromopentanoate (MDPB) was chosen, *figure 55* shows the revised scheme. I did not agitate the reaction mixtures, as this could encourage aggregation.¹⁸⁹

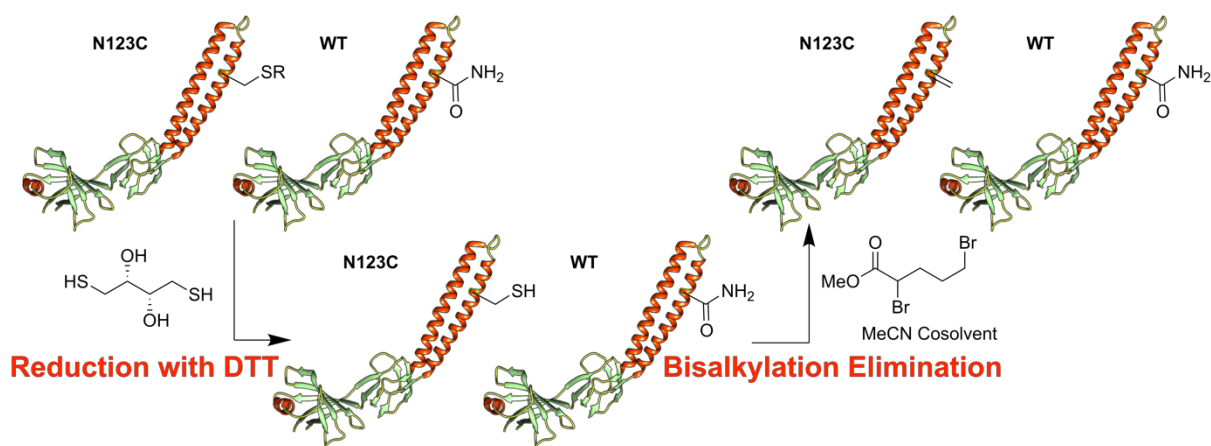


Figure 55 - Scheme for the second attempt at making AcrA N123Dha: simultaneous AcrA N123C and AcrA WT (wildtype) are reduced with DTT (dithiothreitol) to cleave any disulfides which had formed followed by treatment with methyl 2,5-diboromopentanoate (MDBP) in acetonitrile (MeCN) to give AcrA N123Dha. AcrA WT should remain unreacted.

In contrast to the DBHDA/DMSO method the MDPB/MeCN method exhibited no visible aggregate in the microcentrifuge tube as had been observed prior (snow type) after addition of the bisalkylation elimination reagents. While monitoring by mass spectrometry showed the formation of the dialkylated intermediate, no elimination was occurring even after 1 day of reaction. Positively there was no significant loss of intensity of ions observed in mass spectrometry, experimental details are given in section 1.4.7.14 and a graphical summary of results can be seen in *Figure 56*.

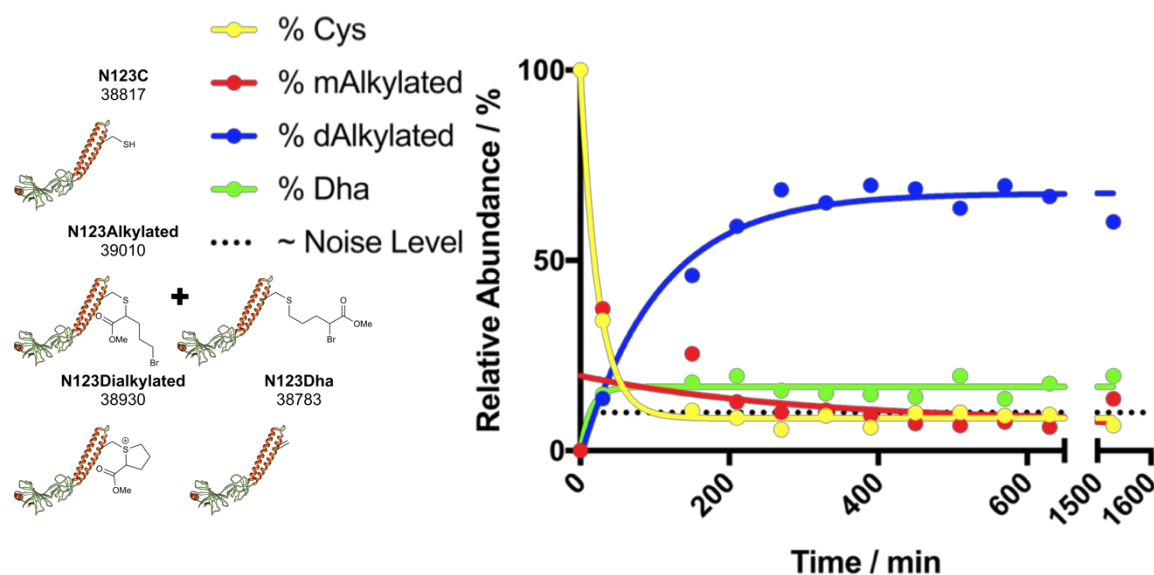


Figure 56 - Graph showing the species present in the bisalkylation elimination step of methyl 2,5diboromopentanoate (MDBP)/acetonitrile (MeCN) Dha formation on AcrA N123C at room temperature. The dialkylated species forms but then no elimination seems to be occurring so very little Dha is seen. The curves are fitted as exponentials. Curves are fitted as exponentials and are the result of a single experiment, the abundance of protein species is obtained by mass spectrometry.

Two options would encourage the elimination step of Dha formation: raising the temperature or raising the pH. Typically, in chemistry a raise in 10 °C will lead to a doubling of the rate of a chemical reaction, while a raise in pH would encourage the sulfonium ylide to form (see *figure 14* section 1.1.3.5). I chose to raise the temperature from room temperature (~20 °C) to 37 °C, experimental details are given in section 1.4.7.9 and a graphical summary results can be seen in *Figure 57*.

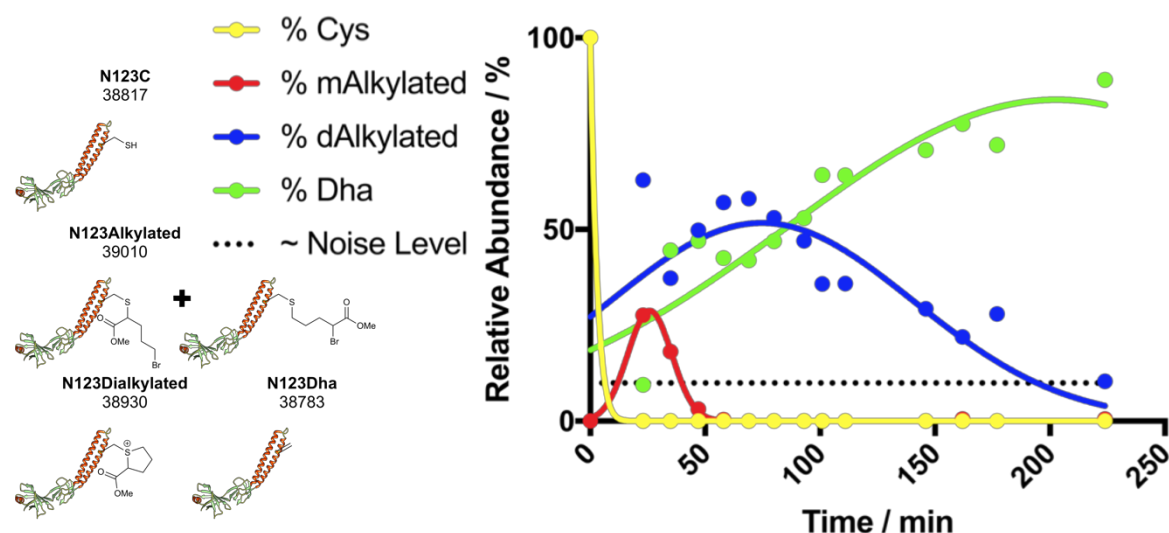
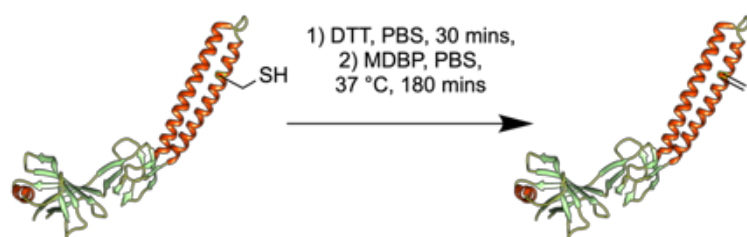
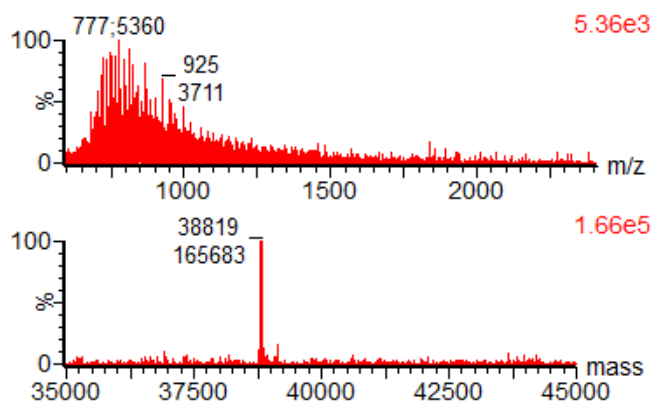


Figure 57 - Graph showing the species present in the bisalkylation elimination step of MDBP/MeCN Dha formation on AcrA N123C at 37 °C. Nearly complete conversion to Dha can be seen at approximately 220 minutes. The curves are fitted as Gaussian and are the result of a single experiment, the abundance of protein species is obtained by mass spectrometry.

After 3 h of reaction with MDBP and MeCN cosolvent at 37 °C I observed almost complete conversion from cysteine to Dha. Reassuringly, the mass spectrum, SEC chromatogram, SDS-PAGE and CD spectrum also looked as desired. Experimental details are given in Section 1.4.7.9, mass spectrometry data can be viewed in Figure 58 and an overview of analytical data can be seen in Figure 59.



Starting Material – AcrA N123C



Product – AcrA N123Dha

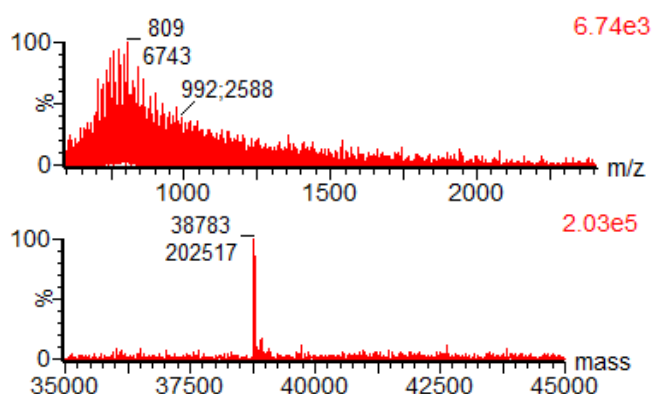


Figure 58 - Scheme for the chemical conversion of AcrA N123C to AcrA N123Dha. AcrA N123C was first treated with dithiothreitol (DTT) in, phosphate buffered saline (PBS). Next the protein was treated with methyl 2,5diboromopentanoate (MDBP) in the same buffer. Note the mass difference between the starting material and product mass spectra corresponding to the loss of SH_2 . Calculated Mass: 38783 Observed Mass: 38783

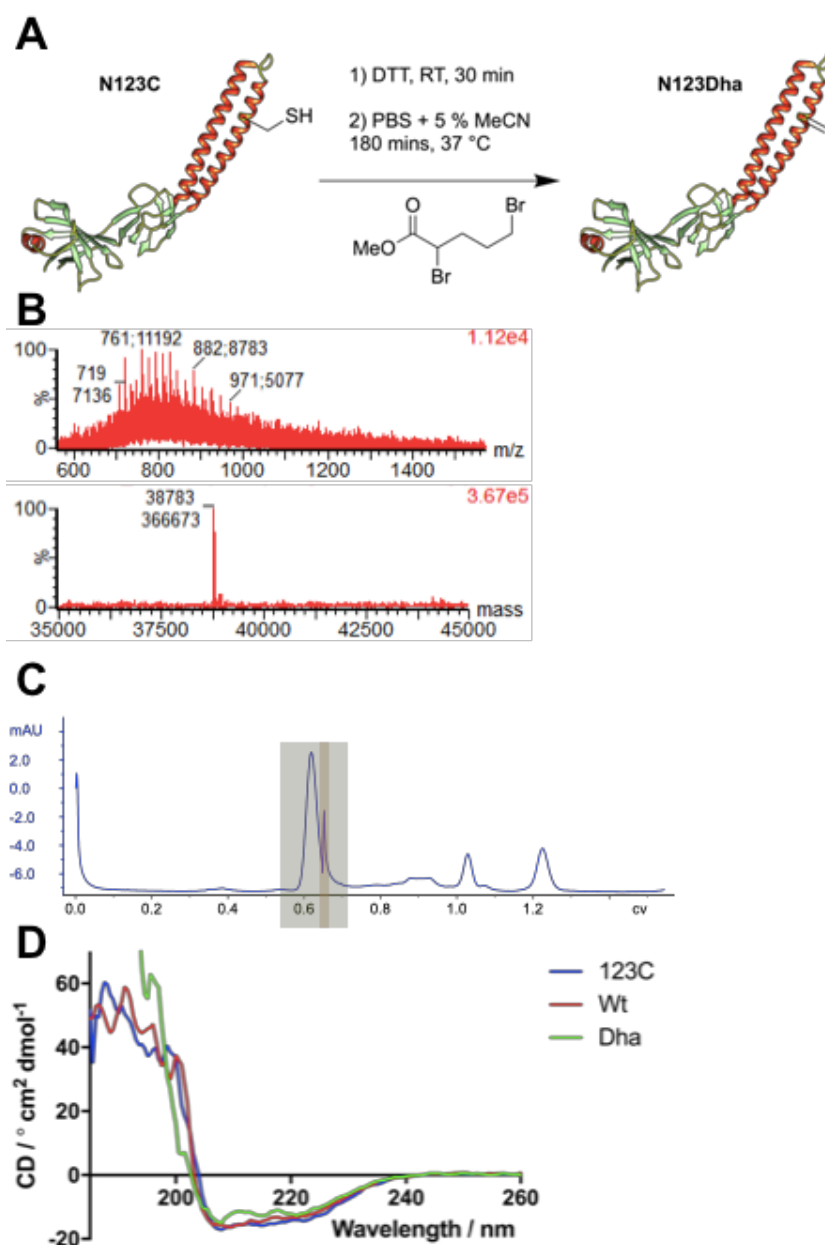


Figure 59 – Summary of Analytical data for AcrA N123Dha prepared by the methyl diboromopentanoate (MDBP)/MeCN method at 37 °C. **A** Scheme for the preparation of AcrA N123Dha. AcrA N123C was first treated with dithiothreitol (DTT) in, phosphate buffered saline (PBS). Next the protein was treated with methyl 2,5diboromopentanoate (MDBP) in the same buffer **B** Mass Spectrometry data of AcrA N123Dha (Calculated Mass: 38783 Observed Mass: 38783) **C** Analytical Size Exclusion Chromatogram of AcrA N123Dha (280 nm, S200 10/300, in phosphate buffered saline (PBS)) note that one main peak is observed (marked in **green**) which corresponds to AcrA N123Dha. A small aggregate peak can be seen slightly before at 0.4 CV. An artefact is present when the wrong line was accidentally removed from the instrument is marked in **red** **D** Circular Dichroism Spectrum of of AcrA N123Dha (ran in fPBS (NaF 150 mM Na₂HPO₄ 10 mM KH₂PO₄ 1.8 mM pH 7.4) at 0.200 mg mL⁻¹).

1.2.5.6 Optimising Radical Addition to AcrA Dha

Freshly prepared AcrA N123Dha was ported into the glove box and was exposed to the nitrogen atmosphere overnight at room temperature. In the morning before attempting reaction the mass spectrum and aSEC chromatogram were recorded, both can be seen in *Figure 60*. No protein peak could be observed in the mass spectrum. The SEC chromatogram indicated significant loss of protein and significant formation of aggregate. It is necessary to perform this step at room temperature as cooling was shown to significantly decrease the rate of gas exchange.¹⁵⁴

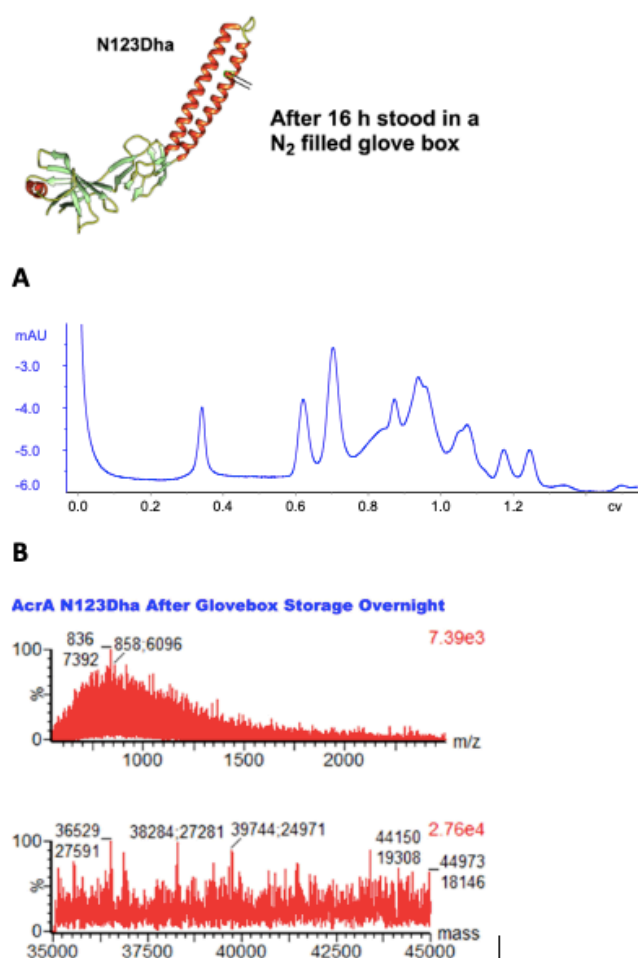


Figure 60 - Analytical data for AcrA N123Dha prepared by the methyl diboromopentanoate (MDBP)/MeCN method at 37 °C after being stored in the glove box overnight at room temperature. **A** Analytical Size Exclusion Chromatogram of AcrA N123Dha (280 nm, S200 10/300, in phosphate buffered saline (PBS)) after overnight standing in the glovebox, a large aggregate peak is present just before 0.4 CV, roughly the same intensity as the peak of AcrA N123Dha just after 0.6 CV, many other unexplained peaks are present **B** Mass Spectrometry data of AcrA N123Dha after standing in the glovebox overnight (Calculated Mass: 38783 Observed Mass: no clear signal).

Against previously established reaction conditions necessary to perform successful radical additions to Dha we attempted to perform the reaction open to air. Surprisingly

no protein cleavage was observed by either SDS-PAGE or mass spectrometry, however very little conversion to product was observed either, this can be seen in *figure 61* experimental data can be viewed in section 1.4.7.11. Earlier studies¹⁵⁴ conducted on the effect of Oxygen on radical addition to Dha were conducted with metal based initiators (Zn(Cu), Zn or In).

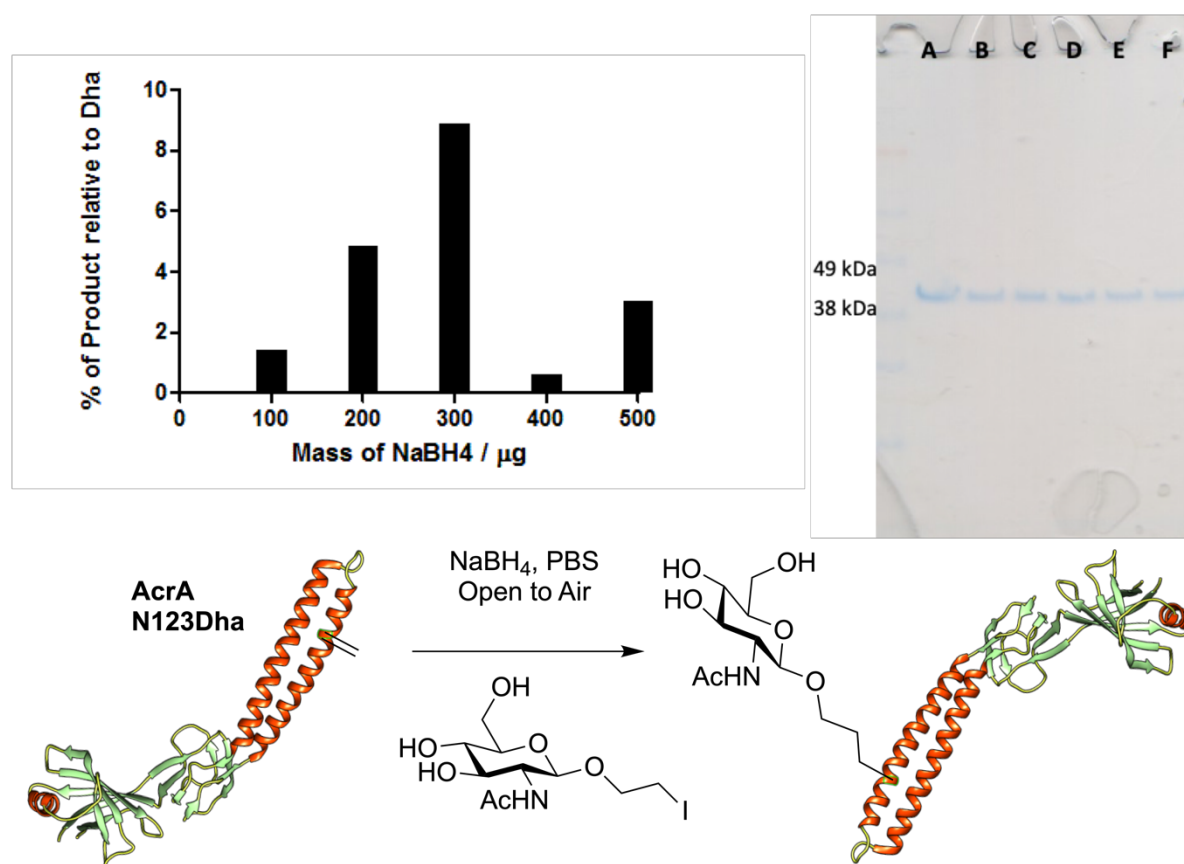


Figure 61 – Data obtained for the attempted reaction of AcrA N123Dha (~39 kDa) in phosphate buffered saline (PBS) on the lab bench open to air with varying amounts of sodium borohydride, all peaks are based on a experiment, conversions were determined by mass spectrometry. While little reaction is observed no cleavage product can be seen on the SDS-PAGE. (4-12% Bis-Tris, MES running buffer, constant potential (200 V), InstantBlueTM development) **A:** Starting Material, **B:** 100 μg NaBH₄, **C:** 200 μg NaBH₄, **D:** 300 μg NaBH₄, **E:** 400 μg NaBH₄, and **F:** 500 μg NaBH₄

It seemed logical to repeat the degassing time course experiment with NaBH_4 and freshly prepared AcrA N123Dha. AcrA N123Dha was ported into the glove box and after given periods of time reactions with 500 μg of NaBH_4 and 3 mg of the *O*-GlcNAc reagent were conducted. The results can be seen in *figure 62*

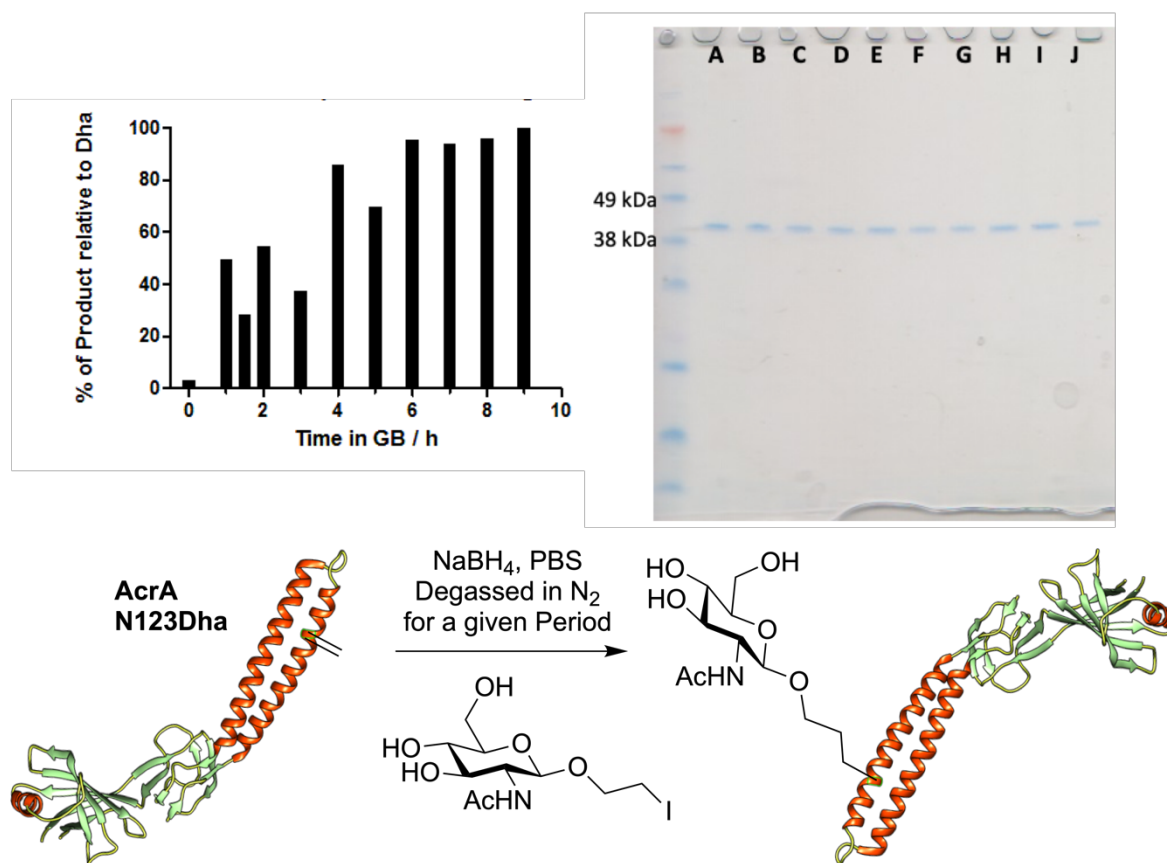
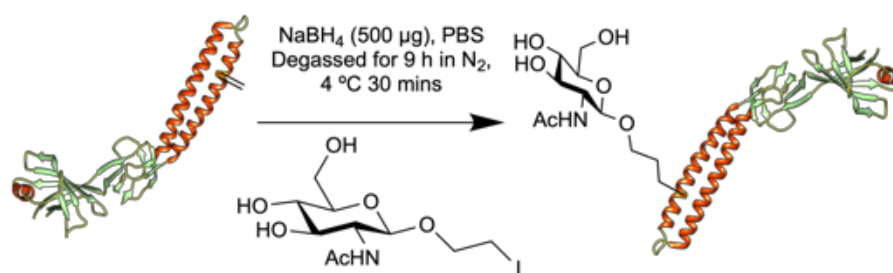
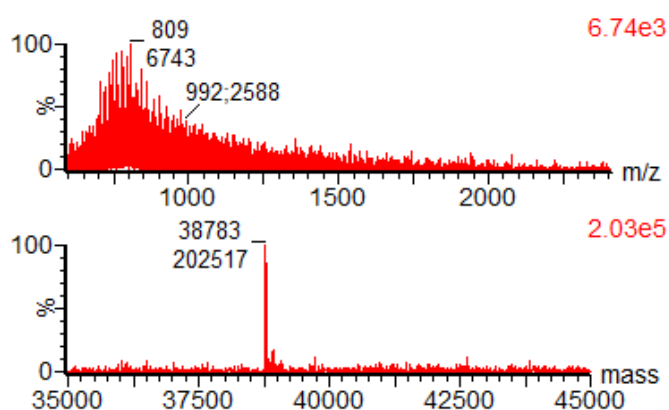


Figure 62 – Graph showing the yield of reactions with AcrA N123Dha (~39 kDa) with NaBH_4 (500 μg) and GlcNAc-1-6-OEtI (3 mg) after given amounts of time degassing all peaks are based on a single experiment, conversions were determined by mass spectrometry. SDS-PAGE (4-12% Bis-Tris, MES running buffer, constant potential (200 V), InstantBlueTM development) of the same samples. **A:** Starting Material, **B:** 1 h degassing, **C:** 2 h degassing, **D:** 3 h degassing, **E:** 4 h degassing, **F:** 5 h degassing, **G:** 6 h degassing, **H:** 7 h degassing, **I:** 8 h degassing, **J:** 9 h degassing.

No cleavage product could be seen in either the mass spectrum or by SDS-PAGE. Most notably after 6 h there is an almost complete conversion to the desired modification, the mass sepectrometry data can be seen in section 1.4.7.9 and *Figure 63*. Concerned that the glycosylated AcrA was more readily ionised by the mass spectrometer was unfounded, plotting the TIC as a function % glycosylated in the mass spectrum showed no correlation ($R^2 = 0.2521$), this is shown in *figure 64*. No correlation was observed indicating that that AcrA N123Dha and AcrA N123hhSGlcNAc are both equally as ionisable by mass spectrometry.



Starting Material – AcrA N123Dha



Product – AcrA N123hhSGlcNAc

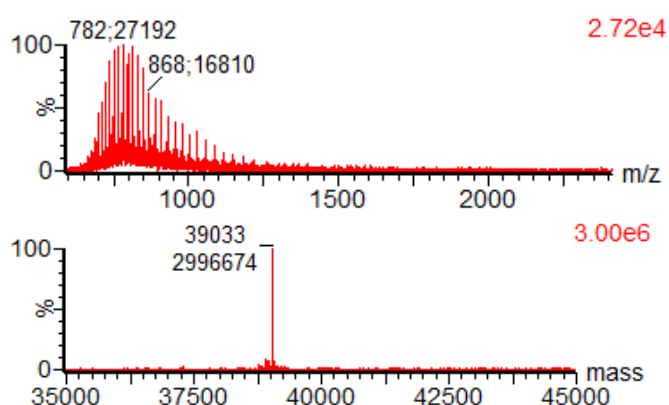


Figure 63 - Scheme for the chemical conversion of AcrA N123Dha to AcrA N123homohomo-Serine-GlcNAc. AcrA N123Dha in phosphate buffered saline (PBS) was sequentially treated with 1-iodoethyl-β-D-GlcNAc (**3**) and NaBH₄. Note the mass difference between the starting material and product mass spectra corresponding to the addition of C₁₀H₁₉NO₆. Calculated Mass: 39033
Observed Mass: 39033

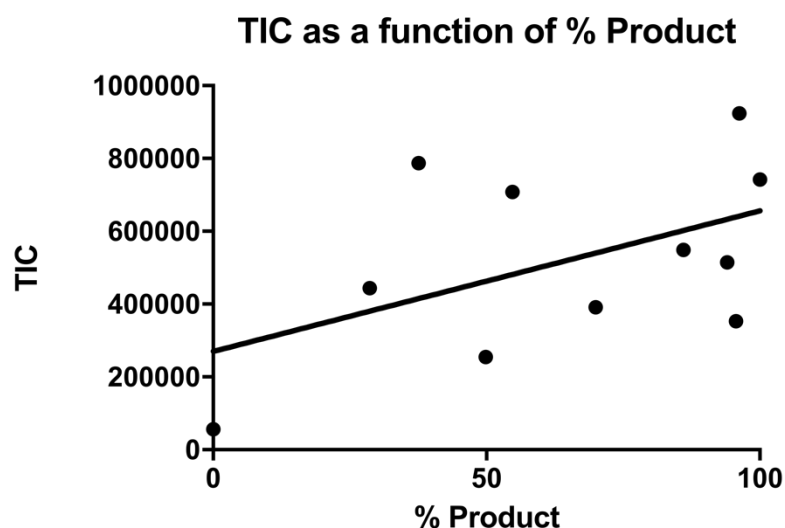


Figure 64 – Total Ion Count (TIC) as a function of the percentage of AcrA present in a mixed sample of AcrA N123Dha and AcrA N123OGlcNAc the line of best fit was determined by linear regression, $R^2 = 0.2521$ showing that there is no (linear) correlation between the TIC and % of glycosylated protein. This result indicates that both AcrA N123Dha and AcrA N123OGlcNAc ionise equally as well during mass spectrometry.

In order to demonstrate how robust our new procedure was for generating AcrA N123OGlcNAc I performed the reaction 10 times (after 9 h degassing), nine out of ten samples showed conversion to a near complete level of modification, while one showed little conversion, this is depicted graphically in *figure 65* experimental details are in section 1.4.7.10. A CD spectrum was obtained and demonstrated protein structure, shown in *figure 66*, it does appear that glycosylation has an effect on the structure of the polypeptide chain.

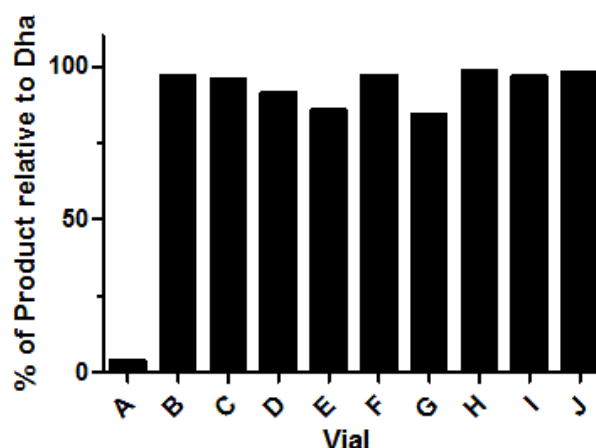


Figure 65 – Graph showing the yields of ten individual reactions of AcrA N123Dha with NaBH_4 (1.0 mg) and $\text{GlcNAc-1-}\beta\text{-OEtI}$ (3mg) after degassing in the glovebox for 9 h.

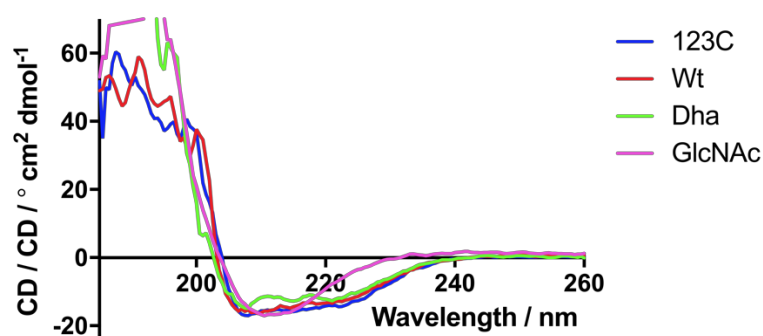


Figure 66 – Circular Dichroism Spectra of various AcrA specimens: **wild type** (wt), **N123C**, **N123Dha**, and **N123OGlcNAc** (ran in fPBS (NaF 150 mM Na_2HPO_4 10 mM KH_2PO_4 1.8 mM pH 7.4) at 0.200 mg mL^{-1}).

To complete my study of the radical addition to AcrA Dha I wanted to observe if a metal initiator would cause the protein to cleave as it had been noted with SBL and Histone H3. I placed freshly produced AcrA N123Dha in the glove box then performed a reaction every hour with 5 mg of *O*-GlcNAc reagent and 20 mg of fine mesh Zn powder. No protein was cleaved, and no obvious trends were observed in the product

mixtures of no reaction, single and double addition that were found, the results can be seen in *figure 67*, and experimental details are given in section 1.4.7.11 The reason why AcrA does not cleave remains a mystery.

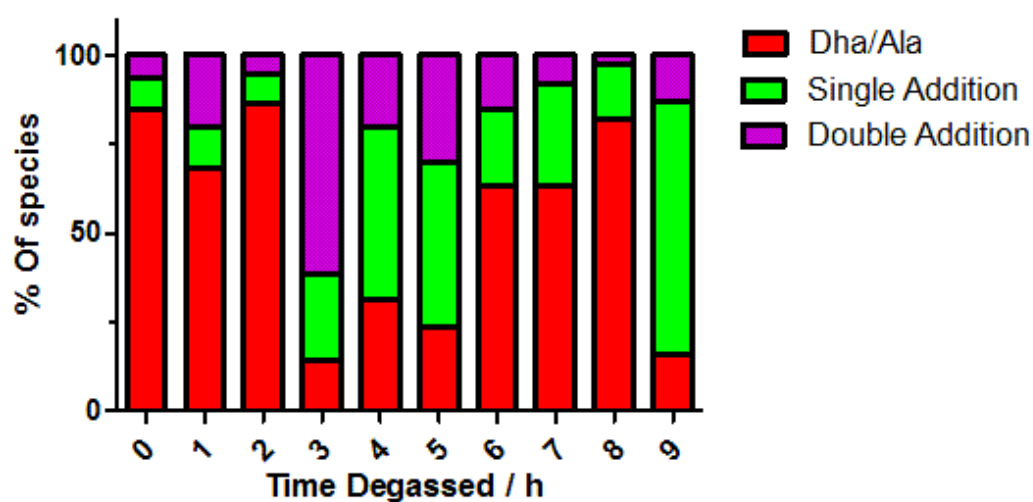
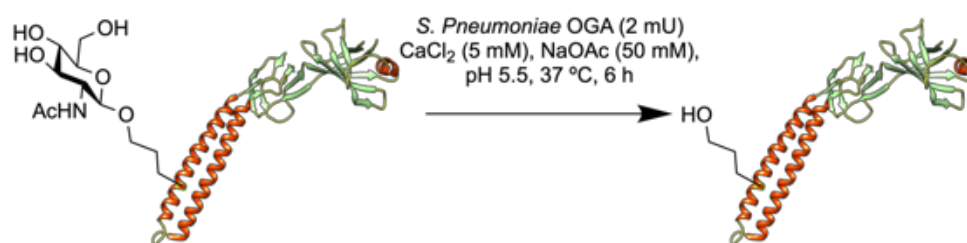


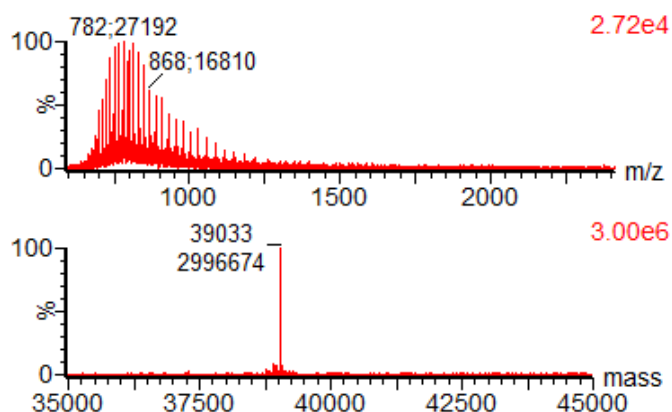
Figure 67 – Bar chart showing the % of protein species in samples of Zn (20 mg) initiated radical addition of the O-GlcNAc reagent (10 mg) to AcrA N123Dha after the stated amount of time degassing in the glovebox. No cleavage was observed. There are no obvious trends. Each is based on the result of a single experiment and conversions based on mass spectrometry.

1.2.5.5 AcrA Enzyme Reactions and Conclusion

To complete our AcrA study we wanted to demonstrate the reaction of the *hh*Ser-*O*-GlcNAc with biocatalysts. We chose to demonstrate the recognition of the *O*-GlcNAc on by OGAs. We demonstrated that AcrA could be deglycosylated by both *h*OGA (*Figure 69* and Section 1.4.10.2) and *Streptococcus pneumoniae* *O*-GlcNAcase (*sp*OGA) (*Figure 68* and Section 1.4.10.1).



Starting Material – AcrA N123hhSGlcNAc



Product – AcrA N123hhS

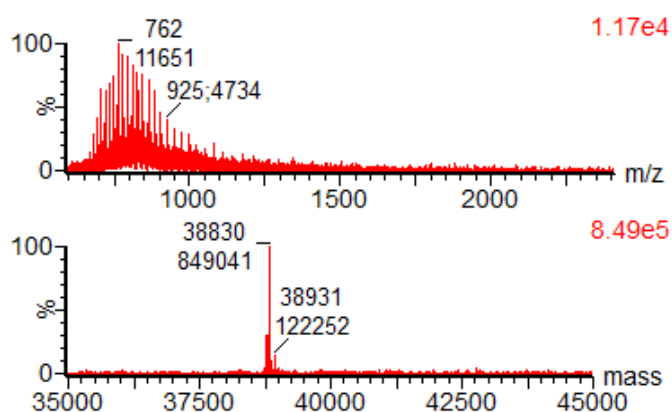
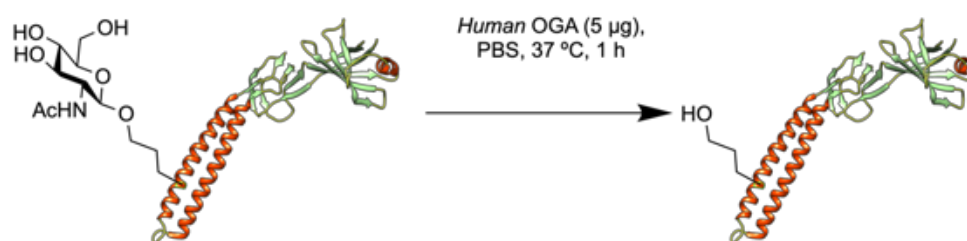
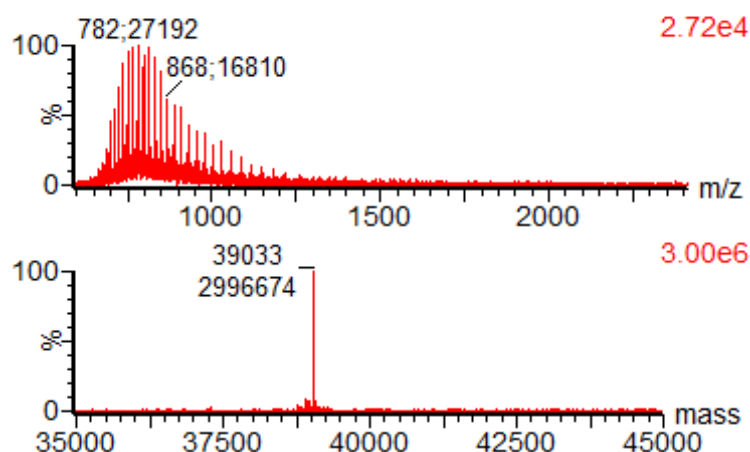


Figure 68 — Scheme for the enzymatic conversion of AcrA N123hhS-GlcNAc to AcrA N123hhS. AcrA N123 hhS-GlcNAc in sodium acetate buffer (50 mM pH 5.5) with CaCl_2 (5 mM) was treated with *Streptococcus pneumoniae* O-GlcNAcase then stood at 37 °C for 6 hours. Note the difference between the starting material and product mass spectra corresponding to the loss of $\text{C}_8\text{H}_{14}\text{NO}_5$. Calculated Mass: 38829, Observed Mass: 38830



Starting Material – AcrA N123hhSGlcNAc



Product – AcrA N123hhS

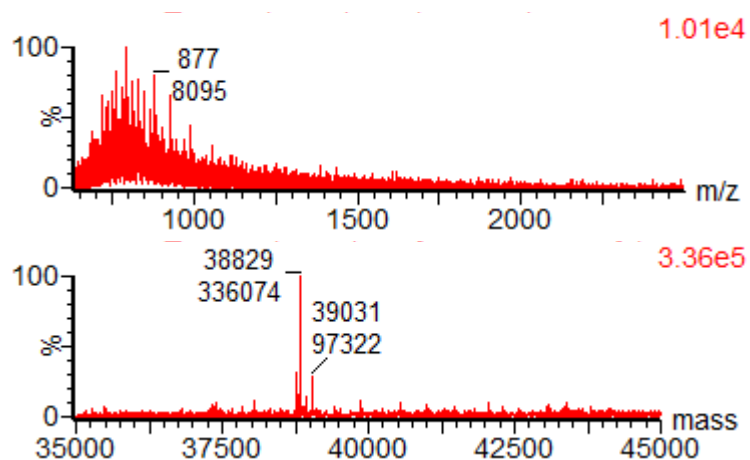


Figure 69 - Scheme for the enzymatic conversion of AcrA N123hhS-GlcNAc to AcrA N123hhS. AcrA N123hhS-GlcNAc in phosphate buffered saline (PBS) was treated with Human O-GlcNAcase (hOGA) then stood at 37 °C for 1 hour. Note the difference between the starting material and product mass spectra corresponding to the loss of $C_8H_{14}NO_5$. Calculated Mass: 38829, Observed Mass: 38829 and 39031 (unreacted starting material)

I think a useful continuation of this work would be to measure the melting temperatures of AcrA with *O*-GlcNAc modification, an *N*-GlcNAc and the naturally glycosylated. The comparison of melting points with the aglycone melting point would demonstrate if the glycosylation is to aid in protein stability, deeper insight into the structural disruption caused by glycosylation would also be useful.

1.3.0.0 Conclusion

1.3.1.0 Summary

A method has been developed to conjugate carbohydrates and proteins which forms a sp^3 - sp^3 bond in the process. We have used this method to incorporate both *N* and *O* glycans and demonstrated that glycans incorporated using our method can be processed by biocatalysts. We have shown how this reaction can be further developed to apply to more challenging protein substrates.

1.3.2.0 Perspective

As my method has great potential there is a lot which can be explored: Personally if I was to extend this I would attempt the addition of poly(saccharides): initially I would explore the addition of *N,N'*-diacetylchitobiose-1-(4-iodobutyramide) and should the addition be successful if it can be cleaved by PNGase F, as discussed in **section 1.2.3.2** and shown in *figure 36*. Eventually I would wish this to be expanded to the addition of the complete $\text{Man}_3\text{GlcNAc}_2$ core.

Unpublished¹⁹⁰ work by co-worker Charlie Fehl has demonstrated the addition of GlcNAc-*N*-AcI to protein dehydroalanine using a sample of the carbohydrate I had stored from my own studies. I suspect the photoredox method of initiation allows reaction due to lower rate of radical generation. The generation of radicals with NaBH₄ is a burst, with a high concentration of radicals being generated for a short period. With photoredox a lower concentration of radicals is generated over longer period (controlled by the power of the irradiation and catalyst loading). A lower concentration of radicals would potentially prevent side reactions and allow more unfavourable reactions to proceed. An overview of this reaction can be seen in *Figure 70*.

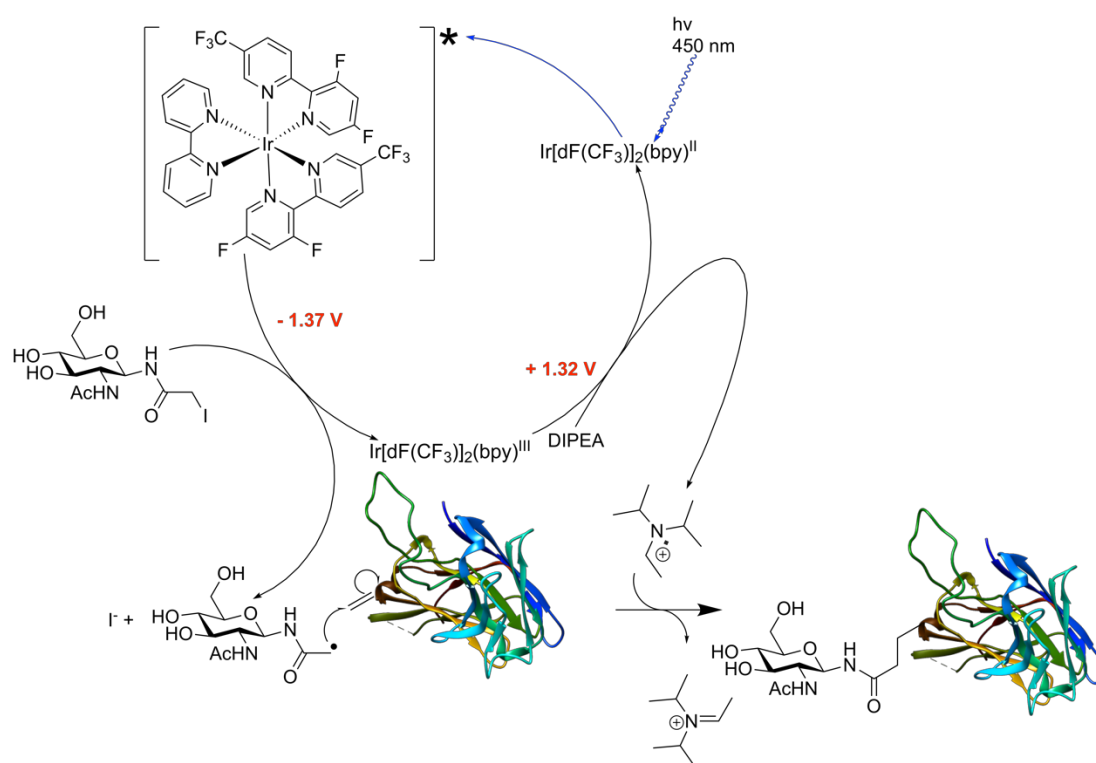


Figure 70 – Speculated Mechanism of photoredox initiation of radicals and their addition to proteins:¹⁹⁰ Irradiation of iridium complex $\text{Ir}[\text{df}(\text{CF}_3)]_2(\text{bpy})^{3+}$ with 450

nm (blue) light promotes an electron from the metal centre to the ligands giving an activated complex which is a strong oxidant and is able to oxidise GlcNAcNAcI to give I^- and the corresponding radical, along with $Ir[df(CF_3)]_2(bpy)^{2+}$. The oxidised iridium complex is now reducing and is able to reduce a molecule of DIPEA releasing which can lose a hydrogen atom through β -scission,¹⁹¹ to restart the cycle. The GlcNAcAc radical add the protein Dha and the hydrogen atom produced by the β -scission can quench the captodative protein radical.

While the kinetics of glycosidases have been extensively studied on model systems their scope has not been widely explored on protein systems, primarily due to the difficulty in obtaining homogenous starting material. To take the example of *O*-GlcNAcases, which have been extensively discussed in this thesis, most kinetic data has been obtained from their action on small molecule fluorogenic substrates, typically coumarin^{192, 193} or *para*-nitrophenol^{194, 195} derivatives.

While small molecule fluorogenic substrates are attractive to the researcher due to their commercial availability and simplicity of assay it is unlikely that that the kinetic parameters obtained translate well to protein substrates due to increased steric hindrance and different electronic properties of the leaving group. The methodology presented in this chapter could be employed to first create homogenous *O*-GlcNAcylated proteins then closely monitor the action of a glycosidases upon protein substrates by mass spectrometry. While *hh*SerGlcNAc would better imitate SerGlcNAc than small molecule fluorogenic substrates due to the closer steric and

electronic properties to the biological molecules, the longer sidechain length and stereochemistry of the amino acid should not be ignored.

The aggregation of α -synuclein is an event associated with Parkinson's disease and other neurodegenerative diseases, it has previously been demonstrated *O*-GlcNAcylation prevents this aggregation however studies have been limited to peptides⁷⁴ or hindered by the availability of starting material available only by solid phase synthesis.⁷³ The methodology presented in this chapter could allow a much reduced work load in the production of these homogenously *O*-GlcNAcylated aggregation prone proteins allowing greater exploration of their biology, including the 7 unexplored^{72, 76} *O*-GlcNAcylation sites on α -synuclein.

1.4.0.0 Experimental

1.4.1.0 General Chemical Experimental Considerations

Chemicals and solvents were purchased from Sigma-Aldrich UK, Acros UK, Alfa Aesar UK, Carbosynth UK, Fischer UK or Fluorochem UK and were used as delivered unless stated otherwise. Thin layer chromatography (TLC) was carried out using plastic 0.20 mm Polygram® Sil g/UV254 plates that were dried using a heat gun and visualised under UV (λ_{max} 254 nm or 366 nm) then by dips described in section 1.4.1.10. Melting points are uncorrected and are recorded using Leica hotstage microscope. Flash column chromatography was performed using Geduran® Si 60.8.2 or a Biotage Flash Purification System with a Kinesis Telos or Biotage Snap columns.

1.4.1.1 NMR Spectroscopy General

Deuterated solvents were used as the lock, and the residual protonated solvent as the internal reference peak. Spectra were analysed using MestReNova, Topspin or iNMR. Correlation spectroscopy (COSY), heteronuclear multiple bond correlation (HMBC), Nuclear Overhauser Effect spectroscopy (NOESY) and heteronuclear single quantum correlation (HSQC) were employed when required for signal assignment, other specific experiments *exempli gratia* 1D TOCSY will be mentioned specifically. Samples were prepared by dissolving ~5 mg of sample (^1H) or ~20 mg (^{13}C) in 500 μL of solvent, then filtered if required. CDCl_3 was stored over ~10 mass % molecular sieves, all other solvents were purchased in ampoules.

1.4.1.2 ^1H NMR Spectroscopy

Chemical shift information (δH) is given in parts per million (ppm) relative to the centre of the solvent signal (CHCl_3 7.26 ppm, DMSO-d_5 2.50 ppm, MeOH-d_3 3.31 ppm, Acetone-d_5 2.05 ppm or D_2O 4.79 ppm). Spectra were recorded on a Bruker AVIIIHD 400 nanobays (400.1 MHz) or AVII 500 (500.1 MHz). The number of protons for a resonance signal is given by $n\text{H}$ based on integral values. Multiplicity is denoted by the bracketed symbol. Coupling constants (J) are reported to the nearest 0.1 Hz and are determined using MestReNova, Topspin or iNMR.

1.4.1.3 ^{13}C NMR Spectroscopy

Chemical shift information (δC) is given in parts per million (ppm) relative to the centre of the solvent signal (CDCl_3 77.16 ppm, DMSO-d_6 39.52 ppm, MeOH-d_4 49.00 or Acetone-d_6 29.84 ppm). Spectra were recorded on a Bruker AVIIIHD 400 nanobays (100.6 MHz) or AVII 500 (125.7 MHz). Signals are assumed to be singlets unless otherwise stated. Distortionless enhancements by polarisation transfer with retention of quaternaries (DEPTQ) pulse sequence with broadband proton decoupling were employed for spectrum recording. In certain cases different distortionless enhancements by polarisation transfer (DEPT) flip angles were employed (135° and 90°) to aid assignment.

1.4.1.4 ^{19}F NMR Spectroscopy

Chemical shift information (δF) is given in parts per million (ppm) relative to the centre of the external standard CFCl_3 . Spectra were recorded on a Bruker AVIIIHD 400 nanobays (376.4 MHz). Signals are assumed to be singlets unless otherwise stated. Pulse sequences with broadband proton decoupling were employed for spectrum recording, unless stated.

1.4.1.5 IR Spectroscopy

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer using an attenuated total reflectance (ATR) attachment – oil or crystals of the compound were placed directly against the stage or a thin film deposited by dissolving the compound in chloroform then allowing the chloroform to evaporate. A background measurement was subtracted from each spectrum and a new background measurement was recorded before each spectrum was recorded.

1.4.1.6 Small Molecule Mass Spectrometry

Low-resolution mass spectra were recorded on an Agilent 6120 Quadrupole spectrometer using electrospray ionisation (ESI), positive and negative mode are described where appropriate. High-resolution mass spectra were recorded on a Bruker MicroTOF ESI mass spectrometer.

1.4.1.7 Optical Rotation

Optical rotation values were recorded on a Schmidt Haensch Unipol polarimeter. Samples were dissolved in spectrophotometric grade chloroform (purchased from Sigma-Aldrich UK) or D₂O (Purchased from Apollo Scientific). The polarimeter was blanked with the corresponding solvent before use.

1.4.1.8 UV-Vis Spectroscopy

UV-Vis spectroscopy was performed in a BMG Labtech Spectrostar Nano. Data was analysed in BMG Labtech Spectrostar Nano MARS. Grenier flat bottom 96 well plates were used for sample containment. Path length corrections were used. Spectra were recorded at 37 °C. para-Nitrophenol (pNP) absorbance values were recorded at 405 nm. For low wavelength readings a quartz 96 well plate was used

1.4.1.9 Glovebox

Anaerobic atmosphere was achieved using a Belle Technology glovebox equipped with 4 x BASF R3-11G catalysts. The oxygen level was measured below 10 ppm during all reactions.

1.4.1.10 TLC Dips

<i>Sulfuric Acid</i>	Sulfuric acid (97 %, 5 mL), methanol (95 mL)
<i>Anisaldehyde</i>	<i>p</i> -Anisaldehyde (0.5 mL), acetic acid (glacial, 50 mL), sulfuric acid (97 %, 1 mL)
<i>Potassium</i>	Potassium permanganate (3.00 g), potassium carbonate
<i>Permanganate</i>	(20.0 g), NaOH (5 % _(aq) , 5 mL), water (300 mL)
<i>Vanillin</i>	Vanillin (1.00 g), 100 mL H ₃ PO ₄ (50 % _(aq))
<i>Phosphomolybdic acid</i>	Phosphomolybdic acid (10.0 g), ethanol (100 mL)

1.4.1.11 Ion Exchange Resin Activation

DOWEX H⁺: DOWEX 50WX8 100-200 Mesh (Acros 100 g) was placed in a chromatography column and sequentially washed with MeOH (1 L), HCl_(aq) (2 M, 1 L) then distilled water (2 L or until neutral), nitrogen was passed through the resin until completely dry. The resin was removed from the column and dried in a CaCl₂ containing vacuum desiccator overnight.

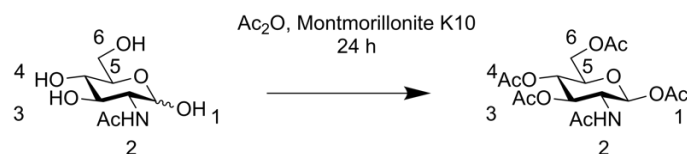
Activity is tested by placing a microspatula tip of resin in both 1 mL of water and 1 mL of brine. Testing with pH paper shows that the water sample is neutral and the brine sample acidic.

1.4.1.12 Molecular Sieves Activation

Molecular sieves were activated by heating to 350 °C furnace overnight then allowing to cool in a 180 °C oven for until required (at least 2 h). The molecular sieves were the placed in an oven dried flask 3-neck round bottom flask and allowed to come to room temperature under continuous flow of nitrogen. This method was employed on both the powder and pellet form of molecular sieves. Reagents were stored with pelleted molecular sieves, and reactions were carried out using the powdered molecular sieves.

1.4.2.0 Small Molecule Compound Characterisation

1.4.2.1 Peracetyl- β -D-GlcNAc 1



To an ice-cold stirred suspension of D-GlcNAc (6.42 g, 29.0 mmol) in Ac₂O (80 mL, 74.0 g, 725 mmol) montmorillonite K-10 (24.0 g) was added in portions over 10 mins. The ice-bath was removed, and the reaction mixture stirred at this temperature for 24 h. The reaction mixture was filtered through Celite and the pad washed with AcOEt until colourless. The combined filtrate was concentrated under reduced. The orange residue was recrystallised from MeOH twice to afford the title product as white needles (2.39 g, 6.11 mmol, 131 °C, 19.5 %).

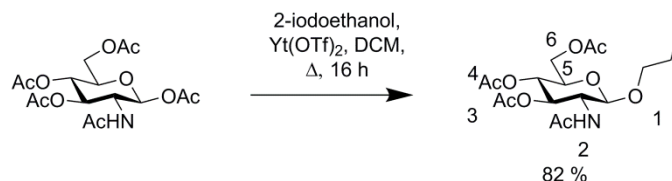
Method adapted from Amorelli¹⁹⁶

^1H NMR δH (400 MHz, Chloroform- d) 5.82 (d, $J = 9.6$ Hz, 1H, **H1**), 5.66 (d, $J = 8.8$ Hz, **NH**), 4.35-4.16 (m, 2H, **H3**, **H4**), 4.35-4.16 (m, 2H, **H6_a**, **H2**), 4.08 (dd, $J = 12.5$, 2.5 Hz, **H6_b**), 3.79 (ddd, $J = 3.79$ Hz, 1H, **H5**), 2.07 (s, 3H, **Ac**), 2.04 (s, 3H, **Ac**), 2.00 (s, 3H, **Ac**), 2.00 (s, 3H, **Ac**) and 1.89 (s, 3H, **Ac**)

^{13}C NMR δC (100.6 MHz, Chloroform- d) 171.2 (**Ac**), 170.7 (**Ac**), 170.2 (**Ac**), 169.6 (**Ac**), 169.3 (**Ac**), 92.5 (**C**), 72.8 (**C**), 72.6 (**C**), 67.8 (**C**), 61.7 (**C**), 52.9 (**C**), 23.2 (**Ac**), 20.9 (**Ac**), 20.7 (**Ac**), 20.7 (**Ac**), 20.7 (**Ac**) and 20.6 (**Ac**)

MS⁺ 413.0 [M + Na⁺]

MS⁻ 434.2 [M + FA]

1.4.2.2 Peracetyl 1-iodoethyl- β -D-GlcNAc 2

To a solution of peracetyl β -D-GlcNAc (500 mg, 1.28 mmol) and 2-iodoethanol (400 μL , 882 mg, 5.12 mmol) in dry DCM (15 mL) under argon, ytterbium (II) triflate (240 mg, 0.387 mmol) was added. The reaction mixture was heated to reflux overnight. At the point where analysis by TLC (100 % EtOAc, Sulfuric Acid development) showed the reaction complete (16 h) by complete consumption of the starting material ($R_f = 0.68$), the red reaction mixture was washed with Sat. Aq. NaHCO_3 (3 x 30 mL) then concentrated. Purification by column chromatography (40 % EtOAc in petroleum ether $R_f = 0.33$) afforded the title product as a colourless amorphous solid (524 mg, 1.04 mmol, 82 %).

Method adapted from Christensen¹⁵⁷

^1H NMR δH (400 MHz, Chloroform- d) 5.74 (d, $J = 8.7$ Hz, 1H, **H1**), 5.31 – 5.24 (m, 1H, **H3**), 5.02 (d, $J = 9.6$ Hz, 1H, NH), 4.73 (d, $J = 8.4$ Hz, 1H, **H4**), 4.21 (dd, $J = 12.3, 4.9$ Hz, 1H, **H5**), 4.12 – 4.02 (m, 2H, **H6_{a,b}**), 3.87 – 3.65 (m, 2H, $\text{OCH}_2\text{CH}_2\text{I}$), 3.29 – 3.17 (m, 2H, $\text{OCH}_2\text{CH}_2\text{I}$), 2.05 (s, 3H, **NAc**), 1.99 (s, 3H **OAc**), 1.98 (s, 3H, **OAc**) and 1.94 (s, 3H, **OAc**).

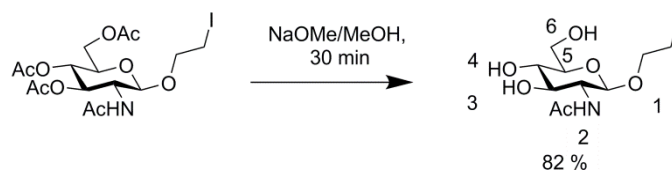
^{13}C NMR δC (100.6 MHz, Chloroform- d) 170.84 (**NAc**), 170.71 (**OAc**), 170.48 (**OAc**), 169.41 (**OAc**), 100.67 ($\text{OCH}_2\text{CH}_2\text{I}$), 72.12 (**C1**), 71.93 (**C2**), 70.19 (**C3**), 68.57 (**C4**), 62.05 (**C6**), 54.70 (**C5**), 23.56 (**NAc**), 20.79 (**OAc**), 20.72 (**OAc**), 20.65 (**OAc**) and 2.95 (**CH₂I**).

MS⁺ 347.1 [$\text{M} + \text{H}^+$], 715.3 [$2\text{M} + \text{Na}^+$]

MS⁻ 381.1 [$\text{M} + \text{Cl}$] and 391.0 [$\text{M} - \text{H}^+ + \text{FA}$]

$[\alpha]^{20}$ - 0.0910

IR ($\nu_{\text{max}}/\text{cm}^{-1}$) 3266, 3097, 1743, 1650, 1566, 1431, 1367, 1223, 1166, 1122, 1078, 1031, 979.0, 946.9, 903.6, 738.0, 695.7 and 631.3

1.4.2.3 1-iodoethyl- β -D-GlcNAc 3

To a solution of peracetyl 1-iodoethyl- β -D-GlcNAc (250 mg, 0.499 mmol) in dry methanol (5 mL), sodium methoxide in methanol (25 %, 100 μ L) was added and stirred until analysis by TLC (100 % EtOAc, sulfuric acid development) had shown the reaction complete by disappearance of the starting material and the appearance of one spot (R_f 0.0). Upon completion (30 min) the reaction mixture was neutralised by addition of DOWEX H⁺ and stirred for 5 minutes, the reaction mixture was filtered and then concentrated to 1 mL, which was washed through a silica plug (which had been thoroughly washed with methanol, water/isopropanol/ethyl acetate 1:2:5), volatiles were removed under reduced pressure to afford the title product as a white amorphous solid (153 mg, 0.409 mmol, 156.8 $^{\circ}$ C (decomp.) 82 %).

Method adapted from Zemplén¹⁵⁸

^1H NMR δH (400 MHz, D_2O) 4.53 (d, $J = 8.4$ Hz, 1H, **H1**), 4.10 (ddd, $J = 11.9$, 5.9, 4.8 Hz, 1H, **H6_b**), 3.90 (dd, $J = 12.4$, 1.6 Hz, 1H, **H6_a**), 3.82 (ddd, $J = 12.1$, 7.7, 4.6 Hz, 1H, **H5**), 3.76 – 3.64 (m, 2H, **H3**, **H4**), 3.58 – 3.47 (m, 1H, **H2**), 3.46 – 3.39 (m, 2H **CH₂I**), 3.39 – 3.19 (m, 2H, **CH₂CH₂I**) and 2.06 (s, **3H**).

^{13}C NMR δC (100.6 MHz, D_2O) 177.24 (**NAc**), 103.45 (**OCH₂CH₂I**), 84.20 (**C1**) 78.40 (**C2**), 76.18 (**C3**), 72.34 (**C4**), 63.17 (**C6**), 57.92 (**C5**), 25.20 (**NAc**) and 5.94 (**CH₂I**).

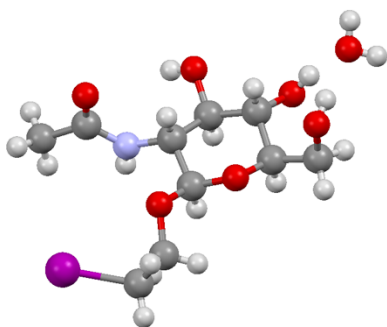
MS⁺ 347.1 [**M** + H^+], 715.3 [**2M** + Na^+]

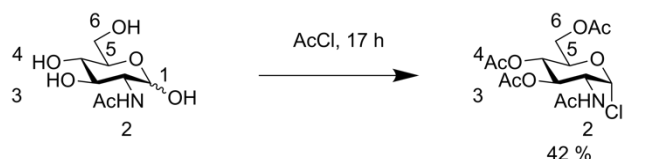
MS⁻ 381.1 [**M** + **Cl**] and 391.0 [**M** - H^+ + **FA**]

$[\alpha]^{20}$ -0.1995

IR ($\nu_{\text{max}}/\text{cm}^{-1}$) 3256 (br), 2849, 1650, 1553, 1422, 1376, 1326, 1281, 1213, 1143, 1112, 1084, 1064, 1035, 961.8, 912.7, 891.4, 750.9, 706.2, 671.0 and 634.0.

XRD A single crystal structure was obtained. Crystals were grown by dissolving 20 mg of sugar in 0.500 mL of hot $\text{D}_4\text{-MeOH}$ then leaving in the dark for a sustained period.



1.4.2.4 Peracetyl 2-chloro- α -D-GlcNAc 5

A suspension of *N*-acetyl-D-glucosamine (25.0 g, 113 mmol) in acetyl chloride (50.0 mL, 55.0 g, 701 mmol) was sealed with a suba seal and balloon then stirred for 17 h at which point analysis by TLC (100 % EtOAc, H₂SO₄ development) indicated complete disappearance of starting material (R_f 0.0) and the appearance of one major product (R_f 0.70) and one side product (R_f 0.47). The scarlet solution was diluted with DCM (500 mL), washed with saturated aqueous NaHCO₃ (3 x 500 mL), dried over MgSO₄, then concentrated under reduced pressure (to 50.0 mL), the crude produce was precipitated with sodium dried Et₂O (1.00 L) to afford beige crystals. Purification by flash column chromatography (Pet Ether/EtOAc 40 % -> 65 % gradient elution) afforded the title product as a white amorphous solid (17.53 g, 47.9 mmol, 42 %).

Method adapted from Ferrier¹⁹⁷

^1H NMR δH (400 MHz, Chloroform- d) 6.15 (d, $J = 3.7$ Hz, 1H, **H1**), 5.82 (d, $J = 8.7$ Hz, 1H, **NH**), 5.29 (dd, $J = 10.7, 9.4$ Hz, 1H, **H3**), 5.18 (t, $J = 9.8$ Hz, 1H, **H4**), 4.50 (ddd, $J = 10.7, 8.8, 3.8$ Hz, 1H, **H2**), 4.28 – 4.19 (m, 2H, **H6_{a+b}**), 4.15 – 4.01 (m, 1H, **H5**), 2.07 (s, 3H, **Ac**), 2.02 (s, 6H, **Ac**, **Ac**) and 1.95 (s, 3H, **Ac**).

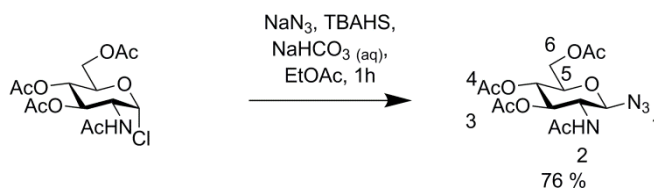
^{13}C NMR δC (100.6 MHz, Chloroform- d) 171.48 (**NAc**), 170.60 (**OAc**), 170.14 (**OAc**), 169.15 (**OAc**), 93.64 (**C1**), 70.88 (**C3**), 70.12 (**C2**), 66.94 (**C4**), 61.14 (**C6**), 53.47 (**C5**), 23.09 (**NAc**), 20.70 (**OAc** **OAc**) and 20.57 (**Ac**).

MS⁺ 366.1 [$\text{M} + \text{H}^+$], 384.1 [$\text{M} + \text{NH}_4^+$] and 388.1 [$\text{M} + \text{Na}^+$]

MS⁻ 364.0 [$\text{M} - \text{H}^+$]

$[\alpha]^{20}$ +129

IR ($\nu_{\text{max}}/\text{cm}^{-1}$) 3239, 2981, 1739, 1641, 1540, 1435, 1375, 1290, 1269, 1250, 1226, 1208, 1167, 1146, 1116, 1094, 1056, 1033, 984.6, 912.0, 893.0, 874.5, 821.0, 760.9, 728.5, 643.0 and 606.6

1.4.2.5 Peracetyl 1-azido- β -D-GlcNAc 6

To a rapidly stirred solution of Peracetyl 2-chloro- α -D-GlcNAc (500 mg, 1.37 mmol) and tetra(ⁿbutyl) ammonium hydrogen sulfate (464 mg, 1.37 mmol) in EtOAc (5 mL) and saturated aqueous NaHCO_3 (5 mL), sodium azide (267 mg, 4.10 mmol) was added in portions. The reaction mixture was stirred for 1 h at which point analysis by TLC (100 % EtOAc, H_2SO_4 Development) indicated complete disappearance of the starting material (R_f 0.70) and appearance of a single product (R_f 0.52). The organic fraction was washed with saturated aqueous NaHCO_3 (3 x 10 mL) and saturated NH_4Cl (10 mL) then volatiles removed under reduced pressure, the white amorphous solid was purified by flash column chromatography (Pet Ether/EtOAc 50 % -> 80 % gradient elution) to give the title product as a white amorphous solid (387 mg, 1.04 mmol, 172.0 °C (decomp.), 76 %).

Previously synthesised by Flitsch,¹⁹⁸ amongst many others¹⁹⁹⁻²⁰³

^1H NMR δH (400 MHz, Chloroform- d) 5.64 (d, $J = 8.9$ Hz, 1H, **NH**), 5.21 (dd, $J = 10.6, 9.3$ Hz, 1H, **H3**), 5.07 (t, $J = 9.7$ Hz, 1H, **H4**), 4.72 (d, $J = 9.2$ Hz, 1H, **H1**), 4.24 (dd, $J = 12.5, 4.8$ Hz, 1H, **H6b**), 4.13 (dd, $J = 12.4, 2.3$ Hz, 1H, **H6a**), 3.88 (dt, $J = 10.5, 9.0$ Hz, 1H, **H2**), 3.76 (ddd, $J = 10.0, 4.8, 2.3$ Hz, 1H, **H5**), 2.07 (s, 3H, **Ac**), 2.01 (s, 6H, **Ac**, **Ac**) and 1.95 (s, 3H, **Ac**).

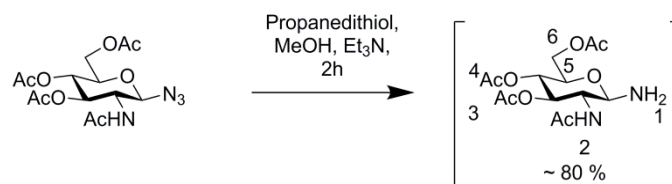
^{13}C NMR δC (100.6 MHz, CDCl_3) 171.1 (**NAc**), 170.7 (**OAc**), 170.4 (**OAc**), 169.3 (**OAc**), 88.45 (**C1**), 74.0 (**C2**), 72.1 (**C3**), 68.0 (**C4**), 61.8 (**C6**), 54.2 (**C5**), 23.3 (**NAc**), 20.8 (**OAc**), 20.7 (**OAc**) and 20.6 (**OAc**).

MS^+ 209.1 [$\text{M} + 2 \text{Na}^+$], 330.1 [$\text{M} - \text{N}_3^-$], 373.1 [$\text{M} + \text{H}^+$] and 767.2 [$2 \text{M} + \text{Na}^+$]

MS^- 371.1 [$\text{M} - \text{H}^+$], 417.1 [$\text{M} - \text{H}^+ + \text{FA}$] and 484.9 [$\text{M} - \text{H}^+ + \text{TFA}$]

$[\alpha]^{20}$ - 45.9

IR ($\nu_{\text{max}}/\text{cm}^{-1}$) 3339, 2982, 2104, 1747, 1660, 1532, 1372, 1320, 1300, 1230, 1109, 1091, 1037, 980.3, 956.8, 897.2, 689.4, 647.1 and 622.4

1.4.2.6 Peracetyl 1-amino- β -D-GlcNAc 7

1,3-Propane Dithiol method: To a rapidly stirred solution of peracetyl 1-azido- β -D-GlcNAc (1.00 g, 2.69 mmol) in anhydrous methanol (16 mL) under Ar, NEt₃ (0.9 mL, 0.653g, 6.45 mmol) and 1,3-propanedithiol (0.6 mL, 0.648 g, 6.00 mmol) were sequentially added. The effervescent reaction mixture was stirred at RT for 2 h, at which point a large amount of white precipitate could be seen suspended in the colourless to pale yellow solution, TLC analysis (10 % MeOH/CHCl₃ Anisaldehyde development) showed complete consumption of the starting material (R_f 0.72) and the formation of a major spot (R_f 0.41). The methanol was removed under reduced pressure and the residue dissolved in chloroform, this was loaded onto a short silica plug and washed with a large amount of chloroform then eluted with MeOH/CHCl₃ (10 %), solvents were removed, then the glassy solid stored under high vacuum overnight. The title product was obtained as a colourless glassy solid (0.744 g, 2.15 mmol, 80 %).

Method adapted from Knowles²⁰⁴

Adam's Catalyst Method: To a rapidly stirred argon degassed solution of peracetyl 1-azido- β -D-GlcNAc (250 mg, 0.671 mmol) and suspended Pd/C (10 %, 30.0 mg) hydrogen was bubbled through for 30 min, at which point analysis by TLC (DCM/MeOH 10 % sulfuric acid development) showed complete consumption of the starting material (R_f 0.42) and the appearance of a single product (R_f 0.34). The reaction mixture was degassed with argon, then washed through a celite pad (DCM/MeOH 20 %), the volatiles were removed under reduced pressure and the resultant off white crystals in oil purified by column chromatography (DCM/MeOH 0% -> 5% gradient elution) to afford the title product as a brown amorphous solid (134 mg, 0.387 mmol, 58 %).

