

Structural Studies of Saccharides and Glycopeptides in Aqueous Solution by ^1H NMR Spectroscopy

Lill Kindahl
Department of Chemistry
Uppsala

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2003

Acta Universitatis Agriculturae Sueciae
Agraria 377

ISSN 1401-6249
ISBN 91-576-6414-5
© 2003 Lill Kindahl, Uppsala
Tryck: SLU Service/Repro, Uppsala 2003

Abstract

Kindahl, L. 2003, *Structural Studies of Saccharides and Glycopeptides in Aqueous Solution by ^1H NMR Spectroscopy*. Doctor's dissertation.
ISSN 1401-6249, ISBN 91-576-6414-5.

The first part of this thesis describes the use of hydroxy protons for ^1H NMR conformational studies of saccharides and small glycopeptide performed in aqueous solution. The conformations of the disaccharide $\beta\text{-D-GlcpNAc-(1}\rightarrow\text{4)-}\beta\text{-D-GlcpNAc}$ and of the glycoside $\beta\text{-D-Galp-(1}\rightarrow\text{3)-}\alpha\text{-D-GalpNAc-O-Me}$ have been investigated and compared to those of the amino acid linked counterparts $\beta\text{-D-GlcpNAc-(1}\rightarrow\text{4)-}\beta\text{-D-GlcpNAc-N-Asn}$ and $\beta\text{-D-Galp-(1}\rightarrow\text{3)-}\alpha\text{-D-GalpNAc-O-Ser}$. For this, the hydroxy proton chemical shifts, vicinal coupling constants, temperature coefficients, exchange rates with water and NOEs were measured. The V[$\beta\text{-D-Galp-(1}\rightarrow\text{3)-}\alpha\text{-D-GalpNAc-(1}\rightarrow\text{)]$ THPGY glycopeptide was also investigated. Information about hydrogen bonding interactions and hydration could be obtained.

The second part of this thesis describes the ^1H NMR studies of the solution conformation for the conotoxin contulakin-G and five analogues. The five analogues had different biological activities, all being less active than contulakin-G. The conformational studies were performed in an attempt to correlate structure to activity. Contulakin-G is a 16 amino acid O-glycosylated glycopeptide, which originally was isolated from the venom of the Cone snail *Conus geographus*. It has the amino acid sequence ZSEEGGSNAT*KKPYIL, with the disaccharide $\beta\text{-D-Galp-(1}\rightarrow\text{3)-}\alpha\text{-D-GalpNAc}$ attached to the threonine residue in position 10. This compound has entered phase II clinical trials for short-term management of post-operative pain. The five analogues are the non-glycosylated peptide, a glycopeptide with the monosaccharide $\alpha\text{-D-GalpNAc}$ attached to Thr10, another glycopeptide with the disaccharide attached at Ser7 and two enantiomeric analogues where the disaccharide was attached at the L-Ser10 and D-Ser10 residues, respectively. The NMR studies showed that, in all compounds, the peptide predominantly existed in extended conformations. Transient populations of folded conformations were found in the glycosylated peptides. The two most active compounds, contulakin-G, and the [D-Ser10] glycosylated analogue, displayed some similar conformational features.

Keywords: hydroxy proton, contulakin-G, conotoxin, conformational analysis, hydrogen bond, O-glycosylation

Author's address: Lill Kindahl, Department of Chemistry, SLU, P.O. Box 7015,
SE-750 07 UPPSALA, Sweden. E-mail: lill@kindahl.net

Till min familj

Contents

List of papers	ix
Abbreviations	x
Amino acid abbreviations	x
Introduction and objectives	xi
 Part 1. ¹H NMR studies on hydroxy protons in conformational analysis of saccharides and glycopeptides (Papers I and II)	
1.1. Introduction	1
1.2. Conformational analysis of oligosaccharides and glycopeptides	3
1.2.1. <i>Techniques used in conformational analysis</i>	3
1.2.2. <i>NMR in conformational analysis</i>	3
1.3. Why use hydroxy protons as conformational probes?	5
1.4. Practical considerations when studying hydroxy protons	7
1.4.1. <i>Water disturbs signal detection</i>	7
1.4.2. <i>Water suppression</i>	7
1.4.3. <i>Sample preparation</i>	9
1.5. ¹ H NMR methods used in papers I and II	10
1.5.1. <i>Proton exchange rates</i>	10
1.5.2. <i>Coupling constants of hydroxy protons signals</i>	10
1.5.3. <i>Temperature coefficients of hydroxy proton signals</i>	11
1.5.4. <i>Chemical shifts of hydroxy proton signals</i>	11
1.5.5. <i>NOEs of exchangeable protons</i>	11
1.6. Present work: papers I and II	12
1.6.1. <i>Hydroxy protons in disaccharides linked to amino acids (Paper I)</i>	12
1.6.2. <i>Hydroxy protons in a glycopeptide (Paper II)</i>	15
1.6.3. <i>Conclusion</i>	18
 Part 2. ¹H NMR studies on the solution conformation of contulakin-G and analogues (Papers III and IV)	
2.1. Introduction	21
2.2. The conotoxin contulakin-G and analogues	23
2.2.1. <i>Isolation and initial characterisation</i>	23
2.2.2. <i>Characterisation by the sluggishness-loss of motor control assay</i>	23
2.2.3. <i>Characterisation by the tail flick-acute pain assay</i>	24
2.3. The solution conformation of glycopeptides	25

2.4. Experimental methods used in papers III and IV	26
2.4.1. <i>NOEs in determination of peptide structure</i>	26
2.4.2. <i>Chemical shifts in determination of peptide structure</i>	27
2.4.3. <i>Temperature coefficients of amide protons</i>	28
2.4.4. <i>Coupling constants of amide protons</i>	28
2.4.5. <i>Determination of absolute conformation</i>	28
2.5. Present work: contulakin-G and three analogues (Paper III)	29
2.5.1. <i>A trial to study conformations in TFE/water solution</i>	32
2.5.2. <i>Conclusion</i>	32
2.6. A study of proteolytic degradation	32
2.7. Present work: the [D-Ser10] and [L-Ser10] analogues (Paper IV)	33
2.7.1. <i>Conclusion</i>	38
Concluding remarks	41
Acknowledgement	43
References	45

List of papers

Papers I-IV

The present thesis is based on the following papers, which will be referred to by their Roman numerals:

- I. Kindahl, L.; Sandström, C.; Norberg, T.; Kenne, L.; ^1H NMR studies of hydroxy protons of Asn- and Ser-linked disaccharides in aqueous solution, *J. Carbohydr. Chem.* **2000**, *19*, 1291-1303.
- II. Kindahl, L.; Sandström, C.; Norberg, T.; Kenne, L.; ^1H NMR studies of hydroxy protons of the V[β -Gal(1 \rightarrow 3)- α -GalNAc(1 \rightarrow O)]THPGY glycopeptide, *Carbohydr. Res.* **2001**, *336*, 319-323.
- III. Kindahl, L.; Sandström, C.; Craig, A. G.; Norberg, T.; Kenne, L.; ^1H NMR studies on the solution conformation of Contulakin-G and analogues, *Can. J. Chem.* **2002**, *80*, 1022-1031.
- IV. Kindahl, L.; Sandström, C.; Craig, A. G.; Norberg, T.; Kenne, L.; Structural studies by ^1H NMR of the [L-Ser10] and [D-Ser10] analogues of contulakin-G, *manuscript*.

Paper I was reprinted by courtesy of Marcel Dekker, Inc., Paper II with permission from Elsevier Science, and Paper III with permission from NRC Research Press.

Abbreviations

1D, 2D, 3D	One dimensional, two dimensional, three dimensional
^1H , ^2H	Proton, deuterium. The number denotes the isotope.
^{13}C	Carbon. The number denotes the isotope.
$^3J_{\text{Hx,Hy}}$	Vicinal coupling constant between protons x and y
Å	Ångström
ADC	Analog to digital converter
C α , C β , C γ	Carbon. The α , β , and γ identifies the atom.
COSY	Correlation spectroscopy
CSI	Chemical shift index
DMSO	Dimethyl sulphoxide
DQF-COSY	Double quantum filtered COSY
D-Xaa	D-configuration of amino acid Xaa
f_1 , f_2	The axes in 2D NMR spectra
G	Gradient (NMR) or <i>geographus</i> (in the name contulakin-G)
GC	Gas chromatography
Galp	Galactopyranose
GalpNAc	N-Acetyl galactosamine, pyranose form
GlcP	Glucopyranose
GlcPNAc	N-Acetyl glucosamine, pyranose form
H α , H β , H γ	Proton. The α , β , and γ identifies the atom.
HMBC	Heteronuclear multiple bond correlation
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum coherence
L-Xaa	L-configuration of amino acid Xaa
MALDI-TOF	Matrix-assisted laser desorption/ionisation-time of flight
Manp	Mannopyranose
MS	Mass spectrometry
NMR	Nuclear Magnetic Resonance
NOE	Nuclear Overhauser enhancement
NOESY	Nuclear Overhauser enhancement spectroscopy
NSS	Neutral saline solution
NTR	Neurotensin receptor
PFG	Pulsed field gradients
ROESY	Rotating frame Overhauser enhancement spectroscopy
RP	Reverse phase
TFE	2,2,2-trifluoroethanol
t_1 , t_2	Time delays (consecutively numbered)
t_1 -noise	Bands of noise in 2D spectra parallel to the f_1 axis
TLC	Thin layer chromatography
t_m	Mixing time
TOCSY	Total correlation spectroscopy
WATERGATE	Water suppression through gradient tailored excitation
X, Xaa	Any amino acid; a variable

Amino acid abbreviations

A	Ala	Alanine	M	Met	Methionine
C	Cys	Cysteine	N	Asn	Asparagine
D	Asp	Aspartic Acid	P	Pro	Proline
E	Glu	Glutamic Acid	Q	Gln	Glutamine
F	Phe	Phenylalanine	R	Arg	Arginine
G	Gly	Glycine	S	Ser	Serine
H	His	Histidine	T	Thr	Threonine
I	Ile	Isoleucine	V	Val	Valine
K	Lys	Lysine	W	Trp	Tryptophan
L	Leu	Leucine	Y	Tyr	Tyrosine

Introduction and objectives

This thesis is concerned with a study of the three-dimensional characteristics of some oligosaccharides and glycopeptides on the molecular structure level. Oligosaccharides are small- to medium-sized carbohydrate polymers that can be found on the surface of cells, conjugated to proteins or lipids. These compounds have a biological role in *e.g.* immunological recognition and protein stabilisation. An example of carbohydrates as immunodeterminants is found in the A, B and O blood groups, where the labels refer to different oligosaccharides that determine the immunological properties. A glycopeptide consists of a peptide, which is structurally similar to a protein but smaller and linked to an oligosaccharide. In research labs, glycopeptides are often encountered as synthetically produced molecules that serve as models for the larger glycoproteins, however glycopeptides can also be produced by nature.

The experimental data of the thesis were acquired by proton nuclear magnetic resonance (^1H NMR) spectroscopy. More than fifty years ago, the first observations of nuclear magnetic resonance were made. During these years NMR has become the principal technique for structural analysis in chemistry. The tremendous development of NMR is reflected in the Nobel Prize awards of 1952 (Felix Bloch, and E.M. Purcell, Physics award), 1991 (Richard R. Ernst, Chemistry award) and 2002 (Kurt Wütrich, Chemistry award). NMR can be used to determine the molecular structure, the distribution in space (conformation), and the mobility (dynamics) of a compound. Also, NMR can be used to detect inter- and intramolecular interactions of a compound.

The first part of this thesis is devoted to the use of hydroxy protons in conformational analysis of saccharides and glycopeptides. It is part of a larger study, where the NMR characteristics of hydroxy proton are investigated. In carbohydrates, about half of all protons are hydroxy protons, that is to say protons bound to an oxygen. The protons exchange rapidly with the protons of water in their natural environment. This is troublesome, since protons that exchange rapidly do not give rise to detectable NMR signals, and thus cannot be studied. Since it is desirable to study molecules in an environment that resembles their natural environment, it is a drawback that only half of the protons are detectable. This can, however, be overcome by adjusting the sample conditions. How this is done is explained later in this thesis.

The second part of this thesis deals with ^1H NMR studies of the solution conformation of the conotoxin contulakin-G and analogues. The conformations of contulakin-G and analogues were investigated, and an attempt was made to relate the conformations to biological properties. Contulakin-G is a glycopeptide toxin isolated from the marine snail *Conus geographus*. Each *Conus* species produces a mixture of different compounds in a venomous cocktail, which is used for predation, although glycosylation is unusual among these compounds. This work is part of a larger project, where the significance of the glycosylation of the glycopeptide contulakin-G has been investigated by a combination of synthesis, NMR, biochemical, and biological methods.

The general aims of this thesis were to develop additional ^1H NMR methods to study saccharides and glycopeptides, and if possible validate the observations with other methods. If the three-dimensional structures and dynamical properties could be better characterized, an improved understanding of the recognition processes these compounds are involved in could be achieved. Also, the aim was to gain knowledge of the importance of glycopeptide conformation to biological activity. The hydroxy protons in saccharides and glycopeptides are not well studied, and the desire was to improve the understanding of what role the hydroxy protons play in oligosaccharides and glycopeptides.

The specific aims in part 1 were, as part of a broad ongoing study on the use of hydroxy protons of saccharides in structural and conformational analysis, to,

- Contribute data to determine the utility and of hydroxy protons as conformational probes.
- Contribute data to the investigation of the causes to the chemical shift value of a certain hydroxy proton.
- Develop a method to monitor the hydroxy protons of glycopeptides in aqueous solution.
- Extend the investigation of hydroxy protons in saccharides to glycopeptides.

The specific aims in part 2 were further,

- To investigate the impact of glycosylation to peptide and glycan conformation.
- To investigate if conformational features in glycopeptides can be correlated to biochemical and biological activity.

Part 1

¹H NMR studies on hydroxy protons in conformational analysis of saccharides and glycopeptides (Papers I and II)

Many molecules in biological systems, for example oligosaccharides, are dissolved in water in their natural environment. It is therefore desirable to study these biological molecules in an environment that resembles their natural environment. Hydroxy protons are difficult to observe by ¹H NMR in water solution, since they are exchanged too rapidly with the bulk solvent. The exchange rate can, however, be reduced by lowering the temperature and optimising sample conditions so that the hydroxy protons can be observed. Experimental data on hydroxy protons in a water solution are nonetheless scarce. Therefore, the studies in papers I and II have been performed in order to gain more information on the utility of hydroxy protons in ¹H NMR studies.

1.1. Introduction

Carbohydrates are, together with nucleic acids and proteins, an important family of biomolecules. They occur in many different systems, and exist as storage polysaccharides, for example as starch in plants and glycogen in animals or as structural polysaccharides such as cellulose and chitin.¹ Starch and glycogen are polymers of α -D-glucose while cellulose is a linear polymer of β -D-glucose. In biological systems, oligosaccharides are normally conjugated to proteins or lipids.²⁻⁶ These conjugates are important to biological recognition processes involving cell-cell interaction such as fertilization, cell growth, inflammatory processes, bacterial infection, circulatory clearance etc.^{2-5,7,8} The significance of oligosaccharides in biological systems has been explored through methods such as enzymatic or chemical removal of the oligosaccharide, prevention of glycosylation, alteration of oligosaccharide processing, and elimination of specific glycosylation sites.⁹ Also, the study of natural variants and genetic mutants has been utilised to investigate the importance of oligosaccharides.⁹

Many proteins and lipids in biological systems are glycosylated. The heterogeneity and branching of oligosaccharides allow glycoconjugates to display a further level of structural and functional diversity compared with linear biomolecules for instance lipids, nucleic acids and proteins.¹ Differences in saccharide sequence, chain length, anomeric configuration (α or β), position of linkages and branching points as well as covalent attachment of sulphate, phosphate, acetyl or methyl groups contribute to this great structural variation.

Based on the type of linkage between the oligosaccharide and the protein, most oligosaccharide moieties of glycoproteins fall in one of two major classes: *N*-glycans where the oligosaccharide is linked to the side chain amide group of an

asparagine residue through an *N*-glycosyl bond, and *O*-glycans where the oligosaccharide is linked to the side chain hydroxy group of a serine or threonine residue through an *O*-glycosyl bond. The main types of *N*-glycans (oligomannose, complex-, and hybrid-type) can be generated by enzymatic extension of the processed precursor. They have in common the core structure Man₃GlcNAc₂ (Figure 1).⁶ High-mannose-type *N*-glycans are common products of yeasts, but are also present in animal glycoproteins such as ovalbumin. *N*-glycosylation is diverse and common, with extensive descriptions that can be found elsewhere.^{2,3}

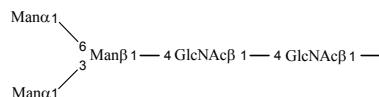


Figure 1. The pentasaccharide core common to all *N*-linked glycans.

