Copyright Statement

The digital copy of this thesis is protected by the Copyright Act 1994 (New Zealand).

This thesis may be consulted by you, provided you comply with the provisions of the Act and the following conditions of use:

- Any use you make of these documents or images must be for research or private study purposes only, and you may not make them available to any other person.
- Authors control the copyright of their thesis. You will recognize the author's right to be identified as the author of this thesis, and due acknowledgement will be made to the author where appropriate.
- You will obtain the author's permission before publishing any material from their thesis.

General copyright and disclaimer

In addition to the above conditions, authors give their consent for the digital copy of their work to be used subject to the conditions specified on the Library Thesis Consent Form and Deposit Licence.
An Investigation into the Key Structural Role of Glycosylated Lysine Residues in the Collagenous Domain of Adiponectin, a Potential Therapeutic for Type II Diabetes

Katherine Rachel Lutteroth

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy,
School of Biological Sciences, University of Auckland, February 2017
This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

The introduction for "Chapter Two: Towards the synthesis of native glycosylated hydroxylysine" is largely extracted from the publication "Synthesis of glycosylated 5-hydroxylysine, an important amino acid present in collagen-like proteins such as adiponectin" which was published in the journal Organic and Biomolecular Chemistry in 2012 with the DOI 10.1039/c1ob06994d

<table>
<thead>
<tr>
<th>Nature of contribution by PhD candidate</th>
<th>First and primary author of the publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extent of contribution by PhD candidate (%)</td>
<td>90</td>
</tr>
</tbody>
</table>

### CO-AUTHORS

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Margaret A. Brimble</td>
<td>Principal editor, corresponding author and main supervisor of the student</td>
</tr>
<tr>
<td>Garth J.S. Cooper</td>
<td>Editor and supervisor of the student</td>
</tr>
<tr>
<td>Geoffrey M. Williams</td>
<td>Editor and advisor of the student</td>
</tr>
</tbody>
</table>

### Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate’s contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Margaret A. Brimble</td>
<td>[Signature]</td>
<td>27/10/16</td>
</tr>
<tr>
<td>Garth J.S. Cooper</td>
<td>[Signature]</td>
<td>27/10/16</td>
</tr>
<tr>
<td>Geoffrey M. Williams</td>
<td>[Signature]</td>
<td>27/04/16</td>
</tr>
</tbody>
</table>
Co-Authorship Form

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. Please include one copy of this form for each co-authored work. Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

The thesis "Chapter Four: The use of collagen model peptides as structural scaffolds for Adipn" has been submitted for publication under the title "An optimal level of glycosylation for collagen model peptides and implications for complications of Type II diabetes" to the Journal of the American Chemical Society in May 2016.

### Nature of contribution by PhD candidate
- First and primary author

### Extent of contribution by PhD candidate (%)
- 80

### CO-AUTHORS

<table>
<thead>
<tr>
<th>Name</th>
<th>Nature of Contribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Margaret A. Brimble</td>
<td>Principal editor, corresponding author and main supervisor of the student</td>
</tr>
<tr>
<td>Garth J.S. Cooper</td>
<td>Editor and supervisor of the student</td>
</tr>
<tr>
<td>Paul W. R. Harris</td>
<td>Editor and advisor to the student</td>
</tr>
<tr>
<td>Harveen Kaur</td>
<td>Editing assistance</td>
</tr>
<tr>
<td>Kevin J. Sparrow</td>
<td>Technical support with peptide synthesis</td>
</tr>
<tr>
<td>Jordi Bella</td>
<td>Structural biology guidance and editing of the manuscript</td>
</tr>
</tbody>
</table>

### Certification by Co-Authors

The undersigned hereby certify that:
- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

<table>
<thead>
<tr>
<th>Name</th>
<th>Signature</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Margaret A. Brimble</td>
<td><img src="signature1" alt="Signature" /></td>
<td>27/4/16</td>
</tr>
<tr>
<td>Garth J.S. Cooper</td>
<td><img src="signature2" alt="Signature" /></td>
<td>27/4/16</td>
</tr>
<tr>
<td>Paul W. R. Harris</td>
<td><img src="signature3" alt="Signature" /></td>
<td>27/4/16</td>
</tr>
<tr>
<td>Harveen Kaur</td>
<td><img src="signature4" alt="Signature" /></td>
<td>27/4/16</td>
</tr>
<tr>
<td>Kevin J. Sparrow</td>
<td><img src="signature5" alt="Signature" /></td>
<td>06/05/2015</td>
</tr>
<tr>
<td>Jordi Bella</td>
<td><img src="signature6" alt="Signature" /></td>
<td>08/05/2015</td>
</tr>
</tbody>
</table>

Last updated: 19 October 2015
Abstract

This thesis explores the role of glycosylated (Gal-Hyl) hydroxylysine residues for the structural stability of the collagenous domain of adiponectin (Adpn), a potential therapeutic for Type II diabetes. These Gal-Hyl residues have been shown to play a key role in the formation of bioactive high molecular weight oligomers (HMW) of Adpn. Notably, Type II diabetics have been shown to have lower levels of HMW of Adpn compared to healthy individuals.

However, due to the problems associated with recombinant synthesis of mammalian Adpn, an in-depth study of the structural importance of the Gal-Hyl residues has not been possible. Furthermore, although the recent total chemical synthesis of Adpn incorporating Gal-Hyl residues has been reported, the published synthesis was long, complex and ultimately low-yielding.

This thesis utilises neo-glycoside mimetics of Gal-Hyl to generate neo-glycopeptide mimetics of the collagenous domain of Adpn through the use of Fmoc-solid phase peptide synthesis. The structural properties and thermal stabilities of the resultant neo-glycopeptides were then studied using circular dichroism techniques.

This thesis describes the syntheses of three neo-glycopeptide families, for which the serine analogues of the lysine neo-glycopeptides were also prepared in order to investigate the effect of the glycan-peptide backbone through-space distance on the secondary structure of the neo-glycopeptide.

Chapter Three outlines the synthesis of short 18-mers corresponding to residues 66-83 of the collagenous domain of Adpn. Chapter Four describes the synthesis of collagen model peptides (CMPs) which were employed to promote triple helical formation of the resultant neo-glycopeptides. Chapter Five describes the synthesis of hybrid neo-glycopeptides incorporating both the short 18-mer Adpn sequence as well as the CMP sequence.

The main conclusions of this thesis are that both unglycosylated and glycosylated lysine residues play a fundamental structural role in the collagenous domain of Adpn. Furthermore, the degree of glycosylation of the lysine residues is integral to the structural stability of Adpn. There may be an ‘optimal level’ of glycosylation for collagenous peptides, above which the secondary structure of the peptide is significantly altered. These conclusions could have major implications on the understanding of the forces that underpin the structure and function of many collagenous peptides.
Preface

All synthetic work described in this thesis was carried out by the author under the supervision of D/Professor Margaret A. Brimble and Professor Garth J.S. Cooper in the School of Chemical Sciences and the School of Biological Sciences at the University of Auckland, except for the synthesis of the crude peptides which were carried out by research fellows due to the author’s acquired allergy to coupling agents. The crude peptides were purified and characterised by the author. Guidance with peptide design and synthesis was given by Dr Paul Harris.

Some parts of this thesis have been previously published:


- Herbert, Katherine R.; Harris, Paul W. R.; Sparrow Kevin J; Kaur, Harveen; Bella, Jordi; Cooper, Garth J. S.; Brimble, Margaret A. JACS, 2017, submitted.
For Christof

“Persistence beats resistance”
Acknowledgements

Completion of this thesis could not have been possible without the tremendous support I received through the university and my family and friends. The last five years have taught me how to be a scientist and I am very grateful to have had this opportunity to conduct research in a fascinating project in very beautiful part of the world.

As always, it is the people that make an experience so special. Many thanks to my main supervisor D/Professor Margaret Brimble and my co-supervisor Professor Garth Cooper for the opportunity to research adiponectin and for your support for continuing with this project when faced with a significant allergy to coupling agents. Thank you, Margaret, for giving me the technical support for the peptide synthesis and thus enabling me to complete this PhD project. Thank you, Garth, for the stimulating biology conversations and for your introduction to Dr Jordi Bella, whose insight into the structure of collagen has been of great importance and interest for this thesis.

Many thanks to my lab mates past and present, especially Sandhya, Mathilde, Morgan, Dave, Jared, Megan, Nora, Michelle, Lisa and Roger. Thank you also to Janice Choi and Tim Layt for their help and assistance with chemicals and equipment. Many thanks also to all the teaching lab staff for letting me use their lab space over the summer and for many crucial cups of tea! A huge thanks to Dr Hannah Hogben, Dr Andy Herbert, Dr Harveen Kaur, Dr Michelle Van Rensburg and Ho Yeung for proof-reading my thesis and to Dr Paul Hume for his assistance with naming my novel triazole compounds.

I am also very grateful to the fantastic post-doctoral fellows who have offered me guidance and support at various stages of this PhD project including my PhD advisors Dr Paul Harris and Dr Geoffrey Williams as well as Dr Tom Woods, Dr Manuel Johannes and Dr Renata Kowalczyk. A huge thanks also to the research fellows who conducted peptide synthesis on my behalf: Dr Paul Harris, Dr Kevin Sparrow, Dr Sung-Hyun Yang and Tom Wright.

Thank you also to the extraordinarily capable Cynthia Tse for her continued advice and support for this project. I am also very grateful to the Cooper group for helping me understand the biological aspects of this PhD project, especially Dr Kate Lee and Dr Lance Xu. I am also very grateful to my group of SBS post-doctoral friends who have offered structural biology advice, guidance and friendship during this PhD, especially Genevieve, Jodie, Harriet and Esther.
I would also like to acknowledge the support of my high school chemistry teacher Mrs Gina Reese-Boughton who inspired me to study chemistry at university and also to my uncle Dr John Maund for his continued interest in my project and all things chemistry!

I am incredibly lucky to have a fantastic family who have supported me in every step along this journey. Thank you mum, dad and Phil for coming over to New Zealand so many times to visit and for all the care parcels of tea you have sent over! Thank you also to my grandparents for keeping in touch, especially to my grandmother Daphne for her weekly letters. Thank you to my aunts and uncles for their support and an overwhelming thank you to Andy and Carol for inviting me to visit you in New Zealand. You were the first step of this fantastic journey- thank you! Thank you also to my soon-to-be parents-in-law Gottfried and Birgit for their support along the way.

I have also been lucky enough to have an amazing set of friends supporting me from the UK including Rachel, Kat, Hannah, Georgie, Emma, Mary, Claire and Nadine. Thanks for keeping in touch and see you all soon! Thank you to all the wonderful friends I have made through AUTC including Annie, Sarah, Richard, Hannah, Jarra and Tom.

Finally many thanks to my Kiwi whanau. Chris and Marijke, I am grateful to you a thousand times over for welcoming me into your lives and supporting me though this rollercoaster of a ride. Christof, you have been an incredible tower of support for completion of this PhD. Thank you so much for all the love and support you have given me and for always having faith that the thesis would be finished in the end!
Abbreviations

3Å MS = Molecular sieves of size 3Å
4Å MS = Molecular sieves of size 4Å
9-BBN= 9-Borabicyclo[3.3.1]nonane
AA = Amino acids
Adpn = Adiponectin
Adpn(66-83) = Small peptides synthesised corresponding to residue numbers 66 – 83 of Adpn
AgOTf = Silver triflate
AGPs = Arabinogalactans, common plant-cell wall glycoproteins
Ala = Alanine
AOP = 7-Azabenzotriazol-1-yloxy)tris(di-methylamino)phosphonium hexafluorophosphate
AQN = Anthraquinone
Ar = Argon
Asp = Aspartic acid
BF$_3$.$\text{Et}_2$O = Boron trifluoride diethyl etherate
BOC/Boc = tert-Butyloxycarbonyl protecting group
BOMP = 2-(Benzotriazol-1-yloxy)-1,3-dimethyl-2-pyrrolidin-1-yl-1,3-diazaphospholidinium hexafluorophosphate
BOP = Benzotriazol-1-yl- N\text{-}oxy-tris(dimethylamino)phosphonium hexafluorophosphate
c. = Circa
CD = Circular dichroism
CHO = Chinese hamster ovary cell line
CIP = Contact ion pair
CMPs = Collagen model peptides
COS-7 = African Green Monkey cell line
CPG = Controlled pore glass support for oligonucleotide
CuAAC = Cu(I)-catalysed Azide-Alkyne Cycloaddition reaction
d.e. = Diastereomeric excess
DCE = 1,2-Dichloroethene
DHQ = Dihydroquinine
(DHQ)$_2$AQN = 1,4-\textit{bis}(9-O-Dihydroquinine) anthraquinone
(DHQ)$_2$PHAL = 1,4-\textit{bis}(9-O-Dihydroquinine) phthalazine
DHQD = Dihydroquinidine
(DHQD)$_2$PHAL = 1,4-\textit{bis}(9-O-Dihydroquinidine) phthalazine
DIPEA = \textit{N},\textit{N}-Diisopropylethylamine
DODT = 2,2'-(Ethylenedioxy)diethanethiol
DMAP = \textit{N},\textit{N}'-Dimethylaminopyridine
DMF = Dimethylformamide
DMSO = Dimethyl sulfoxide
EDCI = 1-\textit{Ethyl}-3-(3-dimethylaminopropyl)carbodiimide
\textit{e.e.} = Enantiomeric excess
\textit{epi}-Hyl = Epimeric-Hyl
ESI = Electron spray ionisation
fcc = Flash column chromatography
Fmoc = 9-Fluorenylmethylxycarbonyl protecting group
Fmoc-OSu = \textit{N}-(9-Fluorenylmethoxycarbonyloxy)succinimide
FI-MS = Flow-inject mass spectrometry
Gln = Glutamine
Glu = Glutamic acid
Gly = Glycine
Gn.HCl = Guanidine hydrochloride
HATU = 1-[\textit{Bis}(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate \textit{N}-oxide
HBTU = \textit{N},\textit{N},\textit{N}',\textit{N}'-Tetramethyl-O-(1H-benzotriazol-1-yl)uronium hexafluorophosphate \textit{N}-oxide
HCTU = (2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate) \textit{N}-oxide
HOAt = 7-aza-1-hydroxybenzotriazole
HOBt = 1-hydroxybenzotriazole
HMBA = 4-Hydroxymethylbenzoic acid linker
HMPP = Hydroxymethylphenoxy propionic acid linker
HPLC = High performance liquid chromatography
Hyp = (4R)-Hydroxyproline
'PrOH = Isopropyl alcohol
LC-MS = Liquid chromatography mass spectrometry
LH3 = Lysyl hydroxylase isoform 3
LP = Lone Pair
Lys-PTM = Lysine residue which has been post-translationally modified
Lys = Lysine residue
M.S. = Molecular sieves
Mass spec. = Mass spectrometry
MBHA = para-Methylbenzhydrylamine
MCPBA = Meta-Chloroperoxybenzoic acid
MeOH = Methanol
Ms = Milliseconds
NGP = Neighbouring group participation
NOESY = Nuclear Overhauser Spectroscopy
NMM = N-Methylmorpholine
NMP = N-Methylpyrrolidone
NMR = Nuclear magnetic resonance
NP-HPLC = Normal phase HPLC
Oxyma = Ethyl-2-cyano-2-(hydroxyimino)acetate
o/n = Overnight
PEGA = Poly(acryloyl-bis(aminopropyl)polyethylene glycol) resin
POC = Peptide-oligonucleotide conjugates
POG = Pro-Hyp-Gly, a common triplet found in collagen model peptides
PS = Polystyrene resin utilised in Fmoc-SPPS
PyAOP = 7-Azabenzo[d]tetrazol-1-yloxy)tris(pyrrolidino)phosphonium hexafluorophosphate
RP-HPLC = Reverse phase high performance liquid chromatography
Rt = Retention time
Rt = Room temperature
RuAAC = Ru(II)-catalysed Azide-Alkyne Cycloaddition reaction
SAH = Sharpless aminohydroxylation
SEC-MALS = Size exclusion chromatography with multiangle laser light scattering
SEM = Scanning electron microscopy
SIP = Solvent shared ion pair
SPPS = Solid phase peptide synthesis
tBuOH = tert-Butyl alcohol
T3 = Type III collagen
TCEP.HCl = Tris(2-carboxyethyl)phosphine hydrochloride
TFA = Trifluoroacetic acid
Thr = Threonine
TIPS = Triisopropylsilane
TLC = Thin layer chromatography
Tm = Thermal melting point of a protein
TMP = 2,2,6,6-tetramethylpiperidine
TMSOTf = Trimethylsilyl trifluoromethanesulfonate
TZDs = Thiazolidinediones
μw = Microwave
# Table of Contents

Abstract ..................................................................................................................................................... i  

Preface...................................................................................................................................................... ii  

Acknowledgements ................................................................................................................................... v  

Abbreviations ......................................................................................................................................... vii  

Table of Contents .................................................................................................................................... xi  

Chapter One: Adiponectin as a Potential Therapeutic for Type II Diabetes ................................. 1  
1.1 Prevalence of Type II diabetes in New Zealand ............................................................................ 1  
1.2 The role of adiponectin in obesity-related diseases ................................................................. 1  
  1.2.1 Background ............................................................................................................................. 1  
  1.2.2 Hypoadiponectinemia ............................................................................................................. 1  
  1.2.3 Adiponectin receptors ............................................................................................................. 3  
1.3 Adpn species cloned to date ........................................................................................................... 4  
1.4 Oligomeric states of adiponectin .................................................................................................... 4  
1.4 Different recombinant forms of adiponectin, bacterial versus mammalian .............................. 5  
1.5 Adiponectin as a therapeutic target for the treatment of T2D ................................................... 7  
  1.5.1 Lifestyle modifications ........................................................................................................... 7  
  1.5.2 Administration of anti-diabetic agents .................................................................................... 7  
  1.5.3 Other medications known to affect Adpn levels ..................................................................... 8  
  1.5.4 Expression of adiponectin using recombinant techniques ................................................... 8  
  1.5.5 The total chemical synthesis of adiponectin ....................................................................... 9  
  1.5.6 Synthesis of small molecule therapeutics to stimulate adiponectin production .................. 9  
  1.5.7 Targeting of a suitable intracellular signalling pathway ....................................................... 10  
1.6 Aims of this PhD project ............................................................................................................. 10  
1.7 Outline of this doctoral thesis ..................................................................................................... 12  

Chapter Two: Towards the Synthesis of Native Glycosylated Hydroxylysine ............................... 13  
2.1 The chemical structure of native, glycosylated hydroxylysine ............................................... 13  
2.2 Historical isolation of Hyl and Gal-Glu-Hyl ............................................................................. 14  
2.3 Previous synthesis of (2S, 5R)-Hyl .............................................................................................. 15  
  2.3.1 The different approaches employed to date .......................................................................... 15
2.3.1.1 Epimeric synthesis ......................................................................................................... 15
2.3.1.2 Epimeric synthesis followed by separation using silica chromatography ................. 16
2.3.1.3 Epimeric synthesis with use of a chiral resolving agent for separation of diastereoisomers ......................................................................................................................... 17
2.3.1.4 Epimeric synthesis of epi-Hyl utilising a chiral auxiliary ............................................. 18

2.3.2 Asymmetric synthesis using chiral catalysis ........................................................................ 20

2.3.3 Substrate directed asymmetric synthesis: use of piperidinone intermediates ....................... 21

2.4 Previous syntheses of Gal-Hyl 1.9 ............................................................................................... 22
2.4.1 Use of an α-1,2-anhydro sugar as the galactose sugar donor ............................................... 23
2.4.2 Use of an acetyl bromide as the galactose sugar donor ........................................................ 23
2.4.3 Use of a trichloroacetimidate as the galactose sugar donor .................................................. 25

2.5 Previous syntheses of Glu-Gal-Hyl 1.8 ....................................................................................... 27
2.5.1 Use of a trichloroacetimidate as the glucose sugar donor .................................................... 27
2.5.2 Use of an activated thioethyl glycoside as the glucose sugar donor ..................................... 27

2.6 Summary of reported syntheses of Gal-Hyl 1.9 and Glu-Gal-Hyl 1.8 to date............................ 29

2.7 Selected synthetic strategy towards (2S,5R)-Hyl 2.1 utilised in this PhD thesis ....................... 30
2.7.1 Synthesis of (2S,5R)-Hyl 2.1 utilising selective oxidation of a chiral piperidinone enolate .......................................................................................................................... 30

2.8 Revised strategy; use of Belokon complex and Sharpless aminohydroxylation reaction for the synthesis of (2S,5R)-Hyl ................................................................................................... 31
2.8.1 Introduction to the use of the Belokon complex for the stereoselective synthesis of amino acids .......................................................................................................................... 32
2.8.2 The use of the Belokon complex as a chiral auxiliary for the synthesis of protected (2S)-hex-5-enoic acid ................................................................................................................ 34
2.8.2.1 The key alkylation reaction of Belokon complex 2.93 .................................................. 34
2.8.3 Use of the Sharpless aminohydroxylation reaction for furnishing of the (5R)-stereocentre of Hyl............................................................................................................................. 37
2.8.3.1 Introduction to the Sharpless aminohydroxylation reaction........................................ 38
2.8.3.2 Proposed reaction mechanism for the Sharpless aminohydroxylation reaction............ 39
2.8.3.3 Mode of alkene addition for the Sharpless aminohydroxylation reaction .................... 40
2.8.3.4 Use of chiral cinchona alkaloids to control both the regioselectivity and stereoselectivity of the Sharpless aminohydroxylation reaction ......................................................... 41
2.8.3.5 Use of tethered substrates to improve the regioselectivity of the Sharpless aminohydroxylation reaction ...................................................................................................................... 44
2.8.3.6 Conditions investigated for the SAH reaction of model starting materials and (2S)-hex-5-enoic acid .................................................................................................... 44
2.8.3.7 Effect of a chiral ligand on SAH reaction of protected-(2S)-hex-5-enoic acid derivatives ...................................................................................................................... 48
2.8.3.8 Regiochemistry observed for the achiral reaction conditions ........................................ 50
2.8.4 Lactonisation studies of the SAH products ........................................................................... 51
2.9 Use of epoxidation reactions to afford the desired (5R) stereocentre of Hyl ............................... 53
2.9.1 Revised retrosynthetic analysis for (2S,5R)-Hyl, 2.1 ........................................................... 53
2.9.2 Similar synthetic strategies employed by other research groups ........................................ 53
2.9.3 Chiral epoxidation reactions ................................................................................................. 55
2.10 Use of 9-BBN as a protecting group for (2S,5R)-Hyl .............................................................. 57
2.11 Summary of our approach towards the synthesis of (2S,5R)-Hyl ............................................... 60
2.12 Retrosynthetic strategy towards Gal-Hyl 1.9 ........................................................................ 62
2.13 Synthesis of the galactosyl donor .......................................................................................... 62
2.14 Coupling of the galactose donor to epi-Hyl derivatives .............................................................. 63
2.14.1 Glycosylation of epi-BBN-Hyl with galactose donor ........................................................... 63
2.14.3 Glycosylation studies of epi-Fmoc/OMe-Hyl and epi-Fmoc/OAll-Hyl ............................... 68
2.11 Conclusion for Chapter Two ........................................................................................................ 69

Chapter Three: The Synthesis of Neoglycopeptides Corresponding to the Collagenous Domain of Adiponectin. ............................................................................................................................................ 69
3.2 Overview of the biological activity of Adpn ........................................................................... 69
3.3 Previous chemical syntheses of Adpn ...................................................................................... 69
3.4 Utilisation of neoglycosides as robust glycoside mimetics ........................................................ 71
3.5 Introduction to the Copper-Catalysed 1-3-Dipolar Cycloaddition Reaction, (CuAAC) ........... 72
3.5.1 The development of 1,3-dipolar addition reactions .................................................................. 72
3.5.2 The pericyclic reaction mechanism for 1,3-dipolar cycloaddition reactions ...................... 74
3.5.3 Introduction of a copper catalyst to 1,3-dipolar cycloaddition reactions; the birth of the CuAAC reaction ............................................................................................................ 74
3.5.4 Proposed mechanisms for the thermal and Cu(I) catalysed cycloaddition ......................... 74
3.5.5 Mono-copper versus bis-copper pathways for the CuAAC reaction ................................ 75
3.5.6 Ruthenium catalysis of 1,3-dipolar cycloaddition reactions ................................................ 76
3.5.7 Other transition metal catalysis of 1,3-dipolar cycloaddition reactions .............................. 78
3.6 Use of the CuAAC reaction to generate ‘clicked’ peptides .................................................... 78
3.7 Use of the CuAAC reaction to afford neoglycopeptides ........................................................ 80
3.7.1 Previous synthetic studies on the MUC-1 neoglycopeptides ............................................. 83
3.14.3 Incorporation of neogalactolysine and neogalactoserine building blocks into Adpn peptides using Fmoc SPPS .................................................................................................................. 130

3.15 Examination of secondary structure of neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 using circular dichroism ........................................................................................................... 137

3.16 The molecular basis of triple helical conformation of collagenous peptides .......... 138

3.16.1 The effect of steric constraints upon triple helical conformation of proteins .......... 138

3.16.2 The role of hydrogen bonding networks in formation of a triple-helical conformation ......... 139

3.16.3 The importance of electrostatic interactions for triple-helical stability ..................... 139

3.17 Hypothesis for the random coil conformation exhibited by deacetylated neoglycopeptides 
   Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.61b ......................... 141

3.18 Conclusion for Chapter Three ......................................................................................... 142

Chapter Four: The Use of Collagen Model Peptides as Structural Scaffolds for Neoglycopeptides 
of Adpn ...................................................................................................................................................... 145

4.1 Introduction to collagen model peptides ........................................................................... 145

4.2 Structural features of CMPs ............................................................................................. 147

4.3 Previous studies of glycosylated CMPs ........................................................................... 148

4.3.1 CMPs bearing glycosylated hydroxyproline residues .............................................. 148

4.3.2 Structural models of glycan-peptide backbone interactions .................................... 150

4.3.3 Incorporation of glycans into a CMP sequence using the copper-catalysed 
   1,3-cycloaddition ‘click’ reaction ..................................................................................... 150

4.3.4 CMPs containing glycosylated threonine residues ................................................... 151

4.4 Chosen CMP sequences for synthesis ............................................................................. 152

4.5 Synthesis of Family 2 CMPs .......................................................................................... 153

4.6 Purification and deacetylation of Family 2 CMPs .......................................................... 169

4.7 Key observations from CD studies of Family 2 CMPs .................................................... 176

4.7.1 Structural differences observed between acetylated and deacetylated CMPs .......... 176

4.7.1.1 The role interstrand hydrogen bonding networks for structural stability of CMPs ..... 176

4.7.1.2 The importance of intrastrand hydrogen bonding networks for stabilisation of the 
   triple helical conformation of CMPs ........................................................................ 179

4.7.2 Other reported uses of an acetyl protecting group to investigate hydrogen bonding 
   networks in CMPs ........................................................................................................ 179

4.7.2.1 The use of facilitating hydration networks for extension of hydrogen bonding ....... 180

4.7.3 Comparison of control non-glycosylated CMP with glycosylated CMP ................. 181

4.7.4 Investigation of different levels of glycosylation of Family 2: CMPs ......................... 182

4.7.4.1 The effect of steric interactions for triple helical conformation of CMPs ............. 183
4.7.4.2 The role of hydrogen bonding networks in glycosylated CMPs: monomeric single
stranded state versus triple helical conformation ......................................................... 184
4.7.4.3 Role of guest domain length in structural stabilisation of CMPs ......................... 185
4.8 Proposal of an 'optimal level' of glycosylation for collagenous peptides ..................... 185
4.8.1 Comparison of lysine and serine CMP series .......................................................... 186
4.9 Thermal stability of Family 2: CMPs .......................................................................... 188
4.9.1 Comparison of thermal stabilities of Family 2: CMPs with previously reported similar
peptides ............................................................................................................................ 189
4.10 Summary of results for Chapter Four ....................................................................... 190
4.11 Conclusion for Chapter Four .................................................................................... 190

Chapter Five: Lysine, the Keystone of the Triple Helical Structure of the Collagenous Domain of
Adiponectin .......................................................................................................................... 189
5.1 Summary of work presented in previous chapters .................................................... 189
5.2 The synthesis of a ‘host-guest’ CMP-Adpn family of peptides .................................... 191
5.3 Circular dichroism studies of CMP-Adpn family of peptides .................................... 200
5.4 Thermal stability of the CMP-Adpn peptides 5.1b and 5.2 ......................................... 201
5.4.1 Comparison of thermal stability of CMP-Adpn peptides 5.1b and 5.2 with similar
literature glycopeptides .................................................................................................... 201
5.5 Conclusions drawn from the circular dichroism studies of the CMP-Adpn peptides .... 204
5.6 Role of electrostatic interactions for the formation and stability of the triple helical
conformation of collagens ................................................................................................. 205
5.6.1 Observations of the relative positioning of acidic and basic resides in the peptide
sequence ........................................................................................................................... 205
5.6.2 Synthesis of ‘host-guest’ CMPs to investigate the role of ion-pairs on the thermal
stability of such peptides ............................................................................................... 205
5.7 Effect of pH on thermal stability of collagens and CMPs ............................................ 207
5.8 Crystal structures of CMPs illustrating adjacent chain ion-pairs ............................... 209
5.9 The role of hydrogen bonding in the structural stability of collagenous peptides ...... 211
5.10 The use of water molecules to extend the hydrogen bonding network and for electrostatic
ion pair formation .......................................................................................................... 212
5.11 Proposed molecular structure of the non-glycosylated, control CMP peptides 5.2 and 5.4 ... 213
5.12 Proposed molecular structure of glycosylated CMP-Adpn peptides 5.1b and 5.3b ....... 214
5.13 Glycosylated peptides in the context of diabetes ...................................................... 216
5.14 Conclusion for Chapter Five ..................................................................................... 218
Chapter Six: Summary and Conclusions

6.1 Summary of doctoral studies reported in this thesis

6.2 Comparison of Family 2 CMP and Family 3 Adpn-CMP peptide families

6.2.1 Observed structural similarities between the Family 2 CMP and Family 3 Adpn-CMP neoglycopeptide families

6.2.2 Observed structural differences between the Family 2 CMP and Family 3 Adpn-CMP neoglycopeptide families

6.2.3 Proposed difference in triple helical pitch between Family 2 CMPs and Family 3 Adpn-CMPs

6.3 Proposed model for triple helical secondary structure of collagenous peptides

6.4 Application of the triple helical secondary structure model to CMP and CMP-Adpn peptides

6.5 Relevance of conclusions to other collagen based systems

6.6 Concluding remarks

6.7 Future work

Chapter Seven: Experimental Methods and Product Characterisation

1.1 Section A: Details for organic compounds

1.2 Section B: Details for peptide synthesis

Appendix NMR Spectra For Novel Organic Compounds

References
Chapter One

Adiponectin as a Potential Therapeutic for Type II Diabetes
Chapter One: Adiponectin as a Potential Therapeutic for Type II Diabetes

1.1 Prevalence of Type II diabetes in New Zealand

According to the 2014/15 New Zealand Health Survey\(^1\) there are approximately 257,000 diagnosed cases of diabetes in New Zealand, of which approximately 203,000 are Type II diabetes (T2D). Furthermore, the prevalence of T2D is much higher in Maori (8%) and Pacifica (17%) populations compared to the overall prevalence (6%), Figure 1.1. According to Diabetes NZ, the expected growth rate for diagnosed cases of T2D in New Zealand is 8% per annum, causing T2D to be a major health concern in New Zealand. As T2D patients typically present with low levels of adiponectin (Adpn) compared to non-diabetics (Section 1.2.2),\(^2\) this protein could hold promise as a potential new therapeutic for T2D.

![Figure 1.1 Prevalence of Type II diabetes in New Zealand.](image)

1.2 The role of adiponectin in obesity-related diseases

1.2.1 Background

Adiponectin (Adpn, also known as ACRP 30) was first isolated from mouse blood plasma by Scherer \(\text{et al.}\)\(^3\) in 1995 and has since been identified in humans as the primary protein secreted by white adipose tissue.\(^3\)-\(^5\) Numerous biological studies have extensively proven the key role of adiponectin in diseased states including obesity-related diseases collectively known as ‘metabolic syndrome’,\(^6\) Type II diabetes mellitus (T2D),\(^2,\)\(^7,\)\(^8\) cardiovascular disease\(^4,\)\(^9\) and most recently melanoma.\(^10\)

1.2.2 Hypoadiponectinemia

These disease states have all been shown to be associated with low levels of plasma adiponectin known as hypoadiponectinemia. Normal plasma levels of Adpn are classified as 5 - 25 µg/mL\(^8,\)\(^11\) with blood plasma levels lower than 5 µg/mL categorised as hypoadiponectinemic. There is a clear inverse relationship between body mass index (BMI) and/or body fat levels and Adpn levels, with multiple studies showing increased obesity levels being closely correlated with decreased plasma levels of Adpn.\(^12\)-\(^14\) In the first such study published in 1999, Arita \(\text{et al.}\)\(^12\) examined plasma concentrations of
Adpn in 87 non-obese subjects and 57 obese subjects and found a strong negative correlation between increased BMI and reduced blood plasma levels of Adpn. Indeed, the mean of the non-obese subjects was calculated to be 8.9 µg/mL (+/- standard deviation of 5.4) considerably higher than the average of the obese subjects which was found to be 3.8 µg/mL (+/- standard deviation of 3.2).

A later study, with a cohort of 62 weight-stable adult subjects, published by Ranganatha et al.\textsuperscript{13} clearly illustrated the relationship between increased body fat levels, Figure 1.2a), or BMI, Figure 1.2b), and lower blood plasma levels of Adpn. Also, interestingly the results of this study also showed women to have higher blood adiponectin levels compared to their male equivalents of similar percent body fat. The authors hypothesise that this gender bias may be due to differences in body composition between genders.

This same study also indicated the possible causal relationship of lower plasma levels of Adpn leading to reduced insulin sensitivity, Figure 1.3. Insulin sensitivity was recorded using the insulin sensitivity index (\(S_I\)) as described by Pedersen et al.\textsuperscript{15} whereby \(S_I\) represents the net increase in glucose clearance per unit of serum insulin following intravenous glucose injection. In contrast to the observed relationships between BMI and/or body fat and plasma Adpn, the relationship between decreased plasma Adpn levels and corresponding reduction in insulin sensitivity was shown to be gender monomorphic.

\[\text{Figure 1.2 }\text{Plasma adiponectin levels are correlated to}\ a)\ \text{percent body fat and}\ b)\ \text{BMI. Regression lines in Figure 1.2 are drawn using a least-squares regression.}\textsuperscript{13}\]

\[\text{Figure 1.3}\ \text{The relationship between plasma adiponectin and insulin sensitivity (}\text{S}_i\text{).}\textsuperscript{13}\]
These clear links between obesity, low plasma levels of Adpn and subsequent insulin resistance were also observed by Weyer et al.\textsuperscript{14} Furthermore, in their study Weyer et al.\textsuperscript{14} proved that these relationships occur in different populations as they examined results from both Pima Indian and Caucasian populations. Frampton et al.\textsuperscript{16} have conducted a similar study in New Zealand comparing Adpn plasma concentrations between Maori and Caucasian populations with a study size of 111 participants. Their results indicated that for matched participants (gender, age, BMI, waist circumference and insulin sensitivity) Adpn plasma concentration was significantly lower in Maori participants 7.3 µg/mL (+/- standard deviation of 4.0) compared to Caucasian participants 8.3 µg/mL (+/- standard deviation of 4.2) and hypoadiponectinemia was found to be twice as prevalent in the Maori population.

As the Frampton study\textsuperscript{16} was carried out with matched participants in terms of BMI, the decreased Adpn levels observed across the Maori population indicates a genetic disposition to low Adpn levels. Indeed mutations in the Adpn gene located on chromosome 3q27 have been shown to lead to increased susceptibility to diabetes,\textsuperscript{17-19} although gene mapping for specific populations has not yet been published.

### 1.2.3 Adiponectin receptors

As described above, Adpn has been shown to be linked to insulin sensitivity. Adpn has also been identified as a key player in glucose homeostasis and protection against the development of atherosclerosis. These different biological activities of Adpn will now be considered in terms of identified Adpn receptors. Two main receptors for Adpn have been identified to date; the first in skeletal muscle (AdipoR1) and the second in the liver (AdipoR2).\textsuperscript{20} The vastly different locations of these two receptors, point to two different modes of action for Adpn \textit{in vivo}. The first mode of action of Adpn, known as the muscular skeletal pathway occurring \textit{via} AdipoR1, is reduction of blood plasma glucose and fatty acid levels. Blood glucose levels are reduced by increased glucose uptake by the muscle\textsuperscript{21} and fatty acid levels are decreased by increased lipid oxidation.\textsuperscript{21,22} The second mode of action of Adpn takes place in the liver tissue occurring \textit{via} AdipoR2, where Adpn has been shown to both decrease hepatic glucose output as well as increase insulin sensitivity. These liver-based activities occur by reducing gluconeogenesis and promoting fatty acid catabolism \textit{via} the peroxisome proliferator-activated receptor (PPAR)-α signalling pathway.\textsuperscript{4,23-25}

The cardiovascular protective properties of Adpn are thought to relate to the downstream effect of Adpn raising nitrogen oxide secretion in the endothelium,\textsuperscript{4} thereby promoting inhibition of adhesion factors linked to atherosclerosis. There is also some evidence towards a third receptor for Adpn, which is known as T-cadherin and is located in vascular tissue. It is thought that this possible third receptor for Adpn selectively binds only hexameric and high molecular weight oligomers (HMW) of Adpn.\textsuperscript{26} The pathways associated with T-cadherin are yet to be defined, however as the receptor is found only in cardiovascular, nervous and skeletal muscular tissue, the pathways associated with this third receptor are likely to be
aligned with the cardiovascular protective functions of Adpn rather than the anti-diabetic properties of Adpn.

1.3 Adpn species cloned to date

To date eleven Adpn species have been cloned including human, mouse, bovine and monkey Adpn. \(^\text{27}\) The length of the Adpn peptide sequence varies between species between 194 - 292 amino acids (AA); with human (\textit{homo sapiens}) Adpn consisting of 244 AA residues and murine (\textit{mus musculus}) Adpn consisting of 247 AA residues. \(^\text{27,28}\) Murine Adpn is the most extensively studied clone with both complete protein sequence data as well as exact mapping of post-transitional modification (PTM) sites available. \(^\text{29}\) Furthermore, the identified Lys-PTM (lysine post-translationally modified) residues identified by Wang \textit{et al.} \(^\text{28}\) as important for bioactivity of the protein (see Section 1.5) and occurring at positions 68, 71, 80 and 104 in murine Adpn have also been shown to be conserved in other mammalian species, Figure 1.4. Therefore, murine Adpn was shown as a suitable model for Adpn for this thesis, and murine Adpn will be referred to as Adpn hereafter.

\textbf{Figure 1.4} Conservation of the PTM-Lys residues found in the collagenous domain of Adpn across four different species; mouse, human, bovine, monkey and dog, as reported by Wang \textit{et al.} \(^\text{28}\)

1.4 Oligomeric states of adiponectin

Adpn has been shown to occur in blood plasma in four main oligomerisation states, namely monomers, trimers, low molecular weight oligomers (LMW, c. 180 KDa) and high molecular weight oligomers (HMW, c.400 KDa), Figure 1.5b). \(^\text{30-32}\) The HMW oligomers of Adpn have been shown to be the most bioactive form of the protein, especially for the protein’s anti-diabetic properties. \(^\text{23}\) The monomeric form of Adpn is a 30 KDa peptide consisting of four different peptide segments; the \(N\)-terminal signal peptide (SP, residues 1 – 21), the variable domain (residues 22 – 45), the collagenous domain (residues 46 – 107) and a \(C\)-terminal globular domain (residues 108 – 247), \(^\text{28}\) Figure 1.5a).
1.4 Different recombinant forms of adiponectin, bacterial versus mammalian

The biological mode of action of these different oligomeric states of Adpn has been widely investigated. Adpn produced recombinantly using bacterial cell-lines can only reach the trimer oligomeric state, due to the lack of post-translational modifications (PTMs) present in the peptide. These recombinant Adpn trimers have been shown to be biologically active and to utilise the muscular skeletal pathway. The Adpn trimers are able to lower blood glucose, fatty acid and triglyceride levels in murine models via AMPK (adenosine monophosphate protein kinase)-activation causing increased glucose uptake by muscular tissue and oxidation of available fatty acids.

On the other hand, Adpn isolated from or produced using a mammalian expression system has been shown to form high levels of the higher order structures, namely LMW and HMW oligomers. The primary target for these larger oligomers has been shown to be the liver tissue (AdipoR2) with the primary function of these larger oligomers being increasing insulin sensitivity. This enhanced insulin sensitivity was shown to lead to decreased hepatic glucose output via reduction of the expression of gluconeogenic enzymes.

Cooper et al. have shown that PTMs present in mammalian-produced Adpn are crucial for the formation of HMW oligomers of Adpn and subsequent insulin-sensitising properties of the protein. They employed site-directed mutagenesis to substitute the lysine residues (K) at positions 68, 71, 80 and 104 of the collagenous domain of murine Adpn with structurally similar arginine residues (R), Figure 1.6a). This K/R substitution resulted in decreased levels of bioactive HMW oligomers for the Adpn-K/R substituted product (Adpn-K/R, 1.2) compared to the native mammalian adiponectin (Adpn-WT, 1.1), Figure 1.6b).
Richards et al.\textsuperscript{35} have conducted a similar K/R substitution study of murine Adpn and also observed that their Adpn-K/R variant was unable to form HMW oligomers. Furthermore, Richards et al.\textsuperscript{35} also investigated the role of enzymes involved in the post translational modifications of lysine residues, such as lysyl-hydroxylases. They observed that when these enzymes were inhibited, HMW oligomers were also unable to be formed. Taken together, the observations from Cooper et al.\textsuperscript{34} and Richards et al.\textsuperscript{35} indicate that post-translationally modified lysine residues may play a crucial role in the formation of bioactive HMW oligomers of Adpn. Furthermore, research conducted by Cooper et al.\textsuperscript{36} indicates the possibility of an optimal level of these post-translational modifications in adiponectin. This idea will be explored in detail later in the context of this thesis.
1.5 Adiponectin as a therapeutic target for the treatment of T2D

Due to the central role of Adpn in the regulation of blood glucose and insulin sensitivity, multiple therapeutic approaches have been undertaken to increase blood plasma levels of Adpn in T2D patients. These therapeutic approaches include lifestyle modifications, the administration of anti-diabetic agents, the expression of adiponectin using recombinant techniques, the total chemical synthesis of Adpn, the synthesis of small molecule therapeutics to stimulate adiponectin production and targeting of a suitable intracellular signalling pathway. These different approaches are described in detail below.

1.5.1 Lifestyle modifications

The clear relationship between hypoadiponectinemia and multiple disease states has prompted research into methods to raise blood plasma Adpn levels. Adpn levels are closely linked with obesity (see Section 1.2.2) and lifestyle choices. Factors such as increased exercise, the adoption of a Mediterranean diet and increased fish consumption have been shown to raise plasma Adpn levels in diabetic patients. Increased coffee consumption amongst both diabetic and non-diabetic women has also been shown to lead to increased plasma Adpn concentration, probably due to the metabolic enhancing effect of caffeine. More drastic measures, such as reduction in body weight via bariatric surgery, have also been shown to be effective in raising blood plasma levels of Adpn.

1.5.2 Administration of anti-diabetic agents

Several anti-diabetic agents including glimepiride 1.3 have shown clinical effectiveness towards increasing Adpn levels, Figure 1.8. Interestingly, administration of thiazolidinediones such as troglitazone 1.4 and pioglitazone 1.5 (TZDs, known PPARγ agonists) to both rodent and human subjects have also been shown to be effective in raising Adpn levels, whereas administration of metformin, a PPARα agonist, was shown to have no effect.

![Functional group of the thiazolidinediones (TZDs)](image)

**Figure 1.7** Chemical structures of glimepiride 1.3, troglitazone 1.4 and pioglitazone 1.5 (linked to increased Adpn production). Stereochemistry of compounds used in the clinical studies were not reported.

In New Zealand, a comparative study of the effects of administering pioglitazone on T2D patients from Maori-Polynesian and Caucasian populations was carried out by Frampton et al. with a cohort of 81 patients over a six month period. In both ethic groups, pioglitazone was shown to more than double
plasma adiponectin concentration and reduce fasting blood glucose levels, blood plasma cholesterol concentration and blood plasma triglyceride content, Table 1.1. Interestingly, through-out the study plasma Adpn levels were found to be higher in the Caucasian participants compared to the Maori-Polynesian participants, despite similar ages and BMI values.

Table 1.1 Blood indicator changes in Caucasian and Maori-Polynesian ethnic groups following administration of pioglitazone (45 mg/day for 6 months).

<table>
<thead>
<tr>
<th>Blood Indicator</th>
<th>Percent change observed for Caucasian participants</th>
<th>Percent change observed for Maori-Polynesian participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasma adiponectin concentration (µg/mL)</td>
<td>+116%</td>
<td>+118%</td>
</tr>
<tr>
<td>Fasting glucose levels (mmol/L)</td>
<td>-21%</td>
<td>-24%</td>
</tr>
<tr>
<td>Plasma cholesterol concentration (mmol/L)</td>
<td>-4%</td>
<td>-2%</td>
</tr>
<tr>
<td>Plasma triglyceride concentration (mmol/L)</td>
<td>-28%</td>
<td>-16%</td>
</tr>
</tbody>
</table>

1.5.3 Other medications known to affect Adpn levels

Inhibition of Adpn production and associated insulin resistance have been linked to catecholamines such as isoproterenol 1.6; a recent study by Delporte et al.\(^45\) revealed a 30% decrease in Adpn mRNA after culturing in human visceral tissue with isoproterenol for 8 h, Figure 1.8. Isoproterenol is commonly used to treat bradycardia and heart block.\(^46\)

![Functional group of the catecholamines](Image)

**Figure 1.8** Chemical structure of isoproterenol 1.6, which has been associated with Adpn inhibition.

1.5.4 Expression of adiponectin using recombinant techniques

Expression of Adpn using recombinant methods has proven to be problematic, due to the non-homogenous nature of the recombinant product and subsequent separation of the different oligomers within this mix. To date, the most successful recombinant technique to produce an Adpn biological therapeutic has centred on a trimeric structure of the globular domain of Adpn (gAdpn) linked using a GAS (Gly-Ala-Ser, denoted in study as sc) linker sequence to a crystallisable fragment of human immunoglobulin G1 (Fc) to afford Fe-sc-gAdpn, Figure 1.9.\(^47\)

The Fc fragment was attached to enhance the pharmokinetic profile of the complex. Fe-sc-gAdpn was expressed using CHO (Chinese hamster ovary) cell lines. The cell medium was concentrated and filtered...
prior to capture of the protein using a Protein A FF column (GE Healthcare), subsequent elution using ImmunoPure Gentle Ag/Ab Elution Buffer (Pierce) and purification using high-performance liquid chromatography (HPLC). The purified protein, Fc-sc-gAdpn, was found to have an excellent plasma half-life of 2 weeks compared to the non-Fc linked protein sc-dAdpn which had a much shorter half-life of 2 h. Fc-sc-gAdpn was also found to be bioactive, and able to reduce blood glucose levels in ob/ob mice models, Figure 1.9.

**Figure 1.9** Trimer globular adiponectin therapeutic produced by Yang et al.\textsuperscript{47}

### 1.5.5 The total chemical synthesis of adiponectin

As demonstrated in the above Section, recombinant synthesis of the globular domain of Adpn has been successfully reported by Ge et al.\textsuperscript{47}. However, recombinant synthesis of the collagenous domain of Adpn is not possible due to the PTM-Lys residues present in this section of the protein. Therefore, the total chemical synthesis of the collagenous domain has been attempted.

In 2014 Nishiuchi et al.\textsuperscript{48} reported the first total chemical synthesis of the collagenous domain of Adpn. Their synthetic approach involved chemical synthesis of the PTM-Lys residues and incorporation of the PTM-Lys residues into the Adpn sequence using Fmoc- solid phase peptide synthesis (Fmoc-SPPS) to afford four different PTM-Adpn segments of the collagenous domain of Adpn corresponding to residues 19 - 36, 37 - 63, 64 – 87 and 89 – 107. Conversion of these PTM-Adpn segments into Boc- functionalised peptides followed by a series of native chemical ligation reactions, afforded the glycopeptide of Adpn corresponding to residues 19 – 107. This Adpn-glycopeptide was shown by Ohkubo et al.\textsuperscript{49} to be able to self-order and form HWM oligomers, similar to those observed in native, full length Adpn. Although impressive, this chemical synthesis of the collagenous domain of Adpn afforded the final Adpn-glycopeptide in less than 0.12% overall yield. Therefore, this chemical synthesis is not commercially viable and so could not be used to raise blood plasma Adpn levels therapeutically.

### 1.5.6 Synthesis of small molecule therapeutics to stimulate adiponectin production

There have been promising leads in the field of small molecule pharmaceuticals which target either the Adpn receptors or the key signalling partners of Adpn. However, due to the different receptors and different tissue types targeted by Adpn, the identification of a pan-receptor target has been difficult. The
most successful small Adpn receptor agonist developed to date is AdipoRon 1.7, published by Kadowaki et al. in 2013, Figure 1.10.

![AdipoRon 1.7](image)

**Figure 1.10** AdipoRon 1.7 small molecular therapeutic developed by Kadowaki et al. Stereochemistry of the compound was not reported.

AdipoRon was shown to activate AMPK via both the AdipoR1 and AdipoR2 receptors in a dose-dependent manner and also has the additional advantage of being able to be administered orally. Although introduction of AdipoRon to high-fat-diet-induced mice was not able to reduce the obesity of the mice, significant reduction in blood plasma sugar levels and improved insulin sensitivity was observed, making AdipoRon a promising lead therapeutic.

### 1.5.7 Targeting of a suitable intracellular signalling pathway

Due to the large number of pathways associated with Adpn bioactivity, identification of a suitable intracellular signalling molecular target has proved challenging. The AMPK pathway, which has been repeatedly shown to be up-regulated by increased Adpn levels, is also involved in many other metabolic processes. Therefore, a therapeutic which targets the AMPK pathway could have adverse side-effects for the patient. Due to the complex nature of the interactions of Adpn in vivo, the focus of Adpn based therapeutics remains based on the synthesis of the protein itself.

### 1.6 Aims of this PhD project

Due to the significant interest in the use of Adpn as a possible therapeutic for T2D, the overall aim of this PhD project was to increase understanding as to how the structural properties of Adpn relate to the protein’s anti-diabetic properties in vivo. As discussed above in Section 1.4, post translationally modified lysine residues may play a crucial role in the formation of bioactive HMW oligomers of Adpn. Therefore, the scope of this doctoral project was limited to an investigation of the structural role of the post-translationally modified lysine residues of Adpn.

The post-translational modifications of the lysine residues of native mammalian Adpn have been shown to be hydroxylation followed by glycosylation. However, native mammalian Adpn has been shown to exist in multiple glycosylated isoforms, namely a mixture of Glu-Gal-Hyl 1.8 and Gal-Hyl 1.9,
Figure 1.11. For simplicity, this PhD project focused on Adpn isoforms bearing either singly glycosylated Gal-Lys residues 1.9 or the neoglycoside analogues neoGal-Lys 1.10 and neoGal-Ser 1.11, Figure 1.11.

![Chemical structures](image)

**Figure 1.11** The chemical structures of the native Adpn glycosides Glu-Gal-Hyl 1.8, Gal-Hyl 1.9 and neoglycoside analogues neoGal-Lys 1.10 and neoGal-Ser 1.11.

These glycosylated and neoglycosylated residues will be incorporated into an 18-amino acid segment of residues 66 – 83 of the collagenous domain of Adpn, where three of the four native glycosylated lysine residues are located, Figure 1.12. An 18-mer peptide was chosen as an initial synthetic target as short peptide sequences are easily synthesised in high yields using solid-phase peptide synthesis (SPPS).

Furthermore, extension of the chosen 18-mer peptide sequence (Adpn 66-83) to include the fourth PTM-Lys residue at position 104, would have required addition of a further 23 amino acids, thus necessitating the synthesis of a second peptide fragment and ligation of the two peptide fragments using native chemical ligation methods thus reducing the overall yield of the synthetic Adpn neoglycopeptide.
Figure 1.12 The post-translation modifications of the four lysine residues found in the collagenous domain of murine-Adpn and the 18-mer peptide chosen for this PhD project.\(^{28}\)

The secondary structure of the resulting neoglycopeptides will then be examined using circular dichroism at physiological pH, with the aim of elucidating the structural role of the glycosylated lysine residues in the formation of the triple helical secondary structure of the collagenous domain of Adpn.

1.7 Outline of this doctoral thesis

The aims outlined above will be discussed in turn in the following Chapters. In addition, a comprehensive introduction to the relevant literature will be given at the start of each Chapter.

**Chapter Two**: Towards the synthesis of native, glycosylated hydroxylysine.

**Chapter Three**: The synthesis of neoglycopeptides corresponding to the collagenous domain of adiponectin.

**Chapter Four**: The use of collagen model peptides as structural scaffolds for Adpn.

**Chapter Five**: Lysine, the keystone of the triple helical structure of Adpn.

**Chapter Six**: Summary and conclusions of this doctoral study.

**Chapter Seven**: Experimental methods and product characterisation.
Chapter Two

Towards the Synthesis of Native Glycosylated Hydroxylysine
Chapter Two: Towards the Synthesis of Native Glycosylated Hydroxylysine

2.1 The chemical structure of native, glycosylated hydroxylysine

As previously discussed in Section 1.5, the key post-translational modifications (PTMs) present in human adiponectin are hydroxylation and subsequent glycosylation of four key lysine residues located at positions 68, 71, 80 and 104 of the collagenous domain of adiponectin (Adpn). These PTMs have been shown to be crucially important for the oligomerisation and subsequent bioactivity of Adpn. The chemical structure of the PTM-lysine residues present in collagen type peptides, such as Adpn, was elucidated by Spiro et al. in 1967 to be the disaccharide unit 2-O-α-D-glucopyranosyl-D-galactose bound to (2S,5R)-hydroxylysine by a β-glycosidic bond, Figure 2.1 Glu-Gal-Hyl 1.8.53,54

![Chemical structures](image)

Figure 2.1 The chemical structure of main compounds discussed in this Chapter namely (2S,5R)-hydroxylysine (2S,5R)-Hyl 2.1, galactosylated hydroxylysine Gal-Hyl 1.8 and glucosylgalactosyl hydroxylysine Glu-Gal-Hyl 1.9.

In this Chapter, two different post-translational modifications will be considered in turn. Firstly, the initial hydroxylation of lysine to afford (2S,5R)-hydroxylysine, (2S,5R)-Hyl, 2.1 will be examined. The subsequent glycosylation to produce initially the monosaccharide Gal-Hyl 1.8 followed by further glycosylation to afford Glu-Gal-Hyl 1.9 will then be explored. Both chemical and biological methods for generating these PTM-lysine building blocks will be discussed.

It should be noted that hydroxylysine is commercially available. However, the price of the enantiopure (2S,5R)-Hyl 2.1 is prohibitively expensive for use in academic research; $780, 1 g, Sigma Aldrich. The price for the single enantiomer is very high, as it must be either synthesised chemically or purified from an epimeric mixture using lengthy physical separation techniques such as fractional crystallisation. In addition, although Glu-Gal-Hyl 1.8 has been previously synthesised by Allevi et al.58,59, Kihlbeg et al.60 and Nishiuchi et al.48, these published methods are long, complex and ultimately low yielding. These synthetic methods are discussed in detail in Section 2.5 and an overview is given below in Table 2.1.
Table 2.1 Overview of methods for synthesis of (2S,5R)-Hyl 2.1 reported to date.

<table>
<thead>
<tr>
<th>Author</th>
<th>Commercially available starting materials</th>
<th>Number of synthetic steps to yield protected Glu-Gal-Hyl 1.8</th>
<th>Overall yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allevi et al.</td>
<td>L-Glutamic acid, pentabenzylgalactose, tetrabenzylglucose</td>
<td>21</td>
<td>0.6%</td>
</tr>
<tr>
<td>Kihlberg et al.</td>
<td>Commercial Hyl 2.1, tri-O-acetyl-D-glucal, ethyl1-thio-β-D-glucopyranoside</td>
<td>10, but started with commercial Hyl 2.1</td>
<td>5%</td>
</tr>
<tr>
<td>Nishiuchi et al.</td>
<td>Synthetic details of Hyl 2.1 and sugar donors not reported</td>
<td>Not reported</td>
<td>Addition Gal- 61% yield Addition Glu- 47% yield</td>
</tr>
</tbody>
</table>

Furthermore, whilst the use of enzymatic techniques for the production of post-translationally modified proteins such as Adpn has been explored by Myllyla et al.,\textsuperscript{61,62} the overall yields of these methods are low and therefore ultimately not commercially viable. Therefore, due to both the high cost of commercial (2S,5R)-Hyl 2.1 and current low yielding synthetic methods, the task of producing multigram quantities of both (2S,5R)-Hyl 2.1 and Glu-Gal-Hyl 1.8 presents an opportunity for development of a better chemical and/or biological synthetic pathway.

The aim of this Chapter is to explore potentially shorter and higher yielding methods for the synthesis of both (2S,5R)-Hyl 2.1 and Gal-Glu-Hyl 1.8. The chemical synthesis will focus on installation of the (2S)-stereocentre through use of the Belokon complex\textsuperscript{63,64} and the (5R)-hydroxyl functionality through either Sharpless aminohydroxylation\textsuperscript{65,66} or asymmetric epoxidation\textsuperscript{67} reactions. Various glycosylation methods will be probed to optimise conditions for the selective synthesis of the β-glycosidic linkage. Enzymatic methods, similar to those employed by Myllyla et al.,\textsuperscript{61,62} will also be investigated for production of the desired Glu-Gal-Hyl 1.8.

### 2.2 Historical isolation of Hyl and Gal-Glu-Hyl

Hydroxylation of lysine residues in proteins is an important post-translational modification. This is well illustrated by the variety of genetic diseases which occur when this modification is perturbed e.g. Ehlers-Danlos syndrome whereby the patient exhibits connective tissue dysfunction. The degree of hydroxylation of lysine resides within bone tissue has been shown to govern the cross-linking of α-helical peptide chains, and thereby the tertiary and quaternary structure of the collagen.\textsuperscript{68} Crucial evidence for the key role of Hyl 2.1 in cross-linking of peptide chains can be seen within the bone tissue of patients with Ehlers-Danlos syndrome, which contains only 17% of the normal hydroxyllysine content.\textsuperscript{68}
As a naturally occurring compound, Hyl 2.1 can be isolated from native collagens. In 1941, Van Slyke et al. first introduced the technique of hydrolysis of collagens using phosphotungstic acid to afford a mixture of diamino acids. Dosage of the reaction mixture with periodic acid followed by measurement of the liberated ammonia facilitated estimation of the amount of Hyl 2.1 within the diamino acid mixture. However, although this method enabled accurate quantification of the Hyl content of the collagen, the Hyl contained in the diamino mixture was destroyed by the experiment, thus this method is not a viable option to afford Hyl. Later work by Napoli et al. demonstrated the isolation of Gal-Hyl 1.9 and Glu-Gal-Hyl 1.8 from both human placental collagen and bovine anterior lens collagen. The extraction of the desired glycoproteins was conducted by initial hydrolysis of the lyophilised collagens followed by gel filtration to isolate both Gal-Hyl 1.9 and Glu-Gal-Hyl 1.8. Ion-paired reverse phase high performance liquid chromatography (RP-HPLC) was then employed to separate and isolate both Gal-Hyl 1.9 and Glu-Gal-Hyl 1.8 from the reaction mixture. Although this isolation technique appears promising in its ability to afford pure Gal-Hyl 1.9 and Glu-Gal-Hyl 1.8, unfortunately as no product yields were reported the method is assumed to be low yielding.

### 2.3 Previous synthesis of (2S, 5R)-Hyl, 2.1

#### 2.3.1 The different approaches employed to date

There have been six different approaches to the chemical synthesis of Hyl 2.1 reported to date. These different methods can be categorised as: an epimeric synthesis, an epimeric synthesis followed by silica chromatographic separation of diastereoisomers, an epimeric synthesis using a chiral resolving agent, a stereoselective synthesis using a chiral auxiliary, a stereoselective synthesis using a chiral catalyst and a substrate-controlled stereoselective synthesis. These different methods will be discussed in detail below.

##### 2.3.1.1 Epimeric synthesis

The first synthesis of Hyl was an epimeric synthesis reported by Sheehan et al. in 1950. In this synthesis epi-Hyl 2.6 was synthesised from L-glutamic acid 2.4 via a methyl α,ε-diphthalimido-δ-keto-DL-caproate intermediate 2.5, Scheme 2.1. However, as this synthetic pathway was non-stereoselective and low yielding with 9% overall yield, it will not be further considered in this thesis.
Reagents and conditions: a) Potassium phthalimide (1.05 eq.), DMF, rt, 30 min; b) Al(OiPr)₃/iPrOH (1.8 eq.), 80 h, mild heating; b) HCl (6 N, 1 eq.), 10 h, reflux.

Scheme 2.1 Synthesis of epimeric Hyl by Sheehan et al.⁷¹

2.3.1.2 Epimeric synthesis followed by separation using silica chromatography

Subsequent epimeric syntheses have made use of silica chromatography to separate the diastereomers formed at the C-5 stereocentre. The use of flash column chromatography on silica gel as a purification technique in these syntheses, has enabled chemists to produce diastereomerically pure samples of the desired (2S,5R)-Hyl 2.1 stereoisomer.

A good example of this strategy was published by Allevi et al.⁵⁸ who reported the synthesis of both (2S,5R)-Hyl 2.1 and (2S,5S)-Hyl 2.7 in which the key diastereomeric lactone intermediates (2.8 and 2.9) were synthesised via a known diazo ketone 2.1₀⁷⁷ derived from L-glutamic acid 2.4. These lactone intermediates were identified as having different polarities by normal phase-HPLC and then separated using flash silica column chromatography (fcc) to afford two different diastereoisomers, Scheme 2.2. This route afforded protected derivatives of both isomers of 5-Hyl but unfortunately was low yielding for the desired (2S,5R) Hyl 2.1 product with only 7.6% overall yield reported. The main drawback for this synthetic protocol is the low-yielding lactonisation step, Scheme 2.2 step e).
Reagents and conditions: a) NaHCO₃ (2.2 eq.), MeOH, reflux, 10 mins; b) i) HBr 33%, AcOH, THF, -5 °C, until pH = 1 by thymol blue indicator violet solution colouration, ii) NaHCO₃ (sat. soln.), until pH = 7; c) NaN₃ (1.4 eq.), DMF, rt, 1 h, d) NaBH₄ (1.3 eq.), MeOH, 0 °C, 20 min; e) TFA (6.5 eq.), rt, 40 min; f) α-chymotrypsin, phosphate buffer (0.1 M, pH 7.4), Me₂CO, 25 °C, 12 h; g) and i) H₂, Pd/C, MeOH: H₂O (1:1, v/v) then pH 6.7 – 7.0; h) Cs₂CO₃ (0.27 M, 2 eq.), MeOH: H₂O (1:1 v/v), rt, 35 min.

Scheme 2.2 Use of rapid silica chromatography to separate lactones 2.8 and 2.9 in order to isolate the desired product (2S,5R)-Hyl 2.1 as demonstrated by Allevi et al.⁵⁸

2.3.1.3 Epimeric synthesis with use of a chiral resolving agent for separation of diastereoisomers

Adamczyk et al.⁷² utilised a thyroxine-derived chiral resolving agent to enable separation of diastereoisomers within their synthesis of (2S,5R)-Hyl 2.1, Scheme 2.3. Their approach entailed elongation and conversion of the side-chain of Boc-protected glutamic acid 2.15 to the 1,2-bromoalcohol 2.17, followed by esterification using a thyroxine derivative as a chiral resolving agent, Scheme 2.3, step e). The resulting diastereoisomers were then successfully separated by preparative scale reverse phase high performance liquid chromatography (RP-HPLC). The desired (2S,5R) diastereoisomer 2.18 was then heated with sodium azide, followed by base-catalysed hydrolysis of the thyroxine ester group. The resulting azide 2.19 was then reduced via catalytic hydrogenation using Pd/C catalyst and the Boc protecting group subsequently removed by treatment with TFA to afford (2S,5R)-Hyl 2.1, Scheme 2.3. Although a good enantiomeric excess (>97% ee for 2.1) was achieved for step e), the yield for this step was low (28%), resulting in a low overall yield of 5% for this seven step synthesis.
Chapter Two

Reagents and conditions: a) Isobutyl chloroformate (1.0 eq.), 4-methylmorpholine (1.1 eq.), THF, 0 °C, 25 min; b) CH$_2$N$_2$ (10 eq.), Et$_2$O, 0 °C, 10 min, rt; c) 48% HBr (1.0 eq.), ether, -20 °C, 25 min; d) NaBH$_4$ (1.0 eq.), MeOH, 0 °C, 3 h; e) O-methoxy-N-tertbutoxycarbonyl-L-thyroxine (2.20, 1.0 eq.), DCC (2.0 eq.), DMAP (cat.), CH$_2$Cl$_2$, rt, 24 h then RP-HPLC separation; f) i) NaN$_3$, DMF, 80 °C, 6 h; ii) LiOH, THF:H$_2$O, rt, 4 h; g) i) Pd/C, H$_2$, EtOH, 4.5 h; ii) TFA:H$_2$O, rt, 2 h. Further synthetic details not available.

Scheme 2.3: Use of the thyroxine auxiliary 2.20 as a chiral auxiliary for the synthesis of (2S,5R)-Hyl 2.1.

2.3.1.4 Epimeric synthesis of epi-Hyl utilising a chiral auxiliary

Kihlberg et al.$^{73}$ reported an interesting synthesis of a (2S,5R)-Hyl derivative 2.31 from D-malic acid 2.21 using a chiral auxiliary developed by Williams et al.$^{78}$ to control the key alkylation step of the synthesis, Scheme 2.4, step i). Alkylation of iodide 2.29 with Williams’ chiral glycine reagent 2.32 afforded the desired single diastereoisomer 2.30. Subsequent reduction of the azide functionality using triphenylphosphine, followed by Boc-protection and hydrogenolysis of the chiral auxiliary and subsequent Fmoc protection lead to (2S,5R)-Hyl derivate 2.31. This synthesis required 11 linear steps and afforded the optically pure (2S,5R)-Hyl derivate 2.31 in 32% overall yield.
Reagents and conditions: a) BH$_3$·SMe$_2$ (2 M, 3.2 eq.), B(OMe)$_3$ (3.4 eq.), THF, 0 °C → rt, 16 h; b) p-methoxybenzaldehyde (1.5 eq.), pyridinium toluene-4-sulfonate (0.01 eq.), CH$_2$Cl$_2$, reflux, o/n; c) NaCNBH$_3$ (2 eq.), TBDMSCl (2 eq.), CH$_3$CN, rt, 30 min; d) NaN$_3$ (5.0 eq.), DMF, 55 °C, 24 h; e) MsCl (1.5 eq.), Et$_3$N (1.5 eq.), DMAP (0.1 eq.), CH$_2$Cl$_2$, 0 °C, 10 min; f) DDQ (1.2 eq.), CH$_3$CN:H$_2$O (8:1, v/v), rt, 20 min; g) NaI (2 eq.), acetone, reflux, 1 h; i) PPh$_3$ (1.5 eq.), THF/H$_2$O (ratio not reported), microwave, 130 °C, 5 min, ii) Boc$_2$O (1.5 eq.), THF/H$_2$O (ratio not reported), 1:1, rt, o/n; j) Fmoc-OSu (1.04 eq.), Na$_2$CO$_3$ (10% aq. soln, 8 eq.), acetone/ H$_2$O, rt, o/n.

Scheme 2.4: Use of Williams’ chiral auxiliary 2.32 to guide alkylation of iodide intermediate 2.26 in the synthesis of (2S,5R)-Hyl derivate 2.31.

Allevi et al.$^{59}$ have also reported a convenient 4 step synthesis of a zwitterionic (2S,5R)-Hyl derivative 2.37 utilising the Williams’ glycine template. However, this synthesis was also found to be low yielding, with an overall yield 9.5% over four steps, Scheme 2.5a). Their synthesis hinged upon the stereoselective alkylation of the chiral template with homoallyl iodide 2.33 after which epoxidation of the olefin took place to afford the epimeric epoxide 2.35. Azide induced ring opening of the epimeric epoxide 2.35 and final hydrogenation of azido intermediate 2.36 afforded the zwitterionic (2S,5R)-Hyl derivative 2.37. Van den Nieuwendij et al.$^{79}$ have also used Williams’ glycine template to affect stereoselective alkylation of the iodide 2.39 in an excellent yield of 80%. Subsequent hydrolysis to remove the template afforded intermediate 2.41 also in an excellent yield of 84%. Acidification of intermediate 2.41 then afforded the
desired unprotected compound (2S,5R)-Hyl 2.1 in a good overall yield of 21% over eight steps, Scheme 2.5b).

\[
\text{desired unprotected compound (2S,5R)-Hyl 2.1 in a good overall yield of 21\% over eight steps, Scheme 2.5b).}
\]

Reagents and conditions: a) \((\text{Me}_3\text{Si})_2\text{NLi} (0.3 \text{ eq.}), \text{William's Boc chiral glycine template 2.38 (0.2 eq.)}, \text{THF:HMPA (10:1, v/v)}, -78 ^\circ\text{C} \rightarrow -40 ^\circ\text{C}, 4 \text{ h}); b) m-CPBA (1.2 eq.), 1,2-dichloroethane:phosphate buffer, pH = 7, rt, 12 h; c) Na\text{N}_3 (2.8 eq.), Na\text{HCl} (0.75 eq.), MeO\text{CH}_2\text{CH}_2\text{OH}, 50 ^\circ\text{C}, 4.5 \text{ h}; d) i) TFA:H\text{O} (v/v, 95:5, excess), rt, 1 h, ii) H\text{O}, Pd/C (10\%), H\text{O}:\text{MeOH:HCl (8:3:1)}, 25 ^\circ\text{C}, 24 \text{ h}; e) Na\text{HMDS (1.25 eq.)}, \text{William's chiral glycine template 2.32 (1.25 eq.)}, 15\text{-crown-ether (7 eq.)}, \text{THF, -70 ^\circ\text{C} \rightarrow -35 ^\circ\text{C}, 3 \text{ h}}; f) H\text{O}, Pd/C (10\%), \text{THF:MeOH (v/v, 1:3)}, \text{rt, 24 h}; g) 1.5 \text{M DCl (13 eq.)}, \text{rt, 2 h}.

\text{Scheme 2.5 Further examples of use of William's chiral glycine template as a chiral auxiliary for the synthesis of (2S,5R)-Hyl derivatives as reported by a) Allevi et al.}\text{59 and b) Van den Nieuwendij et al.79}

\text{2.3.2 Asymmetric synthesis using chiral catalysis}

The first stereoselective synthesis of a Hyl derivative was published by Kunz et al.\text{74 in 1999. The synthesis started from D-valine 2.42 and L-glycine and used the Schöllkopf method\text{80 to form the \alpha-stereogenic centre observed in intermediate 2.44 via selective alkylation of the cyclic dipeptide on the} }
si-face. An asymmetric Sharpless aminohydroxylation\(^8\) of the resulting terminal olefin 2.47 then yielded the (2S,5R)-Hyl derivative 2.48, Scheme 2.6.

