Investigating Glycosaminoglycan Conformations by Ion Mobility Mass Spectrometry and Nuclear Magnetic Resonance

Lan Jin

A thesis submitted for the degree of Doctor of Philosophy

The University of Edinburgh
2007

Declaration

Unless otherwise stated, the work described in this thesis is my own work and has not been submitted in whole or as part for a degree or other qualification at this or any other University.

Lan Jin

Acknowledgement

I would like to thank all the people who have helped and collaborated with me during my PhD project.

I would like to thank Dr. Dušan Uhrín, my supervisor, who led me into the NMR world. I really appreciate his patience and instruction throughout the whole project. Prof. David W H Rankin, my second supervisor, received my application to the School of Chemistry and gave me the opportunity to contact Dušan. Thanks for his introduction. Dr. Perdita E Barran provided with me the IMMS experimental results and much help and useful discussions during my first year of research, from which I successfully transferred to my PhD project.

Dr. Malcolm Lyon kindly provided heparin oligosaccharides for my research. Dr. Andrew P. Herbert kindly provided with me the NMR structures of factor H modules. Astrid Jahnke in Dr. Alison N. Hulme's group provided the synthesized TEMPO-disaccharide. Thank you for all your help.

I would like to thank my parents, who gave me support during my studies abroad. They showed their care to me through daily E-mail with their support. I would also like to thank all my friends here in Edinburgh and at home.

Finally, I would like to acknowledge funding from the ORS awards and the Linked ORS awards from the University of Edinburgh, and the School of Chemistry.

Abstract

Glycosaminoglycans (GAGs) constitute a considerable fraction of the glycoconjugates found on cellular membranes and in the extracellular matrix of virtually all mammalian tissues. Their ability to bind proteins and alter protein-protein interactions or enzymatic activities has identified them as biologically important molecules. The binding specificity of GAGs is encoded in their primary structures, but ultimately depends on how their functional groups are presented to a protein in the three-dimensional space. The aim of this project is to develop and test methodologies for conformational studies of free and bound GAGs.

Gas phase conformations of heparin-derived oligosaccharides were studied by using Ion Mobility Mass Spectrometry (IMMS) and molecular modelling. Gas phase conformations were modelled using the AMBER force field. Their theoretical collision cross-sections were compared with the IMMS data and a good agreement was obtained. The gas phase conformations of heparin oligosaccharides were more compact than those observed in solution or crystal structures of heparin-protein complexes. This was attributed to the effects of sodium cations interacting with the negatively charged sulfate and carboxyl groups of oligosaccharides. Adiabatic maps of dihedral angles vs. potential energy of disaccharide fragments of tetrasaccharides were calculated in the absence of the sodium cations.

New NMR methods for the measurement of scalar and dipolar ^{1}H-^{1}H coupling constants and ^{13}C-^{13}C coupling constants in natural abundance ^{13}C samples were developed. Performance of these methods was tested extensively.

Solution conformation of the heparin-derived fully sulfated tetrasaccharide was studied by NMR spectroscopy. ^{1}H-^{1}H scalar coupling constants were used to characterize the dynamic equilibria of flexible monosaccharide rings. ^{1}H-^{1}H and ^{1}H-^{13}C RDCs were used in the study of the conformations of the glycosidic linkages. RDC-refined structures were obtained from molecular dynamics with sodium cations in explicit water.

Interactions of the heparin-derived fully sulfated tetrasaccharide with factor H modules, fH~19, 20 and fH~7, were studied using AUTODOCK. Conformation of a spin-labelled heparin-derived fully sulfated disaccharide was studied by NMR and molecular modelling.

