

Aus dem Institut für Chemie
in der
Technisch-Naturwissenschaftlichen Fakultät
der Universität zu Lübeck

Direktor:
Prof. Dr. rer. nat. Thomas Peters

**Binding of Carbohydrate Ligands to Siglecs, Sialoadhesin and
Myelin-associated Glycoprotein (MAG): Studied by NMR
Experiments and Computational Modelling**

Inauguraldissertation
zur
Erlangung der Doktorwürde
der Universität zu Lübeck
- Aus der Technisch-Naturwissenschaftlichen Fakultät -

Vorgelegt von
Anirban Bhunia
aus Indien

Lübeck, 2005

Dekan : Prof. Dr. E. Hartmann

1. Berichterstatter : Prof. Dr. T. Peters

2. Berichterstatter : Prof. Dr. P. K. Müller

Tag der mündlichen Prüfung: 26.05.2005

Zum Druck genehmigt : Lübeck, den 26.05.2005

To

*Maa and Baba
whom I owe a lot.....*

Zusammenfassung

Siglecs (sialic acid binding immunoglobulin like lectins) sind eine Familie von I-Typ Lektinen, die an Sialinsäuren auf Zelloberflächen binden. Das Sialoadhesin, Siglec-1, wird besonders stark in entzündeten Makrophagen exprimiert und bindet spezifisch an α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-GlcNAc Reste von Zuckerketten. Es wird angenommen, dass das Disaccharid α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal die Haupteckennungsstruktur für das Sialoadhesin ist. Um die Grundlagen dieses biologischen Erkennungsprozesses besser zu verstehen, wurden Sättigungs-Transfer-Differenz (STD) NMR Experimente durchgeführt und die Bindungs epitope von Liganden an Sialoadhesin mit atomarer Auflösung charakterisiert. Durch Kombination mit *Molecular Modelling* wurde ein zuverlässiges Modell für das Sialoadhesin in Komplex mit den Zuckern in wässriger Lösung erstellt.

Das myelin-assoziierte Glycoprotein (MAG), ein weiteres Mitglied der Siglecs (Siglec-4), befindet sich auf der Oberfläche von Nervenzellen und ist ein spezifischer Inhibitor des Neuritenwachstums nach einer Verletzung des Zentralen Nervensystems (ZNS). Diese inhibitorische Funktion kann durch Ganglioside blockiert werden, die terminal α -(2 \rightarrow 3)-verknüpfte Sialinsäuren tragen, z.B. GQ1ba oder GT1b. Deshalb wurde die Konformation von synthetischen GQ1ba Fragmenten gebunden an MAG mittels *transferred* NOE (trNOE) Analysen studiert, sowie zusätzlich die Bindungs epitope mittels STD Experimenten bestimmt. Ein Homologie-Modell von MAG erlaubte das Docking der Saccharid-Liganden in die hypothetische Bindungstasche von MAG. Auf der Basis dieser Modelle wurde mit CORCEMA Berechnungen eine quantitative Analyse der gebundenen Konformationen der Oligosaccharide durchgeführt. Diese Studie zeigt, dass alle Saccharid-Liganden in der gleichen Weise an MAG binden, wie Sialyl Lewis^x an E-Selektin. Die zusätzliche Carboxylgruppe an der α -(2 \rightarrow 6)-Bindung verstärkt die Bindung um den Faktor 4. Docking-Studien mit AutoDock 3.0 lassen vermuten, dass die verstärkte Bindungsaffinität auf einer starken elektrostatischen Interaktion zwischen der Carboxylgruppe an der α -(2 \rightarrow 6)-Bindung des Oligosaccharides und den Protonen des N^ε von Lys 67 des MAG basiert. Die Mutation des Lys 67 zu Alanin (K67A) führte zu einer signifikanten Reduktion der Bindung. Diese starke Abnahme der Aktivität ist vermutlich durch den Verlust der elektrostatischen Wechselwirkungen zwischen den Protonen des N^ε von Lys 67 und der Carboxylgruppe vom α -(2 \rightarrow 6)-verknüpften Neu5Ac/LAc zu erklären.

Somit war das Ziel dieser Arbeit die systematische Bestimmung der bioaktiven Konformationen und der Bindungs epitope verschiedener Oligosaccharide in der Gegenwart von Siglecs, Sialoadhesin und MAG.

Abstract

The siglecs (sialic acid binding immunoglobulin like **lectins**) are a family of I-type lectins that bind to sialic acid on the cell surface. Sialoadhesin, siglec-1 is expressed at much higher levels in inflammatory macrophages and specifically binds to α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-GlcNAc residues of glycan chains; the terminal disaccharide of which α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal is thought to be the main epitope recognized by sialoadhesin. To understand the basis of this biological recognition reaction, saturation transfer difference (STD) NMR experiments were employed to characterize the binding epitope of ligands to sialoadhesin at atomic resolution. Consequently, NMR experiments in combination with molecular modelling yields a reliable structural model for the complex of sialoadhesin with the sugars in aqueous solution.

Myelin-associated glycoprotein (MAG), another siglecs member (siglec-4) is present on the surface of nerve cells, and is a specific inhibitor of neurite outgrowth after injury of the central nervous system (CNS). This inhibitory function of MAG can be blocked by gangliosides containing terminal α -(2 \rightarrow 3)-linked sialic acids, e.g. GQ1b α or GT1b. Therefore, the bound conformation of synthetic fragments of GQ1b α to MAG were studied employing transferred NOE (trNOE) analysis as well the group epitope mapping of the same ligands by STD NMR experiments. A homology model of MAG allowed the docking of the saccharide ligands into the hypothetical binding pocket of MAG. On the basis of these models a quantitative analysis of the bound conformations of the oligosaccharides were performed using CORCEMA calculations. This study indicates that all saccharides bind to MAG in the same way as Sialyl Lewis^x binds to E-selectin. The additional carboxyl group at the α -(2 \rightarrow 6)-linkage of increases the binding affinity by a factor of 4. Docking with AutoDock 3.0 suggest this increase in binding affinity is due to a strong electrostatic interaction between carboxyl group of α -(2 \rightarrow 6)-linkage of oligosaccharides and the protons of N^ε of Lys 67 of MAG. Mutation of Lys 67 to an alanine (K67A), results in a significant reduction in binding, suggesting that this loss of activity is due to the loss of the electrostatic interaction between protons of N^ε of Lys 67 and the carboxylate group of the α -(2 \rightarrow 6)-linked Neu5Ac/LAc.

Consequently, the goal of this thesis is the systematic determination of bioactive conformation and binding epitopes of different oligosaccharides in the presence of siglecs, sialoadhesin and MAG.

Acknowledgements

I take this opportunity to express my heartiest gratitude to Professor Thomas Peters for his precious and innovative guidance, invaluable suggestions and creative encouragement. His dedication and devotion will always be a source of inspiration for me.

I must give thanks to Dr. Milos Hricovíni for his time and patience and explaining the concept of trNOE and CORCEMA.

I especially thank Dr. Andrew J. Benie for providing me valuable suggestions and ideas, which certainly helped in shapping up the present work.

I sincerely thank my colleagues Dr. Thorsten Biet, Dr. Christian Plath, Dr. Jesus Angulo Alvarez, Dr. Hannelore Peters and Mr. Richard Szilaghi for help, support and useful discussions, which made my stay in the laboratory a memorable experience.

I must thank to the other member of the laboratory Mr. Thies Köhli, Dr. E. Bäuml, Dr. Rosemarie Pulz for providing me the necessary support.

I am grateful to Dr. Astrid Blume for her suggestions and corrections to this manuscript and also translation of the Abstract in German.

I must thank “Graduiertenkolleg 464” for providing me the stipendium without which this work would not have been possible.

I sincerely thank all my friends who made me feel homely here in Lübeck.

I also thank my brother, Tutun and sister, Piu for their love and support. Also, I want to thank my other family members who supported me in their own way.

At last but not the least, I would like to thank my wife for her mental support during writing of this manuscript.

CONTENTS

ZUSAMMENFASSUNG.....	IV
ABSTRACT	V
ACKNOWLEDGEMENTS.....	VII
ABBREVIATIONS & SYMBOLS	X
ABBREVIATIONS.....	X
SYMBOLS.....	XII
1 INTRODUCTION	1
1.1 SIGLECS.....	1
1.1.1 <i>Siglec-1</i>	2
1.1.2 <i>Siglec-2</i>	3
1.1.3 <i>Siglec-4</i>	3
1.1.4 <i>Siglec-3 and related siglecs</i>	4
1.2 GANGLIOSIDES.....	5
1.3 THE NERVOUS SYSTEM	5
1.3.1 <i>Mechanism of suppression of axonal regeneration</i>	7
1.4 PROTEIN-LIGAND INTERACTION STUDIED BY NMR SPECTROSCOPY	8
SCOPE OF THE THESIS.....	10
2 EXPERIMENTAL SECTION	12
2.1 NMR EXPERIMENT.....	12
2.1.1 <i>Sample preparation</i>	12
2.1.1.1 α -(2→3)-sialyl lactose 1 and saccharides 2-6.....	12
2.1.1.2 TOCSY experiment	13
2.1.1.3 HSQC experiment.....	14
2.1.1.4 HMBC experiment	14
2.1.1.5 NOESY experiment.....	14
2.1.1.6 ROESY experiment with $T_{1\rho}$ spinlock.....	15
2.1.1.7 Saturation Transfer Difference experiment	15
2.1.1.8 Integration of NOESY spectra	16
2.1.2 <i>Complete Relaxation and Conformational Exchange Matrix Analysis</i>	17
2.2 MOLECULAR DYNAMICS SIMULATION	18
2.3 HOMOLOGY MODEL	20
2.4 AUTOMATED DOCKING	20

3	METHODS.....	22
3.1	NUCLEAR MAGNETIC RESONANCE SPECTROSCOPY	22
3.1.1	<i>The Nuclear Overhauser Effect</i>	22
3.1.1.1	transferred Nuclear Overhauser Effect	26
3.1.1.2	transferred Rotating frame Overhauser Effect	28
3.1.1.3	Saturation Transfer Difference NMR experiment.....	30
3.2	COMPLETE RELAXATION AND CONFORMATIONAL EXCHANGE MATRIX ANALYSIS	32
3.3	MOLECULAR DYNAMICS SIMULATION	34
3.4	AUTOMATED DOCKING	36
4	RESULTS AND DISCUSSION	38
4.1	SIALOADHESIN	38
4.1.1	<i>Group epitope mapping of α-(2\rightarrow3)-sialyl lactose 1 in presence of sialoadhesin.....</i>	38
4.1.1.1	STD NMR experiments.....	39
4.1.1.2	Flexible Dock calculations	43
4.1.1.3	Discussion	45
4.1.2	<i>Sialoadhesin-Trisaccharide 2 complex.....</i>	47
4.1.2.1	Free conformational analysis of trisaccharide 2.....	48
4.1.2.1.1	NMR experiments.....	48
4.1.2.1.2	Molecular mechanics	50
4.1.2.2	Conformational analysis of trisaccharide 2 bound to sialoadhesin	51
4.1.2.2.1	NOESY experiment of sialoadhesin-trisaccharide 2 complex	51
4.1.2.2.2	ROESY experiment of trisaccharide 2 in the presence of sialoadhesin	53
4.1.2.2.3	CORCEMA calculation.....	54
4.1.2.2.4	Epitope mapping	56
4.1.2.2.5	Modelling	59
4.1.2.3	Discussion	60
4.2	MYELIN-ASSOCIATED GLYCOPROTEIN.....	63
4.2.1	<i>Free conformational analysis of saccharides 2-6</i>	66
4.2.1.1	¹ H NMR experiments	66
4.2.1.2	NOESY experiments	70
4.2.1.3	Simulation	73
4.2.1.4	Discussion	78
4.2.2	<i>Bound conformations of saccharides 2-6</i>	80
4.2.2.1	NOESY experiments of saccharides 2-6 in complex with MAG	81
4.2.2.2	ROESY experiments of saccharides 2-6 in complex with MAG	94
4.2.2.3	AutoDock	96
4.2.2.4	CORCEMA calculation.....	98
4.2.2.5	Epitope mapping of 5 by STD NMR experiments.....	104
4.2.2.6	Ligand-protein contacts.....	106
4.2.2.7	Discussion	109

5	CONCLUSIONS	114
	APPENDIX	117
A.1	CHEMICAL SHIFTS FOR OLIGOSACCHARIDES 1-6	117
A.1.1	<i>¹H Chemical shift assignments for α-(2\rightarrow3)-sialyl lactose 1 at 303 K and 500 MHz</i>	<i>117</i>
A.1.2	<i>¹H Chemical shift assignments for trisaccharide 2 at 277 K and 500 MHz</i>	<i>117</i>
A.1.3	<i>¹H Chemical shift assignments for tetrasaccharide 3 at 288 K and 700 MHz</i>	<i>118</i>
A.1.4	<i>¹H Chemical shift assignments for tetrasaccharide 4 at 280 K and 500 MHz</i>	<i>118</i>
A.1.5	<i>¹H Chemical shift assignments for pseudo tetrasaccharide 5 at 288 K and 500 MHz</i>	<i>119</i>
A.1.6	<i>¹H Chemical shift assignments for trisaccharide 6 at 288 K and 500 MHz</i>	<i>119</i>
A.2	PULSEPROGRAM.....	120
A.2.1	<i>1D reference pulse sequence</i>	<i>120</i>
A.2.2	<i>1D STD</i>	<i>121</i>
A.2.3	<i>STD TOCSY.....</i>	<i>123</i>
A.2.4	<i>NOESY/trNOESY</i>	<i>128</i>
A.3	CORCEMA	130
A.3.1	<i>CORCEMA “command” file</i>	<i>130</i>
A.3.1.1	<i>Compilation for Mac OS X 10.3.....</i>	<i>131</i>
A.3.2	<i>CORCEMA atom list file.....</i>	<i>132</i>
A.3.3	<i>Comparison between experimental and theoretical trNOEs at different mixing times (ms)</i>	<i>133</i>
A.3.3.1	<i>Trisaccharide 2 – MAG complex.....</i>	<i>133</i>
A.3.3.2	<i>Tetrasaccharide 4 – MAG complex</i>	<i>133</i>
A.3.3.3	<i>Pseudo tetrasaccharide 5 – MAG complex.....</i>	<i>133</i>
A.3.3.4	<i>Trisaccharide 2 – sialoadhesin complex</i>	<i>134</i>
	BIBLIOGRAPHY	135

Abbreviations & Symbols

Abbreviations

1D/2D	One, Two dimensional
AutoDock	Automated Docking
CNS	Central Nervous System
CORCEMA	Complete Relaxation and Conformational Exchange Matrix Analysis
Cyc	Cyclic ether
DPFGSE	Double Pulsed Field Gradient Spin Echo
ER	Endoplasmic Reticulum
FlexiDock	Flexible Docking
Gal	Galactose
GalNAc	N-acetyl galactosamine
GEM	Group Epitope Mapping
Glc	Glucose
H3ax	H3axial
H3eq	H3equatorial
HMBC	Heteronuclear Multiple Bond Coherence
HSEA	Hard Sphere <i>Exo</i> -Anomeric Approach
HSQC	Heteronuclear Single Quantum Coherence
IgSF	Immunoglobulin Super Family
ITIM	Immunoreceptor Tyrosine based Inhibitory Motif
IL	Interleukin
K _D	Dissociation constant
LAc	Lactic acid
LGA	Lamarckian Genetic Algorithm
MAG	Myelin-associated Glycoprotein
MAG _{ct}	MAG Cytoplasmic Domain
MCS	Multilevel Coordinate Search
MD	Molecular Dynamics
ms	millisecond
Neu5Ac	N-acetylneuraminic acid
Ngr	Nogo-receptor
NOESY	Nuclear Overhauser and Exchange Spectroscopy
NMR	Nuclear Magnetic Resonance
OMgp	Oligodendrocyte Myelin Glycoprotein
p75 ^{NTR}	p75 Neutrophin Receptor
ppm	Parts Per Million
pH*	pH Uncorrected for Deuterium Isotope Effect
PNS	Peripheral Nervous System
rIP	Relative Inhibitory Potential
ROESY	Rotational Overhauser Effect Spectroscopy
s	second
Sn	Sialoadhesin
SCR	Structurally Conserved Region

siglecs	Sialic Acid Binding Ig like Lectins
SMP	Schwann Cell Myelin Protein
STD	Saturation Transfer Difference
SVR	Structurally Variable Region
TOCSY	Total Correlation Spectroscopy
trNOE	transferred Nuclear Overhauser Effect
trROE	transferred Rotating frame Overhauser Effect

Symbols

a_i	Accelaration of a particle
D	Dynamic Matrix
δ	Chemical Shift
F_i	Force Field
ϵ	Dielectric Constant
γ_x	Gyromagnetic Ratio of nucleus x
m_i	Mass of a particle
η_{\max}	Maximum NOE
K	Kinetic Matrix
k_{on}	On-rate Constant
k_{off}	Off-rate Constant
r	Distance
R	Relaxation Matrix
ρ	Dipolar Relaxation Rate Constant
σ	Cross Relaxation Rate Constant
τ_c	Rotational Correlation Time of a Molecule
τ_m	Mixing Time in NOESY and TOCSY experiments
$T_{1\text{sel}}$	Selective Spin Lattice Relaxation
V	Potential Energy Function
$W_{0\text{IS}}$	Zero Quantum Transition
$W_{1\text{I}}$	Single Quantum Transition
$W_{2\text{IS}}$	Double Quantum Transition
ω	Frequency
X_i	Coordinate

1 Introduction

All living cells are surrounded by a plasma membrane, which represents a barrier by its structure and permeability characteristics, against the external world. At the same time, it also controls the exchange of materials. Membranes are very complex in their structure with different multiple functions. They are involved in numerous metabolic and energy conversion processes such as photosynthesis and oxidative phosphorylation, movement procedures, growth and cell divisions and control the information flow between the cells and their environment. They also play an important role for the signal transduction. Membranes contain numerous lipids and proteins with different carbohydrate structures. These glycoconjugates interact with biomolecules and hence the signal flows between cells and their environment. Often the terminal components of glycoconjugate are sialic acid/*N*-acetyl neuraminic acid (Neu5Ac) that are involved in many biological processes such as the formation or masking of recognition determinants (Schauer et al., 1995) and stabilisation of glycoprotein structures (Rens-Domiano and Reisine, 1991). In mammalian cell-cell interactions, generally two major groups of proteins, selectins and siglecs (sialic acid binding immunoglobulin like lectins), interact with sialic acid in physiological and pathological processes.

1.1 Siglecs

Cell-cell interactions provide an important means of communication within the immune system. They are instrumental in the regulation of proliferation, differentiation and migration of immunocompetent cells (Varki et al., 1999). These interactions involve cell surface adhesion receptors, many of which have now been identified (van den Berg et al., 1992). Animal lectins belong to this class of receptors that mediate cell-cell and cell-matrix interactions (Rini and Lobsanov, 1999). The ligands for these lectins are generally oligosaccharides. Mammalian lectins have been classified into discrete families based on similarities in primary amino acid sequence. Among these are Ca^{2+} -dependent (C-type) lectins, galectins, mannose-6-phosphate binding lectins (P-type lectins) and immunoglobulin like lectins (I-type lectins). Many aspects of their structural and functional properties have already been established

(Crocker and Feizi, 1996). Siglecs, sialic acid binding immunoglobulin like lectins provide a group of cellular recognition molecules characterised by sequence homology with members of the immunoglobulin super family (Kelm et al., 1994; Crocker et al., 1998). This is a subgroup of I-type lectins characterised by an N-terminal V-set Ig domain that mediates sialic acid binding, followed by varying number of C2-set Ig domains (Kelm et al., 1994a; Kelm et al., 1996; Crocker et al., 1998). To date 11 different siglecs are identified in humans and are named as siglec-1, siglec-2 etc. depending on the date of their discovery (Figure 1.1). The siglecs can be divided into two subgroups; one represented by siglec-1 (sialoadhesin), siglec-2 (CD 22), MAG (siglec-4a) and SMP (siglec-4b), the other subgroup consists of siglec-3 (CD 33) and related siglecs (Crocker, 2002). No siglec has been found so far to recognise any cell-surface ligand other than Neu5Ac (Crocker and Varki, 2001) but each siglec exhibit a characteristic preference for their recognition of sub terminal Neu5Ac linkage. For instance, sialoadhesin, CD 33 and MAG bind preferentially to Neu5Ac in α -(2 \rightarrow 3)-linkages (Kelm et al., 1994a) whereas CD 22 binds preferentially to Neu5Ac in α -(2 \rightarrow 6)-linkages (Kelm et al., 1994b). These differences in sugar binding specificities are likely to be important in the cellular recognition functions of these proteins.

1.1.1 Siglec-1

Sialoadhesin (previously referred to as the sheep erythrocyte receptor) had been originally defined as a cell surface receptor (van den Berg et al., 1992) with an apparent molecular weight of 185 kDa. Sialoadhesin (Sn) is found in bone marrow macrophages and plays a role in haematopoiesis. There are several differences between Sn and the other siglecs. Firstly, it has 17 Ig like domains compared to a maximum of 7 Ig like domains found in other siglecs (Figure 1.1). Secondly, the Ig like domains of human and mouse Sn are 60-80 % identical, whereas the intracellular domains are only 30 % identical (Hartnell et al., 2001). Thirdly, it lacks immunoreceptor tyrosine based inhibitory motifs (ITIMs) in the cytoplasmic tail (Figure 1.1).

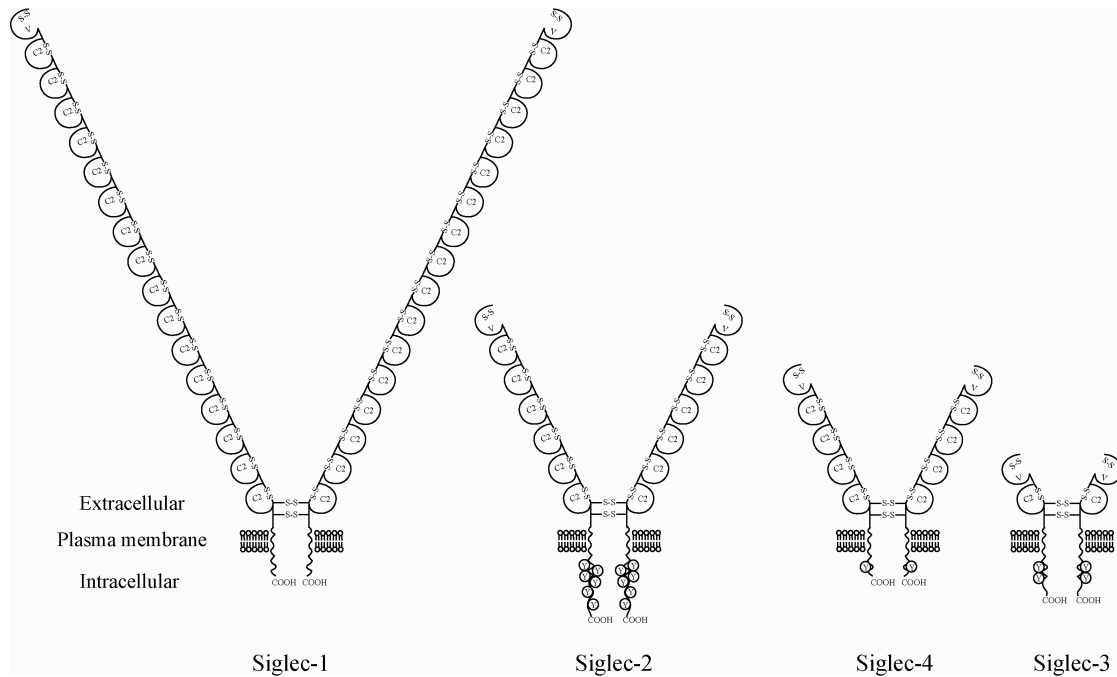


Figure 1.1: The siglec family. Members of the siglec family have a similar domain organisation and consist of an N-terminal V-set domain followed by differing numbers of C2-set domains. All siglecs except sialoadhesin contain immunoreceptor tyrosine based inhibitory motifs (ITIMs) Y in the cytoplasmic tail (adapted and modified from Crocker and Varki, 2001b).

1.1.2 Siglec-2

CD 22 or siglec-2 is a B-lymphocyte specific transmembrane protein. CD 22 functions as a modulator of intracellular signalling through the B cell receptor (sIgM) complex and as a cell surface adhesion molecule (Razi and Varki, 1998). It is specific for α -(2 \rightarrow 6)-linked Neu5Ac and carries total six ITIMs on its intracellular tail. CD 22 contains one V-set and six C2-set Ig like domains (Figure 1.1) (Crocker and Varki, 2001b). Potentially, CD 22 can bind to target cells such as lymphocytes or cytokine activated endothelial cells that express large amounts of α -(2 \rightarrow 6)-linked Neu5Ac on the surface (Nitschke et al., 2001). The CD 22 deficient B cells have a shorter life span than the CD 22 positive B cells.

1.1.3 Siglec-4

Myelin associated glycoprotein (MAG) (siglec-4a) is a glycoprotein that is expressed in oligodendrocytes (in CNS) and schwann cells (in PNS), the cells which form myelin sheaths around neural fibres. Schwann cell myelin protein (SMP) (siglec-

4b) is identified only in birds and is homologous to MAG (Dulac et al., 1992). Both MAG and SMP have one V-set Ig like domain and four C2-set Ig like domains (Crocker and Varki, 2001) (Figure 1.1). The V-set domain contains the Neu5Ac binding site (Brummendorf and Rathjen, 1994). The two distinctive properties of this domain are the presence of an intrasheet disulfide bridge, and the linkage of the first and second Ig domains via another disulfide bridge (Pedraza et al., 1990). Two major isoforms of MAG, namely L-MAG (72 kDa) and S-MAG (67 kDa) have been reported, differing only by the carboxyl terminal portion of their cytoplasmic domain (Lai et al., 1987; Tropak et al., 1988). Recent studies by Kursula and co-workers (in 2001) have focused light on the structure and function of the MAG cytoplasmic domain (MAG_{ct}). The S-MAG_{ct} has a random coil structure in solution (Kursula et al., 2001). It binds to zinc, which induces a change in the surface hydrophobicity of the protein (Kursula et al., 1999), but apparently, no folding occurs upon binding to Zn (Kursula et al., 2001). On the other hand, secondary structure predictions and circular dichroism indicates that the L-MAG_{ct} folds into a structure containing both helices and sheets (Sadoul et al., 1990). Evidence also exists on a tendency of the L-MAG_{ct} to homodimerise, which may be important in myelin-related signal transduction events (Sadoul et al., 1990).

Several studies reveal that MAG plays a role in myelination, axonal growth regulation and signal transduction. *In vitro* experiments indicated that MAG plays crucial role in the early steps of myelination (Owens and Bunge, 1989; Owens and Bunge, 1991; Trapp, 1990). MAG can influence neuronal growth in opposite ways *in vitro*. On one hand, MAG promotes neurite outgrowth in newborn dorsal root ganglion neurons (Johnson et al., 1989; Mukhopadhyay et al., 1994) while on the other hand, MAG exhibits an inhibitory effect on neurite outgrowth of neurons (Mukhopadhyay et al., 1994; McKerracher et al., 1994). Therefore, an interesting question, which remains unanswered, is whether the same ligand for MAG transmits these signals or is different signal transduction molecules are involved.

1.1.4 Siglec-3 and related siglecs

CD 33 is the smallest (67 kDa) of the siglec family to date with only two extracellular Ig like domains (Freeman et al., 1995; Simmons and Seed, 1988) (Figure 1.1). It is exclusively produced by myelomonocytic progenitors, monocytes and tissue

macrophages (Pierelli et al., 1993). Due to its production pattern, CD 33 became an important marker for the diagnosis of acute myeloid leukaemia (Del Poeta et al., 1994; Knapp et al., 1994).

The seven other siglecs have also been identified and characterised in humans. They all share a high degree of sequence similarity (~50-80 %) with CD 33, particularly in their extracellular and intracellular regions, and hence are collectively referred as 'CD 33 related siglecs' (Crocker and Varki, 2001a; Crocker and Varki, 2001b).

1.2 Gangliosides

Gangliosides are anionic glycosphingolipids that carry one or more Neu5Ac residues, in addition to other sugar residues. The presence of Neu5Ac in gangliosides indicates that these lipids are typical cell-surface constituents. Normally, free Neu5Ac in aqueous solution is present in the β -anomeric form. But it is known that the naturally occurring sialo-compounds contain Neu5Ac in the α -glycosidic linkage, except CMP-Neu5Ac. Gangliosides of human origin usually contain Neu5Ac or 9-O-acetyl-*N*-acetyl neuraminic acid (Schauer and Kamerling, 1997).

The carbohydrate portion of the ganglioside molecules is oriented toward the outer environment, and this strategic position enhances the diversity of biological events in which gangliosides are implicated. The biological significance of gangliosides is connected with molecules, most probably proteins that are able to interact more or less significantly with them. The ganglioside profile depends significantly and sensitively upon the developmental state of the cell and growth conditions (Hakomori and Zhang, 1997; Karlsson, 1995). Studies of gangliosides have revealed that they are involved in a range of different biological events such as the interactions with siglecs (Kelm et al., 1994; Mcever, 1994), toxins, hormones, differentiation factors (Walz et al., 1990) and neurotropic agents (Polley et al., 1991).

1.3 The nervous system

The nervous system is a complex tissue, which allows the coordinated function of all organs of the animal body by the transmission of excitatory and inhibitory signals. In higher animals, the nervous system is divided into two parts—the central

nervous system (CNS), which consists of the brain and spinal cord and the peripheral nervous system (PNS). Nerve fibres within both CNS and PNS may be either myelinated or unmyelinated. While unmyelinated nerve fibres are associated with somatic and visceral functions, myelinated nerve fibres assure the rapid and focused signal transmission. The rapid impulse conduction by myelinated nerve fibres is made possible by the presence of the myelin sheath (Figure 1.2), a lipid-rich multilamellar membrane enveloping the axons of nerve cells (Virchow, 1854). It has the same function in both the CNS and PNS, acts as an insulator around the axons, which guides the propagation of action potential (Vabnick et al., 1996; Kaplan et al., 1997) and provides mechanical support.

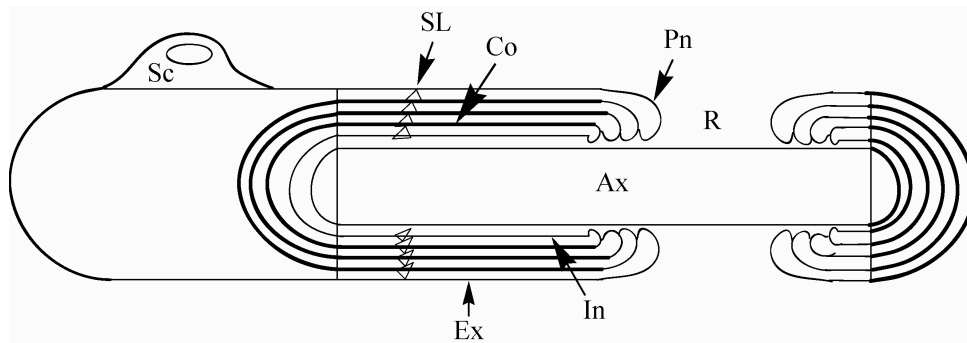


Figure 1.2: Schematic representation of a peripheral nerve myelin sheath. The axon (Ax) is surrounded by compact myelin (Co). The cytoplasm of a myelinating glial cell is confined to the narrow canals of the sheath (SL, Schmidt-Lanferman incisures), which is composed of internal (In) and external (Ex) mesaxons, the lateral edges (Pn, paranodal loops) and the Schwann cell body (Sc). The region of axons between two myelin segments is called nodes of Ranvier (R) (adapted and modified from Morell et al., 1994).

Regeneration is a process that occurs in many tissues, but in the nervous system it has the special meaning of axon growth. Axon regeneration is a motile process, and growth cones at the elongation axon tip express receptors that mediate response to environmental signals. In the PNS, regeneration occurs spontaneously after nerve injury, whereas in the CNS, damaged nerves do not regenerate. This is why brain and spinal cord injuries are so devastating. Schwab and co-workers (in 1993) first discovered the growth inhibitory activity in the CNS and also showed that much of this inhibitory activity is associated with myelin. Subsequently, two important myelin derived growth inhibitory proteins were identified, MAG and Nogo-A (McKerracher et al., 1994; Mukhopadhyay et al., 1994; Chen et al., 2000; Prinjha et al., 2000; GrandPre et al., 2000). Recently, a third myelin-associated growth inhibitor has been

identified, which is called oligodendrocyte myelin glycoprotein (OMgp) (Wang et al., 2002).

1.3.1 Mechanism of suppression of axonal regeneration

After identification of these inhibitors (MAG, Nogo-A and OMgp), the next step was to uncover the receptors that transduce the inhibitory signals across the membrane to the neuron. All the three inhibitors interact with Ngr (Nogo receptor) with about the same affinity, which is relatively high (Wang et al., 2002; Fournier et al., 2001; Liu et al., 2002; Domeniconi et al., 2002) (Figure 1.3). The contribution of each inhibitor to the receptor depends on two factors-(i) their relative abundance in myelin and (ii) the likelihood of a growth cone encountering each inhibitor as it attempts to regenerate. MAG represents 1 % and 0.1 % of the total myelin protein in the CNS and PNS respectively. It is found mostly in the inner myelin loop (Salzer et al., 1990; Trapp, 1990). Nogo-A is located in both the outer and inner myelin membrane in the CNS, although the majority is found in the endoplasmic reticulum (ER) (Huber et al., 2002; Wang et al., 2002). In one of the studies (Wang et al., 2002) it was reported that Nogo-A is upregulated in the CNS after injury but in another study it was not (Huber et al., 2002). OMgp, which is concentrated in the paranodal regions (Apostolski et al., 1994), is a relatively minor component of CNS and PNS myelin. Of these three inhibitors, only MAG seems to be myelin specific whereas Nogo-A and OMgp are expressed by neurons (Chen et al., 2000; Habib et al., 1998; Josephson et al., 2001). Therefore, it is still unclear whether the expression of Nogo-A and OMgp inhibit other axons, which attempt to regenerate.

Although Ngr is essential for the inhibitors to exert their inhibitory effects, it cannot transduce the signal across the membrane, because it is a glycosylphosphatidylinositol (GPI)-linked protein. Therefore, it has no transmembrane or cytoplasmic domains. A transmembrane 'transducing' partner p75 neurotrophin receptor (p75^{NTR}) must then interact with Ngr (Yamashita et al., 2002) to transduce the inhibitory signals across the membranes. The Ngr-p75^{NTR} complex activates 'Rho', which immediately suppresses the regeneration process (Niederost et al., 2002; Lehmann et al., 1999). By contrast, it has also been reported that the gangliosides, GT1b, which is one of the binding partners of MAG, specifically interacts with p75^{NTR}

and form a receptor complex for MAG. The GTPase, Rho associates with $p75^{\text{NTR}}$ and becomes active upon binding of MAG to receptor complex (Yamashita et al., 2002).

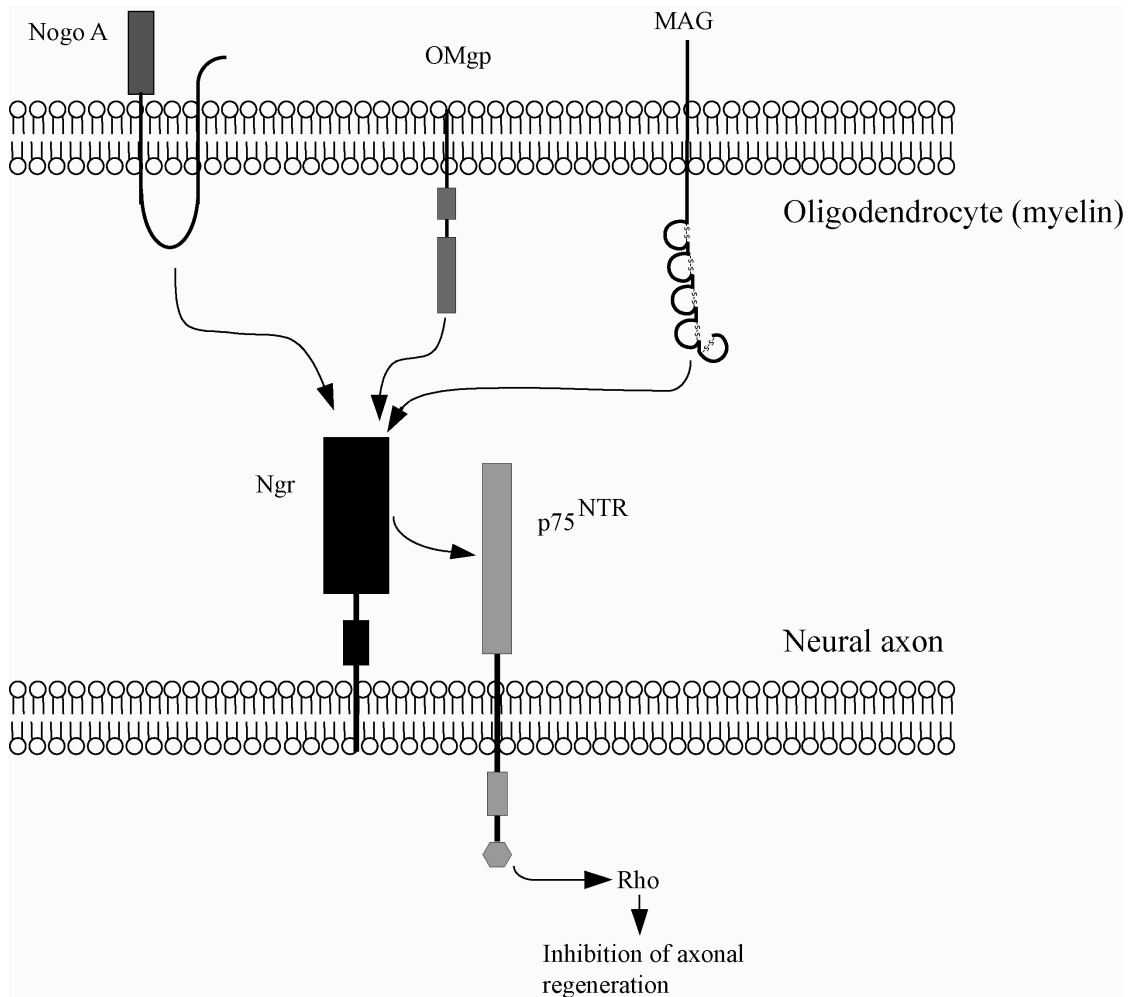


Figure 1.3: Schematic representation of the inhibition of axon regeneration. The proteins-MAG, Nogo-A and OMgp all bind to the Nogo-receptor (Ngr), which interacts with $p75^{\text{NTR}}$ to transduce the signal across the membrane. All the three inhibitors activate “Rho”, which in turn suppress the regeneration process (adapted and modified from Filbin, 2003).

1.4 Protein-ligand interaction studied by NMR spectroscopy

In general, the biological function of a protein depends on its interaction with ligand molecules. Several NMR spectroscopic techniques have emerged as powerful techniques to understand the binding process at a molecular level and utilise them for identification of new bioactive substances (Meyer and Peters, 2003; Peters et al., 2001; Dierks et al., 2001; Shapiro, 2001; Craik and Scalton, 2000; Fry and Emerson, 2000; Roberts, 2000; Hajduk et al., 1999, 1997a; Moore, 1999; Shapiro and Gounarides,

1999; Schriemer and Hindsgaul, 1998). There are two major NMR experimental approaches for the protein-ligand interactions. In the first approach, changes of chemical shifts of protein are observed in the absence as well as in the presence of ligands. The proteins labelled with ^{15}N can be used for determination of the change of chemical shift of protein in the absence and presence of ligand using ^1H - ^{15}N HSQC experiment (Shuker et al., 1996). This is an effective method to investigate the interactions of backbone NH groups of arginine or guanidino groups with charged ligands (Pascal et al., 1995; Feng et al., 1996; Gargaro et al., 1996; Morgan et al., 1999). A variation of this method is the selective labelling of proteins with ^{13}C (mainly the δ^1 protons of methyl groups of valine, leucine and isoleucine) and then observing the chemical shift change of protein using ^1H - ^{13}C HSQC experiment in the absence as well as in the presence of the ligand (Hajduk et al., 2000). Normally proteins with molecular weights ranging from 12-110 kDa are used for screening in this method. The information about the ligand binding parts of the protein can be derived from both experiments.

The second approach is based on the observation of signal changes of the ligand. Generally, in this approach, the protein molecular weight should be above 10 kDa and ligand size should not exceed 1-2 kDa. But exceptions are also reported in the literature (Kooistra et al., 2002). Analysis of the relaxation and diffusion edited NMR method is an important method for determining the bound ligand (Hajduk et al., 1997b). For weakly binding ligands i.e., ligand with dissociation constants (K_D) in the millimolar (mM) regime, the bioactive conformation of ligand can be determined by a popular method known as transferred nuclear overhauser effect (trNOE) (Balaram et al., 1972; Clore and Gronenborn, 1982; Feng, 1994). In principle, *intramolecular* and *intermolecular* trNOEs are observed using this method. The *intramolecular* trNOEs are the key to determine the bound conformation of ligands in the presence of protein (Meyer et al., 1997; Haselhorst et al., 2001) whereas the *intermolecular* trNOEs determines the conformation of ligands in connection to the receptor protein (Krishna and Moseley, 1999). An alternative to the trNOE technique is saturation transfer difference (STD) NMR method. This method is most commonly used to study which parts of a ligand bind to the protein or to screen a library of compound for binders (Mayer and Meyer, 1999). Other techniques are WaterLOGSY (Dalvit et al., 2000) and NOE pumping (Chen and Shapiro, 1998, 2000).

Scope of the thesis

The discussion of the preceding sections suggests that the axon regeneration in CNS could be feasible in two different ways: (i) by blocking the multiple axon regeneration inhibitors simultaneously; or (ii) by blocking the common signal pathway of axon regeneration of inhibitors (Dergham et al., 2002; Spencer and Filbin, 2004). The blocking of each axon regeneration inhibitors by small molecules contributes to the first approach. Thus in this context, inhibitors of MAG could be beneficial for the treatment of injuries in the CNS. The ganglioside, GQ1b α N (Figure 1.4) is the most potent MAG antagonist identified so far (Yang et al., 1996; Vyas et al., 2002) and consists of an octasaccharide with a ceramide at its reducing end (Figure 1.4). Since ganglioside carry 75-80 % of the Neu5Ac in the brain (Tettamanti et al., 1973), this could be targeted as a MAG ligand. Based on a structural activity relationship (SAR) study with numerous gangliosides, the branched tetrasaccharide α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)-[α -D-Neu5Ac-(2 \rightarrow 6)]- β -D-GalNAc (Figure 1.4) was found to make the major contribution to binding of MAG (Collins et al., 1999). To date no bound conformation of this ligand has been published.

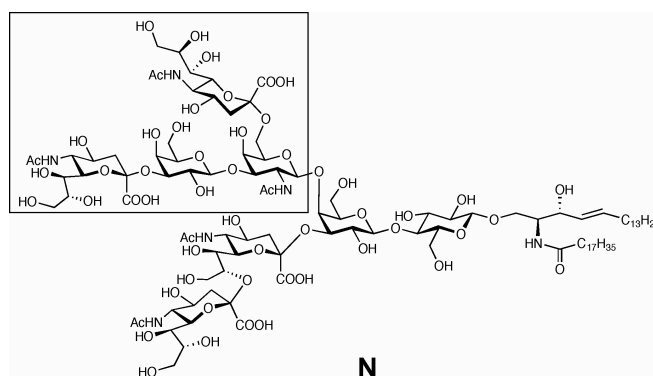


Figure 1.4: Natural ligand of MAG, GQ1b α N. The α -(2 \rightarrow 3)-linked Neu5Ac is a prerequisite for affinity to MAG (Yang et al., 1996; Vyas et al., 2002) and α -(2 \rightarrow 6)-linked Neu5Ac improves the binding affinity (Collins et al., 1999).

Aim of the present study is, therefore, the determination of the bioactive conformation of branched tetrasaccharide α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)-[α -D-Neu5Ac-(2 \rightarrow 6)]- β -D-GalNAc (Figure 1.4) in the presence of MAG using a “fragment based approach”. In this context a detailed knowledge of the three dimensional structure of MAG is necessary. In order to gain such information, a homology model

of MAG is needed to be determined on the basis of the crystal structure of sialoadhesin (which is also a member of siglec) since the crystal structure of MAG is not known yet. The main attention of this thesis is paid to the identification of the major conformational features of the tetrasaccharide ligand and its fragments using transferred NOE experiments and molecular modelling. On this basis the development of better inhibitors of MAG became possible.

2 Experimental section

2.1 NMR experiment

All NMR spectra were recorded either on Bruker DRX 500 MHz (Insitut für Chemie, Universität zu Lübeck) or on Bruker DRX 700 MHz (Institut für Organische Chemie, Universität Hamburg) spectrometer. Data acquisition and processing were performed with XWINNMR 3.1 software of Bruker. All of the experiments were carried out using standard 5 mm-triple resonance probes (TX1 HCN) equipped with gradients along the z-axis. The 2D experiments were performed with TPPI (*Time Proportional Phase Incrementation*; Marion and Wüthrich, 1983) quadrature detection. All the measurements were referenced to -OSE at 0 ppm, which is attached to terminal Gal/GalNAc residue of the investigated saccharides.

2.1.1 Sample preparation

2.1.1.1 α -(2→3)-sialyl lactose 1 and saccharides 2-6

α -(2→3)-sialyl lactose 1.

The α -(2→3)-sialyl lactose 1 (MW 633.55 g/mol) (Dorland et al., 1986) was purchased from Dextra Laboratories, Germany. The 2.0 mg of α -(2→3)-sialyl lactose 1 was lyophilized three times from 1.0 ml of 99.9 % D₂O. It was then dissolved in 0.6 ml (5.26 mM) of 25 mM deuterated Tris buffer, pH* 8.0.

Fragments of GQ1b α N, 2-6.

The saccharides 2-6, which were the fragments as well the mimics of GQ1b α N were a kind gift from Prof. Beat Ernst, Institut für Molekulare Pharmazie, Universität Basel, Switzerland. Firstly, all saccharides 2-6 were lyophilized at least three times from 1.0 ml of 99.9 % D₂O and then dissolved in 10 mM deuterated phosphate buffer, pH* 7.4 (with 150 mM NaCl) (Table 2.1) for assignments and conformational analysis.

Table 2.1: The concentration of the saccharides **2-6** used for assignments and structural elucidation. The sample was dissolved in 10 mM deuterated phosphate buffer, pH* 7.4 (with 150 mM NaCl).

Saccharides	MW (g/mol)	[Saccharides] (mM)
2	773.785	4.95
3	1110.056	1.12
4	1069.004	1.72
5	717.579	3.25
6	755.767	2.64

Protein-ligand complex.

For the preparation of NMR sample, the concentration of sialoadhesin (MW 185 kDa) and MAG (MW 150 kDa) was determined using UV absorbance at 280 nm with extinction coefficient of 1.43 and 1.44 M⁻¹ cm⁻¹ respectively. The exchangeable protons of the protein were exchanged with a buffer by repeated washing using a micro-concentrator with a 10 kDa molecular weight exclusion limit (Sartorius, Germany). Two types of buffer were used in these study, **a** containing 25 mM deuterated Tris buffer, pH* 8.0 and **b** consisting of 10 mM deuterated phosphate buffer, pH* 7.4 (with 150 mM NaCl). Molar ratios are shown in Table 2.2.

Table 2.2: The concentration of the saccharides **1-6** in the presence of protein. The sample was prepared in respected buffer.

Saccharides	Buffer	Protein	[Protein] (μM)	[Saccharide] (mM)	Molar ratio [Protein]: [Saccharide]
1	a	sialoadhesin	5	5.00	1:1000
2	b	sialoadhesin	18	2.50	1:140
2	b	MAG	30	0.54	1:18
3	b	MAG	46	0.92	1:20
4	b	MAG	40	0.80	1:20
5	b	MAG	37	0.81	1:22
6	b	MAG	60	0.84	1:14

2.1.1.2 TOCSY experiment

The TOCSY spectra were acquired using the “MLEV17” (Bax and Davids, 1985) for saccharide **1** and **3** and “dipsi” for trisaccharide **2** with a spin lock time of 80, 110 and 80 ms respectively. The spectral width was 11 ppm in both dimensions and a TOCSY spin lock field (Braunschweiler and Ernst, 1983) of 12 dB was applied.

The TOCSY spectrum was recorded with 512 increments in t1 and 4K data points in t2 with a relaxation delay of 2 s. The WATERGATE (Sklenar et al., 1993) procedure was used for water suppression.

2.1.1.3 HSQC experiment

The ^1H - ^{13}C HSQC was performed using the spectral width of 2K (t2) \times 128 (t1) for saccharides **2** and **4-6** whereas 4K (t2) \times 256 (t1) for tetrasaccharide **3** and 4K (t2) \times 128 (t1) for α -(2 \rightarrow 3) sialyl lactose **1**. All spectra were recorded with relaxation delays between 1.0-2.5 seconds and GARP pulse train was used for heteronuclear decoupling during acquisition. The delay for the evolution of heteronuclear coupling in the ^1H - ^{13}C HSQC experiment was set to 1.67 ms ($\approx \frac{1}{4J_{\text{CH}}}$, where J= 150 Hz).

2.1.1.4 HMBC experiment

The ^1H - ^{13}C HMBC was performed using the spectral width of 2K (t2) \times 128 (t1) for saccharides **5** and **6**, 4K (t2) \times 512 (t1) for tetrasaccharide **3** and 4K (t2) \times 128 (t1) for α -(2 \rightarrow 3) sialyl lactose **1**. All spectra were recorded with relaxation delays between 1.0-2.5 seconds and QF for magnitude mode. The delay between the first ^1H $\frac{\pi}{2}$ pulse and the first ^{13}C π pulse in the ^1H - ^{13}C HMBC experiment was set to 3.3 ms ($\approx \frac{1}{2J_{\text{CH}}}$, where J= 150 Hz) and the delay for evolution of long range coupling was between 70-80 ms.

2.1.1.5 NOESY experiment

The 2D NOESY (Jeener et al., 1979) spectra were recorded with 256 increments in t1 and 4K data points in t2 dimensions. The resulting NOEs were negative for saccharides **2-4** and **6** whereas the NOEs of pseudo tetrasaccharide **5** was close to zero crossing at 288 K. Phase sensitive NOESY experiments were performed using TPPI with presaturation of the HDO signal for **4**, **5** and **6** whereas WATERGATE (Sklenar et al., 1993) procedure was used for water suppression in case of **2** and **3**. For all 2D NOESY spectra, a $\frac{\pi}{2}$ shifted squared sine bell-window function (Scherf and Anglister, 1993) was applied in both dimensions prior to the Fourier transformation. After zero

filling in t_1 , $4K(t_2) \times 512(t_1)$ data matrices were obtained. For trisaccharide **2** and tetrasaccharide **3**, eight mixing times viz. 50, 75, 100, 150, 250, 350, 500 and 750 ms were chosen to generate the NOE build up curves whereas 500 ms mixing time was used for tetrasaccharide **4**, pseudo tetrasaccharide **5** and trisaccharide **6**.

In the presence of protein, a spin lock pulse with strength of 18 dB and duration of 15 ms was applied after the first $\frac{\pi}{2}$ pulse to suppress protein 1H NMR signals. After 16 dummy scans, 64 scans were recorded per t_1 increment. The residual HDO signal was presaturated with a weak rf field (74-76 dB) during relaxation and mixing time. A gradient pulse (1 ms) at the end of the mixing time was applied to remove the transverse magnetization. Six mixing times viz. 50, 75, 100, 200, 300 and 500 ms were chosen to generate trNOE build up curves for saccharides **2** and **4-6** in the presence of protein. Each experiment was performed with approximate 18 h measuring time. After zero filling in t_1 , $4K(t_2) \times 1K(t_1)$ data matrices were obtained.

2.1.1.6 ROESY experiment with $T_{1\rho}$ spinlock

ROESY experiments of the saccharides **2-6** in the presence of protein were performed with a phase-alternated 180° pulse (Bax and Davis, 1985). The mixing times in the ROESY experiments were varied from 150-350 ms. A relaxation delay of 1.5 s and ROESY spin lock field of 20 dB was applied. The residual HDO signal was presaturated with a weak rf field (74-76 dB) during relaxation and mixing time.

2.1.1.7 Saturation Transfer Difference experiment

For 1D Saturation transfer difference (STD) NMR spectra, the spectral width was 10 ppm and all spectra were acquired using digital quadrature detection. Selective irradiation of the protein was achieved by a train of Gaussian shaped pulses each with a 1 % truncation and 50 ms in duration and separated by a 1 ms delay. The minimum irradiation time required to achieve the maximum signal to noise ratio in the STD spectra, in a given total experimental time, was determined by performing a series of difference experiments with the irradiation times 0.25, 0.5, 1.0, 2.0, 2.5, 3.75 and 5 s. The protein was irradiated at 0 or -2.0 ppm (on resonance) and at 40 ppm (off resonance).

Data analysis was performed using the dual display mode of XWINNMR 3.1 program suite (Bruker). The 1D reference and STD spectra were measured under perfectly identical conditions except the number of scans where the reference spectrum was half of those of STD spectrum. The individual signal intensities (% of STD) were quantified from the overlay of 1D STD and reference spectra.

$$\% \text{ STD} = \left(\frac{I_{\text{STD}} - I_{\text{Ref}}}{I_{\text{Ref}}} \right) \times 100 \quad (2.1)$$

where, I_{STD} and I_{Ref} are the intensity of a signal in the on-resonance and off-resonance spectrum respectively and $(I_{\text{STD}} - I_{\text{Ref}})$ represent the intensity of the STD NMR spectrum.

2.1.1.8 Integration of NOESY spectra

All the NOESY spectra were processed with the program XWINNMR 3.1 software (Bruker). The integration of the NOESY spectra was also performed with XWINNMR 3.1 software (Bruker). For the calculation of absolute NOEs, the decay curves of the diagonal signals (Figure 2.1) were fitted to an exponential function with the following form:

$$f(\tau_m) = C \exp\left(-\frac{\tau_m}{T_{\text{Isel}}}\right) \quad (2.2)$$

where, C is the scaling factor, τ_m is the mixing time and T_{Isel} is the selective spin-lattice relaxation time for extrapolating to a mixing time (τ_m) of 0 ms.

The volume of the diagonal signal at zero mixing time ($\tau_m = 0$) was defined as 100 % and absolute NOEs were given in percentage.

$$\% \text{ NOE} = \left(\frac{I_{\text{crosspeak}}}{C} \right) \times 100 \quad (2.3)$$

where, $I_{\text{crosspeak}}$ is the integration of the cross peak intensity and C is the scaling factor. The build up curves was developed by fitting the experimental NOE percentages to the following double exponential function using Origin 7.0 software.

$$f(t) = p_0 * e^{-p_2 \cdot t} [1 - e^{-p_1 \cdot t}] \quad (2.4)$$

where, p_0 , p_1 and p_2 are adjustable parameters.