Method adapted from Shibata²⁰⁵

^1H NMR δH (400 MHz, Chloroform- d) 5.65 (d, $J = 9.1$ Hz, 1H, **NH**), 5.19 – 4.85 (m, 2H, **H3** and **H4**), 4.18 (dd, $J = 12.3, 4.9$ Hz, 1H, **H6_b**), 4.08 (d, $J = 9.18$ Hz, 1H, **H1**), 4.07 (dd, $J = 12.31, 2.14$, 1H, **H6_a**) 4.02 – 3.93 (m, 2H, **H2** and **H5**), 2.06 (s, 3H, **Ac**), 2.01 (s, 3H, **Ac**), 2.00 (s, 3H, **Ac**) and 1.94 (s, 3H, **Ac**).

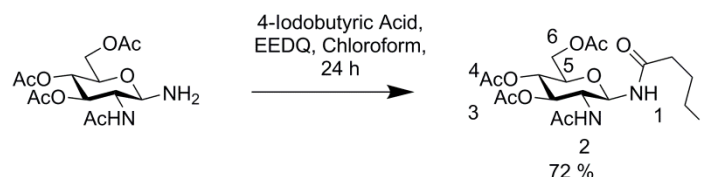
^{13}C NMR δC (100.6 MHz, Chloroform- d) 171.7 (**NAc**), 170.8 (**OAc**), 170.8 (**OAc**), 169.4 (**OAc**), 86.5 (**C1**), 73.4 (**C2**), 72.8 (**C3**), 68.4 (**C4**), 62.4 (**C6**), 54.9 (**C5**), 23.4 (**NAc**), 20.8 (**OAc**), 20.7 (**OAc**) and 20.6 (**OAc**).

MS⁺ 347.1 [**M** + **H⁺**], 715.3 [**2M** + **Na⁺**]

MS⁻ 381.1 [**M** + **Cl**] and 391.0 [**M** - **H⁺** + **FA**]

$[\alpha]^{20}$ - 0.0920

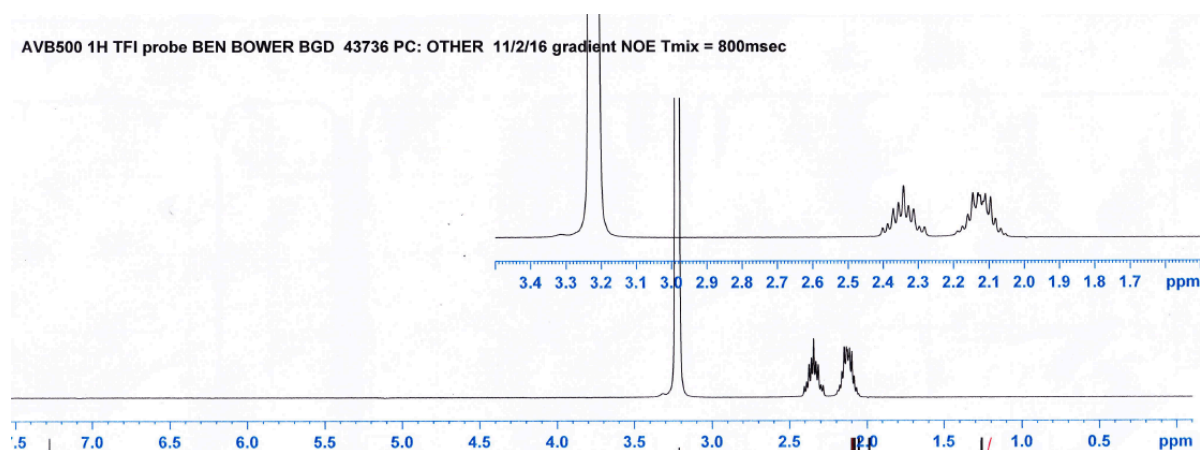
IR ($\nu_{\text{max}}/\text{cm}^{-1}$) 3400, 3340, 2938, 1749, 1734, 1659, 1537, 1429, 1367, 1231, 1112, 1087, 1039, 1017, 955.2, 902.4, 690.6 and 640.2

1.4.2.7 Peracetyl 1-(4-iodobutyramido)- β -D-GlcNAc 10

To a stirred solution of EEDQ (486 mg, 1.96 mmol) and 4-iodobutyric acid (420 mg, 1.96 mmol) in chloroform (10 mL) peracetyl 1-amino- β -D-GlcNAc (600 mg, 1.73 mmol) was added. The reaction mixture was stirred at room temperature for 24 h at which point analysis by TLC (5 % MeOH/ CHCl_3 development with 254 nm and anisaldehyde dip) had shown significant development of product ($R_f = 0.41$) The reaction mixture was loaded directly onto a chromatography column, which had been pre-equilibrated with chloroform then purified by column chromatography (0 $\rightarrow$ 6 % MeOH/ CHCl_3 gradient elution) to afford the title product as a white amorphous solid (679 mg, 1.25 mmol, 225.2 $^\circ\text{C}$ (decomp.) 72 %), which turned yellow if exposed to light for sustained periods.

^1H NMR δH (500 MHz, Chloroform-d) 6.99 (dd, $J = 10.0, 8.3$ Hz, 1H, **NHBu**), 6.11 (d, $J = 8.2$ Hz, 1H, **NHAc**), 5.05 (t, $J = 9.6$ Hz, 1H, **H1**), 5.03 (tt, $J = 4.9, 3.2$ Hz, 1H, **H4**), 4.99 (dd, $J = 9.7, 6.8$ Hz, 1H, **H3**), 4.23 (dd, $J = 12.6, 4.3$ Hz, 1H, **H6a**), 4.08 (dd, $J = 9.9, 8.4$ Hz, 1H, **H2**), 4.02 (dd, $J = 12.6, 2.2$ Hz, 1H, **H6b**), 3.71 (ddd, $J = 9.6, 4.3, 2.2$ Hz, 1H), 3.14 (appears t, $J = 6.7$ Hz, 2H, **Bu γ**), 2.35 (tt, $J = \text{Hz}$, 2H, **Bu β**), 2.2 – 2.0 (m, 2H, **Bu α**), 2.02 (s, 3H, **Ac**), 2.01 (s, 3H, **Ac**), 1.98 (s, 3H, **Ac**) and 1.91 (s, 3H, **Ac**)

In order to visualise the Bu chain peaks obscured by Ac groups a 1D TOCSY was performed irradiating the 3.2 ppm triplet with a NOE TMix = 800 msec



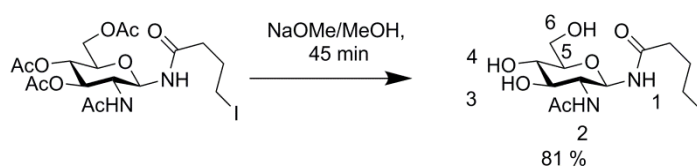
^{13}C NMR δC (100.6 MHz, Chloroform-d) 172.2 (**NAcC1**), 172.0 (**NAcC2**), 172.0 (**OAc**), 170.7 (**Ac**), 169.3 (**Ac**), 80.3 (**C1**), 73.5 (**C5**), 72.9 (**C3**), 67.7 (**C4**), 61.7 (**C6**), 53.1 (**C2**), 36.7 (**Bu α**), 28.4 (**Bu β**), 23.4 (**NAc**), 20.8 (**Ac**), 20.7 (**Ac**), 20.6 (**Ac**) and 5.5 (**Bu γ**)

MS⁺ [M + Na⁺] 656.0 and [M ¹³C + Na⁺] 566.1

MS⁻ [M + ³⁵Cl] 577.0, [M ¹³C + ³⁵Cl] 578.0, [M + FA] 587.0 and [M ¹³C+ FA] 588.0

[α]²⁰ -0.09845

IR (ν_{max}/cm⁻¹) 3294 (NH Str), 1744.51 (C=O Str Ester), 1657 (C=O Str Amide), 1534 (N-H Bend), 1435 (C-H Bend), 1412 (C-H Bend), 1370 (C-H Bend), 1305 (C-O Str), 1224, 1212, 1176, 1105, 1036, 966, 916, 900, 817, 754, 667, 640 and 610.

1.4.2.8 1-(4-iodobutyramido)- β -D-GlcNAc 11

Peracetyl 1-(4-iodobutyramido)- β -D-GlcNAc (400 mg, 0.739 mmol) was treated with, sodium methoxide in methanol (5 mM, 25 mL, 0.125 mmol) and stirred until analysis by TLC (MeOH/ CHCl_3 30 %, anisaldehyde) had shown the reaction complete by disappearance of the starting material (R_f 0.9) and the appearance of one spot (R_f 0.5). Upon completion the reaction mixture was neutralised by addition of DOWEX H^+ , the reaction mixture was filtered and the reaction mixture concentrated to dryness to afford the title product as a white to pale orange amorphous solid (248 mg, 5.96 mmol, 188.5 $^\circ\text{C}$ (decomp.), 81 %), which turned yellow if exposed to light for sustained periods.

Method adapted from Zemplén¹⁵⁸

^1H NMR δH (400 MHz, D_2O) 4.99 (d, $J = 9.7$ Hz, 1H, **H1**), 3.79 (dd, $J = 12.4, 1.9$ Hz, 1H, **H6_a**), 3.73 (t, $J = 9.9$ Hz, 1H, **H2**), 3.66 (dd, $J = 12.3, 4.7$ Hz, 1H, **H6_b**), 3.51 (dd, $J = 10.2, 8.3$ Hz, 1H, **H3**), 3.44-3.35 (m, 2H, **H4**, **H5**), 3.20-3.05 (m, 2H, **Bu γ**), 2.31 (td, $J = 7.1, 2.7$ Hz, 2H, **Bu α**), 2.01-1.95 (m, 2H, **Bu β**) and 1.94 (s, 3H, **Ac**)

^{13}C NMR δC (100.6 MHz, D_2O) 176.11 (**Bu C=O**), 174.59 (**NAc C=O**), 78.20 (**C1**), 77.55 (**C5**), 74.18 (**C3**), 69.37 (**C4**), 60.40 (**C6**), 54.22 (**C2**), 36.30 (**Bu α**), 28.59 (**Bu β**), 22.18 (**Ac**) and 5.28 (**Bu γ**)

MS⁺ 439.0 [**M** + Na^+] and 440.0 [**M** ^{13}C + Na^+]

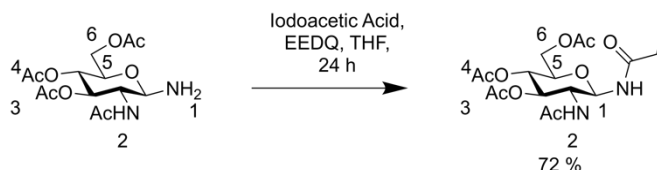
MS⁻ 455.0 [**M** + ^{35}Cl], 457.0 [**M** + ^{37}Cl], 461.0 [**M** + **FA**] and 462.0 [**M** ^{13}C + **FA**]

$[\alpha]^{20}$ + 0.0534

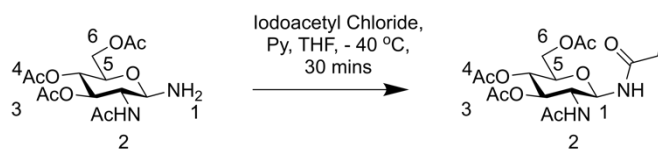
IR ($\nu_{\text{max}}/\text{cm}^{-1}$) 3286.4 (NH str, br), 3254 (OH str, br), 2933.1 (CH str), 2852.9 (CH str), 1651.4 (C=O str), 1543.4 (C-C Str), 1437.4, 1364.5 (C-H wag), 1321.2, 1276.3, 1258.2, 1206.5, 1169.73 (C-O Str), 1109.14, 1087.6, 1071.3, 1040.1, 980.42, 957.0 (NH wag), 915.75, 893.75, 755.7, 720.2, 669.5 and 634.1

1.4.2.9 Peracetyl 1-(iodoacetamide)- β -D-GlcNAc 8

EEDQ method:



To a stirred solution of EEDQ (427 mg, 1.73 mmol) and iodoacetic acid (323 mg, 1.73 mmol) in THF (10 mL) peracetyl 1-amino- β -D-GlcNAc (500 mg, 1.73 mmol) was added. The reaction mixture was stirred at room temperature for 24 h at which point analysis by TLC (5 % MeOH/ CHCl_3 development with 254 nm and anisaldehyde dip) had shown significant development of product ($R_f = 0.38$) The reaction mixture was concentrated to dryness on diatomaceous earth (5.00 g) then loaded onto a chromatography column which had been pre-equilibrated with chloroform then purified by column chromatography (0 $\rightarrow$ 10 % MeOH/ CHCl_3 gradient elution) to afford the title product as a white amorphous solid (410 mg, 951 μmol , 200 $^\circ\text{C}$ (decomp.), 55 %), which turned yellow if exposed to light for sustained periods.

Acid chloride method:

To a rapidly stirred solution of peracetyl 1-amino-β-D-GlcNAc (500 mg, 1.73 mmol) in anhydrous DCM (10 mL) with anhydrous pyridine (150 μL, 147 mg, 1.86 mmol) precooled in a dry ice/MeCN bath freshly distilled 2-iodoacetyl chloride (380 mg, 1.86 mmol) was added. The reaction mixture was stirred at this temperature for 30 mins. The black reaction mixture was rapidly poured into chloroform (100 mL) then quenched by washing with sat. aq. NaHCO₃ (3 x 100 mL). The organic phase was purified by column chromatography chromatography (0 -> 10 % MeOH/CHCl₃ gradient elution) to afford the title product as a white amorphous solid (4.4 mg, 8.65 μmol, 0.5 %), which turned yellow if exposed to light for sustained periods.

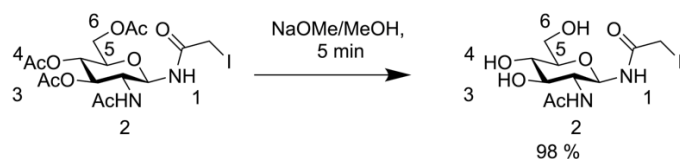
Method adapted from Flitsch¹⁹⁸

^1H NMR δH (400 MHz, d_6 -DMSO) 8.99 (d, $J = 9.2$ Hz, 1H, **NHAcI**), 7.96 (d, $J = 9.4$ Hz, 1H, **NHAc**), 5.09 (m (appears q, $J = 10.1$ Hz), 2H, **H1**, **H4**), 4.81 (t, $J = 9.8$ Hz, 1H, **H3**), 4.17 (dd, $J = 12.5, 4.4$ Hz, 1H, **H6_a**), 3.95 (dd, $J = 12.5, 2.3$ Hz, 1H, **H6_b**), 3.93 – 3.87 (q, $J = 10.1$ Hz, 1H, **H2**), 3.84 (ddd, $J = 10.1, 4.2, 2.3$ Hz, 1H, **H5**), 3.69 (d, $J = 9.7$ Hz, 1H, **AcI_a**), 3.58 (d, $J = 9.7$ Hz, 1H, **AcI_b**), 2.00 (**Ac**), 1.96 (**Ac**), 1.91 (**Ac**) and 1.76 (**Ac**)

^{13}C NMR δC (100.6 MHz, D_2O) 170.0 (**Ac**), 169.4 (**Ac**), 169.2 (2 x **Ac**), 168.6 (**Ac**), 78.2 (**C1**), 73.1 (**C4**) 72.3 (**C5**) 68.3 (**C3**), 61.7 (**C6**), 51.8 (**C4**), 22.67 (**NAc**), 20.5 (**OAc**), 20.4 (**OAc**), 20.3 (**OAc**) and -0.5 (**OAcI**)

MS^+ 516.2 [**M** + H^+], 517.2 [**M** ^{13}C + H^+] and 537.0 [**M** + Na^+]

MS^- 513.0 [**M** - H^+], 559.0 [**M** + FA], 627.0 [**M** + TFA]

1.4.2.10 1-(iodoacetamide)- β -D-GlcNAc 9

To a solution of peracetyl 1-(iodoacetamide)- β -D-GlcNAc (410 mg, 0.796 mmol) in methanol (8 mL) sodium methoxide in methanol (25 %, 200 μ L) was added and stirred for 5 minutes at which point analysis by TLC (MeOH/ CHCl_3 30 %, anisaldehyde) had shown the reaction complete by disappearance of the starting material (R_f 0.9) and the appearance of one spot (R_f 0.45). The reaction mixture was neutralised by addition of DOWEX H^+ (352 mg) and stirred for 5 minutes, the reaction mixture was filtered and then concentrated to dryness to afford the title product as a white to pale orange amorphous solid (303 mg, 774 μ mol, 188.5 $^\circ\text{C}$ (decomp.), 98 %), which turned brown if exposed to light for sustained periods.

Method adapted from Zemplén¹⁵⁸

^1H NMR δH (600 MHz, d_6 -DMSO) 8.46 (d, $J = 9.2$ Hz, 1H, **NHAcI**), 7.80 (d, $J = 9.0$ Hz, 1H, **NHAc**), 4.76 (t, $J = 9.5$ Hz, 1H, **H1**), 3.71 (d, $J = 9.7$ Hz, **AcI_a**), 3.66 (dd, $J = 11.9$ Hz, 2.0 Hz, 1H, **H6_a**), 3.57 (d, $J = 9.7$ Hz, **AcI_b**), 3.58-3.53 (m, 1H, **H2**) 3.43 (dd, $J = 12.0, 5.5$, 1H, **H6_b**), 3.29 (dd, $J = 10.2, 7.9$ Hz, **H3**), 3.13-3.09 (m, 1H, **H5**), 3.07 (dd, $J = 9.9$ Hz, 8.3 Hz, 1H, **H4**) and 1.82 (s, 3H, **Ac**)

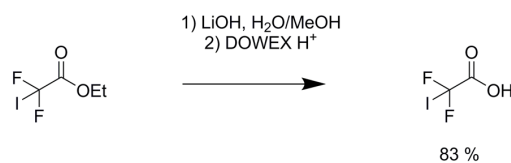
The experiment must be conducted under water suppression conditions in order to observe all signals

^{13}C NMR δC (100.6 MHz, d_6 -DMSO) 169.2 (**NHAcI**), 168.0 (**NHAc**), 78.7 (**C1**), 78.5 (**C5**), 74.0 (**C3**), 70.0 (**C4**), 60.4 (**C6**), 53.8 (**C2**), 22.6 (**NHAc**) and -0.5 (**NHAcI**)

MS⁺ 390.2 [**M** + **H⁺**] and 411.0 [**M** + **Na⁺**]

HRMS *Expected*: 389.02041 *Found*: 389.02049

1.4.2.11 2,2-Difluoro-2-iodoacetic Acid 18

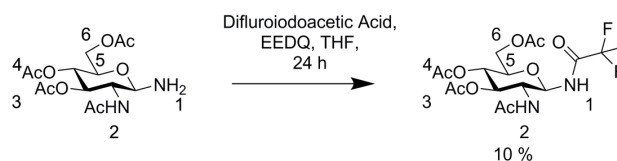


To a methanolic (85 mL) solution of ethyl 2,2-difluoro-2-iodoacetate (1.00 g, 4.00 mmol), aqueous lithium hydroxide (2.5 M, 15 mL) was added. The orange pink reaction mixture immediately turned colourless. The reaction mixture was stirred at room temperature for 2 h at which point analysis by TLC (100 % EtOAc, permanganate development) shows the reaction complete by complete loss of the starting material ($R_f = 0.6$) and the formation of a single product ($R_f = 0.1 - 0.0$). The reaction mixture was neutralised with DOWEX H^+ , filtered then concentrated to afford a thick colourless oil, which was dried azeotropically with toluene (4 x 20 mL) to afford the title product as an off white amorphous solid (738 mg, 3.33 mmol, 83 %).

^1H NMR δH (500 MHz, $\text{D}_4\text{-MeOH}$)

^{13}C NMR δC (100.6 MHz, $\text{D}_4\text{-MeOH}$) 163.2 (COOH), 124.9 (CF_2I)

^{19}F NMR δF (376.3 MHz, $\text{D}_4\text{-MeOH}$) -58.6

1.4.2.12 Peracetyl 1-(2,2-difluoro-2-iodoacetamido)- β -D-GlcNAc 19

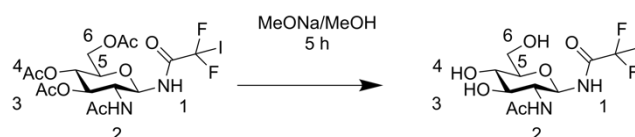
To a stirred solution of EEDQ (427 mg, 1.73 mmol) and 2,2-difluoro-2-iodoacetic acid (384 mg, 1.73 mmol) in THF (20 mL) peracetyl 1-amino- β -D-GlcNAc (500 mg, 1.44 mmol) was added. The reaction mixture was stirred at room temperature for 26 h at which point analysis by TLC (10 % MeOH/ CHCl_3 development with 254 nm and anisaldehyde dip) had shown significant development of product ($R_f = 0.41$). The reaction mixture was loaded directly onto a chromatography column, which had been pre-equilibrated with chloroform then purified by column chromatography (0 $\rightarrow$ 8 % MeOH/ CHCl_3 gradient elution) to afford the title product as an off white amorphous solid (80 mg, 0.145 mmol, 10 %), which turned yellow if exposed to light for sustained periods.

^1H NMR δH (500 MHz, CHCl_3) 7.97 (1H, d, $J = 7.9$ Hz, **NH1**), 6.25 (1H, d, $J = 7.9$ Hz, **NH2**), 5.11-5.01 (m, 2H, **H3**, **H4**), 4.94 (1H, dd, $J = 9.8, 7.9$ Hz, **H1**), 4.25 (1H, dd, $J = 12.5, 4.5$ Hz, **H6a**), 4.21-4.12 (m, 1H, **H2**), 4.04 (1H, dd, $J = 12.5, 2.3$ Hz, **H6b**), 3.75 Hz (1H, ddd, $J = 9.6, 4.4, 2.3$ Hz, **H5**), 2.03 (3H, s, **Ac6**), 2.03 (3H, s, **Ac4**), 1.99 (3H, s, **Ac3**) and 1.94 (3H, s, **Ac2**)

^{13}C NMR δC (100.6 MHz, CHCl_3) 172.4 (**C=O 6**), 172.1 (**C=O 4**), 170.7 (**C=O 3**), 169.3 (**C=O 2**), 162.0 (**C=O 1**), 89.8 (**CIF₂**), 81.0 (**C1**), 73.9 (**C5**), 72.5 (**C3**), 67.6 (**C4**), 61.6 (**C6**), 53.3 (**C2**), 23.1 (**Ac2**), 20.7 (**Ac3**, **Ac4**) and 20.6 (**Ac 6**)

^{19}F NMR δF (376.3 MHz, CHCl_3) -58.31 (d, $J = 31.7$ Hz)

MS 549.0 [**M** - H^+] and 550.0 [**M** ^{13}C + H^+]

1.4.2.13 1-(2,2-difluoro-2-iodoacetamido)- β -D-GlcNAc 20

Peracetyl-1-(2,2-difluoro-2-iodoacetamido)- β -D-GlcNAc (120 mg, 218 μ mol) in anhydrous MeOH (2 mL) was treated with sodium methoxide in methanol (25 %, 50 μ L) and stirred until analysis by TLC (MeOH/ CHCl_3 30 %, anisaldehyde) had shown the reaction complete by disappearance of the starting material (R_f 0.9, 15 mins) and the appearance of one spot (R_f 0.4). Upon completion the reaction mixture was neutralised by addition of DOWEX H^+ (88 mg) and stirred for 5 minutes, the reaction mixture was filtered and then concentrated to dryness to afford the title product as pale yellow amorphous solid (85 mg, 200 μ mol, 92 %)

Method adapted from Zemplén¹⁵⁸

^1H NMR δH (400 MHz, $\text{d}_4\text{-MeOH}$) 4.91 (d, $J = 9.8$ Hz, 1H, **H1**) 3.76 (dd, $J = 12.1$, 1.6 Hz, 1H, **H6_a**) 3.76 (t, $J = 10.0$, 1H, **H2**), 3.62-3.58 (m, 1H, **H6_b**), 3.43-3.38 (m, 1H, **H5**), 3.28-3.26 (m, 2H, **H3**, **H4**) and 1.90 (s, 3H, **NHAc**).

^{13}C NMR δC (100.6 MHz, $\text{d}_4\text{-MeOH}$) 173.0 (**NHAc**), 89.6 (**NHAcF₂I**), 79.5 (**C1**), 78.8 (**C3**), 74.5 (**C4**), 70.2 (**C5**), 61.1 (**C6**), 54.6 (**C2**), and 21.6(**NHAc**)

^{19}F NMR δF (376.3 MHz, $\text{d}_4\text{-MeOH}$) -61.9 (d, $J = 15.7$ Hz)

MS^+ 447.0 [**M** + **Na⁺**] and 448.0 [**M** ^{13}C + **Na⁺**]

MS^- 423.0 [**M** - **H⁺**], 424.0 [**M** ^{13}C + **H⁺**]

1.4.2.14 General method for sugar peracetylation

To a stirred ice-cold suspension of monosaccharide (21.6 g, 120 mmol) in Ac_2O (62.3 mL, 660 mmol) and pyridine (240 mL), DMAP (1.46 gm 12.0 mmol) was added cautiously. The reaction mixture was stirred at room temperature until analysis by TLC (100 % EtOAc) showed complete disappearance of the starting material and visual inspection of the reaction vessel showed all suspended sugar had dissolved. The reaction mixture was partitioned between Et_2O (250 mL) and aqueous CuSO_4 (250 mL, 10 %w/w). The organic fraction was washed a further five times with aqueous CuSO_4 (150 mL, 10 %w/w, until the organic fraction remains colourless) and once with aqueous EDTA (pH 8.0 saturated). The organic fraction was dried on anhydrous MgSO_4 , filtered then concentrated under reduced pressure. Peracetylated sugars were used without further purification.

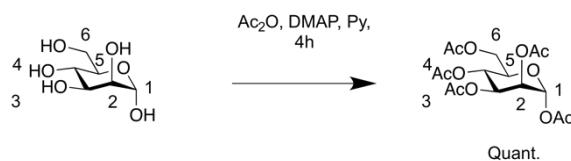
1.4.2.15 General method for sugar iodoethanol glycosylation

Peracetylated sugar (8.00 g, 20.5 mmol) and 2-iodoethanol (4.8 mL, 61.5 mmol) were dissolved in anhydrous and DCM (80 mL) and stirred with freshly activated 4 Å molecular sieves (15 g) for 2 h under a continuous flow of Ar. The solution was cooled with an ice bath then $\text{BF}_3 \cdot \text{OEt}_2$ (15.5 mL, 123 mmol) was added dropwise. The reaction mixture was stirred under Ar (the continuous flow was halted to prevent solvent evaporation) for 16 h. The reaction was quenched by the addition of K_2CO_3 (3.50 g, 25.0 mmol). Solids were removed by filtration through glass wool then the reaction mixture washed with water (2 x 200 mL). The organic fraction was dried on Na_2SO_4 , filtered and concentrated under reduced pressure. The resultant oil was purified by column chromatography (30 -> 50 % EtOAc/Hexane)

1.4.2.16 General Method for 1-(2-Iodoethoxy)aldohexose deacetylation

To a stirred suspension of peractyl 1-(2-Iodoethoxy)aldohexose (0.500 g), in anhydrous MeOH (10 mL), MeONa in MeOH (25 %, 80 μL) was added. The reaction mixture was stirred until complete, with TLC analysis (30 % MeOH/ CHCl_3 , 254 nm and anisaldehyde development) being performed every 5 minutes. Upon completion the reaction mixture is quenched by the addition of DOWEX H^+ . The resin is removed by filtration and solvents removed under reduced pressure to afford the deprotected product, which was used without further purification.

Method adapted from Zemplén¹⁵⁸

1.4.2.17 Peracetyl α,β -D-Mannose 15

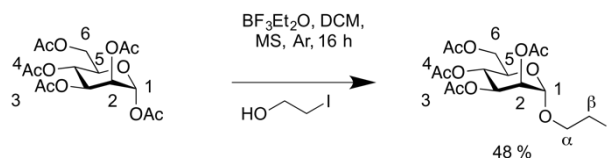
Isolated as a colourless oil, which started to crystallise when left for prolonged periods.

NMR characterisation reported for α anomer, a small amount of the β anomer can also be seen but the abundance is < 10 %

^1H NMR δH (400 MHz, Chloroform- d) 6.02 (1H, d, J = 1.8, H1), 5.29-5.27 (2H, m, H3 + H4), 5.20 (1H, td, J = 2.4, 0.6 Hz, H2), 4.21 (1H, dd, J = 12.4, 5.0 Hz, H6a), 4.03 (1H, dd, J = 12.5, 2.2 Hz, H6b), 4.02-3.78 (1H, m, H5), 2.12 (3H, s, Ac), 2.11 (3H, s, Ac), 2.03 (3H, s, Ac), 1.99 (3H, s, Ac) and 1.95 (3H, s, Ac)

^{13}C NMR δC (100.6 MHz, Chloroform- d) 170.6 (**Ac**), 170.0 (**Ac**), 169.7 (**Ac**), 169.5 (**Ac**), 168.1 (**Ac**), 90.6 (**C1**), 70.6 (**C5**), 68.7 (**C3**), 68.3 (**C2**), 65.5 (**C4**), 62.1 (**C6**), 20.9 (**Ac**), 20.8 (**Ac**), 20.7 (**Ac**), 20.7 (**Ac**) and 20.7 (**Ac**)

MS $^+$ 413.0 [$\text{M} + \text{Na}^+$], 414.0 [$^{13}\text{C} \text{ M} + \text{Na}^+$] 469.4 [$\text{M} + \text{DMSO} + \text{H}^+$]

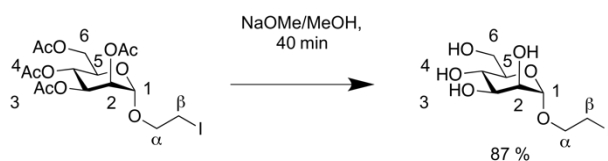
1.4.2.18 Peracetyl 1-(2-iodoethoxy)- α -D-Mannose 16

Isolated as a colourless solid, which turned yellow when exposed to light for prolonged periods.

^1H NMR δH (400 MHz, Chloroform- d) 5.29 (1H, dd, $J = 10.0, 3.3$ Hz, H3), 5.23 (1H, d, $J = 9.5$ Hz, H4), 5.21 – 5.19 (1H, m, H2), 4.81 (1H, d, $J = 1.6$ Hz, H1), 4.21 (1H, dd, $J = 12.8, 6.0$, H6a), 4.09 – 4.05 (2H, m, H6b + H5), 3.87 (1H, dt, $J = 11.0, 7.0$ Hz, H α a), 3.77 (1H, dt, $J = 11.0, 6.5$ Hz, H α b), 3.24 (1H, t, $J = 6.7$, H β), 2.10 (3H, s, Ac), 2.05 (3H, s, Ac), 1.99 (3H, s, Ac) and 1.94 (3H, s, Ac)

^{13}C NMR δC (100.6 MHz, Chloroform- d) 169.1 (Ac), 168.5 (Ac), 168.3 (Ac), 168.2 (Ac), 96.0 (C1), 67.9 (C6), 67.7 (C5), 67.6 (C2), 67.4 (C3), 64.5 (C4), 60.9 (C α), 19.4 (Ac), 19.3 (Ac), 19.2 (Ac), 19.2 (Ac) and 0.00 (C β)

MS $^+$ 525.0 [M + H $^+$], 526.0 [^{13}C M + H $^+$]

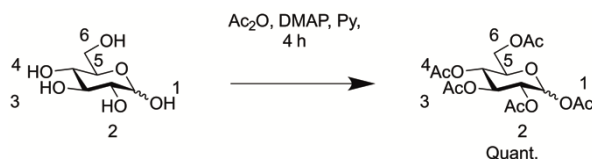
1.4.2.19 1-(2-iodoethoxy)- α -D-Mannose 17

Isolated as a colourless solid, which turned yellow when exposed to light for prolonged periods.

^1H NMR δH (400 MHz, D_2O) 4.81 (1H, d, $J = 1.5$ Hz, H1), 3.85 (1H, dd, $J = 3.5, 1.5$ Hz, H2), 3.84 – 3.80 (1H, m, H β b), 3.78–3.74 (1H, m, H6a), 3.71 – 3.65 (2H, m, H3 + H β b), 3.64–3.60 (2H, m, H5 + H6b). 3.53 (1H, t, $J = 9.5$ Hz, H4) and 3.32 – 3.23 (2H, m, H α)

^{13}C NMR δC (100.6 MHz, D_2O) 99.5 (C1), 73.1 (C5), 70.4 (C3), 69.9 (C2), 68.0 (C α), 66.6 (C4), 60.8 (C6) and 3.3 (C β)

MS 369.0 [$\text{M} + ^{35}\text{Cl}^-$], 371.0 [$\text{M} + ^{37}\text{Cl}^-$]

1.4.2.20 Peracetyl α,β -D-Glucose 12

Isolated as a colourless amorphous solid

^1H NMR δH (400 MHz, Chloroform- d) 5.65 (d, $J = 9.1$ Hz, 1H, NH), 5.19 – 4.85 (m, 2H, H3 and H4), 4.18 (dd, $J = 12.3, 4.9$ Hz, 1H, H6_b), 4.08 (d, $J = 9.18$ Hz, 1H, H1), 4.07 (dd, $J = 12.31, 2.14$, 1H, H6_a) 4.02 – 3.93 (m, 2H, H2 and H5), 2.06 (s, 3H, Ac6), 2.01 (s, 3H, Ac4), 2.00 (s, 3H, Ac3) and 1.94 (s, 3H, Ac2).

^{13}C NMR δC (100.6 MHz, Chloroform- d) 171.7 (NC(=O)Me), 170.8 (OC(=O)Me3), 170.8 (OC(=O)Me4), 169.4 (OC(=O)Me6), 86.5 (C1), 73.4 (C2), 72.8 (C3), 68.4 (C4), 62.4 (C6), 54.9 (C5), 23.4 (NC(=O)Me), 20.8 (OC(=O)Me3), 20.7 (OC(=O)Me4) and 20.6 (OC(=O)Me 6).

MS⁺ 347.1 [M + H⁺], 715.3 [2M + Na⁺]

MS⁻ 381

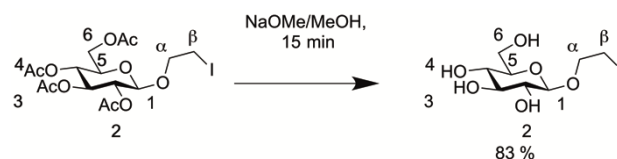
1.4.2.21 Peracetyl 1-(2-iodoethoxy)- β -D-Glucose 13

Isolated as a colourless solid, which turned yellow when exposed to light for prolonged periods.

^1H NMR δH (400 MHz, Chloroform- d) 5.15 (1H, t, $J = 9.5$ Hz, **H3**), 5.02 (1H, t, $J = 9.9$ Hz, **H4**), 4.95 (1H, dd, $J = 9.5, 8.0$ Hz, **H2**), 4.50 (1H, d, $J = 8.0$, **H1**), 4.19 (1H, dd, $J = 12.4, 4.8$ Hz, **6a**), 4.07 (1H, dd, $J = 12.4, 2.5$ Hz, **H6b**), 4.03 (1H, m, **H α a**), 3.71 (1H, dt, $J = 11.2, 7.3$ Hz, **H α b**), 3.64 (1H, ddd, $J = 10.0, 4.8, 2.4$ Hz, **H5**), 3.19 (2H, m, **H β**), 2.03 (3H, s, **Ac**), 2.02 (3H, s, **Ac**), 1.96 (3H, s, **Ac**) and 1.95 (3H, s, **Ac**)

^{13}C NMR δC (100.6 MHz, Chloroform- d) 168.6 (**Ac**), 168.2 (**Ac**), 167.3 (**Ac**), 167.3 (**Ac**), (**C1**), 70.5 (**C3**), 69.8 (**C5**), 68.9 (**C2**), 68.4 (**C α**), 66.2 (**C4**), 59.7 (**C6**), 18.5 (**Ac**) and 0.0 (**C β**)

MS^+ 525.0 [$\text{M} + \text{H}^+$], 526.0 [$^{13}\text{C} \text{ M} + \text{H}^+$]

1.4.2.22 1-(2-iodoethoxy)- β -D-Glucose 14

Isolated as a colourless solid, which turned yellow when exposed to light for prolonged periods.

^1H NMR δH (400 MHz, D_2O) 4.40 (1H, d, $J = 8.0$ Hz, **H1**), 4.02 (1H, dt, $J = 11.6, 6.4$ Hz, **H $\alpha\alpha$**), 3.86 (1H, dt, $J = 11.6, 6.5$ Hz, **H αb**), 3.79 (1H, dd, $J = 12.4, 2.2$ Hz, **H6a**), 3.59 (1H, dd, $J = 12.3, 5.6$ Hz, **H6b**), 3.34 (2H, m, **H3 + H5**), 3.26 (3H, m, **H α + H4**) and 3.18 (1H, dd, $J = 9.2, 8.0$ Hz, **H2**)

^{13}C NMR δC (100.6 MHz, D_2O) 102.1 (**C1**), 75.9 (**C5**), 75.6 (**C3**), 73.0 (**C2**), 70.7 (**C α**), 69.5 (**C4**), 60.6 (**C6**) and 2.59 (**C β**)

MS $^+$ 357.0 [**M** + Na^+], 358.0 [^{13}C **M** + Na^+], 413.0 [**M** + DMSO + H^+] and 414.0 [^{13}C **M** + DMSO + H^+]

MS $^-$ 369.0 [**M** + $^{35}\text{Cl}^-$], 371.0 [**M** + $^{37}\text{Cl}^-$], 379.0 [**M** + $\text{FA} - \text{H}^+$] and 380.0 [**M** ^{13}C + $\text{FA} - \text{H}^+$], 667.0 [**2M** - H^+] and 668.0 [^{13}C **M** + **M** - H^+]

1.4.3.0 General Protein Experimental Considerations

1.4.3.1 General

Media and equipment used in the handling of microbial cultures were sterilised in a Falcon LTE Touchclave-R autoclave at 121 °C for 20 mins. Biological work was carried out in a Bassaire IS Clean Air laminar flow hood. *E. coli* cultures were incubated in New Brunswick Scientific Innova 42 incubator at 37 °C 200 RPM, unless mentioned otherwise. *E. coli* cultures were grown in 2 L baffled flasks with 500 mL of media inside and air exchange sponges plugging the necks.

Centrifugation was performed in: a Beckman Coulter Avanti JXN centrifuge using a JA-30.60 TI rotor for 50 mL samples or a JLA 9.1000 for 1 L samples, a Heraeus Megafuge 40R centrifuge with a F15-6x100y rotor for 50 mL samples, a Beckman Coulter Microfuge 18 Centrifuge for microcentrifugation or a technico mini.

Ultrapure water was prepared using a Millipore Direct-Q 3 UV system. The heating bath used for heat shocking was a Grant Sub Aqua 2 Plus. PCR was performed on Applied Biosystems GeneAmp PCR System 9700.

Gel electrophoretic analysis was performed using a Biorad Powerpac Basic powerpack in an Innovex XCell Sure Lock tank with MES running buffer. Novex Invitrogen NuPAGE 4-12 % bis-tris gels were used. Gels were developed with Sigma Aldrich ProteoSilver or Expedeon Instant Blue.

1.4.3.2 Common Protein Solutions

Expression Stocks:

The following solutions are prepared as 1000X stocks:

Ampicillin: 100 mg mL⁻¹ in ultrapure water

Carbenicillin: 100 mg mL⁻¹ in ultrapure water

Chloramphenicol: 34 mg mL⁻¹ in ethanol

Kanamycin: 50 mg mL⁻¹ in ultrapure water

IPTG: 300 mg mL⁻¹ in ultrapure water

Broths Employed:

<i>Lauria (LB):</i>	Tryptone 10 g L ⁻¹ , Yeast Extract 5 g L ⁻¹ , NaCl 5 g L ⁻¹
<i>2 x Yeast Tryptone (2xYT)</i>	Tryptone 16 g L ⁻¹ , Yeast Extract 10 g L ⁻¹ , NaCl 5 g L ⁻¹
<i>Super (SB):</i>	Tryptone 35 g L ⁻¹ , Yeast Extract 20 g L ⁻¹ , NaCl 5 g L ⁻¹ , 1 M NaOH 5 mL L ⁻¹
<i>Terrific (TB):</i>	Tryptone 12 g L ⁻¹ , Yeast Extract 24 g L ⁻¹ , K ₂ HPO ₄ 2.2 g L ⁻¹ , KH ₂ PO ₄ 2.2 g L ⁻¹ , Glycerol 8 mL L ⁻¹
<i>Muller-Hinton (MHB):</i>	Beef infusion solids 2.0 g L ⁻¹ , Casein 17.5 g L ⁻¹ , Starch 1.5 g L ⁻¹
<i>Super Optimal Broth with</i>	Tryptone 20 g L ⁻¹ , Yeast Extract 5.0 g L ⁻¹ , NaCl
<i>Catabolite Repression (SOC):</i>	0.6 g L ⁻¹ , KCl 0.2 g L ⁻¹ , MgCl ₂ 1.0 g L ⁻¹ , MgSO ₄ ·(H ₂ O) ₆ , Glucose 3.6 g L ⁻¹

Common Buffers:

<i>PBS</i>	NaCl 137 mM KCl 2.7 mM Na ₂ HPO ₄ 10 mM KH ₂ PO ₄ 1.8 mM pH 7.4
<i>fPBS</i>	NaF 150 mM Na ₂ HPO ₄ 10 mM KH ₂ PO ₄ 1.8 mM pH 7.4
<i>SAU-100</i>	Urea 7 M, NaOAc 20 mM, NaCl 100 mM, BME 5 mM, EDTA 1 mM, pH 5.2
<i>SAU-1000</i>	Urea 7 M, NaOAc 20 mM, NaCl 1 M, BME 5 mM, EDTA 1 mM, pH 5.2
<i>MES Running</i>	MES 50 mM, Tris 50 mM, SDS 0.1 %, EDTA 1 mM, pH 7.3
<i>Laemmli</i>	SDS 2 %, Glycerol 10 %, Tris 60 mM, bromophenol blue 0.01 % pH 6.8
<i>TBS</i>	100 mM NaCl 50 mM Tris pH 7.6
<i>Wash Buffer</i>	Tris 50 mM, NaCl 100 mM, BME 5 mM, EDTA 1 mM, pH 7.5
<i>TW Buffer</i>	As wash buffer + 1 % Triton X-100
<i>Unfolding Buffer</i>	Guanidinium chloride 7 M, Tris 10 mM, DTT 10 mM, Benzamidine HCl 1 mM, EDTA 1 mM, pH 7.5
<i>Refolding Buffer</i>	NaCl 2 M, Tris 10 mM, BME 5 mM, EDTA 1 mM, pH 7.5

Lysis Buffer TBS + 25 mM imidazole and 1 complete protease inhibitor
cocktail tablet per 50 mL

H3 Reaction Buffer Guanidinium Chloride 5 M, NH₄OAc 500 mM pH 6.0

H3 Dha Buffer Guanidinium Chloride 3 M, Na₂H₃PO₄ 100 mM pH 8.0

Glycobuffer 1 10 x 50 mM CaCl₂, 500 mM NaOAc, pH 5.5

Storage Buffer 50 mM NaCl, 20 mM Tris, 1 mM EDTA, pH 7.5

HBS-P 10 mM HEPES, 150 mM NaCl, P-20 Detergent 5 ppm, pH 7.4

Acetate 10 mM Sodium Acetate pH 4.0

1.4.3.3 Protein Mass Spectrometry

Protein LCMS was conducted on either:

1. A Waters LCT with Shimadzu 20 HPLC Thermo Scientific and ProswiftTM RP-2H (4.6x50 mm SS) column. *Capillary Voltage*: 3 kV, *Cone Voltage*: 100 V. Nitrogen was used as the desolvation gas at a flow rate of 600 L h⁻¹. A gradient of 5 % MeCN in H₂O with 0.1 % FA to 60 % MeCN in H₂O with 0.1 % FA was ran over 2 minutes.
2. A Waters LCT Premier XE with a Waters 1525 Micro HPLC and RP18e (2 X 5 mm) guard cartridge. *Capillary Voltage*: 3.2 kV, *Cone Voltage*: 25 V. Nitrogen was used as the desolvation gas at a flow rate of 300 L h⁻¹. A gradient of 5 % MeCN in H₂O with 0.1 % FA to 40 % MeCN in H₂O with 0.1 % FA was ran over 6 minutes.
3. A Waters Xevo G2-S QTof with a Waters Acquity UPLC and ProswiftTM RP-2H (4.6x50 mm SS) column. *Capillary Voltage*: 3.2 kV, *Cone Voltage*: 160 V. Nitrogen was used as the desolvation gas at a flow rate of 600 L h⁻¹. A gradient of 5 % MeCN in H₂O with 0.1 % FA to 100 % MeCN in H₂O with 0.1 % FA was ran over 8 minutes.
4. An Agilent 6230 TOF MS with a 1260 Infinity HPLC and poroshell 5 µm 3000 B C3 column. *Capillary Voltage*: 4 kV, *Cone Voltage*: 2 kV. Nitrogen was used as the desolvation gas at a flow rate of 600 L h⁻¹. A gradient of 10 % MeCN in

H₂O with 0.1 % FA to 90 % MeCN in H₂O with 0.1 % FA was ran over 4 minutes. The TOF was ran in 1 GHz extended mass mode (m/z range 20000) for the analysis of intact antibodies.

Data was analysed using Waters MassLynx for data acquired on Waters instruments and Agilent Technologies MassHunter BioConfirm for data acquired on Agilent instruments. Maximum entropy deconvolution was employed in both cases.

1.4.3.4 Circular Dichroism

CD was conducted with an Applied Photophysics Chirascan Circular Dichroism Spectrometer in a Peltier cell holder, which controlled the temperature. Experiments were performed in triplicate and the average taken, data was recorded at 20 °C with the path length of 1 mm. Protein concentration was 0.1 mg mL⁻¹ in fPBS buffer, described above. CD was recorded at 0.5 nm intervals between 185 and 260 nm. The average of 3 spectra is used in all cases.

1.4.3.5 Fast Liquid Protein Chromatography

FPLC was performed using a GE Healthcare ÄKTA purifier system, with prepacked columns from GE Healthcare. Size exclusion was performed on superdex s200 resin.

1.4.3.6 Thermal Florescence

Thermal florescence samples were prepared by myself then samples ran by James Heads at UCB. Experiments were ran in quadruplet.

1.4.3.7 Transformation Protocol

1 μ L of plasmid and 20 μ L of cells were incubated on ice for 40 minutes. The mixture was heated to 42 °C for 45 s then cooled on ice for 10 minutes. 200 μ L of SOC media was added and then shaken (200 RPM) at 37 °C for 1 h. The cells are split (180 μ L and 20 μ L) between two LB/agar plates containing the appropriate antibiotic and dispersed using glass beads. Plates were then incubated at 37 °C until required typically overnight but not longer than one day.

1.4.3.8 Protein Concentration Measurement

Protein concentration was typically measured using A_{280} from a NanoDrop ND-1000 or an Expedeon Versawave. However, for experiments where concentration plays a critical role (such as CD) a BCA assay was used, ran against *bovine* serum albumin standards.

1.4.4.0 Mutagenesis

Mutagenesis was performed with an Agilent Technologies QuikChange kit according to instructions provided. Primers were ordered from Sigma-Aldrich custom Oligos service.

1.4.4.1 AcrA wt to AcrA N123C

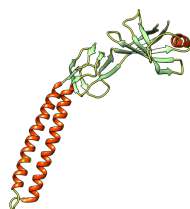
Primer: gccgtatttgataactgtaactctacacttttaccg

1.4.5.0 Images

Images of protein structures were taken from RCSB PDB²⁰⁶ and processed using UCSF Chimera.²⁰⁷ Individual references for structures are: AcrA,²⁰⁸ AcrA-AcrB-AcrZ-TolC System,¹⁸⁴ Histone H3,²⁰⁹ EndoA,²¹⁰ hOGA,²¹¹ GalT,²¹² cAB-Lys²¹³ and Fab PCT64_LMCA²¹⁴ (used as a 'general protein' for figures). Any chemical sidechains were added using Perkin Elmer Chemdraw Professional (16.0.1.4). Abstract glycan diagrams were drawn according the ACS protocol using OMICtools Glycan Builder.

1.4.6.0 Protein Expression

1.4.6.1 Multidrug efflux pump subunit AcrA Expression (wt and 123 Cys)



Adapted from Davis²¹⁵

Plasmids (pET24) containing the appropriate AcrA gene were transformed into BL-21(DE3) *E. coli* cells and plated on Kanamycin/LB/agar plates. Single colonies were picked and 4 x 5 mL starter cultures (LB/Kanamycin) of were grown over night as 37 °C 200 RPM then transferred into 4 x 500 mL of media (LB/Kanamycin) in 2 L baffled conical flasks. The cultures were grown at 37 °C until an optical density of 0.6 – 0.8 was achieved (all between 40 – 70 min) at which point the cells were induced by the addition of IPTG (to give a concentration of 3 mM) and incubated for a further 4 h. The cells were then pelleted for 10 minutes at 8 kRPM 9.6 kG at $r_{average}$.

Cell pellets were resuspended in buffer (50 mL of 50 mM tris, 100 mM NaCl, 10 mM imidazole, 1 mg mL⁻¹ lysozyme and 0.1 mg mL⁻¹ DNase) and stirred on ice for 1 h. The suspension was then subjected to sonication (50 % power, 30 s sonication 1 min rest, four times), with the resultant mixtures treated by centrifugation (20 kRPM, 33.2 kG at $r_{average}$, 45 min).

The supernatant was purified using Ni-NTA resin (50 mL of 50 mM tris, 100 mM NaCl, 5 mM imidazole binding buffer and 20 mL of 50 mM tris, 100 mM NaCl, 250 mM imidazole elution buffer). The fractions containing the desired protein (visualised by SDS-PAGE analysis), were concentrated to 4 mL using a viva spin concentrator (10 k cutoff) then immediately passed through a gel filtration column (s200, 36/60) in 50 mM tris, 100 mM NaCl pH 7.5 buffer. The fractions containing the desired protein (visualised by SDS-PAGE analysis) were pooled, aliquoted into 1 mL samples then flash frozen in liquid nitrogen before being stored at -80 °C

1.4.6.2 Family 84 human *O*-GlcNAcase (hOGA) Expression



Adapted from Vocadlo²¹⁶

Plasmids (pET28a(+)) containing the hOGA gene were transformed into BL-21(DE3) *E. coli* cells and plated on Kanamycin/LB/agar plates. Single colonies were picked and 4 x 5 mL starter cultures (LB/Kanamycin) were grown overnight at 37 °C 200 RPM then transferred into 4 x 1.0 L of media (LB/Kanamycin) in 2 L baffled conical flasks. The cultures were grown at 37 °C until an optical density of 0.6 – 0.8 was achieved (all between 40 – 70 min) at which point the cells were induced by the addition of IPTG (to give a final concentration of 3 mM) and incubated for a further 16 h at 25 °C. The cells were then pelleted for 10 minutes at 8 kRPM 9.6 kG at $r_{average}$.

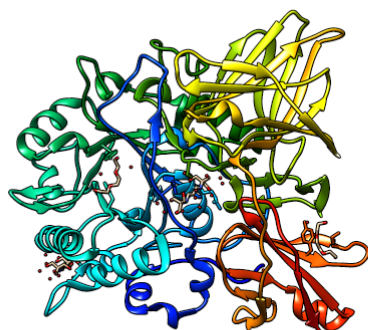
Cell pellets were resuspended in buffer (50 mL of 50 mM tris, 100 mM NaCl, 10 mM imidazole and 1 mg mL⁻¹ lysozyme) and stirred on ice for 1 h. The pellets were then subjected to sonication (50 % power, 30 s sonication 1 min rest, four times), with the resultant mixtures treated by centrifugation (20 krpm, 33.2 kG at $r_{average}$, 45 min).

The supernatant was purified using Ni-NTA resin and then (50 mL of 50 mM tris, 100 mM NaCl, 5 mM imidazole binding buffer and 20 mL of 50 mM tris, 100 mM NaCl,

250 mM imidazole elution buffer). The fractions containing the desired protein were concentrated to 4 mL using a vivaspin concentrator (10 k cutoff) then immediately passed through a gel filtration column (s200, 36/60) in PBS buffer. The fractions containing the desired protein (visualised by SDS-PAGE analysis) were pooled, aliquoted into 1 mL samples then flash frozen in liquid nitrogen before being stored at -80 °C

1.4.6.3 Endo- β -*N*-acetylglucosaminidase from *Arthrobacter protophormiae*

Y299F (EndoA Y299F) Expression



Adapted from Davis¹⁶³

Plasmids (pGEX-2T) containing the EndoA Y299F gene were transformed into BL-21(DE3) *E. coli* cells and plated on ampicillin/LB/agar plates. Single colonies were picked and 2 x 5 mL starter cultures (LB/Ampicillin) were grown overnight at 37 °C 200 RPM then transferred into 2 x 1.0 L of media (LB/ampicillin) in 2 L baffled conical flasks. The cultures were grown at 37 °C until an optical density of 0.6 – 0.8 was achieved (all between 40 – 70 min) at which point the cells were induced by the addition of IPTG (to give a final concentration of 3 mM) and incubated for a further 4 h at 37 °C 200 RPM. The cells were then pelleted for 10 minutes at 8 krpm 9.6 kG at r_{average} .

Cell pellets were resuspended in buffer (40 mL of PBS, 2 complete protease inhibitor tablet and 1 mg mL⁻¹ lysozyme) and stirred on ice for 2 h. The pellets were then subjected to sonication (50 % power, 30 s sonication 1 min rest, four times), with the resultant mixtures treated by centrifugation (20 krpm, 33.2 kG at r_{average} , 45 min). The

supernatant filtered (0.2 μm) then purified using FPLC (Buffer A: PBS Buffer B: 30 mM Tris HCl, 10 mM reduced L-glutathione pH 8.0 5 % B for 2 CV to 100 % B over 2 CV at 1 mL/min) with a GSTrap HP column (2 x 5 mL). The fractions containing the desired protein (visualised by SDS-PAGE analysis) were pooled, aliquoted into 1 mL samples then flash frozen in liquid nitrogen before being stored at -80 °C.

1.4.6.4 Histone H3 K27C



Adapted from Chromatin Protocols (Methods in Molecular Biology)²¹⁷

Plasmids (pET3d) containing the histone H3 K27C gene were transformed into BL-21(DE3) pLys *E. coli* cells and plated on ampicillin/chloramphenicol/LB/agar plates. Single colonies were picked and two 5 mL starter cultures (LB/ampicillin/chloramphenicol) were grown overnight at 37 °C 200 RPM then transferred into 2 x 500 mL of media (2xYT/ ampicillin) in 2 L baffled conical flasks. The cultures were grown at 37 °C until an $OD_{600} = 0.6 - 0.8$ was achieved (typically between 100 – 140 min) at which point the cells were induced with the addition of IPTG to reach a final concentration of 0.3 mM and incubated for a further 3 h at 37 °C. The cells were then pelleted for 10 minutes at 8 kRPM 9.6 kG at $r_{average}$ then stored at -80 °C overnight.

Cell pellets were thawed then resuspended in buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, 1 Roche cOmplete EDTA free protease inhibitor tablet, 1 mg mL⁻¹ lysozyme, 0.1 mg mL⁻¹ DNase, pH 7.5) and stirred on ice for 30 min. The suspension was subjected to sonication (50 % Amplitude, 30 s sonication 1 min

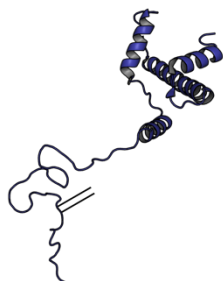
rest, five times) followed by centrifugation (20 kRPM, 33.2 kG at r_{average} 30 mins). The supernatant was discarded, and the pellet suspended in TW buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, 1 % Triton X, pH 7.5) using an IKA Ultra-turrax T-8 and the suspension was subjected to sonication (50 % power, 30 s sonication 1 min rest, two times) followed by centrifugation (20 krpm, 33.2 kG at r_{average} , 20 mins). The pellet was washed twice more using TW buffer and once with wash buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, pH 7.5). The pellet was treated with DMSO (500 μ L) and allowed to stand for 10 mins. Unfolding buffer (Gdm.HCl 7M, Tris 10 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM, pH 7.5, 4 mL) was applied to the pellet and the suspension vigorously shaken (400 RPM) for 1 h. The suspension was subject to centrifugation (20 kRPM, 33.2 kG at r_{average} , 20 mins) and the supernatant filtered sequentially through a 1.2 μm^2 syringe filter and 0.5 μm^2 syringe filter.

The filtered supernatant was purified by gel filtration (s200 26/60) using SAU-100 (Urea 7 M, NaCl 100 mM, NaOAc 20 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM pH 5.2) the fractions containing the desired protein (visualised by SDS-PAGE analysis) were pooled and subject to a 2nd purification by strong cation exchange chromatography (using a GE Healthcare HiTrap SP HP 5 mL column). The sample was loaded in SAU-100 buffer then a linear gradient 70 % SAU-1000 buffer (Urea 7 M, NaCl 1 M, NaOAc 20 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM pH 5.2) was ran over 20 CV. The fractions containing the desired protein (visualised by

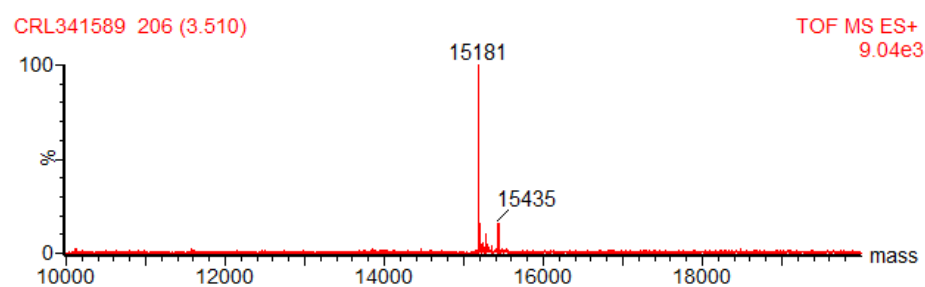
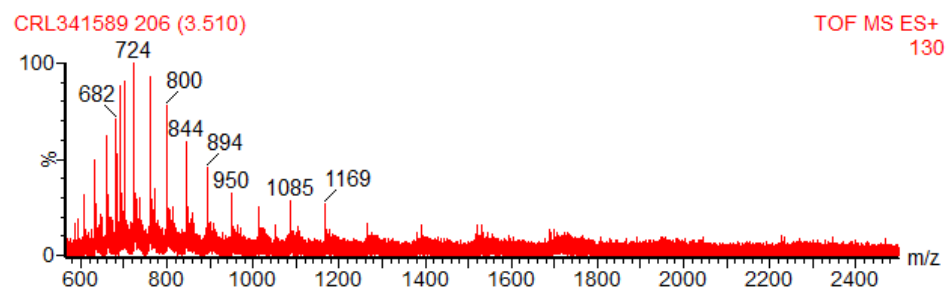
SDS-PAGE analysis) were pooled, dialysed against aqueous BME (2 mM, 3 x 4 L, at least 4 hours for each step), then lyophilised and stored at -20 °C.

1.4.7.0 Chemical Reactions of Proteins

1.4.7.1 Histone H3 K27Dha



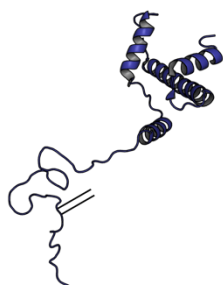
To a solution of Histone H3 K27C (10 mg) in Gdm.Cl (3 M) and $\text{Na}_2\text{H}_3\text{PO}_4$ (100 mM) buffer (pH 8, 500 μL), DTT (30 mg) was added. The reaction mixture was shaken for 30 mins at room temperature then passed through a G-25 PD10 minitrapp desalting column equilibrated with the above buffer and eluted with 1000 μL of the above buffer. To a sample of the eluent (900 μL) DHDBA stock solution (100 μL , 164 mg mL^{-1}) was added, the reaction mixture as shaken, for 45 mins at room temperature then 90 mins at 37 °C. The reaction mixture was concentrated to 500 μL using a vivaspin 3 kDa and then passed through a G-25 PD10 minitrapp desalting column pre-equilibrated with NH_4OAc (500 mM) and Gdm.Cl (5 M) buffer (pH 8.0) and eluted with 1000 μL of the above buffer.



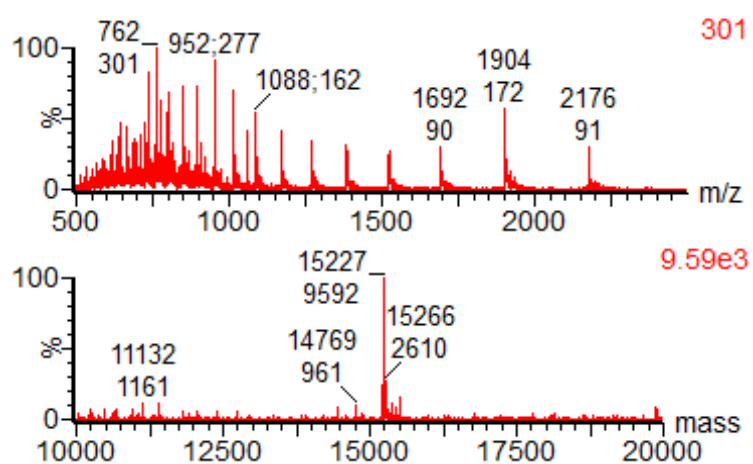
Calculated mass: 15181

Observed mass: 15181

1.4.7.2 Histone H3 S10Dha

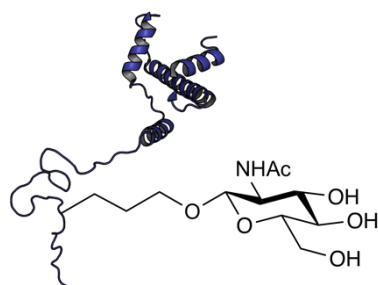


To a solution of Histone H3 S10C (10 mg) in Gdm.Cl (3 M) and $\text{Na}_2\text{H}_3\text{PO}_4$ (100 mM) buffer (pH 8, 500 μL) DTT (30 mg) was added. The reaction mixture was shaken for 30 mins at room temperature then passed through a G-25 PD-10 minitrapp desalting column equilibrated with the above buffer and eluted with 1000 μL of the above buffer. To a sample of the eluent (900 μL) DHDBA stock solution (100 μL , 164 mg mL^{-1}) was added, the reaction mixture as shaken, for 45 mins at room temperature then 90 mins at 37 °C. The reaction mixture was concentrated to 500 μL using a vivaspin 3 kDa and then passed through a G-25 PD10 minitrapp desalting column pre-equilibrated with NH_4OAc (500 mM) and Gdm.Cl (5 M) buffer (pH 8.0) and eluted with 1000 μL of the above buffer.



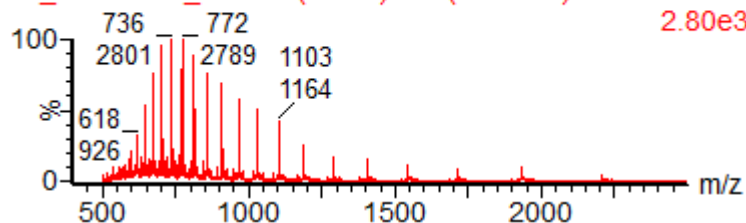
Calculated mass: 15221

Observed mass: 15227

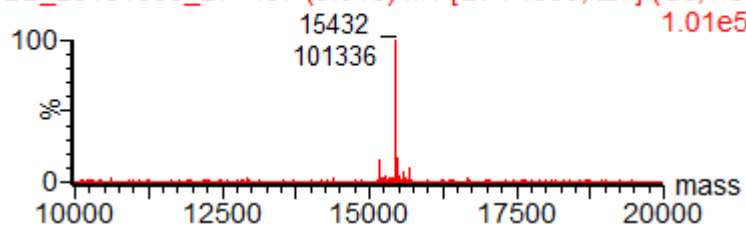
1.4.7.3 H3 K27 hh SGlcNAc

A solution of H3 K27Dha (1 mg mL^{-1} 200 μL) and 1-(2-iodoethoxy)- β -D-GlcNAc (3 mg) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 $^{\circ}\text{C}$ and pipetted directly over NaBH_4 (1 mg). The reaction mixture was shaken at 4 $^{\circ}\text{C}$ for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.

BB_20151003_27 437 (8.016) Cm (427:446) TOF MS ES+ 2.80e3

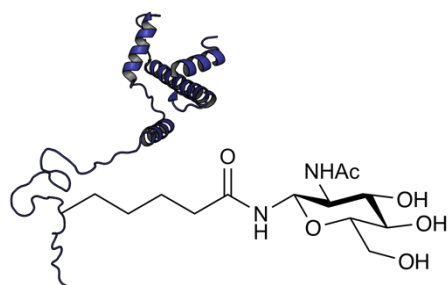


BB_20151003_27 437 (8.016) M1 [Ev-74930,It27] (Gs,1.5) 1.01e5

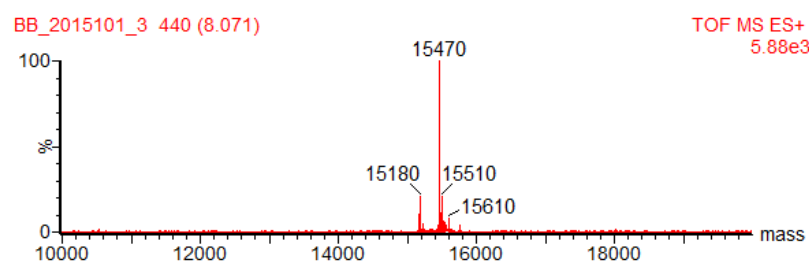
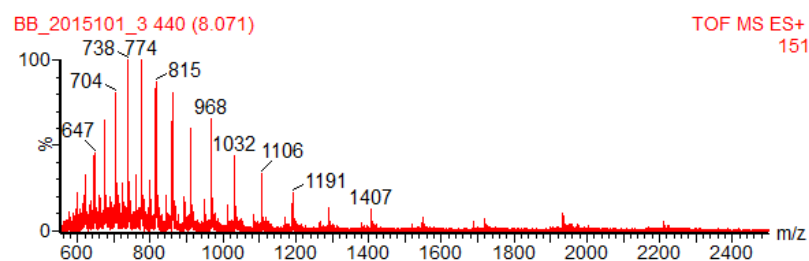


Calculated mass: 15430

Observed mass: 15432

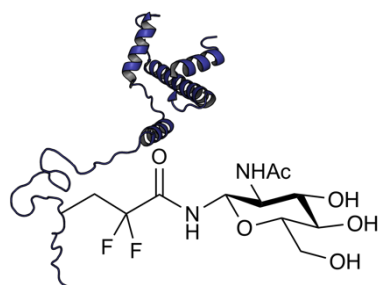
1.4.7.4 H3 K27 hh QGlcNAc

A solution of H3 K27Dha (1 mg mL^{-1} $200 \text{ }\mu\text{L}$) and 1-(4-iodonuttyramido)- β -D-GlcNAc in buffer (5 M Gdm.Cl , $500 \text{ mM NH}_4\text{OAc}$ pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to $4 \text{ }^\circ\text{C}$ and pipetted directly over NaBH_4 (1 mg). The reaction mixture was shaken at $4 \text{ }^\circ\text{C}$ for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.



Calculated mass:15471

Observed mass: 15470

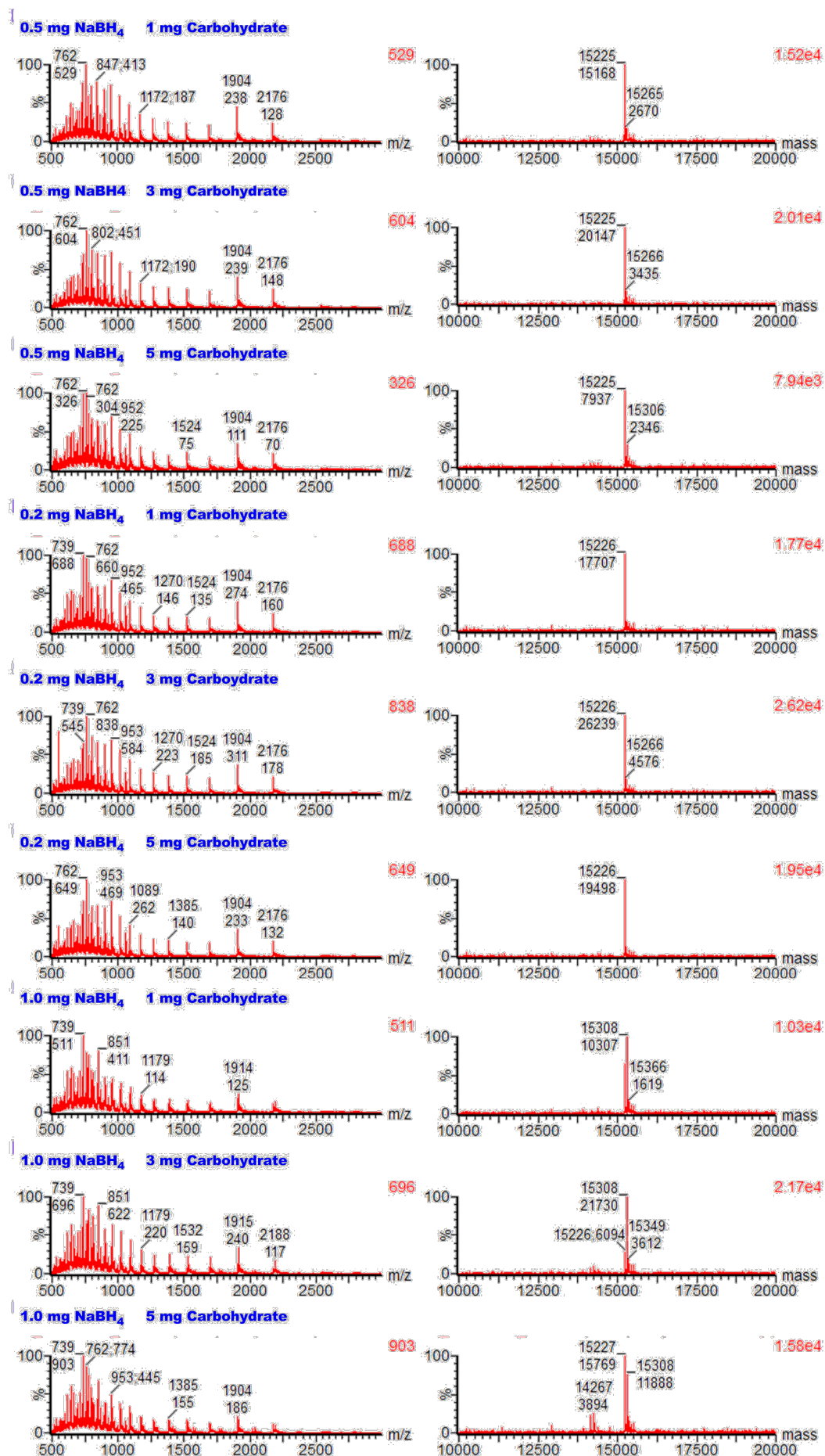
1.4.7.5 H3 S10NF₂GlcNAc

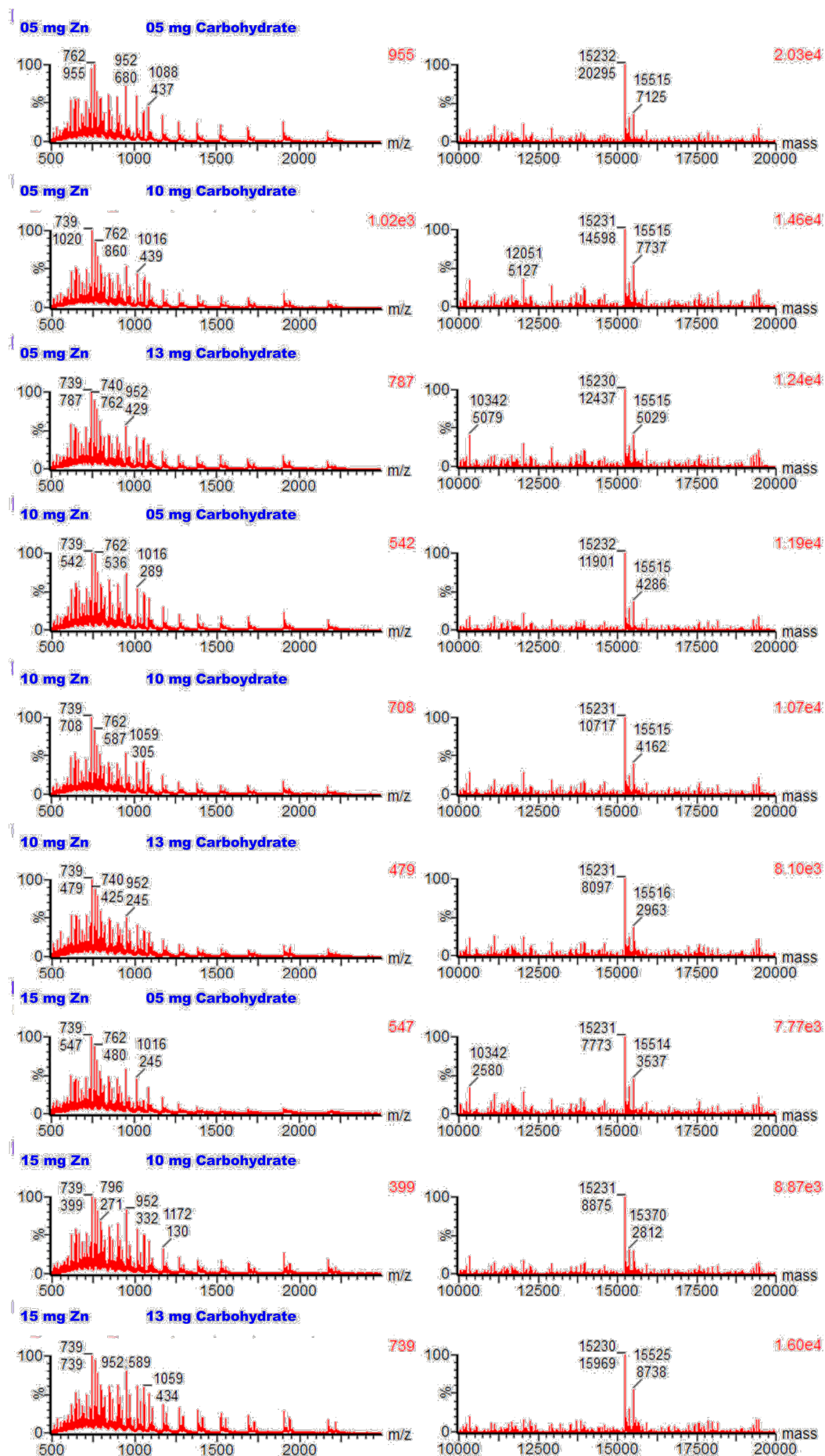
A solution of H3 S10Dha (1 mg mL^{-1} 200 μL) and 1-(2,2-difluoro-2-iodoacetamido)- β -D-GlcNAc (specified amount) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 $^{\circ}\text{C}$ and pipetted directly over the initiator (specified and specified amount) The reaction mixture was shaken at 4 $^{\circ}\text{C}$ for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS

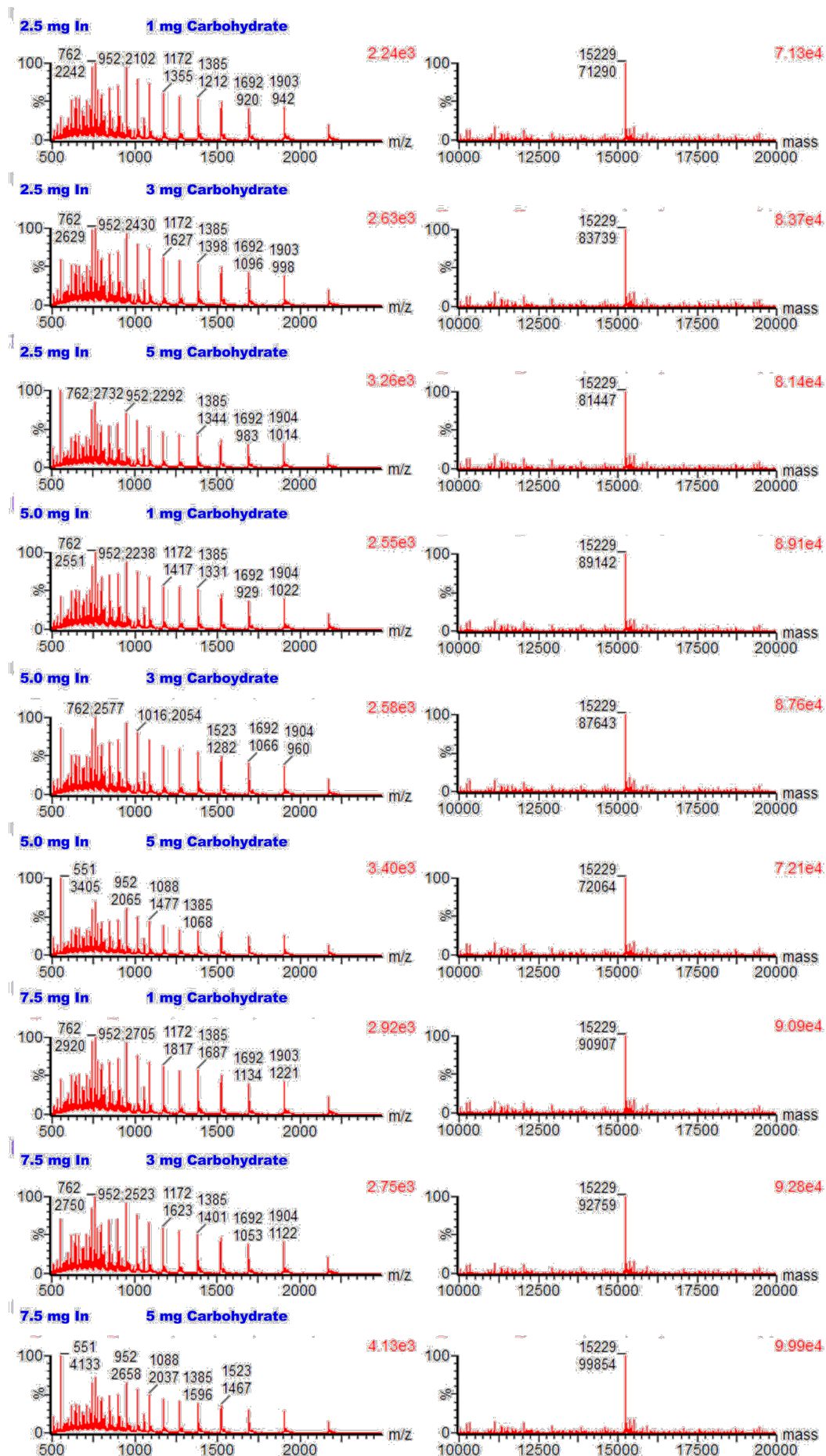
Three sets of MS data are given on the following pages corresponding to reactions tried with the indicated amounts of carbohydrate and initiator: NaBH_4 initiation (Page 165), Zn initiation (Page 166) and In initiation (Page 167).

