

**An Investigation into the Effects of Glycosylation
On the Properties of L-Proline in Peptides**

by

Neil Wayne Owens

A Thesis submitted to the Faculty of Graduate Studies of
The University of Manitoba
in partial fulfilment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Chemistry
University of Manitoba
Winnipeg, Manitoba, Canada

Copyright © 2009 by Neil W. Owens

Abstract

The amino acid L-proline plays a critical role in many biological processes. Therefore, efforts have been made to understand and control its influence. Since glycosylation is a common post-translational modification known to affect the characteristics of peptides and proteins, in a series of experiments, the effects of glycosylation on the properties of L-proline in peptides have been investigated.

A conformationally constrained C-glucosyl proline hybrid is introduced, which has the capacity to vary the N-terminal amide equilibrium in model peptides through derivatization of the carbohydrate scaffold.

For the first time, a comprehensive study of the effects of O-glycosylation on the kinetics and thermodynamics of prolyl amide isomerization is reported. The O-glycosylation of 4-hydroxy-L-proline has different effects on amide isomerization depending on the stereochemistry of the 4-hydroxyl group, which alters the orientation of the glycan with respect to the prolyl side chain. 4*S*-Galactosylation of 4-hydroxy-L-proline affects both the amide isomer equilibrium and the rate of amide isomerization, whereas 4*R*-galactosylation does not measurably influence either parameter. However, close contacts between the carbohydrate and prolyl rings lead to changes in the conformation and stability of longer peptides.

As an expansion on these initial model studies, the effects of prolyl O-glycosylation on the properties of model peptides of two extremely important structural proteins are investigated. O-Galactosylation of 4*R*-hydroxy-L-proline residues in collagen model peptides does not preclude formation of the collagen triple helix, where the anomeric linkage of the Hyp O-glycan has slightly different influences on the

conformational stability of the peptides. Also, the *O*-galactosylation of Hyp residues in polyproline model peptides causes a large increase in conformational stability. In both cases, interactions between the glycan and the peptide backbone and changes in hydration are implicated in contributing to the conformational stabilization of the peptides.

These studies demonstrate that both natural and unnatural glycosylation of L-proline can be used as a means to control amide isomerization, and can increase the conformational stability of peptides, properties that will likely contribute to the development of new biomaterials. Also, these experiments provide further insight into the broad role glycosylation plays in affecting peptide and protein structure.

Acknowledgements

I would like to thank my supervisor Dr. Frank Schweizer for sharing his enthusiasm for research, which has inspired me to exceed my own expectations, and for his support throughout this project.

Thank you to my examining committee, Dr. Joe O'Neil, Dr. Philip Hultin, Dr. Elizabeth Worobec, and Dr. Robert Ben for your teachings, comments, and advice.

Thank you to my colleagues, especially Dr. Kaidong Zhang, Siddharth Sikchi, Xi Hao, Julia Campbell, Robel Teklebrhan, and Meghan Gallant for your friendship and support.

Thank you to members of the department of chemistry, especially Dr. Kirk Marat for exceptional help with NMR analysis, Dr. Erika Lattová for running MALDI analysis, Dr. James Xidos for help with statistical analysis, as well as Dr. Jörg Stetefeld and Dr. Frank Hruska for many helpful and enjoyable discussions.

I would also like to thank the members of the Schweizer group: Dr. Smritilekha Bera, Dr. Dhananjoy Mondal, Adrian Lee, Zhizhi Sun, Craig Braun, Danielle Desautels, David Taylor, Malcolm Lucy, Marlin Penner, Dung-Huang Nguyen, Natali Serrano, Jason Zhang, and Brandon Findlay; I have really enjoyed working with all of you.

I would like to thank several people for inspiring me to carry out this work: my high school science teachers: Mr. Albertson and Ms. Stoeber, as well as Dr. David McKinnon, and my grandfather Douglas MacFarlane.

Finally to my family and friends, especially Marie, my parents Thompson and Audrey, and my brother Kyle, thank you for your continued support, it has made all the difference.

Dedication

For my parents, Thompson and Audrey

Table of Contents

Abstract	i
Acknowledgements	iii
Dedication	iv
Table of Contents	v
List of Abbreviations	x
Chapter 1: Introduction and Background	1
1.1 Introduction.....	1
1.2 Properties and functions of carbohydrates found in nature	2
1.3 The development of synthetic glycopeptides.....	13
1.4 C-Glycosyl amino acids.....	21
1.5 L-Proline.....	25
1.6 Proline analogues are used to control prolyl amide isomerization .	41
1.7 Collagen	55
1.8 Hydroxyproline-rich glycoproteins.....	77
1.9 References.....	92
Chapter 2: Thesis Objectives	119
2.1 References.....	122
Chapter 3: Tuning of the Prolyl <i>Trans/Cis</i> Amide Rotamer Population Using C- Glycosyl Proline Hybrids	123
3.1 Abstract.....	123

3.2 Introduction.....	124
3.3 Results.....	127
3.4 Discussion.....	135
3.5 Conclusions.....	138
3.6 Experimental section.....	139
3.7 References.....	152
Chapter 4: Effects of Glycosylation of (2<i>S</i>,4<i>R</i>)-4-Hydroxyproline on the Conformation, Kinetics and Thermodynamics of Prolyl Amide Isomerization	158
4.1 Abstract.....	158
4.2 Introduction.....	159
4.3 Results.....	160
4.4 Discussion.....	168
4.5 Conclusions.....	169
4.6 Experimental section.....	169
4.7 References.....	177
Chapter 5: The Implications of (2<i>S</i>,4<i>S</i>)-Hydroxyproline 4-<i>O</i>-Glycosylation on Prolyl Amide Isomerization	180
5.1 Abstract.....	180
5.2 Introduction.....	181
5.3 Results.....	185
5.4 Discussion.....	196

5.5 Conclusions.....	199
5.6 Experimental section.....	201
5.7 References.....	206
Chapter 6: The Effects of 4<i>R</i>-Hydroxy-L-Proline <i>O</i>-Glycosylation on the Stability of Collagen Model Peptides	214
6.1 Abstract.....	214
6.2 Introduction.....	215
6.3 Results.....	219
6.4 Discussion.....	228
6.5 Conclusions.....	230
6.6 Experimental section.....	231
6.7 References.....	231
Chapter 7: Contiguous <i>O</i>-Galactosylation of 4<i>R</i>-Hydroxy-L-Proline Residues Forms Hyper-Stable Polyproline II Helices	263
7.1 Abstract.....	263
7.2 Introduction.....	264
7.3 Results.....	267
7.4 Discussion.....	273
7.5 Conclusions.....	278
7.6 Experimental section.....	279
7.7 References.....	285

Chapter 8: Conclusions and Future Work	295
8.1 References.....	302
Chapter 9: Supporting Information for Chapter 3	304
9.1 NMR assignments.....	305-310
9.2 Tables.....	311-314
9.3 NMR spectra.....	315-368
Chapter 10: Supporting Information for Chapter 4.....	369
10.1 Tables.....	370-373
10.2 Van't Hoff plots	374-376
10.3 Eyring plots.....	377-379
10.4 Spectra.....	380-388
10.5 nOe spectra.....	389-396
10.6 Summary of nOe interactions	397
10.7 Magnetization transfer NMR experiments.....	398-427
Chapter 11: Supporting Information for Chapter 5.....	428
11.1 General details	429
11.2 General procedures	429
11.3 Synthesis and characterization.....	430-436
11.4 Tables.....	437-438
11.5 Van't Hoff plots	439-441
11.6 Spectra.....	442-455

11.7 nOe spectra.....	456-470
11.8 Magnetization transfer NMR experiments.....	471-479
Chapter 12: Supporting Information for Chapter 6.....	480
12.1 MALDI and analytical chromatograms	481-483
12.2 Table of ellipticity as a function of temperature.....	484
12.3 Melting curves and CD spectra.....	485-488
12.4 NMR spectra	489-506
Chapter 13: Supporting Information for Chapter 7.....	507
13.1 Mass spectra and analytical chromatograms.....	508-510
13.2 CD spectra.....	511-512
13.3 Table of ellipticity as a function of temperature.....	513
13.4 CD melting curves	514-515
13.4 NMR spectra	516-517
List of Publications and Patents Related to Thesis Work	518

List of Abbreviations

[θ]	mean residue ellipticity
[α]	specific rotation
Ac	acetyl
Ala	L-alanine
Asn	L-asparagine
Asp	L-aspartic acid
aq	aqueous
<i>t</i> Bu	<i>tert</i> -butyl
Bn	benzyl
Boc	<i>tert</i> -butoxycarbonyl
<i>c</i>	concentration in g/100 mL
CD	circular dichroism
CDCl ₃	deuterated chloroform
COSY	correlation spectroscopy
δ	chemical shift in parts per million
D ₂ O	deuterium oxide
DCC	1,3-dicyclohexylcarbodiimide
DCM	dichloromethane
DIPEA	diisopropylethylamine
DMF	<i>N,N</i> -dimethylformamide
EtOAc	ethyl acetate
Fmoc	9-fluorenylmethoxycarbonyl

Flp	4 <i>R</i> -fluoro-L-proline
flp	4 <i>S</i> -fluoro-L-proline
Gal	D-galactose
Glc	D-glucose
GlcNAc	<i>N</i> -acetyl-D-glucosamine
Gly	glycine
GOESY	gradient enhanced nuclear Overhauser effect spectroscopy
HRGP	hydroxyproline-rich glycoprotein
HPLC	high performance liquid chromatography
HSQC	heteronuclear single quantum coherence
Hyp	4 <i>R</i> -hydroxy-L-proline
hyp	4 <i>S</i> -hydroxy-L-proline
Ile	L-isoleucine
<i>J</i>	coupling constant (in NMR)
k_{ct}	rate constant from <i>cis</i> to <i>trans</i>
k_{tc}	rate constant from <i>trans</i> to <i>cis</i>
$K_{trans/cis}$	equilibrium constant of <i>trans:cis</i> amide isomers
Leu	L-leucine
Lys	L-lysine
MALDI	matrix-assisted laser desorption ionization
MeOH	methanol
MS	mass spectrometry
NHMe	<i>N</i> -methylamide

nOe	nuclear Overhauser effect
NMR	nuclear magnetic resonance
OMe	methoxyl
P4H	prolyl-4-hydroxylase
Pfp	pentafluorophenyl
Phe	L-phenylalanine
PP	polyproline
ppm	parts per million
Pro	L-proline
PyBOP	benzotriazol-1-yl oxytripyrrolidinophosphonium hexafluorophosphate
SE	standard error
Ser	L-serine
SPPS	solid-phase peptide synthesis
TBDMS	<i>tert</i> -butyldimethylsilyl
TBTU	<i>O</i> -(1H-benzotriazole-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium tetrafluoroborate
TEA	triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
Thr	L-threonine
TMSOTf	trimethylsilyl trifluoromethanesulfonate
Tyr	L-tyrosine

Chapter 1: Introduction and Background

1.1 Introduction

The purpose of this thesis is to study the effects of natural and unnatural glycosylation on the properties of L-proline in peptides, towards a greater understanding of their role in biology, and for the development of unnatural glycopeptides. This requires an understanding of two broad topics: how natural and unnatural glycosylation affects the properties of peptides and proteins, and the characteristics of L-proline and its role in nature. A background to the structure and function of carbohydrates in nature will be presented, with examples of the effects of carbohydrates on the conformation and stability of glycopeptides. The development of unnaturally linked glycopeptides, with examples of *C*-glycosyl amino acids and their use for improving glycopeptide stability, will also be discussed.

Similarly, an introduction to the unique properties of L-proline and its role in nature will be presented, along with examples of how the modification of L-proline can be used to understand biological phenomena. Finally, the role of 4-hydroxy-L-proline analogues in understanding and improving upon the stability of collagen will be discussed, as well as the implications of 4-hydroxy-L-proline glycosylation on the polyproline II conformation, as found in hydroxyproline-rich glycoproteins.

1.2 Properties and Functions of Carbohydrates Found in Nature

Carbohydrates play important and complex roles in nature, and while a great deal of information has been accumulated about their function, there is much left to be discovered. Understanding the function of carbohydrates in nature begins by considering their structure.¹

1.2.1 Carbohydrate Structure

Carbohydrates, sugars, or saccharides, have the formal stoichiometry of $C_n(H_2O)_n$, and thus are worthy of their title derived from 'hydrates of carbon'. The most common carbohydrates are hexoses, meaning they are composed of six carbon atoms, with an aldehyde functional group at position 1 in the carbon chain (Figure 1.2.1.1).¹

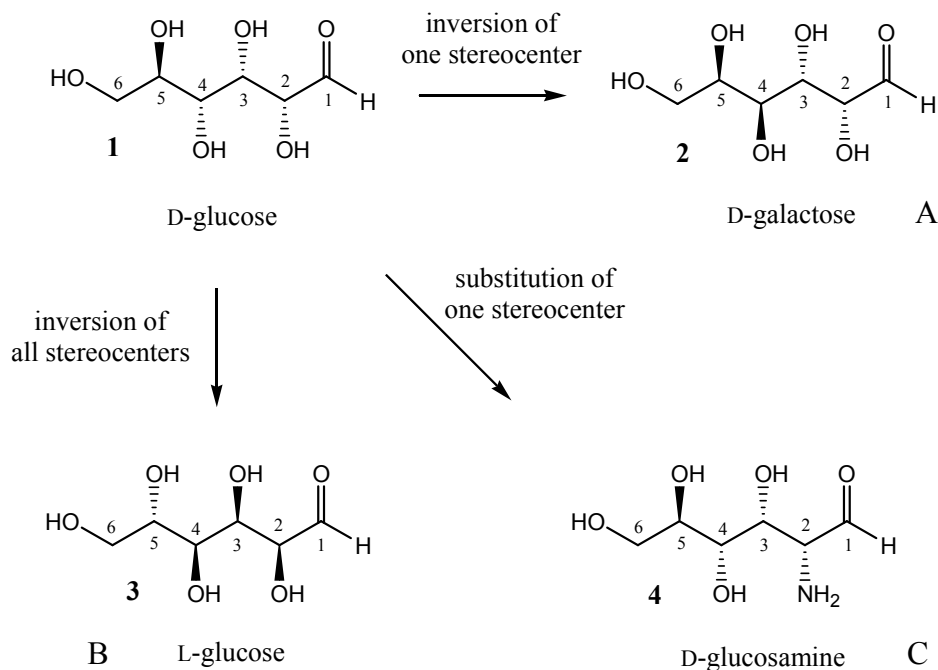


Figure 1.2.1.1: Inversion of stereocenters leads to diversity in carbohydrates.

The remaining carbon atoms, being hydroxylated, create four stereogenic centers. Since each can exist in two possible configurations, there are a total of 16 possible hexoses. The inversion of one chiral center leads to a different hexose, for example, glucose (Glc) and galactose (Gal) differ only in the stereochemistry at carbon 4 (Figure 1.2.1.1A). If all of the chiral centers are switched, then the two hexoses are enantiomers of each other, leading to the nomenclature of L- and D-sugars (Figure 1.2.1.1B). In some hexoses, certain hydroxyl groups are substituted by other functional groups, for example, glucosamine has an amino group attached to carbon 2 (Figure 1.2.1.1C). This is commonly modified in nature to an acetamide to form *N*-acetyl glucosamine (GlcNAc).¹

1.2.2 Carbohydrate Conformation

In nature, hexose saccharides are usually found in a cyclised or ‘pyranose’ form.¹ This creates an additional stereogenic center at carbon 1, referred to as the anomeric carbon. The anomeric hydroxyl group can adopt either an axial orientation (α) or an equatorial orientation (β) (Figure 1.2.2.1).

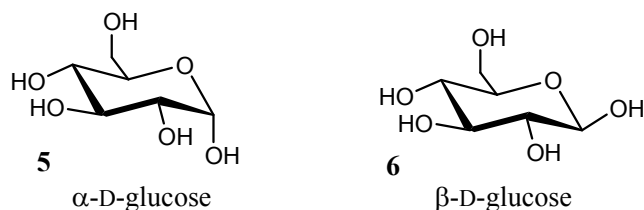


Figure 1.2.2.1: The α - and β -anomers of D-glucose in a 4C_1 chair conformation.

The preferred conformation of a pyranose carbohydrate ring is a ‘chair’ conformation, where four ring atoms form a plane, and the other ring atoms are placed above and below the plane; this conformation is preferred thermodynamically since it

minimizes steric interactions between hydrogen atoms or hydroxyl groups and the number of eclipsed ring dihedral angles.¹

1.2.3 Functions of Carbohydrates in Humans

It is of great delight to glycobiologists that carbohydrates were long considered to only have a minor role in biology (for energy storage), but in the past 20 years they have been found to be involved in many fundamental biological processes, such as cellular communication and protein function (Table 1.2.3.1).² However, carbohydrates do not act independently, and often are covalently-linked with other biomolecules, such as lipids and proteins. The resultant glycolipids and glycoproteins are referred to as glycoconjugates;¹ here the focus will be on glycopeptides and glycoproteins.

Table 1.2.3.1: Examples of the roles of carbohydrates in biological processes.^{2,4,5,6}

Carbohydrate Linkage	Biological Roles
<i>N</i> -linked glycans	-protein trafficking
	-cellular communication
	-initiation of inflammation
	-host defense
<i>O</i> -linked glycans	-components of the extracellular matrix
	-molecular lubricants

The post-translational modification of proteins and peptides by glycosylation causes a myriad of changes to their structure and function.² This process is still poorly understood, but several underlying elements have been uncovered. It has been found that

carbohydrates are normally *O*-linked to the amino acids serine and threonine or *N*-linked to asparagine.² Less common are *O*-linked 5-hydroxylysine, 4-hydroxyproline and tyrosine glycosides.³ The *O*-linked and *N*-linked glycans induce diverse functional roles, but can be broadly categorized: the *O*-linked glycopeptides typically have very large glycans, and are secreted outside the cell to form components of the extracellular matrix, as well as functioning as molecular lubricants (Table 1.2.3.1).⁴ The *N*-linked glycoproteins have smaller glycan structures, typically with a conserved core structure, but high variation in the outer regions.⁵ Examples of their function include protein trafficking, cellular communication, initiation of inflammation, and host defense.⁶

Many biological responses induced by carbohydrates are caused by direct interaction between the glycan, or carbohydrate portion of a glycopeptide or protein, and a protein receptor.⁷ Interestingly, in these cases, the glycan composition is more important than the site of carbohydrate attachment, or glycosylation.⁸ Carbohydrates can also have effects on the physical properties of proteins: such as solubility, viscosity, charge, conformation and stability.⁶ In these cases, the site of glycosylation has a much greater effect than the glycan composition.⁶ Here the focus will be on the effects of carbohydrates on peptide and protein conformation and stability.

1.2.4 Effects of Carbohydrates on Protein Conformation

The process of understanding the role of carbohydrates on peptide or protein conformation begins by knowing the location of the glycans on the structure of the peptide or protein structure. Surveys of *N*-linked glycoproteins by Dwek and coworkers have shown that glycans are more likely to be found projecting outwards from proteins

and not buried within them.⁶ Also, there is a greater chance of a glycosylation site being at a transition point between secondary structure types.⁶ These are signs that carbohydrates exert their influence on the protein surface, and not at sites located within a protein.

In some cases, glycosylation has no apparent effect on protein function, while in other cases glycosylation is critical for maintaining the native conformation, and thus function, of a protein.¹ Being able to explain and predict these functional consequences from glycosylation remains a major pursuit of glycobiology (the study of the structure, biosynthesis, and biology of saccharides and glycoconjugates in nature).⁹ To this end, it is common practice to synthesize and study fragments of proteins, *i.e.* peptides, to model and understand the conformational influence of glycosylation.

The influenza virus contains an antigenic glycoprotein, hemagglutinin, on its surface. An *N*-glycosylation site is found at a β -turn region of the protein, and is implicated in correct folding of the protein.¹⁰ Imperiali and coworkers were interested in understanding if the glycan (a tetradecasaccharide) induces the β -turn conformation. Instead of studying the full length protein, peptide fragments (9 amino acids) from the β -turn region were used to model the effects of glycosylation on peptide conformation.¹¹ The first two residues of the native glycan, $\beta(1-4)$ -linked *N*-acetylglucosamine (GlcNAc), were substituted by glucose (Glc) in various ways: GlcNAc disaccharides (**7**), Glc disaccharides (**8**), mixtures of the two (**9,10**), and a GalNAc monosaccharide (**11**) (Figure 1.2.4.1).¹²

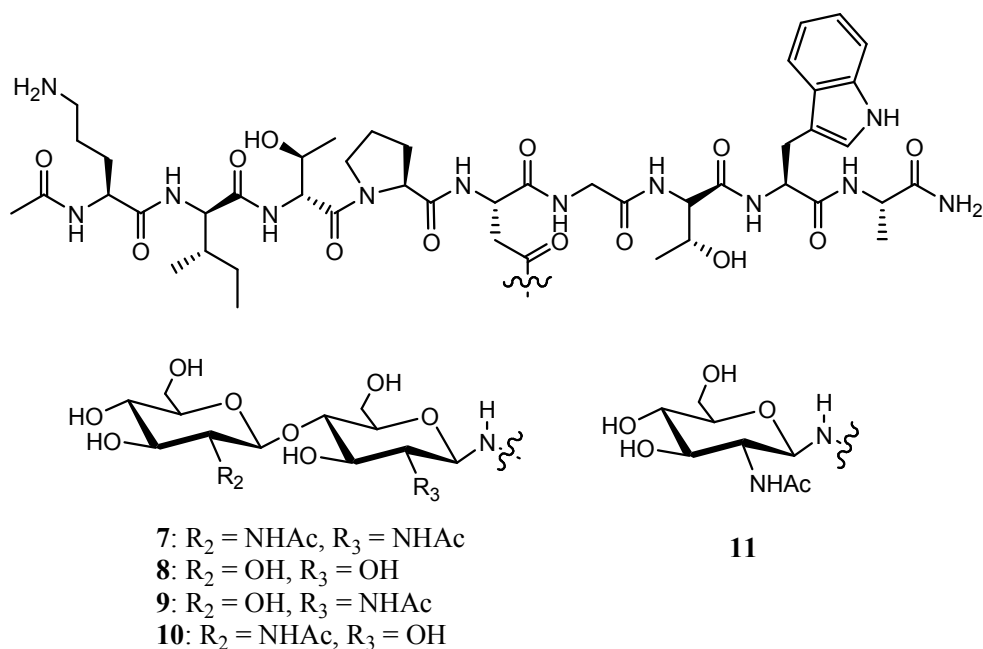


Figure 1.2.4.1: Glycosylated hemagglutinin model peptides.

The averaged conformation of the peptides was assigned based on nuclear Overhauser effect (nOe) nuclear magnetic resonance (NMR) data and peptide coupling constants ($^3J_{\text{HN}\alpha}$). The authors found that overall, glycosylation caused a conformational change in the peptide, from an extended conformation to a β -turn type conformation. The *N*-acetyl group of the sugar directly attached to asparagine was found to be critical for maintaining this β -turn conformation. Replacement of GlcNAc by Glc resulted in a loss of the β -turn conformation as identified by changes in nOe contacts. Similarly, the distal sugar was implicated in stabilizing the β -turn type conformation based on changes in backbone amide proton temperature coefficients and molecular modelling. Interestingly, there were no nOe contacts observed between the sugar and the peptide backbone. Thus, the influence of the carbohydrate on peptide conformation may be steric in origin, by limiting the conformational space available to the peptide, or the *N*-acetyl group may induce unique solvation properties which favour a more compact peptide conformation.

Further studies investigated the effect of the glycosidic linkage attached to asparagine.¹³ It was found that while the β -linked disaccharides induced a β -turn, the α -linked disaccharides did not. Clearly, small changes in the glycan can affect peptide conformation and rigidity. This example likely reflects a broader trend in biology; it has been shown that nearly one quarter of *N*-glycosylated asparagines are located at β -turn regions.^{6,14}

Mucin glycopeptides are characterized by having large, serially *O*- α -*N*-acetyl galactosamine (GalNAc)-linked oligosaccharides, which are primarily used by the body to retain water at surfaces exposed to the environment for lubrication and protection.^{4,15} Danishefsky and coworkers were interested in understanding the effects of these serially *O*-linked glycans on peptide conformation.¹⁶⁻¹⁸ The authors synthesized model pentapeptides (Ser-Thr-Thr-Ala-Val) with α -*O*-GalNAc-linked mono- (**13**) and disaccharides (**14** and **15**) at the serine and threonine residues (Figure 1.2.4.2).¹⁶

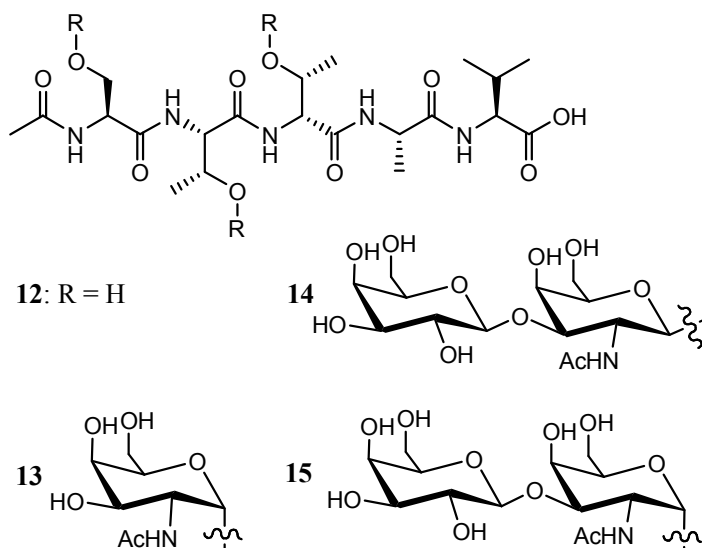


Figure 1.2.4.2: Clustered *O*-mucin glycosylated model peptides.

It was found by a combination of NMR experiments and computational calculations that the clustering of these *O*-linked glycans induced an extended conformation on the peptide. This was based on the large number of nOe contacts between the proximal *O*- α -GalNAc residue and the peptide backbone, a reduction in GalNAc amide proton chemical shift temperature dependencies, and $^3J_{\text{NH-H}\alpha}$ coupling constants inconsistent with extensive conformational averaging. Computational calculations indicated that the structural rigidity could be attributed to hydrogen-bonding interactions between the sugar and the peptide backbone. The authors postulated that the extended structure induced by *O*- α -GalNAc facilitates recognition events involving the glycan.

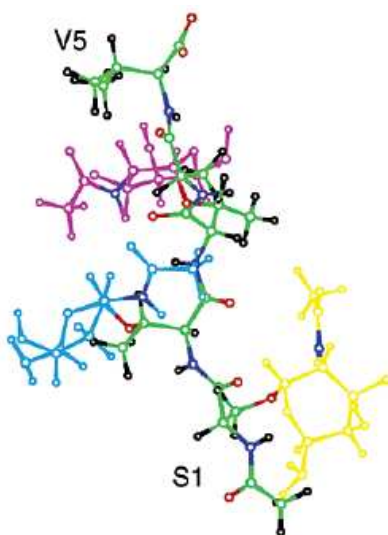


Figure 1.2.4.3: Clustering of α -*O*-GalNAc glycans (coloured purple, light blue, and yellow) caused a stable, extended structure in the peptide backbone (coloured green, dark blue, black, and red) of **13**; Ser1 and Val5 are labelled. [Reproduced with permission, *J. Am. Chem. Soc.* **2002**, *124*, 9833. Copyright 2002 American Chemical Society.]

This study also provides insight into the different influences from the inner and outer glycan residues and the native α - and abnormal β -linked sugars on peptide structure and stability. Interestingly, it was found that sugars attached beyond the first α -*O*-GalNAc residue had no influence on the structure of the peptide. The distinct nOe fingerprint did not change as the number of sugar residues was increased from one to three. Also, while α -*O*-GalNAc was found to be critical for structural organization, the β -*O*-linked sugars did not seem to induce any rigidity in the peptide. Based on uniformity of ^1H and ^{15}N chemical shifts, temperature dependence of amide proton chemical shifts, and comparison of backbone and side-chain coupling constants, the β -*O*-linked model peptide **14** resembled the non-glycosylated form **12** with a high degree of flexibility. Therefore, the anomeric linkage has a large influence on the structural properties of the peptide.

These studies provide insight into the changes that can occur in protein secondary structure as a result of small changes in the glycan structure, such as the number of residues in the glycan, or a change in the anomeric linkage to the peptide or protein. It seems carbohydrates can both induce turns and rigidify the residues proximal to the glycosylation site. Because of these traits, carbohydrates may well play a crucial role in the protein folding pathway. The control of local environment has also proven to be of importance for the overall stability of a protein.

1.2.5 Effects of Carbohydrates on Protein Stability

Glycans affect not only the conformation, but also the stability of glycoproteins. Studies have found that removal of a glycan from a glycoprotein can make the protein

more susceptible to thermal denaturation¹⁹⁻²¹ and proteolytic degradation.²²⁻²⁶ As with the influence on conformation, even small changes in the glycan composition can have a large effect on protein stability. For example, PMP-C is a small protein consisting of 36 amino acids, which is normally folded into a three-stranded β -sheet, and has a single L-fucose residue *O*-linked to threonine in the sequence -Gly⁷-Lys-Thr*-Phe-Lys¹¹- at the start of a β -strand.¹⁹ The protein acts by inhibiting the function of proteases.²⁷

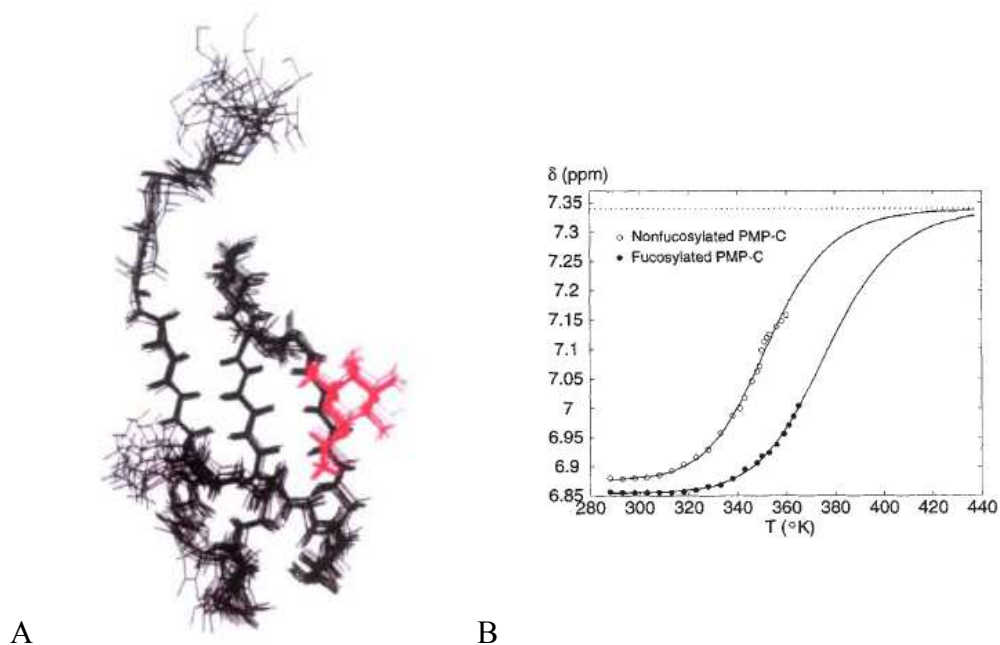


Figure 1.2.5.1: A) Sum of 14 refined structures of fucosylated PMP-C, where the L-fucose residue is in red and the polypeptide backbone is in black; B) Thermal denaturation profiles of fucosylated and non-fucosylated PMP-C according to the chemical shift of Phe₁₀ H₄. [Reproduced with permission, *Nature Structural Biology* **1996**, 3, 45. Copyright 1996 Nature Publishing Group.]

The effects of the L-fucose residue on the protein were studied by comparing synthetic proteins with and without the attached L-fucose residue. Using NMR measurements and computational calculations, it was found that removal of the sugar had

almost no impact on the protein's conformation. Careful analysis of nOe interactions showed that the single L-fucose residue had a short range of influence on the protein (~6 Å); the L-fucose residue had nOe contacts to the adjacent amino acids phenylalanine₁₀ and lysine₁₁ as well as arginine₁₈ in a neighbouring β -strand (Figure 1.2.5.1A). Slower proton-deuteron exchange rates of amide protons involved in inter-strand hydrogen bonds in the vicinity of the L-fucose residue were indicative that fucosylation stabilizes the folded state of PMP-C. Interestingly, removal of the sugar did have an impact on the thermal stability of the protein, as determined by monitoring the change in chemical shift of H4 of Phe₁₀ as a function of temperature (Figure 1.2.5.1B). The H4 resonance of Phe₁₀ was chosen because it is buried in the hydrophobic core of PMP-C in the folded state. It was observed that the non-glycosylated protein denatured at a temperature 20 degrees lower than the fucosylated form, which is a significant difference in thermal stability. Therefore, despite the seemingly short range of influence from the L-fucose residue, stabilization of one region of the β -sheet through local interactions affects the stability of the whole protein, by shifting the equilibrium to the folded state.

Proteins have a finite lifetime *in vivo*, and are constantly being recycled through the action of proteases. There are indications that glycosylation may serve to protect proteins from proteolytic degradation and act as a mechanism by which organisms can regulate protein function.^{28,29} The human chorionic gonadotropin hormone (hCG) is a 36.7 kDa heterodimeric glycoprotein composed of an α - and β -subunit.^{30,31} It has critical functions during pregnancy to maintain progesterone production.^{32,33} Analysis of hCG has found that the α -subunit is *N*-glycosylated at Asn52 and Asn78, where the glycan attached to Asn52 appears to be involved in signal transduction and the subunit

association process.³⁴ In contrast, the glycan attached to Asn78 has not been shown to play any direct biological role, but in a similar approach to PMP-C, comparison of the glycosylated and non-glycosylated forms of the α -subunit has shown that removal of the glycan attached to Asn78 leads to rapid degradation *in vivo*, and also decreases the thermal stability of the α -subunit.³⁵ NMR experiments indicated that the glycan linked to Asn78 forms extensive interactions with the protein.^{36,37} The majority of these interactions are hydrophobic in nature, and involve the side chains of the neighbouring amino acids leucine, isoleucine, and valine. Furthermore, the sugar appears to shield a hydrophobic pocket between two β -hairpins.³⁷ Therefore, the carbohydrate may act to disrupt the association of proteases to key sites on the protein. These studies serve as examples for the manner in which glycosylation, even by a single carbohydrate residue, can increase the thermal stability of a protein and its resistance to proteolytic degradation through local interactions to the protein surface. Thus, this evidence suggests the possibility of using glycosylation to modulate the stability of proteins that are not normally glycosylated.

1.3 The Development of Synthetic Glycopeptides

Interest in the development of synthetic glycopeptides has arisen for several reasons. As the science of glycobiology has matured, so has a desire to understand the roles of specific glycoproteins. Synthetic peptides are required for several reasons: one, glycopeptides are present in very minute quantities *in vivo*, making isolation of a given glycoprotein in sufficient quantities quite difficult. Second, each glycopeptide can exist in many different glycoforms, which exacerbates the problem of isolation of specific

glycopeptides.¹ For example, erythropoietin is a clinically useful stimulant of red blood cells for treatment of anaemia, but recombinant expression in Chinese hamster ovary cells produces 13 different glycoforms; this complicates the isolation and characterization process.³⁸ Thus, emphasis has shifted from isolation to the chemical synthesis of glycoproteins and glycopeptides;³⁹⁻⁴³ this can ensure the production of sufficient quantities of glycopeptides with well-defined glycan structures, in order to correlate a specific glycan with a physiological outcome.

1.3.1 C-linked Glycopeptides Increase Proteolytic Stability

Since glycopeptides are involved in many important biological processes, there is also interest in developing synthetic glycopeptide-based therapeutics. However, all glycopeptides are susceptible *in vivo* to the action of glycosidases, which are enzymes that cleave the glycosidic bond.^{2,7} Also, the glycosidic bond is vulnerable to the often-harsh conditions used in the synthetic chemistry. These drawbacks have spurred interest in developing unnaturally linked glycopeptides.^{44,45} Replacement of the oxygen or nitrogen atom in a glycosidic linkage by a carbon atom should effectively reduce all chance of hydrolysis. The so-called C-glycopeptides are expected to be completely stable to enzymatic cleavage by glycosidases under physiological conditions. Interestingly however, there are very few studies that have actually addressed this assumption.

Glycoamidase, or N-glycanase is an enzyme known to cleave intact glycans from N-glycopeptides and N-glycoproteins through cleavage of the β -aspartyl-glucosylamine amide linkage.^{46,47} In a study of N-glycanase, Wang and coworkers compared the activity of the enzyme against a high-mannose N-glycopentapeptide and its C-linked

counterpart.⁴⁸ The structural features of the selected *N*-glycopeptide and its *C*-glycopeptide counterpart were identical, except for the replacement of the nitrogen atom in the glycosidic linkage with a methylene group (Figure 1.3.1.1).

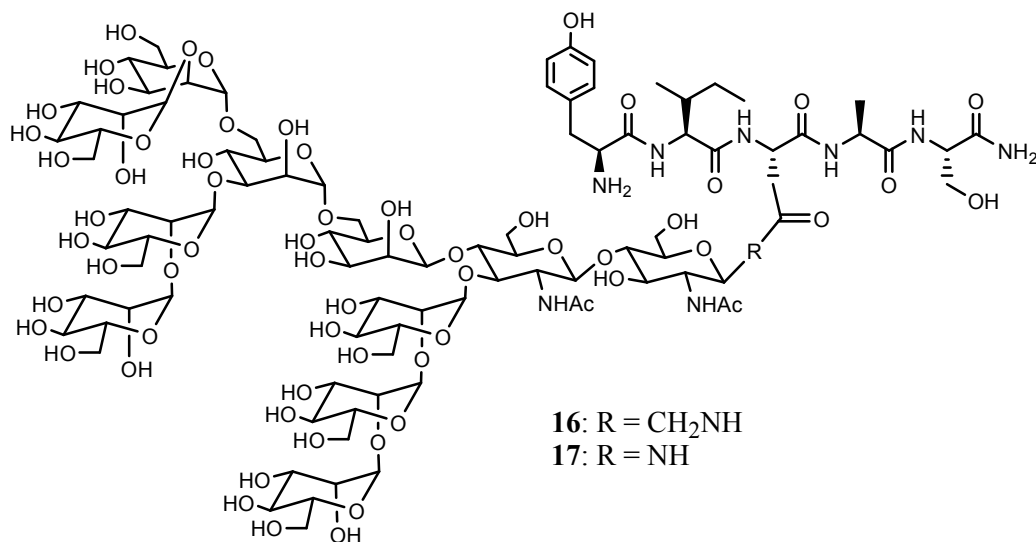
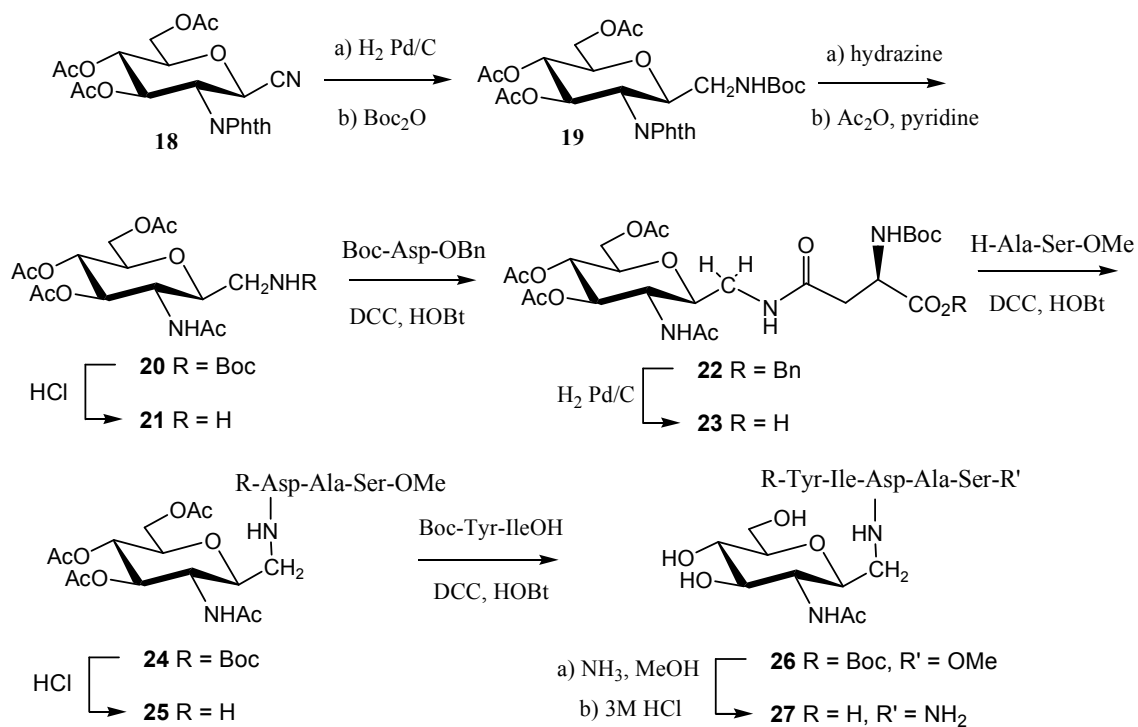


Figure 1.3.1.1: High-mannose *N*-glycopeptide **17** and its *C*-linked counterpart **16**.

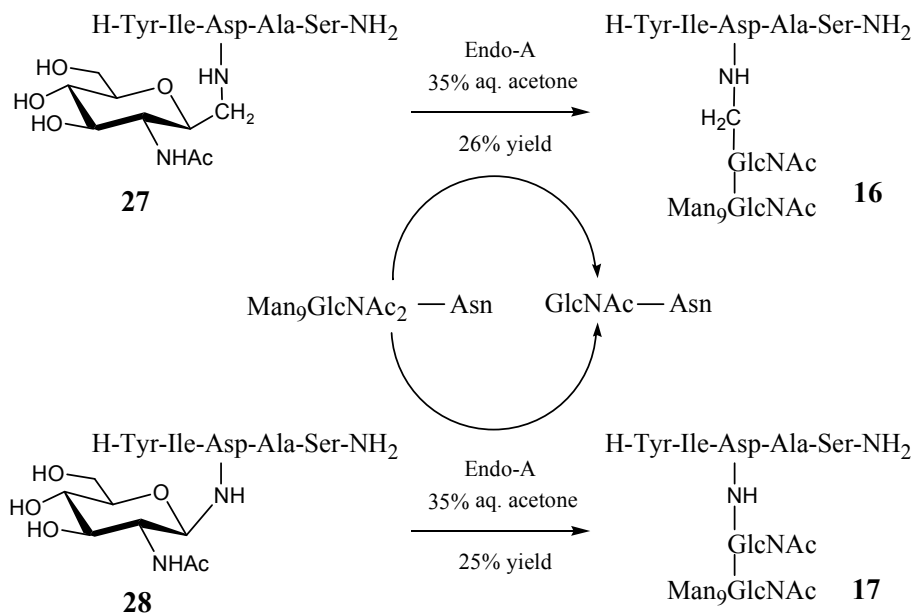
The synthetic approach for the creation of the *C*-linked glycopeptide was based on synthesis of a *C*-glycosyl amino acid building block for incorporation into a model pentapeptide (H-Tyr-Ile-Asp-Ala-Ser-NH₂) using fragment condensations (Scheme 1.3.1.1). Catalytic hydrogenation of β -glycosyl cyanide **18** followed by *N*-Boc protection formed the protected glycosylmethylamine **19**. Installation of the 2-position *N*-acetyl group and *N*-Boc deprotection was followed by coupling to Boc-Asp-OBn to give the *C*-glycosyl amino acid **22**. Successive deprotection and fragment condensation steps at the *C*- and *N*-termini of the *C*-glycosamino acid building block formed the desired *C*-linked GlcNAc pentapeptide **27**.



Scheme 1.3.1.1: Synthesis of *C*-glycopeptide **27**.⁴⁸

In contrast, for the *N*-glycosyl pentapeptide, the glycosylation step occurred after the pentapeptide was synthesized: The free β -carboxyl group of the aspartic acid residue was coupled to 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy- β -D-glucopyranosylamine using HBTU to form the protected β -linked GlcNAc-*N*-pentapeptide. Deprotection afforded the naturally-linked model pentapeptide **28**.

For both the *N*- and *C*-glycopeptides, the high mannose Man₉GlcNAc oligosaccharide was installed in a single enzymatic transfer using the enzyme endo- β -N-acetylglucosaminidase (Endo-A) isolated from *Arthrobacter protophormiae*. The enzyme is capable of transferring Man₅₋₉GlcNAc glycans selectively to the 4-position hydroxyl group of a terminal GlcNAc residue. To their delight, the authors found that the transglycosylation of Man₉GlcNAc proceeded in 26% and 25% yield for the *C*- and *N*-glycopeptides **16** and **17** after purification, respectively (Scheme 1.3.1.2).



Scheme 1.3.1.2: Transglycosylation step of **27** and **28** to make **16** and **17**, respectively.⁴⁸

To determine the stability of the *C*-glycopeptide toward *N*-glycanases, the glycopeptide was incubated with various *N*-glycanases isolated from plant, bacteria and animal sources. In all cases, the authors found no changes to the *C*-glycopeptide after enzymatic incubation, or any presence of any hydrolytic products according to HPLC chromatography. In contrast, and as expected, the high mannose *N*-glycopeptide was found to be rapidly hydrolysed. Furthermore, the *C*-glycopeptide showed inhibitory activity (K_i) ranging from 1 to 160 μM for the different *N*-glycanases, indicating the glycopeptides can still bind to the enzyme binding site, but the cleavage step cannot occur. This study confirms that replacement of the natural *N*-glycosidic linkage by a *C*-glycosidic linkage confers enzymatic resistance to glycosidases, such as *N*-glycanase.

1.3.2 C-Linked Glycopeptides Provide Insight into Biological Activity

While *C*-glycosides provide increased resistance to enzymatic degradation, there are slight differences between *O*- and *C*-glycosides: a difference in dipole moment, a loss of the anomeric effect, and the loss of the ability to form hydrogen bonds.⁴⁹⁻⁵⁴ These discrepancies can either be viewed as a drawback of the analogues, or as in the following example, another means to understand the influence of the *O*-glycosidic linkage on biological activity.

Rheumatoid arthritis is an autoimmune disease in which collagen, a key structural protein, is degraded, causing painful swelling of the joints.⁵⁵ This results in bone erosion in peripheral joints. It is believed that collagen-derived peptide fragments are involved in stimulating the autoimmune response. The sequence of amino acids required for stimulating an immune response, or epitope, includes a β -*O*-D-galactosyl-5*R*-hydroxy-L-lysine residue (**29** in Figure 1.3.2.1).⁵⁶⁻⁵⁸ While uncommon for most proteins, this is a common post-translational modification in collagen.

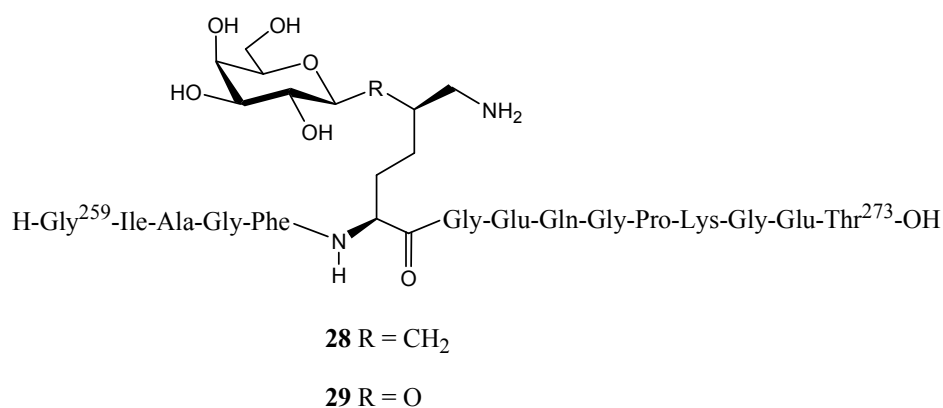
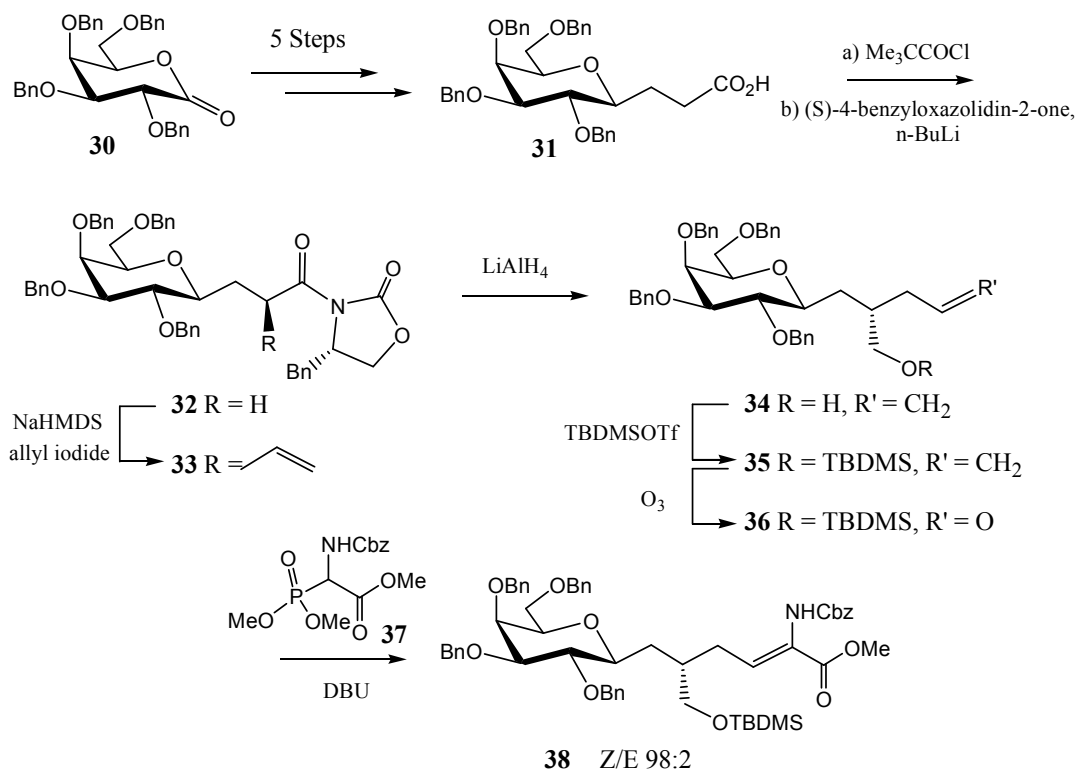


Figure 1.3.2.1: Model collagen epitope *C*- and *O*-glycopeptides **28** and **29**, respectively.

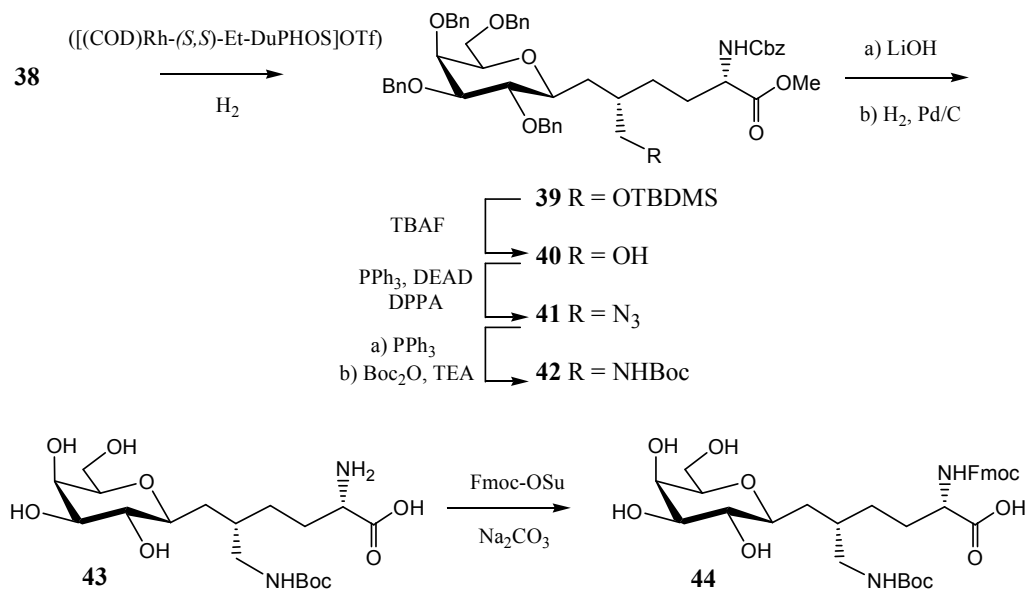
Having first demonstrated that the galactosylated hydroxylysine residue is critical for recognition by the T-cell receptor, Kihlberg and coworkers were interested in the impact of switching the *O*-linkage with a *C*-linked sugar (Figure 1.3.2.1).⁵⁹ Thus, they synthesized a *C*-galactosyl hydroxylysine building block for incorporation into a model peptide in order to further evaluate the requirements for T-cell recognition, specifically related to the carbohydrate moiety.

After consideration of several synthetic routes, the authors began by reacting the benzyl protected galactopyranosyl lactone **30** to form the carboxylic acid **31** in five steps (Scheme 1.3.2.1).⁶⁰ This formed solely the β -anomer product, which is the same as the natural linkage.



Scheme 1.3.2.1: Synthesis of protected *C*-hydroxylysine building block **38**.

Ozonolysis of the β -linked terminal alkene, followed by Jones oxidation of the resulting aldehyde gave the desired carboxylic acid **31**, which was converted to a mixed anhydride derivative on treatment with pivaloyl chloride. Reaction with Evans chiral (*S*)-4-benzyl-oxazolidin-2-one in the presence of *n*-butyllithium gave **32**, which allowed a single diastereomer **33** to be formed in 88% yield upon treatment with allyl iodide and sodium hexamethyldisilazane. The Evans chiral auxiliary was subsequently removed using lithium aluminumhydride, and the primary alcohol was protected using the TBDMS group to give **35**. Ozonolysis of the terminal alkene followed by Horner-Emmons olefination using a protected glycine phosphonate **37** gave a 98:2 mixture of *Z/E* isomers **38**. Asymmetric hydrogenation of **38** using Burk's catalyst ($[\text{Rh}-(S,S)\text{-Et-DuPHOS}]\text{OTf}$) gave >99:1 stereoselectivity for the 2*S* stereoisomer **39** (Scheme 1.3.2.2). Installation of the Boc-protected primary amino group over four steps to give **42** and further protecting group manipulation of the α -amino acid portion, produced the target *C*-glycosylated building block **44**.



Scheme 1.3.2.2: Synthesis of *C*-hydroxylysine building block **44**.

Incorporation of the *C*-glycosylated building block **44** into the peptide epitope H-Gly-Ile-Ala-Gly-Phe-Hyl*-Gly-Glu-Gln-Gly-Pro-Lys-Gly-Glu-Thr-OH **28** derived from collagen, known to induce a T-cell response, was accomplished by solid-phase peptide synthesis using an Fmoc-strategy. The naturally *O*-linked galactosylated hydroxylysine glycopeptide counterpart **29** was also synthesized in this manner.

Using engineered cells (hybridomas) that express T-cells specific for collagen peptides as a model for collagen-induced arthritis, the authors were able to quantitatively monitor the immune response elicited upon exposure to the *O*- and *C*-glycopeptides. It was found that in a panel of hybridomas, 10 to 20-fold higher concentrations of the *C*-glycopeptide **28** were required to elicit the same immune response compared to the naturally *O*-linked glycopeptide **29**. The authors concluded that the *O*-glycosidic linkage is important for T-cell binding, and demonstrates the sensitivity of the T-cell receptor binding site in distinguishing between a methylene and glycosidic oxygen atom. Despite the lower activity, the *C*-glycopeptide is expected to be more resistant to enzymatic breakdown, and so an outcome of this work may be the development of robust collagen analogues as competitive inhibitors of T-cells for the therapeutic treatment of rheumatoid arthritis.⁵⁸ As can be seen by these examples, *C*-glycopeptides have shown their merit, and this has resulted in an increased demand for non-natural linkages of carbohydrates and amino acids from glycobiologists as leads for developing bioactive compounds.

1.4 C-Glycosyl Amino Acids

1.4.1 The Extension of the Repertoire of C-Glycosylated Amino Acids

In order to meet the demand for unnaturally glycosylated amino acids, the past 20 years have seen a great increase in the number of *C*-glycosylated amino acids, beyond the previous examples in which a methylene group has replaced the *O*- or *N*-glycosidic linkage.⁶¹⁻⁶⁴ Dondoni defines *C*-glycosyl amino acids as an α -amino acid moiety carbon-linked to the anomeric carbon of a carbohydrate, with part or all of the amino acid side chain being incorporated into the carbohydrate frame (Figure 1.4.1.1).⁶¹

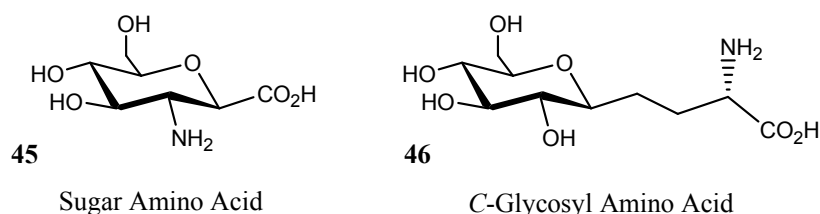


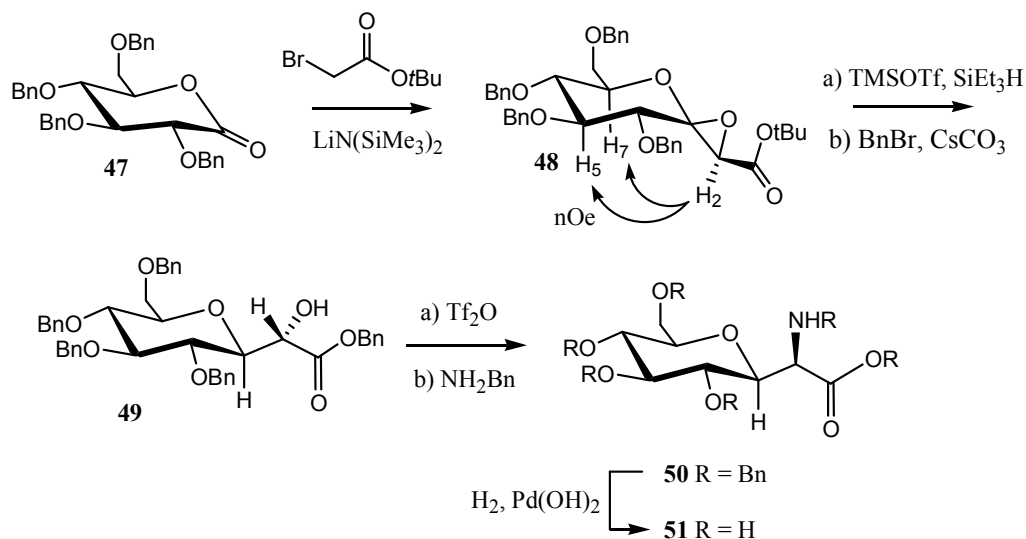
Figure 1.4.1.1: Comparison of SAAs and *C*-glycosyl amino acids

Kessler has expanded on this definition, such that an amino and carboxyl group directly attached to a carbohydrate frame is referred to as a sugar amino acid (SAA),⁶² or using Schweizer's terminology, a sugar amino acid hybrid (SAAH).⁶⁵ Since this leads to a large class of compounds, the focus here will be on Dondoni's classification of *C*-glycosyl amino acids.

1.4.2 Synthesis of a *C*-Glycosyl Glycine Amino Acid

In 2001, Schweizer and Inazu developed a method for the synthesis of *C*-glycosyl glycines, which relied on an enolate addition of *tert*-butyl bromoacetate to protected *D*-glycopyranonolactones, followed by amino substitution.⁶⁶ For example, 2,3,4,6-tetra-*O*-benzyl-*D*-glucono-1,5-lactone **47** was reacted with the enolate formed by the

deprotonation of *tert*-butyl bromoacetate with lithium bis(trimethylsilyl)amide at $-78\text{ }^{\circ}\text{C}$, in 81% yield (Scheme 1.4.2.1).

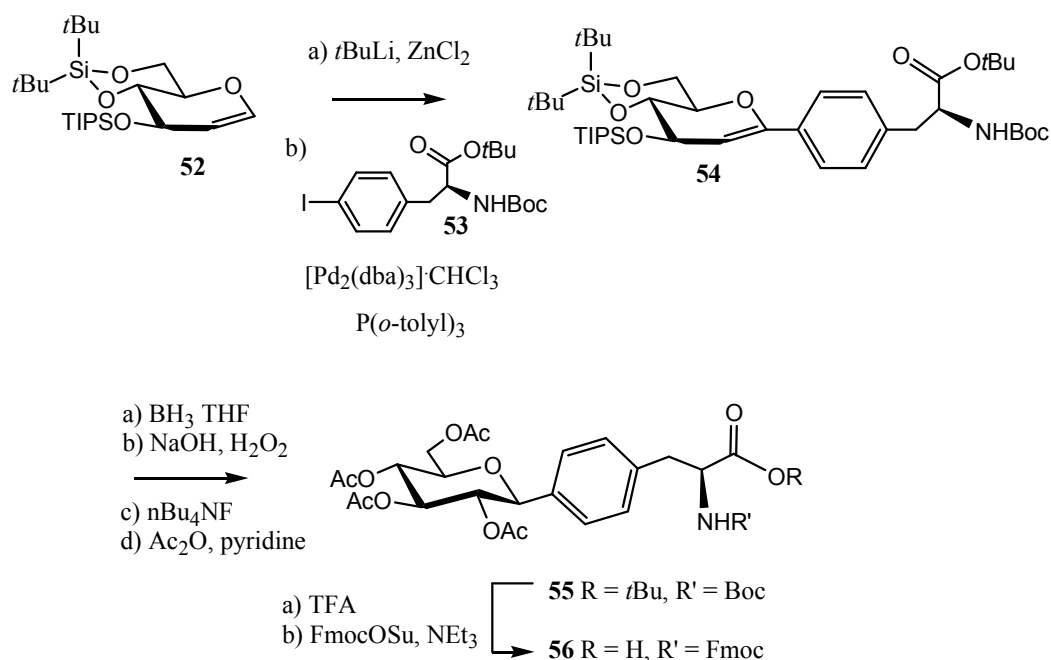


Scheme 1.4.2.1: Synthesis of Schweizer and Inazu's C-glycosyl glycine **51**.

The stereochemistry of the spirocyclic product **48** was assigned using *nOe* contacts between the H_2 to the H_5 and H_7 sugar protons. Regioselective epoxide opening using triethylsilane under TMSOTf-promoted conditions at $-78\text{ }^{\circ}\text{C}$ afforded the functionalized C-glycoside **49** in 50% yield. Conversion to the C-glycosyl glycine ester **50** in 80% yield was accomplished using trifluoromethanesulfonic anhydride in pyridine to activate the secondary alcohol, followed by nucleophilic displacement of the sulfonate ester using benzylamine. The '*R*' stereochemistry of the α -amino group was deduced using a combination of molecular mechanics calculations, circular dichroism (CD), and NMR experiments. Deprotection of the amine and sugar hydroxyl groups was carried out by hydrogenation using Pearlman's catalyst ($\text{Pd}(\text{OH})_2$) in acidified methanol in quantitative yield to give the target building block **51**.

1.4.3 Synthesis of a C-Glycosyl Phenylalanine Amino Acid

In 2005, Ousmer and coworkers developed a C-glycosyl phenylalanine derivative, which relied on a Negishi cross-coupling reaction of a silyl protected D-glucal **52** and *N*-Boc-p-iodo-L-phenylalanine *tert*-butyl ester **53** (Scheme 1.4.3.1);⁶⁷ both are accessible using established high-yielding methods.^{68,69} Use of silyl protecting groups was required for compatibility with the cross-coupling conditions.



Scheme 1.4.3.1: Synthesis of Ousmer's C-glycosyl phenylalanine building block.

Treatment of the glucal **52** with *tert*-butyl lithium and zinc dichloride gave the transient organozinc derivative, which was coupled to Boc-Phe(4-I)-*Ot*Bu **53** using [Pd₂(dba)₃]·CHCl₃ and the bulky tri-*o*-tolylphosphane to give the desired product **54** in 90% yield. Regio- and stereoselective hydroxylation of the double bond was carried out by hydroboration, which was followed by oxidative quenching. The β-configuration of the C-glucoside was confirmed by a coupling constant of 9.6 Hz from the H₁ and H₂

protons of the carbohydrate ring. Protecting group manipulation of the sugar hydroxyl groups and the amino acid moiety gave Fmoc-protected *p*-(*C*-glucopyranosyl)-phenylalanine **56**, suitable for use in SPPS.

These, and other *C*-glycosyl amino acids provide the opportunity to probe several concepts: first, to study the effects of glycosylating amino acids in peptides and proteins that are not normally glycosylated; second, the opportunity to study the functional consequences of incorporating an amino acid side chain into the carbohydrate structure, thereby restricting its conformational freedom.

Interestingly, while there are many examples of *C*-glycosyl glycines, alanines, tyrosines, phenylalanines and tryptophans, almost no proline-based *C*-glycosyl amino acids exist.^{70,71} This is somewhat surprising as proline is unique among amino acids and plays specialized roles in protein function. The remainder of this chapter will be focused on proline: both its importance in biology and how proline analogues are used to understand and affect its properties.

1.5 L-Proline

1.5.1 L-Proline is Unique Among Amino Acids

Consider the structure of the 20 proteinogenic amino acids: they are defined by an amino and carboxyl group attached to an α -carbon. A side chain also projects from the α -carbon, which aside from glycine, makes the amino acid α -carbon a stereogenic center. Therefore, they can have '*R*' or '*S*' stereochemistry, which corresponds to the D- and L-amino acids respectively. The L-amino acids predominate in nature, with few examples of D-amino acids.

Proline is the only proteinogenic amino acid in which the side chain is cyclised onto the backbone nitrogen. This formally makes proline a secondary amine, and confers special properties to proline when incorporated into a polypeptide. While flexible, the pyrrolidine ring of proline does restrict the movement of the atoms in the proline side chain. This is in part because the ϕ -angle, which can be freely rotated in all other amino acids, is fixed at -75° (Figure 1.5.1.1).⁷²

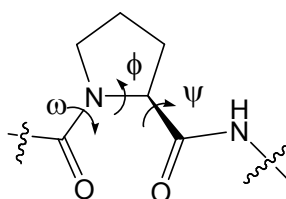


Figure 1.5.1.1: Proline backbone dihedral angles.

A survey of proline residues in the crystal structures of 50 different proteins indicated that the proline side chain commonly adopts two distinct conformations.⁷² Proline is puckered in an envelope conformation, where the β - or γ -carbon project out of the plane defined by the other atoms in the pyrrolidine ring (Figure 1.5.1.2).

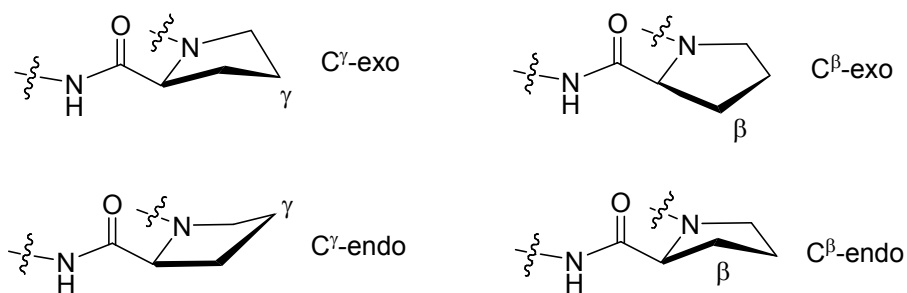


Figure 1.5.1.2: Proline γ - and β -exo and endo puckers; the γ - and β -atoms are labelled.

If the γ -carbon projects from the opposite face to the prolyl carboxyl carbon, it is referred to as C^γ -exo. If the γ -carbon is projected onto the same face as the prolyl

carboxyl carbon then it is referred to as C^γ-endo. The same concept applies to the β-puckers. L-Proline has a slight preference for adopting a C^γ-endo conformation.⁷³⁻⁷⁹

The amide bonds that link amino acids in a polypeptide chain have a pseudo-double bond character; through resonance stabilization, π-electrons are delocalised across the amide bond, which forces a planar ω-torsional angle of the amide bond. This creates two distinct amide conformations; the amino acid α-carbons can either be *cis* (ω is 0°) or *trans* (ω is 180°) with respect to each other across the amide bond (Figure 1.5.1.3). While the amide bond has double bond character, the term ‘conformation’ seems to be more appropriate than ‘configuration’ since the amide C-N bond is formally a single bond.

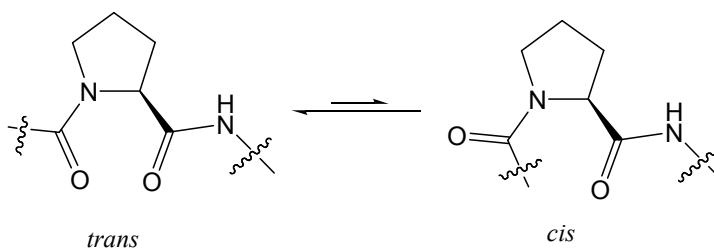


Figure 1.5.1.3: Proline *cis* and *trans* amide isomers.

It has been calculated that for most amide bonds in peptides and proteins, the *trans* amide conformation is 8.3 kcal·mol⁻¹ lower in energy than the *cis* amide conformation.⁸⁰ This results in a strong bias for the *trans* conformation for amide bonds, with only 0.3% of amide bonds existing in the *cis* conformation in peptides at any one time (this corresponds to an amide equilibrium value $K_{trans/cis}$ of 332.3).⁸¹⁻⁸² In contrast, for proline, with a cyclized backbone, the *trans* amide isomer is only 1.9 kcal·mol⁻¹ lower in energy compared to the *cis* amide isomer.⁸⁰ This corresponds to nearly 5% *cis* amide conformation at any one time ($K_{trans/cis}$ of 19.0), which is a significant difference.^{82,83} The

origin of the smaller energetic difference between *trans* and *cis* amide bonds in proline compared to other amino acids is a result of several factors.

1.5.2 Thermodynamics of Prolyl Amide Isomerization

Comparative studies of model peptides in the form of Ac-Gly- X_{aa} -NHMe have shown that when X_{aa} = glycine, the *trans* conformation is approximately 8 kcal·mol⁻¹ lower in energy relative to the *cis* conformer as a result of four factors: steric repulsion between the α -carbons (1 kcal·mol⁻¹), steric repulsion between the hydrogen atoms attached to the α -carbons (4-5 kcal·mol⁻¹) (Figure 1.5.2.1), favourable hydrogen bonds only possible in the *trans* form (1-2 kcal·mol⁻¹) and a conformational entropy difference (1-2 kcal·mol⁻¹).⁸⁰ The entropy difference arises from different hydration shells for the *cis* and *trans* isomers, with the *cis* conformer having a slightly more ordered solvent environment.

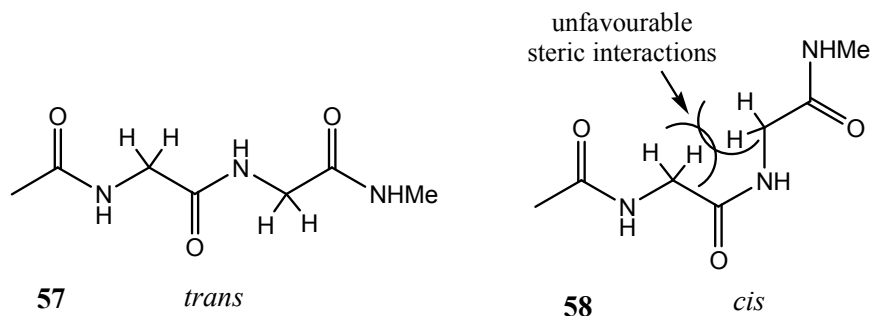


Figure 1.5.2.1: The *trans* and *cis* amide isomers of Ac-Gly-Gly-NHMe; the *cis* amide form suffers from an unfavourable steric interactions.

Conversely, when X_{aa} = proline, there is less energetic discrepancy between the *cis* and *trans* forms since the cyclized side chain means that the α -carbon and δ -carbon

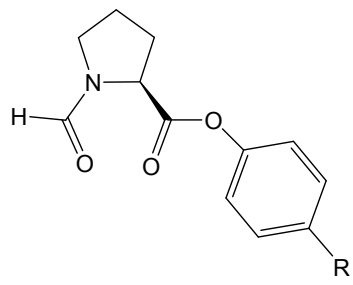
provide equivalent steric environments for the *N*-terminal amino acid α -carbon. The only energetic difference between the proline *cis* and *trans* isomers arises from a favourable $n \rightarrow \pi^*$ interaction from the lone pair on the prolyl *N*-terminal amide carbonyl oxygen to antibonding orbital of the prolyl *C*-terminal carbonyl carbon (Figure 1.5.2.2).⁸⁰ This interaction is only possible for the *trans* amide conformation and accounts for the 2 kcal·mol⁻¹ energy difference between the *cis* and *trans* isomers, and is aligned because the pyrrolidine ring fixes the C _{α} -N bond ϕ -torsional angle at -75 °.⁷²

See Figure 1 on page 1189 of Hinderaker, M. P.; Raines, R.T. *Prot. Sci.* **2003**, *12*, 1188-1194.

Figure 1.5.2.2: A) Proposed $n \rightarrow \pi^*$ electrostatic interaction present in the *trans* amide of *N*-formyl-L-proline methyl ester; B) Depiction of the n and π^* molecular orbitals.

This effect was further explored by Hodges and Raines, who proposed that the correct orientation for optimal $n \rightarrow \pi^*$ interaction from the prolyl *N*-terminal amide carbonyl oxygen to the *C*-terminal carbonyl carbon is analogous to the approach of a nucleophile to the electrophilic carbon of an acyl group, and is strongest at the Bürgi-Dunitz trajectory.⁸⁵ They reasoned that if the *trans* isomer is stabilized through such an electrostatic interaction, it should be possible to modulate the population of the *trans*

amide isomer by changing the electrophilicity of the *C*-terminal carbonyl carbon using para-substituted phenyl esters of *N*-formyl proline (Figure 1.5.2.3).



R	$K_{trans/cis}$
NO ₂	4.13
CN	3.95
H	2.12
OMe	2.01
NMe ₂	1.67

Figure 1.5.2.3: Effect of substituted phenyl esters on *trans/cis* amide ratio ($K_{trans/cis}$) in CDCl₃ at 25 °C.

Through their study, Hodges and Raines confirmed that the nature of the para-substituent does affect the *N*-terminal *trans/cis* amide isomer equilibrium ($K_{trans/cis}$): when X = NO₂, an electron-withdrawing group, $K_{trans/cis}$ was 4.13, whereas an electron donating group such as X = NMe₂, gave a $K_{trans/cis}$ of 1.67. The electron withdrawing substituent is implicated in lowering the energy of the π^* orbital, bringing it closer in energy to the *n* orbital, thereby enhancing *n* \rightarrow π^* electron donation and stabilization of the *trans* amide isomer. The opposite is true for the electron donating substituents. Therefore, this work supports the role of an *n* \rightarrow π^* electrostatic interaction in influencing the *N*-terminal amide equilibrium.

In peptides and proteins, the amino acid *N*-terminal to proline has a significant influence on the *cis/trans* amide equilibrium.⁸⁵ Analysis of crystallographic databases supports much higher prolyl *cis* amide populations when aromatic residues are *N*-terminal to proline.⁸¹ Taylor and coworkers used model peptides of the form Ac-X_{aa}-Pro-

OMe to investigate the origins of a stabilizing effect on the *cis* amide conformation from aromatic residues in the X_{aa} position.⁸⁶ Originally thought to arise through steric effects, Taylor instead proposed favourable overlap between the aromatic side chain and prolyl pyrrolidine ring in **59** (Figure 1.5.2.4).

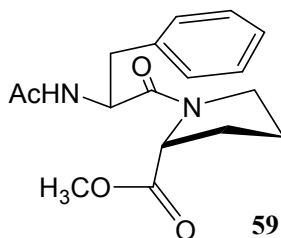


Figure 1.5.2.4: Proposed aromatic overlap of proline side chain that stabilizes the prolyl *cis* *N*-terminal amide conformation in Taylor's model peptide.

A large difference in the spatial orientation of the side chain exists between the *cis* and *trans* amide conformations, and this means that this overlap is only possible in the *cis* amide conformation. Evidence of the aromatic overlap of the proline side chain is provided by NMR studies, which shows that there is a much higher upfield shift of the proline α -hydrogen for the *cis* amide isomer relative to the *trans* isomer. This is explained as arising through shielding of the α -hydrogen by the aromatic ring.

Polar amino acids, such as serine or threonine, *N*-terminal to proline have less of an effect on the prolyl amide equilibrium.⁸¹ However, glycosylation or phosphorylation of serine or threonine *N*-terminal to proline is known to have an effect on $K_{trans/cis}$.^{87,88} Pao and coworkers believed that glycosylation of serine *N*-terminal to proline could induce high prolyl *cis* amide content in a similar fashion to aromatic amino acids, which they perceived to arise via a steric influence.⁸⁷ This study was done to help further define the role of glycosylation on peptide and protein conformational properties.

Using analogues of a Sendai Viral Nucleoprotein antigen, peptides of the form H-Phe-Ala-Pro-Ser-Asn- X_{aa} -Pro-Ala-Leu-OH were synthesized using solid phase peptide synthesis, where the X_{aa} was Ser, Ser(α -GalNAc) or Ser(β -GlcNAc).⁸⁷ Both anomers were studied to determine if the anomeric linkage affected the *trans/cis* equilibrium. Analysis by NMR found that the peptides adopted solely a *trans* amide conformation *N*-terminal to proline. Therefore glycosylation *N*-terminal to proline greatly stabilizes the *trans* amide conformation, however, the authors provided little explanation for this phenomenon.

In summary, as opposed to other amino acids, the *N*-terminal amide equilibrium of proline is governed only by a favourable $n \rightarrow \pi^*$ electrostatic interaction. Amino acids *N*-terminal to proline can also greatly affect this equilibrium, though a full understanding of their influence has not been achieved.

1.5.3 The Kinetics of Prolyl Amide Isomerization

The *N*-terminal amide isomerization of proline in peptides and proteins is known to be very slow compared to other amino acids. Analysis of model peptides has found that for the typical secondary amide bond, such as an alanine-phenylalanine amide bond, the rate constants for *cis-trans* isomerization are 0.05 sec^{-1} ($k_{trans \rightarrow cis}$) and 2.3 sec^{-1} ($k_{cis \rightarrow trans}$) in water at 25 °C.⁸⁹ In contrast, the typical isomerization rates for an alanine-proline amide bond are 0.001 sec^{-1} ($k_{trans \rightarrow cis}$) and 0.005 sec^{-1} ($k_{cis \rightarrow trans}$) in water at 25 °C.⁹⁰ As will be seen, this relatively slow amide isomerization has important implications for proline in protein folding.

Stein and coworkers investigated the energetic contributions to the isomerization energy barrier of prolyl amide isomerization.⁹¹ It was found that when the kinetic parameters of proline isomerization in several short model peptides is averaged, the energetic contributions to amide C-N bond rotation *N*-terminal to proline are entirely enthalpic in nature ($\Delta H^\ddagger = 19.0 \pm 1.0 \text{ kcal}\cdot\text{mol}^{-1}$), while the entropic contribution is negligible ($\Delta S^\ddagger = 0.0 \pm 1.0 \text{ cal}\cdot\text{K}\cdot\text{mol}^{-1}$) at 25 °C. The enthalpic contribution to the activation energy is believed to arise solely from the loss of the resonance-stabilized amide bond; whereas the very small entropic contribution to proline isomerization indicates that the transition state has no solvent participation and/or little solvent reorganization, and therefore isomerization can be considered to be a unimolecular process.

Several studies have given insight into the transition state structure with regards to prolyl amide isomerization, in which the effects of pH and solvent were considered.⁹² It was found that pH has no effect on prolyl isomerization between pH 5 and 9. This indicates that at physiological pH, prolyl isomerization proceeds without general acid-base catalysis or involvement of the solvent. However, at extremes of pH, there is an increase in the rate of amide isomerization. It was also found that non-polar solvents increase the rate of prolyl amide isomerization, which indicates that the transition state for amide C-N bond rotation is non-polar relative to the ground state.⁹¹ It is now believed that the transition state is characterized by partial rotation about the amide bond, in which the polar amide bond resonance structures no longer apply, making the transition state less polar (Figure 1.5.3.1).

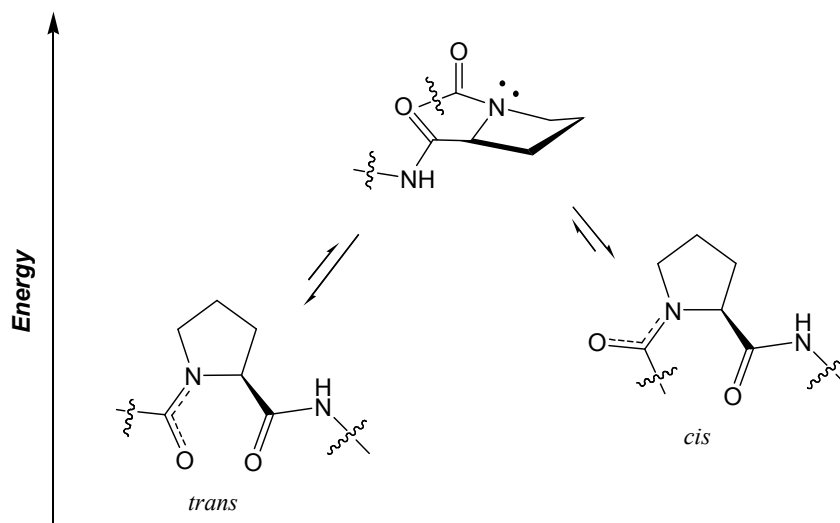


Figure 1.5.3.1: Relative energies of *trans* and *cis* amide isomers, and the transition state characterized by partial rotation of the amide bond and increased pyramidalization of the proline nitrogen atom.

Insight into important biological processes can be obtained by considering the rate of amide isomerization *N*-terminal to proline. An example can be found in the study of Oxytocin and Arginine Vasopressin peptide hormones, carried out by Larive and Rabenstein.⁹³ Both Oxytocin and Vasopressin are nonapeptides in which the first and 6th amino acids are linked through a disulfide bridge; proline is the first of three amino acids in a tripeptide tail attached to the macrocyclic hexapeptide ring (Figure 1.5.3.2).

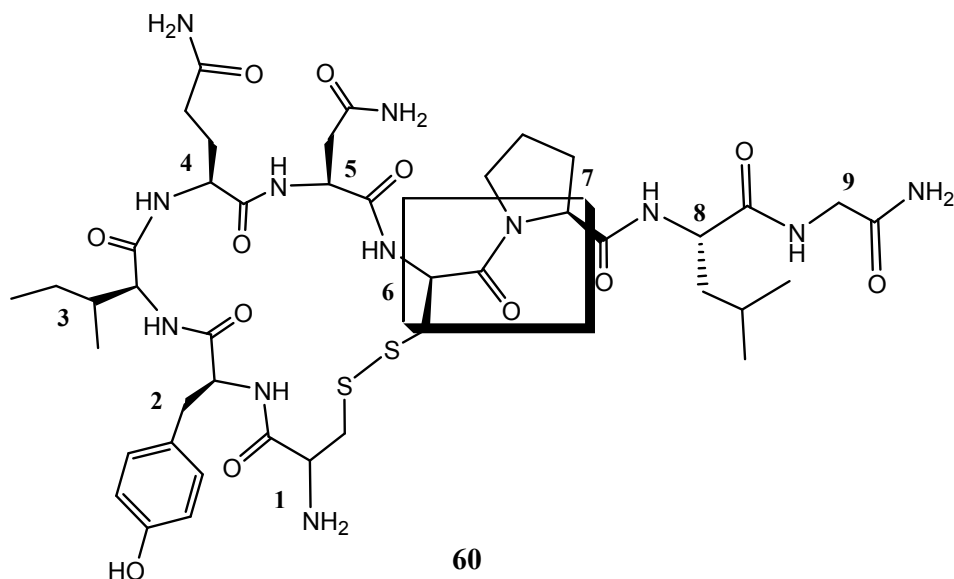


Figure 1.5.3.2: Structure of Oxytocin with the *trans* Cys⁶-Pro⁷ amide bond highlighted.

Oxytocin causes uterine contraction and milk ejection, while Vasopressin causes moderate constriction of blood vessels and acts on the kidneys to conserve water.⁹⁴ The proline residue at position 7 is known to be required for optimal biological activity. Early work had established that the amide bond *N*-terminal to this proline residue is strongly biased towards a *trans* conformation ($K_{trans/cis}$ of 12 for Oxytocin, 15 for Arginine Vasopressin).⁹⁵⁻⁹⁷ The authors characterized the thermodynamics and kinetics of proline isomerization for both Oxytocin and Arginine Vasopressin in an attempt to understand the underlying factors that favour the *trans* amide conformation, which they postulated to be an intramolecular interaction between the hexapeptide ring and the tripeptide tail.

The isomerization of the cysteine-proline amide bond is on a time-scale that is amenable to analysis by NMR magnetization transfer experiments, thus these experiments were used to calculate the rates of proline isomerization (Figure 1.5.3.3).

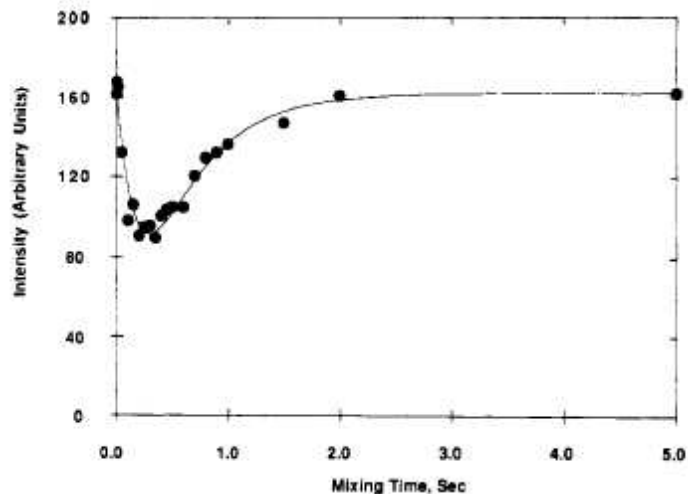


Figure 1.5.3.3: Integrated resonance intensity of the Cys⁶-NH proton of the *cis* isomer as a function of mixing time; the smooth line is the calculated least-squares fit of the data. [Reproduced with permission, *J. Am. Chem. Soc.* **1993**, *115*, 2833. Copyright 1993 American Chemical Society.]

The calculated isomerization rate from the *trans* to the *cis* isomer ($k_{trans \rightarrow cis}$) for Oxytocin (0.0035 sec^{-1}) and Arginine Vasopressin (0.0046 sec^{-1}) were quite similar to X_{aa}-Pro isomerization rates in other peptides (ranging from 0.00064 to 0.0048 sec^{-1}). However, the rate of isomerization from the *cis* to the *trans* isomer ($k_{cis \rightarrow trans}$) was much faster for Oxytocin (0.042 sec^{-1}) and Arginine Vasopressin (0.067 sec^{-1}) compared to X_{aa}-Pro isomerization rates in other peptides (0.0013 to 0.0079 sec^{-1}).⁹⁸ This indicates that instead of a stabilizing effect on the *trans* conformer, there is a destabilizing effect on the *cis* amide conformer that lowers the activation energy barrier, and causes faster isomerization back to the *trans* isomer. The authors did not attempt to identify the destabilizing effects on the *cis* amide conformer in Oxytocin or Arginine Vasopressin. Regardless, the significant finding was that the rate of prolyl amide isomerization was shown to impact the biologically active conformation of two important peptide hormones.

Furthermore, their work supports the theory of a non-polar transition state for the amide bond *N*-terminal to proline. In methanol, the rates of isomerization for Oxytocin and Arginine Vasopressin in both directions ($k_{trans \rightarrow cis}$ and $k_{cis \rightarrow trans}$) were much faster than in water. A less polar solvent such as methanol would be expected to lower the enthalpic penalty for creating a non-polar transition state, thus lowering the free energy of the transition state.

Because of the relatively slow isomerization rate of a proline *N*-terminal amide bond (second timescale), proline isomerization becomes the rate-determining step of protein folding.⁹⁹ To aid in the protein folding process, Nature has developed a broad class of enzymes that are responsible for isomerizing the *N*-terminal amide bond of proline in proteins that are found ubiquitously, which perhaps is the best indication of the importance of proline in biology.^{99,100} The peptidylprolyl *cis-trans* isomerase (PPIase) enzymes catalyze the *cis* to *trans* isomerization of proline. There are many forms of PPIases, with distinct catalytic mechanisms and substrates.¹⁰¹ A common feature of PPIases is a hydrophobic pocket specific for binding the pyrrolidine side chain.¹⁰² It has been found that the enzymes can accommodate some variation in the pyrrolidine side chain.¹⁰³ There is much work being done to understand the exact enzymatic mechanism for accelerating/catalyzing amide isomerization.¹⁰³ It seems that PPIases can act through protonation of the prolyl nitrogen atom, nucleophilic attack of the prolyl *C*-terminal carbonyl group, and electrostatic stabilization of the transition state structure.¹⁰² The unique properties of proline are exploited by Nature to cause changes in protein structure, which has an impact on protein function.

1.5.4 Conformational Changes Induced by Proline Affect Protein Function

Proline is often found in β -turn regions of proteins, which represent the most common type of reverse turn and an important structural motif.¹⁰⁴ The β -turn consists of four amino acids, designated as i , $i + 1$, $i + 2$ and $i + 3$, from the N - to C -terminus, which induce a complete 180° turn in the direction of the polypeptide (Figure 1.5.4.1).

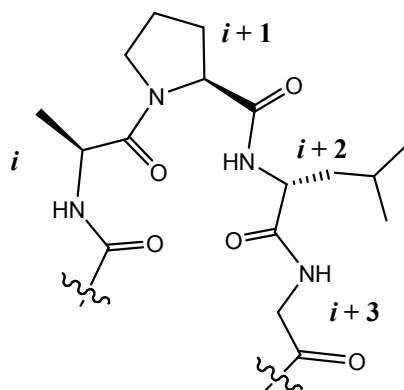


Figure 1.5.4.1: Proline at the $i + 1$ position of a β -turn.

There are more than 10 types of β -turns, which differ mostly on the ϕ - and ψ -torsional angles of the $i + 1$ and $i + 2$ residues.¹⁰⁴ Typically, the turns exhibit an intramolecular hydrogen bond between the i and $i + 1$ residue. Proline is by far the most common amino acid found at the $i + 1$ position in type I and II β -turns. Also, proline is often found in the $i + 2$ position in Type VI β -turns, which are unique in requiring a *cis* amide conformation N -terminal to the $i + 2$ residue. Overall, these turns are induced by proline because it has reduced flexibility around its ϕ - and ψ -angles, which are predisposed to the ideal angles for β -turn formation at that position.

In many cases, it is reverse turn regions of a peptide or protein that are directly involved in protein-protein interactions.¹⁰⁵ Proline is often found either in a flanking

position of these binding sites, or within the turn region itself.¹⁰⁶ For example, small peptide hormones, such as Angiotensin II and Bradykinin, have been shown to adopt β -turn conformations when bound to their receptors. Proline is found in these binding regions. Since the receptors for these peptides are integral membrane proteins, this structural requirement has only been inferred through indirect binding studies.

In 1992, Garcia and coworkers attempted to uncover the bioactive conformation of the vasoconstrictor Angiotensin II (Asp-Arg-Val-Try-Ile-His-Pro-Phe-His-Leu), which is extremely flexible in solution.¹⁰⁷ Using the crystal structure of the peptide bound to a monoclonal antibody, acting as a surrogate receptor, they confirmed that Angiotensin II binds to the protein receptor through a turn region involving the Ile, His and Pro residues. The center of the turn was lodged in the deepest region of the binding site, which reflects its importance in the binding process. Similarly, extensive molecular modelling and NMR studies by Kyle and coworkers of the vasodilator Bradykinin (Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg) found that the four terminal residues Ser-Pro-Phe-Arg in fact form a β -turn.¹⁰⁸ This turn involves a hydrogen bond from the arginine backbone amide proton to the serine backbone carbonyl group. Later, Bradykinin analogues developed by Thuriereau and coworkers indicated that this β -turn region is critical for high affinity binding to the Bradykinin B2 receptor.¹⁰⁹ Since it is imperative to know the conformation of the bound peptide for developing potent analogues, these studies were of the utmost importance in the development of Angiotensin II and Bradykinin analogues for therapeutic use. In these instances, proline was found to induce β -turns in the binding regions of these important peptide hormones.

Apart from a role in binding interactions through reverse turns, proline itself can induce different biological responses simply through isomerization of its *N*-terminal amide bond. Because the isomerization of proline is slow, both *cis* and *trans* X_{aa}-Pro amide isomers effectively exist as different structural isomers in peptides and proteins. In certain cases, only one isomer is the active conformation of a peptide or protein. For instance, the proteins: staphylococcal nuclease,¹¹⁰ insulin,¹¹¹ and salmon calcitonin¹¹² are known to contain a *cis* X_{aa}-Pro amide bond in their active conformation. In other cases, both amide isomers can demonstrate biological activity; sometimes with different effects. For instance, it has been suggested that both the *cis* and *trans* Xaa-Pro amide isomers of the protein interleukin-2 tyrosine kinase have distinct biological activities.¹¹³

Perhaps most excitingly, it recently has been found that isomerization of a single proline residue causes opening and closing of an ion gate. The 5-hydroxytryptamine type 3 (5-HT₃) receptors are known to associate and form ion pores upon neurotransmitter binding.¹¹⁴ These proteins contain a neurotransmitter-binding region separated from two antiparallel transmembrane helices (Figure 1.5.4.2). A single proline residue joins the antiparallel helices. Lummis and coworkers found that this proline residue is crucial for pore formation by altering the orientation of the two helices through prolyl *cis-trans* amide isomerization.¹¹⁵ When the proline residue adopts a *trans*-amide conformation, the helices are aligned parallel to each other, and no ion flow through the membrane is observed. If the same proline residue is in a *cis*-amide conformation, the helices are pinched together to form a pore, and ion flow is observed. It is believed that the 5-HT₃ neurotransmitter binding region may lock the crucial proline residue in a *trans* amide

conformation; neurotransmitter binding causes a conformation change, thus removing the constraint and permitting prolyl amide isomerization and thus, ion flow, to take place.

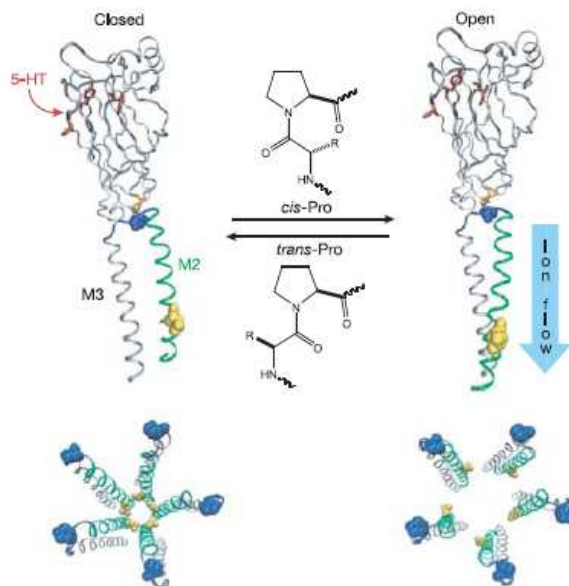


Figure 1.5.4.2: Side and top-down views of the proposed ‘closed’ and ‘open’ forms of 5-HT₃ receptor subunits; 5-HT binding causes *cis-trans* isomerization of a key proline residue (coloured dark blue), and movement of the M3 and M2 helices, thereby allowing ion flow to take place. [Reproduced with permission from Macmillan Publishers Ltd: *Nature* **2005**, 438, 248. Copyright 2005.]

With such diverse effects on peptides and proteins, perhaps it is no surprise that proline has been called the most important amino acid when it comes to determining protein function and structure,^{116,117} and has led to numerous synthetic analogues of proline used for understanding biological processes.

1.6 Proline Analogues Are Used to Control Prolyl Amide Isomerization

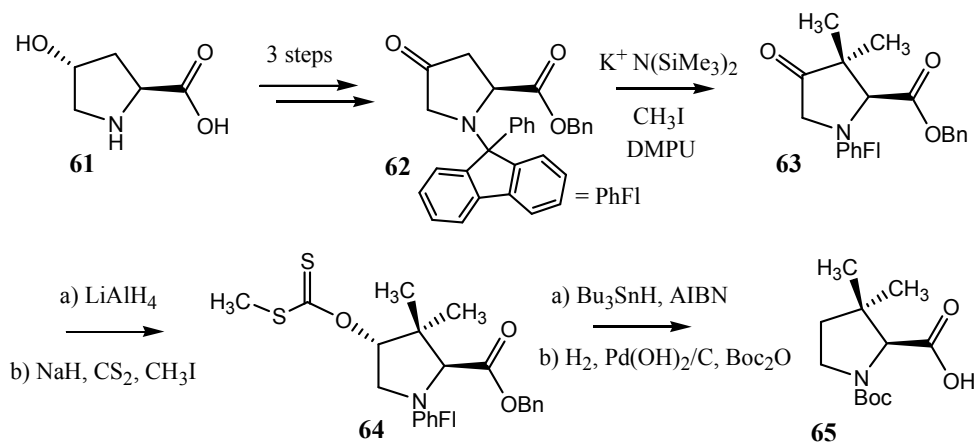
Since proline demonstrates fundamental influences on protein folding, structure and function, there is a significant interest in creating synthetic proline analogues to expand the understanding of proline in biology, and to use this knowledge to create therapeutics for treating disease. Specifically, incorporation of non-natural structural motifs or unnatural amino acids into peptides can be used for therapeutic development to prevent enzymatic breakdown of peptide bonds and thus improve bioavailability.¹¹⁸ Also, incorporation of non-natural amino acids is sometimes necessary to probe structure-function relationships of proteins that are not accessible by the natural 20 amino acids. In order to study protein-protein interactions mediated through regions containing proline, the focus generally centers on controlling *N*-terminal amide isomerization.

There are many approaches to controlling the *N*-terminal amide isomerization of proline. Generally, a peptide can be locked in either the *cis* or *trans* conformation through covalent linkages¹¹⁹⁻¹²³ or simply biased towards one conformation through modification of the peptide backbone or proline side chain.^{124,125} Covalently linked analogues of prolyl *cis* and *trans* amide isomers usually involve larger structural changes to the peptide, and this approach has been criticized for creating peptide analogues with reduced receptor binding affinity.¹²⁶ Furthermore, replacement of the X_{aa} -Pro motif by dipeptide isosteres has been criticized for limiting the potential development of combinatorial libraries of X_{aa} -Pro compounds.¹²⁶ In other cases, instilling a bias in the *cis/trans* amide equilibrium is desirable because it still allows *cis-trans* isomerization to occur and can be used to study more nuanced protein folding and binding situations. This can be accomplished through modification of the proline side chain or the peptide backbone. However, only proline analogues in which the side chain has been modified are amenable for automated

solid phase peptide synthesis and protein engineering.¹²⁶ Here, the focus will be on examples of synthetic analogues of proline in which the side-chain of proline has been substituted at the β - and δ -positions with alkyl substituents, and at the γ -position with electron withdrawing groups in order to influence the *N*-terminal amide *cis/trans* equilibrium, and for use in studying biological phenomena.

1.6.1 Proline β -Alkyl Substitution: 3,3-Dimethylproline

Lubell and coworkers studied the effects of alkylation at the β -position of proline on its *N*-terminal amide equilibrium and rate of amide isomerization.¹²⁷ The 3,3-dimethylproline analogues were synthesized from the protected 4-oxoproline **62**,¹²⁸ which is accessible from 4*R*-hydroxy-L-proline **61** in three steps (Scheme 1.6.1.1).¹²⁹



Scheme 1.6.1.1: Synthesis of protected a 3,3-dimethylproline building block.^{128,129}

The 4-oxoproline **62** can be regioselectively dialkylated at the 3-position upon treatment with potassium bis(trimethylsilyl)amide and iodomethane.¹²⁸ Removal of the 4-position ketone was carried out by reduction and deoxygenation steps. Finally, manipulation of the protecting groups gave the desired 3,3-dimethylproline building

block **65** which was incorporated into *N*-acetyl proline *N'*-methanamide model peptides for analysis of the effects of proline β -alkylation on the thermodynamics and kinetics of *N*-terminal amide isomerization.

Lubell and coworkers found that in D₂O, prolyl β -alkylation has no impact on the *cis/trans* amide equilibrium (Table 1.6.1.1).¹²⁷ The model peptide containing L-proline exhibited a $K_{trans/cis}$ of 2.6, while the 3,3-dimethylproline model peptide had $K_{trans/cis}$ of 2.3 as determined by integration of the *cis* and *trans* isomer ¹H NMR signals.

Table 1.6.1.1: Thermodynamic and kinetic parameters of 3,3-dimethylproline.

Compound ^a	Pucker	$K_{trans/cis}$ ^b	$k_{cis-to-trans}$ (sec ⁻¹) ^c
L-proline	C ^{γ} -endo	2.6 \pm 0.3	2.0 \pm 0.1
3,3-dimethylproline	C ^{γ} -endo	2.3 \pm 0.2	0.32 \pm 0.07

^aIn the model peptide Ac-X_{aa}-NHMe; ^bCarried out in D₂O at 25 °C; ^cCarried out in D₂O at 60 °C.

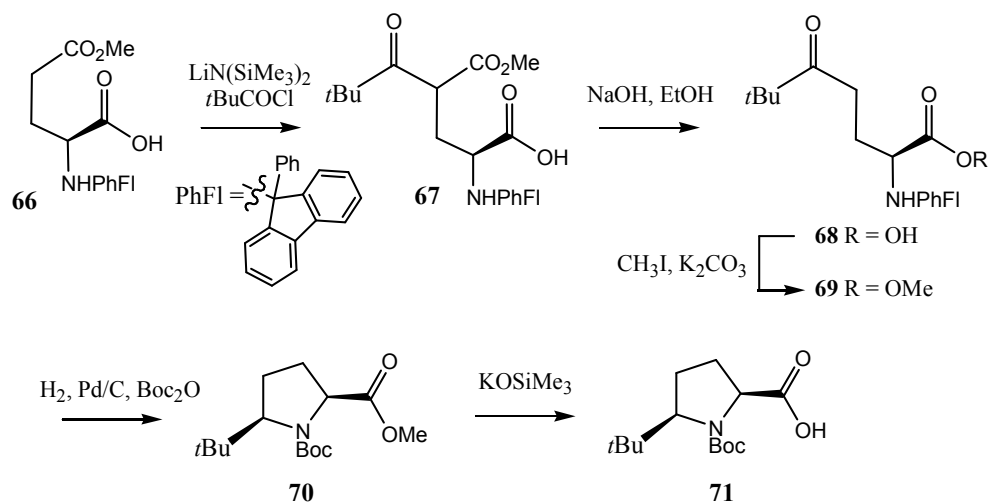
X-ray crystallographic analysis revealed that the proline analogue adopts a C ^{γ} -endo conformation. Therefore, the ring pucker of proline is not changed by β -alkylation. However, according to ¹³C NMR magnetization transfer experiments, 3,3-dimethylproline significantly lowers the rate of amide isomerization compared to L-proline. The *cis-to-trans* amide isomerization rate constant (k_{ct}) for the L-proline model peptide was 2.10 \pm 0.09 sec⁻¹ in D₂O at 60 °C, while the rate constant (k_{ct}) for the 3,3-dimethylproline model peptide was found to be 0.32 \pm 0.07 sec⁻¹ under the same conditions. The lower rate of isomerization for 3,3-dimethylproline was attributed to an unfavourable ψ angle of 150° caused by steric repulsion from the *C*-terminal amide group

with the β -position alkyl groups. This would place the *C*-terminal carbonyl group in a position to disfavour sp^3 pyramidalization of the prolyl nitrogen atom, which is formed in the transition state of isomerization. Therefore, while β -alkylation has no impact on the prolyl *cis/trans* amide equilibrium, it does significantly lower the rate of amide isomerization.

1.6.2 Proline δ -Alkylation: 5-*tert*-butylproline

Lubell and coworkers looked to modify prolyl *N*-terminal amide isomerization through installation of a bulky alkyl substituent, a *tert*-butyl group, at the δ -position of the prolyl side chain.¹²⁶

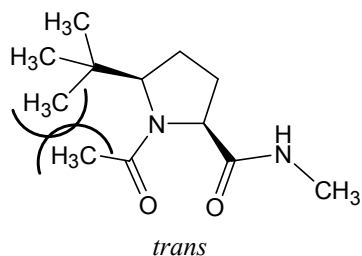
The synthesis of 5*R*- and 5*S*-*tert*-butylproline (*t*BuPro) was carried out in an enantioselective process starting from glutamic acid via an acylation/diastereoselective reductive amination sequence.¹³⁰ Deprotonation of *N*-protected L-glutamic acid **66** using lithium bis(trimethylsilyl)amide was followed by *C*-acylation with pivaloyl chloride to give **67** (Scheme 1.6.2.1).



Scheme 1.6.2.1: Synthesis of (2*S*,5*R*)-*tert*-butylproline building block **71**.

Hydrolysis and decarboxylation of the γ -methyl ester using sodium hydroxide at high temperature followed by esterification of the acid using iodomethane gave the protected ester **69**. Catalytic hydrogenation of the *N*-protecting group allowed both cyclization and subsequent reductive amination to occur. Carrying out the reaction in the presence of di-*tert*-butyl dicarbonate allowed simultaneous *N*-Boc protection; giving the target (2*S*,5*R*)-*N*-Boc-5-*tert*-butylproline methyl ester **70** with 90% diastereoselectivity and 50% overall yield. Hydrolysis provides the *N*-Boc protected building block **71**. Access to 5*S*-*tert*-butylproline could be obtained by stereoselective reduction of the imine instead of using catalytic hydrogenation.

Lubell and coworkers used model peptides of the form *N*-acetyl-proline-*N'*-methylamide to show that 5-*tert*-butylproline increases the prolyl *N*-terminal *cis* amide content ($K_{trans/cis}$ of 1.0 and 0.52 for 5*R*-*t*BuPro and 5*S*-*t*BuPro, respectively) in D₂O at 25 °C compared to the model peptide containing L-proline ($K_{trans/cis}$ of 2.7), according to integration of ¹H NMR signals (Figure 1.6.2.1).¹²⁶



Ac-X _{aa} -NHMe	$K_{trans/cis}$	ΔG^\ddagger (kcal mol ⁻¹)
L-proline	2.7	20.4
2 <i>S</i> ,5 <i>R</i> - <i>tert</i> -butylproline	1.0	16.5
2 <i>S</i> ,5 <i>S</i> - <i>tert</i> -butylproline	0.52	20.2

Figure 1.6.2.1: Unfavourable steric clash in *trans* amide form of 2*S*,5*R*-*tert*-butylproline; corresponding amide isomer equilibrium at 25 °C in D₂O ($\pm 10\%$) and transition state free energy difference determined in D₂O.

The authors attributed this effect to the 5-position *tert*-butyl group destabilizing the *trans* amide isomer through steric repulsion. Of note is that the effect of the *tert*-butyl group on the pucker of the pyrrolidine ring was not investigated.

The energy barrier to amide isomerization (ΔG^\ddagger) was also determined using magnetization transfer experiments and amide proton chemical shift coalescence experiments. It was found that the ΔG^\ddagger for 5*R*-*t*BuPro was 3.9 ± 0.2 kcal·mol⁻¹ lower in energy compared to L-proline, while there was no energy difference found between 5*S*-*t*BuPro and L-proline in the model peptides (20.2 and 20.4 kcal·mol⁻¹, respectively). This result was justified by the fact that the ψ -angle for 5*R*-*t*BuPro ($\sim 0^\circ$) was found to be different from 5*S*-*t*BuPro and L-proline ($\sim 125^\circ$) according to molecular mechanics calculations. A ψ -angle of 0° could allow the *C*-terminal NH group to interact with the prolyl nitrogen lone pair or the *C*-terminal carbonyl group, thereby affecting the pyramidalized transition state; this could account for the difference in ΔG^\ddagger .

In 2000, Lubell and coworkers extended the scope of their work by replacing the L-proline residue in Oxytocin with 5-*tert*-butylproline.¹³¹ Oxytocin, as described earlier (Section 1.5.3), is a peptide hormone in which the first and 6th amino acids are linked through a disulfide bridge and proline is the first of three amino acids in a tripeptide tail attached to the macrocyclic ring. Oxytocin is used clinically during childbirth to induce rhythmic uterine contractions as well as to ease breast-feeding.¹³² It has been proposed that antagonists of Oxytocin function could be used to prevent pre-term labour.^{133,134} Because of significant interest in developing synthetic analogues of Oxytocin, efforts have been made to further understand the biologically active conformation of Oxytocin, especially the conformation of the proline-linked tripeptide appendage.¹³⁵⁻¹³⁷

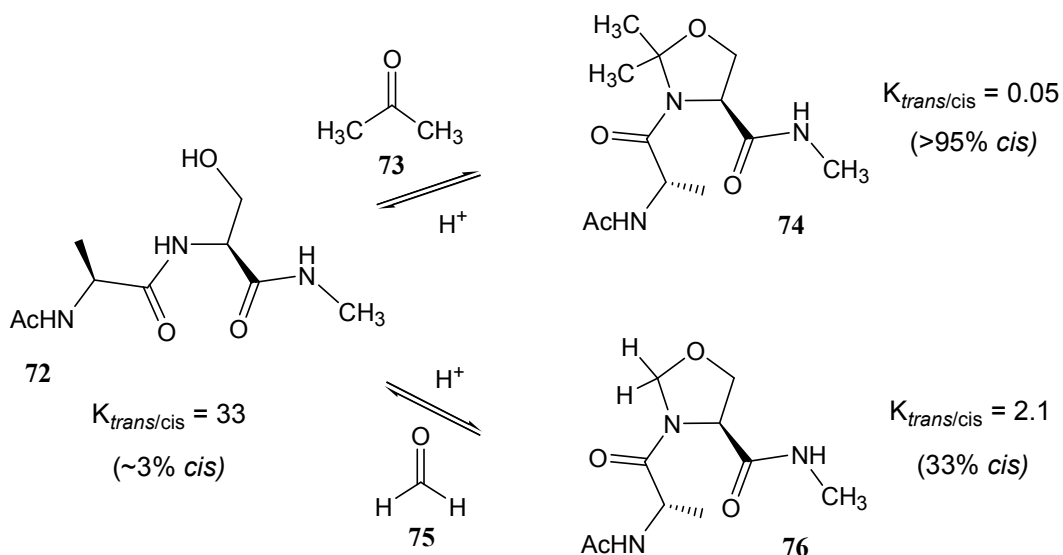
Lubell and coworkers found that replacing the L-proline residue in Oxytocin with 5-*tert*-butylproline caused an increase in the cysteine-proline *cis* amide conformation (from 10% to 35% in 9:1 H₂O/D₂O at 30 °C, which corresponds to a decrease in $K_{trans/cis}$ from 9 to 1.9), based on NMR assignment. This is a smaller stabilization of the *cis* amide conformation based on previous studies with model peptides, which was attributed to inherent properties of Oxytocin that favour the *trans* amide isomer. By studying the biological activity of these Oxytocin analogues in rat models, Lubell found that *t*BuPro lowered both agonist activity and binding affinity relative the natural peptide by nearly 400 times! This could indicate that the *trans* amide conformation is required for optimal agonist activity, or that the *tert*-butyl group interferes with receptor binding, or both.

In a complementary study, when incorporated into inhibitors of Oxytocin function, *t*BuPro caused a slight increase in the antagonist activity with a 3-fold increase in binding affinity.¹³⁸ In response to these results, the authors proposed that a *cis* amide conformation results in antagonist activity, while a *trans* amide conformation results in agonist activity of Oxytocin. In this way, *t*BuPro has been used to understand the activity of a biologically important peptide.

1.6.3 Proline δ -alkylation: The Pseudoprolines

Originally developed as protected serine, threonine or cysteine building blocks, the pseudoprolines (denoted as ψ Pro) were developed to reduce α -epimerization, increase solubility, prevent self aggregation and improve coupling yields in solid phase peptide synthesis.¹⁴⁰⁻¹⁴² The pseudoprolines H,H- ψ Pro **76** and Me,Me- ψ Pro **74** are accessible

from cyclocondensation of Ser, Thr, or Cys with formaldehyde **75** and acetone **73**, respectively (Scheme 1.6.3.1).



Scheme 1.6.3.1: Serine in a Ac-Ala- X_{aa} -NHMe model peptide can be converted to Ser($\psi^{H,H}$ Pro) **76** and Ser($\psi^{Me,Me}$ Pro) **74** using formaldehyde **75** and acetone **73**, respectively, under acidic conditions; this leads to large changes in $K_{trans/cis}$.

Mutter and coworkers found that the amide bond *N*-terminal to a pseudoproline is highly influenced by the δ,δ -substituents on the pseudoproline side chain; the *cis* amide content of model peptides of the form Ac-Ala- ψ Pro-NHMe was varied from 33% for H,H- ψ Pro ($K_{trans/cis} = 2.1$), which is similar to L-proline (18% or $K_{trans/cis} = 4.6$), to >95% ($K_{trans/cis} = 0.05$) for Me,Me- ψ Pro (in DMSO at 27 °C) according to NMR measurements.¹⁴³ The authors proposed that in a similar fashion to 5-*tert*-butylproline, steric interactions play a major role in the ψ Pro's influence on *N*-terminal amide isomerization.

According to ^1H NMR coupling constants, the pseudoprolines adopt a C^β -exo pucker, regardless of the substitution, which is different from the C^γ -endo conformation

normally observed for L-proline.¹⁴³ A key advantage of the pseudoproline is that deprotection under acidic conditions restores the *N*-terminal *trans* amide conformation.¹⁴⁴ Thus, the pseudoproline represents an accessible way to greatly modify *cis* amide content and peptide conformation.

Mutter and coworkers incorporated the pseudoproline Cys($\psi^{\text{H,H}}$ Pro) and Cys($\psi^{\text{Me,Me}}$ Pro) into position 7 of the Oxytocin nonapeptide using SPPS to understand its bioactive conformation (Section 1.5.3).¹⁴⁵ For the Cys($\psi^{\text{Me,Me}}$ Pro) analogue **77**, NMR experiments were used to identify the major isomer (92%) as having a cysteine-Cys($\psi^{\text{Me,Me}}$ Pro) *cis* amide bond ($K_{\text{trans/cis}}$ of 0.09). In biological tests, they found that when a *cis* amide conformation is fixed at this position, the Oxytocin analogue showed no agonist activity *in vitro* compared to the native peptide (Figure 1.6.3.1).

Interestingly, the ψ Pro-Oxytocin analogue did exhibit equal binding affinity to the Oxytocin receptor compared to the native peptide. This indicates that as opposed to *t*BuPro, the methyl groups do not interfere with peptide binding, nor does a prolyl *cis* amide conformation preclude receptor binding. Mutter proposed that the initial binding of the Oxytocin peptide might have a *cis* amide conformation, but then isomerize to the *trans* conformation to produce the agonist effect.

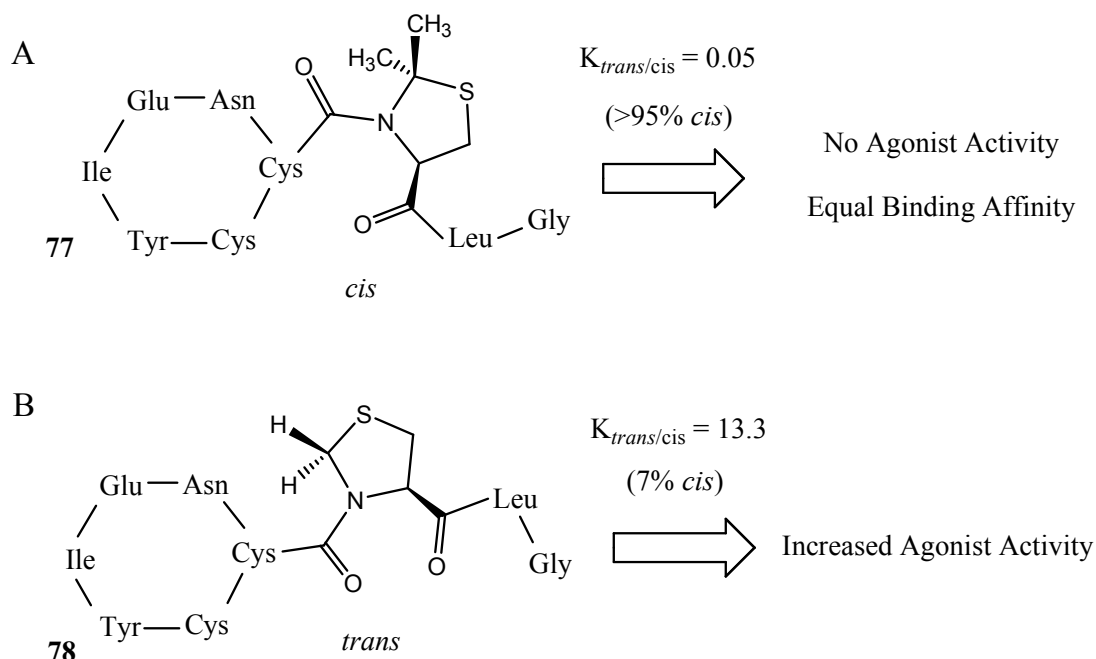


Figure 1.6.3.1: A) Incorporation of Cys($\psi^{\text{Me,Me}}\text{Pro}$) into Oxytocin stabilizes the *cis* amide conformation, leading to no agonist activity *in vitro*; B) Incorporation of Cys($\psi^{\text{H,H}}\text{Pro}$) causes little change in $K_{trans/cis}$, but leads to increased agonist activity *in vitro*.

The H,H- ψPro analogue **78**, which had a lower *cis* amide content (7%) ($K_{trans/cis}$ of 13.3), exhibited a higher agonist activity compared to the native peptide (with *cis* amide content of 10% in D_2O) ($K_{trans/cis}$ of 9.0) (Figure 1.6.3.1).¹⁴⁵ This increase in activity was attributed to a lower barrier to *cis-trans* isomerization for H,H- ψPro compared to L-proline. This evidence supports the hypothesis that the binding and agonist activity of Oxytocin requires *cis-to-trans* amide isomerization of proline. A *cis-to-trans* conformational switch upon receptor binding to agonist activity has also been suggested for another peptide hormone: Thyroliberin.¹⁴⁶ Thus, these examples likely indicate a broader role of proline isomerization in which both isomers, and the isomerization process play a role in biology. In summary, the pseudoproline were useful in

demonstrating the importance of the conformation of the prolyl *N*-terminal amide bond on the biological activity of Oxytocin. This example also highlights the fact that this subtle process would not have been understood using covalently locked proline analogues; demonstrating the advantage of proline analogues in which only the side chain has been modified.

Criticisms of pseudoproline (oxazolidines and thiazolidines) are that they are acid labile, which limits their use in peptide synthesis.¹⁴⁷ Also, the presence of a heteroatom alters the proline side-chain conformation and hydrophobicity compared to proline, which limits their use when retaining the conformation of the proline side chain is required. Also, as with *t*BuPro, the steric bulk on the prolyl side chain could affect binding interactions involving the proline side chain.

The proline analogues developed so far have limited potential to be further derivatized once incorporated into a peptide for modulation of effect or development of combinatorial libraries. Furthermore, few proline analogues maintain the native pucker of L-proline, which, as will be seen, is critical for the function of certain proteins. A derivative of proline that comes close to satisfying these requirements is created via the naturally-occurring γ -position modification of the prolyl side chain.

1.6.4 Proline γ -Modification: 4-Hydroxyproline

Hydroxylation at the 4-position of proline can be considered to be a naturally occurring modification of proline; it has effects on both the pucker and *N*-terminal amide equilibrium of proline.

Proline 4-hydroxylation in vertebrates occurs post-translationally through the action of prolyl-4-hydroxylase (P4H), of which there are two closely related forms (I and II).¹⁴⁸ The enzyme, located in the lumen of the endoplasmic reticulum, acts on single-stranded peptides (formation of a triple helix completely prevents hydroxylation by P4H). Poly-L-proline peptides are highly competitive inhibitors of P4H, which suggests the polyproline conformation is required for enzyme binding. The minimum sequence required for hydroxylation of Pro has been found to be the tripeptide sequence (X_{aa} -Pro-Gly).¹⁴⁹ It is known that the enzyme requires Fe^{2+} , O_2 , 2-oxoglutarate and ascorbate as cosubstrates.¹⁵⁰ The current model for catalysis is that the Fe^{2+} ion located in a hydrophobic pocket, is coordinated to the side chains of three amino acids (His412, His483 and Asp414) (Figure 1.6.4.1).

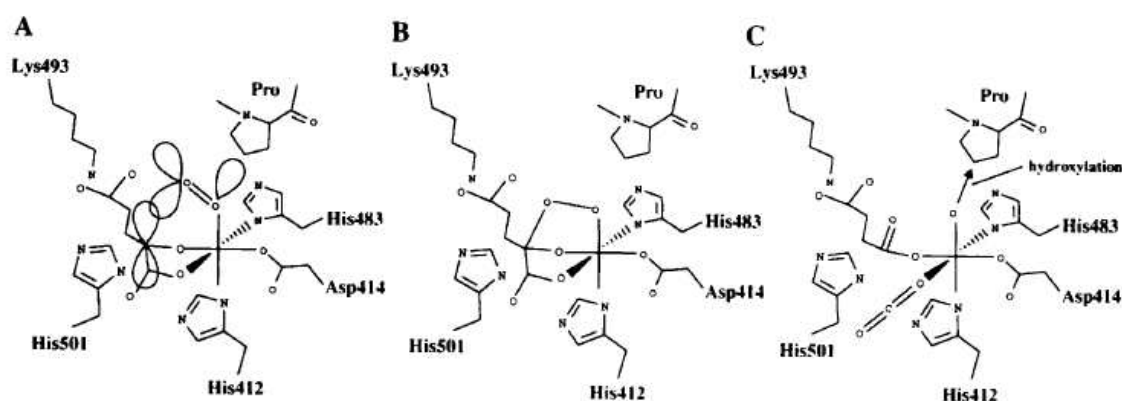


Figure 1.6.4.1: Proposed mechanism of proline 4-hydroxylation: A) Fe^{2+} is coordinated by the side chains of His412, Asp414, and His 483; B) Fe^{2+} forms a complex with O_2 and 2-oxoglutarate; C) Formation of CO_2 coincides with regio- and stereospecific hydroxylation of proline. [Reproduced from *Matrix Biology* **1998**, 16, Kari I. Kivrikko and Johanna Myllyharju, pages 357-368. Copyright 1998 with permission from Elsevier]

During catalysis, 2-oxoglutarate is decarboxylated to form succinate; followed by one atom of O₂ being incorporated into succinate, and the other onto the proline side-chain.^{151,152} The ferryl ion (Fe³⁺O⁻) produced is reduced by ascorbate in order to regenerate the catalytic cycle. The hydroxylation of L-proline in the (X_{aa}-Pro-Gly) sequence occurs regio- and stereospecifically, presumably because of a selective binding pocket;¹⁵³ hence, 4*S*-hydroxyproline is not produced by P4H in vertebrates.

Hydroxylation of L-proline modifies two important physical parameters of proline. First, hydroxylation tends to stabilize the proline side chain pucker, which depends on the stereochemistry of the 4-position oxygen atom (Figure 1.6.4.2).¹⁵⁴ While 4*R*-hydroxyproline (Hyp) adopts a C^γ-exo pucker, 4*S*-hydroxyproline (hyp) adopts a C^γ-endo type conformation.^{73-79,154-157} Second, hydroxylation affects the *N*-terminal amide *cis/trans* equilibrium, which also depends on the stereochemistry of the attached oxygen atom. It has been observed that Hyp stabilizes the *N*-terminal amide *trans* conformation ($K_{trans/cis}$ of 6.1) relative to proline ($K_{trans/cis}$ of 4.6), and hyp stabilizes the *cis* amide conformation ($K_{trans/cis}$ of 2.4).¹⁵⁴

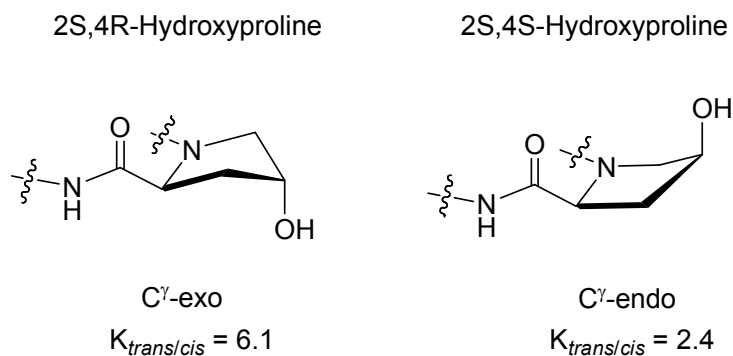


Figure 1.6.4.2: Differences in the pucker and amide isomer ratio ($K_{trans/cis}$) between 4*R*- and 4*S*-hydroxyproline.

Raines and coworkers found that hydroxylation does not seem to influence the rate of *N*-terminal amide isomerization compared to L-proline in model peptides (in D₂O at 37 °C) according to NMR magnetization inversion transfer experiments.¹⁵⁵ Interestingly, evidence indicates that the change in proline pucker and prolyl amide equilibrium that occurs upon 4-hydroxylation are directly related; this fact was discovered as a part of understanding an important protein rich in Hyp: collagen.

1.7 Collagen

1.7.1 Understanding the Structure of Collagen

Collagen is the most abundant protein found in vertebrate species, a subphylum that includes animals, birds, and humans.¹⁵⁸⁻¹⁶⁰ In fact, collagen is found in nearly every human tissue and is the major protein component of skin, bones, ligaments, tendon, and cartilage.¹⁶¹ Collagen exists in the extracellular matrix as extended fibrillar structures, in which tropocollagen molecules have associated into characteristic fibrils, which then bundle into larger collagen fibers (Figure 1.7.1.1).^{162,163}

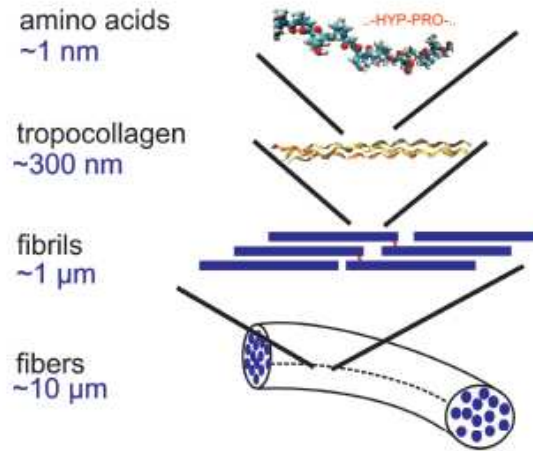


Figure 1.7.1.1: Overview of collagen structure: tropocollagen molecules, composed of a characteristic amino acid sequence, combine to form fibrils, which form larger collagen fibers. [Reproduced with permission, *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 12285. Copyright 2006 the National Academy of Sciences, U.S.A.]

On a molecular level, tropocollagen molecules are composed of three polypeptide chains, each in a left-handed polyproline II type conformation, twisted together into a right-handed triple helix.¹⁶⁴⁻¹⁶⁹ The amino acid sequence of each strand is a repeating tripeptide motif of $(\text{Gly-X}_{\text{aa}}\text{-Y}_{\text{aa}})_n$, where glycine is found at every third position in the amino acid sequence. Glycine is required at every third position to allow tight packing of the three polypeptide strands; an amino acid with a larger side chain would preclude triple helix formation. The amino acid in the X_{aa} position is often L-proline, while the Y_{aa} position amino acid is often (4R)-4-hydroxy-L-proline. There are normally approximately 300 tripeptide repeats per collagen molecule, which gives a length of ~ 300 nm.¹⁷⁰ Of note is that all of the amide bonds in the collagen molecule exist in a *trans* amide conformation. Because collagen is such a prevalent structural protein, and is implicated as a target of several disease states including osteogenesis imperfecta,¹⁷¹ rheumatoid

arthritis,¹⁷² and diabetes mellitus,¹⁷³⁻¹⁷⁴ much effort has been made to understand its structural properties and the origin of its thermal stability.

In 1951, Pauling and Corey first proposed that collagen is composed of three parallel polypeptide chains that coiled into a helix.¹⁷⁵ Their model was based on careful analysis of the fibre X-ray diffraction data available at the time. Three years later, Kartha and Ramachandran improved upon Pauling's model by proposing that the polypeptide strands each adopt a left-handed helix and twist together to form a right-handed helix.¹⁶⁴ They also predicted that the glycine residue, known to be at every third position, would have to be placed at the interior of the triple helix axis to allow for the close association of the three strands. A final refinement to the model was made a year later in 1955 by Rich and Crick, who proposed a single systematic hydrogen bond between the polypeptide strands.¹⁶⁷

It was not until 1994, nearly 40 years later, that the structural model established by these researchers was confirmed through the work of Berman, Brodsky and coworkers, who solved the single crystal structure of collagen for high resolution analysis (0.19 nm).¹⁶⁹ They affirmed that there are three left-handed strands twisted into a right-handed triple helix, with a single hydrogen bond between the strands per tripeptide repeat (from the glycine *NH* to the proline carbonyl group). What was not expected was that a hydration network surrounded the triple helices involving hydrogen bonds with the 4-hydroxyl groups of hydroxyproline residues (Figure 1.7.1.2).

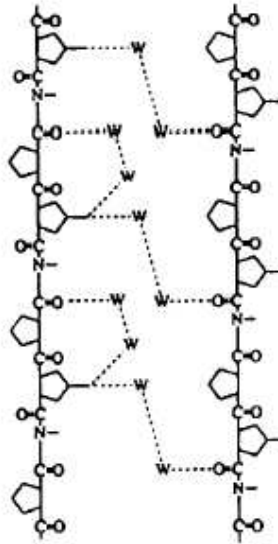


Figure 1.7.1.2: Illustration of water mediated inter- and intrachain hydrogen bonds to the 4-hydroxyl group of hydroxyproline residues observed in the crystal structure of Bella's collagen model peptides. [Reproduced with permission, *Science* **1994**, 266, 75. Copyright 1994 the American Association for the Advancement of Science.]

As it was known that collagen stability is highly correlated to hydroxyproline content,¹⁷⁶⁻¹⁷⁹ and that collagen folding is primarily an enthalpic process,¹⁸⁰ it was postulated that hydroxyproline stabilizes the collagen triple helix through water-mediated hydrogen bonds. This spurred interest in understanding the role of 4-hydroxyproline in stabilizing the collagen triple helix. It has since been found that the stability of collagen is highly sensitive to hydroxyproline configuration¹⁸¹ and location¹⁸² in the tripeptide sequence. In a similar manner to the study of glycopeptides and glycoproteins, much of the understanding of collagen has arisen from the study of collagen model peptides.

1.7.2 Understanding the Role of 4-Hydroxyproline in Collagen Stability

Brodsky and coworkers varied all 20 natural amino acids in the X_{aa} and Y_{aa} positions of collagen model 'guest' peptides of the form $Ac-(Gly-Pro-Hyp)_3-Gly-X_{aa}-Y_{aa}-(Gly-Pro-Hyp)_4-Gly-Gly-NH_2$.¹⁸³ They found that of all the amino acid combinations, L-proline in the X_{aa} position and 4-hydroxy-L-proline in the Y_{aa} positions of the guest peptides provided the most stable collagens according to thermal melting experiments using circular dichroism (CD), where ellipticity is monitored as a function of temperature. Of note is that arginine in the Y_{aa} position is nearly as effective at stabilizing the collagen triple helix as Hyp. However, for most species, large amounts of hydroxyproline indicate it is more relied upon for the stabilization of collagen, with the exception of several species of bacteria that form stable collagen molecules without the presence of hydroxyproline.¹⁸⁴

In 1976, Inouye, Sakakibara, and Prockop found that exchange of 4*S*-hydroxyproline in place of 4*R*-hydroxyproline in the Y_{aa} position of model peptides ($Pro-Y_{aa}-Gly$)₁₀ precludes formation of the collagen triple helix according to similar thermal melting experiments (Table 1.7.2.1).¹⁸¹ Additionally, exchanging proline and 4-hydroxyproline in the X_{aa} and Y_{aa} positions in collagen model peptides of the form ($X_{aa}-Y_{aa}-Gly$)₁₀, greatly reduced their melting temperature (T_m).¹⁸²

Table 1.7.2.1: Moving 4*R*-hydroxyproline (Hyp) from the Y_{aa} to the X_{aa} position or replacement with 4*S*-hydroxyproline (hyp) affects triple helix formation in collagen model peptides.

Model Peptide	T _m (± 2 °C) ^a
[Pro-Pro-Gly] ₁₀	41
[Pro-(4 <i>R</i>)-Hyp-Gly] ₁₀	69
[(4 <i>R</i>)-Hyp-Pro-Gly] ₁₀	<5
[Pro-(4 <i>S</i>)-hyp-Gly] ₁₀	n.t. ^b

^aCarried out in 50 mM acetic acid; ^bNo thermal transition observed.

Clearly, the position and stereochemistry of 4-hydroxyproline is very important for stabilizing collagen. For many years, this was attributed to a water network mediated through hydrogen bonds to 4-hydroxyproline.¹⁸⁵ However, based on this theory, small changes in position and stereochemistry should not greatly affect the stability of the collagen triple helix as is observed. The work of Raines and coworkers gave some insight into this apparent contradiction.

1.7.3 Insight into the Stabilization of Collagen from 4-Fluoroproline

Raines and coworkers have significantly contributed to understanding the role of hydroxyproline in collagen and related peptides through the use of 4-fluoroprolines. Fluorine and hydrogen are considered to be isosteric despite minor differences in van der Waals radius, bond length and dipole moment of the C-F bond.¹⁸⁶ However, organic fluorine is not expected to form hydrogen bonds and thus has little effect on aqueous

solvation.^{187,188} Using model peptides of the form *N*-acetyl-proline-*N'*-methyl ester, it was found that in a similar fashion to 4*R*-Hyp, 4*R*-fluoroproline (Flp) **79** induces a C^γ-exo pucker and shifts the *N*-terminal amide equilibrium to favour the *trans* amide conformation ($K_{trans/cis}$ of 6.1 versus 6.7 in D₂O at 25 °C, respectively) (Figure 1.7.3.1).¹⁵⁴

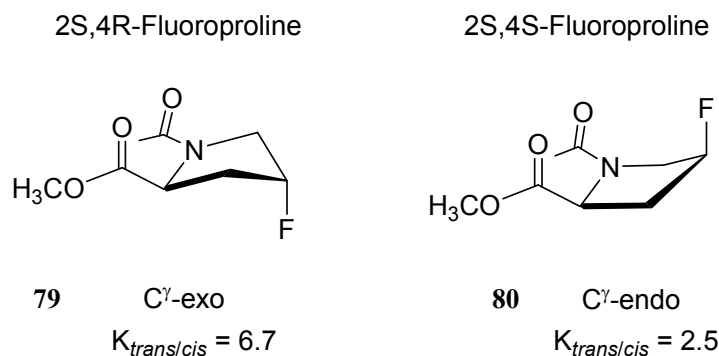


Figure 1.7.3.1: 4-Fluorination of proline has similar effects to 4-hydroxylation.

Also, similar to 4*S*-hyp, 4*S*-fluoroproline **80** induces a C^γ-endo pucker and a higher *cis* amide content ($K_{trans/cis}$ of 2.4 versus 2.5 in D₂O at 25 °C, respectively) relative to L-proline (with a $K_{trans/cis}$ of 4.6). Raines justified the similar changes in proline pucker as arising through ‘gauche effects’. When two electronegative atoms are vicinal to each other, they prefer to adopt a gauche conformational relationship.¹⁸⁹ This allows for maximal σ - σ^* electron donation through neighbouring hydrogen atoms, and favours a given pucker depending on the stereochemistry at the 4-position group (Figure 1.7.3.2).

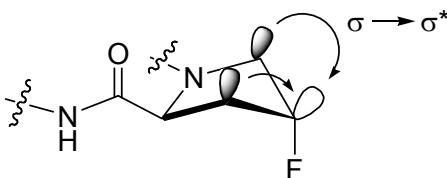


Figure 1.7.3.2: Proposed ‘gauche effect’: The proline side chain will adjust in order to place electronegative atoms in a gauche orientation to maximize σ - σ^* electron donation.

This stereoelectronic effect has different implications for $K_{trans/cis}$ depending the alignment of the favourable $n \rightarrow \pi^*$ interaction from the lone pair on the prolyl *N*-terminal amide carbonyl oxygen to antibonding orbital of the prolyl *C*-terminal carbonyl carbon caused by the different puckers (Section 1.5.1).¹⁵⁴ The C^γ -exo pucker induced by 4*R*-hydroxyproline causes a ψ -angle of 141° , which favours the $n \rightarrow \pi^*$ interaction, and stabilizes the *trans* amide isomer, raising $K_{trans/cis}$. In contrast, for 4*S*-hydroxyproline, the C^γ -endo pucker disfavours the optimal ψ -angle, perhaps of electrostatic repulsion between the 4-hydroxy group and the *C*-terminal carbonyl oxygen atom, thereby disfavours the $n \rightarrow \pi^*$ interaction, which destabilizes the *trans* amide isomer and lowers $K_{trans/cis}$.¹⁹⁰

The proline analogue 4-fluoroproline also seems to cause an increase in the rate of *N*-terminal amide isomerization.¹⁵⁵ This was explained in several ways, which centered on the fact that fluorine is a more electronegative atom than oxygen, and therefore can inductively withdraw more electron density from the prolyl nitrogen atom (Figure 1.7.3.3).

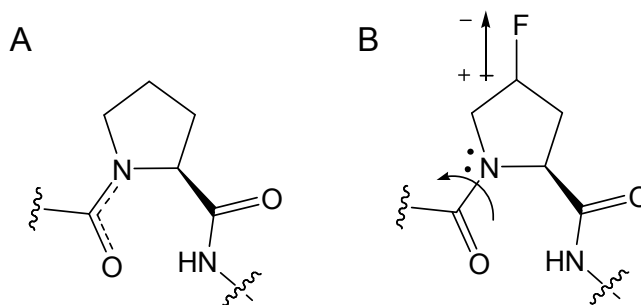


Figure 1.7.3.3: Inductive effect on prolyl isomerization: A) Delocalized electron density in the prolyl *N*-terminal amide bond; B) Fluorination causes an inductive effect which is proposed to shift electron density onto the prolyl nitrogen, thereby increasing C-N single bond character and the rate of isomerization.

Evidence to support this theory came from finding a lower pKa of the prolyl nitrogen for 4-fluoroproline in model peptides of the form *N*-acetyl-proline-*N'*-methyl ester. A decrease in pKa indicates a shift of electron density from the C-N bond onto the prolyl nitrogen.^{191,192} Also, the *N*-terminal amide carbonyl amide I vibrational mode was shifted to higher wavenumbers, which indicates an increase in carbonyl double bond character and implies an increase in single bond character of the amide C-N bond.^{193,194} Thus, through inductive electron-withdrawing effects, fluorine decreases the C-N double bond character, and lowers the barrier to amide isomerization. This effect was further supported by the work of Moroder and coworkers, who studied 4,4-difluoroproline model peptides.¹⁹⁵ They found that the inductive effect is additive, as the difluorinated proline analogues have even higher isomerization rates compared to the 4*R*- or 4*S*-fluoroproline model peptides.

Apart from demonstrating the differences between 4-fluoroproline and 4-hydroxyproline, perhaps most importantly, these studies highlight the important difference between L-proline and 4-hydroxy-L-proline. These results give insight into one reason as to why 4*R*-hydroxyproline stabilizes collagen; it greatly favours the *trans* amide conformation, which is required for triple helix formation. The contribution of this effect to the stability of collagen becomes significant when multiple hydroxyproline residues are present. Raines boldly predicted that 4-fluoroproline incorporated into collagen peptides would greatly stabilize the triple helix, which was somewhat optimistic considering the small increase in the stabilization of the *trans* conformer from 4*R*-fluoroproline ($K_{trans/cis}$ of 6.7 in D₂O at 25 °C) compared to 4*R*-hydroxyproline ($K_{trans/cis}$ of 6.1 in D₂O at 25 °C).¹⁵⁵ However, he was proven correct.

In 1998, Raines and coworkers published a seminal paper in the journal *Nature* in which he confirmed that in model peptides of the form (Pro- Y_{aa} -Gly)₁₀, the replacement of 4*R*-hydroxyproline by 4*R*-fluoroproline in the Y_{aa} position greatly increased the stability of the collagen triple helix.¹⁹⁶ The transition midpoint (T_m) between the triple helix and single strand states for the model peptides progressed from 41 °C for (Pro-Pro-Gly)₁₀ to 69 °C for (Pro-Hyp-Gly)₁₀ to 91 °C for (Pro-Flp-Gly)₁₀ (Table 1.7.3.1).

Table 1.7.3.1: Incorporating 4*R*-fluoroproline (Flp) in the Y_{aa} position stabilizes the collagen triple helix.

Model Peptide	T_m (± 1 °C) ^a
(Pro-Pro-Gly) ₁₀	41
(Pro-(4 <i>R</i>)-Hyp-Gly) ₁₀	69
(Pro-(4 <i>R</i>)-Flp-Gly) ₁₀	91

^aCarried out in 50 mM acetic acid.

Therefore, not only did fluoroproline provide a very stable form of collagen, but it also required that the model established for the reasoning behind the stability of collagen be modified. Since fluoroproline does not participate in hydrogen bonding,¹⁸⁷⁻¹⁸⁸ it means the stability that arises through hydroxyproline content does not arise from hydrogen bonding through water networks alone. Rather, Raines attributed the stability of the collagen triple helix to inductive effects arising from the 4-position fluorine substituent, which favours the *N*-terminal *trans* amide conformation. Raines speculated that chemical modification of the Hyp hydroxyl groups with electron-withdrawing substituents could enhance the stability of natural collagen, allowing new biomaterials to be produced.

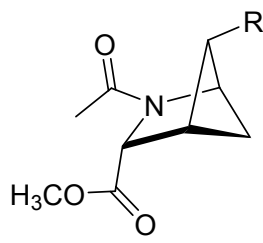
However, the theory that inductive effects are the sole factor in the stabilization of the prolyl *trans* amide conformation now seems unlikely. The study by Moroder and coworkers of 4,4-difluoroproline found that while an inductive effect seems to be additive for the rate of isomerization, it is not for the prolyl amide equilibrium.¹⁹⁵ The $K_{trans/cis}$ for Ac-4,4-F₂Pro-OMe in water at 20 °C is 3.8, which is between that of Ac-4*R*-Flp-OMe ($K_{trans/cis}$ of 8.0) and Ac-4*S*-Flp-OMe ($K_{trans/cis}$ of 2.8) under similar conditions. If an inductive effect was directly responsible for the stabilization of the *trans* amide isomer, then a change in the stereochemistry at the 4-position would not drastically change the $K_{trans/cis}$ as was observed. Clearly, there are other factors mediating inductive effects and the prolyl amide equilibrium.

1.7.4 Contributions of the Proline Pucker to the Stabilization of Collagen

Further studies by Raines found that stabilization of the *trans* amide conformation arises indirectly from inductively electron-withdrawing groups favouring a particular proline pucker, which depends on the stereochemistry of the substituent.¹⁵⁴ This influence of the stereochemistry of side-chain substituents on proline pucker has been known for many years, but it was Raines who reasoned that certain puckers allow for optimal alignment of the *N*-terminal amide carbonyl group to the *C*-terminal amide carbonyl to donate electron density in an $n \rightarrow \pi^*$ fashion. It is known that the ψ -angle in a C γ -exo conformation is 141°, which is near the ideal ψ -angle of 150° for alignment of the C=O to C=O bonds in the Burgi-Dunitz trajectory for a nucleophile approaching an electrophilic center (Section 1.5.1).¹⁹⁷ Thus, the C γ -exo conformation would be expected to favour the *N*-terminal *trans* amide conformation to exploit this interaction. Conversely,

in the C^γ -endo conformation, unfavourable electrostatic clashes between the non-bonded electrons on the fluoro and ester groups would be expected to disfavour this $n \rightarrow \pi^*$ interaction.¹⁵⁴

A study by Raines using 2-acetyl-2-azabicyclo[2.1.1]hexane-3-carboxylic acid methyl ester (Ac-methano-Pro-OMe) was an attempt to tease apart the influence of inductive effects from pucker on $K_{trans/cis}$.¹⁹⁸ Since the pucker of 2-azabicyclo[2.1.1]hexane is fixed, attaching fluoro and hydroxyl groups at the *pseudo-C $^\gamma$* -position revealed the effects of induction without any change in the pucker (Figure 1.7.4.1).



R	$K_{trans/cis}$
H	3.5
OH	3.6
F	3.5

Figure 1.7.4.1: Amide isomer ratio of bridged model peptides in D_2O at 25 °C.

Interestingly, no change in the amide equilibrium was observed upon modification of the side chain by fluoro or hydroxyl groups in the bridged model peptides (all had a $K_{trans/cis}$ of 3.5 in D_2O at 25 °C). This study confirmed that changes in $K_{trans/cis}$ arise from changes in the proline pucker, and not from inductive effects. Furthermore, this work implicates the stabilization of the proline pucker as a key determinant in the stability of the collagen triple helix.

Recent high-resolution crystal structures of collagen model peptides grown in microgravity have confirmed that proline residues in the X_{aa} position of the collagen

(Gly- X_{aa} - Y_{aa}) tripeptide sequence adopt a C^γ -endo conformation, while hydroxyproline in the Y position adopts a C^γ -exo conformation.¹⁹⁹ This reinforces the theory that for optimal packing of the polypeptide strands, adoption of the correct pucker is important.^{76,190} The authors modelled 4*S*-hydroxyproline in the X_{aa} position of the (Gly- X_{aa} - Y_{aa}) repeat, and found it created unfavourable steric interactions to the neighbouring chain.¹⁹⁹ Therefore, it can be extrapolated that 4*R*-hydroxylation of proline favours the *trans* amide conformation, but also a C^γ -exo conformation, which explains why this residue stabilizes collagen in the Y_{aa} position and destabilizes collagen when in the X_{aa} position.

As a natural progression, by predisposing the amino acids in the (Gly- X_{aa} - Y_{aa}) sequence to their favoured puckers for triple helix formation, one would expect to create a more stable triple helix. Raines and coworkers synthesized (Gly-4*S*-flp-4*R*-Flp)₇ model peptides, since 4*S*-fluoroproline (denoted with the lower-case acronym flp) is known to adopt a C^γ -endo conformation and 4*R*-fluoroproline (Flp) is known to favour a C^γ -exo conformation.²⁰⁰ Interestingly, this approach actually caused a decrease in the stability of the collagen peptides. The CD spectrum of the model peptide (Gly-4*S*-flp-4*R*-Flp)₇ had a T_m that was 8 °C lower than both (Gly-4*S*-flp-Pro)₇ (T_m of 33 °C) and (Gly-Pro-4*R*-Flp)₇ (T_m of 45 °C) as determined by CD thermal melting experiments. This result was attributed to unfavourable steric interactions between the X_{aa} position 4*S*-fluoro group to a Y_{aa} position 4*R*-fluoro group in an adjacent chain (Figure 1.7.4.2). Therefore, it seems the steric environment around the amino acid in the X_{aa} position is very crowded in the triple helix.

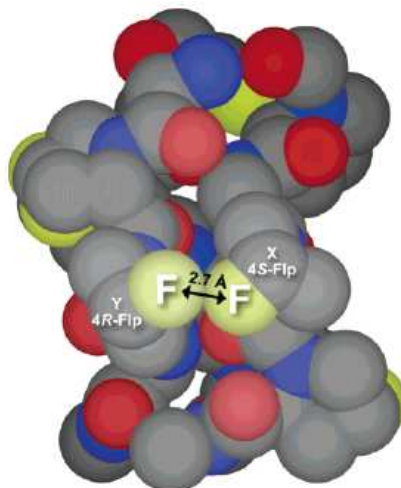


Figure 1.7.4.2: Space filling model of (Gly-4S-flp-4R-Flp)₇ with the proposed unfavourable F...F interstrand interaction precluding triple helix formation. [Reproduced with permission, *J. Am. Chem. Soc.* **2005**, *127*, 15923. Copyright 2005 American Chemical Society.]

1.7.5 Hydroxyproline Analogues Developed to Modulate the Stability of Collagen

Raines and coworkers further explored the effect of proline pucker on collagen stability by using a steric influence rather than a stereoelectronic one to pre-organize the puckers of the X_{aa} and Y_{aa} position residues to create optimal alignment of the collagen polypeptide strands. The proline analogue 2*S*,4*R*-methylproline (denoted with the lower-case acronym mep) **82** was used to induce a C^γ-endo conformation and 2*S*,4*S*-methylproline (Mep) **81** to induce a C^γ-exo conformation (Figure 1.7.5.1).²⁰¹ The methyl groups prefer to adopt a *pseudo*-equatorial orientation on the pyrrolidine ring, which favours a given proline pucker in absence of any stereoelectronic effects.

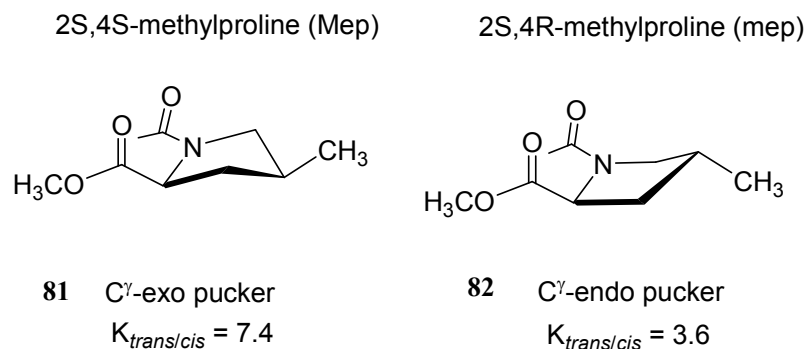


Figure 1.7.5.1: 4-Methylgroups affect the prolyl pucker by adopting a *pseudo*-equatorial orientation; this causes a change in the amide equilibrium in D₂O at 25 °C.

The proline analogues were incorporated into collagen model peptides of the form (X_{aa}-Y_{aa}-Gly)₇ to mirror the preferred puckers in normal collagen. In the (X_{aa}-Y_{aa}-Gly) sequence, 4*R*-methylproline (mep) and 4*S*-methylproline (Mep) were incorporated in the X_{aa} and Y_{aa} positions to induce C^γ-endo and C^γ-exo puckers, respectively. It was found that the collagen model peptide (mep-Mep-Gly)₇ adopted a triple helix based on its CD maxima (Figure 1.7.5.2A) and melting curve showing a two-state transition (Figure 1.7.5.2B), which was equally stable compared to the collagen peptide (Pro-Hyp-Gly)₇; both had a T_m of 36 °C. In this case, the methyl groups are predicted to radiate outward from the triple helix, and not cause any destabilizing steric interactions.

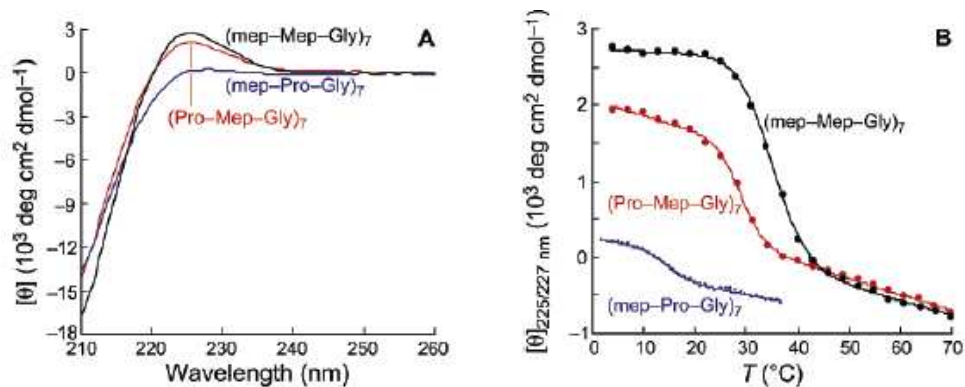


Figure 1.7.5.2: A) CD curve of the collagen model peptides incorporating Mep and mep; the positive maximum at 225 nm is indicative of the polyproline II helix and a triple helical fold; B) Thermal melting curves of (Pro-Mep-Gly)₇ (red), (mep-Pro-Gly)₇ (blue), and (mep-Mep-Gly)₇ (black). [Reproduced with permission, *J. Am. Chem. Soc.* **2006**, *128*, 8112. Copyright 2006 American Chemical Society.]

The stabilization incurred upon incorporation of 4-methylproline seemed to be an additive process. The model peptides (Pro-Mep-Gly)₇ and (mep-Pro-Gly)₇ had melting temperatures of 29 °C and 13 °C, respectively (Figure 1.7.5.2). This could be a result of the large disparity in the stabilization of the *trans* amide isomer from the two proline analogues. In *N*-acetyl proline methylester peptide models, the $K_{trans/cis}$ for Mep was 7.4 while the $K_{trans/cis}$ for mep was only 3.6 in D₂O at 25 °C (Figure 1.7.5.1). Thus, even when the puckers in the X_{aa} and Y_{aa} positions are predisposed to triple helix formation, stabilization of a *trans* amide conformation still has an impact on collagen stability. Therefore, there seems to be a duality (or overlap) from the influence of proline pucker and *N*-terminal *trans* amide stabilization.

In this case, the authors proved that steric effects can replicate stereoelectronic effects on pucker conformation. Therefore, attachment of a heteroatom to the prolyl side

chain is not required for triple helix formation. Furthermore, previous work showed that stereoelectronic effects are not additive in (Gly-4*S*-flp-4*R*-Flp)₇ peptides for affecting collagen stability, but the steric influence on pucker from mep and Mep is additive. The authors predict that application of the combined knowledge of both steric and stereoelectronic effects on influencing collagen stability will lead to the development of highly robust collagen for use in biomedicine and biotechnology.

Several other notable approaches have been used to develop synthetic collagens, which include incorporating charged amino acids and acylation of hydroxyproline. Babu and Ganesh demonstrated that by replacement of 4-hydroxyproline by an ionizable group, such as 4*R*-aminoproline (Amp), the stability of the collagen triple helix in model peptides of the form Ac-Phe-(Pro-Amp-Gly)₆-NH₂ could exhibit pH dependence (Table 1.7.5.1).²⁰² At pH 3.0, the T_m for the model peptide Ac-Phe-(Pro-(4*R*)-Amp-Gly)₆-NH₂ was found to be 60.0 °C compared to 27.0 °C for the model peptide with Hyp at the Y_{aa} position. Similarly, at pH 12.0, the T_m was much higher for the model peptide which had Amp in the Y_{aa} position (T_m of 56.6 °C compared to 39.6 °C); while at pH 9.0, there was no difference in the stability of the model peptide triple helices.

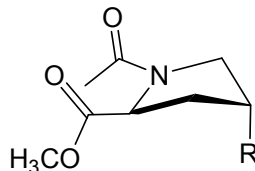
Table 1.7.5.1: Effect of pH on the stability of collagen model peptides incorporating 4-aminoproline and 4-hydroxyproline in the Y_{aa} position.

Model Peptide	T_m (± 0.5 °C) ^a			
	pH 3.0	pH 7.0	pH 9.0	pH 12.0
Ac-Phe-(Pro-(4 <i>R</i>)-Hyp-Gly) ₆ -NH ₂	27.0	28.0	27.0	27.0
Ac-Phe-(Pro-(4 <i>R</i>)-Amp-Gly) ₆ -NH ₂	60.0	56.5	26.0	49.0

^apH 3.0: 20 mM acetate buffer; pH 7.0: phosphate buffer; pH 9.0 and 12.0: 20 mM borate buffer.

The authors reasoned that at low pH, the protonated amine could enhance electron-withdrawing effects similar to 4-fluoroproline and also increase hydrogen bonding potential. At high pH, stabilization could also arise through hydrogen bonding interactions and the lack of charged repulsion between amino groups. Therefore, hydrogen bonding interactions from the Y_{aa} position amino acids in collagen model peptides was attributed to having a large impact on the stability of the collagen triple helix. The Amp proline analogue was proposed to be an improvement over 4-fluoroproline, since replacement of the hydroxyproline 4-hydroxyl group by a 4-amino group is more feasible chemically for naturally-occurring tropocollagen molecules.

Raines and coworkers investigated the effects of acylation of hydroxyproline on collagen stability. Using model peptides of the form *N*-acetyl proline methylester, comparison of the proline C^γ -carbon ^{13}C chemical shifts was done to estimate the electron-withdrawing effect from the acyl group (Figure 1.7.5.3).



R	$K_{trans/cis}$	$^{13}C^\gamma$ (ppm)
OH	4.0	70.4
OCH ₃	3.3	73.7
OCF ₃	3.5	78.5

Figure 1.7.5.3: 4-*O*-Acetylation of proline causes little change in the amide equilibrium, but does affect the chemical shift of the C^γ -atom at 25 °C in 1,4-dioxane- d_8 .

It was found that there was an incremental increase in C^γ chemical shift upon acylation, from hydroxyproline (70.4 ppm) to acylated hydroxyproline (73.7 ppm) to trifluoroacetyl hydroxyproline (78.5 ppm). These shifts were all less than was observed for

4-fluoroproline (93.3 ppm). Also, the stereochemistry of the 4-position substituent had no effect on the chemical shift, which indicates that it is an inductive electron-withdrawing effect.

Acylation did not appear to have a significant effect of the *N*-terminal amide equilibrium: (4*R*)-*O*-acetylproline exhibited a $K_{trans/cis}$ of 3.3 ± 0.3 compared to 4.0 ± 0.4 for 4*R*-hydroxyproline in model peptides in 1,4-dioxane at 25 °C (Figure 1.7.5.3). Trifluoroacetyl modification of 4-hydroxyproline had no further effect on prolyl *N*-terminal amide equilibrium (3.3 ± 0.4). It was reasoned that the relatively small electron-withdrawing effect would not translate into a measurable difference in pyramidalization of the prolyl nitrogen atom. Finally, acylation did not affect the preferred C^γ -exo and C^γ -endo puckers of Hyp and hyp, respectively. Based on these results, the authors asserted that acylation would not preclude triple helix formation and could enhance the stability of the collagen triple helix.

The model peptide (Pro-Hyp-Gly)₁₀ was reacted with acetyl chloride in 6M HCl to form the corresponding acylated collagen model peptide (Pro-Hyp(OAc)-Gly)₁₀. Mass spectrometry was used to confirm complete conversion to the target peptide. Raines and coworkers found that the acetylated collagen model peptide (Pro-Hyp(OAc)-Gly)₁₀ had a T_m of 57.5 °C compared to 69 °C for (Pro-Hyp-Gly)₁₀. Based on simple modelling, the decrease in stability of the acylated triple helix was attributed to steric repulsion between acyl groups and neighbouring chains. In summary, despite a small decrease in thermal stability, acylation of hydroxyproline does not preclude formation of the collagen triple helix.

1.7.6 Effects of Carbohydrates on Collagen Structure and Stability

In nature, hydroxyproline residues in collagen are not known to be glycosylated; the effects of this modification have never been explored. However, 5*R*-hydroxy-L-lysine residues in the Y_{aa} position of the collagen (Gly-X_{aa}-Y_{aa})_n repeat are known to be β-*O*-glycosylated with the glycans D-galactose and D-glucosylgalactose (Section 1.3.2).²⁰³⁻²⁰⁸ These *O*-linked glycans have been implicated in the secretion and assembly of collagen fibrils,²⁰⁹⁻²¹⁴ embryonic development and cell viability,²¹⁵⁻²¹⁶ and the interaction of collagen with protein receptors,²¹⁷⁻²²¹ however, a complete understanding of their function has yet to be achieved. Interest in understanding the effects of glycosylation of collagen also stems from the knowledge that non-enzymatic glycosylation (glycation) of collagen has been implicated in diseases such as diabetes mellitus^{173,174,222-226} and rheumatoid arthritis.^{172,227} Several studies have looked at the implications of non-specific glycation of collagen fibres.

Brodsky and coworkers used solutions of D-ribose (0.2 M) to study the structural implications of non-specific glycation on collagen isolated from rat-tail tendon.²²² They found this process had several effects, notably an expansion of the space between collagen fibres by ~10 % perpendicular to the axis of the collagen fibres. This was attributed to cross-link formation between collagen fibres mediated by the sugars, which pushes the collagen molecules apart. It was also found that the collagen fibres became increasingly insoluble with prolonged incubation with D-ribose. Furthermore, the authors observed disruption of the crystalline packing of the collagen fibres, which is associated to a change in the mechanical properties of collagen and compromises normal interactions with other components of the extracellular matrix. Therefore, there were

significant structural changes that resulted from glycation, which have implications for the effects of disease on collagen physiology. However, in this case the chemical modification of collagen was poorly controlled and characterized.

Interestingly, in more detailed studies of collagen model peptides, glycosylation has been shown to stabilize the collagen triple helix. The cuticle collagen of a deep-sea worm, *Riftia pachyptila*, which lives near hydrothermal vents, has a β -*O*-galactosylated threonine residue in place of hydroxyproline in Y_{aa} position of the collagen (Gly- X_{aa} - Y_{aa}) repeat (Figure 1.7.6.1).^{228,229}

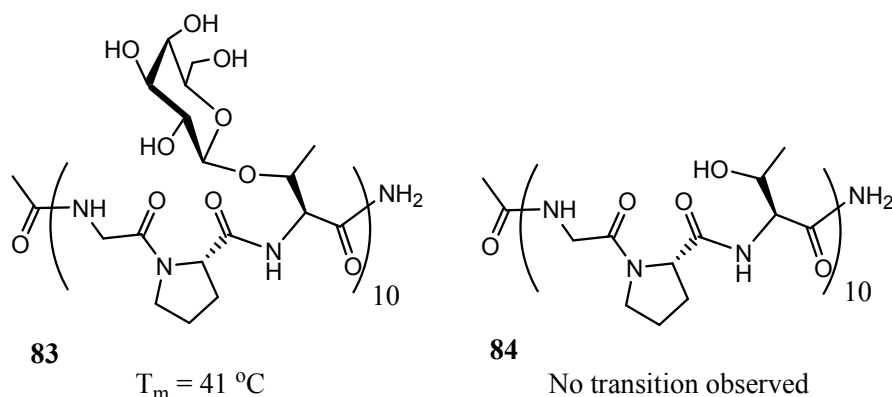


Figure 1.7.6.1: The glycosylated collagen model peptide Ac-(Gly-Pro-Thr(β -D-Gal))₁₀-NH₂ forms a stable triple helix, while Ac-(Gly-Pro-Thr)₁₀-NH₂ does not.

A study carried out by Bächinger and coworkers found that the carbohydrate was essential for formation of the collagen triple helix.^{230,231} While the model peptide Ac-(Gly-Pro-Thr(β -D-Gal))₁₀-NH₂ had a T_m of 41 °C, the peptide Ac-(Gly-Pro-Thr)₁₀-NH₂ did not exhibit a thermal transition (Figure 1.7.6.2).

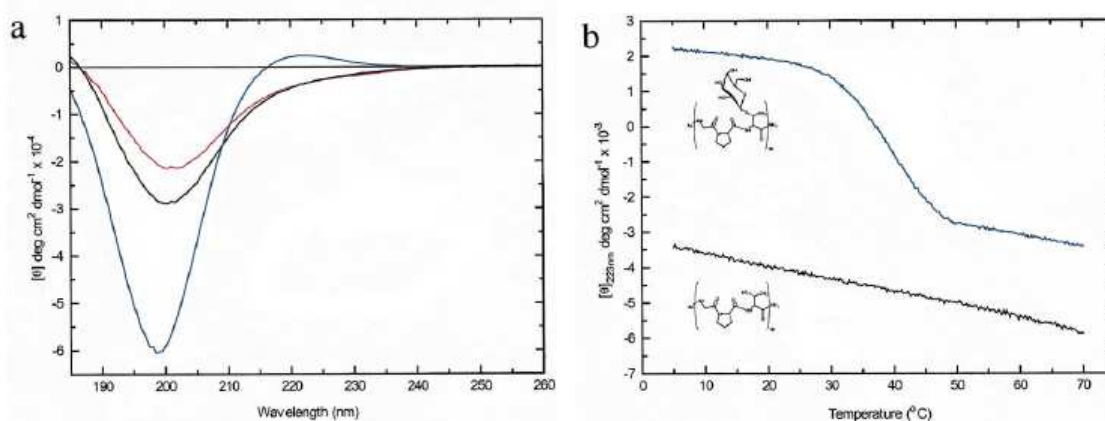


Figure 1.7.6.2: A) CD curves of Ac-(Gly-Pro-Thr)₁₀-NH₂ (black), Ac-(Gly-Pro-Thr(β-D-Gal))₁₀-NH₂ (blue) at 5 °C, and Ac-(Gly-Pro-Thr(β-D-Gal))₁₀-NH₂ (red) at 70 °C; B) Thermal melting curves shows a two-state transition for Ac-(Gly-Pro-Thr(β-D-Gal))₁₀-NH₂ (blue) and no transition for Ac-(Gly-Pro-Thr)₁₀-NH₂ (black). [Reproduced with permission, *FEBS letters*, **2000**, 473, 237. Copyright 2000 Elsevier Science B. V.]

Therefore, it seems that the glycosylated threonine residue somehow stabilizes the triple helix in place of hydroxyproline. The authors proposed that in a similar manner to other *O*-glycosylated peptides, such as the mucins, the galactose residue instils an extended, rigid structure on the polypeptide strands. The galactose residue may also stabilize the triple helix through hydrogen bonding to the polypeptide backbone. This theory was supported by low exchange rates of backbone NH protons, which indicates that the sugar may shield the polypeptide backbone from the solvent. Bächinger suggested that in order to clarify the mechanism of stabilization incurred through glycosylation, further studies were required to determine if glycosylation affects the *cis-trans* isomerization of the neighbouring proline residue and/or affects the conformation of the individual collagen strands.

1.8 Hydroxyproline-Rich Glycoproteins

1.8.1 Hydroxyproline is Glycosylated in Plants Cell Wall Proteins

Hydroxyproline is *O*-glycosylated in plants and algae, a fact first discovered by Derek Lamport in 1967.²³² It is now known that hydroxyproline-rich glycoproteins (HRGPs) are extremely widespread in the plant kingdom and are the major protein components of plant cell walls, which form protective extracellular networks.^{233,234} These proteins are characterized by large amounts of hydroxyproline and serine, a polyproline II (PPII) type conformation (see section 1.8.2) and glycosylation by D-arabinose and D-galactose.²³²⁻²³⁴ Hydroxyproline arabinosides are usually small, linear β -linked homooligosaccharides of 1-4 residues. Galactosylation of hydroxyproline results in the addition of larger β -linked heteropolysaccharides of D-arabinose and D-galactose.²³⁵ Characterization of HRGPs by Homer and Roberts in 1979 isolated from *Chlamydomonas reinhardtii* revealed that the proteins have a rigid, rod-like polyproline II helical backbone with short oligosaccharide side-chains sticking out like a test tube brush.²³⁶ The functional consequences of glycosylation on HRGPs are still unclear, but efforts have been made to understand this modification.

In 1980, molecular modelling by Lamport showed that β -linked tetra-arabinosides attached to hydroxyproline in polyproline peptides would 'nest' with the polyproline II type helix and form three hydrogen bonds to the carbonyl groups of the peptide backbone.²³⁷ This indicated that the carbohydrate moiety may provide HRGPs with a rod-like structure and high tensile strength, in a similar fashion to the stability provided to collagen from the triple helix structure.

It was Lamport who suggested in 1976 that glycosylation of HRGPs increases their solubility, thermal stability, and resistance to proteolytic degradation through steric hindrance to proteases.²³² These attributes may have been the selection pressure responsible for the evolutionary origin of HRGPs.^{238,239} Evidence of this was exemplified by the work of Esquerré-Tugayé, who found that hydroxyproline and glycosylation levels in plant cell walls are known to increase up to 10-fold after bacterial infection. While it has been accepted that glycosylation is important to the integrity of the plant cell wall and contributes to its stability, few studies have explored in detail the contributions to stability of HRGPs upon hydroxyproline glycosylation.

In 2001, Ferris and coworkers compared the CD spectra of the glycosylated and non-glycosylated GP1 protein isolated from *Chlamydomonas reinhardtii*; an HRGP from the cell wall that carries a complex array of arabinogalactan residues, many of which are branched.²⁴⁰ They found that upon chemical deglycosylation using hydrogen fluoride/pyridine according to the protocol of van Holst and Varner,²⁴¹ analysis by CD showed that there was a significant decrease in the net intensities of the spectrum extrema, which was correlated to a transition from a structured to a less structured molecule (Figure 1.8.1.1). The authors concluded that the carbohydrate side chains reinforce the PPII conformation of GP1.

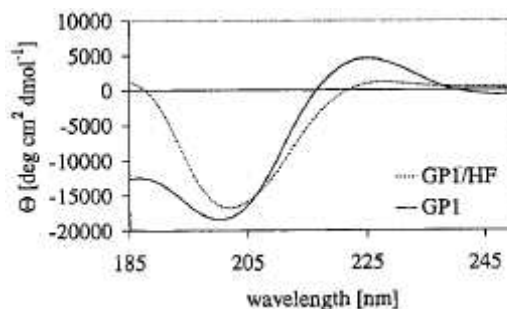


Figure 1.8.1.1: CD spectra of native GPI (solid line) and deglycosylated GPI using HF (dashed line) at pH 7.0 in water. [Reproduced with permission, *Biochemistry* **2001**, *40*, 2978. Copyright 2001 American Chemical Society.]

The short oligoarabinosides and larger arabinogalactan polysaccharides that characterize HRGPs seem to have different effects on PPII conformation. In 2001, Shpak and coworkers studied the extensins, which are a major subgroup of HRGPs identifiable by a repetitive Ser-Hyp₄ sequence, to explore the relationship between contiguous and non-contiguous hydroxyproline sequences on the glycoforms produced.²³⁵ Synthetic genes were delivered into cultured tobacco cells (*Nicotiana tabacum*) to produce polypeptides with sequences of (Ser-Pro₁₋₄)_n. These sequences were hydroxylated by P4H to form contiguous and non-contiguous hydroxyproline residues. It was found that the (Ser-Pro)_n and (Ser-Pro₃)_n sequences led to non-contiguous hydroxylation, and were glycosylated with arabinogalactan polysaccharides. Similarly, (Ser-Pro₂)_n sequences also led to non-contiguous hydroxylation, but were glycosylated with both arabinogalactan polysaccharides and arabinose trisaccharides. Finally, (Ser-Pro₄)_n sequences led to contiguous hydroxylation, and were glycosylated predominantly with arabinose trisaccharides.

The effects of the carbohydrates on the stability of the model peptides was measured using CD by comparing the intensity of the extrema of the glycosylated and deglycosylated peptides. Deglycosylation was carried out using anhydrous hydrogen fluoride. It was found that the polysaccharides intermittently attached to $(\text{Ser-Pro})_n$ and $(\text{Ser-Pro}_3)_n$ sequences caused a slight decrease in the intensity of the CD maxima (Figure 1.8.1.2A and B). Similarly, the mixture of glycans attached to $(\text{Ser-Pro}_2)_n$ caused a slight decrease in CD maxima (Figure 1.8.1.2C). The authors correlated these decreases in intensity to destabilization of the polyproline conformation by intermittent *O*-D-galactose-linked polysaccharides.

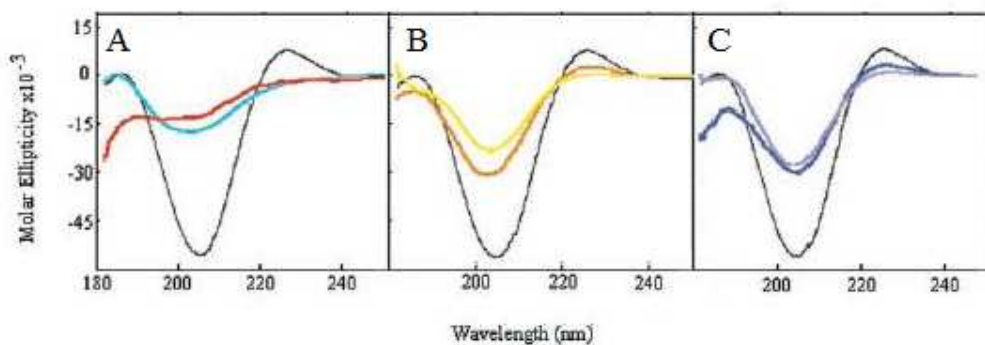


Figure 1.8.1.2: CD curves of the extensin model peptides $(\text{Ser-Pro})_n$, $(\text{Ser-Pro}_2)_n$, and $(\text{Ser-Pro}_3)_n$ in comparison to an ideal polyproline CD curve (black line); A) Comparison of deglycosylated $(\text{Ser-Pro})_n$ (blue) to $(\text{Ser-Pro})_n$ (red); B) Comparison of deglycosylated $(\text{Ser-Pro}_3)_n$ (yellow) to $(\text{Ser-Pro}_3)_n$ (orange); C) Comparison of deglycosylated $(\text{Ser-Pro}_2)_n$ (lavender) to $(\text{Ser-Pro}_2)_n$ (blue) [Reproduced with permission, *J. Biol. Chem.* **2001**, 276, 11276. Copyright 2001 American Society for Biochemistry and Molecular Biology.]

In contrast, the arabinose oligosaccharides attached to the sequence (Ser-Hyp₄)_n caused a marked increase in the CD maxima, which was correlated with a stabilizing effect on the polyproline conformation (Figure 1.8.1.3).

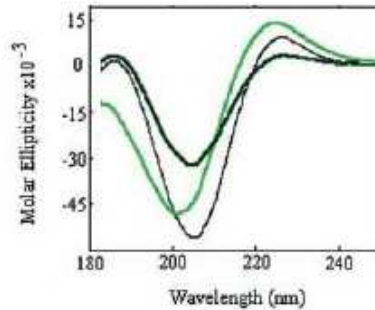


Figure 1.8.1.3: CD curves of the extensin model peptides (Ser-Pro₄)_n in comparison to an ideal polyproline CD curve (black line); deglycosylated (Ser-Pro₄)_n (dark green) to (Ser-Pro₄)_n (light green). [Reproduced with permission, *J. Biol. Chem.* **2001**, 276, 11276. Copyright 2001 American Society for Biochemistry and Molecular Biology.]

Also, a shift from 206 nm to 201 nm in the minima of the negative absorption band was correlated with a direct interaction between the arabinooligosaccharides and the polypeptide backbone, which was attributed to stabilizing the PPII conformation. However, this was not corroborated by any other techniques. The authors concluded that the polypeptide sequence determines the type of glycan attached, which in turn affects the stability of the polyproline conformation of the HRGP and perhaps by extrapolation, the integrity of the plant cell wall. What remains to be investigated in well defined model peptides is if it is the first arabinose or galactose residue that is responsible for the change in polyproline conformation, or the larger glycan structure.

1.8.2 HRGPs Adopt a Polyproline II Conformation

The polyproline conformation is an interesting secondary structure type. Two distinct conformational forms are known to exist: polyproline I (PPI) and polyproline II (PPII).^{242,243} Polyproline I is a more compact right-handed helix characterized by having all *cis* amide bonds, a helical pitch of 5.6 Å/turn, and 3.3 residues per turn.²⁴² By comparison, polyproline II is a more extended left-handed helix with all *trans* amide bonds, a helical pitch of 9.3 Å/turn, and 3 residues per turn (Figure 1.8.2.1).²⁴³

See Figure 1 on page 219 of Creamer, T. P. *Proteins* **1998**, 33, 218-226.

Figure 1.8.2.1: Illustration of an ideal PPII conformation ($\phi = -75^\circ$, $\psi = +145^\circ$); A) Side-view showing the extended conformation; B) Top-down view showing three-fold symmetry.

Therefore, every fourth residue in the PPII helix is stacked on top of each other. Another distinct structural property of the polyproline conformation is that in the PPI

conformation, the carbonyl oxygen atoms are pointed inwards and are shielded by the pyrrolidine rings, while in the PPII conformation, they are pointed outwards and are exposed to the solvent. This means that interconversion between the polyproline I and II conformations can occur upon change in solvent polarity.²⁴⁴⁻²⁴⁶ Polar solvents favour the PPII conformation, while non-polar solvents favour the PPI conformation.

To date, few examples exist of the PPI conformation in biology;²⁴⁷ however, the PPII conformation is common in both folded²⁴⁷⁻²⁵¹ and unfolded proteins,²⁵²⁻²⁶¹ including collagen (Section 1.7.1) and HRGPs (Section 1.8.1). The PPII conformation has now been implicated in numerous protein-protein recognition interactions.²⁶² For example, Src homology 3 (SH3) domain binding sites are found in numerous intracellular proteins, such as kinases, lipases, and GTPases, and mediate protein-protein interactions through a proline-rich 7-9 amino acid sequence that adopts a PPII conformation.²⁶³ Therefore, stabilization of such a PP conformation is of interest in the pursuit of therapeutic applications of specific SH3 ligands. This reflects a broader interest in controlling PPII conformation to understand ligand-binding interactions in the process of developing new medicines.

The rigid nature of the polyproline conformation also lends itself to applications as functional materials, such as molecular ‘rulers’,²⁶⁴⁻²⁶⁶ or as cell-penetrating compounds.²⁶⁷⁻²⁶⁹ However, in order to construct proline oligomers having enhanced biological activity and tailored physical characteristics, a greater understanding of the factors controlling polyproline conformation is required.

In contrast to α -helices and β -sheets, the polyproline conformation exists without the formation of any systematic hydrogen bonding. This presents two problems: first,

assignment of the PP conformation by X-ray crystallography or NMR spectroscopy can be problematic, often leading to its incorrect assignment as a random coil.^{250,270} Circular dichroism is a useful method for identifying the PP conformations. The PPII conformation is characterized by a weak positive band between 220 and 230 nm and a strong negative band between 200 and 210 nm (Figure 1.8.3.1).^{253,271} The PPI conformation is almost a mirror opposite: a weak negative band between 230 and 240 nm and a strong positive band between 210 and 220 nm.

See Figure 2 on page 454 of Kakinoki, S.; Hirano, Y.; Oka, M. *Pept. Sci.* **2005**, *41*, 453-456.

Figure 1.8.3.1: Typical CD spectra of PPII (solid line) and PPI (dashed line) conformations.

Second, the origin for the significant stability of the polyproline conformation remains unclear. Most amino acids can adopt the PP conformation, but have different propensities to do so.²⁷²⁻²⁷⁷ Proline and hydroxyproline seem to be the most well predisposed to induce the PP conformation, because of favourable ϕ and ψ backbone angles; the ϕ -angle of Pro and Hyp is fixed at -75° because the prolyl side chain is

cyclised onto the polypeptide backbone. Furthermore, a ψ -angle of 150° leads to a favourable $n \rightarrow \pi^*$ interaction from the amide carbonyl oxygen atom (O_{i-1}) to the amide carbonyl group ($C'_i=O_i$) that stabilizes the PP helix. Aside from proline and hydroxyproline, glutamine (Gln), Lysine (Lys) and Arginine (Arg) are the most adept amino acids for stabilizing the PPII helix; this has been attributed to hydrogen bonding interactions between side chains and the peptide backbone.^{247,254} Similarly, solvent-mediated hydrogen bonding has been implicated in hydroxyproline stabilizing the PPII helix in (Hyp-Hyp-Gly)₁₀-NH₂ peptides,²⁷⁸ although the role of water in stabilizing the PP conformation has been debated.^{252,279-283} Proline analogues have been used to learn more about factors affecting the stability of the PP conformation, both for understanding biological processes and in the development of potential functional materials and therapeutics.

1.8.3 Proline Analogues Used to Tune the Polyproline Conformation

The study of polyproline peptides is generally carried out using model peptides, which can give insight into the effects of conformational change even when they are extremely short. It has been shown by spectroscopic and computational methods that polyproline peptides begin to adopt a helical conformation in a sequence with as few as three proline residues,²⁸⁴ however most studies seem to use model peptides of 6-10 residues.

The effects of synthetic modification of polyproline peptides are generally measured by changes in the intensity of the CD maxima. Typically, stronger absorbance is correlated to a stronger propensity to form a specific conformation. Several proline

analogues modified at the γ -position have been used to stabilize the PPII conformation to understand its role in biological processes and for the development of molecular scaffolds.

Protein kinases are responsible for protein phosphorylation, which is a common post-translational modification responsible for regulating protein function. What is still unclear is the conformation that is adopted by substrates that bind to the active site of protein kinases. In 2005, Madalengoitia and coworkers looked to prove that the cGMP-dependent protein kinase (PKG) binds to a PPII type conformation using proline-templated arginine (PTR) amino acids.²⁸⁵ Previous studies by the group had shown that oligoPTR peptides adopt a PPII conformation.²⁸⁶ Using the peptide model Arg-Arg-Gln-Ala-Ile, one, two, or three PTR's were inserted in the first three positions of the peptide. It was found that the control peptide displayed extremely low inhibitory activity (K_i of 10 mM), while successive insertions of PTR increased the K_i to $76 \pm 19 \mu\text{M}$, which is similar to natural peptide substrates (Figure 1.8.3.1).