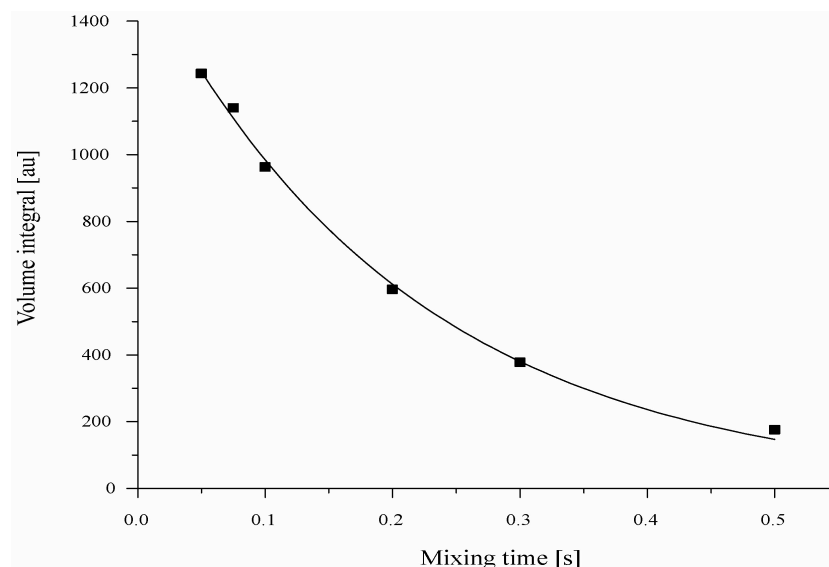


Figure 2.1: Decay curve of the diagonal signals of H3eq''' of pseudo tetrasaccharide **5** in the presence of MAG at different mixing times. ■ = Integration of the experimental volume, — = experimental fittings. (AU= arbitrary unit).

2.1.2 Complete Relaxation and Conformational Exchange Matrix Analysis

Theoretical trNOEs were obtained using a complete relaxation and conformational exchange matrix (CORCEMA) (Moseley et al., 1995, Curto et al., 1996) program. For the CORCEMA calculations a two state equilibrium involving a protein (P) and a ligand (L) forming a protein-ligand complex (PL) was assumed.

$$K_D = \frac{k_{off}}{k_{on}} = \frac{[P][L]}{[PL]} \quad (2.5)$$

The program requires (1) the co-ordinates (in pdb format) of the ligands in its free as well as in the bound forms, (2) overall rotational correlation times of the complex and the free ligand and (3) the exchange rates, i.e., off- (k_{off}) and on-rates (k_{on}). Thus the NOEs were based on the ligand geometry. As the structures of the

complexes of oligosaccharides **2-5** with MAG are not known, the protein protons were not evaluated. The following amino acids were used for CORCEMA calculation: Trp 22, Tyr 60, Ser 63-Tyr 69 and Arg 118-Ser 130 for MAG and Trp 2, Thr 37, Ala 38, Ile 39, Tyr 41, Tyr 44, Ser 45, Arg 48, Arg 97-Ser 103 and Asn 104-Val 109 for sialoadhesin.

For the calculations an overall isotropic motion of the complex was assumed, but internal motion was neglected. The group of Prof. Bernd Meyer, Universität Hamburg have reported the dissociation constants (K_D) of the saccharides **2** and **3** from STD titration curves. Thus the calculations were performed using the reported K_D values. A grid search was performed in which the off-rate k_{off} was incremented from 10 to 40 s^{-1} by 5 steps. These calculations were performed at a correlation time of 40 ns in the bound state. For the free oligosaccharides **2**, **4** and **5**, a correlation time of 0.3, 0.5 and 0.35 ns respectively yielded a good approximation of experimental NOEs. The program calculates the cross peak intensities separately for the direct NOESY cross-peaks and for the exchange-mediated cross-peaks. For the comparison with experimental data these contributions were added. The deviation between theoretical and experimental trNOEs were calculated using R-factors (equation 2.6) (Hricovini et al., 1999; Bhunia et al., 2004):

$$\text{R - factor} = \sqrt{\frac{\sum (|\text{NOE}_{\text{exp}} - \text{NOE}_{\text{cal}}|)^2}{\sum (|\text{NOE}_{\text{exp}}|)^2}} \quad (2.6)$$

NOE_{exp} and NOE_{cal} denote experimental and calculated NOEs, respectively. After matching of experimental and theoretical *intraglycosidic* trNOE curves, CORCEMA calculations for the *interglycosidic* trNOEs of the ligands were performed.

2.2 Molecular Dynamics simulation

Molecular dynamics were performed on Silicon Graphics workstations using the SYBYL 6.8 software suite (Tripos, USA). The 3D structure of the trisaccharide **2**, tetrasaccharide **3** and pseudo tetrasaccharide **7** had been computed using the molecular mechanics calculation (grid search) routine. The conformations at the each glycosidic

linkages were systematically incremented by 10° and the conformations were plotted in the relaxed type ϕ , ψ map (see “Results and discussion” for details). Dihedral angles ϕ , ψ at the glycosidic linkages were defined as follows: ϕ_1 , C1'''-C2'''-O2''-C3'' (Neu5Ac-Gal); ψ_1 , C2'''-O2'''-C3''-H3'' (Neu5Ac-Gal); ϕ_2 , H1''-C2''-O3'-C3' (Gal-Gal); ψ_2 , C1''-O1''-C3'-H3' (Gal-Gal); ϕ_3 , C1-C2-O6'-C6' (Neu5Ac-GalNAc); ψ_3 , C2-O2-C6'-C5' (Neu5Ac-GalNAc) and ω , O6'-C6'-C5'-O5'. In the next step, the lowest energy conformers, which were obtained from the above relaxed type ϕ , ψ energy map, was further optimised either in the vacuum (for trisaccharide **2**) or in the presence of water (for tetrasaccharide **3** and pseudo tetrasaccharide **7**) using the Tripos force field (Tripos, USA). The carboxyl group of Neu5Ac residue was treated as ionised. No counter ions were included in the energy minimisation; *ab-initio* charges were used at the carboxyl group (COO^-) whereas Gasteiger Hückel charges were used in the rest of the molecule. The non-bonding cut-off sphere was 8 Å and the dielectric constant (ϵ) was distance dependent (r) ($\epsilon = 4.r$). Energy minimisation of each conformation was carried out with 2000 steps each. The lowest energy conformers of **3** and **7** were placed in the centre of a rectangular box of water molecules corresponding to a density of 1.3 g cm^{-3} . The solute-to-wall distances in the X-, Y- and Z- directions were different for different combination to determine the dimensions of the rectangular box. An equilibrated cubic box of edge length 33.82 Å was used to generate the required rectangular box of water molecules and the periodic boundary conditions were not applied.

Following the energy minimisation, the resulting optimised conformations of tetrasaccharide **3** and pseudo tetrasaccharide **7** were used as a starting conformation for MD simulation. The initial configurations were then heated to 300 K from 10 K over a period of 1 ps. During the MD simulations a constant temperature of 300 K was maintained using the thermal bath coupling value of 100 fs. All hydrogen-containing bond lengths were constrained to their equilibrium values through application of the SHAKE algorithm. Newton's equations of motion were integrated using a Verlet algorithm with a 2 fs time step. During the dynamics a constant pressure of 1 atm was maintained with isotropic position scaling and a pressure relaxation. The non-bonded cut-off was set to 8 Å and the non-bonded pair list was updated every 5 steps. The system was equilibrated for 5 ps and this was followed by a 1 ns productive run. The Cartesian co-ordinates were stored after every ps.

2.3 Homology model

On the basis of the primary structure, the extracellular domain of MAG is predicted to be composed of five separate Ig domains (Lai et al., 1987). At the amino acid sequence level, the first two Ig domains of MAG share a significant degree of homology with other Ig family members such as sialoadhesin (Kelm et al., 1996). The first extracellular domain of all the family members is a V-type Ig domain, containing the Neu5Ac binding site. The structure of this domain from sialoadhesin had recently been determined by X-ray crystallography in complex with the ligand α -(2 \rightarrow 3)-sialyl lactose (May et al., 1998). This crystal structure was used as a template for homology modelling of the binding domain (N-terminal V-set domain) of MAG.

Homology modelling was conveniently implemented in the COMPOSER option (Blundell et al., 1988; Sutcliffe et al., 1987) of the SYBYL software suite (Tripos, USA). The crystal structure of sialoadhesin and the primary sequence of MAG were aligned using the program, COMPOSER with a default gap penalty of 8. The program identified the groups of topologically equivalent residues, which generated an optimal structural alignment of the sequences. A framework for the model was created based on the structurally conserved regions (SCRs). Following the SCRs and structurally variable regions (SVRs), the program examined the geometric constraints required to insert the loops between SCRs. The loops were recruited from the database of the sialoadhesin structure and were inserted into the SCRs to complete the 3D structure of MAG.

Energy minimisation routine was performed using the SYBYL software (Tripos, USA) with the Tripos force field (Tripos, USA) and Kollmann charges to relieve the strain between SCRs of the framework and the inserted variable loops.

2.4 Automated Docking

Computational docking simulations were performed with the program AutoDock 3.0 (Morris et al., 1998). Partial structure, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal of tetrasaccharide **4** and pseudo tetrasaccharide **5** were extracted from the crystal structure of α -(2 \rightarrow 3)-sialyl lactose (acquisition code, 1QFO). The protein models were used by Kollmann charges using SYBYL 6.9/Biopolymer (Tripos, USA). Grid maps or affinity maps representing the proteins were constructed using 127 \times 127 \times 127

points, with a grid spacing of 0.375 Å, centred on the ligand, which was manually positioned within the binding site. During the docking, no torsions of the ligand were allowed to vary. Protein-ligand complexes were generated from this starting point using a Lamarckian genetic algorithm (LGA) (Morris et al., 1998). Parameters for LGA docking were used as recommended in the AutoDock manual. For each protein-ligand pair, 100 LGA docking runs were performed. The maximum number of accepted and rejected trials per cycle was set to 10^6 . The binding modes were clustered using a RMS deviation cutoff of 1.0 Å with respect to the starting position. The intermolecular energy of the lowest energy cluster was used for all plots in this study. Intermolecular energy refers to the potential energy of the interaction between the protein and ligand (van-der-Waals, electrostatic, hydrogen bonding and desolvation free energy components (Morris et al., 1998)); it does not include the internal energies of the protein and ligand themselves.

3 Methods

3.1 Nuclear Magnetic Resonance spectroscopy

Nuclear magnetic resonance (NMR) is one of the most powerful techniques to the study of molecules. The great advantage of this technique is that it can work with molecules in the gas, liquid and solid states or any mixture of these states. It has a large potential to determine the structure, motions and reactions of the molecules.

3.1.1 The Nuclear Overhauser Effect

When one of the two spins (I and S, I and S refer respectively insensitive and sensitive spin) (lets say S) is disturbed by saturation or inversion with radio frequency pulses, the intensity of spin I change from the equilibrium. This change in intensity arising from the cross relaxation by dipolar interaction between the two spins is called the nuclear overhauser effect (NOE). The change of intensity of spin I is governed by the three transition probabilities: zero-, single- and double-quantum transitions, namely W_{0IS} , W_{1I} and W_{2IS} respectively. Figure 3.1 describes the cross relaxation pathways in an idealised two-spin system. The intensity change of spin I with time is defined by the following Solomon equation (equation 3.1):

$$\frac{dI_z}{dt} = -(I_z - I_z^0)(W_{0IS} + 2W_{1I} + W_{2IS}) - (S_z - S_z^0)(W_{2IS} - W_{0IS}) \quad (3.1)$$

Here, S_z and I_z are the longitudinal components of the magnitude of spin S and I respectively; S_z^0 and I_z^0 are at time zero. Solomon equation is strictly applicable for an idealized spin system i.e., for two isolated spins that are not scalar coupled and exist in a rigid and isotropically tumbling molecule.

At steady state, $\frac{dI_z}{dt} = 0$ and $S_z = 0$;

$$\therefore \frac{I_z - I_z^0}{S_z^0} = \frac{W_{2IS} - W_{0IS}}{W_{0IS} + 2W_{1I} + W_{2IS}} \quad (3.2)$$

Hence,

$W_{2IS} - W_{0IS} = \sigma_{IS}$ = Cross relaxation rate constant

$W_{0IS} + 2W_{1I} + W_{2IS} = \rho_{IS}$ = Dipolar relaxation rate constant

where, σ_{IS} defines how fast an NOE is transferred between two spins I and S whereas, ρ_{IS} is responsible for restoring the equilibrium state of spin I. At the beginning of experiment,

$$S_z^0 = \left(\frac{\gamma_S}{\gamma_I}\right) I_z^0 \quad (3.3)$$

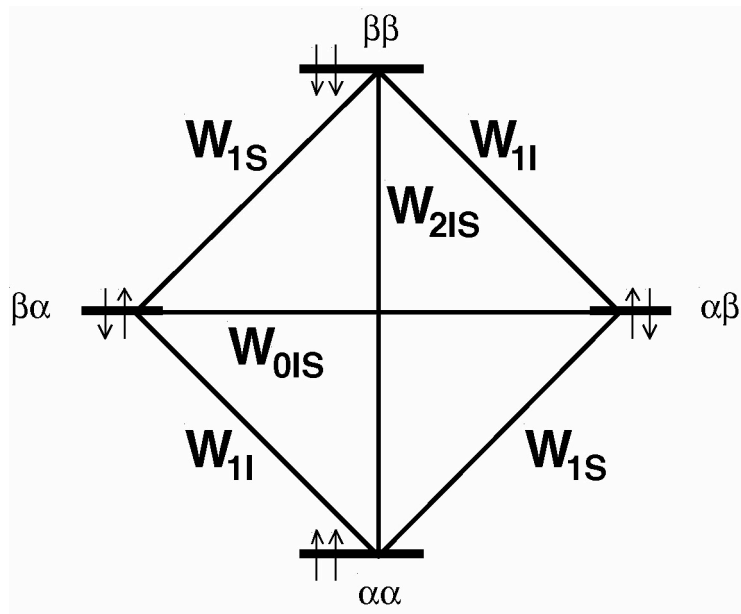


Figure 3.1: Energy level diagram for a two-spin system, showing the definitions of transition probabilities and spin states (Neuhaus and Williamson, 1989).

After saturation of spin S, the maximum steady state NOE enhancement for spin I is

$$\eta_{\max} = \frac{I_z - I_z^0}{I_z^0} = \left(\frac{\gamma_S}{\gamma_I}\right) \frac{\sigma_{IS}}{\rho_{IS}} \quad (3.4)$$

For a homonuclear ^1H - ^1H NOE, $\gamma_S = \gamma_I$. The transition probabilities and σ_{IS} and ρ_{IS} depend strongly on the precession frequencies (ω) of the spins and the overall

correlation times (τ_c) of the molecule. Therefore, the maximum homonuclear steady state NOE enhancement (equation 3.5) becomes

$$\eta_{\max} = \frac{5 + \omega^2 \tau_c^2 - 4\omega^4 \tau_c^4}{10 + 23\omega^2 \tau_c^2 + 4\omega^4 \tau_c^4} \quad (3.5)$$

Now, three cases may arise (Figure 3.2):

(1) $\omega \tau_c \ll 1$

This region of the NOE curve is often referred to as the “extreme narrowing limit”. Small molecules that tumble rapidly fall in this region. As W_{2IS} is effective in this region, maximum NOE enhancement for the steady state NOE is 50 %. However, for the transient NOE experiment the maximum value reached is 38.5 %.

(2) $\omega \tau_c \approx 1$

Molecules with molecular weight of 1000-2000 Da are located in this region. The NOE curve passes through zero, causing the maximum theoretical NOE enhancement to be zero.

$\eta_{\max} = 0$ so,

$$\begin{aligned} 5 + \omega^2 \tau_c^2 - 4\omega^4 \tau_c^4 &= 0 \\ \therefore \tau_c &= \sqrt{\frac{5}{4}} = 1.12 \end{aligned} \quad (3.6)$$

(3) $\omega \tau_c \gg 1$

Large molecules such as proteins or protein-ligand complex fall in this region. Due to the long correlation time, the W_{0IS} transition is dominant.

$$\therefore \eta_{\max} = -\frac{W_{0IS}}{W_{0IS}} = -1 \quad (3.7)$$

In reality the idealized spin systems do not exist, therefore, the longitudinal dipolar relaxation between spins I and S is not the only relaxation mechanism present in the system. In such real molecules the NOE enhancement is inversely proportional to the distance (r^6) between the nuclei.

$$\eta_{\max} \propto \frac{\tau_c}{r^6} \quad (3.8)$$

If NOE between a pair of nuclei (say, C and D) are separated by a known distance (r_{CD}), it is possible to determine the distance between two other nuclei (say, A and B) using the following simple relationship. This method is known as two-spin approximation.

$$\frac{\text{NOE}_{AB}}{\text{NOE}_{CD}} = \frac{r_{AB}^{-6}}{r_{CD}^{-6}} \quad (3.9)$$

In oligosaccharides the maximum distance between two hydrogen nuclei for which a NOE effect is detectable is approximately 4 Å whereas in either very large or very small molecules this limit may extend to 5 or 6 Å.

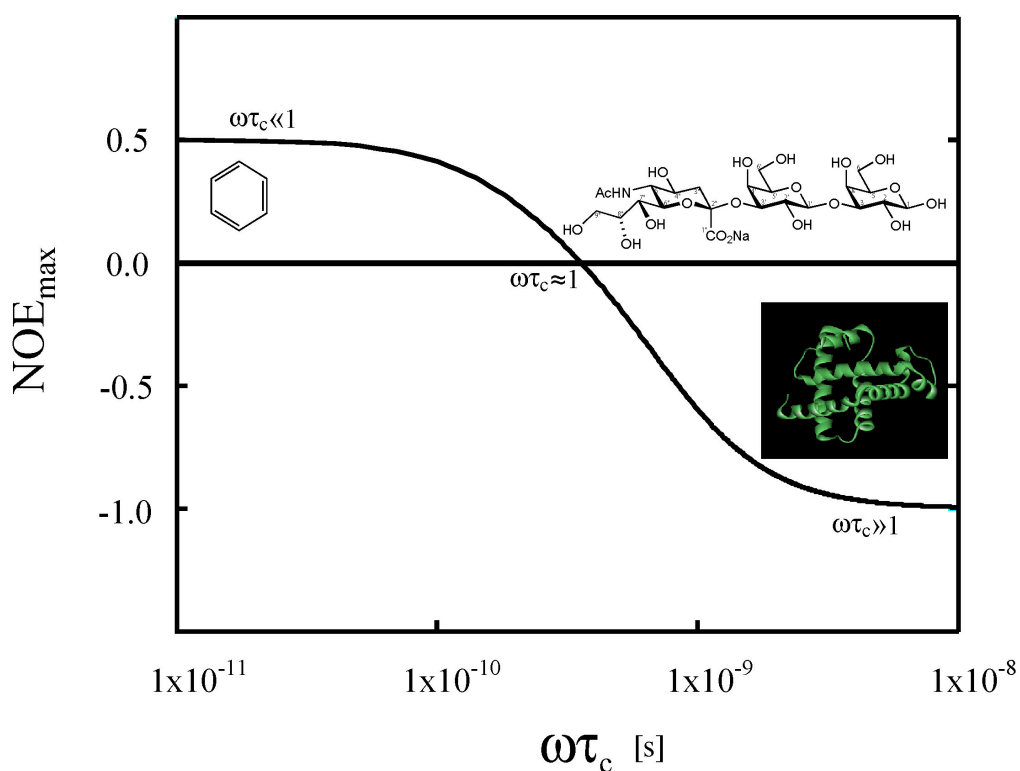
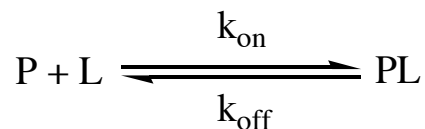


Figure 3.2: Dependence of the maximum homonuclear NOE enhancement on the product of the spectrometer frequency (ω) with the overall correlation time (τ_c) of the molecule $\omega\tau_c$ for a two-spin system (Neuhaus and Williamson, 1989).

3.1.1.1 transferred Nuclear Overhauser Effect

Let us consider a system where a receptor P binds to a ligand L to form a complex PL.



The dissociation constant (K_D) and association constant (K_A) for this reaction is:

$$K_D = \frac{k_{\text{off}}}{k_{\text{on}}} = \frac{[P][L]}{[PL]} = \frac{1}{K_A} \quad (3.10)$$

where, k_{on} and k_{off} are the forward and backward rate constants respectively.

The nuclear overhauser effect (NOE) is an extremely useful technique to determine the 3D structure of molecules in solution. When a small molecule is bound to a macromolecule, it behaves as a part of the macromolecule and adopts the corresponding NOE behaviour. This method is known as transeferred NOE (trNOE) (Balaram et al., 1972; Clore and Gronenborn, 1983; Feng, 1994). In this method, cross-relaxation (NOE) between two protons in the bound ligand is transferred to the free molecule by exchange between bound and free species (Figure 3.3) (Usually this is the case for K_D values ranging from μM to mM). Therefore, this method (trNOE) is widely used to determine the 3D structures of ligand bound to macromolecule.

Figure 3.4 shows a two-spin system undergoing two-site exchange between free (I^F and S^F) and bound states (I^B and S^B) of the ligand whereas I and S are the spins of the protons of the ligand. σ_{IS}^B and σ_{IS}^F are the cross relaxation rates for the bound and the free ligand respectively.

According to the principle of trNOE experiment,

$$k_{\text{off}} \gg \sigma^B \quad (3.11)$$

Under this condition it can be stated that

$$\left\langle \sigma_{IS} \right\rangle = N^F \sigma_{IS}^F + N^B \sigma_{IS}^B \quad (3.12)$$

where, N^F and N^B are the fractions of free and bound ligand molecules.

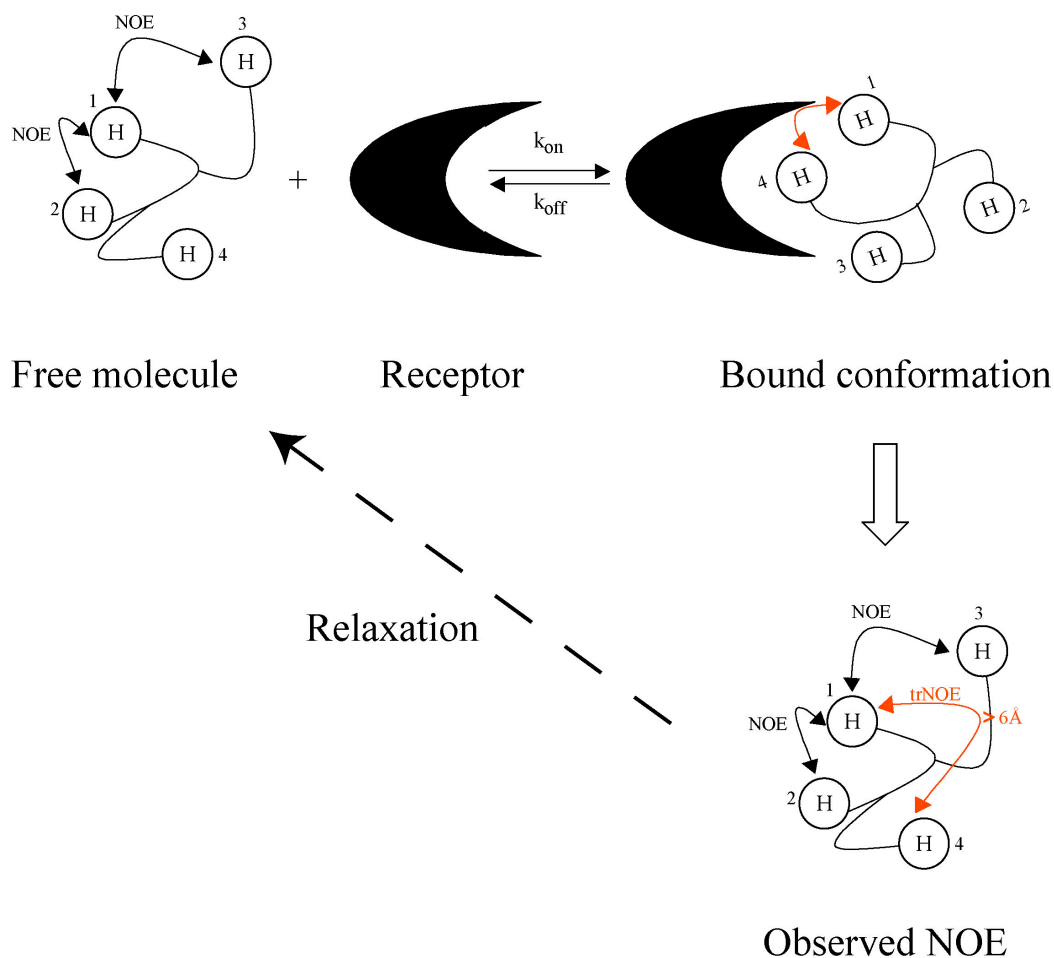


Figure 3.3: Schematic representation of trNOE, which take place during the NOESY experiments of a small molecule in the presence of macromolecule. (1) Initial state: free molecule. (2) Formation of the complex: bound conformation of the molecule. NOEs between protons close in space are developed. (3) Dissociation of the complex. The free molecule maintains the information acquired in the bound state for a given period of time, which depends on its relaxation times. After this time the system reverts to the initial state.

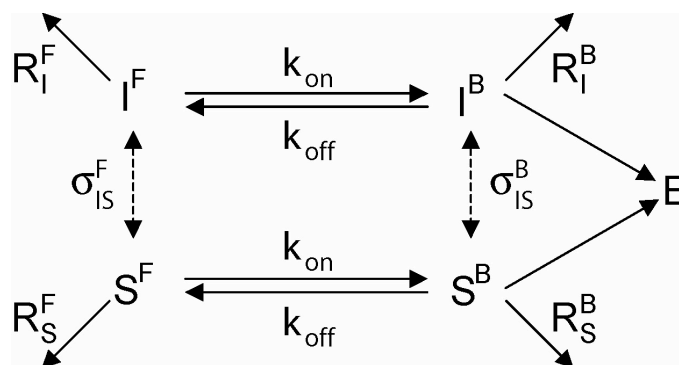


Figure 3.4: Model of two-spin exchanging system involving a ligand molecule in free and bound states. Spin diffusion via the receptor E is also indicated (Neuhaus and Williamson, 1989).

The cross relaxation rate for a large molecule (such as protein) with long correlation time (τ_c) is generally larger than a small ligand (such as oligosaccharide). But the cross relaxation rate of small molecule in the presence of large molecule is significantly larger than the cross relaxation rate of the free ligand (σ^F).

Therefore, for the measurement of trNOE effects, the correlation time (τ_c) for the bound form should dominate over the free form of the ligand.

$$\left| N^B \sigma_{IS}^B \right| \gg \left| N^F \sigma_{IS}^F \right| \quad (3.13)$$

The difference between trNOEs originating from the bound state and NOEs of the ligand in solution can be achieved by analyzing the build up rate, i.e., the time required to achieve maximum NOE intensity. For the trNOEs, this is in the range of 100-200 ms, whereas for the non-binding molecule, it is four to ten times longer. Therefore, the maximum intensity for trNOEs is observed significantly at shorter mixing times than for the isolated molecules in solution. The quantitative interpretations of *inter*- and *intramolecular* trNOEs yield reliable information about the conformation of bound ligand. The *intramolecular* trNOEs, which generally occur at distances shorter than 6 Å, are key in defining the bound ligand conformation, whereas the *intermolecular* trNOE allows to determine the orientation of bound ligand in the macromolecular binding pocket.

3.1.1.2 transferred Rotating frame Overhauser Effect

Major complication of the determination of bioactive conformation of a ligand using trNOE analysis is the spin diffusion by protein protons (Figure 3.5). It was first reported by Glaudemans and co-workers in 1990 for their analysis of β -Gal-(1 \rightarrow 3)-Gal- β -OMe-antibody complex. The problem was solved using transferred rotating frame overhauser effect (trROE) analysis (Arepalli et al., 1995). In trROE analysis, as in the regular ROESY experiment (Bothner-By et al., 1984; Bax and Davids, 1985), direct cross peaks is always positive (Figure 3.6) i.e., they show the opposite sign relative to the diagonal signals. Spin diffusion via one relay proton leads to negative cross peaks that can easily be distinguished from direct interaction.

The experiment was designed using the T-ROESY scheme (Hwang and Shaka, 1992) with a phase-alternated 180° pulse to generate the spin lock field. The T-ROESY sequence efficiently suppresses the Hartman-Hahn effects. Therefore, if a cross peak is observed in the trNOE analysis spectrum and no cross peak appear in the trROE analysis spectrum then it can be postulated that the cross peak in the trNOE analysis spectrum is originated from spin diffusion.

One drawback of trROE analysis experiment is that the bound and free ligand both give positive signals (Figure 3.6), therefore, the spin diffusion effect via protein protons can be distinguished only in combination with trNOE analysis experiment.

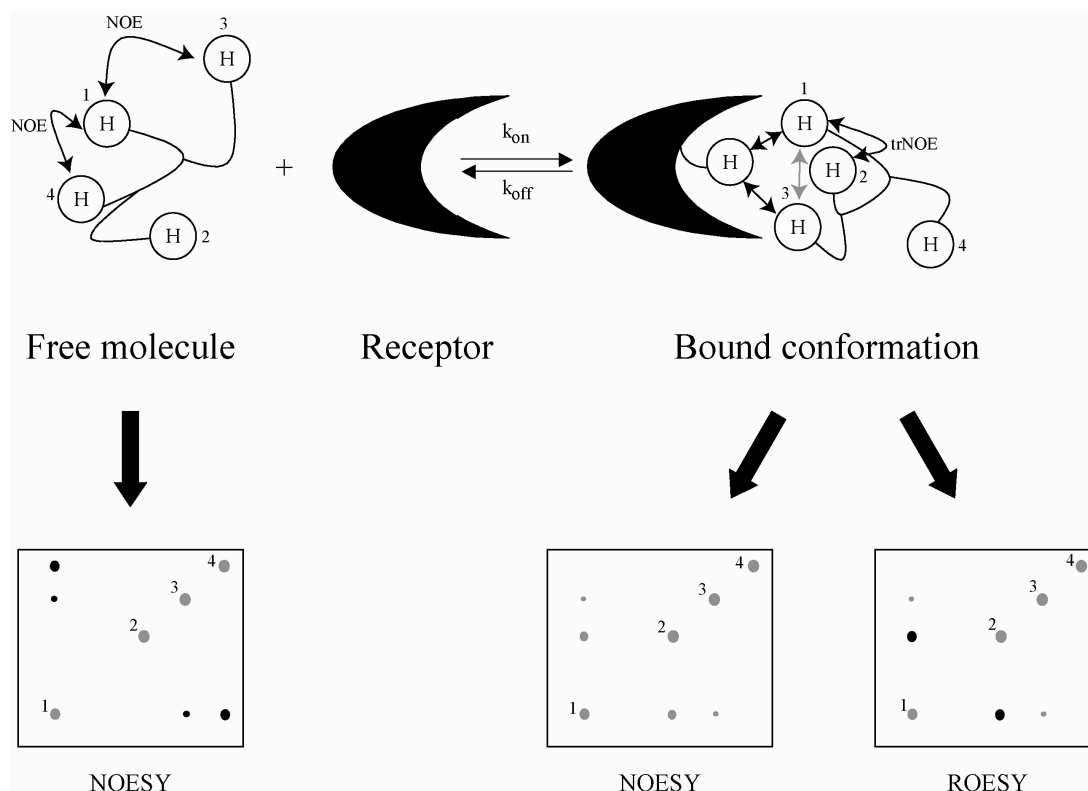


Figure 3.5: (Left) Schematic representation of a NOESY spectrum of a free small molecule. Cross peaks and diagonal peaks have different signs. (Right) Schematic representation of NOESY and ROESY spectrum of a small molecule in the presence of protein. In the NOESY spectrum of the complex, the direct and spin diffusion-mediated cross peaks have same sign as diagonal peaks. The relative sizes of the peaks and the appearance of new ones may be used to detect conformational variations. In the ROESY spectrum of the complex, the spin diffusion (three spin effects) cross peaks (i.e., 1/3) and diagonal peaks has same signs. On the other hand, direct cross peaks (1/2) show different sign to diagonal peaks. The trROE help to differentiate the direct NOEs and three spin effects.

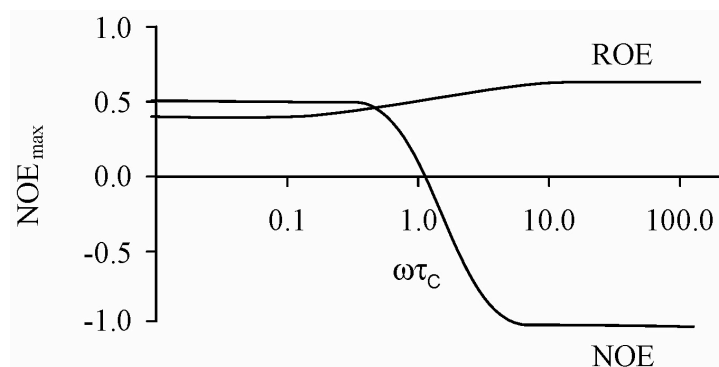


Figure 3.6: Schematic illustration of the dependence of the ROE and NOE for an isolated homonuclear two-spin system as a function of molecular tumbling rates ($\omega\tau_c$).

3.1.1.3 Saturation Transfer Difference NMR experiment

The large receptor molecules such as protein has characteristic spin diffusion phenomenon due to its fast T_2 relaxation. A further characteristic of these molecules is the line broadening of the signals in the ^1H NMR spectra whereas the signals of the small molecules of molecular weight of 1-2 kDa have very sharp signals. This phenomenon can be termed, however, into an advantage for protein-ligand interaction, if one uses it for the saturation transfer.

The saturation transfer difference (STD) NMR is an alternative method of the trNOE technique. The principles of the STD NMR was discovered over three decades ago, but only recently it has been used to study the binding activity of ligands with protein and to screen libraries of compounds for binders (Mayer and Meyer, 1999). In the STD NMR technique (Mayer and Meyer, 1999), irradiation of the protein is performed at a frequency where no ligand signals are present. This leads to a selective and very efficient saturation of the entire protein by spin diffusion. If a ligand binds to saturated protein, then it will also be saturated. The degree of saturation depends on the residence time of the ligand in the protein-binding pocket. The acquired spectrum is referred to as the “on resonance” spectrum and the contained signal intensities are referred as I_{sat} . In a second experiment, the irradiation frequency is set to a value that is far from any signal e.g., 40 ppm. Thus neither the protein nor the ligand is saturated by this selective irradiation. This spectrum is referred to as the “off resonance” spectrum (I_0). Subtraction of these two spectra leads to a difference spectra of signal intensities,

$I_{\text{sat}} - I_0$, called STD NMR spectrum. All signals of the ligand proton that are observed in the STD NMR spectrum are due to the transfer of saturation from protein to ligand.

One of the applications of this technique is to identify which ligand binds to a protein/receptor from a compound library of the ligands (Mayer and Meyer, 1999; Schuster, 2000; Vogtherr and Peters, 2000, Biet, 2003). This is visualized in Figure 3.7. The compounds, which bind to the protein, show signals, whereas the non-binding compound does not appear in the STD NMR spectrum.

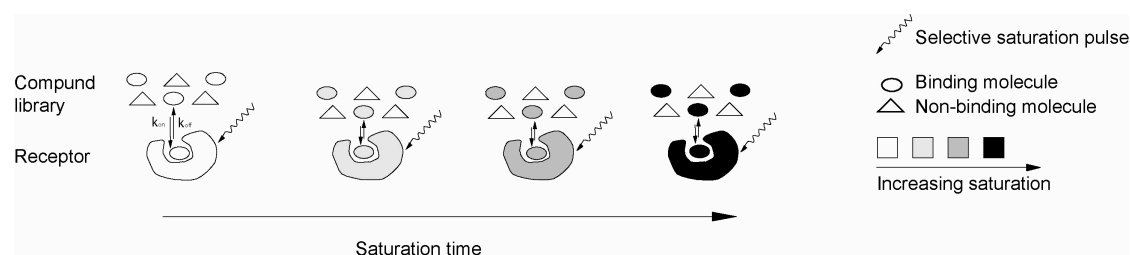


Figure 3.7: Schematic representation of the screening of the ligands using STD experiments (Mayer and Meyer, 1999). In general, the saturation pulse consists of a cascade of gaussian shaped pulses. The duration of saturation times ranges from 1 to 2 s. The ligand is used in a ca. 100 fold molar excess over the protein, allowing to work with μM protein concentrations.

Another application of this technique is the epitope mapping (Mayer and Meyer, 2001) (Figure 3.8) of a ligand binding to a receptor molecule. This method distinguishes which the part of a ligand is in close contact with the protein/receptor. As STD is a distance dependent phenomenon, the intensity of the ligand resonances in the STD spectrum decreases with increasing distance (r) between protein and ligand ($\frac{1}{r^6}$ relationship).

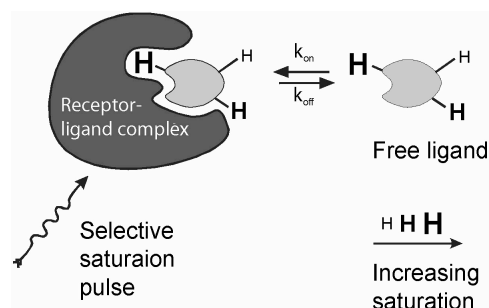


Figure 3.8: Schematic representation of group epitope mapping (GEM) using STD experiment (Mayer and Meyer, 2001). Here, groups represented by the large proton are in close contact with the protein whereas the medium sized proton symbolizes groups with less interaction. The smallest proton represents a group with almost no contact with the protein.

3.2 Complete Relaxation and Conformational Exchange Matrix Analysis

The analysis of trNOE in multispin systems by the use of isolated spin pair approximation (ISPA) is often inadequate since it neglects multispin effects i.e., spin diffusion effects for large molecules with longer correlation time (Krishna et al., 1978, Ni, 1992). The CORCEMA algorithm (Moseley et al., 1995, 1997; Curto et al., 1996) has been developed to analyse the 2D NOESY spectra of interacting systems undergoing multistate conformational exchange. It explicitly incorporates *intermolecular* dipolar cross relaxation between the molecules when they are in complex form. It also permits an analysis of NOESY intensities for the *intra*- as well as *intermolecular* contacts between ligand and receptor under a variety of binding conditions, e.g., dissociation constant (K_D), off rate (k_{off}), on rate (k_{on}) and correlation time (τ_c). Figure 3.9 shows the schematic representation of the program.

The model is characterised by free molecules L (ligand) and E (enzyme/receptor) and the bound molecule E'L' (Figure 3.10). The ligand adopts the correlation time of the receptor when it is in the bound state. Similarly, in the bound state the conformation of the ligand may also change from its free state. CORCEMA produces the relaxation matrix (**R**) and kinetic matrix (**K**) from the given pdb file of receptor-ligand complex.

The relaxation matrix **R** describes the relaxation behavior of the entire system, i.e., it consists of ligand L, receptor E and the complex E'L' in the equilibrium. For the two-state system the relaxation matrix **R** is defined as (Moseley et al., 1995; Curto et al., 1996):

$$\mathbf{R} = \begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix} \quad (3.14)$$

where, R_1 and R_2 represent the relaxation matrix for the free and bound state conditions.

The kinetic matrix \mathbf{K} is represented as:

$$\mathbf{K} = \begin{bmatrix} K_{12} & -K_{21} \\ -K_{12} & K_{21} \end{bmatrix} \quad (3.15)$$

$$K_{12} = \begin{bmatrix} k_{12}^L & 0 \\ 0 & k_{12}^E \end{bmatrix} \quad \text{and} \quad K_{21} = \begin{bmatrix} k_{21}^L & 0 \\ 0 & k_{21}^E \end{bmatrix}$$

where, k_{12}^L and k_{12}^E are the forward rate constant for the ligand and receptor molecules respectively. k_{21} is the backward rate constant.

The dynamic matrix \mathbf{D} ($= \mathbf{R} + \mathbf{K}$) is symmetrical and diagonalised by transformation. The dynamic matrix \mathbf{D} describes the development of the cross signal intensities in a 2D NOESY experiment (Ernst et al., 1987).

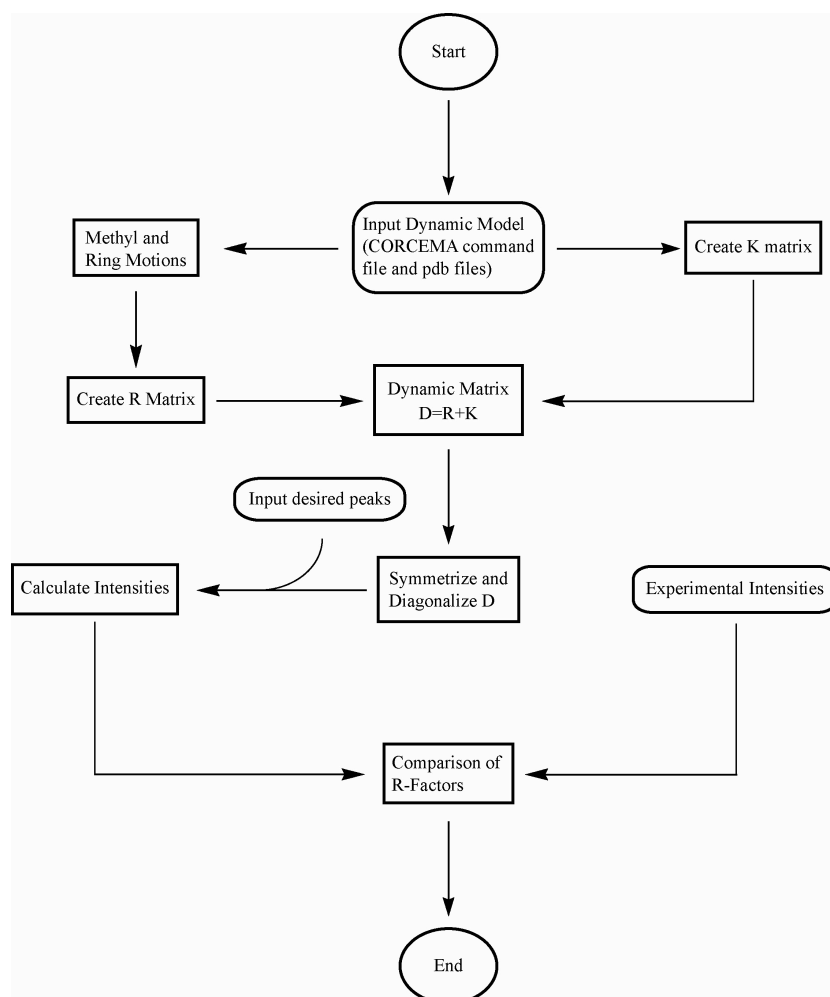


Figure 3.9: Flow chart for CORCEMA calculations. CORCEMA execution involves creation of generalized rate matrices for relaxation (\mathbf{R}) and kinetics (\mathbf{K}) from the input parameters. Next, the dynamic matrix ($\mathbf{D}=\mathbf{R}+\mathbf{K}$) is created. The experimental and theoretical NOE values are compared based on R factors.

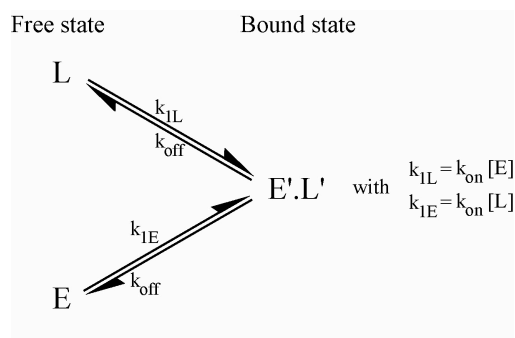


Figure 3.10: A “two-state” system involving a ligand (L) and an enzyme (E) forming a ligand-enzyme complex.

Four different species are included in the CORCEMA calculation, i.e., free and bound ligand and free and bound receptor. The NOE R-factor (Xu et al., 1995; Hricovini et al., 1999) describes the quality of the theoretical model in comparison to experimental data. Therefore, lower the R-factor better the model explains the observations.

$$R - \text{factor} = \sqrt{\frac{\sum (| \text{NOE}_{\text{exp}} - \text{NOE}_{\text{cal}} |)^2}{\sum (| \text{NOE}_{\text{exp}} |)^2}} \quad (3.16)$$

NOE_{exp} and NOE_{cal} denote the experimental and calculated NOEs respectively.

3.3 Molecular Dynamics simulation

Historically, modeling has been associated with the search of the “global minimum energy” of a molecule. Several factors are responsible for minimizing the energy. This can be described by a mathematical relationship (equation 3.17) (Leach, 1996), which is referred as force field (Figure 3.11).

$$E_{\text{pot}} = \sum_{\text{bonds}} \frac{k_i}{2} (l_i - l_{i,0})^2 + \sum_{\text{angles}} \frac{k_i}{2} (\theta_i - \theta_{i,0})^2 + \sum_{\text{torsions}} \frac{V_n}{2} (1 + \cos(n\omega - \gamma)) + \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}} \right) \quad (3.17)$$

There are currently two general classes of carbohydrate force fields. The first and earliest class of force field is known as hard sphere *exo*-anomeric (HSEA) (Thøgersen et al., 1982) force fields, which predicts the conformation of oligosaccharides based solely on the energy due to van-der-Waals interactions and the *exo*-anomeric effect. The classical force fields fall into the second class e.g., AMBER (Weiner et al., 1984), CHARMM (Brooks et al., 1983; Reiling et al., 1996), TRIPOS (Clark et al., 1989) and MM3 (Allinger et al., 1989).

In a system, all molecules are always in motion except at absolute zero. The overall motion of the molecules in the system is composed of the translational, rotational and vibrational motions, which are in turn related to the temperature and pressure of the system. The molecular dynamics (MD) generates the configurations of the system with time using Newton's second law of motion. For a mono-dimensional system, the acceleration of each atom is calculated as,

$$a_i = -\frac{\left(\frac{\delta V}{\delta x_i}\right)}{m_i} = \frac{F_i}{m_i} \quad (3.18)$$

V = potential energy function

F_i = force field

m_i = mass of particle

x_i = one coordinate

a_i = acceleration.

Using the initial atom position, velocity and acceleration, a new position for each atom is determined after a certain time step (Δt), typically 1 fs. In this way, MD simulation generates a trajectory that describes the dynamic variables of the system with time. The total time of simulation are typically in the order of hundreds of picoseconds (ps) to several nanoseconds (ns). From the MD trajectories one can obtain information regarding the thermodynamic properties of the molecule.

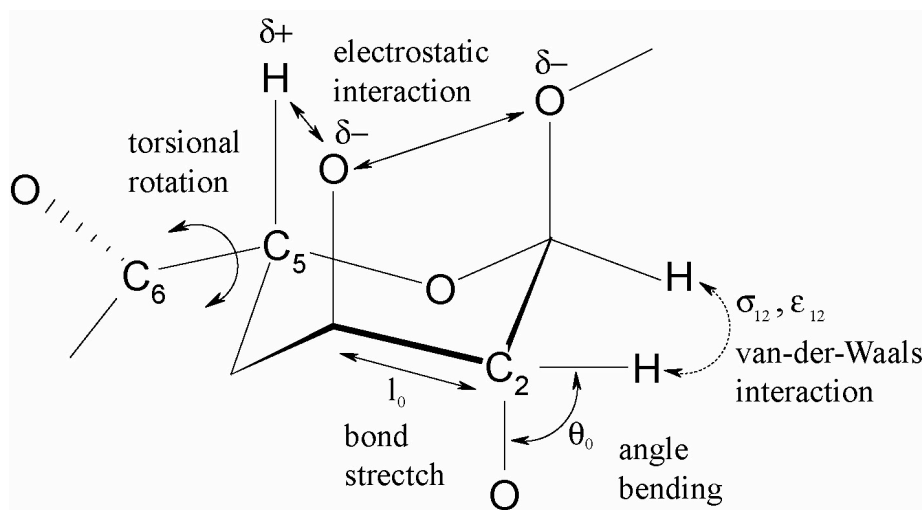


Figure 3.11: Schematic representation of the four key contributions to a molecular mechanics force field.

3.4 Automated Docking

For the last three decades the molecular docking and computer-aided-drug designing has become very important field. There are many different docking programs e.g., DOCK, FlexiDock, AutoDock and LUDI. All have a common feature—optimisation of the binding conformations between two molecules. All of these methods utilise techniques like simulated annealing, genetic algorithms, flexible molecular bonds and grid boxes.

AutoDock uses a search method called Lamarckian genetic algorithm (LGA) (Morris et al., 1998), which is a hybrid of local search (LS) and genetic algorithm (GA). This algorithm is named after Jean Baptiste de Lamarck who believed that phenotypic characteristics one obtained during a lifetime could become heritable and incorporated in ones genotype. Most of the Lamarckian algorithms employ an inverse mapping function to map a minimized phenotype back into genotype. The Lamarckian aspect of this algorithm pertains to the LS performed on the offspring to minimize the phenotype (structure), and then in essence turn the phenotype into the genotype of the new offspring. In case of AutoDock, the ligand's state corresponds to the genotype whereas its atomic coordinates correspond to the phenotype.

In this case, the macromolecule is rigid and fixed while the ligand is flexible and can have both translational and rotational freedom. A rapid grid-based method is used for finding the lowest binding energy of the bound conformation. The grids are

calculated for each atom type and the size of the box can be set manually and placed at a specific location. These boxes create maps over the molecules that are used for the docking.

The AutoTors, modules of AutoDock defines which bonds in the ligands are rotatable, affecting the degree of freedom (DOF) of the ligand. Each rotatable torsional angle adds an extra DOF.

The AutoGrid, modules of AutoDock pre-calculates a three dimensional grid of interaction energies based on the target of macromolecule using AMBER force field. Since the structure of the receptor protein is assumed to be rigid and known, interaction energies between the ligand and surrounding amino acids can be calculated at each point in the grid and stored in a table. Additional tables are made for each atom type in the ligand, taking into account of dispersion/repulsion and hydrogen bond energies.

After the generation of the grid with AutoGrid, the AutoDock simulation itself can be performed with LGA docking runs, which consists of a series of generations. AutoDock counts the number of energy evaluations and the number of generations as the docking runs proceeds. The run terminates when the limit is reached. *Intermolecular energy* refers to the potential energy of the interaction between the protein and ligand, which is represented by clusters with respect to the RMS deviation cut-off.

4 Results and Discussion

4.1 Sialoadhesin

The siglec (sialic acid immunoglobulin like lectins) family comprises of a group of cell surface molecules (Crocker et al., 1998; Kelm et al., 1994a; Kelm et al., 1994b), which is classified by sequence homology with the members of immunoglobulin super family (IgSF). To date, there are 11 different siglecs known that can be divided into two subgroups. One subgroup represented by sialoadhesin (siglec-1), CD 22 (siglec-2), MAG (siglec-4a) and SMP (siglec-4b) whereas the other subgroup consists of CD 33 (siglec-3) related siglecs (Crocker, 2002). Sialoadhesin is the largest siglec of seventeen Ig like domains that is predominantly expressed in specialised subsets of macrophages within haemopoietic and secondary lymphoid tissues (Crocker et al., 1991a; Crocker et al., 1991b) under non-inflammatory conditions. In chronic inflammatory conditions such as atherosclerosis and rheumatoid arthritis, evidence indicates that sialoadhesin is expressed at high levels on inflammatory macrophages (Crocker et al., 1997).

A crystal structure (May et al., 1998) of sialoadhesin (co-crystallised with α -(2 \rightarrow 3)-sialyl lactose diffracted to 1.85 Å resolution) was obtained from Brookhaven Protein Data Bank (Acquisition code, 1QFO). In Chapter 4.1.1 the *in vivo* interaction of α -(2 \rightarrow 3)-sialyl lactose **1** with sialoadhesin will be described by NMR experiments under physiological conditions. The results will be compared with the crystal data (May et al., 1998). The NMR derived 3D structure of the trisaccharide **2** in the bound state is discussed in Chapter 4.1.2 and compared with α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)-D-Glc **1** structure in the crystal.

4.1.1 Group epitope mapping of α -(2 \rightarrow 3)-sialyl lactose **1** in presence of sialoadhesin

A complete assignment of ^1H NMR resonance signals of α -(2 \rightarrow 3)-sialyl lactose **1** (Figure 4.1) (Dorland et al., 1986) was achieved on the basis of TOCSY, NOESY, HSQC and HMBC experiments. All experiments were performed on a Bruker DRX

500 MHz spectrometer at 303 K. As the sugar has a non-reducing terminal glucose, it is present as both α and β anomers. Figure 4.2 shows a ^1H NMR spectrum of the trisaccharide **1**. The chemical shift values are summarised in A.1.1.

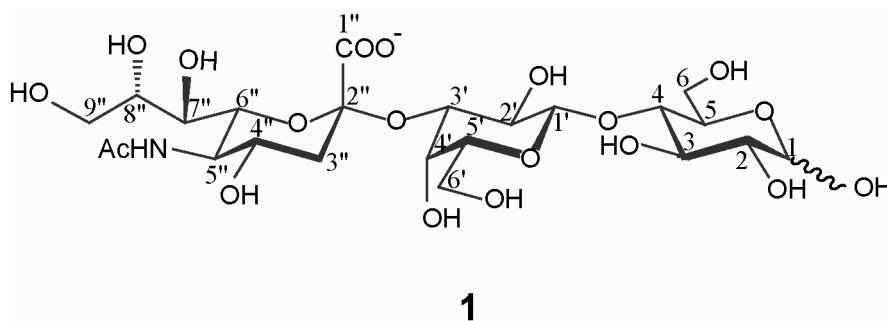


Figure 4.1: Structure of α -(2 \rightarrow 3)-sialyl lactose **1**, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)-D-Glc.