Kunz et al.\(^74\) investigated the use of both (DHQ)\(_2\)PHAL and (DHQD)\(_2\)PHAL as chiral ligands with chloroamine T, chloroamine B and N-halocarbamate salts as the nitrogen source for the Sharpless aminohydroxylation reaction. Optimised reaction conditions used (DHQD)\(_2\)PHAL and chloroamine T for the Sharpless aminohydroxylation step (step f), Scheme 2.6), which afforded the (2S,5R)-Hyl derivative 2.48 in a good 67% yield but unfortunately with a low diastereomeric excess (d.e. 29%).

\begin{equation}
\text{Reagent and conditions: a) Gly-OEt, DIC/HOBt; b) i) TFA, ii) reflux (toluene); c) (Et}_3\text{O})^+ (BF}_4^-; d) 4-bromobutene, }^\text{BuLi; e) i) HCl, ii) Boc}_2\text{O; f) chloroamine T, (DHQD)_2PHAL, K}_2\text{OsO}_2(\text{OH})_2. (further synthetic details not reported).}
\end{equation}

\textbf{Scheme 2.6} Use of the Schöllkopf method\(^8\) to furnish the (2S) stereogenic centre of Hyl, followed by an asymmetric Sharpless aminohydroxylation\(^8\) to install the (5R) stereocentre of Hyl, as reported by Kunz et al.\(^74\)

\textbf{2.3.3 Substrate directed asymmetric synthesis: use of piperidinone intermediates}

Piperidinone derivatives of Meldrum’s acid and aspartic acid are also useful intermediates for the synthesis of (2S,5R)-Hyl 2.1 as demonstrated by Guichard et al.\(^75,76\), Scheme 2.7. The key step in this synthesis is the stereoselective oxidation of the chiral piperidinone intermediate 2.50. Use of an alkaline metal base such as LiHMDS induces chelation control over the C2 and C3 substituents by blocking the upper face of the generated enolate from electrophilic attack.

Thus, upon addition of an electrophilic oxidising agent such as (+)-CSO (10-camphorsulfonyloxaziridine), the lower face of the enolate is attacked preferentially. Guichard et al.\(^76\) reported an excellent d.e. of over 80% for the key oxidation step (Scheme 2.7, step e) with many different electrophilic oxidising agents, including MoOPH (MoO\(_5\)·HMPA·pyridine), PPO (\textit{trans}-2-(phenylsulfonyl)-3-phenyloxaziridine) and (+)-CSO. Of these oxidising agents, (+)-CSO 2.55 was found to be the most effective oxidising agent with a reported yield of 92% and d.e. of 98%.

Following addition of (+)-CSO (2.55) to the piperidinone 2.50 a reductive ring opening using NaBH\(_4\) takes place, affording the desired 5R-α-hydroxy-δ-lactam 2.51 selectively in a high yield of 92%. This
induced stereochemistry translates into the desired (2S,5R) stereochemistry of the final product, affording protected (2S,5R)-Hyl 2.54 in an excellent overall yield of 55% over nine steps.

\[
\begin{align*}
\text{Boc-Asp-OBu} & \quad 2.49 \\
\text{Piperidinone} & \quad 2.50 \\
\text{a) Meldrum's acid (1 eq.), EDC (1.5 eq.), DMAP (1.5 eq.), CH}_2\text{Cl}_2, \text{0 °C → rt, 3 h; b) NaBH}_4 (3.0 eq.), CH}_2\text{Cl}_2, \text{AcOH (10% eq. soln.), rt, 48 h; c) toluene, reflux, o/n; d) LiHMDS (1 N in THF, 1 eq.), THF, -78 °C, 2.5 h; e) (+)-CSO (2.56, 1.5 eq.), -78 °C, 16 h; f) NaBH}_4 (5 eq.), EtOH, 0 °C, 4 h; g) collidine (10 eq.), MsCl (1.1 eq.), CH}_2\text{Cl}_2, \text{0 °C, 20 h; h) NaN}_3 (2 eq.), DMF, 80 °C, 8 h; i) H}_2, \text{Pd/C (10%), EtOH, rt, 4 h.}
\end{align*}
\]

Reagents and conditions: a) Meldrum’s acid (1 eq.), EDC (1.5 eq.), DMAP (1.5 eq.), CH₂Cl₂, 0 °C → rt, 3 h; b) NaBH₄ (3.0 eq.), CH₂Cl₂, AcOH (10% eq. soln.), rt, 48 h; c) toluene, reflux, o/n; d) LiHMDS (1 N in THF, 1 eq.), THF, -78 °C, 2.5 h; e) (+)-CSO (2.56, 1.5 eq.), -78 °C, 16 h; f) NaBH₄ (5 eq.), EtOH, 0 °C, 4 h; g) collidine (10 eq.), MsCl (1.1 eq.), CH₂Cl₂, 0 °C, 20 h; h) NaN₃ (2 eq.), DMF, 80 °C, 8 h; i) H₂, Pd/C (10%), EtOH, rt, 4 h.

Scheme 2.7 Use of stereoselective oxidation of the chiral piperidinone intermediate 2.50 using (+)-CSO, 2.55, to install desired (5R)-stereocentre of desired final product protected (2S,5R)-Hyl 2.54 reported by Guichard et al.⁷⁶

In summary, early syntheses of hydroxylsine were epimeric in nature, with the different diastereomers being separated through physical methods such as flash silica chromatography. Chiral resolving agents such as thyroxine derivatives have also been used to further enable the physical separation of the resulting diastereomeric pairs. Later, more sophisticated syntheses have made use of asymmetric synthesis techniques to furnish the desired (2S,5R)-Hyl 2.1. These techniques included: the use of a chiral auxiliary (Williams’ glycine template)⁷⁸, asymmetric catalysts (Sharpless aminohydroxylation)⁷⁴,⁸¹ and substrate controlled stereoselectivity (chelation controlled enolate formation).⁷⁵,⁷⁶

The most stereoselective approach to the synthesis of (2S,5R)-Hyl 2.1 reported to date is the substrate-directed stereoselective approach published by Guichard et al.⁷⁵,⁷⁶ (stereoselective oxidation of a chiral piperidinone derived enolate). This synthesis produced Boc/Bu protected (2S,5R)-Hyl 2.54 in the highest overall yield of 55% with a reported d.e. of 98%.

2.4 Previous syntheses of Gal-Hyl 1.9

The glycosylation of hydroxylated lysine plays a fundamental role in collagen proteins present in the human body such as Type II collagen,⁸² a key component of joint cartilage, and adiponectin, a key regulator of blood sugar levels.⁴⁹ Kihlberg et al.⁶⁰,⁸³ have shown that removal of the disaccharide
moieties from lysine residues causes the T-cells within the human body to be unable to recognise Type II collagen. Furthermore, Cooper et al.\textsuperscript{28} and others\textsuperscript{49} have shown that glycosylation of lysine residues of adiponectin are crucial for bioactivity of the protein.

As mentioned previously, Spiro et al.\textsuperscript{53} elucidated the structure of the disaccharide moiety bound to Hyl in many collagens to be Glu-Gal-Hyl \textbf{1.8}. Since this structural report, there has been great interest in the synthesis of the Glu-Gal-Hyl building block \textbf{1.8} and its subsequent incorporation into peptide synthesis. In this section, the previous syntheses of Gal-Hyl \textbf{1.9} will be discussed, which will be subsequently be followed by a discussion of the synthesis of Glu-Gal-Hyl \textbf{1.8} in Section 2.5. One of the challenges in the synthesis of Gal-Hyl \textbf{1.9} is the need to form the $\beta$-O-galactosyl bond in preference to the anomerically favoured $\alpha$-linkage. Different groups have responded to this challenge in a variety of ways, which are outlined below.

### 2.4.1 Use of an $\alpha$-1,2-anhydro sugar as the galactose sugar donor

Kihlberg et al.\textsuperscript{84} reported the coupling of protected Hyl \textbf{2.56} with $\alpha$-1,2-anhydro sugar \textbf{2.57} under zinc chloride promoted coupling conditions resulting in protected Gal-Hyl derivative \textbf{2.58} in a reasonable yield of 30\%, Scheme 2.8.

![Scheme 2.8 Use of $\alpha$-1,2-anhydro sugar \textbf{2.57} for glycosylation of (2S,5R)-Hyl derivative \textbf{2.56} to afford Gal-Hyl derivative \textbf{2.58} as reported by Kihlberg et al.\textsuperscript{84}](image)

**Reagents and conditions:** a) i) molecular sieves (AW-300, excess), rt, 30 min, ii) ZnCl$_2$ (1 M in Et$_2$O, 1.5 eq.), THF, -50 °C → rt, 19 h.

**Scheme 2.8** Use of $\alpha$-1,2-anhydro sugar \textbf{2.57} for glycosylation of (2S,5R)-Hyl derivative \textbf{2.56} to afford Gal-Hyl derivative \textbf{2.58} as reported by Kihlberg et al.\textsuperscript{84}

### 2.4.2 Use of an acetyl bromide as the galactose sugar donor

Later work by Kihlberg et al.\textsuperscript{85} utilised silver silicate to promote coupling of acetobromogalactose \textbf{2.60} with an $N^\varepsilon$-Cbz protected Hyl derivate \textbf{2.59} to afford Gal-Hyl derivative \textbf{2.61} in an excellent reported yield of 82\%, Scheme 2.9.
Reagents and conditions: a) acetobromogalactose (2.60, 1.5 eq.), silver silicate (Ag₂O₃Si, 5 eq.), molecular sieves (3 Å, excess), CH₂Cl₂, 0 °C, 2 h.

Scheme 2.9 Use of acetylbromide galactose donor 2.60 for the synthesis of Gal-Hyl derivative 2.61, as reported by Kihlberg et al.⁸⁵

Fields et al.⁸⁶ have reported the use of copper complexation of the N- and C-termini of the amino group alongside Boc-protection of the ε-amine group of Hyl, as a useful protecting group strategy for the galactosylation of Hyl. In their synthesis protected Hyl 2.62 is coupled with acetobromogalactose 2.60 using sodium hydride and dry acetonitrile to afford protected Gal-Hyl in an excellent yield of 77% yield, Scheme 2.10. Subsequent deprotection of the amino group using Chelex acidic beads, followed by Fmoc protection of the amino group afforded Gal-Hyl 2.63 in 29% overall yield from protected Hyl 2.62. Gal-Hyl derivative 2.63 was then used in solid phase peptide synthesis (SPPS) to product the glycopeptide corresponding to residues 1263-1277 of type IV collagen.

Reagents and conditions: a) D-acetobromogalactose (2.60, 2.5 eq.), NaH (2.5 eq.), CH₃CN, reflux, 2 h then rt o/n, b) i) Chelex 100 (H⁺ form, excess), H₂O:CH₃OH (6:5, v/v), 20 °C, 3 h, ii) Fmoc-OSu (1.1 eq.), NaHCO₃ (1.14 eq.), H₂O: acetone (1:1), 30 °C, o/h.

Scheme 2.10 Use of copper chelation to protect the (2S)-amino functionality of (2S,5R)-Hyl derivative 2.62 during glycosylation, as reported by Fields et al.⁸⁶

Adamczyk et al.⁸⁷ have also coupled a Boc-protected (2S,5R)-Hyl derivative to acetobromogalactose 2.60 using mercury cyanide promoted glycosylation conditions. Aside from being environmentally unfriendly, this method was also found to be low yielding (22%).
Reagents and conditions: a) Acetobromogalactose (2.60, 2 eq.), Hg(CN)$_2$ (2.5 eq.), toluene, 75 °C, 24 h.

Scheme 2.11 Use of d-acetobromogalactose donor 2.60 for glycosylation of (2S,5R)-Hyl derivative 2.64 to afford Gal-Hyl derivative 2.65, as reported by Adamczyk et al.87

Guichard et al.76 have performed galactosylation of Hyl 2.66 utilising the bulkier tetra-pivaloylated galactosyl bromide 2.67 as a glycosylation partner to afford Gal-Hyl derivative 2.69. In their study, they found the orthoester side-product 2.68 was still the dominant glycosylation product afforded in 79% yield, but gratifyingly 2.68 could easily be rearranged using trimethylsilyl trifluoromethanesulfonate (TMSOTf) to the desired Gal-Hyl derivative 2.69 in an excellent 76% yield, Scheme 2.12. Guichard et al.82 later reported use of their Gal-Hyl derivative 2.69 for the synthesis of the glycopeptide corresponding to residues 256-270 of bovine type II collagen using SPPS.

Reagents and conditions: a) Tetra-pivaloylated galactosyl bromide (2.67, 0.1 M in CH$_2$Cl$_2$, 1.5 eq.), molecular sieves (4 Å, excess), silver silicate (5 eq.), in the absence of light, Ar, rt, 8 h; b) TMSOTf (0.1 M in CH$_2$Cl$_2$, 0.1 eq.), molecular sieves (4 Å, excess), 0 → 4 °C, o/n.

Scheme 2.12 Use of tetra-pivaloylated galactosyl donor 2.67 for glycosylation of (2S,5R)-Hyl derivative 2.66 to afford Gal-Hyl derivative 2.69, as reported by Guichard et al.76

2.4.3 Use of a trichloroacetimidate as the galactose sugar donor

Allevi et al.88 have reported the use of a trichloroacetimidate 2.71 as the sugar donor for the glycosylated of Hyl derivative 2.70. The coupling was found to proceed in an average yield affording Gal-Hyl derivative 2.72 in 52%. Notably, in this synthesis, the C2-hydroxyl group is selectively protected as a chloroacetyl group in readiness for a second glycosylation at the C2 position (see Section 2.5.1).
Reagents and conditions: a) tBuMe₂SiSO₃CF₃, molecular sieves (3 Å, excess), Et₂O, 25 °C, 1 h (further synthetic details not reported).

Scheme 2.13 Use of trichloroacetimidate donor 2.71 to glycosylate Hyl derivative 2.70, as reported by Allevi et al.⁸⁸

In a recent publication, Nishiuchi et al.⁴⁸ detailed the total chemical synthesis of the variable and collagenous domain of Adpn (residues 19-107). This reported synthesis included a similar synthetic method to that reported by Allevi et al.⁸⁸ used to synthesise Gal-Hyl derivative 2.72. The two key differences between the methods employed by Allevi et al.⁸⁸ and Nishiuchi et al.⁴⁸ were the temperature at which the glycosylation was carried out and the glycosylation catalyst used. Nishiuchi et al.⁴⁸ conducted their glycosylation of Hyl derivative 2.73 using the trichloroacetimidate galactose donor 2.71 at the lower temperature of -42 °C using trimethylsilyl triflate (TMSOTf) as the glycosylation catalyst in preference to tert-butyldimethylsilyl triflate employed by Allevi et al.⁸⁸. Using the lower temperature with the more active catalyst afforded a slightly improved yield of Gal-Hyl derivative 2.74 of 61%, Scheme 2.14.

Reagent and conditions: a) TMSOTf, molecular sieves (3 Å, excess), -42 °C, CH₂Cl₂, 61% (further synthetic details not reported).

Scheme 2.14 Glycosylation conditions reported by Nishiuchi et al.⁴⁸ to afford Gal-Hyl derivative 2.74.
2.5 Previous syntheses of Glu-Gal-Hyl 1.8

In native Adpn, the disaccharide unit 2-O-α-D-glucopyranosyl-D-galactose is often found linked to the hydroxylysine residues in the collagenous domain,\(^{28}\) giving rise to the need to synthesize Glu-Gal-Hyl 1.8. Selective protection of the C2 hydroxyl group of the galactose donor is necessary for this synthesis, as is the need to construct an α-glycosidic linkage between the Glu and Gal moieties. The different synthetic strategies employed to date for the synthesis of Glu-Gal-Hyl 1.8 are outlined below.

2.5.1 Use of a trichloroacetimidate as the glucose sugar donor

As mentioned previously in Section 2.4.3, Allevi et al.\(^{88}\) reported a high yielding chemical synthesis of Glu-Gal-Hyl in 2004. Following the aforementioned initial glycosylation which afforded protected Gal-Hyl 2.72 in a good yield of 52%, selective deprotection of the C2-chloroacetyl group using zinc acetate proceeded in an excellent yield of 93%. The second glycosylation then took place with α-glucosyltrichloroacetimidate donor 2.75 to afford desired protected Glu-Gal-Hyl 2.76 in a good yield of 42%, Scheme 2.15.

\[\text{Reagents and conditions: a) Zn(OAc)}_2\cdot2\text{H}_2\text{O, MeOH, reflux, 9 h, 93\% yield; b) } \text{BuMe}_2\text{SiOTf, molecular sieves (3 Å, excess), Et}_2\text{O, 25 °C, 15 min, 45\% yield (further synthetic details not reported).}\]

\text{Scheme 2.15 Glycosylation strategy employed by Allevi et al.}^{88} \text{ for the synthesis of protected Glu-Gal-Hyl derivative 2.76.}

2.5.2 Use of an activated thioethyl glycoside as the glucose sugar donor

As mentioned previously in Section 2.4.1, Kihlberg et al.\(^{84}\) have previously reported the synthesis of protected Gal-Hyl derivative 2.58 via an α-1,2-anhydro galactose donor. However, glycosylation of protected Gal-Hyl derivative 2.58 using the α-glycoside donor 2.77 and N-iodosuccinimide/silver triflate as promoting agents, to afford protected Glu-Gal-Hyl derivative 2.78 was unsuccessful, despite the glycosylation of a structurally similar model Gal-norvaline 2.79 proceeding using the same conditions in an excellent yield of 71% to afford Glu-Gal-norvaline derivative 2.80, Scheme 2.16.
Reagents and conditions: a) \( N \)-iodosuccinimide, AgOTf, CH\(_2\)Cl\(_2\), molecular sieves (4 Å, excess), -45 \( \rightarrow \) -15 °C (further synthetic details not reported).

Scheme 2.16 The attempted synthesis of Glu-Gal-Hyl derivative 2.78 as reported by Kihlberg et al.\(^84\)

As the acid lability of the \( N \)-\( \epsilon \) Boc protecting group of protected Gal-Hyl derivative 2.58 was noted by Kihlberg et al.\(^84\) as the key problem for the glycosylation, subsequent change of the \( N \)-\( \epsilon \) protecting group to the acid stable Cbz group in Gal-Hyl derivative 2.81 resulted in successful glycosylation. Protected Gal-Hyl derivative 2.81 was glycosylated using the thioethyl glycoside 2.82 to afford Glu-Gal-Hyl derivative 2.83 in a reasonable yield of 47% and reasonable overall yield of 14%, Scheme 2.17. The protected Glu-Gal-Hyl derivative 2.83 was then used in SPPS to afford the glycopeptide corresponding to residues 256-270 of type II collagen in 23% yield.

Reagents and conditions: a) Thioethyl glycoside 2.82 (1.2 eq.), NIS (1.2 eq.), AgOTf (0.5 eq.), CH\(_2\)Cl\(_2\), molecular sieves (3 Å, excess), -45 °C, protected from light, 80 min.

Scheme 2.17 The successful synthesis of Glu-Gal-Hyl derivative 2.83 as reported by Kihlberg et al.\(^84\)

As mentioned above in Section 2.4.3, the recent report by Nishiuchi et al.\(^48\) detailing the total chemical synthesis of the variable and collagenous domain of Adpn (residues 19-107) contains the synthesis of Gal-Hyl derivative 2.74. In their publication, Nishiuchi et al.\(^48\) report the synthesis of Glu-Gal-Hyl
derivative 2.86 by selective deprotection of the C2-chloroacetyl protecting group of Gal-Hyl derivative 2.74 using hydrazinedithiocarbonate (HDTC) followed by a second glycosylation reaction using thioglycoside donor 2.85 (as opposed to using a second trichloroacetimidate favoured by Allevi et al.)88, Scheme 2.18. Subsequent removal of the allyl protecting group ready for SPPS afforded the desired protected Glu-Gal-Hyl derivative 2.86 in 52% yield from Gal-Hyl derivative 2.74; the highest reported yield to date.

Reagents and conditions: a) Hydrazinedithiocarbonate (HDTC), pH 7; b) NIS, TMSOTf, DMF; c) Pd(PPh₃)₄, morpholine (further synthetic details not reported).

Scheme 2.18 Glycosylation conditions reported by Nishiuchi et al.48 to afford Glu-Gal-Hyl derivative 2.86.

2.6 Summary of reported syntheses of Gal-Hyl 1.9 and Glu-Gal-Hyl 1.8 to date

There has been significant interest in different methods used to affect glycosylation of (2S,5R)-Hyl 2.1 to afford Glu-Gal-Hyl 1.884-87,89 The key challenge for these glycosylations are to effect a β-O-glycoside link to the galactosyl residue followed by an α-O-glycoside link to the subsequent glycosyl residue. The highest yielding method to produce a protected derivative of Glu-Gal-Hyl 1.8 was published in 2014 by Nishiuchi et al.48 This method used a thioglycoside with C2-selectively protected with a chloroacetyl group for the first glycosylation reaction and a benzyl protected glycosyltrichloroacetimidate for the second glycosylation, affording protected Glu-Gal-Hyl derivative 2.86 in 29% overall yield.
2.7 Selected synthetic strategy towards (2S,5R)-Hyl 2.1 utilised in this PhD thesis

Based on the above literature overview, the method reported by Guichard et al.\textsuperscript{75} utilising stereoselective oxidation of a chiral piperidinone derived enolate to synthesise (2S,5R)-Hyl 2.54 was chosen as this method afforded Hyl derivative in an excellent overall yield of 55% over nine steps.

2.7.1 Synthesis of (2S,5R)-Hyl 2.1 utilising selective oxidation of a chiral piperidinone enolate

Initial attempts to synthesise (2S,5R)-Hyl 2.1 followed the reported synthesis by Guichard et al.,\textsuperscript{75} utilising stereoselective oxidation of a chiral piperidinone derived enolate to furnish the key (5R)-stereocentre (see Section 2.3.3 above). However, during this doctoral study, the key reduction step of ketone intermediate 2.87 (Scheme 2.19, step b) afforded a crude reaction mixture, which \textsuperscript{1}H NMR indicated contained the partially reduced β-hydroxyester intermediate 2.88 in preference to the desired triester intermediate 2.89. The \textsuperscript{1}H NMR spectrum of the crude reaction mixture following step b) contained a new broad doublet resonance corresponding to one proton, observed at δ 5.23, \(J = 7.50\) Hz (400 MHz), which was felt to correspond to CH-OH proton of β-hydroxyester intermediate 2.88. Intriguingly, in the COSY spectrum recorded for the crude reaction mixture, coupling of the new broad doublet was only observed to the adjacent CH of the lactam ring, and not to the adjacent CH\(_2\) group, Figure 2.2.

![Scheme 2.19](image)

\textit{Reagents and conditions:} a) Meldrum's acid (1 eq.), EDC (1.5 eq.), DMAP (1.5 eq.), CH\(_2\)Cl\(_2\), rt, 2 h (crude product carried forwards without further purification); b) NaBH\(_4\) (3 eq.), CH\(_2\)Cl\(_2\): AcOH (9:1, v/v), rt, o/n.

\textbf{Scheme 2.19} Incomplete reduction of the ketone intermediate 2.87 afforded undesired β-hydroxyester side product 2.88 in preference to the desired ester product 2.89.
Despite further addition of the reductant (further 3 eq. of NaBH₄) and extended reaction time (72 h at room temperature), the reduction failed to go to completion. Use of alternate mild nucleophilic reductants such as NaBH(OAc)₃ and NaBH₃CN (up to 6 eq.) were also investigated to effect reduction at the ketone site preferentially over the two ester sites, but unfortunately, they were also unsuccessful in generating the desired triester product. Harsher reductants such as LiAlH₄, diborane and the ‘superhydride’ LiBHEt₃ were not trialed for this reduction as these reagents would reduce both the ketone and ester sites.

2.8 Revised strategy; use of Belokon complex and Sharpless aminohydroxylation reaction for the synthesis of (2S,5R)-Hyl

Due to the unsuccessful replication of the synthesis published by Guichard et al.⁷⁵ in this PhD thesis, a revised novel synthetic pathway to prepare protected (2S,5R)-Hyl derivative 2.90 was devised, Scheme 2.20. This retrosynthetic strategy utilises the Sharpless aminohydroxylation reaction⁶₆,⁹₀ to install the (5R) stereocentre of (2S,5R)-Hyl derivative 2.90 from protected (2S)-hex-5-enoic acid derivative 2.91. The protected (2S)-hex-5-enoic acid derivative 2.91 is in turn disconnected to the stereoselectively alkylated nickel complex 2.92, easily accessed from the Belokon complex⁹¹ 2.93 and bromobutene.
**Scheme 2.20** Retrosynthetic strategy devised for the synthesis of (2S,5R)-Hyl derivative 2.90.

### 2.8.1 Introduction to the use of the Belokon complex for the stereoselective synthesis of amino acids

The first stage of this revised strategy requires the synthesis and stereoselective alkylation of Belokon complex 2.93. This popular method of preparing stereogenic amino acids was first published by Belokon et al.\(^{63}\) in 1985 and utilises the nickel(II) Schiff base complex of (S)-O-[(N-benzylpropyl)amino]benzophenone and glycine 2.93 which acts as a chiral auxiliary for alkylation of the complex. The subsequent hydrolysis of (2S)-alkylated nickel complex 2.92 using either aqueous HCl or acetic acid was shown to afford stereogenic amino acids, as reported by Belokon et al.\(^{91}\) in 1988.

Belokon et al.\(^{91}\) reported alkaline reaction conditions for the alkylation of Belokon complex 2.93 using sodium hydroxide as the base, a range of alkyl halides as alkylating agents in aprotic solvents such as DMF and acetonitrile. These alkylation conditions resulted in 70 - 95% e.e. for the alkylation of the Belokon complex 2.93 with a range of alkyl halides including MeI (to afford alanine 2.94, 70% e.e.), PhCH\(_2\)Br (to afford phenylalanine 2.95 (86% e.e.) and Me\(_2\)CHBr (to afford valine 2.96, 92% e.e.) in excellent yields of over 75%, Scheme 2.21.
**Reagents and conditions:** a) Alkyl halide (1 eq.), NaOH (2.5 eq.), DMF, 10 - 30 min, rt; b) aqueous AcOH (0.1 mmol).

**Scheme 2.21** Alkylation conditions of Belokon complex 2.93 as reported by Belokon et al.\(^91\) to afford alanine 2.94, phenylalanine 2.95 and valine 2.96.

To prevent bisalkylation, the ratio of alkyl halide: Belokon complex 2.93 was set between 1.0 - 1.5. The concentration of the reagents was also found to be of importance for overall reaction time, with increased concentrations of both (20 mmol Belokon complex 2.93, 20 mmol alkyl halide, 15 mL DMF) rendering the alkylation complete in 30 minutes, as opposed to observed reaction times of 5 – 7 hours for lower concentration of reagents. Furthermore, the resulting bright-red alkylated complexes were reported to be easily purified using silica chromatography.

Since these seminal papers, Belokon complex 2.93 has been widely adopted as a reliable method for the facile asymmetric synthesis of amino acids. Belokon et al.\(^92,93\) have developed a robust, large scale synthesis of the chiral auxiliary for the Belokon complex, \((2S)-[N-'benzylprolyl]amino]benzophenone (BPB, 2.98, Scheme 2.22), which is easily recovered upon hydrolysis of the alkylated material. Nádvorník et al.\(^94\) have used this method to construct novel PET diagnostics. Soloshonok et al.\(^95\) have published novel alternative chiral auxiliaries for the complex including \((2S)-[N-(N'-'benzylprolyl)amino]-4-methylbenzophenone (4-Me-BPB), which have been shown to be cheaper to produce and more scalable than the traditional BPB ligands. Brimble et al.\(^64\) have also reported success utilising Belokon complex 2.93 for the synthesis of methyl-N-Boc-(2S,4R)-4-methylpipecolate. Furthermore, Brimble et al.\(^96\) have also developed a novel general method for the thesis of chiral α-amino acids using Belokon complex 2.93 in a Mitsunobu-Tsunoda reaction.
2.8.2 The use of the Belokon complex as a chiral auxiliary for the synthesis of protected (2S)-hex-5-enoic acid

Given the widely reported successful use of Belokon complex 2.93 as a suitable chiral auxiliary for the synthesis of stereogenic amino acids, this method was hence chosen and Belokon complex 2.93 was synthesised in a good yield of 55% over three steps, Scheme 2.22. Alkylation of Belokon complex 2.93 with 4-bromo-1-butene proceeded well affording the (2S)-alkylated complex 2.92 in an excellent yield of 98% with an excellent d.e. of 92%. (2S)-Hex-5-enoic acid 2.91 was then released in quantitative yield from the (2S)-alkylated complex 2.92 by hydrolysis using HCl in methanol, Scheme 2.22. The optical activity of the (2S)-hex-5-enoic acid 2.91 synthesised was found to be in accordance with the literature data reported by Gu et al.97, \([\alpha]_D^{22} +15.7, c 0.14, H_2O, (lit +13.1, c 1.30, H_2O)\)97.

Reagents and conditions: a) KOH (3.9 eq.), iPrOH, 50 °C, 6 h; b) MeSO_2Cl (1 eq.), N-methylimidazole (4 eq.), 2-aminobenzophenone (0.90 eq.), CH_2Cl_2, 50 °C, 24 h, c) Ni(NO_3)_2.6H_2O ( 2 eq.), glycine (5 eq.), KOH (3 eq.), MeOH, N_2, reflux, 2 h; d) 4-bromobut-1-ene (0.98 eq.), KOH (3 eq.), NaI (0.3 eq.), CH_3CN, N_2, rt, 2 h; e) i) 2 M HCl (excess), MeOH, reflux, 1 h, ii) aq. NH_3, purification using a Dowex H\(^+\) column.

Scheme 2.22 Use of the Belokon complex 2.93 towards the synthesis of (2S,5R)-Hyl.

2.8.2.1 The key alkylation reaction of Belokon complex 2.93

The alkylation reaction of Belokon complex 2.93 and 4-bromo-1-butene was extensively investigated in order to optimise the reaction conditions to those reported above. The different conditions investigated are listed below in Table 2.1. The percent (2S) of the alkylated complex 2.92 was determined by the ratio of the integrals of (2S): (2R) peaks in the \(^1\)H NMR spectrum (found between δ 8.0 – 8.5 ppm) of the alkylated product. The desired (2S)-alkylated complex 2.92 was isolated from the product mix using fractional recrystallisation (CH_2Cl_2: n-hexane), Figure 2.3.
Reagents and conditions: see Table 2.2 below.

**Scheme 2.23** Reaction scheme for alkylation of Belokon complex 2.93 with 4-bromo-1-butene.

**Table 2.2** Different reaction conditions trialled for alkylation of Belokon complex 2.93 with 4-bromo-1-butene (0.98 eq.). All reactions were conducted at room temperature (c. 25 °C).

<table>
<thead>
<tr>
<th>Entry</th>
<th>Scale (g)</th>
<th>Solvent</th>
<th>Reaction Time</th>
<th>Base (eq.)</th>
<th>Nal (eq.)</th>
<th>Yield of 2.92 (both 2R and 2S)</th>
<th>Percent (2S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3</td>
<td>CH₃CN</td>
<td>2 h</td>
<td>NaOH, (2.5)</td>
<td>0.0</td>
<td>42%</td>
<td>96%</td>
</tr>
<tr>
<td>2</td>
<td>0.1</td>
<td>CH₃CN</td>
<td>2 h</td>
<td>NaOH, (2.5)</td>
<td>0.3</td>
<td>20%</td>
<td>92%</td>
</tr>
<tr>
<td>3</td>
<td>0.5</td>
<td>CH₃CN</td>
<td>3 h</td>
<td>NaOH, (2.5)</td>
<td>0.3</td>
<td>94%</td>
<td>89%</td>
</tr>
<tr>
<td>4</td>
<td>1.0</td>
<td>CH₃CN</td>
<td>2 h</td>
<td>KOH, (3.0)</td>
<td>0.3</td>
<td>98%</td>
<td>92%</td>
</tr>
<tr>
<td>5</td>
<td>2.0</td>
<td>CH₃CN</td>
<td>2 h</td>
<td>KOH, (2.5)</td>
<td>0.3</td>
<td>72%</td>
<td>86%</td>
</tr>
<tr>
<td>6</td>
<td>4.0</td>
<td>CH₃CN</td>
<td>3 h</td>
<td>NaOH, (2.5)</td>
<td>0.3</td>
<td>92%</td>
<td>95%</td>
</tr>
<tr>
<td>7</td>
<td>0.1</td>
<td>DMF</td>
<td>5 min</td>
<td>NaOH, (10)</td>
<td>0.0</td>
<td>28%</td>
<td>92%</td>
</tr>
<tr>
<td>8</td>
<td>0.5</td>
<td>DMF</td>
<td>5 min</td>
<td>NaOH, (10)</td>
<td>0.3</td>
<td>97%</td>
<td>82%</td>
</tr>
</tbody>
</table>
All of the alkylation reactions were run at room temperature, as previous literature studies have shown alkylation of Belokon complex to be most effective at room temperature, with elevated temperatures leading to decomposition of the Belokon complex and lower temperatures resulting in lower diastereoselectivity of the alkylation reaction. Both sodium hydroxide and potassium hydroxide were evaluated as bases for the alkylation reaction. Both bases were found to be effective with yields obtained for both in excess of 90%.

Acetonitrile and DMF (dimethylformamide) were used as solvents for the alkylation reaction, according to the methods reported by Brimble et al. and Hruby et al. respectively. For this system acetonitrile was found to be a better solvent than DMF, probably due to the fast reaction in DMF resulting in loss of stereoselectivity of the alkylated product. Addition of sodium iodide was found to increase reaction yields significantly in both solvents, due to the in situ formation of the more reactive alkyl iodide. In DMF the increase in reaction yield upon addition of sodium iodide was very marked (from 28% to 97%) although the d.e. of the alkylated complex was lower (92% to 82%), Table 2.2, Entry 7 and 8. In acetonitrile, the addition of sodium iodide gratifyingly resulted in consistently higher yields of the desired alkylated product with excellent d.e. (89 – 95%), Table 2.2, Entries 3 – 6.

The most successful conditions were found to be use of potassium hydroxide as the base (3 eq.), sodium iodide (0.3 eq.) as the catalyst, resulting in a reaction time of 2 h, 98% yield and 93% d.e. (Table 2.2, Entry 4).

The (2S)-alkylated Belokon complex was hydrolysed by heating methanoic solution of 2.92 with 2M HCl for one hour. Subsequent purification using Dowex H⁺ ion-exchange column afforded desired
(2S)-hex-5-enoic acid 2.91 in quantitative yield. The optical activity of the (2S)-hex-5-enoic acid 2.91 synthesised was found to be in accordance with the literature data reported by Hruby et al.,97 \([\alpha]_D^{22} +15.7, c \ 0.14, \ H_2O, \ (\text{lit} \ +13.1, \ c \ 1.30, \ H_2O)\).97

Reagents and conditions: i) 2 M HCl (excess), MeOH, reflux, 1 h, ii) NH₃ (aq.), purification using Dowex H⁺ column.

**Scheme 2.24** Hydrolysis of (2S)-alkylated Belokon complex 2.92 to afford (2S)-hex-5-enoic acid 2.91.

### 2.8.3 Use of the Sharpless aminohydroxylation reaction for furnishing of the (5R)-stereocentre of Hyl

With the (2S)-stereocentre in hand, in the form of protected (2S)-hex-5-enoic acid 2.91, focus then shifted to introduction of the desired (5R) stereocentre into protected (2S,5R)-Hyl 2.90. The Sharpless aminohydroxylation (SAH) reaction was an attractive choice for the introduction of the (5R)-stereocentre, as the aminohydroxylation reaction results in simultaneous syn addition of an amino and alcohol group across the olefin; both of which are functionalities required in the desired (2S,5R)-Hyl 2.1, Scheme 2.25a). Furthermore, the stereoselectivity of the SAH can be controlled using chiral Cinchona alkaloids such as DHQ (dihydroquinine) and DHQD (dihydroquinidine), both of which are commercially available. A similar strategy has been reported by Kunz et al.74 with Boc-protected (2S)-hex-5-enoic acid using potassium osmate(VI) hydrate as the oxidant and choroamine-T as the nitrogen source, Scheme 2.25b). It is worth noting, that although a reasonable yield of 67% was reported, the d.e. for the reaction performed by Kunz et al.74 was unacceptably low at 29%.

**Scheme 2.25** a) Proposed approach for the introduction of the (5R)-stereocenter of protected (2S,5R)-Hyl 2.90 using SAH reaction and b) similar reported synthesis by Kunz et al.74
2.8.3.1 Introduction to the Sharpless aminohydroxylation reaction

The first example of an oxyamination reaction to afford a 1,2-amino-alcohol was published by Sharpless et al. in 1975. This publication detailed the non-catalytic reaction of olefins (such as cyclohexane, styrene and 1-decene) using OsO₃(N-Bu) to afford 1,2-amino-alcohols in excellent yields (74 - 85%). In all cases, the amino functionality was found to be bound selectively at the least hindered carbon, providing a basis for regioselectivity for the reaction.

Over the years, other nitrogen sources have been used in the Sharpless aminohydroxylation reaction (SAH) including chloroamine-T, N-chlorosodiocarbamates such as and most recently alkyl 4-chlorobenzyloxycarbamates such as Figure 2.4. Commercially available chloroamine-T remains a popular choice for the SAH, although the harsh conditions required for deprotection of the resulting tosyl-protected amino-alcohols (reductive, Li/NH₃) or strongly acidic conditions (HBr/AcOH) makes this nitrogen source unsuitable for many organic compounds.

The introduction of carbamates such as N-chlorosodiocarbamates such as which contain the more easily cleavable protecting groups such as Cbz, Boc and Troc within the nitrogen source, have opened up the scope of the SAH to a wider selection of olefins. However, the need to re-oxidise the N-chlorosodiocarbamates in situ (typically using tert-butyl hypochlorite and sodium hydroxide) which generates undesired chlorinated side-products, as well as the inherent instability of N-chloro compounds has led to the development of additional new nitrogen sources for the SAH.

Alkyl 4-chlorobenzyloxycarbamates such as benzyl 4-chlorobenzyloxycarbamate have been recently reported by Luxemburger et al. as novel nitrogen sources for the SAH. These alkyl-4-chlorobenzyloxycarbamates also have the advantage of being compatible with base-free media, thus widening the scope of the SAH reaction to substrates containing base-labile protecting groups such as the Fmoc amino protecting group.

![Figure 2.4 Common nitrogen sources used in the Sharpless aminohydroxylation reaction.](image)

In this doctoral thesis investigations of the SAH reaction of protected (2S)-hex-5-enoic acid followed the method developed by Luxenburger et al., as this method is compatible with the Fmoc protecting group on the amine functionality needed for incorporation of protected (2S)-hex-5-enoic acid derivative into the adiponectin sequence via SPPS.
**Reagents and conditions:** a) Carbamate 2.103 (1.4 eq.), \( \text{OsO}_4/\text{tBuOH} \) (4% mol), chiral ligand (5% mol), \( \text{CH}_3\text{CN/H}_2\text{O} \) (8:1, v/v), rt, 3 h.

**Scheme 2.26** General reaction scheme for the Sharpless aminohydroxylation reaction of protected (2S)-hex-5-enoic acid derivatives 2.91.

Benzyl 4-chlorobenzoyloxy carbamate 2.103 was chosen as the nitrogen source for the SAH reaction. The following mechanism was proposed by Luxenburger et al.\(^99\) for the formation of catalytic imido-trioxo osmium intermediate 2.104. Initial attack of the carbamate 2.101 onto osmium tetraoxide and subsequent proton transfer results in loss of a water molecule and formation of trioxo osmium intermediate 2.103. The lone pair now available on the osmium centre of intermediate 2.103 is then able to form an imide bond with the carbamate to afford imido-trioxo osmium intermediate 2.104, Scheme 2.27.

**Scheme 2.27** Proposed mechanism for formation of imido-trioxo osmium intermediate 2.106 using benzyl 4-chlorobenzoyloxy carbamate 2.103 as the nitrogen source, as proposed by Luxenburger et al.\(^99\)

### 2.8.3.2 Proposed reaction mechanism for the Sharpless aminohydroxylation reaction

The proposed mechanism for the SAH reaction\(^101\) involves a fast primary cycle affording the desired Sharpless AH product in high e.e. and a slow secondary cycle which yields product with low e.e, as shown in Scheme 2.28.
As outlined above in Section 2.8.3.1, initial complexation of osmium tetroxide with the chosen nitrogen source leads to imido-trioxoosmium intermediate 2.106. Addition of the ligand (usually a Cinchona alkaloid ligand) followed by addition of the alkene (see Section 2.8.3.3) results in osmium (VI) azaglycolate 2.107. Oxidation of 2.107 by second addition of nitrogen source and release of the ligand results in complex 2.108. At this point in the mechanism competing primary and secondary cycles come into play. In the faster primary cycle, hydrolysis of complex 2.108 results in release of the desired aminohydroxylated product in high e.e. and regeneration of imido-trioxoosmium intermediate 2.106. Alternatively, in the slower secondary cycle, addition of a second alkene molecule without directing influence of a chiral ligand results in osmium (VI) bis(azaglycolate) 2.109. Hydrolysis of 2.109 results in formation of an aminohydroxylated product with much lower e.e. In order to promote the primary pathway and thus afford SAH products with high e.e., SAH reactions are often run in 50% water: solvent mixes to increase the rate of hydrolysis of complex 2.106 and thus promote the faster primary cycle.

2.8.3.3 Mode of alkene addition for the Sharpless aminohydroxylation reaction

The exact nature of alkene addition to imido-trioxoosmium intermediate 2.106 has not been well-defined to date. As addition of a chiral ligand leads to both stereochemical and regiochemical control of the final product (see Section 2.8.3.4), it is thought that addition of the ligand to imido-trioxoosmium intermediate 2.106 takes place prior to addition of the alkene substrate. There are two possible modes of addition for the alkene; either a [3+2] cycloaddition or a [2+2] cycloaddition followed by a ring expansion migration to afford osmium (VI) azaglycolate 2.107, Scheme 2.29. Recent quantum mechanical calculations show that the activation energy for the stepwise [2+2] pathway is very high (>39 kcal/mol), whereas the concerted [3+2] pathway is much lower (<10 kcal/mol). Therefore, the concerted [3+2] mechanism of alkene addition is most likely to occur.
Scheme 2.29 Possible modes of attack for the alkene with imido-trioxoosmium intermediate 2.106.

2.8.3.4 Use of chiral cinchona alkaloids to control both the regioselectivity and stereoselectivity of the Sharpless aminohydroxylation reaction

Cinchona alkaloids such as (DHQ)$_2$PHAL (1,4-bis(9-O-dihydroquinine) phthalazine) 2.110, (DHQD)$_2$PHAL (1,4-bis(9-O-dihydroquinidine) phthalazine) 2.111 and (DHQ)$_2$AQN (1,4-bis(9-O-dihydroquinine) anthraquinone 2.112 have been shown to be useful ligands for the SAH reaction, Figure 2.5. These chiral ligands are able to increase the reaction rate and guide both the stereochemical and regiochemical outcome of the SAH reaction.\textsuperscript{102} The addition of Cinchona ligands to the SAH reaction has the advantage of supressing formation of the undesired diol side-product.

![Chemical structures of the main Cinchona ligands used to control regio- and stereoselectivity in the SAH reaction.](image)

Figure 2.5 Chemical structures of the main Cinchona ligands used to control regio- and stereoselectivity in the SAH reaction.

In 1996, Sharpless et al.\textsuperscript{102} demonstrated the remarkable ability of these Cinchona ligands to induce facial selectivity to the SAH of methyl cinnamate 2.113. When (DHQD)$_2$PHAL 2.111 was added to the
reaction mix, 71% e.e. was observed favouring the β-product \textbf{2.114} (upper face), whereas when (DHQ)$_2$PHAL \textbf{2.110} was added the α-product \textbf{2.115} (lower face) was favoured with 81% e.e., Scheme 2.30. The overall yield for the reaction was also reasonable (66%), and the enantiomers were separable by recrystallisation from methanol.

Scheme 2.30 Induced facial selectivity of SAH of methyl cinnamate \textbf{2.113} using Cinchona ligands.$^{102}$

Cinchona alkaloid ligands have also been shown to guide the regioselectivity of the SAH reaction. In general terms the regioselectivity of the SAH reaction is guided by a slight preference for the amino functionality of the SAH product to be placed at the least hindered carbon of the olefin.$^{98}$ This is likely to be due to steric considerations, due to more sterically encumbered substitution present on the protected amino group (e.g. NHTs as compared to free OH group). Therefore, the substituted bulkier amino group will be preferentially added to the least sterically hindered end of the alkene.

O’Brien \textit{et al.}$^{107}$ investigated the effect of altering the chiral ligand upon the regioselectivity of the SAH of styrene \textbf{2.116}. They found that switching the chiral ligand from (DHQ)$_2$PHAL \textbf{2.110} to its diastereomeric partner (DHQD)$_2$PHAL \textbf{2.111} reduced the regioselectivity of the reaction with the ratio of the major regioisomer \textbf{2.117} reduced from 80% to 64% (determined by $^1$H NMR), Scheme 2.31. They also explored different alkene substrates and found that changing the substrate structure had little effect on the regioselectivity of the SAH reaction.

Scheme 2.31 Reduced regioselectivity of the SAH reaction of styrene upon switching of chiral ligand from (DHQ)$_2$PHAL \textbf{2.110} to (DHQD)$_2$PHAL \textbf{2.111}, as demonstrated by O’Brien \textit{et al.}$^{107}$
Sharpless et al.\textsuperscript{108} have also demonstrated the ability of the chiral ligand to affect the regiochemistry of SAH reactions. During their investigations on the SAH of cinnamates 2.119, they found that changing the ‘spacer’ of the ligand could reverse the regioselectivity of the reaction. Remarkably, by switching from (DHQ)\textsubscript{2}PHAL 2.110\textsuperscript{109} to (DHQ)\textsubscript{2}AQN 2.112,\textsuperscript{108} the regioselectivity and stereoselectivity of the reaction could be reversed, Scheme 2.32. When (DHQ)\textsubscript{2}PHAL 2.110 was used, SAH product 2.120 was favoured in a 20:1 ratio, whereas when (DHQ)\textsubscript{2}AQN 2.112 was employed, SAH product 2.121 was favoured instead in a 79:21 ratio (ratios determined by \textsuperscript{1}H NMR). The reasons for the dramatic change in regioselectivity and stereoselectivity are not known, but are hypothesised to involve the substrate binding to different pockets of the chiral ligands.\textsuperscript{108}

\chemstart
\begin{align*}
\text{Ar-}C=CH_2 &\xrightarrow{\text{a)} } \text{Ar-CH=C(O)Me} + \text{Ar-CH=C(O)Me} \\
2.119, R = \text{Ac or Me} &\quad 2.120, R = \text{Ac, } R' = \text{Pr} \\
&\quad \text{Favoured with (DHQ)\textsubscript{2}PHAL} \\
&\quad 20.1, 81\% \text{ yield} \\
2.121, R = \text{Me, } R' = \text{Cbz} &\quad \text{Favoured with (DHQ)\textsubscript{2}AQN} \\
&\quad 79.21, 58\% \text{ yield}
\end{align*}
\chemend

\textbf{Scheme 2.32} Switching of regioselectivity of the SAH reaction of cinnamates by changing the chiral ligand from (DHQ)\textsubscript{2}PHAL 2.110\textsuperscript{109} to (DHQ)\textsubscript{2}AQN 2.112\textsuperscript{108}

Furthermore, Bodkin et al.\textsuperscript{110} have explored the factors governing regioselectivity during the SAH reaction, in particular the reversal in regiochemistry observed when switching from the PHAL to the AQN spacer in the chiral cinchona ligand family. By using space-filling models they were able to illustrate that the additional steric constraints imposed by the larger AQN spacer could result in a change in binding orientation of the olefin substrate and thus a change in the regioselectivity of the SAH reaction.
2.8.3.5 Use of tethered substrates to improve the regioselectivity of the Sharpless aminohydroxylation reaction

Donohoe et al.\textsuperscript{111,112} have devised a novel approach to improve the regioselectivity of the SAH reaction by utilising ‘tethered’ substrates for the SAH, whereby the nitrogen source is attached to the starting material and therefore delivered in an intramolecular manner during the course of the reaction. This approach allows for complete control of the regiochemistry of the SAH product, such as the tethered SAH reaction of compound 2.122 to produce SAH product 2.123 in an excellent 80\% yield, Scheme 2.33.

\[
\text{Reagents and conditions: } \text{K}_2\text{OsO}_2(\text{OH})_4 (1 \text{ mol\%}), \text{tBuOH/H}_2\text{O} (3:1, \text{v/v}), \text{rt, 24 h}
\]

\textbf{Scheme 2.33} Use of a tethered substrate in the SAH reaction.\textsuperscript{111,112}

However, additional synthetic steps are necessary to produce the ‘tethered’ substrate and strongly basic conditions are needed for cleavage of the carbamate nitrogen source from the SAH product, so this approach is not compatible with substrates containing base-labile protecting groups (such as Fmoc protection of the amino group). In addition, the ‘tethered’ substrates do not respond well to chiral ligands, often leading to lower levels of enantioselectivity.

2.8.3.6 Conditions investigated for the SAH reaction of model starting materials and (2S)-hex-5-enoic acid

As mentioned previously, the intention to incorporate synthesised Gal-Hyl derivatives into peptides of the Adpn sequence using Fmoc-SPPS (thus requiring the amino acids to be protected with the base-labile Fmoc protecting group), directed this study using the method reported by Luxenburger et al.\textsuperscript{99} that used base-free conditions for the SAH reaction. The achiral SAH conditions using benzyl 4-chlorobenzoyloxycarbamate 2.103 as the nitrogen source were first tested on model substrate styrene 2.116 on a 50 mg scale and proceeded well resulting in an excellent 83\% yield of the desired AH product, Scheme 2.34. \textsuperscript{1}H NMR analysis revealed a 10:1 mix of regioisomers, 2.124 and 2.125 with the terminal NHCbz regioisomer 2.124 (\textit{rac}-benzyl 2-hydroxy-2-phenylethylcarbamate) being the dominant product.
Scheme 2.34  Test achiral SAH reaction carried out by author using styrene 2.116 as a model substrate and the base-free reaction conditions reported by Luxenburger et al.99

Further trial reactions using the chiral ligand (DHQ)$_2$PHAL 2.110 were conducted successfully on cyclohexene 2.126 and methyl cinnamate 2.113 with reasonable (47%) and good (89%) yields obtained respectively, Scheme 2.35. In the case of cyclohexene 2.123, as the reaction substrate is symmetrical, there were no issues of regiochemistry to be resolved. The regiochemistry of the major product produced from the SAH of methyl cinnamate 2.111 was determined by $^1$H analysis. The optical activity of the SAH products 2.124 and 2.125 were compared to literature values to determine the enantiomeric excess (e.e.) of the SAH products.

Reagents and conditions: a) Carbamate 2.103 (1.4 eq.), OsO$_4$/BuOH (4 mol%), (DHQ)$_2$PHAL 2.110 (5 mol %), CH$_3$CN/H$_2$O (8:1, v/v), rt, 3 h

Scheme 2.35  Trialled chiral SAH reaction conditions published by Luxenburger et al.99 conducted on cyclohexane 2.126 and methyl cinnamate 2.113.

Following on from these model reactions, both achiral and chiral methods were investigated for the SAH reaction of Fmoc/OMe ester derivative of (2S)-hex-5-enoic acid 2.128, such as those shown in
Scheme 2.36. Disappointingly, in all cases, it was found that chiral reaction conditions using cinchona alkaloid ligands such as (DHQ)$_2$PHAL 2.110 resulted in poor to low reaction yields of 20 - 41% of the desired SAH product 2.129. Under these chiral conditions the major products isolated were the side products benzyl carbamate 2.130 and the diol 2.131. The regio stereochemistry of the SAH product 2.129 is discussed later in this section.

Reagents and conditions: a) Carbamate 2.103 (1.4 eq.), OsO$_4$/BuOH (4 mol%), (DHQ)$_2$PHAL 2.110 (5 mol%), CH$_3$CN/H$_2$O (8:1, v/v), rt, 3 h.

Scheme 2.36 Trialled chiral SAH reaction conditions with Fmoc/OMe-ester derivative of (2S)-hex-5-enoic acid 2.128 affording the desired AH product 2.129 alongside benzyl carbamate 2.130 and diol 2.131 side products.

A number of different protecting group strategies were trialled in order to optimise the reaction conditions. An acetonitrile/water mix (8:1) was chosen as the reaction solvent in preference to the traditional tert-butyl alcohol/water mix, as acetonitrile afforded a more homogeneous reaction mixture. Both OsO$_4$/BuOH and the less toxic and more environmentally friendly osmium reagent K$_2$OsO$_4$.2H$_2$O were trialled as catalysts for the SAH.

Unfortunately, K$_2$OsO$_4$.2H$_2$O was not successful in affording the desired aminohydroxylation product, with unreacted starting materials being isolated instead. A brief summary of the reaction conditions for the SAH reactions conducted, along with the different protecting group strategies employed, is shown in Table 2.3 below.
Table 2.3 Different reaction conditions trialled for the Sharpless aminohydroxylation reaction of differently protected forms of (2S)-hex-5-enoic acid. AH = aminohydroxylation product, DH = dihydroxylation product. Benzyl-4-chlorobenzoxyloxy carbamate 2.103 (1.4 eq.) was used as the nitrogen source for all reactions, Reactions run in acetonitrile/water (8:1) under N₂, rt was approximately 25 °C for all reactions. *Test reaction carried out on a TLC scale, products not isolated.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Scale (mg)</th>
<th>Osmium reagent</th>
<th>Olefin protection</th>
<th>Temp, reaction time</th>
<th>Ligand</th>
<th>AH product yield</th>
<th>DH product yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>(DHQ)₂PHAL</td>
<td>41%</td>
<td>Not isolated</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHBoc, OMe</td>
<td>Rt, o/n</td>
<td>(DHQ)₂PHAL</td>
<td>5%</td>
<td>Not isolated</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>(DHQ)₂AQN Test reaction*</td>
<td>Test reaction*</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>(DHQD)₂PHAL Test reaction*</td>
<td>Test reaction*</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>(DHQ)₂PHAL Test reaction*</td>
<td>Test reaction*</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>50</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>None- achiral</td>
<td>16%</td>
<td>21%</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>(DHQ)₂PHAL</td>
<td>27%</td>
<td>22%</td>
</tr>
<tr>
<td>8</td>
<td>45</td>
<td>OsO₄/tBuOH (4 mol%), extra catalyst loading per hour</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>None- achiral</td>
<td>20%</td>
<td>Not isolated</td>
</tr>
<tr>
<td>9</td>
<td>45</td>
<td>K₂OsO₄.2H₂O</td>
<td>NHFmoc, OMe</td>
<td>Rt, o/n</td>
<td>None- achiral</td>
<td>None-unreacted SM</td>
<td>None-unreacted SM</td>
</tr>
<tr>
<td>10</td>
<td>23</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>35 °C, 5 h</td>
<td>None- achiral</td>
<td>73%</td>
<td>22%</td>
</tr>
<tr>
<td>11</td>
<td>100</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>35 °C, 5 h</td>
<td>None- achiral</td>
<td>99%</td>
<td>Not isolated</td>
</tr>
<tr>
<td>12</td>
<td>29</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OMe</td>
<td>35 °C, 5 h</td>
<td>None- achiral</td>
<td>99%</td>
<td>Not isolated</td>
</tr>
<tr>
<td>13</td>
<td>30</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OtBu</td>
<td>35 °C, o/n</td>
<td>(DHQ)₂PHAL</td>
<td>3%</td>
<td>90%</td>
</tr>
<tr>
<td>14</td>
<td>30</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OtBu</td>
<td>35 °C, o/n</td>
<td>None- achiral</td>
<td>70%</td>
<td>Not isolated</td>
</tr>
<tr>
<td>15</td>
<td>30</td>
<td>OsO₄/tBuOH (4 mol%)</td>
<td>NHFmoc, OtBu</td>
<td>35 °C, o/n</td>
<td>(DHQ)₂PHAL</td>
<td>21%</td>
<td>57%</td>
</tr>
</tbody>
</table>

Mild heating of the reaction mixtures was found to be effective in increasing the yield of the SAH product using achiral reaction conditions (i.e in the absence of chiral Cinchona alkaloid ligands) with a 10 °C increase in reaction temperature leading to a dramatic increase in the yield of the SAH product from 20% to 99%. Disappointingly, increasing the reaction temperature did not improve the yields for the chiral reaction mixtures, with only low yields of the SAH product obtained.
2.8.3.7 Effect of a chiral ligand on SAH reaction of protected-(2S)-hex-5-enoic acid derivatives

It was observed that increased reaction temperature and exclusion of bulky chiral ligands resulted in increased yields of the desired AH product 2.129. This observation is in agreement with results published by Castle et al.\textsuperscript{66} who found that introduction of the chiral ligand (DHQD)$_2$PHAL 2.111 both dramatically reduced yield of the AH product (epimeric yield 62%, chiral yield 18%) as well as increased the production of the undesired by-product benzyl carbamate 2.130.

These observations point towards the formation of a highly congested osmium complex upon addition of the chiral ligand. Furthermore, π-stacking between the Cbz and Fmoc group of the Fmoc/OMe ester derivative of protected-(2S)-hex-5-enoic acid 2.128, could reduce reactivity of alkene by increased steric hindrance, especially when bulky chiral ligand used in the reaction. A congested osmium centre would result in the SAH pathway being less favoured, with competing reactions taking precedence, such as hydrolysis of the nitrogen source to produce undesired by-product benzyl carbamate 2.130. Indeed, it is possible that the formation of the desired AH product in the chiral reaction mixes may result from reaction with a less hindered, non-ligand bound osmium centre. This would explain the lack of stereoselectivity observed in the chiral reaction mixtures. A simple diagram, based on the assumptions of ligand binding reported by Bodkin et al.\textsuperscript{110}, that illustrates the highly congested osmium centre upon complexation of the chiral ligand is shown below in Figure 2.6.

The Fmoc/O\textsuperscript{t}Bu SAH products isolated from both the achiral and chiral reaction conditions were examined by normal-phase high performance liquid chromatography (NP-HPLC), Figure 2.7a) and b). Both sets of reaction conditions produced one dominant stereoisomer eluting at a retention time of between 21.613- 21.080 min, referred to hereafter as Isomer X (the slight difference in retention times is due to the manual nature of sample injection).
Chapter Two

Figure 2.7a) NP-HPLC trace for the Fmoc/O\textsubscript{t}Bu SAH product formed in achiral conditions.

Figure 2.7b) NP-HPLC trace for the Fmoc/O\textsubscript{t}Bu SAH product formed in chiral reaction conditions using the ligand (DHQ\textsubscript{2})\textsuperscript{2}PHAL ligand \textsuperscript{2.110}.

Interestingly, the NP-HPLC of the product mix resulting from the chiral reaction conditions Figure 2.7b) showed a 10\% increase in formation of Isomer X, with Peak 1 having a relative area of 45\% following the chiral reaction conditions employed, compared to 35\% for the achiral reaction conditions. However, major impurities were present in both SAH product mixes, despite these products appearing clean by thin layer chromatography. Attempts to further purify Isomer X by NP-HPLC were unfortunately not successful.

Analysis of the \textsuperscript{1}H, COSY, DEPT and HSQC spectra of the product mixture obtained from the chiral reaction conditions revealed that the regiochemistry of the major product, Isomer X to be the desired Fmoc/O\textsubscript{t}Bu terminal amide \textsuperscript{2.132} due to key coupling observed between the \(\varepsilon\)-\(CH\textsubscript{2}\) and \(NH\textsubscript{Cbz}\) groups, Figure 2.8.
2.8.3.8 Regiochemistry observed for the achiral reaction conditions

The achiral reactions proceeded in good - excellent yields of 70 - >90% at an elevated temperature of 35 °C, Scheme 2.37. The regiochemistry of the major SAH product was also elucidated by \(^1\)H and \(^{13}\)C NMR, in a similar manner described above for Isomer X. The example illustrated below is for the Fmoc/OMe SAH product 2.129, Figure 2.9. The \(\varepsilon\)-CH\(_2\) group were identified through the \(^1\)H, \(^{13}\)C, HSQC and DEPT experiments. Analysis of the 2D COSY spectrum of the major regioisomer showed a key correlation between the \(\varepsilon\)-CH\(_2\) group and the NH of the NHCbz group. This indicated that the major AH product was the desired terminal Fmoc/OMe NHCbz isomer 2.129 in preference to the unwanted terminal OH isomer 2.133. This regiochemistry matches with the structure of the desired molecule (2S,5R)-(Hyl), 2.1.

Reagents and conditions: a) Carbamate 2.103 (1.4 eq.), OsO\(_4\)/tBuOH (4 mol%), (5 mol%), CH\(_3\)CN/H\(_2\)O (8:1, v/v), 35 °C, 3 h.

Scheme 2.37 Achiral SAH reaction conditions at the elevated temperature of 35 °C with Fmoc/OMe-ester derivative of (2S)-hex-5-enoic acid 2.129 affording the desired AH product 2.133.
Figure 2.9a) Possible regioisomers of the SAH reaction of Fmoc/OMe-protected-(2S)-hex-5-enoic acid and b) COSY spectrum of major regioisomer isolated, compound 2.129.

2.8.4 Lactonisation studies of the SAH products

Lactonisation of the epimeric Fmoc/OtBu SAH product 2.132 using trifluoroacetic acid was attempted in an attempt to separate the epimeric mixture into (S)-lactone 2.134 and (R)-lactone 2.135, Scheme 2.38. Despite mild heating and lengthy reaction times (35 °C, o/n), a significant quantity of starting material was still present in the reaction mixture. Further addition of trifluoroacetic acid along with the carbocation scavenger\textsuperscript{113} triethylsilane was therefore added to the reaction mixture, in a similar method to that optimised by Mehta \textit{et al.}\textsuperscript{113}. Unfortunately, the lactonisation reaction only yielded a complex mixture, with significant quantities of starting material still present in the reaction mixture, as indicated by thin layer chromatography and \textsuperscript{1}H NMR analysis.

\textit{Reagents and conditions:} a) CF$_3$CO$_2$H (excess), CH$_2$Cl$_2$, 4Å M.S., 35 °C, o/n; b) CF$_3$CO$_2$H (excess), Et$_3$SiH, CH$_2$Cl$_2$, 4Å M.S., 35 °C, 2 h.

\textbf{Scheme 2.38} Attempted lactonisation of SAH product 2.132 with trifluoroacetic acid and triethylsilane to afford (2S,5S)-lactone 2.134 and (2S,5R)-lactone 2.135.

It is curious that the epimeric Fmoc/OtBu SAH product 2.132 should be resistant towards lactonisation, as lactonisation has been a common method to separate derivatives of (S)-Hyl and (R)-Hyl for decades.
Indeed, other published Hyl derivatives appear to lactonise with ease. Examples of the use of lactonisation to separate diastereoisomers of derivatised Hyl include the 1953 paper by Fones et al.

Koeners et al. (1981) and later papers by Allevi et al. have also employed lactones as useful intermediates following achiral epoxidation reactions, which are described in more detail in Section 2.9.3.

In summary, the Sharpless aminohydroxylation reaction was successfully performed on differently protected forms of (2S)-hex-5-enoic acid. Gratifyingly, the SAH selectively produce the desired regioisomers of Hyl (5-Hyl) (Fmoc/OMe AH product) and 2.132 (Fmoc/O'Bu AH product).

Despite the use of various chiral Cinchona ligands, we were unable to control the absolute stereochemistry of the SAH reaction, resulting in an epimeric mixture being produced rather than the desired (2S,5R)-Hyl, 2.1. The epimeric mixture was not able to be separated into its constituent

**Scheme 2.39** Successful lactonisation methods applied to separate epi-Hyl derivatives 2.136, 2.139 and 2.10 by Fones et al., Koeners et al. and Allevi et al. respectively. Fcc= silica flash column chromatography.
molecules (2S,5R)-Hyl and (2S,5S)-Hyl using either silica chromatography, chiral NP-HPLC or lactonisation techniques. Therefore, the epi-Hyl was carried forward and used in glycosylation reactions, as outlined in Section 2.14. As native Hyl has been found to be exclusively (2S,5R)-Hyl 2.1, there may be biological complications arising from bringing an epimeric mixture into Adpn peptide synthesis. However, it was hoped that the Gal-residue installed upon glycosylation may act as a chiral auxiliary and facilitate separation of the epimeric mixture.

2.9 Use of epoxidation reactions to afford the desired (5R) stereocentre of Hyl

2.9.1 Revised retrosynthetic analysis for (2S,5R)-Hyl, 2.1

Due to the difficulties encountered with the separation of epi-Hyl, alternative strategies to introduce the hydroxyl group to the olefin centre of protected (2S)-hex-5-enoic acid 2.91 to afford protected (2S,5R)-Hyl 2.90 were considered. A revised retrosynthetic analysis of (2S,5R)-Hyl 2.1 revealed epoxidation to be another possible method for installation of the desired (5R)-OH stereocentre, Scheme 2.40.

It was envisaged that protected (2S,5R)-Hyl 2.90 could be disconnected to the corresponding chiral azido alcohol 2.142 via reduction (Pd/C, H₂) of the terminal azide functionality. It was then hoped that the chiral azido alcohol 2.142 could be separated from the corresponding epimeric mixture 2.143 using silica gel chromatography. The epimeric mixture 2.143 could be disconnected to the corresponding epimeric epoxide 2.144, which in turn could be accessed from the previously synthesised starting material; protected derivatives of (2S)-hex-5-enoic acid 2.91.

![Scheme 2.40 Revised retrosynthetic analysis for the synthesis of protected (2S,5R)-Hyl 2.90 from previously synthesised protected (2S)-hex-5-enoic acid 2.91.](image)

2.9.2 Similar synthetic strategies employed by other research groups

A similar synthetic strategy was employed by Chida et al.115 in their reported synthesis of Bengamide A. They used the achiral peracid epoxidation method to epoxidise alkene 2.145 in excellent 99% yield. They found the resulting epoxide 2.146 to be an inseparable epimeric mixture, and so conducted azidolysis of epimeric epoxide 2.146 to afford the epimeric azido compound 2.147. Gratifyingly, the
epimeric azido compound 2.147 could be separated into the corresponding diastereoisomers 2.148 and 2.149 using silica gel chromatography both in 48% yield, Scheme 2.41.

![Scheme 2.41](image-url)

Reagents and conditions: a) mCPBA, phosphate buffer (pH 8), 50 °C; b) NaN₃, NH₄Cl, 2-methoxyethanol/ H₂O (9:1, v/v), 50 °C; c) silica gel chromatography (further synthetic details not available).

Scheme 2.41 Functionalisation of the alkene moiety of compound 2.145 by epoxidation followed by azidolysis.¹¹⁵

The absolute configuration of epimeric compounds 2.148 and 2.149 was determined by conversion of these compounds into their corresponding known lactones 2.150 and 2.151, as shown in Scheme 2.42. The optical rotation of lactones 2.150 and 2.151 were then compared to known literature values.

![Scheme 2.42](image-url)

Reagents and conditions: a) NaH, BnBr, n-Bu₄NI, THF, rt; b) TsOH (5 mol%), MeOH, rt; c) Jones’ reagent (CrO₃ in dil. H₂SO₄), acetone, 0 °C; d) 20% Ph(OH)₂, H₂, EtOH/dil HCl; e) Cbz-Cl, NaOH, H₂O/ EtOH/THF; f) TFA/ CH₂Cl₂; g) as per e); toluene, reflux (further experimental details including yields for steps d) – h) not reported).

Scheme 2.42 Determination of absolute configuration of azido compounds 2.148 and 2.149 by conversion into known lactones 2.150 and 2.151.¹¹⁵
2.9.3 Chiral epoxidation reactions

It should be noted that as protected (2S)-hex-5-enoic acid 2.91 is a terminal alkene, it is not a suitable substrate for classical enantioselective epoxidation methods such as Jacobson117-Katsuki118 epoxidation methods which use chiral salen-manganese complexes or Shi epoxidation119 which uses dioxirane and a fructose derived ketone to direct the stereoselectivity of the reaction. Indeed the low reactivity of terminal olefins towards electrophilic oxidising reagents renders the enantioselective oxidation of terminal alkenes a significant challenge in organic chemistry.120

Late in candidature, Strukul et al.120 reported a new asymmetric epoxidation method of terminal alkenes using hydrogen peroxide as the oxidant and chiral Pt(II) complexes to guide the stereoselectivity of the reaction. The most similar alkene to our substrate tested by Strukul et al.120 was hexene 2.158. Whilst very interesting, commitment had already been made to the alternative route with mCPBA outlined below and so this synthetic pathway was not explored.

Scheme 2.43 The asymmetric epoxidation of hexene 2.158 to (R)-1,2-epoxyhexene 2.159 using hydrogen peroxide and a chiral Pt(II) complex, as published by Strukul et al.120

Therefore, for the synthesis of (2S,5R)-Hyl 2.1 achiral epoxidation reactions of previously synthesised terminal alkenes 2.128 (Fmoc/OMe), 2.160 (Fmoc/OtBu) and 2.161 (Boc/OMe) using the peracid mCPBA (meta-chloroperoxybenzoic acid) were conducted, Scheme 2.44. These reactions produced epimeric epoxides 2.162 – 2.164 with varied degrees of success, 34 – 64% yield. The Boc/OMe protected epimeric epoxide 2.164 was taken as a test epoxide and converted into the corresponding azido alcohol 2.165 using sodium azide in refluxing DMF. Unfortunately, attempts to resolve either the epoxide epimers 2.162 – 2.164 or the azido epimeric mix 2.165 to afford the desired diastereomer azide 2.166 using silica gel chromatography were unsuccessful and this route was abandoned in favour of the higher yielding SAH reaction followed by glycosylation (see Section 2.12).
\[\text{Reagents and conditions: } a) \text{i) mCPBA (2 eq.), CH}_2\text{Cl}_2, \text{rt, o/n}; b) \text{ii) NaN}_3 (2 \text{ eq.}, \text{DMF, 2 h, reflux}.\]

**Scheme 2.44** Attempted synthesis of (2S,5R)-Hyl 2.1 via epoxide intermediates 2.162 – 2.164.

Interestingly, Allevi et al.\textsuperscript{116} also reported technical issues in separating both the epimeric epoxide 2.167 and the epimeric bromohydrin 2.168 generated in their published route for the synthesis of (2S,5R)-Hyl 2.1. Allevi et al.\textsuperscript{116} solved their separation problem by conversion of epimeric bromohydrin 2.168 into the corresponding lactones, which were easily separated into the constituent epimeric lactones 2.169 and 2.170 by silica gel chromatography, Scheme 2.45.

\[\text{Reagents and conditions: } a) \text{LiBr (1.6 eq.), THF: AcOH (20:0.34, v/v), rt, o/n}; b) \text{CF}_3\text{CO}_2\text{H (15 eq.), rt, 40 min.}\]

**Scheme 2.45** Section of the reported synthesis by Allevi et al.\textsuperscript{116} of (2S,5R)-Hyl 2.1 via an inseparable epimeric epoxide intermediate 2.167 and epimeric bromohydrin 2.168.

Lactonisation of the epimeric azide 2.167 was not attempted during the course of this work as previous lactonisation of a similar substrate SAH product 2.132 (Section 2.8.4) had been unsuccessful. Furthermore, the initial aim of this work was to develop a shorter and higher yielding pathway towards (2S,5R)-Hyl 2.1. Use of the lactonisation technique to separate epimeric azide 2.165 into its constituent
diastereoisomers would add further steps to our synthetic pathway and in addition render this synthesis too similar to the route described by Allevi et al.\textsuperscript{58,116} Therefore, alternative pathways for the synthesis of (2S,5R)-Hyl 2.1 were sought and investigated including the use of the 9-borabicyclononane (9-BBN) protecting group for native lysine.