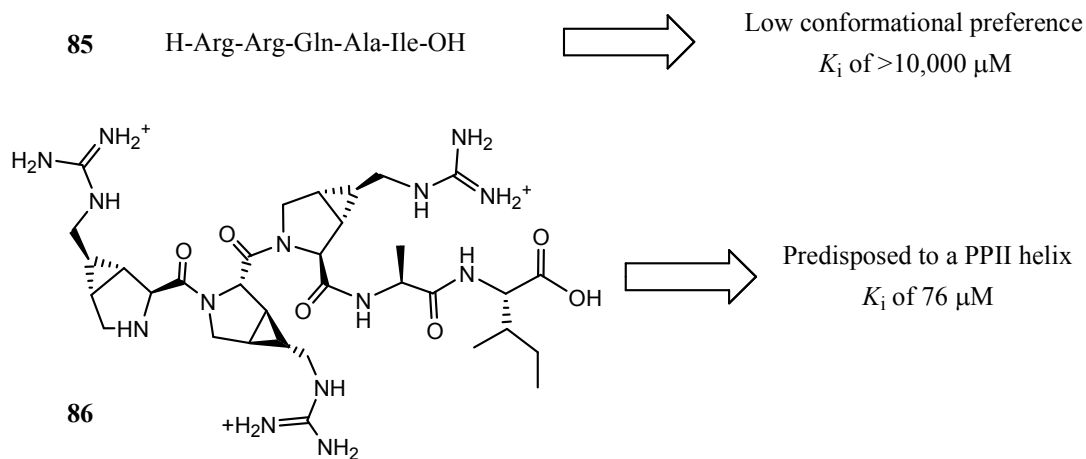


Figure 1.8.3.1: Proline-templated arginine amino acids are predisposed to form a PPII helix, and exhibit increased binding affinity compared to the native pentapeptide.

The authors attributed the increase in binding affinity to a lower net entropy loss upon binding from the conformationally constrained proline analogues, which have backbone torsional angles that approximate the ideal PPII type conformation. The authors concluded that protein kinases such as PKG bind to peptides in a PPII conformation. Other proline analogues, such as 4-azidoproline and 4-fluoroproline have also been shown to stabilize the PPII conformation, but different techniques have been used to establish this fact.

In 2007, Kümin, Sonntag, and Wennemers constructed oligomers of 4-azidoproline (Azp) of the form Ac-(Azp)₉-OH, in order to stabilize the PPII conformation.²⁸⁷ The advantage of Azp is that the azide group can be functionalized using ‘click’ chemistry,²⁸⁸ and so are considered to be a proline-based molecular scaffold. Interestingly, 4-azidoproline exhibits a ‘gauche’ effect similar to 4-hydroxyproline and 4-fluoroproline; therefore, the stereochemistry of the 4-position azido group affects the *N*-terminal amide ratio ($K_{trans/cis}$) (Figure 1.8.3.2).²⁸⁹

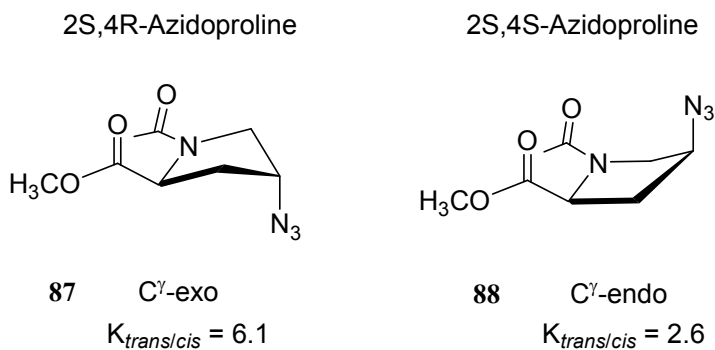


Figure 1.8.3.2: 4*R*- and 4*S*-Azidoproline instil similar changes in the pucker and amide isomer ratio ($K_{trans/cis}$) as 4-hydroxy and 4-fluoroproline in D₂O at 25 °C.

In model peptides of the form *N*-acetyl-proline methylester, the 4*R*-Azp stereoisomer **87** stabilized the *N*-terminal *trans* amide conformation ($K_{trans/cis}$ of 6.1 in D₂O at 25 °C), while the 4*S*-Azp stereoisomer **88** stabilized the *cis* amide conformation relative to the unmodified amino acid ($K_{trans/cis}$ of 2.6 and 4.6 in D₂O at 25 °C, respectively).

In phosphate buffer, the CD spectrum of Ac-(4*R*-Azp)₉-OH exhibited a maximum at 225 nm and a minimum at 207 nm, which is consistent with a PPII conformation.²⁸⁷ The authors also measured the ability of 4*R*-Azp to stabilize the PPII conformation by monitoring its CD spectrum as the % *n*-propanol in water was progressively increased from 0 to 100% (Figure 1.8.3.3).

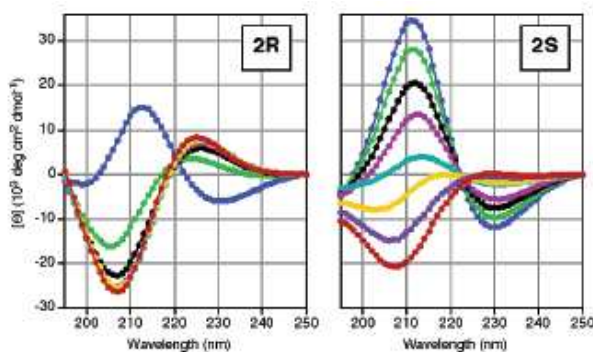


Figure 1.8.3.3: CD spectra of Ac-(4*R*-Azp)₉-OH (left) shows more resistance to changing from PPII to PPI, while Ac-(4*S*-Azp)₉-OH (right) exhibits a more proportional change; phosphate buffer (red); 25% *n*-propanol in buffer (dark purple); 50% *n*-propanol in buffer (yellow); 75% *n*-propanol in buffer (aqua); 85% *n*-propanol in buffer (light purple); 90% *n*-propanol in buffer (black); 95% *n*-propanol in buffer (green); pure *n*-propanol (blue). [Reproduced with permission, *J. Am. Chem. Soc.* **2007**, *129*, 466. Copyright 2007 American Chemical Society.]

As *n*-propanol favours the PPI conformation, this experiment effectively tests the ability of a peptide to resist conformational change. Interestingly, it was found that the CD spectrum of Ac-(4*R*-Azp)₉-OH did not change until the peptide was dissolved in pure *n*-propanol. Thus, the peptide is highly resistant to deviation from the PPII conformation. In contrast, based on the CD spectra, the conformation of Ac-(4*S*-Azp)₉-OH changed immediately towards a PPI conformation with increasing % of *n*-propanol. Thus, the 4*S*-Azp stereoisomer is easily converted to the PPI conformation. The unmodified peptide Ac-(Pro)₉-OH exhibited a more gradual conversion to the PPI type helix with increasing amounts of *n*-propanol. The authors saw their azidoproline peptides as a means to tune the conformational stability of the PPII helix, while retaining the ability to derivatize the azido functional groups for the construction of functional materials.

Horng and Raines used 4-fluoroproline to determine if stereoelectronic effects can be used to stabilize polyproline conformations.²⁹⁰ They incorporated both 4*R*-fluoroproline (Flp) and 4*S*-fluoroproline (denoted with the lower-case acronym flp) into the model peptides H-(Flp)₁₀-Gly-Tyr-OH (denoted (Flp)₁₀) and H-(flp)₁₀-Gly-Tyr-OH (denoted (flp)₁₀). The poly-L-proline homooligomer H-(Pro)₁₀-Gly-Tyr-OH (denoted (Pro)₁₀) and 4*R*-hydroxyproline homooligomer H-(Hyp)₁₀-Gly-Tyr-OH (denoted (Hyp)₁₀) were also made for comparison. Raines and Horng found that the peptides (Pro)₁₀, (Hyp)₁₀, and (Flp)₁₀ exhibited CD spectra consistent with a PPII type conformation in water, while flp seemed to adopt a conformation intermediate between PPII and PPI conformations (Figure 1.8.3.4). This was expected, since 4*R*-fluoroproline stabilizes the *N*-terminal *trans* amide bond, while 4*S*-fluoroproline stabilizes the *N*-terminal *cis* amide bond, and so should favour the PPII and PPI conformations, respectively. Interestingly,

the less intense maxima exhibited by (Flp)₁₀ compared to both (Pro)₁₀ and (Hyp)₁₀ indicates it has less PPII character. However, this does not translate into decreased stability of the PPII conformation.

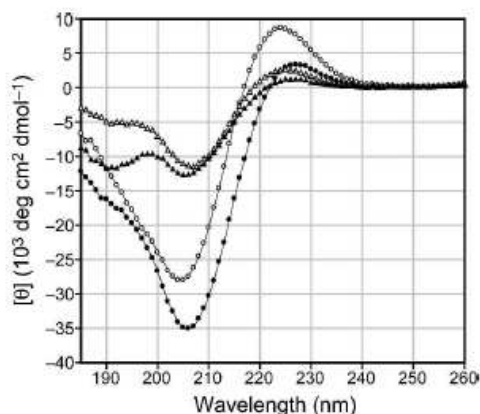


Figure 1.8.3.4: CD spectra of (Pro)₁₀ (closed circles), (Hyp)₁₀ (open circles), (Flp)₁₀ (open triangles), and (flp)₁₀ (closed triangles) in 20 mM phosphate buffer. [Reproduced with permission, *Protein Sci.* **2006**, *15*, 74. Copyright 2006 The Protein Society.]

The ability of (Flp)₁₀ to stabilize the PPII conformation was assessed using temperature-dependent CD measurements. Raines and Horng found an apparent decrease in PPII content with increasing temperature for all the peptides, but without a clear transition to a high temperature state, which is suggested to be an extended β -strand conformation (Figure 1.8.3.5A).²⁹¹⁻²⁹⁸

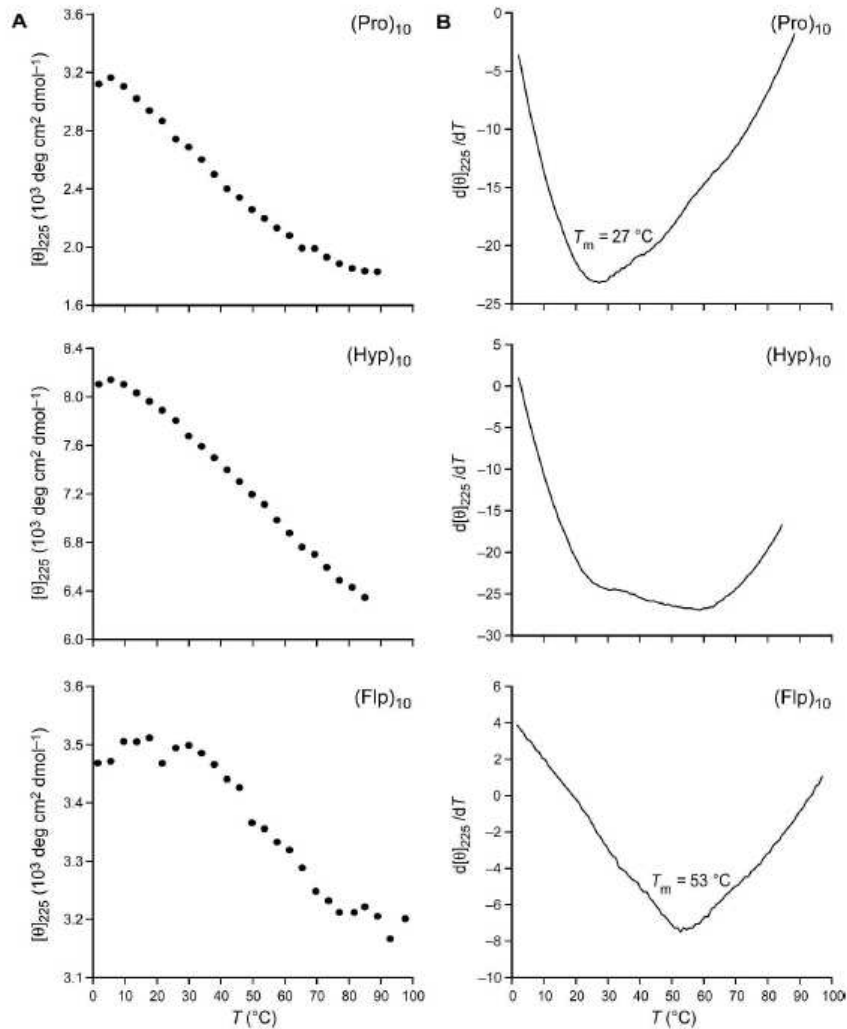


Figure 1.8.3.5: A) Thermal melting curves for (Pro)₁₀, (Hyp)₁₀, and (Flp)₁₀ in 20 mM phosphate buffer; B) First derivative curves of the ellipticity versus temperature in A, and corresponding melting temperatures. [Reproduced with permission, *Protein Sci.* **2006**, *15*, 74. Copyright 2006 The Protein Society.]

Hinderaker and Raines have suggested that the barrier to this transition is due to differential hydration effects and/or hyperconjugation. Since no clear transition was observed, the first derivative curves were used to compare the conformational stability of each PPII helix (Figure 1.8.3.5B). The estimated T_m values were found to be 27 °C for

(Pro)₁₀, and 53 °C for (Flp)₁₀. Because of a highly non-cooperative transition for (Hyp)₁₀, the results were less clear, but the T_m was attributed to being higher than for (Pro)₁₀. Since 4-fluoroproline cannot hydrogen bond, another factor must be responsible for stabilizing the PPII helix. The authors attributed the stabilization of the PPII helix from 4-fluoroproline to stereoelectronic effects in which the 4-position electron-withdrawing group stabilizes the prolyl backbone angles of ϕ (-75°) and ψ (150°), and the *N*-terminal *trans* amide bond, which predisposes the peptide to a PPII conformation.

These examples serve to illustrate that proline analogues can be used to modify the polyproline conformation and its conformational stability; the ability to further derivatize the polyproline peptide is seen as an advantage for developing functional materials such as molecular scaffolds.

1.9 References

- 1) Taylor, M. E.; Drickamer, K. *Introduction to Glycobiology*; Oxford University Press: New York, 2003.
- 2) a) Varki, A. *Glycobiology* **1993**, *3*, 97-130; b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
- 3) a) Ratner, D. M.; Adams, E. W.; Disney, M. D.; Seeberger, P. H. *ChemBioChem* **2004**, *5*, 1375-1390; b) Gabius, H.-J.; Siebert, H.-C.; André, S.; Jiménez-Barbero, J.; Rüdiger, H. *ChemBioChem* **2004**, *5*, 740-745; c) Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *291*, 2370-2371.

4) a) van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 151-208; b) Perez-Villar, J.; Hill, R. L. *J. Biol. Chem.* **1999**, *274*, 31751-31754.

5) Kornfeld, R.; Kornfeld, S. *Ann. Rev. Biochem.* **1985**, *54*, 631-664.

6) Petrescu, A.-J.; Milac, A.-L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103-114.

7) Drickamer, K.; Taylor, M. E. *Annu. Rev. Cell Biol.* **1993**, *9*, 237-264; b) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637-674.

8) Page: 93
Gagneux, P.; Varki, A. *Glycobiology*, 1999, *9*, 747-755.

9) Varki, A.; Cummings, R.; Esko, J.; Freeze, H.; Hart, G.; Marth, J. *Essentials of Glycobiology*; Cold Spring Harbour Laboratory Press, 1999.

10) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 97-101.

11) O'Connor, S. E.; Imperiali, B. *J. Am. Chem. Soc.* **1997**, *119*, 2295-2296.

12) O'Connor, S. E.; Imperiali, B. *Chem. Biol.* **1998**, *5*, 427-437.

13) Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421-8428.

14) Imberty, A.; Pérez, S. *Protein Eng.* **1995**, *8*, 699-703.

- 15) Brockhausen, I. in *Glycoproteins*; (Eds. Montreuil, J.; Vliegthart, J. F. G.; Schachter, H.); Elsevier Science, New York, 1995; pp 201-259.
- 16) Coltart, D. M.; Royyuru, A. K.; Williams, L. J.; Glunz, P. W.; Sames, D.; Kuduk, S. D.; Schwarz, J. B.; Chen, X. T.; Danishefsky, S. J.; Live, D. H. *J. Am. Chem. Soc.* **2002**, *124*, 9833-9840.
- 17) Live, D. H.; Williams, L. J.; Kuduk, S. D.; Schwarz, J. B.; Glunz, P. W.; Chen, X. T.; Sames, D.; Kumar, R. A.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3489-3493.
- 18) Sames, D.; Chen, X.-T.; Danishefsky, S. J. *Nature* **1997**, *389*, 587-591.
- 19) Mer, G.; Hietter, H.; Lefèvre, J.-F. *Nat. Struc. Biol.* **1996**, *3*, 45-53.
- 20) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 97-101.
- 21) Wang, C.; Eufemi, M.; Turano, C.; Giartosio, A. *Biochemistry* **1996**, *35*, 7299-7307.
- 22) Chu, F. K.; Trimble, R. B.; Maley, F. *J. Biol. Chem.* **1978**, *253*, 8691-8693.
- 23) Bernard, E. R.; Sheila, A. N.; Olden, K. *J. Biol. Chem.* **1983**, *258*, 12198-12202.
- 24) Fisher, J. F.; Harrison, A. W.; Bundy, G. L.; Wilkinson, K. F.; Rush, B. D.; Ruwart, M. J. *J. Med. Chem.* **1991**, *34*, 3140-3143.
- 25) Rudd, P. M.; Joao, H. C.; Coghill, E.; Fiten, P.; Saunders, M. R.; Opdenakker, G.; Dwek, R. A. *Biochemistry* **1994**, *33*, 17-22.

- 26) Mehta, S.; Meldal, M.; Duus, J. O.; Bock, K. *J. Chem. Soc., Perkin Trans. 1*, **1999**, 1445-1451.
- 27) Nakakura, N.; Hietter, H.; van Dorsselaer, A.; Luu, B. *Eur. J. Biochem.* **1992**, *204*, 147-153.
- 28) Berger, S.; Menudier, A.; Julien, R.; Karamanos, Y. *Biochimie* **1995**, *77*, 751-754.
- 29) Suzuki, T. in *Glycosciences*; (Eds. Gabius, H. J., Gabius, S.); Chapman & Hall, 1997; pp 121-131.
- 30) Blithe, D. L. *Endocrinology* **1990**, *126*, 2788 -2799.
- 31) Weisshaar, G.; Hiyama, J.; Renwick, A. G. C. *Glycobiology* **1991**, *1*, 393-404.
- 32) Sairam, M. R. *FASEB J.* **1989**, *3*, 1915-1926.
- 33) Iles, R. K.; Chard, T. *J. Mol. Endocrinol.* **1993**, *10*, 217-234.
- 34) de Beer, T.; Van Zuylen, C. W. E. M.; Leeftang, B. R.; Hard, K.; Boelens, R.; Kaptein, R.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1996**, *241*, 229-242.
- 35) Matzuk, M. M.; Boime, I. *J. Cell Biol.* **1988**, *106*, 1049-1059.
- 36) van Zuylen, C.W.; Kamerling, J.P.; Vliegthart, J.F. *Biochem. Biophys. Res. Commun.* **1997**, *232*, 117-120.
- 37) Erbel, P. J. A.; Karimi-Nejad, Y.; de Beer, T.; Boelens, R.; Kamerling, J. P.; Vliegthart, J. F. G. *Eur. J. Biochem.* **1999**, *260*, 490-498.

- 38) Delorme, E.; Lorenzini, T.; Griffin, J.; Martin, F.; Jacobsen, F.; Boone, T.; Elliot, S. *Biochemistry* **1992**, *31*, 9871-9876.
- 39) Meiniohanns, E.; Meldal, M.; Paulsen, H.; Dwek, R. A.; Bock, K. *J. Chem. Soc. Perkin Trans.* **1998**, *7*, 549-560.
- 40) Roberge, J. Y.; Beebe, X.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1998**, *120*, 3915-3927.
- 41) Meldal, M.; St. Hilaire, P. M. *Curr. Opin. Chem. Biol.* **1997**, *1*, 552-563.
- 42) Rodriguez, E. C.; Winans, K. A.; King, D. S.; Bertozzi, C. R. *J. Am. Chem. Soc.* **1997**, *119*, 9905-9906.
- 43) Witte, K., Sears, P., Martin, R. & Wong, C.H. *J. Am. Chem. Soc.* **1997**, *119*, 2114-2118.
- 44) Specker, D.; Wittmann, V. *Top. Curr. Chem.* **2007**, *267*, 65-107.
- 45) Zhu, X.; Schmidt, R. R. *Chem. Eur. J.* **2004**, *10*, 875-887.
- 46) Takahashi, N. *Biochem. Biophys. Res. Commun.* **1977**, *76*, 1194-1199.
- 47) Plummer, T. H., Jr.; Elder, J. H.; Alexander, S.; Phelan A. W.; Tarentino, A. L. *J. Biol. Chem.* **1984**, *259*, 10700-10705.
- 48) Wang, L.-X.; Tang, M.; Suzuki, T.; Kitajima, K.; Inoue, Y.; Inoue, S.; Fan, J.-Q.; Lee, Y. C. *J. Am. Chem. Soc.* **1997**, *119*, 11137-11146.

- 49) Espinosa, J. F.; Montero, E.; Vian, A.; Garcia, J. L.; Dietrich, H.; Schmidt, R. R.; Martin-Lomas, M.; Imberty, A.; Canada, F. J.; Jimenez-Barbero, J. *J. Am. Chem. Soc.* **1998**, *120*, 1309-1318.
- 50) Jimenez-Barbero, J.; Espinosa, J. F.; Asensio, J. L.; Canada, F. J.; Poveda, A. *Adv. Carbohydr. Chem. Biochem.* **2001**, *56*, 235-284.
- 51) Chaulagain, M. R.; Postema, M. H. D.; Valeriote, F.; Pietraszkewicz, H. *Tetrahedron Lett.* **2004**, *45*, 7791-7794.
- 52) Toba, T.; Murata, K.; Yamamura, T.; Miyake, S.; Annoura, H. *Tetrahedron Lett.* **2005**, *46*, 5043-5047.
- 53) Michael, K.; Wittmann, V.; Konig, W.; Sandow, J.; Kessler, H. *Int. J. Pept. Protein Res.* **1996**, *48*, 59-70.
- 54) Peri, F.; Cipolla, L.; Rescigno, M.; La Ferla, B.; Nicotra, F. *Bioconjugate Chem.* **2001**, *12*, 325-328.
- 55) Arnett, F. C.; Edworthy, S. M.; Bloch, D. A.; McShane, D. J.; Fries, J. F.; Cooper, N. S.; Healey, L. A.; Kaplan, S. R.; Liang, M. H.; Luthra, H. S.; Medsger, T. A.; Mitchell, D. M.; Neustadt, D. H.; Pinals, R. S.; Schaller, J. G.; Sharp, J. T.; Wilder, R. L.; Hunder, G. G. *Arthritis Rheum.* **1988**, *31*, 315-324.
- 56) Broddefalk, J.; Bäcklund, J.; Almqvist, F.; Johansson, M.; Holmdahl, R.; Kihlberg, J. *J. Am. Chem. Soc.* **1998**, *120*, 7676-7683.

57) Corthay, A.; Bäcklund, J.; Broddefalk, J.; Michaelsson, E.; Goldschmidt, T. J.; Kihlberg, J.; Holmdahl, R. *Eur. J. Immunol.* **1998**, *28*, 2580-2590.

58) Dzhambazov, B.; Nandakumar, K. S.; Kihlberg, J.; Fugger, L.; Holmdahl, R.; Vestberg, M. *J. Immunol.* **2006**, *26*, 1525-1533.

59) Gustafsson, T.; Hedenström, M.; Kihlberg, J. *J. Org. Chem.* **2006**, *71*, 1911-1919.

60) a) Gustafsson, T.; Saxin, M.; Kihlberg, J. *J. Org. Chem.* **2003**, *68*, 2506-2509; b) Wellner, E.; Gustafsson, T.; Bäcklund, J.; Holmdahl, R.; Kihlberg, J. *ChemBioChem* **2000**, *1*, 272-280.

61) Dondoni, A.; Marra, A. *Chem Rev.* **2000**, *100*, 4395-4421.

62) Gruner, S. A. W.; Locardi, E.; Lohof, E.; Kessler, H. *Chem. Rev.* **2002**, *102*, 491-514.

63) Schweizer, F. *Angew. Chem. Int. Ed.* **2002**, *41*, 230-235.

64) Chakraborty, T. K.; Ghosh, S.; Jayaprakash, S. *Curr. Med. Chem.* **2002**, *9*, 421-435.

65) Schweizer, F. *Trends Glycoscience Glycotechnology* **2003**, *15*, 315-328.

66) Schweizer, F.; Inazu, T. *Org. Lett.* **2001**, *3*, 4115-4118.

67) Page: 98
Ousmer, M.; Boucard, V.; Lubin-Germain, N.; Uziel, J.; Augé, J. *Eur. J. Org. Chem.* **2006**, 1216-1221.

68) Shull, B. K.; Wu, Z.; Koreeda, M. *J. Carbohydr. Chem.* **1996**, *15*, 955-964.

- 69) Lei, H.; Stoakes, M. S.; Herath, K. P. B.; Lee, J.; Schwabacher, A. W. *J. Org. Chem.* **1994**, *59*, 4206-4210.
- 70) Zhang, K.; Schweizer, F. *Synlett* **2005**, *20*, 3111-3115.
- 71) Cipolla, L.; Redaelli, C.; Nicotra, F. *Lett. Drug Des. Discovery* **2005**, *2*, 291-293.
- 72) Thomasson, K.A.; Applequist, J. *Biopolymers* **1990**, *30*, 437-450.
- 73) Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc.* **1973**, *95*, 5725-5729.
- 74) Cai, M.; Huang, Y.; Liu, J.; Krishnamoorthi, R. *J. Biomol. NMR* **1995**, *6*, 123-128.
- 75) Haasnoot, C. A. G.; De Leeuw, F. A. A. M.; De Leeuw, H. P. M.; Altona, C. *Biopolymers* **1981**, *20*, 1211-1245.
- 76) Vitagliano, L.; Berisio, R.; Mastrangelo, A.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2001**, *10*, 2627-2632.
- 77) Song, I. K.; Kang, Y. K. *J. Phys. Chem. B* **2006**, *110*, 1915-1927.
- 78) Benzi, C.; Improta, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 341-350.
- 79) Lam, J. S. W.; Koo, J. C. P.; Hudáky, I.; Varro, A.; Papp, J. G.; Penke, B.; Csizmadia, I. G. *J. Mol. Struct.* **2003**, *666*, 285-289.
- 80) a) Zimmerman, S. S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 408-416; b) Hinderaker, M. P.; Raines, R. T. *Prot. Sci.* **2003**, *12*, 1188-1194.

- 81) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286*, 291-299.
- 82) Weiss, M. S.; Jabs, A.; Hilgenfeld, R. *Nat. Struct. Biol.* **1998**, *5*, 676-677.
- 83) Pal, D.; Chakrabarti, P. *J. Mol. Biol.* **1999**, *294*, 271-275.
- 84) Hodges, J. A.; Raines, R. T. *Org. Lett.* **2006**, *8*, 4695-4697.
- 85) Dugave, C.; Demange, L. *Chem Rev.* **2003**, *103*, 2475-2532.
- 86) Taylor, C. M.; Hardré, R.; Edwards, P. J. B.; Park, J. H. *Org. Lett.* **2003**, *5*, 4413-4416.
- 87) Pao, Y.-L.; Wormarld, M. R.; Dwek, R. A.; Lellouch, A. C. *Biochem. Biophys. Res. Commun.* **1996**, *219*, 157-160.
- 88) Schutkowski, M.; Bernhardt, A.; Zhou, X. Z.; Shen, M.; Reimer, U.; Rahfeld, J.-U.; Lu, K. P.; Fischer, G. *Biochemistry* **1998**, *37*, 5566-5570.
- 89) Scherer, G.; Kramer, M. L.; Schutkowski, M.; Reimer, U.; Fischer, G. *J. Am. Chem. Soc.* **1998**, *120*, 5568-5570.
- 90) Schiene, C.; Reimer, U.; Schutkowski, M.; Fischer, G. *FEBS Lett.* **1998**, *432*, 202-206.
- 91) Stein, R. L. *Adv. Prot. Chem.* **1993**, *44*, 1-24.
- 92) Harrison, R. K.; Stein, R. L. *Biochemistry* **1990**, *29*, 1684-1689.
- 93) Larive, C. K.; Rabenstein, D. L. *J. Am. Chem. Soc.* **1993**, *115*, 2833-2836.

- 94) Caldwell HK, Young WS III *Oxytocin and Vasopressin: Genetics and Behavioral Implications*; in *Handbook of Neurochemistry and Molecular Neurobiology: Neuroactive Proteins and Peptides* (Eds. Lajtha, A.; Lim, R.) 3rd ed. Berlin, Springer; 2006; pp. 573-607.
- 95) Hruby, V. J.; Brewster, A. I.; Glasel, J. A. *Proc. Nat. Acad. Sci. U. S. A.* **1971**, *68*, 45-53.
- 96) Brewster, A. I.; Hruby, V. J.; Spatola, A. F.; Bovey, F. A. *Biochemistry* **1973**, *12*, 1643-1649.
- 97) Deslauriers, R.; Walter, R.; Smith, I. C. P. *Biochem. Biophys. Res. Commun.* **1972**, *48*, 854-859.
- 98) Mariappan, S. V. S.; Rabenstein, D. L. *J. Org. Chem.* **1992**, *57*, 6675-6678.
- 99) Fischer, G.; Schmid, F. X. *Biochemistry* **1990**, *29*, 2205-2212.
- 100) Wedemeyer, W. J.; Welker, E.; Scheraga, H. A. *Biochemistry* **2002**, *41*, 14637-14644.
- 101) Schmid, F. X.; Mayr, L. M.; Mücke, M.; Schönbrunner, E. R. (1993) *Adv. Protein Chem.* **1993**, *44*, 25-66.
- 102) Dugave, C. *Curr. Org. Chem.* **2002**, *6*, 1397-1402.
- 103) Kern, D.; Schutkowski, M.; Drakenberg, T. *J. Am. Chem. Soc.* **1997**, *119*, 8403-8405.

- 104) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221-232.
- 105) Burgess, K. *Acc. Chem. Res.* **2001**, *34*, 826-835.
- 106) Kini, R. M.; Evans, H. J. *Biochem. Biophys. Res. Comm.* **1995**, *212*, 1115-1124.
- 107) Garcia, K. C.; Ronco, P. M.; Verroust, P. J.; Brünger, A. T.; Amzel, L. M. *Science* **1992**, *257*, 502-507.
- 108) Kyle, D. J.; Blake, P. R.; Smithwick, D.; Green, L. M.; Martin, J. A.; Sinsko, J. A.; Summers, M. F. *J. Med. Chem.* **1993**, *36*, 1450-1460.
- 109) Thurieau, C.; Féléto, M.; Hennig, P.; Raimbaud, E.; Canet, E.; Fauchère, J.-L. *J. Med. Chem.* **1996**, *39*, 2095-2101.
- 110) Fox, R. O.; Evans, P. A.; Dobson, C. M. *Nature* **1986**, *320*, 192-194.
- 111) Higgins, K. A.; Craik, D. J.; Hall, J. G.; Andrews, P. R. *Drug. Des. Delivery* **1988**, *3*, 159-170.
- 112) Amodeo, P.; Morelli, M. A. C.; Motta, A. *Biochemistry* **1994**, *33*, 10754-10762.
- 113) Brazin, K. N.; Mallis, R. J.; Fulton, D. B.; Andreotti, A. H. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 1899-1903.
- 114) Lester, H. A.; Dibas, M. I.; Dahan, D. S.; Leite, J. F.; Dougherty, D. A. *Trends Neurosci.* **2004**, *27*, 329-336.
- 115) Lummis, S. C. R.; Beene, D. L.; Lee, L. W.; Lester, H. A.; Broadhurst, R. W.; Dougherty, D. A. *Nature* **2005**, *438*, 248-252.

- 116) Vanhoof, G.; Goossens, F.; De Meester, I.; Hendriks, D.; Scharpé, S. *FASEB* **1995**, *9*, 736-744.
- 117) Reiersen, H.; Rees, A. R. *TIBS* **2001**, *26*, 679-684.
- 118) a) Gante, J. *Angew. Chem. Int. Ed.* 1994, *33*, 1699-1720; b) Vagner, J.; Qu, H.; Hruby, V. J. *Curr. Op. Chem. Biol.* **2008**, *12*, 292-296.
- 119) Hoffman, T.; Lanig, H.; Waibel, R.; Gmeiner, P. *Angew. Chem. Int. Ed.* **2001**, *40*, 3361-3362.
- 120) Crisma, M.; Moretto, A.; Toniolo, C.; Kaczmarek, K.; Zabrocki, J. *Macromolecules* **2001**, *34*, 5048-5055.
- 121) Gosselin, F.; Lubell, W. D. *J. Org. Chem.* **2000**, *65*, 2163-2165.
- 122) Konopelski, J.; Wei, Y.; Olmstead, M. M. *J. Org. Chem.* **1999**, *64*, 5148-5152.
- 123) Curran, T. P.; Marcaurelle, L. A.; O'Sullivan, K. M. *Org. Lett.* **1999**, *1*, 1225-1229.
- 124) Delaney, N. G.; Madison, V. *J. Am. Chem. Soc.* **1982**, *104*, 6635-6642.
- 125) Montelione, G. T.; Hughes, P.; Clardy, J.; Scheraga, H. A. *J. Am. Chem. Soc.* **1986**, *108*, 6765-6768.
- 126) Beausoleil, E.; Lubell, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 12902-12908.
- 127) Beausoleil, E. Sharma, R.; Michnick, S. W.; Lubell, W. D. *J. Org. Chem.* **1998**, *63*, 6572-6578.

- 128) Sharma, R.; Lubell, W. D. *J. Org. Chem.* **1996**, *61*, 202-209.
- 129) Lubell, W.; Rapoport, H. *J. Org. Chem.* **1989**, *54*, 3824-3829.
- 130) Beausoleil, E.; L'Archevque, B.; Blec, L.; Atfani, M.; Lubell, W. D. *J. Org. Chem.* **1996**, *61*, 9447-9454.
- 131) Bélec, L.; Slaninová, J.; Lubell, W. D. *J. Med. Chem.* **2000**, *43*, 1448-1455.
- 132) López-Zeno, J. A.; Peaceman, A. M.; Adashek, J. A.; Socol, M. L. *N. Engl. J. Med.* **1992**, *326*, 450-454.
- 133) Manning, M.; Cheng, L. L.; Klis, W. A.; Balaspiri, L.; Olma, A.; Sawyer, W. H.; Wo, N. C.; Chan, W. Y. *J. Med. Chem.* **1995**, *38*, 1762-1769.
- 134) Williams, P. D.; Anderson, P. S.; Ball, R. G.; Bock, M. G.; Carroll, L. A.; Chiu, S. H. L.; Clinesmidt, B. V.; Culberson, J. C.; Erb, J. M.; Evans, B. E.; Fitzpatrick, S. L.; Freidinger, R. M.; Kaufman, M. J.; Lundell, G. F.; Murphy, J. S.; Pawluczyk, J. M.; Perlow, D. S.; Pettibone, D. J.; Pitzenberger, S. M.; Thompson, K. L.; Verber, D. F. *J. Med. Chem.* **1994**, *37*, 565-573.
- 135) Du Vigneaud, V.; Ressler, C.; Swan, J. M.; Roberts, C. W.; Katsoyannis, P. G. *J. Am. Chem. Soc.* **1954**, *76*, 3115-3121.
- 136) du Vigneaud, V.; Ressler, C.; Swan, C. W.; Katsoyannis, P. G.; Gordon, S. *J. Am. Chem. Soc.* **1953**, *75*, 4879-4880.
- 137) Hruby, V. J.; Chow, M.-S.; Smith, D. D. *Annu. Rev. Pharmacol. Toxicol.* **1990**, *30*, 501-534

- 138) Bélec, L.; Maletinska, L.; Slaninová, J.; Lubell, W. D. *J. Pep. Sci.* **2001**, *58*, 263-273.
- 139) Mutter, M.; Nefzi, A.; Sato, T.; Sun, X.; Wöhr, T. *Pept. Res.* **1995**, *8*, 145-153.
- 140) Haack, T.; Mutter, M. *Tetrahedron Lett.* **1992**, *33*, 1589-1592.
- 141) Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. *J. Am. Chem. Soc.* **1996**, *118*, 9218-9227.
- 142) Wöhr, T.; Mutter, M. *Tetrahedron Lett.* **1995**, *36*, 3847-3848.
- 143) Dumy, P.; Keller, M.; Ryan, D. E.; Rohwedder, B.; Wöhr, T.; Mutter, M. *J. Am. Chem. Soc.* **1997**, *119*, 918-925.
- 144) Wöhr, T.; Wahl, F.; Nefzi, A.; Rohwedder, B.; Sato, T.; Sun, X.; Mutter, M. *J. Am. Chem. Soc.* **1996**, *118*, 9218-9227.
- 145) Wittelsberger, A.; Patiny, L.; Slaninova, J.; Barberis, C.; Mutter, M. *J. Med. Chem.* **2005**, *48*, 6553-6562.
- 146) Tong, Y.; Olczak, J.; Zabrocki, J.; Gershengorn, M. C.; Marshall, G. R. *Tetrahedron* **2000**, *56*, 9791-9800.
- 147) An, S. S. A.; Lester, C. C.; Peng, J.-L.; Li, Y.-J., Rothwarf, D. M.; Welker, E.; Thannhauser, T. W.; Zhang, L. S.; Tam, J. P.; Scheraga, H. A. *J. Am. Chem. Soc.* **1999**, *121*, 11558-11566.
- 148) Kivirikko, K. I.; Myllyharju, J. *Matrix Biology* **1998**, *16*, 357-368.

- 149) Kivirikko, K. I.; Myllylfi, R.; Pihlajaniemi, T. *Post-Translational Modifications of Proteins*; (Eds. Harding, J. J.; Crabbe, M. J. C.) CRC Press, Boca Raton, 1992; pp. 1-51.
- 150) Kivirikko, K.I.; Myllylfi, R.; Pihlajaniemi, T. *FASEB J.* **1989**, *3*, 1609-1617.
- 151) Myllyharju, J.; Kivirikko, K.I. *EMBO* **1997**, *16*, 1173-1180.
- 152) Hanauske-Abel, H. M.; Gtinzler, V. *Theor. Biol.* **1982**, *94*, 421-455.
- 153) Gorres, K. L.; Edupuganti, R.; Krow, G. R.; Raines, R. T. *Biochemistry* **2008**, *47*, 9447-9455.
- 154) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777-778.
- 155) Eberhardt, E. S.; Panasik, N. Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266.
- 156) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497-2505.
- 157) Cadamuro, S. A.; Reichold, R.; Kusebauch, U.; Musiol, H.-J.; Renner, C.; Tavan, P.; Moroder, L. *Angew. Chem. Int. Ed.* **2008**, *47*, 2143-2146.
- 158) Nimni, M. E. *Collagen*; CRC Press: Boca Raton, FL, 1988.
- 159) Brodsky, B.; Shah N. K. *FASEB J.* **1995**, *9*, 1537-1546.
- 160) Myllyharju, J.; Kivirikko, K. I. *Ann. Med.* **2001**, *33*, 7-21.

- 161) Bateman, J. F.; Lamande, S. R.; Ramshaw, J. A. M. *Extracellular Matrix*; Vol. 2, (Ed. Comper, W. D.) Harwood Academic Publishers, Amsterdam, 1996; p. 27.
- 162) Kadler, K. *Protein Profile* **1994**, *1*, 519-638.
- 163) Graham, H. K.; Holmes, D. F.; Watson, R. B.; Kadler, K. E. *J. Mol. Biol.* **2000**, *295*, 891-902.
- 164) Ramachandran, G. N.; Kartha, G. *Nature* **1954**, *174*, 269-270.
- 165) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593-595.
- 166) Cowan, P. M.; McGavin, S.; North, A. C. *Nature* **1955**, *176*, 1062-1064.
- 167) Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915-916.
- 168) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483-506.
- 169) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75-81.
- 170) Kadler, K. E.; Holmes, D. F.; Graham, H.; Starborg, T. *Matix Biology* **2000**, *19*, 359-365
- 171) Radmer, R. J.; Klein, T. E. *Biochemistry* **2004**, *43*, 5314-5323.
- 172) Courtenay, J. S.; Dallman, M. J.; Dayan, A. D.; Martin, A.; Mosedale, B. *Nature* **1980**, *283*, 666-668.
- 173) Bailey, A. J.; Sims, T. J.; Avery, N. C.; Halligan, E. P. *Biochem. J.* **1995**, *305*, 385-390.

- 174) Fu, M. X.; Wells-Knecht, K. J.; Blackledge, J. A.; Lyons, T. J.; Thorpe, S. R.; Baynes, J. W. *Diabetes* **1994**, *43*, 676-683.
- 175) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 272-281.
- 176) Berg, R. A.; Prockop, D. J. *Biochem. Biophys. Res. Commun.* **1973**, *52*, 115-120.
- 177) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta.* **1973**, *303*, 198-202.
- 178) Kobayashi, Y.; Sakai, R.; Kakiuchi, K.; Isemura, T. *Biopolymers* **1970**, *9*, 415-425.
- 179) Burjanadze, T. V. *Biopolymers* **1992**, *32*, 941-949.
- 180) Privalov, P. L. *Adv. Protein Chem.* **1982**, *35*, 1-104.
- 181) Inouye, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1976**, *420*, 133-141.
- 182) Inouye, K.; Kobayashi, Y.; Kyogoku, Y.; Kishida, Y.; Sakakibara, S.; Prockop, D. J. *Arch. Biochem. Biophys.* **1982**, *219*, 198-203.
- 183) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960-14967.
- 184) Mohs, A.; Silva, T.; Yoshida, T.; Amin, R.; Lukomski, S.; Inouye, M.; Brodsky, B. *J. Biol. Chem.* **2007**, *282*, 29757-29765.
- 185) Suzuki, E.; Fraser, R. D. B.; MacRae, T. P. *J. Biol. Macromol.* **1980**, *2*, 54-56.

186) Jackel, C.; Kokschi, B. *Eur. J. Org. Chem.* **2005**, 4483-4503.

187) Page: 108

Howard, J. A. K.; Hoy, V. J.; O'Hagan, D.; Smith, G. T. *Tetrahedron* **1996**, 52, 12613-12622.

188) Dunitz, J. D.; Taylor, R. *Chem. Eur. J.* **1997**, 3, 89-98.

189) O'Hagan, D.; Bilton, C.; Howard, J. A. K.; Knight, L.; Tozer, D. J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 605-607.

190) Improta, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, 123, 12568-12577.

191) Fasman, G. D. *Practical Handbook of Biochemistry and Molecular Biology*; CRC Press: Boca Raton, FL, 1989.

192) Hall, H. K. *J. Am. Chem. Soc.* **1957**, 79, 5441-5444.

193) Krimm, S.; Bandekar, J. *Advances in Protein Chemistry*; Academic Press: New York, 1986; Vol. 38, pp 183-364.

194) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Molec. Biol.* **1995**, 30, 95-120.

195) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem. Int. Ed.* **2001**, 40, 923-925.

196) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, 392, 666-667.

197) Bürgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am. Chem. Soc.* **1973**, 95, 5065-5067.

- 198) Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A.; Raines, R. T.; Krow, G. R.; *J. Org. Chem.* **2004**, *69*, 8565-8573.
- 199) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *58*, 459-464.
- 200) Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2005**, *127*, 15923-15932.
- 201) Shoulders, M. D.; Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2006**, *128*, 8112-8113.
- 202) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079-2080.
- 203) Marin, J.; Blaton, M.-A.; Briand, J.-P.; Chiocchia, G.; Fournier, C.; Guichard, G. *ChemBioChem* **2005**, *6*, 1796-1804.
- 204) Marin, J.; Didierjean, C.; Aubry, A.; Briand, J.-P.; Guichard, G. *J. Org. Chem.* **2002**, *67*, 8440-8449.
- 205) Butler, W. T.; Cunningham, L. W. *J. Biol. Chem.* **1966**, *241*, 3882-3888;
- 206) Spiro, R. G. *J. Biol. Chem.* **1969**, *244*, 602-612.
- 207) Miller, E. J. *Biochemistry* **1971**, *10*, 1652-1659.
- 208) Kivirikko, K. I.; Myllylä, R. *Int. Rev. Connect. Tissue Res.* **1979**, *8*, 23-72.
- 209) Notbohm, H.; Nokelainen, M.; Myllyharju, J.; Fietzek, P. P.; Müller, P. K.; Kivirikko, K. I. *J. Biol. Chem.* **1999**, *274*, 8988-8992.

- 210) Brinckmann, J.; Notbohm, H.; Tronnier, M.; Acil, Y.; Fietzek, P. P.; Schmeller, W.; Müller, P. K.; Bätge, B. *J. Invest. Dermatol.* **1999**, *113*, 617-621.
- 211) Norman, K. R.; Moerman, D. G. *Dev. Biol.* **2000**, *227*, 690-705.
- 212) Ruotsalainen, H.; Sipilä, L.; Vapola, M.; Sormunen, R.; Salo, A. M.; Uitto, L.; Mercer, D. K.; Robins, S. P.; Risteli, M.; Attila, A.; Fässler, R.; Myllylä, R. *J. Cell. Sci.* **2006**, *119*, 625–635.
- 213) Sipilä, L.; Ruotsalainen, H.; Sormunen, R.; Baker, N. L.; Lamandé, S. R.; Vapola, M.; Wang, C.; Sado, Y.; Aszodi, A.; Myllylä, R. *J. Biol. Chem.* **2007**, *282*, 33381-33388.
- 214) Myllylä, R.; Wang, C.; Heikkinen, J.; Juffer, A.; Lampela, O.; Risteli, M.; Ruotsalainen, H.; Salo, A.; Sipilä, L. *J. Cell. Physiol.* **2007**, *212*, 323-329.
- 215) Myllylä, R.; Wang, C.; Heikkinen, J.; Juffer, A.; Lampela, O.; Risteli, M.; Ruotsalainen, H.; Salo, A.; Sipilä, L. *J. Cell. Physiol.* **2007**, *212*, 323-329.
- 216) Salo, A. M.; Cox, H.; Farndon, P.; Moss, C.; Grindulis, H.; Risteli, M.; Robins, S. P.; Myllylä, R. *Am. J. Hum. Genet.* **2008**, *83*, 495-503.
- 217) Wang, C.; Kovanen, V.; Raudasoja, P.; Eskelinen, S.; Pospiech, H.; Myllylä, R. *J. Cell. Mol. Med.* **2009**, *13*, 508-521.
- 218) Lauer-Fields, J. L.; Malkar, N. B.; Richet, G.; Drauz, K.; Fields, G. B. *J. Biol. Chem.* **2003**, *278*, 14321-14330.
- 219) van den Steen, P. E.; Proost, P.; Brand, D. D.; Kang, A. H.; Van Damme, J.; Opdenakker, G. *Biochemistry* **2004**, *43*, 10809-10816.

- 220) Broddefalk, J.; Bäcklund, J.; Almqvist, F.; Johansson, M.; Holmdahl, R.; Kihlberg, J. *J. Am. Chem. Soc.* **1998**, *120*, 7676-7683.
- 221) Bäcklund, J.; Carlsen, S.; Höger, T.; Holm, B.; Fugger, L.; Kihlberg, J.; Burkhardt, H.; Holmdahl, R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9960-9965.
- 222) Tanaka, S.; Avigad, G.; Eikenberry, E. F.; Brodsky, B. *J. Biol. Chem.* **1988**, *263*, 17650-17657.
- 223) Grandhee, S. K.; Monnier, V. M. *Biol. Chem.* **1991**, *266*, 11649-11655.
- 224) Bailey, A. J.; Sims, T. J.; Avery, N. C.; Miles, C. A. *Biochem. J.* **1993**, *296*, 489-496.
- 224) Sady, C.; Khosrof, S.; Nagaraj, R. *Biochem. Biophys. Res. Commun.* **1995**, *214*, 793-797.
- 225) Paul, R. G.; Bailey, A. J. *Int. J. Biochem. Cell Biol.* **1996**, *28*, 1297-1310.
- 226) Sell, D. R.; Kleinman, N. R.; Monnier, V. M. *FASEB J.* **2000**, *14*, 145-156.
- 227) Andersson, I. E.; Dzhambazov, B.; Holmdahl, R.; Linusson, A.; Kihlberg, J. *J. Med. Chem.* **2007**, *50*, 5627-5643.
- 228) Gaill, F., Mann, K., Wiedemann, H., Engel, J. and Timpl, R. *J. Mol. Biol.* **1995**, *246*, 284-294.
- 229) Mann, K., Mechling, D.E., Bachinger, H.P., Eckerskorn, C., Gaill, F. and Timpl, R. *J. Mol. Biol.* **1996**, *261*, 255-266.

- 230) Bann, J. G.; Peyton, D. H.; Bächinger, H. P. *FEBS Lett.* **2000**, *473*, 237-240.
- 231) Bann, J. G.; Bächinger, H. P.; Peyton, D. H. *Biochemistry* **2003**, *42*, 4042-4048.
- 232) Lamport, D. T. A. *Nature* **1967**, *216*, 1322-1324.
- 233) Lamport, D. T. A.; Miller, D. H.; *Plant Physiol.* **1971**, *48*, 454-456.
- 234) Lamport, D. T. A. *Recent Adv. Phytochem.* **1977**, *11*, 79-115.
- 235) Shpak, E.; Barbar, E.; Leykam, J. F.; Kieliszewski, M. J. *J. Biol. Chem.* **2001**, *276*, 11272-11278.
- 236) Homer, R. B.; Roberts, K. *Planta* **1979**, *146*, 217-222.
- 237) Lamport, D. T. A. *Structure and Function of Plant Glycoproteins in The Biochemistry of Plants*; Vol.3, Academic Press, 1980; pp501-540.
- 238) a) Esquerré-Tugayé, M. T.; Mazau, D. *J. Exp. Biol.* **1974**, *25*, 509-513; b) Esquerré-Tugayé, M. T.; Lamport, D. T. A. *Plant Physiol.* **1979**, *64*, 314-319.
- 239) Raggi, V. *Plant Pathology* **2000**, *49*, 179-186.
- 240) Ferris, P. J.; Woessner, J. P.; Waffenschmidt, S.; Kilz, S.; Drees, J.; Goodenough, U. *W. Biochemistry* **2001**, *40*, 2978-2987.
- 241) van Holst, G.-J.; Varner, J. E. *Plant Physiol.* **1984**, *74*, 247-251.
- 242) Traub, W.; Shmueli, U. *Nature* **1963**, *198*, 1165-1167.
- 243) Cowan, P. M.; McGavin, S. *Nature* **1955**, *176*, 501-503.

- 244) Torchia, D. A., Bovey, F. A. *Macromolecules* **1971**, *4*, 246-250.
- 245) Gornick, F.; Mandelkern, L.; Diorio, A. F.; Roberts, D. E. *J. Am. Chem. Soc.* **1964**, *86*, 2549-2555.
- 246) Engel, J. *Biopolymers* **1966**, *4*, 945-957.
- 247) Adzhubei, A. A.; Sternberg, M. J. E. *J. Mol. Biol.* **1993**, *229*, 472-493.
- 248) Siligardi, G.; Drake, A. F. *Biopolymers* **1995**, *37*, 281-292.
- 249) Kleywegt, G. T.; Jones, T. A. *Structure* **1996**, *4*, 1395-1400.
- 250) Bochicchio, B.; Tamburro, A. M. *Chirality* **2002**, *14*, 782-792.
- 251) Berisio, R.; Loguercio, S.; De Simone, A.; Zagari, A.; Vitagliano, L. *Prot. Pept. Lett.* **2006**, *13*, 847-854.
- 252) Tiffany, M. L.; Krimm, S. *Biopolymers* **1968**, *6*, 1379-1382.
- 253) Woody, R. W. *Adv. Biophys. Chem.* **1992**, *2*, 37-39.
- 254) Stapley, B. J.; Creamer, T. P. *Prot. Sci.* **1999**, *8*, 587-595.
- 255) Rucker, A. L.; Creamer, T. P. *Protein Sci.* **2002**, *11*, 980-985.
- 256) Shi, Z.; Woody, R. W.; Kallenbach, N. R. *Adv. Protein Chem.* **2002**, *62*, 163-240.
- 257) Hamburger, J. B.; Ferreon, J. C.; Whitten, S. T.; Hilser, V. J. *Biochemistry* **2004**, *43*, 9790-9799.

- 258) Ramakrishnan, V.; Ranbhor, R.; Durani, S. *J. Am. Chem. Soc.* **2004**, *126*, 16332-16333.
- 259) Tran, H. T.; Wang, X.; Pappu, R. V. *Biochemistry* **2005**, *44*, 11369-11380.
- 260) Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. *Biochemistry* **2005**, *44*, 6269-6275.
- 261) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877-1897.
- 262) Rath, A.; Davidson, A. R.; Deber, C. M. *Biopolymers* **2005**, *80*, 179-185.
- 263) Zarrinpar, A.; Bhattacharyya, R. P.; Lim, W. A. *Sci. STKE* **2003**, *179*, re8., 1-10.
- 264) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U. S. A.* **1967**, *58*, 719-726.
- 265) Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B. *J. Am. Chem. Soc.* **2002**, *124*, 13067-13071.
- 266) Arora, P. S.; Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743-747.
- 267) Sadler, K.; Eom, K. D.; Yang, J.-L.; Dimitrova, Y.; Tam, J. P. *Biochemistry* **2002**, *41*, 14150-14157.
- 268) Fernández-Carneado, J.; Kogan, M. J.; Pujals, S.; Giralt, E. *Biopolymers* **2004**, *76*, 196-203.
- 269) Fernández-Carneado, J.; Kogan, M. J.; Castel, S.; Giralt, E. *Angew. Chem. Int. Ed.* **2004**, *43*, 1811-1814.

- 270) Lam, S. L.; Hsu, V. L. *Biopolymers* **2003**, *69*, 270-281.
- 271) Ronish, E. W.; Krimm, S. *Biopolymers* **1974**, *13*, 1635-1651.
- 272) Creamer, T. P.; Campbell, M. N. *Adv. Protein Chem.* **2002**, *62*, 263-282.
- 273) Chellgren, B. W.; Creamer, T.P. *Biochemistry* **2004**, *43*, 5864-5869.
- 274) Kelly, M. A.; Chellgren, B. W.; Rucker, A. L.; Troutman, J. M.; Fried, M. G.; Miller, A.-F.; Creamer, T. P. *Biochemistry* **2001**, *40*, 14376-14383.
- 275) Rucker, A. L.; Pager, C. T.; Campbell, M. N.; Qualls, J. E.; Creamer, T. P. *Proteins* **2003**, *53*, 68-75.
- 276) Vila, J. A.; Baldoni, H. A.; Ripoll, D. R.; Ghosh, A.; Scheraga, H. A. *Biophys. J.* **2004**, *86*, 731-742.
- 277) Jha, A. K.; Colubri, A.; Zaman, M. H.; Koide, S.; Sosnick, T. R.; Freed, K. F. *Biochemistry* **2005**, *44*, 9691-9702.
- 278) Schumacher, M.; Mizuno, K.; Bächinger, H.P. *J. Biol. Chem.* **2005**, *280*, 20397-20403.
- 279) Page: 116
Chen, K.; Liu, Z.; Kallenbach, N. R.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.*, **2004**, *101*, 15352-15357.
- 280) Chen, K.; Liu, Z.; Zhou, C.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2005**, *127*, 10146-10147.

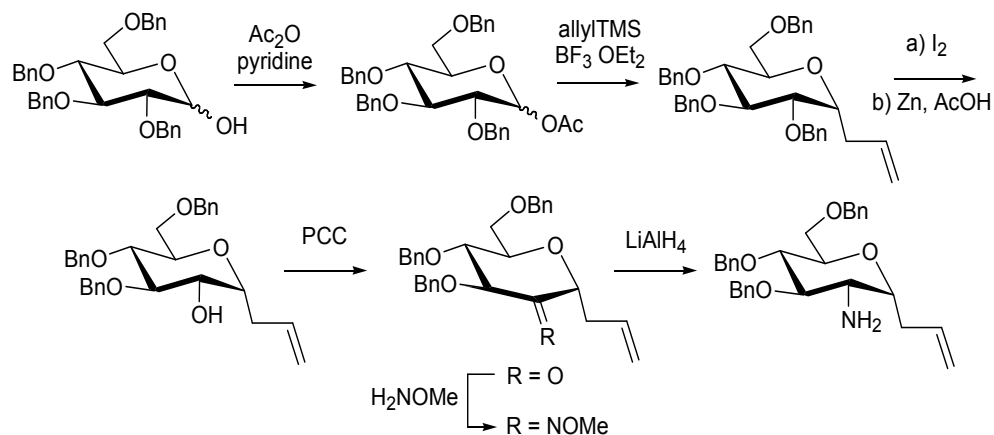
- 281) Pettitt, B. M.; Karplus, M. *Chem. Phys. Lett.* **1985**, *121*, 194-201.
- 282) Han, W.-G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587-2602.
- 283) Mezei, M.; Fleming, P. J.; Srinivasan, R.; Rose, G. D. *Proteins* **2004**, *55*, 502-507.
- 284) Helbecque, N.; Loucheux-Lefebvre, M. H. *Int. J. Pept. Prot. Res.* **1982**, *19*, 94-101.
- 285) Zhang, R.; Nickl, C. K.; Mamai, A.; Flemer, S.; Natarajan, A.; Dostmann, W. R.; Madalengoitia, J. S. *J. Pept. Res.* **2005**, *66*, 151-159.
- 286) Mamai, A.; Hughes, N. E.; Wurthmann, A.; Madalengoitia, J. S. *J. Org. Chem.* **2001**, *66*, 6483-6486.
- 287) Kümin, M.; Sonntag, L.-S.; Wennemers, H. *J. Am. Chem. Soc.* **2007**, *129*, 466-467.
- 288) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004-2021.
- 289) Sonntag, L.-S.; Schweizer, S.; Ochsenfeld, C. Wennemers, H. *J. Am. Chem. Soc.* **2006**, *128*, 14697-14703.
- 290) Horng, J.-C.; Raines, R. T. *Prot. Sci.* **2006**, *15*, 74-83.
- 291) Chen, K.; Liu, Z.; Kallenbach, N. R.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.*, **2004**, *101*, 15352-15357.
- 292) Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2003**, *125*, 8178-8185.

- 293) Greenfield, N. J.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108-4116.
- 294) Sreerama, N.; Woody, R. W. *Prot. Struct. Funct. Genet.* **1999**, *36*, 400-406.
- 295) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9190-9195.
- 296) Ding, L.; Chen, K.; Santini, P. A.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2003**, *125*, 8092-8093.
- 297) Avbelj, F.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.*, **2003**, *100*, 5742-5747.
- 298) Eker, F.; Griebenow, K.; Cao, X.; Nafle, L. A.; Schweitzer-Stenner, R. *Proc. Natl. Acad. Sci. U. S. A.*, **2004**, *101*, 10054-10059.

Chapter 2: Thesis Objectives

Towards an understanding of the effects of glycosylation on the properties of L-proline, this thesis has two main goals. First, design and synthesize a C-glycosyl proline hybrid and investigate its ability to control *N*-terminal amide isomerization through derivatization of the carbohydrate scaffold. Second, to study the effects of both α - and β -*O*-glycosylation on 4-hydroxy-L-proline (Hyp) in model compounds, as well as its influence on the conformational stability of appropriate model peptides.

Many approaches exist for the synthesis of *C*-glycosides (Section 1.4.1), however, we chose to use the approach of 2,3,4,6-tetra-*O*-benzyl- α -D-glucopyranose which can be converted in excellent yield and stereoselectivity to the α -allyl *C*-glycoside using allyltrimethylsilane (Scheme 2.1).¹



Scheme 2.1: Synthesis of 2-amino *C*-glycoside.^{1,2}

According to the procedure of Cipolla and coworkers, the benzyl protected *C*-glycoside can be selectively deprotected at the 2-position using I_2 followed by treatment with zinc under acidic conditions.² Regioselective installation of the 2-amino group can be accomplished by consecutive oxidative and reductive amination steps to give the

gluco-stereoisomer in good yield. This provides a suitable starting point for incorporation of the prolyl side chain into the carbohydrate frame.

Once the *C*-glycosyl proline hybrid has been developed, the aim will be to incorporate the building block into model peptides, and determine the effects of modification of the carbohydrate hydroxyl groups on the conformation and amide isomer equilibrium (Chapter 3).

There are also many approaches for the synthesis of *O*-glycosylated hydroxyproline building blocks; however, few methods give both α - and β -anomeric products as desired in this case.³ The anomeric configuration is known to have different effects on peptide characteristics (Section 1.2.4). The use of *O*-benzyl protected thioglycosides provides both α - and β -anomers in high yields, although separation of the anomers becomes an issue.

With the glycosylated hydroxyproline building block in hand, it will be incorporated into the conventional model amides: Ac- X_{aa} -NHMe or Ac- X_{aa} -OMe. Use of the former provides an amide at both *N*- and *C*-termini, but has been criticized for interference from a possible intramolecular hydrogen bond from the amide N-H group.⁴ The latter does not suffer from this problem, but provides less peptide character (the contributions of the NH group to hydrogen bonds and the amide bond characteristics).

The effects of *O*-glycosylation on the conformation, thermodynamics, and kinetics of amide isomerization for both *4R*- (Chapter 4) and *4S*-hydroxy-L-proline (hyp) (Chapter 5) will be investigated, with attempts to explain the basis of these effects. In accord with previous studies, measurement of these parameters will be carried out

primarily by NMR experiments: integration of amide isomer signals, van't Hoff plots, nOe experiments, and magnetization inversion transfer experiments.

While Hyp is found in important structural proteins in plants and animals, hyp is rarely found in nature,⁵ and has a different influence on the conformation of the pyrrolidine ring, as well as on amide isomerization (Section 1.6.4); it will provide an opportunity to study the effect of *O*-linked glycans projected from the opposite face of the prolyl side chain relative to Hyp on amide isomerization.

Since Hyp is found in the important structural proteins: collagen (Section 1.7.1) and hydroxyproline-rich glycoproteins (HRGPs) (Section 1.8.1), model peptides incorporating *O*-glycosylated hydroxyproline residues will be synthesized to probe the effects of Hyp *O*-glycosylation on their conformational stability.

Typically, model peptides of the form Ac-(Pro-Hyp-Gly-)₇₋₁₀-NH₂ have been used to probe the effects of Hyp substitution on collagen stability, since these peptides are known to form triple helices in solution. The effects of both α - and β -*O*-glycosylation of Hyp on the conformational stability of collagen model peptides will be investigated (Chapter 6); primarily using CD spectrometry, in accordance with conventional approaches.

Similarly, model peptides of the form Ac-(Pro)₆₋₉-NH₂ have been used to study the effects of proline modification on the polyproline II conformation, as found in HRGPs and molecular scaffolds. Since HRGPs are characterized by glycans β -*O*-linked to Hyp through a D-galactose residue, the effects of β -*O*-glycosylation of Hyp on the conformational stability of polyproline model peptides will be investigated (Chapter 7); also primarily using CD spectroscopy, in accordance with conventional approaches.

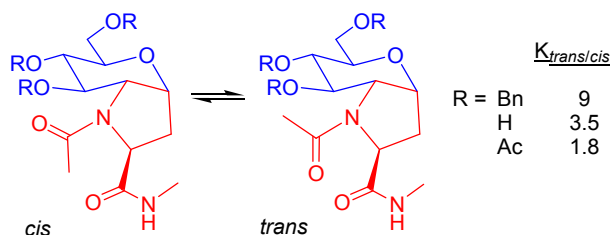
2.1 References

- 1) Brenna, E.; Fuganti, C.; Grasselli, P.; Serra, S.; Zambotti, S. *Chem. Eur. J.* **2002**, *8*, 1872-1878.
- 2) Cipolla, L.; Lay, L. Nicotra, F. *J. Org. Chem.* **1997**, *62*, 6678-6681.
- 3) Arsequell, G.; Sarries, N.; Valencia, G. *Tetrahedron Lett.* **1995**, *36*, 7323-7326.
- 4) Eberhardt, E. S.; Panasik, N. Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266.
- 5) Mauger, A. B. *J. Nat. Prod.* **1996**, *59*, 1205-1211.

Chapter 3: Tuning of the Prolyl *Trans/Cis* Amide Rotamer Population Using C-Glucosyl Proline Hybrids

Neil W. Owens,^{*,#,\$∞} Craig Braun,[#] and Frank Schweizer^{*,\$∞}

^{*}Designed research; [#]Performed research; ^{\$}Analyzed data; [∞]Wrote the paper



3.1 Abstract

We describe the synthesis of a fused bicyclic C-glucosyl proline hybrid (GlcProH) from commercially available 2,3,4,6-tetra-O-benzyl-D-glucopyranose. The GlcProH was incorporated into the model peptides Ac-GlcProH-NHMe and Ac-Gly-GlcProH-NHMe. Postsynthetic modifications can be introduced via derivatization of the carbohydrate scaffold. Conformational analysis of the GlcProH-modified model peptides shows that while the conformation of the GlcProH remains fixed, the prolyl N-terminal amide equilibrium ($K_{trans/cis}$) can be varied with different modifications of the carbohydrate scaffold. Simple N-acyl derivatives studied by NMR spectroscopy showed that in CD₃OD, there was an increase in the cis-amide content as the sugar substituents changed from benzyl (10%) to hydroxyl (22%) to acetate (36%). Similar effects were observed in DMSO-D₆. The exact nature of the influence is unclear, but most likely arises through intramolecular interactions between sugar groups and the peptidic amide backbone. Overall, our GlcProH demonstrates variation in ($K_{trans/cis}$) through tuning of the carbohydrate scaffold: a new concept in proline peptidomimetics.

3.2 Introduction

Proline is uniquely endowed with a side chain that is fused onto the peptide backbone. This trait restricts the rotation about its ϕ dihedral angle, thereby reducing the energy difference between the prolyl amide *cis*- and *trans*-isomers, making them nearly isoenergetic. Thus, while most peptide amide bonds exist almost exclusively in the *trans* form, proline has a much greater propensity to form *cis* amide bonds; this makes proline crucial for inducing a reversal in peptide backbone conformation.¹ Also, *cis-trans* isomerization of proline is of importance because it becomes the rate-determining step in the folding pathways of peptides and proteins.²

Variation of the *trans/cis* ratio ($K_{trans/cis}$) is of interest for understanding the behavior of peptides and proteins. Over the years a number of proline analogues have been developed to study the structural and biological properties of proline surrogates in peptides. This is done by controlling prolyl *N*-terminal amide isomerization, and thus reverse turn formation. Examples of amide geometry controlled through steric or stereoelectronic influences includes C^β -, C^γ -, C^δ -substituted prolines,³⁻⁶ including 4-fluoroproline,⁷ azaproline,⁸ and pseudoproline.⁹

While these analogues have proved useful for inducing specific constraints on prolyl *N*-terminal amide isomerization, no *one* analogue can lay claim to the ability to shift the prolyl amide equilibrium towards both the *cis* and *trans* isomers. In other words, different proline analogues are required to induce a desired bias in $K_{trans/cis}$. This is in part because none of the present building blocks have strategic functional groups positioned for further derivatization in order to alter the amide equilibrium. This led us towards the

development of a proline analogue building block in which prolyl *cis-trans* isomerization could be tuned through simple chemistry, even after incorporation into a peptide.

Our concept for a proline analogue was derived from glycosyl amino acids (GAAs), which are defined by an α -amino acid group ($-\text{CH}(\text{NH}_2)\text{CO}_2\text{H}$) either directly attached or carbon-linked to the anomeric carbon of a carbohydrate scaffold (furan- or pyran-).¹⁰ The relative rigidity of the furan or pyran ring combined with the polyfunctional nature of the carbohydrate scaffold, has inspired the design of unusual and conformationally constrained amino acids and novel peptidomimetics. While there are many examples of C-glycosyl glycine, alanine, serine, and asparagine,^{10c} few proline-based GAAs exist.¹¹

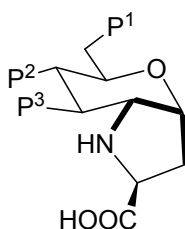


Figure 3.2.1: General structure of glucose proline hybrid (GlcProH). The carbohydrate scaffold is used as a template to constrain C_β , C_γ , C_δ , and N of the pyrrolidine ring (shown in bold) of L-proline or (2*S*,4*R*)-hydroxyproline; P^1 , P^2 , and P^3 are hydroxy-derived substituents used to manipulate the steric and electronic properties of the GlcProH.

We report here on the resulting novel fused bicyclic gluco-proline hybrid (GlcProH) GAA for use as a peptidomimetic to tune prolyl *N*-terminal amide isomerization. Our bicyclic GlcProH rigidly combines the molecular elements of carbohydrates (pyran-based polyol) with the unique features of proline (Figure 3.2.1).

The resulting GlcProH is a rigid, polyfunctional building block, which may find use as a proline mimetic, glycomimetic, or scaffold for combinatorial synthesis. Many conformationally constrained L-proline analogues have been developed^{12,13} but recently, fused bicyclic prolines have attracted interest due to their increased rigidity, which may permit better control of the *trans/cis* ratio.¹⁴⁻¹⁶ For instance, bicyclic proline analogues have been studied as angiotensin¹⁷ and thyroliberin¹⁸ analogues, which have served as building blocks for the synthesis of peptidomimetics.¹⁹ Because of its stable chair conformation, glucose provides an ideal scaffold to template proline since it freezes the orientation of four proline atoms (C_β, C_γ, C_δ, N). Furthermore, the sugar hydroxyl groups lend themselves to derivatization of the building block as potential sites for influencing the peptide backbone conformation. It has been shown that hydroxylated amino acids can induce novel secondary structures in small peptides. For instance, incorporation of unprotected sugar amino acids into small peptides (opioid peptides and gramicidin S), prohibited the formation of the targeted secondary structural motif.^{20,21} Appearing instead were unusual turn structures stabilized by intramolecular hydrogen bonds between sugar hydroxyl groups and the peptidic amide backbone.²²

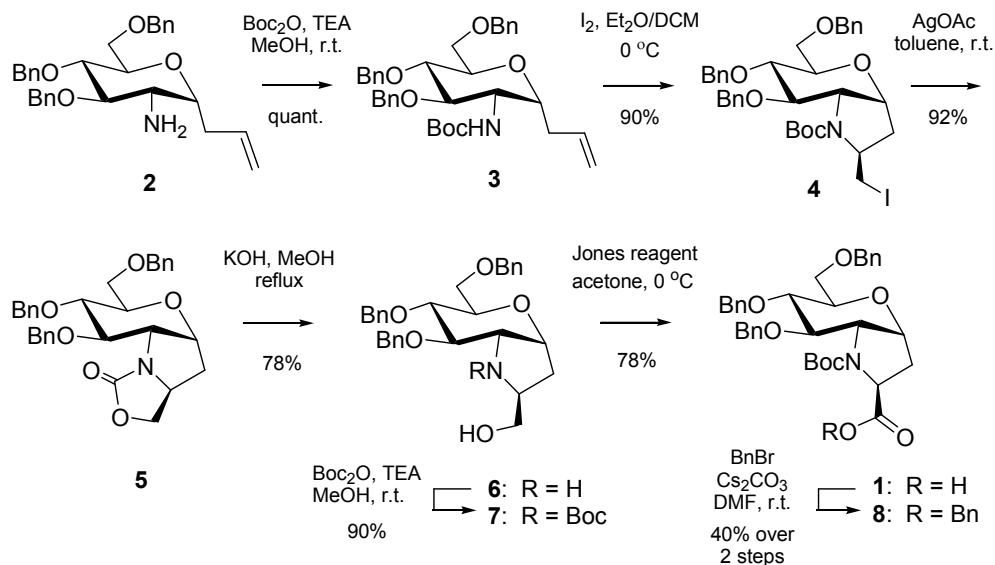
We envision that similar effects may also be observed with the GlcProH. For instance, we hope to influence peptide geometry via intramolecular polar interactions (H-bonding) between the sugar substituents and the peptide backbone. Furthermore, through derivatization of the polyhydroxylated carbohydrate scaffold we hope to demonstrate the capacity to tune prolyl *N*-terminal amide isomerization. To test the effects of chemically modifying the carbohydrate hydroxyl groups on prolyl *cis-trans* isomerization, we will study simple *N*-acyl and *N*-glycyl model compounds. Modifications of the glycine-

proline peptide motif are of particular interest since it is present in collagen,²³ as well as small peptides found to have neuroprotective properties²⁴ and the ability to inhibit HIV replication.²⁵ Furthermore, the Gly-Pro-Gly peptide sequence is known to induce reverse turns.²⁶

3.3 Results

3.3.1 Synthesis of *N*-Boc-GlcPro-Carboxylic Acid **1**

The synthesis of GlcProH **1** (Scheme 3.3.1.1) started from known amine **2**²⁷ which was synthesized in seven steps from commercially available 2,3,4,6 tetra-*O*-benzyl-D-glucopyranose in an overall yield of 40%. Protection of the amino function as *tert*-butyloxycarbamate (Boc) afforded **3** in quantitative yield. Installation of the pyrrolidine ring was achieved via amino-iodocyclization in 50% CH₂Cl₂/ether to afford the bicyclic iodo-derivative **4** in 90% yield as a single stereoisomer together with 5% unreacted **3**.²⁸ Attempts to substitute the iodo-function in **4** by hydroxide ion (KOH) or acetate failed, and produced complex mixtures containing tricyclic carbamate **5**. However, high yields (92%) of **5** could be obtained by exposure of **4** to silver acetate in toluene.²⁹ Hydrolysis of **5** using potassium hydroxide in methanol at elevated temperature provided the amino alcohol **6** in 78% isolated yield. Protection of the amino function in **6** was accomplished using di-*tert*-butyl dicarbonate to yield the Boc-protected proline analogue **7** in 90% yield. Subsequently, the alcohol **7** was subjected to Jones oxidation to afford protected GlcProH **1** in 78% yield. To assign the stereochemistry of **1**, esterification using benzyl bromide and cesium carbonate in DMF afforded the protected GlcProH **8** in 40% yield from **7**.



Scheme 3.3.1.1: Synthesis of *N*-Boc-GlcPro-Carboxylic Acid **1**

3.3.2 Assignment of Stereochemistry for *N*-Boc-GlcPro-Benzyl Ester **8**

The *S*-configuration at the 2-position of protected GlcProH **8** was identified on the basis of observed/unobserved nOe contacts shown in Figure 3.3.2.1. For instance, subtraction of the H-2 signal in **8** to a one-dimensional GOESY³⁰ experiment showed inter-proton effect to H-5 (0.8% nOe³¹ relative to the H-2 (dd) signal). By comparison, no interproton effects were observed between H-2 and H-9 or H-2 and H-8.

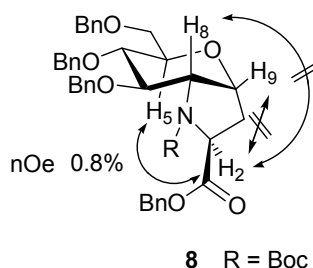
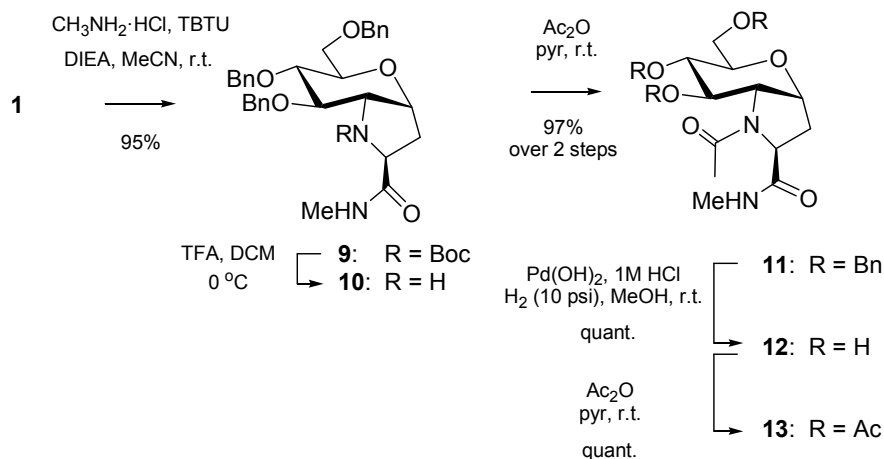


Figure 3.3.2.1: Configurationally relevant nOe interactions observed for **8**.

3.3.3 Synthesis of *N*-Acetyl-GlcPro-*N'*-Methylamides **11-13** and Determination of $K_{trans/cis}$

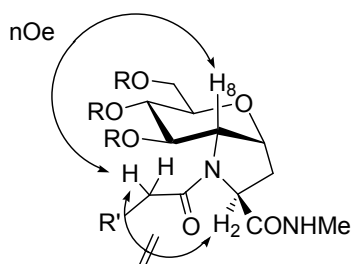
With GlcProH **1** in hand we then explored the use of the building block in peptide chemistry. Initially, we focused on analogues for comparison with Ac-Pro-NHMe **17** that has frequently served as a model to study the *cis-trans* isomerization of proline.^{4,32} The synthesis of GlcProH-modified amides **11-13** is outlined in Scheme 3.3.3.1. Acid **1** was directly coupled to methylamine using *O*-benzotriazol-yl-*N,N,N',N'*-tetramethyluronium tetrafluoroborate (TBTU) as coupling reagent in MeCN to produce amide **9** in 95% yield. Deprotection of the amine giving **10** followed by acylation produced diamide **11** in 97% yield over 2 steps. In addition, removal of the benzylether protecting groups by catalytic hydrogenolysis in methanol provided the polyhydroxylated diamide **12**, which was further acylated to afford diamide **13** in quantitative yield.



Scheme 3.3.3.1: Synthesis of *N*-Acetyl-GlcPro-*N'*-Methylamides **11-13**.

For each of **11-13**, the ratio of *trans/cis* isomers was calculated by integrating as many well-resolved peaks as possible for each isomer, and taking the average of over all

peaks for respective isomers.^{32a} The assignment of *N*-terminal amide geometry for both major and minor isomers of **11-16** was based on multiple GOESY experiments.³⁰ For example, selective inversion of the H-8 proton in **11** (CD₃OD) showed inter-proton effect to the *N*-terminal acetate singlet (5.2%) (Figure 3.3.3.1). By comparison, a very small (0.4%) inter-proton effect was observed between H-2 and the acetate methyl singlet, and weak or no interproton effects were observed between H-8 and H-2. As both H-8 and H-2 lie on opposite faces of the pyrrolidine ring, nOe contact from H-8 to the *N*-terminal acetate singlet can be assigned as the *trans* isomer. Analysis was carried out in different NMR solvents to understand how the model compounds behaved in different solvent environments. While it would be most desirable to study the GlcProHs in water, these results represent a proof of concept, and there is precedence for studying modifications of prolyl isomerization in non-aqueous environments.^{13,33}



11-13 R = Bn, H, Ac; R' = H
14-16 R = Bn, H, Ac; R' = CH₃CONH

Figure 3.3.3.1: Representative conformationally relevant nOe interactions observed for GlcProH derived model compounds. Inter-proton effects for the *trans* amide rotamer were observed between H-8 and the *N*-acyl methyl group (compounds **11-13**) or glycinyl protons (compounds **14-16**). By comparison, very small or no inter-proton effects were observed between H-2 and the *N*-acyl methyl group (compounds **11-13**) or glycinyl protons (compounds **14-16**).

We found a significant variation of the *cis* content as the sugar substituents were varied (Table 3.3.3.1). In CD₃OD, there was an increase in the *cis*-content as the sugar substituents changed from benzyl **11** (10%) to hydroxyl **12** (22%) to acetate **13** (36%). Analysis in DMSO-D₆ produced very similar results, with the *N*-terminal amide *cis* content varying from benzyl **11** (13%) to hydroxyl **12** (20%) to acetate **13** (33%).

Table 3.3.3.1: *Trans/cis* ratio^[a] ($K_{trans/cis}$) and % *cis* isomer of **11–13** in various solvents.

Compound	Solvent			
	CDCl ₃	D ₂ O	CD ₃ OD	DMSO-D ₆
11 (R = Bn)	>30 (<3%)	n.s. ^[b]	9 (10%)	6.7 (13%)
12 (R = H)	n.s. ^[b]	5.7 (15%)	3.5 (22%)	4 (20%)
13 (R = Ac)	6.7 (13%)	n.s. ^[b]	1.8 (36%)	2 (33%)

^[a]Determined by 500 MHz NMR at 25 °C; ^[b]not soluble.

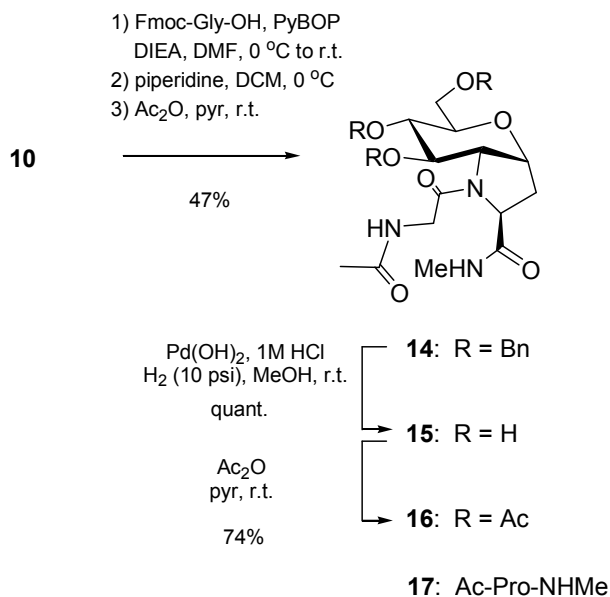
Overall, the $K_{trans/cis}$ for the unprotected GlcProH **12** (3.5 in CD₃OD) was very similar to the L-proline model compound **17** ($K_{trans/cis} = 4$),^{32b} leaving the per-*O*-benzylated **11** favoring the *trans* conformer ($K_{trans/cis} = 9$ in CD₃OD), and the per-*O*-acylated **13** favoring the *cis* conformer ($K_{trans/cis} = 1.8$ in CD₃OD). Also, CDCl₃ gave higher $K_{trans/cis}$ than the more polar solvents ($K_{trans/cis} = 30$ when R = OBn, $K_{trans/cis} = 6.7$ when R = OAc).

3.3.4 Synthesis of *N*-Acetyl-Glycyl-GlcPro-*N'*-Methylamides **14-16** and Determination of $K_{trans/cis}$

GlcProH **10** was coupled to Fmoc-Gly-OH using benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) as coupling reagent, followed

by deprotection of the amino function and acylation providing the triamide **14** in 47% yield over 3 steps (Scheme 3.3.4.1). Deblocking of the carbohydrate moiety afforded polyhydroxylated peptide **15**, which was further acylated to afford triamide **16**.

Again, the average of as many well-resolved peaks for the major and minor isomers was recorded to find the $K_{trans/cis}$ for respective compounds in each solvent. One-dimensional GOESY experiments helped to assign the *N*-terminal amide geometry. The major isomer was assigned as the *trans* isomer based on interproton effect observed between H-8 and the glycylic α -protons (Figure 3.3.3.1) (for example a 6.1% nOe was observed for **14** in CD₃OD). Furthermore, selective inversion of H-2 of the major isomer showed no interproton effect to either H _{α 1(Gly)} or H _{α 2(Gly)} (for example no such nOe was observed for **14** in CDCl₃). Additionally, selective inversion of H _{α 2(Gly)} of the minor isomer showed no interproton effect to H-8 of the minor isomer (for example no such nOe was observed for **16** in CD₃OD).



Scheme 3.3.4.1. Synthesis of *N*-Acetyl-Glycyl-GlcPro-*N'*-Methylamides **14-16**.

We found a very similar profile of $K_{trans/cis}$ for the Ac-Gly-GlcPro-NHMe model compounds **14-16** compared to the Ac-GlcPro-NHMe model compounds **11-13** (Table 3.3.4.1). In CD₃OD, there was an increase in the *cis*-content as the sugar substituent changed from benzyl **14** (<3%), to hydroxyl **15** (15%), to acetate **16** (25%). These results confirm that the substituents on the sugar are influencing $K_{trans/cis}$.

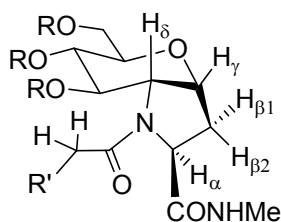
Table 3.3.4.1: *Trans/cis* ratio^[a] ($K_{trans/cis}$) and % *cis* isomer of **14-16** in various solvents.

Compound	Solvent		
	CDCl ₃	D ₂ O	CD ₃ OD
14 (R = Bn)	19 (5%)	n.s. ^[b]	>30 (<3%)
15 (R = H)	n.s. ^[b]	9 (10%)	5.7 (15%)
16 (R = Ac)	4 (20%)	n.s. ^[b]	3 (25%)

^[a]Determined by 500 MHz NMR at 25 °C; ^[b]not soluble.

Analysis of the coupling constants of the pyranose ring showed it exists in a chair conformation. For example, for **12** in CD₃OD, $J_{5,6}$ was 9.4Hz, while $J_{6,7}$ was 9.6Hz and $J_{7,8}$ was 9.0Hz (see Figure 3.3.2.1 for numbering). The average values of each coupling constant in comparison to the literature values match best to a C γ -endo conformation for the pyrrolidine ring (Table 3.3.4.2).

Table 3.3.4.2: Comparison of Average Coupling Constants (Hz) for **11-16** Major and Minor isomers^[a] with C γ -endo and C γ -exo puckers of 4-Fluoroproline and L-Proline.



11-13 R = Bn, H, Ac; R' = H

14-16 R = Bn, H, Ac; R' = CH₃CONH

Compound	$^3J_{\alpha,\beta1}$	$^3J_{\alpha,\beta2}$	$^3J_{\gamma,\beta1}$	$^3J_{\gamma,\beta2}$	$^3J_{\delta,\gamma1}$
11-16 Major isomers	9.9 ± 0.3	1.0 ± 0.1	11.8 ± 0.3	7.4 ± 0.2	7.2 ± 0.1
11-16 Minor isomers	9.3 ± 0.2	1.7 ± 0.3	10.9 ± 0.4	6.8 ± 0.2	7.1 ± 0.3
4 <i>S</i> -Fluoroproline ³⁴ (C γ -endo)	10	3	4	1	4
4 <i>R</i> -Fluoroproline ³⁴ (C γ -exo)	8	10	1	4	3
L-Proline ³⁵ (C γ -endo)	6-10	2-3	5-9	8-12	6-10
L-Proline ³⁵ (C γ -exo)	7-10	7-11	5-9	2-3	5-9

^[a] ± standard error.

Also, coupling constants indicate that the prolyl 4-position hydroxyl group (pyrano endocyclic oxygen) is oriented in an equatorial position relative to the pyrrolidine ring, which is not its preferred axial orientation.³⁶ Perhaps most importantly, all coupling constant values changed very little as the sugar substituents were varied, even as the solvent was varied, and between the major and minor isomers in each case. Together, these results indicate that the rigid pyranose ring is restricting the conformational freedom of the pyrrolidine ring.

3.4 Discussion

After developing a synthetic route to a novel sugar-proline analogue, we found through NMR experiments that *N*-terminal amide isomerization is dependent on the modification of the gluco-scaffold, but cannot be correlated to a significant variation in conformation.

There could be a steric explanation for the trend in $K_{trans/cis}$, as work by Lubell *et al.* has shown that the $K_{trans/cis}$ for prolyl *N*-terminal amide bonds can be influenced through steric effects.^{3,37} Alkyl-substituted proline derivatives such as 5-*tert*-butyl proline, force a very high *cis* *N*-terminal amide content (80%+) through steric repulsion of the *N*-terminal amino acid side chain and the 5-*tert*-butyl substituent on proline. Here, we chose model compounds that provide a minimal steric contribution to influence the $K_{trans/cis}$. Thus, the *N*-terminal acetate or glycyl groups should provide little to no steric bias between the *cis* and *trans* isomers. If there were an explanation for the trend in $K_{trans/cis}$ based on sterics, then the deprotected sugar-proline hybrid ($K_{trans/cis}$ of 3.5 in CD₃OD for **12**) should provide the lowest steric influence, while the benzyl ($K_{trans/cis}$ of 9.0 in CD₃OD for **11**) and acetate ($K_{trans/cis}$ of 1.8 in CD₃OD for **13**) substituents would provide greater steric bulkiness. However, the trend in $K_{trans/cis}$ does not follow the trend in increase in steric bulk of the sugar substituents (from 3.5 to 1.8 to 9.0 for **12**, **13** and **11** respectively). This led us to consider other factors.

A large amount of interest lies in studying the conformation of proline,^{4,8,23,32,38} and modifications which alter the preferred conformation of the pyrrolidine ring.^{3-9,36,37,39} We analyzed the coupling constants of the sugar and pyrrolidine rings in an attempt to consider the change in $K_{trans/cis}$ as a function of the change in the pucker of the pyrrolidine

ring. Raines and Krow *et al.* proved,^{7,36} through work with 4-fluoroproline model compounds, that as the stereochemistry of the 4-position EWG is varied, so is the pucker of the pyrrolidine ring (in an attempt to maximize the gauche effect). Consequently, as the pucker is changed the degree of $n \rightarrow \pi^*$ donation from the *N*-terminal C=O to the *C*-terminal C=O is also changed, thus affecting $K_{trans/cis}$. In our case, one could imagine that as the substituents on the sugar are varied, the pucker of the pyrrolidine ring also changes, influencing $n \rightarrow \pi^*$ donation and $K_{trans/cis}$. Through NMR analysis of **11-16** we found several results; as the sugar substituents are varied, and even as the solvent is varied, the coupling constants are held within a very narrow range (Table 3.3.4.2). These results indicate that the pucker of the pyrrolidine ring does not vary as the substituents on the sugar scaffold are varied, and the change in $K_{trans/cis}$ cannot be explained in terms of a change in the pyrrolidine pucker.

Most likely, simple intramolecular interactions between the sugar substituents and the peptide backbone are responsible for the observed variation in the prolyl *N*-terminal amide isomerization. For compounds **12** and **15** where R = H, there is the potential for a hydrogen bond between the 7-position hydroxyl group and the prolyl *N*-terminal amide carbonyl (a in Figure 3.5.1). This would stabilize the *cis* *N*-terminal amide isomer to some extent, and would not be possible for the *trans* isomer. For compounds **13** and **16** where R = Ac, there is the potential for $n \rightarrow \pi^*$ donation from the prolyl *N*-terminal amide carbonyl to the sugar 7-position acetate ester group (b in Figure 3.5.1). Molecular modeling analysis would be useful in determining the distances between the pertinent functional groups.

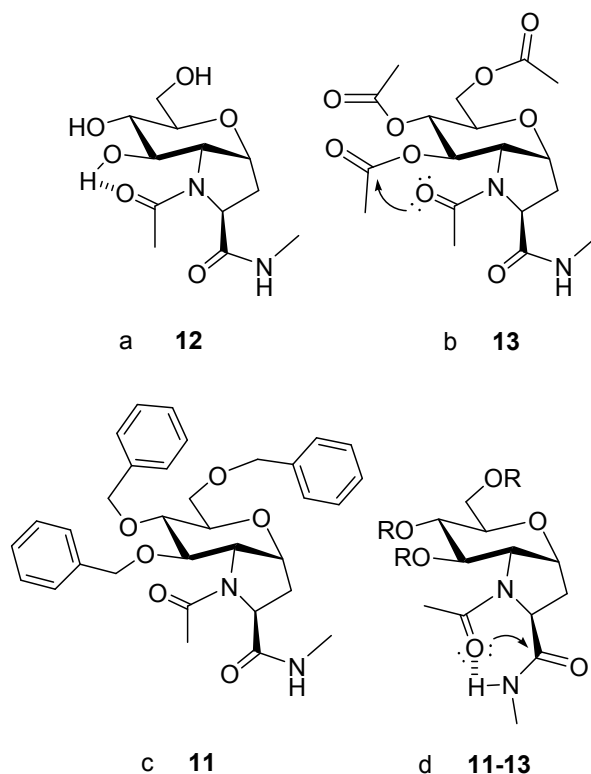


Figure 3.4.1: Proposed intramolecular interactions for **11**, **12** and **13**: (a) *Cis* amide rotamer isomer can be stabilized through an intramolecular hydrogen bond in **12**; (b) *Cis* amide rotamer can also be stabilized through $n \rightarrow \pi^*$ donation from the prolyl *N*-terminal amide to the acetate ester group at the 7-position of GlcProH in **13**; (c) The interactions that stabilize the *cis* amide rotamer in **12** and **13** are not possible for **11**; (d) *Trans* amide rotamer for **11-13** can be stabilized through $n \rightarrow \pi^*$ donation or a hydrogen bond to the *C*-terminal amide; Analogous interactions can be envisaged for **14-16**.

This would stabilize the *cis* *N*-terminal amide isomer. The relative decrease in $K_{trans/cis}$ for **13** and **16** ($K_{trans/cis}$ of 1.8 and 3.0 respectively in CD_3OD) compared to **12** and **15** ($K_{trans/cis}$ of 3.5 and 5.7 respectively in CD_3OD) may reflect a greater energetic stabilization of the *cis* amide ground state from $n \rightarrow \pi^*$ donation relative to a hydrogen bond. Finally, for compounds **11** and **14** where $R = Bn$ (c in Figure 3.5.1), the very high

trans content is of interest (In CD₃OD, $K_{trans/cis}$ is 9 for compound **11**, and $K_{trans/cis}$ is 30 for compound **14**). The interactions that stabilize the *cis* amide geometry for **12** and **13** would not be possible with an ether-linked protecting group. Most likely there is a steric influence which orients the *N*-terminal amide such that it can maximize the $n \rightarrow \pi^*$ donation to the *C*-terminal amide, thereby stabilizing the *trans* isomer. This $n \rightarrow \pi^*$ donation is also probable for the *trans* isomers of **12** and **13** (d in Figure 3.4.1). There is also the potential for an intramolecular hydrogen bond from the *N*-terminal carbonyl to the *C*-terminal amide. Therefore, the overall $K_{trans/cis}$ must represent the equilibrium established between these competing interactions.

3.5 Conclusions

In order to extend the molecular repertoire of proline analogues, we became interested in the design and synthesis of a bicyclic sugar-proline hybrid. The resulting GlcProH is a rigid, polyfunctional building block, which may find use as a proline mimetic, glycomimetic, or scaffold for combinatorial synthesis. The GlcProH served as a building block in peptide synthesis and was incorporated into model peptides. Simple modifications of the sugar scaffold permitted tuning of the *trans/cis* amide rotamer population in various solvents. The *cis* rotamer population increased as the substituents on the sugar scaffold were varied (*O*-benzyl < OH (unprotected) < *O*-acetyl) in CD₃OD and DMSO-D₆. Until now, the ability to vary the *N*-terminal prolyl amide isomer ratio required the synthesis of proline surrogates by independent synthetic routes. However, our results demonstrate that one-step modifications of the hydroxyl groups on the

carbohydrate scaffold can be used to tune the conformational properties of GlcProH-containing model peptides.

3.6 Experimental Section

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-(*tert*-Butyloxycarbonyl)-6,7-di-*O*-benzyl-octahydro-5-

(benzyl hydroxymethyl)-pyrano[3,2-*b*]pyrrole-2-carboxylic acid (1**):** Compound **7**

(0.125 g, 0.21 mmol) was dissolved in acetone (6 mL). The solution was cooled to 0°C. Fresh Jones' reagent (0.75 mL, 6.3 mmol) was prepared, and was added dropwise. The reaction mixture was stirred for 30 minutes, before adding water (5 mL), then aq. saturated sodium bicarbonate (5 mL). The acetone was removed under reduced pressure. The product was extracted into CH₂Cl₂ (3 x 15 mL) then dried (Na₂SO₄), concentrated under reduced pressure and was normally used directly in the next reaction. Purification by flash chromatography using 3:1 ethyl acetate/methanol yielded **1** as a clear oil (0.100 g, 0.17 mmol) (78.4%): $[\alpha]_D^{25} = -12.8^\circ$ (*c* 0.4 CH₃OH); ¹H NMR (500 MHz, CD₃OD, 298K) $\delta = 7.09$ -7.37 (m, 15H, aromatic), 4.38-4.76 (m, 7H, -OCH₂Ph, H₄), 4.25 (broad m, 1H, H₂, H₂ minor), 4.20 (dd apparent t, 0.8H, H₇, J_{7,8} = 6.0Hz, J_{6,7} = 6.3Hz), 4.10 (dd apparent, 1H, H₈, J_{8,9} = 6.6Hz), [4.06-4.10, m, 0.2H, H₈], [4.01, dd apparent t, 0.2H, H₇], 3.80-3.89 (m, 1H, H₅, H₅ minor), 3.71 (dd, 0.8H, H_{10a}, J_{10a,10b} = 10.6Hz, J_{5,10a} = 6.0Hz), [3.67-3.73, m, 0.2H, H_{10a}], 3.59 (dd, 0.8H, H_{10b}, J_{5,10b} = 3.4Hz), [3.57-3.62, m, 0.2H, H_{10b}], [3.55, dd apparent t, 0.2H, H₆], 3.46 (dd apparent t, 1H, H₆, J_{5,6} = 6.4Hz), 2.53 (ddd, 1H, H_{3a}, J_{3a,3b} = 12.7Hz, H_{3a} minor), 1.90-1.97 (m, 1H, H_{3b}, H_{3b} minor), 1.40 (s, 7.2H, *tert*-butyl), [1.32, s, 1.8H, *tert*-butyl]; ¹³C NMR (75 MHz, CD₃OD, 298K) $\delta = 168.7, 156.1, 139.8, 139.7, 139.5, (81.9), 81.5, (79.8), 78.4, (76.9), 76.6, 74.9, 74.7,$

(74.6), (74.5), 74.3, (74.2), 74.1, (73.9), 72.9, (70.4), 70.2, (59.9), 59.8, 33.9, (28.8), 28.6 ppm; HRMS (ES) calc. for $C_{35}H_{40}NO_8$ (M - H)⁻: 602.2759. Found (M - H)⁻: 602.2755.

1-(2'-(*tert*-Butyloxycarbonyl)-amino-3',4',6'-tri-*O*-benzyl-2'-deoxy- α -D-

glucopyranosyl)-2-propene (3): Compound **2** (0.202 g, 0.43 mmol) was dissolved in methanol (4 mL). Addition of triethylamine (0.6 mL, 4.3 mmol) was followed by addition of di-*tert*-butyl dicarbonate (0.46 g, 2.1 mmol). The reaction mixture was stirred for 16 hours. All reagents and solvent were removed under reduced pressure to provide **3** as a white solid (0.244 g, 0.43 mmol) (quant.): $[\alpha]_D^{25} = +11.4^\circ$ (*c* 3.7 $CHCl_3$); mp 98-101 °C; ¹H NMR (300 MHz, $CDCl_3$, 298K) $\delta = 7.20$ -7.40 (m, 15H, aromatic), 5.85 (dddd, 1H, -CH=CH₂, J = 6.9Hz, J = 7.0Hz, J = 10.1Hz, J = 17.0Hz), 5.61 (d, 1H, *NHBoc*, J = 9.8Hz), 5.02-5.17 (m, 2H, -CH=CH₂), 4.46-4.87 (m, 6H, -OCH₂Ph), 4.20 (dd, 1H, H₅, J = 6.1Hz, J = 6.2Hz), 3.95 (ddd, 1H, H₁, J = 2.0Hz, J = 5.6Hz, J = 7.8Hz), 3.80-3.89 (m, 2H, H₂, H_{6a}), 3.68-3.79 (m, 2H, H₃, H_{6b}), 3.55-3.60 (m, 1H, H₄), 2.17-2.39 (m, 2H, allylic), 1.45 (s, 9H, *tert*-butyl); ¹³C NMR (75 MHz, $CDCl_3$, 298K) $\delta = 155.8$, 138.3, 137.8, 137.6, 134.6, 127.4-128.6 (aromatic carbons), 117.1, 79.1, 74.9, 74.9, 73.4, 73.2, 72.1, 71.8, 68.3, 68.0, 48.9, 35.5, 28.4 ppm; MS (ES) calc. for $C_{35}H_{43}NO_6$ (M + Na)⁺: 596.30. Found (M + Na)⁺: 596.30; Anal. Calcd for $C_{35}H_{43}NO_6$: 73.27 C, 7.55 H, 2.44 N. Found: 73.43 C, 7.75 H, 2.19 N.

2-iodomethyl-(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-(*tert*-Butyloxycarbonyl)-6,7-di-*O*-benzyl-

octahydro-5-(benzyl hydroxymethyl)-pyrano[3,2-*b*]pyrrole (4): Compound **3** (0.29 g, 0.51 mmol) was dissolved in 1:1 CH_2Cl_2 /diethyl ether (10 mL). The solution was cooled to 0 °C before addition of iodine (0.39 g, 1.5 mmol). After 1 hour, the reaction mixture

was warmed to ambient temperature before being worked-up by the addition of 20 mL saturated aq. sodium thiosulphate. With shaking, the solution became colorless. The product was extracted into CH₂Cl₂ (3 x 20 mL), dried (Na₂SO₄), concentrated and purified by flash chromatography using 5:1 hexanes/ethyl acetate. The product **4** was isolated as a single stereoisomer as a pale yellow oil (0.32 g, 0.46 mmol) (89.6%): $[\alpha]_D^{25} = -23.9^\circ$ (*c* 2.5 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.13-7.40$ (m, 15H, aromatic), 4.32-4.78 (broad m, 8H), 3.64-4.08ppm (broad m's, 5H), 3.55 (broad m, 1H), 3.45ppm (broad t, 2H), 2.40 (broad m, 1H, H_{3a}), 1.9 (m, 1H, H_{3b}), 1.5ppm (s, 9H, *tert*-butyl); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta = 153.8, 138.8, 138.6, 138.4, 127.4-128.7$ (aromatic carbons), 80.7, 77.6, 74.2, 74.0, 73.7, 73.5, 72.5, 70.2, 69.4, 59.5, 56.8, 36.5, 28.8, 14.8 ppm; MS (ES) calc. for C₃₅H₄₃INO₆ (M + H)⁺: 700.21. Found (M + H)⁺: 700.08. Calc. for C₃₅H₄₂INNaO₆ (M + Na)⁺: 722.20. Found (M + Na)⁺: 722.01; Anal. Calcd for C₃₅H₄₂INO₆: 60.09 C, 6.05 H, 2.00 N. Found: 60.02 C, 5.74 H, 2.39 N.

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-(*tert*-Butyloxycarbonyl)-6,7-di-*O*-benzyl-octahydro-5-(benzyl hydroxymethyl)-pyrano[3,2-*b*]pyrrolo[1,2-*c*]-oxazol-3-one (5**):** Compound **4** (0.28 g, 0.41 mmol) was dissolved in toluene (6 mL). Addition of silver acetate (0.68 g, 4.1 mmol) made the solution instantly become colorless. The reaction was stirred for 16 hours at ambient temperature. The reaction mixture was diluted with 10 mL ethyl acetate then was filtered through celite. The product was concentrated under reduced pressure then purified using flash chromatography using 1:1 hexanes/ethyl acetate to yield **5** as a white solid (0.19 g, 0.37 mmol) (92.1%): $[\alpha]_D^{25} = +21.9^\circ$ (*c* 2.0 CHCl₃); mp 106-111 °C; ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.14-7.48$ (m, 15H, aromatic), 4.95 (d, 1H, -OCH₂Ph, *J* = 11.1Hz), 4.88 (d, 1H, -OCH₂Ph, *J* = 11.4Hz), 4.75 (d, 1H, -OCH₂Ph, *J* =

11.4Hz), 4.68 (ddd, 1H, H₉, J_{3a,9} = 1.0Hz, J_{3b,9} = 5.7Hz, J_{8,9} = 5.7Hz), 4.42-4.53 (m, 3H, -OCH₂Ph, H_{11a}), 4.38 (d, 1H, -OCH₂Ph, J = 12.2Hz), 4.08-4.20 (m, 2H, H_{11b}, H₂), 3.95 (dd, 1H, H₈, J_{7,8} = 6.3Hz), 3.72-3.82 (m, 1H, H₅), 3.64-3.72 (m, 2H, H₆, H₇), 3.53-3.59 (dd, 1H, H_{10a}, J_{10b,10a} = 10.6Hz, J_{5,10a} = 4.7Hz), 3.47-3.53 (dd, 1H, H_{10b}, J_{5,10b} = 2.8Hz), 2.10 (ddd, 1H, H_{3a}, J_{2,3a} = 5.3Hz, J_{3a,3b} = 13.2Hz), 1.59 (ddd, 1H, H_{3b}, J_{2,3b} = 10.6Hz); ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 161.1, 138.4, 138.1, 138.0, 127.6-128.4 (aromatic carbons), 80.7, 76.9, 75.4, 74.7, 74.4, 73.5, 73.2, 69.5, 66.8, 65.7, 57.5, 38.1 ppm; MS (ES) calc. for C₃₁H₃₃NNaO₆ (M + Na)⁺: 538.22. Found (M + Na)⁺: 538.22; Anal. Calcd for C₃₁H₃₃NO₆: 72.21 C, 6.45 H, 2.72 N. Found: 72.17 C, 6.69 H, 2.57 N.

2-hydroxymethyl-(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-6,7-di-*O*-Benzyl-octahydro-5-(benzyl-hydroxymethyl)-pyrano[3,2-*b*]pyrrole (6): Compound **5** (0.18 g, 0.35 mmol) was dissolved in methanol (8 mL). After addition of potassium hydroxide (1.5 g, 26.2 mmol) the solution was heated to reflux for 4 hours. The reaction mixture was then cooled to 0°C followed by acidification by addition of 5 mL 3 M aq. HCl. The methanol was removed under reduced pressure. The reaction mixture was brought to pH 9 by addition of 20 mL aq. saturated sodium bicarbonate. The product was extracted into CH₂Cl₂ (4 x 10 mL) then dried (Na₂SO₄), concentrated under reduced pressure and purified by flash chromatography using first 1:1 hexanes/ethyl acetate then 5:1 ethyl acetate/methanol to yield **6** as a pale yellow solid (0.15 g, 0.27 mmol) (78.4%): [α]_D²⁵ = +38.9° (c 2.8 CHCl₃); mp 91-94 °C; ¹H NMR (300 MHz, CDCl₃, 298K) δ = 7.15-7.42 (m, 15H, aromatic), 4.45-4.65 (m, 6H, -OCH₂Ph), 4.25-4.35 (m, 1H, H₉), 4.05-4.11 (m, 1H, H₅), 3.82 (dd, 1H, H_{10a}, J_{10a,10b} = 10.2Hz, J_{5,10a} = 6.7Hz), 3.77 (dd apparent t, 1H, H₇, J_{7,8} = 4.5Hz, J_{6,7} = 4.4Hz), 3.61 (dd, 1H, H_{10b}, J_{5,10b} = 5.1Hz), 3.56 (dd apparent t, 1H, H₆, J_{5,6} =

4.2Hz), 3.41-3.51 (m, 2H, H_{11a}, H₂), 3.29 (broad s, 0.3H, NH), 3.25 (dd, 1H, H_{11b}, J_{11a,11b} = 12.1Hz, J_{11b,2} = 7.3Hz), 3.09 (dd, 1H, H₈, J_{8,9} = 3.7Hz), 2.0 (ddd, 1H, H_{3a}, J = 2.1Hz, J = 8.0Hz, J_{3a,3b} = 14.0Hz), 1.59-1.69 (m, 1H, H_{3b}, J = 6.0 Hz, J = 7.0Hz); ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 138.1, 137.8, 137.5, 127.7-128.6 (aromatic carbons), 74.4, 73.9, 73.4, 73.3, 72.8, 72.7, 71.8, 67.7, 64.1, 60.2, 57.4, 33.9 ppm; MS (ES) calc. for C₃₀H₃₆NO₅ (M + H)⁺: 490.26. Found (M + H)⁺: 490.39. Calc. for C₃₀H₃₅NNaO₅ (M + Na)⁺: 512.24. Found (M + Na)⁺: 512.36; Anal. Calcd for C₃₀H₃₅NO₅: 73.59 C, 7.21 H, 2.86 N. Found: 73.23 C, 7.52 H, 2.86 N.

2-hydroxymethyl-(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-(*tert*-Butyloxycarbonyl)-6,7-di-*O*-benzyl-octahydro-5-(benzyl-hydroxymethyl)-pyrano[3,2-*b*]pyrrole (7): Compound **6** (0.048 g, 0.098 mmol) was dissolved in methanol (4 mL). Addition of triethylamine (0.68 mL, 4.9 mmol) was followed by addition of di-*tert*-butyl dicarbonate (0.11 g, 0.49 mmol). The reaction mixture was stirred for 16 hours. All reagents and solvent were removed under reduced pressure to provide **7** as a colorless oil (0.052 g, 0.088 mmol) (89.7%): [α]_D²⁵ = -10.4° (*c* 2.0 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 313K) δ = 7.14-7.41 (m, 15H, aromatic), 4.36-4.90 (m, 8H, -OCH₂Ph, H₂, H_{11a}), 3.95-4.14 (m, 3H, H_{11b}, H₉, H₇), 3.85-3.95 (m, 1H, H₈), 3.52-3.88 (m, 4H, H₅, H₆, H_{10a}, H_{10b}), 2.31 (broad s, 1H, H_{3a}), 1.75 (broad singlet, 1H, H_{3b}), 1.45 (s, 9H, *tert*-butyl); ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 155.8, 138.4, 138.0, 137.9, 126.9-128.7 (aromatic carbons), 80.9, 75.2, 73.4, 73.2, 72.7, 68.9, 66.8, 60.2, 59.8, 32.1, 29.7, 28.4, 22.7, 14.2 ppm; MS (ES) calc. for C₃₅H₄₄NO₇ (M + H)⁺: 590.31. Found (M + H)⁺: 590.30; Anal. Calcd for C₃₅H₄₃NO₇: 71.28 C, 7.35 H, 2.38 N. Found: 71.03 C, 7.59 H, 2.33 N.

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-(*tert*-Butyloxycarbonyl)-6,7-di-*O*-benzyl-octahydro-5-(benzyl-hydroxymethyl)-pyrano[3,2-*b*]pyrrole-2-carboxylic acid benzyl ester (8):

Compound **7** (0.018 g, 0.029 mmol) was dissolved in DMF (3 mL). Addition of cesium carbonate (0.015 g, 0.045 mmol) was followed by addition of benzyl bromide (0.011 mL, 0.089 mmol). The reaction mixture was stirred for 1 hour, and then the solvent was removed under reduced pressure. The reaction mixture was diluted with 10 mL water followed by extraction into CH₂Cl₂ (3 x 10 mL), dried over anhydrous Na₂SO₄, and concentrated under reduced pressure before being purified by flash chromatography using 4:1 hexanes/ethyl acetate to yield **8** as a colorless oil (0.11 g, 0.18 mmol) (83.6%): $[\alpha]_D^{25} = -14.2^\circ$ (*c* 0.4 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.05$ -7.41 (m, 20H, aromatic), 5.19 (d, 1H, -OCH₂Ph, *J* = 12.1Hz), 5.11 (d, 1H, -OCH₂Ph, *J* = 12.3Hz), 4.35-4.80 (m, 8H, -OCH₂Ph, H₂, H₉), 4.28 (dd apparent t, 1H, H₇, *J*_{7,8} = 5.6Hz, *J*_{6,7} = 5.7Hz), 4.17 (dd apparent t, 1H, H₈, *J*_{8,9} = 5.8Hz), 3.83-3.93 (m, 1H, H₅), 3.70 (dd, 1H, H_{10a}, *J*_{10a,10b} = 10.4Hz, *J*_{5,10a} = 6.3Hz), 3.57 (dd, 1H, H_{10b}, *J*_{5,10b} = 3.4Hz), 3.48 (dd apparent t, 1H, H₆, *J*_{5,6} = 5.8Hz), 2.40-2.55 (m, 1H, H_{3a}), 1.80-2.00 (m, 1H, H_{3b}), 1.38 (s, 9H, *tert*-butyl); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta = 172.9, 153.6, 138.4, 138.1, 138.1, 135.3, 127.0$ -128.7 (aromatic carbons), 80.4, 76.0, 74.8, 73.6, 73.5, 73.4, 72.8, 71.2, 69.0, 66.8, 58.8, 58.2, 33.0, 28.1 ppm; MS (ES) calc. for C₄₂H₄₇NNaO₈ (M + Na)⁺: 716.32. Found (M + Na)⁺: 716.11; Anal. Calcd for C₄₂H₄₇NNaO₈: 70.37 C, 6.61 H, 1.95 N. Found: 70.44 C, 6.71 H, 1.88 N.

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-(*tert*-Butyloxycarbonyl)-6,7-di-*O*-benzyl-octahydro-5-(benzyl-hydroxymethyl)-pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methyleamide (9):

Compound **1** (0.10 g, 0.16 mmol) was dissolved in acetonitrile (6 mL). Addition of

diisopropylethylamine (0.11 mL, 0.64 mmol) was followed by addition of TBTU (0.10 g, 0.32 mmol) and methylamine hydrochloride (0.02 g, 0.32 mmol). The reaction mixture was stirred at ambient temperature for 4 hours. The reaction mixture was diluted with 15 mL water followed by extraction into CH₂Cl₂ (3 x 15 mL), dried (Na₂SO₄), then concentrated under reduced pressure and purified by flash chromatography using 3:1 ethyl acetate/hexanes to provide **9** as a clear oil (0.086 g, 0.14 mmol) (84.3%): $[\alpha]_D^{25} = -3.5^\circ$ (*c* 0.4 CHCl₃); ¹H NMR (300 MHz, acetone-D₆, 298K) $\delta = 7.15$ -7.45 (m, 15H, aromatic), 4.45-4.82 (m, 8H, -OCH₂Ph, H₂, H₉), 4.39 (dd apparent t, 1H, H₇, J_{7,8} = 5.3Hz, J_{6,7} = 5.3Hz), 4.28 (dd, 1H, H₂, J = 5.9Hz, J = 7.9Hz), 4.11 (dd apparent t, 1H, H₈, J_{8,9} = 5.5Hz), 3.86-3.95 (m, 1H, H₅), 3.78 (dd, 1H, H_{10a}, J_{10a,10b} = 10.5Hz, J_{5,10a} = 6.1Hz), 3.68 (dd, 1H, H_{10b}, J_{5,10b} = 3.8Hz), 3.55 (dd apparent t, 1H, H₆, J_{5,6} = 5.7Hz), 2.80 (d, 3H, -NHCH₃), 2.42 (ddd, 1H, H_{3a}, J = 6.3Hz, J = 8.1Hz, J = 13.4Hz), 1.89 (ddd, 1H, H_{3b}, J = 5.8Hz, J = 6.2Hz, J = 12.0Hz), 1.37 (s, 9H, *tert*-butyl); ¹³C NMR (75 MHz, acetone-D₆, 298K) $\delta = 173.8$, 154.4, 139.7, 139.5 (2), 128.1-129.0 (aromatic carbons), 79.4, 76.5, 75.9, 74.0, 73.4 (2), 72.9, 71.5, 69.9, 60.8, 59.3, 34.6, 28.2, 25.9 ppm; MS (ES) calc. for C₃₆H₄₄N₂NaO₇ (M + Na)⁺: 639.30. Found (M + Na)⁺: 639.27; Anal. Calcd for C₃₆H₄₄N₂O₇: 70.11 C, 7.19 H, 4.54 N. Found: 70.19 C, 7.43 H, 4.47 N.

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-6,7-di-*O*-Benzyl-octahydro-5-(benzyl-hydroxymethyl)-pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methylamide (10**):** Compound **9** (0.035 g, 0.057 mmol) was dissolved in CH₂Cl₂ (1.5 mL). The reaction mixture was then cooled to 0°C. Trifluoroacetic acid (0.5 mL, 6.73 mmol) was added slowly. After 1 hour the solution was co-distilled with toluene (2 x 5 mL), and was normally used directly in the next reaction. Purification by flash chromatography using 10:1 ethyl acetate/methanol

provided **10** as a clear oil (0.027 g, 0.14 mmol) (93.1%): $[\alpha]_D^{25} = -11.2^\circ$ (c 1.0 CHCl_3); ^1H NMR (500 MHz, CDCl_3 , 298K) $\delta = 7.58$ (broad q, 1H, $-\text{NHCH}_3$), 7.20-7.38 (m, 15H, aromatic), 4.47-4.62 (m, 6H, $-\text{OCH}_2\text{Ph}$), 4.18-4.27 (m, 1H, H_2), 4.04-4.14 (m, 1H, H_5), 3.78-3.91 (m, 2H, H_9 , H_{10a}), 3.73 (broad dd, 1H, H_7 , $J_{7,8} = 3.9\text{Hz}$, $J_{6,7} = 3.8\text{Hz}$), 3.65 (dd, H_{10b} , $J_{10a,10b} = 10.2\text{Hz}$, $J_{5,10b} = 5.6\text{Hz}$), 3.59 (broad dd, 1H, H_6 , $J_{5,6} = 3.8\text{Hz}$), 2.89 (broad dd, 1H, $-\text{NH}$, $J = 3.0\text{Hz}$, $J = 3.3\text{Hz}$), 2.78 (d, 3H, $-\text{NHCH}_3$, $J = 5.0\text{Hz}$), 2.36 (ddd, 1H, H_{3a} , $J = 1.96\text{Hz}$, 9.6Hz , 14.4Hz), 1.95 (ddd, 1H, H_{3b} , $J = 5.0\text{Hz}$, 7.0Hz , 14.4Hz); ^{13}C NMR (75 MHz, CDCl_3 , 298K) $\delta = 175.3$, 138.1, 137.8, 137.6, 127.6-128.6 (aromatic carbons), 74.2, 73.5, 73.3, 72.8, 72.6, 72.4, 71.6, 67.6, 60.9, 58.8, 36.7, 25.7 ppm; MS (ES) calc. for $\text{C}_{31}\text{H}_{37}\text{N}_2\text{O}_5$ ($\text{M} + \text{H}$) $^+$: 517.27. Found ($\text{M} + \text{H}$) $^+$: 517.30; Anal. Calcd for $\text{C}_{31}\text{H}_{36}\text{N}_2\text{O}_5$: 72.07 C, 7.02 H, 5.42 N. Found: 72.11 C, 7.13 H, 5.36 N.

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-Acetyloctahydro-6,7-di-*O*-benzyl-5-(benzyl-hydroxymethyl)-pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methanamide (11**):**

Compound **10** (0.029 g, 0.056 mmol) was dissolved in pyridine (2 ml) followed by addition of acetic anhydride (0.053 mL, 0.56 mmol). The reaction mixture was stirred for 15 hours, then the solvent and reagents were removed under reduced pressure and the product was purified by flash chromatography using 10:1 ethyl acetate/methanol to provide **11** as a white solid (0.030 g, 0.054 mmol) (97% over 2 steps): $[\alpha]_D^{25} = +49.6^\circ$ (c 1.0 CHCl_3); decomposed at 142-147 $^\circ\text{C}$; ^1H NMR (500 MHz, CDCl_3 , 298K, 0.036M) $\delta = 7.23$ -7.41 (m, 13H, aromatic), 7.11-7.20 (m, 2H, aromatic), 5.97 (broad q, 1H, $-\text{NHCH}_3$), 4.98 (ddd, 1H, H_9 , $J_{3a,9} = 11.7\text{Hz}$, $J_{3b,9} = 7.3\text{Hz}$, $J_{8,9} = 7.4\text{Hz}$), 4.93 (d, 1H, $-\text{OCH}_2\text{Ph}$, $J = 11.2\text{Hz}$), 4.80 (d, 1H, $-\text{OCH}_2\text{Ph}$, $J = 10.8\text{Hz}$), 4.52-4.64 (m, 4H, $-\text{OCH}_2\text{Ph}$), 4.25 (dd apparent d, 1H, H_2 , $J_{2,3a} = 9.5\text{Hz}$, $J_{2,3b} = 0.8\text{Hz}$), 4.03 (dd, 1H, H_8 , $J_{7,8} = 9.2\text{Hz}$), 3.63-3.78

(m, 4H, H₅, H₆, H_{10a}, H_{10b}), 3.57 (dd apparent t, 1H, H₇, J_{6,7} = 9.0Hz), 2.82 (d, 3H, -NHCH₃, J = 4.2Hz), 2.35 (ddd, 1H, H_{3a}, J_{3a,3b} = 12.3Hz), 2.15 (s, 3H, -COCH₃), 2.05 (ddd, 1H, H_{3b}); ¹³C NMR (75 MHz, CDCl₃, 298K) (major conformer) δ = 172.1, 171.6, 137.8, 137.7, 137.5, 127.8-128.5 (aromatic carbons), 83.0, 78.0, 75.9, 74.9, 73.7, 73.6, 73.3, 68.8, 60.1, 57.8, 28.4, 26.4, 22.9 ppm; HRMS (ES) calc. for C₃₃H₃₉N₂O₆ (M + H)⁺: 559.2802. Found (M + H)⁺: 559.2801.

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-Acetyloctahydro-6,7-dihydroxy-5-(hydroxymethyl)-

pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methanamide (12): Compound **11** (0.027 g, 0.048 mmol) was dissolved in methanol (10 mL). Addition of Pearlman's catalyst (20% palladium hydroxide on carbon) (0.030 g, approx. 0.028 mmol) was followed by addition of 1M aq. HCl (0.072 mL, 0.072 mmol). The reaction mixture was stirred vigorously under hydrogen atmosphere (10 psi) for 4.5 hours, after which it was flushed with nitrogen and filtered. The product was then concentrated under reduced pressure to provide **12** as a clear oil (0.014 g, 0.048 mmol) (quant.): [α]²⁵_D = +20.5° (c 1.0 CHCl₃); ¹H NMR (500 MHz, D₂O, 298K, 0.035M) δ = 4.55-4.64 (m, 0.85H, H₉), 4.32 (dd apparent d, 0.85H, H₂, J_{2,3a} = 10.2Hz, J_{2,3b} = 1.3Hz), [4.11, dd, 0.15H, H₈, J_{8,9} = 7.7Hz, J_{7,8} = 8.0Hz], 3.99 (dd, 0.85H, H₈, J_{8,9} = 7.2Hz, J_{7,8} = 9.2Hz), 3.70 (dd, 0.85H, H_{10a}, J_{5,10a} = 2.2Hz, J_{10a,10b} = 12.4Hz), [3.67-3.72, m, 0.15H, H_{10a}], 3.62 (dd, 0.85H, H_{10b}, J_{5,10b} = 5.0Hz), [3.60, dd, 0.15H, H_{10b}, J = 1.7Hz, J = 12.5Hz], 3.51-3.56 (m, 0.85H, H₅), 3.51 (dd apparent t, 0.85H, H₇, J_{6,7} = 9.9Hz, H₇ minor), [3.45-3.50, m, H₅], [3.35, dd, 0.15H, H₆, J_{5,6} = 9.9Hz], 3.32 (dd apparent t, 0.85H, H₆, J_{5,6} = 9.7Hz), [2.63, s, 0.45H, -COCH₃], 2.58 (s, 2.55H, -COCH₃), 2.55 (ddd, 0.85H, H_{3a}, J_{3a,9} = 11.7Hz, J_{3b,3a} = 13.5Hz, H_{3a} minor), 2.10 (s, 2.55H, -NHCH₃), [1.93, ddd, 0.15H, H_{3b}, J_{2,3b} = 1.8Hz, J_{3b,9} = 6.7Hz,

$J_{3a,3b} = 12.9\text{Hz}$], [1.83, s, 0.45H, $-\text{COCH}_3$], 1.79 (ddd, 0.85H, H_{3b} , $J_{3b,9} = 7.8\text{Hz}$, H_{3b} minor); ^{13}C NMR (75 MHz, D_2O , 298K) (major conformer) $\delta = 174.6, 174.4, 74.5, 73.8, 73.3, 68.6, 61.5, 61.0, 58.1, 28.3, 26.1, 22.3$ ppm; HRMS (ES) calc. for $\text{C}_{12}\text{H}_{20}\text{N}_2\text{O}_6\text{Na}$ ($\text{M} + \text{Na}$) $^+$: 311.1212. Found ($\text{M} + \text{Na}$) $^+$: 311.1214.

(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-1-Acetyloctahydro-6,7-*O*-acetyl-5-(hydroxymethyl-*O*-acetyl)-

pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methanamide (13**):** Compound **12** (0.014 g, 0.048 mmol) was dissolved in pyridine (1 mL). Acetic anhydride (0.046 mL, 0.48 mmol) was added and the reaction mixture was stirred at ambient temperature for 15 hours. All solvent and reagent were removed under reduced pressure providing compound **13** as a clear oil (0.020 g, 0.048 mmol) (quant.): $[\alpha]_{\text{D}}^{25} = +53.3^\circ$ (c 0.3 CHCl_3); ^1H NMR (500 MHz, CDCl_3 , 298K, 0.036 M) $\delta = 6.37$ (broad q, 0.87H, $-\text{NHCH}_3$), [5.97, broad q, 0.13H, $-\text{NHCH}_3$], [5.25, dd, 0.13H, H_7], 5.22 (dd apparent t, 0.87H, H_7 , $J_{7,8} = 9.8\text{Hz}$, $J_{6,7} = 10.0\text{Hz}$), 5.10 (ddd, 0.87H, H_9 , $J_{3a,9} = 11.8\text{Hz}$, $J_{3b,9} = 7.1\text{Hz}$, $J_{8,9} = 7.3\text{Hz}$), [5.03, dd, 0.13H, H_6 , $J_{6,7} = 9.5\text{Hz}$, $J_{5,6} = 9.2\text{Hz}$], 4.99 (dd apparent t, 0.87H, H_6 , $J_{5,6} = 9.9\text{Hz}$), [4.70, ddd, 0.13H, H_9 , $J_{3a,9} = 11.6\text{Hz}$, $J_{3b,9} = 6.8\text{Hz}$, $J_{8,9} = 7.0\text{Hz}$], [4.54, dd, 0.13H, H_8 , $J_{7,8} = 7.8\text{Hz}$], [4.41, dd, 0.13H, H_2 , $J_{2,3a} = 9.4\text{Hz}$, $J_{2,3b} = 1.5\text{Hz}$], 4.27-4.35 (m, 1.87H, H_2 , H_{10a} , H_{10a} minor), 4.19 (dd, 0.87H, H_8), 4.08 (dd, 1H, H_{10b} , $J_{5,10b} = 2.4\text{Hz}$, $J_{10a,10b} = 12.5\text{Hz}$, H_{10b} minor), 3.98 (ddd, 0.87H, H_5 , $J_{5,10a} = 4.3\text{Hz}$), [3.89, ddd, 0.13H, H_5 , $J_{5,10b} = 2.6\text{Hz}$, $J_{5,10a} = 4.9\text{Hz}$], [2.84, d, 0.39H, $-\text{NHCH}_3$, $J = 4.8\text{Hz}$], 2.81 (d, 2.61H, $-\text{NHCH}_3$, $J = 4.8\text{Hz}$), [2.73, ddd, 0.13H, H_{3a} , $J_{3a,3b} = 12.8\text{Hz}$], 2.48 (ddd, 0.87H, H_{3a} , $J_{2,3a} = 9.8\text{Hz}$, $J_{3a,3b} = 12.3\text{Hz}$), 2.13 (s, 2.61H, $-\text{COCH}_3$), 2.09 (s, 3H, $-\text{COCH}_3$, $-\text{COCH}_3$ minor), 2.06 (s, 2.61H, $-\text{COCH}_3$), [2.04, s, 0.39H, $-\text{COCH}_3$], [2.03, s, 0.39H, $-\text{COCH}_3$], [2.02, s, 0.39H, $-\text{COCH}_3$], 2.00-2.05 (m, 0.87H, H_{3b}), 2.01 (s, 2.61H, $-\text{COCH}_3$), [1.86, s, 0.39H, -

COCH₃]; ¹³C NMR (75 MHz, CDCl₃, 298K) (major conformer) δ = 171.8, 170.7, 170.3, 169.9, 169.5, 73.9, 73.8, 69.7, 68.0, 62.2, 58.6, 58.1, 28.2, 26.4, 22.0, 20.9, 20.7, 20.6 ppm; HRMS (ES) calc. for C₁₈H₂₇N₂O₉ (M + H)⁺: 415.1711. Found (M + H)⁺: 415.1711.

***N*-Acetyl-glycyl-(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-6,7-di-*O*-benzyl-octahydro-5-(benzyl-hydroxymethyl)-pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methanamide (14):**

Compound **10** (0.105 g, 0.203 mmol) was dissolved in *N,N*-dimethylformamide (6 mL) and the solution was cooled to 0°C. The reaction was stirred under inert atmosphere. Diisopropylethylamine (0.212 mL, 1.218 mmol) and PyBOP (0.317 g, 0.609 mmol) were added and the solution was stirred for 10 min. Fmoc-Gly-OH (0.181 g, 0.609 mmol) was added and the reaction mixture was stirred for a further 5 min. before being allowed to warm to ambient temperature where it stirred for 18 hours. The red solution was diluted with ethyl acetate, washed with 1 M HCl (10 mL), then brine (10mL), dried and evaporated giving a red oil. The product was purified by flash chromatography using ethyl acetate, giving a white solid product (0.085g) (53%), along with unreacted starting material (0.025 g) (24%). The coupled product was dissolved in 4 mL dichloromethane, cooled to 0°C, and treated with piperidine (1 mL). The reaction mixture was stirred for 1 hour before the solvent and reagents were removed under reduced pressure, leaving a white solid. The intermediate was dissolved in 4 mL pyridine followed by addition of acetic anhydride (0.5 mL). The reaction mixture was stirred for 15 hours, then the solvent and reagents were removed under reduced pressure and the product was purified by flash chromatography using 10:1 ethyl acetate/methanol to provide **14** as a clear oil (0.059 g, 0.096 mmol) (47.1%) (over 3 steps): [α]_D²⁵ = +23.5° (*c* 1.0 CHCl₃); ¹H NMR (500 MHz, CDCl₃, 298K, 0.033 M) δ = 7.10-7.38 (m, 15H, aromatic), 6.30 (t, 1H, -NH_(Gly), -NH_(Gly))

minor), 6.22 (q, 1H, -NHCH₃, -NHCH₃ minor), 4.95 (d, 0.95H, -OCH₂Ph, J = 11.3Hz), 4.83 (ddd, 0.95H, H₉, J_{3b,9} = 7.6Hz, J_{3a,9} = 12.0Hz, J_{8,9} = 7.2Hz), 4.75 (d, 0.95H, -OCH₂Ph, J = 10.8Hz), 4.51-4.64 (m, 4.15H, -OCH₂Ph, -OCH₂Ph minor, H₉ minor), [4.22, dd, 0.05H, H₈, J_{8,9} = 6.5Hz, J_{7,8} = 8.2Hz], 4.17 (dd apparent d, 0.95H, H₂, J_{2,3b} = 1.0Hz, J_{2,3a} = 9.7Hz), 4.05-4.13 (m, 1.9H, H₈, H_{α1(Gly)}, H₂ minor), 4.02 (dd, 0.95H, H_{α2(Gly)}, J_{Hα2(Gly),NH} = 4.1Hz, J_{Hα1(Gly),Hα2(Gly)} = 17.1Hz), 3.68-3.76 (m, 2H, H₇, H_{10a}, H₇ minor, H_{10a} minor), 3.61-3.67 (m, 2.1H, H₅, H_{10b}, H₅ minor, H_{10b} minor, H_{α2(Gly)} minor, H_{α1(Gly)} minor), 3.53 (dd apparent t, 0.95H, H₆, J_{6,7} = 9.7Hz, J_{5,6} = 9.4Hz), [3.52-3.58, m, 0.05H, H₆], [2.80, d, 0.15H, -NHCH₃, J = 5.0Hz], 2.74 (d, 2.85H, -NHCH₃, J = 5.0Hz), 2.36 (ddd, 1H, H_{3a}, J_{3a,3b} = 12.5Hz, H_{3a} minor), 1.99 (ddd, 1H, H_{3b}, H_{3b} minor), 1.92 (s, 3H, -COCH₃, -COCH₃ minor); ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 171.5, 170.5, 169.5, 137.8, 137.6, 137.3, 127.7-128.8 (aromatic carbons), 81.8, 78.2, 75.7, 74.8, 73.8, 73.7, 73.4, 68.7, 59.0, 58.4, 42.9, 28.6, 26.4, 22.7 ppm; HRMS (ES) calc. for C₃₅H₄₁N₃O₇Na (M + Na)⁺: 638.2837. Found (M + Na)⁺: 638.2841.

***N*-Acetyl-glycyl-(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-6,7-dihydroxy-5-(hydroxymethyl)-**

octahydro-pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methanamide (15): Compound **14** (0.020 g, 0.032 mmol) was dissolved in methanol (5 mL). Addition of Pearlman's catalyst (20% palladium hydroxide on carbon) (0.020 g, approx. 0.019 mmol) was followed by addition of 1 M aq. HCl (0.010 mL, 0.010 mmol). The reaction mixture was stirred vigorously under hydrogen atmosphere (10 psi) for 4.5 hours, after which it was flushed with nitrogen and filtered. The product was then concentrated under reduced pressure to provide **15** as a yellow oil (0.012 g, 0.035 mmol) (quant.): [α]²⁵_D = -2.5° (c 0.6 CHCl₃); ¹H NMR (500 MHz, D₂O, 298K, 0.035M) δ = 4.57-4.66 (m, 0.9H, H₉), 4.47 (d,

0.9H, $H_{\alpha 1(\text{Gly})}$, $J_{H_{\alpha 1(\text{Gly})}, H_{\alpha 2(\text{Gly})}} = 17.2\text{Hz}$), 4.35 (dd apparent d, 1H, H_2 , $J_{2,3a} = 10.3\text{Hz}$, $J_{2,3b} = 1.0\text{Hz}$, H_2 minor), [4.14, dd apparent t, 0.1H, H_8 , $J_{8,9} = 7.2\text{Hz}$, $J_{7,8} = 7.7\text{Hz}$], 4.01 (dd, 0.9H, H_8 , $J_{8,9} = 7.2\text{Hz}$, $J_{7,8} = 9.0\text{Hz}$), 3.97 (d, 0.9H, $H_{\alpha 2(\text{Gly})}$), [3.87, d, 0.1H, $H_{\alpha 1(\text{Gly})}$, $J_{H_{\alpha 1(\text{Gly})}, H_{\alpha 2(\text{Gly})}} = 17.1\text{Hz}$], 3.70 (dd, 1H, H_{10a} , $J_{5,10a} = 2.3\text{Hz}$, $J_{10a,10b} = 12.2\text{Hz}$, H_{10a} minor), 3.62 (dd, 1H, H_{10b} , $J_{5,10b} = 5.0\text{Hz}$, H_{10b} minor), [3.45-3.51, m, 0.1H, H_5], 3.50-3.58 (m, 1.9H, H_5 , H_7 , H_7 minor), [3.36, dd, 0.1H, H_6], 3.32 (dd apparent t, 0.9H, H_6 , $J_{6,7} = 9.9\text{Hz}$, $J_{5,6} = 9.5\text{Hz}$), [2.64, s, 0.3H, $-\text{NHCH}_3$], 2.58 (s, 2.7H, $-\text{NHCH}_3$), 2.54 (ddd, 1H, H_{3a} , $J_{3a,9} = 11.2\text{Hz}$, $J_{3a,3b} = 12.7\text{Hz}$, H_{3a} minor), 1.90 (s, 2.7H, $-\text{COCH}_3$), 1.79 (ddd, 1H, H_{3b} , $J_{3b,9} = 7.7\text{Hz}$, H_{3a} minor); ^{13}C NMR (75 MHz, D_2O , 298K) $\delta = 175.0, 174.4, 171.3, 74.5, 73.8, 73.5, 68.8, 61.1, 60.4, 58.6, 42.4, 27.9, 26.2, 22.0$ ppm; HRMS (ES) calc. for $\text{C}_{14}\text{H}_{23}\text{N}_3\text{O}_7\text{Na}$ ($\text{M} + \text{Na}$) $^+$: 368.1428. Found ($\text{M} + \text{Na}$) $^+$: 368.1427.

***N*-Acetyl-glycyl-(2*S*,3*aR*,5*R*,6*R*,7*S*,7*aS*)-6,7-di-*O*-acetyl-5-(hydroxymethyl-*O*-acetyl)-octahydro-pyrano[3,2-*b*]pyrrole-2-carboxamide *N'*-methylamide (**16**):** Compound **15** (0.012g, 0.035 mmol) was dissolved in pyridine (1 mL). Acetic anhydride (0.034 mL, 0.35 mmol) was added and the reaction mixture was stirred at ambient temperature for 15 hours. All solvent and reagent were removed under reduced pressure and the product was purified by flash chromatography using 10:1 ethyl acetate/methanol to provide **16** as a clear oil (0.014 g, 0.030 mmol) (74%): $[\alpha]_D^{25} = +37.3^\circ$ (c 0.3 CHCl_3); ^1H NMR (500 MHz, CDCl_3 , 298K, 0.03 M) $\delta = 6.56$ [broad q, 0.2H, $-\text{NHCH}_3$, $J = 4.6\text{Hz}$], [6.39, broad dd, 0.2H, $-\text{NH}_{(\text{Gly})}$], 6.31 (dd, 0.8H, $-\text{NH}_{(\text{Gly})}$), 6.16 (q, 0.8H, $-\text{NHCH}_3$, $J = 4.6\text{Hz}$), [5.27, dd, 0.2H, H_7 , $J_{7,8} = 8.3$, $J_{6,7} = 9.1\text{Hz}$], 5.19 (dd apparent t, 0.8H, H_7 , $J_{7,8} = 9.4\text{Hz}$, $J_{6,7} = 9.9\text{Hz}$), 5.03 (ddd, 0.8H, H_9 , $J_{3a,9} = 11.8\text{Hz}$, $J_{3b,9} = 7.5\text{Hz}$, $J_{8,9} = 7.2\text{Hz}$), [5.00-5.05, m, 0.2H, H_6], 4.99 (dd apparent t, 0.8H, H_6 , $J_{5,6} = 9.8\text{Hz}$), [4.84, ddd, 0.2H, H_9 , $J_{3a,9} =$

10.9Hz, $J_{3b,9} = 6.7\text{Hz}$, $J_{8,9} = 7.3\text{Hz}$), [4.52, dd apparent t, 0.2H, H_8 , $J = 7.7\text{Hz}$], 4.44 (dd, 1H, $H_{\alpha 1(\text{Gly})}$, $J_{H_{\alpha 1(\text{Gly})},\text{NH}} = 6.3\text{Hz}$, $J_{H_{\alpha 1(\text{Gly})},H_{\alpha 2(\text{Gly})}} = 17.4\text{Hz}$, H_2 minor), 4.26-4.35 (m, 1.8H, H_2 , H_{10a} , H_{10a} minor), 4.23 (dd, 0.8H, H_8), 4.08 (dd, 1H, H_{10b} , $J_{5,10b} = 2.3\text{Hz}$, $J_{10a,10b} = 12.2\text{Hz}$, H_{10b} minor), 3.97 (ddd, 0.8H, H_5 , $J_{5,10a} = 4.2\text{Hz}$), [3.86-3.92, m, 0.4H, H_5 , $H_{\alpha 1(\text{Gly})}$], 3.86 (dd, 0.8H, $H_{\alpha 2(\text{Gly})}$, $J_{H_{\alpha 2(\text{Gly})},\text{NH}} = 3.1\text{Hz}$), [3.56, dd, 0.2H, $H_{\alpha 2(\text{Gly})}$, $J_{H_{\alpha 2(\text{Gly})},\text{NH}} = 3.5\text{Hz}$, $J_{H_{\alpha 1(\text{Gly})},H_{\alpha 2(\text{Gly})}} = 16.9\text{Hz}$], [2.83, d, 0.6H, $-\text{NHCH}_3$, $J = 4.6\text{Hz}$], 2.80 (d, 2.4H, $-\text{NHCH}_3$, $J = 4.6\text{Hz}$), [2.68, m, 0.2H, H_{3a} , $J_{2,3a} = 9.5\text{Hz}$, $J_{3a,3b} = 12.6\text{Hz}$], 2.55 (ddd, 0.8H, H_{3a} , $J_{2,3a} = 10.0\text{Hz}$, $J_{3a,3b} = 12.5\text{Hz}$), 2.13 (s, 2.4H, $-\text{COCH}_3$), 2.09 (s, 3H, $-\text{COCH}_3$), 1.98-2.04 (m, 1H, H_{3b} , H_{3b} minor), [2.02, s, 0.6H, $-\text{COCH}_3$], 2.01 (s, 3H, $-\text{COCH}_3$, $-\text{COCH}_3$ minor), 1.99 (s, 3H, $-\text{COCH}_3$, $-\text{COCH}_3$ minor); ^{13}C NMR (75 MHz, CDCl_3 , 298K) δ = (major conformer) 171.4, 170.7, 170.4, 170.4, 169.4, 168.8, 73.6, 73.4, 69.8, 68.1, 62.1, 58.6, 57.1, 41.9, 28.1, 26.5, 22.9, 20.9, 20.7, 20.5 ppm; HRMS (ES) calc. for $\text{C}_{20}\text{H}_{29}\text{N}_3\text{O}_{10}\text{Na}$ ($\text{M} + \text{Na}$) $^+$: 494.1745. Found ($\text{M} + \text{Na}$) $^+$: 494.1743.

3.8 References

- 1) Wilmot, C. M.; Thornton, J. M. *J. Mol. Biol.* **1988**, *203*, 221-232.
- 2) Fischer, G.; Schmid, F. X. *Biochemistry* **1990**, *29*, 2205-2212.
- 3) Beausoleil, E.; Lubell, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 12902-12908 and references therein.
- 4) Delaney, N. G.; Madison, V. *J. Am. Chem. Soc.* **1982**, *104*, 6635-6641.

- 5) Samanen, J.; Zuber, G.; Bean, J.; Eggleston, D.; Romoff, T.; Kopple, K.; Saunders, M.; Regoli, D. *Int. J. Pept. Protein. Res.* **1990**, *35*, 501-509.
- 6) Sharma, R.; Lubell, W. D. *J. Org. Chem.* **1996**, *61*, 202-209.
- 7) Eberhardt, E. S.; Panasik Jr., N.; Raines, R. T.; *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266.
- 8) Che, Y.; Marshall, G. R. *J. Org. Chem.* **2004**, *69*, 9030-9042 and references therein.
- 9) Tam, J. P.; Miao, Z. *J. Am. Chem. Soc.* **1999**, *121*, 9013-9022 and references therein.
- 10) For recent reviews see: a) Dondoni, A.; Marra, A. *Chem. Rev.* **2000**, *100*, 4395-4421. b) Schweizer, F. *Angew. Chem. Intl. Ed.* **2002**, *41*, 230-264; For recent reviews on the use of glycosyl amino acids as peptidomimetics see; c) Gruner, S. A. W.; Locardi, E.; Lohof, E.; Kessler, H. *Chem. Rev.* **2002**, *102*, 491-514; d) Chakraborty, T. K.; Ghosh, S.; Jayaprakash, S. *Curr. Med. Chem.* **2002**, *9*, 421-435.
- 11) Two examples have been reported: a) Zhang, K.; Schweizer, F. *Synlett* **2005**, *20*, 3111-3115; b) Cipolla, L.; Redaelli, C.; Nicotra, F. *Letters in Drug Design & Discovery* **2005**, *2*, 291-293.
- 12) Manzoni, L.; Colombo, M.; May, E.; Scolastico, C. *Tetrahedron* **2001**, *57*, 249-255.
- 13) Petter, R. C. *Tetrahedron Lett.* **1989**, *30*, 399-402.
- 14) Jeannotte, G.; Lubell, W. D. *J. Org. Chem.* **2004**, *69*, 4656-4662.

- 15) a) Wagaw, S.; Rennels, R. A.; Buchwald, S. L. *J. Am. Chem. Soc.* **1997**, *119*, 8451-8458; b) Kuwano, R.; Sato, K.; Kurokawa, T.; Karube, D.; Ito, Y. *J. Am. Chem. Soc.* **2000**, *122*, 7614-7615; c) Viswanathan, R.; Prabhakaran, E. N.; Plotkin, M. A.; Johnston, J. N. *J. Am. Chem. Soc.* **2003**, *125*, 163-168.
- 16) Koep, S.; Gais, H.-J.; Raabe, G. *J. Am. Chem. Soc.* **2003**, *125*, 13243-13251 and references therein.
- 17) Blankley, C. J.; Kaltenbronn, J. S.; DeJohn, D. E.; Werner, A.; Bennett, L. R.; Bobowski, G.; Krolls, U.; Johnson, D. R.; Pearlman, W. M.; Hoefle, M. L.; Essenberg, A. D.; Cohen, D. M.; Kaplan, H. R. *J. Med. Chem.* **1987**, *30*, 992-998.
- 18) Li, W.; Moeller, K. D. *J. Am. Chem. Soc.* **1996**, *118*, 10106-10112.
- 19) Hanessian, S.; Mcnaughton-Smith, G.; Lombart, H.-G.; Lubell, W. D. *Tetrahedron* **1997**, *53*, 12789-12854 and references 3a-h and 4a-k in reference 17.
- 20) Grotenbreg, G. M.; Timmer, M. S. M.; Llamas-Saiz, A. L.; Verdoes, M.; van der Marel, G. A.; van Raaij, M. J.; Overkleeft, H. S.; Overhand, M. *J. Am. Chem. Soc.* **2004**, *126*, 3444-3446.
- 21) Chakraborty, T. K.; Jayaprakash, S.; Diwan, P. V.; Nagaraj, R.; Jampani, S. R. B.; Kunwar, A. C. *J. Am. Chem. Soc.* **1998**, *120*, 12962-12963.
- 22) It has been reported that substitution of D-proline by cis-3-hydroxy-D-proline (HypC3-OH) in the sequence Boc-Leu-Pro-Gly-NHMe resulted in novel pseudo β -turn like nine-membered ring structure involving an intramolecular hydrogen bond between

LeuNH. Chakraborty, T. K.; Srinivasu, P.; Vengal Rao, R.; Kiran Kumar, S.; Kunwar, A. C.; *J. Org. Chem.* **2004**, *69*, 7399-7402.

23) Vitagliano, L.; Beriso, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *58*, 459-464.

24) a) Brimble, M. A.; Trotter, N. S.; Harris, P. W. R.; Sieg, F. *Bioorg. Med. Chem.* **2005**, *13*, 519-532; b) Alonso De Diego, S. A.; Munoz, P.; Gonzalez-Muniz, R.; Herranz, R.; Martin-Martinez, M.; Cenarruzabeitia, E.; Frechilla, D.; Del Rio, J.; Jimeno, M. L.; Garcia-Lopez, M. T. *Bioorg. Med. Chem.* **2005**, *15*, 2279-2283; c) Ioudina, M.; Uemura, E. *Exp. Neuro.* **2003**, *184*, 923-929.

25) a) Su, J.; Andersson, E.; Horal, P.; Naghavi, M. H.; Palm, A.; Wu, Y.-P.; Eriksson, K.; Jansson, M.; Wigzell, H.; Svennerholm, B.; Vahlne, A. *J. Human Virology* **2001**, *4*, 1-7; b) Laczko, I.; Hollosi, M.; Urge, L.; Ugen, K. E.; Weiner, D. B.; Mantsch, H. H.; Thurin, J.; Otvos Jr., L. *Biochemistry*, **1992**, *31*, 4282-4288.

26) a) Hayashi, T.; Asai, T.; Ogoshi, H. *Tetrahedron Lett.* **1997**, *38*, 3039-3042; b) Zerkout, S.; Dupont, V.; Aubry, A.; Vidal, J.; Collet, A.; Vicherat, A.; Marraud, M. *Int. J. Peptide Protein Res.* **1994**, *44*, 378-387.

27) Cipolla, L.; Lay, L.; Nicotra, F. *J. Org. Chem.* **1997**, *62*, 6678-6681.

28) NMR-analysis of the crude reaction mixture containing **5** did not indicate the formation of the other diastereoisomer.

29) Davies, S. G.; Nicholson, R. L.; Price, P. D.; Roberts, P. M.; Smith, A. D. *Synlett* **2004**, *5*, 901-903.

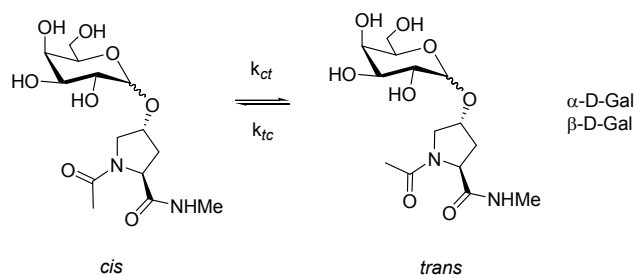
- 30) GOESY is a 1D selective NOESY experiment. It provides quantitative information and much higher sensitivity compared to NOESY'. First appeared in: Stonehouse, J.; Adell, P.; Keeler, J.; Shaka, A. J. *J. Am. Chem. Soc.* **1994**, *116*, 6037-6038.
- 31) A 40ms gaussian pulse with a 560 ms mixing time was used.
- 32) a) Taylor, C. M.; Hardre, R.; Edwards, P. J. B.; Park, J. H. *Org. Lett.* **2003**, *5*, 4413-4416; b) Madison, V.; Kopple, K. D. *J. Am. Chem. Soc.* **1980**, *102*, 4855-4863.
- 33) Trabocchi, A.; Potenza, D.; Guarna, A. *Eur. J. Org. Chem.* **2004**, *22*, 4621-4627.
- 34) Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc.* **1973**, *95*, 5725-5729.
- 35) Cai, M.; Huang, Y.; Liu, J.; Krishnamoorthi, R. *J. Biomolecular NMR.* **1995**, *6*, 123-128.
- 36) a) Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A.; Raines, R. T.; Krow, G. *R. J. Org. Chem.* **2004**, *69*, 8565-8573; b) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777-778.
- 37) a) Halab, L.; Lubell, W. D.; *J. Org. Chem.* **1999**, *64*, 3312-3321; b) Halab, L.; Lubell, W. D. *J. Peptide Science* **2001**, *7*, 92-104; c) Halab, L.; Lubell, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2474-2484.
- 38) Madison, V. *Biopolymers* **1977**, *16*, 2671-2692.
- 39) a) Montelione, G. T.; Hughes, P.; Clardy, J.; Scheraga, H. A. *J. Am. Chem. Soc.* **1986**, *108*, 6765-6773; b) Avenoza, A.; Busto, J. H.; Peregrina, J. M.; Rodriguez, F. J. *Org. Chem.* **2002**, *67*, 4241-4249; c) Hanessian, S.; Papeo, G. Angiolini, M. Fettes, K.;

Beretta, M.; Munro, A. *J. Org. Chem.* **2003**, *68*, 7204-7218; d) Taylor, C. M.; Hardre, R.; Edwards, P. J. B. *J. Org. Chem.* **2005**, *70*, 1306-1315; e) Koskinen, A. M. P.; Helaja, J.; Kumpulainen, E. T. T.; Koivisto, J.; Mansikkamaeki, H.; Rissanen, K. *J. Org. Chem.* **2005**, *70*, 6447-6453.

Chapter 4: Effects of Glycosylation of (2*S*,4*R*)-4-Hydroxyproline on the Conformation, Kinetics and Thermodynamics of Prolyl Amide Isomerization

Neil W. Owens,^{*,#§∞} Craig Braun,^{*,#§∞} Joe D. O'Neil,^{§∞} Kirk Marat,[#] and Frank Schweizer^{*,§∞}

^{*}Designed research; [#]Performed research; [§]Analyzed data; [∞]Wrote the paper



4.1 Abstract

*Glycosylation (galactosylation) of (2*S*,4*R*)-4-hydroxyproline (Hyp) in the peptide mimic *N*-acetyl-Hyp-*N'*-methylamide does not influence the isomer equilibrium nor the rate of isomerization between the *cis*- and *trans*-prolyl amides in water. However, glycosylation of Hyp results in a downfield shift (6.5–8 ppm) of the *C* γ atom of proline that is consistent with an enhanced inductive effect. Moreover, *nOe* experiments on the α - and β -linked glycosylated *Ac*-Hyp-NHMe amides indicate proximity between the galactose and pyrrolidine rings. This study strongly suggests that glycosylation of Hyp will have important implications on peptide backbone conformation.*

4.2 Introduction

Glycosylation is a common post-translational modification of proteins implicated in cellular recognition processes and controlling protein conformation.¹ Typically, carbohydrates are *O*-linked to serine (Ser) and threonine (Thr) or *N*-linked to asparagine (Asn). Glycosylation of (2*S*,4*R*)-4-hydroxyproline (Hyp) is widespread in the plant kingdom and occurs in Hyp-rich glycoproteins (HRGPs) that are associated with the cell walls of algae and flowering plants.² HRGPs are characterized by extensively glycosylated Hyp sequences that contain *O*-glycosidic linkages between the pyranose D-galactose and the furanose L-arabinose.² Although HRGPs are broadly implicated in many aspects of plant growth and development³ and cell wall stability,² no information is available about the structural and conformational implications of Hyp-glycosylation on peptide backbone conformation.

Hyp and proline (Pro) are unique among the proteinogenic amino acids since they are characterized by limited rotation of the ϕ dihedral angle (Figure 4.2.1) as their side chain is fused to the peptide backbone.

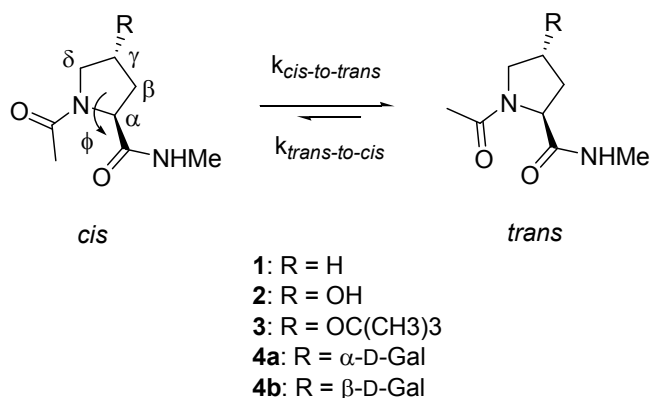


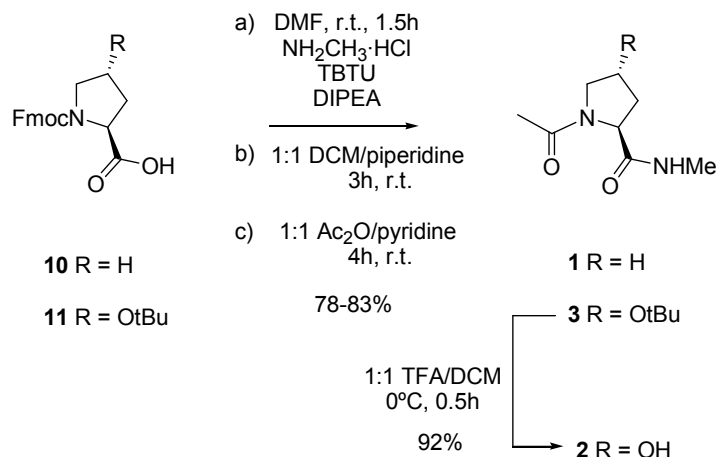
Figure 4.2.1: *Cis-trans* isomerization of reference diamides **1-3** and the galactosylated Hyp model amides **4a** and **4b**.

As a consequence, there is a reduction in the energy difference between the prolyl amide *cis*- and *trans*-isomers, making them nearly isoenergetic; this leads to higher *cis* *N*-terminal amide content relative to the other amino acids. Moreover, the isomerization of the prolyl amide bond has been shown to be the rate-determining step in the folding pathways of many peptides and proteins.⁴

4.3 Results

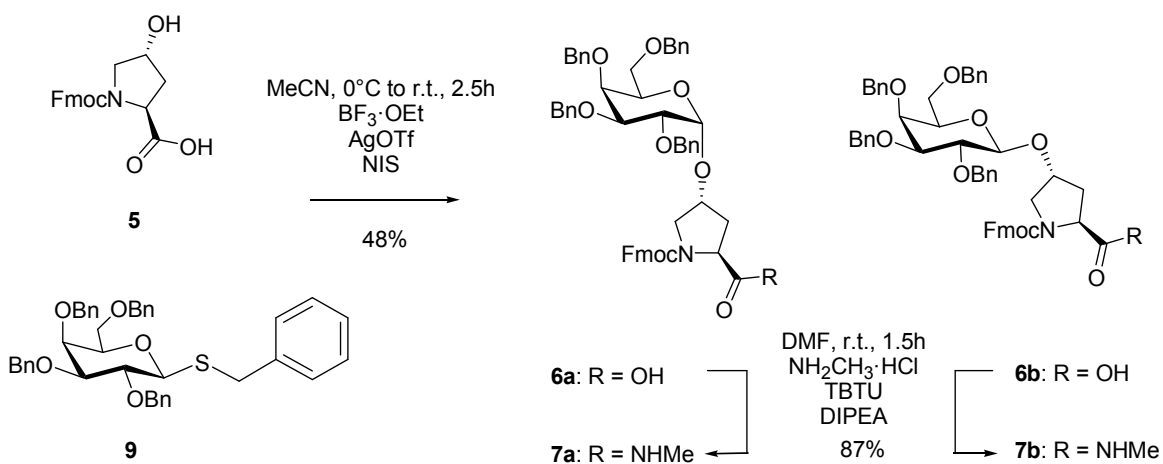
Herein we describe the effects of galactosylation of Hyp on the conformation as well as the thermodynamics and kinetics of prolyl *N*-terminal amide isomerization. Compounds **4a** Ac-Hyp(α -D-Gal)-NHMe and **4b** Ac-Hyp(β -D-Gal)-NHMe were selected as glycopeptide mimics, while Ac-Pro-NHMe **1**, Ac-Hyp-NHMe **2** and Ac-Hyp(*O*-*tert*-butyl)-NHMe **3** served as non-glycosylated reference compounds (Figure 4.2.1).

The synthesis of **1** was carried out by coupling Fmoc-Hyp-OH **10** to methylamine hydrochloride using TBTU in the presence of *N,N*-diisopropylethylamine, followed by Fmoc-deprotection under basic conditions, and *N*-terminal acylation using acetic anhydride in pyridine (Scheme 4.3.1). The synthesis of **3** was carried out in a similar fashion starting from commercially available Fmoc-Hyp(*O*-*tert*-butyl)-OH **11**. The model peptide Ac-Hyp-NHMe **2** was synthesized in 92% yield by *O*-*tert*-butyl deprotection of **3** under acidic conditions.



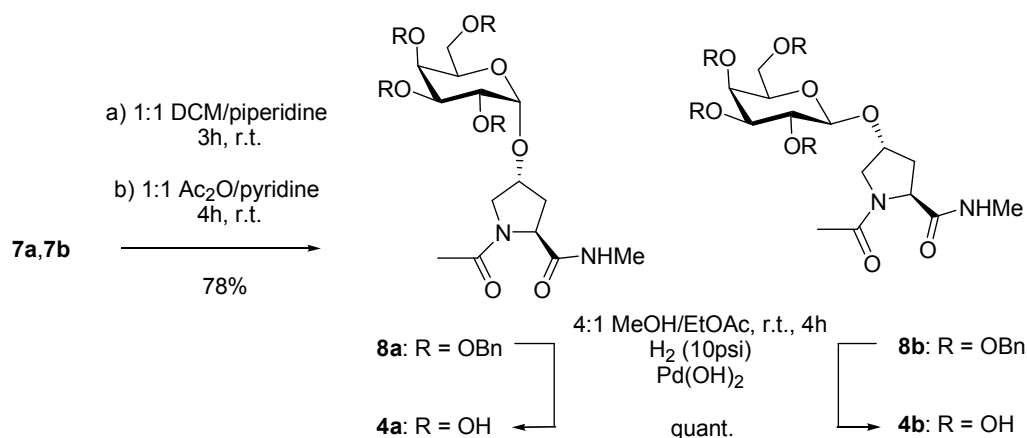
Scheme 4.3.1: Synthesis of diamides **1**, **2** and **3**.

The synthesis of **4a** and **4b** began was carried out by glycosylation of Fmoc-Hyp-OH **5** using 2,3,4,6-tetra-*O*-benzyl-thio- β -D-galactopyranoside¹⁶ **9** in the presence of boron trifluoride diethyletherate, *N*-iodosuccinimide and catalytic amounts of silver triflate to yield a mixture of diastereomers **6a** and **6b** in 48% yield (Scheme 4.3.2). Separation of the anomers could be carried out after *C*-terminal amidation using TBTU in the presence of *N,N*-diisopropylethylamine to yield nearly a 1:1 mixture of **7a** and **7b** in 87% yield.



Scheme 4.3.2: Synthesis of the glycosylated hydroxyproline building blocks **7a** and **7b**.

Installation of the *N*-acetyl group was carried out by *N*-Fmoc deprotection using piperidine in dichloromethane followed by *N*-acylation using acetic anhydride in pyridine to give **8a** and **8b** in 78% yield (Scheme 4.3.3). Removal of the benzyl ether protecting groups by catalytic hydrogenation in methanol/ethyl acetate gave the model amides **4a** and **4b** in 17% and 16% overall yield from **5**, respectively.



Scheme 4.3.3: Synthesis of the diamides **4a** and **4b**.

The *trans* rotamers in compounds **1-4b** were assigned on the basis of higher C- δ NMR chemical shifts relative to the *cis* rotamer⁵ and nOe transfer between H- δ of proline to the *N*-acyl protons in selective 1D GOESY experiments.⁶ The relative amounts of *cis* and *trans* isomers were determined by integrating and averaging as many distinct proton signals as possible for both the major and minor isomers in the ¹H NMR spectra.⁷ In D₂O at 37 °C, the *trans/cis* isomer ratio equilibrium constants ($K_{trans/cis}$) for **4a** (3.41 ± 0.30) and **4b** (3.37 ± 0.28) are nearly identical to those of **2** (3.52 ± 0.05) and **3** (3.34 ± 0.15) and the observed differences are within the experimental errors (Table 4.3.2).

Table 4.3.1: Rates of prolyl amide isomerization for **1-4b**.

Amide	$k_{cis-to-trans}^{[a]}$ (sec ⁻¹)	$k_{trans-to-cis}^{[b]}$ (sec ⁻¹)
1	0.81 ± 0.02	0.31 ± 0.02
2	0.73 ± 0.01	0.25 ± 0.01
3	0.77 ± 0.02	0.27 ± 0.01
4a	0.83 ± 0.05 (0.82 ± 0.06)	0.27 ± 0.02 (0.27 ± 0.03)
4b	0.61 ± 0.04 (0.57 ± 0.05)	0.21 ± 0.02 (0.19 ± 0.02)

^[a]Carried out in D₂O at 67°C; ^[b]Calculated from k_{ct} and equilibrium ($K_{t/c}$); ^[c]Carried out in phosphate buffer pH = 7.4 at 67 °C.

The kinetics of *cis/trans* isomerization for compounds **1-4b** were determined by ¹H-NMR spectroscopy inversion transfer experiments⁸ in D₂O at elevated temperature. At 67°C, the *cis-to-trans* rate constant of isomerization (k_{ct}) of the α -glycosylated Hyp model peptide **4a** ($k_{ct} = 0.83 \text{ sec}^{-1}$) is very similar to that of the hydroxyproline model peptide **2** (k_{ct} of 0.73 sec^{-1}) and **3** ($k_{ct} = 0.77 \text{ sec}^{-1}$), while the β -anomer **4b** gave slightly lower rates ($k_{ct} = 0.61 \text{ sec}^{-1}$). A similar trend was observed for the *trans-to-cis* rate constants of compounds **1-4b** (Table 4.3.1). At physiological temperature the rates are too slow to be differentiated by this assay.

The effects of temperature on k_{ct} and k_{tc} were analyzed by Eyring plots⁹ (Figures 4.3.1 and 10.7-10.11) and values for ΔH^\ddagger and ΔS^\ddagger were calculated from linear least squares fits of the data in these plots and are presented in Table 4.3.2. The activation parameters demonstrate that the free energy barriers to isomerization of compounds **1-4b** are enthalpic in origin. The effects of temperature on the values of $K_{trans/cis} = (k_{ct}/k_{tc})$ were measured directly by NMR spectroscopy and the resulting data were analyzed by van't

Hoff plots (Figure 4.3.2). Values for ΔH° and ΔS° were calculated from linear least-squares fits of these plots (Table 4.3.3). In all cases studied, the *trans* isomer of **1-4b** is more stable than the *cis* isomer. Moreover, the values of $K_{t/c}$ for **1-4b** are dependent on temperature such that the *trans* isomer becomes increasingly favoured as the temperature decreases.

Table 4.3.2: Activation enthalpies (ΔH^\ddagger) and entropies (ΔS^\ddagger) as derived from Eyring plots in D₂O for **1-4b**.

Amide	<i>cis-to-trans</i> ^[a]			<i>trans-to-cis</i> ^[a]		
	ΔH^\ddagger (kcal·mol ⁻¹)	ΔS^\ddagger (cal·mol ⁻¹ ·K ⁻¹)	ΔG^\ddagger_{300K} ^[b] (kcal·mol ⁻¹)	ΔH^\ddagger (kcal·mol ⁻¹)	ΔS^\ddagger (cal·mol ⁻¹ ·K ⁻¹)	ΔG^\ddagger_{300K} ^[b] (kcal·mol ⁻¹)
1	20.6	1.2	20.2	21.3	1.4	20.9
2	20.2	0.1	20.2	21.1	0.5	21.0
3	21.8	4.7	20.4	22.7	5.4	21.1
4a	20.4	0.6	20.2	20.6	1.0	20.3
4a	22.4	6.2	20.6	23.3	6.6	21.3

^[a]Error limits obtained from the residuals of the linear least squares fitting of the data to equation $\ln(k/T) = (-\Delta H^\ddagger/R)(1/T) + \Delta S^\ddagger/R + \ln(k_B/h)$ were 1-2% for compounds **1**, **2** and 3-6% for compounds **3**, **4a**, and **4b**; where k_B is the Boltzmann constant and h is Planck's constant; ^[b]Calculated using the equation $\Delta G^\ddagger = \Delta H^\ddagger - T\Delta S^\ddagger$.

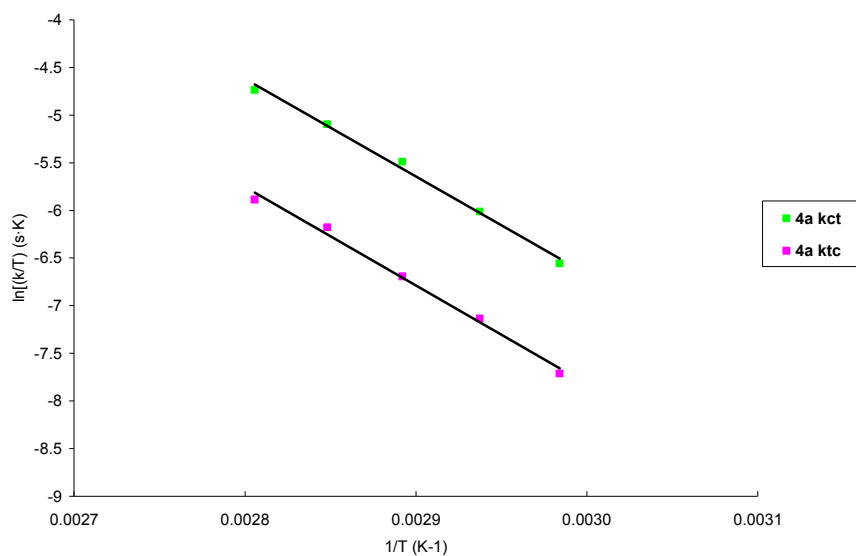


Figure 4.3.1: Representative Eyring plot of the rates of *cis-trans* isomerization (*kct*) (green squares) and *trans-cis* isomerization (*ktc*) (pink squares) relative to temperature for **4a**, with best fit lines (black line).

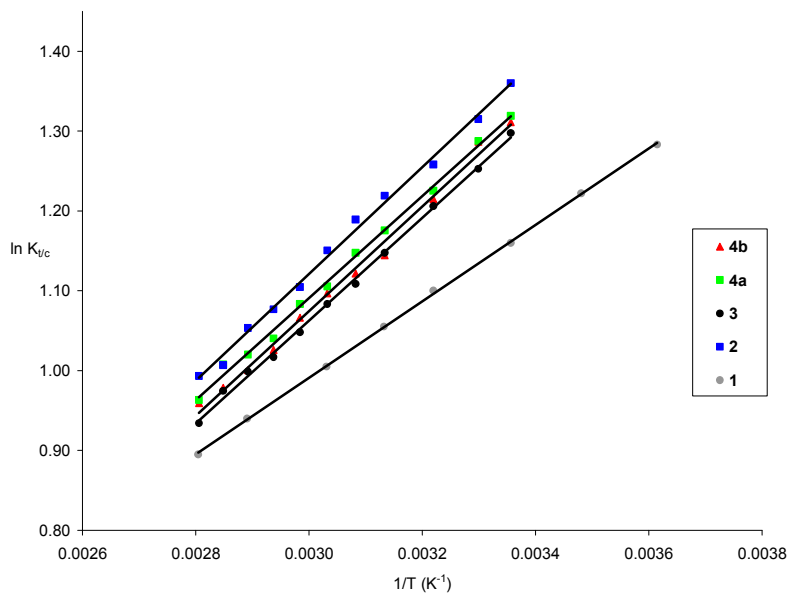


Figure 4.3.2: Comparison of van't Hoff plots of the model diamides **1** (grey circles), **2** (blue squares), **3** (black circles), **4a** (green squares), **4b** (red triangles), with best fit lines (black line).

Table 4.3.3: Thermodynamic parameters for *cis-trans* isomerization of **1-4b**.

Amide	$\Delta H^{\circ[a]}$ (kcal·mol ⁻¹)	$\Delta S^{\circ[a]}$ (cal·mol ⁻¹ ·K ⁻¹)	$K_{trans/cis}^{[b]}$ (37 °C)
1	-0.95 ± 0.01	-0.87 ± 0.02	3.03 ± 0.08
2	-1.33 ± 0.03	-1.76 ± 0.09	3.52 ± 0.05
3	-1.29 ± 0.02	-1.72 ± 0.06	3.34 ± 0.15
4a	-1.27 ± 0.02	-1.64 ± 0.07	3.41 ± 0.30
4a	-1.30 ± 0.03	-1.77 ± 0.10	3.37 ± 0.28

^[a]Error limits obtained by linear least-squares fitting the data of the van't Hoff plots to equation $\ln K_{t/c} = (-\Delta H^{\circ}/R)(1/T) + \Delta S^{\circ}/R$; ^[b]Carried out in D₂O; ± SE determined by integration of the ¹H NMR signals from two or more sets of *trans/cis* isomers.

The pucker of Hyp (2*S*,4*R*) in model peptide **2** in solution has been previously assigned to the C^γ-exo conformation on the basis of observed ³*J*-coupling constants.¹⁰ The prolyl ring pucker for compounds **3**, **4a**, and **4b** were similarly established as the C^γ-exo conformation on the basis of ¹H NMR coupling constants by comparison to literature values. For example, in compound **4a** we observed both ³*J*_{α,β1} and ³*J*_{α,β2} couplings constants to be 8.2 Hz. Based on the dihedral angles of respective puckers, expected coupling constants for the C^γ-exo conformer are 7-10 Hz and 7-11 Hz, respectively, whereas those for C^γ-endo are 6-10 Hz and 2-3 Hz, respectively, due to the Karplus equation. Other couplings are characteristic for the C^γ-exo pucker (see Table 10.2).

Previous reports have shown that inductive effects in the γ-position of proline exert a significant influence on the thermodynamics and kinetics of prolyl amide bond

isomerization.¹¹ To assess the inductive effect caused by glycosylation of hydroxyproline we determined the ¹³C-NMR chemical shifts of the C^γ atom, which can indicate electron withdrawal by pendant functional groups,¹² and have previously been used to correlate the electron-withdrawing effects in various C^γ-substituted proline analogs.^{11c} The observed ¹³C-NMR chemical shifts ($\delta_{C\gamma(\text{trans})}$) indicate that electron withdrawal increases in the order hydroxyl ($\delta_{C\gamma} = 69.9$ ppm) (**2**) < *tert*-butoxyl ($\delta_{C\gamma} = 70.1$ ppm) (**3**) < α -galactosyl ($\delta_{C\gamma} = 76.5$ ppm) (**4a**) < β -galactosyl ($\delta_{C\gamma} = 77.9$ ppm) (**4b**) (see Table 10.1). It should be noted that chemical shift can be affected by other factors, including anisotropy, steric effects, and Van der Waals effects. However, anisotropy has little influence on ¹³C chemical shift, and if steric effects play a role, then the *tert*-butyl group of **3** should have a similar greater influence on the ¹³C chemical shift than **1** and **2**, which is not observed.

To determine the extent of interaction between the galactose and prolyl rings, we performed nOe transfer experiments with compounds **4a** and **4b** in D₂O. Selective inversion of one of the H- β protons in **4a** by a selective GOESY⁶ experiment resulted in a 1.5% resonance transfer to a peak at $\delta = 3.83$ ppm corresponding to the overlapped signals of both H-4 and H-5 of galactose (Figure 4.3.3A). By comparison, selective inversion of H-2 in **4b** produced nOe transfer (3.0%) to H- α of Hyp (Figure 4.3.3B). These results suggest that galactosylation of Hyp induces close contacts between distant positions in the carbohydrate and pyrrolidine rings.

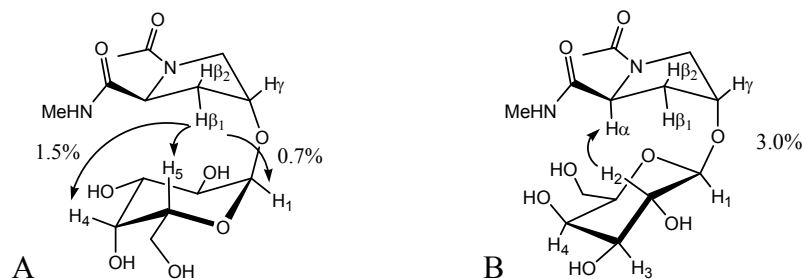


Figure 4.3.3: Relevant nOe interactions between the prolyl ring and the carbohydrate ring for A) Ac-Hyp(α -D-Gal)-NHMe (**4a**) and B) Ac-Hyp(β -D-Gal)-NHMe (**4b**).

4.4 Discussion

We have found that glycosylation of Hyp in compounds **4a** and **4b** has no apparent effect on the isomer equilibrium ($K_{trans/cis}$) nor on the rate of isomerization (k_{tc} , k_{ct}) in water between the *cis* and *trans* isomers when compared to unglycosylated reference compounds **1-3**. However, glycosylation of Hyp provides an inductive electron withdrawing effect on the prolyl ring. The magnitude of the change in ^{13}C -NMR chemical shift (6.5-8 ppm) has been correlated to the strength of the inductive effect, and is similar to a downfield shift of 8 ppm observed when a trifluoroacetate group was attached to Hyp in Ac-Hyp-OMe.^{11c} The effect is slightly larger for Ac-Hyp(β -D-Gal)-NHMe **4b** (by 1.5 ppm) compared to the α -anomer **4a**. It is known that (4*R*)-electronegative substituents stabilize the C^γ -exo pucker of proline in peptide mimics¹¹ and contribute to enhanced stability of the triple helix in collagen.¹³ Moreover, it has been established both from computational¹⁴ and experimental studies¹⁵ that stabilization of the C^γ -exo ring pucker favours the prolyl *trans* amide isomer. Our results therefore suggest that glycosylation of Hyp may lead to additional stabilization of the C^γ -exo ring pucker of

Hyp. However, this stabilization does not translate in a measurable increase on $K_{trans/cis}$ for peptide mimics **4a** and **4b** when compared to unglycosylated **2**. Most likely the stabilization of the *trans* isomer in compounds **4a** and **4b** is too small to be differentiated and remains within the experimental error. Any stabilization of the C^γ -exo ring pucker may be more apparent in larger glycopeptides that contain two or more glycosylated Hyp units; where the effect may be additive.

Perhaps more importantly, nOe experiments show that the glycosylation of **2** results in distant contacts between the proline and galactose rings, suggesting that galactosylation of Hyp induces conformational constraint into glycopeptides. Very likely this is the result of increased steric strain induced upon glycosylation. Additionally, glycosylation of Hyp may also affect other properties including solvation, solubility and thermostability.

4.5 Conclusions

While there is no significant influence on prolyl *N*-terminal amide isomerization, the presence of both an enhanced inductive effect and Gal-Pro contacts between distant positions in **4a** and **4b** suggests that glycosylation of Hyp will have important implications on peptide backbone conformation in HRGPs and glycosylated Hyp-containing peptides.

4.6 Experimental Section

General

Thin-Layer chromatography was performed on Si250F precoated plates of silica gel (250 μ m). Column chromatography was performed on SilicaFlash P60 silica gel (40-63 μ m). Reagent grade solvents were used without further purification. Spectra were assigned based on 2D COSY and 2D HSQC experiments. For ^1H NMR, minor isomers are listed between square brackets. For ^{13}C NMR, when assigned, carbon peaks for the minor isomer are listed in brackets.

General Procedures

Preparation of *N*-Fmoc-*N'*-methylamide amino acids

The amino acid (1.0 eq) was dissolved in anhydrous DMF, followed by addition of TBTU (1.5 eq), diisopropylethylamine (3.0 eq) and methylamine hydrochloride (1.5 eq). The solution was stirred at ambient temperature for 1.5 hrs. The reaction mixture was diluted with 20 mL water, followed by extraction (EtOAc), drying (Na_2SO_4), and concentration under reduced pressure for purification by flash chromatography.

Preparation of *N*-Acetyl-*N'*-methylamide amino acids

The amino acid (1.0 eq) was dissolved in 1:1 (v/v) DCM/piperidine, stirred for 3 hours at ambient temperature, then co-distilled with toluene (3 x 10 mL). The crude product was then dissolved in 1:1 Ac_2O /pyridine, and the mixture was stirred at ambient temperature for 4 hours. The solution was concentrated under reduced pressure, and then was co-distilled with toluene (3 x 10 mL) before purification by flash chromatography.

***N*-Acetyl-*L*-proline *N'*-methylamide (1)**: The general preparation methods were followed: *C*-terminal amidation of (2*S*)-Fmoc-Proline **10** (0.2 g, 1.175 mmol, 1.0 eq) using TBTU (0.566 g, 1.76 mmol, 1.5 eq), DIPEA (0.614 mL, 3.52 mmol, 3.0 eq), and methylamine hydrochloride (0.119 g, 1.76 mmol, 1.5 eq) in anhydrous DMF (5 mL). *N*-terminal acylation of the crude product was performed using 1:1 (v/v) DCM:piperidine (4 mL), then 1:1 Ac₂O/pyridine (4 mL). The reaction mixture was purified by flash chromatography using first ethyl acetate, then 9:1 ethyl acetate/methanol to yield **1** as a clear oil (0.181 g, 1.063 mmol) (83.0%): $[\alpha]_D^{25} = -1.4^\circ$ (*c* 1.0 CHCl₃); ¹H NMR (300 MHz, D₂O, 298K) $\delta = [4.46, \text{dd}, 0.25\text{H}, \text{Pro}_\alpha^{\text{cis}}, J = 3.0\text{Hz}, J = 8.7\text{Hz}], 4.30 (\text{dd}, 0.75\text{H}, \text{Pro}_\alpha^{\text{cis}}, J = 4.1\text{Hz}, J = 8.8\text{Hz}), 3.52\text{-}3.70 (\text{m}, 1.5\text{H}, \text{Pro}_{\gamma 1}^{\text{trans}}, \text{Pro}_{\gamma 2}^{\text{trans}}), [3.39\text{-}3.54, \text{m}, 0.5\text{H}, \text{Pro}_{\gamma 1}^{\text{cis}}, \text{Pro}_{\gamma 2}^{\text{cis}}], [2.75, \text{s}, 0.75\text{H}, \text{-NHCH}_3^{\text{cis}}], 2.70 (\text{s}, 2.25\text{H}, \text{-NHCH}_3^{\text{trans}}), [2.26\text{-}2.35, \text{m}, 0.25\text{H}, \text{Pro}_{\beta 1}^{\text{cis}}], 2.15\text{-}2.27 (\text{m}, 0.75\text{H}, \text{Pro}_{\beta 1}^{\text{trans}}), 2.09 (\text{s}, 2.25\text{H}, \text{-COCH}_3^{\text{trans}}), [1.95, \text{s}, 0.75\text{H}, \text{-COCH}_3^{\text{cis}}], 1.85\text{-}2.01 (\text{m}, 3\text{H}, \text{Pro}_{\beta 1}^{\text{trans}}, \text{Pro}_{\beta 1}^{\text{cis}}, \text{Pro}_{\delta 1}^{\text{trans}}, \text{Pro}_{\delta 1}^{\text{cis}}, \text{Pro}_{\delta 2}^{\text{trans}}, \text{Pro}_{\delta 2}^{\text{cis}}]$; ¹³C NMR (75 MHz, D₂O, 298K) $\delta = 175.3, (175.2), (173.9), 173.7, (62.3), 60.8, 49.1, (47.4), (31.9), 30.3, (26.3), 26.2, 24.5, (22.9), 21.8, (21.6)$.