Calculated mass: 15521

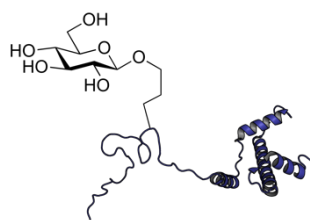
Observed mass: ~ 15521 (Zn+In – although never exclusively) + 15308 (NaBH_4)







1.4.7.6 H3 K27OGlc



A solution of H3 K27Dha (1 mg mL^{-1} 200 μL) and 1-(2-iodoethoxy)- β -D-Glucose (specified amount) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 $^\circ\text{C}$ and pipetted directly over NaBH_4 (1 mg). The reaction mixture was shaken at 4 $^\circ\text{C}$ for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS

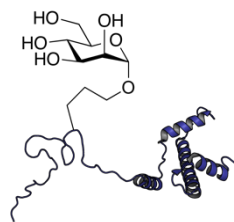
MS data is given on the following page corresponding to reactions tried with the indicated amount of carbohydrate.

Calculated mass: 15389

Observed mass: 15389



1.4.7.7 H3 K27OMAn

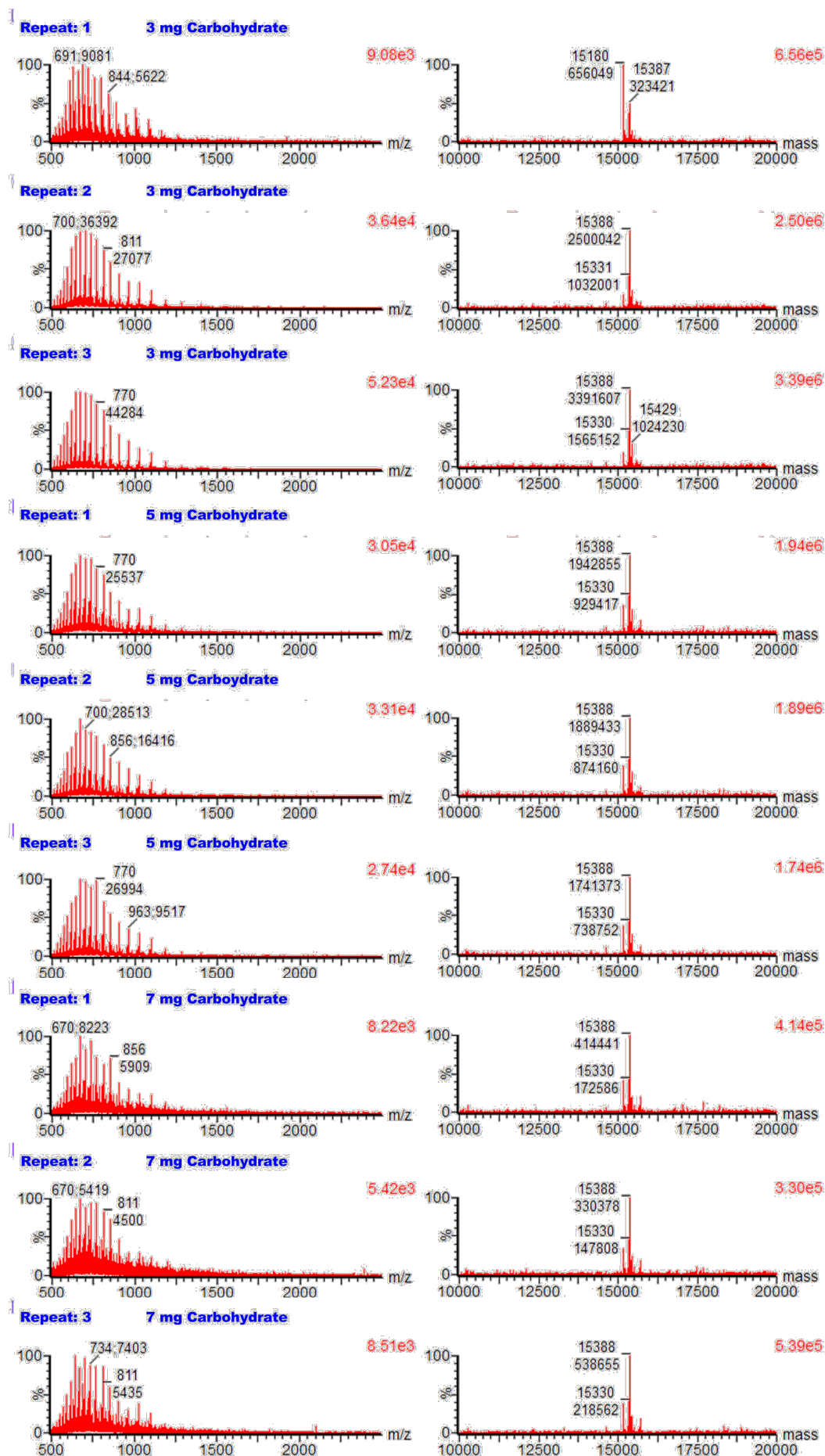


A solution of H3 K27Dha (1 mg mL^{-1} 200 μL) and 1-(2-iodoethoxy)- α -D-mannose (specified amount) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 $^\circ\text{C}$ and pipetted directly over NaBH_4 . The reaction mixture was shaken at 4 $^\circ\text{C}$ for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.

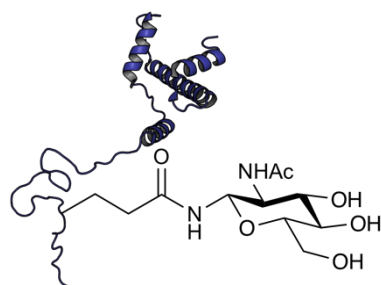
MS data is given on the following page corresponding to reactions tried with the indicated amount of carbohydrate.

Calculated mass: 15389

Observed mass: 15388



1.4.7.8 H3 S10NGlcNAc



A solution of H3 S10Dha (1 mg mL^{-1} 200 μL) and 1-(2-iodoacetamido)- α -D-GlcNAc (specified amount) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 $^{\circ}\text{C}$ and pipetted directly over NaBH_4 (specified amount). The reaction mixture was shaken at 4 $^{\circ}\text{C}$ for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.

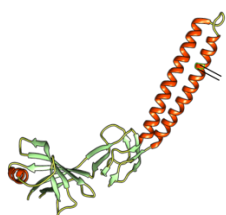
MS data is given on the following page corresponding to reactions tried with the indicated amounts of carbohydrate and initiator

Calculated mass: 15483

Observed mass: 15225 (no reaction)



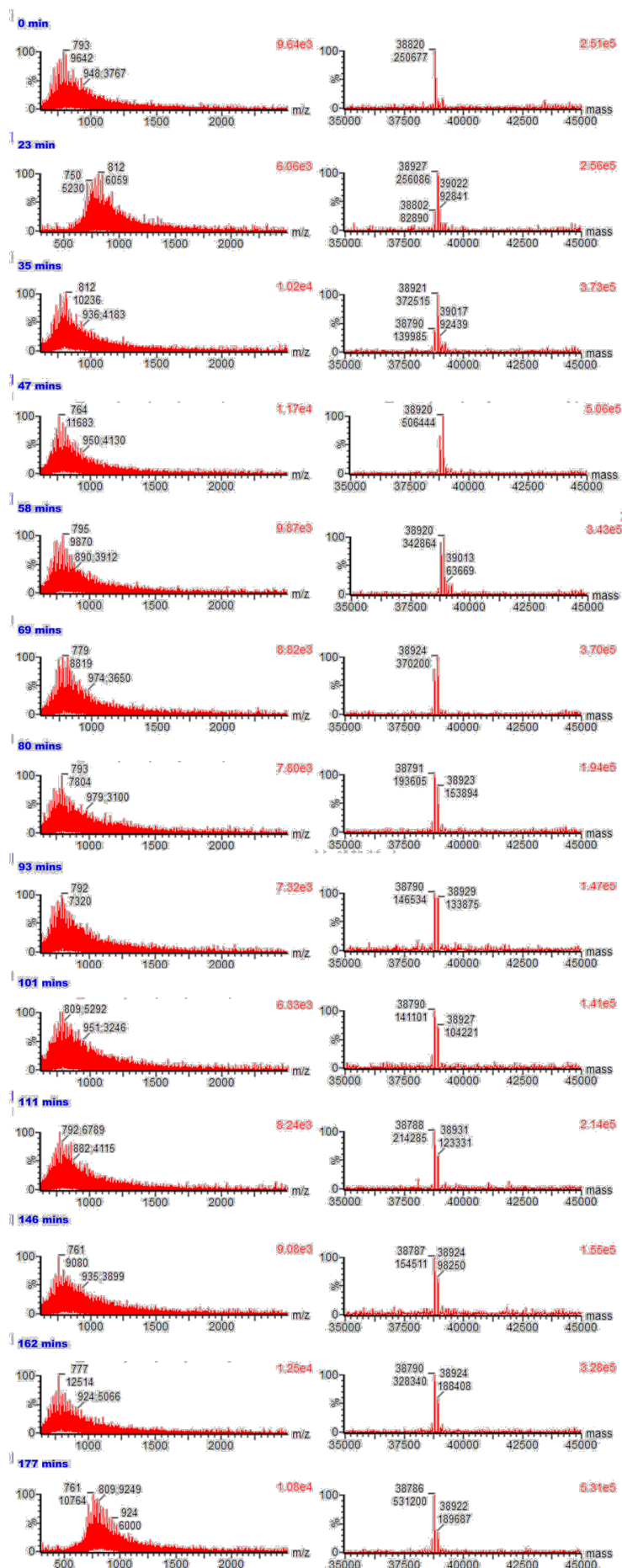
1.4.7.9 AcrA N123Dha



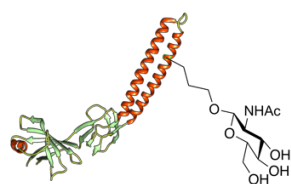
A solution AcrA N123C (1.00 ml, 1.0 mg mL in PBS with 5 mM BME and 1 mM Benzamidine HCl) was treated with DTT (30 mg) then gently pipetted up and down until homogenous, then stood for 30 mins at room temperature. The buffer was exchanged by using G-25 Minitraps (2 x 500 μ L of sample elute with 1 mL of buffer) preequilibrated in PBS. A solution of MDBP in MeCN (5 μ L MDBP is added to 45 μ L of MeCN) was added to each sample of freshly reduced AcrA N123C (now 2 x 1.00 mL of 0.50 mg mL⁻¹ in PBS) and the reaction mixtures stood in a 37 °C incubator for 300 minutes. The buffer was exchanged by G-25 Minitraps (4 x add 500 μ L of sample elute with 1 mL of buffer) preequilibrated in PBS. Analytics were then performed. For the recording of a CD spectrum the AcrA Dha was buffer exchanged directly into 100 mM NaF 50 mM NaP_i pH 7.40. The AcrA Dha must be used as soon as possible. Avoid keeping longer than strictly necessary. Do not freeze. MS data is given on the following pages corresponding to the indicated time point.

Calculated mass: 38788

Observed mass: 38786



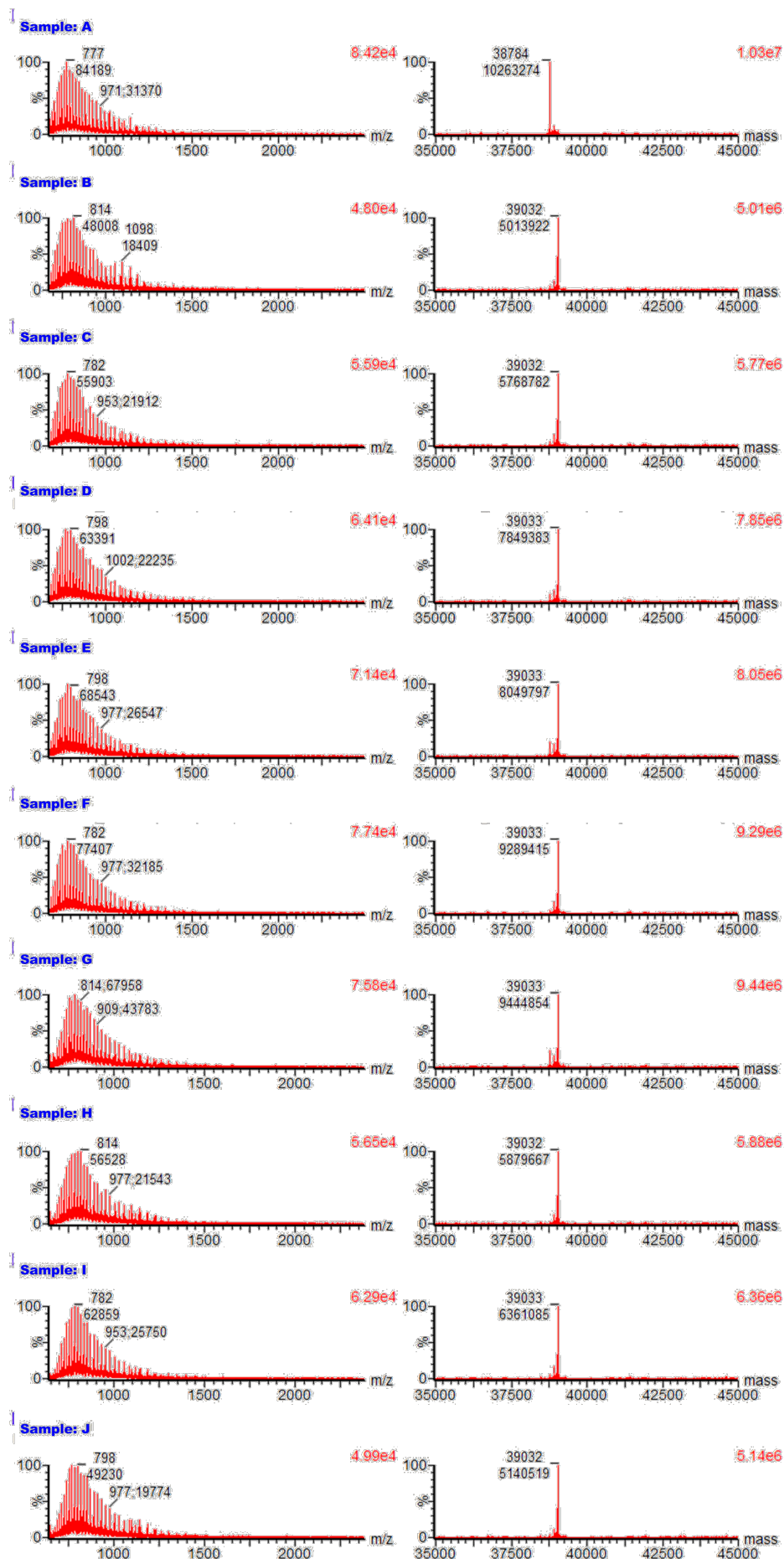
1.4.7.10 AcrA N123 O-GlcNAc



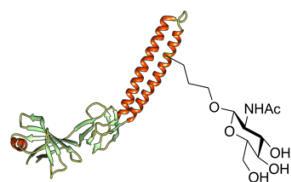
Samples of a PBS solution of AcrA 123Dha (10 x 200 μL , $\sim 250 \mu\text{L mL}^{-1}$, the preparation is previously described in this thesis in section 1.4.7.7) in microcentrifuge tubes (1.5 mL) was ported into the glovebox ($[\text{O}_2] = 3.4 \text{ ppm}$ at the time of porting) and stood for 9 h. 3 mg of the *O*-GlcNAc reagent is dissolved in the protein solution. The protein solution was pipetted directly over a single crystal of NaBH_4 (mass no more than 550 μg and no less than 450 μg) in a different microcentrifuge tube (1.5 mL). The reaction mixture was moved immediately to a 4 $^\circ\text{C}$ cooling block and stood for 30 mins. The microcentrifuge tube was sealed and removed from the glovebox, 5 μL of neat formic acid is added and the solution analysed by LCMS, the sample is placed in a 4 $^\circ\text{C}$ refrigerator in order to analyse by SDS-PAGE when all samples are complete. Nine of ten reactions showed complete or almost complete conversion to the desired product, MS data is given on the following page.

Calculated mass: 39032

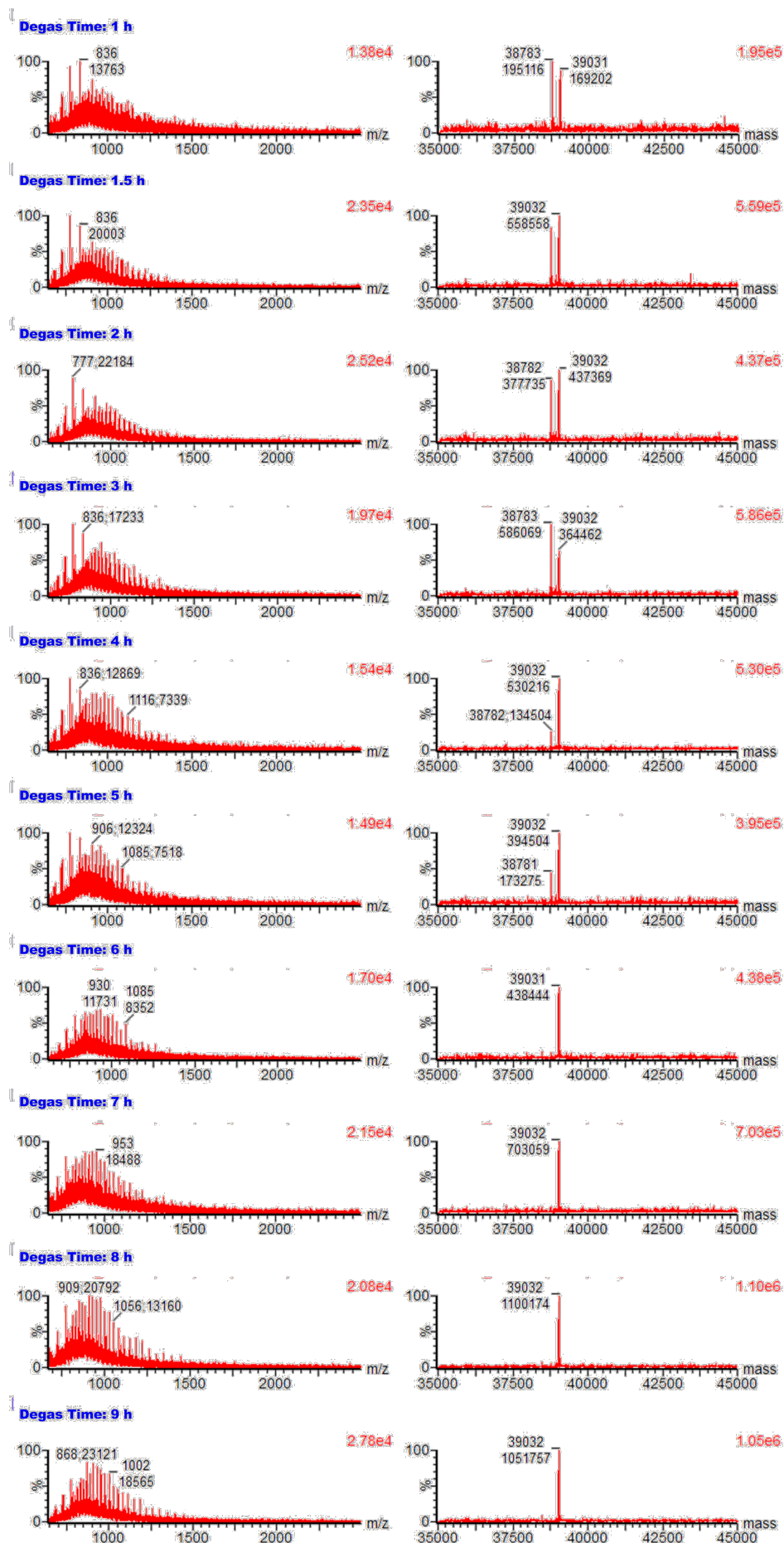
Observed mass: 39032



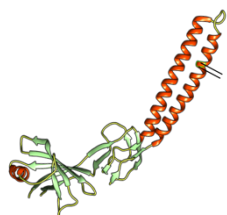
1.4.7.11 AcrA N123 OGlcNAc Degas Time Screen



Samples of a PBS solution AcrA 123Dha (10 x 200 μL , $\sim 250 \mu\text{L mL}^{-1}$, the preparation is previously described in this thesis in section 1.4.7.7) in microcentrifuge tubes (1.5 mL) was ported into the glovebox ($[\text{O}_2] = 3.4 \text{ ppm}$ at the time of porting) and stood, a stop watch was started. When the stopwatch meets the indicated time 3 mg of the O-GlcNAc reagent was dissolved in the protein solution. The protein solution was pipetted directly over a single crystal of NaBH_4 (mass no more than 550 μg and no less than 450 μg) in a different microcentrifuge tube (1.5 mL). The reaction mixture was moved immediately to a 4 $^\circ\text{C}$ cooling block and stood for 30 mins. The microcentrifuge tube is sealed and removed from the glovebox, 5 μL of neat formic acid was added and the solution analysed by LCMS, the sample was placed in a 4 $^\circ\text{C}$ refrigerator in order to analyse by SDS-PAGE when all samples are complete. MS data is given on the following pages corresponding to the indicated time point.

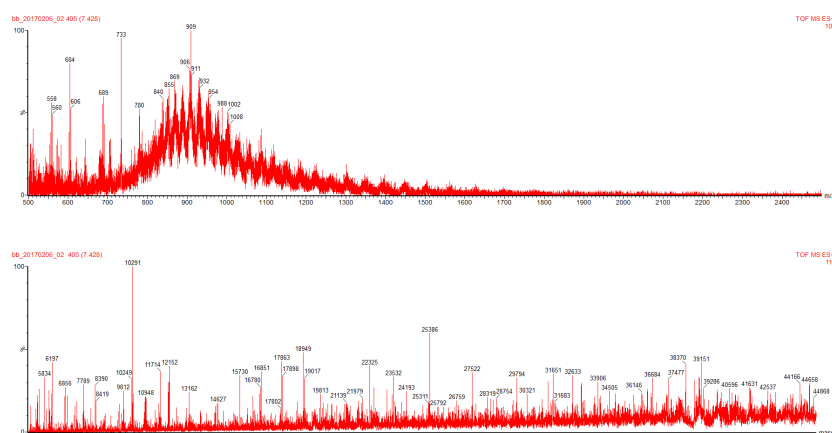


1.4.7.12 AcrA N123Dha with DBHDA – Failed

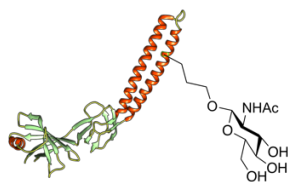


A solution AcrA N123C (1.00 ml, 1.0 mg mL in PBS with 5 mM BME and 1 mM Benzamidine HCl) was treated with DTT (30 mg) and gently pipetted up and down until homogenous, then stood for 30 mins at room temperature. The buffer was exchanged by using G-25 Minitraps (2 x add 500 μ L of sample elute with 1 mL of buffer) preequilibrated in PBS. A solution of DBHDA in DMSO (100 μ L 164 mg mL⁻¹) was added to each sample of freshly reduced AcrA N123C (now 2 x 1.00 mL of 0.50 mg mL⁻¹ in PBS) and the reaction mixtures stood at room temperature with samples being periodically analysed by LCMS.

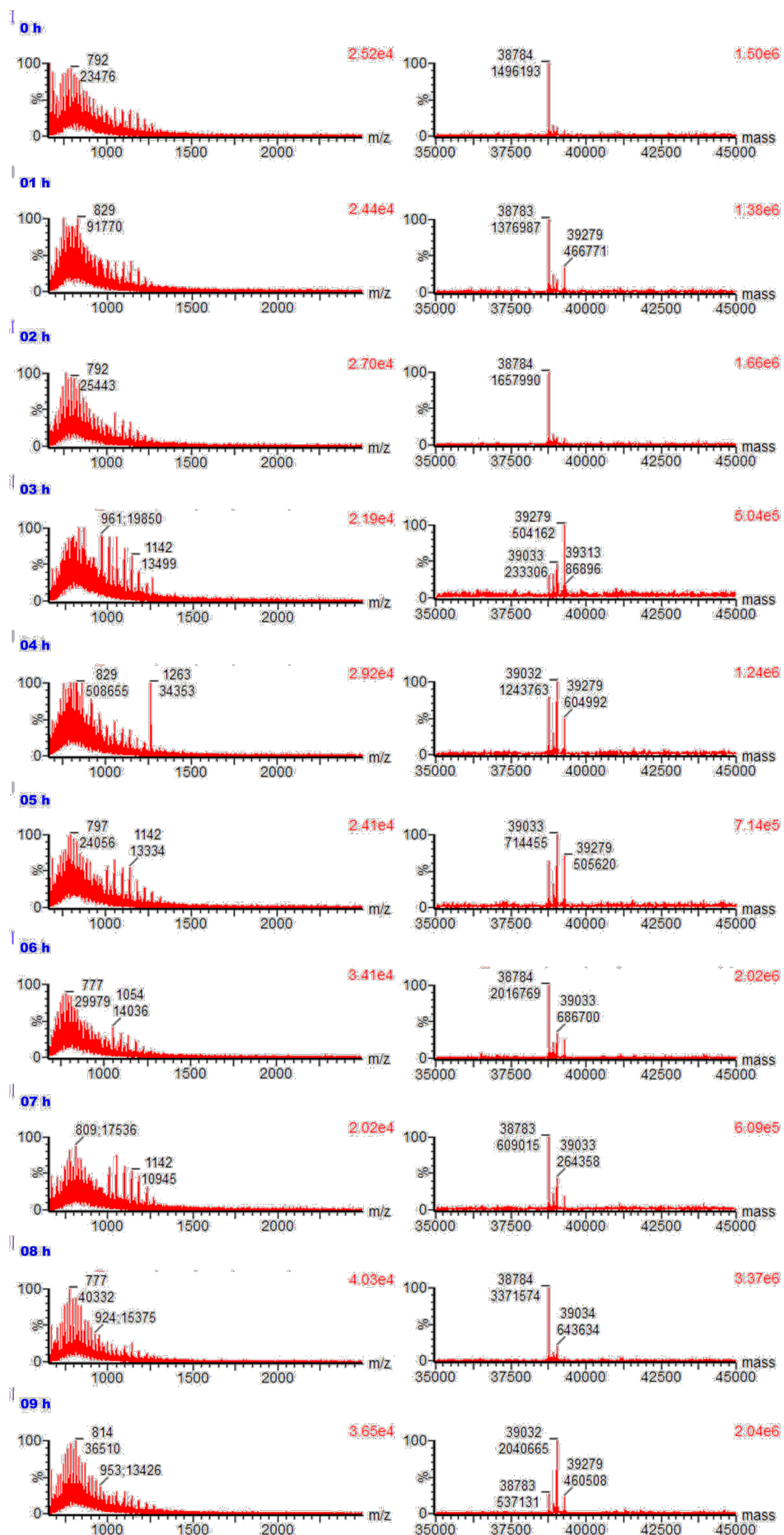
MS After 2.5 h:



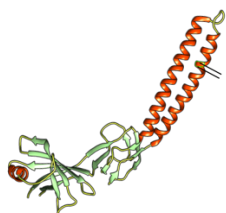
1.4.7.13 AcrA N123OGlcNAc Zn Degas Time Screen



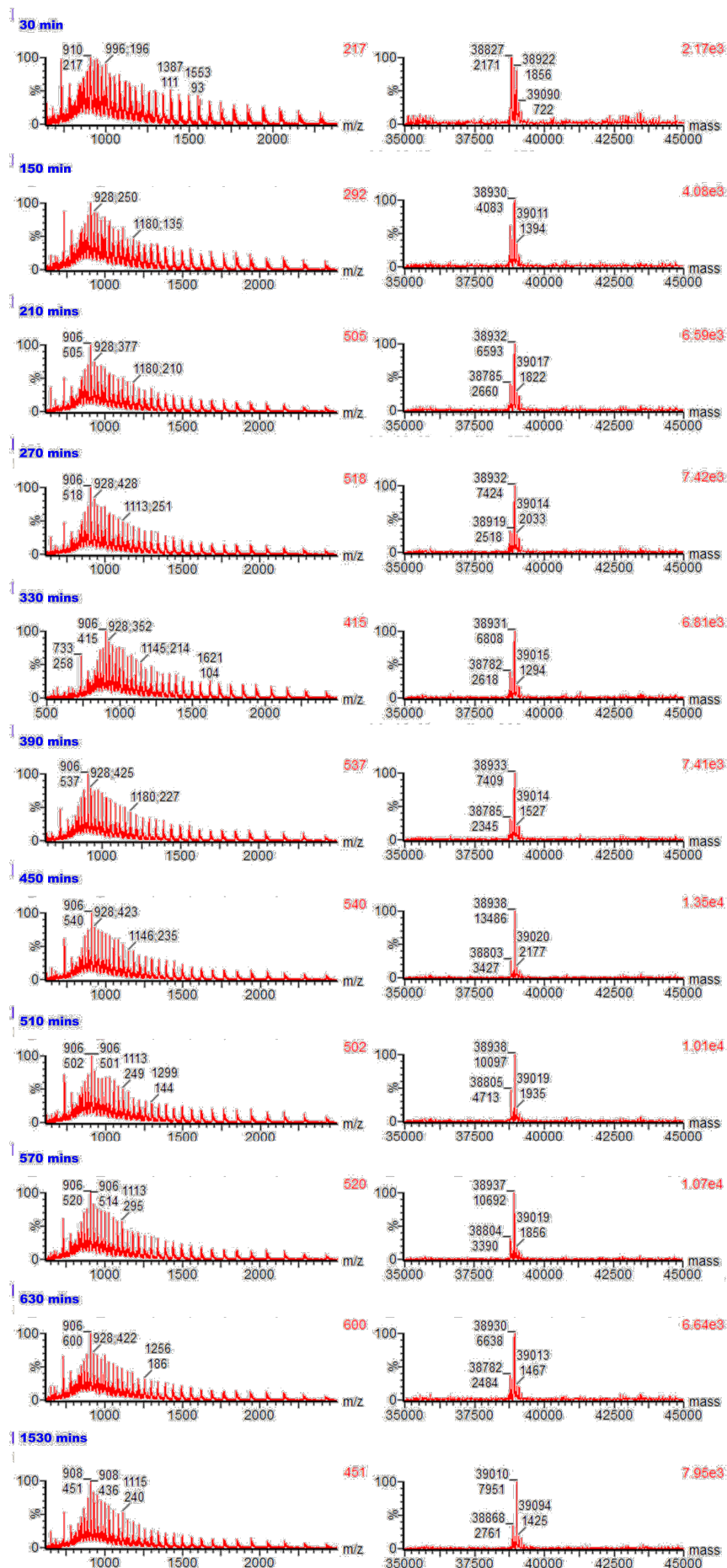
Samples of a PBS solution AcrA 123Dha (10 x 200 μL , $\sim 250 \mu\text{L mL}^{-1}$, the preparation is previously described in this thesis section 1.4.7.7) in microcentrifuge tubes (1.5 mL) was ported into the glovebox ($[\text{O}_2] = 3.4 \text{ ppm}$ at the time of porting) and stood, a stop watch is started. When the stopwatch meets the indicated time 15 mg of the *O*-GlcNAc reagent was dissolved in the protein solution. The protein solution is pipetted directly over a Zn (20 mg) in a different microcentrifuge tube (1.5 mL). The reaction mixture was moved immediately to a 4 $^{\circ}\text{C}$ cooling block and shaken for 30 mins. The microcentrifuge tube was sealed and removed from the glovebox, 5 μL of neat formic acid was added and the solution analysed by LCMS, the sample was placed in a 4 $^{\circ}\text{C}$ refrigerator in order to analyse by SDS-PAGE when all samples are complete. The protein ion series and mass spectra are significant 'messier' than with NaBH_4 initiations, this is likely a result of shaking the samples and the heterogenous reaction mixture. MS data is given on the following pages corresponding to the indicated time point.



1.4.7.14 AcrA N123Dha - Failed



A solution AcrA N123C (1.00 ml, 1.0 mg mL in PBS with 5 mM BME and 1 mM Benzamidine HCl) was treated with DTT (30 mg) and gently pipetted up and down until homogenous, then stood for 30 mins at room temperature. The buffer was exchanged by using G-25 Minitraps (2 x add 500 μ L of sample elute with 1 mL of buffer) preequilibrated in PBS. A solution of MDBP in MeCN (5 μ L MDBP is added to 45 μ L of MeCN) was added to each sample of freshly reduced AcrA N123C (now 2 x 1.00 mL of 0.50 mg mL⁻¹ in PBS) and the reaction mixtures stood at room temperature with samples being periodically analysed by mass spec. MS data is given on the following pages corresponding to the indicated time point.



1.4.8.0 Glyco-engineering of histone H3

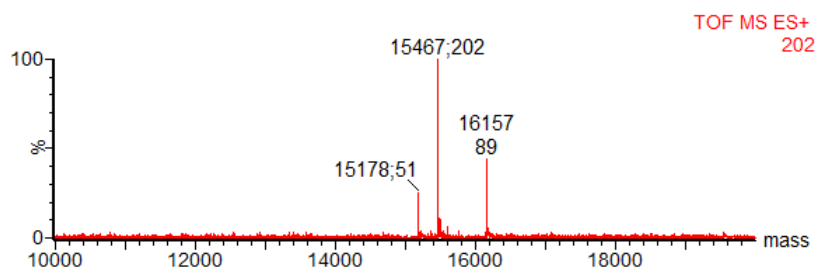
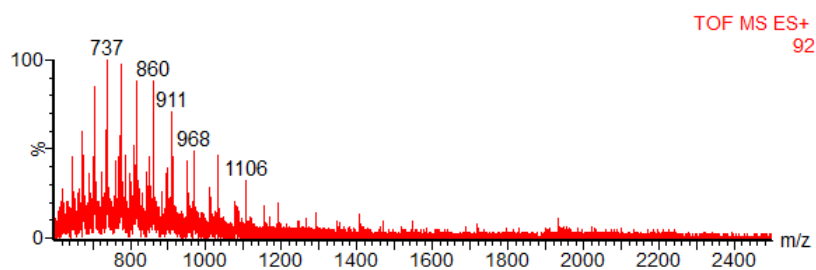
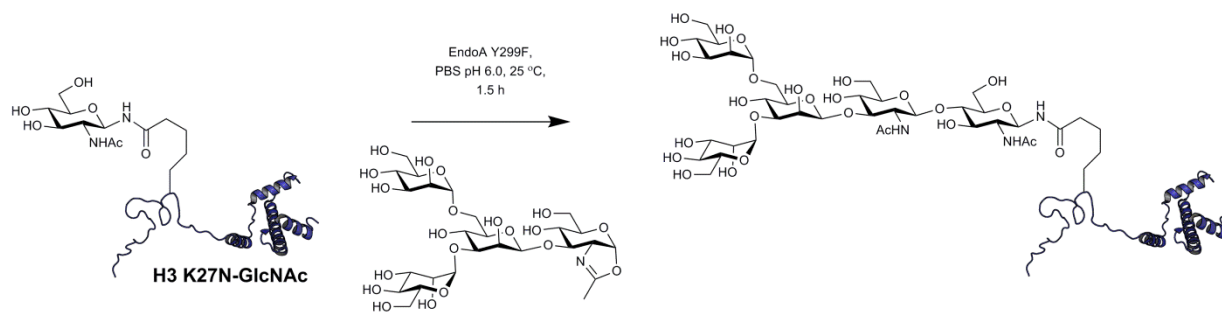
Many thanks to Tom Parsons who kindly donated the carbohydrate oxazoline, the procedure for its synthesis was described by Davis¹⁶³

1.4.8.1 EndoA

Glycosylated H3 (10.0 μL , 1.0 mg mL^{-1}), EndoA Y299F (0.5 μL , 0.5 mg mL^{-1}) and the tetrasaccharide oxazoline 1 (3.6 μL , 5 mg mL^{-1}) were combined in PBS pH 6.0 and shaken gently at 25 °C for two 30 minute periods. After each period an aliquot of the tetrasaccharide oxazoline (3.6 μL , 5 mg mL^{-1}) was added, the reaction mixture was then shaken for a further 30 minutes.

Method adapted from Wang¹⁷¹

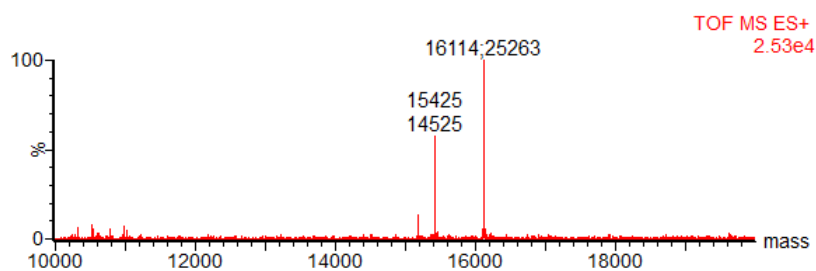
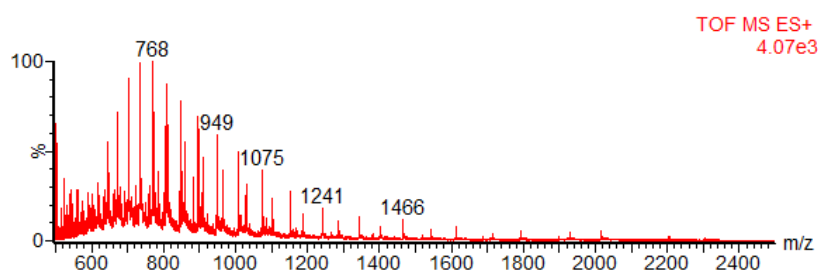
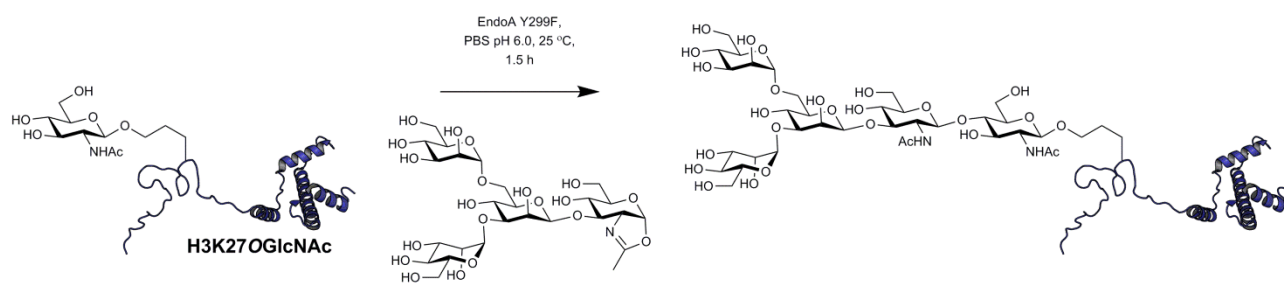
1.4.8.1.1 N-Linked



Calculated mass (H3 K27OMan₃GlcNAc₂): 16160

Observed mass: 16157 and 15467 (unreacted starting material)

1.4.8.1.2 O-Linked



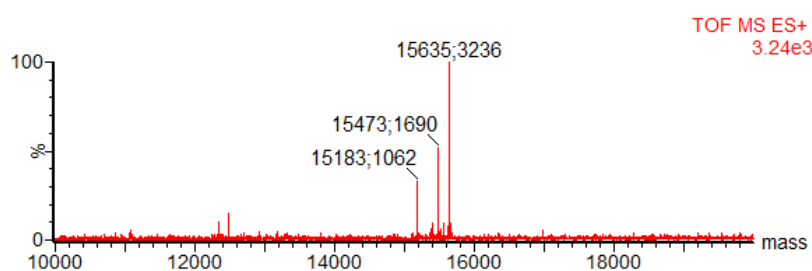
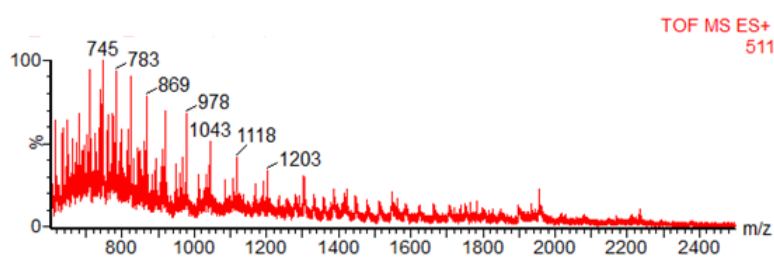
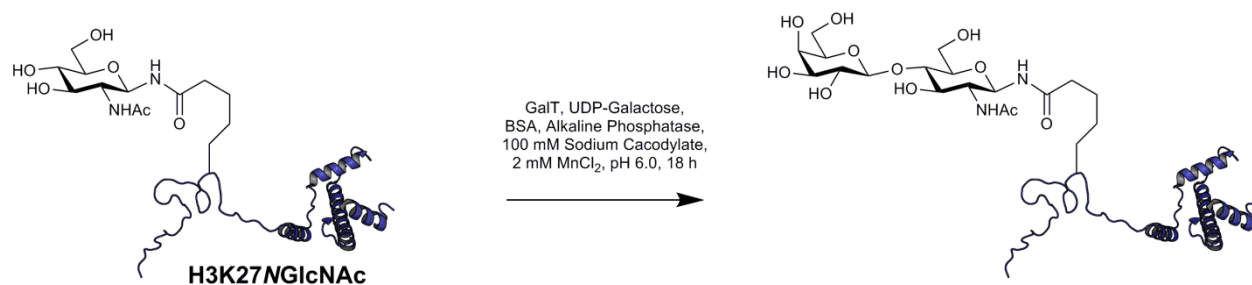
Calculated mass (H3 K27OMan₃GlcNAc₂): 16119

Observed mass: 16114 and 15425 (unreacted starting material)

1.4.8.2 GalT

Glycosylated H3 (20 μL , 1.0 mg mL^{-1}), β -(1,4)-Galactosyltransferase (1 U mL^{-1} , 6 μL), bovine alkaline phosphatase ($>100 \text{ U mL}^{-1}$, 2 μL), bovine serum albumin (2.1 μL) and UDP-galactose (60 mg mL^{-1} , 1 μL) were combined in sodium cacodylate (100 mM), manganese (II) chloride (2 mM) buffer pH 6.0. The reaction mixture was shaken gently for 9 h at 25 $^{\circ}\text{C}$, at which point more β -(1,4)-Galactosyltransferase (1 U mL^{-1} , 6 μL) and UDP-galactose (60 mg mL^{-1} , 1 μL) was added followed by shaking for a further 9 h.

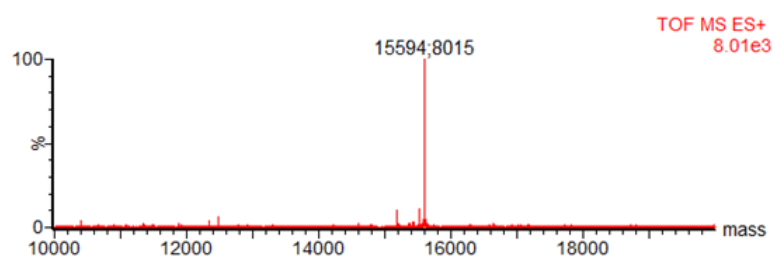
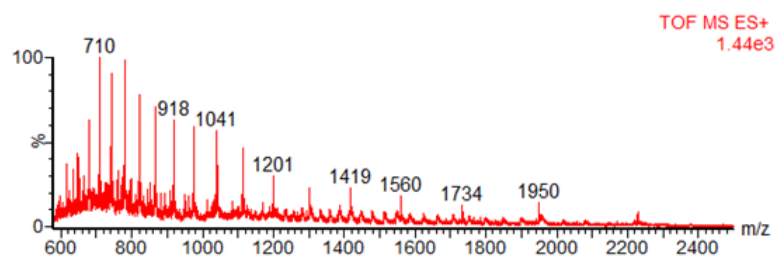
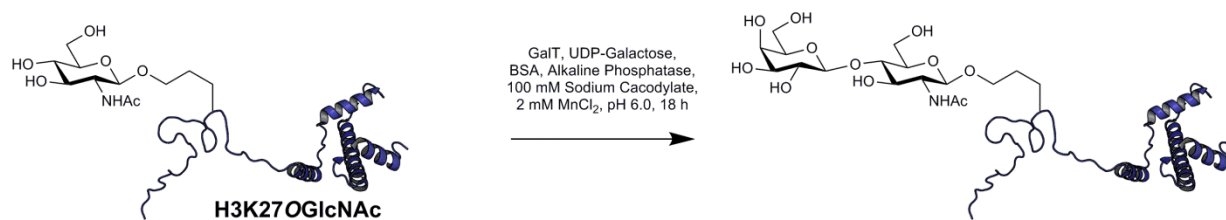
Method adapted from Weiser.²¹⁸

1.4.8.2.1 N-Linked

Calculated mass (H3 N-(beta-1,4-GalGlcNAc)): 15633

Observed mass: 15635 and 15473 (unreacted starting material)

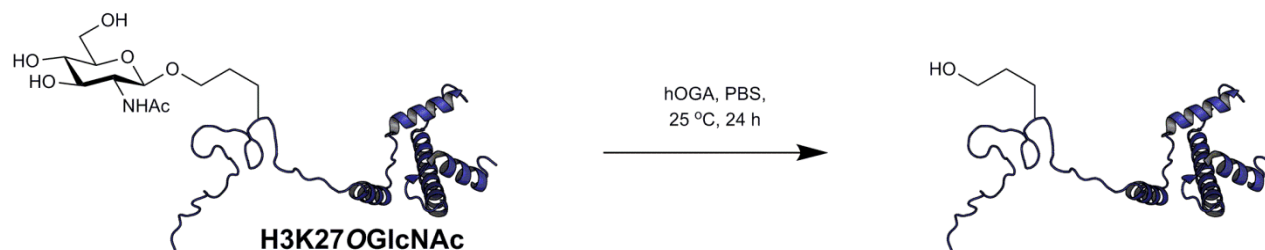
1.4.8.2.2 O-Linked



Calculated mass (H3 K27 O-(β -1,4-GalGlcNAc)): 15592

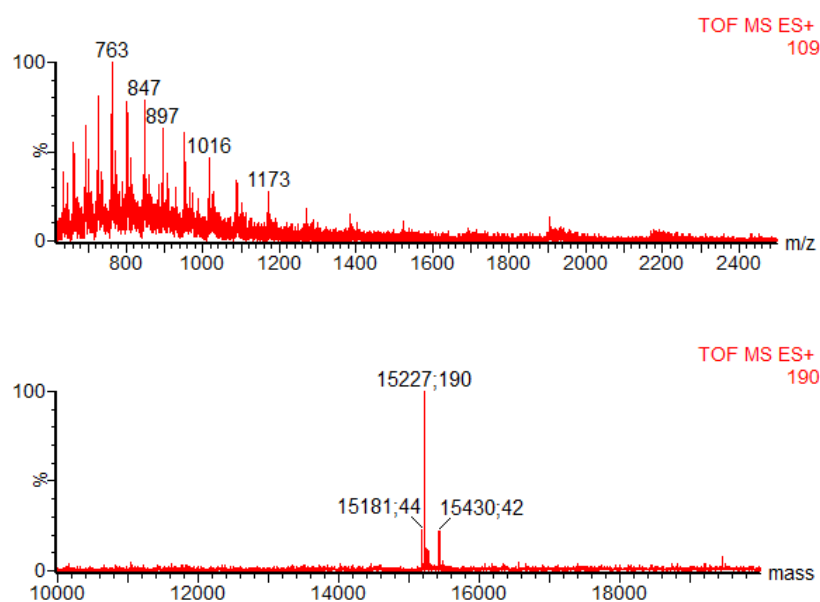
Observed mass: 15594

1.4.8.3 hOGA (human *O*-GlcNAcase)



O-Glycosylated H3 (20.0 μL , 1.0 mg mL^{-1}) was treated with Family 84 hOGA (0.5 μL , 10 mg mL^{-1}) in PBS pH 7.4. The reaction mixture was shaken at 25 °C for 2 h. A second dose of hOGA was added (0.5 μL , 10 mg mL^{-1}) then shaken at 25 °C until completion or no further reaction was observed by mass spectrometry (> 85% cleavage of the sugar).

Method adapted from Vocadlo.²¹⁶

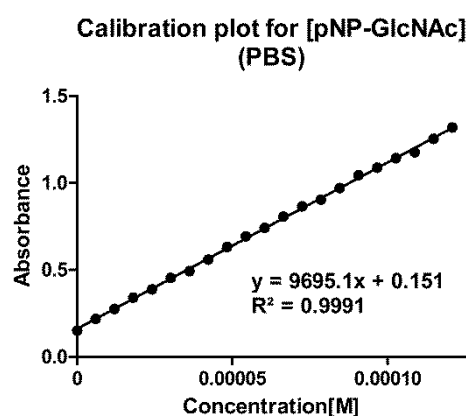


Calculated mass (H3 K27 bis-homoSer): 15227

Observed mass: 15227

pNP-GlcNAc hOGA Test

A calibration curve was set up from precisely prepared samples of pNP in the reaction buffer (PBS), by measuring absorbance at 405 nm.



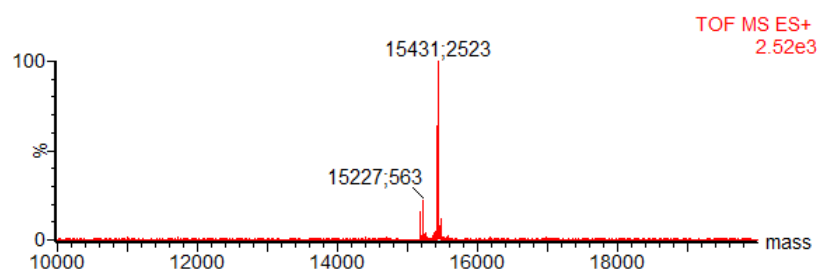
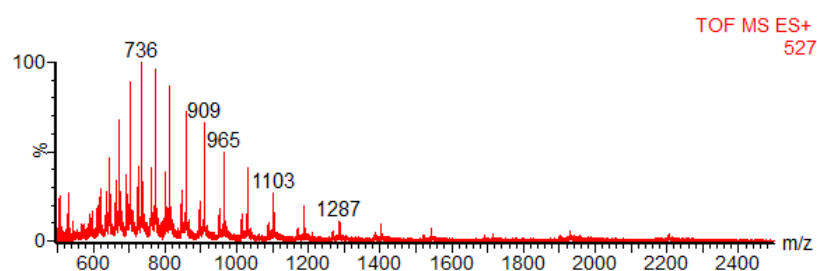
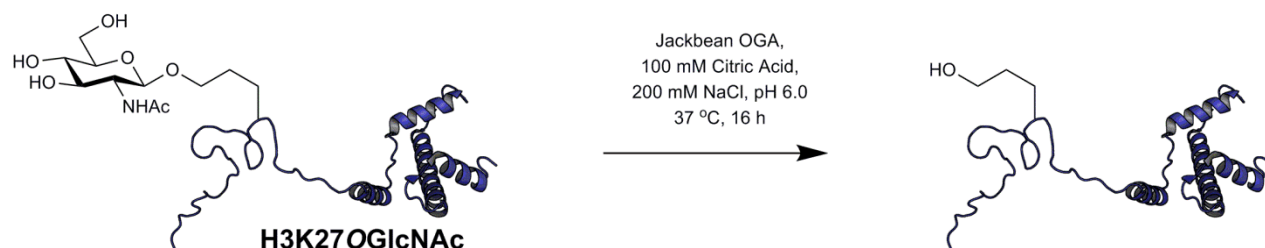
In triplicate hOGA ($2.5 \mu\text{L}$, 10 mg mL^{-1}) in PBS was then added to a carefully prepared stock solution of pNP-GlcNAc ($58.4 \mu\text{M}$, $100 \mu\text{L}$ in PBS) and heated to 25°C for 2 h,

more hOGA was added ($2.5\ \mu\text{L}$, $10\ \text{mg mL}^{-1}$) and the reaction heated to $25\ ^\circ\text{C}$ for 1 h. The absorbance was measured and % conversion calculated (98 %, 93%, 114 %). The reaction was heated to $25\ ^\circ\text{C}$ for a further 21 h. The absorbance was measured and % conversion calculated (98 %, 93%, 115 %).

1.4.8.4 Jackbean *O*-GlcNAcase

O-Glycosylated H3 (20.0 μL , 1.0 mg mL^{-1}) was treated with *Canavalia ensiformis* β -*N*-Acetylglucosaminidase (0.1 U mL^{-1} , 30 μL) in citric acid (100 mM), NaCl (200 mM) buffer pH 6.0. The reaction mixture was shaken at 37 °C overnight.

Method adapted from Sigma-Aldrich²¹⁹

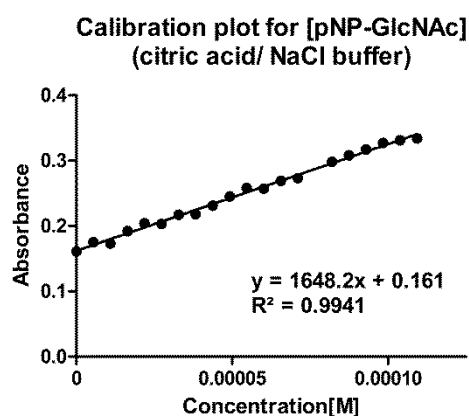


Calculated mass (H3K27 bis-homoSer): 15227

Observed mass: 15227 and 15431 (unreacted starting material)

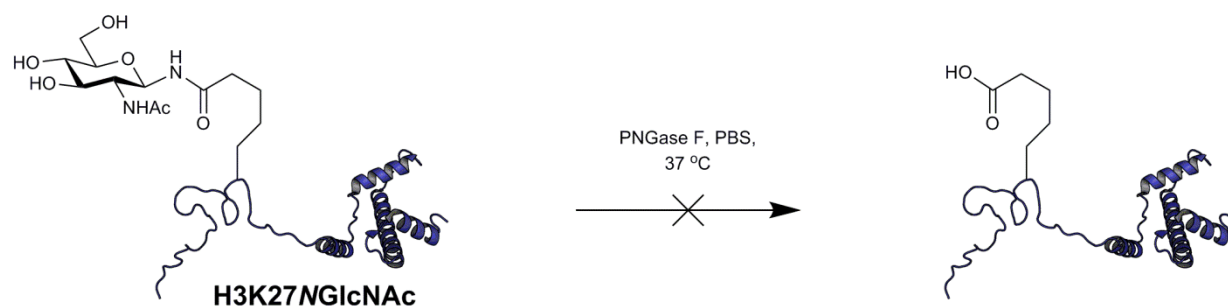
pNP-GlcNAc jbOGA Test

A calibration curve was set up from precisely prepared samples of pNP in the reaction buffer (citric acid 100 mM, NaCl 200 mM, pH 6.0), by measuring absorbance at 405 nm.



In triplicate jbOGA ($6.3 \mu\text{L}$, 8 U mL^{-1}) in the above buffer was added to a carefully prepared stock solution of pNP-GlcNAc ($58.4 \mu\text{M}$, $100 \mu\text{L}$ in reaction buffer) and heated to 25°C for 3 h. The absorbance was measured and % conversion calculated (97 %, 77%, 90 %). The reaction was heated to 25°C for a further 21 h. The absorbance was measured and % conversion calculated (102 %, 105%, 90.4 %).

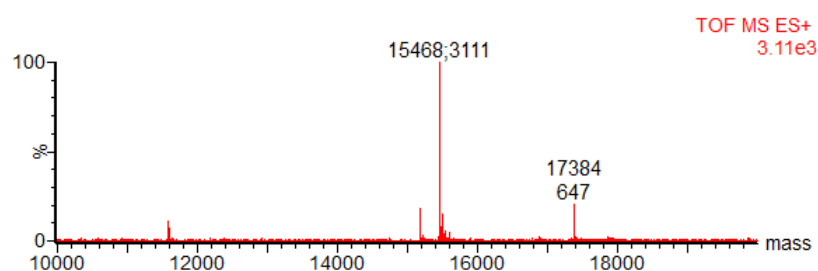
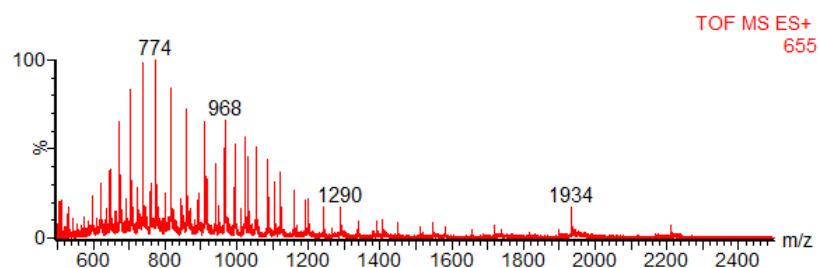
1.4.8.5 PNGase F



Methods adapted from New England Biolabs²²⁰

1.4.8.5.1 Native

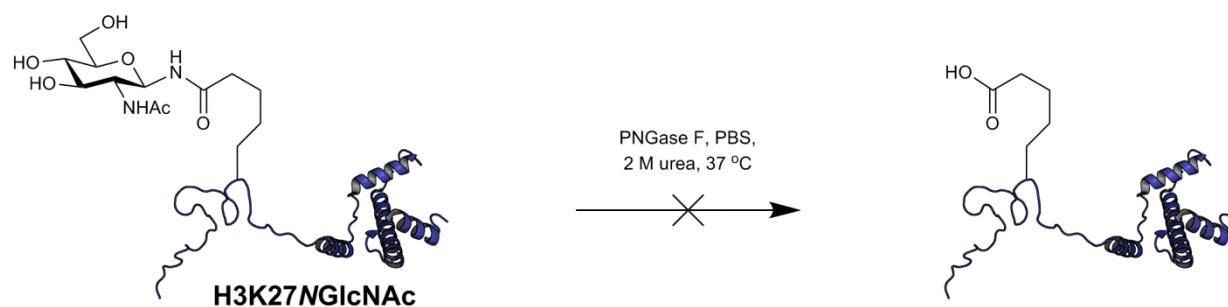
N-Glycosylated H3 (10.0 μL , 1.0 mg mL^{-1}) in PBS was mixed with PNGase F (0.1 μL , 500 U μL^{-1}) then shaken at 37 $^\circ\text{C}$ for 2 h. A second dose of PNGase F was added (1.0 μL , 500 U μL^{-1}) and shaken at 37 $^\circ\text{C}$ for 22 h.



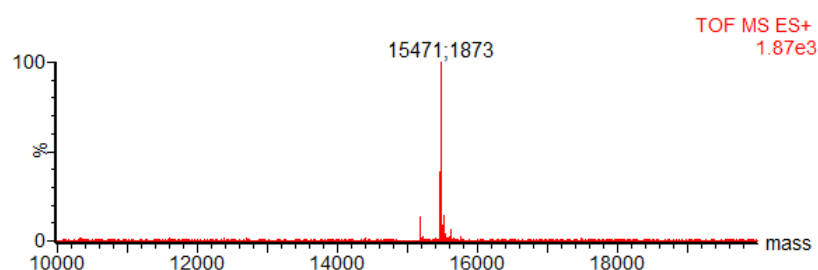
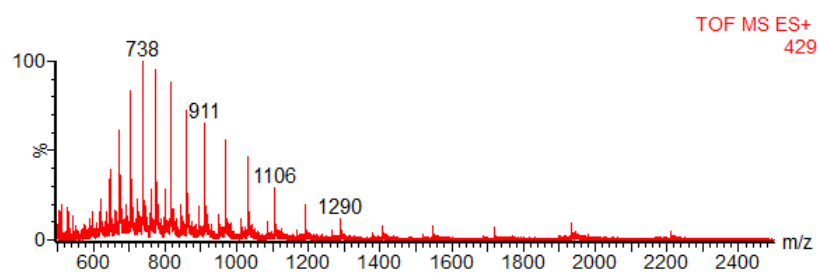
Calculated mass (H3K27 bis-homoAsp): 15268

Observed mass: 15468 (Starting Material)

1.4.8.5.2 Denatured



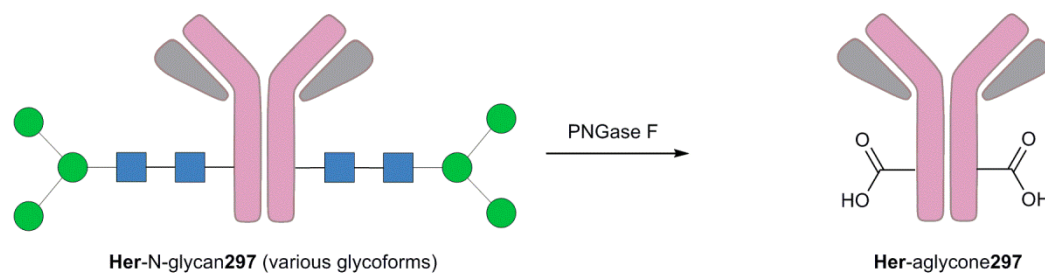
N-Glycosylated H3 (10.0 μL , 1.0 mg mL^{-1}) was mixed with urea solution (8 M, 2.5 μL) heated to 95 °C for 2 mins then PNGase F (0.1 μL , 500 U μL^{-1}) was added and the reaction mixture shaken at 37 °C for 2 h



Calculated mass (H3K27 bis-homoAsp): 15268

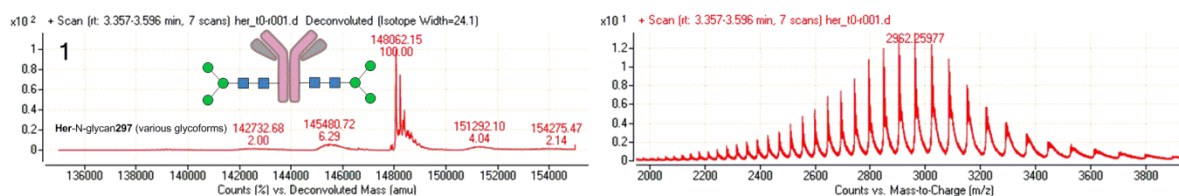
Observed mass: 15471 (Starting Material)

1.4.8.5.3 PNGase F activity control



Herceptin (5.0 μL , 4.0 mg mL^{-1}) was mixed with GlycoBuffer 2 (2.0 μL), Milli-Q H_2O (11.0 μL) and PNGase F (2 μL , 500 $\text{U } \mu\text{L}^{-1}$) and incubated at 37 °C for 24 hours. Samples were taken at given time-points and analysed by intact LCMS-ESI⁺.

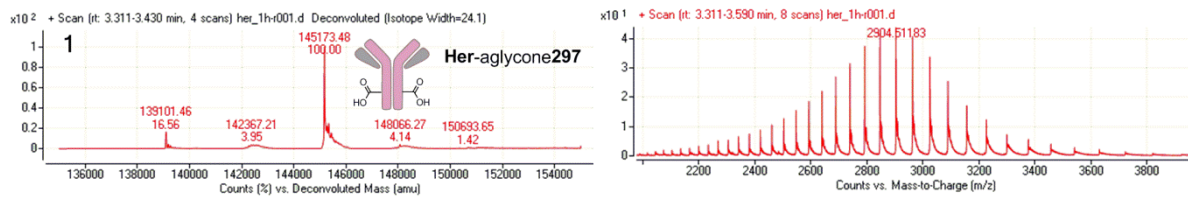
t = 0 (before PNGase F added)



Her-N-glycan297 calculated mass: 148057

observed mass: 148062 + other glycoforms

t = 1h

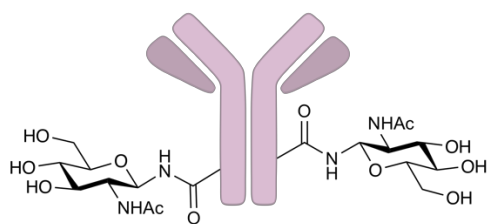


Her-aglycone297 calculated mass: 145166

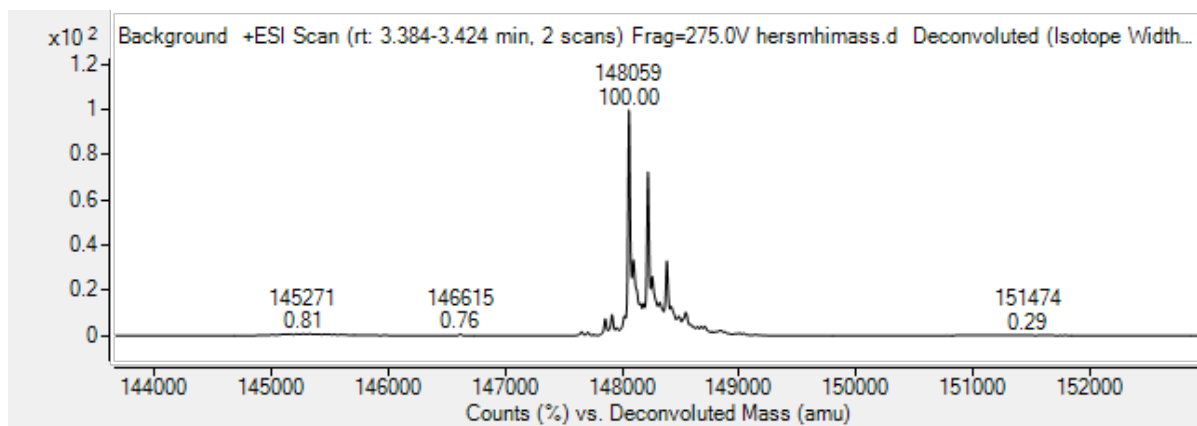
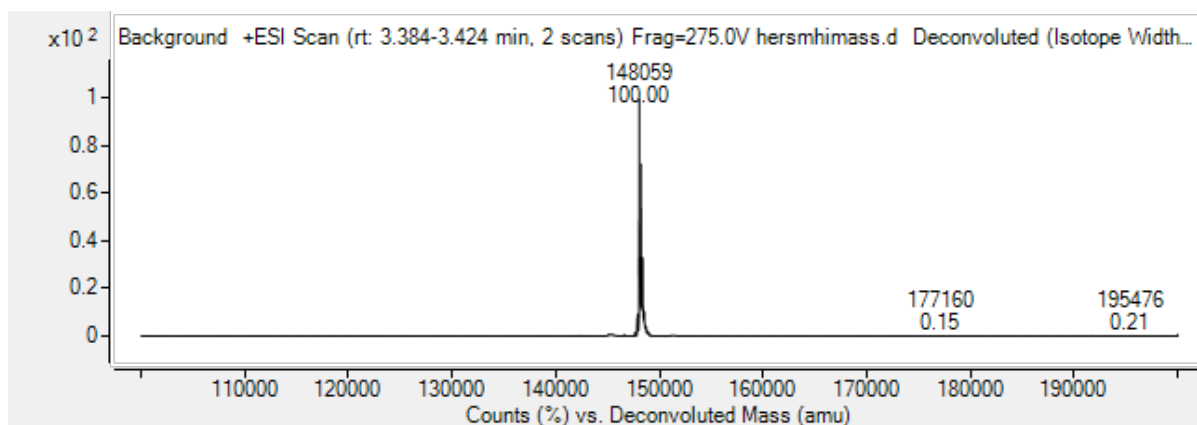
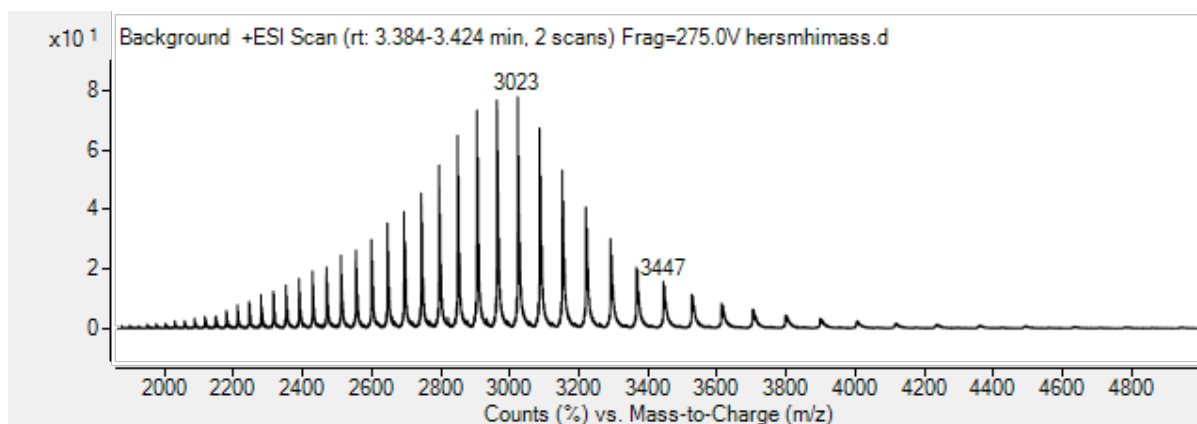
observed mass: 145174

1.4.8.6.4 Her EndoS Cleavage

Many thanks to Cristina Marculescu who kindly donated the sample of Endo S D233Q which was used in this reaction, the method for its preparation and use has been described by Davis.¹⁶³

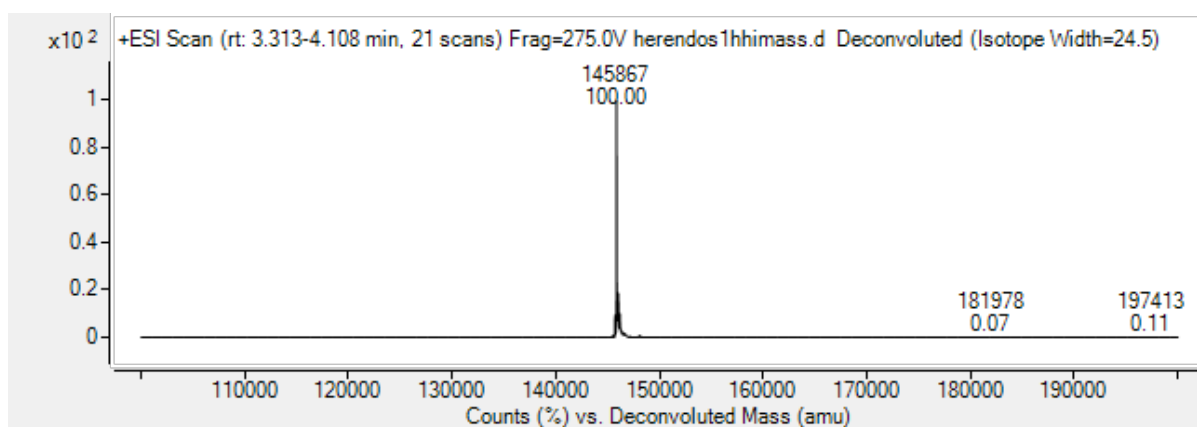
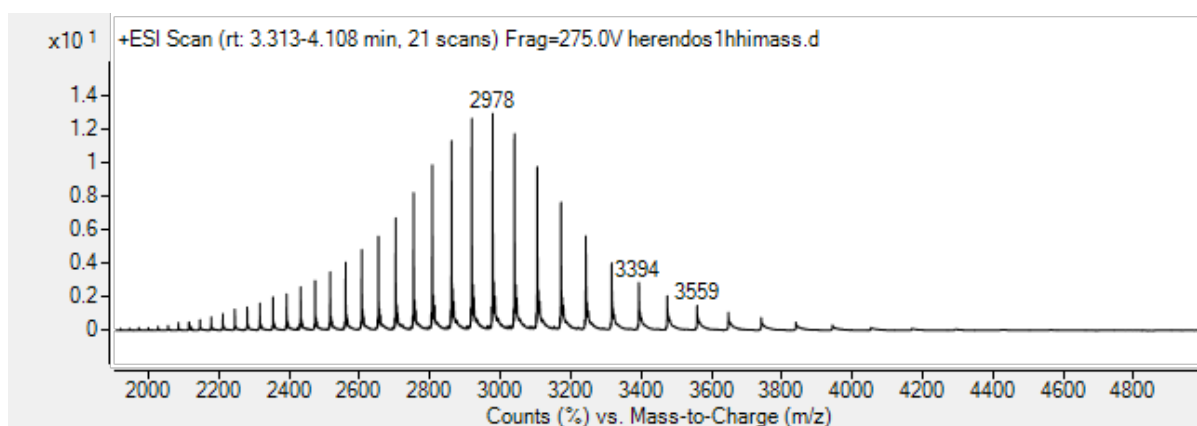


Herceptin (100 μL , 1.0 mg mL^{-1}) in PBS was treated with EndoS D233Q (10 μL , 0.5 mg mL^{-1}) and heated to 37 $^{\circ}\text{C}$. After 1 h complete conversion to the trimmed Her was observed by MS. The reaction mixture was then treated with PNGase F (0.1 μL , 500 $\text{U } \mu\text{L}^{-1}$) then heated to 37 $^{\circ}\text{C}$ for 29 h and analysed by MS.