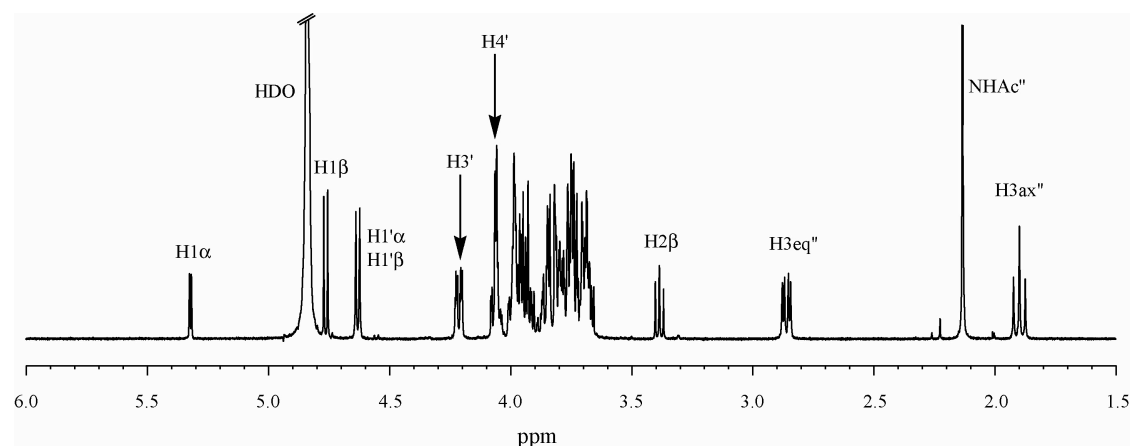


Figure 4.2: ^1H NMR spectra of α -(2 \rightarrow 3)-sialyl lactose **1** at 303 K and 500.13 MHz in D_2O tris-buffer, pH* 8.0. The spectrum was referenced to NHAc'' at 2.13 ppm.

4.1.1.1 STD NMR experiments

To determine the binding epitope of α -(2 \rightarrow 3)-sialyl lactose **1** STD NMR experiments was used and saturation times were 0.25, 0.5, 1, 2, 2.5, 3.75 and 5 s. From the titration experiments it was found that the intensity of *N*-acetyl methyl group was larger than any other signals. Figure 4.3 shows the intensity of *N*-acetyl methyl group of α -(2 \rightarrow 3)-sialyl lactose **1** at different saturation times. The STD amplification factor is very low ca. 0.1 % at 0.25 s whereas the factor is ca. 0.9 % at 2 s and 1.1 % at 5 s.

So, at 2 mM concentration, 2 s saturation time was sufficient. Therefore, in all of the subsequent experiments the saturation time was set to 2 s.

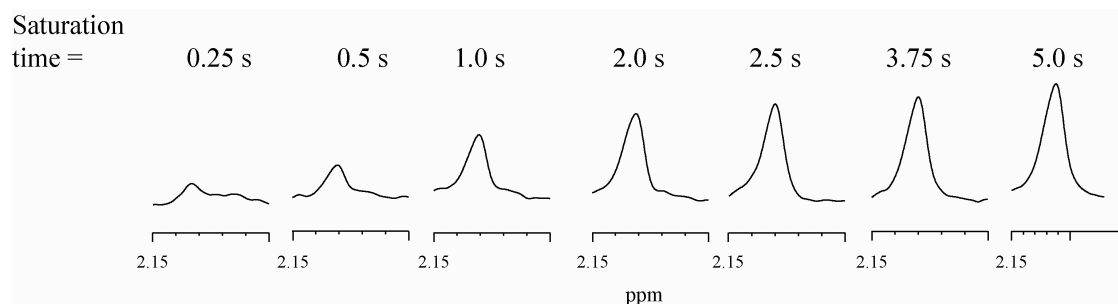


Figure 4.3: Selected portion of ^1H STD spectrum of the α -(2 \rightarrow 3)-sialyl lactose **1** (2 mM) in the presence of sialoadhesin (5 μM binding sites) at 500 MHz and 303 K in D_2O tris-buffer, pH* 8.0 showing the *N*-acetyl methyl group of **1**. For the STD NMR spectrum the protein was saturated with 0.25, 0.5, 1, 2, 2.5, 3.75 and 5 s (on-resonance: 0 ppm and off-resonance: 40 ppm).

The ^1H NMR spectrum of the complex (1:1000 excess of ligand) using the DPGSE pulse sequence (Stott et al., 1995) for water suppression is shown in Figure 4.4 A. Since α -(2 \rightarrow 3)-sialyl lactose **1** is a weakly binding ligand with a dissociation constant (K_D) of 0.8 mM (Crocker et al., 1999), the off-rate constant, k_{off} , should be large, assuming a diffusion controlled on-rate (k_{on}) of the ligand, ideal conditions for STD NMR experiments are fulfilled. The protein was irradiated at 0 ppm (on resonance) and 40 ppm (off resonance). The ^1H difference (Figure 4.4 B) and STD TOCSY (Figure 4.4 C) spectrum contained only signals of α -(2 \rightarrow 3)-sialyl lactose **1** that was saturated exclusively from sialoadhesin.

The largest STD effect was observed for the *N*-acetyl methyl group of Neu5Ac (Figure 4.4 B). Therefore, this signal was used as a reference and was set to 100 %. The relative degree of saturation for the individual protons is displayed in Figure 4.5. The contribution of the H8'' was around 16 % whereas the remaining protons of Neu5Ac were attributed to a mean value of 30 % per proton. It shows that the Neu5Ac residue has an intimate contact to the sialoadhesin-binding site. The H4', H5' and H6' protons of Gal show a STD effect of 15 % each. Interestingly, the H3' proton of Gal was saturated over 24 %, again demonstrating that individual protons within a single residue can show variations (Figure 4.6). No STD response was observed from the anomeric protons of Gal and Glc. Only H2 β and H6 β of Glc showed STD effects of 17 % and 30 % respectively. Thus saturation transfer to the protons of Gal was significantly less effective and for the terminal Glc residue almost no STD response

was detected. Thus a clear distinction between protons with a close contact to the protein and others could be made.

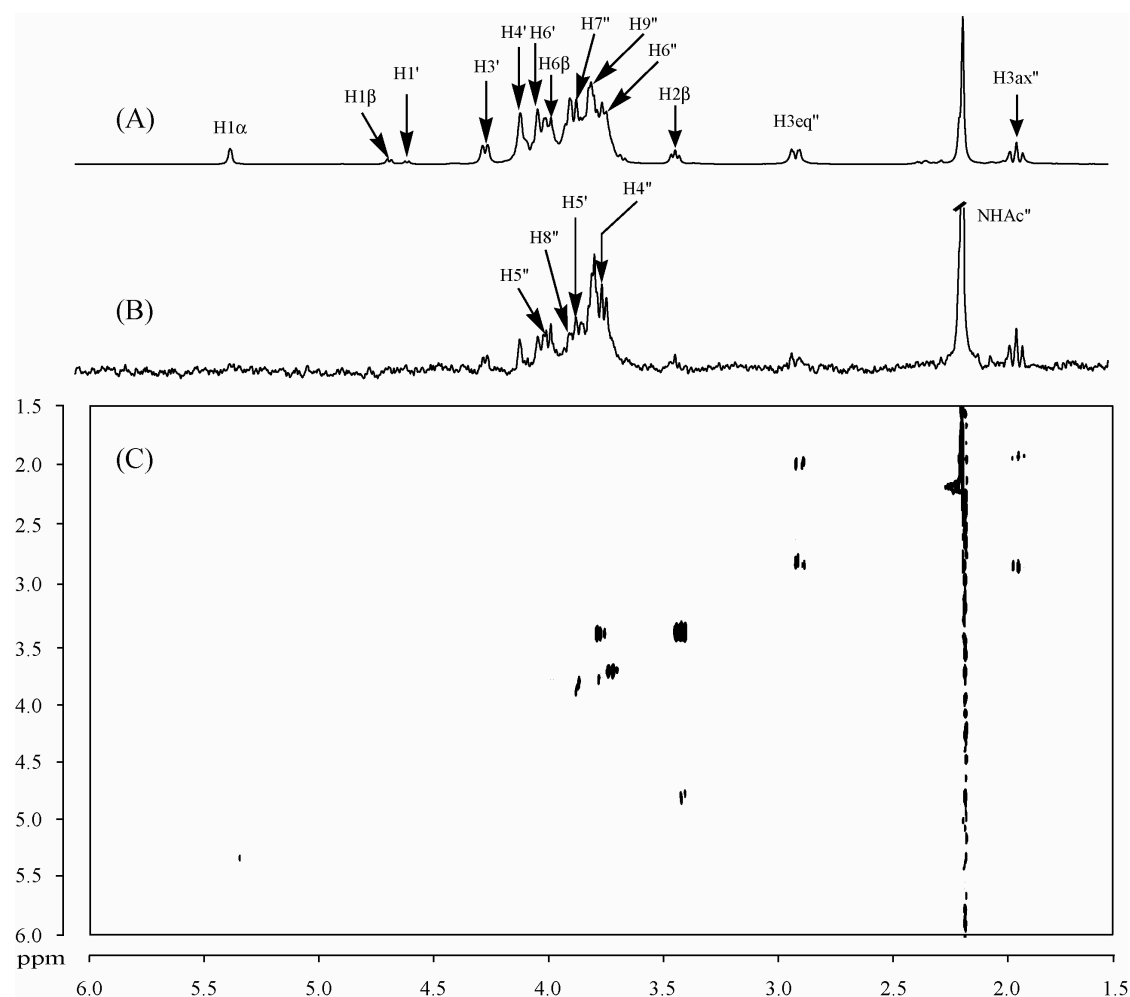


Figure 4.4: (A) Reference ^1H NMR spectrum of α -(2 \rightarrow 3)-sialyl lactose **1** (5 mM) in the presence of sialoadhesin (5 μM) using the DPGSE for water suppression. Corresponding (B) 1D STD NMR and (C) STD TOCSY spectrum of α -(2 \rightarrow 3)-sialyl lactose **1** in the presence of sialoadhesin at 500 MHz and 303 K in D_2O tris-buffer, pH* 8.0. For the TOCSY spin lock field, a DIPSI sequence was used. The mixing time was 60 ms. For the saturation of the protein (on resonance: 0 ppm, off resonance: 40 ppm) a cascade of 40 selective Gaussian pulses (50 ms each) were applied resulting in a total saturation time of 2 s.

The principle of STD NMR is easily continued with any multidimensional NMR experiment. In comparison to 1D STD NMR spectra, STD TOCSY has the advantage that even in regions of severe signal overlap a discrimination of individual resonances is feasible. Figure 4.4 C shows the region of the STD TOCSY spectrum. Therefore, it was easily possible to distinguish between protons with a close contact to the protein and others (Figure 4.6). Highlighted cross peaks were corresponding to NHAc'' ,

H3ax'', H3eq'', H4'', H6'' and H9'' of Neu5Ac, H5' of Gal and H2 β of Glc protons. In agreement with the 1D STD experiment, the strongest signals from the Neu5Ac moiety of α -(2 \rightarrow 3)-sialyl lactose **1** were clearly visible.

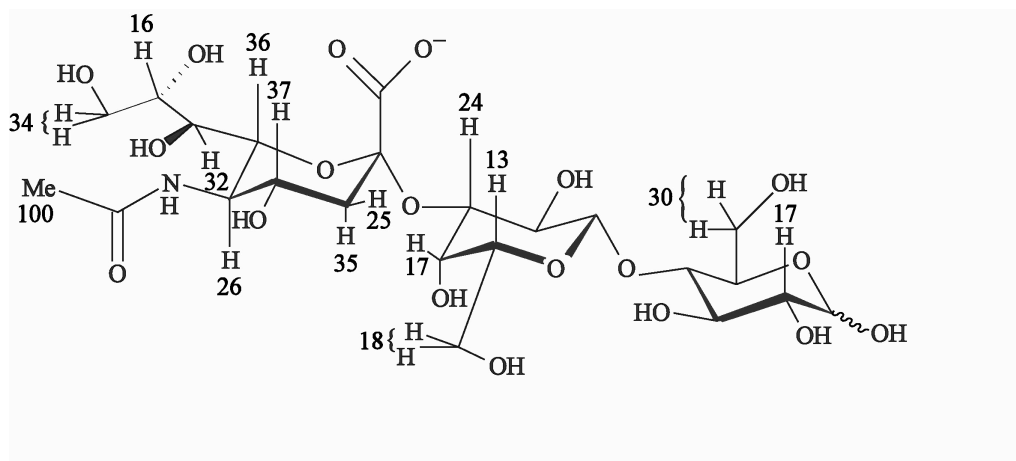


Figure 4.5: The relative STD effects of α -(2 \rightarrow 3)-sialyl lactose **1** bound to sialoadhesin. The values were calculated by determining the individual signal intensities in the STD spectrum (I_{STD}), and in the reference ^1H NMR spectrum (I_0). The ratios of the intensities ($I_{STD}-I_0/I_0$) were normalised using the largest STD effect (*N*-acetyl methyl group, 100 %) as a reference. The number indicates the percentages of STD effects experienced by protons at a 1000 fold excess. The concentration of sialoadhesin was 5 μM and that of α -(2 \rightarrow 3)-sialyl lactose **1** was 5 mM.

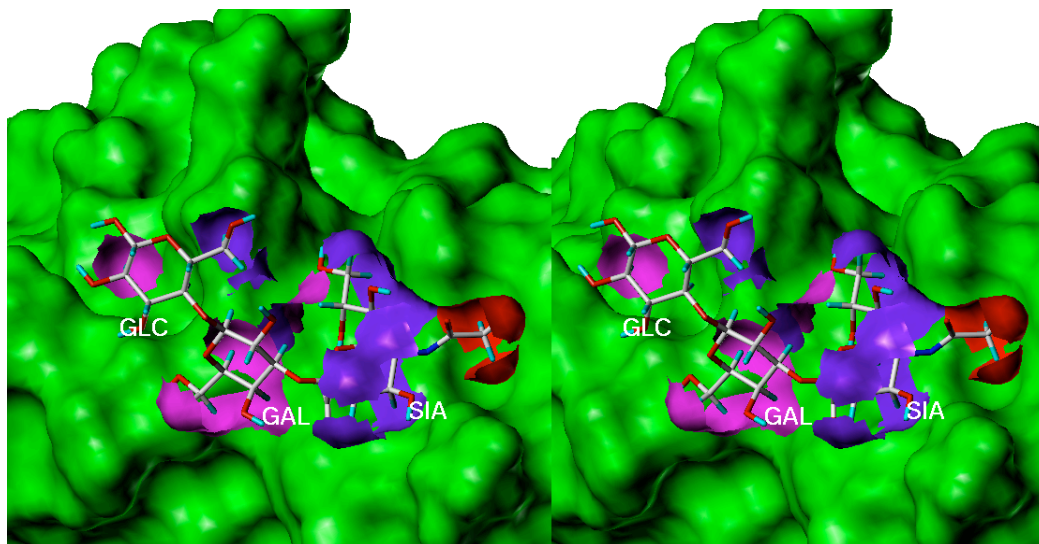


Figure 4.6: Experimental STD effects shown in the X-ray structure of α -(2 \rightarrow 3)-sialyl lactose **1** bound to sialoadhesin. The green MOLCAD surface shows the surface of the protein. Differentially coloured MOLCAD surfaces of sialyl lactose **1** are used to distinguish the size of STD effects. A red colour indicates strong STD effects whereas purple and magenta show the medium and small STD effects respectively.

4.1.1.2 Flexible Dock calculations

The α -(2 \rightarrow 3)-sialyl lactose **1** was first manually docked into the active site of sialoadhesin using the geometry of crystal structure (May et al., 1998) as a guide. The optimisation of the binding geometry was then performed using the program Flexible Dock (FlexiDock), part of SYBYL software package (Tripos, USA). The protein was kept rigid whereas the ligand was fully flexible and allowed to move in the binding site. The interaction energy was calculated using van-der-Waals, electrostatic and torsional energy terms of the Tripos force field (Tripos, USA). The energy of the entire complex was then minimised in 30 cycles using FlexiDock. The result of the docking is shown in Figure 4.7. The dihedral angles ϕ and ψ at the α -(2 \rightarrow 3)-glycosidic linkages in **1** obtained from the Flexidock model were almost identical to the values from the X-ray structure (Table 4.1). The comparison of the torsional angles in the *exo*-cyclic side chains of the pyranose rings of **1** (Table 4.2) also corresponds well with the values from the crystal structure. The global root mean square deviation (RMSD) of the FlexiDock model compared to the crystal structure was 0.75 Å.

Table 4.1: The dihedral angles (ϕ and ψ) (in degrees) at the glycosidic linkages were defined as ϕ 1, C1''-C2''-O2'-C3' (Neu5Ac-Gal); ψ 1, C2''-O2''-C3'-H3' (Neu5Ac-Gal) and ϕ 2, H1'-C1'-O4-C4 (Gal-Glc); ψ 2, C1'-O1'-C4-H4 (Gal-Glc).

	ϕ 1	ψ 1	ϕ 2	ψ 2
X-ray	– 70	– 18	34	– 11
FlexiDock	– 68	– 19	49	– 43

The FlexiDock results (Figure 4.7) for the complex of α -(2 \rightarrow 3)-sialyl lactose **1** with sialoadhesin lead to a model that was very similar to the crystal structure. The model revealed that the carboxyl function of Neu5Ac formed a salt bridge with the guanidine group of Arg 97. The acetamido methyl group showed a van-der-Waals contact with the indole ring of Trp 2 and C9'' of Neu5Ac made a hydrophobic contact to the aromatic side chain of Trp 106. The OH8'' and OH9'' of Neu5Ac formed hydrogen bonds with the main chain amide carbonyl of Leu 107 (~ 1.8 and ~ 1.5 Å respectively). The NH5'' of Neu5Ac formed a hydrogen bond with the main chain carbonyl function of Arg 105 and OH4'' interacted with the main chain carbonyl of Ser 103. The Gal formed only one hydrogen bond with the protein between OH6' and the hydroxyl group of Tyr 44. The H5' proton of Gal made van-der-Waal contacts with

Leu 107. Therefore, the key interaction for the sialoadhesin was predicted to be a hydrophobic interaction between the aromatic ring of Trp 2 and *N*-acetyl methyl group of Neu5Ac on the basis of X-ray structure and Flexidock model and was confirmed by STD NMR experiments.

Table 4.2: The torsional angles (ω) (in degrees) for side chains are defined as ω (C7''), C8''-C7''-C6''-O6'' (Neu5Ac); ω (C8''), C9''-C8''-C7''-C6'' (Neu5Ac); ω (C9''), O9''-C9''-C8''-C7'' (Neu5Ac); ω (C6'), O6'-C6'-C5'-O5' (Gal); ω (C6), O6-C6-C5-O5 (Glc).

	ω (C9'')	ω (C8'')	ω (C7'')	ω (C6')	ω (C6)
X-ray	-56	-173	-176	77	69
FlexiDock	-54	-164	179	60	174

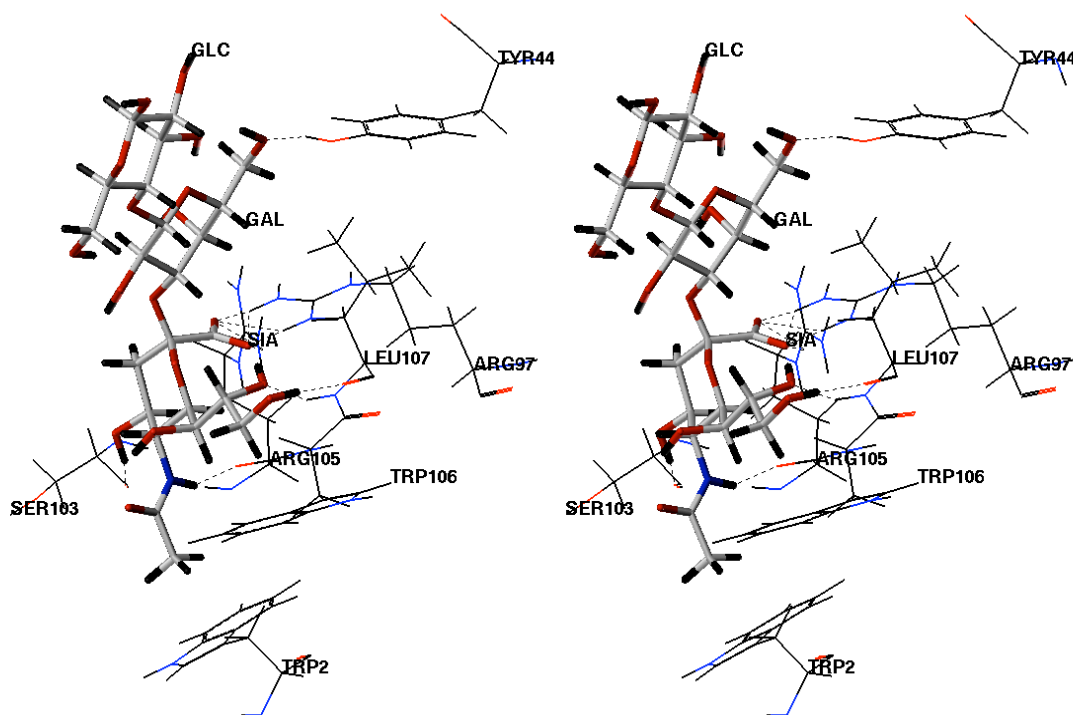


Figure 4.7: Stereo picture (relaxed view) of α -(2 \rightarrow 3)-sialyl lactose **1** in the binding pocket of sialoadhesin showing the hydrogen bonds between protein and trisaccharide **1** obtained from the FlexiDock calculations. The ligand is represented in stick mode whereas lines represent the amino acids. Dotted lines denote hydrogen bonds between the ligand and the protein.

4.1.1.3 Discussion

Quantitative STD analysis allows one to characterise the binding epitope of α -(2→3)-sialyl lactose **1** bound to sialoadhesin at an atomic resolution. The experimental STD effects are compared with the crystal structure of the complex. A docking study is performed on the basis of this crystal structure using the FlexiDock algorithm and substantiated the assumption that the overall conformation of the complex is very similar in the crystal and in aqueous solution.

Generally, the samples containing the ligand at lower concentration do not show any STD effect. But at higher concentrations (5 mM in this case) of ligand a correction is required because spurious effects were observed in the STD spectra of the ligand only. Thus the correction was achieved by subtracting the STD effects of ligand alone from the final STD spectrum of the ligand with protein. The experimentally determined STD effects (cf. Figure 4.4 and 4.5) are depicted in Figure 4.6. The degrees of saturation of the ligand are dependent on the distance between ligand and protein. A colour coding has been employed to illustrate the results of the group epitope mapping. It is clear that those protons of **1**, which are closer to protein in the binding pocket of sialoadhesin, show larger STD effects. The strongest STD effect was observed for the *N*-acetyl methyl group of α -(2→3)-sialyl lactose **1**. This is in excellent agreement with the X-ray structure of **1** bound to sialoadhesin where a strong hydrophobic interaction between Trp 2 of sialoadhesin and *N*-acetyl methyl group of sialyl lactose **1** is observed (May et al., 1998).

The dramatic effect on the binding affinity of sialoadhesin for α -(2→3)-sialyl lactose **1** upon mutation of Trp 2 indicated that this interaction makes a significant contribution to the binding energy (May et al., 1998; Crocker et al., 1999). All of the protons of Neu5Ac residue give a response in the STD spectrum, indicating that Neu5Ac is in close contact with the protein surface (Figure 4.4 and 4.6). From comparison with the crystal structure it is seen that the H4'' and H9'' protons are oriented towards the protein's interior and the corresponding hydroxyl functions OH4'' and OH9'' are hydrogen bonded to Ser 103 and Leu 107 respectively. In particular, H9'' is in close proximity of Trp 106 and shows a hydrophobic interaction with it.

The protons of the Gal residue show smaller STD effects (Figure 4.4 and 4.6) indicating that this residue has a less intimate contact with the protein surface. This finding is also in a good agreement with the site directed mutagenesis studies of

sialoadhesin (May et al., 1998). In the crystal structure a hydrogen bond between OH6' of Gal and Tyr 44 of sialoadhesin was observed. Mutation of Tyr 44 does not significantly affect adhesion of sialoadhesin to **1** (May et al., 1998). Therefore, this mutagenesis result implies that this hydrogen bond is not critically important for binding of α -(2 \rightarrow 3)-sialyl lactose **1** to sialoadhesin.

All of the D-Glc protons, except those of H2 β and H6 β do not show any STD response. This suggests that the Glc moiety is not critical for the binding process.

The experimental STD effects were compared with theoretical STD effects generated by the group of N. Rama Krishna, University of Alabama at Barmingham, USA (V. Jayalakshmi, personal communication) using the CORCEMA-STD calculation (Jayalakshmi and Krishna, 2002) based on the crystal structure of the sialoadhesin-**1** complex. The calculated STD effects are large for Neu5Ac residues whereas H3', H4', H5' and H6' of Gal show smaller STD effects (Figure 4.8). The calculated STD values for all the protons of Glc are very small, indicating that this residue is remote from the protein surface. These calculations are in good agreement with the experimental data. A comparison between calculated and experimental STD values for the individual protons revealed that the calculated STD values for the *N*-acetyl methyl group of Neu5Ac and the H6 β of Glc are significantly smaller than the experimentally determined values. This is not surprising as the T1 relaxation of *N*-acetyl methyl group of Neu5Ac can lead to an overestimation of the STD effect (Yan et al., 2003). In the crystal structure, the H6 β of Glc is too far from the protein surface to experience a saturation transfer. The R-factor is calculated from the quantitative comparison between the experimental and the calculated STD values. Its value was different for different protein to ligand ratios. Exclusion of STD values for the *N*-acetyl methyl group of Neu5Ac and H6 β of Glc, the R-factor varied from 0.4 (at 1:300) to 0.49 (at 1:500). The overall correlation between crystal and solution state data are obtained within an estimated error of 10-20 %.

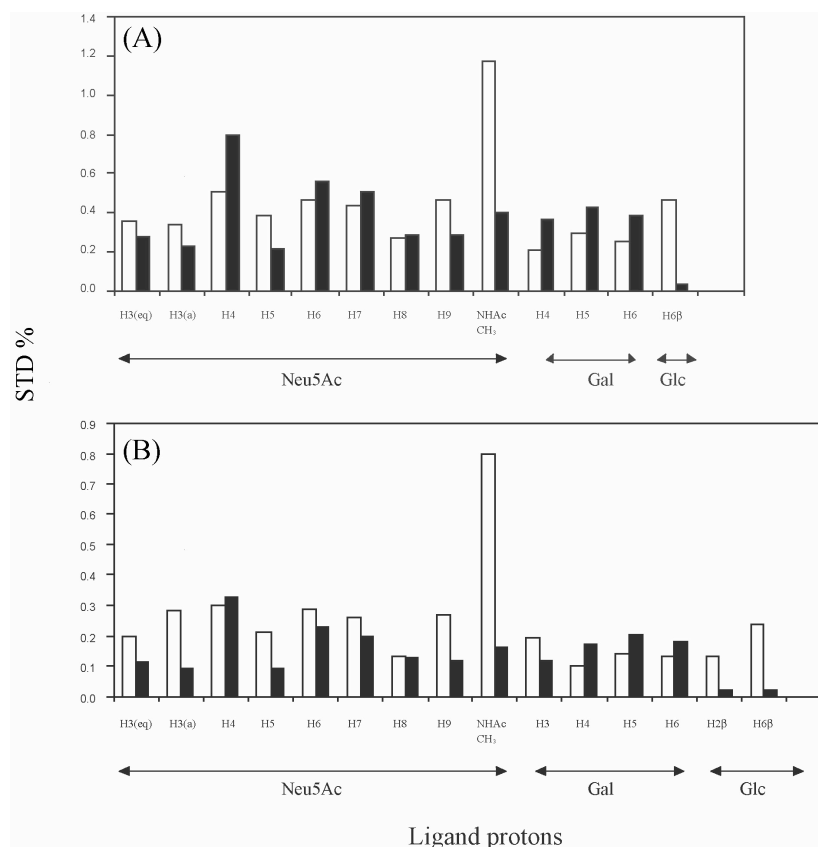


Figure 4.8: Comparison of experimental STD values (white bars) and predicted values from the CORCEMA-STD for the crystallographic structure (black bars). (A) Protein/ligand= 1:300 and (B) protein/ligand= 1:1000. STD values were calculated as $[(I_{0(k)} - I(t)_{(k)}) / (I_{0(k)})] \times 100$, with $I_{0(k)}$ being the intensity of the signal of proton k without saturation at time $t = 0$, and $I(t)_{(k)}$ being the intensity of proton k after a saturation transfer during the saturation time t .

4.1.2 Sialoadhesin-Trisaccharide 2 complex

In the previous chapter (Chapter 4.1.1) the solution state data was compared with crystal structure (May et al., 1998). In this chapter the 3D structure of the trisaccharide α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)- β -D-Gal-OSE **2** (Figure 4.9), a modified analogue of the natural ligand α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 4)- β -D-GlcNAc is studied in the presence of sialoadhesin using NMR spectroscopy. The analysis is based on the STD NMR and two-dimensional NOESY of sialoadhesin-trisaccharide **2** complex.

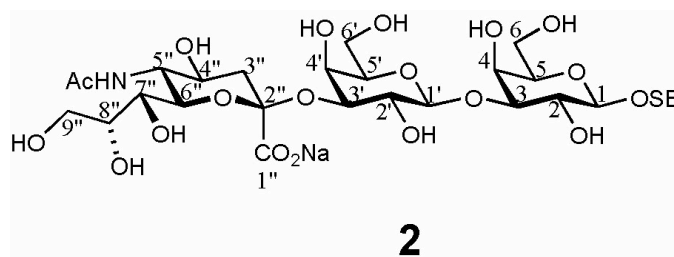


Figure 4.9: Structure of trisaccharide **2**, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)- β -D-Gal-OSE^a. ^a Key: -OSE, -O-(CH₂)₂-SiMe₃.

4.1.2.1 Free conformational analysis of trisaccharide **2**

4.1.2.1.1 NMR experiments

A complete assignment of ¹H NMR resonance signals (Figure 4.10) was achieved on the basis of NOESY, TOCSY and HSQC experiments. The chemical shift values are reported in Appendix (A.1.2). At 700 MHz, most of the proton resonances of trisaccharide **2** were clearly resolved.

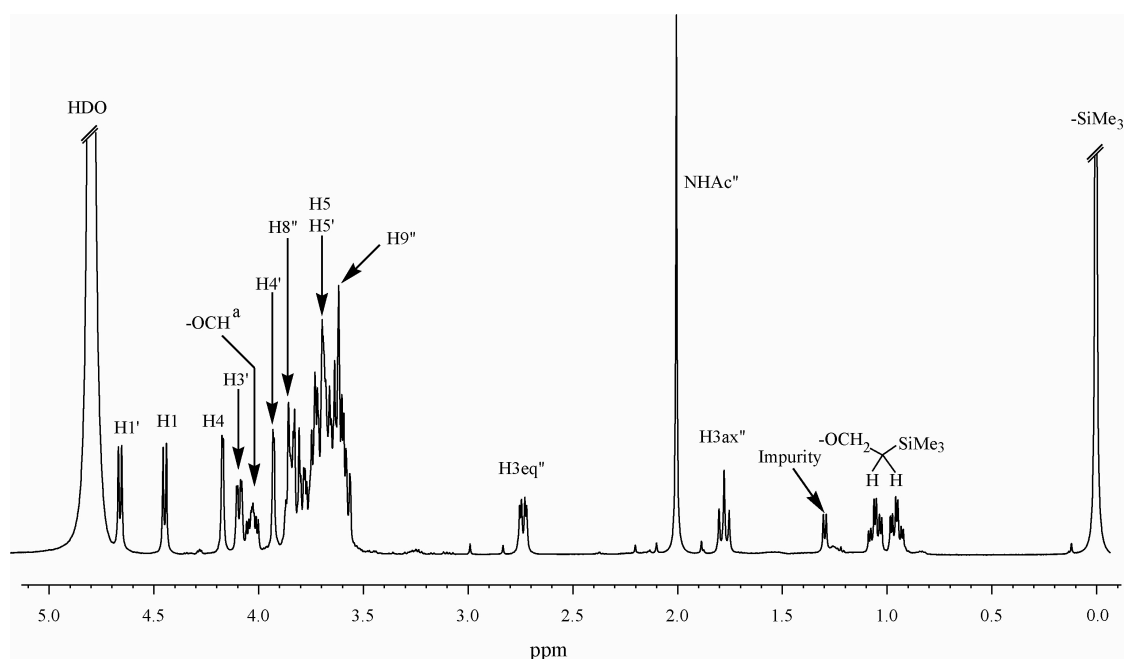


Figure 4.10: ¹H NMR spectra of trisaccharide **2** at 288 K and 700 MHz in 10 mM phosphate buffer, pH* 7.4 (with 150 mM NaCl). Chemical shifts were referenced to the attached -OSE groups at 0 ppm.

The observed NOEs were negative at 288 K and 700 MHz. *Interglycosidic* NOEs provide the information about the conformation of trisaccharide **2** in aqueous

solution. Five *interglycosidic* NOEs were observed: H1'-H3, H1'-H4, H3'-H3ax'', H3'-H3eq'' and H3'-H8'' (Figure 4.11). To complete the build up curves of different *intra-* and *interglycosidic* NOEs, the 2D NOESY spectra (Figure 4.11) were measured for seven mixing times. At the α -(2 \rightarrow 3)-glycosidic linkage, the *interglycosidic* H3'-H3ax'' NOE is very strong, indeed same strong as the corresponding H3ax''-H5'' NOE. A very weak H3'-H8'' and even weaker H3'-H3eq'' NOEs were also detected, although the latter was just above the noise level as shown in Figure 4.11.

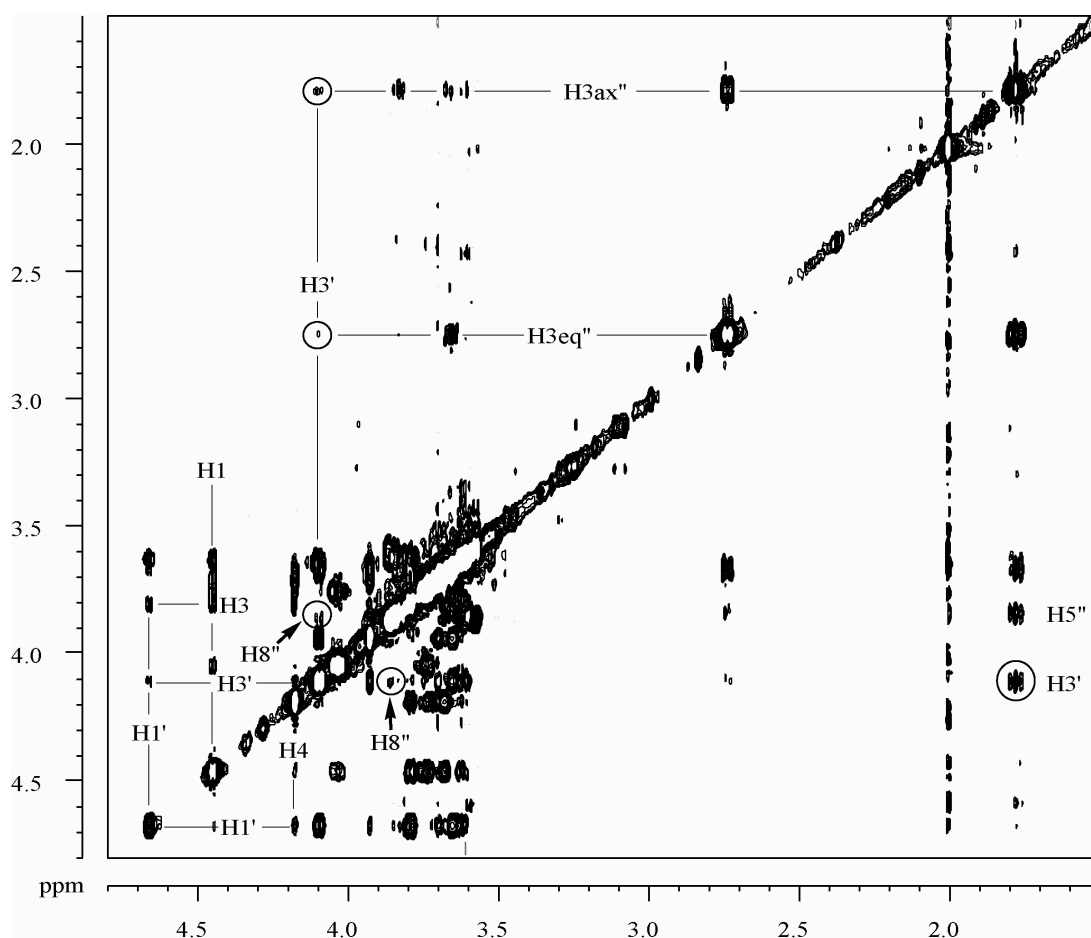


Figure 4.11: 2D NOESY spectrum of trisaccharide **2** (288 K, 700 MHz) in 10mM phosphate buffer, pH* 7.4 (with 150 mM NaCl). Mixing time was 500 ms. All the five *interglycosidic* NOEs, H1'-H3, H1'-H4, H3'-H3ax'', H3'-H3eq'' and H3'-H8'' are clearly visible in the spectrum.

The orientation around the β -(1 \rightarrow 3)-linkage should be defined by the NOEs between H1' of Gal and the protons of the another Gal moiety, especially H3, H4 and H2. Very strong H1'-H3 and weak H1'-H4 NOEs were observed (Figure 4.11) in the NOESY spectrum, indicating that the distance between H1'-H3 and H1'-H4 were very

small and large respectively. The corresponding H1'-H2 cross peak was just above the noise level.

4.1.2.1.2 Molecular mechanics

As the saccharides can adopt multiple conformations in aqueous solution, a search of conformational space was carried out using the molecular mechanics search routine. The conformations at the glycosidic linkages were systematically incremented by 10° . Three lowest minimum energy conformers (Table 4.3) were found at the relaxed ϕ , ψ surfaces (Figure 4.12). The conformations of the lowest minima were further energy minimised on a Silicon Graphics O2 R 10000 workstation with 2000 steps (dielectric constant, $\epsilon = 4$; r = distance) using Gasteiger-Hückel charge and Tripos force field (Tripos, USA).

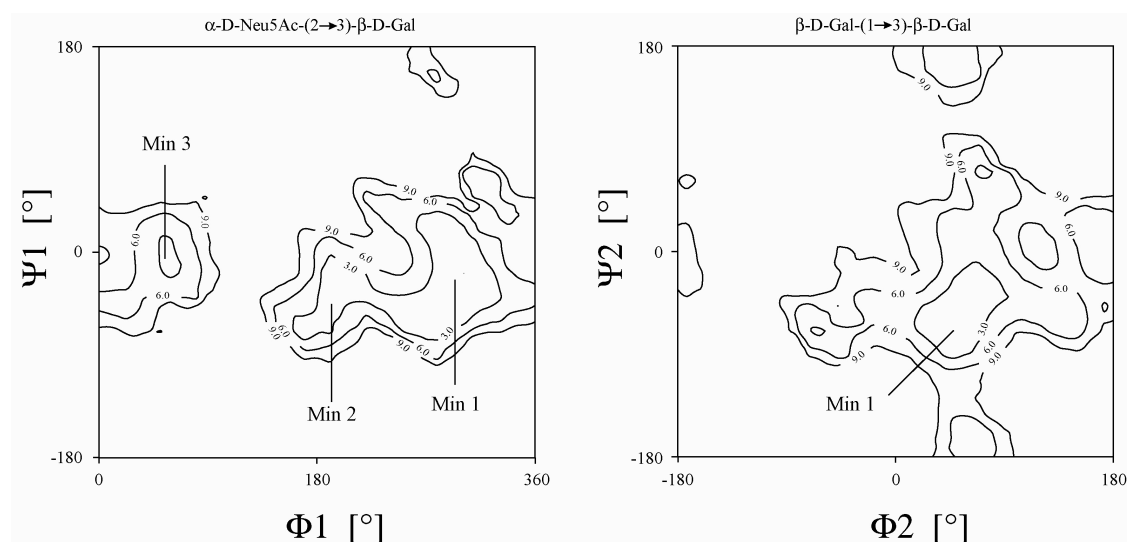


Figure 4.12: Relaxed energy maps of each glycosidic linkage of **2**. Three energy minima were found, designated as Min 1, 2 and 3 and the relative energies are listed in Table 4.3.

The flexibility of the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage is well known. Three different conformers- $\phi_1, \psi_1 = -58^\circ, -30^\circ$; $-156^\circ, -33^\circ$ and $64^\circ, -21^\circ$ were found at this linkage from the grid search analysis that were very similar to the values obtained by others (Breg et al., 1989; Scheffler et al., 1995; Poppe et al., 1997; Wu et al., 1999; Vasudevan and Balaji, 2002; Suresh and Veluraja, 2003). In accordance with the above calculations a global minimum, Min 1 was identified (Table 4.3). A local minimum, Min 2 has $\Delta E \approx 1.9 \text{ kcal mol}^{-1}$ higher in energy than global minima, Min 1.

It is this local minimum, Min 2, that accounts for large distances between H3'-H8'' and H3'-H3eq'' around 4.0 and 3.5 Å respectively, whereas the H3'-H3ax'' distance was around 2.2 Å. For the β -(1 \rightarrow 3)-linkage, the minima were located at ϕ 2, ψ 2= 50°, - 65°, corresponded well with the literature data (ϕ , ψ = 45°, - 65° and - 65°, - 18°) (Weimar et al., 2000). Therefore, the conformational properties of trisaccharide **2** in aqueous solution are described by a conformational equilibrium of at least Min 1 and Min 2.

Table 4.3: Relative energy (in kcal mol⁻¹) and dihedral angles (in degrees) at glycosidic linkages of **2**. The dihedral angles (ϕ and ψ) at the glycosidic linkages are defined as ϕ 1, C1''-C2''-O2'-C3' (Neu5Ac-Gal); ψ 1, C2''-O2''-C3'-H3' (Neu5Ac-Gal) and ϕ 2, H1'-C1'-O3-C3 (Gal-Gal); ψ 2, C1'-O1'-C3-H3 (Gal-Gal).

Min	ϕ 1	ψ 1	ϕ 2	ψ 2	ΔE (kcal mol ⁻¹)
1	- 58	- 29	50	- 65	0.0
2	- 156	- 33	50	- 65	1.9
3	64	- 21	49	- 65	3.6

4.1.2.2 Conformational analysis of trisaccharide **2** bound to sialoadhesin

4.1.2.2.1 NOESY experiment of sialoadhesin-trisaccharide **2** complex

The trNOEs of trisaccharide **2** at 288 K were also negative in the presence of sialoadhesin (Figure 4.13). Therefore, complete build up curves were necessary to distinguish properly between NOEs obtained from the free trisaccharide **2** and trNOEs originating from bound form (Figure 4.14). The NOESY spectrum of trisaccharide **2** in the presence of sialoadhesin showed significantly different trNOEs as compared to those in aqueous solution (Figure 4.13 A). An *interglycosidic* NOE between H3'-H3ax'' was negligible in the presence of protein (Figure 4.13 B). At the same time a strong trNOE was observed between H3'-H8'' (Figure 4.13 B), which was detected in very small amount in the free state. The same effect had been observed previously for the binding of sLe^x to E-selectin (Scheffler et al., 1997). This result indicates that the α -(2 \rightarrow 3)-glycosidic linkage undergoes a conformational reflection upon binding to sialoadhesin and the global minima, Min 1 is the one that is recognised by the protein (Figure 4.13).

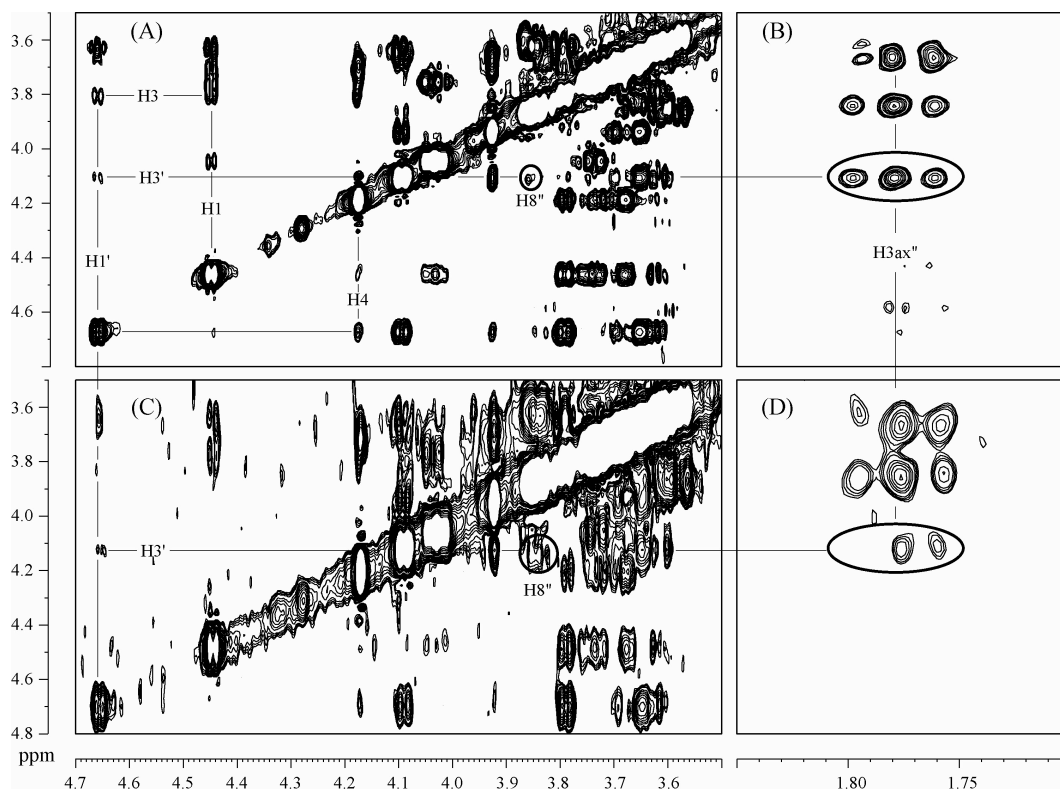


Figure 4.13: Parts of NOESY spectra of trisaccharide **2**: (A and B) in aqueous solution (288 K, 700 MHz, mixing time 500 ms); (C and D) in the presence of sialoadhesin (288 K, 700 MHz, mixing time 150 ms). The cross peak between H3'-H3ax'' is very weak in the complex whereas the cross peak between H3'-H8'' is of increased intensity.

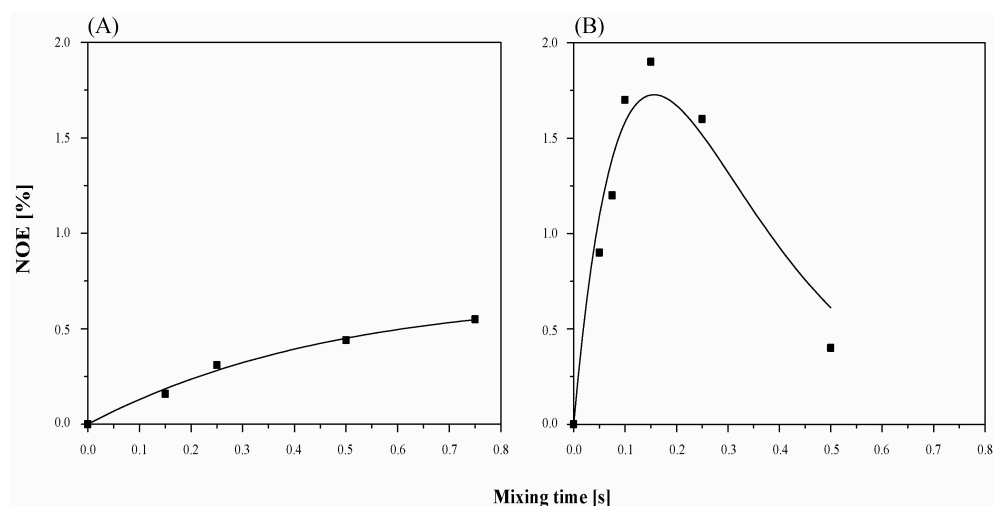


Figure 4.14: The NOE build up curves of trisaccharide **2** at different mixing times (A) in the absence of protein and (B) in the presence of sialoadhesin. The sialoadhesin:trisaccharide **2** molar ratio was 1:20 and both spectrums were processed with XWINNMR software (Bruker). The curve compares the NOE and trNOE that obtained from the H3'-H8'' cross peak. The maximum NOE enhancement for the free trisaccharide **2** was at mixing times of ca. 0.8-1.0 s, whereas for **2** bound to sialoadhesin the maximum trNOE enhancement was ca. at 0.2 s.

The *interglycosidic* trNOE, H1'-H3 at the β -(1 \rightarrow 3)-glycosidic linkage between the Gal and Gal residues was very large and similar to the NOE in the free state. A very weak H1'-H4 cross peak was also observed (Figure 4.13). No difference was observed between the free state and bound state conformations of the β -D-Gal-(1 \rightarrow 3)- β -D-Gal fragment. As the H1'-H4 cross peak was weak so ψ_2 angle deviated from the free state. Thus, global minimum energy conformation, Min 1 with ϕ_2 , $\psi_2 = 50^\circ$, -65° at this linkage is in very good agreement with the literature data (Bukowski et al., 2001). Additionally, a very weak trNOE, between H1-H4 was observed around 4.18 ppm in the NOESY spectrum of the sialoadhesin-trisaccharide **2** complex.

4.1.2.2.2 ROESY experiment of trisaccharide **2** in the presence of sialoadhesin

It is well known that indirect magnetisation transfer (spin diffusion) plays an important role for false distance constraints while generating a bioactive conformation of a ligand. Although, it is very unlikely that the strong trNOEs e.g., H3'-H8'', H1'-H3 and H1'-H4 arise entirely through spin diffusion, ROESY (Haselhorst et al., 1999; Maaheimo et al., 2000) experiments of trisaccharide **2** in the presence of sialoadhesin were performed which verified that the above *interglycosidic* trNOEs were not generated by spin diffusion via protein proton (Figure 4.15). For H1-H4 *intraglycosidic* trNOE the situation was different which was identified by spin diffusion artefact.

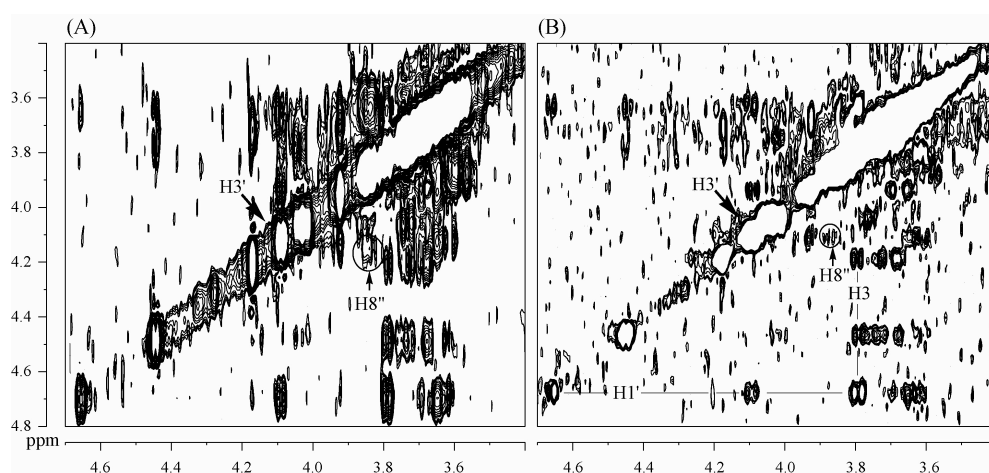


Figure 4.15: (A) NOESY (288 K, 700 MHz) and (B) ROESY (288 K, 500 MHz) spectrum of trisaccharide **2** in the presence of sialoadhesin. Mixing time was 150 and 100 ms for NOESY and ROESY experiments respectively. The trNOEs: H1'-H3, H1'-H4 and H3'-H8'' are not due to spin diffusion as they are present in the ROESY spectrum.

4.1.2.2.3 CORCEMA calculation

For the quantitative comparison of trNOEs, theoretical build-up curves were calculated using the program, CORCEMA (Moseley et al., 1995; Rinnbauer et al., 1998; Hricovini et al., 1999). The magnitude of the trNOEs depends on various parameters (in addition to geometry of the molecular complex) such as dissociation constant (K_D), off-rate constant (k_{off}) and the molecular correlation times of ligand (τ_c^{ligand}) and protein (τ_c^{protein}). The global minimum energy conformation, Min 1 of trisaccharide **2** was further optimised within the sialoadhesin-binding site using the program DOCK (part of the Sybyl software package) (Tripos, USA). The trNOEs were subsequently computed from the docking conformation of **2** in the sialoadhesin-binding site with the program CORCEMA that applies full relaxation and exchange matrix. A search was carried out in which all three parameters, K_D , k_{off} and τ_c^{ligand} were varied systematically. Table 4.4 shows the structure whose glycosidic linkage conformations best fits to experimental data. The computed trNOEs were then compared with experimental values (Figure 4.16) and R-factors were calculated (Figure 4.17). The theoretical trNOEs agreed well with the experimental trNOEs and the trNOE H1'-H3 was used as a reference. The magnitude of H1' diagonal peak was affected by presaturation of the residual HDO signal. As a result, all the cross-peaks related to H1' such as H1'-H3', H1'-H3 and H1'-H4 were also affected. The distance constraints that were derived from the computed trNOEs with the values of $K_D = 0.4$ mM, $k_{\text{off}} = 250$ s⁻¹ and $\tau_c^{\text{ligand}} = 0.3$ ns are summarised in Table 4.5.

Table 4.4: The torsional angles (in degree) at the glycosidic linkages for bound conformation of trisaccharide **2** found during the docking method and then compared with the CORCEMA program. The ϕ and ψ values are shown for trisaccharide **2** in the free state for comparison. The torsional angles derived from the crystal structure of sialyl lactose **1** in the presence of sialoadhesin (May et al., 1998) are also shown. The torsional angles (ϕ and ψ) at the glycosidic linkages are defined as follows: ϕ , C1''-C2''-O2'-C3' (Neu5Ac-Gal); ψ , C2''-O2''-C3'-H3' (Neu5Ac-Gal) and ϕ , H1'-C1'-O3-C3 (Gal-Gal); ψ , C1'-O1'-C3-H3 (Gal-Gal).

^afor X-ray structure, ϕ , H1'-C1'-O4-C4 (Gal-Glc); ψ , C1'-O1'-C4-H4 (Gal-Glc).