***N*-Acetyl-*trans*-4-hydroxy-*L*-proline *N'*-methylamide (2)**: Compound **3** (0.095 g, 0.392 mmol) was dissolved in 1:1 (v/v) TFA/DCM (4 mL), stirred for 30 minutes at 0°C, then co-distilled with toluene (3 x 6 mL). The product was purified by flash chromatography using 10:1 ethyl acetate/methanol to yield **2** as a clear oil (0.067 g, 0.36 mmol) (92.0%): $[\alpha]_D^{25} = -26.7^\circ$ (*c* 0.3 CH₃OH); ¹H NMR (300 MHz, D₂O, 298K) $\delta = [4.57, \text{dd}, 0.23\text{H}, \text{Pro}_\alpha^{\text{cis}}, J_{2,3a} = 8.1\text{Hz}], 4.50\text{-}4.56 (\text{m}, 0.77\text{H}, \text{Pro}_\gamma^{\text{trans}}), [4.45\text{-}4.50, \text{m}, 0.23\text{H}, \text{Pro}_\gamma^{\text{cis}}], 4.40 (\text{dd}, 0.77\text{H}, \text{Pro}_\alpha^{\text{trans}}, J_{2,3a} = 8.4\text{Hz}, J_{2,3b} = 8.5\text{Hz}), 3.79 (\text{dd}, 0.77\text{H}, \text{Pro}_{\delta 1}^{\text{trans}}, J_{\gamma,\delta 1} = 4.1\text{Hz}, J_{\delta 1,\delta 2} = 11.8\text{Hz}), [3.69, \text{dd}, 0.23\text{H}, \text{Pro}_{\delta 1}^{\text{cis}}, J_{\gamma,\delta 1} = 3.9\text{Hz}, J_{\delta 1,\delta 2} =$

12.7Hz], 3.61 (dd, 0.77H, Pro $_{\delta 2}^{trans}$, $J_{\gamma, \delta 2} = 3.5\text{Hz}$), [3.53, dd, 0.23H, Pro $_{\delta 2}^{cis}$, $J_{\gamma, \delta 2} = 4.5\text{Hz}$], [2.75, s, 0.69H, -NHCH $_3^{cis}$], 2.71 (s, 2.31H, -NHCH $_3^{trans}$), [2.35-2.48, m, 0.23H, Pro $_{\beta 1}^{cis}$], 2.29 (ddd, 0.77H, Pro $_{\beta 1}^{trans}$, $J_{\beta 1, \gamma} = 4.4\text{Hz}$, $J_{\beta 1, \beta 2} = 13.7\text{Hz}$), [2.17, 0.23H, Pro $_{\beta 2}^{cis}$, $J_{\beta 2, \gamma} = 4.7\text{Hz}$, $J_{\alpha, \beta 2} = 7.9\text{Hz}$, $J_{\beta 1, \beta 2} = 13.7\text{Hz}$], 2.09 (s, 2.31H, -COCH $_3^{trans}$), 2.17 (0.77H, Pro $_{\beta 2}^{trans}$, $J_{\beta 2, \gamma} = 4.5\text{Hz}$), [1.96, s, 0.69H, -COCH $_3^{cis}$]; ^{13}C NMR (75 MHz, D $_2\text{O}$, 298K) $\delta = 174.7$, (174.6), (174.3), 173.8, 69.9, (68.6), (60.5), 59.3, 56.5, (54.7), (39.6), 38.0, (26.3), 26.2, 21.9, (21.0).

4-*O*-(*tert*-Butyl)-*N*-acetyl-*trans*-4-hydroxy-*L*-proline *N'*-methylamide (3): The general preparation methods were followed: *C*-terminal amidation of (2*S*,4*R*)-Fmoc-*tert*-butylhydroxyproline **11** (0.263 g, 0.642 mmol, 1.0 eq) using TBTU (0.205 g, 0.963 mmol, 1.5 eq), DIPEA (0.336 mL, 1.926 mmol, 3.0 eq), and methylamine hydrochloride (0.119 g, 1.76 mmol, 1.5 eq) in anhydrous DMF (5 mL). *N*-terminal acylation of the crude product was performed using 1:1 DCM:piperidine (6 mL), then 1:1 Ac $_2\text{O}$ /pyridine (6 mL). The reaction mixture was purified by flash chromatography using 40:1 ethyl acetate/methanol to yield **3** as a clear oil (0.050 g, 0.206 mmol) (78.0%): $[\alpha]_D^{25} = -21.6^\circ$ (*c* 0.9 CHCl $_3$); ^1H NMR (300 MHz, D $_2\text{O}$, 298K) $\delta =$ [4.55, dd, 0.23H, Pro $_{\alpha}^{cis}$, $J = 5.4\text{Hz}$, $J = 8.4\text{Hz}$], 4.42-4.52 (m, 0.77H, Pro $_{\gamma}^{trans}$), [4.42-4.48, m, 0.23H, Pro $_{\gamma}^{cis}$], 4.39 (dd, 0.77H, Pro $_{\alpha}^{trans}$, $J = 6.5\text{Hz}$, $J = 8.3\text{Hz}$), 3.88 (dd, 0.77H, Pro $_{\delta 1}^{trans}$, $J_{\gamma, \delta 1} = 5.8\text{Hz}$, $J_{\delta 1, \delta 2} = 11.1\text{Hz}$), [3.68, dd, 0.23H, Pro $_{\delta 1}^{cis}$, $J_{\gamma, \delta 1} = 6.0\text{Hz}$, $J_{\delta 1, \delta 2} = 12.3\text{Hz}$], 3.43 (dd, 0.77H, Pro $_{\delta 2}^{trans}$, $J_{\gamma, \delta 2} = 4.6\text{Hz}$), [3.41, dd, Pro $_{\delta 2}^{cis}$, $J_{\gamma, \delta 2} = 4.8\text{Hz}$], [2.75, s, 0.69H, -NHCH $_3$], 2.70 (s, 2.31H, -NHCH $_3$), 2.09-2.37 (m, 2H, Pro $_{\beta 1}^{trans}$, Pro $_{\beta 1}^{cis}$, Pro $_{\beta 2}^{trans}$, Pro $_{\beta 2}^{cis}$), 2.07 (s, 2.31H, -COCH $_3^{trans}$), [1.94, s, 0.69H, -COCH $_3^{cis}$], 1.21 (s, 6.93H, *tert*-butyl trans), [1.20, s, 2.07H, *tert*-butyl cis]; ^{13}C NMR (75 MHz, D $_2\text{O}$, 298K) $\delta = 174.7$, (174.5), (174.0), 173.6, 76.6, (76.5), 70.1,

(68.9), (60.6), 59.2, 55.4, (53.7), (39.1), 37.6, 27.6, (26.3), 26.2, 21.9, (20.9); HRMS (ES) calc. for $C_{12}H_{22}N_2NaO_3$ ($M + Na$)⁺: 265.1528. Found ($M + Na$)⁺: 265.1530.

4-*O*-(2,3,4,6-Tetra-*O*-benzyl- α,β -D-galactopyranosyl)-*N*-fluorenylmethoxycarbonyl-*trans*-4-hydroxy-*L*-proline *N'*-methanamide (7a,b**):** (2*S*,4*R*)-Fmoc-hydroxyproline **5** (1.092 g, 3.08 mmol, 1.0 eq) was dissolved in MeCN (10 mL), followed by addition of 2,3,4,6-tetra-*O*-benzyl-thio- β -D-galactopyranoside **9** (2.398 g, 3.70 mmol, 1.2 eq) and $BF_3 \cdot OEt$ (0.389 mL, 3.08 mmol, 1.0 eq). The solution was cooled to 0°C and stirred for ten minutes. Then, AgOTf (0.154 g, 0.617 mmol) and NIS (0.973 g, 4.32 mmol, 1.4 eq) were added, at which point the solution turned red. The mixture was allowed to return to ambient temperature, and stirred for an additional 2.5 hours before being concentrated under reduced pressure. The resulting residue was re-dissolved in ethyl acetate (20 mL). This solution was washed with 1:1 saturated sodium bicarbonate (aq)/saturated sodium thiosulphate (aq) (5 x 20 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The product was purified by flash chromatography using 4:4:1 CH_2Cl_2 /ethyl acetate/methanol to yield a mixture of diastereomers **6a,b** recovered as a white solid (1.284 g, 1.466 mmol) (48.0%).

The general procedure for *C*-terminal amidation was carried out on the mixture of diastereomers **6a,b** (0.200 g, 0.228 mmol, 1.0 eq) using TBTU (0.109 g, 0.342 mmol, 1.5 eq), DIPEA (0.120 mL, 0.684 mmol, 2.0 eq), and methylamine hydrochloride (0.022 g, 0.342 mmol, 1.0 eq) in anhydrous DMF (5 mL). The products were purified by flash chromatography using 2:1 ethyl acetate/hexanes to yield first **7a** (0.084 g, 0.098 mmol) and then **7b** (0.088 g, 0.100 mmol) (87.0%):

7a: $[\alpha]_D^{25} = +14.2^\circ$ (*c* 2.0 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.72$ -7.81 (m, 2H, aromatic), 7.47-7.64 (m, 2H, aromatic), 7.18-7.46 (m, 24H, aromatic), 6.31 (broad q, 0.65H, -NHCH₃), [5.42, broad q, 0.35H, -NHCH₃], 4.96 (d, 1H, -OCH₂Ph, *J* = 11.6Hz), 4.91 (broad d, 1H, H₁), 4.67-4.87 (m, 3H, -OCH₂Ph), 4.62 (d, 1H, -OCH₂Ph, *J* = 12.3Hz), 4.58 (d, 1H, -OCH₂Ph, *J* = 11.6Hz), 4.11-4.53 (m, 7H, -OCH₂Ph, H₄, H₅, H_{6a}, H_{6b}, Pro _{γ} , Pro _{α}), 4.05 (dd, 1H, H₂, *J* = 3.6Hz, *J* = 10.1Hz), 3.94-4.01 (m, 2H, -C(O)OCH₂CH), 3.90 (broad dd, 1H, H₃), 3.44-3.84 (m, 4H, Pro _{δ 1}, Pro _{δ 2}, -C(O)OCH₂CH), 2.73 (d, 2H, -NHCH₃^{*trans*}, *J* = 4.3Hz), [2.58, broad d, 1H, -NHCH₃^{*cis*}], 2.42-2.54 (m, 1H, Pro _{β 1}), 2.26-2.40 (m, 1H, Pro _{β 2}); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta = 171.6$, 155.9, 143.9, 143.7, 141.3, 138.8, 138.6, 137.9, 127.0-128.6 (aromatic carbons), 125.1, 120.0, 97.5, 78.8, 76.4, 75.0, 74.8, 73.6, 73.3, 73.1, 70.0, 69.5, 67.7, 67.3, 59.4, 51.8, 47.2, 35.0, 26.3; HRMS (ES) calc. for C₅₅H₅₆N₂NaO₉ (M + Na)⁺: 911.3884. Found (M + Na)⁺: 911.3887.

7b: $[\alpha]_D^{25} = -23.7^\circ$ (*c* 1.4 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.70$ -7.81 (m, 2H, aromatic), 7.47-7.64 (m, 2H, aromatic), 7.14-7.45 (m, 24H, aromatic), 6.28 (broad q, 0.75H, -NHCH₃), [5.37, broad q, 0.25H, -NHCH₃], 4.95 (d, 1H, -OCH₂Ph, *J* = 11.7Hz), 4.68-4.86 (m apparent d, 4H, -OCH₂Ph), 4.58 (d, 1H, -OCH₂Ph, *J* = 11.7Hz), 4.50-4.58 (m, 1H, Pro _{γ}), 3.98-4.50 (m, 7H, -OCH₂Ph, H₁, H₅, H_{6a}, H_{6b}, Pro _{α}), 3.87-4.96 (m, 1H, -C(O)OCH₂CH), 3.73-3.86 (m, 2H, H₃, Pro _{δ 1}), 3.45-3.72 (m, 5H, H₂, H₄, Pro _{δ 2}, -C(O)OCH₂CH), 2.75 (d, 2.25H, -NHCH₃^{*trans*}, *J* = 4.7Hz), [2.61, broad d, 0.75H, -NHCH₃^{*cis*}], 2.41-2.55 (m, 1H, Pro _{β 1}), 2.08-2.30 (m, 1H, Pro _{β 2}); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta = 171.5$, 156.1, 143.8, 143.7, 141.3, 138.7, 138.5, 138.4, 137.8, 127.4-128.6 (aromatic carbons), 127.2, 125.2, 120.0, 102.4, 82.2, 79.2, 75.2, 74.7, 73.5, 73.4,

73.3, 72.9, 68.6, 67.8, 58.9, 52.9, 47.1, 33.8, 26.4; HRMS (ES) calc. for $C_{55}H_{56}N_2NaO_9$ ($M + Na$)⁺: 911.3884. Found ($M + Na$)⁺: 911.3880.

4-O-(2,3,4,6-Tetra-O-benzyl- α,β -D-galactopyranosyl)-N-acetyl-*trans*-4-hydroxy-L-proline N'-methylamide (8a,b): The general preparation method was followed for N-terminal acylation of **7a** (0.080 g, 0.090 mmol) using 1:1 DCM:piperidine (4 mL), then 1:1 Ac₂O/pyridine (4 mL). The reaction mixture was purified by flash chromatography using 10:1 ethyl acetate/methanol to yield **8a** as a clear oil (0.050 g, 0.071 mmol) (78.0%) (a similar procedure was used for **7b**):

8a: $[\alpha]_D^{25} = +0.84^\circ$ (*c* 1.0 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.22$ -7.40 (m, 20H), 6.76 (broad q, 1H, -NHCH₃, *J* = 4.6Hz), 4.94 (d, 1H, -OCH₂Ph, *J* = 11.4Hz), 4.71-4.88 (m, 4H, H₁, -OCH₂Ph), 4.52-4.67 (m, 3H, Pro _{γ} , -OCH₂Ph), 4.35-4.52 (m, 3H, Pro _{α} , -OCH₂Ph), 4.04 (dd, 1H, H₂, *J* = 3.6Hz, *J* = 10.1Hz), 3.93-4.00 (m, 2H, H₄, H₅), 3.88 (dd, 1H, H₃, *J* = 2.2Hz, *J* = 10.1Hz), 3.65 (dd, 1H, Pro _{δ 1}, *J* = 5.5Hz, *J* = 10.5Hz), 3.56-3.45 (m, 3H, H_{6a}, H_{6b}, Pro _{δ 2}), 2.72 (d, 3H, -NHCH₃, *J* = 4.6Hz), 2.58-2.62 (m, 1H, Pro _{β 1}), 2.11-2.15 (m, 1H, Pro _{β 2}), 1.99 (s, 3H, -COCH₃); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta = 171.3$, 170.8, 138.6, 138.5 (2), 137.9, 127.3-128.6 (aromatic carbons), 98.1, 78.9, 76.9, 76.4, 74.9, 74.8, 73.8, 73.6, 72.9, 70.1, 69.2, 58.4, 53.2, 33.9, 26.2, 22.5; HRMS (ES) calc. for $C_{42}H_{48}N_2NaO_8$ ($M + Na$)⁺: 731.3308. Found ($M + Na$)⁺: 731.3312.

8b: $[\alpha]_D^{25} = -0.11^\circ$ (*c* 1.0 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.20$ -7.40 (m, 20H), 6.89 (broad q, 1H, -NHCH₃, *J* = 4.7Hz), 4.94 (d, 1H, -OCH₂Ph, *J* = 11.7Hz), 4.67-4.84 (m, 4H, -OCH₂Ph), 4.62 (d, 1H, -OCH₂Ph, *J* = 11.6Hz), 4.48-4.58 (m, 2H, Pro _{α} , Pro _{γ}), 4.36-4.47 (m, 3H, H₁, -OCH₂Ph), 3.87 (dd apparent d, 1H, H₄, *J* = 3.0Hz, *J* =

3.0Hz), 3.77 (dd, 1H, H₂, J = 7.8Hz, J = 9.7Hz), 3.60-3.71 (m, 2H, Pro_{δ1}, Pro_{δ2}), 3.47-3.60 (m, 4H, H₃, H₅, H_{6a}, H_{6b}), 2.74 (d, 3H, -NHCH₃, J = 4.6Hz), 2.57-2.68 (m, 1H, Pro_{β1}), 1.97-2.08 (m, 1H, Pro_{β2}), 2.04 (s, 3H, -COCH₃); ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 171.3, 170.9, 138.6, 138.5, 138.4, 137.8, 127.3-128.7 (aromatic carbons), 103.1, 82.2, 79.2, 77.5, 75.2, 74.6, 73.5 (2), 73.4, 73.1, 68.9, 57.7, 54.0, 33.0, 26.2, 22.5; HRMS (ES) calc. for C₄₂H₄₈N₂NaO₈ (M + Na)⁺: 731.3308. Found (M + Na)⁺: 731.3309.

4-O-(α,β-D-Galactopyranosyl)-N-acetyl-trans-4-hydroxy-L-proline N'-methylamide

(4a,b): Compound **8a** (0.050 g, 0.048 mmol, 1.0 eq) was dissolved in 4:1 methanol/ethyl acetate (5 mL). After addition of Pearlman's catalyst (20% palladium hydroxide on carbon) (0.050 g, approx. 0.047 mmol, 1.0 eq), the reaction flask was flushed with N₂ for 5 minutes. The reaction mixture was then stirred vigorously under hydrogen atmosphere (10 psi) for 4 hours, after which it was flushed with nitrogen and filtered. The product was then concentrated under reduced pressure to provide **4a** as a clear oil (0.024 g, 0.069 mmol) (quant.) (a similar procedure was used for **8b**):

4a: [α]_D²⁵ = +82.0° (c 0.3 CH₃OH); ¹H NMR (500 MHz, D₂O, 298K) δ = 4.93 (d, 0.8H, H₁^{trans}, J = 3.2Hz), [4.89, d, 0.2H, H₁^{cis}, J = 3.3Hz], [4.52, dd apparent t, 0.2H, Pro_α^{cis}, J = 7.7Hz, J = 7.9Hz], 4.38-4.42 (m, 0.8H, Pro_γ^{trans}), [4.33-4.37, m, 0.2H, Pro_γ^{cis}], 4.32 (dd apparent t, 0.8H, Pro_α^{trans}, J = 8.3Hz, J = 8.4Hz), 3.80-3.86 (m, 2H, H₄^{trans}, H₄^{cis}, H₅^{trans}, H₅^{cis}), [3.72-3.77, m, 0.2H, Pro_{δ1}^{cis}], 3.63-2.71 (m, 3.6H, H₂^{trans}, H₂^{cis}, H₃^{trans}, H₃^{cis}, Pro_{δ1}^{trans}, Pro_{δ2}^{trans}), 3.57-3.63 (m, 2H, H_{6a}^{trans}, H_{6a}^{cis}, H_{6b}^{trans}, H_{6b}^{cis}), [3.45, dd, 0.2H, Pro_{δ2}^{cis}, J = 4.4Hz, J = 12.8Hz], [2.64, s, 0.6H, -NHCH₃^{cis}], 2.60 (s, 2.4H, -NHCH₃^{trans}), [2.44-2.53, m, 0.2H, Pro_{β1}^{cis}], 2.39 (ddd, 0.8H, Pro_{β1}^{trans}, J = 2.5Hz, J = 8.2Hz, J =

13.4Hz), [2.14, ddd, 0.2H, Pro β_2 ^{cis}, J = 5.5Hz, J = 7.3Hz, J = 13.4Hz], 1.94-2.04 (m, 0.8H, Pro β_2 ^{trans}), 1.99 (s, 2.4H, -COCH₃^{trans}), [1.85, s, 0.6H, -COCH₃^{cis}]; ¹³C NMR (75 MHz, D₂O, 298K) δ = 174.7, (174.6), (174.3), 173.8, 98.3, (98.1), 76.5, (74.8), 71.87, (71.85), 69.65, (69.62), 68.44, (68.38), 61.7, (60.8), 59.6, 53.9, (52.0), (38.2), 36.5, (26.3), 26.2, 21.9, (21.1); HRMS (ES) calc. for C₁₄H₂₄N₂NaO₈ (M + Na)⁺: 371.1430. Found (M + Na)⁺: 371.1438.

4b: [α]_D²⁵ = -22.5° (c 0.7 CH₃OH); ¹H NMR (500 MHz, D₂O, 298K) δ = 4.62-4.67 (m, 0.8H, Pro γ ^{trans}), [4.60, dd apparent t, 0.2H, Pro α ^{cis}, J = 7.9Hz, J = 8.1Hz], [4.57-4.61, m, Pro γ ^{cis}], 4.48 (d, 0.8H, H₁^{trans}, J = 8.1Hz), [4.46, d, 0.2H, H₁^{cis}, J = 8.1Hz], 4.40 (dd apparent t, 0.8H, Pro α ^{trans}, J = 8.3Hz, J = 8.3Hz), [3.94-3.98, m, 0.2H, Pro δ_1 ^{cis}], 3.89-3.94 (m, 1H, H₄^{trans}, H₄^{cis}), 3.79-3.89 (m, 1.6H, Pro δ_1 ^{trans}, Pro δ_2 ^{trans}), 3.66-3.78 (m, 3H, H₅^{trans}, H₅^{cis}, H_{6a}^{trans}, H_{6a}^{cis}, H_{6b}^{trans}, H_{6b}^{cis}), 3.63 (dd, 0.8H, H₃^{trans}, J = 3.1Hz, J = 10.0Hz), [3.61-3.65, m, 0.2H, H₃^{cis}], [3.58, dd, 0.2H, Pro δ_2 ^{cis}, J = 4.3Hz, J = 12.9Hz], 3.46 (dd, 0.8H, H₂^{trans}, J = 7.9Hz, J = 10.0Hz), [3.43-3.48, m, 0.2H, H₂^{cis}], [2.77, s, 0.6H, -NHCH₃^{cis}], 2.72 (s, 2.4H, -NHCH₃^{trans}), [2.55-2.62, m, 0.2H, Pro β_1 ^{cis}], 2.44-2.51 (m, 0.8H, Pro β_1 ^{trans}), [2.18-2.25, m, 0.2H, Pro β_2 ^{cis}], 2.11 (s, 2.4H, -COCH₃^{trans}), 2.03-2.10 (m, 0.6H, Pro β_2 ^{trans}), [1.96, s, 0.6H, -COCH₃^{cis}]; ¹³C NMR (75 MHz, D₂O, 298K) δ = 174.6, (174.5), (174.3), 173.7, (102.1), 102.0, 77.9, (76.7), (75.7), 75.6, (73.04), 73.02, (70.93), 70.90, 68.9, 61.3, (60.4), 59.2, 55.0, (53.2), (49.2), (37.5), 35.7, (26.3), 26.2, 21.9, (21.0); HRMS (ES) calc. for C₁₄H₂₄N₂ NaO₈ (M + Na)⁺: 371.1430. Found (M + Na)⁺: 371.1432.

4.7 References

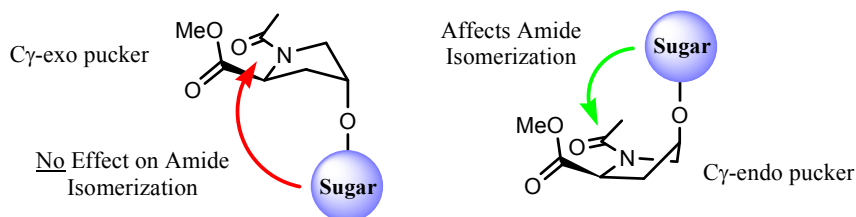
- 1) a) Varki, A. *Glycobiology* **1993**, *3*, 97-130; b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
- 2) Lamport, D. T. A. *Recent Adv. Phytochem.* **1977**, *11*, 79-115.
- 3) a) Knox, R. B.; Clarke, A.; Harrison, S.; Smith, P.; Marchalonis, J. J. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 2788-2792; b) Wang, H.; Wu, H. M.; Cheung, A. Y. *Plant Cell* **1993**, *5*, 1639-1650; c) Keller, B.; Lamb, C. J. *Genes Dev.* **1989**, *3*, 1639-1646; d) Sadava, D.; Chrispeels, M. J. *Biochem. Biophys. Acta.* **1973**, *227*, 278-287; e) Esquerre-Tugaye, M. T.; Lamport, D. T. A. *Plant Physiol.* **1979**, *64*, 314-319.
- 4) Reviewed in Fischer, G.; Schmid, F. X. *Biochemistry* **1990**, *29*, 2205-2212.
- 5) Beausoleil, E.; Sharma, R.; Michnick, S. W.; Lubell, W. D. *J. Org. Chem.* **1998**, *63*, 6572-6578.
- 6) GOESY is a selective 1D NOESY experiment, see: Stonehouse, J.; Adell, P.; Keeler, J.; Shaka, A. J. *J. Am. Chem. Soc.* **1994**, *116*, 6037-6038.
- 7) Taylor, C. M.; Hardre, R.; Edwards, P. J. B.; Park, J. H. *Org. Lett.* **2003**, *5*, 4413-4416.
- 8) Larive, C. K.; Rabenstein, D. L. *J. Am. Chem. Soc.* **1993**, *115*, 2833-2836.
- 9) Eyring, H. *J. Chem. Phys.* **1935**, *3*, 107-115.
- 10) Cai, M., Huang, Y., Liu, J., Krishnamoorthi, R. *J. Biomolecular NMR* **1995**, *6*, 123-128;

- 11) a) Eberhardt, E. S.; Panasik Jr., N.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266; b) Renner, C.; Alefelder, S.; Bae, J. H.; Budisa, N.; Huber, R.; Moroder, L. *Angew. Chem. Int. Ed.* **2001**, *40*, 923-925; c) Jenkins, C. L.; McCloskey, A. I.; Guzei, I. A.; Eberhardt, E. S.; Raines, R. T. *Biopolymers*, **2005**, *80*, 1-8.
- 12) Friebolin, H. *Basic One- and Two-Dimensional NMR Spectroscopy*; 3rd ed.; Wiley-VCH: New York, 1998.
- 13) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *J. Biochem. Biophys. Acta* **1973**, *303*, 198-202.
- 14) Improta, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, *123*, 12568-12577.
- 15) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497-2505.
- 16) Griffin, F. K.; Paterson, D. E.; Murphy, P. V.; Taylor, R. J. K. *Eur. J. Org. Chem.* **2002**, 1305-1322.

Chapter 5: The Implications of (2*S*,4*S*)-Hydroxyproline 4-*O*-Glycosylation on Prolyl Amide Isomerization

Neil W. Owens,^{*#§∞} Adrian Lee,^{#§∞} Kirk Marat,[#] and Frank Schweizer^{*§∞}

^{*}Designed research; [#]Performed research; [§]Analyzed data; [∞]Wrote the paper



5.1 Abstract

The conformation of peptides and proteins is often influenced by glycans *O*-linked to serine (Ser) and threonine (Thr). (2*S*,4*R*)-4-Hydroxyproline (Hyp), together with *L*-proline (Pro), are interesting targets for *O*-glycosylation as they have a unique influence on peptide and protein conformation. In previous work, we found that glycosylation of Hyp does not affect the *N*-terminal amide *trans/cis* ratio ($K_{trans/cis}$) nor the rate of amide isomerization in model amides. The stereoisomer of Hyp, (2*S*,4*S*)-4-hydroxyproline (hyp), is rarely found in nature, and has a different influence on the conformation of the pyrrolidine ring as well as on $K_{trans/cis}$. Glycans attached to hyp would be expected to be projected from the opposite face of the prolyl side chain relative to Hyp; the impact this would have on $K_{trans/cis}$ was unknown. Measurement of 3J coupling constants indicates that the glycan has little impact on the C γ -endo conformation instilled by hyp. As a result, it was found that the *D*-galactose residue oriented from a C γ -endo pucker affects both $K_{trans/cis}$ and the rate of isomerization, which is not found to occur when projected from a

C^γ-exo pucker; this reflects the different environments delineated by the proline side chain. The enthalpic contributions to the stabilization of the trans amide isomer may be due to disruption of intramolecular interactions present in hyp; the change in enthalpy is balanced by a decrease in entropy incurred upon glycosylation. Since the different stereoisomers Hyp and hyp, project the O-linked carbohydrates in opposite spatial orientations, these glycosylated amino acids may be useful for understanding how the projection of a glycan from the peptide or protein backbone exerts its influence.

5.2 Introduction

Glycosylation is a common post-translational modification of proteins in eukaryotes, as well as for bacteria and archaea.¹ Aside from direct involvement in various biological processes involving protein-carbohydrate recognition,² glycosylation also affects protein conformation and folding,^{2a,b} receptor binding and signalling,^{2a,b} enhancement of the thermal stability of proteins,³ protection against proteolytic degradation,⁴ hydration and hydrophilicity,⁵ and may facilitate membrane penetration.⁶ In order to explore the effect of glycosylation on peptide backbone conformation, many small model glycopeptides and glycopeptide mimics have been prepared over the years.⁷ These studies have concluded that the nature of the glycosidic linkage not only influences the presentation of the carbohydrate moiety, but also influences peptide backbone conformation.^{2d,8} Typically, carbohydrates are *O*-linked to serine (Ser) and threonine (Thr) or *N*-linked to asparagine (Asn). The *O*-glycosylation of (2*S*,4*R*)-4-hydroxyproline (Hyp) is widespread in the plant kingdom, and occurs in hydroxyproline-rich glycoproteins (HRGPs) that are associated with the cell walls of algae and flowering

plants.⁹ The stereoisomer of Hyp, (2*S*,4*S*)-4-hydroxyproline (hyp), is rarely found in nature, having been isolated from extracts of the sandalwood tree *Santalum album*, several species of fungi, and the cyanobacteria *Lyngbya majuscula*.¹⁰

Proline (Pro), along with Hyp and hyp, exhibit unique properties among proteinogenic amino acids. First, these amino acids are characterized by limited rotation of the ϕ dihedral angle, as their side chain is fused to the peptide backbone. As a result, there is a reduction in the energy difference between the prolyl amide *cis* ($\omega = 0^\circ$) and *trans* ($\omega = 180^\circ$) isomers making them nearly isoenergetic (Figure 5.2.1);¹¹ this leads to higher *cis* *N*-terminal amide isomer content relative to the other amino acids.

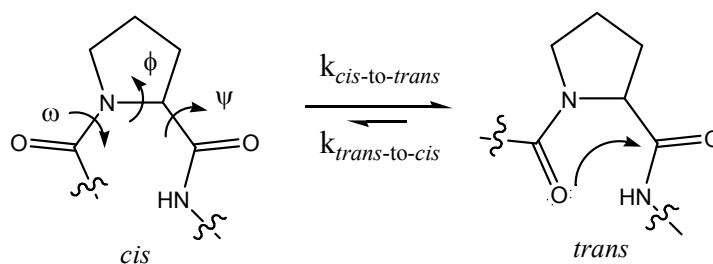


Figure 5.2.1: Proline *cis* amide isomers showing relevant backbone torsional angles as well as the *trans* amide isomer with the $n \rightarrow \pi^*$ interaction between the *N*-terminal amide carbonyl oxygen to the *C*-terminal carbonyl carbon.

Second, prolyl *cis-trans* isomerization is often the rate-determining step in the folding pathways of many peptides and proteins^{11a,12} and third, Pro and Hyp induce β -turns and extended helical structures (polyproline helix) in peptides that are crucial in protein-protein and protein-peptide interactions.^{13,14} Moreover, both Pro and Hyp play an important role in the stability of structural proteins such as collagen,¹⁵ and have been implicated in contributing to the stability of HRGPs, such as the extensins.¹⁶ As a result,

there is a growing interest in understanding and controlling prolyl *N*-terminal amide isomerization in biological processes.

The prolyl *cis-trans* isomerization equilibrium has been proposed to be governed by an $n \rightarrow \pi^*$ interaction between the oxygen lone pair from the prolyl *N*-terminal amide C=O to the antibonding orbital of the *C*-terminal C=O (Figure 5.2.1).^{11,17-20} This interaction is analogous to the approach of an amino nucleophile to a carbonyl group.⁵⁹ There is a significant amount of evidence that this interaction does occur,⁶⁰ however the expected changes in bond angle and bond length in the *C*-terminal C=O are not observed in the crystal structure of proline model peptides,³² therefore the $n \rightarrow \pi^*$ interaction remains a model that is only analogous to the Bürgi-Dunitz approach of a nucleophile.

Recent studies have shown that 4-hydroxylation and attachment of other electron-withdrawing groups have a further influence on *cis-trans* isomerization through inductive and stereoelectronic effects,^{17,21-25} which affect the $n \rightarrow \pi^*$ interaction through changes to the conformation of the pyrrolidine ring^{26,27} and the prolyl backbone ψ dihedral angle.^{17,21-25} It has also been shown that carbohydrates can affect *cis-trans* isomerization, as Ser *O*-glycosylated with either α -linked *N*-acetyl galactosamine or β -linked *N*-acetyl glucosamine *N*-terminal to Pro stabilizes the *trans* amide conformation.²⁸ Recently, our group has developed several unnaturally *C*-glycosylated Pro analogues, which have demonstrated a strong ability to vary the prolyl *N*-terminal amide equilibrium ($K_{trans/cis}$).²⁹

In terms of clarifying the effects of 4-*O*-glycosylation of Hyp, as found in HRGPs, we demonstrated in earlier work that both α - and β -glycosylation of Hyp in model amides had no apparent effect on the isomer equilibrium constant or the rate of amide isomerization.³⁰ However, our results indicated that galactosylation of Hyp may

cause an inductive electron-withdrawing effect on the prolyl ring. It is known that (4*R*)-electronegative substituents stabilize the C^γ-exo pucker of proline.^{17,21-25} Moreover, nOe experiments indicated that glycosylation of Hyp resulted in distant contacts between the proline and galactose rings, which indicates that glycosylation induces conformational constraint into glycopeptides. Based on these results, we became interested in studying how galactosylation of hyp, the stereoisomer of Hyp, influences the thermodynamics and kinetics of prolyl amide *cis-trans* isomerization. We anticipated that *O*-glycosylation of hyp may have a different impact on *N*-terminal amide isomerization since 4*S*-hydroxylation causes the hydroxyl group to be projected from the opposite face of the prolyl side chain relative compared to its stereoisomer (Figure 5.2.2).

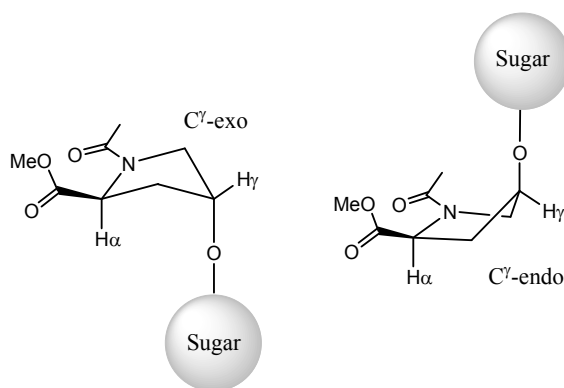


Figure 5.2.2: The C^γ-exo pucker of Hyp would place the sugar below the plane of the proline ring, while the C^γ-endo pucker of hyp would place the sugar above the plane of the proline side chain.

While Hyp has been found to adopt a C^γ-exo conformation, hyp is associated with a C^γ-endo conformation;³¹⁻³⁵ the closely related C^β-exo conformation has also been suggested.^{36,37} This conformational switch has been attributed to a stereoelectronic effect, in which the 4-hydroxyl group prefers to adopt a gauche orientation with the prolyl

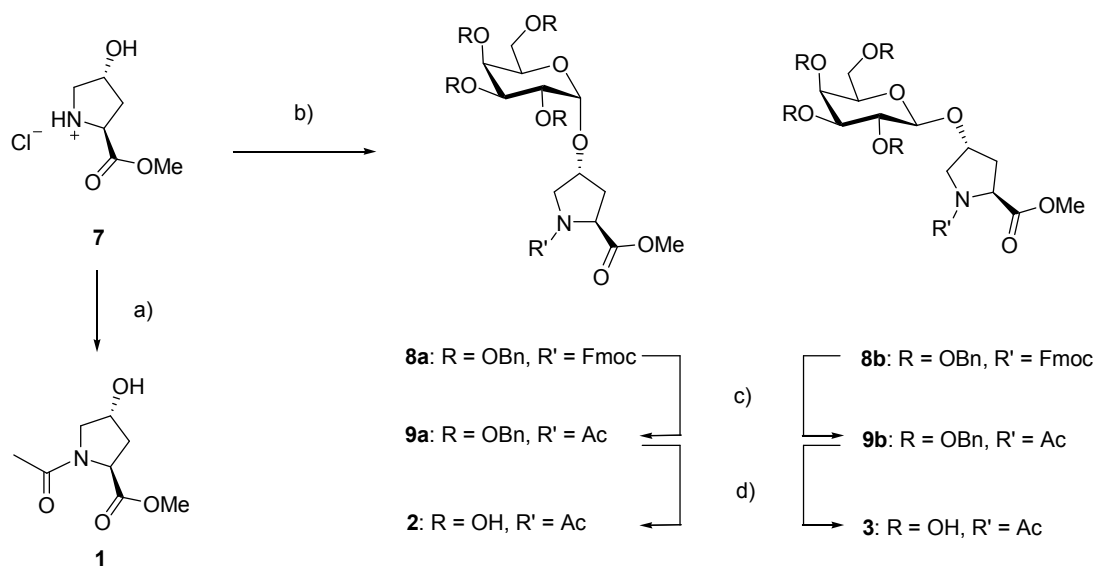
nitrogen atom.^{17,21-25} This gauche orientation is stabilized by hyperconjugative $\sigma(\text{C}^\beta\text{-H}) \rightarrow \sigma^*(\text{C}^\gamma\text{-O})$ and $\sigma(\text{C}^\delta\text{-H}) \rightarrow \sigma^*(\text{C}^\gamma\text{-O})$ interactions.⁶¹ As a result, Hyp favours the *trans* amide conformation relative to Pro because the $\text{C}^\gamma\text{-exo}$ pucker forces a ψ -angle of 150° , which is ideal for a favourable $n \rightarrow \pi^*$ interaction; this interaction between the *N*-terminal amide carbonyl oxygen to the *C*-terminal carbonyl carbon has been shown to stabilize the *trans* amide (Figure 5.3.1).^{17,21-25} In contrast, the $\text{C}^\gamma\text{-endo}$ conformation associated with hyp has been shown to disfavour the *trans* amide conformation because of an unfavourable ψ dihedral angle for the same $n \rightarrow \pi^*$ interaction, which has been attributed to several factors: experimental and computational methods indicate that a hydrogen bond likely exists between the 4-hydroxyl group and the *C*-terminal carbonyl oxygen atom in hyp, as well as electrostatic repulsion between 4-position oxygen atom and the *C*-terminal carbonyl oxygen atom;^{17,35} both of these factors likely force the prolyl ψ -angle into a poor orientation for the $n \rightarrow \pi^*$ interaction, resulting in hyp favouring the *cis* amide conformation relative to Pro and Hyp.

Since glycosylation of hyp effectively places the carbohydrate moiety on a different face of the proline side chain relative to Hyp, it may have a different impact on proline isomerization. We explored the effects of glycosylation of (2*S*,4*S*)-4-hydroxyproline on prolyl conformation, and the kinetics and thermodynamics of prolyl amide isomerization using the monosaccharide D-galactose attached in both α - and β -anomeric configurations. Galactose was selected due to the occurrence of Hyp-Gal linkages in HRGPs.^{9,16}

5.3 Results

5.3.1 Synthesis

Model peptides of the form, *N*-acetyl-Pro-*N'*-methyl ester are well established for studying subtle effects of modification of the proline side chain on *N*-terminal amide isomerization.^{17,21-25,26,38} This prevents intramolecular hydrogen bonding and γ -turn formation observed in *N*-acetyl proline *N*-methylamide model peptides.³⁹ Therefore, the model amides Ac-Hyp-OMe (**1**), Ac-[Hyp(α -D-Gal)]-OMe (**2**), Ac-[Hyp(β -D-Gal)]-OMe (**3**), Ac-hyp-OMe (**4**), Ac-[hyp(α -D-Gal)]-OMe (**5**), and Ac-[hyp(β -D-Gal)]-OMe (**6**) were synthesized using different strategies depending on the commercial availability of each proline derivative (Schemes 5.3.1.1 and 5.3.1.2).

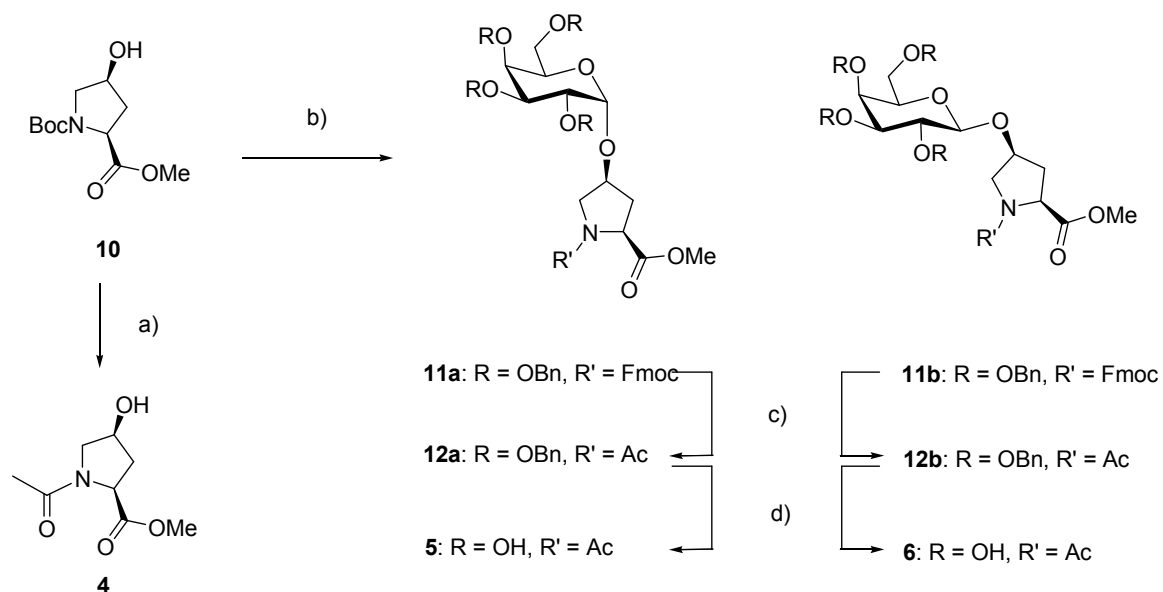


Scheme 5.3.1.1: Synthesis of **1**, **2**, and **3**. a) Ac₂O, Et₃N, MeOH, 25 °C, 2h, 95%; b) i) Fmoc-OPfp, NaHCO₃, acetone/H₂O (3:1), 25 °C, 3h. ii) 2,3,4,6-Tetra-*O*-benzyl-thio- β -D-galactopyranoside, BF₃·OEt₂, AgOTf, NIS, CH₃CN, 0 to 25 °C, 2h, 35% overall yield; c) i) CH₂Cl₂/piperidine (1:1), 25 °C, 3h. ii) Ac₂O, Py, 25 °C, 4h, 78% overall yield; d) H₂/Pd(OH)₂-C, MeOH/ethyl acetate (4:1), 25 °C, 4h, quantitative.

Model amides having both α - and β -anomeric linkages were made for comparison to explore how the nature of the glycosidic linkage impacts prolyl amide *cis-trans* isomerization. The synthesis of **1** was carried out by selective *N*-acylation of **7** using acetic anhydride and triethylamine in methanol in 95% yield (Scheme 5.3.1.1). The glycosylated 4*R*-hydroxyproline model peptides **2** and **3** were obtained through *N*-Fmoc protection of **7** using 9-fluorenylmethyl pentafluorophenyl carbonate under mild basic conditions, followed by glycosylation using 2,3,4,6-tetra-*O*-benzyl-thio- β -D-galactopyranoside⁴⁰ in the presence of boron trifluoride dietherate, *N*-iodosuccinimide and catalytic amounts of silver triflate to yield both the α - (**8a**) and β - (**8b**) linked products in 35% yield over two steps. Slightly more of the α -linked product is formed under these conditions. Installation of the *N*-acetyl group was carried out by *N*-Fmoc deprotection using piperidine in dichloromethane followed by *N*-acylation using acetic anhydride in pyridine to give **9a** and **9b** in 78% yield. Removal of the benzyl ether protecting groups by catalytic hydrogenation in methanol gave the model amides **2** and **3** in 27% overall yield from **7**. Assignment of the α - and β -anomers was carried out by measuring the ¹H NMR ³*J*_{H1,H2} coupling constant, which for **2** was 1.9 Hz indicative of the gauche dihedral angle expected for the α -anomer, while ³*J*_{H1,H2} was 8.0 Hz for **3** indicative of the *trans* diaxial relationship expected for the β -anomer.

The synthesis of the 4*S*-model amides was carried out in a similar fashion beginning from **10** (Scheme 5.3.1.2). *N*-Boc deprotection using trifluoroacetic acid in dichloromethane at 0 °C gave **4** in 94% yield after *N*-acylation using standard conditions. Glycosylation of **10** was carried out using the same conditions as for **7**. It was found that separation of the α - and β -glycosides required replacement of the *N*-Boc group with the

N-Fmoc protecting group, which was carried out using standard conditions to give **11a** and **11b** in 32% yield, with slightly more of the α -anomer formed. Installation of the *N*-acetyl group to give **12a** and **12b** followed by removal of the benzyl ether protecting groups was carried out using the same conditions as for **8a** and **8b**, to give the model peptides **5** and **6** in 25% overall yield over three steps.



Scheme 5.3.1.2: Synthesis of **4**, **5**, and **6**. a) i) TFA/CH₂Cl₂ (1:1), 0 °C, 1.5h. ii) Ac₂O, Et₃N, MeOH, 25 °C, 2h, 94% overall yield; b) i) 2,3,4,6-Tetra-*O*-benzyl-thio- β -D-galactopyranoside, BF₃·OEt₂, AgOTf, NIS, CH₃CN, 0 to 25 °C, 2.5h. ii) TFA/CH₂Cl₂ (1:1), 0 °C, 1.5h. iii) Fmoc-OPfp, NaHCO₃, acetone/H₂O (3:1), 25 °C, 2h, 32% overall yield; c) i) CH₂Cl₂/piperidine (1:1), 25 °C, 3h. ii) Ac₂O, Py, 25 °C, 4h, 78% overall yield; d) H₂/Pd(OH)₂-C, MeOH/ethyl acetate (4:1), 25 °C, 4h, quantitative.

5.3.2 IR spectroscopic study

The frequency of the amide I vibrational mode (ν_{amide}), which results primarily from the amide C=O stretching vibration, has been correlated to changes in the bond order of the amide C=O group.⁴¹ In D₂O, ν_{amide} for the model compounds **1**, **2**, and **3** were nearly identical, with maxima at 1612, 1611.5 and 1611.5 cm⁻¹, respectively (see Table 11.1). Similarly, nearly identical maxima of 1612, 1611, and 1613 cm⁻¹ were exhibited by **4**, **5**, and **6**, respectively. Therefore, for compounds **1-6**, no change in the amide carbonyl I vibrational mode was found to occur upon α - or β -glycosylation or by inversion of stereochemistry at the 4-position. A slightly more pronounced effect was observed in a study of 4-fluoroproline (Flp) in Ac-Flp-OMe model compounds, which showed a ν_{amide} maxima shift of 3 cm⁻¹ in D₂O relative to 4*R*-hydroxyproline in similar model compounds.²¹

5.3.3 NMR spectroscopic studies

Full assignment of ¹H NMR spectra of **1-6** was carried out using COSY and HSQC experiments. Assignment of the major isomer as the *trans* amide isomer for **1-6** was established using selective one-dimensional GOESY⁴² experiments, which showed interproton effects from the *N*-acetyl methyl group to the prolyl δ -protons (1.1-2.4% nOe relative to the *N*-acetate singlet signal). By comparison, this interaction was not observed in the minor isomer.

5.3.4 Prolyl side-chain conformation

The prolyl ring pucker of the major isomers was established for **1-6** by comparison of ^1H NMR coupling constants to literature values (see Table 11.2).⁴³ The prolyl pucker for **1-3** were assigned to a C^γ -exo pucker based on $^3J_{\alpha,\beta 1}$ of 8.2-8.3 Hz, and $^3J_{\alpha,\beta 2}$ of 8.5-8.7 Hz. The coupling constants for the C^γ -exo pucker are expected to be 7-10 and 7-11 Hz, respectively. In contrast, the prolyl pucker for **4-6** were assigned to a C^γ -endo pucker based on $^3J_{\alpha,\beta 1}$ of 7.8-9.8 Hz, and $^3J_{\alpha,\beta 2}$ of 2.4-4.3 Hz. The coupling constants for the C^γ -endo pucker are expected to be 6-10 and 2-3 Hz, respectively. These results are as expected relative to other experimental^{17,21-25,31-33,43} and computational^{34,44-46} measurements. The coupling constants for **4** and **6** were nearly identical at 25 °C, while **5** deviated slightly.

5.3.5 C^γ Inductive effect

Changes in ^{13}C chemical shifts have been used to estimate the electron-withdrawing effect of a substituent on the prolyl side-chain.⁴⁷ Therefore, measurement of the C^γ -carbon chemical shifts of the model compounds **1-6** was used to assess the relative changes in electron-withdrawing ability incurred upon glycosylation. A significant change in the shift of the C^γ -carbon (~9 ppm) was found to occur upon glycosylation for the 4R model compounds (78.9, 77.6 ppm compared to 69.9 ppm for **2**, **3** and **1**, respectively) and the 4S model compounds (80.3, 80.6 ppm compared to 69.9 ppm for **5**, **6** and **4**, respectively). Therefore, glycosylation appears to cause a local electron-withdrawing effect. This is similar in magnitude to the $^{13}\text{C}^\gamma$ chemical shift observed from

the attachment of a trifluoroacetate group to similar model peptides of ~ 8 ppm, but more than a simple acyl group which caused only a shift of ~ 3 ppm.³³

5.3.6 nOe experiments

To determine the extent of interaction between the galactose and prolyl rings, selective nOe transfer experiments were performed on the galactosylated model peptides **2**, **3**, **5** and **6** in D₂O (Figure 5.3.6.1).

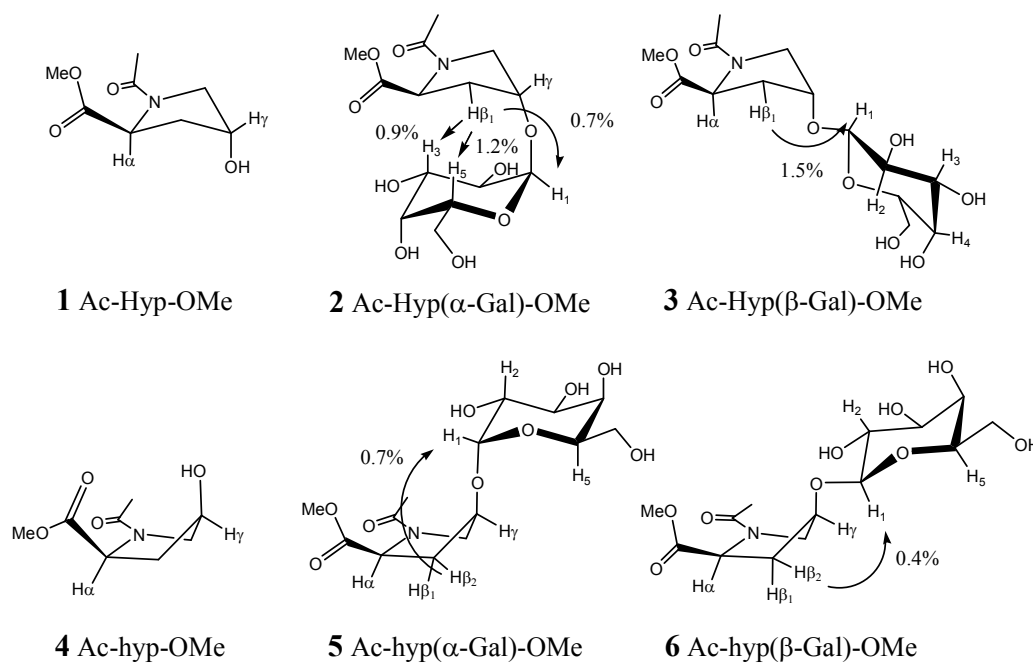


Figure 5.3.6.1: Relevant nOe interactions between the prolyl side chain and the carbohydrate ring for **1-6**. These experiments indicate that close contacts exist between the rings when D-galactose is α -linked to Hyp, but not when α -linked to hyp. Furthermore, both α - and β -linked D-galactose seem to be oriented away from the proline side chain when *O*-linked to hyp. Thus, Hyp and hyp have different impacts on the orientation of the glycans.

For compound **2**, it was found that selective inversion of Hyp $_{\beta 1}$ resulted in a 0.9% and 1.2% resonance transfer to the peaks at $\delta = 3.96$ ppm and $\delta = 3.99$ ppm, which correspond to the H₃ and H₅ protons of galactose, respectively. This suggests that α -galactosylation of Hyp results in close contact between distant positions in the galactose and prolyl rings (Figure 5.3.6.1). The overlap of the hydrophobic α -face of D-galactose with the pyrrolidine ring of proline in **2** resembles other hydrophobic galactose-protein interactions found in several crystal structures.^{48,49} In contrast, selective inversion of Hyp $_{\beta 1}$ in **3** only showed resonance transfer to H₁ of galactose (1.5%), and no other sugar protons. However, there was greater overlap of the galactose protons in the ¹H spectrum of **3**, which made assignment of nOe contacts with certainty more difficult. Selective inversion of the Hyp $_{\beta 1}$ /Hyp $_{\beta 2}$ signal in **5** and **6** only showed an nOe contact to H₁ of galactose (0.7 and 0.4%, respectively), but not other sugar protons. Therefore, the galactose ring in the glycosylated 4*S*-hyp model peptides **5** and **6** is likely located distally to the proline ring (Figure 5.3.6.1).

5.3.7 Measurement of $K_{trans/cis}$

The ratio of *trans/cis* isomers ($K_{trans/cis}$) was established by integrating as many well-resolved peaks as possible for each isomer, and taking the average for all peaks for each respective isomer (Table 5.3.7.1).²⁰ As previously found, glycosylation of 4*R*-Hyp does not affect $K_{trans/cis}$ in D₂O (**1-3** had $K_{trans/cis}$ of 5.9-6.2).³⁰ However, it was found that the glycosylation of 4*S*-hyp does affect $K_{trans/cis}$. The α - and β -linked sugars in **5** and **6**, respectively, were found to stabilize the prolyl *N*-terminal *trans* amide conformation to an equal extent ($K_{trans/cis}$ of 2.9 in D₂O at 25 °C) compared to **4** ($K_{trans/cis}$ of 2.4). This is in

contrast to previous studies of 4-*O*-modification of hyp, which saw a slight stabilization of the *cis* amide isomer. Taylor *et al.* found that *O*-methylation of hyp in Ac-Phe-4*S*-hyp-NHMe model peptides caused a decrease in $K_{trans/cis}$ from 1.9 to 1.4 in D₂O at 25 °C.³⁶ However, this result is complicated by the prolyl *N*-terminal L-phenylalanine residue, which is known to stabilize the *cis* amide isomer through interaction with the prolyl side chain.^{14b} Similarly, Jenkins *et al.* found that *O*-acylation of hyp in Ac-4*S*-hyp-OMe model compounds caused a decrease in $K_{trans/cis}$ from 1.8 ± 0.4 to 1.6 ± 0.2 in D₂O at 25 °C, although these values were within experimental error.³³

Table 5.3.7.1: Thermodynamic parameters for 1-6.

Compound	$K_{trans/cis}$ ^[a]	$\Delta H^\circ_{cis \rightarrow trans}$ (kcal·mol ⁻¹) ^[b]	$\Delta S^\circ_{cis \rightarrow trans}$ (cal·mol ⁻¹ ·K ⁻¹) ^[b]	ΔG°_{300K} (kcal·mol ⁻¹) ^[c]
1	6.2 ± 0.1	-1.43 ± 0.04	-1.17 ± 0.11	-1.08 ± 0.07
2	6.0 ± 0.1	-1.42 ± 0.12	-1.14 ± 0.37	-1.08 ± 0.23
3	5.9 ± 0.2	-1.45 ± 0.13	-1.38 ± 0.41	-1.04 ± 0.25
4	2.4 ± 0.1	-0.29 ± 0.08	0.79 ± 0.25	-0.53 ± 0.16
5	2.9 ± 0.3	-0.94 ± 0.13	-1.03 ± 0.42	-0.63 ± 0.26
6	2.9 ± 0.1	-0.91 ± 0.05	-0.96 ± 0.16	-0.62 ± 0.10

^[a]Carried out in D₂O at 24.8 °C; ±SE determined by integration of two or more sets of *trans/cis* isomers; ^[b]Error limits obtained by linear least-squares fitting of the van't Hoff plots to equation $\ln K_{ct} = (-\Delta H^\circ/R)(1/T) + \Delta S^\circ/R$; ^[c]Calculated using $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ$.

5.3.8 Thermodynamics

The effect of temperature on $K_{trans/cis}$ of compounds **1-6** was measured by NMR spectroscopy and the resulting van't Hoff plots are shown in Figure 5.3.8.1. This assumes that the enthalpic and entropic energy differences between the *cis* and *trans* amide isomers are independent of temperature; the linear van't Hoff plots indicate this assumption is likely valid.^{21,33,38,50}

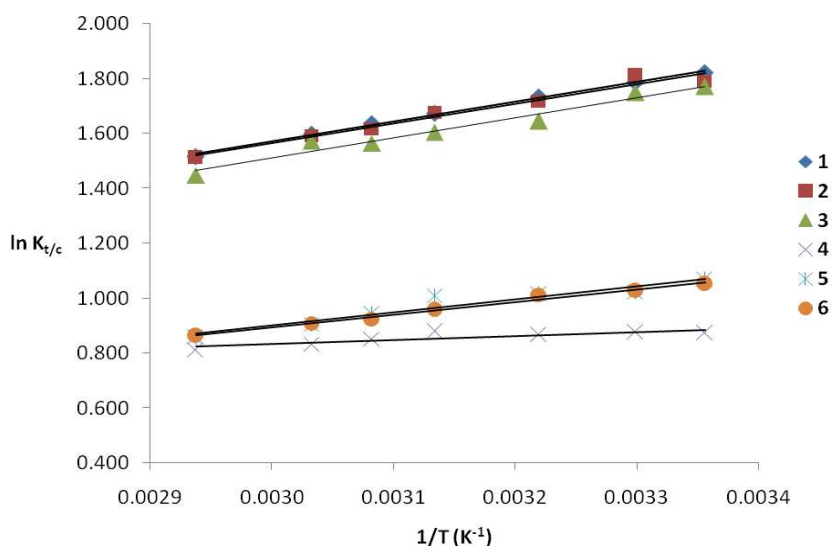


Figure 5.3.8.1: van't Hoff plots for **1-6** in D₂O.

We found that in all cases, the $K_{trans/cis}$ was found to decrease with increasing temperature. Compound **4** was slightly anomalous by exhibiting less temperature dependence on $K_{trans/cis}$ relative to the other model compounds. Accordingly, ΔH° and ΔS° could be calculated from the least-squares fits of the van't Hoff plots. In all cases, ΔH° was < 0 , which correlates well to other studies of proline model peptides (Table 5.3.7.1).^{21,38,51,52} The differences in $K_{trans/cis}$ were reflected in ΔH° and ΔS° , with no significant differences in ΔH° between **1-3** (ΔH° of -1.43 ± 0.04 , -1.42 ± 0.12 , and -1.45

$\pm 0.13 \text{ kcal}\cdot\text{mol}^{-1}$, respectively) or ΔS° (ΔS° of -1.17 ± 0.11 , -1.14 ± 0.37 , and $-1.38 \pm 0.41 \text{ cal}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$, respectively). In contrast, there was an increase of $0.6 \text{ kcal}\cdot\text{mol}^{-1}$ in the enthalpic difference in **5** and **6** (ΔH° of -0.94 ± 0.13 and $-0.91 \pm 0.05 \text{ kcal}\cdot\text{mol}^{-1}$, respectively) relative to **4** (ΔH° of $-0.29 \pm 0.08 \text{ kcal}\cdot\text{mol}^{-1}$). Also, there seemed to be different entropic contributions to the ground state energy difference (ΔG°) between the *cis* and *trans* amide isomers, as glycosylation seemed to cause the system in the *trans* amide conformation to become more ordered in **5** and **6** (ΔS° of -1.03 ± 0.42 and $-0.96 \pm 0.16 \text{ cal}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$, respectively), whereas the system when **4** adopted a *trans* amide conformation was more disordered (ΔS° of $+0.79 \pm 0.25 \text{ cal}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$).

5.3.9 Kinetics

The rates of *cis-trans* amide isomerization were calculated for **1-6** using ^1H NMR magnetization inversion transfer experiments in 0.1 M phosphate buffer (pH 7.2).^{21,30,31,53} These experiments were performed at elevated temperature, since at physiological temperature the kinetic rates are too slow to be determined by this assay. It was found that the glycosylated hyp model compounds **5** and **6** exhibited an increase in the rate of isomerization ($k_{cis\rightarrow trans}$ of $0.59 \pm 0.06 \text{ sec}^{-1}$ and $0.71 \pm 0.04 \text{ sec}^{-1}$, respectively) relative to the unglycosylated model peptide **4** ($k_{cis\rightarrow trans}$ of $0.44 \pm 0.04 \text{ sec}^{-1}$) at 67 °C (Table 5.3.9.1). This is in contrast to the glycosylated Hyp compounds **2** and **3** which showed almost no difference in the rate of isomerization ($k_{cis\rightarrow trans}$ of $0.85 \pm 0.02 \text{ sec}^{-1}$ and $0.77 \pm 0.01 \text{ sec}^{-1}$, respectively) compared to the unglycosylated model compound **1** ($k_{cis\rightarrow trans}$ of $0.81 \pm 0.01 \text{ sec}^{-1}$) at 67 °C, which is consistent with previous work.³⁰

Table 5.3.9.1: Kinetic parameters for **1-6**.

Compound	$k_{cis \rightarrow trans}^{[a]}$ (sec ⁻¹)	$k_{trans \rightarrow cis}^{[b]}$ (sec ⁻¹)
1	0.81 ± 0.01	0.18 ± 0.01
2	0.85 ± 0.01	0.19 ± 0.01
3	0.77 ± 0.02	0.18 ± 0.02
4	0.44 ± 0.04	0.20 ± 0.01
5	0.59 ± 0.06	0.25 ± 0.03
6	0.71 ± 0.04	0.30 ± 0.02

^[a]Carried out in 0.1 M phosphate buffer pH = 7.2 at 67.3 °C, the temperature was calibrated using an ethylene glycol standard, error values obtained using the linear least-squares fit of the data (in Figures 11.7-11.12);

^[b]Calculated from $k_{cis \rightarrow trans}$ and amide isomer equilibrium ($K_{trans/cis}$) at 67.3 °C.

5.4 Discussion

The change in the ¹³C^γ chemical shift of **5** and **6** relative to **4** of ~10 ppm (and similarly for **2** and **3** relative to **1**) is indicative that glycosylation may cause a local electron-withdrawing effect. An increased electron-withdrawing ability has been attributed to stabilizing a given pucker through a stereoelectronic effect.^{17,21-25,54} This likely explains why glycosylation does not significantly change the prolyl pucker for **1-3** and **4-6** as determined by the measurement of ³J coupling constants. However, since changes in $K_{trans/cis}$ have been correlated to a change in the prolyl pucker, the stabilization

of the *trans* amide isomer in **5** and **6** relative to **4** cannot be explained by such a conformational change in this case.^{26,27}

The local electron-withdrawing effect caused by glycosylation may diminish the intramolecular electrostatic repulsion between 4-hydroxyl group and the C-terminal carbonyl oxygen atom in **5** and **6**.³⁵ This would allow the prolyl ψ -angle to relax from 180° closer to the optimal angle of 150° for the favourable $n \rightarrow \pi^*$ interaction that has been estimated to contribute 0.7 kcal·mol⁻¹ to the stability of the *trans* amide isomer.^{22,55} As this is specific for the C ^{γ} -endo conformation, it would explain the lack of effect in compounds **1-3**, which have C ^{γ} -exo puckers.

The effect of glycosylation on the intramolecular hydrogen bond in hyp on $K_{trans/cis}$ should also be considered as it has been estimated to contribute 1.5 kcal·mol⁻¹ to stabilizing the C ^{γ} -endo conformation, and thus to the *cis* amide isomer.³⁵ The glycosidic linkage would be expected to eliminate this intramolecular interaction, and since nOe experiments indicated the sugar is not proximal to the prolyl side chain, it is unlikely to be restored through the sugar hydroxyl groups. The removal of the intramolecular hydrogen bond should destabilize the C ^{γ} -endo conformation in **5** and **6** and thereby destabilize the *cis* amide isomer. This would also explain why there is no impact from glycosylation on the $K_{trans/cis}$ in **1-3**, since it has no equivalent intramolecular hydrogen bond. However, Taylor *et al.* have downplayed the effects of the intramolecular hydrogen bond as *O*-methylation of hyp had little effect on $K_{trans/cis}$.³⁶

The proposed reduction of the intramolecular electrostatic repulsion and loss of the intramolecular hydrogen bond, which have an influence on the enthalpic contributions (ΔH°) to the ground state energy difference (ΔG°) between the *trans* and *cis* amide

isomers, seem to be counteracted by a decrease in entropy (ΔS°) incurred upon glycosylation in compounds **4-6** (Table 5.3.7.1). Here, glycosylation had an impact on the entropic difference beyond what was established between the prolyl *trans* and *cis* amide isomers in **4-6**, which itself has been attributed to differences in solvation.^{52,56} Interestingly, this effect is less apparent for **1-3**, therefore the influence of solvation differences seem to be specific to the face of the proline side chain.

NMR magnetization inversion transfer experiments indicated that the hyp model compounds **1-3** have faster amide isomerization rates overall compared to **4-6**. Improta *et al.* have calculated that the prolyl nitrogen is more pyramidalized in the C^γ-exo pucker relative to the C^γ-endo pucker,³⁵ which would facilitate isomerization for **1-3** that have a C^γ-exo pucker, compared to **4-6** with a C^γ-endo pucker. This is in contrast to the findings of Beausoleil *et al.* who found the reverse effect: hyp had a faster rate than Hyp in Ac-X_{aa}-NHMe model amides in D₂O at 60 °C (2.05 ± 0.50 and 1.46 ± 0.13 sec⁻¹, respectively).³¹ This was attributed to the intramolecular hydrogen bond in hyp reducing coulombic repulsion between the C-terminal carbonyl oxygen atom and the prolyl nitrogen, however the values were within experimental error.

Glycosylation caused an increase in the rate of isomerization of **5** and **6** ($k_{cis \rightarrow trans}$ of 0.59 ± 0.04 and 0.71 ± 0.06 sec⁻¹, respectively) relative to **4** (0.44 ± 0.04 sec⁻¹), whereas an increase in isomerization rate was not observed in **1-3** (Table 5.3.9.1). In contrast, the acylation of hyp by Jenkins *et al.* caused no observable increase in the rate of amide isomerization.³³ Changes in *cis-trans* amide isomerization have been correlated to an inductive electron-withdrawing effect from prolyl γ -substituents,^{21,23} in which the γ -position group withdraws electron density from the peptide bond thereby increasing *N*-

pyramidalization,^{17,35} and reducing the C-N bond order;⁵⁷ this effectively weakens the amide bond and enables isomerization to occur.³⁶ While glycosylation does appear to cause a local electron-withdrawing effect, it does so for both 4*S*- and 4*R*-stereoisomers, and therefore does not explain the relative increase in isomerization rate for **5** and **6** compared to **4**, with no change between **1-3**.

An increase in C=O bond order is indicative of a lower C-N bond order, and has been used to explain the increase in the rate of isomerization of 4-fluoroproline.²³ Here, perhaps because of a weaker electron-withdrawing effect, glycosylation of both the 4*R*- and 4*S*-hydroxyproline model compounds did not affect the amide I vibrational mode maxima, and therefore also cannot be used to explain the change in the rate in isomerization of **5** and **6** compared to **4**. Therefore, the basis for the increase in the rate of amide isomerization in **5** and **6** relative to **4**, while there is no effect in **1-3** remains unclear.

5.5 Conclusions

While the glycosylation of 4*R*-hydroxyproline was not found to affect amide isomerization, the glycosylation of 4*S*-hydroxyproline affects both the prolyl *N*-terminal amide equilibrium and the rate of amide isomerization. Here, we found that both α - and β -anomeric linkage to 4*S*-hydroxyproline in **5** and **6** stabilizes the *trans* amide conformation relative to the unglycosylated model compound **4**. Glycosylation does not significantly alter the C ^{γ} -endo conformation of **4** induced by the 4*S*-hydroxylation, but does cause a local electron-withdrawing effect; this likely reduces repulsion between the 4-oxygen atom and the *C*-terminal carbonyl group and allows the ψ -angle to relax closer

to 150° to restore the $n \rightarrow \pi^*$ interaction that stabilizes the *trans* amide isomer. The change in $K_{trans/cis}$ for **5** and **6** relative to **4** may also be due to the loss of an intramolecular hydrogen bond that is specific to the 4*S*-stereoisomer. This interaction is likely not restored in **5** and **6** since the sugar was not found to form close contacts to the prolyl side chain. Regardless of their origin, the enthalpic contributions to ΔG° seem to be offset by entropic changes, which create a more ordered environment for the *trans* amide isomer in **5** and **6** relative to **4**; resulting in only a small net change in $K_{trans/cis}$. The glycosylation of 4*S*-hydroxyproline does seem to cause a small increase in rate of amide isomerization, but it is not reflected in the amide I vibrational mode and thus the bond order of the *N*-terminal amide group. Therefore, the cause of the increase in the rate remains unclear.

The different stereoisomers of 4-hydroxyproline provide an opportunity to understand how each face of the prolyl ring has an influence on $K_{trans/cis}$. Furthermore, Hyp and hyp can be used to rigidly project a glycan in opposite spatial orientations; therefore, these building blocks may be useful for studying carbohydrate binding interactions and influences of glycans on peptide and protein structure in which the orientation of the glycan is important.⁵⁸

5.6 Experimental section

General procedures: Reagent grade solvents were used without further purification. Thin-layer chromatography was performed on Si250F precoated plates of silica gel (250 μ m). Column chromatography was performed on SilicaFlash P60 silica gel (40-63 μ m). NMR spectra were assigned based on 2D COSY and 2D HSQC experiments. For

^1H NMR, minor isomers are listed between square brackets. For ^{13}C NMR, when assigned, carbon peaks for the minor isomer are listed in brackets.

General Preparation of *N*-Acetyl-*N'*-methyl ester amino acids: The amino acid (1.0 eq) was dissolved in 3:1 (v/v) DCM/piperidine, stirred for 2 hours at ambient temperature, then co-distilled with toluene (3 x 10 mL). The crude product was then dissolved in 1:1 acetic anhydride/pyridine, and the mixture was stirred at ambient temperature for 4 hours. The solution was concentrated under reduced pressure, and then was co-distilled with toluene (3 x 10 mL) before purification by flash chromatography.

Preparation of 4-*O*-Galactopyranosyl-*N*-acetyl-*N'*-methyl ester amino acids: The amino acid (1.0 eq) was dissolved in methanol. The catalyst, (20% palladium hydroxide on carbon) was added (approx. 0.5 eq), and the flask was flushed with N_2 for 5 minutes. The reaction mixture was then stirred under hydrogen gas (10 psi) for 6 hours, after which it was flushed with nitrogen and filtered. The product was then concentrated under reduced pressure.

***N*-Acetyl-*trans*-4-hydroxy-*L*-proline methyl ester (1):** Compound **7** (0.250 g, 1.38 mmol,) was dissolved in methanol (2 mL), followed by the addition of triethylamine (0.77 mL, 5.52 mmol, 4.0 eq) and acetic anhydride (0.78 mL, 8.28 mmol, 6.0 eq). The reaction was stirred for 2 hours at ambient temperature before being concentrated under reduced pressure. The product was purified by flash chromatography using 9:1 ethyl acetate/methanol to yield **1** as a white solid (0.244 g, 1.31 mmol) (95.0%): $[\alpha]_{\text{D}}^{25} = -$

83.3° (*c* 0.7 CH₃OH); mp 75-78°C; ¹H NMR (500 MHz, D₂O, 298K) δ = [4.83, dd, 0.14H, Pro_α^{cis}, J_{α, β1} = 8.0Hz, J_{α, β2} = 7.6Hz], 4.61 (dddd, 0.86H, Pro_γ^{trans}, J_{β1, γ} = 2.1Hz, J_{β2, γ} = 4.4, J_{γ, δ1} = 4.4Hz, J_{γ, δ2} = 1.9Hz), 4.54 (dd, 0.86H, Pro_α^{trans}, J_{α, β1} = 8.4Hz, J_{α, β2} = 8.6Hz), [4.51-4.57, m, 0.14H, Pro_γ^{cis}], 3.85 (dd, 0.86H, Pro_{δ1}^{trans}, J_{δ1, δ2} = 13.7Hz), [3.83, s, 0.4H, -COCH₃^{cis}], 3.79 (s, 2.6H, -COCH₃^{trans}), [3.71, dd, 0.14H, Pro_{δ1}^{cis}, J_{γ, δ1} = 2.2Hz, J_{δ1, δ2} = 12.6Hz], 3.66 (dd, 0.86H, Pro_{δ2}^{trans}), [3.56, dd, 0.14H, Pro_{δ2}^{cis}, J_{γ, δ2} = 4.5Hz], [2.50, ddd, 0.14H, Pro_{β1}^{cis}, J_{β1, γ} = 1.7Hz, J_{β1, β2} = 13.7Hz], 2.39 (ddd, 0.86H, Pro_{β1}^{trans}, J_{β1, β2} = 13.7Hz), [2.34-2.40, m, 0.14H, Pro_{β2}^{cis}], 2.18 (ddd, 0.86H, Pro_{β2}^{trans}), 2.14 (s, 2.6H, -NCOCH₃^{trans}), [2.03, s, 0.4H, -NCOCH₃^{cis}]; ¹³C NMR (75 MHz, D₂O, 298K) δ = 172.9, (172.7), (170.8), 170.2, 69.6, (68.0), (58.7), 57.5, 55.9, (54.3), (52.7), 52.2, (39.6), 37.8, 22.1, (21.5) ppm; MS (ES) calc. for C₈H₁₃NNaO₄ (M + Na)⁺: 210.07. Found (M + Na)⁺: 209.73.

4-O-(α-D-Galactopyranosyl)-N-acetyl-trans-4-hydroxy-L-proline methyl ester (2):

The general preparation method was followed for *O*-debenzylation of **9a** (0.153g, 0.215 mmol) to yield **2** as a clear oil (0.075 g, 0.215 mmol) (quant.): [α]_D²⁵ = +105.9° (*c* 0.3 CH₃OH); ¹H NMR (500 MHz, D₂O, 338K) δ = 5.07 (d, 0.8H, H₁, J = 1.9Hz), [5.03, broad d, 0.2H, H₁], [4.83, dd, 0.2H, Pro_α^{cis}, J = 7.1Hz, J = 7.6Hz], 4.54-4.61 (m, 1.6H, Pro_γ^{trans}, Pro_α^{trans}), [4.51, dddd, 0.2H, Pro_γ^{cis}, J = 1.4Hz, J = 5.0Hz, J = 5.0Hz, J = 7.5Hz], 3.98-4.01 (m, 1H, H₅), 3.93-3.98 (m, 1H, H₃, J = 6.0Hz, J = 6.2Hz), 3.71-3.89 (m, 8.8H, H₂, H₄, H_{6a}, H_{6b}, Pro_{δ1}, Pro_{δ2}^{trans}, -CO₂CH₃), [3.59, dd, 0.2H, Pro_{δ2}^{cis}, J = 5.0Hz, J = 12.9Hz], [2.67, ddd, 0.2H, Pro_{β1}^{cis}, J = 7.5Hz, J = 7.6Hz, J = 13.4Hz], 2.58 (ddd, 0.8H, Pro_{β1}^{trans}, J = 5.3Hz, J = 7.9Hz, J = 13.3Hz), [2.43, ddd, 0.2H, Pro_{β2}^{cis}, J = 5.0Hz, J = 7.1Hz, J = 13.4Hz], 2.23 (ddd, 0.8H, Pro_{β2}^{trans}, J = 5.0Hz, J = 8.3Hz, J = 13.3Hz), 2.13 (s, 2.4H, -

NCOCH_3^{trans}), [2.01, s, 0.6H, $-\text{NCOCH}_3^{cis}$]; ^{13}C NMR (75 MHz, D_2O , 298K) δ = 179.6, (179.2), (178.8), 178.2, 102.9, (102.8), 81.3, (79.4), 76.6, (76.5), 74.3, (74.2), 73.0, (72.9), 66.3, (64.1), 62.9, (58.4), 58.3, 58.0, (56.2), 53.9, (41.8), 40.4, 26.3, (25.7) ppm; MS (ES) calc. for $\text{C}_{14}\text{H}_{23}\text{NNaO}_9$ (M + Na) $^+$: 372.13. Found (M + Na) $^+$: 372.03.

4-O-(β -D-Galactopyranosyl)-N-acetyl-*trans*-4-hydroxy-L-proline methyl ester (3):

The general preparation method was followed for *O*-debenzylation of **9b** (0.100g, 0.141 mmol) to yield **3** as a clear oil (0.049 g, 0.141 mmol) (quant.): $[\alpha]_{\text{D}}^{25} = -33.6^\circ$ (*c* 0.5 CH_3OH); ^1H NMR (500 MHz, D_2O , 338K) δ = [4.81, dd, 0.2H, Pro_α^{cis} , $J = 7.3\text{Hz}$, $J = 7.8\text{Hz}$], 4.66-4.73 (m, 0.8H, $\text{Pro}_\gamma^{trans}$), [4.59-4.65, m, 0.2H, Pro_γ^{cis}], 4.55 (dd, 0.8H, $\text{Pro}_\alpha^{trans}$, $J = 8.2\text{Hz}$, $J = 8.2\text{Hz}$), 4.49 (d, 0.8H, H_1^{trans} , $J = 8.0\text{Hz}$), [4.47, d, 0.2H, H_1^{cis} , $J = 8.3\text{Hz}$], 3.92-3.99 (m, 1.2H, H_4 , $\text{Pro}_{\delta 1}^{cis}$), 3.86-3.92 (m, 1.6H, $\text{Pro}_{\delta 1}^{trans}$, $\text{Pro}_{\delta 2}^{trans}$), [3.85, s, 0.6H, $-\text{CO}_2\text{CH}_3^{cis}$], 3.74-3.83 (m, 4.4H, H_{6a} , H_{6b} , $-\text{CO}_2\text{CH}_3^{trans}$), 3.64-3.74 (m, 2H, H_3 , H_5), [3.61, dd, 0.2H, $\text{Pro}_{\delta 2}^{cis}$, $J = 4.7\text{Hz}$, $J = 12.9\text{Hz}$], 3.52 (dd, 0.8H, H_2^{trans} , $J = 8.6\text{Hz}$, $J = 9.1\text{Hz}$), [3.82-3.53, m, 0.2H, H_2^{cis}], [2.60-2.69, m, 0.2H, $\text{Pro}_{\beta 1}^{cis}$], 2.55 (ddd, 0.8H, $\text{Pro}_{\beta 1}^{trans}$, $J = 2.4\text{Hz}$, $J = 8.2\text{Hz}$, $J = 13.4\text{Hz}$), [2.42, ddd, 0.2H, $\text{Pro}_{\beta 2}^{cis}$, $J = 5.6\text{Hz}$, $J = 6.6\text{Hz}$, $J = 13.4\text{Hz}$], 2.21 (ddd, 0.8H, $\text{Pro}_{\beta 2}^{trans}$, $J = 8.2\text{Hz}$, $J = 10.0\text{Hz}$, $J = 13.4\text{Hz}$), 2.14 (s, 2.4H, $-\text{NCOCH}_3^{trans}$), [2.02, s, 0.6H, $-\text{NCOCH}_3^{cis}$]; ^{13}C NMR (75 MHz, D_2O , 298K) δ = 179.5, (179.1), (178.8), 178.2, (106.8), 106.6, 82.6, (81.3), (80.3), 80.2, (77.65), 77.61, (75.54), 75.48, 73.5, 65.9, (63.7), 62.6, 59.3, (58.4), 58.0, (57.3), (53.9), (41.2), 39.6, 26.3, (25.7) ppm; MS (ES) calc. for $\text{C}_{14}\text{H}_{23}\text{NNaO}_9$ (M + Na) $^+$: 372.13. Found (M + Na) $^+$: 372.03.

***N*-Acetyl-*cis*-4-hydroxy-*L*-proline methyl ester (4):** Compound **10** (0.050 g, 0.204 mmol, 1.0 eq) was dissolved in DCM (5 mL) and cooled to 0°C, before addition of TFA (5 mL). After stirring for 1.5 hrs, the solution was co-distilled with toluene (3 x 6 mL) and concentrated under reduced pressure. The residue was then re-dissolved in methanol (2 mL), followed by the addition of triethylamine (0.112 mL, 0.816 mmol, 4.0 eq) and acetic anhydride (0.113 mL, 1.224 mmol, 6.0 eq). The reaction was stirred for 2 hours before concentration, and purification with 9:1 ethyl acetate/methanol to yield **4** as a clear oil (0.036 g, 0.192 mmol) (94.7%): $[\alpha]_D^{25} = +89.9^\circ$ (*c* 0.7 CH₃OH); ¹H NMR (500 MHz, D₂O, 298K) $\delta = [4.82, \text{dd}, 0.3\text{H}, \text{Pro}_\alpha^{\text{cis}}, J_{\alpha, \beta 2} = 3.2\text{Hz}, J_{\alpha, \beta 1} = 6.7\text{Hz}], 4.65 (\text{dd}, 0.7\text{H}, \text{Pro}_\alpha^{\text{trans}}, J_{\alpha, \beta 2} = 2.4\text{Hz}, J_{\alpha, \beta 1} = 9.6\text{Hz}), 4.57 (\text{dddd}, 0.7\text{H}, \text{H}_\gamma^{\text{trans}}, J_{\gamma, \delta 2} = 1.0\text{Hz}, J_{\beta 2, \gamma} = 2.5\text{Hz}, J_{\gamma, \delta 1} = 4.6\text{Hz}, J_{\beta 1, \gamma} = 4.6\text{Hz}), [4.51-4.54, \text{m}, 0.3\text{H}, \text{H}_\gamma^{\text{cis}}], 3.84 (\text{dd}, 0.7\text{H}, \text{Pro}_{\delta 1}^{\text{trans}}, J_{\delta 1, \delta 2} = 11.6\text{Hz}), [3.80, \text{s}, 0.9\text{H}, -\text{CO}_2\text{CH}_3^{\text{cis}}], 3.77 (\text{s}, 2.1\text{H}, -\text{CO}_2\text{CH}_3^{\text{trans}}), [3.64, \text{dd}, 0.3\text{H}, \text{Pro}_{\delta 1}^{\text{cis}}, J_{\gamma, \delta 1} = 4.4\text{Hz}, J_{\delta 1, \delta 2} = 13.4\text{Hz}], 3.61 (\text{dd}, 0.7\text{H}, \text{Pro}_{\delta 2}^{\text{trans}}), [3.47, 0.3\text{H}, \text{Pro}_{\delta 2}^{\text{cis}}, J_{\gamma, \delta 2} = 1.0\text{Hz}], 2.41-2.52 (\text{m}, 1.3\text{H}, \text{Pro}_{\beta 1}, \text{Pro}_{\beta 2}^{\text{cis}}), 2.22 (\text{ddd}, 0.7\text{H}, \text{Pro}_{\beta 2}^{\text{trans}}, J_{\beta 1, \beta 2} = 14.2\text{Hz}), 2.14 (\text{s}, 2.1\text{H}, -\text{NCOCH}_3^{\text{trans}}), [2.08, \text{s}, 0.9\text{H}, -\text{NCOCH}_3^{\text{cis}}]; ¹³C NMR (75 MHz, D₂O, 298K) $\delta = (174.7), 174.5, (174.4), 173.8, 69.9, (68.8), (59.8), 57.9, 56.1, (54.9), (53.5), 53.3, (38.8), 37.3, 21.7, (21.5)$ ppm; MS (ES) calc. for C₈H₁₃NNaO₄ (M + Na)⁺: 210.07. Found (M + Na)⁺: 209.73.$

4-*O*-(α -D-Galactopyranosyl)-*N*-acetyl-*cis*-4-hydroxy-*L*-proline methyl ester (5): The general preparation method was followed for *O*-debenzylation of **12a** (0.087g, 0.123 mmol) to yield **5** as a clear oil (0.043 g, 0.123 mmol) (quant.): $[\alpha]_D^{25} = +93.6^\circ$ (*c* 0.3 CH₃OH); ¹H NMR (500 MHz, D₂O, 298K) $\delta = 4.90 (\text{d}, 0.75\text{H}, \text{H}_1^{\text{trans}}, J = 3.8\text{Hz}), [4.88, \text{d}, 0.25\text{H}, \text{H}_1^{\text{cis}}, J = 3.8\text{Hz}], [4.69, \text{dd}, 0.25\text{H}, \text{Pro}_\alpha^{\text{cis}}, J = 1.1\text{Hz}, J = 9.2\text{Hz}], 4.53 (\text{d}, 0.75\text{H},$

Pro $_{\alpha}^{trans}$, $J = 4.3\text{Hz}$, $J = 7.8\text{Hz}$), 4.34-4.38 (m, 0.75H, Pro $_{\gamma}^{trans}$), [4.30-4.33, m, 0.25H, Pro $_{\gamma}^{cis}$], 3.78-3.87 (m, 2.5H, H $_3^{trans}$, H $_5$, Pro $_{\delta 1}^{trans}$), [3.71-3.75, m, 0.25H, H $_3^{cis}$], 3.45-3.71 (m, 8.25H, H $_2$, H $_4$, H $_{6a}$, H $_{6b}$, Pro $_{\delta 1}^{cis}$, Pro $_{\delta 2}$, -CO $_2$ CH $_3$), [2.48-2.55, m, 0.25H, Pro $_{\beta 1}^{cis}$], 2.26-2.37 (m, 1.75H, Pro $_{\beta 1}^{trans}$, Pro $_{\beta 2}$), 1.98 (s, 2.25H, -NCOCH $_3^{trans}$), [1.92, s, 0.75H, -NCOCH $_3^{cis}$]; ^{13}C NMR (75 MHz, D $_2$ O, 298K) $\delta =$ (174.39), (174.34), 174.2, 173.8, 99.1, (98.7), 77.9, (76.3), 72.1, (71.6), 69.75, 69.69, (69.5), 68.4, (68.3), 61.7, (61.3), (59.8), 57.9, 54.5, (53.6), 53.4, (36.4), 35.1, 21.64, (21.59) ppm; MS (ES) calc. for C $_{14}$ H $_{23}$ NNaO $_9$ (M + Na) $^+$: 372.13. Found (M + Na) $^+$: 372.03.

4-*O*-(β -D-Galactopyranosyl)-*N*-acetyl-*cis*-4-hydroxy-*L*-proline methyl ester (6**):** The general preparation method was followed for *O*-debenzylation of **12b** (0.047g, 0.066 mmol) to yield **6** as a clear oil (0.023 g, 0.066 mmol) (quant.): $[\alpha]_{\text{D}}^{25} = -40.6^{\circ}$ (*c* 0.7 CH $_3$ OH); ^1H NMR (500 MHz, D $_2$ O, 338K) $\delta =$ [5.09, dd, 0.3H, Pro $_{\alpha}^{cis}$, $J = 1.3\text{Hz}$, $J = 9.2\text{Hz}$], 4.86-4.98 (m, 1.7H, Pro $_{\gamma}$, Pro $_{\alpha}^{trans}$), 4.73 (d, 0.7H, H $_1^{trans}$, $J = 8.2\text{Hz}$), [4.71, d, 0.3H, H $_1^{cis}$, $J = 8.2\text{Hz}$], 4.20-4.25 (m, 1H, H $_4$), 4.16 (dd, 0.7H, Pro $_{\delta 1}^{trans}$, $J = 4.8\text{Hz}$, $J = 11.8\text{Hz}$), 4.02-4.13 (m, 5.7H, H $_5$, H $_{6a}$, Pro $_{\delta 2}^{trans}$, -CO $_2$ CH $_3$), 3.90-3.99 (m, 2.6H, H $_3$, H $_{6b}$, Pro $_{\delta 1}^{cis}$, Pro $_{\delta 1}^{cis}$), 3.75 (dd, 0.7H, H $_2^{trans}$, $J = 7.5\text{Hz}$, $J = 9.6\text{Hz}$), [3.71, dd, 0.3H, H $_2^{cis}$, $J = 8.0\text{Hz}$, $J = 10.0\text{Hz}$], [2.90-2.96, m, 0.3H, Pro $_{\beta 1}^{cis}$], 2.79-2.90 (m, 1H, Pro $_{\beta 1}^{trans}$, Pro $_{\beta 2}^{cis}$), 2.67-2.74 (m, 0.7H, Pro $_{\beta 2}^{trans}$), 2.43 (s, 2.1H, -NCOCH $_3^{trans}$), [2.34, s, 0.9H, -NCOCH $_3^{cis}$]; ^{13}C NMR (75 MHz, D $_2$ O, 298K) $\delta =$ (175.0), 174.7, (174.3), 173.8, 102.3, (102.0), 78.3, (77.0), 75.6, (75.5), 72.9, (71.2), 71.1, 68.9, 61.3, (59.9), 58.0, 54.2, (53.7), 53.5, (52.9), (37.6), 35.9, 21.7, (21.5) ppm; MS (ES) calc. for C $_{14}$ H $_{23}$ NNaO $_9$ (M + Na) $^+$: 372.13. Found (M + Na) $^+$: 372.03.

5.7 References

- 1) a) Weerepana, E.; Imperiali, B. *Glycobiology* **2006**, *16*, 91R-101R; b) Messner, P. *J. Bacteriol.* **2004**, *186*, 2517-2519; c) Eichler, J.; Adams, M. W. *Microbiol. Mol. Biol. Rev.* **2005**, *69*, 393-425.
- 2) a) Varki, A. *Glycobiology* **1993**, *3*, 97-130; b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720; c) Reuter, G.; Gabius, H. J. *Cell. Mol. Life Sci.* **1999**, *55*, 368-422; d) Petrescu, A. J.; Milac, A. L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103-114.
- 3) a) Mer, G.; Hietter, H.; Lefèvre, J.-F. *Nat. Struc. Biol.* **1996**, *3*, 45-53; b) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 97-101.
- 4) a) Fisher, J. F.; Harrison, A. W.; Bundy, G. L.; Wilkinson, K. F.; Rush, B. D.; Ruwart, M. J. *J. Med. Chem.* **1991**, *34*, 3140-3143; b) Mehta, S.; Meldal, M.; Duus, J. O.; Bock, K. *J. Chem. Soc., Perkin Trans. 1*, **1999**, 1445-1451.
- 5) Imperiali, B. *Acc. Chem. Res.* **1997**, *30*, 452-459.
- 6) Bilsky, E. J.; Egleton, R. D.; Mitchell, S. A.; Palian, M. A.; Davis, P.; Huber, J. D.; Jones, H.; Yamamura, H. I.; Janders, J.; Davis, T. P.; Porreca, F.; Hraby, V. J.; Polt, R. *J. Med. Chem.* **2000**, *43*, 2586-2590.
- 7) a) Meyer, B.; Moller, H. *Top. Curr. Chem.* **2007**, *267*, 187-251; b) Specker, D.; Wittman, V. *Top. Curr. Chem.* **2007**, *267*, 65-107.

8) a) Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421-8425; b) O'Connor, S. E.; Imperiali, B. *Chem. Biol.* **1998**, *5*, 427-437; c) Pao, Y.-L.; Wormarld, M. R.; Dwek, R. A.; Lellouch, A. C. *Biochem. Biophys. Res. Commun.* **1996**, *219*, 157-162; d) Liang, R.; Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1995**, *117*, 10395-10396.

9) a) Lamport, D. T. A. *Nature* **1967**, *216*, 1322-1324; b) Lamport, D. T. A. *Recent Adv. Phytochem.* **1977**, *11*, 79-115; c) Knox, R. B.; Clarke, A.; Harrison, S.; Smith, P.; Marchalonis, J. J. *Proc. Natl. Acad. Sci. U.S.A.* **1976**, *73*, 2788-2792; d) Showalter, A. M. *Plant Cell* **1993**, *5*, 9-23; e) Bowles, D. J. *Annu. Rev. Biochem.* **1990**, *59*, 873-907; f) Josè-Estanyol, M.; Puigdomènech, P. *Plant Physiol. Biochem.* **2000**, *38*, 97-108; g) Sommer-Knudsn, J.; Bacic, A.; Clarke, A. E. *Phytochemistry* **1998**, *47*, 483-497; h) Kieliszewski, M. J. *Phytochemistry* **2001**, *57*, 319-323; i) Khashimova, Z. S. *Chem. Nat. Comput.* **2003**, *39*, 229-236.

10) Mauger, A. B. *J. Nat. Prod.* **1996**, *59*, 1205-1211.

11) a) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953-4963; b) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 397-412; c) Stein, R. L. *Adv. Protein Chem.* **1993**, *44*, 1-24.

12) Fischer, G.; Schmid, F. X. *Biochemistry* **1990**, *29*, 2205-2212.

13) a) Sapse, A. M.; Mallah-Levy, L.; Daniels, S. B.; Erickson, B. W. *J. Am. Chem. Soc.* **1987**, *109*, 3526-3529; b) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 698-705; c) Rose, G. D.; Gierasch, L. M.; Smith, J. A. *Adv. Protein Chem.* **1985**, *37*, 1-109.

14) a) Halab, L.; Gosselin, F.; Lubell, W. D. *Biopolymers* **2000**, *55*, 101-122; b) Halab, L.; Lubell, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2474-2484; c) Kakinoki, S.; Hirano, Y.; Oka, M. *Polym. Bull.* **2005**, *53*, 109-118.

15) a) Berg, R. A.; Prockop, D. J. *Biochem. Biophys. Res. Commun.* **1973**, *52*, 115-120; b) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *Biochemistry* **2000**, *39*, 14960-14967.

16) a) Van Holst, G. J.; Varner, J. E. *Plant Physiol.* **1984**, *74*, 247-251; b) Shpak, E.; Barbar, E.; Leykam, J. F.; Kieliszewski, M. J. *J. Biol. Chem.* **2001**, *276*, 11272-11278; c) Ferris, P. J.; Waffenschmidt, S.; Umen, J. G.; Ishida, K.; Kubo, T.; Lau, J.; Goodenough, U. W. *Plant Cell* **2005**, *17*, 597-615.

17) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777-778.

18) Hodges, J. A.; Raines, R. T. *Org. Lett.* **2006**, *8*, 4695-4697.

19) a) Zimmerman, S. S.; Scheraga, H. A. *Macromolecules* **1976**, *9*, 408-416; b) Zimmerman, S. S.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 811-843.

20) Taylor, C. M.; Hardré, R.; Edwards, P. J. B.; Park, J. H. *Org. Lett.* **2003**, *5*, 4413-4416.

21) Eberhardt, E. S.; Panasik, N.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266.

- 22) DeRider, M. L.; Wilkens, S. J.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497–2505.
- 23) Renner, S.; Alefelder, J. H.; Bae, N.; Budisa, R.; Huber, L.; Moroder, L. *Angew. Chem. Int. Ed.* **2001**, *40*, 923-925.
- 24) Cadamuro, S. A.; Reichold, R.; Kusebauch, U.; Musiol, H.-J.; Renner, C.; Tavan, P.; Moroder, L. *Angew. Chem. Int. Ed.* **2008**, *47*, 2143-2146.
- 25) Sonntag, L.-S.; Schweizer, S.; Ochsenfeld, C.; Wennemers, H. *J. Am. Chem. Soc.* **2006**, *128*, 14697-14703.
- 26) Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A.; Raines, R. T.; Krow, G. R. *J. Org. Chem.* **2004**, *69*, 8565-8573.
- 27) Shoulders, M. D.; Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2006**, *128*, 8112-8113.
- 28) Pao, Y.-L.; Wormarld, M. R.; Dwek, R. A.; Lellouch, A. C. *Biochem. Biophys. Res. Commun.* **1996**, *219*, 157-162.
- 29) a) Owens, N. W.; Braun, C.; Schweizer, F. *J. Org. Chem.* **2007**, *72*, 4635-4643; b) Zhang, K.; Schweizer, F. *Synlett* **2005**, *20*, 3111-3115; c) Zhang, K.; Schweizer, F. *Carb. Res.* **2009** (In Press); d) Zhang, K.; Teklebrhan, R. B.; Schreckenbach, G.; Wetmore, S.; Schweizer, F. *J. Org. Chem.* **2009** (In Press).
- 30) Owens, N. W.; Braun, C.; O’Neil, J. D.; Marat, K.; Schweizer, F. *J. Am. Chem. Soc.* **2007**, *129*, 11670-11671.

- 31) Beausoleil, E.; Sharma, R.; Michnick, S. W.; Lubell, W. D. *J. Org. Chem.* **1998**, *63*, 6572-6578.
- 32) Panasik Jr., N.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262-269.
- 33) Jenkins, C. L.; McCloskey, A. I.; Guzei, I. A.; Eberhardt, E. S.; Raines, R. T. *Biopolymers* **2005**, *80*, 1-8.
- 34) Vitagliano, L.; Berisio, R.; Mastrangelo, A.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2001**, *10*, 2627-2632.
- 35) Improta, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, *123*, 12568-12577.
- 36) Taylor, C. M.; Hardré, R.; Edwards, P. J. B. *J. Org. Chem.* **2005**, *70*, 1306-1315.
- 37) Shamala, N.; Row, T. N. G.; Venkatesan, K. *Acta Crystallogr. B* **1976**, *B32*, 3267-3270.
- 38) a) Eberhardt, E. S.; Loh, S. N.; Hinck, A. P.; Raines, R. T. *J. Am. Chem. Soc.* **1992**, *114*, 5437-5439; b) Eberhardt, E. S.; Loh, S. N.; Raines, R. T. *Tetrahedron Lett.* **1993**, *34*, 3055-3056.
- 39) a) Matsuzaki, T.; Iitaka, Y. *Acta Crystallogr.* **1971**, *B27*, 507-516; b) Liang, G.-B.; Rito, C. J.; Gellman, S. H. *Biopolymers* **1992**, *32*, 293-301.
- 40) Griffin, F. K.; Paterson, D. E.; Murphy, P. V.; Taylor, R. J. K. *Eur. J. Org. Chem.* **2002**, 1305-1322.

- 41) a) Krimm, S.; Bandekar, J. in *Advances in Protein Chemistry*, Vol. 38 (Eds: Anfinsen, C. B.; Edsall, J. T.; Richards, F. M.), Academic Press, New York, **1986**, pp. 181-364; b) Jackson, M.; Mantsch, H. H. *Crit. Rev. Biochem. Molec. Biol.* **1995**, *30*, 95-120.
- 42) GOESY is a selective 1D NOESY experiment, see: Stonehouse, J.; Adell, P.; Keeler, J.; Shaka, A. J. *J. Am. Chem. Soc.* **1994**, *116*, 6037-6038.
- 43) a) Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc.* **1973**, *95*, 5725-5729; b) Cai, M.; Huang, Y.; Liu, J.; Krishnamoorthi, R. *J. Biomol. NMR* **1995**, *6*, 123-128; c) Haasnoot, C. A. G.; De Leeuw, F. A. A. M.; De Leeuw, H. P. M.; Altona, C. *Biopolymers* **1981**, *20*, 1211-1245.
- 44) Song, I. K.; Kang Y. K. *J. Phys. Chem. B* **2006**, *110*, 1915-1927.
- 45) Benzi, C.; Improtà, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 341-350.
- 46) Lam, J. S. W.; Koo, J. C. P.; Hudáky, I.; Varro, A.; Papp, J. G.; Penke, B.; Csizmadia, I. G. *J. Mol. Struct.* **2003**, *666*, 285-289.
- 47) Friebolin, H. in *Basic One- and Two-Dimensional NMR Spectroscopy*, 3rd ed., Wiley-VCH, New York, 1998.
- 48) Quijcho, F. A. *Ann. Rev. Biochem.* **1986**, *55*, 287-315.
- 49) Vyas, N. K.; Vyas, M. N.; Quijcho, F. A. *Science* **1988**, *242*, 1290-1295.

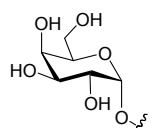
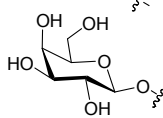
- 50) Stauffer, D. A.; Barrans Jr., R. E.; Dougherty, D. A. *J. Org. Chem.* **1990**, *55*, 2762-2767.
- 51) Radzicka, A.; Pedersen, L.; Wolfenden, R. *Biochemistry* **1988**, *27*, 4538-4541;
- 52) Troganis, A.; Gerothanassis, I. P.; Athanassiou, Z.; Mavromoustakos, T.; Hawkes, G. E.; Sakarellos, C. *Biopolymers* **2000**, *53*, 72-83.
- 53) Beausoleil, E.; Lubell, W. D. *J. Am. Chem. Soc.* **1996**, *118*, 12902-12908.
- 54) However, there may be a slight destabilization of the pucker as Alabugin *et al.* have proposed that there is an inverse relationship between hyperconjugative acceptor ability of the C-X bond and the electronegativity of X: Alabugin, I. V.; Zeidan, T. A. *J. Am. Chem. Soc.* **2002**, *124*, 3175-3185.
- 55) Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188-1194.
- 56) Bagger, H. L.; Fuglsang, C. C.; Westh, P. *Eur. Biophys. J.* **2006**, *35*, 367-371.
- 57) Pauling, L. in *The Nature of the Chemical Bond*, 3rd ed., Cornell University Press, Ithaca, NY, **1960**, pp. 281-282.
- 58) a) Wormald, M. R.; Dwek, R. A. *Structure* **1999**, *7*, R155-R160; b) White, K. D.; Cummings, R. D.; Waxman, F. J. *J. Immunol.* **1997**, *158*, 426-435.
- 59) Burgi, H. B.; Dunitz, J. D.; Shefter, E. *J. Am. Chem. Soc.* **1972**, *95*, 5065-5067.
- 60) Hinderaker, M. P.; Raines, R. T. *Prot. Sci.* **2003**, *12*, 1188-1194.

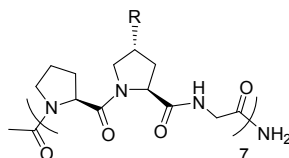
61) O'Hagan, D.; Bilton, C.; Howard, J. A. K.; Knight, L.; Tozer, D. J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 605-607.

Chapter 6: The Effects of 4*R*-Hydroxy-L-Proline *O*-Glycosylation on the Stability of Collagen Model Peptides

Neil W. Owens^{*#§∞} and Frank Schweizer^{*§∞}

^{*}Designed research; [#]Performed research; [§]Analyzed data; [∞]Wrote the paper

Collagen Model Peptide	T_m
1: R = OH	42
2: R = 	44
3: R = 	40



6.1 Abstract

The modification of 4*R*-hydroxy-L-proline (Hyp) in collagen model peptides has frequently been used to probe the role of Hyp in the stability of collagen and for the development of collagen-based biomaterials. Here, we report on the effects of Hyp *O*-glycosylation on the stability of collagen model peptides. Glycosylation is a common post-translational modification of proteins known to affect protein hydration, conformation, and thermal stability. Collagen model peptides were synthesized using a fragment condensation approach. We found Hyp glycosylation does not preclude triple helix formation, and the melting temperature (T_m) of the collagen model peptides was influenced by the anomeric linkage of the glycan; α -*O*-linked galactosylation led to a slight stabilization of the triple helix, while β -*O*-galactosylation had a slight destabilizing effect. The reasons behind this effect remain unclear, but are likely due to differences in hydration.

6.2 Introduction

Collagen is the predominant structural protein in humans and other animals,¹ where it forms extended fibrous networks that give tissues their structural integrity.² These collagen fibres are composed of bundles of tropocollagen molecules, which have a characteristic tertiary structure of three polypeptide strands, each in a left-handed polyproline II (PPII) conformation, twisted into a right-handed triple helix.³⁻⁵ The polypeptide strands have a repeating X_{aa} - Y_{aa} -Gly sequence, where X_{aa} is often L-proline (Pro) and Y_{aa} is often 4*R*-hydroxy-L-proline (Hyp). Typically, there are ~300 repeats of the X_{aa} - Y_{aa} -Gly sequence per strand in the collagen molecule.

While much information has been garnered about the structure of collagen, a complete understanding of the stabilization of the collagen triple helix has yet to be achieved. However, analysis of collagen model peptides has shed light on the critical requirements for collagen stability, especially on the role of Hyp.⁶⁻⁸ Hyp content correlates well to the stability of the collagen triple helix,⁹⁻¹¹ for several reasons: it seems that through a stereoelectronic effect Hyp stabilizes the *N*-terminal *trans* amide conformation,¹² which in turn stabilizes the PPII conformation present in each strand of the triple helix. Also, experimental^{12,13} and computational^{14,15} methods indicate that Hyp adopts a C^γ -exo conformation, and recent studies have emphasized that this predisposes the backbone dihedral angles in the Y_{aa} position for optimal packing of the three polypeptide strands.¹⁶⁻¹⁸ Overall, it has been estimated that Hyp in the Y_{aa} position contributes $0.2 \text{ kcal}\cdot\text{mol}^{-1}$ per tripeptide repeat to the stability to the collagen triple helix,^{9,10,19} such that its presence can offset an adverse pucker in the X_{aa} position.²⁰⁻²²

The contributions of hydrogen bonds formed through the 4-hydroxyl group of Hyp to a water network surrounding the collagen triple helix,²⁴⁻²⁶ became marginalized for several reasons: there is rapid exchange of water molecules bound to collagen with bulk solvent,²⁷ collagen stability increases in anhydrous environments,²⁸ and replacement of Hyp with 4-fluoroproline, which forms very weak hydrogen bonds,²⁹ forms hyperstable triple helices.^{19,30} However, the role of hydration in the stabilization of collagen remains controversial.³¹ Crystal structures of collagen model peptides consistently show that the peptides are significantly hydrated,^{5,16,21-23,25,26,32-35} which has been shown to increase when Hyp replaces Pro in the Y_{aa} position.^{36,37} Recently, it has been proven experimentally that this hydration does contribute thermodynamically to the stabilization of collagen.³⁶ Furthermore, hydrogen bonding interactions involving solvent molecules have been used to explain the increased stability of collagen model peptides in which Hyp has been replaced with threonine (Thr) and aminoproline (Amp).^{38,39} Similarly, the β -D-galactosylation of Thr in Ac-(Gly-Pro-Thr)₁₀-NH₂ collagen model peptides was found to induce triple helix formation, where shielding the triple helix from strong interactions with water molecules was attributed to the carbohydrates, thereby stabilizing interchain N-H to C=O hydrogen bonding interactions.⁴⁰

Glycosylation is a common post-translational modification of proteins^{41,42} known to increase protein hydration and solubility,⁴³ as well as influence protein folding and conformation,⁴¹ protect against proteolytic degradation,⁴⁴ and enhance the thermal stability of proteins.⁴⁵⁻⁴⁶ While protein glycosylation has been broadly attributed to increasing the free energy difference between the unfolded and folded state of the protein,⁴⁷ there is controversy as to whether glycans affect the entropy or enthalpy of the

system and whether these effects mainly influence the folded or unfolded state.⁴⁸ A recent *in silico* thermodynamic analysis of engineered SH3 domain variants by Shental-Bechor and Levy found that with no direct contacts between the glycan and the polypeptide backbone, glycosylation did not change the overall free energy of folded state, but rather enthalpically destabilized the unfolded state;⁴⁷ the bulky glycans reduced the number of favourable intermolecular interactions possible. Overall, the thermodynamic stabilization was more strongly correlated to the number of glycans attached to the protein rather than the size of the glycan.

Hydration plays an important role in the protein folding process of peptides and proteins and the stability of their native structure.⁸⁴ As mentioned above, glycosylation can further affect the hydration of peptides and proteins by changing the hydration levels of the peptide backbone and increasing the solubility of hydrophobic regions; these effects are dependent on the glycan size and stereochemistry.⁸⁵ It is often differences in hydration that can lead to entropic and enthalpic differences between the folded and unfolded states of a glycoprotein and thus their thermodynamic stability.

Typically, glycans are *O*-linked to serine (Ser) and Thr or *N*-linked to asparagine (Asn). The *O*-glycosylation of 5*R*-hydroxy-L-lysine (Hyl) residues,⁴⁹ while relatively uncommon,⁵⁰ is known to occur in collagen molecules.⁵¹ The glycans D-galactose and D-glucosylgalactose β -*O*-linked to Hyl have been implicated in the secretion and assembly of collagen fibrils,^{52,53} embryonic development and cell viability,^{53,54} and the interaction of collagen with protein receptors,^{55,56} however their role is still not fully understood. The *O*-glycosylation of Hyp is not known to occur in humans at all, but is widespread in the plant kingdom in the form of hydroxyproline-rich glycoproteins (HRGPs).^{57,58-60} These

glycoproteins are associated with forming the protective extracellular network of plant cell walls for algae and flowering plants. The HRGPs are characterized by a PPII conformation and short homooligomers of L-arabinofuranose and larger heteropolymers of L-arabinofuranose and D-galactopyranose.^{57,58,61} The functional consequences of glycosylation are still unclear, but indications are that glycosylation contributes to the stability of the PPII conformation in HRGPs.^{58,61,62}

Since Hyp *O*-glycosylation has been associated with stabilizing the PPII conformation in HRGPs, a similar effect may be observed for collagen. While glycosylation has been attributed with increasing the stability of many proteins including variants of collagen,^{40,63} *O*-modification of Hyp has rarely been investigated.⁶⁴ Hyp *O*-glycosylation should have an interesting effect on the hydration of collagen in a similar fashion to other glycoproteins. In previous work, we demonstrated that *O*-galactosylation of Hyp in model amides of the form Ac-X_{aa}-NHMe did not affect the C^γ-exo conformation of Hyp or reduce its *trans* amide isomer content (Table 6.2.1).⁶⁵ We report here on the effects of Hyp *O*-glycosylation on the stability of collagen model peptides.

Table 6.2.1: Effect of Glycosylation on the *Trans/Cis* Amide Isomer Ratio ($K_{trans/cis}$) and Prolyl Pucker^[a]

Model Amide	$K_{trans/cis}$ (25 °C)	Pucker
Ac-L-Pro-NHMe	3.2 ± 0.1	C ^γ -endo
Ac-Hyp-NHMe	3.9 ± 0.1	C ^γ -exo
Ac-Hyp(α-D-Gal)-NHMe	3.7 ± 0.1	C ^γ -exo
Ac-Hyp(β-D-Gal)-NHMe	3.7 ± 0.1	C ^γ -exo

^[a]From Chapter 4.

6.3 Results

We chose to use D-galactose (Gal) as the *O*-glycan because of the similar Hyp-Gal linkage found in HRGPs.^{57,58,61} Monosaccharides as glycans are rarely found in nature, but have shown the capacity to significantly contribute to protein stability.^{45,66} We built the following model peptides: Ac-(Pro-Hyp-Gly)₇-NH₂ (**1**), Ac-[Pro-Hyp(α -D-Gal)-Gly]₇-NH₂ (**2**), and Ac-[Pro-Hyp(β -D-Gal)-Gly]₇-NH₂ (**3**) (Figure 6.3.1).

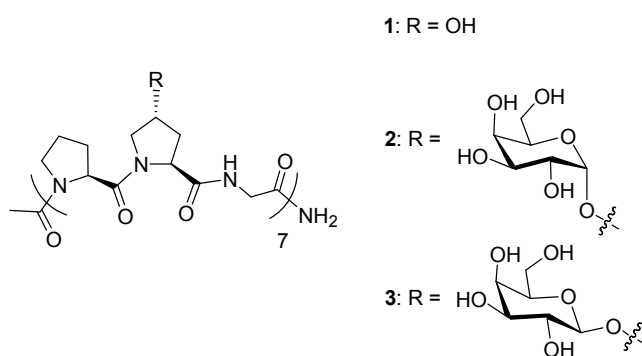
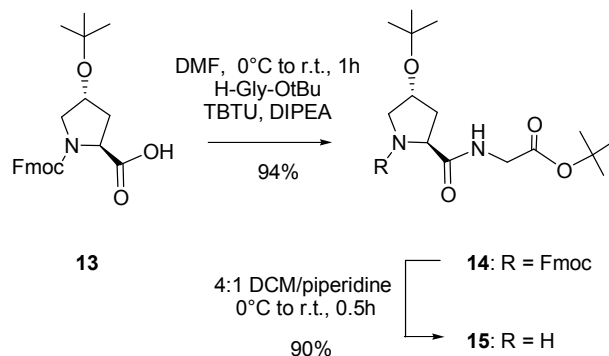


Figure 6.3.1: Collagen model peptides **1-3**.

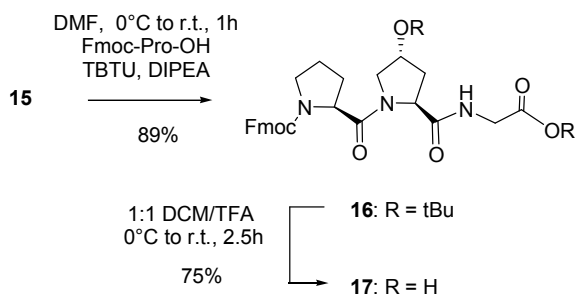
Collagen model peptides of the form (Pro-Hyp-Gly)_n have been shown to have essentially the same 7/2-helix symmetry and 20 Å axial repeat as native collagen, and so represent an appropriate model for studying the triple-helix folding transition of collagen.^{32,34,67} Model peptides with both α - and β -anomeric linkages were made for comparison to explore how the nature of the glycosidic linkage impacts the stability of the triple helix; the anomeric linkage is known to affect peptide and protein conformation.⁶⁸

6.3.1 Synthesis of Ac-(Pro-Hyp-Gly)₇-NH₂ (**1**)



Scheme 6.3.1.1: Synthesis of protected dipeptide **15**.

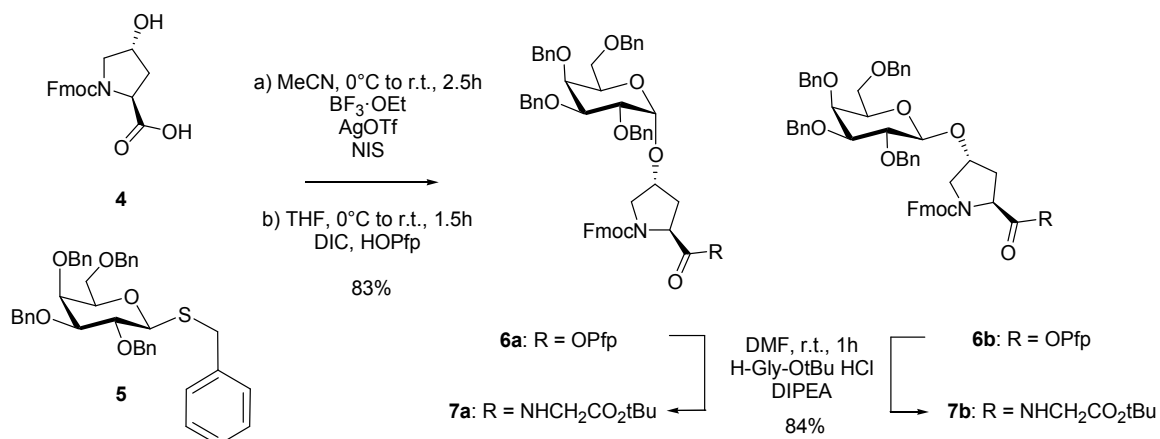
Model peptide **1** was synthesized by condensation of units of Fmoc-Pro-Hyp-Gly-OH (**17**) using rink amide MBHA resin, TBTU, and DIPEA. Tripeptide **17** was synthesized by coupling Fmoc-L-Hyp(*t*Bu)-OH (**13**) with H-Gly-*O**t*Bu using TBTU and DIPEA in DMF to give the dipeptide **14** in 94% yield (Scheme 6.3.1.1). *N*-Fmoc deprotection using 4:1 DCM/piperidine gave the dipeptide **15** in 90% yield. This was followed by coupling of **15** to Fmoc-L-Pro-OH under similar conditions to form the protected tripeptide **16** in 80% yield over two steps (Scheme 6.3.1.2). Then, simultaneous *N*-Boc and *O*-*t*Bu deprotection using 1:1 DCM/TFA gave the tripeptide **17** in 56% overall yield from **13**, suitable for solid-phase peptide synthesis of **1**.



Scheme 6.3.1.2: Synthesis of the building block Fmoc-Pro-Hyp-Gly-OH **17**.

6.3.2 Synthesis of Ac-[Pro-Hyp(α -D-Gal)-Gly]₇-NH₂ (**2**) and Ac-[Pro-Hyp(β -D-Gal)-Gly]₇-NH₂ (**3**)

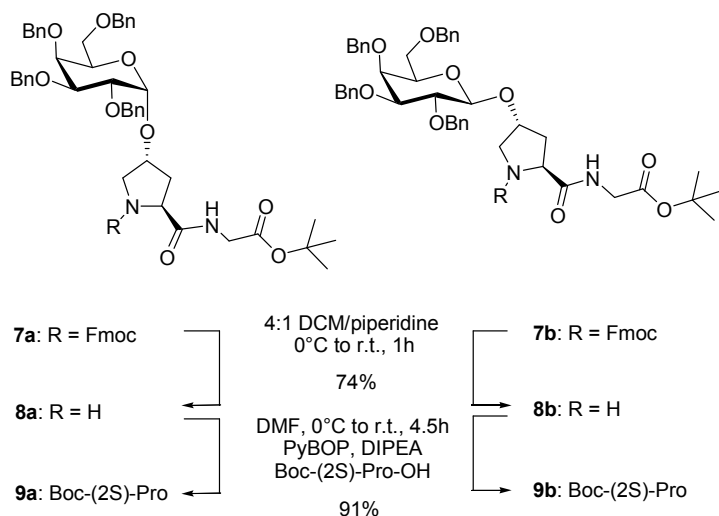
Model peptides **2** and **3** were synthesized by condensation of units of Fmoc-Pro-Hyp(α -D-Gal)-Gly-OH (**12a**) and Fmoc-Pro-Hyp(β -D-Gal)-Gly-OH (**12b**), respectively, using rink amide MBHA resin, TBTU, and DIPEA. Synthesis of tripeptides **12a** and **12b** began with the glycosylation of Fmoc-L-Hyp-OH (**4**) using benzyl 1-thio-2,3,4,6-tetra-*O*-benzyl- β -D-galactopyranoside⁶⁹ (**5**) in the presence of boron trifluoride diethyletherate, *N*-iodosuccinimide and catalytic amounts of silver triflate to yield a mixture of the α - and β -anomers, which could not readily be separated (Scheme 6.3.2.1).



Scheme 6.3.2.1: Synthesis of glycosylated dipeptides **7a** and **7b**.

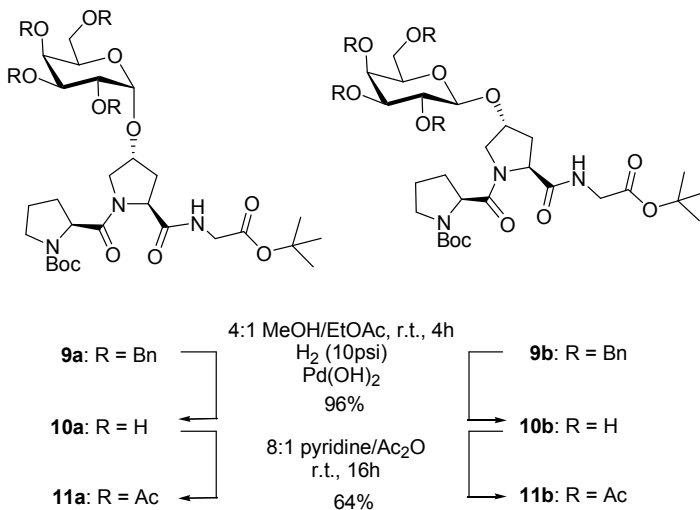
Treatment with *N,N'*-diisopropylcarbodiimide and pentafluorophenol gave the corresponding pentafluorophenyl esters,⁷⁰ which allowed separation of **6a** and **6b** in 83% yield over two steps. Nearly twice as much of the α -anomer as the β -anomer is made under these conditions. Coupling of **6a** and **6b** to H-Gly-*O**t*Bu using DIPEA in DMF gave the corresponding glycosylated dipeptides **7a** and **7b** in 44% and 40% yield.

In separate pathways, *N*-Fmoc deprotection under standard conditions to give **8a** and **8b** was followed by coupling to Boc-L-Pro-OH using PyBOP and DIPEA in DMF to give the protected tripeptides **9a** and **9b** in 67% yield over two steps (Scheme 6.3.2.2).



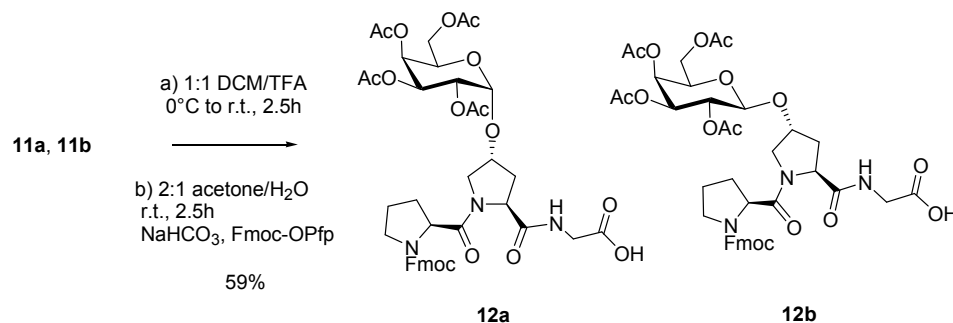
Scheme 6.3.2.2: Synthesis of glycosylated tripeptides **9a** and **9b**.

Removal of the benzyl ether protecting groups by catalytic hydrogenation in methanol gave **10a** and **10b**, which was followed by *O*-acetylation using pyridine/acetic anhydride (8:1) to give **11a** and **11b** (Scheme 6.3.2.3).



Scheme 6.3.2.3: Synthesis of protected tripeptides **11a** and **11b**.

Simultaneous *N*-Boc and *O*-*t*Bu deprotection using 1:1 DCM/TFA followed by installation of the *N*-Fmoc group using Fmoc-OPfp under mild basic conditions gave the target α - (**12a**) and β -*O*-galactosylated (**12b**) tripeptides in 59% yield over two steps and 17% yield overall from **4** (Scheme 6.3.2.4).



Scheme 6.3.2.4: Synthesis of glycosylated tripeptide fragments **12a** and **12b**.

6.3.3 Circular dichroism spectra of 1-3

The far-ultraviolet CD spectra of **1-3** at 25 °C indicate a maximum positive ellipticity near 225 nm and negative ellipticity at 202 nm, which is indicative of a PPII helix and a triple helical fold (Figure 6.3.3.1).³⁰ Both glycosylated model peptides **2** and **3** showed less intensity in the maxima at 225 nm compared to **1**, with ellipticity of 2800 and 3500 compared to 4400 deg·cm²·dmol⁻¹, respectively. Therefore, glycosylation seems to cause some distortion of the PPII conformation, with Ac-[Pro-Hyp(α -D-Gal)-Gly]₇-NH₂ deviating more than Ac-[Pro-Hyp(β -D-Gal)-Gly]₇-NH₂ compared to Ac-(Pro-Hyp-Gly)₇-NH₂.

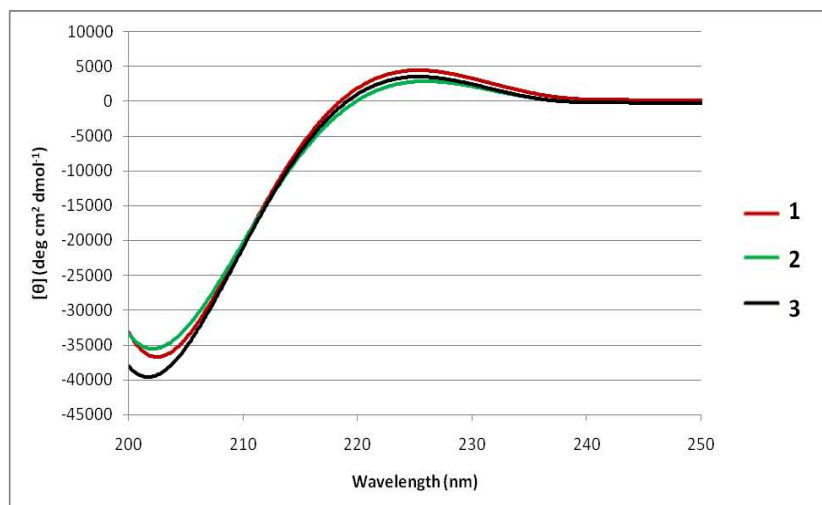


Figure 6.3.3.1: Smoothed CD curves of 1-3 in water at 25 °C.

Comparison of the CD spectra at 25 °C to that of 70 °C shows that the CD curves for 1-3 change significantly, with a decrease in the intensity of both positive and negative maxima to the point that there is no positive band at all (Figures 6.3.3.2-6.3.3.4). The higher temperature CD spectra are indicative of a random coil conformation.³⁰

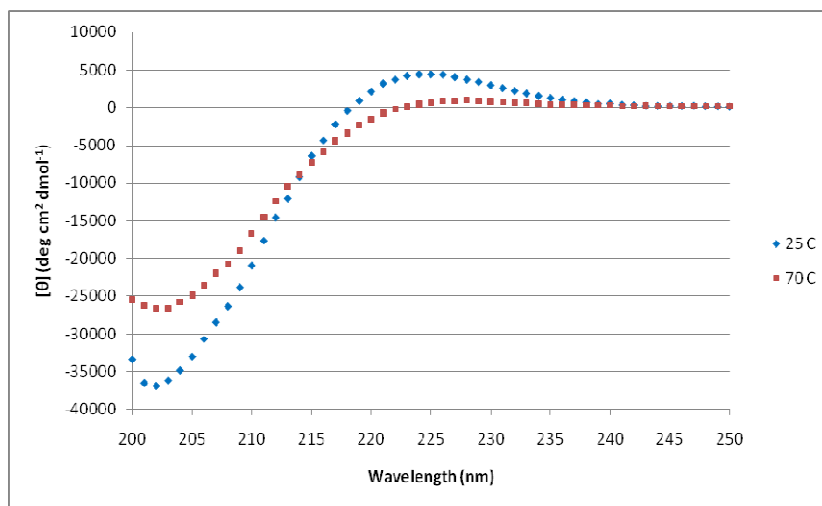


Figure 6.3.3.2: CD curve of 1 in water at 25 °C and 70 °C at 1 nm intervals.

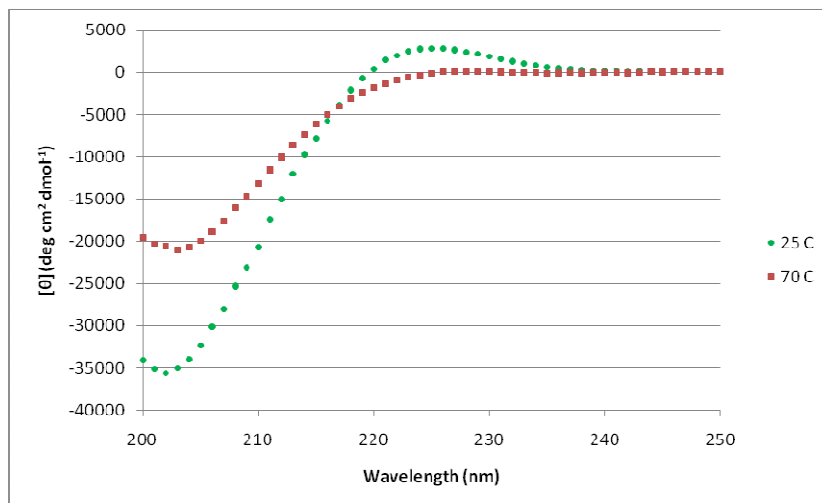


Figure 6.3.3.3: CD curve of **2** in water at 25 °C and 70 °C at 1 nm intervals.

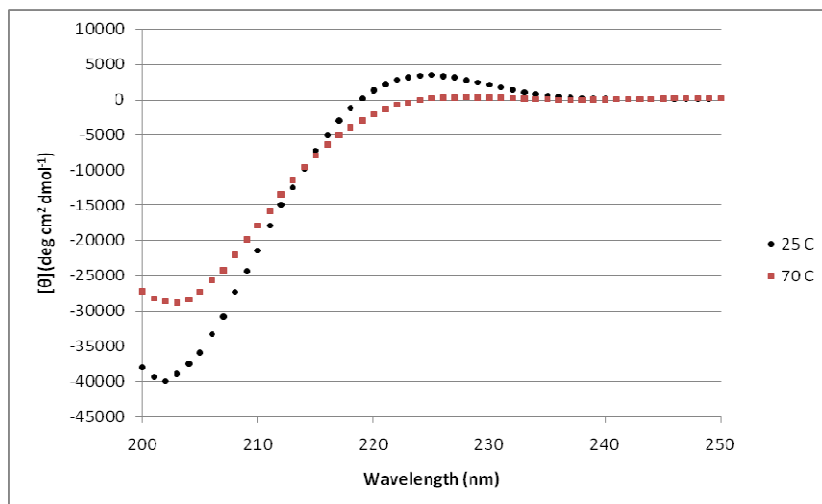


Figure 6.3.3.4: CD curve of **3** in water at 25 °C and 70 °C at 1 nm intervals.

6.3.4 Conformational stability of 1-3

The molar ellipticity as a function of temperature was monitored for **1-3** at 225 nm in 5 °C intervals (Figure 6.3.4.1).

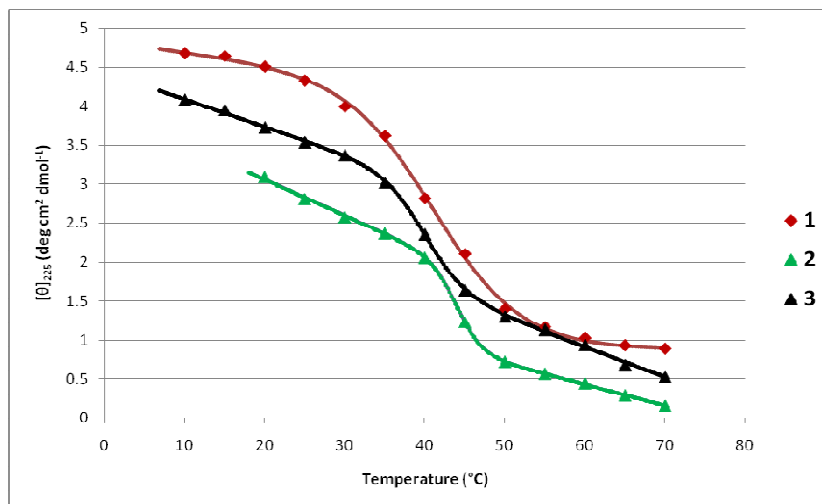


Figure 6.3.4.1: Melting curves of **1-3** in water. The best fit curves of the data points (each point was recorded in duplicate, giving error values of ± 0.2 -5.0%) for **1-3** were calculated using a two-state transition model (See Section 6.6.3).

The melting curve of **1** shows a cooperative transition with a calculated T_m of 42 ± 1 °C, which is consistent with previous work.^{12,71} The melting curves of **2** and **3** showed less cooperative transitions compared to **1**, but are similar to other collagen model peptides.^{22,38,40,72} The T_m values for **2** and **3** are 44 and 40 ± 1 °C, respectively (Table 6.3.4.1). Therefore, it seems that galactose α -linked to Hyp residues in the Y_{aa} position leads to a slight stabilization of the collagen triple helix, while galactose β -linked to Hyp causes a slight destabilization.

Table 6.3.4.1: Effect of α - and β -Galactose on the conformational stability of the collagen triple helix.

Peptide	T_m (± 1 °C)	ΔH_{VH} (KJ·mol ⁻¹)	ΔS_{VH} (J·deg·mol ⁻¹)
Ac-(Pro-Hyp-Gly) ₇ -NH ₂ (1)	42	152 ± 35	482 ± 112
Ac-(Pro-Hyp(α -Gal)-Gly) ₇ -NH ₂ (2)	44	529 ± 193	1670 ± 608
Ac-(Pro-Hyp(β -Gal)-Gly) ₇ -NH ₂ (3)	40	322 ± 80	1030 ± 255

The van't Hoff enthalpy (ΔH_{VH}) and entropy (ΔS_{VH}) values for the triple-helix to single-coil transition were calculated by fitting the unfolding curves to a two-state transition model. The enthalpy difference between the folded and unfolded states increased significantly upon glycosylation (Table 6.3.4.1). The ΔH_{VH} was much higher for both Ac-[Pro-Hyp(α -D-Gal)-Gly]₇-NH₂ (**2**) (ΔH_{VH} of 529 ± 193 KJ·mol⁻¹) and Ac-[Pro-Hyp(β -D-Gal)-Gly]₇-NH₂ (**3**) (ΔH_{VH} of 322 ± 80 KJ·mol⁻¹) compared to Ac-(Pro-Hyp-Gly)₇-NH₂ (ΔH_{VH} of 152 ± 35 KJ·mol⁻¹) (**1**). The ΔH_{VH} value for **1** is within experimental error of similar collagen model peptides,^{22,36,73,74} large error values are known to be associated with parameters derived from the van't Hoff equation.⁷⁴ The anomeric linkage also had an effect on ΔH_{VH} as α -*O*-galactosylation in **2** lead to a larger enthalpy difference relative to β -*O*-galactosylation in **3**. These increases in ΔH_{VH} are mirrored by large increases in ΔS_{VH} for **2** and **3** (ΔS_{VH} of 1670 ± 608 and 1030 ± 255 J·deg·mol⁻¹, respectively) compared to **1** (ΔS_{VH} of 482 ± 112 J·deg·mol⁻¹). Therefore, the solvated folded states of **2** and **3** appear to be much more ordered than the corresponding solvated unfolded peptides than is the case for their unglycosylated counterpart **1**.

6.4 Discussion

Based on the positive CD maxima at 225 nm and negative maxima at 202 nm, as well as the transition observed in the thermal melt experiments, it seems glycosylation of 4*R*-hydroxy-L-proline in collagen model peptides does not preclude triple helix formation. Indication that *O*-glycosylation would not affect collagen triple helix formation came from previous work using Ac-Hyp(α -D-Gal)-NHMe and Ac-Hyp(β -D-Gal)-NHMe model peptides, where it was shown that glycosylation does not affect two major determinants of collagen stability: the prolyl *N*-terminal *trans-cis* amide equilibrium ($K_{trans/cis}$) and the C^{γ} -exo pucker of Hyp that is required in the Y_{aa} position of collagen (Table 6.2.1).⁶⁵ Similarly, Jenkins and coworkers found that *O*-acylation of hydroxyproline in (Pro-Hyp-Gly)₁₀ model peptides did not affect $K_{trans/cis}$ or the C^{γ} -exo pucker and did not preclude triple helix formation.⁶⁴ However, in contrast to *O*-glycosylation, *O*-acylation caused a decrease in T_m (57.5 °C) relative to its unmodified counterpart (69 °C). This was attributed to destabilizing steric interactions between the *O*-acyl group and the X_{aa} position side chain in neighbouring strands. In host-guest model peptide studies, amino acids with bulky side chains such as phenylalanine caused greater destabilization in the Y_{aa} position relative to the X_{aa} position, which could be attributed to similar unfavourable steric interactions.^{7,75} Therefore, it would be predicted that the *O*-glycans linked to Hyp in the Y_{aa} position should cause some steric interaction with the neighbouring strand, which may be responsible for the decreased intensity of the CD positive maxima at 225 nm indicative of a decrease in PPII character (Figure 6.3.3.1). However, any steric repulsion caused by Hyp *O*-glycosylation does not translate into a significant decrease in T_m .

Here, the anomeric linkage of D-galactose to Hyp in the collagen model peptides **1-3** had different effects on the T_m of the triple helix (Figure 6.3.4.1). While Ac-[Pro-Hyp(α -D-Gal)-Gly]₇-NH₂ (**2**) had a slight stabilizing effect (T_m of 44 ± 1 °C) relative to Ac-(Pro-Hyp-Gly)₇-NH₂ (**1**) (T_m of 42 ± 1 °C), Ac-(Pro-Hyp(β -D-Gal)-Gly)₇-NH₂ (**3**) had a slight destabilizing effect (T_m of 40 ± 1 °C). The anomeric linkage alters the orientation of glycans, and therefore may cause different intra- and intermolecular interactions with the collagen peptide backbone or with surrounding solvent molecules. Study of Ac-Hyp(α -D-Gal)-NHMe and Ac-Hyp(β -D-Gal)-NHMe model amides indicates that the α -linked sugar may interact preferentially with the proline side chain, while the β -linked sugar may extend away from the proline side chain into the solvent.⁶⁵ Additional studies are required to understand the effect of the anomeric configuration on the orientation of the glycan with respect to the collagen triple helix structure.

The van't Hoff enthalpy (ΔH_{VH}) and entropy (ΔS_{VH}) differences between the folded and unfolded states were calculated to be much larger for the glycosylated collagen model peptides **2** and **3** compared to the unmodified peptide **1** (Table 6.3.4.1). This is likely due to a combination of factors, which includes changes in the hydration sphere surrounding the triple helix as well as possible interplay between the glycans and the polypeptide backbone in the folded and unfolded states. While the role of hydration in stabilizing the collagen triple helix remains controversial, there is renewed interest in understanding its overall contribution.^{31,34,35-37} Recently, Nishi and coworkers found that replacement of Pro with Hyp in the Y_{aa} position of (Pro-Y_{aa}-Gly)₁₀ caused an increase in hydration upon triple-helix formation according to an increased partial molar volume (V); this translated into a higher enthalpy (ΔH) and entropy difference (ΔS) between the

folded and unfolded states of the model peptides.³⁶ This, together with crystal structures of collagen model peptides which show Hyp participating in interchain water bridges, indicated that the larger ΔH for (Pro-Hyp-Gly)₁₀ was due to an increased number of hydrogen bonds to hydrated water molecules. Glycosylation would be expected to amplify this effect since carbohydrates have multiple hydroxyl groups. Therefore, the increase in ΔH_{VH} in **2** and **3** relative to **1** may be due to an increased number of intermolecular interactions involving water molecules in the folded state. The restricted motion of the water molecules surrounding the triple helix would be expected to offset the increase in ΔH_{VH} because of a large entropic (ΔS_{VH}) penalty, presumably from a much more ordered hydration shell in the folded state;^{36,37} such an increase in ΔS_{VH} is observed (Table 6.3.4.1). Additionally, glycosylation may destabilize the unfolded state of the triple helix by precluding formation of favourable intermolecular interactions or by causing the unfolded state to become more ordered. Computational studies of model glycoproteins by Shental-Bechor and Levy found that glycosylation stabilizes the folded state of a protein mainly by acting on its unfolded state.⁴⁷ Further study of these *O*-glycosylated collagen model peptides should provide more insight into the exact nature of the influence from the Hyp *O*-glycans on the folding of the collagen triple helix.

6.5 Conclusions

Hyp is known to play a critical role in the stability of collagen: the most prevalent protein found in humans and animals. The *O*-glycosylation of Hyp is not known to occur in humans, but does occur in plant HRGPs, where it is implicated in stabilizing the PPII conformation. Here, we found that *O*-glycosylation of Hyp in collagen model peptides

did not preclude formation of the triple helix, but may slightly alter the PPII conformation of each coil. The anomeric linkage of the glycan influenced the stability of the triple helix: α -*O*-galactosylation had a slight stabilizing effect, while β -*O*-galactosylation slightly destabilized the triple helix. The origin of this altered stabilization remains unclear. Glycosylation seems to have a significant effect on the thermodynamics of collagen folding; differences in hydration are likely involved.

The study of collagen model peptides has led to the development of triple helical scaffolds with useful properties: inhibitors of metalloproteases,⁷⁶ thermostable biomaterials,^{12,19,30} and the ability to mimic and modulate the interactions of collagen with its many binding partners.^{77,78} Further study of Hyp *O*-glycosylated collagen model peptides and similar derivatives may lead to the development of novel functionalized collagens which may find use as biomaterials in the fields of biomedicine⁷⁹ and bioengineering,⁸⁰ and could give insight into collagen-related disorders attributed to non-enzymatic glycosylation and glycation.⁸¹

6.6 Experimental Procedures

6.6.1 General Experimental

Protected amino acids were purchased from Bachem (Bachem Bioscience, Inc., King of Prussia, PA). Reagent grade solvents were used without further purification. Thin-Layer chromatography was performed on Si250F pre-coated glass plates of silica gel (250 μ m). Column chromatography was performed on SilicaFlash P60 silica gel (40-63 μ m).

NMR Spectra were obtained using a Bruker AMX-500 NMR spectrometer equipped with a triple resonance (^1H , ^{13}C , ^{15}N) gradient inverse probehead. Spectra were assigned based on 2D COSY and HSQC experiments. For ^1H NMR, minor isomers are listed between square brackets. When multiple isomers are observed, the isomers are listed as H, H' and H''. For ^{13}C NMR, when assigned, carbon peaks for the minor isomer are listed in brackets.

A Waters Micromass ZQ 2000 mass spectrometer was used for electrospray ionization (ESI) mass spectrometry measurements. The matrix assisted laser desorption/ionization (MALDI) data was acquired using a prototype quadrupole-quadrupole-TOF (QqTOF) mass spectrometer with photon pulses from a 20-Hz nitrogen laser (VCL 337ND, Spectra-Physics, Mountain-View, CA) with 300 mJ energy/pulse. The sample solution (~0.5 μL) was loaded onto a stainless steel target with the same volume of the saturated matrix solution (DHB in 1:1 (v/v) acetonitrile-water) and allowed to air dry.

4-O-(2,3,4,6-Tetra-O-benzyl- α,β -D-galactopyranosyl)-N-fluorenylmethoxycarbonyl-*trans*-4-hydroxy-L-proline *N'*-pentafluorophenyl ester (6a,b). Initially, (2*S*,4*R*)-Fmoc-hydroxyproline **4** (1.00 g, 2.83 mmol, 1.0 eq) was dissolved in MeCN (10 mL), followed by addition of 2,3,4,6-tetra-*O*-benzyl-thio- β -D-galactopyranoside **5** (2.201 g, 3.40 mmol, 1.2 eq) and $\text{BF}_3\cdot\text{OEt}$ (0.355 mL, 2.83 mmol, 1.0 eq). The solution was cooled to 0 $^\circ\text{C}$ and stirred for ten minutes. Then, AgOTf (0.141 g, 0.566 mmol, 0.2 eq) and NIS (0.891 g, 3.96 mmol, 1.4 eq) were added, at which point the solution turned red. The mixture was allowed to return to ambient temperature, and stirred for an additional 2.5 hours before

being concentrated under reduced pressure. The resulting residue was re-dissolved in ethyl acetate (20 mL). This solution was washed with 1:1 saturated sodium bicarbonate (aq)/saturated sodium thiosulphate (aq) (5 x 20 mL), dried (Na_2SO_4) and concentrated under reduced pressure. The resultant mixture of diastereomers could not be readily separated, but were crudely purified by flash chromatography using first 10:10:1 DCM/EtOAc/MeOH then 4:4:1 DCM/EtOAc/MeOH to yield the mixture of diastereomers as a clear oil (1.58 g, 1.81 mmol) (64.0%). The mixture of diastereomers (1.58 g, 1.81 mmol, 1.0 eq) were dissolved in anhydrous THF (20 mL), and the solution was cooled to 0 °C before addition of *N,N'*-diisopropylcarbodiimide (0.841 mL, 5.43 mmol, 3.0 eq) and pentafluorophenol (1.00 g, 5.43 mmol, 3.0 eq). The solution was stirred for 30 minutes at 0 °C, followed by one additional hour at ambient temperature. The THF was removed under reduced pressure, after which the reaction mixture was redissolved in DCM and was washed with water (30 mL), and brine (30 mL) before drying (Na_2SO_4), and concentration under reduced pressure. The reaction mixture was purified by flash chromatography using 5:1 hexanes/EtOAc to yield first **6a** as a clear oil (1.058 g, 1.02 mmol) (56.1%) then **6b** also as a clear oil (0.508 g, 0.488 mmol) (26.9%) (83.0% overall).

6a: $[\alpha]_{\text{D}}^{25} = +2.5^\circ$ (*c* 1.5 CHCl_3); ^1H NMR (300 MHz, CDCl_3 , 298K) $\delta = 7.71\text{-}7.81$ (m, 2H, aromatic), 7.49-7.64 (m, 2H, aromatic), 7.15-7.46 (m, 24H, aromatic), 4.67-5.03 (m, 6H, H_1 , Hyp_α , $-\text{OCH}_2\text{Ph}$), 4.16-4.66 (m, 8H, H_5 , Hyp_γ , $\text{Hyp}_{\delta 1}$, $\text{Hyp}_{\delta 2}$, $-\text{OCH}_2\text{Ph}$), 4.06 (dd, 1H, H_2 , $J = 3.8\text{Hz}$, $J = 10.1\text{Hz}$), 3.82-4.01 (m, 3H, H_3 , H_4 , $-\text{C}(\text{O})\text{OCH}_2\text{CH}$), 3.70-3.82 (m, 2H, H_{6a} , H_{6b}), 3.37-3.60 (m, 2H, $-\text{C}(\text{O})\text{OCH}_2\text{CH}$), 2.68-2.84 (m, 1H, $\text{Hyp}_{\beta 1}$), 2.24-2.44 (m, 1H, $\text{Hyp}_{\beta 2}$); ^{13}C NMR (75 MHz, CDCl_3 , 298K) $\delta = (168.8)$, 168.5, 154.8,

(154.2), (144.1), 143.80, 143.77, (143.5), 141.33, 141.31, (141.2), (138.7), 138.6, (138.52), 138.44, 137.9, 124.8-128.6 (aromatic carbons), 120.0, (119.9), 97.9, (97.2), 78.8, (77.2), 76.4, (76.3), 76.2, 75.0, 74.8, (74.3), (73.6), 73.5 (2), (73.4), (73.3), 73.2, 70.42, (70.40), (69.42), 69.36, (68.4), 67.9, 57.9, (57.6), (51.9), 51.8, 47.1, (37.7), 36.5; MS (ES) calc. for C₆₀H₅₂F₅NNaO₁₀ (M + Na)⁺: 1065.04. Found (M + Na)⁺: 1065.18.

6b: $[\alpha]_D^{25} = -21.6^\circ$ (*c* 1.2 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.70$ -7.80 (m, 2H, aromatic), 7.49-7.65 (m, 2H, aromatic), 7.14-7.45 (m, 24H, aromatic), 4.87-5.01 (m, 1H, -OCH₂Ph), 4.69-4.87 (m, 5H, Hyp _{α} , -OCH₂Ph), 4.49-4.69 (m, 2H, Hyp _{γ} , -OCH₂Ph), 4.31-4.48 (m, 5H, H₁, H₅, Hyp _{δ_1} , -OCH₂Ph), 4.17-4.31 (m, 1H, Hyp _{δ_2}), 3.74-4.00 (m, 4H, H₂, H₃, H₄, -C(O)OCH₂CH), 3.45-3.64 (m, 4H, H_{6a}, H_{6b}, -C(O)OCH₂CH), 2.52-2.72 (m, 1H, Hyp _{β_1}), 2.15-2.39 (m, 1H, Hyp _{β_2}); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta = (168.8)$, 168.5, 154.9, (154.1), (144.1), 143.9, 143.7, (143.5), 141.3, (141.2), (138.50), 138.46, 138.32 (2), (138.29) (2), (137.8), 137.7, 127.0-128.7 (aromatic carbons), 125.2, (125.0), 120.0, (119.9), (102.8), 102.5, 82.4, (82.3), 79.2, 76.5, (75.9), 75.7, (75.5), 74.7, (74.6), 73.7, (73.64), 73.60, 73.3, (73.2), 73.1, (72.9), 68.8, (68.7), (68.4), 68.0, 57.6, (57.4), (53.5), 53.1, (47.14), 47.08, (36.9), 35.5; MS (ES) calc. for C₆₀H₅₂F₅NNaO₁₀ (M + Na)⁺: 1065.04. Found (M + Na)⁺: 1065.10.

4-O-(2,3,4,6-Tetra-O-benzyl- α,β -D-galactopyranosyl)-N-fluorenylmethoxycarbonyl-trans-4-hydroxy-L-proline N'-glycyl tert-butyl ester (7a,b). Compound **6a** (0.100 g, 0.096 mmol, 1.0 eq) was dissolved in anhydrous DMF (6 mL), before addition of DIPEA (0.067 mL, 0.38 mmol, 4.0 eq) and H-Gly-OtBu • HCl (0.032 g, 0.19 mmol, 2.0 eq). The solution was stirred for one hour at ambient temperature, before addition of water (15

mL) and ethyl acetate (15 mL). The organic layer was separated and dried (Na_2SO_4), then concentrated under reduced pressure for purification by flash chromatography using 3:2 hexanes/EtOAc to yield **7a** as a clear oil (0.080 g, 0.081 mmol) (84.2%) (a similar procedure was used for **7b**).

7a: $[\alpha]_D^{25} = +9.5^\circ$ (*c* 5.0 CHCl_3); ^1H NMR (300 MHz, CDCl_3 , 298K) $\delta =$ 7.70-7.85 (d, 2H, aromatic, $J = 7.4\text{Hz}$), 7.48-7.68 (m, 2H, aromatic), 7.17-7.46 (m, 24H, aromatic), 6.83 (broad t, 0.6H, Gly-NH, $J = 4.6\text{Hz}$), [6.13, broad t, 0.4H, Gly-NH], 4.96 (d, 1H, - OCH_2Ph , $J = 11.5\text{Hz}$), 4.93 (d, 1H, H_1 , $J = 3.8\text{Hz}$), 4.69-4.89 (m, 3H, - OCH_2Ph), 4.14-4.68 (m, 9H, Hyp_α , Hyp_γ , - OCH_2Ph , - $\text{C}(\text{O})\text{OCH}_2\text{CH}$, - $\text{C}(\text{O})\text{OCH}_2\text{CH}$), 3.45-4.12 (m, 10H, H_2 , H_3 , H_4 , H_5 , H_{6a} , H_{6b} , $\text{Hyp}_{\delta 1}$, $\text{Hyp}_{\delta 2}$, -Gly- CH_2), 2.34-2.56 (m, 1.6H, $\text{Hyp}_{\beta 1}$, $\text{Hyp}_{\beta 2}^{\text{trans}}$), 2.17-2.33 (m, 0.4H, $\text{Hyp}_{\beta 2}^{\text{cis}}$), 1.47 (s, 6H, *tert*-butyl), [1.44, s, 3H, *tert*-butyl]; ^{13}C NMR (75 MHz, CDCl_3 , 298K) $\delta =$ (171.89), 171.4, 168.6, 155.9, (155.1), 143.89, 143.81, 141.3, 138.8, 138.56, 138.55, 137.9, 124.9-128.6 (aromatic carbons), 120.0, 97.5, 82.2, 78.8, 77.3, 76.4, (76.2), 75.0, 74.8, 73.6, (73.3), 73.2, 70.1, 69.3, 67.9, 59.4, 51.9, 47.2, 42.2, (41.8), (37.6), 35.4, (29.7), 28.1; MS (ES) calc. for $\text{C}_{60}\text{H}_{64}\text{N}_2\text{NaO}_{11}$ ($\text{M} + \text{Na}$) $^+$: 1012.15. Found ($\text{M} + \text{Na}$) $^+$: 1012.21.

7b: $[\alpha]_D^{25} = -19.6^\circ$ (*c* 3.0 CHCl_3); ^1H NMR (300 MHz, CDCl_3 , 298K) $\delta =$ 7.70-7.81 (d, 2H, aromatic, $J = 7.3\text{Hz}$), 7.47-7.66 (m, 2H, aromatic), 7.18-7.43 (m, 24H, aromatic), 6.84 (broad t, 0.8H, Gly-NH, $J = 5.4\text{Hz}$), [6.03, broad t, 0.2H, Gly-NH], 4.95 (d, 1H, - OCH_2Ph , $J = 11.8\text{Hz}$), 4.68-4.87 (m, 4H, - OCH_2Ph), 4.63 (d, 1H, - OCH_2Ph , $J = 11.8\text{Hz}$), 4.52-4.60 (m, 0.8H, $\text{Hyp}_\alpha^{\text{trans}}$), 4.13-4.52 (m, 7.2H, H_1 , H_2 , $\text{Hyp}_\alpha^{\text{cis}}$, Hyp_γ , - OCH_2Ph , - $\text{C}(\text{O})\text{OCH}_2\text{CH}$), 3.43-4.04 (m, 10H, H_3 , H_4 , H_5 , H_{6a} , H_{6b} , $\text{Hyp}_{\delta 1}$, $\text{Hyp}_{\delta 2}$, - $\text{C}(\text{O})\text{OCH}_2\text{CH}$, -

Gly-CH₂), 2.13-2.58 (m, 2H, Hyp_{β1}, Hyp_{β2}), 1.46 (s, 7.2H, *tert*-butyl), [1.28, s, 1.8H, *tert*-butyl]; ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 171.2, 168.5, 156.1, 143.9, 143.8, 141.3, 138.6, 138.4, 137.8, 127.0-128.6 (aromatic carbons), 125.2, 119.9, 102.6, 82.2, 82.1, 79.2, 77.3, 75.3, 74.7, 73.55, 73.48, 73.43, 73.1, 68.8, 67.9, 58.9, 53.1, 47.1, 42.2, 33.9, (29.7), 28.1; MS (ES) calc. for C₆₀H₆₄N₂NaO₁₁ (M + Na)⁺: 1012.15. Found (M + Na)⁺: 1012.17.

4-O-(2,3,4,6-Tetra-O-benzyl- α,β -D-galactopyranosyl)-*trans*-4-hydroxy-L-proline N'-glycyl *tert*-butyl ester (8a,b**).** Compound **7a** (0.975 g, 0.99 mmol, 1.0 eq) was dissolved in DCM (40 mL), then cooled to 0 °C before slow addition of piperidine (10 mL). The solution was allowed to warm to ambient temperature over one hour, before concentration under reduced pressure. The product was purified by flash chromatography using 2:1 EtOAc/hexanes to yield **8a** as a clear oil (0.560 g, 0.73 mmol) (74.1%) (a similar procedure was used for **8b**).

8a: [α]_D²⁵ = +36.0° (c 0.5 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) δ = 8.04 (t, 1H, Gly-NH, J = 5.3Hz), 7.20-7.43 (m, 20H, aromatic), 4.35-4.98 (m, 9H, H₁, -OCH₂Ph), 4.22-4.30 (m, 1H, Pro_γ), 3.82-4.08 (m, 7H, H₂, H₃, H₄, H₅, Hyp_α, Gly-CH₂), 3.47-3.57 (m, 2H, H_{6a}, H_{6b}), 3.05 (dd, 1H, Hyp_{δ1}, J = 1.5Hz, J = 12.9Hz), 2.67 (dd, 1H, Hyp_{δ2}, J = 3.0Hz, J = 12.9Hz), 2.36 (ddd, 1H, Hyp_{β1}, J = 0.5Hz, J = 9.1Hz, J = 14.0Hz), 2.01 (ddd, 1H, Hyp_{β2}, J = 5.3Hz, J = 7.6Hz, J = 14.0Hz), 1.48 (s, 8H, *tert*-butyl), [1.43, s, 1H, *tert*-butyl]; ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 174.9, 169.1, 138.71, 138.65, 138.64, 137.9, 127.3-128.6 (aromatic carbons), 96.3, 82.0, 78.9, 78.4, 76.5, 74.9, 74.8, 73.8, 73.5,

72.9, 69.7, 68.9, 59.9, 52.1, 41.5, 37.2, 28.1; MS (ES) calc. for C₄₅H₅₅N₂O₉ (M + H)⁺: 767.93. Found (M + H)⁺: 767.89.

8b: [α]_D²⁵ = -12.0° (c 1.2 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) δ = [8.08, t, 0.3H, Gly-NH, J = 5.4Hz], 7.52 (t, 0.7H, Gly-NH, J = 5.6Hz), 7.18-7.40 (m, 20H, aromatic), 4.54-4.98 (m, 6H, -OCH₂Ph), 4.26-4.48 (m, 4H, H₁, Pro _{γ} , -OCH₂Ph), 3.67-4.02 (m, 4H, H₂, H₅, Gly-CH₂), 3.25-3.62 (m, 6H, H₃, H₄, H_{6a}, H_{6b}, Hyp _{α} , Hyp _{δ 1}), 2.67-2.83 (m, 1H, Hyp _{δ 2}), 2.23-2.46 (m, 1H, Hyp _{β 1}), 2.02-2.16 (m, 0.7H, Hyp _{β 2}^{trans}), [1.89-2.02, m, 0.3H, Hyp _{β 2}^{cis}], [1.47, s, 2.7H, *tert*-butyl], 1.42, (s, 6.3H, *tert*-butyl); ¹³C NMR (75 MHz, CDCl₃, 298K) δ = 174.8, (173.8), (169.1), 168.9, 138.60, (138.58), 138.51, (138.47), 138.45, (138.41), 137.9, (137.8), 127.4-128.7 (aromatic carbons), (102.9), 102.4, 82.3, (82.0), 81.9, (81.5), (79.4), 79.3, 77.2, 75.3, 74.6, (74.5), 73.6, (73.51), 73.48, 73.3, (73.1), 73.0, (68.9), 68.6, 64.6, 59.7, (58.9), (53.7), (41.51), 41.46, (36.8), 36.5, (29.7), 28.1; MS (ES) calc. for C₄₅H₅₅N₂O₉ (M + H)⁺: 767.93. Found (M + H)⁺: 767.95.

***N*-*tert*-Butyloxycarbonyl-*L*-prolyl-[4-*O*-(2,3,4,6-tetra-*O*-benzyl- α , β -D-galactopyranosyl)]-*trans*-4-hydroxy-*L*-proline *N'*-glycyl *tert*-butyl ester (9a,b).**

Compound **8a** (0.550 g, 0.072 mmol, 1.0 eq) was dissolved in anhydrous DMF (20 mL), and was cooled to 0 °C before addition of DIPEA (0.500 mL, 2.87 mmol, 4.0 eq), PyBOP (0.746 g, 0.143 mmol, 2.0 eq) and Boc-(2*S*)-Pro-OH (0.308 g, 0.143 mmol, 2.0 eq). The solution stirred for 15 minutes at 0 °C and then was stirred at ambient temperature for an additional 4 hours. The reaction mixture was diluted with water (20 mL) and the product was extracted into EtOAc (2 x 20 mL). The organic layer was dried (Na₂SO₄) and concentrated under reduced pressure for purification by flash chromatography using 2:1

EtOAc/hexanes to yield **9a** as a clear oil (0.630 g, 0.081 mmol) (91.1%) (a similar procedure was used for **9b**).

9a: $[\alpha]_D^{25} = -3.5^\circ$ (*c* 1.3 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta =$ [8.39, t, 0.16H, -Gly-NH, *J* = 5.9Hz], 7.18-7.45 (m, 20H, aromatic), 7.14 (t, 0.42H, -Gly-NH, *J* = 5.4Hz), [6.99, t, 0.42H, -Gly-NH, *J* = 5.4Hz], 4.21-5.03 (m, 12H, H₁, Hyp _{α} , Hyp _{γ} , Pro _{α} , -OCH₂Ph), 3.79-4.11 (m, 6.16H, H₂, H₃, H₄, H₅, H_{6a}, -Gly-CH₂), 3.29-3.79 (m, 5.84H, H_{6b}, Pro _{δ 1, δ 2}, Hyp _{δ 1, δ 2}, -Gly-CH₂), 2.36-2.67 (m, 1H, Hyp _{β 1}), 1.66-2.30 (m, 5H, Hyp _{β 2}, Pro _{β 1, β 2}, Pro _{γ 1, γ 2}), [1.45, s, 3H, *tert*-butyl], 1.43 (s, 9H, *tert*-butyl), [1.42, s, 3H, *tert*-butyl], [1.38, s, 3H, *tert*-butyl]; ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta =$ (172.8), 172.3, (172.2), (171.5), 171.2, (170.9), (168.9), 168.6, (168.3), (154.8), 154.5, (153.9), (138.81), (138.78), (138.73), 138.7 (3), (138.53), (138.47), (138.1), 137.95, (137.92), 127.2-128.6 (aromatic carbons), (98.9), 98.4, (96.4), (82.1), 81.7, (81.4), (79.9), 79.8, (79.6), 78.9, (78.7), (77.9), 77.5, (77.2), 76.5, (75.8), (75.1), 74.9, (74.84), 74.79, (73.7), 73.55, (73.48), (73.43), (73.11), 73.07, 73.05, (72.9), (72.5), (70.2), 70.0, (69.8), (69.3), 69.0, (68.7), (59.6), 58.85, (58.75), 58.5, (57.9), (57.4), (51.8), 51.6, (50.8), (47.1), 46.9, (46.6), (42.2), 41.89, (41.83), (37.8), 33.77, (33.72), (29.9), 29.2, (29.1), (28.49), 28.46, (28.42), 28.0, (24.6), 24.1, (23.5); MS (ES) calc. for C₅₅H₆₉N₃NaO₁₂ (M + Na)⁺: 987.14. Found (M + Na)⁺: 987.24.

9b: $[\alpha]_D^{25} = -26.3^\circ$ (*c* 1.0 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta =$ [8.44, t, 0.15H, -Gly-NH, *J* = 6.1Hz], 7.19-7.46 (m, 20.3H, aromatic, -Gly-NH), 7.14 (t, 0.55H, -Gly-NH, *J* = 5.5Hz), 4.25-5.03 (m, 12H, H₁, Hyp _{α} , Hyp _{γ} , Pro _{α} , -OCH₂Ph), 3.30-4.07 (m, 12H, H₂, H₃, H₄, H₅, H_{6a}, H_{6b}, Pro _{δ 1, δ 2}, Hyp _{δ 1, δ 2}, -Gly-CH₂), [2.62-2.73, m, 0.15H, Hyp _{β 1}],

[2.51-2.62, m, 0.3H, Hyp $_{\beta 1}$], 2.25-2.49 (m, 0.7H, Hyp $_{\beta 1}$, Hyp $_{\beta 2}$), 1.66-2.24 (m, 6H), [1.44, s, 4.05H, *tert*-butyl], 1.43 (s, 4.95H, *tert*-butyl), [1.42, s, 2.7H, *tert*-butyl], [1.41, s, 1.35H, *tert*-butyl], 1.37 (s, 4.95H, *tert*-butyl); ^{13}C NMR (75 MHz, CDCl_3 , 298K) δ = 172.9, (172.2), (171.7), (171.38), (171.34), 170.9, 168.7, (168.3), (154.8), (154.5), 153.9, (138.64), (138.59), 138.53, 138.47, 138.46, (138.44), (138.3), (137.9), (137.81), 137.79, 127.4-128.6 (aromatic carbons), 103.7, (102.8), (102.7), 82.16, (82.12), (82.08), 81.9, (81.6), (81.4), (79.9), 79.8, (79.3), (79.23), 79.19, 78.2, (77.3), (75.4), 75.3, (75.1), (74.8), 74.7, (74.6), (73.6), 73.5 (3), (73.3), 73.2, (72.9), 68.9, (68.7), (68.5), 59.3, (58.5), 58.1, (57.4), 52.7, (52.5), (52.2), (47.1), (46.8), 46.5, (42.2), 41.9, (41.7), (36.9), 33.0, (32.8), 29.9, (29.3), (28.9), 28.5, (28.4), 28.0, (24.5), (24.2), 23.4; MS (ES) calc. for $\text{C}_{55}\text{H}_{69}\text{N}_3\text{NaO}_{12}$ (M + Na) $^+$: 987.14. Found (M + Na) $^+$: 987.18.

***N-tert*-Butyloxycarbonyl-*L*-prolyl-(4-*O*- α,β -D-galactopyranosyl)-*trans*-4-hydroxy-*L*-proline *N'*-glycyl *tert*-butyl ester (10a,b).** Compound **9a** (0.250 g, 0.26 mmol, 1.0 eq) was dissolved in 4:1 methanol/ethyl acetate (25 mL). After addition of Pearlman's catalyst (20% palladium hydroxide on carbon) (0.150 g, approx. 0.15 mmol, 0.6 eq), the reaction flask was flushed with $\text{Ar}_{(\text{g})}$ for 5 minutes. The reaction mixture was then stirred vigorously under hydrogen atmosphere (10 psi) for 4 hours, after which it was flushed with nitrogen and filtered. The product was then concentrated under reduced pressure to provide **10a** as a clear oil (0.150 g, 0.25 mmol) (96.2%) (a similar procedure was used for **10b**).

10a: $[\alpha]_{\text{D}}^{25} = -1.8^\circ$ (*c* 1.4 CH_3OH); ^1H NMR (300 MHz, D_2O , 298K) δ = 5.07 (d, 0.9H, H_1 , $J = 3.7\text{Hz}$), [5.01, d, 0.1H, H_1], 4.48-4.68 (m, 3H, Hyp $_{\alpha}$, Hyp $_{\gamma}$, Pro $_{\alpha}$), 3.67-4.03 (m,

10H, H₂, H₃, H₄, H₅, H_{6a,6b}, Hyp_{δ1,δ2}, -Gly-CH₂), 3.33-3.54 (m, 2H, Pro_{δ1,δ2}), 2.53-2.74 (m, 1H, Hyp_{β1}), 2.22-2.41 (m, 1H, Pro_{β1}), 2.02-2.19 (m, 1H, Hyp_{β2}), 1.76-1.97 (m, 3H, Pro_{β2}, Pro_{γ1,γ2}), 1.43 (s, 8.1H, *tert*-butyl), [1.42, s, 0.9H, *tert*-butyl], 1.37 (s, 8.1H, *tert*-butyl), [1.34, s, 0.9H, *tert*-butyl]; ¹³C NMR (75 MHz, D₂O, 298K) δ = (174.5), 174.2, 174.1, (173.9), 170.7, (156.5), 156.1, 99.4, (98.7), 84.2, (82.1), 82.0, 78.6, (77.3), 72.1, (71.9), (69.69), 69.66, 69.59, (68.6), 68.5, 61.7, 59.8, 58.9, (58.7), 53.3, (53.1), (47.5), 46.9, 42.7, (42.6), 35.9, 29.8, (29.4), (28.13), 28.09, (27.6), 27.5, (24.4), 23.8; MS (ES) calc. for C₂₇H₄₅N₃NaO₁₂ (M + Na)⁺: 626.65. Found (M + Na)⁺: 626.69.

10b: [α]_D²⁵ = -54.1° (c 2.5 CH₃OH); ¹H NMR (300 MHz, D₂O, 298K) δ = 4.40-4.76 (m, 4H, Hyp_α, Hyp_γ, Pro_α, H₁), 3.55-4.09 (m, 9H, H₃, H₄, H₅, H_{6a,6b}, Hyp_{δ1,δ2}, -Gly-CH₂), 3.30-3.54 (m, 3H, H₂, Pro_{δ1,δ2}), 2.39-2.66 (m, 1H, Hyp_{β1}), 2.21-2.38 (m, 1H, Pro_{β1}), 2.00-2.21 (m, 1H, Hyp_{β2}), 1.75-1.96 (m, 3H, Pro_{β2}, Pro_{γ1,γ2}), 1.43 (s, 9H, *tert*-butyl), [1.42, s, 3H, *tert*-butyl], 1.37 (s, 6H, *tert*-butyl); ¹³C NMR (75 MHz, D₂O, 298K) δ = (174.4), 174.3, 174.2, (173.8), 170.7, (168.7), (156.5), 156.1, 102.2, (101.8), 84.2, 82.3, (82.1), (77.8), 77.5, (75.7), 75.5, 73.3, (72.9), 70.9, (68.98), 68.92, 61.4, (59.21), 59.18, 58.8, (58.7), 53.9, (47.5), 46.9, 42.64, (42.59), 35.2, (34.8), 29.9, (29.3), 28.1, 27.6, (24.4), 23.9; MS (ES) calc. for C₂₇H₄₅N₃NaO₁₂ (M + Na)⁺: 626.65. Found (M + Na)⁺: 626.77.

***N*-*tert*-Butyloxycarbonyl-*L*-prolyl-[4-*O*-(2,3,4,6-tetra-*O*-acetyl- α,β -D-galactopyranosyl)]-*trans*-4-hydroxy-*L*-proline *N'*-glycyl *tert*-butyl ester (11a,b).**

Compound **10a** (0.125 g, 0.207 mmol) was dissolved in 8:1 (v/v) pyridine/Ac₂O (4.5 mL), and the mixture was stirred at ambient temperature for 16 hours. The solution was concentrated under reduced pressure, and then was co-distilled with toluene (3 x 10 mL)

before purification by flash chromatography using 4:1 ethyl acetate/hexanes to yield **11a** as a clear oil (0.102 g, 0.13 mmol) (63.8%) (a similar procedure was used for **11b**).

11a: $[\alpha]_D^{25} = +23.1^\circ$ (*c* 1.0 CHCl₃); ¹H NMR (500 MHz, CDCl₃, 298K) δ = [8.46, t, 0.3H, -Gly-NH', J = 6.0Hz], 7.27 (t, 0.35H, -Gly-NH, J = 5.5Hz), [7.24, t, 0.35H, -Gly-NH', J = 5.5Hz], 5.38-5.47 (m, 1H, H₄), 5.16-5.32 (m, 1.7H, H₂, H₃, H₃'), 4.99-5.14 (m, 1.3H, H₁, H₃''), 4.78 (dd, 0.35H, Hyp _{α} , J = 3.4Hz, J = 8.2Hz), [4.71, dd, 0.3H, Hyp _{α} '', J = 2.8Hz, J = 8.2Hz], 4.52-4.61 (m, 0.65H, Hyp _{γ} , Hyp _{γ} ''), [4.47, dd, 0.35H, Hyp _{α} ', J = 4.5Hz, J = 8.7Hz], 4.41 (dd, 0.35H, Pro _{α} , J = 4.2Hz, J = 8.4Hz), 4.17-4.38 (m, 2H, H₅, Hyp _{γ} ', Pro _{α} ', Pro _{α} ''), 3.99-4.16 (m, 2H, H_{6a}, H_{6b}), 3.31-3.98 (m, 6H, Hyp _{δ 1,2}, Pro _{δ 1,2}, -Gly-CH₂), [2.56-2.70, m, 0.65H, Hyp _{β 1}', Hyp _{β 1}''], 2.48-2.56 (m, 0.35H, Hyp _{β 1}), [2.38-2.48, m, 0.35H, Hyp _{β 2}'], 1.74-2.24 (m, 16.65H, Hyp _{β 2}, Hyp _{β 2}'', Pro _{β 1,2}, Pro _{γ 1,2}, -COCH₃), 1.32-1.50 (m, 18H, *tert*-butyl); ¹³C NMR (75 MHz, CDCl₃, 298K) δ = (172.8), (172.5), 172.1, 171.1, (170.9), (170.65), (170.62), 170.61, (170.55), 170.50, (170.4), (170.18), (170.15), 170.12, 169.93, (169.89), (169.7), (168.6), (168.5), 168.3, 154.8, (154.5), (153.7), (96.8), (96.0), 95.8, (82.0), (81.8), 81.5, 80.1, (79.8), (79.7), 77.3, (76.5), (74.6), (68.1), 68.0, (67.9), (67.8), 67.7, (67.3), (67.26), 67.24, 66.9, (66.8), 61.9, (61.8), (61.7), (59.5), 58.58, (58.52), (58.2), (57.9), 57.3, (51.4), 51.2, (51.0), (47.1), 46.8, (46.6), (42.1), (41.9), 41.8, 37.7, (33.5), (33.2), (30.0), 29.3, (29.2), 28.4, 27.9, 24.6, (24.2), (23.5), 20.8, 20.7, 20.65, 20.59; MS (ES) calc. for C₃₅H₅₃N₃NaO₁₆ (M + Na)⁺: 794.80. Found (M + Na)⁺: 794.74.

11b: $[\alpha]_D^{25} = -37.1^\circ$ (*c* 1.0 CHCl₃); ¹H NMR (500 MHz, CDCl₃, 298K) δ = [8.45, t, 0.15H, -Gly-NH', J = 6.0Hz], 7.48 (t, 0.45H, -Gly-NH, J = 5.3Hz), [7.33, t, 0.4H, -Gly-

NH , $J = 5.3\text{Hz}$], 5.27-5.38 (m, 1H, H_4), 5.02-5.14 (m, 1H, H_2), 4.90-4.99 (m, 1H, H_3), 4.70 (dd, 0.45H, Hyp_α , $J = 3.1\text{Hz}$, $J = 8.4\text{Hz}$), [4.65, dd, 0.4H, Hyp_α' , $J = 2.8\text{Hz}$, $J = 8.3\text{Hz}$], 4.47-4.58 (m, 1.7H, H_1 , H_1' , Hyp_γ , Hyp_γ'), [4.45, d, 0.15H, H_1'' , $J = 7.8\text{Hz}$], 4.33-4.39 (m, 0.45H, Pro_α), [4.28-4.33, m, 0.55H, Hyp_α'' , Pro_α'], [4.23-4.28, m, 0.3H, Hyp_γ'' , Pro_α''], 4.01-4.15 (m, 2H, H_{6a} , H_{6b}), 3.96 (dd, 0.45H, $-Gly-CH_{2a}$, $J = 6.7\text{Hz}$, $J = 17.9\text{Hz}$), 3.29-3.90 (m, 6.55H, H_5 , $Hyp_{\delta 1}$, $Hyp_{\delta 2}$, $Pro_{\delta 1}$, $Pro_{\delta 2}$, $-Gly-CH_{2a}'$, $-Gly-CH_{2a}''$, $-Gly-CH_{2b}$), 2.51-2.62 (m, 0.6H, $Hyp_{\beta 1}$, $Hyp_{\beta 1}''$), [2.35-2.45, m, 0.4H, $Hyp_{\beta 1}'$], [2.20-2.30, m, 0.15H, $Hyp_{\beta 2}''$], 1.70-2.20 (m, 16.85H, $Hyp_{\beta 2}$, $Hyp_{\beta 2}'$, $Pro_{\beta 1,\beta 2}$, $Pro_{\gamma 1,\gamma 2}$, $-COCH_3$), 1.27-1.46 (m, 18H, *tert*-butyl); ^{13}C NMR (75 MHz, CDCl_3 , 298K) $\delta =$ (173.0), 172.3, (171.7), 171.2, (170.8), (170.4), 170.3, 170.2, (170.06), 170.01, 169.22, (169.18), 168.5, (168.4), (168.2), (154.8), 154.5, (153.7), (101.2), (100.6), 100.0, (81.9), 81.6, (81.4), (79.99), 79.87, (79.7), (78.4), 77.3, (77.2), (75.8), (70.9), (70.8), 70.7 (2), (68.7), 68.5, 67.0, (66.8), 61.3, (61.2), 58.9, (58.1), (57.9), 57.8, (57.3), (52.13), (52.08), 51.9, (47.1), 46.8, (46.5), (42.1), (41.9), 41.6, (37.1), (32.6), 32.3, 29.9, (29.6), (29.2), (28.9), (28.41), 28.37, 27.9, (24.5), 24.2, (23.4), (20.74), (20.69), (20.66), 20.62 (3), 20.5; MS (ES) calc. for $\text{C}_{35}\text{H}_{53}\text{N}_3\text{NaO}_{16}$ ($M + \text{Na}$) $^+$: 794.80. Found ($M + \text{Na}$) $^+$: 794.77.

***N*-Fluorenylmethoxycarbonyl-*L*-prolyl-[4-*O*-(2,3,4,6-tetra-*O*-acetyl- α,β -D-galactopyranosyl)]-*trans*-4-hydroxy-*L*-proline *N'*-glycine (12a,b).** Compound **11a** (0.110 g, 0.14 mmol, 1.0 eq) was dissolved in DCM (4 mL), and then cooled to 0 °C. After slow addition of TFA (4 mL), the solution was stirred for 15 minutes at 0 °C followed by an additional 2 hours at ambient temperature before being co-distilled with toluene (2 x 10 mL). The resultant pale yellow oil was then suspended in 2:1 acetone/water (6 mL). Addition of NaHCO_3 (0.115 g, 1.4 mmol, 10.0 eq) caused

evolution of gas and a pH of 9. This was followed by addition of 9-fluorenylmethyl pentafluorophenyl carbonate (0.167 g, 0.42 mmol, 3.0 eq) dissolved in acetone (2 mL). After stirring for 2 hours, the solvent was removed under reduced pressure and the mixture was re-dissolved in water (10 mL) and EtOAc (10 mL). After the pH was adjusted to 5 using formic acid, the product was extracted into EtOAc (4 x 10 mL). The organic layer was dried (Na_2SO_4) and concentrated under reduced pressure for purification by flash chromatography using first 6:1 DCM/MeOH then 5:1 DCM/MeOH to yield **12a** as a clear oil (0.070 g, 0.084 mmol) (59.0% over 2 steps) (a similar procedure was used for **12b**).

12a: $[\alpha]_{\text{D}}^{25} = +21.9^\circ$ (*c* 2.0 CH_3OH); ^1H NMR (300 MHz, CDO_3D , 298K) $\delta = 7.75\text{-}7.89$ (m, 2H, aromatic), 7.55-7.70 (m, 2H, aromatic), 7.25-7.51 (m, 4H, aromatic), 5.02-5.53 (m, 4H, H_1 , H_2 , H_3 , H_4), 3.64-4.71 (m, 12H, H_5 , H_{6a} , H_{6b} , Hyp_α , Hyp_γ , $\text{Hyp}_{\delta 1}$, Pro_α , $-\text{C}(\text{O})\text{OCH}_2\text{CH}$, $-\text{C}(\text{O})\text{OCH}_2\text{CH}-\text{Gly}-\text{CH}_2$), 3.32-3.64 (m, 3H, $\text{Hyp}_{\delta 2}$, $\text{Pro}_{\delta 1, \delta 2}$), 1.75-2.69 (m, 18H, $\text{Hyp}_{\beta 1, \beta 2}$, $\text{Pro}_{\beta 1, \beta 2}$, $\text{Pro}_{\gamma 1, \gamma 2}$, $-\text{COCH}_3$); ^{13}C NMR (75 MHz, CDO_3D , 298K) $\delta = 174.3$, (174.2), (173.9), 173.8, (173.6), 173.4, 172.31, (172.2), (172.1), 172.0, 171.9, (171.7), 171.5, (171.4), 156.5, (156.2), 145.5, (145.3), (142.7), 142.6, (129.0), 128.8, (128.3), 128.2, 126.3, (126.12), (121.1), 120.9, (98.3), 96.8, (78.9), 77.7, 69.7, (69.6), 69.1, 68.9, (68.8), (68.6), (68.4), 68.2, (63.3), 63.1, 60.6, (60.5), 59.8, (59.3), 53.1, (52.6), 42.97, 42.79, 36.7, (36.5), (30.9), 30.3, 25.4, (24.1), 20.8, (20.74), 20.71, 20.64, 20.55, (20.51); MS (ES) calc. for $\text{C}_{41}\text{H}_{47}\text{N}_3\text{NaO}_{16}$ ($\text{M} + \text{Na}$) $^+$: 860.29. Found ($\text{M} + \text{Na}$) $^+$: 860.32.

12b: $[\alpha]_D^{25} = -26.5^\circ$ (*c* 1.0 CH₃OH); ¹H NMR (500 MHz, CDO₃D, 298K) $\delta = 7.74$ -7.87 (m, 2H, aromatic), 7.55-7.67 (m, 2H, aromatic), 7.25-7.48 (m, 4H, aromatic), 5.31-5.40 (m, 1H, H₄), 5.00-5.18 (m, 2H, H₂, H₃), 4.68-4.78 (m, 1H, H₁), 4.31-4.63 (m, 4.3H, Hyp _{α} , Hyp _{γ} ', Pro _{α} , -C(O)OCH₂CH), 4.02-4.30 (m, 4.7H, H₅, H_{6a}, H_{6b}, Hyp _{γ} , -C(O)OCH₂CH), 3.33-3.97 (m, 6H, Hyp _{δ 1,2}, Pro _{δ 1,2}, -Gly-CH₂), 1.68-2.34 (m, 18H, Hyp _{β 1,2}, Pro _{β 1,2}, Pro _{γ 1,2}, -COCH₃); ¹³C NMR (75 MHz, CDO₃D, 298K) $\delta = (173.8), 173.6, (173.4), 170.0, 172.0 (2), 171.9, 171.5, 171.3, (156.6), 156.3, 145.8, (145.1), 142.6, (142.5), (128.94), 128.90, (128.5), 128.3, (126.4), 126.3, (121.1), 121.0, 102.0, (101.5), 79.58, (79.51), (72.2), 72.1, 70.33, (70.31), (68.97), 68.91, (68.6), 68.2, 63.0, (60.45), 60.37, (60.1), 59.3, (54.3), 53.9, 43.93, 43.87, 36.1, 30.9, (30.1), (25.3), 23.9, 20.8, 20.7, 20.6, 20.3$; MS (ES) calc. for C₄₁H₄₇N₃NaO₁₆ (M + Na)⁺: 860.29. Found (M + Na)⁺: 860.37.