Contents

List of Figures

Abbreviations

COSY	correlation spectroscopy
CPMG	Carr-Purcell-Meiboom-Gill
CSSF	chemical shift selective filter
DEPT	distortionless enhanced polarization transfer
DNP	dynamic nuclear polarization
DQ	double-quantum
FID	free induction decay
GAGs	glycosaminoglycans
HMBC	herteronuclear multiple-bond correlation
HOHAHA	homonuclear Hartmann-Hahn
HSQC	heteronuclear single quantum coherence
IMMS	ion mobility mass spectrometry
INADEQUATE	incredible natural abundance double-quantum transfer experiment
INEPT	insensitive nuclei enhanced by polarization transfer
MD	molecular dynamics
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect (enhancement)
NOESY	nuclear Overhauser effect spectroscopy
PAS	principle axis system
PEG	polyethylene glycol
PFG	pulsed field gradient
ppm	parts per million
RDC	residual dipolar coupling
RF	radiofrequency
RMSD	root mean square deviation
ROESY	rotating frame nuclear Overhauser effect spectroscopy
SPFGSE	single pulse field gradient spin-echo
SNR	signal-to-noise ration
SVD	singular value decomposition
TOCSY	total correlation spectroscopy

Chapter I

Introduction

I. 1 Glycosaminoglycans

Glycosaminoglycans (GAGs) constitute a considerable fraction of the glycoconjugates found on cellular membranes and in the extracellular matrix of virtually all mammalian tissues. Their ability to bind proteins and alter protein-protein interactions or enzymatic activities has identified them as important determinants of cellular responsiveness in development, homeostasis and disease. GAGs usually exist as the O-linked side-chains of proteoglycans [1]. Members of the GAG family include heparin, heparan sulfate (HS), dermatan sulfate (DS), chondroitin sulfate (CS), hyaluronan (HA) and keratan sulfate (KS), whose unsulfated forms consist of repeating disaccharide units containing hexosamine and uronic acid residues (Table I.1.1). These polysaccharide precursors are extensively modified (epimerisation, N-deacetylation and N-, O-sulfation) creating a wide heterogeneity to GAG structures. The binding specificity of GAGs is encoded in their primary structures, but ultimately depends on how their functional groups are presented to a protein in the three-dimensional space. In recent years, GAGs have attracted considerable attention from researchers across a spectrum of disciplines.

Table I.1.1 Structures of the main repeating units of different GAGs [2]

GAGs	Structure of main repeating disaccharide units
Heparin	-4)-α-L-IdoA2(OSO$_3^-$)-(1-4)-α-D-GlcNSO$_3^-$6(OSO$_3^-$)-(1-
Unsulfated domain of HS Sulfated domain of HS	-4)-β-D-GlcA-(1-4)-α-D-GlcNAc-(1- -4)-β-D-IdoA2(OH/OSO$_3^-$)-(1-4)-α-D-GlcNSO$_3^-$6(OH/OSO$_3^-$)-(1-
Dermatan Sulfate (DS)	-4)-α-L-IdoA-(1-3)-β-D-GalNAc4(OSO$_3^-$)-(1-
Chondroitin Sulfate CS4	-4)-β-D-GlcA-(1-3)-β-D-GalNAc4(OSO$_3^-$)-(1-
Chondroitin Sulfate CS6	-4)-β-D-GlcA-(1-3)-β-D-GalNAc6(OSO$_3^-$)-(1-
Keratan Sulfate (KS)	-3)-β-D-Gal-(1-4)-β-D-GlcNAc6(OSO$_3^-$)-(1-
Hyaluronan (HA)	-4)-β-D-GlcA-(1-3)-β-D-GlcNAc-(1-

In the GAG family, heparin and HS attract major research interests because of their biological and pharmaceutical importance. Numerous heparin-binding proteins have been discovered and their binding characteristics described [3]. Among these are heparin-binding proteins such as basic Fibroblast Growth Factor (bFGF or FGF-2) [4], acidic FGF (FGF-1) [5], transforming Growth Factor β-1 (TGF β-1) [6], Hepatocyte Growth Factor (HGF) [7], Cardiotoxins (CTXs [8] and Chemokines (CXC) [2]. HS has an important role as a co-receptor in activation of many growth factor receptors. By binding directly to both thrombin and antithrombin III, heparin inhibits blood coagulation - a property that leads to its clinical application [9]. Heparin also has a regulatory role in limiting inflammation, which is unrelated to its anticoagulant activity [10]. It has also been discovered that heparin can minimize angiogenesis, which enhances the growth of solid tumours [9].