For the *O*-glycans, elongation of the reducing end GalNAc has been found to generate eight different core structures (Table 1). These core structures may be further elongated and modified.^{4,10}

Table 1. The Ser/Thr linked oligosaccharide core region sequences of serum, cell membrane and mucin glycoproteins (adopted from Hounsell, E. F.; Davies, M. J.; Renouf, D. V., *Glycoconjugate J.* **1996**, *13*, 19-26, and Hanisch, F. A. *Biol. Chem.* **2001**, *382*, 143-149).^{4,10}

Sequence	Core type ^a
Galβ1-3GalNAcα1- GlcNAcβ1	1 2
<div style="margin-left: 40px;"> \ <div style="margin-left: 20px;">6</div> <div style="margin-left: 20px;">GalNAcα1-</div> <div style="margin-left: 20px;">3</div> / </div>	
<div style="margin-left: 40px;">Galβ1</div>	
GlcNAcβ1-3GalNAcα1- GlcNAcβ1	3 4
<div style="margin-left: 40px;"> \ <div style="margin-left: 20px;">6</div> <div style="margin-left: 20px;">GalNAcα1-</div> <div style="margin-left: 20px;">3</div> / </div>	
GlcNAcβ1	
GalNAcα1-3GalNAcα1-	5
GlcNAcβ1-6GalNAcα1-	6
GalNAcα1-6GalNAcα1-	7
Galα1-3GalNAcα1-	8

^a. For a complete reference listing of the primary sources of these core structures, see the original publications.^{4,10}

Glycosylation based on the linkage α -GalpNAc(1 \rightarrow O)-Ser/Thr is found on serum and cell membrane glycoproteins and high molecular weight mucins, which line the gastrointestinal tract and bronchial airways.^{4,5,10} Other types of *O*-glycosylation are rare or restricted to certain species, tissues or proteins. These include mainly the *O*-linked fucose and *O*-linked glucose sugars found in the epidermal growth factor (EGF) domains of different proteins, and *O*-linked GlcpNAc on cytosolic and nuclear proteins.^{5,11}

In order for *N*-glycosylation to take place, a peptide sequence of Asn-Xaa-Ser/Thr (a sequon) is required, where Xaa refers to any amino acid with the exception of Pro.⁶ No particular peptide sequence requirements have been described so far for classical mucin-type *O*-glycosylation, and the rules by which occupied *O*-glycosylation sites can be predicted are only beginning to be understood.^{12,13} Glycosylation is species- and cell specific.¹⁴ It has been discovered that the efficiency of glycosylation is decreased or hindered by factors such as the amino acid sequences at or near the putative glycosylation site, the availability of glycosyltransferases, and stable secondary and tertiary structures.^{2,3,5,6,10}

Oligosaccharides can change the physical characteristics of the protein to which they are attached by modifying the local structure and overall dynamics.^{6,15} *N*-linked glycosylation occurs co-translationally,¹⁶ and can therefore have impact on the protein folding process. Both *N*- and *O*-glycosylation can stabilise structures in the protein.^{5,15} Extensive *O*-linked glycosylation of repetitive serine and threonine residues elongates the peptide structure and results in a filamentous protein, while sections of peptide with a high Pro-content as well as many sugars attached to serines or threonines may assume a 'bottle-brush'-like structure.⁵ Extensive *O*-glycosylation has also been found to result in protease resistance, which is important since this can prolong the lifetime of the protein in biological systems.⁵

1.2. Conformational analysis of oligosaccharides and glycopeptides

Determination of the three-dimensional structures and the dynamical properties of oligosaccharides is a prerequisite for a better understanding of recognition processes. For biomolecules such as enzymes, the arrangement of atoms in space is crucial for the recognition of substrates. One important way of investigating the significance of carbohydrate solution conformation to biological recognition processes is to compare conformations of a carbohydrate in its free and bound states, respectively.¹⁷

1.2.1. Techniques used in conformational studies

To determine the conformation and dynamics of oligosaccharides and glycoproteins, complementary data from NMR spectroscopy, X-ray crystallography and molecular modelling is often used.^{18,19} NMR has been the main experimental technique used to determine conformations of carbohydrates. However, since it only provides time-averaged conformational data, this has led to

the use of molecular modelling as an aid in determining linkage conformations.^{20,21} By using X-ray crystallography, complete oligosaccharide conformations can be obtained, but since the glycan moiety is often poorly resolved, only one or two carbohydrate residues can generally be seen in glycoprotein crystal structures. The conformation that is obtained by X-ray crystallography also represents only one of the many conformations possible in solution. The main limitation, however, is that few oligosaccharides in their native form crystallise.^{21,22} The reason to this is not obvious, but it could be due to the high polarity and/or inherent flexibility of most oligosaccharides. Glycans that are covalently attached to proteins or ligands to proteins form crystals easier than underivatized glycans, since the formation of crystals is dominated by protein interactions.

1.2.2. NMR in conformational analysis

NMR has been a major technique for structural determination of oligosaccharides, being the only technique that can determine both the anomeric configuration and linkages of a novel glycan.^{23,24} There are a number of different NMR techniques developed for structural determination. The assignment of ^1H and ^{13}C chemical shifts is often achieved using the two dimensional ^1H - ^1H COSY, ^1H - ^{13}C HTOCSY, ^1H - ^1H NOESY, ^1H - ^{13}C HSQC and ^1H - ^{13}C HMBC techniques.²³ NMR has also been the main experimental technique for investigating the conformation of oligosaccharides. The conformation of a molecule is the arrangements of its atoms in space. In the case of carbohydrates, the global shape or conformation depends mainly on rotations around the glycosidic linkages. The relative orientations of saccharide units are expressed in terms of glycosidic linkage torsional angles ϕ and ψ , which are defined as H1-C1-O1-CX and C1-O1-CX-HX, respectively (Figure 2). Many saccharides also have an exocyclic primary alcohol function at the 6-position, and the corresponding torsion angle ω is given by O5-C5-C6-O6 (Figure 2). It is the non-restricted glycosidic bonds that render the carbohydrates highly flexible.

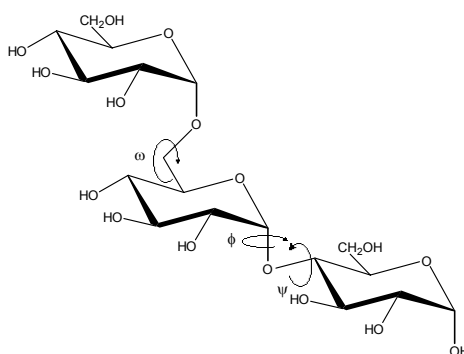


Figure 2. The ϕ , ψ , and ω angles of rotation around the glycosidic bonds in a saccharide.

An alternative to determine the torsion angles is to determine the conformation by measuring distances between atoms on each side of the glycosidic linkage,

since rotation around the linkage will change the distance between atoms. These data, torsion angles and distances can be obtained by NMR spectroscopy. Scalar coupling constants (3J -values) can be used to give information on the values of the torsion angles. The relation between a certain 3J -value and the corresponding torsion angle is given by the Karplus curve.^{25,26} However, some 3J -values may correspond to several different torsion angles. Determination of ϕ and ψ requires the measurement of three-bond carbon-proton coupling constants ($^3J_{CH}$). To determine ω in 1,6-linkages three-bond proton-proton couplings ($^3J_{HH}$) can be used.²⁷

The major source of conformational information comes from the nuclear Overhauser effect (NOE or ROE), which gives information about protons that are close in space.²⁸ The magnitude of the NOE between two nuclei depends on the distance between them ($\propto r^{-6}$). The NOE is short range (up to about 5 Å), and is usually only observed between nuclei within a monosaccharide or across a glycosidic linkage. The NOEs are measured on a 50 ms to 1 s time scale. A flexible molecule such as an oligosaccharide will convert between several conformations during this time, and thus only average NOEs will be observed. Usually at least one and not more than three NOEs are measurable across the glycosidic linkage, and two interresidual three-bond heteronuclear ^{13}C - 1H spin-spin coupling constants can be determined. In the case of 1,4-linkages, this would be between atoms H1' and C4, and between C1' and H4, respectively. The number of conformational constraints provided by the NOEs and spin-spin coupling constants are not sufficient to define the conformation around the glycosidic bond. This means that the conformation of saccharides is generally not well defined due to a lack of constraints.

New approaches to extract saccharide conformational data by NMR include the measurement of trans-glycosidic heteronuclear NOEs involving ^{13}C and 1H ,²⁹ the measurements of trans-glycosidic 1H - ^{13}C and ^{13}C - ^{13}C coupling constants,^{26,30-32} and the use of residual dipolar couplings.³³ However, the measurement of ^{13}C - ^{13}C coupling constants requires labelled compounds that are not always available.

1.3. Why use hydroxy protons as conformational probes in 1H NMR?

Despite their abundance in oligosaccharides, and their potential importance in hydrogen bond interactions with for example enzymes, the exchangeable hydroxy protons have been used very little for conformational analysis in aqueous solution. The lack of data is mainly due to the fact that hydroxy protons are in rapid exchange with the bulk water, making them difficult to observe by NMR. But instrumental limitations, which today are overcome, also lie behind this lack of data. Due to the improvements in experimental techniques and procedures, a picture of the utility of hydroxy protons has started to form.

In the last few years, it has been shown that it is possible to observe hydroxy protons in H_2O solutions.^{34,35} This has been achieved by using low temperatures to reduce the rate of exchange with water, and pulse sequences that efficiently suppress the water signal without affecting the resonances of the exchangeable

protons.³⁶ Structural information in terms of stereochemistry,³⁷ inter-proton distances,³⁸⁻⁴⁵ and hydrogen bonding has been obtained.^{38-41,46-50}

Intramolecular hydrogen bonds may be an important means of stabilising carbohydrate conformations, however, how important is still unclear.⁵¹ In order to evaluate hydrogen bond interactions where hydroxy protons are involved, vicinal coupling constants ($^3J_{\text{CH,OH}}$), exchange rates with water, and temperature dependence of the chemical shifts of hydroxy proton signals are measured.^{38-41,44,48,50} The observation of hydroxy protons has showed that persistent hydrogen bond interactions also can exist in aqueous solution,⁴⁵ and not only in DMSO solutions.⁴⁷ Using NMR, hydroxy protons in the solvent DMSO are readily observed, but then the influence of this aprotic solvent on the conformation should be considered. It has been shown, for example, that strong hydrogen bonds that were found to exist in DMSO solutions do not persist in aqueous solutions.^{40,46} Furthermore, it has been found that different intramolecular bonds can be formed in DMSO as compared with H₂O.⁴⁹ An example of conformational information that can be retrieved from the observation of hydroxy protons is the conformation around the C(5)-C(6) bond.⁴⁵ Here, the NOEs between the exocyclic hydroxy protons and the other exchangeable or non-exchangeable protons, together with the hydroxy proton chemical shift values, allow for the determination of the favoured conformation, even if spectral overlap precludes measurement of the $^3J_{\text{H5,H6}}$ coupling constants.

Deuterium isotope effects on the ^{13}C chemical shifts have also been used as a possible means to determine hydrogen bonding in saccharides in both aqueous solution,⁴⁰ and DMSO.^{52,53} Data obtained in aqueous solution can, however, be difficult to interpret since other carbons than the one substituted with the hydrogen involved in interresidual hydrogen bonds show splitting due to the deuterium isotope effects.

1.4. Practical considerations when studying hydroxy protons

In order to be able to study the hydroxy protons in H₂O solution, some practical concerns have to be solved. Firstly, the water gives rise to an enormous water signal in the NMR machine, which has to be suppressed. Secondly, the sample conditions have to be adjusted in order to minimise the hydrogen exchange, thereby ensuring detectable hydroxy proton peaks.

1.4.1. Water disturbs signal detection

In the NMR tube a sample concentration of 1-20 mM is commonly used, while the concentration of the protons of water is 110 M. This means that the signal from the water peak in the NMR spectrum is ten thousand to one hundred thousand times more intense than the intensity of the signals from the molecule that is investigated. The principal reason to suppress a large solvent resonance in the presence of far smaller solute resonances is to ensure that the dynamic range of the NMR signals lie within the dynamic range of the receiver and the ADC. Further concerns are the baseline distortions, t_1 -noise in 2D experiments, radiation damping and potential spurious responses that are associated with very intense signals. Furthermore, an unsuppressed water peak will overlap resonances that are close, as in the case of hydroxy proton resonances (Figure 3a). In order to diminish the water signal, one can use deuterated water (D₂O) instead of H₂O, since D₂O is not observable by ¹H NMR spectroscopy. However, since the deuterium of D₂O also will exchange with the hydroxy protons of the molecule that is to be investigated, the intensity of the hydroxy protons will also be reduced to zero. This excludes D₂O as a solvent in NMR studies of hydroxy protons, and leaves no other choice than H₂O as a solvent. The solution to the dilemma with the huge water peak lies in the use of NMR water suppression techniques.⁵⁴

1.4.2. Water suppression

Water suppression can be achieved by using water presaturation, but this will result in suppression of not only the water peak, but also of the exchangeable hydroxy proton resonances (Figure 3b).

Pulsed field gradients (PFGs) is an effective approach to solvent suppression. This technique ensures that nothing of the net solvent magnetisation remains observable immediately prior to acquisition. The pulse sequence used in this work is the WATERGATE experiment (water suppression by gradient tailored excitation).³⁶, which makes use of a single PFG spin-echo, G₁-S-G₁, (Figure 4). The result of applying WATERGATE to a sample containing hydroxy protons can be seen in figure 3c, where, the water peak is small and narrow enough not to overlap with the neighbouring resonances from the hydroxy protons, while the intensity of the hydroxy protons is not affected, as in the case of presaturation (Figure 3b).

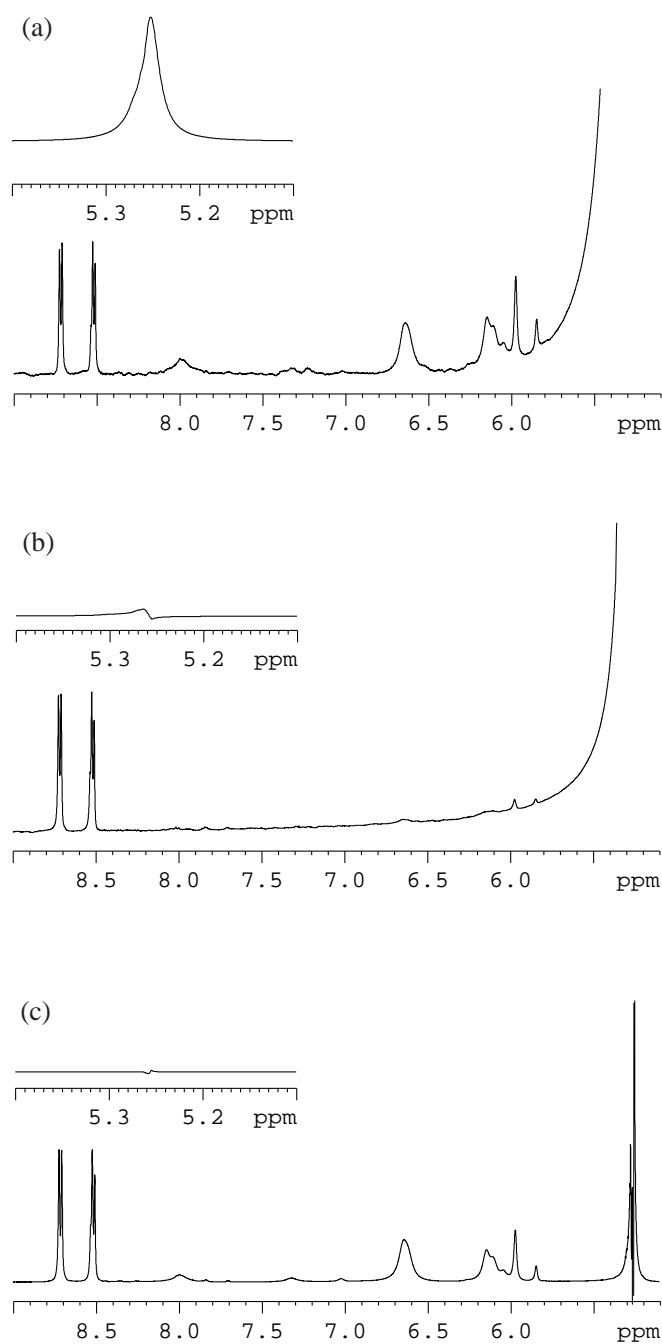


Figure 3. 1D ^1H NMR spectra of chitobiose in 85% H_2O /15% $(\text{CD}_3)_2\text{CO}$ solution, showing the amide, and hydroxy proton ($\delta \sim 5\text{-}7$ ppm) regions together with the water peak region. The inclusion shows the water peak scaled down 512 times. a. No water suppression is applied. b. Presaturation is used for water suppression. c. The WATERGATE pulse sequence is used for water suppression.

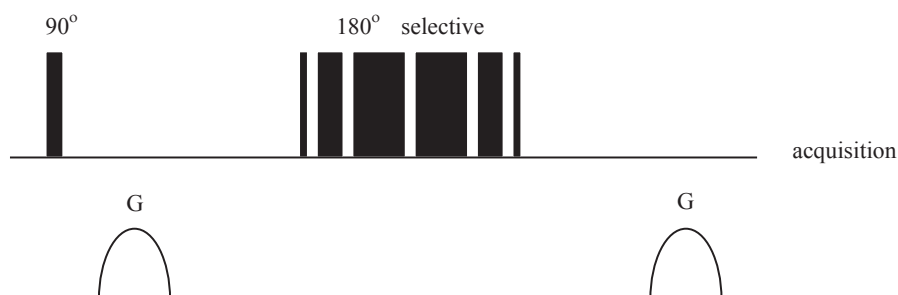


Figure 4. The WATERGATE pulse sequence.

1.4.3. Sample preparation

There is however another problem associated with the use of water as a solvent. At room temperature, the hydroxy protons are normally not observable due to fast exchange with the protons of water. In order to be able to observe the hydroxy protons, the rate of exchange must be reduced. It has been shown that by careful preparation of the NMR samples, it is possible to observe hydroxy groups by NMR. The preparation involves washing the NMR tube with phosphate buffer mainly to get rid of borate ions.⁴⁶ To purify the samples from ionic impurities that catalyse hydrogen exchange, the sample solution can be passed through a bed of ion exchange resins. However, this does not apply to ionic samples that would adhere to the ion-exchange resin, and thereby causing sample loss. Hydrogen exchange can also be catalysed by a high or low pH. To ensure the pH is neutral, either a buffer can be used, or small amounts of a sodium hydroxide and/or hydrogen chloride solution can be added to adjust the pH. A decrease of the sample temperature to sub-zero temperatures can further reduce the rate of exchange. Sub-zero sample temperatures of aqueous samples has been achieved by utilising a delicate cooling procedure,^{48,55} and by using highly concentrated solutions.⁵⁶ The most common strategy to avoid freezing of the sample, however, is to add organic solvent such as acetone to the sample.

1.5. ^1H NMR methods used in papers I and II

1.5.1. Proton exchange rates

Reduced proton exchange rates (k_{ex}) are indicative of involvement in hydrogen bonding or prevention from hydrogen exchange by steric shielding. The 2D exchange experiment proposed by Jeener *et al.* allows the entire matrix of all the exchange processes in the system to be obtained in a single experiment.⁵⁷ In this experiment the NOESY pulse sequence is utilised (Figure 5).