***N*-Fluorenylmethoxycarbonyl-(4-*O*-*tert*-butyl)-*trans*-4-hydroxy-*L*-proline *N'*-glycyl *tert*-butyl ester (**14**)**. Compound **13** (3.04 g, 7.42 mmol, 1.0 eq) was dissolved in anhydrous DMF (40 mL), and was cooled to 0 °C before addition of DIPEA (3.88 mL, 22.3 mmol, 3.0 eq), TBTU (3.58 g, 11.1 mmol, 1.5 eq) and H-Gly-OtBu • HCl (1.56 g, 9.28 mmol, 1.25 eq). The solution stirred for 1 hour at 0 °C and then was diluted with water (100 mL) and the product was extracted into EtOAc (3 x 100 mL). The organic layer was dried (Na₂SO₄) and concentrated under reduced pressure for purification by flash chromatography using 1:1 EtOAc/hexanes to yield **14** as a white foam (3.63 g, 6.95 mmol) (93.6%): $[\alpha]_D^{25} = -22.5^\circ$ (*c* 1.4 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.70$ -7.83 (m, 2H, aromatic), 7.48-7.67 (m, 2H, aromatic), 7.21-7.47 (m, 4H, aromatic), 7.01 (broad t, 0.65H, -Gly-NH), [6.04, broad t, 0.35H, -Gly-NH], 4.06-4.62 (m, 5H, Hyp _{α} , Hyp _{γ} , -C(O)OCH₂CH, -C(O)OCH₂CH), 3.72-4.00 (m, 2H, -Gly-CH₂), 3.66 (dd,

1H, Hyp $_{\delta 1}$, J = 6.5Hz, J = 10.7Hz), 3.24-3.45 (m, 1H, Hyp $_{\delta 2}$), 2.34-2.56 (m, 0.65H, Hyp $_{\beta 1}$), 1.87-2.29 (m, 1.35H, Hyp $_{\beta 2}$, H $_{\beta 1}$ '), 1.45 (s, 9H, *tert*-butyl), 1.21 (s, 9H, *tert*-butyl); ^{13}C NMR (75 MHz, CDCl_3 , 298K) δ = 171.5, 168.6, 156.1, 143.8, 141.3, 127.7, 127.1, 125.1, 119.9, 82.2, 77.2, 74.1, 69.5, 69.3, 67.8, 59.1, 53.2, 47.2, 42.1, (38.8), 36.3, 28.3, 28.0; MS (ES) calc. for $\text{C}_{30}\text{H}_{38}\text{N}_2\text{NaO}_6$ (M + Na) $^+$: 545.26. Found (M + Na) $^+$: 545.31.

(4-*O-tert*-Butyl)-*trans*-4-hydroxy-*L*-proline *N'*-glycyl *tert*-butyl ester (15). Compound **14** (3.50 g, 6.70 mmol, 1.0 eq) was dissolved in DCM (40 mL), and then was cooled to 0 °C before slow addition of piperidine (10 mL). The solution was stirred for 30 min before being co-distilled with toluene (3 x 15 mL). The product was purified by flash chromatography using 12:1 EtOAc/MeOH to yield **15** as a white solid (1.80 g, 5.99 mmol) (89.6%); $[\alpha]_D^{25} = -23.3^\circ$ (*c* 0.2 CHCl_3); mp 83-87°C; ^1H NMR (300 MHz, CDCl_3 , 298K) δ = 7.96 (t, 1H, -Gly-NH, J = 5.4Hz), 4.03-4.12 (m, 1H, Hyp $_{\gamma}$), 3.74-3.95 (m, 3H, Hyp $_{\alpha}$, -Gly- CH_2), 2.75-2.90 (m, 2H, Hyp $_{\delta 1, \delta 2}$), 2.34 (broad s, 1H, -Pro-NH), 1.89-2.13 (m, 2H, Hyp $_{\beta 1, \beta 2}$), 1.40 (s, 9H, *tert*-butyl), 1.10 (s, 9H, *tert*-butyl); ^{13}C NMR (75 MHz, CDCl_3 , 298K) δ = 175.1, 169.1, 81.8, 73.5, 72.2, 59.8, 54.9, 41.4, 39.1, 28.4, 28.0; MS (ES) calc. for $\text{C}_{15}\text{H}_{29}\text{N}_2\text{O}_4$ (M + H) $^+$: 301.21. Found (M + H) $^+$: 301.10; MS (ES) calc. for $\text{C}_{15}\text{H}_{28}\text{N}_2\text{NaO}_4$ (M + Na) $^+$: 323.19. Found (M + Na) $^+$: 323.07.

***N*-Fluorenylmethoxycarbonyl-*L*-prolyl-(4-*O-tert*-butyl)-*trans*-4-hydroxy-*L*-proline *N'*-glycyl *tert*-butyl ester (16).** Initially, Fmoc-Pro-OH (0.362 g, 1.07 mmol, 1.5 eq), TBTU (0.459 g, 1.43 mmol, 2.0 eq) and DIPEA (0.500 mL, 2.86 mmol, 4.0 eq) were dissolved in anhydrous DMF (4 mL). The mixture was cooled to 0 °C, and the flask was

flushed with Ar_(g). In a separate flask, compound **15** (0.215 g, 0.716 mmol, 1.0 eq) was dissolved in anhydrous DMF (4 mL), and was also flushed with Ar_(g). The solution containing **15** was then added dropwise by syringe to the mixture containing Fmoc-Pro-OH, TBTU and DIPEA. The resulting solution was stirred for 30 minutes at 0 °C followed by an additional 2 hours at ambient temperature before being diluted with water (40 mL). The product was extracted into EtOAc (3 x 40 mL), and then the organic layer was dried (Na₂SO₄) and concentrated under reduced pressure for purification by flash chromatography using 3:2 EtOAc/hexanes to yield **16** as a clear oil (0.394 g, 0.636 mmol) (88.7%): $[\alpha]_D^{25} = -42.9^\circ$ (*c* 3.1 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta =$ [8.09, t, 0.25H, -Gly-NH', J = 5.9Hz], 7.70-7.81 (m, 2H, aromatic), 7.49-7.68 (m, 2H, aromatic), 7.23-7.44 (m, 4.4H, aromatic, -Gly-NH), [7.20, t, 0.35H, -Gly-NH', J = 5.6Hz], 4.76 (dd, 0.4H, Hyp _{α} , J = 2.2Hz, J = 8.2Hz), 4.14-4.63 (m, 5.6H, Hyp _{α'} , Hyp _{α''} , Hyp _{γ} , Pro _{α} , -C(O)OCH₂CH, -C(O)OCH₂CH), 3.41-4.04 (m, 5.75H, Hyp _{δ_1} , Hyp _{δ_2} , Hyp _{δ_2'} , Pro _{δ_1, δ_2} , -Gly-CH₂), [3.15, dd, 0.25H, Hyp _{δ_2''} , J = 7.3Hz, J = 9.4Hz], [2.60, ddd, 0.25H, Hyp _{β_1''} , J = 2.5Hz, J = 6.0Hz, J = 12.5Hz], 2.37-2.52 (m, 0.75H, Hyp _{β_1} , Hyp _{β_1'}), 1.68-2.32 (m, 5H, Hyp _{β_2} , Pro _{β_1, β_2} , Pro _{γ_1, γ_2}), 1.43 (s, 6.75H, *tert*-butyl), [1.36, s, 2.25H, *tert*-butyl], 1.14-1.21 (m, 9H, *tert*-butyl); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta =$ 172.3, (172.1), (171.5), (171.4), 171.3, (171.1), 168.6, (168.5), (168.3), (155.2), 154.9, (154.4), (144.4), 144.2, (143.9), 141.3, (141.2), (141.1), (127.74), (127.68), 127.6, (127.08), 127.05, (127.02), 125.3, (125.1), (125.0), (119.97), 119.92, (81.9), 81.8, (81.4), (74.1), (74.0), 73.9, 70.2, (70.1), (67.8), 67.5, (67.0), (59.5), 58.5, 58.4, (57.9), (57.7), (52.9), 52.8, (52.5), (47.4), 47.3, (47.12), (47.09), 46.8, (42.1), 41.9, (39.1), 34.7, (34.6), (30.3),

(29.4), 29.2, 28.3, (28.2), 28.0, (27.9), (24.7), 24.3, (22.9); MS (ES) calc. for $C_{35}H_{45}N_3NaO_7$ ($M + Na$)⁺: 642.32. Found ($M + Na$)⁺: 642.26.

***N*-Fluorenylmethoxycarbonyl-*L*-prolyl-*trans*-4-hydroxy-*L*-proline *N'*-glycine (17).**

Compound **16** (0.218 g, 0.352 mmol, 1.0 eq) was dissolved in DCM (4 mL), and then cooled to 0 °C. After slow addition of TFA (4 mL), the solution was stirred for 1 hour at 0 °C followed by an additional 1.5 hours at ambient temperature before being co-distilled with toluene (2 x 10 mL). The resultant white solid was then suspended in water (6 mL), before adjusting the pH to 9-10 using 5% (v/v) NaHCO₃ (1.5 mL). The aqueous layer was washed with EtOAc (3 x 6 mL) before the pH was adjusted to 3 using formic acid. After leaving the solution at 4 °C overnight, large amounts of white precipitate were visible in the flask, which were filtered off and vacuum dried to yield **17** as a white solid (0.134 g, 0.264 mmol) (74.9%): $[\alpha]_D^{25} = -64.5^\circ$ (*c* 0.6 CH₃OH); mp 176-181 °C; ¹H NMR (500 MHz, CD₃OD, 298K) $\delta = 7.72$ -7.86 (m, 2H, aromatic), 7.51-7.68 (m, 2H, aromatic), 7.23-7.44 (m, 4H, aromatic), 4.27-4.67 (m, 5H, Hyp_α, Hyp_γ, Pro_α, -C(O)OCH₂CH), 4.16-4.26 (m, 1H, -C(O)OCH₂CH), 4.02 (dd, 1H, -Gly-CH_{2a}, *J* = 6.2Hz, *J* = 17.8Hz), 3.76 (dd, 1H, -Gly-CH_{2b}, *J* = 2.1Hz, *J* = 17.8Hz), 3.64-3.74 (m, 1.5H, Hyp_{δ1}, Hyp_{δ2}), 3.36-3.62 (m, 2.5H, Hyp_{δ2'}, Pro_{δ1,δ2}), 1.77-2.43 (m, 6H, Hyp_{β1,β2}, Pro_{β1,β2}, Pro_{γ1,γ2}); ¹³C NMR (75 MHz, CD₃OD, 298K) $\delta = 174.6$, (174.5), 173.43, (173.41), 172.83, (172.81), 156.7, (156.4), 145.4, (145.2), 142.7, (142.6), 128.8, (128.4), (128.3), 128.2, 126.2, (126.1), 120.9, 71.1, (70.7), 68.7, 60.5, (60.0), (59.8), 59.5, 56.0, (55.7), 48.6, 47.9, 41.9, 38.6, (38.5), (31.1), 30.1, 25.2, (24.1); MS (ES) calc. for $C_{27}H_{30}N_3O_7$ ($M + H$)⁺: 508.21. Found ($M + H$)⁺: 508.00. MS (ES) calc. for $C_{27}H_{29}N_3NaO_7$ ($M + Na$)⁺: 529.18. Found ($M + Na$)⁺: 529.22.

6.6.2 Peptide Synthesis and Purification

Peptides **1-3** were synthesized by segment condensation of their corresponding Fmoc-protected tripeptides using solid-phase chemistry on an Argonaut Quest 210 peptide synthesizer (Argonaut Technologies, San Carlos, CA). Peptide couplings were carried out in 5 mL reaction vessels using Rink Amide MBHA resin for **1-3** on a 40, 12, and 16 μmol scale, respectively. The Fmoc-protected resin was allowed to swell in dry DCM for 1 hour, and then was washed with DMF. Fmoc deprotection was carried out using piperidine in DMF (1:4). Coupling steps were carried out by pre-activation of the tripeptide fragment (3.0 eq) with TBTU (4.0 eq) and DIPEA (8.0 eq) in dry DMF. All couplings were allowed to react for 2.5 hours at room temperature. After coupling of the first and last tripeptides, acetylation was performed using acetic anhydride in pyridine (1:9) for capping of any unreacted sites and of the *N*-terminus of the peptide, respectively. *O*-Acetate deprotection of **2** and **3** was carried out using methanol/hydrazine hydrate (6:1) (2 x 3 hours).⁸² This deprotection step was carried out for **3** after the peptide was cleaved from the resin in order to limit acid-catalyzed de-glycosylation of the β -linked sugars. Cleavage from the resin was accomplished using TFA/triisopropylsilane/water (18:1:1). The resin was washed with TFA and DCM before concentration by co-distillation with toluene.

The peptides were purified by semi-preparative HPLC using a Waters HPLC system and a Vydac C-18 reversed-phase 5 μm particle size (250 mm x 10 mm) column (W. R. Grace & Co., Deerfield, IL). A linear gradient of 0 to 40% solvent B (CH_3CN) in solvent A (H_2O) was used at a flow rate of 1 mL/min. Detection of the products was achieved using a photodiode array monitoring at $\lambda = 210$ nm. Analytical HPLC was

carried out using an Atlantis C-18 5 μm particle size (150 mm x 4.6 mm) column (Waters Corp., Milford, MA) using the same solvent gradient. Peptides **1-3** were $\geq 90\%$ pure according to both analytical HPLC and MALDI mass spectrometry. Peptide **1** Ac-(Pro-Hyp-Gly)₇-NH₂: [M + H⁺] Calc. for C₈₆H₁₂₅N₂₂O₂₉: 1929.8983. Found: 1929.8962; Peptide **2** Ac-(Pro-Hyp(α -D-Gal)-Gly)₇-NH₂: [M + H⁺] Calc. for C₁₂₈H₁₉₅N₂₂O₆₄: 3064.2681. Found: 3064.2949; Peptide **3** Ac-(Pro-Hyp(β -D-Gal)-Gly)₇-NH₂: [M + Na⁺] Calc. for C₁₂₈H₁₉₄N₂₂NaO₆₄: 3086.2500. Found: 3086.3054.

6.6.3 Circular Dichroism Spectroscopy

Peptides were dried under vacuum for 48 hours prior to weighing, and then were dissolved in pure water at an approximate concentration of 0.1-0.2 mg/mL (0.12-0.26 mM). The solutions were left to equilibrate at 4 °C for 6 days. The pH of the samples of **1-3** was determined at 25 °C after a two-point calibration (pH 4.0 and 10.0) to be 5.6, 6.1, and 5.2, respectively. Relative concentrations were confirmed by UV-Vis absorbance measurements at 210 nm.¹² Thermal melt experiments of **1-3** were carried out on a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD) equipped with a Peltier thermoelectric temperature control system, by monitoring the ellipticity at 225 nm from 5 to 70 °C. Fluctuation in the temperature contributes to an error of ± 1.0 °C to measurement of the temperature and represents the greatest source of error. The CD intensity and wavelength of the spectropolarimeter were calibrated using solutions of *d*-10-camphorsulphonic acid.⁸³ A scan was run from 200 to 250 nm in duplicate at 5 °C intervals using a scanning rate of 10 nm/min, response time of 4 sec, data pitch of 0.1 nm, and bandwidth of 2 nm. The baseline was corrected by subtraction of the solvent spectra, and for **2** and **3** the

spectrum of D-galactose (0.84 mM) was subtracted. After a 15 min heating period between temperatures, the sample was allowed to equilibrate for 15 min before data acquisition; giving an approximate heating rate of 0.16 °C/min. The reversibility of the thermal denaturation was judged by recovery of the signal at 25 °C at 225 nm (by 85-92%) at the conclusion of the experiment. The molar ellipticity per mean residue ($[\theta]$) was calculated from the average spectrophotometer output (millidegrees) at 225 nm using the formula:

$$[\theta] = (M \cdot \theta) / (10 \cdot c \cdot l \cdot n)$$

where M is the molecular weight of the sample ($\text{g}\cdot\text{mol}^{-1}$), θ is the CD signal (millidegrees), c is the peptide concentration in $\text{g}\cdot\text{L}^{-1}$, l is the path length of the cell (cm), and n is the number of amide bonds in the peptide (22). The temperature at the midpoint of the thermal transition (T_m) as well as the van't Hoff enthalpy (ΔH_{VH}) and entropy (ΔS_{VH}) values were determined by fitting the molar ellipticity at 225 nm to a two-state model using Mathematica v6.0 (Wolfram Research, Inc) according to the equation:

$$y = \frac{(\Delta N + mN \cdot x) + (\text{Exp}[-(\Delta H_m + x \cdot \Delta S_m)/(R \cdot x)]) \cdot (\Delta U + mU \cdot x)}{1 + \text{Exp}[-(\Delta H_m + x \cdot \Delta S_m)/(R \cdot x)]}$$

where y is the optical parameter (mdeg), x is the temperature (Kelvin), ΔH_m and ΔS_m are the differences between the folded and unfolded enthalpy and entropy at the transition midpoint (T_m), respectively, mN and mU are the slopes of the curve before and after the transition, respectively, and ΔN and ΔU are the zero temperature optical values of the fully folded and unfolded forms of the triple helical model peptide, respectively. The error values calculated for ΔH and ΔS are reported based on the confidence interval calculated for each parameter. The value of T_m was calculated according to:

$$T_m = \Delta H / \Delta S$$

The corresponding error value in T_m was calculated using the asymptotic standard error for ΔH and ΔS according to the equation:

$$\sigma_T = [(\sigma_{\Delta H}^2 / \Delta H^2) + (\sigma_{\Delta S}^2 / \Delta S^2)]^{1/2}$$

Where σ_T is the relative uncertainty in T_m , $\sigma_{\Delta H}$ is the uncertainty in ΔH , and $\sigma_{\Delta S}$ is uncertainty in ΔS . The calculated error in T_m (± 0.15 - 0.25 °C) was less than the instrumental error.

The cooperativity of the transitions was quantified by calculating the temperature range (ΔT) over which 80% of the unfolding occurs and corresponds to a K_{eq} range of 0.25 to 4.0. For **1-3**, the ΔT values were found to be 15.1 °C, 4.4 °C, and 7.0 °C, respectively.

6.7 References

- 1) a) Nimni, M. E. *Collagen*; CRC Press: Boca Raton, FL, 1988; b) Brodsky, B.; Shah N. K. *FASEB J.* **1995**, *9*, 1537-1546; c) Myllyharju, J.; Kivirikko, K. I. *Ann. Med. (Helsinki)* **2001**, *33*, 7-21.
- 2) Kadler, K. *Protein Profile* **1994**, *1*, 519-638.
- 3) a) Ramachandran, G. N.; Kartha, G. *Nature* **1954**, *174*, 269-270; b) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593-595; c) Cowan, P. M.; McGavin, S.; North, A. C. *Nature* **1955**, *176*, 1062-1064.

- 4) a) Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915–916; b) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483–506.
- 5) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75-81.
- 6) a) Inouye, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1976**, *420*, 133-141; b) Inouye, K.; Kobayashi, Y.; Kyogoku, Y.; Kishida, Y.; Sakakibara, S.; Prockop, D. J. *Arch. Biochem. Biophys.* **1982**, *219*, 198-203.
- 7) Persikov, A. V.; Ramshaw, J. A. M.; Kirkpatrick, A.; Brodsky, B. *Biochemistry*, **2000**, *39*, 14960-14967.
- 8) Reviewed in: a) Moroder *J. Pept. Sci.* **2005**, *11*, 258-261; b) Brodsky, B.; Persikov, A. V. *Adv. Prot. Chem.* **2005**, *70*, 301-339; c) Jenkins, C. L.; Raines, R. T. *Nat. Prod. Rep.* **2002**, *19*, 49-59; d) Engel, J.; Bächinger, H. *Top Curr Chem.* **2005**, *247*, 7-33.
- 9) Berg, R. A.; Prockop, D. J. *Biochem. Biophys. Res. Commun.* **1973**, *52*, 115-120.
- 10) Sakakibara, S.; Inouye, K.; Shudo, K.; Kishida, Y.; Kobayashi, Y.; Prockop, D. J. *Biochim. Biophys. Acta.* **1973**, *303*, 198-202.
- 11) a) Kobayashi, Y.; Sakai, R.; Kakiuchi, K.; Isemura, T. *Biopolymers* **1970**, *9*, 415-425; b) Burjanadze, T. V. *Biopolymers* **1992**, *32*, 941-949.
- 12) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777-778.
- 13) a) Haasnoot, C. A. G.; De Leeuw, F. A. A. M.; De Leeuw, H. P. M.; Altona, C. *Biopolymers* **1981**, *20*, 1211-1245; b) Panasik, N., Jr.; Eberhardt, E. S.; Edison, A. S.;

Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262-269; c) Cadamuro, S. A.; Reichold, R.; Kusebauch, U.; Musiol, H.-J.; Renner, C.; Tavan, P.; Moroder, L. *Angew. Chem. Int. Ed.* **2008**, *47*, 2143-2146; d) Garbay-Jaureguiberry, C.; Arnoux, B.; Prangé, T.; Wehri-Altenburger, S.; Pascard, C.; Roques, B. P. *J. Am. Chem. Soc.* **1980**, *102*, 1827-1837.

14) a) DeRider, M. L.; Wilkens, S. T.; Waddell, M. J.; Bretscher, L. E.; Weinhold, F.; Raines, R. T.; Markley, J. L. *J. Am. Chem. Soc.* **2002**, *124*, 2497-2505; b) Song, I. K.; Kang, Y. K. *J. Phys. Chem. B* **2006**, *110*, 1915-1927; c) Benzi, C.; Improta, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2002**, *23*, 341-350.

15) Mooney, S. D.; Kollman, P. A.; Klein, T. E. *Biopolymers* **2002**, *64*, 63-71.

16) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *58*, 459-464.

17) Vitagliano, L.; Berisio, R.; Mastrangelo, A.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2001**, *10*, 2627-2632.

18) Improta, R.; Benzi, C.; Barone, V. *J. Am. Chem. Soc.* **2001**, *123*, 12568-12577.

19) Holmgren, S. K.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T. *Chem. Biol.* **1999**, *6*, 63-70.

20) a) Berisio, R.; Granata, V.; Vitagliano, L.; Zagari, A. *J. Am. Chem. Soc.* **2004**, *126*, 11402-11403; b) Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nishio, H.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *J. Pept. Sci.* **2005**, *11*, 609-616.

- 21) Schumacher, M.; Mizuno, K.; Bächinger, H. P. *J. Biol. Chem.* **2005**, *280*, 20397-20403.
- 22) Mizuno, K.; Hayashi, T.; Peyton, D. H.; Bächinger, H. P. *J. Biol. Chem.* **2004**, *279*, 38072-38078.
- 23) Kawahara, K.; Nishi, Y.; Nakamura, S.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *Biochemistry* **2005**, *44*, 15812-15822.
- 24) a) Ramachandran, G. N.; Doyle, B. B.; Bloot, E. R.; *Biopolymers* **1968**, *6*, 1771-1775; b) Ramachandran, G. N.; Bansal, M.; Bhatnagar, R. S. *Biochim. Biophys. Acta* **1973**, *322*, 166-171.
- 25) Bella, J.; Brodsky, B.; Berman, H. M. *Structure* **1995**, *3*, 893-906.
- 26) Kramer, R. Z.; Vitagliano, L.; Bella, J.; Berisio, R.; Mazzarella, L.; Brodsky, B.; Zagari, A.; Berman, H. M. *J. Mol. Biol.* **1998**, *280*, 623-638.
- 27) Melacini, G.; Bonvin, A. M.; Goodman, M.; Boelens, R.; Kaptein, R. *J. Mol. Biol.* **2000**, *300*, 1041-1048.
- 28) Engel, J.; Chen, H.-T.; Prockop, D. J.; Klump, H. *Biopolymers* **1977**, *16*, 601-622.
- 29) a) Howard, J. A. K.; Hoy, V. J.; O'Hagan, D.; Smith, G. T. *Tetrahedron* **1996**, *52*, 12613-12622; b) Dunitz, J. D.; Taylor, R. *Eur. J. Chem.* **1997**, *3*, 89-98; c) Hiyama, T.; Kanie, K.; Kusumoto, T.; Morizawa, Y.; Shimizu, M. in *Organofluorine Compounds: Chemistry and Applications*; Hiyama, T., Ed.; Springer, Berlin, Germany, 2000; pp 10-12.

- 30) Holmgren, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666-667.
- 31) De Simone, A.; Vitagliano, L.; Berisio, R. *Biochem. Biophys. Res. Commun.* **2008**, *372*, 121-125.
- 32) a) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem.* **1999**, *125*, 310–318; b) Berisio, R.; Vitagliano, L.; Mazzarella, L.; Zagari, A. *Biopolymers* **2001**, *56*, 8-13.
- 33) Kramer, R. Z.; Venugopal, M. G.; Bella, J.; Mayville, P.; Brodsky, B.; Berman, H. M. *J. Mol. Biol.* **2000**, *301*, 1191-1205.
- 34) Okuyama, K.; Hongo, C.; Fukushima, R.; Wu, G.; Narita, H.; Noguchi, K.; Tanaka, Y.; Nishino, N. *Biopolymers* **2004**, *76*, 367–377.
- 35) Okuyama, K.; Hongo, C.; Wu, G.; Mizuno, K.; Noguchi, K.; Ebisuzaki, S.; Tanaka, Y.; Nishino, N.; Bächinger, H. P. *Biopolymers* **2009**, *91*, 361-372.
- 36) Nishi, Y.; Uchiyama, S.; Doi, M.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *Biochemistry* **2005**, *44*, 6034-6042.
- 37) Terao, K.; Mizuno, K.; Murashima, M.; Kita, Y.; Hongo, C.; Okuyama, K.; Norisuye, T.; Bächinger, H. P. *Macromolecules* **2008**, *41*, 7203-7210.
- 38) a) Bann, J. G.; Bächinger, H. P. *J. Biol. Chem.* **2000**, *275*, 24466–24469; b) Mizuno, K.; Hayashi, T.; Bächinger, H. P. *J. Biol. Chem.* **2003**, *278*, 32373–32379.
- 39) Babu, I. R.; Ganesh, K. N. *J. Am. Chem. Soc.* **2001**, *123*, 2079-2080.

- 40) a) Bann, J. G.; Peyton, D. H.; Bächinger, H. P. *FEBS Lett.* **2000**, *473*, 237-240; b) Bann, J. G.; Bächinger, H. P.; Peyton, D. H. *Biochemistry* **2003**, *42*, 4042-4048.
- 41) a) Varki, A. *Glycobiology* **1993**, *3*, 97-130; b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
- 42) a) Petrescu, A. J.; Milac, A. L.; Petrescu, S. M.; Dwek, R. A.; Wormald, M. R. *Glycobiology* **2004**, *14*, 103-114; b) Weerapana, E.; Imperiali, B. *Glycobiology* **2006**, *16*, 91R-101R.
- 43) a) Imperiali, B. *Acc. Chem. Res.* **1997**, *30*, 452-459; b) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1-27; c) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R-56R; d) Paul, G.; Lottspeich, F.; Wielland, F. *J. Biol. Chem.* **1986**, *261*, 1020-1024; e) Kern, G.; Kern, D.; Jaenicke, R.; Seckler, R. *Protein Sci.* **1993**, *2*, 1862-1868.
- 44) a) Chu, F. K.; Trimble, R. B.; Maley, F. *J. Biol. Chem.* **1978**, *253*, 8691-8693; b) Bernard, E. R.; Sheila, A. N.; Olden, K. *J. Biol. Chem.* **1983**, *258*, 12198-12202; c) Fisher, J. F.; Harrison, A. W.; Bundy, G. L.; Wilkinson, K. F.; Rush, B. D.; Ruwart, M. J. *J. Med. Chem.* **1991**, *34*, 3140-3143; d) Rudd, P. M.; Joao, H. C.; Coghill, E.; Fiten, P.; Saunders, M. R.; Opdenakker, G.; Dwek, R. A. *Biochemistry* **1994**, *33*, 17-22; e) Mehta, S.; Meldal, M.; Duus, J. O.; Bock, K. *J. Chem. Soc., Perkin Trans. 1*, **1999**, 1445-1451.
- 45) Mer, G.; Hietter, H.; Lefèvre, J.-F. *Nat. Struc. Biol.* **1996**, *3*, 45-53.
- 46) a) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 97-101; b) Wang, C.; Eufemi, M.; Turano, C.; Giartosio, A. *Biochemistry* **1996**, *35*, 7299-7307.

- 47) Shental-Bechor, D.; Levy, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8256-8261.
- 48) a) Riederer, M. A.; Hinnen, A. *J. Bacteriol.* **1991**, *173*, 3539-3546; b) Rudd, P. M.; Joao, H. C.; Coghill, E.; Fiten, P.; Saunders, M. R.; Opdenakker, G.; Dwek, R. A. *Biochemistry* **1994**, *33*, 17-22; c) Wang, C.; Eufemi, M.; Turano, C.; Giartosio, A. *Biochemistry* **1996**, *35*, 7299-7307; d) Helenius, A.; Aebi, M. *Annu. Rev. Biochem.* **2004**, *73*, 1019-1049; e) Sinha, S.; Surolia, A. *Biophys. J.* **2007**, *92*, 208-216.
- 49) Marin, J.; Blaton, M.-A.; Briand, J.-P.; Chiocchia, G.; Fournier, C.; Guichard, G. *ChemBioChem* **2005**, *6*, 1796-1804.
- 50) Marin, J.; Didierjean, C.; Aubry, A.; Briand, J.-P.; Guichard, G. *J. Org. Chem.* **2002**, *67*, 8440-8449.
- 51) a) Butler, W. T.; Cunningham, L. W. *J. Biol. Chem.* **1966**, *241*, 3882-3888; b) Spiro, R. G. *J. Biol. Chem.* **1969**, *244*, 602-612; c) Miller, E. J. *Biochemistry* **1971**, *10*, 1652-1659; d) Kivirikko, K. I.; Myllylä, R. *Int. Rev. Connect. Tissue Res.* **1979**, *8*, 23-72.
- 52) a) Notbohm, H.; Nokelainen, M.; Myllyharju, J.; Fietzek, P. P.; Müller, P. K.; Kivirikko, K. I. *J. Biol. Chem.* **1999**, *274*, 8988-8992; b) Brinckmann, J.; Notbohm, H.; Tronnier, M.; Acil, Y.; Fietzek, P. P.; Schmeller, W.; Müller, P. K.; Bätge, B. *J. Invest. Dermatol.* **1999**, *113*, 617-621; c) Norman, K. R.; Moerman, D. G. *Dev. Biol.* **2000**, *227*, 690-705; d) Ruotsalainen, H.; Sipilä, L.; Vapola, M.; Sormunen, R.; Salo, A. M.; Uitto, L.; Mercer, D. K.; Robins, S. P.; Risteli, M.; Attila, A.; Fässler, R.; Myllylä, R. *J. Cell. Sci.* **2006**, *119*, 625-635; e) Sipilä, L.; Ruotsalainen, H.; Sormunen, R.; Baker, N. L.;

Lamandé, S. R.; Vapola, M.; Wang, C.; Sado, Y.; Aszodi, A.; Myllylä, R. *J. Biol. Chem.* **2007**, *282*, 33381-33388.

53) Myllylä, R.; Wang, C.; Heikkinen, J.; Juffer, A.; Lampela, O.; Risteli, M.; Ruotsalainen, H.; Salo, A.; Sipilä, L. *J. Cell. Physiol.* **2007**, *212*, 323-329.

54) a) Salo, A. M.; Cox, H.; Farndon, P.; Moss, C.; Grindulis, H.; Risteli, M.; Robins, S. P.; Myllylä, R. *Am. J. Hum. Genet.* **2008**, *83*, 495-503; b) Wang, C.; Kovanen, V.; Raudasoja, P.; Eskelinen, S.; Pospiech, H.; Myllylä, R. *J. Cell. Mol. Med.* **2009**, *13*, 508-521.

55) a) Lauer-Fields, J. L.; Malkar, N. B.; Richet, G.; Drauz, K.; Fields, G. B. *J. Biol. Chem.* **2003**, *278*, 14321-14330; b) Van den Steen, P. E.; Proost, P.; Brand, D. D.; Kang, A. H.; Van Damme, J.; Opdenakker, G. *Biochemistry* **2004**, *43*, 10809-10816.

56) a) Broddefalk, J.; Bäcklund, J.; Almqvist, F.; Johansson, M.; Holmdahl, R.; Kihlberg, J. *J. Am. Chem. Soc.* **1998**, *120*, 7676-7683; b) Bäcklund, J.; Carlsen, S.; Höger, T.; Holm, B.; Fugger, L.; Kihlberg, J.; Burkhardt, H.; Holmdahl, R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9960-9965.

57) a) Lamport, D. T. A. *Nature* **1967**, *216*, 1322-1324; b) Lamport, D. T. A.; Miller, D. H.; *Plant Physiol.* **1971**, *48*, 454-456.

58) Lamport, D. T. A. in *Recent Advances in Phytochemistry*; Loewus, F. A.; Runeckles, V. C., Eds.; Plenum Publishing Corp., New York, 1977; pp. 79-115.

- 59) a) Knox, R. B.; Clarke, A.; Harrison, S.; Smith, P.; Marchalonis, J. J. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 2788-2792; b) Showalter, A. M. *Plant Cell* **1993**, *5*, 9-23; c) Bowles, D. J. *Annu. Rev. Biochem.* **1990**, *59*, 873-907; d) Sommer-Knudsen, J.; Bacic, A.; Clarke, A. E. *Phytochemistry* **1998**, *47*, 483-497; e) José-Estanyol, M.; Puigdomènech, P. *Plant Physiol. Biochem.* **2000**, *38*, 97-108; f) Khashimova, Z. S. *Chem. Nat. Comput.* **2003**, *39*, 229-236.
- 60) Kieliszewski, M. J. *Phytochemistry* **2001**, *57*, 319-323.
- 61) Homer, R. B.; Roberts, K. *Planta* **1979**, *146*, 217-222.
- 62) a) Lamport, D. T. A. in *The Biochemistry of Plants*; Preiss, J., Ed.; Academic Press, New York, 1980; Vol. 3, pp. 501-541; b) van Holst, G.-J.; Varner, J. E. *Plant Physiol.* **1984**, *74*, 247-251; c) Shpak, E.; Barbar, E.; Leykam, J. F.; Kieliszewski, M. J. *J. Biol. Chem.* **2001**, *276*, 11272-11278; d) Ferris, P. J.; Woessner, J. P.; Waffenschmidt, S.; Kilz, S.; Drees, J.; Goodenough, U. W. *Biochemistry* **2001**, *40*, 2978-2987; e) Ferris, P. J.; Waffenschmidt, S.; Umen, J. G.; Ishida, K.; Kubo, T.; Lau, J.; Goodenough, U. W. *Plant Cell* **2005**, *17*, 597-615.
- 63) Yamada, H.; Sasaki, T.; Niwa, S.; Oishi, T.; Murata, M.; Kawakami, T.; Aimoto, S. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 5677-5680.
- 64) Jenkins, C. L.; McCloskey, A. I.; Guzei, I. A.; Eberhardt, E.; Raines, R. T. *Biopolymers* **2005**, *80*, 1-8.
- 65) Owens, N. W.; Braun, C.; O'Neil, J. D.; Marat, K.; Schweizer, F. *J. Am. Chem. Soc.* **2007**, *129*, 11670-11671.

- 66) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269*, 1273-1278.
- 67) Okuyama, K.; Wu, G.; Jiravanichanun, N.; Hongo, C.; Noguchi, K. *Biopolymers* **2006**, *84*, 421-432.
- 68) a) Live, D. H.; Williams, H. J.; Kuduk, S. D.; Schwarz, J. B.; Glunz, P. W.; Chen, X.-T.; Sames, D.; Kumar, R. A.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3489-3493; b) Coltart, D. M.; Royyuru, A. K.; Williams, L. J.; Glunz, P. W.; Sames, D.; Kuduk, S. D.; Schwarz, J. B.; Chen, X.-T.; Danishefsky, S. J.; Live, D. H. *J. Am. Chem. Soc.* **2002**, *124*, 9833-9844; c) Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421-8425.
- 69) Griffin, F. K.; Paterson, D. E.; Murphy, P. V.; Taylor, R. J. K. *Eur. J. Org. Chem.* **2002**, 1305-1322.
- 70) a) Kisfaludy, L.; Schon, I. *Synthesis* **1983**, *4*, 325-327; b) Atherton, E.; Cameron, L. R.; Sheppard, R. C.; *Tetrahedron* **1988**, *44*, 843-857; c) Meldal, M.; Jensen, K. J. *J. Chem. Soc., Chem. Commun.* **1990**, 483-485.
- 71) Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. *J. Am. Chem. Soc.* **1996**, *118*, 10351-10358.
- 72) a) Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2005**, *127*, 15923-15932; b) Shoulders, M. D.; Hodges, J. A.; Raines, R. T. *J. Am. Chem. Soc.* **2006**, *128*, 8112-8113.

- 73) a) Long, C. G.; Braswell, E.; Zhu, D.; Apigo, J.; Baum, J.; Brodsky, B. *Biochemistry* **1993**, *32*, 11688-11695; b) Venugopal, M. G.; Ramshaw, J. A. M.; Braswell, E.; Zhu, D.; Brodsky, B. *Biochemistry* **1994**, *33*, 7948-7956; c) Engel, J.; Bächinger, H. P. *Matrix Biol.* **2000**, *19*, 235-244.
- 74) Doi, M.; Nishi, Y.; Uchiyama, S.; Nishiuchi, Y.; Nakazawa, T.; Ohkubo, T.; Kobayashi, Y. *J. Am. Chem. Soc.* **2003**, *125*, 9922-9923.
- 75) Ramshaw, J. A. M.; Shah, N. K.; Brodsky, B. *J. Struct. Biol.* **1998**, *122*, 86-91.
- 76) a) Ottl, J.; Gabriel, D.; Murphy, G.; Knauper, V.; Tominaga, Y.; Nagase, H.; Kroger, M.; Tschesche, H.; Bode, W.; Moroder, L. *Chem. Biol.* **2000**, *7*, 119-132; b) Lauer-Fields, J.; Brew, K.; Whitehead, J. K.; Li, S.; Hammer, R. P.; Fields, G. B. *J. Am. Chem. Soc.* **2007**, *129*, 10408-10417.
- 77) a) Renner, C.; Saccà, B.; Moroder, L. *Biopolymers* **2004**, *76*, 34-47; b) Paramonov, S. E.; Gauba, V.; Hartgerink, J. D. *Macromolecules* **2005**, *38*, 7555-7561; c) Kishimoto, T.; Morihara, Y.; Osanai, M.; Ogata, S.; Kamitakahara, M.; Ohtsuki, C.; Tanihara, M. *Biopolymers* **2005**, *79*, 163-172; d) Kotch, F. W.; Raines, R. T. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 3028-3033; e) Leitinger, B.; Hohenester, E. *Matrix Biol.* **2007**, *26*, 146-155; f) Berisio, R.; De Simone, A.; Ruggiero, A.; Improta, R.; Vitagliano, L. *J. Pept. Sci.* **2009**, *15*, 131-140.
- 78) Zong, Y.; Xu, Y.; Liang, X.; Keene, D. R.; Hook, A.; Gurusiddappa, S.; Hook, M.; Narayana, S. V. A. *EMBO J.* **2005**, *24*, 4224-4236.
- 79) Lee, C. H.; Singla, A.; Lee, Y. *Int. J. Pharma.* **2001**, *221*, 1-22.

80) a) Wang, A. Y.; Mo, X.; Chen, C. S.; Yu, S. M. *J. Am. Chem. Soc.* **2005**, *127*, 4130-4131; b) Badylak, S. F. *Biomaterials* **2007**, *28*, 3587-3593; c) Koide, T. *Phil. Trans. R. Soc. B* **2007**, *362*, 1281-1291.

81) a) Tanaka, S.; Avigad, G.; Eikenberry, E. F.; Brodsky, B. *J. Biol. Chem.* **1988**, *263*, 17650-17657; b) Grandhee, S. K.; Monnier, V. M. *Biol. Chem.* **1991**, *266*, 11649-11655; c) Bailey, A. J.; Sims, T. J.; Avery, N. C.; Miles, C. A. *Biochem. J.* **1993**, *296*, 489-496; d) Fu, M. X.; Wells-Knecht, K. J.; Blackledge, J. A.; Lyons, T. J.; Thorpe, S. R.; Baynes, J. W. *Diabetes* **1994**, *43*, 676-683; e) Bailey, A. J.; Sims, T. J.; Avery, N. C.; Halligan, E. P. *Biochem. J.* **1995**, *305*, 385-390; f) Sady, C.; Khosrof, S.; Nagaraj, R. *Biochem. Biophys. Res. Commun.* **1995**, *214*, 793-797; g) Paul, R. G.; Bailey, A. J. *Int. J. Biochem. Cell Biol.* **1996**, *28*, 1297-1310; h) Sell, D. R.; Kleinman, N. R.; Monnier, V. M. *FASEB J.* **2000**, *14*, 145-156.

82) Yang, Y.-Y.; Ficht, S.; Brik, A.; Wong, C.-H. *J. Am. Chem. Soc.* **2007**, *129*, 7690-7701.

83) Wu, C. S.; Chen, G. C. *Anal. Biochem.* **1989**, *177*, 178-182.

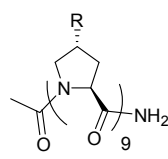
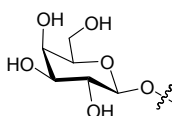
84) a) Sorenson, J. M.; Hura, G.; Soper, A. K.; Pertsemlidis, A.; Head-Gordon, T. J. *Phys. Chem. B* **1999**, *103*, 5413-5426; b) Cheung, M. S.; Garcia, A. E.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 685-690.

85) Czechura, P.; Tam, R. Y.; Dimitrijevic, E.; Murphy, A. V.; Ben, R. N. *J. Am. Chem. Soc.* **2008**, *130*, 2928-2929.

Chapter 7: Contiguous *O*-Galactosylation of 4*R*-Hydroxy-L-Proline Residues Forms Hyper-Stable Polyproline II Helices

Neil W. Owens^{*#§∞} and Frank Schweizer^{*§∞}

^{*}Designed research; [#]Performed research; [§]Analyzed data; [∞]Wrote the paper

	Polyproline Peptide	T _m (+/-3 °C)
	1: R = H	22
	2: R = OH	38
	3: R = 	70

7.1 Abstract

The polyproline II (PPII) conformation, present in the major structural proteins of both plants and animals, has inspired the development of rigid molecular scaffolds for understanding the role of the PPII conformation in biology, and for the development of functional materials. Here, we synthesized a series of model polyproline peptides Ac-(Pro)₉-NH₂ (**1**), Ac-(Hyp)₉-NH₂ (**2**), and Ac-[Hyp(β-D-Gal)]₉-NH₂ (**3**) to investigate the effects of Hyp *O*-glycosylation on the conformational stability of the PPII helix, as found in the hydroxyproline-rich glycoproteins found in plant cell walls. We found that contiguous *O*-galactosylation of Hyp residues causes a decrease in the intensity of CD maxima, which may be due to shielding of the amide groups from the carbohydrate rings. Thermal melting experiments show that *O*-galactosylation causes the transition from the PPII to β-strand conformation to occur at a much higher temperature; this has implications for the role of Hyp *O*-glycosylation as found in nature, and for the development of conformationally rigid scaffolds that allow for further functionalization.

7.2 Introduction

The polyproline type II (PPII) conformation, characterized as a left-handed helix ($\phi = -78^\circ$ and $\psi = 146^\circ$) with 3 amino acid residues per turn and all *trans* amide bonds ($\omega = 180^\circ$),¹⁻³ has attracted attention because of the discovery of its presence in both folded⁴⁻⁶ and unfolded⁷⁻¹⁶ protein structures. The PPII conformation has now been implicated in numerous protein-protein recognition interactions,^{6,17} as well as playing an important role in the protein folding process.^{12,15,18} Improved understanding of the characteristics of the PPII conformation has come from crystallographic,^{1,2,19} spectroscopic,^{2,20-22} and computational²³ analysis of model peptides. Both L-proline (Pro) and (4*R*)-hydroxy-L-proline (Hyp) have demonstrated the highest propensity for forming the PPII conformation among the naturally occurring amino acids,²⁴⁻²⁸ likely because their ϕ - and ψ -angles are ideally predisposed to do so; the ϕ -angle of Pro and Hyp is fixed at -75° because the prolyl side chain is cyclised onto the polypeptide backbone. Furthermore, a ψ -angle of 150° has been theorized to lead to a favourable $n \rightarrow \pi^*$ interaction from the amide carbonyl oxygen atom (O_{i-1}) to the amide carbonyl group ($C'_i=O_i$).²⁹ Therefore, Pro and Hyp-rich domains of polypeptides bestow the extended PPII conformation with considerable structural stability.

Nature has exploited the structural rigidity of the PPII conformation, as exemplified by its prevalence in both animal and plant structural proteins. Collagen is the most abundant structural protein found in humans and animals,³⁰ and is characterized by three polypeptide strands of the repeating sequence ($X_{aa}-Y_{aa}-Gly$), where X_{aa} is often Pro and Y_{aa} is often Hyp; each strand adopts a PPII helix, which associate to form a right-handed triple helix.³¹ Also, the hydroxyproline-rich glycoproteins (HRGPs) are extremely

widespread in the plant kingdom;^{32,33} they are the major protein components of the extracellular matrix of algae and land plants and have been found to be critical for proper cell wall formation.³⁴ Of several subclasses of HRGPs, the extensins are the main insoluble proteins associated with the cell wall;³⁵ they are characterized by a rigid PPII conformation, large amounts of Hyp and serine (Ser), and extensive *O*-glycosylation of Hyp residues, which is a unique post-translational modification of proteins.³² The Hyp *O*-linked glycans are usually small (1-4 residues), linear β -linked homooligosaccharides of D-arabinose and larger heteropolysaccharides of D-galactose, D-arabinose, and smaller amounts of other sugars.³² Glycosylation patterns of HRGPs seem to be based on peptide sequence: contiguous Hyp residues lead to Hyp *O*-arabinosylation, while non-contiguous Hyp residues lead to a mixture of Hyp *O*-arabinosylation and *O*-galactosylation.^{36,37} The functional consequences of glycosylation on HRGPs remains unclear, but efforts have been made to understand this modification. It has been suggested that the oligosaccharides increase HRGP solubility,³² resistance to proteolytic degradation,³² and thermal stability,³⁸⁻⁴¹ however the extent of this stabilization has never been quantified or studied in well-defined peptides.

The structural rigidity of the PPII conformation has inspired the development of molecular scaffolds that adopt predictable and well-defined conformations.⁴² Proline-rich peptides that adopt a PPII conformation have the added advantages of being capable of crossing cell membranes⁴³⁻⁴⁵ and being non-antigenic.⁴⁶ Therefore, these peptides have been used as molecular spacers,⁴⁷ molecular rulers,⁴⁸ cell-penetrating compounds,^{43,45,49} and as antibiotics.^{43,50} However, there are indications that the polyproline conformation is not completely rigid;⁵¹ proline is susceptible to isomerization from the *trans* to the *cis*

amide conformation, which causes local variation in the PPII structure. Thus, there is interest in developing polyproline derivatives in order to influence the conformational stability of the PPII conformation; both for understanding the role of PPII in biological contexts, such as in protein folding, and for the development of biocompatible biomaterials that have improved conformational stability.⁵² Recently, polyproline derivatives have been developed by incorporating proline analogues such as (2*S*,4*R*)-4-fluoroproline (Flp) and (4*R*)-4-azidoproline (Azp). Horng and Raines found that homooligomers of Flp increased the thermal stability of the PPII conformation relative to Pro.⁵³ The estimated melting temperatures (T_m) were found to be 27 °C for (Pro)₁₀, and 53 °C for (Flp)₁₀. Similarly, homooligomers of Azp prepared by Kümin and coworkers stabilized the PPII conformation as determined by spectroscopic analysis.⁵⁴ The ability of these proline analogues to stabilize the PPII conformation has been attributed to the 4-electron withdrawing group stabilizing the C^γ-exo conformation of the proline side chain through a stereoelectronic effect, which in turn stabilizes the *trans* amide configuration.

Recent work by our group found that Hyp *O*-galactosylation as found in HRGPs, seemed to cause an increased electron withdrawing effect in model diamides, as determined by an increase in the ¹³C chemical shift of the C^γ atom.⁵⁵ This did not translate into a measurable increase in the stability of the *trans* amide bond as determined by measurement of the *trans/cis* amide equilibrium ($K_{trans/cis}$) or the prolyl C^γ-exo pucker, but does indicate that glycosylation of Hyp residues will not preclude PPII helix formation. This provides the opportunity to study the effects of Hyp *O*-glycosylation on the stability of the PPII conformation in well-defined model peptides. We report here on the synthesis and characterization of *O*-glycosylated hydroxyproline model peptides. Our

aim is to gain further insight into the extent of stabilization of the PPII conformation brought about by a naturally-occurring Hyp *O*-glycosylation and for the development of structurally rigid scaffolds that can be applied as functional materials.

7.3 Results

7.3.1 Synthesis of model peptides 1-3

In a similar fashion to Horng⁵³ and Kümin,⁵⁴ we constructed model polyproline homooligomers: Ac-(Pro)₉-NH₂ (**1**), Ac-(Hyp)₉-NH₂ (**2**) and Ac-[Hyp(β-D-Gal)]₉-NH₂ (**3**) (Figure 7.3.1.1).

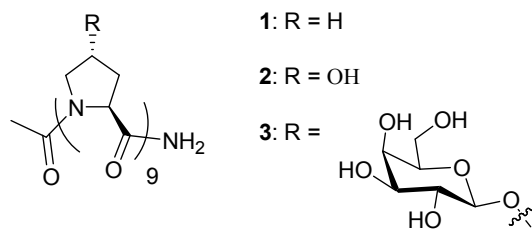
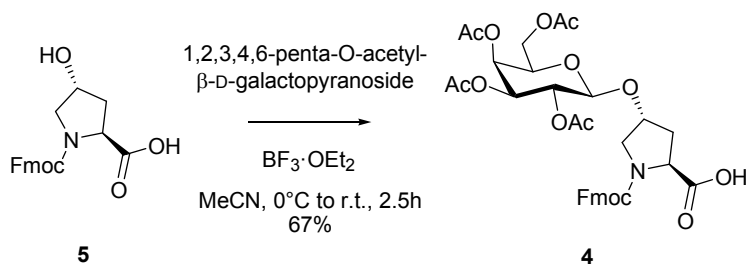


Figure 7.3.1.1: Model polyproline peptides **1-3**.

The *N*- and *C*-termini were amidated to prevent interference from charged groups. Contiguous glycosylation of the Hyp residues allow for maximum overlap of the sugar residues; with 3 residues per turn in the PPII helix, every fourth sugar residue is stacked on top of each other.¹⁻³

The model peptides **1-3** were synthesized using solid-phase coupling of their Fmoc-protected building blocks on a 30-100 μM scale. The Fmoc-[Hyp(β-D-Gal)]-OH building block was synthesized based on the procedure of Arsequell, Sàrries, and Valencia using the Lewis acid BF₃·OEt₂ to promote the addition of 1,2,3,4,6-penta-*O*-acetyl-β-D-galactopyranoside⁵⁶ to commercially available Fmoc-Hyp-OH, which gives

only the β -anomer in 67% yield (Scheme 7.3.1.1).⁵⁷ The assignment of the β -anomer was confirmed by the $^3J_{\text{H1,H2}}$ coupling constant of 8.0 Hz.



Scheme 7.3.1.1: Synthesis of Fmoc-[Hyp(β -D-Gal)]-OH (4).

7.3.2 Circular dichroism spectra of model peptides 1-3

Analysis of the far-UV circular dichroism (CD) spectra of **1-3** in water at 25 °C shows that all the peptides exhibit positive maxima between 220 and 230 nm and negative maxima between 200 and 210 nm, which are characteristic of the PPII conformation (Figure 7.3.2.1).^{8,58}

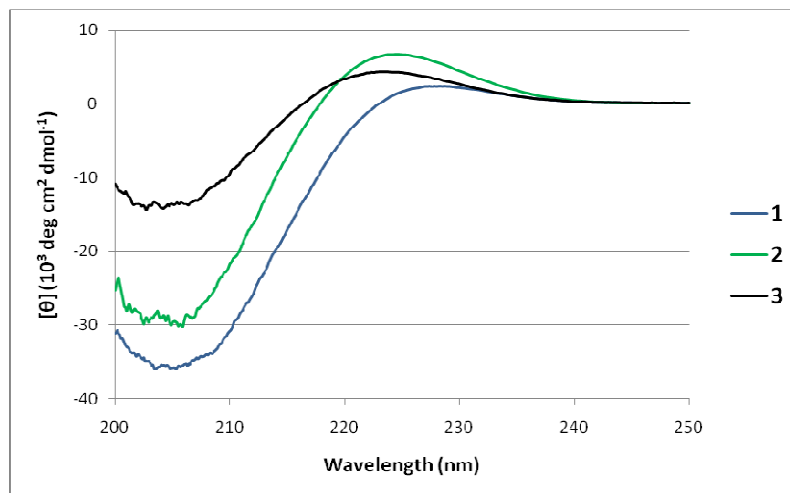


Figure 7.3.2.1: CD spectra of **1-3** in water at 25 °C at an approximate concentration of 0.4 mM, however all samples have the same relative concentration based on far ultraviolet (UV) absorbance measurements at 210 nm.

The peptide Ac-(Hyp)₉-NH₂ (**2**) showed a more intense positive band (λ_{\max} at 225 nm, $[\theta]$ of 6600 deg·cm²·dmol⁻¹) relative to Ac-(Pro)₉-NH₂ (**1**) (λ_{\max} at 228 nm, $[\theta]$ of 2300 deg·cm²·dmol⁻¹) but a weaker negative band (λ_{\min} at 204 nm, $[\theta]$ of -31,000 and -37,000 deg·cm²·dmol⁻¹, respectively), which is consistent with previous work.^{53,59} The *O*-glycosylated model peptide Ac-[Hyp(β -D-Gal)]₉-NH₂ (**3**) also exhibited a stronger positive band (λ_{\max} at 224 nm of 4200 deg·cm²·dmol⁻¹) relative to Ac-(Pro)₉-NH₂ (**1**), but a weaker negative band (λ_{\min} at 204 nm, $[\theta]$ of -14,500 deg·cm²·dmol⁻¹) relative to both **1** and **2**. The relative band strength ρ is the ratio of the maximum positive ellipticity to the maximum negative ellipticity, and has been used to interpret changes in CD spectra of poly(proline) and poly(hydroxyproline).⁵⁹ The ρ values for Ac-(Pro)₉-NH₂ and Ac-(Hyp)₉-NH₂ were found to be 0.06 and 0.21 respectively, which correlates well to previous studies of poly(proline) (0.06) and poly(hydroxyproline) (0.19),⁵⁹ while ρ for Ac-[Hyp(β -D-Gal)]₉-NH₂ was higher than both **1** and **2** at 0.29. Changes in ρ value have previously been attributed to conformational changes or changes in interactions between solvent molecules and peptide backbone carbonyl groups.⁵⁹

7.3.3 Conformational stability of model peptides 1-3

The CD spectra of **1-3** were monitored at 225 nm between 5 and 90 °C in pure water. Similar to other model peptides that exhibit PPII curves, there was a decrease in ellipticity at 225 nm as the temperature was increased (Figures 7.3.3.1-7.3.3.3).^{7,21,22,25,60-}

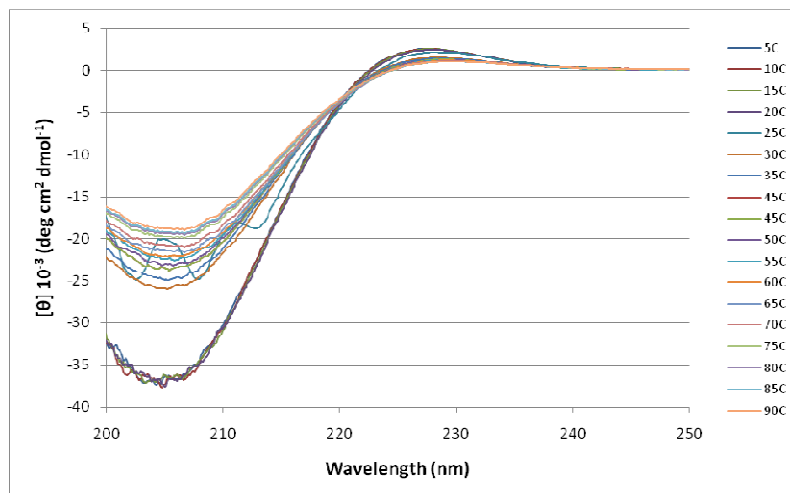


Figure 7.3.3.1: Overlay of CD spectra of **1** from 5 to 90 °C in 5 °C intervals in water from 200 to 250 nm. An isodichroic point is apparent at 220 nm.

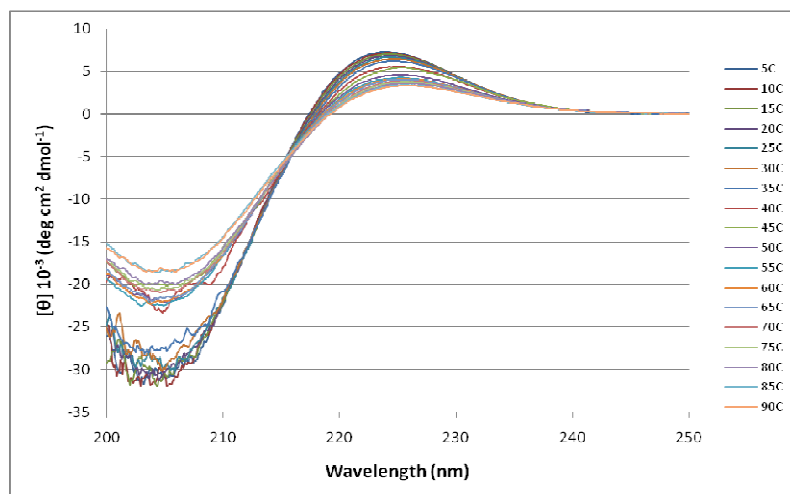


Figure 7.3.3.2: Overlay of CD spectra of **2** from 5 to 90 °C in 5 °C intervals in water from 200 to 250 nm. An isodichroic point is apparent at 215 nm.

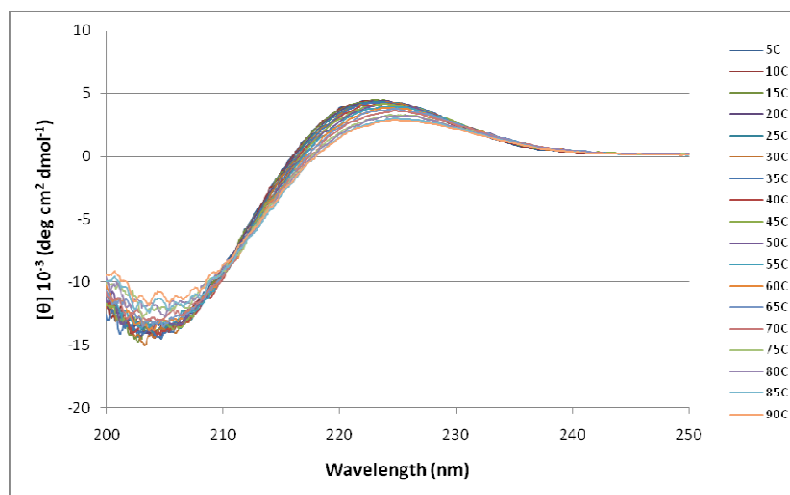


Figure 7.3.3.3: Overlay of CD spectra of **3** from 5 to 90 °C in 5 °C intervals in water from 200 to 250 nm. An isodichroic point is apparent at 211 nm.

An isodichroic point was found for peptides **1-3** at 220, 215, and 211 nm, respectively, which has been observed for related model peptides and is indicative of a coexistence of two conformations (Figures 7.3.3.1-7.3.3.3).⁶¹⁻⁶³ Plots of ellipticity ($[\theta]$) at 225 nm as a function of temperature for **1-3** exhibit subtle sigmoidal-type curves, which is further evidence of a coexistence of two conformations (Figure 7.3.3.4).

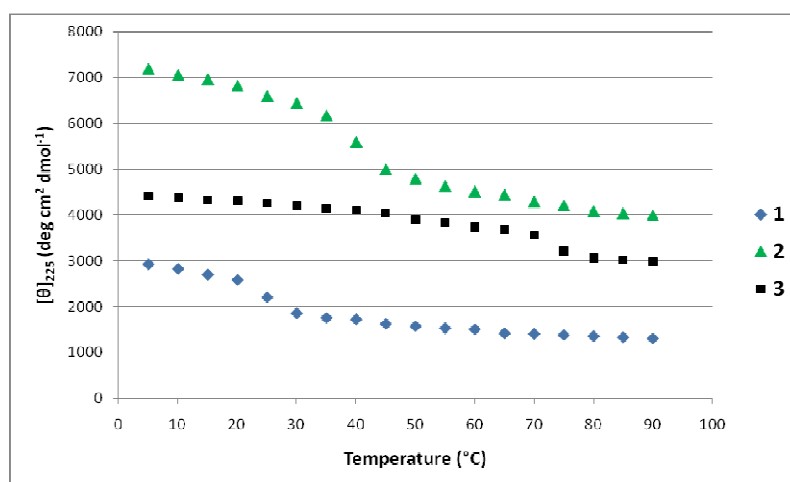


Figure 7.3.3.4: Ellipticity at 225 nm of **1-3** in water from 5 to 90 °C in 5 °C intervals; Each data point was recorded in duplicate, giving error values of ± 0.2 -6.0%.

The melting temperature (T_m) of **1-3** was estimated by taking the first-derivative of consecutive data points of molar ellipticity at 225 nm at various temperatures as seen in Figures 7.3.3.5-7.3.3.7.⁵³ The first-derivative data points were then fit using a spline curve to show the region of greatest change in slope. The T_m for Ac-(Hyp)₉-NH₂ (T_m of 38 ± 3 °C) was estimated to be higher compared to Ac-(Pro)₉-NH₂ (T_m of 22 ± 3 °C), while the conformational transition occurred at a much higher temperature for Ac-[Hyp(β -D-Gal)]₉-NH₂ (T_m of 70 ± 3 °C).

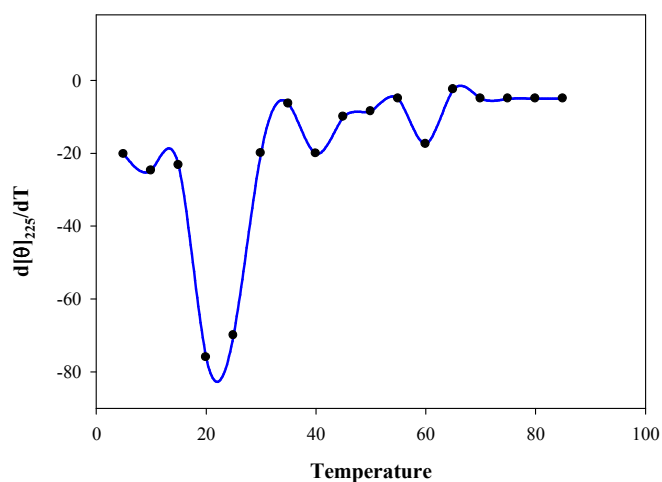


Figure 7.3.3.5: Spline curve of the first-derivative of the plot of **1** in Figure 7.3.3.4.

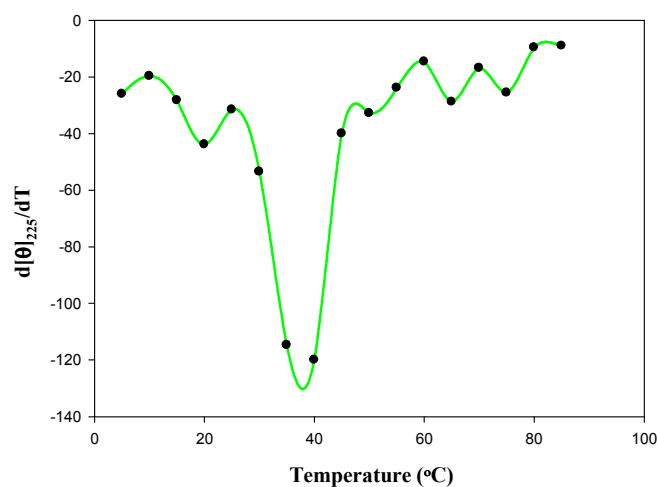


Figure 7.3.3.6: Spline curve of the first-derivative of the plot of **2** in Figure 7.3.3.4.

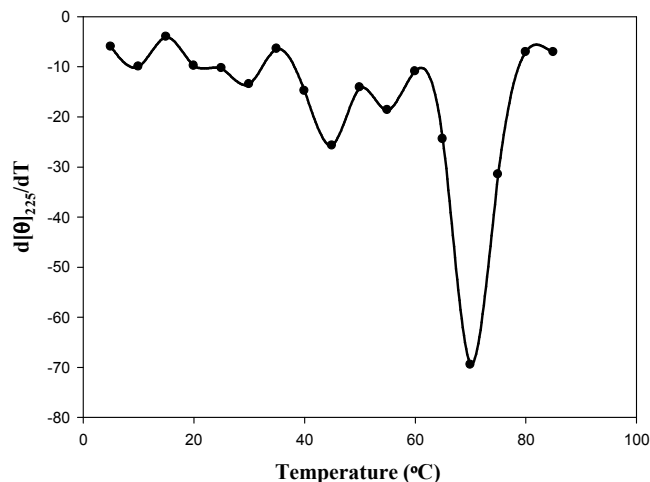


Figure 7.3.3.7: Spline curve of the first-derivative of the plot of **3** in Figure 7.3.3.4.

7.4 Discussion

7.4.1 Relative differences in the CD curves of the model peptides

The changes in the CD maxima of Ac-(Pro)₉-NH₂ (**1**) relative to Ac-(Hyp)₉-NH₂ (**2**) (Figure 7.3.2.1) may be accounted for by differences in their PPII conformation as determined by X-ray diffraction.⁵⁹ Brahmachari and coworkers found two different crystalline forms of polyHyp, one very similar to the crystal structure of polyproline,¹ and a different form that involved intramolecular hydrogen bonds from the 4-hydroxyl group of the Hyp_i residue to the Hyp_{i-1} carbonyl oxygen atom. Evidence for the second form was supported by IR studies that found water molecules are not strongly bound to polyHyp and a similar viscosity for poly(Pro) and poly(Hyp),⁶⁴ which would be expected if the 4-OH group was involved in intramolecular hydrogen bonding interactions. Furthermore, while the proline side chains of Ac-(Pro)₉-NH₂ are expected to be a mixture of C^γ-endo and C^γ-exo conformations,⁶⁵ the prolyl side chains of Ac-(Hyp)₉-NH₂ are

expected to be fixed solely in the C^γ-exo conformation.⁶⁶ Since CD is sensitive to the side-chain conformation of proline, the differences in the CD spectra may be due to this difference.⁶⁷ However, oligoproline model peptides incorporating 4*R*- and 4*S*-azidoproline⁵⁴ and fluoroproline,⁵³ which fix the prolyl pucker in a C^γ-exo and C^γ-endo conformation, respectively, did not exhibit similar relative differences in CD maxima. This indicates that the pucker change is not likely responsible for the differences in the intensity of the CD bands.

The CD spectra of Ac-[Hyp(β-D-Gal)]₉-NH₂ (**3**) relative to Ac-(Hyp)₉-NH₂ (**2**) shows a decrease in the relative intensity of both the positive and negative maxima (Figure 7.3.2.1), which may be attributed to a distortion of the PPII conformation.^{10,27,68} Analysis of PPII model peptides by Kelly and coworkers has shown that incorporation of non-proline residues in polyproline peptides (from PPP to PAAP) leads to a similar decrease in the CD maxima of **3** to **2**.²⁵ While this was correlated to a change in PPII conformation, the peptides were still considered to have significant PPII character. Overall, the molar ellipticities of **3** at λ_{max} and λ_{min} are consistent with naturally occurring HRGPs, such as GP1 (isolated from *Chlamydomonas reinhardtii*) and tomato extensin, which suggests that **3** adopts a similar PPII conformation to HRGPs.⁴⁰ The galactose residue may play some role in changing the PPII conformation of **3** as contiguous *O*-arabinylation of extensin Ser-(Hyp)₄ model peptides has been found to increase the intensity of CD maxima.³⁷ Therefore, it seems that either galactose residues inherently alter the PPII conformation, or the glycan size is important.

Interpretation and rationalization of changes in CD spectra of the PPII conformation is not a straightforward process.⁶⁹ Horng and Raines found that when Hyp

was replaced with 4-fluoroproline (Flp) in model polyproline peptides, a significant decrease in the intensity of the CD maxima was also observed.⁵³ Flp is known to stabilize the *trans* amide conformation and the C^γ-exo pucker relative to proline in a similar fashion to Hyp *O*-glycosylation.⁵⁵ Stabilization of these parameters that would be expected to maintain the PPII conformation, but caused an apparent decrease in PPII content based on the CD spectra. The authors gave no explanation for this apparent contradiction. Insight into the decrease in the intensity of the CD maxima of **3** may come from analysis of the ρ value, which is higher for **3** compared to **2** (0.29 and 0.21, respectively). An increase in ρ has been correlated to decreased solvation of the peptide backbone carbonyl groups.⁷⁰ Therefore, glycosylation may shield the peptide backbone carbonyl groups from solvent water molecules, which may account for the changes in the CD bands due to changes in the absorption of the amide groups. Regardless of the apparent change in PPII conformation in Ac-[Hyp(β -D-Gal)]₉-NH₂ relative to Ac-(Hyp)₉-NH₂, it may be misleading to interpret the decrease in CD maxima as destabilization of the PPII conformation.

7.4.2 Relative differences in the thermal melting curves of the model peptides

Monitoring the decrease in ellipticity as a function of temperature showed the presence of both a slight sigmoidal-type curve and an isodichroic point, which indicates that the polyproline peptides **1-3** underwent a conformational transition as the temperature was raised in aqueous conditions (Figures 7.3.3.1-7.3.3.4). Previous studies have shown that based on the measurement of $^3J_{\alpha N}$ of polyproline model peptides, while the PPII conformation is favoured at low temperatures, it exists in equilibrium with the

energetically similar⁷¹ β -strand conformation ($\phi = -120^\circ$ and $\psi = 120^\circ$), which becomes more favoured at high temperature.^{60,61,72-76} The β -strand conformation is characterized by a weak to no positive band near 220 nm and a negative band at 200 nm.^{60,61,72-76,89,90} The higher temperature state is unlikely to be a random coil type conformation, which does not exhibit a net positive cotton effect.⁷⁷

The melting of the PPII helix has previously been attributed to a non-cooperative transition, where destruction of the hydration shell leads to alteration or loosening of the PPII helix.⁶² Here, the peptides **1-3** showed a distinctive transition based on the CD melting curves. The sharpness of the transition seems to depend on the peptide sequence.⁷⁵ Model peptides incorporating non-proline residues: alanine (Ala), glycine (Gly), and lysine (Lys) have exhibited both linear^{25,26,60,61,73,78} and non-linear^{21,62,68,74} types of conformational transitions. Horng and Raines found a linear decrease in ellipticity with no clear thermal transition for the polyproline peptides: H-(Pro)₁₀-Gly-Tyr-OH and H-(Flp)₁₀-Gly-Tyr-OH, but approximated T_m values of 27 °C and 53 °C based on first-derivatives curves of the temperature-dependent CD spectra.⁵³ The derivative of the H-(Hyp)₁₀-Gly-Tyr-OH curve showed no clear inflection point, but was assumed to be between 27 and 53 °C. Therefore, while Flp exhibited a decrease in the intensity of its CD maxima, it was attributed to thermally stabilizing the PPII conformation relative to Hyp and Pro. Here, a similar effect is observed since Ac-[Hyp(β -D-Gal)]₉-NH₂ showed less intense CD maxima relative to Ac-(Hyp)₉-NH₂, but according to first-derivatives curves of the temperature-dependent CD spectra, the thermal transition occurred at a much higher temperature (T_m of 70 ± 3 °C compared to

38 ± 3 °C, respectively). Therefore, it seems contiguous *O*-glycosylation of Hyp significantly stabilizes the PPII conformation.

While a complete understanding of the stabilization of the PPII helix has yet to be achieved, hydrogen bonding interactions between the amino acids side chain to backbone nitrogen and carbonyl groups have been implicated.^{4,9} This is based on the study of model peptides, which has found that after Pro and Hyp, glutamine (Gln), arginine (Arg) and lysine (Lys) have the highest propensity to form PPII helices. This may explain the origin of the increase in thermal stabilization of Ac-(Hyp)₉-NH₂ relative to Ac-(Pro)₉-NH₂, which is supported by experimental study, and Ac-[Hyp(β-D-Gal)]₉-NH₂ compared to Ac-(Hyp)₉-NH₂; the multiple sugar groups may form stabilizing hydrogen bonds to the peptide backbone in the PPII conformation, leading to considerable conformational stability. Stabilization of the *trans* amide conformation and the prolyl pucker has been attributed to stabilizing the PPII conformation.⁵³ Previous work has shown that Hyp *O*-glycosylation does not lead to a measurable increase in the stabilisation of the *trans* amide conformation,⁵⁵ but this may require longer peptides to become apparent. Hydration has also been proposed to contribute to the stability of the PPII conformation;^{60,63,79} The increase in stabilization upon contiguous *O*-glycosylation in Ac-[Hyp(β-D-Gal)]₉-NH₂ relative to Ac-(Hyp)₉-NH₂ may be due to the change in solvation. Glycosylation is known to affect the hydration of peptides and proteins,⁸⁰ and restrict the conformational freedom of peptides and proteins,^{81,82} however the exact mechanism by which this is accomplished remains unclear.

7.5 Conclusions

Analysis of β -*O*-glycosylated polyproline model peptides indicates that Hyp *O*-galactosylation leads to a decrease in PPII character based on a decrease in the intensity of CD maxima. This may be due to a loss of PPII character, or may be attributed to shielding of the amide groups by the carbohydrate structures. Monitoring ellipticity as a function of temperature has shown the presence of isodichroic points for **1-3**, which indicates that equilibrium exists between the PPII conformation and a high-temperature state; believed to be a β -strand conformation. The origin of stabilization of **3** is likely due to stronger hydrogen bonding interactions from the multiple sugar residues to backbone C=O groups, or due to changes in hydration of the peptide.

Previously, it was shown that glycosylation of HRGPs causes a change in the intensity of CD maxima, which was correlated to changes in the PPII stability, but is more accurately attributed to a change in PPII conformation. Here, we have shown with well-defined model peptides that contiguous *O*-glycosylation leads to an increase in conformational stability; similar increases in the conformational stability of HRGPs likely also occur. Whether non-contiguous glycosylation or larger glycan structures affect this stability remains to be explored.

Since glycosylation leads to considerable conformational stability, these peptides may lead to the development of scaffolds and biomaterials that have improved conformational stability. The glycans also provide the opportunity for further derivatization, which can lead to facile modification of the scaffold. Furthermore, the use of polyproline as a well-defined secondary structure scaffold for attaching glycans⁸³ may

find use for the study of carbohydrate-carbohydrate⁸⁴ interactions and carbohydrate-lectin binding interactions.⁸⁵

7.6 Experimental Procedures

7.6.1 General Experimental

Protected amino acids were purchased from Bachem (Bachem Bioscience, Inc., King of Prussia, PA). Reagent grade solvents were used without further purification. Thin-Layer chromatography was performed on Si250F pre-coated glass plates of silica gel (250 μm). Column chromatography was performed on SilicaFlash P60 silica gel (40-63 μm).

Automated peptide synthesis was carried out using an Argonaut Quest 210 peptide synthesizer (Argonaut Technologies, San Carlos, CA) in 5 mL reaction vessels.

NMR Spectra were obtained using a Bruker AMX-500 NMR spectrometer equipped with a triple resonance (^1H , ^{13}C , ^{15}N) gradient inverse probehead. Spectra were assigned based on 2D COSY and HSQC experiments. For ^1H NMR, minor isomers are listed between square brackets or as H'. A Waters Micromass ZQ 2000 mass spectrometer was used for electrospray ionization (ESI) mass spectrometry measurements.

The matrix assisted laser desorption ionization (MALDI) data was acquired using a prototype quadrupole-quadrupole-TOF (QqTOF) mass spectrometer with photon pulses from a 20-Hz nitrogen laser (VCL 337ND, Spectra-Physics, Mountain-View, CA) with 300 mJ energy/pulse. The sample solution (~0.5 μL) was loaded onto a stainless steel

target with the same volume of the saturated matrix solution (DHB in 1:1 (v/v) acetonitrile-water) and allowed to air dry.

CD measurements were carried out using a Jasco J-810 spectropolarimeter (Jasco Inc., Easton, MD) equipped with a Peltier thermoelectric temperature control system.

Ac-[Pro]₉-NH₂ (1): was synthesized using Fmoc-protected Rink Amide MBHA resin on a 80 μM scale from Fmoc-L-Pro-OH. The synthesis was accomplished using the general procedures for peptide synthesis. Analytical HPLC: t_R : 11.8 min; ¹H NMR (300 MHz, D₂O, 298K) δ = 4.77-4.62 (m, 8H, Pro _{α}), [4.31-4.41, m, 1H, Pro _{α}], 3.73-3.92 (m, 8H, Pro _{δ_1}), 3.43-3.72 (m, 10H, Pro _{δ_2} , Pro _{δ_1'}), 2.19-2.47 (m, 9H, Pro _{β_1}), 1.75-2.15 (m, 30H, Pro _{β_2} , Pro _{$\gamma_{1,2}$} , -COCH₃); MS (ES) calc. for C₄₇H₆₉N₁₀O₁₀ (M + H)⁺: 933.52. Found (M + H)⁺: 933.49; MS (ES) calc. for C₄₇H₆₈N₁₀NaO₁₀ (M + Na)⁺: 955.50. Found (M + Na)⁺: 955.47;

Ac-[Hyp]₉-NH₂ (2): was synthesized using Fmoc-protected Rink Amide MBHA resin on a 100 μM scale from Fmoc-(2*S*,4*R*)-4-*O*-*tert*-butyl-Hyp-OH. The peptide cleavage step from the resin also removes the *O*-*tert*-butyl protecting groups. Analytical HPLC: t_R : 9.1 min; ¹H NMR (500 MHz, D₂O, 298K) δ = 4.84-4.94 (m, 8H, Hyp _{α}), 4.57-4.70 (m, 9H, Hyp _{γ}), [4.49-4.56, m, 1H, Hyp _{α}], 3.86-4.00 (m, 8H, Hyp _{δ_1}), 3.63-3.86 (m, 10H, Hyp _{δ_2} , Hyp _{δ_1'}), 2.30-2.49 (m, 9H, Hyp _{β_1}), 1.94-2.17 (m, 12H, Hyp _{β_2} , -COCH₃); MS (ES) calc. for C₄₇H₆₉N₁₀O₁₉ (M + H)⁺: 1077.47. Found (M + H)⁺: 1077.60;

Ac-[Hyp(β -Gal)]₉-NH₂ (3): was synthesized using Fmoc-protected Rink Amide MBHA resin on a 30 μ M scale from Fmoc-(2*S*,4*R*)-4-*O*-[2,3,4,6-tetra-*O*-acetyl- β -D-galactopyranose]-Hyp-OH (4). Analytical HPLC: t_R : 2.6 min; ¹H NMR (500 MHz, D₂O, 298K) δ = 4.84-4.96 (m, 8H, Hyp $_{\alpha}$), [4.70-4.84, m, 1H, Hyp $_{\alpha}$], [4.59-4.68, m, 2H, Hyp $_{\gamma}$], 4.45-4.58 (m, 16H, H₁, Hyp $_{\gamma}$), 4.07-4.22 (m, 9H, H₅), 3.60-4.00 (m, 54H, H₂, H₃, H₄, H_{6a}, H_{6b}, Hyp $_{\delta 1}$), 3.44-3.59 (m, 9H, Hyp $_{\delta 2}$), 2.50-2.67 (m, 9H, Hyp $_{\beta 1}$), 1.95-2.16 (m, 12H, Hyp $_{\beta 2}$, -COCH₃); MS (MALDI) calc. for C₁₀₁H₁₅₈N₁₀NaO₆₄ (M + Na)⁺: 2557.93. Found (M + H)⁺: 2557.89;

4-*O*-[(2,3,4,6-Tetra-*O*-acetyl)- β -D-galactopyranosyl]-*N* ^{α} -fluoren-9-yl-

methoxycarbonyl-(2*S*,4*R*)-4-hydroxyproline (4). This synthesis was based on the procedure of Arsequell, Sàrries, and Valencia.⁵⁷ Initially, 4-*N* ^{α} -fluoren-9-yl-methoxycarbonyl-(2*S*,4*R*)-4-hydroxyproline (5) (2.00 g, 5.66 mmol, 1.0 eq) was dissolved in dry acetonitrile (30 mL), cooled to 0 °C, and flushed with Ar_(g). This was followed by the addition of 1,2,3,4,6-penta-*O*-acetyl- β -D-galactopyranoside (3.31 g, 8.49 mmol, 1.5 eq) and BF₃·OEt (2.13 mL, 16.9 mmol, 3.0 eq). After stirring for 30 minutes, the mixture was allowed to warm to ambient temperature, and was stirred for an additional 2 hours before being concentrated under reduced pressure. The resulting residue was suspended in water (30 mL) and was extracted into ethyl acetate (3 x 20 mL), then was dried (Na₂SO₄) and concentrated under reduced pressure. The product was purified by flash chromatography using first 10:10:1 DCM/EtOAc/MeOH then 4:4:1 DCM/EtOAc/MeOH to yield (4) as a clear oil (2.61 g, 3.81 mmol) (67.4%): $[\alpha]_D^{25} = -43.3^\circ$ (*c* 1.1 CH₃OH); ¹H NMR (300 MHz, CD₃OD, 298K) δ = 7.73-7.87 (m, 2H,

aromatic), 7.57-7.70 (m, 2H, aromatic), 7.25-7.47 (m, 4H, aromatic), 5.33-5.44 (m, 1H, H₃), 5.03-5.21 (m, 2H, H₂, H₄), 4.71-4.81 (m, 1H, H₁), 4.01-4.55 (m, 8H, H₅, H_{6a}, H_{6b}, Hyp_α, Hyp_γ, -C(O)OCH₂CH, -C(O)OCH₂CH), 3.73-3.85 (m, 1H, Hyp_{δ1}), 3.61-3.73 (m, 1H, Hyp_{δ2}), 2.33-2.54 (m, 1H, Hyp_{β1}), 1.87-2.27 (m, 13H, Hyp_{β2}, -COCH₃); ¹³C NMR (75 MHz, CD₃OD, 298K) δ = (173.8), (173.6), (173.4), (173.0), 172.0, 171.9, 171.5, 171.3, (156.6), 156.3, (145.8), 145.1, 142.6, (142.5), 129.1, (128.9), (128.5), 128.3, (126.4), 126.3, (121.1), 121.0, 102.0, (101.5), 79.6, (79.5), (72.2), 72.1, 70.33, (70.31), (68.97), 68.91, (68.6), 68.2, 63.0, (60.5), 60.4, (60.1), 59.3, (54.3), 53.9, 30.9, (30.1), (25.3), 23.9, 20.8, 20.7, 20.6, 20.3; MS (ES) calc. for C₃₄H₃₇NNaO₁₄ (M + Na)⁺: 706.21. Found (M + Na)⁺: 706.26

7.6.2 Peptide synthesis and purification

Peptide couplings were carried out using Fmoc-Rink Amide MBHA resin, which was initially allowed to swell in DCM for 1 hour, then was washed with DMF (5x). *Fmoc deprotections*: Deprotection of the Fmoc group was carried out using piperidine/DMF (1:4) (2 x 15 min), followed by a washing step with DMF (3x), DCM (3x) and again DMF (3x). *Peptide coupling*: Amino acids (Fmoc-Xaa-OH) were coupled after pre-activation (4 min) in DMF (4 mL) along with TBTU (4 eq) and DIPEA (8 eq). After mixing for 2.5 hours, the resin was washed with DMF (3x), DCM (3x) and again DMF (3x). Coupling and deprotection reactions were monitored qualitatively by the Chloranil test.⁸⁶ *Capping*: After coupling of the first amino acid, the resin was capped using acetic anhydride/pyridine (1:9) (2 x 15 min), followed by a washing step with DMF (3x), DCM (3x) and DMF (3x). The same procedure was used to cap the *N*-terminus of the peptide.

Cleavage: After a washing step of MeOH (3x) and DCM (6x), the peptides were cleaved from the resin by adding TFA/triisopropylsilane/water (18:1:1) (v/v/v) and mixing for 2 hours. The resin was washed with TFA (2 x 5 mL) and DCM (2 x 5 mL) before concentration by co-distillation with toluene (3 x 10 mL). *Acetate deprotection:* The sugar *O*-acetate groups were removed using methanol/hydrazine hydrate (6:1) (v/v) (2 x 3 hours). This deprotection step was carried out after the peptide was cleaved from the resin in order to limit acid-catalyzed de-glycosylation of the β -linked sugars.

Semi-preparative HPLC was performed on a Waters HPLC system (Waters Corp., Milford, MA) using a Vydac C₁₈ 5 μ m particle size (10 mm x 250 mm) column (W. R. Grace & Co., Deerfield, IL). Analytical HPLC was carried out using an Atlantis C₁₈ 5 μ m particle size (4.6 mm x 150 mm) column (Waters Corp., Milford, MA). Detection of the products was achieved using a photodiode array monitoring at $\lambda = 210$ nm. Preparative HPLC was carried out using a gradient of 0% to 40% acetonitrile in water (0.1% TFA) over 30 min at a flow rate of 1 mL/min. Analytical HPLC was carried out using a gradient of 0% to 10% acetonitrile in water (0.1% TFA) over 30 min at a flow rate of 1 mL/min.

7.6.3 Circular dichroism spectroscopy

Purified samples of **1-3** were dried *in vacuo* for 24 hours prior to weighing, then were dissolved in pure water at an approximate concentration of 0.4 mM, and left to equilibrate for at least 24 hours. Relative concentrations were established by absorbance measurements at 210 nm.⁸⁷ The pH of the samples of **1-3** was determined at 25 °C after a two-point calibration (pH 4.0 and 10.0) to be 5.6, 7.5, and 6.1, respectively. For analysis,

300 μl of the sample were loaded in a rectangular quartz cell with a 0.1 cm pathlength. CD spectra were recorded using a scanning rate of 10 nm/min, response time of 4 sec, data pitch of 0.1 nm, and band width of 2 nm. The CD intensity and wavelength of the spectropolarimeter were calibrated using solutions of *d*-10-camphorsulphonic acid.⁸⁸ A blank spectrum of the solvent was subtracted, and for **3** the spectrum of D-galactose (3.6 mM) was subtracted. For thermal melt experiments, a scan from 200 to 250 nm was performed at 5 °C intervals between 5 and 90 °C. Fluctuation in the temperature contributes to an error of ± 1.0 °C to measurement of the temperature. The sample was allowed to equilibrate at each temperature for 10 min before the measurement was taken. The molar ellipticity per mean residue ($[\theta]$) was calculated from the average spectrophotometer output (millidegrees) at 225 nm using the formula:

$$[\theta] = (M \cdot \theta) / (10 \cdot c \cdot l \cdot n)$$

where M is the molecular weight of the sample ($\text{g}\cdot\text{mol}^{-1}$), θ is the CD signal (millidegrees), c is the peptide concentration in $\text{g}\cdot\text{L}^{-1}$, l is the path length of the cell (cm), and n is the number of amide bonds in the peptide (10). The reversibility of the thermal denaturation was judged by the recovery of the signal at the conclusion of the experiment (by 80-89%) at 25 °C. The temperature at the midpoint of the thermal transition (T_m) was estimated by calculating the first-derivative of the molar ellipticity at 225 nm as a function of temperature. Values of T_m were estimated by the minimum of the spline curves of the corresponding data points using SigmaPlot software v11.0 (Systat Software, Inc). Error in T_m can be assumed to be half the value of consecutive data points (± 2.5 °C) along with the instrumental error (± 0.5 °C).

7.7 References

- 1) Cowan, P. M.; McGavin, S. *Nature* **1955**, *176*, 501-503.
- 2) Steinberg, I. Z.; Harrington, W. F.; Berger, A.; Sela, M.; Katchalski, E. *J. Am. Chem. Soc.* **1960**, *82*, 5263-5279.
- 3) Traub, W.; Shmueli, U. *Nature*, **1963**, *198*, 1165-1166.
- 4) Adzhubei, A. A.; Sternberg, M. J. E. *J. Mol. Biol.* **1993**, *229*, 472-493.
- 5) a) Siligardi, G.; Drake, A. F. *Biopolymers* **1995**, *37*, 281-292; b) Kleywegt, G. T.; Jones, T. A. *Structure* **1996**, *4*, 1395-1400; c) Bochicchio, B.; Tamburro, A. M. *Chirality* **2002**, *14*, 782-792; d) Berisio, R.; Loguercio, S.; De Simone, A.; Zagari, A.; Vitagliano, L. *Prot. Pept. Lett.* **2006**, *13*, 847-854.
- 6) Rath, A.; Davidson, A. R.; Deber, C. M. *Biopolymers* **2005**, *80*, 179-185.
- 7) Tiffany, M. L.; Krimm, S. *Biopolymers* **1968**, *6*, 1379-1382.
- 8) Woody, R. W. *Adv. Biophys. Chem.* **1992**, *2*, 37-39.
- 9) Stapley, B. J.; Creamer, T. P. *Prot. Sci.* **1999**, *8*, 587-595.
- 10) Rucker, A. L.; Creamer, T. P. *Protein Sci.* **2002**, *11*, 980-985.
- 11) Shi, Z.; Woody, R. W.; Kallenbach, N. R. *Adv. Protein Chem.* **2002**, *62*, 163-240.
- 12) Hamburger, J. B.; Ferreon, J. C.; Whitten, S. T.; Hilser, V. J. *Biochemistry* **2004**, *43*, 9790-9799.

- 13) Ramakrishnan, V.; Ranbhor, R.; Durani, S. *J. Am. Chem. Soc.* **2004**, *126*, 16332-16333.
- 14) Tran, H. T.; Wang, X.; Pappu, R. V. *Biochemistry* **2005**, *44*, 11369-11380.
- 15) Whittington, S. J.; Chellgren, B. W.; Hermann, V. M.; Creamer, T. P. *Biochemistry* **2005**, *44*, 6269-6275.
- 16) Shi, Z.; Chen, K.; Liu, Z.; Kallenbach, N. R. *Chem. Rev.* **2006**, *106*, 1877-1897.
- 17) a) Gertler, F. B.; Niebuhr, K.; Reinhard, M.; Wehland, J.; Soriano, P. *Cell* **1996**, *87*, 227-239; b) Jardetzky, T. S.; Brown, J. H.; Gorga, J. C.; Stern, L. J.; Urban, R. G.; Strominger, J. L.; Wiley, D. C. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 734-738; c) Mahoney, N. M.; Janmey, P. A.; Almo, S. C. *Nat. Struct. Biol.* **1997**, *4*, 953-960; d) Freund, C.; Kühne, R.; Yang, H.; Park, S.; Reinherz, E. L.; Wagner, G. *Embo. J.* **2002**, *21*, 5985-5995; e) Parrot, I.; Huang, P. C.; Khosla, C. *J. Biol. Chem.* **2002**, *277*, 45572-45578; f) Wiesner, S.; Stier, G.; Sattler, M.; Macias, M. J. *J. Mol. Biol.* **2002**, *324*, 807-822; g) Srinivasan, M.; Lu, D.; Eri, R.; Brand, D. D.; Haque, A.; Blum, J. S. *J. Biol. Chem.* **2005**, *280*, 10149-10155; h) Raghavan, B.; Skoblenick, K. J.; Bhagwanth, S.; Argintaru, N.; Mishra, R. K.; Johnson, R. L. *J. Med. Chem.* **2009**, *52*, 2043-2051.
- 18) Ferreon, J. C.; Hilser, V. J. *Prot. Sci.* **2003**, *12*, 447-457.
- 19) a) Saskisekharan, V. *Acta Cryst.* **1959**, *12*, 903-909; b) Cartier, L.; Lotz, B. *Macromolecules* **1998**, *31*, 3049-3054.

- 20) a) Isemura, T.; Okabayashi, H.; Sakakibara, S. *Biopolymers* **1968**, *6*, 307-321; b) Okabayashi, H.; Isemura, T.; Sakakibara, S. *Biopolymers* **1968**, *6*, 323-330; c) Grathwohl, C.; Wuthrich, K. *Biopolymers* **1981**, *20*, 2623-2633; d) Kakinoki, S.; Hirano, Y.; Oka, M. *Peptide Sci.* **2004**, 453-456; e) Yuge, M.; Kakinoki, S.; Kitamura, M.; Hirano, Y.; Oka, M. *Peptide Sci.* **2005**, 403-406.
- 21) Helbecque, N.; Loucheux-Lefebvre, M. H. *Int. J. Pept. Prot. Res.* **1982**, *19*, 94-101.
- 22) Lam, S. L.; Hsu, V. L. *Biopolymers* **2003**, *69*, 270-281.
- 23) a) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1974**, *7*, 698-705; b) Bour, P.; Kubelka, J.; Keiderling, T. A. *Biopolymers* **2002**, *65*, 45-59.
- 24) a) Creamer, T. P.; Campbell, M. N. *Adv. Protein Chem.* **2002**, *62*, 263-282; b) Chellgren, B. W.; Creamer, T.P. *Biochemistry* **2004**, *43*, 5864-5869.
- 25) Kelly, M. A.; Chellgren, B. W.; Rucker, A. L.; Troutman, J. M.; Fried, M. G.; Miller, A.-F.; Creamer, T. P. *Biochemistry* **2001**, *40*, 14376-14383.
- 26) Rucker, A. L.; Pager, C. T.; Campbell, M. N.; Qualls, J. E.; Creamer, T. P. *Proteins* **2003**, *53*, 68-75.
- 27) Vila, J. A.; Baldoni, H. A.; Ripoll, D. R.; Ghosh, A.; Scheraga, H. A. *Biophys. J.* **2004**, *86*, 731-742.
- 28) Jha, A. K.; Colubri, A.; Zaman, M. H.; Koide, S.; Sosnick, T. R.; Freed, K. F. *Biochemistry* **2005**, *44*, 9691-9702.

- 29) a) Hinderaker, M. P.; Raines, R. T. *Protein Sci.* **2003**, *12*, 1188-1194; b) Jenkins, C. L.; Lin, G.; Duo, J.; Rapolu, D.; Guzei, I. A.; Raines, R. T.; Krow, G. R. *J. Org. Chem.* **2004**, *69*, 8565-8573.
- 30) Nimni, M. E. *Collagen*; CRC Press: Boca Raton, FL, 1988.
- 31) a) Ramachandran, G. N.; Kartha, G. *Nature* **1954**, *174*, 269-270; b) Ramachandran, G. N.; Kartha, G. *Nature* **1955**, *176*, 593-595; c) Rich, A.; Crick, F. H. C. *Nature* **1955**, *176*, 915-916; d) Rich, A.; Crick, F. H. C. *J. Mol. Biol.* **1961**, *3*, 483-506; e) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75-81.
- 32) a) Lamport, D. T. A. *Nature* **1967**, *216*, 1322-1324; b) Lamport, D. T. A.; Miller, D. H.; *Plant Physiol.* **1971**, *48*, 454-456; c) Lamport, D. T. A. *Recent Adv. Phytochem.* **1977**, *11*, 79-115.
- 33) a) Knox, R. B.; Clarke, A.; Harrison, S.; Smith, P.; Marchalonis, J. J. *Proc. Natl. Acad. Sci. U. S. A.* **1976**, *73*, 2788-2792; b) Bowles, D. J. *Annu. Rev. Biochem.* **1990**, *59*, 873-907; c) Showalter, A. M. *Plant Cell* **1993**, *5*, 9-23; d) Sommer-Knudsen, J.; Bacic, A.; Clarke, A. E. *Phytochemistry* **1998**, *47*, 483-497; e) Josè-Estanyol, M.; Puigdomènech, P. *Plant Physiol. Biochem.* **2000**, *38*, 97-108; f) Kieliszewski, M. J. *Phytochemistry* **2001**, *57*, 319-323; g) Khashimova, Z. S. *Chem. Nat. Comput.* **2003**, *39*, 229-236.
- 34) Cannon, M. C.; Terneus, K.; Hall, Q.; Tan, L.; Wang, Y.; Wegenhart, B. L.; Chen, L.; Lamport, D. T. A.; Chen, Y.; Kieliszewski, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 2226-2231.
- 35) Josè, M.; Puigdomènech, P. *New Phytol.* **1993**, *125*, 259-282.

- 36) Shpak, E.; Leykam, J. F.; Kieliszewski, M. J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 14736-14741.
- 37) Shpak, E.; Barbar, E.; Leykam, J. F.; Kieliszewski, M. J. *J. Biol. Chem.* **2001**, *276*, 11272-11278.
- 38) Homer, R. B.; Roberts, K. *Planta* **1979**, *146*, 217-222.
- 39) van Holst, G.-J.; Varner, J. E. *Plant Physiol.* **1984**, *74*, 247-251.
- 40) Ferris, P. J.; Woessner, J. P.; Waffenschmidt, S.; Kilz, S.; Drees, J.; Goodenough, U. W. *Biochemistry* **2001**, *40*, 2978-2987.
- 41) Ferris, P. J.; Waffenschmidt, S.; Umen, J. G.; Ishida, K.; Kubo, T.; Lau, J.; Goodenough, U. W. *Plant Cell* **2005**, *17*, 597-615.
- 42) Zhang, R.; Brownwell, F. E.; Madalengoitia, J. S. *J. Am. Chem. Soc.* **1998**, *120*, 3894-3902.
- 43) Sadler, K.; Eom, K. D.; Yang, J.-L.; Dimitrova, Y.; Tam, J. P. *Biochemistry* **2002**, *41*, 14150-14157.
- 44) a) Crespo, L.; Sanclimens, G.; Montaner, B.; Perez-Tomas, R.; Royo, M.; Pons, M.; Albericio, F.; Giralt, E. *J. Am. Chem. Soc.* **2002**, *124*, 8876-8883; b) Farrera-Sinfreu, J.; Giralt, E.; Castel, S.; Albericio, F.; Royo, M. *J. Am. Chem. Soc.* **2005**, *127*, 9459-9468; c) Geisler, I.; Chmielewski, J. *Bioorg. Med. Chem. Lett.* **2007**, *17*, 2765-2768.

- 45) a) Fernández-Carneado, J.; Kogan, M. J.; Pujals, S.; Giralt, E. *Biopolymers* **2004**, *76*, 196-203; b) Fernández-Carneado, J.; Kogan, M. J.; Castel, S.; Giralt, E. *Angew. Chem. Int. Ed.* **2004**, *43*, 1811-1814.
- 46) Jasin, H. E.; Glynn, L. E. *Immunology* **1965**, *8*, 260-269.
- 47) a) Ungar-Waron, H.; Gurari, D.; Hurwitz, E.; Sela, M. *Eur. J. Immunol.* **1973**, *3*, 201-205; b) Maryanski, J. L.; Verdini, A. S.; Weber, P. C.; Salemme, F. R.; Corradin, G. *Cell* **1990**, *60*, 63-72.
- 48) a) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U. S. A.* **1967**, *58*, 719-726; b) Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B. *J. Am. Chem. Soc.* **2002**, *124*, 13067-13071; c) Arora, P. S.; Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743-747.
- 49) Fillon, Y. A.; Anderson, J. P.; Chmielewski, J. *J. Am. Chem. Soc.* **2005**, *127*, 11798-11803.
- 50) Otvos Jr., L. *Cell. Mol. Life Sci.* **2002**, *59*, 1138-1150.
- 51) a) Wu, C. C.; Komoroski, R. A.; Mandelkern, L. *Macromolecules* **1975**, *8*, 635-637; b) Jacob, J.; Baker, B.; Bryant, R. G.; Cafiso, D. S. *Biophys. J.* **1999**, *77*, 1086-1092; c) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2754-2759; d) Watkins, L. P.; Chang, H.; Yang, H. *J. Phys. Chem. A* **2006**, *110*, 5191-5203; e) Doose, S.; Neuweller, H.; Barsch, H.; Sauer, M. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 17400-17405; f) Sahoo, H.; Roccatano, D.; Hennig, A.; Nau, W. M. *J. Am. Chem. Soc.* **2007**, *129*, 9762-9772.

- 52) a) Väkiparta, M.; Forsback, A.-P.; Lassila, L. V.; Jokinen, M.; Yli-Urpo, A. U. O.; Vallittu, P. K. *J. Mat. Sci.* **2005**, *16*, 873-879; b) Väkiparta, M.; Vallittu, P. K. *Macromol. Symp.* **2007**, *253*, 88-93.
- 53) Horng, J.-C.; Raines, R. T. *Prot. Sci.* **2006**, *15*, 74-83.
- 54) Kūmin, M.; Sonntag, L.-S.; Wennemers, H. *J. Am. Chem. Soc.* **2007**, *129*, 466-467.
- 55) Owens, N. W.; Braun, C.; O'Neil, J. D.; Marat, K.; Schweizer, F. *J. Am. Chem. Soc.* **2007**, *129*, 11670-11671.
- 56) Erwig, E.; Koenigs, W. *Ber. Dtsch. Chem. Ges.* **1889**, *22*, 2207-2213.
- 57) Arsequell, G.; Sàrries, N.; Valencia, G. *Tetrahedron Lett.* **1995**, *36*, 7323-7326.
- 58) Ronish, E. W.; Krimm, S. *Biopolymers* **1974**, *13*, 1635-1651.
- 59) Brahmachari, S. K.; Bansal, M.; Ananthanarayanan, V. S.; Sasisekharan, V. *Macromolecules* **1979**, *12*, 23-28.
- 60) Chen, K.; Liu, Z.; Kallenbach, N. R.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.*, **2004**, *101*, 15352-15357.
- 61) Eker, F.; Griebenow, K.; Schweitzer-Stenner, R. *J. Am. Chem. Soc.* **2003**, *125*, 8178-8185;
- 62) Makarov, A. A.; Adzhubei, I. A.; Protasevich, I. I.; Lobachov, V. M.; Esipova, N. G. *J. Prot. Chem.* **1993**, *12*, 85-91.

- 63) Chen, K.; Liu, Z.; Zhou, C.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2005**, *127*, 10146-10147.
- 64) Clark, D. S.; Mattice, W. L. *Macromolecules* **1977**, *10*, 369-374.
- 65) a) Torchia, D. A. *Macromolecules* **1971**, *4*, 440-442; b) Kapitán, J.; Baumruk, V.; Bour, P. *J. Am. Chem. Soc.* **2006**, *128*, 2438-24; c) Kang, Y. K.; Jhon, J. S.; Park, H. S. *J. Phys. Chem. B* **2006**, *110*, 17645-17655; d) Zhong, H.; Carlson, H. A. *J. Chem. Theory Comput.* **2006**, *2*, 342-353.
- 66) a) Panasik, N., Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262-269; b) Eberhardt, E. S.; Panasik, N., Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 12261-12266.
- 67) Bhatnagar, R. S.; Gough, C. A. in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Ed. Fasman, G. D.), Plenum Press, New York, 1996; p.193.
- 68) a) Ma, K.; Kan, L.-S. Kan, and K. Wang *Biochemistry* **2001**, *40*, 3427-3438; b) Kitamura, M.; Yamauchi, T.; Oka, M.; Hayashi, T. *Polym. Bull.* **2003**, *51*, 143-150.
- 69) Woody, R. W. in *Circular Dichroism and the Conformational Analysis of Biomolecules* (Ed. Fasman, G. D.), Plenum Press, New York, 1996; p.56.
- 70) Pysh, E. S. *Biopolymers* **1974**, *13*, 1563-1571.
- 71) Dyson, H. J.; Wright, P. E. *Annu. Rev. Biophys. Chem.* **1991**, *20*, 519-538.
- 72) a) Greenfield, N. J.; Fasman, G. D. *Biochemistry* **1969**, *8*, 4108-4116; b) Sreerama, N.; Woody, R. W. *Prot. Struct. Funct. Genet.* **1999**, *36*, 400-406.

- 73) Shi, Z.; Olson, C. A.; Rose, G. D.; Baldwin, R. L.; Kallenbach, N. R. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 9190-9195.
- 74) Ding, L.; Chen, K.; Santini, P. A.; Shi, Z.; Kallenbach, N. R. *J. Am. Chem. Soc.* **2003**, *125*, 8092-8093.
- 75) Avbelj, F.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U. S. A.*, **2003**, *100*, 5742-5747.
- 76) Eker, F.; Griebenow, K.; Cao, X.; Nafle, L. A.; Schweitzer-Stenner, R. *Proc. Natl. Acad. Sci. U. S. A.*, **2004**, *101*, 10054-10059.
- 77) Ronish, E. W.; Krimm, S. *Biopolymers* **1972**, *11*, 1919-1928.
- 78) a) Drake, A. F.; Siligardi, G.; Gibbons, W. A. *Biophys. Chem.* **1988**, *31*, 143-146; b) Kitamura, M.; Kakinoki, S.; Hirano, Y.; Oka, M. *Polym. Bull.* **2005**, *54*, 303-310.
- 79) a) Tiffany, M. L.; Krimm, S. *Biopolymers* **1968**, *6*, 1767-1770; b) Pettitt, B. M.; Karplus, M. *Chem. Phys. Lett.* **1985**, *121*, 194-201; c) Han, W.-G.; Jalkanen, K. J.; Elstner, M.; Suhai, S. *J. Phys. Chem. B* **1998**, *102*, 2587-2602; d) Mezei, M.; Fleming, P. J.; Srinivasan, R.; Rose, G. D. *Proteins* **2004**, *55*, 502-507.
- 80) a) Imperiali, B. *Acc. Chem. Res.* **1997**, *30*, 452-459; b) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1-27; c) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R-56R; d) Paul, G.; Lottspeich, F.; Wielland, F. *J. Biol. Chem.* **1986**, *261*, 1020-1024; e) Kern, G.; Kern, D.; Jaenicke, R.; Seckler, R. *Protein Sci.* **1993**, *2*, 1862-1868.
- 81) a) Varki, A. *Glycobiology* **1993**, *3*, 97-130; b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.

- 82) a) Live, D. H.; Williams, H. J.; Kuduk, S. D.; Schwarz, J. B.; Glunz, P. W.; Chen, X.-T.; Sames, D.; Kumar, R. A.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 3489-3493; b) Coltart, D. M.; Royyuru, A. K.; Williams, L. J.; Glunz, P. W.; Sames, D.; Kuduk, S. D.; Schwarz, J. B.; Chen, X.-T.; Danishefsky, S. J.; Live, D. H. *J. Am. Chem. Soc.* **2002**, *124*, 9833-9844; c) Bosques, C. J.; Tschampel, S. M.; Woods, R. J.; Imperiali, B. *J. Am. Chem. Soc.* **2004**, *126*, 8421-8425.
- 83) Norgren, A. S.; Arvidsson, P. I. *J. Org. Chem.* **2008**, *73*, 5272-5278.
- 84) Simpson, G. L.; Gordon, A. H.; Lindsay, D. M.; Promsawan, N.; Crump, M. P.; Mulholland, K.; Hayter, B. R.; Gallagher, T. *J. Am. Chem. Soc.* **2006**, *128*, 10638-10639.
- 85) Norgren, A. S.; Geitmann, M.; Danielson, U. H.; Arvidsson, P. I. *J. Mol. Recognit.* **2007**, *20*, 132-138.
- 86) Vojkovsky, T. *Pept. Res.* **1995**, *8*, 236-237.
- 87) Bretscher, L. E.; Jenkins, C. L.; Taylor, K. M.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777-778.
- 88) Wu, C. S.; Chen, G. C. *Anal. Biochem.* **1989**, *177*, 178-182.
- 89) Wu, J.; Yang, J.-T.; Wu, C.-S. *C. Anal. Biochem.* **1992**, *200*, 359-364.
- 90) Sreerama, N.; Woody, R. W. *Prot. Sci.* **2003**, *12*, 384-388.

Chapter 8: Conclusions and Future Work

This work provides the first in depth analysis of the influence of a proline-linked carbohydrate residue on prolyl amide isomerization, the properties of the proline side chain, and the conformational stability of proline-containing peptides.

The synthesis of a *C*-glucosyl proline hybrid (GlcProH) is the first example of a proline-based *C*-glycosyl amino acid, and the use of a carbohydrate as a scaffold to fix the conformation of a proline side chain (Chapter 3). Recently, several other examples have emerged.¹ In model peptides, the GlcProH demonstrated the ability to vary the *N*-terminal amide equilibrium ($K_{trans/cis}$) by modification of the substituents on the sugar scaffold. Until now, the ability to vary $K_{trans/cis}$ required the synthesis of proline surrogates by independent synthetic routes. However, our results demonstrate that one-step modification of the hydroxyl groups on the carbohydrate scaffold can be used to tune the conformational properties of GlcProH-containing model peptides. As opposed to many δ -modified proline analogues (Sections 1.6.2 and 1.6.3), the origin of the influence on $K_{trans/cis}$ is more than a sterically based, and likely involves intramolecular interactions between the sugar substituents and the peptide backbone.

Limitations of the GlcProH building block include: a lengthy synthetic route, the loss of the hydrophobic character of the prolyl side chain, and the crowded steric environment around the prolyl nitrogen atom, which limits coupling to amino acids with sterically bulky side chains. While this proline analogue may not be appropriate for studying binding interactions involving the prolyl side chain, the ability to vary $K_{trans/cis}$ may find use for studying reverse turn formation, where the multiple hydroxyl groups can be further devitalized in a combinatorial fashion. Future work should include further

variation of substituents on the carbohydrate scaffold in order to vary $K_{trans/cis}$, and gain a better understanding of the origin of the factors responsible for such changes. For example, attachment of mono-, di- and trifluoroacetate groups may confirm the proposed $n \rightarrow \pi^*$ intramolecular interaction. Developing a range of biases between all *trans* to all *cis* amide content using different functional groups on the sugar would be very useful.

The effects of *O*-galactosylation of 4*R*-hydroxy-L-proline (Hyp) in model amides of the form Ac-Hyp(α -D-Gal)-NHMe and Ac-Hyp(β -D-Gal)-NHMe on the conformation, kinetics, and thermodynamics of amide isomerization were studied in detail (Chapter 4). It was found that there was no apparent effect on $K_{trans/cis}$ or on the rate of isomerization (k_{tc} and k_{ct}) when compared to unglycosylated reference compounds. These were reflected in both the ground state enthalpy and entropy differences, and activation energies as calculated from van't Hoff and Eyring plots, respectively. These results suggest the *O*-glycosylation of Hyp is a poor method for influencing the prolyl *N*-terminal amide equilibrium and rate of isomerization. However, these results also have implications for the naturally-occurring Hyp *O*-galactosylation found in hydroxyproline-rich glycoproteins (HRGPs) (Section 1.8.1); if glycosylation affects the stability of HRGPs, it is not through an apparent influence of the isomer equilibrium or the rate of amide isomerization.

However, *O*-glycosylation of Hyp provides an inductive electron withdrawing effect on the prolyl ring according to a change in $^{13}\text{C}^\gamma$ -NMR chemical shift, which may lead to additional stabilization of the C^γ -exo ring pucker. Also, nOe experiments show that α -galactosylation results in distant contacts between the proline and galactose rings, suggesting that galactosylation of Hyp induces a conformational constraint into

glycopeptides. These results were indications that influences from Hyp *O*-glycosylation may become more apparent in longer peptides. Further work could include investigating the effects of different monosaccharides, such as D-galactosamine or D-arabinose, and larger di- or trisaccharide glycans *O*-linked to Hyp on amide isomerization. Also, the *N*-terminal amino acid could be varied by amino acids such as L-phenylalanine or L-serine, to determine if the *N*-terminal amino acid can influence either parameter.

The stereoisomer of Hyp, (2*S*,4*S*)-4-hydroxyproline (hyp), is rarely found in nature,² and has a different influence on the conformation of the pyrrolidine ring as well as on $K_{trans/cis}$ (Section 1.6.4); it provided an opportunity to study the effect of *O*-linked glycans projected from the opposite face of the prolyl side chain relative to Hyp on the thermodynamics and kinetics of amide isomerization (Chapter 5).

It was found that in a similar fashion to 4*R*-*O*-galactosylation, 4*S*-*O*-galactosylation of hyp in model peptides of the form Ac-hyp(α -D-Gal)-OMe and Ac-hyp(β -D-Gal)-OMe caused a local electron-withdrawing effect, which likely explains why 4*S*-*O*-glycosylation did not significantly alter the C ^{γ} -endo prolyl pucker. While the glycosylation of 4*R*-hydroxyproline was not found to affect amide isomerization, the glycosylation of 4*S*-hydroxyproline affected both the prolyl *N*-terminal amide equilibrium and the rate of amide isomerization.

A change in $K_{trans/cis}$ observed only for the 4*S*-stereoisomer may be due to reduced electrostatic repulsion between the 4-oxygen atom and the *C*-terminal carbonyl group, or the loss of an intramolecular hydrogen bond that is specific to the 4*S*-stereoisomer. Interestingly, the enthalpic contributions to ΔG° seem to be offset by entropic changes, which means a more ordered environment for the *trans* amide isomer in the 4*S*-*O*-

glycosylated model peptides, resulting in only a small net change in $K_{trans/cis}$. The glycosylation of 4*S*-hydroxyproline does seem to cause a small increase in the rate of amide isomerization, but it is not reflected in the amide I vibrational mode (ν_{amide}) and thus the bond order of the *N*-terminal amide group. Therefore, the cause of the increase in the rate remains unclear. Insight could come from measurement of ν_{amide} in a non-protic solvent, which may reveal any differences in hydrogen bonding interactions.

Furthermore, since both α - and β -anomers *O*-linked to Hyp and hyp are accessible in high yields, and the glycans can be rigidly projected in opposite spatial orientations. An interesting study would be the incorporation of glycosylated hyp and Hyp building blocks into a peptide to study the effect of the orientation of the glycan on a carbohydrate-lectin binding interaction. Similar to 4*R*-hydroxyproline, variation of the glycan and *N*-terminal amino acid could also be undertaken to further understand the influence of the glycan projected from the endo face of the prolyl side chain on *N*-terminal amide isomerization. Analysis of these model amides provided an important baseline for studying longer model peptides.

The effects of Hyp *O*-glycosylation on the stability of collagen, an extremely important structural protein with an ongoing debate as to the role of Hyp in its conformational stability, was carried out by synthesizing collagen peptides using a fragment condensation approach (Chapter 6). This represents one of the first attempts to studying both the effects of glycosylation,³ as well as *O*-modification of Hyp on the stability of collagen,⁴ and the first to explore Hyp *O*-glycosylation on the stability of collagen.

Based on the CD spectra of the model peptides, glycosylation of 4*R*-hydroxy-L-proline in collagen model peptides does not preclude triple helix formation; this is despite a crowded steric environment in the triple helix, which has been attributed to lowering the stability of *O*-acyl collagen peptides.⁴ The knowledge that Hyp *O*-glycosylation did not affect $K_{trans/cis}$ or the C^γ -exo pucker in model peptides was important for understanding the effect of glycosylation on the triple helix. Interestingly, the anomeric linkage of D-galactose to Hyp in the collagen model peptides had different effects on the T_m of the triple helix, which may be due to different intra- and intermolecular interactions with the collagen peptide backbone or with surrounding solvent molecules. This represents an important distinction, and provides additional insight into the role of the anomeric linkage in nature (Section 1.2.4).

An increase in the enthalpy (ΔH) and entropy differences (ΔS) measured between the folded and unfolded states likely reflects an increase in hydration from the multiple sugar groups, but this does not lead to an overall increase in conformational stability of collagen. The fact that amplification of Hyp-mediated hydrogen bonds to a hydration network surrounding the collagen triple helix has little contribution to the overall conformational stability of collagen is an interesting contribution to the debate of the role of Hyp in the stability of collagen (Section 1.7.1). More accurate measurements of the ΔH and ΔS between the folded and unfolded states could be obtained by analysis of the collagen model peptides using differential scanning calorimetry.

Understanding the orientation of the glycan with respect to the triple helix could be accomplished through further computational and experimental studies. Computational calculations using the crystal structure of collagen model peptides in the protein data

bank may give an approximation of the different interactions caused by the α - and β -anomeric linkages. Modification of the glycan with diagnostic functional groups would facilitate analysis of through-space interactions. For example, the *N*-acetyl group of GalNAc has been useful for understanding the orientation of monosaccharides with respect to a polypeptide backbone using NOESY NMR experiments.⁵ Also, the influence of different or larger glycans could be explored; however indications from previous work are that the thermodynamic stabilization of a protein is more strongly correlated to the number of glycans attached to the protein rather than the size of the glycan.⁶

The development of a *C*-glycosyl hydroxyproline building block could have interesting effects on the stability of the collagen triple helix. 4*S*-Methylproline analogues are known to place the 4-methyl group in an equatorial orientation while maintaining a C^{γ} -exo pucker and favouring the *trans* amide isomer (Section 1.7.5); this may place the glycan in such a manner as to have a different impact for interactions with the peptide backbone, and would have high resistance to chemical and biological degradation. To date, no such *C*-glycosyl proline analogue exists.

Finally, the study of the effects of Hyp *O*-galactosylation in contiguous polyproline peptides (Chapter 7) has implications for the functional consequences of glycosylation of the naturally-occurring HRGPs (Section 1.8.1). This is the first quantitative study of the effects of glycosylation of HRGPs in well-defined peptides.

A decrease in the relative intensity of both the positive and negative maxima in CD spectra may be attributed to a distortion of the PPII conformation, but the peptides are still considered to have significant PPII character and are consistent with naturally occurring HRGPs.⁷ Based on the increase in the ρ value, which has been correlated to

decreased solvation of the peptide backbone carbonyl groups,⁸ glycosylation may shield the peptide backbone carbonyl groups from solvent water molecules; this may account for the changes in the CD bands due to changes in the absorption of the amide groups. Therefore, it may be misleading to interpret the decrease in CD maxima as destabilization of the PPII conformation.

The presence of an isodichroic point indicates that the model polyproline peptides underwent a conformational transition from a PPII to a β -strand conformation. Contiguous *O*-glycosylation of Hyp appears to significantly stabilize the PPII conformation since the thermal transition occurred at a much higher temperature relative to the unmodified peptides; this is likely due to the multiple sugar groups forming stabilizing hydrogen bonds to the peptide backbone in the PPII conformation. Additionally, the study of model amides indicated that Hyp *O*-glycosylation does not lead to a measurable increase in the stabilization of the *trans* amide conformation, but this may require longer peptides to become apparent. Comparison of the polyproline peptides in different solvents, such as dimethyl sulfoxide, may give insight into intramolecular hydrogen bonding interactions. In a similar fashion to the collagen model peptides, similar computational and experimental studies will facilitate understanding the orientation of the glycan with respect to the PPII helix, and any intramolecular interactions present.

Whether non-contiguous glycosylation or larger glycan structures will mediate this conformational stabilization remains to be explored. The galactose residue may play some role in changing the PPII conformation as contiguous *O*-arabinylation of extensin model peptides has been found to increase the intensity of CD maxima.⁹ Therefore, it

seems that either galactose residues inherently alter the PPII conformation, or the glycan size is important.

Since glycosylation leads to considerable conformational stability, these peptides may lead to the development of scaffolds and biomaterials that have improved conformational stability. The glycans also provide the opportunity for further derivatization, which can lead to facile modification of the scaffold. Furthermore, the use of polyproline as a well-defined secondary structure scaffold for attaching glycans may find use for the study of carbohydrate-carbohydrate interactions and carbohydrate-lectin binding interactions.

8.1 References

- 1) a) Zhang, K.; Schweizer, F. *Carb. Res.* **2009**, *344*, 576-585; b) Zhang, K.; Teklebrhan, R.; Schreckenbach, G.; Wetmore, S.; Schweizer, F. *J. Org. Chem.* **2009**, *74*, 3735-3743.
- 2) Mauger, A. B. *J. Nat. Prod.* **1996**, *59*, 1205-1211.
- 3) a) Bann, J. G.; Peyton, D. H.; Bächinger, H. P. *FEBS Lett.* **2000**, *473*, 237-240; b) Bann, J. G.; Bächinger, H. P.; Peyton, D. H. *Biochemistry* **2003**, *42*, 4042-4048.
- 4) Jenkins, C. L.; McCloskey, A. I.; Guzei, I. A.; Eberhardt, E.; Raines, R. T. *Biopolymers* **2005**, *80*, 1-8.
- 5) a) Imperiali, B.; Rickert, K. W. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 97-101; b) Coltart, D. M.; Royyuru, A. K.; Williams, L. J.; Glunz, P. W.; Sames, D.; Kuduk, S. D.;

Schwarz, J. B.; Chen, X. T.; Danishefsky, S. J.; Live, D. H. *J. Am. Chem. Soc.* **2002**, *124*, 9833-9842.

6) Shental-Bechor, D.; Levy, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 8256-8261.

7) Ferris, P. J.; Woessner, J. P.; Waffenschmidt, S.; Kilz, S.; Drees, J.; Goodenough, U. W. *Biochemistry* **2001**, *40*, 2978-2987.

8) Pysh, E. S. *Biopolymers* **1974**, *13*, 1563-1571.

9) Shpak, E.; Barbar, E.; Leykam, J. F.; Kieliszewski, M. J. *J. Biol. Chem.* **2001**, *276*, 11272-11278.

**Chapter 9: Tuning of the Prolyl *Trans/Cis* Amide Rotamer Population
Using C-Glucosyl Proline Hybrids**

Supporting Information for Chapter 3

Table of Contents

NMR assignments	305-310
Tables	311-314
NMR spectra	315-368

(11): ^1H NMR (500 MHz, CD_3OD , 298K, 0.036 M) $\delta = 7.12\text{-}7.37$ (m, 15H, aromatic), 4.91 (d, 1H, $-\text{OCH}_2\text{Ph}$, $J = 11.0\text{Hz}$), 4.76 (d, 1H, $-\text{OCH}_2\text{Ph}$, $J = 10.3\text{Hz}$), 4.73 (ddd, 0.9H, H_9 , $J_{3a,9} = 11.7\text{Hz}$, $J_{3b,9} = 7.3\text{Hz}$, $J_{8,9} = 7.1\text{Hz}$), 4.47-4.64 (m, 4H, $-\text{OCH}_2\text{Ph}$), [4.40-4.47, m, 0.1H, H_9], 4.25 (dd apparent d, 1H, H_2 , $J_{2,3a} = 9.7\text{Hz}$, $J_{2,3b} = 1.0\text{Hz}$), 4.07 (dd, 1H, H_8 , $J_{7,8} = 9.3\text{Hz}$, H_8 minor), [3.90-3.95, m, 0.1H, H_7], 3.62-3.80 (m, 3.9H, H_5 , H_7 , H_{10a} , H_{10b} , H_5 minor, H_{10a} minor, H_{10b} minor), 3.59 (dd apparent t, 0.9H, H_6 , $J_{6,7} = 9.3\text{Hz}$, $J_{5,6} = 9.3\text{Hz}$), [3.43-3.46, m, 0.1H, H_6], [2.73, s, 0.3H, $-\text{NHCH}_3$], 2.70 (s, 2.7H, $-\text{NHCH}_3$), 2.60 (ddd, 1H, H_{3a} , $J_{3a,3b} = 12.3\text{Hz}$, H_{3a} minor), 2.07 (s, 2.7H, $-\text{COCH}_3$), 1.76 (ddd, 1H, H_{3b} , H_{3b} minor), [1.72, s, 0.3H, $-\text{COCH}_3$]

(11): ^1H NMR (500 MHz, $\text{DMSO-}D_6$, 298K, 0.036 M) $\delta = [8.12$, q, 0.13H, $-\text{NHCH}_3$, $J = 4.1\text{Hz}$], 7.75 (q, 0.87H, $-\text{NHCH}_3$, $J = 4.1\text{Hz}$), 7.11-7.34 (m, 15H, aromatic), 4.81 (d, 0.87H, $-\text{OCH}_2\text{Ph}$, $J = 10.9\text{Hz}$), 4.68 (d, 0.87H, $-\text{OCH}_2\text{Ph}$, $J = 10.9\text{Hz}$), 4.38-4.62 (m, 5.13H, H_9 , $-\text{OCH}_2\text{Ph}$, $-\text{OCH}_2\text{Ph}$ minor), [4.39-4.42, m, 0.13H, H_2], [4.30-4.38, m, 0.26H, H_9 , H_7], 4.17 (dd, 0.87H, H_2 , $J_{2,3a} = 9.8\text{Hz}$, $J_{2,3b} = 1.0\text{Hz}$), [4.08-4.12, m, 0.13H, H_8], 3.99 (dd, 0.87H, H_8 , $J_{8,9} = 7.2\text{Hz}$, $J_{7,8} = 9.4\text{Hz}$), [3.81-3.85, m, 0.13H, H_5], 3.73-3.79 (m, 0.87H, H_5), 3.70 (dd apparent t, 0.87H, H_7 , $J_{6,7} = 9.8\text{Hz}$), 3.63 (dd, 1H, H_{10a} , $J_{5,10a} = 4.5\text{Hz}$, $J_{10a,10b} = 10.7\text{Hz}$, H_{10a} minor), 3.57 (dd, 0.87H, H_{10b} , $J_{5,10b} = 1.5\text{Hz}$), [3.52, dd, 0.13H, H_{10a} , $J_{5,10a} = 3.8\text{Hz}$, $J_{10a,10b} = 10.5\text{Hz}$], 3.46 (dd apparent t, 0.87H, H_6 , $J_{5,6} = 9.4\text{Hz}$), [3.37-3.41, m, 0.13H, H_6], 2.53-2.63 (m, 1.39H, H_{3a} , H_{3a} minor, $-\text{NHCH}_3$ minor), 2.49 (d, 2.61H, $-\text{NHCH}_3$), 1.89 (s, 2.61H, $-\text{COCH}_3$), [1.73-1.81, m, 0.13H, H_{3b}], [1.61, s, 0.39H, $-\text{COCH}_3$], 1.54 (ddd, 0.87H, H_{3b} , $J_{3b,9} = 7.5\text{Hz}$, $J_{3a,3b} = 12.8\text{Hz}$)

(12): ^1H NMR (500 MHz, CD_3OD , 298K, 0.035 M) δ = 4.74 (ddd, 0.78H, H_9 , $J_{3b,9} = 7.2\text{Hz}$, $J_{3a,9} = 12.0\text{Hz}$, $J_{8,9} = 7.2\text{Hz}$), [4.65, ddd, 0.22H, H_9 , $J_{8,9} = 7.1\text{Hz}$, $J_{3b,9} = 7.3\text{Hz}$, $J_{3a,9} = 10.9\text{Hz}$], [4.55, dd, 0.22H, H_2 , $J_{2,3a} = 9.0\text{Hz}$, $J_{2,3b} = 1.6\text{Hz}$], 4.37 (dd apparent d, 0.78H, H_2 , $J_{2,3a} = 9.8\text{Hz}$, $J_{2,3b} = 1.0\text{Hz}$), [4.11, dd apparent t, 0.22H, H_8 , $J_{7,8} = 8.4\text{Hz}$], 3.95 (dd, 0.78H, H_8 , $J_{7,8} = 9.0\text{Hz}$), 3.80 (dd, 0.78H, H_{10a} , $J_{5,10a} = 2.4\text{Hz}$, $J_{10a,10b} = 12.0\text{Hz}$), [3.77-3.82, m, 0.22H, H_{10a}], [3.67, dd, 0.22H, H_{10b} , $J_{5,10b} = 5.6\text{Hz}$, $J_{10a,10b} = 11.8\text{Hz}$], 3.67 (dd, 1H, H_{10b} , $J_{5,10b} = 5.4\text{Hz}$), 3.62 [dd, 0.22H, H_7 , $J_{6,7} = 9.2\text{Hz}$], 3.48-3.55 (m, 0.78H, H_5), 3.51 (dd, 0.78H, H_7 , $J_{6,7} = 9.6\text{Hz}$), [3.43-3.48, m, 0.22H, H_5], [3.39, dd apparent t, 0.22H, H_6 , $J_{5,6} = 9.8\text{Hz}$], 3.30 (dd, 1H, H_6 , $J_{5,6} = 9.4\text{Hz}$), [2.76, s, 0.66H, $-\text{NHCH}_3$], 2.71 (s, 2.34H, $-\text{NHCH}_3$), [2.64-2.69, m, 0.22H, H_{3a}], 2.60 (ddd, 78H, H_{3a} , $J_{3a,3b} = 12.2\text{Hz}$), 2.27 (s, 2.34H, $-\text{COCH}_3$), [1.95, s, 0.66H, $-\text{COCH}_3$], [1.92-1.98, m, 0.22H, H_{3b}], 1.77 (ddd, 0.78H, H_{3b})

(12): ^1H NMR (500 MHz, DMSO-D_6 , 298K, 0.035 M) δ = [8.16, q, 0.2H, $-\text{NHCH}_3$, $J = 4.5\text{Hz}$], 7.71 (q, 0.8H, $-\text{NHCH}_3$, $J = 4.3\text{Hz}$), 4.52 (ddd, 0.8H, H_9 , $J_{3b,9} = 7.4\text{Hz}$, $J_{3a,9} = 12.0\text{Hz}$, $J_{8,9} = 7.0\text{Hz}$), [4.45, ddd, 0.2H, H_9 , $J_{3b,9} = 7.1\text{Hz}$, $J_{3a,9} = 10.9\text{Hz}$, $J_{8,9} = 7.2\text{Hz}$], [4.42, dd, 0.2H, H_2 , $J_{2,3a} = 9.0\text{Hz}$, $J_{2,3b} = 1.7\text{Hz}$], 4.11 (dd apparent d, 0.8H, H_2 , $J_{2,3a} = 9.7\text{Hz}$, $J_{2,3b} = 1.0\text{Hz}$), [3.85, dd apparent t, 0.2H, H_8 , $J_{7,8} = 8.3\text{Hz}$], 3.70 (dd, 0.8H, H_8 , $J_{7,8} = 9.2\text{Hz}$), 3.57 (dd, 0.8H, H_{10a} , $J_{5,10a} = 1.9\text{Hz}$, $J_{10a,10b} = 11.6\text{Hz}$), [3.54-3.60, m, 0.2H, H_{10a}], 3.39-3.46 (m, 1.2H, H_{10b} , $J_{5,10b} = 5.2\text{Hz}$, H_7 minor, H_{10b} minor), 3.30-3.36 (m, 0.8H, H_5), 3.30 (dd apparent t, 0.8H, H_7 , $J_{6,7} = 9.4\text{Hz}$), [3.23-3.28, m, 0.2H, H_5], [3.13, dd apparent t, 0.2H, H_6 , $J_{6,7} = 9.4\text{Hz}$, $J_{5,6} = 9.4\text{Hz}$], 3.06 (dd apparent t, 0.8H, H_6 , $J_{5,6} = 9.5\text{Hz}$), [2.57, d, 0.6H, $-\text{NHCH}_3$], 2.48 (d, 2.4H, $-\text{NHCH}_3$), 2.43 (ddd, 1H, H_{3a} , $J_{3a,3b} =$

12.3Hz, H_{3a} minor), 2.06 (s, 2.4H, -COCH₃), [1.79, s, 0.6H, -COCH₃], [1.72, ddd, 0.2H, H_{3b}, J_{3a,3b} = 12.5Hz], 1.47 (ddd, 0.8H, H_{3b})

(13): ¹H NMR (500 MHz, CD₃OD, 298K, 0.036 M) δ = [5.31, dd apparent t, 0.36H, H₇, J_{7,8} = 9.3Hz, J_{6,7} = 9.4Hz], 5.26 (dd apparent t, 0.64H, H₇, J_{7,8} = 9.6Hz, J_{6,7} = 9.8Hz), 5.00 (dd apparent t, 0.64H, H₆, J_{5,6} = 9.8Hz), 4.98 [dd, 0.36H, H₆], 4.81-4.89 (m, 0.64H, H₉), [4.69, ddd, 0.36H, H₉, J_{3a,9} = 10.9Hz, J_{3b,9} = 6.8Hz, J_{8,9} = 7.2Hz], [4.52, dd, 0.36H, H₂, J_{2,3a} = 9.3Hz, J_{2,3b} = 1.9Hz], [4.44, dd apparent t, 0.36H, H₈], 4.38 (dd, 0.64H, H₈, J_{8,9} = 7.1Hz), 4.36 (dd apparent d, 0.64H, H₂, J_{2,3a} = 9.7Hz, J_{2,3b} = 1.3Hz), 4.26-4.33 (m, 1H, H_{10a}, H_{10a} minor), 3.99-4.12 (m, 2H, H₅, H_{10b}, H₅ minor, H_{10b} minor), [2.87, ddd, 0.64H, H_{3a}, J_{3a,3b} = 12.9Hz], 2.76 [d, 1.1H, -NHCH₃, J = 4.3Hz], 2.72-2.80 (m, 0.64H, H_{3a}), 2.71 (d, 1.9H, -NHCH₃, J = 4.3Hz), 2.11 (s, 1.9H, -COCH₃), 2.06 (s, 1.9H, -COCH₃), 2.03 [s, 1.1H, -COCH₃], 2.03 (s, 1.9H, -COCH₃), 2.00 (s, 1.9H, -COCH₃), 1.97 [s, 1.1H, -COCH₃], [1.93-1.99, m, 0.36H, H_{3b}], 1.85 (ddd, 0.64H, H_{3b}, J_{3b,9} = 7.4Hz, J_{3a,3b} = 13.1Hz), 1.81 [s, 1.1H, -COCH₃]

(13): ¹H NMR (500MHz, DMSO-D₆, 298K, 0.036 M) δ = [8.11, q, 0.33H, -NHCH₃], 7.82 (q, 0.67H, -NHCH₃), [5.21, dd apparent t, 0.33H, H₇, J_{7,8} = 8.5Hz, J_{6,7} = 8.9Hz], 5.13 (dd apparent t, 0.67H, H₇, J_{7,8} = 9.6Hz, J_{6,7} = 9.6Hz), 4.82 (dd apparent t, 0.67H, H₆, J_{5,6} = 9.8Hz), [4.78, dd apparent t, 0.33H, H₆, J_{5,6} = 9.4Hz], 4.66 (ddd, 0.67H, H₉, J_{3a,9} = 11.5Hz, J_{3b,9} = 7.1Hz, J_{8,9} = 7.1Hz), [4.53, ddd, 0.33H, H₉, J_{3a,9} = 10.8Hz, J_{3b,9} = 6.6Hz, J_{8,9} = 6.6Hz], 4.36 (dd, 0.67H, H₈), [4.32, dd apparent d, 0.33H, H₂, J_{2,3a} = 9.6Hz, J_{2,3b} = 1.4Hz], [4.22, dd apparent t, 0.33H, H₈], 4.09-4.19 (m, 2H, H₂, H₅, H_{10a}), [4.05, m, 0.33H, H₅], 3.87-3.96 (m, 1H, H_{10b}, H_{10b} minor), [2.84, ddd, 0.33H, H_{3a}, J_{3a,3b} = 12.1Hz],

2.77 (ddd, 0.67H, H_{3a}, J_{2,3a} = 10.0Hz, J_{3a,3b} = 12.8Hz), [2.57, d, 1H, -NHCH₃, J = 4.4Hz], 2.49 (d, 2H, -NHCH₃, J = 4.6Hz), [2.07, s, 1H, -COCH₃], [2.04, s, 1H, -COCH₃], 1.98 (s, 2H, -COCH₃), 1.95 (s, 2H, -COCH₃), 1.93 (s, 2H, -COCH₃), 1.91 (s, 2H, -COCH₃), [1.85, s, 1H, -COCH₃], [1.73, ddd, 0.33H, H_{3b}], [1.61, s, 1H, -COCH₃], 1.57 (ddd, 0.67H, H_{3b}, J_{2,3b} = 1.0Hz)

(14): ¹H NMR (500 MHz, CD₃OD, 298K, 0.033 M) δ = 7.85 (q, 1H, -NHCH₃), 7.71 (t, 1H, -NH_(Gly)), 7.12-7.40 (m, 15H, aromatic), 4.91 (d, 1H, -OCH₂Ph, J = 11.3Hz), 4.47-4.77 (m, 6H, -OCH₂Ph, H₉), 4.40 (dd, 1H, H_{α1(Gly)}, J_{Hα1(Gly),NH} = 4.1Hz, J_{Hα1(Gly),Hα2(Gly)} = 17.5Hz), 4.19 (dd apparent d, 1H, H₂, J_{2,3a} = 10.0Hz, J_{2,3b} = 1.0Hz), 4.10 (dd, 1H, H₈, J_{8,9} = 7.0Hz, J_{7,8} = 9.0Hz), 4.03 (dd, 1H, H_{α2(Gly)}, J_{Hα2(Gly),NH} = 3.4Hz), 3.56-3.80 (m, 5H, H₅, H₆, H₇, H_{10a}, H_{10b}), 2.68 (d, 2.7H, -NHCH₃), 2.55 (ddd, 1H, H_{3a}, J_{3a,9} = 12.0Hz, J_{3a,3b} = 13.0Hz), 1.92 (s, 3H, -COCH₃), 1.76 (ddd, 1H, H_{3b}, J_{3b,9} = 7.9Hz)

(15): ¹H NMR (500 MHz, CD₃OD, 298K, 0.035 M) δ = 4.72 (ddd, 0.85H, H₉, J_{3b,9} = 7.2Hz, J_{3a,9} = 12.3Hz, J_{8,9} = 7.1Hz), [4.60-4.68, m, 0.15H, H₉], 4.62 (d, 0.85H, H_{α1(Gly)}, J_{Hα1(Gly),Hα2(Gly)} = 17.1Hz), 4.41 (dd apparent d, 0.85H, H₂, J_{2,3a} = 9.9Hz, J_{2,3b} = 1.0Hz), [4.21, dd, 0.15H, H₂, J_{2,3a} = 5.3Hz, J_{2,3b} = 2.8Hz], [4.12, dd apparent t, 0.15H, H₈, J_{8,9} = 7.1Hz, J_{7,8} = 8.3Hz], 4.07 (d, 0.85H, H_{α2(Gly)}), [3.97-4.03, m, 0.15H, H_{α1(Gly)}], 4.00 (dd, 0.85H, H₈, J_{7,8} = 9.0Hz), 3.80 (dd, 0.85H, H_{10a}, J_{5,10a} = 2.0Hz, J_{10a,10b} = 11.8Hz), [3.77-3.83, m, 0.15H, H_{10a}], 3.67 (dd, 0.85H, H_{10b}, J_{5,10b} = 5.4Hz), [3.61-3.71, m, 0.3H, H_{10b}, H_{α2(Gly)}], 3.49-3.56 (m, 1.7H, H₅, H₇), [3.43-3.48, m, 0.15H, H₅], [3.40, dd apparent t, 0.15H, H₇, J_{6,7} = 9.5Hz, J_{5,6} = 9.5Hz], 3.33 (dd, 0.85H, H₇), [2.77, s, 0.45H, -NHCH₃],

2.70 (s, 2.55H, -NHCH₃), 2.61 (m, 1H, H_{3a}, H_{3a} minor), 1.99 [s, 0.45H, -COCH₃], 1.98 (s, 2.55H, -COCH₃), 1.79 (m, 1H, H_{3b}, H_{3b} minor)

(16): ¹H NMR (500 MHz, CD₃OD, 298K, 0.03 M) δ = [5.34, dd apparent t, 0.25H, H₇, J_{7,8} = 8.3Hz, J_{6,7} = 9.1Hz], 5.27 (dd apparent t, 0.75H, H₇, J_{7,8} = 9.5Hz, J_{6,7} = 9.9Hz), 4.99 (dd apparent t, 0.75H, H₆, J_{5,6} = 9.5Hz), [4.94-4.99, m, 0.25H, H₆], 4.81-4.89 (m, 0.75H, H₉), [4.69, ddd, 0.25H, H₉, J_{3a,9} = 11.1Hz, J_{3b,9} = 6.8Hz, J_{8,9} = 6.8Hz], [4.58, dd, 0.25H, H₂, J_{2,3a} = 9.5Hz, J_{2,3b} = 1.1Hz], 4.48 (d, 0.85H, H_{α1(Gly)}, J_{Hα1(Gly),Hα2(Gly)} = 17.1Hz), [4.44-4.49, m, 0.25H, H₈], 4.36-4.44 (m, 1.5H, H₂, H₈), [4.26-4.34, m, 0.25H, H_{10a}], 4.30 (dd, 0.75H, H_{10a}, J_{5,10a} = 4.0Hz, J_{10a,10b} = 12.1Hz), 4.06-4.13 (m, 0.75H, H₅), 4.05 (dd, 0.75H, H_{10b}, J_{5,10b} = 2.0Hz), [4.00-4.12, m, 0.5H, H₅, H_{10b}], 3.87 (d, 0.75H, H_{α2(Gly)}), [3.80, d, 0.25H, H_{α1(Gly)}, J_{Hα1(Gly),Hα2(Gly)} = 16.7Hz], [3.58, d, 0.25H, H_{α2(Gly)}], [2.87, ddd, 0.25H, H_{3a}, J_{3a,3b} = 12.7Hz], 2.68-2.80 (m, 3.75H, H_{3a}, -NHCH₃, -NHCH₃ minor), 2.11 (s, 2.25H, -COCH₃), 2.03 (s, 3H, -COCH₃, -COCH₃ minor), 1.99 (s, 3H, -COCH₃, -COCH₃ minor), [1.98, s, 0.75H, -COCH₃ minor], 1.97 (s, 3H, -COCH₃, -COCH₃ minor), [1.95-2.01, m, 0.25H, H_{3b}], 1.84 (ddd, 0.75H, H_{3b}, J_{2,3b} = 1.0Hz, J_{3b,9} = 7.4Hz, J_{3a,3b} = 13.0Hz)

(16): ¹H NMR (500 MHz, DMSO-D₆, 298K, 0.03 M) δ = [8.20, q, 0.17H, -NHCH₃, J = 4.5Hz], 7.96 (dd, 0.83H, -NH_(Gly), J_{NH,Hα1(Gly)} = 7.5Hz, J_{NH,Hα2(Gly)} = 3.7Hz), 7.88 (q, 0.83H, -NHCH₃, J = 4.5Hz), [7.79, dd, 0.17H, -NH_(Gly), J_{NH,Hα1(Gly)} = 6.0Hz, J_{NH,Hα2(Gly)} = 4.5Hz], [5.28, dd apparent t, 0.17H, H₇, J_{7,8} = 8.2Hz, J_{6,7} = 8.6Hz], 5.14 (dd apparent t, 0.83H, H₇, J_{7,8} = 9.7Hz, J_{6,7} = 10.1Hz), 4.82 (dd apparent t, 0.83H, H₆, J_{5,6} = 9.7Hz), [4.78, dd apparent t, 0.17H, H₆, J_{5,6} = 9.0Hz], 4.65 (ddd, 0.83H, H₉, J_{3a,9} = 11.5Hz, J_{3b,9} = 7.4Hz, J_{8,9} = 7.1Hz), [4.54, ddd, 0.17H, H₉, J_{3a,9} = 10.1Hz, J_{3b,9} = 6.7Hz, J_{8,9} = 7.1Hz],

4.43 (dd, 0.83H, H₈), [4.39, dd, 0.17H, H₂, J_{2,3a} = 9.2Hz, J_{2,3b} = 2.1Hz], 4.33 (dd, 0.83H, H_{α1(Gly)}, J_{Hα1(Gly),Hα2(Gly)} = 17.6Hz), [4.24, dd apparent t, 0.17H, H₈], 4.10-4.20 (m, 2.66H, H₂, H₅, H_{10a}, H_{10a} minor), [4.04, ddd, 0.17H, H₅, J = 2.6Hz, J = 5.6Hz], 3.87-3.96 (m, 1H, H_{10b}, H_{10b} minor), 3.67 (dd, 0.83H, H_{α2(Gly)}), [3.61, dd, 0.17H, H_{α1(Gly)}, J_{Hα1(Gly),Hα2(Gly)} = 16.9Hz], [3.34, dd, 0.17H, H_{α2(Gly)}], [2.75-2.84, m, 0.17H, H_{3a}], 2.78 (ddd, 0.83H, H_{3a}, J_{2,3a} = 10.8Hz, J_{3a,3b} = 12.5Hz), [2.57, d, 0.51H, -NHCH₃, J =], 2.49 (d, 2.49H, -NHCH₃), 2.03 (s, 2.49H, -COCH₃), [1.96, s, 0.51H, -COCH₃], 1.95 (s, 2.49H, -COCH₃), 1.92 (s, 3H, -COCH₃, -COCH₃ minor), [1.87, s, 0.51H, -COCH₃], [1.78, s, 0.51H, -COCH₃], 1.78 (s, 2.49H, -COCH₃), 1.57 (ddd, 0.83H, H_{3b}, J_{2,3b} = 1.0Hz, H_{3b} minor)

Table 9.1: Assigned Coupling Constants of the Major Isomer for Compounds **11-16** (Hz)

		$J_{2,3a}$	$J_{2,3b}$	$J_{3a,9}$	$J_{3b,9}$	$J_{5,6}$	$J_{6,7}$	$J_{7,8}$	$J_{8,9}$
Compound	Solvent								
11	CDCl ₃	9.5	0.8	11.7	7.3	Ovlp ^[a]	9.0	9.2	7.4
	CD ₃ OD	9.7	1.0	11.7	7.3	9.3	9.3	9.3	7.1
	DMSO-D ₆	9.8	1.0	Ovlp ^[a]	7.5	9.4	9.8	9.4	7.2
12	D ₂ O	10.2	1.3	Ovlp ^[a]	7.8	9.7	9.9	9.2	7.2
	CD ₃ OD	9.8	1.0	12.0	7.2	9.4	9.6	9.0	7.2
	DMSO-D ₆	9.7	1.0	12.0	7.4	9.5	9.4	9.2	7.0
13	CDCl ₃	9.8	Ovlp ^[a]	11.7	7.1	9.9	10.0	9.8	7.3
	CD ₃ OD	9.7	1.3	Ovlp ^[a]	7.4	9.8	9.8	9.6	7.1
	DMSO-D ₆	10.0	1.0	11.5	7.1	9.8	9.6	9.6	7.1
14	CDCl ₃	9.7	1.0	12.0	7.6	9.4	9.7	Ovlp ^[a]	7.2
	CD ₃ OD	10.0	1.0	12.0	7.9	Ovlp ^[a]	Ovlp ^[a]	9.0	7.0
15	D ₂ O	10.3	1.0	11.2	7.7	9.5	9.9	9.0	7.2
	CD ₃ OD	9.9	1.0	12.3	7.2	Ovlp ^[a]	Ovlp ^[a]	9.0	7.1
16	CDCl ₃	10.0	Ovlp ^[a]	11.8	7.5	9.8	9.9	9.4	7.2
	CD ₃ OD	Ovlp ^[a]	1.0	Ovlp ^[a]	7.4	9.5	9.9	9.5	Ovlp ^[a]
	DMSO-D ₆	10.8	1.0	11.5	7.4	9.7	10.1	9.7	7.1
Average		9.9	1.0	11.8	7.4	9.6	9.7	9.3	7.2
Standard Error		0.32	0.13	0.30	0.24	0.20	0.30	0.27	0.11

^[a]not assigned since peak was overlapped in ¹H spectrum.

Table 9.2: Assigned Coupling Constants of the Minor Isomer for Compounds **12-16** (Hz)

Compound	Solvent	$J_{2,3a}$	$J_{2,3b}$	$J_{3a,9}$	$J_{3b,9}$	$J_{5,6}$	$J_{6,7}$	$J_{7,8}$	$J_{8,9}$
12	D ₂ O	n.a. ^[a]	1.8	n.a. ^[a]	6.7	9.9	n.a. ^[a]	8.0	7.7
	CD ₃ OD	9.0	1.6	10.9	7.3	9.8	9.2	8.4	7.1
	DMSO-D ₆	9.0	1.7	10.9	7.1	9.4	9.4	8.3	7.2
13	CDCl ₃	9.4	1.5	11.6	6.8	9.2	9.5	7.8	7.0
	CD ₃ OD	9.3	1.9	10.9	6.8	n.a. ^[a]	9.4	9.3	7.2
	DMSO-D ₆	9.6	1.4	10.8	6.6	9.4	8.9	8.5	6.6
14	CDCl ₃	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	8.2	6.5
15	D ₂ O	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	n.a. ^[a]	7.7	7.2
	CD ₃ OD	9.3	1.8	n.a. ^[a]	n.a. ^[a]	9.5	9.5	8.3	7.1
16	CDCl ₃	9.5	n.a. ^[a]	10.9	6.7	n.a. ^[a]	9.1	8.3	7.3
	CD ₃ OD	9.5	1.1	11.1	6.8	n.a. ^[a]	9.1	8.3	6.8
	DMSO-D ₆	9.2	2.1	10.1	6.7	9.0	8.6	8.2	7.1
Average		9.3	1.7	10.9	6.8	9.5	9.2	8.3	7.1
Standard Error		0.21	0.30	0.41	0.22	0.32	0.30	0.40	0.32

^[a]not assigned since peak was not observed or overlapped in ¹H spectrum

Table 9.3: % Interproton Effect Observed From nOe Experiments Under Conditions A-F
For Compounds **11-13***

Compound	Solvent	A	B	C	D	E	F
11	CDCl ₃	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]
	CD ₃ OD	5.2	0.4	n.r. ^[a]	n.r. ^[a]	7.8	n.r. ^[a]
	DMSO-D ₆	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]
12	D ₂ O	5.5	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	6.9	n.r. ^[a]
	CD ₃ OD	3.9	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	6.0	n.r. ^[a]
	DMSO-D ₆	3.5	0.6	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]
13	CDCl ₃	4.5	n.r. ^[a]	4.2	3.5	7.9	n.r. ^[a]
	CD ₃ OD	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	6.9	n.r. ^[a]
	DMSO-D ₆	3.4	n.r. ^[a]	3.3	n.r. ^[a]	n.r. ^[a]	<0.4

^[a]experiment not recorded because of peak overlap in ¹H NMR spectrum.

*nOe experiment	Description	Result
A	Selective inversion of H-8 (major) showed interproton effect to the terminal N-acyl singlet (major)	Major isomer established as <i>trans</i> isomer
B	Selective inversion of H-2 (major) showed very little interproton effect to the terminal N-acyl singlet (major)	Major isomer established as <i>trans</i> isomer
C	Selective inversion of the terminal N-acyl singlet (minor) showed interproton effect to H-2 (minor)	Minor isomer established as <i>cis</i> isomer
D	Selective inversion of H-2 (minor) showed interproton effect to the terminal N-acyl singlet (minor)	Minor isomer established as <i>cis</i> isomer
E	Selective inversion of terminal N-acyl singlet (major) showed interproton effect to H-8 (major)	Major isomer established as <i>trans</i> isomer
F	Selective inversion of H-8 (minor) showed no interproton effect to the terminal N-acyl singlet (minor)	Minor isomer established as <i>cis</i> isomer

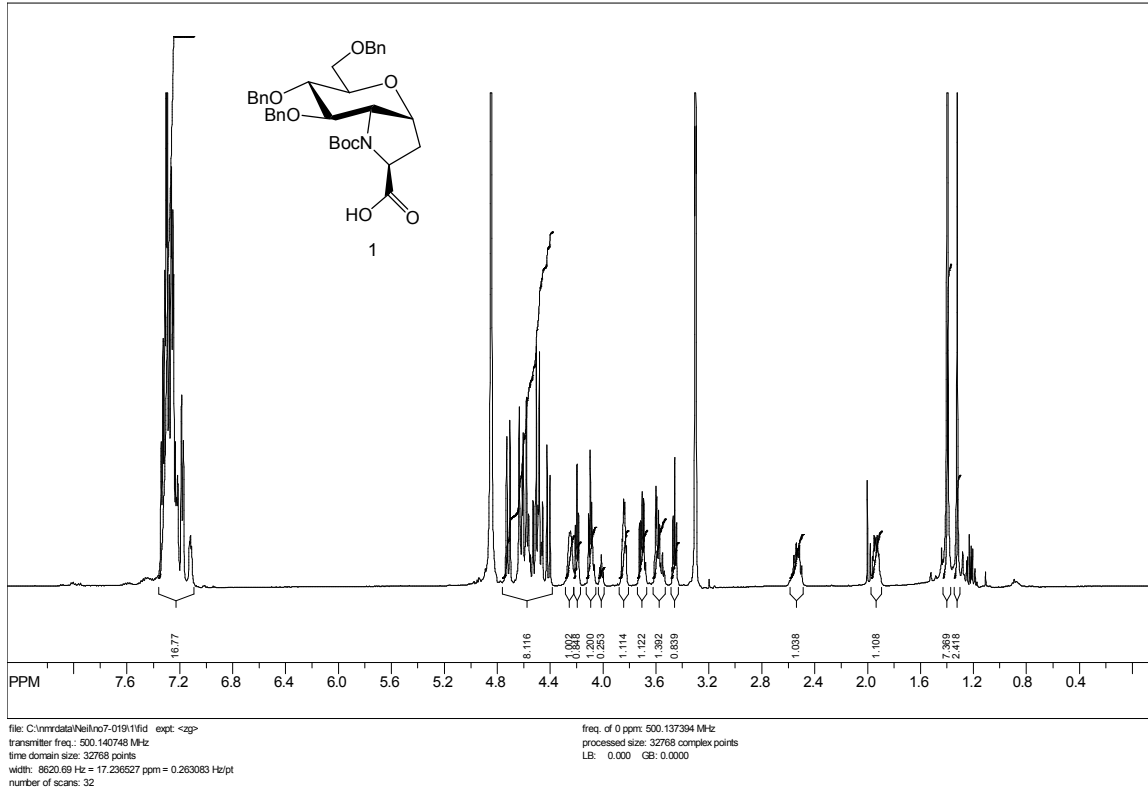
Table 9.4: % Interproton Effect Observed From nOe Experiments Under Conditions A-D
For Compounds **14-16***

Compound	Solvent	A	B	C	D
14	CDCl ₃	n.r. ^[a]	none	n.r. ^[a]	n.r. ^[a]
	CD ₃ OD	6.1	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]
15	D ₂ O	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]
	CD ₃ OD	4.9	none	n.r. ^[a]	n.r. ^[a]
16	CDCl ₃	4.3	n.r. ^[a]	n.r. ^[a]	n.r. ^[a]
	CD ₃ OD	n.r. ^[a]	n.r. ^[a]	6.7	<0.4

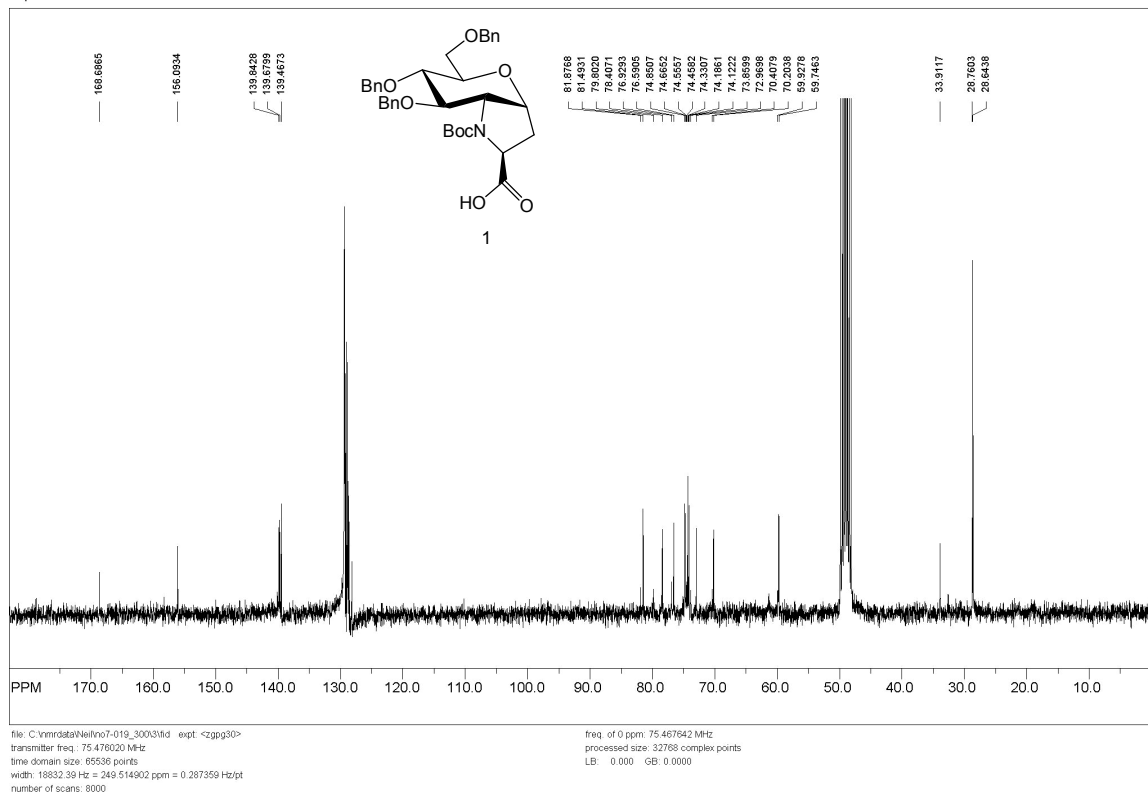
^[a]experiment not recorded because of peak overlap in ¹H NMR spectrum.

*nOe experiment	Description	Result
A	Selective inversion of H-8 (major) showed interproton effect to H _{α(Gly)} (major)	Major isomer established as <i>trans</i> isomer
B	Selective inversion of H-2 (major) showed very little interproton effect to H _{α(Gly)} (major)	Major isomer established as <i>trans</i> isomer
C	Selective inversion of H _{α(Gly)} (major) showed interproton effect to H-8 (major)	Major isomer established as <i>trans</i> isomer
D	Selective inversion of H _{α(Gly)} (minor) showed interproton effect to H-8 (minor)	Minor isomer established as <i>cis</i> isomer

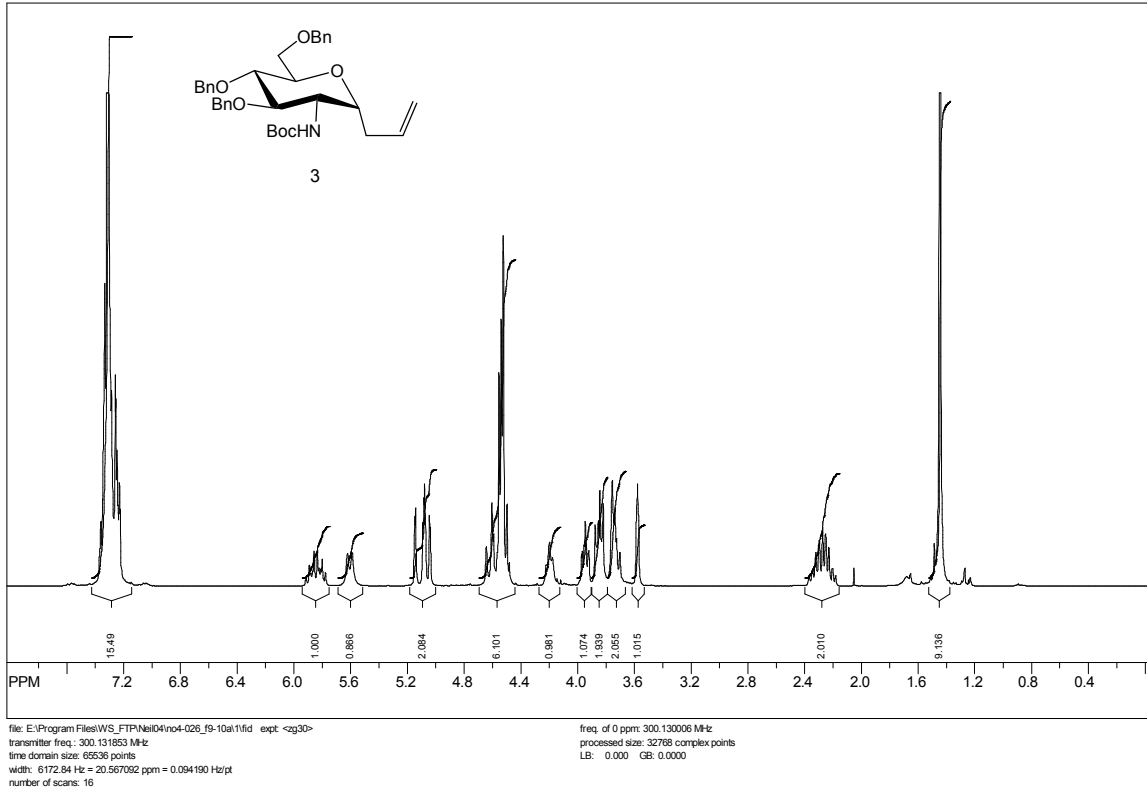
SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



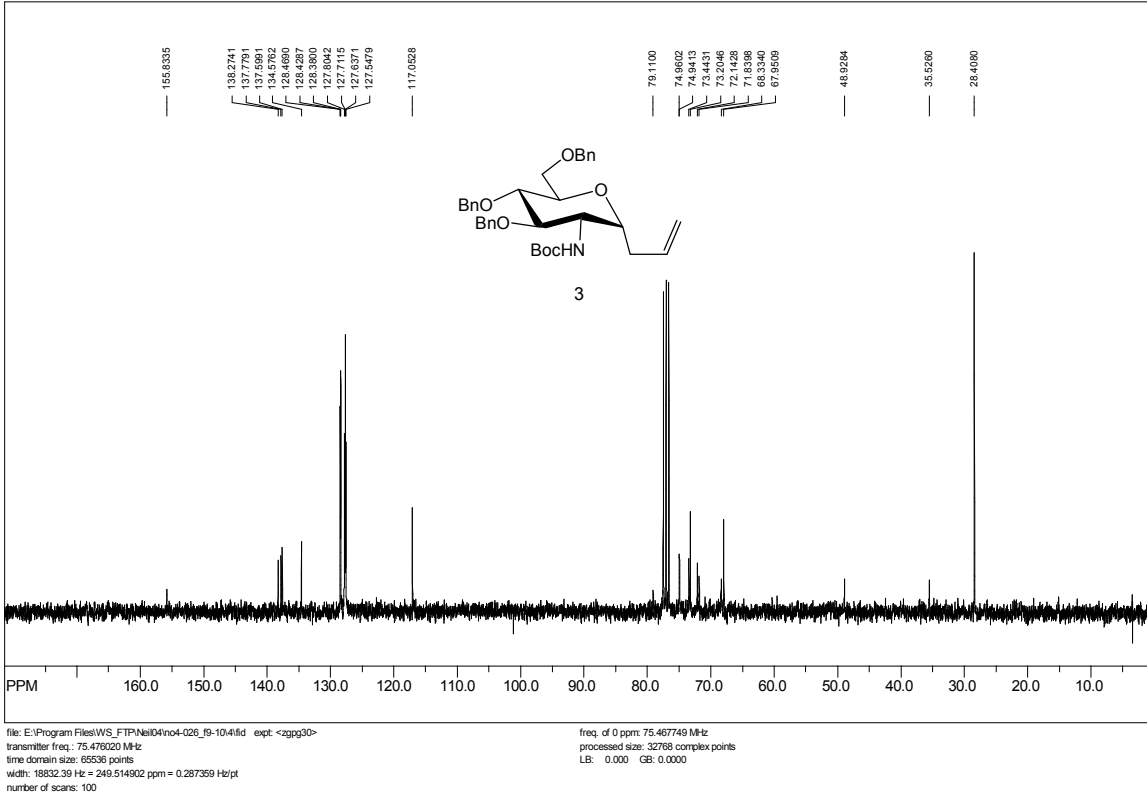
SpinWorks 2.5: C13CPD MeOH u schweiz 1



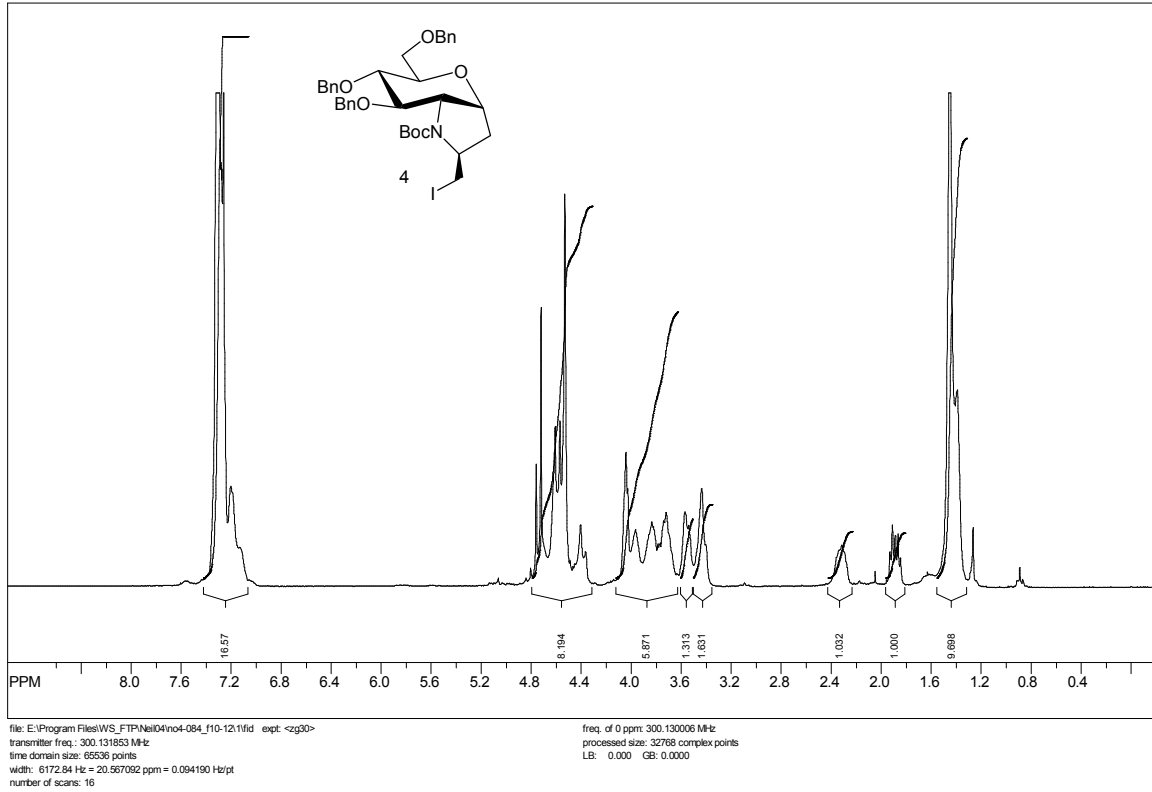
SpinWorks 2.5: PROTON CDCI3 u schweiz 1



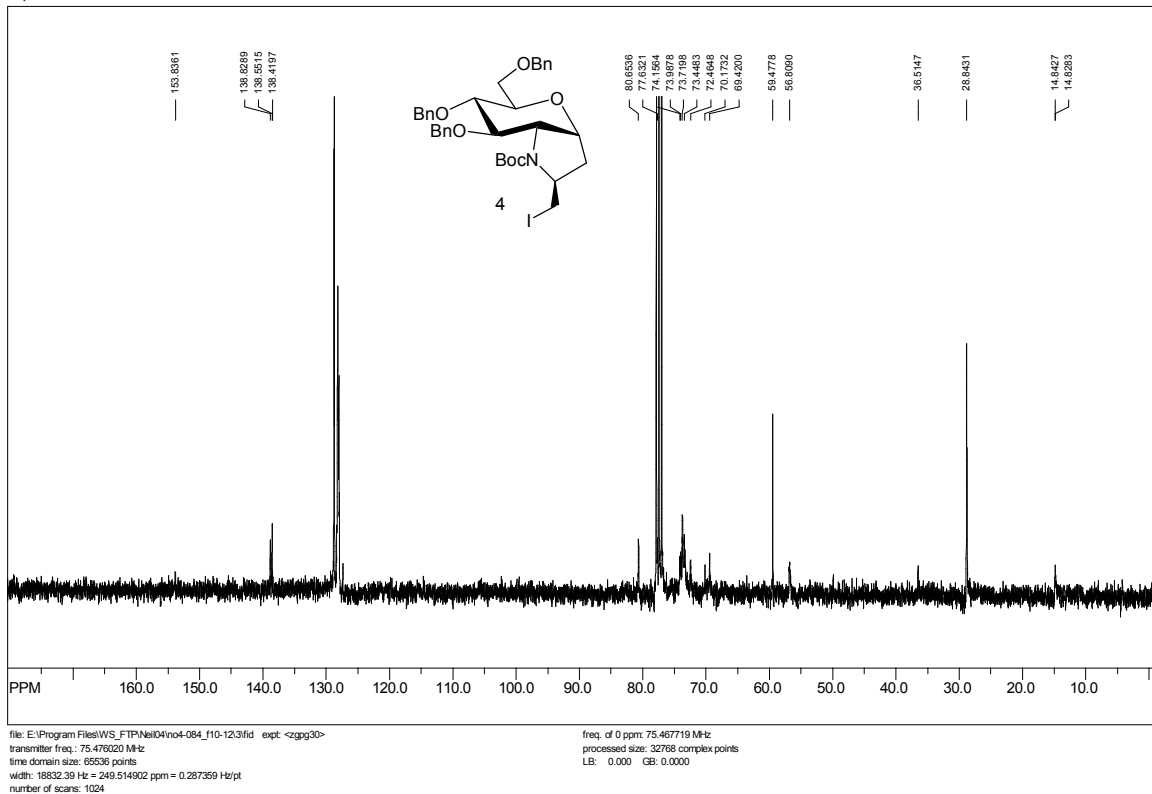
SpinWorks 2.5: C13CPD CDCI3 u schweiz 2



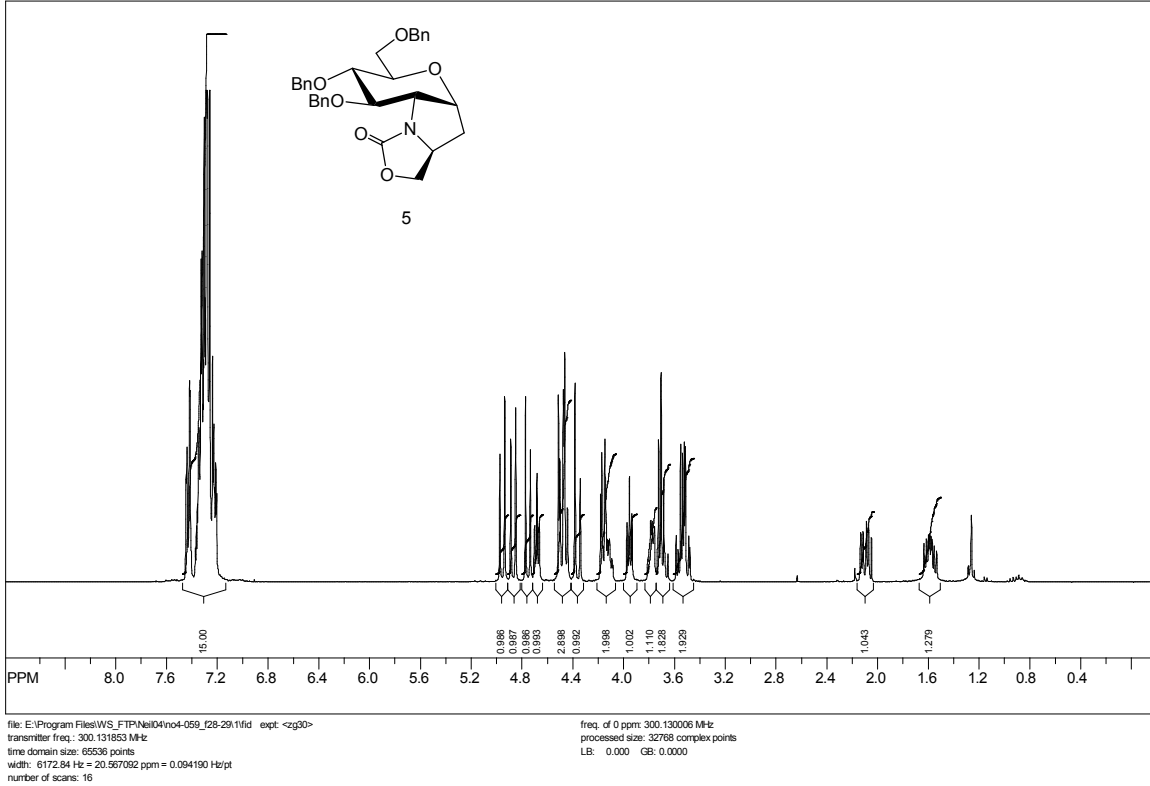
SpinWorks 2.5: PROTON CDCI3 u schweiz 1



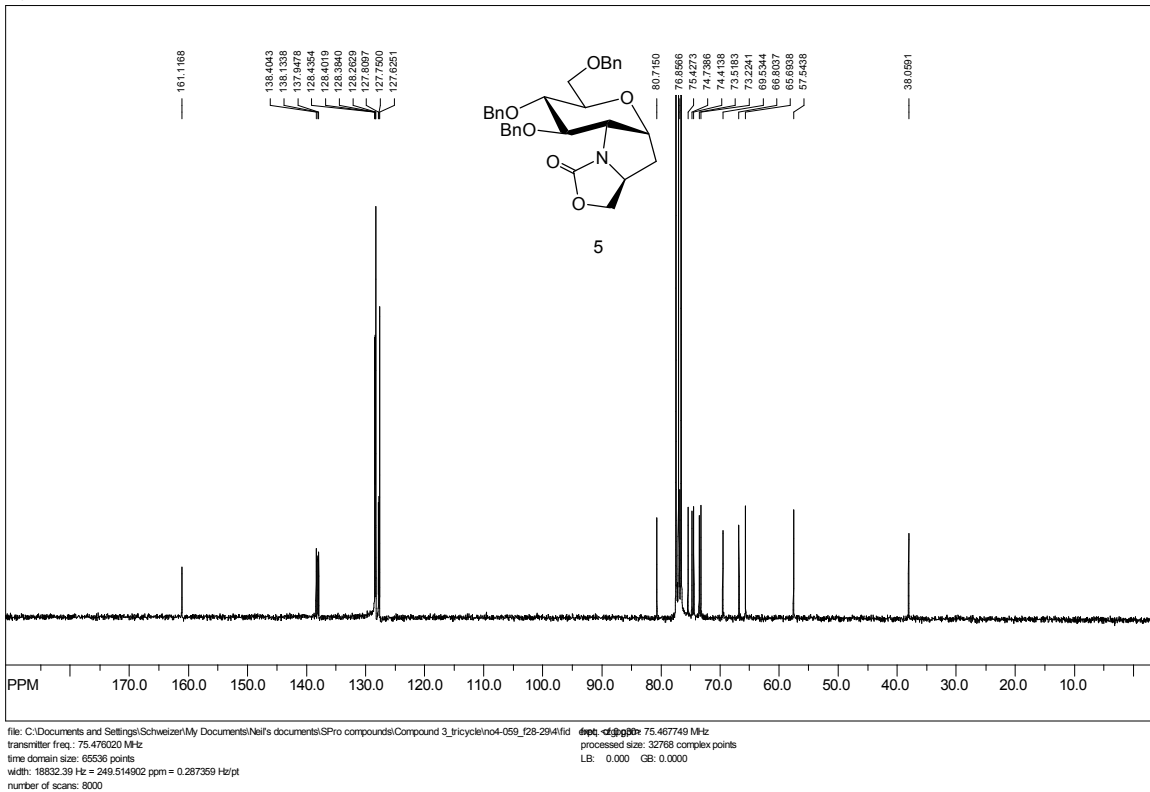
SpinWorks 2.5: C13CPD CDCI3 u schweiz 1



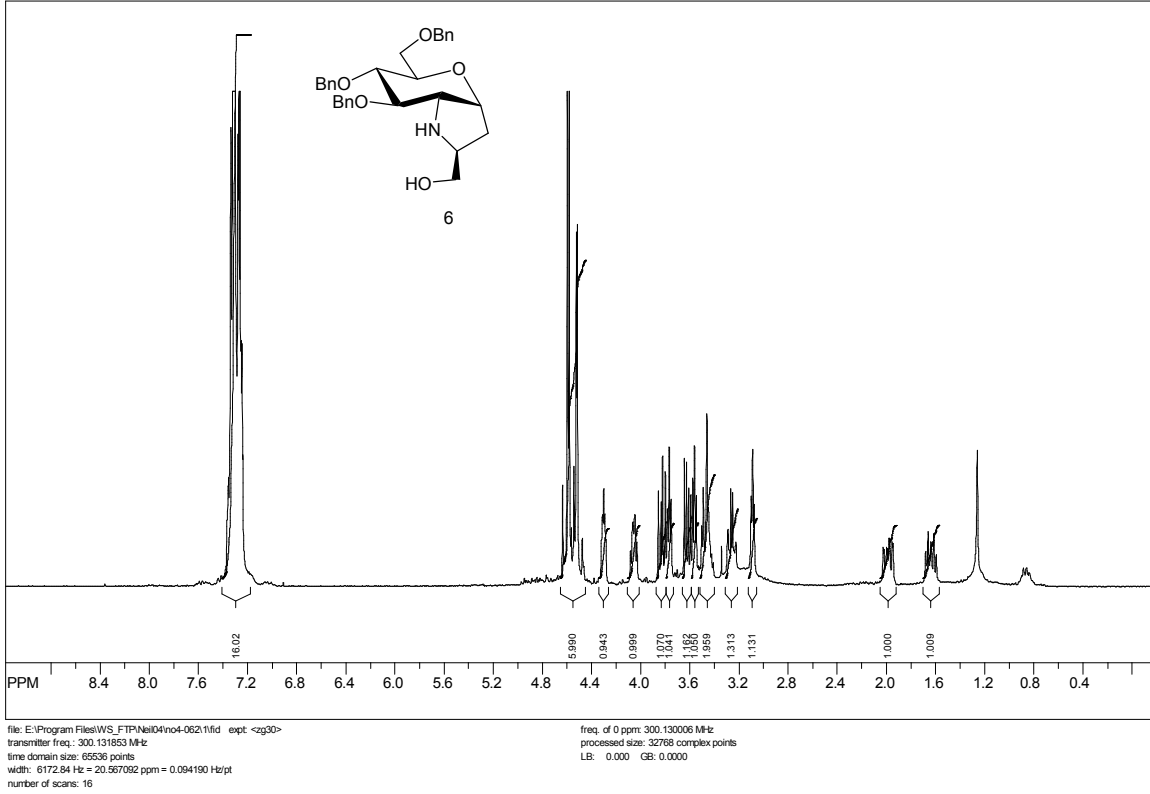
SpinWorks 2.5: PROTON CDCI3 u schweiz 1



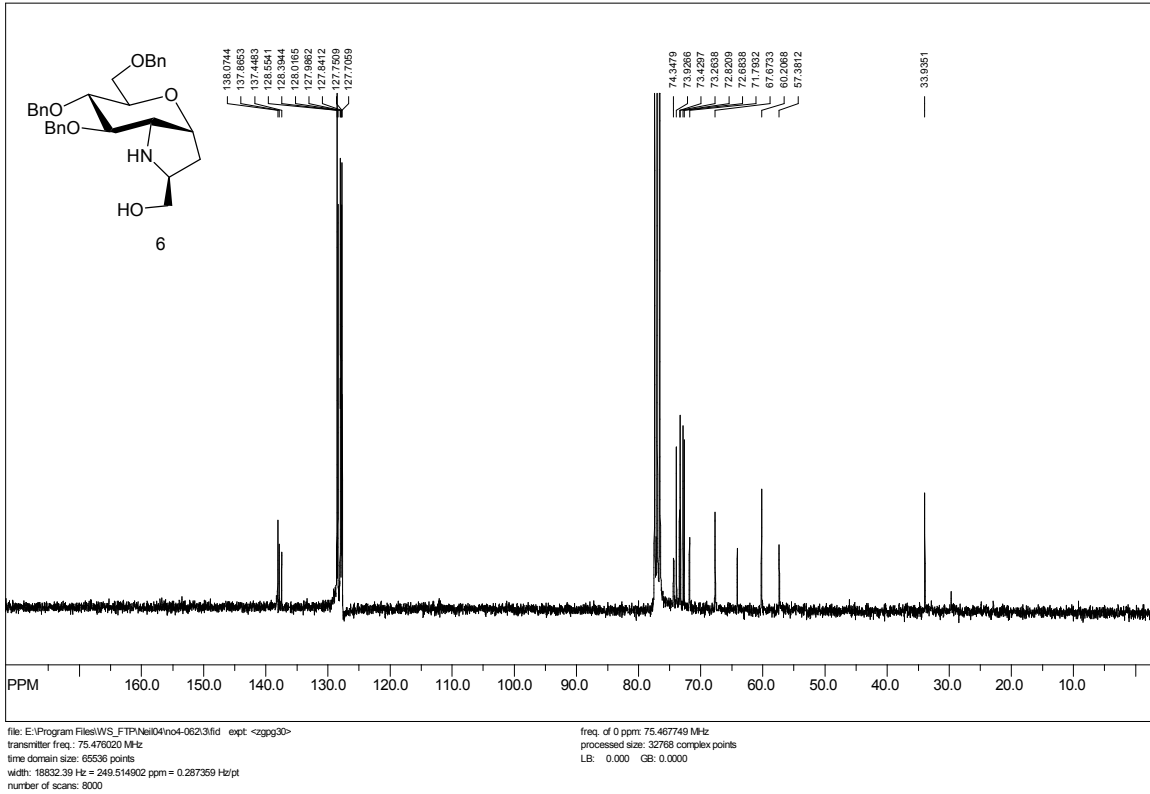
SpinWorks 2.5: C13CPD CDCI3 u schweiz 1



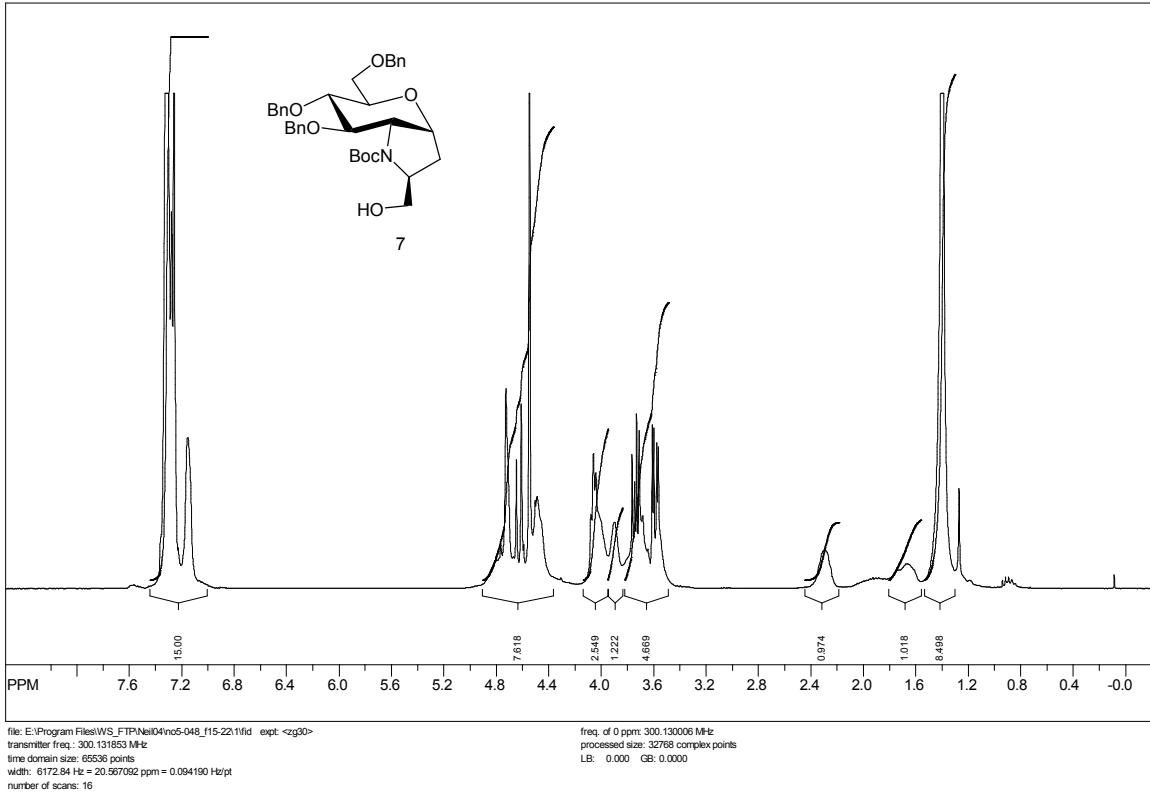
SpinWorks 2.5: PROTON CDCI3 u schweiz 1



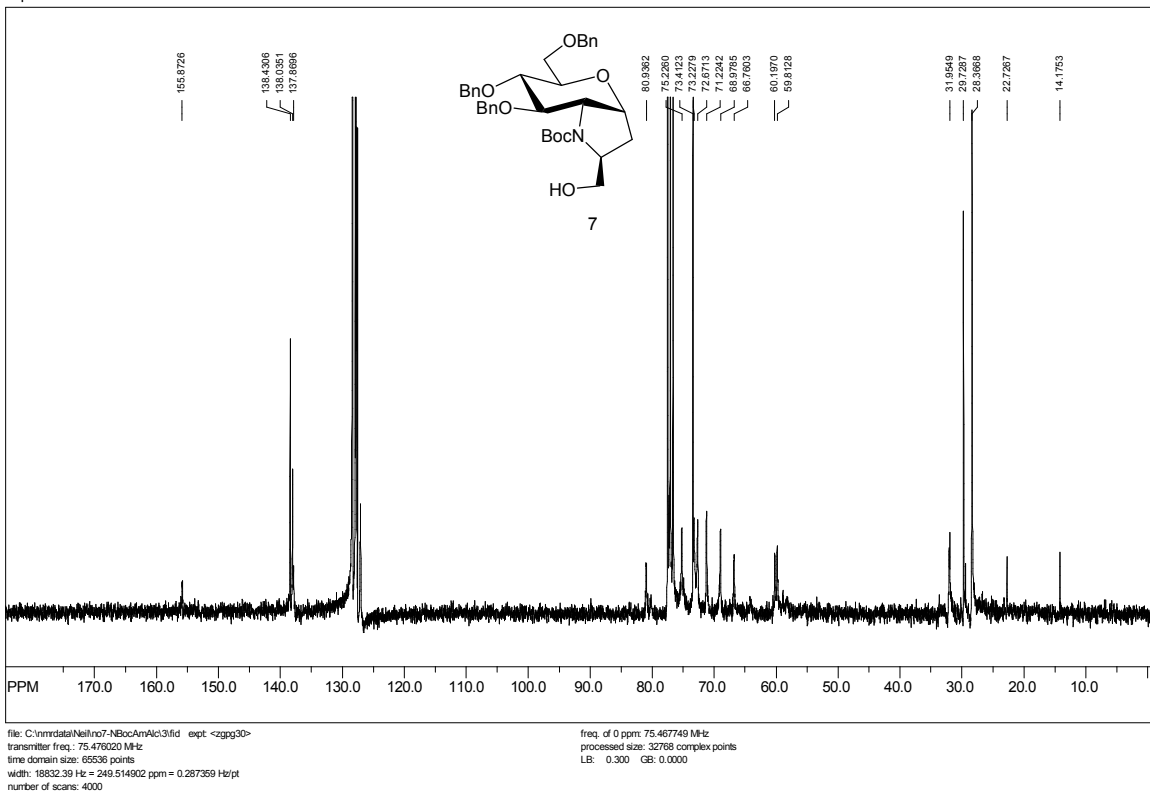
SpinWorks 2.5: C13CPD CDCI3 u schweiz 1



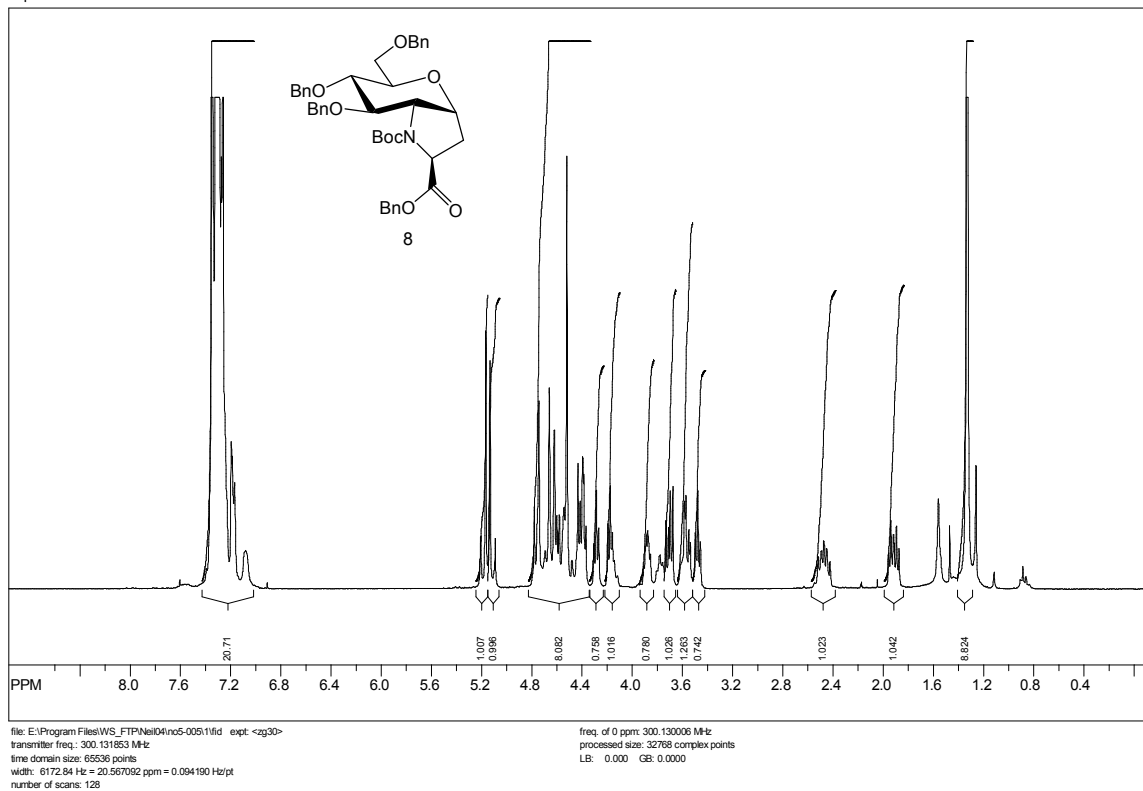
SpinWorks 2.5: PROTON CDCI3 u schweiz 1



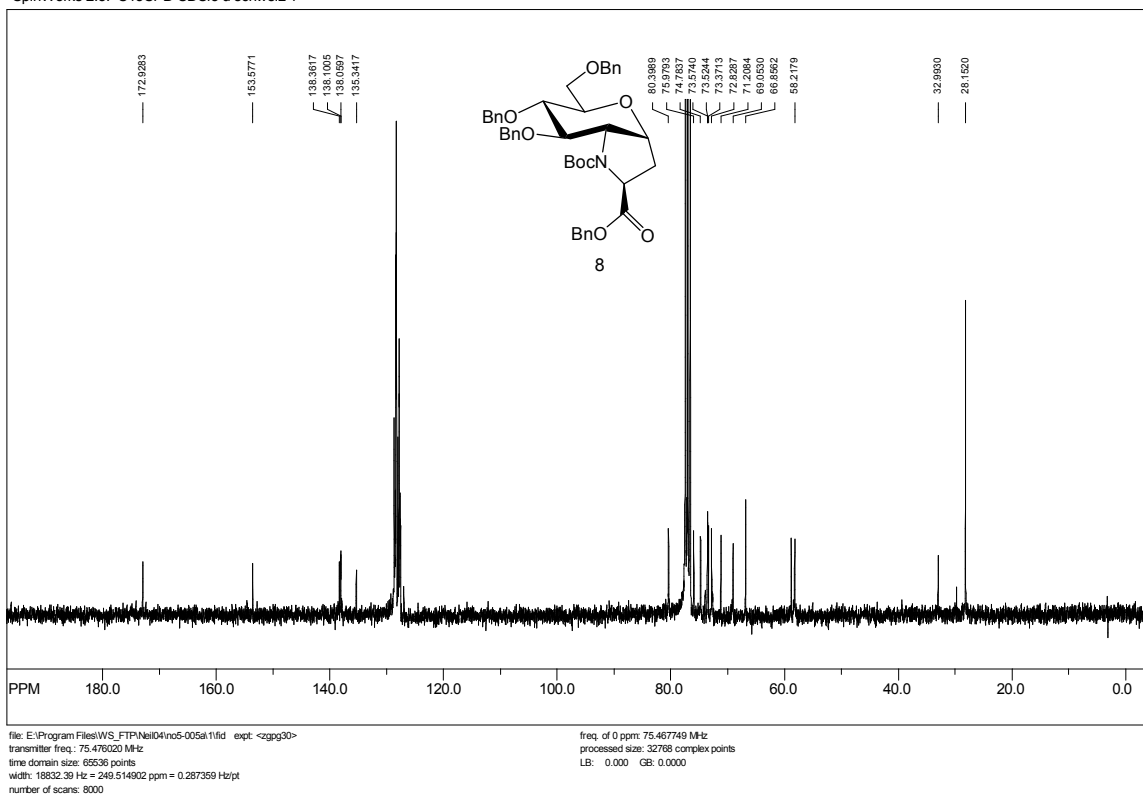
SpinWorks 2.5: C13CPD CDCI3 u schweiz 1



SpinWorks 2.5: PROTON CDCI3 u schweiz 1

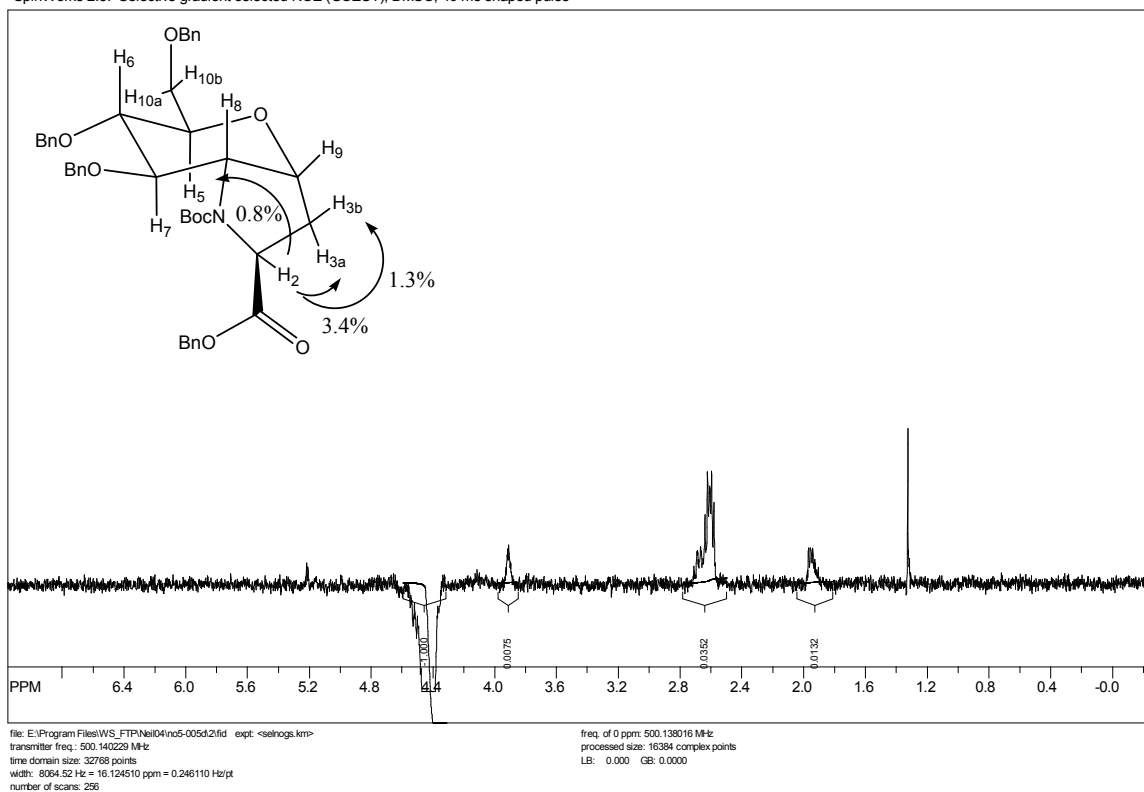


SpinWorks 2.5: C13CPD CDCI3 u schweiz 1

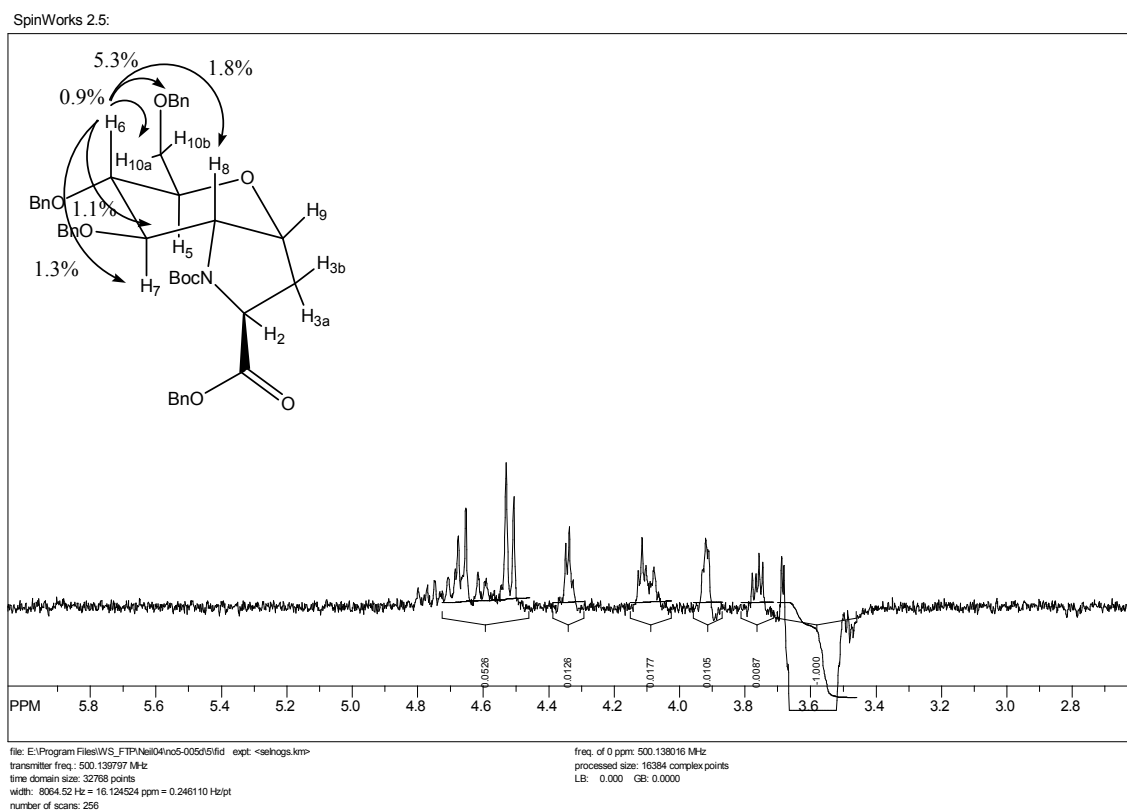


Subjection of the H₂ signal (4.42 ppm, 1 H) to a one-dimensional GOESY experiment in acetone-D₆ showed inter-proton effect to H₅ (3.92 ppm, 1 H) (0.8%) along with observed nOe contact to H_{3a} (2.62 ppm, 1 H) (3.4%) and H_{3b} (1.94 ppm, 1 H) (1.3%), while there was no nOe contact observed between H₂ and H₉ (4.58 ppm, 1 H) or H₈ (4.10 ppm, 1 H).