### 2.10 Use of 9-BBN as a protecting group for (2S,5R)-Hyl

During the last decade, the use of the 9-BBN (9-borabicyclononane) group has gained popularity in organic synthesis. The ability of the 9-BBN group to simultaneously protect both the carboxylic acid and the amine group of the amino functionality has led to increased use of this moiety as a protecting group for amino acids. In 2002, Dent \textit{et al.}\textsuperscript{121} demonstrated that 9-BBN can be used as a protecting group for many amino acids including lysine, serine and arginine. Furthermore, they showed addition of the 9-BBN groups to these amino acids increased the solubility of these compounds in organic solvents. The removal of the 9-BBN group was also shown to be facile, proceeding well in either acidic (aq. HCl) or basic (ethylenediamine) conditions, Scheme 2.46.

**Reagents and conditions:** a) 9-BBN (1.1 eq.), MeOH, reflux, 3 h; b) conc. HCl (1 mL per mmol 2.170), MeOH (2 mL per mmol 2.172), rt, 10 min; c) b) ethylenediamine (excess), reflux, 1 min.

**Scheme 2.46** The synthesis of BBN-lysine 2.172 from native lysine 2.171 and subsequent removal of protecting group by either acidolysis or basic conditions, as demonstrated by Dent \textit{et al.}\textsuperscript{121}

In 2004, Kol \textit{et al.}\textsuperscript{122} utilised 9-BBN as a protecting group for the synthesis of 3-amino alcohols. During the same year, Kihlberg \textit{et al.}\textsuperscript{123} successfully employed 9-BBN to protect the amino group of commercially available (very expensive) (2S,5R)-Hyl.HCl 2.173 to afford BBN-Hyl 2.172. The ε-amine group of BBN-Hyl 2.172 was then Cbz-protected to afford borane complex 2.174, with an excellent yield of 92% reported for both protecting group steps. Furthermore, Kihlberg \textit{et al.}\textsuperscript{123} demonstrated that BBN-Cbz-Hyl 2.174 could be glycosylated in 62% yield using acetobromogalactose 2.61 as the galactose donor to afford the glycosylated borane complex 2.175. The BBN group was then easily removed using a mixture of chloroform and methanol to yield glycosylated Gal-Hyl derivtaive 2.176 in an excellent yield of 91%, Scheme 2.47.
Reagents and Conditions: a)i) NH₃ (aq., excess), 0 °C, 30 min, ii) 9-BBN (1.2 eq.), MeOH, reflux, 4 h; b) Cbz-Cl (1.2 eq.), NaHCO₃ (1.5 eq.), dioxane/water, 0 °C → rt, 5 h, c) acetobromogalactose (2.61, 1.5 eq.), silver silicate (5 eq.), molecular sieves (3 Å, excess), CH₂Cl₂, absence of light, 0 °C, 8 h; d) i) CHCl₃/MeOH (5:1, v/v), rt, 12 h, ii) Fmoc-OSu (1.2 eq.), NaHCO₃ (1.1 eq.), H₂O:acetone (5:2), rt, 5 h.

Scheme 2.47 The synthesis of BBN-Cbz-(2S,5R)-Hyl 2.174 and subsequent glycosylation to afford BBN-Cbz-Gal-Hyl 2.175 as reported by Kihlberg et al.²²³

Following the methodology published by Kihlberg et al.,²²³ commercially available epi-Hyl 2.177 was successfully protected using the 9-BBN protecting group for the amino functionality and the Cbz protecting group for N-ε amine group, affording borane complex 2.178 in a good overall yield of 43% (over three steps), Scheme 2.48. The resultant borane complex 2.178 was unable to be separated into its constituent diastereoisomers, but was carried through and subjected to a range of glycosylation conditions as detailed in Section 2.14. Borane complex 2.178 was also converted into Fmoc/OAll-epi-Hyl 2.179 in a moderate yield of 18% over three steps, Scheme 2.48, step c). Borane complex 2.178 and Fmoc/OAll-epi-Hyl 2.179 were tested in glycosylation reactions outlined in Section 2.14.
Reagents and Conditions: a)i) NH$_3$ (aq., excess), 0 °C, 30 min, ii) 9-BBN (1.1 eq.), MeOH, reflux, 5 h; b) Cbz-Cl (1.4 eq.), NaHCO$_3$ (1.5 eq.), dioxane/water, 0 °C → rt, 2 h; c)i) ethylenediamine (4 eq.), THF, reflux, 10 min, ii) Fmoc-OSu (1.4 eq.), K$_2$CO$_3$ (1 eq.), rt, 2 h, iii) tetrabutylammonium bromide (1 eq.), allyl bromide (5 eq.), rt, o/n.

Scheme 2.48 Synthesis of BBN-Cbz-epi-Hyl complex 2.178 from commercially available epi-Hyl 2.177 and subsequent change of protecting groups to afford Fmoc/OAll-epi-Hyl 2.179.

Using the same methodology, borane complex 2.180 of previously prepared (2S)-hex-5-enoic acid 2.91 was also synthesised in excellent yield (83%), Scheme 2.48. However, subsequent epoxidation of borane complex 2.180 with mCPBA to afford borane epoxide complex 2.181 was not successful, with a complex, inseparable mixture resulting from the reaction conditions, Scheme 2.49. Mass spectrometric analysis of the reaction mixture indicated that the epoxidation conditions trialled resulted in deprotection of the 9-BBN group in preference to the desired epoxidation reaction.

Reagents and Conditions: a)i) NH$_3$ (aq., excess), 0 °C, 30 min, ii) 9-BBN (1.2 eq.), MeOH, reflux, 3 h; b) mCPBA (2.5 eq.), THF, 0 °C → rt, o/n.

Scheme 2.49 The attempted synthesis of the borane epoxide complex 2.180 from (2S)-hex-5-enoic acid 2.91 and attempted subsequent epoxidation of borane complex 2.180.
2.11 Summary of our approach towards the synthesis of (2S,5R)-Hyl

In summary, Sections 2.7 – 2.10 have detailed the attempted synthesis of (2S,5R)-Hyl \textbf{2.1} through four main synthetic pathways. Firstly, in Section 2.7 selective oxidation of a chiral piperidinone enolate following the synthetic route published by Guichard \textit{et al.}\textsuperscript{75} was attempted as the key chiral reaction to form the desired 5R-stereocentre. However, this first synthetic pathway was not successful due to incomplete reduction of ketone intermediate \textbf{2.87}, affording the undesired β-hydroxyester intermediate \textbf{2.88} in preference to the desired diester intermediate \textbf{2.89}.

The second synthetic pathway, outlined in Section 2.8 utilised stereoselective alkylation of the Belokon complex\textsuperscript{91} to afford (2S)-hex-5-enoic acid \textbf{2.91}. The Sharpless aminohydroxylation reaction\textsuperscript{90} with key modifications by Luxenburger \textit{et al.}\textsuperscript{99} was then used to install the second (5R)-hydroxyl stereocentre on the terminal alkene. Although high yields (73\%) and good regioselectivity were observed, only epimeric-Hyl was produced. Unfortunately, methods for separation of the epimeric mixture into its constituent diastereoisomers including silica gel chromatography, chiral NP-HPLC and lactonisation failed.

Therefore, a third synthetic pathway was attempted utilising the (2S)-hex-5-enoic acid \textbf{2.91} produced in the second synthesis. In this method, achiral epoxidation followed by azidolysis was carried out. Unfortunately, neither separation of the epimeric epoxides \textbf{2.162} – \textbf{2.164} nor the subsequent azido alcohol epimeric mixture \textbf{2.165} was successful.

Therefore, as detailed in Section 2.10, using methodology developed by Kihlberg \textit{et al.}\textsuperscript{123} commercially available \textit{epi}-Hyl \textbf{2.177} was suitably protected to afford BBN-Cbz-\textit{epi}-Hyl complex \textbf{2.178} and Fmoc/OAll-\textit{epi}-Hyl \textbf{2.179}, Scheme 2.50.
Reagents and conditions: a) Meldrum’s acid (1 eq.), EDC (1.5 eq.), DMAP (1.5 eq.), CH₂Cl₂, rt, 2 h (crude product carried forwards without further purification); b) NaBH₄ (3 eq.), CH₂Cl₂:AcOH (9:1, v/v), rt, o/n; c) carbamate 2.101 (1.4 eq.), OsO₄/‘BuOH (4 mol%), (DHQ)₂PHAL 2.110 (5 mol%), CH₃CN/H₂O (8:1, v/v), rt, 3 h; d) i) mCPBA (2 eq.), CH₂Cl₂, rt, o/n; e) ii) NaN₃ (2 eq.), DMF, 2 h, reflux; f) i) NH₃ (aq., excess), 0 °C, 30 min, ii) 9-BBN (1.1 eq.), MeOH, reflux, 5 h; g) Cbz-Cl (1.4 eq.), NaHCO₃ (1.5 eq.), dioxane/water, 0 °C → rt, 2 h; h) i) ethylenediamine (4 eq.), THF, reflux, 10 min, ii) Fmoc-OSu (1.4 eq.), K₂CO₃ (1 eq.), rt, 2 h, iii) tetrabutlammonium bromide (1 eq.), allyl bromide (5 eq.), rt, o/n.

Scheme 2.50 The four different synthetic pathways attempted for the synthesis of (2S,5R)-Hyl 2.1.
Although the four attempted synthetic pathways were not successful in affording \((2S,5R)\)-Hyl, significant quantities of differently protected epimeric-Hyl (\(epi\)-Hyl) were produced. It was decided that glycosylation of these protected \(epi\)-Hyl derivatives would be attempted, in the hope that the addition of a sugar moiety to \(epi\)-Hyl might act as a chiral auxiliary and thus help to resolve the epimeric mixture. The attempted synthesis of Gal-Hyl 2.2 is detailed in the following Section 2.12.

### 2.12 Retrosynthetic strategy towards Gal-Hyl 1.9

Protected Gal-Hyl derivative 2.182 was disconnected to the previously synthesised protected \(epi\)-Hyl derivatives (Section 2.7 – 2.11, namely Fmoc/OMe-\(epi\)-Hyl 2.129, BBN-\(epi\)-Hyl 2.178 and Fmoc/OAll-\(epi\)-Hyl 2.179) and the Schmidt\(^{124}\) glycosidic donor: 2,3,4,6-tetra-\(O\)-acetyl-\(\alpha,\beta\)-\(D\)-galactopyranosyl-trichloroacetimidate 2.183, Scheme 2.51. The Schmidt glycosylation method was chosen in preference to the Koenigs-Knorr\(^{125}\) glycosylation pathway or use of a thioglycoside donor\(^{126}\) as the Schmid method has already been successful applied to the synthesis of Gal-Hyl 2.2 by Allevi \textit{et al.} \(^{88}\) and Nishiuchi \textit{et al.} \(^{48}\) Therefore, this technique was known to work well for similarly protected Hyl substrates. In addition, the Schmidt method has been previously successfully employed by the Brimble group for the synthesis of glycosides including fish antifreeze glycopeptides.\(^{127,128}\)

![Scheme 2.51 Retrosynthetic analysis conducted for Gal-Hyl 2.181 utilising previously synthesised \(epi\)-Hyl derivatives 2.129, 2.178 and 2.179.](image)

#### 2.13 Synthesis of the galactosyl donor

The methodology used for the synthesis of 2.183 had previously been developed by Brimble \textit{et al.}\(^{127}\) Commercially available \(D\)-galactose was peracetylated using acetic anhydride/ perchloric acid to afford peracetylgalactose 2.184 in quantitative yield. The anomic acid protecting group was then selectively removed using glacial acetic acid/ ethylenediamine using the method outlined by Lindhorst \textit{et al.}\(^{129}\) to afford tetraacetylgalactose 2.185 in 70% yield. Trichloroacetonitrile along with potassium carbonate were then added to tetraacetylated-\(D\)-galactose 2.185 to afford tetraacteyl-\(D\)-galactopyranosyl trichloroacetimidate galactose donor 2.183 in quantitative yield.
Chapter Two

Scheme 2.52 Synthesis of tetraacetyl-d-galactopyranosyl trichloroacetimidate galactose donor 2.183.

2.14 Coupling of the galactose donor to epri-Hyl derivatives

2.14.1 Glycosylation of epri-BBN-Hyl with galactose donor

As described in Section 2.10, Kihlberg et al.\textsuperscript{123} have published the successful glycosylation of BBN-Cbz-(5R)-Hyl 2.174 with tetraacetylated-galactosyl bromide donor 2.60 using silver silicate/zeolite as the reaction promoter in 68\% yield. Due to the high light- and moisture sensitivity of tetraacetylated-galactosyl bromide donor 2.60, standard Schmidt conditions\textsuperscript{130} were used to couple epri-BBN-Hyl 2.178 and tetraacteyl-d-galactopyranosyl-trichloroacetimidate galactose donor 2.183 using trimethylsilyl trifluoromethanesulfonate (TMSOTf) as the reaction promoter, Scheme 2.53.

\[
\text{Reagents and Conditions: a) 2.182 (3 eq.), TMSOTf (0.5 eq.), CH}_{2}\text{Cl}_{2}, \text{mol. sieves (4 Å, excess), 0 °C → rt, 45 min.}
\]

Scheme 2.53 Attempted glycosylation of epri-BBN-Hyl 2.178 with tetraacteyl-d-galactopyranosyl trichloroacetimidate as the galactose donor 2.183 to afford BBN-Gal-Hyl 2.186.

Although these glycosylation conditions did afford the desired product 2.186 the product was found to be an inseparable mixture of different glycosides; (5R)-β-BBN-Gal-Hyl 2.186a, (5R)-α-BBN-Gal-Hyl 2.186b, (5S)-β-BBN-Gal-Hyl 2.186c and (5S)-α-BBN-Gal-Hyl 2.186b, Figure 2.10.
Figure 2.10 The four different stereoisomers 2.186a–d obtained from the glycosylation of epi-BBN-Hyl 2.178 with tetraactylo-D-galactopyranosyl-trichloroacetimidate 2.183.

Mass spectrometry confirmed the presence of the desired BBN-Gal-Hyl 2.186 product, HRMS-ESI calculated for C_{36}H_{51}BN_{2}O_{14}Na^{+} 769.3332, found 769.3308, Figure 2.11 and NMR analysis of the glycosidic mixture indicated the presence of the desired (5R)-β-BBN-Gal-Hyl 2.186a, albeit as a complex mixture.

Figure 2.11 Mass spectrum of the complex mixture containing BBN-Gal-Hyl 2.186 with [M+Na]^{+} found 769.3308, calculated 769.3332.

Unfortunately, separation of (5R)-β-BBN-Gal-Hyl 2.186a from the complex product mixture using silica chromatography was unsuccessful. However, the optical rotation value recorded for the complex mixture of epi-BBN-Gal-Hyl 2.186 ([α]_{D}^{23} +12.14, c 1.9, CHCl_{3}) is similar to the value recorded by
Kihlberg et al.\textsuperscript{123} for (5R)-α-BBN-Gal-Hyl 2.186\textsuperscript{a} ([α]\textsubscript{D}\textsuperscript{20} +4.9, c 1.9, CHCl\textsubscript{3}\textsuperscript{123}) thus implying that (5R)-β-BBN-Gal-Hyl 2.186\textsuperscript{a} was the major stereoisomer present in the complex product mixture.

In addition, the major side-products isolated from the glycosylation reactions were the orthoester 2.187 (26% yield) and C-1 deprotected tetraacetylgalactose 2.185 (14% yield), Figure 2.12.

**Figure 2.12** Orthoester product 2.187 and tetraacetylgalactose 2.185, the major side-products isolated from small-scale glycosylation reactions of epi-BBN-Hyl 2.178 with tetraacteyl-D-galactopyranosyl trichloroacetimidate 2.183.

### 2.14.2 Mechanism of glycosylation

The mechanism for the glycosylation reaction of epi-BBN-Hyl 2.178 with tetraacteyl-D-galactopyranosyl-trichloroacetimidate 2.183 was considered to optimise the reaction conditions for the glycosylation reaction. As the acetylated trichloroacetimidate 2.183 was chosen as the glycosylation donor, the mechanism for glycosylation involves initial elimination of the anomic trichloroacetimidate leaving group of 2.183 followed by neighbouring group participation (NGP) of the C2-acetyl group to stabilise the intermediate oxocarbenium ion 2.188, resulting in a more stable acyloxonium ion 2.189.

The glycosylation acceptor, in this case epi-BBN-Hyl 2.178, then attacks the acyloxonium ion 2.189 affording the desired BBN-Gal-Hyl 2.186. Alternatively, water can attack the acyloxonium ion 2.189 resulting in the undesired orthoester side product 2.187, Scheme 2.54. Mildly acidic reaction conditions cause the orthoester product 2.187 to be formed reversibly, thus increasing reaction yields of the desired β-glycoside 2.186.
Scheme 2.54 Mechanism for the glycosylation reaction of \( \text{epi-BBN-Hyl} \) \( 2.178 \) with acetylated trichloroacetimidate \( 2.183 \) invoking neighbouring group participation (NGP) from C2-acetyl group.

In the glycosylation conditions trialled for the glycosylation of \( \text{epi-BBN-Hyl} \) \( 2.178 \), the hydrolysis side reaction generating the unwanted orthoester side product \( 2.187 \) was disfavoured due to the use of activated molecular sieves and a nitrogen atmosphere to keep the reaction conditions as dry as possible. Different Lewis acids were trialled as reaction promoters for this glycosylation reaction including pyridinium \( \text{para-toluenesulfonate (PPTS)} \) and boron trifluoride diethyl etherate (BF\(_3\).Et\(_2\)O).

The different reaction conditions trialled are listed below in Table 2.4. Given that all the glycosylation conditions resulted in low yields of the desired BBN-Gal-Hyl product \( 2.186 \), \( \text{epi-BBN-Hyl} \) \( 2.178 \) was abandoned as a glycosylation substrate in favour of \( \text{epi-Fmoc-OAll-OMe} \) \( 2.129 \) and \( \text{epi-Fmoc-OAll-Hyl} \) \( 2.179 \), discussed in the following Section.
Reagents and Conditions: a) 2.183 (3 eq.), TMSOTf (0.5 eq.), CH$_2$Cl$_2$, mol. sieves (4 Å, excess), 0 °C → rt, 45 min.

Scheme 2.55 Attempted glycosylation of epi-BBN-Hyl 2.178 with tetraacteyl-D-galactopyranosyl-trichloroacetimidate as the galactose donor 2.183 to afford rac-BBN-Gal-Hyl 2.186.

Table 2.4 Different reaction conditions trialed for the glycosylation of epi-BBN-Hyl 2.177 with tetraacteyl-D-galactopyranosyl trichloroacetimidate 2.183. Reactions carried out under a nitrogen atmosphere with flamed dried glassware and with activated molecular sieves (4 Å). Reaction mixtures were quenched with Et$_3$N (20 µL). Purification was carried out using flash column chromatography on deactivated silica (1% Et$_3$N).

*Minor desired product epi-Gal-Hyl separated from product mixture using fcc. The unwanted side products were not isolated.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Scale (mg, epi-BBN-Hyl)</th>
<th>Lewis acid (0.5 eq.)</th>
<th>Temp (°C)</th>
<th>Time</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>TMSOTf</td>
<td>-10 → 0</td>
<td>5 h</td>
<td>Major product isolated: 4OAc-Gal.</td>
</tr>
<tr>
<td>2</td>
<td>14</td>
<td>TMSOTf</td>
<td>-10 → rt</td>
<td>2 h</td>
<td>Major product: 4OAc Gal Minor product: Gal-Hyl (2 mg, 8% yield)*</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>TMSOTf</td>
<td>-10 → rt</td>
<td>45 min</td>
<td>Minor product: Gal-Hyl (8 mg, 13% yield)</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>TMSOTf</td>
<td>-10 → rt</td>
<td>1 h</td>
<td>Minor product: Gal-Hyl (10 mg, 14% yield).</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>PPTS</td>
<td>-10 → rt</td>
<td>3.5 h</td>
<td>Major products: orthoester and 4OAc product</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>BF$_3$.EtO</td>
<td>-10 → rt</td>
<td>1 h</td>
<td>Decomposition</td>
</tr>
</tbody>
</table>
2.14.3 Glycosylation studies of \textit{epi}-Fmoc/OMe-Hyl and \textit{epi}-Fmoc/OAll-Hyl

Due to the failure to achieve good yields of Gal-BBN-Hyl 2.186, attention shifted to the glycosylation of previously synthesised \textit{epi}-Hyl derivatives \textit{epi}-Fmoc/OMe-Hyl 2.129 and \textit{epi}-Fmoc/OAll-Hyl 2.179 with trichloroacetimidate 2.183, Scheme 2.56 to prepare Gal-Fmoc/OMe-Hyl 2.190 and Gal-Fmoc/OAll-Hyl 2.191 respectively.

![Scheme 2.56](image_url)

\textit{Reagents and Conditions:} a) 2.183 (1.7 eq.), TMSOTf (0.5 eq.), CH$_2$Cl$_2$, mol. sieves (4 Å, excess), rt, 1 h.

Scheme 2.56 Glycosylation conditions attempted for the synthesis of Gal-Fmoc/OMe-Hyl 2.191 and Gal-Fmoc/OAll-Hyl 2.192.

Unfortunately, these reaction conditions did not promote formation of the desired β-glycoside as the major product. Instead, deprotection of both amino groups (α-NHFmoc and ε-NHCbz) was observed along with degradation products. Optimisation of the glycosylation reaction conditions was attempted but was ultimately unsuccessful. A summary of the different glycosylation conditions attempted is given in the Table 2.5.

Table 2.5 Different reaction conditions trialed for the glycosylation of \textit{epi}-Fmoc/OMe-Hyl 2.129 and \textit{epi}-Fmoc/OAll-Hyl 2.179 and with tetaacteyl-D-galactopyranosyl trichloroacetimidate 2.183. Reactions quenched with Et$_3$N (20 µL). Purification was carried out using flash column chromatography on deactivated silica (1% Et$_3$N).

<table>
<thead>
<tr>
<th>Entry</th>
<th>Scale (mg \textit{epi}-Hyl)</th>
<th>\textit{epi}-Hyl protection</th>
<th>Lewis acid (0.5 eq.)</th>
<th>Temp (°C)</th>
<th>Time</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>NHFmoc, OAll</td>
<td>TMSOTf</td>
<td>rt</td>
<td>1 h</td>
<td>Complex product mix</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>NHFmoc, OAll</td>
<td>TMSOTf</td>
<td>rt</td>
<td>2 h</td>
<td>Degradation products</td>
</tr>
<tr>
<td>3</td>
<td>26</td>
<td>NHFmoc, OMe</td>
<td>TMSOTf</td>
<td>rt</td>
<td>2 h</td>
<td>Complex product mix</td>
</tr>
<tr>
<td>4</td>
<td>100</td>
<td>NHFmoc, OMe</td>
<td>TMSOTf</td>
<td>rt</td>
<td>1 h</td>
<td>Complex product mix</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>NHFmoc, OMe</td>
<td>TMSOTf</td>
<td>rt</td>
<td>3 h</td>
<td>Complex product mix</td>
</tr>
<tr>
<td>6</td>
<td>100</td>
<td>NHFmoc, OMe</td>
<td>TMSOTf</td>
<td>rt</td>
<td>2 h</td>
<td>Complex product mix</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>NHFmoc, OMe</td>
<td>TMSOTf</td>
<td>rt</td>
<td>4 h</td>
<td>Complex product mix</td>
</tr>
</tbody>
</table>
2.11 Conclusion for Chapter Two

In conclusion, the attempts to glycosylate differently protected $epi$-Hyl compounds BBN-$epi$-Hyl 2.178, Fmoc/OMe-$epi$-Hyl 2.129 and Fmoc/OAll-$epi$-Hyl 2.179 were not successful. Further studies for this glycosylation could include dilution of the sulfonate catalyst, examination of different protecting groups, such as OBz, for the sugar donor which could discourage orthoester formation and use of alternate glycosylation strategies, such as the Koenigs-Knorr and the Helferich glycosylation methods. However, focus of this PhD study shifted to use of ‘click’ neoglycoside mimetics as a model for Gal-Hyl 1.9, which are discussed in detail in Chapter 3.
Chapter Three

The Synthesis of *Neoglycopeptides* Corresponding to the Collagenous Domain of Adiponectin
Chapter Three

Chapter Three: The Synthesis of Neoglycopeptides Corresponding to the Collagenous Domain of Adiponectin.

3.2 Overview of the biological activity of Adpn.

As discussed in Section 1.1, Adpn has been shown to be a key player in obesity related diseases such as Type II diabetes\(^2\) and cardiovascular disease.\(^4\) Studies have revealed a clear inverse relationship between body mass index (BMI) and Adpn. With these studies, overweight and obese patients exhibited lower blood plasma levels of Adpn than observed for patients within the ‘normal’ BMI range.\(^{12}\) Of particular note is the role of Adpn in controlling blood glucose levels in Type II diabetic patients. Scherer \textit{et al.}\(^{23}\) demonstrated that injections of murine Adpn into Type II diabetic mice (\textit{ob/ob} mouse model) resulted in a direct decrease of blood glucose levels in the mice. An abdominal injection of murine Adpn (84 \(\mu\)g per gram of mouse body weight) into the \textit{ob/ob} mice led to a six fold increase in blood Adpn levels from 20 \(\mu\)g/mL to 125 \(\mu\)g/mL after four hours. This spike in blood Adpn levels was accompanied by a corresponding decrease in blood glucose levels from 150 mg/dL to 115 mg/dL, close to the considered ‘normal’ murine blood glucose level of 117 mg/dL.\(^{131}\) This data strongly suggests a causative relationship between blood Adpn levels and blood glucose levels.

In addition, simple lifestyle changes such as the adoption of a Mediterranean diet\(^{37}\) and increased exercise\(^{38}\) have also been shown to stimulate increase blood plasma levels of Adpn. Furthermore, administration of several diabetic medications such as glimepiride\(^{41}\) and troglitazone\(^{42}\) have been shown to stimulate Adpn production. Given the wealth of evidence linking Adpn as a key mediator involved in the regulation and control of blood glucose levels, there is widespread interest in developing a therapeutic version of human Adpn for the treatment of obesity-related diseases such as Type II diabetes.\(^6\)

3.3 Previous chemical syntheses of Adpn

Expression of a therapeutic form of Adpn through recombinant techniques has proved to be challenging and ultimately not commercially viable, due to the different oligomeric forms of Adpn produced in a mammalian cell line and the difficulties encountered in subsequent purification.\(^{36}\) Therefore, a chemical approach to the synthesis of human Adpn has been adopted by numerous groups as discussed in Section 1.7. Brimble \textit{et al.}\(^{132}\) published the first synthesis of the collagenous domain of Adpn and in 2014 Nishiuchi \textit{et al.}\(^{48}\) succeeded in the first total chemical synthesis of the collagenous domain of Adpn (19 - 107) bearing the key post-translational modifications at lysine residues 68, 71, 80 and 104.

These post-translationally modified lysine (PTM-Lys) residues have been shown by Spiro \textit{et al.}\(^{53}\) to have the chemical structure Gal-Glu-Hyl 1.8, Figure 3.1. Many other groups such as Allevi \textit{et al.}\(^{88}\) and Kihlberg \textit{et al.}\(^{85}\) have also reported the synthesis of the Gal-Glu-Hyl 1.8. As highlighted in Section 1.4,
the four PTM-Lys residues found in human Adpn have been shown to be crucial for the bioactivity of the protein.\textsuperscript{28}

Figure 3.1 The structure of Gal-Glu-Hyl \textbf{1.8} (PTM-Lys), as elucidated by Spiro \textit{et al.}\textsuperscript{53}

Although impressive, the chemical synthesis of the post-translationally modified collagenous domain (19 – 107) of native human Adpn by Nishiuchi \textit{et al.}\textsuperscript{48} involves two low yielding glycosylation reactions for preparation of the Glu-Gal-Hyl residue, steps a) and c) Scheme 3.1.

\textit{Reagents and conditions:} a) TMSOTf, molecular sieves (3 Å, excess), -42 °C, CH\textsubscript{2}Cl\textsubscript{2}, 61%; b) hydrazinedithiocarbonate (HDTC), pH 7; c) NIS, TMSOTf, DMF; d) Pd(PPh\textsubscript{3})\textsubscript{4}, morpholine (further synthetic details not reported).

\textbf{Scheme 3.1} Glycosylation conditions reported by Nishiuchi \textit{et al.}\textsuperscript{48} to afford Glu-Gal-Hyl derivative \textbf{2.86}. 

70
3.4 Utilisation of neoglycosides as robust glycoside mimetics

In addition to being low yielding, the classical glycosylation methods utilised by Nishiuchi et al. are known to be very air and moisture sensitive and therefore must be conducted quickly, at low temperatures and under expensive inert atmospheres. Conversely, the Cu(I)-catalysed azide-alkyne cycloaddition reaction (CuAAC, ‘click’ reaction) has been shown to be neither air nor moisture sensitive, produces neoglycosides such as neogal-Lys 1.10 in high yields and is also more flexible in terms of solvent and protecting group choice. Furthermore the resultant neoglycosides have been widely shown to be good mimetics for native glycosidic linkages, as well as more stable to both oxidation and hydrolysis than native glycosides, making the ‘click’ variants easier to handle in the laboratory. Furthermore, neoglycopeptides have been shown to be more resistant to digestive enzymes, thus making the triazole moiety an attractive linker for potential biopharmaceuticals.

Neoglycosides differ from native glycosides by replacement of the native O-linker with a triazole group, for example the native Adpn glycoside Gal-Hyl 1.9 and its neoglycoside analogue 1.10, Figure 3.2. For comparison purposes, neogal-Ser 1.11 was also synthesised and incorporated into Adpn peptides.

![Chemical structures](image)

**Figure 3.2** The chemical structures of the native Adpn glycoside Gal-Hyl 1.9 and the neoglycoside analogue neogal-Lys 1.10 alongside neogal-Ser 1.11.

This thesis concentrates on the synthesis of Adpn peptides incorporating the robust neogal-Hyl functionalised amino acids neogal-Lys 1.10 and neogal-Ser 1.11. As mentioned previously, in native Adpn the PTM-Lys residues are located at positions 68, 71, 80 and 104 in the peptide sequence. Therefore, in the context of this thesis peptides were synthesised corresponding to residues 66 – 83 of Adpn [Adpn(66-83)], thus incorporating three of the four PTM-Lys, Figure 3.3.
Peptides containing 18 amino acids can be synthesised in good yield using Fmoc-SPPS (solid phase peptide synthesis). The addition of a further 20 amino acids needed to incorporate the fourth PTM-Lys residue found natively at position 104 would have required the synthesis of a second peptide fragment necessitating the use of native chemical ligation (NCL) to link the two chains together. These extra synthetic steps would have increased the complexity of the synthesis and reduced the overall yield of the desired ‘clicked’ Adpn peptides and so was not attempted. Instead, the author concentrated on initial synthesis of 18-mer peptides corresponding to residues 66 – 83 of Adpn [Adpn(66-83)].

3.5 Introduction to the Copper-Catalysed 1,3-Dipolar Cycloaddition Reaction, (CuAAC)

The synthesis of neoglycopeptides utilises the Cu(I)-catalysed Azide-Alkyne Cycloaddition reaction (CuAAC, ‘click’ chemistry) to afford a 1,2,3-triazole ring as a model for the native O-glycosidic linkage. The term ‘click’ reaction was defined by Sharpless et al.\textsuperscript{148} as denoting a reaction which is regioselective, high yielding and takes place in easily-removable or biologically-compatible solvents with few side-products.\textsuperscript{148} The CuAAC reaction is now the most commonly used ‘click’ reaction in organic synthesis.\textsuperscript{149} Therefore, the term ‘click’ chemistry is now frequently employed to indicate the CuAAC reaction and will be used as such in this thesis.

3.5.1 The development of 1,3-dipolar addition reactions

The 1,3-dipolar cycloaddition reaction between azides and alkynes was first discovered by A. Michael in 1893.\textsuperscript{150} A. Michael investigated the reaction between phenylazide \textsuperscript{3.1} and dimethyl acetylenedicarboxylate \textsuperscript{3.2} in diethyl ether at 100 °C to afford the 1,2,3-triazole \textsuperscript{3.3}, Scheme 3.2 (reaction yield not reported).
Chapter Three

Reagents and Conditions: Et₂O, 8 h, 100 °C.

Scheme 3.2 The 1,3-dipolar cycloaddition reaction between phenyl azide 3.1 and dimethyl acetylenedicarboxylate 3.2 to afford 1,2,3-triazole 3.3, as reported by A. Michael.

The popularity of 1,3-dipolar additions increased from 1950 onwards, with multiple publications citing reactions of various 1,3-dipoles with unsaturated systems. Examples include the reactions between ozone and alkynes reviewed in 1953 by Lederer et al. including the reaction between acetylene 3.4 and ozone initially conducted by Wohl et al. in 1920, affording formic acid 3.5 as the main product in a reasonable 46% yield, presumably through an ozonide 3.6 followed by a formic anhydride intermediate 3.7, Scheme 3.3.

Reagents and conditions: i) Acetylene 3.4 (1.3 eq.), O₃ (5.0 mmol), gas phase, 15 h, ii) water vapour used to extract product via condensation. Further synthetic details not reported.

Scheme 3.3 The 1,3-dipolar reaction between acetylene and ozone to afford formic acid 3.5 as reported by Wohl et al.

Also notably, in 1950 Herbst et al. reported a series of high temperature reactions between hydrogen azide and various nitriles including acetonitrile to afford monosubstituted tetrazoles such as 5 methyl-tetrazole 3.8 in an excellent yield of 76%, Scheme 3.4.

Reagents and conditions: i) MeCN (0.9 eq.), benzene/iPrOH, 150 °C, 96-120 h.

Scheme 3.4 An example of the reaction between lithium azide and acetonitrile reported by Herbst et al.
3.5.2 The pericyclic reaction mechanism for 1,3-dipolar cycloaddition reactions

In 1963 Huisgen et al.\textsuperscript{154} published a key paper investigating numerous 1,3 cycloaddition reactions and proposing a concerted pericyclic mechanism (see Section 3.4.3) for the reaction. However, although Huisgen et al.\textsuperscript{154} fundamentally increased understanding of 1,3-dipolar cycloaddition reactions, their reaction conditions generated (1:1) mixtures of the 1,4- and 1,5-substituted products, Scheme 3.5. Therefore, the reaction was not widely adopted for use in synthetic organic chemistry.

![Scheme 3.5 General reaction scheme for the Huisgen (uncatalysed) thermal 1,3-cycloaddition reaction affording both 1,4- isomer 3.9 and 1,5- isomer 3.10.\textsuperscript{154}]

3.5.3 Introduction of a copper catalyst to 1,3-dipolar cycloaddition reactions; the birth of the CuAAC reaction

A major breakthrough in the field of 1,3-dipolar cycloaddition reactions came in 2002 when Sharpless et al.\textsuperscript{155} and Meldal et al.\textsuperscript{156} simultaneously pioneered the use of Cu(I) as a catalyst for the reaction. The Cu(I) catalyst used by Sharpless et al. was a mixed CuSO\textsubscript{4}.5H\textsubscript{2}O/sodium ascorbate system, generating the Cu(I) catalyst \textit{in situ}, whereas Meldal et al. employed CuCl, CuBr–dimethyl sulphide complex or CuI as catalysts. Remarkably, use of the Cu(I) catalyst in low catalyst loading (0.25 - 2 mol\%) not only afforded products in high yields (82 -94\%)\textsuperscript{155} and high purity (75 – 99\%)\textsuperscript{155,156} but also produced 1,4-regioisomer 3.9 exclusively, Scheme 3.6.

![Scheme 3.6 General reaction scheme for the Cu(I) catalysed reaction 1,3-cycloaddition reaction affording 1,4-isomer 3.9 exclusively, as reported by Sharpless et al.\textsuperscript{155} and Meldal et al.\textsuperscript{156}]

3.5.4 Proposed mechanisms for the thermal and Cu(I) catalysed cycloaddition

The increased regioselectivity and reaction yields afforded by the CuAAC reaction in comparison to the thermal uncatalysed Huisgen 1,3-cycloaddition reaction can be explained by the different mechanisms for these two types of reaction. The CuAAC reaction occurs through a stepwise, Cu(I) catalysed pathway, whereas the thermal uncatalysed Huisgen 1,3-cycloaddition reaction proceeds through a concerted, pericyclic pathway,\textsuperscript{157} Scheme 3.7.
In the concerted pericyclic mechanism, there is no alignment preference for the cycloaddition, hence the 1,4-tetrazole and 1,5-tetrazole products are formed in roughly a 1:1 ratio. However, in the stepwise Cu(I) pathway, copper acetylide \(3.11\) is formed prior to azide attack. The reagent azide then attacks copper acetylide \(3.11\), displacing a ligand and forming intermediate \(3.12\). In this way, the regiochemistry for the cycloaddition is set. The reaction proceeds onwards via copper metallocycle \(3.13\) which undergoes a ring contraction to afford, after protolysis, the desired 1,4-tetracycle \(3.9\).

The Cu(I) stepwise pathway has also been shown to have a lower activation barrier than the thermal Huisgen pericyclic pathway, illustrated by Himo et al.\(^ {157}\) who found that, in a range of examples, the CuAAC was \(10^7\) times faster than the corresponding uncatalysed thermal Huisgen 1,3-cycloaddition reaction. Meldal et al.\(^ {158}\) proposed that the slow rate of reaction for the uncatalysed thermal cycloaddition reaction is due to the low dipole of the alkyne reagent rendering the alkyne a poor electrophile. In contrast, in the CuAAC reaction the alkyne is converted into copper acetylide \(3.11\) which has a higher dipole moment and is therefore a better electrophile.

### 3.5.5 Mono-copper versus bis-copper pathways for the CuAAC reaction

Computational studies by Fokin et al.\(^ {159}\) and Meldal et al.\(^ {158}\) proposed a mono-copper mechanism via monomeric copper-acetylide \(3.16\) for the CuAAC reaction to afford 1,4-triazole product \(3.17\), Scheme 3.8a). However, (bis) copper intermediates for the CuAAC reaction have been proposed by Fokin et al.\(^ {159}\) and recently isolated by Bertrand et al.\(^ {160}\) adding further insight to the proposed reaction mechanism. In their reported study, Bertrand et al.\(^ {160}\) isolated both (bis)copper-acetylide \(3.18\) as well as (bis)triazole complex \(3.19\), suggesting further complexation of a second copper centre to monomeric copper-acetylide complex \(3.16\) and thus a (bis)copper pathway in operation for the CuAAC reaction, Scheme 3.8b).
Scheme 3.8 The two reaction mechanisms proposed for the Cu(I) catalysed 1,3-cycloaddition reaction between an azide and an alkyne a) mono-copper pathway and b) (bis)copper pathway as reported Bertrand et al.\textsuperscript{160}

3.5.6 Ruthenium catalysis of 1,3-dipolar cycloaddition reactions

In contrast to the CuAAC reaction which affords 1,4-triazole products, such as \textit{3.17}, other catalytic mixes based on Ru(II) catalysts have been developed for selective generation of the 1,5-triazole products, such as \textit{3.20}, Scheme 3.9 as reported by Liskamp et al.\textsuperscript{161}. Liskamp et al.\textsuperscript{161} demonstrated that on the same alkyne substrate \textit{3.21} use of Cu(I) catalysis formed the 1,4-triazole cyclised vancomycin-derivative \textit{3.22}, whereas the use of Ru(II) catalysis formed the 1,5-triazole product \textit{3.20}, Scheme 3.9.

The production of either the 1,4-triazole linked product or the 1,5-diazole linked product selectively according to metal catalyst type allowed Liskamp \textit{et al.} to investigate the effect of different ring sizes resulting from the cycloaddition-cyclisation reaction, as the Cu(I) route afforded the 1,4-triazole 17-atom ring cycle product \textit{3.22}, whereas the Ru(II) route (known as RuAAC, Ru(II)-catalysed Azide-Alkyne Cycloaddition reaction) afforded the 1,5-triazole 16-atom ring cycle product \textit{3.20}. 

\textit{Chapter Three}
Reagents and conditions: i) Cu(CH$_3$CN)$_4$PF$_6$, toluene/MeOH, rt, 24 h; ii) [Cp*RuCl]$_4$, THF/MeOH, 50 °C, 24 h.

Scheme 3.9 Illustration of 1,4-triazole $\text{3.22}$ product produced by Cu(I) catalysis and 1,5-triazole product $\text{3.20}$ generated by Ru(II) catalysis from alkyne substrate $\text{3.21}$, as reported by Liskamp et al.$^{161}$

The catalyst-dependency of the regioisomerism of the 1-3-cycloaddition reaction points to significantly different reaction pathways for the RuAAC and CuAAC reactions. In addition, the CuAAC reaction is limited to terminal alkynes, whereas RuAAC can take place for both terminal and internal alkynes.$^{162}$ These two observations combined lead to the conclusion that the RuAAC reaction goes through a different mechanism to the proposed CuAAC copper-acetylide mechanism. Recent studies of the RuAAC mechanism suggest the reaction proceeds via oxidative coupling of the azide and the alkyne to the Ru(II) centre to form a six-membered ruthenacycle $\text{3.23}$, which collapses by reductive elimination to afford the 1,5-triazole product $\text{3.24}$, Scheme 3.10.$^{163}$
3.5.7 Other transition metal catalysis of 1,3-dipolar cycloaddition reactions

It should also be noted that other transition metal complexes based on Ag(I), Pd(I), Pt(II) and Au(I/II) have been investigated as sources of new catalysts for 1,3-cycloaddition reactions between alkynes and azides, but unfortunately these studies did not afford triazole products in good yields. Cu(I) appears unique in its ability to successfully catalyse these cycloaddition reactions to afford 1,4-triazole selectively. This is perhaps due the ability of Cu(I) to both rapidly exchange ligands as well as coordinate with the alkyne in both a $\sigma$ and $\pi$ manner.

3.6 Use of the CuAAC reaction to generate ‘clicked’ peptides

The robust nature of the CuAAC reaction and subsequent flexibility in terms of solvent and protecting group choice as well as reaction pH, makes the CuAAC reaction very popular within both the fields of organic and as well as peptide synthesis. The CuAAC reaction is especially attractive for the generation of neoglycopeptides. In 2002, Meldal et al. demonstrated the CuAAC reaction of resin-bound alkynes 3.30, with free azides 3.31 at room temperature utilising Cu(I) catalysis to afford resin-bound triazole products such as 3.32. Subsequent cleavage from resin using sodium hydroxide afford free 1,4-triazole products including 3.33 - 3.35 in excellent yields and purity, Scheme 3.11.

**Scheme 3.10** Proposed mechanism for RuAAC reaction via six-membered ruthenacycle 3.23 as reported by Fokin et al.\textsuperscript{165}
Reagents and conditions: PEGA(poly[acryloyl-bis(aminopropyl)polyethylene glycol) resin was utilised with HMBA (4-hydroxymethylbenzoic acid) linker; a) Resin (5 mg) swollen in 200 µL THF, R-N 3.31 (2 eq.), DIPEA (50 eq.), CuI (2 eq.), 25°C, 16 h; bi) Peptidyl resin washed with THF, H2O, THF, ii) 0.1 M NaOH (cleavage agent), rt, 3 h. Products 3.33 – 3.35 examined by HPLC and LCMS to calculate yield and purity.

Scheme 3.11 The first example of an on-resin CuAAC reaction to afford on-resin triazole product 3.32 using Cu(I) as a catalyst, as reported by Meldal et al.

Since this seminal paper by Meldal et al. numerous other research groups have investigated the use of this remarkable reaction in peptide chemistry, which will be discussed in detail below. To date, the CuAAC ‘click’ reaction has been used for numerous applications including:

- The synthesis of conjugate oligonucleotides such as 3.36 and peptides such as 3.37 together to form clicked’ peptide-oligonucleotide conjugates (POC) 3.38, as reported by Strömberg et al.

- Ligation of two or more peptide fragments together for example the synthesis of neurotensin-phosphopeptide 3.39 by ‘clicking’ alkyne-neurotensin derivative 3.40 with Pik1-PBD binding hexaphosphopeptide azide derivative 3.41 as reported by Wuest et al.

- The use of the triazole moiety as a cyclisation handle for the cyclisation of antibiotic vancomycin mimics 3.20 and 3.22 using a triazole bridge by Liskamp et al. (previously illustrated in Scheme 3.9).
Chapter Three

Reagents and conditions: Scheme 3.12a) CuI [3 + 2] dipolar cycloaddition of oligonucleotide 3.36 (0.1 µmol scale) with a peptide-azide derivative 3.37 (4 eq.), CuSO₄ (2.4 eq.), Na ascorbate (6 eq.), tBuOH/H₂O (5 µL, 1:1, v/v), o/n, rt; Scheme 3.12b) alkyne peptide 3.39 (1 eq.), azide peptide (1 eq.), CuSO₄ (0.5 eq.), Na ascorbate (5.0 eq.), buffer pH 8.4, 40 °C, 48 h.

Scheme 3.12 a) Conjugation of oligonucleotide 3.36 and with azide peptide 3.37 using CuAAC to afford ‘clicked’ peptide-oligonucleotide conjugates (POC) 3.38 as reported by Strömberg et al.¹⁶⁴ PS (polystyrene), CPG (controlled pore glass); b) the synthesis of neurotensin-phosphopeptide 3.39 by ‘clicking’ alkyne-neurotensin derivative 3.40 with Pik1-PBD binding hexaphosphopeptide azide derivative 3.41, as reported by Wuest et al.¹⁶⁵

3.7 Use of the CuAAC reaction to afford neoglycopeptides

Rutjes et al.¹⁶⁶ have advanced the field of HIV-1 entry inhibitors by synthesising triazole containing neoglycopeptides of C34 3.42, a segment of the HIV-1 inner envelope glycoprotein gp41, which has been shown to have potent anti-HIV activity. The triazole moiety in neoglycoside 3.43 was used to link
the initial glycan GlcNAc 3.44 with Boc-propargylglycine 3.45. Neoglycoside 3.43 was then incorporated into the peptide chain of C34 using Boc-SPPS to afford neoglycopeptide 3.46, which was subjected to glycosylation conditions using the enzyme Endo-A and glycosylation substrate Man₃ClcNAC-oxazoline 3.47 to afford the glycosylated isostere of C34 3.42, Scheme 3.13. Glycosylated isostere 3.42 was shown to exhibit increased water solubility and more potent anti-HIV activity when compared to the native C34-analogue. This study highlighted the potential of triazole-containing neoglycopeptides in biopharmaceutical discovery programs.

Reagents and conditions: a) GlcNAc 3.44 (2.5 mmol, 1 eq.), Boc-propargylglycine 3.45 (1 eq.), Cu(OAc)₂ (2 eq.), Na ascorbate (4 eq.), t-BuOH (5 mL), rt, o/n; b) Boc-SPPS performed using MBHA (p-methylbenzhydrylamine) resin using HBTU as a coupling agent and N-terminal acetyl capping; c) glycosylation affected using Endo-A enzyme, phosphate buffer, 23 °C, 2 h, 95 %. Further synthetic details not reported.

Scheme 3.13 The use of a triazole linked neoglycopeptide 3.46 to generate C34-anti-HIV therapeutic 3.42.

Other notable examples of the use of the CuAAC reaction to generate biologically active neoglycopeptides include advances by Danishefsky et al.¹⁶⁷ in neoglycoside anti-cancer vaccines by associating a number of carbohydrate targeting antigens (Tn and STn) on to a model carrier peptide via lysine residues using triazole linkers such as in neoglycopeptide 3.48, Scheme 3.14 and the chemoenzymatic approach employed by Walsh et al.¹⁶⁸ for the synthesis of neoglycopeptide antibiotics.
such as 3.51 by enzymatically producing the cyclic decapptide tyrocidine 3.52 with alkyne ‘handles’ followed by ‘clicking’ azido sugar 3.53 to the construct, Scheme 3.15.

\[
\text{Scheme 3.14 The use of CuAAC reaction to form trisazole product 3.48, a potential anti-cancer vaccine, as reported by Danishefsky et al.}^{167}
\]

\[\text{Reagents and conditions: Alkyne substrate 3.49, azido sugar 3.50, Cu wire (as reducing agent), CuSO}_4,\ CH_3CN/H_2O (1:1, v/v), rt, 40 h. Further synthetic details not reported.}\]
Reagents and conditions: Cyclic alkyne peptide 3.52 (27 µL, 5 nM in MeOH), azido sugar 3.53 (10 µL, 50 mM in MeOH), CuSO₄ (5 µL, 20 mM in H₂O), Na ascorbate (10 µL, 20 mM in H₂O), rt, 3 h.

Scheme 3.15 General reaction scheme for the synthesis of neo glycopeptide antibiotics 3.51 by CuAAC reaction between alkyne substrate 3.52 and azido sugar 3.53, as reported by Walsh et al.¹⁶⁸

3.7.1 Previous synthetic studies on the MUC-1 neo glycopeptides

The Brimble group have also developed techniques for the use of a microwave-enhanced CuAAC ‘click’ reaction to produce a range of neo glycopeptide mimics of naturally occurring proteins including fish antifreeze glycopeptides,¹²⁷,¹²⁸,¹³⁴ MUC-1 glycoproteins,¹³⁶,¹³⁹ Pam₂Cys-based lipopeptides,¹³⁸,¹⁴¹ NDPK phosphocarrier peptides,¹⁴⁶ human erythropoietin (EPO)¹⁴⁰ and human amylin.¹⁶⁹,¹⁷⁰ The Brimble group has also been instrumental in combining the CuAAC and native chemical ligation (NCL) reactions into a ‘one-pot’ procedure, which promises to enhance the expedient synthesis of neo glycopeptide libraries.¹³⁵

Of particular relevance to this thesis is the synthesis of a 20-mer neo glycopeptide derivate of MUC-1 3.54, a glycoprotein produced excessively by tumour cells, recently reported by Brimble et al.¹³⁶ This work entailed synthesis of propylargylated peptide 3.55, with ‘alkyne-hooks’ using microwave-enhanced Fmoc-SPPS on ‘in-house’ prepared aminopolystyrene resin utilising the acid-labile hydroxymethylphenoxypropionic acid (HMPP) linker and DIPEA/HBTU as coupling reagents. The resultant propylargylated peptide 3.55 was then subjected to ‘click’ conditions, coupling GalNAc-azide derivative 3.56 to the ‘alkyne hooks’ affording ‘clicked’ neo glycopeptide 3.54, Scheme 3.16.
Reagents and conditions: alkyne peptide 3.55 (3 mM), azido sugar 3.56 (22.5 mM), CuSO₄ (100 mM); TCEP (100 mM), 6 M GnHCl/ 0.2 M Na₂HPO₄ buffer (pH 7), µw, 25 W, 50 °C, 4 h. Synthetic yield not reported.

**Scheme 3.16** Neoglycopeptide synthesis of a 20-mer MUC-1 derivative 3.54 utilising alkyne ‘hooks’ on the peptide 3.55 as ‘click’ handles reported by Brimble *et al.*

### 3.8 Synthetic approach to ‘clicked’ neoglycopeptides of Adpn

#### 3.8.1 Adpn neoglycopeptide design

As discussed in Section 1.4.2, the key PTMs necessary for the bioactivity of adiponectin are hydroxylation and subsequent glycosylation of four key Lys residues located at positions 68, 71, 80 and 104 of the collagenous domain of the protein, illustrated in murine Adpn 3.57. This thesis focuses on the chemical synthesis of the 18-mer neoglycopeptide Adpn(66-83)-Lys-3Gal-OH 3.58, which consists of residues 66 – 83 of murine Adpn 3.57, with neoglycoside neoGly-Lys 1.10 incorporated at Lys positions 68, 71 and 80, Scheme 3.17.
The through-space distance between the sugar moiety and the peptide backbone was also felt to be a possible contributing factor for protein secondary structure and possible bioactivity of the neo-glycopeptides. Modelling of native Gal-Hyl 1.9, neoGal-Lys 1.10 and neoGal-Ser 1.11 using ChemBioDraw Ultra 3D (CambridgeSoft Corporation, subsidiary of PerkinElmer Inc, USA) on MM2-minimised structures allowed calculation of the through-space distance between the anomeric position of the sugar residue and the α-carbon of the amino acid. The native Gal-Hyl 1.9 was found to have a through-space distance of 4.6 Å, whereas neoGal-Lys 1.10 was calculated to have a substantially longer through-space distance of 10.5 Å. In contrast, neoGal-Ser 1.11 was found to have a more similar through-space distance to native Gal-Hyl 1.9 of 7.4 Å, Scheme 3.18. However, it should be noted that simple ChemDraw calculations do not take into account all of the possible conformations of the molecules and thus should only be treated as an approximation.
Scheme 3.18 Through-space distance calculations for native Gal-Hyl 1.9 and neoglycopeptide analogues neoGal-Lys 1.10 and neoGal-Ser 1.11 calculated using ChemDraw 3D on MM2 minimised structures.

Therefore, the serine neoglycoside derivative Adpn(66-83)-Ser-3Gal-OH 3.59 was also selected as a synthetic target, Figure 3.4.

Figure 3.4 Selected neoglycopeptide Adpn(66-83)-Ser-3Gal-OH 3.59.
3.8.2 Retrosynthetic analysis for planned Adpn neoglycopeptides

Drawing inspiration from the microwave-assisted CuAAC reactions reported by Brimble et al.\textsuperscript{136}, discussed in Section 3.6, the initial approach chosen for the synthesis of Adpn neoglycopeptides Adpn(66-83)-Lys-3Gal-OH \textit{3.58} and Adpn(66-83)-Ser-3Gal-OH \textit{3.59} was to carry out ‘on-peptide’ CuAAC reactions, as illustrated in Scheme 3.19. In brief, ‘clicked’ Adpn 18-mers Adpn(66-83)-Lys-3Gal-OH \textit{3.58} and Adpn(66-83)-Ser-3Gal-OH \textit{3.59} were disconnected back to ‘click’ substrates peptides Adpn(66-83)-Lys-azide \textit{3.60} and Adpn(66-83)-Ser-alkyne \textit{3.61} and the corresponding sugar ‘click’ partners HO-Gal-C1-O-alkyne \textit{3.62} and HO-Gal-N\textsubscript{3} \textit{3.63} respectively. It was envisaged that Adpn(66-83)-Lys-azide \textit{3.60} and Adpn(66-83)-Ser-alkyne \textit{3.61} could be synthesised using microwave-enhanced Fmoc-SPPS incorporating Fmoc-Lys-azide \textit{3.64} and Fmoc-Ser-alkyne \textit{3.65} respectively.

\begin{scheme}
\centering
\includegraphics[width=\textwidth]{scheme3.19.png}
\caption{Initial retrosynthetic strategy towards neoglycopeptides Adpn(66-83)-Lys-3Gal-OH \textit{3.58} and Adpn(66-83)-Ser-3Gal-OH \textit{3.59}.}
\end{scheme}
3.9 The synthesis of the amino acid building blocks 3.64 and 3.65 and sugar ‘click’ partners 3.62 and 3.63

As illustrated above in Scheme 3.18, the key amino acid and sugar ‘building blocks’ needed for synthesis of neo-glycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 were identified to be Fmoc-Lys-azide 3.64, Fmoc-Ser-alkyne 3.65, HO-Gal-C1-O-alkyne 3.62 and HO-Gal-N3 3.63. Short, high-yielding synthetic pathways were developed for these amino acid and sugar building blocks. The synthesis of each building block, along with a discussion of the necessary optimisation of the synthetic pathway, will now be outlined in turn in the following sections.

3.9.1 Choice of sugar protecting group

Acetyl protecting groups were chosen for the sugar alcohol groups, due to the ease of removal of acetyl groups using the Zemplén deacetylation procedure using sodium methoxide.\(^{171}\) Therefore, the acetylated versions of the sugar ‘click’ partners HO-Gal-C1-O-alkyne 3.62 and HO-Gal-N3 3.63 were planned; namely AcO-Gal-C1-O-alkyne 3.66 and AcO-Gal-N3 3.67, Scheme 3.20. Notably though, it was envisaged that the on-peptide ‘click’ reactions would be performed with the deacetylated sugar ‘click’ partners HO-Gal-C1-O-alkyne 3.62 and HO-Gal-N3 3.63 as this approach had been shown by Brimble et al.\(^{136}\) to be successful for the synthesis of the 20-mer MUC-1 derivative 3.54.

\[\text{Scheme 3.20 Use of acetyl protecting groups for the synthesis of sugar ‘click’ partners.}\]

3.9.2 Preparation of Fmoc-Lys-azide 3.64

Fmoc-Lys-azide 3.64 was prepared using a known diazotransfer reaction\(^{172}\) from Fmoc-lysine 3.68 under basic reaction conditions, employing imidazole-1-sulfonyl azide hydrochloride as the azido-transfer reagent. The reaction was found to go to completion after five hours at room temperature, with the desired Fmoc-Lys-azide 3.64 product afforded in an excellent yield of 78%, Scheme 3.21.
Reagents and conditions: a) ISA.HCl (imidazole-1-sulfonyl azide hydrochloride, 1.2 eq.), K$_2$CO$_3$ (2.1 eq.), CuSO$_4$ (cat.), MeOH/H$_2$O (50 mL, 49:1, v/v), rt, 5 h.

Scheme 3.21 Synthetic pathway for Fmoc-Lys-azide 3.64.

$^1$H and $^{13}$C NMR data of 3.64 were found to be in agreement with literature values,$^{142}$ with the characteristic $^1$H CH$_2$N$_3$ triplet resonance observed at $\delta$ 3.24, $J = 6.6$ Hz (300 MHz, CDCl$_3$), in accordance with the literature triplet resonance observed at $\delta$ 3.28, $J = 6.8$ Hz (400 MHz, CDCl$_3$)$^{142}$ and the corresponding CH$_2$N$_3$ $^{13}$C resonance was observed at $\delta$ 51.1 (75 MHz), in accordance with the literature resonance observed at $\delta$ 51.1 (100 MHz).$^{142}$

The stereochemistry of Fmoc-Lys-azide 3.64 was confirmed by optical rotation measurement of the purified compound, which was found to be in agreement with the corresponding reported literature value; 3.64 $[\alpha]_D^{23}$ -2.3, $c$ 0.07, methanol; literature$^{172}$ $[\alpha]_D^{23}$ -1.8, $c$ 0.03, methanol. A melting point of 64.1-64.8 °C was recorded for Fmoc-Lys-azide 3.64, but purity of the product could not be assessed on this basis due to lack of a literature reference melting point.

The mechanism for the diazotransfer reaction was established using $^{15}$N NMR labelling experiments by Stevens $et$ $al.$$^{173}$. The mechanism takes place by nucleophilic attack of deprotonated Fmoc-Lys intermediate 3.69 onto the azide group of diazotransfer reagent 3.70. Subsequent tautomerisation of resultant intermediate 3.71 leads to unstable intermediate 3.72, which easily decomposes to afford Fmoc-Lys-azide 3.64, Scheme 3.22.
3.9.3 Preparation of Fmoc-Ser-alkyne 3.65

Fmoc-Ser-alkyne 3.65 was prepared using a modified four step synthesis based on the synthesis reported by Pedersen et al.\textsuperscript{172} outlined below in Scheme 3.23. In brief, treatment of the alcohol in Boc-Ser 3.73 with sodium hydride and propargyl bromide afforded Boc-Ser(O-propargyl)-OH 3.74.\textsuperscript{172} The Boc protecting group was then removed using trifluoroacetic acid (TFA) and the amine re-protected using fluorenylmethoxycarbonyloxy)succinimide (Fmoc-OSu) and potassium carbonate in a solvent mix of water and dioxane to afford 3.65 suitable for incorporation into solid phase peptide synthesis (SPPS) in an excellent 70% yield.

Reagents and conditions: a) NaH (2.2 eq.), propargyl bromide/ toluene (1.1 eq.), DMF, 0°C → rt, o/n, 94%; b)) TFA (neat, excess), 0°C, 1 h, iii) Fmoc-OSu (1 eq.), K$_2$CO$_3$ (10% w/v aq. soln, eq. not reported), dioxane, 0°C → rt, 4 h, 74%.

Scheme 3.23 Synthesis of Fmoc-Ser-alkyne 3.65 as reported by Pedersen et al.\textsuperscript{172}
The main difference between the reported synthesis Fmoc-Ser-alkyne 3.65 by Pedersen et al.\textsuperscript{172} and the synthetic pathway followed in this study is the methyl protection of the carboxylic acid group of intermediate 3.74 using methyl iodide prior to removal of the Boc amine protecting group and re-protection with the Fmoc group. The methylation of the carboxylic acid group of intermediate 3.74 was felt to be necessary to block substitution at the carboxylic acid site during the Fmoc protection step. Deprotection of the methyl ester group of Fmoc-OMe intermediate 3.75 was easily conducted using the LiOH/CaCl\textsubscript{2} mixture reported by Pascal et al.\textsuperscript{174} to afforded the desired Fmoc-Ser-alkyne 3.65 in a good overall yield of 51%, Scheme 3.24. Furthermore, in this project it was found 1.5 equivalents of propargyl bromide/ toluene was needed to obtain good alkylation yields, Scheme 3.24 step a).

\begin{center}
\includegraphics[width=\textwidth]{synthesis_diagram.png}
\end{center}

\textit{Reagents and conditions:} a) NaH (2.3 eq.), propargyl bromide/ toluene (1.5 eq.), DMF, 1.5 h, rt, 75%; b) K\textsubscript{2}CO\textsubscript{3} (1.5 eq.) Mel (1.5 eq.), DMF, rt, o/n, 75%, c)) TFA (excess), CH\textsubscript{2}Cl\textsubscript{2}, rt, 2 h, ii) Fmoc-OSu (1 eq.), Na\textsubscript{2}CO\textsubscript{3} (1.2 eq.), dioxane/H\textsubscript{2}O, rt, o/n, 73% (steps ii and iii); d) 0.8 M CaCl\textsubscript{2}: 0.1 M LiOH (10 eq.), H\textsubscript{2}O, 2 h, rt, 97%.

\textbf{Scheme 3.24 Synthesis of Fmoc-Ser-alkyne 3.65 conducted in this study.}

\textsuperscript{1}H and \textsuperscript{13}C NMR data of Fmoc-Ser-alkyne 3.65 were found to be in agreement with those reported by Pedersen et al.\textsuperscript{172} with the characteristic \textsuperscript{1}H CHC≡ broad singlet resonance observed at δ 2.86 (400 MHz, MeOD), in accordance with the literature triplet resonance observed at δ 2.86, J = 2.5 Hz (400 MHz, MeOD).\textsuperscript{172} Unfortunately, the broad singlet was unable to be resolved into the literature reported triplet, possibly due to poor solvation of the product in deuterated methanol. The two characteristic alkyne resonances were also clearly observed in the \textsuperscript{13}C spectrum of Fmoc-Ser-alkyne 3.65, with the C≡CH resonance observed at δ 78.7 and the C≡CH resonance observed at δ 75.1 (75 MHz, CDCl\textsubscript{3}), within a similar range to the corresponding literature resonances observed at δ 79.8 and 77.4 (100 MHz, DMSO) given the differences in magnetic field and deuterated solvent employed for the \textsuperscript{13}C NMR spectrum.

The stereochemistry of Fmoc-Ser-alkyne 3.65 was confirmed by optical rotation measurement of the purified compound, which was found to be in agreement with the corresponding reported literature value; 3.65 [\alpha]_{D}^{24} +6.5, c 0.62, methanol; literature [\alpha]_{D}^{23} +6.2, c 1.00, methanol.\textsuperscript{172} A melting point of
144.6 – 145.2 °C was recorded for Fmoc-Ser-alkyne 3.65 but the purity of the product could not be assessed on this basis due to lack of a literature reference melting point.

Interestingly, it was observed that if the deprotonation time of Boc-Ser 3.75 was reduced to 10 min, selective alkylation at the carboxylic acid could be effected affording Boc-Ser-O-alkyne, 3.77. The different alkylated products 3.74 and 3.77 could be distinguished by different R_f values (3.74 R_f 0.34\textsuperscript{172}, 3.77 R_f 0.54\textsuperscript{175}, ethyl acetate: acetic acid, 1: 0.01) as well as significant differences in the \textsuperscript{1}H NMR spectra of the compounds. The characteristic CH\textsubscript{2} of the propargyl group resonated at δ 4.13 - 4.17 in 3.74 and at δ 4.77 - 4.78 in 3.77. Figure 3.5.

\textbf{Figure 3.5} \textsuperscript{1}H NMR spectra of a) alkoxide-alkylated product 3.74 and b) carboxylate-alkylated product 3.77.
3.9.4 Preparation of AcO-Gal-N$_3$ 3.67

AcO-Gal-N$_3$ 3.67 was prepared by a standard method involving a three step synthesis; initial acetyl protection of D-galactose 3.78 to afford AcO-Gal 2.184 was followed by bromination at the anomeric position using hydrogen bromide/ acetic acid to produce AcO-Gal-Br 2.60. Acetylated α-galactobromide 2.60 was then treated with sodium azide to afford AcO-Gal-N$_3$ 3.67 in 38% overall yield, Scheme 3.25.

$\text{Conditions and reagents: a)}$ Ac$_2$O (excess), HClO$_4$ (cat.), rt, o/n; b) HBr/AcOH (33%, 10 eq.), CH$_2$Cl$_2$, rt, 4 h, 68%; c) NaN$_3$ (5 eq.), NaHCO$_3$ (excess), Bu$_4$NHSO$_4$ (1 eq.), H$_2$O/CH$_2$Cl$_2$ (1:1, v/v), 56%.

Scheme 3.25 Synthetic route for AcO-Gal-N$_3$ 3.67.

$^1$H and $^{13}$C NMR data of AcO-Gal-N$_3$ 3.67 were found to be in agreement with literature values$^{177}$ with the characteristic doublet resonance corresponding to the anomeric proton observed in the $^1$H spectrum at $\delta$ 4.58, $J$ = 6.6 Hz (300 MHz, CDCl$_3$), in accordance with the doublet resonance observed in the literature at $\delta$ 4.57, $J$ = 8.6 Hz (300 MHz, CDCl$_3$).$^{177}$ The large coupling value of 6.6 Hz also gave confirmation for the stereochemistry of the desired β-anomer. Furthermore, there was also good agreement with between observed and reported $^{13}$C resonances for the anomeric carbon, with an observed resonance at $\delta$ 88.2 (75 MHz, CDCl$_3$) and the reported literature value of $\delta$ 88.3 (125 MHz, CDCl$_3$).$^{177}$

In addition to confirmation by $^1$H NMR, the stereochemistry of AcO-Gal-N$_3$ 3.67 was also established by measurement of the optical rotation of the purified compound, which was found to be in agreement with the corresponding reported literature value; 3.67 $[\alpha]_D$$^{21}$ -14.7, c 2.35, CHCl$_3$; literature $[\alpha]_D$$^{21}$ -10.3, c 1.00, CHCl$_3$.$^{176}$

3.9.5 Preparation of AcO-Gal-C1-O-alkyne 3.66

AcO-Gal-C1-O-alkyne 3.66 was prepared by initial conversion to the anomeric-deprotected intermediate 2.185 and subsequent conversion to trichloroacetimidate 2.183 using trichloroacetonitrile and potassium carbonate in dichloromethane, as discussed in Section 2.13.$^{127}$ Schimdt glycosylation conditions were then employed to convert trichloroacetimidate AcO-Gal-C1-TCA 2.183 to the desired propargylated product AcO-Gal-C1-O-alkyne 3.66$^{178}$ in 61% overall yield, Scheme 3.26.
Reagents and conditions: a) Ethylene diamine (1.1 eq.), AcOH (1.4 eq.), THF, 70%; b) Cl3CCN (4 eq.), K2CO3 (3 eq.), CH2Cl2, rt, o/n, quant.; c) propargyl alcohol (10 eq.), TMSOTf (0.2 eq.), CH2Cl2 (dry), 4Å MS, Ar, -40 °C, 10 min, 87%.

Scheme 3.26 Synthetic route for AcO-Gal-C1-O-alkyne 3.66.

1H and 13C NMR data of AcO-Gal-C1-O-alkyne 3.66 were found to be in agreement with literature values177 with the characteristic doublet resonance corresponding to the anomeric proton observed in the 1H spectrum at δ 4.72, J = 7.9 Hz (400 MHz, CDCl3), in accordance with the doublet resonance observed in the literature at δ 4.72 (300 MHz, CDCl3, coupling constant not reported).179 The large coupling value of 7.9 Hz also gave confirmation for the stereochemistry of the desired β-anomer. Furthermore, there was also good agreement between observed and reported 13C resonances for the anomeric carbon, with an observed resonance at δ 98.6 (100 MHz, CDCl3) and the reported literature value of δ 98.5 (125 MHz, CDCl3).179

In addition to the 1H NMR data, the stereochemistry of AcO-Gal-C1-O-alkyne 3.66 was also established by measurement of the optical rotation of the purified compound, which was found to be in agreement with the corresponding reported literature value; 3.66 [α]D24 -18.6, c 0.61, CHCl3; lit [α]D25 -23.0, c 1.0, CHCl3.179

3.9.5.1 The anomeric effect

For the final step in the synthesis of AcO-Gal-C1-O-alkyne 3.66 several different glycosylation conditions were explored in order to optimise the reaction. Reaction conditions were chosen that enabled selective formation of the desired β-anomer glycoside 3.66 over α-anomer glycoside 3.66b, Figure 3.6. Therefore, it was necessary to use glycosylation conditions that disfavoured the anomeric effect,180 leading to preferential formation of desired β-anomer glycoside 3.66.

The anomeric effect can be explained by two models; the electrostatic model and the molecular orbital model.181 These two models work together to form a complete understanding of the anomeric effect. The electrostatic model involves dipole minimisation and reduced lone pair (LP) repulsion in the α-anomer relative to the β-anomer. The opposite alignment of the dipoles in the α-anomeric configuration minimises the overall dipole of the glycoside, thus lowering the energy of the α-anomer 3.66b relative to β-anomer 3.66.181 Furthermore, there is decreased lone pair electron repulsion in the α-anomeric configuration. The lone pairs on the ring oxygen (O-5) and the lone pairs on the substituted alcohol (C-1)
are aligned away from each other, thus leading to increased stability for α-anomer 3.66b relative to the β-anomer 3.66, Figure 3.6.

α-anomer 3.66b: stabilised due to the anomeric effect

- i) Dipole minimisation
- ii) LP repulsion minimised
- iii) Favourable \( n \rightarrow \sigma^* \) interaction

β-anomer 3.66

- i) Additive dipole
- ii) Increased LP repulsion
- iii) \( N_c \) possible \( n \rightarrow \sigma^* \) interaction

Figure 3.6 The anomeric effect stabilising the α-glycoside 3.66b over desired β-glycoside 3.66 through i) dipole minimisation, ii) LP repulsion minimisation and iii) favourable \( n \rightarrow \sigma^* \) interaction.\(^{181}\)

The molecular orbital model is based on hyperconjugation that takes place in the α-anomer; lone pair delocalisation from the ring oxygen (O-5) to the antibonding orbital of C-1 (\( n \rightarrow \sigma^* \)). This (\( n \rightarrow \sigma^* \)) hyperconjugation lowers the energy of the lone pair and therefore is a stabilising interaction for the α-anomer relative to the β-anomer.

3.9.5.2 Glycosylation techniques that favour the β-anomer

As mentioned previously in Section 2.14.2, the Schmidt glycosylation method\(^{182}\) utilising trichloroacetimidate AcO-Gal-C1-TCA 2.183 as the sugar donor and TMSOTf as the reaction promoter promotes selective formation of the desired β-anomer due the formation of acyloxonium ion 2.189, which blocks the lower face of the glycoside from attack thus promoting formation of the β-anomeric product 3.66 along with the orthoester side product 2.187, Scheme 3.27.
Scheme 3.27 Mechanism for the glycosylation reaction invoking neighbouring group participation (NGP) from C2-acetyl group.