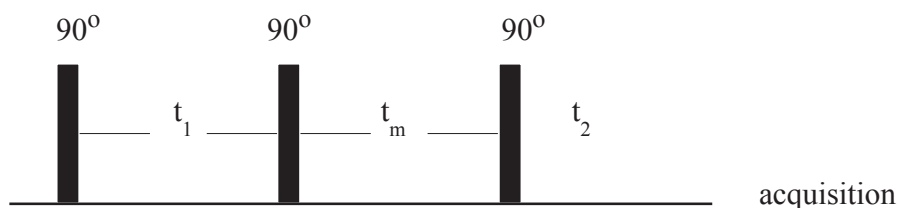


Figure 5. The 2D NOESY pulse sequence

At short mixing times, the relationship between 2D peak intensities and rate constants can be expressed as in equation 1.^{58,59} Here M_j^0 is the equilibrium magnetisation of the nuclei in site j and \mathbf{R} has off diagonal elements $\mathbf{R}_{ij} = -k_{ji}$, where k_{ji} is the first order rate constant for chemical exchange from site j to site i .

$$I_{ij}(t_m) \sim -t_m \mathbf{R}_{ij} M_j^0 = k_{ji} t_m M_j^0 \quad i \neq j \quad (1)$$

Equation 1 permits solving for the rate constants, but it is inherently inaccurate, since all cross-peak intensities are low at short t_m . At longer t_m there is a risk that indirect cross-peaks will be mistaken as evidence of a direct exchange process. A practical solution is to use several short t_m values and determine k_{ij} as the slope of a plot of $I_{ij}(t_m)$ vs $M_j^0 t_m$.

1.5.2. Coupling constants of hydroxy protons signals

Vicinal spin-spin coupling constants ($^3J_{\text{CH,OH}}$) can be measured both in 1D and 2D (DQF-COSY) spectra, and provide useful additional conformational information. The $^3J_{\text{CH,OH}}$ can, together with chemical shift values and proton exchange rates, provide indications whether a hydroxy proton is involved in a hydrogen bond or not. The magnitude of the vicinal $^3J_{\text{CH,OH}}$ coupling constant shows a stereochemical dependence according to the Karplus equation.⁶⁰ Since a hydroxy proton involved in a hydrogen bond will have a restricted rotation, this means that the corresponding $^3J_{\text{CH,OH}}$ coupling constant value will deviate from that of conformational averaging. For example would a $^3J_{\text{CH,OH}}$ -value of 1.8 Hz would indicate a conformational preference for the *syn* orientation, while a value of 10.5 Hz would indicate a preference for the *trans* orientation.⁴⁵ If there is free rotation about the C-O bond connecting the hydroxy group to the ring, the 3J -value should be an average value of *ca.* 5-7 Hz.

1.5.3. Temperature coefficients of hydroxy proton signals

The proton temperature coefficient is a measure of how much the proton chemical shift is changed upon changing the sample temperature ($-\text{d}\delta/\text{d}T$). This can give an indication to whether an exchangeable proton is involved in a hydrogen bond. Exchangeable protons that are involved in hydrogen bonds are expected to have lower temperature coefficients than those that are not. In aqueous solution, hydroxy protons that display a temperature coefficient as low as -1 ppb/ $^{\circ}\text{C}$ have been measured for hydroxy protons involved in hydrogen bonding.³⁸ In DMSO solution, temperature coefficients <-3 ppb/ $^{\circ}\text{C}$ has been measured for OH protons involved in hydrogen bonds.⁴⁷

1.5.4. Chemical shifts of hydroxy proton signals

Hydroxy proton chemical shifts have not been frequently used in conformational analysis. This is not only due to experimental difficulties, but also due to the difficulties in evaluating the significance of a certain value. Hydroxy proton chemical shifts may be sensitive to many different effects such as hydrogen bonding and solvent effects, and these effects have not been well investigated. To evaluate the hydroxy proton chemical shifts of the different disaccharides, the chemical shifts of these compounds were compared to those measured for the corresponding monosaccharide methyl glycosides. The comparison is expressed as the chemical shift of the signals of the hydroxy proton in question minus the chemical shift of the corresponding hydroxy proton signal in the monosaccharide methyl glycoside, or $\Delta\delta$. Hydroxy protons involved in intramolecular hydrogen bonds are normally deshielded, as compared to the corresponding protons in the monosaccharide methyl glycoside.^{44,45,61} However, shielding of the hydroxy proton signals close to non-protonated oxygen atoms, such as ring or glycosidic oxygens, or hydroxy protons situated in crowded regions, were measured.^{44,45,48} This shielding also included hydroxy protons hydrogen bonded to ring protons, which were expected to be shifted downfield due to their hydrogen bonding. Protons experiencing an upfield shift also often showed a slower rate of exchange with water. The upfield shift of hydroxy protons has recently been attributed to solvent effects.⁶²

1.5.5. NOEs of exchangeable protons

The characteristic feature of the 2D NOESY experiment, which utilises the nuclear Overhauser effect, is the mixing time, t_m (Figure 5). Magnetisation transfer that takes place during the mixing time generates the NOE crosspeaks. The NOESY pulse sequence can, however, not distinguish between magnetisation transfer caused by cross-relaxation and chemical exchange, since both kinds of cross-peaks have the same sign as the diagonal cross peaks. However, in the closely related 2D ROESY experiment, cross-peaks caused by cross-relaxation and chemical exchange have opposite signs in the spectra that are generated.⁶³

1.6. Present work: papers I and II

Papers I and II are part of a continuing study on hydroxy protons as conformational probes performed at our department. The aim of the study is to determine the utility and limitations of hydroxy protons as conformational probes. The previous research at our department includes studies of the hydroxy protons of a series of di- and trisaccharides,^{44,45,61,64} as well as the tetrasaccharide Lewis b.⁶⁵

1.6.1. ¹H NMR studies of hydroxy protons in disaccharides linked to amino acids (paper I)

Glycosylation has been found to influence peptide structure.^{66,67} This influence could either be exerted by a direct interaction through, for example, a hydrogen bond, or by indirect interaction, *i.e.* the sugar acts like a bulk group. In paper I, NMR studies of hydroxy protons of asparagine and serine linked disaccharides were performed. This is an extension of previous studies of hydroxy protons in saccharides to amino acid linked saccharides.

The hydroxy protons of the disaccharide β -D-GlcpNAc-(1 \rightarrow 4)- β -D-GlcpNAc, **1**, and the glycoside β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-*O*-Me, **3**, were investigated in aqueous solution by NMR spectroscopy. The chemical shifts, coupling constants, temperature coefficients, exchange rates and NOEs were measured. The data was compared to those of their amino acid linked counterparts β -D-GlcpNAc-(1 \rightarrow 4)- β -D-GlcpNAc-*N*-Asn, **2**, and β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-*O*-Ser, **4**. These compounds were chosen due to their biological importance. The disaccharide β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc is a Ser/Thr-linked core structure in *O*-glycans, while the β -D-GlcpNAc(1 \rightarrow 4)- β -D-GlcpNAc disaccharide is part of the Asn-linked core structure in *N*-glycans. Conformational studies using NMR have already been reported for disaccharides **1**^{68,69} and **3**,^{70,71} and for glycopeptides structurally related to **2**,^{68,69} and **4**.^{72,73} However, no information on the NMR data of the hydroxy protons has been reported for these compounds.

Hydrogen bond interactions between the peptide and sugar moieties have been shown to exist in DMSO solutions of small glycopeptides.⁷⁴ No such interaction was observed in the compounds investigated in paper I, and the presence of one amino acid did not affect the conformation of the disaccharides. These results were not surprising, since the introduction of a single amino acid is not expected to significantly affect the conformational flexibility of the peptide.

A lowering of the temperature down to -12 °C reduced the hydroxy protons rate of exchange with water, and allowed for their observation. To achieve this temperature without freezing of the sample, acetone-*d*₆ was added to the solutions of compounds **1-4** to give a ratio of H₂O:(CD₃)₂CO of 85:15, respectively. The exchange rate (*k*_{ij}) was determined as the slope of a plot of *I*_{ij}(*t*_m) vs. *M*_j⁰*t*_m. A series of 2D NOESY spectra with short mixing times was recorded, and the peak volumes were determined by using the program AURELIA (Bruker, Germany).⁷⁵ The rate plots for the hydroxy proton exchange can be seen below.

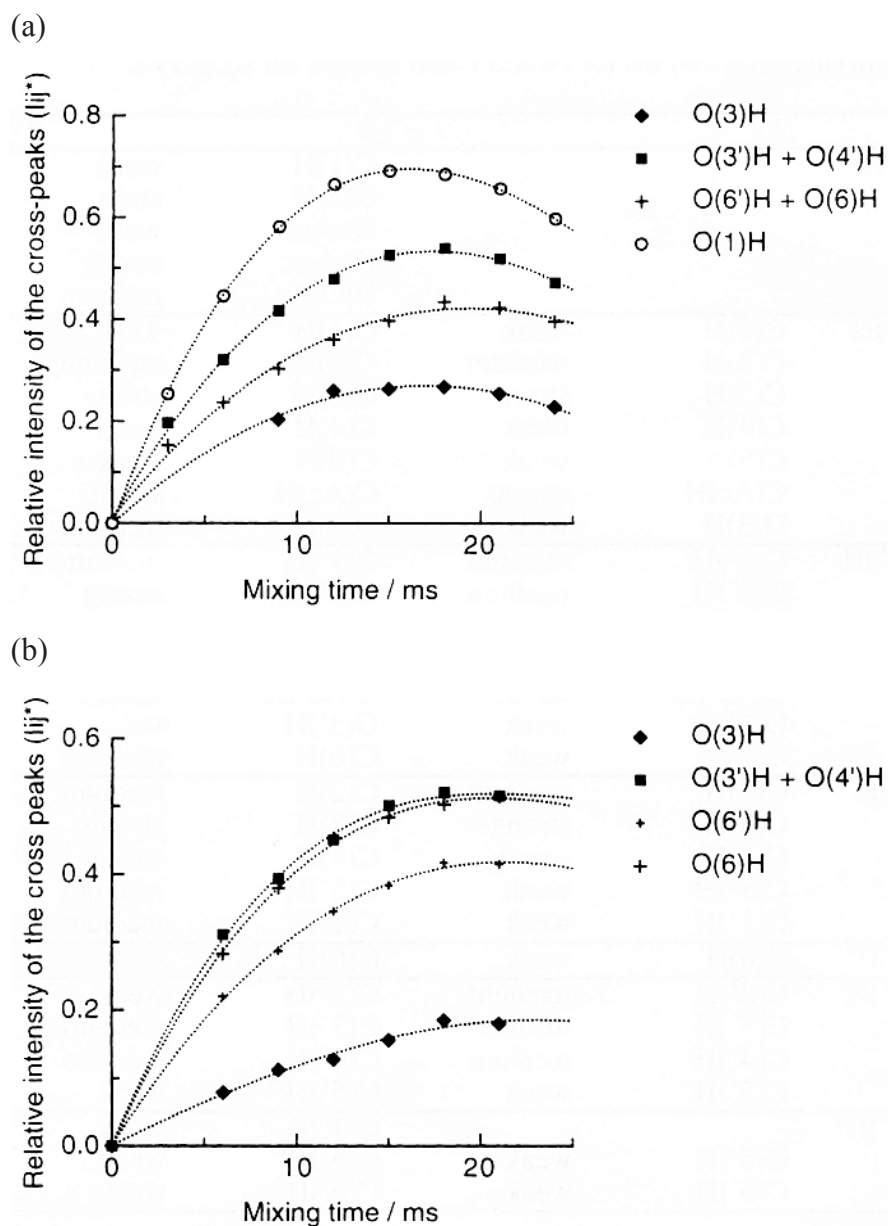


Figure 6. Initial rate constant plots from paper I.

In both β -D-GlcpNAc-(1 \rightarrow 4)- β -D-GlcpNAc, **1**, and β -D-GlcpNAc-(1 \rightarrow 4)- β -D-GlcpNAc-*N*-Asn, **2**, the signals for all protons, with the exception of O(3)H have chemical shifts very similar ($|\Delta\delta| \leq 0.140$ ppm) to those measured for the corresponding monosaccharide methyl glycosides. The O(3)H, on the other hand, is shielded by 0.459 and 0.415 ppm in **1** and **2** (Figure 7). Distance measurements on energy-minimized conformations showed that the O(3)H-O(5')H distance was

short, and the upfield shift was therefore attributed to spatial proximity, in analogy to earlier findings by our research group. The small $^3J_{\text{CH,OH}}$ -values for O(3)H of 2.2 Hz in **1** and 2.6 Hz in **2** (Figure 7), indicate a *gauche* orientation that places the hydroxy proton in a favourable position for hydrogen bonding. This together with the fact that O(3)H in **1** and **2** has the lowest rate of exchange with water suggested the presence of hydrogen bonding between O(5')H and O(3)H. The O(3)H had a temperature coefficient similar to the other hydroxy protons, which suggests that the hydrogen bond interaction with O(5')H is weak. MD studies have shown that this hydrogen bond exists *in vacuo*,⁶⁸ but is replaced by hydrogen bonds with the solvent water molecules in aqueous solution. The data extracted in paper I showed that some transient hydrogen bond interaction must be present.

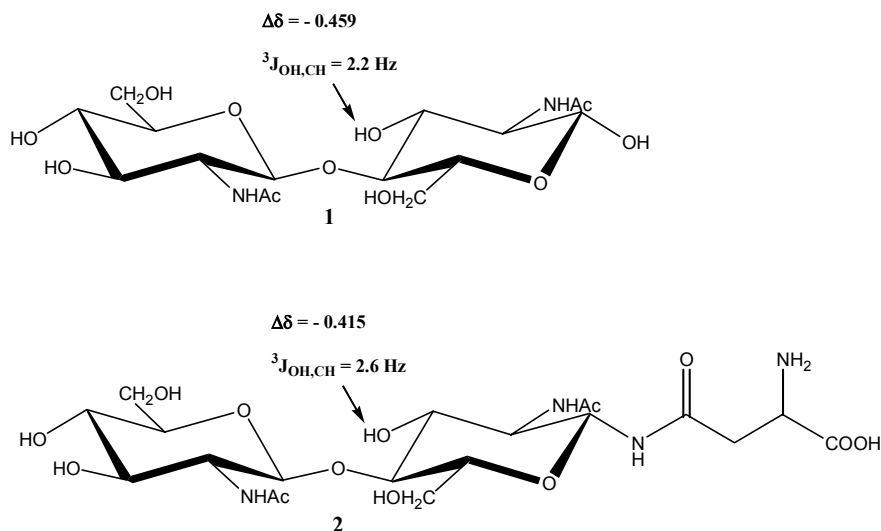


Figure 7. Schematic representation of β -D-GlcpNAc-(1 \rightarrow 4)- β -D-GlcpNAc, **1**, and β -D-GlcpNAc-(1 \rightarrow 4)- β -D-GlcpNAc-N-Asn, **2**. The hydroxy protons with large chemical shift differences of $|\Delta\delta| \leq 0.140$ ppm (see text) are indicated, as well as $^3J_{\text{CH,OH}}$ -values that indicate a restricted rotation around the C-O bond.

In β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-*O*-Me, **3**, and β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-*O*-Ser, **4**, all hydroxy proton signals have chemical shifts very similar ($|\Delta\delta| \leq 0.150$ ppm) to those measured for the corresponding monosaccharide methyl glycosides, with the exception of the O(2')H signal, which was shielded with 0.387 ppm in **3**, and with 0.420 ppm in **4** (Figure 8). Inspection of the MM2 minimum energy conformations of **3** and **4** showed that O(2')H was close to the 2-acetamido group of the neighbouring sugar residue. In the corresponding disaccharides with a 2-hydroxy group instead of a 2-acetamide group, the O(2')H proton has either a very small or a positive $\Delta\delta$.⁶¹ The upfield shift experienced by O(2')H is thus probably due to the influence of the 2-acetamido group. The coupling constant-values and the temperature coefficients suggest that no hydrogen bond interaction exists in **3** and **4**. The low exchange rate of O(2')H is thus probably due to a limited accessibility of water because of the proximity to the 2-acetamido group of the neighbouring sugar.

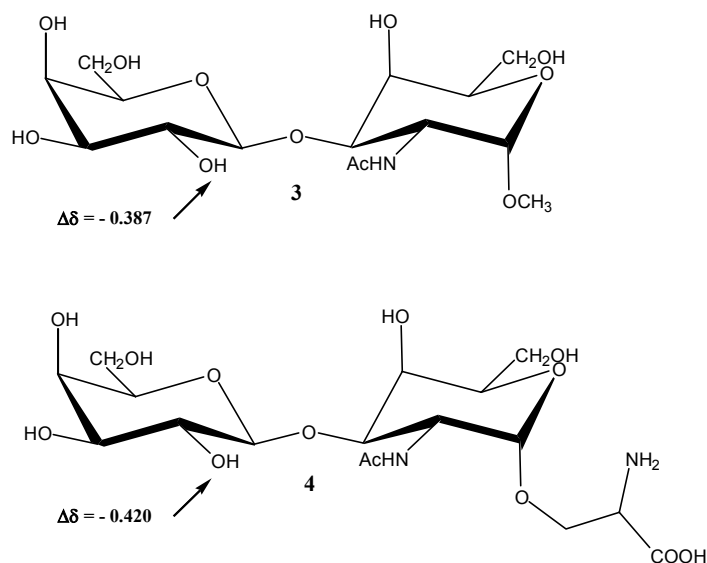


Figure 8. Schematic representation of β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-O-Me, **3**, and β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-O-Ser, **4**. The hydroxy protons with large chemical shift differences of $|\Delta\delta| \leq 0.150$ ppm (see text) are indicated.

In the two disaccharides **1** and **3**, and in their amino acid linked counterparts **2** and **4**, the hydroxy proton, which experiences an upfield shift, also showed lower rate of exchange with water. It is possible that the steric interference caused by the oxygen ring in **1** and **2** and by the 2-acetamido group in **3** and **4**, competes with solvation and leads to the upfield shift of the hydroxy proton resonances.

The work in paper I agreed well with preceding studies,^{44,45,61,64} where it was shown that when the chemical shift of a hydroxy proton significantly differs from that in the monosaccharide methyl glycoside, it is an indication of a particular conformational feature. This also confirmed the potential use of hydroxy proton shifts in structural analysis.

1.6.2. ^1H NMR studies of hydroxy protons in a glycopeptide (paper II)

The V[β -Galp(1 \rightarrow 3)- α -GalpNAc(1 \rightarrow)]THPGY glycopeptide, **5**, which is a partial structure of oncofetal fibronectin,⁷⁶ was chosen in order to extend the work in paper I to larger glycopeptides where interactions between the sugars and the amino acids are more likely. The hydroxy protons of an antifreeze glycopeptide in aqueous solution previously were observed by NMR spectroscopy.⁷⁷ A temperature of 5 °C, and a concentration of 17 mg/mL in water was employed, but the signals were not assigned and no attempt was made to use them as an aid in the interpretation of glycopeptide conformation.