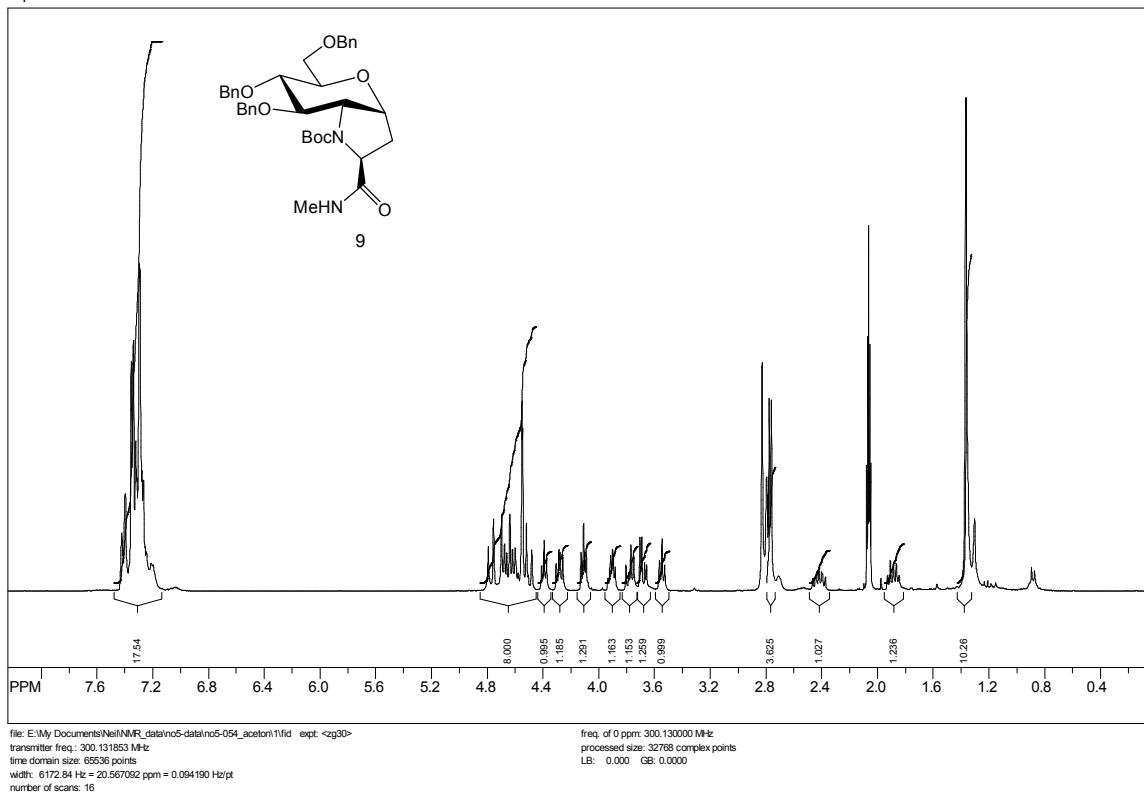
SpinWorks 2.5: Selective gradient selected NOE (GOESY), DMSO, 40 ms shaped pulse



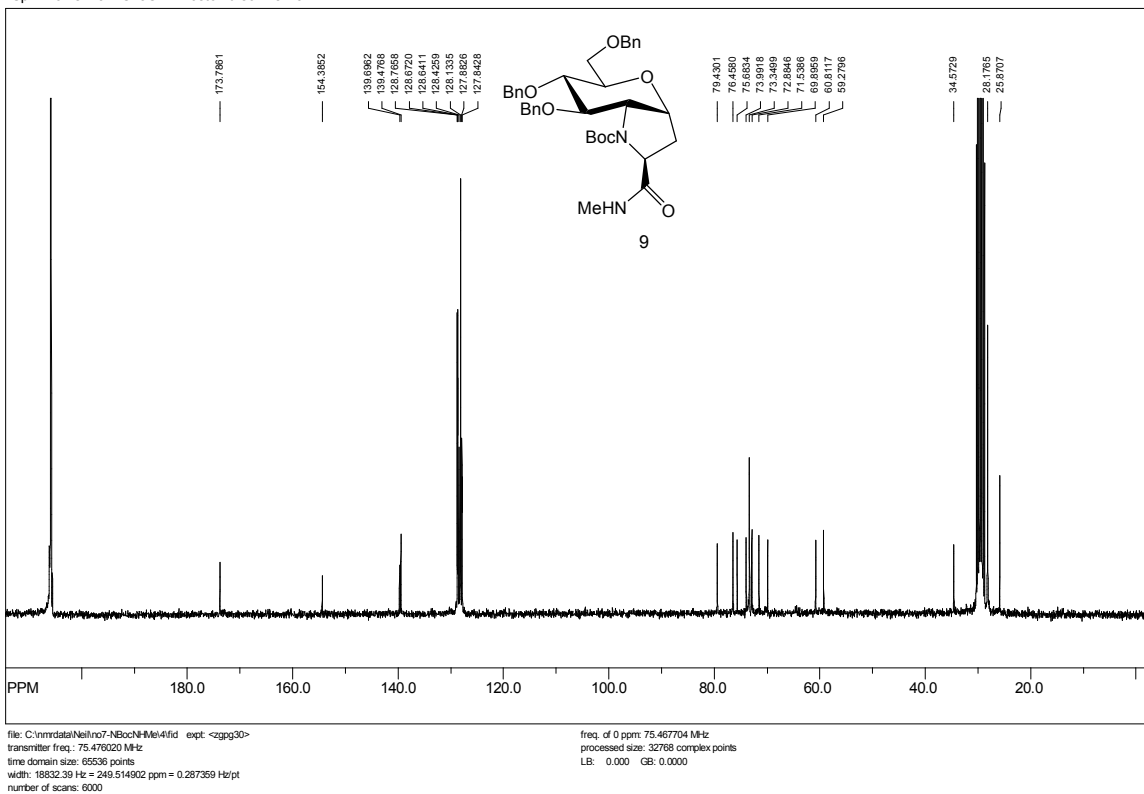
Subjection of the H₆ signal (3.56 ppm, 1 H) to a one-dimensional GOESY experiment in acetone-D₆ showed inter-proton effect to H₈ (4.12 ppm) (1.8%) along with observed nOe contact to H₇ (4.34 ppm) (1.3%), H₅ (3.92 ppm, 1 H) (1.1%), H_{10a} (3.76 ppm, 1 H) (0.9%), and benzylic protons (4.52 ppm, 4.67 ppm, 2 H) (5.3%) while there was no nOe contact observed between H₆ and H₂ (4.42 ppm, 1 H).



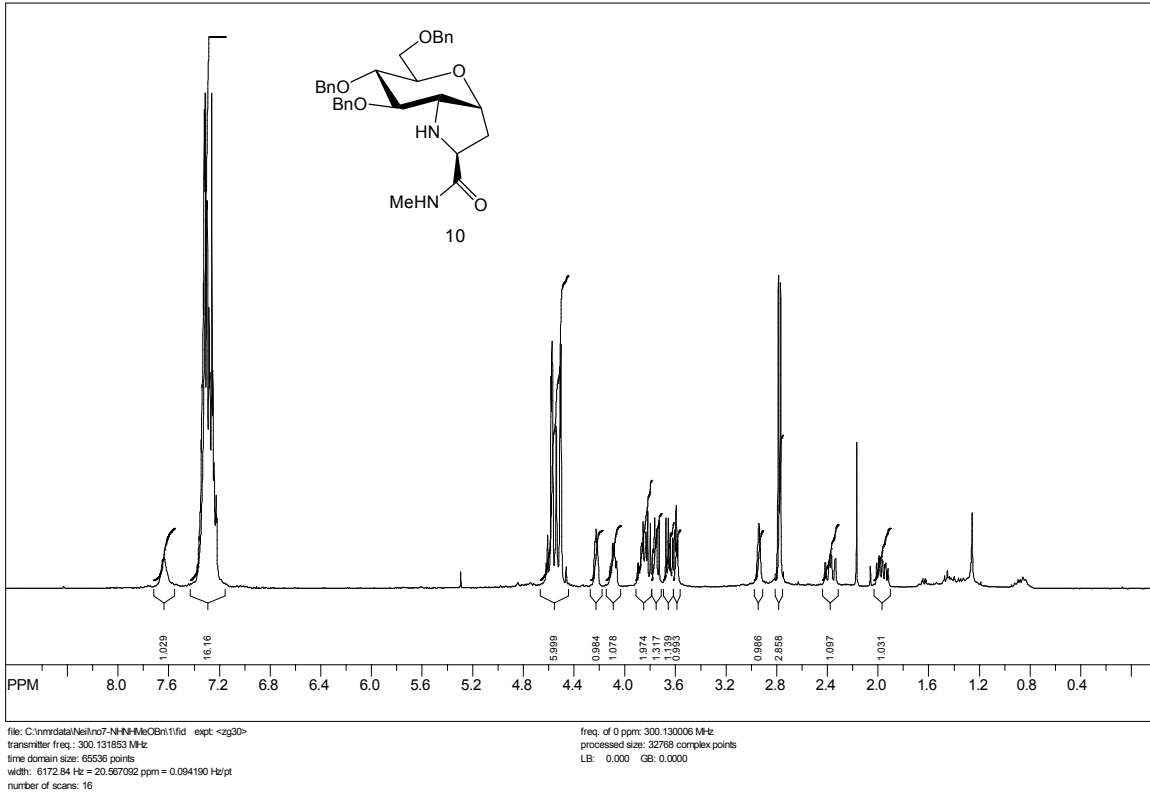
SpinWorks 2.5: PROTON Aceton u schweiz 1



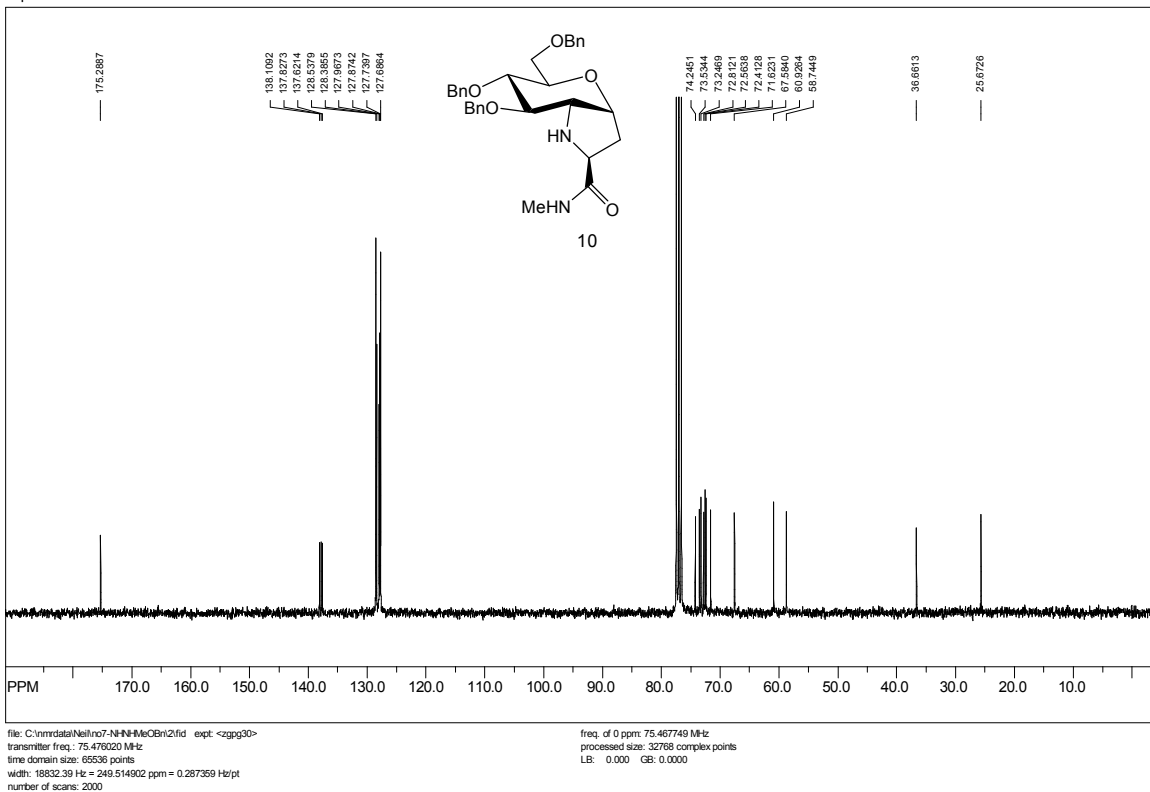
SpinWorks 2.5: C13CPD Aceton u schweiz 3



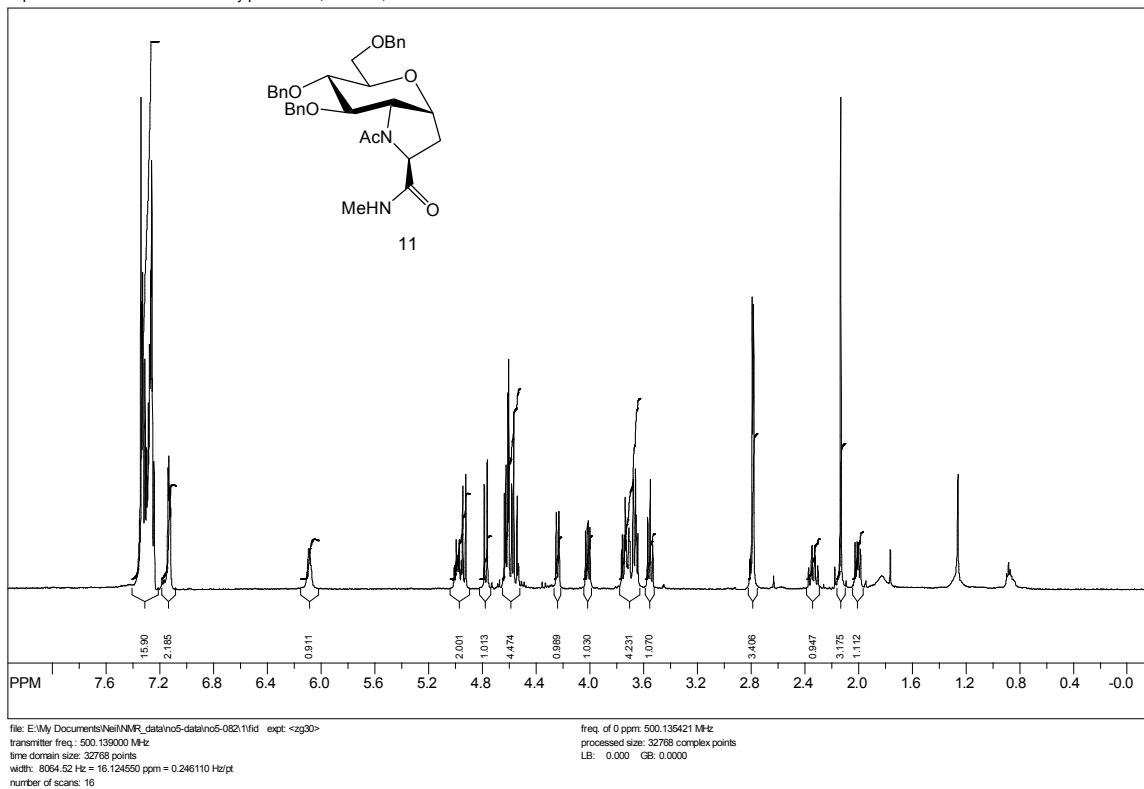
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



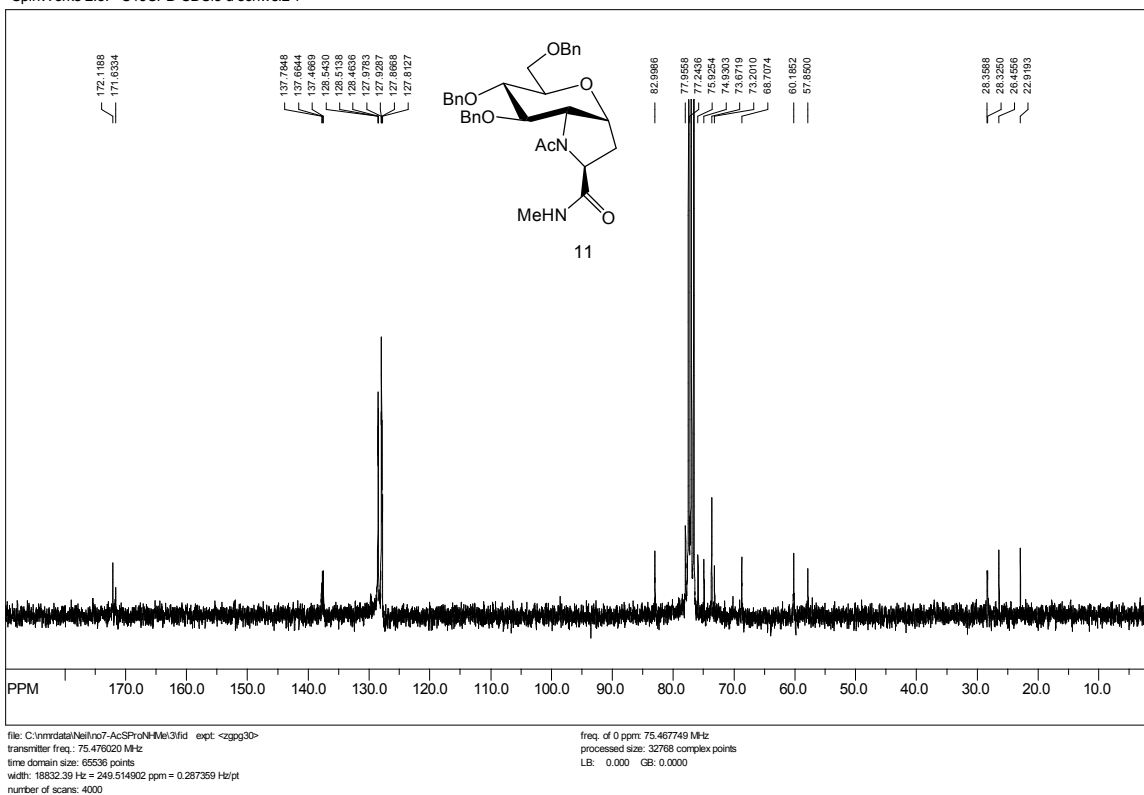
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



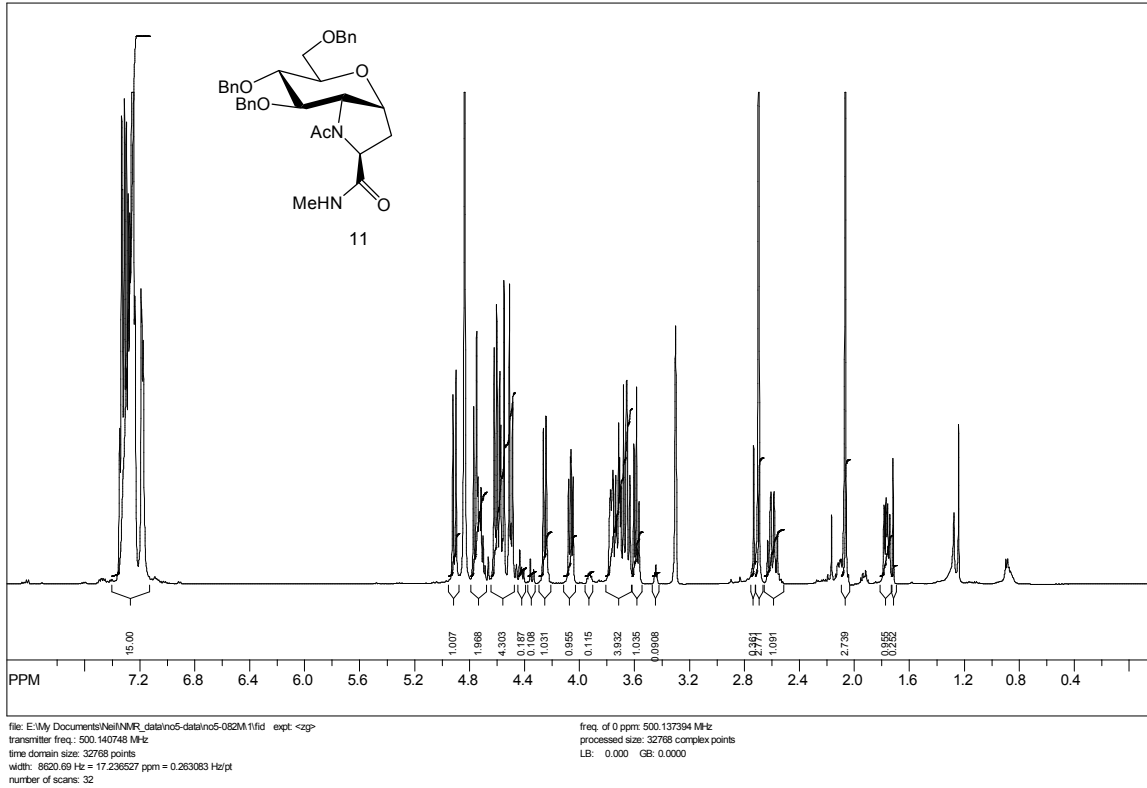
SpinWorks 2.5: standard 1-H survey parameters, AMX500, CDCl3



SpinWorks 2.5: C13CPD CDCl3 u schweiz 1

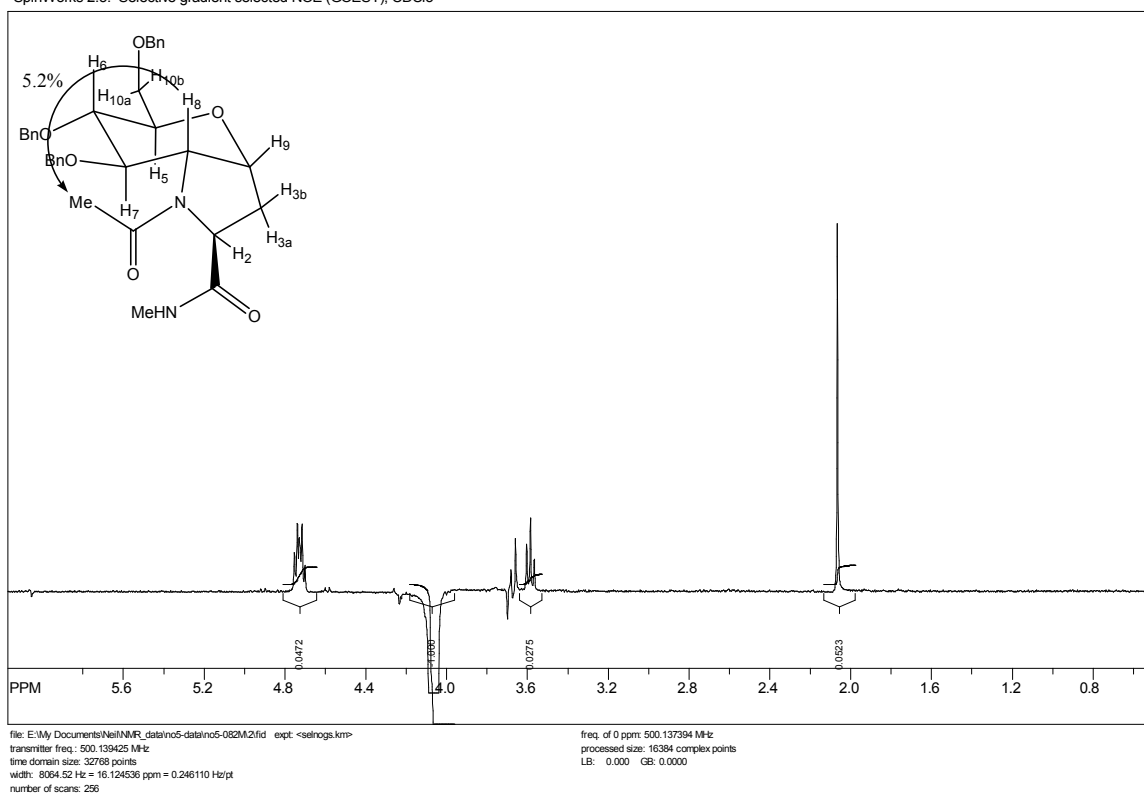


SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



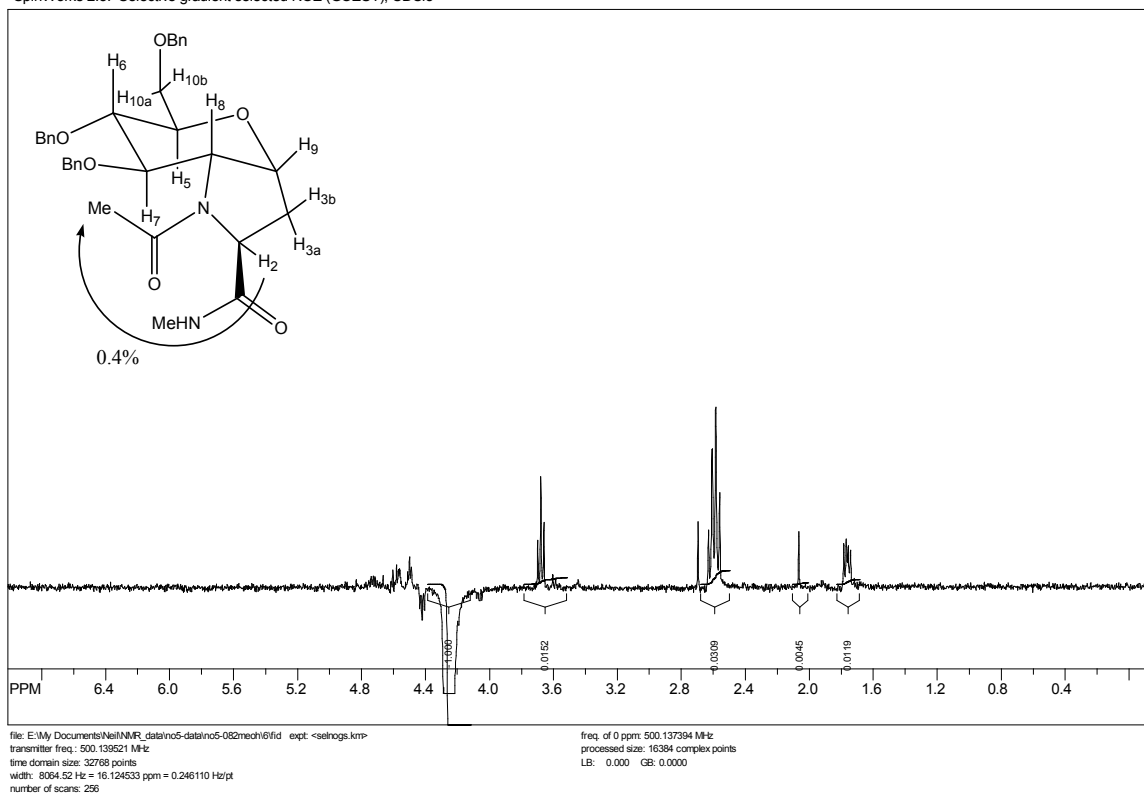
Subjection of the H₈ signal of the major conformer (4.07 ppm, 0.9 H) to a one-dimensional GOESY experiment in CD₃OD showed inter-proton effect to the acetate methyl singlet -COCH₃ of the major conformer (2.05 ppm, 2.7 H) (5.2%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃



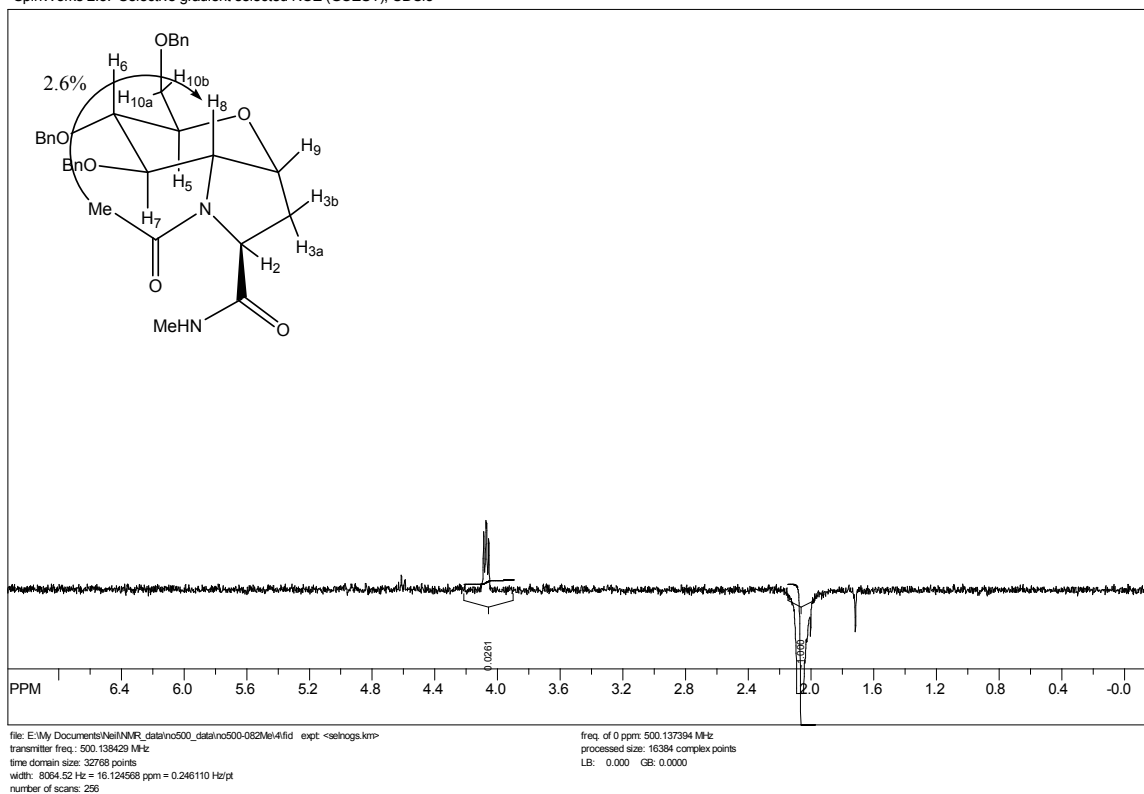
Subjection of the H₂ signal of the major conformer (4.25 ppm, 0.9 H) to a one-dimensional GOESY experiment in CD₃OD showed a small inter-proton effect to the acetate methyl singlet -COCH₃ of the major conformer (2.05 ppm, 2.7 H) (0.4%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃

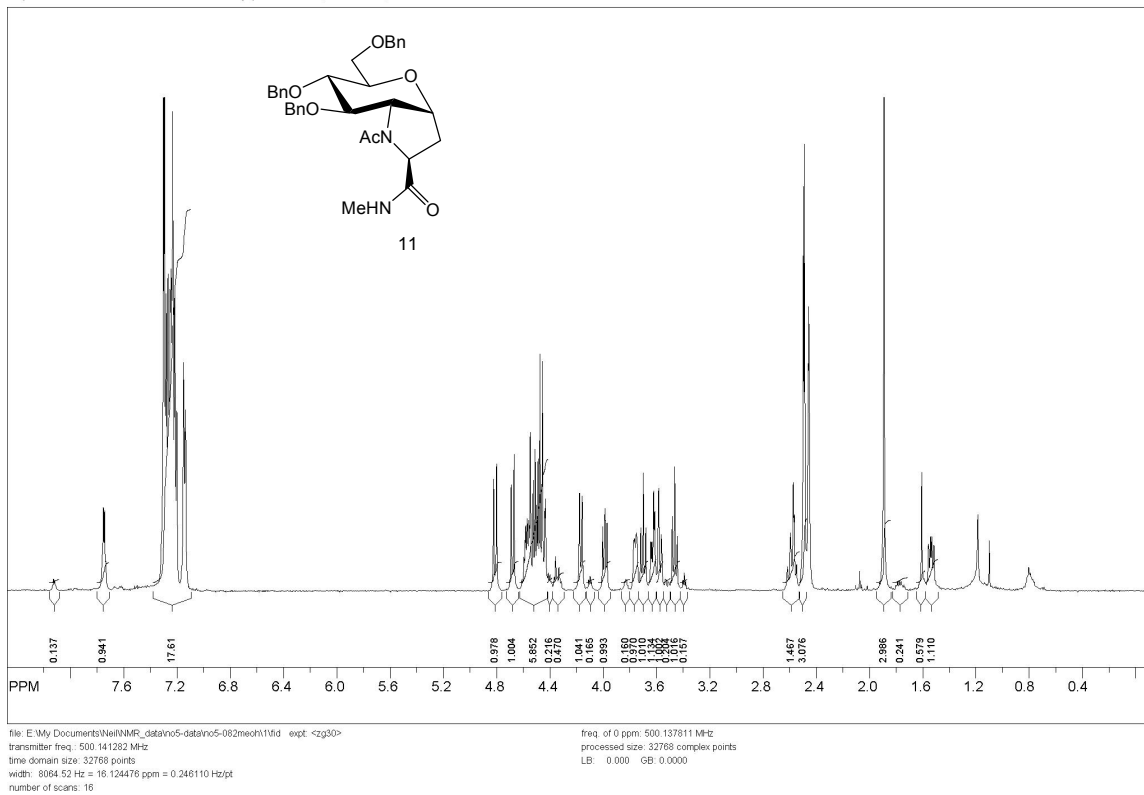


Subjection of the acetate methyl singlet $-\text{COCH}_3$ of the major conformer (2.05 ppm, 2.7 H) to a one-dimensional GOESY experiment in CD_3OD showed a large inter-proton effect to the H_8 signal of the major conformer (4.07 ppm, 0.9 H) (7.8%)

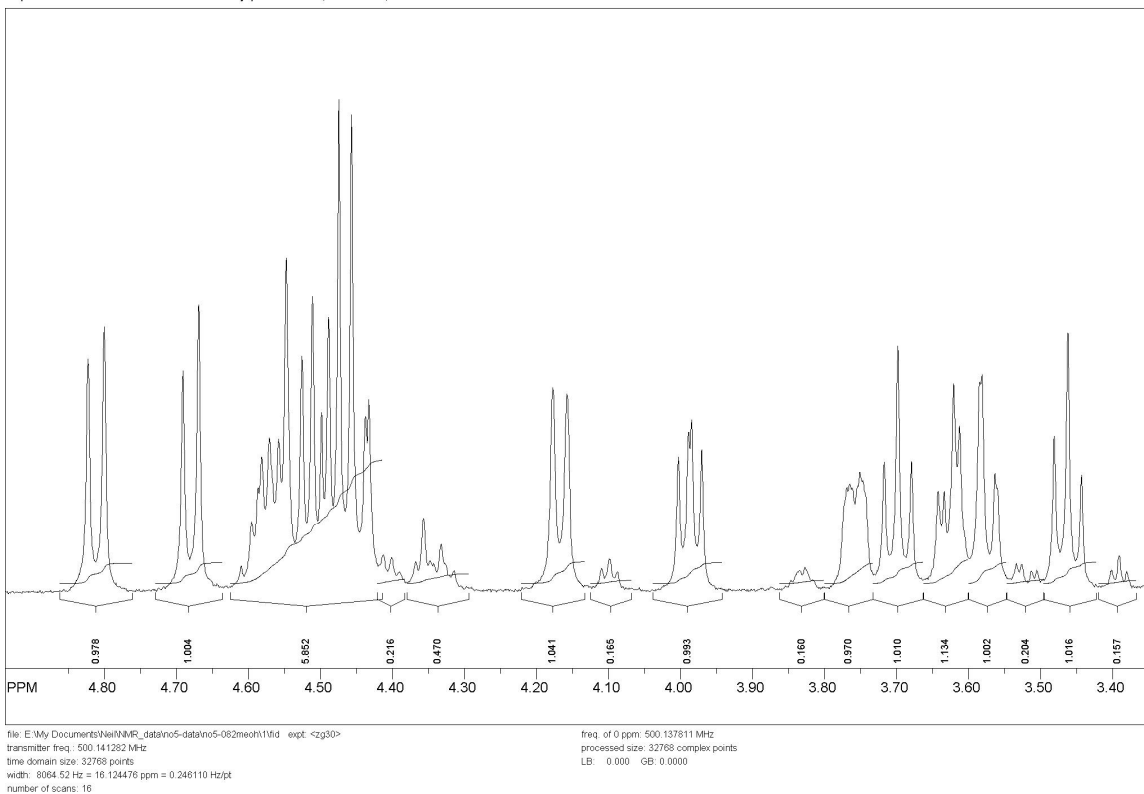
SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl_3



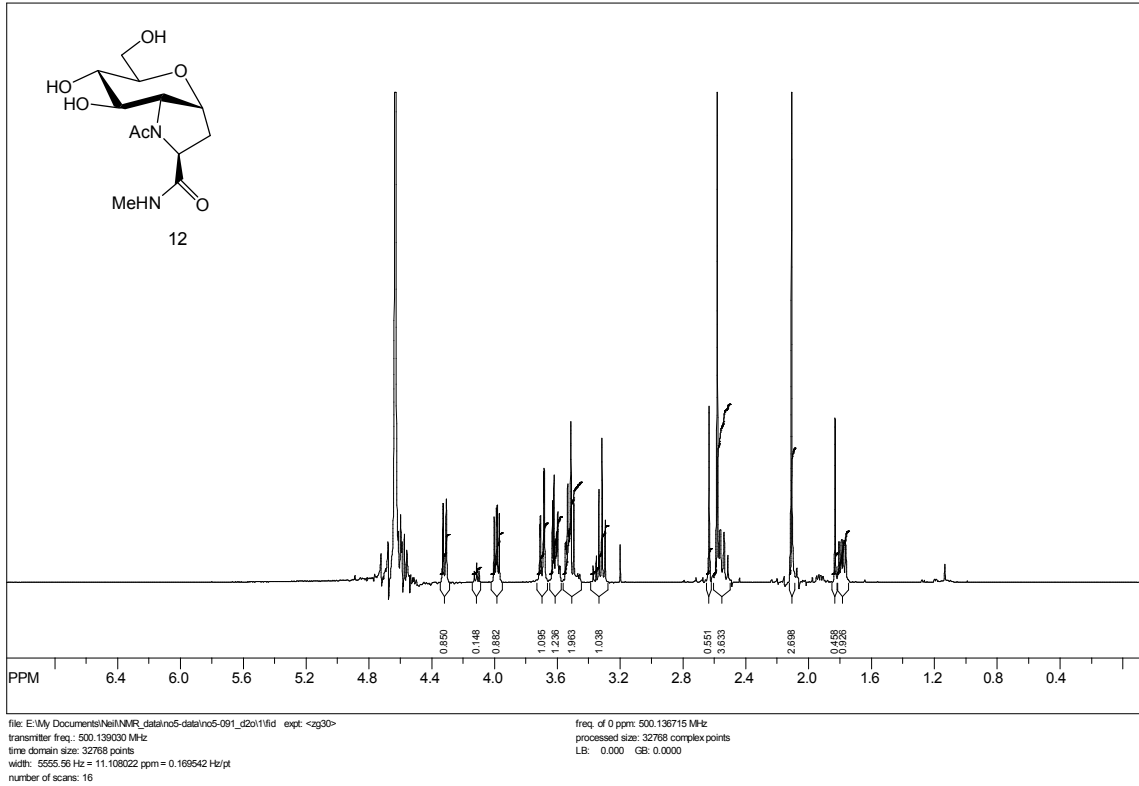
SpinWorks 2.5: standard 1-H survey parameters, AMX500, dms0-d6



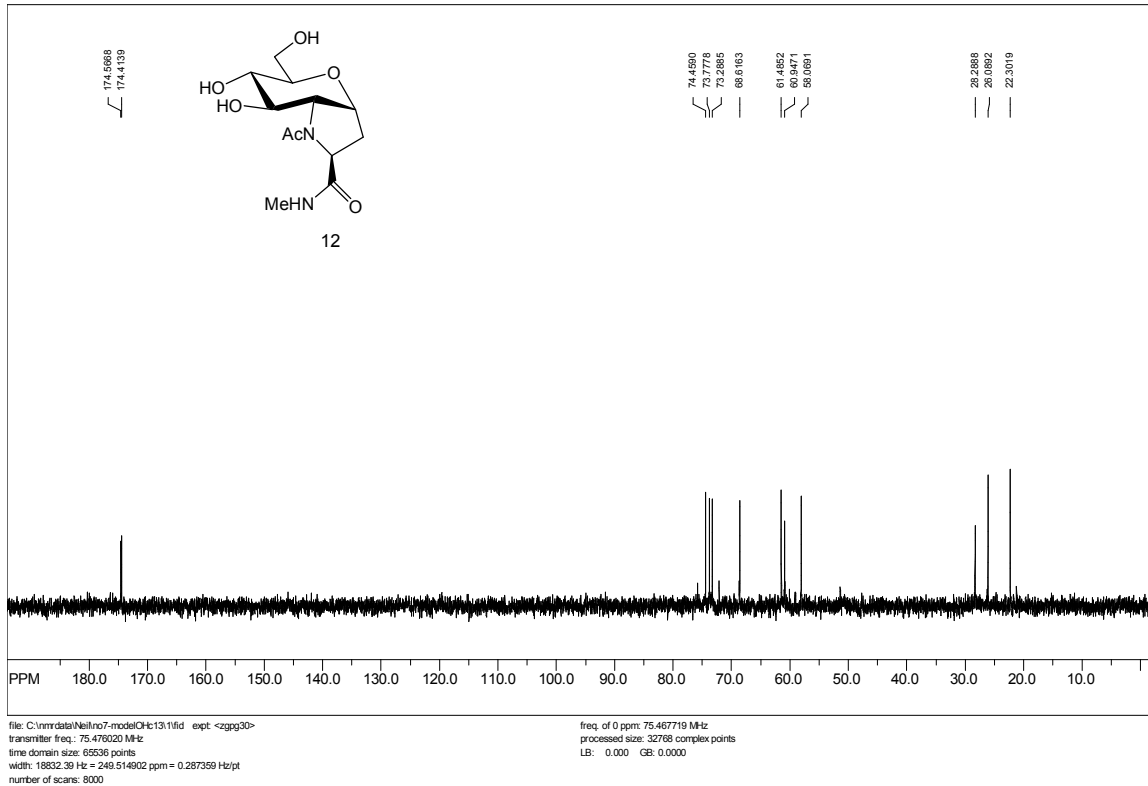
SpinWorks 2.5: standard 1-H survey parameters, AMX500, dms0-d6



SpinWorks 2.5: standard 1-H survey parameters, AMX500, D2O

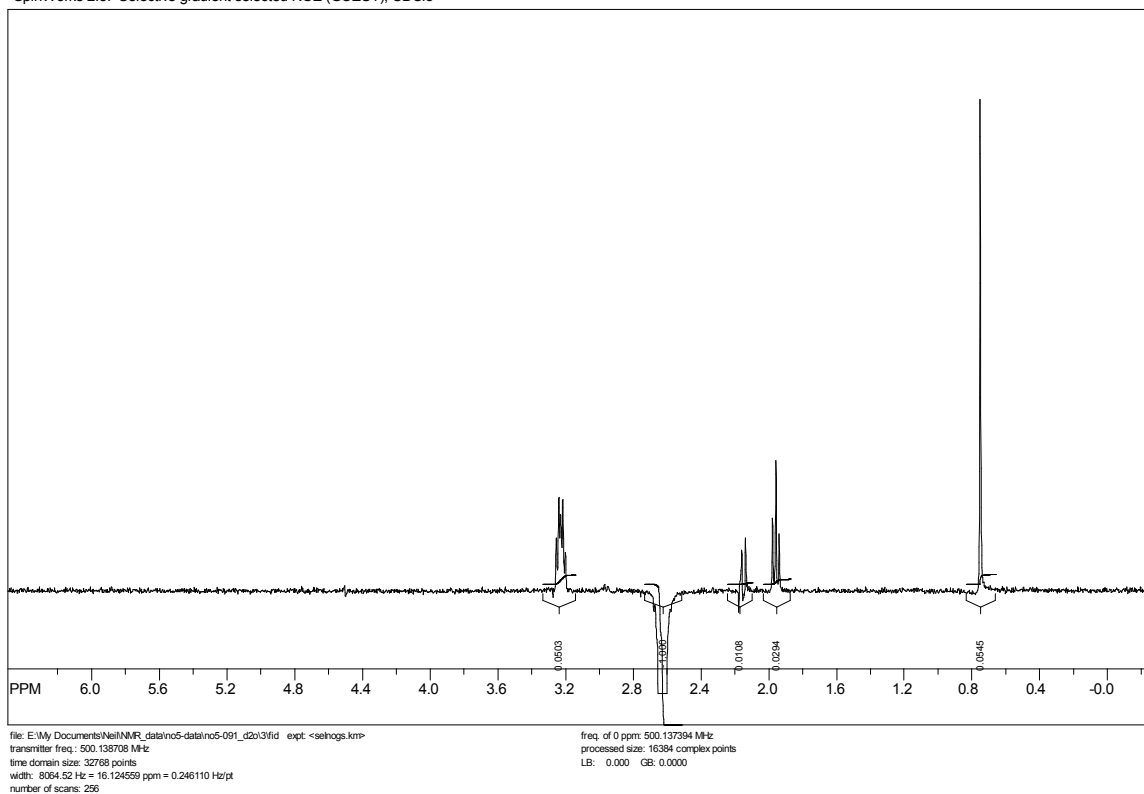


SpinWorks 2.5: C13CPD D2O u schweiz 1



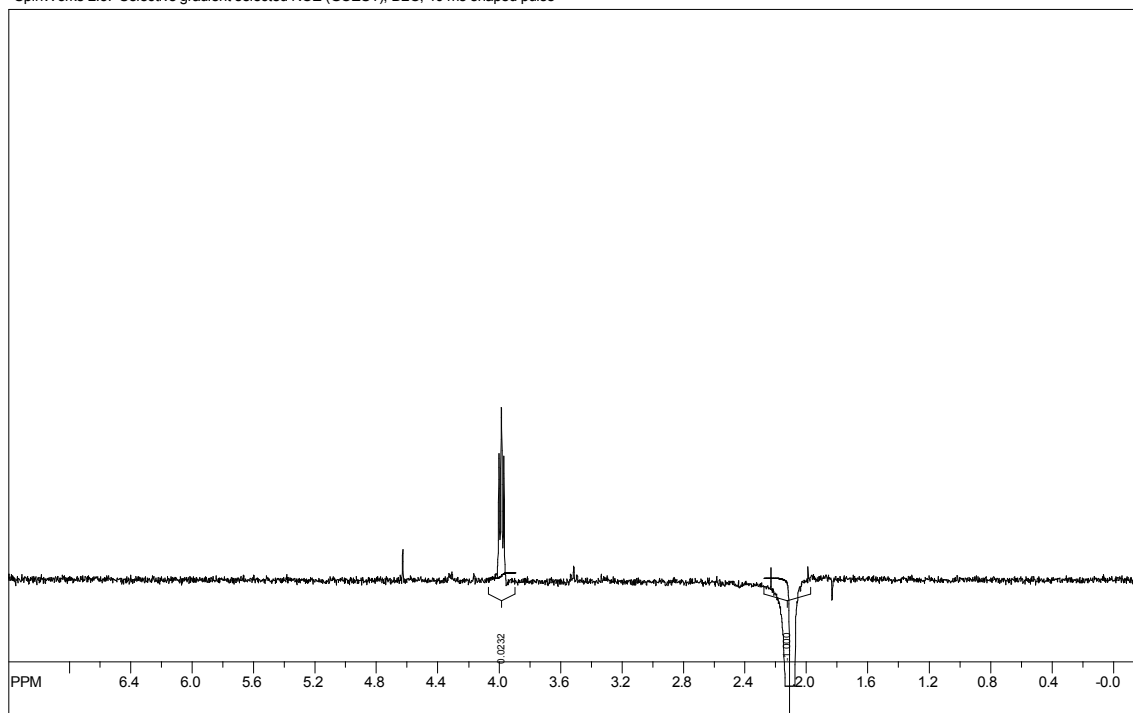
Subjection of the H₈ signal of the major conformer (3.99 ppm, 0.85 H) [H₈, 4.11 ppm, 0.15 H] to a one-dimensional GOESY experiment in D₂O showed inter-proton effect to the acetate methyl singlet –COCH₃ of the major conformer (2.11 ppm, 2.55 H) (5.5%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃



Subjection of the acetate methyl singlet $-\text{COCH}_3$ of the major conformer (2.11 ppm, 2.55 H) to a one-dimensional GOESY experiment in D_2O showed inter-proton effect to the H_8 of the major conformer (3.99 ppm, 0.85 H) (6.9%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), D_2O , 40 ms shaped pulse

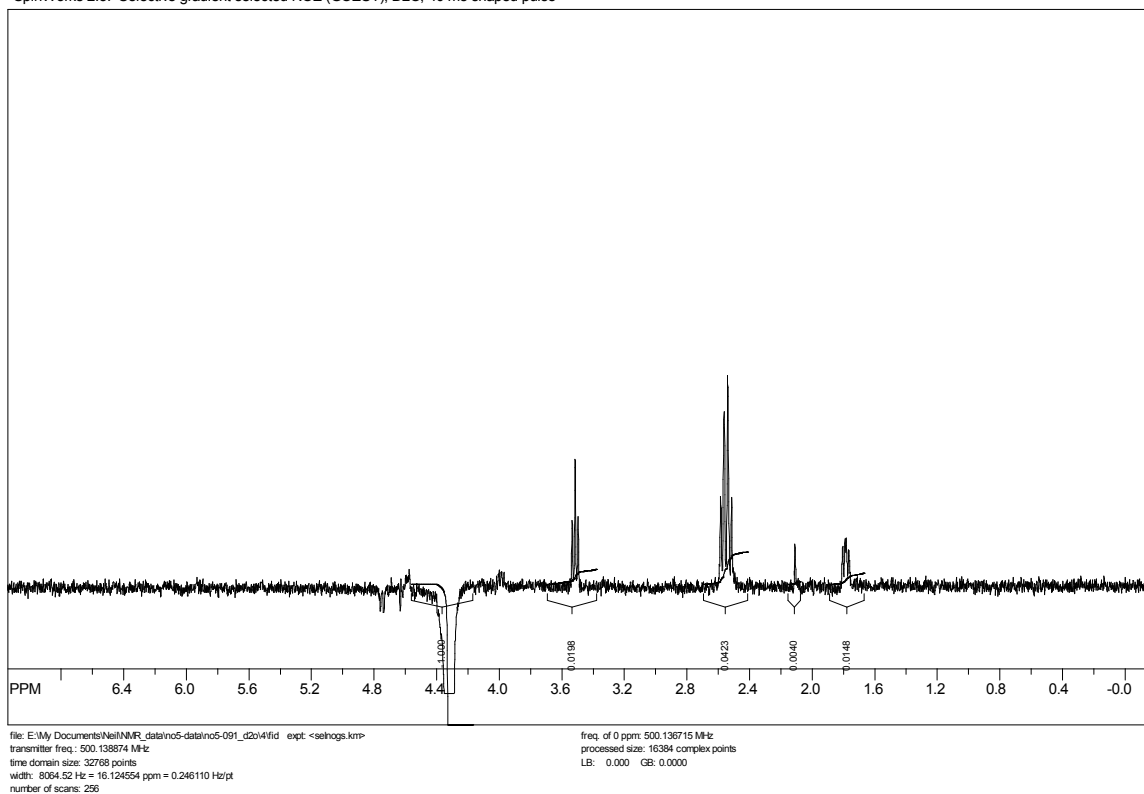


file: E:\My Documents\NMR_data\nc5-data\nc5-091_d2o5fid exp: -selhogs.kmp
transmitter freq.: 500.137770 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124589 ppm = 0.246110 Hz/pt
number of scans: 256

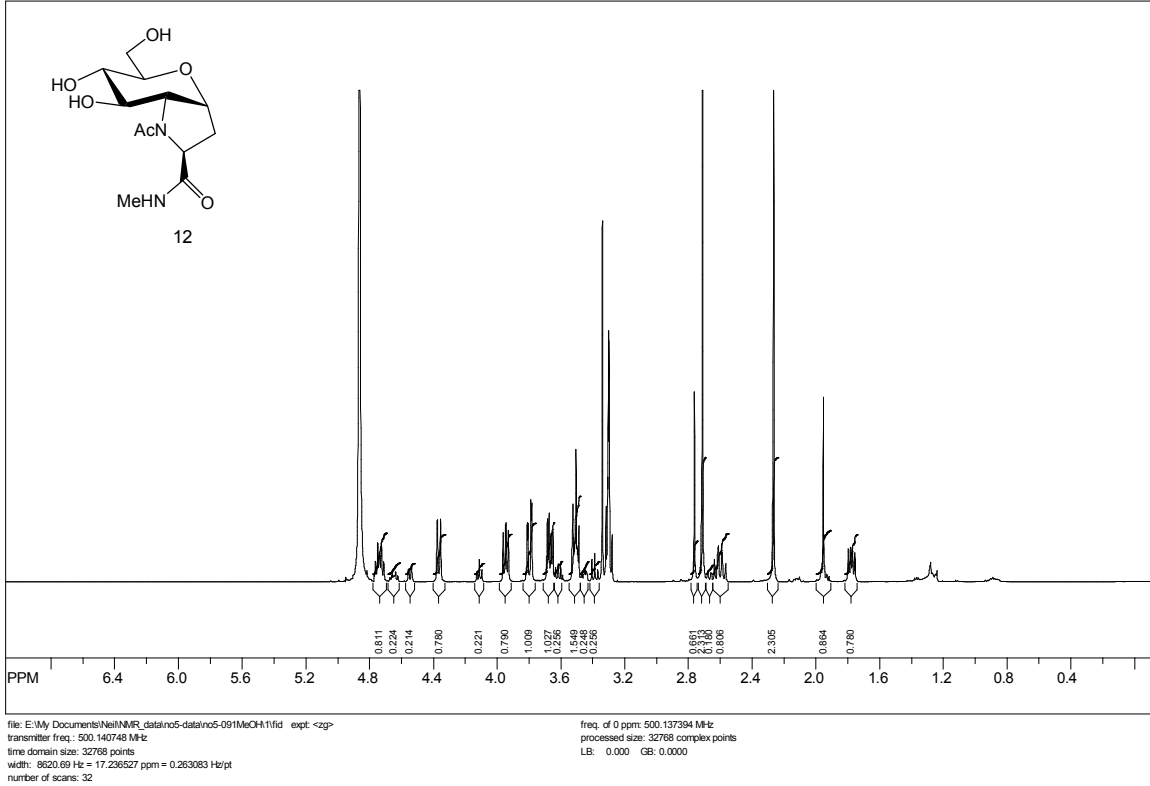
freq. of 0 ppm: 500.136715 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

Subjection of the H₂ signal of the major conformer (4.32 ppm, 0.85 H) to a one-dimensional GOESY experiment in D₂O showed a small inter-proton effect to the acetate methyl singlet –COCH₃ of the major conformer (2.11 ppm, 2.55 H) (0.4%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), D2O, 40 ms shaped pulse

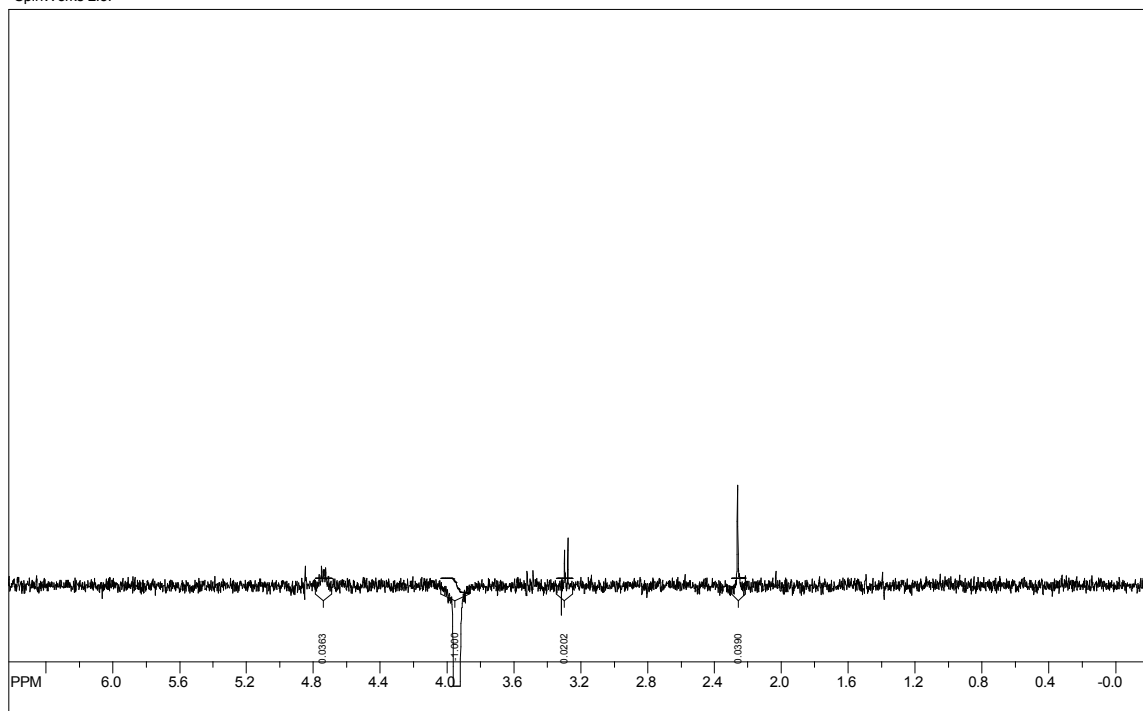


SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



Subjection of the H₈ signal of the major conformer (3.95 ppm, 0.78 H) [H₈, 4.11 ppm, 0.22 H] to a one-dimensional GOESY experiment in CD₃OD showed inter-proton effect to the acetate methyl singlet –COCH₃ of the major conformer (2.27 ppm, 2.34 H) (3.9%)

SpinWorks 2.5:

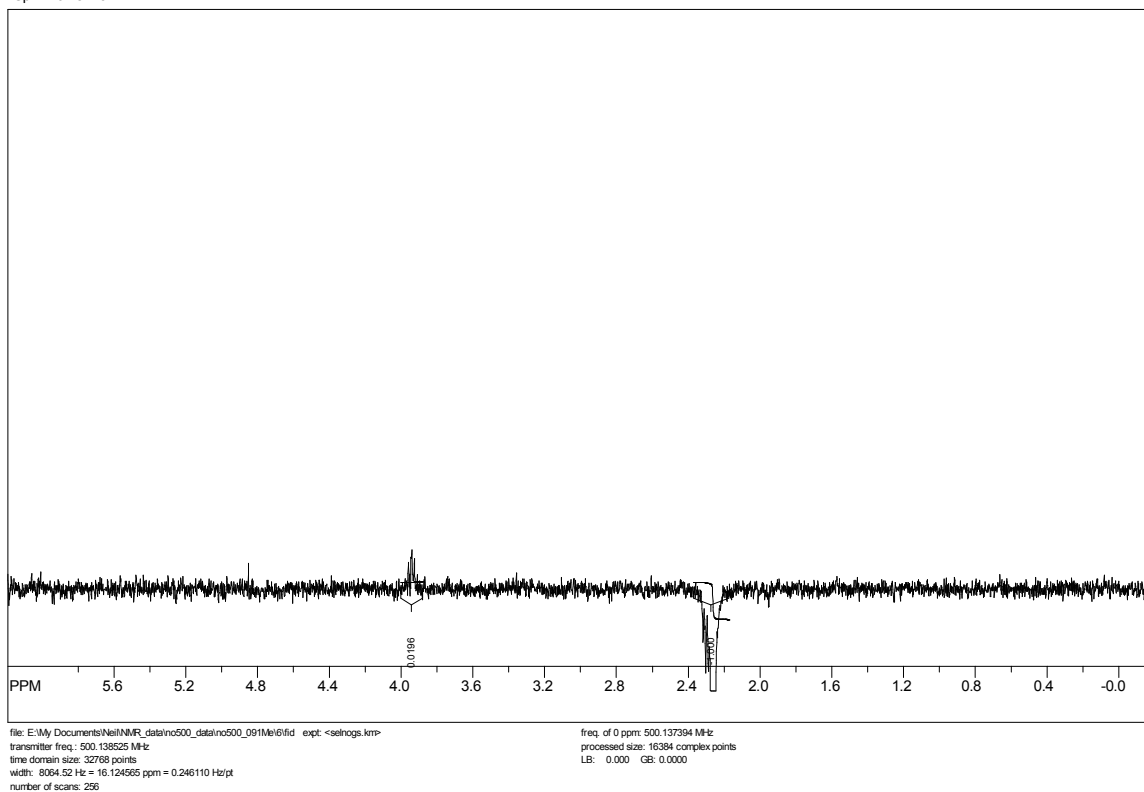


file: E:\My Documents\NMR_data\nc500_data\nc500_091MeV4fid exp: -selhogs.kmp
transmitter freq.: 500.139366 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124538 ppm = 0.246110 Hz/pt
number of scans: 256

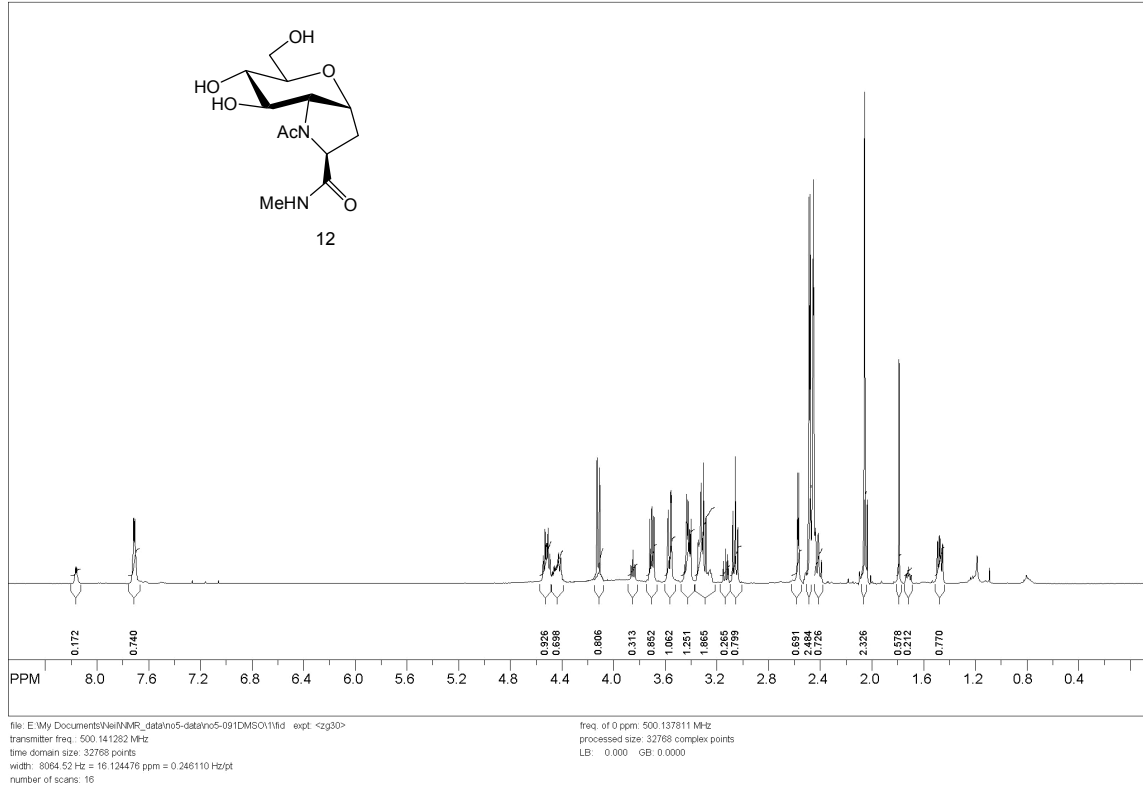
freq. of 0 ppm: 500.137394 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

Subjection of the acetate methyl singlet $-\text{COCH}_3$ of the major conformer (2.27 ppm, 2.34 H) to a one-dimensional GOESY experiment showed inter-proton effect to the H_8 signal of the major conformer (3.95 ppm, 0.78 H) (6.0%)

SpinWorks 2.5:

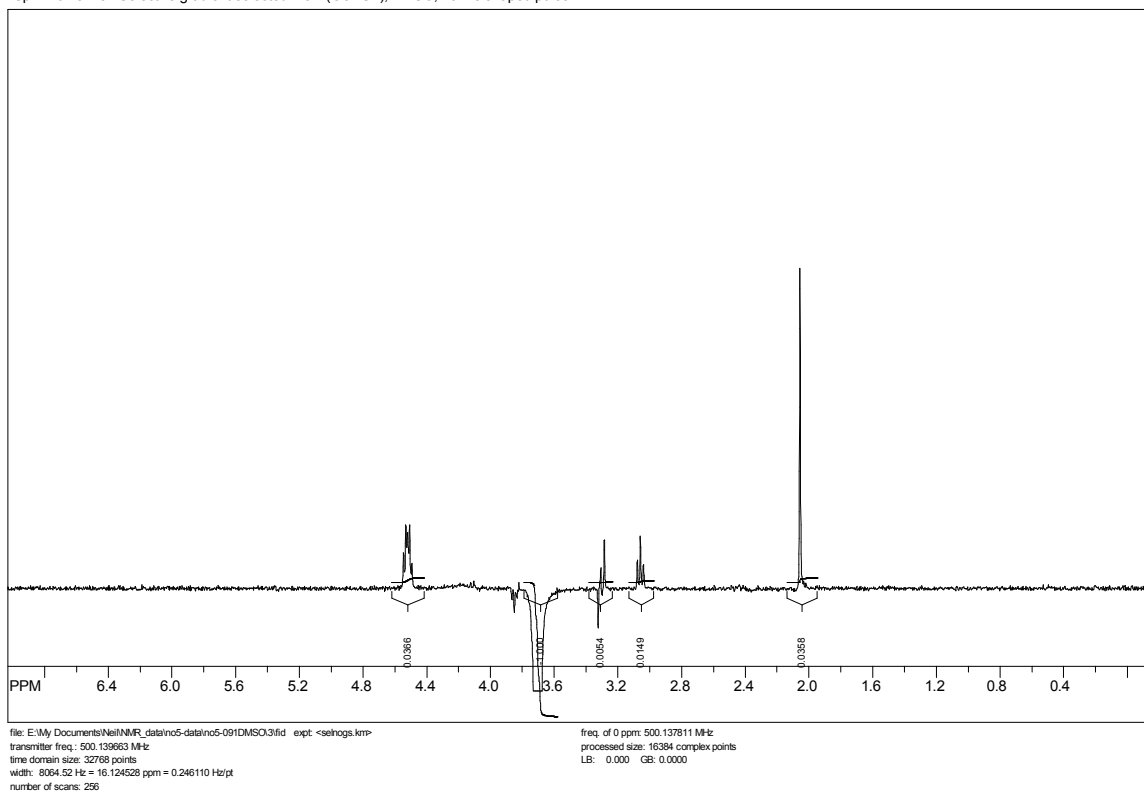


SpinWorks 2.5: standard 1-H survey parameters, AMX500, dms0-d6



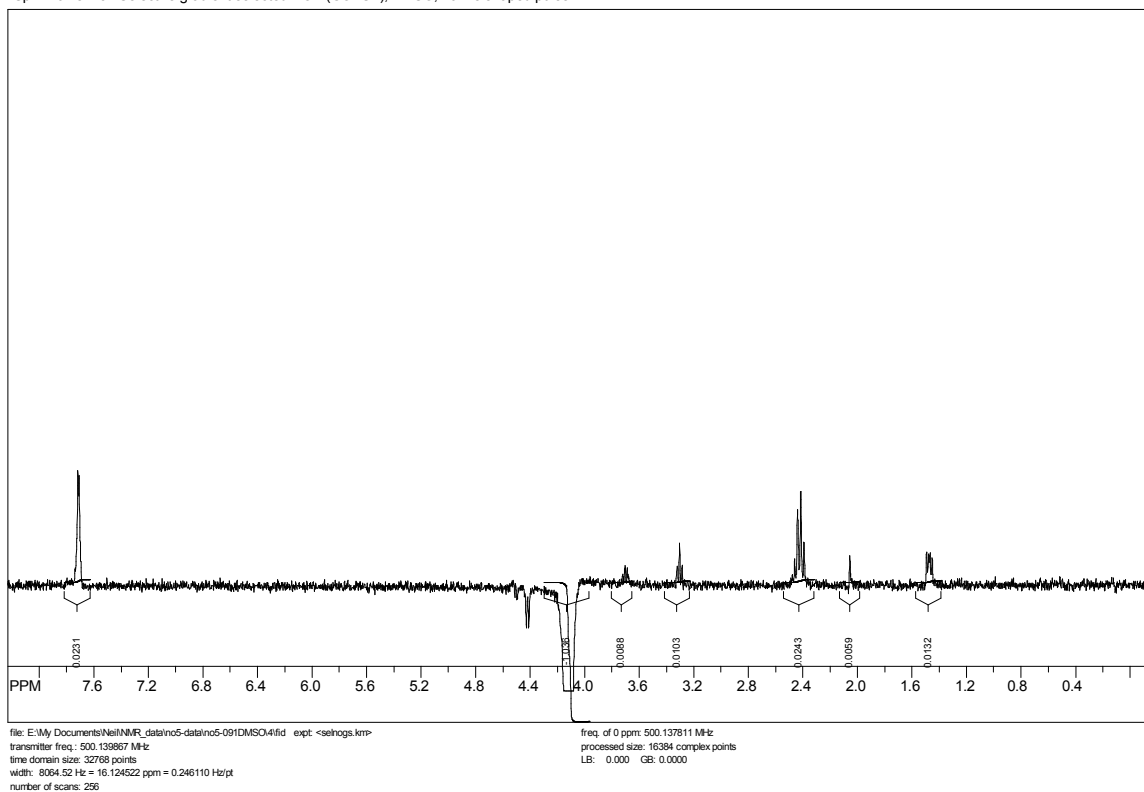
Subjection of the H₈ signal of the major conformer (3.70 ppm, 0.8 H) [H₈, 3.85 ppm, 0.2 H] to a one-dimensional GOESY experiment in DMSO-D₆ showed inter-proton effect to the acetate methyl singlet –COCH₃ of the major conformer (2.06 ppm, 2.4 H) (3.5%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), DMSO, 40 ms shaped pulse

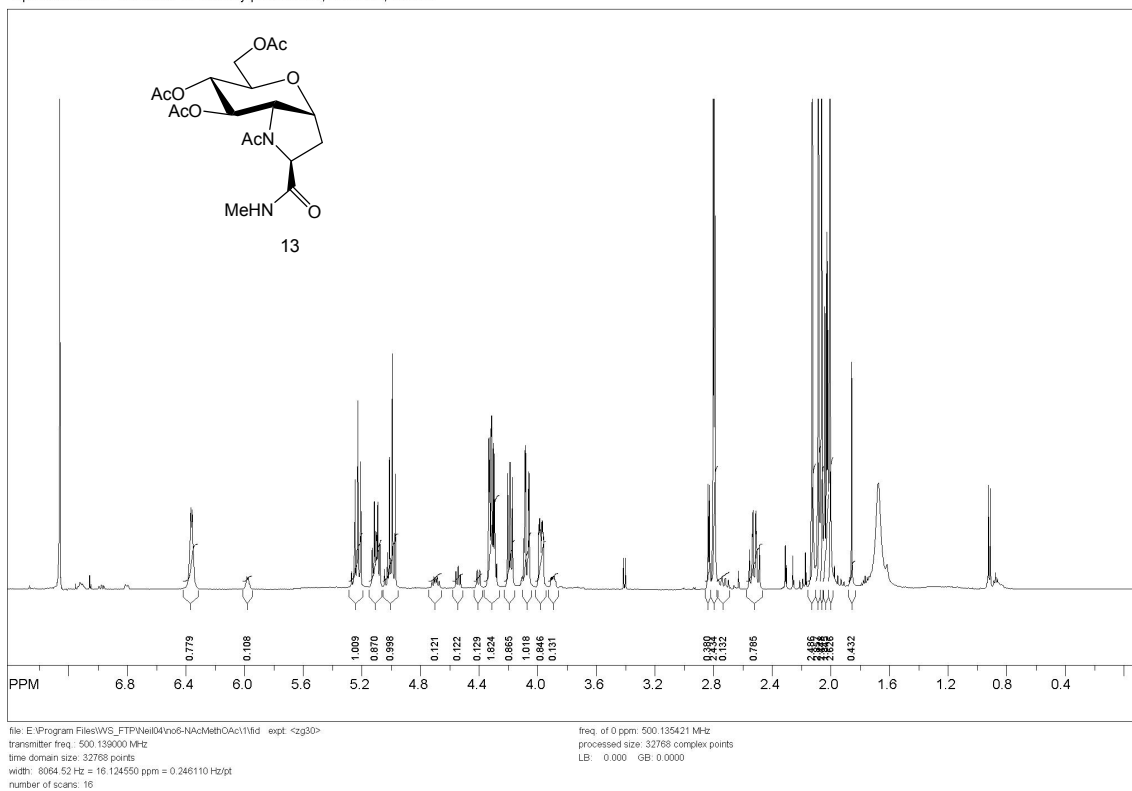


Subjection of the H₂ signal of the major conformer (4.12 ppm, 0.8 H) to a one-dimensional GOESY experiment showed a much smaller inter-proton effect to the acetate methyl singlet –COCH₃ of the major conformer (2.06 ppm, 2.4 H) (0.6%)

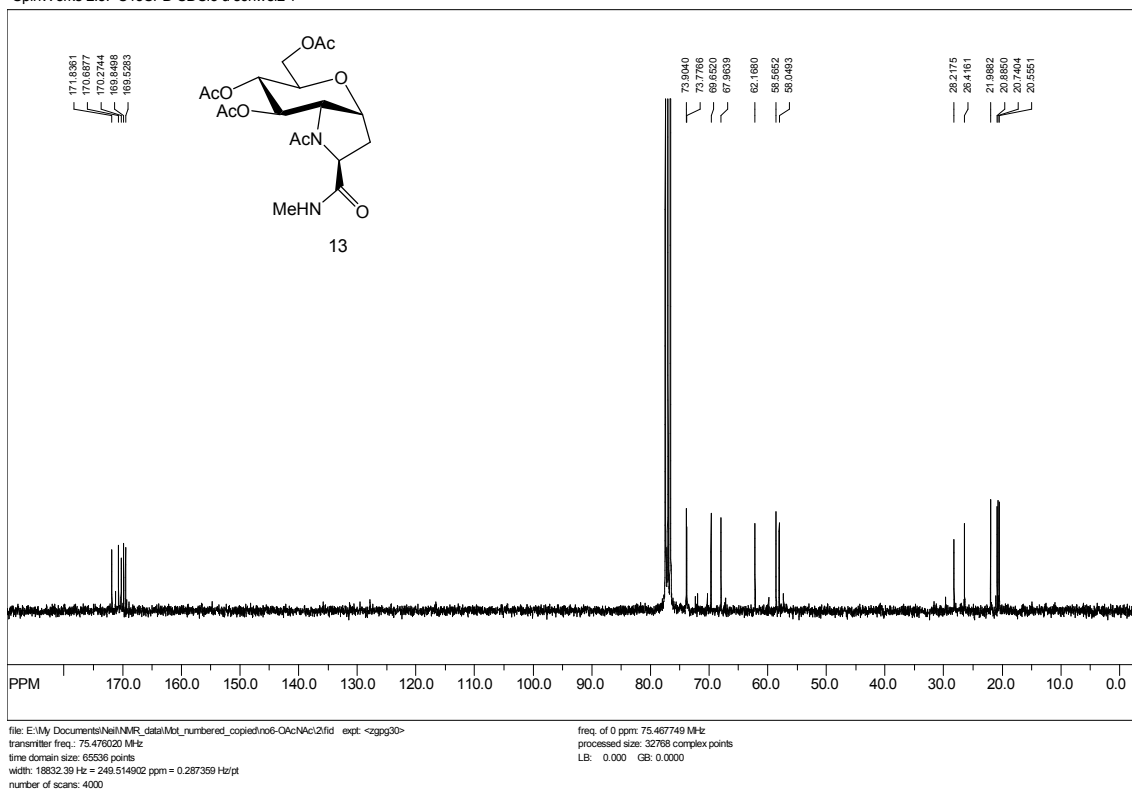
SpinWorks 2.5: Selective gradient selected NOE (GOESY), DMSO, 40 ms shaped pulse



SpinWorks 2.5: standard 1-H survey parameters, AMX500, CDCl3

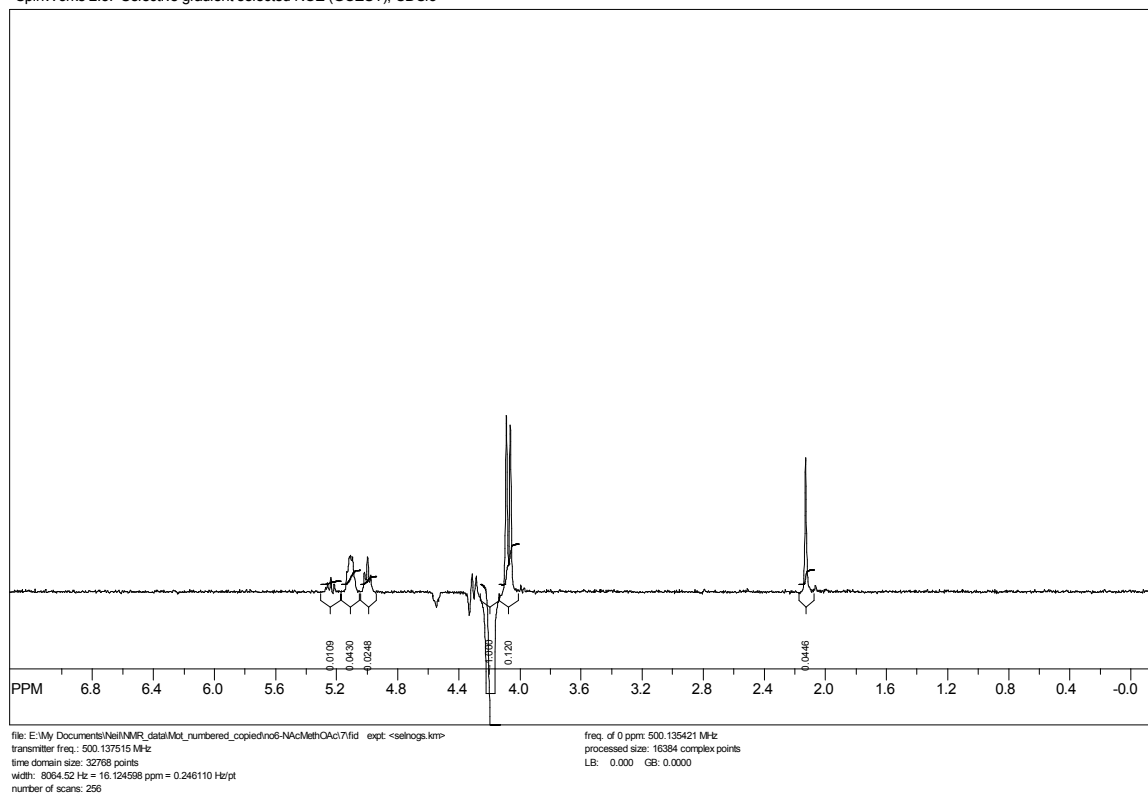


SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



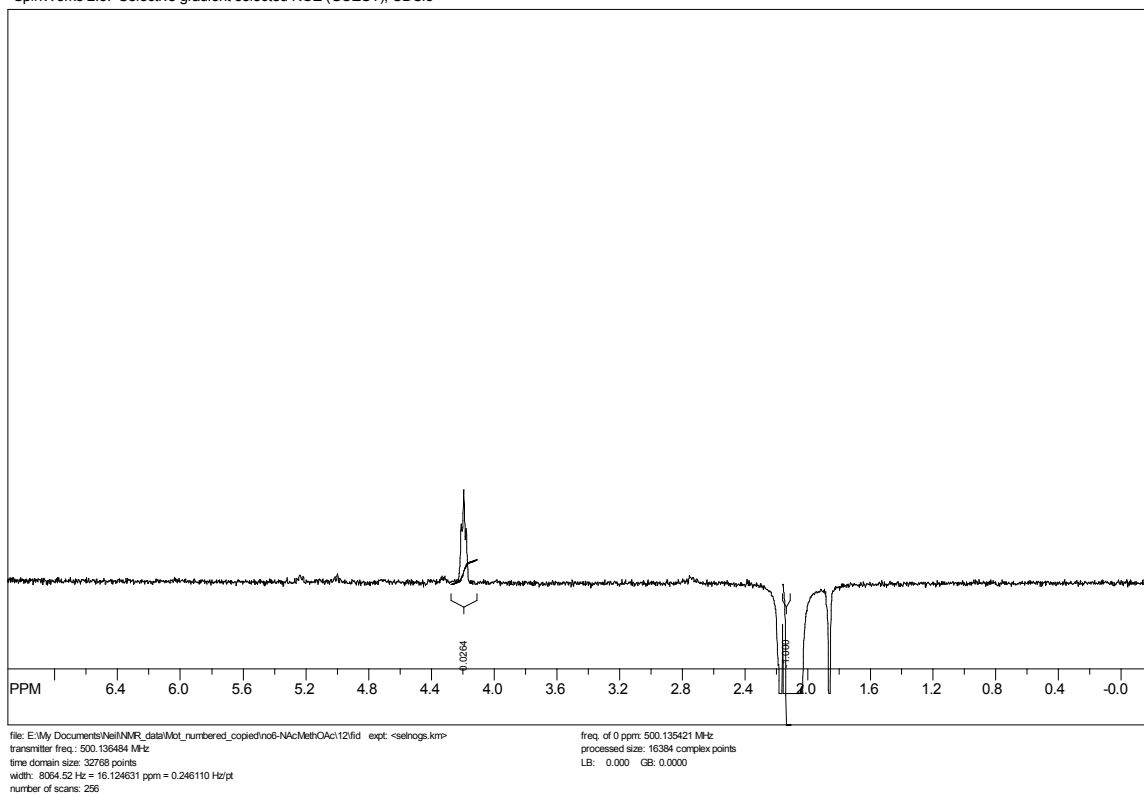
Subjection of the H₈ signal of the major conformer (4.19 ppm, 0.85 H) [H₈, 4.54 ppm, 0.15 H] to a one-dimensional GOESY experiment in CDCl₃ showed inter-proton effect to the acetate methyl singlet –COCH₃ of the major conformer (2.13 ppm, 2.55 H) (4.5%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃



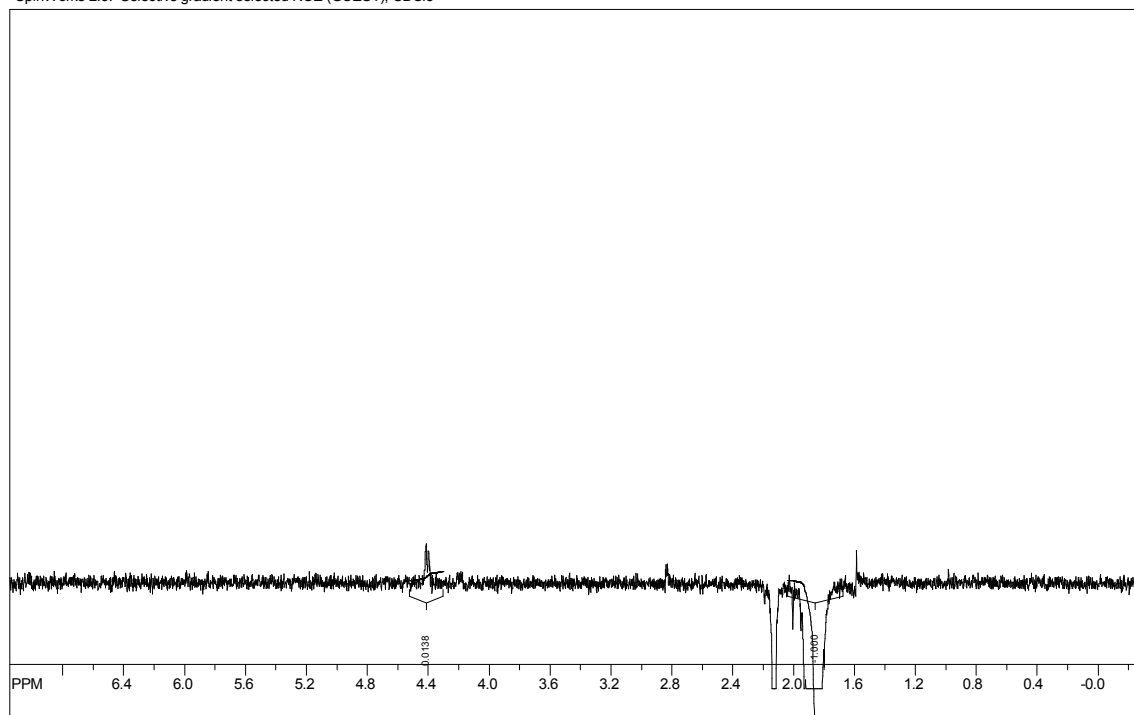
Subjection of the acetate methyl singlet $-\text{COCH}_3$ of the major conformer (2.13 ppm, 2.55 H) to a one-dimensional GOESY experiment in CDCl_3 showed inter-proton effect to the H_8 signal of the major conformer (4.19 ppm, 0.85 H) (7.9%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl_3



Subjection of the acetate methyl singlet $-\text{COCH}_3$ of the minor conformer [1.86 ppm, 0.45 H] to a one-dimensional GOESY experiment in CDCl_3 showed inter-proton effect to the H_2 signal of the minor conformer [4.41 ppm, 0.15 H] (4.2%)

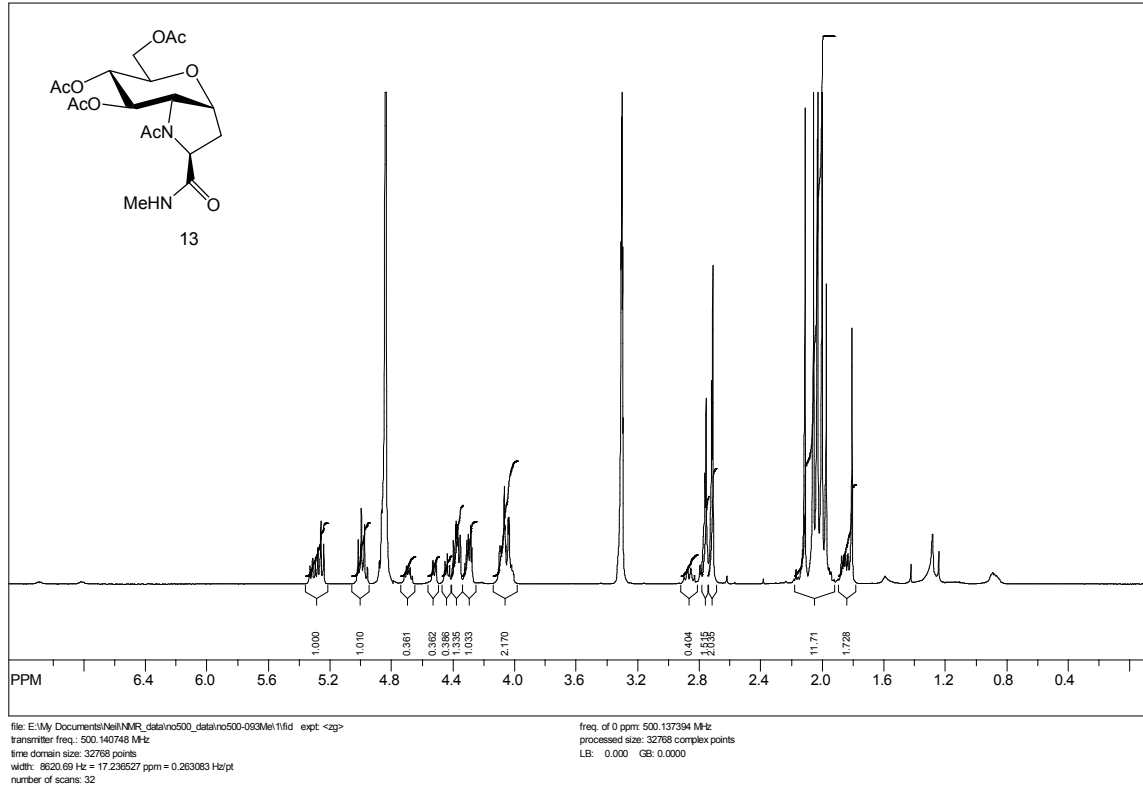
SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl_3



file: E:\My Documents\NMR_data\no500_data\no500-093CD12fid expt: <selhogs.km>
transmitter freq.: 500.136350 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124635 ppm = 0.246110 Hz/pt
number of scans: 1024

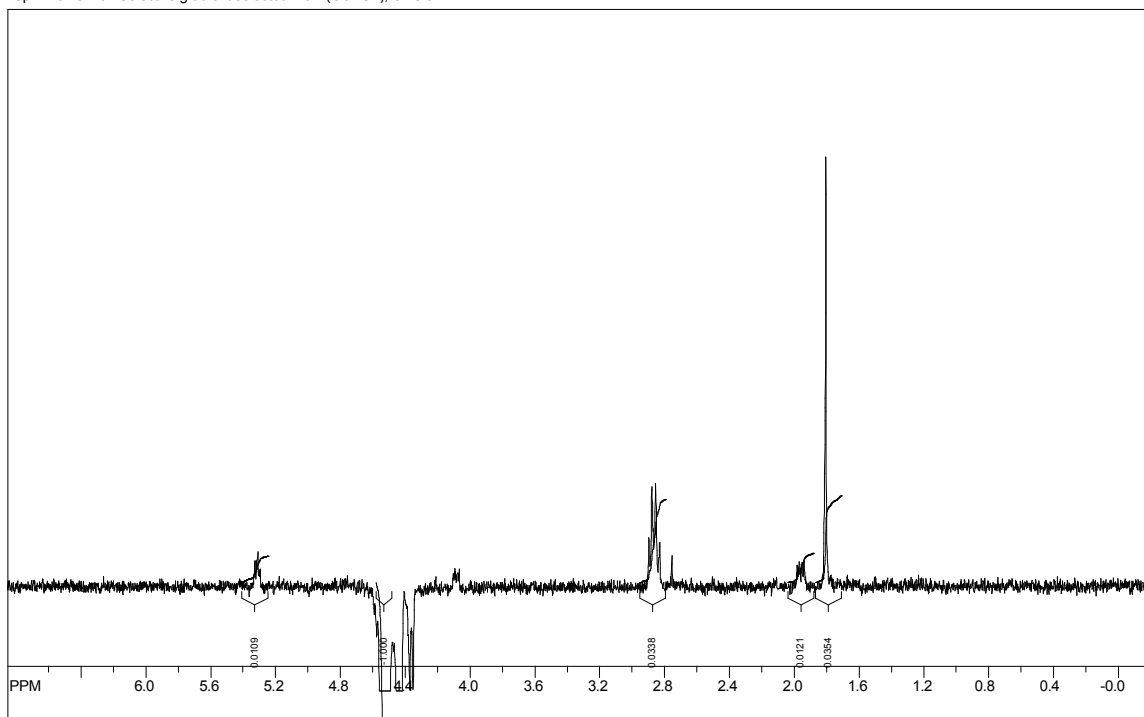
freq. of 0 ppm: 500.135421 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



Subjection of the H₂ signal of the minor conformer [4.52 ppm, 0.33 H] to a one-dimensional GOESY experiment in CD₃OD showed inter-proton effect to the acetate methyl singlet –COCH₃ of the minor conformer [1.81 ppm, 1 H] (3.5%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃

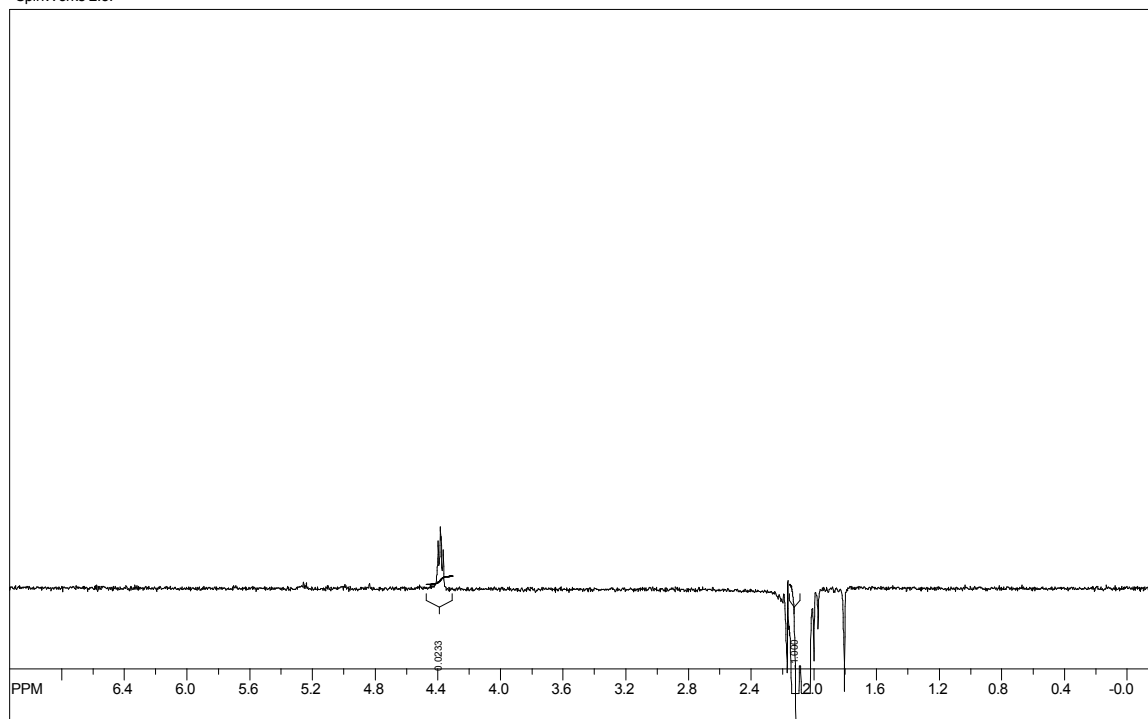


file: E:\My Documents\Neil\NMR_data\inc500_data\inc500-093Me\3\fid exp1 <selhogs.km>
transmitter freq.: 500.1379654 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124529 ppm = 0.246110 Hz/pt
number of scans: 1024

freq. of 0 ppm: 500.137994 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

Subjection of the acetate methyl singlet $-\text{COCH}_3$ of the major conformer (2.11 ppm, 2 H) to a one-dimensional GOESY experiment in CD_3OD showed inter-proton effect to the H_8 signal of the major conformer (4.38 ppm, 0.67 H) (6.9%)

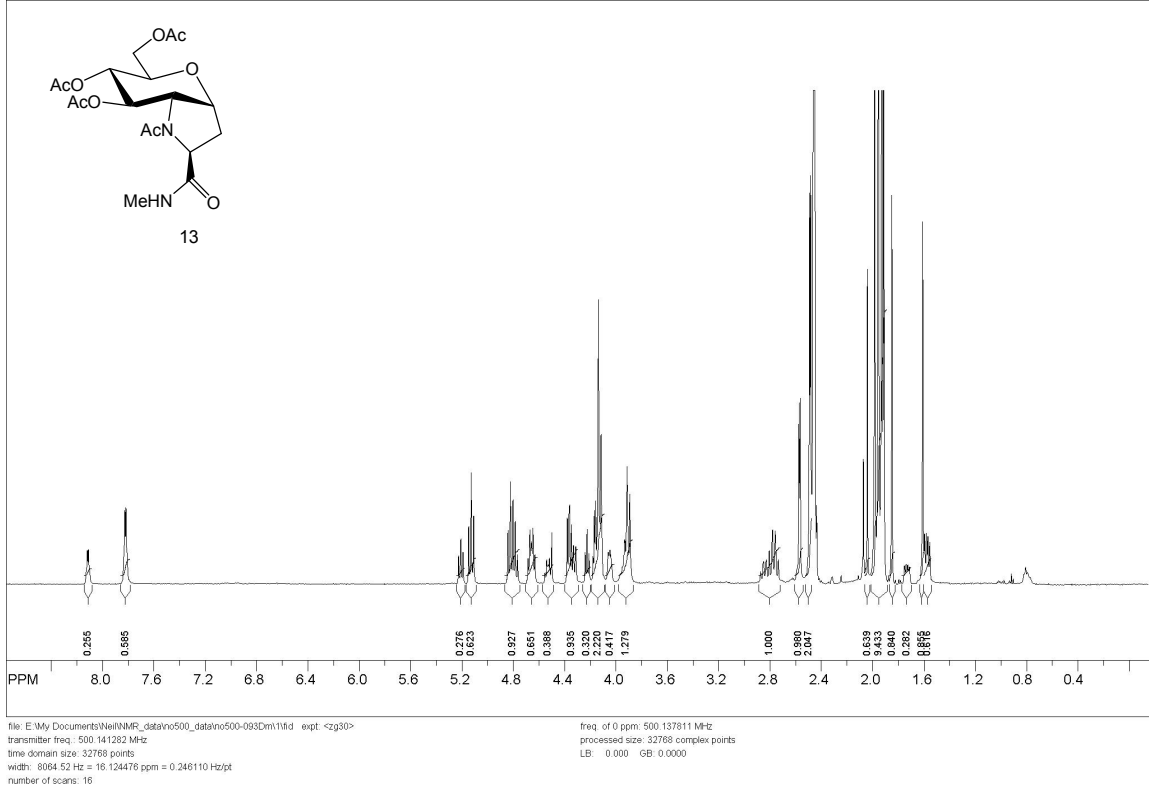
SpinWorks 2.5:



file: E:\My Documents\NMR\data\n500_data\n500-060Me151fid exp: <sehnogs.km>
transmitter freq.: 500.136450 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124567 ppm = 0.246110 Hz/pt
number of scans: 256

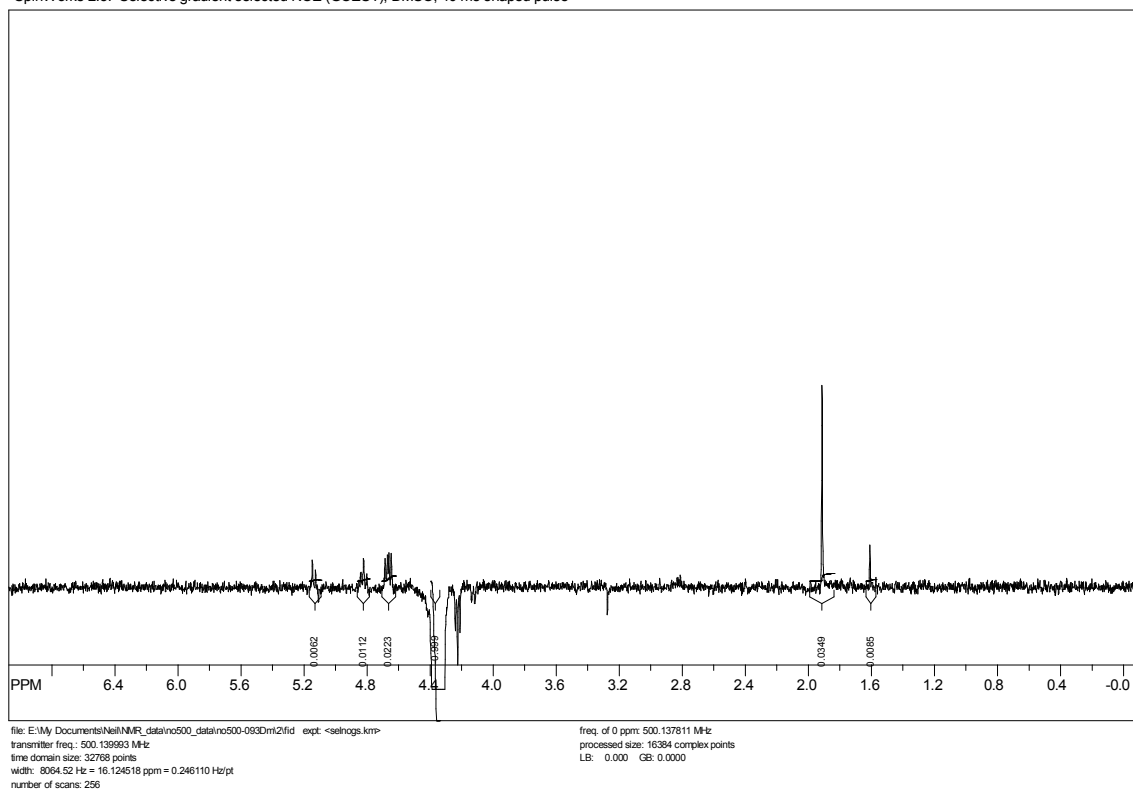
freq. of 0 ppm: 500.137394 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

SpinWorks 2.5: standard 1-H survey parameters, AMX500, dms0-d6



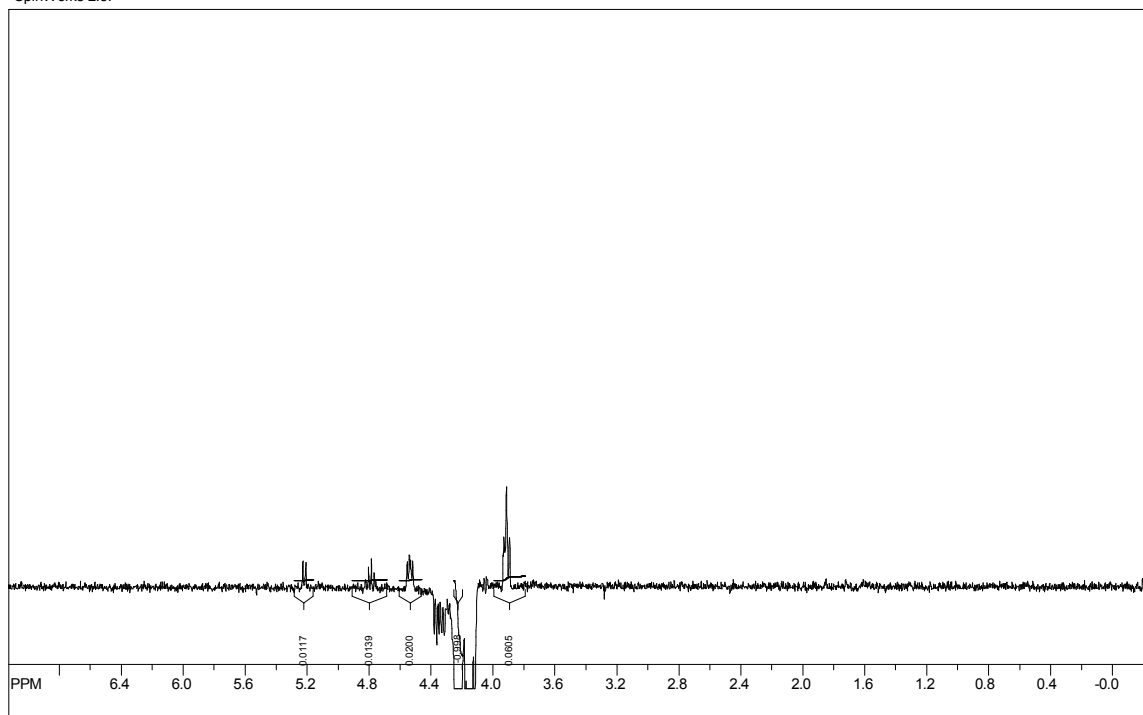
Subjection of the H₈ signal of the major conformer (4.36 ppm, 0.67 H) [H₈, 4.22 ppm, 0.33 H] to a one-dimensional GOESY experiment in DMSO-D₆ showed inter-proton effect to the acetate methyl singlet –COCH₃ of the major conformer (1.91 ppm, 2 H) (3.4%)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), DMSO, 40 ms shaped pulse



Subjection of the H₈ signal of the minor conformer [4.22 ppm, 0.33 H] to a one-dimensional GOESY experiment in DMSO-D₆ showed no nOe contact to the acetate methyl singlet –COCH₃ of the minor conformer [1.61 ppm, 1 H]

SpinWorks 2.5:

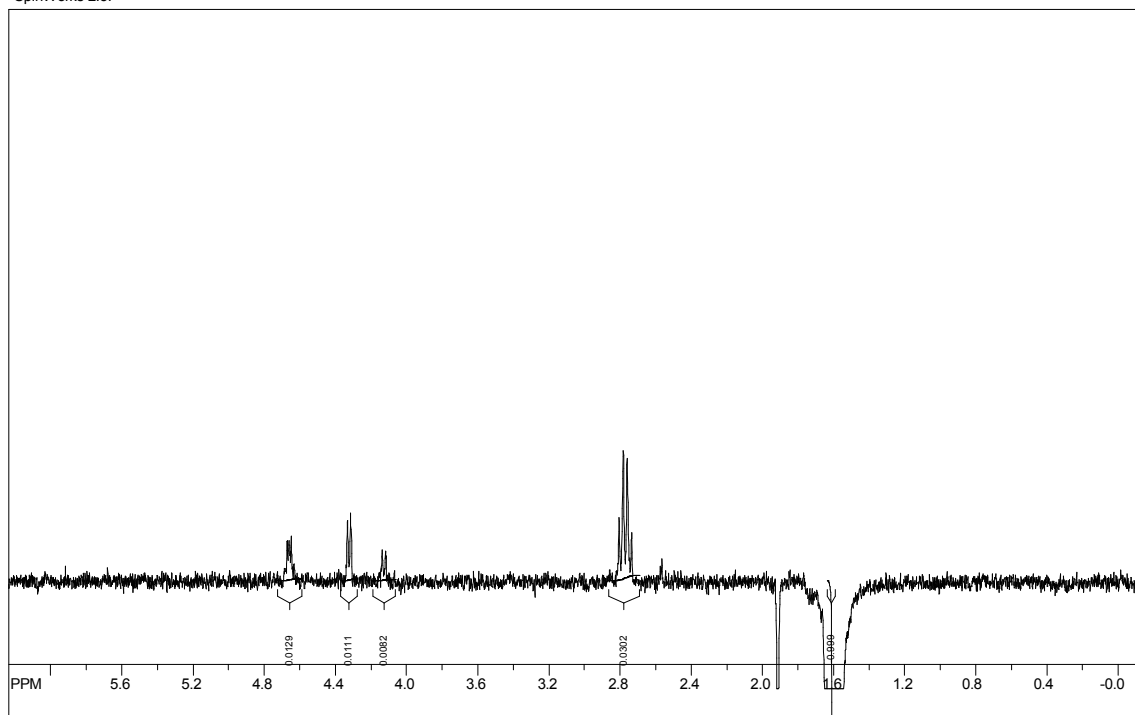


file: E:\My Documents\NMR\data\500_data\500-093Dm14.fid expt: <selhogs.km>
transmitter freq.: 500.139924 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124520 ppm = 0.246110 Hz/pt
number of scans: 1024

freq. of 0 ppm: 500.137811 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

Subjection of the acetate methyl singlet $-\text{COCH}_3$ of the minor conformer [1.61 ppm, 1 H] to a one-dimensional GOESY experiment in DMSO-D_6 showed inter-proton effect to the H_2 signal of the minor conformer [4.32 ppm, 0.33 H] (3.3%)

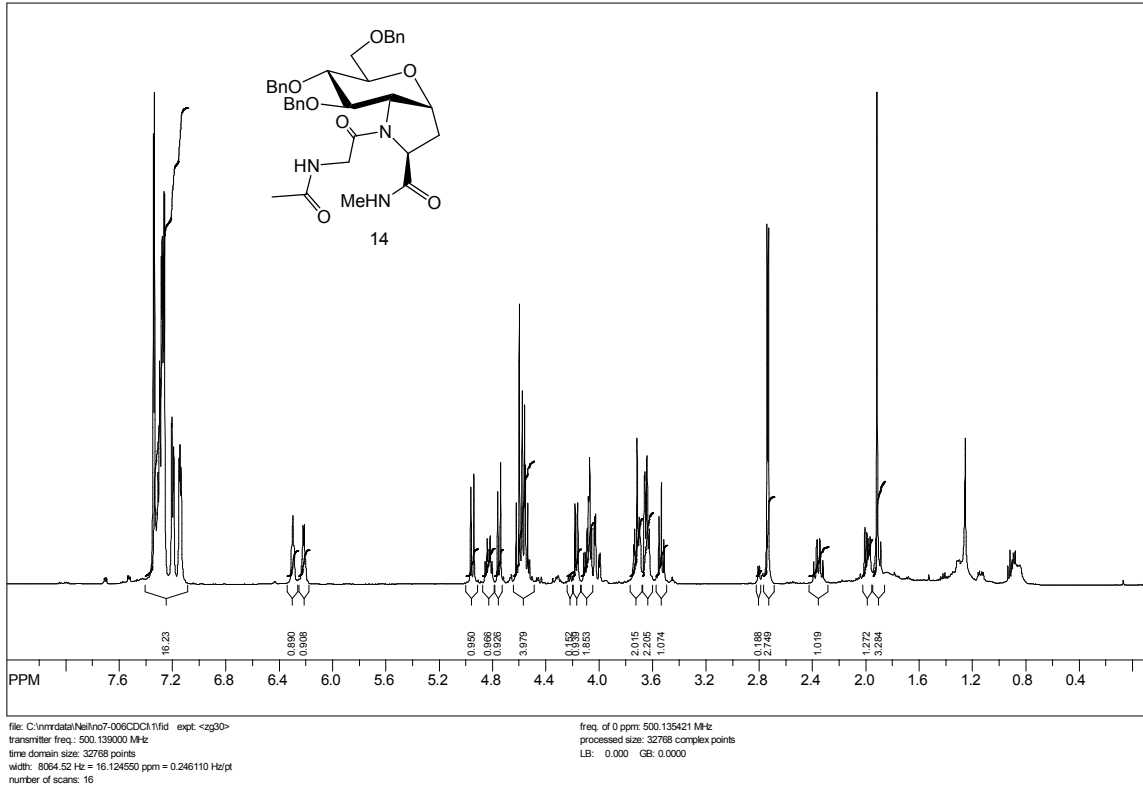
SpinWorks 2.5:



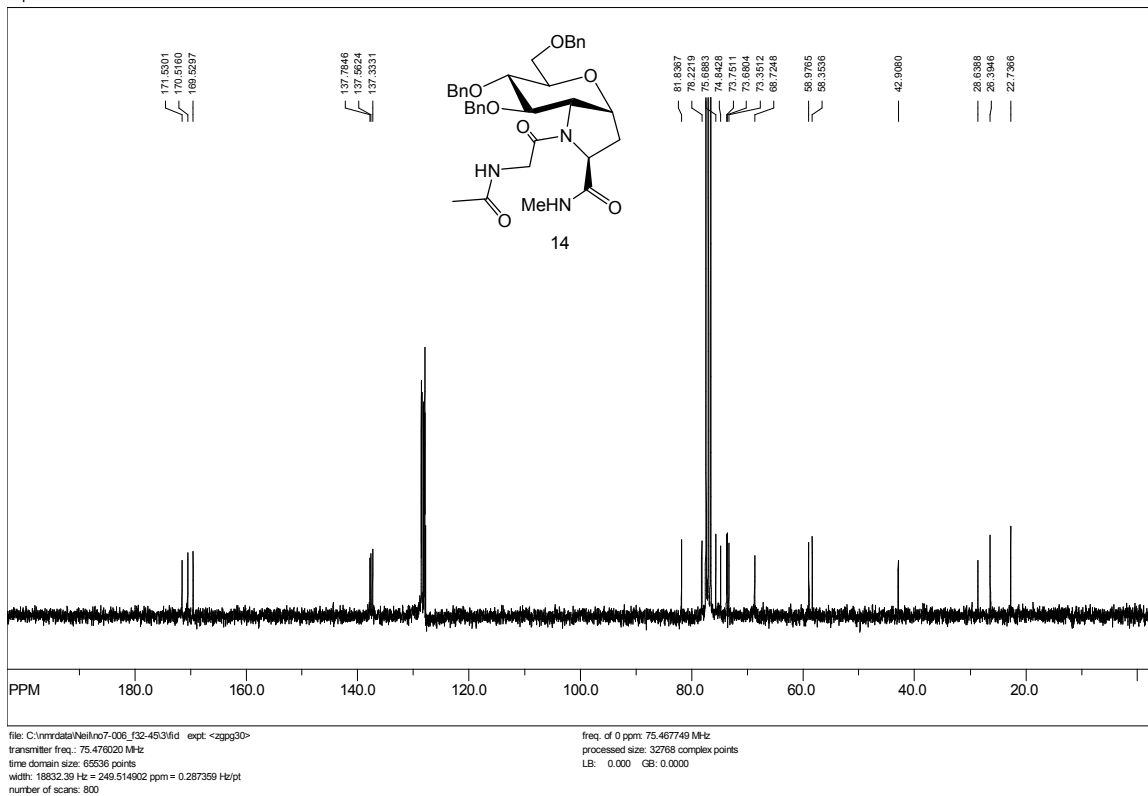
file: E:\My Documents\NMR_data\500_data\500-093Dml6\fid expt -selhogs.km
transmitter freq.: 500.138616 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124562 ppm = 0.246110 Hz/pt
number of scans: 1024

freq. of 0 ppm: 500.137811 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

SpinWorks 2.5: standard 1-H survey parameters, AMX500, CDCl3

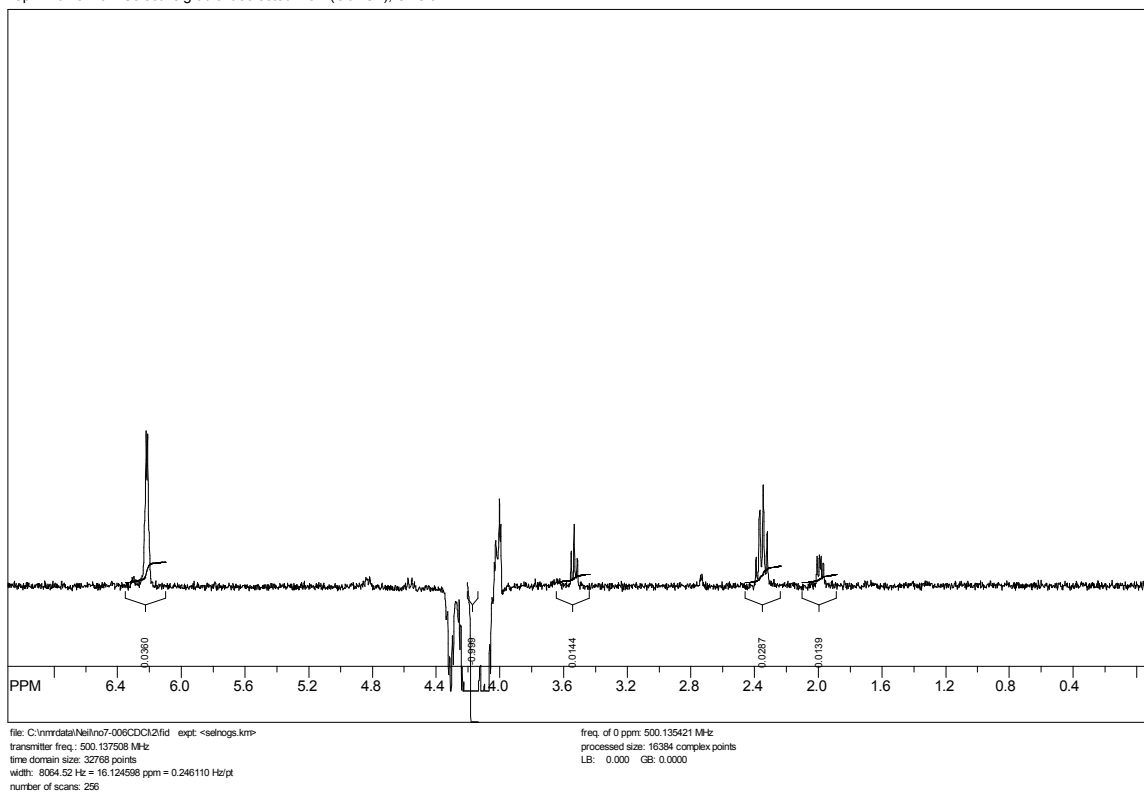


SpinWorks 2.5: C13CPD CDCl3 u schweiz 1

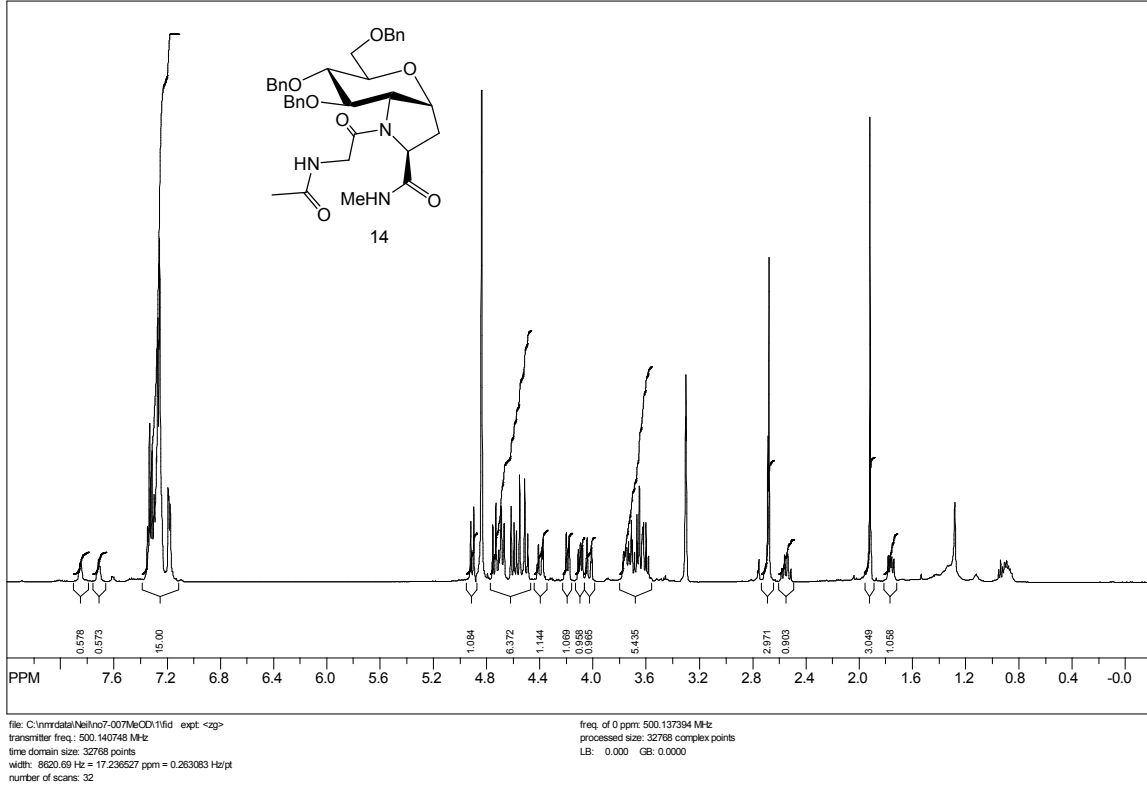


Subjection of the H₂ signal of the major conformer (4.17 ppm, 0.9 H) to a one-dimensional GOESY experiment in CDCl₃ showed no inter-proton effect either H_{α1}(Gly) (4.09 ppm, 0.9 H) or H_{α2}(Gly) (4.02 ppm, 0.9 H)

SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃

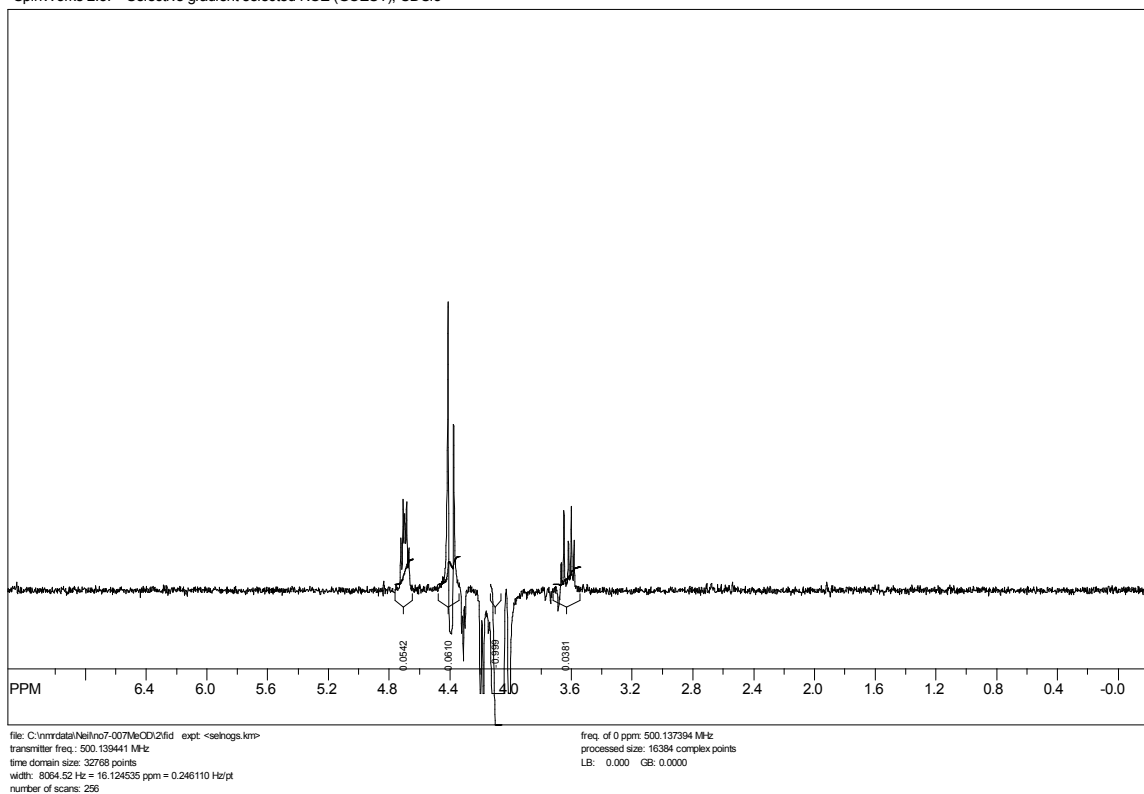


SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD

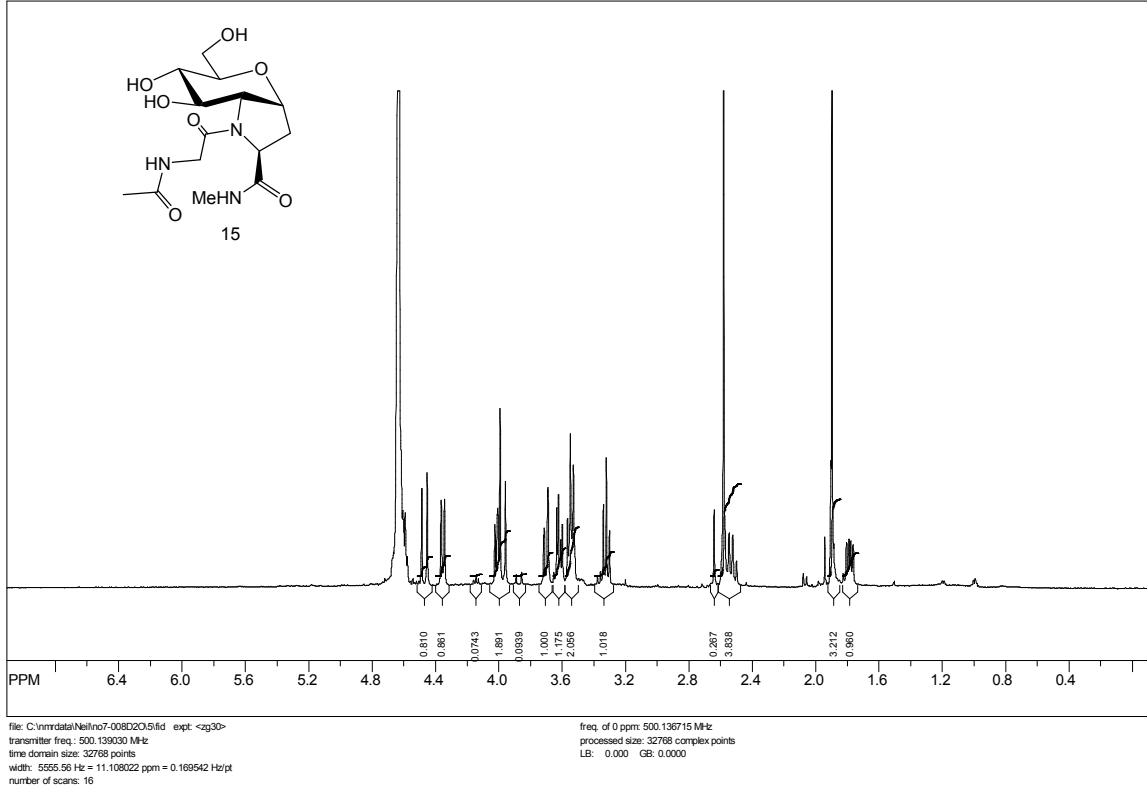


Subjection of the H₈ signal of the major conformer (4.10 ppm, 1 H) to a one-dimensional GOESY experiment in CD₃OD showed inter-proton effect to H_{α1}(Gly) of the major conformer (4.40 ppm, 1 H) (6.1%)

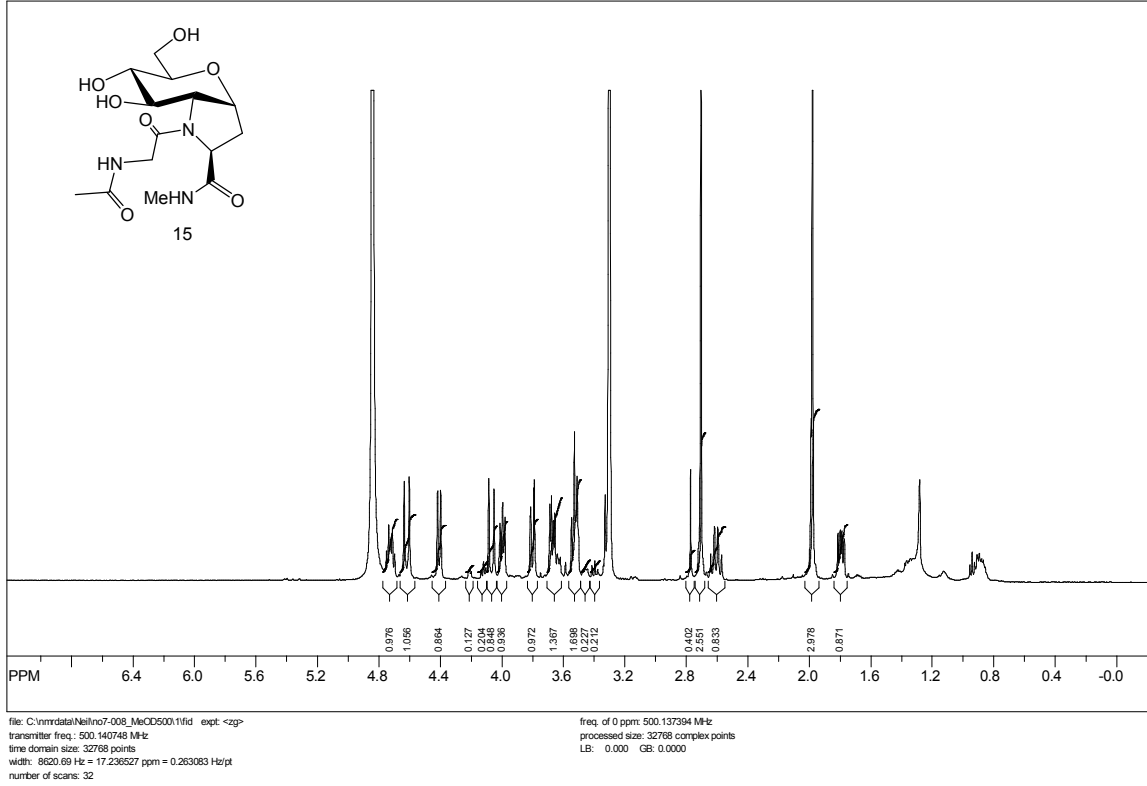
SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃



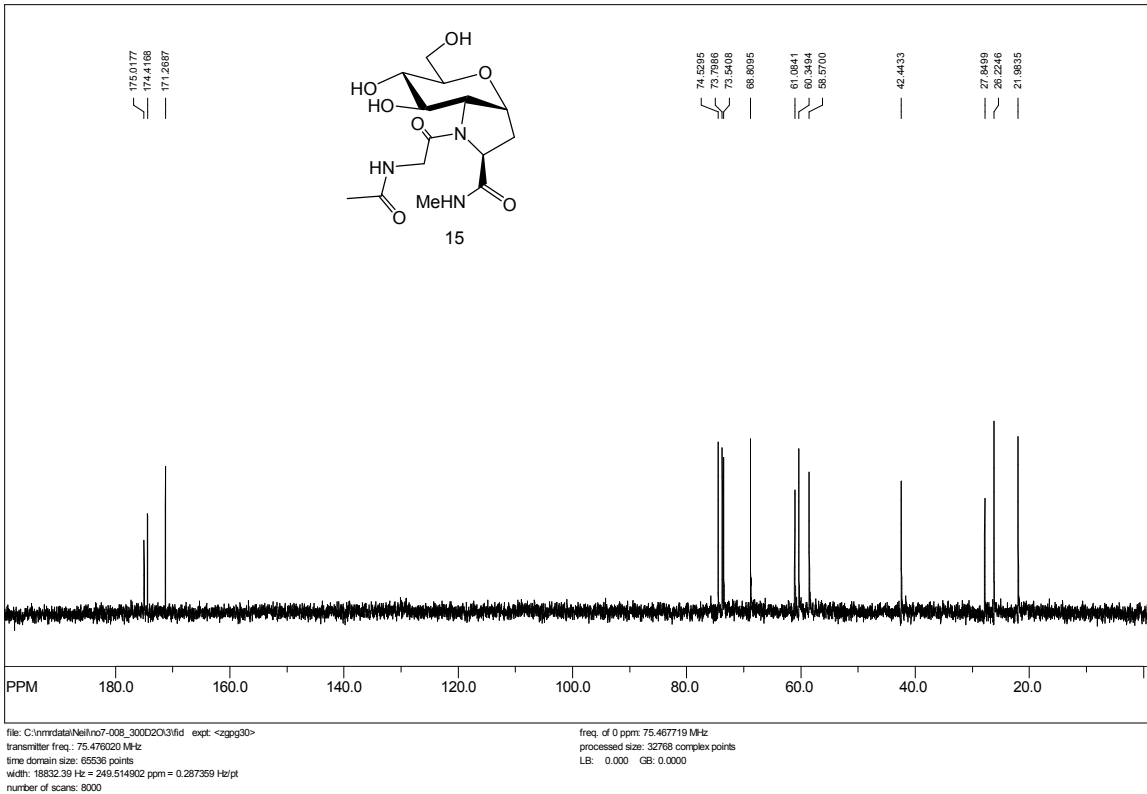
SpinWorks 2.5: standard 1-H survey parameters, AMX500, D2O



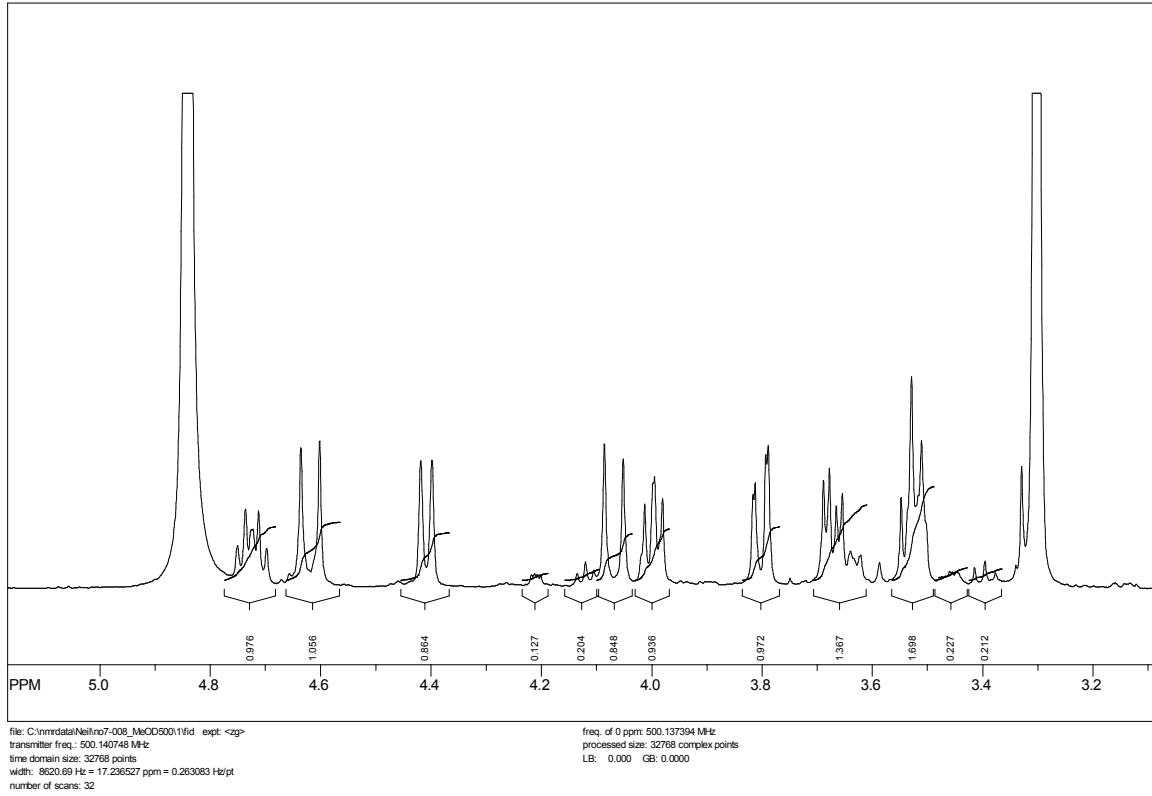
SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



SpinWorks 2.5: C13CPD D2O u schweiz 1

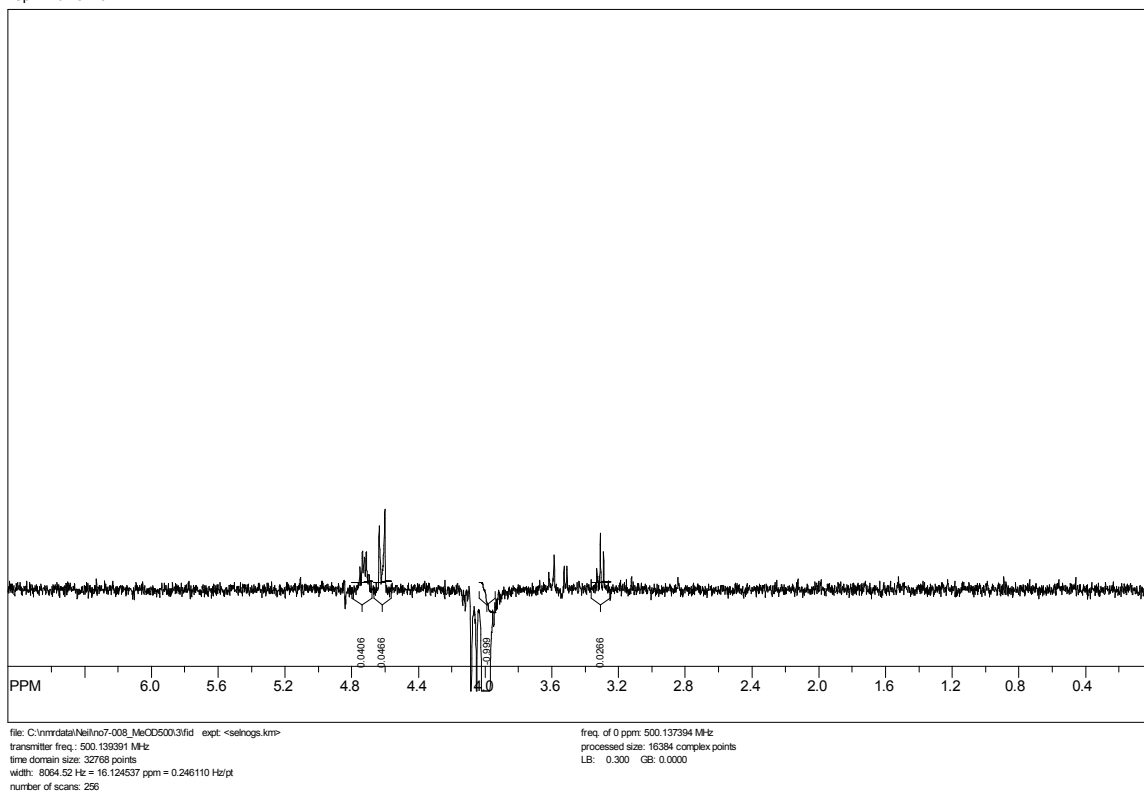


SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



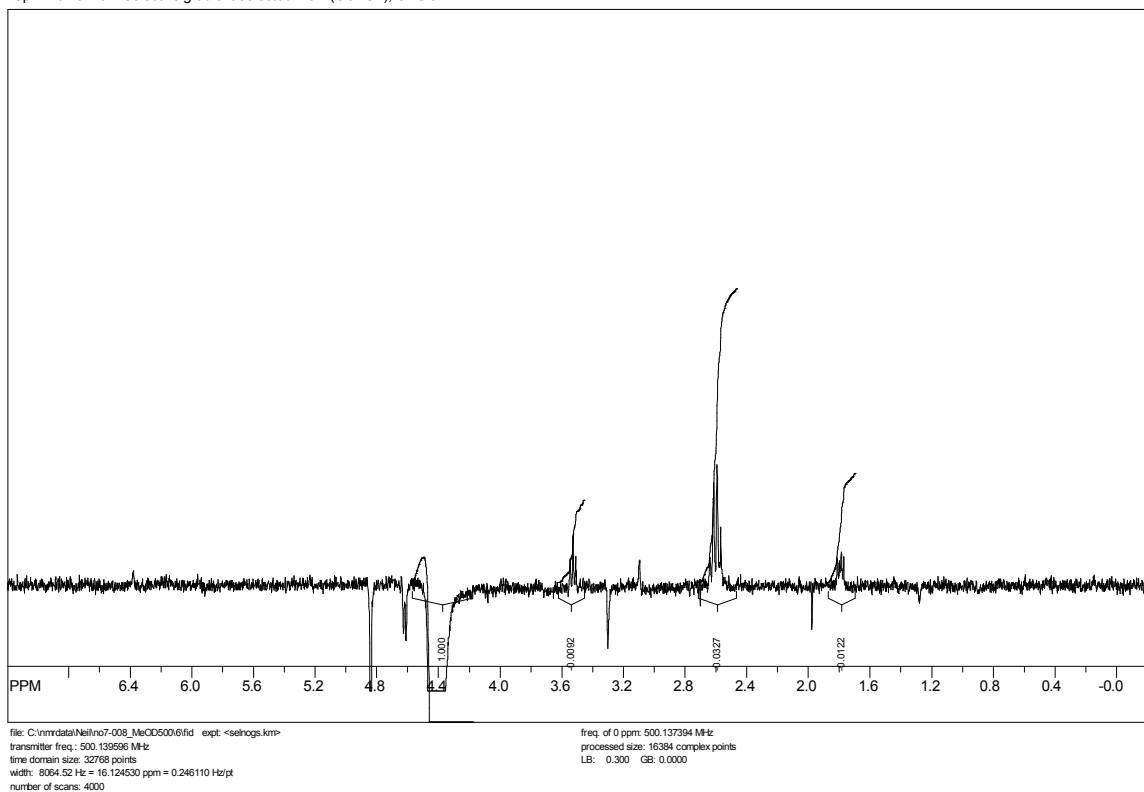
Subjection of the H₈ signal of the major conformer (4.00 ppm, 0.85 H) [H₈, 4.12 ppm] to a one-dimensional GOESY experiment in CD₃OD showed inter-proton effect to H_{α1}(Gly) of the major conformer (4.62 ppm, 0.85 H) (4.9%)

SpinWorks 2.5:

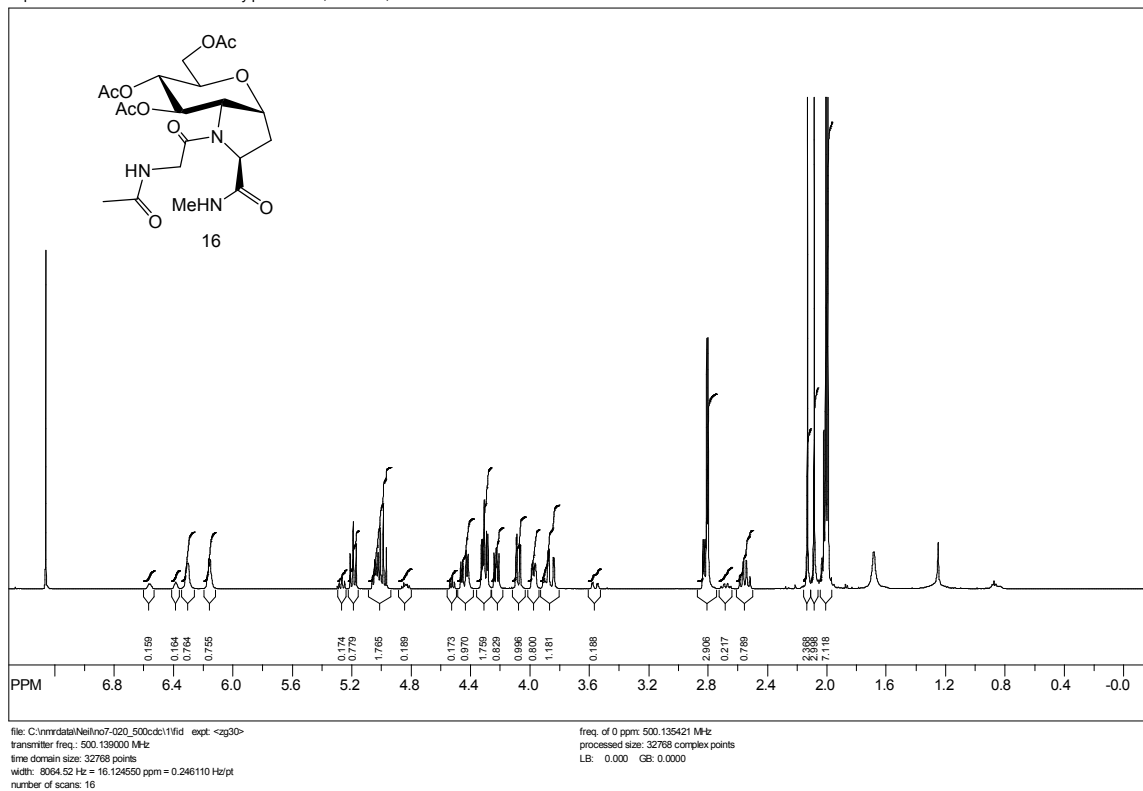


Subjection of the H₂ signal of the major conformer (4.41 ppm, 0.85 H) to a one-dimensional GOESY experiment in D₂O showed no inter-proton effect either H_{α1}(Gly) (4.62 ppm, 0.85 H) or H_{α2}(Gly) (4.07 ppm, 0.85 H)

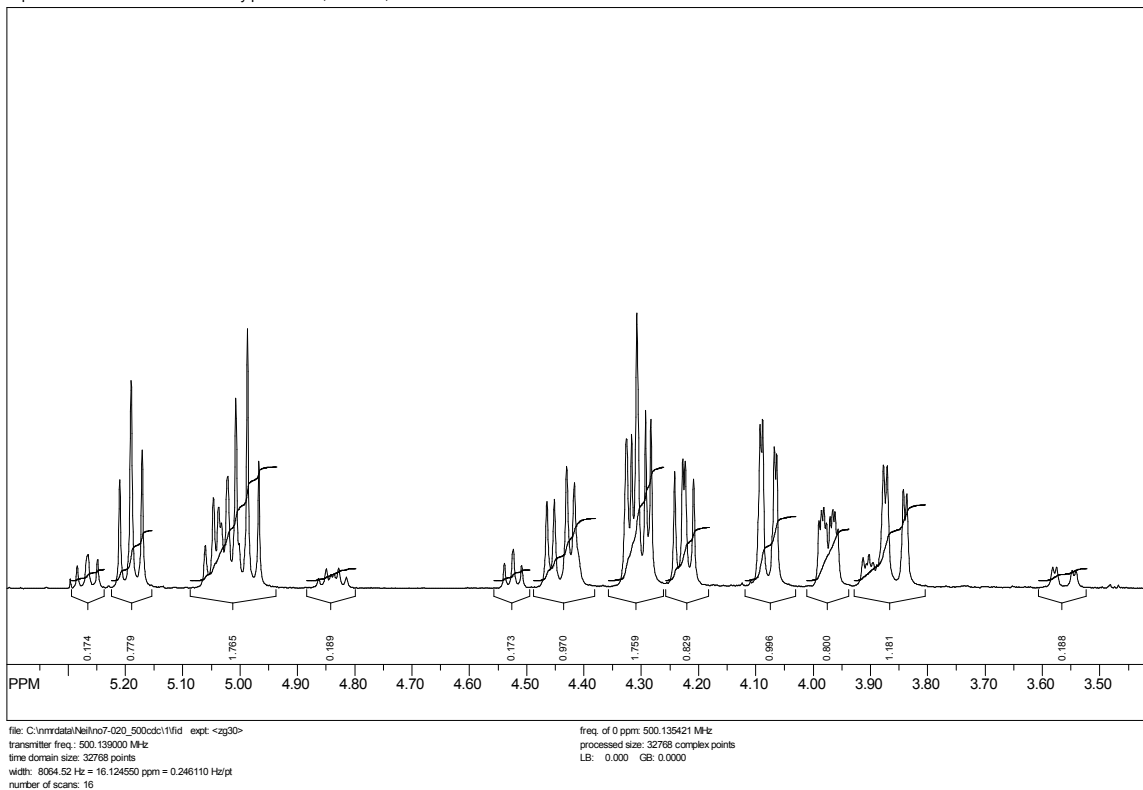
SpinWorks 2.5: Selective gradient selected NOE (GOESY), CDCl₃



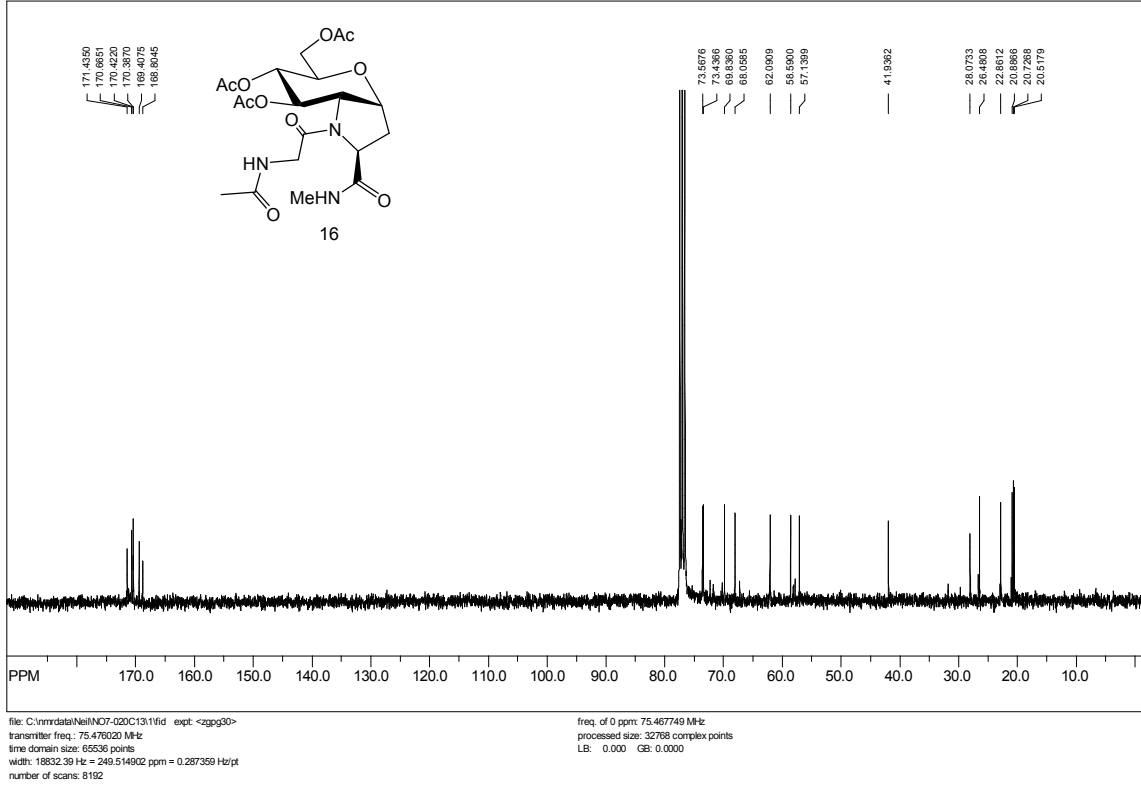
SpinWorks 2.5: standard 1-H survey parameters, AMX500, CDCl3



SpinWorks 2.5: standard 1-H survey parameters, AMX500, CDCl3

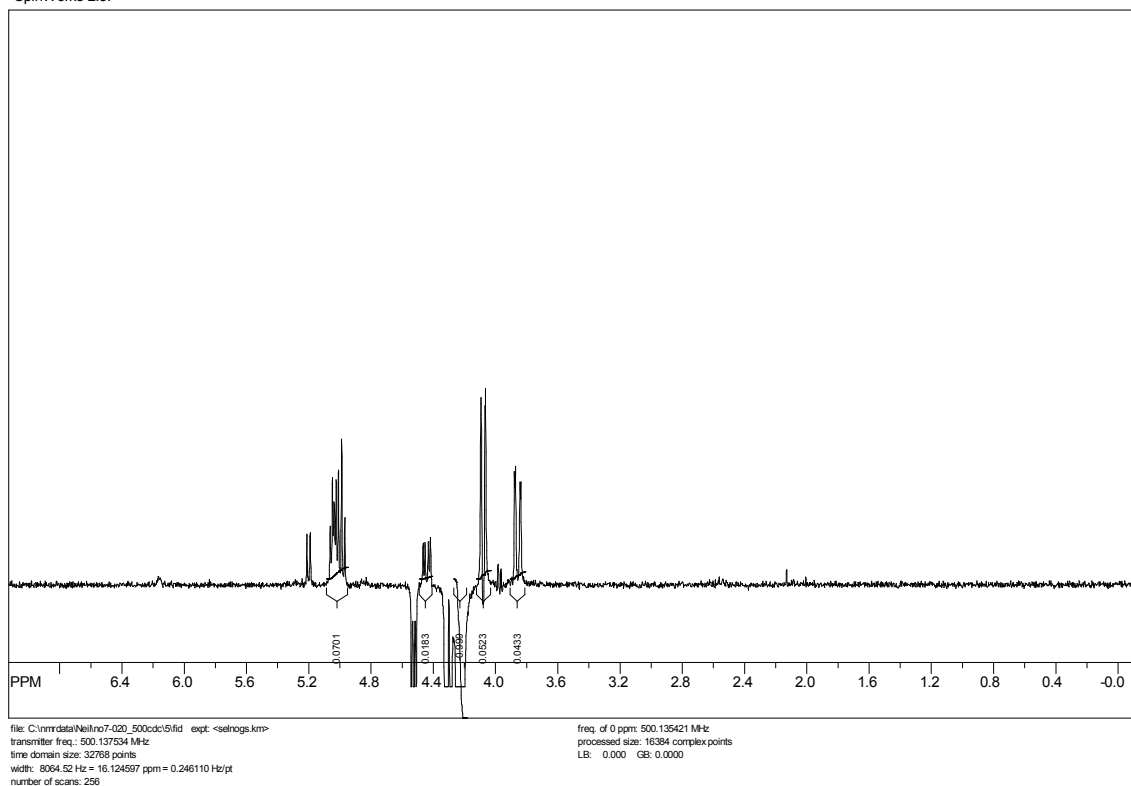


SpinWorks 2.5: NO7-020C13

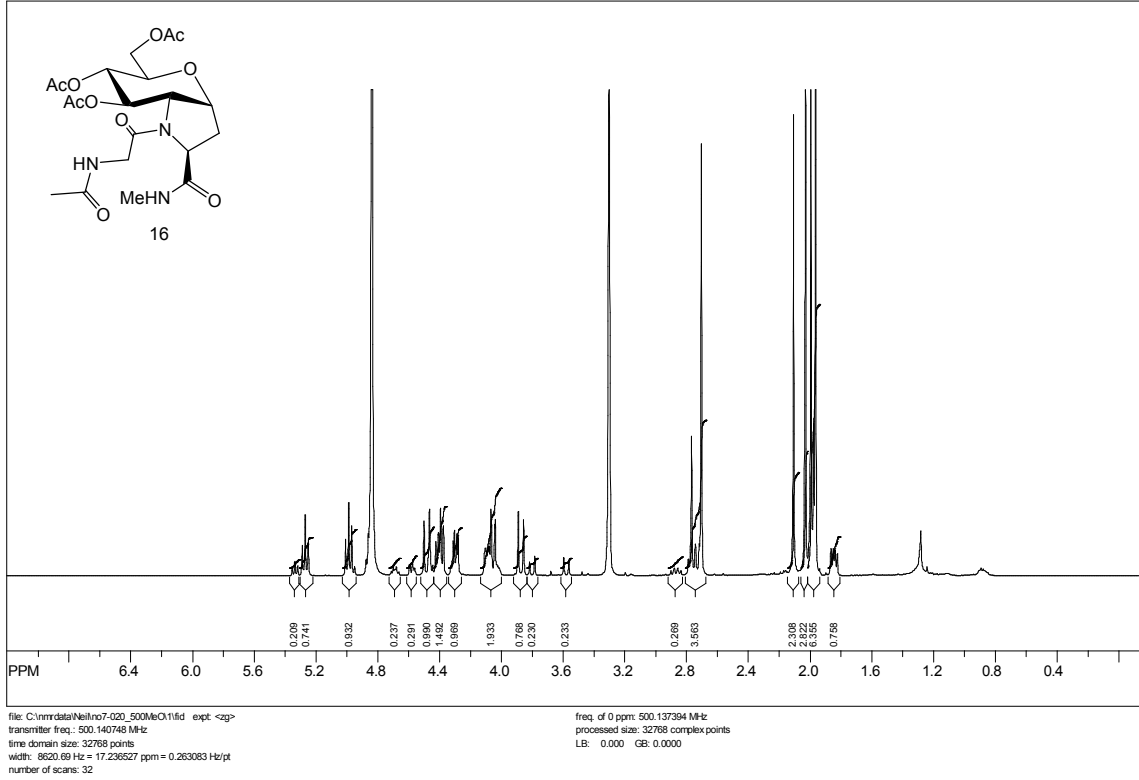


Subjection of the H₈ signal of the major conformer (4.23 ppm, 0.8 H) [H₈, 4.52 ppm, 0.2 H] to a one-dimensional GOESY experiment in CDCl₃ showed inter-proton effect to H_{α1}(Gly) of the major conformer (4.44 ppm, 0.8 H) (1.8%), and H_{α2}(Gly) of the major conformer (3.86 ppm, 0.8 H) (4.3%)

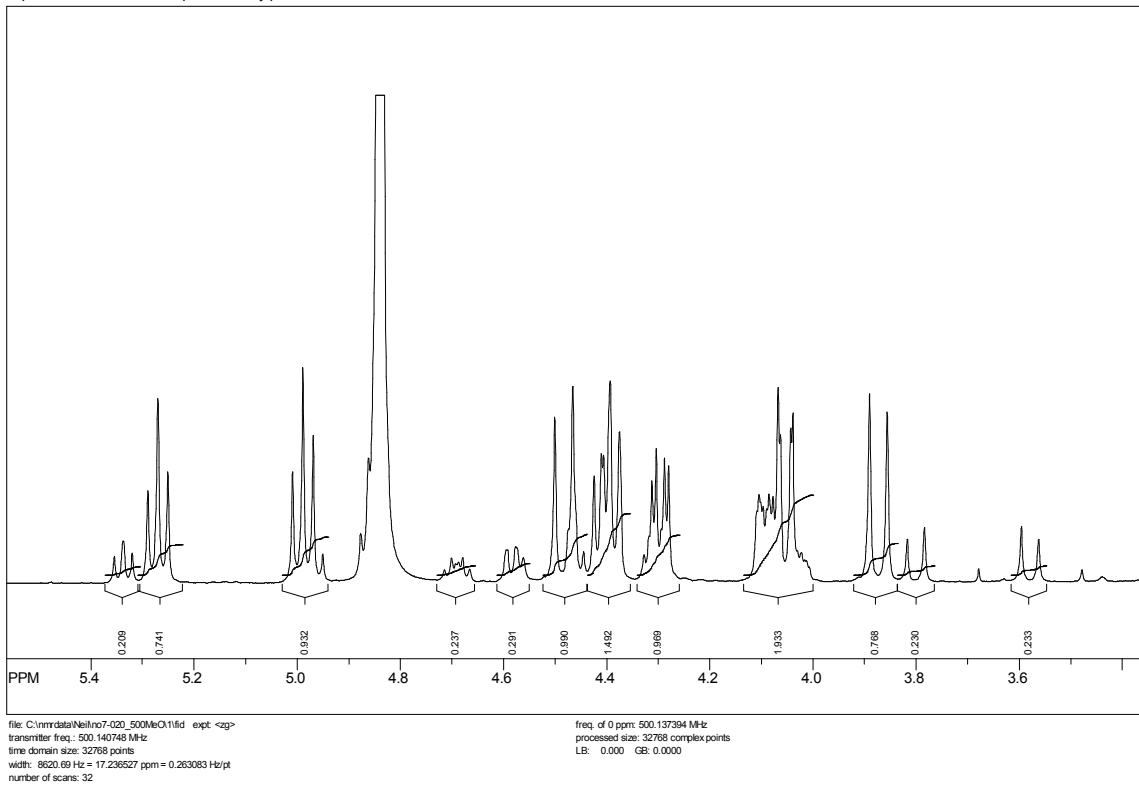
SpinWorks 2.5:



SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD

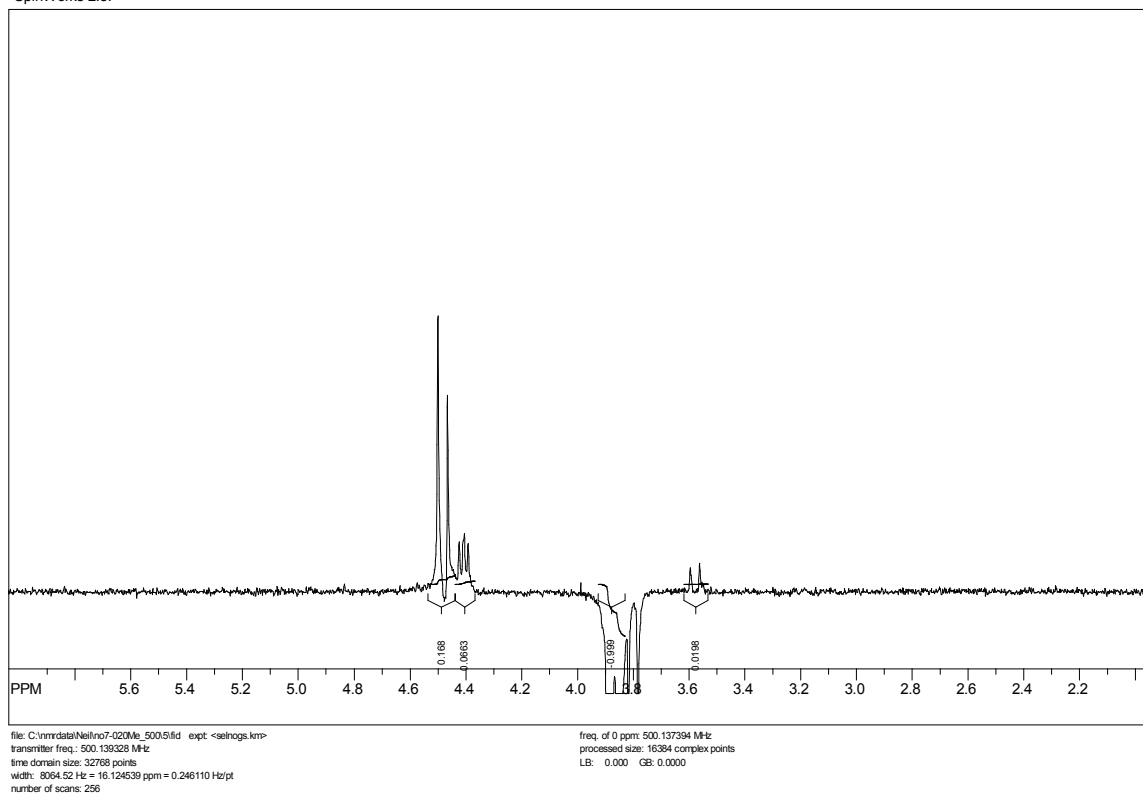


SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



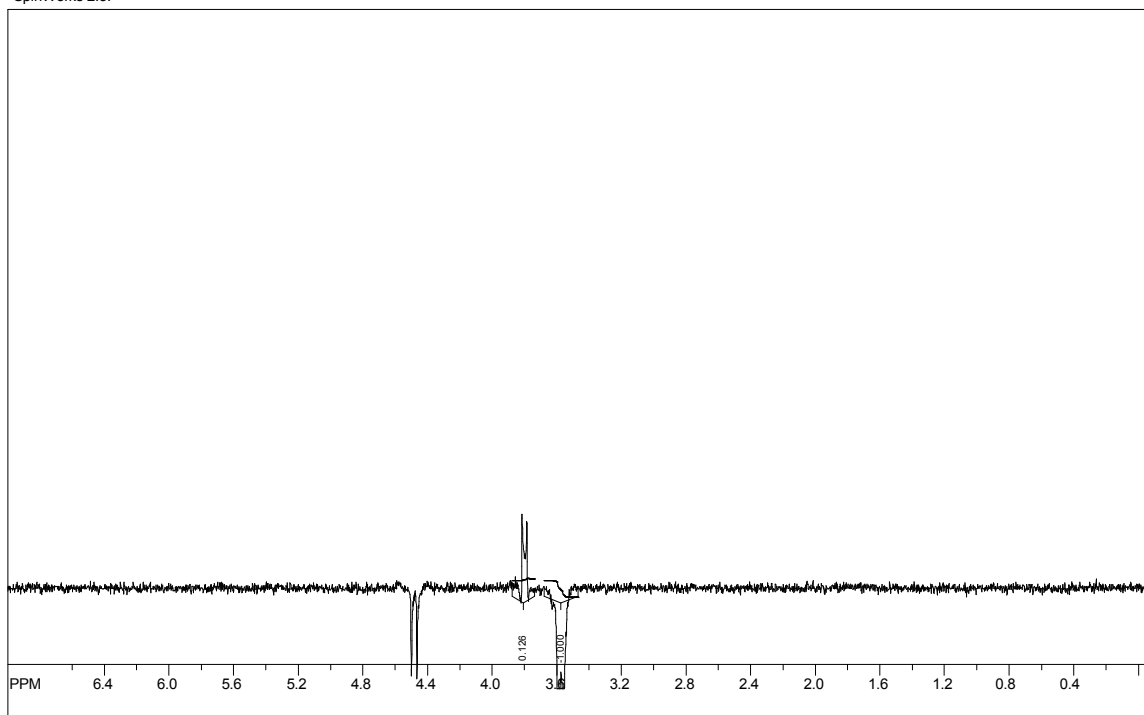
Subjection of the $H_{\alpha 2(\text{Gly})}$ signal of the major conformer (3.87 ppm, 0.83 H) [H_{β} , 4.48 ppm, 0.17 H] to a one-dimensional GOESY experiment in CD_3OD showed inter-proton effect to H_{β} of the major conformer (4.41 ppm, 0.83 H) (6.7%)

SpinWorks 2.5:



Subjection of the $H_{\alpha 2(\text{Gly})}$ signal of the minor conformer (3.58 ppm, 0.17 H) to a one-dimensional GOESY experiment in CD_3OD showed no inter-proton effect to the H_{β} of the minor conformer (4.48 ppm, 0.25 H)

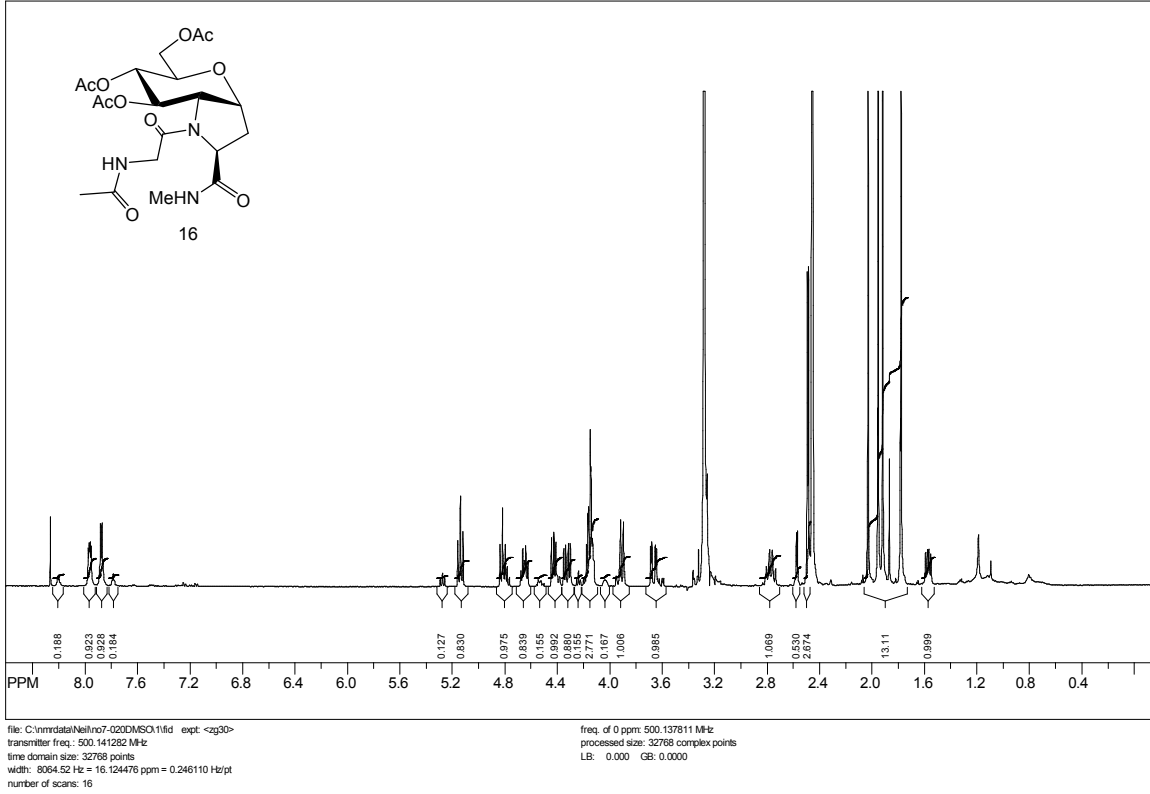
SpinWorks 2.5:



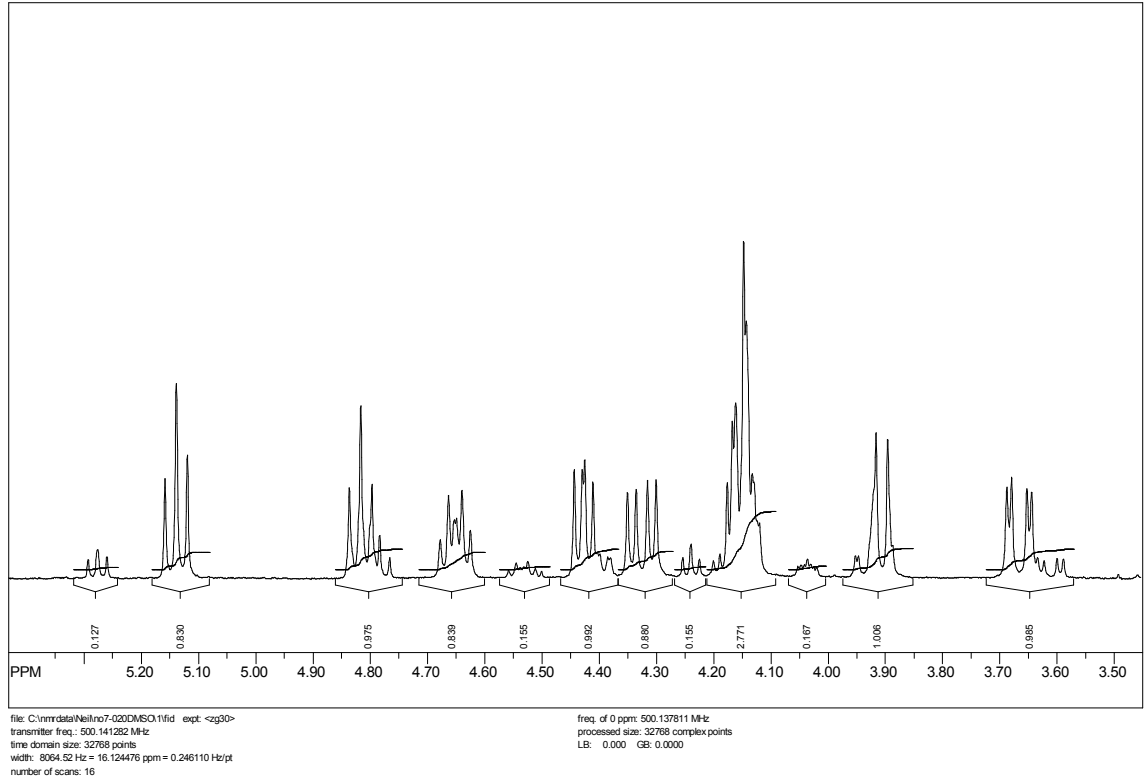
file: C:\nmrdata\Nailno7-020Me_5004.fid ept: <sethogs.km>
transmitter freq.: 500.139182 MHz
time domain size: 32768 points
width: 8064.52 Hz = 16.124544 ppm = 0.246110 Hz/pt
number of scans: 256

freq. of 0 ppm: 500.137394 MHz
processed size: 16384 complex points
LB: 0.000 GB: 0.0000

SpinWorks 2.5: standard 1-H survey parameters, AMX500, dms0-d6



SpinWorks 2.5: standard 1-H survey parameters, AMX500, dms0-d6



**Chapter 10: Effects of Glycosylation of (2*S*,4*R*)-4-Hydroxyproline on
the Conformation, Kinetics and Thermodynamics of Prolyl Amide**

Isomerization

Supporting Information For Chapter 4

Table of Contents

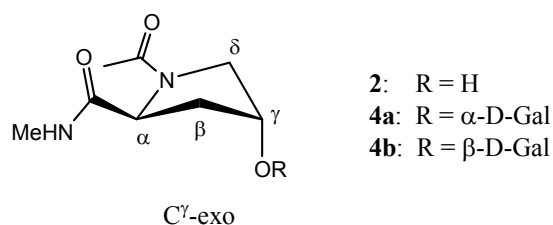
Tables	370-373
Van't Hoff Plots	374-376
Eyring Plots	377-379
Spectra	380-388
nOe Spectra	389-396
Summary of nOe interactions for 4a , 4b	397
Magnetization Transfer NMR Experiments	398-427

Table 10.1: Chemical Shifts (δ) for $^{13}\text{C}^\alpha$, $^{13}\text{C}^\delta$ and $^{13}\text{C}^\gamma$ of *cis* and *trans* Prolyl Amide Isomers^[a]

Compound	$\delta(^{13}\text{C}^\alpha)$ <i>cis</i>	$\delta(^{13}\text{C}^\alpha)$ <i>trans</i>	$\delta(^{13}\text{C}^\delta)$ <i>cis</i>	$\delta(^{13}\text{C}^\delta)$ <i>trans</i>	$\delta(^{13}\text{C}^\gamma)$ <i>cis</i>	$\delta(^{13}\text{C}^\gamma)$ <i>trans</i>
1	62.3	60.8	47.4	49.1	22.8	24.5
2	60.5	59.3	54.7	56.5	68.6	69.9
3	60.6	59.2	53.7	55.4	68.9	70.1
4a	60.8	59.6	52.0	53.9	74.8	76.5
4b	60.4	59.2	53.2	54.9	76.7	77.9

^[a]Determined by 75 MHz NMR in D₂O at 25 °C.