	$\phi 1$	$\psi 1$	$\phi 2$	$\psi 2$
Free state 2	– 156	– 33	50	– 65
Bound conformation	– 55	– 24	51	– 54
X-ray^a	– 70	– 18	34 ^a	– 11 ^a

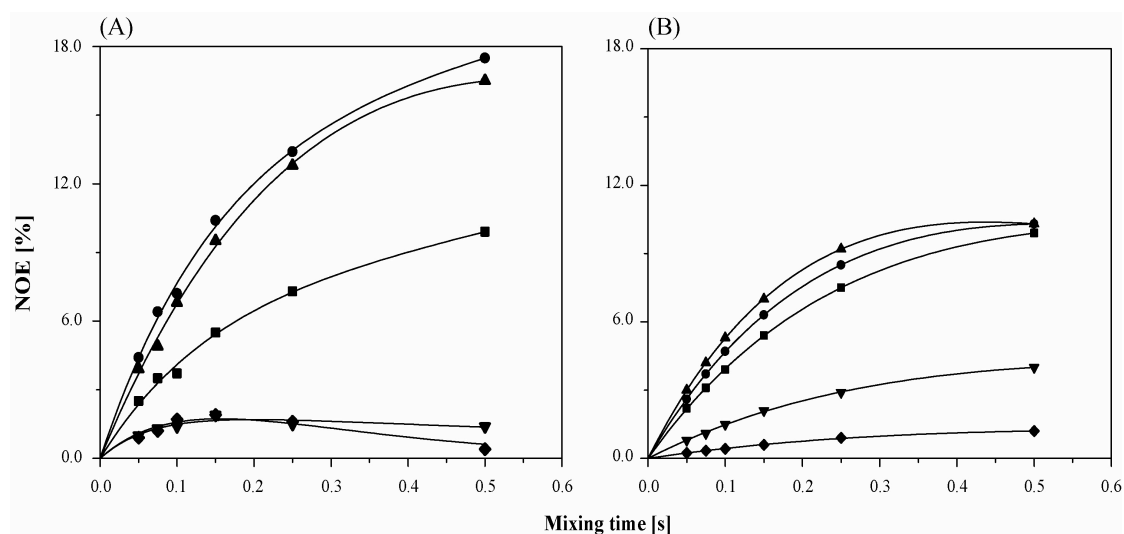


Figure 4.16: (A) Experimental and (B) CORCEMA calculated trNOE (%) build up curves of trisaccharide **2** in the presence of sialoadhesin at different mixing times. The sialoadhesin:trisaccharide **2** molar ratio was 1:20. Computed trNOEs were obtained with $K_D = 0.4$ mM, $k_{off} = 250$ s⁻¹, $\tau_c^{ligand} = 0.3$ ns and $\tau_c^{protein} = 40$ ns. The $\phi_1, \psi_1 = -55^\circ, -24^\circ$ and $\phi_2, \psi_2 = 51^\circ, -54^\circ$ was used for the CORCEMA calculated theoretical trNOEs. The curves represent following trNOEs: H1-H3 (■), H1'-H3' (●), H1'-H4 (▲), H1'-H4 (▼) and H3'-H8'' (◆).

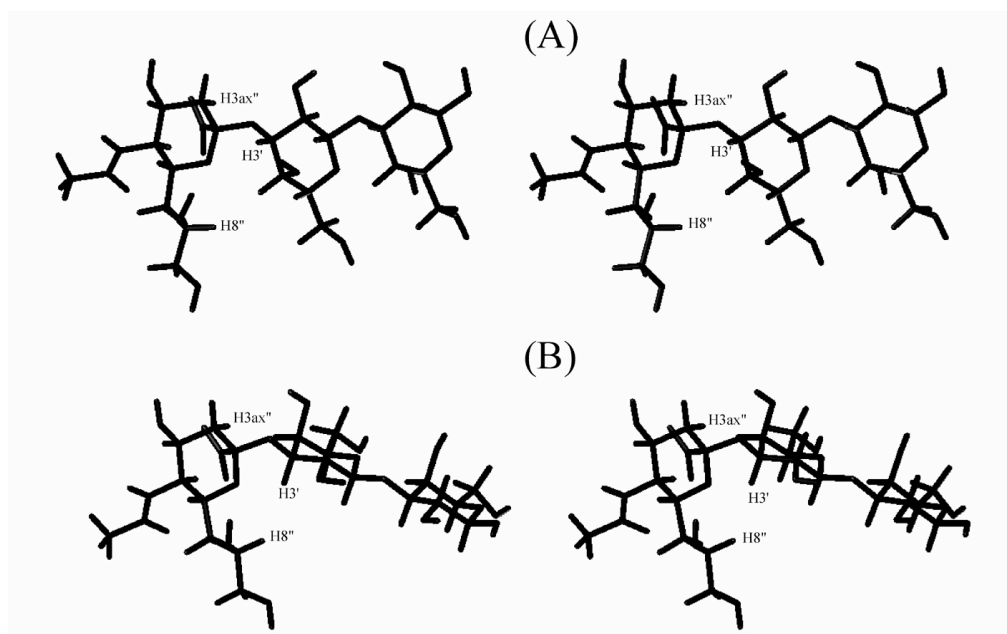


Figure 4.17: Free conformation of trisaccharide **2** is the mixture of (A) Min 2 and (B) Min 1 but Min2 is predominant in the conformational equilibrium due to the strong interaction between H3'-H3ax'' (Figure 4.13). (B) The bound conformation shows the main interaction between H3' of Gal and H8'' of Neu5Ac. In the bound conformation, Neu5Ac orientation is similar to the one found for α -(2→3)-sialyl lactose bound to sialoadhesin (May et al., 1998).

Good correspondence between experimental and calculated *interglycosidic* trNOEs (Table 4.5) shows that the bound conformation of the trisaccharide **2** is very close to the conformation that was used in the present calculation.

Table 4.5: Computed interproton distances and R-factors from the structures that interpreted NOESY spectra of **2** bound to sialoadhesin using CORCEMA program.

Restrains (symbols)	Distance (Å)	R-factors
H1-H3 (■)	2.60	0.04
H1'-H3' (●)	2.50	0.39
H1'-H3 (▲)	2.42	0.32
H1'-H4 (▼)	3.11	0.14
H3'-H8'' (◆)	2.93	0.70

4.1.2.2.4 Epitope mapping

To verify the binding mode of trisaccharide **2** bound by sialoadhesin in aqueous solution, 1D STD NMR experiments (Mayer and Meyer, 1999; Meyer and Peters, 2003; Bhunia et al., 2004) were performed. This type of NMR experiment creates a map of proton resonances of the ligand that are in close contact with the protein. The STD experiment transfers the effective magnetisation via spin diffusion from the protein to the ligand protons that are in close proximity to protein protons during the time the ligand spends in the complex. During this experiment, protein resonances were saturated at a position in the spectrum where no ligand resonances appeared (on resonance spectrum, I_{sat}). Another spectrum is recorded with the irradiation frequency set to a value that is significantly apart from all resonance frequencies of either of the protein or of the ligand (off resonance spectrum, I_0). Subtraction of the two spectra (on resonance–off resonance, $I_{\text{sat}}-I_0$) leads to a difference spectrum that contains only signals resulting from the saturation transfer. Figure 4.18 shows the saturation profile of *N*-acetyl methyl group of Neu5Ac at two different ligand concentrations and predicts that the longer saturation time is required to reach the maximum observable STD effect at high excess of ligand. At 0.36 mM concentration, a 2 s saturation time was more than sufficient while at a 1.8 mM concentration; ca. 2.5 s was required to reach the maximum STD effects.

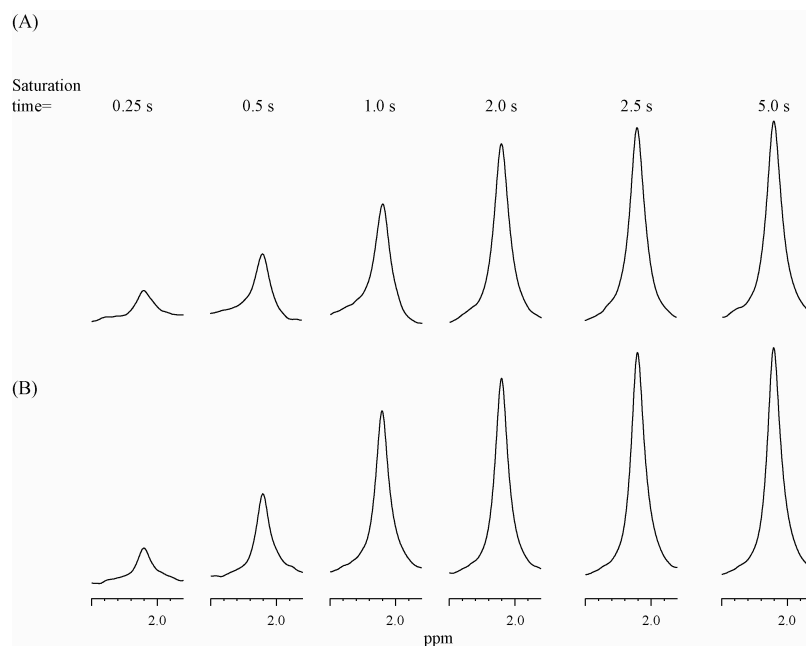


Figure 4.18: Selected part of 1D STD spectrum of the *N*-acetyl methyl group of trisaccharide **2** at (A) 0.36 mM and (B) 1.8 mM in the presence of sialoadhesin (18 μ M binding sites) in 10 mM phosphate-buffer, pH* 7.4 (with 150 mM NaCl). The spectrum was measured at 500 MHz and 288 K. For the STD NMR spectrum the protein was saturated with 0.25, 0.5, 1, 2, 2.5 and 5 s (on resonance: -2 ppm and off resonance: 40 ppm).

The 1D STD NMR spectra of trisaccharide **2** bound to sialoadhesin is shown in Figure 4.19. The *N*-acetyl methyl group of Neu5Ac gave a large STD response, and therefore, it was in more intimate contact with the protein-binding site. The Neu5Ac residue led to the most prominent STD signals indicating that this residue was mainly responsible for the binding reaction. The β -D-Gal-(1 \rightarrow 3)- β -D-Gal moieties also play an important role for binding to sialoadhesin.

The different signal intensities of the individual protons were best analysed from the integral values, obtained from the reference and STD spectra respectively (Figure 4.19). The relative degrees of saturation for the individual protons are displayed in Figure 4.20. The H5'', *N*-acetyl methyl group (NHAc''), H6'', H7'', H9'', H3', H1', H3 and H1 protons all have STD intensities between 60 % and 100 %. On the other hand H3ax'', H3eq'', H4'', H8'', H4' and H4 have medium STD intensities ranging from 41 % to 58 %. The lowest intensities corresponded to H5', H6', H5 and H6, which reached the values of only 35 % each. Thus a clear distinction between protons with a strong contact to the protein and the others could be made.

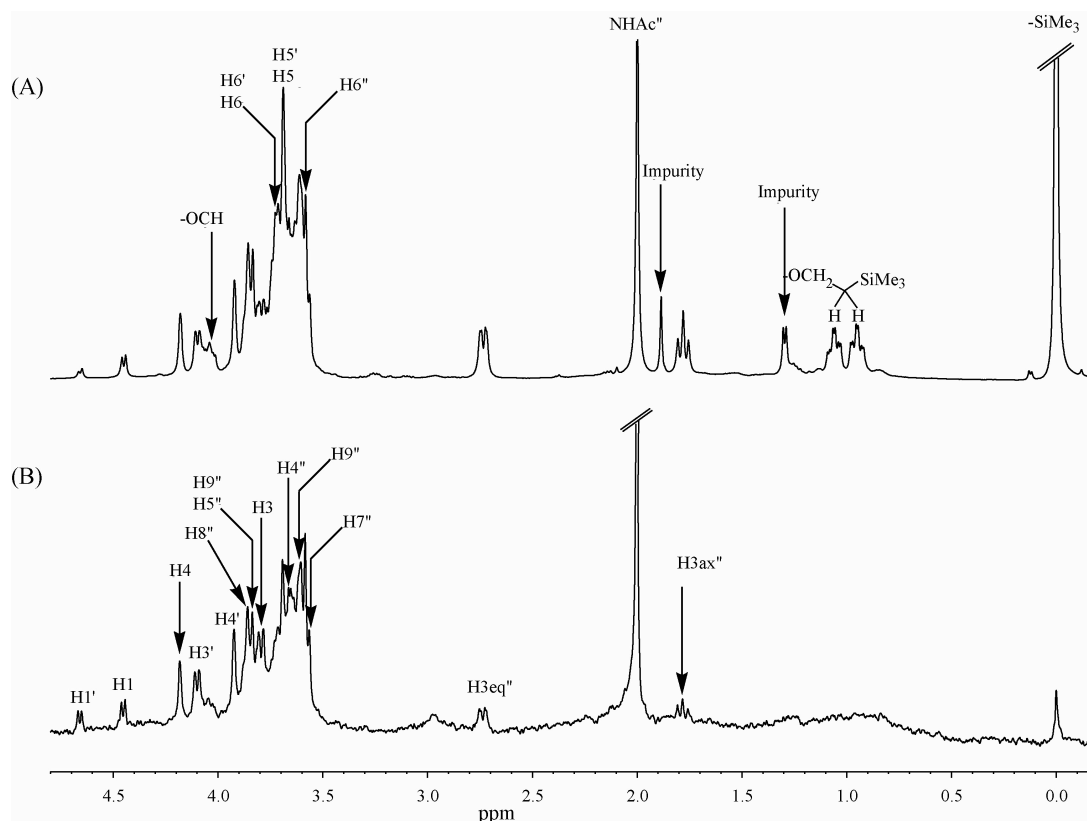


Figure 4.19: (A) Reference ¹H NMR spectrum of a mixture of sialoadhesin (18 μM binding sites) and trisaccharide **2** in a ratio of 1:140. (B) STD NMR spectra of the same sample. For the STD NMR spectrum the protein was saturated with 40 selective Gaussian-shaped pulses (50 ms each) resulting in a total saturation time of ~ 2 s (on resonance: – 2 ppm and off resonance: 40 ppm). The characterisation of the binding epitope of trisaccharide **2** can be performed with analysis of the relative intensities of the signals in spectrum B (cf. Figure 4.20).

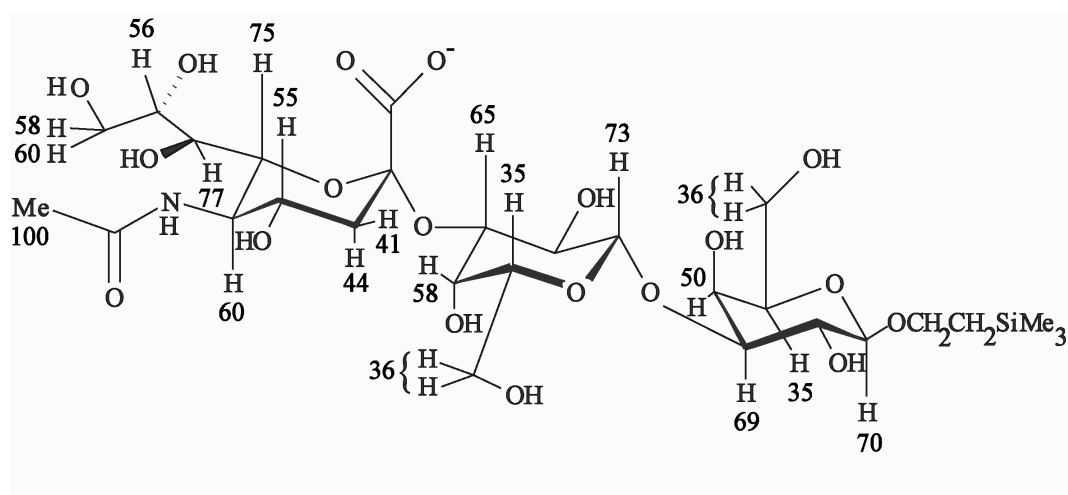


Figure 4.20: Structure of trisaccharide **2** and the relative degrees of saturation of the individual protons normalised to that of the *N*-acetyl methyl group as determined from 1D STD NMR spectra at a 140-fold excess. The concentration of sialoadhesin was 18 μM (binding sites) and that of trisaccharide **2** was 2.5 mM.

4.1.2.2.5 Modelling

The bioactive conformation of trisaccharide **2** was manually docked into the active site of sialoadhesin using the geometry of the crystal structure (acquisition code, 1QFO) (May et al., 1998) as a model (Figure 4.21). The docking was performed using the program DOCK (part of the Sybyl software package) (Tripos, USA) and Tripos force field (Tripos, USA).

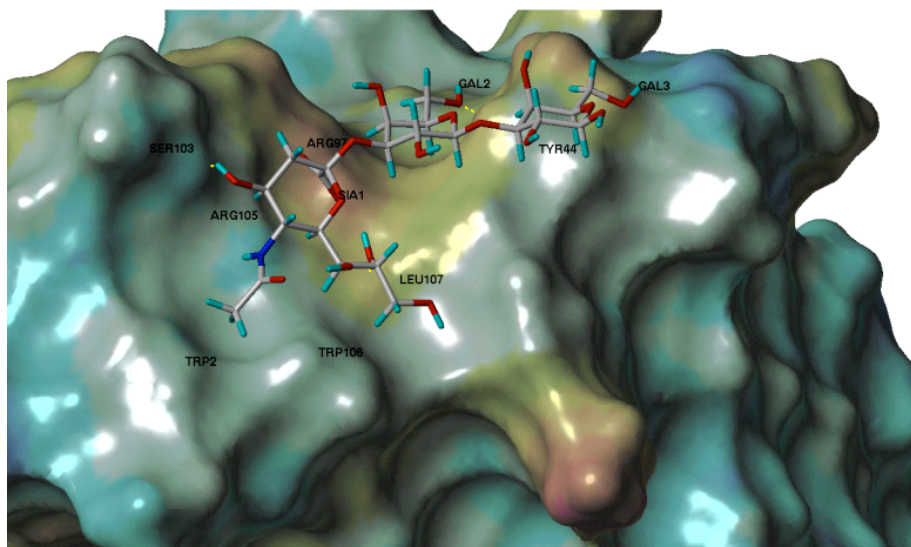


Figure 4.21: 3D views of trisaccharide **2** bound to sialoadhesin. The electrostatic MOLCAD surface shows the surface of the protein. The ligand, **2** is represented in stick mode and the dotted lines (yellow colour) denote the hydrogen bonds between ligand and protein.

The putative binding of the trisaccharide **2** in the active site of sialoadhesin is shown in Figure 4.22. The Dock results for the complex of trisaccharide **2** with sialoadhesin lead to a model that is very similar to the crystal structure for sialoadhesin (May et al., 1998) and FlexiDock results (Bhunia et al., 2004). The carboxyl function of Neu5Ac formed a salt bridge with the guanidine group of Arg 97 and the acetamido methyl group of Neu5Ac showed van-der-Waals contact with the indole ring of Trp 2. The terminal carbon (C9) of the Neu5Ac side chain made a hydrophobic contact to the aromatic side chain of Trp 106. Two hydrogen bonds were formed by the OH8'' and OH9'' of Neu5Ac and the main chain carbonyl of Leu 107 (~ 1.8 and 1.6 Å, respectively). The amide nitrogen of the Neu5Ac formed a hydrogen bond with the main chain carbonyl of Arg 105. The OH6' formed only one hydrogen bond with the hydroxyl group of Tyr 44. The H4' was close to Arg 118 whereas the H1', H3' and H5'

made van-der-Waals contacts with Leu 107 (~ 3 , 2.5 and 3.8 Å respectively). The H5 proton was also close to the ring of protons of Tyr 44 (~ 2.7 Å) and H3 and H4 were ~ 3.9 and 4.2 Å respectively away from Leu 107. The H6 of terminal Gal formed van-der-Waals contact with Arg 48 (~ 2.9 Å) of sialoadhesin.

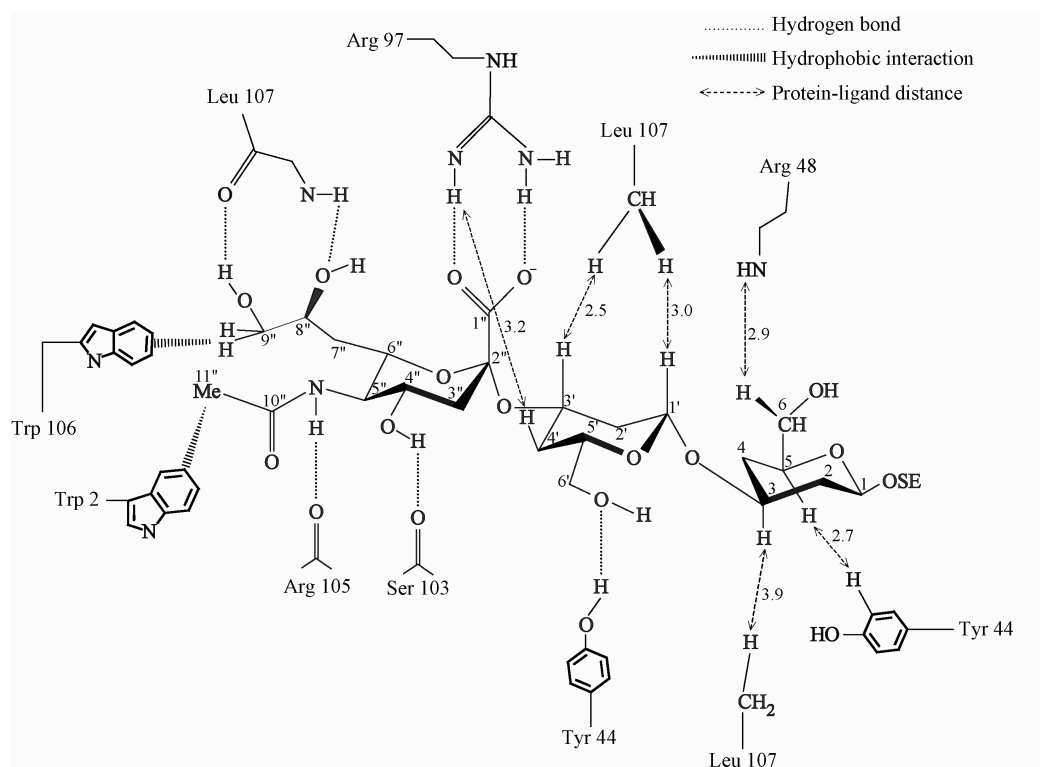


Figure 4.22: Schematic representation of sialoadhesin-trisaccharide **2** complex.

4.1.2.3 Discussion

TrNOE and STD NMR were employed as key techniques to characterise the binding of trisaccharide **2** to sialoadhesin at atomic resolution. For the trNOE techniques, the protein-mediated spin diffusion was largely excluded by performing ROESY experiments of **2** in the presence of protein (Haselhorst et al, 1999; Maaheimo et al., 2000). The possibility of protein-leakage that is not removed by the above experiment leads to slightly expanded structures for the bound ligands. Therefore, the experimental build up curves was compared with the theoretical build up curves, which were determined using the CORCEMA program, which utilise full relaxation and conformational exchange matrix (Moseley et al., 1995; Rinnbauer et al., 1998; Hricovini et al., 1999).

The geometry of the trisaccharide **2** in the free state was interpreted using NOESY experiments in combination with the molecular mechanics analysis. The α -(2 \rightarrow 3)-glycosidic linkage between Neu5Ac and Gal adopt three major conformations (Table 4.4) in the aqueous solution of **2**. This result is in good agreement with others obtained either by MD simulations or by MMC calculations (Breg et al., 1989; Sabesan et al., 1991; Scheffler et al., 1995; Poppe et al., 1997, Wu et al., 1999; Haselhorst et al., 2001). The result was even same either in the vacuum or in aqueous solution. The best fit of experimental NOEs of trisaccharide **2** was obtained with geometry of ϕ_1 , $\psi_1 = -156^\circ, -33^\circ$ and ϕ_2 , $\psi_2 = 50^\circ, -65^\circ$ (Table 4.4). The bioactive conformation of trisaccharide **2** in the presence of sialoadhesin was focused mainly on the glycosidic linkage conformation of **2**. It is observed from the bound state of trisaccharide **2** that the α -(2 \rightarrow 3)-glycosidic linkage is severely distorted in comparison to the free state (Figure 4.13). Computed trNOEs with global minima, Min 1 (ϕ_1 , $\psi_1 = -58^\circ, -29^\circ$) give theoretical values in a good agreement with those found experimentally (Table 4.5 and Figure 4.16). This bound conformation is very similar to the one that had been previously observed for sLe^x bound to E-selectin (Scheffler et al., 1997). The bioactive conformation of trisaccharide **2** at this linkage also corresponded well with the crystal structure of α -(2 \rightarrow 3)-sialyl lactose in the presence of sialoadhesin (Table 4.4) (Acquisition code, 1QFO) (May et al., 1998). Distorsions of the glycosidic linkages upon binding to proteins have already been observed before (Haselhorst et al., 1999, Hricovini et al., 2001), but it is difficult to estimate the energy that is required to allow the binding of such distorted conformation. This is not surprising as sialoadhesin is Neu5Ac binding lectin and the reorientation takes place around this α -(2 \rightarrow 3)-glycosidic linkage (Figure 4.23). The experimental data indicated that the bound conformation of β -(1 \rightarrow 3)-glycosidic linkage is close to the free conformation of this linkage (Table 4.4).

From the STD NMR experiments (Figure 4.19) it is clear that sialoadhesin provide a binding pocket for Neu5Ac. The strongest STD effect (Figure 4.20) is observed for the *N*-acetyl methyl group of trisaccharide **2**, as methyl protons are closest to the binding site (Figure 4.22). This is in excellent agreement with the X-ray crystal structure of sialyl lactose bound to sialoadhesin (May et al., 1998) where a strong hydrophobic interaction between Trp 2 of sialoadhesin and *N*-acetyl methyl group of α -(2 \rightarrow 3)-sialyl lactose was observed. The H4'', H5'', H8'' and H9'' have

about 55 % integral whereas H6'' and H7'' shows about 75 % STD effect. The comparison with the crystal structure show that the H4'' and H9'' are oriented towards the protein interior and the corresponding OH4'' and OH9'' are hydrogen bonded to Ser 103 and Leu 107 respectively (Figure 4.22). The H8'' proton is approximately 3.5 Å away from the aromatic ring of Trp 106 and sialoadhesin-2 model (Acquisition code, 1QFO) show that the OH8'' form a hydrogen bond with the Leu 107. The H7'' of Neu5Ac also pointing towards the protein interior, which explains its high STD effect. Apart from the Neu5Ac moiety, the H1', H3' and H4' show the largest intensity in the 1D STD spectrum (Figure 4.19). The H1' proton has a STD signal of ca. 70 %, and H3'-H4' has a STD effect of 55 %. This is not surprising as H3' and H4' are around 2.5 and 3.2 Å away from Leu 107 and Arg 97 (Figure 4.21 and 4.22) respectively. Unfortunately, the H5' and H6' protons of middle Gal overlapped with H5 and H6 protons of terminal Gal respectively. Therefore, at this time it is difficult to say conclusively whether H5' or H5 shows the STD effect. The same is true for H6' and H6 protons. The intensity decreases gradually with the increasing distance of the protons from the binding site. The H1 and H3 of Gal has about 70 % STD integral while H4 shows ca. 50 % due to its location far away from the protein. Fine details of molecular recognition for trisaccharide **2** as deduced from the trNOE experiments in combination with CORCEMA program are also in excellent agreement with the data from STD NMR experiments.

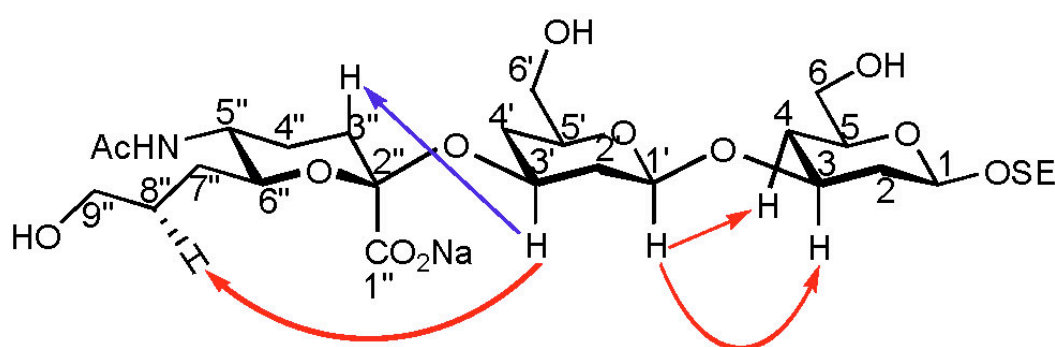


Figure 4.23: Main *interglycosidic* NOEs of trisaccharide **2** in the absence (blue arrow) and presence of sialoadhesin (red arrow) as inferred from 2D NOESY spectrum respectively. For clarity, all the functional groups except important protons are removed.

4.2 Myelin-Associated Glycoprotein

The myelin-associated glycoprotein (MAG), a siglecs (siglec-4a) (Kelm et al., 1994a) is composed of five separate immunoglobulin (Ig) domains (Lai et al., 1987). At the amino acid sequence level, the first two Ig domains of MAG share a 27.6 % homology with another siglecs member, sialoadhesin. The first extracellular domain of all the family members is a V-type Ig domain, containing the Neu5Ac binding site. The structure of N-terminal V-set domain of sialoadhesin in complex with α -(2 \rightarrow 3)-sialyl lactose (May et al., 1998) has been determined by X-ray crystallography (acquisition code, 1QFO). This crystal structure has three asymmetric subunits-A, B and C. But the subunit B was used as a template for homology model of N-terminal V-set domain of MAG due to its high electron density and minimal temperature factors (B factor). The structural alignment of the N-terminal V-set domain sequences of sialoadhesin and MAG are shown in Figure 4.24.

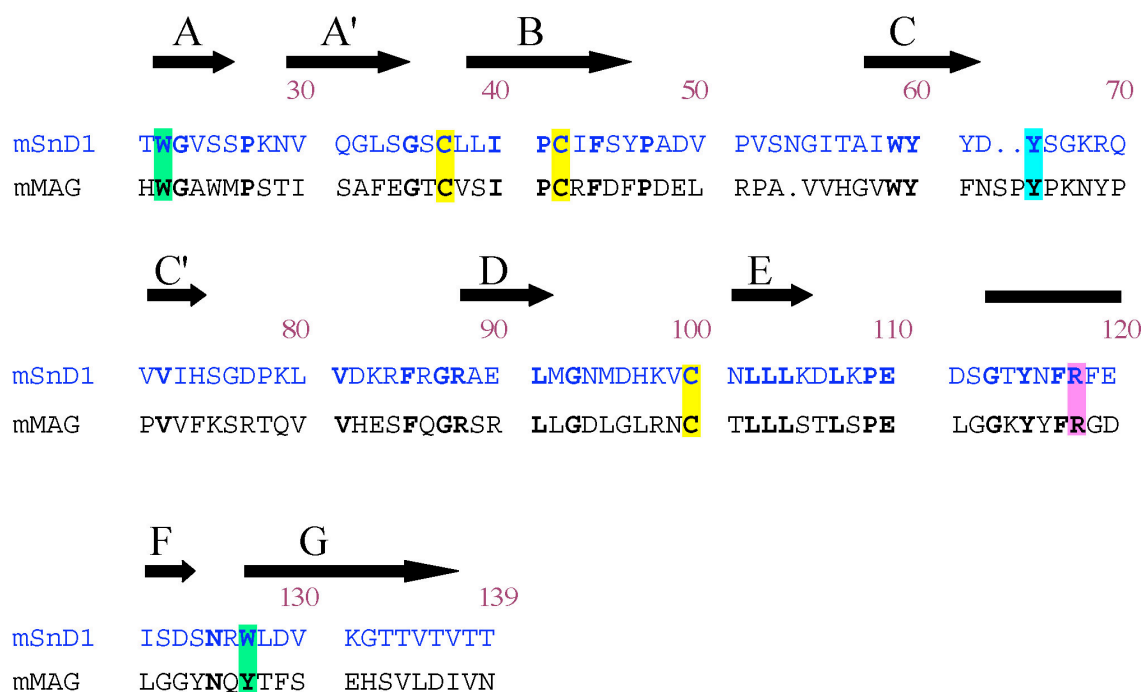


Figure 4.24: Amino acid sequence alignment of the N-terminal V-set domains of sialoadhesin and MAG. The aromatic residue putatively important in binding the ligands is shown in cyan box, while the siglec hallmark arginine and tryptophans are highlighted here in pink and green colour box respectively. The cysteines are shown in the yellow colour. The secondary structure elements are shown as arrows above the sequence. The sequences shown are of: mSnD1, mouse sialoadhesin and mMAG, mouse myelin-associated glycoprotein.

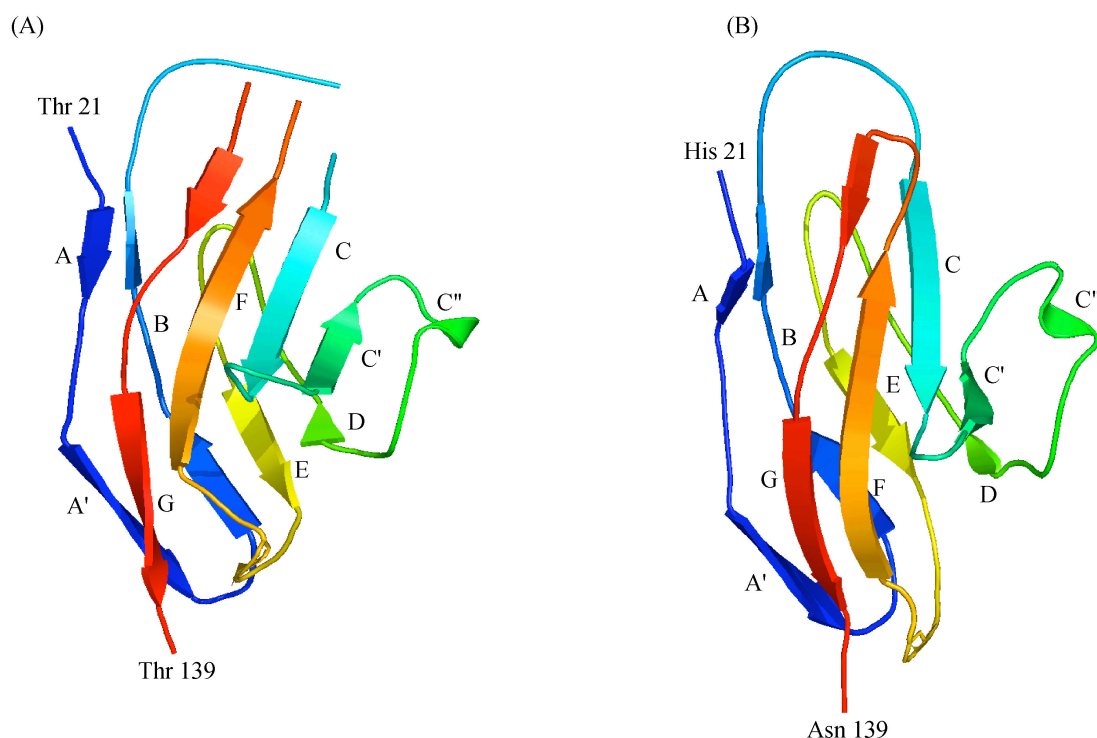


Figure 4.25: The 3D structure of the N-terminal V-set domain of (A) sialoadhesin and (B) MAG. MAG was modelled using the functionally related siglec, sialoadhesin as a template employing the program COMPOSER as provided by the Sybyl program package (Tripos, USA). The length of the two proteins is 118 residues each. MAG shares 27.6 % sequence identity with sialoadhesin. Each strand is labelled. The RMS deviation of the C α atoms of MAG was 1.7 Å with respect to those of the known crystal structure of sialoadhesin.

The 3D homology model of MAG consists of a β sandwich made of two β sheets termed ABED and A'GFCC' (Figure 4.25). Similar to the sialoadhesin crystal structure (Acquisition code, 1QFO) (May et al., 1998), the G strand of MAG was split up by a short insert into two shorter strands. The distinctive “kink” in the G strand occurred at residues 128-129 because of the stabilisation of hydrogen bonds between Thr 128 and Asn 99 and the strong hydrophobic interaction between Thr 128 and Tyr 116. The MAG model contains an intrasheet disulfide bond between Cys 42 and Cys 100, which is located on the B and E strands of the ABED sheet. The loop between C' and D in the main chain conformation occurred by formation of a short strand C'' in the variable domains. The COMPOSER software (Tripos, USA) found the fragments from the known crystal structure of sialoadhesin (Acquisition code, 1QFO) that are compatible with the rest of the model. From the comparison of siglecs family members

it was clear that the Neu5Ac binding residues are completely conserved throughout the family (Tang et al., 1997). The crucial residue for the binding of carboxyl group of Neu5Ac by MAG is Arg 118, which is located on the F strand. The missing parts, in the loop between B and C and in the loop between F and G were reconstructed in the model by COMPOSER software and Tripos force field (Tripos, USA). The overall geometric quality of the model was assessed by the program PROCHECK (Laskowski et al., 1993). In the homology model, 71.1 % amino acid residues were found in the most favoured regions of the Ramachandran plot (Figure 4.26), 2.1 % were in generously allowed regions and the rest 26.8 % were in additional allowed regions. No residues were found in the disallowed regions of the Ramachandran plot. The accuracy of the model was measured by calculating the root mean square (RMS) deviation of the C α atoms of MAG with respect to their position in the structure of sialoadhesin.

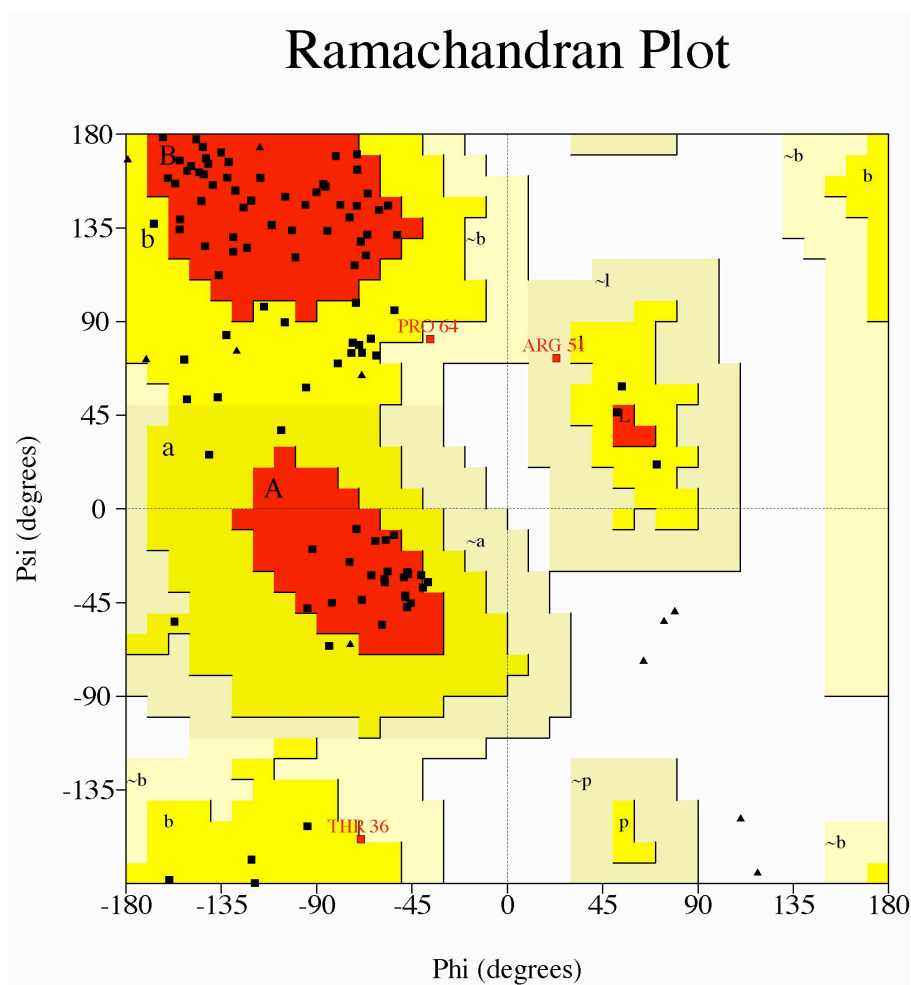


Figure 4.26: Ramachandran plot.

4.2.1 Free conformational analysis of saccharides 2-6

The myelin-associated glycoprotein (MAG) is a member of the siglecs family (Kelm et al., 1994a) expressed exclusively by oligodendrocytes and Schwann cells in the CNS and PNS respectively. Under normal physiological conditions, MAG is highly enriched at the contact sites between the myelin and axon (Schachner and Bartsch, 2000) and is responsible for neurite outgrowth inhibition (Mckerracher et al., 1994; Mukhopadhyay et al., 1994). Therefore, inhibitors of MAG could be beneficial for the treatment of injuries in the CNS. Generally, the gangliosides carry 75-80 % of the Neu5Ac in the brain (Tettamanti et al., 1973) and the major brain gangliosides GQ1b α , GD1 α and GT1b express the preferred α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc terminal determinant for MAG ligand (Collins et al., 1997). The synthetic and biological studies show that a second Neu5Ac residue attached to the β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc core of the natural ligand (GQ1b α N) enhances the binding affinity to MAG (Collins et al., 1999). For a better understanding of the binding of GQ1b α N to MAG, the qualitative and quantitative analysis of synthetic fragments 2-6 of GQ1b α N (Figure 4.27) in the context of their binding ability to MAG has been discussed systematically *in vivo*. In the upcoming chapters the free as well as the bound conformation of the synthetic saccharides 2-6, fragments of GQ1b α N, (Figure 4.27) will be discussed on the basis of NMR experimental results in combination with molecular modelling.

4.2.1.1 ^1H NMR experiments

In the ^1H NMR spectrum of the oligosaccharides 2-6, the majority of the ring protons fall within a narrow chemical shift region of ca. 3.4-4.1 ppm. The complete assignment of ^1H and ^{13}C NMR spectra was achieved using 2D methods-NOESY, TOCSY, HSQC and HMBC. The ^1H chemical shifts (A.1.2-A.1.6) were referenced relative to -OSE group, which is attached to the terminal Gal or GalNAc moiety. Most of the proton resonances (Figure 4.28) of the saccharides 2-6 (Figure 4.27) were well resolved at Bruker DRX 700 MHz as well as at 500 MHz.

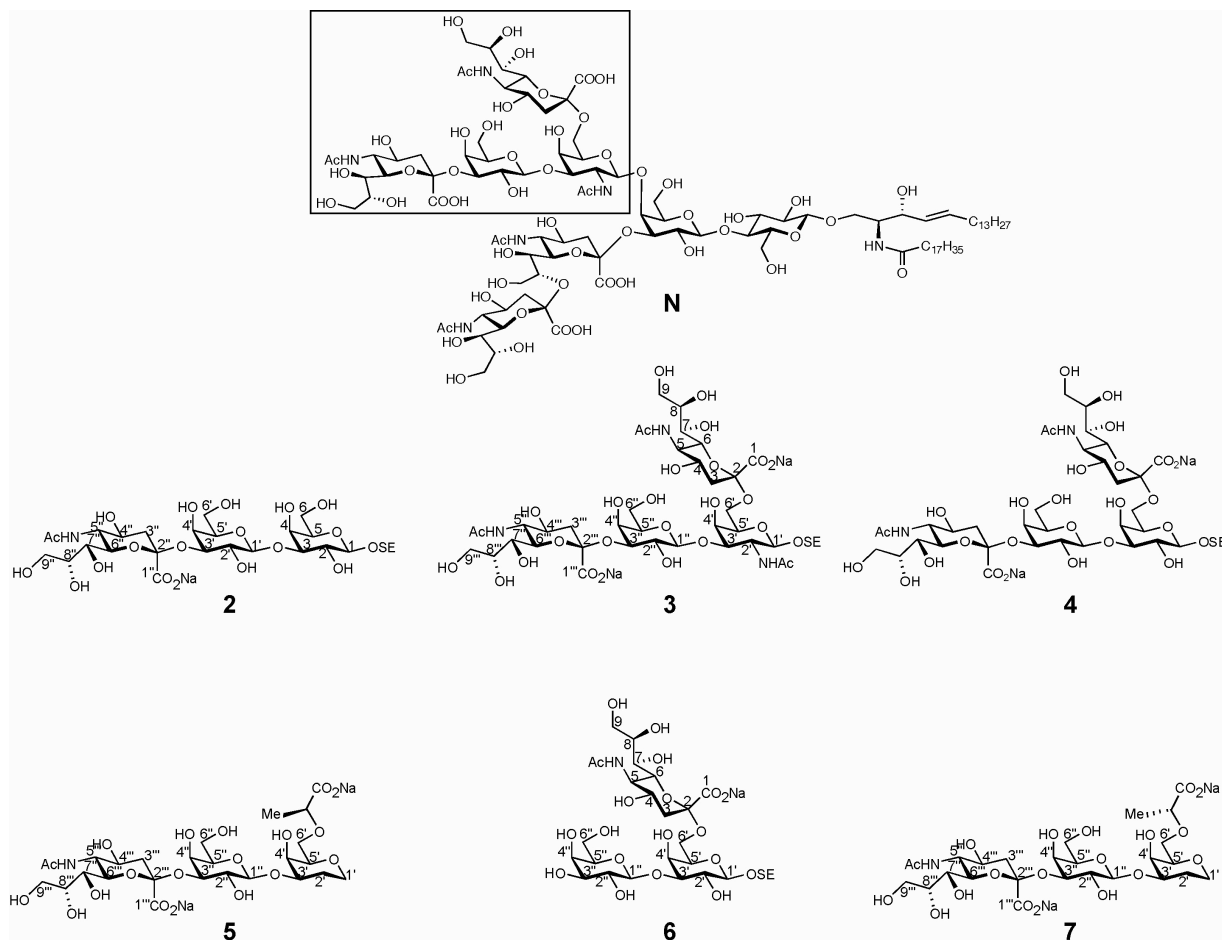


Figure 4.27: Structures of the investigated saccharides **2-6**. The saccharides **2-7** are the fragments of the natural ligand of MAG, GQ1bα N.^a Key: - OSE, -O-(CH₂)₂-SiMe₃.

Extension of **2** with Neu5Ac at the α -(2 \rightarrow 6)-linkage leading to **3**, results in coinciding the anomeric protons of Gal and GalNAc. This linkage affected the chemical shifts of H1', H2', H4', H5', H6'a, H1'' and H2'' by + 0.06, + 0.36, + 0.07, + 0.08, + 0.01, - 0.17 and - 0.12 ppm respectively. Most of the proton resonances of **3** were also overlapped. The H6'a chemical shift value of **3** were shifted strongly downfield in comparison to **2**, indicating the deshielding effects of H6'a of **3**. This suggested that the orientation of the Neu5Ac residue is specific with respect to the GalNAc H6 protons. As the tetrasaccharide **3** is the core of GQ1b α N and is mainly responsible for the binding to MAG, therefore, to overcome the anomeric proton-overlapping problem, the GalNAc residue of **3** was modified to Gal residue in tetrasaccharide **4**. The tetrasaccharide **3** was measured at 288 K and 700 MHz whereas the modified tetrasaccharide **4** was measured at 280 K and 500 MHz. Comparison of the ^1H NMR data for **3** and **4** showed that the modification of GalNAc to Gal residue induced the chemical shift of H1', H2', H1'' and H2'' by - 0.08, - 0.39, + 0.14 and + 0.10 ppm respectively. The signals for all anomeric protons were well resolved. The pseudo tetrasaccharide **5** is a derivative of **3**, where the α -D-Neu5Ac-(2 \rightarrow 6)- β -D-GalNAc linked Neu5Ac was replaced by (S)-LAc and the GalNAc moiety was replaced by Gal mimic cyclic ether (Cyc) residue. Removal of Neu5Ac residue from the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage of **4**, yielded trisaccharide **6** which affected the chemical shifts of H1'', H2'' and H3'' by + 0.08, + 0.05 and + 0.45 ppm respectively. The pseudo tetrasaccharide **5** and the trisaccharide **6** were measured at 288 K and 500 MHz.

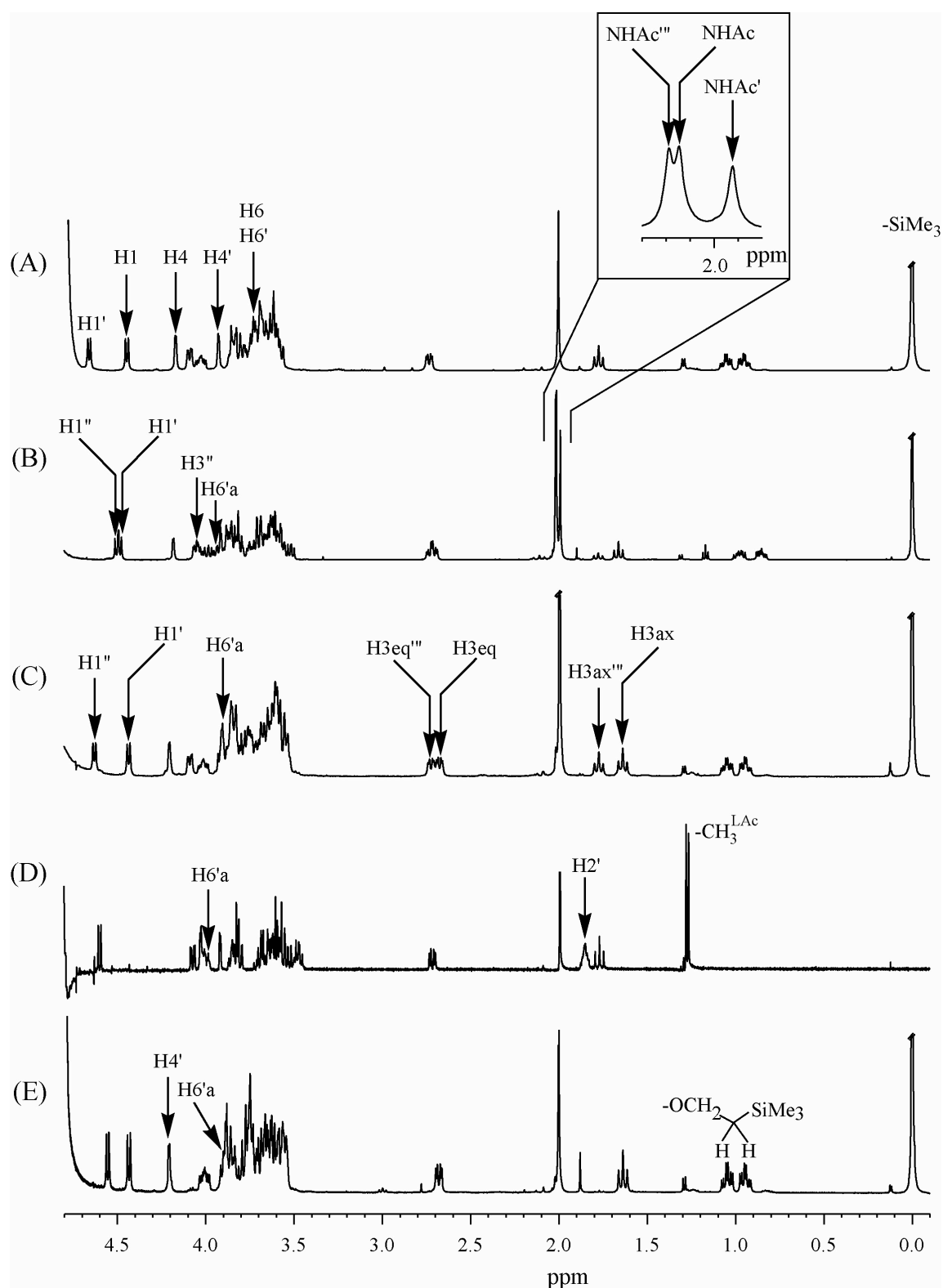


Figure 4.28: ^1H NMR spectrum of the synthetic saccharides **2-6** in 10 mM phosphate buffer, pH* 7.4 (with 150 mM NaCl). Here, (A) α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)- β -D-Gal-*OSE* **2** (700 MHz), (B) α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)-[α -D-Neu5Ac-(2 \rightarrow 6)]- β -D-GalNAc-*OSE* **3** (700 MHz), (C) α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)-[α -D-Neu5Ac-(2 \rightarrow 6)]- β -D-Gal-*OSE* **4** (500 MHz), (D) α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal-(1 \rightarrow 3)-[(S)-LAc-(2 \rightarrow 6)]-D-Cyc **5** (500 MHz) and (E) β -D-Gal-(1 \rightarrow 3)-[α -D-Neu5Ac-(2 \rightarrow 6)]- β -D-Gal-*OSE* **6** (500 MHz). Chemical shifts are referenced relative to -*OSE* group at 0 ppm, which is attached to the terminal Gal-GalNAc or Gal-Gal residue.

4.2.1.2 NOESY experiments

The NOESY spectrums of the saccharides **2-5** are shown in Figure 4.29. The saccharides **2** and **3** were measured at 700 MHz whereas the rest of the saccharides **4** and **5** were measured at 500 MHz. The NOE effects of the free saccharides **2**, **3** and **4** were negative at 288 K (tetrasaccharide **4** was measured at 280 K) whereas the pseudo tetrasaccharide **5** showed the NOEs were very close to zero crossing at 288 K.

From a set of seven NOESY spectrums of **2** and **3** with mixing times between 50 and 750 ms, it was possible to quantify both *interglycosidic* and *intraglycosidic* NOEs. All the above addressed saccharides **2-5** have a common back bone linkage, α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal. The NOESY spectrum conveys information about the distribution of different conformational families at the glycosidic linkage. It was earlier mentioned in case of trisaccharide **2** that H3' of Gal strongly coupled with the H3ax'' of Neu5Ac and this cross peak was as stronger as the corresponding *intraresidual* H3ax''-H5'' cross peak. The tetrasaccharide **3** and **4** and pseudo tetrasaccharide **5** also show the same effect (Figure 4.29). At the same time two very weak NOEs between H3''-H3eq''' and between H3''-H8''' were also observed just above the noise level at higher mixing times. The experimental NOEs at this linkage are in good agreement with the aqueous solution conformation of sLe^x (Haselhorst et al., 2001). Therefore, the above *interglycosidic* NOE between H3'' of Gal and H3ax'' of Neu5Ac suggested that the distance is very short of ca. 2.2 Å (Haselhorst et al., 2001). The relative intensity at the *interglycosidic* linkages of the saccharides **2-5** is shown in Table 4.6.