In order to be able to observe the hydroxy protons of the glycopeptide, the sample was cooled to -13 °C to reduce the rate of exchange with water. To avoid

freezing, acetone- d_6 was added to give an 85% $H_2O/15\%$ $(CD_3)_2CO$ solution. Also, the pH was adjusted by adding minute amounts of NaOH, in order to avoid acid- or base-catalysed proton exchange that would reduce the intensity of the hydroxy proton signals. The importance of sample pH can be seen in figure 9, where a series of 1D spectra of the hydroxy proton region of the V[β -Galp(1 \rightarrow 3)- α -GalpNAc(1 \rightarrow)]THPGY glycopeptide at various pH are displayed. The hydroxy proton signals are not observable at a low pH, but the signals are visible at neutral pH. The exchange rate (k_{ij}) was determined as the slope of a plot of $I_{ij}(t_m)$ vs $M_j^0 t_m$. A series of 2D NOESY spectra with short mixing times was recorded, and the peak volumes were determined by using the program AURELIA (Bruker, Germany).⁷⁵

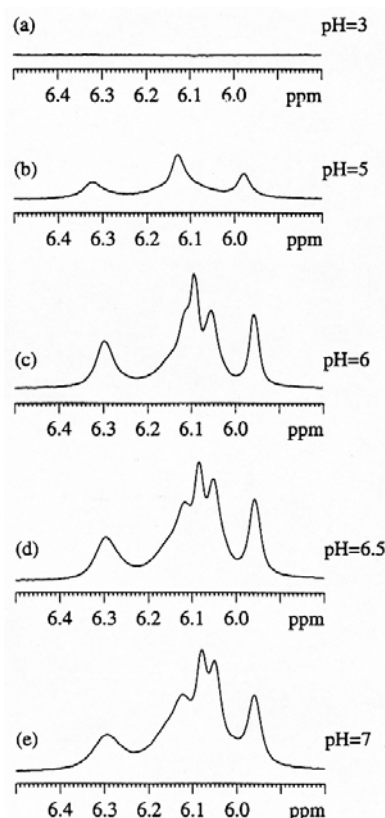


Figure 9. A series of 1D 1H NMR spectra of the hydroxy proton region of the V[β -Galp(1 \rightarrow 3)- α -GalpNAc(1 \rightarrow)]THPGY glycopeptide recorded at various pH values.

In this work, the chemical shifts, coupling constants ($^3J_{CH,OH}$), temperature coefficients, exchange rates, and NOEs of the hydroxy protons were measured. Analogous to the previous findings the data showed that the O(2')H was shielded with 0.5 ppm as compared with the corresponding methyl monosaccharide (Figure 10, Table 2). Also, O(2')H had the lowest rate of exchange with water (Table 2). The conclusion, analogous to paper I, was made that this was due to a reduced

contact with water due to steric interference caused by the 2-acetamido group of GalpNAc.

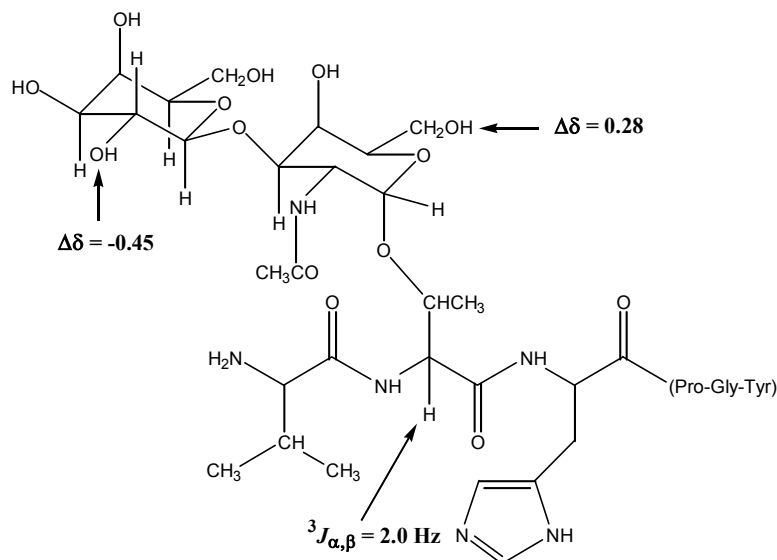


Figure 10. Schematic representation of the V[β-Galp(1→3)-α-GalpNAc(1→)]THPGY glycopeptide. The hydroxy protons with large chemical shift differences $|\Delta\delta| \leq 0.15$ ppm (see text) are indicated, as well as the $^3J_{\text{H}\alpha, \text{H}\beta}$ -value that indicates a restricted rotation around the Thr Cα-Cβ bond

In addition to O(2')H, O(4)H, and O(4')H also showed a lower rate of exchange than that of the other hydroxy protons (Table 2). Hydroxy protons whose contacts with the solvent are decreased should have lower exchange rates than hydroxy protons in full contact with the solvent. The lower rate of exchange measured for O(4')H and O(4)H in **1** was also observed for the monosaccharide methyl glycoside.⁴⁴ This might be due to reduced solvation of an axial hydroxy group relative to an equatorial group.^{78,79} The data was not as conclusive for β-D-Galp-(1→3)-α-D-GalpNAc-*O*-Me and β-D-Galp-(1→3)-α-D-GalpNAc-*O*-Ser of paper I in this thesis.⁸⁰

Table 2. The chemical shift differences ($\Delta\delta$)^a, temperature coefficients ($d\delta/dT$), and exchange rates ($k_{\text{ex}}/\text{s}^{-1}$), of the hydroxy protons in the V[β-Galp(1→3)-α-GalpNAc(1→)]THPGY glycopeptide.

	$\Delta\delta$	$d\delta/dT$	k_{ex}
O(2')H	-0.45	-13.7	25
O(3')H	-0.07	-12.9	41
O(4')H	0.06	-11.2	25
O(6')H	0.05	-15.2	59
O(4)H	0.04	-14.9	19
O(6)H	0.28	-17.2	62

^a Chemical shift of the hydroxy proton signal in the disaccharide minus that of the corresponding monosaccharide (methyl β-D-Galp and D-GalpNAc, respectively)

The O(6)H was found to be deshielded with 0.28 ppm (Figure 10, Table 2). A large positive $\Delta\delta$ has previously been observed for hydroxy protons, which are close to hydroxy protons of neighbouring sugars.^{45,65} The large positive $\Delta\delta$ of O(6)H cannot, however, be explained by the proximity to a neighbouring sugar and the reason for this upfield shift is still unclear. Due to broad resonance lines, only two $^3J_{\text{CH,OH}}$ -values, for O(2')H and O(4')H could be measured. These values, around 4 Hz, indicated free rotation of both hydroxy protons. The temperature coefficients had large values, in the range -11.2 to -17.2 ppb/ $^{\circ}\text{C}$ (Table 2), suggesting that the hydroxy protons are not involved in hydrogen bond interaction. The small $^3J_{\text{H}\alpha,\text{H}\beta}$ -value (2.0 Hz) measured for the threonine residue indicated a restricted rotation around the sugar-peptide linkage. The NOEs between the sugar and peptide moieties, which were only to one face of the sugar, also confirmed the restricted rotation around the sugar-peptide linkage (Figure 11).

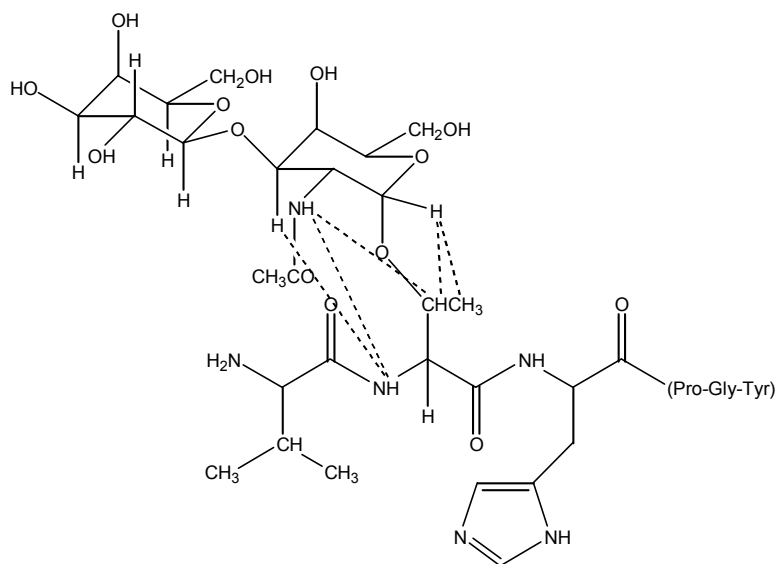


Figure 11. Schematic representation of the inter-residue NOEs, between the disaccharide and the peptide, at the site of glycosylation.

1.6.3. Conclusion

The work in papers I and II has shown that it is possible to observe hydroxy protons of saccharides that are conjugated to amino acids and peptides. The importance of careful sample preparation is shown in paper II, where the effect of different pH-values on the viability of the hydroxy protons is illustrated. Information could be obtained concerning the existence of hydrogen bonds through monitoring $^3J_{\text{CH,OH}}$ -values, chemical shifts, temperature coefficients and exchange rates of hydroxy protons. The data extracted also showed that information on the hydration of the hydroxy groups could be obtained through monitoring chemical shifts and the exchange rate of the hydroxy protons. It also

showed that the chemical shift values are quite reliable, since very similar chemical shifts are measured for signals of hydroxy protons with similar electronic surroundings. The investigations performed in papers I and II showed promising results for the future use of hydroxy protons as conformational probes in ^1H NMR studies of glycoproteins.

Part 2

¹H NMR studies on the solution conformation of contulakin-G and analogues (Papers III and IV)

2.1. Introduction to conotoxins

Conotoxins are disulfide-rich peptide toxins produced by marine cone snails.⁸¹⁻⁸³ These snails are widely distributed in tropical waters, and the venom is utilised primarily for prey capture and defence. The prey includes fish, worms and other molluscs. Cone snails inject the venom into their prey by means of a disposable harpoon-like radular tooth, and the venom rapidly causes immobilization. The prey is subsequently engulfed and digested (Figure 12).

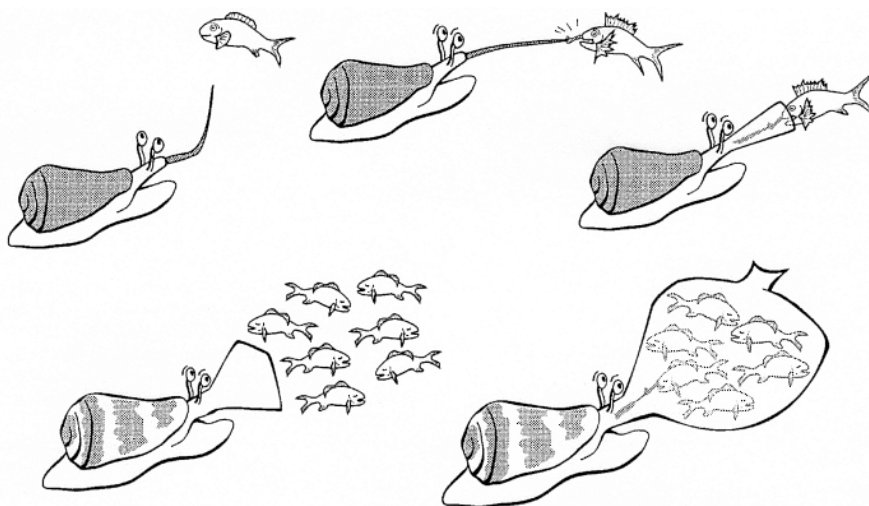


Figure. 12. A cartoon representing the hook-and-line (top panel) and the net strategy of fish-hunting cone snails. *Conus striatus*, *magus* and *purpurascens* are examples of hook-and-line piscivores. Species such as *Conus tulipa* and *Conus geographus* use a net strategy. Reprinted from *Molecular Biology of the Cell*, (1997, vol 8, pgs. 2101-2109) with permission by the American Society for Cell Biology.

The venom contains a mixture of various conotoxins that block, or in some instances activate, a broad spectrum of ion channels and neuronal receptors in the cone snail's prey.^{84,85} The venom of each *Conus* species has its own set of typically 50-200 different peptides.⁸⁶ In all, there may be over 50,000 unique peptide sequences present in the approximately 500 cone species, and these sequences may be grouped into superfamilies. The major superfamilies of conotoxins are the A-, M-, O-, S-, T-, and the P-superfamily.^{85,87} These superfamilies are subdivided into families that share a conserved disulfide bridging pattern and common molecular target (Table 3).^{83 85}

Table 3. Conotoxin classification and molecular targets. (Modified from McIntosh, J.M.; Jones R.M., *Toxicon* **2001**, 39, 1447-1451, and Jones R.M.; Bulaj, G, *Curr. Opin. Drug Discov. Dev.* **2000**, 3, 141-154)^{83,84}

Superfamily	Cysteine arrangement ^a	Family	Molecular target
A	CC-C-C (2)	α	nAChR (competitive antagonist)
	CC-C-C (2)	ρ	α_1 -adrenoreceptors (antagonist)
	CC-C-C-C-C (3)	α	nAChR (competitive antagonist)
	CC-C-C-C-C (3)	κ	K ⁺ channel (antagonist)
M	CC-C-C-CC (3)	μ	Na ⁺ channel (blocker)
	CC-C-C-CC (3)	ψ	nAChR (non-competitive antagonist)
O	C-C-CC-C-C (3)	δ	Na ⁺ channel (delays inactivation)
	C-C-CC-C-C (3)	μ	Na ⁺ channel (blocker)
	C-C-CC-C-C (3)	ω	Ca ²⁺ channel (blocker)
	C-C-CC-C-C (3)	κ	Shaker K ⁺ channel (blocker)
	C-C-CC-C-C (3)	γ	T-type Ca ²⁺ channels (blocker)
P	C-C-C-C-C-C (3)	Spastics	Unknown
S	C-C-C-C-C-C-C-C-C (5)	σ	5-HT ₃ (antagonist)
T	CC-CC (2)	τ	Pre-synaptic Ca ²⁺ channel (blocker)
		χ	Noradrenaline transporter inhibitor
- ^b	C-C (1)	Conopressin	Vasopressin receptors (agonist)-GPCR
- ^b	C-C (1)	Contryphans	Unknown
- ^b	Linear	Conantokins	NMDA channel (non-competitive antagonist)
		Contulakins	Neurotensin receptors (agonist)

^a. The number within parantheses refers to the number of disulfide bridges in the compound.

^b. The peptides have not been assigned to a superfamily. Abbreviations: NACHR (nicotinic acetyl choline receptor), 5-HT₃ (a serotonin receptor subtype, another name for serotonin is 5-hydroxytryptamine, 5-HT), GPCR (G-protein coupled receptors). NMDA is a receptor for the neurotransmitter glutamate. The receptor is named after the agonist N-methyl-D-aspartate.

The conotoxins are typically small (8-40 amino acids) peptides with a high content of disulfide bridges. Many contain multiple post-translational modifications, such as C-terminal amidation, glycosylation, hydroxylation of proline, γ -carboxylation of glutamate, L- to D-epimerisation of naturally occurring amino acids and regioisomeric bromination at C-6 of L-tryptophan.^{85,88} Although they are in the minority, *Conus* peptides with one or no disulfide bridge possess novel pharmacological properties and are of significant medicinal interest.⁸⁵ One of these is contulakin-G, an *O*-glycosylated peptide isolated from *Conus*



geographus (Figure 13).⁸⁹ The conformation of contulakin-G was investigated in paper III.⁹⁰

Figure 13. A shell from the marine snail *Conus geographus*.

2.2. The conotoxin contulakin-G and analogues

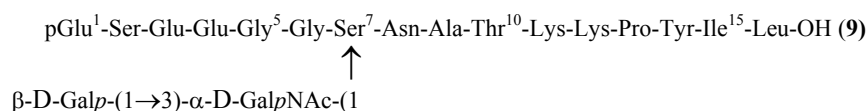
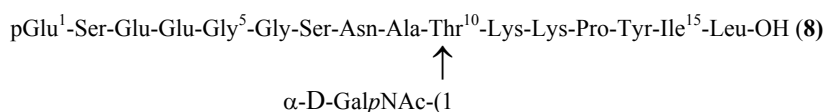
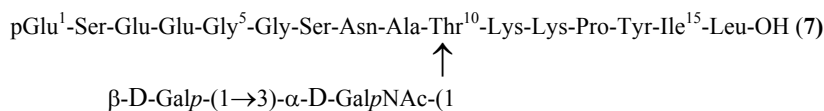
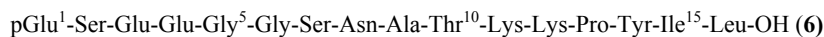
2.2.1. Isolation and initial characterisation

Craig *et al.* purified and characterised contulakin-G, in 1999 (Scheme 1).⁸⁹ Contulakin-G was found to be a 16 amino acid *O*-linked glycopeptide with the sequence ZSEEGGSNAT*KKPYIL, where the N-terminus consisted of a pyroglutamic acid residue (Z). The disaccharide β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-(1 \rightarrow was attached to the threonine residue in position 10 (see Scheme 1). *O*-glycosylated threonine had previously not been found in *Conus* peptides. The carboxy-terminus of contulakin-G displayed significant sequence identity with members of the neurotensin family of neurotransmitters, and contulakin-G also proved to be a neurotensin receptor agonist at physically relevant concentrations.⁸⁹ When contulakin-G was compared with the non-glycosylated contulakin-G and with neurotensin, which is known to be an analgesic, contulakin-G was two orders of magnitude more potent *in vivo*.⁹¹ From these results, Craig *et al.* concluded that *O*-glycosylation was a strategy to enhance the binding affinity to neurotensin receptors. Concerning the mechanism by which the sugar residues affect the biological activity, it was speculated that the significantly enhanced potency of glycosylated contulakin-G could be due to *e.g.* an alternative mode of action, a modulation of the kinetics of the interaction between ligand and receptor, an enhanced supply of the substance to the site of action due to inhibition of proteolytic degradation, or an enhanced uptake through the blood brain barrier.⁹² Other suggestions were that it could be due to a conformational change of the peptide, or a change of the binding to the receptor in a way that is favourable for activity.