Table 10.2: *J* Coupling Constants (Hz) for Compounds **2**, **4a**, **4b**^[a]



Compound	$J_{\text{H}\alpha,\text{H}\beta 1}$	$J_{\text{H}\alpha,\text{H}\beta 2}$	$J_{\text{H}\beta 1,\text{H}\beta 2}$	$J_{\text{H}\beta 1,\text{H}\gamma}$	$J_{\text{H}\beta 2,\text{H}\gamma}$	$J_{\text{H}\gamma,\delta 1}$	$J_{\text{H}\gamma,\delta 2}$	$J_{\text{H}\delta 1,\delta 2}$	$J_{\text{H}\beta 1,\delta 1}$
2 (<i>trans</i>) ^[c]	8.4	8.4	13.6	<2	4.5	n.d. ^[b]	4.0	11.8	<2
2 (<i>cis</i>) ^[c]	7.9	7.9	13.7	~2	6.4	n.d. ^[b]	4.3	12.9	~2
4a (<i>trans</i>)	8.3	8.5	13.2	<2	4.5	n.d. ^[b]	3.9	12.0	<2
4a (<i>cis</i>)	8.0	8.0	13.2	n.d. ^[b]	n.d. ^[b]	n.d. ^[b]	4.2	12.8	n.d. ^[b]
4b (<i>trans</i>)	8.2	8.2	13.6	2.5	5.0	n.d. ^[b]	4.8	12.7	<1.5
4b (<i>cis</i>)	7.8	7.8	14.1	n.d. ^[b]	5.0	n.d. ^[b]	4.5	12.8	n.d. ^[b]

^[a]Determined in D₂O at 25 °C; The observed coupling constants are consistent with C ^{γ} -exo conformation of the pyrrolidine ring. ^[b]Not determined due to signal overlap; ^[c]Previously assigned to exist in a C ^{γ} -exo conformation.

Table 10.3: Amide Isomer Equilibrium ($K_{t/c}$) of **2-4b** at Various Temperatures^[a]

Cpd	Temperature (°C) ^[b]										
	24.8	30.0	37.0	46.0	51.3	56.6	62.0	67.3	72.6	78.0	83.3
2	3.89 ±0.11	3.73 ±0.09	3.52 ±0.05	3.39 ±0.13	3.29 ±0.11	3.22 ±0.09	3.16 ±0.09	2.94 ±0.08	2.87 ±0.08	2.74 ±0.08	2.70 ±0.04
3	3.66 ±0.15	3.50 ±0.07	3.34 ±0.15	3.15 ±0.20	3.03 ±0.17	2.96 ±0.18	2.85 ±0.2	2.77 ±0.01	2.72 ±0.05	2.65 ±0.03	2.55 ±0.07
4a	3.74 ±0.20	3.62 ±0.46	3.41 ±0.30	3.24 ±0.31	3.22 ±0.22	3.02 ±0.35	2.96 ±0.21	2.83 ±0.17	2.77 ±0.2	2.74 ±0.13	2.62 ±0.57
4b	3.71 ±0.21	3.62 ±0.21	3.37 ±0.28	3.14 ±0.25	3.07 ±0.17	2.99 ±0.17	2.91 ±0.25	2.79 ±0.19	2.72 ±0.18	2.66 ±0.18	2.61 ±0.11

^[a]Carried out in D₂O, ± SE determined by integration of two or more sets of *trans/cis* isomers;

^[b]Calibrated using an ethylene glycol standard;

Table 10.4: Amide Isomer Equilibrium ($K_{t/c}$) of **1** at Various Temperatures^[a]

Compound	Temperature (°C) ^[b]							
	3.5	14.2	24.8	37.0	46.0	56.6	72.6	83.3
1	3.61 ±0.11	3.39 ±0.11	3.19 ±0.06	3.03 ±0.08	2.87 ±0.11	2.73 ±0.11	2.56 ±0.11	2.45 ±0.11

^[a]Carried out in D₂O; ^[b]Calibrated using an ethylene glycol standard.

Table 10.5: Thermodynamic Parameters for Amide Isomerization of **1-4b**^[a]

Compound	ΔH° (kcal·mol ⁻¹)	ΔS° (cal·mol ⁻¹ ·K ⁻¹)
1	-0.95 ± 0.01	-0.87 ± 0.02
2	-1.33 ± 0.03	-1.76 ± 0.09
3	-1.29 ± 0.02	-1.72 ± 0.06
4a	-1.27 ± 0.02	-1.64 ± 0.07
4b	-1.30 ± 0.03	-1.77 ± 0.10

^[a]Carried out in D₂O; Values and error limits were obtained using the linear least-squares fits of the data in Tables 10.3 and 10.4 to the equation the $\ln K_{ct} = (-\Delta H^\circ/R)(1/T) + \Delta S^\circ/R$.

Table 10.6: Activation Enthalpy (ΔH^\ddagger) and Entropy (ΔS^\ddagger) of Isomerization of **1-4b**^[a]

Compound	<i>cis to trans</i>			<i>trans to cis</i>		
	ΔH^\ddagger (Kcal· mol ⁻¹)	ΔS^\ddagger (cal·mol ⁻¹ ·K ⁻¹)	ΔG^\ddagger_{300K} (Kcal· mol ⁻¹)	ΔH^\ddagger (Kcal· mol ⁻¹)	ΔS^\ddagger (cal·mol ⁻¹ ·K ⁻¹)	ΔG^\ddagger_{300K} (Kcal· mol ⁻¹)
1	20.6 ± 0.4	1.2 ± 0.1	20.2 ± 0.4	21.3 ± 0.2	1.4 ± 0.1	20.9 ± 0.2
2	20.2 ± 0.3	0.08 ± 0.01	20.2 ± 0.3	21.1 ± 0.3	0.54 ± 0.05	21.0 ± 0.3
3	21.8 ± 0.7	4.7 ± 0.2	20.4 ± 0.7	22.7 ± 0.9	5.4 ± 0.3	21.1 ± 0.9
4a	20.4 ± 0.9	0.60 ± 0.03	20.2 ± 0.9	20.6 ± 1.0	1.0 ± 0.1	20.3 ± 1.0
4b	22.4 ± 0.8	6.2 ± 0.3	20.6 ± 0.8	23.3 ± 1.4	6.6 ± 0.4	21.3 ± 1.0

^[a]Values and error limits were obtained using the linear least-squares fits of the data in Table 10.7 to the equation $\ln(k/T) = (-\Delta H^\ddagger/R)(1/T) + \Delta S^\ddagger/R + \ln(k_B/h)$; Additionally the free energy of activation at 300 K (ΔG^\ddagger_{300K}) is given.; The error in ΔG^\ddagger was calculated using the equation $\sigma_k = k[(\sigma^2 \Delta H/R^2 T^2) + (\sigma^2 \Delta S/R^2)]^{1/2}$.

Table 10.7: Rate of Prolyl Amide Isomerization of **1-4b** at Various Temperatures

Compound	Rate ^[a] (s ⁻¹)	Temperature (°C) ^[b]					
		62.0	67.3	72.6	78.0	83.3	37.0 ^[e]
1	k _{ct}	0.461 ±0.012	0.810 ±0.022	1.262 ±0.035	2.000 ±0.071	3.171 ±0.081	0.039 ±0.024
	k _{tc} ^[d]	0.180 ±0.009	0.312 ±0.015	0.507 ±0.022	0.815 ±0.046	1.314 ±0.053	0.013 ±0.010
2	k _{ct}	0.455 ±0.017	0.725 ±0.014	1.149 ±0.022	1.852 ±0.048	2.941 ±0.097	0.039 ±0.005
	k _{tc} ^[d]	0.147 ±0.007	0.247 ±0.008	0.398 ±0.014	0.634 ±0.024	1.053 ±0.038	0.011 ±0.001
3	k _{ct}	0.455 ±0.024	0.769 ±0.022	1.395 ±0.029	2.071 ±0.036	3.414 ±0.054	0.032 ±0.009
	k _{tc} ^[d]	0.163 ±0.009	0.272 ±0.008	0.532 ±0.015	0.792 ±0.016	1.302 ±0.040	0.010 ±0.004
4a	k _{ct}	0.476 ±0.042	0.833 ±0.047	1.428 ±0.089	2.149 ±0.075	3.124 ±0.077	0.043 ±0.015
	k _{tc} ^[d]	0.150 ±0.016	0.271 ±0.021	0.429 ±0.047	0.728 ±0.043	0.988 ±0.181	0.013 ±0.005
4a ^[c]	k _{ct}		0.820 ±0.059				
	k _{tc} ^[d]		0.274 ±0.027				
4b	k _{ct}	0.385 ±0.038	0.606 ±0.037	1.064 ±0.075	1.682 ±0.053	3.071 ±0.066	0.022 ±0.007
	k _{tc} ^[d]	0.127 ±0.021	0.207 ±0.018	0.350 ±0.032	0.578 ±0.040	1.113 ±0.051	0.007 ±0.004
4b ^[b]	k _{ct}		0.571 ±0.052				
	k _{tc} ^[d]		0.189 ±0.022				

^[a]Carried out in D₂O as described on page 398. Error values were obtained using the student's t test of the linear least-squares fit of the data in Figures 10.15-10.41; ^[b]Calibrated using an ethylene glycol standard; ^[c]Carried out in phosphate buffer pH = 7.4; ^[d]Calculated from k_{ct} and equilibrium ($K_{trans/cis}$) at respective temperatures as described on page 398; ^[e]Rate ±SE calculated using equation $\ln(k/T) = (-\Delta H^\ddagger/R)(1/T) + \Delta S^\ddagger/R + \ln(k_B/h)$ and values from Table 10.7.

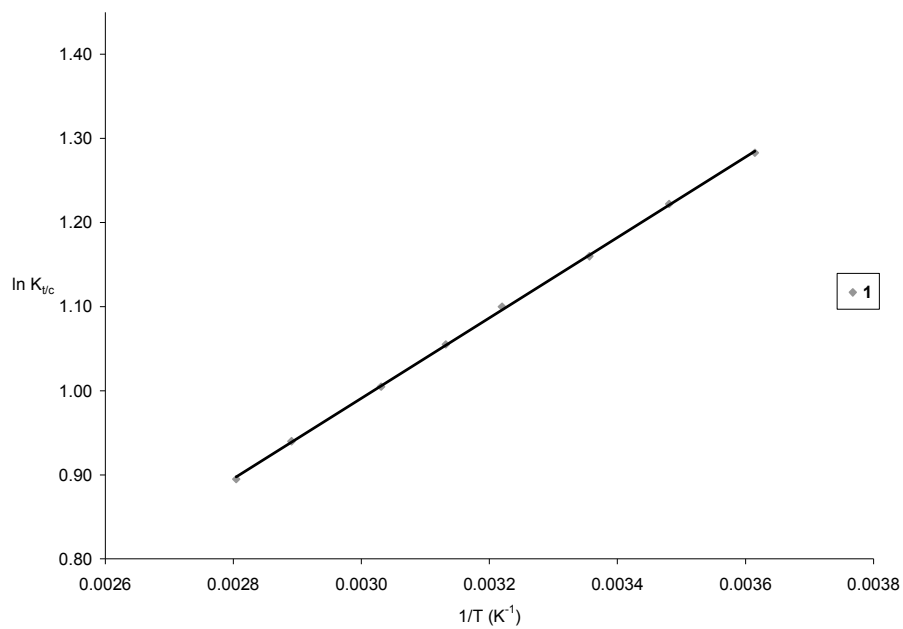
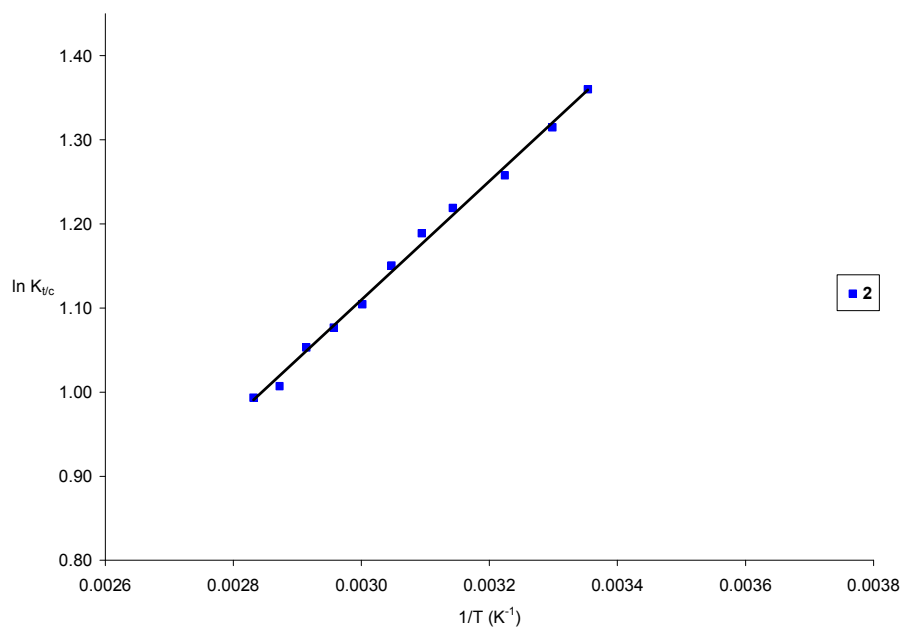
Figure 10.1: Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **1** in D₂O**Figure 10.2:** Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **2** in D₂O

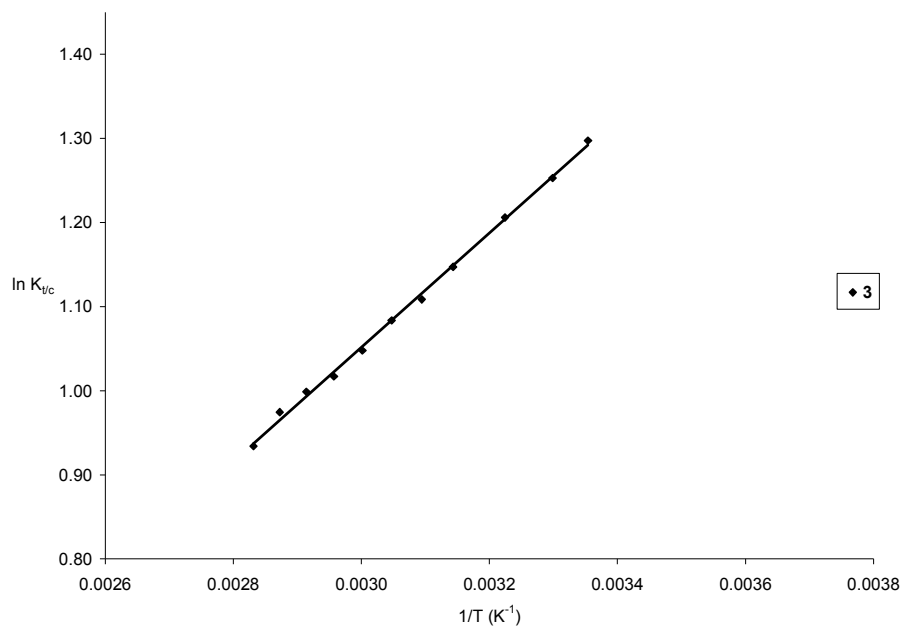
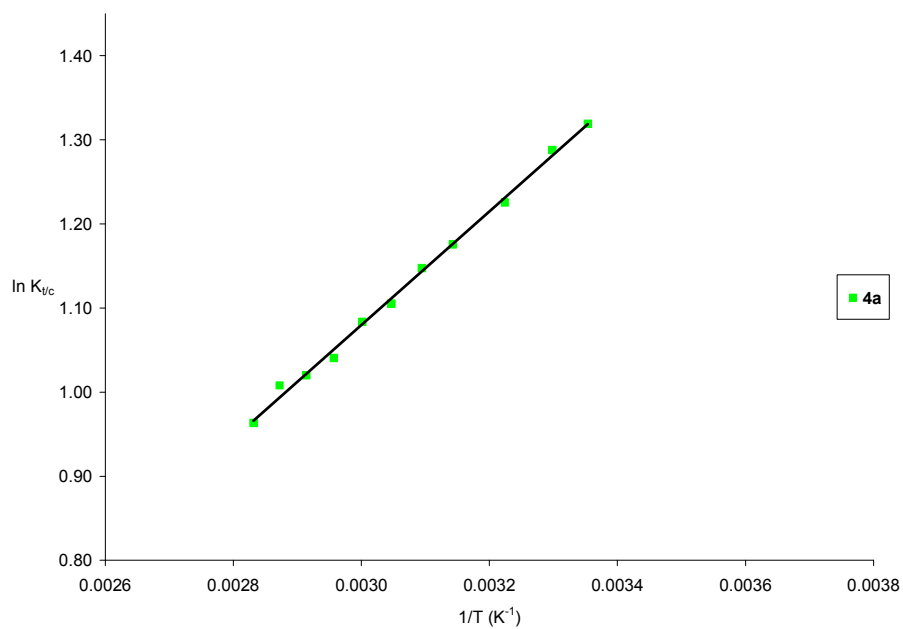
Figure 10.3: Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **3** in D₂O**Figure 10.4:** Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **4a** in D₂O

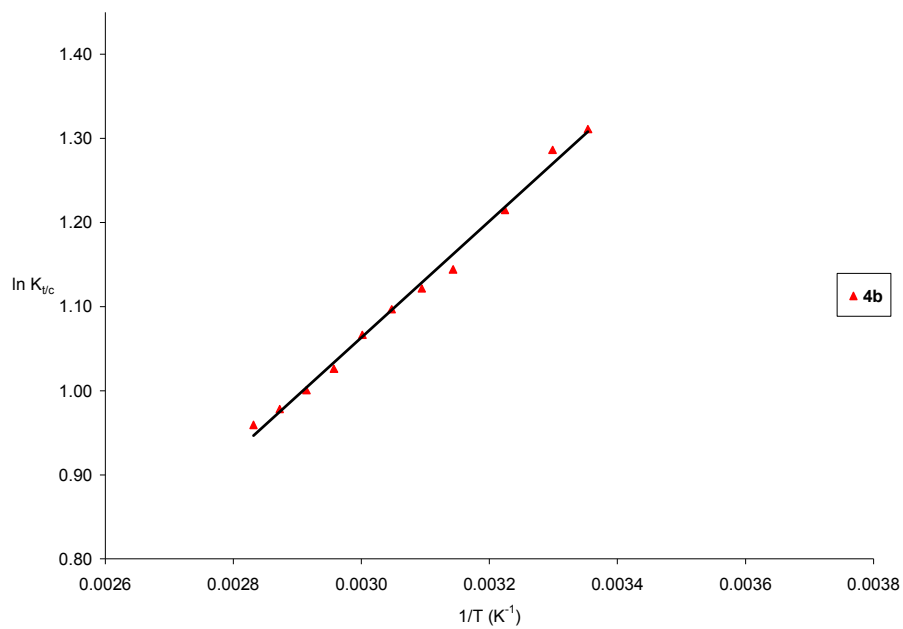
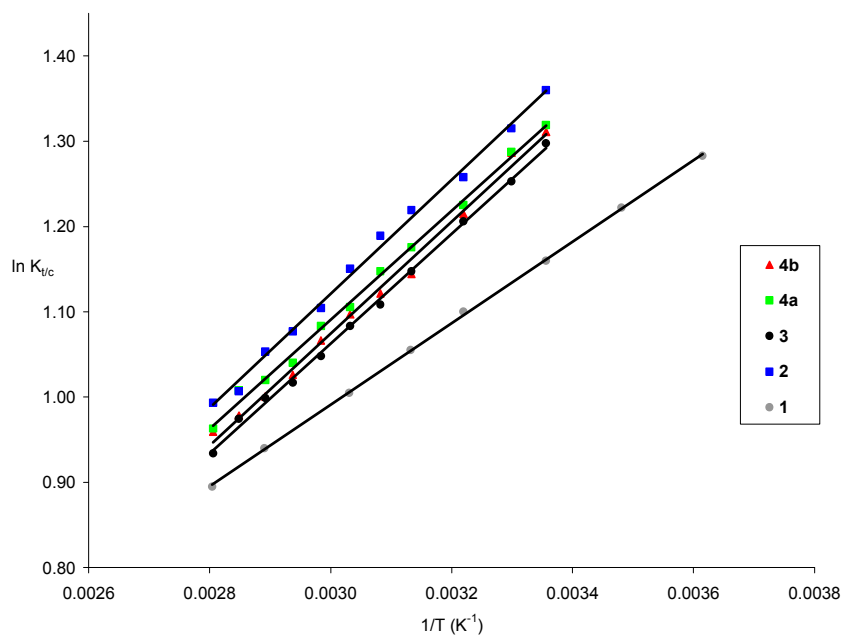
Figure 10.5: Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **4b** in D₂O**Figure 10.6:** Comparison of Van't Hoff Plots for **1-4b** in D₂O

Figure 10.7: Eyring Plot for the *cis*-to-*trans* Isomerization (k_{ct}) and *trans*-to-*cis* Isomerization (k_{tc}) for **1** in D₂O

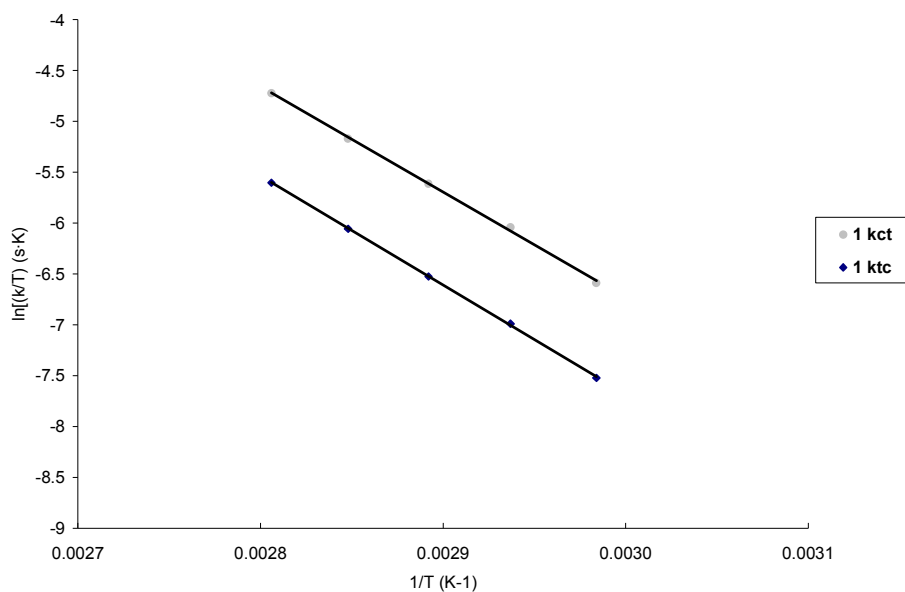


Figure 10.8: Eyring Plot for the *cis*-to-*trans* Isomerization (k_{ct}) and *trans*-to-*cis* Isomerization (k_{tc}) for **2** in D₂O

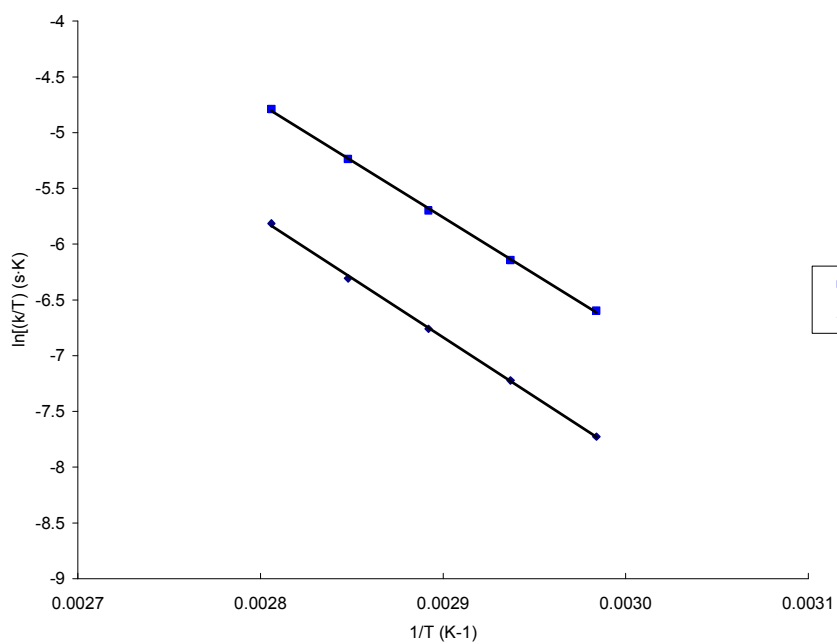


Figure 10.9: Eyring Plot for the *cis*-to-*trans* Isomerization (k_{ct}) and *trans*-to-*cis* Isomerization (k_{tc}) for **3** in D₂O

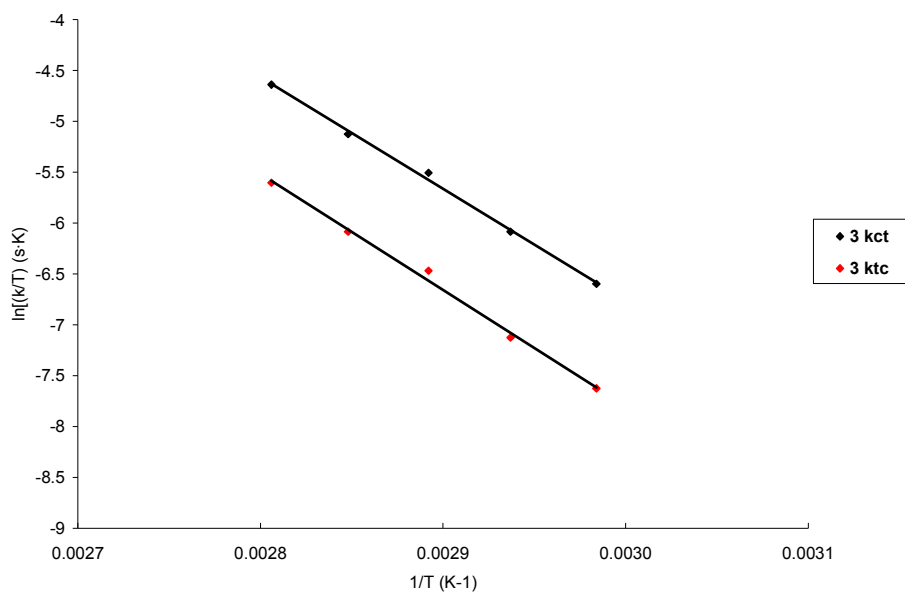


Figure 10.10: Eyring Plot for the *cis*-to-*trans* Isomerization (k_{ct}) and *trans*-to-*cis* Isomerization (k_{tc}) for **4a** in D₂O

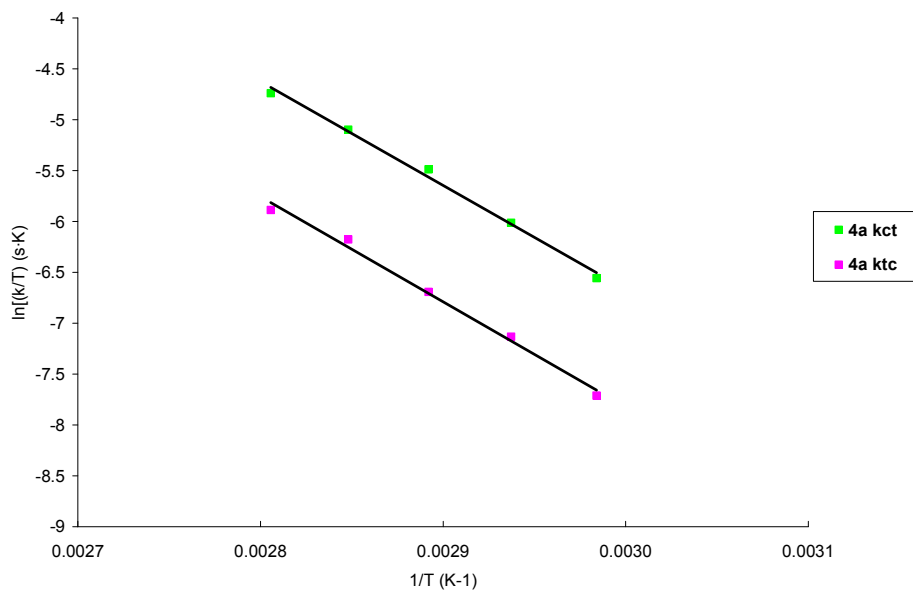


Figure 10.11: Eyring Plot for the *cis*-to-*trans* Isomerization (k_{ct}) and *trans*-to-*cis* Isomerization (k_{tc}) for **4b** in D₂O

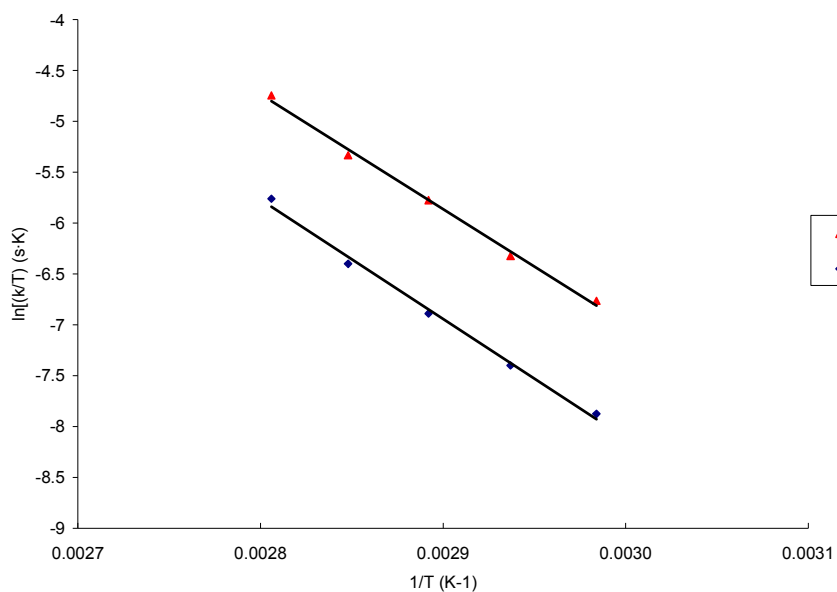
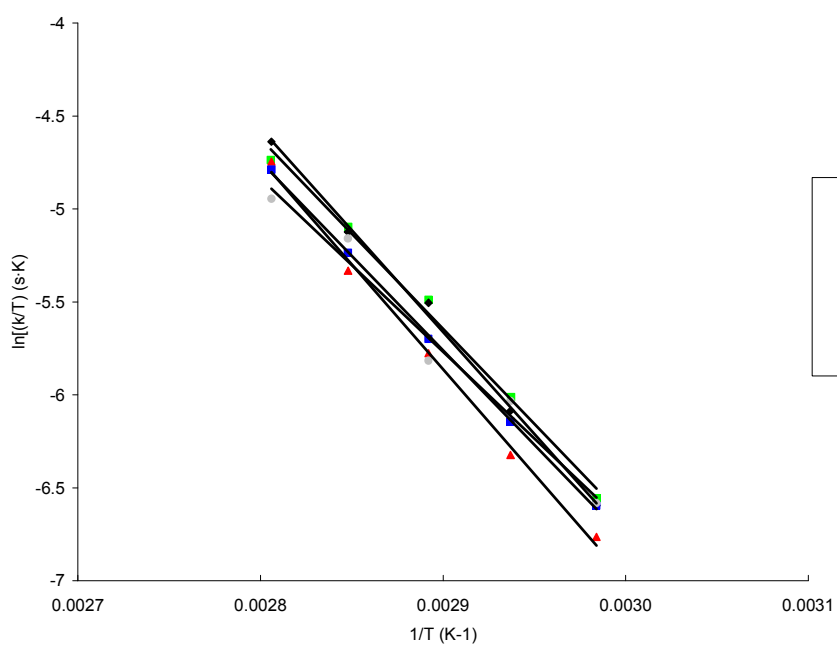
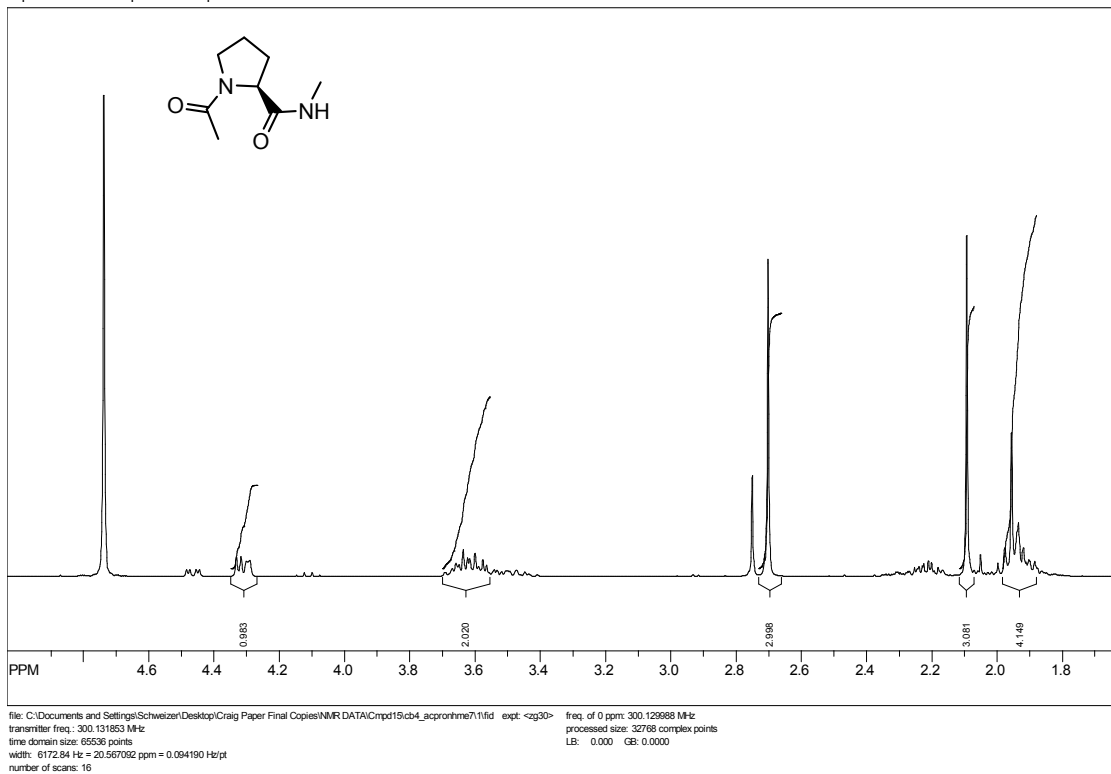


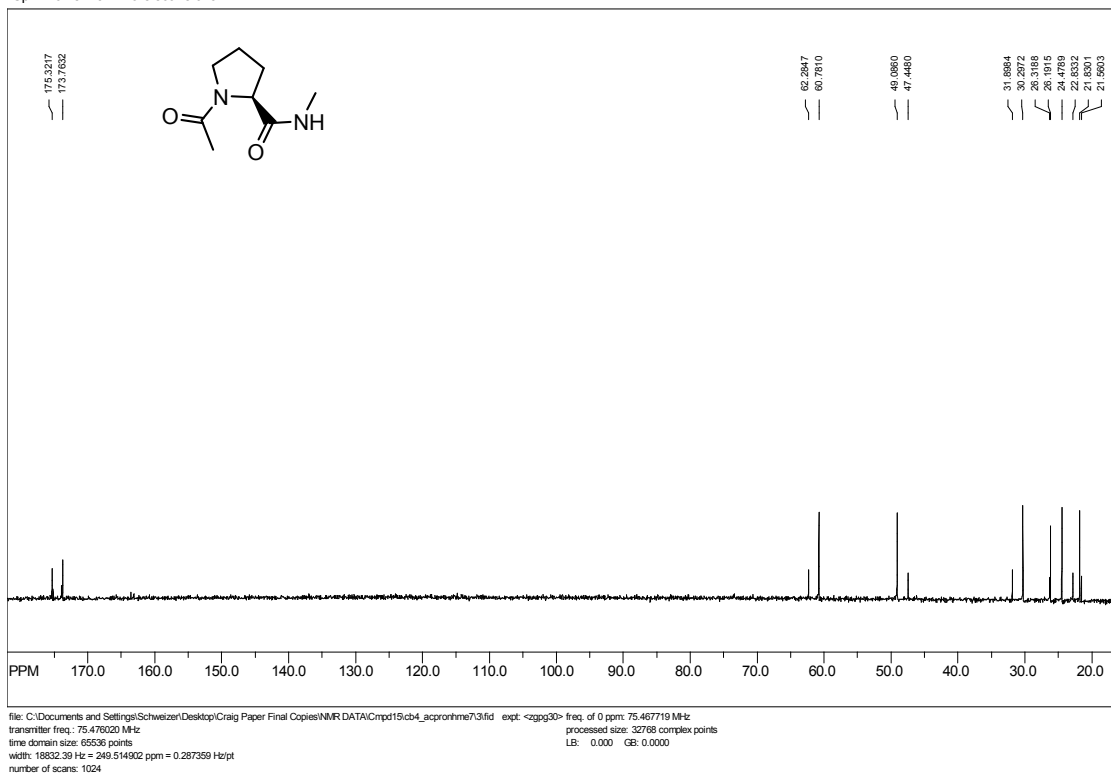
Figure 10.12: Comparison of Eyring Plot of *cis*-to-*trans* isomerization for **1-4b**



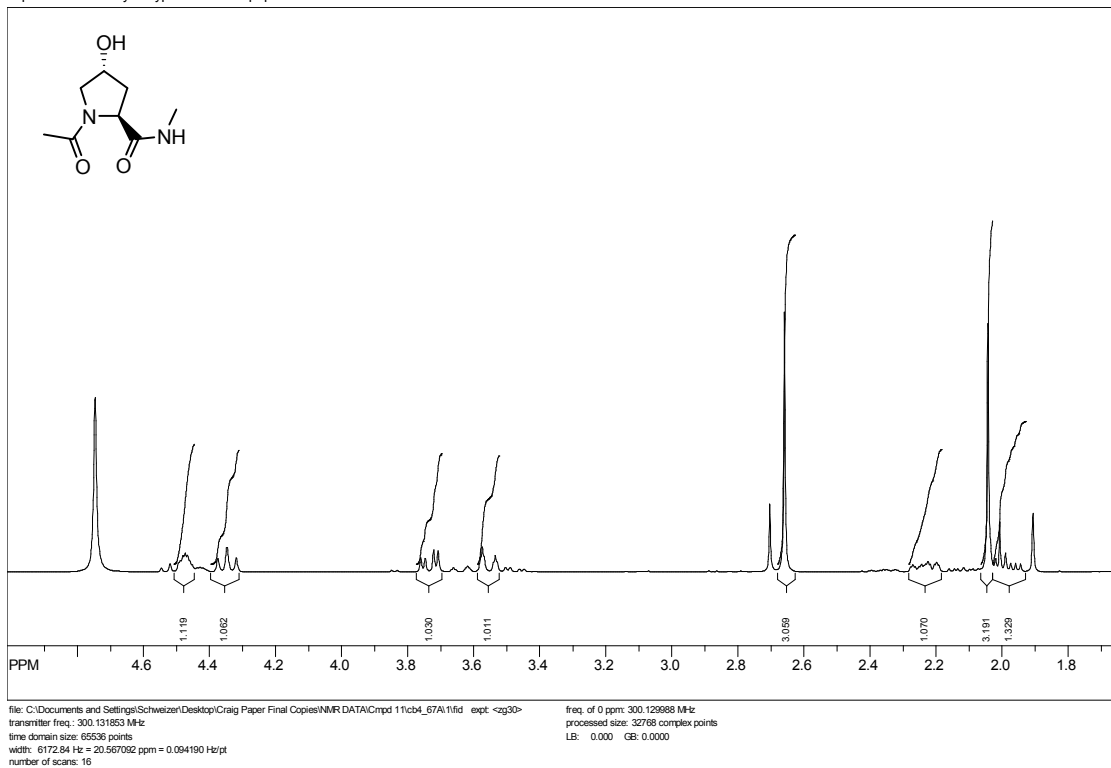
SpinWorks 2.5: repurified sample



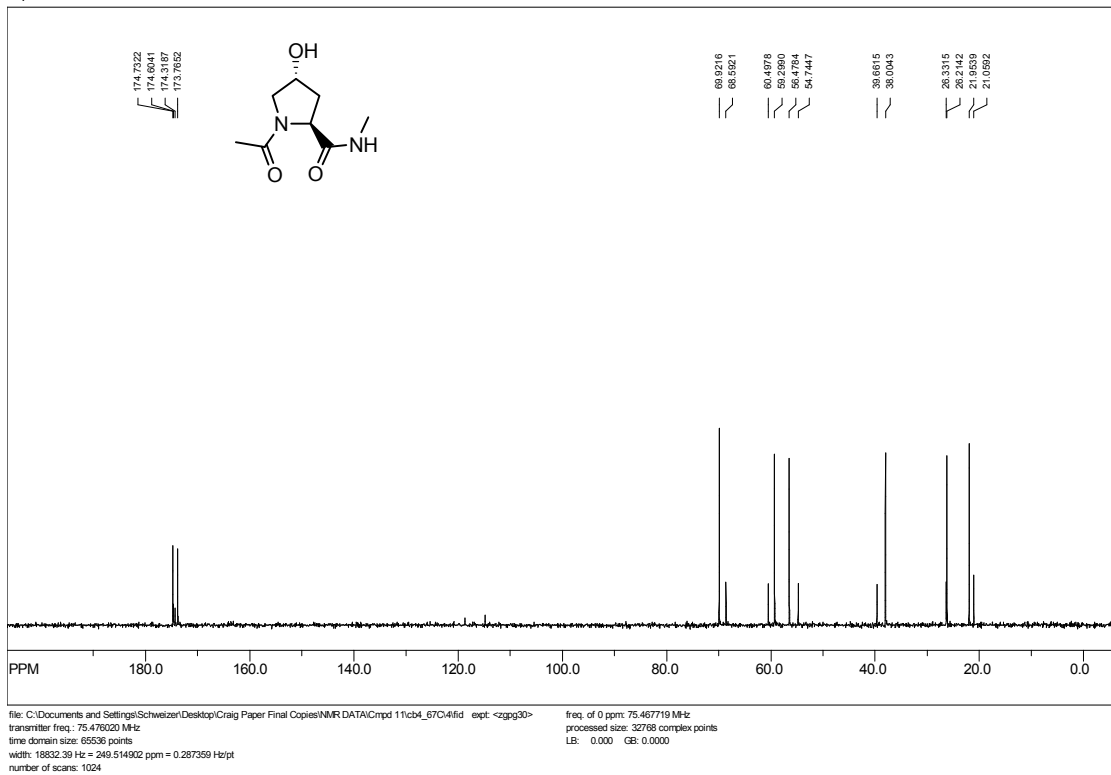
SpinWorks 2.5: more scans c13



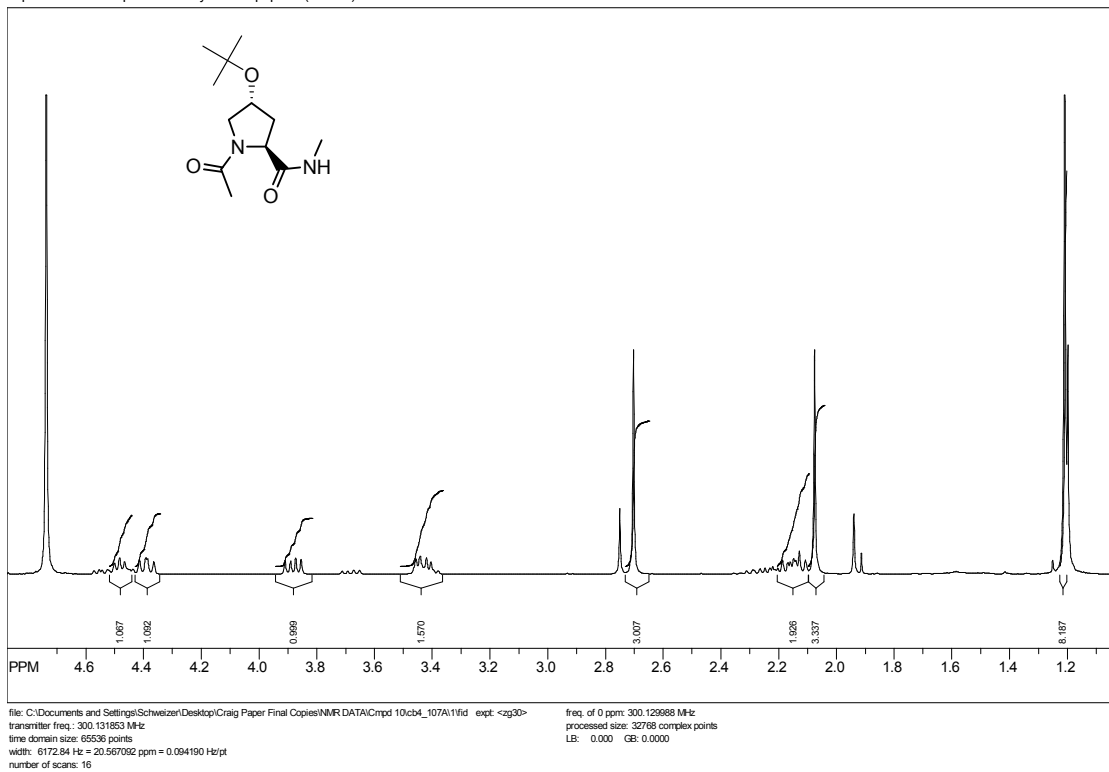
SpinWorks 2.5: hydroxyproline model peptide



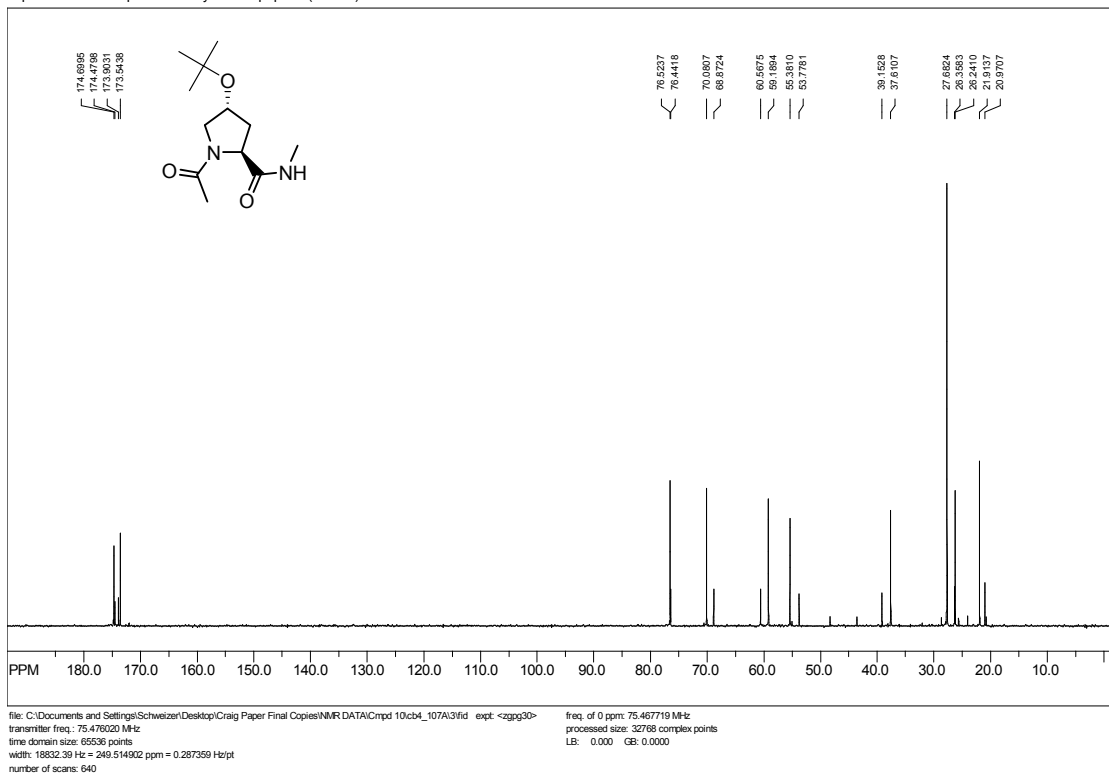
SpinWorks 2.5: C13CPD D2O u schweiz 2



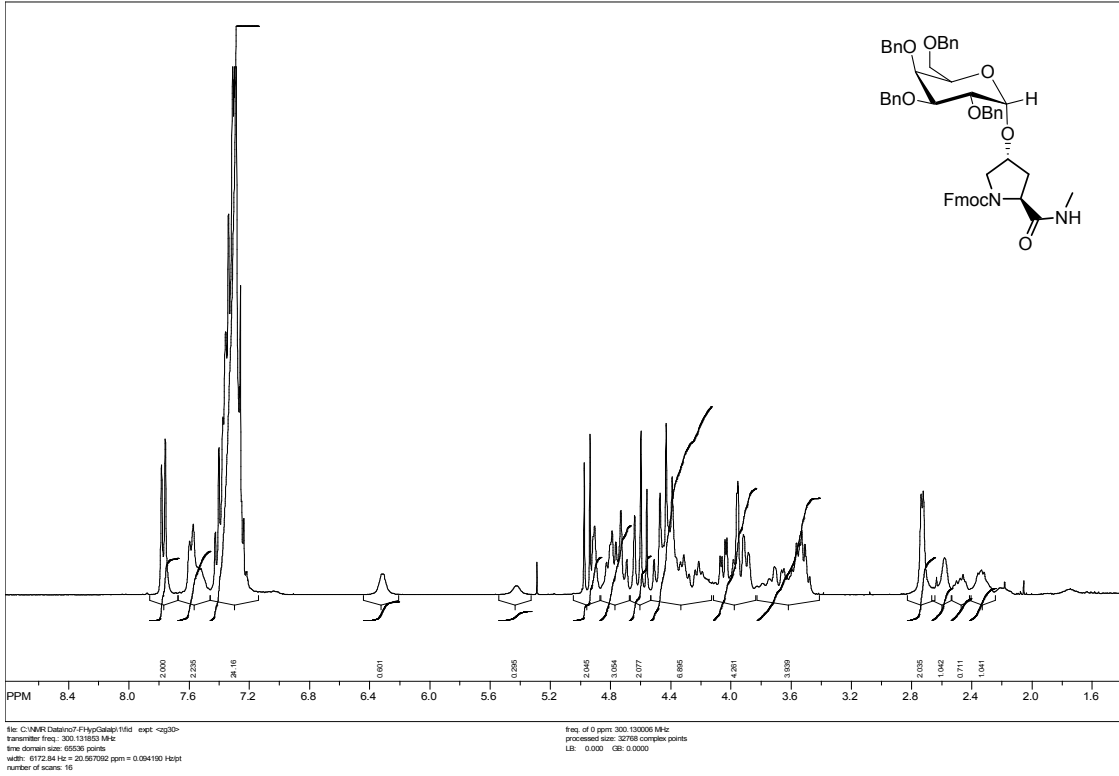
SpinWorks 2.5: purified t-butyl model peptide (in D2O)



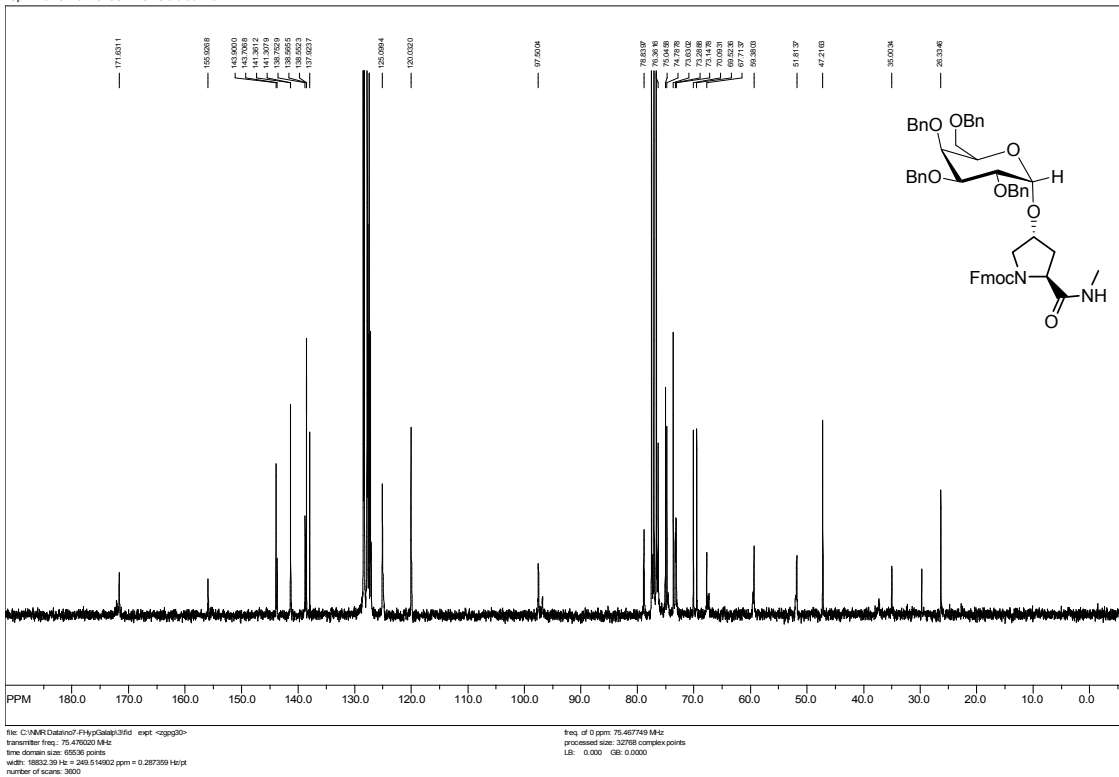
SpinWorks 2.5: purified t-butyl model peptide (in D2O)



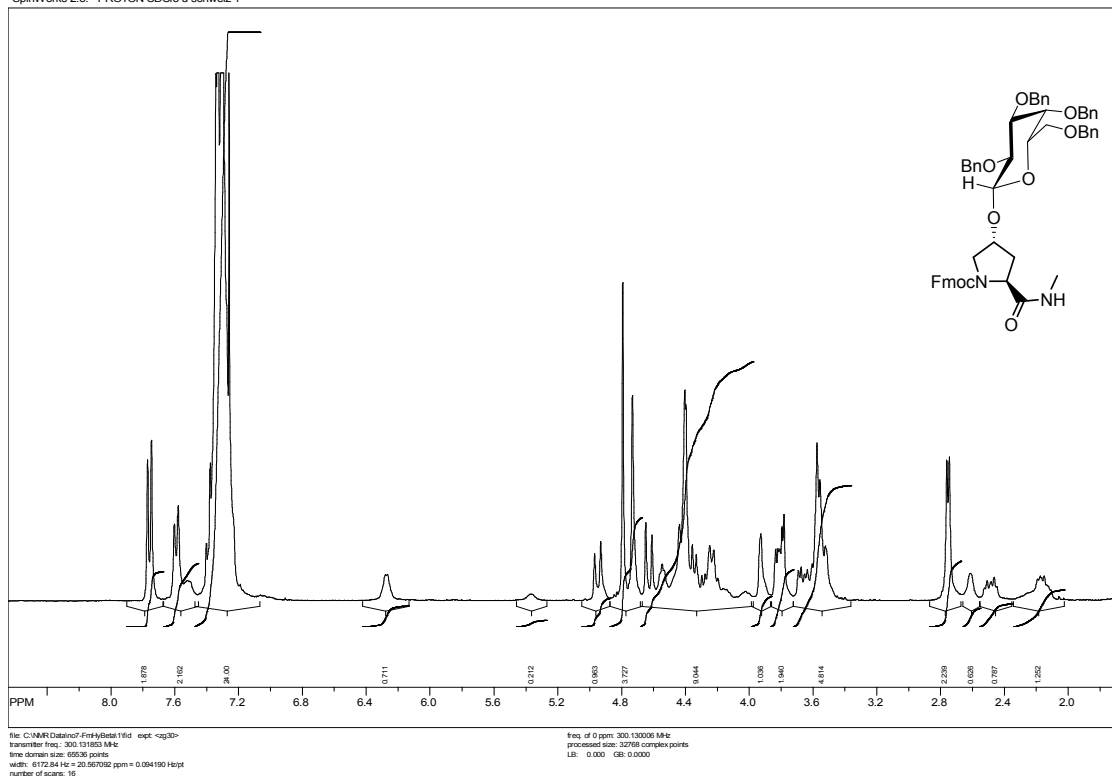
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



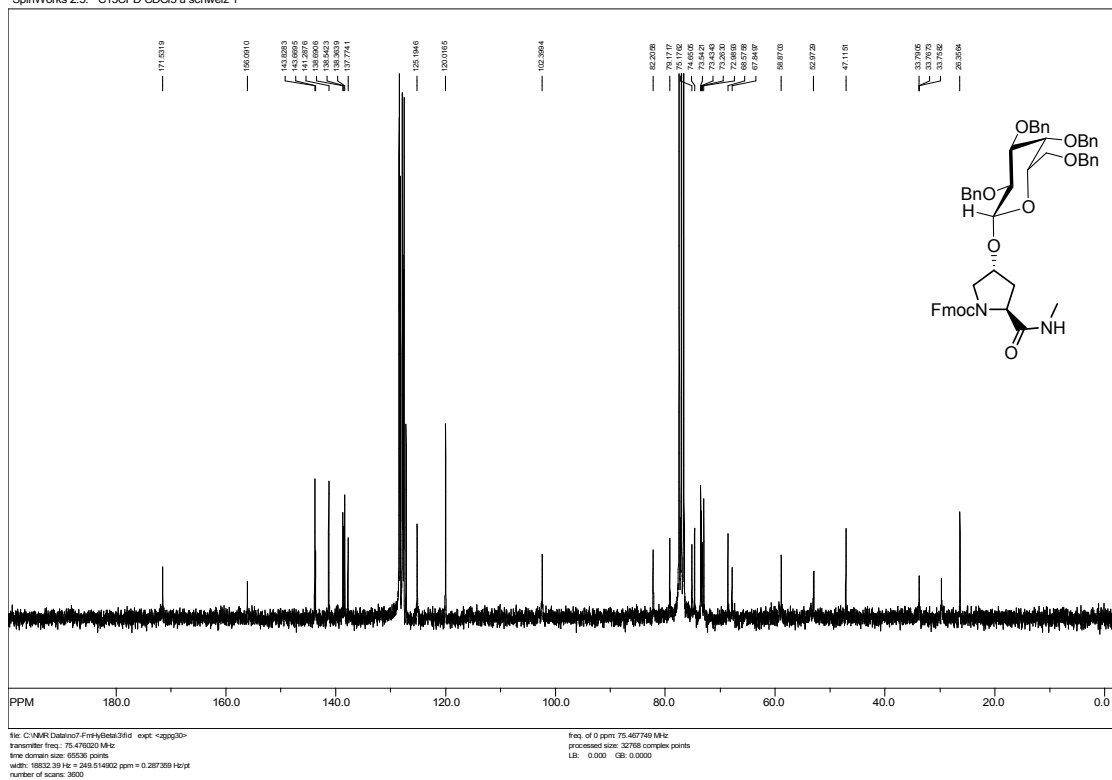
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



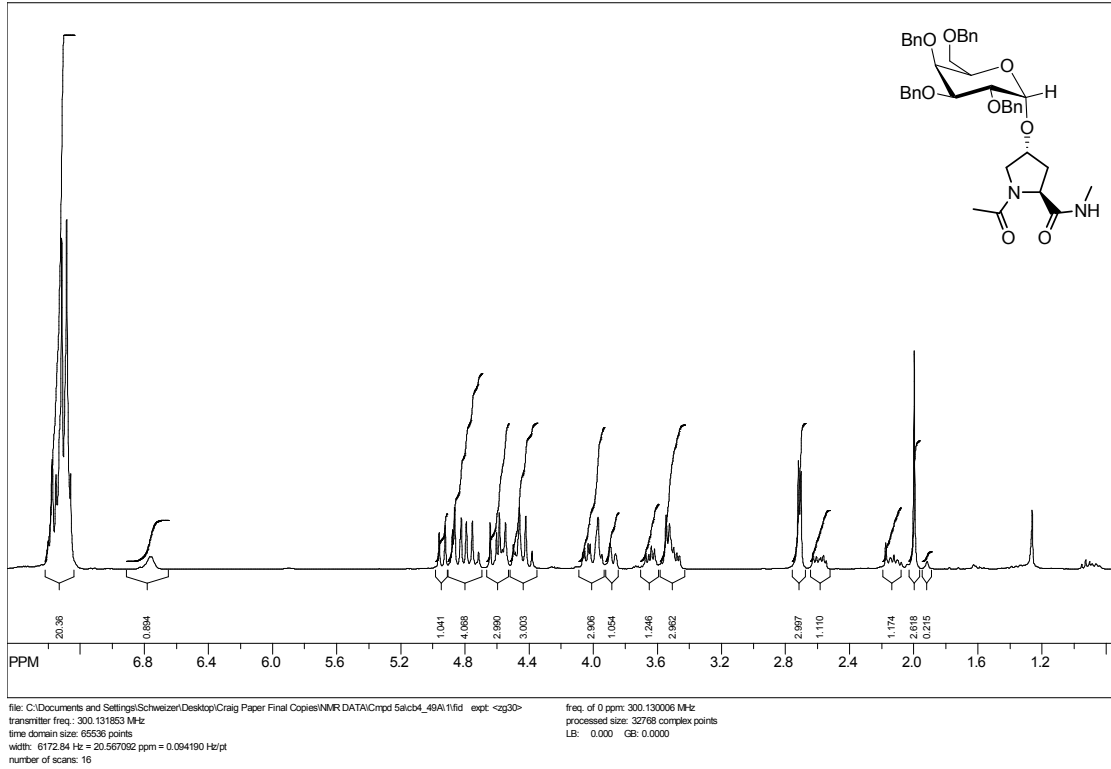
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



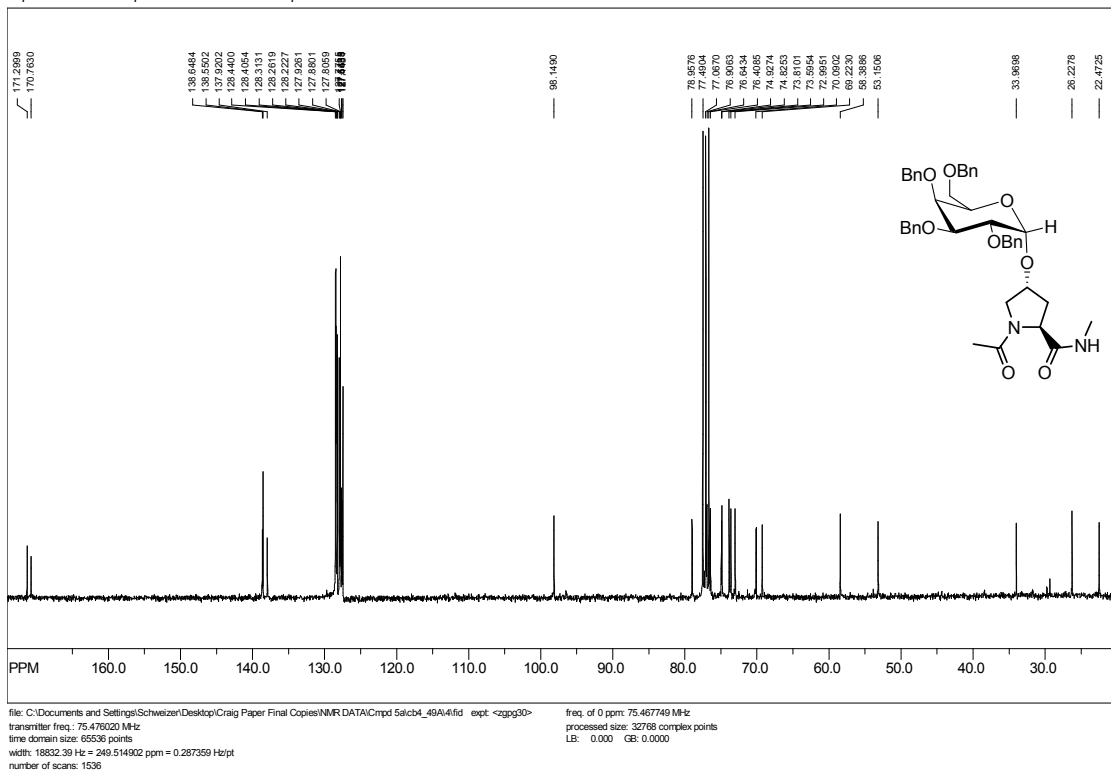
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



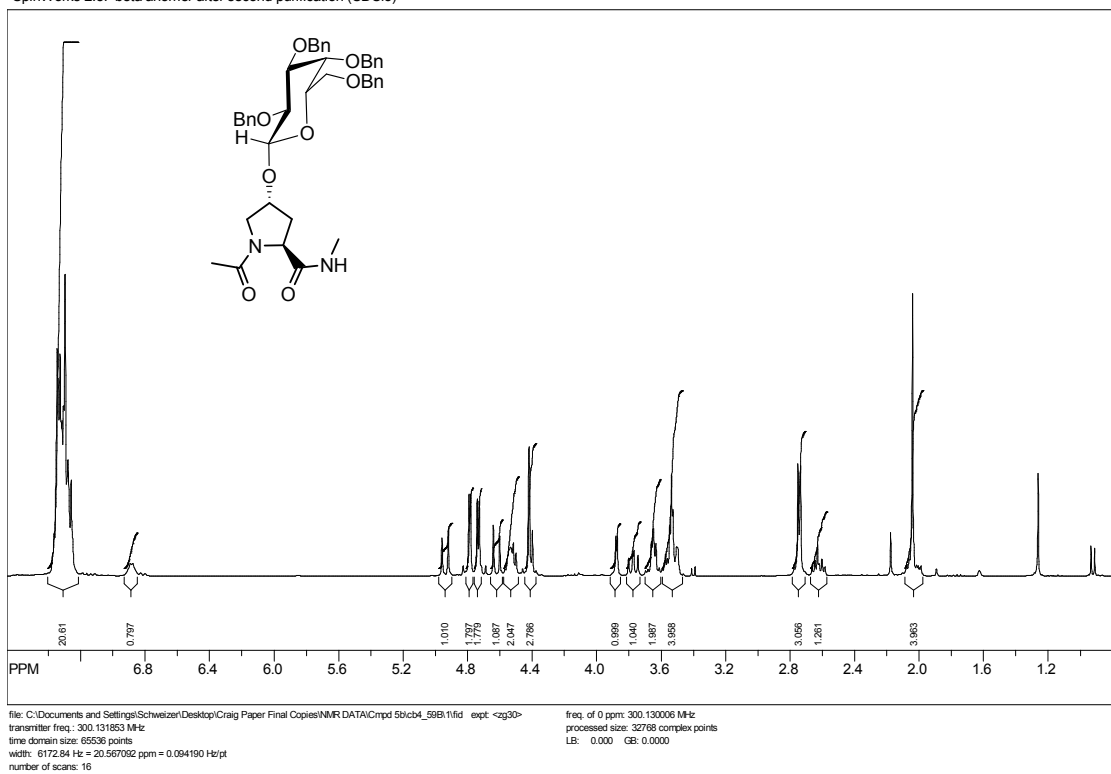
SpinWorks 2.5: alpha anomer after second purification



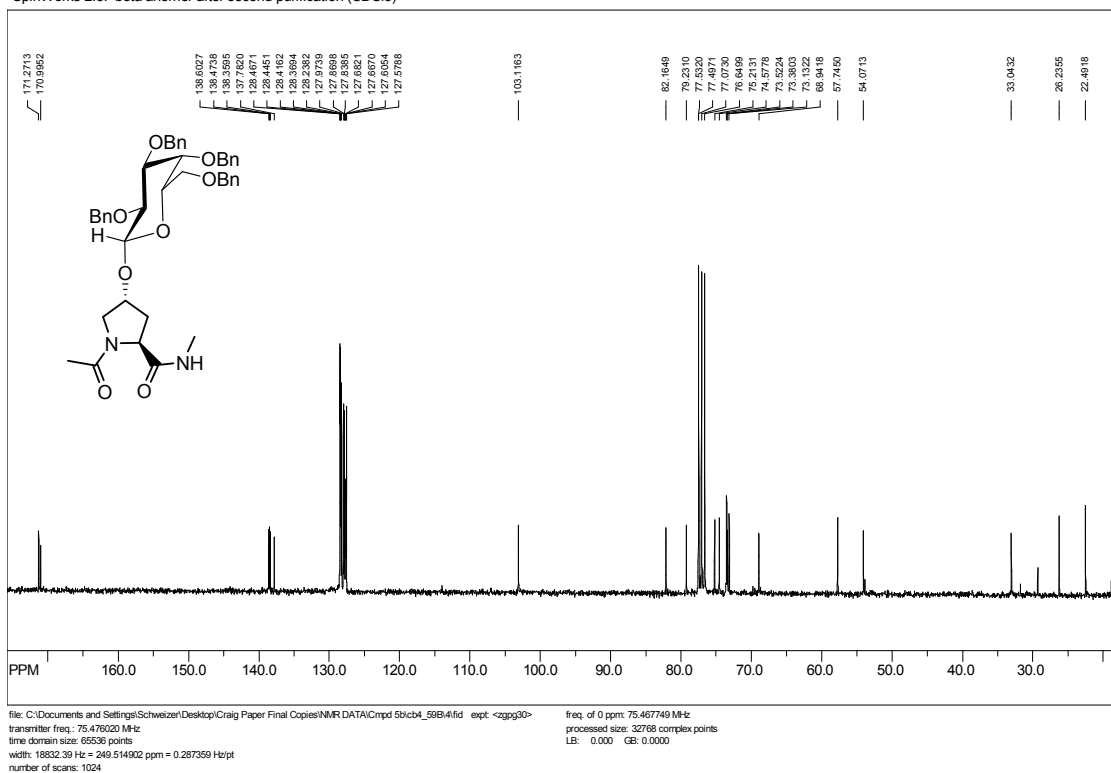
SpinWorks 2.5: alpha anomer after second purification



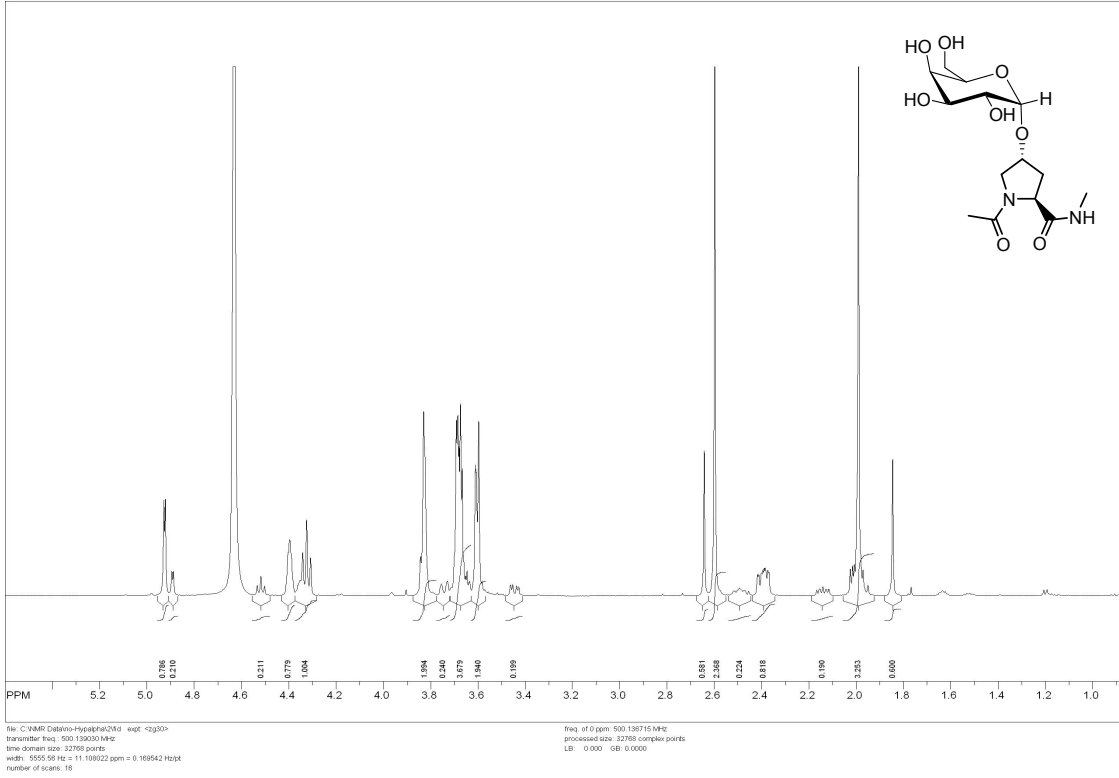
SpinWorks 2.5: beta anomer after second purification (CDCl3)



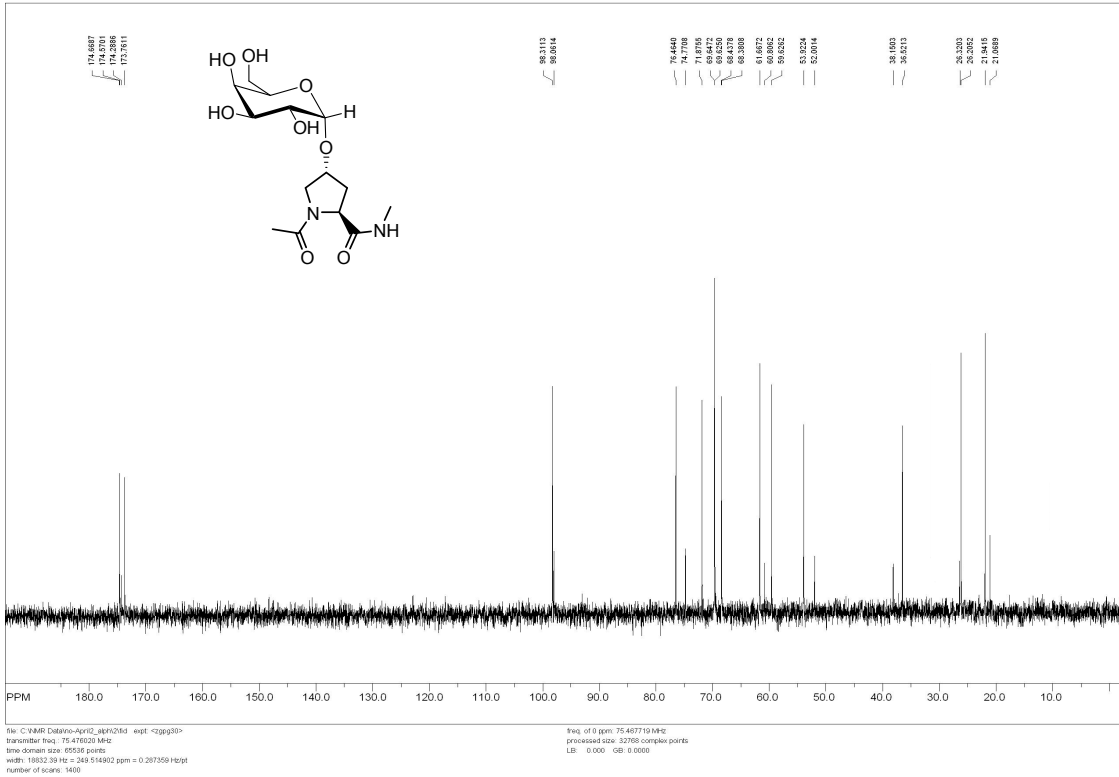
SpinWorks 2.5: beta anomer after second purification (CDCl3)



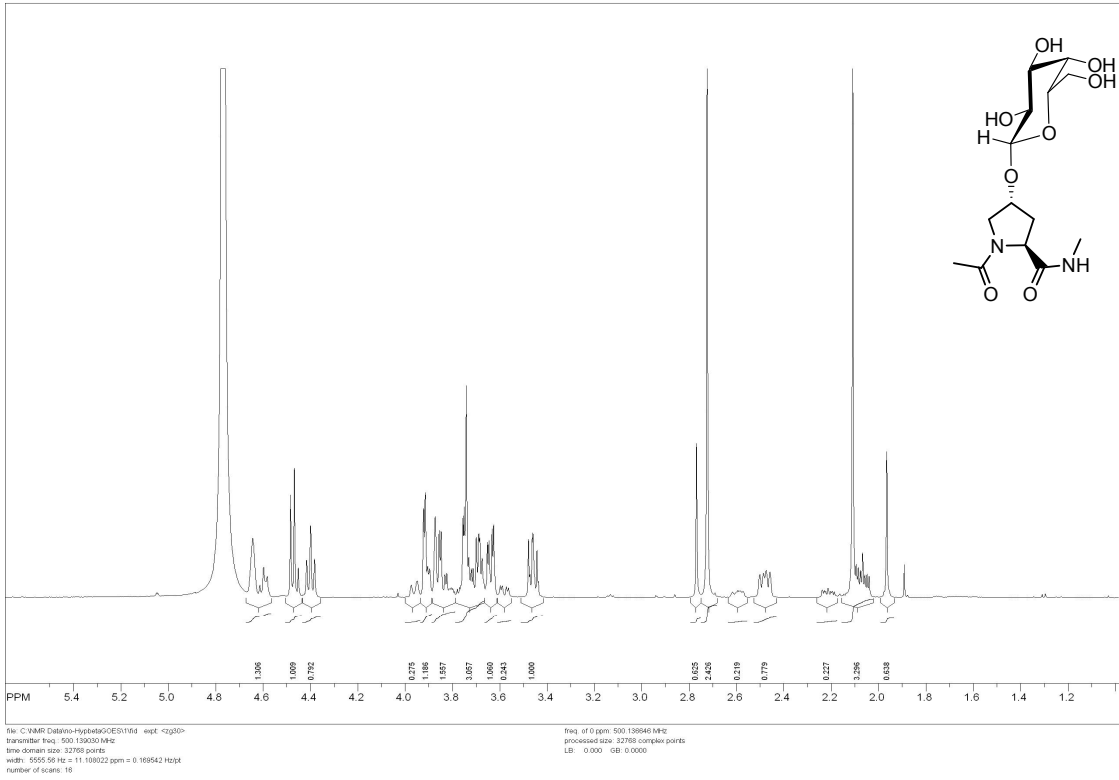
SpinWorks 2.5: standard 1-H survey parameters, AMX500, D2O



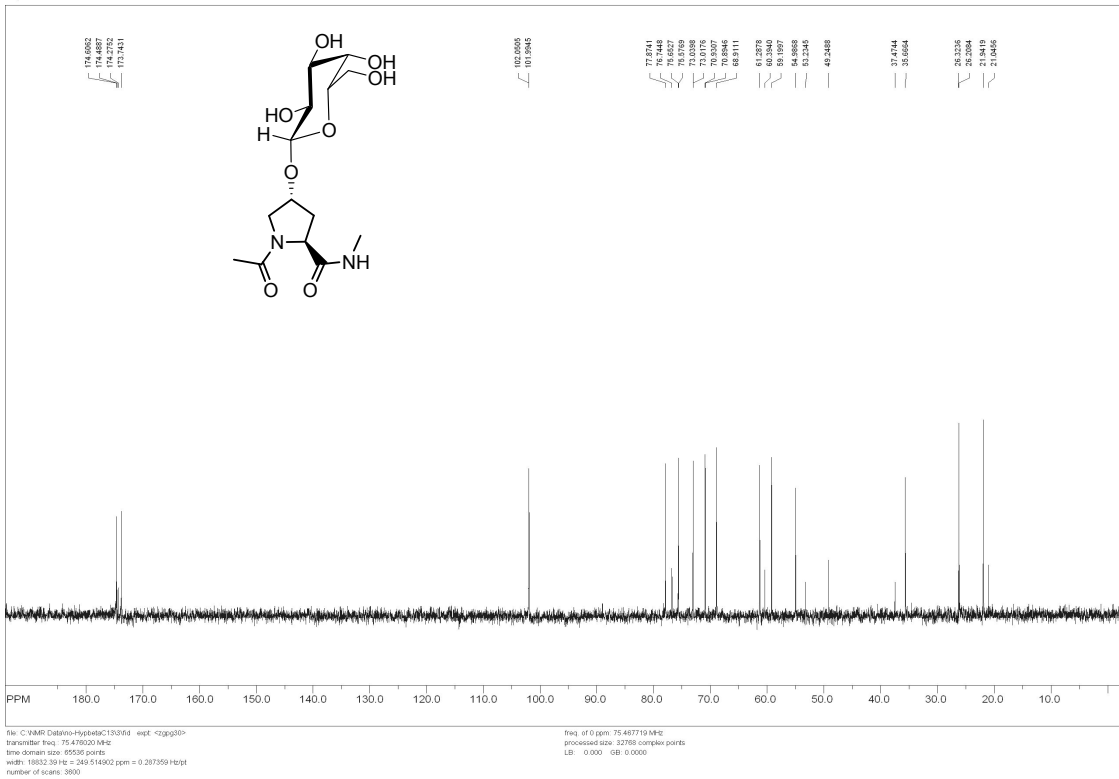
SpinWorks 2.5: C13CPD D2O u schweiz 1



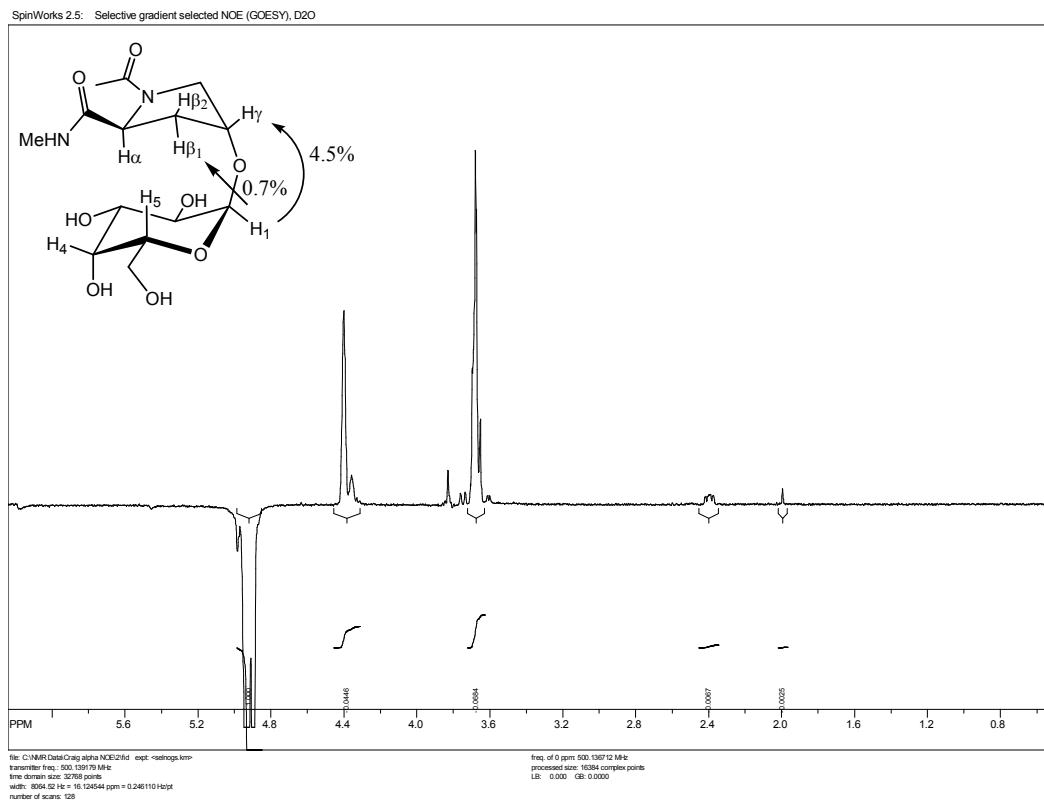
SpinWorks 2.5: standard 1-H survey parameters, AMX500, D2O



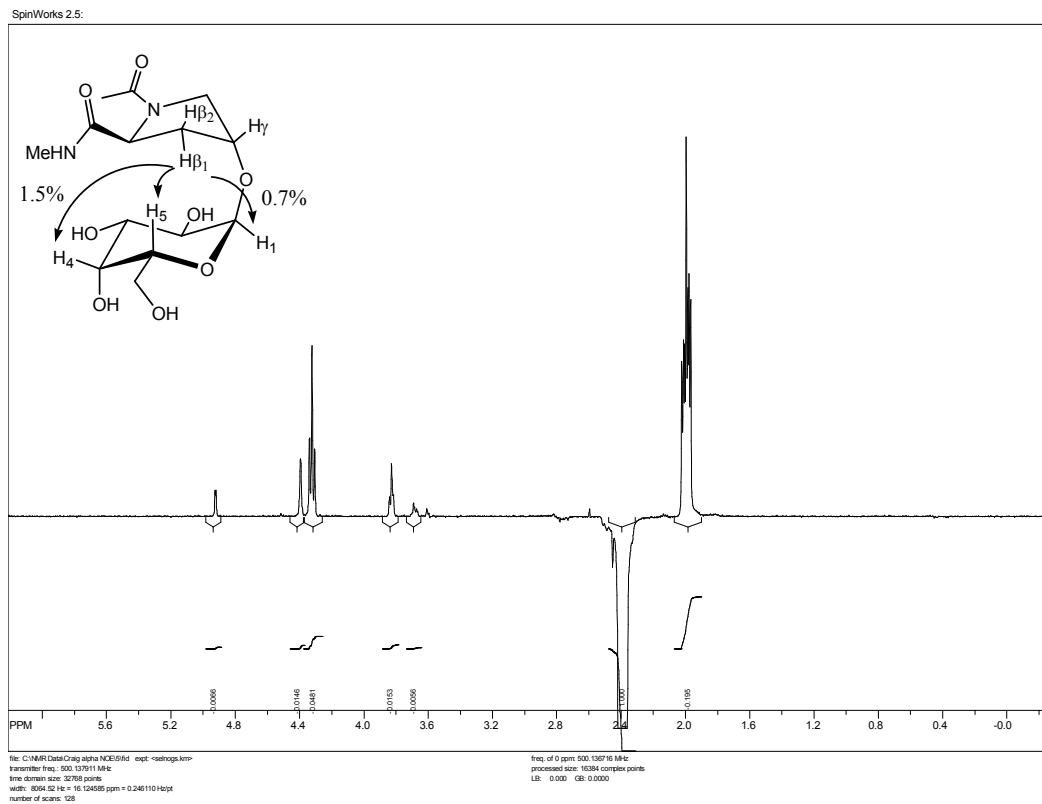
SpinWorks 2.5: C13CPD D2O u schweiz 1



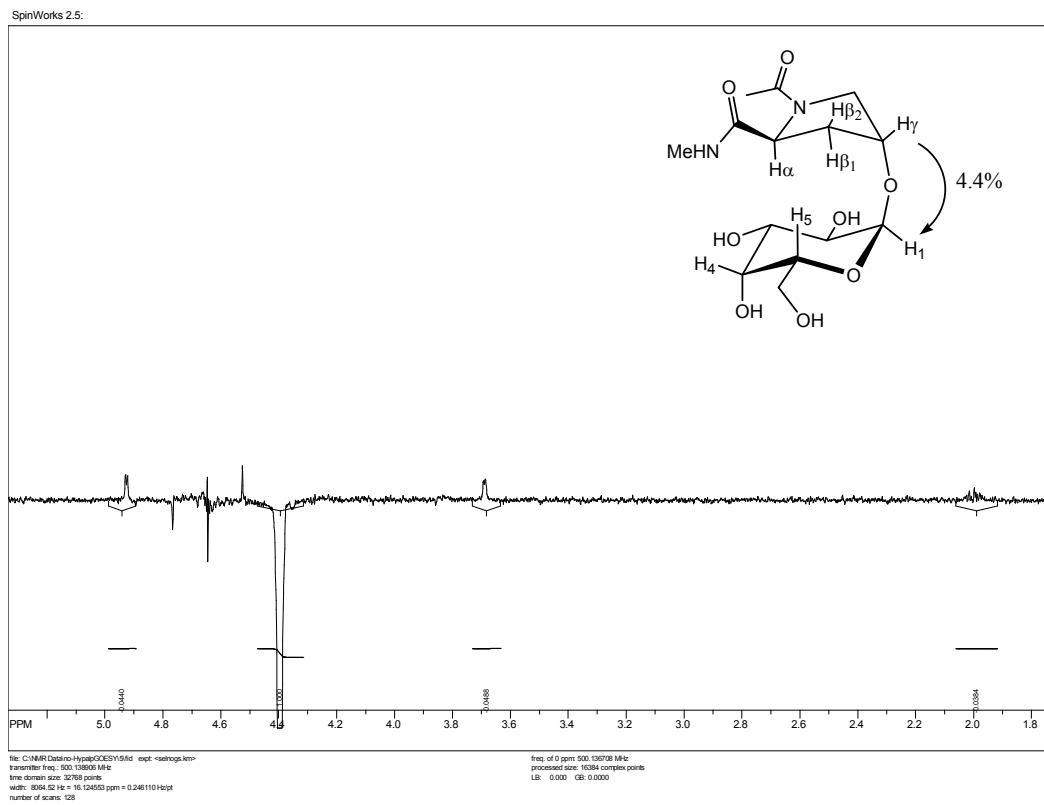
A one-dimensional GOESY experiment in D₂O irradiating the H₁^{trans} (4.93 ppm, 0.8H) [H₁^{cis}, 4.89 ppm, 0.2H] in **4a** showed inter-proton effect to Pro_γ^{trans} (4.40 ppm, 0.8H) (4.5%), Pro_{β1}^{trans} (2.39 ppm, 0.8H) (0.7%) and the *N*-acyl methyl singlet –COCH₃^{trans} (1.99 ppm, 2.4H) (0.3%).



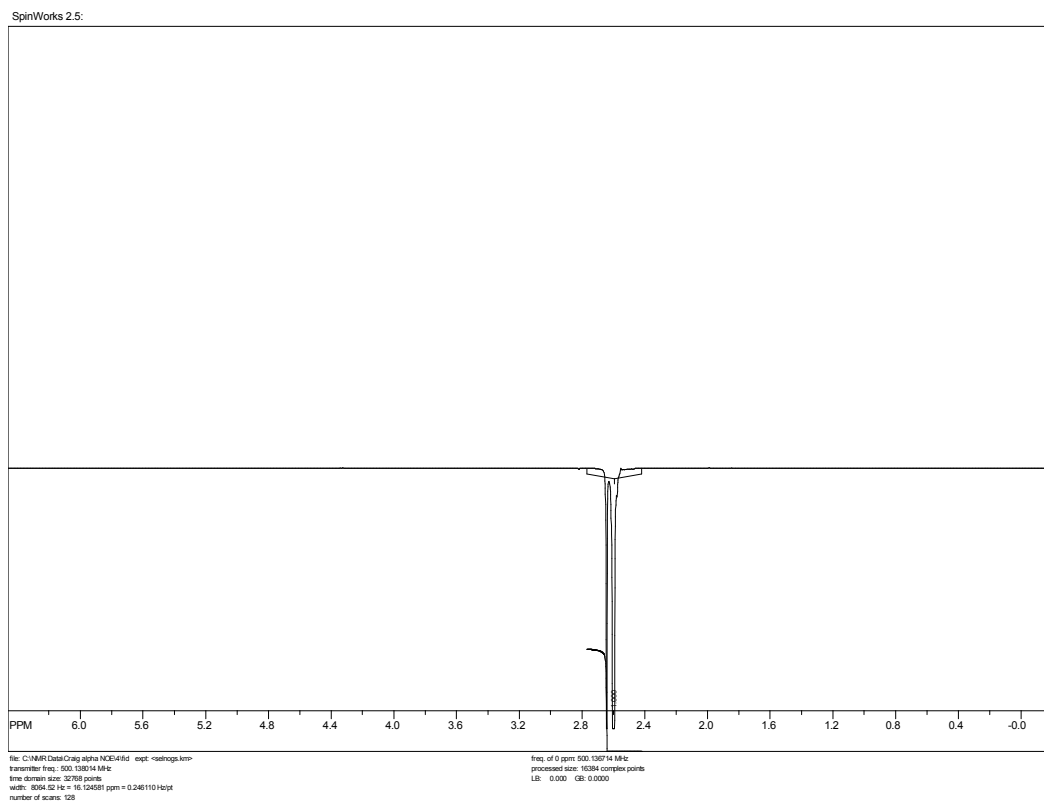
A one-dimensional GOESY experiment in D₂O irradiating the Pro_{β1}^{trans} (2.39 ppm, 0.8H) in **4a** showed inter-proton effect to H₁^{trans} (4.93 ppm, 0.8H) (0.7%) and the H₄^{cis,trans}/H₅^{cis,trans} peak (3.82 ppm, 2H) (1.5%).



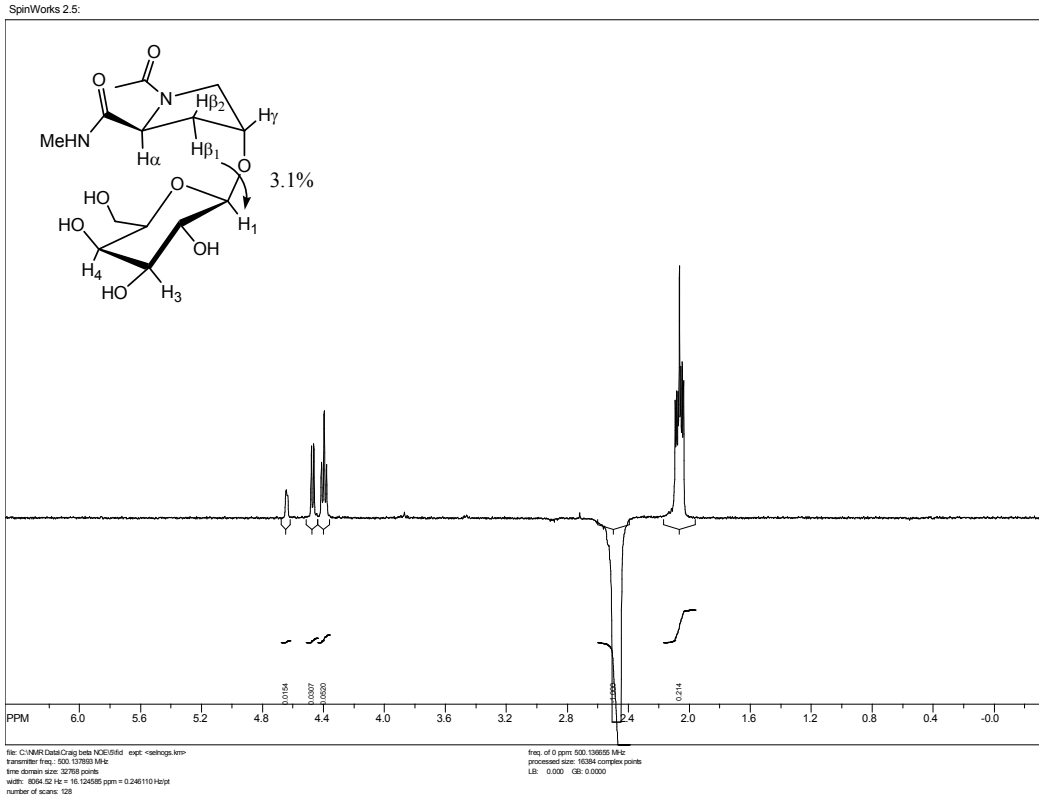
A one-dimensional GOESY experiment in D₂O irradiating the Pro_γ^{trans} (4.40 ppm, 0.8H) in **4a** showed inter-proton effect to H₁^{trans} (4.93 ppm, 0.8H) (4.4%).



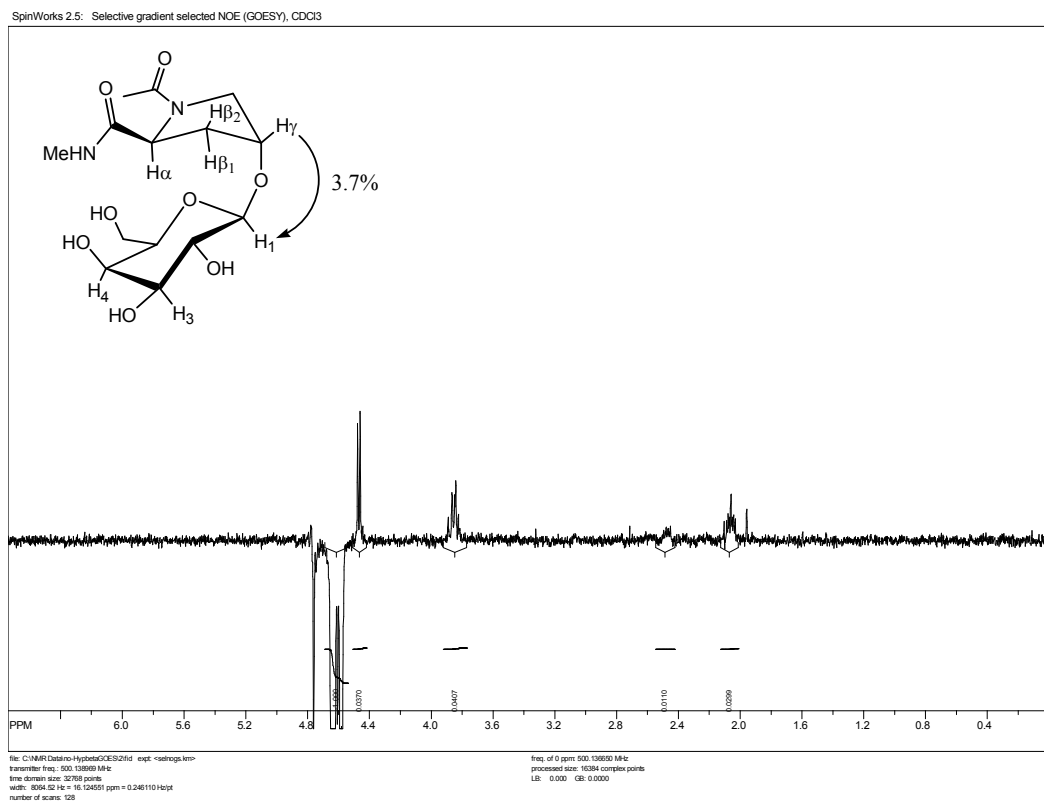
A one-dimensional GOESY experiment in D₂O irradiating the *N*-methylamide singlet -NHCH₃^{trans} (2.60 ppm, 2.4H) [-NHCH₃^{cis}, 2.64 ppm, 0.6H] in **4a** showed no inter-proton effects.



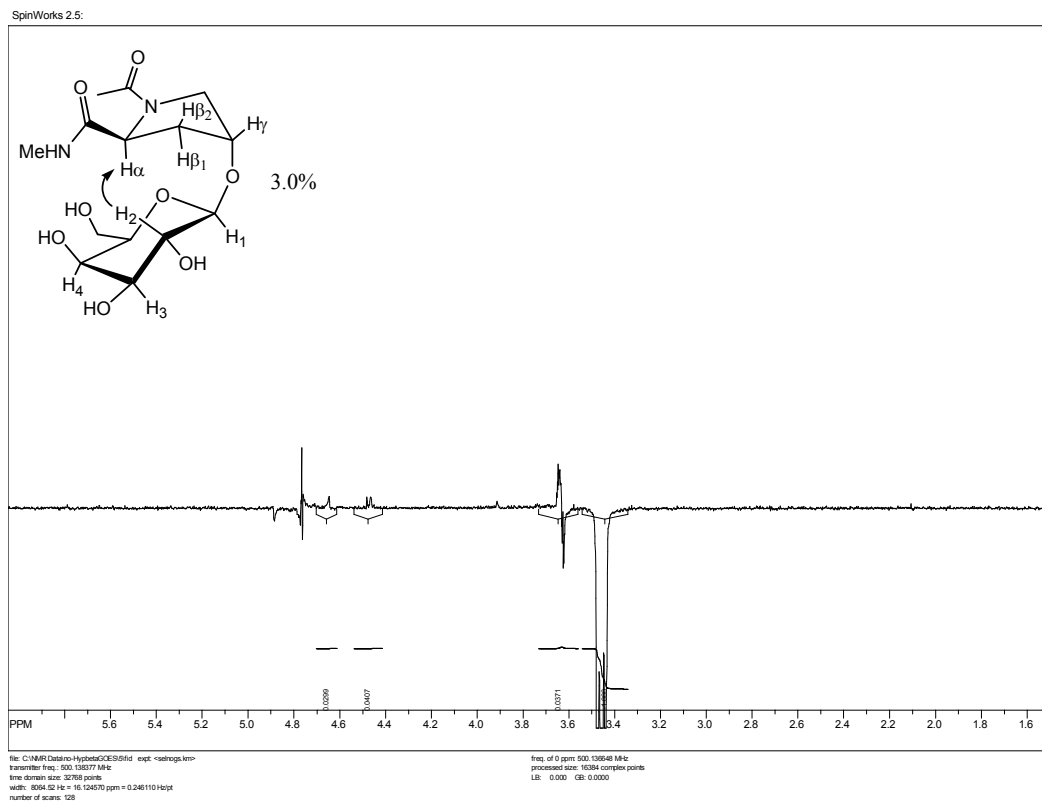
A one-dimensional GOESY experiment in D₂O irradiating the Pro_{β₁}^{trans} (2.46 ppm, 0.8H) in **4b** showed inter-proton effect to H₁^{trans} (4.48 ppm, 0.8H) (3.1%).



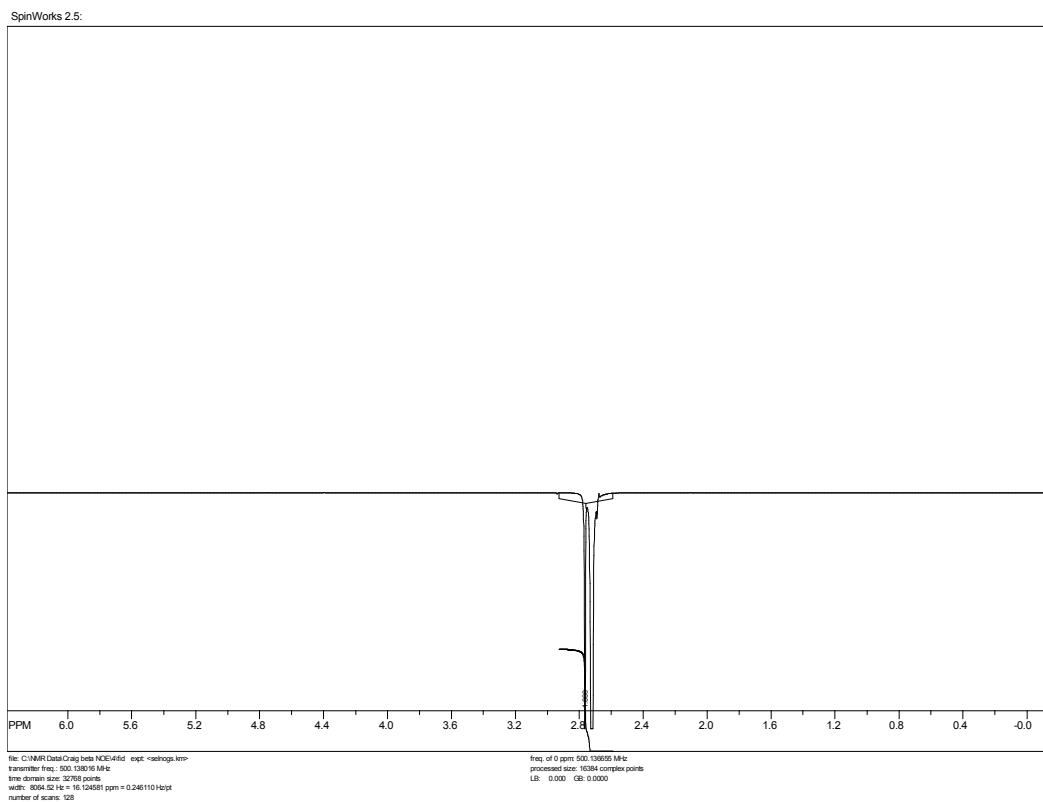
A one-dimensional GOESY experiment in D₂O irradiating the Pro_γ^{trans} (4.65 ppm, 0.8H) [Pro_γ^{cis}, 4.59 ppm 0.2H] in **4b** showed inter-proton effect to H₁^{trans} (4.48 ppm, 0.8H) (3.7%).



A one-dimensional GOESY experiment in D₂O irradiating the H₂^{trans} (3.46 ppm, 0.8H) in **4b** showed inter-proton effect to Pro_α^{trans} (4.65 ppm, 0.8H) (3.0%).



A one-dimensional GOESY experiment in D₂O irradiating the *N*-methylamide singlet -NHCH₃^{trans} (2.72 ppm, 2.4H) [-NHCH₃^{cis}, 2.77 ppm, 0.6H] in **4b** showed no inter-proton effects.



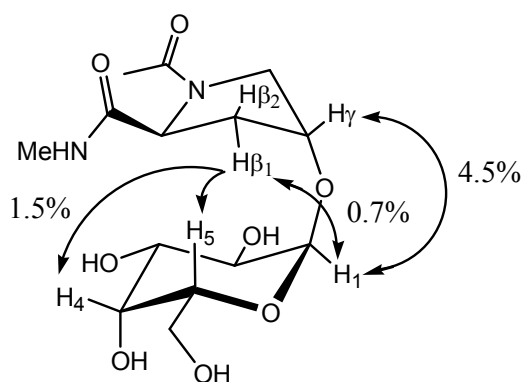


Figure 10.13: Summary of nOe Interactions Between the Prolyl and Galactosyl Rings in

4a

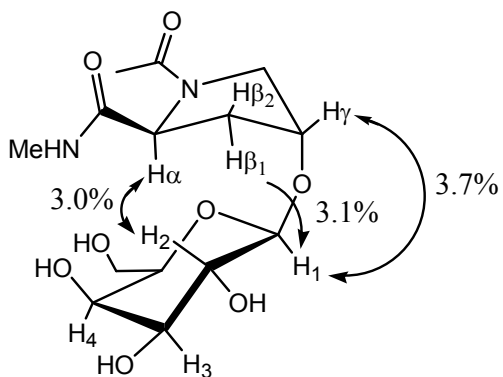


Figure 10.14: Summary of nOe Interactions Between the Prolyl and Galactosyl Rings in

4b

Magnetization Transfer NMR Experiments

Samples of **1-4b** were prepared by dissolving compounds in D₂O at concentrations between 0.05 M and 0.1 M. All experiments were performed on a Bruker AMX500 NMR spectrometer equipped with a triple resonance (¹H, ¹³C, ¹⁵N) gradient inverse probehead.

Selective inversion of the *N'*-methylamide singlet of the *trans* isomer was done using an 80 msec pulse centered on resonance. Relaxation delays of 20 s and inversion recovery delays between 1 ms and 15 s were used. Experiments were performed over several temperatures from 333.15 to 353.15 K for each compound. The temperature was calibrated using an ethylene glycol standard. For each compound at an individual temperature, inversion-recovery data were collected with selective inversion of the ¹H *N'*-methylamide singlet of the major *trans* peak. ¹H NMR spectroscopy was used in preference to ¹³C NMR since over the course of an experiment, the signal/noise ratio is much higher for ¹H than ¹³C, and heating effects caused by decoupling for ¹³C causes uncertainty in the temperature of the sample.

The time-dependent magnetization transfers of the *cis* ($M_c(t)$) and *trans* ($M_t(t)$) NMR signals as a function of the inversion transfer time (t) were simultaneously fit to equations 1 and 2 (Alger and Prestegard, 1977; Mariappan and Rabenstein, 1992) below for compounds **1-4b** using *Mathematica* (v. 5.0). In the following pulse sequence, the ¹H *trans* resonance is selectively inverted using a shaped pulse. Its recovery during t is determined by its intrinsic T_{1t} , magnetization transfer to and from the *cis* resonance, and the T_{1c} of the *cis* resonance:

$$\pi(x) \text{sel} \text{-----} t \text{-----} \pi/2(x, y, -x, -y) \text{---acquire}$$

The resonances of the *trans* and *cis* isomers show the following time dependences:

$$M_t(t) = (c_1)(\tau_t)(\lambda_1 + 1/\tau_{1c})\exp(\lambda_1 * t) + (c_2)(\tau_c)(\lambda_2 + 1/\tau_{1c})\exp(\lambda_2 * t) + M_{c\infty} \text{-----} 1$$

$$M_t(t) = (c_1)\exp(\lambda_1 * t) + (c_2)\exp(\lambda_2 * t) + M_{t\infty} \text{-----} 2$$

T_{1c} and T_{1t} are the longitudinal relaxation times of the resonances in the absence of exchange.

τ_c and τ_t are the lifetimes of the *cis* and *trans* conformers and k_{ct} and k_{tc} are the corresponding rate constants.

τ_{1c} and τ_{1t} are the effective relaxation times of the *cis* and *trans* resonances when relaxation and exchange are both occurring and are defined below in terms of T_{1c} and τ_c , T_{1t} and τ_t .

λ_1 and λ_2 are related to the time constants τ_c , τ_t , τ_{1c} , and τ_{1t} , and are defined below.

c_1 and c_2 are defined below.

$M_{c\infty}$ and $M_{t\infty}$ are determined experimentally from the magnetization measured after 5 T_1 periods for the *cis* and *trans* resonances, respectively.

Mathematica then calculates τ_t from τ_c , $M_{c\infty}$, and $M_{t\infty}$ as: $\tau_t = \tau_c * M_{t\infty}/M_{c\infty}$

Thus,

$$k_{ct} = 1/\tau_c$$

$$k_{tc} = 1/\tau_t$$

$$K_{eq} = M_{t\infty}/M_{c\infty}$$

$$\tau_{1c} = (T_{1c} * \tau_c) / (\tau_c + T_{1c})$$

$$\tau_{1t} = (T_{1t} * \tau_t) / (\tau_t + T_{1t})$$

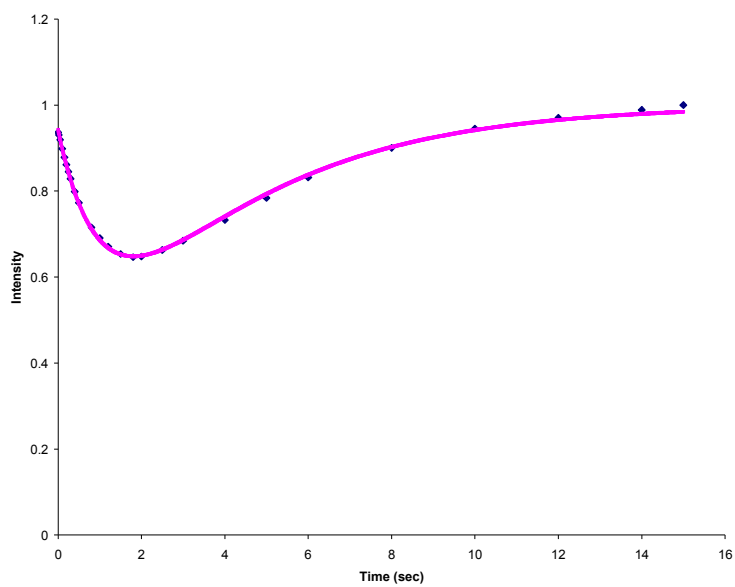
$$\lambda_1 = (0.5) * (-((1/\tau_{1c}) + (1/\tau_{1t})) + (((((1/\tau_{1c}) + (1/\tau_{1t})) / 2) - 4) * (((1/\tau_{1c}) * (1/\tau_{1t})) - ((1/\tau_c) * (1/\tau_t)))))) / 2$$

$$\lambda_2 = (0.5) * (-((1/\tau_{1c}) + (1/\tau_{1t})) - (((((1/\tau_{1c}) + (1/\tau_{1t})) / 2) - 4) * (((1/\tau_{1c}) * (1/\tau_{1t})) - ((1/\tau_c) * (1/\tau_t)))))) / 2$$

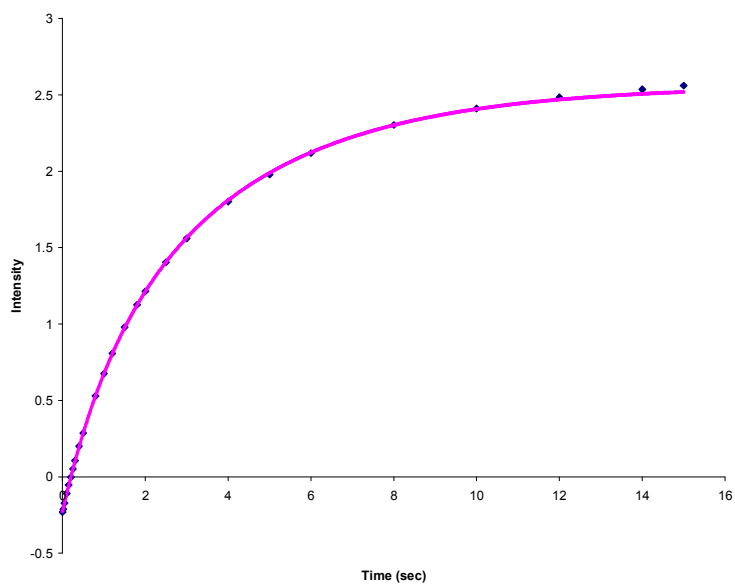
$$c_2 = ((1/(\tau_t)) * (\lambda_1 - \lambda_2)) * ((\tau_t * (\lambda_1 + (1/\tau_{1c}))) * ((M_{0c} - M_{c\infty}) + (M_{t\infty} - M_{0t})))$$

$$c_1 = M_{0c} - M_{c\infty} - c_2$$

Figure 10.15: Inversion recovery of non-inverted (*cis*) *N'*-methamide singlet (A) and inverted (*trans*) *N'*-methamide singlet (B) of **1** at 62.0°C in D₂O

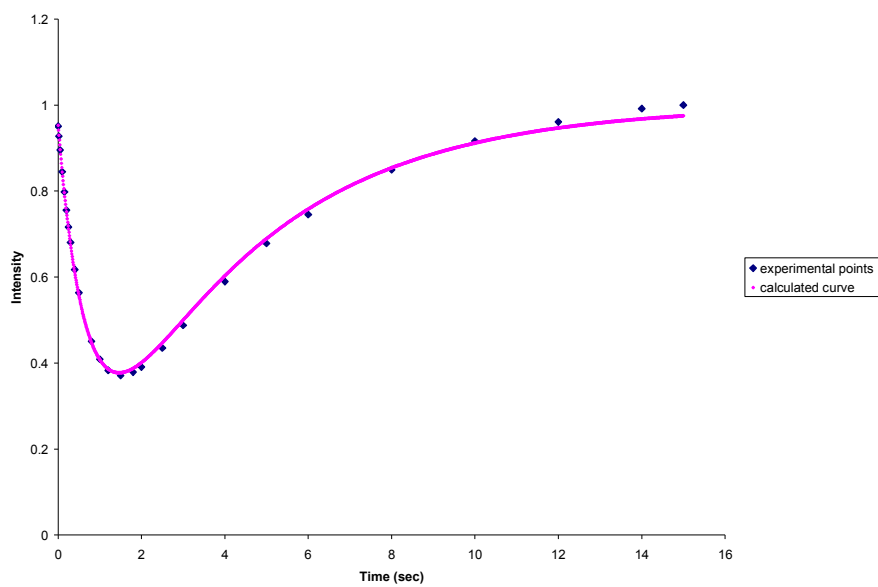


A

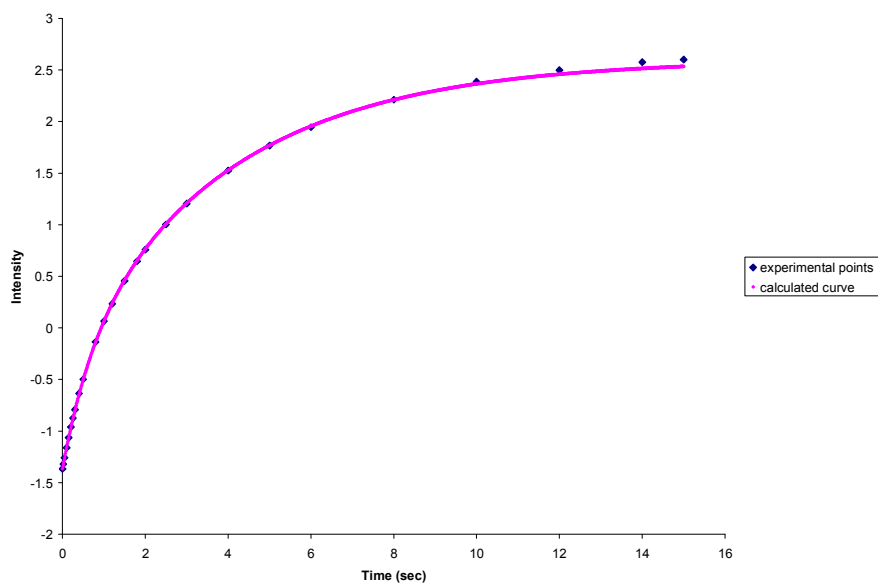


B

Figure 10.16: Inversion recovery of non-inverted (*cis*) *N'*-methyamide singlet (A) and inverted (*trans*) *N'*-methyamide singlet (B) of **1** at 67.3°C in D₂O

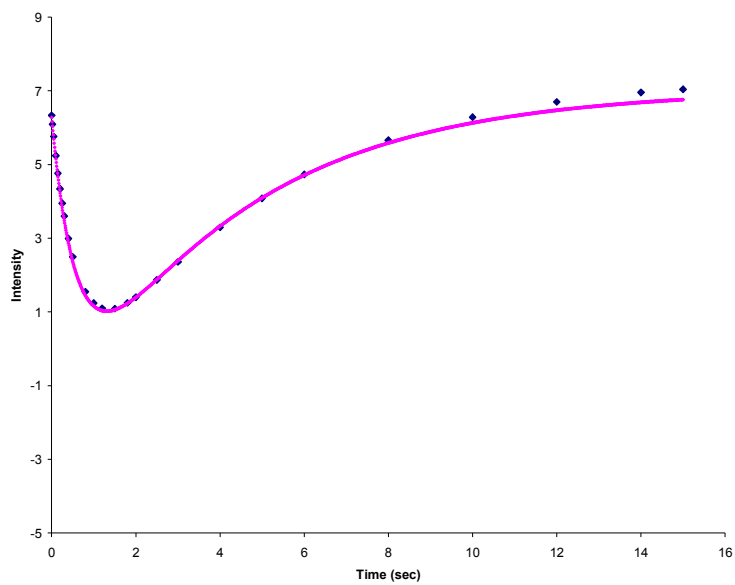


A

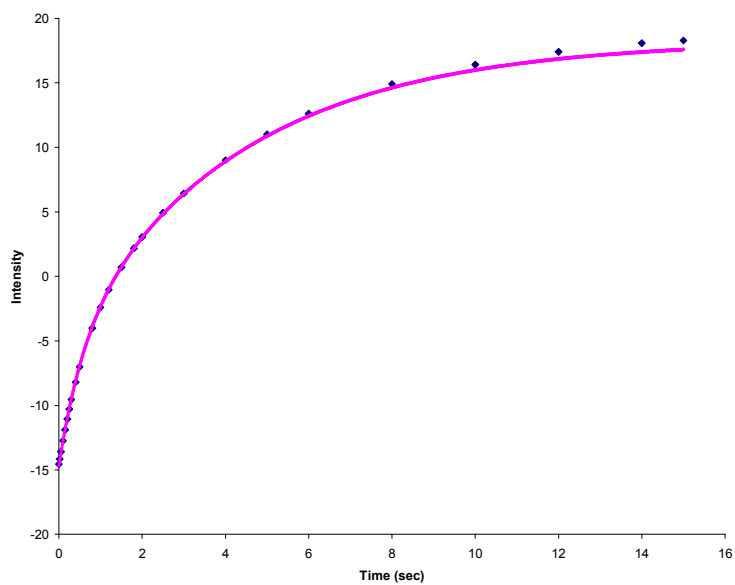


B

Figure 10.17: Inversion recovery of non-inverted (*cis*) *N'*-methanamide singlet (A) and inverted (*trans*) *N'*-methanamide singlet (B) of **1** at 72.6°C in D₂O

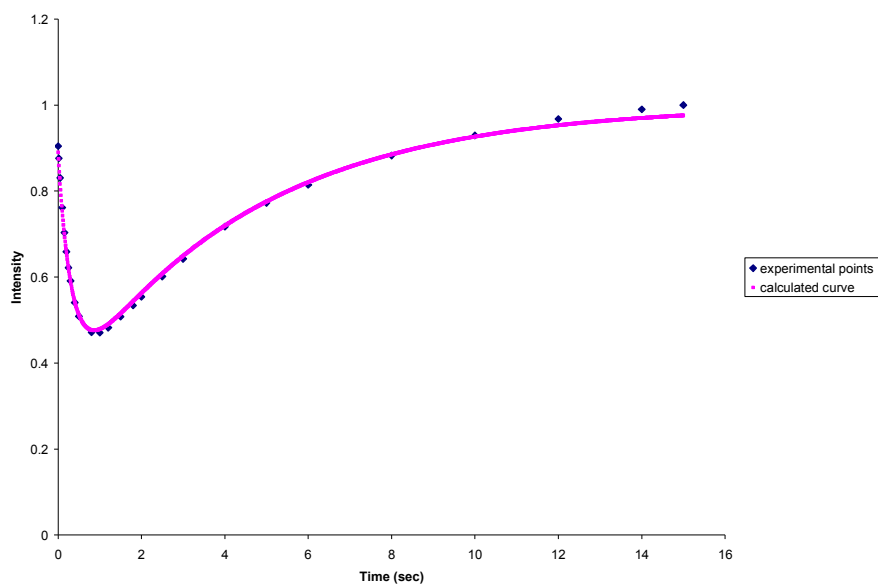


A

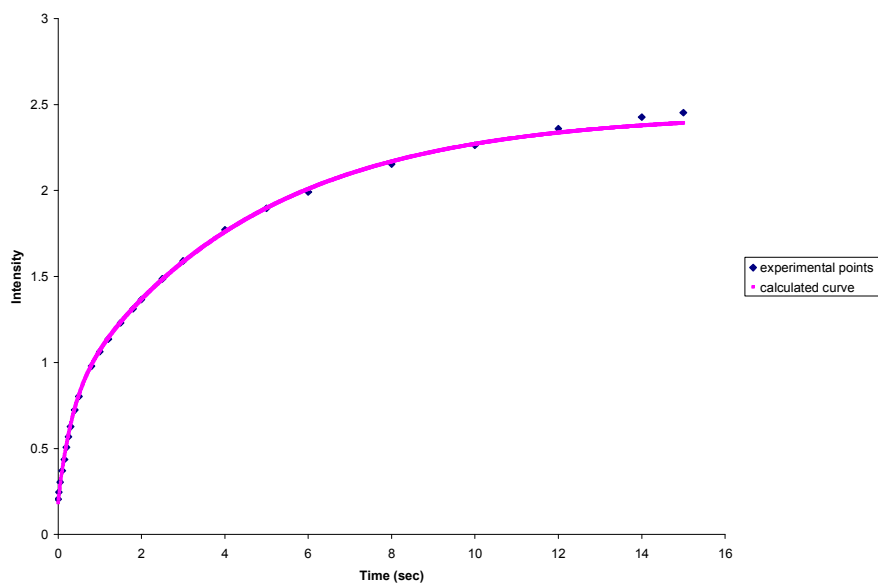


B

Figure 10.18: Inversion recovery of non-inverted (*cis*) *N'*-methanamide singlet (A) and inverted (*trans*) *N'*-methanamide singlet (B) of **1** at 78.0°C in D₂O

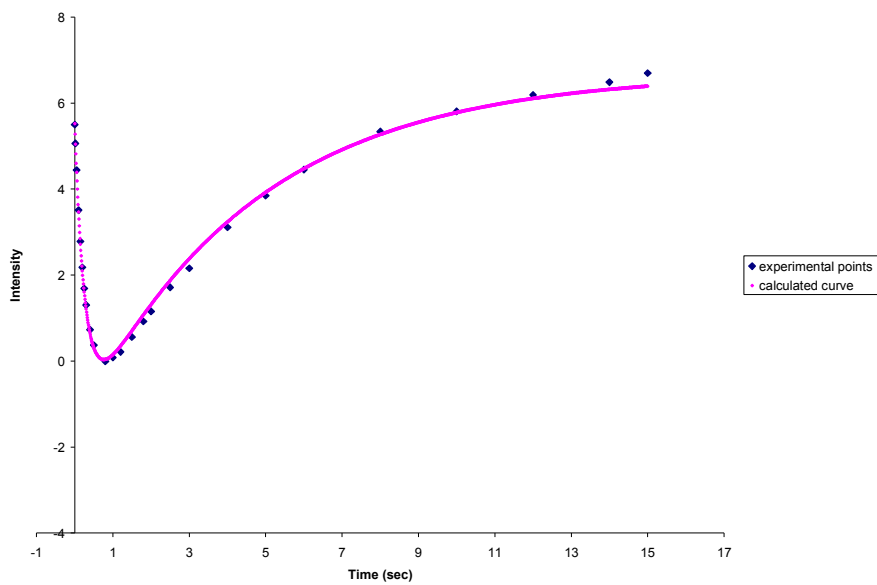


A

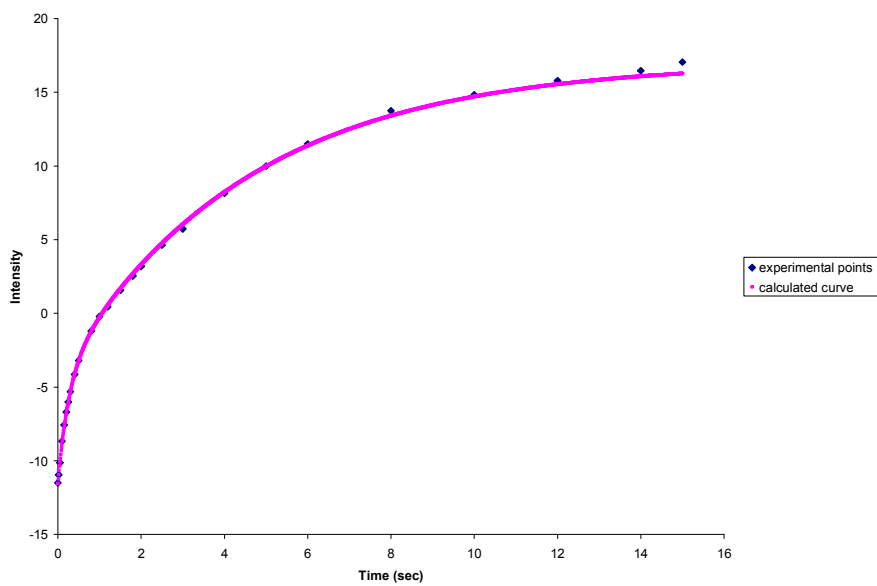


B

Figure 10.19: Inversion recovery of non-inverted (*cis*) *N'*-methanamide singlet (A) and inverted (*trans*) *N'*-methanamide singlet (B) of **1** at 83.0°C in D₂O

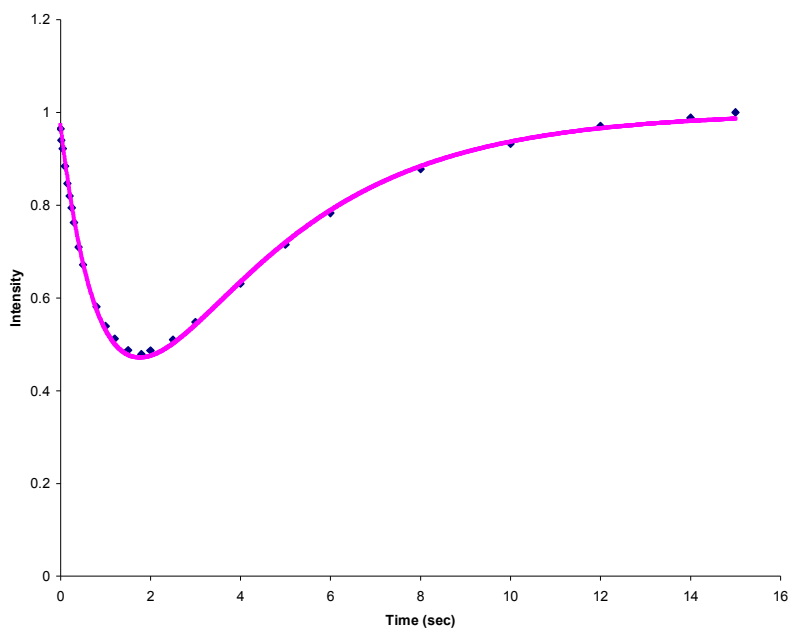


A

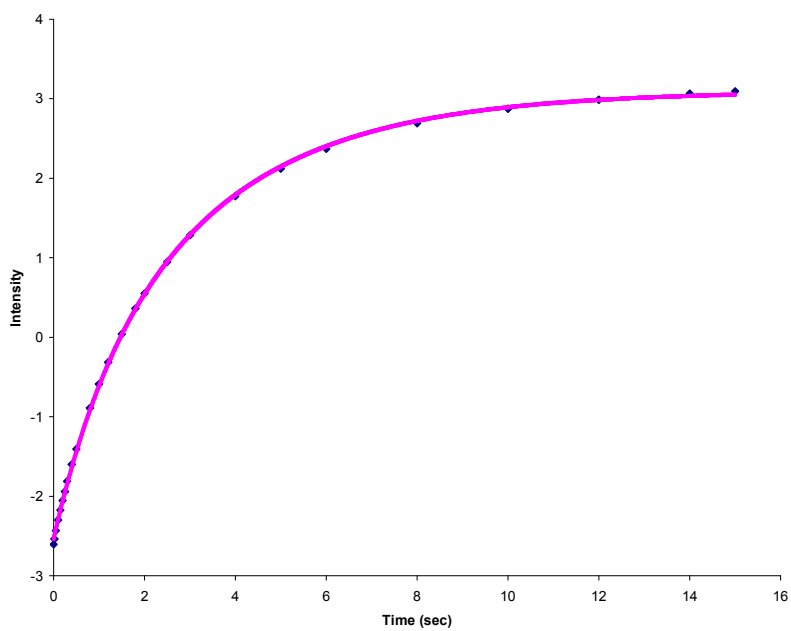


B

Figure 10.20: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **2** at 62.0°C in D₂O

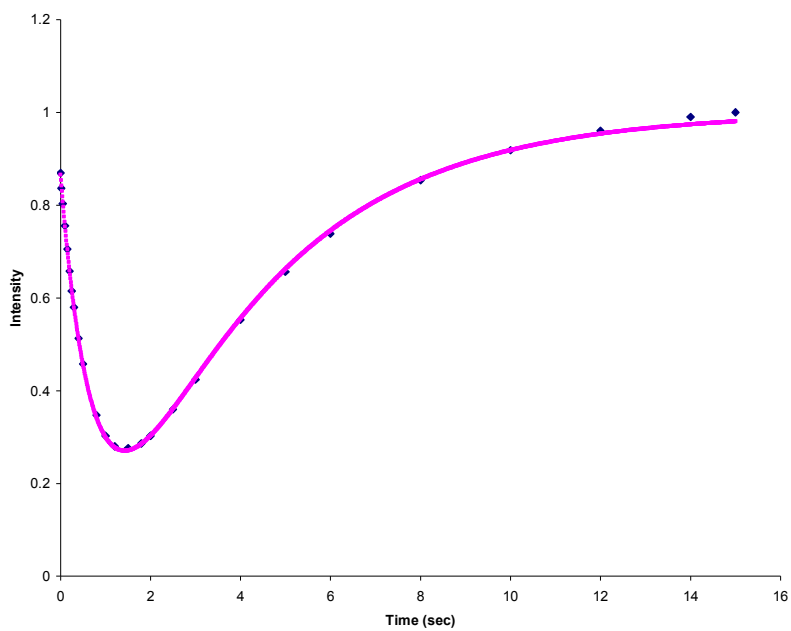


A

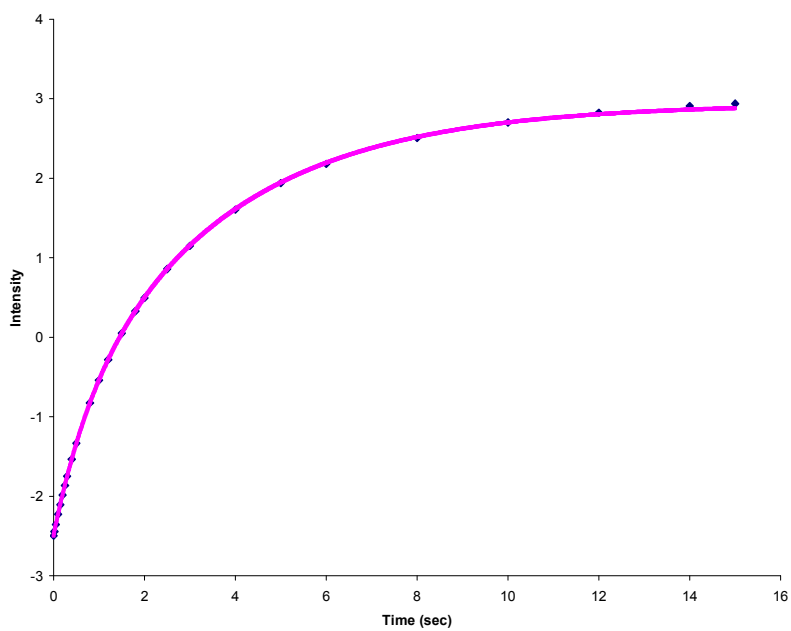


B

Figure 10.21: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **2** at 67.3°C in D₂O

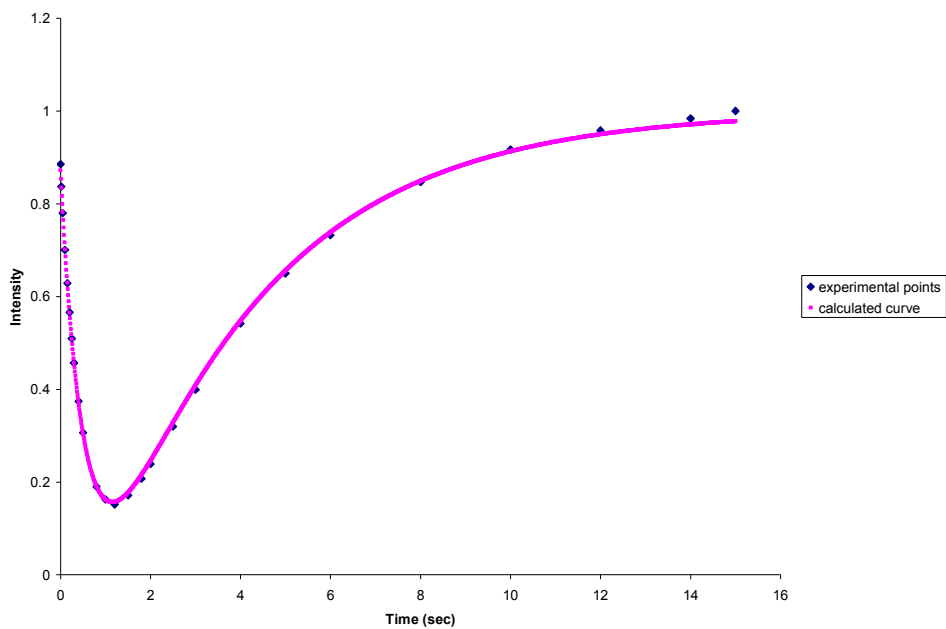


A

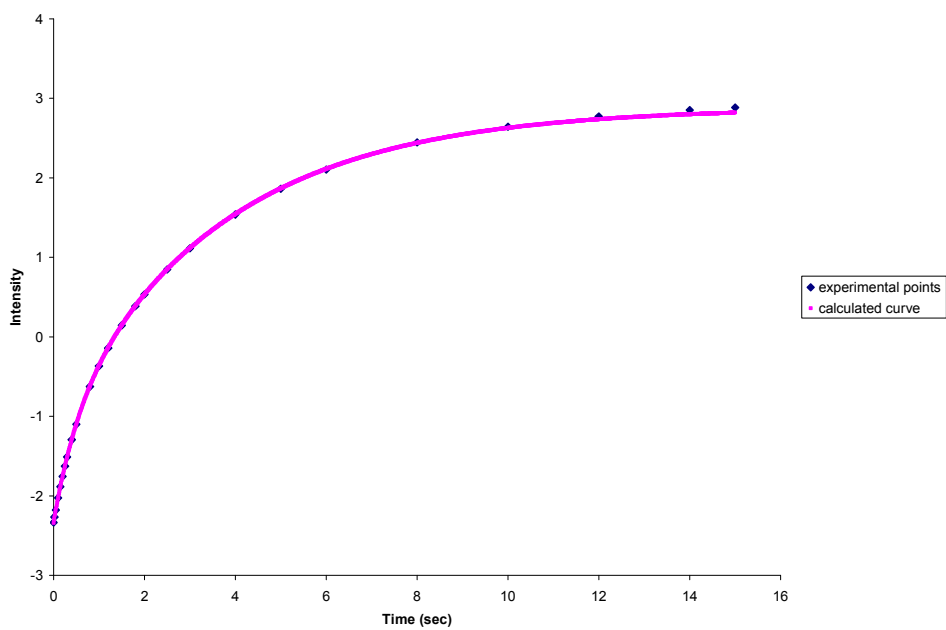


B

Figure 10.22: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **2** at 72.6°C in D₂O

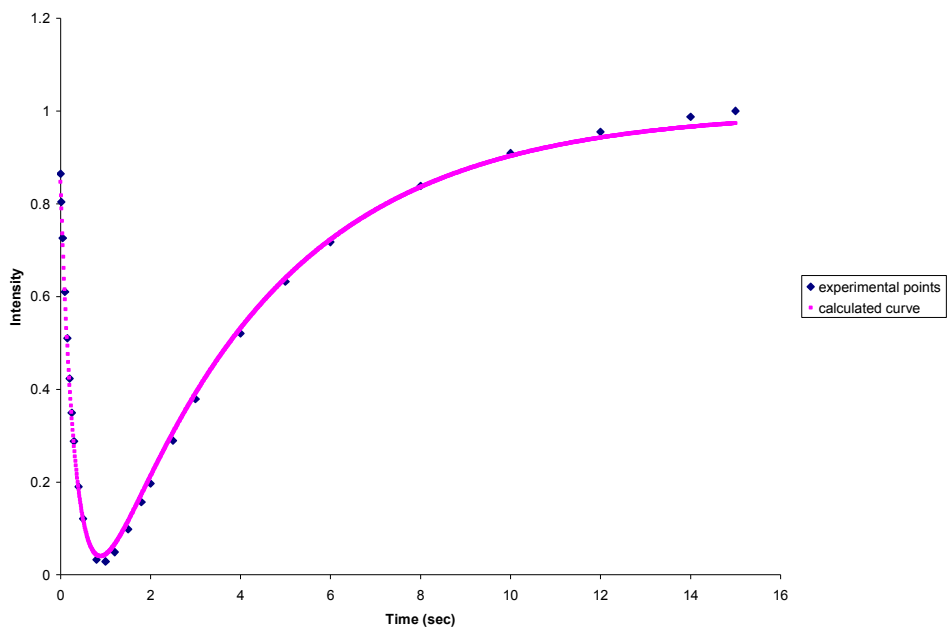


A

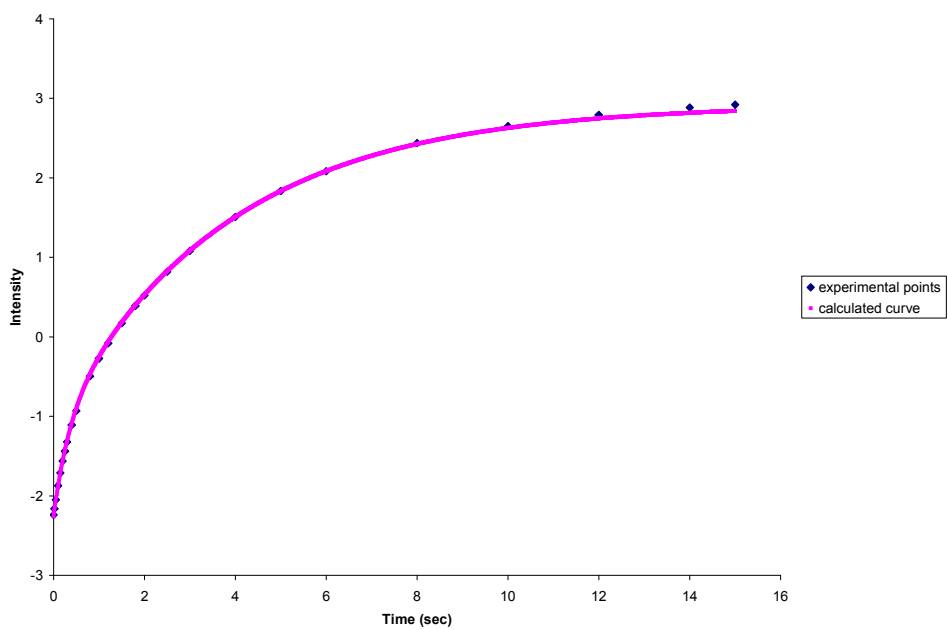


B

Figure 10.23: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **2** at 78.0°C in D₂O

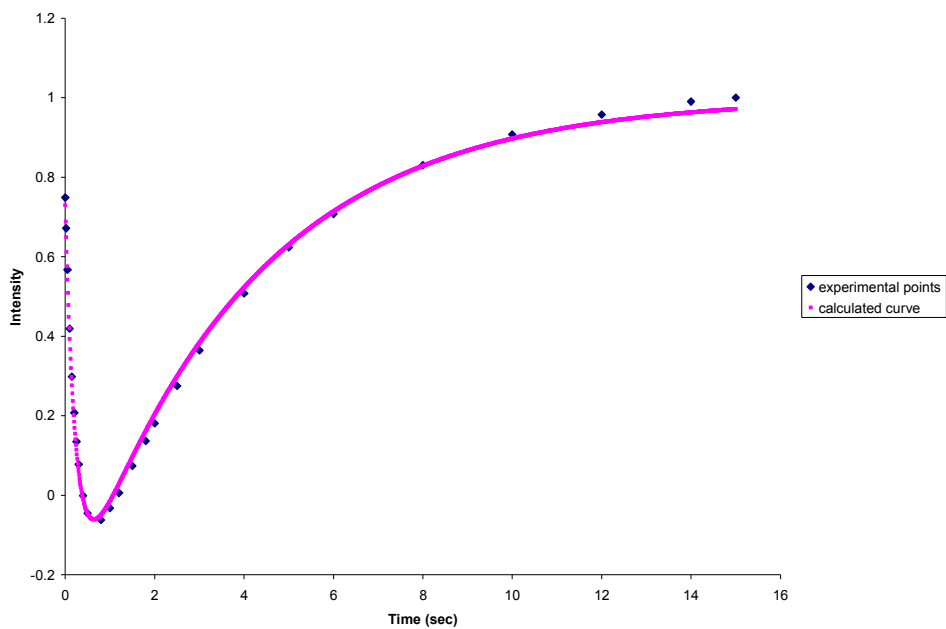


A

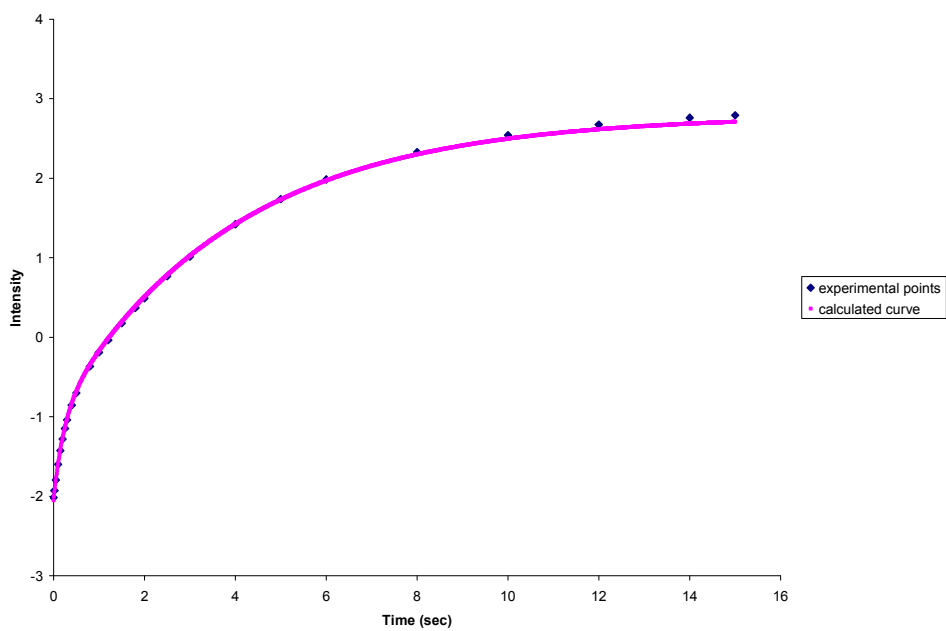


B

Figure 10.24: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **2** at 83.0°C in D₂O

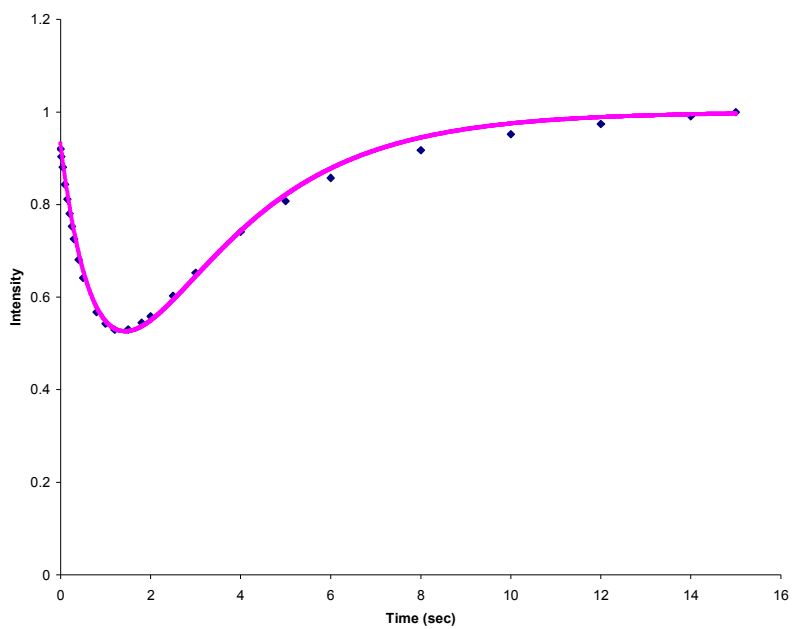


A

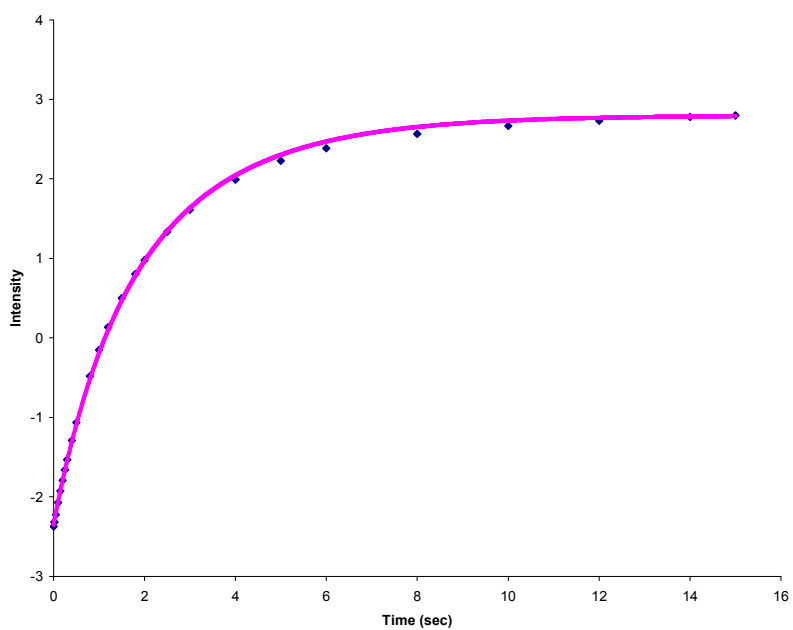


B

Figure 10.25: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **3** at 62.0°C in D₂O

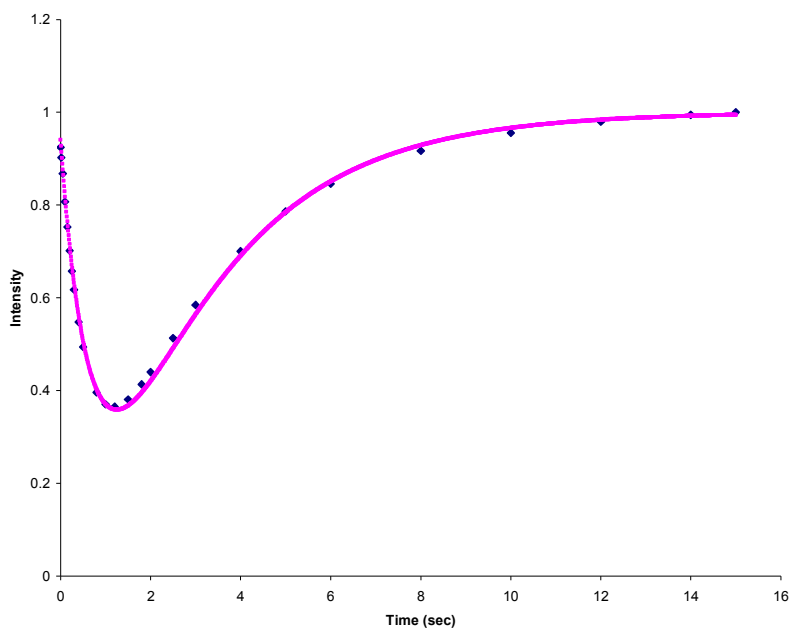


A

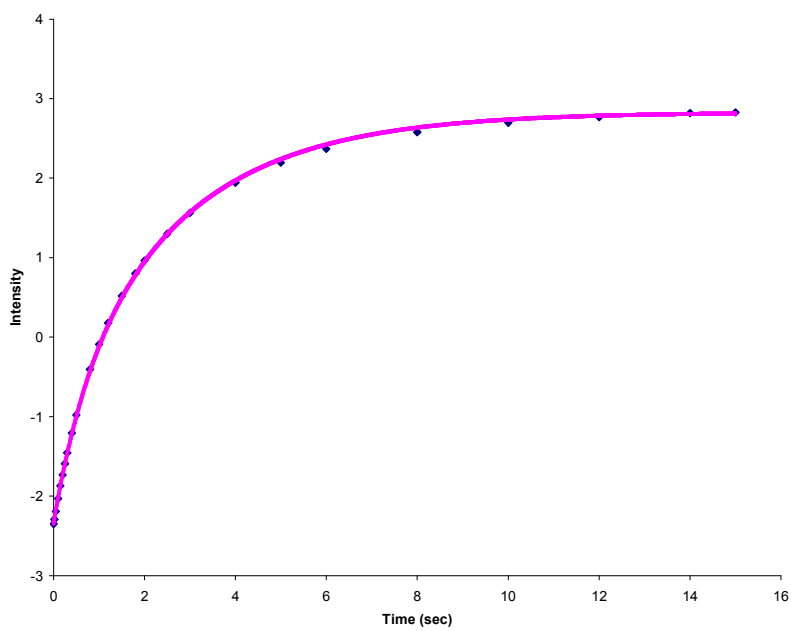


B

Figure 10.26: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **3** at 67.3°C in D₂O

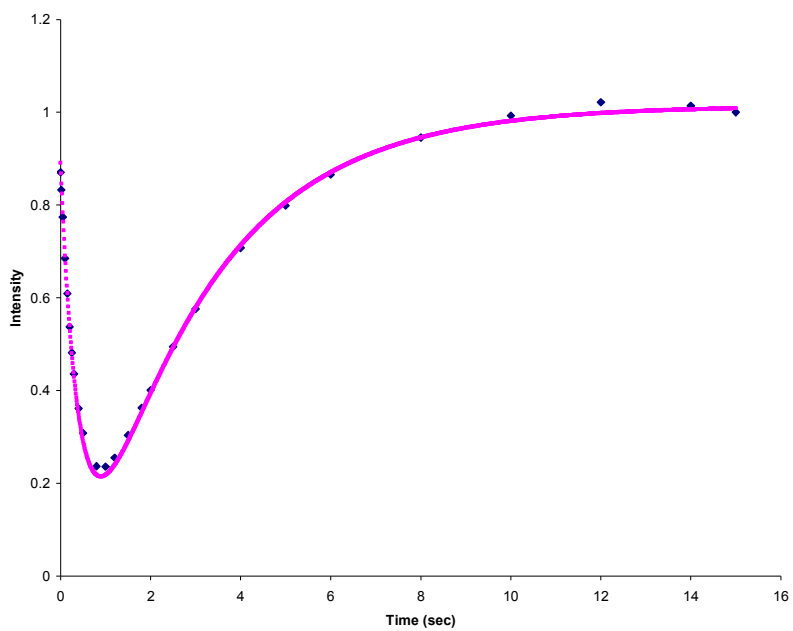


A

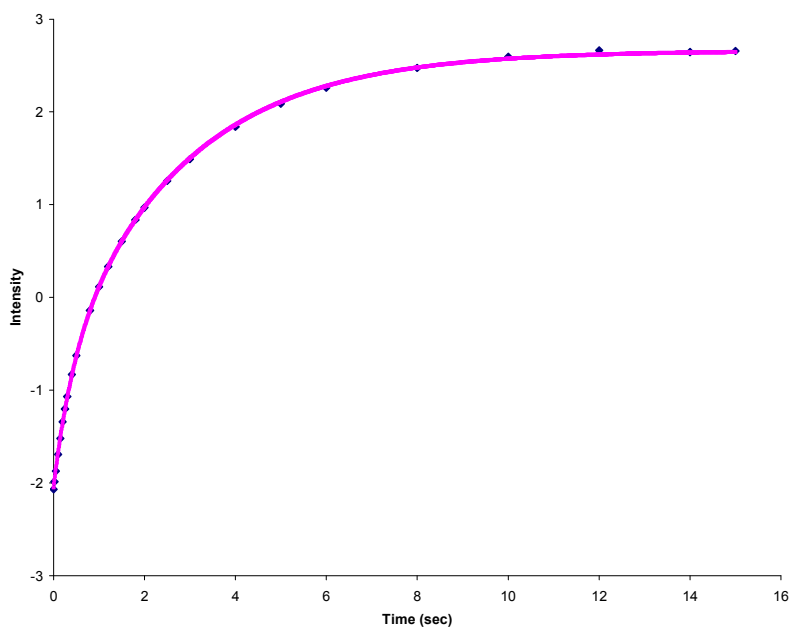


B

Figure 10.27: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **3** at 72.6°C in D₂O

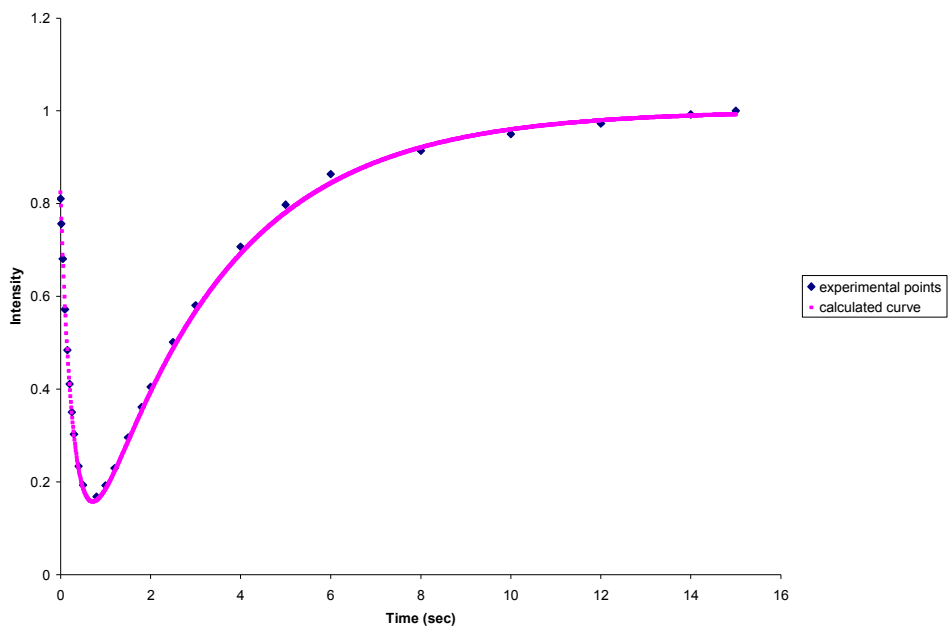


A

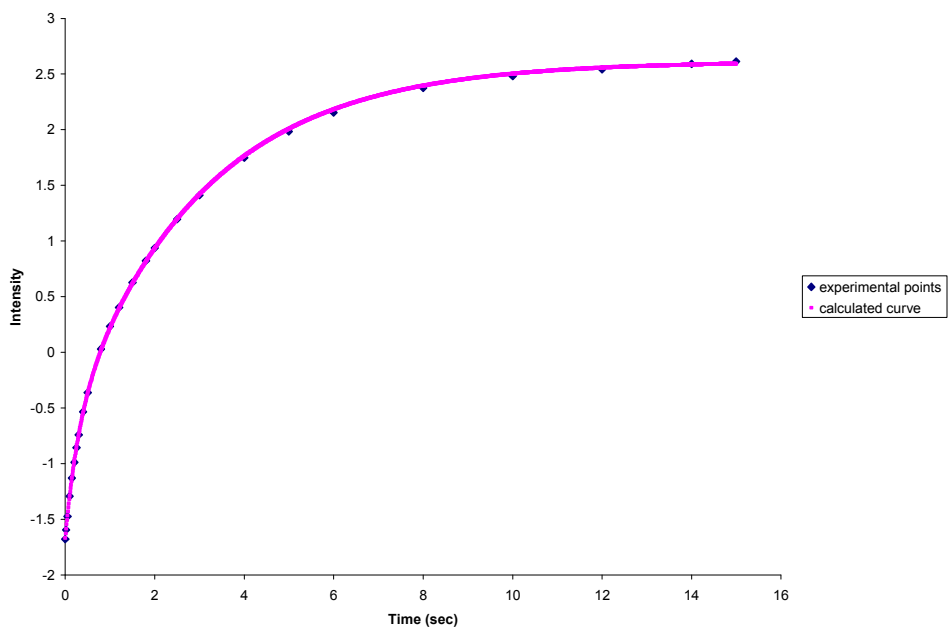


B

Figure 10.28: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **3** at 78.0°C in D₂O

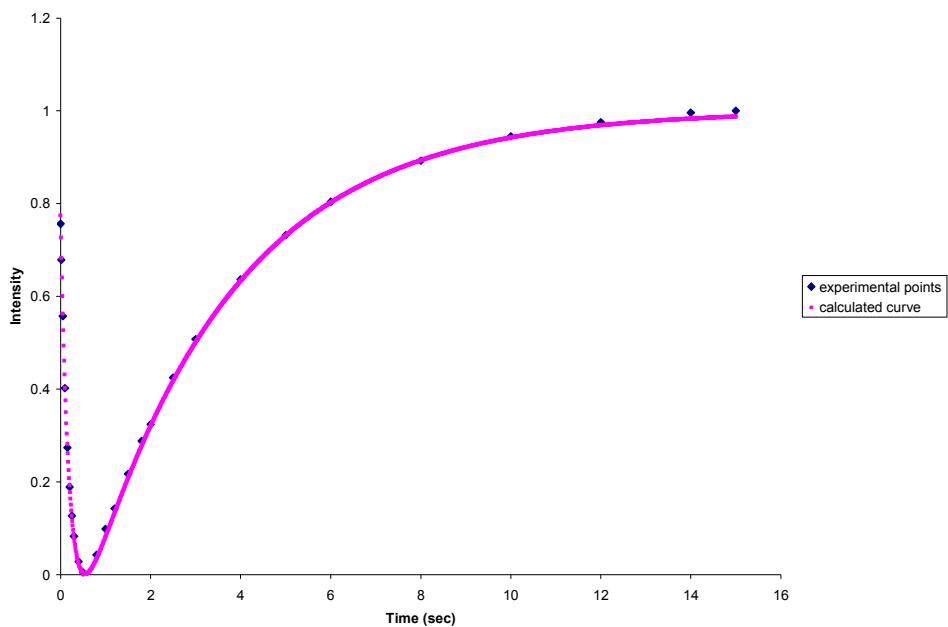


A

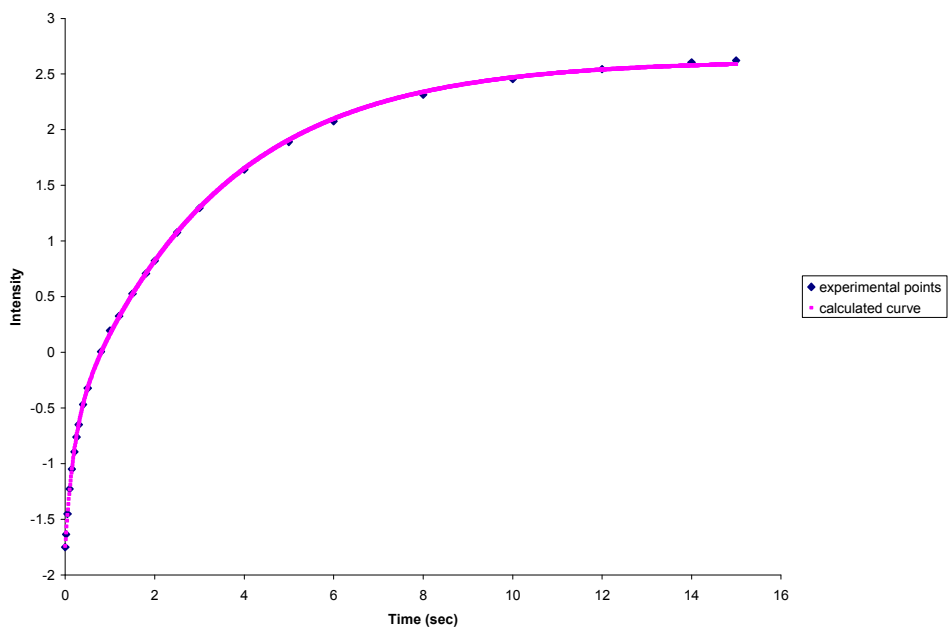


B

Figure 10.29: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **3** at 83.0°C in D₂O

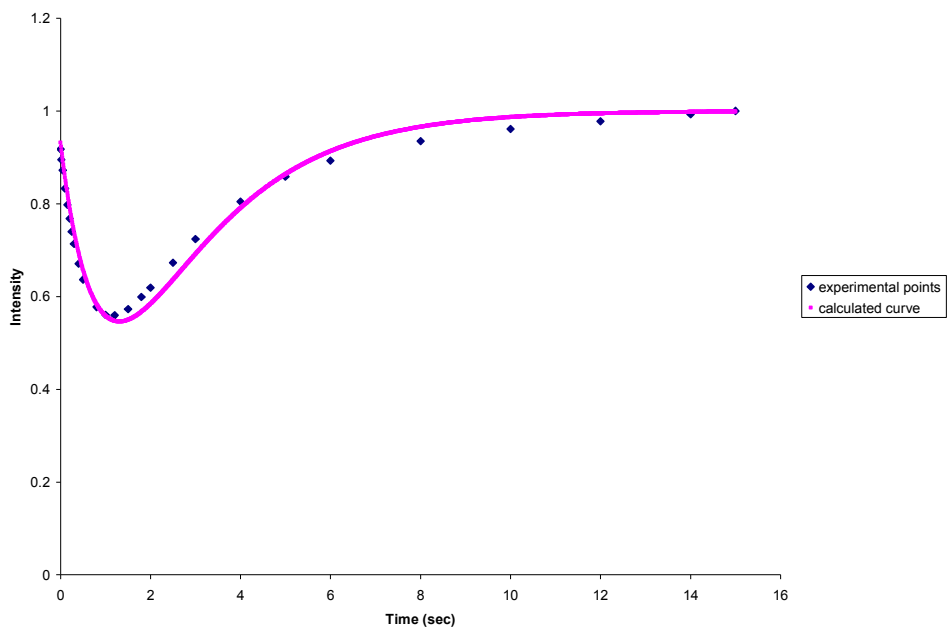


A

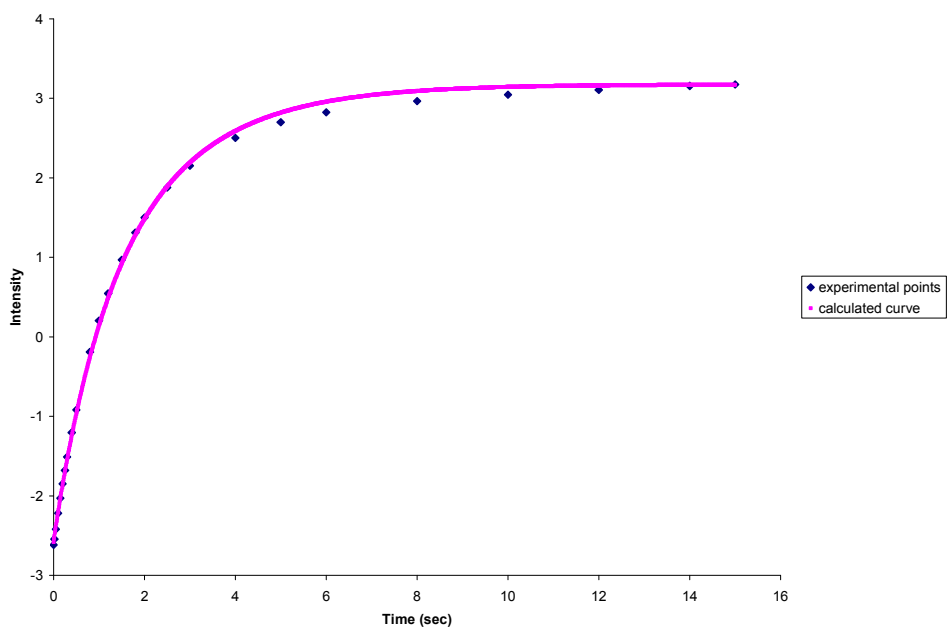


B

Figure 10.30: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4a** at 62.0°C in D₂O

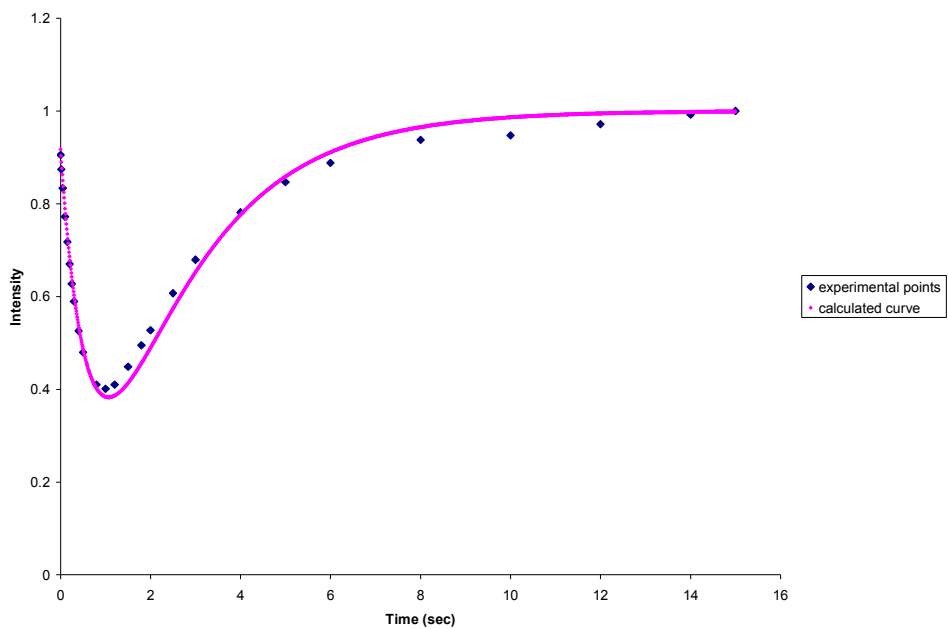


A

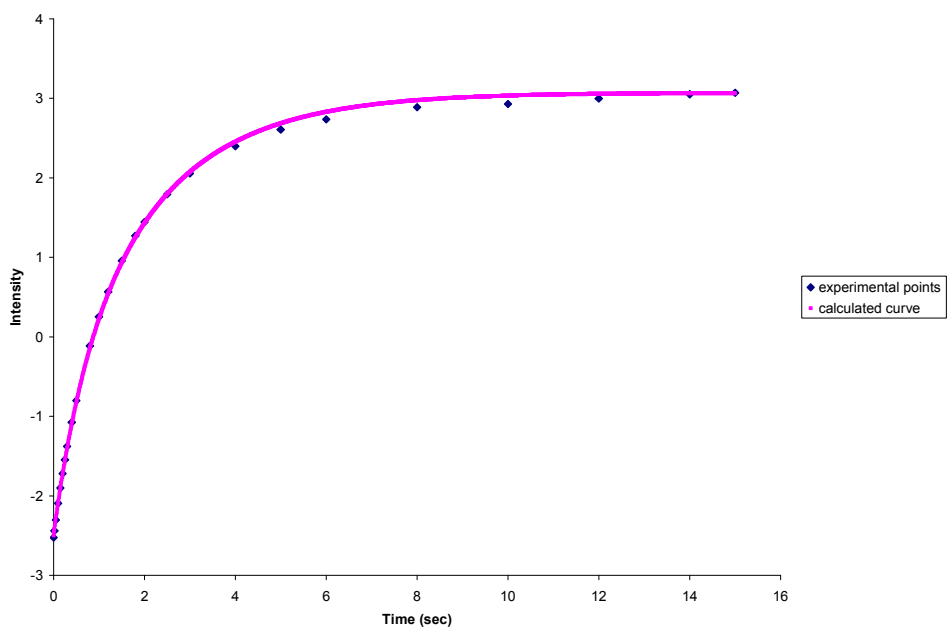


B

Figure 10.31: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4a** at 67.3°C in D₂O