Therefore, Schmidt glycosylation conditions\textsuperscript{182} were tested for the synthesis of AcO-Gal-C1-O-alkyne 3.66, Table 3.1 with the optimised reaction conditions found to be using ten equivalents of propargyl alcohol, 0.2 equivalents of TMSOTf with a reaction temperature of -40 °C and a short reaction time of 10 minutes. These conditions afforded the desired product AcO-Gal-C1-O-alkyne 3.66 in an excellent 87% yield, Scheme 3.28.

Reagents and conditions: a) Propargyl alcohol (10 eq.), TMSOTf (0.2 eq.), CH\textsubscript{2}Cl\textsubscript{2} (dry), 4Å MS, Ar, -40 °C, 10 min, 87%.

Scheme 3.28 Schmidt glycosylation conditions\textsuperscript{182} for the synthesis of AcO-Gal-C1-O-alkyne 3.66.

Table 3.1 Overview of glycosylation conditions attempted for the synthesis of AcO-Gal-C1-O-alkyne 3.66.*

<table>
<thead>
<tr>
<th>Entry</th>
<th>Galactose donor</th>
<th>Glycosylation Promoter</th>
<th>Reaction conditions</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Ac-Gal-TCA, 2.183</td>
<td>TMSOTf, 2 eq.</td>
<td>RT</td>
<td>28%</td>
</tr>
<tr>
<td>2</td>
<td>Ac-Gal-TCA, 2.183</td>
<td>TMSOTf, 0.2 eq.</td>
<td>60 min, -40 °C → -20 °C</td>
<td>53%</td>
</tr>
<tr>
<td>3</td>
<td>Ac-Gal-TCA, 2.183</td>
<td>TMSOTf, 0.2 eq.</td>
<td>15 min, -40 °C → -20 °C</td>
<td>71%</td>
</tr>
<tr>
<td>4</td>
<td>Ac-Gal-TCA, 2.183</td>
<td>TMSOTf, 0.2 eq.</td>
<td>10 min, -40 °C</td>
<td>87%</td>
</tr>
</tbody>
</table>

*In all cases 10 equivalents of propargyl alcohol were used in the reaction.
Other glycosylation methods that were also investigated included those reported by Gurrula et al.,\(^{179}\) which used BF\(_3\).Et\(_2\)O as a reaction promoter for the glycosylation of the β-anomer only of pentacetylgalactose 3.80 (i.e. not the anomeric mixture) with propargyl alcohol to afford AcO-Gal-C1-O-alkyne 3.66 in an excellent yield of 92%, Scheme 3.29. It is reasonable to assume that preference for the β-glycoside 3.66 in these reaction conditions is due to the formation of the acyloxonium ion 2.189, which blocks the lower face of the glycoside from attack thus forming the β-anomer selectively. However, when this method was trialled in this study, only a low 34% yield of AcO-Gal-C1-O-alkyne 3.66 was obtained and therefore this route was abandoned, Table 3.2 Entry 1.

\[
\text{Scheme 3.29} \quad \text{Reported synthesis of AcO-Gal-C1-PA 3.66 from β-AcO-Gal 3.81 as reported by Gurrula et al.}^{179}
\]

Reagents and conditions: a) Propargyl alcohol (1.2 eq.), BF\(_3\).Et\(_2\)O (1.5 eq.), CH\(_2\)Cl\(_2\), rt, 2 h.

The Koenigs-Knorr glycosylation method,\(^{125}\) in which glycosyl halides such as Ac-Gal-Br 2.60 together with silver catalysts such as silver triflate (AgOTf), were also investigated for the synthesis of AcO-Gal-C1-O-alkyne 3.66. The ‘soft’ silver centre of the reaction catalyst forms a bond to the halide leaving group of the glycosyl halide thus enabling facile precipitation of silver halide and facilitating fast formation of the oxocarbenium ion intermediate 2.188. Neighbouring group participation then quickly converts the oxocarbenium ion 2.188 into the acyloxonium ion 2.189 and the reaction progresses as shown in Scheme 3.30. Although this method was found to produce the β-glycoside 3.66 preferentially over the α-glycoside 3.82, the reaction yield was low due to large quantities of the orthoester side-product 2.187 being generated, Table 3.2 Entry 2. This route was also therefore discontinued.

\[
\text{Scheme 3.30} \quad \text{Use of the Koenigs-Knorr glycosylation method}^{125} \text{ for synthesis of AcO-Gal-C1-0-alkyne 3.66 from AcO-Gal-Br 2.60.}
\]

Reagents and conditions: a) Propargyl alcohol (10 eq.), AgOTf (2 x 0.2 eq.), CH\(_2\)Cl\(_2\), 10 min, -40 °C.

The Helferich method\(^{183}\) was also attempted as a possible synthetic method to afford AcO-Gal-C1-O-alkyne 3.66. The Helferich method is very similar to the Koenigs-Knorr glycosylation method,\(^{125}\) except that mercury cyanide is employed as the reaction promoter in preference to a silver
halide. The cyanide ligand of the catalyst is able to attack the acyloxonium ion intermediate 2.189 affording stabilised acyloxonium ion 3.83, which is only vulnerable to attack in the anomeric position, thus avoiding orthoester side product formation, Scheme 3.31.

**Scheme 3.31** The reaction mechanism for the Helferich method\(^\text{183}\) of glycosylation via stabilised acyloxonium ion 3.83.

In this thesis, the Helferich method\(^\text{183}\) of glycosylation was successful affording AcO-Gal-C1-O-alkyne 3.66 in an excellent 82\% yield, Scheme 3.32, Table 3.2; Entries 3 - 6. However, due to the unstable nature of sugar halides such as Ac-Gal-Br 2.60 and the near equimolar quantities of toxic heavy metal salt catalysts required to ensure good reaction yields using the Helferich method,\(^\text{183}\) the Schmidt\(^\text{182}\) glycosylation method is often preferred as it uses the more stable trichloroacetimidate sugar derivatives and only catalytic amounts of mild reaction promoters such as TMSOTf.

**Reagents and conditions:** a) Propargyl alcohol (10 eq.), Hg(CN)\(_2\) (2 eq.), CH\(_2\)Cl\(_2\), o/n, rt.

**Scheme 3.32** Use of the Helferich method\(^\text{183}\) for synthesis of AcO-Gal-C1-O-alkyne 3.66 from AcO-Gal-Br 2.60.

The most successful method for preparing AcO-Gal-C1-O-alkyne 3.66 was therefore found to be the Schmidt\(^\text{182}\) glycosylation method using trichloroacetimidate AcO-Gal-C1-TCA 2.183 as the sugar donor, TMSOTf as the catalyst and a reaction time of 10 min at -40 °C, which afforded AcO-Gal-C1-O-alkyne 3.66 in an excellent 87\% yield.
Table 3.2 Overview of glycosylation conditions attempted for the synthesis of AcO-Gal-C1-O-alkyne 3.66.*

<table>
<thead>
<tr>
<th>Entry</th>
<th>Galactose donor</th>
<th>Glycosylation Promoter</th>
<th>Reaction conditions</th>
<th>Yield</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>β-Ac-Gal-OAc, 3.81</td>
<td>BF₃·Et₂O, 2 eq.</td>
<td>o/n, 0 °C → rt</td>
<td>34%</td>
</tr>
<tr>
<td>2</td>
<td>Ac-Gal-Br, 2.60</td>
<td>AgOTf, 2 x 0.2 eq.</td>
<td>o/n, rt</td>
<td>62%</td>
</tr>
<tr>
<td>3</td>
<td>Ac-Gal-Br, 2.60</td>
<td>Hg(CN)₂, 2 eq.</td>
<td>o/n, rt</td>
<td>8%</td>
</tr>
<tr>
<td>4</td>
<td>Ac-Gal-Br, 2.60</td>
<td>Hg(CN)₂, 2 eq.</td>
<td>Reflux, o/n</td>
<td>60%</td>
</tr>
<tr>
<td>5</td>
<td>Ac-Gal-Br, 2.60</td>
<td>Hg(CN)₂, 2 eq.</td>
<td>µw, 1 h, 80 °C</td>
<td>78%</td>
</tr>
<tr>
<td>6</td>
<td>Ac-Gal-Br, 2.60</td>
<td>Hg(CN)₂, 2 eq.</td>
<td>o/n, rt</td>
<td>82%</td>
</tr>
</tbody>
</table>

*In all cases 10 equivalents of propargyl alcohol were used in the reaction.

3.10 Deprotection of the sugar building blocks prior to on-peptide ‘click’ reaction

With acetylated sugar building blocks AcO-Gal-C1-O-alkyne 3.66 and AcO-Gal-N₃ 3.67 in hand, the Zemplén deacetylation procedure, a very efficient method to effect removal of carbohydrate acetyl groups using freshly prepared sodium methoxide in methanol, was employed to afford ‘click’ building blocks HO-Gal-C1-O-alkyne 3.62 and HO-Gal-N₃ 3.63 respectively. In this manner, AcO-Gal-C1-O-alkyne 3.66 and AcO-Gal-N₃ 3.67 were successfully deacetylated afford HO-Gal-C1-O-alkyne 3.62 in 75% yield and HO-Gal-N₃ 3.63 in 51% yield respectively, Scheme 3.33.

Reagents and conditions: a) Freshly prepared NaOMe (1M) in MeOH, pH 10, rt, 2 h, then Dowex H⁺ beads pH 2, 5 min.


1H and 13C NMR data of HO-Gal-C1-O-alkyne 3.62 and HO-Gal-N₃ 3.63 were found to be in agreement with literature values. For HO-Gal-C1-O-alkyne 3.62, the characteristic anomeric proton resonated as a doublet in the 1H spectrum at δ 4.53, J = 7.9 Hz (400 MHz, D₂O) in agreement with the doublet reported in the literature at δ 4.57, J = 7.8 Hz (300 MHz, D₂O). The large coupling value of 7.9 Hz also confirmed the stereochemistry of the desired β-anomer. There was also good agreement between observed and reported 13C resonances for the anomeric carbon, with an observed resonance at δ101.2 (100 MHz, CDCl₃) and the reported literature value of δ 101.0 (75 MHz, D₂O).

The stereochemistry of HO-Gal-C1-O-alkyne 3.62 was also confirmed by measurement of the optical rotation of the purified compound. The optical rotation value recorded for 3.62 was found to be in
agreement with the reported literature value; 3.62 $[\alpha]_D$ -18.6, $c$ 0.61, CH$_3$OH; literature $[\alpha]_D$ -24.6, $c$ 0.85, CH$_3$OH.$^{187}$

For HO-Gal-N$_3$ 3.63, the characteristic anomic proton resonated as a doublet in the $^1$H spectrum at $\delta$ 4.65, $J$ = 8.6 Hz (400 MHz, D$_2$O) in agreement with the doublet reported in the literature by Percec et al.$^{176}$ at $\delta$ 4.51, $J$ = 8.6 Hz (300 MHz, D$_2$O) and Hackenberger et al.$^{188}$ at $\delta$ 4.63, $J$ = 8.7 Hz (400 MHz, D$_2$O). The large coupling constant of 8.6 Hz also confirmed retention of configuration of the desired $\beta$-anomer. There was also good agreement between the observed and reported $^{13}$C resonances for the anomeric carbon, with an observed resonance at $\delta$ 90.6 (100 MHz, CDCl$_3$) and the literature values of $\delta$ 90.2 (75 MHz, D$_2$O) reported by Percec et al.$^{176}$ and of $\delta$ 90.6 (100 MHz, D$_2$O) reported Hackenberger et al.$^{188}$

Although $^1$H NMR, spectroscopy confirmed the stereochemistry of HO-Gal-N$_3$ 3.63 to be the desired desired $\beta$-anomer, the optical rotation value recorded for 3.63 was found to be in disagreement with the reported literature value 3.63 $[\alpha]_D$ -0.8, $c$ 1.09, CH$_3$OH; literature $[\alpha]_D$ -5.3, $c$ 1.0, CH$_3$OH.$^{176}$ This may be due to possible degradation of the purified azide compound in the tested methanolic solution.

### 3.11 Summary of the synthesis of amino acid and sugar building blocks Fmoc-Lys-azide 3.64, Fmoc-Ser-alkyne 3.65, HO-Gal-C1-O-alkyne 3.62 and HO-Gal-N$_3$ 3.63

In brief, Fmoc-Lys-azide 3.64 was prepared directly from Fmoc-Lys-OH through a known azido-transfer reaction$^{172}$ using imidazole-1-sulfonyl azide hydrochloride (ISA.HCl) in 78% yield.

Fmoc-Ser-alkyne 3.65 was prepared using a four step synthesis from Boc-Ser 3.73 with an overall yield of 51%. Treatment of the alcohol in Boc-Ser 3.73 with sodium hydride and propargyl bromide afforded Boc-Ser(O-propargyl)-OH, 3.74, which was then re-protected as the desired Fmoc derivative Fmoc-Ser-alkyne 3.65.

HO-Gal-C1-O-alkyne 3.62 was synthesised using an optimised five step synthesis from D-galactose in 46% overall yield, with the key reaction being glycosylation of trichloroacetimidate 2.183 with propargyl alcohol utilising TMSOTf as a reaction promoter.

HO-Gal-N$_3$ 3.63 was synthesised using a literature four step synthesis$^{176}$ from D-galactose in 20% overall yield via treatment of AcO-Gal-Br 2.60 with sodium azide.
Reagents and conditions: a) ISA.HCl (imidazole-1-sulfonyl azide hydrochloride, 1.2 eq.), \(K_2CO_3\) (2.1 eq.), \(CuSO_4\) (cat.), MeOH/H\(_2\)O (50 mL, 49:1, v/v), rt, 5 h; b) NaH (2.3 eq.), propargyl bromide/ toluene (1.5 eq.), DMF, 1.5 h, rt, 75%; c) \(K_2CO_3\) (1.5 eq.) Mel (1.5 eq.), DMF, rt, o/n, 75%, d)i) TFA (excess), CH\(_2\)Cl\(_2\), rt, 2 h, ii) Fmoc-OSu (1 eq.), \(Na_2CO_3\) (1.2 eq.), dioxane/H\(_2\)O, rt, o/n, 73% (steps c and d); e) 0.8 M CaCl\(_2\): 0.1 M LiOH (10 eq.), H\(_2\)O, 2 h, rt, 97%; f) ethylene diamine (1.1 eq.), AcOH (1.4 eq.), THF, 70%; g) Cl\(_3\)CCN (4 eq.), \(K_2CO_3\) (3 eq.), CH\(_2\)Cl\(_2\), rt, o/n, quant.; h) propargyl alcohol (10 eq.), TMSOTf (0.2 eq.), CH\(_2\)Cl\(_2\) (dry), 4Å MS, Ar, -40 °C, 10 min, 87%; i) freshly prepared NaOMe (1M) in MeOH, pH 10, rt, 2h, then Dowex H\(^+\) beads pH 2, 5 min; j) HBr/AcOH (33%, 10 eq.), CH\(_2\)Cl\(_2\), rt, 4 h, 68%; k) NaN\(_3\) (5 eq.), NaHCO\(_3\) (excess), Bu\(_4\)NHSO\(_4\) (1 eq.), H\(_2\)O/CH\(_2\)Cl\(_2\) (1:1, v/v), 56%.

**Scheme 3.34** Synthetic schemes for the preparation of amino acid and sugar building blocks Fmoc-Lys-azide 3.64, Fmoc-Ser-alkyne 3.65, HO-Gal-C1-O-alkyne 3.62 and HO-Gal-N\(_3\) 3.63.
3.12 On-resin ‘click’ reactions to generate neo glycopeptides of adiponectin

3.12.1 Synthesis of the ‘clicked’ peptides

3.12.1.1 Introduction to solid phase peptide synthesis

The ‘clicked’ peptides Adpn(66-83)-Lys-3Gal-OH \textbf{3.58} and Adpn(66-83)-Ser-3Gal-OH \textbf{3.59}, Figure 3.7, were prepared using solid phase peptide synthesis (SPPS), a technique developed by Merrifield et al.\textsuperscript{189} in 1963.

\textbf{Figure 3.7} Target neo glycopeptides Adpn(66-83)-Lys-3Gal-OH \textbf{3.58} and Adpn(66-83)-Ser-3Gal-OH \textbf{3.59}.
The guiding principle of SPPS is that the peptide sequence is sequentially built up using coupling reactions onto a solid support, which is known as the ‘resin’. The coupling reactions occur in a single reaction vessel as a two-phase reaction mixture, with the solid phase being the resin support and the solution phase containing the coupling reagents. Excess quantities of the coupling reagents in the solution phase are used to reduce coupling times and increase product yield and are easily removed from the reaction mixture by filtration.

The $\text{N}^\alpha$-amino group of the amino acids can be either Boc\textsuperscript{189} or Fmoc\textsuperscript{190} protected with the amino acid side groups protected either as benzyl (Bn) or tert-butyl (tBu) groups respectively. The Fmoc/tBu protecting group strategy is most commonly employed in preference to the Boc/Bn method, due to the extremely acidic conditions (TFA/ HF) needed for deprotection of final peptide product when the Boc/Bn method is employed.\textsuperscript{191} Therefore, for the synthesis of the peptides discussed in this thesis, the Fmoc/tBu protecting group strategy was chosen.

In brief, a linker molecule (see Section 3.11.1.3) is coupled onto the resin (see Section 3.11.1.2) to functionalise the resin for coupling with the activated carboxyl group of the reacting amino acid. The protected incoming amino acid is activated by a coupling agent (see Section 3.11.1.4) to increase its reactivity towards the $\text{N}$-terminus of the functionalised resin. The activated amino acid \textbf{3.84} undergoes a condensation reaction with the functionalised resin \textbf{3.85} resulting in formation of a new amide bond and the amino acid being joined to the resin, resulting in peptidyl resin \textbf{3.86}. The protecting group of the amino acid is then removed to afford \textbf{3.87} and the coupling cycle repeats until the desired peptide sequence is completed. Upon completion of the peptide sequence, any remaining protecting groups are cleaved from the peptidyl resin \textbf{3.89} and the fully assembled peptide \textbf{3.90} is cleaved from the resin. An illustration of the general method of Fmoc-SPPS is shown below in Scheme 3.35.
3.12.1.2 Solid resin supports utilised in SPPS

The first reported SPPS by Merrifield et al.\textsuperscript{189} utilised a solid support based on polystyrene (PS) cross-linked with divinylbenzene (DVB), commercially known as aminomethylated PS. PS is still commonly used as a solid support for SPPS along with other functionalised solid resins including polyethylene glycol (PEG) based resins such as ChemMatrix and polyethylene glycol-polystyrene (PEG-PS) functionalised resins such as Amino TentaGel.\textsuperscript{192} These different resins have various loading and swelling properties in different solvents, as reported by Shelton et al.\textsuperscript{192} and illustrated in Table 3.3. For the peptides discussed in this thesis aminomethylated PS resin was chosen due to high loading and good swelling in DMF, the chosen reaction solvent.
Table 3.3 Loading and swelling properties of different solid support resins utilised in SPPS as reported by Shelton et al.\textsuperscript{192}. Swelling properties are given for 0.2 mmol/\textit{g} resins.

<table>
<thead>
<tr>
<th>Resin subgroups</th>
<th>Commercial name</th>
<th>Initial loading (mmol/\textit{g})</th>
<th>Approximate swelling (mL/\textit{g})</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS (1 % DVB)</td>
<td>Aminomethylated PS</td>
<td>0.4–1.5</td>
<td>DCM 4, DMF 4, Ether N.A, TFA 2, THF 9, MeOH 1.6</td>
</tr>
<tr>
<td>PEG-PS</td>
<td>Amino TentaGel (TG)\textsuperscript{a}</td>
<td>0.15–0.3</td>
<td>DCM 6, DMF 2, Ether 3.6, N.A 5.0, THF 3.6</td>
</tr>
<tr>
<td>PEG based</td>
<td>Amino PEGA\textsuperscript{b}</td>
<td>0.2–0.4</td>
<td>DCM 13, DMF 11, Ether N.A 16, N.A 13, THF 13</td>
</tr>
<tr>
<td></td>
<td>ChemMatrix (CM)</td>
<td>0.4–0.6</td>
<td>DCM 11, DMF 8, Ether N.A 11, TFA 14, MeOH N.A 9</td>
</tr>
</tbody>
</table>

3.12.1.3 Common linker moieties utilised in SPPS

‘Linker’ molecules are attached to the solid resin and are used to facilitate anchoring of the amino acids to the resin and cleavage of the peptide from the resin upon completion of the peptide sequence. Common linkers employed in Fmoc-SPPS include Wang 3.91, HMPP (hydroxymethylphenoxy propionic acid) 3.92, HMBA (4-hydroxymethylbenzoic acid) 3.93, Rink-amide 3.94 and 2-chlorotrityl chloride 3.95\textsuperscript{192} Figure 3.8. In Fmoc-SPPS, the linker releases the peptide upon treatment with TFA, with linkers 3.91 - 3.93 and 3.95 producing the peptide as a C-terminal acid and the Rink-amide linker 3.94 releasing the peptide as a C-terminal amide.\textsuperscript{192} For the synthesis of Family 1 Adpn peptides Adpn-Lys-azide 3.60 and Adpn-Ser-alkyne 3.61, HMPP linker attached to polystyrene resin 3.92 was used, in which the first residue (Ile) was pre-loaded onto the resin.

\textbf{Figure 3.8} Common linker molecules utilised in Fmoc-SPPS; Wang 3.91, HMPP 3.92, HMBA 3.93, Rink amide 3.94 and 2-chlorotrityl 3.95.\textsuperscript{192}
3.12.1.4 Common coupling agents employed in SPPS

Coupling agents are used to activate the unreactive carbonyl groups of the reacting amino acids present in the solution phase of the reaction mixture, by converting the unreactive carboxyl group to a more reactive moiety such as an anhydride (in the case of carbodiimides 3.96 – 3.97) or an active ester (in the case of benzotriazoles such as 3.98 – 3.104).\textsuperscript{193} The use of a more reactive group on the amino acid to be coupled leads to shorter coupling times and consequently higher yields of the desired peptide sequence.

Common coupling agents include the classical carbodiimides such as DIC 3.96 (\(N,N^{\prime}\)-diisopropylcarbodiimide) and DCC 3.7 (\(N,N^{\prime}\)-dicyclohexylcarbodiimide) as well as more modern benzotriazole coupling agents such as HATU 3.98 ((1-[bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate N-oxide, HBTU 3.99 (1H-benzotriazolium-1-[bis(dimethylamino)methylene]-5chloro-hexafluorophosphate N-oxide, BOP 3.100 (Benzotriazol-1-yl-N-oxyl-tris(dimethylamino)phosphonium hexafluorophosphate), PyBOP 3.101 (benzotriazol-1-yloxytri(pyrrolidino)phosphonium hexafluorophosphate) and more recently BOMP 3.102 (2-(benzotriazol-1-yloxy)-1,3-dimethyl-2-pyrrolidin-1-yl-1,3-diazaphospholidinium hexafluorophosphate), AOP 3.103 (7-azabenzotriazol-1-ylxy)tris(di-methylamino)phosphonium hexafluorophosphate), pyAOP 3.104 (7-azabenzotriazol-1-yloxy)tris(pyrrolidino)phosphonium hexafluorophosphate) and COMU 3.105, Figure 3.9.\textsuperscript{191,194} The use and mechanism of action of both types of coupling agent will be outlined in detail in the following sections.
3.12.1.5 Carbodiimide coupling agents

With the use of carbodiimide coupling agent an intermediate O-acylisourea activated ester, such as 3.106, is formed when the deprotonated amino acid 3.107 reacts with the coupling agent such as DIC 3.96. The highly reactive activated ester 3.106 is much more susceptible to attack by the nucleophilic incoming amino acid 3.107 thus forming coupled product 3.108 more quickly, resulting in higher yields of the desired peptide product.

Scheme 3.36 Activation of deprotonated amino acid 3.105 by DIC 3.99 to form activated ester 3.106.\textsuperscript{191}
However, use of such a highly active ester such as 3.106 can result in epimerisation of the coupled amino acid. The mechanism for racemisation involves base induced formation of oxazolone intermediate 3.109, which be readily deprotonated to form intermediate 3.110 and subsequent loss of stereochemistry in intermediate 3.111,\textsuperscript{194} Scheme 3.37.

Scheme 3.37 Mechanism for racemisation of coupled amino acid in Fmoc-SPPS as reported by Al-Warhi et al.\textsuperscript{194}

3.12.1.6 The use of hydroxylamine auxiliary nucleophiles to reduce epimerisation

Therefore, in order to reduce the risk of epimerisation of the coupled amino acid, auxiliary nucleophiles such as HOBt 3.114 (1-hydroxybenzotriazole), HOAt 3.115 (7-aza-1-hydroxybenzotriazole) and more recently Oxyma 3.116 (ethyl-2-cyano-2-(hydroxyimino)acetate) are often used alongside carbodiimide coupling agents, Figure 3.10.\textsuperscript{194}

![Common hydroxylamine auxiliary nucleophiles used to reduce epimerisation during carbodiimide coupling reactions.\textsuperscript{194}]

Due to the incorporation of a nitrogen atom in the benzene ring, HOAt 3.115 is a more effective coupling reagent than HOBt 3.114 due to the electron withdrawing nature of the nitrogen atom enhancing the leaving group ability of the benzotriazole, as well as additional hydrogen bonding with the incoming
amino acid enabling the coupling reaction.\textsuperscript{194} The new hydroxylamine additive Oxyma \textsuperscript{3.116} was reported by Albericio \textit{et al.}\textsuperscript{195} in 2009, and was shown to have excellent coupling efficiency, to reduce racemisation during coupling as well as a lower explosive risk than HOBt \textsuperscript{3.114} and HOAt \textsuperscript{3.115}.

During the coupling reaction, auxiliary nucleophiles such as HOBt \textsuperscript{3.114}, HOAt \textsuperscript{3.115} and Oxyma \textsuperscript{3.116} react with the highly reactive carbodiimide activated ester \textsuperscript{3.106} to afford less reactive benzotriazole activated esters, such as \textsuperscript{3.117}. The resulting benzotriazole active ester \textsuperscript{3.117} does not easily form the oxazolone intermediate \textsuperscript{3.109} and thus epimerisation is reduced. Furthermore, benzotriazole active ester \textsuperscript{3.117} is able to hydrogen bond through the triazole group with the \(N^\alpha\)-H of the incoming amino group thus facilitating the coupling reaction,\textsuperscript{192} Scheme 3.38.

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Scheme_3.38.png}
\caption{Scheme 3.38 Use of the nucleophilic auxiliary HOBt \textsuperscript{3.112} to direct the stereochemistry of the new amide bond through hydrogen bonding as described by Shelton \textit{et al.}\textsuperscript{192}}
\end{figure}

\textbf{3.12.1.7 Aminium and phosphonium salt coupling agents}

As mentioned previously in Section 3.11.1.4, most modern coupling reagents are based on a benzotriazole core structure such as HOBt \textsuperscript{3.114}. Broadly, these benzotriazole coupling reagents can be categorised into two main groups; aminium salts such as HATU \textsuperscript{3.98} and HBTU \textsuperscript{3.99} and phosphonium salts such as BOP \textsuperscript{3.100}, PyBOP \textsuperscript{3.101}, BOMP \textsuperscript{3.102}, AOP \textsuperscript{3.103} and pyAOP \textsuperscript{3.104}, Figure 3.11.
Figure 3.11 Aminium and phosphonium salt coupling reagents with the common benzotriazole core.\textsuperscript{194}

These triazole containing coupling agents 3.98 – 3.104 facilitate low epimerisation in coupling reactions in a similar manner to that illustrated for HOBr 3.112. Pyrroldidine derivatives PyBOP 3.101 and pyAOP 3.104 were developed to avoid the carcinogenic side-product hexamethylphosphotriamide (O=P(NMe\textsubscript{2})_3) afforded when the BOP 3.100 and AOP 3.103 coupling agents are used.

Aminium salt coupling agents HATU 3.98 and HBTU 3.99 facilitate coupling reactions in a similar manner to the carbodimide coupling reagents, by formation of an activated ester such as 3.119, which is more susceptible to attack by the incoming amino acid 3.106 thus increasing the rate and efficiency of the coupling reaction, Scheme 3.39.
Scheme 3.39 Reaction mechanism for aminium salt coupling agents such as HATU 3.98 via activated ester derivatives such as 3.119.

The reaction mechanism for phosphonium based coupling reagents such as BOP 3.100, PyBOP 3.101, BOMP 3.102, AOP 3.103 and pyAOP 3.104 proceeds via a charged phosphonium intermediate such as 3.120, which eliminates generating the HOBr activated ester 3.117, Scheme 3.40. The reaction mechanism then proceeds as per the HOBr reaction mechanism described in Scheme 3.38.

Scheme 3.40 The reaction mechanism for peptide coupling using phosphonium based coupling reagents such as BOP 3.100.

3.12.1.1 Uronium based coupling agents

The uronium based coupling reagent based coupling agent COMU 3.105 has recently been reported by Albericio et al.\textsuperscript{196}. COMU 3.105 is based on the nucleophilic auxiliary Oxyma 3.116, which was also also developed by Albericio et al.\textsuperscript{196} Figure 3.12. The mechanism of action is thought to be similar to that of BOP 3.100, shown above in Scheme 3.40.

Figure 3.12 The nucleophilic auxiliary Oxyma 3.116 and the coupling agent COMU 3.105 developed by Albericio et al.\textsuperscript{195,196}
In a similar manner to Oxyma 3.116, COMU 3.105 has been shown to have a lower explosive risk and higher solubility in DMF than the benzotriazole coupling agents.196 Furthermore, when COMU 3.105 is used as the coupling agent in conjunction with DIPEA (N,N-diisopropylethylamine) or TMP (2,2,6,6-tetramethylpiperidine) as the base, a simple colour change can be used to monitor the course of the reaction, with a red-yellow colour change observed with DIPEA and pink-colourless with TMP.196

3.12.2 Fmoc-SPPS conditions employed for the synthesis of Family 1 Adpn peptides

Preparation of the Adpn(66-83)-Lys-azide 3.60 and Adpn(66-83)-Ser-alkyne 3.61 peptides were carried out using automated microwave-assisted Fmoc-SPPS136 with incorporation of the Fmoc-Lys-azide 3.64 /Fmoc-Ser-alkyne 3.65 functionalised amino acids to provide the desired functionality. Peptides were prepared on 0.98 g/ mmol loading on “in house”197 synthesised aminomethyl polystyrene (AM-PS) and immobilised via HMPP linker 3.92 with the first amino acid Ile pre-loaded onto the resin. Fmoc deprotection and coupling reactions were carried out with microwave-irradiation using an automated Liberty 12 Microwave Peptide Synthesiser (CEM Corporation, Mathews, NC) and with HATU/DIPEA and piperazine as the coupling and Fmoc deprotection reagents, respectively. For the coupling of building blocks Fmoc-Lys-azide 3.64 /Fmoc-Ser-alkyne 3.65 a longer coupling time of 20 minutes was employed, rather than the usual coupling time of 5 minutes. On completion of synthesis, cleavage of the peptidyl-resin with TFA, TIPS (triisopropylsilane), H2O (95: 2.5: 2.5, v/v) afforded crude peptides, which were purified by RP-HPLC to afford the desired peptides Adpn(66-83)-Lys-azide 3.60 and Adpn(66-83)-Ser-alkyne-3.61 in >95% purity, identified by LC-MS, Scheme 3.41.

The Adpn(66-83)-Lys-azide 3.60 peptide was afforded in 9% yield and was characterised by the [M+H]+ (calculated 1877.9, observed 1876.6) and [M+2H]2+ (calculated 939.6, observed 939.2) charged states, Figure 3.13a.

The Adpn(66-83)-Ser-alkyne 3.61 peptide was afforded in 6% yield and was characterised by the [M+H]+ (calculated 1790.6, observed 1790.4) and [M+2H]2+ (calculated 895.8, observed 895.5) charged states, Figure 3.13b.
Reagents and Conditions: i) 5% (w/v) piperazine + 0.1 M 6-Cl-HOBt in DMF, \( \mu w \), 60 W, 75 °C, 180 s; ii) Fmoc-amino acid (5 eq.), HATU in DMF (0.5 M, 4.6 eq.), DIPEA in NMP (2 M, 10 eq.), 5 min, \( \mu w \), 25 W, 73 °C except for Fmoc-Lys-azide 3.64 or Fmoc-Ser-alkyne 3.65 functionalised amino acids when the following method was employed; functionalised Fmoc-amino acid (2 eq.), HATU in DMF (0.5 M, 2 eq.), DIPEA in NMP (2 M, 4 eq.), 20 min, \( \mu w \), 25 W, 73 °C; iii) 5% (w/v) piperazine + 0.1 M 6-Cl-HOBt in DMF, \( \mu w \), 60 W, 75 °C, 180 s; iv) TFA/TIS/H\( \text{H}_2\text{O} \) (95:2.5:2.5, v/v/v), \( \mu w \), 10 W, 35 °C, 20 min.

**Scheme 3.41** Synthesis of Adpn-Lys-azide peptide 3.60 and Adpn-Ser-alkyne 3.61 peptides using Fmoc-SPPS where \( Y^* \) = Fmoc-Lys-azide 3.64 or Fmoc-Ser-alkyne 3.65 and \( X^* \) = activated ester leaving group.
Adpn(66-83)-Lys-azide 3.60

Figure 3.13 a) Adpn(66-83)-Lys-azide 3.60 chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Agilent C3 SB-300, 150 mm x 3.0 mm) at a flow rate of 0.3 mL/min, linear gradient 5% B – 65% B over 21 min, where buffer A = 0.1% TFA in H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN.
Figure 3.13 b) Adpn(66-83)-Ser-alkyne 3.61 chemical structure and RP-HPLC trace, mass spectrum. The analytical RP-HPLC was performed using an analytical column (Agilent C3 SB-300, 150 mm x 3.0 mm) at a flow rate of 0.3 mL/min, linear gradient 5% B – 65% B over 21 min, where buffer A = 0.1% TFA in H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN.
3.12.3 Studies for the on-peptide ‘click’ reactions

As outlined in Section 3.7.2, the synthetic strategy to prepare the target neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 used ‘click’ handles on the peptide fragments to perform ‘on-peptide’ CuAAC ‘click’ reactions. The azide ‘click’ handles present in Adpn(66-83)-Lys-azide 3.60 were then used to prepare Adpn(66-83)-Lys-3Gal-OH 3.58 and the alkyne ‘click’ handles present in Adpn(66-83)-Ser-alkyne 3.61 were then used to afford Adpn(66-83)-Ser-3Gal-OH 3.59, Schemes 3.42 - 3.43.

Scheme 3.42 Planned synthetic strategy to prepare neoglycopeptide Adpn(66-83)-Lys-3Gal-OH 3.58 using the CuAAC ‘click’ reaction.
Initial studies using copper sulphate/sodium ascorbate catalyst systems to ‘click’ the sugar moieties onto the azide/alkyne handles present in Adpn(66-83)-Lys-azide peptide 3.60 and Adpn(66-83)-Ser-alkyne peptide 3.61 afforded only trace quantities of the ‘mono-clicked’ product in preference to the desired ‘triple-clicked’ product, even after using extended reaction times of up to 16 hours and thermal heating of reaction mixtures with temperatures up to 50 °C, Table 3.4. Both DMSO (dimethyl sulfoxide) and 6 M aq. guanidine hydrochloride (Gn.HCl) were evaluated as solvents for the CuAAC reaction. In this initial study, the small traces of mono-‘clicked’ products were observed by flow-inject mass spectrometry, Figure 3.14.

The ‘mono-clicked’ Adpn(66-83)-Lys-1Gal-OH peptide 3.120 was characterised by the [M+2H]\(^{2+}\) (calculated 1048.5, observed 1047.9) and [M+3H]\(^{3+}\) (calculated 699.3, observed 699.4) charged states, Figure 3.14a).

The ‘mono-clicked’ Adpn(66-83)-Ser-1Gal-OH peptide 3.121 was characterised by the [M+2H]\(^{2+}\) (calculated 998.4, observed 997.8) and [M+3H]\(^{3+}\) (calculated 665.9, observed 665.4) charged states, Figure 3.14b).
Table 3.4 Reaction conditions trialled for the CuAAC reaction of Adpn(66-83)-Lys-azide peptide 3.60 with HO-Gal-C1-O-alkyne 3.62 and Adpn(66-83)-Ser-alkyne peptide 3.61 with HO-Gal-N₃ 3.63. Numbers of equivalents of reagents given per ‘click’ handle of the corresponding peptide.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Adpn peptide</th>
<th>Solvent</th>
<th>Sugar</th>
<th>Eq. sugar</th>
<th>Eq. CuSO₄</th>
<th>Eq. Na ascorbate</th>
<th>Reaction conditions</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adpn(66-83)-Lys-azide peptide 3.60</td>
<td>Gn.HCl (aq. 6 M)</td>
<td>HO-Gal-C1-O-alkyne 3.62</td>
<td>1.1</td>
<td>1.6</td>
<td>1.6</td>
<td>o/n, thermal heating, 40 °C</td>
<td>Trace ‘mono-clicked’ product</td>
</tr>
<tr>
<td>2</td>
<td>Adpn(66-83)-Lys-azide peptide 3.60</td>
<td>DMSO/H₂O (1:1, v/v)</td>
<td>HO-Gal-C1-O-alkyne 3.62</td>
<td>1.1</td>
<td>3.2</td>
<td>1.6</td>
<td>1 h, thermal heating 50 °C</td>
<td>Trace ‘mono-clicked’ product</td>
</tr>
<tr>
<td>3</td>
<td>Adpn(66-83)-Lys-azide peptide 3.60</td>
<td>DMSO/H₂O (1:1, v/v)</td>
<td>HO-Gal-C1-O-alkyne 3.62</td>
<td>1.3</td>
<td>6.6</td>
<td>3.2</td>
<td>1 h, thermal heating 50 °C</td>
<td>Trace ‘mono-clicked’ product</td>
</tr>
<tr>
<td>4</td>
<td>Adpn(66-83)-Ser-alkyne peptide 3.61</td>
<td>Gn.HCl (aq. 6 M)</td>
<td>HO-Gal-N₃ 3.63</td>
<td>1.1</td>
<td>1.6</td>
<td>1.6</td>
<td>o/n, thermal heating, 40 °C</td>
<td>Trace ‘mono-clicked’ product</td>
</tr>
<tr>
<td>5</td>
<td>Adpn(66-83)-Ser-alkyne peptide 3.61</td>
<td>DMSO/H₂O (1:1, v/v)</td>
<td>HO-Gal-N₃ 3.63</td>
<td>1.1</td>
<td>3.2</td>
<td>1.6</td>
<td>1 h, thermal heating, 50 °C</td>
<td>Trace ‘mono-clicked’ product</td>
</tr>
<tr>
<td>6</td>
<td>Adpn(66-83)-Ser-alkyne peptide 3.61</td>
<td>DMSO/H₂O (1:1, v/v)</td>
<td>HO-Gal-N₃ 3.63</td>
<td>1.3</td>
<td>6.6</td>
<td>3.2</td>
<td>1 h, thermal heating, 50 °C</td>
<td>Trace ‘mono-clicked’ product</td>
</tr>
</tbody>
</table>

Figure 3.14 Flow-inject mass spectra for a) ‘mono-clicked’ Adpn(66-83)-Lys-1Gal-OH peptide 3.120 and b) ‘mono-clicked’ Adpn(66-83)-Ser-1Gal-OH peptide 3.121.
Due to the failure of the CuSO$_4$/ Na ascorbate system to afford the desired ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH $3.58$ and Adpn(66-83)-Ser-3Gal-OH $3.59$, the reductant tris(2-carboxyethyl)phosphine hydrochloride (TCEP.HCl) was then used as the reductant for the CuAAC reactions of Adpn(66-83)-Lys-azide peptide $3.62$ and Adpn(66-83)-Ser-alkyne peptide $3.63$

Following the conditions previously reported by Brimble et al.$^{136}$ for the successful synthesis of MUC-1 neoglycopeptides, a CuSO$_4$/ TCEP.HCl system was used to generate the required Cu(I) catalyst, with the solvent system of aq. 6 M GnHCl/ Na$_2$HPO$_4$ at pH 7.1, microwave irradiation and an elevated temperature. Gratifyingly, elevation of the reaction temperature to 80 °C and use of additional Cu(I) reagent (50 eq.) and sugar reagent (5 eq.) afforded the desired ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH $3.58$ and Adpn(66-83)-Ser-3Gal-OH $3.59$ after 3 hours of heating with microwave irradiation; Table 3.5, Entries 3 and 6. The neoglycopeptides Adpn(66-83)-Lys-3Gal-OH $3.58$ and Adpn(66-83)-Ser-3Gal-OH $3.59$ were characterised by electrospray ionisation (ESI) mass spectrometry, Figures 3.15 – 3.16.

The structure of the ‘triple-clicked’ Adpn(66-83)-Lys-3Gal-OH peptide $3.58$ was confirmed by ESI mass spectrometry; [M+2H]$^{2+}$ (calculated 1266.47, observed 1266.56) and [M+3H]$^{3+}$ (calculated 844.71, observed 844.71) charged states, Figure 3.15.

The structure of the ‘triple-clicked’ Adpn(66-83)-Ser-3Gal-OH peptide $3.59$ was confirmed by ESI mass spectrometry; [M+2H]$^{2+}$ (calculated 1203.40, observed 1203.64) charged state, Figure 3.16.

However, due to the excess quantities of both Cu(I) reagent (in total 150 eq.) and sugar reagents (in total 25 eq.) required for successful synthesis of the desired neoglycopeptides Adpn(66-83)-Lys-3Gal-OH $3.58$ and Adpn(66-83)-Ser-3Gal-OH $3.59$ as well as the large number of side-products observed during the purification process, this synthetic method was abandoned in favour of using ‘pre-clicked’ building blocks and their subsequent incorporation into the desired Adpn sequence as discussed in Section 3.13.
**Chapter Three**

Table 3.5 Reaction conditions trialled for the CuAAC reaction of Adpn(66-83)-Lys-azide peptide 3.60 with HO-Gal-C1-O-alkyne 3.62 and Adpn(66-83)-Ser-alkyne peptide 3.61 with HO-Gal-N3 3.63 utilising microwave irradiation and elevated temperatures. Numbers of equivalents of reagents given per ‘click’ handle of the corresponding peptide.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Adpn peptide</th>
<th>Solvent</th>
<th>Sugar</th>
<th>Eq. sugar</th>
<th>Eq. CuSO4</th>
<th>Eq. TCEP.HCl</th>
<th>Reaction conditions</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adpn(66-83)-Lys-azide peptide 3.60</td>
<td>Gn.HCl, Na2HPO4, pH 7.1</td>
<td>HO-Gal-C1-O-alkyne 3.62</td>
<td>1.5</td>
<td>6.6</td>
<td>6.6</td>
<td>3 h, µw, 70 °C</td>
<td>Trace ‘triple-clicked’ product</td>
</tr>
<tr>
<td>2</td>
<td>Adpn(66-83)-Lys-azide peptide 3.60</td>
<td>Gn.HCl, Na2HPO4, pH 7.1</td>
<td>HO-Gal-C1-O-alkyne 3.62</td>
<td>3</td>
<td>30</td>
<td>30</td>
<td>3 h, µw, 70 °C</td>
<td>Trace ‘triple-clicked’ product</td>
</tr>
<tr>
<td>3</td>
<td>Adpn(66-83)-Lys-azide peptide 3.60</td>
<td>Gn.HCl, Na2HPO4, pH 7.1</td>
<td>HO-Gal-C1-O-alkyne 3.62</td>
<td>15</td>
<td>50</td>
<td>50</td>
<td>3 h, µw, 80 °C, additional sugar 3.62 (5 eq.) and catalyst (50 eq.) loading each hour</td>
<td>‘Triple-clicked’ product</td>
</tr>
<tr>
<td>4</td>
<td>Adpn(66-83)-Ser-alkyne peptide 3.61</td>
<td>Gn.HCl, Na2HPO4, pH 7.1</td>
<td>HO-Gal-N3 3.63</td>
<td>1.5</td>
<td>6.6</td>
<td>6.6</td>
<td>3 h, µw, 70 °C</td>
<td>‘Mono-clicked’ product</td>
</tr>
<tr>
<td>5</td>
<td>Adpn(66-83)-Ser-alkyne peptide 3.61</td>
<td>Gn.HCl, Na2HPO4, pH 7.1</td>
<td>HO-Gal-N3 3.63</td>
<td>3</td>
<td>30</td>
<td>30</td>
<td>2 h, µw, 80 °C</td>
<td>‘Di-clicked’ product</td>
</tr>
<tr>
<td>6</td>
<td>Adpn(66-83)-Ser-alkyne peptide 3.61</td>
<td>Gn.HCl, Na2HPO4, pH 7.1</td>
<td>HO-Gal-N3 3.63</td>
<td>15</td>
<td>50</td>
<td>50</td>
<td>3 h, µw, 80 °C, additional sugar 3.63 (5 eq.) and catalyst (50 eq.) loading each hour</td>
<td>‘Triple-clicked’ product</td>
</tr>
</tbody>
</table>
Chapter Three

Figure 3.15 a) Overlaid HPLC traces for 1h, 2h and 3 h aliquots of the CuAAC reaction of Adpn(66-83)-Lys-azide peptide 3.60 with HO-Gal-C1-O-alkyne 3.62 to afford triple-clicked’ Adpn(66-83)-Lys-3Gal-OH peptide 3.58 using Cu(I) (150 eq.), sugar reagent 3.62 (15 eq.), 6 M GnHCl/ 0.2 M Na$_2$HPO$_4$, 80 °C, µw, Ar, 3 h, additional catalyst and sugar donor loading each hour, HPLC run using analytical column Gemini 5 µm C18 110 Å, 150 mm x 4.6 mm, 0.2 mL/min, linear gradient 5% B – 65% B over 20 min, where buffer A = 0.1% TFA in H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN; b) ESI-mass spectrum for ‘triple-clicked’ Adpn(66-83)-Lys-3Gal-OH peptide 3.58.

a) Triple-clicked’ Adpn(66-83)-Lys-3Gal-OH peptide 3.58
b) ‘Triple-clicked’ Adpn(66-83)-Ser-3Gal-OH peptide 3.59

Figure 3.16 a) Overlaid HPLC traces for 1 h, 2 h and 3 h aliquots of the CuAAC reaction of Adpn(66-83)-Ser-alkyne peptide 3.61 with HO-Gal-N₃ 3.63 to afford triple-clicked’ Adpn(66-83)-Ser-3Gal-OH peptide 3.59 using Cu(I) (150 eq.), sugar reagent 3.63 (15 eq.), 6 M GlnHCl/ 0.2 M Na₂HPO₄, 80 °C, µw, Ar, 3 h, additional catalyst and sugar donor loading each hour, HPLC run using analytical column Gemini 5 µm C18 110 Å, 150 mm x 4.6 mm, 0.2 mL/min, linear gradient 5% B – 65% B over 20 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN; b) ESI-mass spectrum for ‘triple-clicked’ Adpn(66-83)-Ser-3Gal-OH peptide 3.59.
Key conclusions drawn from the on-peptide ‘click’ approach to the synthesis of ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 will be discussed in the following Section 3.12. Section 3.13 will then outline the synthetic pathway used to prepare the desired ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 and the subsequent incorporation of these ‘pre-clicked’ building blocks into neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 using Fmoc-SPPS.

3.13 Conclusions drawn from the on-peptide ‘click’ approach to the synthesis of neoglycopeptide derivatives of Adpn

Although the on-peptide ‘click’ synthetic method was successful in affording ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59, the method was found to be inefficient and low yielding, particularly due to the large excess of both the Cu(I) catalyst and sugar reagents required to push the reaction to completion. These extreme conditions could have been necessitated due to the inaccessibility of the ‘click’ reaction sites due to peptide folding in solution and/or the carbon chain length of the ‘click’ handle. These two possible explanations are discussed in turn below.

3.13.1 Accessibility of ‘click’ reactions sites

Considerably harsher reaction conditions were required to afford the ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 from their corresponding azide/alkyne peptide starting materials Adpn(66-83)-Lys-azide peptide 3.60 and Adpn(66-83)-Ser-alkyne peptide 3.61, compared to the synthesis of ‘triple-clicked’ MUC-1 neoglycopeptide 3.54 previously reported by Brimble et al.\textsuperscript{136} in Section 3.6. Whilst ‘triple-clicked’ MUC-1 neoglycopeptide 3.54 could be successfully synthesised using 1.5 equivalents of sugar azide 3.54 and 6 equivalents of Cu (I) catalyst per ‘click’ site after 3 hours of heating at 50 °C under microwave irradiation, synthesis of the equivalent ‘triple-clicked’ neoglycopeptide Adpn(66-83)-Ser-3Gal-OH 3.59 required considerably harsher reaction conditions using 25 eq. of sugar azide 3.63, 150 eq. of Cu (I) catalyst per ‘click’ site and heating at 80 °C for 3 hours under microwave irradiation.

One possible explanation for the harsher reactions needed to afford the ‘triple-clicked’ neoglycopeptide Adpn(66-83)-Ser-3Gal-OH 3.59 compared to the ‘triple-clicked’ MUC-1 neoglycopeptide 3.54 could be the accessibility of the alkyne ‘click’ sites of the corresponding alkyne peptides to CuAAC reaction. Indeed, the MUC-1 alkyne peptide 3.55 and Adpn(66-83)-Ser-alkyne peptide 3.61 have different locations and frequency of proline/ hydroxyproline residues as illustrated in Figure 3.17.
As proline and hydroxyproline are uniquely able to form either a \textit{cis}- or \textit{trans}- peptide bond, the location and frequency of these two amino acids in a peptide sequence are often fundamental in guiding the secondary structure for the protein.\textsuperscript{133} As Figure 3.17 illustrates, the MUC-1 alkyne peptide \textbf{3.55} contains five proline residues, whereas the Adpn(66-83)-Ser-alkyne peptide \textbf{3.61} contains only two hydroxyproline residues. Therefore, one could expect that the secondary structures adopted by these two alkyne-peptides could be sufficiently different in the chosen reaction solvent (6 M aq. Gn.HCl, 0.2 M Na\textsubscript{2}HPO\textsubscript{4}, pH 7.1). In order to test this theory, both the MUC-1 alkyne peptide \textbf{3.55} and the Adpn(66-83)-Ser-alkyne peptide \textbf{3.61} were modelled using MM2 minimised structures in ChemBioDraw Ultra 3D (v15.0, Cambridge Soft Corporation, Perkin Elmer Informatics Inc), Figure 3.18.

As Figure 3.18 illustrates, the three alkyne groups present in the MUC-1 alkyne peptide \textbf{3.55} point away from the peptide chains and are not sterically encumbered by neighbouring atoms. In contrast, two out of the three alkyne groups present in the Adpn(66-83)-Ser-alkyne peptide \textbf{3.61} are close to neighbouring atoms. Therefore, the alkyne ‘click’ handles of the Adpn(66-83)-Ser-alkyne peptide \textbf{3.61} could be less vulnerable to attack in the performed CuAAC reaction. Thus, the CuAAC reaction to afford ‘triple-clicked’ MUC-1 neoglycopeptide \textbf{3.54} could proceed more easily than the corresponding CuAAC reaction to afford ‘triple-clicked’ \textit{neoglycopeptide} Adpn(66-83)-Ser-3Gal-OH \textbf{3.59}. 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.17.png}
\caption{The chemical structures of MUC-1 alkyne peptide \textbf{3.55} and Adpn(66-83)-Ser-alkyne peptide \textbf{3.61} with the proline and hydroxyproline residues highlighted in blue.}
\end{figure}
Figure 3.18 MM2 minimised ChemBioDraw Ultra 3D structures of the 'triple-clicked' substrates for a) MUC-1 alkyne peptide 3.55 and b) Adpn(66-83)-Ser-alkyne peptide 3.61 with the alkyne 'click' handles highlighted using white ellipses.
3.13.2 Comparison of lysine-azide and serine-propargyl ‘click’ substrates

During this study, it was noted that the Adpn(66-83)-Lys-azide peptide 3.60 was more reactive towards the ‘click’ conditions employed than the corresponding Adpn(66-83)-Ser-alkyne peptide 3.61. A reasonable explanation for this observation could be the longer carbon chain length of the lysine azide ‘click’ handle compared to the serine-propargyl ‘click’ handle, thus rendering the Adpn(66-83)-Lys-azide peptide 3.60 more reactive than the Adpn(66-83)-Ser-alkyne peptide 3.61, Figure 3.19.

![Adpn(66-83)-Lys-azide peptide](image1)

![Adpn(66-83)-Ser-alkyne peptide](image2)

**Figure 3.19** Comparison of the chemical structures of Adpn(66-83)-Lys-azide peptide 3.60 and Adpn(66-83)-Ser-alkyne peptide 3.61 illustrating the longer carbon chain length of azide ‘click’ handle.

3.14 Neoglycopeptides of adiponectin generated through ‘pre-clicked’ amino acid-sugar building blocks prior to peptide synthesis

3.14.1 Previous literature examples

Due to the harsh reaction conditions required to afford the ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 a new synthetic approach was sought. It was decided to generate ‘pre-clicked’ building blocks of the amino acid and sugar donors and to subsequently incorporate these ‘pre-clicked’ building blocks into the adiponectin sequence using Fmoc-SPPS. It is worth noting that other groups have also reported problems with on-peptide ‘click’ reactions. Rutjes *et al.* 166 who, although previously successful with several on-peptide ‘click’ reactions, were unable to effect the on-peptide ‘click’ reaction of the anti-HIV-C34-alkyne peptide 3.121 with GalNAc-azide derivative 3.122 using a CuI/sodium ascorbate catalyst system, Scheme 3.44.
Scheme 3.44 Failed on-peptide ‘click’ reaction with anti-HIV-C34 alkyne peptide 3.121 GalNAc-azide 3.122 as reported by Rutjes et al.\textsuperscript{166}

Therefore, Rutjes et al.\textsuperscript{198} developed a simple synthetic technique for ‘pre-clicking’ amino acids and sugar moieties together using Cu(OAc)$_2$ and sodium ascorbate as the catalyst system. Brimble et al.\textsuperscript{128} have also reported the use of a CuSO$_4$/Na ascorbate catalyst system to generate ‘pre-clicked’ building blocks joining amino acids and sugar moieties together.

3.14.2 The synthesis of ‘clicked’ neogalactolysine and neogalactoserine

The revised synthetic strategy towards the synthesis of ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 therefore focused on the synthesis of ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 and their subsequent incorporation into the Adpn peptide sequence using Fmoc-SPPS.

The ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 were easily disconnected back to previously prepared reagents Fmoc-Lys-azide 3.64 and AcO-Gal-C1-O-alkyne 3.66 and Fmoc-Ser-alkyne 3.65 and AcO-Gal-N$_3$ 3.67 respectively, Scheme 3.45.

Based on the reaction conditions published by Brimble et al.\textsuperscript{134}, a CuSO\(_4\)/sodium ascorbate catalytic system was chosen for the CuAAC reactions to afford ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124. The CuAAC reactions were run in a degassed ethanol/water mixture at 80 °C with 1.1 equivalents of the sugar reagent and afforded the desired ‘pre-clicked’ product in excellent yield after thermal heating for 1 h.
Using these conditions, Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 was produced in 92% yield from Fmoc-Lys-azide 3.64 and AcO-Gal-C1-O-alkyne 3.66. Similarly, Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 was afforded in quantitative yield from Fmoc-Ser-alkyne 3.65 and AcO-Gal-N₃ 3.67, Scheme 3.46. The ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 were purified by flash column chromatography and were found to be stable at room temperature and soluble in a range of organic and protic solvents.

Reagents and conditions: Sugar donor AcO-Gal-N₃ 3.67 or AcO-Gal-C1-O-alkyne 3.66 (1.1 eq.), CuSO₄·5H₂O (0.3 eq.), sodium ascorbate (0.6 eq.), ethanol: water (2:5, v/v, degassed), 80 °C, 1 h.


The ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 were characterised using NMR, IR, optical rotation, melting point and high resolution ESI-mass spectrometry measurements.

For Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 the [M+H]⁺ charged state was observed at m/z 803.2757 (calc. 803.2746) and the optical rotation recorded as [α]D²⁴⁻4.9 (c 0.30, MeOH). In the ¹H NMR spectrum, the characteristic triazole proton was observed as a singlet at δ 7.53 – 7.60. The anomeric proton was observed as a doublet at δ 4.62, with a coupling constant J = 7.8 Hz, establishing formation...
of the desired β-anomer. In the $^{13}$C NMR spectrum for Fmoc-Lys-(Click-Gal-OAc)-OH 3.123, the characteristic triazole carbons resonated at δ 122.9 (C=CHN) and δ 143.7 (C=CHN) and the anomeric carbon resonated at δ 100.4.

For Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 the [M+H]$^+$ charged state was observed at m/z 761.2268 (calc. 761.2277) and the optical rotation recorded as $[\alpha]_D^{24}$ -3.3 (c 1.30, MeOH). In the $^1$H NMR spectrum, the characteristic triazole proton was observed as a singlet at δ 7.82. The anomeric proton was observed as a doublet at δ 5.81, with a coupling constant $J = 9.2$ Hz, establishing formation of the desired β-anomer. In the $^{13}$C spectrum for Fmoc-Ser-(Click-Gal-OAc)-OH 3.124, the triazole carbons resonated at δ 121.5 (C=CHN) and δ 141.2 (C=CHN) and the anomeric carbon was resonated at δ 86.3.

### 3.14.3 Incorporation of neogalactolysine and neogalactoserine building blocks into Adpn peptides using Fmoc SPPS

‘Triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 were next prepared in an analogous fashion to the Adpn(66-83)-Lys-azide 3.60 and Adpn(66-83)-Ser-alkyne 3.61 peptides discussed in Section 3.11.2.

In brief, the synthesis was carried out using automated microwave-assisted Fmoc-SPPS$^{136}$ with incorporation of the Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 / Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 functionalised amino acids to provide the desired functionality. Peptides were prepared utilising 0.98 g/mmol loading on “in house”$^{197}$ synthesised aminomethyl polystyrene (AM-PS) and immobilised via HMPP linker 3.92. Fmoc deprotection and coupling reactions were carried out with microwave-irradiation using an automated Liberty 12 Microwave Peptide Synthesiser (CEM Corporation, Mathews, NC) and with HATU/DIPEA and piperazine as the coupling and Fmoc deprotection reagents, respectively. For the coupling of the ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 / Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 a longer coupling time of 20 minutes was employed, rather than the usual coupling time of 5 minutes. Upon completion of the synthesis, cleavage of the peptidyl-resin with TFA/ DODT (2,2′-(ethylenedioxy)diethanethiol)/ H$_2$O/ TIPS (94:2.5:2.5:1.5, v/v) afforded the crude peptides, which were purified by RP-HPLC to afford the desired O-acetylated neoglycopeptides Adpn(66-83)-Lys-3Gal-OAc 3.58a and Adpn(66-83)-Ser-3Gal-OAc 3.59a in >95% purity, identified by HPLC and LC-MS, Scheme 3.47.

The O-acetylated Adpn(66-83)-Lys-3Gal-OAc 3.58a neoglycopeptide was afforded in 1.0% overall yield and was characterised by the [M+2H]$^{2+}$ (calculated 1518.5, observed 1518.4) charged state, Figure 3.20.
The O-acetylated Adpn(66-83)-Ser-3Gal-OAc 3.59a neoglycopeptide was afforded in 1.4% overall yield and was characterised by $[\text{M+2H}]^{2+}$ (calculated 1455.9, observed 1455.5) and $[\text{M+3H}]^{3+}$ (calculated 970.9, observed 970.6) charged states, Figure 3.21.

\begin{equation}
\text{Fmoc-NH-Ile} \quad \text{HMPP} \quad 3.125
\end{equation}

\begin{equation}
\text{i) Fmoc deprotection}
\end{equation}

\begin{equation}
\text{NH}_2\text{-Ile} \quad \text{HMPP} \quad 3.126
\end{equation}

\begin{equation}
\text{FmocNH-AA} \quad X^* \quad 3.85
\end{equation}

\begin{equation}
\text{ii) Coupling cycle}
\end{equation}

\begin{equation}
\text{FmocNH-AA} \quad \text{NH-Ile} \quad \text{HMPP} \quad 3.127
\end{equation}

Repeat steps i) - ii)

\begin{equation}
\text{iii) Final Fmoc deprotection}
\end{equation}

\begin{equation}
\text{iv) Cleavage from resin using TFA}
\end{equation}

\begin{equation}
\end{equation}

\begin{equation}
Z' = \frac{\text{AcC}}{\text{AcC}} \quad 3.128, \text{Adpn(66-83)-Lys-3Gal-OAc}
\end{equation}

\begin{equation}
Z' = \frac{\text{AcC}}{\text{AcC}} \quad 3.129, \text{Adpn(66-83)-Ser-3Gal-OAc}
\end{equation}

\textit{Reagents and Conditions:} i) 5% (w/v) piperazine + 0.1 M 6-Cl-HOBt in DMF, $\mu w$, 60 W, 75 °C, 180 s; ii) Fmoc amino acid (5 eq.), HATU in DMF (0.5 M, 4.6 eq.), DIPEA in NMP (2 M, 10 eq.), 5 min, $\mu w$, 25 W, 73 °C except for Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 / Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 functionalised amino acids when ii) functionalised Fmoc-amino acid (2 eq.), HATU in DMF (0.5 M, 2 eq.), DIPEA in NMP (2 M, 4 eq.), 20 min, $\mu w$, 25 W, 73 °C; iii) 5% (w/v) piperazine + 0.1 M 6-Cl-HOBt in DMF, $\mu w$, 60 W, 75 °C, 180 s; iv) TFA/TIS/H$_2$O (95:2.5:2.5, v/v/v), $\mu w$, 10 W, 35 °C, 20 min.

Adpn(66-83)-Lys-3Gal-OAc 3.58a

Figure 3.20 Acetylated neoglycosides Adpn(66-83)-Lys-3Gal-OAc 3.58a structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using a Gemini 5µm column (C18, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5% - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
Chapter Three

Adpn(66-83)-Ser-3Gal-OAc 3.59a

Figure 3.21 Acetylated neoglycoside Adpn(66-83)-Ser-3Gal-OAc 3.59a structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using a Gemini 5µm column (C18, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5% - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
Crude ‘triple-clicked’ O-acetylated neoglycopeptides Adpn(66-83)-Lys-3Gal-OAc 3.58a and Adpn(66-83)-Ser-3Gal-OH 3.59a were deacetylated using freshly prepared sodium methoxide solution (1 M) at pH 11 overnight at room temperature. The crude peptides were then purified by RP-HPLC to afford O-deacetylated ‘triple-clicked’ neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b in >95% purity.

The deacetylated neoglycopeptide Adpn(66-83)-Lys-3Gal-OH 3.58b was afforded in 10.6% overall yield (including synthesis of crude acetylated peptide) and was characterised by the [M+2H]²⁺ (calculated 1266.5, observed 1265.6) and [M+3H]³⁺ (calculated 844.7, observed 844.4) charged states Figure 3.22.

The deacetylated neoglycopeptide Adpn(66-83)-Ser-3Gal-OH 3.59b was afforded in 8.6% overall yield (including synthesis of crude acetylated peptide) and was characterised by [M+2H]²⁺ (calculated 1203.4, observed 1203.1) and [M+3H]³⁺ (calculated 802.6, observed 802.5) charged states, Figure 3.23.
Figure 3.22 Deacetylated neoglycoside Adpn(66-83)-Lys-3Gal-OH \textbf{3.58b} structure, RP-HPLC trace, mass spectrum. The analytical RP-HPLC was performed using a Gemini 5\(\mu\)m column (C18, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5\% - 40\% B over 35 min, where buffer A = 0.1\% TFA in DI H\(_2\)O, Buffer B = 0.1\% TFA in CH\(_3\)CN.
Adpn(66 83)-Ser-3Gal-OH 3.59b

Figure 3.23 Deacytlated neoglycoside Adpn(66-83)-Ser-3Gal-OH 3.59b structure, RP-HPLC trace, mass spectrum. The analytical RP-HPLC was performed using a Gemini 5µm column (C18, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5% - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
3.15 Examination of secondary structure of neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58 and Adpn(66-83)-Ser-3Gal-OH 3.59 using circular dichroism

Buffer solutions of O-deacetylated neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b were prepared in potassium phosphate buffer at pH 7.4 and incubated overnight at 0 °C before being examined at 6 °C using circular dichroism spectroscopy (CD). CD is a useful tool to investigate the secondary structure of proteins in solution, as known secondary structures such as the triple helix, β-sheet and α-helix have characteristic CD spectra. Therefore, by comparing the CD spectra of synthetic peptides against the characteristic CD spectra of known secondary structures, researchers can gain insight into the secondary structure of the synthetic peptide.

The CD spectra of the collagen triple helix is characterised by maximum at 225 nm and a minimum at 198 nm, as demonstrated by Wennemers et al. who recently reported the CD spectra of the collaginous peptide Ac-(Pro-(4R)Azp-Gly)-NH₂ (Azp = azidoprolne) 3.130, Figure 3.24.

As neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b synthesised in this thesis correspond to residues 66-83 of the collaginous domain of adiponectin, one would expect the secondary structure of neoglycopeptides 3.58b and 3.59b to conform to the triple helical secondary structure exhibited by other collaginous proteins. However, neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b were shown to adopt a random coil conformation in the tested buffer solution rather the desired triple helical conformation, Figure 3.25. The random coil conformation is identifiable by a minimum close to 200 nm but an absence of a maximum at 225 nm.
3.16 The molecular basis of triple helical conformation of collagenous peptides

In order to understand the preference of the neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b to form a random coil conformation in contrast to the desired triple helical conformation common to collagenous proteins, one must consider the molecular basis for the formation and stability of the triple helix. The ability of a collagen-like peptide to form a triple helix is known to be influenced by three main factors; sterical constraints, hydrogen bonding networks and electrostatic interactions, which will be discussed in turn below.

3.16.1 The effect of steric constraints upon triple helical conformation of proteins

The triple helical conformation adopted by collagenous peptides involves close-packing of three polyproline II (PPII) strands around a central helical axis, as illustrated by Berman et al.203 in their crystal structure of H-(POG)$_4$POA(POG)$_5$-OH 3.131, Figure 3.26a.203 In order for close-packing to take place, the individual peptide chains must not be too sterically encumbered. Therefore, collagenous proteins always have the smallest amino acid glycine placed at each third residue to facilitate the close-packing of the individual PPII strands to form a triple helix.204 The triple helical conformation is held together by interstrand hydrogen bonding between the NH of a glycine residue and the C=O of a proline residue on an adjacent chain, Figure 3.26b.
3.16.2 The role of hydrogen bonding networks in formation of a triple-helical conformation

The imino acid proline (Pro, P) and its post-translationally modified partner hydroxyproline (Hyp, O) commonly observed in collagenous peptide sequences comprising approximately 20% of all residues in collagenous proteins, are also integral to the structure and stability of the triple helical conformation of collagenous proteins. Hydroxyproline residues extend the interstrand hydrogen bonding network often by solvent mediated interactions thus stabilising the triple helical conformation relative to the monomeric single stranded peptide. Interstrand glycan-glycan interactions have also been shown to extend the interstrand hydrogen bonding network, thus facilitating triple helical formation.

3.16.3 The importance of electrostatic interactions for triple-helical stability

Electrostatic interactions between peptide chains have also been shown by Brodsky et al. to influence triple helix and the stability of synthetic collagenous proteins. The role of both intrastrand and interstrand electrostatic interactions for stabilisation of the triple helical conformation of collagenous peptides is discussed in depth in Section 5.5, but here the role of capping of the N-termini of synthetic collagenous proteins will be discussed. As part of their study, Brodsky et al. synthesised the non-capped peptide T3-497 (based on human Type 3 collagen) and the N-terminal acetylated variant 3.133, Figure 3.27.

\[
\begin{align*}
\text{H}_2\text{N-GKOGEOGPKDAGOGAO(GPO)}_4\text{GV-ÖH} & \quad \text{Ac-GKOGEOGPKDAGOGAO(GPO)}_4\text{GV-ÖH} \\
\text{T3-487} & \quad \text{Ac-T3-487}
\end{align*}
\]

Figure 3.27 Collagenous peptides 3.132 and 3.133 synthesised by Brodsky et al., with expected charges shown on ionisable residues at pH 7.4.
At pH 7, non-capped peptide T3-497 3.132 was found to have a thermal melting temperature ($T_m$) of 26 °C, whereas the acetylated variant 3.133 was found to be more thermally stable with a $T_m$ of 28 °C. This can be explained by the fact that at pH 7.4 the N-termini of peptide 3.132 will be positively charged and thus repel each other in solution, reducing the thermal stability of the non-capped peptide 3.132. Brodsky et al. \textsuperscript{204} also investigated the effect of pH on the thermal stability of peptides 3.132 and 3.133, Figure 3.28.

![Figure 3.28 pH dependency of non-capped peptide T3-497 3.132 (□) and acetylated variant 3.133 (●) as reported by Brodsky et al.\textsuperscript{204}.](image)

At pH values below 9, the acetylated T3 peptide 3.133 was found to be more stable than the native T3-analogue 3.132 indicating interstrand destabilising N-termini electrostatic interactions to be of importance for peptide stability. In addition, both T3 peptides experienced a steep increase in stability between pH 3 – 4.5, corresponding to ionisation of glutamic acid (Glu, E) and aspartic acid (Asp, D) residues leading to stabilising interstrand electrostatic interactions (for further details see Section 5.5).

Furthermore, Shoulders et al.\textsuperscript{206} noted that in native collagen self-assembly N- and C-terminal telopeptides are needed for the formation of the tropocollagen triple helix. The use of an N-terminal acetyl group, therefore is a good mimic for native tropocollagen triple helical formation, Figure 3.29.
Figure 3.29 Diagram of native self-assembly of collagen illustrating the role of N- and C-terminal propeptides for tropocollagen triple helical formation, as reported by Shoulders et al.\textsuperscript{206}.

The inability of neoglycopeptides Adpn(66-83)-Lys-3Gal-OH \textbf{3.58b} and Adpn(66-83)-Ser-3Gal-OH \textbf{3.59b} to form a triple helical conformation could have been due to steric constraints imposed by the ‘clicked’ glycan moieties, hydrogen bonding restrictions linked to the Pro/Hyp residues in the peptide sequence and/or electrostatic interactions due to charged residues within the peptide sequence, as well as the charged, uncapped N- and C- termini of the neoglycopeptides.

\textbf{3.17 Hypothesis for the random coil conformation exhibited by deacetylated neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.61b}

The secondary structures of similar neoglycopeptides, such as MUC-1 neoglycopeptide \textbf{3.54} have not yet been reported. Therefore, a review of similar length collagenous peptides with known crystal structures was conducted, in order to establish a suitable explanation for the random coil conformation adopted by neoglycopeptides Adpn(66-83)-Lys-3Gal-OH \textbf{3.58b} and Adpn(66-83)-Ser-3Gal-OH \textbf{3.59b}, Table 3.6.

Baumann et al.\textsuperscript{207} have reported the synthesis and CD spectrum of the collagenous 18-mer Ac-(PPG)\textsubscript{2}PTGPRG(PPG)\textsubscript{2}-OH \textbf{3.134}, a mimic of the ER-resident chaperone collagen Hsp47, and the longer 27-mer Ac-(PPG)\textsubscript{2}PTGPRGPPPP(GPP)\textsubscript{2}CPPG-NH\textsubscript{2} \textbf{3.135}\textsuperscript{207} The 18-mer \textbf{3.134} was found to
adopt a random coil conformation at 20 °C, whereas the longer 27-mer CMP 3.135 was found to adopt a triple helical conformation under these conditions.