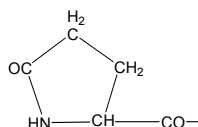
2.2.2. Characterisation by the sluggishness - loss of motor control assay

To study the effect of glycosylation on the biological activity, the unglycosylated peptide, **6** (Scheme 1), was synthesised together with two analogues; one glycosylated with the monosaccharide α -D-GalpNAc-(1 \rightarrow at Thr10, **8**, and one glycosylated with the disaccharide β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc-(1 \rightarrow at Ser7, **9** (Scheme 1). In order to determine the potency of compounds **6** - **9**, Craig *et al.* performed a sluggishness assay,⁹³ in which mice were injected intracerebral ventricular with the compound in solution and their behaviour was observed. After administration, the mouse was placed onto a bench top (by lifting it by the tail for a second) and the activity was observed, *e.g.*, whether the mouse showed any activity or rested with its legs spread out. In addition, the time elapsed after the administration prior to the mouse regaining control (so that the hind part when at rest resembled that of the NSS controls) was noted, as well as the time elapsed when the mouse

moved freely when released. Contulakin-G, **7**, was found to be the most active, and the nonglycosylated peptide, **6**, the least active. The activity of the glycosylated Thr10 analogue, **8**, and the glycosylated Ser7 analogue, **9**, were somewhat in between that observed for contulakin-G, **7**, and the non-glycosylated peptide **6**. However, the relative potency of all four compounds could not be assessed.



pGlu = pyroglutamic acid:



Scheme 1. Structure of contulakin-G (**7**), the non-glycosylated (**6**), and glycosylated (**7** and **8**) analogues.

2.2.3. Characterisation by the tail flick - acute pain assay

Since the relative potency of compounds **6** - **9** could not be determined with the sluggishness assay, the tail flick assay of antinociception was employed.⁹³ In the tail flick-acute pain assay, mice were administered intrathecally according to the method of Hylden and Wilcox with the compound in solution,⁹⁴ and then gently wrapped in a towel with the tail exposed. At various time points following the intrathecal injection, the tail was dipped in a water bath maintained at 54° C and the time taken for the mouse to withdraw its tail recorded. If there was no withdrawal by 8 seconds, the tail was withdrawn to avoid tissue damage. The data showed that shifting the position of the glycosylation from residue 10 to 7, and reducing the size of the glycan decreased the activity compared with the native contulakin-G. In relative terms, **7** was the most potent, **8** was more potent than **9**, which had approximately the same potency as **6**. In recognition of the promising test results achieved, in the year 2000 contulakin-G entered preclinical trials for

short term management of post-operative pain.⁸⁵ Today, contulakin-G has reached phase II clinical trials.

2.3. The solution conformation of glycopeptides

NMR studies of glycopeptides provide insight into the influence of glycosylation on the conformation of the polypeptide backbone. The peptide itself is a very flexible structure, but it has been found that there are peptide sequences that could give rise to secondary structures, such as turns, which display some intrinsic stability.^{95,96} Glycosylation has been shown to have varying effects on the structure of peptides, and it seems to depend on both the amino acid sequence and the glycan.⁹⁷ Some studies have shown that the saccharide may have no or marginal influence on the peptide.⁹⁸⁻¹⁰⁰ In a few cases, the carbohydrate moiety has been shown to form stabilising contacts with the protein.^{101,102} These stabilisations can occur as hydrogen bonds between the saccharide and peptide, and have been observed in DMSO.^{74,103,104} Observations of hydrogen bonds between the saccharide and protein parts in aqueous medium are however few.^{74,97,105} Glycosylation has been found to induce turns in peptide backbones,^{15,66,67} and evidence for a hydrogen bond in aqueous medium between the sugar and the glycosylated threonine residue, which presumably helps in forming a turn in the peptide backbone, has been obtained.^{97,105} However, a study has also shown that glycosylation is not always turn-inducing, but can in some cases result in local straightening of the peptide backbone.¹⁰⁶ Where there are no persistent interactions between the saccharide and the protein, it is more difficult to understand how the saccharide influences the protein. It has been proposed that oligosaccharides may influence the local structure of the glycosylation site either by altering the solvation properties of the peptide, or acting as a bulk group, where the saccharide limits the conformational space of the peptide.^{101,107,108}

Most conformational studies of glycopeptides concern the influence that the saccharide has on the peptide, but the influence of the peptide on saccharide conformation has also been studied. The results include examples where the peptide influenced the conformation of the saccharide moiety (Upon glycosylation with Man₅GlcNAc₂, the conformation of the NAc-group closest to the backbone changed, and the strengths of the individual interresidual ROEs were affected).¹⁰⁷ Also, there are examples where the conformation of the carbohydrate moiety was not affected (the conformation of *O*-linked sialyl-Lewis-X attached to an octapeptide was the same as that of free sialyl-Lewis-X).¹⁰⁹

One general conclusion that has been drawn from investigations of *O*-glycosylated peptides is that the effects of glycosylation primarily are due to the sugar closest to the backbone.¹¹⁰ However, differences in average conformation for peptide backbones substituted with carbohydrate chains of different length have been found.^{66,109} Also, a sequential reorientation of the saccharide unit closest to the peptide, as compared with the peptide, has been observed upon elongation of the saccharide chain.¹⁰⁹ Results indicated that *O*-glycosylation could affect the peptide conformation near the glycosylation site,^{66,67} as well as the conformation of the peptide backbone at distant sites.^{111,112} One consequence of

O-glycosylation in small peptides is to limit the conformational space for an otherwise very flexible peptide.^{66,67,108,111-115}

In studies of peptide and protein conformation, 2,2,2-trifluoroethanol (TFE), and TFE/water solutions have been used for decades to stabilise secondary structures.¹¹⁶ The effects of TFE on peptide chain conformation are diverse, and appear to depend on the particular amino acid sequences, the cosolvent concentration and other solution conditions as well as on the structures involved. The molecular interactions of TFE with peptides that cause this stabilisation are not well understood, though models have been proposed.^{117,118} In most cases, addition of TFE (10-50%, v/v) enhances the helicity of peptides that are already helical in solution or induces helicity in disordered peptides.¹¹⁶ There have also been reports of β -sheet, β -hairpin, and β -turn stabilisation by TFE.¹¹⁶ TFE solutions have also been used to stabilise glycopeptides.¹¹⁹ However, the effect of TFE on saccharide conformations is little investigated,¹²⁰ and should therefore be used with precaution in conformational studies of glycopeptides.

2.4. ¹H NMR methods used in papers III and IV

2.4.1. NOEs in determination of peptide structure

By convention, the distance between protons A and B, located in amino acid residues *i* and *j* is denoted by $d_{AB}(i,j)$. For example, the distance between the H α proton of residue number *i* and the NH proton of residue number *j* is denoted $d_{\alpha N}(i,j)$. The $d_{\alpha N}(i,j)$ NOEs are utilised to do the sequential assignment of the amino acids that together constitute the primary structure. The main secondary structures that are found in polypeptides and proteins are the α -helix, the β -sheet, and the β -turn.¹²¹ Each of these structural elements has a characteristic pattern of NOE connectivities (Figure 14), since each of them has a characteristic set of corresponding short (and long) inter-proton distances (Table 4).¹²¹

Table 4. Short (≤ 4.5 Å) sequential and medium-range ¹H-¹H distances in polypeptide secondary structures. (Modified from Wütrich, K. *NMR of Proteins and Nucleic acids* **1986**).¹²¹

Distance	α -helix	β	β_n	turn I ^a	turn II ^a
$d_{\alpha N}(i,i+1)$	3.5	2.2	2.2	3.4	2.2
				3.2	3.2
$d_{\alpha N}(i,i+2)$	4.4			3.6	3.3
$d_{\alpha N}(i,i+3)$	3.4			3.1-4.2	3.8-4.7
$d_{\alpha N}(i,i+4)$	4.2				
$d_{NN}(i,i+1)$	2.8	4.3	4.2	2.6	4.5
				2.4	2.4
$d_{NN}(i,i+2)$	4.2			3.8	4.3
$d_{\beta N}(i,i+1)$	2.5-4.1	3.2-4.5	3.7-4.7	2.9-4.4	3.6-4.6
				3.6-4.6	3.6-4.6
$d_{\alpha\beta}(i,i+1)$	2.5-4.4				

^a For the turns, the first of two numbers applies to the distance between residues 2 and 3, the second to that between residues 3 and 4 (A turn incorporates four residues, see the original publication).¹²¹

	β, β_p	α -Helix	Turn I	Turn II
$d_{\alpha N}(i, i+3)$				
$d_{NN}(i, i+2)$				
$d_{\alpha N}(i, i+2)$				
$d_{NN}(i, i+1)$				
$d_{\alpha N}(i, i+1)$				
$^3J_{HN\alpha}(\text{Hz})$	9 9 9 9 9 9	4 4 4 4 4 4 4	4 9	4 5
	1 2 3 4 5 6	1 2 3 4 5 6 7	1 2 3 4	1 2 3 4

Figure 14. Survey of the sequential and medium range ^1H - ^1H NOEs and the spin coupling constants $^3J_{\text{NH},\text{H}\alpha}$ in the parallel or antiparallel β -sheet, the α -helix, and turns of types I and II. The numbers at the bottom represent the amino acid residues in the secondary structure elements and the $^3J_{\text{NH},\text{H}\alpha}$ -values. Short ^1H - ^1H distances are indicated by lines, and the thickness of the lines represent the NOESY cross-peak intensities. (Modified from Wagner, G.; Neuhaus, D.; Worgotter, E.; Vasak, M.; Kagi, J. H. M.; Wütrich, K. *J. Mol. Biol.* **1986**, 187, 131-135).¹²²

These patterns of NOE connectivities are used to identify the different secondary structures. Regular secondary structures are also characterised by the three bond $^3J_{\text{NH},\alpha\text{H}}$ coupling constant (Figure 14).

2.4.2. Chemical shifts in determination of peptide structure

Chemical shifts have become a powerful diagnostic tool for the detection of secondary and tertiary structure in peptides and proteins.^{123,124} To utilise the structural information that can be found in the chemical shifts of a peptide or protein, random coil values for peptides and proteins have been worked out as reference sets of values.¹²⁵ The difference in chemical shift between the observed and random coil values, $\Delta\delta$, are useful for identifying secondary structures, such as α -helices and β -sheets.^{124,126} The αH experience on average an upfield shift in α -helices, while it experiences a downfield shift in β -strand or extended configurations. This is utilised in the chemical shift index (CSI) method,¹²⁷ where a random coil value is subtracted from the αH chemical shift. If the difference in chemical shift is larger than 0.1, the corresponding amino acid is assigned an index value of +1 or -1, depending on the sign of the difference of the subtraction. The resulting pattern of the index values is then analysed to detect the presence of secondary structures. A CSI method for αC has also been developed.¹²⁸

2.4.3. Temperature coefficients of amide protons

Slowed exchange kinetics and low temperature coefficients can indicate protection of the amide proton from the solvent by *e.g.* hydrogen bonding.⁹⁵ It has become common practice to assume that amide protons with temperature coefficients ($d\delta/dT$) less than -4 ppb/ $^{\circ}\text{C}$ are protected from solvent contact.^{129,130} Random coil peptides in the series GGXGG, which are not likely to have secondary structures that protect amides from exchange, displayed temperature gradients of -6.43 to -9.32 ppb/ $^{\circ}\text{C}$ in aqueous medium.¹³¹ In proteins, hydrogen-bonded exchange-protected amide protons were characterised by $d\delta/dT$ values of -2.0 ± 1.4 ppb/ $^{\circ}\text{C}$ while exposed amide protons typically display gradients of -6.0 to -8.5 ppb/ $^{\circ}\text{C}$.¹³² However, exceptions to this rule occur. In partially folded peptides, for example, exceptions are common due to a decrease in the population of the structured state upon warming.¹³² To predict hydrogen bonding in proteins with more reliability, methods such as plotting NH chemical shift deviations from random coil reference values *versus* the $d\delta/dT$,¹³² or combining the amide temperature coefficients with exchange rate data have been used.¹³³ In smaller molecules, amide proton exchange rates are combined with temperature coefficients and coupling constants to provide an indication if the amide proton is involved in hydrogen bonding or not. NOESY distance measurements and modelling can give further information on the likeliness that there are possible hydrogen bond donors or acceptors within hydrogen bond distance.

2.4.4. Coupling constants of amide protons

In peptides and proteins, The value of the coupling constant between the αH and amide protons, $^3J_{\text{NH,H}\alpha}$, can provide information on the backbone conformation of peptides and proteins; extended strand displays high ($^3J_{\text{NH,H}\alpha}=7.6\text{-}9.3$ Hz), α -helices low (~ 4 Hz) and random coils intermediate ($\sim 6\text{-}7$ Hz) $^3J_{\text{NH,H}\alpha}$ -values.⁹⁶ The interpretation of a certain coupling constant value is however often not straightforward, since it may represent multiple angles as well as conformational averaging. In highly flexible short peptides, a random coil structure is often encountered, where the peptide bonds are subject to conformational averaging.

2.4.5. Determination of absolute conformation

To characterise the absolute configuration of the serine residues of the peptides in paper IV, the GC-MS method of Ström *et al.* was used.¹³⁴ A small amount (< 0.1 mg) of each peptide were hydrolysed by using 6 M hydrochloric acid. Prior to injection on the column, the acid function was esterified with (S)-(+)-butanol, and perfluoropropanoic acid was allowed to react with the amine and alcohol functions, in order to make the amino acids more volatile. Reference samples were prepared by derivatisation of commercially available D- and L- serine as outlined above.

2.5. Present work: Contulakin-G and three analogues (Paper III)

As a first step towards understanding the effect of glycosylation of the biological activity of contulakin-G, an NMR study on the solution conformation of contulakin-G, **7**, and analogues **6**, **8** and **9** was performed (see Scheme 1, section 2.2.2.). The goal of the NMR study was to investigate if differences in conformation could be correlated to the observed differences in biological activity.

The ^1H NMR chemical shifts of the amino acid residues in the glycosylated peptides **7**, **8**, and **9**, were found to be similar to those in the corresponding non-glycosylated peptide **6**, with the exception of signals from NH and αH protons at and around the glycosylation sites. This indicated that the glycosylation only affected the peptide backbone at and around the glycosylation site. The $^3J_{\text{NH,H}\alpha}$ values in the peptide chain, in the range of 5.7-9.0 Hz, and the $d_{\alpha\text{N}}(i,i+1)$ and $d_{\text{NN}}(i,i+1)$ NOEs (Figure 15), indicated that the peptide backbones were mainly in random coil or extended conformation.^{95,96,121} The presence of sequential NH-NH NOEs from Ser7 to Thr10 in **7**, from Gly5 to Thr10 in compound **8**, and from Ser7 to Ala9 in compound **9**, however indicated that the glycopeptides have some population of folded structures in these regions. In contulakin-G, **7**, NH-NH NOEs from Tyr14 to Leu16 suggested that an additional turn was formed at the C-terminal end of the peptide chain. This result showed that glycosylation not only lead to local changes in conformations at the junction site, but also induced long distance conformational changes. In compounds **8** and **9**, only the NH-NH Tyr14-Ile15 NOE was observed. An increase in the $^3J_{\text{NH},\alpha\text{H}}$ value from 7.5 Hz to 8.6 Hz was observed for Thr10 in **7** and **8** compared to **6** and **9**. This larger value of $^3J_{\text{NH},\alpha\text{H}}$ in **7** and **8** is in agreement with the existence of some populations of ordered secondary structures.¹³⁵ The $^3J_{\text{NH,H2}}$ of 10.1 Hz measured for GalpNAc demonstrated a trans orientation of NH and C(2)H.

In **7** and **8**, *O*-glycosylated at Thr10, the small $^3J_{\text{H}\alpha,\text{H}\beta}$ value of 1.5 Hz and 2 Hz measured in Thr 10 of **7** and **8**, respectively, indicated a preference for an orthogonal orientation between the α and β protons of Thr10.⁷⁴ If the α -D-GalpNAc was undergoing motion relative to the peptide backbone through rotation around the C α -C β bond of Thr10, a value of 5 Hz, as obtained in the non-glycosylated peptide, **6**, should be expected. In **7** and **8**, five NOEs were found between the α -D-GalpNAc residue and Thr10, and Lys11 (Figure 16). All NOEs from the peptide chain were to only one side of the sugar, supporting a restricted rotation around the sugar-peptide linkage. In the peptide glycosylated at Ser7, **9**, only three NOEs were found between the sugar and the peptide chain (Figure 16). The αH signal of Ser7 appears in the NMR spectrum as a triplet of 9.9 Hz ($^3J_{\text{H}\alpha,\text{H}\beta} + ^3J_{\text{H}\alpha,\text{H}\beta'}$), as in the non-glycosylated peptide, **6**, suggesting a free rotation of the sugar around the sugar-peptide linkage.

The temperature coefficients ($d\delta/dT$) for the NH proton signals from compounds **6** - **9** were measured in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (95:5) solution from a series of one-dimensional spectra recorded between 5 °C and 45 °C. The signals from the NH protons in the peptide chain displayed temperature coefficients above -6.0 ppb/°C, confirming that the peptide chain in compounds **6** - **9** mainly adopted random coil or



Figure 15. Intra- and inter-residual NOEs involving the backbone NH protons in **6** - **9**. Thickness of the bars indicate the relative intensity of the NOE cross-peaks. A greyish NOE denotes that the signal was found to be overlapping or ambiguous. N.B. This is a later version of figure 2 in paper III, where some corrections have been implemented. (a) **6**, NOESY, 25° C, mixing time 300 ms; (b) **7**, NOESY, 25° C, mixing time 300 ms; (c) **8**, NOESY, 5° C, mixing time 300 ms; (d) **9**, NOESY, 5° C, mixing time 300 ms.

extended conformations. The lowest $d\delta/dT$ -values were found for NH of α -D-GalpNAc in compound **7** and **8** (-5.8 and -5.7 ppb/°C, respectively). This is of particular interest, since these values indicate that the torsional rigidity between the sugar and the peptide, suggested by the NOEs and by the value of the $^3J_{H\alpha, H\beta}$ of Thr10 (*vide infra*), could be enforced by a hydrogen bond interaction between NH of α -D-GalpNAc and probably the carbonyl oxygen of the *O*-linked Thr10 residue. In contrast, the large value for the temperature coefficient of NH of α -D-GalpNAc in the Ser7 glycosylated peptide, **9**, (-11.1 ppb/°C) indicated that this proton was not protected from exchange with the solvent through hydrogen bond formation.