Her-*N*-glycan297 calculated mass: 148060

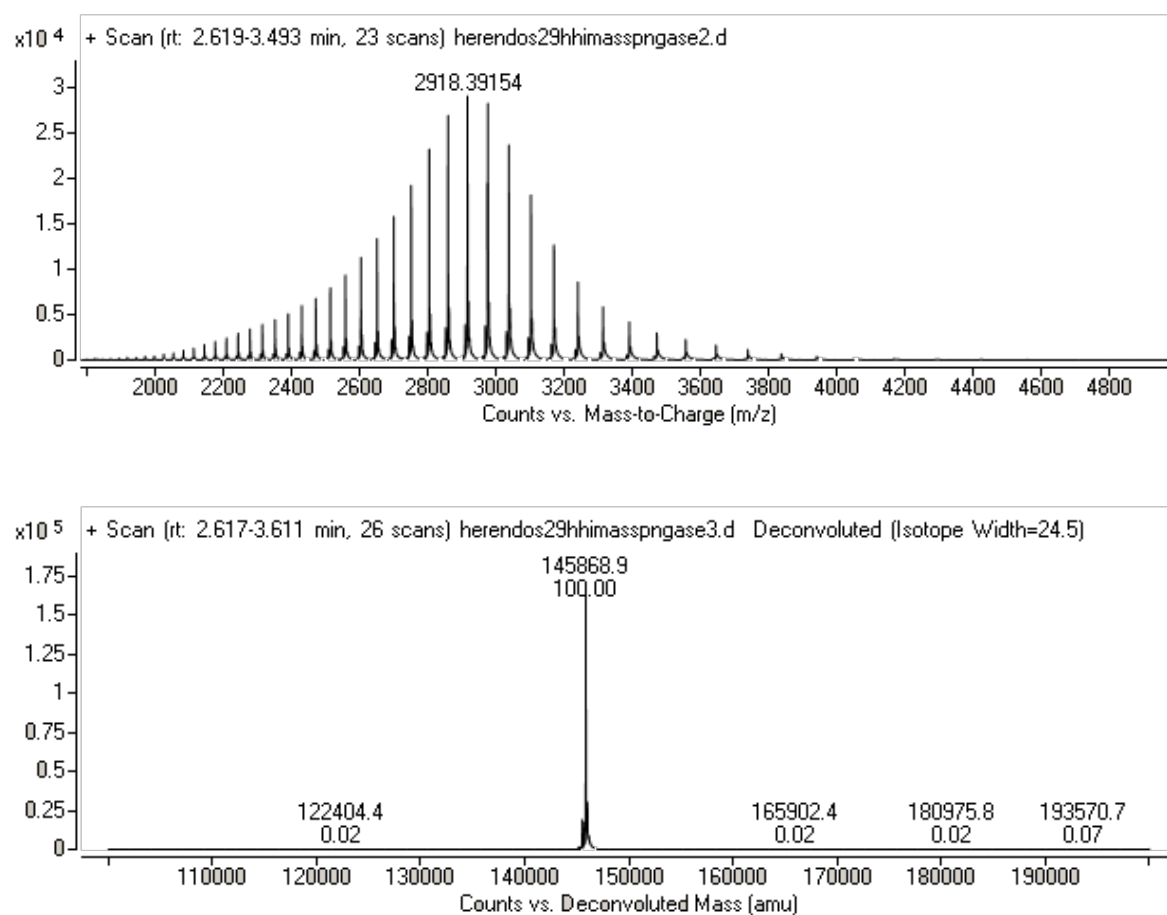
observed mass: 148059



Her-(GlcNAc-Fuc)₂ calculated mass: 145868

observed mass: 145867

1.4.8.5.5 Trimmed Her PNGase F cleavage



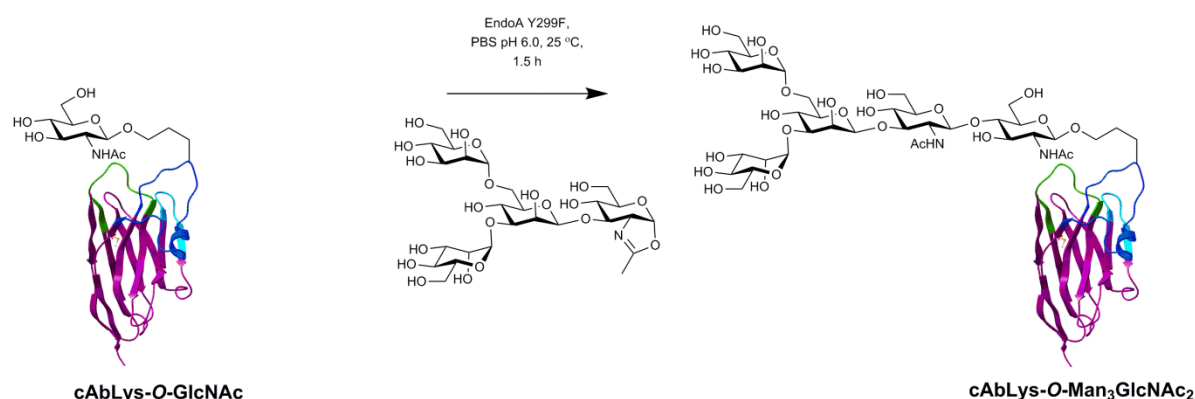
Her-aglycone297 calculated mass: 145166

observed mass: 145869 (starting material)

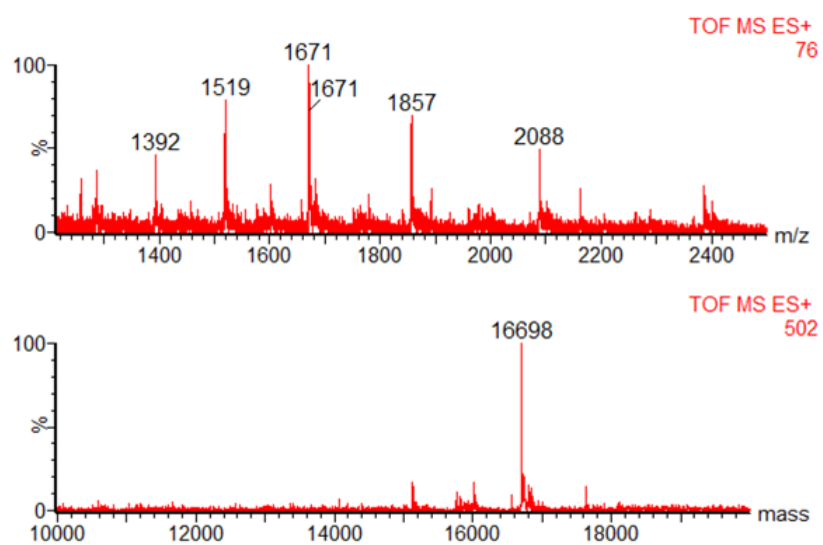
1.4.9.0 Glyco-engineering of cAb-Lys

Many thanks to Tom Wright who prepared the cAb-Lys glycoconjugate and Oliver Coleman who expressed the protein used in the following section, their preparation is described by Davis¹⁵⁴ and to Tom Parsons who kindly donated the carbohydrate oxazoline, the procedure for its synthesis is described by Davis¹⁶³

1.4.9.1 EndoA extension with tetrasaccharide oxazoline

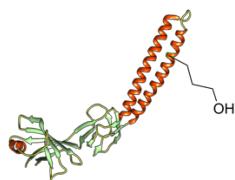


cAb-Lys A104-*O*-GlcNAc (10 μ L of a 0.3 mg mL⁻¹ solution in PBS pH 6.0 buffer) and tetrasaccharide oxazoline (3.6 μ L of a 5 mg mL⁻¹ aqueous solution) were combined and diluted to a final reaction volume of 25 μ L with PBS pH 6.0. EndoA Y299F (0.50 μ L of a 0.5 mg mL⁻¹ solution in PBS buffer) was added. The mixture was vortexed, incubated at 37 °C and aliquots taken for LC-MS analysis at regular time intervals. Full conversion to the desired extended glycoprotein product was observed after 1.5 hours. Method adapted from Wang¹⁷¹



1.4.10.0 Glyco-engineering of AcrA

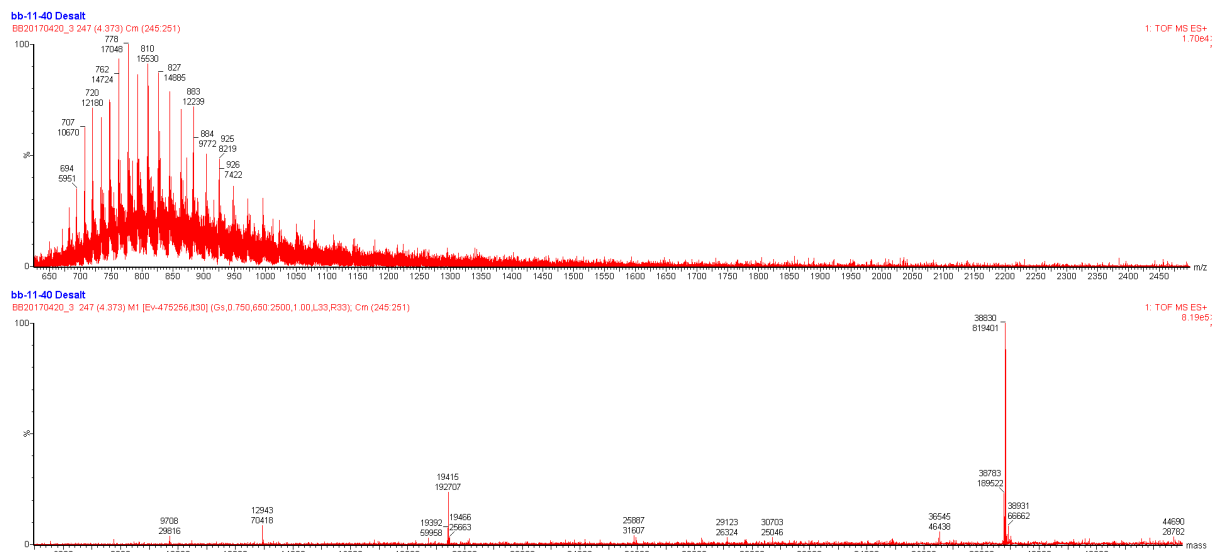
1.4.10.1 *Streptococcus pneumoniae* O-GlcNAcase cleavage test



O-Glycosylated H3 (18.0 μL , 0.15 mg mL^{-1}) in PBS was treated with Glycobuffer 1 10 x (2 μL) and *S. Pneumoniae* OGA (2 μL , 4 U mL^{-1}) in storage buffer. The reaction mixture was stood at 37 °C for 6 h. (97 % cleavage of the sugar). It is essential that the protein solution is free of 1-iodoethyl- β -D-GlcNAc as this inhibits the action of the OGA.

Method adapted from New England Biolabs.²²¹

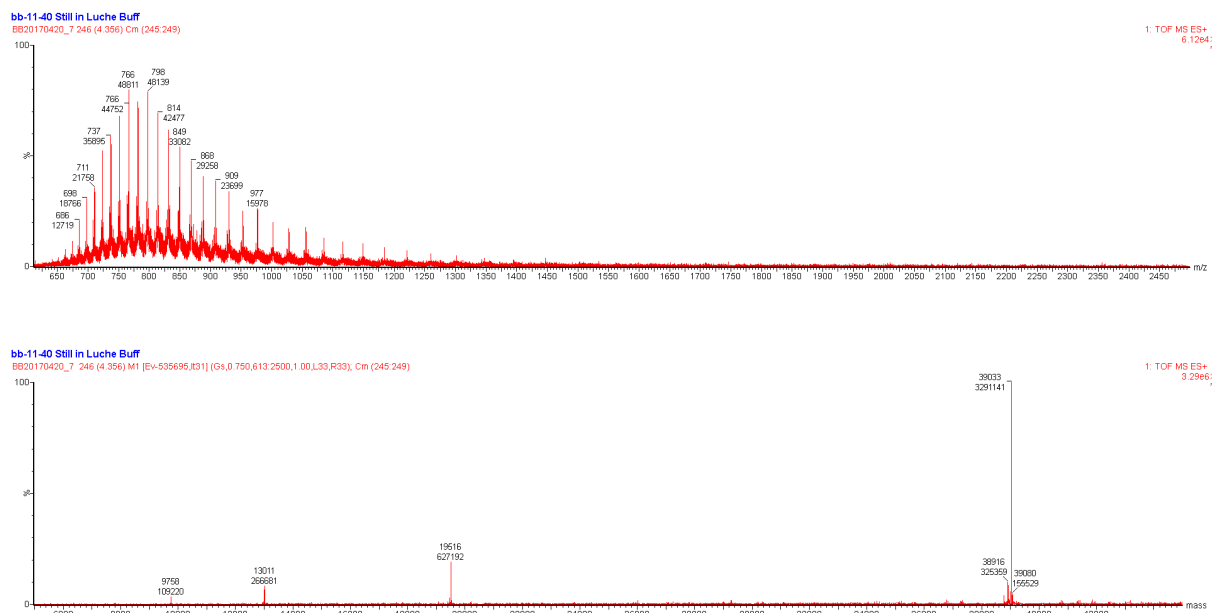
Desalted:



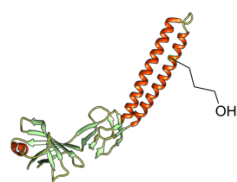
Calculated mass: 38829

Observed mass: 38830

Not Desalted:

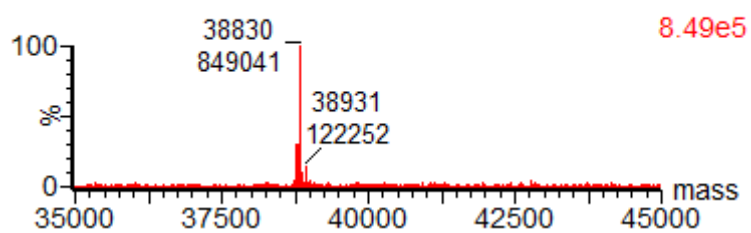
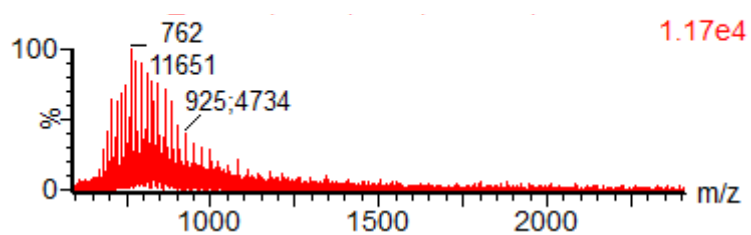


1.4.10.2 Human O-GlcNAcase cleavage test



O-Glycosylated H3 (20.0 μL , 0.15 mg mL^{-1}) in PBS was treated with Family 84 hOGA (0.5 μL , 10 mg mL^{-1}) in PBS pH 7.4. The reaction mixture was stood at 37 $^{\circ}\text{C}$ for 1 h. (79 % cleavage of the sugar).

Method adapted from Vocadlo.²¹⁶



Calculated mass: 38829

Observed mass: 38830

1.4.11.0 Protein Sequences of Expressed Proteins:

1.4.11.1 His₆-AcrA wt:

SKEEAPKIQMPPQPVTMTSAKSEDLPLSFTYPAKLVSDYDVIIKPQVSGVIVNKLFKAGDKV
KKGQTLFIIEQDKFKASVDSAYGQALMAKATFENASKDFNRSKALFSKSAISQKEYDSSLAT
FNNSKASLASARAQLANARIDLHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPI
YADFFISDTDKLNLVRNTQSGKWDLDSIHANLNLNGETVQGKLYFIDSVIDANSGTVKAKAV
FDNNNSTLLPGAFATITSEGFIQKNGFKVPQIGVKQDQNDVYVLLVKNGKVEKSSVHISYQN
NEYAIIIDKGLQNGDKIILDNFKKIQVGSEVKEIGAQLEHHHHHH

1.4.11.2 His₆-AcrA N123C:

SKEEAPKIQMPPQPVTMTSAKSEDLPLSFTYPAKLVSDYDVIIKPQVSGVIVNKLFKAGDKV
KKGQTLFIIEQDKFKASVDSAYGQALMAKATFENASKDFCRSKALFSKSAISQKEYDSSLAT
FNNSKASLASARAQLANARIDLHTEIKAPFDGTIGDALVNIGDYVSASTTELVRVTNLNPI
YADFFISDTDKLNLVRNTQSGKWDLDSIHANLNLNGETVQGKLYFIDSVIDANSGTVKAKAV
FDNNNSTLLPGAFATITSEGFIQKNGFKVPQIGVKQDQNDVYVLLVKNGKVEKSSVHISYQN
NEYAIIIDKGLQNGDKIILDNFKKIQVGSEVKEIGAQLEHHHHHH

1.4.11.3 Histone H3 K27C, C110A:

ARTKQTARKSTGGKAPRKQLATKAARCSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLI
RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVALFEDTNLAAIHAKRVTIMPKDI
QLARRIRGERA

1.4.11.4 Histone H3 S10C, C110A:

ARTKQTARKCTGGKAPRKQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLI
RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVALFEDTNLAAIHAKRVTIMPKDI
QLARRIRGERA

1.4.11.5 His₆-Family 84 *human O*-GlcNAcase (hOGA):

GSSHHHHHHSSGLVPRGSHMVQKESQATLEERESELSNPAAASAGASLEPPAAPAPGEDNPA
GAGGA AVAGAAGGARFLCGVVEGFYGRPWVMEQRKELFRRLQKWELNTYLYAPKDDYKHRM
FWREMYSVEEAEQLMTLISAAREYEIEFIYAI SPGLDITFSNPKEVSTLKRKLDQVSQFGCR
SFALLFDDIDHNMCAADKEVFSSFAHAQVSITNEIYQYLGE PETFLFCPTEYCGTFCYPNVS
QSPYLRTVGEKLLPGIEVLWTGPKVVSKEIPVESIEEVSKI IKRAPVIWDNIHANDYDQKRL
FLGPYKGRSTELIPRLKGVLTNP NCEFEANYVAIHTLATWYKSNMNGVRKDVVMTDSEDSTV
SIQIKLENEGSDEDIETDVL YSPQMALKLALTEWLQEFVPHQYSSRQVAHSGAKASVVDGT
PLVAAPSLNATTVVTTVYQEPIMSQGAALS GEPTTLTKEEEKKQPDEEPMDMVVEKQEETDH
KNDNQILSEIVEAKMAEELKPMDTDKESIAESKSPMSMQEDCISDIAPMQTDEQTNKEQFV
PGPNEKPLYTAEPVTLEDLQLLADLFYLPYEHGPKGAQMLREFQWLRANSSVSVNCKGKDS
EKIEEWSRAAKFEEMCGLVMGMFTRLSNCANRTILYDMYSYVWDIKSIMSMVKS FVQWLGC
RSHSSAQFLIGDQEPWAFRGGLAGEFQRLLPIDGANDLFFQPPPLTPTSKVYTIRPYFPKDE
ASVYKICREMYDDGVGLPFQSQPDLIGDKLVGGLLSLSLDYCFVLEDEDGICGYALGTVDVT
PFIKKCKISWIPFMQEKYTKPNGDKELSEAEKIMLSFHEEQEVLPETFLANFPSLIKMDIHK
KVTDPSVAKSMMACLLSSLKANGSRGAFCEVRPDDKRILEFYSKLGC FEIAKMEGF PKDVVI
LGRSL

1.4.11.6 GST-Endoglycosidase from *Streptococcus pyogenes* D233Q

MSPILGYWKIKGLVQPTRL LLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD
VKLTQSMAIIRYIADKHNLG GCPKERA EISMLEGAVLDIRYGVSR IAYSKDFETLKVDFLS
KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAI
PQIDKYLKSSKYIAWPLQGWQATFGGGDHPPKSDLVPRGSMDKHL LVKRTLGCVCAATLMGA
ALATHHDSLNTVKAEEKTVQVQKGLPSIDSLHYLSSENSKKEFKKEELSKAGQESQKVKEILAK
AQQADKQAQELAKMKIPEKIPMKPLHGPLYGGYFRTWHDKTS DPTEKDKVNSMGELPKVDL
AFIFHDWTKDYSLFWKELATKHVPKLNKQGTRVIRTIPWRFLAGGDNSGIAEDTSKYPNTPE
GNKALAKAIVDEYVYKYNLDGLDVQVEHDSIPKVDKKEDTAGVERSIQVFEEIGKLIGPKGV
DKSRLFIMDSTYMADKNPLIERGAPYINLLLVQVYGSQGEKGGWEPVSNRPEKTMEERWQGY
SKYIRPEQYMIGFSFYEENAQEGNLWYDINSRKDEDKANGINTDITGTRAERYARWQPKTGG
VKG GIFS YAIDRDGVAHQPKKYAKQKEFKDATDNI FHSDYSVSKALKTVMLKDKSYDLIDEK
DFPDKALREAVMAQVGRKGD LERFNGTLRLDNPAIQSLEGLNKFKKLAQLDLIGLSRITKL
DRSVLPANMKPGKDTLETVLETYKKDNKEEPATIPPVSLKVSGLTGLKELDLSGFDRETLAG
LDAATLTSLEKVDISGNKLDLAPGTENRQIFDTMLSTISNHVGSNEQTVKFDKQKPTGHYPD
TYGKTSRLRPVANEKVDLQS QLLFGTVTNQGT LINSEADYKAYQNHKIAGRSFVDSNYHYNN
FKVSYENYTVKVTDSTLGT TTDKTLATDKEETYKV DFFSPADKTKAVHTAKVIVGDEKTMV
NLAEGATVIGGSADPVNARKVFDGQLGSETDNISLGWDSKQSIIFKLKEDGLIKHWRFFNDS
ARNPETTNKPIQEASLQIFNIKDYNLDNLLNPNKFDD EKYWITVD TYSAQGERATAFSNTL
NNITSKYWRVVFDTKGDRYSSPVPELQILGYPLPNADTIMKTVTTAKELSQQKDKFSQKML
DELKIKEMALETSLNSKIFDVTAINANAGVLKDCIEKRQLLKK

1.4.11.7 GST-Endo- β -N-acetylglucosaminidase from *Arthrobacter*
protophormiae Y299F:

MSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELGLEFPNLPYYIDGD
VKLTQSMAIIRYIADKHNMLGGCPKERAETSMLEGAVLDIRYGVSRAYS KDFETLKVDFLS
KLPEMLKMFEDRLCHKTYLNGDHVTHPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAI
PQIDKYLKSSKYIAWPLQGWAQATFGGGDHPPKSDLVPRGSSTYNGPLSSHWFPEELAQWEPD
SDPDAPFNRS HVPLEPGRVADRVNANADKDAHLVSLALNRHTSGVPSQGAPVFYENTFSYW
HYTDLMVYWAGSAGEGIIVPPSADVIDASHRNGVPILGNVFFPPTVYGGQLEWLEQMLEQEE
DGSFPLADKLLEVADYYGFDGWFINQETEGADEGTAEAMQAFLVYLQEQQPEGMHIMWYDSM
IDTGAIWQNHLTDRNKMYLQNGSTRVADSMFLNFWWRDQRQSNELAQALGRSPYDLYAGVD
VEARGTSTPVQWEGLFPEGEKAHTSLGLFRPDWAFQSSETMEAFYEKELQFWVGSTGNPAET
DGQSNWPGMAHWFP AKSTATSVPFVTHFNTGSGAQFSAEGKTVSEQEWNNRSLQDVLPTWRW
IQHGGDLEATFSWEEAFEGGSSLQWHGSLAEGEHAQIELYQTELP ISEGTSLTWTFKSEHVN
DLNVGFRLDGEEDFRYVEGEQRESINGWTQWTLPLDAFAGQTITGLAFAAEGNETGLAEFYI
GQLAVGADSEKPAAPNVNVRQYDPDPSGIQLVWEKQSNVHHYRVYKETKHGKELIGTSAGDR
IYLEGLVEESKQNDVRLHIEALSETFVPSDARMIDIKSGS

1.5.0.0 References

11. P. Wang, S. Dong, J.-H. Shieh, E. Peguero, R. Hendrickson, M. A. S. Moore and S. J. Danishefsky, *Science*, 2013, **342**, 1357-1360.
16. J. S. Chauhan, A. Rao and G. P. S. Raghava, *PLoS One*, 2013, **8**, e67008.
17. K.-Y. Huang, M.-G. Su, H.-J. Kao, Y.-C. Hsieh, J.-H. Jhong, K.-H. Cheng, H.-D. Huang and T.-Y. Lee, *Nucleic Acids Res.*, 2016, **44**, D435-D446.
18. A. Eichholz, *J. Physiol.*, 1898, **23**, 163-177.
19. J. Cabezas, *Biochem. Ed.*, 1994, **22**, 3-7.
20. A. Neuberger, *Biochem. J.*, 1938, **32**, 1435-1451.
21. R. A. Laine, *Glycobiology*, 1994, **4**, 759-767.
22. R. D. Cummings, *Mol. Biosyst.*, 2009, **5**, 1087-1104.
23. R. G. Spiro, *Annu. Rev. Biochem.*, 1970, **39**, 599-638.
24. P. V. d. Steen, P. M. Rudd, R. A. Dwek and G. Opdenakker, *Crit. Rev. Biochem. Mol. Biol.*, 1998, **33**, 151-208.
25. F. Schwarz and M. Aebi, *Curr. Opin. Struct. Biol.*, 2011, **21**, 576-582.
26. J. Hofsteenge, D. R. Müller, T. de Beer, A. Löffler, W. J. Richter and J. F. Vliegthart, *Biochem.*, 1994, **33**, 13524-13530.
27. J. Stepper, S. Shastri, T. S. Loo, J. C. Preston, P. Novak, P. Man, C. H. Moore, V. Havlíček, M. L. Patchett and G. E. Norris, *FEBS Lett.*, 2011, **585**, 645-650.
28. G. A. Cross, *Cell*, 1987, **48**, 179-181.
29. R. Singh, A. Barden, T. Mori and L. Beilin, *Diabetologia*, 2001, **44**, 129-146.
30. R. J. Koenig, S. H. Blobstein and A. Cerami, *J. Biol. Chem.*, 1977, **252**, 2992-2997.
31. H. F. Bunn, K. H. Gabbay and P. M. Gallop, *Science*, 1978, **200**, 21.
32. V. M. Monnier and A. Cerami, *Science*, 1981, **211**, 491.
33. R. J. Koenig, C. M. Peterson, R. L. Jones, C. Saudek, M. Lehrman and A. Cerami, *N. Engl. J. Med.*, 1976, **295**, 417-420.
34. K. Zarschler, B. Janesch, M. Pabst, F. Altmann, P. Messner and C. Schäffer, *Glycobiology*, 2010, **20**, 787-798.
35. P. S. Chen, H. K. Mitchell and M. Neuweg, *Insect Biochem.*, 1978, **8**, 279-286.
36. M. A. Aon and J. A. Curtino, *Biochem. J.*, 1985, **229**, 269-272.
37. I. R. Rodriguez and W. J. Whelan, *Biochem. Biophys. Res. Commun.*, 1985, **132**, 829-836.
38. W. T. Butler and L. W. Cunningham, *J. Biol. Chem.*, 1966, **241**, 3882-3888.
39. H. Zhao, J. Sagert, D. S. Hwang and J. H. Waite, *Journal of Biological Chemistry*, 2009.
40. C. M. Taylor, C. V. Karunaratne and N. Xie, *Glycobiology*, 2012, **22**, 757-767.

41. D. T. A. Lamport and L. Clark, *Biochem.*, 1969, **8**, 1155-1163.
42. G. L. Gustafson and J. E. Gander, in *Methods Enzymol.*, Academic Press, 1984, vol. 107, pp. 172-183.
43. J. Roth, Y. Wang, A. E. Eckhardt and R. L. Hill, *Proc. Natl. Acad. Sci. U.S.A.*, 1994, **91**, 8935.
44. M. C. Field and L. J. Wainwright, *Glycobiology*, 1995, **5**, 463-472.
45. S. Kitazume-Kawaguchi, S. Inoue, Y. Inoue and W. J. Lennarz, *Proc. Natl. Acad. Sci. U.S.A.*, 1997, **94**, 3650.
46. Y. Tobisawa, Y. Imai, M. Fukuda and H. Kawashima, *J. Biol. Chem.*, 2009.
47. W.-X. Shi, R. Chammas and A. Varki, *J. Biol. Chem.*, 1996, **271**, 31517-31525.
48. W.-X. Shi, R. Chammas and A. Varki, *Glycobiology*, 1998, **8**, 199-205.
49. C. Cappellaro, K. Hauser, V. Mrša, M. Watzele, G. Watzele, C. Gruber and W. Tanner, *EMBO J.*, 1991, **10**, 4081-4088.
50. M. J. Kuranda and P. W. Robbins, *J. Biol. Chem.*, 1991, **266**, 19758-19767.
51. C. A. Wilcox and R. S. Fuller, *J. Cell Biol.*, 1991, **115**, 297.
52. J. Trueheart and G. R. Fink, *Proc. Natl. Acad. Sci. U.S.A.*, 1989, **86**, 9916-9920.
53. T. Krusius, J. Finne, R. K. Margolis and R. U. Margolis, *J. Biol. Chem.*, 1986, **261**, 8237-8242.
54. R. K. Margolis and R. U. Margolis, *Experientia*, 1993, **49**, 429-446.
55. A. Chiba, K. Matsumura, H. Yamada, T. Inazu, T. Shimizu, S. Kusunoki, I. Kanazawa, A. Kobata and T. Endo, *J. Biol. Chem.*, 1997, **272**, 2156-2162.
56. R. J. Harris, V. T. Ling and M. W. Spellman, *J. Biol. Chem.*, 1992, **267**, 5102-5107.
57. S. Hase, S.-i. Kawabata, H. Nishimura, H. Takeya, T. Suevoshi, T. Mivata, S. Iwanaga, T. Takao, Y. Shimonishi and T. Ikenaka, *J. Biochem*, 1988, **104**, 867-868.
58. G. D. Holt and G. W. Hart, *J. Biol. Chem.*, 1986, **261**, 8049-8057.
59. G. W. Hart, R. S. Haltiwanger, G. D. Holt and W. G. Kelly, *Annu. Rev. Biochem.*, 1989, **58**, 841-874.
60. C. Slawson, P. Housley Michael and W. Hart Gerald, *J. Cell. Biochem.*, 2005, **97**, 71-83.
61. C. R. Torres and G. W. Hart, *J. Biol. Chem.*, 1984, **259**, 3308-3317.
62. G. W. Hart, *Annu. Rev. Biochem.*, 1997, **66**, 315-335.
63. G. W. Hart, M. P. Housley and C. Slawson, *Nature*, 2007, **446**, 1017.
64. W. A. Lubas, D. W. Frank, M. Krause and J. A. Hanover, *J. Biol. Chem.*, 1997, **272**, 9316-9324.

65. R. J. Dennis, E. J. Taylor, M. S. Macauley, K. A. Stubbs, J. P. Turkenburg, S. J. Hart, G. N. Black, D. J. Vocadlo and G. J. Davies, *Nat. Struct. Mol. Biol.*, 2006, **13**, 365.
66. R. S. Haltiwanger, M. A. Blomberg and G. W. Hart, *J. Biol. Chem.*, 1992, **267**, 9005-9013.
67. I. Braidman, M. Carroll, N. Dance and D. Robinson, *FEBS Lett.*, 1974, **41**, 181-184.
68. D. L. Dong and G. W. Hart, *J. Biol. Chem.*, 1994, **269**, 19321-19330.
69. G. Manning, D. B. Whyte, R. Martinez, T. Hunter and S. Sudarsanam, *Science*, 2002, **298**, 1912.
70. L. Wells, L. K. Kreppel, F. I. Comer, B. E. Wadzinski and G. W. Hart, *J. Biol. Chem.*, 2004, **279**, 38466-38470.
71. W. B. Dias and G. W. Hart, *Mol. Biosyst.*, 2007, **3**, 766-772.
72. Y. E. Lewis, A. Galesic, P. M. Levine, C. A. De Leon, N. Lamiri, C. K. Brennan and M. R. Pratt, *ACS Chem. Biol.*, 2017, **12**, 1020-1027.
73. N. P. Marotta, Y. H. Lin, Y. E. Lewis, M. R. Ambroso, B. W. Zaro, M. T. Roth, D. B. Arnold, R. Langen and M. R. Pratt, *Nat. Chem.*, 2015, **7**, 913.
74. P. Marotta Nicholas, A. Cherwien Carli, T. Abeywardana and R. Pratt Matthew, *ChemBioChem*, 2012, **13**, 2665-2670.
75. M. Rechsteiner and S. W. Rogers, *Trends in Biochem. Sci.*, 1996, **21**, 267-271.
76. P. M. Levine, C. A. De Leon, A. Galesic, A. Balana, N. P. Marotta, Y. E. Lewis and M. R. Pratt, *Bioorg. Med. Chem.*, 2017, **25**, 4977-4982.
77. S. Rogers, R. Wells and M. Rechsteiner, *Science*, 1986, **234**, 364.
78. M. Sümegi, É. Hunyadi-Gulyás, K. F. Medzihradszky and A. Udvardy, *Biochem. Biophys. Res. Commun.*, 2003, **312**, 1284-1289.
79. F. Zhang, K. Su, X. Yang, D. B. Bowe, A. J. Paterson and J. E. Kudlow, *Cell*, 2003, **115**, 715-725.
80. N. E. Zachara and G. W. Hart, *Biochim. Biophys. Acta Gen. Subj.*, 2004, **1673**, 13-28.
81. N. E. Zachara, N. O'Donnell, W. D. Cheung, J. J. Mercer, J. D. Marth and G. W. Hart, *J. Biol. Chem.*, 2004, **279**, 30133-30142.
82. T. Lefebvre, C. Cieniewski, J. Lemoine, Y. Guerardel, Y. Leroy, J.-P. Zanetta and J.-C. Michalski, *Biochem. J.*, 2001, **360**, 179.
83. Emilyn U. Alejandro, N. Bozadjieva, D. Kumusoglu, S. Abdulhamid, H. Levine, L. Haataja, S. Vadrevu, Leslie S. Satin, P. Arvan and E. Bernal-Mizrachi, *Cell Rep.*, 2015, **13**, 2527-2538.
84. N. E. Zachara and G. W. Hart, *Biochim. Biophys. Acta Mol. Cell Biol. Lipids*, 2006, **1761**, 599-617.

85. N. Selvan, S. George, F. J. Serajee, M. Shaw, L. Hobson, V. M. Kalscheuer, N. Prasad, S. E. Levy, J. Taylor, S. Afitmos, C. E. Schwartz, A. M. Huq, J. Gecz and L. Wells, *J. Biol. Chem.*, 2018.
86. K. Sakabe, Z. Wang and G. W. Hart, *Proc. Natl. Acad. Sci. U.S.A*, 2010, **107**, 19915.
87. M. Chavan and W. Lennarz, *Trends in Biochem. Sci.*, 2006, **31**, 17-20.
88. M. Chavan and W. Lennarz, *Trends in Biochemical Sciences*, 2006, **31**, 17-20.
89. A. Yan and W. J. Lennarz, *Journal of Biological Chemistry*, 2005, **280**, 3121-3124.
90. R. D. Marshall, *Annu Rev Biochem*, 1972, **41**, 673-702.
91. L. Lamriben, B. Graham Jill, M. Adams Benjamin and N. Hebert Daniel, *Traffic*, 2015, **17**, 308-326.
92. H. Schachter, *Glycoconj. J.*, 2000, **17**, 465-483.
93. M. Kowarik, S. Numao, M. F. Feldman, B. L. Schulz, N. Callewaert, E. Kiermaier, I. Catrein and M. Aebi, *Science*, 2006, **314**, 1148.
94. P. Burda and M. Aebi, *Biochim. Biophys. Acta Gen. Subj.*, 1999, **1426**, 239-257.
95. C.-Y. Chung, N. I. Majewska, Q. Wang, J. T. Paul and M. J. Betenbaugh, *Cell*, 2017, **171**, 258-258.e251.
96. Y. Y. Zhao, M. Takahashi, J. G. Gu, E. Miyoshi, A. Matsumoto, S. Kitazume and N. Taniguchi, *Cancer Science*, 2008, **99**, 1304-1310.
97. S. S. Mendiratta, N. Sekulic, F. G. Hernandez-Guzman, B. E. Close, A. Lavie and K. J. Colley, *J. Biol. Chem.*, 2006, **281**, 36052-36059.
98. J. S. Li, L. Cui, D. L. Rock and J. Li, *Journal of Biological Chemistry*, 2005, **280**, 38513-38521.
99. S. Zou, S. Guo, I. Kaleem and C. Li, *J. Chem. Technol. Biotechnol.*, 2013, **88**, 1913-1919.
100. D. G. Singh, J. Lomako, W. M. Lomako, W. J. Whelan, H. E. Meyer, M. Serwe and J. W. Metzger, *FEBS Lett.*, 1995, **376**, 61-64.
101. S. Li, L. Zhang, Q. Yao, L. Li, N. Dong, J. Rong, W. Gao, X. Ding, L. Sun, X. Chen, S. Chen and F. Shao, *Nature*, 2013, **501**, 242.
102. J. S. Pearson, C. Giogha, S. Y. Ong, C. L. Kennedy, M. Kelly, K. S. Robinson, T. W. F. Lung, A. Mansell, P. Riedmaier, C. V. L. Oates, A. Zaid, S. Mühlen, V. F. Crepin, O. Marches, C.-S. Ang, N. A. Williamson, L. A. O'Reilly, A. Bankovacki, U. Nachbur, G. Infusini, A. I. Webb, J. Silke, A. Strasser, G. Frankel and E. L. Hartland, *Nature*, 2013, **501**, 247.

103. J. Lassak, E. C. Keilhauer, M. Fürst, K. Wuichet, J. Gödeke, A. L. Starosta, J.-M. Chen, L. Søgaaard-Andersen, J. Rohr, D. N. Wilson, S. Häussler, M. Mann and K. Jung, *Nat. Chem. Biol.*, 2015, **11**, 266.
104. G. Pergolizzi, S. Dedola and R. A. Field, *Carbohydrate Chemistry: Volume 42*, 2017, **42**, 1-46.
105. R. B. Merrifield, *Journal of the American Chemical Society*, 1963, **85**, 2149-2154.
106. P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. Kent, *Science*, 1994, **266**, 776.
107. G. R. Gray, *Arch. Biochem. Biophys.*, 1974, **163**, 426-428.
108. L. Malaprade, *Bull. Soc. Chim.*, 1926, **39**, 325.
109. P. F. L. Fleury, J., *Compt. Rend.*, 1932, **195**, 1395.
110. J. M. Bobbitt, in *Adv. Carbohydr. Chem.*, eds. M. L. Wolfrom and R. S. Tipson, Academic Press, 1956, vol. 11, pp. 1-41.
111. P. Bulpitt and D. Aeschlimann, *J. Biomed. Mater. Res.*, 1999, **47**, 152-169.
112. J. Porath, R. AxÉN and S. Ernback, *Nature*, 1967, **215**, 1491.
113. G. S. Bethell, J. S. Ayers, W. S. Hancock and M. T. Hearn, *J. Biol. Chem.*, 1979, **254**, 2572-2574.
114. V. V. Rostovtsev, L. G. Green, V. V. Fokin and K. B. Sharpless, *Angew. Chem. Int. Ed.*, 2002, **41**, 2596-2599.
115. C. W. Tornøe, C. Christensen and M. Meldal, *J. Org. Chem.*, 2002, **67**, 3057-3064.
116. S. Dedola, S. A. Nepogodiev and R. A. Field, *Org. Biomol. Chem.*, 2007, **5**, 1006-1017.
117. T. Lipinski, T. Luu, P. I. Kitov, A. Szpacenko and D. R. Bundle, *Glycoconj. J.*, 2011, **28**, 149-164.
118. D. P. Nguyen, H. Lusic, H. Neumann, P. B. Kapadnis, A. Deiters and J. W. Chin, *J. Am. Chem. Soc.*, 2009, **131**, 8720-8721.
119. B. T. Worrell, J. A. Malik and V. V. Fokin, *Science*, 2013, **340**, 457.
120. C. Amatore, G. Le Duc and A. Jutand, *Chem. Eur. J.*, 2013, **19**, 10082-10093.
121. C. D. Spicer and B. G. Davis, *ChemComm*, 2011, **47**, 1698-1700.
122. J. M. Chalker, C. S. C. Wood and B. G. Davis, *J. Am. Chem. Soc.*, 2009, **131**, 16346-16347.
123. M. B. Soellner, B. L. Nilsson and R. T. Raines, *J. Am. Chem. Soc.*, 2006, **128**, 8820-8828.
124. E. Saxon, J. I. Armstrong and C. R. Bertozzi, *Org. Lett.*, 2000, **2**, 2141-2143.
125. J. L. Bernardes Gonçalo, L. Linderöth, J. Doores Katie, O. Boutureira and G. Davis Benjamin, *ChemBioChem*, 2011, **12**, 1383-1386.

126. H. Liu, L. Wang, A. Brock, C.-H. Wong and P. G. Schultz, *J. Am. Chem. Soc.*, 2003, **125**, 1702-1703.
127. N. Floyd, B. Vijayakrishnan, R. Koeppe Julia and G. Davis Benjamin *Angew. Chem. Int. Ed.*, 2009, **48**, 7798-7802.
128. A. Bairoch, Amino acid scale: Amino acid composition (%) in the UniProtKB/Swiss-Prot data bank., <https://web.expasy.org/protscale/pscale/A.A.Swiss-Prot.html>, (accessed 7-8-18, 2018).
129. T. W. Bentley, *Org. Biomol. Chem.*, 2011, **9**, 6685-6690.
130. D. R. Goddard and L. Michaelis, *J. Biol. Chem.*, 1935, **112**, 361-371.
131. N. J. Davis and S. L. Flitsch, *Tetrahedron Lett.*, 1991, **32**, 6793-6796.
132. S. Y. Wong, G. R. Guile, R. A. Dwek and G. Arsequell, *Biochem. J.*, 1994, **300**, 843-850.
133. I. Shin, H.-j. Jung and M.-r. Lee, *Tetrahedron Lett.*, 2001, **42**, 1325-1328.
134. J. Ni, S. Singh and L.-X. Wang, *Bioconjugate Chem.*, 2003, **14**, 232-238.
135. M. Elzinga and J. J. Phelan, *Proc. Natl. Acad. Sci. U.S.A*, 1984, **81**, 6599.
136. S. M. Bachrach and D. C. Mulhearn, *The Journal of Physical Chemistry*, 1996, **100**, 3535-3540.
137. D. Russell, N. J. Oldham and B. G. Davis, *Carbohydr. Res.*, 2009, **344**, 1508-1514.
138. K. Omura, A. K. Sharma and D. Swern, *The Journal of Organic Chemistry*, 1976, **41**, 957-962.
139. S. R. G. Galan, J. R. Wickens, J. Dadova, W.-L. Ng, X. Zhang, R. A. Simion, R. Quinlan, E. Pires, R. S. Paton, S. Caddick, V. Chudasama and B. G. Davis, *Nat. Chem. Biol.*, 2018, **14**, 955-963.
140. L. Polgar and M. L. Bender, *J. Am. Chem. Soc.*, 1966, **88**, 3153-3154.
141. C. Zioudrou, M. Wilchek and A. Patchornik, *Biochem.*, 1965, **4**, 1811-1822.
142. K. E. Neet and D. E. Koshland, *Proc. Natl. Acad. Sci. U.S.A*, 1966, **56**, 1606-1611.
143. M. Fernández-González, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C. Errey and B. G. Davis, *Chem. Sci.*, 2010, **1**, 709-715.
144. G. J. L. Bernardes, J. M. Chalker, J. C. Errey and B. G. Davis, *J. Am. Chem. Soc.*, 2008, **130**, 5052-5053.
145. C. A. De Leon, P. M. Levine, T. W. Craven and M. R. Pratt, *Biochem.*, 2017, **56**, 3507-3517.
146. R. Raj, L. Lercher, S. Mohammed and B. G. Davis, *Angewandte Chemie*, 2016, **128**, 9064-9068.
147. L. Lercher, R. Raj, N. A. Patel, J. Price, S. Mohammed, C. V. Robinson, C. J. Schofield and B. G. Davis, *Nature Communications*, 2015, **6**, 7978.

148. E. J. Corey and N. W. Boaz, *Tetrahedron Lett.*, 1985, **26**, 6015-6018.
149. C. Petrier, C. Dupuy and J. L. Luche, *Tetrahedron Lett.*, 1986, **27**, 3149-3152.
150. J. Wang, H. Lundberg, S. Asai, P. Martín-Acosta, J. S. Chen, S. Brown, W. Farrell, R. G. Dushin, C. J. O'Donnell, A. S. Ratnayake, P. Richardson, Z. Liu, T. Qin, D. G. Blackmond and P. S. Baran, *Proc. Natl. Acad. Sci. U.S.A.*, 2018, **115**, E6404.
151. C. L. Hawkins and M. J. Davies, *Biochim. Biophys. Acta Bioenerg.*, 2001, **1504**, 196-219.
152. Q. Wan and J. Danishefsky Samuel *Angew. Chem. Int. Ed.*, 2007, **46**, 9248-9252.
153. H. G. Viehe, Z. Janousek, R. Merenyi and L. Stella, *Accounts Chem. Res.*, 1985, **18**, 148-154.
154. T. H. Wright, B. J. Bower, J. M. Chalker, G. J. L. Bernardes, R. Wiewiora, W.-L. Ng, R. Raj, S. Faulkner, M. R. J. Vallée, A. Phantumruiwath, O. D. Coleman, M.-L. Thézénas, M. Khan, S. R. G. Galan, L. Lercher, M. W. Schombs, S. Gerstberger, M. E. Palm-Espling, A. J. Baldwin, B. M. Kessler, T. D. W. Claridge, S. Mohammed and B. G. Davis, *Science*, 2016, **354**.
155. D. R. Lide, *Handbook of Chemistry and Physics*, CRC Press, Taylor & Francis Group, Boca Raton, 90th edn., 2009.
156. J. Z. Saavedra, A. Resendez, A. Rovira, S. Eagon, D. Haddenham and B. Singaram, *J. Org. Chem.*, 2012, **77**, 221-228.
157. H. Christensen, M. S. Christiansen, J. Petersen and H. H. Jensen, *Org. Biomol. Chem.*, 2008, **6**, 3276-3283.
158. G. Zemplén and E. Pacsu, *Ber. Dtsch. Chem. Ges.*, 1929, **62**, 1613-1614.
159. K. Mori, M. Tominaga, T. Takigawa and M. Matsui, *Synthesis*, 1973, **1973**, 790-791.
160. F. G. Bordwell and W. T. Brannen, *J. Am. Chem. Soc.*, 1964, **86**, 4645-4650.
161. F. M. Welle, H.-D. Beckhaus and C. Rüchardt, *J. Org. Chem.*, 1997, **62**, 552-558.
162. T. Foster, D. Klapstein and P. R. West, *Can. J. Chem.*, 1974, **52**, 524-526.
163. J. J. Goodfellow, K. Baruah, K. Yamamoto, C. Bonomelli, B. Krishna, D. J. Harvey, M. Crispin, C. N. Scanlan and B. G. Davis, *J. Am. Chem. Soc.*, 2012, **134**, 8030-8033.
164. F. Schwarz, W. Huang, C. Li, B. L. Schulz, C. Lizak, A. Palumbo, S. Numao, D. Neri, M. Aebi and L.-X. Wang, *Nat. Chem. Biol.*, 2010, **6**, 264.
165. K. Witte, P. Sears, R. Martin and C.-H. Wong, *J. Am. Chem. Soc.*, 1997, **119**, 2114-2118.
166. C. S. Bennett and C.-H. Wong, *Chem. Soc. Rev.*, 2007, **36**, 1227-1238.
167. J. R. Rich and S. G. Withers, *Nat. Chem. Biol.*, 2008, **5**, 206.

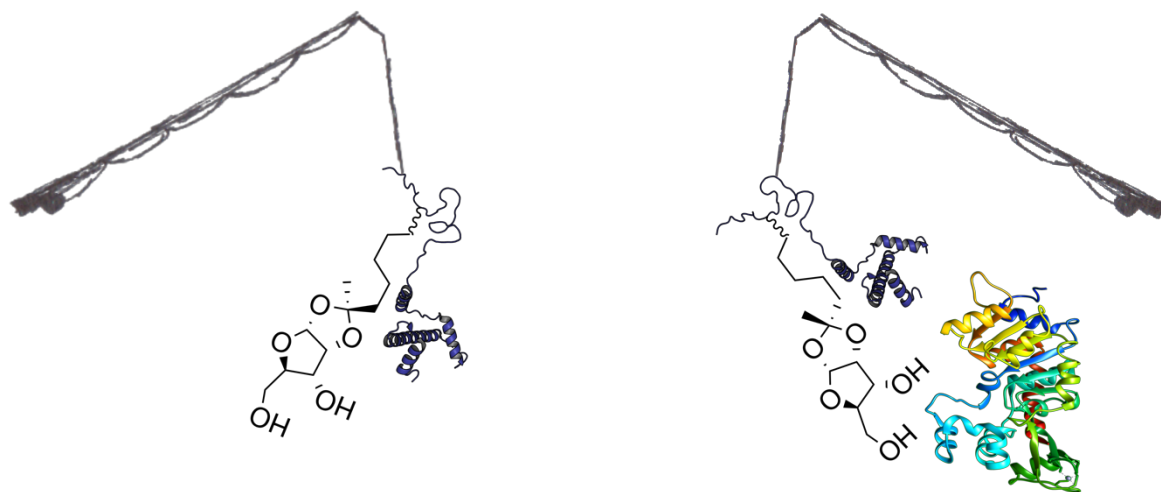
168. L.-X. Wang, *Carbohydr. Res.*, 2008, **343**, 1509-1522.
169. J. D. McIntosh, M. A. Brimble, A. E. S. Brooks, P. R. Dunbar, R. Kowalczyk, Y. Tomabechi and A. J. Fairbanks, *Chem. Sci.*, 2015, **6**, 4636-4642.
170. C. D. Heidecke, Z. Ling, N. C. Bruce, J. W. B. Moir, T. B. Parsons and A. J. Fairbanks, *ChemBioChem*, 2008, **9**, 2045-2051.
171. J. Yin, L. Li, N. Shaw, Y. Li, J. K. Song, W. Zhang, C. Xia, R. Zhang, A. Joachimiak, H.-C. Zhang, L.-X. Wang, Z.-J. Liu and P. Wang, *PLoS One*, 2009, **4**, e4658.
172. J. T. Powell and K. Brew, *Eur. J. Biochem.*, 1974, **48**, 217-228.
173. T. M. Gloster and D. J. Vocadlo, *Curr. Signal Transd. T.*, 2010, **5**, 74-91.
174. Y. Gao, L. Wells, F. I. Comer, G. J. Parker and G. W. Hart, *J. Biol. Chem.*, 2001, **276**, 9838-9845.
175. S.-C. Li and Y.-T. Li, *J. Biol. Chem.*, 1970, **245**, 5153-5160.
176. A. L. Tarentino, C. M. Gómez and T. H. Plummer, *Biochem.*, 1985, **24**, 4665-4671.
177. J. H. Elder and S. Alexander, *Proc. Natl. Acad. Sci. U.S.A.*, 1982, **79**, 4540.
178. A. L. Tarentino and T. H. Plummer, in *Methods Enzymol.*, Academic Press, 1994, vol. 230, pp. 44-57.
179. C. W. N. Damen, W. Chen, A. B. Chakraborty, M. van Oosterhout, J. R. Mazzeo, J. C. Gebler, J. H. M. Schellens, H. Rosing and J. H. Beijnen, *J. Am. Soc. Mass Spectrom.*, 2009, **20**, 2021-2033.
180. J.-Q. Fan and Y. C. Lee, *J. Biol. Chem.*, 1997, **272**, 27058-27064.
181. M. Imiolek, G. Karunanithy, W.-L. Ng, A. J. Baldwin, V. Gouverneur and B. G. Davis, *Journal of the American Chemical Society*, 2018, **140**, 1568-1571.
182. K. Lewis, *Trends in Biochem. Sci.*, 1994, **19**, 119-123.
183. M. F. Symmons, E. Bokma, E. Koronakis, C. Hughes and V. Koronakis, *Proc. Natl. Acad. Sci. U.S.A.*, 2009, **106**, 7173.
184. Z. Wang, G. Fan, C. F. Hryc, J. N. Blaza, I. I. Serysheva, M. F. Schmid, W. Chiu, B. F. Luisi and D. Du, *eLife*, 2017, **6**, e24905.
185. H. I. Zgurskaya and H. Nikaido, *J. Mol. Biol.*, 1999, **285**, 409-420.
186. M. Wacker, D. Linton, P. G. Hitchen, M. Nita-Lazar, S. M. Haslam, S. J. North, M. Panico, H. R. Morris, A. Dell, B. W. Wren and M. Aebi, *Science*, 2002, **298**, 1790.
187. M. Kowarik, N. M. Young, S. Numao, B. L. Schulz, I. Hug, N. Callewaert, D. C. Mills, D. C. Watson, M. Hernandez, J. F. Kelly, M. Wacker and M. Aebi, *EMBO J.*, 2006, **25**, 1957.
188. D. Linton, N. Dorrell, P. G. Hitchen, S. Amber, A. V. Karlyshev, H. R. Morris, A. Dell, M. A. Valvano, M. Aebi and B. W. Wren, *Mol. Microbiol.*, 2005, **55**, 1695-1703.

189. S. Kiese, A. Pappenberg, W. Friess and H.-C. Mahler, *J. Pharm. Sci.*, 2008, **97**, 4347-4366.
190. C. Fehl, personal communication.
191. J. W. Beatty and C. R. J. Stephenson, *Accounts Chem. Res.*, 2015, **48**, 1474-1484.
192. Helge C. Dorfmüller, Vladimir S. Borodkin, M. Schimpl and Daan M. F. van Aalten, *Biochemical Journal*, 2009, **420**, 221.
193. B. Li, H. Li, L. Lu and J. Jiang, *Nature Structural & Molecular Biology*, 2017, **24**, 362.
194. N. L. Elsen, S. B. Patel, R. E. Ford, D. L. Hall, F. Hess, H. Kandula, M. Kornienko, J. Reid, H. Selnick, J. M. Shipman, S. Sharma, K. J. Lumb, S. M. Soisson and D. J. Klein, *Nature Chemical Biology*, 2017, **13**, 613.
195. C. Su, Y. Lu and H. Liu, *Nature Communications*, 2016, **7**, 12916.
196. S. H. Knapp, R. A.; Amorelli, B., *Org. Synth.*, 2007, **84**, 68-75.
197. A. C. Cunha, L. O. R. Pereira, R. O. P. de Souza, M. C. B. V. de Souza and V. F. Ferreira, *Nucleosides Nucleotides Nucleic Acids*, 2001, **20**, 1555-1569.
198. D. Macmillan, A. M. Daines, M. Bayrhuber and S. L. Flitsch, *Org. Lett.*, 2002, **4**, 1467-1470.
199. K. K. G. Ramakrishna, R. K. Thakur, V. R. Pasam, J. Pandey, R. Mahar, S. K. Shukla, A. K. Tamrakar and R. P. Tripathi, *Tetrahedron*, 2017, **73**, 187-203.
200. U. Sirion, H. J. Kim, J. H. Lee, J. W. Seo, B. S. Lee, S. J. Lee, S. J. Oh and D. Y. Chi, *Tetrahedron Lett.*, 2007, **48**, 3953-3957.
201. E. D. Soli and P. DeShong, *J. Org. Chem.*, 1999, **64**, 9724-9726.
202. S. Dutta, S. J. Gupta and A. K. Sen, *Tetrahedron Lett.*, 2016, **57**, 3086-3090.
203. M. Mizuno, T. Shioiri and M. Mizuno, *ChemComm*, 1997, DOI: 10.1039/A705768G, 2165-2166.
204. H. Bayley, D. N. Standring and J. R. Knowles, *Tetrahedron Lett.*, 1978, **19**, 3633-3634.
205. T. Takeda, Y. Sugiura, Y. Ogihara and S. Shibata, *Can. J. Chem.*, 1980, **58**, 2600-2603.
206. H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235-242.
207. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J. Comput. Chem.*, 2004, **25**, 1605-1612.
208. J. Mikolosko, K. Bobyk, H. I. Zgurskaya and P. Ghosh, *Structure*, 2006, **14**, 577-587.
209. K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature*, 1997, **389**, 251.

210. Z. Ling, M. D. L. Suits, R. J. Bingham, N. C. Bruce, G. J. Davies, A. J. Fairbanks, J. W. B. Moir and E. J. Taylor, *J. Mol. Biol.*, 2009, **389**, 1-9.
211. C. Roth, S. Chan, W. A. Offen, G. R. Hemsworth, L. I. Willems, D. T. King, V. Varghese, R. Britton, D. J. Vocadlo and G. J. Davies, *Nat. Chem. Biol.*, 2017, **13**, 610.
212. B. Ramakrishnan, A. J. Moncrief, T. A. Davis, L. A. Holland and P. K. Qasba, *Glycoconj. J.*, 2013, **30**, 835-842.
213. A. Desmyter, T. R. Transue, M. A. Ghahroudi, M.-H. Dao Thi, F. Poortmans, R. Hamers, S. Muyldermans and L. Wyns, *Nat. Struct. Biol.*, 1996, **3**, 803.
214. K. Rantalainen, Z. T. Berndsen, S. Murrell, L. Cao, O. Omorodion, J. L. Torres, M. Wu, J. Umotoy, J. Copps, P. Poignard, E. Landais, J. C. Paulson, I. A. Wilson and A. B. Ward, *Cell Rep.*, 2018, **23**, 3249-3261.
215. F. Liu, B. Vijayakrishnan, A. Faridmoayer, T. A. Taylor, T. B. Parsons, G. J. L. Bernardes, M. Kowarik and B. G. Davis, *J. Am. Chem. Soc.*, 2014, **136**, 566-569.
216. M. S. Macauley, K. A. Stubbs and D. J. Vocadlo, *J. Am. Chem. Soc.*, 2005, **127**, 17202-17203.
217. P. B. Becker, *Chromatin Protocols (Methods in Molecular Biology)*, Humana Press, New York City, 1999.
218. D. K. Podolsky and M. M. Weiser, *Biochemical and Biophysical Research Communications*, 1975, **65**, 545-551.
219. S. A. UK, β -N-Acetylglucosaminidase from *Canavalia ensiformis* (Jack bean), <https://www.sigmaaldrich.com/catalog/product/sigma/a2264?lang=en®ion=GB>, 2018).
220. N. E. B. UK, PNGase F, <https://international.neb.com/products/p0704-pngase-f#Product%20Information>, 2018).
221. N. E. Biolabs, β -N-Acetylglucosaminidase S Usage Data Card, https://www.neb.com/products/p0744-b-n-acetylglucosaminidase-s#Protocols,%20Manuals%20&%20Usage_Datacards).

Chapter 2: Determining the Stereochemistry of a Sirt2

Intermediate



2.1.1.0 Introduction

The data encoded in DNA cannot explain all of the complexity of life. An additional layer of regulation exists on top genetics which is propagated through the modification of the DNA bases and histone proteins which make up chromatin, known as epigenetics.²²² Histone acetylation is an epigenetic marker²²³ which has been demonstrated to be involved in upregulating protein expression by switching between repressive and permissive chromatin states.²²⁴ In 1993 Broach *et al.* discovered²²⁵ that overexpression of Sirt2 in yeast cells (*Saccharomyces cerevisiae*) lead to a significant reduction in core histone acetylation, the process is depicted in *Figure 1*.

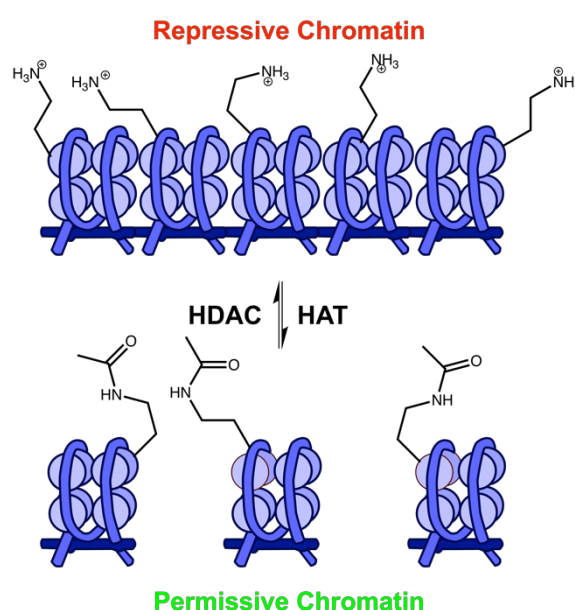


Figure 1 - Chromatin can exist in two states: permissive euchromatin, which allows transcription enzymes access to the DNA; and repressive heterochromatin, where nucleosomes are packed tightly, and transcription enzymes cannot access the DNA. The acetylation of histone lysines is associated with permissive chromatin, the acetylation can be removed by the action of a histone deacetylase (HDAC), while histone lysines can be acetylated by the action of a histone acetyltransferase (HAT).

2.1.2.0 Sirt2 Structure and Function

Sirtuins are members of the Class III Histone Deacetylases (HDACs), characterised by a 260 amino acid domain²²⁶ which is highly conserved across numerous species²²⁷ in all kingdoms of life,^{228, 229} including viruses.²³⁰ Sirtuins broadly possess NAD⁺ dependent deacylase^{231, 232} or ADP-ribosyltransferase activity,²³³ and function on numerous proteins.²³⁴

For the rest of this review I shall focus on Sirtuin2 (Sirt2). Sirt2 deacetylates proteins including: histones H3 (at K18 and K56) and H4 (at K16),²³¹ α -tubulin,²³⁵ p53²³⁶ and K-RAS²³⁷ amongst others.²³⁴ While the most studied target of Sirt2 is acetylation, other acyl groups such as benzoate²³⁸ and myristoleate²³⁹ can also be hydrolysed.

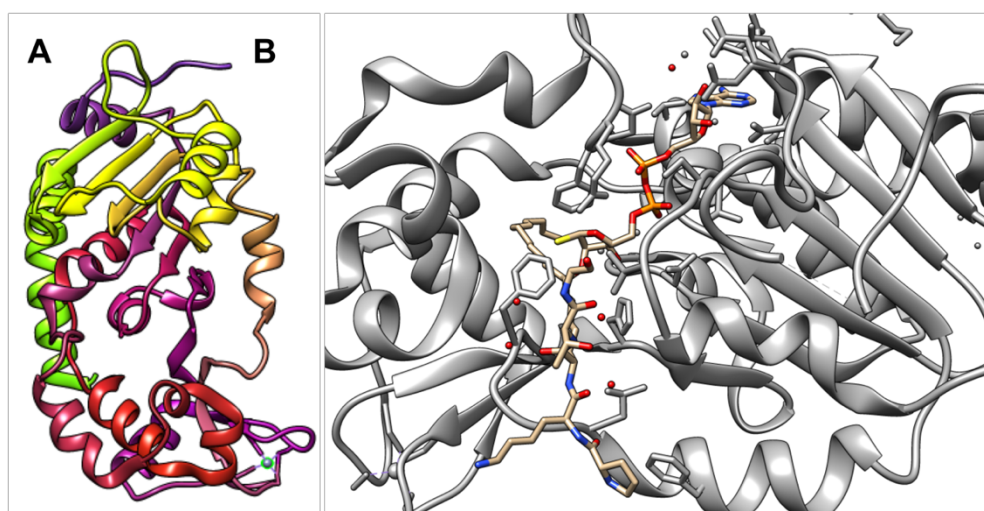


Figure 2 – Structures of Sirt2 **A)** Apo Sirt2 crystallography data taken from Pavletich et al²⁴⁰ and processed using UCSF Chimera **B)** The Sirt2 active site binding to myristoyl peptide crystallography data taken from Yoshida²⁴¹ and processed using UCSF Chimera

The first structure of Sirt2²⁴⁰ was solved by crystallography in 2001 in the *apo* form, shown in *Figure 2*. Numerous structures have since been solved by crystallography which include: Sirt2 complexed to cofactor fragments,²⁴² inhibitors^{242, 243} and peptides.^{239, 244, 245}

2.1.3.0 Sirt2 medical implications

While Sirt2 is found primarily in the cytosolic tubulin network,²³⁵ it is also be present in the nucleus being used for events such as global H4 K16Ac deacetylation during mitosis.^{246, 247} As Sirt2 is intimately involved with the cell cycle, both through the tubulins²⁴⁸ and histones,²⁴⁷ it should come as no surprise that it has been implicated in tumorigenesis, although evidence towards its role is often contradictory.

Sirt2 deficient mice have been demonstrated to grow sex specific tumours: females gain mammary tumours and males gain hepatocellular carcinoma.²⁴⁹ The *murine* experiment complemented *human* studies which found samples from corresponding *human* tumours deficient in Sirt2.²⁴⁹ Sirt2 deacetylation of BubR1, a mitotic checkpoint enzyme, has been shown to prevent cancer and aneuploidy by preventing its proteasomal degradation.²⁵⁰

In contrast to the antitumor properties in the prior paragraph, Sirt2 was shown to: deacetylate and therefore activate K-RAS,²⁴⁸ deacetylate p53 which reduces tumour suppression,²³⁶ and act as a tumour promoter in C6 glioma cells.²⁵¹

Sirt2 has also been associated in numerous non-cancer conditions: the deacetylation of FOXO3a transcription factor increases the levels of antioxidant manganese superoxide dismutase which lowers the concentration of reactive oxygen species in the cell,²⁵² and the deacetylation of glucose-6-phosphate dehydrogenase aids cell survival during oxidative stress by keeping NADPH concentrations steady.²⁵³

Sirt2 is directly involved in cell metabolism. The deacetylation of the gluconeogenesis rate limiting enzyme²⁵⁴ phosphoenolpyruvate carboxykinase I inhibits its proteasomal degradation.²⁵⁵ High levels of phosphoenolpyruvate carboxykinase I have been associated with the development of diabetes mellitus type 2.^{256, 257}

Aging central nervous system cells often show an accumulation of Sirt2,²⁵⁸ although it is not known if the high Sirt2 levels cause harm to nerve cells.²⁵⁹ Increased Sirt2 levels can cause the build-up of toxic intracellular protein aggregates by slowing protein degradation.²⁶⁰ The inhibition of Sirt2 has been shown to be neuroprotective with Alzheimer's disease associated α -synuclein,²⁶¹ and Huntington's disease.^{262, 263}

2.1.4.0 Sirt2 Mechanism

With being implicated in many diseases it follows that the mechanism of Sirt2 is well characterised. Sirtuins are Zn^{2+} dependent enzymes which employ NAD^+ as a cofactor. While Zn^{2+} is not directly involved in the deacylation mechanism, the catalytic core features both a Zn binding domain and a Rossmann fold.^{264, 265}

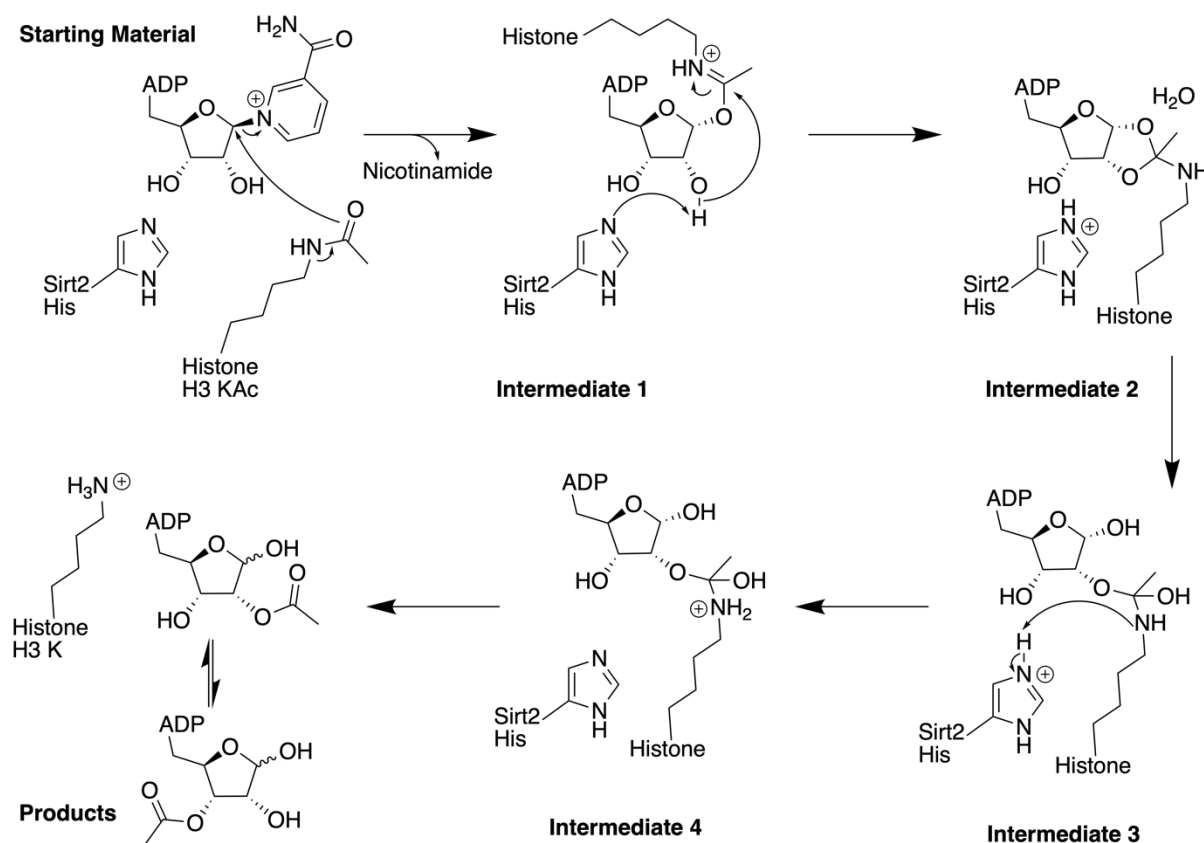


Figure 3 – The adenosine diphosphate ribose (ADPR)-peptidyl-imidate mechanism of Sirt2 mediated deacetylation: Amide activation of the acetyl group's carbonyl allows the nucleophilic attack of nicotinamide adenine dinucleotide NAD^+ with the ejection of nicotinamide to form intermediate 1. Sirt2's histidine (H135 in human Sirt2) then deprotonates the 2 hydroxyl of the ribose ring which attacks the imidate to form a quaternary intermediate, known as intermediate 2. The quaternary intermediate is cleaved via a carbocationic hydrolysis, to give intermediate 3. Intermediate 3 is protonated on the nitrogen atom of the amidoorthoester to give intermediate 4, which undergoes intramolecular decomposition to give the products of this reaction an unmodified lysine (K) residue and 2'-O-acetyl-ADP-ribose which can isomerise to 3'-O-acetyl-ADP-ribose.²⁶⁶

The currently accepted mechanism of Sirt2 mediated deacetylation is known as the ADPR-peptidyl-imidate mechanism, shown in *Figure 3*.²⁶⁷ Deacetylation of proteins

by sirtuin2 proceeds by activation of the amide oxygen from the nitrogen atom, which performs a nucleophilic attack on the ribose anomeric position expelling nicotinamide, to form the namesake ADPR-peptidyl-imidate, known as intermediate 1.²⁶⁶ An unproductive backwards reaction can occur to the ADPR-peptidyl-imidate. Alternatively, a histidine (H135 in *human* Sirt2)²⁶⁸ residue on Sirt2 can deprotonate the 2-position hydroxyl which cyclises to form a bicyclic aminoacetal, forming intermediate 2.²⁶⁶ Rapid solvolysis of aminoacetal yields the unmodified protein lysine residue and a hydroxylated bicycle (via the dioxonium ion).²⁶⁶ The hydroxylated bicycle then decomposes to exclusively give 2'-*O*-acetyl-ADP-ribose.^{269, 270} The acetate group can shift from the 2'-position to the 3'-position via intramolecular transesterification.²⁶⁶

Evidence for the ADPR-peptidyl-imidate mechanism was provided by Schramm *et al.*²⁶⁶ who compared chemically synthesised 1- β -*O*-acetyladenosine diphosphate ribose (1- β -OAADPR), which would result from the previously speculated mechanism,^{271, 272} to a sample of metabolite which results from the action of Sirt2 on the JB12 peptide (KKGQSTSRHKKA**Ac**LMFKTEG). Comparison by ¹H NMR spectroscopy, MS² mass spectrometry and HPLC showed that the product was not the expected 1- β -OAADPR, but an anomeric mixture of 2-OAADPR. Schramm also demonstrated the 2-OAADPR can convert to 3-OAADPR under ambient conditions.

Further proof²⁶⁶ that the reaction mechanism goes through ADPR-peptidyl-imidate mechanism came from performing the experiment in H₂¹⁸O buffer. It was demonstrated

that the ^{18}O label went to the acetate carbonyl oxygen, which shows that the original carbonyl bond is cleaved during the transformation.

Zheng *et al.*²⁷³ synthesised peptides containing *N*^ε-thioacetyl-lysine and treated them with sirtuins. The study was demonstrated that Sirt3 was incapable of cleaving the thioacetate. Later crystallographic evidence²⁷⁴ showed the formation of a thioimide, consistent with the ADPR-peptidyl-imide mechanism. Denu *et al* demonstrated²⁷⁵ that while thioacetate could be cleaved, the rate was much reduced, as the result of an increased C-S bond length.^{233, 274, 275} An increased bond length means the nucleophilic attack for (5-*exo-trig*)²⁷⁶ cyclisation must occur over a greater distance, which is responsible for the reduced rate.

2.1.5.0 Aims

While the mechanism of Sirtuin 2 is largely understood, the geometry around the quaternary centre formed has not been demonstrated experimentally (but has been predicted to be **R** computationally, *Figure 5*). I speculated that exposure of Sirt2 to a histone protein modified with the biological geometry around the quaternary centre would bind to it stronger than its diastereomer. Chemistry developed in previous chapters of this thesis gives us the means to have atomic control when posttranslationally modifying proteins, therefore I set out to synthesise both diastereomers of the predicted intermediates. Due to the biological aminoacetal instability I set out to replace the amino group with a methylene group, speculating that the amino group was not essential for the histone Sirtuin 2 interaction, shown in *figure 4*.

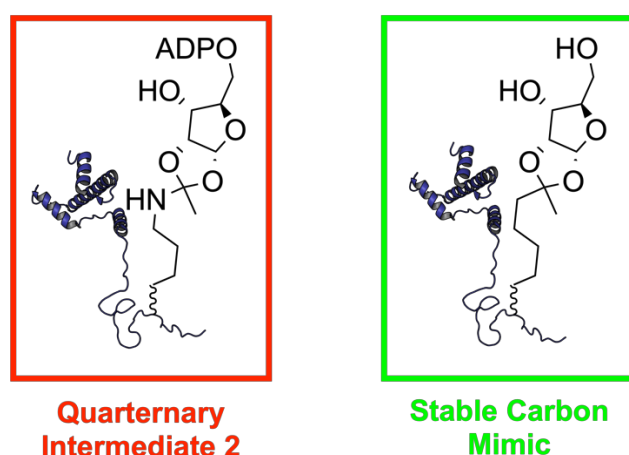


Figure 4 - The structures of the biological quaternary intermediate 2 and the stable carbon mimic prepared by chemical mutagenesis. Replaching the amine (NH) group with a methylene (CH₂) group will prevent the hydrolysis of intermediate 2, either by solvolysis or the action of Sirt2. The methodology developed in chapter 1 of this thesis gives us a method of introducing this modification diastereospecifically, with respect to the quaternary acetal centre.

A potential problem would be if the Sirt2 released the aminoacetal to allow a spontaneous degradation in reaction media, making it unlikely to bind the aminoacetal in order to prevent product inhibition of the enzyme. I also assumed that the amine group would not be essential for the histone deacetylation intermediate binding to Sirt2.