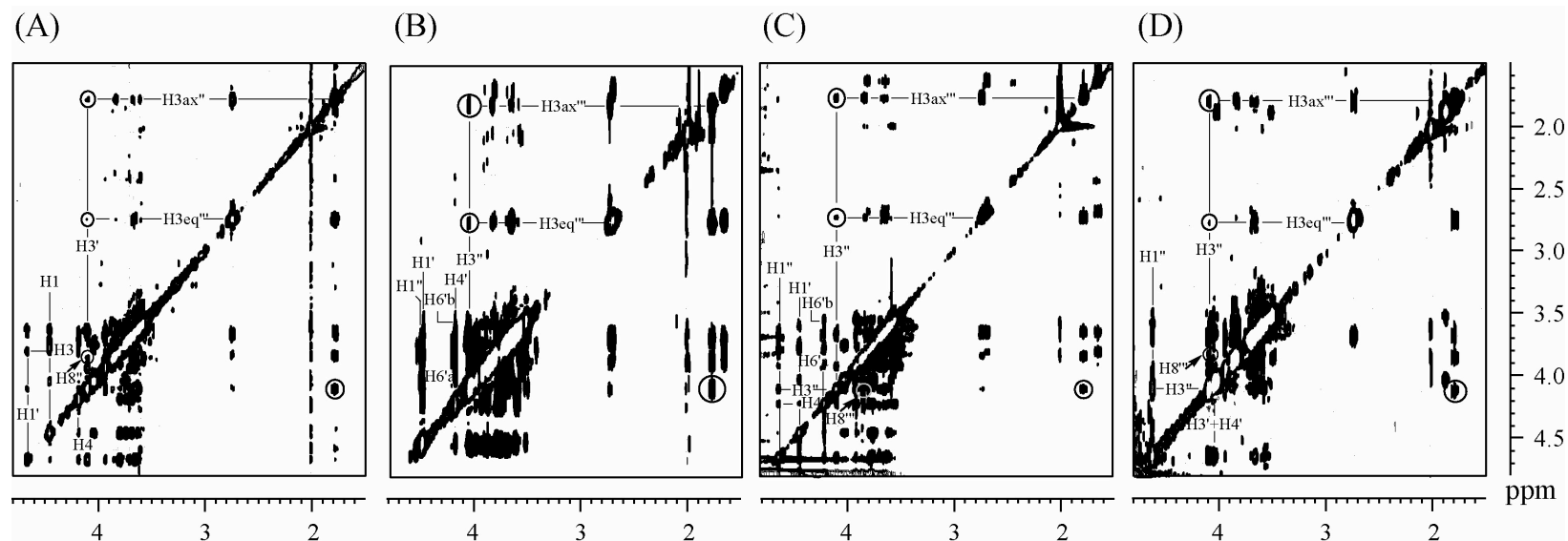


Figure 4.29: The NOESY spectra of the (A) trisaccharide **2** (700 MHz, 288 K), (B) tetrasaccharide **3** (700 MHz 288 K), (C) tetrasaccharide **4** (500 MHz, 280 K) and (D) pseudo tetrasaccharide **5** (500 MHz, 288 K). Mixing time of all the NOESY spectrums was same, 500 ms. In all spectra the strong NOE was observed either between H3' and H3ax'' (trisaccharide **2**) or between H3''-H3ax''' (saccharides **3-5**). The corresponding H3'-H3eq'' and H3'-H8'' (for trisaccharide **2**) or H3''-H3eq''' and H3''-H8''' (for saccharides **3-5**) were comparatively smaller.

The central disaccharide β -D-Gal-(1 \rightarrow 3)- β -D-Gal (or β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc of **3**) has much less conformational freedom in comparison to a α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage. Two *interglycosidic* NOEs: H1'- H3 and H1'- H4 of **2** or H1''- H3' and H1''- H4' of tetrasaccharides **3** and **4** were observed (Figure 4.29). The H1'-H3 *interglycosidic* NOE of **2** or the H1''-H3' NOE of **3** and **4** was very large whereas the H1''-H4' cross peak was relatively weak. The other *interglycosidic* NOEs: H1''-H2' and H1''-H5' of tetrasaccharide **3** were of negligible intensities.

Table 4.6: *Interglycosidic* NOEs of the saccharides **2-5** in 10 mM Phosphate buffer, pH* 7.4 (with 150 mM NaCl). Relative NOE intensities are defined as: strong (s), medium (m) and weak (w). (LAc: (S)-Lactic acid residue).

Saccharides	Interglycosidic NOEs		Relative intensity
2 (700 MHz, 288 K)	H3'	H3ax''	s
		H3eq''	w-m
		H8''	w
	H1'	H3	s
		H4	s
3 and 4 (3 : 700 MHz, 288 K and 4 : 500 MHz, 280 K)	H3''	H3ax'''	s
		H3eq'''	m
		H8'''	w
	H1''	H3'	s
		H4'	m
	H6'	H3ax	w
		H4	w
		H4'	m
5 (500 MHz, 288 K)	H3''	H3ax'''	s
		H3eq'''	w
		H8'''	w
	H1''	(H3'+H4')	s
		(H2ax'+H2eq')	m (from ROESY)
	H6'	-CH ^{LAc}	s
		-CH ₃ ^{LAc}	w (from ROESY)

For the β -D-Gal-(1 \rightarrow 3)-D-Cyc moiety of pseudo tetrasaccharide **5**, a strong overlap was observed for the H3' and the H4' resonances. Therefore, it was not possible to compare the corresponding *interglycosidic* NOEs, H1''-H3' and H1''-H4' in size. A ROESY experiment revealed an additional medium size ROE between H1''

and H2' (Table 4.6). The corresponding NOE was difficult to observe, as all NOEs were very weak due to an unfavorable correlation time of pseudo tetrasaccharide **5**. The two protons H2eq' and H2ax' display the same chemical shift, and therefore, distinction between the NOEs of H1''-H2eq' and H1''-H2ax' could not be made (Figure 4.29).

The conformational analysis of the tetrasaccharide **3** at the α -D-Neu5Ac-(2 \rightarrow 6)- β -D-GalNAc or tetrasaccharide **4** at the α -D-Neu5Ac-(2 \rightarrow 6)- β -D-Gal linkage could not be analysed due to severe signal overlap. However, the H4'-H6'a and H4'-H6'b of tetrasaccharides **3** and **4** were undisturbed by other cross peaks. From the volume integral, it is clear that the H4' of GalNAc (for tetrasaccharide **3**) or Gal (tetrasaccharide **4**) is closer to H6'b than H6'a.

For the pseudo tetrasaccharide **5**, the NOE intensity of the H6'-CH^{LAc} was more intense than that of the H6'-CH₃^{LAc} cross peak (Table 4.6). No other *interglycosidic* cross peaks were observed at the (S)-LAc-(2 \rightarrow 6)-D-Cyc linkage.

4.2.1.3 Simulation

The conformation of the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage has been extensively studied previously in different oligosaccharides in both aqueous and non-aqueous solvents (Breg et al., 1989; Sabesan et al., 1991; Scheffler et al., 1995; Poppe et al., 1997, Wu et al., 1997; Haselhorst et al., 2001). Most of the results suggested that there is an inherent flexibility of this linkage, leading to exchange amongst several distinct conformations. Calculations using the hard sphere *exo*-anomeric approach (HSEA) (Veluraja and Rao, 1983, 1984) and Metropolis Monte Carlo (MMC) (Haselhorst et al., 2001) suggested that three conformational regions possible for the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage. In other work, Mukhopadhyay and Bush (1994) reported four local minima for this linkage. Molecular mechanics calculations were carried out using the Tripos force field (Tripos, USA) and the resulting relaxed energy maps at the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage for the saccharides **2**, **3** and **7** are shown in Figure 4.30. For all the α -(2 \rightarrow 3)-linkages, the global minimum is found for values of the glycosidic torsions around ϕ , $\psi = -60^\circ, -30^\circ$ (Min 1) and two other minima with torsional angles ϕ , $\psi = -160^\circ, -30^\circ$ (Min 2 for **2** and **3** or Min 3 for **7**) and ϕ , $\psi = 60^\circ, -20^\circ$ (Min 3 for **2** or Min 4 for **3** and Min 5 for **7**) are also predicted.

The relative energies of the conformations corresponding to the global and local minima of saccharides **2**, **3** and **7** are shown in Table 4.7, indicating that Min 1 is the global minimum with ϕ , ψ values of -60° and -30° and Min 2 and Min 3 are the local minima with energy differences around 1.9 and 3.6 kcal mol⁻¹ respectively for trisaccharide **2**. In case of tetrasaccharide **3** and pseudo tetrasaccharide **5**, approximately of the same energy difference was also predicted (Table 4.7).

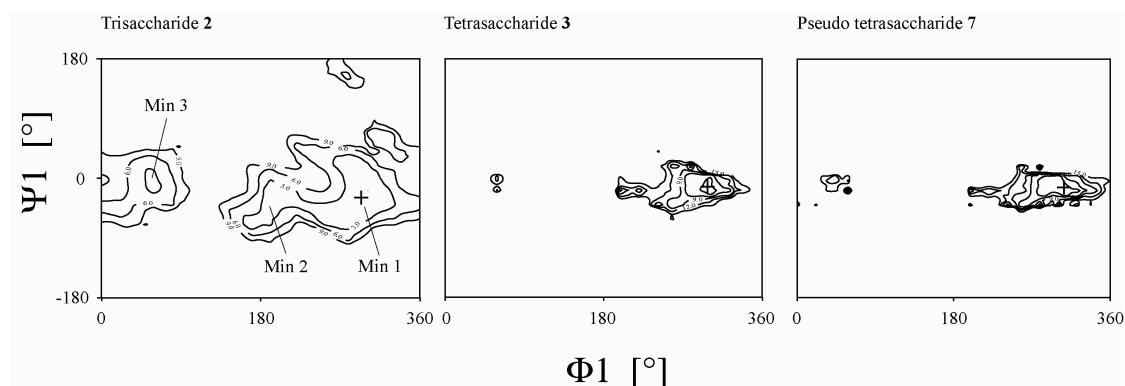


Figure 4.30: Relaxed energy maps for the glycosidic linkage α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal. The torsional angles are defined as: ϕ_1 , C1'-C2'-O2-C3 (Neu5Ac-Gal); ψ_1 , C2'-O2'-C3-H3 (Neu5Ac-Gal).

It is clear from the molecular mechanics calculations that the α -(2 \rightarrow 3)-glycosidic linkage is not mostly populated with its global minimum conformation, Min 1 in aqueous solution. But a local minimum, Min 2, is the conformation, which is mostly populated in solution. Hence, it is essential to find the conformational space accessible for the glycosidic linkages. MD simulation, either by itself or in conjunction with NMR spectroscopy, is the technique of choice for such a study. MD simulation were carried out for the tetrasaccharide **3** and pseudo tetrasaccharide **7** with explicit water over a period of 1 ns and the resulting structures were stored in every 1 ps, which led to a total of 1000 structures. In the trajectory, transitions were observed only between Min 1 and Min 2 (Figure 4.31) more frequently at the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage.

Table 4.7: Relative energy (in kcal mol⁻¹) and dihedral angles (in degrees) at the glycosidic linkages of (A) trisaccharide **2**, (B) tetrasaccharide **3** and (C) pseudo tetrasaccharide **7**. The dihedral angles were defined as: $\phi 1$, C1'''-C2'''-O2'''-C3'' (Neu5Ac-Gal); $\psi 1$, C2'''-O2'''-C3''-H3'' (Neu5Ac-Gal); $\phi 2$, H1''-C2''-O3'-C3' (Gal-Gal); $\psi 2$, C1''-O1''-C3'-H3' (Gal-Gal); $\phi 3$, C1-C2-O6'-C6' (Neu5Ac-GalNAc); $\psi 3$, C2-O2-C6'-C5' (Neu5Ac-GalNAc). The torsional angle, ω (O6'-C6'-C5'-O5') = 60° and 180° was fixed for energy calculation of tetrasaccharide **3** and pseudo tetrasaccharide **7** respectively.

(A)

Min	$\phi 1$	$\psi 1$	$\phi 2$	$\psi 2$	ΔE
1	-58	-29	50	-65	0.0
2	-156	-33	50	-65	1.9
3	64	-21	49	-65	3.6

(B)

Min	$\phi 1$	$\psi 1$	$\phi 2$	$\psi 2$	$\phi 3$	$\psi 3$	ΔE
1	-59	-29	50	-67	-71	177	0.0
2	-156	-31	49	-68	-71	175	2.0
3	-58	-30	50	-65	-167	175	2.1
4	64	-19	48	-68	-73	175	3.6
5	-156	-34	50	-68	177	166	4.8

(C)

Min	$\phi 1$	$\psi 1$	$\phi 2$	$\psi 2$	ΔE
1	-71	-18	54	-62	0.0
2	-60	-33	62	48	0.2
3	-154	-29	53	-63	1.0
4	-156	-30	62	52	1.8
5	64	-19	53	-63	2.5

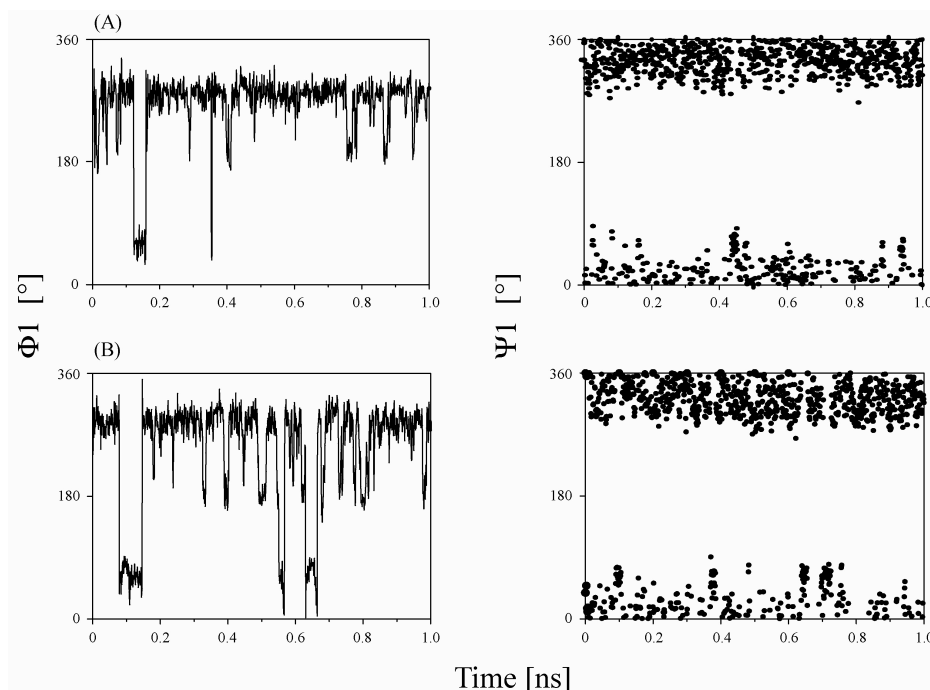


Figure 4.31: Histories of the α -(2 \rightarrow 3)-glycosidic torsions $\phi 1$, $\psi 1$ for the (A) tetrasaccharide **3** and (B) pseudo tetrasaccharide **7** fragments during 1 ns MD simulation.

The NOE data of the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage was analysed on the basis of molecular mechanics and the MD simulations. In the conformation corresponding to the global minima, Min 1, the interatomic distance between H3'-H8'' (for trisaccharide **2**) or H3''-H8''' (for tetrasaccharide **3** or pseudo tetrasaccharide **7**) is about 2.8 Å, whereas in the local minima, Min 2, the conformation shows a distance of the H3'-H3ax'' (for trisaccharide **2**) or H3''-H3ax''' (for tetrasaccharide **3** or pseudo tetrasaccharide **7**) is about 2.2 Å. On the other hand, Min 3 predicts the H3 of Gal is closer to H3eq'' of Neu5Ac moiety (for **2**). The same is true in case of tetrasaccharide **3** and pseudo tetrasaccharide **7** in Min 4 and Min 5 respectively (Figure 4.32). As the weak NOEs, H3''-H3eq''' and H3''-H8''' were observed only at higher mixing time, this excluded the appreciable contribution of Min 3 (for trisaccharide **2**) or Min 4 (for tetrasaccharide **3**) or Min 5 (pseudo tetrasaccharide **7**) and Min 1 respectively in solution (Figure 4.29). The strong cross peak between H3 of Gal and H3ax of Neu5Ac indicated the preferable conformation of Min 2 with $\phi, \psi = -160^\circ, -30^\circ$ in aqueous solution. The above conformation of the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage is in good agreement with those obtained by Poppe and co-workers (1989) and Hasselhorst and co-workers (2001) for the ganglioside GM4 and sLe^x respectively.

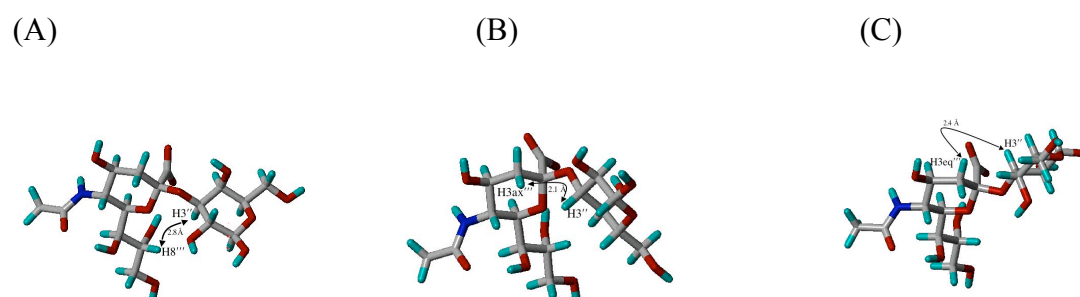


Figure 4.32: Calculated distances between (A) H3''-H8''' (Min 1) (B) H3''-H3ax''' (Min 2) and (C) H3''-H3eq''' (Min 4) of tetrasaccharide **3** from the molecular mechanics calculations.

Analysis of the relaxed energy maps of the saccharides **2** and **3** (Figure 4.33) indicated the rigidity of the linkage β -D-Gal-(1 \rightarrow 3)- β -D-Gal (or β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc) (Figure 4.33). The region of the conformational space accessed by the β -(1 \rightarrow 3)-linkage between Gal and Gal or Gal and GalNAc was very similar to that found in the literature ($\phi/\psi = 45^\circ/-65^\circ$ and $-65^\circ/-18^\circ$) (Bukowski et al., 2001; Weimar et al., 2000). The *interglycosidic* NOEs between H1'-H3 and H1'-H4 or

H1''-H3' and H1''-H4' were very consistent with the global minimum energy, Min 1, obtained from the energy maps analysis. MD simulations of tetrasaccharide **3** supported a slight flexible dynamic behaviour of this linkage. Majority of the structures belonged to a region, which displayed ϕ_2 angles consistent with the *exo*-anomeric effect ($\approx +60^\circ$) (Figure 4.33). However, the ψ_2 fluctuated from $+60^\circ$ to -60° . This result was in good agreement with the MD simulations of T-antigen (Bukowski et al., 2001; Caffarena and Bisch, 2000). As the *interglycosidic* NOEs H1''-H2' and H1''-H5' of **3** at this linkage were just above the noise level, so the conformation with $\psi_2 = 60^\circ$ were excluded. A different situation occurs for the β -D-Gal-(1 \rightarrow 3)-D-Cyc moiety of pseudo tetrasaccharide **5**. In this case the medium ROE cross peak between H1'' of Gal and H2' of Cyc was detectable along with those of H1''-H3' and H1''-H4' NOEs. This result confirms the flexibility of this linkage and that ψ varied between $+40^\circ$ to -60° .

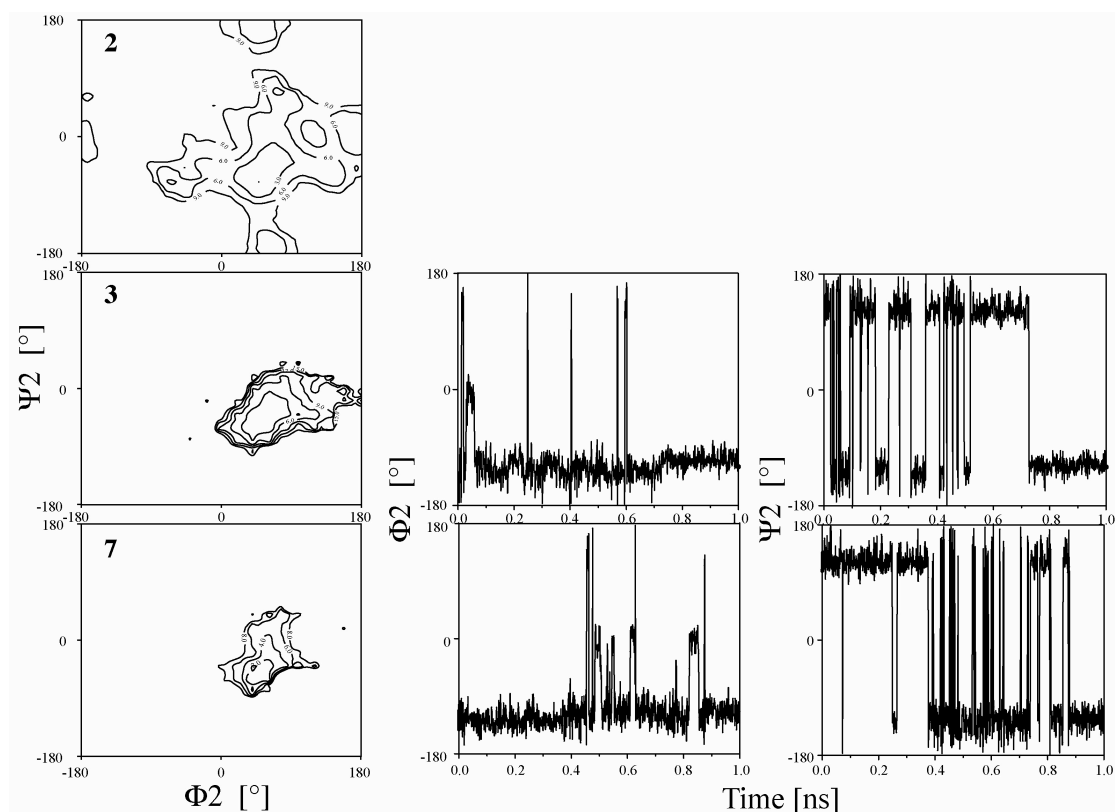


Figure 4.33: (Left) Relaxed energy maps for the glycosidic linkage β -D-Gal-(1 \rightarrow 3)- β -D-Gal (trisaccharide **2**) or β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc (tetrasaccharide **3**) or β -D-Gal-(1 \rightarrow 3)-D-Cyc (pseudo tetrasaccharide **7**). (Right) Histories of ϕ_2 and ψ_2 during 1 ns MD simulation. The torsional angles are defined as: ϕ_2 , H1'-C1'-O3-C3 (Gal-Gal); ψ_2 , C1'-O1'-C3-H3 (Gal-Gal).

Molecular mechanics followed by MD simulations of tetrasaccharide **3** reveal a number of interesting features. Figure 4.34 depicts the dynamic behaviour of the glycosidic torsions of α -D-Neu5Ac-(2 \rightarrow 6)- β -D-GalNAc over a period of 1 ns MD simulation in explicit water. From the relaxed energy maps (Figure 4.34) it was inferred that the regions were roughly similar to that found for the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage. Each of the conformations supported some of the NOEs, but it was unfortunately difficult to quantify the NOEs due to severe signal overlap.

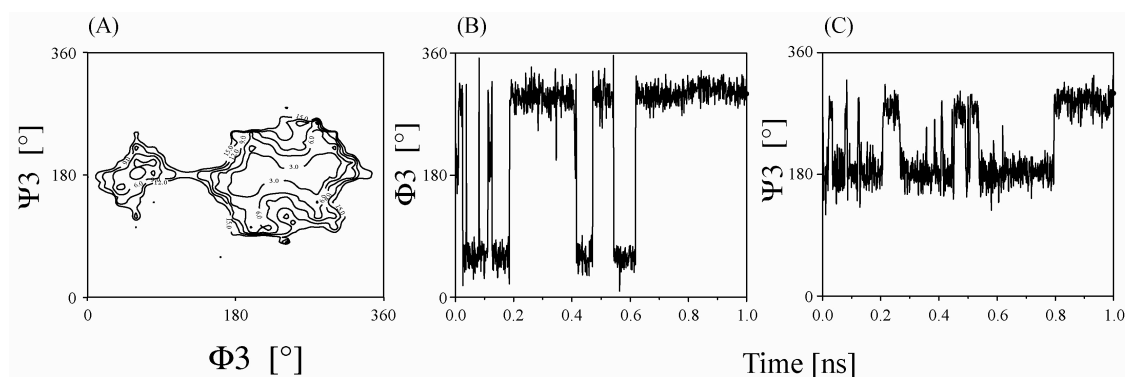


Figure 4.34: (A) Relaxed energy maps for the glycosidic linkage α -D-Neu5Ac-(2 \rightarrow 6)- β -D-GalNAc of tetrasaccharide **3**. (B) and (C) Histories of the $\phi 3$ and $\psi 3$ during 1 ns MD simulation. The torsional angles are defined as: $\phi 3$, C1'-C2'-O6-C6; $\psi 3$, C2'-O2'-C6-C5.

4.2.1.4 Discussion

The saccharides can adopt multiple conformations in aqueous solution. The conformational properties of the saccharides depend on the glycosidic linkages that are determined by molecular mechanics calculation and MD simulation. The NMR studies have shown that the higher energy conformers do co-exist in solution along with lower energy conformers. This result confirmed that oligosaccharides exist as a mixture of conformers in the aqueous solution.

Flexibility of the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage is well known (Breg et al., 1989; Sabesan et al., 1991; Scheffler et al., 1995; Poppe et al., 1997, Wu et al., 1997; Haselhorst et al., 2001). Three local minimum conformations are obtained from the molecular mechanics calculation with ϕ , ψ values of $(-60^\circ, -30^\circ)$, $(-160^\circ, -30^\circ)$ and $(60^\circ, -20^\circ)$, denoted by Min 1, Min 2 and Min 3 (Table 4.7). For trisaccharide **2** the relative energies are 0, 1.9 and 3.6 kcal mol $^{-1}$ respectively. The NOESY spectrum

of the saccharides **2-5**, obtained at 288 K or 280 K with a mixing time of 500 ms, points to many dipolar interactions within each residue (Figure 4.29). A strong dipolar interaction is observed between H3 of Gal and H3ax of Neu5Ac and at the same time two weak interactions are observed between H3''-H3eq''' and H3''-H8''' (Figure 4.29). This experimental observation clearly indicates that the higher energy conformer Min 2 (for trisaccharide **2** and tetrasaccharide **3**) or Min 3 (for pseudo tetrasaccharide **7**) are predominant in the solution. The NOE result of sLe^x at α -(2 \rightarrow 3)-linkage is in good agreement with the results obtained for the saccharides **2-5** (Haselhorst et al., 2001). From the MMC simulation (Haselhorst et al., 2001) it was highlighted that Min 2 is ~ 1.3 kcal mol⁻¹ higher in energy than the global minimum conformer, Min A was predominant in solution.

It is known that the β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc linkage is rigid (Bukowski et al., 2001; Weimer et al., 2000; Caffarena and Bisch, 2000). MD simulations were performed to develop models to interpret the NMR results. This experimental results indicates that the conformation of this linkage exists at the global minimum energy, Min 1 with ϕ , $\psi = 50^\circ$, -65° . But the β -D-Gal-(1 \rightarrow 3)-D-Cyc linkage of pseudo tetrasaccharide **5** is quite flexible than the central disaccharide β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc linkage of **3**. The medium size ROE is observed between H1''-H2' (Table 4.6) and it is already stated that the H2ax' and H2eq' display the same chemical shift. At the same time the NOEs between H1''-H3' and between H1''-H4' were also observed in the NOESY spectrum (Figure 4.29) of pseudo tetrasaccharide **5**. This NMR results have shown that the higher energy conformers, Min 4 (H1''-H2') with ϕ , ψ of 62° and 52° do co-exist in the solution with Min 3 (H1''-H3' and H1''-H4') conformer with ϕ , ψ of 53° and -63° . Therefore, the experimental result in combination with MD simulation clearly indicates the flexibility of pseudo tetrasaccharide **5** or pseudo tetrasaccharide **7** at the β -(1 \rightarrow 3)-glycosidic linkage.

As Cumming and Carver (1987) have mentioned about the problem of conformational properties for the (1 \rightarrow 6)-linked mannosidase, conformational properties for the α -(2 \rightarrow 6)-linkage are more elusive and experimentally difficult due to the flexibility of the torsional angles (ϕ , ψ , ω). Nevertheless, the volume integral of the H4'-H6' cross peaks confirms that the H4'-H6'b is slightly higher in population than H4'-H6'a (Figure 4.29). As there are lots problems of the Tripos force field (Tripos, USA) e.g., one of them is that the Tripos force field do not consider

exo-anomeric effect, this can be one reason why the torsional angles in the different glycosidic linkages do not match perfectly with the literature (Poppe et al., 1992).

4.2.2 Bound conformations of saccharides 2-6

The ^1H NMR spectrum of the oligosaccharides **2-6** in the presence of protein, MAG at 288 K is shown in Figure 4.35 and Table 4.8 shows the molar ratios of saccharides in the presence of MAG. In this section the bioactive conformation of the saccharides **2-6** will be discussed systematically. In tetrasaccharide **4** the ^1H chemical shift of anomeric protons were near to water; therefore, the experiment was performed at 280 K to move the water signal more downfield. The influence of the protein on the structure of the ligand molecule can be monitored qualitatively by observing the chemical shift changes in the ligand e.g., upfield as well as downfield protein-induced chemical shift changes which could be seen in the spectra of the ligand in the presence of protein. In trisaccharide **2**, the H1 and H4 of Gal were shifted to the upfield from the free to the bound state (-0.00596 and -0.00874 ppm respectively) whereas the H3' and H4' of middle Gal shifted to the downfield ($+0.066$ and $+0.00516$ ppm respectively). Because of the strong signal overlap, it was difficult to quantify precisely the chemical shift changes of tetrasaccharide **3** in the presence of MAG.

Table 4.8: Molar ratios used in the NOESY investigated saccharides **2-6** in the presence of MAG.

Saccharides	Protein:ligand
Trisaccharide 2	1:18
Tetrasaccharide 3	1:20
Tetrasaccharide 4	1:20
Pseudo tetrasaccharide 5	1:22
Trisaccharide 6	1:14

To overcome the anomeric proton-overlapping problem, the GalNAc residue of **3** was modified to Gal residue in tetrasaccharide **4**. No significant difference was obtained from the relative inhibitory potential (rIP) studies of tetrasaccharides **3** and **4** (Heiko G  thje personal communication) in the presence of MAG, indicating that the terminal moiety (GalNAc or Gal) of the tetrasaccharides **3** or **4** are not responsible for binding to MAG. The anomeric protons of tetrasaccharide **4** were clearly resolved. The

H1' and H4' of Gal proton resonances of **4** were shifted to upfield (-0.00238 and -0.00198 ppm respectively) in the presence of the protein whereas the chemical shift of the remaining anomeric proton remains unchanged. Strong interaction between H3'' of Gal and MAG was also manifested by changes in line width.

Both anomeric protons of trisaccharide **6** were clearly resolved in the presence of protein. The ^1H chemical shift values of anomeric proton resonances (H1'' and H1') were shifted to upfield from the free to the bound state (-0.0039 and -0.0012 ppm respectively), and H4' of Gal was shifted to the downfield ($+0.0046$ ppm) in the presence of MAG.

The differences in the ring protons of the saccharides **2-6** were less detectable due to the complexity of the ^1H NMR spectra of the saccharides in the presence of MAG.

4.2.2.1 NOESY experiments of saccharides 2-6 in complex with MAG

All the NOESY spectrums of saccharides **2-6** were measured at 500 MHz in 10 mM phosphate buffer, pH* 7.4 (contain 150 mM NaCl) except tetrasaccharide **3**, which was measured at 700 MHz under the same buffer condition. Addition of MAG to an aqueous solution of saccharides **2-4** led to a strong negative trNOEs (Figure 4.36). TrNOE effects of trisaccharide **2** were larger in the case of MAG than in the case of sialoadhesin, reflecting its slightly higher affinity for MAG. As stated earlier the NOEs of the free trisaccharide **2** and tetrasaccharide **3** were negative, therefore, only on the basis of complete trNOE build up curves it is possible to distinguish the NOEs of the free saccharides **2** and **3** from trNOEs (Figure 4.37). In case of pseudo tetrasaccharide **5** the distinction was straightforward since NOEs of the free saccharide **5** were very close to zero. The build up curves unambiguously shows that trNOEs are observed in all cases. The trNOE patterns in the presence of MAG were clearly different from the NOE patterns observed for the free saccharides, e.g., the maximum NOE enhancements for the free tetrasaccharide **3** was at mixing time of ca. 300-400 ms, whereas for tetrasaccharide **3** bound to MAG the maximum trNOE enhancement was near to 200 ms (Figure 4.37). The relative intensities at the *interglycosidic* linkages of the saccharides **2-6** in the presence of MAG are shown in Table 4.9.

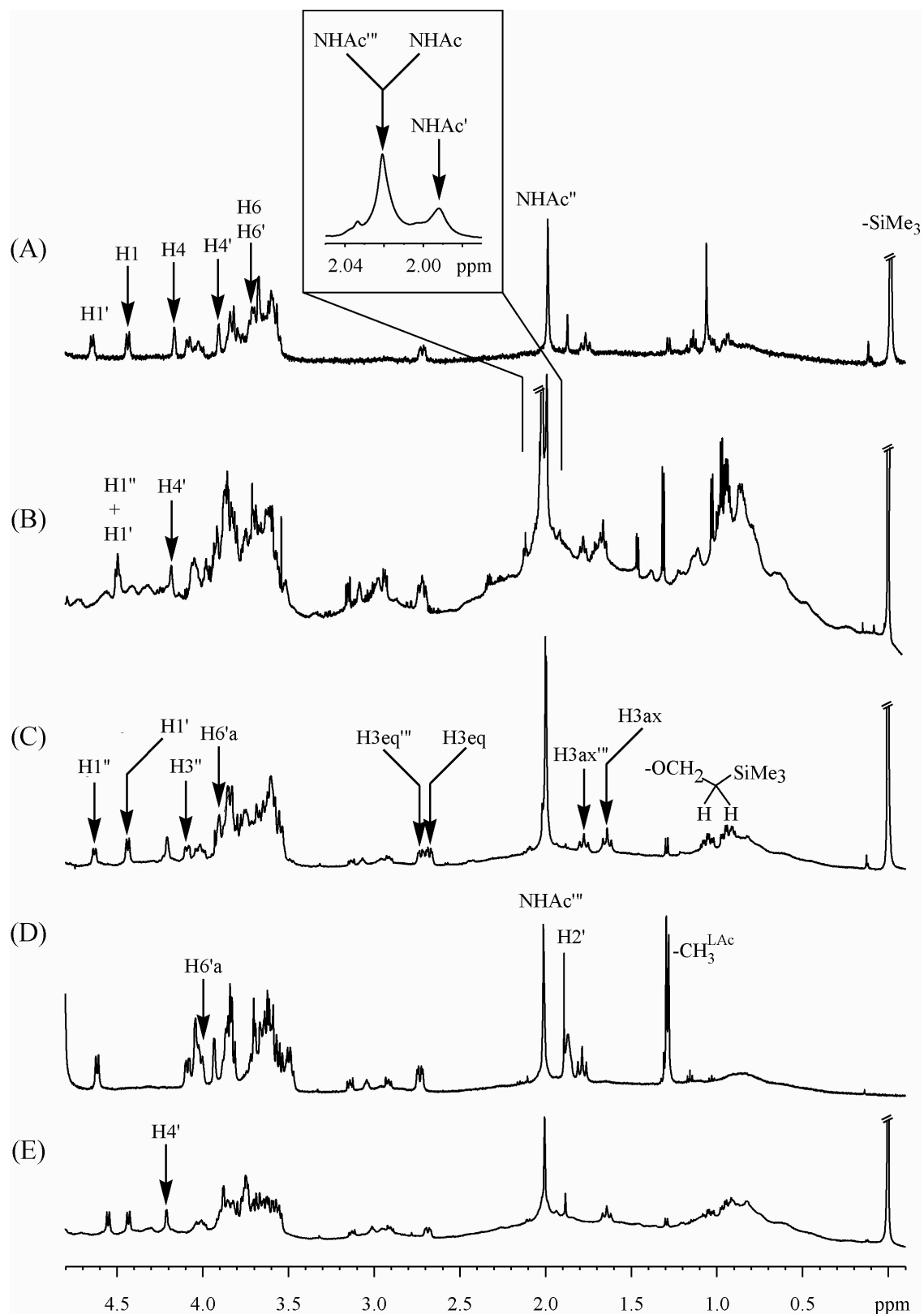


Figure 4.35: 700 and 500 MHz ^1H NMR spectrum of the synthetic saccharides 2-6 in 10 mM phosphate buffer, pH* 7.4 (with 150 mM NaCl) in the presence of MAG. The representations of the spectra are same as in Figure 4.28. Chemical shifts are referred relative to standard -OSE group at 0 ppm, which is attached to the terminal moiety.

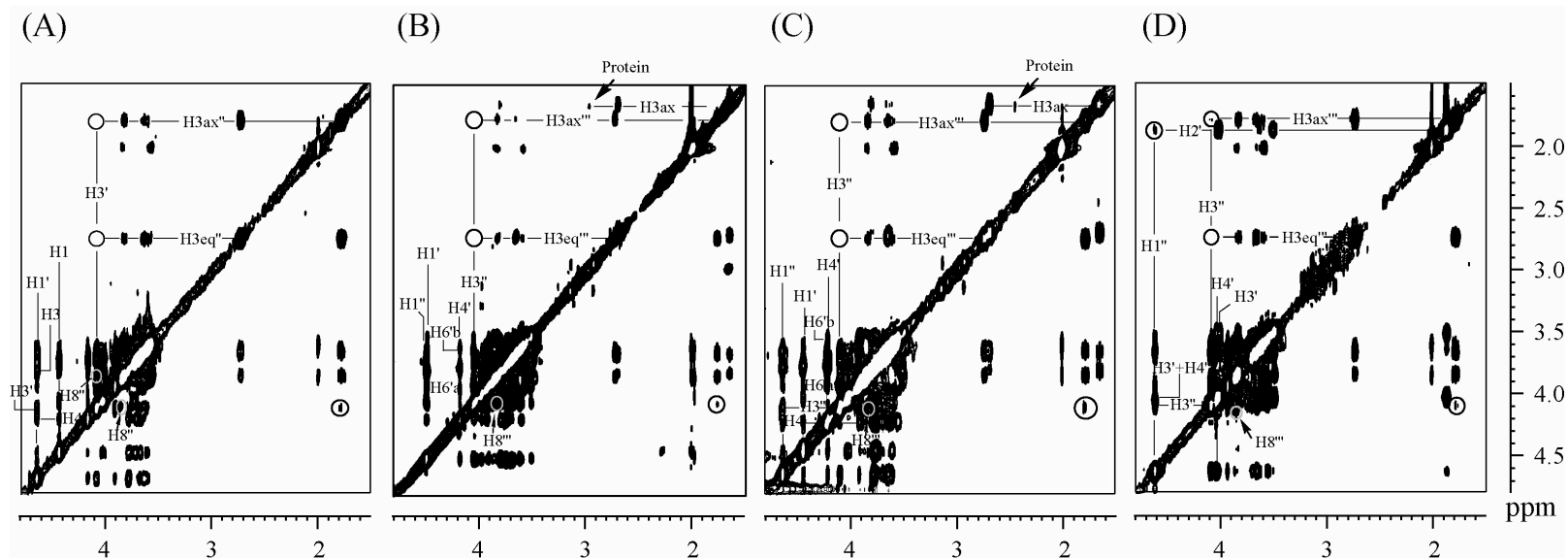


Figure 4.36: The NOESY spectra of the saccharides **2-5** in the presence of MAG (**3**: 700 MHz, 288 K; **4**: 500 MHz, 280 K; **2**, **5**: 500 MHz, 288 K). Mixing time of all the NOESY spectra were same, 200 ms except **3** (mixing time: 250 ms). The representations of the spectra are same as in Figure 4.29. The trNOE between H3' and H3ax'' or between H3'' and H3ax''' of saccharides **2-5** is nearly absent or very weak, and the trNOE between H3 of Gal and H3eq of the α -(2 \rightarrow 3)-linked Neu5Ac is completely absent. In contrast, the trNOEs between H3' and H8'' (trisaccharide **2**) or between H3'' and H8''' (saccharide **3-5**) are stronger in NOESY spectra in the presence of MAG. Open circles highlight these cross peaks.

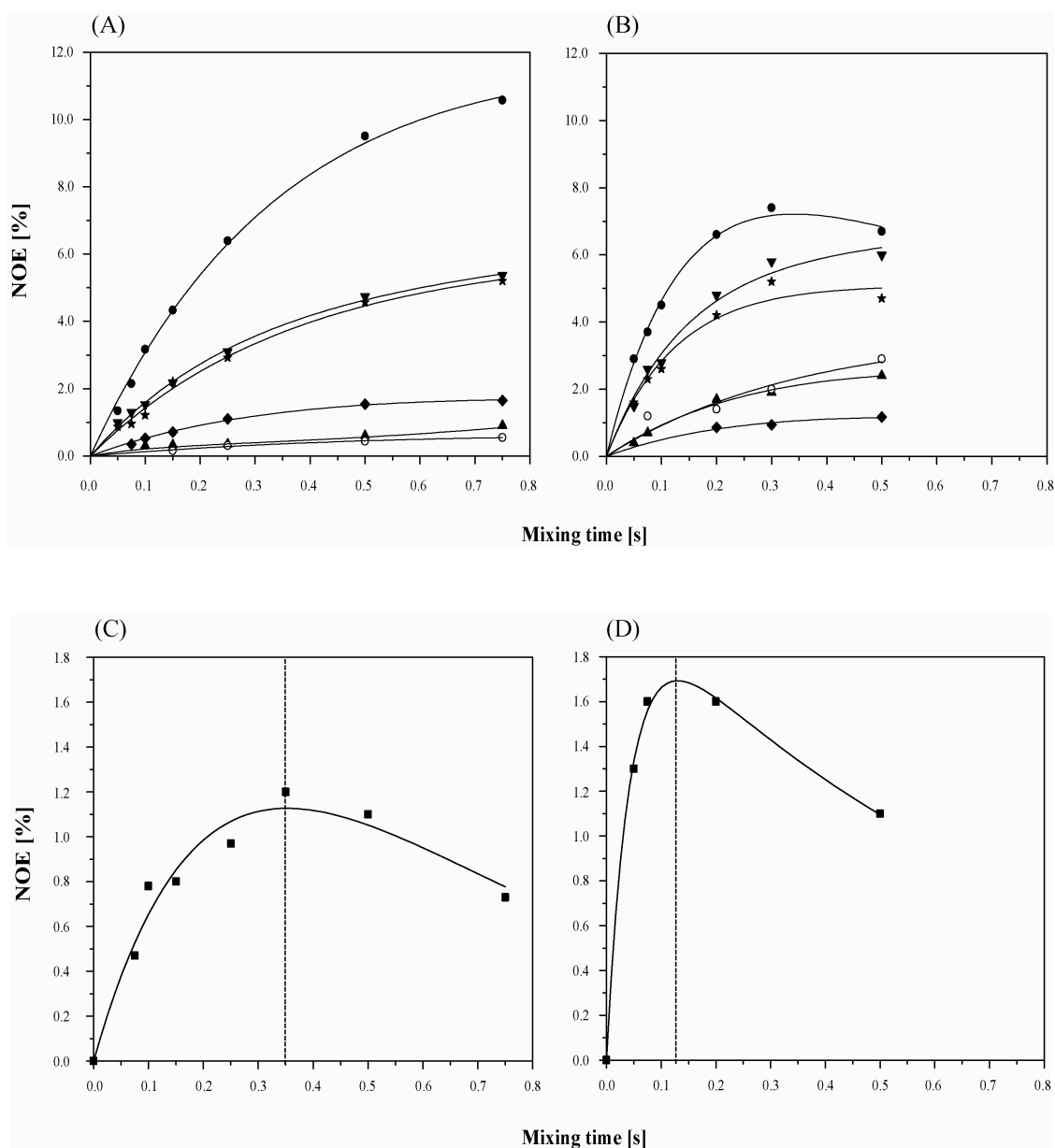


Figure 4.37: In the upper panel, build up curves of trisaccharide **2** (A) in the aqueous solution (288 K, 700 MHz) and (B) in the presence of MAG (288 K, 500 MHz) at different mixing times. The anomeric proton of Gal was close to residual water; it was partially affected by presaturation of residual water. The curves represent following trNOEs: H1-H3 (▼), H1'-H3' (★), H3'-H3ax'' (◆), H1'-H3 (●), H1'-H4 (▲) and H3'-H8'' (○). Both spectrums were processed with XWINNMR software (Bruker). In the lower panel, the build up curve of H3''-H8''' cross peak of tetrasaccharide **3** (C) in aqueous solution (700 MHz, 288 K) and (D) in the presence of MAG (700 MHz, 288 K) at different mixing times. It is clearly seen that the maximum trNOE of H3''-H8''' of tetrasaccharide **3** bound to MAG is significantly shifted towards the lower mixing times as compared to the corresponding NOE curve. Instead of that the maximum intensity of H3''-H8''' trNOE is larger than its corresponding NOE. The curves have been normalised using the decay of the corresponding diagonal signal as a reference.

Table 4.9: Interglycosidic trNOEs of the saccharides **2-6** in 10 mM Phosphate buffer, pH* 7.4 (with 150 mM NaCl). Relative NOE intensities are defined as: strong (s), medium (m), weak (w) and not observable (n.o.). (LAc: (S)-Lactic acid residue).

Saccharides	Interglycosidic NOEs		Relative intensity
2 (500 MHz, 288 K)	H3'	H3ax''	w
		H3eq''	n.o.
		H8''	s
	H1'	H3	s
		H4	m-w
3 and 4 (3 : 700 MHz, 288 K and 4 : 500 MHz, 280 K)	H3''	H3ax'''	m
		H3eq'''	w
		H8'''	s
	H1''	H3'	s
		H4'	m-w
	H6'	H3ax	w
		H4'	s-m
5 (500 MHz, 288 K)		H6''	m
	H3''	H3ax'''	w
		H3eq'''	n.o.
		H8'''	s
	H1''	(H3'+H4')	s
		(H2ax'+H2eq')	m
	H6'	-CH ^{LAc}	m

Interglycosidic trNOEs are important as they provide the information about the conformation in the bound state (Figure 4.36). Hence, the trNOEs across the α -(2 \rightarrow 3)-glycosidic linkages were most interesting since at this linkage upon binding to MAG the selection of one bioactive conformation was expected. In all cases a strong interglycosidic trNOE, H3'-H8'' or H3''-H8''' and a negligibly small H3'-H3ax'' or H3''-H3ax''' trNOE in the presence of MAG was found (Figure 4.38). At a glance, Figure 4.39 showed no difference in trNOE between H3'' of Gal and H3ax''' for tetrasaccharide **3**. But a close inspection utilising NOE/trNOE build up curve (Figure 4.40) revealed that the cross peak between H3'' of Gal and H3ax''' of Neu5Ac was solely due to NOE and not due to trNOE. The maximum trNOE enhancement between H3''-H8''' was observed at lower mixing time of ca. 200 ms (Figure 4.37) whereas for H3''-H3ax''' cross peak the maximum enhancement was observed at 300 ms mixing time in the absence as well as in the presence of MAG (Figure 4.40). Not only that but also the relative intensity of the trNOE between H3'' of Gal and H3ax''' of Neu5Ac is much smaller than the intraglycosidic trNOE H3ax'''-H5''' whereas in aqueous

solution of **3** both NOEs $H3''$ - $H3ax'''$ and $H3ax'''$ - $H5'''$ were of similar in intensity (Figure 4.41). So, the saccharides **2-5** prefer “syn” conformation with $\phi_1, \psi_1 = -60^\circ, -20^\circ$ in the presence of MAG. This result indicates that the α -(2 \rightarrow 3)-glycosidic linkage undergoes a major conformational change upon binding to protein (Figure 4.42). This result was not surprising. The same effect had been observed by Scheffler and co-workers (1997) for the binding of sLe^x to E-selectin.

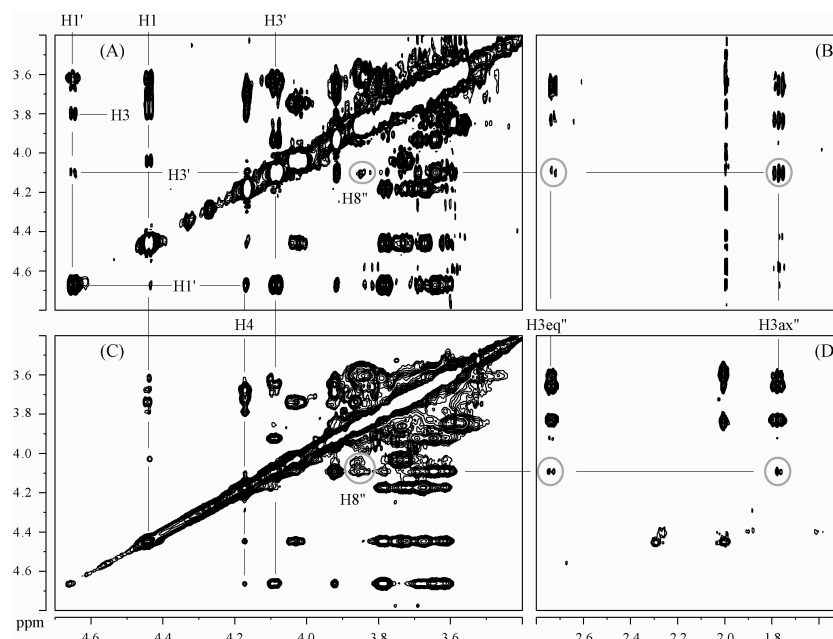


Figure 4.38: (A and B), selected parts of NOESY spectrum (700 MHz, 288 K, 500 ms) of free trisaccharide **2** in 10 mM phosphate buffer, pH* 7.4 (with 150 mM NaCl). (C and D), the corresponding cross peaks of the trisaccharide **2** in the presence of MAG (700 MHz, 288 K, 250 ms). The strong cross peak between $H3'$ - $H3ax''$ of free trisaccharide **2** is weak in the presence of MAG while, the weak $H3'$ - $H8''$ NOE is stronger in trNOESY spectrum.

On the other hand, the trNOEs observed across the β -(1 \rightarrow 3)-glycosidic linkage between the central Gal-GalNAc (tetrasaccharide **3**) or Gal-Gal (**2** and **4**) were approximately of same strength as NOEs for the free saccharides **2-4**. This result indicates that the β -(1 \rightarrow 3)-glycosidic linkage is less affected upon binding to the protein (Figure 4.36). This experimental result again confirmed the rigidity of the linkage and its existence in the global minimum conformer, Min 1 for free trisaccharide **2** and also in the presence of the protein. As it was stated earlier that the anomeric protons of tetrasaccharide **3** were overlapped i.e., $H1''$ coincided with $H1'$ (Figure 4.28) so it was difficult to quantify the trNOEs at the β -D-Gal-(1 \rightarrow 3)- β -D-GalNAc linkage. But the NOESY spectrum of **3** in the presence of MAG revealed the

following trNOEs at this linkage: H1''-H3', H1''-H4', H6''-H6'. Three major *interglycosidic* trNOEs, namely H1''-H3', H1''-H4' and H6''-H4' were observed (Figure 4.36 and Figure 4.42) at the β -D-Gal-(1 \rightarrow 3)- β -D-Gal of tetrasaccharide **4** in the presence of MAG. A weak *intraglycosidic* trNOE was also observed between H1-H4 of trisaccharide **2** or H1'-H4' of tetrasaccharides **3** and **4**.

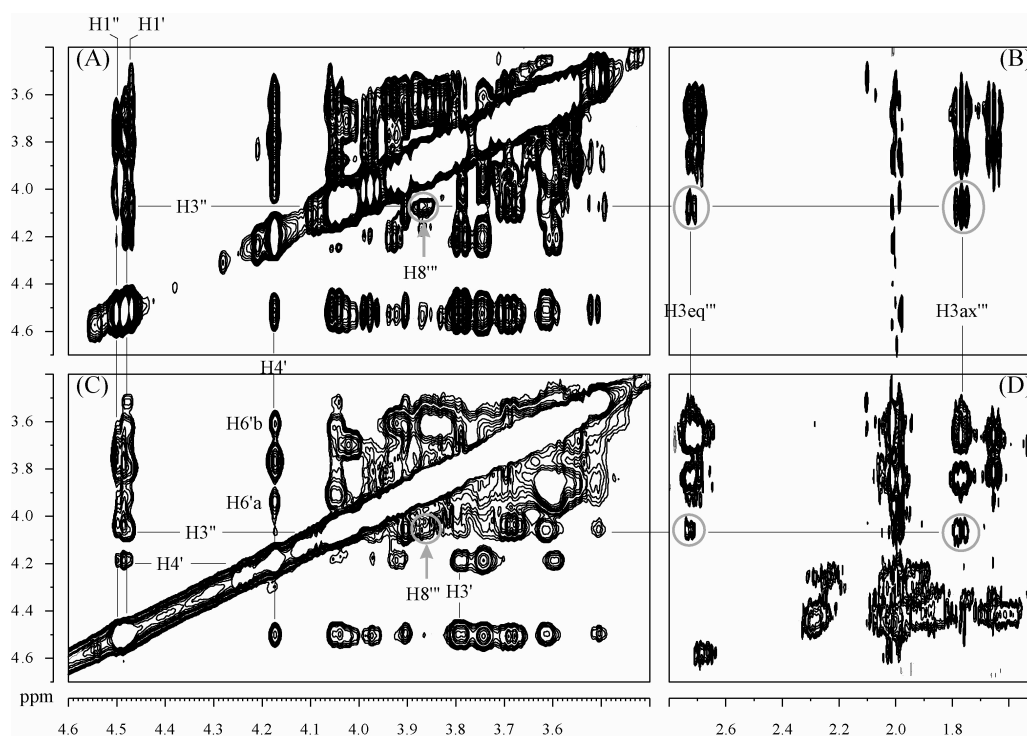


Figure 4.39: (A and B) The NOESY spectrum of tetrasaccharide **3** in aqueous solution (700 MHz, 288 K, 500 ms), (C and D) the corresponding NOESY spectrum of **3** in the presence of MAG (700 MHz, 288 K, 250 ms). The *interglycosidic* NOEs and trNOEs that define the conformation at the α -(2 \rightarrow 3)-glycosidic linkage are indicated. It is seen that trNOEs are observed between H3'' and H8''' as well as between H3'' and H3ax'''. The spectra also show that the relative size of the H3''-H3ax''' trNOE is much smaller than the corresponding NOE.

In the pseudo tetrasaccharide **5**, the cyclic ether (Cyc) mimics the GalNAc residue. The NOESY spectrum (Figure 4.36) of the pseudo trisaccharide **5** in the presence of MAG show the following *interglycosidic* trNOEs: H1''-H3', H1''-H4' and H1''-H2'. Interestingly, the *interglycosidic* trNOE between H1'' and H2' had gained significant intensity. It was noticed that the H3' and H4' of Cyc were nearly coinciding and H2ax' and H2eq' displayed identical chemical shifts, and hence they together formed a strongly coupled spin system. This makes it difficult to distinguish the corresponding NOEs, and impedes a detailed conformational analysis of this linkage.