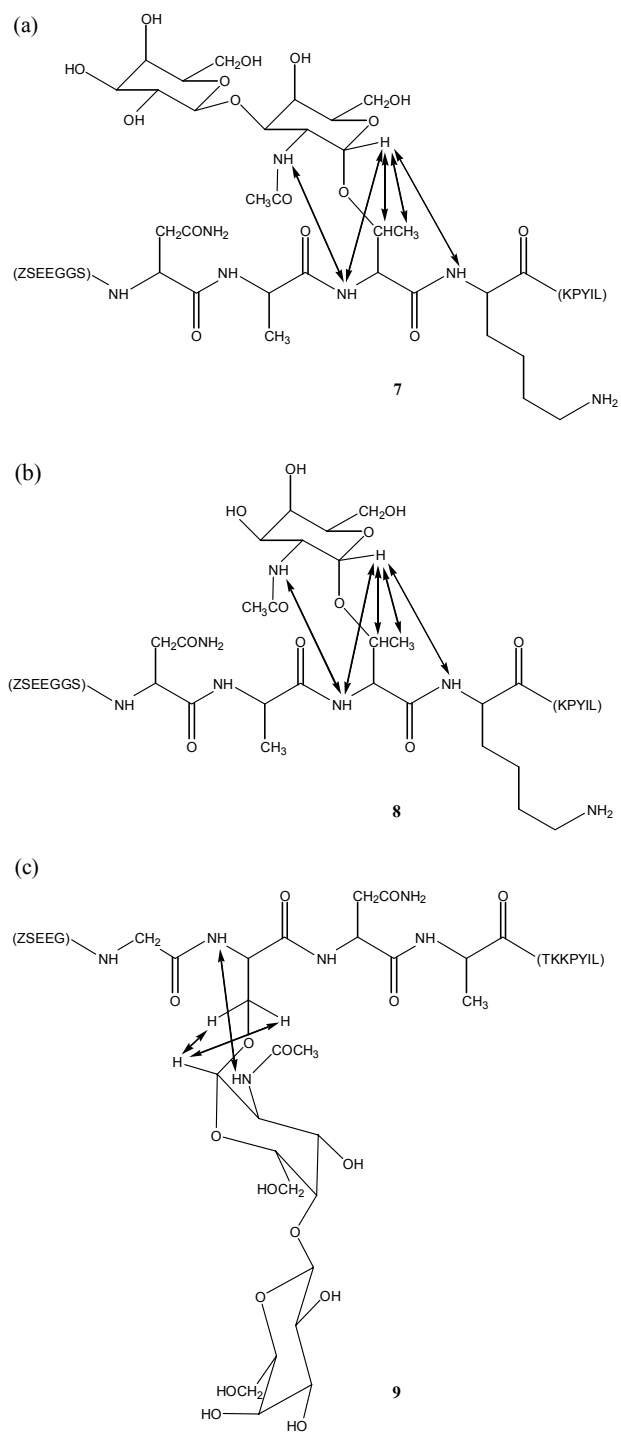


Figure 16. Schematic representation of interresidual NOEs observed at the junction between the peptide and the saccharides in **7**, **8**, and **9**.

2.5.1. A trial to study conformations in TFE/water solution

Since a large number of NMR studies on peptides have shown that the use of 2,2,2-trifluoroethanol (TFE) as a solvent can stabilise secondary structures,¹¹⁶ NMR spectra were also recorded in 50% TFE/50% H₂O. When the peptide backbone of **6** - **9** were analysed in TFE/H₂O solutions, very similar ¹H NMR chemical shifts of the amino acids in the glycosylated and non-glycosylated peptides were observed, except for the residues at the junction site. The shifts due to glycosylation were, however, not as large as in water. The ³J_{NH,H α} -values were very similar to those measured in water. The temperature coefficients, d δ /dT, were slightly smaller than those obtained for the compounds in water. As in water, the *d*_{αN(i,i+1)} NOEs characteristic of extended conformations were observed all along the peptide chains in the glycosylated and non-glycosylated peptides. The *d*_{NN(i,i+1)} NOEs present in water were also present in TFE/H₂O, but additional *d*_{NN(i,i+1)} NOEs were also found. The hydrogen bond interaction between Thr10 and α-D-GalpNAc found for **7** and **8** in water was also present in TFE/H₂O solution.

2.5.2. Conclusion

The NMR study showed that the four analogues mainly existed in random coil conformations, with some transient populations of folded conformation present around the site of glycosylation in the glycosylated peptides. In contulakin-G, **7**, indications of an additional turn at the C-terminal end of the peptide chain were found, which showed that glycosylation induced long distance conformational changes. However, the sites of local structure most likely occur in a subpopulation of the predominantly disordered solution ensemble of conformers. Since the six C-terminal amino acids of contulakin-G, **7**, are similar to several members of the neurotensin class of peptides, it might be that the terminal turn, which exists only in contulakin-G, **7**, the most active of the four analogues **6** - **9**, is an additional conformational feature important for the biological activity. However, previous NMR studies of neurotensin showed that no preferred conformation existed in solution.^{136,137} In compounds **7** and **8**, which are *O*-glycosylated at the Thr10 residue, an intramolecular hydrogen bond between the sugar and the peptide stabilised the structure. Since compounds **7** and **8**, which are *O*-glycosylated at the Thr10 residue, are more active than the non-glycosylated peptide, **6**, and the peptide *O*-glycosylated at Ser7, **9**, it was speculated that the stabilisation of conformation by hydrogen bonding to the sugar could contribute to the high biological activity.

2.6. A study of proteolytic degradation

There is significant sequence identity between the carboxyterminus of contulakin-G and members of the neurotensin family of neurotransmitters. A number of studies have characterized the structural requirements for neurotensin receptor agonists. It has been determined that the neurotensin 8-13 fragment (RRPYIL) retains a significant potency relative to neurotensin in a variety of *in vivo* and *in*

vitro assays.^{138,139} In fact, the neurotensin 8-13 fragment has been shown to have significant transfer across the blood brain barrier¹⁴⁰ and therefore the hydrolysis of the Pro7-Arg8 bond may represent a biologically relevant mechanism for allowing an active form of the peptide to be transferred into the brain. In contrast, cleavage of the Pro10-Tyr11 bond generating C-terminal tri-peptide (YIL) will inactivate neurotensin.¹⁴¹

If enzymatic degradation were responsible for the reduced antinociceptive activity *in vivo* of analogues **6**, **8** and **9** as compared with contulakin-G, **7**, one indication would be to do a proteolysis study to see if the rates of degradation parallel the biological activity. Therefore, Craig *et al.* investigated the dependence of the proteolytic cleavage of the Pro13-Tyr14 bond on the size, position, and presence of the glycan.⁹³ The presence of the β -D-Galp-(1 \rightarrow 3)- α -D-GalpNAc disaccharide attached to Thr10 in contulakin-G, **7**, significantly inhibited the enzymatic degradation. The monosaccharide α -D-GalpNAc in Thr10-contulakin-G, **8**, is somewhat less protected against proteolytic degradation than contulakin-G, **7**. In contrast, it was observed that both Ser7-contulakin-G, **9**, and the non-glycosylated analogue, **6**, the protection against enzymatic degradation afforded by the glycan appears to be completely abolished when the glycan is moved three residues towards the N-terminus of the peptide (analogue **9**) (*i.e.*, the level of protection against enzymatic degradation is equal to that of the non-glycosylated analogue, **6**). These findings are consistent with the proposal that the glycan sterically inhibits the enzyme from hydrolysing the Pro13-Tyr14 peptide bond and thereby inactivating the peptide. The stability to enzymatic degradation of these four compounds was hence found to correlate well with the *in vivo* potencies of the analogues as measured in the biological assays.

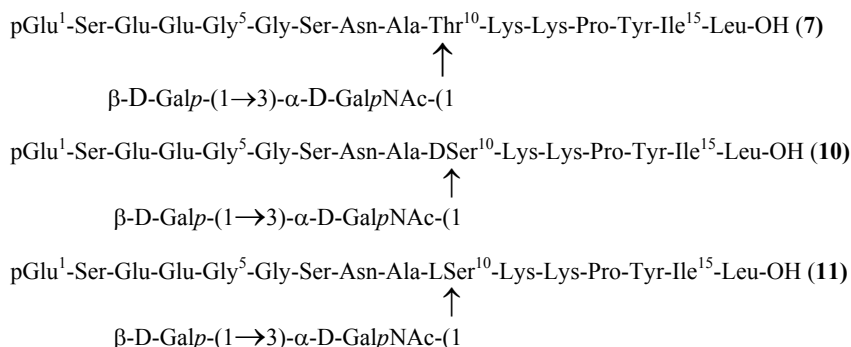
2.7. Present work: the [D-Ser10] and [L-Ser10] analogues (Paper IV)

Since the data obtained in the proteolysis study indicated that the main role of the glycan is a steric inhibition of enzymatic degradation, it was of interest to study the Ser10 analogue of contulakin-G. Whereas the site of glycosylation at residue number 10 consists of a Thr in contulakin-G, this site is a Ser residue in the Ser10 analogue. Previous studies have reported differences in glycan mobility and orientation relative to the peptide backbone depending on if the glycosylated residue were a Thr or a Ser.^{97,112}

As the synthetic product was purified with reverse phase HPLC, two components with different retention times were found; one fraction with shorter retention time containing compound **10**, and one fraction with longer retention time containing compound **11** (Scheme 2). The mass spectrum showed that both compounds had the correct mass, and that both were glycopeptides. In the antinociceptive assay of acute pain of the two samples it was found that glycopeptide **10** was active at about the same potency as contulakin-G, **7**, whereas glycopeptide **11** only was slightly active at more than 100 times the dose. (Scheme 2).⁹³

The 2D COSY, TOCSY and NOESY spectra of compounds **10** and **11** were very similar; the chemical shift differences between the signals from the amino acid protons in **10** and **11** were < 0.04 ppm. An exception was found for the α H signal of Ser10, which showed a difference of 0.06 ppm between compounds **10** and **11**. Also, the NH signals of Lys11 and α -D-GalpNAc showed differences between compounds **10** and **11** of 0.16 and 0.17 ppm, respectively. A major difference was found in the fine structure of the H α -H β and H α -H β' cross-peaks of Ser10 in the phase-sensitive DQF-COSY spectra. In the spectrum of compound **10**, the downfield proton had a $^3J_{H\alpha,H\beta} = 6$ Hz, whereas the upfield proton had a $^3J_{H\alpha,H\beta} = 4$ Hz. (No stereospecific assignments of the geminal H β protons of Ser10 were made). In the spectrum of compound **11**, the peak pattern of these two cross-peaks was reversed; the downfield proton had a J -value of 4 Hz, while the upfield proton had a J -value of 6 Hz. These data indicated that the structure of the two compounds differed by the configuration of C α of the Ser10 amino acid, which means that the serine residue substituted by the glycan would have been racemized during peptide synthesis.

The absolute configuration of the serine residues in the peptide chains of **10** and **11** was determined using the GC-MS method of Ström et al.¹³⁴ The analysis showed that **10** contained L-Ser and D-Ser in the ratio 2:1, while **11** only contained Ser with the L-configuration. This indicated that compound **10** had a D-Ser at the site of glycosylation, while **11** had a L-Ser.



Scheme 2. Contulakin-G, **7**, and the glycosylated [D-], **10**, and [L-Ser10], **11**, analogues of contulakin-G. The abbreviation pGlu stands for pyroglutamic acid (see Scheme 1).

The $^3J_{NH,H\alpha}$ values of **10** and **11**, in the range of 5.7-9.0 Hz, and the $d_{\alpha N(i,i+1)}$ and $d_{NN(i,i+1)}$ NOEs (Figure 17), indicated that the peptide backbones were mainly in random coil or extended conformation.^{95,96,121} The presence of sequential NH-NH NOEs of residues 7-12 and 14-16 in **10**, and from 7-11 in compound **11**, indicated that the glycopeptides have some population of folded structures in these regions. However, none of the $d_{NN(i,i+1)}$ NOEs found in **10** and **11** were strong, suggesting a high degree of flexibility in glycopeptides **10** and **11**. The [D-Ser10] glycopeptide **10** had a Gly6/Asn8 $d_{NN(i,i+2)}$ NOE indicative of a

turn, which could also be found, in contulakin-G, **7**, but not in the [L-Ser10] analogue **11**.

In paper III, an increase in J -value of **7** (8.6 Hz) and **8** (9 Hz) as compared to **6** (7.5 Hz) was attributed to the presence of ordered secondary structures in **7** and **8**. A larger $^3J_{\text{NH},\alpha\text{H}}$ -value was also measured for [D-Ser10] in **10** (8.5 Hz) as compared to [L-Ser10] in **11** (6.6 Hz). It should be noted that in analogue **9**,⁹⁰ a value of 6.4 Hz was measured for the glycosylated Ser7. Further does all of the non-glycosylated Ser residues in **7** – **11** have J -values around 6.0 to 6.8 Hz. Thus, the higher $^3J_{\text{NH},\alpha\text{H}}$ -value of [D-Ser10] in **10** would indicate a higher degree of secondary structures around [D-Ser10] in **10** as compared to [L-Ser10] in **11**.

In the [D-Ser10] analogue, **10**, and the [L-Ser10] analogue, **11**, the large $d\delta/dT$ -value obtained for NH of α -D-GalpNAc (-11.3 ppb/ $^{\circ}\text{C}$ in **10**, and 8.3 ppb/ $^{\circ}\text{C}$ in **11**) showed that these protons are not protected from exchange with the solvent through hydrogen bond formation. Also, in **10** and **11**, the two $^3J_{\text{H}\alpha,\text{H}\beta}$ of Ser7 were approximately equal, with values in the range 4 to 6 Hz as measured from 2D DQF-COSY, and the relative intensities of the two H α -H β NOEs were also approximately equal. This indicated that the C α -C β bond is subject to conformational averaging, and that there is motion around the sugar-peptide linkage. The $^3J_{\text{H}\alpha,\text{H}\beta}$ -value characteristic for conformational averaging ($^3J_{\text{H}\alpha,\text{H}\beta} \approx 6$ –8 Hz) is, however, not easy to distinguish from the corresponding value for the gauche-rotamer ($^3J_{\text{H}\alpha,\text{H}\beta} \approx 4$ Hz) only from the fine structure of the DQF-COSY cross-peaks, since there is a limitation in digital resolution and linewidths. A few NOEs were found between the peptide and the carbohydrate moiety in **10** and **11** (Figure 18). Two additional NOEs, NH(Ser10)-NH(α -D-GalpNAc) and NH(Ser10)-CH₃CO(α -D-GalpNAc), were found for **10** as compared to **11**. These additional NOEs together with a larger $^3J_{\text{NH},\alpha\text{H}}$ (Ser10)-value of 8.5 Hz indicated that the average orientations of the glycan in the [D-Ser10]-analogue **10** and in the [L-Ser10]-analogue **11** were different. The $^3J_{\text{NH},\text{H}2}$ of 9.7 Hz measured for GalpNAc of **10** and **11** demonstrated a trans orientation of NH relative to C(2)H.

The previous NMR study on contulakin-G and analogues **6** – **9** (paper III) together with the present work on **10** and **11** suggest that there is more local rigidity around the Thr segment, as in **7** and **8**, than near the Ser position, as in **9** – **11**. These different conformational preferences of Thr and Ser have previously been shown for other *O*-linked glycopeptides by NMR and molecular modelling studies.^{97,112} The galactosamine units at the Thr sites in these studies were in close contact with several peptide atoms, whereas the galactosamine units attached to Ser residues assumed more apical positions with the surrounding peptide backbones distanced from the sugar, allowing more rotational freedom around the glycosidic bond. On the other hand did another study find a Thr-linked glycan that was positioned perpendicular to the peptide backbone, which did not have a hydrogen bond to the peptide.¹¹⁵

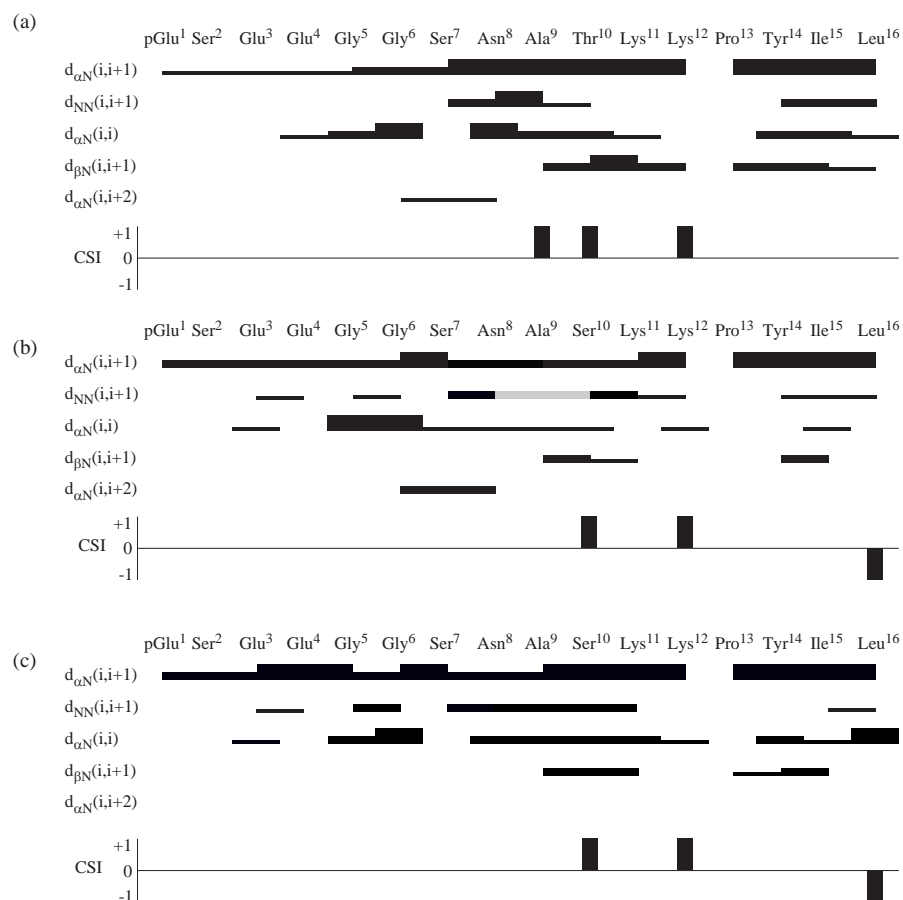


Figure 17. Intra- and inter-residual NOEs involving the backbone NH protons in contulakin-G, **7**, the analogue **10** and **11**, respectively. The thickness of the bars indicates the relative intensity of the NOE cross-peaks. A greyish NOE denotes that the signals were found to be overlapping. (a) **10**, NOESY, 5 °C, mixing time 300 ms; (b) **11**, NOESY, 5 °C, mixing time 300 ms

The NMR data for **10** and **11** together with models obtained by MM2 (Figure 19), suggested that, while there was no hydrogen bond interaction in the L- and D-serine analogues, the contacts between the peptide chain and the glycan were closer with a D-serine than with a L-serine. Because of the difference in chirality of the Ser10 residue of **10** as compared to **11**, the backbone from Ser10 to the amino acid 16 extended in opposite directions relative to the glycan in these two compounds. The disaccharide was bent away from the peptide chain in the [L-Ser7]-enantiomer **11**, whereas in the [D-Ser7]-enantiomer **10**, the glycan was positioned over the Lys10-Leu16 chain.