Studies show that CS and DS chains, GlcA/IdoA-GalNAc, have intriguing functions in the central nervous system (CNS) development, wound repair, infection, growth factor signalling, morphogenesis and cell division [11]. CSs are used as chondroprotective and antirheumatic drugs in the treatment of tibiofibular osteoarthritis [12]. The presence of a sulfate group at C4 in GalNAc is thought to be important in the cytoadherence of malaria-infected red blood cells [13] and in the binding of the malaria parasite to placenta [14]. The over-sulfated CS with two or three sulfate groups per disaccharide unit exhibits enhanced antithrombotic activity [15]. CS and DS chains can specifically bind to some heparin-binding proteins. E.g. the interaction of DS with FGF-2 and FGF-7 is implicated in wound repair [6a] and DS also binds to HGF [16]. Sulfates at the C-2 of the iduronic acid residues and C-4 of the galactosamines constitute active regions that bind with HCII [17], an inhibitor of thrombin, which gives DS its mild anticoagulant activity. The key roles in regulating cell migration, recognition and tissue morphogenesis have been reviewed [18].

KS has been investigated as a potential early marker for osteoarthritis disease because osteoarthritis is associated with elevated KS levels, which can be determined by the monoclonal antibody [19]. Some structures or components of KS, derived from different tissues such as sheep brain [20], bovine [21] and bonefish [22] tissues have been studied by NMR.

HA have many physiological functions, such as water homeostasis, regulation of capillary growth, cell recognition and cell migration [23]. Some evidence suggests that HA plays an essential role in the regulation of tumour growth, tissue invasion and metastasis [24]. The structure of a regulatory HA binding domain of the inflammatory Leukocyte Homing Receptor CD44 has been studied [25] by X-ray and NMR. HA-binding proteins such as TSG-6, BRAL1 have been reviewed [26].

In this project, heparin-derived oligosaccharides are used for the conformational study. Therefore only heparin and HS conformations are discussed in the following sections.

I. 2 Conformations of heparin and heparan sulfate (HS)

Heparin and HS are both linear polysaccharides containing similar structural units of 1-4 linked pyranosyluronic acid and 2-amino-2-deoxyglucopyranose (glycosamine) residues (Fig.I.2.1). The differences between their structures are in the level of sulfate substitution and the occurrence of the predominant acid - iduronic acid in heparin and glucuronic acid in HS. An average heparin disaccharide contains 2.7 sulfate groups, whilst in HS there is only about one sulfate group per disaccharide unit [2]. HS contains all of the structural variations found in heparin but with more minor sequence variants, which makes HS much more complex than heparin [2].