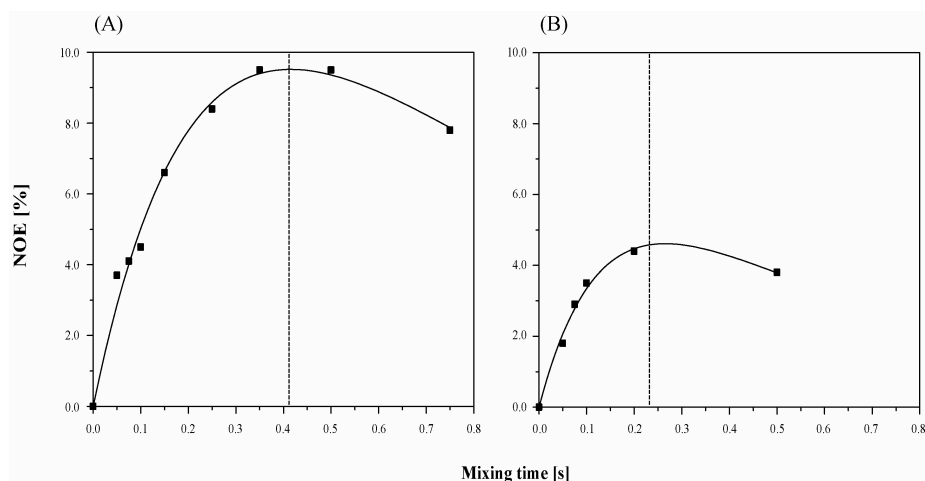


Figure 4.40: The build up curve of H3''-H3ax''' cross peak of tetrasaccharide **3** (A) in aqueous solution and (B) in the presence of MAG (700 MHz, 288 K) at different mixing times. The maximum trNOE of H3''-H3ax''' of tetrasaccharide **3** is only slightly displaced compared to the corresponding NOE. This clearly indicates that in the bound state the distance between H3''-H3ax''' is relatively larger than the H3''-H8'''. The curves have been normalised using the decay of the corresponding diagonal signal as a reference.

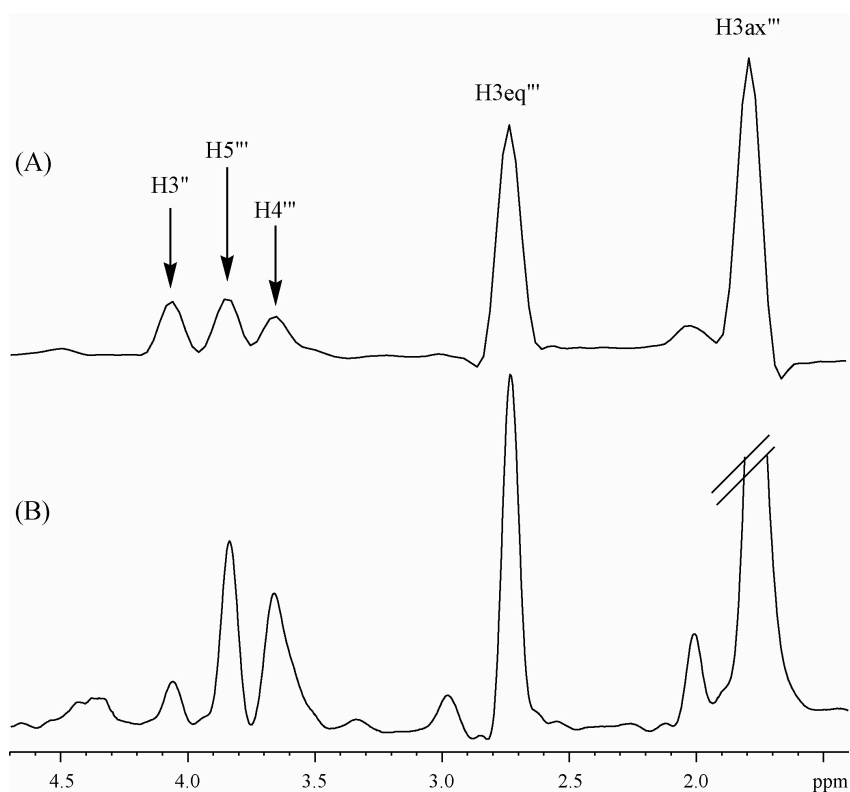


Figure 4.41: Selected cross section from the NOESY spectrum (cf. Figure 4.29 and Figure 4.36) of **3**. In the aqueous solution of tetrasaccharide **3** the H3''-H3ax''' cross peak was similar in intensity as H5''-H3ax''' but in the presence of MAG the intensity of H3''-H3ax''' was much less than H5''-H3ax'''. This result clearly shows that the distance between H3''-H3ax''' is small in aqueous solution whereas in the presence of MAG the distance is relatively larger.

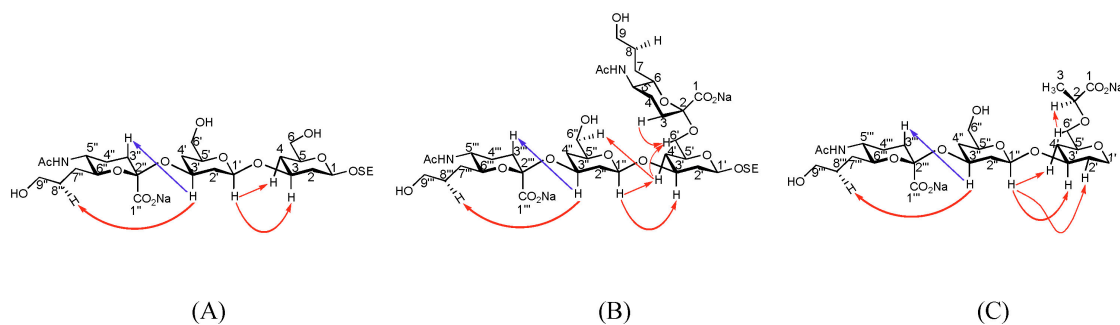


Figure 4.42: Schematic representation of the NOEs of (A) trisaccharide **2** (B) tetrasaccharide **4** and (C) pseudo tetrasaccharide **5** in the aqueous solution (blue arrow) and in the presence of the protein (red arrow). For clarity, all the functional groups except important protons are removed.

Though the α -D-Neu5Ac-(2 \rightarrow 6)- β -D-GalNAc (for tetrasaccharide **3**) or α -D-Neu5Ac-(2 \rightarrow 6)- β -D-Gal (for tetrasaccharide **4**) linkage was very flexible in the aqueous solution yet the fixed bound conformation of this linkage was expected in the presence of MAG. Due to severe signal overlap the analysis of this linkage was extremely difficult. A very weak but important cross peak between H1' of Gal and H8 of Neu5Ac at the α -(2 \rightarrow 6)-glycosidic linkage was observed. Unfortunately, this cross peak was difficult to integrate. Absence of this cross peak in the ROESY experiment of **4** in the presence of MAG indicated that this trNOE could be due to spin diffusion by protein protons. Further conformational properties of this linkage were investigated using selective 1D NOESY experiment. As H3ax of α -(2 \rightarrow 6)-linked Neu5Ac was clearly resolved so it was saturated during the mixing time with low power pulse resonance frequency. The weak signals were observed from the H6'a and H6'b of GalNAc or Gal moiety along with the *intraglycosidic* trNOEs (Figure 4.43). Often only trNOEs involving the proton resonances of the ligand molecule were observed; however, while investigating the complex, the *intermolecular* trNOE effects between protein and ligand were also observed. The same was observed in the 2D NOESY spectrum of **3** and **4** in the presence of MAG (Figure 4.36). Therefore, Some control experiments (Figure 4.44 and Figure 4.45) were carried out to have an idea about this *intermolecular* trNOEs between ligand and protein. In one of the control experiments (Figure 4.44), STD experiment was performed by saturating the protein at aliphatic region (2.25 ppm) to know the epitope of the ligand, **3**. The *intermolecular* trNOE effect between protein and ligand was also investigated using selective 1D NOESY experiment. The protein signal was saturated at 2.25 ppm during the mixing time with low power pulse. The STD and selective 1D NOESY experiments of **3** in the presence

of MAG point out that the H3ax proton of α -(2 \rightarrow 6)-linked Neu5Ac is in close proximity to the protein protons. Another control experiment (Figure 4.45) was performed to verify the *intramolecular* protein protons. The same protein signal was saturated at 2.25 ppm in the absence of ligand during the mixing time with low power pulse. All the above results clearly reveal the existence of short distances between protein and H3ax of α -(2 \rightarrow 6)-linked Neu5Ac. At this time it is not the aim of the work to determine which amino acid is near to H3ax of α -(2 \rightarrow 6)-linked Neu5Ac by NMR experiment.

The α -(2 \rightarrow 6)-glycosidic linkage between (S)-LAc and Cyc residue of pseudo tetrasaccharide **5** showed two *interglycosidic* trNOEs, H6'-CH^{LAc} and H6''-CH^{LAc} in the presence of MAG (Figure 4.36). Additionally, a very weak *interglycosidic* trNOE between H6' of Cyc and methyl group of (S)-LAc was also observed only at 500 ms mixing time.

The trNOEs of trisaccharide **6** were of lower absolute intensity than the corresponding oligosaccharides **2-5** binding to MAG (Figure 4.46). Therefore, it was very difficult to distinguish between the contribution from trNOEs and NOEs. This result indicates that the binding of trisaccharide **6** to MAG is of lower affinity than that of oligosaccharides **2-5**. Therefore, no attempts were made to further analyze the bound conformation of **6**.

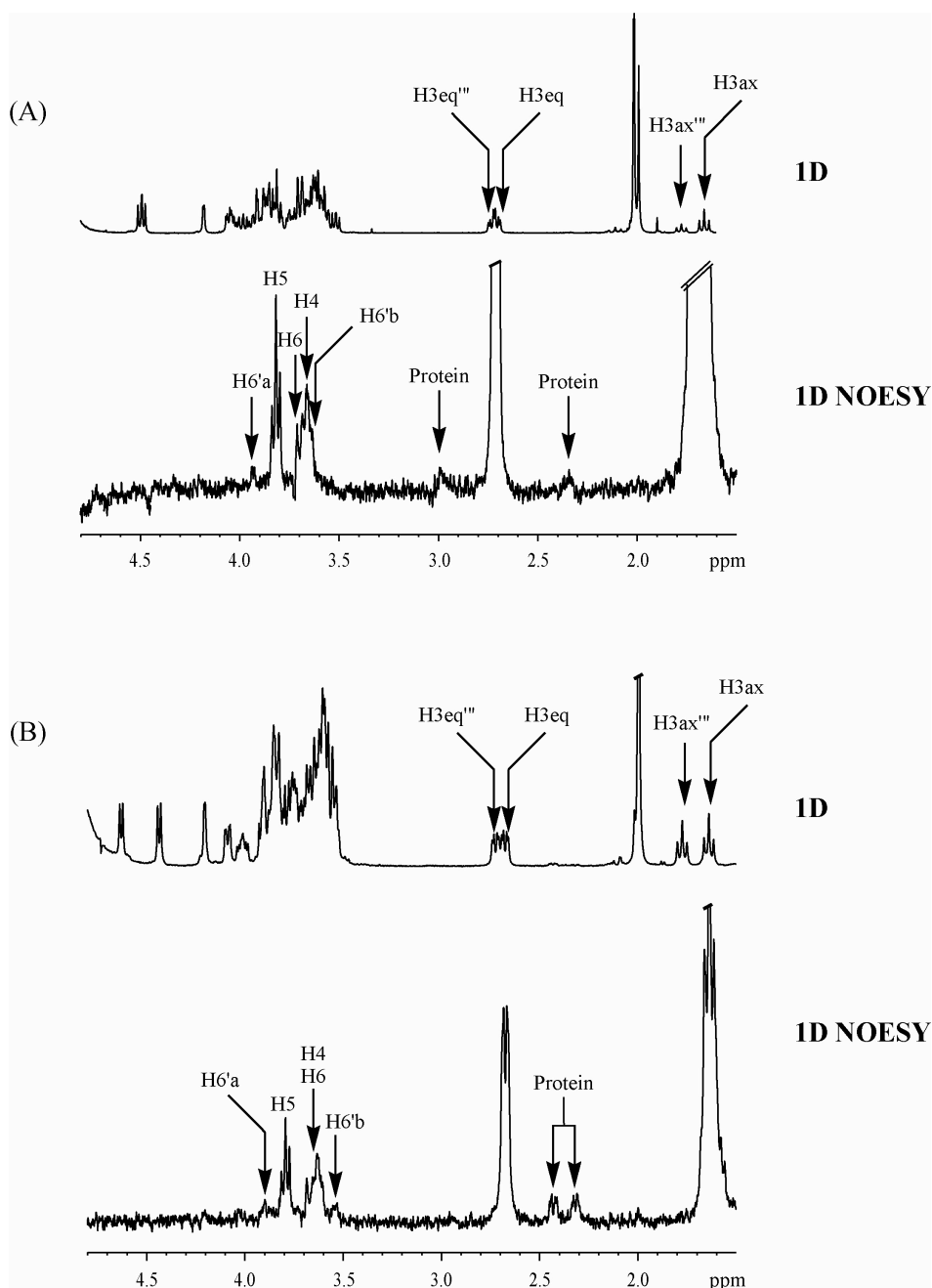


Figure 4.43: ^1H NMR and 1D NOESY spectra of **3** and **4** in the absence and presence of MAG respectively. The upper two projections (A) show the ^1H and selective 1D NOESY of tetrasaccharide **3** in the presence of MAG (molar ratio 1:20) respectively (288 K, 500 MHz, 300ms). For the selective 1D NOESY spectrum of the complex, the relaxation delay was 2 s and the number of scans were 24 K. The lower two projections (B) refer to tetrasaccharide **4**. Here, the mixing time was 300 ms, relaxation delay was 2 s and numbers of scans were 18 K. The selective 1D NOESY was acquired for both tetrasaccharides in the presence of MAG with a selective transfer from H3ax to H6'a and H6'b. Both the selective 1D NOESY spectra also show the *intermolecular* trNOEs between ligand and protein as it was observed in the 2D NOESY spectrum.

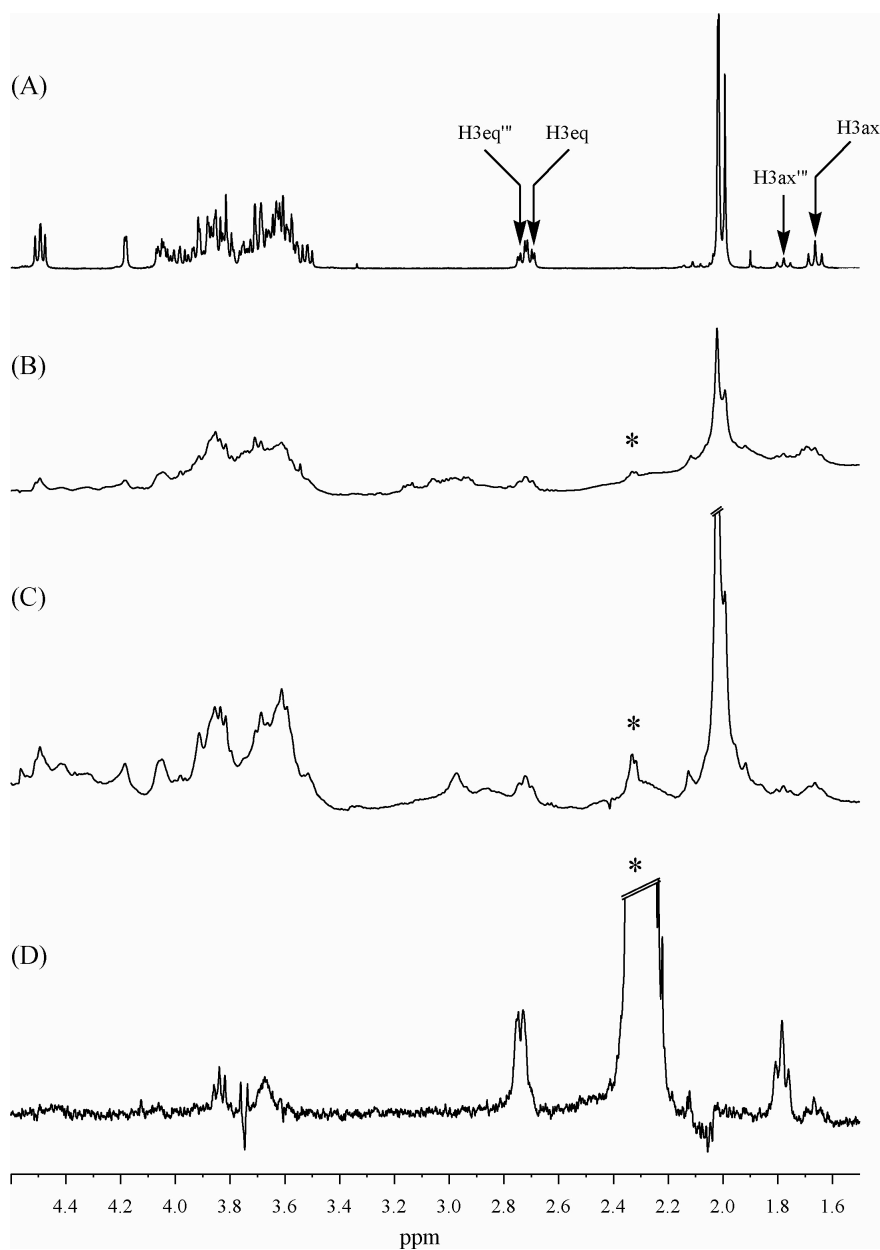


Figure 4.44: Control experiment: 1. (A) ^1H NMR of free tetrasaccharide **3** in aqueous solution and (B) in the presence of MAG (molecular ratio 1:20) at 288 K and 500 MHz. The same condition was applied for the next two experiments. (C) STD NMR spectra of the same sample. For the STD NMR spectrum the protein was saturated with 40 selective Gaussian-shaped pulses (50 ms each) resulting in a total saturation time of ~ 2 s (on resonance: 2.25 ppm and off resonance: 40 ppm). (D) Selective 1D NOESY spectrum of sample (B) (14 K scans). Mixing time was 300 ms. For 1D selective NOESY experiment, the protein was saturated at 2.25 ppm, which is shown by asterisk (*).

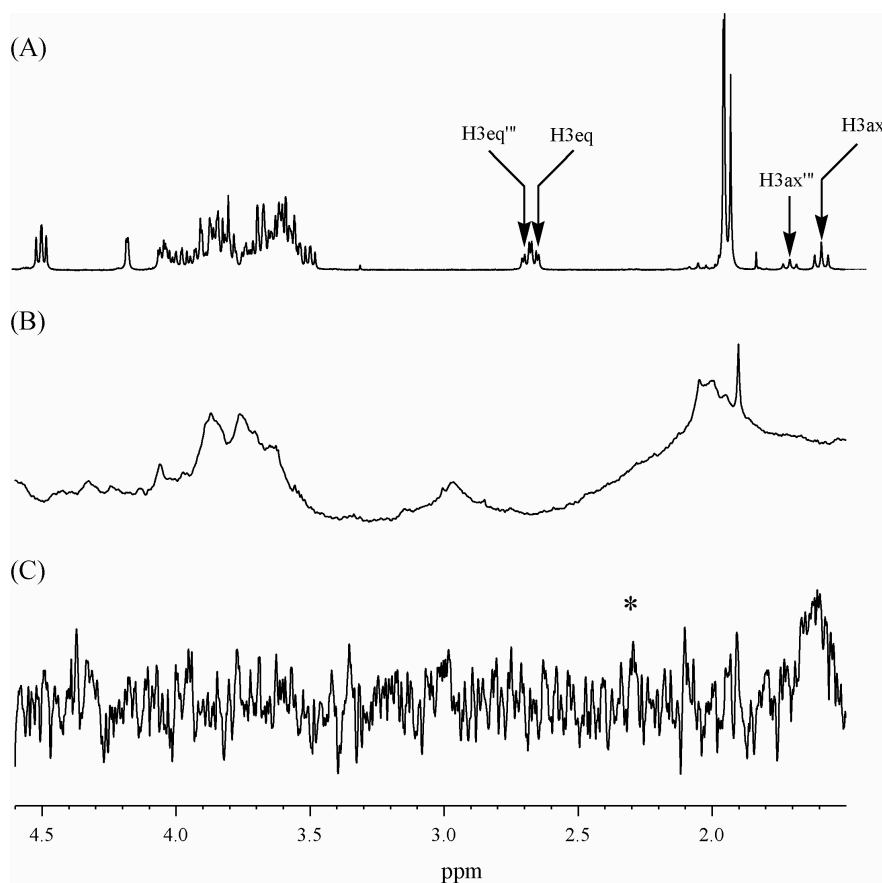


Figure 4.45: Control experiment: 2. (A) ^1H NMR spectra of free tetrasaccharide **3** in 10 mM phosphate buffer, pH* 7.4 (with 150 mM NaCl) at 288 K and 500 MHz. (B) ^1H NMR and (C) selective 1D NOESY spectrum of MAG alone (288 K, 500 MHz, 300 ms). Protein was saturated at 2.25 ppm (*).

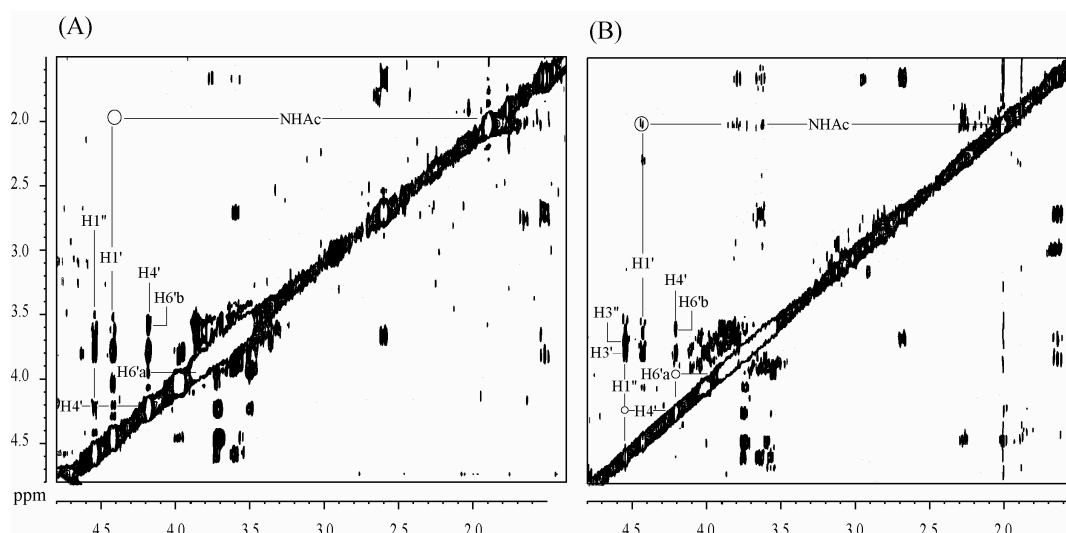
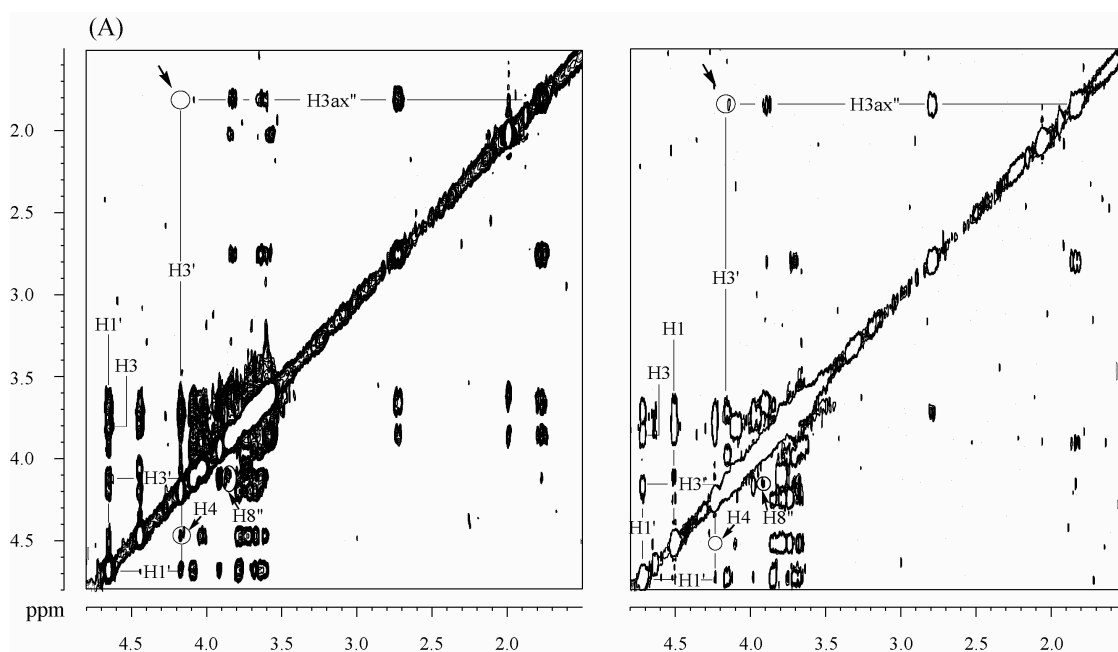


Figure 4.46: NOESY spectrum of (A) trisaccharide **6** (500 MHz, 288 K, 500 ms) in aqueous solution and (B) in the presence of MAG (Molar ratio 1:14) (500 MHz, 288 K, 200 ms). The cross peak intensities in the presence of MAG are very weak, indicating that binding affinity is weaker.

4.2.2.2 ROESY experiments of saccharides 2-6 in complex with MAG

Spin diffusion by protein protons plays an important role for false distance constraints while generating a bioactive conformation of a ligand (Poveda et al., 1997; Casset et al., 1996; Scheffler et al., 1997; Haselhorst et al., 1999; Maaheimo et al., 2000). Therefore, if a cross peak is observed in the trNOE analysis spectrum and no cross peak in the trROE analysis spectrum then it can be postulated that the cross peak in the trNOE analysis spectrum originated from spin diffusion. The prominent trNOE between H3 of Gal and H8 of Neu5Ac was present in all ROESY spectra (Figure 4.47), confirming that it was not generated by spin diffusion via protein protons. The trNOE between H1 and H4 of **2** and H1''-H4'' of **4** was absent in ROESY spectra, and therefore, these trNOEs were due to spin diffusion from *intraresidual* protons. The negative cross peaks H6''-CH^{LAc} and H6'-CH₃^{LAc} were also observed in the ROESY spectrum of **5** in the presence of MAG, indicating that these trNOEs were contaminated via spin diffusion from protein protons.



(Figure continued)

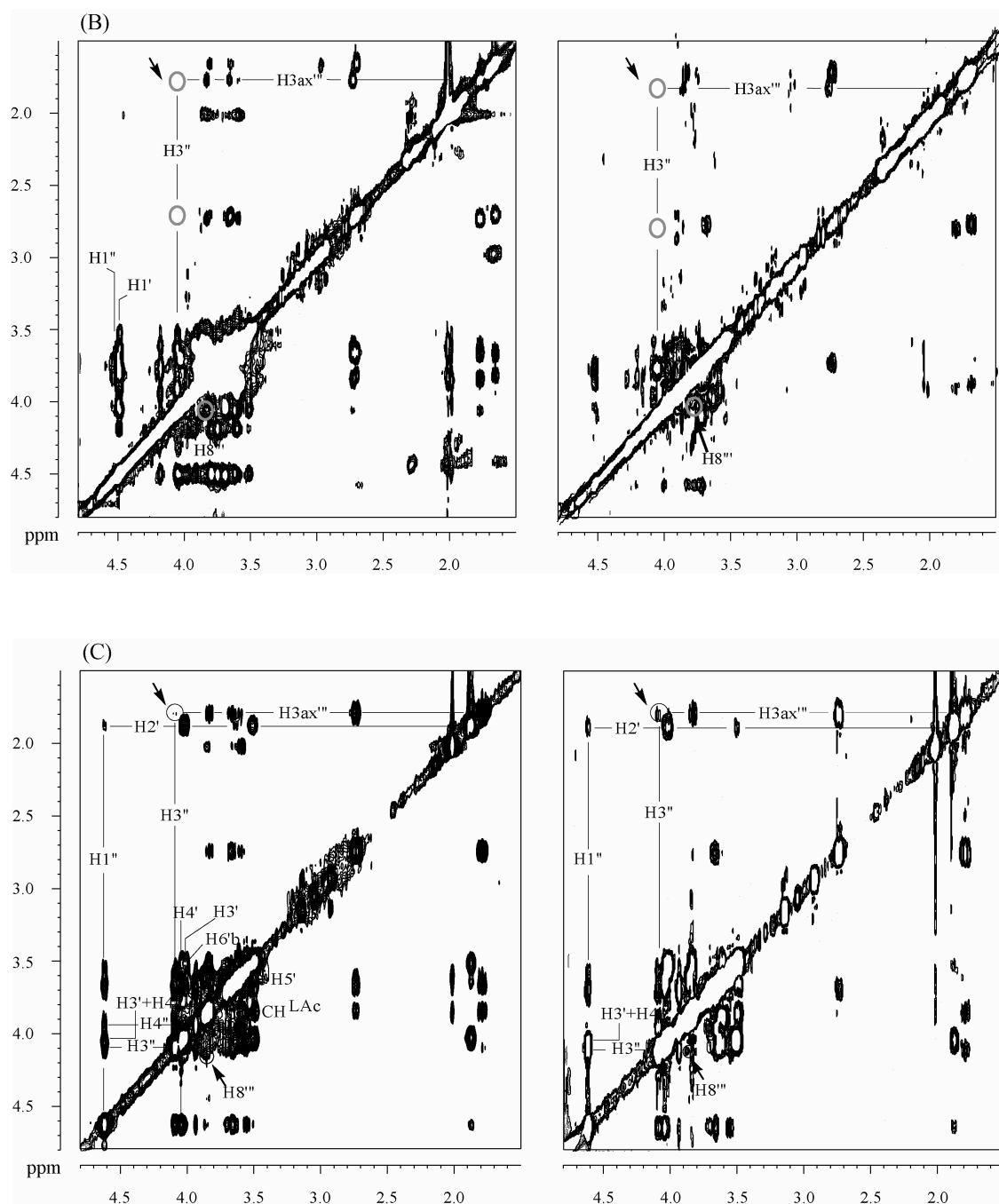


Figure 4.47: In the presence of MAG, the NOESY (288 K) (left) and ROESY (288 K) (right) spectrum of (A) trisaccharide **2** and (B) tetrasaccharide **3** and (C) pseudo tetrasaccharide **5**. In case of **2** and **5**, NOESY and ROESY spectra were measured at 500 MHz whereas for tetrasaccharide **3**, NOESY and ROESY spectra were acquired at 700 MHz and 500 MHz respectively. The mixing time for trisaccharide **2** and pseudo tetrasaccharide **5** was 200 ms (NOESY) and 250 ms (ROESY) whereas, for tetrasaccharide **3** the mixing time was 200 ms (NOESY) and 150 ms (ROESY). The arrow indicates the changes between NOESY and ROESY of saccharides **2**, **3** and **5** in the presence of MAG. The trNOE, H3''-H8'' (for **3** and **5**) or H3'-H8'' (for **2**) was present in both spectra so this trNOE was not generated via spin diffusion from protein protons.

4.2.2.3 AutoDock

The initial docking study of tetrasaccharide **4** and pseudo tetrasaccharide **5** into the protein was performed using AutoDock 3.0 (Morris et al., 1998). AutoDock 3.0 uses a Lamarckian genetic algorithm (LGA) to find the most favourable ligand binding orientation. The force field, which is a linear combination of van-der-Waals, hydrogen bonding, hydrophobic desolvation, electrostatic and torsional free energy represents the atomic affinity grid maps to compute the docking energy of ligand.

The co-ordinates of α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage of tetrasaccharide **4** was taken from the crystal structure of α -(2 \rightarrow 3)-sialyl lactose (acquisition code, 1QFO) bound to sialoadhesin (May et al., 1998). In the second step of ligand modelling, a monosaccharide unit e.g., β -D-Gal for **4** and D-Cyc for **5** was added to the above linkage and the resulting structure was energy minimised using Tripos force field (Tripos, USA). It was stated earlier that α -(2 \rightarrow 6)-linkage of tetrasaccharide **3** can adopt several conformation in aqueous solution. But a medium STD effect was observed from the α -(2 \rightarrow 6)-linked Neu5Ac of tetrasaccharide **3** in the presence of MAG (So-Young Shin, personal communication), indicating that the side chain of α -(2 \rightarrow 6)-linked Neu5Ac is in close contact with the protein surface. Though free conformational analysis of α -(2 \rightarrow 6)-linkage show that the ψ value is predominant at a value of 180° (Peppe et al., 1992), manual docking of tetrasaccharide **4** with Sybyl (Tripos, USA) found that the ψ dihedral angle at the α -(2 \rightarrow 6)-glycosidic linkage cannot attain a value of 180° because the “trans” orientation at this linkage points the side chain of Neu5Ac towards the solution. But the “gauche” orientation for ψ turned the side chains of α -(2 \rightarrow 6)-linked Neu5Ac close to the protein surface without effecting the α -(2 \rightarrow 3)-linked Neu5Ac in its binding site. Therefore, a model was constructed for tetrasaccharide **4** with the ϕ , ψ and ω values of $\phi_3 = -159^\circ$, $\psi_3 = 81^\circ$ and $\omega = 60^\circ$ at the α -(2 \rightarrow 6)-linkage and this conformation was then utilized for the AutoDock runs. This structure of tetrasaccharide **4** was constructed according to the qualitative trNOE data shown in Figure 4.36 and Figure 4.42). The program, AutoDock searched the potential binding pockets on the surface of MAG. The docking calculations were performed with the constrained ligand, while its orientation within the binding site was varied. 100 docking experiments were initiated for both the tetrasaccharide **4** and pseudo tetrasaccharide **5** but only cluster was obtained in each

case within a tolerance of 1 Å RMS deviation (Figure 4.48) and the cluster was represented by its docking energy of ca. -60.7 and -58.0 kcal mol⁻¹ respectively. In all cases AutoDock placed the ligands in the binding pocket in an orientation that is similar to the orientation of the sialyl lactose ligand co-crystallized with sialoadhesin (PDB: 1QFO) (May et al., 1998). Each conformation of the cluster identified the guanidine group of Arg 118 to form a salt bridge with the carboxyl group of α -(2 \rightarrow 3)-linked Neu5Ac. As it is well known from the literature (Collins et al., 1997) that Arg 118 is conserved in all the family members of siglec and recognise Neu5Ac at this site, the conformation of AutoDock cluster is in very good agreement with the literature. All conformations of the cluster also predicted the hydrophobic interaction between Trp 22 and the *N*-acetyl methyl group of α -(2 \rightarrow 3)-linked Neu5Ac. The α -(2 \rightarrow 3)-linked Gal moiety was nearer to Thr 128 as well as to Tyr 65. Additionally, AutoDock predicted the strong electrostatic interaction between carboxyl group of α -(2 \rightarrow 6)-linked Neu5Ac or (S)-LAc and Lys 67 of MAG. At the same time, the *N*-acetyl methyl group of α -(2 \rightarrow 6)-linked Neu5Ac was oriented in such a way that the favourable hydrophobic interaction with the Tyr 69 was possible. The side chain of α -(2 \rightarrow 6)-linked Neu5Ac was also close to Tyr 65, Pro 66 and Lys 67. The complexes obtained from AutoDock were energy minimised using Tripos force field (Tripos, USA) and Gasteiger Hückel charge. The energy-minimised structures were then further optimised using the program DOCK (part of the Sybyl software package) (Tripos, USA).

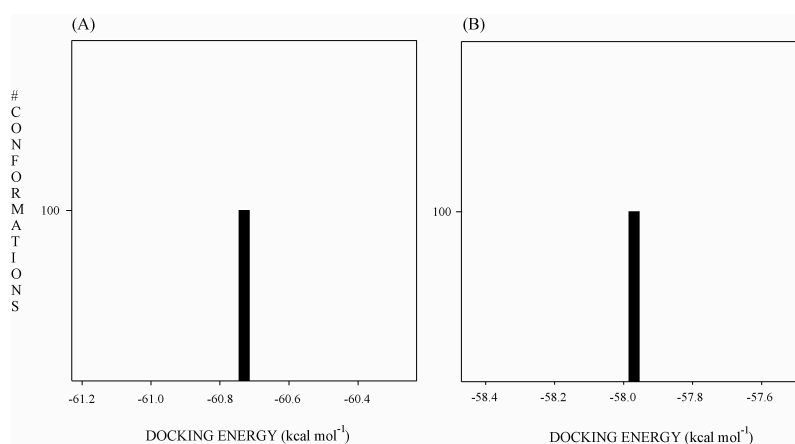


Figure 4.48: Cluster of (A) tetrasaccharide **4** and (B) pseudo tetrasaccharide **5**, obtained from the AutoDock 3.0 program. The cluster was further optimised in the binding site of MAG using Sybyl Dock (Tripos, USA) to compare the experimental and theoretical trNOE build up curves. In both cases, the carboxyl group of α -(2 \rightarrow 3)-linked Neu5Ac identify the guanidine group of Arg 118 similar to the crystal structure of sialoadhesin-sialyl lactose complex. AutoDock also predicts the strong electrostatic interaction between the carboxyl group of α -(2 \rightarrow 6)-linked Neu5Ac/(S)-LAc and Lys 67 of MAG.

4.2.2.4 CORCEMA calculation

In order to quantitatively assess the bioactive conformations of saccharides **2**, **4** and **5**, CORCEMA program (Moseley et al., 1995; Curto et al., 1996; Rinnbauer et al., 1998; Hricovini et al., 1999) was employed using full relaxation and conformational exchange matrix. Experimental and calculated trNOE build up curves are compared in Figure 4.49. Apart from the bioactive conformation of the ligand and its orientation in the binding pocket the magnitude of the trNOEs depends on a number of parameters such as the dissociation constant K_D , the off-rate constant k_{off} , and the molecular correlation times of the ligand (τ_c^{ligand}) and the protein ($\tau_c^{protein}$). For the calculation of theoretical trNOEs overall isotropic motion of the complex a correlation time of $\tau_c^{protein} = 40$ ns ($0.25 \times MW$ of the protein) was assumed. In general, free ligands, such as oligosaccharides, have correlation times in the range of $\tau_c^{ligand} = 10^{-9} - 10^{-10}$ s, whereas bound ligands adapt the correlation time of the receptor protein. The correlation times for the free and bound ligands are summarized in Table 4.10.

The theoretical and experimental trNOE data of trisaccharide **2** agreed well with the *intraglycosidic* protons of H1-H3 ($R = 0.10$), H1'-H3' ($R = 0.20$) and H3'-H4' ($R = 0.04$). Table 4.11 summarizes inter proton distances that correspond to experimental trNOEs. In all cases, the theoretical H3'-H8'' or H3''-H8''' trNOE data based on the docking model with ca. $\phi 1, \psi 1 = -60^\circ, -20^\circ$ could interpret the experimental trNOE data satisfactorily (Table 4.11 and Table 4.12). As a control experiment, the local minimum energy, Min 2 ($\phi 1 = -174^\circ, \psi 1 = -8^\circ$), which was predominant in the aqueous solution of trisaccharide **2**, was also docked into the active site of MAG and theoretical trNOEs (Figure 4.50) were calculated using the program, CORCEMA (Moseley et al., 1995; Curto et al., 1996; Rinnbauer et al., 1998; Hricovini et al., 1999). The docking model showed that the β -D-Gal-(1 \rightarrow 3)- β -D-Gal linkage pointing towards the solution (Figure 4.51). Keeping the same parameter (K_D and k_{off}) constant (Table 4.10) it was found that R-factor, which indicates the deviation from experimental and theoretical trNOE, was even high in case of *intraglycosidic* linkages (Table 4.13). The only satisfactory agreement was obtained in the case of *intraglycosidic* H1'-H3' trNOE with R-factor = 0.12. The main difference between the theoretical and experimental trNOE was between H3' and H8'' protons, on the α -(2 \rightarrow 3)-glycosidic linkage. Satisfactory agreement was also obtained for the

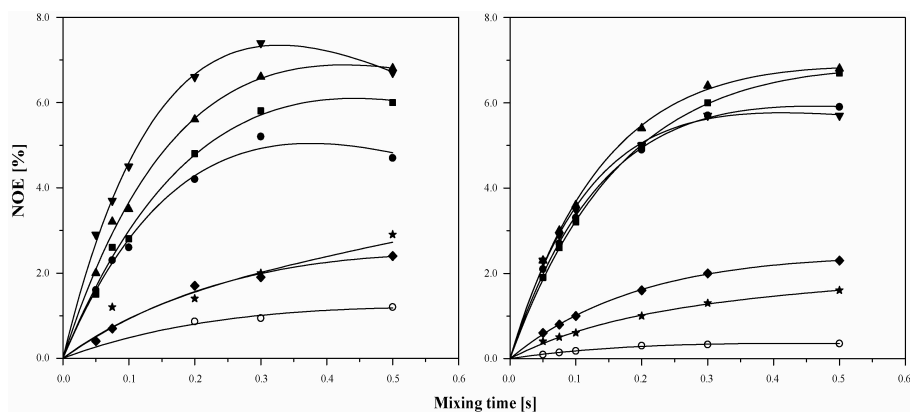
intraglycosidic trNOEs between H4'-H6'a and H4'-H6'b of tetrasaccharide **4** with R= 0.32 and 0.36 respectively. Therefore, “gt” conformation (O5'-C5'-C6'-O6') that was present in larger population (Poppe et al., 1992) in the aqueous solution of tetrasaccharide **3** could be the major conformation of tetrasaccharide **4** in the presence of protein.

It was earlier mentioned that the β -(1 \rightarrow 3)-glycosidic linkage between Gal and Cyc of pseudo tetrasaccharide **5** was flexible in the aqueous solution, it was interesting to know the conformation present in the bound state. Satisfactory agreement was obtained with the geometry of ϕ_2 , $\psi_2 = 63^\circ$, -36° where H1''-(H2ax'+H2eq') (R= 1.00) as well as H1''-(H3'+H4') (R= 0.24) cross peak was observed (Table 4.11 and 4.12). The only trNOE between H6' and CH^{LAc} at the (S)-LAc- α -(2 \rightarrow 6)-D-Cyc linkage was also in good agreement with the calculated trNOE (R= 0.56). Good correspondence between experimental and calculated *interglycosidic* trNOEs show that the bound conformation of tetrasaccharide **4** and pseudo tetrasaccharide **5** is very close to the conformation that obtained by AutoDock 3.0 model. The bioactive conformations of **2**, **4** and **5** are shown in Figure 4.52.

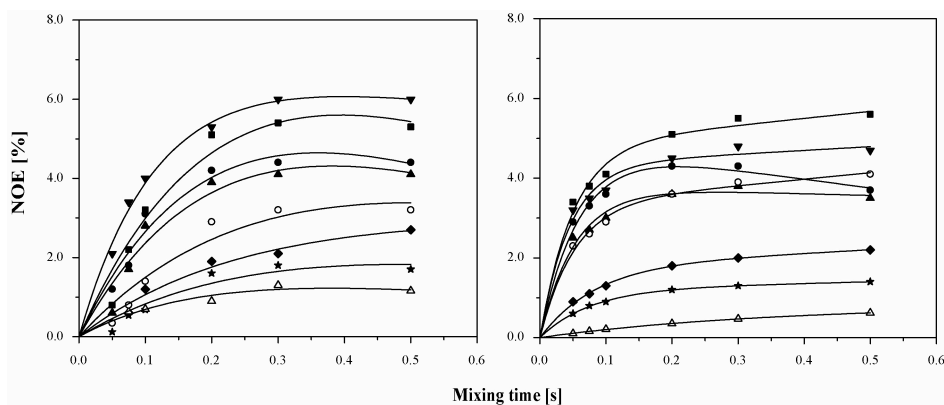
Table 4.10: The K_D , k_{off} , τ_c^{ligand} and $\tau_c^{protein}$ values of the saccharides **2**, **4** and **5** are used for the theoretical trNOE (CORCEMA) calculation.

Ligand	K_D (μM)	k_{off} (s^{-1})	k_{on} ($M^{-1} s^{-1}$)	τ_c^{ligand} (ns)	$\tau_c^{protein}$ (ns)
Trisaccharide 2	390	40	10^5	0.3	40
Tetrasaccharide 4	180	10	5.10^4	0.5	40
Pseudo tetrasaccharide 5	180	12	6.10^4	0.35	40

(A)



(B)



(C)

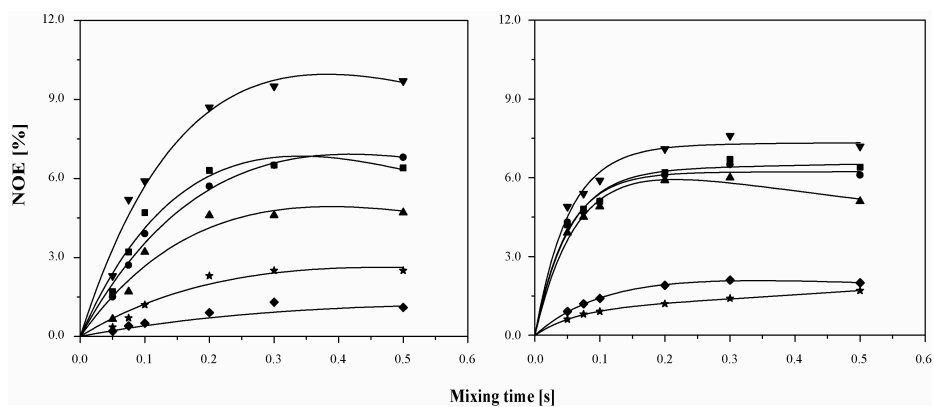


Figure 4.49: Experimental (left) and theoretical (right) trNOEs (%) build up curves of (A) trisaccharide **2**, (B) tetrasaccharide **4** and (C) pseudo tetrasaccharide **5** in the presence of MAG at different mixing times (s). The molar ratio was 1:18, 1:20 and 1:22 for MAG: **2**, MAG: **4** and MAG: **5** respectively. The initial parameters (K_D , k_{off} , τ_c^{ligand} and $\tau_c^{protein}$) are reported in Table 4.10. The curves represent following trNOEs: (A) H1-H3 (■), H1'-H3' (●), H3'-H4' (▲), H1'-H3 (▼), H1'-H4 (◆), H3'-H8'' (★) and H3'-H3ax'' (○); (B) H1''-H3'' (■), H4'-H6'b (●), H4'-H6'a (▲), H1''-H3' (▼), H1''-H4' (◆), H4'-H6'' (★), H3''-H8''' (○) and H3''-H3ax''' (△); (C) H1''-H3'' (■), H1''-H5'' (●), H6'-CH^{LAc} (▲), H1''-(H3'+H4') (▼), H1''-(H2ax'+H2eq') (◆) and H3''-H8''' (★).

Table 4.11: Using CORCEMA calculation, interproton distances for the bioactive conformations (Figure 4.49) of the saccharides **2-5** (A) trisaccharide **2**, (B) tetrasaccharide **4** and (C) pseudo tetrasaccharide **5**. R-factor indicates the deviation between observed and calculated trNOEs.

(A)

Restrains	Distance (Å)	R-factor
H1-H3	2.6	0.10
H1'-H3'	2.5	0.20
H3'-H4'	2.5	0.04
H1'-H3	2.4	0.21
H1'-H4	3.4	0.08
H3'-H3ax''	4.2	0.67
H3'-H8''	2.6	0.42

(B)

Restrains	Distance (Å)	R-factor
H1''-H3''	2.6	0.39
H1''-H3'	2.5	0.18
H1''-H4'	3.7	0.20
H4'-H6'b	2.5	0.36
H4'-H6'a	3.5	0.32
H4'-H6''	3.3	0.28
H3''-H3ax'''	4.3	0.57
H3''-H8'''	2.7	0.71

(C)

Restrains	Distance (Å)	R-factor
H1''-H3''	2.5	0.24
H1''-H5''	2.4	0.30
H6'b-CH ^{LAc}	2.4	0.56
H1''-(H3'+H4')	H1''-H3': 2.4 H1''-H4': 3.6	0.24
H1''-H2'	H1''-H2ax': 4.4 H1''-H2eq': 3.7	1.00
H3''-H8'''	2.8	0.40

Table 4.12: Dihedral angles (in degrees) at the glycosidic linkages for the bound conformation of saccharides **2**, **4** and **5** in the presence of MAG. The dihedral angles (ϕ , ψ , ω) at the glycosidic linkages were defined as: ϕ_1 , C1'''-C2'''-O2''-C3'' (Neu5Ac-Gal); ψ_1 , C2'''-O2'''-C3''-H3'' (Neu5Ac-Gal); ϕ_2 , H1''-C1''-O3'-C3' (Gal-Gal); ψ_2 , C1''-O1''-C3'-H3' (Gal-Gal); ϕ_3 , C1-C2-O6'-C6' (Neu5Ac-Gal); ψ_3 , C2-O2-C6'-C5' (Neu5Ac-Gal) and ω = O5'-C5'-O6'-C6'.

Saccharides	ϕ_1/ψ_1	ϕ_2/ψ_2	ϕ_3/ψ_3	ω
2	- 64/ - 17	60/ - 46		
4	- 58/ - 25	68/ - 40	- 159/ 87	81
5	- 59/ - 26	63/ - 36		67

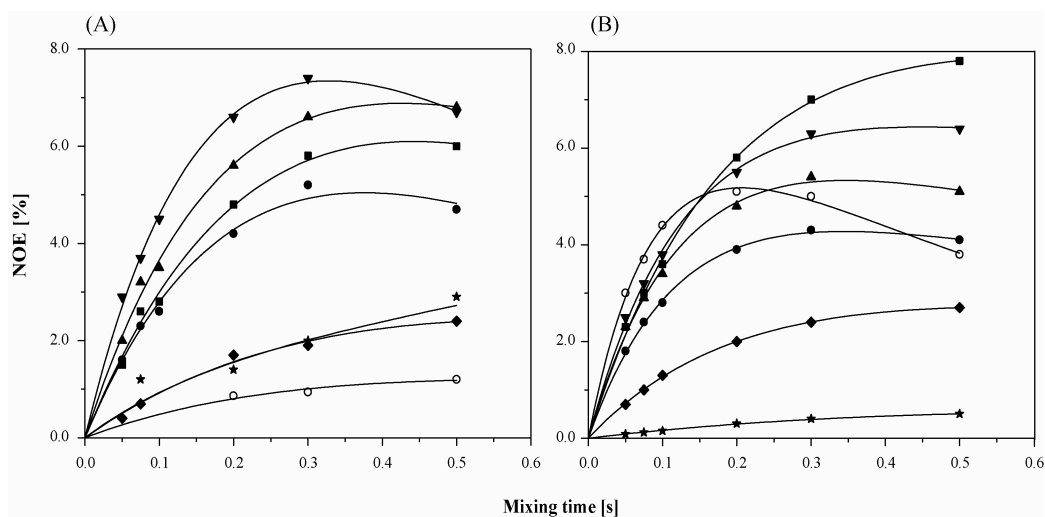


Figure 4.50: (A) Experimental and (B) theoretical trNOEs (%) build up curves trisaccharide **2** in the presence of MAG at different mixing times. The $\phi_1, \psi_1 = -174^\circ, -8^\circ$ and $\phi_2, \psi_2 = 58^\circ, -51^\circ$ was used for the calculation of theoretical trNOEs using CORCEMA program. All the parameters (K_D , k_{off} , τ_c^{ligand} and τ_c^{protein}) and symbols are same as in Figure 4.49.

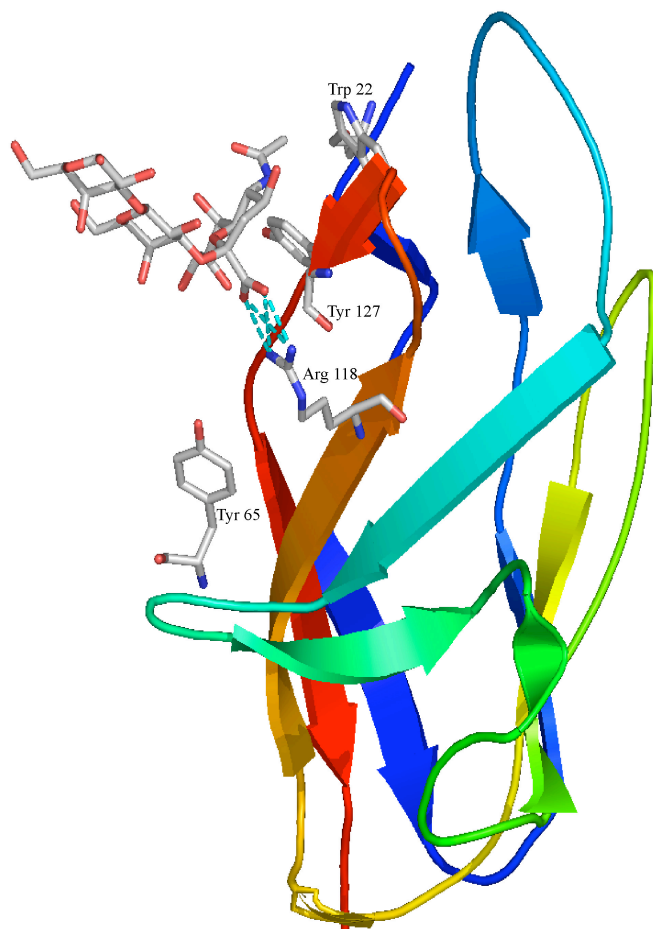
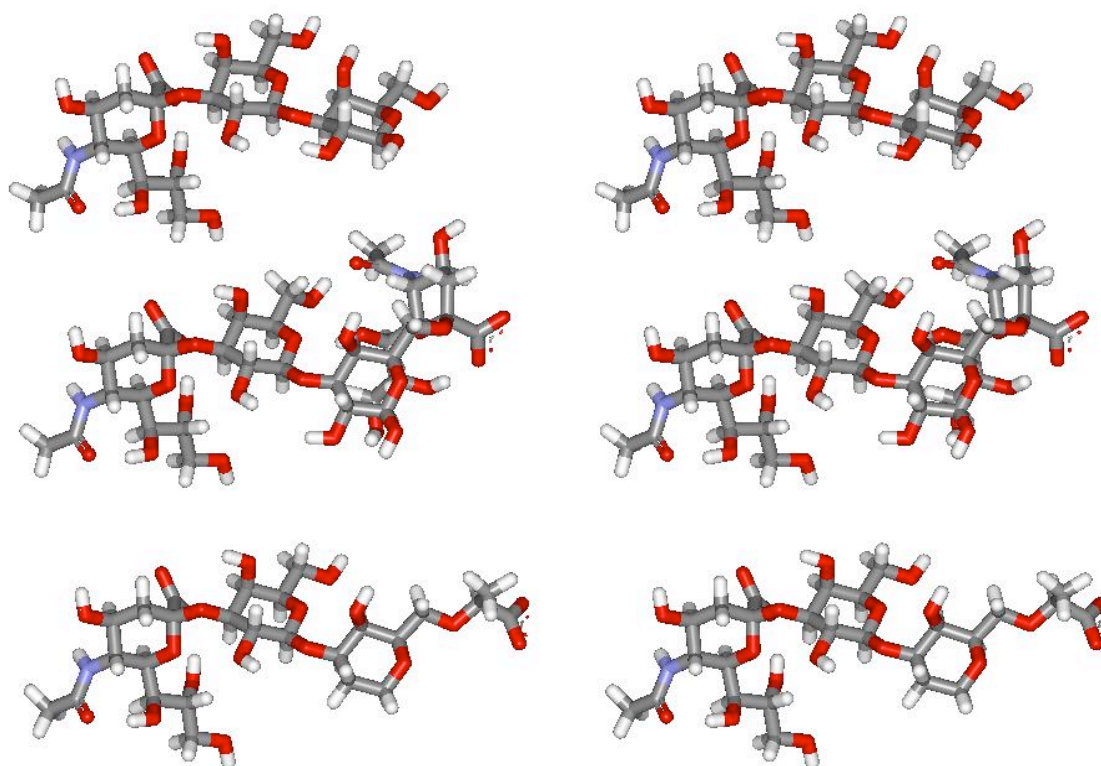


Figure 4.51: Trisaccharide **2**- MAG complex in H3'-H3ax'' mode.