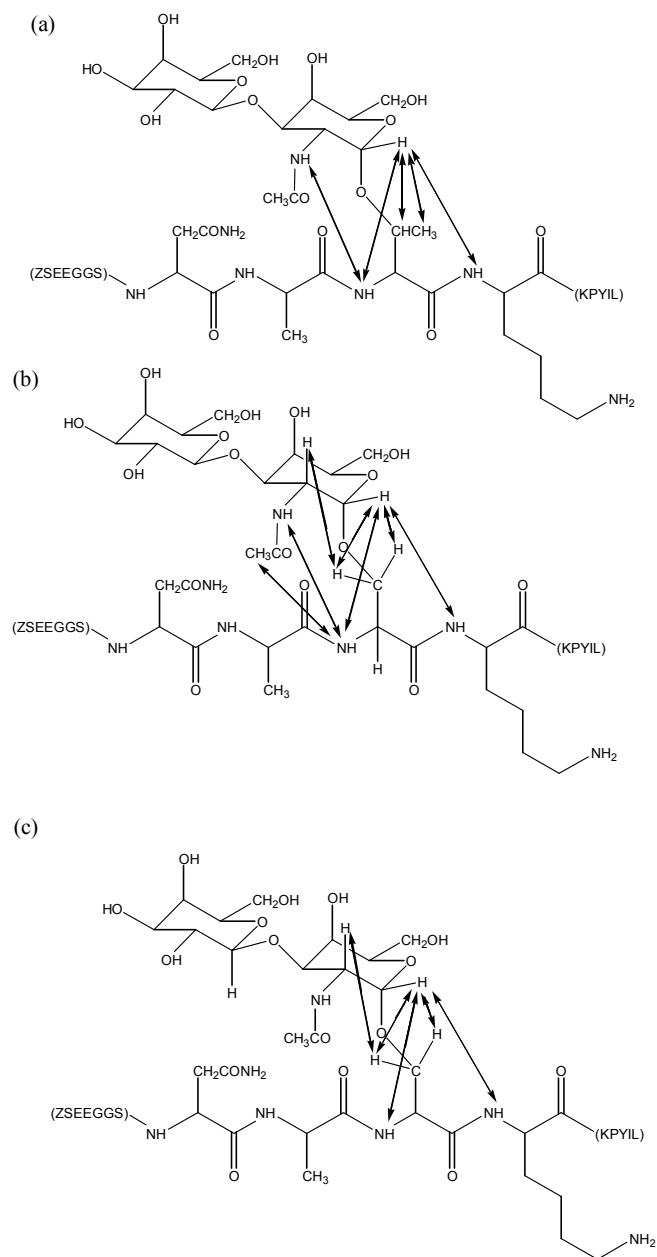


Figure 18. Schematic representation of interresidual NOEs observed at the junction between the peptide and the saccharides in contulakin-G, **7**, the [D-Ser10] analogue, **10**, and the [L-Ser10] analogue, **11**.

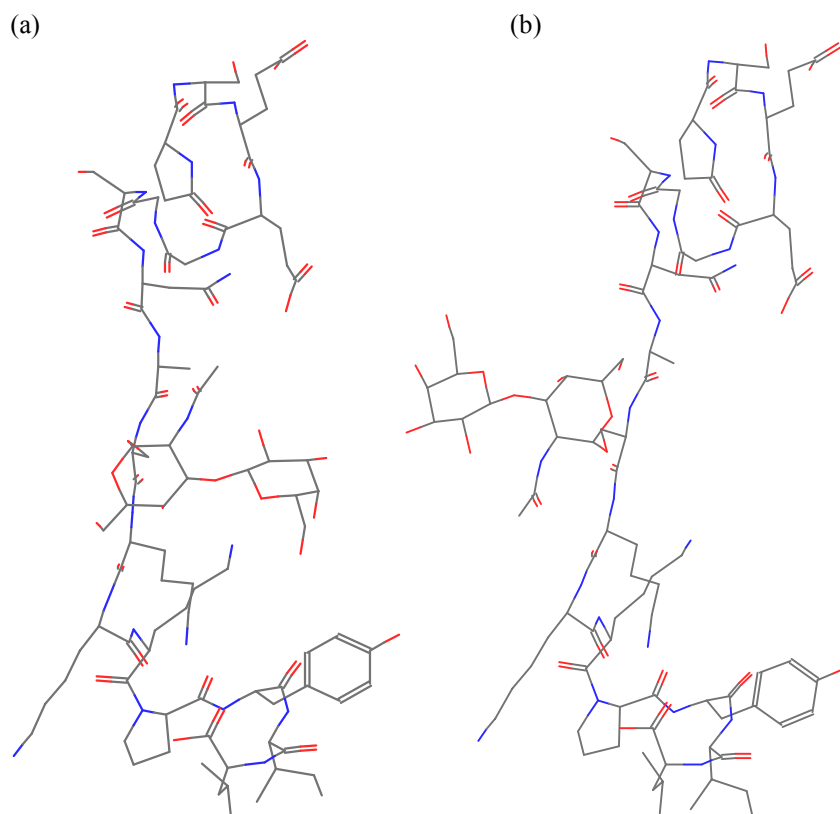


Figure 19. A 3D representation of (a) [D-Ser10]-contulakin-G, **10**, and (b) [L-Ser10]-contulakin-G, **11**.

2.7.1. Conclusion

Due to racemization, the synthesis of [Ser10]-contulakin-G yielded two enantiomers that have been characterised by ^1H NMR spectroscopy and GC-MS as the [D-Ser10]- and [L-Ser10]-contulakin-G analogues. Preliminary biological data indicated that [D-Ser10]-contulakin-G, **10**, had the same biological activity as contulakin-G, **7**, whereas the [L-Ser10]-contulakin-G, **11**, only was slightly active at 100 times the dose. The absence of strong $d_{\text{NN}}(i,i+1)$ NOEs together with the proton temperature coefficients and $^3J_{\text{NH},\alpha\text{H}}$ -values indicated that the peptide chain in glycopeptides **10** and **11** predominantly existed in random coil or extended conformations. The NMR data also showed that there was no hydrogen bond between the peptide chain and the glycan. Thus, the hydrogen bond previously found in contulakin-G,⁹⁰ does not seem to be of crucial importance for the biological activity. In contulakin-G, **7**, and in the [D-Ser10]-analogue **10**, the NH(Ser10)-NH(GalpNAc), $\alpha\text{H}(\text{Glu6})\text{-NH}(\text{Asn8})$, and consecutive NH(Tyr14)-NH(Ile15) and NH(Ile15)-NH(Leu16) NOEs were found, but these NOEs were not found in the [L-Ser10] analogue, **11**. These differences in NOEs, together with the larger $^3J_{\text{NH},\alpha\text{H}}$ -value of 8.6 Hz for Thr10 in contulakin-G, **7**, and of 8.5 Hz for D-Ser10 in **10**, suggest that the ensemble average conformation of the [D-Ser10]-

enantiomer **10** resembles the conformation of contulakin-G, **7**, more than the [L-Ser10]-enantiomer **11** does. Since a transient turn between residues 14 to 16 only was found in contulakin-G, **7**, and in [D-Ser10]-contulakin-G, **10**, and there is a correlation with a high biological activity for these substances, this indicated that this transient turn might be of importance for the biological activity. Previously, Craig *et al.* concluded that *O*-glycosylation was a strategy of contulakin-G to enhance the binding affinity to neurotensin receptors.⁸⁹ If the biological activity were due to binding of the glycopeptides to neurotensin receptors, and this transient C-terminal turn would be beneficial in the binding to these receptors, the turn would probably also have been found in conformational studies of this neurotensin. This is, however, not the case.^{136,137} Previous proteolytic degradation studies showed that the glycan sterically inhibited the enzyme from hydrolysing the Pro13-Tyr14 bond, and that this was a likely explanation to the high biological activity of contulakin-G.⁹³ Since the difference in stereochemistry of the glycosylated Ser10-residues, which yielded a difference in orientation of the sugar relative to the backbone, was the only difference between [D-Ser10]-contulakin-G, with high biological activity, and [L-Ser10]-contulakin-G, with lower biological activity, this indicates that the orientation of the sugar relative to the peptide is important in order to protect against proteolytic degradation.

Concluding remarks

NMR investigations of hydroxyl protons in aqueous solution have so far been devoted almost exclusively to the hydroxyl protons of saccharides. The work in papers I and II has shown that it is possible to observe hydroxy protons of saccharides that are conjugated to amino acids and peptides. Information could be obtained on the existence of hydrogen bonds, and on the hydration of the hydroxy groups. The investigations performed in papers I and II showed promising results for the future use of hydroxy protons as conformational probes in ^1H NMR studies of glycopeptides and glycoproteins.

The NMR study in paper III was initiated to try to correlate the increased activity of the glycopeptide contulakin-G, as compared to its unglycosylated counterpart, with its structure. A transient intramolecular hydrogen bond between the sugar and the peptide that stabilised the structures of the biologically most active compounds was found. Later, a study of proteolytic degradation was performed by Craig *et al.*, which showed that the size and position of the glycan was responsible for the differences in rate of degradation by the enzyme. The rate of degradation was paralleled by the biological activity. Hence it was proposed that the glycan reduced the rate of degradation, thereby causing a higher biological activity of contulakin-G as compared to its analogues. In paper IV, two analogues with different configuration of the amino acid at the glycosylation site were studied. Preliminary biological data, together with NMR data suggested that the hydrogen bond of contulakin-G did not contribute significantly to its high biological activity. Since the glycans of the two analogues have different orientations with respect to the peptide chain, and different degrees of biological activity, it was suggested that the orientation of the glycan to the peptide chain is that which is recognised by the proteolytic enzyme. The studies in paper II and IV showed interesting results on the structure-activity relationship for contulakin-G, which is currently in phase II clinical trials for short-term management of post-operative pain.

The research performed in this thesis has yielded ideas for further research that are listed below.

Suggestions for further research

- Perform molecular modelling of the [D-Ser10] and [L-Ser10] analogues of contulakin-G, as a further support for the conformational data from NMR.
- Utilise the hydroxy protons as an aid in conformational studies of contulakin-G and analogues.
- Synthesise a contulakin-G [L-Ser] and [D-Ser] analogue where the Pro13-Tyr14 is non-hydrolysable. This analogue could then be used for a trNOE or STD NMR study of how the proteolytic enzyme binds to the analogues, in order to explain the differences in biological activity of **11** and **12**.

- Investigate how TFE/water and TFE solutions affect the conformational preferences of glycopeptides and saccharides.

Acknowledgement

I would like to thank **all former and present colleagues** for an enjoyable time at the Department of Chemistry. I would especially like to acknowledge some persons who in different ways have contributed to realizing this thesis.

I am deeply obliged to my assisting supervisor, **Professor Lennart Kenne**, for the opportunity to work with NMR spectroscopy in the field of carbohydrate chemistry. By providing the department with the necessary instrumentation, and by teaching and administrating Ph.D. courses, he has supplied the basic necessities to accomplish this thesis. His experience in carbohydrate chemistry has been much appreciated, especially his careful proof-reading of all of my manuscripts, and valuable advice on research projects.

I'm privileged to be the first Ph.D. student under the supervision of **Associate Professor Corine Sandström**. Her enthusiasm and sincere interest in science has been a source of inspiration. Her disciplined and structured way of working, combined with curiosity and technical skills, yielded projects that were well defined, realisable and interesting. She has always been available for discussions—including evenings, holidays, and even during her maternity leave. Her personal concern for me has made us more than colleagues. I would especially like to thank her for the loads of small pyjamas and other baby clothes that she donated to my son Jon.

I am grateful to my first supervisor, **Professor Thomas Norberg**, for accepting me as a Ph.D. student and introducing me to carbohydrate synthesis.

Ph.D. Anthony Craig introduced research on conotoxins to our lab and was responsible for the work performed at the SALK institute. He also put an effort into learning how to synthesise the glycopeptides. It has been most enjoyable to work in this field, not the least the structure-activity part, which I truly find interesting.

I would also like to thank the following people for their support and help in different ways during the years

Former head of department, **Lennart Lundgren** helped me find an excellent solution to a difficult problem. Without him there would not have been a thesis. **Bernt Andersson** has been of invaluable assistance in the lab when teaching undergraduate students. **Lasse Andersson** funded The Pub, a bright move. **Rolf Andersson** introduced me to the Bruker NMR machines and has always been available for help with technical difficulties and with the interpretation of spectra. **Anders Broberg** helped with the GC-MS method. **Jan Eriksson** read and commented part one of the thesis manuscript. **Lena Johansson**, has been indispensable in assisting with administrative issues, and in the coordination of social activities. **Christian Marcu** and **Tom Karlsson**, have kept network, computers, servers, and printers running. **My room mates** through the years; **Susanna Broberg**, **Helena Almer**, **Ianric Ivarsson** and **Somer Bekiroglu** have been good and amusing company. **Lars Nord** has always been helpful in teaching matters and with tips and tricks with instrumentation and programs. **Nils** and

Annika Nyberg gave helpful tips on Adobe Illustrator, and **Nils** also gave me an introduction to gradient shimming.

Sean Daugherty helped me with linguistic corrections (abstract, introduction, part 1, concluding remarks, and acknowledgement).

I also would like to thank:

The management of IAR Systems AB for the modern attitude towards parenthood, thereby providing a ‘dad-friendly’ working environment. **Mikael Sjödin** for advice on, and a trial of, speech recognition software. All **Members of Budohuset** (especially KFUM-KFUK Uppsala Kendo and KFUM-KFUK Uppsala Kyokushinkai) for showing budo good spirit, and giving me loads of inspiration. *Domo arigato!* **Agneta Wallmon** and **Ingmar Holmberg** for spare time tango distractions. **My Friends** for being friends. **My family** for support during the years, especially **Mats**, my love, for good food and distraction, and **Jon**, my son, for being my joy.

References

- (1) Lindhorst, T. K. *Essentials of Carbohydrate Chemistry and Biochemistry*; WILEY-VCH Verlag GmbH: Weinheim, 2000.
- (2) Reuter, G.; Gabius, H. J. *Cell. Mol. Life Sci.* **1999**, *55*, 368-422.
- (3) Opdenakker, G.; Rudd, P. M.; Ponting, C. P.; Dwek, R. A. *FASEB J.* **1993**, *7*, 1330-1337.
- (4) Hounsell, E. F.; Davies, M. J.; Renouf, D. V. *Glycoconjugate J.* **1996**, *13*, 19-26.
- (5) Van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. *Crit. Rev. Biochem. Mol. Biol.* **1998**, *33*, 151-208.
- (6) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
- (7) Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, *218*, 1-27.
- (8) Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, 637-674.
- (9) Varki, A. *Glycobiology* **1993**, *3*, 97-130.
- (10) Hanisch, F. A. *Biol. Chem.* **2001**, *382*, 143-149.
- (11) Hanover, J. A. *FASEB J.* **2001**, *15*, 1865-1876.
- (12) Elhammer, A. P.; Poorman, R. A.; Brown, E.; Maggiora, L. L.; Hoogerheide, J. G.; Kezdy, F. J. *J. Biol. Chem.* **1993**, *268*, 10029-10038.
- (13) Hansen, J. E.; Lund, O.; Engelbrecht, J.; Bohr, H.; Nielsen, J. O.; Hansen, J. E. S.; Brunak, S. *Biochem. J.* **1995**, *308*, 801-813.
- (14) Fukuda, M. In *Molecular Glycobiology*; Fukuda, M., Hindsgaul, O., Eds.; Oxford University Press: Oxford, 1994.
- (15) Imperiali, B.; O'Connor, S. E. *Curr. Opin. Chem. Biol.* **1999**, *3*, 643-649.
- (16) Kornfeld, R.; Kornfeld, S. *Ann. Rev. Biochem.* **1985**, *54*, 631-64.
- (17) Homans, S. W. In *Molecular Glycobiology*; Fukuda, M., Hindsgaul, O., Eds.; Oxford University Press: Oxford, 1994.
- (18) Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371-386.
- (19) Bush, C. A.; Martin-Pastor, M.; Imberty, A. *Annu. Rev. Biophys. Biomolec. Struct.* **1999**, *28*, 269-293.
- (20) Woods, R. J. *Glycoconjugate J.* **1998**, *15*, 209-216.
- (21) Imberty, A.; Perez, S. *Chem. Rev.* **2000**, *100*, 4567-4588.
- (22) Pérez, S.; Gauthier, C.; Imberty, A. In *Oligosaccharides in Chemistry and Biology: A Comprehensive Handbook*; Ernst, B., Hart, G., Sinay, P., Eds.; Wiley/VCH: Weinheim, 2000.
- (23) Duus, J. O.; Gotfredsen, C. H.; Bock, K. *Chem. Rev.* **2000**, *100*, 4589-4614.
- (24) Agrawal, P. K. *Phytochemistry* **1992**, *31*, 3307-3330.
- (25) Tvaroska, I.; Taravel, F. R. In *Advances in Carbohydrate Chemistry and Biochemistry, Vol 51*, 1995; Vol. 51.
- (26) Bose, B.; Zhao, S.; Stenutz, R.; Cloran, F.; Bondo, P. B.; Bondo, G.; Hertz, B.; Carmichael, I.; Serianni, A. S. *J. Am. Chem. Soc.* **1998**, *120*, 11158-11173.
- (27) Wooten, E. W.; Bazzo, R.; Edge, C. J.; Zamze, S.; Dwek, R. A.; Rademacher, T. W. *Eur. Biophys. J.* **1990**, *18*, 139-148.
- (28) Neuhaus, D.; Williamsson, M. P. *The nuclear Overhauser effect in structural and conformational analysis*; VCH Publishers, Inc., 1989.