Notably, all of the other triple helical peptides reported in the literature, except that reported by Chen et al.\textsuperscript{142}, and shown in Table 3.6 consist of at least 7 collagen repeat units, i.e. are at least 21 amino acids long. Therefore, one could postulate that the 18-mers neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b were too short to be able to form the triple helical conformation and therefore adopted the random coil conformation in the tested buffer solution. The CD study of the 18-mer CMP Ac-(GPO)\textsubscript{2}GLOGEA(GPO)\textsubscript{2}-NH\textsubscript{2} 3.136 reported by Chen et al.\textsuperscript{142} were run at 2 °C, lower than the CD studies reported by others. Therefore, the CD spectrum of 3.136 is not comparable with the other collagenous peptides in Table 3.6.

A study of similar collagenous peptides also revealed that many synthetic collagenous proteins including 3.135, 3.136, 3.130 and 3.138 were synthesised with N-terminal acetyl capping to avoid N-termini repulsion, as described in Section 3.15. Therefore, the lack of N-terminal acetyl capping for neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b could also have been a contributing factor for the inability of the neoglycopeptides to form a triple helical conformation.

\textbf{3.18 Conclusion for Chapter Three}

To conclude, the random coil conformation exhibited by neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b at 6 °C is likely to be due to both the short nature of the peptide sequence and N-termini interchain repulsion. These conclusions form the basis for the design of the next two families of peptides described in Chapters 4 and 5.
Table 3.6 Peptide sequences of small collagen-like peptides and their correspondingly reported secondary structures. All CD studies run between pH 7.2-7.4.

<table>
<thead>
<tr>
<th>Peptide Sequence</th>
<th>Peptide Number</th>
<th>Number of collagen repeat units</th>
<th>Secondary Structure</th>
<th>Temperature of CD study (°C)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>H-GEK<em>GEK</em>GDOGLGPK*GDI-OH</td>
<td>3.58b</td>
<td>6</td>
<td>Random coil</td>
<td>6</td>
<td>This thesis</td>
</tr>
<tr>
<td>H-GES<em>GES</em>GDOGLGPS*GDI-OH</td>
<td>3.59b</td>
<td>6</td>
<td>Random coil</td>
<td>6</td>
<td>This thesis</td>
</tr>
<tr>
<td>Ac-(PPG)2PTGPRG(PPG)2-OH</td>
<td>3.134</td>
<td>6</td>
<td>Random coil</td>
<td>20</td>
<td>Baumann et al.207</td>
</tr>
<tr>
<td>Ac-(PPG)2PTGPRGPPP(GPP)2CPPG-NH2</td>
<td>3.135</td>
<td>8</td>
<td>Triple helix</td>
<td>20</td>
<td>Baumann et al.207</td>
</tr>
<tr>
<td>Ac-(GPO)2GLOGEA(GPO)2-NH2</td>
<td>3.136</td>
<td>6</td>
<td>Triple helix</td>
<td>2</td>
<td>Chen et al.142</td>
</tr>
<tr>
<td>Ac-(POG)2PXG(POG)2-NH2 where X is a ‘clicked’ proline derivative</td>
<td>3.130</td>
<td>7</td>
<td>Triple helix</td>
<td>10</td>
<td>Wennemers et al.143</td>
</tr>
<tr>
<td>H-(POG)3PYG(POG)3-OH, where Y is an AGE (advanced glycation endproduct)</td>
<td>3.140</td>
<td>7</td>
<td>Triple helix</td>
<td>5</td>
<td>Brimble et al.201</td>
</tr>
<tr>
<td>H-(POG)3POA(POG)3-OH</td>
<td>3.131</td>
<td>10</td>
<td>Triple helix</td>
<td>n/a, Xtal structure</td>
<td>Berman et al.203</td>
</tr>
<tr>
<td>H-GKOGEGPKDGADAOOGAOG(DPO)3GV-OH</td>
<td>3.137</td>
<td>10</td>
<td>Triple helix</td>
<td>10</td>
<td>Brodsky et al.204</td>
</tr>
<tr>
<td>Ac-GKOGEGPKDGADAOOGAOG(DPO)3GV-OH</td>
<td>3.138</td>
<td>10</td>
<td>Triple helix</td>
<td>10</td>
<td>Brodsky et al.204</td>
</tr>
<tr>
<td>H-(POG)3EKG(POG)3-OH</td>
<td>3.139</td>
<td>10</td>
<td>Triple helix</td>
<td>10</td>
<td>Brodsky et al.204</td>
</tr>
</tbody>
</table>
Chapter Four

The Use of Collagen Model Peptides as Structural Scaffolds for Neoglycopeptides of Adpn
Chapter Four: The Use of Collagen Model Peptides as Structural Scaffolds for Neoglycopeptides of Adpn

4.1 Introduction to collagen model peptides

As discussed in Chapter 3, the deacetylated ‘triple-clicked’ Adpn-neoglycopeptides 3.58b and 3.59b, Figure 4.1, were shown to adopt a random coil conformation in solution in preference to the desired triple helical conformation. Reasonable explanations for the 18-mer Adpn peptides 3.58b and 3.59b not adopting a triple helical conformation could be the short length of the peptide as well as the lack of acetyl capping of the N-termini thus leading to interstrand repulsion.

Therefore, the next logical step was to design a peptide system which would support the Adpn segment to form a triple helical conformation. The triple helical conformation is crucially important for the synthetic peptides generated, as the aim of this thesis is to better understand the structural role of Lys and glycosylated Lys residues found within the collagenous domain of Adpn as Cooper et al.,\textsuperscript{28} Myllylä et al.\textsuperscript{208} Ohkubo et al.\textsuperscript{49} have demonstrated that glycosylation of lysine residues found in the collagenous domain of Adpn play a crucial role in facilitating the self-assembly and bioactivity of this protein.\textsuperscript{49} The secondary structure of the peptides generated to address this topic of structural stability

**Figure 4.1** Prepared neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b.
must therefore match the secondary structure of the collagenous domain of Adpn, which is a triple helix. An alternative synthetic peptide system was sought, which could incorporate the 18-mer Adpn peptide sequence of 3.58b and 3.59b into a longer peptide that exhibited triple helical properties. These longer peptides should also closely mimic the natural structure of the collagenous domain of Adpn.

Collagen model peptides (CMPs) are synthetic equivalents of native collagen, where the XYG native collagen repeat unit is commonly ProHypGly (where Hyp = O = 4R-hydroxyproline, POG), the most abundant triplet of native collagen. CMPs serve as structural mimics of collagen and as such have numerous medical and biotechnological applications, such as wound healing agents and electrically conducting nanowires. CMPs have aided the understanding of the structural basis of collagen, with the first high resolution (1.9 Å) X-ray crystal structure of a collagen peptide H-(POG)₄POA(POG)₅-OH 4.1 being reported by Berman et al. in 1994. This structure confirmed the current structural model for collagen as a right-handed triple helix consisting of three staggered, left-handed PPII (polyproline II) helices held together by interstrand hydrogen bonding between the NH groups of glycine residues and the C=O groups of the residue at the X position on the adjacent chain, Figure 4.2.

![Figure 4.2](image-url)

**Figure 4.2** a) Illustration of interstrand hydrogen bonding found in collagen and collagen model peptides as reported by Bella et al. with green lines indicating the hydrogen bonding and b) diagram illustrating the interstrand hydrogen bonding network.

Moreover, CMPs have been used to investigate different physicochemical features of collagen, including the endo/exo ring pucker of the hydroxyproline ring and the cross-linking properties of collagen peptides using both disulfide bridges and lysyl advanced glycation end products (AGEs). Furthermore, multiple diabetic complications such as ventricular dysfunction, diabetic cardiomyopathy and diabetic nephropathy result from increased rigidity of the tissue collagens.
4.2 Structural features of CMPs

An important structural feature of collagen is the network of interstrand hydrogen bonds which support the collagen triple helical structure. CMPs which incorporate a glycine ester (Gle) derivative in place of a native glycine residue (Gly) have been used by Raines et al.\textsuperscript{220} to probe the strength of the interstrand hydrogen bonding between Gly and Pro residues, Figure 4.3, CMPs 4.2 and 4.3. Their study found that the CMPs containing the glycine ester in place of native glycine were not able to form triple helices in solution, thereby demonstrating the importance of interstrand hydrogen bonding for structure formation.

\[
\text{Gle} = \text{glycine ester derivative}
\]

\[
\text{H-(ProProGly)}_{10}\text{-OH}
\]

\[
\text{H-(ProProGly)}_{4}\text{ProProGle(ProProGly)}_{5}\text{-OH}
\]

\[
\text{Hyp(OAc)} = \text{N-\(\text{OAc}\)OHN}
\]

\[
\text{H-[ProHyp(OAc)Gly]}_{10}\text{-OH}
\]

\[
\text{H-(ProHypGly)}_{10}\text{-OH}
\]

Figure 4.3 a) Schematic detail of interstrand hydrogen bonding in H-(ProProGly)\textsubscript{10}-OH 4.2 and c) interstrand hydrogen bonding prevented in H-(ProProGly)\textsubscript{4}ProProGle(ProProGly)\textsubscript{5}-OH 4.3 as reported by Raines et al.\textsuperscript{220}

Raines et al.\textsuperscript{221,222} have also investigated the importance of the hydroxyl group of Hyp residues, which are usually present in CMPs. In that study, the structural properties of the 4-O-acetylated-Hyp variant (O\textsuperscript{*}) of the CMP H-(PO\textsuperscript{*}G)\textsubscript{10}-OH 4.4 was compared with corresponding non-acetylated variant 4.5, Figure 4.4. Both the acetylated and non-acetylated Hyp variants were able to form triple helices in solution (as the Hyp residue is not involved in interstrand hydrogen bonding as shown by Figure 4.2b), although the acetylated peptide 4.4 exhibited lower thermal stability, with a T\textsubscript{m} of 58 °C reported for CMP 4.4 compared to the corresponding non-acetylated peptide 4.5 which had a T\textsubscript{m} of 69 °C. In this case, the lower thermal stability of the acetylated peptide was attributed to interchain steric clashes arising from the increased bulk of the acetyl group compared to the smaller hydroxyl group of the Hyp residue.

Figure 4.4 Acetylated and deacetylated Hyp based CMPs as reported by Raines et al.\textsuperscript{221,222}
4.3 Previous studies of glycosylated CMPs

4.3.1 CMPs bearing glycosylated hydroxyproline residues

Glycosylation of Hyp residues has been shown to alter the secondary structure of CMPs in differing ways. Schweizer \textit{et al.}\textsuperscript{223} demonstrated that incorporation of a single glycosylated Hyp residue in an Ac-(POG)\textgreek{7}-NH\textsubscript{2} peptide was tolerated, but incorporation of three or more glycans forced the CMP into a random coil conformation. On the other hand, Wetmore \textit{et al.}\textsuperscript{205} showed that increased levels of glycosylated Hyp in the peptide Ac-(O)\textgreek{7}-NH\textsubscript{2} elevated the thermal stability of these CMPs. These two differing reports are discussed in detail below.

The thesis of E. Naziga\textsuperscript{223} highlights a series of unpublished CMPs synthesised by Schweizer \textit{et al.}, Figure 4.5. These CMPs contain varying degrees of glycosylation of the middle Hyp (O) amino acid together with the control CMP variant \textbf{4.6} containing no glycosylation of Hyp, CMPs \textbf{4.7} and \textbf{4.8} bearing one glycosylated Hyp as \textalpha{} and \textbeta{} anomers respectively, CMP \textbf{4.9} incorporates three \textbeta{}-glycosylated Hyp residues and CMP \textbf{4.10} contains seven \textbeta{}-glycosylated Hyp residues.

Schweizer \textit{et al.}\textsuperscript{223} showed that a single \textbeta{}-glycosylation decreased the thermal stability of the triple helix from 39.1 °C for unglycosylated control CMP \textbf{4.6} to 37.1 °C for glycosylated CMP \textbf{4.8}. CMPs \textbf{4.9} and \textbf{4.10} bearing three or more sugar moieties were unable to from a triple helix, with increased steric hindrance of these peptides cited as a possible explanation. Interestingly, the \textalpha{}-anomer CMP \textbf{4.7} was also unable to form a triple helix, indicating that the stereochemistry of the glycosidic linkage also affects the secondary structure of the peptide.

\begin{figure}
\centering
\includegraphics[width=0.5\textwidth]{figure4.5.png}
\caption{Control non-glycosylated CMP 4.6 and glycosylated CMPs 4.7 – 4.10 reported by Schweizer \textit{et al.}\textsuperscript{223,224}}
\end{figure}
Furthermore, in conference proceedings Schweizer et al.\textsuperscript{224} have also indicated that the inability of the more highly glycosylated CMPs 4.9 and 4.10 to form a triple helical conformation may provide a rationalisation for the observed structural differences between plant and animal collagen. Plant collagens are single-stranded and contain high levels of glycosylated Hyp residues, whereas animal collagens are triple-stranded and bear no glycosylated Hyp residues. These results would point towards a mechanism whereby modulation of glycosylation in plant collagens is key to their structural stability and function.

Wetmore et al.\textsuperscript{205} have investigated the structural stability of a series of oligoprolines, namely Ac-(Pro)\textsubscript{9}-NH\textsubscript{2} 4.11, Ac-(Hyp)\textsubscript{9}-NH\textsubscript{2} 4.12 and Ac-(Hyp-β-Gal)\textsubscript{9}-NH\textsubscript{2} 4.13, with T\textsubscript{m} values of these CMPs reported as 22 °C, 38 °C and 70 °C respectively. The more stable trans-conformation of the imino acid is favoured in the Hyp-containing CMP 4.12 relative to mixed cis- and trans- conformations of the imino acid of Pro-containing CMP 4.11. This could explain the increased thermal stability upon hydroxylation of the Pro residues and thus the increased thermal stability observed for Hyp-containing CMP 4.12 compared to Pro-containing CMP 4.11. The trans-conformation of imino acids has been shown to allow for the favourable n→π* interactions between neighbouring carbonyl groups, thereby stabilising the trans-conformation relative to the cis-conformation,\textsuperscript{205} Figure 4.6.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{trans_conformation.png}
\caption{Increased stability of the trans-conformation of proline due to favourable n→π* interaction between neighbouring carbonyl groups as demonstrated by Wetmore et al.\textsuperscript{205}}
\end{figure}

Wetmore et al.\textsuperscript{205} reported that the hydroxylated CMP 4.12 and the glycosylated CMP 4.13 exhibited the same number of trans-conformation of the imino acids, although the thermal stabilities of the CMPs were very different, with hydroxylated CMP 4.12 recording a T\textsubscript{m} value of 38 °C and glycosylated CMP 4.13 a T\textsubscript{m} value of 70 °C. Therefore, a different explanation is needed to explain the increased stability of glycosylated CMP 4.13 compared to hydroxylated CMP 4.12. Computational analysis demonstrated that glycosylated CMP 4.13 contained an extended hydrogen bonding network relative to the unglycosylated, hydroxylated CMP 4.12. Both glycan-glycan and glycan-peptide backbone interactions were shown to present in glycosylated CMP 4.13, Figure 4.7. Free energy estimates for glycosylated CMP 4.13 indicated that solvent interactions were integral to this extensive hydrogen bonding network.
Figure 4.7 The two different types of hydrogen bonding proposed by Wetmore et al.\textsuperscript{205} for glycosylated CMP 4.13: a) glycan-glycan interactions between O-6 and O-2 of adjacent glycan moieties and b) glycan-peptide backbone interactions between glycan O-6 and the carbonyl group of the peptide backbone.

4.3.2 Structural models of glycan-peptide backbone interactions

Further evidence for the structural role of glycans in CMPs has been reported by Fincher et al.\textsuperscript{225} and Qi et al.\textsuperscript{226} in their studies of plant-cell wall glycoproteins, commonly known as arabinogalactans (AGPs). Through examination of electron-microscopy images of different AGPs both groups have observed key interactions between the glycans and the peptide backbone of these AGPs. However, the level of these interactions is disputed by the two groups with different proposed structural models for the positioning of glycans in AGPs. Fincher et al.\textsuperscript{225} proposed the ‘wattle blossom’ model, whereas Qi et al.\textsuperscript{226} proposed the ‘twisted-hairy-rope’ model, Figure 4.8. In the ‘wattle blossom’ model\textsuperscript{225} the glycans are positioned pointing away from the triple helical core, thus minimising any glycan-peptide interactions. On the other hand in the ‘twisted-hairy-rope’ model\textsuperscript{226} the glycans are arranged along the peptide backbone, thereby allowing interactions to take place between the glycan and the peptide backbone. The ‘twisted-hairy-rope’ model could be used as a suitable model for the glycan-peptide interactions observed by Wetmore et al.\textsuperscript{205}

Figure 4.8 a) The ‘wattle blossom’ model proposed by Fincher et al.\textsuperscript{225} and b) the ‘twisted-hairy-rope’ model proposed by Qi et al.\textsuperscript{226} for positioning of glycans in AGPs.

4.3.3 Incorporation of glycans into a CMP sequence using the copper-catalysed 1,3-cycloaddition ‘click’ reaction

Wennemers et al.\textsuperscript{143,227} have also reported a series of glycosylated CMPs 4.14 – 4.17 in which the glycan is bonded to a (4R)-Hyp residue, Figure 4.9. In their work the copper-catalysed 1,3-cycloaddition ‘click’
reaction\textsuperscript{155} is used to attach the glycan to the Hyp via a 1,2,3 triazole ring to afford glycosylated CMPs \textbf{4.15} and \textbf{4.16} along with the control, unglycosylated CMPs \textbf{4.14} and \textbf{4.17}.

Circular dichroism studies revealed that both the glycosylated and unglycosylated peptides were able to form triple helices in solution. Intriguingly, glycosylation by a single glycan did not affect the thermal stability of CMPs \textbf{4.15} and \textbf{4.16} relative to unglycosylated, triazole CMP \textbf{4.14}, with similar thermal melt values between 36 - 37 °C reported for all three CMPs. Notably the thermal stability of these three triazole containing CMPs were found to be considerably lower than the native Hyp analogue CMP \textbf{4.17}, for which a T\textsubscript{m} of 43 °C was recorded, Figure 4.9.

![CMP structures](image)

\textbf{Figure 4.9} Glycosylated CMPs \textbf{4.15} and \textbf{4.16}, control CMP \textbf{4.14} and native Hyp CMP analogue \textbf{4.17}, as reported by Wennemers \textit{et al.}\textsuperscript{143,227}

\section*{4.3.4 CMPs containing glycosylated threonine residues}

Bächinger \textit{et al.}\textsuperscript{228} have investigated the role of threonine glycosylation in the cuticle collagen of deep sea giant tube worms through the synthesis of glycosylated CMP Ac-[Gly-Pro-Thr(β-Gal)]\textsubscript{10}-NH\textsubscript{2} \textbf{4.18} and unglycosylated control CMP variant Ac-(Gly-Pro-Thr)\textsubscript{10}-NH\textsubscript{2} \textbf{4.19}. Through circular dichroism (CD) studies, they were able to demonstrate that only the glycosylated CMP \textbf{4.18} was able to adopt a triple helical conformation. Therefore, in the case of these CMPs, threonine glycosylation was shown to be essential for triple helix formation, with the glycosylated CMPs forming a highly thermally-stable triple helix (T\textsubscript{m} of 40 °C) and the unglycosylated CMP unable to adopt a triple helical conformation.
4.4 Chosen CMP sequences for synthesis

As discussed above, CMPs have been shown to be suitable models for collagenous proteins. Therefore, it was decided to utilise a ‘host-guest’ CMP system with three ‘host’ POG (Pro-Hyp-Gly) triplets encasing a central collagenous 18-mer ‘guest’ Adpn domain. Before incorporation of the 18-mer Adpn sequence was attempted (see Chapter 5), a model system was synthesised. This model CMP series 4.20 - 4.27, incorporated one to three \( X^* \)EG ‘guest’ triplets containing the ‘click’ glycosylated amino acid of interest (\( X^* = \) neoglycosylated amino acid) between (POG)\(_3\) head and tail ‘host’ structures to afford two series of ‘host-guest’ CMPs, Figure 4.10.

*Wennemers et al.*

**Figure 4.10** Structure of ‘host-guest’ collagen model peptides reported by Wennemers *et al.*,\(^ {143,227}\) and chosen for this series of model CMPs 4.20 – 4.27.

The \( X^* \)EG ‘guest’ sequences were chosen to mirror the KEGKEG sequence shown to be of structural importance when the Lys residues are glycosylated, in the collagenous domain of adiponectin.\(^ {34}\) It should be noted that as shown in Figure 1.11, the native Adpn sequence 66-83 is GEKGEKDOGLIGOKGDI (\( N \rightarrow C \) terminus). For the lysine and serine model CMPs 4.20 – 4.27, the ‘host’ sequence of (KEG)\(_n\) or (SEG)\(_n\) was chosen, to be as close as possible to the ‘clicked’ CMPs 4.14 – 4.17 reported by Wennemers *et al.*,\(^ {143,227}\) which were shown to adopt the desired triple helical conformation. Thus, a ‘host’ sequence of (KEG)\(_n\) or (SEG)\(_n\) was chosen in preference to (GEK)\(_n\). Similar to Wennemers *et al.*,\(^ {143,227}\), the chosen CMP peptides 4.20 – 4.27 were \( N \)-terminus acetylated and \( C \)-terminus amidated to improve thermal stability of the peptides and encourage adoption of the desired triple helical conformation.

Furthermore, acetyl groups were used as masking groups on the hydroxyl groups of the glycans to directly investigate hydrogen bonding interactions involving the glycans. Similar to the neoglycopeptides synthesised in Chapter 3 glycosylated lysine CMPs 4.20 - 4.23 were synthesised alongside glycosylated serine CMPs 4.24 - 4.27, due to the differences observed in the glycan-peptide backbone through-space distances using modelling studies (see Section 3.7). In order to establish the effect of incorporation of a triazole linker into the CMP sequence, the control non-glycosylated CMPs analogues 4.23 (Lys) and 4.27 (Ser) were also synthesised, Figure 4.11.
Figure 4.11 Lysine CMPs 4.20 – 4.23 and serine analogues 4.24 – 4.27 incorporating ‘click’ glycosylated amino acids $X_1$ and $X_2$ respectively.

4.5 Synthesis of Family 2 CMPs

Preparation of CMPs 4.20a – 4.27 were carried out using Fmoc-SPPS$^{136}$ with incorporation of the previously synthesised ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 functionalised amino acids to provide the desired functionalities, Figure 4.12.
Figure 4.12 Chemical structure of ‘pre-clicked’ building blocks Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 and Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 which were incorporated into Family 2 peptides, lysine CMPs 4.20a – 4.23a and serine CMPs 4.24a – 4.27a respectively.

CMPs 4.20a – 4.27 were prepared on 0.98 g/ mmol loading on “in house”197 synthesised aminomethyl polystyrene (AM-PS) and immobilised via Rink amide linker 3.94. Due to equipment availability, different microwave synthesisers were employed for different CMPs. An overview of the different synthesisers and optimised methods utilised for the synthesis of CMPs 4.20a – 4.27 is given in Table 4.1. These synthetic methods are described in full in Chapter Seven: Experimental Methods and Product Characterisation. Unless otherwise stated, after purification using RP-HPLC with a Dionex Ultimate 3000 equipped with a 4 channel UV detector, CMPs 4.20a – 4.22a and 4.24a – 4.27 were afforded in >95% purity. The yields stated are with respect to the crude peptide.

For the glycosylated Family 2 CMPs, coupling of the appropriate ‘pre-clicked’ neoglycosides Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 or Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 was performed using the powerful coupling agent HATU for an extended period of 30 minutes, as opposed to the normal coupling time of 20 minutes employed for incorporation of the other amino acid residues into the peptide sequence. An overview of the synthetic method employed for the synthesis of 4.20a - 4.27 is given in Scheme 4.1.
For reagents and conditions please see Table 4.1.

**Scheme 4.1** Synthesis of Family 2 CMP peptides 4.20a - 4.27 using Fmoc-SPPS, where $X^*$ = activated ester leaving group.

In brief, CMP-Lys-1Gal-OAc 4.20a and CMP-Ser-Control 4.27 were synthesised using an automated Liberty 12 Microwave Peptide Synthesiser (CEM Corporation, Mathews, NC). Fmoc deprotection and coupling reactions were carried out with microwave-irradiation with HCTU/DIPEA and piperidine as the coupling and Fmoc deprotection reagents, respectively. On completion of synthesis, the peptidyl-resin was Fmoc-deprotected (as above) and $N$-terminal acetylated using acetic anhydride and
rotary mixing. Cleavage of the acetylated peptidyl-resin with TFA, TIPS (triisopropylsilane), H₂O (95:2.5:2.5, v/v) afforded crude peptides, which were purified by RP-HPLC to afford the desired peptides CMP-Lys-1Gal-OAc 4.20a and CMP-Ser-Control 4.27 identified by LC-MS.

The CMP-Lys-1Gal-OAc 4.20a peptide was afforded as a mixture of differently acetylated products in 7.6% overall yield. The differently acetylated products were identified by LC-MS to be the K2OAc and K1OAc variants:

Where K1OAc = and K2OAc = in the peptide sequence.

The K2OAc variant of CMP-Lys-1Gal-OAc 4.20a was characterised by the [M+2H]²⁺ (calculated 1153.2, observed 1153.1) and [M+3H]³⁺ (calculated 769.1, observed 769.3) charged states, Figure 4.13.

The K1OAc variant CMP-Lys-1Gal-OAc 4.20a was characterised by the [M+2H]²⁺ (calculated 1132.2, observed 1132.4) and [M+3H]³⁺ (calculated 755.1, observed 755.3) charged states, Figure 4.14.

The CMP-Ser-Control 4.27 peptide was afforded in 16.3% overall yield and was characterised by the [M+2H]²⁺ (calculated 968.5, observed 968.5) charged state, Figure 4.21.

For CMP-Lys-2Gal-OAc 4.21a, CMP-Lys-3Gal-OAc 4.22a, Fmoc deprotection and coupling reactions were carried out for the first eleven residues using a Tribute™ synthesiser (Protein Technologies, Tuscon, Az) with HCTU/DIPEA and piperidine as the coupling and Fmoc deprotection reagents, respectively to afford Fmoc-EG(POG)₃-Rink-ϕ 4.32. Subsequent elongation of the peptide was carried out using a CEM™ microwave synthesiser (CEM Corporation, Mathews, NC) with HATU/DIPEA and piperidine as the coupling and Fmoc deprotection reagents respectively. On completion of synthesis, the peptidyl-resin was Fmoc-deprotected, N-terminal acetylated and the peptide cleaved from resin using the method described above for CMP-Lys-1Gal-OAc 4.20a and CMP-Ser-Control 4.27. The crude peptides were the purified by RP-HPLC to afford the desired peptides CMP-Lys-2Gal-OAc 4.21a and CMP-Lys-3Gal-OAc 4.22a identified by LC-MS. There was also evidence of spontaneous deacetylation of CMP-Lys-2Gal-OAc 4.21a and CMP-Lys-3Gal-OAc 4.22a, visible in the mass spectrum data of these compounds.
The CMP-Lys-2Gal-OAc 4.21a peptide was afforded in 4.2% overall yield and was characterised by the [M+2H]^{2+} (calculated 1558.3, observed 1558.6) and [M+3H]^{3+} (calculated 1039.17, observed 1039.4) charged states, Figure 4.15.

The CMP-Lys-3Gal-OAc 4.22a peptide was afforded in 0.6% overall yield and was characterised by the [M+2H]^{2+} (calculated 1921.4, observed 1922.2) and [M+3H]^{3+} (calculated 1281.3, observed 1281.7) charged states, Figure 4.16.

For CMP-Lys-Control 4.23, CMP-Ser-1Gal-OAc 4.24a, CMP-Ser-2Gal-OAc 4.25a and CMP-Ser-3Gal-OAc 4.26a, Fmoc deprotection and coupling reactions were carried out for the first eleven residues using a Tribute™ synthesiser (Protein Technologies, Tuscon, Az) with either HATU/DIPEA (final ‘guest’ (POG) triplets after the ‘host’ sequence) or HCTU/DIPEA (all other residues) and piperidine as the coupling and Fmoc deprotection reagents, respectively.

On completion of synthesis, the peptidyl-resin was Fmoc-deprotected (as above) and N-terminal acetylated using acetic anhydride and a Tribute™ synthesiser (Protein Technologies, Tuscon, Az). Cleavage of the acetylated peptidyl-resin with TFA, TIPS (triisopropylsilane), H_{2}O (95:2.5:2.5, v/v) afforded crude peptides, which were purified by RP-HPLC to afford the desired peptides CMP-Lys-Control 4.23, CMP-Ser-1Gal-OAc 4.24a, CMP-Ser-2Gal-OAc 4.25a and CMP-Ser-3Gal-OAc 4.26a identified by LC-MS.

The CMP-Lys-Control 4.23 peptide was afforded in 13% overall yield and was characterised by the [M+H]^{+} (calculated 1997.1, observed 1997.8), [M+2H]^{2+} (calculated 989.1, observed 989.0) and [M+3H]^{3+} (calculated 659.7, observed 659.9) charged states, Figure 4.20.

The CMP-Ser-1Gal-OAc 4.24a peptide was afforded in 5.8% overall yield and was characterised by the [M+2H]^{2+} (calculated 1174.0, observed 1174.1) and [M+3H]^{3+} (calculated 783.0, observed 783.2) charged states, Figure 4.17.

The CMP-Ser-2Gal-OAc 4.25a peptide was afforded in 5.8% overall yield and was characterised by the [M+2H]^{2+} (calculated 1516.2, observed 1516.6) and [M+3H]^{3+} (calculated 1011.1, observed 1011.4) charged states, Figure 4.18.

The CMP-Ser-3Gal-OAc 4.26a peptide was afforded in 0.6% overall yield and was characterised by the [M+2H]^{2+}(calculated 1858.4, observed 1858.7.6) and [M+3H]^{3+} (calculated 1239.3, observed 1239.5) charged states, Figure 4.19.
<table>
<thead>
<tr>
<th>Peptide #</th>
<th>Peptide Acronym</th>
<th>Peptide Sequence</th>
<th>Coupling conditions employed*</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.20a</td>
<td>CMP-Lys-1Gal-OAc</td>
<td>Ac-(POG)_3X_1EG(POG)_3-NH_2</td>
<td>Liberty 12™ µw synthesiser, single coupling cycles, HCTU (4.5 eq.), DIPEA/NMP (4 eq.).</td>
</tr>
<tr>
<td>4.21a</td>
<td>CMP-Lys-2Gal-OAc</td>
<td>Ac-(POG)_3X_1EG-X_1EG(POG)_3-NH_2</td>
<td>Fmoc-EG(POG)_3-Rink-ϕ synthesised using the Tribute™ synthesiser, single coupling cycles, HCTU (4.6 eq.), DIPEA/NMP (10 eq.). Remainder of the peptide synthesised using the CEM™ µw synthesiser, single coupling cycles, HATU (4.6 eq.), DIPEA/DMF (10 eq.).</td>
</tr>
<tr>
<td>4.22a</td>
<td>CMP-Lys-3Gal-OAc</td>
<td>Ac-(POG)_3X_1EG-X_2EGX_1EG(POG)_3-NH_2</td>
<td>Fmoc-EG(POG)_3-Rink-ϕ synthesised using the Tribute™ synthesiser, single coupling cycles using HCTU (4.6 eq.), DIPEA/NMP (10 eq.). Remainder of the peptide synthesised using the CEM™ µw synthesiser, single coupling, HATU (4.6 eq.), DIPEA/DMF (10 eq.).</td>
</tr>
<tr>
<td>4.23</td>
<td>CMP-Lys-Control</td>
<td>Ac-(POG)_3KEG(POG)_3-NH_2</td>
<td>Fmoc-(POG)_3-Rink-ϕ synthesised using the Tribute™ synthesiser, single coupling cycles using HCTU (4.6 eq.), NMM/NMP (10 eq.). Remainder of the peptide synthesised using the Tribute™ synthesiser, double coupling, HATU (1.9 eq.), DIPEA/NMP (4 eq.).</td>
</tr>
<tr>
<td>4.24a</td>
<td>CMP-Ser-1Gal-OAc</td>
<td>Ac-(POG)_3X_2EG(POG)_3-NH_2</td>
<td>Fmoc-(POG)_3-Rink-ϕ synthesised using the Tribute™ synthesiser, single coupling cycles using HCTU (4.6 eq.), NMM/NMP (10 eq.). Remainder of the peptide synthesised using the Tribute™ synthesiser, double coupling, HCTU (1.9 eq.), DIPEA/NMP (4 eq.).</td>
</tr>
<tr>
<td>4.25a</td>
<td>CMP-Ser-2Gal-OAc</td>
<td>Ac-(POG)_3X_2EG-X_2EG(POG)_3-NH_2</td>
<td>Fmoc-(POG)_3-Rink-ϕ synthesised using the Tribute™ synthesiser, single coupling cycles using HCTU (4.6 eq.), NMM/NMP (10 eq.). Remainder of the peptide synthesised using the Tribute™ synthesiser, double coupling, HATU (1.9 eq.), DIPEA/NMP (4 eq.).</td>
</tr>
<tr>
<td>4.26a</td>
<td>CMP-Ser-3Gal-OAc</td>
<td>Ac-(POG)_3X_2EG-X_2EGX_2EG(POG)_3-NH_2</td>
<td>Fmoc-(POG)_3-Rink-ϕ synthesised using the Tribute™ synthesiser, single coupling cycles using HCTU (4.6 eq.), NMM/NMP (10 eq.). Remainder of the peptide synthesised using the Tribute™ synthesiser, double coupling, HATU (1.9 eq.), DIPEA/NMP (4 eq.).</td>
</tr>
<tr>
<td>4.27</td>
<td>CMP-Ser-Control</td>
<td>Ac-(POG)_3SEG(POG)_3-NH_2</td>
<td>Liberty 1™ µw synthesiser, single coupling cycles using HCTU (4.5 eq.), DIPEA/NMP (4 eq.).</td>
</tr>
</tbody>
</table>
CMP-Lys-1Gal-OAc 4.20a

a) Characterising data for the 2KOAc variant

![Chemical structure and RP-HPLC trace]

\[ R_t = 12.9 \text{ min} \]

\[ [M+3H]^3+ \]

769.3

\[ [M+2H]^2+ \]

1153.1

**Figure 4.13.** The 2KOAc variant of CMP-Lys-1Gal-OAc 4.20a; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
CMP-Lys-1Gal-OAc 4.20a

b) Characterising data for the 1KOAc variant

\[ \text{Figure 4.14} \] The 1KOAc variant of CMP-Lys-1Gal-OAc 4.20a; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H\(_2\)O, Buffer B = 0.1% TFA in CH\(_3\)CN.
Figure 4.15 CMP-Lys-2Gal-OAc, 4.21a; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 – 40% B, 1% B per min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.16 CMP-Lys-3Gal-OAc, 4.22a; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.17 CMP-Ser-1Gal-OAc, **4.24a**; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
CMP-Ser-2Gal-OAc 4.25a

Figure 4.18 CMP-Ser-2Gal-OAc 4.25a; chemical structure, RP-HPLC trace (92% purity) and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
CMP-Ser-3Gal-OAc 4.26a

Figure 4.19 CMP-Ser-3Gal-OAc 4.26a; chemical structure, RP-HPLC trace (82% purity) and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.20 CMP-Lys-control 4.23; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.21 CMP-Ser-control, 4.27; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
4.6 Purification and deacetylation of Family 2 CMPs

Using the same deacetylated methods described in Section 3.13 for Family 1 neoglycopeptides Adpn(66-83)-Lys-3Gal-OAc 3.128 and Adpn(66-83)-Ser-3Gal-OH 3.129, crude O-acetylated CMPs 4.20a - 4.22a and 4.24a – 4.26a were deacetylated using freshly prepared sodium methoxide solution (1 M) at pH 11 overnight at room temperature. The crude peptides were purified by RP-HPLC to afford O-deacetylated neoglycopeptides 4.20b – 4.22b and 4.24b – 4.26b in >95% purity, identified by HPLC and LC-MS.

The deacetylated neoglycopeptide CMP-Lys-1Gal-OH, 4.20b was afforded in 73% yield (for deacetylation and purification step, 5.3% overall yield including synthesis of acetylated peptide) and was characterised by the [M+2H]^{2+} (calculated 111.1, observed 1111.0) and [M+3H]^{3+} (calculated 741.0, observed 741.3) charged states, Figure 4.22.

The deacetylated neoglycopeptide CMP-Lys-2Gal-OH, 4.21b was afforded in 15% yield (for deacetylation and purification step, 0.9% overall yield including synthesis of acetylated peptide) and was characterised by the [M+2H]^{2+} (calculated 1390.4, observed 1390.6) and [M+3H]^{3+} (calculated 927.3, observed 927.3) charged states, Figure 4.23.

The deacetylated neoglycopeptide CMP-Lys-3Gal-OH, 4.22b was afforded in 46% yield (for deacetylation and purification step, 1.5% overall yield including synthesis of acetylated peptide) and was characterised by the [M+2H]^{2+} (calculated 1113.3, observed 1113.6) and [M+3H]^{3+} (calculated 835.2, observed 835.5) charged states, Figure 4.24.

The deacetylated neoglycopeptide CMP-Ser-1Gal-OH, 4.24b was afforded in 11% yield (for deacetylation and purification step, 0.5% overall yield including synthesis of acetylated peptide) and was characterised by the [M+2H]^{2+} (calculated 1090.0, observed 1090.4) and [M+3H]^{3+} (calculated 727.0, observed 727.0) charged states, Figure 4.25.

The deacetylated neoglycopeptide CMP-Ser-2Gal-OH, 4.25b was afforded in 2% yield (for deacetylation and purification step, 0.4% overall yield including synthesis of acetylated peptide) and was characterised by the [M+2H]^{2+} (calculated 1348.2, observed 1348.6) and [M+3H]^{3+} (calculated 899.1, observed 899.3) charged states, Figure 4.26.

The deacetylated neoglycopeptide CMP-Ser-3Gal-OH, 4.26b was afforded in 5% yield (for deacetylation and purification step, 0.4% overall yield including synthesis of acetylated peptide) and was characterised by the [M+2H]^{2+} (calculated 1606.3, observed 1606.9) and [M+3H]^{3+} (calculated 803.7, observed 804.0) charged states, Figure 4.27.
Figure 4.22 CMP-Lys-1Gal-OH 4.20b chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.23 CMP-Lys-2Gal-OH 4.21b RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.24 CMP-Lys-3Gal-OH 4.22b RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN.
Figure 4.25 CMP-Ser-1Gal-OH 4.24b RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.26 CMP-Ser-2Gal-OH 4.25b RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 4.27  CMP-Ser-3Gal-OH 4.26b RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
4.7 Key observations from CD studies of Family 2 CMPs

Buffer solutions of \( O \)-acetylated CMPs \(4.20a - 4.22a\) and \(4.24a - 4.27\), \( O \)-deacetylated \textit{neo}glycopeptides \(4.20b - 4.22b\) and \(4.24b - 4.26b\) and control, unglycosylated CMPs \(4.23\) and \(4.27\) were prepared in potassium phosphate buffer at pH 7.4 and incubated overnight at 0 °C before being examined at 6 °C using circular dichroism spectroscopy (CD).

Three key observations were made during the CD studies of the Family 2 CMP family. Firstly, acetylation of the glycans reduced the triple helical properties of the peptide. Secondly, increasing levels of glycosylation decreased the triple helical content of the peptide and finally the serine-containing CMPs adopted a higher degree of triple helical content compared to their lysine counterparts. Further details of these observations and the proposed hypotheses to explain these trends are outlined in more detail below.

4.7.1 Structural differences observed between acetylated and deacetylated CMPs

The first key observation in Family 2 CMPs was the difference in the CD spectrum between acetylated monoglycan-serine CMP \(4.24a\) (CMP-Ser-1Gal-OAc) and its deacetylated counterpart \(4.24b\) (CMP-Ser-1Gal-OH). The deacetylated CMP \(4.24b\) (CMP-Ser-1Gal-OH) exhibited the characteristic features of a triple helix (with a weak maximum near 225 nm and a strong minimum near 200 nm), whereas the acetylated \(4.24a\) peptide (CMP-Ser-1Gal-OAc) was less able to adopt a triple helical conformation, with considerably smaller minimum and maximum values observed, Figure 4.28. The reduction in magnitude of the maximum at 225nm has been shown by Wallace \textit{et al.}\textsuperscript{229} to correspond to an uncoiled triple helical conformation.
A similar trend was also observed, to a lesser extent, with the acetylated diglycan-serine CMP 4.25a (CMP-Ser-2Gal-OAc) and its deacetylated counterpart 4.25b (CMP-Ser-2Gal-OH), Figure 4.29.

Figure 4.29 CD spectra comparing acetylated (4.25a) and deacetylated variants (4.25b) of a CMP bearing a double glycosylation recorded at 6 °C at pH 7.4 using potassium phosphate buffer.

The trend was not observed for the acetylated triglycan-serine CMP 4.26a and the deactylated CMP 4.26b, with both of these CMPs exhibiting a random coil conformation in the tested buffer solution, Figure 4.30. This is possibly due to steric hinderance resulting from closely spaced glycan groups in the triglycan CMPs.

Unfortunately, due to spontaneous deacetylation of acetylated lysine CMPs 4.20a (CMP-Lys-1Gal-OAc), 4.21a (CMP-Lys-2Gal-OAc) and 4.22a (CMP-Lys-2Gal-OAc) accurate circular dichroism data could not be gathered for these peptides, therefore the trend for the lysine CMPs 4.20a - 4.22a could not be established. The observed differences between the acetylated and deacetylated
serine CMPs indicate that a more extensive hydrogen bonding network is likely to be present in glycosylated CMPs than is currently accounted for in current models.\textsuperscript{203}

4.7.1.1 The role interstrand hydrogen bonding networks for structural stability of CMPs

A reasonable explanation for this observed difference could be that the acetyl group is likely to act as a ‘mask’ for interstrand hydrogen bonding between adjacent glycans thus destabilising the triple helical conformation formed by the acetylated peptide. Indeed, the free hydroxyl groups of the glycans in deacetylated peptides 4.24b (CMP-Ser-1Gal-OH) and 4.25b (CMP-Ser-2Gal-OH) are likely to contribute to both glycan-glycan and glycan-peptide hydrogen bonding networks of the CMP, thus stabilising the triple helix formed by the deacetylated peptides 4.24b (CMP-Ser-1Gal-OH) and 4.25b (CMP-Ser-2Gal-OH) compared to the acetylated variants 4.24a (CMP-Ser-1Gal-OAc) and 4.25a (CMP-Ser-2Gal-OAc) respectively.

In order to more easily visualise the possible interstrand hydrogen bonding interactions between the glycans, a simple triple helical wheel diagram has been constructed for peptides 4.24a (CMP-Ser-1Gal-OAc) and 4.24b (CMP-Ser-1Gal-OH), Figure 4.31, based on that described by Goh et al.\textsuperscript{230}, which was derived using published X-ray crystal structures of CMPs. The triple helical wheel is a projection down the long-molecular axis in which a triple helix containing GXY repeats is marked out by dashed lines and numbered accordingly. Given that X-ray crystal structures of CMPs have shown 7/2 symmetry,\textsuperscript{203} (seven residues per two turns of the triple helix leading to a 360° rotation), we included seven repeat units in the triple helical wheel diagram. This approach also illustrates the low probability that the acetyl modifying groups on the glycans destabilise the acetylated CMPs via steric hindrance, as the glycans are well spaced around the triple helix with a low propensity for steric interactions.
The importance of intrastrand hydrogen bonding networks for stabilisation of the triple helical conformation of CMPs

In addition to interstrand hydrogen bonding networks, intrastrand hydrogen bonding must also be taken into account. Wetmore et al.\textsuperscript{205} illustrated two types of intrastrand hydrogen bonding which can occur in glycosylated CMPs: namely glycan-glycan 2-OH to 6-OH interactions and glycan-peptide backbone interactions from glycan 6-OH to the carbonyl group of the peptide backbone as illustrated in Figure 4.32.

4.7.2 Other reported uses of an acetyl protecting group to investigate hydrogen bonding networks in CMPs

The utilisation of an acetyl group to mask and thus investigate the nature of interstrand hydrogen bonding of CMPs has previously been reported by Raines et al.\textsuperscript{221,222} who used acetylation to probe the importance of the free hydroxyl group of Hyp residues in the CMP H-(POG)\textsubscript{10}-OH. Their study provided evidence that O-acetylation of Hyp residues of the CMP did not disrupt the ability of the acetylated
peptide to form a triple helix in solution. This result indicates that the Hyp residues do not participate in interstand hydrogen bonding contributing to triple helical formation of the CMPs. Indeed, X-ray crystal structures of CMPs place the free hydroxyl group of Hyp in the periphery of the triple helix, which would prevent these hydroxyl groups from participating in direct interstrand hydrogen bonding. Thus, acetylation of Hyp in CMPs may not disrupt triple helical formation by these peptides.

In contrast, the difference in secondary structure between the acetylated and deacetylated CMPs shown for CMP pairs 4.24a (CMP-Ser-1Gal-OAc) and 4.24b (CMP-Ser-1Gal-OH) and 4.25a (CMP-Lys-1Gal-OAc) and 4.25b (CMP-Lys-1Gal-OH) demonstrate that acetylation of glycans does significantly impair the CMPs investigated in the present study from forming a triple helix. It is probable that the glycans contribute to crucial structural, probably inter-strand hydrogen bonding, thus further stabilising the triple helix. In essence, the mechanism of stabilisation by the free hydroxyl groups on the glycans must then differ from that of the free hydroxyl groups on the Hyp side chains.

4.7.2.1 The use of facilitating hydration networks for extension of hydrogen bonding

The role of facilitating hydration networks for interstrand hydrogen bonding for the CMPs investigated herein must also be considered. Bridging water molecules have been shown to aid interstrand hydrogen bonding, with key structural changes observed upon dehydration of collagens.231 X-ray crystal structures have also illustrated an extensive hydration network mediating interstrand hydrogen bonding between Hyp groups and carbonyl groups of the peptide backbone.203 It is curious that such a hydration network was not disrupted upon Hyp acetylation in the studies reported by Raines et al.221,222 Water mediated hydrogen bonding between carbonyl and amide groups of adjacent chains has also been observed in X-ray crystal structures.232

As mentioned previously in Section 4.3.1, Wetmore et al.205 have established using thermodynamic calculations that bridging water molecules are key for the increased thermal stability of their Hyp-glycosylated CMP Ac-(Hyp-β-Gal)-NH₂ 4.13, which was reported to have T_m of 70 °C compared to the unglycosylated CMP Ac-(Hyp)-NH₂ 4.12 with a T_m of 38 °C. Using molecular dynamic simulations, an extensive hydrogen bonding network was shown to exist between glycan hydroxyl groups, bridging water molecules and the peptide backbone. Therefore, the possibility of bridging water molecules facilitating interstrand hydrogen bonding must also be considered of key structural importance for the CMPs reported in this thesis. Indeed, the key structural differences observed between acetylated and deacetylated CMPs reported herein indicate that such a hydration network would be disturbed by acetylation of glycan hydroxyl groups, leading to the reduced triple helical propensity of acetylated CMPs 4.24a (CMP-Ser-1Gal-OAc) and 4.25a (CMP-Ser-2Gal-OAc) compared to their deacetylated analogues 4.24b (CMP-Ser-1Gal-OH) and 4.25b (CMP-Ser-2Gal-OH).
4.7.3 Comparison of control non-glycosylated CMP with glycosylated CMP

Examination of the control CMPs 4.23 (CMP-Lys-Control) and 4.27 (CMP-Ser-Control) revealed a strikingly similar profile to that of mono-deacetylated CMPs 4.20b (CMP-Lys-1Gal-OH) and 4.24b (CMP-Ser-1Gal-OH), Figure 4.33.

![CD spectra comparing triazole containing CMPs](image)

**Figure 4.33** CD spectra comparing triazole containing CMPs a) 4.24a (CMP-Ser-1Gal-OAc) and 4.24b (CMP-Ser-1Gal-OH) with CMP-Ser-Control 4.27 and b) 4.20b (CMP-Lys-1Gal-OH) with CMP-Lys-Control 4.23 recorded at 6 °C at pH 7.4 using potassium phosphate buffer.

This observation is in agreement with the observations reported by Wennemers et al.\textsuperscript{227} previously mentioned in Section 4.3.2, whereby glycosylated ‘clicked’ CMPs 4.15, 4.16 were shown to have similar CD profiles to their non-glycosylated CMP analogues 4.14 and 4.28, Figure 4.34.
These combined observations indicate that use of a triazole linker to synthesise glycosylated CMPs is a suitable mimetic to ‘natural’ CMPs. In addition, this result shows us that a similar interstrand hydrogen bonding network must exist in the Lys and Ser control CMPs 4.23 and 4.27 as occurs in the singly glycosylated, deacetylated CMPs 4.20b (CMP-Lys-1Gal-OH) and 4.24b (CMP-Ser-1Gal-OH). In the case of the control unglycosylated CMPs 4.23 (CMP-Lys-Control) and 4.27 (CMP-Ser-Control), the additional interchain hydrogen bonds could exist between the Lys amine groups or serine side-chain hydroxyl groups whereas in the singly glycosylated, deacetylated CMPs 4.20b (CMP-Lys-1Gal-OH) and 4.24b (CMP-Ser-1Gal-OH), the additional interchain hydrogen bonds could exist between the glycan hydroxyl groups.

4.7.4 Investigation of different levels of glycosylation of Family 2: CMPs

The effect of increased levels of glycosylation on peptide secondary structure was then investigated. The secondary structure of CMPs containing one, two or three ‘click’ deacetylated glycosyl amino acid residues (4.20b - 4.22b and 4.24b – 4.26b) were investigated using CD. The CD studies of the deacetylated glycosylated CMP series 4.20b (CMP-Lys-1Gal-OH), 4.21b (CMP-Lys-2Gal-OH), 4.22b (CMP-Lys-3Gal-OH) and 4.24b (CMP-Ser-1Gal-OH), 4.25b (CMP-Ser-2Gal-OH), 4.26b (CMP-Ser-3Gal-OH) provided key insight into the correlation between the degree of glycosylation and peptide secondary structure. It was observed that increased glycosylation (greater number of glycosylated amino acid residues) of the CMPs resulted in decreased triple helix formation, Figure 4.35.

Figure 4.34 Chemical structures and CD spectra of glycosylated CMPs 4.15 and 4.16 and control CMPs 4.14 and 4.28 as reported by Wennemers et al.\textsuperscript{143,227}
In both the serine-containing and lysine-containing CMP families, it was observed that although incorporation of a single ‘clicked’ galactolysine residue 3.123 or a ‘clicked’ galactoserine residue 3.124 into the CMP sequence was readily tolerated, incorporation of two or more glycosylated residues resulted in a corresponding decrease of the triple helical content of the peptide.

4.7.4.1 The effect of steric interactions for triple helical conformation of CMPs

One potential explanation for this observation involves steric interactions between the bulky glycan groups: increased glycosylation could lead to a more congested space incorporating the peptide structure,
thus decreasing the ability of the resulting peptide to coil into a triple helical conformation. This hypothesis can be illustrated by use of a triple helical wheel diagram, Figure 4.36.

![Triple helical wheel diagrams](image)

**Figure 4.36** Triple helical wheel diagrams of lysine-containing CMPs 4.20b, 4.21b and 4.22b to illustrate the possible increased steric hindrance encountered upon incorporation of an increasing number of ‘click’ galactolysine 3.123 residues.

Furthermore, monoglycosylated CMPs 4.20b and 4.24b may be able to pre-order in solution to a greater extent than their di- or tri-glycosylated counterparts, due to the presence of fewer flexible glycans and therefore present a lower entropic cost for both pre-ordering in solution as well as triple helix formation.

### 4.7.4.2 The role of hydrogen bonding networks in glycosylated CMPs: monomeric single stranded state versus triple helical conformation

One must also consider the role of both interstand and intrastrand hydrogen bonding for the stability of both the monomeric single-strand state and the triple helical conformation of the CMP. As discussed earlier in Section 4.7.1, Wetmore et al.\(^{205}\) proposed that increased glycosylation of imino-rich CMPs (commonly found in plant collagens) leads to additional intrastrand hydrogen bonding both between glycans and between the glycans and the carbonyl groups of the peptide backbone. This extension of the hydrogen bonding network leads to increased folding of the monomeric peptide chain thus stabilising the monomeric single-strand conformation of the CMP relative to the triple helical conformation. Therefore, equilibrium between the triple-helical conformation and the monomeric single strand conformation is shifted towards the single-stranded species, which becomes the dominant species in solution.

The observation for both the lysine CMP series 4.20b – 4.22b and the serine CMP series 4.24b - 4.26b that increased levels of glycosylation lead to a decreased propensity of the CMP to adopt a triple helical conformation, indicates that a similar mechanism to that proposed by Wetmore et al.\(^{205}\) may be operating for animal collagen. We have shown that increased levels of glycosylation of CMPs disfavour adoption...
of a triple helical structure. It is possible that as the Family 2 CMPs become increasingly glycosylated, the CMPs are able to participate in extended intrastrand hydrogen bonding networks, such as glycanc-glycan interactions and glycan-peptide interactions. These extended hydrogen bonding networks may thereby stabilise the monomeric single strand conformation relative to the triple strand, hence the more highly glycosylated CMPs such as 4.22b (CMP-Lys-3Gal-OH) and 4.26b (CMP-Ser-3Gal-OH) show less triple helical nature than the lesser glycosylated CMPs such as 4.20b (CMP-Lys-1Gal-OH) and 4.24b (CMP-Ser-1Gal-OH).

4.7.4.3 Role of guest domain length in structural stabilisation of CMPs

One must also take into account the observation that the extension of the length of the ‘guest’ domain from 4.20b \( \rightarrow \) 4.21b \( \rightarrow \) 4.22b and 4.24b \( \rightarrow \) 4.25b \( \rightarrow \) 4.26b could also account for the conformational changes observed in CD spectra of these CMPs. An increase in the number of XEG triplets lowers the overall relative imino acid content of the CMP. This could cause a shift from the tight 7/2 conformation favoured by imino acid rich peptides towards the looser 10/3 conformation adopted by imino acid poor peptides.\(^{232}\) This change in conformation upon increased length of the ‘guest’ domain could disrupt both the intrastrand and interstand hydrogen bonding networks of the CMP, thus leading to reduced propensity of the peptide to adopt a triple helical conformation.

4.8 Proposal of an ‘optimal level’ of glycosylation for collagenous peptides

The observed difference in the nature of the conformation adopted the CMPs upon increased glycosylation as discussed in Section 4.7 could aid the understanding of the cause of collagen defects observed in diabetes.\(^{233}\) In diabetes, blood glucose and tissue glucose levels are elevated, which could lead to increased glycosylation and glycation of collagens present in these environments. Glycosylation of such collagens could lead to uncoiling of the protein’s triple helix leading to a dominant single-strand conformation. This proposed change in collagen structure could account for the collagen defects observed in diabetes.

Indeed, the results presented in Section 4.7 indicate that there could be an ‘optimal level’ of glycosylation of collagens, in the case of CMPs 4.20b \( \rightarrow \) 4.26b the singly glycosylated CMPs 4.20b (CMP-Lys-1Gal-OH) and 4.24b (CMP-Ser-1Gal-OH). Above this ‘optimal level’ of glycosylation, the triple helical nature of the protein collapses. In the case of the CMPs synthesised and discussed in this thesis, there was a significant reduction in triple helical conformation for the di- and tri-glycosylated CMPs 4.22b (CMP-Lys-2Gal-OH), 4.23b (CMP-Lys-3Gal-OH), 4.25b (CMP-Ser-2Gal-OH) and 4.26b (CMP-Ser-3Gal-OH). When this hypothesis is applied to the pathogenesis of diabetes, one could hypothesise that the increased tissue sugar levels and subsequent glycosylation of collagens above the optimal level, could lead to uncoiling of the triple helical conformation of the collagens resulting in the
collagen complications observed in diabetes, such as ventricular dysfunction,217 diabetic cardiomyopathy218 and diabetic nephropathy.219

This hypothesis is in agreement with previous work reported by Kivirikko et al.234 on the correlation of glycosylated hydroxylysine (Glyco-Hyl) content and the fibril formation of Type II human collagen. Kivirikko et al.234 demonstrated that proteins with high Glyco-Hyl (0.89%) content were only able to form thin fibrils, compared to those with lower Glyco-Hyl (0.13%) content which were able to form thick fibrils. These observations provide further evidence towards an ‘optimal level’ level of glycosylation for collagenous peptides. Furthermore, Dominguez et al.235 have reported that overglycosylation of collagen leads to reduced cross-link formation resulting in a decreased fibril diameter. They hypothesise that three clinical conditions may be associated with this overglycosylation, namely osteogenesis imperfecta, diabetes and bone loss observed during the menopause.

4.8.1 Comparison of lysine and serine CMP series

Comparison of the CD spectra between the lysine CMP series 4.20b – 4.22b and the serine CMP series 4.24b – 4.26b led to a surprising result. In all cases, the serine-containing CMPs displayed a higher degree of triple helical content than the lysine-containing CMPs. For example, see Figure 4.37 which compares the CD spectra lysine-containing CMP 4.20b (CMP-Lys-1Gal-OH) with its serine-containing analogue 4.24b (CMP-Ser-1Gal-OH).

Figure 4.37 CD spectra comparing lysine-containing CMP 4.20b and the serine-containing counterpart 4.24b, recorded at 6 °C at pH 7.4 using potassium phosphate buffer.

On the surface, this result may seem counter-intuitive, as one may have expected the longer carbon skeleton of the ‘click’ galactolysine residue X1 in 4.20b (compared to the shorter ‘click’ galactoserine X2 present in analogue 4.24b) could more effectively decrease steric interactions between the large glycan and core of the triple helix. However, on the other hand, by virtue of the longer carbon skeleton of X1, the ‘click’ galactolysine residues are more conformationally flexible than their shorter ‘click’
galactoserine $X_2$ counterparts. This greater conformational flexibility leads to a greater entropic cost for the galactolysine residues to form interstrand hydrogen bonding networks that stabilise the triple helix. Therefore, CMPs containing the longer galactolysine residues, such as CMP 4.20b, may form a triple helix to a lesser extent than CMPs containing the shorter galactoserine residue, such as CMP 4.24b. In order to test this hypothesis CMPs 4.20b and 4.24b were computationally modelled using the 1QSU (unique identifier code) crystal structure of Ac-(POG)$_3$EKG(POG)$_5$-NH$_2$, the most structurally similar published sequence to 4.20b and 4.24b. Using an Amber 99 Force Field model, the potential rotamers for the glycosylated serine and lysine residues were modelled and there from the probable conformations of the resulting peptides were derived, Figure 4.38.

Figure 4.38 a) CMP sequences of model 1QSU crystal structure and the structures of CMPs 4.20b and 4.24b; b) Force Field model of lysine-containing CMP 4.20b; c) Force Field model of serine-containing CMP 4.24b.
Both CMPs 4.20b and 4.24b were found to fit well into the 7/2 triple helical conformation derived from the 1QSU model. No additional hydrogen bonding was observed due to the oxygen heteroatom present in the galactoserine analogue 4.24b as compared to the corresponding galactolysine analogue 4.20b. Additionally, both model peptides showed the glycans to be interacting with the core of the triple helix. Thus, these results tie in with the published the ‘twisted hairy-rope’ model\textsuperscript{226} discussed in Section 4.3.2.

With the computational models of 4.20b and 4.24b in hand, one can more easily visualise the triple helical conformation of these CMPs. As mentioned above, the propensity of the CMPs to form a triple helix is likely to depend on the entropic cost for formation of the rigid triple helical structure, as opposed to a looser single strand monomeric conformation. The role of potential pre-ordering of the peptide strands prior to triple helix formation must also be taken into account. In a cellular environment, collagen chains pre-order by association and alignment of the three peptide chains prior to triple helix formation.\textsuperscript{236} More extensive pre-ordering of polypeptide chains prior to triple helix formation could be also a contributing factor to the extent of triple helix formation exhibited by our CMPs. If the serine-containing collagen CMP 4.24b was able to pre-order to a greater extent than its lysine-containing CMP counterpart 4.20b, this could also reduce the entropic cost of triple helix formation and therefore account for the higher degree of triple helix formation of serine CMP 4.24b compared the lysine CMP 4.20b.

4.9 Thermal stability of Family 2: CMPs

Glycosylation has been shown to affect the thermal stability of triple helical CMPs.\textsuperscript{143,228} Therefore, the thermal stability of triple helical CMPs 4.20b and 4.24b alongside the equivalent non-glycosylated control peptides 4.23 and 4.27 were measured, Figure 4.39.

![Thermal melt data of the CMP series for control and CMP-1Gal-OH CMPs recorded at pH 7.4 using potassium phosphate buffer.](image)

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
<th>$T_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.23</td>
<td>Ac(POG)$_3$KEG(POG)$_3$NH$_2$</td>
<td>35.8 °C</td>
</tr>
<tr>
<td>4.20b</td>
<td>Ac(POG)$_3$X$_1$EG(POG)$_3$NH$_2$</td>
<td>28.8 °C</td>
</tr>
<tr>
<td>4.27</td>
<td>Ac(POG)$_3$SEG(POG)$_3$NH$_2$</td>
<td>26.2 °C</td>
</tr>
<tr>
<td>4.24b</td>
<td>Ac(POG)$_3$X$_2$EG(POG)$_3$NH$_2$</td>
<td>25.7 °C</td>
</tr>
</tbody>
</table>
For both the lysine-containing and serine CMP series, the glycosylated peptides 4.20b and 4.24b displayed lower thermal stability than their non-glycosylated analogues 4.23 and 4.27. Also of note, is the observation that for both the glycosylated and non-glycosylated CMPs, the lysine-containing analogues displayed higher thermal stability compared to the serine-containing analogues.

4.9.1 Comparison of thermal stabilities of Family 2: CMPs with previously reported similar peptides

As discussed above in Section 4.3.4, Bächinger et al.\textsuperscript{228} reported that threonine glycosylation was shown to be essential for triple helical formation of cuticle collagens found in deep sea worms. Their glycosylated CMP Ac-[Gly-Pro-Thr(β-Gal)]\textsubscript{10}-NH\textsubscript{2} \textbf{4.18} was found to have high thermal stability (T\textsubscript{m} of 40 °C), whereas the non-glycosylated CMP Ac-(Gly-Pro-Thr)\textsubscript{10}-NH\textsubscript{2} \textbf{4.19} was found to be unable to adopt a triple helical conformation. In addition, as discussed above in Section 4.3.1, Wetmore et al.\textsuperscript{205} have reported that the glycosylated CMP Ac-(Hyp-β-Gal)\textsubscript{9}-NH\textsubscript{2} \textbf{4.13} had higher thermal stability (T\textsubscript{m} of 70 °C) compared to the non-glycosylated variant Ac-(Hyp)\textsubscript{9}-NH\textsubscript{2} \textbf{4.12} (T\textsubscript{m} of 38 °C).

These two reports which demonstrated increased glycosylation leads to increased thermal stability of CMPs are in contrast with the observed thermal stabilities of the Family 2 CMPs 4.20b – 4.22b and 4.24b – 4.26b, which showed decreased thermal stability with increasing glycosylation of the peptide. This conflicting observation can be explained by the differences in the amino acids glycosylated in each study and the positional differences of the glycosylated residue. The glycosylated CMP studied by Bächinger et al.\textsuperscript{228} \textbf{4.18} was based on threonine glycosylation and CMPs \textbf{4.12} and \textbf{4.13} investigated by Wetmore et al.\textsuperscript{205} were based on Hyp glycosylation. In contrast, the Family 2 CMPs discussed in this thesis incorporated ‘click’ glycosylated mimetics of serine and lysine residues into the peptide sequence. In addition, there are key positional differences between the Family 2 CMPs and CMP \textbf{4.18} synthesised by Bächinger et al.\textsuperscript{228} \textbf{4.18} incorporated the glycosylated threonine residues at the Y position of the collagen sequence, whereas the Family 2 CMPs incorporate the ‘click’ glycosylated residue at the X position. Such positional variation of residues leading to different thermal stabilities of CMPs has been reported by Wennemers et al.\textsuperscript{213} who showed that incorporation of (4S)-acetamidoproline into the X position of a CMP decreased the thermal stability of the CMP, compared to the same residue being present at the Y position.

Furthermore, the observation of increased levels of glycosylation leading to decreased thermal stability of the Family 2 CMPs is consistent with a similar study also utilising neoglycosides as reported by Wennemers et al.\textsuperscript{227}. In agreement with the results reported herein for Family 2 CMPs, Wennemers et al.\textsuperscript{227} reported that introduction of ‘click’ glycans decreased the thermal stability of the CMPs relative to the unglycosylated peptides.
4.10 Summary of results for Chapter Four

In summary, the trial CMP series of neo
glycopeptides 4.20 – 4.27 demonstrated that the collagen model peptide system characterised by GOP repeat units, is a good model for glycosylated collagenous peptides. Neo
glycosides Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 or Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 were successfully incorporated into CMP neo
glycopeptides using Fmoc-SPPS to afford O-acetylated CMPs 4.20a – 4.22a (lysine-based CMPs) and 4.24a - 4.26a (serine-based CMPs) in moderate yields after purification using RP-HPLC. Deacetylation of the O-acetylated CMPs afforded deacetylated CMPs 4.20b – 4.22b and 4.24b - 4.26b in moderate yields after purification using RP-HPLC.

Examination of secondary structures formed by the O-acetylated CMPs 4.20a – 4.22a and 4.24a - 4.26b, deacetylated CMPs 4.20b – 4.22b and non-glycosylated, control CMPs 4.23 and 4.27 using circular dichroism generated two key insights into the role of glycans on the structure and function of collagens.

The first key insight was the difference observed in the CD spectra of the O-acetylated and O-deacetylated CMPs. The O-acetylated CMPs were less able to form triple helices than their deacetylated counterparts, which suggest existence of a far more extensive network of interstrand hydrogen bonding in glycosylated collagens than has hitherto been recognised.

Secondly, the investigations into different levels of glycosylation in CMP series 4.20 – 4.47 and the subsequent effect on collagen structure revealed that there could be an ‘optimal level’ of glycosylation for collagens. This result indicates that although low levels of glycosylation are tolerated in the collagen triple helix, high levels of glycosylation destabilise triple helix formation. These conclusions could have major implications on the understanding of the forces that underpin the structure and function of collagens.

4.11 Conclusion for Chapter Four

In conclusion, the trial CMP system was able to support the incorporation of neo
glycosides Fmoc-Lys-(Click-Gal-OAc)-OH 3.123 or Fmoc-Ser-(Click-Gal-OAc)-OH 3.124 into a ‘host-guest’ CMP sequence. The resultant O-acetylated CMPs 4.20a – 4.22a and 4.24a - 4.26a were readily deacetylated to afford the desired deacetylated CMPs 4.20b – 4.22b and 4.24b – 4.26b. Examination of the structural properties of the resultant neo
glycopeptides revealed triple helical conformations for the control, non-glycosylated and mono-glycosylated CMPs 4.23, 4.27 and 4.20b and 4.24b respectively.

The di- and triply-glycosylated CMPs 4.21b, 4.22b, 4.25b and 4.26b were found to adopt a random coil conformation, probably due to a combination of steric clashes between the bulky glycan residues and a
reduced hydrogen bonding capabilities in the triple helical conformation as opposed to the monomeric single-stranded state.

On the basis of this successful trial synthetic study, the conclusion can be reached that the CMP system is a good model for native glycosylated collagen systems. The next Chapter will detail the synthesis and examination of secondary structures of a CMP-Adpn hybrid family of neoglycopeptides.
Chapter Five

Lysine, the Keystone of the Triple Helical Structure of the Collagenous Domain of Adiponectin
Chapter Five: Lysine, the Keystone of the Triple Helical Structure of the Collagenous Domain of Adiponectin

5.1 Summary of work presented in previous chapters

The synthesis and circular dichroism (CD) investigation of the secondary structures adopted by ‘triple-clicked’ 18-mer neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser-3Gal-OH 3.59b was outlined in Chapter Three. The CD study of 18-mer Adpn neoglycopeptides 3.58b and 3.59b showed that these two neoglycopeptides adopted a random coil conformation, in preference to the desired triple helical conformation common to collagenous peptides. It is probable that these 18-mer Adpn neoglycopeptides were unable to adopt a triple helical conformation due to both the short nature of the 18-mer peptide sequence as well as N-termini interchain repulsion.

![Figure 5.1](image)

Figure 5.1 The chemical structures of Family 1 neoglycopeptides Adpn(66-83)-Lys-3Gal-OH 3.58b and Adpn(66-83)-Ser 3Gal OH 3.59b, which were unable to adopt the triple helical conformation common to collagenous peptides.

Therefore, ‘host-guest’ collagen model peptide sequences (CMPS) were investigated as a mimetic for glycosylated collagenous peptides. Chapter Four describes the synthesis of a series of trial ‘host-guest’ CMP neoglycopeptides 4.20 – 4.27 with differing levels of glycosylation of the ‘guest’ GEX domain where \( X \) donates the neoglycoside incorporated into the peptide sequence, Figure 5.2. These CMP peptides were based on the ‘clicked’ CMPs 4.14 – 4.17 with sequence Ac-(POG)₃PO*G(POG)₃-NH₂ (where O* is a neoglycosylated hydroxyproline residue) reported by Wennemers et al.¹⁴³,²²⁷ which were...
shown to adopt the desired triple helical conformation. The ‘host’ domain of (POG) triplets was employed to lend structural support to the GEX ‘guest’ domain and thereby facilitate formation of a triple helical secondary structure. Notably, the CMP neo-glycopeptides 4.20 – 4.27 are acetylated at the $N$-terminus and amidated at the $C$-terminus to lend extra stability to the compounds.

Examination of the CMP neo-glycosides using CD revealed that the control, non-glycosylated and mono-glycosylated CMPs 4.23, 4.27 and 4.20b and 4.24b were able to adopt a triple helical conformation, whereas the di- and triply-glycosylated CMPs 4.21b, 4.22b, 4.25b and 4.26b were found to adopt a random coil conformation. Is it probable that a combination of steric clashes between the bulky glycan residues and a reduced hydrogen bonding capabilities in the triple helical conformation as opposed to the monomeric single-stranded state lead to the di- and triply-neo-glycosylated CMPs being unable to adopt the triple helical conformation common to collagenous peptides.

Figure 5.2 The chemical structures of Family 2 CMP neo-glycopeptides 4.20 – 4.27 with the propensity to adopt a triple helical conformation shown to decrease with increasing number of glycans incorporated into the peptide sequence.

This Chapter combines the synthetic methods discussed in Chapter Three (Family 1 Adpn neo-glycopeptides) and Chapter Four (Family 2 CMP neo-glycopeptides) to afford a family of hybrid CMP-Adpn neo-glycopeptides, Family 3, Figure 5.3. Similar to Family 2, the Family 3 CMP-Adpn neo-glycopeptides also use the (POG) repeat unit as a ‘host’ domain to facilitate the neo-glycopeptides towards adoption of a triple helical secondary structure. In the case of Family 3, the ‘host’ repeat unit takes the form of (GPO) triplets to match the GXY sequence of the 18-mer Adpn ‘guest’ sequence.
Chapter Five

The synthesis of Family 3 CMP-Adpn neoglycopeptides \(5.1a - 5.4\) will be discussed and the secondary structures of the resultant peptides will be examined using circular dichroism studies. In line with the previously discussed neoglycopeptide families, both lysine- and serine-containing peptides will be synthesised to investigate the effect of carbon chain length linker to the glycan on the triple helical properties of the peptide.