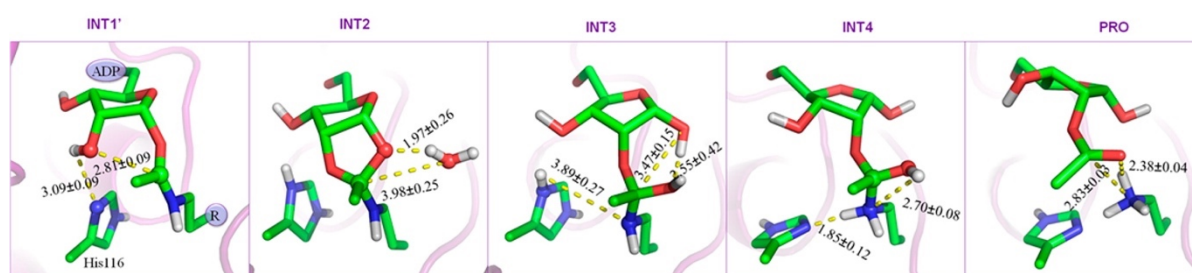


Figure 5 - Adapted with permission from²⁷⁷ Copyright 2013 American Chemical Society. Computational simulations of Sirtuin 2 intermediates, note the **R** geometry around INT2. The computation was carried out using a Born-Oppenheimer *ab initio* quantum mechanics/molecular dynamic simulation.

2.2.0.0 Results and Discussion

2.2.1.0 Synthesis of Aminoacetal Mimic

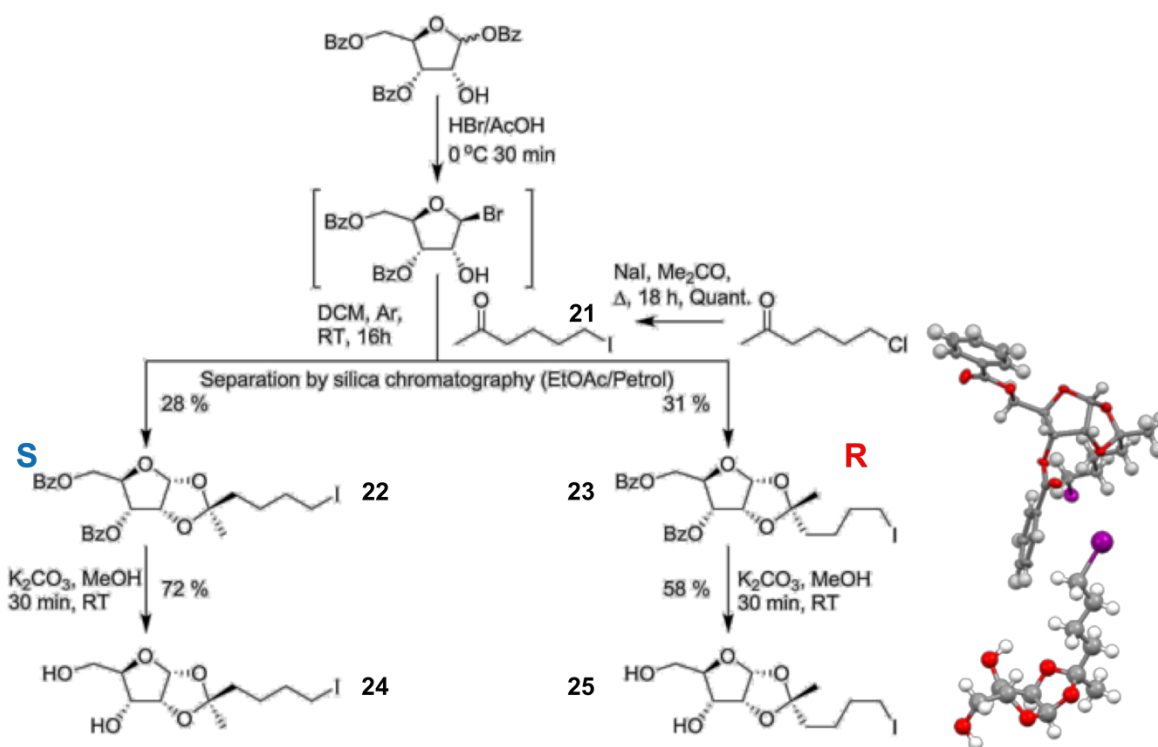


Figure 6 - Scheme of the divergent synthesis of radical precursors **24** and **25** which have **S** and **R** geometry around the quaternary centre respectively. 1,3,5-Tribenzoyl ribose is brominated to give intermediate **21**, which is treated with 1-iodohexan-5-one in dichloromethane (DCM). The reaction leads to a mixture of products including protected **S** and **R** (**22** and **23** respectively) which were separated using column chromatography on silica using an ethyl acetate/petrol gradient. Deprotection of **22** and **23** with potassium carbonate in methanol leads to desired products **24** and **25**.

The synthesis of the **S** and **R** intermediate reagent mimics was conducted divergently, the scheme is shown in Figure 6 experimental details are given in section 2.4.1. Divergent synthesis has two benefits, it produces both reagents from one

reaction and I do not need to worry about performing the reaction in such a way to produce only one diastereomer. 1-iodohexan-5-one was prepared by a Finkelstein reaction from the corresponding chloro- compound. The carbohydrate fragment of the molecule was prepared by brominated 2,3,5-tribenzoyl ribofuranose with HBr/AcOH. The brominated sugar proved difficult to isolate but a change could be observed by thin layer chromatography. Using the brominated ribose crude had the advantage that no further acid was required to catalyse the reaction, therefore when treated with the ketone the protected reagents were formed.

Separation of the diastereomers was accomplished by two rounds of silica chromatography. Initially a gradient 10 -> 30 % EtOAc in Petroleum Ether removed the excess of the ketone, then a more prolonged gradient between 10 -> 20 % EtOAc in Petroleum Ether was used to separate the diastereomers (experimental detail and chromatograms given in section 2.4.1.2). Deprotection with K_2CO_3 in MeOH gave reagents **24** and **25**. The identification of the reagents was aided by the fact that **R** reagent readily crystallised in both the protected and deprotected forms, while the **S** reagent, which eluted first was always an oil. Once I had confirmed the geometry around this stereocenter by X-ray crystallography, 1H NMR spectroscopy could be used to differentiate them as there is a large difference in the chemical shift of the acetal methyl group (R = 1.48 ppm, S= 1.28 ppm), shown in *Figure 7*. Experimental details for the synthesis of radical precursors can be found in section 2.4.1.

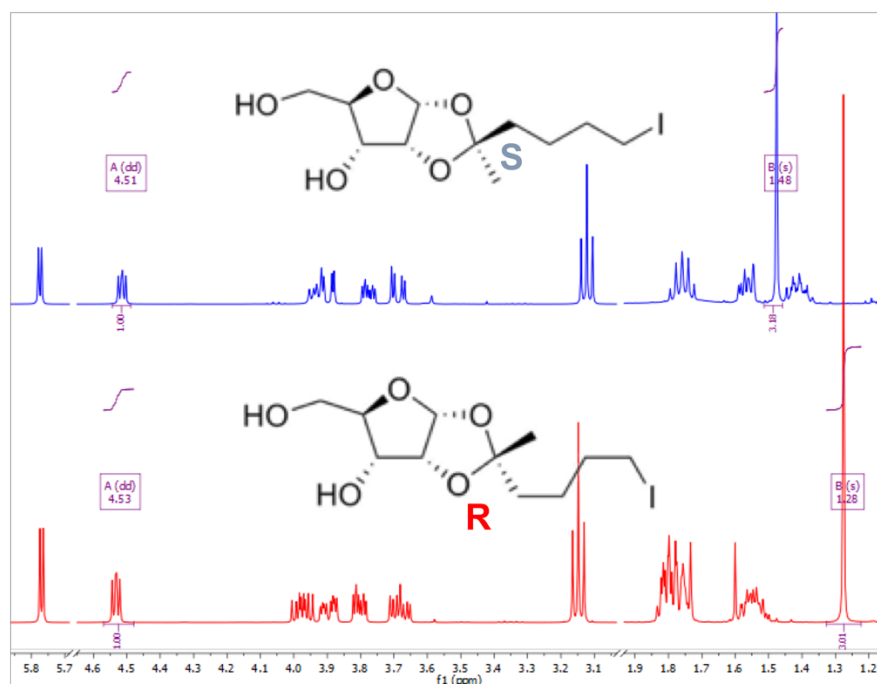


Figure 7 - The **R** and **S** isomers of the radical precursors can be readily distinguished by the ^1H NMR spectroscopy (400 MHz, CDCl_3) chemical shift of methyl group protons indicated in this figure. The **S** isomer has a shift of $\delta_{\text{H}} = 1.48$ ppm while the **R** isomer is $\delta_{\text{H}} = 1.28$ ppm.

2.2.2.0 Conjugation to Histones

Addition of **R** and **S** riboacetals **25** and **24** to histone H3 K18Dha and H4 K16Dha was accomplished with refreshingly little optimisation. However, in order to weigh the requisite 3 mg of each reagent it was found easier to first dissolve the compound in water then lyophilise. Lyophilisation allowed **S** reagent to form a solid (rather than a viscous oil) and lowered the density of both reagents, which simplified weighing the small quantities required. With enough effort, both reagents (3 mg of each) dissolved entirely in 200 μL of protein solution which was pipetted directly onto NaBH_4 . This

reproducibly gave the desired modified histones. The experimental details for glycoconjugation can be found in section 2.4.5 and seen in *Figure 8*.

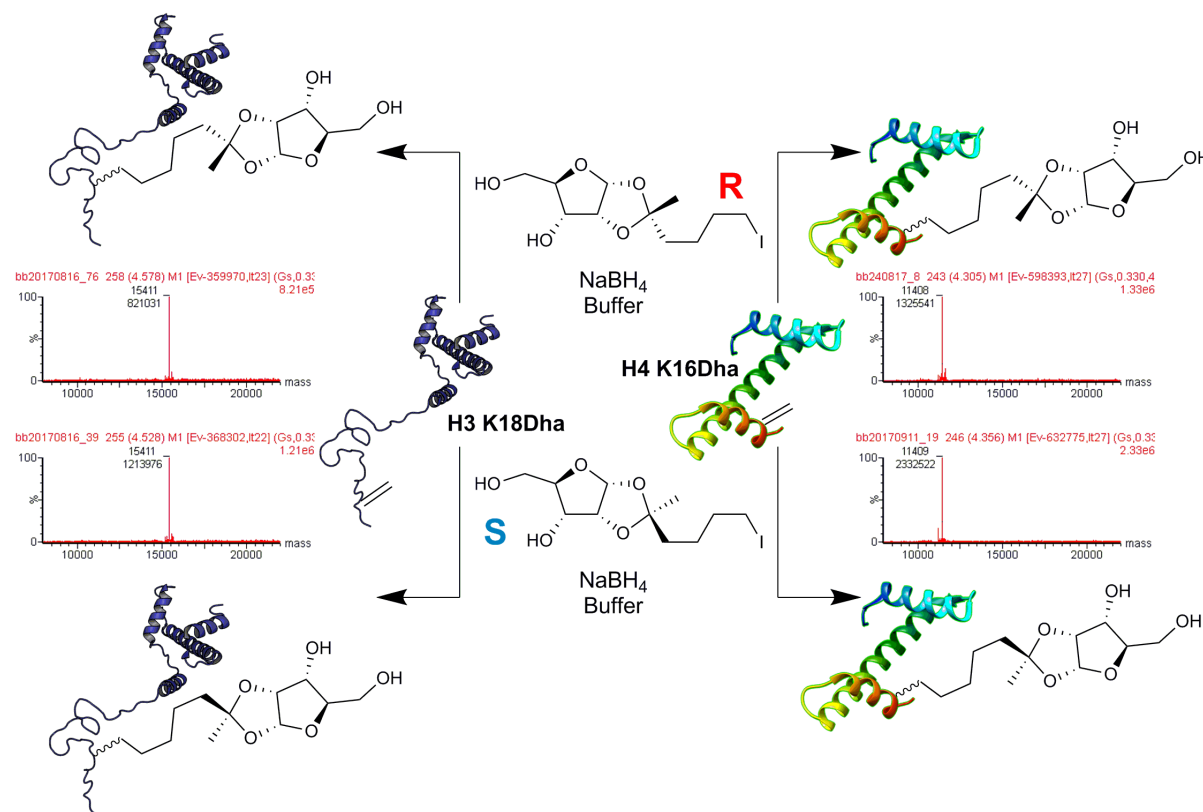
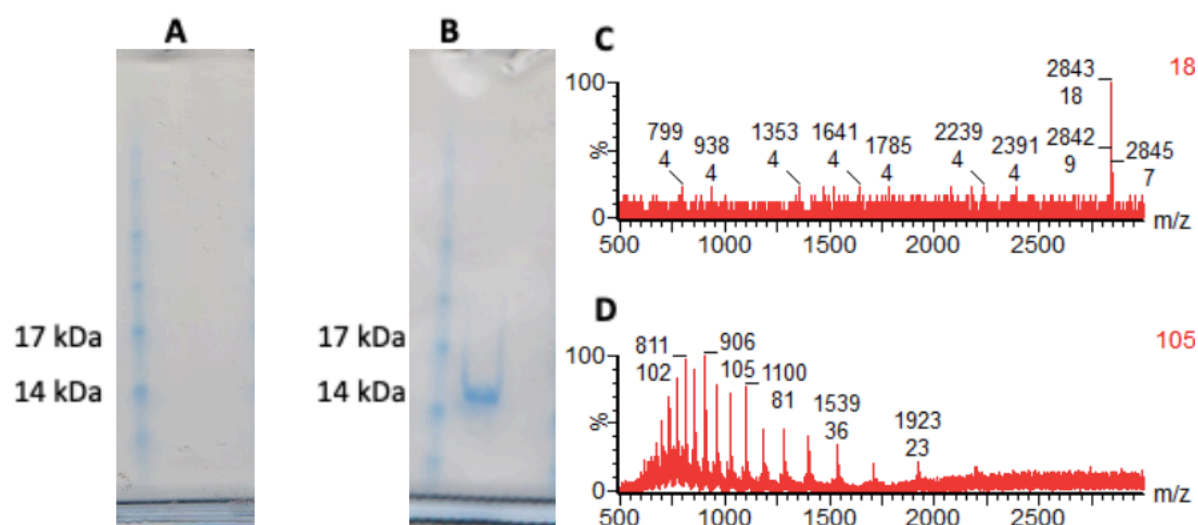


Figure 8 - Histones H3 K18Dha and H4 K16Dha can be reacted with the radical precursors and sodium borohydride to give the corresponding stable carbon intermediate mimics. Calculated Mass (H3): 15413 Observed Mass: 15411(R and S), Calculated Mass (H4): 11409 Observed Mass: 11409 (S) and 11408 (R)

2.2.3.0 – Histone Refolding

As histones are purified²¹⁷ and reacted¹⁵⁴ in the denatured state the protein must be refolded before any binding studies take place. Initially I attempted a method taken from nucleosome core reconstitution:²⁷⁸ 1 mg mL⁻¹ modified H3 in 5 M Gdm.HCl 500 mM NH₄OAc pH 6.0 was dialysed into 2 M NaCl 10 mM Tris 1 mM EDTA pH 7.5,

over three 4 h steps. The protein aggregated and precipitated, despite reports²⁷⁹ that this process is much slower. Centrifugation of the suspension yielded a clear colourless supernatant and a white pellet. The pellet was resuspended in denaturing buffer. Analysis by SDS-PAGE and mass spectrometry indicated that all of the histone H3 had precipitated, data from this experiment can be seen in *Figure 9*.



*Figure 9 - Comparison of the - Sodium dodecylsulfate poly(acrylamide) gel (SDS-PAGE) (4-12% Bis-Tris, MES running buffer, constant potential (200 V), InstantBlue™ development) and Mass Spectrum of the supernatant and pellet, when an attempt at refolding Histone H3 K18^R (~15 kDa) by exhaustive dialysis from H3 reaction buffer (Guanidinium Chloride 5 M, NH₄OAc 500 mM, pH 6.0) to refolding buffer (NaCl 2 M, Tris 10 mM, BME 5 mM, EDTA 1 mM, pH 7.5) precipitated; SDS-PAGE samples **A** and **B** represent the supernatant and pellet respectively. **C** and **D** represent the ion serise given by LC-MS analysis of the supernatant (**C**) and resuspended pellet (**D**). In both examples no protein can be seen in the supernatant, while protein is present in the pellet.*

I tried a rapid exchange into the refolding buffer by passing through a size exclusion column (Minitrap G-25). No precipitation was observed, and circular dichroism revealed the protein was folded. A second buffer exchange through the same column to fPBS was then performed in order to obtain a better low wavelength circular dichroism spectrum, the spectra can be seen in *Figure 10*. Experimental details are given in section 2.4.8.0.

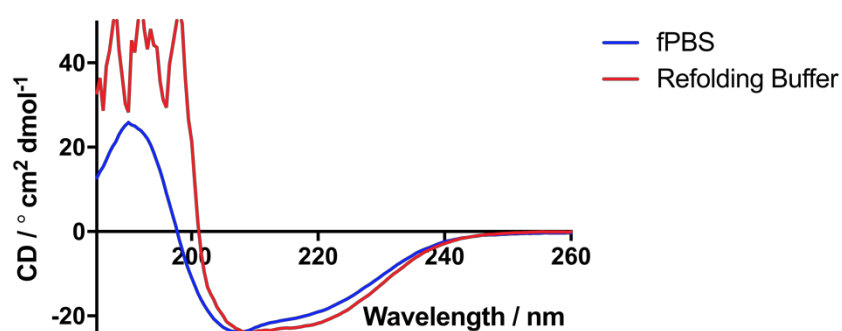


Figure 10 - Circular dichroism spectrum of refolded H3 K18R in fPBS (NaF 150 mM, Na_2HPO_4 10 mM, KH_2PO_4 1.8 mM, pH 7.4) and refolding buffer (NaCl 2 M, Tris 10 mM, BME 5 mM, EDTA 1 mM, pH 7.5). The low wavelength data is poor in refolding buffer due to 2 M Cl^- ions. Both samples are recorded at 0.2 mg mL⁻¹.

2.2.4.0 Expression of Sirt2

During my trial expression of Sirt2 in conditions reported Jung *et al.* I found the yield to be extremely low ($\sim 2 \text{ mg L}^{-1}$). As I intended to use a single batch of expressed protein for all subsequent experiments this low amount was not satisfactory.

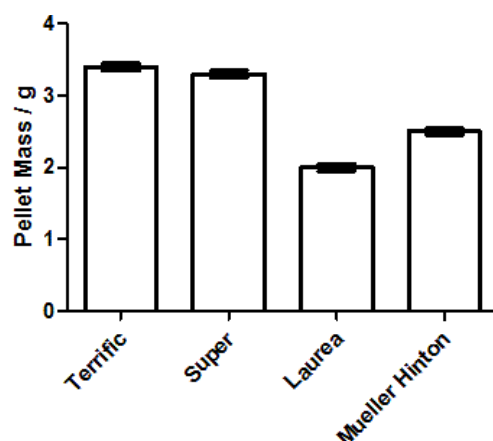


Figure 11 – Pellet Mass of 500 mL expressions of Human Sirt2 50-356 in BL21 (DE3) *E. coli* cells for 16 h at 291 K varying the broth employed. Cultures were induced by 0.3 mM IPTG at $OD_{600} = 0.6$. As expected the richer Terrific (Tryptone 12 g L^{-1} , Yeast Extract 24 g L^{-1} , K_2HPO_4 2.2 g L^{-1} , KH_2PO_4 2.2 g L^{-1} , Glycerol 8 mL L^{-1}) and Super broth (Tryptone 35 g L^{-1} , Yeast Extract 20 g L^{-1} , NaCl 5 g L^{-1} , 1 M NaOH 5 mL L^{-1}) yielded a higher pellet mass than less rich Laurea (Tryptone 10 g L^{-1} , Yeast Extract 5 g L^{-1} , NaCl 5 g L^{-1}) and Mueller Hinton (Beef infusion solids 2.0 g L^{-1} , Casein 17.5 g L^{-1} , Starch 1.5 g L^{-1}) broths. Error bars are determined from the precision of the balance used, after a single experiment.

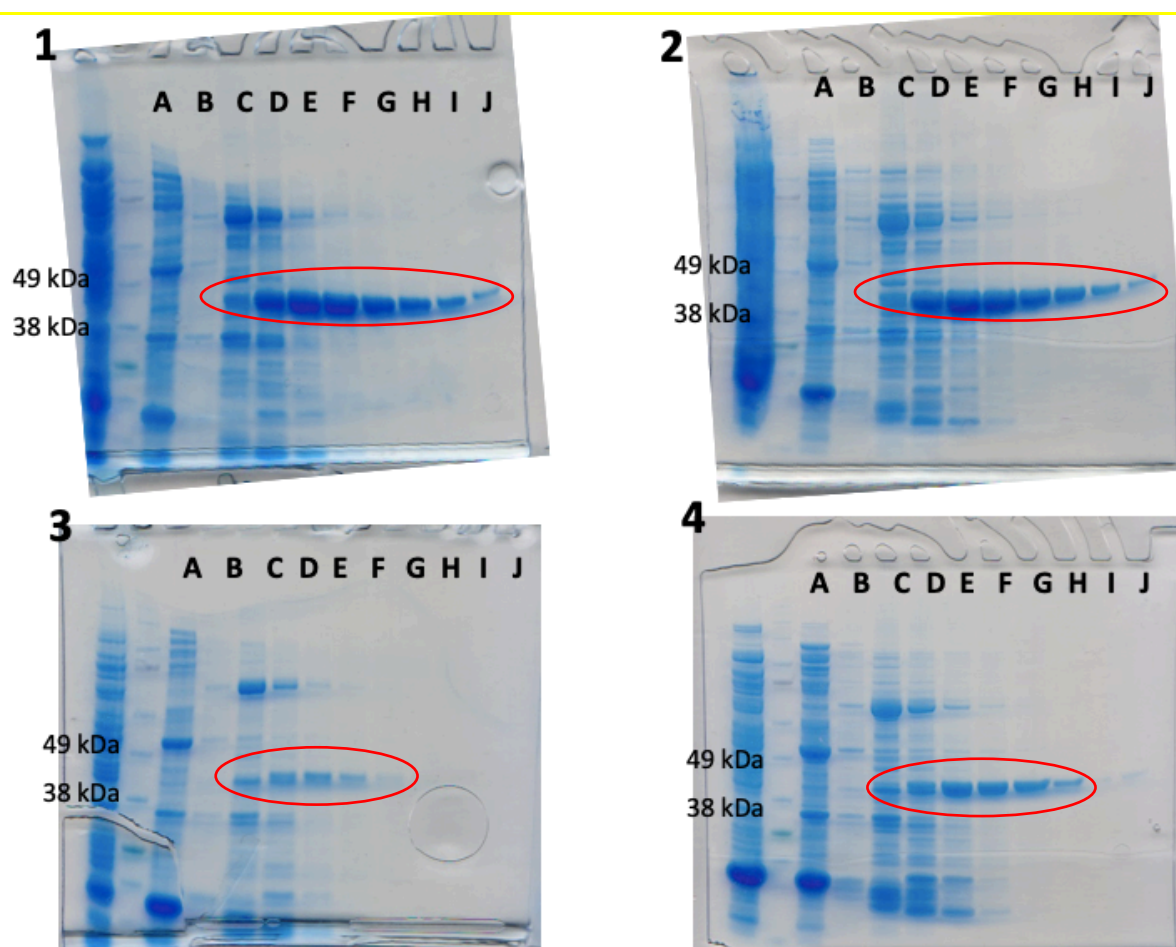


Figure 12 – Comparison of Sodium dodecylsulfate poly(acrylamide) gel analysis (SDS-PAGE) (4-12% Bis-Tris, MES running buffer, constant potential (200 V), InstantBlue™ development) of Ni affinity fast protein liquid chromatography (FPLC, A: NaCl 500 mM, Tris 50 mM, Glycerol 5%, BME 5 mM, Imidazole 25 mM pH of 7.5, B: A + 225 mM Imidazole pH 7.5, 5 CV A 10 CV B step gradient, HisTrap FF 1 mL) fractions from 500 mL expressions of Human Sirt2 50-356 (~37 kDa) in BL21 (DE3) *E. coli* cells for 16 h at 291 K varying the broth employed. Induced by 0.3 mM IPTG at $OD_{600} = 0.6$. As expected the richer Terrific (**1**) (Tryptone 12 g L⁻¹, Yeast Extract 24 g L⁻¹, K₂HPO₄ 2.2 g L⁻¹, KH₂PO₄ 2.2 g L⁻¹, Glycerol 8 mL L⁻¹) and Super broth (**2**) (Tryptone 35 g L⁻¹, Yeast Extract 20 g L⁻¹, NaCl 5 g L⁻¹, 1 M NaOH 5 mL L⁻¹) show more desired protein (circled in red) mass than less rich Laurea (**4**) (Tryptone 10 g L⁻¹, Yeast Extract 5 g L⁻¹, NaCl 5 g L⁻¹) and Mueller Hinton (**3**) (Beef infusion solids 2.0 g L⁻¹, Casein 17.5 g L⁻¹, Starch 1.5 g L⁻¹) broths.

A variety of different media were screened with the hope that a richer broth would allow the protein to be expressed more readily. Trial expressions on 500 mL scale

yielded a higher pellet mass for the richer Super Broth and Terrific Broth, depicted graphically in *Figure 11*. Sirt2 was purified from the lysate by Ni-affinity chromatography (SDS-PAGE samples can be viewed in *Figure 12*) and the concentration measured, graphed in *Figure 13*. While the concentration is not a perfect representation due to impurities still present, it can be used as a guideline.

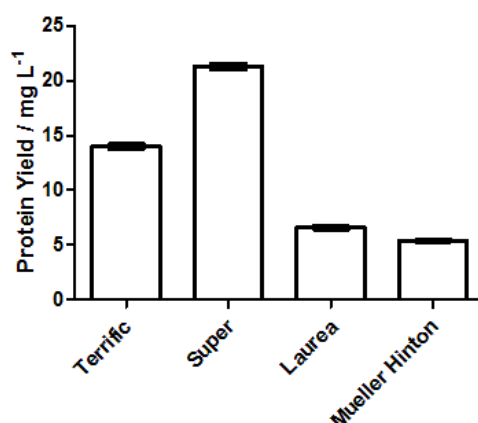
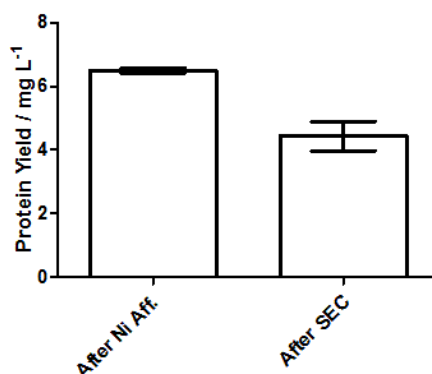


Figure 13 - Protein yield measured by A_{280} following Ni-affinity fast protein liquid chromatography (FPLC, A: NaCl 500 mM, Tris 50 mM, Glycerol 5%, BME 5 mM, Imidazole 25 mM pH of 7.5, B: A + 225 mM Imidazole pH 7.5, 5 CV A 10 CV B step gradient, HisTrap FF 1 mL) of 500 mL expressions of Human Sirt2 50-356 in BL21 (DE3) *E. coli* cells for 16 h at 291 K varying the broth employed. Induced by 0.3 mM IPTG at $OD_{600} = 0.6$. As expected the richer Terrific (Tryptone 12 g L⁻¹, Yeast Extract 24 g L⁻¹, K₂HPO₄ 2.2 g L⁻¹, KH₂PO₄ 2.2 g L⁻¹, Glycerol 8 mL L⁻¹) and Super broths (Tryptone 35 g L⁻¹, Yeast Extract 20 g L⁻¹, NaCl 5 g L⁻¹, 1 M NaOH 5 mL L⁻¹) yielded a higher protein concentration than less rich Laurea (Tryptone 10 g L⁻¹, Yeast Extract 5 g L⁻¹, NaCl 5 g L⁻¹) and Mueller Hinton (Beef infusion solids 2.0 g L⁻¹, Casein 17.5 g L⁻¹, Starch 1.5 g L⁻¹) broths. As some impurities still are present after this first round of purification the yields are only a guideline. Error bars are determined from the SEM of three measurements of the same samples.

As Super Broth showed the highest yield, I chose this for an 8 L expression, which, when purified by both Ni-affinity chromatography and size exclusion chromatography (SDS-PAGE analysis shown in Figure 15), gave an improved expression yield of

$\sim 4 \text{ mg L}^{-1}$, the expression yields after each step of chromatography are displayed graphically in *Figure 14*.



*Figure 14 - Protein yield of Human Sirt2 50-356 measured by A280 after purification by Ni-affinity fast protein liquid chromatography (FPLC, A: NaCl 500 mM, Tris 50 mM, Glycerol 5%, BME 5 mM, Imidazole 25 mM pH of 7.5, B: A + 225 mM Imidazole pH 7.5, 5 CV A 10 CV B step gradient, HisTrap FF 1 mL) and size exclusion chromatograph (s200 36/60, 150 mM NaCl, 25 mM Tris pH 8.0 buffer). The protein was expressed in super broth (Tryptone 35 g L⁻¹, Yeast Extract 20 g L⁻¹, NaCl 5 g L⁻¹, 1 M NaOH 5 mL L⁻¹). 8.0 L expression of Human Sirt2 50-356 in BL21 (DE3) *E. coli* cells for 16 h at 291 K Induced by 0.3 mM IPTG at $OD_{600} = 0.6$. Error bars are determined from the SEM of three measurements of the same samples.*

As the yield measured after Ni-affinity chromatography was significantly reduced to that seen when the expression was performed on a small scale I thought that cell lysis of a larger scale expression may be an issue. However, there was enough structured (CD spectrum provided in *Figure 16*) protein and to continue my studies therefore I did not probe this issue further. Experimental details can be found in section 2.4.3.1.

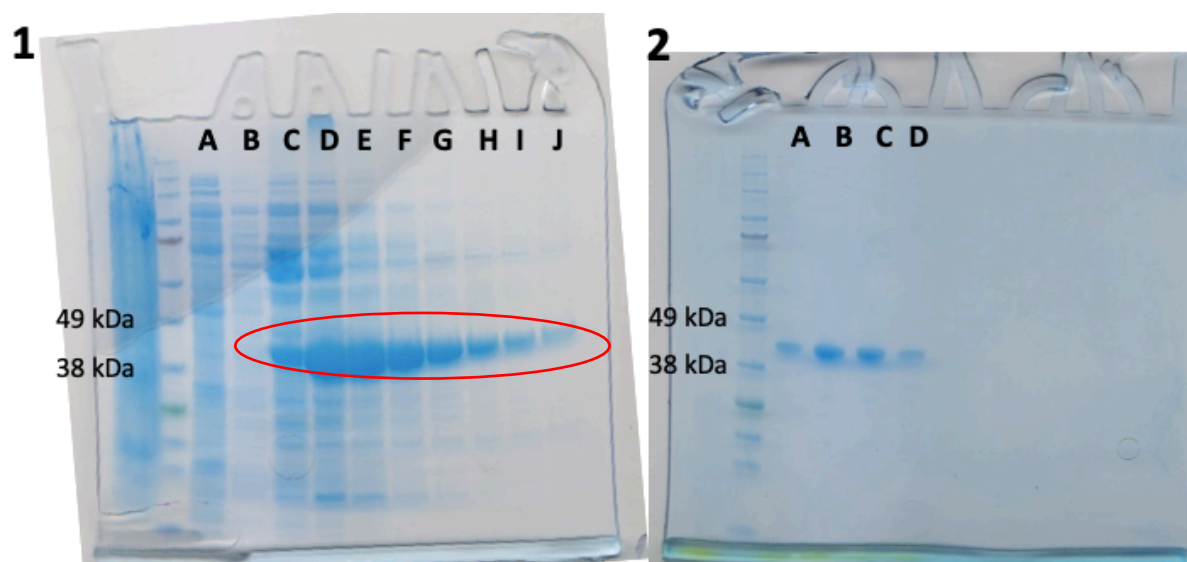


Figure 15 - Comparison of Sodium dodecylsulfate poly(acrylamide) gel (SDS-PAGE) analysis (4-12% Bis-Tris, MES running buffer, constant potential (200 V), InstantBlue™ development) which results after two dimensions of chromatography during purification of Human Sirt2 50-356 (~37 kDa) **1** Ni-affinity(FPLC, A: NaCl 500 mM, Tris 50 mM, Glycerol 5%, BME 5 mM, Imidazole 25 mM pH of 7.5, B: A + 225 mM Imidazole pH 7.5, 5 CV A 10 CV B step gradient, HisTrap FF 1 mL), Sirt2 has been highlighted with a red circle and **2** (s200 36/60, 150 mM NaCl, 25 mM Tris pH 8.0 buffer) from an 8.0 L expression of Human Sirt2 50-356. The protein was expressed in super broth (Tryptone 35 g L⁻¹, Yeast Extract 20 g L⁻¹, NaCl 5 g L⁻¹, 1 M NaOH 5 mL L⁻¹). 8.0 L expression of Human Sirt2 50-356 in BL21 (DE3) E. coli cells for 16 h at 291 K Induced by 0.3 mM IPTG at OD₆₀₀ = 0.6.

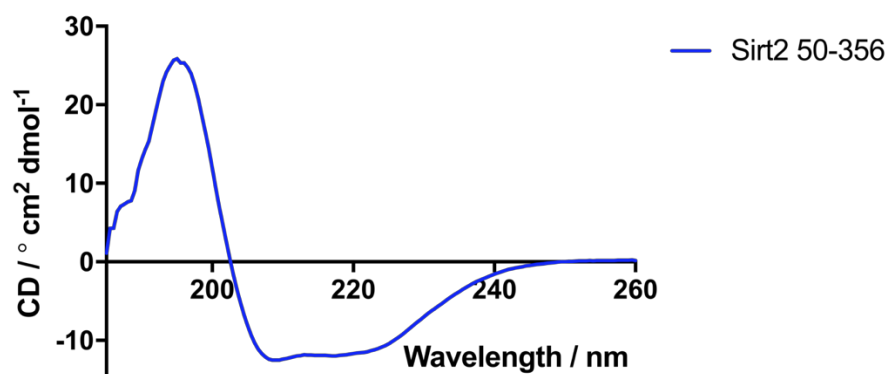


Figure 16 - Circular dichroism spectrum of Human Sirt2 50-356 in fPBS (NaF 150 mM Na_2HPO_4 10 mM KH_2PO_4 1.8 mM pH 7.4) at 0.2 mg mL^{-1}

2.2.5.0 Binding by Surface Plasmon Resonance

My initial strategy to study the binding between Sirt2 and histone H3 was to use surface plasmon resonance spectroscopy²⁸⁰ on the Biacore platform. By binding an α -Poly(His) antibody to the Biacore CM 5 chip (sonogram given in section 2.4.6.1), I immobilised His₆-tagged Sirt2. Immobilisation of Sirt2 allowed us to flow over solutions of the modified H3 variants. A diagram of this is shown in *Figure 17*.

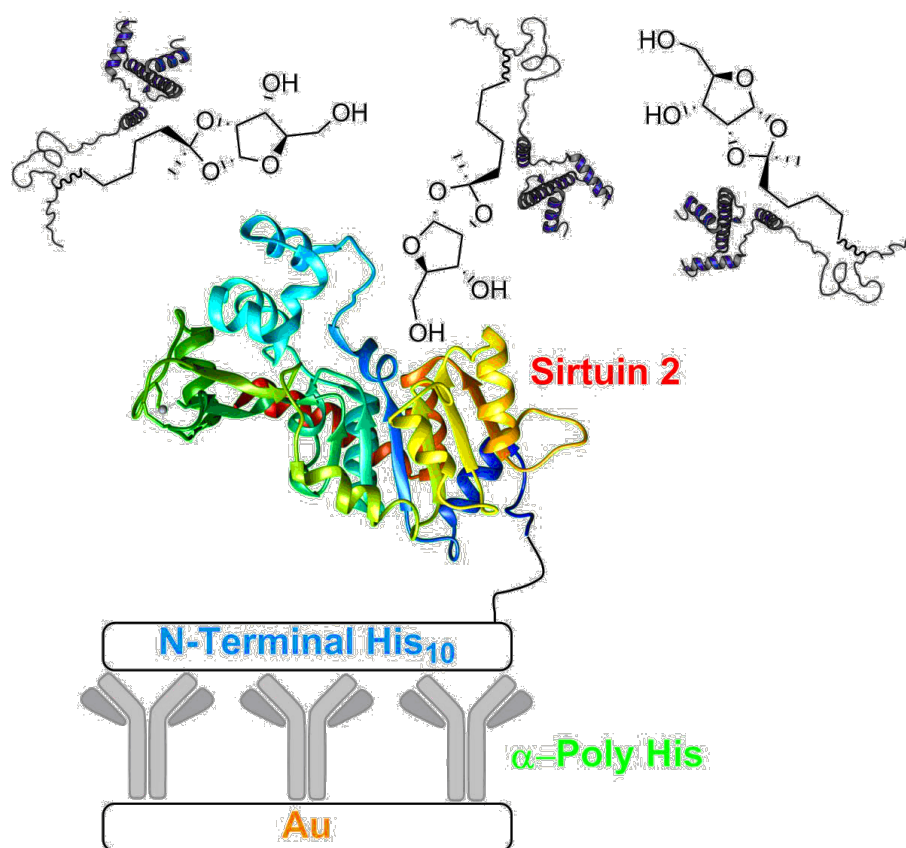


Figure 17 – Diagram of surface plasmon resonance set up - His tagged Sirt2 immobilised on an α -Poly(His) antibody then modified histone will be flowed over and the response measured.

While the antibody coupling proved successful, I found that the Sirt2 was only bound transiently to the chip (sonogram given in section 2.4.6.2). The transient binding alone would not be cause for concern, provided it behaves predictably, but I also found a larger issue: The positive H3 binds to the negative carboxy methyl dextran resin, which forms part of the Biacore chip, a diagram is shown in *Figure 18* and sonogram given in section 2.4.6.3.

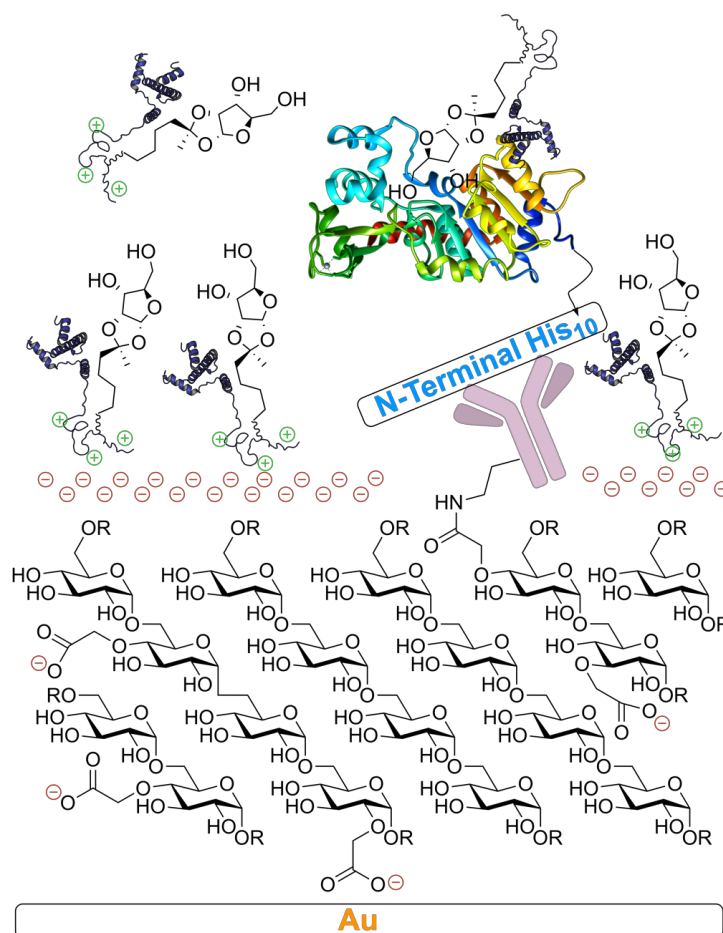


Figure 18 - Diagram displaying why attempts at surface plasmon resonance experiments failed as we did not account for the negative carboxy methyl (CM) dextran resin binding to the positive histone protein

I tested a CM4 chip (sonograms given in sections 2.4.6.6 and 2.4.6.7) – which has lower levels of carboxymethyl functionalisation; however, this proved unsuccessful as well. I found raising [NaCl] to 1 M prevented unspecific binding, (sonogram given in section 2.4.6.8) however this is far from physiological conditions and no binding between histone variants and Sirt2 could be observed (sonogram given in section 2.4.6.9). Lowering the pH had little effect on unspecific binding (sonogram given in section 2.4.6.8).

I tried to remove the negative charge of the resin by reacting iteratively with, EDC and NHS, then carbohydrazide. A different coupling strategy was then required, so I chose to oxidise the antibody glycan with periodate and then perform reductive amination with the carbohydrazide functionalised CM4 chip. A scheme for the reductive amination strategy is shown in *Figure 19* and sonograms given in section 2.4.6.10.

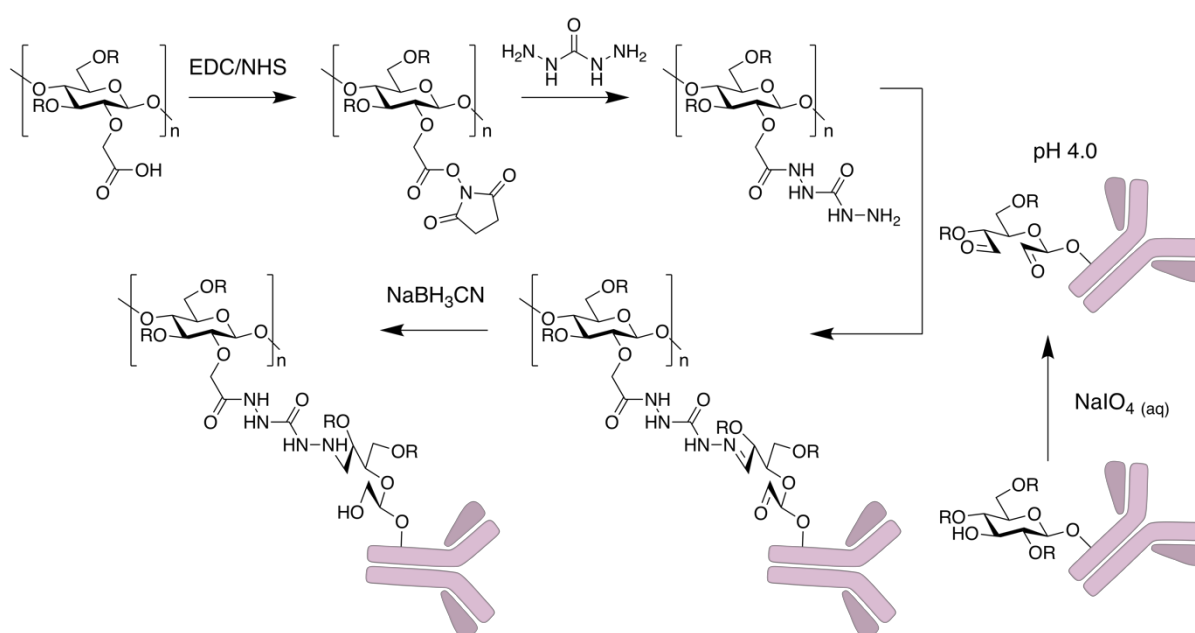


Figure 19 – Attempted strategy for a positively charged resin - The carboxymethyl resin is iteratively treated with 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and N-Hydroxysuccinimide (NHS) followed by carbohydrazide for three cycles then a reductive amination reaction is used to conjugate an antibody with oxidised glycans to the carbohydrazide

After three rounds of coupling, I added the antibody. The H3 still bound to the resin in an unspecific manner, I imagine this is because I have only managed to functionalise the surface layer of the resin and negative charge remains under the surface. At this

point I felt I had exhausted my options for SPR spectroscopy, short of reversing my strategy and immobilising the histone, so a different approach was taken. All sonograms can be found in section 2.4.6.

2.2.6.0 Binding by Size Exclusion Chromatography

With no luck using SPR; my next approach was monitor if a new peak, with a larger mass emerges in a size exclusion chromatogram²⁸¹ when Sirt2 and a modified H3 sample are mixed. Control injections of 40 μ L of 10 μ M Sirt2, H3 **S** and H3 **R** were carried out followed by a 1:0.5, 1:1 and 1:1.5 mix of Sirt2 and of a H3 riboacetal. All injections were monitored for absorption at 214 nm and 280 nm. A multi angle light scattering detector (MALS)²⁸² was connected inline to the chromatography system.

Inspection of samples of **R** and **S** control samples show little difference between them, these can be seen in *Figure 20* with experimental details being given in sections 2.4.7.2 and 2.4.7.3. Inspection of the 1:1 samples showed a new peak appearing in only the **R** sample, this can be seen in *Figure 21*. Analysis of all three **R** samples showed that the peak gets larger as more histone is added, which is to be expected and can be seen in *Figure 22*. Although a small peak is visible with 1.5 eq of the **S** sample, this is barely above the baseline, this can be seen in *Figure 23* and section 2.4.7.

Unfortunately, the MALS signal was too weak to provide any real insight into the mass of the complex formed, which is explained in more detail in *figure 24*.

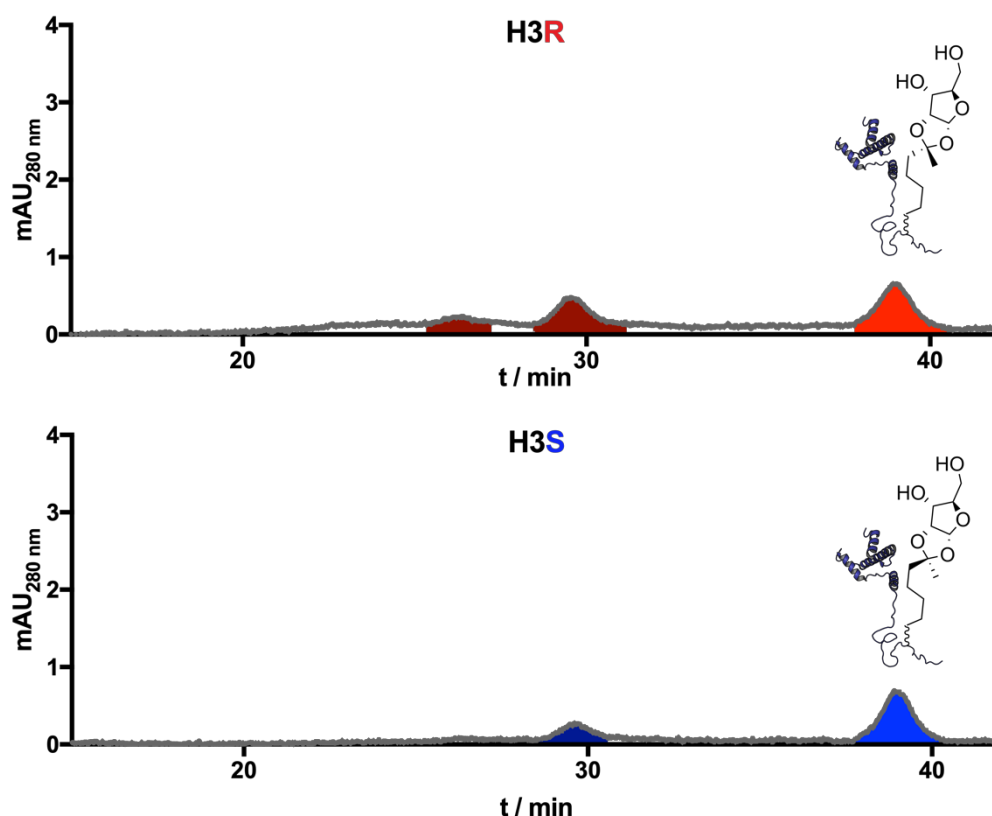


Figure 20 – Analytical size exclusion chromatograms (280 nm, S200 10/300, in HBS-P (Hepes 10 mM pH 7.4, NaCl 150 mM, p20 surfactant 0.005 % v/v)) to compare control samples of H3 K18S and H3 K18R. Both chromatograms are near identical with the main protein peak being visible at ~40 mins being coloured **red** for **R** and **blue** for **S** and higher order oligomers being seen at lower time coloured **dark red** for **R** and **dark blue** for **S**. aSEC key: **Blue**: Histone H3 K18S monomer, **Dark Blue**: Histone H3 K18S oligomers, **Red**: Histone H3 K18R monomer, **Dark Red**: Histone H3 K18R oligomers, **Yellow**: Sirt2, **Green**: Sirt2/Histone H3 K18S complex and **Orange**: Sirt2/Histone H3 K18R complex.

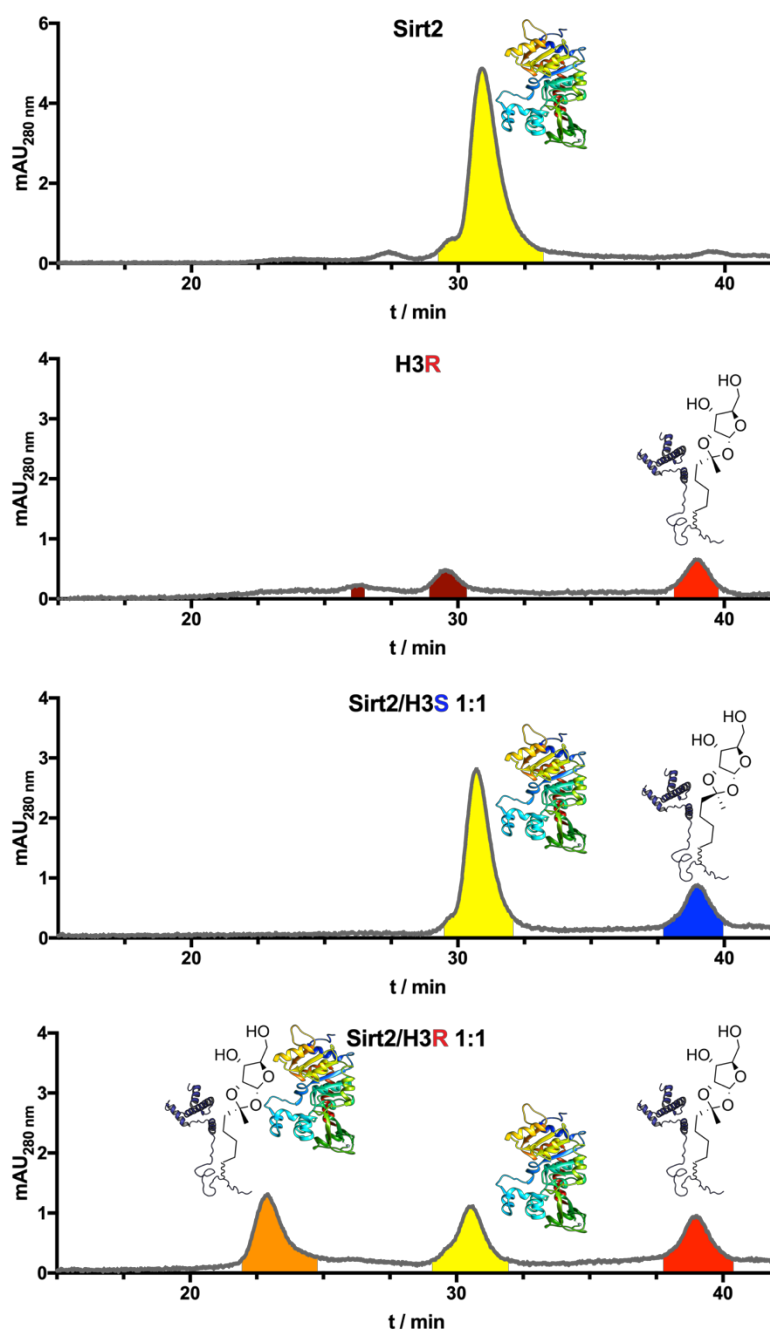


Figure 21 – Analytical size exclusion chromatograms (280 nm, S200 10/300, in HBS-P (Hepes 10 mM pH 7.4, NaCl 150 mM, p20 surfactant 0.005 % v/v)) showing (top to bottom) Sirt2, H3 K18R, Sirt2/H3 K18S 1:1 and Sirt2/H3 K18R 1:1. In the chromatogram for Sirt2/H3 K18R 1:1 a new higher molecular mass peak appears at ~22 mins which is not present in any of the previous samples, likely a result of a complex formation between Sirt2 and H3 K18R. aSEC key: Blue: Histone H3 K18S monomer, Dark Blue: Histone H3 K18S oligomers, Red: Histone H3 K18R monomer, Dark Red: Histone H3 K18R oligomers, Yellow: Sirt2, Green: Sirt2/Histone H3 K18S complex and Orange: Sirt2/Histone H3 K18R complex.

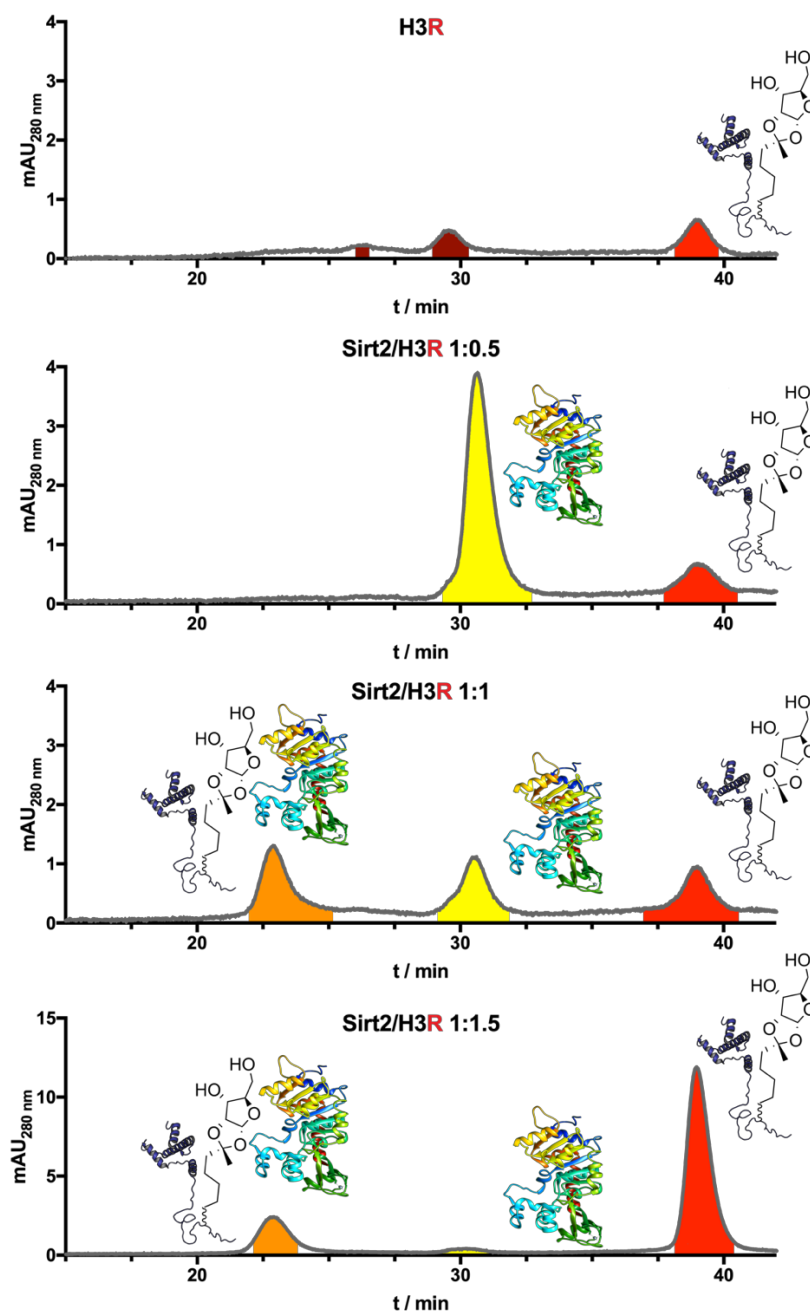


Figure 22 – Analytical size exclusion chromatograms (280 nm, S200 10/300, in HBS-P (Hepes 10 mM pH 7.4, NaCl 150 mM, p20 surfactant 0.005 % v/v)) to compare chromatograms of H3 K18R with varying ratios of Sirt2 showing (top to bottom) Sirt2, H3 K18R, Sirt2/H3 K18R 1:0.5, Sirt2/H3 K18R 1:1, Sirt2/H3 K18R 1:1.5. The new peak which forms at ~22 mins is responsive to increasing amounts of H3 K18R, as is the Sirt2 peak at ~30 mins aSEC key: Blue: Histone H3 K18S monomer, Dark Blue: Histone H3 K18S oligomers, Red: Histone H3 K18R monomer, Dark Red: Histone H3 K18R oligomers, Yellow: Sirt2, Green: Sirt2/Histone H3 K18S complex and Orange: Sirt2/Histone H3 K18R complex.

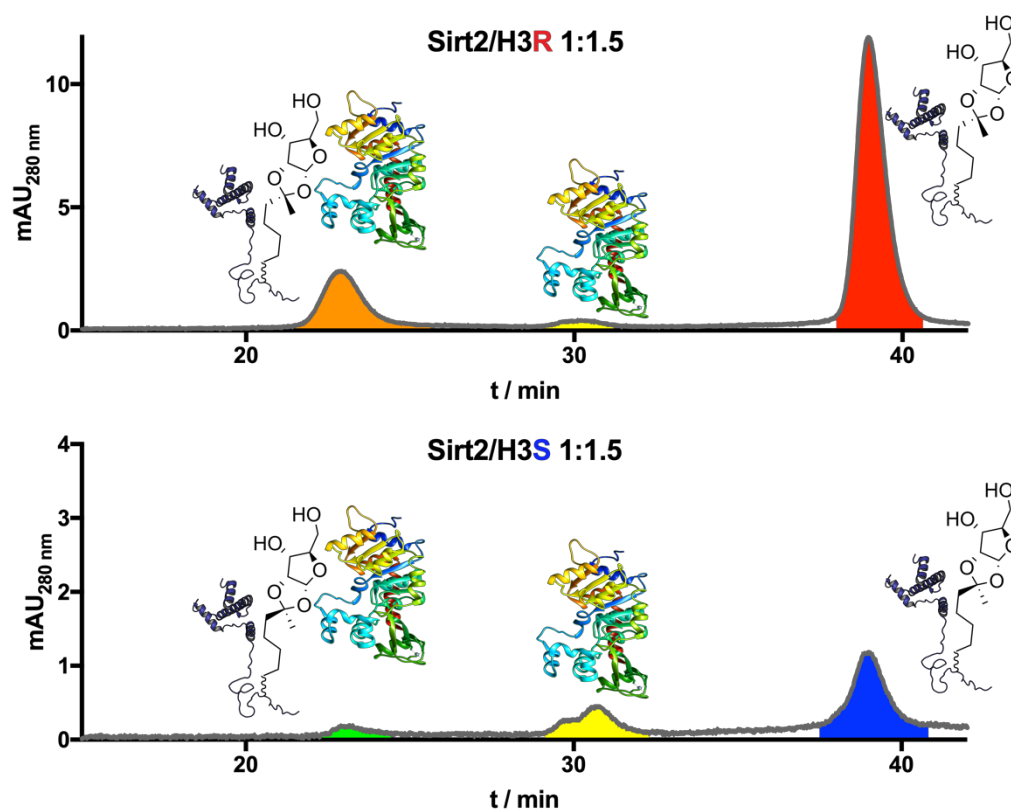
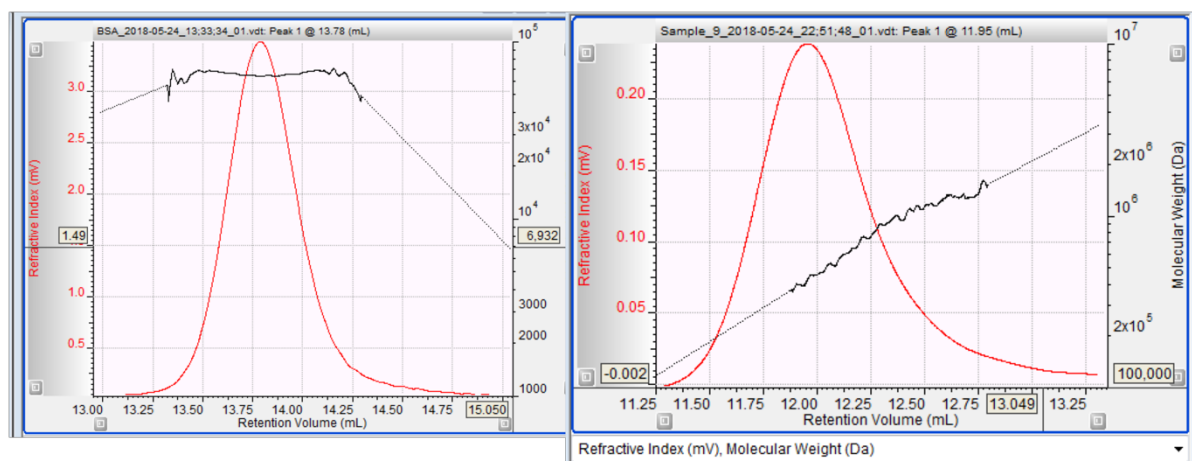


Figure 23 - Analytical size exclusion chromatograms (280 nm, S200 10/300, in HBS-P (Hepes 10 mM pH 7.4, NaCl 150 mM, p20 surfactant 0.005 % v/v) shown) to compare chromatograms of Sirt2/H3 K18^R 1:1.5 and Sirt2/H3 K18^S 1:1.5. Note the complex peak in the **S** chromatogram is greatly suppressed when compared to **R**. Likewise the Sirt2 peak is larger in the **S** when compared to **R**. The monomeric histone peak appears much larger in the **R** chromatogram, I suspect this is because having a complexing agent such as Sirt2 in solution will retain the Histone H3 K18^R in its monomeric state, rather than forming oligomers. aSEC key: Blue: Histone H3 K18^S monomer, Dark Blue: Histone H3 K18^S oligomers, Red: Histone H3 K18^R monomer, Dark Red: Histone H3 K18^R oligomers, Yellow: Sirt2, Green: Sirt2/Histone H3 K18^S complex and Orange: Sirt2/Histone H3 K18^R complex.



Good MALS Signal

Bad MALS Signal

Figure 24 - Comparison of a good multiangle light scattering (MALS) signal provided from our bovine serum albumin (BSA) calibration and the bad data which we obtained, the MALS signal needs to be horizontal in order to obtain a good mass approximation, when it is at a diagonal the mass is approximated over several log units.

2.3.0.0 Conclusion

Considering the computational²⁷⁷ evidence of Shi *et al.* and the physical evidence provided in this thesis, I think a fair conclusion would be that the **R** intermediate is the biologically relevant diastereomer. Due to time constraints this is all the evidence I am currently able to provide. The next step in this investigation is to take the qualitative evidence from this thesis and compliment it with quantitative data. It would also be useful to perform experiments on reconstituted nucleosomes, rather than histone monomers, as *in vivo* Sirt2 is functional against the histone tail when part of the chromatin structure.²³⁴ As the aSEC experiments cannot give quantitative results, I do not think it is worth pursuing further.

The first approach I wish to employ for quantification is NMR spectroscopy. By enriching the small molecule fragment of the intermediate2 mimics **24** and **25** (to give H3**S** and H3**R** respectively) with ¹³C on the methyl group (structure shown in Figure 25) I can avoid the costly endeavour of isotopically labelling the complete protein. A titration²⁸³ between each ¹³C-labeled intermediate2 mimic and Sirt2 could then be undertaken, from which I could determine the binding constant of each diastereomer to Sirt2. A second approach using NMR spectroscopy would be to replace a methylene group in the Lys-side chain mimic fragment of the molecule with C¹⁹F₂ (Figure 25). ¹⁹F NMR has the advantages of having no background signal from elsewhere in the protein, 100 % natural abundance¹⁵⁵ and is very sensitive to its micro environment.¹⁸¹

I would not recommend directly labeling the methyl group, which I am assuming is key for these interactions due to major electronic differences between CF_3 and CH_3 groups.

Isotope Tagged Histone H3 Variants for Sirt2 Intermediate2 Binding Quantification

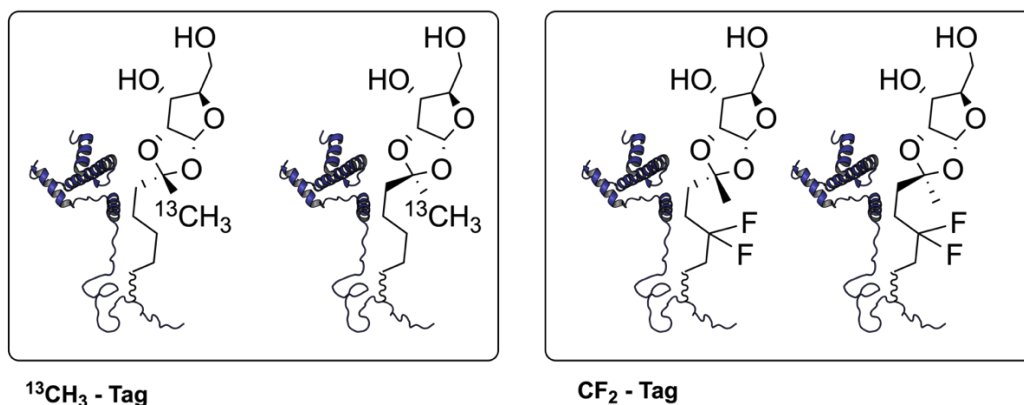


Figure 25 – Diagrams of Isotopically labelled Histone H3 variants enriched in specific NMR active nuclei, which could be used in titration with Sirt2 to determine dissociation constants (K_d) for each intermediate 2 mimic diastereomer.

A simpler but still quantitative technique which could be employed is to exploit the Sirt2 binding of the intermediate2 mimic H3 R and demonstrate kinetically that H3 R is much stronger inhibitor competitive inhibitor of Sirt2 than H3 S . Analysis of both diastereomers as inhibitors has the advantage that no novel molecules need to be prepared and that much simpler equipment is required than the NMR methods suggested above. Isothermal titration calorimetry²⁸⁴ could be employed to determine dissociation constants (K_d) and the enthalpy of binding (ΔH_b), but typically require a very large protein, perhaps unsuitable for this application.

Quantitative interaction proteomics using mass spectrometry^{146, 285} could be employed by treating immobilised nucleosome samples containing one of the three histone H3 variants (intermediate2 mimics H3^R and H3^S and wt histone H3) with cellular lysate (either whole cell or nuclear), followed by tryptic digest and analysis by mass spectrometry. Use of this quantitative interaction proteomics method has the advantage of linking the modification to cells. A similar unpublished technique pioneered by my colleague B. Josephson allows the incorporation of synthetic histones into the chromatin of mammalian cells. We plan to use this technique to introduce epitope tagged H3^R and H3^S histones into the chromatin of HeLa cells then (after *in vivo* cross linking and purification) use interaction proteomic techniques to analyse if Sirt2 is enriched with isomer compared to wild-type histone and the opposite diastereomer, a short workflow can be seen in *Figure 26*.

As it appears that we can isolate the Histone ^R/Sirt2 complex we plan to send a sample of this isolated complex to the Jung group in order to conduct crystallisation screens a scheme can be seen in *Figure 26*. While I expect difficulty in both obtaining crystals and determining their structure of a histone H3/Sirt2 complex by x-ray crystallography, a simpler method would be to prepare peptide mimics of intermediate 2 and co-crystallise as has been performed in literature with different peptides and Sirt2.^{245, 286, 287}

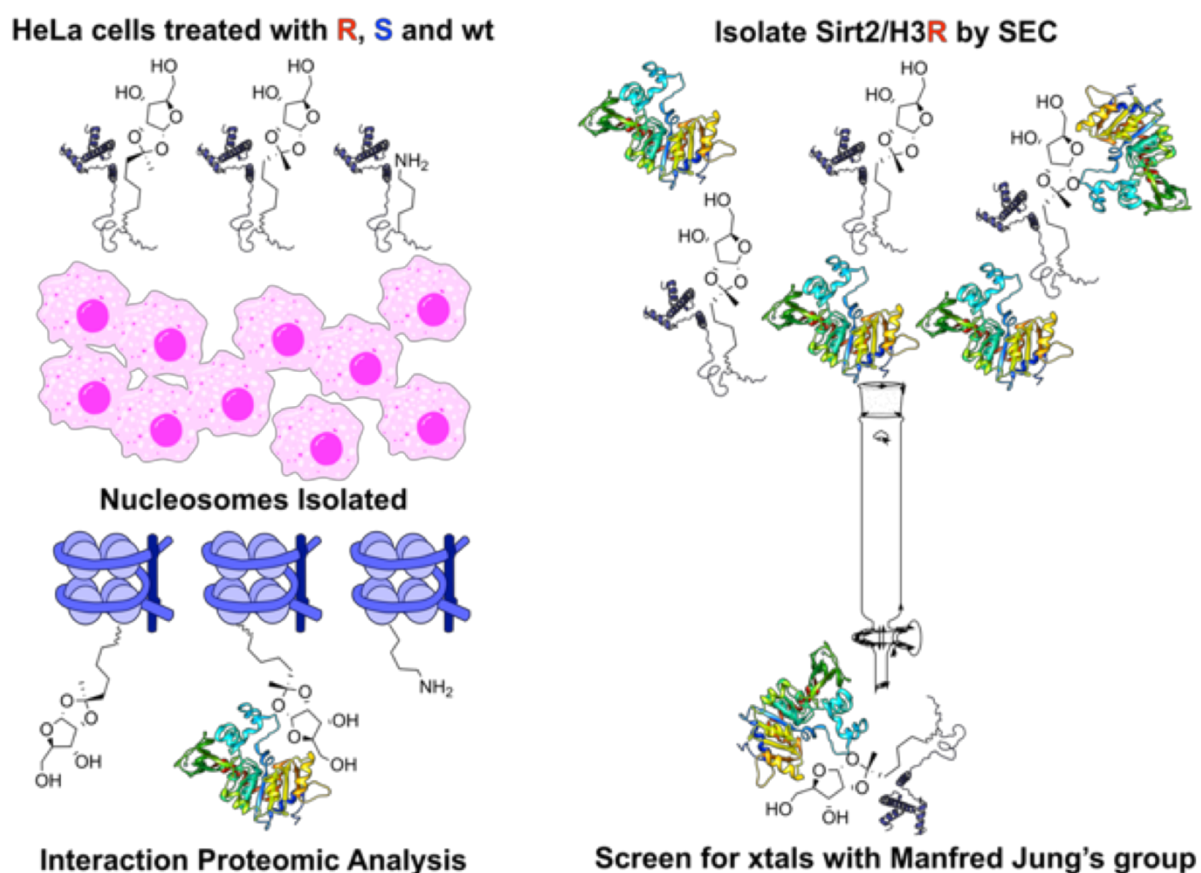


Figure 26 – Plans for the future experiments - left shows a new technique which will be performed in collaboration with B. Josephson, HeLa cells will be treated with epitope tagged histone H3 K18^R, histone H3 K18^S or histone wt, following an incubation period the histones will be isolated by immunoprecipitation and interaction proteomics used to detect enrichment of Sirt2. Right depicts a second plan where I plan to isolate the Sirt2/H3^R complex (as detailed in section **2.2.6.0** of this these) via size exclusion chromatography then send a sample to the Jung group, who plan to screen for crystals.

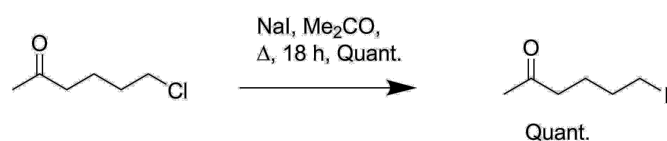
The work presented in this chapter has been used for a specific application, however more broadly I have developed a new technique to determine the geometry of unstable reaction intermediates. Work in this thesis demonstrates that by bypassing the ribosomal filter using chemical techniques we can construct effective and stable mimics

of otherwise inaccessible protein reaction intermediates. Furthermore, the proteins produced can then be employed to discover new biological phenomena.

2.4.0.0 Experimental

2.4.1.0 Compound Characterisation

2.4.1.1 1-Iodohexan-5-one 21



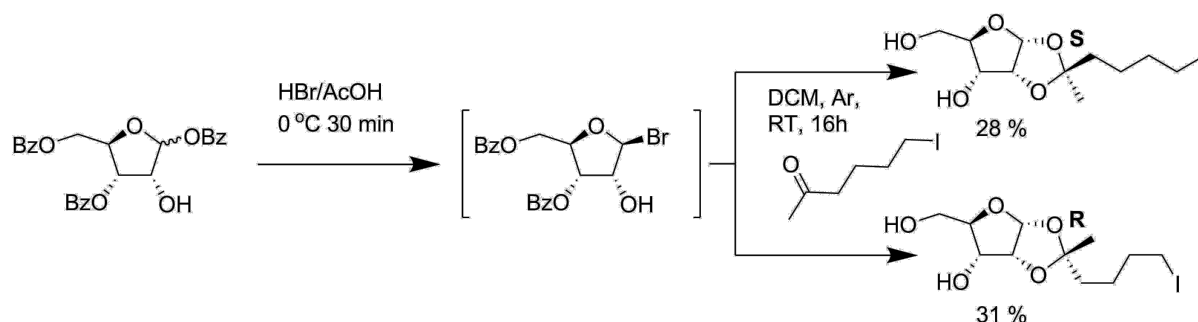
To a solution of 1-chlorohexan-5-one (5.00 g, 37.1 mmol) in molecular sieve dried acetone (150 mL) sodium iodide (8.35 g, 55.6 mmol) was added. The reaction mixture was heated to reflux for 18 h. The two compounds are inseparable by TLC, so it is important that this reaction goes to completion. The reaction mixture was allowed to cool, then partitioned between DCM (450 mL) and brine (450 mL). The aqueous phase was extracted once with DCM (100 mL). The combined organic layers were washed twice with brine (250 mL), dried over MgSO₄ then concentrated under reduced pressure to yield the title compound as a thin pale-yellow oil (8.38 g, 37.1 mmol, quant.), which was used without further purification.

¹H NMR δH (400 MHz, D₂O) 3.17 (2H, t, *J* = 6.9 Hz, CH₂I), 2.45 (2H, t, *J* = 7.2 Hz, H_γ), 2.14 (3H, s, H_α), 1.83-1.81 (2H, m, H_ε) and 1.68-1.63 (2H, m, H_δ)

^{13}C NMR δC (100.6 MHz, D_2O) 208.2 ($\text{C}\beta$), 42.3 ($\text{C}\gamma$), 32.8 ($\text{C}\epsilon$), 30.0 ($\text{C}\alpha$), 24.6 ($\text{C}\delta$) and 6.2 ($\text{C}\zeta$)

MS^+ 249.0 ($\text{M} + \text{Na}^+$)

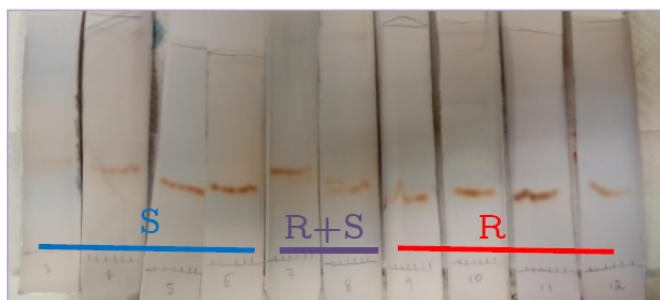
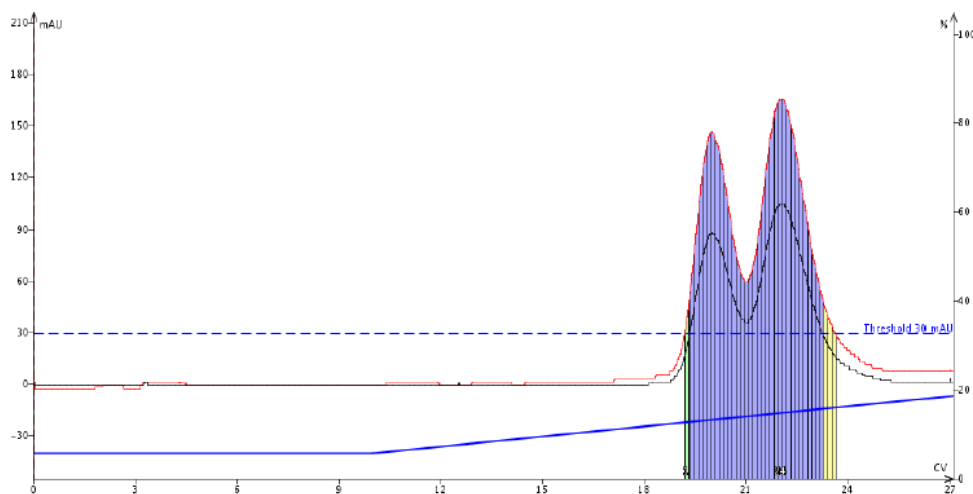
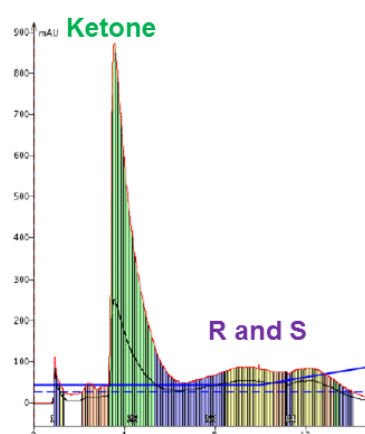
2.4.1.2 5,3-Dibenzoylriboacetal



To an ice cold solution of 1,3,5-tribenzoylribose (3.00 g, 6.49 mmol) in DCM (30 mL) under Ar, hydrogen bromide in acetic acid (33 %, 2.5 mL, ~9.74 mmol) (it is essential that this is anhydrous, ACROS HBr/AcOH in a sealed bottle under Ar), the reaction mixture was stirred on ice for 30 minutes at which point the starting material had disappeared by TLC (20 % EtOAc/Petrol). The reaction mixture was diluted with DCM (120 mL) and washed with ice cold brine (250 mL) followed by ice cold saturated aqueous NaHCO_3 (250 mL). The organic layer was dried on MgSO_4 , filtered, then concentrated to dryness under reduced pressure. The viscous, slightly orange oil was put under high vacuum for 30 minutes (Do not leave longer than this as decomposition sets in).