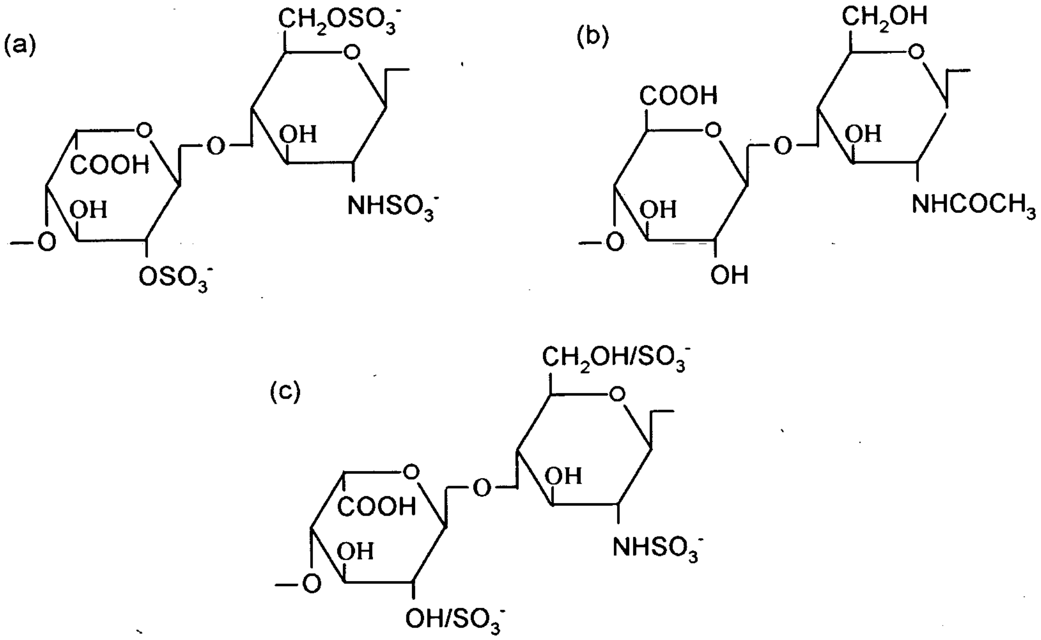


Fig.I.2.1 The main repeating disaccharide units of (a) heparin, (b) unsulfated domains of HS and (c) sulfated domains of HS

The solution conformation of heparin-derived oligosaccharide from NMR studies shows a helical structure with clustering of sulfate groups on two sides of the molecule [27]. The distance between the sulfate groups in the same position in the two consecutive repeating disaccharide units and at the same face of the helix is about 17 Å [24]. This ribbon-like structure exists in both the free heparin in solution and the bound heparin in crystal structures [1]. When interacting with polylysine, heparin adopts a right-handed α-helical form, in which the sulfate groups are spaced at intervals of 17 Å and match the spacing of lysine side-chains very well [27]. But a different geometry [3c] of heparin exists in the ternary complex of FGF1-FGFR2-heparin [4c], where a 'single, asymmetric' heparin molecule links the FGF-FGFR binary complexes [3c].

In carbohydrates the conformational flexibility is usually associated with glycosidic linkages, while the individual monosaccharide rings form rigid structures. This is not the case for heparin, where additional flexibility is conferred by the IdoA residues [1]. The conformation of the flexible L-iduronic acid (IdoA) residue varies depending on its substitution pattern and relative position in the polymeric chain [2]. Three possible conformations, $^{4}C_{1}$, $^{1}C_{4}$ and $^{2}S_{o}$ (Fig.I.2.2), exist when the IdoA is at the reducing end of the oligosaccharide. However, the internal IdoA adopts equilibrium between $^{1}C_{4}$ and $^{2}S_{o}$ forms, apparently, without causing the whole polysaccharide chain to bend [1]. The $^{2}S_{o}$ form appears to be slightly favoured [28]. One Molecular Dynamics (MD) study of heparin decasaccharide in explicit water was carried out showing that IdoA shares both $^{2}S_{o}$ and $^{1}C_{4}$ conformations during a 3ns simulation [29].

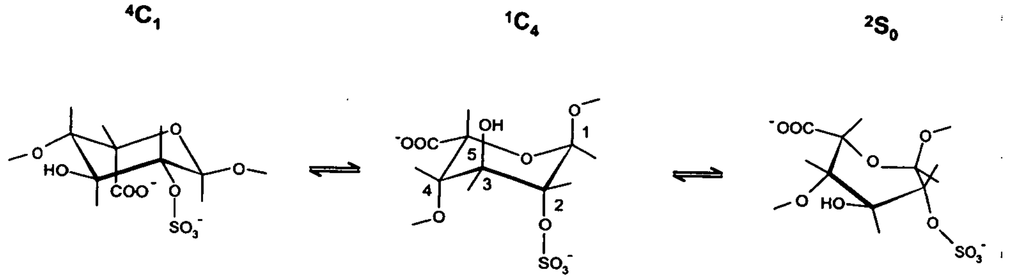


Fig.I.2.2 $^{4}C_{1}$, $^{1}C_{4}$ and $^{2}S_{O}$ conformations of the iduronic acid residue [33]. In the chair forms, atoms C2, C3, C5 and O5 are in one plane (the atom numbering is clockwise); in $^{4}C_{1}$ form C4 is above the plane while C1 is below. In $^{1}C_{4}$ form, C1 is above the plane while C4 is below. In $^{2}S_{O}$ form, a plane can be defined containing atoms C1, C3, C4 and C5, where C2 is above the plane while O5 is below.