Table 4.13: Computed interproton distances and R-factors (c.f., Figure 4.51) from the structures that interpreted NOESY spectra of trisaccharide **2** bound to MAG using CORCEMA program.

Restrains (symbols)	Distance (Å)	R-factors
H1-H3 (■)	2.6	0.25
H1'-H3' (●)	2.5	0.13
H3'-H4' (▲)	2.4	0.19
H1'-H3 (▼)	2.4	0.14
H1'-H4 (◆)	3.3	0.22
H3'-H8'' (★)	4.6	0.83
H3'-H3ax'' (○)	2.1	3.70

**Figure 4.52:** Bioactive conformation of (top) trisaccharide **2**, (middle) tetrasaccharide **4** and (bottom) pseudo tetrasaccharide **5**. In the bioactive conformation, the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal orientation is very similar to the one found for sLe^x bound to E-selectin. The orientation of the α -(2 \rightarrow 6)-linkage of tetrasaccharide **4** is set to “gt” conformation (ω = O5'-C5'-O6'-C6') (Poppe et al., 1992).

4.2.2.5 Epitope mapping of **5** by STD NMR experiments

In general, it is important to know the size and the shape of the binding epitope of a ligand bound to receptor protein. The NOESY experiments deliver the shape of the bound ligand in the presence of receptor protein but provide very limited information about the size of the binding epitope. To map the binding epitope of the pseudo tetrasaccharide **5** bound to MAG, saturation transfer difference (STD) NMR (Mayer and Meyer, 1999; Meyer and Peters, 2003; Bhunia et al., 2004) experiment was performed. As this was a collaboration project with the group of Prof. Bernd Meyer, Universität Hamburg, the STD NMR experiments of **2** and **3** were performed by them but the results are discussed here wherever required. Figure 4.53 show the 1D STD NMR spectra of pseudo tetrasaccharide **5** bound to MAG. It is immediately obvious that the Neu5Ac residue leads to the most prominent STD signals, indicating that this residue was mainly responsible for the binding reaction. The β -D-Gal-(1 \rightarrow 3)-D-Cyc and CH^{LAc} moieties also played an important role for binding to MAG.

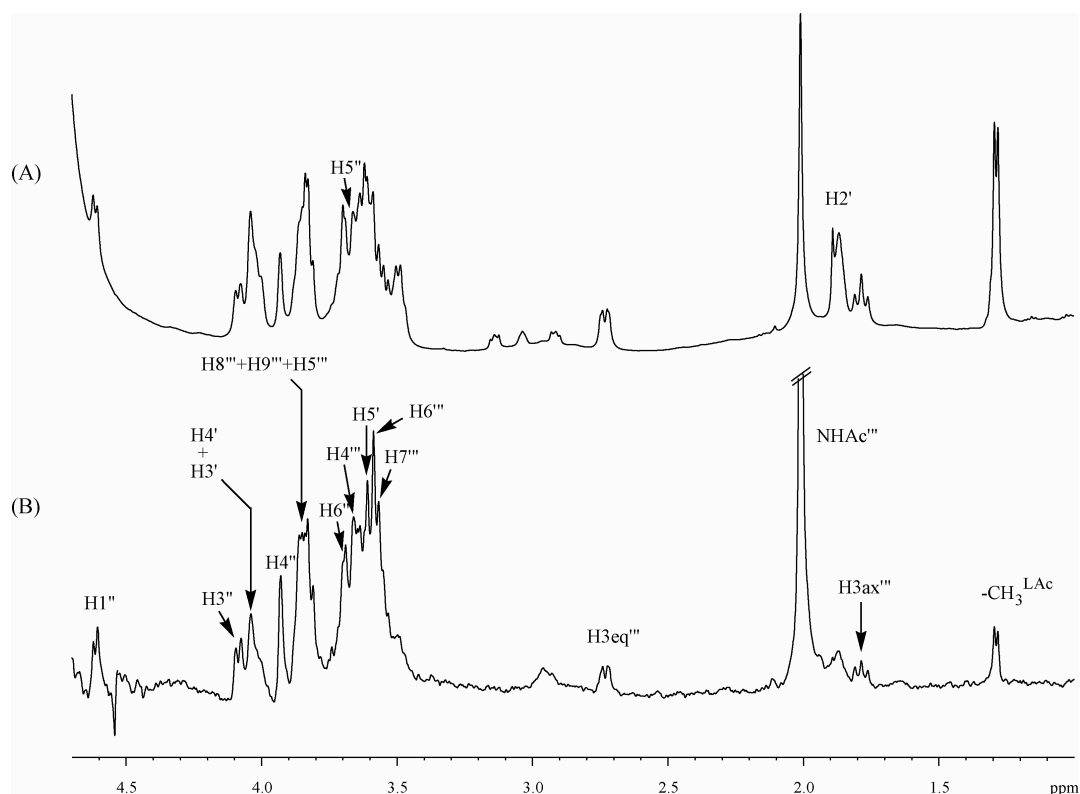


Figure 4.53: (A) ¹H NMR and (B) STD NMR spectra of pseudo tetrasaccharide **5** in the presence of MAG (molar ratio 1:60). For the STD NMR spectrum the protein was saturated with 40 selective Gaussian-shaped pulses (50 ms each) resulting in a total saturation time of ~ 2 s (on resonance 0 ppm and off resonance 40 ppm). STD signals from the Neu5Ac residue are most prominent.

The value of the integral of the signal was largest for *N*-acetyl methyl group of **5**; therefore, according to group epitope mapping (GEM) (Mayer and Meyer, 2001) principle its value was set to 100 %. The intensity of all other protons was calculated as the percentage of this group (Figure 4.54). In the spectral region from 3.50 to 4.10 ppm, 20 protons from the different residues were observed. Only 14 of these made a significant contribution to the STD signal. The H5''', H8''', H9''' of Neu5Ac and CH^{LAc} of LAc were crowded in the region between 3.98-3.99 ppm. So, no individual protons could be assigned in this crowded region. The contribution of the H6''' and H7''' of Neu5Ac was around 60 %, which is in good agreement with the AutoDock result. Tyr 127 made a van-der-Waals contact with these two protons. The H4'' of Gal reached a value of 56 % in comparison to 40 % STD effect by H3'', H5'' or H6'' of Gal. In contrast, the interactions of the H4', H5' or H6' of Cyc with the protein were small, and consequently showed a STD effect of about 40 %. The methyl group proton resonance of LAc residue was saturated to a degree of 38 %.

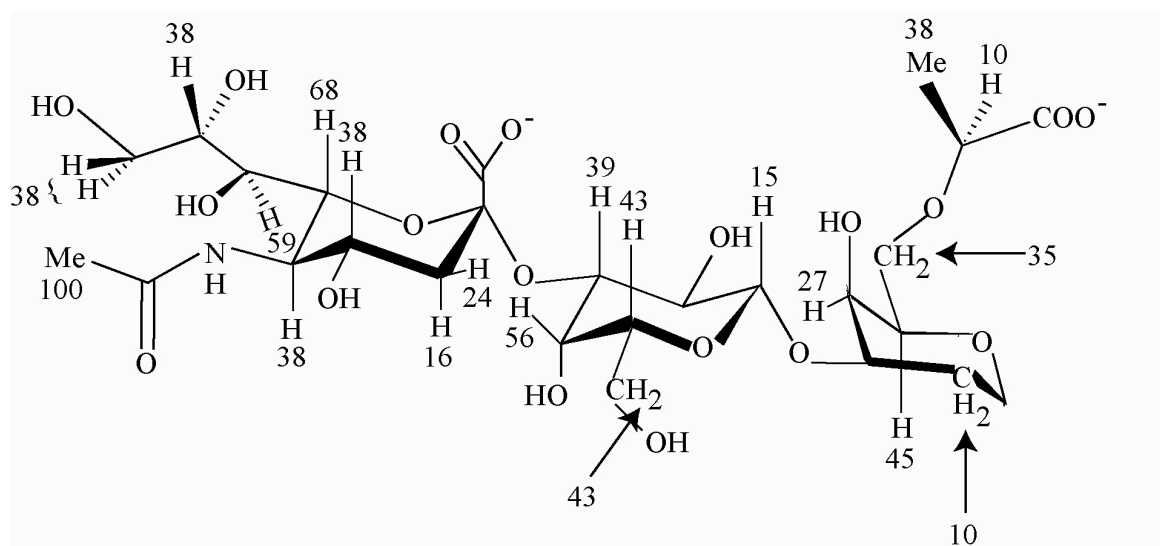
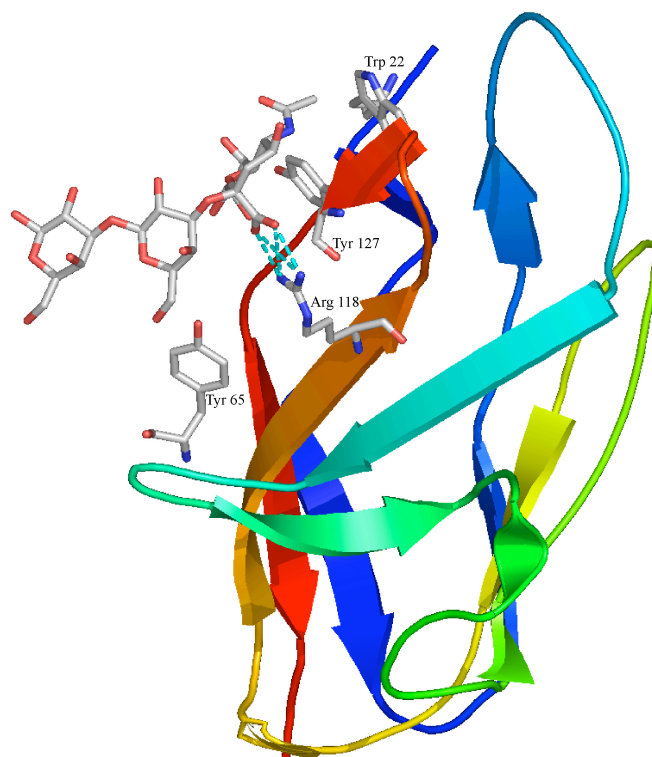


Figure 4.54: Relative STD effects of pseudo tetrasaccharide **5** bound to MAG. The values were calculated from the individual signal intensities in the STD spectrum (I_{STD}) and corresponding reference spectrum (I_0) (Mayer and Meyer, 2001). The ratios of the intensities I_{STD}/I_0 were normalised using the largest STD effect, *N*-acetyl methyl group as a reference. Individual proton intensities were represented with respect to *N*-acetyl methyl group of Neu5Ac.

4.2.2.6 Ligand-protein contacts

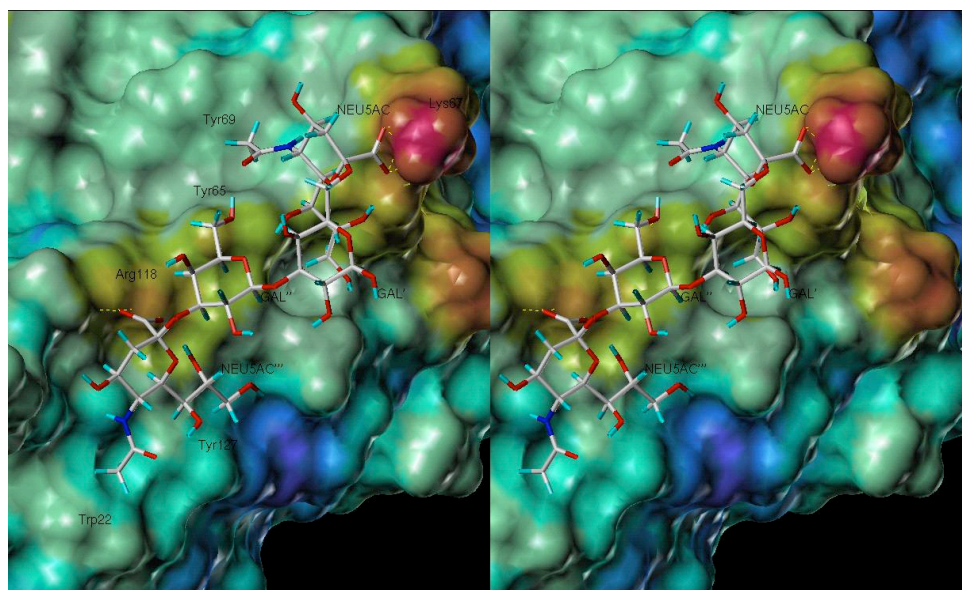
The putative binding of the bioactive conformation of trisaccharide **2**, tetrasaccharide **4** and pseudo tetrasaccharide **5** in the active site of MAG is shown in Figure 4.55 and described in Table 4.14. It is especially interesting that α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal interacts with protein mostly via the *N*-acetyl methyl group of Neu5Ac and its C 6 to C 9 of the side chain. The *N*-acetyl methyl group interacts with Trp 22 residue on the surface of MAG as shown in Figure 4.55. According to the Docking model (Tripos, USA) the carboxyl function of Neu5Ac formed a salt bridge with the guanidine group of Arg 118. The H4' (for **2**) or H4'' (for **4**, **5**) of Gal was also very close to guanidine group of Arg 118. Similar to sialoadhesin-sialyl lactose crystal structure (Acquisition code, 1QFO) (May et al., 1998), the OH6' of trisaccharide **2** also formed a hydrogen bond with the hydroxyl group of Tyr 65. It is clear that **2**, **4** and **5** show the same binding pattern at the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage as found in the crystal structure of sialyl lactose bound to sialoadhesin (May et al., 1998). Additionally, the AutoDock model predicted that the carboxyl group of α -(2 \rightarrow 6)-linked Neu5Ac/Lac formed strong electrostatic interactions with the Lys 67 of MAG.

(A)



(Figure continued)

(B)



(C)

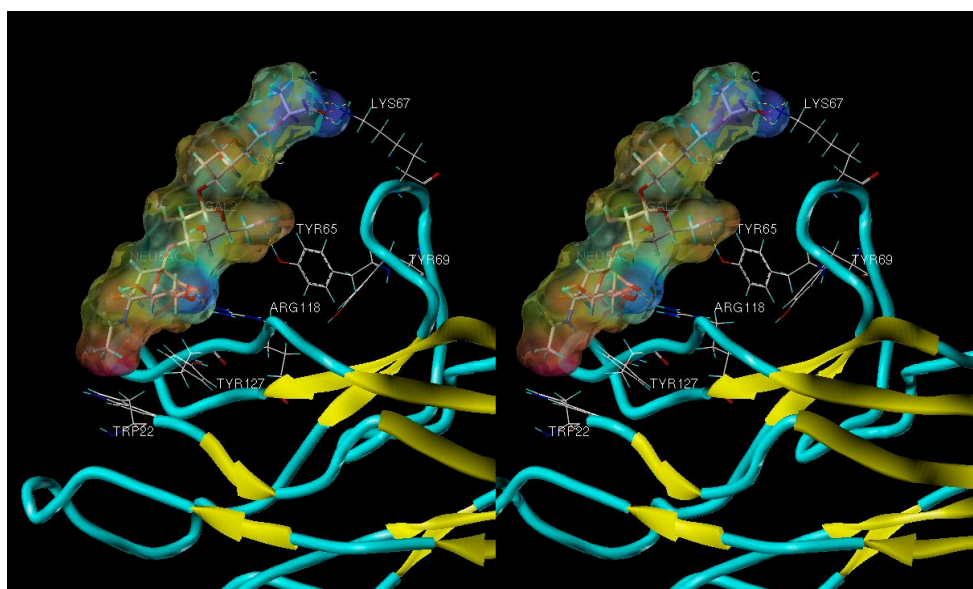


Figure 4.55: The structure of (A) trisaccharide **2**, (B) tetrasaccharide **4** and (C) pseudo tetrasaccharide **5** in the binding site of MAG. The corresponding dihedral angles at the glycosidic linkages of the saccharides are reported for the theoretical trNOE calculations using CORCEMA program (Figure 4.49 and Table 4.12). The α -(2 \rightarrow 3)-glycosidic linkage prefers “syn” conformation. The dotted lines indicates the strong electrostatic interaction between MAG amino acids and saccharides. The strong electrostatic interaction between NH_2 group of Lys 67 and carboxyl group of α -(2 \rightarrow 6)-linked Neu5Ac/(S)-LAc increases the binding affinity of **4** or **5** to MAG.

Table 4.14: Interproton distances between protein and (A) trisaccharide **2**, (B) tetrasaccharide **4** and (C) pseudo tetrasaccharide **5** at the binding site of MAG.

(A)

Type of interaction	Sugar	MAG	Distance (Å)
H-bonding	Neu5Ac''.O1A	Arg118.HH2	1.6
	Neu5Ac''.O1B	Arg118.HH1	1.5
	Neu5Ac''.NH	Gln126.O'	1.8
	Neu5Ac''.O8	Thr128.H ^N	1.9
	Gal'.O6	Tyr65.H ^N	1.9
Proton-Proton	Neu5Ac''.H3eq	Gln126.HB1	3.8
	Neu5Ac''.H4	Gln126.HB1	2.3
	Neu5Ac''.H6	Tyr127.HA	2.4
	Neu5Ac''.H7	Tyr127.HA	2.6
	Neu5Ac''.H8	Thr128.HG23	3.5
	Neu5Ac''.H9A	Tyr127.HD2	3.0
	Neu5Ac''.H9B	Tyr127.HD2	3.6
	Neu5Ac''.CH ₃	Trp22.HB2	3.8
		Trp22.HE1	2.5
		Trp22.HD1	2.8
	Gal'.H3	Arg118.HH2	3.8
	Gal'.H4	Arg118.HH2	3.0
	Gal'.H5	Thr128.HG23	3.8

(B)

Type of interaction	Sugar	MAG	Distance (Å)
H-bonding	Neu5Ac'''.O1A	Arg118.HH2	1.5
	Neu5Ac'''.O1B	Arg118.HH1	1.6
	Neu5Ac'''.NH	Gln126.O'	1.7
	Neu5Ac'''.O8	Thr128.H ^N	1.8
	Gal''.O6	Tyr65.OH	2.0
	Neu5Ac.OH8	Tyr65.OH	2.2
	Neu5Ac.O1A	Lys67.HZ1	1.6
Proton-Proton	Neu5Ac.O2B	Lys67.HZ2	1.6
	Neu5Ac'''.H3eq	Gln126.HB1	4.1
	Neu5Ac'''.H4	Gln126.HB1	2.3
	Neu5Ac'''.H6	Tyr127.HA	2.5
	Neu5Ac'''.H7	Tyr127.HA	2.6
	Neu5Ac'''.H8	Thr128.HG23	3.6
	Neu5Ac'''.H9A	Tyr127.HD2	2.9
	Neu5Ac'''.H9B	Tyr127.HE2	3.3
	Neu5Ac'''.CH ₃	Trp22.HB2	3.7
		Trp22.HE1	2.7
		Trp22.HD1	2.8
	Gal''.H5	Thr128.HG23	3.4
	Neu5Ac.H6	Pro66.HA	2.6

(Table continued)

Neu5Ac.H7	Tyr65.HD1	2.4
Neu5Ac.H9A	Tyr65.HD1	2.4
Neu5Ac.H9B	Tyr65.HE1	2.7
Neu5Ac.CH3	Tyr69.HD1	3.8
	Tyr69.HB2	2.4
	Tyr65.HD1	3.1

(C)

Type of interaction	Sugar	MAG	Distance (Å)
H-bonding	Neu5Ac ^{'''} .O1A	Arg118.HH2	2.0
	Neu5Ac ^{'''} .O1B	Arg118.HH1	1.7
	Neu5Ac ^{'''} .NH	Gln126.O'	1.9
	Neu5Ac ^{'''} .O8	Thr128.H ^N	2.1
	Gal ^{''} .O6	Tyr65.OH	2.2
	LAc.O1A	Lys67.HZ1	2.4
	LAc.O2B	Lys67.HZ2	1.9
Proton-Proton	Neu5Ac ^{'''} .H3eq	Gln126.HB1	4.1
	Neu5Ac ^{'''} .H4	Gln126.HB1	2.3
	Neu5Ac ^{'''} .H6	Tyr127.HA	2.6
	Neu5Ac ^{'''} .H7	Tyr127.HA	2.6
	Neu5Ac ^{'''} .H8	Thr128.HG23	3.5
	Neu5Ac ^{'''} .H9A	Tyr127.HD2	2.5
	Neu5Ac ^{'''} .H9B	Tyr127.HE2	3.7
	Neu5Ac ^{'''} .CH ₃	Trp22.HB2	3.8
		Trp22.HE1	2.6
		Trp22.HD1	2.8
	Gal ^{''} .H5	Thr128.HG23	4.3
	LAc.CH	Lys67.HZ1	4.1

4.2.2.7 Discussion

Recently, Vyas and co-workers (2002) have qualitatively shown that the tetrasaccharide **3** is probably the dominant structure interacting with MAG and regulating its inhibitory function. The previously proposed ligand GQ1b α N is still supposed to be the strongest ligand to MAG. However, due to its scarcity in brain tissue, it does not carry most of the biological function of inhibiting MAG *in vivo*. It is well known from the literature that the small molecule mimics of oligosaccharides are capable of being developed into a drug with good stability and synthetic availability (Jiménez-Barbero et al., 1999). Therefore, to study the biological importance of α -(2 \rightarrow 6)- linked Neu5Ac of tetrasaccharide **3** to MAG, Neu5Ac is replaced by (S)-LAc while GalNAc residue is substituted by its mimic cyclic ether, Cyc residue.

NOESY and STD NMR experiments in conjunction with AutoDock model yield a model that describes the molecular recognition of **2**, **4** and **5** by MAG at an atomic resolution. The interaction between MAG and saccharides **2-6** has been analysed using NOESY experiments, which mainly focus on the characterisation of the *interglycosidic* linkage conformation of **2-6** in the presence of protein.

The most striking trNOE observation is at the flexible α -(2→3)-glycosidic linkage between Neu5Ac and Gal of saccharides **2-5**. It is confirmed by trNOE analysis of the protein that MAG recognises only one conformation, which is characterised by an *interglycosidic* contact between H3 of Gal and H8 of Neu5Ac. Two other conformations exist at the α -(2→3)-glycosidic linkage of **2** and **3** in aqueous solution but the experimental trNOE data are only consistent with the so called “syn” conformation. From this observation alone, it is clear that the α -(2→3)-glycosidic linkage changes its conformation upon binding to MAG and is closely related to global minimum energy, Min 1, which is less populated in aqueous solution. The conformation at this glycosidic linkage is similar to that of the sLe^x bound to E-selectin (Scheffler et al., 1997). Recently, the crystal structure of sialoadhesin co-crystallised with sialyl lactose (May et al., 1998) also provide evidence for the above effect. The reason for this distortion becomes clear on inspecting the molecular model of the local minimum energy, Min 2 and global minimum energy, Min 1. MAG is a Neu5Ac binding lectin, Arg 118 is responsible for binding to carboxyl group of Neu5Ac via salt bridge. Therefore, Neu5Ac residue has to be reoriented in order to come into contact with the protein surface. Obviously, this orientation takes place around the α -(2→3)-glycosidic linkage (Figure 4.42). Distortions of the glycosidic linkages upon binding to proteins have already been observed (Haselhorst et al., 1999, Hricovini et al., 2001), but it is difficult to estimate the energy that is required to facilitate the binding of such distorted conformation. In comparison to this α -(2→3)-glycosidic linkage, β -(1→3)-glycosidic linkage of saccharides **2-5** is more rigid and present only in one conformation with $\phi_2, \psi_2 = 68^\circ, -40^\circ$. The protein mediated spin diffusion was excluded using ROESY experiments (Poveda et al., 1997; Casset et al., 1996; Scheffler et al., 1997; Haselhorst et al, 1999; Maaheimo et al., 2000).

The additional Neu5Ac residue at the α -(2→6)-glycosidic linkage show the flexibility of this linkage in aqueous solution of tetrasaccharide **3**. The only α -(2→6)-

linkage containing crystal structure available in the Protein Data Bank (PDB) is that of the Pertussis Toxin complexed with disaccharide α -D-Neu5Ac-(2 \rightarrow 6)- β -D-Gal (acquisition code, 1PTO) (Stein et al., 1994). In the crystal structure, the disaccharide is observed to adopt several conformations about the ω torsional angle, confirming the flexibility of this α -(2 \rightarrow 6)-linkage. The NMR derived 3D structure of tetrasaccharide **4** agrees not only with the theoretical trNOEs using CORCEMA program but AutoDock model also predicts the electrostatic interaction between positively charged MAG amino acids in the binding site and the negatively charged carboxyl in the tetrasaccharide **4**. The computed distance between the NH₂ group of Lys 67 and the carboxylate oxygen of α -(2 \rightarrow 6)-linked Neu5Ac is about 1.6 Å, corresponding well with the expected strong interaction. Further evidence of these strong electrostatic interactions (Figure 4.55) has been found by the site directed mutagenesis study (Heiko Gächter personal communication). Mutant MAG with an alanine replacing Lys 67 (K67A) show only about 50 % of the biological activity of the tetrasaccharide **3** (Table 4.15). This dramatic effect indicates that this interaction make a significant contribution to the binding energy. A spatial closeness is also observed from the AutoDock model for the aromatic ring of Tyr 69 and the *N*-acetyl methyl group of α -(2 \rightarrow 6)-linked Neu5Ac. This hydrophobic interaction can also be another driving force for the rigidity of the α -(2 \rightarrow 6)-linkage (ω = 60°) of tetrasaccharide **3** or **4** in the presence of MAG. The OH6'' of Gal form only one hydrogen bond with the hydroxyl group of Tyr 65. The AutoDock model is also in good agreement with the STD data obtained from the MAG-tetrasaccharide **3** (Prof. Bernd Meyer, Universität Hamburg) and MAG-pseudo tetrasaccharide **5** complex. The α -(2 \rightarrow 3)-linked Neu5Ac exhibits larger STD effects since this moiety is closest to the binding site (Figure 4.55). In the Gal moiety, the H4'' show a large fraction of saturation transfer due to close proximity of Arg 118. The GalNAc (for tetrasaccharide **3**) or Cyc (for pseudo tetrasaccharide **5**) residue has relatively lower STD values. This is not surprising since the AutoDock model of tetrasaccharide **4** or pseudo tetrasaccharide **5** show that this moiety pointing towards solution. STD-HSQC experiment of **3** in the presence of MAG also show medium STD effects at the side chain of the α -(2 \rightarrow 6)-linked Neu5Ac (Prof. Bernd Meyer, Universität Hamburg). These results indicate that the side chain of the α -(2 \rightarrow 6)-linked Neu5Ac is also responsible for the binding to protein. One fundamental conclusion of these studies is the importance of an additional carboxyl

group at the α -(2→6)-linkage in the structure of saccharides **3-5** (Table 4.15). This substantiates that there are anchoring points for both carboxyl groups that help to fix the saccharides **3-5** on the surface of the MAG.

NOESY experiments of trisaccharide **6** in the presence of MAG show very low intensity, indicating the interaction is very weak. In that case, STD experiment can map the ligand even the K_D value is higher than 1 mM. But *N*-acetyl methyl group of trisaccharide **6** (Figure 4.56) is the only responsible group for intimate contact to the protein, indicating the K_D value is much higher than 1 mM. This experimental result is in good agreement with the relative inhibitory potential values, observed by the group of Prof. Sørge Kelm, Universität Bremen (Table 4.15). Trisaccharide **6** is the least potent inhibitor among all the inhibitor **2-5** tested so far (Table 4.15). This result concludes that the salt bridge between carboxyl group of α -(2→3)-linked Neu5Ac and Arg 118 is prerequisite for binding to MAG.

The K_D values obtained from the STD NMR titrations (Prof. Bernd Meyer, Universität Hamburg) were used to determine the k_{on} and k_{off} rate constants, optimising the fit between experimental and calculated trNOE build up curves. The values of k_{on} and k_{off} rate constants correspond well with the values obtained for the binding of sialyl Lewis^a to a monoclonal antibody by Biacore surface plasmon resonance (Herfurth et al., 2005). The data suggest that association of the saccharides and MAG is significantly slower than this would be expected from a diffusion-controlled process. It is an open question whether this is due to conformational changes of MAG, or of the ligand, or of both during the binding process.

Table 4.15: Relative inhibitory potentials (rIPs) of saccharides **2-6**. The rIPs of each inhibitor are calculated by dividing the IC_{50} of the reference compound **2** by the IC_{50} of the compound of interest. This results in rIPs above 1.0 for saccharides that inhibit better than the reference compound and rIPs lower than 1.0 for saccharides inhibiting weaker. Mutation of Lys 67 (K67A) shows the significant reduction in binding. It is suggested that this loss of activity is due to the loss of electrostatic interaction between Lys 67 and the carboxylate group of α -(2→6)-linked Neu5Ac/LAc. (WT: Wild type).

Saccharides	WT MAG	K67A MAG
Trisaccharide 2	1.0	1.0
Tetrasaccharide 3	3.3	1.4
Tetrasaccharide 4	4.2	not measured
Pseudo tetrasaccharide 5	4.1	not measured
Trisaccharide 6	0.2	not measured

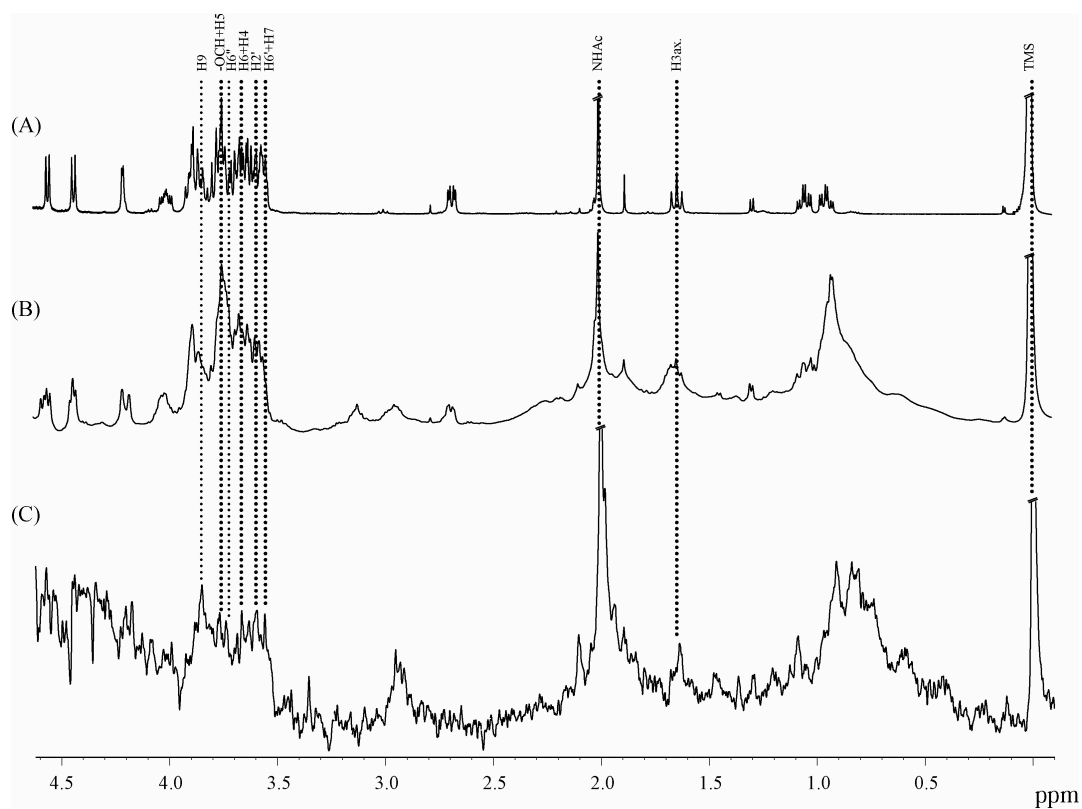


Figure 4.56: (A) ^1H NMR of trisaccharide **6** and (B) in the presence of MAG (protein:ligand::1:50) and (C) corresponding STD NMR spectra. The STD spectra were acquired with selective presaturation at $\delta = -1.5$ ppm for a duration of 2 s utilising 40 cascade of Gaussian pulses (duration 50 ms, spacing 1 ms). 2 K scans were recorded for the STD spectra, and 1 K scans for the reference ^1H NMR spectra.

5 Conclusions

In this work, NMR experiments in conjunction with computer-aided-docking models are used as a key technique to characterise the different protein-carbohydrate interactions. The goal of this thesis is the determination of bioactive conformations and binding epitopes of different oligosaccharides in the presence of siglecs, sialoadhesin or MAG.

In the first part, the binding epitope of α -(2 \rightarrow 3)-sialyl lactose is determined in complex with sialoadhesin by STD NMR experiments. The solution state data in combination with FlexiDock calculation confirm the interactions observed in the crystal structure, demonstrating that the key interaction is between the *N*-acetyl methyl group of α -(2 \rightarrow 3)-sialyl lactose and Trp 2 of sialoadhesin.

In the second part, trisaccharide **2** was subjected to NMR studies in the presence of sialoadhesin. The experimental NOE data indicate that the α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal glycosidic linkage undergoes a major conformational change upon binding to sialoadhesin (Figure 5.1). On the other hand, the β -D-Gal-(1 \rightarrow 3)- β -D-Gal linkage is less affected upon binding to the protein. The data indicate that the Gal-Gal moiety contributes significantly to the binding to sialoadhesin.

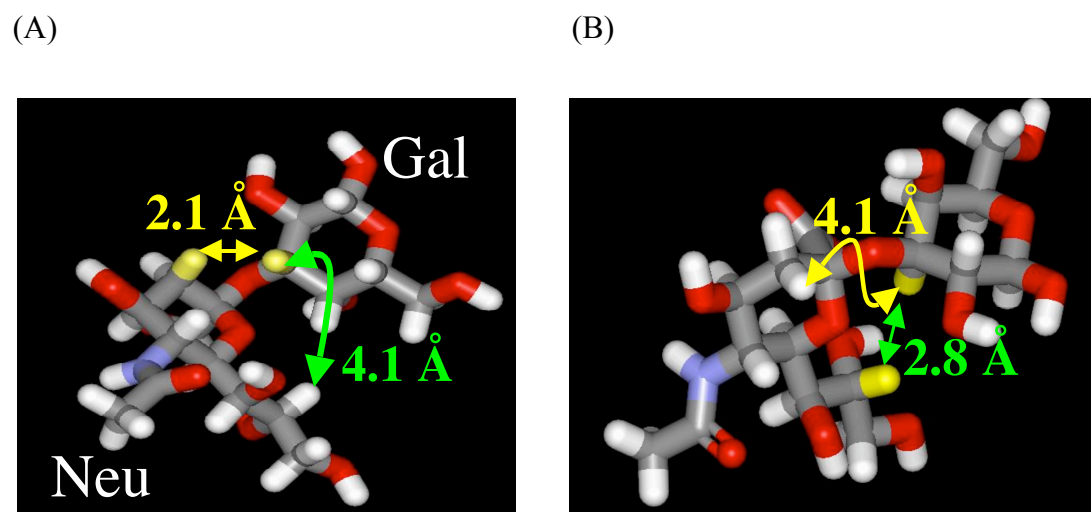


Figure 5.1: The main *interglycosidic* NOE between α -D-Neu5Ac-(2 \rightarrow 3)- β -D-Gal linkage of **2** at (A) aqueous solution and (B) in the presence of sialoadhesin. In aqueous solution, there is a conformational equilibrium between H3 of Gal and H3ax of Neu5Ac and H3 of Gal and H8 of Neu5Ac whereas in the presence of the protein the interaction is solely between H3 of Gal and H8 of Neu5Ac.

In the third part, the conformations of the free saccharides **2-5** are investigated using NOESY experiments in combination with MD simulations. It is observed that

the preferred solution conformations are mostly in accordance with literature data. Since the α -(2 \rightarrow 3)-glycosidic linkage between the terminal Neu5Ac and Gal is flexible, main attention was paid to this linkage. The *interglycosidic* NOEs across this linkage are in very good agreement with previous observations, and therefore, it is assumed that three major conformational families reflect the solution conformation at this linkage. The core β -(1 \rightarrow 3)-linkages are relatively rigid whereas the α -(2 \rightarrow 6)-linkage of **3** or **4** is very difficult to analyse due to severe signal overlap and significant flexibility.

In the last part of the thesis, MAG-oligosaccharide (2-6) interactions are characterised at atomic resolution. For this study the tetrasaccharide **3** is chosen since it is a part of the natural ligand, GQ1b α N. To determine the bioactive conformation of **3** in the presence of MAG, homology model of MAG is necessary, since the crystal structure of MAG is not known yet. Once again, the most striking observation is that the bioactive conformation of tetrasaccharide **3** around the flexible α -(2 \rightarrow 3)-glycosidic linkage between Neu5Ac and Gal (Figure 5.1). Although the α -(2 \rightarrow 6)-linked Neu5Ac significantly enhances the binding affinity towards MAG as compared to the parent trisaccharide **2**, due to severe signal overlap the determination of the conformation of the α -(2 \rightarrow 6)-linked Neu5Ac was not possible.

For the characterisation of ligand binding to MAG in order to facilitate the design of inhibitors, the following inhibitors, modified tetrasaccharide **4**, pseudo tetrasaccharide **5** and trisaccharide **6** were synthesised based on tetrasaccharide **3** by the group of Professor Beat Ernst, Pharmazentrum, Universität Basel, Switzerland.

Both saccharides **4** and **5** indicate the same conformational change at the α -(2 \rightarrow 3)-glycosidic linkage upon binding to MAG (Figure 5.1). In case of tetrasaccharide **4**, modification of the GalNAc to Gal residue at the reducing end did not change the binding affinity. Addition of Neu5Ac at the α -(2 \rightarrow 6)-linkage increased the binding affinity by a factor of 4 (Table 5.1). Pseudo tetrasaccharide **5**, where a replacement of the α -(2 \rightarrow 6)-linked Neu5Ac with (S)-LAc had occurred also showed the same binding affinity as tetrasaccharide **4** (Table 5.1). AutoDock 3.0 explains this by suggesting a strong electrostatic interaction between the carboxyl group of α -(2 \rightarrow 6)-linked residue of **4** or **5** and Lys 67 of MAG that increases the binding affinity. The relative inhibitory potential (rIPs) values of **3** with a mutant MAG (K67A) suggest that the α -(2 \rightarrow 6)-linked carboxyl group has a greater importance for

the specific recognition of MAG than the rest of the Neu5Ac residue. Trisaccharide **6** did not bind to MAG even though it has an α -(2→6)-linked Neu5Ac, indicating that the α -(2→3)-linked Neu5Ac is a prerequisite for affinity. The results of the inhibitor screening allowed to draw the conclusions as to why some compounds function well as inhibitors and others not. The preliminary model for good inhibition derived from this work is that the α -(2→3)-linked Neu5Ac is mandatory to recognise the protein and additional carboxyl group at the C6 position of trisaccharide **2** increases the binding affinity. This information is essential in the future for designing of drugs targeted towards the MAG.

Table 5.1: Relative inhibitory potentials (rIPs) of saccharides **2-6**. The rIPs of each inhibitor are calculated by dividing the IC₅₀ of the reference compound **2** by the IC₅₀ of the compound of interest. This results in rIPs above 1.0 for saccharides that inhibit better than the reference compound and rIPs lower than 1.0 for saccharides inhibiting weaker. Mutation of Lys 67 (K67A) shows the significant reduction in binding. It is suggested that this loss of activity is due to the loss of electrostatic interaction between Lys 67 and the carboxylate group of α -(2→6)-linked Neu5Ac/LAc. (WT: Wild type).

Saccharides	WT MAG	K67A MAG
Trisaccharide 2	1.0	1.0
Tetrasaccharide 3	3.3	1.4
Tetrasaccharide 4	4.2	not measured
Pseudo tetrasaccharide 5	4.1	not measured
Trisaccharide 6	0.2	not measured

Appendix

A.1 Chemical shifts for oligosaccharides 1-6

A.1.1 ¹H Chemical shift assignments for α-(2→3)-sialyl lactose 1 at 303 K and 500 MHz

	Neu5Ac ^{α2,3}	Gal	Glc ^{α-form}	Glc ^{β-form}
H1		4.64	5.32	4.77
H2		3.65	3.68	3.38
H3		4.21	3.99	3.72
H3(ax)	1.91			
H3(eq)	2.86			
H4	3.69	4.06	3.78	3.76
H5	3.95	3.81	3.73	3.75
H6	3.68	3.98	4.07/4.05	3.92/3.90
H7	3.79			
H8	3.84			
H9	3.74/3.96			
N-CH ₃	2.13			

A.1.2 ¹H Chemical shift assignments for trisaccharide 2 at 277 K and 500 MHz

	Neu5Ac ^{α2,3}	Gal	Gal	Spacer
H1		4.66	4.45	
H2		3.64	3.62	
H3		4.10	3.79	
H3(ax)	1.77			
H3(eq)	2.75			
H4	3.66	3.93	4.18	
H5	3.84	3.68	3.68	
H6	3.59	3.74/3.68	3.74/3.68	
H7	3.56			
H8	3.86			
H9	3.61/3.84			
N-CH ₃	2.01			
CH ₂				0.96/1.06
O-CH ₂				3.74/4.04
SiMe ₃				0.00

A.1.3 ^1H Chemical shift assignments for tetrasaccharide 3 at 288 K and 700 MHz

	Neu5Ac $^{\alpha 2,3}$	Gal	GalNAc	Neu5Ac $^{\alpha 2,6}$	Spacer
H1		4.49	4.51		
H2		3.52	3.97		
H3		4.05	3.80		
H3(ax)	1.78			1.66	
H3(eq)	2.72			2.70	
H4	3.66	3.91	4.17	3.65	
H5	3.85	3.63	3.75	3.81	
H6	3.59	3.68/3.72	3.60/3.93	3.70	
H7	3.57			3.56	
H8	3.88			3.86	
H9	3.62/3.83			3.62/3.87	
N-CH ₃	2.01		1.99	2.01	
O-CH ₂					3.69/4.03
CH ₂					0.86/0.97
SiMe ₃					0.00

A.1.4 ^1H Chemical shift assignments for tetrasaccharide 4 at 280 K and 500 MHz

	Neu5Ac $^{\alpha 2,3}$	Gal	Gal	Neu5Ac $^{\alpha 2,6}$	Spacer
H1		4.63	4.43		
H2		3.62	3.59		
H3		4.09	3.78		
H3(ax)	1.77			1.64	
H3(eq)	2.72			2.67	
H4	3.65	3.90	4.20	3.68	
H5	3.83	3.65	3.76	3.80	
H6	3.59	3.68/3.71	3.55/3.90	3.66	
H7	3.55			3.53	
H8	3.86			3.85	
H9	3.60/3.84			3.60/3.83	
N-CH ₃	2.00			2.00	
O-CH ₂					3.74/4.01
CH ₂					0.94/1.04
SiMe ₃					0.00

A.1.5 ^1H Chemical shift assignments for pseudo tetrasaccharide 5 at 288 K and 500 MHz

	Neu5Ac ^{α2,3}	Gal	Cyc	(S)-LAc
H1		4.61	3.50/3.61	
H2		3.64	1.87	
H3		4.08	4.03	
H3(ax)	1.78			
H3(eq)	2.73			
H4	3.66	3.93	4.04	
H5	3.83	3.65	3.63	
H6	3.59	3.68/3.71	3.49/4.00	
H7	3.57			
H8	3.86			
H9	3.62/3.85			
N-CH ₃	2.01			
CH ₃				1.28
CH				3.83

A.1.6 ^1H Chemical shift assignments for trisaccharide 6 at 288 K and 500 MHz

	Neu5Ac ^{α2,6}	Gal	Gal	Spacer
H1		4.55	4.43	
H2		3.57	3.59	
H3		3.64	3.76	
H3(ax)	1.64			
H3(eq)	2.68			
H4	3.64	3.88	4.20	
H5	3.79	3.67	3.75	
H6	3.66	3.69/3.74	3.55/3.88	
H7	3.55			
H8	3.86			
H9	3.61/3.84			
N-CH ₃	2.00			
CH ₂				0.96/1.03
O-CH ₂				3.74/4.00
SiMe ₃				0.00

A.2 Pulseprogram

A.2.1 1D reference pulse sequence

```

;zgpf.mm
;avance-version
;water suppression by dpfgse sequence by A.J. Shaka
;JACS 117, 4199-4200 (1995)

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

"p2=p1*2"
"d12=10u"

1 ze
2 d1 p11:f1
  p1 ph1
  50u setnmr2|0 setnmr0|34
  p17:gp5          ; 600-800 us
  d16 p11:f1       ; 200 us  p11 is the power for the sel. rect. 180
  d12              ;use these three lines for water samples
  p3:f1 ph10:r      ;selective water rectangular 180 2-3ms long
  d12
  d12 p11:f1
  p2 ph12
  p17:gp5
  d16 p11:f1
  p17:gp6
  d16
  d12
  p3:f1 ph20:r      ;selective rectangular 180
  d12
  d12 p11:f1
  p2 ph22
  p17:gp6
  d16
  4u setnmr2^0 setnmr0^34
  go=2 ph31
  wr #0
exit

ph1=0
ph31=0 2 0 2 2 0 2 0
ph10=2 2 2 2 3 3 3 3
ph12=0 0 0 0 1 1 1 1
ph20=2 3 0 1
ph22=0 1 2 3

```

```

;*****Power Level*****
;p1 : f1 channel - power level for pulse (default)
;p11 : f1 channel - soft pulse for water suppression 44 -45 dB
;sp1 : f2 - channel - power level for shaped pulse
;
;*****Pulse*****
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f1 channel - 180 degree soft pulse on o1 (HDO)
;
;*****Delays*****
;d1 : relaxation delay; 1-5 * T1
;d16: delay for homospoil/gradient recovery
;
;NS = 16*n
;DS = 8
;use gradient ratio   gp5   :   gp6
;                      17     27

```

A.2.2 1D STD

```

;stdpfsp
;M. Mayer; B. Meyer, Department of Chemistry
;University of Hamburg, Germany
;email: bernd_meyer@sg1.chemie.uni-hamburg.de
;avance-version
;1D difference sequence with f2 presaturation defined by frequency list
;presaturation by shaped pulses
;frequency alternates after every scan, defined by fq1list
;water suppression by dpfgse sequence by A.J. Shaka, JACS, 117, 4199-4200 (1995)
;define 1H on channel f2 in edasp

```

```

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

```

```

"p2=p1*2"
"d12=10u"

```

```

1 ze
2 d7 p11:f1
  20u fq1:f2
3 p11:sp1:f2
  d11
  lo to 3 times 17
  p1 ph1
  50u setnmr2|0 setnmr0|34
  p17:gp5           ; 600-800 us
  d16 p11:f1        ; 200 us p11 is the power for the sel. rect. 180 ca. 45dB
  d12               ;use these three lines for water samples

```

```

p3:f1 ph20:r      ;selective water rectangular 180 2-3ms long
d12
d12 pl1:f1
p2 ph10
p17:gp5
d16 pl11:f1
p17:gp6
d16
d12
p3:f1 ph22:r      ;selective rectangular 180
d12
d12 pl1:f1
p2 ph8
p17:gp6
d16
4u setnmr2^0 setnmr0^34
go=2 ph31
wr #0
exit

ph1=0
ph31=0 2 0 2 2 0 2 0
ph8=0 0 0 0 3 3 3 3 2 2 2 2 1 1 1 1
ph10=0
ph20=2
ph22=2 2 2 2 1 1 1 1
      0 0 0 0 3 3 3 3

;*****Power Level*****
;
;p11 : f1 channel - power level for pulse (default)
;sp1 : f2 - channel - power level for shaped pulse
;between 50 - 60 dB depending on protein and ligand
;p110: for 3 ms pulse, ca. 44 - 45 dB
;
;*****Pulse*****
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f1 channel - 180 degree soft pulse on water ca. 2- ms
;p11 : f2 channel - presaturation shaped pulse (gauss ca. 50 msec)
;
;*****Delays*****
;d1 : relaxation delay; 1-5 * T1
;d7 : additional delay for complete T1 relaxation
;d11 : delay between shaped pulses [1msec]
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;  d19 = (1/(2*d)), d = distance of next null (in Hz)
;  d19 should be around 150-220 usec.
;
;presaturation = (p11 + d11) * 17 (presaturation should be around 2 sec)

```

```

;
;fq1 : define frequencies for on and off resonance presaturation
;   O 499.87000 off resonance 1x(15-20000 HZ) on resonance 1x(xxx HZ)
;   on frequency list fl.
;NS = 16*n
;DS = 16
;use gradient ratio   gp5   :   gp6
;                   17     27

```

A.2.3 STD TOCSY

```

;std_dipsi19_2f.ab
;avance-version
;homonuclear Hartman-Hahn transfer using DIPSI2 sequence
;  for mixing
;using sensitivity improvement
;phase sensitive using Echo/Antiecho-TPPI gradient selection
;water suppression using 3-9-19 pulse sequence with gradients
;allowing for presaturation during relaxation delay in cases
;  of radiation damping
;J. Cavanagh & M. Rance, J. Magn. Reson. 88, 72-85 (1990)
;M. Piotto, V. Saudek & V. Sklenar, J. Biomol. NMR 2, 661 - 666 (1992)
;V. Sklenar, M. Piotto, R. Leppik & V. Saudek, J. Magn. Reson.,
;  Series A 102, 241 -245 (1993)
;A.J.Benie April 2001
;STD tocsy which produces 2 files as defined by the dslist

```

```

#include <Avance.incl>
#include <Grad.incl>
#include <Delay.incl>

```

```

"p2=p1*2"
"d0=3u"
"d11=30m"
"d12=20u"
"d13=4u"
define list<frequency>fqlist=<$FQ1LIST>

```

```

"DELTA=p16+d16+d0"
"FACTOR1=(d9/(p6*115.112))/2+0.5"
"l1=FACTOR1*2"

```

```

"l3=(td1/2)"

```

```

1 ze
2 d11
   6m
3 d11

```

12m
 4 d12 pl9:f1
 d12 fq1:f1
 5 p11:sp1
 d15
 lo to 5 times l7
 d12 fq1:f1
 d12 pl1:f1
 50u UNBLKGRAD
 p1 ph1
 DELTA
 p2 ph2
 d0
 p16:gp1*EA
 d16
 p1 ph3
 d20 pl10:f1
 6 p6*3.556 ph23
 p6*4.556 ph25
 p6*3.222 ph23
 p6*3.167 ph25
 p6*0.333 ph23
 p6*2.722 ph25
 p6*4.167 ph23
 p6*2.944 ph25
 p6*4.111 ph23
 p6*3.556 ph25
 p6*4.556 ph23
 p6*3.222 ph25
 p6*3.167 ph23
 p6*0.333 ph25
 p6*2.722 ph23
 p6*4.167 ph25
 p6*2.944 ph23
 p6*4.111 ph25
 p6*3.556 ph25
 p6*4.556 ph23
 p6*3.222 ph25
 p6*3.167 ph23
 p6*0.333 ph25
 p6*2.722 ph23
 p6*4.167 ph25
 p6*2.944 ph23
 p6*4.111 ph25
 p6*3.556 ph23
 p6*4.556 ph25
 p6*3.222 ph23
 p6*3.167 ph25
 p6*0.333 ph23
 p6*2.722 ph25

p6*4.167 ph23
 p6*2.944 ph25
 p6*4.111 ph23
 lo to 6 times l1
 d21 pl1:f1
 p1 ph4
 20u pl18:f1
 p16:gp2
 d16
 p28*0.231 ph5
 d19*2
 p28*0.692 ph5
 d19*2
 p28*1.462 ph5
 d19*2
 p28*1.462 ph6
 d19*2
 p28*0.692 ph6
 d19*2
 p0*0.231 ph6
 l6u
 p16:gp3
 d16
 4u BLKGRAD
 go=2 ph31
 d11 wr #1 if #1 zd
 goto 8
 7 d11
 8 6m
 d11
 l2m
 d12 pl9:f1
 d12 fq2:f1
 9 p11:sp1
 d15
 lo to 9 times l7
 d12 fq2:f1
 d12 pl1:f1
 50u UNBLKGRAD
 p1 ph1
 DELTA
 p2 ph2
 d0
 p16:gp1*EA
 d16
 p1 ph3
 d20 pl10:f1
 a,p6*3.556 ph23
 p6*4.556 ph25
 p6*3.222 ph23

p6*3.167 ph25
 p6*0.333 ph23
 p6*2.722 ph25
 p6*4.167 ph23
 p6*2.944 ph25
 p6*4.111 ph23
 p6*3.556 ph25
 p6*4.556 ph23
 p6*3.222 ph25
 p6*3.167 ph23
 p6*0.333 ph25
 p6*2.722 ph23
 p6*4.167 ph25
 p6*2.944 ph23
 p6*4.111 ph25
 p6*3.556 ph25
 p6*4.556 ph23
 p6*3.222 ph25
 p6*3.167 ph23
 p6*0.333 ph25
 p6*2.722 ph23
 p6*4.167 ph25
 p6*2.944 ph23
 p6*4.111 ph25
 p6*3.556 ph23
 p6*4.556 ph25
 p6*3.222 ph23
 p6*3.167 ph25
 p6*0.333 ph23
 p6*2.722 ph25
 p6*4.167 ph23
 p6*2.944 ph25
 p6*4.111 ph23
 lo to a times l1
 d21 pl1:fl
 p1 ph4
 20u pl18:fl
 p16:gp2
 d16
 p28*0.231 ph5
 d19*2
 p28*0.692 ph5
 d19*2
 p28*1.462 ph5
 d19*2
 p28*1.462 ph6
 d19*2
 p28*0.692 ph6
 d19*2
 p0*0.231 ph6

```

16u
p16:gp3
d16
4u BLKGRAD
go=7 ph31
d11 wr #2 if #2 zd
6m ip3*2 igrad EA
lo to 3 times 2
d11 id0
6m ip1*2
6m ip31*2
lo to 4 times l3
exit

ph1=0 2
ph2=0
ph3=0 0 2 2
ph4=2 2 0 0
ph5=0
ph6=2
ph23=3
ph25=1
ph29=0
ph31=0 2

;*****Power level*****
;p11 : f1 channel - power level for pulse (default)
;p19 : f1 channel - power level for std [50 dB]
;p110: f1 channel - power level for TOCSY-spinlock [9 dB]
;p118: f1 channel - power level for 3-9-19-pulse (watergate)
;
;***** Pulse*****
;p0 : f1 channel - 90 degree pulse at p118
; use for fine adjustment
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p6 : f1 channel - 90 degree low power pulse
;p11: f1 channel - 90 degree shaped pulse for STD
;p16: homospoil/gradient pulse
;p28: f1 channel - 90 degree pulse at p118
;sp1: f1 channel - 90 degree shaped pulse power level [50dB]
;
;*****Delays*****
;d0 : incremented delay (2D) [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d9 : TOCSY mixing time [80 msec]
;d11: delay for disk I/O [30 msec]
;d12: delay for power switching [20 usec]
;d13: short delay [4 usec]
;d15: delay between STD pulses [1 msec]