The resultant oil was dissolved in anhydrous DCM (30 mL) under Ar, treated with 1-iodo-hexan-5-one (3.66 g, 16.2 mmol) and stirred in the dark for 16 h. The reaction

mixture was diluted with DCM (70 mL) and washed sequentially with saturated aqueous NaHCO_3 (100 mL) and brine (100 mL), dried over MgSO_4 , filtered, then concentrated to ~5 mL of brown/black liquid. The crude mixture was purified by two rounds of column chromatography to afford a yellow oil (**S** 1.03 g, 1.81 mmol, 28 %) and a colourless oil which crystallised (**R** 1.14 g, 2.01 mmol, 31 %) (59% total).



2.4.1.2.1 *R* 23

Crystals suitable for X-ray diffraction were prepared by dissolving 5 mg of compound in 1.0 mL of chloroform, filtering through a PTFE syringe filter into a mass spectrometry vial then placing inside a larger vial with a layer of hexane in the bottom before sealing the larger vial.

¹H NMR δ H (400 MHz, CDCl₃) 8.07-8.00 (4H, m, *o*-BzH), 7.59 (1H, app. tt, $J = 7.0$, 1.3 Hz, *p*-BzH), 7.54 (1H, app. tt, $J = 7.0$, 1.3 Hz, *p*-BzH), 7.44 (2H, app. tt, $J = 8.1$, 1.5 Hz, *m*-BzH), 7.39 (2H, app. tt, $J = 8.1$, 1.6 Hz, *m*-BzH), 5.96 (1H, d, $J = 3.4$ Hz, H1), 5.02-4.98 (2H, m, H3 and H2), 4.69 (1H, dd, $J = 11.9$, 3.6 Hz, H5a), 4.63 (1H, ddt, $J = 8.3$, 4.8, 3.0 Hz, H4), 4.48 (1H, dd, $J = 11.9$, 4.8 Hz, H5b), 3.15 (2H, t, $J = 7.0$ Hz, H ζ), 1.83-1.76 (2H, m, H ϵ), 1.62-1.58 (2H, m, H γ), 1.55 (3H, s, H α) and 1.48-1.39 (2H, m, H δ)

¹³C NMR δ C (100.6 MHz, CDCl₃) 166.3 (C=O), 165.8 (C=O), 133.6 (*p*-BzC), 133.3 (*p*-BzC), 130.0 (*o*-BzC), 129.9 (*o*-BzC), 129.7 (*i*-BzC), 129.2 (*i*-BzC), 128.6 (*m*-BzC), 128.5 (*o*-BzC), 115.3 (C β), 105.1 (C1), 78.3 (C2), 76.1 (C4), 73.6 (C3), 63.4 (C5), 39.4 (C γ), 33.4 (C ϵ), 25.3 (C α), 25.2 (C δ) and 6.6 (C ζ)

MS⁺ 589 (M + Na⁺) and 590 (¹³C M + Na⁺)

[α]²⁰ +135.2

2.4.1.2.2 *S* 22

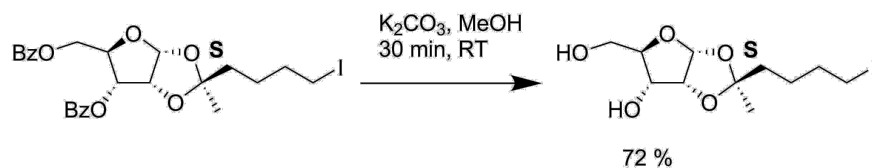
¹H NMR δ H (400 MHz, CDCl₃) 8.08-8.01 (4H, m, *o*-BzH), 7.59 (1H, app. tt, $J = 7.3$, 1.3 Hz, *p*-BzH), 7.54 (1H, app. tt, $J = 7.4$, 1.2 Hz, *p*-BzH), 7.45 (2H, app. tt, $J = 8.0$, 1.5 Hz, *m*-BzH), 7.39 (2H, app. tt, $J = 8.0$, 1.6 Hz, *m*-BzH), 5.96 (1H, d, $J = 3.8$ Hz, H1), 5.05-5.00 (2H, m, H3 and H2), 4.71-4.64 (2H, m, H5a and H4), 4.51-4.49 (1H, m, H5b), 3.82 (2H, t, $J = 7.0$ Hz, H ζ), 1.83-1.76 (4H, m, H ϵ and H γ), 1.61 (2H, m, H δ) and 1.30 (3H, s, H α)

¹³C NMR δ C (100.6 MHz, CDCl₃) 166.3 (C=O), 165.8 (C=O), 133.7 (*p*-BzC), 133.3 (*p*-BzC), 130.1 (*o*-BzC), 129.9 (*o*-BzC), 129.7 (*i*-BzC), 129.1 (*i*-BzC), 128.6 (*m*-BzC), 128.5 (*o*-BzC), 114.2 (C β), 104.5 (C1), 72.2 (C2), 76.2 (C4), 73.6 (C3), 63.3 (C5), 38.5 (C γ), 33.3 (C ϵ), 25.0 (C δ), 24.9 (C α) and 7.0 (C ζ)

MS⁺ 589 (M + Na⁺) and 590 (¹³C M + Na⁺)

[α]²⁰ +83.8

2.4.1.3 S-Riboacetal 24



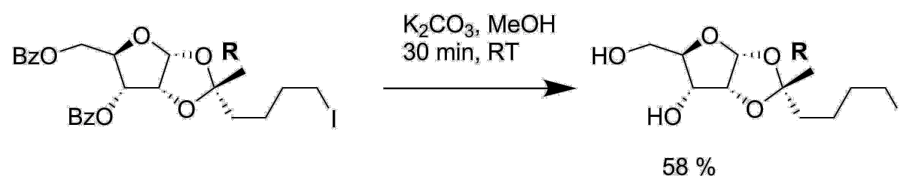
To a solution of S-5,3-Dibenzoylriboacetal (300 mg, 529 μ mol) in DCM/MeOH (1:5, 10 mL) anhydrous K_2CO_3 (292 mg, 2.12 mmol) was added. The reaction mixture was stirred at room temperature open to air for 30 minutes at which point analysis by TLC (10 % DCM/MeOH) showed the reaction complete. The reaction mixture was partitioned between EtOAc and saturated aqueous ammonium chloride (50 mL of each). The organic fraction was dried over $MgSO_4$, concentrated to dryness, then purified by column chromatography (0 $\rightarrow$ 10 % DCM/MeOH) to afford the title compound as a yellow oil (136 mg, 380 μ mol, 72 %).

^1H NMR δH (400 MHz, CDCl_3) 5.84 (1H, d, $J = 3.9$ Hz, **H1**), 4.58 (1H, dd, $J = 5.1$, 3.8 Hz, **H2**), 4.02-3.99 (1H, m, **H3**), 3.96 (1H, dd, $J = 12.3$, 2.8 Hz, **H5a**), 3.84 ($J = 8.8$, 3.7, 2.7 Hz, **H4**), 3.75 (1J, dd, $J = 12.2$, 3.7 Hz, **H5b**), 3.18 (2H, t, $J = 6.9$ Hz, **H ζ**), 1.87-1.76 (2H, m, **H ϵ**), 1.68-1.61 (2H, m, **H γ**), 1.54 (3H, s, **H α**) and 1.52-1.43 (2H, m, **H δ**)

^{13}C NMR δC (100.6 MHz, CDCl_3) 114.9 (**C β**), 104.5 (**C1**), 81.16 (**C4**), 79.7 (**C2**), 71.1 (**C3**), 61.1 (**C5**), 39.3 (**C γ**), 33.2 (**C ϵ**), 25.4 (**C δ**), 25.0 (**C α**) and 6.7 (**C ζ**)

MS $^+$ 381 ($\text{M} + \text{Na}^+$) and 382 ($^{13}\text{C} \text{ M} + \text{Na}^+$)

MS $^-$ 403 ($\text{M} + \text{FA} - \text{H}^+$), 404 ($^{13}\text{C} \text{ M} + \text{FA} - \text{H}^+$), 761 ($2\text{M} + \text{FA} - \text{H}^+$)

2.4.1.4 **R**-Riboacetal 25

To a solution of S-5,3-Dibenzoylriboacetal (300 mg, 529 μ mol) in DCM/MeOH (1:5, 10 mL) anhydrous K_2CO_3 (292 mg, 2.12 mmol) was added. The reaction mixture was stirred at room temperature open to air for 30 minutes at which point analysis by TLC (10 % DCM/MeOH) showed the reaction complete. The reaction mixture was partitioned between EtOAc and saturated aqueous ammonium chloride (50 mL of each). The organic fraction was dried over $MgSO_4$, concentrated to dryness then purified by column chromatography (0 $\rightarrow$ 10 % DCM/MeOH) to afford the title compound as a colourless oil which crystallised (110 mg, 306 μ mol, 72 %). Crystals suitable for X-ray diffraction were prepared by dissolving 5 mg of compound in 1.0 mL of chloroform, filtering through a PTFE syringe filter into a mass spectrometry vial then placing inside a larger vial with a layer of hexane in the bottom before sealing the larger vial.

^1H NMR δH (400 MHz, CDCl_3) 5.83 (1H, d, $J = 3.9$ Hz, **H1**), 4.59 (1H, dd, $J = 5.2$, 3.8 Hz, **H2**), 4.04 (1H, ddd, $J = 10.7$, 8.8, 5.2 Hz, **H3**), 3.96 (1H, ddd, $J = 12.9$, 4.6, 2.7 Hz, **H5a**), 3.86 ($J = 8.9$, 3.6, 2.7 Hz, **H4**), 3.74 (1H, ddd, $J = 12.0$, 8.1, 3.6 Hz, **H5b**), 3.21 (2H, t, $J = 6.9$ Hz, **H ζ**), 2.36 (1H, d, $J = 10.7$ Hz, 3OH) 1.91-1.74 (5H, m, **H ϵ** , γ , 5OH), 1.68-1.55 (2H, m, **H δ**) and 1.52-1.43 (3H, s, **H α**)

^{13}C NMR δC (100.6 MHz, CDCl_3) 113.8 (**C β**), 104.0 (**C1**), 81.2 (**C4**), 78.5 (**C2**), 71.1 (**C3**), 61.0 (**C5**), 38.5 (**C γ**), 33.2 (**C ϵ**), 25.1 (**C δ**), 24.6 (**C α**) and 6.8 (**C ζ**)

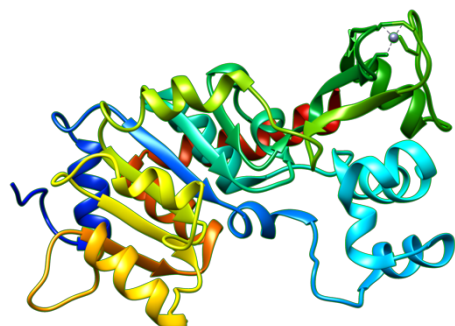
MS $^+$ 381 (M + Na $^+$) 382 (^{13}C M + Na $^+$), 437 (M + DMSO + H $^+$)

2.4.2.0 Images

Individual references for structures are: Sirt2 binding NPD11033,²⁴¹ Sirt2,²⁴² H4

2.4.3.0 Protein Expression

2.4.3.1 *Human* Sirtuin 2 (56-365 and 50-365)



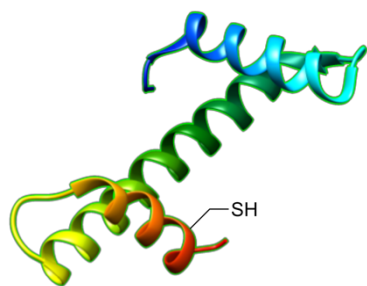
Adapted from Jung²⁴²

The gene for *Human* Sirt2 in pET6 plasmid was transformed in BL21-(DE3) cells and plated on LB/agar/carbenicillin plates. Single colonies were picked into 5 mL of LB/carbenicillin and grown at 37 °C 250 RPM overnight. The starter culture was poured into 500 mL of super broth media and grown to an OD of 0.6 at 37 °C 250 RPM (2 hours). Expression was induced by the addition of IPTG stock to give the final concentration of 0.3 mM then incubated at 37 °C 250 RPM for 4 h. Cells were pelleted at 8 kRPM 9.6 kG at $r_{average}$ for 15 minutes then the pellets frozen at -80 °C.

Pellets were thawed on ice then resuspended in buffer (NaCl 500 mM, Tris 50 mM, Glycerol 5%, BME 5 mM, Imidazole 25 mM and one Roche cOmplete EDTA-free protease inhibitor cocktail tablet at a pH of 7.5, 10 mL). Cells were lysed by sonication on ice (30 %, for 5 minutes on 2 s on 2 s off). The insoluble fraction was removed by centrifugation (25 kRPM 52 kG at $r_{average}$ for 1 h) and the lysate filtered through a 0.2 μ m syringe filter before being applied to a FPLC column.

The protein was purified by 2D-FPLC firstly through a 1 mL ff-Histrap (A: NaCl 500 mM, Tris 50 mM, Glycerol 5%, BME 5 mM, Imidazole 25 mM pH of 7.5, B: A + 225 mM Imidazole pH 7.5, 5 CV A 10 CV B step gradient). Fractions containing the desired protein (analysed by SDS-PAGE) were concentrated to ~5 mL using a 10 kDa GE vivaspin then passed through a s200 36/60 sec column in 150 mM NaCl, 25 mM Tris pH 8.0 buffer to give 3 mL of 0.15 mg mL⁻¹ protein. A 500 µL sample of this was desalted into fPBS using a G-25 minitrapp then concentrated to 500 µL using a 10 kDa vivaspin in order to record higher quality CD spectra. Expression trials in different media were conducted by replacing the broth in which the expression was conducted, but otherwise kept the same. The other broths tested were LB, Muller-Hinton and Terrific Broth.

2.4.3.2 Histone H4 K16C



Adapted from Chromatin Protocols (Methods in Molecular Biology)²¹⁷

Plasmids (pET3d) containing the histone H4 K16C gene were transformed into BL-21(DE3) pLys *E. coli* cells and plated on ampicillin/chloramphenicol/LB/agar plates. Single colonies were picked and two 5 mL starter cultures (LB/ampicillin/chloramphenicol) were grown over night at 37 °C 200 RPM then transferred into 2 x 500 mL of media (2xYT/ ampicillin) in a 2 L baffled conical flask. The cultures were grown at 37 °C until an $OD_{600} = 0.6 - 0.8$ was achieved (typically between 100 – 140 min) at which point the cells were induced with the addition of IPTG to reach a final concentration of 0.3 mM and incubated for a further 3 h at 37 °C. The cells were then pelleted for 10 minutes at 8 kRPM 9.6 kG at $r_{average}$ then stored at -80 °C overnight.

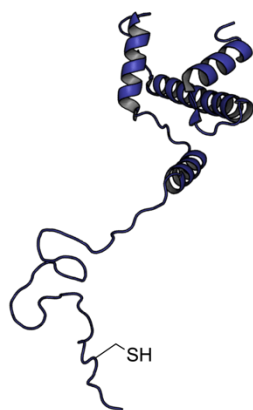
Cell pellets were thawed then resuspended in buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, 1 Roche cOmplete EDTA-free protease inhibitor tablet, 1 mg mL⁻¹ lysozyme, 0.1 mg mL⁻¹ DNase, pH 7.5) and stirred on ice for 30 min. The suspension was subjected to sonication (50 % Amplitude, 30 s sonication 1 min

rest, five times) followed by centrifugation (20 kRPM, 33.2 kG at r_{average} 30 mins). The supernatant was discarded, and the pellet suspended in TW buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, 1 % Triton X, pH 7.5) using an IKA Ultra-turrax T-8 and the suspension was subjected to sonication (50 % power, 30 s sonication 1 min rest, two times) followed by centrifugation (20 krpm, 33.2 kG at r_{average} 20 mins). The pellet was washed twice more using TW buffer and once with wash buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, pH 7.5). The pellet was treated with DMSO (500 μ L) and allowed to stand for 10 mins. Unfolding buffer (Gdm.HCl 7M, Tris 10 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM, pH 7.5, 4 mL) was applied to the pellet and the suspension vigorously shaken (400 RPM) for 1 h. The suspension was subject to centrifugation (20 kRPM, 33.2 kG at r_{average} 20 mins) and the supernatant filtered sequentially through a 1.2 μm^2 syringe filter and 0.5 μm^2 syringe filter.

The filtered supernatant was purified by gel filtration (s200 26/60) using SAU-100 (Urea 7 M, NaCl 100 mM, NaOAc 20 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM pH 5.2) the fractions containing the desired protein (visualised by SDS-PAGE analysis) were pooled and subject to a 2nd purification by strong cation exchange chromatography (using a GE Healthcare HiTrap SP HP 5 mL column). The sample was loaded in SAU-100 buffer then a linear gradient 70 % SAU-1000 buffer (Urea 7 M, NaCl 1 M, NaOAc 20 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM pH 5.2) was ran over 20 CV. The fractions containing the desired protein (visualised by

SDS-PAGE analysis) were pooled, dialysed against aqueous BME (2 mM, 3 x 4 L, at least 4 hours for each step), then lyophilised and stored at -20 °C.

2.4.3.3 Histone H3 K18C



Adapted from Chromatin Protocols (Methods in Molecular Biology)²¹⁷

Plasmids (pET3d) containing the histone H3 K18C gene were transformed into BL-21(DE3) pLys *E. coli* cells and plated on ampicillin/chloramphenicol/LB/agar plates. Single colonies were picked and two 5 mL starter cultures (LB/ampicillin/chloramphenicol) were grown over night at 37 °C 200 RPM then transferred into 2 x 500 mL of media (2xYT/ ampicillin) in a 2 L baffled conical flask. The cultures were grown at 37 °C until an $OD_{600} = 0.6 - 0.8$ was achieved (typically between 100 – 140 min) at which point the cells were induced with the addition of IPTG to reach a final concentration of 0.3 mM and incubated for a further 3 h at 37 °C. The cells were then pelleted for 10 minutes at 8 kRPM 9.6 kG at $r_{average}$ then stored at -80 °C overnight.

Cell pellets were thawed then resuspended in buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, 1 Roche cOmplete EDTA-free protease inhibitor tablet, 1 mg mL⁻¹ lysozyme, 0.1 mg mL⁻¹ DNase, pH 7.5) and stirred on ice for 30 min.

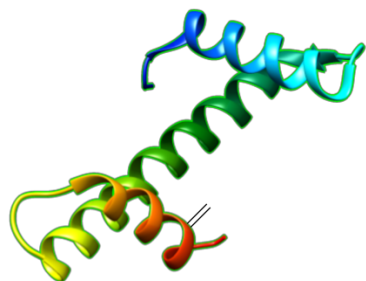
The suspension was subjected to sonication (50 % Amplitude, 30 s sonication 1 min rest, five times) followed by centrifugation (20 kRPM, 33.2 kG at r_{average} , 30 mins). The supernatant was discarded, and the pellet suspended in TW buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, 1 % Triton X, pH 7.5) using an IKA Ultra-turrax T-8 and the suspension was subjected to sonication (50 % power, 30 s sonication 1 min rest, two times) followed by centrifugation (20 krpm, 33.2 kG at r_{average} 20 mins). The pellet was washed twice more using TW buffer and once with wash buffer (40 mL, Tris 50 mM, NaCl 100 mM, EDTA 1mM, BME 5 mM, pH 7.5). The pellet was treated with DMSO (500 μ L) and allowed to stand for 10 mins. Unfolding buffer (Gdm.HCl 7M, Tris 10 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM, pH 7.5, 4 mL) was applied to the pellet and the suspension vigorously shaken (400 RPM) for 1 h. The suspension was subject to centrifugation (20 kRPM, 33.2 kG at r_{average} 20 mins) and the supernatant filtered sequentially through a 1.2 μm^2 syringe filter and 0.5 μm^2 syringe filter.

The filtered supernatant was purified by gel filtration (s200 26/60) using SAU-100 (Urea 7 M, NaCl 100 mM, NaOAc 20 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM pH 5.2) the fractions containing the desired protein (visualised by SDS-PAGE analysis) were pooled and subject to a 2nd purification by strong cation exchange chromatography (using a GE Healthcare HiTrap SP HP 5 mL column). The sample was loaded in SAU-100 buffer then a linear gradient 70 % SAU-1000 buffer (Urea 7 M, NaCl 1 M, NaOAc 20 mM, DTT 10 mM, EDTA 1 mM, benzamidine 1 mM pH

5.2) was ran over 20 CV. The fractions containing the desired protein (visualised by SDS-PAGE analysis) were pooled, dialysed against aqueous BME (2 mM, 3 x 4 L, at least 4 hours for each step), then lyophilised and stored at -20 °C.

2.4.5.0 Chemical Reactions of Proteins

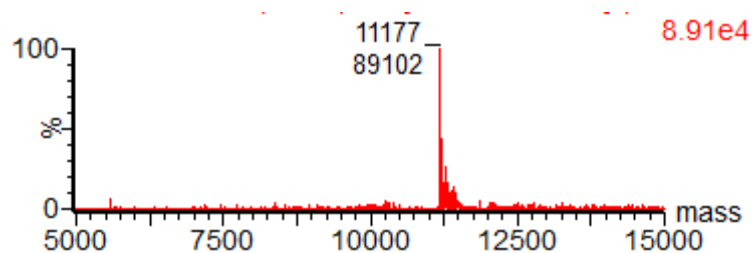
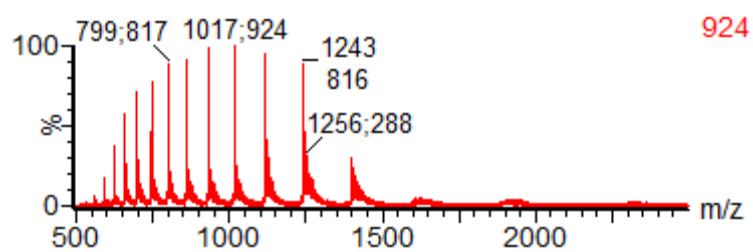
2.4.5.1 Histone H4 K16Dha



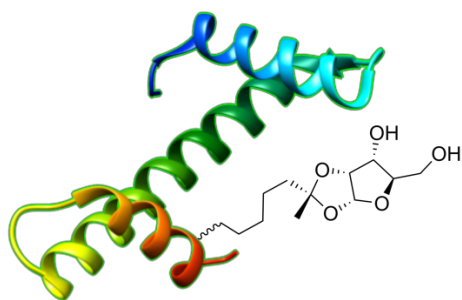
To a solution of Histone H4 K16C (4.3 mg) in guanidinium chloride (5 M) and Tris (50 mM) buffer (pH 8, 500 μ L) DTT (10 mg) was added. The reaction mixture was shaken for 30 mins at room temperature and passed through a G-25 PD-10 minitrap desalting column equilibrated with the above buffer and eluted with 1000 μ L of the above buffer. To a sample of the eluent (900 μ L) DBHDA stock solution (69 μ L, 164 mg/mL in DMF) was added, the reaction mixture was shaken for 45 mins at room temperature then 90 mins at 37 °C. The reaction mixture was concentrated to 500 μ L using a 3 kDa vivaspin and then passed through a G-25 PD10 minitrap desalting column pre-equilibrated with NH_4OAc (500 mM) and guanidinium chloride (5 M) buffer (pH 6) and eluted with 1000 μ L of the above buffer.

Calculated mass: 11177

Observed mass: 11177



2.4.5.2 Histone H4 K16R-Riboacetal

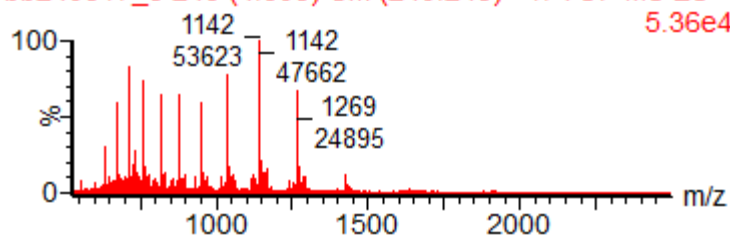


A solution of H4 K16Dha (1 mg mL^{-1} 200 μL) and **R**-Riboacetal (3 mg) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 °C and pipetted directly over NaBH_4 (1 mg). The reaction mixture was shaken at 4 °C for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.

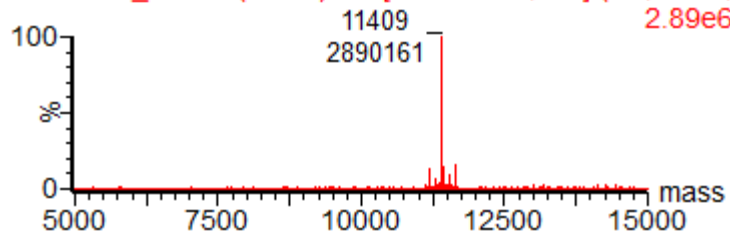
Calculated mass: 11409

Observed mass: 11409

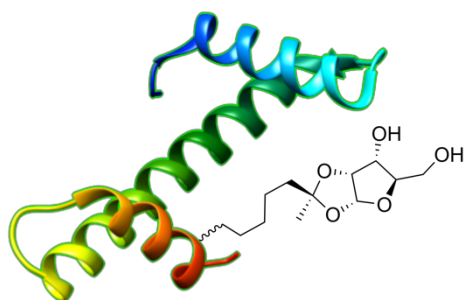
bb240817_8 243 (4.305) Cm (240:248) 1: TOF MS ES+ 5.36e4



bb240817_8 243 (4.305) M1 [Ev-452926,lt31] (Gs,0.750, 2.89e6



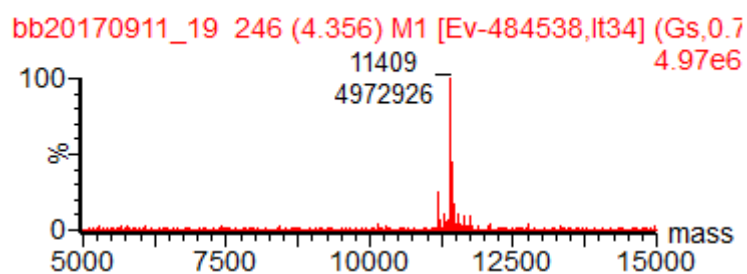
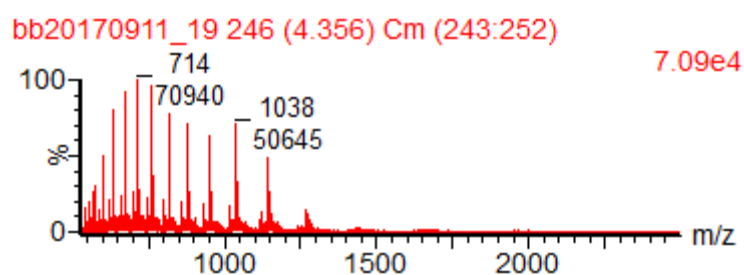
2.4.5.3 Histone H4 K16S-Riboacetal



A solution of H4 K16Dha (1 mg mL^{-1} 200 μL) and S-Riboacetal (3 mg) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 $^{\circ}\text{C}$ and pipetted directly over NaBH_4 (1 mg). The reaction mixture was shaken at 4 $^{\circ}\text{C}$ for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.

Calculated mass: 11409

Observed mass: 11409



2.4.5.4 Histone H3 K18Dha

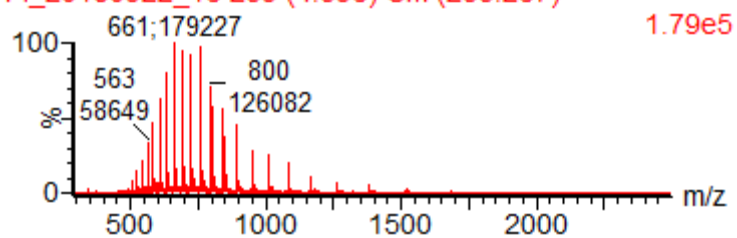


To a solution of Histone H3 K18C (10 mg) in Gdm.Cl (3 M) and $\text{Na}_2\text{H}_3\text{PO}_4$ (100 mM) buffer (pH 8.0, 500 μL) DTT (30 mg) was added. The reaction mixture was shaken for 30 mins at room temperature and passed through a G-25 PD10 minitrapp desalting column equilibrated with the above buffer and eluted with 1000 μL of the above buffer. To a sample of the eluent (900 μL) DBHDA stock solution (100 μL , 164 mg mL^{-1}) was added, the reaction mixture as shaken, for 45 mins at room temperature then 90 mins at 37 °C. The reaction mixture was concentrated to 500 μL using a 3 kDa vivaspin and then passed through a G-25 PD10 minitrapp desalting column pre-equilibrated with NH_4OAc (500 mM) and Gdm.Cl (5 M) buffer (pH 8.0) and eluted with 1000 μL of the above buffer.

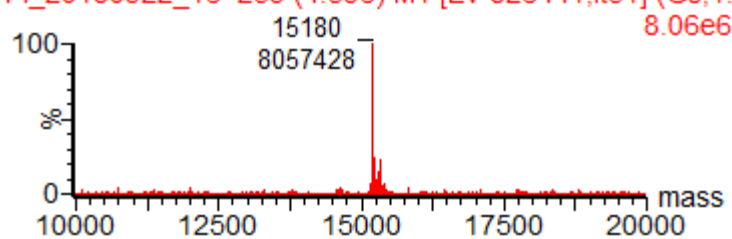
Calculated mass: 15181

Observed mass: 15180

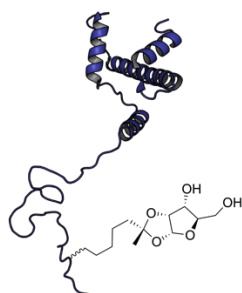
PI_20180922_16 259 (4.596) Cm (255:267)



PI_20180922_16 259 (4.596) M1 [Ev-826441,lt51] (Gs,1.



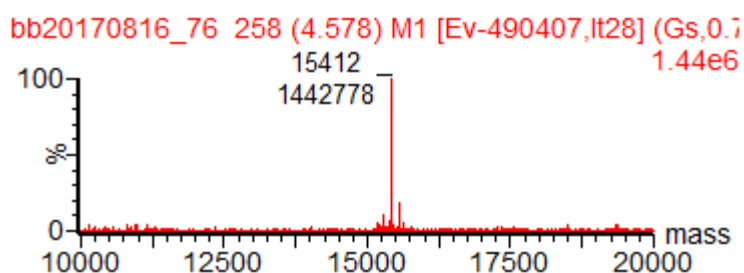
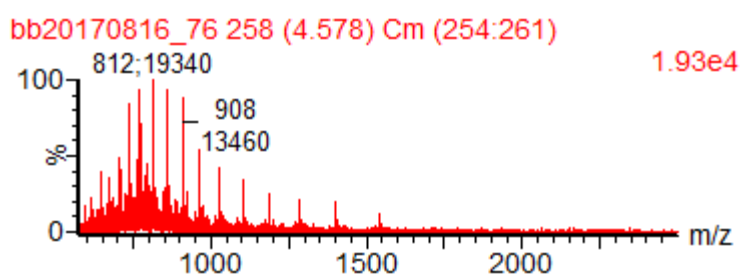
2.4.5.5 Histone H3 K18R-Riboacetal



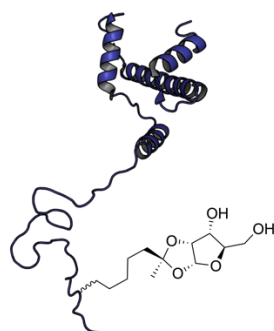
A solution of H3 K18Dha (1 mg mL^{-1} 200 μL) and **R**-Riboacetal (3 mg) in buffer (5 M Gdm.Cl, 500 mM NH_4OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 °C and pipetted directly over NaBH_4 (1 mg). The reaction mixture was shaken at 4 °C for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.

Calculated mass: 15413

Observed mass: 15412



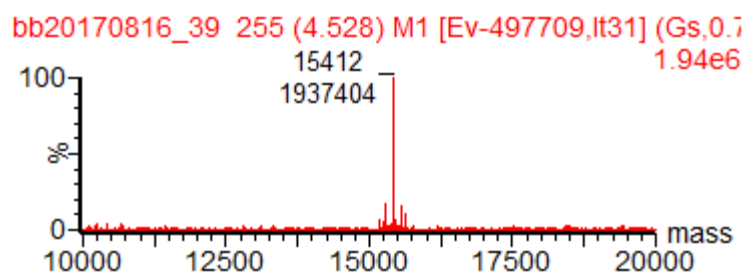
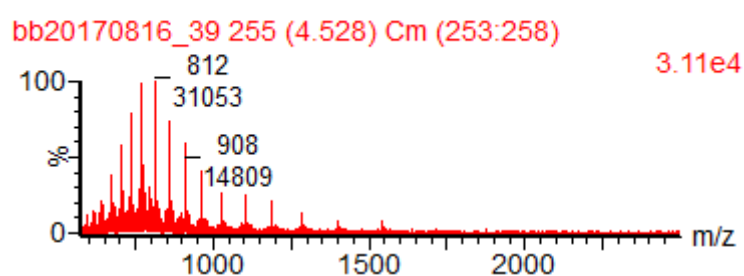
2.4.5.6 Histone H3 K18S-Riboacetal



A solution of H3 K18Dha (1 mg mL⁻¹ 200 μ L) and S-Riboacetal (3 mg) in buffer (5 M Gdm.Cl, 500 mM NH₄OAc pH 6.0) which had been pre-equilibrated in a glovebox overnight was cooled to 4 °C and pipetted directly over NaBH₄ (1 mg). The reaction mixture was shaken at 4 °C for 30 mins then desalted using a G25 Spintrap. The resultant mixture was analysed by LCMS.

Calculated mass: 15413

Observed mass: 15412



2.4.6.0 Biacore Experiments

2.4.6.1 CM5 Coupling

Murine α -Poly(His) was diluted in sodium acetate buffer (1 μ L added to 99 μ L), Stocks of EDC (1 M), NHS (1 M) and ethanolamine (1 M) in HBS-P were prepared.

Flowcell 1: The program takes EDC (50 μ L) and NHS (50 μ L) then mixes in an empty vial before injecting over the GE Healthcare CM5 chip 50 μ L at 10 μ L min⁻¹. Ethanolamine (1 M) was flowed over the CM5 chip (90 μ L at 10 μ L min⁻¹).

Flowcell 2: The program takes EDC (50 μ L) and NHS (50 μ L) then mixes in an empty vial before injecting over the CM5 chip at 50 μ L at 10 μ L min⁻¹. The injection was monitored and when a response of $\sim$ 12 kRU was observed the injection was stopped, this took 3 s. To cap any remaining activated carboxylic acids, ethanolamine (1 M) was flowed over the CM5 chip (90 μ L at 10 μ L min⁻¹). This resulted in $\sim$ 9 kRU of response in flowcell 2 compared to flowcell 1.

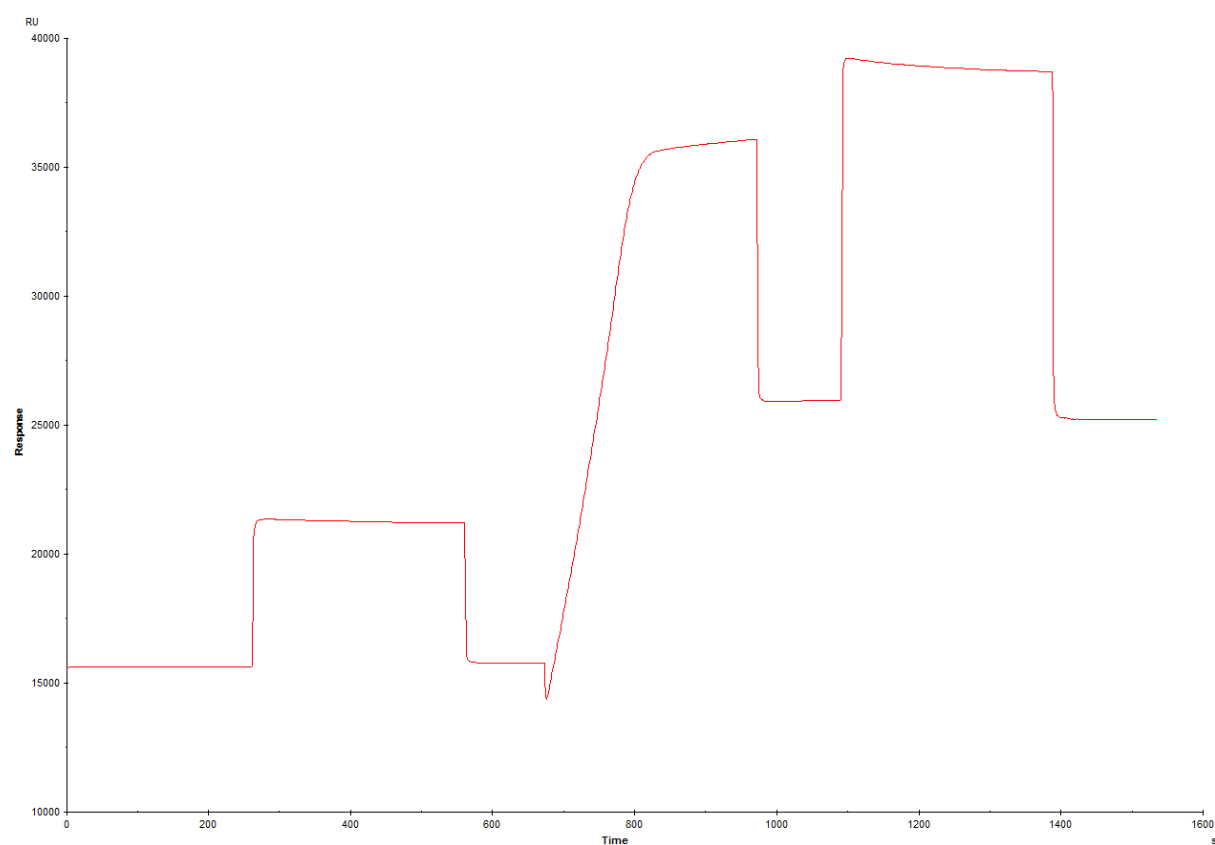


Figure 27 - CM5 Antibody Coupling

2.4.6.2 CM5 Calibration

Serial dilutions of Sirt2 50-356 were carried out in HBS-P (to obtain 30, 10, 3, 1, 0.3 and 0.1 $\mu\text{g mL}^{-1}$). Sirt2 samples were flowed through flowcells 1 and 2 (40 μL at 10 $\mu\text{L min}^{-1}$) followed by HBS-P buffer (10 μL at 10 $\mu\text{L min}^{-1}$) and finally HCl (10 mM in water) 10 μL at 10 $\mu\text{L min}^{-1}$. We found that the binding of Sirt2 to the antibody was somewhat weak causing Sirt2 to be readily washed off the chip.

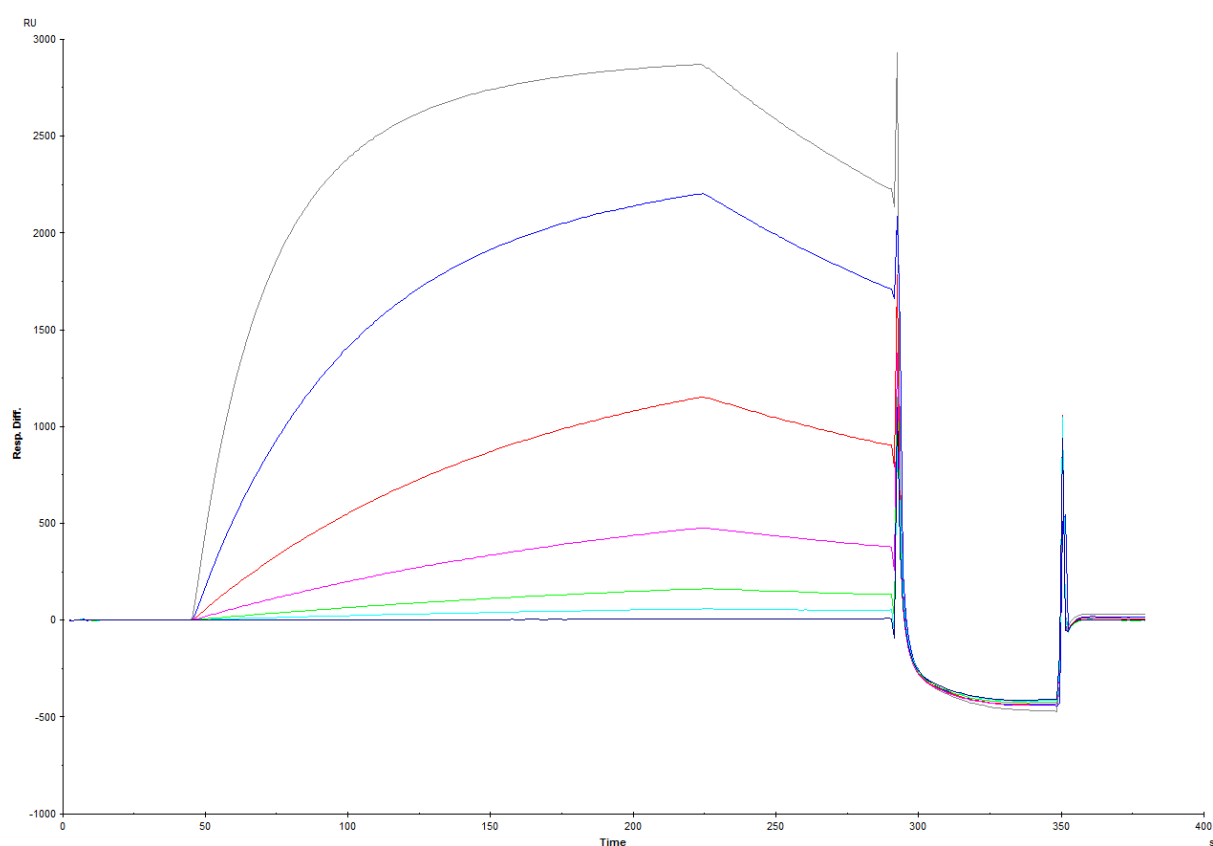


Figure 28 - CM5 Sirt2 Calibration

2.4.6.3 CM5 Histone Test

A 100 nM sample of H3 **R** in HBS-P with the appropriate level of salt added was prepared. Flowcells 3 and 4 (neither have antibody bound). At $30 \mu\text{L min}^{-1}$ (In HBS-P with the indicated level of salt) a sample of H3 **R** typically 90 μL was then flowed over followed by 100 μL of 40 mM HCl in order to remove any bound H3. If significant binding was observed the histone injection was stopped as quickly as possible.

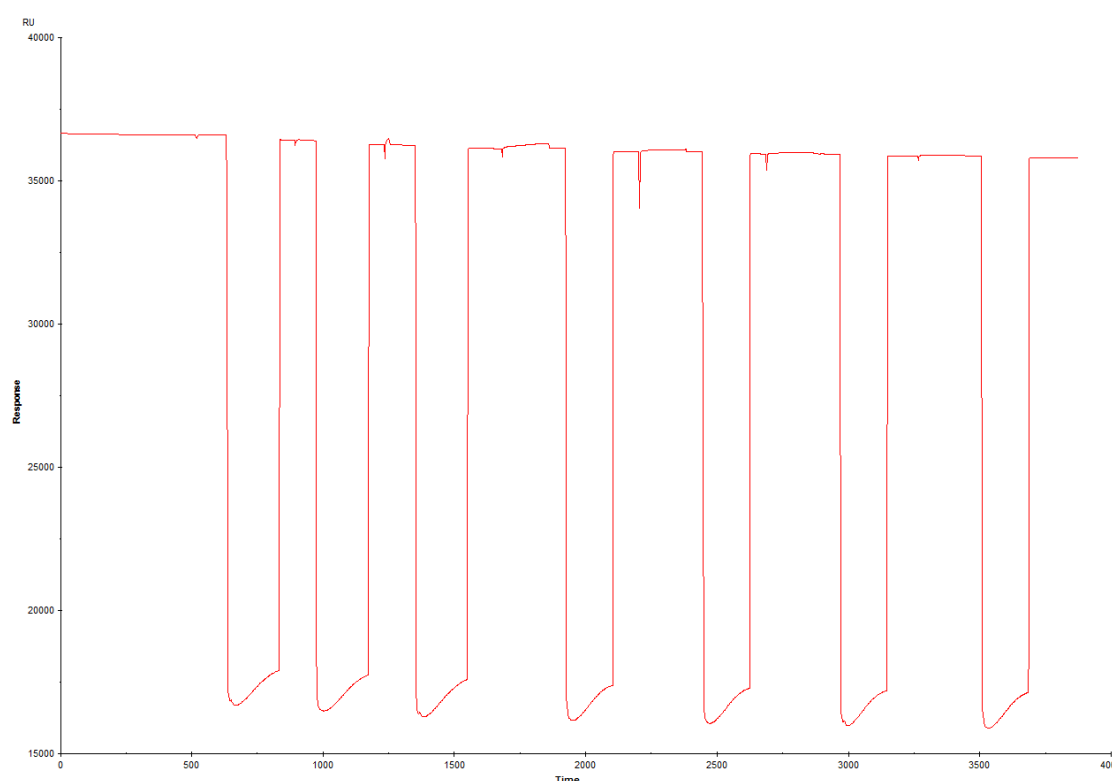


Figure 29 - CM5 Histone test with $[\text{NaCl}] = 2.0 \text{ M}$

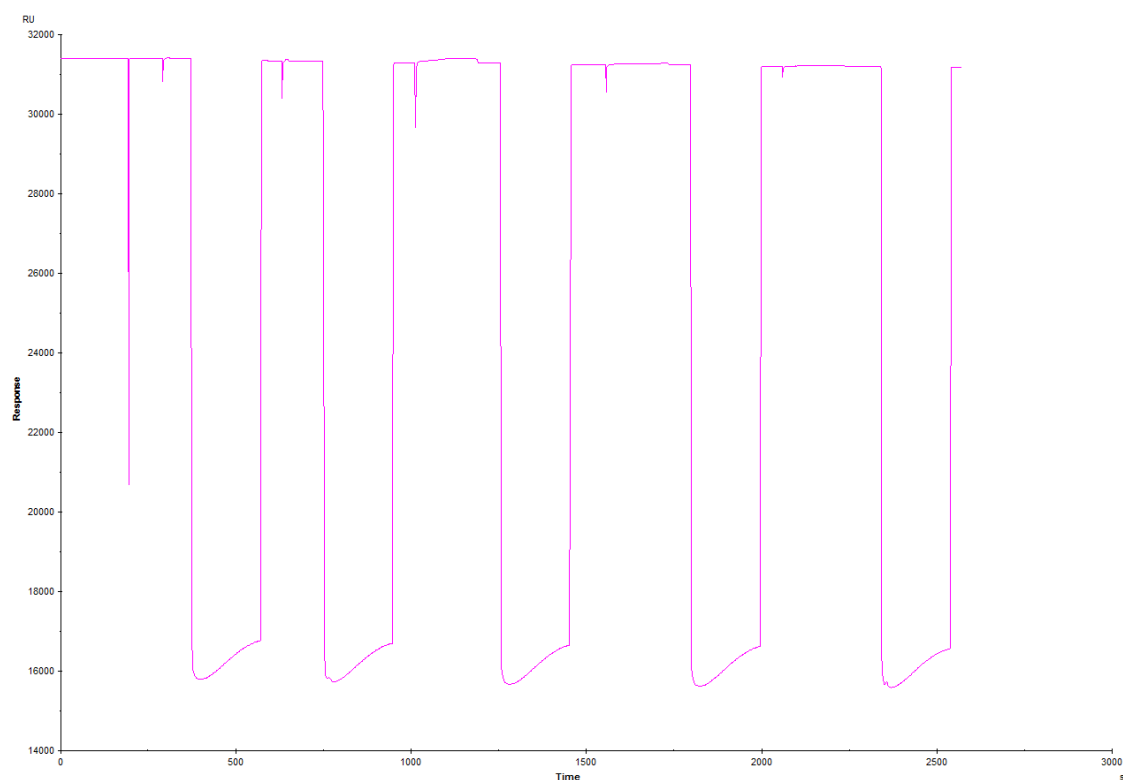


Figure 30 - CM5 Histone test with $[NaCl] = 1.5 M$

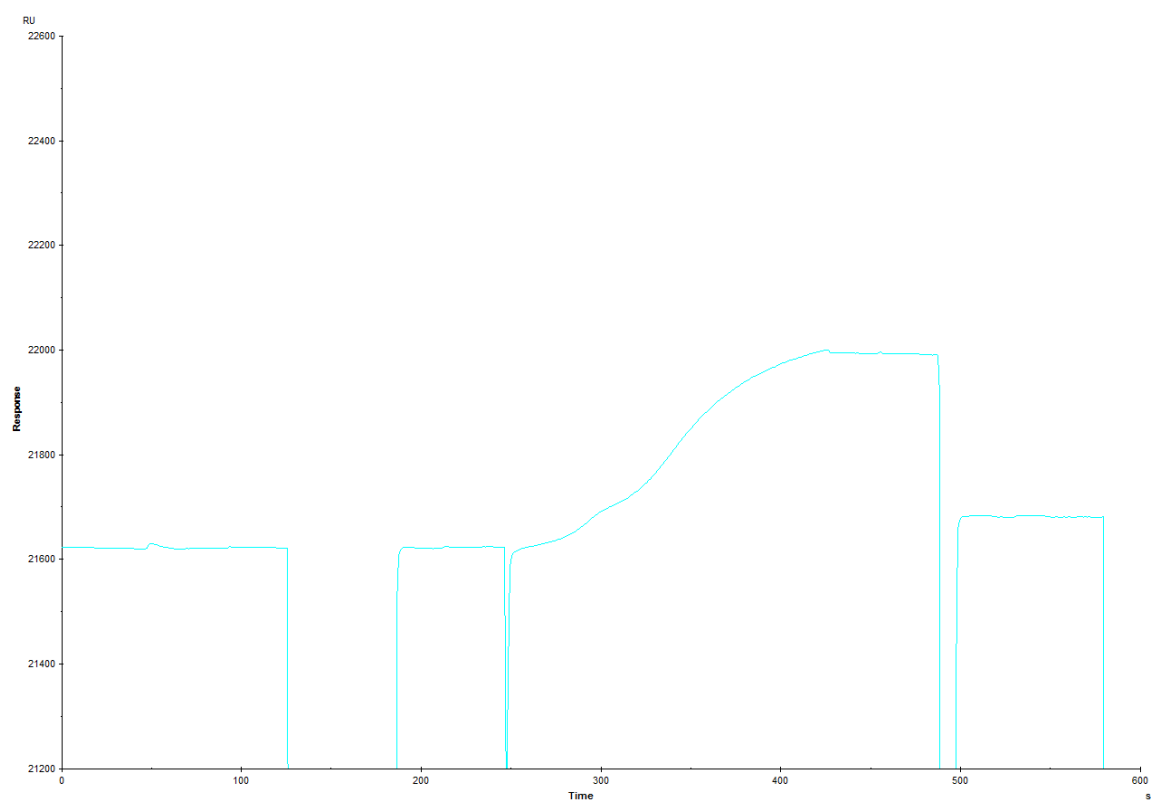


Figure 31 - CM5 Histone test with $[NaCl] = 1.0$

2.4.6.4 CM5 Binding Experiment

All injections were through flowcells 1 and 2 on the previously prepared α -Poly(His) CM 5 chip. A “coinject” was then iteratively carried out using the tabulated sequence.

3 samples of each: H3 **R** at 100 nM, H3 **S** at 100 nM, H3 **R** at 1 μ M and H3 **S** at 1 μ M

Cycle	Injection 1	Injection 2
1	Buffer	Buffer
2	Sirt2	Buffer
3	Sirt2	H3
4	Buffer	H3
5	Buffer	Buffer

The sonograms resulting from 2 and 4 were then subtracted from the sonogram resulting from cycle 3, an unprocessed example can be seen in *figure 21*. The flow rate was set at 30 μ L min⁻¹ and all samples as well as the running buffer were in HBS-P with the [NaCl] altered to 1.5 M

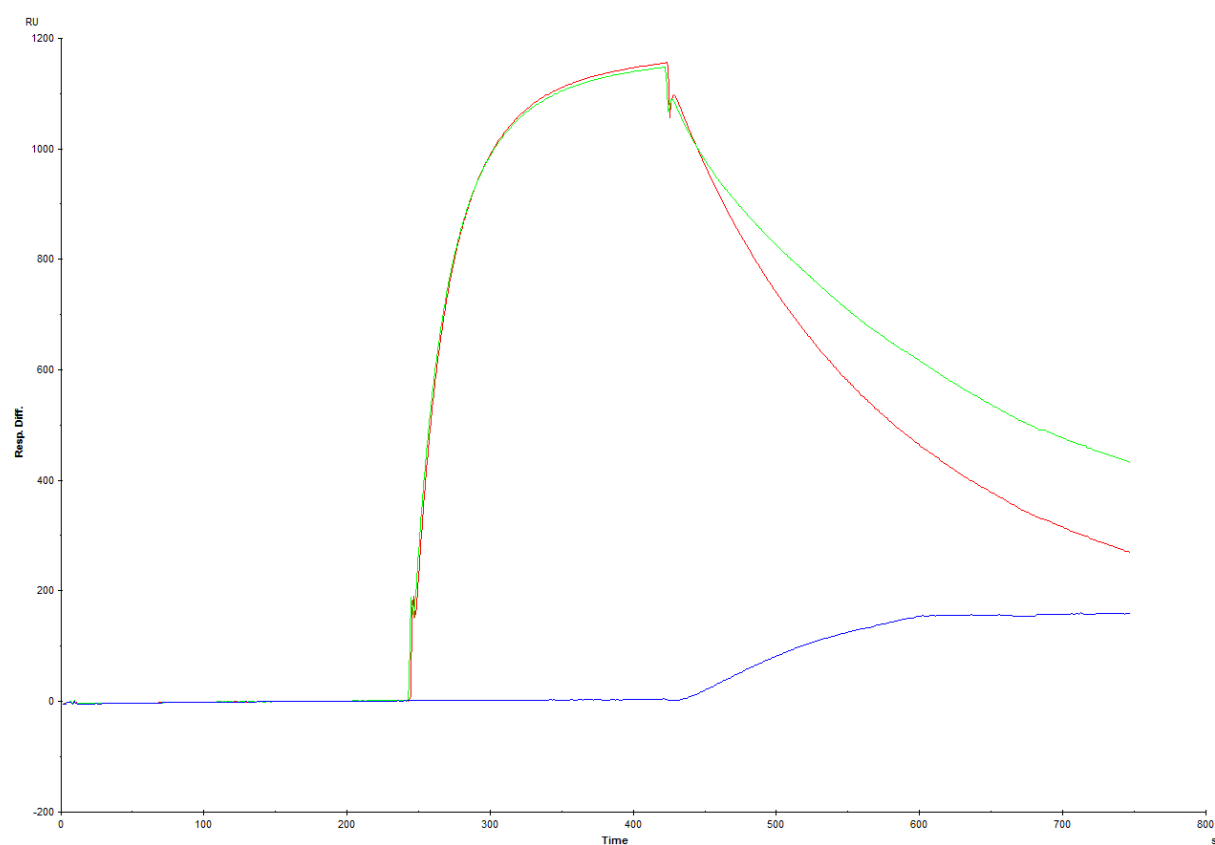


Figure 32 - Example of unprocessed data showing Cycle 2 (Red), Cycle 3 (Green) and Cycle 4 (Blue)

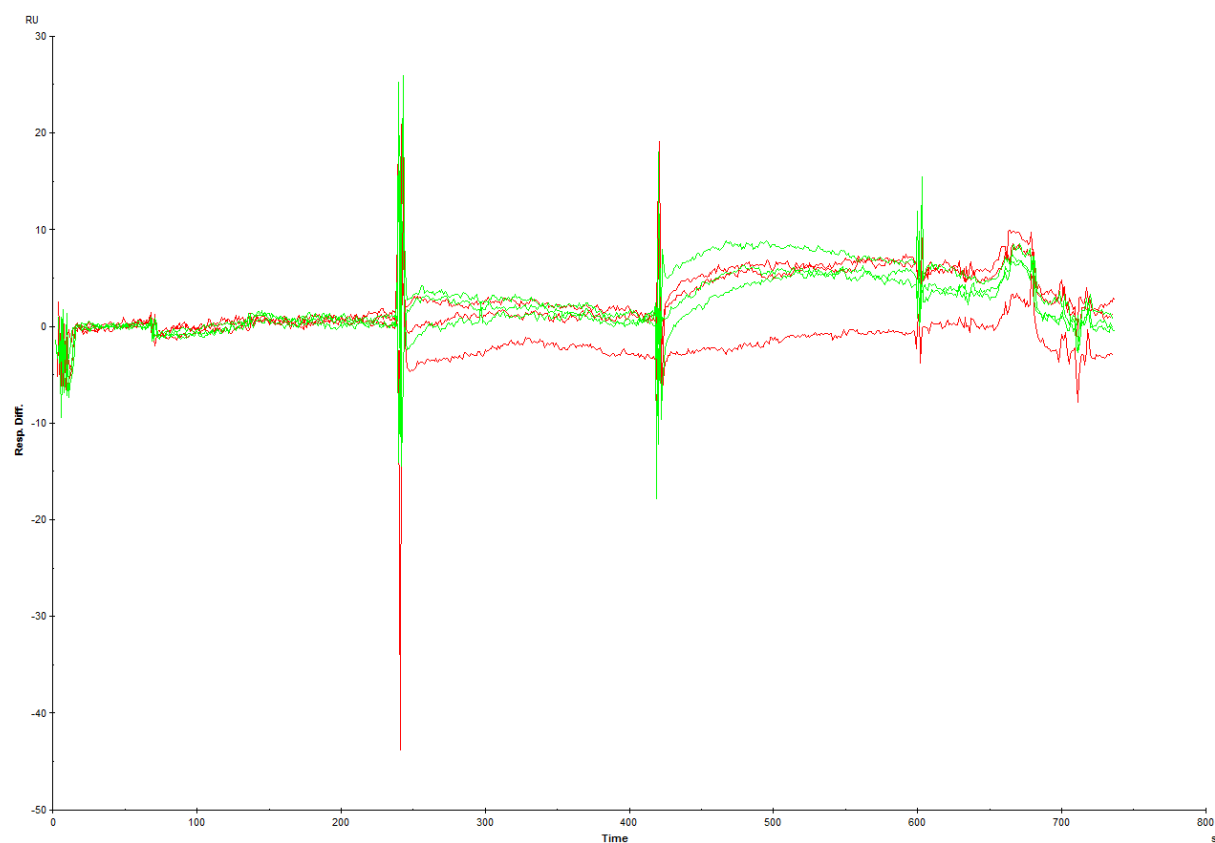


Figure 33 - 100 nM R and S after processing

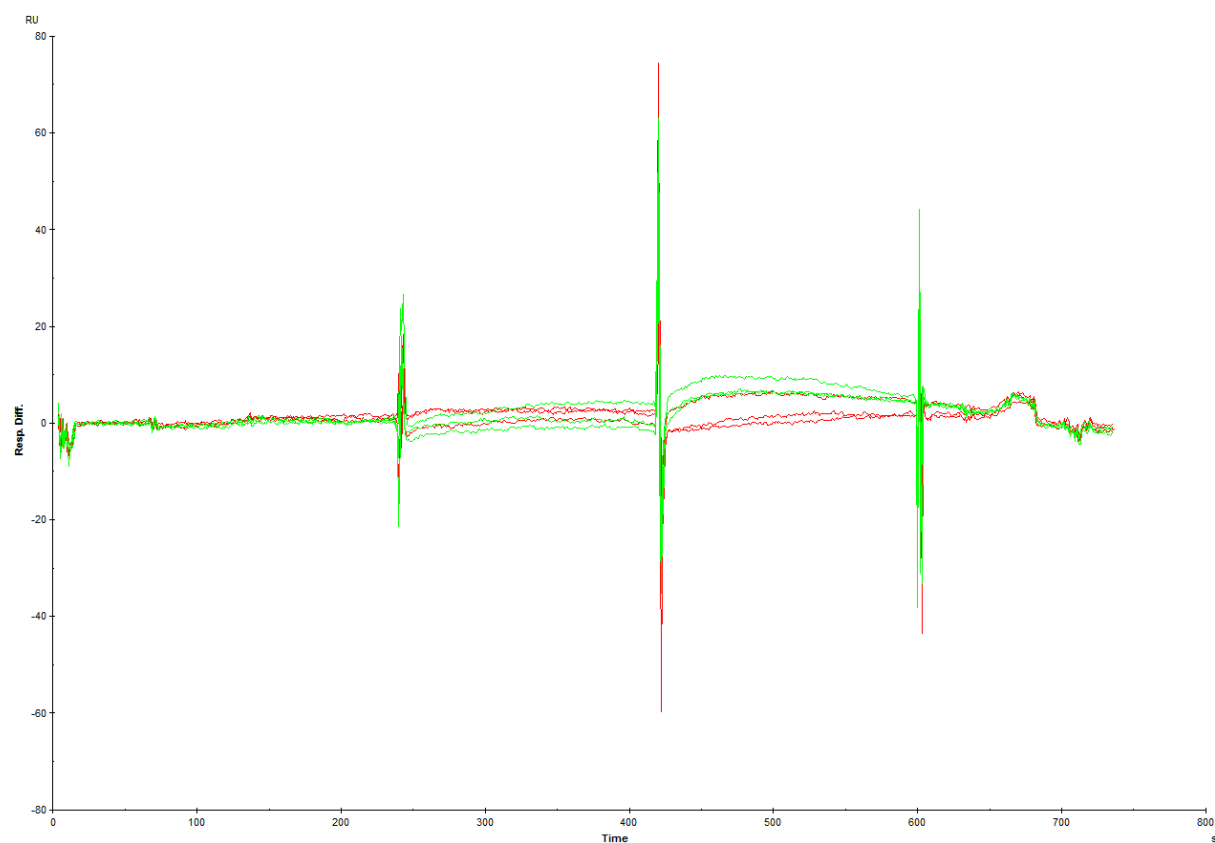


Figure 34 -1 μM **R** and **S** after processing

2.4.6.5 CM5 Histone Test with CM-Dextran

CM Dextran-70 was dissolved in HBS-P to give a concentration of 1 mg mL^{-1} . The biacore was purged with this solvents and a sample of H3 **R** mixed in the dextran buffer to give a concentration of 100 nm . At $30 \text{ } \mu\text{L min}^{-1}$. 2 x $90 \text{ } \mu\text{L}$ injections of the histone was carried out, a large response was observed.

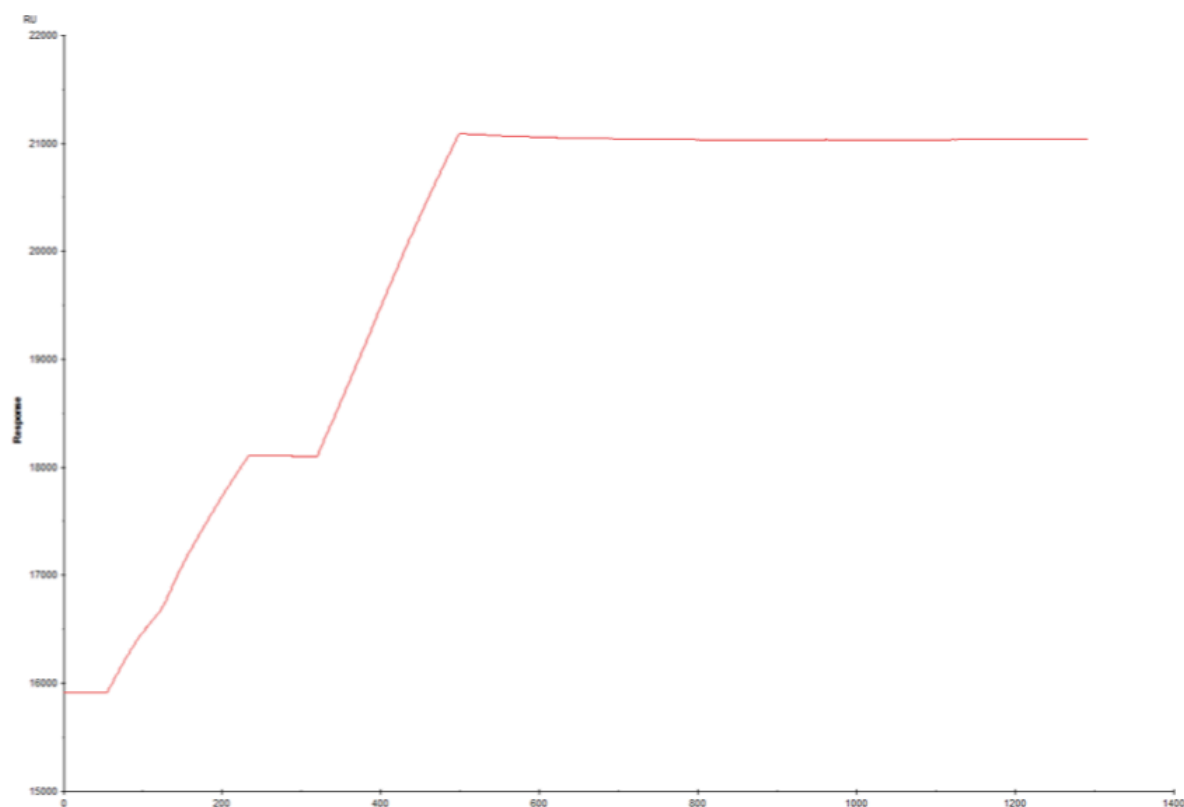


Figure 35 - CM5 Histone test with $[\text{NaCl}] = 150 \text{ mM}$ and 1 mg mL^{-1} CM Dextran 70

2.4.6.6 CM4 Coupling

Murine α -Poly(His) was diluted in sodium acetate buffer (10 μ L added to 90 μ L), stocks of EDC (1 M), NHS (1 M) and ethanolamine (1 M) in HBS-P were prepared.

Flowcell 3: The program takes EDC (50 μ L) and NHS (50 μ L) then mixes in an empty vial before injecting over the GE Healthcare CM4 chip 50 μ L at 10 μ L min⁻¹. Ethanolamine (1 M) was flowed over the CM5 chip (90 μ L at 10 μ L min⁻¹).

Flowcell 4: The program takes EDC (50 μ L) and NHS (50 μ L) then mixes in an empty vial before injecting over the CM5 chip 50 μ L at 10 μ L min⁻¹. The injection was monitored and when a response of $\sim$ 12 kRU was observed the injection was stopped, this took 30 s. To cap any remaining activated carboxylic acids ethanolamine (1 M) was flowed over the CM5 chip (90 μ L at 10 μ L min⁻¹). This resulted in $\sim$ 5 kRU of response in flowcell 2 compared to flowcell 1.

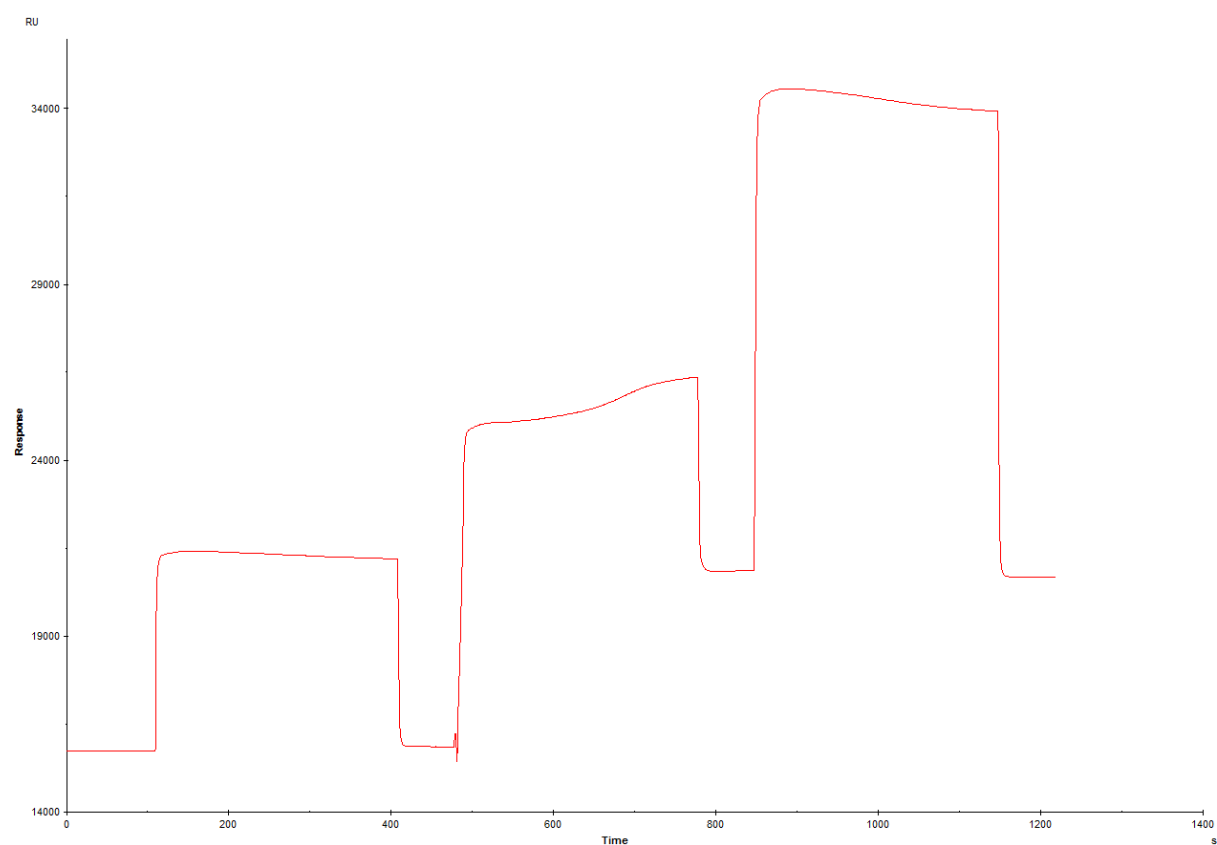


Figure 36 - CM4 Antibody coupling

2.4.6.7 CM4 Calibration

Serial dilutions of Sirt2 50-356 were carried out in HBS-P (to obtain 30, 10, 3, 1, 0.3 and 0.1 $\mu\text{g mL}^{-1}$). Sirt2 samples were flowed through flowcells 1 and 2 (40 μL at 10 $\mu\text{L min}^{-1}$) followed by HBS-P buffer (10 μL at 10 $\mu\text{L min}^{-1}$) and finally HCl (10 mM in water) 10 μL at 10 $\mu\text{L min}^{-1}$. We found that the binding of Sirt2 to the antibody was somewhat weak causing Sirt2 to be readily washed off the chip

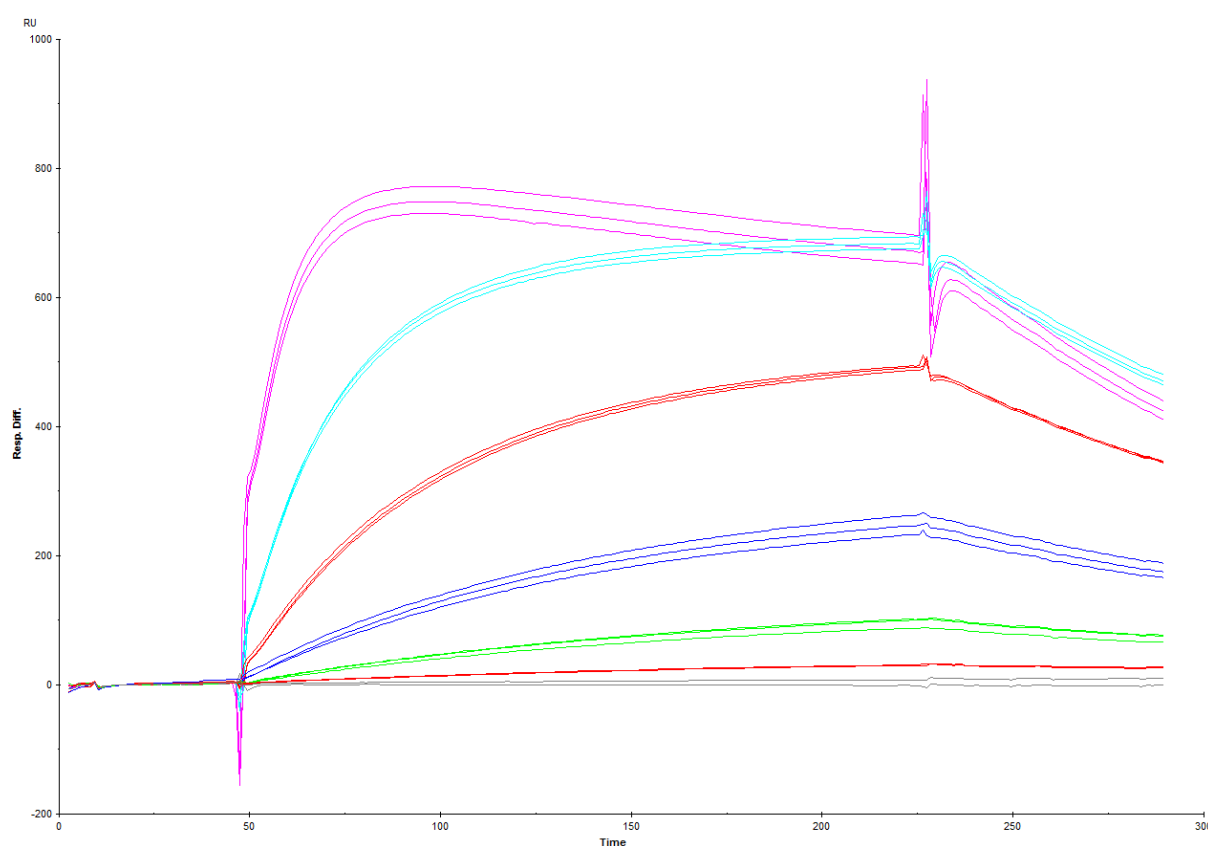


Figure 37 - CM4 Sirt2 Calibration

2.4.6.8 CM4 Histone Test

Samples of H3 **R** in HBS-P (100 nM) with the appropriate level of salt added and indicated pH was prepared. Flowcells 1 and 2 (neither have antibody bound). At 30 $\mu\text{L min}^{-1}$ (In HBS-P with the indicated level of salt) a sample of H3 **R** typically 90 μL was then flowed over followed by 100 μL of 40 mM HCl in order to remove any bound H3. If significant binding was observed the histone injection was stopped as quickly as possible.