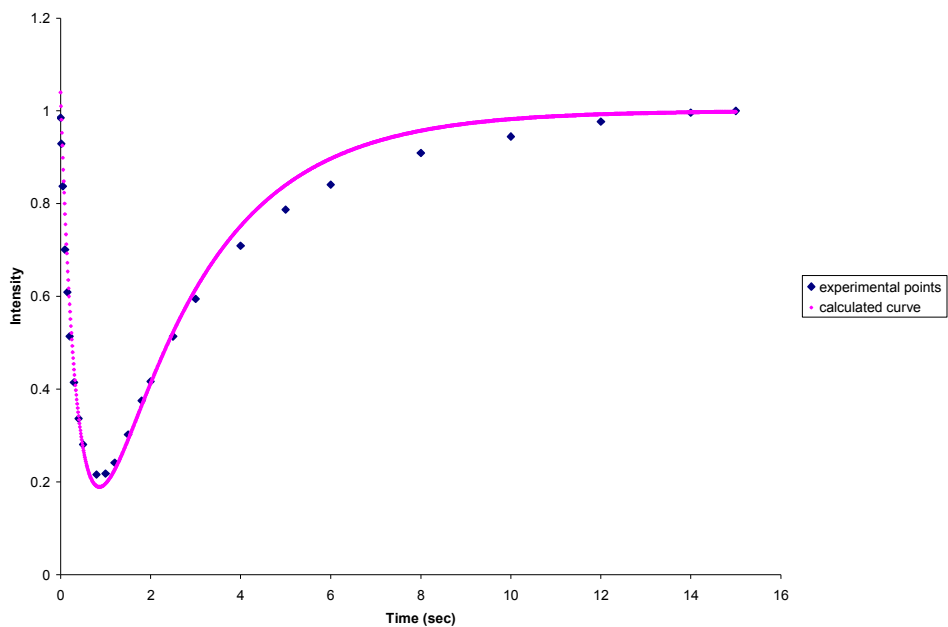


A

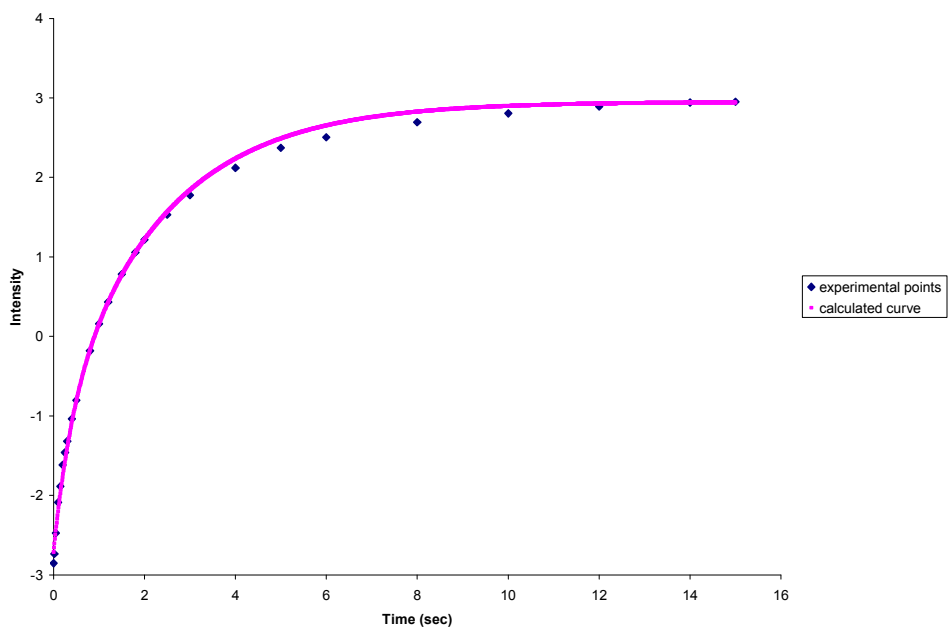


B

Figure 10.32: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4a** at 72.6°C in D₂O

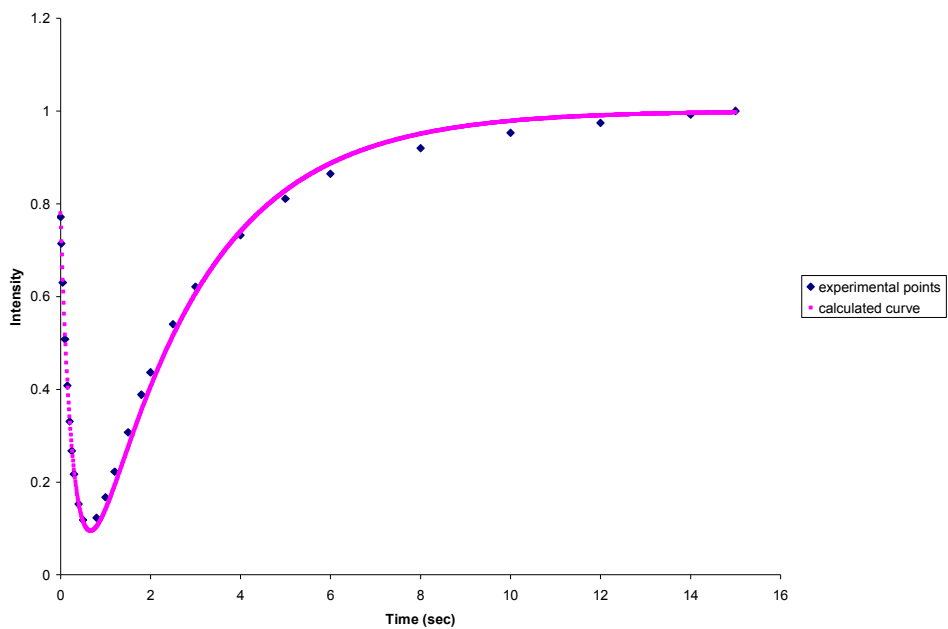


A

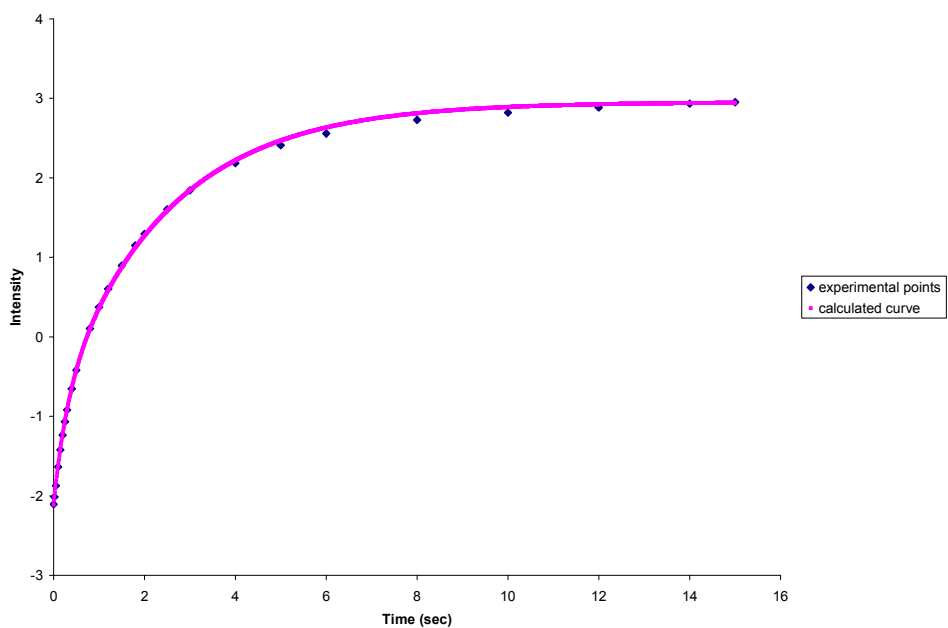


B

Figure 10.33: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4a** at 78.0°C in D₂O

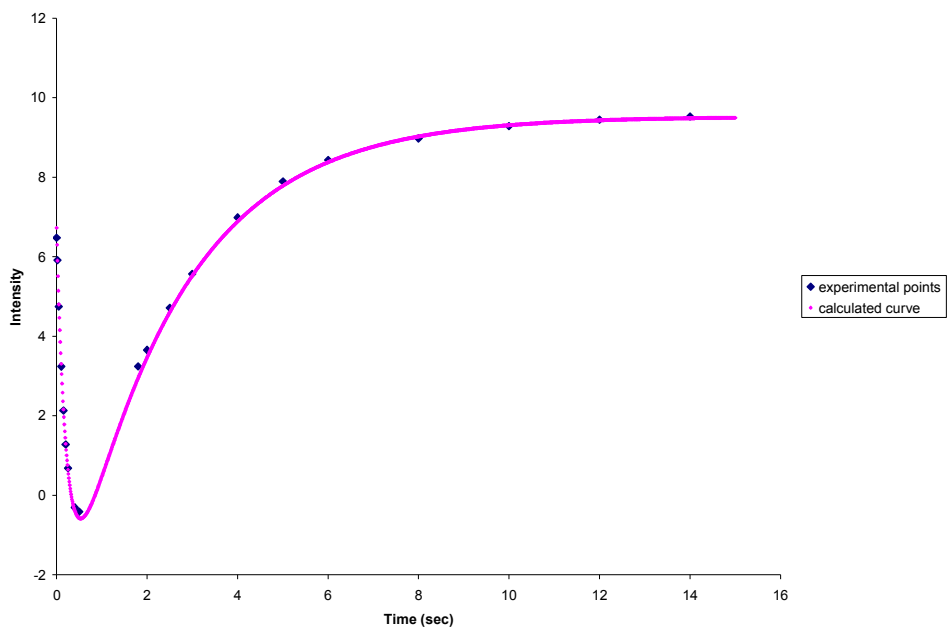


A

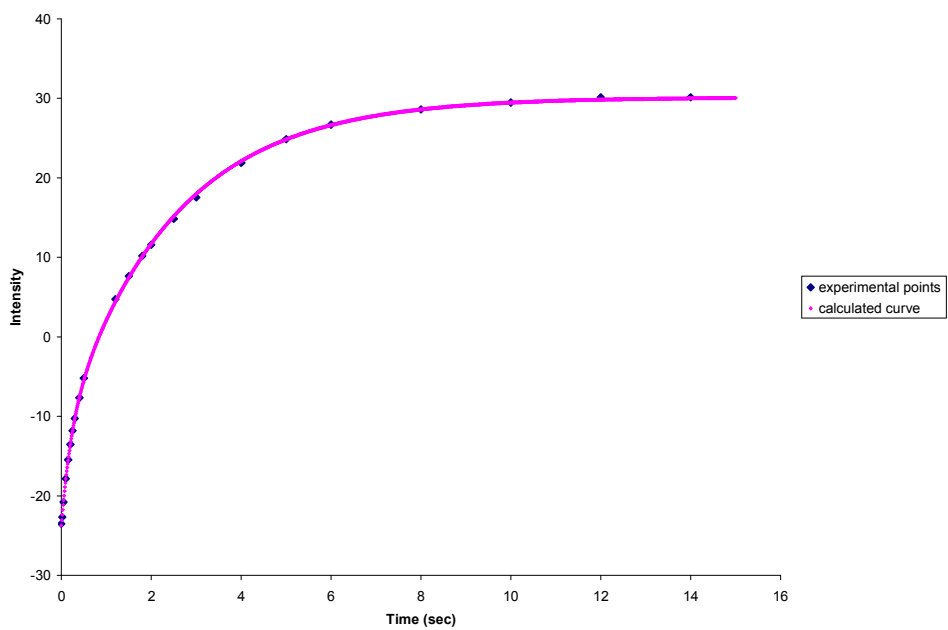


B

Figure 10.34: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4a** at 83.0°C in D₂O

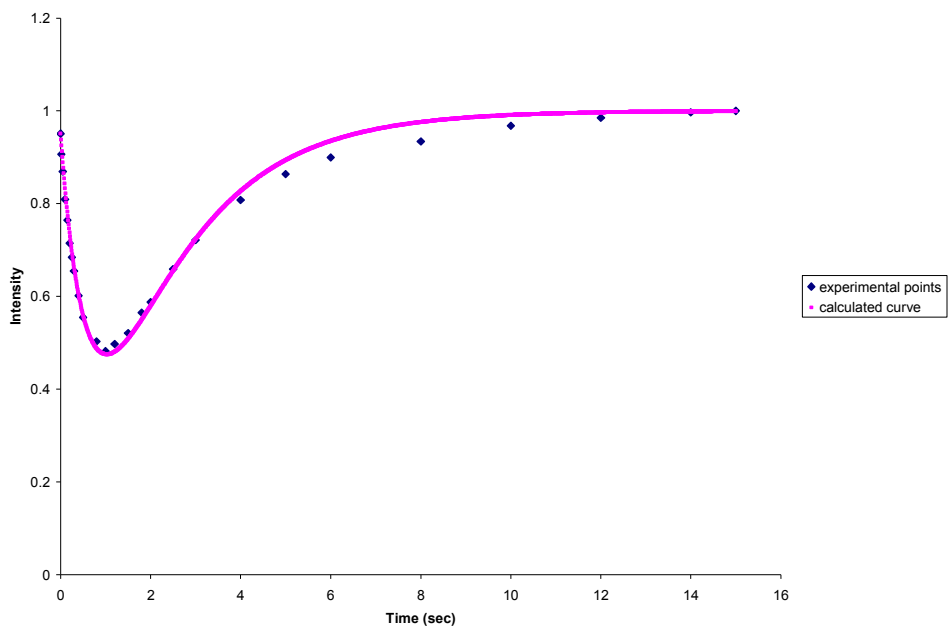


A

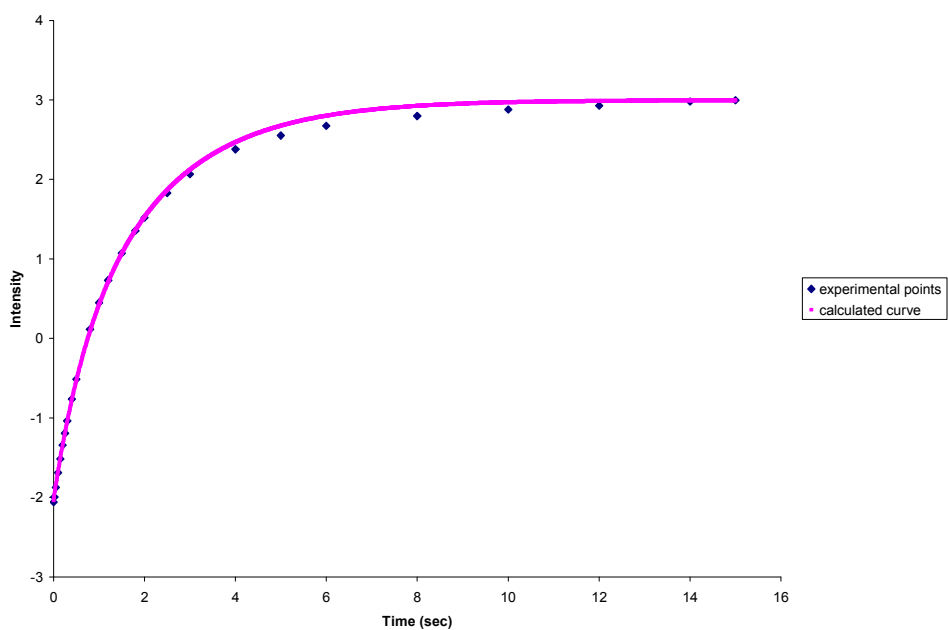


B

Figure 10.35: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4a** at 67.3°C in phosphate buffer pH 7.4

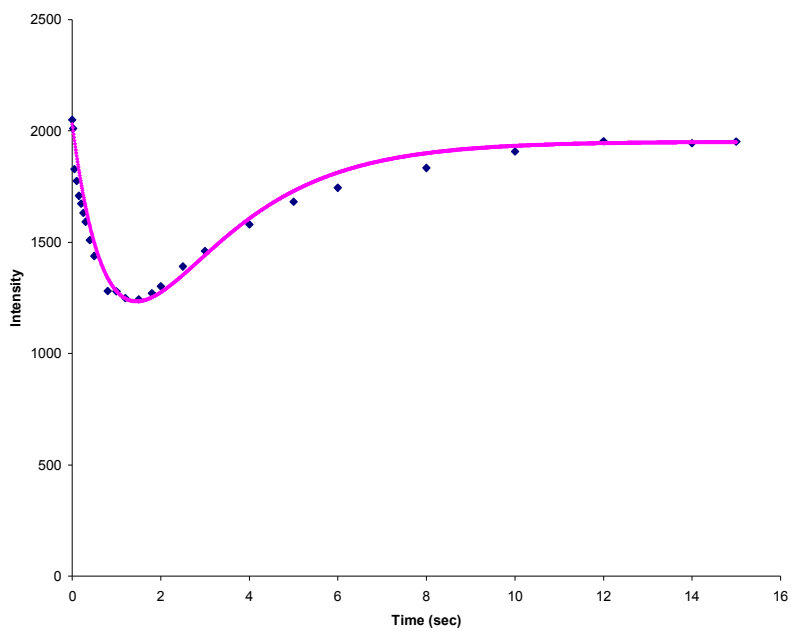


A

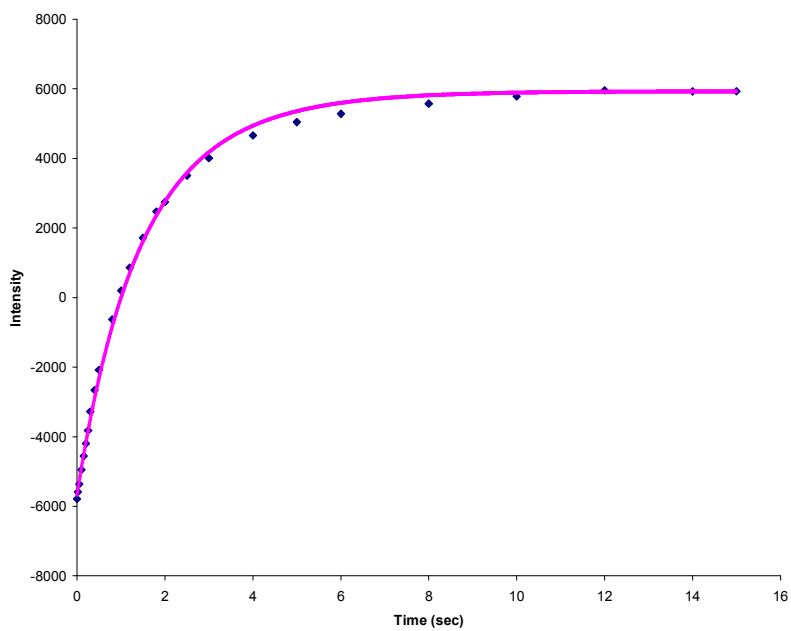


B

Figure 10.36: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4b** at 62.0°C in D₂O

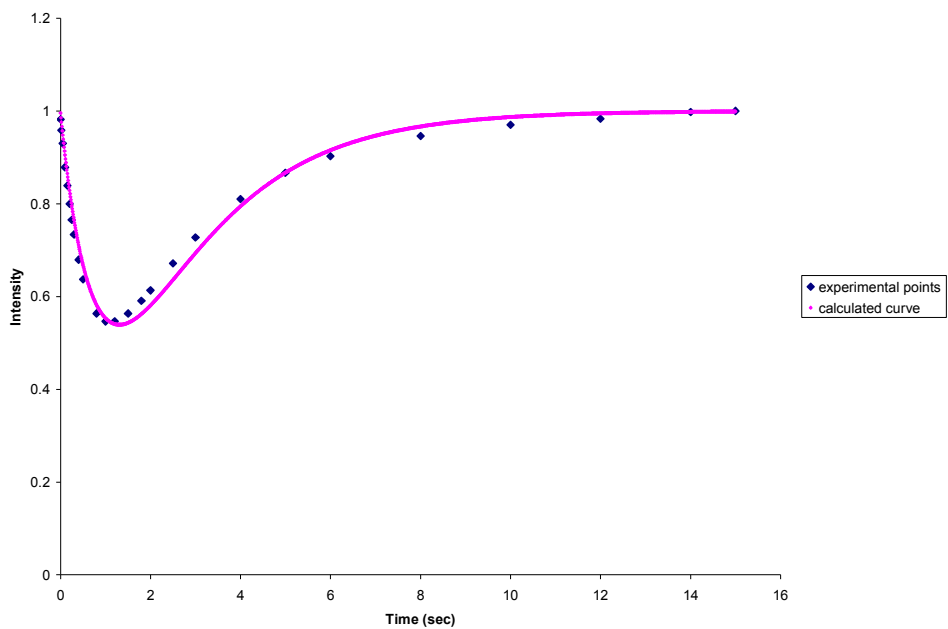


A

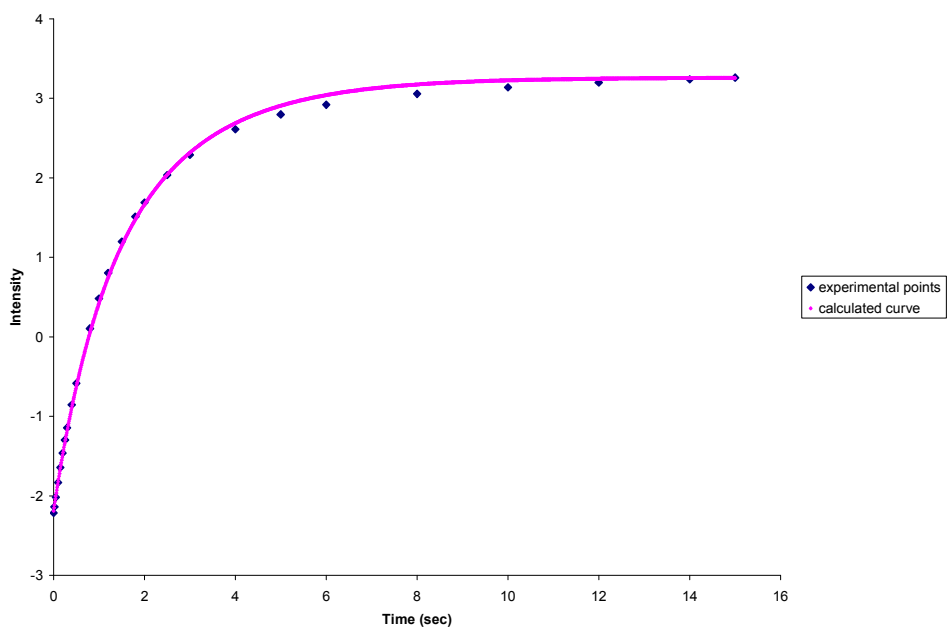


B

Figure 10.37: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4b** at 67.3°C in D₂O

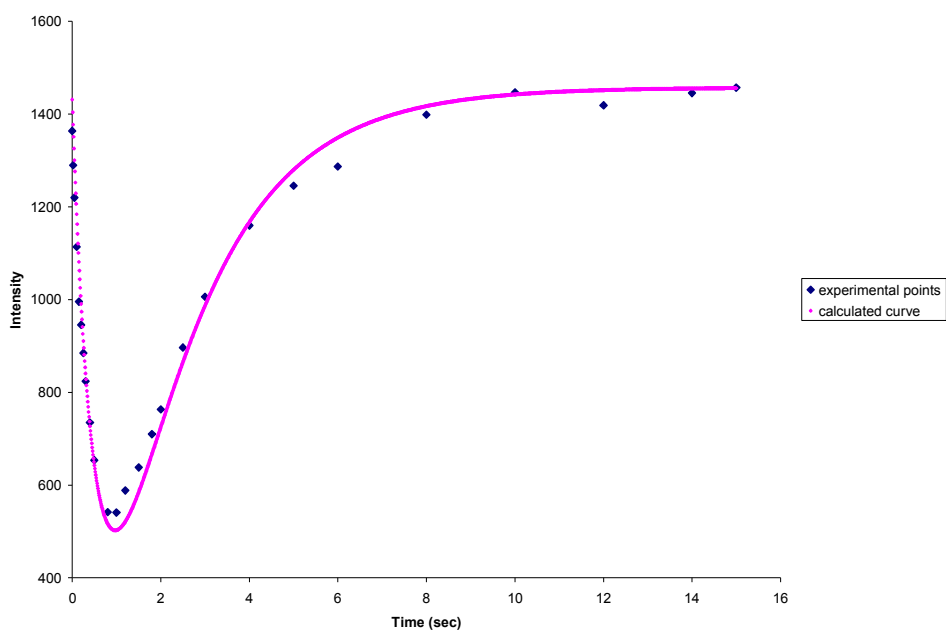


A

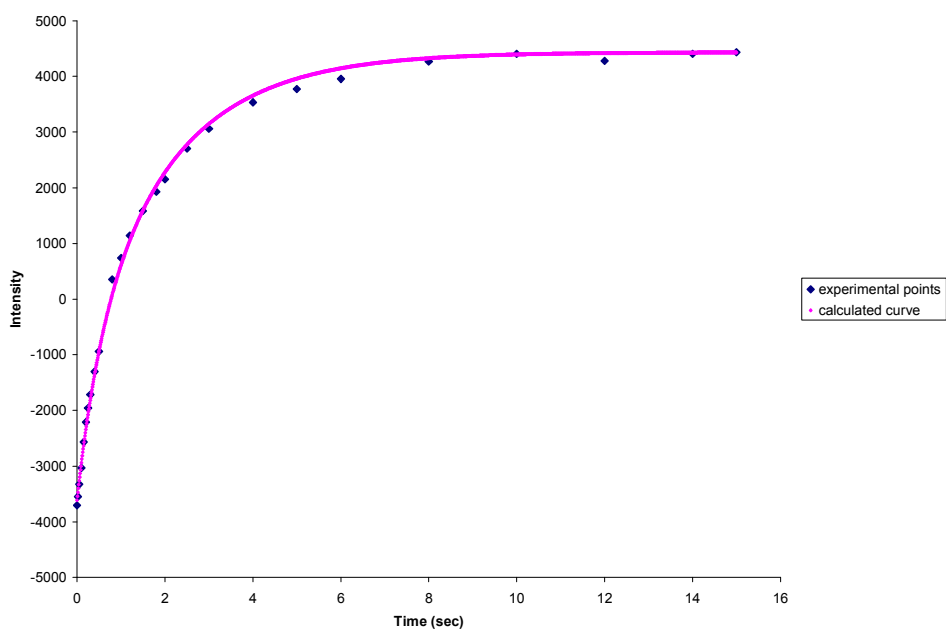


B

Figure 10.38: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4b** at 72.6°C in D₂O

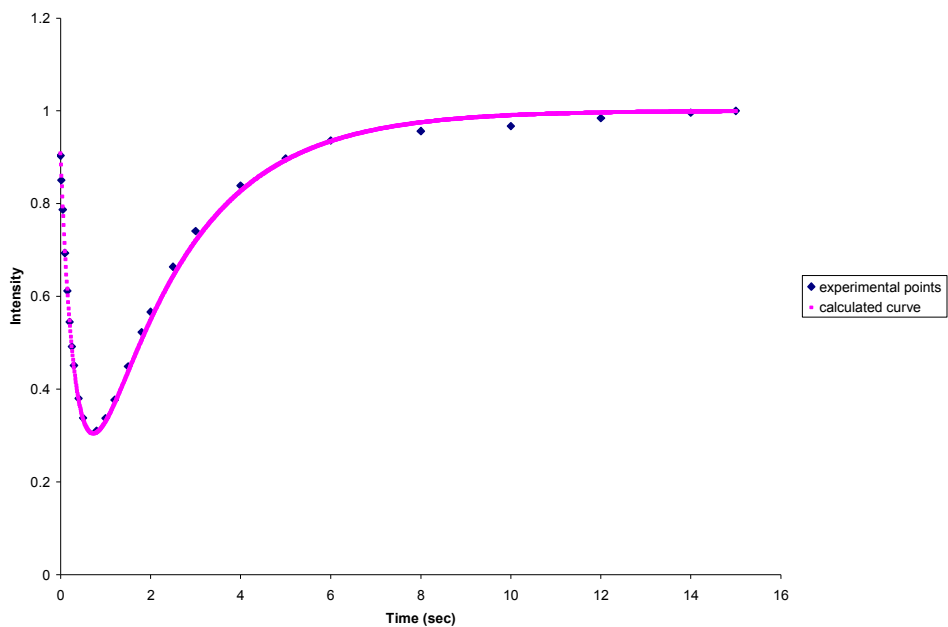


A

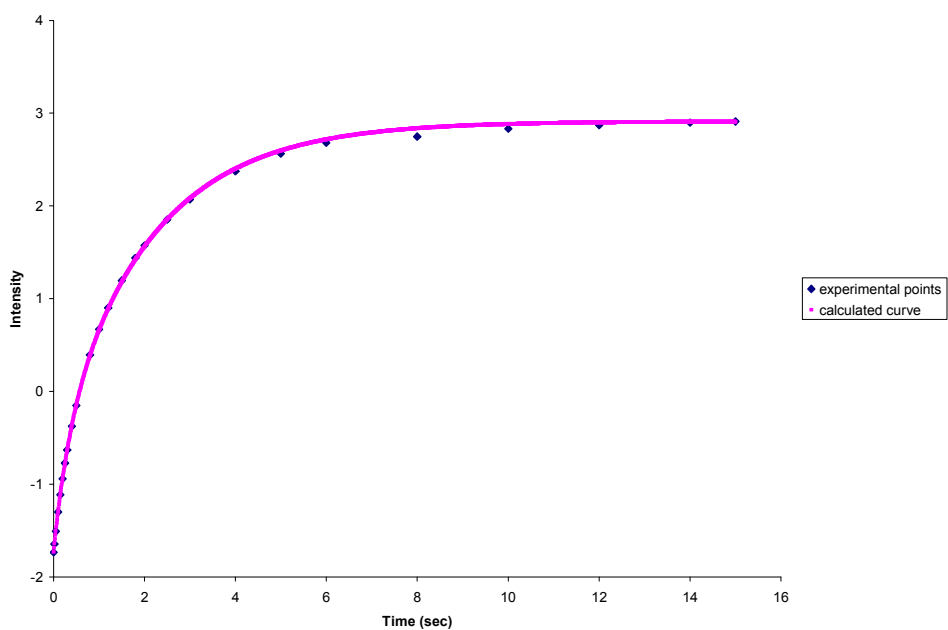


B

Figure 10.39: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4b** at 78.0°C in D₂O

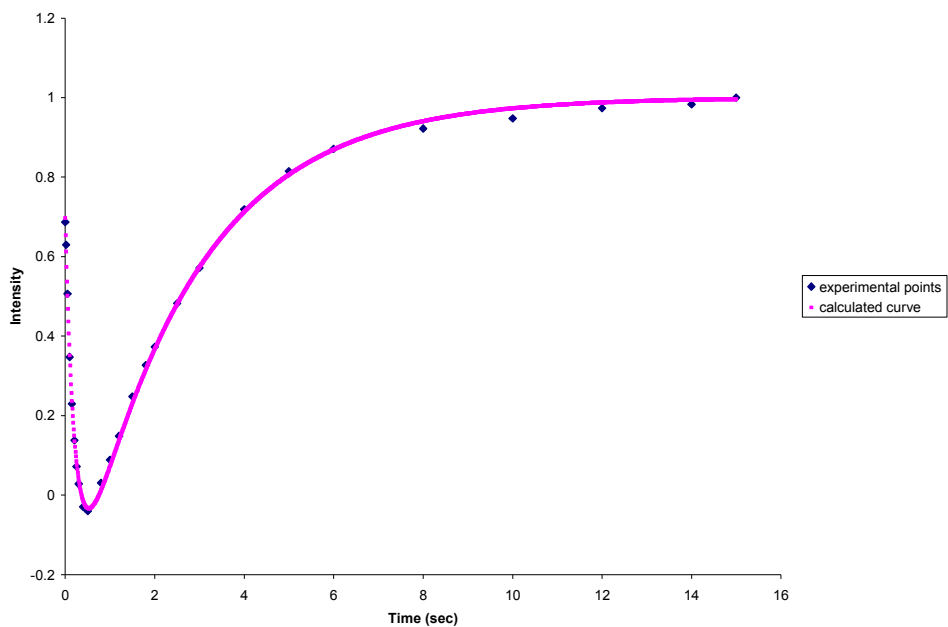


A

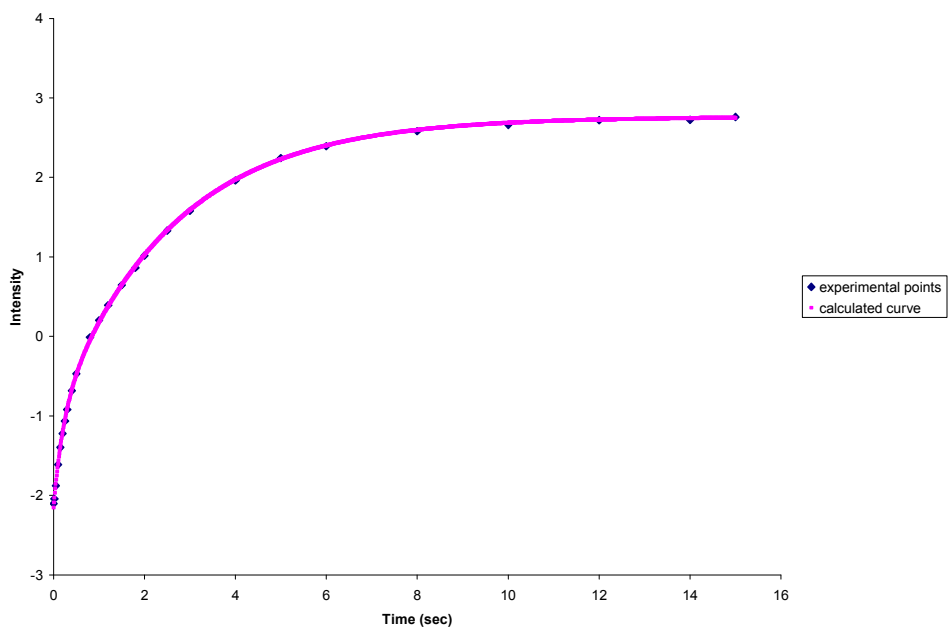


B

Figure 10.40: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4b** at 83.0°C in D₂O

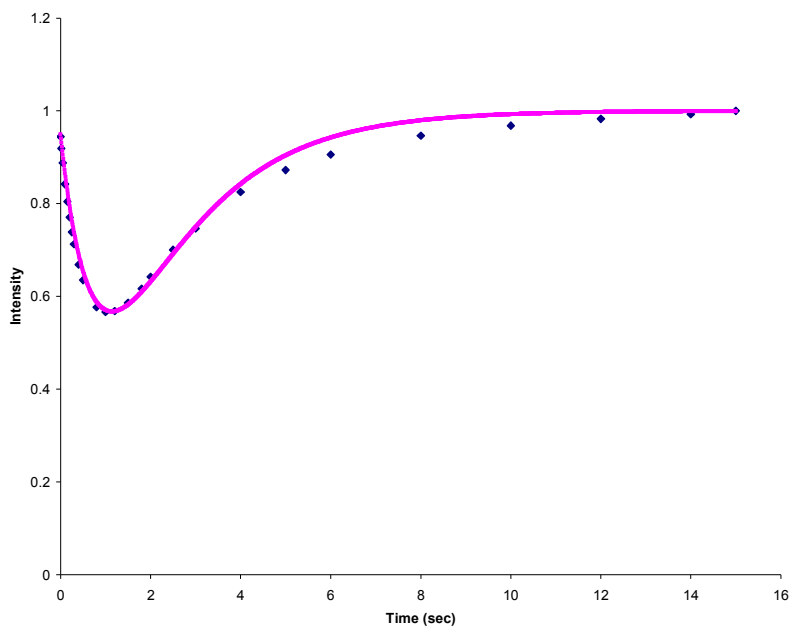


A

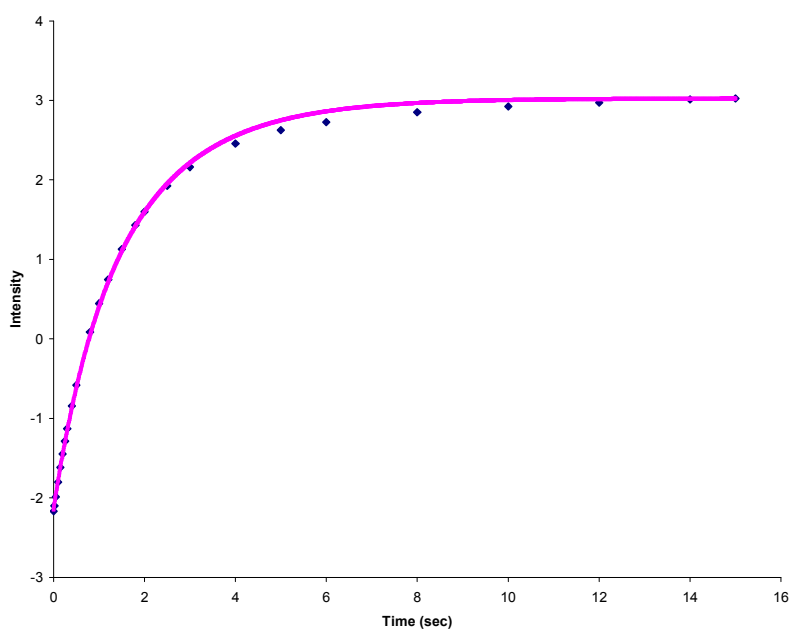


B

Figure 10.41: Inversion recovery of non-inverted (*cis*) methylamide singlet (A) and inverted (*trans*) methylamide singlet (B) of **4b** at 67.3°C in phosphate buffer pH 7.4



A



B

**Chapter 11: The Implications of (2*S*,4*S*)-4-Hydroxyproline 4-O-
Glycosylation on Prolyl Amide Isomerization**

Supporting Information For Chapter 5

Table of Contents

General Details	429
General Procedures	429
Synthesis and Characterization of 8a, 8b, 9a, 9b	430-433
Synthesis and Characterization of 11a, 11b, 12a, 12b	433-436
Tables	437-438
Van't Hoff Plots	439-441
Spectra	442-455
nOe Spectra	456-470
Magnetization Transfer NMR Experiments	471-479

Experimental

General

Protected amino acids were purchased from Bachem (Bachem Bioscience, Inc., King of Prussia, PA). Reagent grade solvents were used without further purification. Column chromatography was performed on SilicaFlash P60 silica gel (40-63 μ m). NMR Spectra were obtained using a Bruker AMX-500 NMR spectrometer equipped with a triple resonance (^1H , ^{13}C , ^{15}N) gradient inverse probehead. Spectra were assigned based on 2D COSY and 2D HSQC experiments. For ^1H NMR, minor isomers are listed between square brackets. For ^{13}C NMR, when assigned, carbon peaks for the minor isomer are listed in brackets. A Waters Micromass ZQ 2000 mass spectrometer was used for electrospray ionization (ESI) mass spectrometry measurements.

General Procedures

Preparation of *N*-Acetyl-*N'*-methyl ester amino acids

The amino acid (1.0 eq) was dissolved in 3:1 (v/v) DCM/piperidine, stirred for 2 hours at ambient temperature, then co-distilled with toluene (3 x 10 mL). The crude product was then dissolved in 1:1 acetic anhydride/pyridine, and the mixture was stirred at ambient temperature for 4 hours. The solution was concentrated under reduced pressure, and then was co-distilled with toluene (3 x 10 mL) before purification by flash chromatography.

Preparation of 4-*O*-Galactopyranosyl-*N*-acetyl-*N'*-methyl ester amino acids

The amino acid (1.0 eq) was dissolved in CH_3OH . The catalyst, (20% palladium hydroxide on carbon) was added (approx. 0.5 eq), and the flask was flushed with N_2 for 5 minutes. The reaction mixture was then stirred under hydrogen gas (10 psi) for 6 hours,

after which it was flushed with nitrogen and filtered. The product was then concentrated under reduced pressure.

4-*O*-(2,3,4,6-Tetra-*O*-benzyl- α,β -D-galactopyranosyl)-*N*-fluorenylmethoxycarbonyl-*trans*-4-hydroxy-*L*-proline methyl ester (8a,b**)**

Compound **7** (0.200 g, 1.1 mmol) was dissolved in 3:1 (v/v) acetone/water. After addition of NaHCO₃ (0.932 g, 11.1 mmol, 10.0 eq) and Fmoc-OPfp (1.34 g, 3.3 mmol, 3.0 eq), the mixture was stirred at ambient temperature for 3 hours. The reaction mixture was then concentrated under reduced pressure before being re-dissolved in ethyl acetate (20 mL). This solution was washed with water (2 x 20 mL) and brine (1 x 20 mL), dried (Na₂SO₄) and concentrated under reduced pressure before being purified by flash chromatography using 4:1 hexanes/ethyl acetate. The intermediate compound was then dissolved in MeCN (10 mL), followed by addition of 2,3,4,6-tetra-*O*-benzyl-thio- β -D-galactopyranoside (0.925 g, 1.43 mmol, 1.3 eq). The solution was allowed to cool to 0 °C and was stirred for ten minutes. Afterwards, NIS (0.322 g, 1.43 mmol, 1.3 eq) and AgOTf (0.056 g, 0.22 mmol, 0.2 eq) were added, and the reaction mixture was stirred for additional 2 hours before being concentrated under reduced pressure. This solution was washed with 1:1 saturated sodium bicarbonate (aq)/saturated sodium thiosulphate (aq) (5 x 20 mL), dried (Na₂SO₄) and concentrated under reduced pressure. The product was purified by flash chromatography using 3:1 hexanes/ ethyl acetate to yield first **8a** (0.225 g, 0.253 mmol) and then **8b** (0.119 g, 0.134 mmol) (35.2%).

8a: $[\alpha]_D^{25} = +80.6^\circ$ (*c* 0.7 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.77$ (d, 2H, aromatic, *J* = 7.5Hz), 7.51-7.66 (m, 2H, aromatic), 7.17-7.47 (m, 24H, aromatic), 4.96 (d, 1H, -OCH₂Ph, *J* = 11.5Hz), 4.89 (d, 1H, H₁, *J* = 3.5Hz), 4.69-4.88 (m, 3H, -OCH₂Ph), 4.62 (d, 1H, -OCH₂Ph, *J* = 12.2Hz), 4.58 (d, 1H, -OCH₂Ph, *J* = 11.7Hz), 4.21-4.58 (m, 6.5H, H₅, H_{6a}, H_{6b}^{trans}, Pro_γ, Pro_α, -OCH₂Ph), 4.16 (dd apparent t, 0.5H, H_{6b}^{cis}), 4.06 (dd, 1H, H₂, *J* = 3.8Hz, *J* = 9.8Hz), 3.87-4.00 (m, 3H, H₃, H₄, -C(O)OCH₂CH), 3.67-3.86 (m, 3.5H, Pro_{δ1}, Pro_{δ2}, -OCH₃^{trans}), 3.64 (s, 1.5H, -OCH₃^{cis}), 3.44-3.59 (m, 2H, -C(O)OCH₂CH), 2.51 (ddd, 0.5H, Pro_{β1}^{trans}, *J* = 4.7Hz, *J* = 8.4Hz, *J* = 13.1Hz), 2.44-2.57 (m, 0.5H, Pro_{β1}^{cis}), 2.11-2.24 (m, 1H, Pro_{β2}); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta =$ (173.0), 172.9, 154.8, (154.4), (144.1), 143.97, 143.93, (143.8), 141.3 (2), (138.8), (138.7), 138.54, (138.47), 137.9, 127.0-128.5 (aromatic), 125.2, 125.14, (125.09), (125.0), 120.0, (119.9), 97.8, (97.0), 78.80, (78.76), 76.4, (76.3), 76.2, 75.0, 74.79, (74.77), 74.3, (73.56), 73.53, 73.4, (73.27), 73.21, 70.21, (70.15), (69.3), 67.7, 58.2, (57.8), 52.38, (52.34), (51.68), 51.64, (47.3), 47.2, (37.4), 36.4; MS (ES) calc. for C₅₅H₅₅NNaO₁₀ (M + Na)⁺: 912.37. Found (M + Na)⁺: 912.55;

8b: $[\alpha]_D^{25} = -23.1^\circ$ (*c* 0.7 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.71$ -7.79 (m, 2H, aromatic), 7.51-7.66 (m, 2H, aromatic), 7.17-7.44 (m, 24H, aromatic), 4.58-5.01 (m, 6H, -OCH₂Ph), 4.29-4.58 (m, 7H, H₁, H₅, H_{6a}, Pro_γ, Pro_α, -OCH₂Ph), 4.25 (dd apparent t, 0.6H, H_{6b}^{trans}, *J* = 7.5Hz, *J* = 7.5Hz), [4.15, dd apparent t, 0.4H, H_{6b}^{cis}, *J* = 6.5Hz, *J* = 7.0Hz], 3.70-3.95 (m, 6H, H₂, Pro_{δ1}, Pro_{δ2}, -C(O)OCH₂CH, -OCH₃^{trans}), [3.65, s, 1H, -OCH₃^{cis}], 3.46-3.62 (m, 4H, H₃, H₄, -C(O)OCH₂CH), 2.33-2.50 (m, 1H, Pro_{β1}), 2.04-2.19 (m, 1H, Pro_{β2}); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta =$ (172.95), 172.89, 154.9, (154.3),

(144.1), 144.0, 143.9, (143.7), 141.3 (2), (138.52), 138.50, 138.37, (138.35), (137.82), 127.0-128.5 (aromatic), (125.26), 125.22, 125.05, (124.98), 119.95, (119.90), (102.72), 102.66, 82.32, (82.26), (79.25), 79.17, (75.98), 75.95, 75.5, (75.4), 74.65, (74.57), 73.61 (2), (73.57), 73.3, (73.2), 73.1, (72.9), 68.75, (68.72), 67.7, (67.6), 57.8, (57.5), (53.3), 52.9, 52.39, (52.35), (47.3), 47.1, (36.5), 35.3; MS (ES) calc. for $C_{55}H_{55}NNaO_{10}$ (M + Na)⁺: 912.37. Found (M + Na)⁺: 912.11;

4-*O*-(2,3,4,6-Tetra-*O*-benzyl- α,β -D-galactopyranosyl)-*N*-acetyl-*trans*-4-hydroxy-*L*-proline methyl ester (9a,b)

The general preparation method was followed for *N*-terminal acylation of **8a** (0.222 g, 0.313 mmol) using 3:1 DCM:piperidine (8 mL). The reaction mixture was purified by flash chromatography using 2:1 ethyl acetate/hexanes to yield **9a** as a clear oil (0.138 g, 0.194 mmol) (78.0%):

9a: $[\alpha]_D^{25} = +31.6^\circ$ (*c* 0.3 $CHCl_3$); 1H NMR (300 MHz, $CDCl_3$, 298K) δ = 7.12-7.49 (m, 20H, aromatic), 4.25-5.05 (m, 11H, H_1 , Pro_γ , Pro_α , $-OCH_2Ph$), 3.82-4.13 (m, 4H, H_2 , H_3 , H_4 , H_5), 3.33-3.82 (m, 7H, H_{6a} , H_{6b} , $Pro_{\delta 1}$, $Pro_{\delta 2}$, $-CO_2CH_3$), [2.50-2.64, m, 0.2H, $Pro_{\beta 1}^{cis}$], 2.34-2.49 (m, 0.8H, $Pro_{\beta 1}^{trans}$), [2.18-2.32, m, 0.2H, $Pro_{\beta 2}^{cis}$], 2.06-2.19 (m, 0.8H, $Pro_{\beta 2}^{trans}$), 2.00 (s, 2.5H, $-NCOCH_3^{trans}$), [1.92, s, 0.5H, $-NCOCH_3^{cis}$]; ^{13}C NMR (75 MHz, $CDCl_3$, 298K) δ = 172.8, 169.6, 138.61, 138.57, 138.49, 137.9, 127.4-128.6 (aromatic), 98.4, 78.9, 76.9, 76.4, 74.9, 74.8, 73.8, 73.5, 73.1, 70.3, 69.2, 57.6, 53.0, 52.3, 35.9, 22.2; MS (ES) calc. for $C_{42}H_{47}NNaO_9$ (M + Na)⁺: 732.31. Found (M + Na)⁺: 732.74;

9b: $[\alpha]_D^{25} = -5.6^\circ$ (*c* 0.3 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.18-7.44$ (m, 20H, aromatic), 4.29-5.04 (m, 11H, H₁, Pro _{γ} , Pro _{α} , -OCH₂Ph), 3.42-3.97 (m, 11H, H₂, H₃, H₄, H₅, H_{6a}, H_{6b}, Pro _{δ 1}, Pro _{δ 2}, -CO₂CH₃), [2.41-2.57, m, 0.1H, Pro _{β 1}^{*cis*}], 2.28-2.41 (m, 0.9H, Pro _{β 1}^{*trans*}), [2.16-2.27, m, 0.1H, Pro _{β 2}^{*cis*}], 1.96-2.15 (m, 3.7H, Pro _{β 2}^{*trans*}, -NCOCH₃^{*trans*}), [1.90, s, 0.3H, -NCOCH₃^{*cis*}]; ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta = 172.8, 169.5, 138.39, 138.35, 138.29, 137.7, 127.5-128.7$ (aromatic), 103.0, 82.3, 79.1, 77.3, 75.5, 74.6, 73.65, 73.60, 73.3, 73.1, 69.0, 57.1, 54.0, 52.3, 35.1, 22.2; MS (ES) calc. for C₄₂H₄₇NNaO₉ (M + Na)⁺: 732.31. Found (M + Na)⁺: 732.58;

4-O-(2,3,4,6-Tetra-O-benzyl- α,β -D-galactopyranosyl)-N-fluorenylmethoxycarbonyl-*cis*-4-hydroxy-L-proline methyl ester (11a,b)

Compound **10** (0.200 g, 0.815 mmol, 1.0 eq) was dissolved in MeCN (10 mL), followed by addition of 2,3,4,6-tetra-*O*-benzyl-thio- β -D-galactopyranoside (0.686 g, 1.06 mmol, 1.3 eq). The solution was allowed to cool to 0 °C and was stirred for 10 minutes. Then, NIS (0.239 g, 1.06 mmol, 1.3 eq) and AgOTf (0.042g, 0.16 mmol, 0.2 eq) were added and the reaction mixture was stirred for 2.5 hours before being concentrated under reduced pressure. The resulting residue was re-dissolved in ethyl acetate (20 mL) and was washed with 1:1 saturated sodium bicarbonate (aq)/saturated sodium thiosulphate (aq) (5 x 20 mL), before being dried (Na₂SO₄) and concentrated under reduced pressure. The residue was dissolved in DCM (5 mL) and cooled to 0°C, before addition of TFA (5 mL). After stirring for 1.5 hrs, the solution was co-distilled with toluene (3 x 6 mL) and concentrated under reduced pressure. The residue was then re-dissolved in 3:1 (v/v) acetone/water (8 mL), followed by the addition of NaHCO₃ (0.328 g, 3.9 mmol, 15 eq)

and Fmoc-OPfp (0.732 g, 0.78 mmol, 3.0 eq). The reaction was stirred for 2 hours before concentration. The resulting residue was re-dissolved in ethyl acetate (20 mL) and was washed with water (2 x 20 mL) and brine (1 x 20 mL), before being dried (Na_2SO_4) and concentrated under reduced pressure for purification by flash chromatography using 3:1 hexanes/ethyl acetate to yield first **11a** (0.141g, 0.158 mmol) then **11b** (0.096g, 0.108 mmol) (32.2%):

11a: $[\alpha]_D^{25} = +72.8^\circ$ (*c* 0.5 CHCl_3); ^1H NMR (300 MHz, CDCl_3 , 298K) $\delta = 7.76$ (d, 2H, aromatic, $J = 7.7\text{Hz}$), 7.50-7.67 (m, 2H, aromatic), 7.18-7.49 (m, 24H, aromatic), 4.56-5.00 (m, 6H, $-\text{OCH}_2\text{Ph}$), 4.30-4.55 (m, 7H, H_1 , H_5 , H_{6a} , Pro_γ , Pro_α , $-\text{OCH}_2\text{Ph}$), 4.15-4.29 (m, 1H, $\text{H}_{6b}^{\text{trans}}$, $\text{H}_{6b}^{\text{cis}}$), 3.84-3.92 (m, 1H, H_4), 3.64-3.83 (m, 4.5H, H_2 , H_3 , $-\text{C}(\text{O})\text{OCH}_2\text{CH}$, $-\text{OCH}_3^{\text{trans}}$), 3.45-3.64 (m, 5.5H, $\text{Pro}_{\delta 1}$, $\text{Pro}_{\delta 2}$, $-\text{C}(\text{O})\text{OCH}_2\text{CH}$, $-\text{OCH}_3^{\text{cis}}$), 2.37-2.56 (m, 2H, $\text{Pro}_{\beta 1}$, $\text{Pro}_{\beta 2}$); ^{13}C NMR (75 MHz, CDCl_3 , 298K) $\delta = 172.7$, 154.5, 143.9, 143.7, 141.3 (2), 138.5, 137.9, 126.3-130.9 (aromatic), 125.2, 120.1, 98.6, (98.2), 79.2, 76.8, 76.5, 75.8, 75.3, 74.6, 74.2, 73.8, 73.7, 70.6, 69.3, 68.1, 53.4, 53.1, 48.1, 37.7; MS (ES) calc. for $\text{C}_{55}\text{H}_{55}\text{NNaO}_{10}$ ($\text{M} + \text{Na}$) $^+$: 912.37. Found ($\text{M} + \text{Na}$) $^+$: 912.89;

11b: $[\alpha]_D^{25} = -19.3^\circ$ (*c* 0.7 CHCl_3); ^1H NMR (300 MHz, CDCl_3 , 298K) $\delta = 7.78$ (d, 2H, aromatic, $J = 7.6\text{Hz}$), 7.49-7.63 (m, 2H, aromatic), 7.15-7.44 (m, 24H, aromatic), 4.29-4.98 (m, 12H, H_1 , H_5 , Pro_γ , Pro_α , $-\text{OCH}_2\text{Ph}$), 4.13-4.27 (m, 2H, H_{6a} , H_{6b}), 3.93-4.08 (m, 3H, H_3 , H_4 , $\text{Pro}_{\delta 1}$), 3.87 (dd, 0.5H, $\text{Pro}_{\delta 2}^{\text{trans}}$, $J = 2.7\text{Hz}$, $J = 10.2\text{Hz}$), 3.81 (dd, 0.5H, $\text{Pro}_{\delta 2}^{\text{cis}}$, $J = 2.7\text{Hz}$, $J = 10.4\text{Hz}$), 3.68-3.78 (m, 2H, H_2 , $-\text{C}(\text{O})\text{OCH}_2\text{CH}$), 3.62 (s, 1.5H, $-\text{OCH}_3^{\text{trans}}$), 3.46-3.56 (m, 3.5H, $-\text{C}(\text{O})\text{OCH}_2\text{CH}$, $-\text{OCH}_3^{\text{cis}}$), 2.26-2.43 (m, 2H, $\text{Pro}_{\beta 1}$,

Pro β 2); ^{13}C NMR (75 MHz, CDCl_3 , 298K) δ = 172.9, 154.9, 144.0, 143.9, 141.3 (2), 138.5, 138.37, 127.1-130.2 (aromatic), 125.5, 120.3, 103.9, 82.7, 79.9, 76.6, 75.9, 75.1, 74.2, 74.0, 73.9, 73.6, 69.4, 68.1, 58.3, 52.8, 52.5, 47.8, 37.8; MS (ES) calc. for $\text{C}_{55}\text{H}_{55}\text{NNaO}_{10}$ ($\text{M} + \text{Na}$) $^+$: 912.37. Found ($\text{M} + \text{Na}$) $^+$: 912.23;

4-O-(2,3,4,6-Tetra-O-benzyl- α,β -D-galactopyranosyl)-N-acetyl-cis-4-hydroxy-L-proline methyl ester (12a,b)

The general preparation method was followed for *N*-terminal acylation of **11a** (0.116 g, 0.129 mmol) using 3:1 DCM:piperidine (6 mL). The reaction mixture was purified by flash chromatography using 2:1 ethyl acetate/hexanes to yield **12a** as a clear oil (0.070 g, 0.247 mmol) (76.1%):

12a: $[\alpha]_{\text{D}}^{25} = +37.3^\circ$ (*c* 1.0 CHCl_3); ^1H NMR (300 MHz, CDCl_3 , 298K) δ = 7.20-7.44 (m, 20H, aromatic), 4.32-4.98 (m, 10H, H₁, Pro α , -OCH₂Ph), [4.19-4.28, 0.2H, Pro γ ^{*cis*}], 4.08-4.19 (m, 0.8H, Pro γ ^{*trans*}), 3.28-4.08 (m, 11H, H₂, H₃, H₄, H₅, H_{6a}, H_{6b}, Pro δ ₁, Pro δ ₂, -CO₂CH₃), [2.46-2.56, m, 0.2H, Pro β ₁^{*cis*}], 2.21-2.40 (m, 1.8H, Pro β ₁^{*trans*}, Pro β ₂), [2.01, s, 0.6H, -NCOCH₃^{*cis*}], 1.88 (s, 2.4H, -NCOCH₃^{*trans*}); ^{13}C NMR (75 MHz, CDCl_3 , 298K) δ = (171.7), 171.6, (169.9), 169.8, 138.85, 138.76, (138.6), 138.4, (138.1), 137.8, 127.3-128.5 (aromatic), 99.1, (97.8), 78.6, 77.9, 76.6, (76.2), 75.4, (74.9), (74.8), 74.7, 73.6, (73.37), (73.30), 73.24, 73.22, (72.9), 70.5, 70.2, (69.7), (68.6), (58.9), 56.8, 53.5, (52.9), (52.5), 52.1, (37.6), 35.9, 22.0, (21.5); MS (ES) calc. for $\text{C}_{42}\text{H}_{47}\text{NNaO}_9$ ($\text{M} + \text{Na}$) $^+$: 732.31. Found ($\text{M} + \text{Na}$) $^+$: 732.98;

12b: $[\alpha]_D^{25} = -12.4^\circ$ (*c* 0.7 CHCl₃); ¹H NMR (300 MHz, CDCl₃, 298K) $\delta = 7.17$ - 7.51 (m, 20H, aromatic), 4.51-5.01 (m, 7H, Pro _{α} , -OCH₂Ph), 4.29-4.51 (m, 4H, H₁, Pro _{γ} , -OCH₂Ph), 3.43-3.95 (m, 11H, H₂, H₃, H₄, H₅, H_{6a}, H_{6b}, Pro _{δ 1}, Pro _{δ 2}, -CO₂CH₃), [2.70-2.79, m, 0.33H, Pro _{β 1}^{*cis*}], 2.33-2.49 (m, 1.67H, Pro _{β 1}^{*trans*}, Pro _{β 2}), [2.03, s, 1H, -NCOCH₃^{*cis*}], 1.98 (s, 2H, -NCOCH₃^{*trans*}); ¹³C NMR (75 MHz, CDCl₃, 298K) $\delta =$ (171.5), 171.4, (170.1), 169.5, 138.8, (138.6), (138.5), 138.4, 138.3, (137.9), 137.8, 127.5-128.6 (aromatic), 103.3, 82.3, (82.1), (79.3), 79.1, 77.2, (76.1), (75.3), 75.1, (74.6), 74.5, 73.64, 73.60, (73.55), (73.51), (73.38), 73.2, 73.1, 68.96, 68.92, (59.0), 53.5, (52.9), (52.5), 52.2, (38.2), 35.9, 22.2, (22.1); MS (ES) calc. for C₄₂H₄₇NNaO₉ (M + Na)⁺: 732.31. Found (M + Na)⁺: 732.39;

Table 11.1: IR Absorbance Bands (cm^{-1}) for Prolyl Amide Isomers^[a]

Compound	Ester	Amide
1	1731	1612
2	1729	1611.5
3	1729	1611.5
4	1730.5	1612
5	1729	1611
6	1726.5	1613

^[a]Determined in D₂O at 25 °C.**Table 11.2:** Selected ³J Coupling Constants (Hz) for **1-6** at 25 °C (and 67 °C)^{[a],[b]}

Compound	J _{Hα,Hβ1}	J _{Hα,Hβ2}	J _{Hβ1,Hγ}	J _{Hβ2,Hγ}	Pucker
1	8.4 (8.3)	8.6 (8.3)	4.4 (4.9)	2.1 (2.6)	C ^{γ} -exo
2	8.3 (7.9)	8.7 (8.3)	4.8 (5.3)	1.4 (2.0)	C ^{γ} -exo
3	8.3 (8.2)	8.5 (8.2)	4.6 (5.6)	1.8 (2.4)	C ^{γ} -exo
4	9.6 (9.4)	2.4 (2.8)	4.6 (4.8)	2.5 (2.8)	C ^{γ} -endo
5	7.8 (9.6)	4.3 (2.8)	4.3 (5.0)	2.8 (2.6)	C ^{γ} -endo
6	9.8 (9.8)	2.5 (2.8)	4.5 (4.7)	2.5 (2.6)	C ^{γ} -endo

^[a]Determined in D₂O; ^[b]These values were calculated with the aid of an iterative simulation using the NUMARIT algorithm (A. R. Quirt, J. S. Martin, *J. Magn. Reson.* **1971**, *5*, 318-327). Interactive line assignment was carried out using Spinworks 3.0 (Copyright © 2009, University of Manitoba, Canada).

Table 11.3: Amide Isomer Equilibrium ($K_{trans/cis}$) of **1-6** at Various Temperatures^[a]

Compound	Temperature (°C) ^[b]						
	24.8	30.0	37.0	46.0	51.3	56.6	67.3
1	6.18 ±0.04	5.97 ±0.03	5.65 ±0.10	5.33 ±0.06	5.14 ±0.05	4.94 ±0.08	4.56 ±0.04
2	5.99 ±0.01	6.12 ±0.11	5.58 ±0.12	5.35 ±0.20	5.04 ±0.08	4.92 ±0.05	4.54 ±0.10
3	5.87 ±0.17	5.74 ±0.20	5.18 ±0.05	4.97 ±0.03	4.78 ±0.04	4.82 ±0.08	4.26 ±0.10
4	2.40 ±0.03	2.40 ±0.04	2.38 ±0.03	2.41 ±0.04	2.35 ±0.06	2.30 ±0.09	2.25 ±0.13
5	2.91 ±0.26	2.77 ±0.22	2.76 ±0.39	2.74 ±0.33	2.58 ±0.22	2.48 ±0.19	2.36 ±0.29
6	2.86 ±0.03	2.79 ±0.01	2.75 ±0.02	2.61 ±0.13	2.52 ±0.08	2.48 ±0.09	2.37 ±0.13

^[a]Carried out in D₂O, ±SE determined by integration of two or more sets of

trans/cis isomers; ^[b]Calibrated using an ethylene glycol standard.

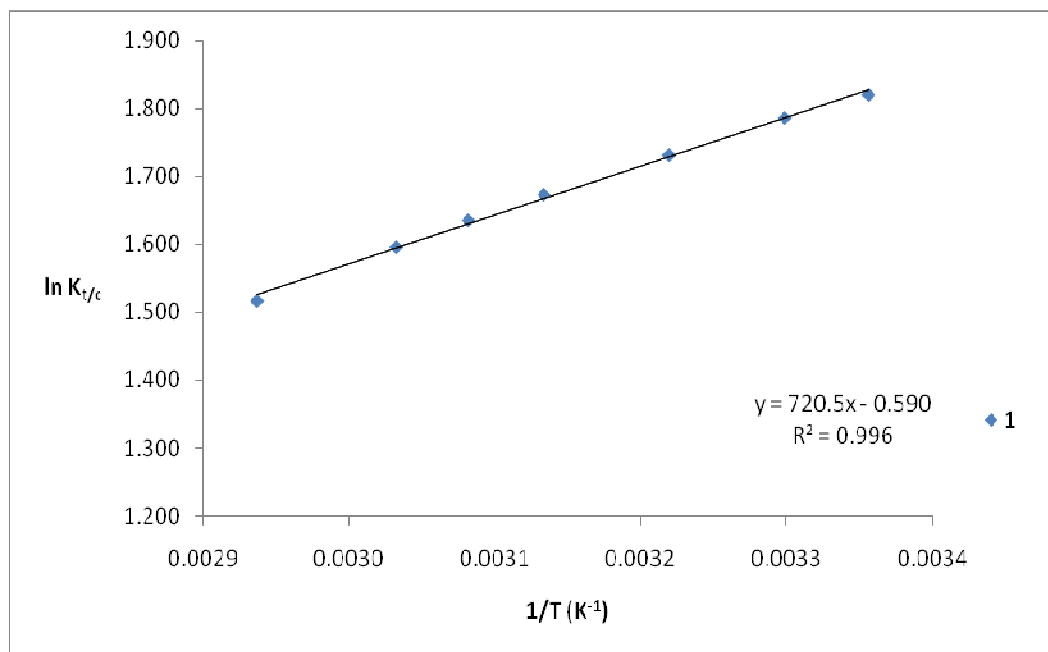
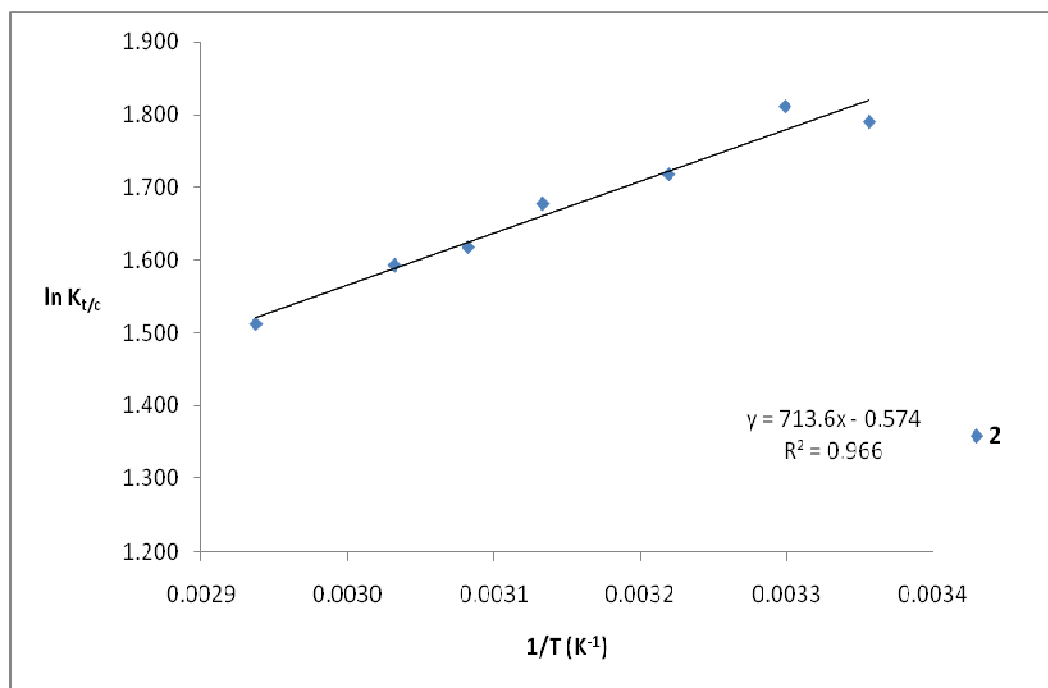
Figure 11.1: Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **1** in D₂O**Figure 11.2:** Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **2** in D₂O

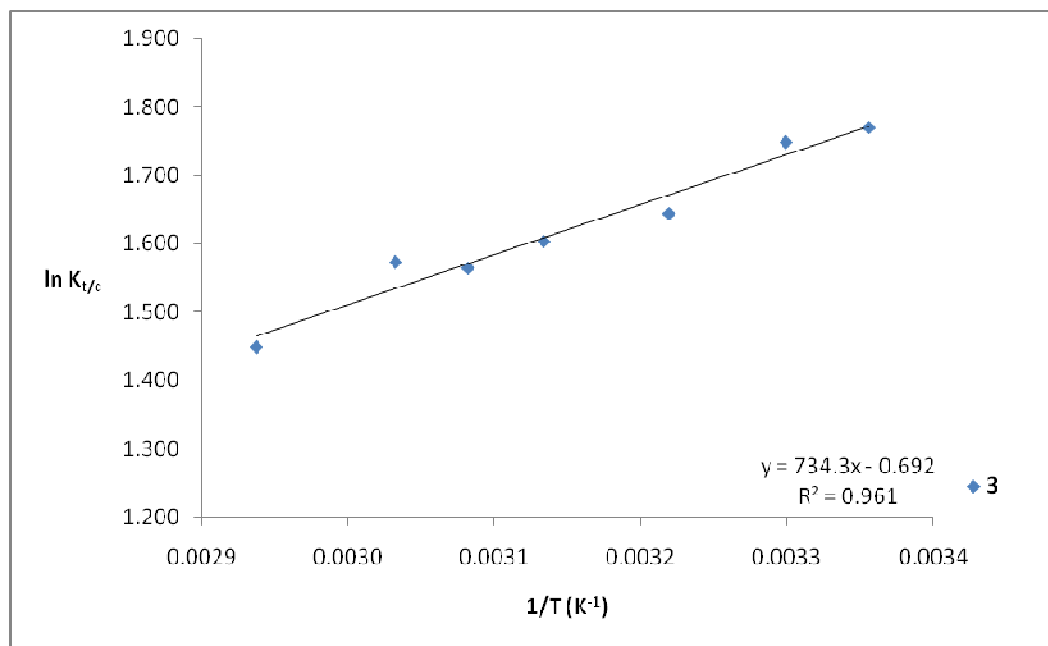
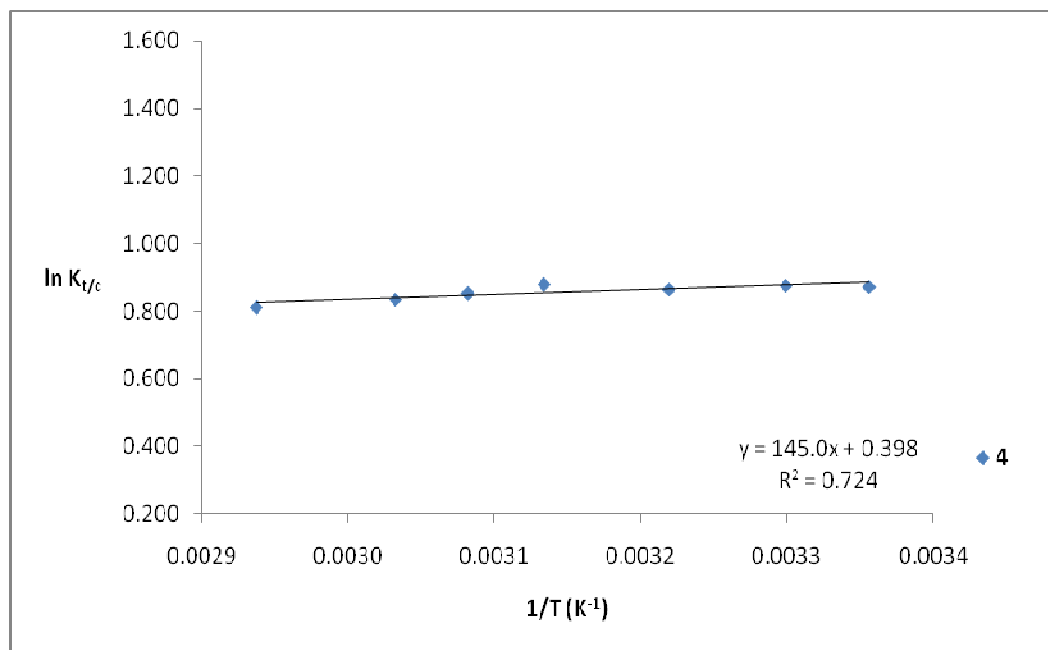
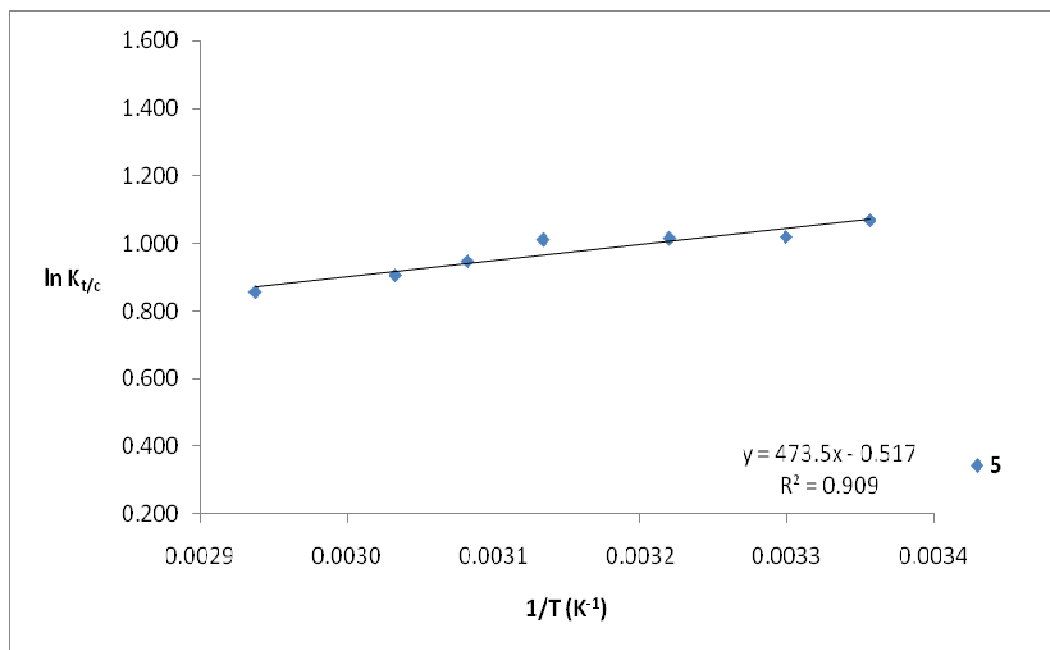
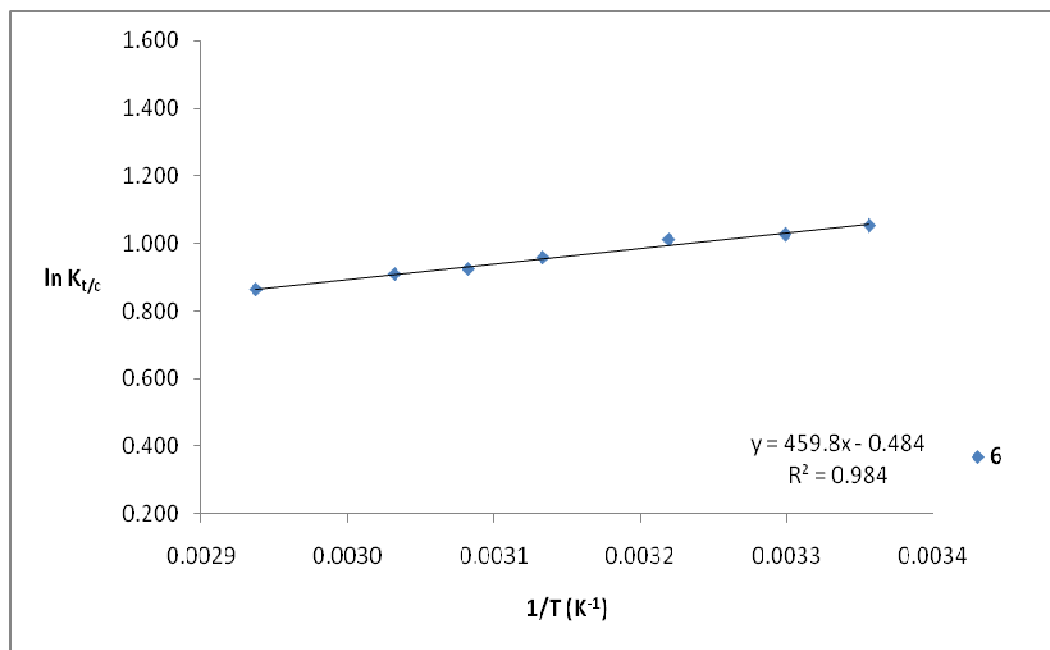
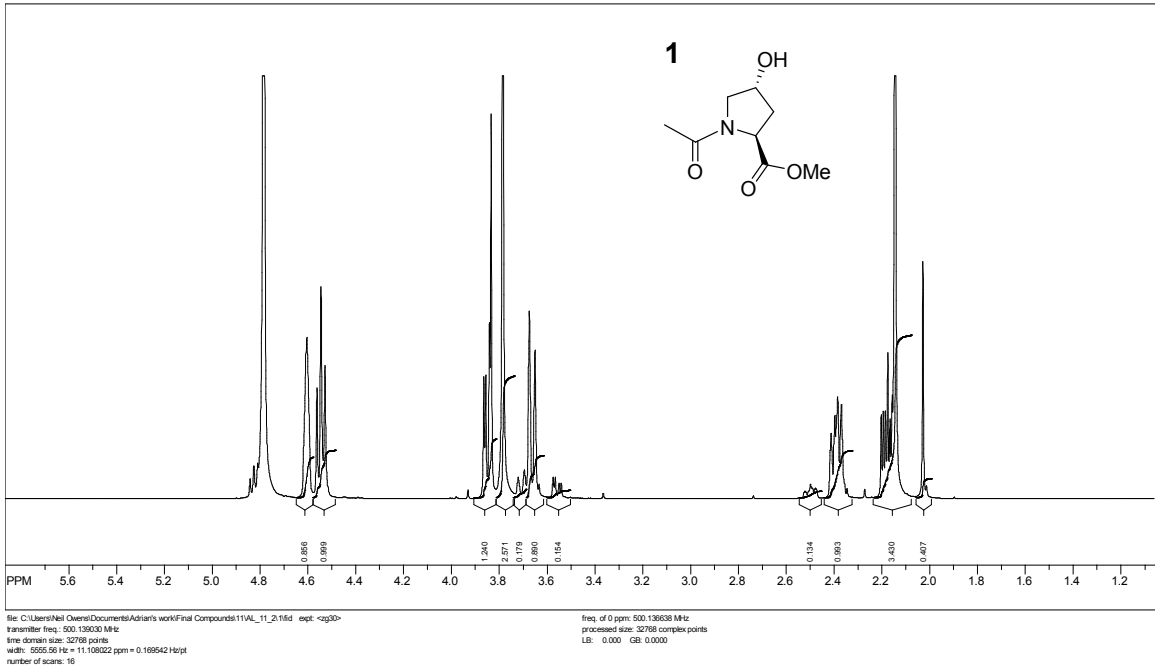
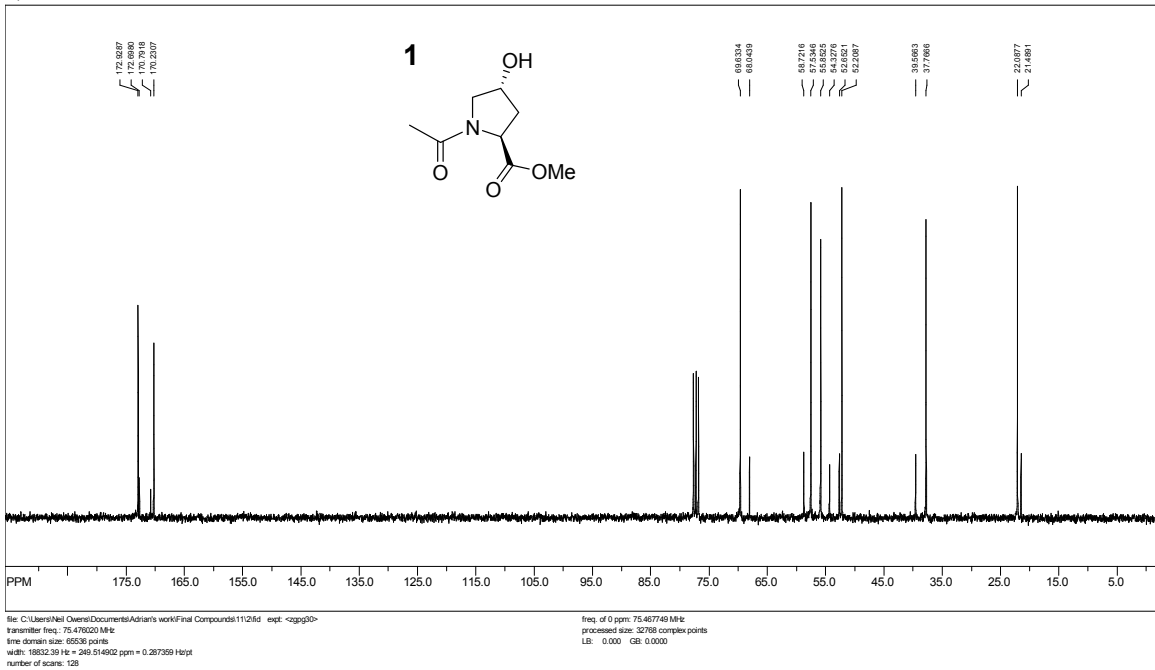
Figure 11.3: Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **3** in D₂O**Figure 11.4:** Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **4** in D₂O

Figure 11.5: Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **5** in D₂O**Figure 11.6:** Van't Hoff Plot for the *cis*-to-*trans* Isomerization of **6** in D₂O

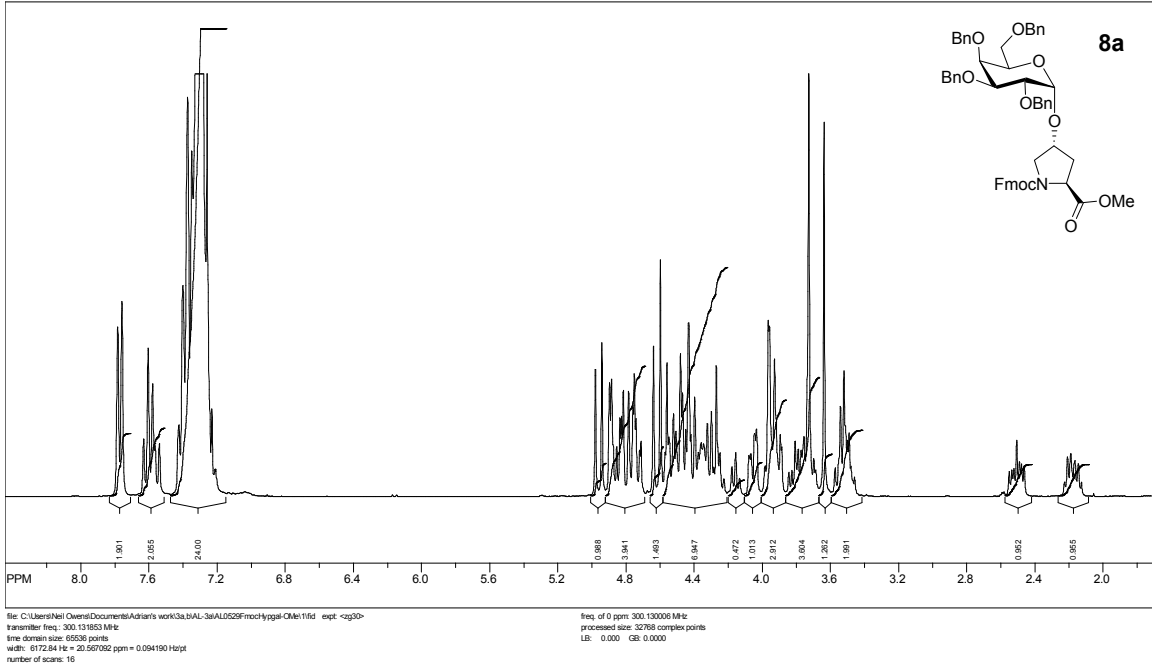
SpinWorks 2.5: AL 4R unglycosylated, 1-H 500 MHz in D2O at 25 C



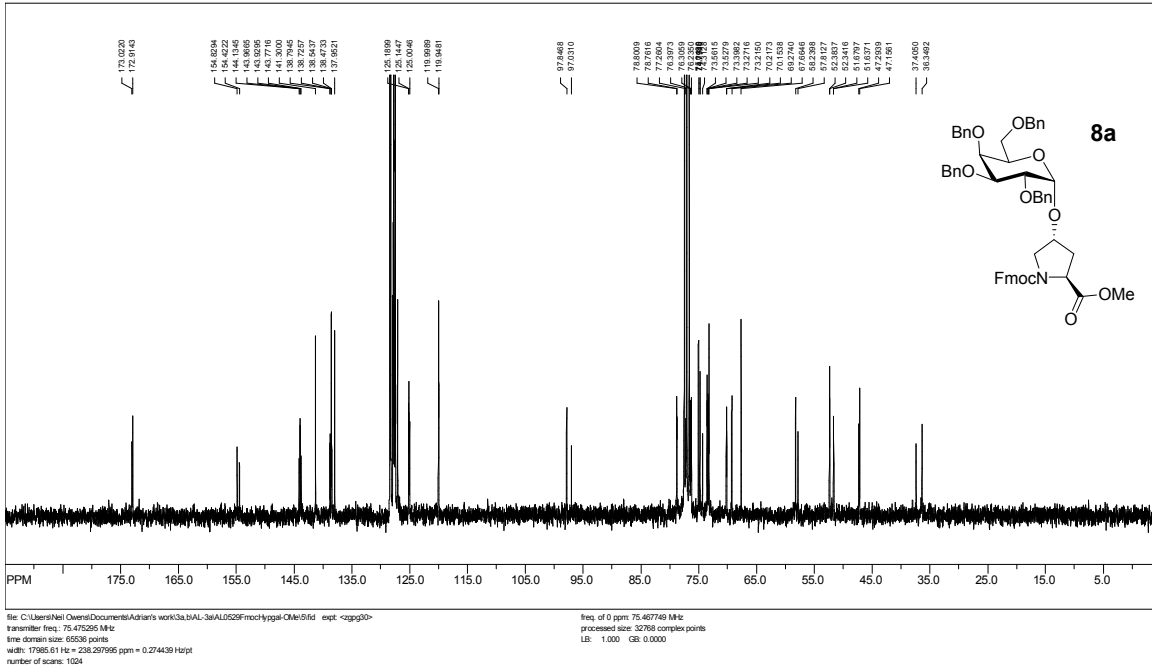
SpinWorks 2.5: C13CPD32 CDCl3 u schweiz 2



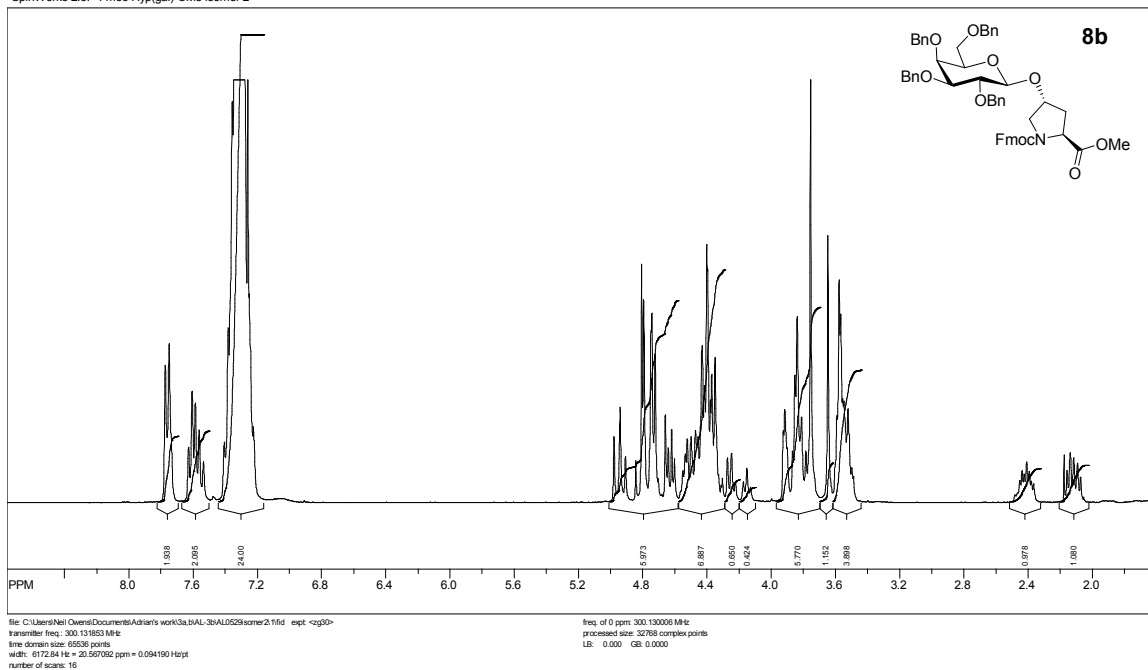
SpinWorks 2.5: Fmoc-Hyp(gal)-OMe



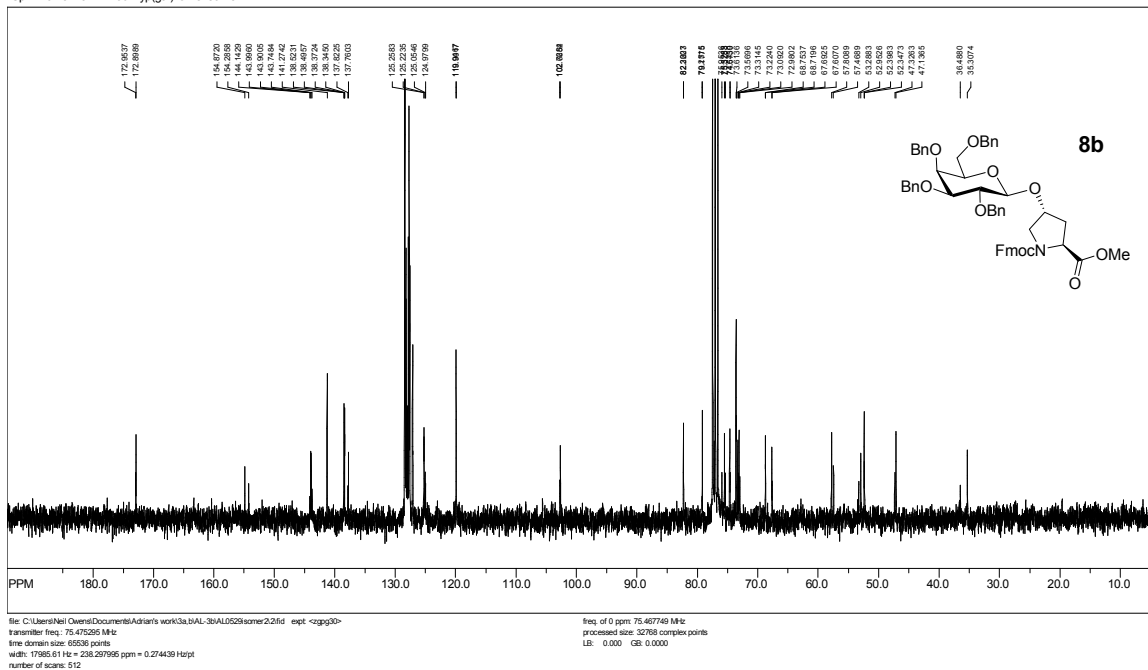
SpinWorks 2.5: Fmoc-Hyp(gal)-OMe



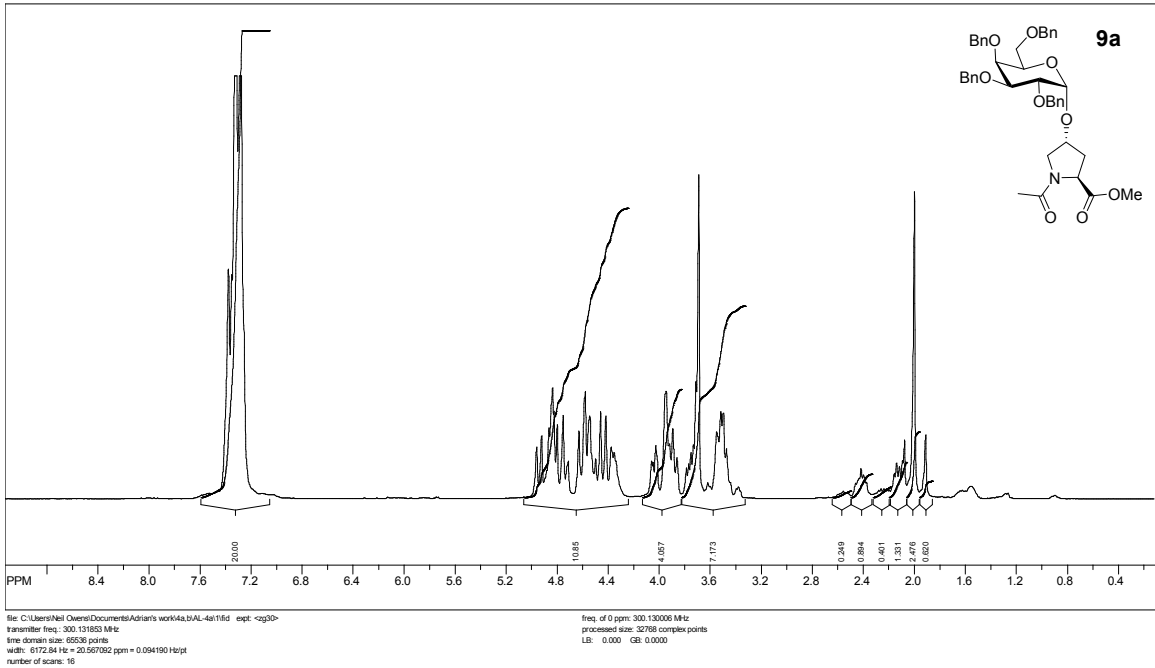
SpinWorks 2.5: Fmoc-Hyp(gal)-OMe isomer 2



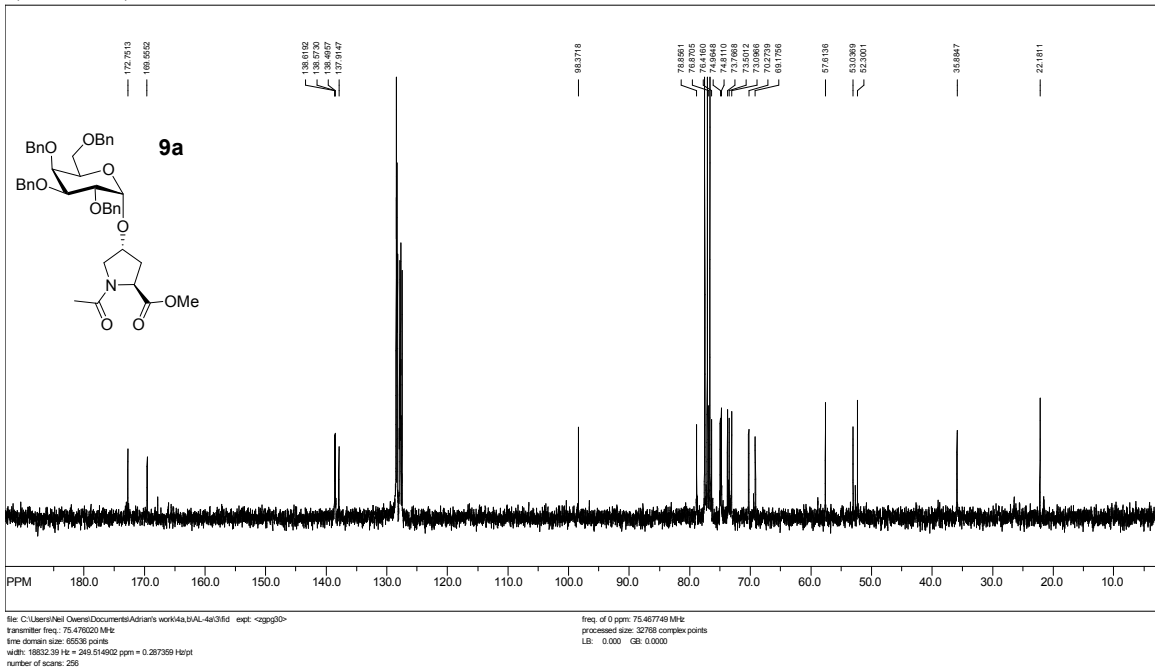
SpinWorks 2.5: Fmoc-Hyp(gal)-OMe isomer 2



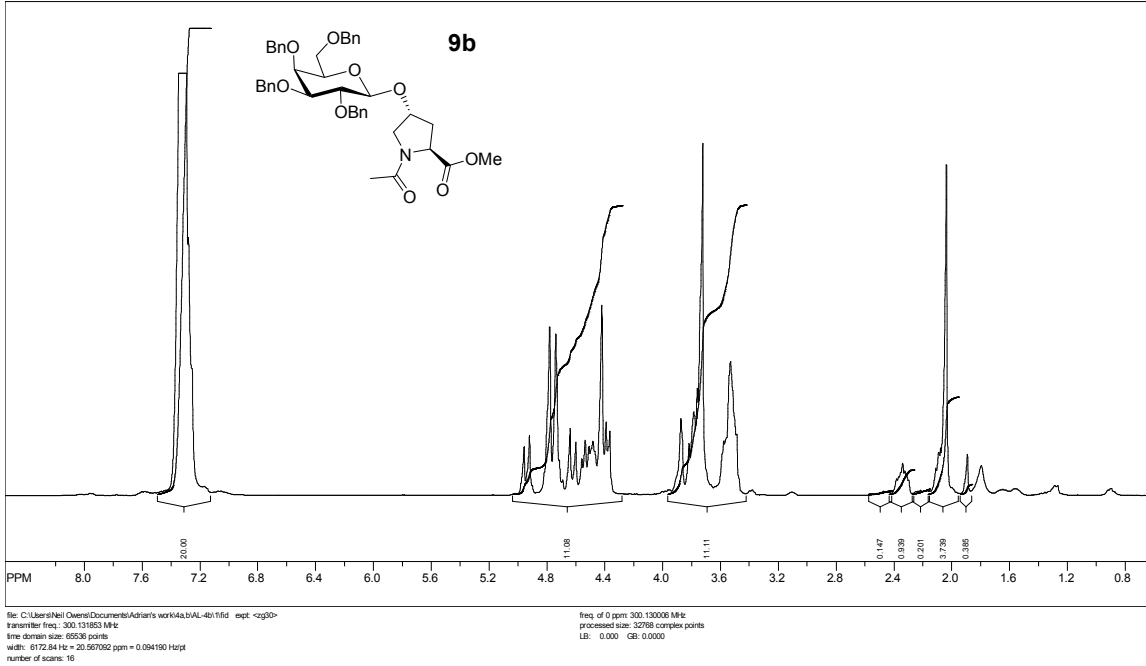
SpinWorks 2.5: 4R-alpha



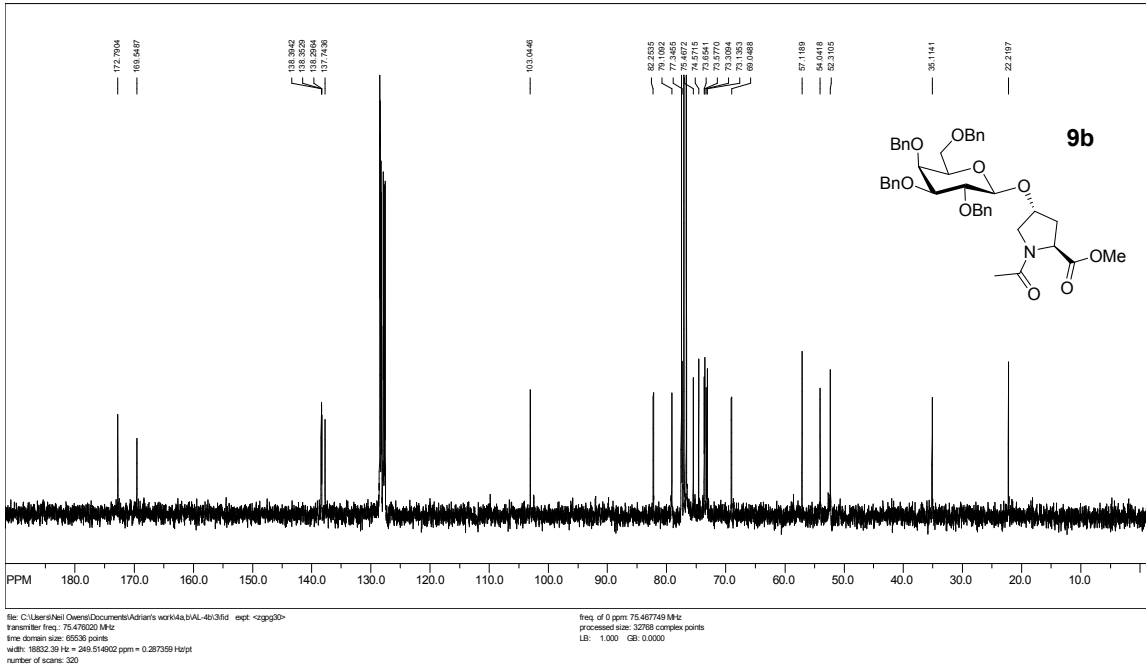
SpinWorks 2.5: 4R-alpha



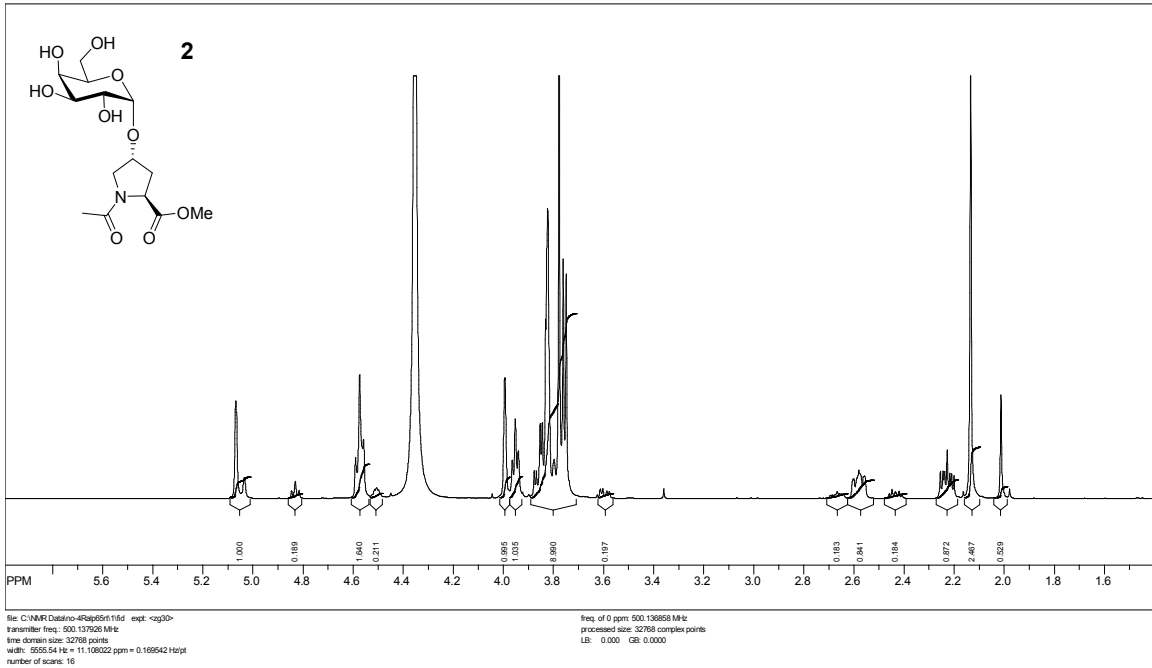
SpinWorks 2.5: 4R Beta



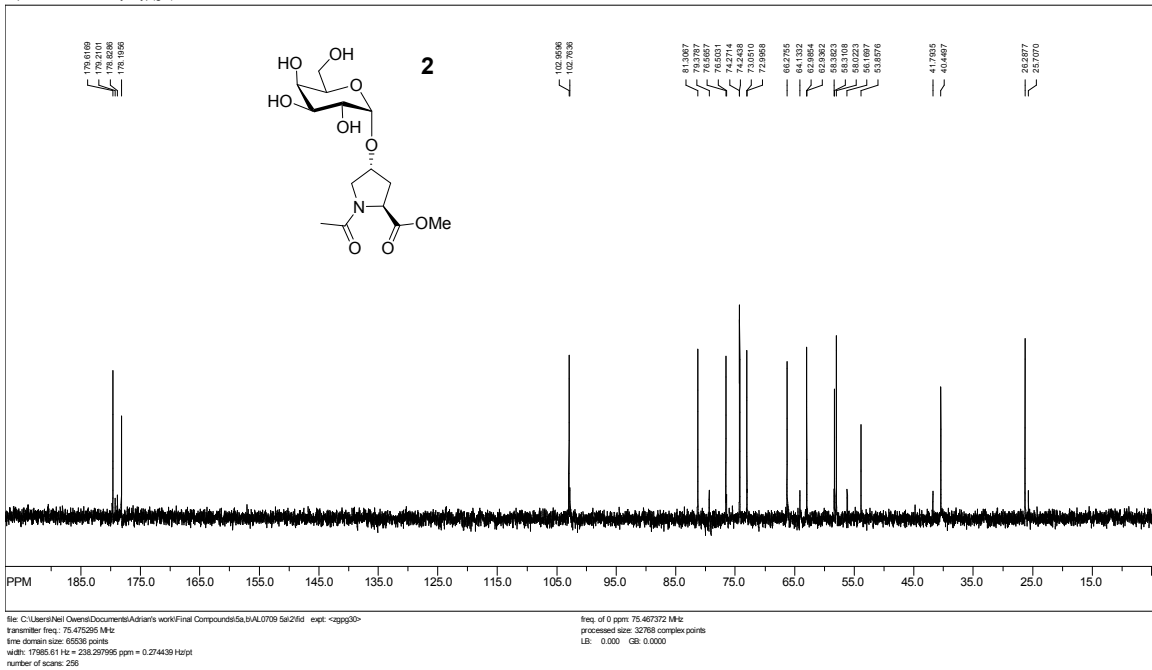
SpinWorks 2.5: 4R Beta



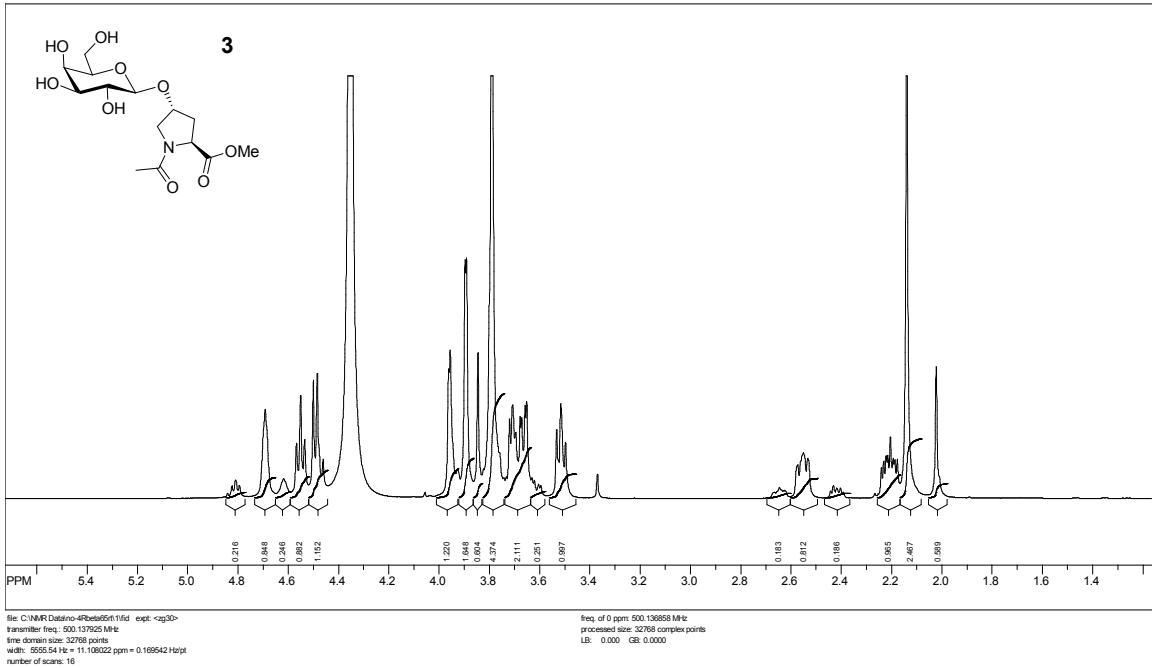
SpinWorks 2.5:



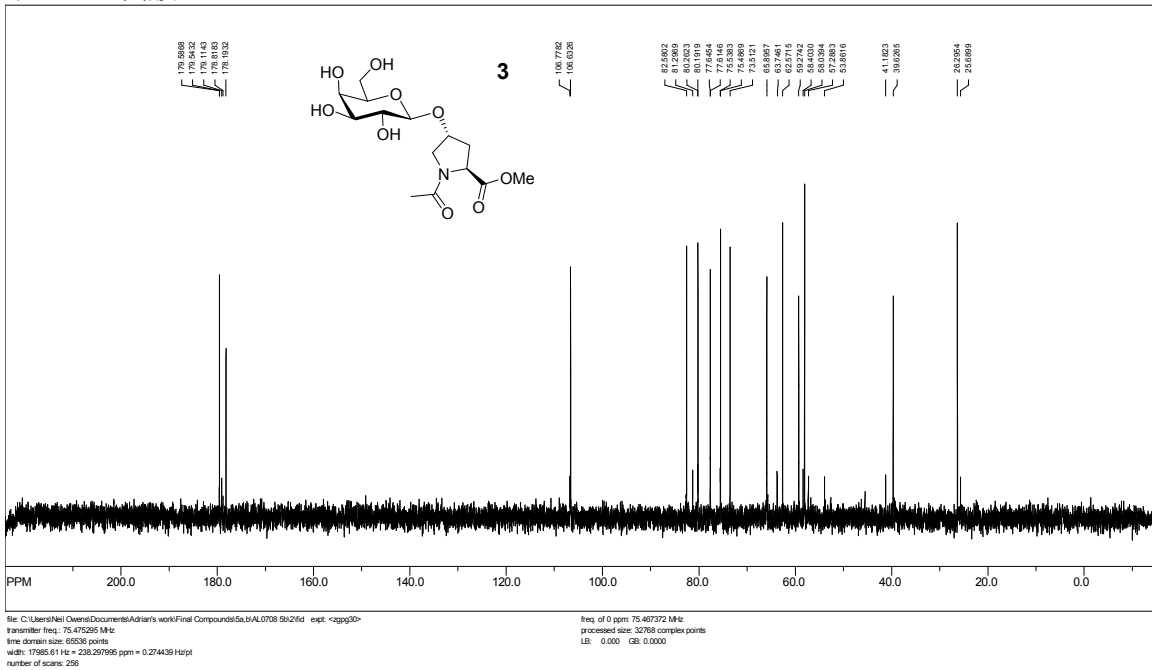
SpinWorks 2.5: Acetyl-Hyp(gal)-OMe



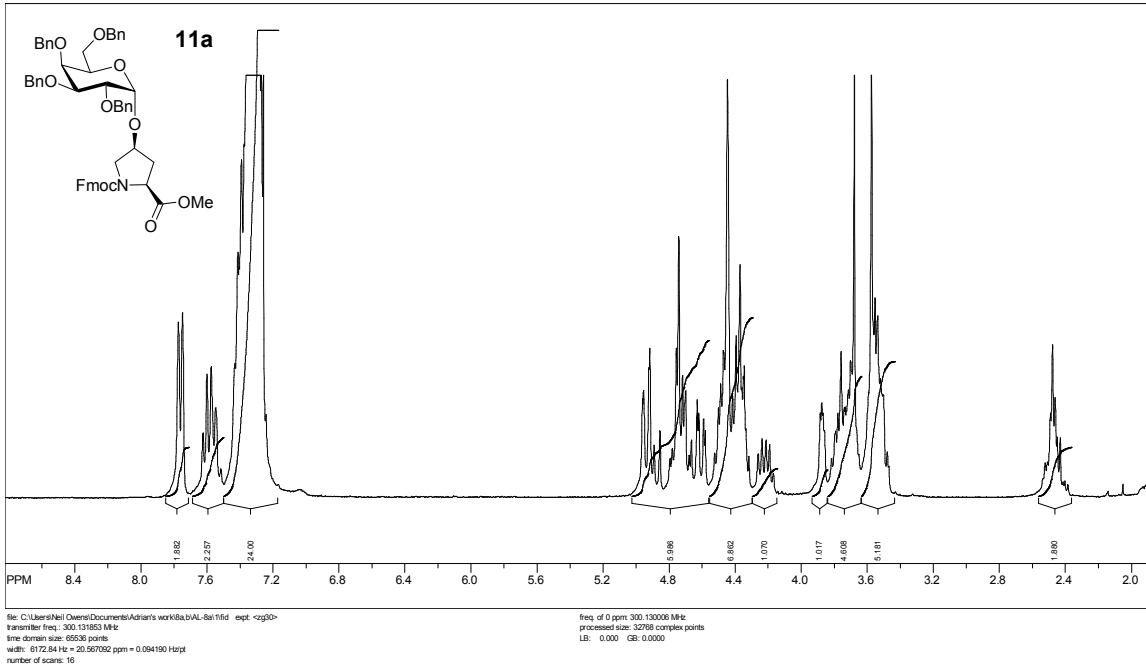
SpinWorks 2.5:



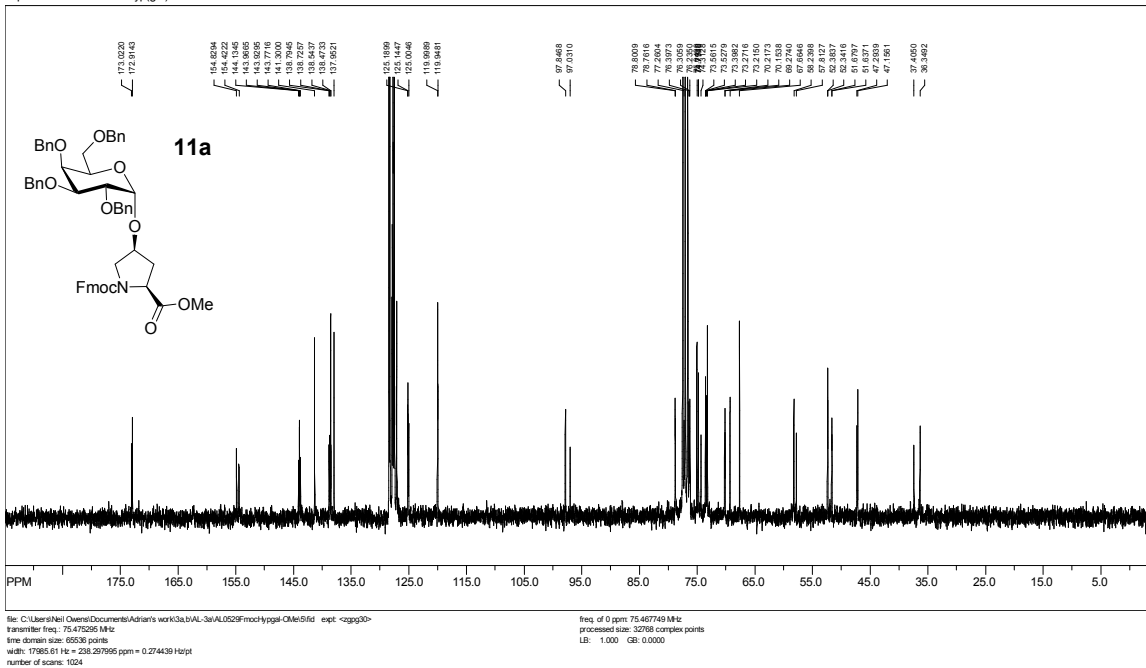
SpinWorks 2.5: Acetyl-Hyp(gal)-OMe debenz isomer 2



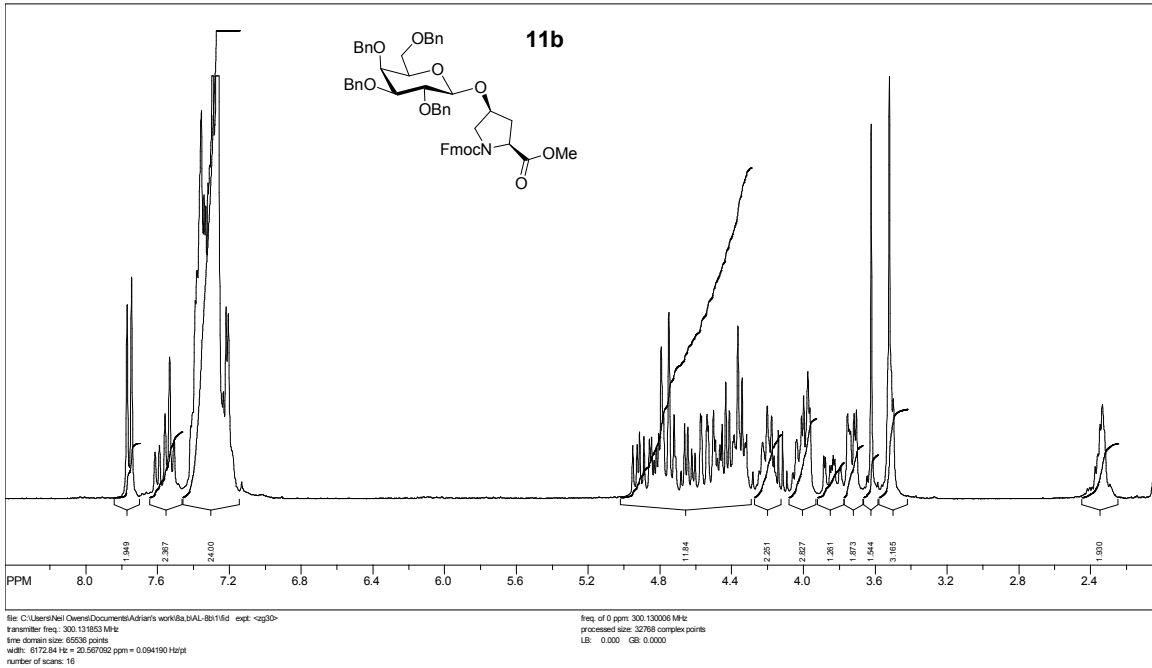
SpinWorks 2.5: Isomer 2



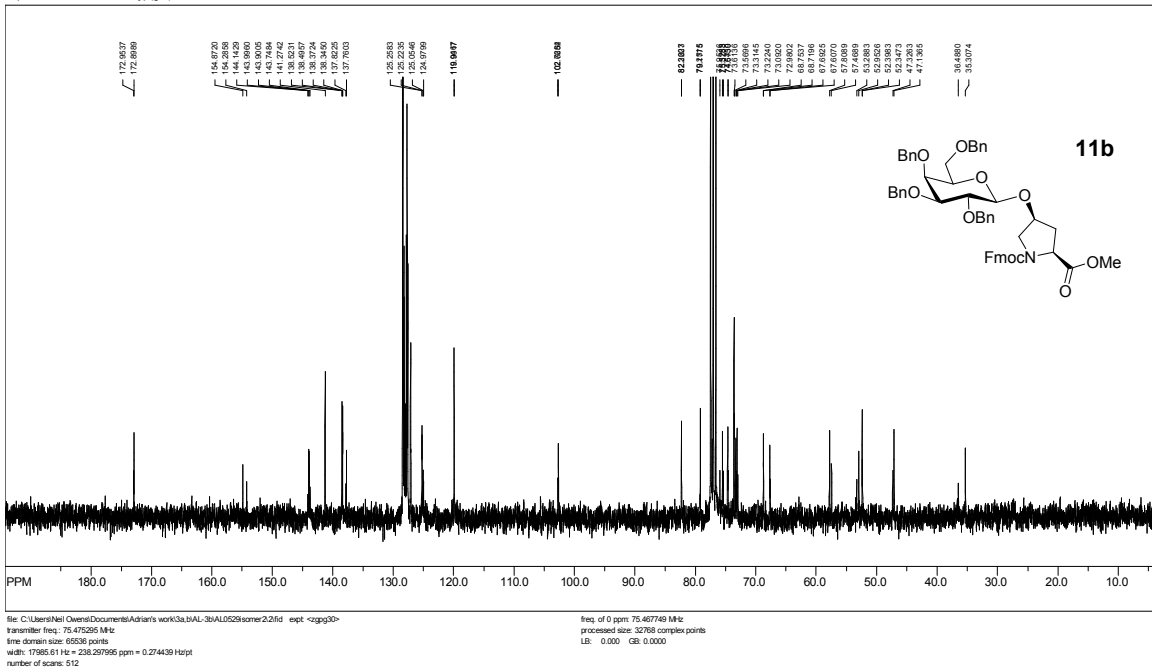
SpinWorks 2.5: Fmoc-Hyp(gal)-OMe



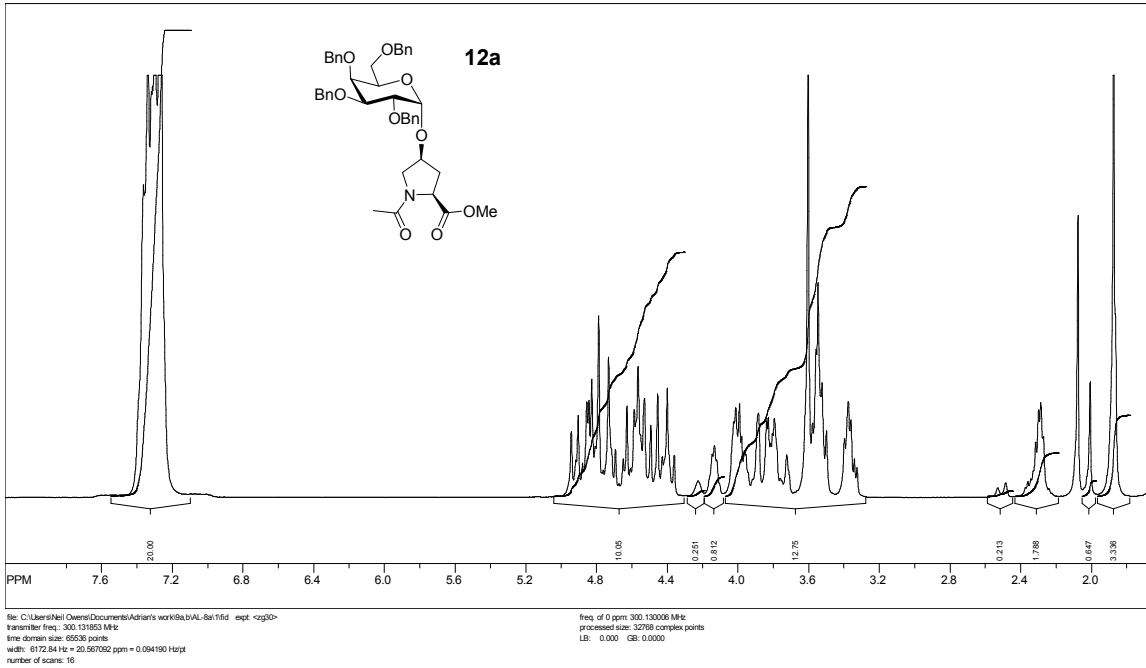
SpinWorks 2.5: Isomer 1



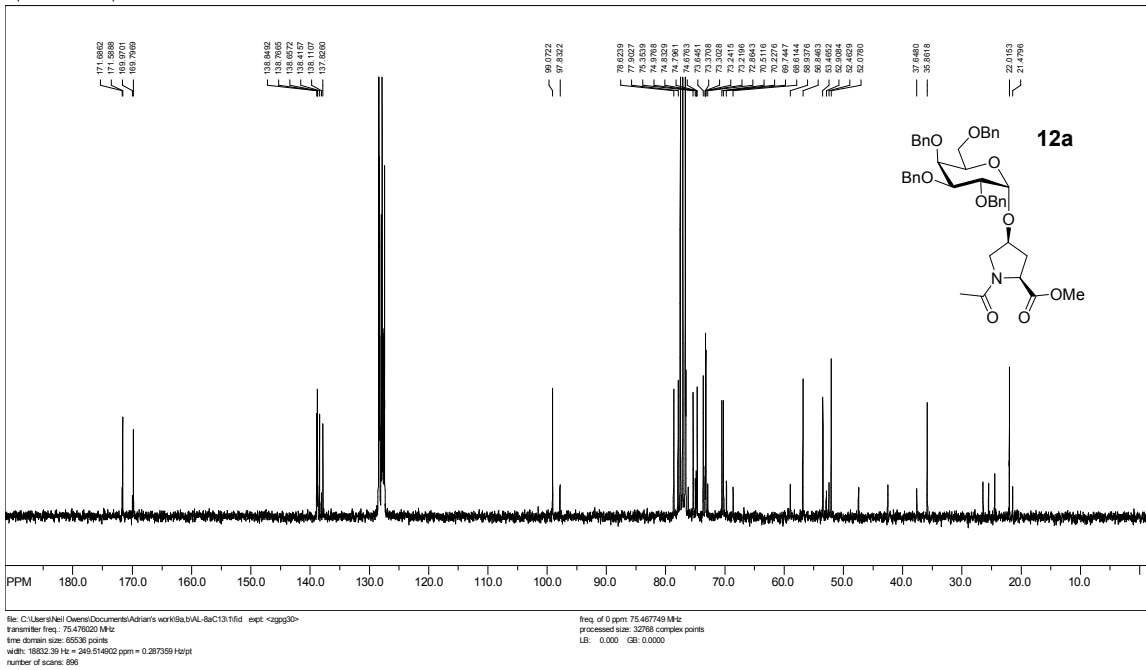
SpinWorks 2.5: Fmoc-Hyp(gal)-OMe isomer 2



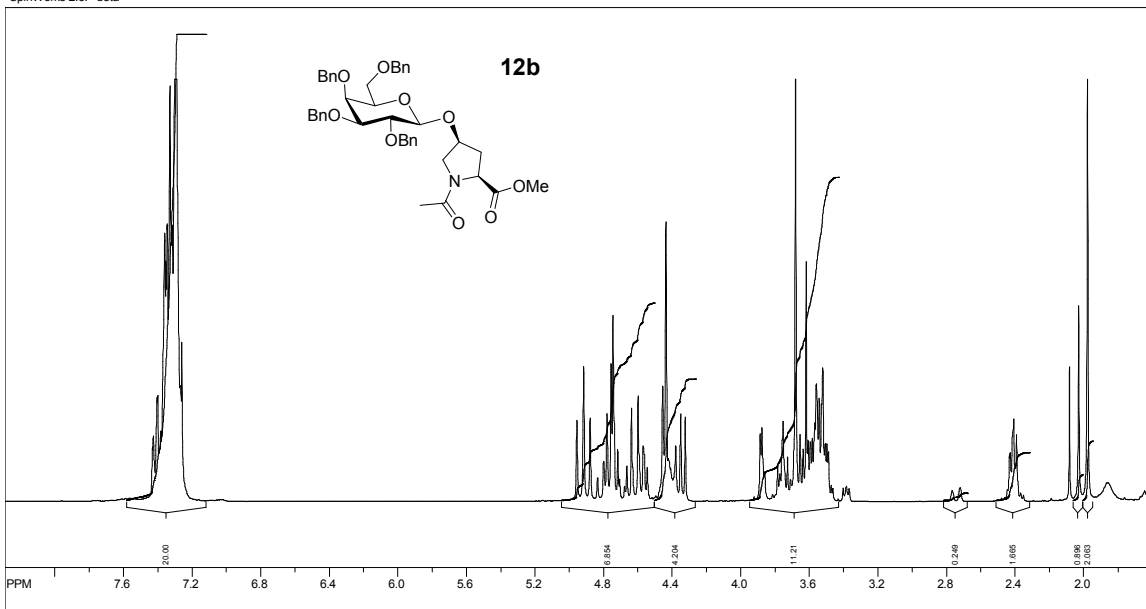
SpinWorks 2.5: Isomer 1



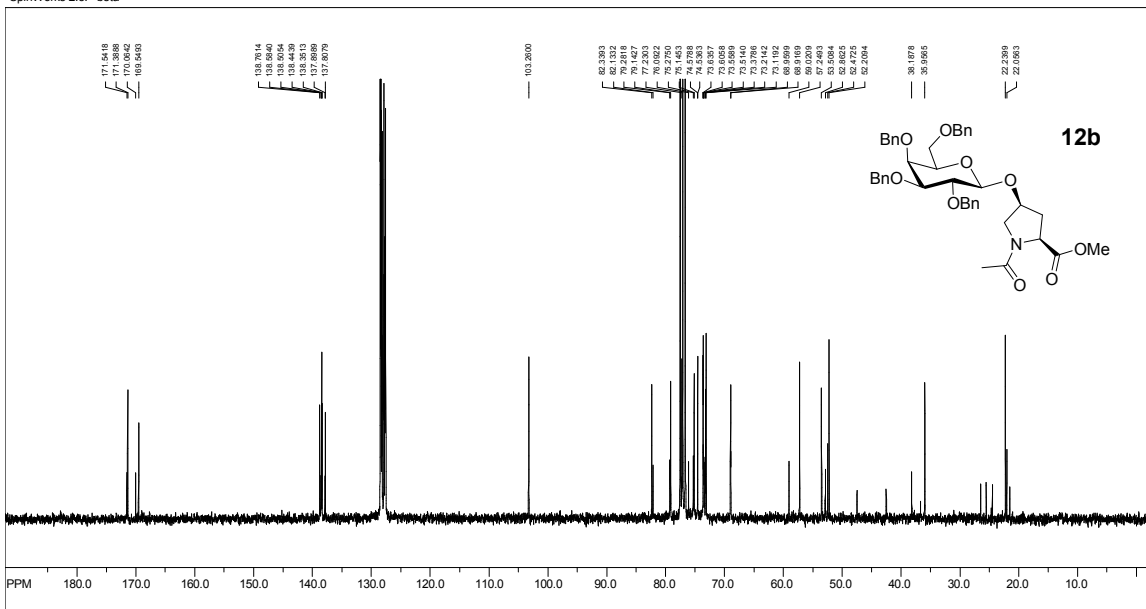
SpinWorks 2.5: alpha



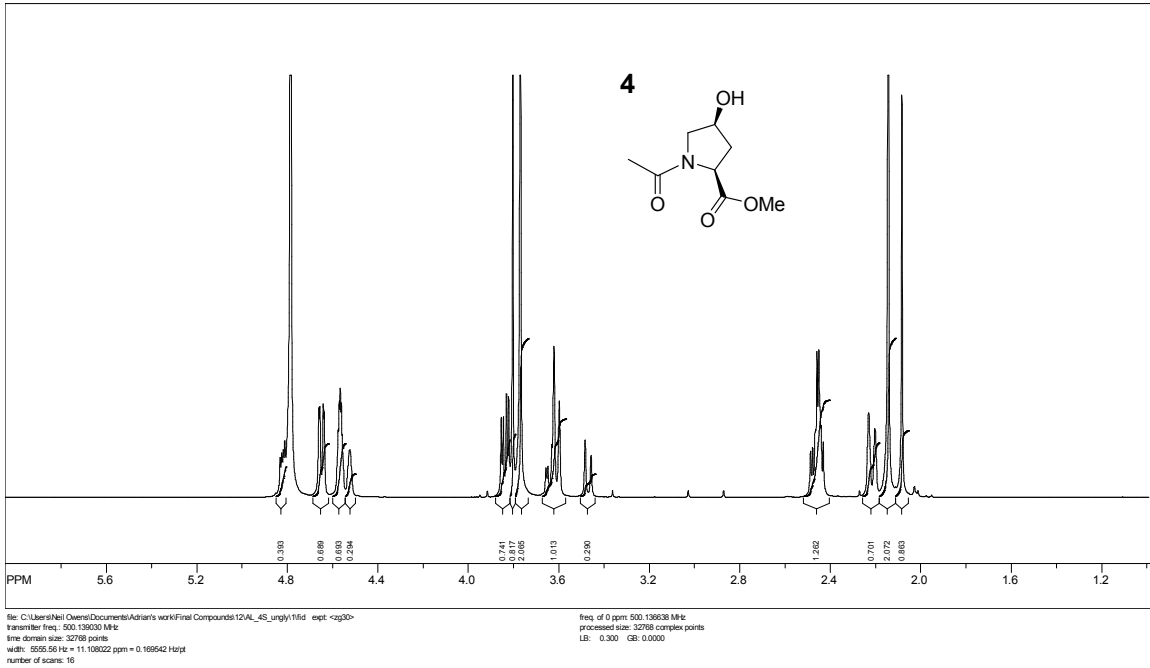
SpinWorks 2.5: beta



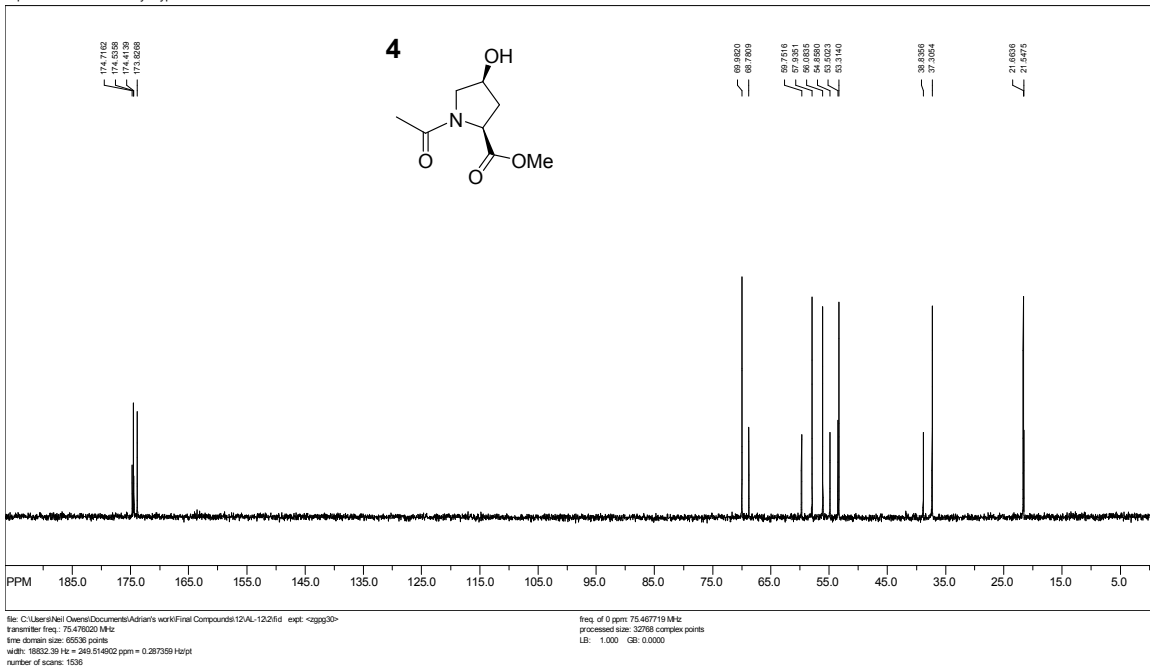
SpinWorks 2.5: beta



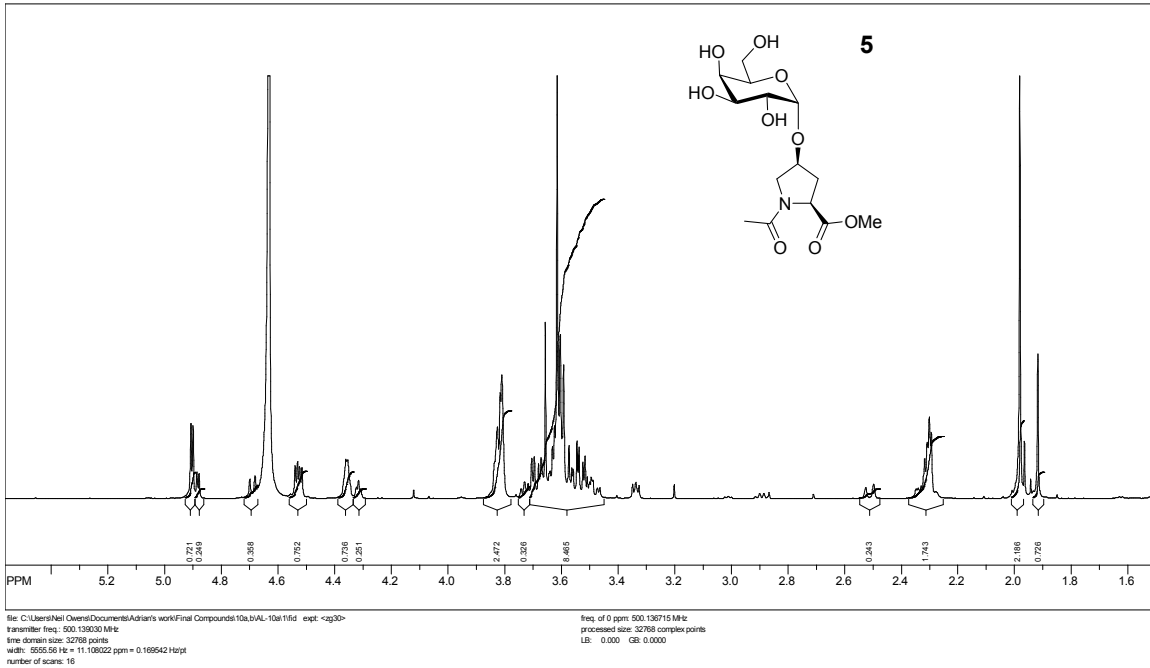
SpinWorks 2.5: AL 4S ungly, 1-H at 500 MHz in D2O, 25 deg C



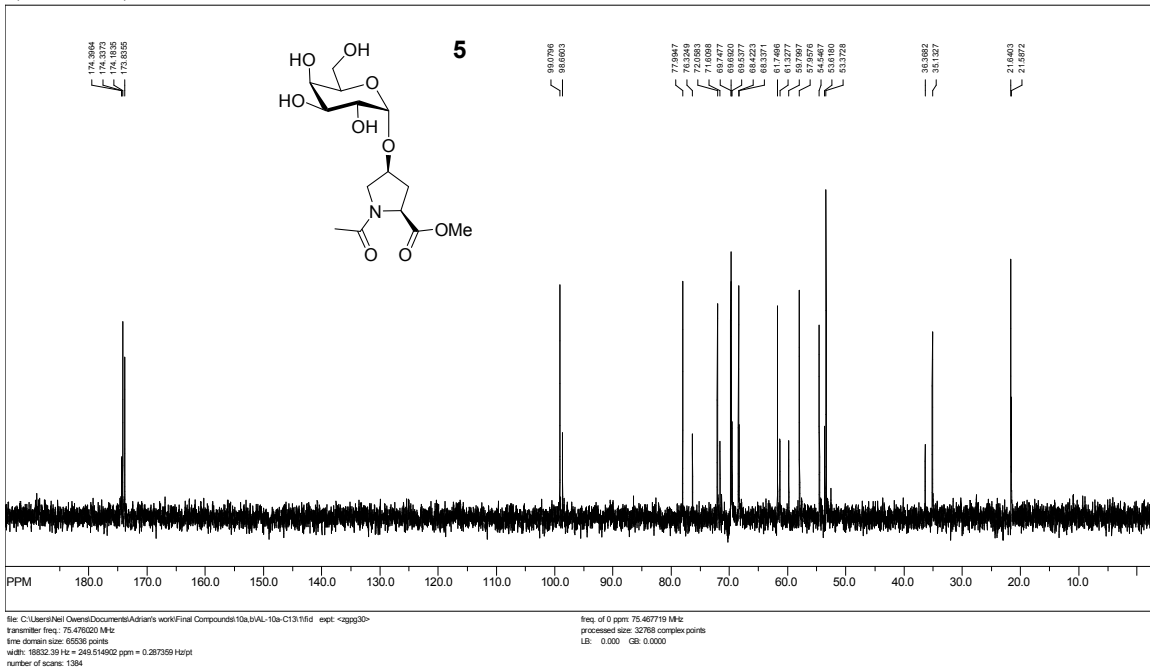
SpinWorks 2.5: 4S Acetyl-Hyp-OMe



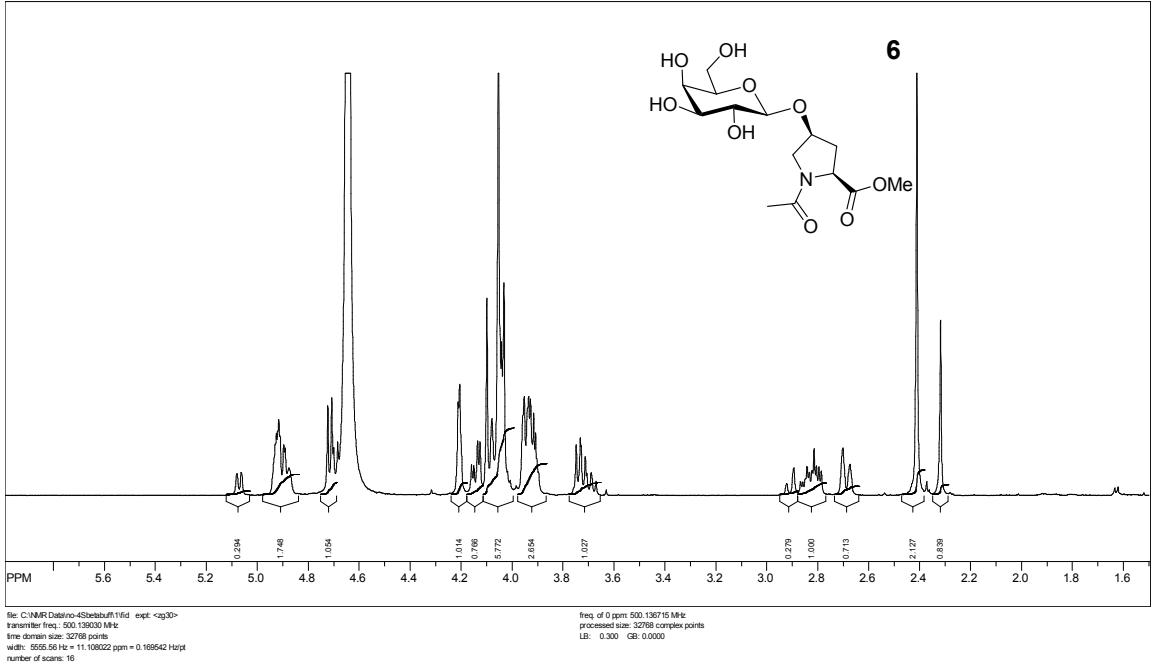
SpinWorks 2.5: standard 1-H survey parameters, AMX500, D2O



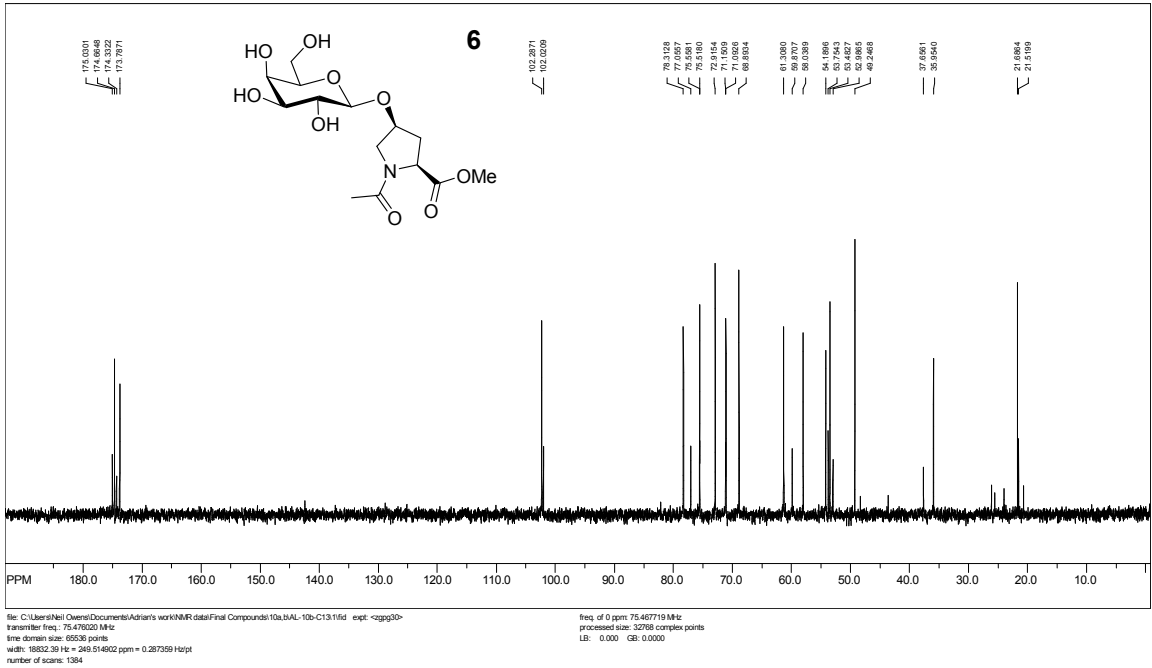
SpinWorks 2.5: 4S alpha



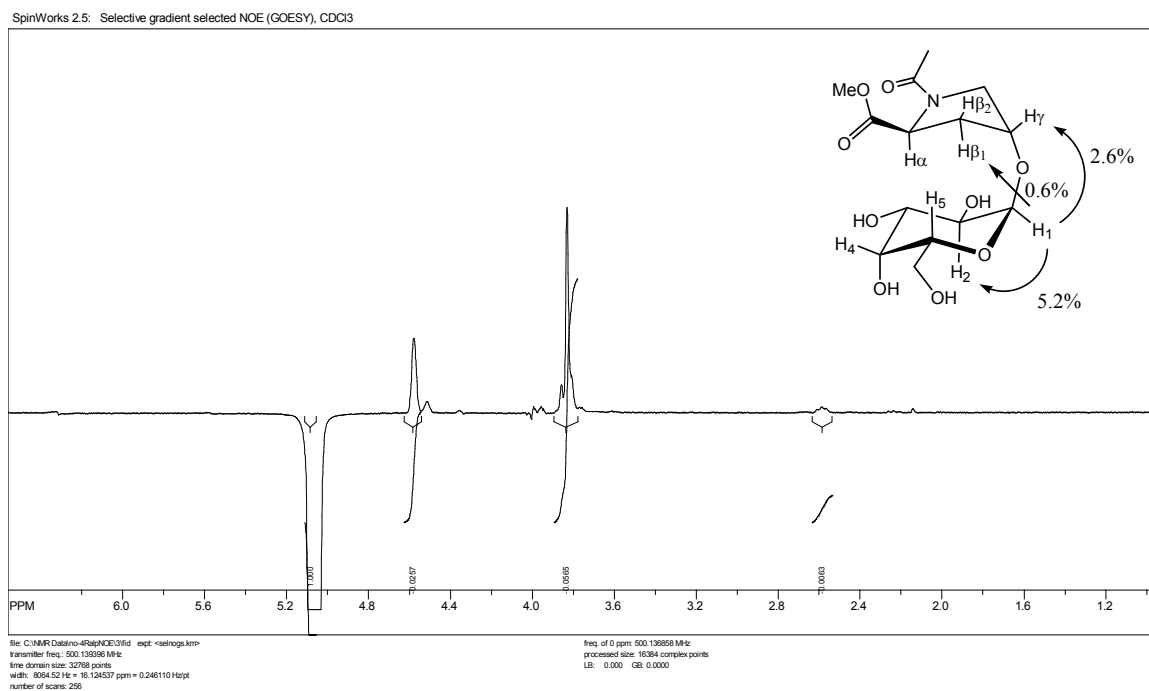
SpinWorks 2.5:



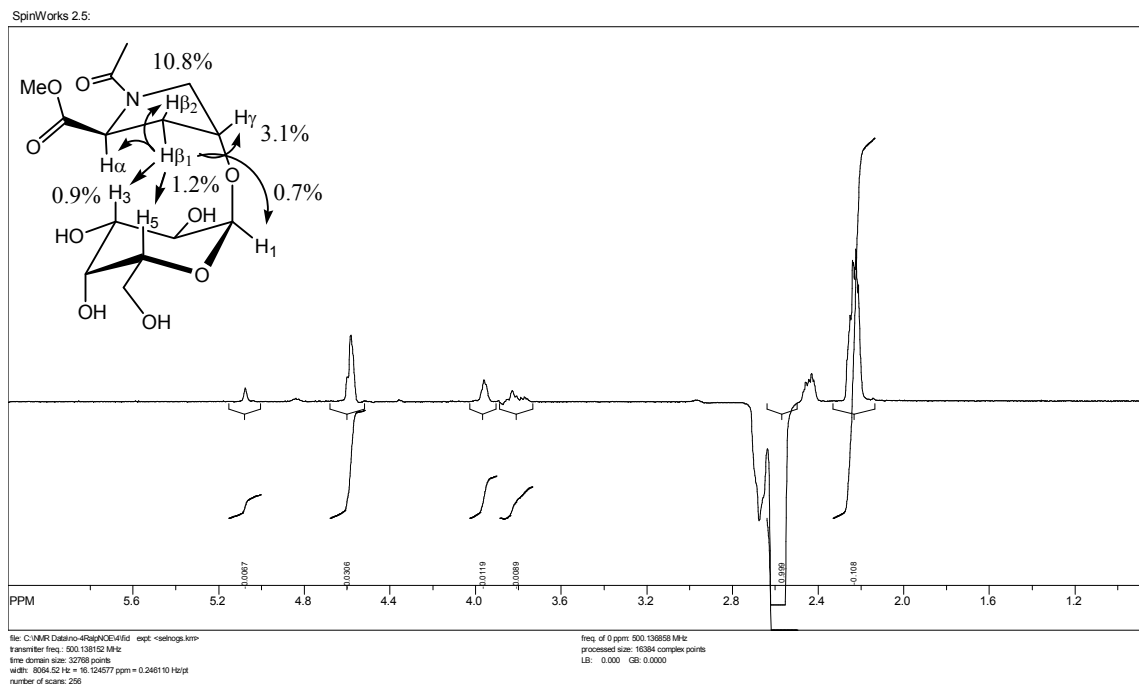
SpinWorks 2.5: 4S Beta



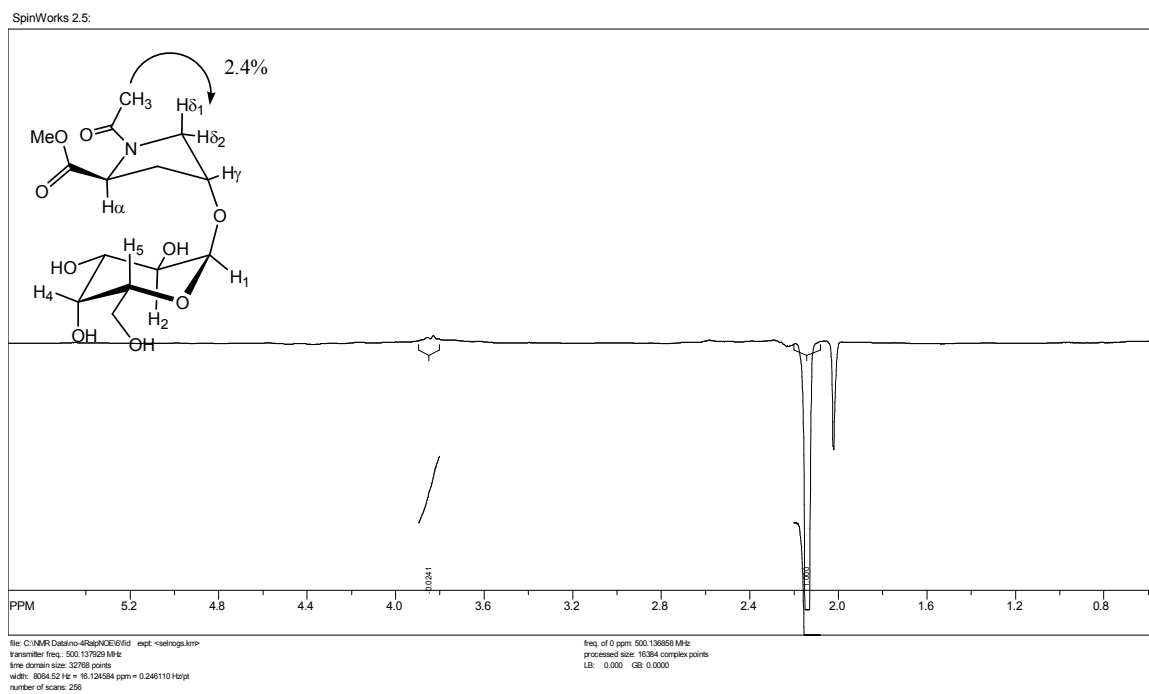
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating H₁^{trans} (5.07 ppm, 0.8H) in **2** showed inter-proton effect to Pro_γ^{trans} (4.57 ppm, 0.8H) (2.6%), H₂^{trans} (3.83 ppm, 0.8H) (5.2%), and Pro_{β1}^{trans} (2.58 ppm, 0.8H) (0.6%). Inversion exchange to H₁^{cis} (5.02 ppm, 0.2H) resulted in inter-proton effect to Pro_γ^{cis} (4.51 ppm, 0.2H) (3.2%) relative to to H₁^{cis}.



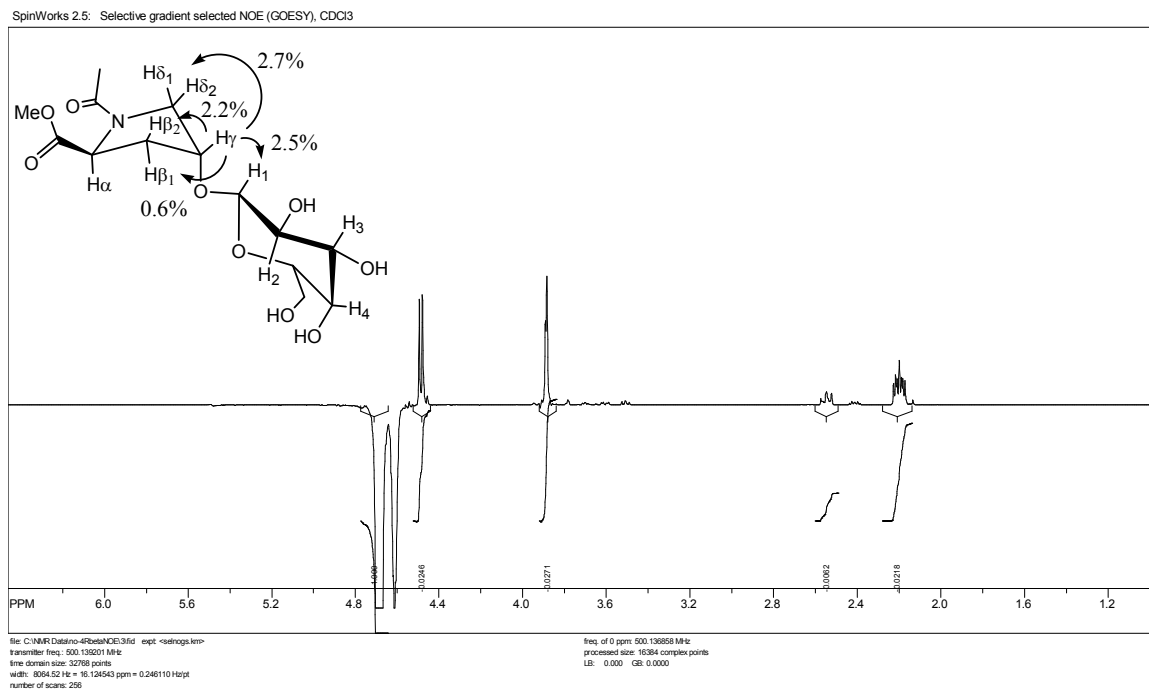
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating $\text{Pro}_{\beta_1}^{trans}$ (2.58 ppm, 0.8H) in **2** showed inter-proton effect to H_1^{trans} (5.07 ppm, 0.8H) (0.7%), $\text{Pro}_{\alpha}^{trans}$ / $\text{Pro}_{\gamma}^{trans}$ (4.54-4.61 ppm, 1.6H) (3.1%), H_5^{trans} (3.98-4.01 ppm, 0.8H) (1.2%), H_3^{trans} (3.93-3.98 ppm, 0.8H) (0.9%), and $\text{Pro}_{\beta_2}^{trans}$ (2.23 ppm, 0.8H) (10.8%). Inversion exchange to $\text{Pro}_{\beta_1}^{cis}$ (2.68 ppm, 0.2H) resulted in inter-proton effect to $\text{Pro}_{\alpha}^{cis}$ (4.83 ppm, 0.2H) (0.4%) and $\text{Pro}_{\beta_2}^{cis}$ (2.44 ppm, 0.2H) (1.9%) relative to to $\text{Pro}_{\beta_1}^{cis}$.



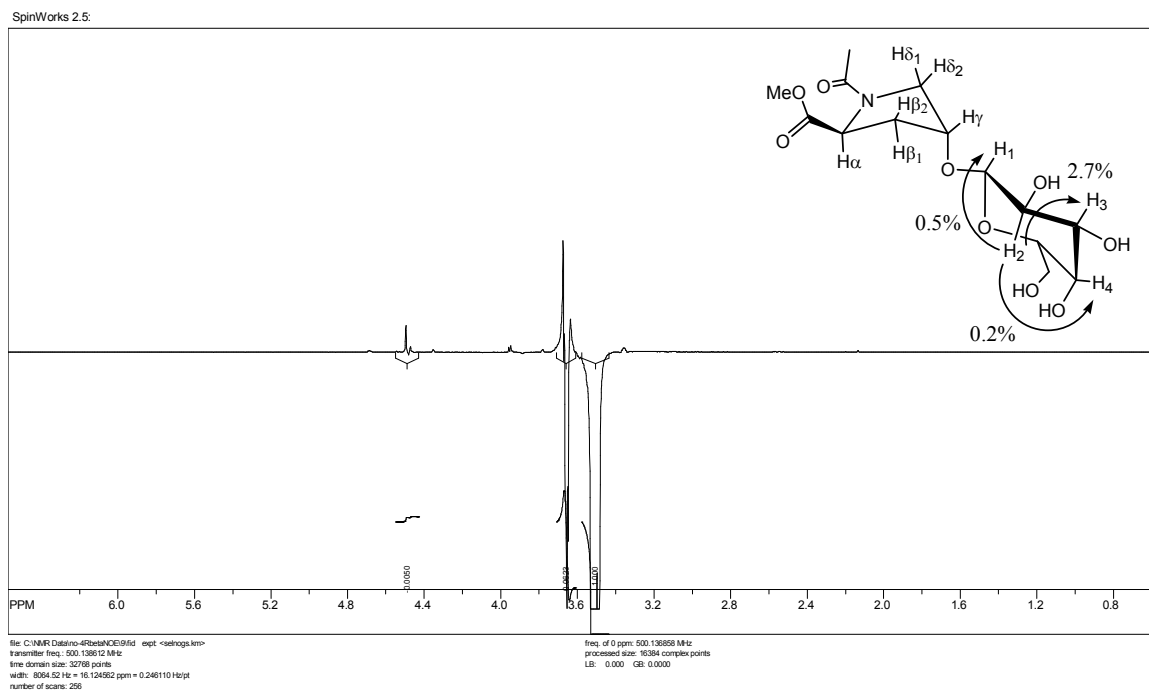
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating -NCOCH₃^{trans} (2.13 ppm, 2.4H) in **2** showed inter-proton effects to Pro_{δ1}^{trans} / Pro_{δ2}^{trans} (3.84 ppm, 1.6H) (2.4%).



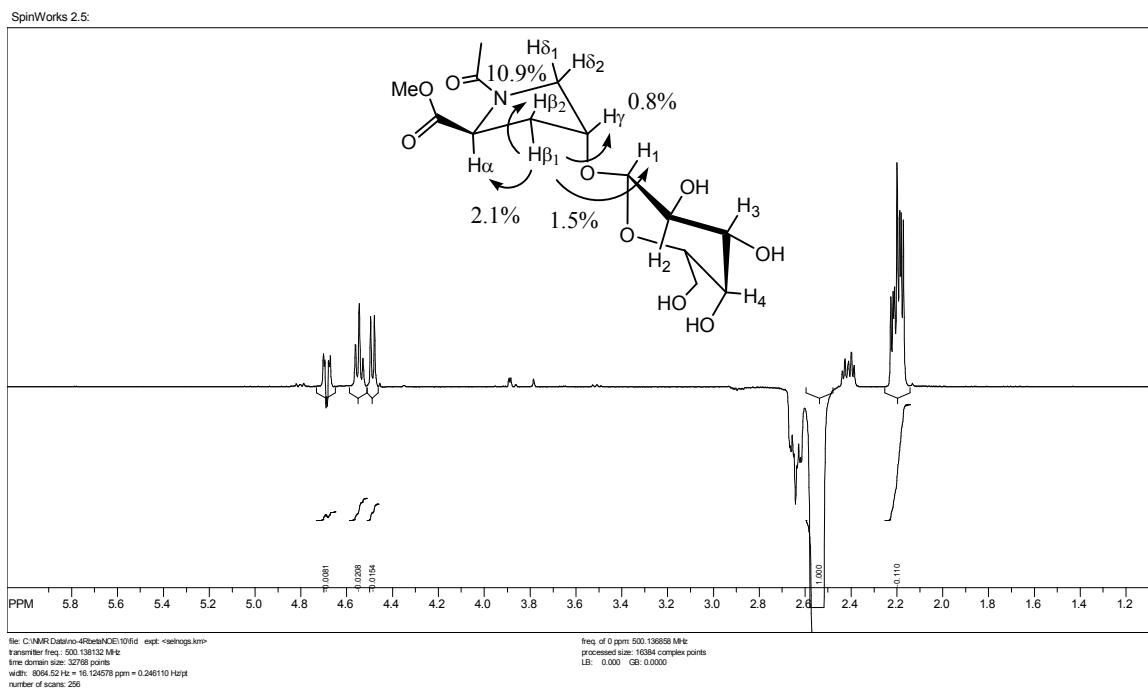
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating Pro_γ^{trans} (4.69 ppm, 0.8H) in **3** showed inter-proton effect to H₁^{trans} (4.49 ppm, 0.8H) (2.5%), Pro_{δ1}^{trans} / Pro_{δ2}^{trans} (3.89 ppm, 1.6H) (2.7%), Pro_{β1}^{trans} (2.55 ppm, 0.8H) (0.6%), Pro_{β2}^{trans} (2.21 ppm, 0.8H) (2.2%).



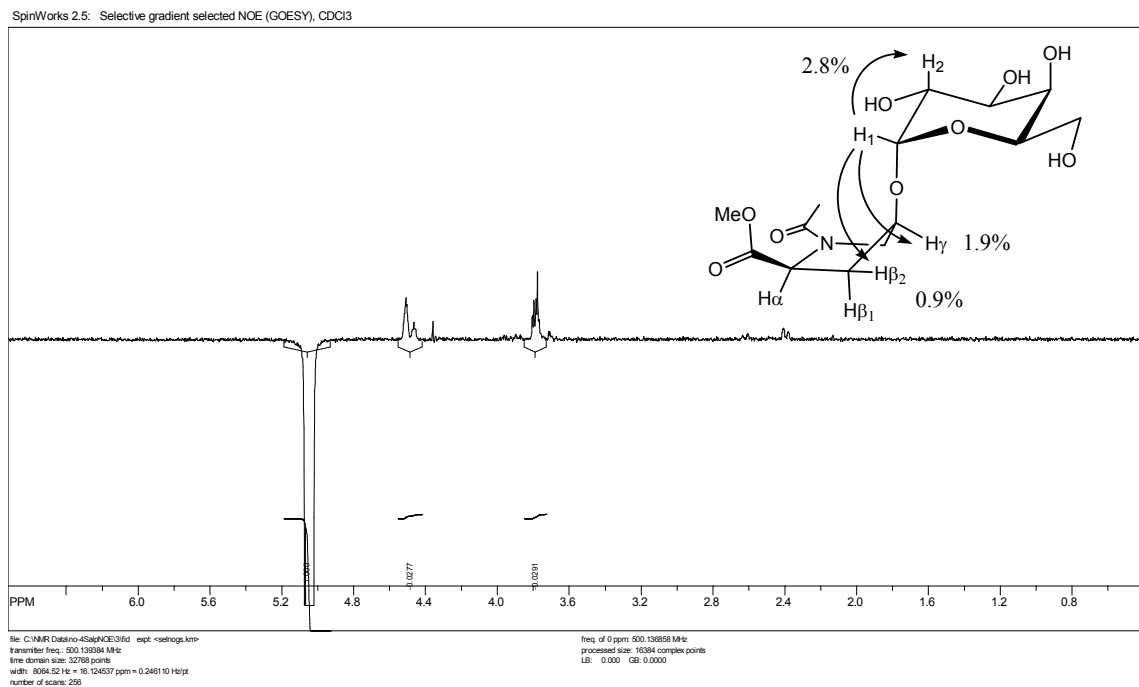
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating H₂^{cis,trans} (3.52 ppm, 1H) in **3** showed inter-proton effect to H₁^{trans} (4.49 ppm, 0.8H) (0.5%), H₄^{trans} (3.96 ppm, 0.8H) (0.2%), and the H₃^{cis,trans}/H₅^{cis,trans} peak (3.69 ppm, 2H) (2.7%).



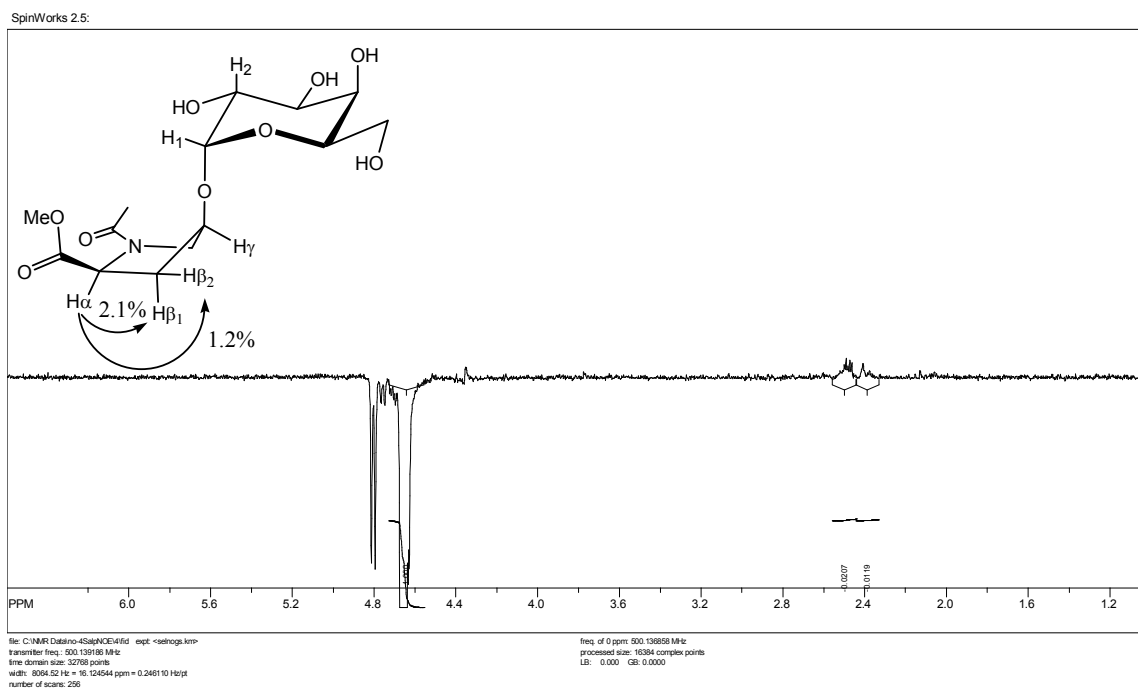
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating Pro_{β₁}^{trans} (2.55 ppm, 0.8H) in **3** showed inter-proton effect to Pro_γ^{trans} (4.69 ppm, 0.8H) (0.8%), Pro_α^{trans} (4.54 ppm, 0.8H) (2.1%), H₁^{trans} (4.49 ppm, 0.8H) (1.5%), and Pro_{β₂}^{trans} (2.21 ppm, 0.8H) (10.9%). Inversion exchange to Pro_{β₁}^{cis} (2.64 ppm, 0.2H) resulted in inter-proton effect to Pro_{β₂}^{cis} (2.41 ppm, 0.2H) (1.5%) relative to Pro_{β₁}^{cis}.



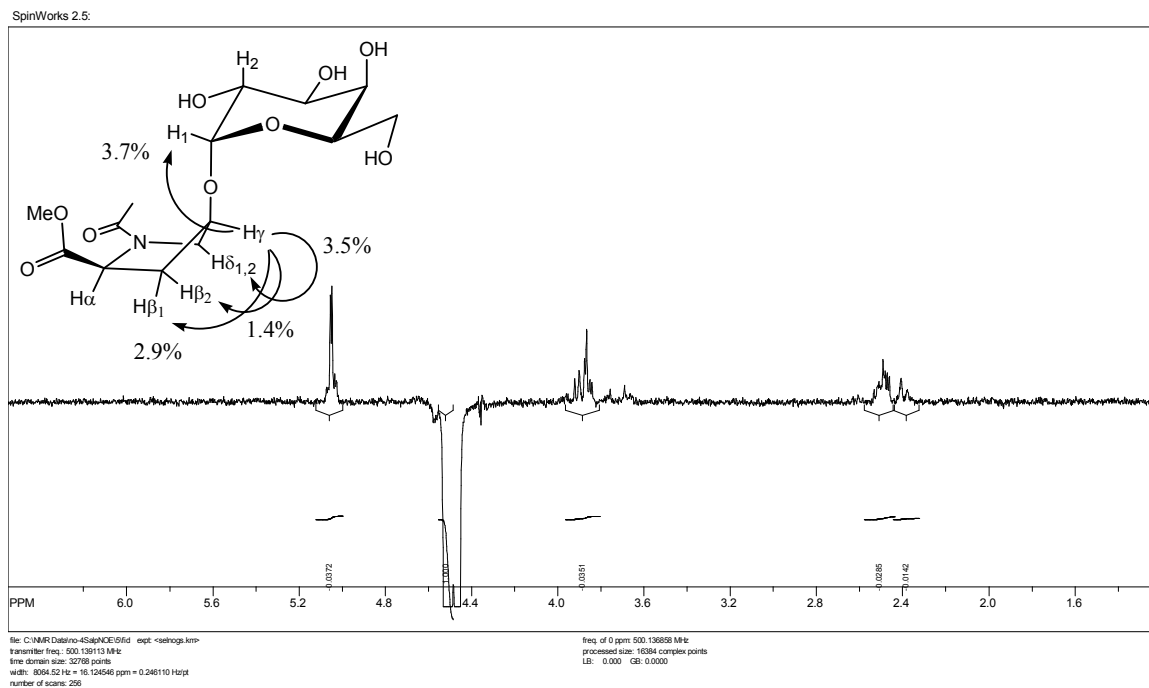
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating H₁^{trans} / H₁^{cis} (5.04 ppm, 0.75H) in **5** showed inter-proton effect to Pro_γ^{trans} (4.51 ppm, 0.75H) (1.9%), Pro_γ^{cis} (4.48 ppm, 0.25H) (0.8%), H₂^{trans} (3.79 ppm, 0.75H) (2.8%), and Pro_{β2}^{trans} (2.41 ppm, 0.75H) (0.9%).



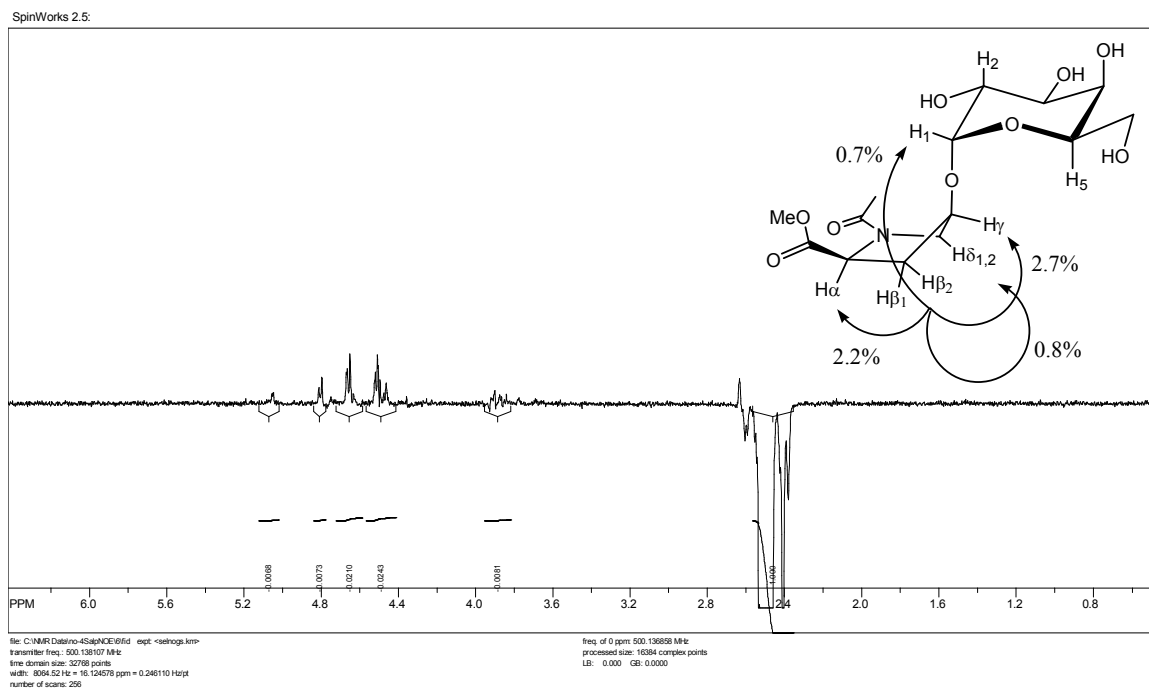
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating Pro_α^{trans} (4.66 ppm, 0.75H) in **5** showed inter-proton effect to Pro_{β1}^{trans} (2.48 ppm, 0.75H) (2.1%) and Pro_{β2}^{trans} (2.41 ppm, 0.75H) (1.2%).



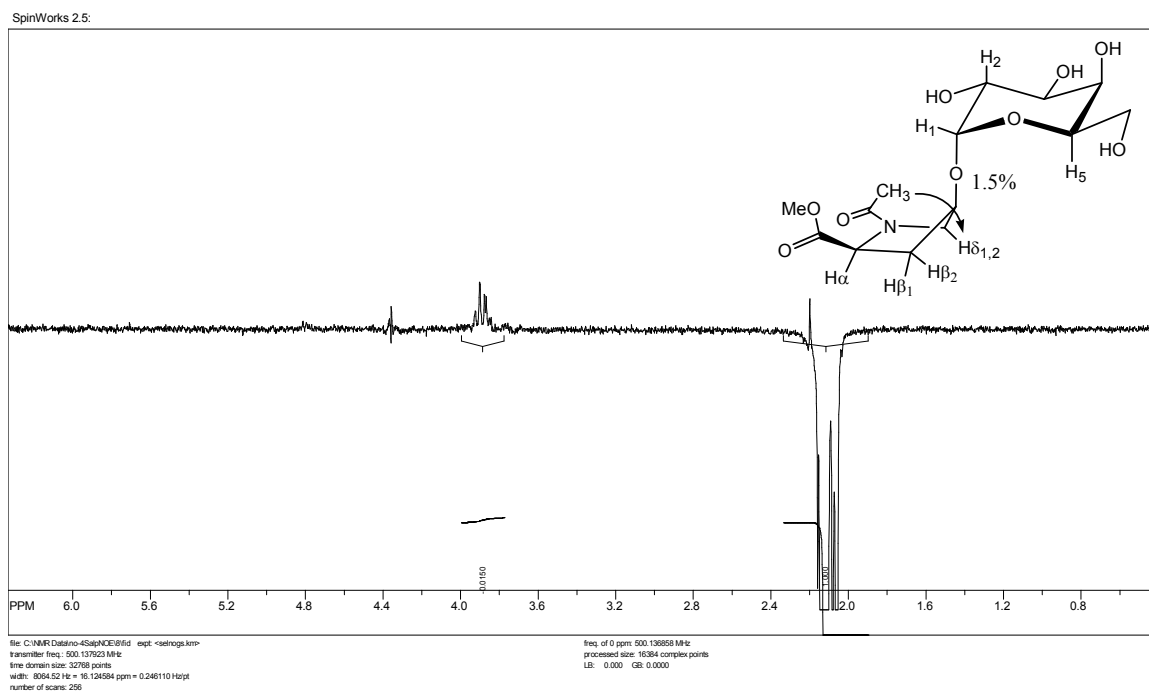
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating Pro_γ^{trans} (4.51 ppm, 0.75H) in **5** showed inter-proton effect to H₁^{trans} (5.04 ppm, 0.75H) (3.7%), Pro_{δ1}^{trans} / Pro_{δ2}^{trans} (3.88 ppm, 1.5H) (3.5%), Pro_{β1}^{trans} (2.48 ppm, 0.75H) (2.9%), and Pro_{β2}^{trans} (2.41 ppm, 0.75H) (1.4%).



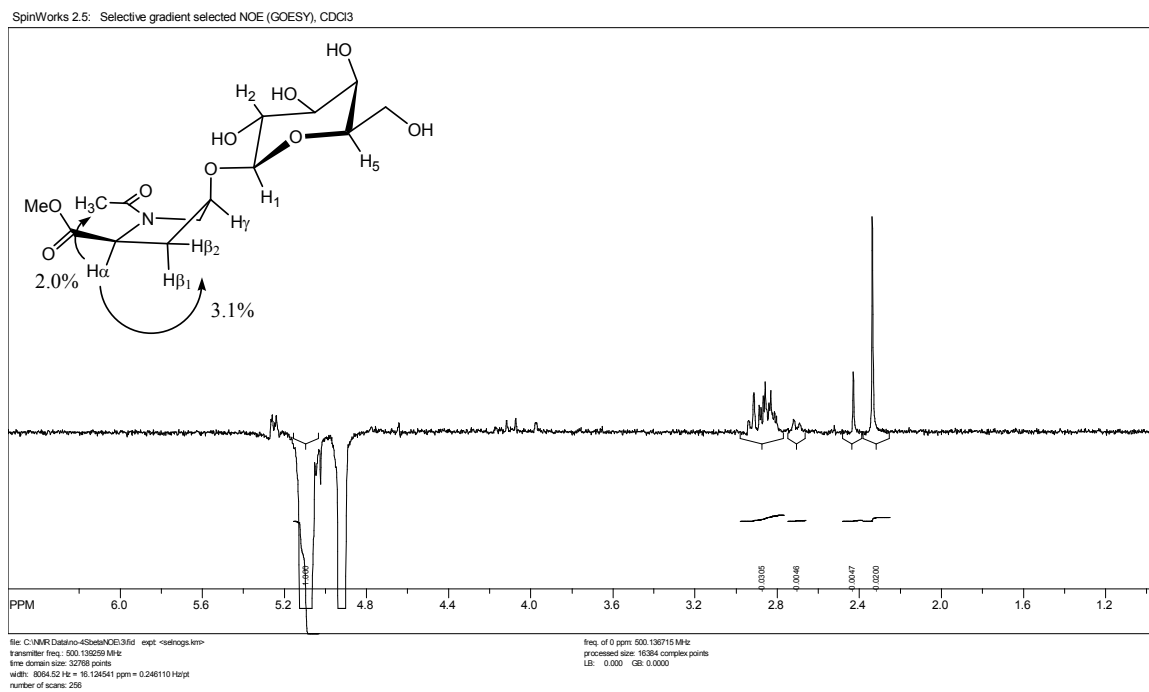
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating the Pro_{β1}^{trans} / Pro_{β2}^{trans} peaks (2.41-2.48 ppm, 1.5H) in **5** showed inter-proton effect to H₁^{trans} (5.04 ppm, 0.75H) (0.7%), Pro_α^{trans} (4.66 ppm, 0.75H) (2.2%), Pro_γ^{trans} (4.50 ppm, 0.75H) (2.7%), and Pro_{δ1}^{trans} / Pro_{δ2}^{trans} (3.88 ppm, 1.5H) (0.8%).



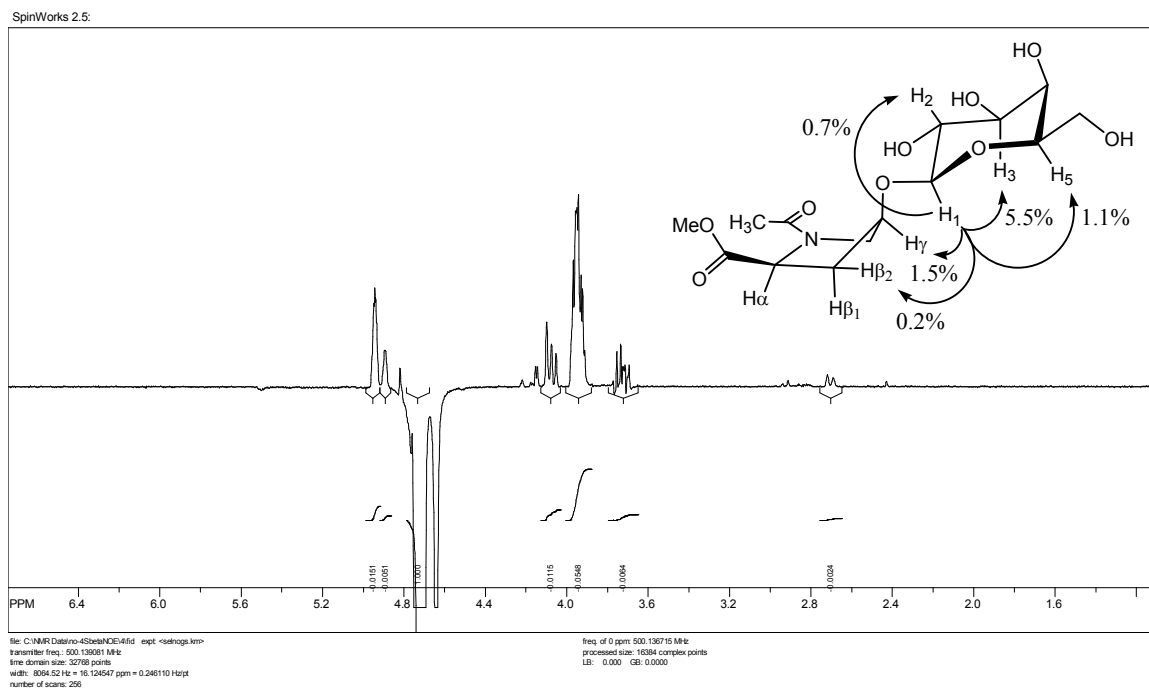
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating $-\text{NCOCH}_3^{\text{trans}}$ (2.13 ppm, 2.25H) in **5** showed inter-proton effect to $\text{Pro}_{\delta 1}^{\text{trans}} / \text{Pro}_{\delta 2}^{\text{trans}}$ (3.88 ppm, 1.5H) (1.5%).