In addition to the triply-glycosylated CMP-Adpn neoglycopeptides CMP-Adpn(66-83)-Lys-3Gal-OH \(5.1b\) and CMP-Adpn(66-83)-Ser-3Gal-OH \(5.3b\), control unglycosylated variants CMP-Adpn-Lys-Control \(5.2\) and CMP-Adpn-Ser-Control \(5.4\) will also be synthesised and the structural properties discussed in order to probe the importance of glycosylation for the secondary structure of the peptide.

5.2 The synthesis of a ‘host-guest’ CMP-Adpn family of peptides

Previously synthesised neoglycosides Fmoc-Lys-(Click-Gal-OAc)-OH \(3.123\) and Fmoc-Ser-(Click-Gal-OAc)-OH \(3.124\) were incorporated into \(O\)-acetylated CMP-Adpn peptides \(5.1a\) and \(5.3a\) using Fmoc-SPPS. CMP-Adpn peptides \(5.1b - 5.4\) were prepared on “in house”\(^{197}\) synthesised aminomethyl polystyrene (AM-PS) and immobilised via Rink amide linker \(3.94\). Due to equipment availability, different microwave synthesisers were employed for synthesis of the CMP-Adpn peptides.
CMP-Adpn(66-83)-Lys-3Gal-OAc $5.1a$ was prepared on a 0.033 mmol scale using the Tribute™ synthesiser (Protein Technologies, Tuscon, Az) was used for all coupling reactions, except for the coupling of the ‘pre-clicked’ neoglycoside Fmoc-Lys-(Click-Gal-OAc)-OH $3.123$ performed using a CEM™ microwave (CEM Corporation, Mathews, NC). Fmoc deprotection and coupling reactions were carried out using HCTU/DIPEA and piperidine as the coupling and Fmoc deprotection reagents, respectively for all amino acids except the ‘pre-clicked’ neoglycoside Fmoc-Lys-(Click-Gal-OAc)-OH $3.123$ when the stronger HATU coupling agent was used.

On completion of peptide sequence, the peptidyl-resin was Fmoc-deprotected (as above) and N-terminal acetylated using acetic anhydride and DIPEA. Cleavage of the acetylated peptidyl-resin with TFA, TIPS, H$_2$O (93: 3.5: 5.5, v/v) afforded the crude O-acetylated neoglycopeptide CMP-Adpn(66-83)-Lys-3Gal-OAc $5.1a$ which was identified by flow-inject mass spectrometry (FI-MS) by the [M+3H]$^3+$ (calculated 1560.6, observed 1561.1) and [M+4H]$^4+$ (calculated 1170.7, observed 1170.9) charged states, Figure 5.4.

For the synthesis of CMP-Adpn(66-83)-Ser-3Gal-OAc $5.3a$, which was conducted on a 0.1mmol scale, the first four triplets were coupled onto the functionalised resin using the Tribute™ Peptide Synthesiser employing HCTU/ NMM in NMP and piperidine as the coupling and Fmoc-deprotection agents respectively to afford Fmoc-GDI(GPO)$_3$Rink-ϕ. Elongation of the peptidyl-resin was then conducted using the CEM™ microwave with HCTU/ DIPEA and piperidine as the as the coupling and Fmoc-deprotection agents respectively, except for incorporation of the pre-clicked’ neoglycoside Fmoc-Lys-(Click-Gal-OAc)-OH $3.124$ when the stronger HATU coupling agent was used. These coupling conditions afforded peptidyl resin Fmoc-GPOGE*GEX*GDOGLIGO*X*GDI(GPO)$_3$Rink-ϕ, where X* is the functionalised serine amino acid. The final two (GPO) triplets were then incorporated using the Tribute™ Peptide Synthesiser and HCTU/ NMM in NMP and piperidine as the coupling and Fmoc-deprotection agents respectively.

On completion of synthesis, the peptidyl-resin was Fmoc-deprotected (as above) and N-terminal acetylated using acetic anhydride and DIPEA. Cleavage of the acetylated peptidyl-resin with TFA, TIPS (triisopropylsilane), H$_2$O (90:5:5, v/v) afforded the crude O-acetylated neoglycopeptide CMP-Adpn(66-83)-Ser-3Gal-OAc $5.3a$, which was identified by FI-MS by the [M+3H]$^3+$ (calculated 1504.6, observed 1505.1) and [M+4H]$^4+$ (calculated 1128.7, observed 1128.7) charged states, Figure 5.5.

The crude O-acetylated peptides $5.1a$ and $5.3a$ were then deacetylated using freshly prepared sodium methoxide solution (1 M) at pH 11 overnight at room temperature. The crude deacetylated peptides were then purified by RP-HPLC to afford neoglycopeptides CMP-Adpn(66-83)-Lys-3Gal-OH $5.1b$ and CMP-Adpn(66-83)-Ser-3Gal-OH $5.3b$ in >95% purity.
Chapter Five

The deacetylated neo-glycopeptide CMP-Adpn(66-83)-Lys-3Gal-OH \textbf{5.1b} was afforded in 15% yield (for deacetylation and purification steps, 1.9% overall yield including synthesis of the acetylated peptide) and was characterised by the $[\text{M}+2\text{H}]^{2+}$ (calculated 1392.6, observed 1392.8) and $[\text{M}+3\text{H}]^{3+}$ (calculated 1044.7, observed 1044.9) charged states, Figure 5.6.

The deacetylated neo-glycopeptide CMP-Adpn(66-83)-Ser-3Gal-OH \textbf{5.3b} was afforded in 6% yield (for deacetylation and purification steps, 0.9% overall yield including synthesis of the acetylated peptide) and was characterised by the $[\text{M}+3\text{H}]^{3+}$ (calculated 1350.7, observed 1350.9) and $[\text{M}+4\text{H}]^{4+}$ (calculated 1013.3, observed 1013.5) charged states, Figure 5.7.

CMP-Adpn(66-83)-Lys-Control peptide \textbf{5.2} and CMP-Adpn(66-83)-Ser-Control peptide \textbf{5.4} were prepared using the Tribute™ Peptide Synthesiser employing HCTU/ NMM in NMP and piperidine as the coupling and Fmoc-deprotection agents respectively. On completion of peptide sequence, the peptidyl-resin was Fmoc-deprotected (as above) and $N$-terminal acetylated using acetic anhydride and rotary mixing. Cleavage of the acetylated peptidyl-resin with TFA, TIPS, H$_2$O (90: 5: 5, v/v) afforded crudes peptides which were purified by RP-HPLC to afford the desired peptides CMP-Adpn(66-83)-Lys-Control \textbf{5.2} and CMP-Adpn(66-83)-Ser-Control \textbf{5.4} in >95% purity, which were identified by LC-MS.

CMP-Adpn(66-83)-Lys-Control \textbf{5.2} was afforded in 1.6% yield and was characterised by the $[\text{M}+3\text{H}]^{3+}$ (calculated 1148.6, observed 1146.6) and $[\text{M}+4\text{H}]^{4+}$ (calculated 861.7, observed 861.8) charged states, Figure 5.8.

CMP-Adpn(66-83)-Ser-Control \textbf{5.4} was afforded in 4.4% yield and was characterised by the $[\text{M}+2\text{H}]^{2+}$ (calculated 1660.7, observed 1660.7), $[\text{M}+3\text{H}]^{3+}$ (calculated 1107.5, observed 1107.5) and $[\text{M}+4\text{H}]^{4+}$ (calculated 830.9, observed 831.0) charged states, Figure 5.9.
Figure 5.4 CMP-Adpn(66-83)-Lys-3Gal-OAc 5.1a; chemical structure and flow-injection mass spectrum.
Figure 5.5 CMP-Adpn(66-83)-Ser-3Gal-OAc 5.3a; chemical structure and flow-injection mass spectrum.
Figure 5.6 CMP-Adpn(66-83)-Lys-3Gal-OH 5.1b; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 5.7 CMP-Adpn(66-83)-Ser-3Gal-OH 5.3b; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 5.8 CMP-Adpn-(66-83)-Lys-Control 5.2: chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
Figure 5.9 CMP-Adpn(66-83)-Ser-Control 5.4; chemical structure, RP-HPLC trace and mass spectrum. The analytical RP-HPLC was performed using an analytical column (Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) at a flow rate of 1 mL/min, linear gradient 5 - 40% B over 35 min, where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN.
5.3 Circular dichroism studies of CMP-Adpn family of peptides

Buffer solutions of O-deacetylated neoglycopeptides CMP-Adpn(66-83)-Lys-3Gal-OH 5.1b and CMP-Adpn(66-83)-Ser-3Gal-OH 5.3b and control unglycosylated CMPs CMP-Adpn(66-83)-Lys-Control 5.2 and CMP-Adpn(66-83)-Ser-Control 5.4 were prepared in potassium phosphate buffer at pH 7.4 and incubated overnight at 0 °C before being examined at 6 °C using circular dichroism spectroscopy (CD).

Examination of the secondary structures of CMP-Adpn peptides 5.1b - 5.4 using CD revealed that the lysine-containing CMP-Adpn peptides 5.1b and 5.2 adopted a triple helical conformation, whereas the serine-containing CMP-Adpn peptides 5.3b and 5.4 did not form triple helices in solution, Figure 5.10. This conclusion can be drawn by the presence of a maximum at 225 nm and a minimum at 195 nm in the CD spectra of lysine-containing CMP-Adpn peptides 5.1b and 5.2 establishing formation of a triple helical secondary structure for these peptides. In contrast, the absence of a maximum at 225 nm but the presence of a minimum at 195 nm observed in the CD spectra of serine-containing CMP-Adpn peptides 5.3b and 5.4 indicates adoption of a random coil secondary structure for these peptides.

![CD spectra of the CMP-Adpn peptides 5.1b – 5.4 recorded at 6 °C at pH 7.4 using potassium phosphate buffer.](image)

The striking observation that only the lysine-containing CMP-Adpn peptides 5.1b and 5.2 were able to adopt a triple helical structure in solution highlights the key role of the lysine (Lys) residues present within the collagenous domain of Adpn for the formation and stability of the protein. This result is in agreement with the aforementioned biological studies by Cooper et al.,28 and the recently reported chemical synthesis of the collagenous domain of Adpn bearing the PTM-lysine residues by Ohkubo et al.49, discussed in detail below in Section 5.4.1.
Chapter Five

5.4 Thermal stability of the CMP-Adpn peptides 5.1b and 5.2

Thermal melt studies were conducted on lysine-containing CMP-Adpn peptides 5.1b and 5.2, which were shown to adopt a triple helix structure. The glycosylated CMP-Adpn peptide 5.1b was found to have a significantly lower melting point ($T_m$ of 20.4 °C) compared to the unglycosylated CMP-Adpn peptide 5.2 ($T_m$ of 40.9 °C), Figure 5.11.

![Thermal melt curves](image)

**Figure 5.11** Thermal melt curves for a) unglycosylated-CMP-Adp peptide 5.2 and b) glycosylated-CMP-Adpn peptide 5.1 recorded at pH 7.4 using potassium phosphate buffer.

5.4.1 Comparison of thermal stability of CMP-Adpn peptides 5.1b and 5.2 with similar literature glycopeptides

The finding that incorporation of a ‘click’ glycoside into the collagenous domain of Adpn lowered the thermal stability of the CMP-Adpn peptides is in contrast to the literature reports that indicate that glycosylation stabilises the structure of collagenous peptides.$^{228}$

A recently published study on glycosylated lysine residues in Adpn by Ohkubo et al.$^{49}$ is also in contrast to the present study. Ohkubo et al.$^{49}$ compared the thermal stabilities of glycosylated and unglycosylated 89-mer peptides comprising the variable and collagenous domain of Adpn. Their studies showed that the glycosylated peptide 5.5 was more thermally stable with a $T_m$ of 33.4 °C compared to the unglycosylated peptide analogue 5.6 which had a $T_m$ of 12.1 °C, Figure 5.12.
Figure 5.12 The chemical structures of the two triple helical Family 3 CMP-Adpn peptides presented in this study neoglycosylated peptide 5.1b and control unglycosylated peptide 5.2 compared to the two reported Adpn peptides by Ohkubo et al. 49 glycosylated peptide 5.5 and unglycosylated peptide 5.6.

As Figure 5.12 illustrates, there are two significant differences between the work published by Ohkubo et al. 49 and the CMP-Adpn neoglycopeptides 5.1b and 5.2 presented in this thesis, which may account for the differences observed in thermal stability upon glycosylation of the peptides, at least in part.
Firstly, the 89-mers synthesised by Ohkubo et al.\textsuperscript{49} correspond to residues 19-107 of Adpn, encompassing the variable and collagenous domains of the protein. In contrast the 36-mer CMP-Adpn peptides \textit{5.1b} and \textit{5.2} contain a middle 18-mer of residues 63 – 80 of Adpn encased by three (GPO) triples on either side. These three (GPO) triplets encasing the Adpn 18-mer, create a CMP ‘scaffold’ stabilising the triple helical secondary structure of the CMP-Adpn peptides. Furthermore, CMPs have been shown to adopt the tighter 7/2 triple helix due to their higher imino acid content,\textsuperscript{232} in contrast to collagenous peptides which favour the looser 10/3 conformation.\textsuperscript{237}

Therefore the 89-mer Adpn peptides \textit{5.16} and \textit{5.17} reported by Ohkubo et al.\textsuperscript{49} would be expected to adopt a 10/3 triple helical confirmation; by contrast the combined CMP-Adpn peptides \textit{5.1b} and \textit{5.2} could be expected to adopt an average conformation closer to the 7/2 triple helix favoured by CMPs (for further discussion of the expected structure of CMP-Adpn peptides see Section 5.13). These expected differences in peptide conformation between the 89-mer Adpn peptides \textit{5.5} and \textit{5.6} reported by Ohkubo et al.\textsuperscript{49} and the 36-mer CMP-Adpn peptides \textit{5.1b} and \textit{5.2} presented in this study, could lead to differences in glycan-glycan interstrand interactions as well as glycan-peptide backbone interactions, which in turn, could affect the thermal stability of the peptides.

Secondly, Ohkubo et al.\textsuperscript{49} incorporated native Glu-Gal-Hyl residues \textit{1.8} into their glycosylated 89-mer Adpn peptide \textit{5.5}, whereas the neoglycosylated CMP-Adpn peptide \textit{5.1b} presented in the present study incorporated \textit{neoGal-Lys 1.10} into the peptide sequence, Figure 5.13.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{Figure5.13.png}
\caption{The chemical structures of the native Adpn glycoside Glu-Gal-Hyl \textit{1.8} and the \textit{neo}glycoside analogue \textit{neoGal-Lys 1.10}.}
\end{figure}

The clear structural differences between the native Glu-Gal-Hyl residues \textit{1.8} and the \textit{neo}glycoside analogue \textit{neoGal-Lys 1.10} could also account for the differences observed in relative thermal stabilities between the glycosylated Adpn 89-mer \textit{5.5} incorporating native Glu-Gal-Hyl residues reported by Ohkubo et al.\textsuperscript{49} and the \textit{neo}glycopeptide CMP-Adpn peptide \textit{5.1b} containing the \textit{neo}glycoside analogue \textit{neoGal-Lys}.  

203
Interestingly, Wennemers et al.\textsuperscript{143} also reported a decrease in thermal stability upon incorporation of a neo\textit{glycoside} into another CMP sequence, peptide 5.7, compared to the unglycosylated peptide 5.8. Their studies focused on glycosylation of hydroxyproline residues in a CMP sequence, with the finding that the T\textsubscript{m} of glycosylated peptide 5.7 was found to be 37 °C, significantly lower than the unglycosylated analogue 5.8 which had a T\textsubscript{m} of 43 °C, Figure 5.14.

\begin{align*}
\text{Ac-}(\text{POG})_3\text{-PO}^\circ\text{-}(\text{POG})_3\text{-NH}_2 & \quad 5.7 \quad \text{T}_\text{m} = 37^\circ \text{C} \\
\text{Ac-}(\text{POG})_3\text{-POG-}(\text{POG})_3\text{-NH}_2 & \quad 5.8 \quad \text{T}_\text{m} = 43^\circ \text{C}
\end{align*}

\textbf{Figure 5.14} The chemical structures of CMPs 5.7 and 5.8 reported by Wennemers et al.\textsuperscript{143}

The significant reduction in thermal stability upon incorporation of glycans into the Adpn sequence in CMP-Adpn neo\textit{glycopeptide} 5.1b, and the striking difference of this property in relation to the native Adpn glycopeptide 5.5 reported by Ohkubo et al.\textsuperscript{49} provides a significant opportunity for the use of neo\textit{glycosides} in collagen-type materials to manipulate the thermal properties of these materials. Raines et al.\textsuperscript{210,211} have already noted the use of synthetic collagen-based materials for wound healing and nanowire applications. Shoulders et al.\textsuperscript{206} also noted the possible uses of CMP materials for collagen imaging and in medical applications as platelet substitutes. The findings discussed in this thesis indicate that incorporation of neo\textit{glycosides} into such materials could be used to fine-tune the thermal stabilities of biomaterials.

\textbf{5.5 Conclusions drawn from the circular dichroism studies of the CMP-Adpn peptides}

In summary, the CD studies of Family 3 CMP-Adpn peptides 5.1b – 5.4 highlighted the crucial structural role of charged lysine residues located in the collagenous domain of Adpn for formation of a triple helical secondary structure of proteins as only the lysine-containing peptides 5.1b and 5.2 were able to adopt a triple helical conformation in this family of peptide.

Furthermore, although neo\textit{glycosylated} lysine-containing CMP-Adpn peptide 5.1b was able to form a triple helical structure, the thermal stability of the neo\textit{glycosylated} peptide was found to be significantly lower than the unglycosylated peptide 5.2.

A reasonable explanation for these observed differences in secondary structure of the Family 3 CMP-Adpn peptides 5.1b – 5.4 could be the different electrostatic interactions present in the peptides, in particular the formation of stabilising ion-pair interactions, as well as the extent of hydrogen bonding
networks present in both the monomeric single stranded conformation and triple helical conformation of the peptides. These two possible explanations will discussed in detail in the following Sections.

5.6 Role of electrostatic interactions for the formation and stability of the triple helical conformation of collagens

The key role of the lysine residues for the observed triple helical structure of the lysine-containing CMP-Adpn peptides 5.1b and 5.2 is likely to be due to a combination of both electrostatic and hydrogen bonding interactions taking place both in an intrastrand and interstrand manner at physiological pH 7.4. In the case of unglycosylated CMP-Adpn-Lys-Control peptide 5.2, these interactions will centre on the charged ammonium group NH$_3^+$ of the Lys residues. In the case of glycosylated CMP-Adpn-Lys-3Gal-OH peptide 5.1b, where the amine group of the Lys residue is replaced by the triazole linker to the glycan, these interactions will centre instead around the free hydroxyl groups of the glycans.

Many groups including Brodsky et al.\textsuperscript{204}, Hartgerink et al.\textsuperscript{238}, Nussinov et al.\textsuperscript{239} and Stultz et al.\textsuperscript{240} have reported the importance of electrostatic interactions between positively and negatively charged residues for the structural stability of collagens. Brodsky et al.\textsuperscript{204} also note that ionisable residues (Lys, Arg, Glu and Asp) account for 15-20% of total residues in collagenous peptides, indicating a key role for such electrostatic interactions in the formation and/or stabilisation of the triple helical structure and subsequent fibril formation of collagenous peptides.

5.6.1 Observations of the relative positioning of acidic and basic resides in the peptide sequence

In addition, Traub et al.\textsuperscript{241} reported in 1975 the frequent placement of basic residues, such as Lys, within one or two residues of acidic residues, such as Glu, thus inferring the possible role of charged ion pairs between Lys and Glu playing a key role in the structure of collagens. Furthermore, Traub et al.\textsuperscript{241} also demonstrated using CPK space-filling models that electrostatic interactions between charged acidic and basic residues in adjacent chains with a one-residue stagger, as in the case of the collagen triple helix, is sterically possible.

5.6.2 Synthesis of ‘host-guest’ CMPs to investigate the role of ion-pairs on the thermal stability of such peptides

The formation of ion-pairs between positively charged lysine (Lys, K) and negatively charged glutamic acid (Glu, E) residues that accounted for the thermal stability of ‘host-guest’ CMPs was demonstrated by Brodsky et al.\textsuperscript{204} in 1994. In their study, Brodsky et al.\textsuperscript{204} investigated the secondary structures of four ‘host-guest’ CMPs; control CMP 5.9 and three ‘host-guest’ CMPs based on an 18-mer of Type III
(T3) collagen; EK-containing peptide 5.10, T3-487 peptide 5.11 and the acetylated variant of T3-487 5.12, Figure 5.15.

\[
\begin{align*}
\text{Control peptide} & \quad 5.9 \\
\text{T3-487 peptide} & \quad 5.10 \\
\text{EK-containing peptide} & \quad 5.11 \\
\text{Ac-T3-487 peptide} & \quad 5.12
\end{align*}
\]

**Figure 5.15** Control CMP 5.9 and ‘host-guest’ CMPs 5.10, 5.11 and 5.12, as reported by Brodsky et al.204

Using CD studies similar to those reported in this thesis, Brodsky et al.204 found that all four peptides 5.9 – 5.12 were able to form stable triple helices in aqueous solution at 10 °C, with the thermal stabilities of the peptides increasing with increased imino acid (proline/ hydroxyproline) content, with the highest \(T_m\) reported for control peptide 5.9 which has an imino acid content of 66%, and the lowest \(T_m\) reported for T3-487 peptide 5.10 with a lower imino acid content of 41%.

Molecular modelling of the short EK-containing peptide 5.11, based on published X ray diffraction data of kangaroo tail bone reported by Fraser et al.,242 revealed the presence of strong ion pairs between the Lys and Glu residues of adjacent chains, Figure 5.16a). Furthermore, molecular modelling of the longer T3-487 peptide 5.10 showed both intrachain and interchain ion pairs to be contributing to the stabilisation of the triple helix, Figure 5.16b).

**Figure 5.16** a) Illustration of interchain ion pairs between oppositely charged Lys and Glu residues of EK-containing peptide 5.11 at neutral pH and b) stabilisation of the triple helix of T3-487 peptide 5.10 by intrastrand electrostatic interactions, figures taken from the publication by Brodsky et al.204
5.7 Effect of pH on thermal stability of collagens and CMPs

The relationship between triple helical stability and pH of the buffer solution provides further evidence for the key role of charged residues such as Lys and Glu for stabilisation of the collagen triple helix.\textsuperscript{204} Dick \textit{et al.}\textsuperscript{243} first published the effect of pH on the stability of collagen extracted from calf-skin in 1966. They found that the collagen had the highest thermal stability at pH 7, where $T_m$ was measured to be 39 \degree C, some seven degrees higher than that recorded in the acidic environments with pH = 0.95 - 2.50. Furthermore, Dick \textit{et al.}\textsuperscript{243} reported a sigmoidal curved relationship between melting temperature and pH with the inflection point occurring around pH 3.5. This sigmoidal curve indicates a gradual increase in stability, which could be explained by increasing ionisation of charged residues such as Lys and Glu and subsequent increase in stabilising electrostatic and hydrogen bonding interactions. Additionally, the inflection point at pH 3.5 correlated with the known pKa of the carboxylic acid group of Glu, providing further evidence for the key role played by charged residues in the stabilisation of collagen.

In their study of CMPs peptides \textit{5.9} – \textit{5.12} previously mentioned in Section 5.4.1.2, Brodsky \textit{et al.}\textsuperscript{204} also investigated the relationship between pH and the thermal stability of the peptides. For the control peptide \textit{5.9}, where the only ionisable groups are the NH$_2$ N-terminus (pKa 9) and the CO$_2$H C-terminus (pKa 2 - 3), the peptide was found to be most stable at the tested pH extremes of pH 1 and pH 11, with a $T_m$ of c.61\degree C recorded at these extreme pHs. The lowest thermal stability of control peptide \textit{5.9} was reported to occur at pH 6, where both the N- and C-termini would be in their ionised forms. This can be explained by charge repulsions occurring between both ionised N- and C-termini at pH 6, whereas at the more extreme pHs either the N-terminus is ionised (in the case of low pH) or the C-terminus is ionised (in the case of high pH), Figure 5.17.

\begin{table}[h]
\begin{tabular}{ccc}
Control peptide & pH 1 & pH 6 & pH 11 \\
\textit{5.9} & *NH$_2$(POG)$_{10}$OH & *NH$_2$(POG)$_{10}$-C$^-$ & NH$_2$(POG)$_{10}$-O$^-$ \\
& $T_m = 61\degree$C & $T_m = 56\degree$C & $T_m = 62\degree$C
\end{tabular}
\caption{Illustration of different ionisation states of control peptide \textit{5.9} at pH 1, 7 and 11 as reported by Brodsky \textit{et al.}\textsuperscript{204}}
\end{table}

In contrast, ‘host-guest’ acetyl-capped CMP \textit{5.12} (Ac-T3-487) was found to be most stable in the pH range 4 - 9 and the thermal stability of non-acetyl capped CMPs \textit{5.10} and \textit{5.11} were observed to increase in a linear fashion with pH up to pH 9. These structure-based pH dependencies provide a better understanding of the role of ionisable residues in collagenous peptides. For all three collagenous type peptides; T3-487 peptide \textit{5.10}, EK-containing peptide \textit{5.11}, and Ac-T3-487 \textit{5.12}, the thermal stability of the triple helix was shown to be influenced by electrostatic interactions between both the N- and
C-termini as well as between the positively-charged Lys residues and the negatively-charged Glu residues.

In the case of the non-acetyl capped CMPs T3-487 peptide 5.10 and EK-containing peptide 5.11, the thermal stability of the peptides is a balance between destabilising repulsive interactions between ionised N- and C-termini of adjacent chains and stabilising salt bridges formed between oppositely charged Lys and Glu residues. At pH > 3, the C-terminus of the peptide is ionised to CO$_2^-$ thus leading to a destabilising repulsive interaction between peptide strands. Above pH 3.5, the Glu residues will be deprotonated and able to form salt bridges with the positively charged Lys residues, Figure 5.18.

The linear increase in thermal stability of peptides 5.10 and 5.11 from pH 3.5 - 9 observed by Brodsky et al.\textsuperscript{204} demonstrates that the stabilising salt-bridge interactions between charged Lys and Glu residues outweigh the destabilising repulsive interactions between the charged C-termini of adjacent chains.

The acetyl-capped CMP 5.12 Ac-T3-487 was found to exhibit maximum stability over a large pH range of 4 – 9 with $T_m = 27^\circ$C recorded in this pH range. This observation leads to the conclusion that partial deprotonation of the N-terminus of non-acetylated peptides 5.10 and 5.11 must occur to some degree above pH 4. This would account for the lower thermal stability of the non-acteylated T3-487 peptide 5.10 compared to its acetyl-capped derivative peptide 5.12 observed between pH 4 - 9. These observations taken together provide robust evidence for the key role of electrostatic interactions between Lys and Glu residues playing a key role in the stability of CMPs.

Furthermore, Chaikov et al.\textsuperscript{244} have reported the role of Arg(+)–Glu(−) ion pairs for the formation of fibrils of the CMP H-(Pro-Arg-Gly)$_4$-(Pro-Hyp-Gly)$_4$-(Glu-Hyp-Gly)$_4$-OH 5.13. By conducting
transmission electron microscopy (TEM) experiments over a range of pH values, Chaikov et al.\textsuperscript{244} demonstrated that the positive domain -(Pro-Arg-Gly)\textsubscript{4} of CMP \textbf{5.13} was able to form zwitterionic ion pairs with the negatively charged domain -(Glu-Hyp-Gly)\textsubscript{4} of the peptide and thus form fibrils, as shown in Figure 5.19 below.

\textbf{Figure 5.19} Zwitterionic ion pairs between positively charged Arg (R) residues and negatively charged Glu (E) residues demonstrated to lead to fibril formation of CMP \textbf{5.13}, figures taken from the publication by Chaikov et al.\textsuperscript{244}

\textbf{5.8 Crystal structures of CMPs illustrating adjacent chain ion-pairs}

In recent years, X-ray crystal structures of CMPs such as Ac-(POG)\textsubscript{3}PKGEOG(POG)\textsubscript{3}-NH\textsubscript{2} \textbf{5.14} have allowed researchers to visualise the ion pairs responsible for electrostatic stabilisation of CMPs.\textsuperscript{238,245} Hartgerink et al.\textsuperscript{238} reported the X-ray crystal structure of CMP \textbf{5.14}, logged in the RCSB Protein Data Bank (http://www.rcsb.org/pdb) with ascension number 3T4F. CMP \textbf{5.14} is N-terminally acetylated and C-terminally amidated in order to increase thermal stability of the peptide. Their structure illustrates the interstrand ion pairs of charged Lys and Glu residues, Figure 5.20.

\textbf{Figure 5.20} Interstrand salt bridges between Lys and Glu residues on adjacent chains of CMP leading to increased structural stability of CMP \textbf{5.14}, figure taken from publication by Hartgerink et al.\textsuperscript{238}

Through the use of NMR spectroscopy, Hartgerink et al.\textsuperscript{246} have also shown that electrostatic interaction play a crucial role for the triple helical formation of CMP Ac-(Pro-Lys-Gly)\textsubscript{10}-NH\textsubscript{2} \textbf{5.15}, Figure 5.21. Through solution phase NMR experiments including 2D NOESY experiments, Hartgerink et al.\textsuperscript{246} demonstrated the presence of ion-pair interactions between lysine (colour-coded red) and aspartic acid
(colour-coded blue) residues, Figure 5.21. The NMR structure for CMP 5.15 is filed in the RCSB Protein Data Bank under ascension number 2KLW.

Figure 5.21 a) The 2D NOESY (Nuclear Overhauser Spectroscopy) solution state spectrum and colour-coded corresponding model of CMP 5.15 illustrating coupling resonances from the Gly N-H to adjacent aC-H protons and b) the presented structural model for CMP 5.15 demonstrating ion-pair electrostatic interactions between red lysine residues and blue aspartic acid residues. Figures taken from the publication by Hartgerink et al.

The crystal structure of aforementioned EK-containing peptide 5.11 was reported by Berman et al. in 2000 and deposited into the RCSB Protein Data Bank under the accession number 1QSU. Interestingly, the Lys and Glu groups within EK-containing peptide 5.11 H-(POG)₄EKG(POG)₂-OH demonstrated extended hydrogen bonding interactions with other residues in preference to formation of an ion pair. The cationic Lys residues were shown to form hydrogen bonds with polar carbonyl groups of Gly residues and the anionic Glu residues displayed hydrogen bonding interactions with the free hydroxyl groups of hydroxyproline, Figure 5.22.

Figure 5.22 Ion-pair formation observed in the crystal structure of EK-containing peptide 5.11. a) Ion-pair formation between cationic Lys and carbonyl groups of Gly residues and b) ion-pair formation between anionic Glu residues and the free hydroxyl groups of hydroxyproline. Figures taken from the publication by Berman et al.
5.9 The role of hydrogen bonding in the structural stability of collagenous peptides

It is also likely that the lysine and glycosylated lysine residues of CMP-Adpn peptides 5.1b and 5.2 have more extended hydrogen bonding networks than the serine and glycosylated serine residues of CMP-Adpn peptide 5.3b and 5.4. These extended hydrogen bonding networks may also be mediated by bridging water molecules, see Section 5.8.205

Creamer et al.247 have reported the ability of glutamine residues to build a strong intramolecular hydrogen bonding network with carbonyl residues in the peptide backbone, Figure 5.23a). Notably, the ability of the glutamine (Gln) residues to form intrastrand hydrogen bonds was attributed to the long arm length of the Gln side chain.247 As the Lys side chain has a length of 5.63 Å, similar to the Gln side chain length of 4.66 Å, one could expect Lys residues to behave in a similar manner to Gln and also contribute to intra-strand hydrogen bonding, Figure 5.23b).

In contrast, the short side-chain length of Ser of 2.45 Å may not be able to adopt the correct conformation to form similar intra-strand hydrogen bonds, Figure 5.23c).

![Figure 5.23](image)

**Figure 5.23** a) Representative illustration of possible intrastrand hydrogen bonding and through-space side arm length of a) glutamine NH$_3^+$ to peptide backbone C=O$_{247}$; b) similar length lysine side arm able to participate in similar hydrogen bonding interactions c) serine side arm unable to participate in hydrogen bonding.

<table>
<thead>
<tr>
<th>Amino acid residue</th>
<th>Through bond length</th>
<th>Through space length</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Cα-end atom (Å)</td>
<td>Cα-end atom (Å)</td>
</tr>
<tr>
<td>Glutamine</td>
<td>5.88</td>
<td>4.66</td>
</tr>
<tr>
<td>Lysine</td>
<td>7.29</td>
<td>5.63</td>
</tr>
<tr>
<td>Serine</td>
<td>2.94</td>
<td>2.45</td>
</tr>
</tbody>
</table>

Table 5.1 Side arm bond lengths calculated for glutamine, lysine and serine residues using the MenD tuberculosis crystal structure provided by Dr Jodie Johnson, School of Biological Sciences, University of Auckland.
The ability of Lys residues to form intra-strand hydrogen bonds with carbonyl group oxygen atoms was also demonstrated by Hartgerink et al.\textsuperscript{248} in their reported synthetic CMP H-(Pro-Lys-Gly)\textsubscript{4}(Pro-Hyp-Gly)\textsubscript{4}(Asp-Hyp-Gly)\textsubscript{4}-OH \textbf{5.16}, which was shown to form very stable triple helices (T\textsubscript{m} of 41 °C), which were able to self-assemble to form fibrils of \(c.1.2\text{nm}\) in length at pH 7, Figure 5.24a). Using NMR spectroscopy, Hartgerink \textit{et al.}\textsuperscript{246,248} were able to demonstrate that Lys residues present in their CMP \textbf{5.16} stabilise the triple helical structure of the peptide and the resulting fibril by both Lys-Asp ion pairs as well as hydrogen bonding to the carbonyl oxygen group on the same peptide chain, Figure 5.24b). Therefore, it can be assumed that unglycosylated Lys residues present in CMP-Adpn peptide \textbf{5.2} could be stabilising the triple helical structure of the peptide in a similar manner.

\textbf{Figure 5.24} a) AFM image of self-assembled fibrils of CMP \textbf{5.16} in phosphate buffer at pH 7 and concentration of 0.5% by weight as reported by Hartgerink \textit{et al.}\textsuperscript{248}; b) model of the proposed electrostatic ion pair interactions and intrastrand hydrogen bonding network facilitated by Lys residues present in CMP \textbf{5.16} figures taken from the publication by Hartgerink \textit{et al.}\textsuperscript{246,248}

\textbf{5.10 The use of water molecules to extend the hydrogen bonding network and for electrostatic ion pair formation}

The X-ray crystal structure of EK-containing peptide \textbf{5.11}\textsuperscript{245} revealed significant additional hydrogen bonding interaction through the use of bridging water molecules, as discussed previously in Chapter 4. Furthermore, Berman \textit{et al.}\textsuperscript{249} examined the hydration structure of CMP \textbf{5.17} H-(POG)\textsubscript{4}POA(POG)\textsubscript{2}-OH revealing, in addition to the known hydrogen bonding between Gly-NH and Pro-C=O residues,\textsuperscript{203} interstrand water bridges between the hydroxyproline-OH and adjacent chain C=O groups and intrastrand water bridges between carbonyl groups. The X-ray crystal structure of CMP T3-785 \textbf{5.18} Ac-(Pro-Hyp-Gly)\textsubscript{3}-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly-(Pro-Hyp-Gly)\textsubscript{3}-NH\textsubscript{2} published by Berman \textit{et al.}\textsuperscript{250} also provided evidence for the role of bridging water molecules connecting Ala-NH with Gly-C=O groups, thus extending the hydrogen bonding network of the peptide.

Furthermore, in order for charged residues to form a direct ion pair, known as a contact ion pair (CIP),\textsuperscript{251} thereby to interact with each other, they must be desolvated and be sufficiently close in space for the
ion-pair to form.\textsuperscript{252} If the energy required for pre-ordering of the protein and desolvation of the charged residues is more than the Columbic energy gained through formation of the ion-pair, the electrostatic interaction may not be favourable. However, if the solvent is involved as a bridging molecule for the ion pair, the ion pair is known as a solvent shared ion pair (SIP)\textsuperscript{251}. Indeed, Iwahara \textit{et al.}\textsuperscript{251} have proposed a dynamic equilibrium between CIP and SIP states for ion-pairs in solution. Furthermore, as charged residues attract a solvation sphere of water molecules, release of these dipole-induced water molecules upon ion-pair formation will lead an entropic gain for the system.

In summary polar solvent molecules, such as water molecules, may assist the formation of both hydrogen bonding networks and electrostatic interactions of proteins. Through the use of bridging solvent molecules, extended hydrogen bonding networks can be formed. Solvent molecules can also act as mediators for ions pairs through the formation of SIPs and the release of solvent spheres surrounding charged residues upon ion-pair formation. Therefore, bridging solvent molecules should also be taken into account when considering the structural stability of triple helical peptides CMP-Adpn 5.1b and 5.2.

5.11 Proposed molecular structure of the non-glycosylated, control CMP peptides 5.2 and 5.4

In light of the discussion of the likely contributions of charged Lys and Glu residues at physiological pH, the following diagram has been constructed to illustrate the possible hydrogen bonding and electrostatic interactions taking place for CMP-Adpn-Lys-Control peptide 5.2, Figure 5.25. This figure is based on the insights from literature as discussed above and most particularly from the graphic recently published by Bella.\textsuperscript{212}
5.12 Proposed molecular structure of glycosylated CMP-Adpn peptides 5.1b and 5.3b

Although the model illustrated in Figure 5.25 of increased electrostatic interactions and hydrogen bonding through Lys/Glu ion pairs can be used to explain the ability of the lysine-containing control CMP-Adpn peptide 5.2 over the serine-containing control CMP-Adpn peptide 5.4, this explanation cannot apply to the equivalent glycosylated CMP-Adpn peptides 5.1b and 5.3b, as the ‘click’ glycans are incorporated via the lysine NH$_2$ group. Therefore, the charged NH$_3^+$ groups of the Lys residues are not present and thus not available for electrostatic or hydrogen bonding interactions in the glycosylated CMP-Adpn peptides 5.1b and 5.3b.

However, as discussed in Chapter 4, glycans have been shown to extend hydrogen bonding networks in many collagen systems and thereby facilitate and support formation of a triple helical conformation in these systems. The free hydroxyl groups present on glycans in native glycosylated systems have
been shown to play significant structural roles both intrastrand and interstrand in collagen systems. Moreover, Cooper et al.\textsuperscript{28}, Myllylä et al.\textsuperscript{208} and Ohkubo et al.\textsuperscript{49} have demonstrated the crucial role of glycosylation of four key Lys residues located in the collagenous domain of adiponectin, in facilitating the self-assembly and bioactivity of Adpn. Thus investigation of the incorporation of glycans into Adpn sequences such as CMP-Adpn peptides 5.1b and 5.3b is key to further the understanding of the molecular forces governing the structure and stability of Adpn.

As discussed in Section 4.3.1, Wetmore et al.\textsuperscript{205} reported that glycosylation of the CMP Ac-(Hyp)\textsubscript{9}-NH\textsubscript{2} 4.12 led to increased levels of intrastrand hydrogen bonding within the triple helix. Using molecular mechanics-based minimisations of solvent-mediated hydrogen bonding with CMP 4.12, Wetmore et al.\textsuperscript{205} were able to identify both glycan-glycan interactions and glycan-peptide intrastrand hydrogen bonding between the free hydroxyl groups of the glycan and a carbonyl group on the peptide backbone, Figure 5.26.

![Figure 5.26 Intrastrand hydrogen bonding a) glycan-glycan and b) glycan 6-OH to peptide backbone C=O as described by Wetmore et al.\textsuperscript{205}](image)

Interstrand hydrogen bonding possibly utilising mediating water molecules may also be possible for glycosylated peptides in solution. Schweizer et al.\textsuperscript{205} have demonstrated using molecular dynamics that glycosylation of hydroxyproline (Hyp) residues leads to an extended interstrand hydrogen bonding network through the use of bridging water molecules. Bächinger et al.\textsuperscript{228} have investigated the role of glycosylated threonine (Thr) in CMPs and showed that Thr glycosylation is necessary for the formation of the triple helix secondary structures for these peptides. Therefore, it is logical to deduce that the glycans incorporated into CMP-Adpn peptides 5.1b and 5.3b may stabilise the triple helix secondary structure of the peptide by extended hydrogen bonding networks, possibly mediated by bridging water molecules.

The observed structural difference between the two click’ glycosylated CMP-Adpn peptides lysine-containing peptide 5.1b ‘and serine-containing peptide 5.3b could be due to differences in side-arm length and thus the distance of the glycan from the peptide backbone. It is possible that the lysine-containing peptide 5.3b with the long side-arm chain length of 5.63 Å is able to adopt a
conformation which allows for multiple intrastrand and interstrand hydrogen bonding interactions to take place, thus stabilising the triple helix, Figure 5.27a). On the other hand, serine-containing peptide 5.1b with the shorter serine side arm length of 2.45 Å, may not be able to adopt the correct conformation for these hydrogen bonds to be formed, thus destabilising the structure and resulting in adoption of a random coil structure in preference to a triple helix, Figure 5.27b).

a) 

b) 

Figure 5.27 Possible intrastrand and interstrand hydrogen bonding interactions for a) lysine-containing peptide 5.1b and b) serine-containing peptide 5.3b. Hydrogen bonds are shown in light blue and for simplicity water bridging interactions have been omitted.

5.13 Glycosylated peptides in the context of diabetes

Although some level of glycosylation of the four Lys residues found in the collagenous domain of Adpn may be necessary for bioactivity of the protein it is possible that there is an optimal level for this glycosylation, which may be different for each of the four Lys residues. Unpublished work by Cooper et al. demonstrates that native mammalian Adpn exists as a number of different glycosidic isoforms at each Lys residue within the collagenous domain. It may be the case that the four different Lys residues play different roles in the structural stability of the triple helical protein, or have signalling roles attuned to the bioactivity of Adpn.

The work published by Ohkubo et al. wherein each of the four Lys residues was fully glycosylated (i.e. the addition of the disaccharide Glu-Gal to the hydroxylysine residue), is therefore a much simpler structure than actually exists in native systems. Thus the results presented by Ohkubo et al. that glycosylated Adpn 5.5 demonstrated significantly higher thermal stability compared the unglycosylated analogue 5.6 may not constitute a true representation of native, mammalian Adpn.

It could be the case that glycosylation of Adpn up to an optimal level leads to formation of bioactive HMW oligomers of Adpn. However, glycosylation beyond this optimal level could lead to breakdown of the triple helical structure of Adpn resulting in reduced bioactivity of the protein, as shown by the model CMP system discussed in Chapter 4. Indeed several studies of diabetic and control patients have reported decreased serum levels of HMW oligomers of Adpn for diabetic patients, with the
molecular mechanism for this change not yet reported. One possible explanation for the decreased level of HMW oligomers observed in the diabetic patients could be over glycosylation of the Adpn protein leading to a breakdown in the higher-order oligomer structure in a similar manner to that observed for the doubly- and triply-glycosylated CMPs 4.21b, 4.22b, 4.25b and 4.26b discussed in Chapter 4.

The link between increased serum glucose level and the level of HMW oligomers of Adpn has also been further explored by Xiang et al.\textsuperscript{256} who reported that administration of insulin-sensitising thiazolidinediones (TZDs) to patients with Type II diabetes led to increased serum levels of HMW oligomers of Adpn, as discussed previously in Chapter 1. A molecular level explanation for this observed structural change could be that administration of the TZDs could cause the blood glucose levels to drop, thus decreasing the levels of glycosylation of Adpn to an optimal level thereby increasing the serum level of HMW oligomers.

Furthermore, Spiro et al.\textsuperscript{52,257} reported that in an alloxan-diabetic rat model, the tissue levels of collagen glucosyl transferases were much higher compared to the control age-matched non-diabetic rats. An increase in levels of collagen glucosyl transferases in diabetic animals may well lead to an increase in (non-physiological) glycosylation of tissue collagens. The study reported by Spiro et al.\textsuperscript{52} provides indirect evidence for increased collagen glycosylation in diabetes. Furthermore, analysis of diabetic renal glomeruli also by Spiro et al.\textsuperscript{258} identified higher glycosylated hydroxylysine content present in the diabetic samples compared to tissue from control subjects, once again linking the diabetic state to an increase in collagen glycosylation.

Later studies have shown that lysyl hydroxylase isoform 3 (LH3) is the key enzyme responsible for the post-translational modifications of Adpn.\textsuperscript{61,259} Cooper et al.\textsuperscript{260} have reported elevated levels of LH3 in diabetic ob/ob mice models compared to control subjects. This observation could lead to the conclusion that in diabetic mice increased levels of LH3 could lead to increased glycosylation of collagenous proteins such as Adpn. The decreased thermal stability of glycosylated CMP-Adpn peptide 5.2 could also provide further evidence that glycosylation of Adpn over a certain optimal level could contribute towards the decreased thermal stability of such collagens.

Taken together the observations discussed above including decreased levels of HMW oligomers of Adpn and increased levels of LH3 observed in diabetic patients could be linked with the known collagenous defects observed in patients with Type II diabetes such as ventricular dysfunction,\textsuperscript{217} diabetic cardiomyopathy\textsuperscript{218} and diabetic nephropathy.\textsuperscript{219}
5.14 Conclusion for Chapter Five

In Chapter Five, the synthesis of hybrid Adpn-CMP neo-glycopeptides incorporating encasing the short Adpn 18-mer sequence in between a collagen model peptide (GPO)₃ sequence using Fmoc-SPPS was described. The synthesis of non-glycosylated, control peptide analogues of the hybrid Adpn-CMPs was also detailed.

The secondary structure and thermal stability of the hybrid Adpn-CMPs was investigated using circular dichroism techniques. Both the neo-glycosylated and non-glycosylated control lysine-based hybrid Adpn-CMPs were shown to adopt a triple helical conformation in buffer solution at neutral pH. The thermal stability of the neo-glycosylated lysine-based peptide was notably reduced compared to the non-glycosylated analogue. The serine-based hybrid Adpn-CMPs were shown to be unable to adopt a triple helical conformation under the same conditions.

Comparisons of the structural properties of this third family of synthetic neo-glycopeptides with the previously described Family 1 Adpn neo-glycopeptides and Family 2 CMPs will be conducted in the next Chapter of this thesis.
Chapter Six

Summary and Conclusions
Chapter Six: Summary and Conclusions

In this Chapter, an overall summary of the synthetic work conducted in this thesis will be given. The key structural conclusions made for each of the three families of neo-glycopeptides will also be discussed. In addition, comparisons between the Family 1 Adpn neo-glycopeptides, Family 2 CMP neo-glycopeptides and Family 3 CMP-Adpn neo-glycopeptides will be drawn. Finally, an overall conclusion for this thesis will be presented.

6.1 Summary of doctoral studies reported in this thesis

In summary, Chapter Two of this thesis described attempts to produce a novel, highly yielding synthetic route towards native Gal-Hyl using a Sharpless aminohydroxylation reaction and the Schmidt glycosylation method. This novel synthetic approach was ultimately unsuccessful due to the inability to resolve the epimers afforded in the Sharpless aminohydroxylation reaction.

Chapter Three of this thesis detailed the synthesis of neo-glycosides of native Gal-Hyl utilising Cu(I)-catalysed azide-alkyne cycloaddition reactions and incorporation of the neo-glycosides into short 18-mers corresponding a section of the collagenous domain of Adpn. Both serine- and lysine-based neo-glycosides were prepared in order to investigate the effect of the glycan-peptide backbone through-space distance on the secondary structure of the neo-glycopeptide. The secondary structures of the Adpn neo-glycopeptides was examined using circular dichroism, but unfortunately the 18-mer neo-glycopeptides were found to adopt a random coil secondary structure in preference to the desired triple helical secondary structure common to collagenous peptides.

Therefore, collagen model peptides (CMPS) containing [Gly-Hyp-Pro] repeat triplets were employed to promote triple helical formation of the neo-glycopeptides. In Chapter Four, trial studies of the collagen model peptide sequences incorporating lysine- and serine-based neo-glycosides were discussed. Key findings from these trial studies included the key role of the hydroxyl groups of the glycans for the formation of triple helix secondary structure, as well as the possibility of an ‘optimal level’ of glycosylation for collagenous proteins.

In Chapter Five, the synthesis of hybrid CMP-Adpn neo-glycopeptides incorporating both the short 18-mer Adpn sequence as well as the CMP [Gly-Hyp-Pro] triplets was detailed. The synthesis of control, non-glycosylated Adpn-CMP analogues was also discussed. Both the neo-glycosylated and non-glycosylated lysine-based hybrid Adpn-CMPs were shown to adopt a triple helical secondary structure, with the thermal stability of the glycosylated peptide notably reduced compared to the non-glycosylated analogue. The serine-based Adpn hybrid peptides were shown to be unable to adopt a triple helical secondary structure.
Figure 6.1 The synthetic peptides presented in this thesis.

6.2 Comparison of Family 2 CMP and Family 3 Adpn-CMP peptide families

The two structural similarities observed between the Family 2 CMPs and Family 3 CMP-Adpn peptides were the ability of the lysine-containing peptides to adopt a triple helix secondary structure and the decrease in thermal stability of the neo-glycopeptide upon increased levels of glycosylation.
Chapter Six

The two structural differences observed between the Family 2 CMPs and Family 3 CMP-Adpn peptides were the ability of the serine-containing peptides and the capability of the triply-glycosylated neoglycopeptides to adopt triple helical secondary structures.

6.2.1 Observed structural similarities between the Family 2 CMP and Family 3 Adpn-CMP neoglycopeptide families

The first structural similarity observed between the Family 2 CMPs and Family 3 CMP-Adpn peptide was that in both peptide Families, the lysine-containing peptides were able to adopt the desired triple helical secondary structure to some extent, although the tolerance for glycosylation in the triple helical secondary structure was different in each Family which is discussed in detail below in Section 6.2.2.

In the case of the Family 2 CMPs, the lysine-containing control unglycosylated CMP 4.23 and the mono-glycosylated CMP 4.22b were found to form the desired triple helix in buffer solution. In the Family 3 CMP-Adpn peptides, both the lysine-containing control unglycosylated CMP-Adpn peptide 5.2 as well as the triply-glycosylated CMP-Adpn peptide 5.1b was able to adopt the desired triple helical secondary structure under the same conditions.

This similarity in the lysine-containing control, unglycosylated CMP and CMP-Adpn peptides may be due to the electrostatic contributions of the Lys residues to the triple helical secondary structure, causing the triple helix to be favoured over the monomeric single stranded state. In the case of the neoglycosylated lysine-containing CMP and CMP-Adpn peptides, the similarity between the two peptide Families may be due to similar levels of additional hydrogen bonding due to glycan based interactions stabilising the triple helical secondary structure over the monomeric state.

The second structural similarity observed between the Family 2 CMPs and Family 3 CMP-Adpn peptides was that for both Families, an increased glycosylation decreased the thermal stability of the resulting neoglycopeptide. Furthermore, this reduction in thermal stability of the neoglycopeptide upon extended glycosylation appeared to be cumulative in nature.

The reduction in thermal melting temperature \( (T_m) \) was much greater between the triply-glycosylated Family 3 CMP-Adpn peptide 5.1b with a \( T_m \) of 40.0 °C and the control unglycosylated CMP-Adpn analogue 5.2 with a \( T_m \) of 20.0 °C compared to the Family 2 mono-glycosylated CMP with a \( T_m \) of 28.8 °C and the control unglycosylated analogue 4.2 with a \( T_m \) of 28.8 °C, Table 6.1. This similarity implies that glycosylation thermally destabilises neoglycopeptides. This observation could be used to modify the thermal properties of functional materials incorporating neoglycopeptides.
Table 6.1 Comparison of thermal melting temperatures of the Family 2 CMPs and Family 3 Adpn-CMPs.

<table>
<thead>
<tr>
<th>Control non-glycosylated peptides</th>
<th>( T_m ) (°C)</th>
<th>Glycosylated peptides</th>
<th>( T_m ) (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMP-Lys-Control, 4.1</td>
<td>35.8</td>
<td>CMP-Lys-1Gal-OH 4.2</td>
<td>28.8</td>
</tr>
<tr>
<td>CMP-Ser-Control, 4.3</td>
<td>26.2</td>
<td>CMP-Ser-1Gal-OH, 4.4</td>
<td>25.7</td>
</tr>
<tr>
<td>CMP-Adpn(66-83)-Lys-Control, 5.2</td>
<td>40.0</td>
<td>CMP-Adpn(66-83)-Lys-3Gal-OH, 5.1b</td>
<td>20.0</td>
</tr>
</tbody>
</table>

6.2.2 Observed structural differences between the Family 2 CMP and Family 3 Adpn-CMP neo-glycopeptide families

The first structural difference observed between the Family 2 CMPs and Family 3 CMP-Adpn peptides was that for the Family 2 CMPs, serine-containing peptides CMP-Ser-1Gal-OH 4.24b and CMP-Ser-Control 4.27 were able to adopt a triple helical secondary structure, whereas in the Family 3 CMP-Adpn peptides, both serine-containing peptides CMP-Adpn(66-83)Ser-3Gal-OH 5.3b and CMP-Adpn(66-83)-Ser-Control 5.4 were unable to adopt the triple helical secondary structure common to collagenous peptides.

The second structural difference observed between the two Families was that Family 3 CMP-Adpn peptides were able to tolerate higher levels of glycosylation and retain the desired triple helical secondary structure. The Family 3 lysine-containing triply-glycosylated neo-glycopeptide CMP-Adpn(66-83)-Lys-3Gal-OH 5.1b was able to adopt the desired triple helical secondary structure, whereas neither the serine- nor the lysine-containing triply-glycosylated CMPs of Family 2 were able adopt a triple helical secondary structure.

6.2.3 Proposed difference in triple helical pitch between Family 2 CMPs and Family 3 Adpn-CMPs

A reasonable explanation for these observed structural differences outlined above between the Family 2 CMPs the Family 3 CMP-Adpn peptides could be an expected variance in the triple helical pitch of the two families of peptides. As mentioned previously, Berman et al.\(^{245}\) reported the crystal structure of EK-containing peptide 5.11 [Section 5.6, sequence Ac-(POG)\(_4\)EKG(POG)\(_3\)-NH\(_2\) and deposited into the Protein Data Bank under the accession number 1QSU]. Through this crystal structure, Berman et al.\(^{245}\) showed that EK-containing peptide 5.11 adopted a tight 7/2 triple helical secondary structure (7 residues per 2 turns of the helix), Figure 6.2a.

Berman et al.\(^{232}\) have also demonstrated that T3-785 peptide 5.18 [Section 5.8, sequence Ac-(Pro-Hyp-Gly)\(_3\)-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly-(Pro-Hyp-Gly)\(_3\)-NH\(_2\)] which consists of a ‘guest’ peptide sequence taken from human Type III (T3) collagen encases inside a CMP ‘host’ domain of POG repeats. The crystal structure of T3-785 peptide 5.18 reported Berman et al.\(^{232}\), stored in the Protein Database under ascension number 1BKV, revealed that the imino-rich ‘host’ POG regions had
a tight 7/2 structure, whereas the imino-poor ‘guest’ T3 region displayed a looser 10/3 structure (10 residues per 3 turns of the helix), closer to the natural helical pitch of native collagen,\textsuperscript{206} Figure 6.2b.

As Family 2 CMPs \textbf{4.20} – \textbf{4.27} have a very similar sequence to EK-containing peptide \textbf{5.11}, it is reasonable to deduce that CMPs \textbf{4.20} – \textbf{4.27} would also adopt a 7/2 triple helical secondary structure. On the other hand, as Family 3 CMP-Adpn peptides \textbf{5.1b} – \textbf{5.4} have a similar structure to T3-785 peptide \textbf{5.18}, with a (GPO) ‘host’ structure encasing the central Adpn ‘guest’ domain, it is likely that CMP-Adpn peptides \textbf{5.1b} – \textbf{5.4} adopt a similar mixed 7/2, 10/3 structure.

This variance in helical structure between the two families could account for both of the observed structural differences between Family 2 and Family 3 peptides. The incorporation of sterically bulky glycans is more disfavoured in the Family 2 CMPs \textbf{4.20} – \textbf{4.27} with the tighter 7/2 structure when compared to the Family 3 CMP-Adpn peptides \textbf{5.1b} – \textbf{5.4} with the looser 10/3 structure.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.2.png}
\caption{a) Crystal structure of EK containing peptide \textbf{5.11} showing a tight 7/2 structure as reported by Berman \textit{et al.}\textsuperscript{245}; b) crystal structure of T3-785 peptide \textbf{5.18} showing hybrid 7/2 and 10/3 structure, as reported by Berman \textit{et al.}\textsuperscript{232}}
\end{figure}
The difference in propensity of the serine-containing peptides from the Family 2 CMPs and the Family 3 CMP-Adpn peptides towards adopting a triple helical secondary structure could also be explained by the hybrid structure of the CMP-Adpn peptides. In contrast to the Family 2 serine-containing CMPs 4.24a - 4.27 which are expected to adopt regular 7/2 helical structure throughout the whole peptide chain, Family 3 serine-containing CMP-Adpn peptides 5.3b and 5.4 display a hybrid structure between 7/2 for the (GPO) ‘host’ sections and 10/3 for the central ‘guest’ domain. This change in structure could affect the structure of the hydrogen bonding network for the Adpn-CMPs, causing a random coil secondary structure to be formed in preference to the otherwise preferred triple helical secondary structure.

6.3 Proposed model for triple helical secondary structure of collagenous peptides

Indeed, the comparison of the triple helical properties between the Family 2 CMPs 4.20 - 4.27 and the Family 3 CMP-Adpn peptides 5.1b - 5.4 allows for the proposal of the following model for the interplay of different factors in the formation of triple helical secondary structure of collagenous peptides. This model illustrates the balance between favourable hydrogen bonding and electrostatic interactions strengthening the triple helical state and the unfavourable entropic costs and steric constraints associated with moving from a monomeric state to the trimeric triple helix structure. The end-result of whether a collagenous peptide will or will not adopt a triple helical secondary structure depends of the weighting of each of these factors and the eventual balance of the ‘see-saw’ illustrated in Figure 6.3.

**Figure 6.3** Proposal for interplay of key factors for the formation of triple helical collagenous peptides.

On the right-hand side of the see-saw, the inherent costs of formation of the triple helix are depicted, which includes the steric constraints for triple helix formation, especially when large, bulky moieties such as glycans are incorporated into the peptide sequence, and the entropic costs associated with formation of a rigid triple helical secondary structure as opposed to three separate monomeric chains. Furthermore, there will be a solvent contribution for the entropic cost for triple helix formation, as solvent molecules often act as interstrand mediators for both hydrogen bonding and charged ion pair formation in the triple helical state. The degree of entropic cost will also depend on the degree of pre-ordering in the monomeric state.
On the left-hand side of the seesaw the inherent gains are illustrated, which are enthalpic in nature. Formation of new hydrogen bonds and electrostatic ion pairs will support the triple helical structure. The strength of these favourable hydrogen bonding and electrostatic interaction is dependent on pH and temperature, illustrated in Figure 6.3 by the dependence of fulcrum position on these two variables.

In summary, the formation of a triple helix is a balance between the inherent costs and gains, which can be visualised as a see-saw with the fulcrum position dependent on pH and temperature. As the right-hand side of the see-saw contains an entropic contribution, one would expect an increase in temperature to shift the fulcrum to the right and reduce the stability of the triple helix. Equally, a change in pH to favour the formation of electrostatic ion pairs will shift the fulcrum position favouring the electrostatic gain in the system upon triple helical formation.

### 6.4 Application of the triple helical secondary structure model to CMP and CMP-Adpn peptides

The see-saw model presented above could be applied to the expected structures of Family 2 CMPs and Family 3 CMP-Adpn peptides. As the Family 2 CMPs have a tighter, 7/2 triple helical structure the steric constraint cost will be larger than for the Family 3 CMP-Adpn peptides, Figure 6.4a) and Figure 6.4b). Therefore, the see-saw for Family 2 CMPs is more heavily weighted towards the left-hand side and so incorporation of bulky glycans may tip the see-saw off balance meaning that it is not energetically favourable for the CMP peptide to adopt a triple helical secondary structure.

This model provides a reasonable explanation for the ability of Family 3 peptide CMP-Adpn(63-80)-Lys-3Gal-OH 5.1b to adopt a triple helical secondary structure with three glycans incorporated into the peptide sequence, in contrast to the random coil secondary structure observed for the Family 2 peptides CMP-Lys-3Gal-OH 4.22b and CMP-Ser-3Gal-OH 4.26b.
Figure 6.4 b) Family 3 CMP-Adpn peptides with incorporated glycan residues with the looser 10/3 structure have reduced steric constraints and so are more able to adopt a triple helical secondary structure.

6.5 Relevance of conclusions to other collagen based systems

The human 18-mer Adpn non-modified sequence GEKGEKGDGPLIGPKGD1 was run through the NCBI BLAST database and matched the Swissprot protein database, which is a database of known protein annotations such as protein function or post-translational modifications.

Aside from other Adpn proteins (i.e from different organisms), the golden hamster (Mesocricetus auratus) macrophage receptor MARCO (Swissprot ID Q9WUB9) reported by Tryggvason et al.\textsuperscript{261} showed a significant sequence match for the first 17-AA (amino acids) of the sequence with 14/17 AA being a direct match with a statistically significant E-value of 2e-04, Figure 6.5. Notably, all three lysine residues were matched between the sequences providing further evidence for the crucial role of lysine residues for the structural stability of collagenous proteins. MARCO is a known macrophage receptor which has a triple helical collagenous domain, which is used to distinguish host and foreign bodies \textit{in vitro}\textsuperscript{261} and thus plays a crucial role in immune defence.

Furthermore, structural similarities were also observed with other collagenous peptides listed in the Swissprot database including collagen-like protein 7 found in mimivirus reported by Claverie \textit{et al.}\textsuperscript{262} (Swissprot ID Q5UNS9.1) with 14/18 AA and a statistically significant E-value of 6e-04. Indeed, Claverie \textit{et al.}\textsuperscript{262} note that glycosylation of lysine residues present in collagen-like protein 7 may be involved in the formation of cross-linked fibrils leading to the ‘hairy’ appearance of the mimivirus surface reported by Raoult \textit{et al.}\textsuperscript{262} in 2003, Figure 6.6.
Similarities in peptide sequence were also observed with other collagenous proteins, Table 6.2. Therefore, due to a number of significantly similar matches of the 18-mer Adpn sequence investigated in this thesis to other collagenous proteins, it is reasonable to suggest that the structural role of both non-glycosylated and glycosylated lysine residues on the structural stability of collagenous proteins shown in this thesis may apply to other collagenous proteins most notably MARCO and collagen-like protein 7.

Table 6.2 Closest amino acid matches observed through NCBI Blast database of Swissprot proteins matched to 18-mer Adpn sequence (run April 2016) with differences in AA sequence highlighted in bold. --Indicates gap in the protein sequence relating to a non-matched protein segment linking the two matched sequences. E-value describes the expectation value or the probability of a matched sequence.

<table>
<thead>
<tr>
<th>Protein Name</th>
<th>AA sequence</th>
<th>AA match</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adpn 18-mer</td>
<td>GEKGEKGDPGLIGPKGD</td>
<td>n/a</td>
<td>n/a</td>
</tr>
<tr>
<td>MARCO</td>
<td>GEKGSKGDBGGLIPGKD</td>
<td>14/17</td>
<td>2e-04</td>
</tr>
<tr>
<td>Collagen-like protein 7, mimivirus</td>
<td>GDKGKEKGDAGLKGNKGDI</td>
<td>14/18</td>
<td>6e-04</td>
</tr>
<tr>
<td>Collagen alpha-2(IV) chain</td>
<td>GKGERGAPGTGPKGDV</td>
<td>12/18</td>
<td>0.003</td>
</tr>
<tr>
<td>Collagen alpha-3(V) chain</td>
<td>GEKERGDPGPKGD</td>
<td>15/18</td>
<td>0.003</td>
</tr>
<tr>
<td>Collagen alpha-1(IV) chain</td>
<td>GERGKEKGERGLI--GP</td>
<td>14/17</td>
<td>0.005</td>
</tr>
</tbody>
</table>

6.6 Concluding remarks

The neo-glycosides discussed in this thesis have investigated the role of glycosylated (Gal-Hyl) hydroxylysine residues for the structural stability of the collagenous domain of adiponectin (Adpn), a potential therapeutic for Type II diabetes.

Through the use of CMP ‘host-guest’ peptides, the effect of chain length and residue charge (Lys versus Ser) and mono, di and tri-glycan incorporation have been investigated. The use of these synthetic model peptides are important to probe the structural properties of collagenous peptides, as in vivo samples of such tissue are heterogeneous in nature and not appropriate for the systematic study of structure-activity effects. As discussed in Chapter 3, the short 18-mer neo-glycosidic Adpn peptides 3.58b and 3.59b were
unable to adopt a triple helical secondary structure at physiological pH. This may be due to electrostatic repulsions of uncapped, charged end-termini and/or the short, non-uniform nature of the 18-mer Adpn sequence.

Therefore, the synthetic strategy was altered to focus on acetyl-capped ‘host-guest’ collagen model peptide systems. The model CMP family 4.20 – 4.27 with a central (GEX)_n (X = ‘pre-clicked’ glycan) ‘guest domain’ and (POG)3 ‘host’ triplets was successfully synthesised using Fmoc-SPPS. Interestingly of the Family 2 CMPs, only the control unglycosylated CMPs 4.23, 4.27 and the mono-glycosylated CMPs 4.21b and 4.24b were able to adopt a triple helical secondary structure. These observations led to the hypothesis that there may be an optimal level of glycosylation for collagenous proteins.

As discussed in Chapter 4, the Family 2 CMPs also provided insight into the different structural roles serine and lysine residues play in formation and stabilisation of the collagen triple helix. Lysine-containing CMPs 4.20b and 4.23 were found to have greater thermal stability than the equivalent serine-containing CMPs 4.24b and 4.27. This result indicates that lysine residues contribute more structural stability the triple helix than Ser residues, possibly due to the formation of charged ion-pairs or an extended hydrogen bonding network.

The Family 3 peptides, CMP-Adpn peptides 5.1b - 5.4, have highlighted the crucial structural role of charged lysine residues located in the collagenous domain of Adpn for formation of a triple helical secondary structure of proteins as only the lysine-containing peptides were able to adopt a triple helical secondary structure in this family of peptide. Furthermore, although neoglycosylated lysine-containing CMP-Adpn peptide 5.1b was able to form a triple helical structure, the thermal stability of the neoglycosylated peptide was found to be significantly lower than the unglycosylated peptide 5.2. This result could have key implications for the use of neoglycosides to modify the thermal properties of collagenous materials for medical and biotechnological applications.

With regards to the structural role of glycosylated hydroxylysine residues in the collagenous domain of Adpn in particular, the Family 3 CMP-Adpn peptides 5.1b - 5.4 indicated that both non-glycosylated lysine residues as well as glycosylated lysine residues are fundamentally important for the structural stability of the triple helical secondary structure of the collagenous domain of Adpn. This could explain the great heterogeneity observed in the degree of glycosylation of hydroxylysine observed in vivo by Cooper, with four different states of the lysine residues observed in human Adpn: unmodified lysine residues, hydroxylation to afford Hyl, addition of galactose to form Gal-Hyl and the addition of the disaccharide to afford Glu-Gal-Hyl. It is possible that there is an ‘optimal level’ of glycosylation for each of the four lysine residues present in the collagenous domain of Adpn and for the formation of the bioactive HMW oligomers these ‘optimal levels’ of glycosylation must be observed for each lysine
residue. Indeed, the decreased levels of bioactive HMW Adpn observed in Type II diabetics,\textsuperscript{2} may be due to glycosylation of these four lysine residues above the ‘optimal level’ needed for HMW formation.

In summary, the work presented in this doctoral thesis has increased understanding of the structural roles of lysine and glycosylated lysine residues in collagenous peptides such as Adpn which can largely be attributed to intrastrand and interstrand electrostatic and hydrogen bonding interactions.

### 6.7 Future work

This thesis provides significant insight into the role of glycans in stabilising the triple helical structure of collagenous proteins. Furthermore, the key role of the lysine residues present in Adpn was also further investigated through the synthesis of analogous serine-containing peptides for comparative purposes.

Future work for this project could involve further investigation of the physical properties of the triple helical peptides synthesised herein namely; CMP-Lys-1Gal-OH \textbf{4.20b}, CMP-Lys-Control \textbf{4.23}, CMP-Ser-1Gal-OH \textbf{4.24b}, CMP-Ser-Control \textbf{4.27}, CMP-Adpn(66-83)-Lys-3Gal-OH \textbf{5.1b} and CMP-Adpn(66-83)-Lys-Control \textbf{5.2}. Further CD experiments could include isothermal CD refolding experiments to determine the kinetics of triple helix formation and investigation of the effect of changing the pH on the triple helix structure. Similar CD experiments examining the relationship between pH and the structure of CMPs, but not Adpn-derived peptides, have already been reported by Brodskey \textit{et al.}\textsuperscript{204} and Wennemers \textit{et al.}\textsuperscript{263}

The structural properties of the peptides could be further studied by attempting to grow single crystals of the peptides to enable detailed examination of the structures formed, in an approach similar to Berman \textit{et al.}\textsuperscript{203}. In particular, this could help to prove or disprove the hypothesis presented in this thesis that the CMP peptides adopt a 7/2 triple helical pitch, whereas the CMP-Adpn peptides tend to adopt a 10/3 triple helix.

NMR studies of the peptides would also provide further insight into the role of charged residues on the stability of the triple helical peptides. Hartgerink \textit{et al.}\textsuperscript{246} reported the solution structures of several CMPs in their 2012 paper. Through these NMR investigations Hartgerink \textit{et al.}\textsuperscript{246} gained a better understanding of the role of charged residues on triple helix formation. Iwahara \textit{et al.}\textsuperscript{251} also report the use of NMR spectroscopy for the determination of ion-pair interactions between glutamic acid and lysine residues, the expected chemical shifts of which are show below, Figure 6.7.
Figure 6.7 Expected chemical shifts for charged Glu and Lys residues, figure taken from the publication by by Iwahara et al.\textsuperscript{251}

The effect of length of spacer of the ‘guest domain’ of the Family 2 CMPs could also be investigated. In this thesis, the control peptides \textbf{4.23} CMP-Lys-Control and \textbf{4.27} CMP-Ser-Control with the ‘guest domain’ of GEK/ GES were synthesised and the structural properties of these CMPs investigated using circular dichroism studies. One could also synthesise control CMPs \textbf{6.1} - \textbf{6.4}, with the ‘guest domain’ of GEKGEK/ GESGES and GEKGEKGEK/GESGESGES respectively to explore the effect of length of ‘guest domain’ on the triple helical properties of Family 2 CMPs, Figure 6.8. Furthermore, synthesis of CMPs \textbf{6.1} - \textbf{6.4} would enable direct comparison of glycosylated and non-glycosylated CMPs with the same length ‘guest domain’ e.g. the circular dichroism spectrum of CMP \textbf{4.21b} CMP-Lys-2Gal-OH could be directly compared to CMP \textbf{6.1} CMP-2Lys-Control.

Figure 6.8 Additional Lys and Ser control Family 2 CMPs which could be synthesised to investigate the effect of spacer length upon triple helical formation.

Further peptide synthesis could also include different levels of glycosylation for Family 3 Adpn-CMPs, i.e. incorporation of one glycan, two glycans etc. Furthermore, positional variation of the glycan residues could also be investigated. For example, CMP-Adpn peptides \textbf{6.5} – \textbf{6.10} could be synthesised in order to establish the effect of glycan position on propensity of the peptide to form a triple helix, Figure 6.9.
Figure 6.9 Additional glyosylated Family 3 CMP-Adpn peptides 6.5 – 6.10 which could be synthesised to investigate the effect of glycan position on triple helical formation of the resultant peptide.