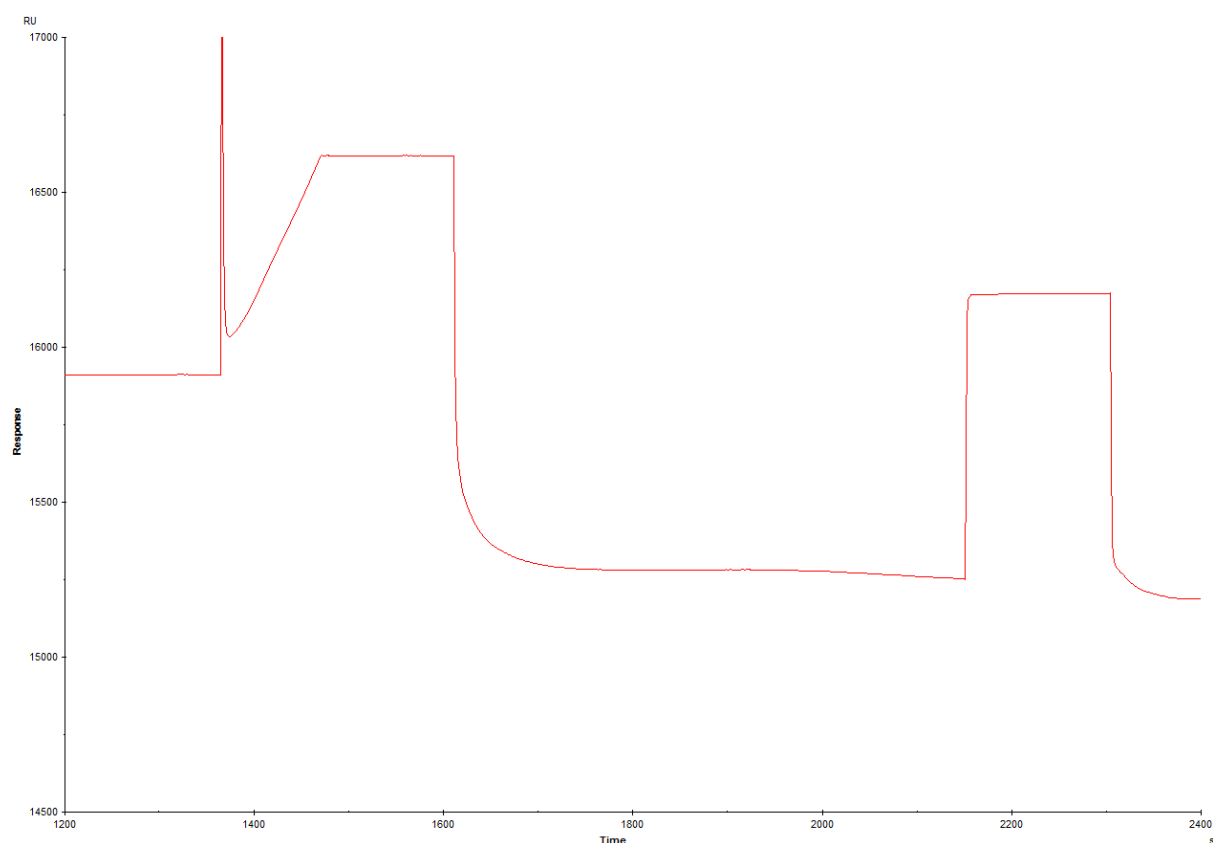


Figure 38 - CM4 Histone test with $[\text{NaCl}] = 150 \text{ mM}$

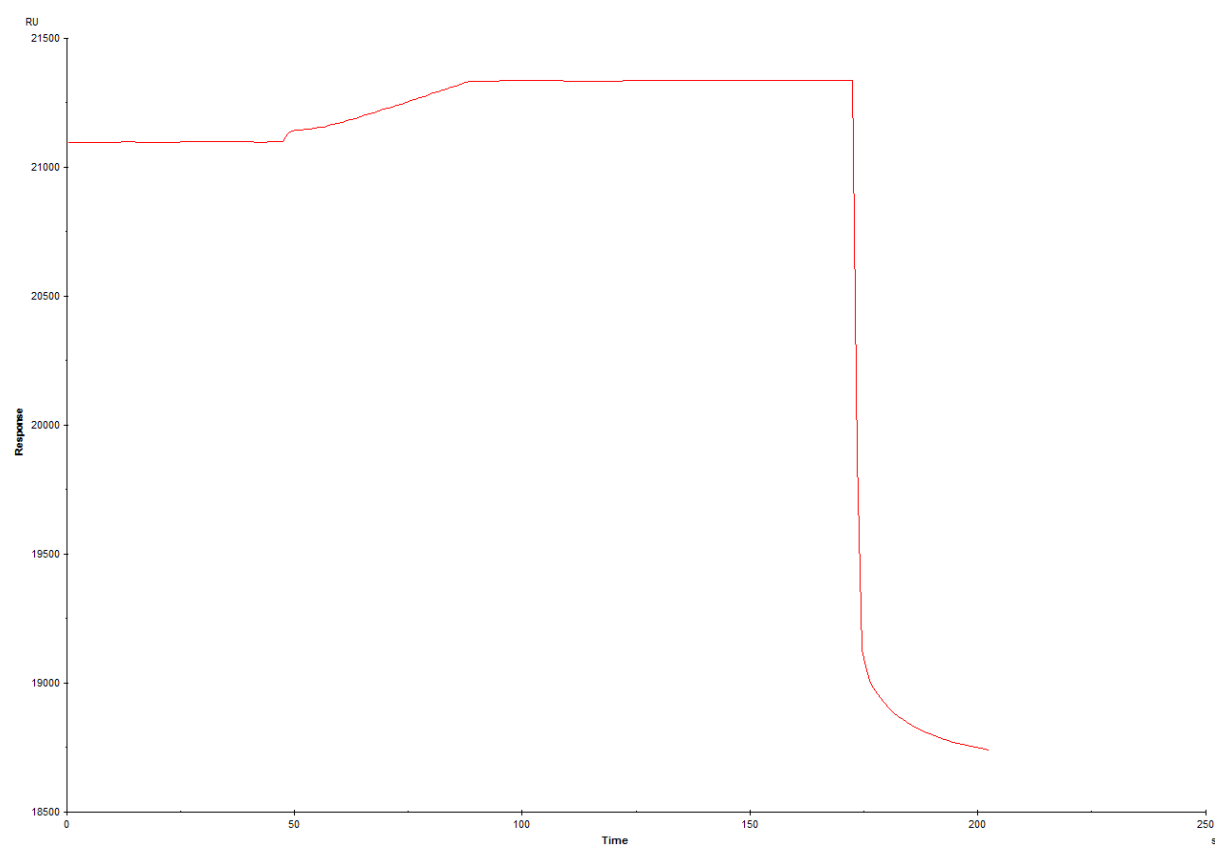


Figure 39 - CM4 Histone test with $[NaCl] = 500 \text{ mM}$

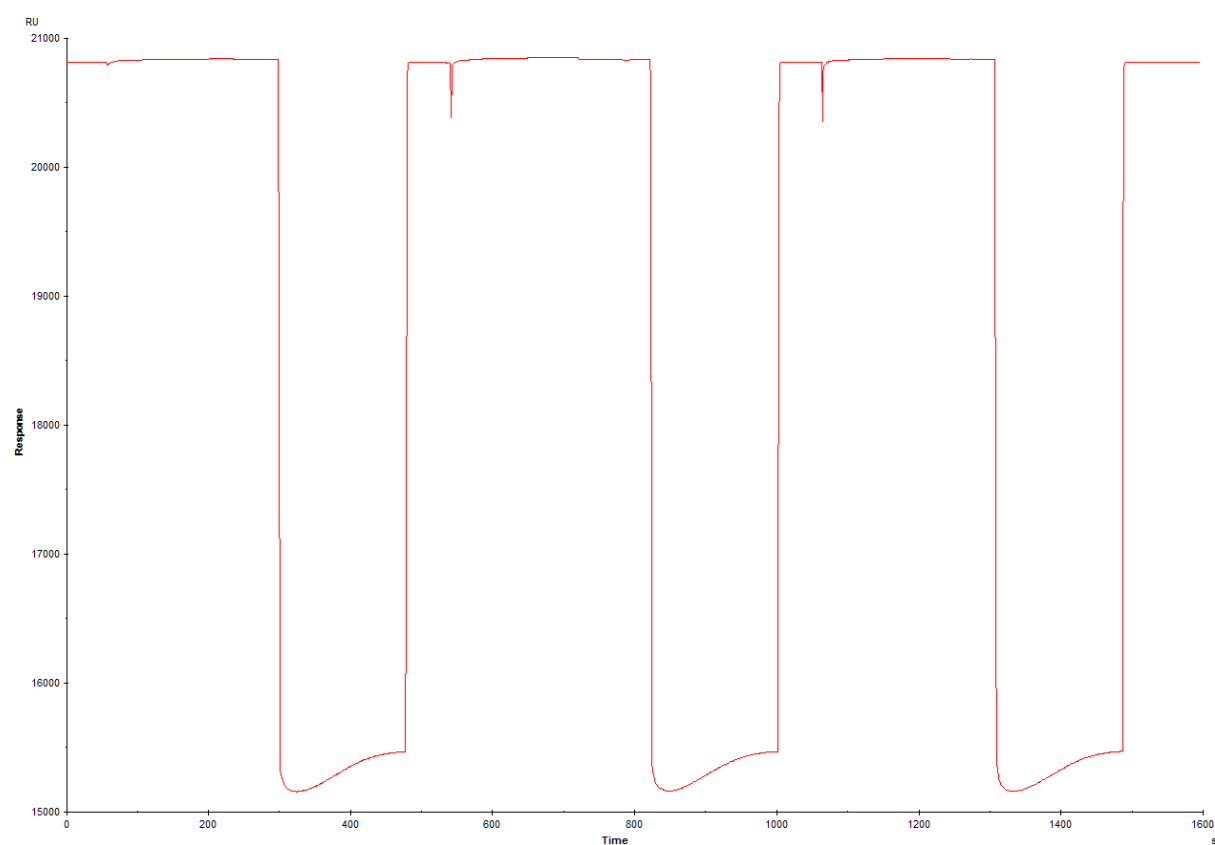


Figure 40 - CM4 Histone test with $[NaCl] = 1.00\text{ M}$

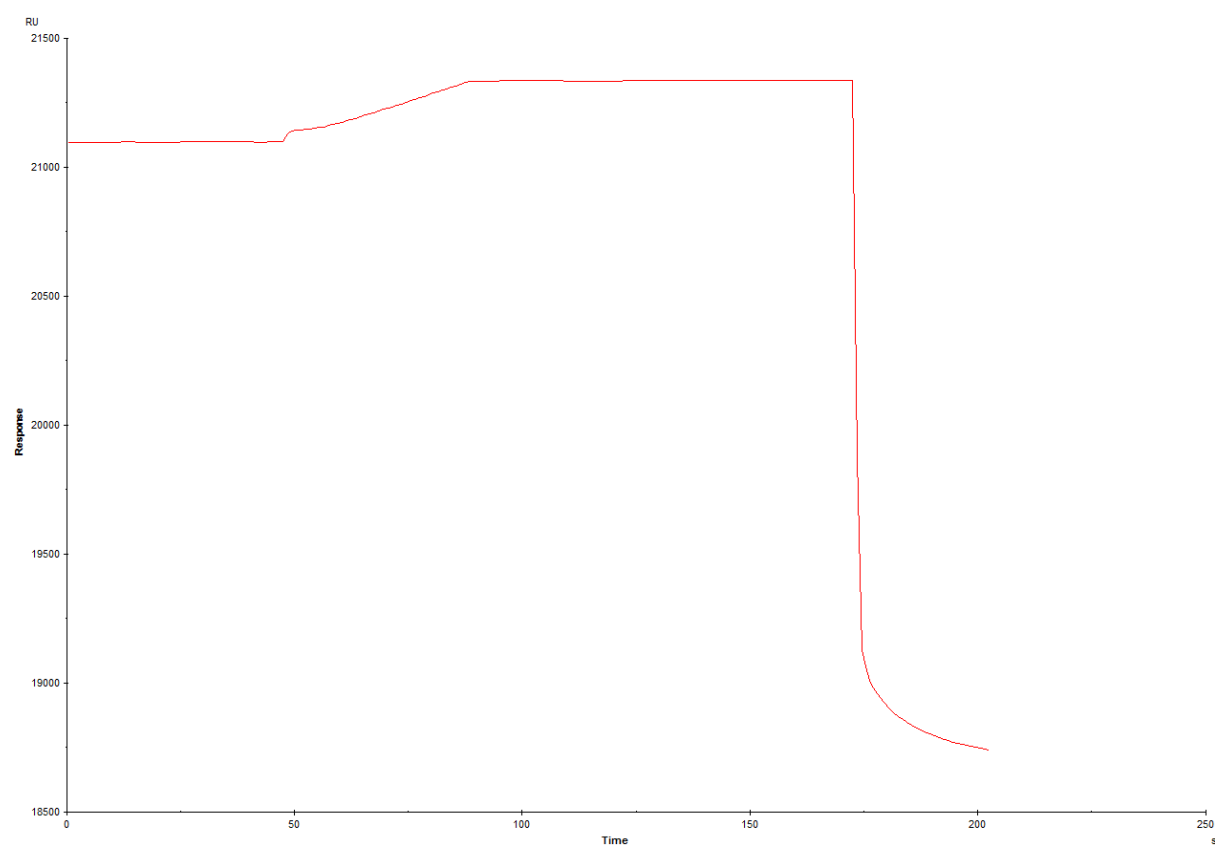


Figure 41 - CM4 Histone test with $[NaCl] = 150 \text{ mM}$ $pH = 3.5$

2.4.6.9 CM4 Binding Experiment

All injections were through flowcells 3 and 4 on the previously prepared α -Poly(His) CM 4 chip. A “coinject” was then iteratively carried out using the tabulated sequence.

3 samples of each: H3 R at 100 nM, H3 S at 100 nM, H3 R at 1 μ M and H3 S at 1 μ M

Cycle	Injection 1	Injection 2
1	Buffer	Buffer
2	Sirt2	Buffer
3	Sirt2	H3
4	Buffer	H3
5	Buffer	Buffer

The sonograms resulting from 2 and 4 were then subtracted from the sonogram resulting from cycle 3, an unprocessed example can be seen in *figure 21*. The flow rate was set at 30 μ L min⁻¹ and all samples as well as the running buffer were in HBS-P with the [NaCl] altered to 1.0 M

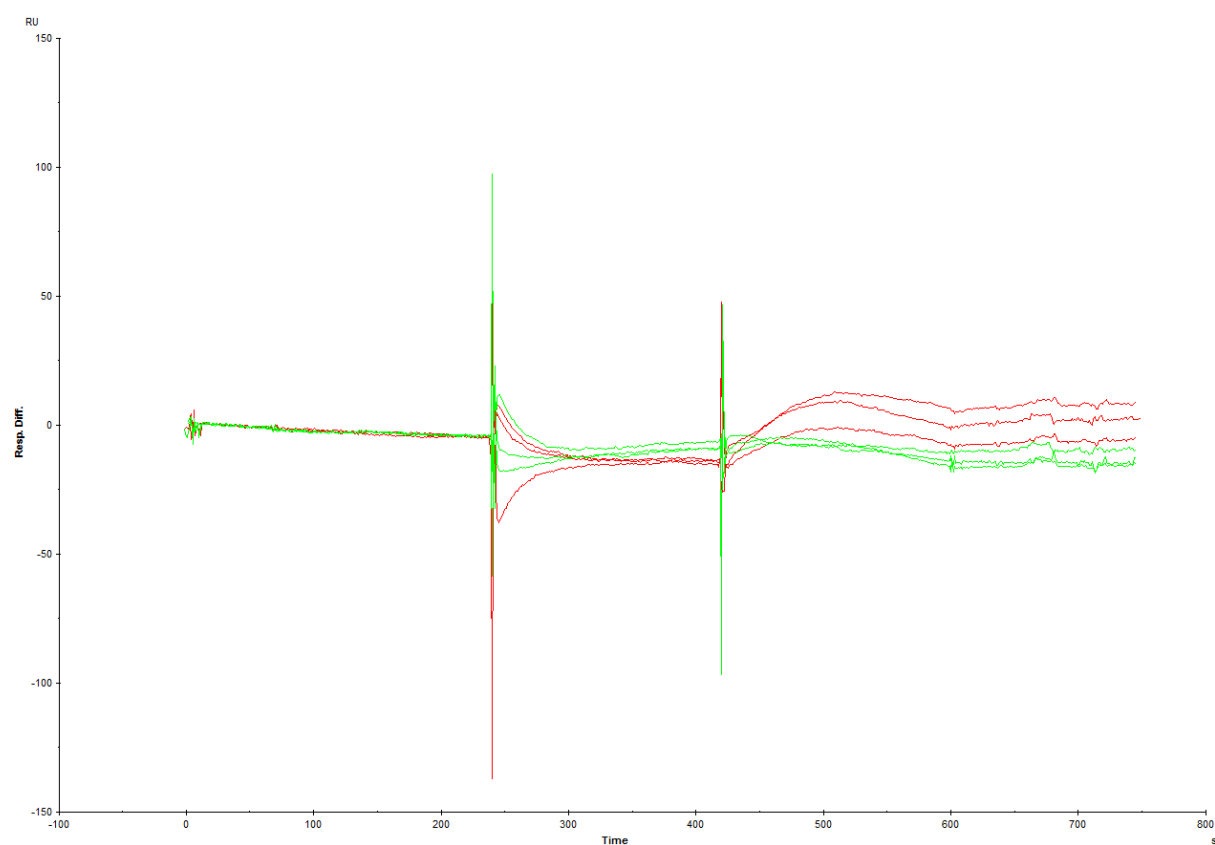


Figure 42- 100 nM H3 **R** and H3 **S** after processing

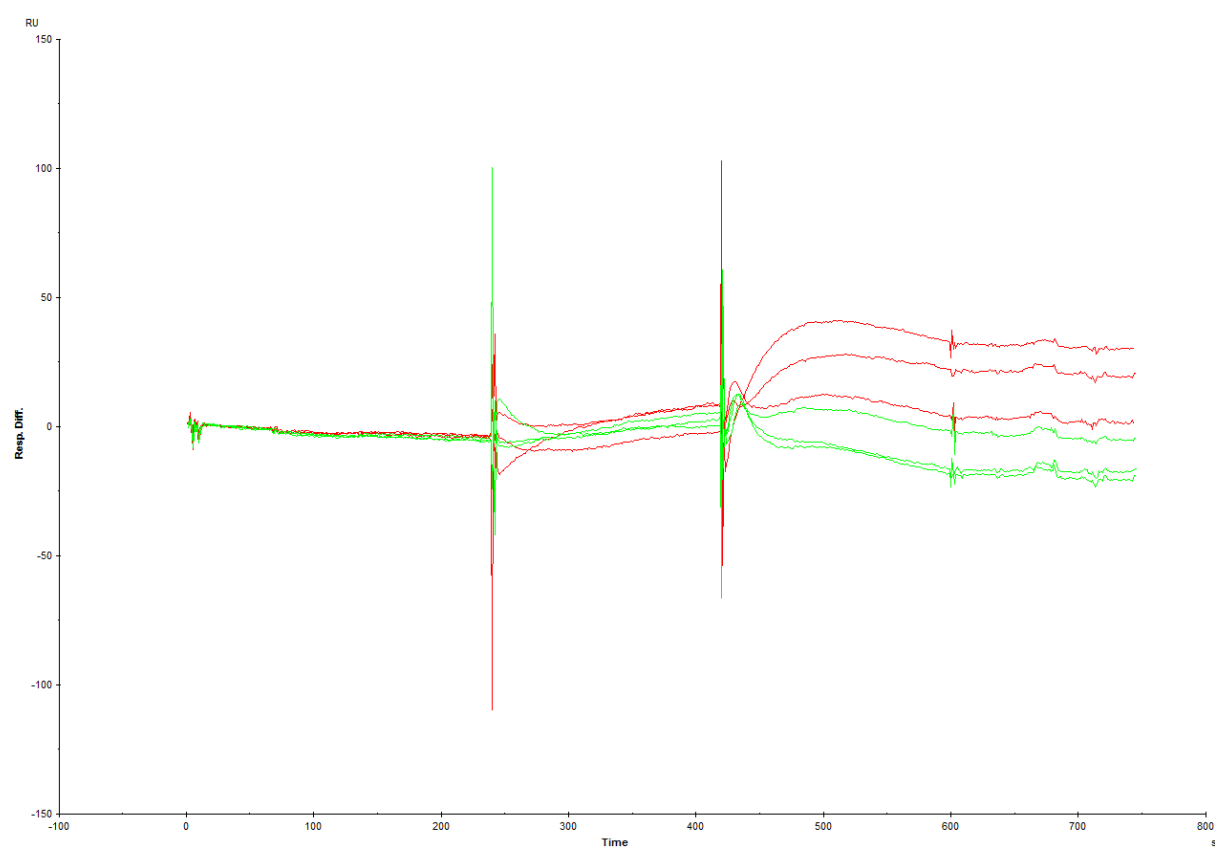


Figure 43 - 1 μ M H3 **R** and H3 **S** after processing

2.4.6.10 Amine Surface Antibody coupling

Preparation of carbonyldrazide surface: A 20 mM solution of carbonyldrazide in water was prepared. At 10 mL min⁻¹ 50 µL of both EDC solution and NHS solution were mixed manually then 50 µL of this mixture injected through flowcells 3 and 4 on a GE Healthcare CM4 chip. Following activation, 150 µL of the carbonyldrazide stock was injected. A 10 µL sample of 100 nM H3 **R** was then injected, the $\frac{dR}{dt}$ of the injection was monitored, showing 5.5 initially. Another round of carbonyldrazide coupling was performed giving $\frac{dR}{dt} = 2.0$ RU s⁻¹. After a third round, $\frac{dR}{dt} = 1.3$ RU s⁻¹. We assumed this would be low enough that no significant unspecific binding would be seen.

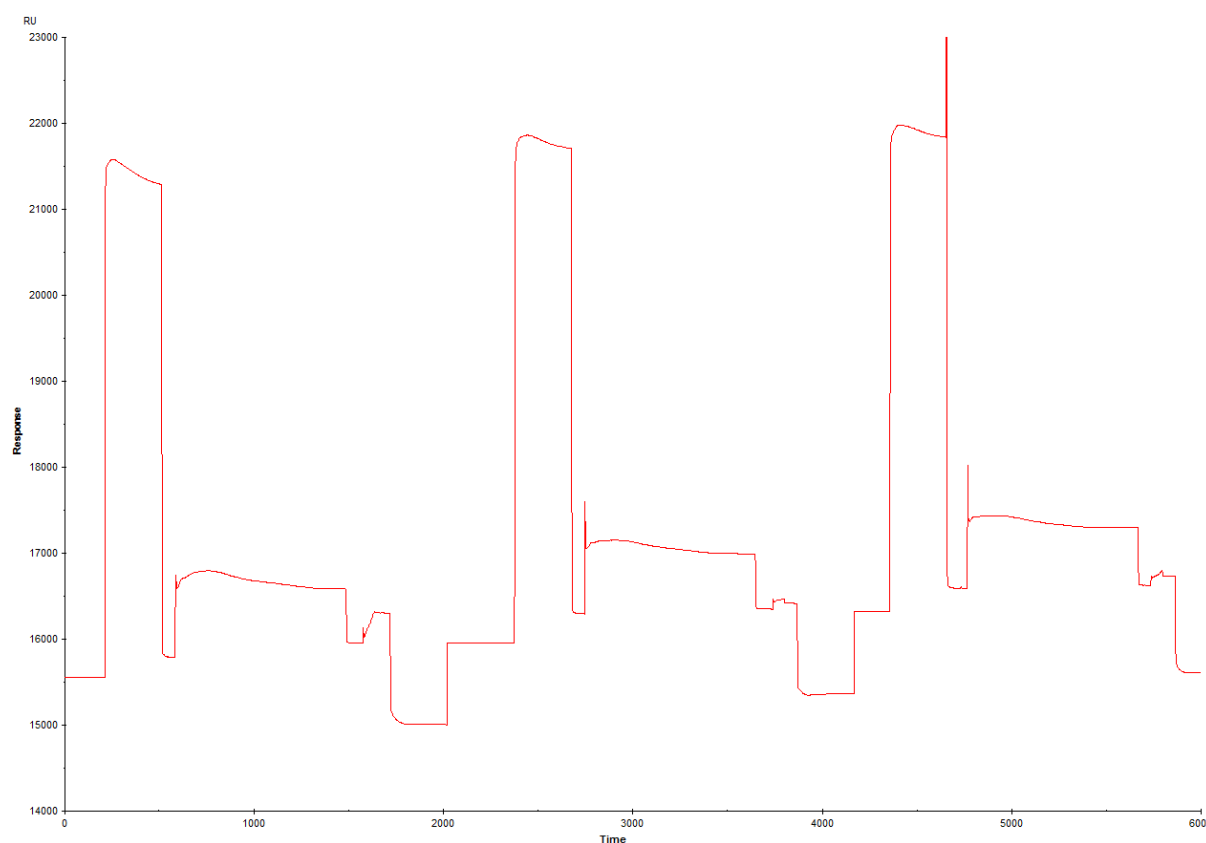


Figure 44 - Carbonyldrazide coupling and histones tests to CM4

Antibody Oxidation: 10 μL of monoclonal murine α -Poly(His) antibody mixed with 90 μL of pH 5.5 acetate buffer and cooled on ice. The antibody was treated with NaIO_4 stock (2 μL of 10.7 mg mL^{-2}) and stood on ice for 30 minutes. After 30 minutes the reaction mixture was quenched by dilution into 4 mL of pH 4.0 acetate buffer then used immediately.

Antibody Coupling: At 10 $\mu\text{L min}^{-1}$ the antibody prepared in the previous step was flowed over flowcell 4 until 9 kRU of response was observed, the flowrate was switched to 2 $\mu\text{L min}^{-1}$ and 40 μL at 10 mg mL^{-1} of NaBH_3CN in acetate buffer pH 4.0 was flowed over, the process had to be repeated 3 times in order to get the desired response, giving a total of 2400 RU.

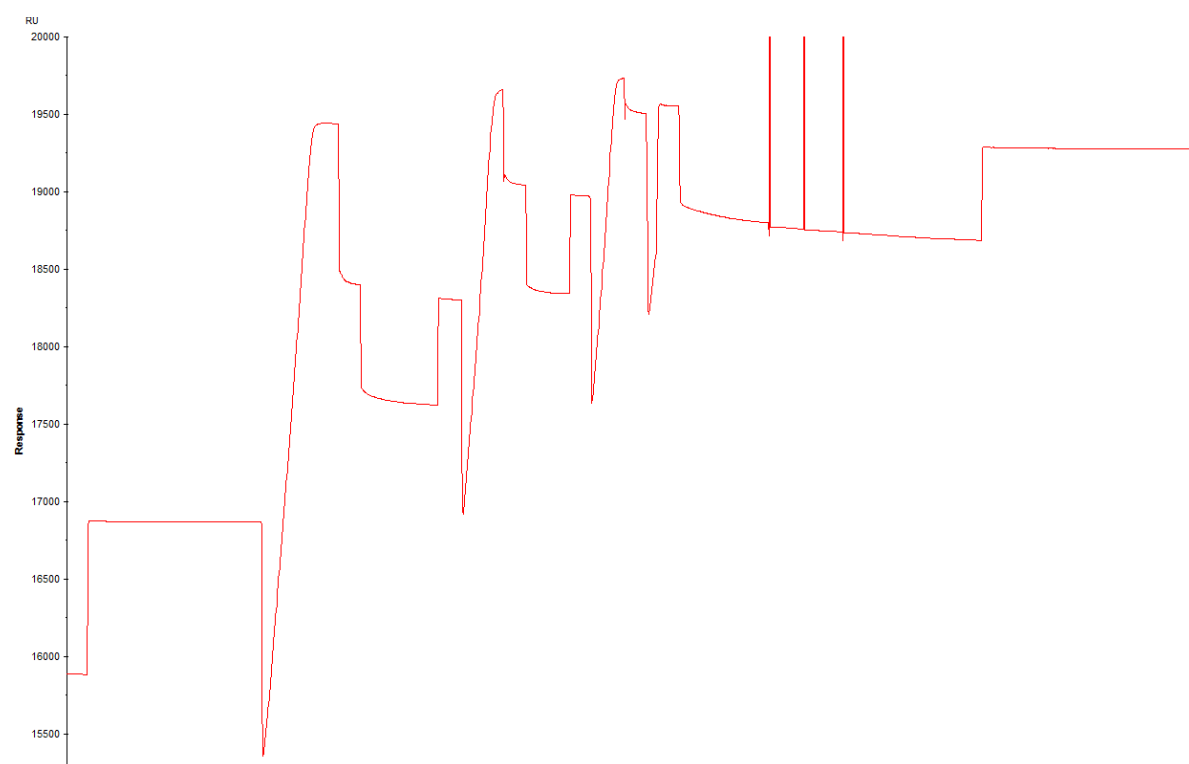


Figure 45 – Antibody coupling and reductive amination

2.4.6.11 Amine Surface Sirt2 Calibration

Serial dilutions of Sirt2 50-356 were carried out in HBS-P (to obtain 30, 10, 3, 1, 0.3 and 0.1 $\mu\text{g mL}^{-1}$). Sirt2 samples were flowed through flowcells 3 and 4 (40 μL at 10 $\mu\text{L min}^{-1}$) followed by HBS-P buffer (10 μL at 10 $\mu\text{L min}^{-1}$) and finally HCl (10 mM in water) 10 μL at 10 $\mu\text{L min}^{-1}$. We found that the binding of Sirt2 to the antibody was somewhat weak causing Sirt2 to be readily washed off the chip

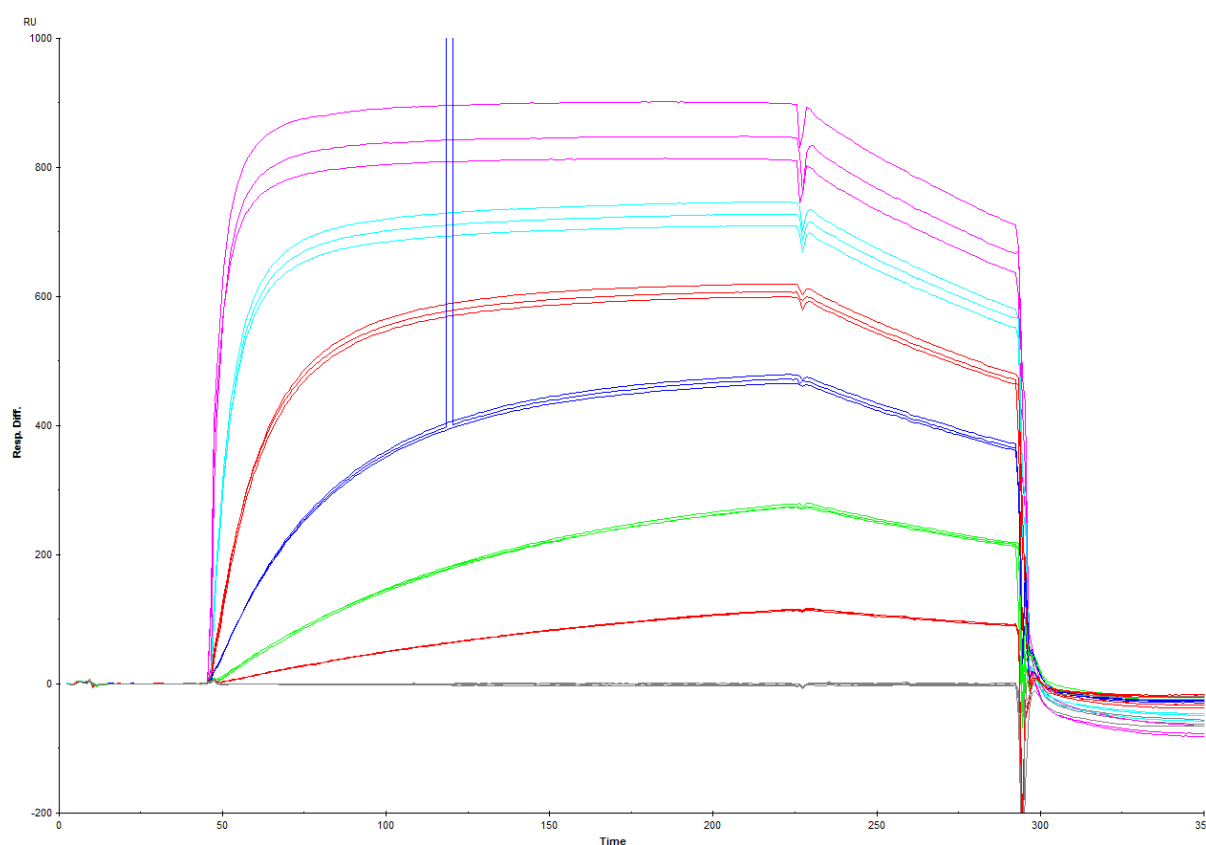


Figure 46 - Sirt2 Calibration to Amine Surface Chip

2.4.6.12 Carbonyhydrazide Oxidised Antibody Histone Test

A 100 nm sample of H3 **R** in HBS-P was flowed at $30 \mu\text{L min}^{-1}$ over the amine linked surface created in 4.6.10 (flowcell 3). Three injections of $90 \mu\text{L}$ were carried out. Significant binding was noted in each case.



Figure 47 - Histone Test on Amine surface chip in HBS-P

2.4.6.13 Noncovalent Histone Test

Flowcell 4 of a GE Healthcare CM5 Biacore chip in HBS-P at $30 \mu\text{L min}^{-1}$ was treated with $10 \mu\text{L}$ of 100 nM H3 **R** in HBS-P, $90 \mu\text{L}$ of 100 nM Sirt2 in HBS-P was then flowed over flowcell 4 and the response monitored. The flowcell was washed twice with $90 \mu\text{L}$ of 40 mM HCl and the process repeated for the H3 **S**. While $\frac{dR}{dt}$ appeared steeper for the H3 **R**, we did not think this was a good experiment and did not continue to explore this angle.



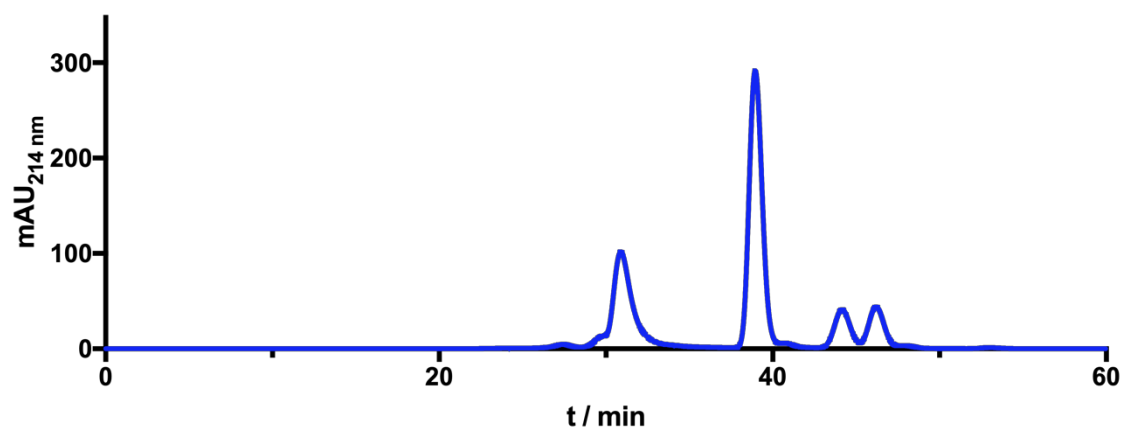
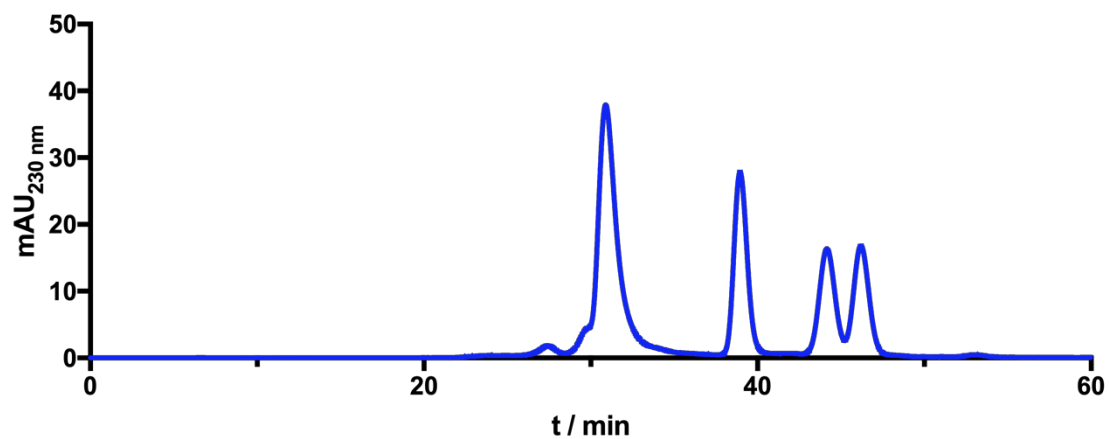
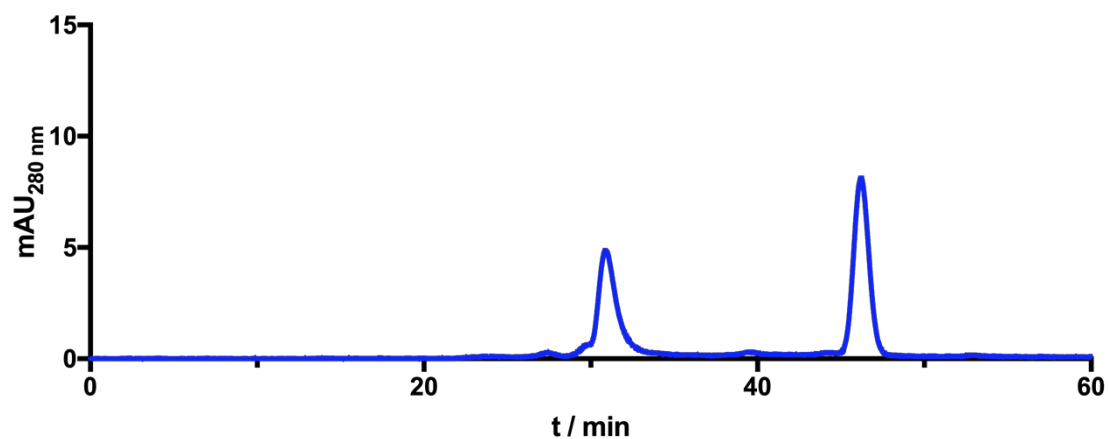
Figure 48 -Noncovalent binding of histones **R** followed by flowing Sirt2 over the chip and acid wash, repeated with **S** variant

2.4.7.0 SEC-MALS Experiments

Sample	Sirt2 / μmol	H3 S / μmol	H3 R / μmol
1	10	0	0
2	0	10	0
3	0	0	10
4	10	5	0
5	10	0	5
6	10	10	0
7	10	0	10
8	10	15	0
9	10	0	15

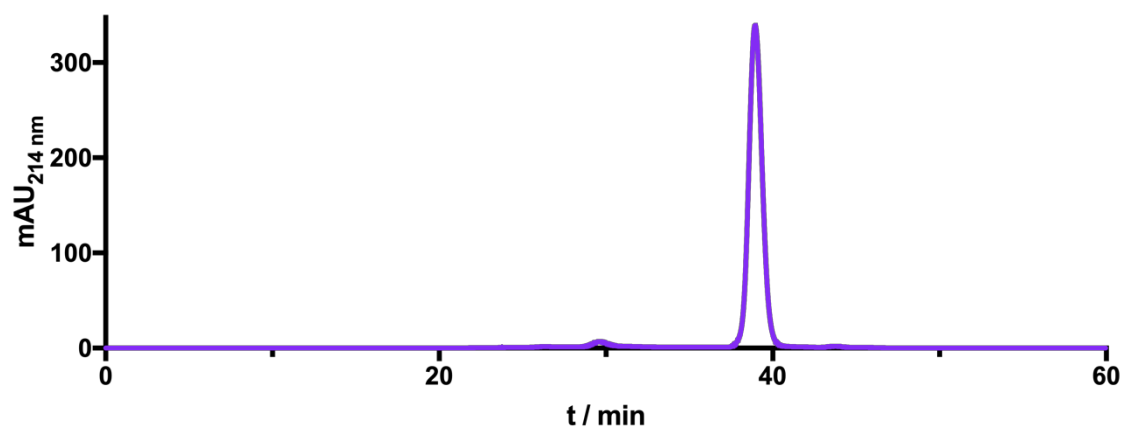
Samples (concentrations tabulated above) were prepared in HBS-P to make a total sample volume of 50 μL , then stood for 1 h at 4 °C prior to sequentially running down (40 μL injections) a superdex s200 10/30 column at 500 $\mu\text{L min}^{-1}$ both 214 nm and 280 nm were monitored.

2.4.7.1 Sirt2 Control

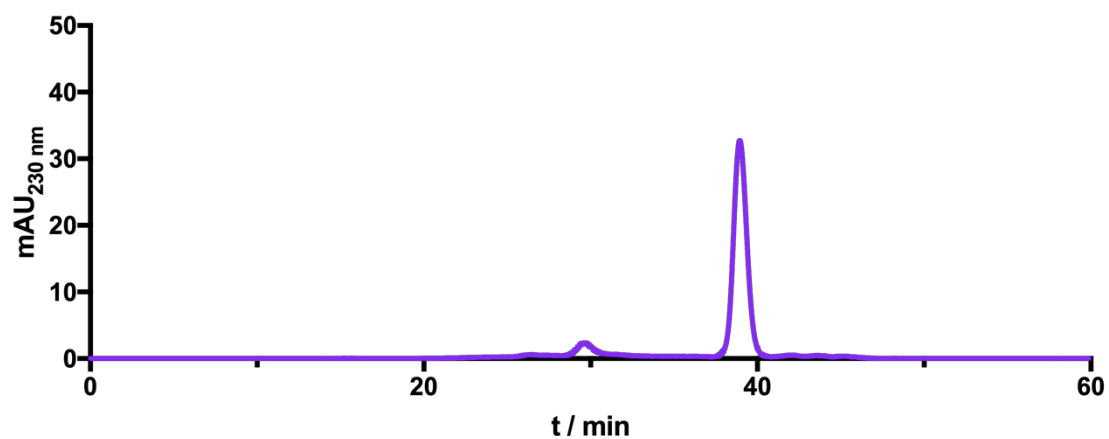
Sample 1 - 214 nm Chromatogram**Sample 1 - 230 nm Chromatogram****Sample 1 - 280 nm Chromatogram**

2.4.7.2 S Control

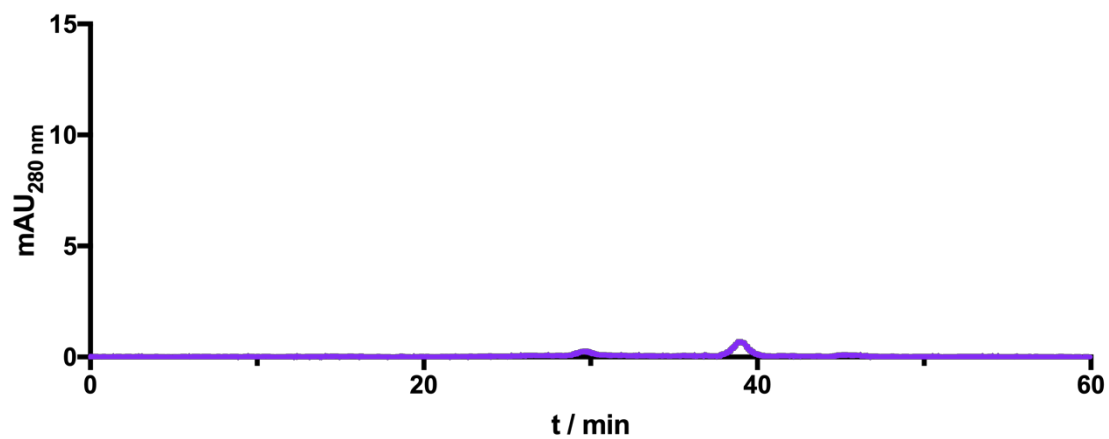
Sample 2 - 214 nm Chromatogram



Sample 2 - 230 nm Chromatogram

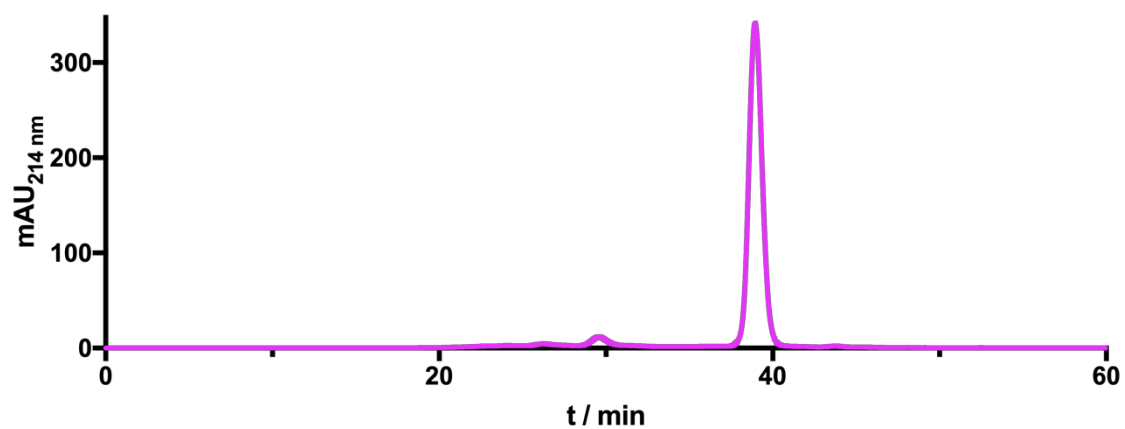


Sample 2 - 280 nm Chromatogram

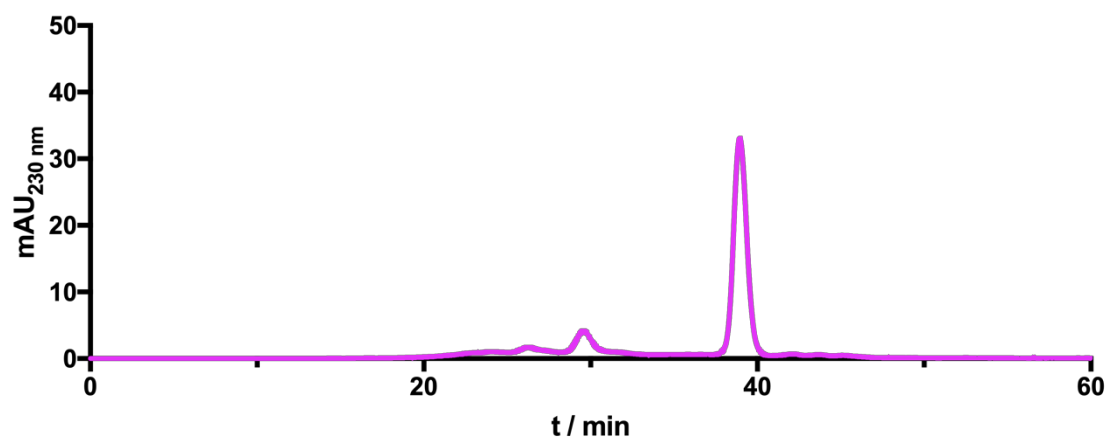


2.4.7.3 R Control

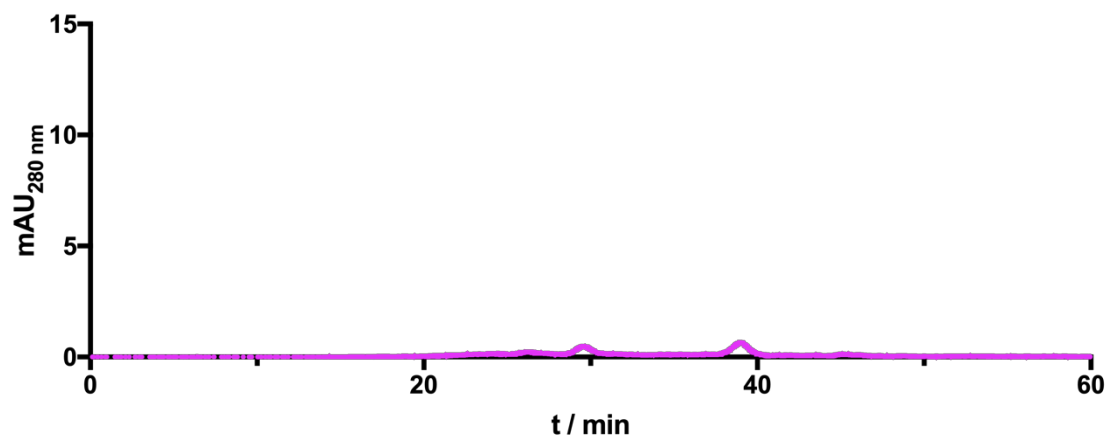
Sample 3 - 214 nm Chromatogram



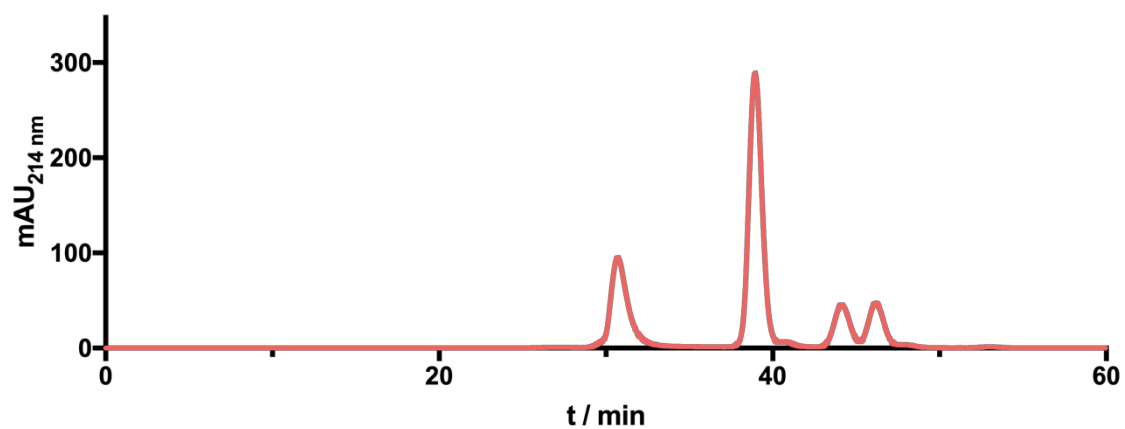
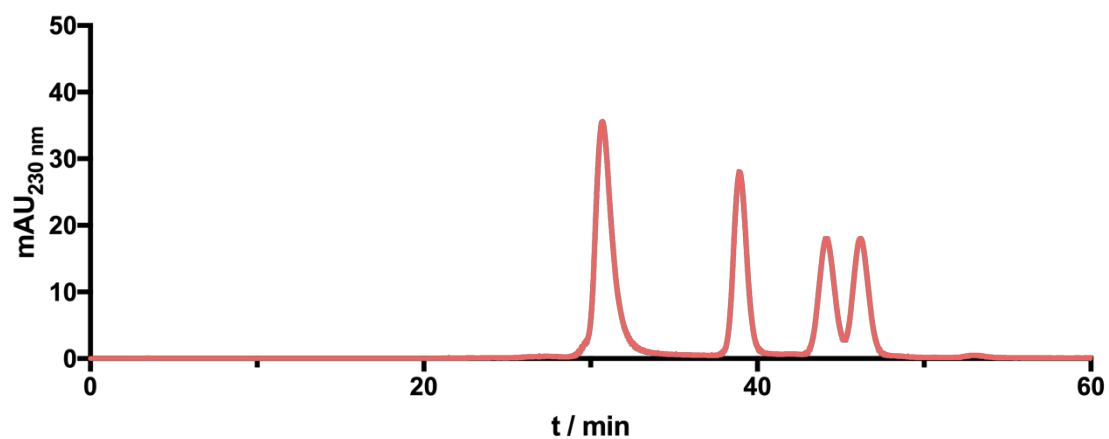
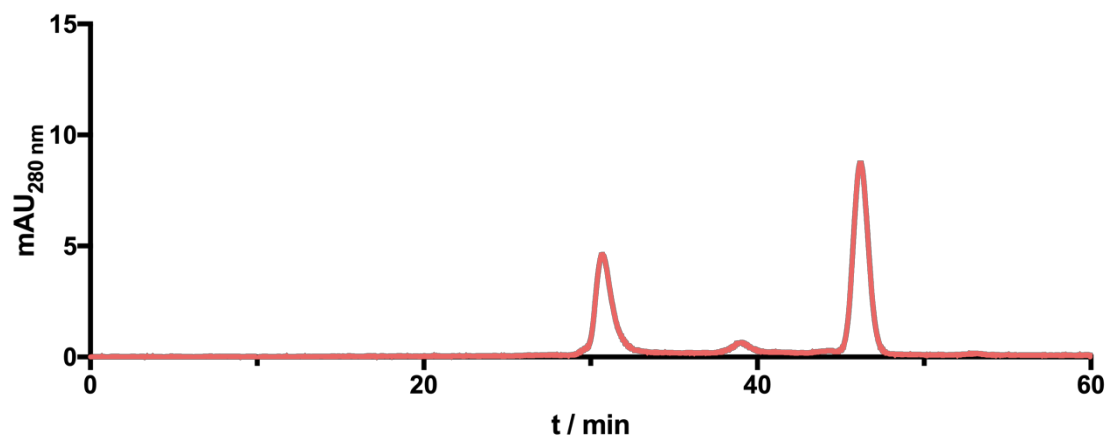
Sample 3 - 230 nm Chromatogram



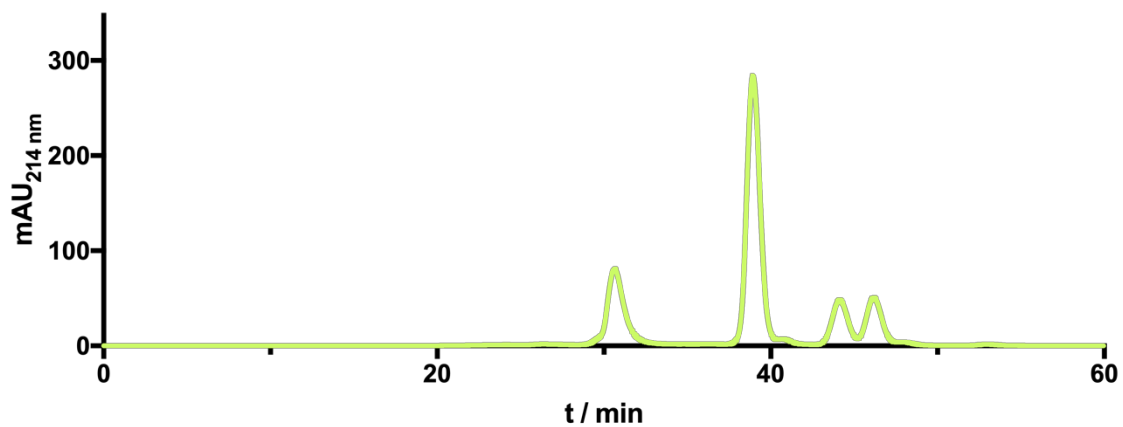
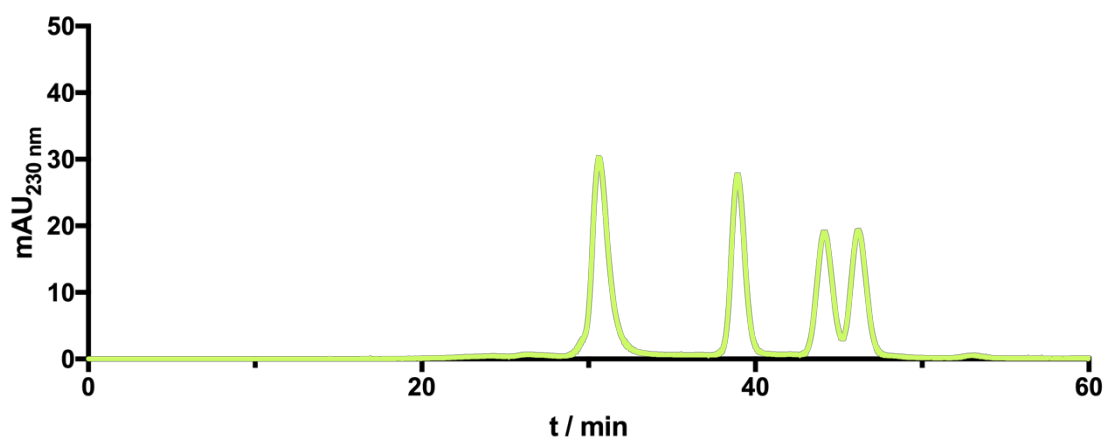
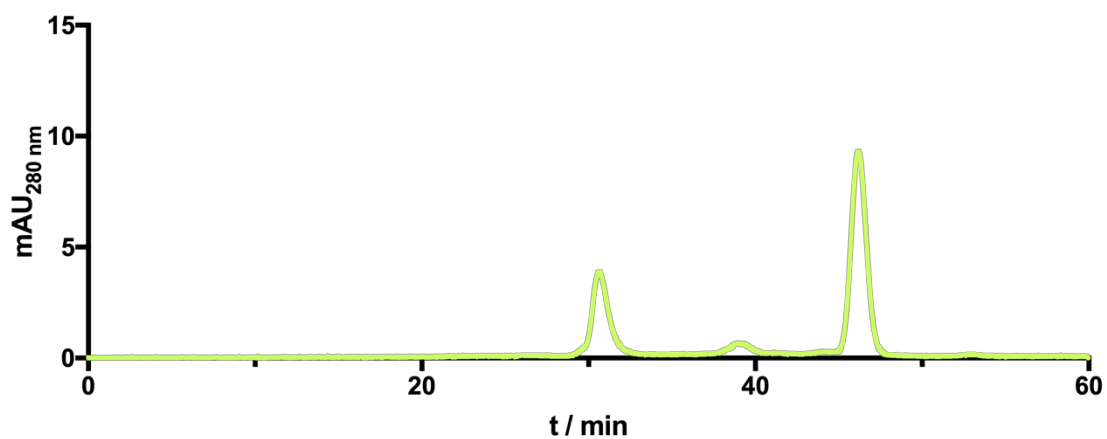
Sample 3 - 280 nm Chromatogram



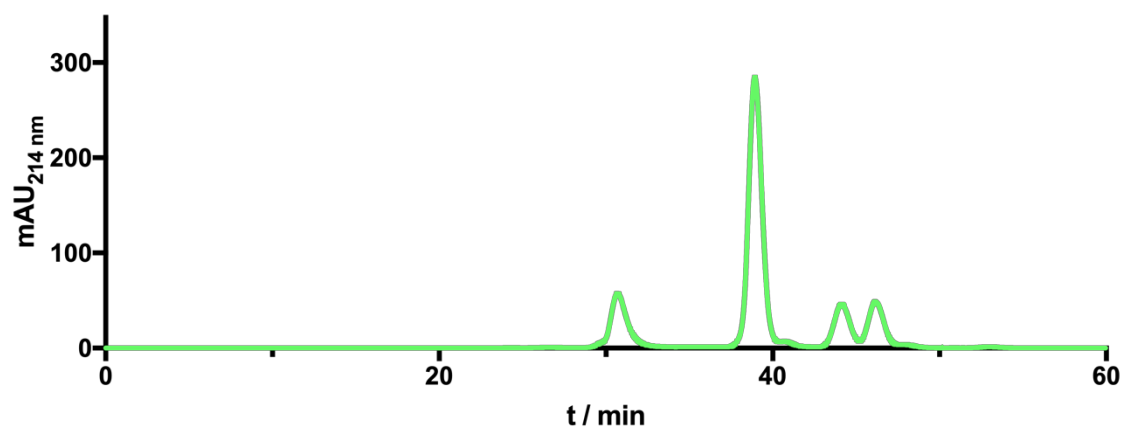
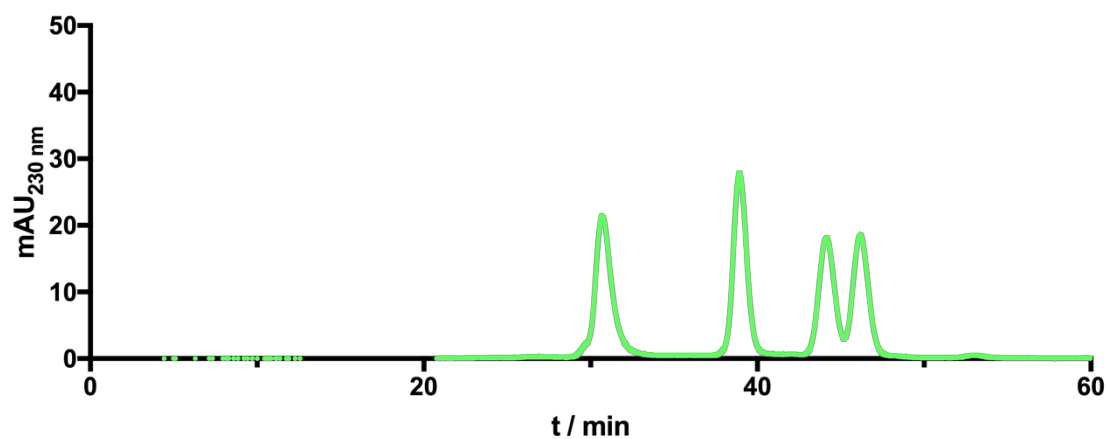
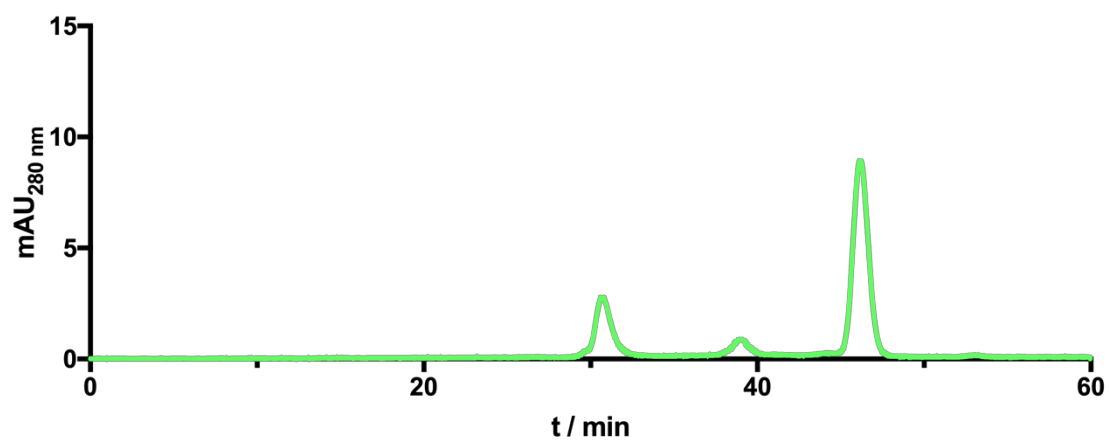
2.4.7.4 S 0.5 eq

Sample 4 - 214 nm Chromatogram**Sample 4 - 230 nm Chromatogram****Sample 4 - 280 nm Chromatogram**

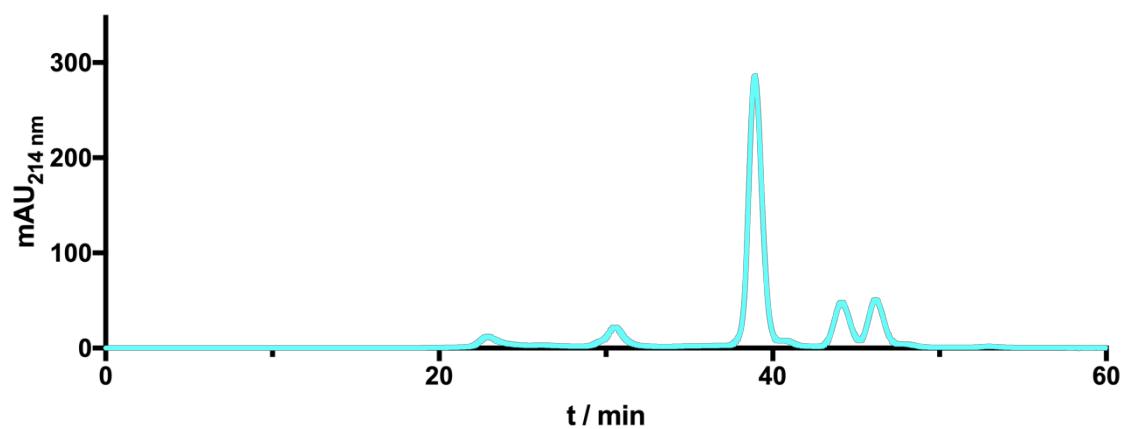
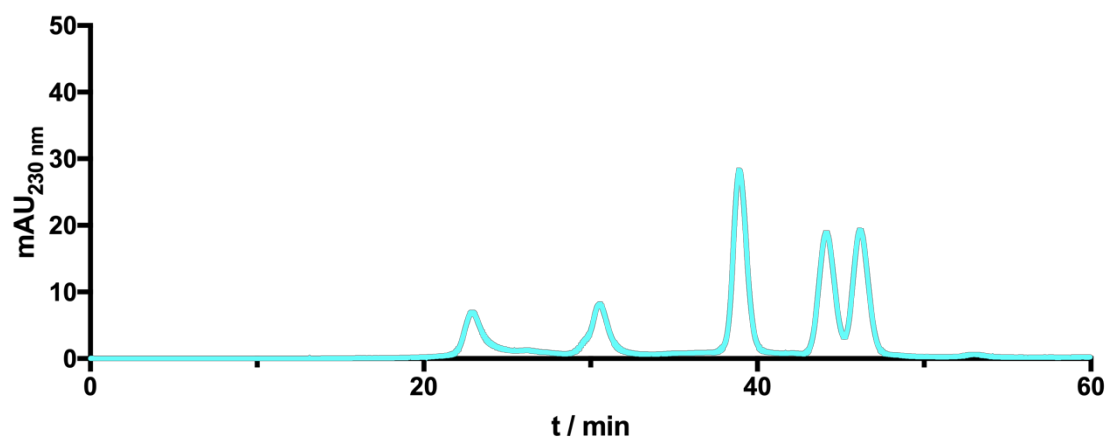
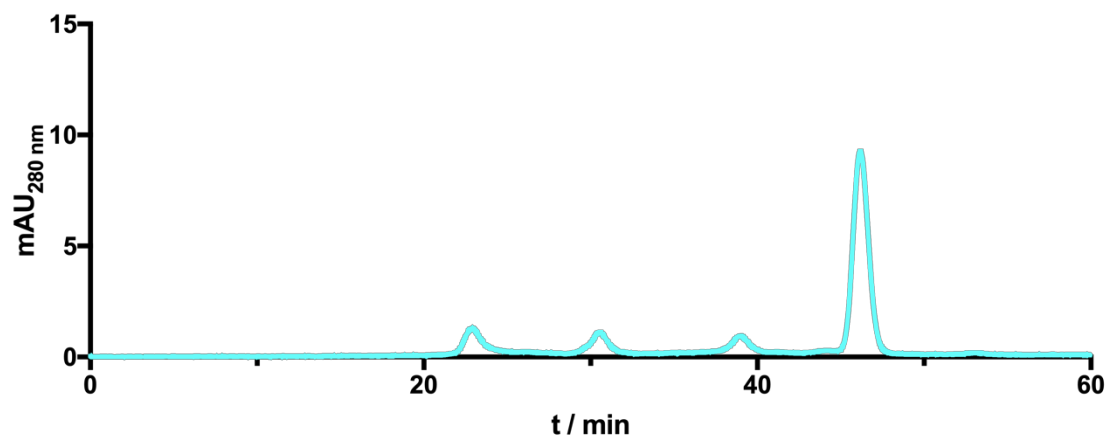
2.4.7.5 R z0.5 eq

Sample 5 - 214 nm Chromatogram**Sample 5 - 230 nm Chromatogram****Sample 5 - 280 nm Chromatogram**

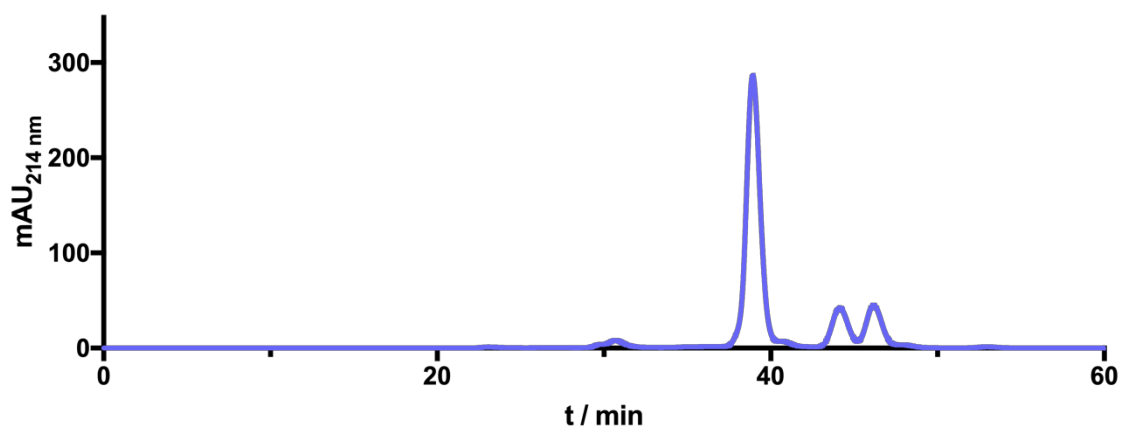
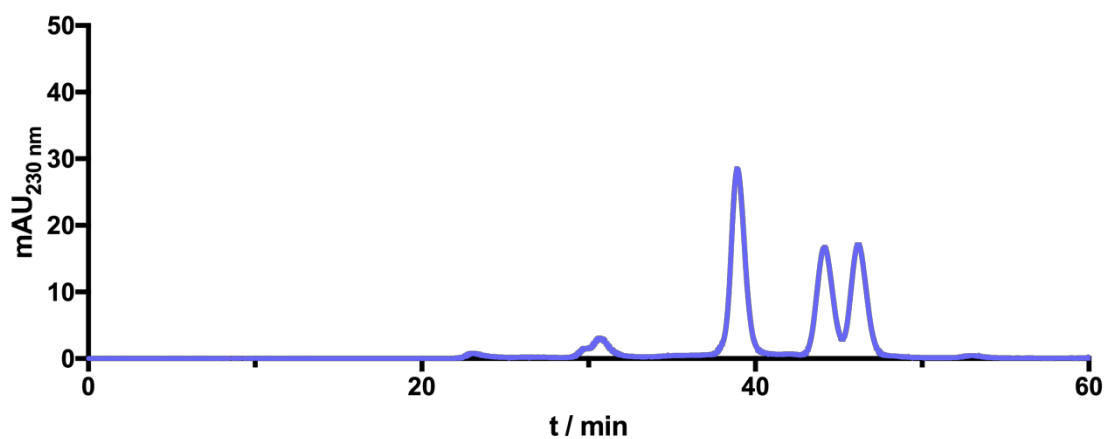
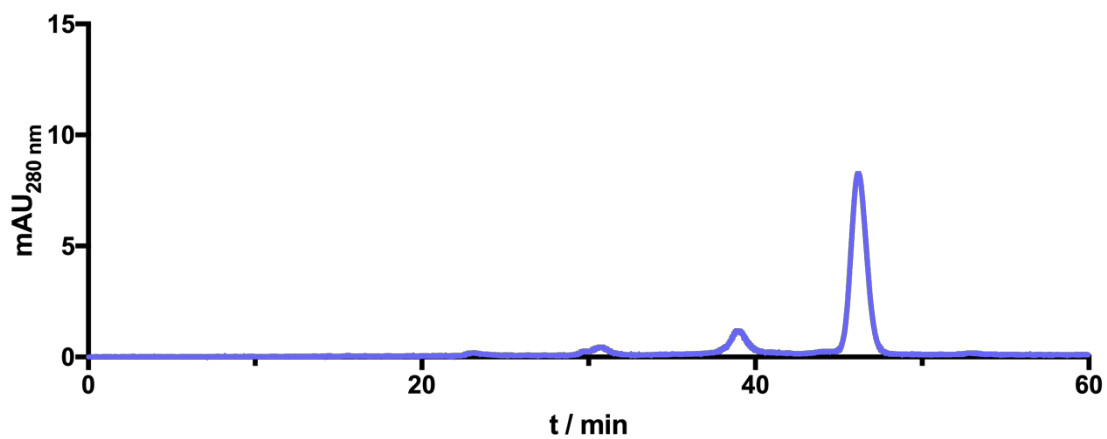
2.4.7.6 S 1.0 eq

Sample 6 - 214 nm Chromatogram**Sample 6 - 230 nm Chromatogram****Sample 6 - 280 nm Chromatogram**

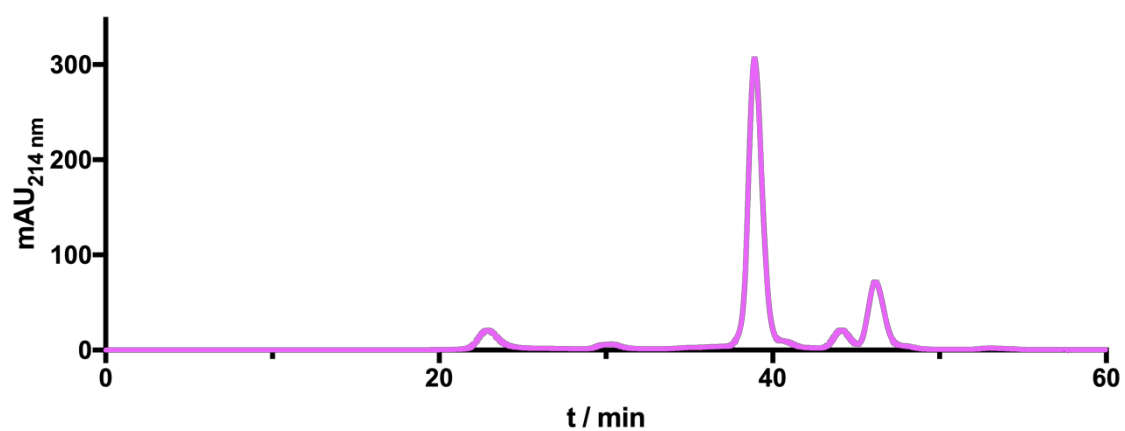
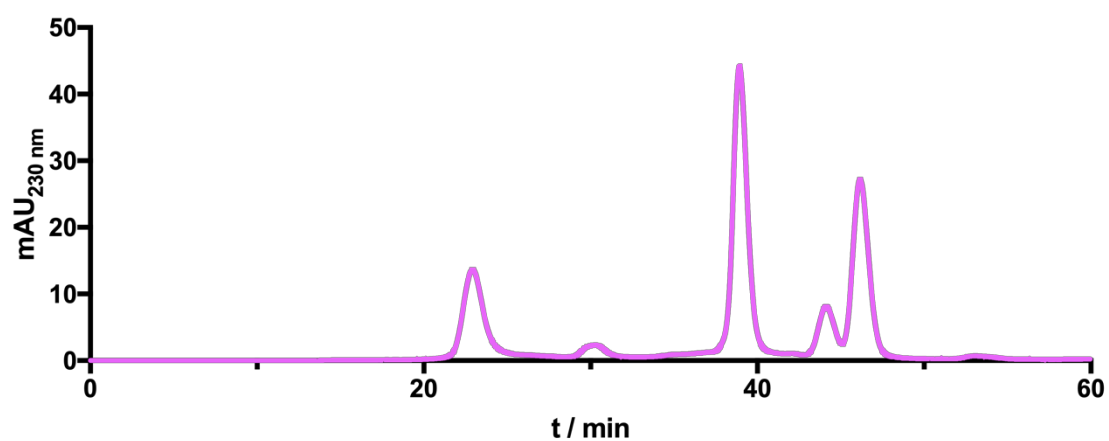
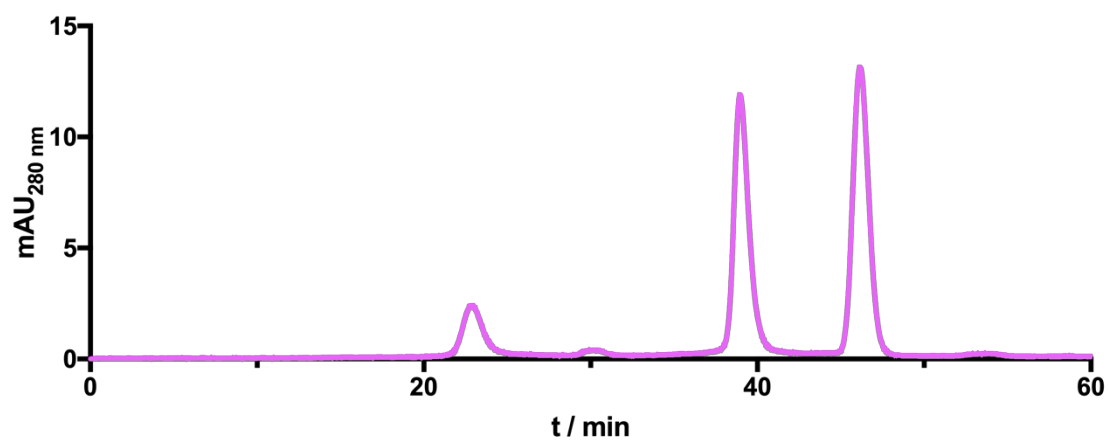
2.4.7.7 R 1.0 eq

Sample 7 - 214 nm Chromatogram**Sample 7 - 230 nm Chromatogram****Sample 7 - 280 nm Chromatogram**

2.4.7.8 S 1.5 eq

Sample 8 - 214 nm Chromatogram**Sample 8 - 230 nm Chromatogram****Sample 8 - 280 nm Chromatogram**

2.4.7.9 R 1.5 eq

Sample 9 - 214 nm Chromatogram**Sample 9 - 230 nm Chromatogram****Sample 9 - 280 nm Chromatogram**

2.4.8.0 Histone Refolding Protocol

A sample of histone (500 μL , $\sim 1 \text{ mg mL}^{-1}$) in H3 Reaction Buffer was passed through a G-25 Minitrap into refolding buffer to give 1 mL, the sample was then passed through 2 G-25 Minitraps into the desired buffer (typically either fPBS to check the protein folding or HBS-P for binding studies), to give a total volume of 2 mL. A vivaspin 500 3 kDa was used to concentrate the protein to a final volume of $\sim 100 \mu\text{L}$. The concentration was measured using A_{280} .