- (29) Kiddle, G. R.; Harris, R.; Homans, S. W. *J. Biomol. NMR* **1998**, *11*, 289-294.
- (30) Poppe, L.; York, W. S.; Vanhalbeek, H. *J. Biomol. NMR* **1993**, *3*, 81-89.
- (31) Milton, M. J.; Harris, R.; Probert, M. A.; Field, R. A.; Homans, S. W. *Glycobiology* **1998**, *8*, 147-153.
- (32) Rundlof, T.; Kjellberg, A.; Damberg, C.; Nishida, T.; Widmalm, G. *Magn. Reson. Chem.* **1998**, *36*, 839-847.
- (33) Rundlof, T.; Landersjö, C.; Lycknert, K.; Maliniak, A.; Widmalm, G. *Magn. Reson. Chem.* **1998**, *36*, 773-776.
- (34) Harvey, J. M.; Symons, M. C. R. *Nature* **1976**, *261*, 435-436.
- (35) Symons, M. C. R.; Benbow, J. A.; Harvey, J. M. *Carbohydr. Res.* **1980**, *83*, 9-20.
- (36) Piotto, M.; Saudek, V.; Sklenar, V. *J. Biomol. NMR* **1992**, *2*, 661-665.
- (37) Adams, B.; Lerner, L. E. *Magn. Reson. Chem.* **1994**, *32*, 225-230.
- (38) Poppe, L.; Vanhalbeek, H. *J. Am. Chem. Soc.* **1991**, *113*, 363-365.
- (39) Vanhalbeek, H.; Poppe, L. *Magn. Reson. Chem.* **1992**, *30*, S74-S86.
- (40) Poppe, L.; Stuikeprill, R.; Meyer, B.; Vanhalbeek, H. *J. Biomol. NMR* **1992**, *2*, 109-136.
- (41) Poppe, L.; Vanhalbeek, H. *J. Am. Chem. Soc.* **1992**, *114*, 1092-1094.
- (42) Brisson, J. R.; Uhrinova, S.; Woods, R. J.; vanderZwan, M.; Jarrell, H. C.; Paoletti, L. C.; Kasper, D. L.; Jennings, H. J. *Biochemistry* **1997**, *36*, 3278-3292.
- (43) Harris, R.; Rutherford, T. J.; Milton, M. J.; Homans, S. W. *J. Biomol. NMR* **1997**, *9*, 47-54.
- (44) Sandström, C.; Baumann, H.; Kenne, L. *J. Chem. Soc. Perkin Trans. 2* **1998**, 809-815.
- (45) Sandström, C.; Baumann, H.; Kenne, L. *J. Chem. Soc. Perkin Trans. 2* **1998**, 2385-2393.
- (46) Adams, B.; Lerner, L. *J. Am. Chem. Soc.* **1992**, *114*, 4827-4829.
- (47) Leeftang, B. R.; Vliegthart, J. F. G.; Kroonbatenburg, L. M. J.; Vaneijck, B. P.; Kroon, J. *Carbohydr. Res.* **1992**, *230*, 41-61.
- (48) Poppe, L.; Vanhalbeek, H. *Nat. Struct. Biol.* **1994**, *1*, 215-216.
- (49) Bundle, D. R.; Baumann, H.; Brisson, J. R.; Gagne, S. M.; Zdanov, A.; Cygler, M. *Biochemistry* **1994**, *33*, 5183-5192.
- (50) Sheng, S. Q.; Vanhalbeek, H. *Biochem. Biophys. Res. Commun.* **1995**, *215*, 504-510.
- (51) Woods, R. J. *Curr. Opin. Struct. Biol.* **1995**, *5*, 591-598.
- (52) Reuben, J. *J. Am. Chem. Soc.* **1984**, *106*, 6180-6186.
- (53) Christofides, J. C.; Davies, D. B.; Martin, J. A.; Rathbone, E. B. *J. Am. Chem. Soc.* **1986**, *108*, 5738-5743.
- (54) Claridge, D. W. C. In *High-Resolution NMR Techniques in Organic Chemistry*; Elsevier Science Ltd.: 1999; Vol. 19.
- (55) Skalicky, J. J.; Sukumaran, D. K.; Mills, J. L.; Szyperski, T. *J. Am. Chem. Soc.* **2000**, *122*, 3230-3231.
- (56) Batta, G.; Kover, K. E. *Carbohydr. Res.* **1999**, *320*, 267-272.
- (57) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546-4553.
- (58) Macura, S.; Ernst, R. R. *Mol. Phys.* **1980**, *41*, 95-117.

- (59) Kumar, A.; Wagner, G.; Ernst, R. R.; Wüthrich, K. *J. Am. Chem. Soc.* **1981**, *103*, 3654-3658.
- (60) Fraser, R. R.; Kaufman, P.; Morand, P.; Govil, G. *Can. J. Chem.* **1969**, *47*, 403-409.
- (61) Ivarsson, I.; Sandström, C.; Sandström, A.; Kenne, L. *J. Chem. Soc.-Perkin Trans. 2* **2000**, 2147-2152.
- (62) Bekiroglu, S.; Sandström, C.; Sandström, A.; Kenne, L. *manuscript*.
- (63) Roberts, G. C. K. *NMR of Macromolecules*; Oxford University Press Inc.: New York, 1993.
- (64) Sandström, C.; Magnusson, G.; Nilsson, U.; Kenne, L. *Carbohydr. Res.* **1999**, *322*, 46-56.
- (65) Bekiroglu, S.; Sandström, C.; Norberg, T.; Kenne, L. *Carbohydr. Res.* **2000**, *328*, 409-418.
- (66) Liang, R.; Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1995**, *117*, 10395-10396.
- (67) Andreotti, A. H.; Kahne, D. *J. Am. Chem. Soc.* **1993**, *115*, 3352-3353.
- (68) Aida, M.; Sugawara, Y.; Oikawa, S.; Umemoto, K. *Int. J. Biol. Macromol.* **1995**, *17*, 227-235.
- (69) Espinosa, J. F.; Asensio, J. L.; Bruix, M.; JimenezBarbero, J. *An. Quim.* **1996**, *92*, 320-324.
- (70) Davis, J. T.; Hirani, S.; Bartlett, C.; Reid, B. R. *J. Biol. Chem.* **1994**, *269*, 3331-3338.
- (71) Ohnishi, H.; Nishida, Y.; Hori, H.; Meguro, H.; Zushi, S. *J. Carbohydr. Chem.* **1988**, *7*, 711-731.
- (72) Homans, S. W.; DeVries, A. L.; Parker, S. B. *FEBS Lett.* **1985**, *183*, 133-137.
- (73) Bush, C. A.; Feeney, R. E. *Int. J. Pept. Protein Res.* **1986**, *28*, 386-397.
- (74) Mimura, Y.; Yamamoto, Y.; Inoue, Y.; Chujo, R. *Int. J. Biol. Macromol.* **1992**, *14*, 242-248.
- (75) Neidig, K. P.; Geyer, M.; Gorler, A.; Antz, C.; Saffrich, R.; Beneicke, W.; Kalbitzer, H. R. *J. Biomol. NMR* **1995**, *6*, 255-270.
- (76) Matsuura, H.; Takio, K.; Titani, K.; Greene, T.; Levery, S. B.; Salyan, M. E. K.; Hakomori, S. *J. Biol. Chem.* **1988**, *263*, 3314-3322.
- (77) Lane, A. N.; Hays, L. M.; Tsvetkova, N.; Feeney, R. E.; Crowe, L. M.; Crowe, J. H. *Biophys. J.* **2000**, *78*, 3195-3207.
- (78) Franks, F. *Cryobiology* **1983**, *20*, 335-345.
- (79) Franks, F. *Pure Appl. Chem.* **1987**, *59*, 1189-1202.
- (80) Kindahl, L.; Sandström, C.; Norberg, T.; Kenne, L. *J. Carbohydr. Chem.* **2000**, *19*, 1291-1303.
- (81) Olivera, B. M.; Rivier, J.; Clark, C.; Ramilo, C. A.; Corpuz, G. P.; Abogadie, F. C.; Mena, E. E.; Woodward, S. R.; Hillyard, D. R.; Cruz, L. J. *Science* **1990**, *249*, 257-263.
- (82) Olivera, B. M. *Mol. Biol. Cell* **1997**, *8*, 2101-2109.
- (83) McIntosh, J. M.; Jones, R. M. *Toxicon* **2001**, *39*, 1447-1451.
- (84) Jones, R. M.; Bulaj, G. *Curr. Opin. Drug Discov.* **2000**, *3*, 141-154.
- (85) Jones, R. M.; Cartier, G. E.; McIntosh, J. M.; Bulaj, G.; Farrar, V. E.; Olivera, B. M. *Expert Opin. Ther. Patents* **2001**, *11*, 603-623.

- (86) Olivera, B. M.; Walker, C.; Cartier, G. E.; Hooper, D.; Santos, A. D.; Schoenfeld, R.; Shetty, R.; Watkins, M.; Bandyopadhyay, P.; Hillyard, D. R. In *Molecular Strategies in Biological Evolution*, 1999; Vol. 870.
- (87) Lirazan, M. B.; Hooper, D.; Corpuz, G. P.; Ramilo, C. A.; Bandyopadhyay, P.; Cruz, L. J.; Olivera, B. M. *Biochemistry* **2000**, *39*, 1583-1588.
- (88) Craig, A. G.; Bandyopadhyay, P.; Olivera, B. M. *Eur. J. Biochem.* **1999**, *264*, 271-275.
- (89) Craig, A. G.; Norberg, T.; Griffin, D.; Hoeger, C.; Akhtar, M.; Schmidt, K.; Low, W.; Dykert, J.; Richelson, E.; Navarro, V.; Mazella, J.; Watkins, M.; Hillyard, D.; Imperial, J.; Cruz, L. J.; Olivera, B. M. *J. Biol. Chem.* **1999**, *274*, 13752-13759.
- (90) Kindahl, L.; Sandström, C.; Craig, A. G.; Norberg, T.; Kenne, L. *Can. J. Chem.* **2002**, *80*, 1022-1031.
- (91) Wagstaff, J. D.; Layer, R. T.; Craig, A. G.; Olivera, B. M.; McCabe, R. T. *Proc. 29th Annu. Meet. Soc. Neurosci.* **1999**, *25*, 1944.
- (92) Polt, R.; Palian, M. M. *Drug Future* **2001**, *26*, 561-576.
- (93) Craig, A. G.; Norberg, T.; Kang, J.; Wagstaff, J.; Layer, R.; Imperial, J.; Olivera, B. M. *manuscript*.
- (94) Hylden, J. L. K.; Wilcox, G. L. *Eur. J. Pharmacol.* **1980**, *67*, 313-316.
- (95) Dyson, H. J.; Rance, M.; Houghten, R. A.; Lerner, R. A.; Wright, P. E. *J. Mol. Biol.* **1988**, *201*, 161-200.
- (96) Dyson, H. J.; Wright, P. E. *Ann. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519-538.
- (97) Naganagowda, G. A.; Gururaja, T. L.; Satyanarayana, J.; Levine, M. J. *J. Pept. Res.* **1999**, *54*, 290-310.
- (98) Agrawal, P. K.; Jacquinet, J. C.; Krishna, N. R. *Glycobiology* **1999**, *9*, 669-677.
- (99) Arsequell, G.; Haurum, J. S.; Elliott, T.; Dwek, R. A.; Lellouch, A. C. *J. Chem. Soc.-Perkin Trans. 1* **1995**, 1739-1745.
- (100) Liu, X. H.; Sejbál, J.; Kotovych, G.; Koganty, R. R.; Reddish, M. A.; Jackson, L.; Gandhi, S. S.; Mendonca, A. J.; Longenecker, B. M. *Glycoconjugate J.* **1995**, *12*, 607-617.
- (101) Wyss, D. F.; Choi, J. S.; Li, J.; Knoppers, M. H.; Willis, K. J.; Arulanandam, A. R. N.; Smolyar, A.; Reinherz, E. L.; Wagner, G. *Science* **1995**, *269*, 1273-1278.
- (102) Dwek, R. A.; Lellouch, A. C.; Wormald, M. R. *J. Anat.* **1995**, *187*, 279-292.
- (103) Braun, P.; Davies, G. M.; Price, M. R.; Williams, P. M.; Tandler, S. J. B.; Kunz, H. *Bioorg. Med. Chem.* **1998**, *6*, 1531-1545.
- (104) Maeji, N. J.; Inoue, Y.; Chujo, R. *Int. J. Pept. Protein Res.* **1987**, *29*, 699-707.
- (105) Simanek, E. E.; Huang, D. H.; Pasternack, L.; Machajewski, T. D.; Seitz, O.; Millar, D. S.; Dyson, H. J.; Wong, C. H. *J. Am. Chem. Soc.* **1998**, *120*, 11567-11575.
- (106) McManus, A. M.; Otvos, L.; Hoffmann, R.; Craik, D. J. *Biochemistry* **1999**, *38*, 705-714.
- (107) Bailey, D.; Renouf, D. V.; Large, D. G.; Warren, C. D.; Hounsell, E. F. *Carbohydr. Res.* **2000**, *324*, 242-254.

- (108) Live, D. H.; Kumar, R. A.; Beebe, X.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 12759-12761.
- (109) Wu, W. G.; Pasternack, L.; Huang, D. H.; Koeller, K. M.; Lin, C. C.; Seitz, O.; Wong, C. H. *J. Am. Chem. Soc.* **1999**, *121*, 2409-2417.
- (110) Jentoft, N. *Trends Biochem. Sci.* **1990**, *15*, 291-294.
- (111) Huang, X. L.; Smith, M. C.; Berzofsky, J. A.; Barchi, J. J. *FEBS Lett.* **1996**, *393*, 280-286.
- (112) Huang, X. L.; Barchi, J. J.; Lung, E. D. T.; Roller, P. P.; Nara, P. L.; Muschik, J.; Garrity, R. R. *Biochemistry* **1997**, *36*, 10846-10856.
- (113) O'Connor, S. E.; Imperiali, B. *Chem. Biol.* **1998**, *5*, 427-437.
- (114) Rickert, K. W.; Imperiali, B. *Chem. Biol.* **1995**, *2*, 751-759.
- (115) Kirnarsky, L.; Prakash, O.; Vogen, S. M.; Nomoto, M.; Hollingsworth, M. A.; Sherman, S. *Biochemistry* **2000**, *39*, 12076-12082.
- (116) Buck, M. *Q. Rev. Biophys.* **1998**, *31*, 297-355.
- (117) Rajan, R.; Balaram, P. *Int. J. Pept. Protein Res.* **1996**, *48*, 328-336.
- (118) CammersGoodwin, A.; Allen, T. J.; Oslick, S. L.; McClure, K. F.; Lee, J. H.; Kemp, D. S. *J. Am. Chem. Soc.* **1996**, *118*, 3082-3090.
- (119) Gobbo, M.; Nicotra, A.; Rocchi, R.; Crisma, M.; Toniolo, C. *Tetrahedron* **2001**, *57*, 2433-2443.
- (120) Molinaro, A.; DeCastro, C.; Lanzetta, R.; Manzo, E.; Parrilli, M. *J. Am. Chem. Soc.* **2001**, *123*, 12605-12610.
- (121) Wütrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley & Sons, Inc., 1986.
- (122) Wagner, G.; Neuhaus, D.; Worgotter, E.; Vasak, M.; Kagi, J. H. R.; Wuthrich, K. *J. Mol. Biol.* **1986**, *187*, 131-135.
- (123) Wishart, D. S.; Nip, A. M. *Biochem. Cell Biol.* **1998**, *76*, 153-163.
- (124) Wishart, D. S.; Sykes, B. D.; Richards, F. M. *J. Mol. Biol.* **1991**, *222*, 311-333.
- (125) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. *J. Biomol. NMR* **1995**, *5*, 67-81.
- (126) Dalgarno, D. C.; Levine, B. A.; Williams, R. J. P. *Bioscience Rep.* **1983**, *3*, 443-452.
- (127) Wishart, D. S.; Sykes, B. D. In *Methods in Enzymology*; publisher:, 1994; Vol. 239.
- (128) Wishart, D. S.; Sykes, B. D. *J. Biomol. NMR* **1994**, *4*, 171-180.
- (129) Jimenez, M. A.; Nieto, J. L.; Rico, M.; Santoro, J.; Herranz, J.; Bermejo, F. *J. J. Mol. Struct.* **1986**, *143*, 435-438.
- (130) Zimmermann, G. R.; Legault, P.; Selsted, M. E.; Pardi, A. *Biochemistry* **1995**, *34*, 13663-13671.
- (131) Merutka, G.; Dyson, H. J.; Wright, P. E. *J. Biomol. NMR* **1995**, *5*, 14-24.
- (132) Andersen, N. H.; Neidigh, J. W.; Harris, S. M.; Lee, G. M.; Liu, Z. H.; Tong, H. *J. Am. Chem. Soc.* **1997**, *119*, 8547-8561.
- (133) Baxter, N. J.; Williamson, M. P. *J. Biomol. NMR* **1997**, *9*, 359-369.
- (134) Ström, K.; Sjogren, J.; Broberg, A.; Schnürer, J. *Appl. Environ. Microbiol.* **2002**, *68*, 4322-4327.
- (135) Marion, D.; Wütrich, K. *Biochem. Biophys. Res. Commun.* **1983**, *113*, 967-974.

- (136) Nieto, J. L.; Rico, M.; Santoro, J.; Herranz, J.; Bermejo, F. J. *Int. J. Pept. Protein Res.* **1986**, *28*, 315-323.
- (137) Xu, G. Y.; Deber, C. M. *Int. J. Pept. Protein Res.* **1991**, *37*, 528-535.
- (138) St-Pierre, S.; J.-M., L.; Gendreau, M.; Quiron, R.; Regoli, D.; Rioux, F. *J. Med. Chem.* **1981**, *24*, 370-376.
- (139) Rivier, J. E.; Lazarus, L. H.; Perrin, M. H.; Brown, M. R. *J. Med. Chem.* **1977**, *20*, 1409-1412.
- (140) Banks, W. A.; Wustrow, D. J.; Cody, W. L.; Davis, M. D.; Kastin, A. J. *Brain Res.* **1995**, *695*, 59-63.
- (141) Checler, F.; Mazella, J.; Kitabgi, P.; Vincent, J. P. *J. Neurochem.* **1986**, *47*, 1742-1748.