Heparin and HS oligosaccharides prepared by enzymatic cleavage of polysaccharides contain a terminal unsaturated uronic acid residue (ΔUA). This residue can exist in either $^{2}H_{1}$ or $^{1}H_{2}$ forms

(Fig.I.2.3), and the equilibration between these two conformers is controlled by the sulfation pattern of this residue [25, 30]. NMR studies show that $^1H_2$ is generally the favoured conformation [31], although both forms appeared in the same unit cell in a crystal structure [32], which indicates that they have nearly the same energy.

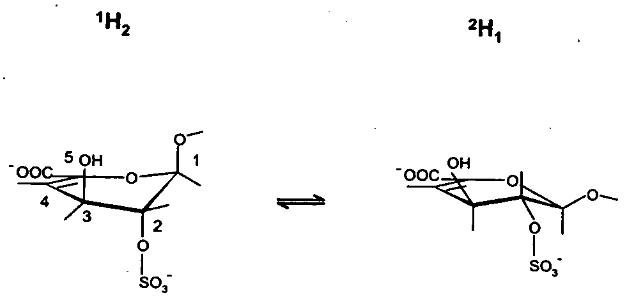


Fig.I.2.3 $^1H_2$ and $^2H_1$ conformations of 4,5-unsatarated uronic acid residue [30]. When the atom numbering is clockwise, the plane is formed by C3, C4, C5 and O5. In the $^1H_2$ form C1 is above the plane while C2 is below. In the $^2H_1$ form C2 is above the plane while C1 is below.

Solution conformations of heparin-derived tetra- [28] and hexasaccharides [34a] and synthesized heparin-like hexasaccharides [34b] have been studied by NMR and molecular modelling. All glucosamine rings in the hexasaccharide showed $^4C_1$ conformation, while the iduronate rings were in the equilibrium between $^1C_4$ and ^2S_o forms [31]. Both forms (for different IdoA residues) were found in the crystal structure of the heparin-derived hexasaccharide bound to FGF [31a, b]. This was also observed in the complex of a synthesized heparin-like hexasaccharide bound to FGF-1 [31c]. The glucosamines of the tetrasaccharide were also in the $^4C_1$ form while the iduronate rings mainly showed the ^2S_o form rather than the $^1C_4$ form [28]. When bound to FGF-2, the iduronate rings were stabilized in the $^1C_4$ form [28].

Alongside the studies of free heparin/HS structures, many protein-heparin/HS complexes have been investigated [2, 3c, 9, 35]. The existence of different sulfation patterns and alternative structural subunits in heparin/HS oligosaccharides [36] has direct consequences for their activity that is manifested through the binding to their protein receptors [37]. Studies that involve determination of residue types and sulfation patterns required for binding are therefore very important. Equally important is the determination of the conformation of GAGs in the bound state. In the complex of Annexin V [38] with a heparin-derived tetrasaccharide, the sugar shows a helical conformation that was also found by NMR in the free dodecasaccharide [39]. The terminal uronate is in $^1H_2$ form and the glucosamine residues are in $^4C_1$ form. The internal IdoA2S show both ^2So and $^1C_4$ forms. In the co-crystal of bFGF with both tetra- and hexasaccharides [4a] containing internal -IdoA($2SO_3$)-GlcNSO_3(6OSO_3) sequences, the internal IdoA residues also show both ^2S_o and $^1C_4$ forms. The IdoA($2SO_3$) that was in contact with the protein was in the $^1C_4$ chair conformation while the interacting GlcNSO_3(6OSO_3) residue was in the $^4C_1$ chair form.

These examples illustrate that a variety of conformations exist in the heparin/HS family, Conformation studies of heparin/HS oligosaccharides mainly deal with the fully sulfated species and only a few studies have been carried out focussing on the effects the different sulfation patterns have on the conformation of these polyanions [40]. The negatively charged groups play important roles in protein binding; nevertheless, it is likely that the conformational changes that accompany the loss of specific sulfate groups are equally important in the recognition processes.