Multimerisation of the triple helical peptides could also be studied using mass spectrometry over a range of temperatures, in a similar method to that described by Mitra et al.\textsuperscript{264} Dynamic light scattering experiments could also be used to examine higher order structures of such peptides. Furthermore, sedimentation velocity analytical ultracentrifugation could also be used to assess the oligomerisation properties of the triple helical peptides, as demonstrated by Ohkubo et al.\textsuperscript{49} for their glycosylated Adpn peptides which reportedly formed octomers in solution at physiological pH and room temperature. Alternatively, size exclusion chromatography with multiangle laser light scattering (SEC-MALS) experiments could be used to investigate the oligomeric state of the CMPs in solution using the method described by Mitra et al.\textsuperscript{265}

The use of enzymatic glycosylation of lysine residues in Adpn sequences could also be explored. Myllylä et al.\textsuperscript{62} have reported that lysyl hydroxylase isoform 3 (LH3) is responsible for both the hydroxylation and glycosylation of lysine residues in collagenous proteins. Initial studies on the use of LH3 primer, kindly provided by Professor Myllylä and further expressed in both COS-7 and CHO cell lines, which were undertaken as part of these thesis studies, were unfortunately unsuccessful and have not been reported here.

Finally, the work presented in this thesis has indicated that there may be an optimal level of glycosylation for collagenous proteins, with glycosylation over this level perhaps contributing towards the pathogenesis of diseases with collagen defects such as Type II diabetes. A study of the glycosylation levels of collagenous proteins of patients with such disorders, for example with diabetic complications, could help to provide an improved molecular understanding for the pathogenesis of these diseases.
Chapter Six
Chapter Seven

Experimental Methods and Product Characterisation
Chapter Seven: Experimental Methods and Product Characterisation

All reagents were purchased as reagent grade from Sigma Aldrich unless otherwise specified, and were used without further purification. All reactions were done under an atmosphere of nitrogen unless otherwise stated. Analytical thin layer chromatography (TLC) was performed using Millipore TLC silica gel 60 F\textsubscript{254} plates and compounds were visualised by ultra-violet fluorescence or by staining with 4% sulphuric acid in ethanol (sugar products) or ethanolic ninhydrin solution (amine products) followed by gentle heating of the plate. Flash column chromatography was performed using Grace Davison Discovery Sciences Davisil\textregistered chromatographic silica (40-63 μ) with the reported solvents. Molecular sieves were ground, dried and stored in 200 °C oven for at least 24 hours prior to use. Infrared spectra were obtained using a Perkin Elmer Spectrum 100 FT-IR spectrometer, with samples placed on a zinc selenide and diamond crystal. Absorption maxima are reported in wavenumbers (cm\textsuperscript{-1}) with the common abbreviations; s = strong, m = medium, w = weak, br = broad. Optical rotations were determined at the sodium D line (598 nm) at 20 °C, unless otherwise stated, with a Rudolph Research Analytical Autopol IV polarimeter and are given in units deg dm\textsuperscript{-1}cm\textsuperscript{3}g\textsuperscript{-1}. \textsuperscript{1}H and \textsuperscript{13}C nuclear magnetic resonance (NMR) spectra were recorded on a Bruker AVANCE 300 MHz or Bruker ADVANCEIII 400 MHz spectrometer. Chemical shifts are reported in parts per million (ppm) relative to the trimethylsilane signal at δH 0.0 ppm in CDCl\textsubscript{3}/SiMe\textsubscript{4} solvent. \textsuperscript{1}H NMR data values are reported as the chemical shift, relative integral, multiplicity (whereby s = singlet, d = doublet, dd = doublet of doublets, t = triplet, q = quartet, m = multiplet), coupling constant (\textit{J} in Hz) and assignment. \textsuperscript{13}C NMR values are reported as chemical shift, degree of hybridisation and assignment. High resolution mass spectrometry (HRMS) samples were recorded on a Bruker micrOTOF-Q2 machine using ESI as an ionisation source in a positive mode, unless otherwise specified. Boc-Ser-OH was purchased from Polypeptides. Fmoc-Lys-OH was purchased from GL Biochem.
1.1 Section A: Details for organic compounds

**Organic General Method A (achiral Sharpless aminohydroxylation)**

Osmium tetroxide solution (0.25 mL, 2.5% w/w in t-BuOH, 0.02 mmol, 0.04 eq.) was added dropwise to a stirred solution of benzyl 4-chlorobenzoyloxy carbamate (20.0 g, 0.67 mmol, 1.40 eq.) in acetonitrile (0.6 mL) and allowed to stir at rt for 10 min. The alkene (0.05 mmol) was then added, followed immediately by distilled water (0.13 mL). The reaction mixture was stirred at rt for 2.5 h, quenched with Na$_2$S$_2$O$_3$ (sat. aq. soln., 1 mL) and then diluted with distilled water (10 mL) and ethyl acetate (10 mL). The organic product was extracted into ethyl acetate (3 x 10 mL), washed with NaHCO$_3$ (sat. aq. soln. 2 x 10 mL) and brine (sat. aq. soln. 2 x 10 mL), dried using Na$_2$SO$_4$ and concentrated *in vacuo* to yield the crude product.

**Organic General Method B (chiral Sharpless aminohydroxylation)**

As per Organic General Method A, but with a reaction time of 3 h and with the addition of the chiral ligand (DHQ)$_2$PHAL (20.0 mg, 0.02 mmol, 0.05 eq.) in acetonitrile (3.5 mL) alongside benzyl 4-chlorobenzoyloxy carbamate.
Chapter Seven

**tert-Butyl-2-((tert-butoxycarbonyl)amino)-4-(2,2-dimethyl-4,6-dioxo-1,3-dioxan-5-yl)-4-oxobutanoate** 2.87

\[
\begin{align*}
\text{Ketone intermediate} & \\
2.87
\end{align*}
\]

1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC\text{I}, 1.00 g, 10.4 mmol, 1.5 eq.) and DMAP (\(N,N\)-dimethylaminopyridine, 1.27 g, 10.4 mmol, 1.5 eq.) was added to an ice-cold solution of Boc-Asp-O\text{Bu} (2.00 g, 6.91 mmol) in dichloromethane (20 mL). Meldrum’s acid (0.99 g, 6.91 mmol, 1 eq.) was added to the ice-cold reaction mixture and the resulting solution was stirred at rt for 2 h. The reaction mixture was washed with ammonium chloride (3 x 50 mL) and concentrated \textit{in vacuo} to afford the desired compound 2.87 as a yellow solid (4.51 g, 6.23 mmol, 90%). This compound was subjected to the subsequent reaction without further purification.

**(S)-1-benzylpyrrolidine-2-carboxylic acid hydrochloride** 2.97

BP.HCl

Ground potassium hydroxide (18.7 g, 333.3 mmol, 3.9 eq.) was added to stirred solution of \(L\)-proline (10.0 g, 86.9 mmol) in isopropanol (50 mL) and the resultant reaction mixture was heated at 40 \(^\circ\text{C}\) until reagents had dissolved. Benzyl chloride (15.4 mL, 133 mmol, 1.5 eq.) was added dropwise to the ice-cold reaction mixture and the resulting solution was then heated at 50 \(^\circ\text{C}\) for 5 h, then cooled to rt. \(\text{CH}_2\text{Cl}_2\) (20 mL) was added to the solution, which was left to stand at 0 \(^\circ\text{C}\) for 16 h. The following day, the solution was acidified to pH 3 (1M HCl) followed by dilution with \(\text{CH}_2\text{Cl}_2\) (100 mL), and the resulting slurry allowed to stand at 0 \(^\circ\text{C}\) for 16 h. The slurry was then filtered and the orange filtrate was concentrated \textit{in vacuo}, washed with ice-cold acetone, filtered and concentrated \textit{in vacuo} to afford the desired compound 2.97 as a colourless solid (38.7 g, 0.08 mmol, 92%).
Chapter Seven

Rf n/a (colourless precipitate), $\left[\alpha\right]_D^{20} -22.2$, c 1.0, EtOH (lit. -25.8, 1g/100ml, EtOH)$^{93}$, the $^1$H and $^{13}$C data were in agreement with the literature values,$^{64}$ mp 172.6 - 173.9 °C (lit. 174.1 - 176.8 °C),$^{64}$ HRMS (EI, [M+H]$^+$) found 206.1181, calc. for C$_{12}$H$_{16}$NO$_2$ 206.1176.

(S)-2-[N-(N’-benzylpropyl)amino]benzophenone hydrochloride$^{64}$ 2.98

BPB.HCl

A solution of (S)-1-benzylpyrrolidine-2-carboxylic acid hydrochloride (2.97, 1.00 g, 4.00 mmol) and N-methylimidazole (1.30 mL, 17.0 mmol, 4 eq.) in dry CH$_2$Cl$_2$ (20 mL) was stirred for 30 min at rt and subsequently cooled to 0 °C. Methanesulfonyl chloride (0.30 mL, 4.10 mmol, 1 eq.) was then added dropwise to the stirred mixture and the resulting pale yellow solution stirred for 5 min at rt. 2-Aminobenzophenone (0.74 g, 3.70 mmol, 0.90 eq.) was then added and the reaction mixture stirred at rt for 2 h followed by stirring at 50 °C for 3 h. The reaction was quenched using NH$_4$Cl (sat. aq. soln., 5 mL) and the organic product extracted into CH$_2$Cl$_2$ (3 x 20 mL), dried using Na$_2$SO$_4$, filtered and concentrated in vacuo. The residue was diluted with acetone (5 mL), acidified to pH 2 (2 M HCl) and stirred at rt for 3 h. The resulting colourless solid was filtered, washed with ice-cold acetone and air-dried to afford the desired compound 2.98 as an off-white solid (1.84 g, 4.0 mmol, quant.).

Rf n/a (colourless precipitate), $\left[\alpha\right]_D^{20} -115.7$, c 0.52, MeOH (lit. -130.8, c 0.15, MeOH)$^{64}$, the $^1$H and $^{13}$C data were in agreement with the literature values,$^{64}$ mp 100.3 - 101.4 °C (lit. 101.2 - 103.5 °C),$^{64}$ HRMS (EI, [M+H]$^+$) found 385.1915, calc. for C$_{25}$H$_{25}$N$_2$O$_2$ 385.1911.
Belokon complex

Potassium hydroxide (0.18 g, 3.05 mmol, 3 eq.), glycine (0.11 g, 1.53 mmol, 5 eq.) and Ni(NO$_3$)$_2$.6H$_2$O (0.18 g, 0.60 mmol, 2 eq.) were added to a stirred solution of (S)-2-[N-(N’-benzylpropyl)amino]benzophenone hydrochloride (2.98, 0.13 g, 0.3 mmol) in methanol (AR grade, 3 mL) and the reaction mixture was heated under reflux for 2 h, cooled to rt and neutralised (glacial acetic acid, 0.15 mL). The reaction mixture was stirred at rt for 16 h, diluted with distilled water (20 mL) and left to stand at rt for 2 h. The organic product was extracted into CH$_2$Cl$_2$ (5 x 20 mL) dried using Na$_2$SO$_4$, filtered, concentrated in vacuo then purified using flash column chromatography (CH$_2$Cl$_2$: acetone, 7:1 → 2:1) followed by recrystallization (CH$_2$Cl$_2$: n-hexanes (1:2) to afford the desired compound 2.93 as blood-red crystals (0.15 g, 0.3 mmol, quant.).

R$_f$ 0.22 (CH$_2$Cl$_2$/acetone, 4:1), [$\alpha$]$_D$ +1800.0, c 0.12, MeOH) (lit. +2008.7, c 0.12, MeOH$^{64}$, the $^1$H and $^{13}$C data were in agreement with the literature values$^{64}$ mp 165.0 – 173.0 °C (lit. 203.0 - 210.0 °C)$^{91}$, HRMS (EI, [M+H]$^+$) found 498.1317, calc. for C$_{27}$H$_{26}$N$_3$NiO$_3$ 498.1322.
(S)-glycine-nickel-(S)-2-[N-(N’-benzylpropyl)amino]benzophenone-but-1-ene complex \( ^{64} 2.92 \)

(2S)-alkylated Belokon complex

Ground sodium hydroxide (0.80 g, 20.0 mmol, 10 eq.) was added to a stirred solution of (S)-glycine-nickel-(S)-2-[N-(N’-benzylpropyl)amino]benzophenone (2.93, 1.00 g, 2.00 mmol) in acetonitrile (60 mL) and the reaction mixture stirred at rt for 5 min. 4-Bromobut-1-ene (0.2 mL, 1.9 mmol, 0.98 eq.) and NaI (0.09 g, 0.6 mmol, 0.3 eq.) were then added and the reaction mixture stirred at rt for 3 h. The reaction mixture was quenched (0.1 M HCl) and stirred at rt for 30 min. The organic product was extracted into CH\(_2\)Cl\(_2\) (3 x 50 mL), washed with brine (50 mL), dried using Na\(_2\)SO\(_4\), concentrated \textit{in vacuo} and purified by flash column chromatography (CH\(_2\)Cl\(_2\): acetone, 4:1). Examination of the blood red solid by NMR revealed a \(2S/2R\) ratio of 94%: 6%. The pure (2S)-enantiomer was obtained by recrystallisation (CH\(_2\)Cl\(_2\): n-hexanes, 1:3.5) to afford the desired product 2.92 as dark red crystals (2S)-enantiomer only, 1.10 g, 1.9 mmol, 94%).

R\(_f\) 0.48 (CH\(_2\)Cl\(_2\): acetone, 4:1); [\(\alpha\)]\(_D\)\(^{20}\) +2600.0, c 0.02, CH\(_2\)Cl (lit. +2471.0, c 0.01, CH\(_3\)Cl)\(^{97}\), the \(^1\)H and \(^13\)C data were in agreement with the literature values,\(^{97}\) mp 218.0 - 220.6 °C (lit. 207.0 - 209.0 °C)\(^{97}\); HRMS (EI, [M+Na]\(^{+}\)) found 574.1601, calc. for C\(_{27}\)H\(_{36}\)N\(_3\)NiO\(_3\) 574.1611.

(2S)-2-Amino-5-hexenoic acid\(^{97}\) 2.91

Hydrochloric acid (2 M, 40 mL) was added to a stirred solution of (S)-glycine-nickel-(S)-2-[N-(N’-benzylpropyl)amino]benzophenone-but-1-ene complex (2.92, 1.66 g, 3.00 mmol) in methanol (AR, 40 mL) and the resulting red solution was heated under reflux at 120 °C for 1 h, during which time the colour changed from red to orange then to yellow. After cooling to rt, ammonium solution (25% in water) was added until the pH of the reaction mixture was pH 9. The
resulting colourless ppt (recovered 2.98, BPB.HCl) which was filtered from the reaction mixture, washed with ice-cold ammonia solution, air-dried and stored. The blue filtrate was washed with CH₂Cl₂ (3 x 50 mL), the aqueous layers isolated, combined, concentrated in vacuo then purified using flash column chromatography (Dowex H⁺ column, eluent as 5% ammonia solution) to afford the desired compound 2.91 as a colourless fluffy solid (0.39 g, 3.0 mmol, quant.).

Rᵣ n/a (TLC plate stained with ninhydrin for presence of free amine), [α]ᴰ22 +15.7, c 0.14, H₂O, (lit. +13.1, c 1.30, H₂O)⁹⁷, the ¹H and ¹³C data were in agreement with the literature values,⁹⁷ mp 258.4 - 259.7 °C (lit. > 270 °C)⁹⁷, HSMS (ESI) found 130.0866, calc. for C₆H₁₁NO₂Na⁺ 130.0863.

Methyl (2S)-(N-tert-butoxycarbonyl)amino)hex-5-enoate 2.161

Thionyl chloride (0.09 mL, 1.20 mmol, 1.20 eq.) was added dropwise to a stirred ice-cold suspension of (S)-2-amino-5-hexenoic acid (2.91, 0.13 g, 1.00 mmol) in methanol (AR, 25 mL). The reaction mixture was then stirred at rt for 3h and excess thionyl chloride was removed by concentration in vacuo. The resulting light pink solid was diluted with acetonitrile (25 mL) and cooled over ice. Triethylamine (0.17 mL, 1.30 mmol, 1.20 eq.) followed by di-tert-butyl dicarbonate (0.27 g, 1.20 mmol, 1.20 eq.) were added to the ice cold mixture and the reaction stirred at rt for 16 h. The reaction mixture was then concentrated in vacuo and ethyl acetate (30 mL) and water (DI, 30 mL) were added to the residue. The organic product was further extracted into ethyl acetate (2 x 30 mL), dried using Na₂SO₄ and concentrated in vacuo to yield the desired compound 2.161 as a light orange oil (0.19 g, 0.80 mmol, 76%). This compound was subjected to the subsequent reaction without further purification.
(2S)-N-(9H-Fluorenlymethoxycarbonyl)-2-amino-5-hexenoic acid \(^{267}\) 7.1

A solution of N-(9H-fluorenlymethoxy carbonyloxy)succinimide (1.49 g, 4.40 mmol, 1.10 eq.) in dioxane (10 mL) was added dropwise to stirred ice-cold solution of (S)-2-amino-5-hexenoic acid (2.91, 0.52 g, 4.00 mmol) in dioxane/Na\(_2\)CO\(_3\) solution (10 mL/ sat. aq. 30 mL). The reaction mixture was stirred at rt for 4 h and concentrated \(in\ vacuo\). The organic product was extracted into ethyl acetate (3 x 30 mL), washed with NaHCO\(_3\) (sat. aq. soln., 30 mL), dried using Na\(_2\)SO\(_4\), concentrated \(in\ vacuo\) then purified using flash column chromatography (ethyl acetate: \textit{n}-hexanes, 1:1) to yield the desired compound 7.1 as a colourless solid (1.32 g, 3.60 mmol, 94%).

\(R_f\) 0.26 (ethyl acetate: \textit{n}-hexanes, 1:1), \([\alpha]_D^{23}\) +20.0, \(c\) 0.12, CHCl\(_3\) (lit. +13.3, \(c\) 1.00, CHCl\(_3\))\(^{267}\), the \(^1\)H and \(^{13}\)C data were in agreement with the literature values\(^{267}\), mp 125.7 - 127.4\(^\circ\)C (no lit. mp); HRMS (ESI) found 374.1355, calc. for C\(_{21}\)H\(_{21}\)NO\(_4\)Na\(^+\) 374.1363.

Methyl (2S)-N-(9H-Fluorenlymethoxycarbonyl)-2-amino-5-hexenoate \(^{268}\) 2.128

Thionyl chloride (0.56 mL, 7.70 mmol, 5.00 eq.) was added dropwise to a stirred ice-cold solution of (S)-2-amino-5-hexenoic acid (2.91, 0.20 g, 1.50 mmol, 1.00 eq.) in methanol (AR, 5 mL). The resulting orange solution was heated under reflux for 2 h, cooled to rt and concentrated \(in\ vacuo\). The reaction mixture was then quenched using NaHCO\(_3\) (sat. aq. soln., 5 mL) diluted with CH\(_2\)Cl\(_2\) (5 mL). N-(9-Fluorenlymethoxy carbonyloxy)succinimide (0.58 g, 1.70 mmol, 1.10 eq.) was added to the orange reaction mixture, and the resulting biphasic solution was stirred at rt for 4 h. After concentration \(in\ vacuo\), the organic product was extracted into ethyl acetate (3 x 5 mL), washed with brine (5 mL), dried using Na\(_2\)SO\(_4\), concentrated \(in\ vacuo\) and purified using flash column chromatography (ethyl acetate: \textit{n}-hexanes: 1:4) to afford the desired compound 2.128 as a colourless solid (0.50 g, 1.40 mmol, 92%).

\(R_f\) 0.1 (ethyl acetate: \textit{n}-hexanes, 1:1), \([\alpha]_D^{24}\) +8.8, \(c\) 0.16, CHCl\(_3\) (no lit. \([\alpha]_D\)), the \(^1\)H and \(^{13}\)C data were in agreement with the literature values\(^{268}\), mp 94.9 - 95.4\(^\circ\)C (lit. mp 98.0 - 00.0\(^\circ\)C)\(^{268}\), HRMS (ESI) found 388.1524, calc. for C\(_{22}\)H\(_{23}\)NO\(_4\)Na\(^+\) 388.1519.
Chapter Seven

**tert-Butyl (2S)-N-(9H-Fluorenlymethoxycarbonyl)-2-amino-5-hexenoate**

\[
\text{\textbf{2.160}}
\]

\[\text{NH_\text{Fmoc}} - \text{CC_2Bu} \]

\[\text{2.160}\]

\[N,N'-(\text{Dimethylamino})\text{pyridine} (6.90 \text{ mg, 0.05 mmol, 0.10 eq.}) \text{ and } \text{\textsc{t}BuOH} (0.06 \text{ mL, 0.62 mmol, 1.10 eq.}) \text{ were added to a stirred solution of (2S)-N-(9H-fluorenlymethoxycarbonyl)-2-amino-5-hexenoic acid (7.1, 0.20 g, 0.50 mmol) in CH}_2\text{Cl}_2 \text{ (dry, 10 mL) and the mixture stirred at rt for 5 min. } \]  
\[N,N'-(\text{Dicyclohexylcarbodiimide (DCC, 0.12 g, 0.62 mmol, 1.10 eq.) was added to the reaction mixture and the suspension was stirred at rt for 16 h. The following day further portions of } \text{\textsc{t}BuOH} (0.06 \text{ mL, 1.1 eq.), DCC (0.11 g, 1.0 eq.) and 0.1 eq. of DMAP (6.9 mg, 0.1 eq.) were added and the reaction mixture was stirred at rt for a further 5 h. The reaction mixture was then filtered, washed (1 x distilled water, 20 mL; 1 x 5% acetic acid solution, 20 mL; 1 x DI H}_2\text{O, 20 mL), dried using Na}_2\text{SO}_4 \text{ and concentrated in vacuo. The crude product was purified using flash column chromatography (ethyl acetate: } n\text{-hexanes, 1:9 } \rightarrow 1:0) \text{ to afford the desired product } \textbf{2.160} \text{ as a colourless solid (0.17 g, 0.41 mmol, 72%).}\]

\[R_f 0.69 \text{ (ethyl acetate: hexane, 1:1), } [\alpha]_\text{D}^{24} +7.0, \text{ c 0.10, CHCl}_3 \text{ (no lit. } [\alpha]_\text{D}) \text{, the } ^1\text{H and } ^{13}\text{C data were in agreement with the literature values,}^{[229]} \text{ mp 61.7-63.8 (no lit. mp), HRMS (ESI) found 430.1980, calc. for C}_{25}\text{H}_{29}\text{NO}_4\text{Na}^+ 430.1989.}\]

**2.180**

\[(2S)-\text{Amino-5-hexenoic acid bicyclo[3.3.1]nonylboron}\]

\[\text{\textbf{2.180}}
\]

\[(\text{S})-\text{2-Amino-5-hexenoic acid (2.91, 0.15 g, 1.13 mmol) was added to a stirred solution of 9-borabicyclo[3.3.1]nonane dimer (0.30 g, 1.23 mmol, 1.2 eq.) in dry methanol (3 mL). The reaction mixture was heated at reflux for 3 h and then concentrated in vacuo. The resulting off-colourless solid residue was triturated with hexane (2 x 10 mL) and diethyl ether (2 x 10 mL) to afford the desired product } \textbf{2.180} \text{ as an off-white solid (0.24 g, 0.94 mmol, 83%). This compound was subjected to the subsequent reaction without further purification.}\]

\[R_f 0.63 \text{ (2% ammonia, 18% methanol, 80% CH}_2\text{Cl}_2), \delta_\text{H}(400 \text{ MHz; CDCl}_3; \text{Me}_4\text{Si}) 0.58 \text{ (2 H, br s, BBN, 2 x CH}_3), 1.42 - 1.88 \text{ (22 H, m, BBN, 6 x CH}_2 \text{ plus excess BBN-dimer), 2.17 - 2.41 \text{ (4H, m, } ^\gamma\text{-CH}_2\text{)}}\]
δ-CH₂), 3.75 - 3.84 (1 H, m, α-CH), 4.01 - 4.06 (1H, t, J 7.44, NH), 4.83 - 4.88 (1H, t, J 7.44, NH), 5.13 - 5.22 (2H, m, ε-CH₂), 5.79 - 5.90 (2H, m, δ-CH₂). δC (400 MHz; CDCl₃; Me₄Si) 23.7, 24.3 (CH, BBN), 29.7, 30.1 (2 x CH₂, β-C, γ-C), 30.9, 31.1, 31.6, 32.0 (CH₂, BBN), 55.3 (CH, α-C), 117.5 (CH₂, ε-C), 136.6 (CH₂, δ-C). HRMS (ESI) found 272.1801, calc. for C₁₄H₂₄BNO₂Na⁺ 272.1795.

±-Benzyl 2-Hydroxy-2-phenylethylcarbamate and ±-Benzyl 2-Hydroxy-1-phenylethylcarbamate

2.124 and 2.125

The synthesis was conducted following the Organic General Method A (achiral Sharpless aminohydroxylation) using styrene (55 µL, 0.05 mmol) as the starting material. The crude material was purified using flash column chromatography (ethyl acetate: n-hexanes, 3:7 → 1:1 → ethyl acetate) to yield the desired aminohydroxylation products as an off-white solid as a 7:1 mixture of regioisomers 2.124:2.125 (0.11 g, 83%).

Rf 0.27 (ethyl acetate: n-hexanes, 3:7), the ¹H and ¹³C data were in agreement with the literature values, HRMS (ESI) found 294.111, calc. for C₁₆H₁₇NO₃Na⁺ 294.1101.

Benzyl (1S,2R)-2-Hydroxycyclohexylcarbamate

2.127

The synthesis was conducted according to General Method B (chiral Sharpless aminohydroxylation) using cyclohexane (45 µL, 0.45 mmol) as the reaction substrate. The crude material was purified by flash column chromatography (ethyl acetate: n-hexanes, 3:7 → 1:1 → ethyl acetate) to yield the desired product 2.127 as an off white solid (0.05 g, 0.21 mmol, 47%, 23% e.e.).

Rf 0.24 (ethyl acetate: n-hexanes, 3:7), [α]D²⁴ +7.1, c 0.45, CHCl₃ (lit. [α]D²³ +29.3, c 0.95, EtOH)²⁷⁰, the ¹H and ¹³C data were in agreement with the literature values, mp 69.0 - 70.0 (75.0 - 77.0 °C),²⁷⁰ HRMS (ESI) found 272.1851, calc. for C₁₄H₁₅NO₃Na⁺ 272.1857.
Methyl (E)-3-Phenylprop-2-enoate (methyl cinnamate)\(^{271}\) 2.113

\[
\text{\begin{tabular}{c}
\begin{tabular}{c}
\text{\includegraphics[width=0.2\textwidth]{cinnamate.png}}
\end{tabular}
\end{tabular}}
\]

2.113

Concentrated sulphuric acid (2 drops, cat.) was added to a stirred solution of \textit{trans}-cinnamic acid (1.00 g, 6.80 mmol) in methanol (20 mL) and the solution was heated at reflux for 16 h. After cooling to rt, the crude mixture was concentrated \textit{in vacuo} and purified using flash column chromatography (ethyl acetate: \textit{n}-hexanes, 3:7 → 1:1) to yield the desired product 2.113 as a colourless, crystalline solid (1.00 g, 6.30 mmol, 93%).

R\(_f\) 0.52 (ethyl acetate: \textit{n}-hexanes, 3:7), the \(1^H\) and \(13^C\) data were in agreement with the literature values\(^{272}\), mp 40.4 - 41.0 (no lit. mp), HRMS (ESI) found 185.0574, calc. for C\(_{10}\)H\(_{10}\)O\(_2\)Na\(^+\) 185.0573.

Methyl (2\(R\),3\(S\))-3-(Benzyloxy carbonylamino)-2-hydroxy-3-phenylpropanoate\(^{99}\) 2.115

\[
\text{\begin{tabular}{c}
\begin{tabular}{c}
\text{\includegraphics[width=0.2\textwidth]{benzylamine.png}}
\end{tabular}
\end{tabular}}
\]

2.115

The synthesis was conducted according to \textbf{General Method B} (chiral Sharpless aminohydroxylation) with the following quantities of reagents: osmium tetroxide solution (0.4 mL, 2.5 % w/w in \textit{t}BuOH, 0.02 mmol, 0.04 eq.), benzyl-4-chlorobenzyloxy carbamate (34.0 mg, 0.11 mmol, 1.4 eq.) and (DHQ)\(_2\)PHAL (31 mg, 0.03 mmol, 0.06 eq.) in acetonitrile (3.5 mL), methyl (E)-3-phenylprop-2-enoate (methyl cinnamate (2.113, 128.0 mg, 0.78 mmol). The crude material was purified by flash column chromatography (ethyl acetate: \textit{n}-hexanes, 1:3 → 1:1) to yield the desired product 2.115 as a colourless solid (229 mg, 0.70 mmol, 89%, 53\% \textit{e.e}).

R\(_f\) 0.17 (ethyl acetate: \textit{n}-hexanes, 1:3), [\(\alpha\)]\(_D\)\(^{23}\) -1.5, c 0.13, CHCl\(_3\) (lit. [\(\alpha\)]\(_D\)\(^{24}\) -2.7, c 1.15, CHCl\(_3\)\(^{99}\)), the \(1^H\) and \(13^C\) data were in agreement with the literature values, \(^{99}\) mp 106.8 - 107.2 (120.0 -121.0 °C),\(^{99}\) HRMS (ESI) found 352.1161, calc. for C\(_{18}\)H\(_{19}\)NO\(_5\)Na\(^+\) 352.1155.
(2S,5R)-Nα-(tert-butoxycarbonyl)-Nε-benzyloxy carbonyl-5-hydroxy-L-lysine methyl-ester 7.2a, (2S,5S)-Nα-(tert-butoxycarbonyl)-Nε-benzyloxy carbonyl-5-hydroxy-L-lysine methyl-ester 7.2b

The synthesis was conducted according to **General Method B** (chiral Sharpless aminohydroxylation) with the following quantities of reagents: osmium tetroxide solution (0.1 mL, 2.5 % w/w in tBuOH, 0.02 mmol, 0.04 eq.), benzyl 4-chlorobenzyloxy carbamate (88.0 mg, 0.29 mmol, 1.40 eq.), (DHQ)2PHAL (8.0 mg, 0.01 mmol, 0.05 eq.) in acetonitrile (3 mL) and methyl (2S)-(N-tert-butoxycarbonyl)amino)hex-5-enoate (2.161, 50.0 mg, 0.21 mmol) in acetonitrile (3 mL). Purification of the crude material by flash column chromatography (ethyl acetate: n-hexanes, 3:7) was attempted, but unfortunately only a complex mixture was afforded.

(2S,5R)-Nα-(tert-butoxycarbonyl)-Nε-benzyloxy carbonyl-5-acetoxy-L-lysine methyl-ester 7.3a, (2S,5S)-Nα-(tert-butoxycarbonyl)-Nε-benzyloxy carbonyl-5-acetoxy L-lysine methyl-ester 7.3b

Triethylamine (6 µL, 0.04 mmol, 1.10 eq.) followed by acetyl chloride (5 µL, 0.07 mmol, 2.00 eq.) were added to a stirred ice-cold solution of methyl (2S)-6-(((benzyloxy)carbonyl)amino)-2-((tert-butoxy carbonyl)amino)-5-hydroxyhexanoate (2.161, 15.0 mg, 0.036 mmol) in CH2Cl2 (0.5 mL). The reaction mixture was stirred at rt for 1 h, diluted with ethyl acetate (2 mL) and neutralised using HCl (1 M). The organic product was extracted into ethyl acetate (3 x 20 mL), dried using Na2SO4 and concentrated in vacuo to afford an inseparable complex mixture.

Rf 0.60 (ethyl acetate: n-hexanes, 1:1), 1H and COSY data indicated a resonance corresponding to δ-CHOH (5.72) to δ-CHOAc (4.90).
Benzyl 4-chlorobenzoyloxycarbamate (407 mg, 1.33 mmol, 1.7 eq.) was added to a solution of methyl (2S,N)-(9H-fluorenylmethoxycarbonyl)-2-amino-5-hexenoate (2.128, 286 mg, 0.78 mmol) in acetonitrile (10 mL) and the reaction mixture was stirred at rt for 10 min. Osmium tetroxide solution (0.6 mL, 2.5% w/w in tBuOH, 0.02 mmol, 0.05 eq.) was added dropwise to the stirred reaction mixture, followed immediately by deionised water (1.25 mL). The reaction mixture was stirred at 35 °C for 16 h, after which time a second portion of osmium tetroxide solution (0.6 mL, 2.5% w/w in tBuOH, 0.02 mmol, 0.05 eq.) was added and the reaction was stirred again at 35 °C for 16 h.

TLC showed incomplete consumption of starting material, hence a third portion of osmium tetroxide solution (0.6 mL, 2.5% w/w in tBuOH, 0.02 mmol, 0.05 eq.) was added and the reaction mixture stirred at 35 °C for 16 h. The reaction mixture was then quenched with Na2S2O3 (sat. aq. soln., 1 mL), diluted with water (DI, 10 mL) and ethyl acetate (10 mL). The organic product was extracted into ethyl acetate (2 x 20 mL), washed with brine (1 x 10 mL, sat. soln.), dried using Na2SO4 and concentrated in vacuo.

The crude material was purified using flash column chromatography (ethyl acetate: n-hexanes, 1:1) to yield the desired product 2.129 as a colourless solid as a mixture of epimers (37.0 mg, 0.06 mmol, 95%).

Rf 0.15 (ethyl acetate: n-hexanes, 1:1), δH (400 MHz; CDCl3; Me4Si) 1.50 - 2.195 (4 H, m, β-CCH2, γ-CCH2), 3.08 (1 H, br s, ε-CCHa), 3.31 (1 H, br s, ε-CCHb), 3.73 (4 H, br s, OCH3, δ-CH), 4.19 (1 H, t, J 6.7, CHFmoc), 4.34 - 4.43 (3 H, m, α-CH, CH2Fmoc), 5.08 (2 H, br s, CH2Cbz), 5.29 (1 H, br s, NHCbz), 5.64 (1 H, br d, J 23.8, NHFmoc) 7.26 - 7.41 (9 H, m, CbzPh, FmocPh), 7.49 - 7.62 (2H, m, FmocPh), 7.75 (2H, d, J 7.4, FmocPh), δC (400 MHz; CDCl3; Me4Si) 29.7, 30.0 (2 x CH2, β-CH2, γ-CH2), 46.9 (CH2, ε-CH2), 47.1 (CH, CHFmoc), 52.5 (CH3, OCH3), 53.5 (CH, α-CH), 67.0, 67.1 (2 x CH2, CH2Fmoc, CH2Cbz), 70.7 (CH, δ-CHOH), 120.0, 125.0, 127.1, 127.7, 128.2, 128.6 (CH, CbzPh, FmocPh), 156.1, 157.2, 173.0 (C=O, Cbz, Fmoc, CO2Me), mp 72.9 – 73.9 °C, (HRMS (ESI) found 555.2100, calc for C30H32N2O7Na+ 555.2102.
Chapter Seven

(2S,5R)-N\textsuperscript{\alpha}-(9H-Fluorenlymethoxycarbonyl)-N\textsuperscript{\epsilon}-benzyloxy carbonyl-5-hydroxy-L-lysine tert-butyl-ester 2.132a and (2S,5S)-N\textsuperscript{\alpha}-(9H-Fluorenlymethoxycarbonyl)-N\textsuperscript{\epsilon}-benzyloxy carbonyl-5-hydroxy-L-lysine tert-butyl-ester 2.132b

Osmium tetroxide solution (0.2 mL, 2.5% w/w in \textsuperscript{t}BuOH, 0.02 mmol, 0.04 eq.) was added dropwise to a solution of benzyl 4-chlorobenzoyloxycarbamate (188 mg, 0.60 mmol, 1.4 eq.) in acetonitrile (3.5 mL) and allowed to stir at rt for 10 min. After 10 minutes, a solution of tert-butyl (2S)-N-(9H-fluorenlymethoxycarbonyl)-2-amino-5-hexenoate (2.160, 30 mg, 0.07 mmol) in acetonitrile (3.5 mL) was added, followed immediately by distilled water (0.85 mL). The reaction mixture was stirred at 35 °C for 16 h, allowed to cool to rt, quenched with Na\textsubscript{2}S\textsubscript{2}O\textsubscript{3} (1 mL, sat. soln.) then diluted with distilled water (10 mL) and ethyl acetate (10 mL). The organic product was extracted into ethyl acetate (2 x 20 mL), washed with brine (1 x 10 mL, sat. soln.), dried using Na\textsubscript{2}SO\textsubscript{4} and concentrated in vacuo. The crude material was purified by flash column chromatography (ethyl acetate: n-hexanes, 1:1) to yield the desired product 2.132 as a colourless solid as a mixture of epimers (28 mg, 0.05 mmol, 70%).

R\textsubscript{f} 0.29 (ethyl acetate: n-hexanes, 1:1), \(\delta\textsuperscript{H} (400 MHz; CDCl\textsubscript{3}; Me\textsubscript{4}Si) 1.46 (9 H, br s, \textsuperscript{t}Bu, 3 x CH\textsubscript{3}), 1.64 - 2.03 (4 H, m, \beta-CH\textsubscript{2}, \gamma-CH\textsubscript{2}), 3.07 (1 H, br s, \varepsilon-CH\textsubscript{2}), 3.34 (1 H, br s, \varepsilon-CH\textsubscript{2}), 3.61 -3.74 (1 H, m, \delta-CH), 4.20 (1 H, t, \textit{J} 7.3, CHFmoc), 4.27 (1 H, br s, \alpha-CH), 4.38 (2 H, d, \textit{J} 7.6, CH\textsubscript{2}Fmoc), 5.08 (2 H, br s, CH\textsubscript{2}Cbz), 5.23 (1 H, br s, NHCbz), 5.54 (1 H, dd, \textit{J} 1.6, \textit{J} 19.0, NHFmoc) 7.26 – 7.41 (9 H, m, CbzPh, FmocPh), 7.58 (2H, d, \textit{J} 7.3, FmocPh), 7.75 (2H, d, \textit{J} 7.3, FmocPh). \(\delta\textsuperscript{C} (400 MHz; CDCl\textsubscript{3}; Me\textsubscript{4}Si) 28.0 (3 x CH\textsubscript{3}, \textsuperscript{t}Bu), 29.5, 30.0 (2 x CH\textsubscript{2}, \beta-CH\textsubscript{2}, \gamma-CH\textsubscript{2}), 47.0 (CH\textsubscript{2}, \varepsilon-CH\textsubscript{2}), 47.1 (CH, CHFmoc), 53.8, 54.0 (2x CH, \alpha-CH), 66.9, 67.0 (2 x CH\textsubscript{2}, CH\textsubscript{2}Fmoc, CH\textsubscript{2}Cbz), 70.7, 70.8 (2 x CH, \delta-CHOH), 82.3, 82.4 (C, \textsuperscript{t}Bu), 119.9, 125.0, 127.8, 127.7, 128.1, 128.1, 128.2, 128.5, 136.3, 1541.3, 143.8 (CH, CbzPh, FmocPh), 156.2, 157.1, 171.3 (C=O, Cbz, Fmoc, CO\textsubscript{2}tBu). HRMS (ESI) calcd for C\textsubscript{33}H\textsubscript{38}N\textsubscript{2}O\textsubscript{7}Na\textsuperscript{+} 597.2568, found 597.2571.
(2S,5R)-5-methylbenzyloxy carbonyl-2S-(9H-Fluorenylmethoxycarbonyl)amino-δ-valerolactone, 2.134 and (2S,5S)-5-methylbenzyloxy carbonyl-2S-(9H-Fluorenylmethoxycarbonyl)amino-δ-valerolactone, 2.135

![Molecules 2.134 and 2.135](image)

Trifluoroacetic acid (neat, 8 drops) was added to a solution of tert-butyl (2S)-N-(9H-fluorenylmethoxycarbonyl)-2-amino-5-hexenoate (2.132, 20 mg, 0.35 mmol) in dry CH<sub>2</sub>Cl<sub>2</sub> over molecular sieves. The resulting reaction mixture was stirred at 35 °C for 16 h. Due to incomplete reaction, a further portion of trifluoroacetic acid (neat, 0.1 mL, 13 eq.) was added along with triethylsilane (as a carbocation scavenger<sup>113</sup>, neat, 0.14 mmol, 25 eq.) and the reaction mixture stirred at 35 °C for 2 h. The organic product was extracted into ethyl acetate (3 x 20 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated in vacuo to afford an inseparable complex mixture.

(2S,5R)-N<sup>α</sup>-(bicyclo[3.3.1]nonylboron)-N<sup>ε</sup>-benzyloxycarbonyl-5-hydroxy-L-lysinate 2.178a and (2S,5S)-N<sup>α</sup>-(bicyclo[3.3.1]nonylboron)-N<sup>ε</sup>-benzyloxycarbonyl-5-hydroxy-L-lysinate 2.178b

![Molecules 2.178a and 2.178b](image)

(±)-5-hydroxy-L-lysine hydrochloride (epi-Hyl, 2.177, 200 mg, 1.01 mmol, purchased from Sigma-Aldrich) was added to ice-cold ammonium hydroxide solution (28% NH<sub>3</sub> in H<sub>2</sub>O, 10 mL, excess) and the reaction mixture stirred for 30 min then concentrated in vacuo. The resulting colourless residue was added to a solution of 9-borabicyclo[3.3.1]nonane dimer (135 mg, 1.11 mmol, 1.1 eq.) in methanol (3 mL) and the reaction heated under reflux for 5 h then at rt for 16 h. The resulting white cloudy solution was then concentrated in vacuo and triturated with ether (2 x 20 mL) to yield a crude colourless solid.

NaHCO<sub>3</sub> (63.4 mg, 7.55 mmol, 1.5 eq.) followed by benzyl chloroformate (0.1 mL, 7 mmol, 1.4 eq.) were added to an ice-cold solution of the crude colourless solid (142 mg, 0.5 mmol) in dioxane: deionised water (dioxane: H<sub>2</sub>O, 1:1, 4 mL) and the reaction mixture stirred for at rt for 1.5 h. The reaction mixture was then concentrated in vacuo and the organic product extracted into ethyl acetate (3 x 5 mL), washed with brine (sat. soln. 10 mL), deionised water (10 mL), dried over Na<sub>2</sub>SO<sub>4</sub> and concentrated in vacuo. The crude material was purified using flash column chromatography (ethyl
acetate: n-hexanes, 7:3) to afford the desired product **2.178** as a colourless solid as an epimeric mixture (94 mg, 0.23 mmol, 43% yield from (±)-5-hydroxy-L-lysine hydrochloride).

Rf 0.29, (ethyl acetate: n-hexanes, 7:3), the $^1$H and $^{13}$C data were in agreement with the literature values$^{123}$, HRMS (ESI) caled for C$_{22}$H$_{33}$BN$_2$O$_5$Na$^+$ 439.2379, found 439.2375.

(2S,5R)-N$^\alpha$-(9H-Fluorenlymethoxycarbonyl)-N$^\varepsilon$-benzyloxycarbonyl-5-hydroxy-L-lysine allyl-ester **2.179a** and (2S,5S)-N$^\alpha$-(9H-Fluorenlymethoxycarbonyl)-N$^\varepsilon$-benzyloxycarbonyl-5-hydroxy-L-lysine allyl-ester **2.179b**

![Chemical structures](image)

Ethylenediamine (0.1 mL, 0.96 mmol, 4 eq.) was added to a solution of epimeric N$^\alpha$-(bicyclo[3.3.1]nonylboron)-N$^\varepsilon$-benzyloxycarbonyl-5-hydroxy-L-lysine (**2.178**, 100 mg, 0.24 mmol) in THF (dry, 10 mL) at 150 °C. After stirring for 10 min, the solution was poured over ice to afford a colourless precipitate which was filtered and washed with ice-cold THF (3 x 5 mL). A solution of N-(9H-fluorenlymethoxycarboxyl)uccinimide (115 mg, 0.34 mmol, 1.4 eq.) in dioxane (1 mL) was added dropwise to an ice-cold solution of the white precipitate in dioxane: K$_2$CO$_3$ solution (3 mL, 10% K$_2$CO$_3$ soln., 1 eq.). The reaction mixture was stirred at rt for 2 h, concentrated in vacuo and diluted with chloroform (10 mL). Citric acid solution (20% w/v) was added to adjust the pH of the reaction mixture until pH 4 was reached. The organic product was then extracted into chloroform (2 x 10 mL), washed with deionised water (10 mL), dried using Na$_2$SO$_4$ and concentrated in vacuo. NaHCO$_3$ (21 mg, 0.25 mmol, 1.05 eq.) added to a solution of the resultant residue in deionised water (5 mL) to form a white slurry.

A suspension of tetrabutylammonium bromide (81 mg, 0.25 mmol, 1.05 eq.) and allyl bromide (0.1 mL, 1.1 mmol, 5 eq.) in CH$_2$Cl$_2$ (5 mL) was added to the white slurry and the reaction mixture stirred at rt for 16 h. The reaction mixture was diluted with deionised water (10 mL), the organic product extracted into CH$_2$Cl$_2$ (3 x 10 mL), dried using Na$_2$SO$_4$ and concentrated in vacuo. The crude material was purified by flash column chromatography (ethyl acetate: n-hexanes, 1:1) to yield the desired product **2.179** as a light yellow oil as an epimeric mixture (24 mg, 0.042 mmol, 18% over three steps).

Rf 0.30 (ethyl acetate: n-hexanes, 1:1), the $^1$H and $^{13}$C data were in agreement with the literature values$^{84}$, HRMS (ESI) caled for C$_{32}$H$_{34}$N$_2$O$_7$Na$^+$ 581.2258, found 581.2258.
Chapter Seven

(2S,5R)-2-(Bicyclo[3.3.1]nonylboron)amino-4-(2-oxiranyl)-butanoate 2.181a and (2S,5S)-2-(Bicyclo[3.3.1]nonylboron)amino)-4-(2-oxiranyl)-butanoate 2.181b

3-Chloroperbenzoic acid (22 mg, 0.13 mmol, 2.5 eq.) was slowly added to an ice-cold solution of (S)-2-amino-5-hexenoic acid bicyclo[3.3.1]nonylboron (2.180, 11 mg, 0.05 mmol) in THF (dry, 2 mL) and the reaction mixture stirred at rt for 16 h. The reaction mixture was quenched using Na₂S₂O₃ (sat. soln, 2 mL) and the organic product was extracted into ethyl acetate (3 x 2 mL), washed with deionised water (2 mL), brine (sat. sol., 2 mL), dried using Na₂SO₄ and concentrated in vacuo to afford a complex mixture.

(2S,4R)-Methyl-2-((tert-butoxycarbonyl)amino)-4-(2-oxiranyl)butanoate 2.164a and (2S,4R)-Methyl-2-((tert-butoxycarbonyl)amino)-4-(2-oxiranyl)butanoate 2.164a

3-Chloroperbenzoic acid (70 mg, 0.41 mmol, 1 eq.) was slowly added to a solution of methyl (2S)-(N-tert-butoxycarbonyl)amino)hex-5-enoate (2.161, 100 mg, 0.41 mmol) in CH₂Cl₂ (dry, 5 mL) and the reaction mixture stirred at rt for 16 h. A further portion of 3-chloroperbenzoic acid (70 mg, 0.41 mmol, 1 eq.) was added, the reaction mixture stirred again at rt for 16 h, quenched using Na₂S₂O₃ (sat. soln., 1 mL). The organic product was extracted into CH₂Cl₂ (3 x 10 mL), dried using Na₂SO₄ and concentrated in vacuo. The crude material was purified by flash column chromatography (ethyl acetate: n-hexanes, 1:4 → 1:1 → 1:0) to yield the desired product 2.164 as an orange oil as an epimeric mixture (36 mg, 0.13 mmol, 34%).

Rᶠ 0.33 (ethyl acetate: n-hexanes, 1:4), δH (400 MHz; CDCl₃; Me₄Si) 1.44 (9 H, br s, Boc, 3 x CH₃), 1.48 - 1.86 (3 H, m, β-CH₂, γ-CH₂), 1.93 – 2.04 (1 H, m, β-CH₂, γ-CH₂) 2.47 – 2.50 (1 H, m, ε-CH₃), 2.74 – 2.78 (1 H, m, ε-CH₃), 2.90 – 2.95 (1 H, m, δ-CH), 3.74 (3 H, br s, OCH₃), 4.35 (1H, br s, α-CH),
5.12 (1H, dd, \(J_1 7.3, J_2 23.4, NH\)), \(\delta_C\) (400 MHz; CDCl\(_3\); Me\(_4\)Si) 28.3 (3 x CH\(_3\), 'Bu), 28.4, 29.1 (2 x CH\(_2\), \(\beta\)-CH\(_2\), \(\gamma\)-CH\(_2\)), 47.0, 47.1 (2 x CH\(_2\), \(\varepsilon\)-CH\(_2\)), 51.4, 51.5 (2 x CH, \(\delta\)-CH), 52.3, 52.3 (2 x CH\(_3\), OCH\(_3\)), 52.9, 53.1 (2x CH, \(\alpha\)-CH), 80.0 (C, 'Bu), 155.4, 172.9, 173.0 (C=O, Boc, \(CO_2Me\)), HRMS (ESI) calculated for C\(_{12}\)H\(_{21}\)NO\(_5\)Na\(^+\) 282.1312, found 282.1306.

\((2S,4R)-\text{Methyl}-2-(((9H-fluoren-9-yl)methoxy)amino)-4-(2-oxiranyl)butanoate 2.162a\) and \((2S,4R)-\text{Methyl}-2-(((9H-fluoren-9-yl)methoxy)amino)-4-(2-oxiranyl)butanoate 2.162b\)

\[
\begin{align*}
\text{2.162a} & \quad \text{NHFmoc} \quad \text{CO}_2\text{Me} \\
\text{2.162b} & \quad \text{NHFmoc} \quad \text{CO}_2\text{Me}
\end{align*}
\]

3-Chloroperbenzoic acid (184 mg, 1.1 mmol, 4 eq.) was added slowly to an ice-cold solution of methyl-(2S)-N-(9H-fluorenylmethoxycarbonyl)-2-amino-5-hexenoate (2.128, 100 mg, 0.27 mmol) in CH\(_2\)Cl\(_2\) (dry, 5 mL) and the reaction mixture stirred at rt for 16 h. The next day the reaction mixture was quenched using NaHCO\(_3\) (sat. soln., 2 mL) and the organic product was extracted into CH\(_2\)Cl\(_2\) (3 x 10 mL), washed with brine (10 mL), dried using Na\(_2\)SO\(_4\) and concentrated in vacuo to afford the crude material as a light yellow oil as an epimeric mixture (66 mg, 0.17 mmol, 64%). This compound was subjected to the subsequent reaction without further purification.

\((2S,4R)-\text{tert-Butyl}-2-(((9H-fluoren-9-yl)methoxy)amino)-4-(2-oxiranyl)butanoate 2.163a\) and \((2S,4R)-\text{tert-Butyl}-2-(((9H-fluoren-9-yl)methoxy)amino)-4-(2-oxiranyl)butanoate 2.163b\)

\[
\begin{align*}
\text{2.163a} & \quad \text{NHFmoc} \quad \text{CO}_2\text{Bu} \\
\text{2.163b} & \quad \text{NHFmoc} \quad \text{CO}_2\text{Bu}
\end{align*}
\]

3-Chloroperbenzoic acid (11 mg, 0.06 mmol, 1.1 eq.) was added to a solution of tert-butyl (2S)-N-(9H-fluorenylmethoxycarbonyl)-2-amino-5-hexenoate (2.160, 20 mg, 0.05 mmol) in CH\(_2\)Cl\(_2\) (dry, 5 mL) and the reaction mixture stirred at rt for 16 h. A further portion of 3-chloroperbenzoic acid (6 mg, 0.03 mmol, 0.5 eq.) was added, the reaction mixture stirred at rt for 8 h then quenched using Na\(_2\)S\(_2\)O\(_3\) (sat. soln., 2 mL). The organic product was extracted into CH\(_2\)Cl\(_2\) (3 x 10 mL), washed with brine (10 mL) dried using Na\(_2\)SO\(_4\) and concentrated in vacuo. The crude material was purified by flash column chromatography (ethyl acetate: \(n\)-hexanes, 1:4 → 1:1 → 1:0) to yield the desired product 2.163 as a colourless oil as an epimeric mixture (12 mg, 0.03 mmol, 62%).

R\(_f\) 0.31 (ethyl acetate: \(n\)-hexanes, 1:4), \(\delta_H\) (400 MHz; CDCl\(_3\); Me\(_4\)Si) 1.47 (9H, br s, Boc, 3 x CH\(_3\)), 1.54 - 1.88 (3 H, m, \(\beta\)-CH\(_2\), \(\gamma\)-CH\(_2\)), 1.96 – 2.08 (1 H, m, \(\beta\)-CH\(_2\), \(\gamma\)-CH\(_2\)), 2.49 (1 H, br s, \(\varepsilon\)-CH\(_3\)), 2.77
(1 H, br s, ε-CH₃), 2.93 (1 H, br s, δ-CH), 4.22 (1H, t, J6.9, CHFmoc), 4.31 (1H, br s, α-CH), 4.39 (2H, d, J 7.7, CH₂Fmoc), 5.39 (1H, dd, J₉ 9.1 J₂ 25.9, NH), 7.28 – 7.35 (2H, m, FmocPh), 7.40 (2H, t, J 7.4, FmocPh), 7.59 (2H, d, J 7.2, FmocPh), 7.70 (2H, d, J 7.3, FmocPh). δC (400 MHz; CDCl₃; Me₄Si) 28.8 (3 x C₃H₇, tBu), 28.2, 29.2 (2 x CH₂, β-CH₂γ-CH₂), 47.1, 47.2 (2 x CH₂, ε-CH₂), 51.5, 51.7 (2 x CH, δ-CH), 53.8, 54.1 (2x CH, α-CH), 66.9 (CH₂, CH₂Fmoc), 82.3 (C, tBu), 119.9, 125.0, 127.0, 127.7, 141.3, 143.9 (FmocPh). HRMS (ESI) calculated for C₂₅H₂₉NO₅Na⁺ 446.1938, found 446.1949.

(2S,5R)-Methyl-6-azido-2-((tert-butoxycarbonyl)amino)-5-hydroxyhexanoate 2.166a and (2S,5S)-Methyl-6-azido-2-((tert-butoxycarbonyl)amino)-5-hydroxyhexanoate 2.166b

Sodium azide (18 mg, 0.28 mmol, 2 eq.) was added to a solution of methyl-2-((tert-butoxycarbonyl)amino)-4-(2-oxiranyl)butanoate (2.165, 34 mg, 0.14 mmol) in DMF (dry, 2 mL) and the reaction mixture heated under reflux for 2 h. The organic product was then extracted into ethyl acetate (3 x 10 mL), copiously washed with deionised water (5 x 5 mL), dried using Na₂SO₄ and concentrated in vacuo to afford a complex mixture.

Minor peak found in HRMS (ESI) calculated for C₁₂H₂₂N₄O₅Na⁺ 325.1488, found 325.1480.

Penta-O-acetyl-D-galactopyranose¹²⁷,²⁷³,²⁷⁴ 2.184

Perchloric acid (60% solution, 120 µL, 3 µmol, cat.), was added dropwise to an ice-cold solution of D-galactose (0.5 g, 2.8 mmol) in acetic anhydride (20 mL, 180 mmol, excess). Further portions of D-galactose (4.5 g, 25 mmol) were then added and the reaction mixture stirred at rt for 2 h. The organic product was extracted into CH₂Cl₂ (2 x 50 mL), washed with NaHCO₃ soln. (sat. soln., 2 x 50 mL), dried using MgSO₄ and concentrated in vacuo to afford the title compound 2.184 as a colourless solid (10.9 g, 27 mmol, quant.).

Rf 0.38 (ethyl acetate: n-hexanes, 1:1), the ¹H and ¹³C data were in agreement with the literature values [α-anomer²⁷⁴, β-anomer²⁷³]. The α/β ratio was calculated from ¹H NMR spectrum to be 5.4/1,
mp 84.4 – 86.5 °C (lit. 88 -90 °C),\textsuperscript{275} HRMS (ESI) calculated for C\textsubscript{16}H\textsubscript{22}O\textsubscript{11}Na\textsuperscript{+} 413.1054, found 413.1070.

2,3,4,6-Tetra-\textit{O}-acetyl-\textalpha,\textbeta-D-galactopyranose\textsuperscript{129} 2.185

Glacial acetic acid (0.5 mL, 9.0 mmol, 1.4 eq.) was added dropwise to a solution of ethylenediamine (0.5 mL, 7.6 mmol, 1.1 eq.) in THF (10 mL). A solution of penta-\textit{O}-acetyl-D-galactopyranose (2.184, 2.5 g, 6.4 mmol) in THF (dry, 10 mL) was then added dropwise to the resulting white suspension. The reaction mixture stirred at rt for 16 h, quenched with deionised water (10 mL) and the organic product was extracted into CH\textsubscript{2}Cl\textsubscript{2} (3 x 10 mL), washed with HCl (2 M, 10 mL), NaHCO\textsubscript{3} soln.(sat. soln., 10 mL) and deionised water (10 mL), dried using Na\textsubscript{2}SO\textsubscript{4} and concentrated \textit{in vacuo}. The crude material was then purified by flash column chromatography (ethyl acetate: \textit{n}-hexanes, 2:3) to afford the title compound 2.185 as a colourless foam (1.5 g, 4.3 mmol, 70%).

R\textsubscript{f} 0.26 (ethyl acetate: \textit{n}-hexanes, 1:1), the \textsuperscript{1}H and \textsuperscript{13}C data were in agreement with the literature values.\textsuperscript{276} The \textalpha/\textbeta ratio was calculated from the \textsuperscript{1}H NMR spectra to be 1.56/1. HRMS (ESI) calculated for C\textsubscript{14}H\textsubscript{20}O\textsubscript{10}Na\textsuperscript{+} 371.0949, found 371.0942.

2,3,4,6-Tetra-\textit{O}-acetyl-\textalpha,\textbeta-D-galactopyranosyl trichloroacetimidate\textsuperscript{127} 2.183

K\textsubscript{2}CO\textsubscript{3} (1.20 g, 8.7 mmol, 3 eq.) followed by trichloroacetonitrile (1.2 mL, 10.0 mmol, 4 eq.) were added to 2,3,4,6-tetra-\textit{O}-acetyl-\textalpha,\textbeta-D-galactopyranose (2.185, 1.0 g, 2.90 mmol) in CH\textsubscript{2}Cl\textsubscript{2} (dry, 10 mL). The reaction mixture was stirred at rt for 16 h, filtered, concentrated \textit{in vacuo} and purified using flash column chromatography (ethyl acetate: \textit{n}-hexanes, 2:3) to afford the title compound 2.183 as a colourless foam (1.42 g, 2.90 mmol, quant.).

R\textsubscript{f} 0.46 (ethyl acetate: \textit{n}-hexanes, 1:1), the \textsuperscript{1}H and \textsuperscript{13}C data were in agreement with the literature values.\textsuperscript{277} The \textalpha/\textbeta ratio was calculated from \textsuperscript{1}H NMR spectra to be 0.58/1. HRMS (ESI) calculated for C\textsubscript{36}H\textsubscript{28}Cl\textsubscript{3}NO\textsubscript{10}Na\textsuperscript{+} 762.0671, found 762.0658.
2,3,4,6-Tetra-\(O\)-acetyl-\(\alpha,\beta\)-d-galactopyranosyl trichloroacetimidate (2.183, 177 mg, 1.19 mmol, 3 eq.) in \(\text{CH}_2\text{Cl}_2\) (dry, 1 mL) was added to a to a flame-dried flask containing molecular sieves (4Å, spatula tip) and the reaction mixture cooled to 0 °C. \(N^{\alpha}\)-(bicyclo[3.3.1]nonylboron)\(-N^{\varepsilon}\)-benzyloxycarbonyl-5-hydroxy-L-lysine (2.178, 50 mg, 0.12 mmol) in \(\text{CH}_2\text{Cl}_2\) (dry, 1 mL) followed by trimethylsilyltrifluoromethanesulfonate (98%, 0.01 mL, 0.5 eq.) were added to the ice-cold reaction mixture. The reaction mixture was stirred at rt for 45 min, quenched using triethylamine (40 μL, to reach pH 7), filtered and concentrated \textit{in vacuo}. Purification by flash column chromatography (ethyl acetate: \(n\)-hexanes: triethylamine, 1:1:0.02) was attempted, but resulted in a complex mixture of products. 

\(R_f\) 0.30 (spot mixture of stereoisomers, ethyl acetate: \(n\)-hexanes, 3:7), the \(^1\text{H}\) and \(^{13}\text{C}\) spectra for the complex mixture shows presence of known (5\(R\))-stereoisomer. HRMS (ESI) calculated for \(\text{C}_{36}\text{H}_{51}\text{BN}_2\text{O}_{14}\text{Na}\) 769.3332, found 769.3308 (within complex mixture).

\(N^{\alpha}\)-(9\(H\)-Fluoren-9-ylmethoxycarbonyl)\(-N^{\varepsilon}\)-benzyloxycarbonyl-5-\(O\)-(2,3,4,6-tetra-\(O\)-acetyl-\(\beta\)-D-galactopyranosyl)-5-hydroxy-L-lysine methyl ester, 2.191

2,3,4,6-Tetra-\(O\)-acetyl-\(\alpha,\beta\)-d-galactopyranosyl trichloroacetimidate (2.183, 12 mg, 0.11 mmol, 1.7 eq.) in \(\text{CH}_2\text{Cl}_2\) (dry, 1 mL) was added to a to a flame-dried flask containing molecular sieves (4 Å, spatula tip) and the reaction mixture cooled to 0 °C. \(N^{\alpha}\)-(9\(H\)-Fluoren-9-ylmethoxycarbonyl)\(-N^{\varepsilon}\)-benzyloxycarbonyl-5-hydroxy-L-lysine methyl-ester (2.129, 26 mg, 0.05 mmol) in \(\text{CH}_2\text{Cl}_2\) (dry, 1 mL) followed by trimethylsilyltrifluoromethanesulfonate (98%, 6 μL, 0.5 eq.) were added to the ice-cold reaction mixture. The reaction mixture was stirred at rt for 45 min, and a further portion of 2,3,4,6-tetra-\(O\)-acetyl-\(\alpha,\beta\)-d-galactopyranosyl trichloroacetimidate (2.183, 20 mg, 0.05 mmol, 1 eq.),
was added and the reaction mixture stirred at rt for a further 45 min. The reaction mixture was then quenched using triethylamine (to reach pH 7), filtered and concentrated \textit{in vacuo}. The organic product extracted into ethyl acetate (3 x 5 mL), dried using Na$_2$SO$_4$ and concentrated \textit{in vacuo}. Purification by flash column chromatography (ethyl acetate: \textit{n}-hexanes: triethylamine, 1:1:0.02) was attempted, but resulted in a complex mixture of products.

$R_f$ 0.26 (spot mixture of stereoisomers, ethyl acetate: \textit{n}-hexanes, 1:1), the $^1$H and $^{13}$C spectra for the complex mixture shows presence of (5R)-Fmoc-OMe-O-Gal-OAc$^{85}$. In the complex mixture, the following characteristic peaks were observed; $\delta$H (400 MHz; CDCl$_3$; Me$_4$Si) 3.08 (2 H, q, $J$ 7.4, $\varepsilon$-CH$_2$), 4.58 (1H, d, $J$ 7.0, H-1), 4.94 (1H, dd, $J_1$ 3.5, $J_2$ 10.5, H-3), 5.06 – 5.17 (2H, m, H-2, H-4). $\delta$C (400 MHz; CDCl$_3$; Me$_4$Si) 45.9 (CH$_2$, $\varepsilon$-CH$_2$), 101.6 (CH, H-1). HRMS (ESI) calculated for C$_{44}$H$_{50}$N$_2$O$_{16}$Na$^+$ 885.3052, found 885.3020 (within complex mixture).

$N^\alpha$-(9H-Fluorenlymethoxycarbonyl)-$N^\varepsilon$-benzyloxy carbonyl-5-O-(2,3,4,6-tetra-O-acetyl-$\beta$-D-galactopyranosyl)-5-hydroxy-L-lysine allyl ester$^{85}$, 2.192

2,3,4,6-Tetra-O-acetyl-$\alpha$,$\beta$-D-galactopyranosyl trichloroacetimidate (2.183, 52 mg, 0.11 mmol, 1.7 eq.) in CH$_2$Cl$_2$ (dry, 1 mL) was added to a to a flame-dried flask containing molecular sieves (4Å, spatula tip) and the reaction mixture cooled to 0 °C. $N^\alpha$-(9H-Fluorenlymethoxycarbonyl)-$N^\varepsilon$-benzyloxy carbonyl-5-hydroxy-L-lysine allyl-ester (2.179, 26 mg, 0.04 mmol) in CH$_2$Cl$_2$ (dry, 2 mL) followed by trimethylsilyltrifluoromethanesulphonate (1 mL, 0.3% w/v in dry CH$_2$Cl$_2$, 0.5 eq.) were added to the ice-cold reaction mixture. The reaction mixture was stirred at rt for 1 h, quenched using triethylamine (to reach pH 7), filtered and concentrated \textit{in vacuo}. The organic product extracted into ethyl acetate (3 x 5 mL), washed with NaHCO$_3$ (sat. soln. 5 mL), dried using Na$_2$SO$_4$ and concentrated \textit{in vacuo}. Purification by flash column chromatography (ethyl acetate: \textit{n}-hexanes: triethylamine, 1:1:0.02) was attempted, but resulted in a complex mixture of products.

$R_f$ 0.24 (spot mixture of stereoisomers, ethyl acetate: \textit{n}-hexanes, 1:1); HRMS (ESI) calculated for C$_{46}$H$_{52}$N$_2$O$_{16}$Na$^+$ 911.3317, found 911.3154 (within complex mixture).
(2S)-((9H-Fluorenlymethoxycarbonyl)amino)-6-azidohexanoic acid\textsuperscript{172} 3.64

Imidazole-1-sulfonyl azide hydrochloride (2.0 g, 9.0 mmol, 1.2 eq.) followed by K\textsubscript{2}CO\textsubscript{3} (2.4 g, 17.0 mmol, 2.1 eq.) were added to Fmoc-L-lysine (3.68, 3.0 g, 8.1 mmol) in MeOH (50 mL, analytical grade). Copper sulphate (CuSO\textsubscript{4}·5H\textsubscript{2}O, 20.0 mg, catalytic) and water (3 drops to aid dissolution) were then added to the reaction mixture, which was stirred at rt for 5 h, concentrated \textit{in vacuo} and acidified using HCl (2 M) to reach pH 2. The organic product was extracted into ethyl acetate (3 x 50 mL), washed with brine (50 mL), NaHCO\textsubscript{3} (sat. soln., 50 mL), dried using Na\textsubscript{2}SO\textsubscript{4} and concentrated \textit{in vacuo}. The crude product was purified by flash column chromatography (ethyl acetate: \textit{n}-hexanes, 3:7 → 1:0) to afford the title compound 3.64 as an off-white solid (2.5 g, 6.3 mmol, 78%).

R\textsubscript{f} 0.28 (ethyl acetate: \textit{n}-hexanes: acetic acid, 1:1:0.01), the \textsuperscript{1}H and \textsuperscript{13}C NMR data were in agreement with the literature values,\textsuperscript{142} [\textit{[\alpha]}\textsubscript{D}\textsuperscript{23} -2.3, \textit{c} 0.07, MeOH (lit [\textit{[\alpha]}\textsubscript{D}\textsuperscript{23} -1.8, \textit{c} 0.03, MeOH)],\textsuperscript{172} mp 64.1 - 64.8 °C (no lit. mp), HRMS (ESI) calculated for C\textsubscript{21}H\textsubscript{22}N\textsubscript{4}O\textsubscript{4}Na\textsuperscript{+} 417.1533, found 417.1536.

Boc-L-Serine\textsuperscript{278} 3.73

Sodium hydroxide (1 M, 28 mL, 3.8 mmol, 2 eq.) followed by di-\textit{tert}-butyl dicarbonate (4.2 g, 1.9 mmol, 1 eq.) were added to an ice-cold solution of L-serine (2.0 g, 1.9 mmol). The reaction mixture was stirred at rt for 16 h, concentrated \textit{in vacuo} and the remaining aqueous layer washed with ethyl acetate (to remove unreacted di-\textit{tert}-butyl dicarbonate). The aqueous layer was then acidified using HCl (1 M) to pH 2, saturated with NaCl ions. The organic product was then extracted into ethyl acetate (3 x 50 mL), dried using Na\textsubscript{2}SO\textsubscript{4} and concentrated \textit{in vacuo} to afford the title compound 3.73 as an off-white solid (0.9 g, 4.7 mmol, 25%). This product was then carried forward in the synthesis without further purification.

The \textsuperscript{1}H and \textsuperscript{13}C NMR data were in agreement with the literature values,\textsuperscript{279} [\textit{[\alpha]}\textsubscript{D}\textsuperscript{23} -1.9, \textit{c} 0.77, CHCl\textsubscript{3} (no lit. [\textit{[\alpha]}\textsubscript{D}]), mp 92.8 - 96.5 °C (lit. mp 90 °C).\textsuperscript{280}
(2S)-(tert-Butoxycarbonyloxy carbonyl)amino)-3-(prop-2-yn-1-yloxy)propanoic acid\(^{172}\) 3.74

![Chemical Structure](image)

Boc-L-serine (3.73, 6.0 g, 29.0 mmol) was added to an ice-cold solution of NaH (2.7 g, 0.1 mmol, 2.3 eq.) in DMF (30 mL) and the reaction mixture stirred for 30 min. Propargyl bromide (80 wt.\% in toluene, 4.7 mL, 0.05 mmol, 1.5 eq.) was added dropwise and the reaction mixture stirred at rt for 1.5 h. The reaction mixture was quenched with ice-cold deionised water (30 mL), concentrated in vacuum and acidified using HCl (1 M) to pH 2. The organic product was extracted into ethyl acetate (3 x 50 mL), dried using Na\(_2\)SO\(_4\) and concentrated in vacuo. The crude product was purified using flash column chromatography (n-hexanes: ethyl acetate: acetic acid, 5: 1: 0.1) to afford the title compound 3.74 as an orange oil (6.9 g, 28.5 mmol, 98%).

R\(_f\) 0.34 (ethyl acetate: acetic acid, 1: 0.01), the \(^1\)H and \(^{13}\)C NMR data were in agreement with the literature values,\(^{172}\) [\(\alpha\)]\(_D\){\(^{24}\)} +5.5, \(c\) 0.52, MeOH (lit. [\(\alpha\)]\(_D\){\(^{23}\)} +0.1, \(c\) 1.0, MeOH).\(^{172}\)

(2S)-Methyl-(tert-butoxycarbonyloxycarbonyl)amino)-3-(prop-2-yn-1-yloxy)propanoate\(^{281,282}\) 3.75

![Chemical Structure](image)

K\(_2\)CO\(_3\) (5.9 g, 42.6 mmol, 1.5 eq.) followed by MeI (2.2 mL, 42.6 mmol, 1.5 eq.) were added to a solution of (2S)-(tert-butoxycarbonyloxycarbonyl)amino)-3-(prop-2-yn-1-yloxy)propanoic acid (3.74, 6.9 g, 28.4 mmol) in DMF (40 mL). The reaction mixture was stirred overnight at room temperature and quenched with brine (50 mL). The organic product extracted with ethyl acetate (2 x 100 mL), washed with water (4 x 50 mL), dried using Na\(_2\)SO\(_4\) and concentrated in vacuo. The crude product was purified using flash column chromatography (ethyl acetate: n-hexanes, 1:4 → 1:1) to afford the title compound 3.75 as a yellow oil (5.5 g, 21.4 mmol, 75%).