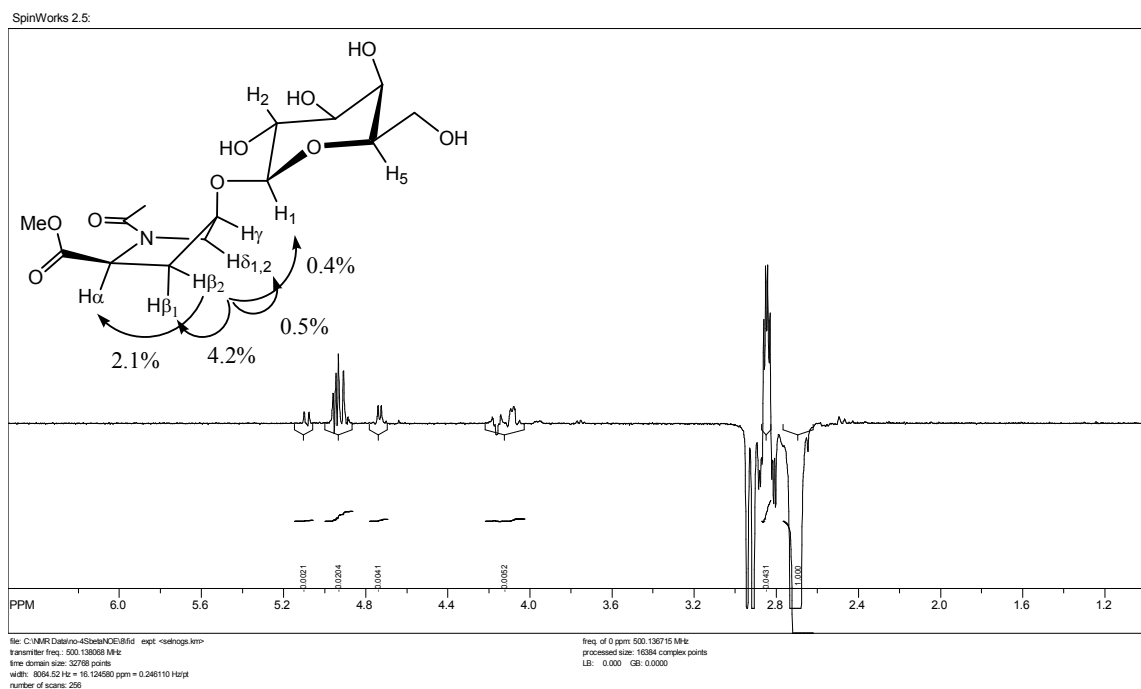
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating the Pro_α^{cis} (5.09 ppm, 0.3H) in **6** showed inter-proton effect to Pro_{β1}^{cis}/ Pro_{β2}^{cis} (2.87 ppm, 0.6H) (3.1%) and -NCOCH₃^{cis} (2.34 ppm, 0.9H) (2.0%). Inversion exchange to Pro_α^{trans} (4.93 ppm, 0.7H) resulted in inter-proton effect to Pro_{β2}^{trans} (2.71 ppm, 0.7H) (0.5%) and -NCOCH₃^{trans} (2.43 ppm, 2.1H) (0.5%).



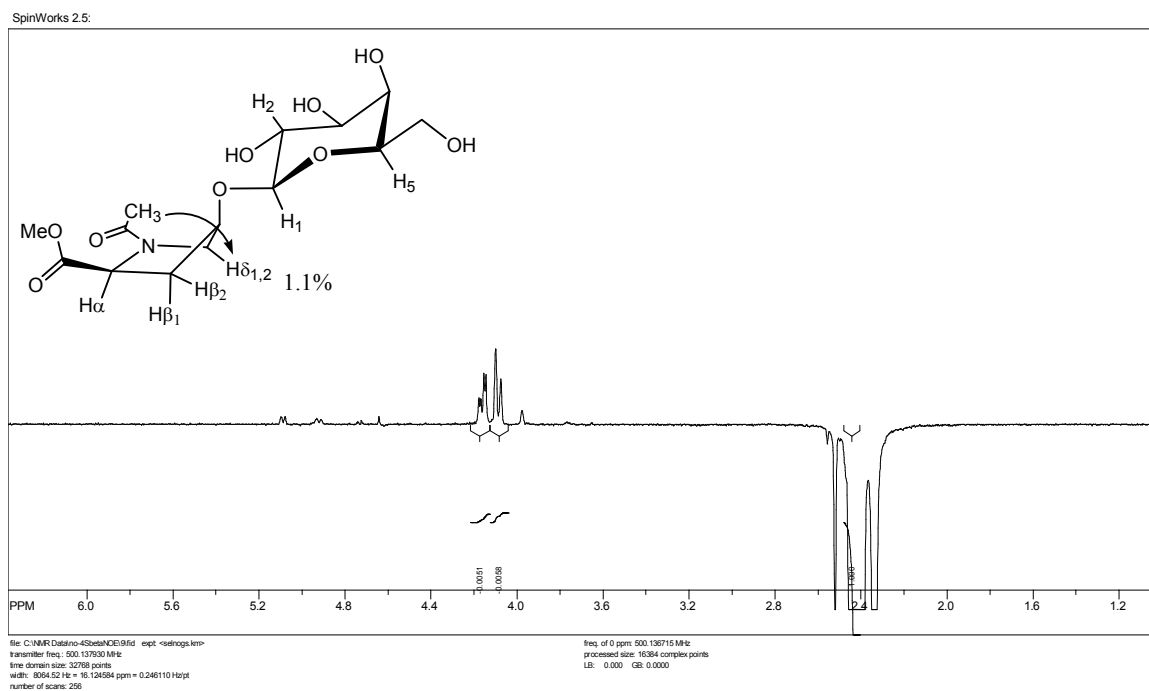
A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating H₁^{trans} / H₁^{cis} (4.74 ppm, 1H) in **6** showed inter-proton effect to Pro_γ^{trans} (4.95 ppm, 0.7H) (1.5%), Pro_γ^{cis} (4.89 ppm, 0.3H) (0.5%), H₅^{trans} / H₅^{cis} (4.08 ppm, 1H) (1.1%), H₃^{trans} / H₃^{cis} (3.88-4.00 ppm, 1H) (5.5%), H₂^{trans} / H₂^{cis} (3.68-3.78 ppm, 1H) (0.7%), and Pro_{β2}^{trans} (2.71 ppm, 0.7H) (0.2%).



A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating Pro β_2^{trans} (2.71 ppm, 0.7H) in **6** showed inter-proton effect to Pro α^{trans} (4.93 ppm, 0.7H) (2.1%), H $_1^{trans}$ (4.72 ppm, 0.7H) (0.4%), Pro δ_2^{trans} (4.09 ppm, 0.7H) (0.5%), and Pro β_1^{trans} (2.87 ppm, 0.7H) (4.2%). Inversion exchange to Pro β_2^{cis} (2.93 ppm, 0.3H) resulted in inter-proton effect to Pro α^{cis} (5.09 ppm, 0.3H) (0.2%).



A one-dimensional GOESY experiment at 340.4 K in D₂O irradiating $-\text{NCOCH}_3^{\text{trans}}$ (2.43 ppm, 2.1H) in **6** showed inter-proton effect to $\text{Pro}_{\delta 1}^{\text{trans}}$ (4.17 ppm, 0.7H) (0.5%) and $\text{Pro}_{\delta 2}^{\text{trans}}$ (4.09 ppm, 0.7H) (0.6%).



Magnetization Transfer NMR Experiments

Samples of **1-6** were prepared by dissolving compounds in D₂O at concentrations between 0.05 M and 0.1 M. All experiments were performed on a Bruker AMX500 NMR spectrometer equipped with a triple resonance (¹H, ¹³C, ¹⁵N) gradient inverse probehead.

Selective inversion of the *N'*-methylamide singlet of the *trans* isomer was done using an 80 msec pulse centered on resonance. Relaxation delays of 20 s and inversion recovery delays between 1 ms and 15 s were used. Experiments were performed at 340.44 K for each compound. The temperature was calibrated using an ethylene glycol standard. For each compound, inversion-recovery data were collected with selective inversion of the ¹H *N'*-methylamide singlet of the major *trans* peak. ¹H NMR spectroscopy was used in preference to ¹³C NMR since over the course of an experiment, the signal/noise ratio is much higher for ¹H than ¹³C, and heating effects caused by decoupling for ¹³C causes uncertainty in the temperature of the sample.

The time-dependent magnetization transfers of the *cis* (M_c(t)) and *trans* (M_t(t)) NMR signals as a function of the inversion transfer time (t) were simultaneously fit to equations 1 and 2 (Alger and Prestegard, 1977; Mariappan and Rabenstein, 1992) below for compounds **1-6** using *Mathematica* (v. 5.0). In the following pulse sequence, the ¹H *trans* resonance is selectively inverted using a shaped pulse. Its recovery during t is determined by its intrinsic T_{1t}, magnetization transfer to and from the *cis* resonance, and the T_{1c} of the *cis* resonance:

$\pi(x)$ sel-----t----- $\pi/2(x,y.-x.-y)$ ---acquire

The resonances of the *trans* and *cis* isomers show the following time dependences:

$$M_i(t) = (c_1)(\tau_t)(\lambda_1 + 1/\tau_{1c})\exp(\lambda_1 * t) + (c_2)(\tau_c)(\lambda_2 + 1/\tau_{1c})\exp(\lambda_2 * t) + M_{c\infty} \text{-----} 1$$

$$M_i(t) = (c_1)\exp(\lambda_1 * t) + (c_2)\exp(\lambda_2 * t) + M_{t\infty} \text{-----} 2$$

T_{1c} and T_{1t} are the longitudinal relaxation times of the resonances in the absence of exchange.

τ_c and τ_t are the lifetimes of the *cis* and *trans* conformers and k_{ct} and k_{tc} are the corresponding rate constants.

τ_{1c} and τ_{1t} are the effective relaxation times of the *cis* and *trans* resonances when relaxation and exchange are both occurring and are defined below in terms of T_{1c} and τ_c , T_{1t} and τ_t .

λ_1 and λ_2 are related to the time constants τ_c , τ_t , τ_{1c} , and τ_{1t} , and are defined below.

c_1 and c_2 are defined below.

$M_{c\infty}$ and $M_{t\infty}$ are determined experimentally from the magnetization measured after $5 T_1$ periods for the *cis* and *trans* resonances, respectively.

Mathematica then calculates τ_t from τ_c , $M_{c\infty}$, and $M_{t\infty}$ as: $\tau_t = \tau_c * M_{t\infty}/M_{c\infty}$

Thus,

$$k_{ct} = 1/\tau_c$$

$$k_{tc} = 1/\tau_t$$

$$K_{eq} = M_{t\infty}/M_{c\infty}$$

$$\tau_{1c} = (T_{1c} * \tau_c) / (\tau_c + T_{1c})$$

$$\tau_{1t} = (T_{1t} * \tau_t) / (\tau_t + T_{1t})$$

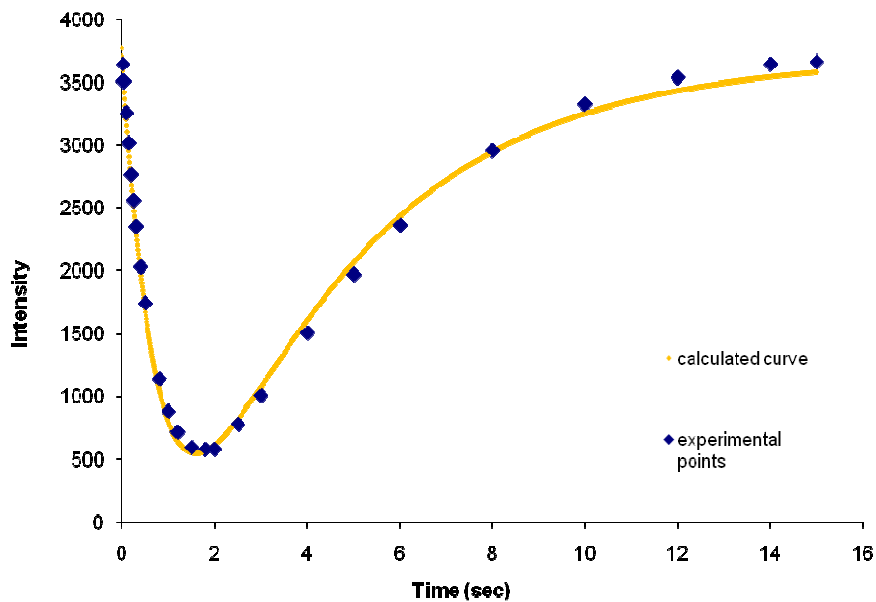
$$\lambda_1 = (0.5) * (-((1/\tau_{1c}) + (1/\tau_{1t})) + (((((1/\tau_{1c}) + (1/\tau_{1t}))1/2) - (4) * (((1/\tau_{1c}) * (1/\tau_{1t})) - ((1/\tau_c) * (1/\tau_t))))))1/2)$$

$$\lambda_2 = (0.5) * (-((1/\tau_{1c}) + (1/\tau_{1t})) - (((((1/\tau_{1c}) + (1/\tau_{1t}))2) - (4) * (((1/\tau_{1c}) * (1/\tau_{1t})) - ((1/\tau_c) * (1/\tau_t))))))1/2)$$

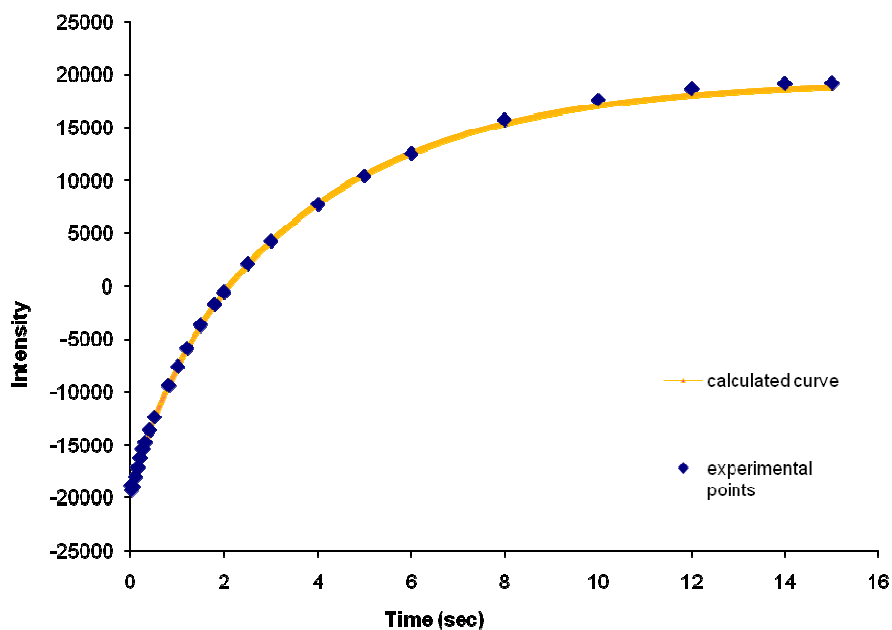
$$c_2 = ((1/(\tau_t)) * (\lambda_1 - \lambda_2)) * ((\tau_t * (\lambda_1 + (1/\tau_{1c}))) * ((M_{0c} - M_{c\infty}) + (M_{t\infty} - M_{0t})))$$

$$c_1 = M_{0c} - M_{c\infty} - c_2$$

Figure 11.7: Inversion recovery of non-inverted (*cis*) *N*-amide singlet (A) and inverted (*trans*) *N*-amide singlet (B) of **1** at 67.3 °C in 0.1M phosphate buffer pH = 7.2



A



B

Figure 11.8: Inversion recovery of non-inverted (*cis*) *N*-amide singlet (A) and inverted (*trans*) *N*-amide singlet (B) of **2** at 67.3 °C in 0.1M phosphate buffer pH = 7.2

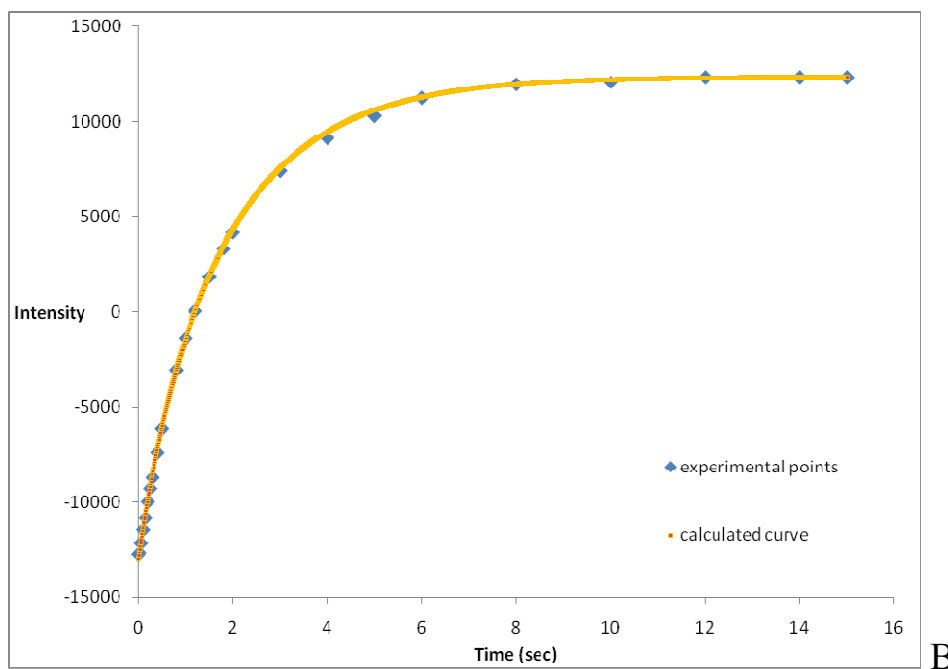
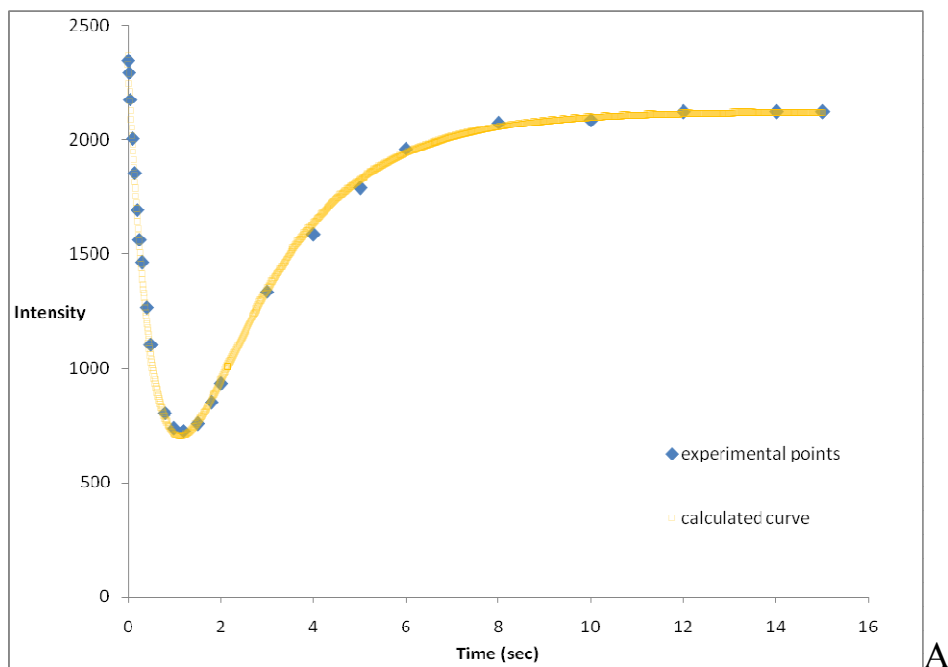


Figure 11.9: Inversion recovery of non-inverted (*cis*) *N*-amide singlet (A) and inverted (*trans*) *N*-amide singlet (B) of **3** at 67.3 °C in 0.1M phosphate buffer pH = 7.2

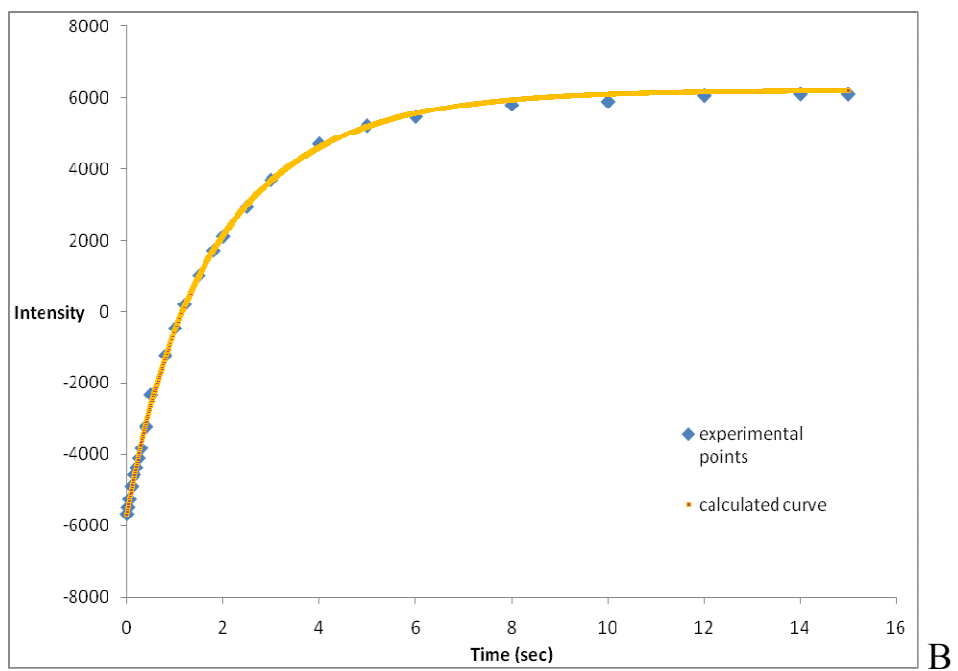
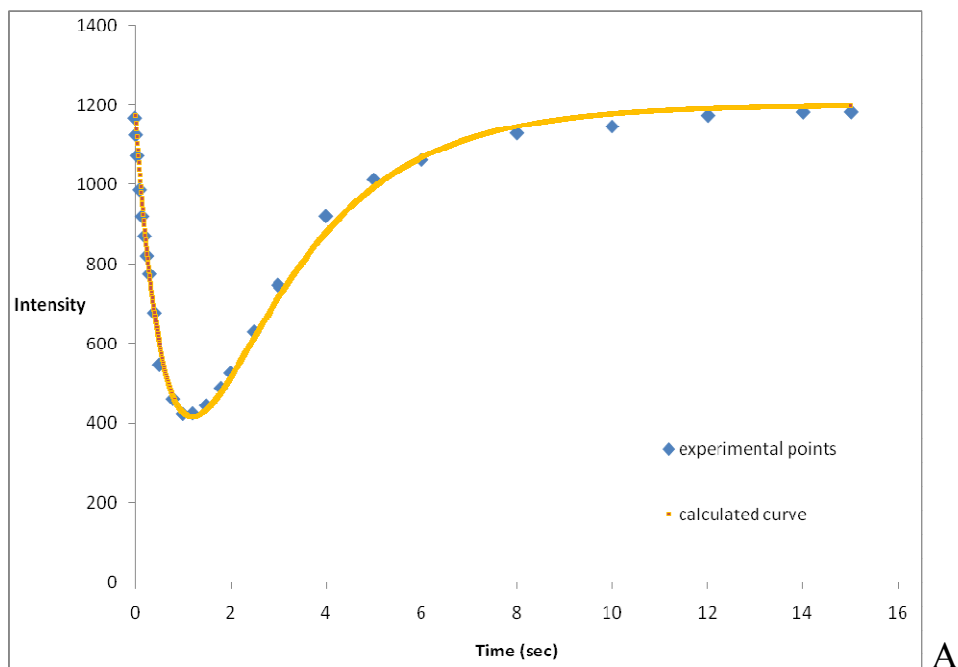


Figure 11.10: Inversion recovery of non-inverted (*cis*) *N*-amide singlet (A) and inverted (*trans*) *N*-amide singlet (B) of **4** at 67.3 °C in 0.1M phosphate buffer pH = 7.2

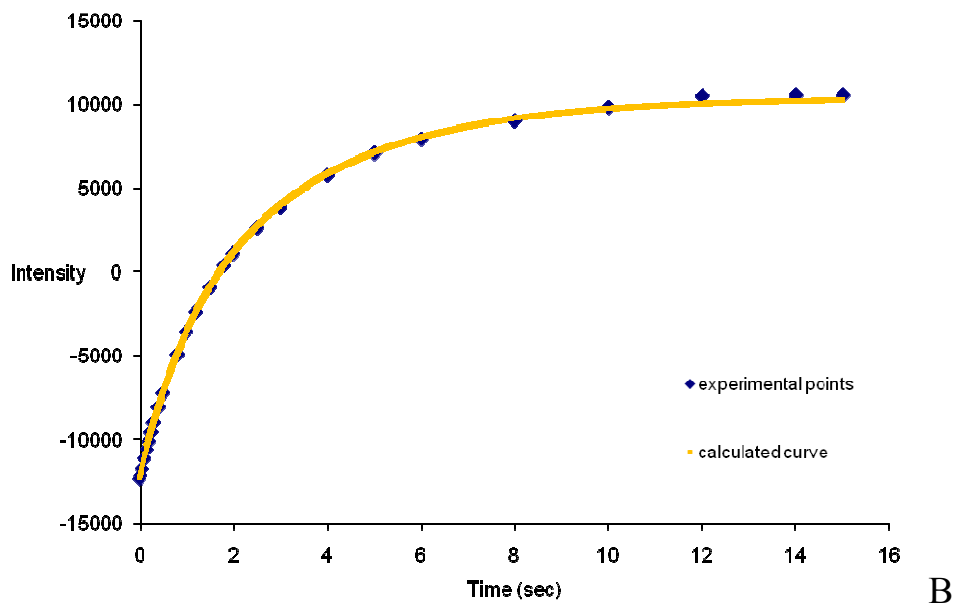
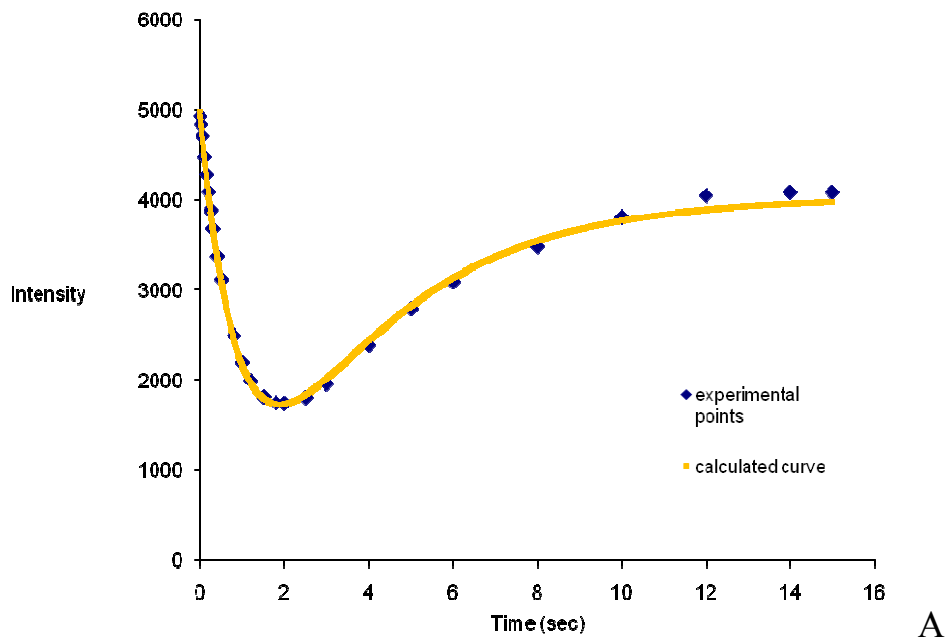
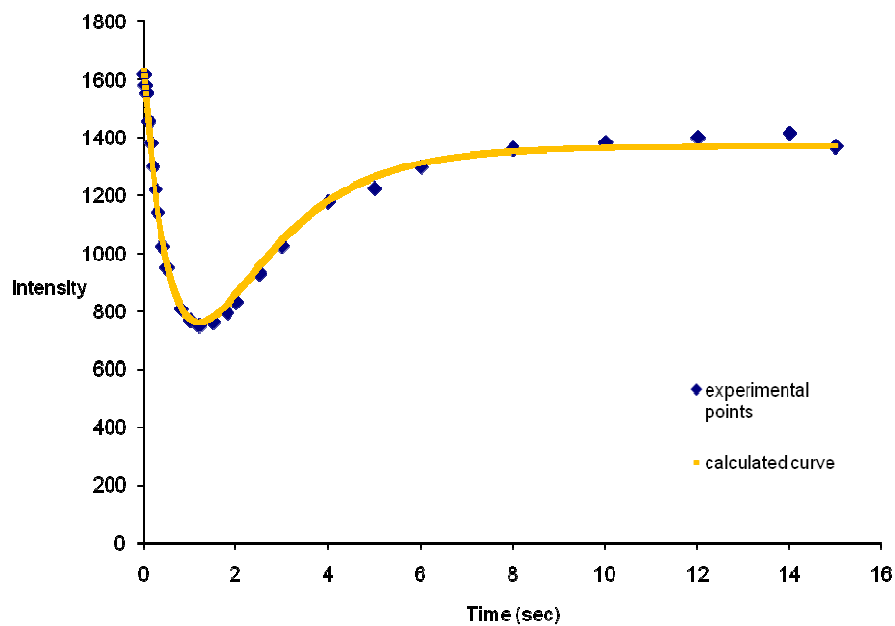
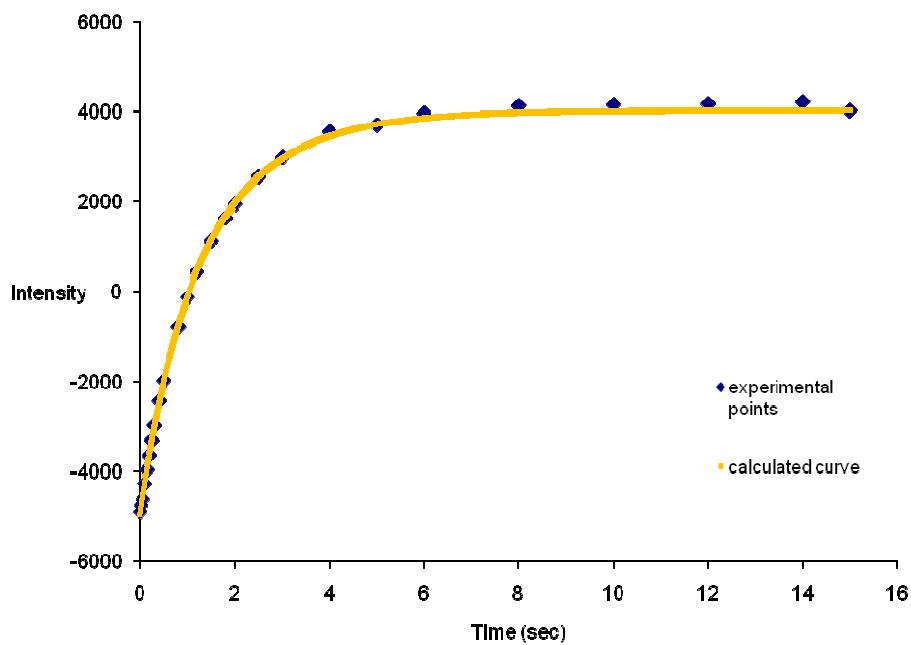


Figure 11.11: Inversion recovery of non-inverted (*cis*) *N*-amide singlet (A) and inverted (*trans*) *N*-amide singlet (B) of **5** at 67.3 °C in 0.1M phosphate buffer pH = 7.2

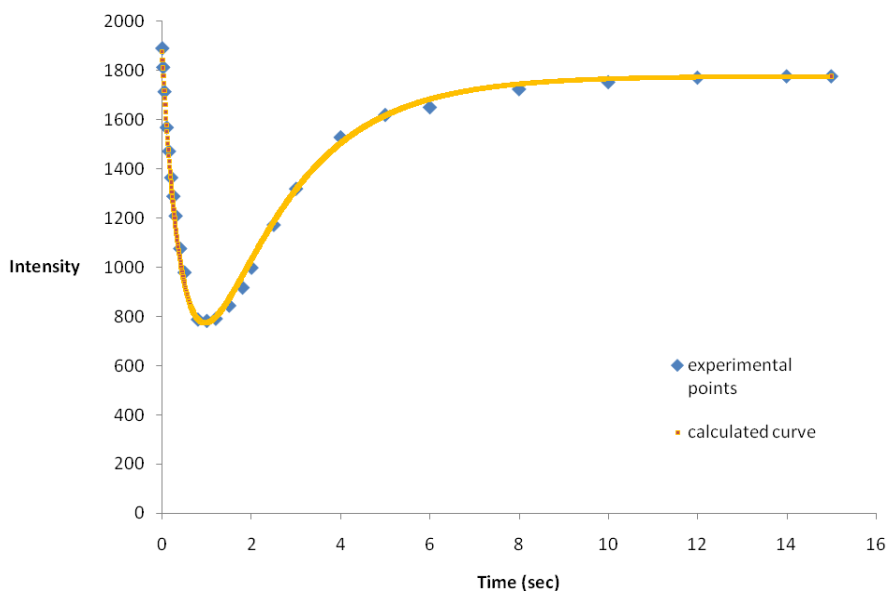


A

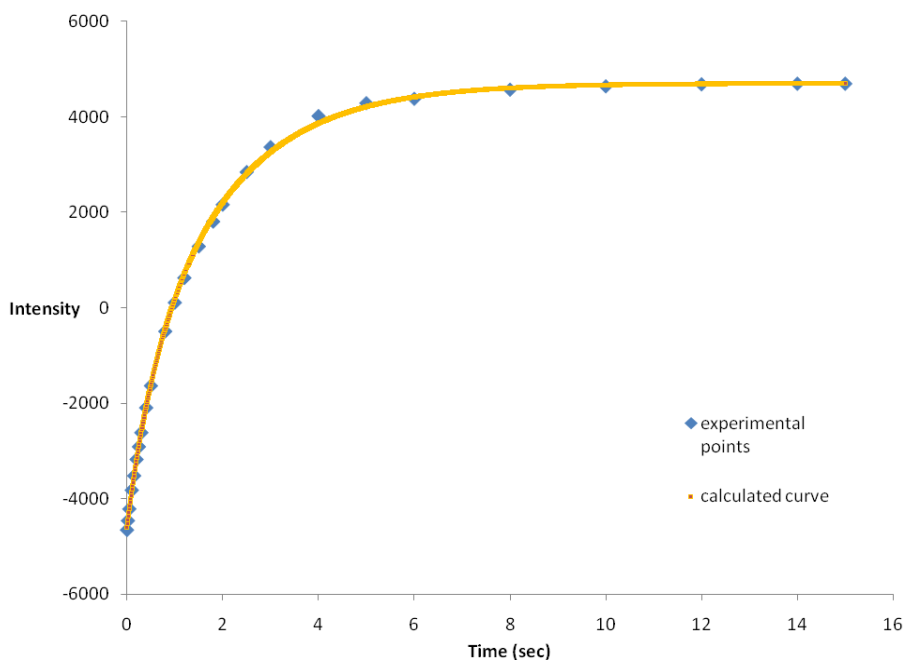


B

Figure 11.12: Inversion recovery of non-inverted (*cis*) *N*-amide singlet (A) and inverted (*trans*) *N*-amide singlet (B) of **6** at 67.3 °C in 0.1M phosphate buffer pH = 7.2



A



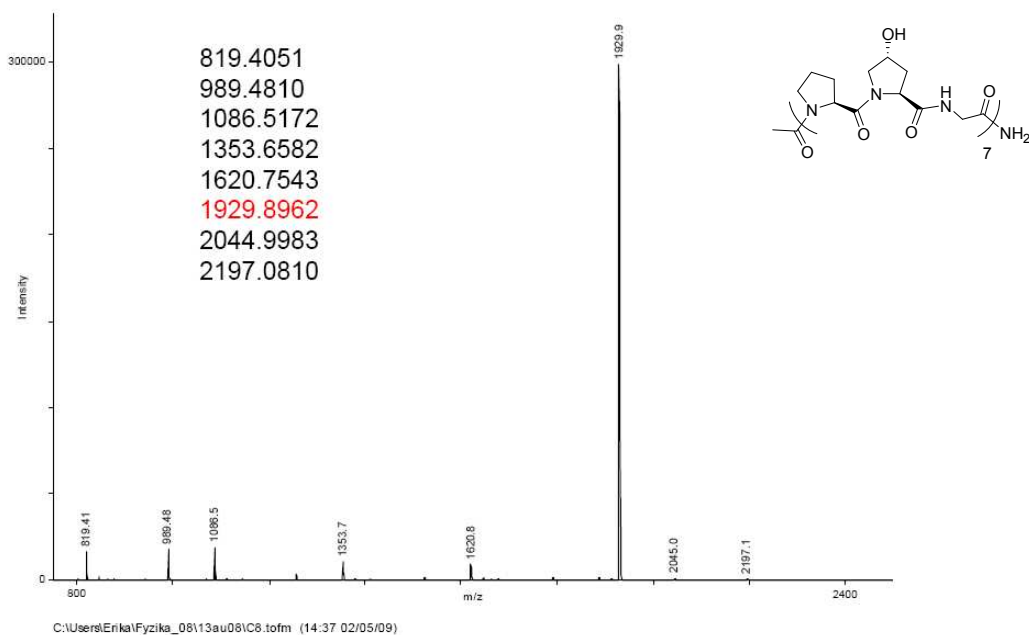
B

**Chapter 12: The Effects of Glycosylation of (2*S*,4*R*)-4-Hydroxyproline
on the Stability of Collagen Model Peptides**

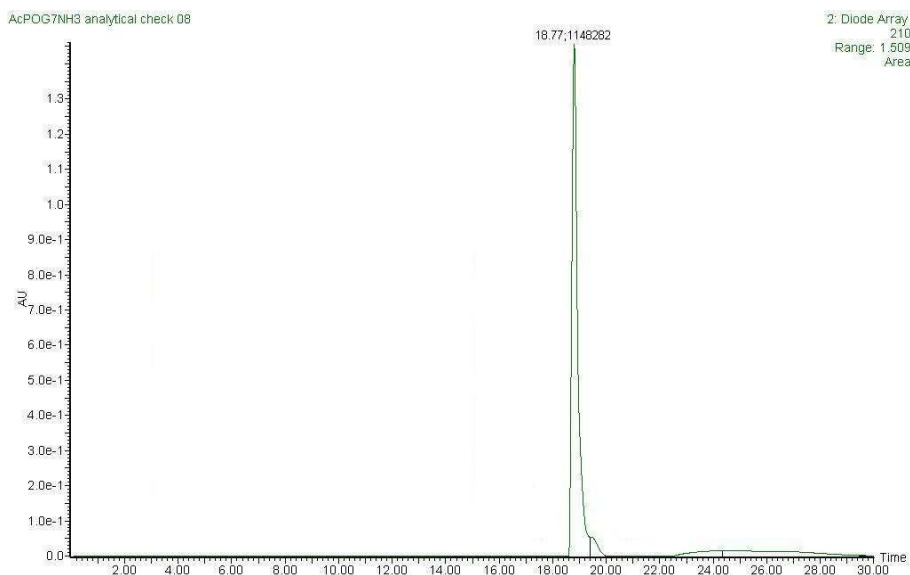
Supporting Information For Chapter 6

Table of Contents

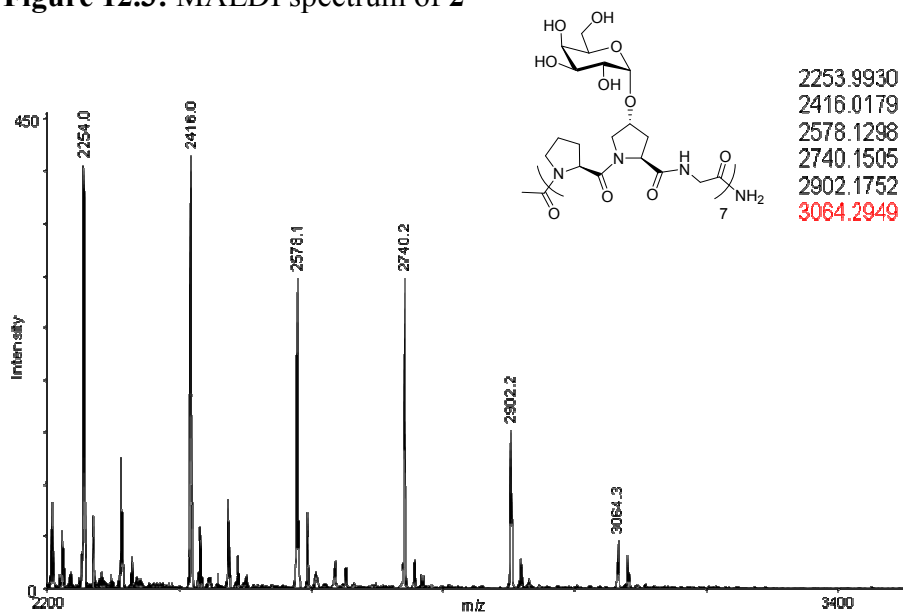
MALDI and Analytical Chromatograms for 1-3	481-483
Table of Ellipticity as a Function of Temperature for 1-3	484
Melting Curves and CD Spectra for 1-3	485-488
NMR Spectra	489-506

Figure 12.1: MALDI spectrum of **1**

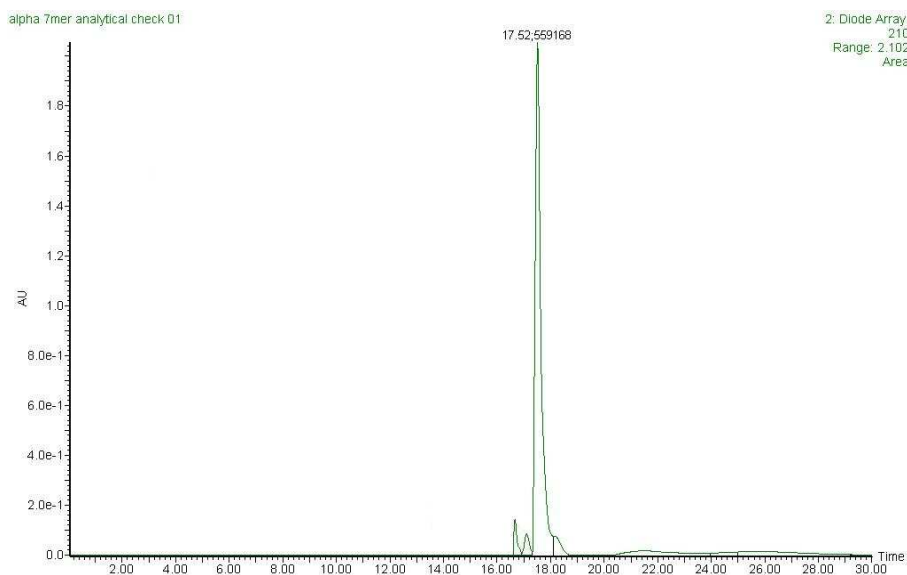
MALDI: Calc. for $C_{86}H_{125}N_{22}O_{29}$: 1929.8983 (M+H)⁺. Found (M+H)⁺: 1929.8962.

Figure 12.2: Analytical chromatogram of **1**

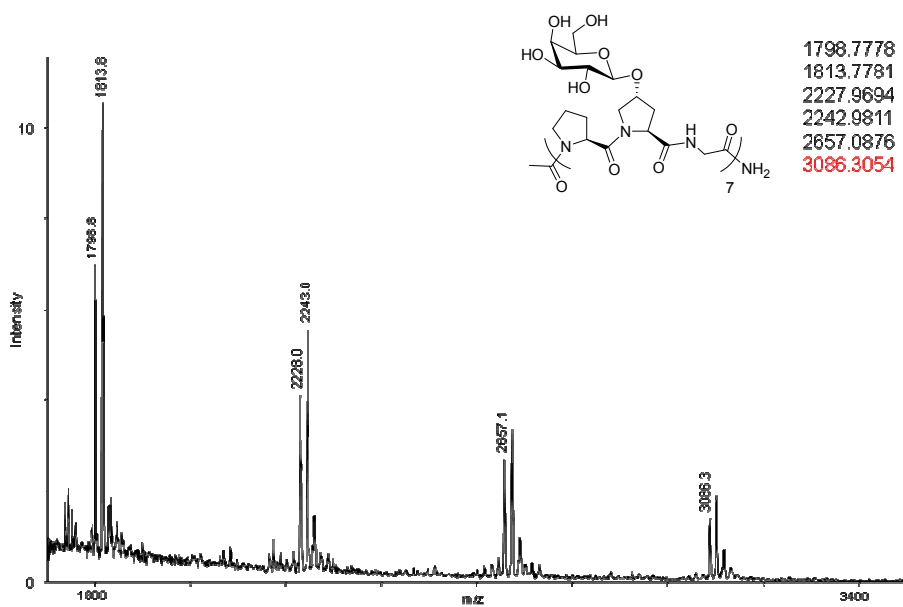
Conditions: 0 to 40% MeCN in H_2O over 30 min, at a flow rate of $1 \text{ mL} \cdot \text{min}^{-1}$.

Figure 12.3: MALDI spectrum of **2**

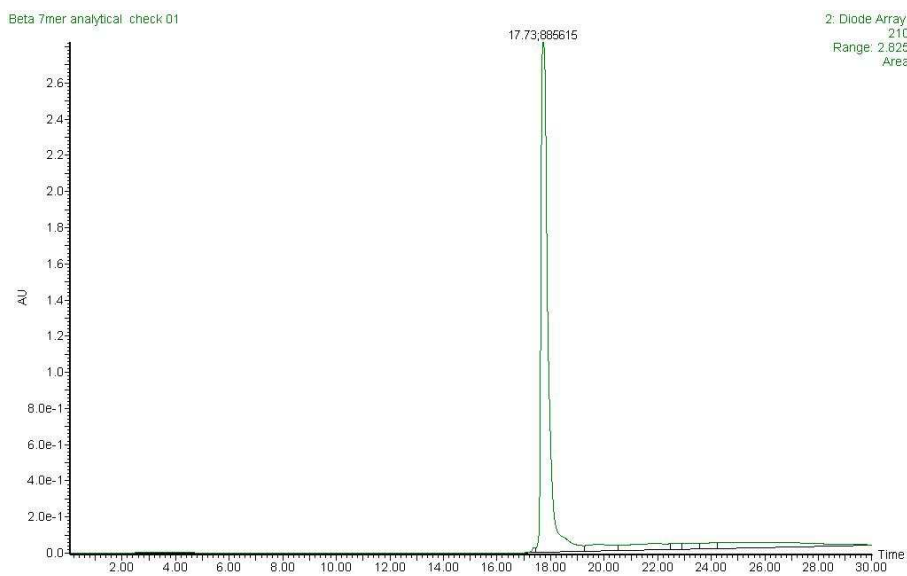
MALDI: Calc. for $C_{128}H_{195}N_{22}O_{64}$: 3064.2681 ($M+H$)⁺. Found ($M+H$)⁺: 3064.2949.

Figure 12.4: Analytical chromatogram of **2**

Conditions: 0 to 40% MeCN in H₂O over 30 min, at a flow rate of 1 mL·min⁻¹.

Figure 12.5: MALDI spectrum of **3**

MALDI: Calc. for $C_{128}H_{194}N_{22}NaO_{64}$: 3086.2500 ($M+Na$)⁺. Found ($M+Na$)⁺: 3086.3054.

Figure 12.6: Analytical chromatogram of **3**

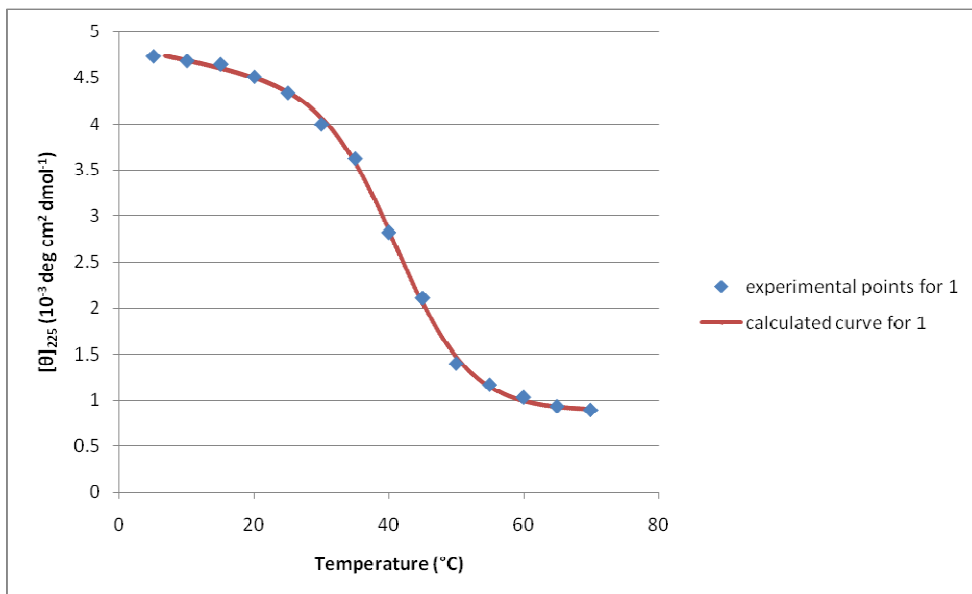
Conditions: 0 to 40% ACN in H_2O over 30 min, at a flow rate of $1\text{ mL}\cdot\text{min}^{-1}$.

Table 13.1: Ellipticity (millidegrees) at 225 nm as a function of temperature for **1-3**^[a]

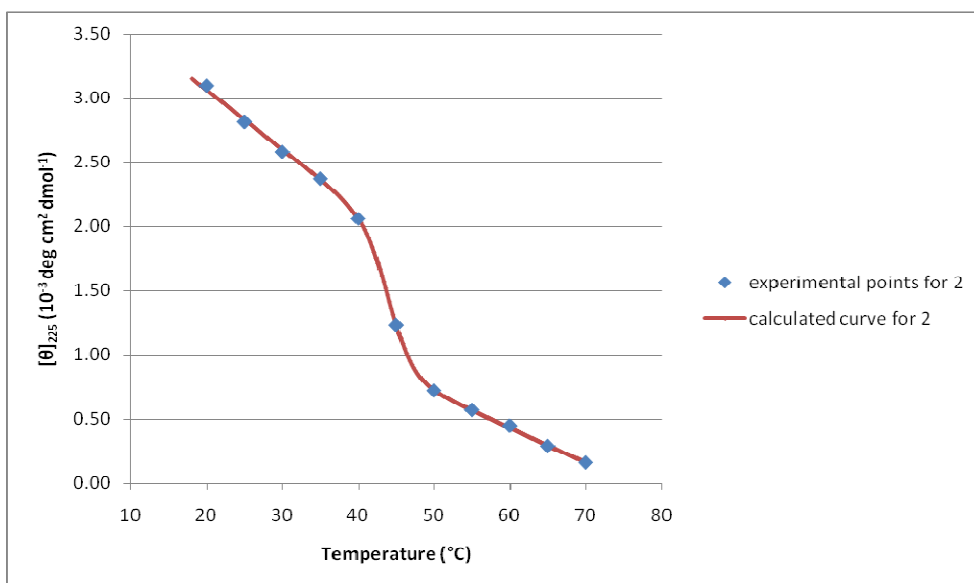
Temperature (±1 °C)	Compound		
	1	2	3
5	26.98 ± 0.05	26.63 ± 0.12	30.50 ± 0.06
10	26.69 ± 0.05	25.60 ± 0.22	29.30 ± 0.08
15	26.48 ± 0.19	23.75 ± 0.04	28.38 ± 0.13
20	25.70 ± 0.11	22.23 ± 0.05	26.75 ± 0.34
25	24.69 ± 0.04	20.19 ± 0.14	25.41 ± 0.28
30	22.78 ± 0.25	18.53 ± 0.25	24.20 ± 0.17
35	20.64 ± 0.16	17.01 ± 0.21	21.68 ± 0.04
40	16.07 ± 0.09	14.78 ± 0.12	16.92 ± 0.25
45	12.00 ± 0.10	8.86 ± 0.09	11.79 ± 0.14
50	7.96 ± 0.11	5.19 ± 0.17	9.36 ± 0.23
55	6.64 ± 0.18	4.08 ± 0.10	8.15 ± 0.21
60	5.86 ± 0.20	3.21 ± 0.05	6.75 ± 0.12
65	5.30 ± 0.05	2.09 ± 0.02	4.90 ± 0.03
70	5.07 ± 0.14	1.14 ± 0.06	3.78 ± 0.10
25 ^[b]	22.47 ± 0.14	17.25 ± 0.36	23.11 ± 0.15

^[a]±Standard error from two or more measurements; ^[b]Post

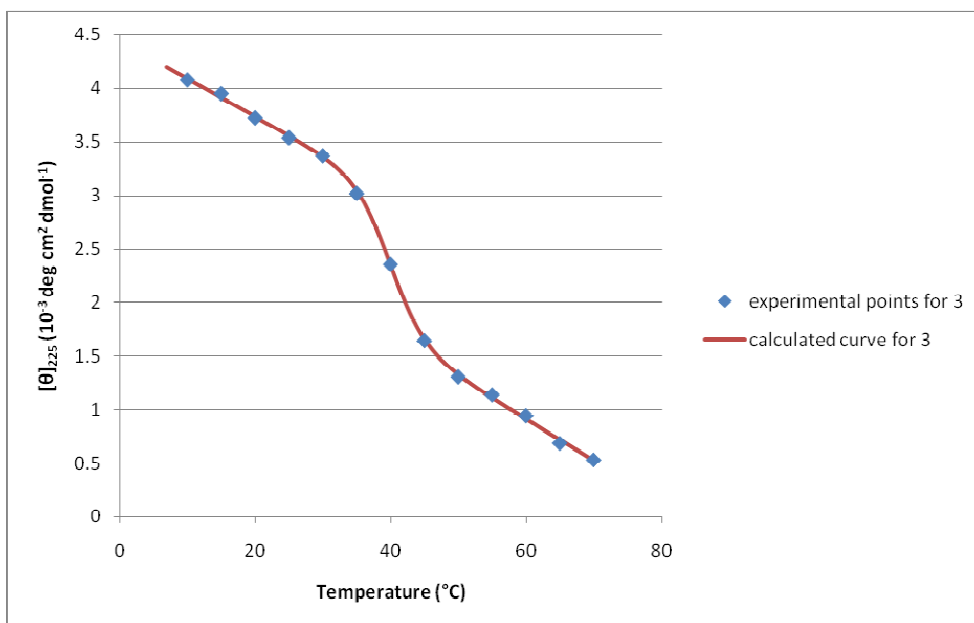
thermal melt experiment after 30 minutes.

Figure 12.7: Melting curve of **1** monitored at 225 nm from 5 to 70 °C in water

T_m calculated to be 41.9 °C as described in section 6.6.

Figure 12.8: Melting curve of **2** monitored at 225 nm from 15 to 70 °C in water

T_m calculated to be 44.1 °C as described in section 6.6.

Figure 12.9: Melting curve of **3** monitored at 225 nm from 10 to 70 °C in water

T_m calculated to be 39.9 °C as described in section 6.6.

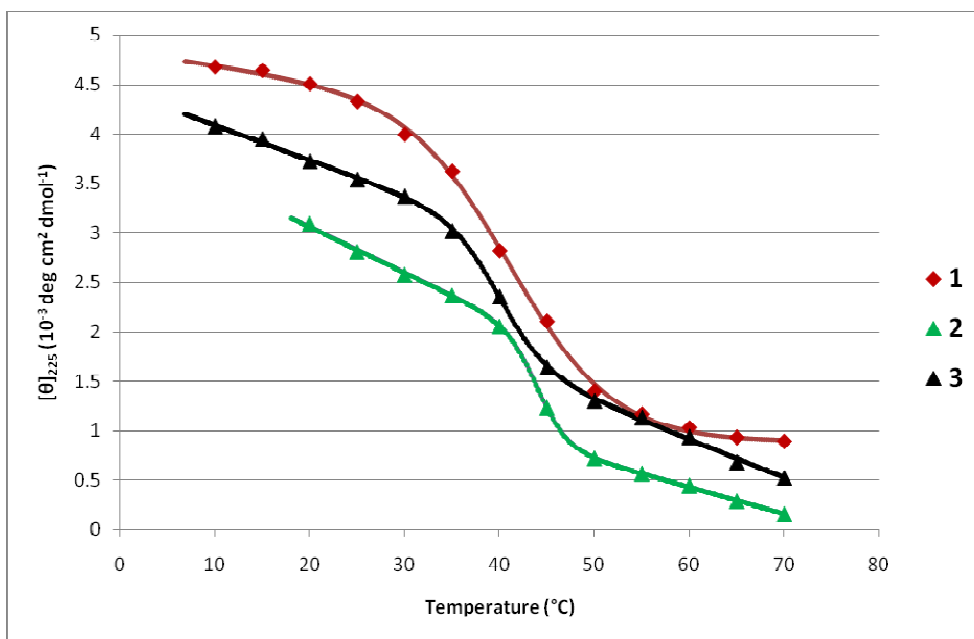
Figure 12.10: Comparison of Melting Curves of **1-3** in water

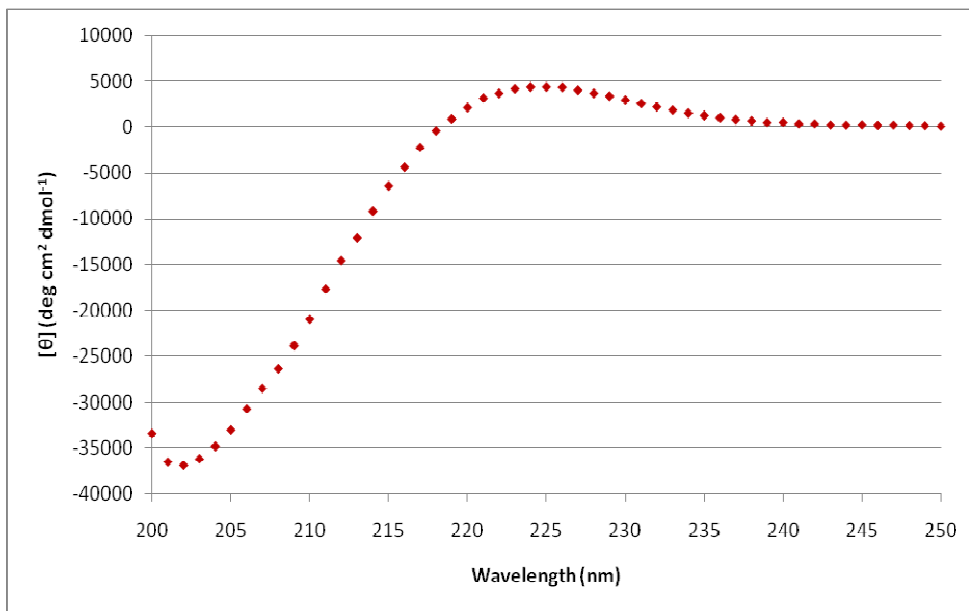
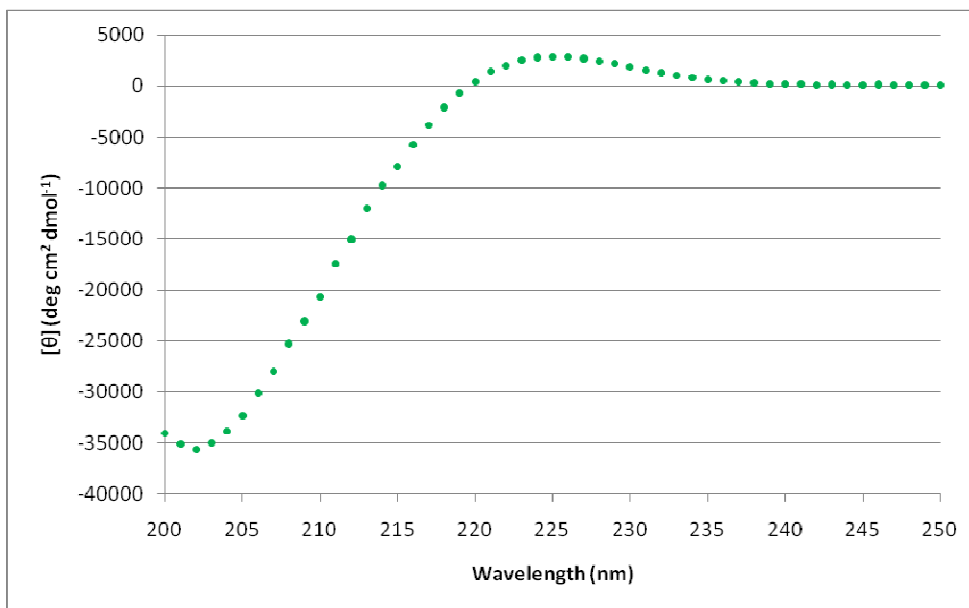
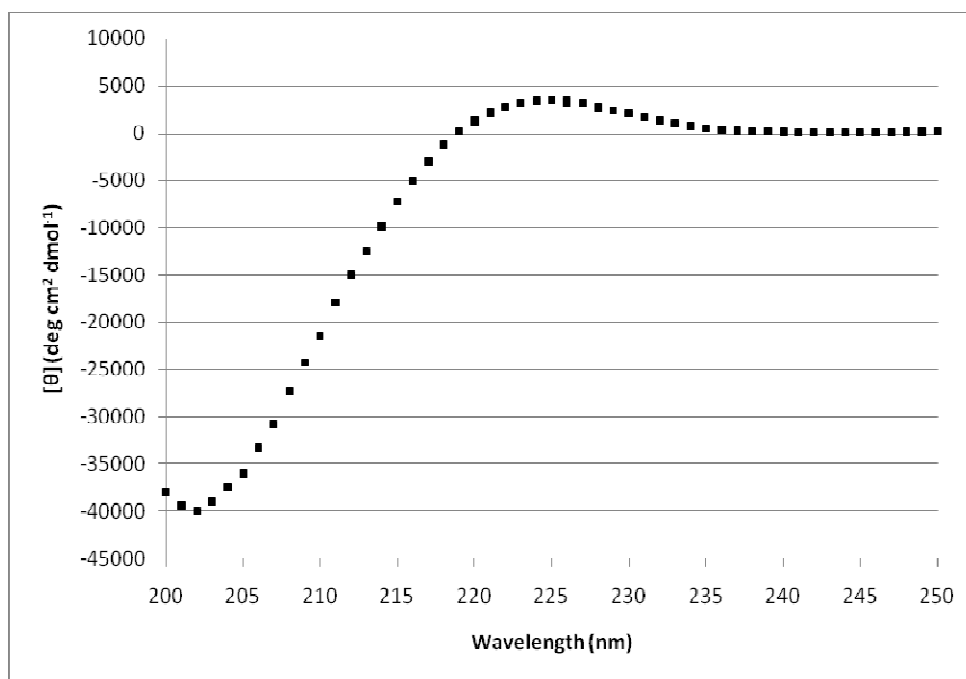
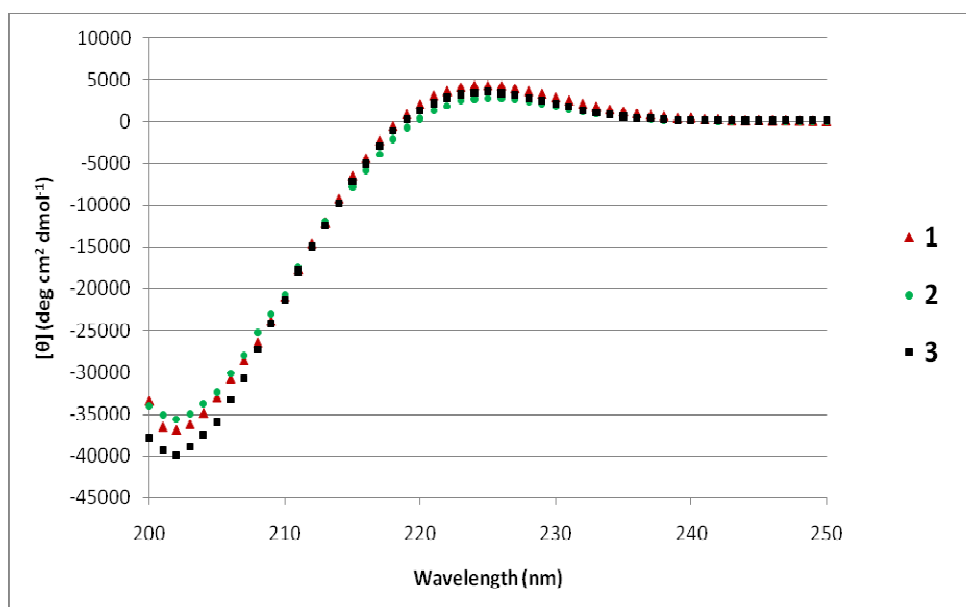
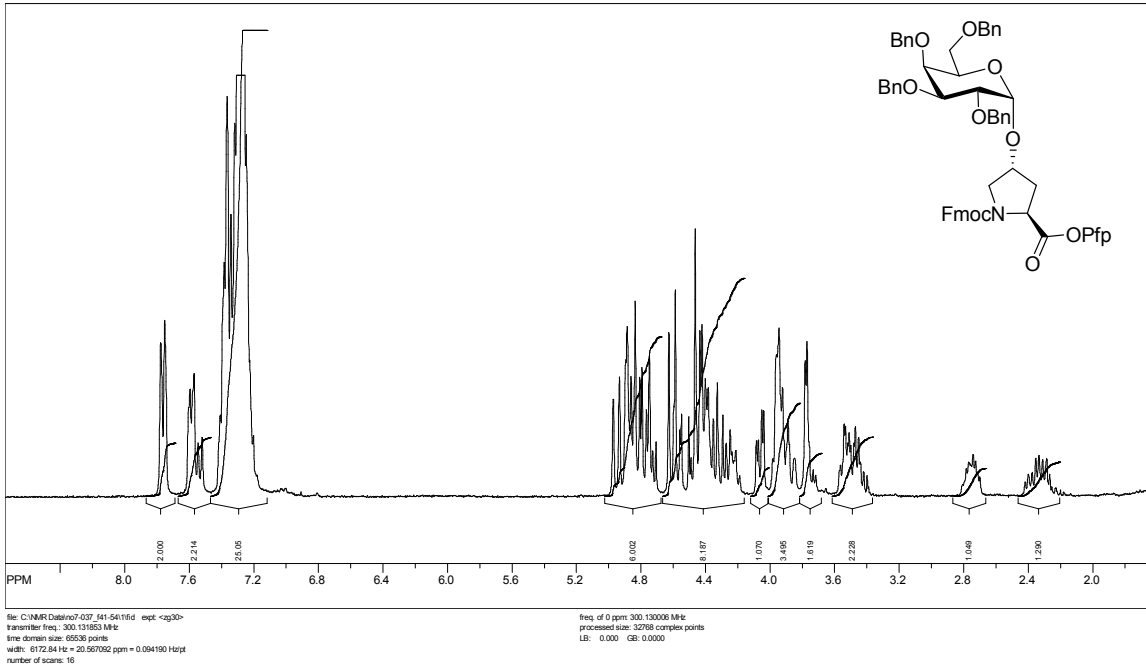
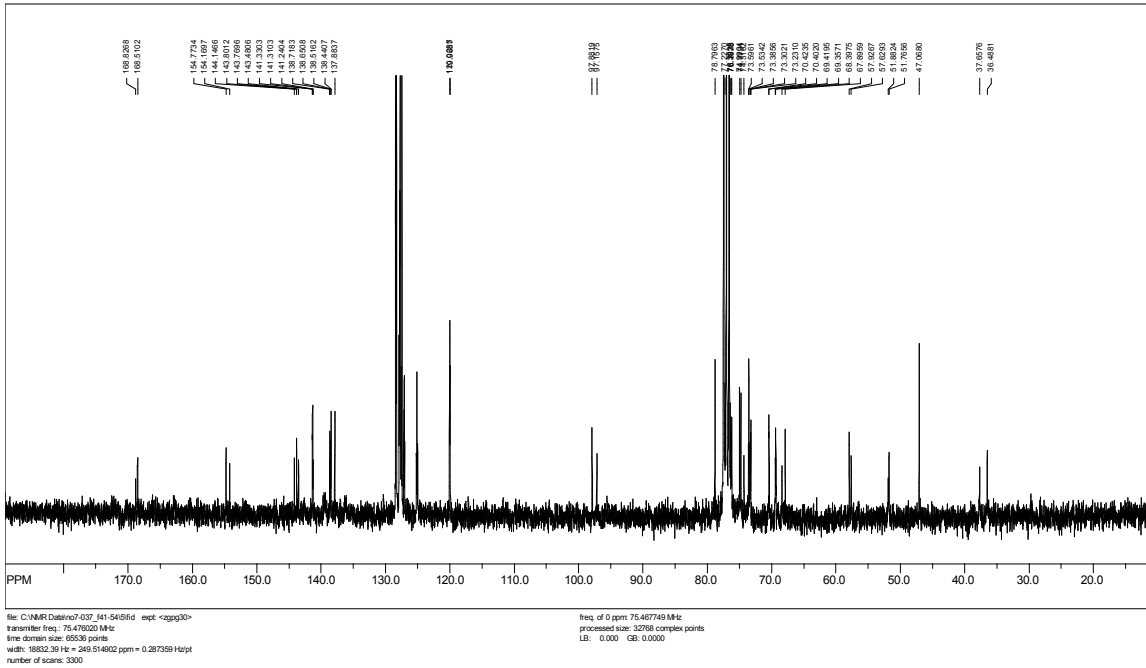
Figure 12.11: CD Curve of **1** in Water at 25 °C**Figure 12.12:** CD Curve of **2** in Water at 25 °C

Figure 12.13: CD Curve of **3** in Water at 25 °C**Figure 12.14:** Comparison of CD Curves of **1-3** in Water at 25 °C

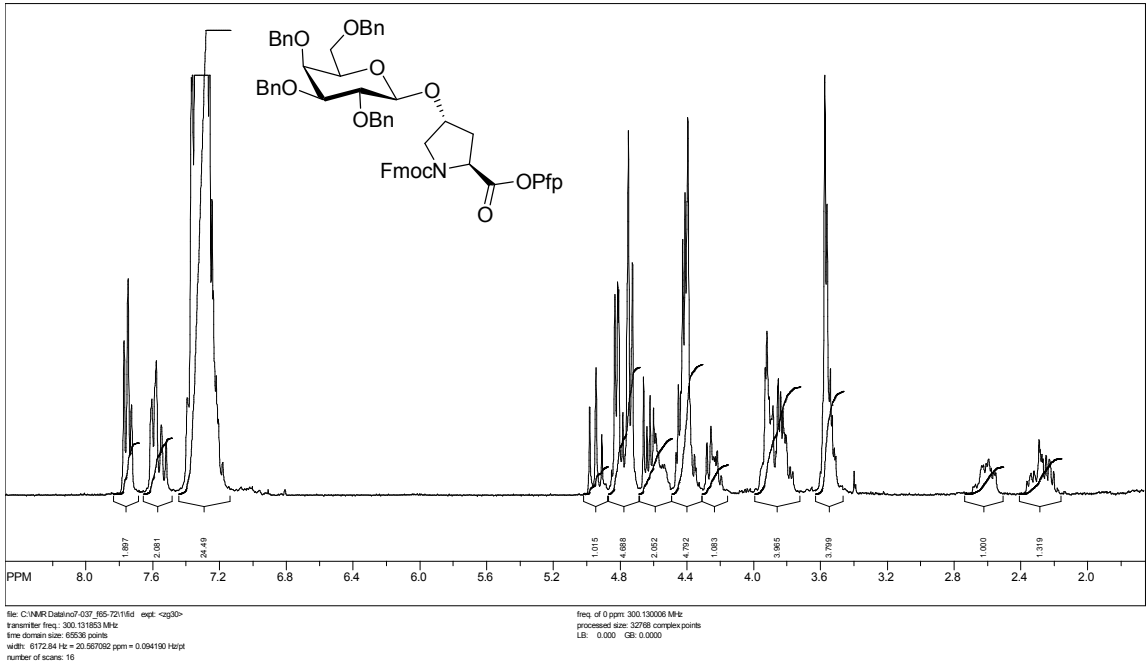
SpinWorks 2.5: PROTON CDCl3 u schweiz 2



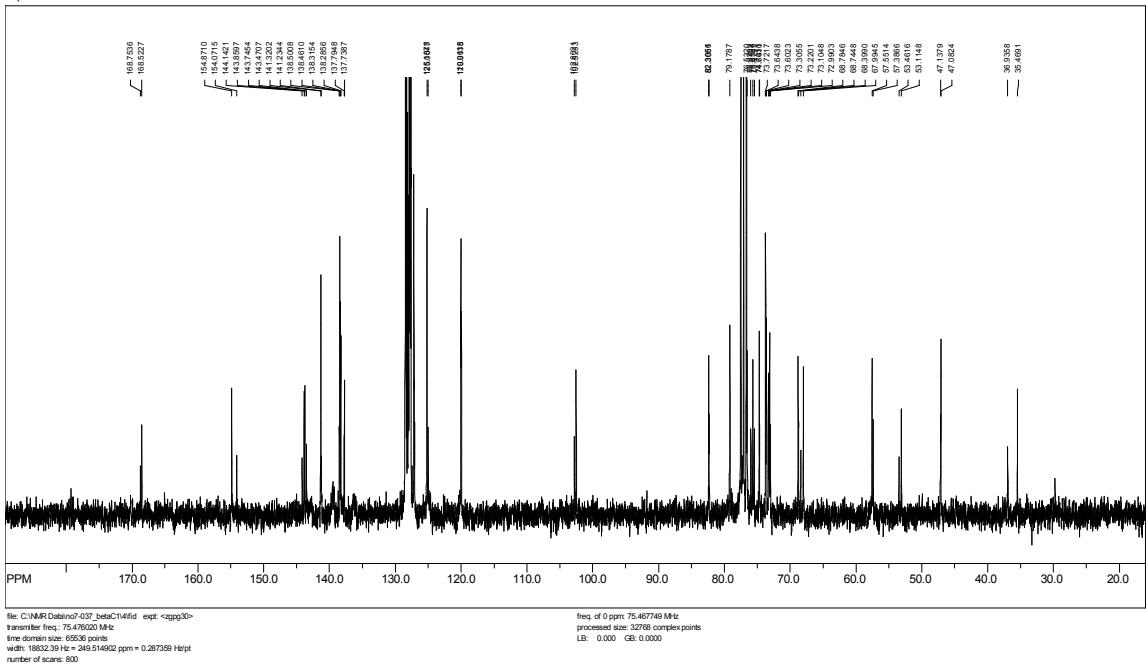
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



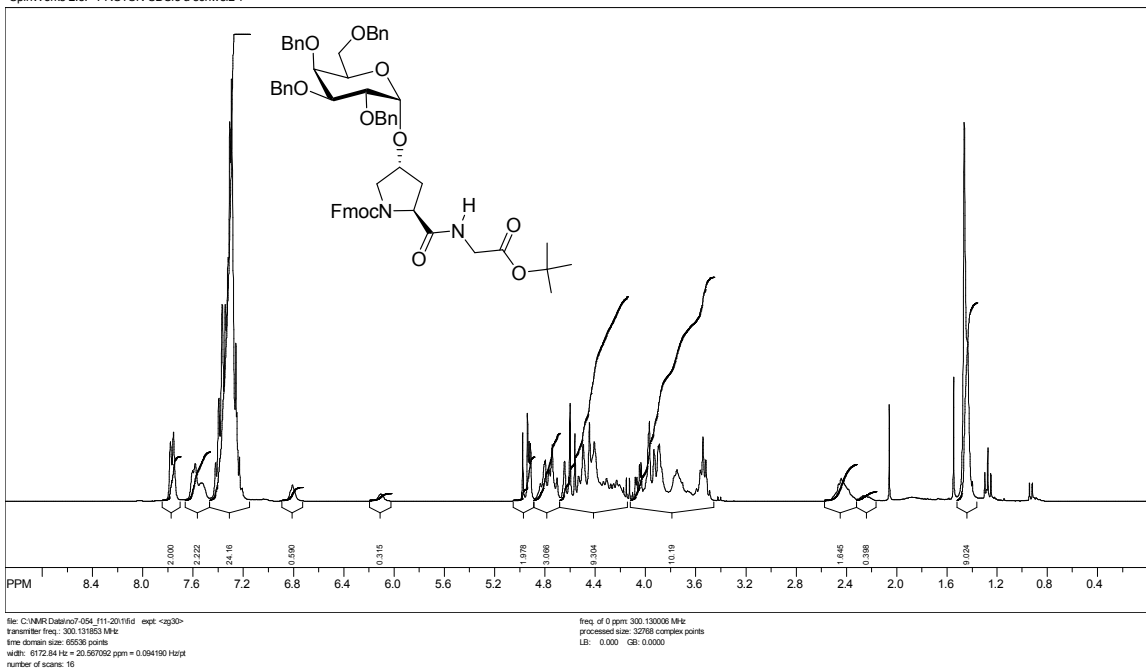
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



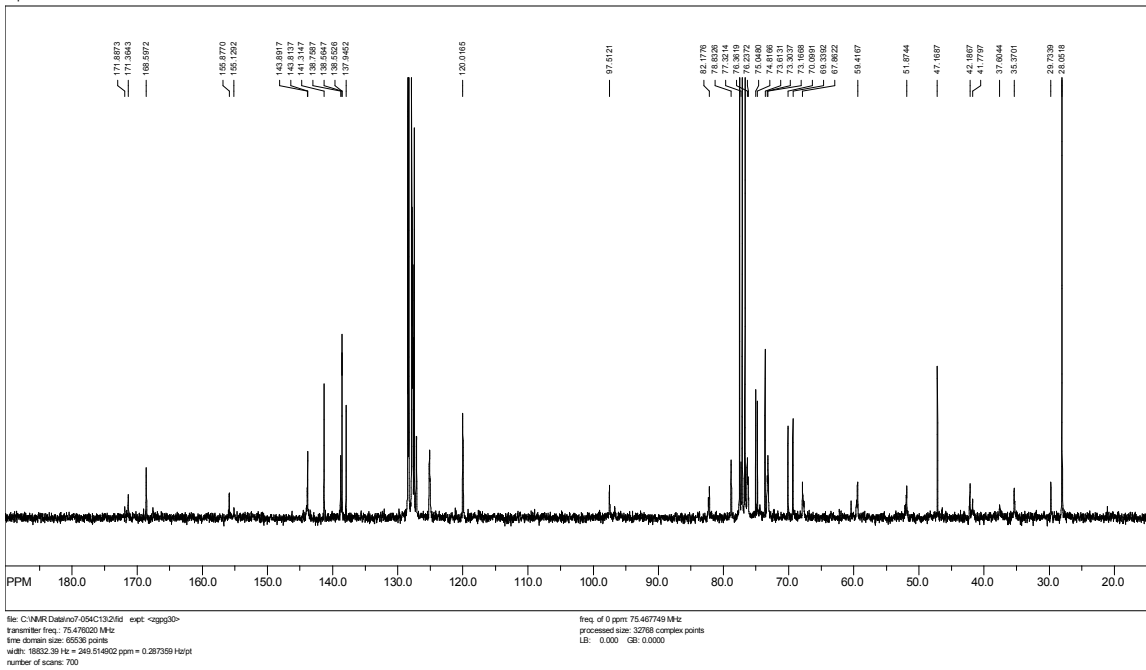
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



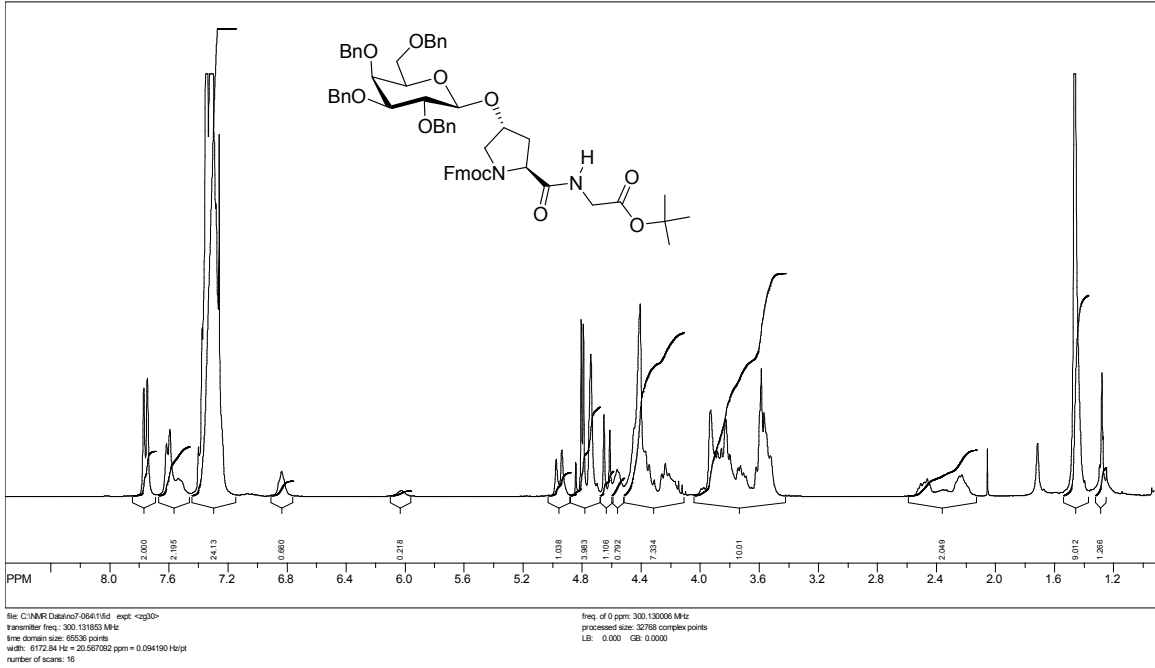
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



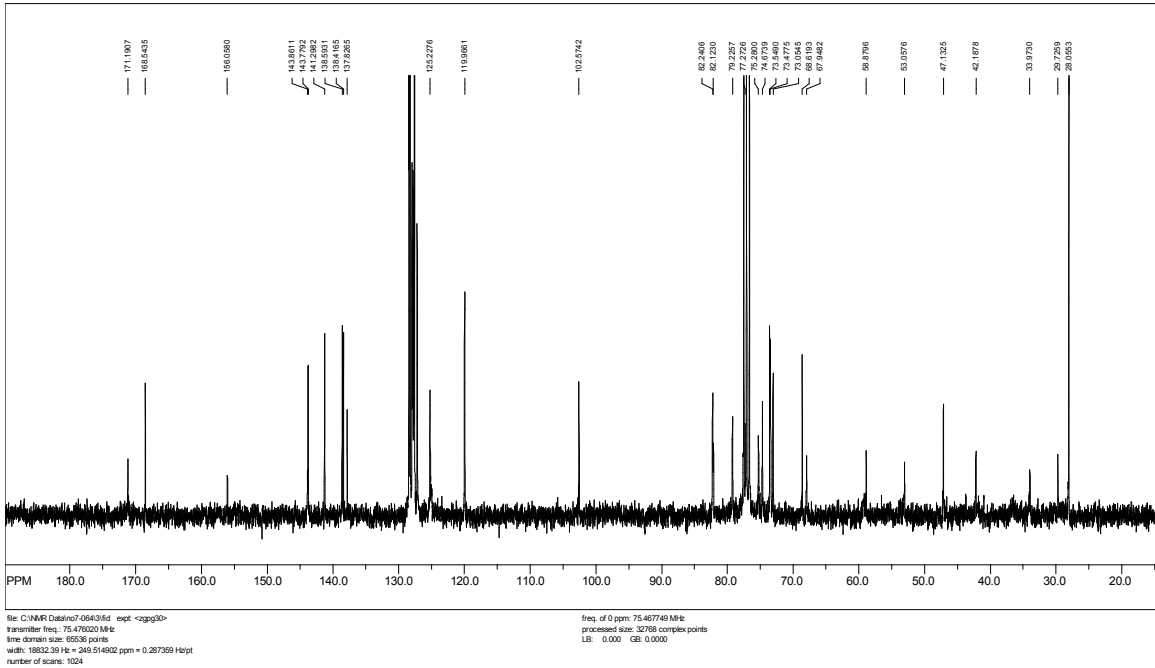
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



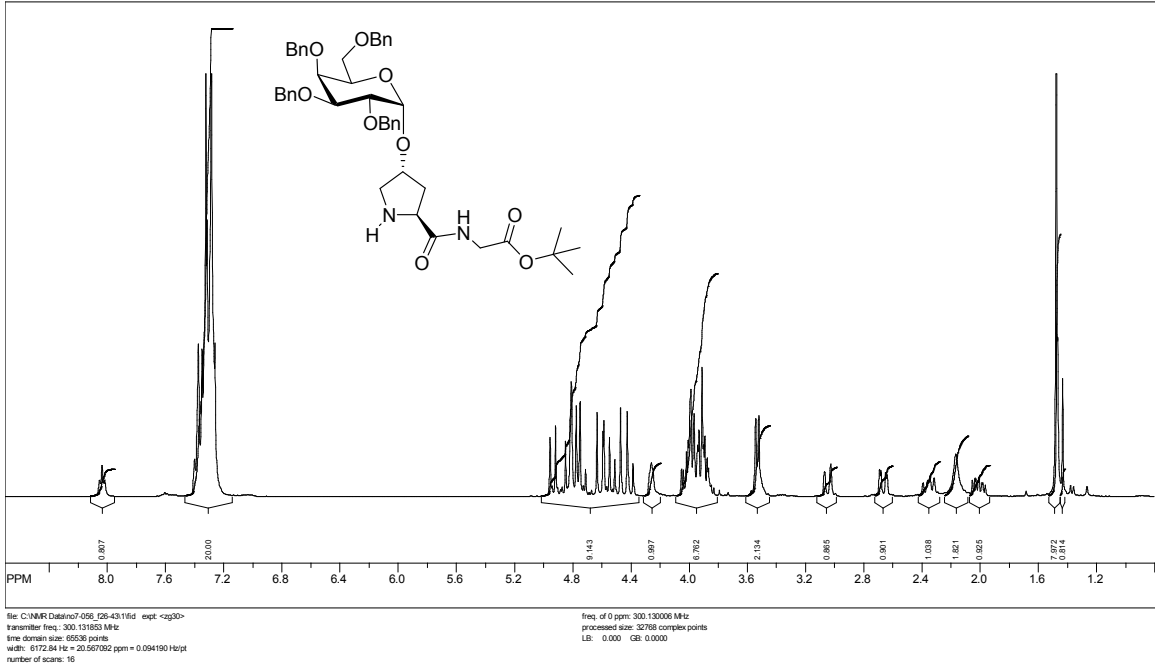
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



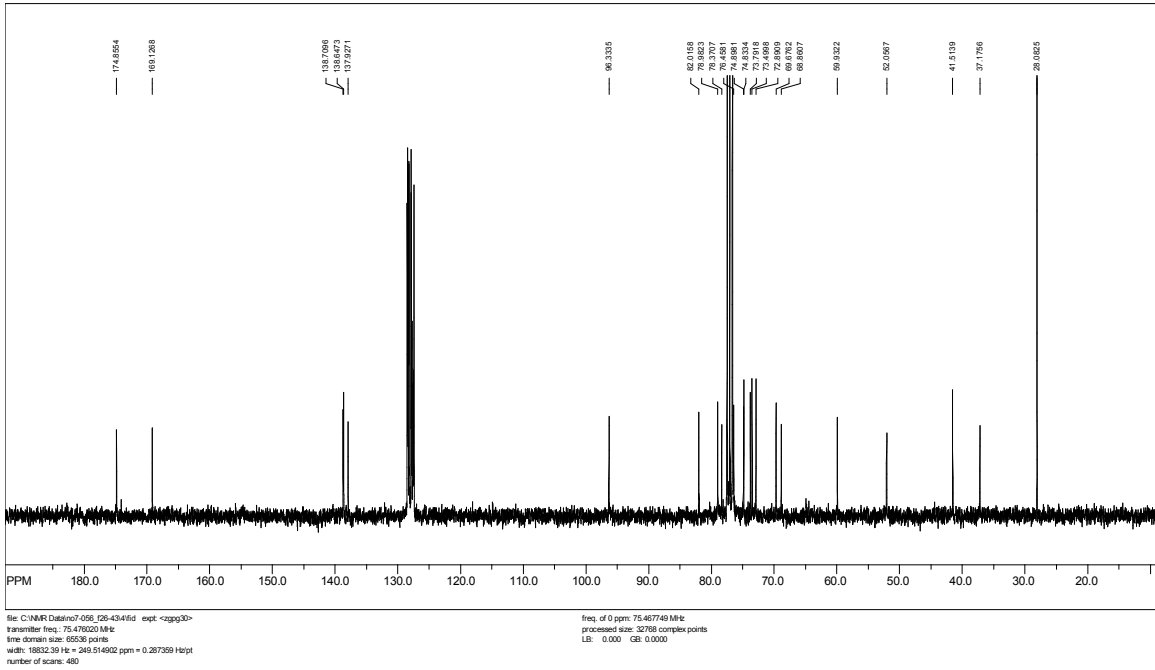
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



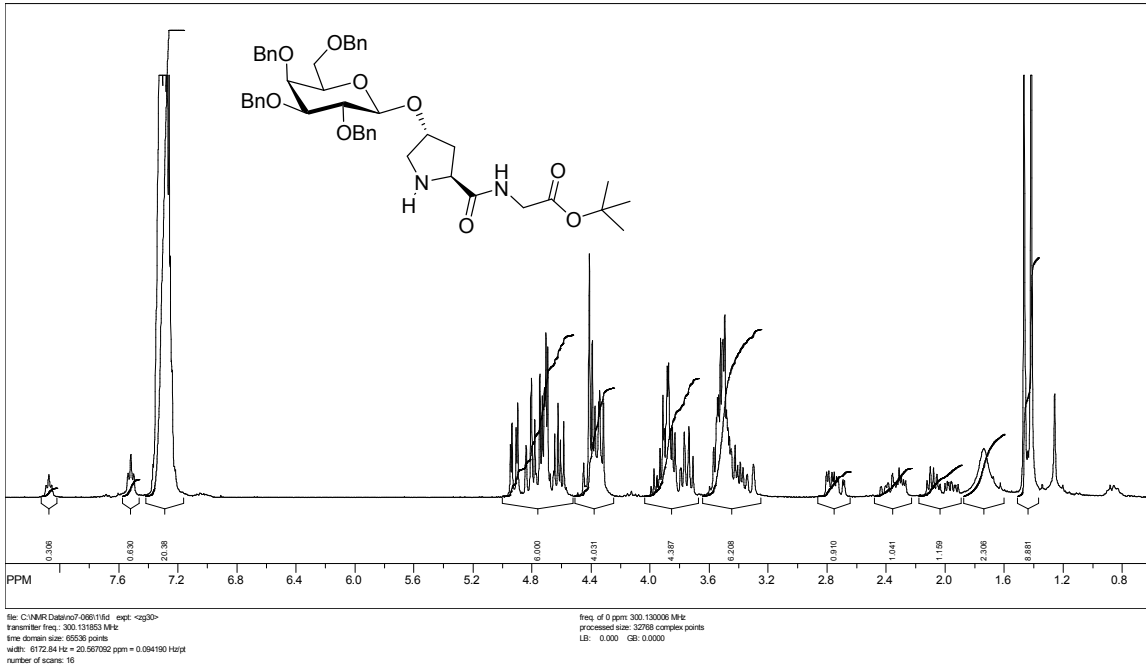
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



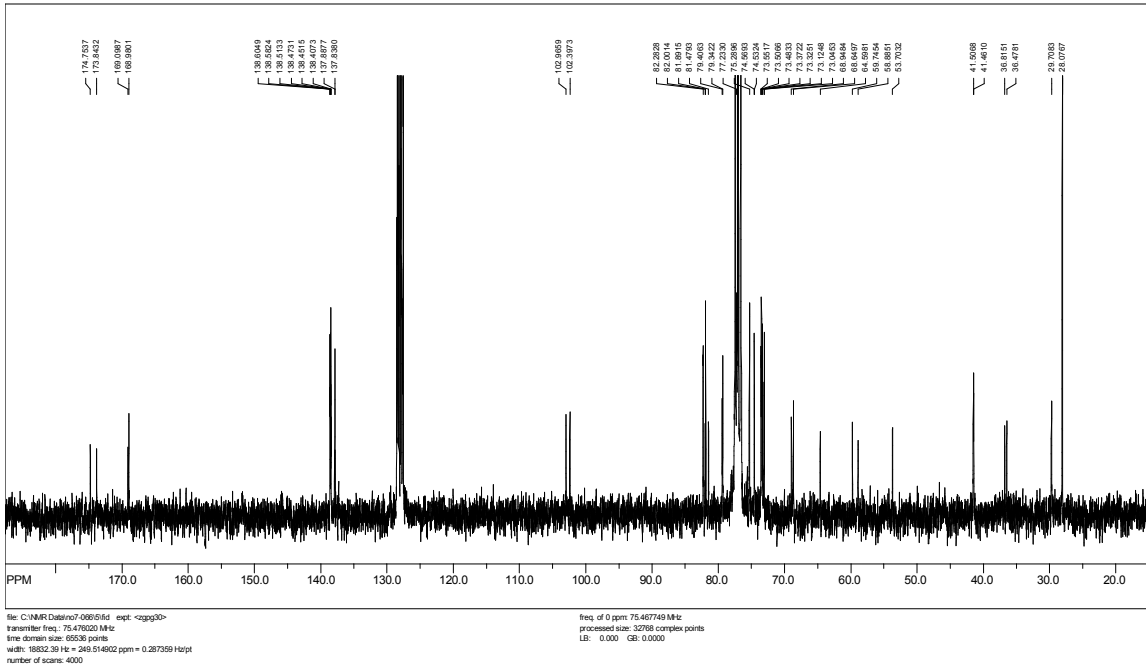
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



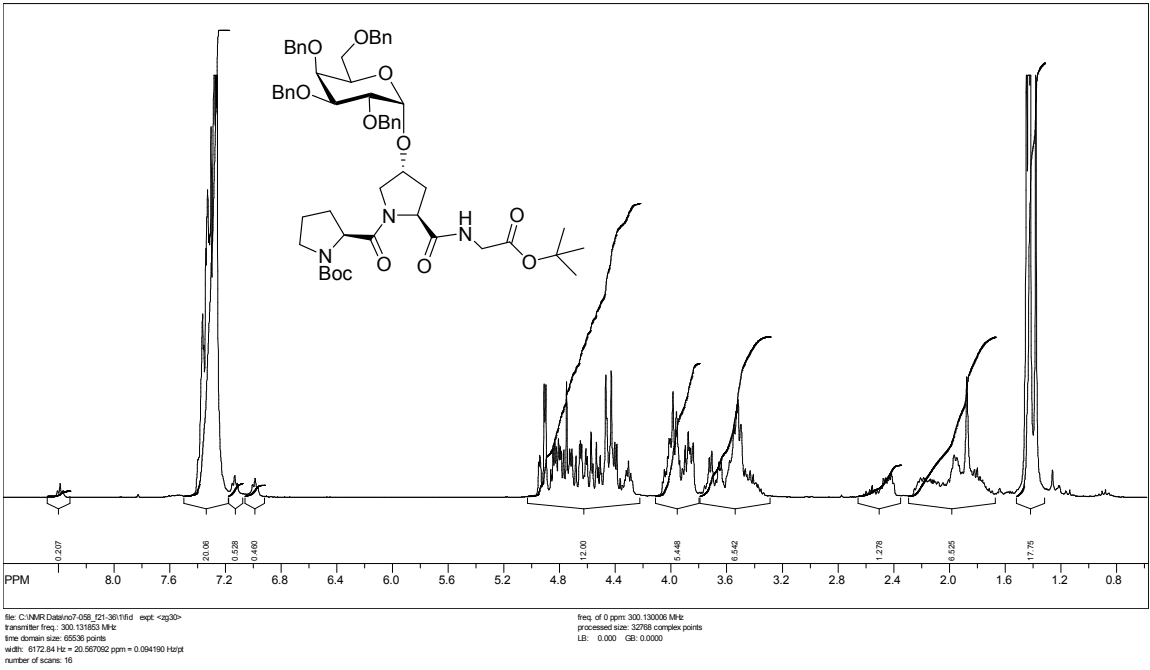
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



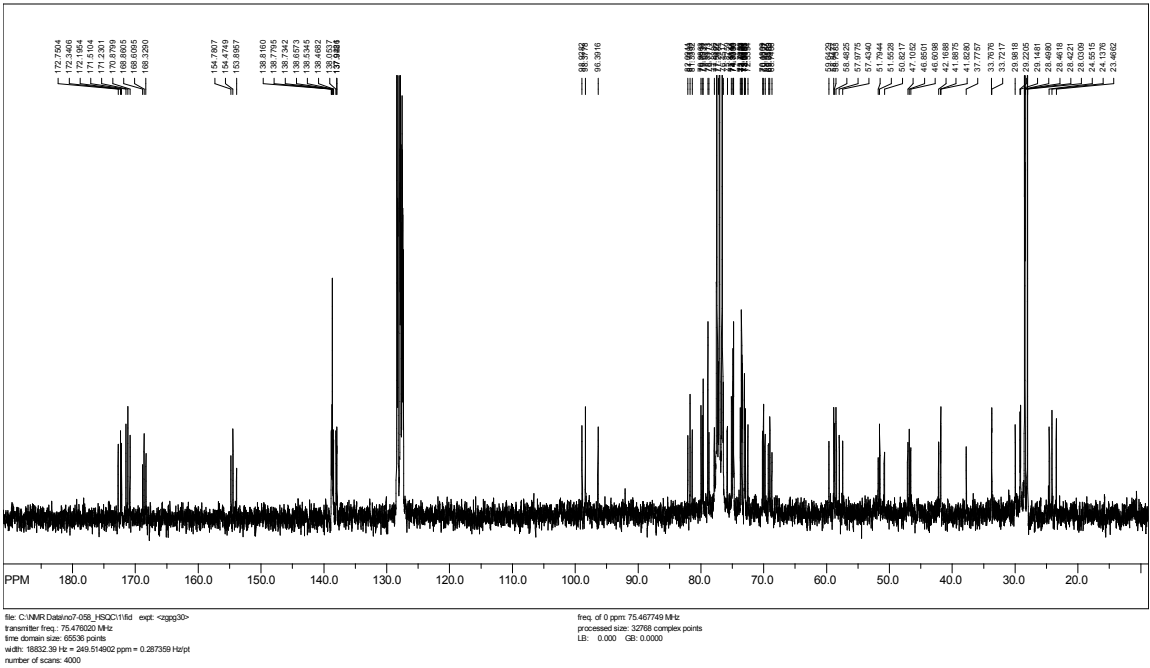
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



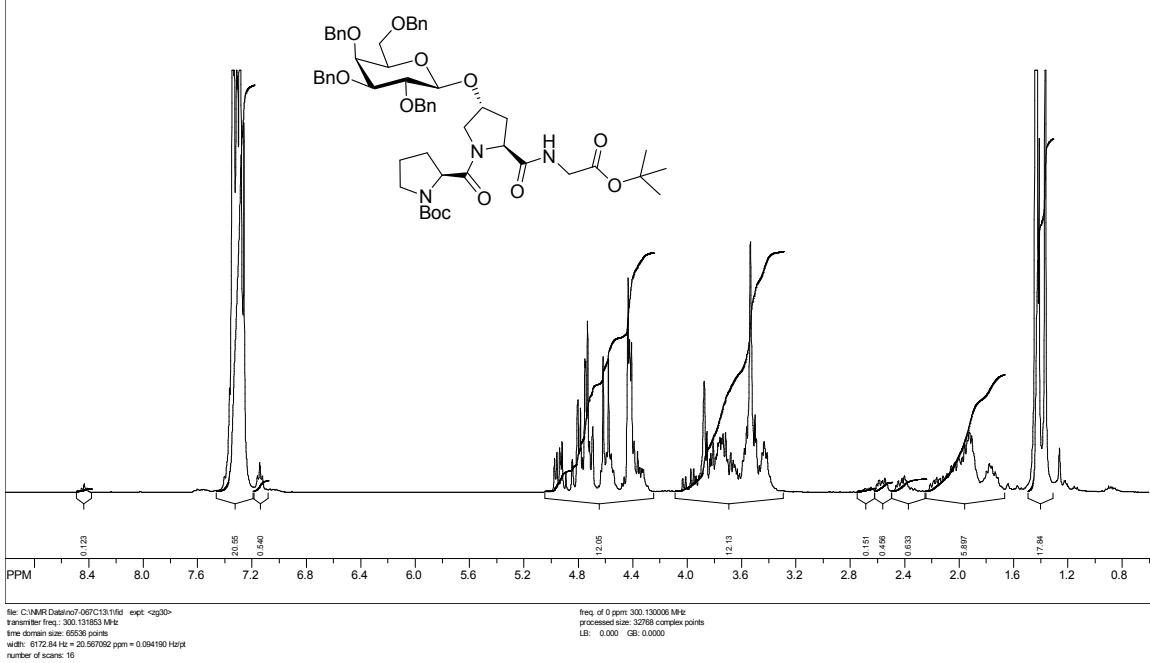
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



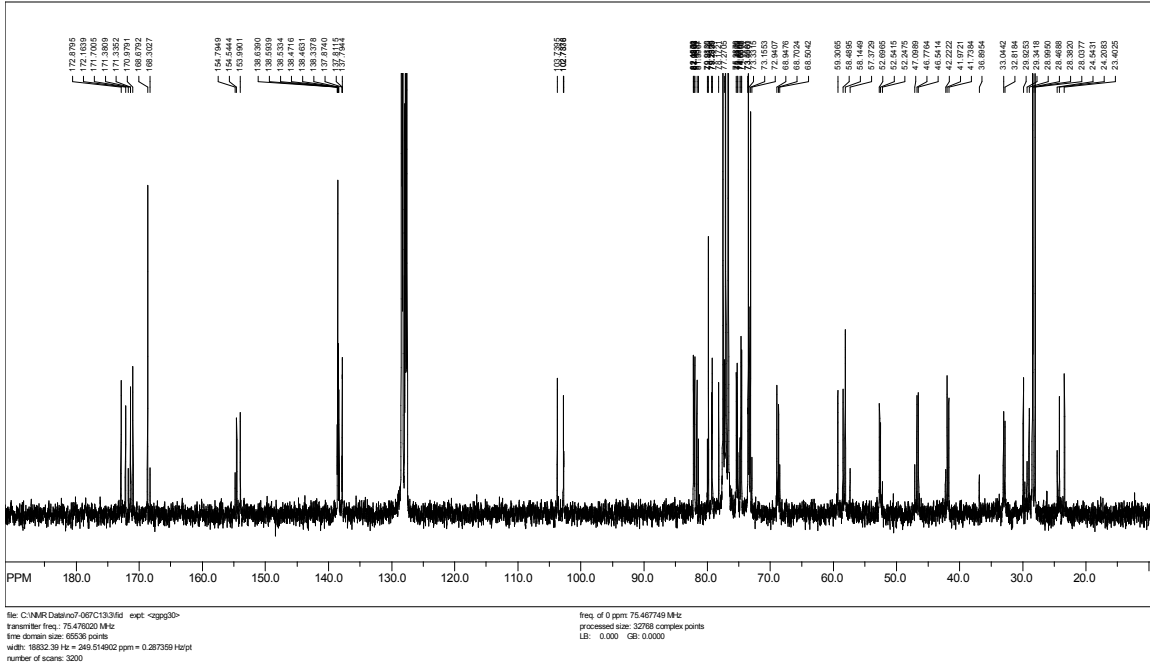
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



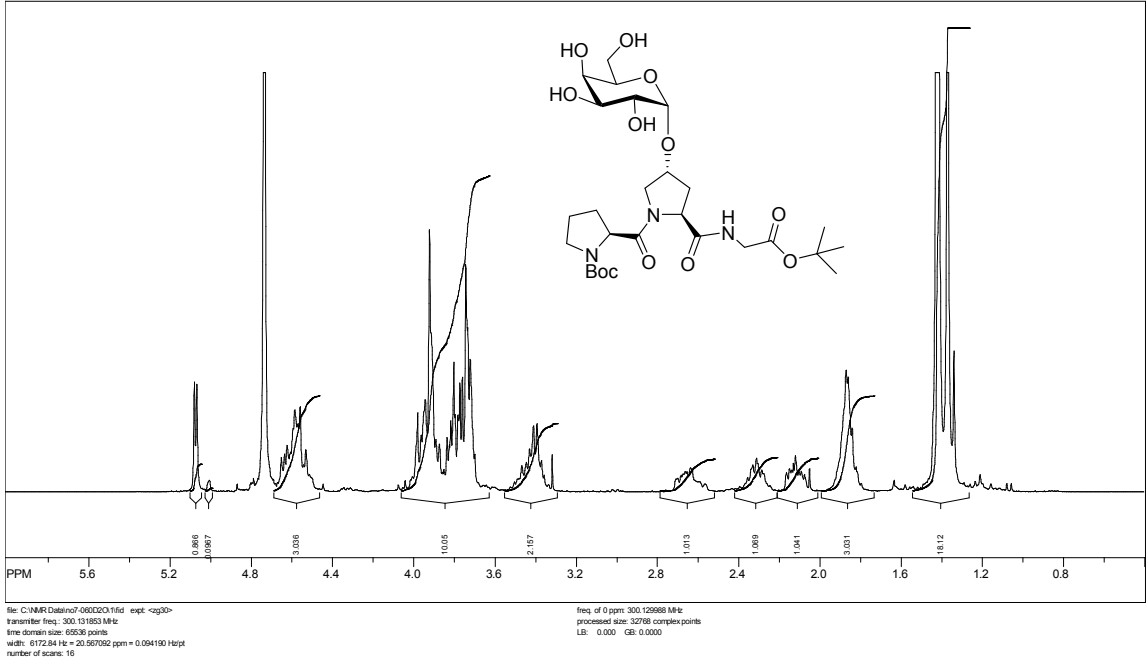
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



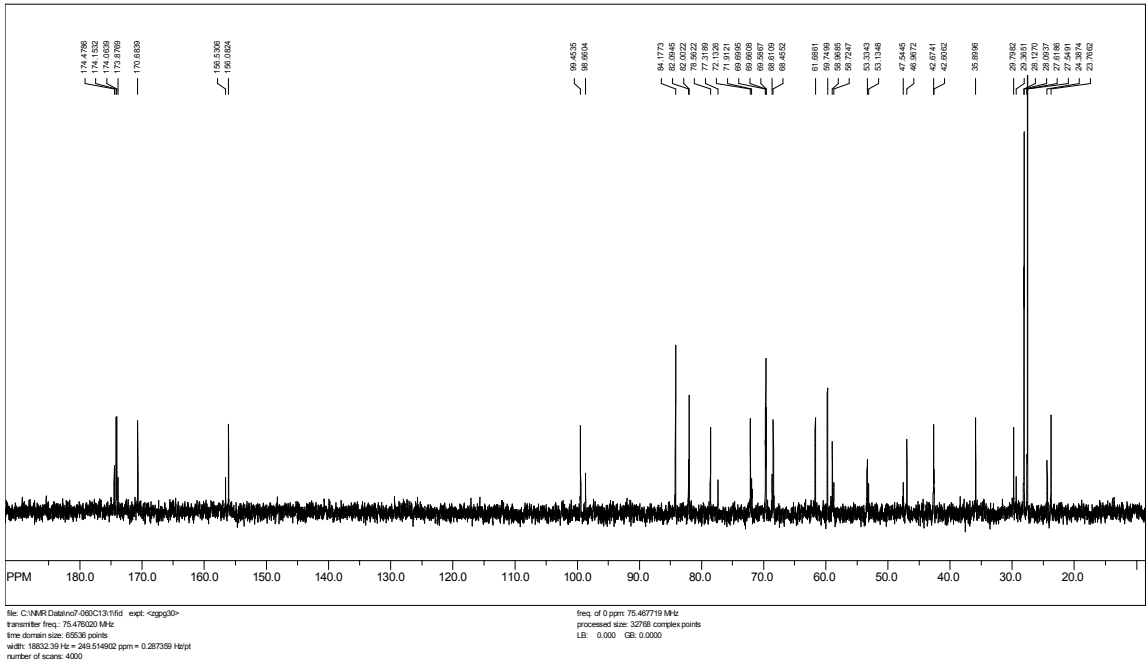
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



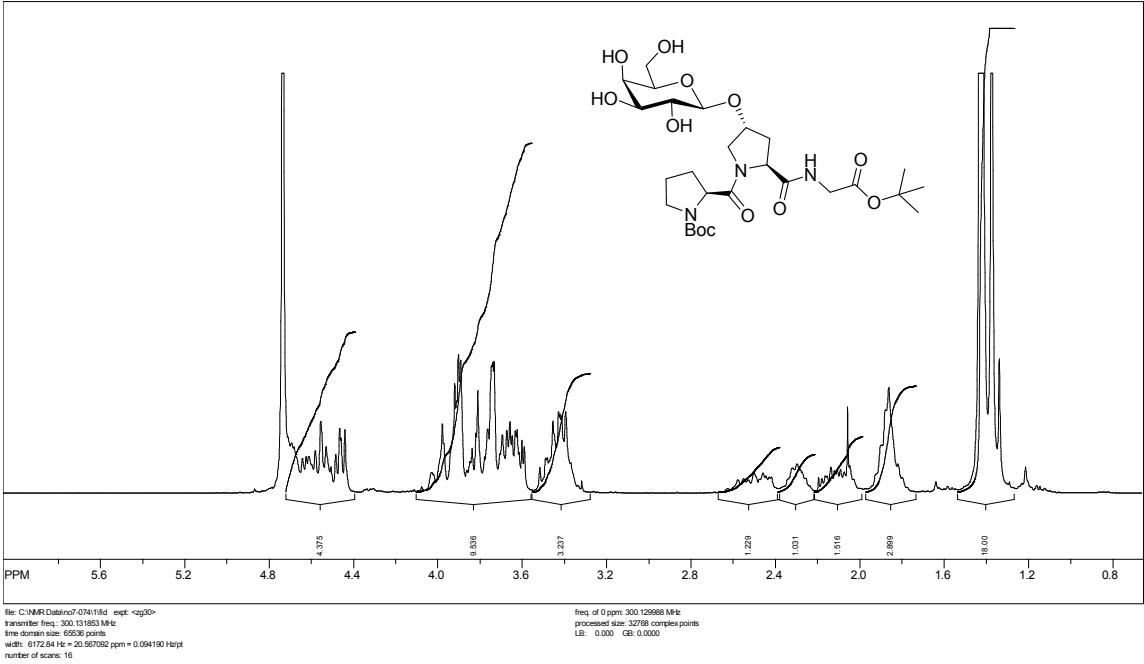
SpinWorks 2.5: PROTON D2O u schweiz 1



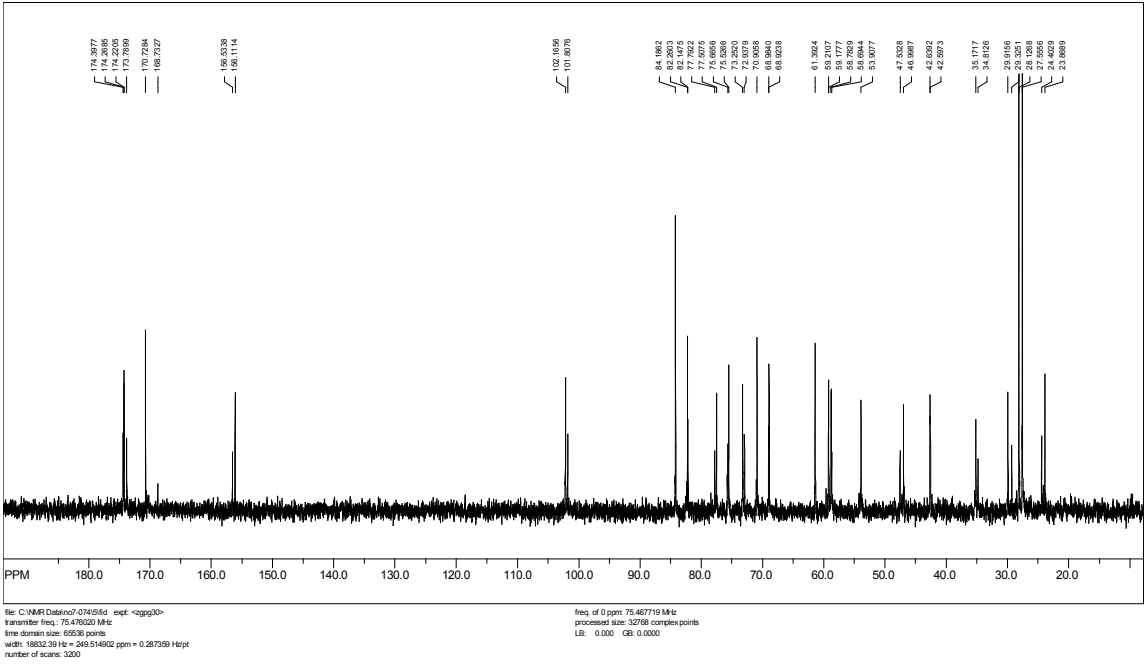
SpinWorks 2.5: C13CPD D2O u schweiz 1



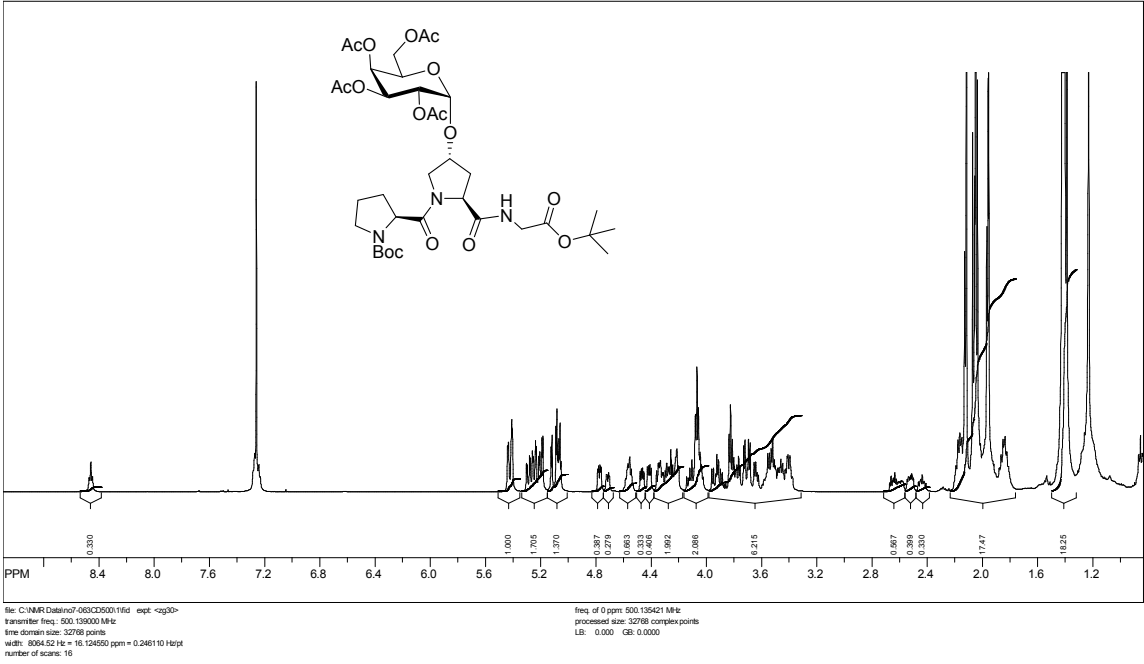
SpinWorks 2.5: PROTON D2O u schweiz 1



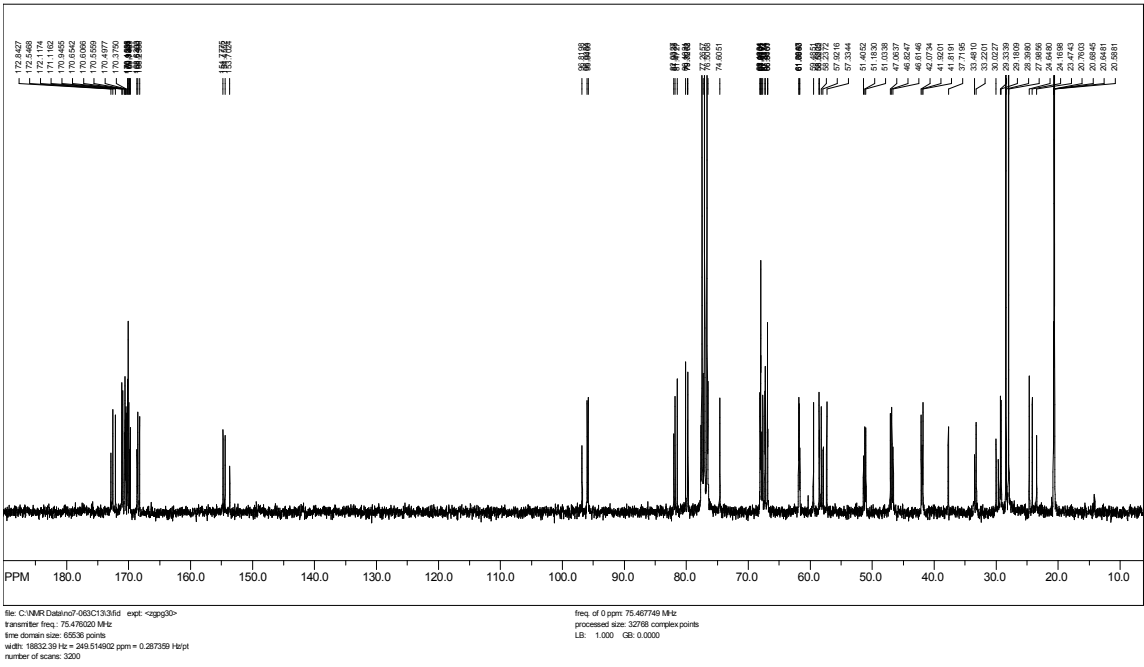
SpinWorks 2.5: C13CPD D2O u schweiz 1



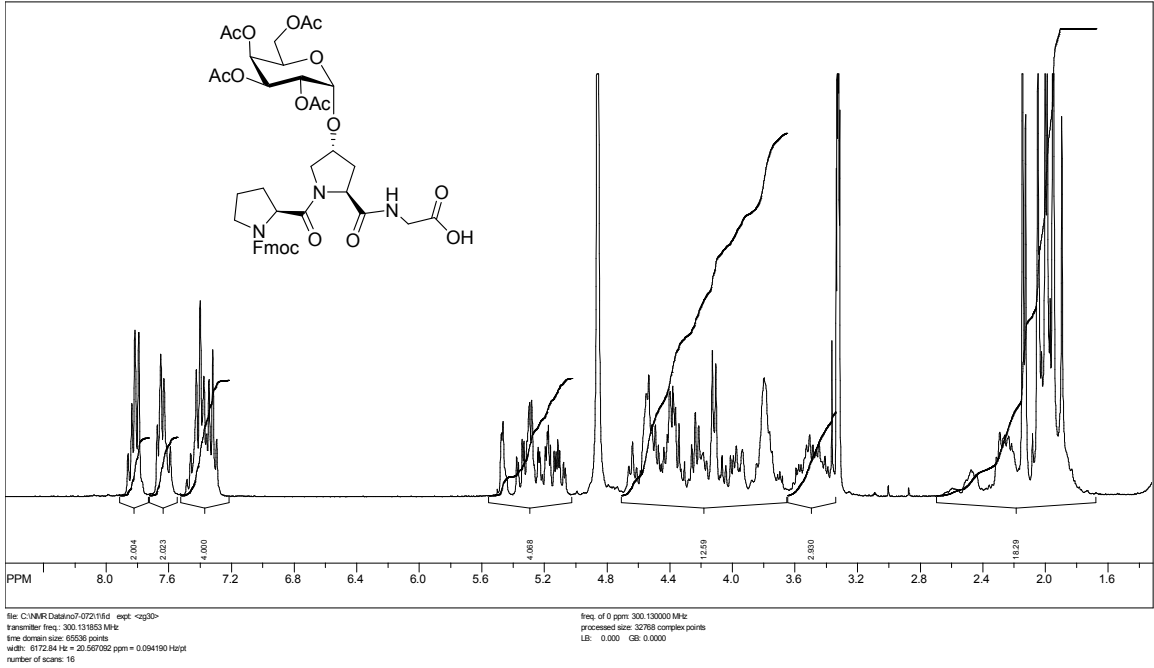
SpinWorks 2.5: standard 1-H survey parameters, AMX500, CDCI3



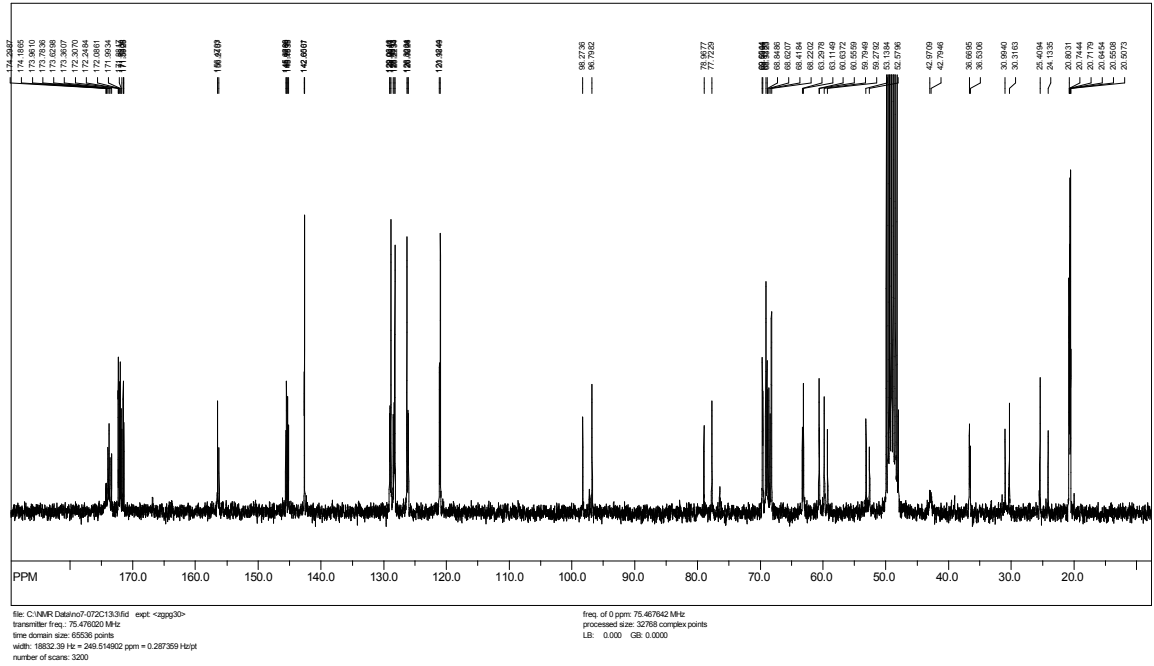
SpinWorks 2.5: C13CPD CDCI3 u schweiz 1



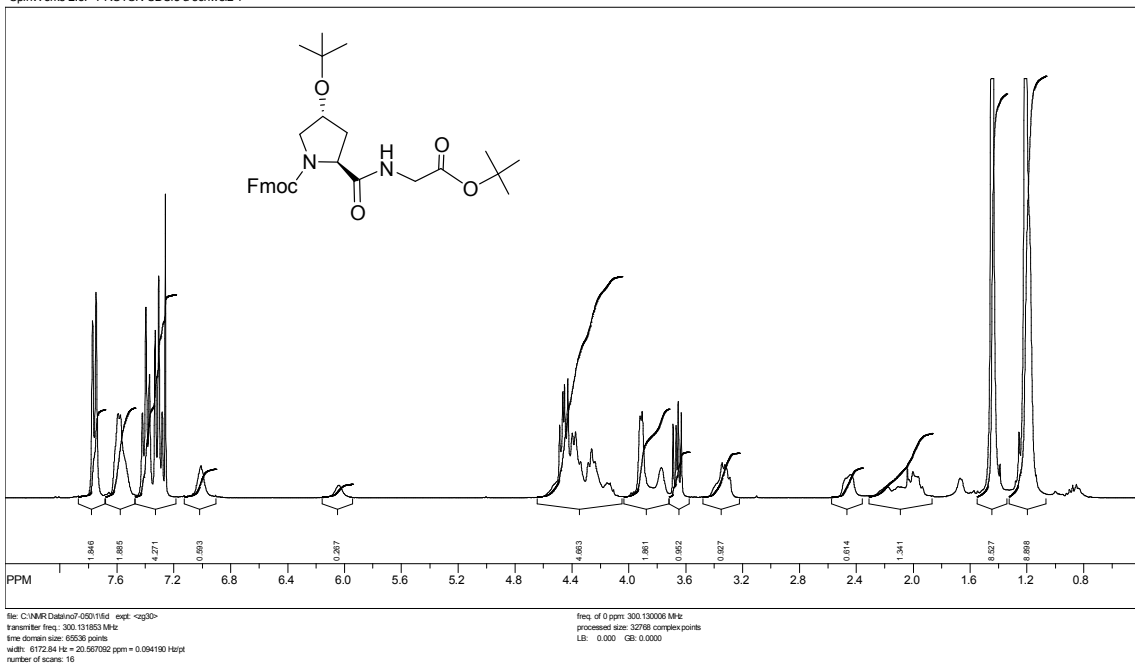
SpinWorks 2.5: PROTON MeOH u schweiz 1



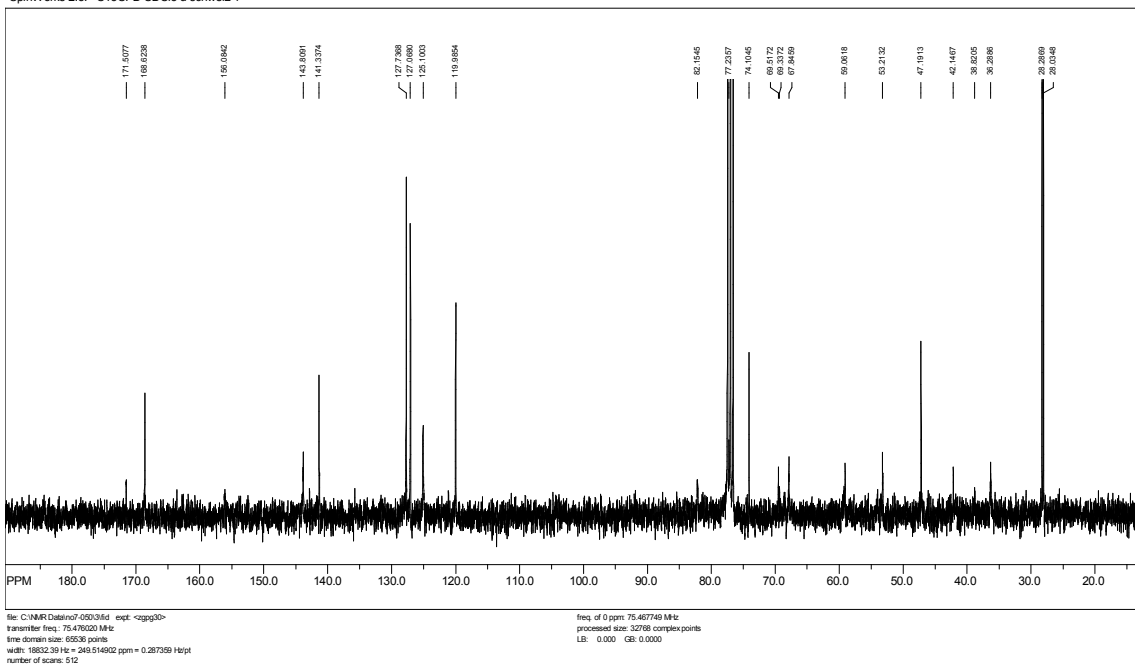
SpinWorks 2.5: C13CPD MeOH u schweiz 1



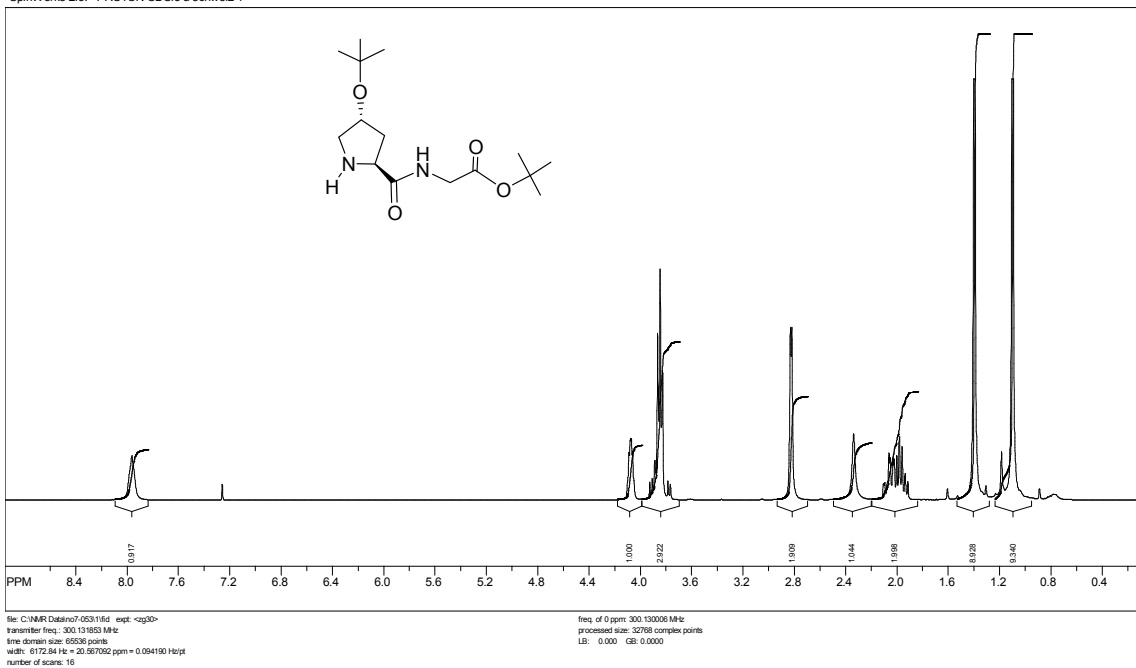
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



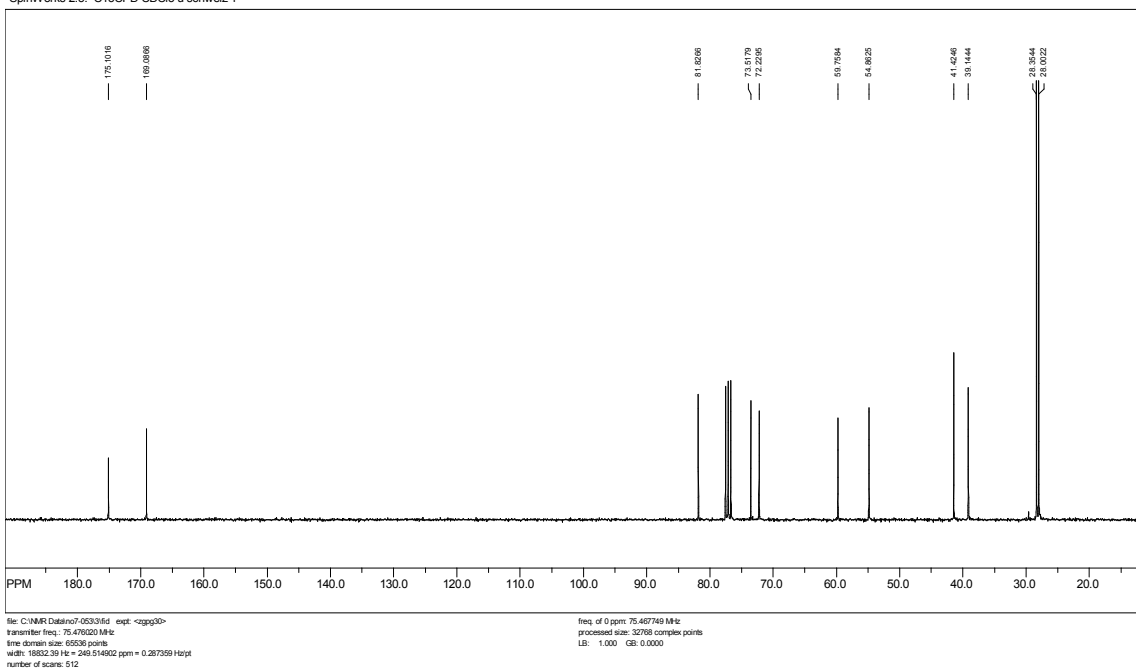
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



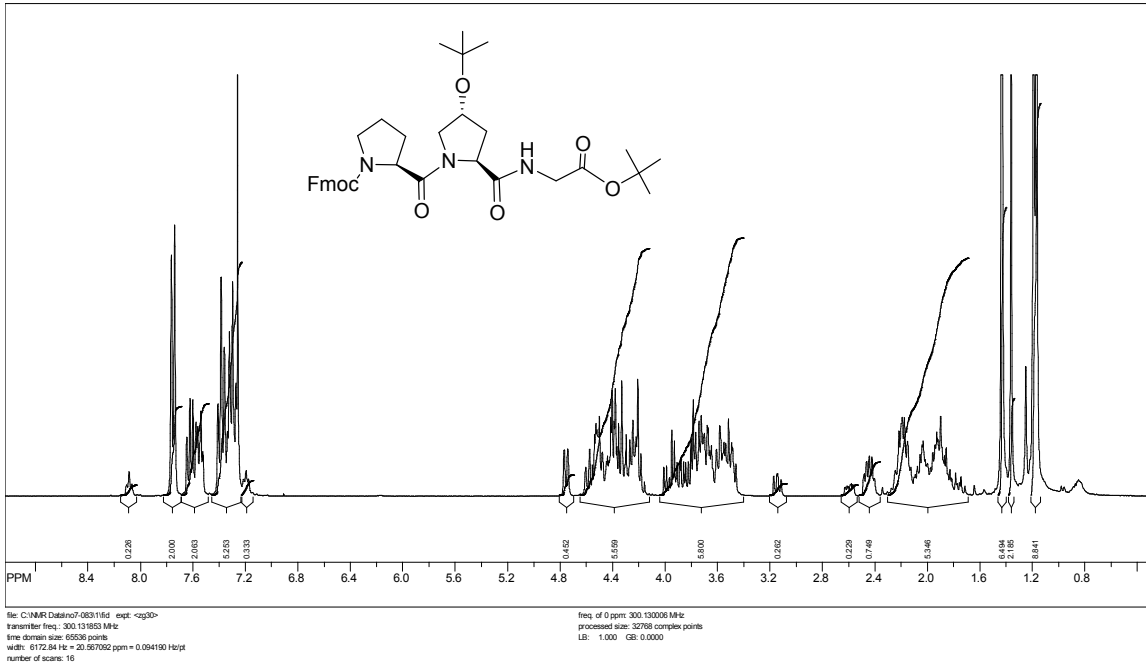
SpinWorks 2.5: PROTON CDCI3 u schweiz 1



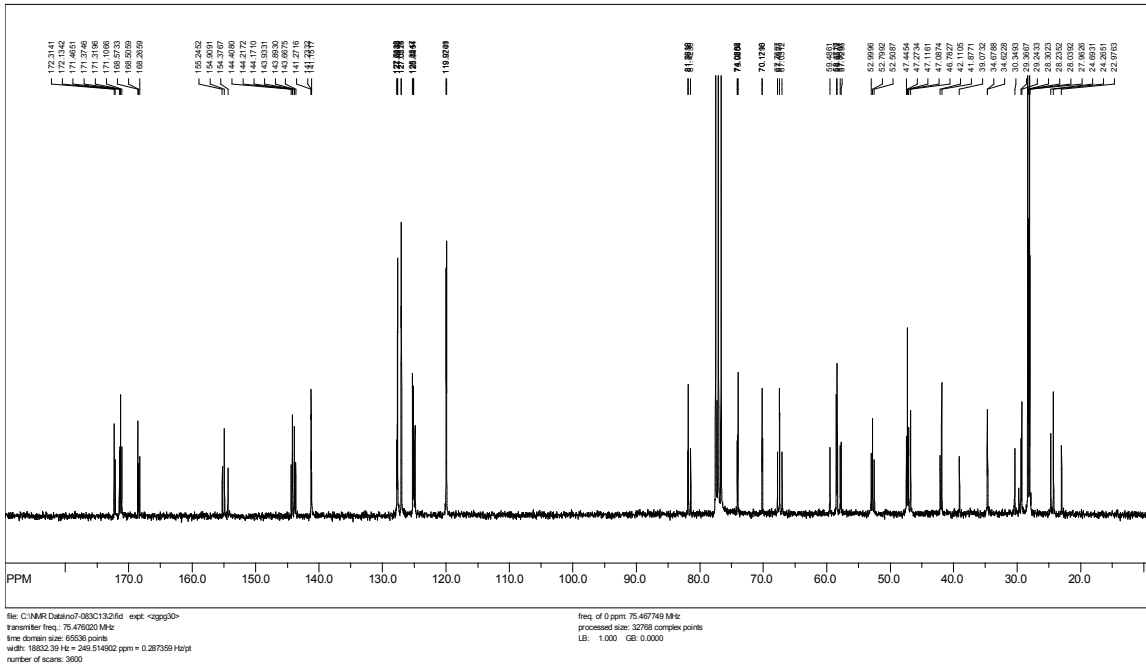
SpinWorks 2.5: C13CPD CDCI3 u schweiz 1



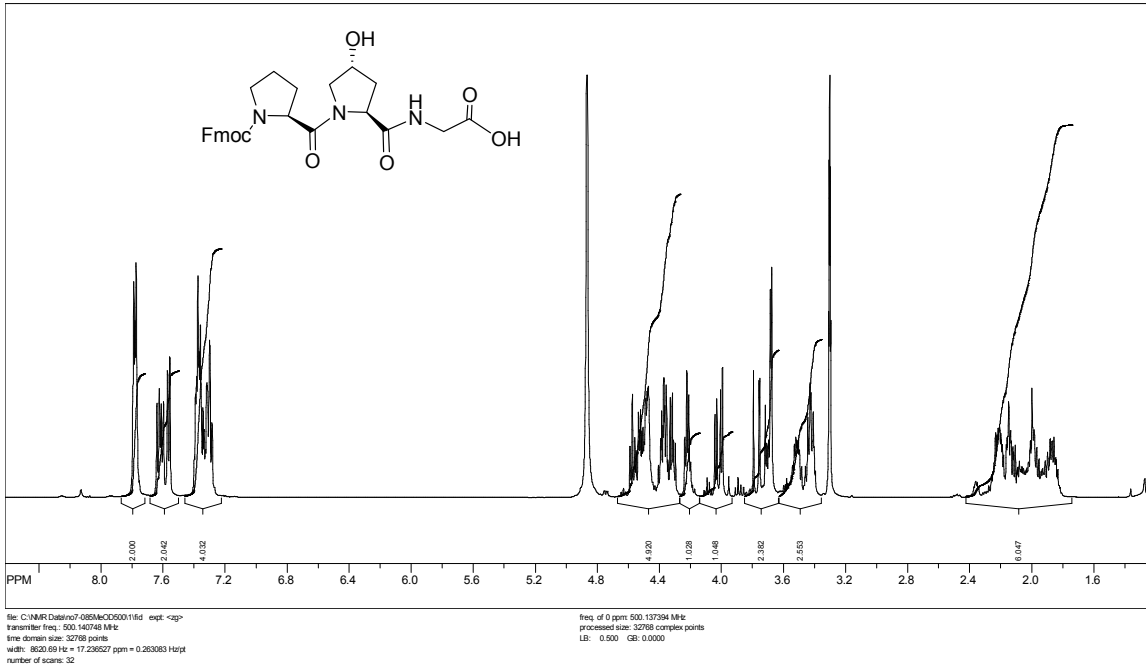
SpinWorks 2.5: PROTON CDCl3 u schweiz 1



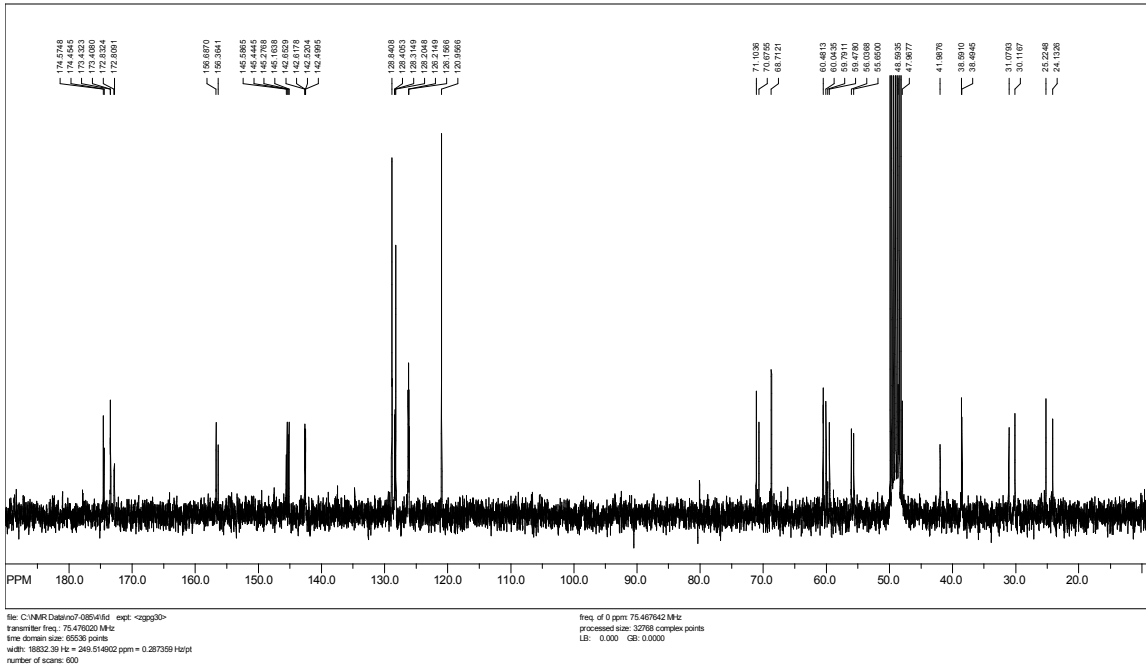
SpinWorks 2.5: C13CPD CDCl3 u schweiz 1



SpinWorks 2.5: AMX 500 proton survey parameters, CD3OD



SpinWorks 2.5: C13CPD MeOH u schweiz 2

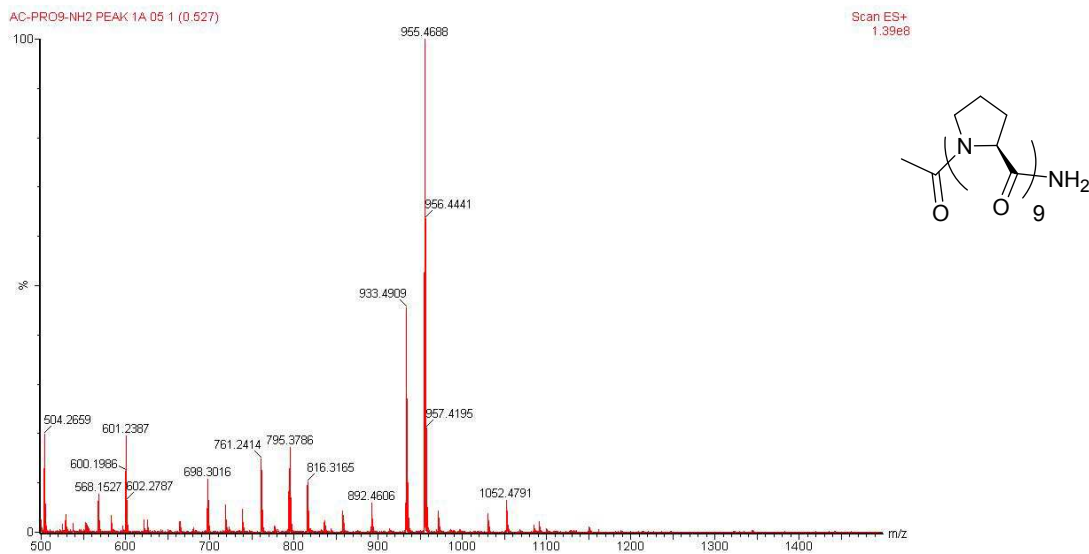


**Chapter 13: Contiguous *O*-Galactosylation of 4*R*-Hydroxy-L-Proline
Residues Forms Hyper-Stable Polyproline II Helices**

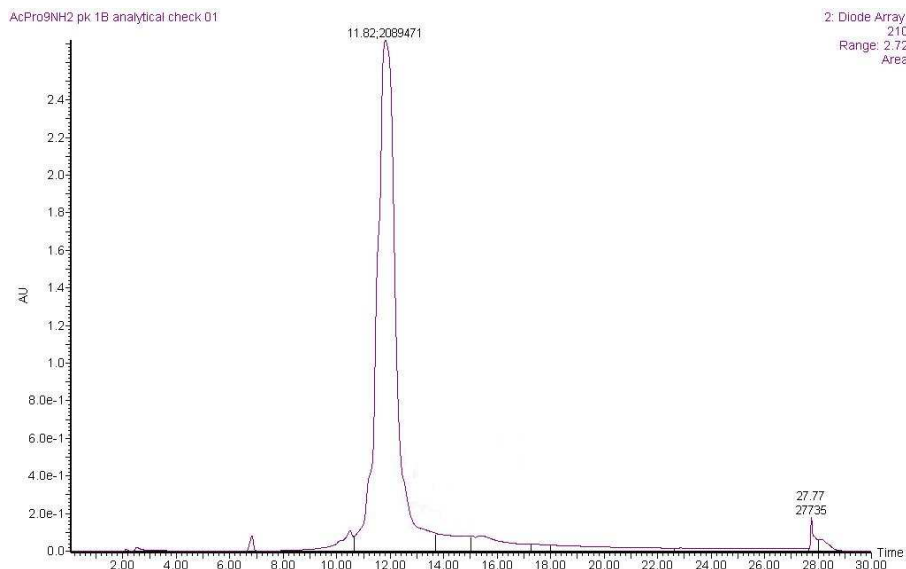
Supporting Information For Chapter 7

Table of Contents

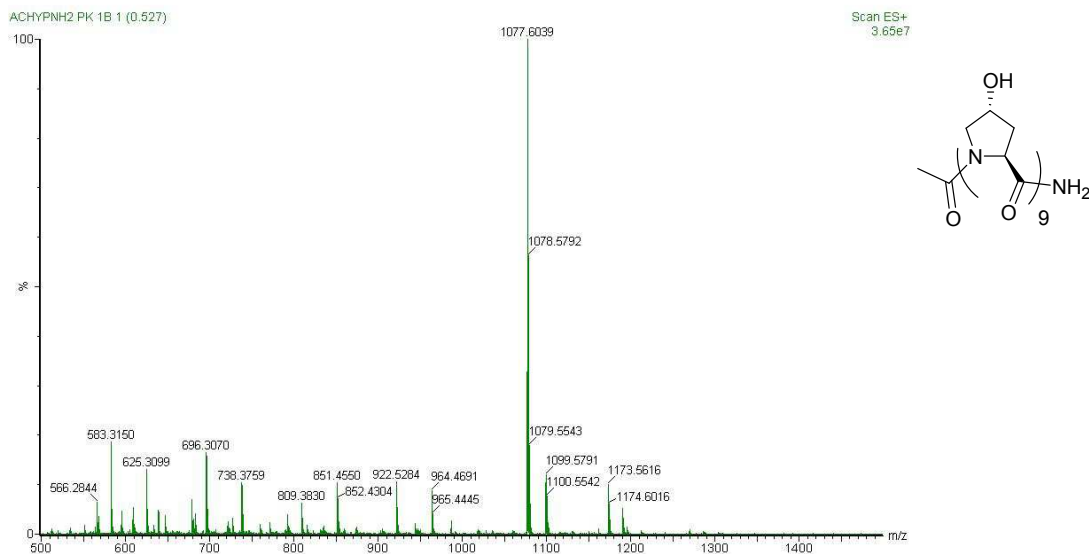
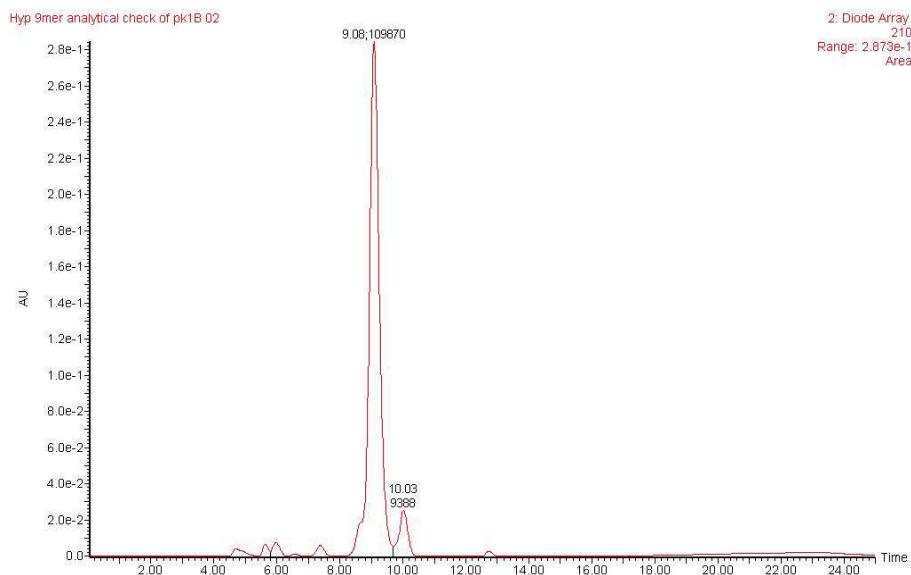
Mass Spectra and Analytical Chromatograms for 1-3	508-510
CD Spectra for 1-3	511-512
Table of Ellipticity as a Function of Temperature for 1-3	513
CD Melting Curves for 1-3	514-515
NMR Spectra for 1-4	516-517

Figure 13.1: Electrospray Ionization Mass Spectrum of **1**

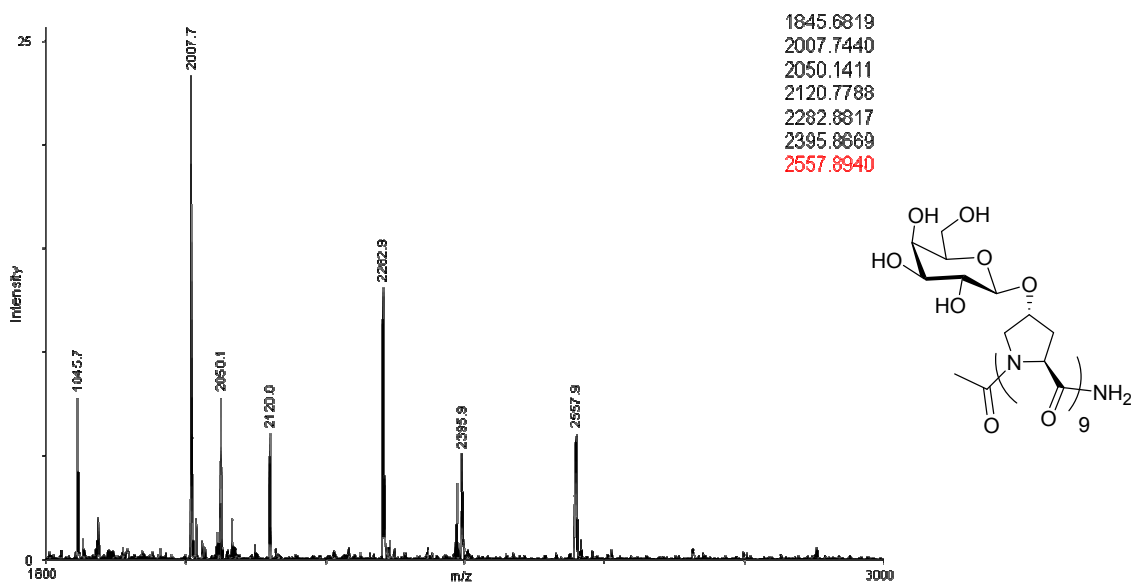
MS (ES) calc. for $C_{47}H_{68}N_{10}NaO_{10}$ ($M + Na$)⁺: 955.50. Found ($M + Na$)⁺: 955.47.

Figure 13.2: Analytical chromatogram of **1** monitored at 210 nm

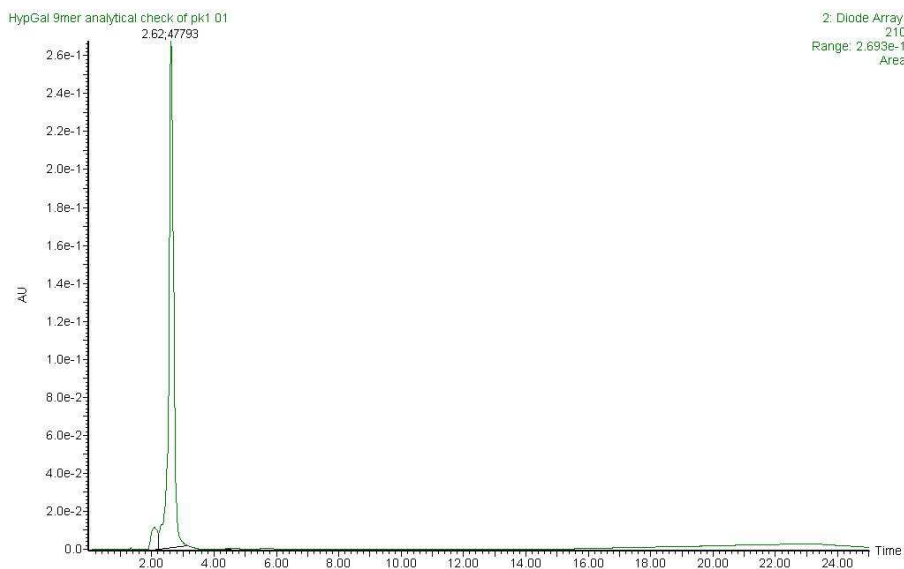
Eluted using 0 to 10% MeCN in H_2O over 30 min, at a flow rate of $1 \text{ mL} \cdot \text{min}^{-1}$.

Figure 13.3: Electrospray Ionization Mass Spectrum of **2****Figure 13.4:** Analytical chromatogram of **2** monitored at 210 nm

Eluted using 0 to 10% MeCN in H_2O over 30 min, at a flow rate of $1 \text{ mL} \cdot \text{min}^{-1}$.

Figure 13.5: MALDI spectrum of **3**

MS (MALDI) calc. for $C_{101}H_{158}N_{10}NaO_{64}$ ($M + Na$)⁺: 2557.93. Found ($M + H$)⁺: 2557.89.

Figure 13.6: Analytical chromatogram of **3** monitored at 210 nm

Eluted using 0 to 10% MeCN in H_2O over 30 min, at a flow rate of $1 \text{ mL} \cdot \text{min}^{-1}$.

Figure 13.7: CD curve of **1** in water at 25 °C at a concentration of 0.4 mM

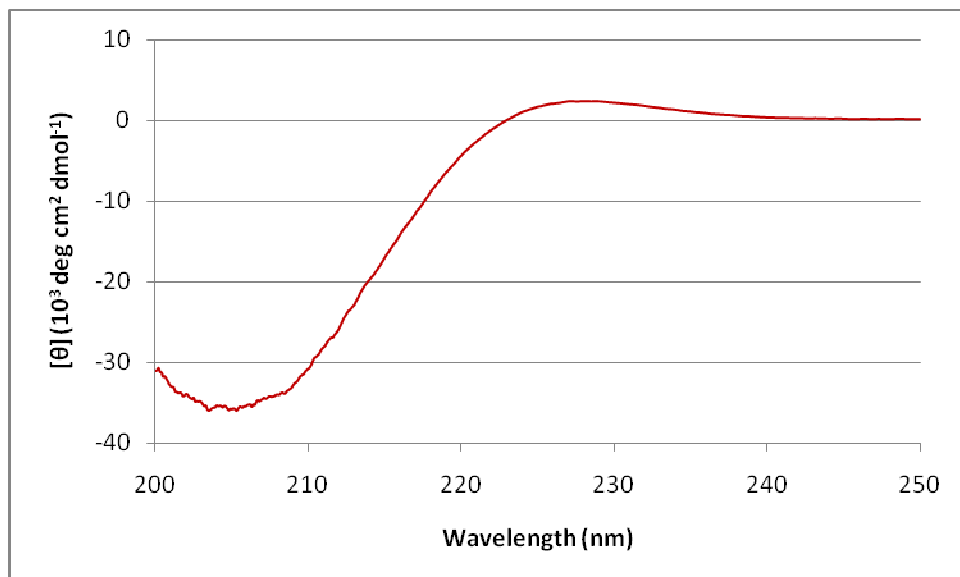


Figure 13.8: CD curve of **2** in water at 25 °C at a concentration of 0.4 mM

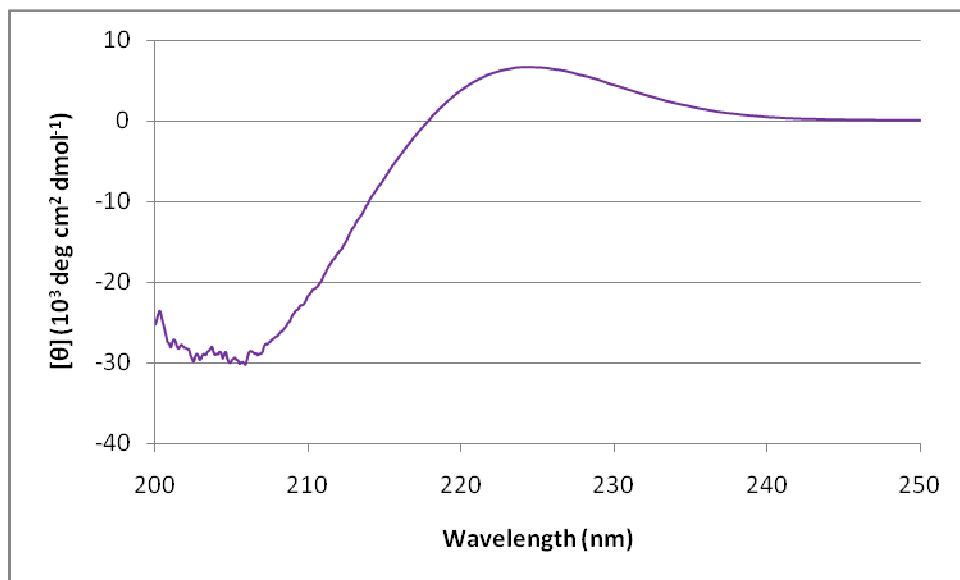


Figure 13.9: CD curve of **3** in water at 25 °C at a concentration of 0.4 mM

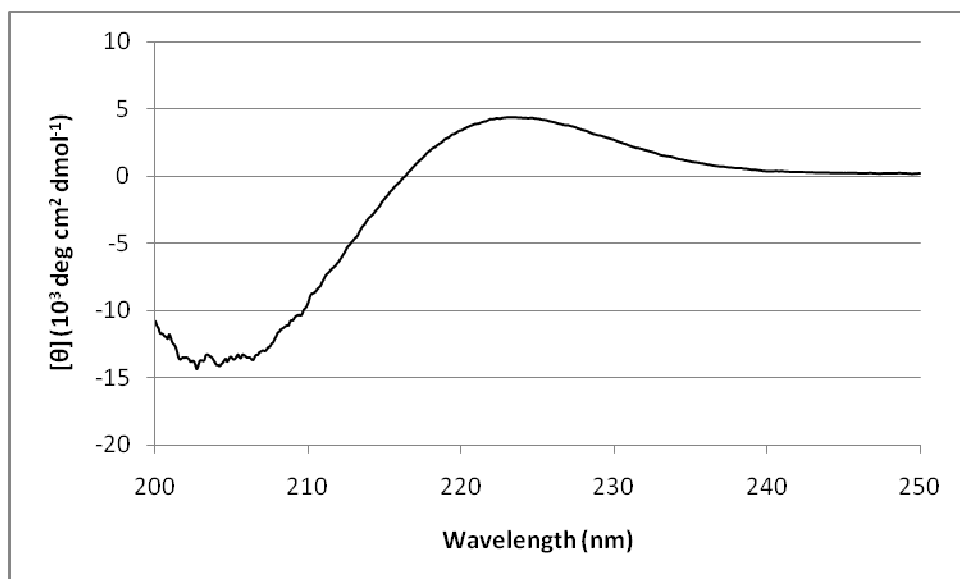
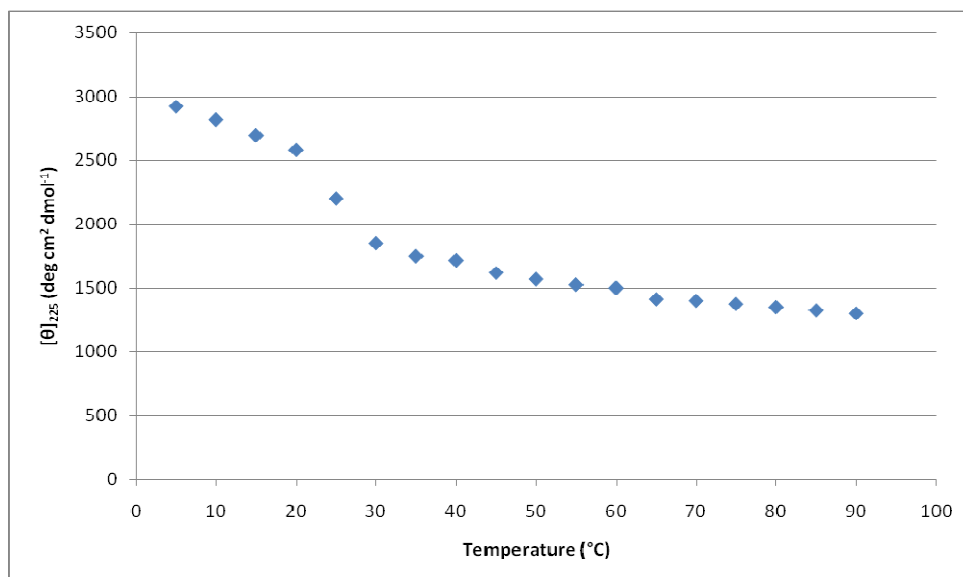


Table 13.1: Ellipticity (millidegrees) at 225 nm as a function of temperature for **1-3**^[a]

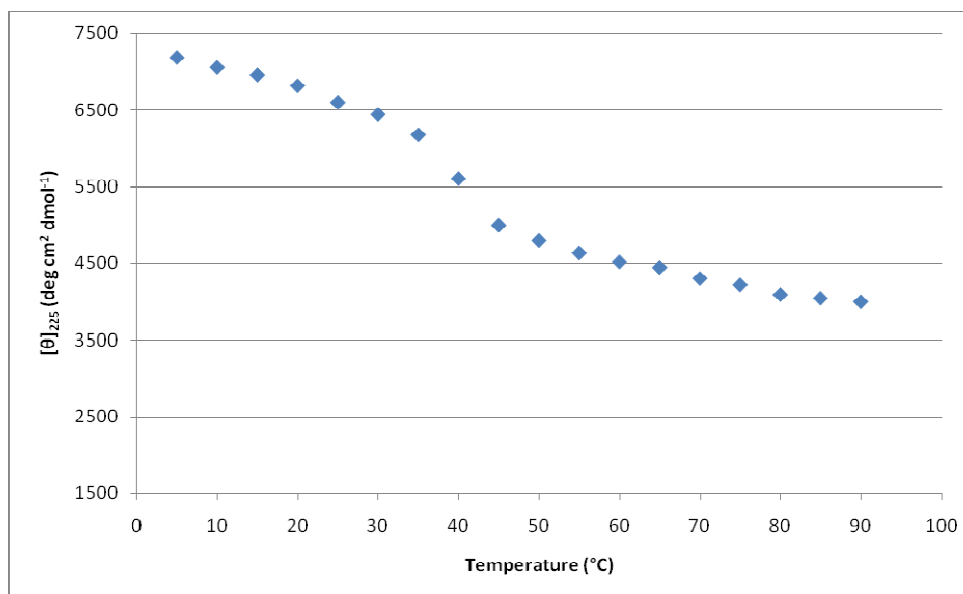
Temperature (±1 °C)	Compound		
	1	2	3
5	7.79 ± 0.25	38.69 ± 0.42	11.57 ± 0.26
10	7.52 ± 0.12	37.99 ± 0.33	11.49 ± 0.18
15	7.19 ± 0.13	37.46 ± 0.27	11.36 ± 0.14
20	6.88 ± 0.03	36.70 ± 0.41	11.31 ± 0.24
25	5.87 ± 0.04	35.52 ± 0.14	11.18 ± 0.19
30	4.93 ± 0.22	34.67 ± 0.25	11.04 ± 0.16
35	4.67 ± 0.13	33.23 ± 0.21	10.87 ± 0.04
40	4.58 ± 0.19	30.14 ± 0.32	10.78 ± 0.15
45	4.31 ± 0.10	26.91 ± 0.29	10.59 ± 0.24
50	4.18 ± 0.11	25.84 ± 0.17	10.25 ± 0.23
55	4.07 ± 0.23	24.95 ± 0.16	10.06 ± 0.21
60	4.00 ± 0.21	24.31 ± 0.25	9.82 ± 0.12
65	3.77 ± 0.05	23.92 ± 0.12	9.67 ± 0.15
70	3.73 ± 0.09	23.14 ± 0.22	9.35 ± 0.09
75	3.67 ± 0.05	22.69 ± 0.27	8.44 ± 0.08
80	3.60 ± 0.04	22.01 ± 0.30	8.02 ± 0.13
85	3.53 ± 0.05	21.75 ± 0.22	7.93 ± 0.05
90	3.47 ± 0.11	21.50 ± 0.15	7.84 ± 0.10
25 ^[b]	4.77 ± 0.22	29.90 ± 0.35	9.48 ± 0.10

^[a]±Standard error from two or more measurements; ^[b]Post

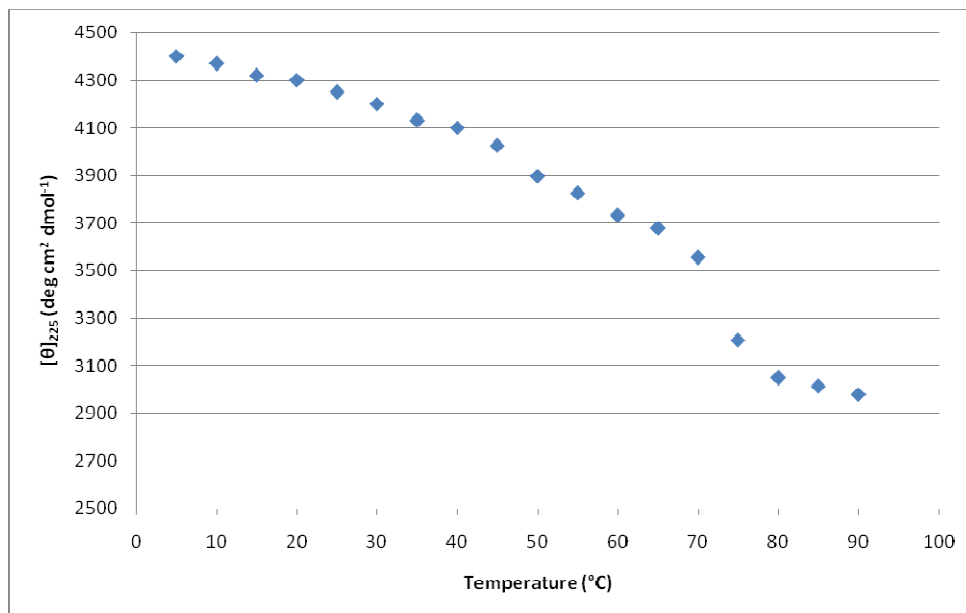
thermal melt experiment after 30 minutes.

Figure 13.10: Thermal melting curve of **1** in H₂O from 5 to 90 °C

T_m calculated to be 25.4 °C as described in section 7.6.

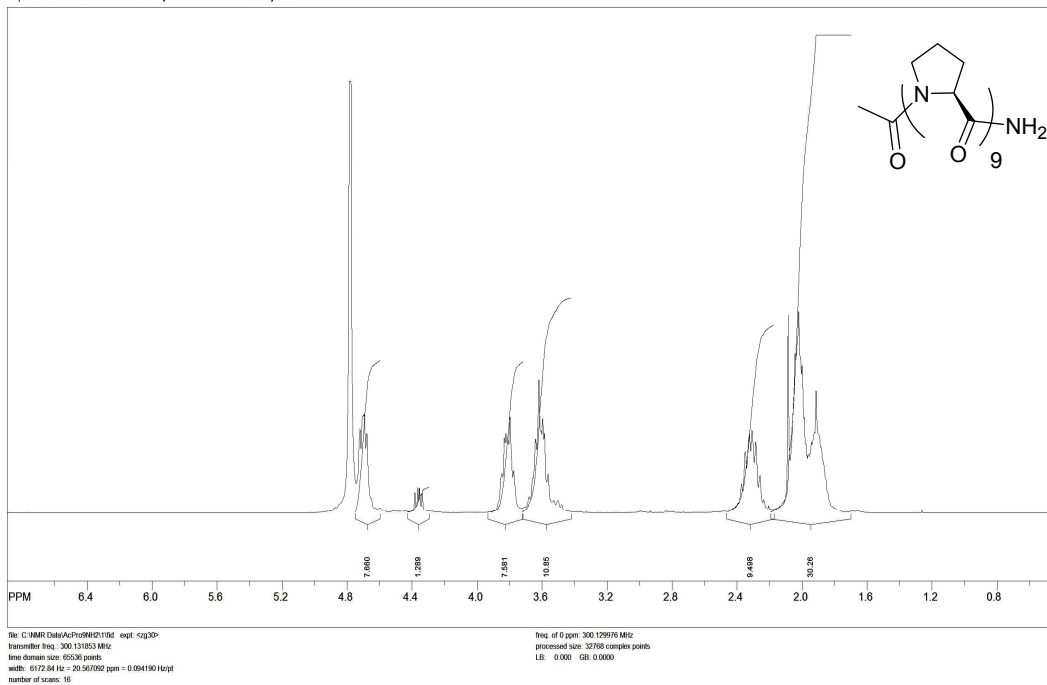
Figure 13.11: Thermal Melting curve of **2** in H₂O from 20 to 90 °C

T_m calculated to be 42.7 °C as described in section 7.6.

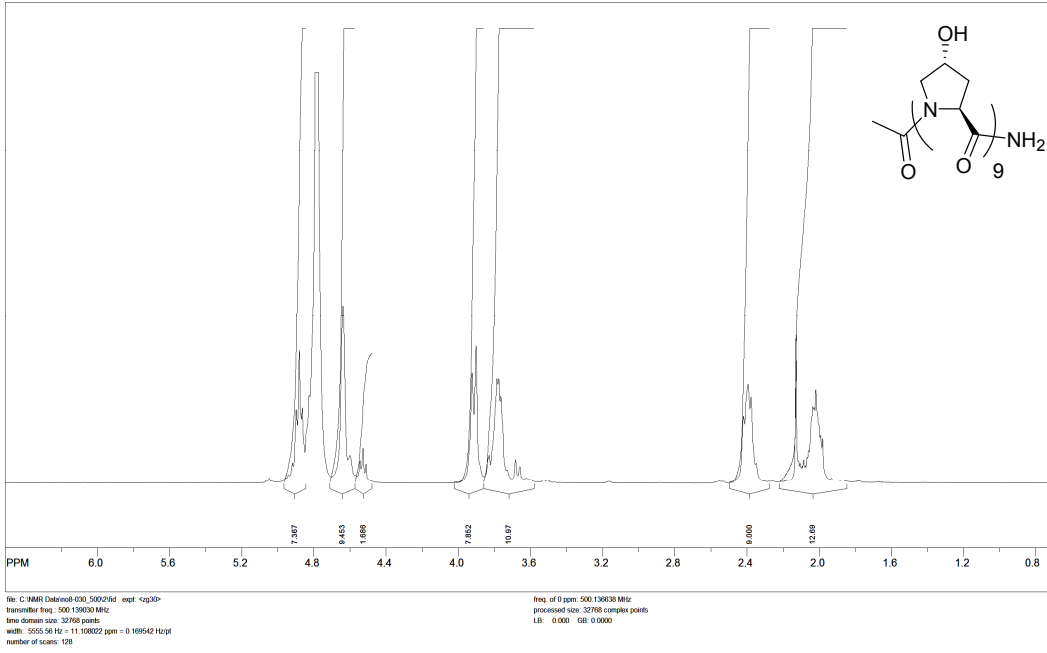
Figure 13.12: Thermal Melting curve of **3** in H₂O from 50 to 90 °C

T_m calculated to be 74.6 °C as described in section 7.6.

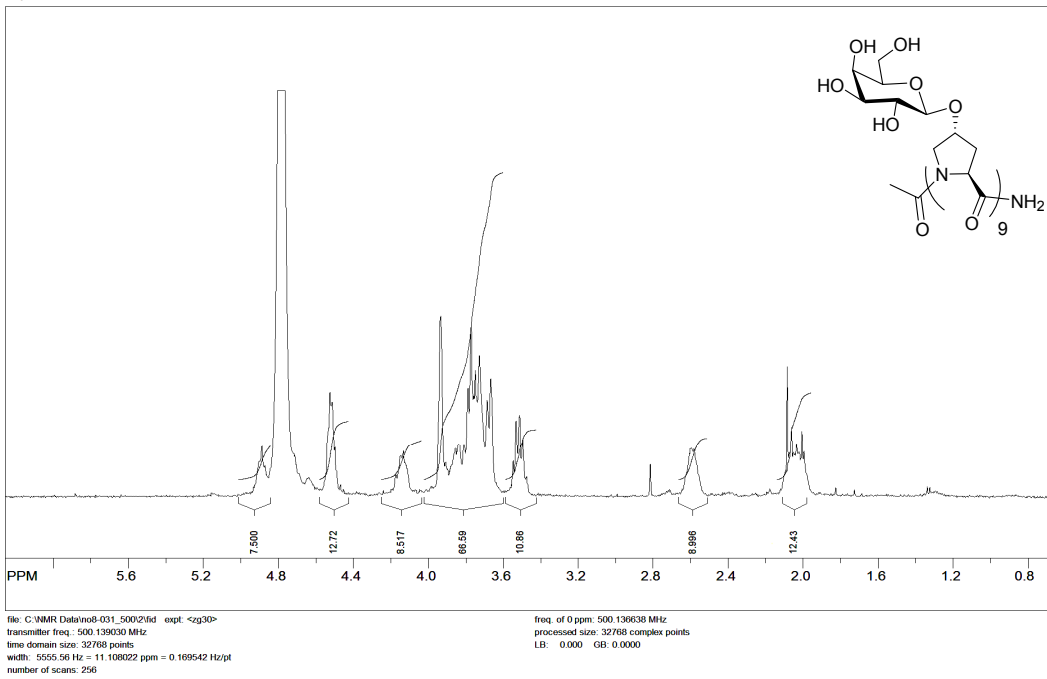
SpinWorks 2.5: PROTON D2O (C:\Bruker\TOPSPIN1.3) schweiz 1



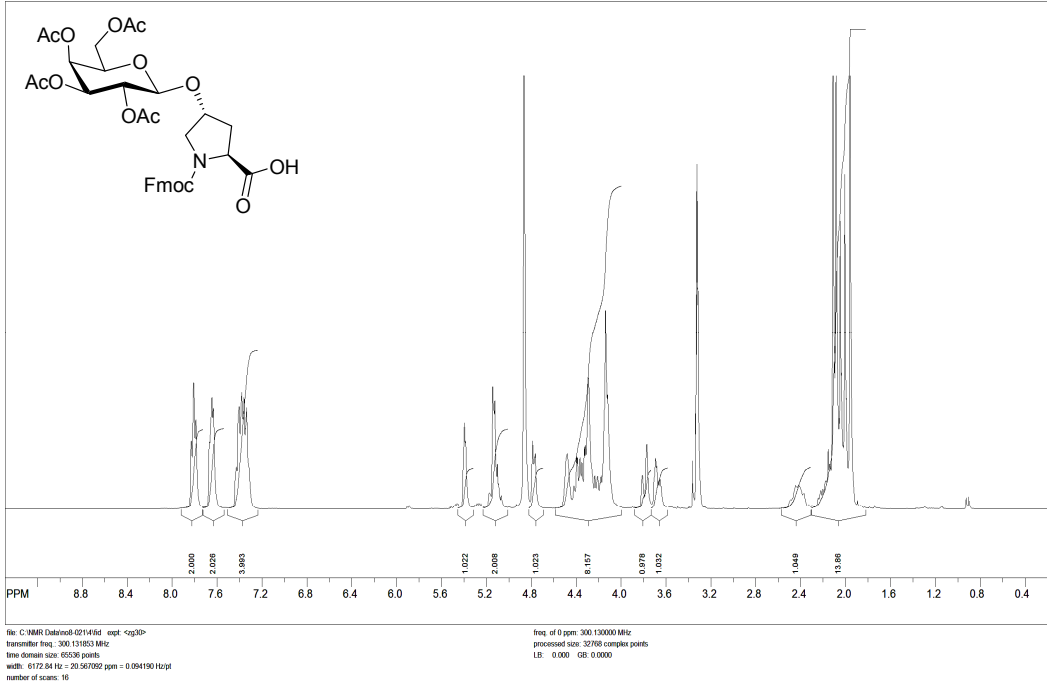
SpinWorks 2.5:



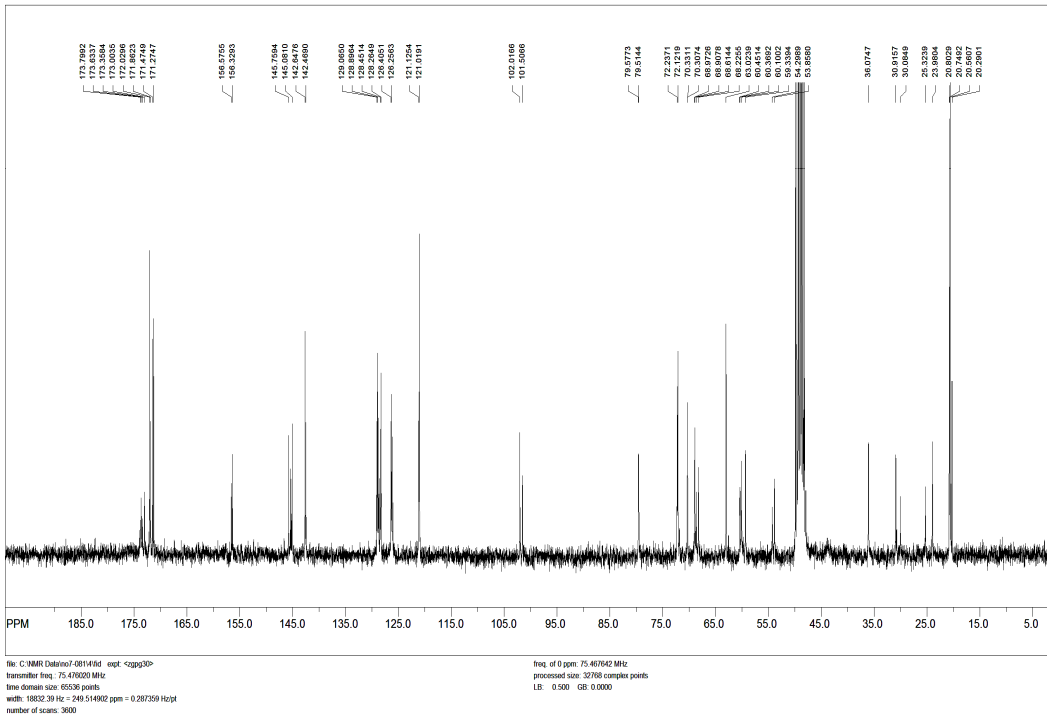
SpinWorks 2.5:



SpinWorks 2.5: PROTON MeOD (C:\Bruker\TOPSPIN1.3) schweiz 1



SpinWorks 2.5: C13CPD MeOH u schweiz 1



List of Publications and Patents Related to Thesis Work

Publications:

1. **Tuning of the Prolyl *Trans/Cis* Amide Rotamer Population Using C-Glucosyl Proline Hybrids** Neil W. Owens, Craig Braun, and Frank Schweizer *Journal of Organic Chemistry* **2007**, 72, 4635-4643.
2. **Effects of Glycosylation of (2*S*,4*R*)-4-Hydroxyproline on the Conformation, Kinetics and Thermodynamics of Prolyl Amide Isomerization** Neil W. Owens, Craig Braun, Joe D. O'Neil, Kirk Marat, and Frank Schweizer *Journal of the American Chemical Society* **2007**, 129, 11670-11671.
3. **The Implications of (2*S*,4*S*)-Hydroxyproline 4-*O*-Glycosylation on Prolyl Amide Isomerization** Neil W. Owens, Adrain Lee, Kirk Marat, and Frank Schweizer (submitted).
4. **The Effects of 4*R*-Hydroxy-L-Proline *O*-Glycosylation on the Stability of Collagen Model Peptides** Neil W. Owens and Frank Schweizer (manuscript in preparation).
5. **Contiguous *O*-Galactosylation of 4*R*-Hydroxy-L-Proline Residues Forms Hyper-Stable Polyproline II Helices** Neil W. Owens and Frank Schweizer (manuscript in preparation).

Patents

1. **Synthesis of Carbohydrate-Templated Amino Acids and Methods of Using Same** Frank Schweizer, Kaidong Zhang, Neil Owens and George Zhanel PCT/US2007/078798.