I. 3 Aims of the project

The aim of this project is to develop and test methodologies that can potentially be useful in the conformational studies of free and bound GAGs. In details the aims of the project are:

- study of the gas phase conformations of sodiated GAG oligosaccharide ions by Ion Mobility Mass Spectrometry (IMMS);

- study of the solution conformations of GAGs by NMR; This part of program lead to researches outlined in the next two points;

- development of new NMR methods for accurate and precise measurement of scalar and dipolar ^1H-^1H coupling constants in aligned samples;

- development of new NMR methods for the measurement of ^{13}C-^{13}C coupling constants in samples with natural abundance of ^{13}C;

- simulation of protein-GAG binding using docking;

- study of the conformation of a spin-labelled disaccharide using NMR and molecular modelling.

I. 4 Molecular Modeling

Computational methods were used to obtain molecular models of heparin-derived oligosaccharides for the interpretation of experimental results. Ab initio techniques consider all electrons for each atom in the molecule making them computationally intensive and therefore restricted to the optimisation of geometries of smaller systems [41]. Semi-empirical methods can, in principle, perform the conformational analysis [42] but they cannot consider the long range interactions between molecules [39]. Again, these methods are good for geometry optimisation of local minima. To study the conformations of oligosaccharides, a large variety of structures representing their dynamic nature are required to get close to the average conformation. Therefore the most efficient approach is to use molecular mechanics.

In molecular mechanics, a molecule is described by simple analytical functions as a collection of atoms that interact with each other. Different implementations of this principle use different force fields such as MM3, CHARM and AMBER. All these force fields represent molecular energy surfaces by empirical and semi-empirical potential energy functions of small molecules [39] derived from experimental data such as bond lengths, bond angles, dihedral angles, strain energies etc. Assuming additive and transferable properties of these functions, parameters derived from small molecules can be used in studies of large and complex molecular systems [39]. AMBER uses the following additive potential energy function [43]:

$$V_{total} = \sum_{bonds} K_r (r-r_{eq})^2 + \sum_{angles} K_\theta(\theta-\theta_{eq})^2 + \sum_{dihedrals} \frac{1}{2} V_n [1+\cos(n\Phi-\gamma)] + \sum_{i<j} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}} \right) + \sum_{H\text{-}bonds} \left(\frac{C_{ij}}{R_{ij}^{12}} - \frac{D_{ij}}{R_{ij}^{10}} \right) \qquad \textbf{(Eq.I.4.1)}$$

V_{total}: total potential energy, which consists of functions of bonds, angles, dihedrals, van der Waals, electrostatic and H-bonds (in AMBER7 H-bonds is neglected);
K_r: bond stretching constant; **r/r_{eq}**: the (equilibrium) bond length; **Φ**: dihedral angle;
K_θ: angle bending constant; **θ/θ_{eq}**: the (equilibrium) angle between two/equilibrium bonds;
V_n: the height of the torsion barrier; **n**: coefficient of symmetry, n= 1, 2, 3…; γ: phase preference
A, B, C, D: the atom-type dependent constants; **q_{ij}**: the atoms' partial charges;
R_{ij}: the (equilibrium) distance separating the atoms' centers;
ε: the relative dielectric coefficient of the medium between the charges;

Different force fields have focused on parameterization of different functional groups. Most force fields are well parameterized for amino acids and nucleic acids. As the interest in carbohydrates has grown, parameterizations specific to this class of molecules appeared over time, e.g. the GLYCAM parameter [44] set has been developed that is compatible with the AMBER force field. The latest version of AMBER9 (2006) has a built-in force field parameterized for carbohydrates. The sulfur parameters were added in for OSO_3^- but not for NSO_3^- groups, which are common in GAGs oligosaccharides. In this work AMBER7 was modified to include the parameters for sulfate carbohydrates (both OSO_3^- and NSO_3^-) and used for the molecular modelling.