The concentrations found for the batch which was used for all SPR and SEC-MALS was found to be **S**: 1.2 mg mL^{-1} **R**: 0.67 mg mL^{-1} .

2.4.9.0 Sequences

2.4.9.1 His₆ TEV *Human Sirtuin 2* 56-365:

GSSHHHHHHSSGENLYFQGSERLLEDLTLEGVARYMQSERCRRVICLVGAGISTSAGIPDFR
SPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYPGQFKPTICHYFMRLKDKG
LLRCYTQNIDTLERIALGLEQEDLVEAHGTFYTSHCVSASCRHEYPLSWMKEKIFSEVTPKC
EDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLVMGTSLQVQPFASLISKAPLSTPRL
LINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAWLGECDQGCLALAEELGWKKELEDLV
RREHASIDAQ

2.4.9.2 His₆ TEV *Human Sirtuin 2* 50-365:

GSSHHHHHHSSGENLYFQGSSSLGSQKERLLEDLTLEGVARYMQSERCRRVICLVGAGISTSA
GIPDFRSPSTGLYDNLEKYHLPYPEAIFEISYFKKHPEPFFALAKELYPGQFKPTICHYFMR
LLKDKGLLLRCYTQNIDTLERIALGLEQEDLVEAHGTFYTSHCVSASCRHEYPLSWMKEKIFS
EVTPKCEDCQSLVKPDIVFFGESLPARFFSCMQSDFLKVDLLVMGTSLQVQPFASLISKAP
LSTPRL LINKEKAGQSDPFLGMIMGLGGGMDFDSKKAYRDVAWLGECDQGCLALAEELGWKK
ELEDLVRREHASIDAQ

2.4.9.3 Histone H3 K18C, C110A:

ARTKQTARKSTGGKAPRCQLATKAARKSAPATGGVKKPHRYRPGTVALREIRRYQKSTELLI
RKLPFQRLVREIAQDFKTDLRFQSSAVMALQEASEAYLVALFEDTNLA^AIAHAKRVTIMPKDI
QLARRIRGERA

2.4.9.4 Histone H4 K16C:

SGRGKGGKGLGKGGA^CRHRKVLRDNIQGITKPAIRRLARRGGVKRISGLIYEETRGLVKVFL
ENVIRDAVTYTEHAKRKTVTAMDVVYALKRQGRTLYGFGG

2.5.0.0 References

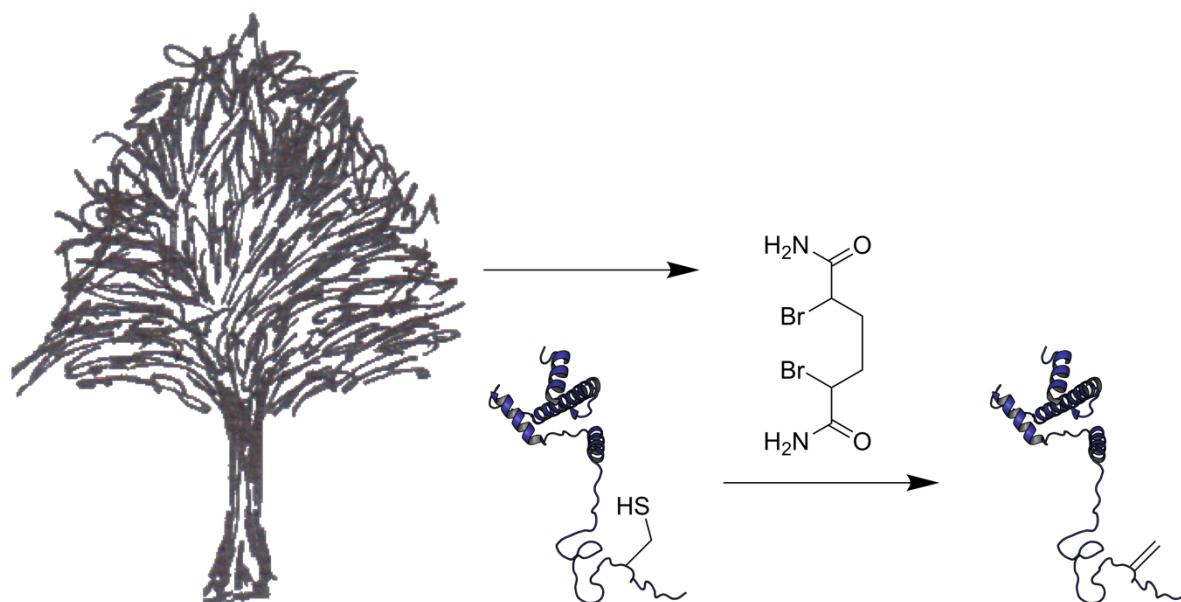
146. R. Raj, L. Lercher, S. Mohammed and B. G. Davis, *Angewandte Chemie*, 2016, **128**, 9064-9068.
154. T. H. Wright, B. J. Bower, J. M. Chalker, G. J. L. Bernardes, R. Wiewiora, W.-L. Ng, R. Raj, S. Faulkner, M. R. J. Vallée, A. Phantumruiwath, O. D. Coleman, M.-L. Thézénas, M. Khan, S. R. G. Galan, L. Lercher, M. W. Schombs, S. Gerstberger, M. E. Palm-Espling, A. J. Baldwin, B. M. Kessler, T. D. W. Claridge, S. Mohammed and B. G. Davis, *Science*, 2016, **354**.
155. D. R. Lide, *Handbook of Chemistry and Physics*, CRC Press, Taylor & Francis Group, Boca Raton, 90th edn., 2009.
181. M. Imiołek, G. Karunanithy, W.-L. Ng, A. J. Baldwin, V. Gouverneur and B. G. Davis, *Journal of the American Chemical Society*, 2018, **140**, 1568-1571.
217. P. B. Becker, *Chromatin Protocols (Methods in Molecular Biology)*, Humana Press, New York City, 1999.
222. T. Jenuwein and C. D. Allis, *Science*, 2001, **293**, 1074.
223. E. Verdin and M. Ott, *Nat. Rev. Mol. Cell Biol.*, 2014, **16**, 258.
224. A. Eberhardter and P. B. Becker, *EMBO Rep.*, 2002, **3**, 224.
225. M. Braunstein, A. B. Rose, S. G. Holmes, C. D. Allis and J. R. Broach, *Genes Dev.*, 1993, **7**, 592-604.
226. R. A. Frye, *Biochem. Biophys. Res. Commun.*, 2000, **273**, 793-798.
227. B. J. North and E. Verdin, *Genome Biol.*, 2004, **5**, 224.
228. J. S. Smith, C. B. Brachmann, I. Celic, M. A. Kenna, S. Muhammad, V. J. Starai, J. L. Avalos, J. C. Escalante-Semerena, C. Grubmeyer, C. Wolberger and J. D. Boeke, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 6658.
229. C. B. Brachmann, J. M. Sherman, S. E. Devine, E. E. Cameron, L. Pillus and J. D. Boeke, *Genes Dev.*, 1995, **9**, 2888-2902.
230. E. S. Miller, J. F. Heidelberg, J. A. Eisen, W. C. Nelson, A. S. Durkin, A. Ciecko, T. V. Feldblyum, O. White, I. T. Paulsen, W. C. Nierman, J. Lee, B. Szczypinski and C. M. Fraser, *J. Bacteriol.*, 2003, **185**, 5220.
231. S.-i. Imai, C. M. Armstrong, M. Kaeberlein and L. Guarente, *Nature*, 2000, **403**, 795.
232. J. Landry, A. Sutton, S. T. Tafrov, R. C. Heller, J. Stebbins, L. Pillus and R. Sternglanz, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 5807.
233. J. Du, H. Jiang and H. Lin, *Biochem.*, 2009, **48**, 2878-2890.
234. M. Schiedel, D. Robaa, T. Rumpf, W. Sippl and M. Jung, *Med. Res. Rev.*, 2018, **38**, 147-200.
235. B. J. North, B. L. Marshall, M. T. Borra, J. M. Denu and E. Verdin, *Mol. Cell*, 2003, **11**, 437-444.

236. Y.-H. Jin, Y.-J. Kim, D.-W. Kim, K.-H. Baek, B. Y. Kang, C.-Y. Yeo and K.-Y. Lee, *Biochem. Biophys. Res. Commun.*, 2008, **368**, 690-695.
237. M. H. Yang, G. Laurent, A. S. Bause, R. Spang, N. German, M. C. Haigis and K. M. Haigis, *Mol. Cancer Res.*, 2013, **11**, 1072.
238. H. Huang, D. Zhang, Y. Wang, M. Perez-Neut, Z. Han, Y. G. Zheng, Q. Hao and Y. Zhao, *Nat. Commun.*, 2018, **9**, 3374.
239. Y.-B. Teng, H. Jing, P. Aramsangtienchai, B. He, S. Khan, J. Hu, H. Lin and Q. Hao, *Sci. Rep.*, 2015, **5**, 8529.
240. M. S. Finnin, J. R. Donigian and N. P. Pavletich, *Nat. Struct. Biol.*, 2001, **8**, 621.
241. N. Kudo, A. Ito, M. Arata, A. Nakata and M. Yoshida, *Philos Trans R. Soc. B Biol. Sci.*, 2018, **373**.
242. T. Rumpf, S. Gerhardt, O. Einsle and M. Jung, *Acta Crystallogr. Sect. F*, 2015, **71**, 1498-1510.
243. M. Gertz, F. Fischer, G. T. T. Nguyen, M. Lakshminarasimhan, M. Schutkowski, M. Weyand and C. Steegborn, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, E2772.
244. T. Rumpf, M. Schiedel, B. Karaman, C. Roessler, B. J. North, A. Lehotzky, J. Oláh, K. I. Ladwein, K. Schmidtkunz, M. Gajer, M. Pannek, C. Steegborn, D. A. Sinclair, S. Gerhardt, J. Ovádi, M. Schutkowski, W. Sippl, O. Einsle and M. Jung, *Nat. Commun.*, 2015, **6**, 6263.
245. K. Yamagata, Y. Goto, H. Nishimasu, J. Morimoto, R. Ishitani, N. Dohmae, N. Takeda, R. Nagai, I. Komuro, H. Suga and O. Nureki, *Structure*, 2014, **22**, 345-352.
246. A. Vaquero, R. Sternglanz and D. Reinberg, *Oncogene*, 2007, **26**, 5505.
247. A. Vaquero, M. B. Scher, D. H. Lee, A. Sutton, H.-L. Cheng, F. W. Alt, L. Serrano, R. Sternglanz and D. Reinberg, *Genes Dev.*, 2006, **20**, 1256-1261.
248. T. Nagai, M. Ikeda, S. Chiba, S.-i. Kanno and K. Mizuno, *J. Cell Sci.*, 2013, **126**, 4369.
249. H.-S. Kim, A. Vassilopoulos, R.-H. Wang, T. Lahusen, Z. Xiao, X. Xu, C. Li, Timothy D. Veenstra, B. Li, H. Yu, J. Ji, Xin W. Wang, S.-H. Park, Yong I. Cha, D. Gius and C.-X. Deng, *Cancer Cell*, 2011, **20**, 487-499.
250. B. J. North, M. A. Rosenberg, K. B. Jeganathan, A. V. Hafner, S. Michan, J. Dai, D. J. Baker, Y. Cen, L. E. Wu, A. A. Sauve, J. M. van Deursen, A. Rosenzweig and D. A. Sinclair, *EMBO J.*, 2014, **33**, 1438.
251. X. He, H. Nie, Y. Hong, C. Sheng, W. Xia and W. Ying, *Biochem. Biophys. Res. Commun.*, 2012, **417**, 468-472.
252. F. Wang, M. Nguyen, F. X.-F. Qin and Q. Tong, *Aging Cell*, 2007, **6**, 505-514.

253. Y. P. Wang, L. S. Zhou, Y. Z. Zhao, S. W. Wang, L. L. Chen, L. X. Liu, Z. Q. Ling, F. J. Hu, Y. P. Sun, J. Y. Zhang, C. Yang, Y. Yang, Y. Xiong, K. L. Guan and D. Ye, *EMBO J.*, 2014, **33**, 1304.
254. I. J. Arinze, A. J. Garber and R. W. Hanson, *J. Biol. Chem.*, 1973, **248**, 2266-2274.
255. W. Jiang, S. Wang, M. Xiao, Y. Lin, L. Zhou, Q. Lei, Y. Xiong, K.-L. Guan and S. Zhao, *Mol. Cell*, 2011, **43**, 33-44.
256. A. Arora and C. S. Dey, *Biochem. Biophys. Res. Commun.*, 2016, **474**, 747-752.
257. A. Arora and C. S. Dey, *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 2014, **1842**, 1372-1378.
258. M. M. Maxwell, E. M. Tomkinson, J. Nobles, J. W. Wizeman, A. M. Amore, L. Quinti, V. Chopra, S. M. Hersch and A. G. Kazantsev, *Hum. Mol. Gen.*, 2011, **20**, 3986-3996.
259. R. Pandithage, R. Lilischkis, K. Harting, A. Wolf, B. Jedamzik, J. Lüscher-Firzlaff, J. Vervoorts, E. Lasonder, E. Kremmer, B. Knöll and B. Lüscher, *J. Cell Biol.*, 2008, **180**, 915.
260. J. Gal, Y. Bang and H. J. Choi, *Neurochem. Int.*, 2012, **61**, 992-1000.
261. T. F. Outeiro, E. Kontopoulos, S. M. Altmann, I. Kufareva, K. E. Strathearn, A. M. Amore, C. B. Volk, M. M. Maxwell, J.-C. Rochet, P. J. McLean, A. B. Young, R. Abagyan, M. B. Feany, B. T. Hyman and A. G. Kazantsev, *Science*, 2007, **317**, 516.
262. R. Luthi-Carter, D. M. Taylor, J. Pallos, E. Lambert, A. Amore, A. Parker, H. Moffitt, D. L. Smith, H. Runne, O. Gokce, A. Kuhn, Z. Xiang, M. M. Maxwell, S. A. Reeves, G. P. Bates, C. Neri, L. M. Thompson, J. L. Marsh and A. G. Kazantsev, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 7927.
263. V. Chopra, L. Quinti, J. Kim, L. Vollor, K. L. Narayanan, C. Edgerly, Patricia M. Cipicchio, Molly A. Lauver, Soo H. Choi, Richard B. Silverman, Robert J. Ferrante, S. Hersch and Aleksey G. Kazantsev, *Cell Rep.*, 2012, **2**, 1492-1497.
264. K. Zhao, X. Chai, A. Clements and R. Marmorstein, *Nat. Struct. Biol.*, 2003, **10**, 864.
265. K. G. Hoff and C. Wolberger, *Nat. Struct. Mol. Biol.*, 2005, **12**, 560.
266. A. A. Sauve, I. Celic, J. Avalos, H. Deng, J. D. Boeke and V. L. Schramm, *Biochem.*, 2001, **40**, 15456-15463.
267. A. A. Sauve, C. Wolberger, V. L. Schramm and J. D. Boeke, *Annu. Rev. Biochem.*, 2006, **75**, 435-465.
268. B. C. Smith and J. M. Denu, *Biochem.*, 2006, **45**, 272-282.
269. S. Lee, L. Tong and J. M. Denu, *Anal. Biochem.*, 2008, **383**, 174-179.

- 270. A. A. Sauve and D. Y. Youn, *Curr. Opin. Chem. Biol.*, 2012, **16**, 535-543.
- 271. K. G. Tanner, J. Landry, R. Sternglanz and J. M. Denu, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 14178.
- 272. J. Landry, J. T. Slama and R. Sternglanz, *Biochem. Biophys. Res. Commun.*, 2000, **278**, 685-690.
- 273. D. G. Fatkins, A. D. Monnot and W. Zheng, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 3651-3656.
- 274. W. F. Hawse, K. G. Hoff, D. G. Fatkins, A. Daines, O. V. Zubkova, V. L. Schramm, W. Zheng and C. Wolberger, *Structure*, 2008, **16**, 1368-1377.
- 275. B. C. Smith and J. M. Denu, *Biochem.*, 2007, **46**, 14478-14486.
- 276. J. E. Baldwin, *J. Chem. Soc. Chem. Commun.*, 1976, DOI: 10.1039/C39760000734, 734-736.
- 277. Y. Shi, Y. Zhou, S. Wang and Y. Zhang, *J. Phys. Chem. Lett.*, 2013, **4**, 491-495.
- 278. P. N. Dyer, R. S. Edayathumangalam, C. L. White, Y. Bao, S. Chakravarthy, U. M. Muthurajan and K. Luger, in *Methods Enzymol.*, Academic Press, 2003, vol. 375, pp. 23-44.
- 279. J. A. D'Anna and I. Isenberg, *Biochem.*, 1974, **13**, 4987-4992.
- 280. B. Liedberg, C. Nylander and I. Lunström, *Sens. Actuators*, 1983, **4**, 299-304.
- 281. P. Andrews, *Br. Med. Bull.*, 1966, **22**, 109-114.
- 282. P. J. Wyatt, *Instrum. Sci. Technol.*, 1997, **25**, 1-18.
- 283. L. R. Dick, A. D. Sherry, M. M. Newkirk and D. M. Gray, *Journal of Biological Chemistry*, 1988, **263**, 18864-18872.
- 284. M. M. Pierce, C. S. Raman and B. T. Nall, *Methods*, 1999, **19**, 213-221.
- 285. A. Wepf, T. Glatter, A. Schmidt, R. Aebersold and M. Gstaiger, *Nature Methods*, 2009, **6**, 203.
- 286. J. L. Feldman, K. E. Dittenhafer-Reed, N. Kudo, J. N. Thelen, A. Ito, M. Yoshida and J. M. Denu, *Biochemistry*, 2015, **54**, 3037-3050.
- 287. P. Knyphausen, S. de Boor, N. Kuhlmann, L. Scislowski, A. Extra, L. Baldus, M. Schacherl, U. Baumann, I. Neundorf and M. Lammers, *Journal of Biological Chemistry*, 2016, **291**, 14677-14694.

Chapter 3: A Greener Route to DBHDA



3.1.0.0 Introduction

2,5-Dibromohexanediamide (DBHDA) is an essential reagent for the chemical mutation of cysteine into dehydroalanine (Dha),²⁸⁸ which has been used extensively in this thesis.²⁸⁹ Incorporation of a Dha residue in a protein allows further reaction to form C-C,²⁹⁰ C-S,^{291, 292} and C-N²⁹³ bonds. DBHDA has also been used as a 'cysteine stapling' reagent²⁹⁴ and a reactive tag to ubiquitinate proteins *in vitro*.²⁹⁵

Despite being a widely used reagent, the synthesis of DBHDA uses the potent ozone depletor²⁹⁶ carbon tetrachloride and while DBHDA is available commercially it is prohibitively expensive (886 £ g⁻¹ at the time of writing).²⁹⁷

3.1.1.0 Past Syntheses

DBHDA was first synthesized in Oxford by W. H. Perkin Jr *et al.*²⁹⁸ Perkin's synthesis²⁹⁸ (*Figure 1*) initially oxidized cyclohexanol to adipic acid as described by Bouveault and Locquin.²⁹⁹ Adipic acid was treated with thionyl chloride to obtain adipyl dichloride,³⁰⁰ followed by bromination with molecular bromine as described by Ingold.³⁰⁰ The α,δ -dibromoadipyl dichloride was then added to methanol to obtain dimethyl α,δ -dibromoadipate, which had been described by Stephen and Weizmann.³⁰¹ Perkin found that shaking dimethyl α,δ -dibromoadipate with concentrated aqueous ammonia readily converts the ester groups to amides to give DBHDA.²⁹⁸

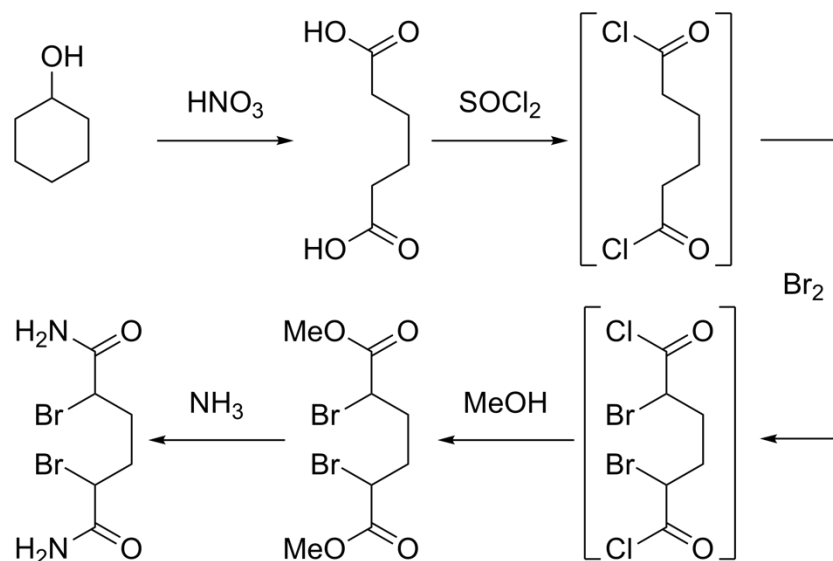


Figure 1 - W. H. Perkins Jr's Synthesis of 2,5-Dibromohexandiamide (DBHDA):²⁹⁸

Oxidative ring opening of cyclohexanol gives adipic acid, which is sequentially chlorinated with thionyl chloride then brominated with molecular bromine. 2,5-dibromoadipyl dichloride (DBADC) is then treated with methanol to give the dimethyl ester and finally transamidated with ammonium hydroxide to give DBHDA.

Davis *et al* took a similar approach to the synthesis²⁸⁸ of DBHDA as Perkin,²⁹⁸ (Figure 2) starting with the chlorination of adipic acid by thionyl chloride. Adipyl dichloride is then brominated using Harpp's method³⁰²⁻³⁰⁴ of acid chloride α -halogenation, which employs *N*-bromosuccinimide (NBS) in CCl_4 with a catalytic amount of hydrobromic acid. Davis²⁸⁸ then reacts the α,δ -dibromoadipyl dichloride directly with ammonium hydroxide to obtain DBHDA.

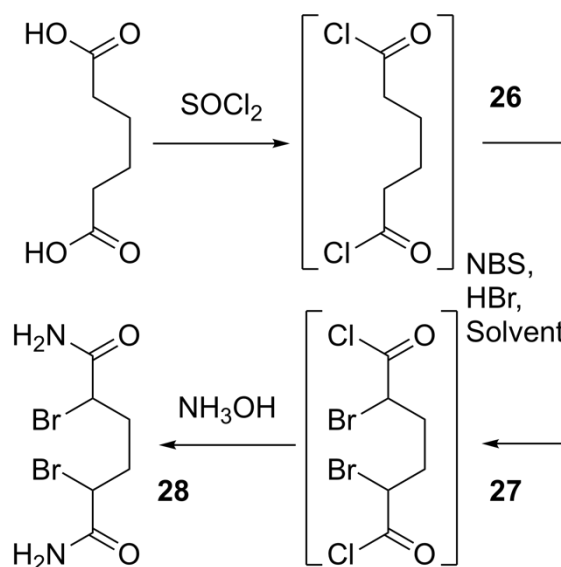


Figure 2 – B. G. Davis' Synthesis of 2,5-Dibromoadipic dihydrazide (DBHDA):²⁸⁸ Adipic acid is chlorinated with thionyl chloride, then brominated using Harpp's method of N-bromosuccinimide (NBS) and catalytic HBr in CCl₄ followed by forming the primary amides with ammonium hydroxide

Davis' method²⁸⁸ cut two steps from Perkin's synthesis: one due to the commercial availability of adipic acid and the second by directly reacting the DBADC with directly with ammonia. However, the synthesis still uses CCl₄ as a solvent in the key bromination step. NBS is a superior source of bromine due to low expense, accessibility and ease of handling when compared to molecular bromine.

3.1.2.0 CCl₄ Impacts

Carbon tetrachloride, a chemical banned under article 2D of the Montreal Protocol on Substances that Deplete the Ozone Layer,³⁰⁵ is currently being phased out of laboratory use and is banned in all industrial labs³⁰⁶ unless used as a feedstock. CCl₄ is also becoming increasingly hard to get hold of in academic labs.³⁰⁷ CCl₄ is of particular

notoriety due to its potent²⁹⁶ and long-lived³⁰⁸ ozone depleting effects when released into the atmosphere. CCl_4 is also a strong and persistent greenhouse gas with a global warming potential 1400 times greater than CO_2 .^{309, 310} Recent estimates put the lifetime of a single molecule of CCl_4 in the atmosphere at ~47 years,³¹¹ this is extended to hundreds of years when the chemical is present in the ground³¹² or water.³¹³ On top of the environmental effects CCl_4 is an extremely potent hepatotoxin.³¹⁴ Prolonged exposure to CCl_4 can lead to comas or even death.³¹⁵ I thought it best to avoid this solvent.

3.1.3.0 Aims

I set out to create a synthesis of DBHDA that is green, cost effective and reproducible. I aimed for this synthesis to be free of column chromatography. I wished that this could be completed by one researcher in one day from chemicals which are readily available in a laboratory.

3.2.0.0 Results and Discussion

3.2.1.0 Planning

The CCl_4 is chosen as a solvent for a variety of reasons:

- It is inert to the bromination conditions used: NBS/HBr and heat
- The 2,5-dibromoadipyl dichloride (DBADC) which results from the reaction is soluble in CCl_4 while NBS and succinimide employed and produced by the reaction are not soluble facilitating isolation

- The medium boiling point (77 °C) is sufficiently high that when at reflux the bromination reaction readily occurs.

A list³¹⁶ of solubilities showed NBS is insoluble in diethyl ether, hexanes and CCl₄. CCl₄ was obviously immediately ruled out, ether has a boiling point of only 35 °C. ⁿHexane has a boiling point of 69 °C while cyclohexane has a boiling point of 81 °C – only 4 °C different from that of CCl₄. On top of the boiling point gains, cyclohexane is both greener (with a Rowan university greenness score³¹⁷ of 3.2³¹⁸ as opposed to 5.4)³¹⁸ and less toxic (with a CHEM21³¹⁹ Health Score of 3³¹⁹ as opposed to 7)³¹⁹ than ⁿhexane.

I chose to try to adapt the current method of DBHDA synthesis by substituting CCl₄ for cyclohexane. Adipic acid seemed like a reasonably green starting point as numerous green syntheses to it have been described including: the oxidation of cyclohexene with hydrogen peroxide^{320, 321} and the degradation of lignin.^{322, 323}

Aside from carbon CCl₄ the only other identified problem with the current route is the inflammability of diethyl ether used in the extraction step, which I switched to methyl ^tbutyl ether (MTBE). The higher boiling point of MTBE reduces the risk of fire when compared to diethyl ether.

3.2.2.0 Developed Procedure

3.2.2.1 2,5-dibromohexanediamide 27

A one necked (B24, FR250/3S) 250 mL flask is equipped with a PTFE coated rugby ball magnetic stir bar (35 x 16 mm) and connected to a Davies condenser (160 mm, C5/12/SC) through a reduction adapter (B24 to B19) and left open to air. The reaction flask is sequentially charged with adipic acid (10.0 g, 68.4 mmol, 1.0 eq) and thionyl chloride (30 mL, 153 mmol, 2.2 eq. per acid group). The reaction mixture is heated to reflux (preheated oil bath, 90 °C) and boiled for 90 minutes. The colourless suspension should now appear as a colourless to pale yellow solution, this can be seen in *figure 5*.

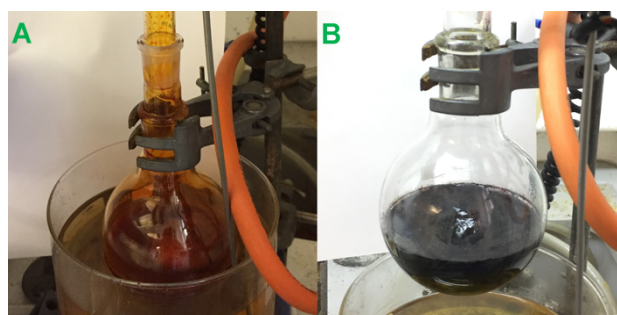


Figure 3 - Bromination of the bis(acid chloride) showing A) bromine vapour formed during the reaction and B) the colourless vapour over a black suspension characteristic of a complete reaction

The reaction mixture is allowed to cool for 10 minutes (from 83 °C to 47 °C). With stirring, cyclohexane (40 mL), *N*-bromosuccinimide (29.2 g, 164 mmol, 1.2 eq per residue) and hydrobromic acid (40 μ L, 9 M, 360 μ mol, 0.5 mol %) are sequentially added by temporarily removing the condenser. The reaction mixture turns

a pale orange upon addition of *N*-bromosuccinimide which deepens dramatically when the hydrobromic acid is added. The reaction mixture is heated to reflux (preheated oil bath, 90 °C) and boiled for 2.5 hours, during which time the deep orange solution and orange vapour will turn to a yellow suspension with colourless vapour and eventually to a black suspension with colourless vapour, this can be seen in *figure 3*. The suspension is then cooled in an ice-and-water bath for 30 minutes (bath temperature 0 °C from 76 °C to 1 °C, Note 11). The reaction mixture is filtered through a sintered glass Hirsch funnel (SF4B32) into a 250 ml round bottom flask (B19, FR250/2S) using a side arm adapter (MF/18/2/3C) using mild vacuum. The filter cake is washed with MTBE (3 x 20 mL). The black filtrate is concentrated by rotary evaporation (40 °C, 200 mbar then 100 mbar) to obtain the 2,5-dibromo bis(acid chloride) as a black, somewhat viscous oil.

The bis(acid chloride) is transferred to a pressure equalising dropping funnel (volume: ~100 mL, top joint: B24, bottom joint: B19, stopcock: NS12.5, 2.5 mm). Ice cold ammonium hydroxide (80 mL, 13.2 M, 1.06 mol, 7.7 eq per residue) is poured into a three necked round bottom flask 250 mL (B24, 2 x B14, diameter: ~250 mm) fitted with a thermometer in one side neck, the dropping funnel in the middle and the third neck remaining open (nitrogen can be blown through this neck in order to see the thermometer if ammonium chloride smoke forms) and equipped with a PTFE coated rugby ball magnetic stir bar (35 x 16 mm). The ammonia filled flask is cooled in an ice-water-sodium chloride bath (-5 °C). The black viscous oil is added dropwise with

rapid stirring, over 1 hour, maintaining a temperature below 15 °C. As soon as the oil is added the reaction mixture turns a vivid blue, which becomes more intense as more is added, which can be seen in *figure 4*. An off-white precipitate forms over time. The reaction mixture is stirred for 1 hour in an ice and water bath after the addition is complete. After 1 hour a green powder is suspended in a brown/black solution. The green solid is filtered through a sintered glass Buchner funnel (55 mm, B24, SF3A33), and the reaction vessel twice washed out with the filtrate.



Figure 4 – A) The initial blue colour formed when the bis(acid chloride) is dropped into the ice cold ammonium hydroxide and B) The green solid suspended in brown liquid that appears towards the end of the reaction.

The crude green powder (~28 g, damp) is transferred to a one-necked (B24, FrR150/3S) 150 mL flask with a PTFE coated rugby ball magnetic stir bar (35 x 16 mm, you can re-use the same one as in the prior reaction) and triturated in aqueous methanol (100 mL, 50 %, diluted from 99 %, Note 20) at 60 °C for 30 minutes, solids are hot filtered through a sintered glass Buchner funnel (55 mm, B24, SF3A33), washing out the reaction vessel once with the brown filtrate. The off-white powder is transferred to a one-necked (B24, FrR150/3S) 150 mL flask with a PTFE coated rugby

ball magnetic stir bar (35 x 16 mm, you can use the same one as in the prior reaction) and triturated in aqueous methanol (100 mL, 50 %, diluted from 99 %) at 60 °C for 30 minutes, solids are hot filtered through a sintered glass Buchner funnel (55 mm, B24, SF3A33), washing out the reaction vessel once with the yellow filtrate. The filter cake is dried in a calcium chloride containing vacuum desiccator overnight (230 mm, polycarbonate, flat-dome top, 850 μ bar, Notes 21 and 22). The off-white powder is crushed with a spatula, and then dried for 3 hours in a vacuum desiccator to yield the title product as a microcrystalline off-white powder (6.44 g, 24.1 mmol, 31 %).

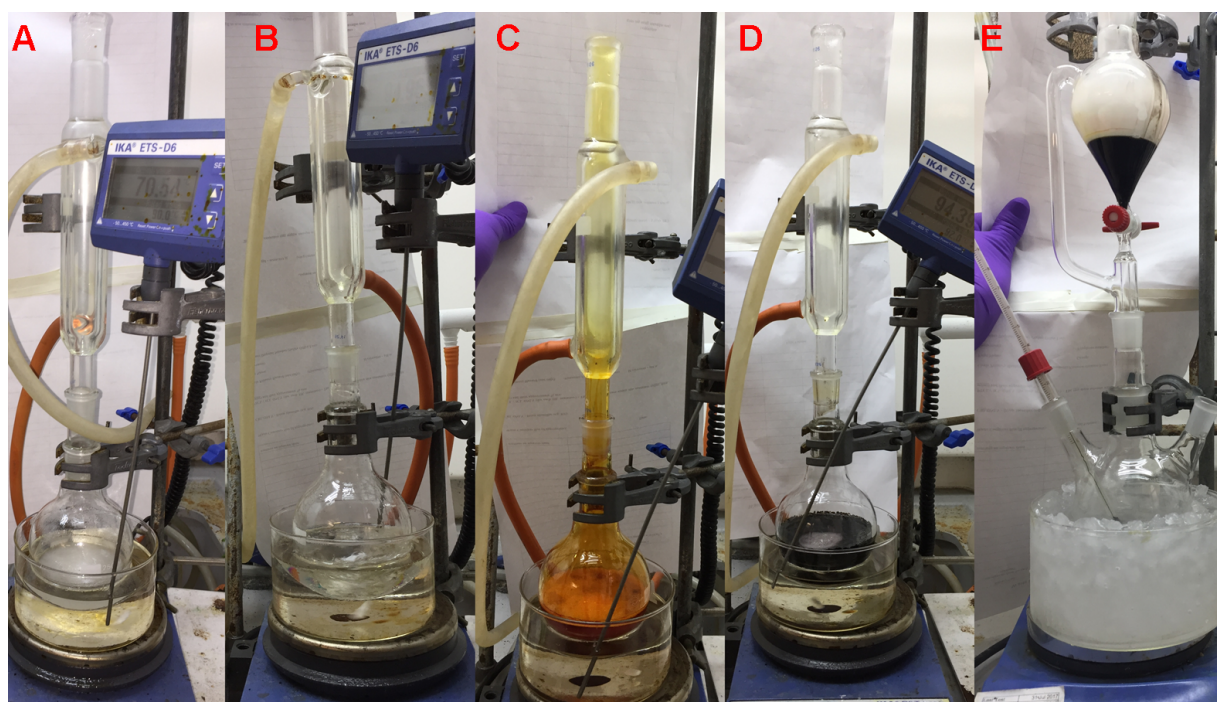


Figure 5 - Set ups at each stage of reaction A) adipic acid suspended in thionyl chloride and B) upon completion of the reaction C) The beginning of the bromination reaction with a deep red suspension and orange vapor and D) when the bromination reaction is complete. E) The reaction set up for the addition of the black DBADC to ammonium hydroxide

3.2.3.0 Discussion

3.2.3.1 Reaction Monitoring

It is difficult to monitor reactions where the products are inherently unstable, such as: the chlorination of a carboxylic acid group. One way of assessing the reaction progress is to inspect the reaction mixture. Thionyl chloride chlorination releases one equivalent of both SO_2 and HCl in every reaction, therefore when proceeding the reaction is effervescent. While the reaction is heated to reflux which means there is an underlying effervescence, the difference is noticeable. Further to this effervescence the reaction is endothermic, so careful monitoring of the reaction temperature can give hints as to when the reaction is complete, more details are given **section 2.3.4**. While the diacid is insoluble in thionyl chloride the reaction media, the product is a liquid, so monitoring for the complete disappearance of solid can also be used.

Unsatisfied with these indirect approaches, I developed a protocol for thin layer chromatography of the reaction media, by first placing a sample of the reaction mixture into neat isopropanol, which yields the diisopropyl ester of adipic acid. While the reaction of the diacid and an alcohol is accompanied by an exotherm I didn't find the exotherm strong enough to be considered a risk, provided the sample quenching is done on the scale I suggest. A photograph of the thin layer chromatogram can be seen in *figure 6*.

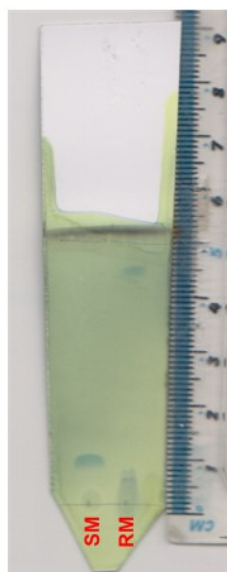


Figure 6 – Thin Layer Chromatography (TLC) Analysis (Silica/Aluminium plate, $\text{CHCl}_3/\text{MeOH}/\text{HCOOH}$ 100:1:1, phosphomolybdic acid stain) of the chlorination step. Reaction mixture (RM) quenched with isopropanol prior to analysis, starting material (SM) added as a suspension in isopropanol.

As with the first step, the product formed by the α -bromination is unstable. Running with the previously used technique I also use the inspection of the reaction mixture or a derivatisation followed by TLC to determine the reaction progress.

When NBS/HBr are added to the reaction mixture the colour changes from colourless to orange to deep red with orange vapour. With the consumption of Br_2 during the reaction both the vapour and reaction mixture lighten in colour, when the reaction is complete the vapour above the reaction mixture is colourless and the reaction mixture a pale yellow, which if left will rapidly turn black. The black colour does not correspond to product decomposition, so this is not a problem.

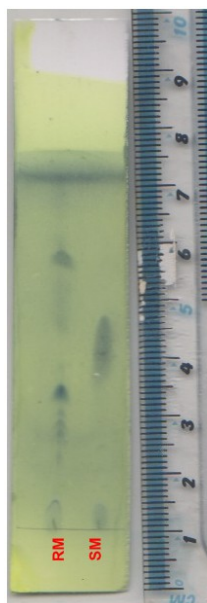


Figure 7 - Thin Layer Chromatography (TLC) Analysis (Silica/Aluminium plate, $\text{CHCl}_3/\text{HCOOH}$ 100:1, phosphomolybdic acid stain) of the bromination step Reaction mixture (RM) quenched with isopropanol prior to analysis, starting material (SM) retained from prior step.

By retaining the isopropanol quenched reaction sample from the first step, I can compare thin layer chromatograms of an isopropanol quenched sample in the second step. As a small amount of mono-brominated 2,5-adipyl dichloride is formed there are 2 major spots and a variety of dilute species. However, complete consumption of the starting material can be observed. A photograph of the thin layer chromatogram can be seen in *Figure 7*.

Monitoring the reaction in the final step is extremely difficult as the product precipitates to form a clumpy solid, therefore taking a representative sample of the reaction mixture is near impossible. Despite not being able to retrieve a sample the reaction between acid chlorides and concentrated ammonium hydroxide is extremely

rapid and I allow a long time for the reaction to complete. If one was concerned about losing product, then retention of the filtrate from the product isolating filtration followed by isolation and analysis of any further precipitate would reassure the chemist.

While my TLC analyses use chloroform as a running solvent I do not consider this a problem: I have provided ample other methods to monitor the reaction progress and if the person wishing to repeat the synthesis adhered to the times outlined in my procedure there should be no need to use TLC analysis.

3.2.3.2 NBS Quality

I found the yield of this reaction to be reduced by ~30 % when impure (orange) NBS was used. Therefore, all *N*-bromosuccinimide was recrystallised using the following protocol: In a 2 L conical flask open to air, water (600 mL) is heated to 90 °C on a hotplate with rapid stirring provided by a rugby ball magnetic stir bar. *N*-bromosuccinimide (60.0 g) is added and stirred at the same temperature until homogenous. The stir bar is removed from the bright orange solution which is subsequently cooled in a refrigerator (4 °C) for at least 3 h but ideally overnight. The colourless plates are isolated by filtration (sintered glass Buchner funnel, 106 mm, SF3B33) then dried in a phosphorus pentoxide containing vacuum desiccator overnight to yield the title product as colourless plates (53.4 g).

3.2.3.3 NBS Quantity

I found that increasing the ratio of NBS to adipyl dichloride lead to a greater selectivity for the dibrominated product over the monobrominated by analysing crude samples of DBHDA by quantitative ^{13}C nuclear magnetic resonance spectroscopy (q ^{13}C NMR), which is depicted graphically in *figure 8*.

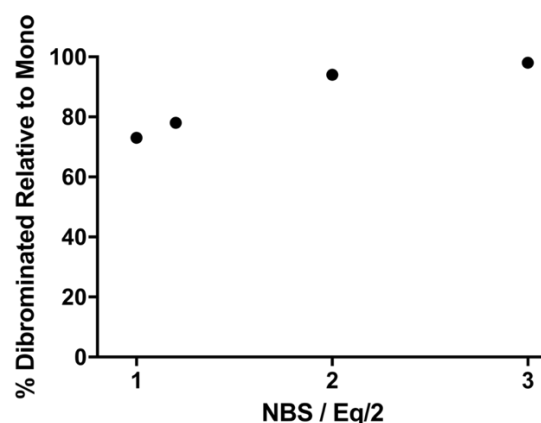


Figure 8 - Percentage of dibrominated product relative to monochlorinated as a function of the equivalence of N-bromosuccinimide (NBS) per reactive residue measured by quantitative ^{13}C nuclear magnetic resonance spectroscopy (q ^{13}C NMR) spectroscopy of crude samples of dibromohexadamide (DBHDA).

Despite the increase in conversion to the desired product I found the isolated yields were considerably lower. I speculate that this is due to the difficulty in isolating 2,5-dibromoadipyl chloride from the succinimide side product of the bromination reaction. I found 1.2 equivalents per acid group to be a good compromise giving an isolated yield of 31 % over three steps.

3.2.3.4 Reaction Temperatures

Temperature profiles of the heated reaction were recorded in order to monitor if the recommended heating and cooling steps were efficient. During the endothermic chlorination I noticed that after the initial heating stage the reaction mixture was cooling, until a local minimum was achieved at ~45 minutes where the temperature started to rise. I assume the reaction is complete when the reaction mixture plateaus at ~80 °C. The reaction temperature profile can be seen in *Figure 9*. I monitored the bromination step but nothing interesting was noted (*Figure 10*). Monitoring the temperature of the amide forming step would not be a useful experiment as the temperature spikes with addition of acid chloride then I wait for it to cool before adding more reagent.

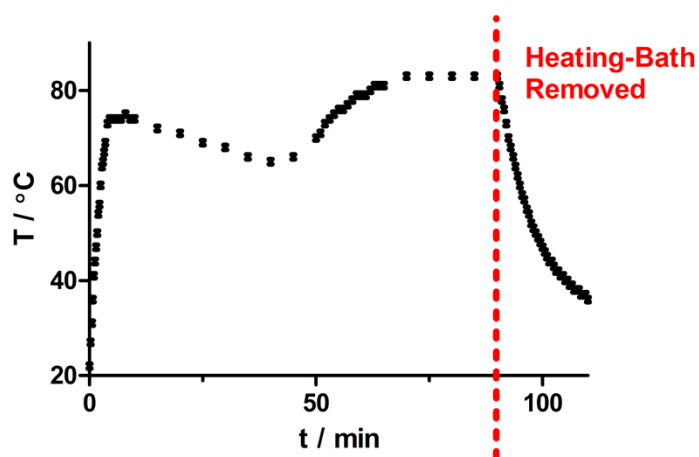


Figure 9 - Temperature Profile of Chlorination Reaction: Adipic acid (10.0 g) is mixed with thionyl chloride and immediately immersed into a 90 °C oil bath, for 90 minutes. The temperature is monitored by reading a thermometer immersed in the reaction mixture, time points were relaxed while the temperature seemed stable. I

speculate that the endothermic reaction is complete when the temperature stabilises at ~80 °C

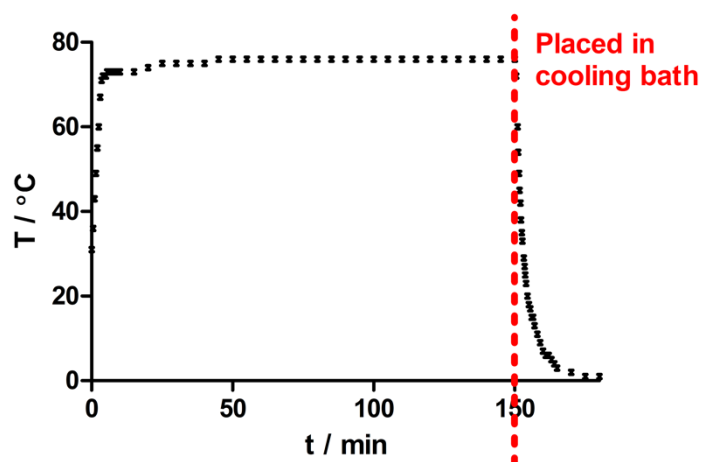


Figure 10 - Temperature profile of the bromination reaction: *N*-bromosuccinimide (29.2 g, 164 mmol, 1.2 eq per residue) and hydrobromic acid (40 μ L, 9 M, 360 μ mol, 0.5 mol %) are added to the reaction mixture from Figure 9 immersed into a 90 °C oil bath, for 90 minutes. The temperature is monitored by reading a thermometer immersed in the reaction mixture, time points were relaxed while the temperature seemed stable.

To ensure complete submersion of the thermometer in the reaction mixture, as well as having an extra neck on the reaction vessel to allow the fitting of a thermometer and adapter slightly different set ups were used from those described in 3.2.2.1 a description of these can be found in 3.4.3.0.

3.2.3.5 Ammonia Vapour

It seems counter intuitive to conduct a reaction in a three-necked flask when only two of the necks are occupied however the use of a two necked flask leads to significant build-up of ammonia vapour which reacts with the droplet of neat 2,5-dibromoadipyl

dichloride to form solid DBHDA. The result of this reaction is the formation of impressive 'stalactites' of DBHDA and acid chloride hanging from the dropping tip above the reaction mixture, a photograph of the stalactite can be seen in *figure 11*.

While the stalactite seems harmless, the reaction between an acid chloride and ammonia is extremely exothermic. The eventual collapse of the stalactite can lead to a dramatic rise in temperature, which can cause the reaction mixture to start boiling.



Figure 11 - Stalactite of acid chloride which forms when ammonia vapour is allowed to build up, its eventual collapse leads to a large spike in temperature.

I found that the use of a three necked flask provided sufficient ventilation to prevent the stalactite from forming. As gas phase water is not particularly reactive towards acid chlorides due to its low concentration and low reactivity compared to ammonia I was not concerned with finding other methods. I speculate that performing the reaction under a blanket of a dense gas such as argon could lead to more problems: ammonia vapour could be forced through the side neck of the pressure equalising dropping funnel

into the reservoir. Continuous flow of nitrogen through the reaction vessel is an option, however it seems unnecessary.

3.2.3.6 Product Purity

Analysis of crude DBHDA precipitated by TLC showed two spots on the chromatogram: one corresponding to DBHDA and the other to 2-bromohexamide. Trituration in 50 % aqueous methanol followed by hot filtration significantly reduced the intensity of this spot on the thin layer chromatogram. The green colour of the crude DBHDA was mostly removed after the initial trituration however, the off-white powder was still contaminated with 2-bromohexanediamide. A second trituration following the same procedure removed the contaminant from the TLC.

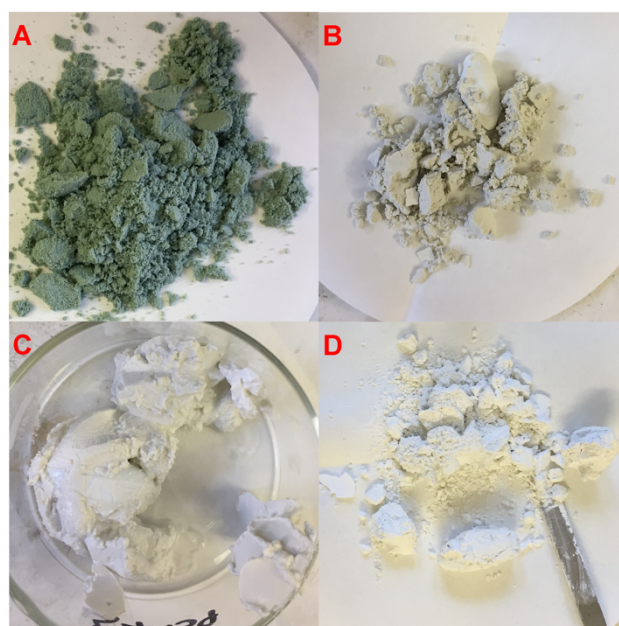


Figure 12 – Comparison of 2,5-dibromohexanediamide powder (DBHDA) A) DBHDA before trituration B) After the first trituration in 1:1 $\text{H}_2\text{O}/\text{MeOH}$ C) After the second trituration trituration in 1:1 $\text{H}_2\text{O}/\text{MeOH}$ and D) After drying the amorphous powder *in vacuo*

Analysis by TLC is not quantitative, traditionally the quantification is provided by quantitative proton nuclear magnetic resonance (qNMR), HPLC or gas chromatography (GC). DBHDA is not volatile enough to perform GC and its solubility in most common HPLC solvents is low. qNMR would be difficult as all proton signals are present as multiplets. While ^{13}C is NMR active ($\text{Spin} = -\frac{1}{2}$) it is not a preferred nucleus for quantitative NMR spectroscopy as it is of low natural abundance (1.11 %) the relaxation time is long, however with a long delay time this experiment is reasonable. The purity was found to be $99.7\% \pm 1.21$ this was determined by three repeats of q ^{13}C NMR spectroscopy against an internal standard of ethylene carbonate (72.76, 73.43, 77.73 mg of compound, 19.07, 20.14, 20.35 mg of internal standard respectively) in d_6 -DMSO (500 μL), using a 60 s delay time. The *syn*- and *anti*-products are roughly equal in abundance.

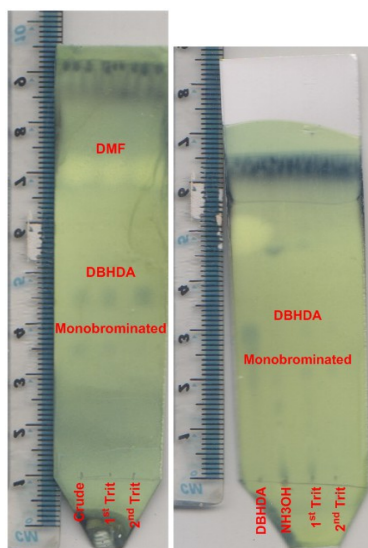


Figure 13 - Thin Layer Chromatography (TLC) Analysis (Silica/Aluminium plate, $\text{CHCl}_3/\text{HCOOH}$ 100:1, phosphomolybdic acid stain) of the of 2,5-dibromohexanediamide (DBHDA) purification showing powder (spotted as a *N,N*-dimethylformamide solution) (left) and mother-liquors (right)

3.3.0.0 Conclusion

I have developed a method of producing DBHDA without the requirement of any chlorinated solvents in order to produce a large amount of high-quality product in a cost effective ($< 5 \text{ £ g}^{-1}$) and time efficient manner (1 day of work for 1 worker) with no column chromatography. I have demonstrated that the method of production can be scaled up to yield $\sim 25 \text{ g}$. I have found that the *syn* and *meso* isomers are present in roughly equal proportions.

3.4.0.0 Experimental

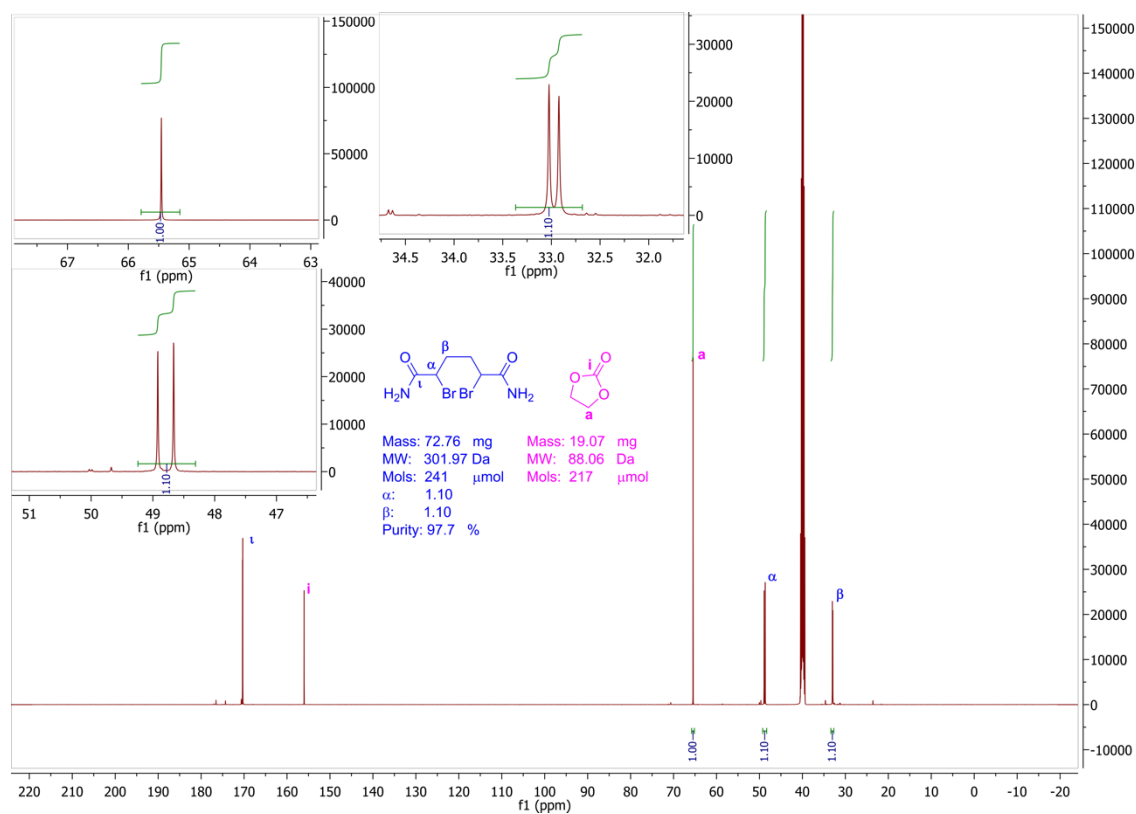
3.4.1.0 TLC Procedures

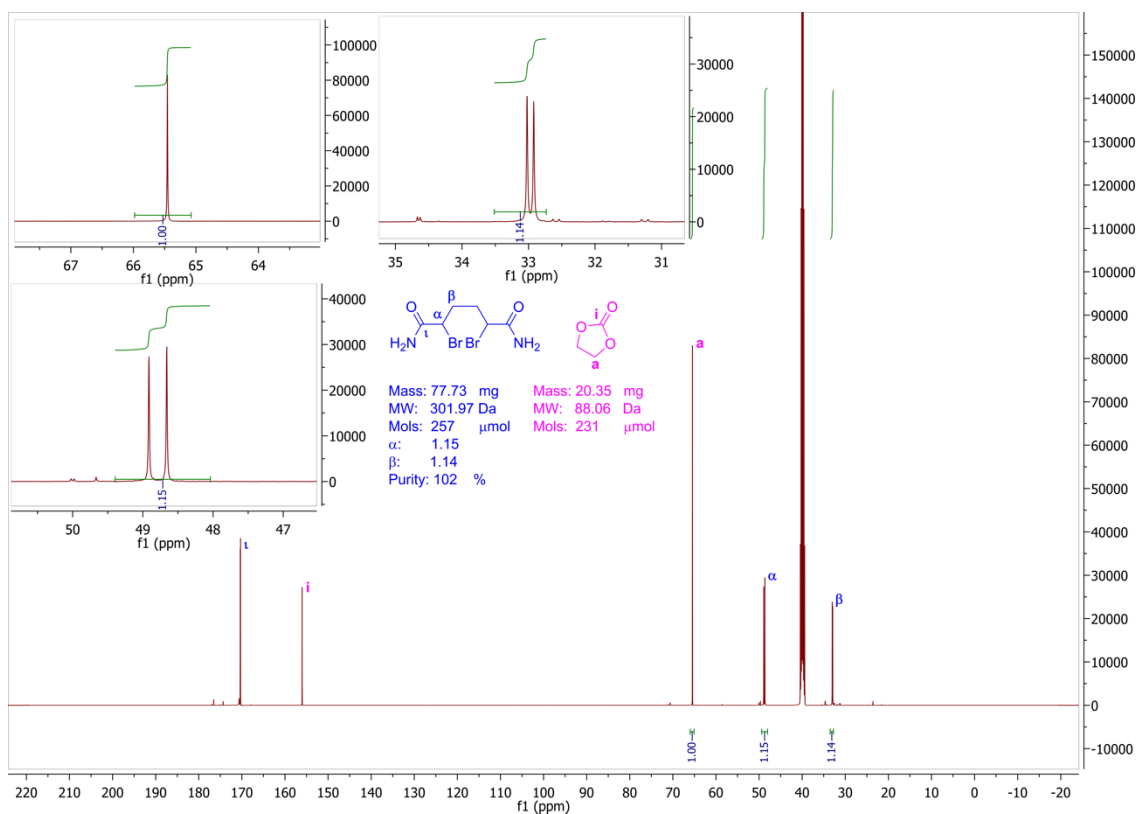
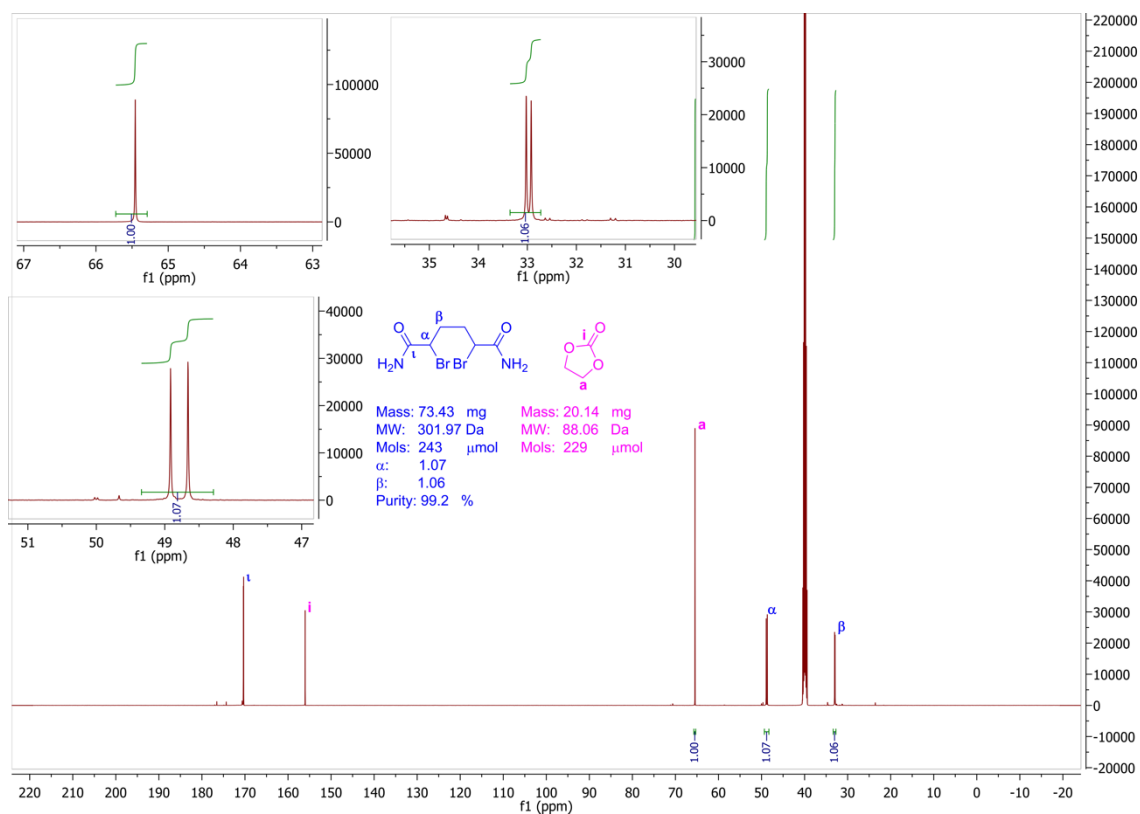
3.4.1.1 Chlorination

50 μ L of reaction mixture is mixed with 500 μ L of isopropanol then stood for 2 minutes, the diluted reaction mixture is spotted on an silica/aluminum TLC plate and run against a standard of adipic acid in isopropanol in a mobile phase made by mixing chloroform (10 mL) with formic acid (100 μ L) and methanol (100 μ L). The plate is dried then visualised using phosphomolybdic acid stain, prepared by dissolving phosphomolybdic acid (10.0 g) in ethanol (10 mL). Adipic acid gives an R_f of 0.18 while the derivatised diacid chloride gives an R_f of 0.88, both spots appear vivid blue on yellow background, impurities in the reaction mixture appear as a brown-blue smear from the baseline to $\sim R_f$ 0.1.

3.4.1.2 Bromination

100 μ L of reaction mixture is mixed with 500 μ L of isopropanol then stood for 2 minutes, the diluted reaction mixture is spotted on an silica/aluminum TLC plate and run against The standard from Note 4 in a mobile phase made by mixing chloroform (10 mL) with formic acid (100 μ L). The plate is dried then visualised using phosphomolybdic acid stain, prepared by dissolving phosphomolybdic acid (10.0 g) in ethanol (10 mL).

3.4.2.0 Quantitative ^{13}C NMR Spectra



3.4.3.0 Notes on Temperature Monitoring

3.4.3.1. Chlorination

The temperature profile of this reaction was recorded in a 100 mL 2 necked round bottom flask (B24, B14, diameter: ~210 mm) fitted with a Davies condenser (160 mm, C5/12/SC) through a reduction adapter (B24 to B19) and a thermometer through a screw cap adapter, this left open to air. A smaller flask was used to allow efficient stirring of the reaction mixture while a thermometer was immersed in the reaction mixture.

3.4.3.2 Bromination

The temperature profile of this reaction was recorded in a 250 mL 2 necked round bottom flask (B24, B14, diameter: ~210 mm) fitted with a Davies condenser (160 mm, C5/12/SC) through a reduction adapter (B24 to B19) and a thermometer through a screw cap adapter, this left open to air. A smaller flask was used to allow efficient stirring of the reaction mixture while a thermometer was immersed in the reaction mixture.

3.4.4.0 DBHDA Characterisation data

Mp: 204-205 °C Decomposition

R_f (EtOAc, 254 nm): 0.20

^1H NMR (400 MHz, $\text{d}_6\text{-DMSO}$) δH 7.70 (s, 2H, NH_2), 7.32 (s, 2H, NH_2), 4.33 (td, J = 5.2, 2.8 Hz, 2H, CBrH) and 2.20 – 1.66 (m, 4H, CH_2)

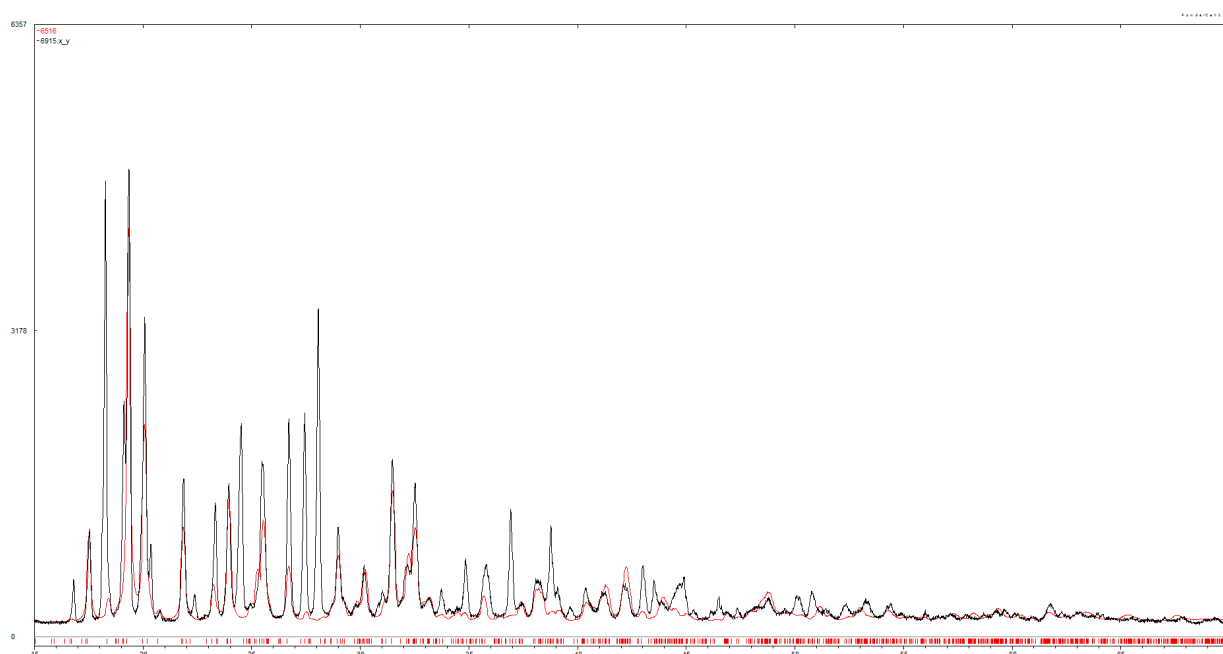
^{13}C NMR (101 MHz, $\text{d}_6\text{-DMSO}$) δC 169.9 (CO), 169.8 (CO), 48.5 (CBr), 48.2(CBr), 32.6 (CH_2) and 32.5 (CH_2)

ATR-FTIR (Neat) 3434 (NH Stretch), 3303 (NH Stretch), 1669 (CO Stretch), 1416 (CH Bend), 1319, 1277 (CH wag), 1251, 1220, 1195, 1158, 979.0, 880.4, 814.0, 770.4. 652.3 (CBr Stretch) and 614.4 (CBr Stretch)

MS (ESI+) 301 ($\text{M}^{79}\text{Br}^{79}\text{Br} + \text{H}^+$) 303 ($\text{M}^{79}\text{Br}^{81}\text{Br} + \text{H}^+$) 305 ($\text{M}^{81}\text{Br}^{81}\text{Br} + \text{H}^+$)

HRMS (ESI+) ($\text{M}^{79}\text{Br}^{79}\text{Br} + \text{H}^+$) *Expected*: 300.9187 (51.2%) *Found*: 300.9195 (50.4 %) ($\text{M}^{79}\text{Br}^{81}\text{Br} + \text{H}^+$) *Expected*: 302.9167 (100%) *Found*: 302.9177 (100%) ($\text{M}^{81}\text{Br}^{81}\text{Br} + \text{H}^+$) *Expected*: 304.9147 (49.1 %) *Found*: 304.9157 (47.3%)

Power Diffraction Spectrum



3.5.0.0 References

288. J. M. Chalker, S. B. Gunnoo, O. Boutureira, S. C. Gerstberger, M. Fernandez-Gonzalez, G. J. L. Bernardes, L. Griffin, H. Hailu, C. J. Schofield and B. G. Davis, *Chem. Sci.*, 2011, **2**, 1666-1676.
289. B. J. Bower, DPhil, The University of Oxford, 2018.
290. T. H. Wright, B. J. Bower, J. M. Chalker, G. J. L. Bernardes, R. Wiewiora, W.-L. Ng, R. Raj, S. Faulkner, M. R. J. Vallée, A. Phantumrithiwath, O. D. Coleman, M.-L. Thézénas, M. Khan, S. R. G. Galan, L. Lercher, M. W. Schombs, S. Gerstberger, M. E. Palm-Espling, A. J. Baldwin, B. M. Kessler, T. D. W. Claridge, S. Mohammed and B. G. Davis, *Science*, 2016, DOI: 10.1126/science.aag1465.
291. K. P. Chooi, S. R. G. Galan, R. Raj, J. McCullagh, S. Mohammed, L. H. Jones and B. G. Davis, *J. Am. Chem. Soc.*, 2014, **136**, 1698-1701.
292. J. M. Chalker, L. Lercher, N. R. Rose, C. J. Schofield and B. G. Davis, *Angew. Chem. Int. Ed.*, 2012, **51**, 1835-1839.
293. J. Dadová, K.-J. Wu, P. G. Isenegger, J. C. Errey, G. J. L. Bernardes, J. M. Chalker, L. Raich, C. Rovira and B. G. Davis, *ACS Cent. Sci.*, 2017, **3**, 1168-1173.
294. P. M. Morrison, P. J. Foley, S. L. Warriner and M. E. Webb, *ChemComm*, 2015, **51**, 13470-13473.
295. R. Meledin, S. M. Mali, S. K. Singh and A. Brik, *Org. Biomol. Chem.*, 2016, **14**, 4817-4823.
296. M. J. Prather and R. T. Watson, *Nature*, 1990, **344**, 729.
297. S. A. UK, DBHDA page, <https://www.sigmaaldrich.com/catalog/product/aldrich/900607?lang=en®ion=GB> (accessed Feb 21, 2018, 2018).
298. A. W. Bernton, H. R. Ing and W. H. Perkin, *J. Chem. Soc. Trans.*, 1924, **125**, 1492-1502.
299. L. Bouveault and R. Locquin, *Bull. Soc. Chim.*, 1908, 432.
300. C. K. Ingold, *J. Chem. Soc. Trans.*, 1921, **119**, 951-970.
301. H. Stephen and C. Weizmann, *J. Chem. Soc. Trans.*, 1913, **103**, 269-276.
302. D. N. Harpp, L. Q. Bao, C. Coyle, J. G. Gleason and S. Horovitch, *Org. Synth.*, 1976, **55**, 27-32.
303. D. N. Harpp, L. Q. Bao, C. J. Black, J. G. Gleason and R. A. Smith, *J. Org. Chem.*, 1975, **40**, 3420-3427.
304. D. N. Harpp, L. Q. Bao, C. J. Black, R. A. Smith and J. G. Gleason, *Tetrahedron Lett.*, 1974, **15**, 3235-3238.
305. U. N. E. Programme, *Journal*, 2016.

306. R. K. Henderson, C. Jiménez-González, D. J. C. Constable, S. R. Alston, G. G. A. Inglis, G. Fisher, J. Sherwood, S. P. Binks and A. D. Curzons, *Green Chem.*, 2011, **13**, 854-862.
307. Q. Liang, P. A. Newman, J. S. Daniel, S. Reimann, B. D. Hall, G. Dutton and L. J. M. Kuijpers, *Geophys. Res. Lett.*, 2014, **41**, 5307-5315.
308. W. M. O. G. O. R. a. M. Project, *Scientific Assessment of Ozone Depletion: 2010* World Meteorological Organization, Geneva, 2010.
309. W. F. J. Evans and E. Puckrin, *Geophys. Res. Lett.*, 1996, **23**, 1769-1772.
310. G. D. S. Myhre, G.; Bréon, F. - M; Collins, W.; Fuglestad, J.; Huang J.; Koch, D.; Lamarque, J. - F.; Lee, D.; Mendoza, B.; Nakajima, T.; Robock, A.; Stephens, G.; Takemura T.; Zhang, H.; , 2013: *Anthropogenic and Natural Radiative Forcing. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA., 2013.
311. M. Valeri, F. Barbara, C. Boone, S. Ceccherini, M. Gai, G. Maucher, P. Raspollini, M. Ridolfi, L. Sgheri, G. Wetzels and N. Zoppetti, *Atmos. Chem. Phys.*, 2017, **17**, 10143-10162.
312. R. C. Rhew and J. D. Happell, *Geophys. Res. Lett.*, 2016, **43**, 2889-2895.
313. J. H. Butler, S. A. Yvon-Lewis, J. M. Lobert, D. B. King, S. A. Montzka, J. L. Bullister, V. Koropalov, J. W. Elkins, B. D. Hall, L. Hu and Y. Liu, *Atmos. Chem. Phys.*, 2016, **16**, 10899-10910.
314. W. F. Seifert, A. Bosma, A. Brouwer, H. F. J. Hendriks, P. J. M. Roholl, R. E. W. van Leeuwen, G. C. F. van Thiel-De Ruiter, I. Seifert-Bock and D. L. Knook, *Hepatology*, 1994, **19**, 193-201.
315. R. O. Recknagel, E. A. Glende, J. A. Dolak and R. L. Waller, *Pharmacol. Ther.*, 1989, **43**, 139-154.
316. S. C. J. Virgil, P. R.; Wilson, A. J.; Garcia Romero, M. D., in *Encyclopedia of Reagents for Organic Synthesis*, Wiley, 2006, DOI: doi:10.1002/047084289X.rb318.pub2.
317. C. S. Slater and M. Savelski, *J. Environ. Sci. Health A*, 2007, **42**, 1595-1605.
318. F. P. Byrne, S. Jin, G. Paggiola, T. H. M. Petchey, J. H. Clark, T. J. Farmer, A. J. Hunt, C. Robert McElroy and J. Sherwood, *Sustain. Chem. Process.*, 2016, **4**, 7.
319. D. Prat, A. Wells, J. Hayler, H. Sneddon, C. R. McElroy, S. Abou-Shehadeh and P. J. Dunn, *Green Chem.*, 2016, **18**, 288-296.
320. K. Sato, M. Aoki and R. Noyori, *Science*, 1998, **281**, 1646.
321. S. M. Reed and J. E. Hutchison, *J. Chem. Ed.*, 2000, **77**, 1627.

322. D. R. Vardon, M. A. Franden, C. W. Johnson, E. M. Karp, M. T. Guarnieri, J. G. Linger, M. J. Salm, T. J. Strathmann and G. T. Beckham, *Energy Environ. Sci.*, 2015, **8**, 617-628.
323. A. Corona, M. J. Bidy, D. R. Vardon, M. Birkved, M. Z. Hauschild and G. T. Beckham, *Green Chem.*, 2018, **20**, 3857-3866.