R\(_f\) 0.65 (ethyl acetate: n-hexanes, 4: 1), the \(^1\)H and \(^{13}\)C NMR data were in agreement with the literature values,\(^{281}\) [\(\alpha\)]\(_D\){\(^{21}\)} +14.7, \(c\) 1.20, CHCl\(_3\) (lit. [\(\alpha\)]\(_D\){\(^{21}\)} +17.8, \(c\) 2.7, CHCl\(_3\)).\(^{283}\)
(2S)-Methyl-(9H-fluorenlymethoxycarbonyl)amino)-3-(prop-2-yn-1-yloxy)propanoate 3.76

TFA (40 mL) was added to an ice-cold solution of (2S) Methyl (tert butoxycarbonyloxy carbonyl) amino) - 3-(prop-2-yn-1-yloxy) propanoate (3.75, 7.7 g, 30.0 mmol) in CH₂Cl₂ (20 mL) and the reaction mixture stirred at rt for 2 h then concentrated in vacuo. The crude product was then dissolved in a mixture of dioxane: 10 % Na₂CO₃ solution (3:5, 80 mL) and cooled to 0 °C. A solution of N-(9H-fluorenlymethoxycarbonyl)succinimide (10.2 g, 30.0 mmol, 1 eq.) in dioxane (20 mL) was added dropwise to the ice-cold reaction mixture, which was then stirred for at rt for 16 h. The organic product was extracted into ethyl acetate (3 x 20 mL), washed with brine (sat. soln. 20 mL), dried using Na₂SO₄ and concentrated in vacuo. The crude product was purified by flash column chromatography (ethyl acetate: n-hexanes, 1:4 → 1:1), followed by recrystallisation from ethyl acetate: n-hexanes (1:2) to yield the title compound 3.76 as a colourless solid (8.3 g, 21.2 mmol, 73%).

Rf 0.28 (ethyl acetate: n-hexanes, 1: 4), δH (400 MHz; CDCl₃; Me₄Si) 2.44 (1 H, t, 4J₃,₁ 2.2, ≡C H), 3.72-3.77-4.01 (2 H, m, β-C H₂), 3.78 (3 H, s, OCH₃), 4.16 (2 H, d, 4J₁,₃ 2.2, CH₂C≡), 4.24 (1 H, t, JCH,CH₂ 7.3, CH Fmoc), 4.35-4.46 (2 H, m, CH₂ Fmoc), 4.53-4.57 (1 H, m, α-C H), 5.66 (1 H, d, JNH,α-H 8.3, NH), 7.32 and 7.39 (2 H, t, J 7.4 CH Fmoc Ar), 7.61 (2 H, m, CH Fmoc Ar), 7.76 (2 H, d, J 7.6, CH Fmoc Ar). δC (100 MHz, CDCl₃) 47.1 (CH, Fmoc-CH), 52.7 (CH₃, OCH₃), 54.2 (CH, α-C), 58.6, (CH₂, CH₂C≡), 67.2 (CH₂, CH₂ Fmoc), 69.6 (CH₂, β-C), 76.7 (CH, ≡CH), 78.8 (quat., CH₂C≡), 120.0, 125.1 and 127.1 (CH, Fmoc Ar), 127.7, 141.3 and 143.9 (quat., Fmoc Ar), 156.0 (C=O), 170.6 (C=O); [α]D 23 −1.7, c 0.70, MeOH (no lit [α]D), mp 83 – 85 °C (from ethyl acetate/ n-hexanes) (no lit. mp). ν max/cm⁻¹ 3402m (N-H), 3263m (C-H, alkyne), 2115w (C=C), 1756s and 1709s (C=O), 1528s (N-H), 1442m (C-C, aromatic), 1340s (C-H, aliphatic), 1203s, 1107s and 1075s (C-O), 1035s (C-N); HRMS (ESI) calculated for C₂₂H₂₁NO₂Na⁺ 402.1312, found 402.1308.
Chapter Seven

(2S)-(9H-Fluorenylmethoxycarbonyl)amino)-3-(prop-2-yn-1-xyloxy)propanoic acid 3.65

A solution of LiOH (0.1 M, 9.4 mL) was added dropwise to a solution of (2S)-methyl-(9H-fluorenylmethoxycarbonyl)amino)-3-(prop-2-yn-1-xyloxy)propanoate (3.76, 1.77 g, 4.8 mmol) in a mixture of 0.8 M CaCl₂: 0.1 M LiOH (70 mL: 9.36 mL, 10 eq.). The reaction mixture was stirred at rt for 2 h, quenched with ice-cold water (30 mL) then acidified using 5% citric acid solution to reach pH 4. The organic product was extracted into ethyl acetate (3 x 10 mL), washed with brine (10 mL), dried using Na₂SO₄ and concentrated in vacuo. The crude product was purified by flash column chromatography (ethyl acetate: n-hexanes 1:1 → 100% ethyl acetate + 2% acetic acid) to afford the title compound 3.65 as a colourless powder (1.6 g, 4.37 mmol, 97%).

Rf 0.14 (ethyl acetate: n-hexanes: acetic acid, 1: 0.02), the ¹H and ¹³C NMR data were in agreement with the literature values,¹⁷² [α]D 24 +6.5, c 0.62, MeOH (lit. [α]D 23 +6.2, c 1.00, MeOH),¹⁷² mp 144.6 - 145.2 °C (lit. mp not reported).

2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl bromide 2.60

AcO-Gal-Br

Hydrogen bromide (33% in acetic acid, 16 mL, 10 eq.) was added dropwise to an ice-cold solution of 1,2,3,4,6-penta-O-acetyl-α,β-D-galactopyranose (2.184, 10.8 g, 27.0 mmol) in CH₂Cl₂ (50 mL). The reaction mixture was stirred at rt for 4 h, quenched with ice-cold water (150 mL) and NaHCO₃ (sat. aq.) was added to reach pH 7. The organic product was extracted into CH₂Cl₂ (3 x 50 mL), washed with NaHCO₃ (sat. aq., 50 mL), dried using Na₂SO₄ and concentrated in vacuo. The crude product was purified by flash column chromatography (ethyl acetate: n-hexanes, 2:1) to afford the title compound 2.60 as a colourless foam (7.8 g, 19.0 mmol, 68%).

Rf 0.39 (ethyl acetate: n-hexanes, 1:2), the ¹H and ¹³C NMR data were in agreement with the literature values,¹⁷⁷ [α]D 21 +237.8, c 0.96 , CHCl₃ (lit. [α]D 21 +225.0, c 0.60, CH₃Cl)²⁸⁴, mp 76.7- 79.5 °C (lit. 86 - 89 °C).²⁸⁴
2,3,4,6-Tetra-O-acetyl-β-D-galactopyranosyl azide\textsuperscript{177} 3.67
AcO-Gal-N\textsubscript{3}

Tetrabutylammonium hydrogen sulfate (6.5 g, 19.0 mmol, 1 eq.) followed by NaHCO\textsubscript{3} (sat. aq., 70 mL) were added to a solution of 2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl bromide (2.60, 7.8 g, 19.0 mmol) in CH\textsubscript{2}Cl\textsubscript{2} (dry, 70 mL). Sodium azide (6.2 g, 95.0 mmol, 5 eq.) was then added, the reaction mixture stirred at rt for 2 d then concentrated \textit{in vacuo}. The organic product was extracted into ethyl acetate (3 x 50 mL), washed with NaHCO\textsubscript{3} (sat. aq., 2 x 10 mL), dried using Na\textsubscript{2}SO\textsubscript{4} and concentrated \textit{in vacuo}. The crude product was then recrystallised from hot ethanol (10 mL) to afford the title compound 3.67 as a colourless foam (4.0 g, 10.7 mmol, 56%).

R\textsubscript{f} 0.30 (ethyl acetate: n-hexanes, 1:2), the \textsuperscript{1}H and \textsuperscript{13}C NMR data were in agreement with the literature values,\textsuperscript{176} [α]\textsubscript{D}\textsuperscript{21} -14.7, c 2.35, CHCl\textsubscript{3} (lit. [α]\textsubscript{D}\textsuperscript{21} -10.3, c 1.00, CHCl\textsubscript{3}).\textsuperscript{176}

1-O-Propargyl-2,3,4,6-tetra-O-acetyl-β-D-galactopyranose\textsuperscript{178} 3.66
AcO-Gal-C1-O-alkyne

2,3,4,6-Tetra-O-acetyl-α,β-D-galactopyranosyl trichloroacetimidate (2.183, 200 mg, 0.4 mmol, 1.7 eq.) in CH\textsubscript{2}Cl\textsubscript{2} (dry, 2 mL) was added to a to a flame-dried flask containing molecular sieves (4Å, spatula tip) and the reaction mixture was stirred at rt for 30 min and then cooled to -40 °C. Propargyl alcohol (0.23 mL, 4.2 mmol, 10 eq) followed by trimethylsilyltrifluoromethanesulfonate (15 µL diluted in 0.2 mL dry CH\textsubscript{2}Cl\textsubscript{2}, 0.08 µmol, 0.2 eq.) were added dropwise to the ice-cold reaction mixture. The reaction mixture was stirred at rt for 30 min, and a further portion of trimethylsilyltrifluoromethanesulfonate (15 µL diluted in 0.2 mL dry CH\textsubscript{2}Cl\textsubscript{2}, 0.08 µmol, 0.2 eq.) was added and the reaction mixture stirred at rt for a further 15 min. The reaction mixture was then quenched using triethylamine (1 mL), filtered through Celite\textsuperscript{®} and concentrated \textit{in vacuo}. The organic product extracted into ethyl acetate (3 x 5 mL), dried using Na\textsubscript{2}SO\textsubscript{4} and concentrated \textit{in vacuo}. The crude product
was purified by flash column chromatography (ethyl acetate: \textit{n}-hexanes, 1:2) to afford the \textit{title compound} 3.66 as a pale yellow oil (0.14 g, 0.35 mmol, 87\%).

R\textsubscript{f} 0.61 (ethyl acetate: methanol, 1:1), the \textsuperscript{1}H and \textsuperscript{13}C NMR data were in agreement with the literature values,\cite{179} \([\alpha]_{D}^{21} -0.8, \text{ c } 1.09, \text{ MeOH (lit. } [\alpha]_{D}^{21} -5.3, \text{ c } 1.09, \text{ MeOH)}.\cite{179}

\textbf{1-O-Propargyl–β-D-galactosylpyranose}\cite{178} 3.62

\text{HO-Gal-C1}-O-alkyne

![Diagram of HO-Gal-C1-O-alkyne 3.62]

Freshly prepared sodium methoxide solution (1 M in MeOH) was added to 1-O-propargyl-2,3,4,6-tetra-O-acetyl-β-D-galactopyranose (3.66, 50 mg, 0.13 mmol) in methanol (AR, 3 mL) to reach pH 10. The reaction mixture was stirred at rt for 2 h, acidified using Dowex H\textsuperscript{+} beads to reach pH 2, filtered and concentrated \textit{in vacuo}. The crude yellow solid was purified using flash column chromatography (ethyl acetate: methanol, 10:1) to afford the \textit{title compound} 3.62 as a fine colourless powder (21 mg, 0.10 mmol, 75\%).

R\textsubscript{f} 0.64 (ethyl acetate: methanol, 1:1), the \textsuperscript{1}H and \textsuperscript{13}C NMR data were in agreement with the literature values,\cite{186} \([\alpha]_{D}^{26} -18.6, \text{ c } 0.61, \text{ CH}_{3}\text{OH; literature } [\alpha]_{D}^{26} -24.6, \text{ c } 0.85, \text{ CH}_{3}\text{OH}.\cite{187}

\textbf{1-O-Azido–β-D-galactosylpyranose}\cite{178} 3.63

\text{HO-Gal-N\textsubscript{3}}

![Diagram of HO-Gal-N\textsubscript{3} 3.63]

Freshly prepared sodium methoxide (1 M solution) was added to 2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl azide (3.67, 50 mg, 0.13 mmol) in methanol (3 mL, AR) to reach pH 10. The reaction mixture was stirred at rt for 2 h, acidified using Dowex H\textsuperscript{+} beads to reach pH 2, filtered and concentrated \textit{in vacuo}. The crude yellow solid was purified using flash column chromatography (ethyl acetate: methanol, 10:1) to afford the \textit{title compound} 3.63 as a fine colourless foam (13.8 mg, 0.07 mmol, 51\%).

R\textsubscript{f} 0.61 (ethyl acetate: methanol, 1:1), the \textsuperscript{1}H and \textsuperscript{13}C NMR data were in agreement with the literature values,\cite{177} \([\alpha]_{D}^{21} -0.8, \text{ c } 1.09, \text{ MeOH; (lit. } [\alpha]_{D}^{21} -5.3, \text{ c } 1.0, \text{ MeOH)}.\cite{176
(2S)-2-[(9H-Flourenyl-9-yl)-methoxy]carbonyl]amino-6-(((2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl(oxy)methyl)-1H-1,2,3-triazol-1-yl)hexanoic acid, 3.123

Fmoc-Lys-(Click-Gal-OAc)-OH

CuSO₄·5H₂O (83 mg, 0.33 mmol, 0.3 eq.) and sodium ascorbate (131 mg, 0.66 mmol, 0.6 eq.) in degassed, deionised water (5 mL) were added to 1-O-propargyl-2,3,4,6-tetra-O-acetyl-β-D-galactopyranose (3.66, 433 mg, 1.09 mmol, 1.1 eq.) in degassed ethanol (2 mL) and the reaction mixture heated at 80 °C for 15 min. (2S)-((9H-Fluorenylmethoxycarbonyl)amino)-6-azidohexanoic acid (3.64, 411 mg, 1.09 mmol, 1.0 eq.) was added to the reaction mixture, which was then heated at 80 °C for 1 h. The reaction mixture was concentrated in vacuo and the crude product was extracted into ethyl acetate (3 x 10 mL), washed with saturated ammonium chloride solution (10 mL), dried using Na₂SO₄ and concentrated in vacuo. The crude product was purified by flash column chromatography (ethyl acetate: n-hexanes: methanol, 1: 1: 0 → 1: 0: 0 → 1: 0: 0.05) to afford the title compound 3.123, as a colourless foam (803 mg, 1.09 mmol, quant.).

Rₕ 0.35 (ethyl acetate: methanol, 9:1), ν_max/cm⁻¹ 2922m and 2853 (C-H, aliphatic), 1744s (C=O), 1531w, 20.6, 20.7 20.8 (CH₃, 4 x OAc), 21.9 (CH₂-γ-C), 29.6 (CH₂-δ-C), 31.6 (CH₂-β-C), 47.1 (CH, Fmoc-CH), 50.1 (CH₂-ε-C), 52.0 (CH, α-C), 61.3 (CH₂, H₂-6), 62.7 (CH₂, β-CH₂OCH₂), 67.0 (CH, C-4), 67.1 (CH₂, Fmoc-CH), 68.0 (CH, C-2), 69.0 (CH, C-3), 70.7 (CH, C-5) 100.4 (C CH, C-1), 120.0 (CH, Fmoc, Ar), 122.9 (CH, C=CHN), 125.1, 127.1, 127.7 (CH, Fmoc Ar), 143.7 (quat., C=CHN), 143.8, 156.1 (quat.,
Chapter Seven

C=O Fmoc Ar), 170.1, 170.2, 170.3, 170.6 (quat., C=O, 4 x OAc – one peak overlapping), [α]D24 -4.9, c 0.30, MeOH, mp 64.1 – 64.8 °C (from ethyl acetate/ n-hexanes); HRMS (ESI+) for C38H44N4O14Na 803.2746, found 803.2757.

(2S)-2-[(9H-Flourenyl-9-yl)methoxy]carbonylamino]-3-(((2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)-1H-1,2,3-triazol-4-yl)methyl)oxy)propanoic acid, 3.124

Fmoc-Ser-(Click-Gal-OAc)-OH

![Chemical Structure of 3.124](image)

CuSO4.5H2O (410 mg, 1.64 mmol, 0.3 eq.) and sodium ascorbate (651 mg, 3.28 mmol, 0.6 eq.) in degassed, deionised water (5 mL) were added to (2S)-((9H-fluorenylmethoxycarbonyl)amino)-3-(prop-2-yn-1-yl)oxy)propanoic acid (3.65, 2.00 g, 5.47 mmol) in degassed ethanol (2 mL) and the reaction mixture was heated at 80 °C for 10 min. 2,3,4,6-Tetra-O-acetyl-α-D-galactopyranosyl azide (3.67, 2.04 mg, 5.47 mmol, 1.0 eq.) was added to the reaction mixture, which was then heated at 50 °C for 1 h. The reaction mixture was concentrated in vacuo and the crude product extracted into ethyl acetate (3 x 10 mL), washed with saturated ammonium chloride solution (10 mL), dried using Na2SO4 and concentrated in vacuo. The crude product was purified by flash column chromatography (ethyl acetate: n-hexanes: methanol, 1: 1: 0 → 1: 0: 0 → 1: 0: 0.05) to afford the title compound 3.124 as a colourless foam (4.03 g, 5.47 mmol, quant.).

Rf 0.32 (ethyl acetate: methanol, 9:1), νmax/cm⁻¹: 3407w (N-H), 2928w (C-H, aliphatic), 1737s (C=O), 1515m and 1450m (C=C, aromatic), 1368m, 1206s (C-O), 1154m, 1107m, 1043s (C-O).

δH (400 MHz; CDCl3; Me4Si) 1.88, 1.99, 2.03, 2.21 (each 3 H, s, OAc), 3.82 - 4.00 (2 H, dd, Jα,β 64.2, Jβαββ 6.9, β-CH2), 4.13 - 4.25 (4 H, m, CH Fmoc, H-5, CH2-6), 4.30 - 4.45 (2 H, m, CH2 Fmoc), 4.56 (1 H, brs, α-CH), 4.63 - 4.78 (2 H, dd, J 42.8, J 12.8, β-CH2OCH2), 5.23 - 528 (1 H, dd, J 10.2, J 3.4, H-3), 5.49 - 5.55 (2H, m, H-2 and H-4), 5.80 - 5.83 (1 H, d, J 9.2, Gal-C1-H), 5.89-5.92 (1 H, d, J 8.1, N-H), 7.28 - 7.32 (2H, m, CH Fmoc Ar), 7.37 - 7.40 (2 H, t, J 7.4, CH Fmoc Ar), 7.60-7.64 (2 H, br t, J 7.5, CH Fmoc Ar), 7.75 - 7.76 (2 H, d, J 7.5, CH Fmoc Ar), 7.82 (1 H, s, C=CHN). δC (400 MHz; CDCl3; Me4Si) 20.2, 20.5, 20.6, 21.0 (CH3, 4 x OAc), 47.1 (CH, Fmoc-CH), 54.2 (CH, α-CH), 61.3 (CH2, C-6), 64.2 (CH2, β-CH2OCH2), 66.9 (CH, C-2 or C-4), 67.3 (CH2, Fmoc-CH2) 68.0 (CH, C-2 or C-4), 69.8 (CH2, β-CH2), 264
Chapter Seven

70.6 (CH, C-3), 74.2 (CH, C-5), 86.3 (CH, Gal-C1H), 119.9 (CH, Fmoc, Ar), 121.5 (CH, C=CHN), 125.3, 127.1, 127.7 (CH, Fmoc, Ar), 141.2 (quat., C=CHN), 143.7, 143.9 (quat., Fmoc Ar), 156.5 (quat. C=O, Fmoc), 169.5, 169.8, 170.0, 170.7, 171.2 (quat. C=O, OAc and CO₂H). mp 80.2 – 81.8 °C (from ethyl acetate/ n-hexanes); [α]_D^24 -3.3, c 1.30, MeOH, HRMS (ESI+) for C₃₅H₃₈N₄O₁₄Na calcd 761.2277, found 761.2268.
1.2 Section B: Details for peptide synthesis

General synthetic details

All solvents and reagents were used as supplied. 1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU), 1H-Benzotriazolium 1-[bis(dimethylamino)methylene]-5chloro-hexafluorophosphate (1-),3-oxide (HCTU) and Fmoc-Rink amide linker were purchased from GL Biochem (Shanghai, China). Dimethylformamide (DMF) (AR grade) and acetonitrile (HPLC grade) were purchased from Scharlau (Barcelona, Spain). N,N'-diisopropylethylamine (DIPEA) and diisopropylcarbodiimide (DIC) were purchased from Aldrich (St Louis, MO). Tris(2-carboxyethyl)phosphine hydrochloride (TCEP.HCl) and triisopropylsilane (TIPS) were purchased from Alfa Aesar (Wardhill, MA). Guanidine hydrochloride (Gn.HCl) was purchased from MP Biomedicals (Auckland, New Zealand). 4-[(R,S)-α-[1-(9H-fluoren-9-yl) methoxycarbonylamino]-2,4-dimethoxy]phenoxyacetic acid (Rink amide) linker was purchased from GL Biochem. Trifluoroacetic acid (TFA) was purchased from Oakwood Chemicals (River Edge, NJ). Hydroxymethylphenoxy propionic acid (HMPP) linker was purchased from Novabiochem(R) (Billerica, MA). Aminomethyl polystyrene (AM-PS) resin was synthesised “in house” according to the method previously published by Brimble et al.\textsuperscript{197} Fmoc-amino acids were purchased from GL Biochem with the following side chain protections: Fmoc-Pro-OH, Fmoc-Gly-OH, Fmoc-Lys(Boc)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Glu(OtBu). Fmoc-Hyp(4R)-OH was purchased from Sigma-Aldrich. Unless otherwise noted, all reactions were performed under an oxygen-free atmosphere of nitrogen using standard peptide techniques.

All peptides were synthesised using solid phase peptide synthesis based on the Fmoc protection strategy on a 0.1 mmol scale unless otherwise stated. Aminomethyl polystyrene resin (AM-PS, 0.98 g/ mmol loading) was employed for all peptide synthesis functionalised using either Rink amide linker (for amide C-terminal peptides) or HMPP linker (for acid C-terminal peptides). Unless otherwise noted, purified peptides were >95% pure by HPLC analysis.

Circular dichroism spectra were recorded using Applied Photophysics PiStar Spectrometer v.4.2.12 using a cuvette of path length 10 mm (106-QS, Hellma Analytics, Müllheim, Germany). The peptide solutions used were 0.2 mM (by weight using a 4.d.p. balance, Sartorius Analytic, A2005) in 10 mM potassium phosphate buffer at pH 7.4, which had been incubated at 6 °C for a minimum of 15 hours. The spectra were obtained at 6 °C between wavelengths of 180 and 310 nm with a bandwidth of 2 nm, steps of 1 nm, with a 3 s averaging time and averaged from 5 scans. The baseline scans were recorded and averaged using 10 mM potassium phosphate buffer and were subtracted from sample scans and mean molar ellipticity was calculated using Pro-Data Viewer (version 4.1.1, Applied Photophysics Ltd.,
UK). The thermal transition experiments recorded were run from 15 to 93 °C in 0.5 °C steps with a smooth ramp rate and set at a wavelength of 215 nM.

Due to a severe allergy against coupling agents such as EDCI and DIC developed during the course of this doctoral study, Fmoc-SPPS was conducted and optimised by senior research fellows within the Brimble group. However, peptide purification using RP-HPLC and subsequent chemical reactions (such as on-peptide CuAAC ‘click’ reactions and deacetylation reactions) and structural biological investigations using circular dichroism techniques were conducted solely by the author of this thesis.

**General Peptide Methods**

**General Method A (‘on-resin’ CuAAC ‘click’ procedure)**

Tris(2-carboxyethyl)phosphine hydrochloride (TCEP.HCl, 22.9 mg, 80.0 mmol, 150 eq.) and CuSO$_4$.5H$_2$O (19.9 mg, 80.0 mmol, 150 eq.) were added to a degassed solution (0.5 mL) of either 6 M GnHCl/ 0.2M Na$_2$HPO$_4$ or 1:1, v/v DMSO/ DI H$_2$O and the pH of the resulting solution adjusted to pH 7.1 using 1 M NaOH (1 M, aq.). The reaction vessel was then gently heated (in order to activate the catalyst), centrifuged for 1 min at 12,000 rpm and the resultant light blue solution added to a solid mixture of peptide (1 mg, 0.53 mmol, 1 eq.) and sugar ‘click’ partner (1.7 mg, 8 mmol, 15 eq.). The reaction vessel was then gently heated (in order to activate the catalyst), centrifuged for 1 min at 12,000 rpm and the resultant light blue solution added to a solid mixture of peptide (1 mg, 0.53 mmol, 1 eq.) and sugar ‘click’ partner (1.7 mg, 8 mmol, 15 eq.). The reaction was heated under microwave irradiation (25 W, 80 °C, 3 h) using Biotage® Initiator microwave (Biotage Sweden AB) under an argon atmosphere. The reaction progress was monitored by analytical reverse phase-high performance liquid (RP-HPLC), [Gemini 5 µm C18 110 Å, 150 mm x 4.6 mm, 1 mL/min, linear gradient 5% B – 40% B over 35 min].

**General Method B (attachment of Fmoc-Ile-HMPP linker to AM-PS resin)**

The resin was swollen in CH$_2$Cl$_2$ (2 mL), to which Fmoc-Ile-HMPP-linker (2 eq.) in CH$_2$Cl$_2$ (1 mL) and DIC (2 eq.) were added. The resulting peptidyl-resin was then shaken at rt for 2 h, after which point the Kaiser test gave a negative result. The resin was dried by washing with methanol and a stream of N$_2$.

**General Method C (attachment of Rink amide linker to AM-PS resin)**

The resin was swollen in CH$_2$Cl$_2$ (2 mL). In a separate vial, DIC was added (2 eq.) was added to a slurried mixture of Rink amide linker (2 eq.) and 6-Cl-HOBt (2 eq.) in DMF (1 mL) and the slurry was vortex-mixed until a clear solution was formed. This clear solution was added to the swollen resin, stirred for 3 h, drained and washed (2 x CH$_2$Cl$_2$, 1 x DMF). The resulting peptidyl-resin was allowed to sit at rt for 1.5 h, at which point the Kaiser test gave a negative result. The resin was then dried by washing with methanol and a stream of N$_2$. 
General Method D (Fmoc-SPPS utilising Liberty 12 Microwave Peptide Synthesiser and piperazine for Fmoc deprotection step)

The dried peptidyl-resin from General method B was transferred to a Liberty 12 Microwave Peptide Synthesiser (CEM Corporation, Mathews, NC). The following coupling cycle was performed for Fmoc-amino acid residues:

*Standard coupling cycle: Deprotection:* 5% Piperazine + 0.1% 6-Cl-HOBt in DMF, microwave (60 W, 75 °C, 180 s); *Coupling:* HATU (1 mL, 0.45 M in DMF, 4.5 eq.) in DMF, DIPEA (0.5 mL, 2 M in NMP, 10 eq.), Fmoc-amino acid (0.5 mmol, 5 eq.), microwave (25 W, 73 °C, 5 min).

For the derivatised Fmoc-amino acids 3.123 and 3.124 the following procedure was used:

Derivatised Fmoc-amino acid (2 eq.), HATU (0.5 M, 1.95 eq.), DIPEA (4 eq.) were dissolved in CH₂Cl₂:DMF (1:1, v/v, 4 mL) and added manually to the microwave cavity. The coupling was conducted as described above for a standard amino acid but using a microwave time of 20 min.

General Method E (Fmoc-SPPS, Liberty™ microwave synthesiser and piperidine for Fmoc deprotection step)

The dried peptidyl-resin from General Method C was transferred to the Liberty 12 Microwave Peptide Synthesiser (CEM Corporation, Mathews, NC) and coupling cycles performed utilising the following conditions for all amino acids:

*Single coupling cycle: Fmoc deprotection:* 40% piperidine/ DMF (3 mL, 2 x 3 min, 75 °C, 25 W), *Resin wash:* DMF (7 mL, 3 x), *Coupling:* HCTU (0.45 M, 4.5 eq.), DIPEA (2 M in NMP, 10 eq.) and Fmoc-amino acid (2.5 mL, 0.2 M in DMF, 5 eq., 5 min, 75 °C, 25 W), *Resin wash:* DMF (3 x 7 mL).

General Method F (single coupling cycle, Tribute™ synthesiser)

The dried peptidyl-resin from General method C was transferred to the Tribute™ synthesiser (Protein Technologies, Tuscon, Az) and the following couple cycle performed utilising the following coupling conditions:

*Single coupling cycle: Fmoc deprotection:* 20% piperidine/ DMF (3 mL, 2 x 5 min), *Resin wash:* DMF (2 mL, 5 x 30 s), *Coupling:* Either HCTU or HATU (2 mL, 0.23 M in DMF, 4.6 eq.) and Fmoc-amino acid (0.5 mmol, 5 eq., N₂, 2 min), DIPEA (0.5 mL, 2 M in NMP, 10 eq, N₂, 45 min); *Resin wash:* DMF (2mL, 2 x 30 s).
General Method G (single coupling cycle, Tribute™ synthesiser using stronger HATU coupling agent)

The dried peptidyl-resin from General method C was transferred to the Tribute™ synthesiser (Protein Technologies, Tuscon, Az) and coupling cycles performed utilising the following conditions:

Single coupling cycle: Fmoc deprotection: 20% piperidine/ DMF (2 mL, 2 x 7 min), Coupling: HATU (2 mL, 0.23 M in DMF, 1.9 eq.), Fmoc-clicked amino acid (2 eq.) and DIPEA/ DMF (0.5 mL, 4 eq.), slurry agitated for 30 min at rt; Resin wash: DMF (2 x 2 mL).

General Method H (double coupling cycle, Tribute™ synthesiser)

The dried peptidyl-resin from General method C was transferred to the Tribute™ synthesiser (Protein Technologies, Tuscon, Az) and coupling cycles performed utilising the following conditions:

Double coupling cycle: As per General method F with coupling cycles performed twice to improve yields.

General Method I (double coupling cycle, Tribute™ synthesiser, increased coupling agent eq.)

The dried peptidyl-resin from General method C was transferred to Tribute™ synthesiser (Protein Technologies, Tuscon, Az) and coupling cycles performed utilising the following conditions:

Double coupling cycle: Deprotection: 20% piperidine/ DMF (3 mL, 2 x 5 min), Resin wash: DMF (2 mL, 5 x 30 s), Coupling: HCTU (2 mL, 0.23 M in DMF, 7 eq., 2 min, performed bubbling under N₂), Base: N-methylmorpholine (NMM, 5.1 M, 10 eq.) in N-methyl-2-pyrrolidinone (NMP, 0.5 mL), and amino acid (10 eq.) (20 min, performed bubbling under N₂); Resin wash: DMF (2 mL, 2 x 30 s).

Second coupling: Base: in N-methylmorpholine (NMM, 5.1 M, 10 eq.) in N-methyl-2-pyrrolidinone (NMP, 0.5 mL), Coupling: HCTU (1 mL, 0.23 M in DMF, 7 eq., 2 min, performed bubbling under N₂), and amino acid (10 eq.) (20 min, performed bubbling under N₂); Resin wash: DMF (2 mL, 2 x 30 s).
General Method J (single coupling cycle, CEM™ microwave)

The dried peptidyl-resin from General method C was transferred to the CEM™ microwave synthesiser (CEM Corporation, Mathews, NC) and coupling cycles performed utilising the following conditions:

- **Fmoc deprotection**: 20% piperidine/ DMF (75 °C, 50 W, 1.5 mL, 2 x 3 min), **Resin wash**: DMF (2 mL, 5 x), **Coupling**: Fmoc-amino acid (5 eq.) with HATU (4.6 eq.), DIPEA (10 eq.) in DMF (1.5 mL) (75 °C, 25 W, 5 min), **Resin wash**: DMF (2 mL, 5 x).

General Method K (acetyl capping of N-terminus using acetic anhydride, Tribute™ synthesiser)

The peptidyl resin was dried by washing with methanol and a stream of N₂. The dried peptidyl resin was then transferred to the Tribute™ synthesiser (Protein Technologies, Tuscon, Az) and the following acetyl capping cycle performed:

- **Fmoc deprotection**: 20% piperidine/ DMF (2 mL, 2 x 7 min), **Resin wash**: DMF (2 mL, 6 x 30 s); **Acetylation**: acetic anhydride (2 mmol, 40 eq.), DIPEA in NMP (0.5 mL, 2 M, 10 eq., N₂, 2 x 20 min); **Resin wash**: DMF (2mL, 2 x 30 s).

General Method L (acetyl capping of N-terminus using acetic anhydride/ DIPEA, rotary mixing)

The peptidyl resin was dried by washing with methanol and a stream of N₂. The dried peptidyl resin was then transferred to a round bottomed flask and rotary mixed with a solution of acetic anhydride/ DIPEA/ DMF (20% v/v, 20 eq.) for 30 min at rt. The acetylated peptidyl resin was then washed with DMF (5mL, 3 x).

General Method M (resin cleavage using microwave irradiation)

The peptidyl-resin was dried (as previously described above) and cleaved from resin using the cleavage cocktail (5 mL, 95 % TFA, 2.5 % TIPS, 2.5 % DI H₂O, v/v/v) in the microwave (10 W, 35 °C, 20 min). The peptide was precipitated using cold diethyl ether (ice bath, 20 min) and the peptide recovered by centrifugation (5 min at 3,000 rpm). A further portion of diethyl ether (2 mL) was added, the solution shaken and the mixture centrifuged. The ether was removed by decanting and the resulting precipitate was dissolved in 50% aq. CH₃CN: containing 0.1% TFA) and lyophilised to yield crude peptide.
General Method N (resin cleavage method 1 using agitation)

The cleavage cocktail (2 mL, 95% v/v TFA, 2.5% TIPS, 2.5% DI H₂O, v/v/v) was added to the peptidyl resin and agitated for 90 min at rt. The resin was then drained, washed with TFA (2 mL x 2) and concentrated under a stream of nitrogen. The peptide was precipitated from the reaction slurry using cold diethyl ether (20 mL, 4 °C, 20 min) and recovered by centrifugation (5 minutes at 3,000 rpm). The peptide precipitate was then washed with diethyl ether (cold, 10 mL), centrifuged (5 min at 3,000 rpm) dried under a stream of stream of N₂, dissolved in CH₃CN: H₂O (1:1 containing 0.1% TFA) and lyophilised to yield the crude peptide.

General Method O (resin cleavage method 2 using agitation)

Following General method N but using the following cleavage cocktail (2 mL, 90% v/v TFA, 5% TIPS, 5% DI H₂O).

General Method P (resin cleavage method 3 using agitation)

Following General method N but using the following cleavage cocktail (2 mL, 93% v/v TFA, 3.5% TIPS, 3.5% DI H₂O).

General method Q (deacetylation)

Freshly prepared 1 M NaOMe solution in methanol was added to a solution of the crude peptide in methanol (AR grade) until the pH of the reaction mixture was pH 11. The reaction mixture was stirred at rt for 4 h after which the pH was measured to be 7. Therefore, a further portion of NaOMe solution (1 M) was added to obtain pH 11, and the reaction mixture was stirred at rt for 16 h, neutralised using Dowex H⁺ beads, filtered and concentrated in vacuo. The pH of the crude peptide solution was adjusted to pH 2 using HCl (2 M) and the product recovered by centrifugation (10 min at 4,000 rpm).

RP-HPLC

Unless otherwise stated, semi-preparative RP-HPLC was performed on a Dionex Ultimate 3000 equipped with a 4 channel UV detector, using a Phenomenex Luna C18 column (5 µm, 10 mm x 250 mm) at a flow rate of 5 mL/min.

Unless otherwise stated, analytical RP-HPLC was performed on a Dionex Ultimate 3000 equipped with a 4 channel UV detector, using a Phenomenex Gemini C18 column (5 µm, 150 mm x 4.6 mm) using buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN at a flow rate of 1 mL/min.
Family 1 18-mer Adpn peptides

Adpn-Lys-azide 3.60

H-GEK$^{N3}$GE K$^{N3}$GDOGLIGO K$^{N3}$GDI-OH

Where K$^{N3} = \text{HN} \text{OHN}$

The peptide was synthesised on 0.1 mmol scale according to General Method B on a using aminomethyl polystyrene resin (AM-PS, 0.98 g/ mmol loading) functionalised with HMPP linker followed by General Method D and then cleaved from resin using General Method M.

The crude peptide was purified using RP-HPLC on a Waters 600 system using an Xterra™ column (19 mm x 300 mm) using a gradient of 1 → 27% B over 10 min (ca. 2.7% B per min) followed by 27 → 38% B for 45 min (0.25% B per min) to afford the title compound 3.60 as a colourless powder (16.7 mg, 0.00902 mmol, 9% yield).

R$_t$ 15.8 min [Agilent C3 SB-300, 150 mm x 3.0 mm, 0.3 mL/min, linear gradient 5% B – 65% B over 21 min, where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+H]$^+$ 1876.6, calculated 1877.9; [M+2H]$^{2+}$ 939.2, calculated 939.6.

Figure 7.1 HPLC trace and low resolution mass spectrometry trace for Adp-Lys-azide 3.62.
Adp-Ser-alkyne 3.61

H-GES*GES*GDOGLIGOS*GDI-OH

Where $S^\equiv = \equiv$ and $O = (2S,4R)-4$-Hydroxyproline

The peptide was synthesised on 0.1 mmol scale according to **General Method B** on a using aminomethyl polystyrene resin (AM-PS, 0.98 g/ mmol loading) functionalised with HMPP linker followed by **General Method D** and then cleaved from resin using **General Method M**.

In brief, Fmoc deprotection and coupling reactions were carried out with microwave-irradiation using an automated Liberty 12 Microwave Peptide Synthesiser (CEM Corporation, Mathews, NC) and with HATU/DIPEA and piperazine as the coupling and Fmoc deprotection reagents, respectively. For the coupling of building block Fmoc-Ser-alkyne 3.65 a longer coupling time of 20 minutes was employed, rather than the usual coupling time of 5 minutes. On completion of synthesis, cleavage of the peptidyl-resin with TFA, TIPS (triisopropylsilane), $H_2O$ (95: 2.5: 2.5, v/v) afforded the crude peptide. The crude peptide was purified using RP-HPLC on a Waters 600 System using an Xterra™ column (19 mm x 300 mm) using a gradient of 1 → 20% B over 14 min (ca. 1.4% B per min) followed by 20 → 50% B over 150 min (0.2% B per min) to afford the **title compound 3.61** as a colourless powder (18.6 mg, 0.00639 mmol, 6.4% yield).

$R_t$ 13.8 min [Agilent C3 SB-300, 150 mm x 3.0 mm, 0.3 mL/min, linear gradient 5% B – 65% B over 21 min, where buffer A = 0.1% TFA in DI $H_2O$, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) $[M+H]^+$ 1790.4, calculated 1790.6; $[M+2H]^{2+}$ 895.5, calculated 895.8.

![Figure 7.2 HPLC trace and low resolution mass spectrometry trace for Adp-Ser-alkyne 3.63.](image-url)
Adpn-Lys-3Gal-OAc 3.58a

H-GE\textsubscript{1}GE\textsubscript{1}GDOGLIGO\textsubscript{1}GDI-OH

Where \( X_1 = \)

The peptide was synthesised on a 0.1 mmol scale according to General Method B followed by General Method D except that the Wang linker was used instead of HMPP linker. Upon completion of the peptide sequence, the peptide was cleaved from resin according to General Method M. The crude peptide was purified by using RP-HPLC on a Dionex Ultimate 3000 with a Phenomenex Luna C18 column (10 mm x 250 mm) using a gradient of 0 → 27% B over 13 min (2% B per min) followed by 27 → 39% B over 23 min (c. 0.5% B per min) to afford the title compound 3.58a as a colourless powder (3.1 mg, 0.00102 mmol, 1.0% yield).

\( R_t \) 23.8 min [Gemini 5 µm C18 110 Å, 150 mm x 4.6 mm, 1 mL/min, linear gradient 5% B – 40% B over 35 min, where buffer A = 0.1% TFA in DI H\textsubscript{2}O, Buffer B = 0.1% TFA in CH\textsubscript{3}CN]; LRMS (ESI+) \([M+2H]^2^+\) 1518.4, calculated 1518.5.

Figure 7.3 HPLC trace for Adpn-Lys-3Gal-OAc 3.58a.
Adpn-Ser-3Gal-OAc 3.59a

H-GE\(\text{X}_2\)GE X\(\text{GDOGLIGO}\)X\(\text{GDI-OH}\)

Where \(\text{X}_2 =\)

The peptide was synthesised on a 0.1 mmol scale according to General Method B followed by General Method D except that the Wang linker was used instead of HMPP linker. Upon competition of the peptide sequence, the peptide was cleaved from resin according to General Method M except that the cleavage cocktail TFA: DODT (2,2′-(Ethyleneoxy)diethanethiol): H\(_2\)O: TIS. (94:2.5:2.5:1.5) was used. The crude peptide was purified using RP-HPLC on a Dionex Ultimate 3000 with a Phenomenex Luna C18 column (10 mm x 250 mm) using a gradient of 0 → 30\% B over 15 min (2\% B per min) followed by 30 → 40\% B over 19 min (c. 0.5\% B per min) to afford the title compound 3.59a as a white powder (4 mg, 0.00137 mmol, 1.4\% yield).

\(R_t\) 22.0 min [Gemini 5 \(\mu\)m C18 110 Å, 150 mm x 4.6 mm, 1 mL/min, linear gradient 5\% B – 40 \% B over 35 min, where buffer A = 0.1\% TFA in DI H\(_2\)O, Buffer B = 0.1\% TFA in CH\(_3\)CN]; LRMS (ESI+) \([\text{M+2H}]^{2+}\) 1455.5, calculated 1455.9; \([\text{M+3H}]^{3+}\) 970.6, calculated 970.9.

**Figure 7.4** HPLC trace and low resolution mass spectrometry trace for Adpn-Ser-3Gal-OAc 3.59a.
Adpn-Lys-3Gal-OH 3.58b

H-GE\textsubscript{3}GEX\textsubscript{3}GDOGLIGO\textsubscript{3}GDI-OH

Where $X_3 =$

Crude Adpn-Lys-3Gal-OAc 3.58a (107 mg) was deacetylated according to General Method Q and then purified using RP-HPLC on a Dionex Ulitmate 3000 using a Phenomenex Luna C18 column (10 mm x 250 mm) using a gradient of 0 → 10% B over 10 min (1% B per min) followed by 10 → 19% B over 70 min (c. 0.1% B per min) to afford the title compound 3.58b as a colourless powder (26.7 mg, 0.01066 mmol, 10.6% overall yield including synthesis of actetylated peptide).

$R_t$ 14.0 min [Gemini 5 µm C18 110 Å, 150 mm x 4.6 mm, 1 mL/min, linear gradient 5% B - 40% B over 35 min, where buffer A = 0.1% TFA in DI H\textsubscript{2}O, Buffer B = 0.1% TFA in CH\textsubscript{3}CN]; LRMS (ESI+) [M+2H]\textsuperscript{2+} 1265.6, calculated 1266.5; [M+3H]\textsuperscript{3+} 844.4, calculated 844.7.

Figure 7.5 HPLC trace and low resolution mass spectrometry trace for Adpn-Lys-3Gal-OH 3.58b.
Adpn-Ser-3Gal-OH 3.59b

H-GE\textsubscript{4}GEX\textsubscript{4}GDOGLIGO\textsubscript{4}GDI-OH

Where \( X_4 = \)

Crude Adpn-Ser-3Gal-OAc, 3.59a (93 mg) was deacetylated according to **General Method Q** and then purified using RP-HPLC on a Dionex Ulitmate 3000 with a Phenomenex Luna C18 column (10 mm x 250 mm) using a gradient of 5 → 10% B over 5 min (1% B per min) followed by 10 → 19% B over 70 min (c. 0.1% B per min) to afford the **title compound 3.59b** as a colourless powder (21 mg, 0.00873 mmol, 8.6% overall yield including synthesis of acetylated peptide).

\( R_t = 13.9 \text{ min} \) [Gemini 5 \textmu m C18 110 Å, 150 mm x 4.6 mm, 1 mL/min, linear gradient 5% B – 40% B over 35 min, where buffer A = 0.1% TFA in DI H\textsubscript{2}O, Buffer B = 0.1% TFA in CH\textsubscript{3}CN]; LRMS (ESI+) \([M+2H]\textsuperscript{2+} 1203.1, \text{ calculated} 1203.4; [M+3H]\textsuperscript{3+} 802.5, \text{ calculated} 802.6

---

**Figure 7.6** HPLC trace and low resolution mass spectrometry trace for Adpn-Ser-3Gal-OH, 3.59b.
Chapter Seven

Family 2 CMP peptides

CMP-Lys-1Gal-OAc 4.20a

Ac-(POG)$_3$X$_1$EG(POG)$_3$-NH$_2$

Where X$_1$ =

The synthesis was carried on a 0.1 mmol scale according to General method E, except for Fmoc-clicked amino acid X$_1$ when the following cycle was used:

*Fmoc deprotection:* 20% piperidine/ DMF (2.3 mL, 2 x 8 min), *Resin wash:* DMF (2 mL, 5 x 30 s),
*Coupling:* 3.123 (2.5 eq.) with HATU (2.4 eq.), DIPEA (0.5 mmol) in DMF (1 mL), (25 W, 75 °C, 20 min). The resin was agitated for 30 min, drained and washed with DMF.

Following the final Fmoc-deprotection, the peptidyl-resin was acetylated using General Method L and cleaved using General Method N to afford 188 mg of crude peptide. The crude peptide was purified using RP-HPLC using a gradient of 0 → 20% B, 0.3% B per min to afford the *title compound 4.20a* as a colourless powder (17.6 mg, 7.4% yield as a mixture of differently acetylated products K1OAc and K2OAc).

For K1OAc, X$_1$ = and K2OAc, X$_1$ = in above peptide sequence.

For K2OAc: R$_t$ 12.9 min, linear gradient 5 - 40% B over 35 min; [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]. LRMS (ESI+) [M+2H]$^{2+}$ 1153.1, calc. 1153.2; [M+3H]$^{3+}$ 769.3, calc. 769.1.
Figure 7.7. HPLC trace and low resolution mass spectrometry trace for K2OAc variant of CMP-Lys-1Gal-OAc 4.20a.

For K1OAc: $R_t = 12.4$ min, linear gradient 5 - 40% B over 35 min; [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]. LRMS (ESI+) $[M+2H]^{2+}$ 1132.4, calc. 1132.2; $[M+3H]^{3+}$ 755.3, calc. 755.1.

Figure 7.8 HPLC trace and low resolution mass spectrometry trace for K1OAc variant of CMP-Lys-1Gal-OAc 4.20a.
Chapter Seven

**CMP-Lys-2Gal-OAc 4.21a**

Ac-(POG)$_3$X$_1$EGX$_1$EG(POG)$_3$-NH$_2$

Resin bound Fmoc-EG(POG)$_3$-Rink-ϕ was synthesised on a 0.1 mmol scale using the Tribute™ synthesiser according to General Method F using HCTU (0.46 M, 4.5 eq.) as the coupling reagent and NMM (0.5 mL 5.1 M, 10 eq.) in NMP (0.5 mL) as the base. The peptidyl-resin was then elongated using General method F using the stronger coupling agent HATU (0.23 M, 2.3 eq.) to afford Fmoc-GX$_1$EGX$_1$EG(POG)$_3$-Rink-ϕ. The ‘clicked’ amino acid building block 3.123 (2.5 eq.) was manually coupled in a CEM™ manual microwave with HATU (2.4 eq.), DIPEA (5 eq.) in DMF (1.5 mL), 75 °C, 25 W, 20 min.

Subsequent elongation using a CEM™ microwave synthesiser (General method J) afforded resin bound Fmoc-(POG)$_3$X$_1$EGX$_1$EG(POG)$_3$-Rink-ϕ. Following the final Fmoc-deprotection, the peptidyl-resin was acetylated (General Method L) and then cleaved (General method M) to afford crude peptide. The crude peptide was purified batch wise using RP-HPLC using a gradient of 5 → 20% B, 2% B per min followed by 20 - 40% B, 0.5% B per min to afford the title compound 4.21a as a colourless powder (13.1 mg, 4.2% yield).

R$_t$ 20.5 min, 5 – 40% B over 35 min [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+2H]$^{2+}$ 1558.6, calc. 1558.3; [M+3H]$^{3+}$ 1039.4, calc. 1039.17.

![HPLC trace and low resolution mass spectrometry trace for CMP-Lys-2Gal-OAc 4.21a.](image)

Figure 7.9 HPLC trace and low resolution mass spectrometry trace for CMP-Lys-2Gal-OAc 4.21a.
The peptide was synthesised on a 0.1 mmol scale using the method described for CMP-Lys-2Gal-OAc 4.21a. The crude peptide was purified batch wise using RP-HPLC using a gradient of 5 → 40% B, 0.5% B per min to afford the title compound 4.22a as a colourless powder (2.1 mg, 0.6% yield).

R\text{t} = 23.3 \text{ min, linear gradient 5 - 40% B over 35 min [where buffer A = 0.1% TFA in DI H}_{2}\text{O, Buffer B = 0.1% TFA in CH}_{3}\text{CN]; LRMS (ESI+) [M+2H]^{2+} 1922.2, calc. 1921.4; [M+3H]^{3+} 1281.7, calc. 1281.3; [M+4H]^{4+} 961.4, calc. 961.2.}

**Figure 7.10** HPLC trace and low resolution mass spectrometry trace for CMP-Lys-3Gal-OAc 4.22a.
CMP-Ser-1Gal-OAc 4.24a

Ac-(POG)_3 X_2 EG(POG)_3-NH_2

Where X_2 =

Resin bound Fmoc-(POG)_3 Rink-φ was synthesised according to General Method F using HCTU on the Tribute™ synthesiser on a 0.1 mmol scale. The resultant resin was divided in half, with one half (0.05 mmol) being used for this synthesis of CMP-Ser-1Gal-OAc (4.24a) and the other half used synthesis of CMP-Lys-Control (4.23). Fmoc-Gly-OH and Fmoc-Glu(O'Bu) were then incorporated using the Tribute™ synthesiser with double coupling cycles using HCTU (General Method G). The clicked amino acid building block 3.124 was manually coupled using HATU (General Method H) and POG assembled using the Tribute™ synthesiser with double coupling cycles using HCTU (General Method G). The remaining amino acids POGPOG were incorporated onto the resin bound peptide with double coupling cycles using HATU (General Method H). Following the final Fmoc-deprotection, the peptidyl-resin was acetyl capped (General Method K) and cleaved from resin (General Method M) to afford the crude peptide. The crude peptide was purified using RP-HPLC using a gradient of 0 → 16% B, 0.3% B per min to afford the title compound 4.24a as a colourless powder (13.7 mg, 5.8% yield)

R_t 14.1 min, linear gradient 5 - 40% B over 35 min [where buffer A = 0.1% TFA in DI H_2O, Buffer B = 0.1% TFA in CH_3CN; LRMS (ESI+) [M+2H]^{2+} 1174.1, calc. 1174.0; [M+3H]^{3+} 783.2, calc. 783.0.

Figure 7.11 HPLC trace and low resolution mass spectrometry trace for CMP-Ser-1Gal-OAc 4.24a.
Chapter Seven

**CMP-Ser-2Gal-OAc 4.25a**

Ac-(POG)_3 X_2EG X_2EG(POG)_3-NH₂

Resin bound Fmoc-EG(POG)_3-Rink-ϕ was synthesised on a 0.1 mmol scale according to **General Method F** using HCTU on the Tribute™ synthesiser. Fmoc-EG(POG)_3-Rink-ϕ was then elongated to afford Fmoc-X_2EG-X_2EG(POG)_3-Rink-ϕ using double couplings with HCTU using the Tribute™ synthesiser (**General Method H**). The ‘clicked’ amino acid 3.124 was manually coupled using HATU using the Tribute™ synthesiser (**General Method F**). The peptidyl-resin was then split in half; with one half being used for the synthesis of CMP-Ser-2Gal-OAc (4.25a) and the other half for the synthesis of CMP-Ser-3Gal-OAc (4.26a, see later).

For CMP-Ser-2Gal-OAc (4.25a), the remaining residues were incorporated onto the peptidyl resin with double coupling cycles using HATU using the Tribute™ synthesiser (**General Method H**). Following the final Fmoc-deprotection, the peptidyl-resin was acetyl capped (**General Method K**) and cleaved from resin (**General Method M**) to afford the crude peptide. The crude peptide was purified batch wise using RP-HPLC using a gradient of 5 - 25% B, 0.5% B per min to afford the **title compound 4.25a** as a colourless powder (17.6 mg, 92% purity, 5.8% yield).

R_t 16.1 min, linear gradient 5 - 40% B over 35 min; [where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN]; LRMS (ESI+) [M+2H]²⁺ 1516.6, calc. 1516.2; [M+3H]³⁺ 1011.4, calc. 1011.1.

![HPLC trace and low resolution mass spectrometry trace for CMP-Ser-2Gal-OAc 4.25a.](image)

**Figure 7.12** HPLC trace and low resolution mass spectrometry trace for CMP-Ser-2Gal-OAc 4.25a.
Chapter Seven

**CMP-Ser-3Gal-OAc 4.26a**

Ac-(POG)$_3$ X$_2$EG X$_2$EG X$_2$EG(POG)$_3$-$\text{NH}_2$

The peptidyl-resin Fmoc-X$_2$EGX$_2$EG(POG)$_3$-Rink-$\phi$ was synthesised on a 0.1mmol scale using the method described for the synthesis of CMP-Ser-2Gal-OAc (4.25a). The peptidyl resin was then elongated to afford Fmoc-X$_2$EGX$_2$EGX$_2$EG(POG)$_3$-Rink-$\phi$ using double coupling cycles with HCTU using the Tribute™ synthesiser (General Method G). The ‘clicked’ amino acid 3.124 was manually coupled with HATU using the Tribute™ synthesiser (General Method H).

The remaining residues were incorporated onto the peptidyl resin with double coupling cycles using HATU (General Method H). Following the final Fmoc-deprotection, the peptidyl-resin was acetyl capped (General Method K) and cleaved from resin (General Method M) to afford the crude peptide. The crude peptide was purified batch wise using RP-HPLC using a gradient of 5 → 25% B, 0.5% B per min to afford the *title compound* 4.26a as a colourless powder (2.2 mg, 82% purity, 0.6% yield).

R$_t$ 17.5 min, linear gradient 5 - 40% B over 35 min [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+2H]$^{2+}$ 1858.7, calc. 1858.4; [M+3H]$^{3+}$ 1239.5, calc. 1239.3.

![HPLC trace and low resolution mass spectrometry trace for CMP-Ser-3Gal-OAc 4.26a.](image)

**Figure 7.13** HPLC trace and low resolution mass spectrometry trace for CMP-Ser-3Gal-OAc 4.26a.
Chapter Seven

CMP-Lys-1Gal-OH 4.20b

Ac-(POG)₃X₃EG(POG)₃-NH₂

Where X₃ =

Crude CMP-Lys-1Gal-OAc 4.20a (17.2 mg) was deacetylated according to General Method Q and purified using RP-HPLC using a gradient of 0 → 30% B, 0.3% B per min to afford the title compound 4.20b as a colourless powder (11.7 mg, 5.3% overall yield including synthesis of acetylated peptide).

Rₜ 12.8 min, linear gradient 5 - 40% B over 35 min [where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN]; LRMS (ESI+) [M+2H]²⁺ 1111.1, calc. 1111.0; [M+3H]³⁺ 741.3, calc. 741.0.

Figure 7.14 HPLC trace and low resolution mass spectrometry trace for CMP-Lys-1Gal-OH 4.20b.
CMP-Lys-2Gal-OH 4.21b
Ac-(POG)$_3$X$_3$EGX$_3$EG(POG)$_3$-NH$_2$

Crude CMP-Lys-2Gal-OAc 4.21a (17.9 mg) was deacetylated according to General Method Q and purified using RP-HPLC using a gradient of 5% - 20% B, 0.3% B per min to afford the title compound 4.21b as a colourless powder (2.4 mg, 0.9% yield including synthesis of acetylated peptide).

R$_t$ 9.0 min, linear gradient 5- 40% B over 35 min [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B =0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+2H]$^{2+}$ 1390.6, calc. 1390.4; [M+3H]$^{3+}$ 927.3, calc. 927.3.

Figure 7.15 HPLC trace and low resolution mass spectrometry trace for CMP-Lys-2Gal-OH 4.21b.
Chapter Seven

CMP-Lys-3Gal-OH 4.22b
Ac-(POG)₃X₃EGX₃EGX₃EG(POG)₃-NH₂

Crude CMP-Lys-3Gal-OAc 4.22a (12.5 mg) was deacetylated according to General Method Q and purified using RP-HPLC using a gradient of 5 - 20% B, 0.3% B per min to afford the title compound 4.22b as a colourless powder (5.0 mg, 1.5 % overall yield including synthesis of acetylated peptide).

Rₜ 13.1 min, linear gradient 5- 40% B over 35 min] [where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN]; LRMS (ESI+) [M+2H]²⁺ 1113.6, calc. 1113.3; [M+3H]³⁺ 835.5, calc. 835.2.

![HPLC trace and low resolution mass spectrometry trace for CMP-Lys-3Gal-OH 4.22b.](image)

**Figure 7.16** HPLC trace and low resolution mass spectrometry trace for CMP-Lys-3Gal-OH 4.22b.
CMP-Ser-1Gal-OH 4.24b

Ac-(POG)₃ X₄EG(POG)₃-NH₂

Where X₄ =

Crude CMP-Ser-1Gal-OAc 4.24a (9.8 mg) was deacetylated according to General Method Q and purified using RP-HPLC using a gradient of 5 → 25% B, 0.5% B per min to afford the title compound 4.24b as a colourless powder (1.0 mg, 0.5% overall yield including synthesis of acetylated peptide).

Rₜ 12.6 min, linear gradient 5 - 40% B over 35 min [where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN]; LRMS (ESI+) [M+2H]²⁺ 1090.4, calc. 1090.0; [M+3H]³⁺ 727.0, calc. 727.0.

Figure 7.17 HPLC trace and low resolution mass spectrometry trace for CMP-Ser-1Gal-OH 4.24b.
CMP-Ser-2Gal-OH 4.25b

Ac-(POG)_3 xEG xEG(POG)_3-NH₂

Crude CMP-Ser-2Gal-OAc 4.25a (56 mg) was deacetylated according to General Method Q and purified using RP-HPLC using a gradient of 5 → 25% B, 0.5% B per min to afford the title compound 4.35b as a colourless powder (1.0 mg, 0.4% overall yield including synthesis of acetylated peptide).

Rt 12.5 min, gradient 5 - 40% B over 35 min [where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN]; LRMS (ESI+) [M+2H]^{2+} 1348.6, calc. 1348.2; [M+3H]^{3+} 899.3, calc. 899.1.

Figure 7.18 HPLC trace and low resolution mass spectrometry trace for CMP-Ser-2Gal-OH 4.25b.
CMP-Ser-3Gal-OH 4.26b

Ac-(POG)₃ X₄EG X₄EG X₄EG(POG)₃-NH₂

CMP-Ser-3Gal-OAc 4.26a (30.1 mg) was deacetylated according to **General Method Q** and purified using RP HPLC using a gradient of 5 → 25% B, 0.5% B per min to afford the **title compound 4.36b** as a colourless powder (1.3 mg, 0.4% overall yield including synthesis of acetylated peptide).

R<sub>t</sub> 12.6 min, linear gradient 5 - 40% B over 35 min; [where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN]; LRMS (ESI+) [M+2H]⁺ 1606.9, calc. 1606.3, [M+3H]³⁺ 1071.4, calc. 1071.2, [M+4H]⁴⁺ 804.0, calc. 803.7.

![HPLC trace and low resolution mass spectrometry trace for CMP-Ser-3Gal-OH 4.26b.](image)

**Figure 7.19** HPLC trace and low resolution mass spectrometry trace for CMP-Ser-3Gal-OH 4.26b.
Chapter Seven

**CMP-Lys-control 4.23**

Ac-(POG)$_3$KEG(POG)$_3$-NH$_2$

Resin bound Fmoc-(POG)$_3$-Rink-ϕ (0.05 mmol) obtained during the synthesis of CMP-Ser-1Gal-OAc (4.24a), ϕ-Rink-(GOP)$_3$-Fmoc was elongated on the Tribute™ synthesiser with double coupling cycles using HCTU (General Method H). Following the final Fmoc-deprotection on the Tribute™ synthesiser (20% piperidine/ DMF, 3 mL, 3 x 5 min), the peptidyl resin was then N-terminal acetyl capped using General Method L. The acetylated peptide was then cleaved from resin (General Method M) to afford the crude peptide, which was purified batch wise using RP-HPLC using a gradient of 0 - 15% B, 0.3% B per min to afford the title compound 4.23 as a colourless powder (26.0 mg, 26% yield).

R$_t$ 11.2 min, linear gradient 5 - 40% B over 35 min; [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+H]$^+$ 1977.8, calc. 1997.1; [M+2H]$^{2+}$ 989.0, calc. 989.1, [M+3H]$^{3+}$ 659.9, calc. 659.7.

![HPLC trace and low resolution mass spectrometry trace for CMP-Lys-control 4.23.](image-url)

**Figure 7.20** HPLC trace and low resolution mass spectrometry trace for CMP-Lys-control 4.23.
CMP-Ser-control 4.27

Ac-(POG)$_3$SEG(POG)$_3$-NH$_2$

This peptide synthesis was conducted according to General Method E, except that a less concentrated solution of piperidine (20% piperidine/DMF in place of 40% piperidine/DMF) was used for Fmoc deprotection. Following the final Fmoc-deprotection, the peptidyl resin was then N-terminal acetyl capped using General Method L and cleaved from resin (General Method M) to afford the crude peptide, which was purified batch wise using RP-HPLC using a gradient of 0 → 15% B, 0.3% B per min to afford the title compound 4.27 as a colourless powder (31.5 mg, 16.3% yield).

R$_t$ 11.6 min, linear gradient 5 - 40% B over 35 min; [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+2H]$^{2+}$ 968.5, calculated 968.5.

Figure 7.21 HPLC trace and low resolution mass spectrometry trace for CMP-Ser-control 4.27.
Family 3 CMP-Adpn peptides

CMP-Adpn-Lys-3Gal-OH 5.1b

This peptide was synthesised on a 0.033 mmol scale instead of a 0.01 mmol scale. The derivatised Fmoc-Lys residue 3,123 was incorporated into the peptide sequence using the following manual coupling conditions in a CEM™ microwave (CEM Corporation, Mathews, NC) according to General Method J.

The other non-functionalised amino acids were synthesised using The Tribute™ synthesiser (Protein Technologies, Tuscon, Az). Resin bound Fmoc- XsGDI(GPO)3NH2-Rink-ϕ was synthesised using General method F but using a larger volume of 20% piperidine/ DMF deprotection solution (2 mL instead of 1.5 mL). The peptidyl resin was then elongated using General Method H to afford Fmoc-GDOGLIG-E XsGDI(GPO)3NH2-Rink-ϕ.

Further elongation of the peptide was then carried out using General Method E to afford Fmoc-EX3-GEX3GDOGLIGE X3GDI(GPO)3NH2-Rink-ϕ but with larger volume of 40% piperidine/ DMF deprotection solution (2 mL instead of 1.5 mL). The final amino acids were then added using General method F using HCTU as the coupling agent and shorter coupling time for each amino acid (10 min instead of 20 min) to afford Fmoc-(GPO)3GEX3GEX3GDOGLIGO X3GDI(GPO)3NH2-Rink-ϕ.

N-terminal acetylation was conducted using General Method K followed by cleavage from resin using General Method P to afford the crude O-acetylated peptide CMP-Adpn-Lys-3Gal-OAc, 5.1a which was identified by Fl-MS as having the the [M+3H]3+ (calculated 1560.6, observed 1561.1) and [M+4H]4+ (calculated 1170.7, observed 1170.9) charged states.
The crude O-acetylated peptide CMP-Adpn-Lys-3Gal-OH, 5.1a (19.2 mg) was then deacetylated according to General Method Q to afford the crude peptide (3.7 mg, 0.88 µmol) which was purified using RP-HPLC using a gradient of 5 - 40% B, 0.5% B per min to afford the title compound 5.1b as a colourless powder (2.6 mg, 1.9% overall yield including synthesis of acetylated peptide).

Rt 12.2 min, linear gradient 5 – 40% B over 35 min; [where buffer A = 0.1% TFA in DI H₂O, Buffer B = 0.1% TFA in CH₃CN]; LRMS (ESI+) [M+3H]³⁺ 1392.8, calc. 1392.6; [M+4H]⁴⁺ 1044.9, calc. 1044.7.

Figure 7.22 Flow-injection mass spectrometry trace for CMP-Adpn-Lys-3Gal-OAc 5.1a.

Figure 7.22 HPLC trace and low resolution mass spectrometry trace for CMP-Adpn-Lys-3Gal-OH 5.1b.
CMP-Adpn-Lys control 5.2

Ac-(GPO)$_3$(GEK)$_2$GDOGLIGEKGDI(GPO)$_3$-NH$_2$

This peptide was synthesised on a 0.05 mmol scale using 0.85 g/mmol loading.

Fmoc-(GPO)$_3$(GEK)$_2$GDOGLIGEKGDI(GPO)$_3$-Rink-φ was synthesised according to General Method F using HCTU/NMM in NMP and piperidine as the coupling and Fmoc-deprotection agents respectively.

The peptidyl-resin was then N-terminal acetylated using General Method K followed by cleavage from resin using General Method O to afford the crude peptide (10.7 mg, 31.3 µmol, 31.3% crude yield) which was purified using RP-HPLC using a gradient of 5 – 20% B, 0.2 %B per min to afford the title compound 5.2 as a colourless powder (2.8 mg, 1.6% yield).

R$_t$ 12.5 min, linear gradient 5 – 40 % B over 35 min; [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+3H]$^{3+}$ 1148.6, calc. 1148.6; [M+4H]$^{4+}$ 861.8, calc. 861.7.

Figure 7.23 HPLC trace and low resolution mass spectrometry trace for CMP-Adpn-Lys-Control 5.2.
This peptide synthesis was conducted on a 0.025 mmol scale using 0.85 g/mmol loading.

Resin bound Fmoc-GDI(GPO)₃Rink-ϕ was synthesised using the Tribute™ Peptide Synthesiser (Protein Technologies Inc.) according to General method F, using HCTU/ NMM in NMP and piperidine as the coupling and Fmoc-deprotection agents respectively.

Elongation of the peptidyl resin was conducted using the CEM™ microwave (CEM Corporation, Mathews, NC) according to General Method J with HCTU/ DIPEA and piperidine as the coupling and Fmoc-deprotection agents respectively. The derivatised Fmoc-Ser residue 3.124 was coupled into the peptide sequence manually using the CEM™ microwave and General Method J. These conditions afforded the peptidy-resin Fmoc-GPOGEₓ₄GEₓ₄GDOGLIGOₓ₄GDI(GPO)₃Rink-ϕ. The final amino acids were incorporated using the Tribute™ Peptide Synthesiser and General Method I with HATU/NMM in NMP and piperidine to afford Fmoc-(GPO)ₓ₄GEₓ₄GEXₓ₄GDOGLIGOₓ₄GDI(GPO)₃Rink-ϕ.

The peptidyl-resin was then N-terminal acetylated using General Method K followed by cleavage from resin using General Method O to afford the crude O-acetylated peptide CMP-Adpn-Ser-3Gal-OH 5.3a, which was identified by FI-MS as having the [M+3H]³⁺ (calculated 1504.6, observed 1505.1) and [M+4H]⁴⁺ (calculated 1128.7, observed 1128.7) charged states.
The crude $O$-acetylated peptide 5.3a (16.7 mg) was then deacetylated according to General Method Q to afford the crude peptide (4 mg, 0.98 µmol), which was purified using RP-HPLC using a gradient of 5 - 40% B, 0.5% B per min to afford the title compound 5.3b as a colourless powder (0.9 mg, 0.9% overall yield including synthesis of acetylated peptide).

$R_t$ 13.6 min, linear gradient 5 – 40 % B over 35 min; [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) $[M+3H]^3+$ 1350.9, calc. 1350.7; $[M+4H]^4+$ 1013.5, calc. 1013.3.
CMP-Adp-Ser control 5.4

Ac-(GPO)$_3$-(GES)$_2$GDOGLIGOSGDI(GPO)$_3$-NH$_2$

This peptide was synthesised on a 0.033 mmol scale using 0.85 g/mmol loading.

Fmoc-O(GES)$_2$GDOGLIGOSGDI(GPO)$_3$-Rink-ϕ was synthesised according to General Method F (single coupling cycles) using HCTU/NMM in NMP and piperidine as the coupling and Fmoc-deprotection agents respectively.

The final amino acids were then incorporated using General Method H (double coupling cycles) and using HCTU/NMM in NMP and piperidine as the coupling and Fmoc-deprotection agents respectively to afford Fmoc-(GPO)$_3$-(GES)$_2$GDOGLIGOSGDI(GPO)$_3$-Rink-ϕ.

The peptidyl-resin was then $N$-terminal acetylated using General Method L and then cleaved from resin using General Method O to afford the crude peptide (13.5 mg, 4.06 µmol), which was purified batch wise using RP-HPLC using a gradient of 5 – 20 % B, 0.5 % B per min to afford the title compound 5.4 as a colourless powder (4.8 mg, 4.4% overall yield including synthesis of acetylated peptide).

$R_t$ 15.8 min, linear gradient 5 – 40 % B over 35 min; [where buffer A = 0.1% TFA in DI H$_2$O, Buffer B = 0.1% TFA in CH$_3$CN]; LRMS (ESI+) [M+2H]$^{2+}$ 1660.7, calc. 1660.7; [M+3H]$^{3+}$ 1107.5, calc. 1107.5, [M+4H]$^{4+}$ 831.0, calc. 830.9.

Figure 7.26 HPLC trace and low resolution mass spectrometry trace for CMP-Adpn-Ser-control 5.4.
Appendix One

NMR Spectra For Novel Organic Compounds
Appendix NMR Spectra For Novel Organic Compounds

(2S)-2-[(9H-Flourenyl-9-yl)-methoxy]carbonyl]amino-6-((2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl(oxy)methyl)-1H-1,2,3-triazol-1-yl)hexanoic acid, 3.123

Fmoc-Lys-(Click-Gal-OAc)-OH

$^1$H NMR (400 MHz, CDCl$_3$)
$^{13}$C NMR (100 MHz, CDCl$_3$)

COSY
Appendix

DEPT-135

HSQC
Appendix

(2S)-2-([(9H-Flourenyl-9-yl-)methoxy]carbonyl)amino]-3-((2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl)-1H-1,2,3-triazol-4-yl)methyl)oxy)propanoic acid, 3.124

Fmoc-Ser-(Click-Gal-OAc)-OH

1H NMR (400 MHz, CDCl3)
Appendix

(2S)-Methyl-((9H-fluorenlymethoxycarbonyl)amino)-3-(prop-2-yn-1-yloxy)propanoate 3.76

\[
\text{NHFmcc}
\]

\[
3.76
\]

\(^1\)H NMR (400 MHz, CDCl\(_3\))
$^{13}$C NMR (100 MHz, CDCl$_3$)

COSY
References

(1) Health, M. o. 2014.
References


(43) Shand, B.; Scott, R.; Connolly, S.; Clarke, R.; Baker, J.; Elder, P.; Frampton, C.; Yeo, J. *Diabetes, Obesity and Metabolism* 2007, 9, 540.


References

(73) Schöllkopf, U.; Groth, U.; Deng, C. Angew. Chem. 1981, 93, 793.
(125) Koenigs, W.; Knorr, E. Berichte der deutschen chemischen Gesellschaft 1901, 34, 957.
References

(184) Zemplén, G.; Kunz, A. Berichte der deutschen chemischen Gesellschaft (A and B Series) 1924, 57, 1357.

316