During a typical molecular modelling, the potential energy of the molecule is calculated by **Eq.I.4.1**, followed by an energy minimization step to drive the molecule to the nearest local minimum. The nature of this minimum depends on the minimization method and the initial conformation of the molecule. This process only explores a small percentage of the total conformational space. In order to sample larger conformational space, molecular dynamics and simulated annealing are used as the next step. Molecular dynamics explores multidimensional conformational spaces and provides a continuous dynamic picture of the molecular system. However, this method is limited to sampling minima with relatively small energy barriers. Simulated annealing can overcome high energy barriers to ensure transitions between local minima by heating the molecule to a high temperature. After the cooling down step, new conformations can be frozen out. However, not all conformations obtained by the simulated annealing correspond to low energy structures as their geometries could have been distorted at high temperatures.

I. 5 Ion Mobility Mass Spectrometry (IMMS) [49]

IMMS has been used in gas phase structure analysis of biomolecules to provide information about their structures in the absence of solvent. Most of the researches in this area are focused on proteins [45] and peptides [46]. Some results were also obtained for carbon or metal clusters [47]. Only few papers have reported IMMS studies of carbohydrates [48]. Simple sodiated oligosaccharide ions were studied by Bowers et al. [45a].

Ion Mobility Mass Spectrometry is based on the theory of ion mobility in the gas phase [49]. The ion mobility of a gas phase ion is the quantity that describes how quickly the ion moves through a drift cell filled with a high pressure buffer gas (such as helium) under the influence of a weak electric field. The shape of ions affects their speed, which results in different drifting times of ions in the cell. Small, compact ions with smaller collision cross-sections drift more quickly than large, extended ions with larger collision cross-sections. In this way, coupled with mass spectrometry, ions with the same mass-to-charge ratio but different shapes or structures are separated.

The ion mobility is measured in a drift cell, which is usually 4-20 cm long, temperature controlled from 80-500 K, filled with a buffer gas at 1-5 Torr and under a weak (5-30 V cm^{-1}), uniform electric field. As the ions drift under the influence of the electric field, they collide with He atoms. As a result of these two opposing forces, the electric field and the retardation caused by collisions with a buffer gas, ions will drift through the cell with a constant velocity. The drifting time of ions can be expressed as:

$$t_d = \frac{L}{V_d} = \frac{L}{KE} = \frac{L^2}{KV} = \frac{L^2 T_0}{K_0 P_0 T} \cdot \frac{P}{V} \qquad \textbf{(Eq.I.5.1)}$$

K: the ion mobility; K_0: the reduced ion mobility;

P: the buffer gas pressure; T: the cell temperature;

P_0: 760 Torr; T_0: 273.15 K;

L: the length of the drift cell; V: the voltage across the cell; E: the electric field

A plot of drifting times, t_d, versus P/V ratio yields a straight line. The reduced ion mobility, K_0, can be calculated from the slope of the line. The reduced ion mobility, K_0, can be expressed as a function of the momentum transfer collision integral, Ω, that depends on the shape of the ion.

$$K_0 = \frac{3e}{16N} \cdot \left(\frac{2\pi}{\mu K_B T}\right)^{1/2} \cdot \frac{1}{\Omega} \qquad \textbf{(Eq.I.5.2)}$$

N: the buffer gas number density; μ: the reduced mass of ion and buffer gas;

K_B: the Boltzman constant; e: the ion charge

Using the experimental K_0, the momentum transfer collision integral, Ω, can be calculated. A good approximation of this integral, obtained using a Monte Carlo routine, is the projection cross-section, σ. The theoretical projection cross-section can easily be calculated for various conformations of studied molecule obtained from molecular modeling. By comparing the experimental and the theoretical collision cross-sections, the gas phase conformations of the ions can be deduced.