```



```
;d16: delay for homospoil/gradient recovery
;d19: delay for binomial water suppression
;   d19 = (1/(2*d)), d = distance of next null (in Hz)
;d20: first z-filter delay [10 usec]
;d21: second z-filter delay [10 usec]
;l1: loop for DIPSI cycle: ((p6*115.112) * l1) = mixing time
;l3: loop for phase sensitive 2D using E/A method : l3 = td1/2
;l7: loop for STD saturation period
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;NS: 8 * n
;DS: 16
;td1: number of experiments
;MC2: echo-antiecho

;use gradient ratio:   gp 1 : gp 2 : gp 3
;                      30 : 20 : 50

;for z-only gradients:
;gpz1: 30%
;gpz2: 20%
;gpz3: 50%

;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
;gpnam3: SINE.100

;set pl9 to between 30-70dB for STD
;fq1list : define on resonance frequency and water
;fq2list : define off resonance frequency and water
```

A.2.4 NOESY/trNOESY

```
;noesygp_h
;avance-version (00/02/07)
;2D homonuclear correlation via dipolar coupling
;dipolar coupling may be due to noe or chemical exchange.
;phase sensitive
;with gradient pulses in mixing time
;J. Jeener, B.H. Meier, P. Bachmann & R.R. Ernst, J. Chem. Phys. 71,
; 4546-4553 (1979)
;R. Wagner & S. Berger, J. Magn. Reson. 123 A, 119-121 (1996)

#include <Avance.incl>
#include <Grad.incl>
```

```
"d0=3u"
"d20=d8*0.5-p16-d16-50u"
```

```
1 ze
2 d11
  d1 pl9:f1
  d9 cw:f1 ph29
  3u do:f1
  d12 pl1:f1
  p1 ph1
  d12 pl11:f1
  p15 ph6
  d0
  d12 pl1:f1
  p1 ph2
  d20 UNBLKGRAD
  p16:gp1
  d16
  3u
  (p2 ph4):f1
  3u
  p16:gp2
  d16
  d20 BLKGRAD
  p1 ph3
  go=2 ph31
  d11 mc #0 to 2 F1PH(ip1 & ip6, id0)
exit
```

```
ph1=0 2
ph2=0 0 0 0 0 0 0 2 2 2 2 2 2 2
ph3=0 0 2 2 1 1 3 3
ph4=0
ph6=1 3
ph29=0
ph31=0 2 2 0 1 3 3 1 2 0 0 2 3 1 1 3
```

```
*****Power level*****
;
;p11 : f1 channel - power level for pulse (default)
;p111: f1 channel - power level for T1rho filter
;
*****Pulse*****
;p1 : f1 channel - 90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p15: spin lock (T1rho) pulse to relax protein during d0
;p16: homospoil/gradient pulse [1 msec]
;
*****Delays*****
;d0 : incremented delay (2D) [3 usec]
```

```
;d1 : relaxation delay; 1-5 * T1
;d9 : short delay [~1.8 s]
;d11: delay for disk I/O [30 msec]
;      ;d12: delay for power switching [20 usec]
;d8 : mixing time
;d16: delay for homospoil/gradient recovery
;d20: d8*0.5 - p16 - d16
;in0: 1/(1 * SW) = 2 * DW
;nd0: 1
;NS: 2 * n
;DS: 16
;td1: number of experiments
;FnMODE: States-TPPI, TPPI, States or QSEC
```

```
;use gradient ratio: gp 1 : gp 2
;                      40 : -40
```

```
;for z-only gradients:
;gpz1: 40%
;gpz2: -40%
```

```
;use gradient files:
;gpnam1: SINE.100
;gpnam2: SINE.100
```

A.3 CORCEMA

A.3.1 CORCEMA “command” file

```
#
stage 1
output complex_loop1_sggs_out1.txt
num_states 2 ;
state 1 "Unbound Complex"
num_groups 2
group 1 "Ligand atoms"
input "complex_loop1_sggs.pdb" resnum = 1
correlation_time 0.5ns
isolate yes
exchange_time 0 67e-02s
end_group
group 2 "Enzyme atoms"
input "complex_loop1_sggs.pdb" resnum <> 1
isolate yes
correlation_time 40.0ns
exchange_time 0 25e-03s
```

```

    end_group
end_state
state 2 "Bound Complex"
    num_groups 1
    group 1 "Complex atoms"
        input "complex_loop1_sggs.pdb"
        correlation_time 40.0ns
        exchange_time 1e-01s 0
    end_group
end_state
end_stage

stage 2
    output complex_loop1_sggs_out2.txt
    frequency 500.25Mhz
    experiment NOESY
    methyl_average yes
    leakage_factor 0.2is
end_stage

stage 3
    output complex_loop1_sggs_out3.txt ; this is the final output file with data.
    mixing_time 50ms 75ms 100ms 200ms 300ms 500ms 750ms
    atom_list tetra_anir.txt exc
    reference_peak " H3 GAL   1.1 " " H3 GAL   1.1 " 100.0
    reference_peak " H3 GAL   1.2 " " H3 GAL   1.2 " 100.0
    reference_peak " H3 GAL   1.1 " " H3 GAL   1.2 " 100.0
    reference_peak " H3 GAL   1.2 " " H3 GAL   1.1 " 100.0
end_stage

```

A.3.1.1 Compilation for Mac OS X 10.3

;edited by Dr. A. J. Benie March 2005

STEP 1: MODIFICATION OF SOURCE CODE

Add the following to corcema17.c (line 100)

```
#include <Accelerate/Accelerate.h>
```

Change the following lines

from (line 475)

```
fprintf(stderr," For help type \"relax -h\"\\n");
```

to

```
fprintf(stderr," For help type \"corcema17 -h\"\\n");
```

from (line 487)

```
fprintf(stderr," For help type \"relax -h\"\\n");
```

to

```
fprintf(stderr," For help type \"corcema17 -h\"\\n");
```

Add the following to metric.h (line 25)

```
#include <float.h>
```

Change the following in matrix2.h

from (line 30)

```
#include <bstring>
```

to

```
/* #include <bstring.h> */
```

from (lines 37-38)

```
int ncol;
```

```
int nrow;
```

to

```
long int ncol;
```

```
long int nrow;
```

STEP 2: COMPIATION OF SOURCE CODE

Compile by typing for G4 and G5 processors

```
cc -faltivec -framework Accelerate corcema17.c -o corcema17
```

STEP 3: INSTALLATION

copy the files "corcema17" and "corcema17.hlp" to wherever you want them then add the following to your .cshrc file.

```
setenv CORCEMADIR /PATH/TO/INSTALLATION
```

or alternatively copy "corcema17" to /usr/local/bin/ and "corcema17.hlp" to /usr/local/doc

A.3.2 CORCEMA atom list file

Num_atoms 11

```
atom " H23GAL  1" atom " H25GAL  1"
atom " H23GAL  1" atom " H3  GAL  1"
atom " H23GAL  1" atom " H36GAL  1"
atom " H3  GAL  1" atom " H33GAL  1"
atom " H3  GAL  1" atom " H5  GAL  1"
atom " H25GAL  1" atom " H50SIA  1"
atom " H25GAL  1" atom " H51SIA  1"
atom " H25GAL  1" atom " H56SIA  1"
atom " H3  GAL  1" atom " H28GAL  1"
atom " H28GAL  1" atom " H15SIA  1"
atom " H29GAL  1" atom " H15SIA  1"
```

A.3.3 Comparison between experimental and theoretical trNOEs at different mixing times (ms)

A.3.3.1 Trisaccharide 2 – MAG complex

	Experimental (% NOE)						Theoretical (% NOE)					
	50	75	100	200	300	500	50	75	100	200	300	500
H1-H3	1.5	2.6	2.8	4.8	5.8	6.0	2.0	2.6	3.2	5.0	6.1	6.8
H1'-H3'	1.6	2.3	2.6	4.2	5.2	4.7	2.2	2.8	3.3	4.9	5.7	5.9
H3'-H4'	2.0	3.2	3.5	5.6	6.6	6.8	2.4	3.1	3.7	5.5	6.5	6.8
H1'-H3	2.9	3.7	4.5	6.6	7.4	6.7	2.3	3.0	3.5	5.0	5.7	5.8
H1'-H4	0.4	0.7	n.d.	1.7	1.9	2.4	0.6	0.8	1.0	1.6	2.0	2.3
H3'-H8''	n.d.	1.2	n.d.	1.4	2.0	2.9	0.4	0.5	0.6	1.0	1.3	1.6
H3'-H3ax''	n.o.	n.d.	n.d.	0.9	1.0	1.2	0.1	0.1	0.2	0.3	0.3	0.4

A.3.3.2 Tetrasaccharide 4 – MAG complex

	Experimental (% NOE)						Theoretical (% NOE)					
	50	75	100	200	300	500	50	75	100	200	300	500
H1''-H3''	0.8	2.2	3.2	5.1	5.4	5.3	3.7	4.1	4.5	5.6	6.0	6.0
H4'-H6'b	1.2	1.8	3.1	4.2	4.4	4.4	3.3	3.7	4.0	4.6	4.6	3.9
H4'-H6'a	0.6	1.7	2.8	3.9	4.1	4.1	2.6	3.0	3.3	4.0	4.2	3.8
H1''-H3'	2.1	3.4	4.0	5.3	6.0	6.0	3.5	3.8	4.1	4.9	5.2	4.9
H1''-H4'	n.o.	0.6	1.2	1.9	2.1	2.7	1.0	1.3	1.5	2.0	2.2	2.4
H4'-H6''	0.1	0.5	0.7	1.6	1.8	1.7	0.6	0.9	1.0	1.3	1.4	1.5
H3''-H3ax'''	0.3	0.6	0.7	0.9	1.3	1.1	0.1	0.2	0.2	0.4	0.5	0.7
H3''-H8'''	0.3	0.8	1.4	2.9	3.2	3.2	2.5	2.9	3.2	3.9	4.3	4.4

A.3.3.3 Pseudo tetrasaccharide 5 – MAG complex

	Experimental (% NOE)						Theoretical (% NOE)					
	50	75	100	200	300	500	50	75	100	200	300	500
H1''-H3''	1.7	3.2	4.7	6.3	6.5	6.4	4.2	4.8	5.1	6.2	6.7	6.4
H1''-H5''	1.5	2.7	3.9	5.7	6.5	6.8	4.3	4.7	5.1	6.1	6.5	6.1
H6'b-CH ^{LAc}	0.6	1.7	3.2	4.6	4.6	4.7	3.9	4.5	4.9	5.9	6.0	5.1
H1''-(H3'+H4')	2.3	5.2	5.9	8.7	9.5	9.7	4.9	5.4	5.9	7.1	7.6	7.2
H1''-(H2ax'+H2eq')	0.2	0.4	0.5	0.9	1.3	1.1	0.9	1.2	1.4	1.9	2.1	2.0
H3''-H8'''	0.3	0.7	1.2	2.3	2.5	2.5	0.6	0.8	0.9	1.2	1.4	1.7

A.3.3.4 Trisaccharide 2 – sialoadhesin complex

	Experimental (% NOE)						Theoretical (% NOE)					
	50	75	100	150	250	500	50	75	100	150	250	500
H1-H3	2.5	3.5	3.7	5.5	7.3	9.9	2.2	3.1	3.9	5.4	7.5	9.9
H'-H3'	4.4	6.4	7.2	10.4	13.4	17.5	2.6	3.7	4.7	6.3	8.5	10.3
H1'-H3	3.9	4.9	6.8	9.5	12.8	16.5	3.0	4.2	5.3	7.0	9.2	10.3
H1'-H4	1.0	1.3	1.4	1.9	2.5	n.d.	0.8	1.1	1.5	2.1	2.9	4.0
H3'-H8''	0.9	1.2	1.7	1.9	1.6	0.4	0.2	0.3	0.4	0.6	0.9	1.2

Bibilography

- Allinger, N. L., Yuh, Y. H. and Lii, J. -H. (1989), Molecular mechanics. The MM3 Force Field for Hydrocarbons. 1. *J. Am. Chem. Soc.*, **111**, 8551-8566.
- Apostolski, S., Sadiq, S. A., Hays, A., Carbo, M., Suturkova-Milosevic, L., Chaliff, P., Stefansson, K., LeBaron, R. G., Ruoslahti, E. and Hays, A. P. (1994), Identification of Gal(beta 1-3)GalNAc bearing glycoproteins at the nodes of Ranvier in peripheral nerve. *J. Neurosci. Res.*, **38** (2), 134-141.
- Arepalli, S. R., Glaudemans, C. P. J., Daves Jr., G. D., Kovák, P. and Bax, A. (1995), Identification of protein-mediated indirect NOE effects in a disaccharide-Fab complex by transferred ROESY. *J. Magn. Reson. B*, **106**, 195-198.
- Balaram, P., Bothner-By, A. A. and Dadok, J. (1972), Negative nuclear overhauser effects as probes of macromolecular structure. *J. Am. Chem. Soc.*, **94** (11), 4015-4017.
- Bax, A. and Davids, D. G. (1985), Practical aspects of two-dimensional transverse NOE spectroscopy. *J. Magn. Reson.*, **63**, 207-213.
- Bhunia, A., Jayalakshmi, V., Benie, A. J., Schuster, O., Kelm, S., Krishna, N. R. and Peters, T. (2004), Saturation transfer difference NMR and computational modelling of a sialoadhesin-sialyl lactose complex. *Carbohydr. Res.*, **339**, 259-267.
- Biet, T. (2003), Neue Erkenntnisse über den Mechanismus von Galactosyltransferasen mit der STD-NMR-Spektroskopie. Ph.D. Thesis, Universität zu Lübeck.
- Blundell, T., Carney, D., Gardner, S., Hayes, F., Howlin, B., Hubbard, T., Overington, J., Singh, D. A., Sibanda, B. L. and Sutcliffe, M. (1988), 18th Sir Hans Krebs lecture. Knowledge-based protein modelling and design. *Eur. J. Biochem.*, **172**, 513-520.
- Bothner-By, A. A., Stephens, R. L. and Lee, J. -M. (1984), Structure determination of a tetrasaccharide: Transient nuclear overhauser effects in the rotating frame. *J. Am. Chem. Soc.*, **106**, 811-813.
- Braunschweiler, L. and Ernst, R. R. (1983), Coherence transfers by isotropic mixing: Application to proton correlation spectroscopy. *J. Magn. Reson.*, **53**, 521-528.
- Breg, J., Kroon-Batenburg, L. M. J., Strecker, G., Montreuil, J. and Vliegthart, J. F. G. (1989), Conformational Analysis of the Sialyl α -(2-3/6)N-acetylglucosamine Structural Element Occurring in Glycoproteins, by Two-dimensional NOE ¹H NMR spectroscopy in combination with Energy Calculations by Hard-Sphere Exo-Anomeric and molecular Mechanics Force-Field with Hydrogen-Bonding Potential. *Eur. J. Biochem.*, **178**, 727-739.

- Brooks, B. R., Brucoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. and Karplus, M. J. (1983), CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.*, **4**, 187-217.
- Brummendorf, T. and Rathjen, F. G. (1994), Cell adhesion molecules. 1: immunoglobulin superfamily. *Protein Profile*, **1** (9), 951-1058.
- Bukowski, R. (2001), NMR- und SPR-Untersuchungen zur biologischen Erkennung von Galactose-haltigen Sacchariden durch lektine. Ph.D. Thesis, Universität zu Lübeck.
- Bukowski, R., Morris, L. M., Woods, R. J. and Weimar, T. (2001), Synthesis and Conformational Analysis of the T-antigen Disaccharide (b-D-Gal-(1→3)-a-D-GalNAc-OMe). *Eur. J. Org. Chem.*, 2697-2705.
- Caffarena, E. R. and Bisch, P. M. (2000), Hydration of T-antigen Galβ(1-3)GalNAc and the isomer Galβ(1-3)GlcNAc by molecular dynamics simulations. *J. Mol. Graph. Model.*, **18**, 119-125.
- Casset, F., Peters, T., Etzler, M., Korchagina, E., Nifant'ev, N., Perez, S. and Imberty, A. (1996), Conformational analysis of blood group A trisaccharide in solution and in the binding site of Dolichos biflorus lectin using transient and transferred nuclear Overhauser enhancement (NOE) and rotating-frame NOE experiments. *Eur. J. Biochem.*, **239** (3), 710-719.
- Chen, A. and Shapiro, M. J. (1998), NOE pumping: A novel NMR technique for identification of compounds with binding affinity to macromolecules. *J. Am. Chem. Soc.*, **120**, 10258-10259.
- Chen, A. and Shapiro, M. J. (2000), NOE pumping 2: A high-throughput method to determine compounds with binding affinity to macromolecules by NMR. *J. Am. Chem. Soc.*, **122**, 414-415.
- Chen, M. S., Huber, A. B., van der Haar, M., Frank, M., Schnell, L., Spillmann, A. A., Christ, F. and Schwab, M. E. (2000), Nogo-A is a myelin-associated neurite outgrowth inhibitor and an antigen for monoclonal antibody IN-1. *Nature*, **403**, 434-438.
- Clark, M., Cramer-III, R. D. and van Opdenbosch, N. (1989), Validation of the general purpose tripos 5.2 force field. *J. Comp. Chem.*, **10**, 982-1012.
- Clore, G. M. and Gronenborn, A. M. (1982), Theory and applications of the transfer-nuclear overhauser effect to the study of the conformations of small ligands bound to proteins. *J. Magn. Reson.*, **48**, 402-417.
- Collins, B. E., Kiso, M., Hasegawa, A., Tropak, M. B., Roder, J. C., Crocker, P. R. and Schnaar, R. L. (1997), Binding Specificities of the Sialoadhesin Family of I-type Lectins. Sialic acid linkage and substructure requirements for binding of myelin-associated glycoprotein, schwann cell myelin protein, and sialoadhesin. *J. Biol. Chem.*, **272**, 16889-16895.

- Collins, B. E., Ito, H., Sawada, N., Ishida, H., Kiso, M. and Schnaar, R. L. (1999), Enhanced binding of the neural siglecs, myelin-associated glycoprotein and schwann cell myelin protein, to Chol-1 (α -series) gangliosides and novel sulphated Chol-1 analogs. *J. Biol. Chem.*, **274** (53), 37637-37643.
- Craik, D. J. and Scanlon, M. J. (2000), Pharmaceutical applications of NMR. *Annu. Rep. NMR Spectrosc.*, **42**, 115-174.
- Crocker, P. R., Kelm, S., Dubois, C., Martin, B., McWilliam, A. S., Shotton, D. M., Paulson, J. C. and Gordon, S. (1991a), Purification and properties of sialiadhesin, a sialic acid-binding receptor of murine tissue macrophages. *EMBO J.*, **10**, 1661-1669.
- Crocker, P. R., Werb, Z., Gordon, S. and Bainton, D. F. (1991b), Ultra structural localization of a macrophage-restricted sialic acid binding hemagglutinin, SER, in macrophage-hematopoietic cell clusters. *Blood*, **76**, 1131-1138.
- Crocker, P. R. and Feizi, T. (1996), Carbohydrate recognition systems: functional triads in cell-cell interactions. *Curr. Opin. Struct. Biol.*, **6** (5), 679-691.
- Crocker, P. R., Hartnell, A., Munday, J. and Nath, D. (1997), The potential role of sialoadhesin as a macrophage recognition molecule in health and disease. *Glycoconj. J.*, **14** (5), 601-609.
- Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le Douarin, N., Powell, L., Roder, J., Schnaar, R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., van den Berg, T. K., van den Merwe, P. A., Watt, S. M. and Varki, A. (1998), Siglecs: a family of sialic-acid binding lectins. *Glycobiol.*, **8** (2), v.
- Crocker, P. R., Vinson, M., Kelm, S. and Drickamer, K. (1999), Molecular analysis of sialoside binding to sialoadhesin by NMR and site-directed mutagenesis. *Biochem. J.*, **341**, 355-361.
- Crocker, P. R. and Varki, A. (2001a), Siglecs in the immune system. *Immunol.*, **103** (2), 137-145.
- Crocker, P. R. and Varki, A. (2001b), Siglecs, sialic acid and innate immunity. *Trends Immunol.*, **22** (6), 337-342.
- Crocker, P. R. (2002), Siglecs: sialic acid-binding immunoglobulin like lectins in cell-cell interactions and signalling. *Curr. Opin. Struct. Biol.*, **12** (5), 609-615.
- Cumming, D. A. and Carver, J. P. (1987), Virtual and Solution Conformations of Oligosaccharides. *Biochemistry*, **26**, 6664-6676.
- Curto, E. V., Moseley, H. N. B. and Krishna, N. R. (1996), CORCEMA evaluation of the potential role of intermolecular transferred NOESY in the characterization of ligand-receptor complexes. *J. Comput. Aided Mol. -Des.*, **10**, 361-371.

- Dalvit, C., Pevarello, P., Tato, M., Veronesi, M., Vulpetti, A. and Sundström, M. (2000), Identification of compounds with binding affinity to proteins via magnetization transfer from bulk water. *J. Biomol. NMR*, **18**, 65-68.
- Del Poeta, G., Stasi, R., Venditti, A., Suppo, G., Aronica, G., Bruno, A., Masi, M., Tabilio, A. and Papa, G. (1994), Prognostic value of cell marker analysis in *de novo* acute myeloid leukemia. *Leukemia*, **8** (3), 388-394.
- Dergham, P., Ellezam, B., Essagian, C., Avedissian, H., Lubell, W. D. and McKerracher, L. (2002), Rho Signalling Pathway Targeted to promote Spinal Cord Repair. *J. Neurosci.*, **22**, 6570-6577.
- Diercks, T., Coles, M. and Kessler, H. (2001), Applications of NMR in drug discovery. *Curr. Opin. Chem. Biol.*, **5**, 285-291.
- Domeniconi, M., Cao, Z., Spencer, T., Sivasankaran, R., Wang, K. C., Nikulina, E., Kimura, N., Cai, H., Deng, K. and Gao, Y. (2002), Myelin-associated glycoprotein interacts with the Nogo 66 receptor to inhibit neurite outgrowth. *Neuron*, **35**, 283-290.
- Dorland, L., van Halbeek, H., Vliegenthart, J. F., Schauer, R. and Wiegandt, H. (1986), A 500-MHz ¹H-n.m.r. study of oligosaccharides derived from gangliosides by ozonolysis-alkaline fragmentation. *Carbohydr. Res.*, **151**, 233-245.
- Dulac, C., Tropak, M. B., Cameron-Curry, P., Rossier, J., Marshak, D. R., Roder, J. and Le Douarin, N. M. (1992), Molecular characterisation of the Schwann cell myelin protein, SMP: structural similarities within the immunoglobulin superfamily. *Neuron*, **8** (2), 323-334.
- Feng, M-H., Philippopoulos, M., MacKerell, A. D. and Lim, C. (1996), Structural characterization of the phosphotyrosine binding region of a high-affinity SH2 domain-phosphopeptide complex by molecular dynamics simulation and chemical shift calculations. *J. Am. Chem. Soc.*, **118** (45), 11265-11277.
- Feng, Ni. (1994), Recent developments in transferred NOE methods. *Prog. In NMR Spectrosc.*, **26**, 517-606.
- Filbin, M. T. (2003), Myelin-associated inhibitors of axonal regeneration in the adult mammalian CNS. *Nat. Rev. Neurosci.*, **4**, 703-713.
- Fournier, A. E., GrandPre, T. and Strittmatter, S. M. (2001), Identification of a receptor mediating Nogo-66 inhibition of axonal regeneration. *Nature*, **409**, 341-346.
- Freeman, S. D., Kelm, S., Barber, E. K. and Crocker, P. R. (1995), Characterisation of CD33 as a new member of sialoadhesin family of cellular interaction molecules. *Blood*, **85**, 2005-2012.
- Fry, D. C. and Emerson, S. D. (2000), Applications of biomolecular NMR to drug discovery. *Drug Des. Discovery*, **17**, 13-33.

- Gargaro, A. R., Frenkiel, T.A., Nieto, P.M., Birdsall, B., Polshakov, V.I., Morgan, W.D. and Feeney, J. (1996), NMR detection of arginine-ligand interactions in complexes of *Lactobacillus casei* dihydrofolate reductase. *Eur. J. Biochem.*, **238**, 435-439.
- Gäthje, H. Universität Bremen, personal communication.
- GrandPre, T., Nakamura, F., Vartanian, T. and Strittmatter, S. M. (2000), Identification of the Nogo inhibitor of axon regeneration as a Reticulon protein. *Nature*, **403**, 439-444.
- Glaudemans, C. P. J., Lerner, L., Daves Jr., G. D., Kovác, P., Venable, R. and Bax, A. (1990), Significant conformational changes in an antigenic carbohydrate epitope upon binding to a monoclonal antibody. *Biochemistry*, **29**, 10906-10911.
- Habib, A. A., Marton, L. S., Allwardt, B., Gulcher, J. R., Mikol, D. D., Hognason, T., Chattopadhyay, N. and Stefansson, K. (1998), Expression of the oligodendrocyte-myelin glycoprotein by neurons in the mouse central nervous system. *J. Neurochem.*, **70** (4), 1704-1711.
- Hajduk, P. J., Meadows, R. P. and Fesik, S. W. (1997a), Discovering high-affinity ligands for proteins. *Science*, **278**, 497-499.
- Hajduk, P. J., Olejniczak, E. T. and Fesik, S. W. (1997b), One-dimensional relaxation- and diffusion edited NMR methods for screening compounds that bind to macromolecules. *J. Am. Chem. Soc.*, **119**, 12257-12261.
- Hajduk, P. J., Meadows, R. P. and Fesik, S. W. (1999), NMR-based screening in drug discovery. *Q. Rev. Biophys.*, **32**, 211-240.
- Hajduk, P. J., Augeri, D. J., Mack, J., Mendoza, R., Yang, J., Betz, S. F. and Fesik, S. W. (2000), NMR based screening of proteins containing ¹³C-labelled methyl groups. *J. Am. Chem. Soc.*, **122**, 7898-7904.
- Hakomori, S. and Zhang, Y. (1997), Glycosphingolipid antigens and cancer therapy. *Chem. Biol.*, **4** (2), 97-104.
- Hartnell, A., Steel, J., Turley, H., Jones, M., Jackson, D. G. and Crocker, P. R. (2001), Characterisation of human sialoadhesin, a sialic acid binding receptor expressed by resident and inflammatory macrophage populations. *Blood*, **97** (1), 288-296.
- Haselhorst, T., Espinosa, J. -F., Jiménez-Barbero, J., Sokolowski, T., Kosma, P., Brade, H., Brade, L. and Peters, T. (1999), NMR-Experiments Reveal Distinct Antibody-Bound Conformations of a Synthetic Disaccharide Representing a General Structural Element of Bacterial Lipopolysaccharide Epitopes. *Biochemistry*, **38**, 6449-6459.
- Haselhost, T., Weimar, T. and Peters, T. (2001), Molecular recognition of sialyl Lewis(x) and related saccharides by two lectins. *J. Am. Chem. Soc.*, **123** (43), 10705-10714.

- Herfurth, L., Ernst, B., Wagner, B., Ricklin, D., Strasser, D. S., Magnani, J. L., Benie, A. J. and Peters, T. (2005), Comprative epitope mapping with STD NMR of sialyl lewis^a and derivatives bound to monoclonal antibody. *Submitted for publication*.
- Hricovíni, M., Guerrini, M. and Bisio, A. (1999), Structure of heparin-derived tetrasaccharide complexed to the plasma protein antithrombin derived from NOEs, J-couplings and chemical shifts. *Eur. J. Biochem.*, **261**, 789-801.
- Hricovíni, M., Guerrini, M., Bisio, A., Torri, G., Petitou, M. and Casu, B. (2001), Conformation of heparin pentasaccharide bound to antithrombin III. *Biochem J.*, **359**, 265-72.
- Huber, A. B., Weinmann, O., Brosamie, C., Oertle, T. and Schwab, M. E. (2002), Patterns of Nogo mRNA and protein expression in the developing and adult rat and after CNS lesions. *J. Neurosci.*, **22**, 3553-3567.
- Hwang, T. -L. and Shaka, A. J. (1992), Cross relaxation without TOCSY: Transverse Rotating Frame Overhauser Effect Spectroscopy. *J. Am. Chem. Soc.*, **114**, 3157-3159.
- Jayalakshmi, V. and Krishna, N. R. (2002), Complete relaxation and conformational exchange matrix (CORCEMA) analysis of intermolecular saturation transfer difference effects in reversibly forming ligand-receptor complexes. *J. Magn. Reson.*, **155**, 106-118.
- Jeener, J., Meier, B. H., Bachmann, P. and Ernst, R. R. (1979), Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.*, **71**, 4546-4563.
- Jiménez-Barbero, J., Asensio, J. L., Cañada, F. J. and Poveda, A. (1999), Free and protein-bound carbohydrate structures. *Curr. Opin. Struct. Biol.*, **9**, 549-555.
- Johnson, P. W., Abramow-Newerly, W., Seilheimer, B., Sadoul, R., Tropak, M. B., Arquint, M., Dunn, R. J., Schachner, M. and Roder, J. C. (1989), Recombinant myelin-associated glycoprotein confers neural adhesion and neurite outgrowth function. *Neuron*, **3**, 377-385.
- Josephson, A., Trifunovski, A., Widmer, H. R., Widenfalk, J., Olson, L. and Spenger, C. (2002), Nogo-receptor gene activity: cellular localization and developmental regulation of mRNA in mice and humans. *J. Comp. Neurol.*, **453** (3), 292-304.
- Karlsson, K. A. (1995), Microbial recognition of target-cell glycoconjugates. *Curr. Opin. Struct. Biol.*, **5**, 622-635.
- Kaplan, M. R., Meyer-Franke, A., Lambert, S., Bennett, V., Duncan, I. D., Levinson, S. R. and Barres, B. A. (1997), Induction of sodium channel clustering by oligodendrocytes. *Nature*, **386**, 724-728.
- Kelm, S., Pelz, A., Schauer, R., Filbin, M., Tang, S., Bellard, M. -E. D., Schnaar, R. L., Mahoney, J. A., Hartnell, A., Bradfield, P. and Crocker, P. R. (1994a), Sialoadhesin, myelin-associated glycoprotein and CD22 define a new family of

- sialic acid-dependent adhesion molecules of the immunoglobulin superfamily. *Curr. Biol.*, **4**, 965-972.
- Kelm, S., Schauer, R., Manuguerra, J. C., Gross, H. J. and Crocker, P. R. (1994b), Modifications of cell surface sialic acids modulate cell adhesion mediated by sialoadhesin and CD22. *Glycoconj. J.*, **11**, 576-585.
- Kelm, S., Schauer, R. and Crocker, P. R. (1996), The Sialoadhesins- a family of sialic acid-dependent cellular recognition molecules within the immunoglobulin superfamily. *Glycoconj. J.*, **13 (6)**, 913-926.
- Kooistra, O., Herfurth, L., Lüneburg, E., Frosch, M., Peters, T. and Zähringer, U. (2002), Epitope mapping of the O-chain polysaccharide of *Legionella pneumophila* serogroup 1 lipopolysaccharide by saturation-transfer-difference NMR spectroscopy. *Eur. J. Biochem.*, **269**, 573-582.
- Knapp, W., Strobl, H. and Majdic, O. (1994), Flow cytometric analysis of cell-surface and intracellular antigens in leukaemia diagnosis. *Cytometry*, **18 (4)**, 187-198.
- Krishna, N. R., Agresti, D. G., Glickson, J. D. and Walter, R. (1978), Solution conformation of peptides by the intramolecular nuclear Overhauser effect experiment. Study of valinomycin-K⁺. *Biophys. J.*, **24 (3)**, 791-814.
- Krishna, N. R. and Moseley, H. N. B., (1999), Complete relaxation and conformational exchange matrix (CORCEMA) analysis of NOESY spectra of reversibly forming ligand-receptor complexes. In "Biological Magnetic Resonance, vol. 17, Structure Computation and Dynamics in Protein NMR", Kluwer Academic/Plenum Press, New York, pp. 223-310.
- Kursula, P., Merilainen, G., Lehto, V. P. and Heape, A. M. (1999), The small myelin-associated glycoprotein is a zinc-binding protein. *J. Neurochem.*, **73 (5)**, 2110-2118.
- Kursula, P., Lehto, V. P. and Heape, A. M. (2001), The small myelin-associated glycoprotein binds to tubulin and microtubules. *Brain Res. Mol. Brain Res.*, **87 (1)**, 22-30.
- Lai, C., Brow, M. A., Nave, K. A., Noronha, A. B., Quarles, R. H., Bloom, F. E., Milner, R. J., and Sutcliffe, J. G. (1987), Two forms of IB236/myelin-associated glycoprotein, a cell adhesion molecule for postnatal neural development, are produced by alternative splicing. *Proc. Natl. Acad. Sci. USA*, **84**, 4337-4341.
- Laskowski, R. A., MacArthur, M. W., Moss, D. S. and Thornton, J. M. (1993), PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, **26**, 283-291.
- Leach, A. L. (1996), Molecular Modelling, Principles and Applications. Addison Wesley Longman Limited.

- Lehmann, M., Fournier, A., Selles-Navarr, I., Dergham, P., Sebok, A., Leclerc, N., Tigyi, G. and McKerracher, L. (1999), Inactivation of Rho signalling pathway promotes CNS axon regeneration. *J. Neurosci.*, **19** (17), 7537-7547.
- Liu, B. P., Fournier, A. E., GrandPre, T. and Strittmatter, S. M. (2002), Myelin-associated glycoprotein as a functional ligand for the Nogo-66 receptor. *Science*, **297**, 1190-1193.
- Maaheimo, H., Kosma, P., Brade, L., Brade, H. and Peters, T. (2000), Mapping the Binding Synthetic Disaccharides Representing Epitopes of Chlamydial Lipopolysaccharide to Antibodies with NMR. *Biochemistry*, **39**, 12778-12788.
- Marion, D. and Wüthrich, K. (1983), Application of phase sensitive two-dimensional correlated spectroscopy for measurement of ^1H - ^1H coupling constants in proteins. *Biochem. Biophys. Res. Commun.*, **113**, 967-974.
- May, A. P., Robinson, R. C., Vinson, M., Crocker, P. R. and Jones, E. Y. (1998), Crystal structure of the N-terminal domain of sialoadhesin in complex with 3' sialyllactose at 1.85 Å resolution. *Mol. Cell*, **1**, 719-728.
- Mcever, R. P. (1994), Selectins. *Curr. Opin. Immunol.*, **6**, 75-84.
- McKerracher, I., David, S., Jackson, D. L., Kottis, V., Dunn, R. J. and Braun, P. E. (1994), Identification of myelin-associated glycoprotein as a major myelin-derived inhibitor of neurite growth. *Neuron*, **13**, 805-811.
- Mayer, M. and Meyer, B. (1999), Characterization of ligand binding by saturation transfer difference spectroscopy. *Angew. Chemie.*, **110**, 1902-1906.
- Mayer, M. and Meyer, B., (2001), Group Epitope Mapping by Saturation Transfer Difference NMR to identify segments of a ligand in direct contact with a protein receptor. *J. Am. Chem. Soc.*, **123**, 6108-6117.
- Meyer, B., Weimar, T., Peters, T. (1997), Screening mixtures for biological activity by NMR. *Eur. J. Biochem.*, **246**, 705-709.
- Meyer, B. and Peters, T. (2003), Review article: NMR techniques for screeing and identifying ligand binding to receptor proteins. *Angew. Chemie*, **115**, 890-918.
- Moore, J. M. (1999), NMR techniques for characterization of ligand binding: utility for lead generation and optimization in drug discovery. *Biopolymers*, **51**, 221-243.
- Morell, P. Quarles, R. H. and Norton, W. T. (1994), Myelin formation, structure and biochemistry. In Siegel GJ (ed.) Basic neurochemistry: Molecular, cellular and medical aspects. Raven Press, New York.
- Morgan, W. D., Birdsall, B., Nieto, P. M., Gargaro, A. R. and Feeney, J. (1999), $^1\text{H}/^{15}\text{N}$ HSQC NMR studies of ligand carboxylate group interactions with arginine residues in complexes of brodimoprim analogues and *Lactobacillus casei* dihydrofolate reductase. *Biochemistry*, **38** (7), 2127-2134.

- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. and Olson, A. J. (1998), Automated docking using a Lamarckian genetic algorithm and empirical binding free energy function. *J. Comput. Chem.*, **19**, 1639-1662. <http://www.scripps.edu/mb/olson/doc/autodock/>
- Moseley, H. N. B., Curto, E. V. and Krishna, N. R. (1995), Complete relaxation and conformational exchange matrix (CORCEMA) analysis of NOESY spectra of interacting systems: Two-dimensional transferred NOESY. *J. Magn. Reson. B*, **108**, 243-261.
- Moseley, H. N. B., Lee, W., Arrowsmith, C. H. and Krishna, N. R. (1997), Quantitative Determination of Conformational Dynamic, and Kinetic Parameters of a Ligand-Protein/DNA Complex from a Complete Relaxation and Conformational Exchange Matrix Analysis of Intermolecular Transferred NOESY. *Biochemistry*, **36**, 5293-5299.
- Mukhopadhyay, C. and Bush, C. A. (1994), Molecular dynamics simulation of oligosaccharides containing *N*-acetyl neuraminic acid. *Biopolymers*, **34**, 11-20.
- Mukhopadhyay, G. Doherty, P., Walsh, F. S., Crocker, P. R. and Filbin, M. T. (1994), A novel role for myelin-associated glycoprotein as an inhibitor of axonal regeneration. *Neuron*, **13**, 757-767.
- Neuhaus, D. and Williamson, M. P. (1989), The Nuclear Overhauser effect in Structural and Conformational Analysis. New York: VCH Publishers Inc., 1st edition.
- Ni, F. (1992), Complete Relaxation Matrix Analysis of Transferred Nuclear Overhauser Effects. *J. Magn. Reson.*, **96**, 651-656.
- Ni, F. (1994), Recent Developments in Transferred NOE Methods. *Prog. Nucl. Magn. Reson. Spectrosc.*, **26**, 517-606.
- Niederost, B., Oertle, T., Fritsche, J., McKinney, R. A. and Bandtlow, C. E. (2002), Nogo-A and myelin-associated glycoprotein mediate neurite growth inhibition by antagonistic regulation of RhoA and Rac1. *J. Neurosci.*, **22**, 10368-10376.
- Nitschke, L., Floyd, H. and Crocker, P. R. (2001), New functions for the sialic acid-binding adhesion molecule CD22, a member of the growing family of siglecs. *Scand. J. Immunol.*, **53**, 227-234.
- Owens, G. C. and Bunge, R. P. (1989), Evidence for an early role for myelin-associated glycoprotein in the process of myelination. *Glia*, **2**, 119-128.
- Owens, G. C. and Bunge, R. P. (1991), Schwann cells infected with a recombinant retrovirus expressing myelin-associated glycoprotein antisense RNA do not form myelin. *Neuron*, **7**, 565-575.
- Pascal, S. M., Yamazaki, T., Singer, A. U., Kay, L. E. and Forman-Kay, J. D. (1995), Structural and dynamic characterization of the phosphotyrosine binding region of a src homology 2 domain-phosphopeptide complex by NMR relaxation,

- proton exchange and chemical shift approaches. *Biochemistry*, **34**, 11353-11362.
- Pedraza, L., Owens, G. C., Green, L. A. and Salzer, J. L. (1990), The myelin-associated glycoproteins: membrane disposition, evidence of a novel disulfide linkage between immunoglobulin-like domains, and posttranslational palmitylation. *J. Cell Biol.*, **111**, 2651-2661.
- Peters, T., Biet, T. and Herfurth, L. (2001), NMR-methods for screening the binding of ligands to proteins-identification and characterization of bioactive ligands. In: *Biological Magnetic Resonance*, (Ed.) Krishna N. R. and Lawrence Berliner, Vol. 22. Kluwer/Plenum Publishers.
- Pierelli, L., Teofili, L., Menichella, G., Rumi, C., Paoloni, A., Iovino, S., Puggioni, P. L., Leone, G. and Bizzi, B. (1993), Further investigations on the expression of HLA-DR, CD33 and CD13 surface antigens in purified bone marrow and peripheral blood CD34⁺ haematopoietic progenitor cells. *Br. J. Haematol.*, **84** (1), 24-30.
- Polley, M. J., Phillips, M. L., Wayner, E., Nudelman, E., Singhal, A. K., Hakomori, S. I. and Paulson, J. C. (1991), CD62 and endothelial-cell leukocyte adhesion molecule-1 (elam-1) recognize the same carbohydrate ligand, sialyl-lewis-x. *Proc. Natl. Acad. Sci. USA*, **88**, 6224-6228.
- Poppe, L., Dabrowski, J., Lieth, C. -W., Numata, M. and Ogawa, T. (1989), Solution conformation of sialosylcerebroside (GM4) and its NeuAc α (2-3)Gal β sugar component. *Eur. J. Biochem.*, **180**, 337-342.
- Poppe, L., Stuike-Prill, R., Meyer, B. and Halbeek, H. (1992), The solution conformation of sialyl- α (2 \rightarrow 6)-lactose studied by modern NMR techniques and Monte Carlo simulations. *J. Biomol. NMR*, **2**, 109-136.
- Poppe, L., brown, G. S., Philo, J. S., Nikrad, P. V. and Shah, B. H. (1997), Conformation of sLe^X Tetrasaccharide, Free and Bound to E-, P- and L-selectin. *J. Am. Chem. Soc.*, **119**, 1727-1736.
- Poveda, A., Asensio, J. L., Martín-Pastor, M. and Jiménez-Barbero, J. (1997), Solution Conformation and Dynamics of a Tetrasaccharide Related to the Lewis^X Antigen Deduced by NMR Relaxation Measurements. *J. Biomol. NMR*, **10**, 29-43.
- Prinjha, R., Moore, S. E., Vinson, M., Blake, S., Morrow, R., Christie, G., Michalovich, D., Simmons, D. L. and Walsh, F. S. (2000), Neurobiology: inhibitor of neurite outgrowth in humans. *Nature*, **403**, 383-384.
- Razi, N. and Varki, A. (1998), Masking and unmasking of the sialic acid-binding lectin activity of CD22 (siglec-2) on B lymphocytes. *Proc. Natl. Acad. Sci. USA*, **95**, 7469-7474.
- Reiling, S., Schlenkrich, M. and Brinkmann, J. (1996), Force field parameters for carbohydrates. *J. Comput. Chem.*, **17**, 450-468.

- Rens-Domiano, S. and Reisine, T. (1991), Structural analysis and functional role of the carbohydrate component of somatostatin receptors. *J. Biol. Chem.*, **266**, 20094-20102.
- Rini, J. M. and Lobsanov, Y. D. (1999), New animal lectin structures. *Curr. Opin. Struct. Biol.*, **9** (5), 578-584.
- Rinnbauer, M., Mikros, E. and Peters, T. (1998), Conformational analysis of a complex between Dolichos biflorus lectin and the Forssman pentasaccharide using transferred NOE build-up curves. *J. Carbohydr. Chem.*, **17**, 217-230.
- Roberts, G. C. K. (2000), Applications of NMR in drug discovery. *Drug Discovery Today*, **5**, 230-240.
- Sabesan, S., Bock, K. and Paulson, J. C. (1991), Conformational analysis of sialyloligosaccharides. *Carbohydr. Res.*, **218**, 27-54.
- Sadoul, R., Fahrig, T., Bartsch, U. and Schachner, M. (1990), Binding properties of liposomes containing the myelin-associated glycoprotein to neural cell cultures. *J. Neurosci. Res.*, **25**, 1-13.
- Salzer, J. L., Pedraza, L., Brown, M., Struyk, A., Afar, D. and Bell, J. (1990), Structure and function of the myelin-associated glycoproteins. *Ann. NY Acad. Sci.*, **605**, 302-312.
- Schachner, M. and Bartsch, U. (2000), Multiple functions of the myelin-associated glycoprotein MAG (siglec-4a) in formation and maintenance of myelin. *Glia*, **29** (2), 154-165.
- Schauer, R., Kelm, S., Reuter, G., Roggentin, P. and Shaw, L. (1995), Biology of the Sialic acids. In: A. Rosenberg (Ed.), Plenum Press, New York, pp. 7-67.
- Schauer, R. and Kamerling, J. P. (1997), Chemistry, biochemistry and biology of sialic acids. In: Montreuil J, Vliegthart JFG, Schachter H (Ed.) Glycoproteins II, Elsevier, Amsterdam, pp. 243-402.
- Scheffler, K., Ernst, B., Katapodis, A., Magnani, J. L., Wang, W. T., Weisemann, R. and Peters, T. (1995), Determination of the Bioactive Conformation of the Carbohydrate Ligand in the E-Selectin/ Sialyl Lewis X Complex. *Angew. Chem. Int. Ed. Engl.*, **34**, 1841-1844.
- Scheffler, K., Brisson, J. -R., Weisemann, R., Magnani, J. L., Wong, W. T., Ernst, B. and Peters, T. (1997), Application of Homonuclear 3D NMR Experiments and 1D Analogs to Study the Conformation of Sialyl Lewis^X Bound to E-Selectin. *J. Biomol. NMR*, **9**, 423-436.
- Scherf, T. and Anglister, J. (1993), A T1ρ-filtered two-dimensional transferred NOE spectrum for studying antibody interactions with peptide antigens. *Biophys. J.*, **64**, 754-761.
- Schriemer, D. C. and Hindsgaul, O. (1998), Deconvolution approaches in screening compound mixtures. *Comb. Chem. High Throughput Screen*, **1**, 155-170.

- Schuster, O. (2000), Identifizierung und Charakterisierung neuer E-Selektin Antagonisten. Dissertation, Universität Hamburg.
- Schwab, M. E., Kapfhammer, J. P. and Bandtlow, C. E. (1993), Inhibitors of neurite growth. *Annu. Rev. Neurosci.*, **16**, 565-595.
- Shapiro, M. J. and Gounarides, J. S. (1999), NMR methods utilized in combinatorial chemistry research. *Prog. Nucl. Magn. Reson. Spectrosc.*, **35**, 153-200.
- Shapiro, M. (2001), Applications of NMR screening in the pharmaceutical industry. *Farmaco*, **56**, 141-143.
- Shin, S.-Y., Universität Hamburg, personal communication.
- Shuker, S. B., Hajduk, P. J., Meadows, R. P. and Fesik, S. W. (1996), Discovering high-affinity ligands for proteins: SAR by NMR. *Science*, **247**, 1531-1534.
- Simmons, D. and Seed, B. (1988), Isolation of a cDNA encoding CD33, a differentiation antigen f myeloid progenitor cells. *J. Immunol.*, **141** (8), 2797-2800.
- Sklenár, V., Piotto, M., Leppik, R. and Saudek, V. (1993), Gradient-tailored Water suppression for ¹H-¹⁵N HSQC experiment optimized to return full sensitivity. *J. Magn. Reson. Ser. A*, **102**, 241-245.
- Spencer, T. and Filbin, M. T. (2004), A role for cAMP in regeneration of the adult mammalian CNS. *J. Anat.*, **204**, 49-55.
- Stein, P. E., Boodhoo, A., Armstrong, G. D., Heerze, L. D., Cockle, S. A., Klein, M. H. and Read, R. J. (1994), Structure of a pertussis toxin-sugar complex as a model for receptor binding. *Nat. Struct. Biol.*, **1** (9), 591-596.
- Suresh, M. X. and Veluraja, K. (2003), Conformations of terminal sialyloligosaccharide fragments- a molecular dynamics study. *J. Theo. Biol.*, **222**, 389-402.
- Sutcliffe, M. J., Haneef, I., Carney, D. and Blundell, T. L. (1987), Knowledge-based modelling of homologous proteins, Part I: Three-dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng.*, **1**, 377-384.
- Sybyl molecular modelling software, version 6.9, Tripos Associates, 1699 South Hanely Rd., St. Louis, MO 63144-2917. <http://www.tripos.com>
- Tang, S., Shen, Y. J., DeBellard, M. E., Mukhopadhyay, G., Salzer, J. L., Crocker, P. R. and Filbin, M. T. (1997), Myelin-associated Glycoprotein Interacts with Neuron via a Sialic acid Binding site at ARG118 and a Distinct Neurite Inhibition Site. *J. Cell Biol.*, **138**, 1355-1366.
- Thøgersen, H., Lemieux, R. U., Bock, K. and Meyer, B. (1982), Further Justification for the exo-Anomeric Effect. Conformational Analysis Based on Nuclear

- Magnetic Resonance Spectroscopy of Oligosaccharides. *Can. J. Chem.*, **60**, 44-57.
- Trapp, B. D. (1990), Myelin-associated glycoprotein: location and potential functions. *Ann. NY Acad. Sci.*, **605**, 29-43.
- Tettamanti, G., Bonali, F., Marchesini, S. and Zambotti, V. (1973), A new procedure for the extraction, purification and fractionation of brain gangliosides. *Biochim. Biophys. Acta*, **296**, 160-170.
- Tropak, M. B., Johnson, P. W., Dunn, R. J. and Roder, J. C. (1988), Differential splicing of MAG transcripts during CNS and PNS development. *Brain Res.*, **464**, 143-155.
- Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G. and Marth, J. (1999), Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Vabnick, I., Novakovik, S. D., Levinson, S. R., Schachner, M. and Shrager, P. (1996), The clustering of axonal sodium channels during development of the peripheral nervous system. *J. Neurosci.*, **16**, 4914-4922.
- van den Berg, T. K., Breve, J. J., Damoiseaux, J. G., Dopp, E. A., Kelm, S., Crocker, P. R., Dijkstra, C. D. and Kraal, G. (1992), Sialoadhesin on macrophages: its identification as a lymphocyte adhesion molecule. *J. Exp. Med.*, **176** (3), 647-655.
- Vasudevan, S. V. and Balaji, P. V. (2002), Comparative analysis of ganglioside conformations by MD simulations: implications for specific recognition by proteins. *J. Mol. Struct. (Theochem)*, **583**, 215-232.
- Veluraja, K. and Rao, V. S. R. (1983), Theoretical studies on the conformation of monosialogangliosides and disialogangliosides. *Carbohydr. Polymers*, **3**, 175-192.
- Veluraja, K. and Rao, V. S. R. (1984), Studies on the conformations of sialyloligosaccharides and implications on the binding specificity of neuraminidases. *J. Biosci.*, **6**, 625-634.
- Virchow, R. (1854), Über das ausserbreitete Vorkommen einer dem Nervenmark analogen Substanz in den tierischen Geweben. Virchows Archiv für pathologische *Anatomie*, **6**, 562.
- Vogtherr, M. and Peters, T. (2000), Application of NMR based binding assay to identify key hydroxyl groups for intermolecular recognition. *J. Am. Chem. Soc.*, **122**, 6093-6099.
- Vyas, A. A., Patel, H. V., Fromholt, S. E., Heffer-Laue, M., Vyas, K. A., Dang, J., Schachner, M. and Schnaar, R. L. (2002), Gangliosides are functional nerve cell ligands for myelin-associated glycoprotein (MAG), an inhibitor of nerve regeneration. *Proc. Natl. Acad. Sci. USA*, **99** (12), 8412-8417.

- Wang, K. C., Koprivica, V., Klm, J. A., Sivasankaran, R., Guo, Y., Neve, R. L. and He, Z. (2002), Oligodendrocyte-myelin glycoprotein is a Nogo receptor ligand that inhibits neurite outgrowth. *Nature*, **417**, 941-944.
- Walz, G., Aruffo, A., Kolanus, W., Bevilacqua, M. and Seed, B. (1990), Recognition by elam-1 of the sialyl-lex determinant on myeloid and tumor-cells. *Science*, **250**, 1132-1135.
- Weimar, T., Bukowski, R. and Young, N. M. (2000), The conformation of the T-antigen disaccharide bound to *Maclura pomifera* Agglutinin in aqueous solution. *J. Biol. Chem.*, **275**, 37006-37010.
- Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta Jr., S. and Weiner, P. (1984), A new force field for molecular mechanical simulation of nucleic acid and proteins. *J. Am. Chem. Soc.*, **106**, 765-784.
- Wu, W. -G., Pasternack, L., Huang, D. -L., Koeller, K. M., Lin, C. -C., Seitz, O. and Wong, C. -H. (1999), Structural Study on *O*-GlcNAc, *O*-LacNAc, *O*-Sialyl-LacNAc, and *O*-Sialyl-Lewis-X Peptide of the Mucin Domain of MAdCAM-1. *J. Am. Chem. Soc.*, **121**, 2409-2417.
- Xu, Y., Krishna, N. R. and Sugár, I. (1995), A global optimization method based on variable target functions for fitting of the experimental and calculated NOESY spectra. *J. Magn. Reson. Ser. B.*, **107**, 201-209.
- Yamashita, T., Higuchi, H. and Tohyama, M. (2002), The p75 receptor transduces the signal from myelin-associated glycoprotein to Rho. *J. Cell Biol.*, **157**, 565-570.
- Yang, L. J. -S., Zeller, C. B., Shaper, N. L., Kiso, M., Hasegawa, A., Shapiro, R. E. and Schnaar, R. L. (1996), Gangliosides are neuronal ligands for myelin-associated glycoprotein. *Proc. Natl. Acad. Sci. USA*, **93** (2), 814-818.

Curriculum vitae

Anirban Bhunia, born 25.02.1977 in Midnapore, India

School

1990 – 1992	Secondary (10 th) from Nirmal Hriday Ashram (Catholic Church)
1992 – 1994	Higher secondary (10+2) from Vidyasagar Vidyapith

College/University

1994 – 1997	B.Sc. (Hons.) in Chemistry from Midnapore College
1997 – 1999	M.Sc. in Biotechnology from Indian Institute of Technology Bombay M.Sc. dissertation project: “Enzyme mediated waste water treatment”
07/1999 – 10/2000	Research Assistant with Prof. Pramod P. Wangikar at Dept. of Chemical Eng., Indian Institute of Technology Bombay
11/2000	Started Ph.D work with Prof. Thomas Peters at Institute of Chemistry, University of Lübeck