I. 6 NMR spectroscopy [50]

When placed in an external, static magnetic field, B_0, magnetic nuclei with a non-zero nuclear spin quantum number, I, will precess at a rate dependent on the field strength and the gyromagnetic ratio, γ, of the spin. An associated nuclear spin magnetic moment, μ, is generated as a consequence of the motion of the charge bearing nuclei with the nuclear spin angular momentum, $\hat{I}$ **(Eq.I.6.1)**. The magnetic moment generates a small local magnetic field, which interacts with the external magnetic field. The energy of the interaction can be calculated with the Hamiltonian operator, $\hat{H}$, from the z-component of the nuclear spin angular momentum, $\hat{I}_z$ **(Eq.I.6.2)**.

$$\mu = -\gamma\hat{I} \qquad \textbf{(Eq.I.6.1)}$$

$$\hat{H} = -\gamma B_0\hat{I}_z \qquad \textbf{(Eq.I.6.2)}$$

The interaction energy between the nuclear magnetic moment and the external field is very much smaller than the energy of the thermal motion of the molecule. Therefore, individual magnetic moments adopt random orientations when first placed in a magnetic field. However, there is a slight energy difference between the two spin states of spin-half nuclei, α and β. The magnetic moments have a slight preference to align parallel with the external magnetic field occupying the lower energy α state. After time comparable to T_1, the spin-lattice relaxation time, magnetic moments will reach equilibrium distribution with a net bulk magnetization vector aligning parallel to the external magnetic field along the z-axis (Fig.I.6.1). The process of the equilibrium formation is referred to as relaxation. When the bulk magnetization is flipped acquiring a non-zero angle with the z-axis, it will execute the precessional motion. The rate of the precession is defined by the angular velocity, ω_0, or frequency υ_0 as:

$$\omega_0 = -\gamma B_0 \qquad \textbf{(Eq.I.6.3)}$$

$$\upsilon_0 = -\frac{\gamma B_0}{2\pi} \qquad \textbf{(Eq.I.6.4)}$$

υ_0 is known as the Larmor frequency (in Hz) of the nucleus. For a single spin-half nucleus, there are two energy levels (spin states), α (spin-up, +1/2, low energy) and β (spin-down, -1/2, high energy). When spinning nuclei absorb a quantum of energy, in the form of electromagnetic radiation, the spin state will change from α to β.

$$\Delta E = h\nu = \frac{\gamma h B_0}{2\pi} \qquad \textbf{(Eq.I.6.5)}$$

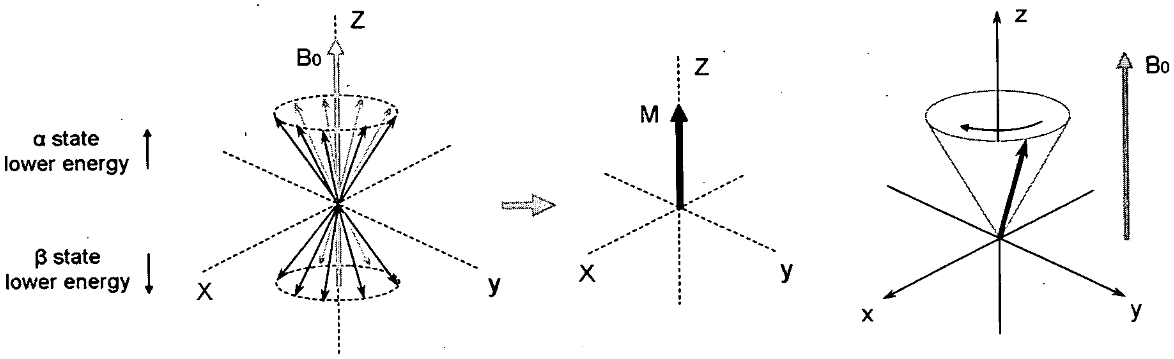


Fig.I.6.1 The formation of the equilibrium magnetization under external magnetic field and the Larmor precession

If the frequency of the electromagnetic radiation matches the Larmor frequency of the nucleus, nuclear magnetic resonance signals are observed with absorption lines at specific positions corresponding to the Larmor frequencies in the NMR spectrum. The resonance frequencies are sensitive to the distribution of the electrons in the chemical bonds within the molecule. Therefore, the position of the absorption lines carries the structural information.

The frequency of the absorption lines is proportional to the external magnetic field strength. This dependency makes it difficult to compare NMR spectra from different magnets. Therefore, the chemical shift scale, which is independent of the magnetic field strength, is used. The unit of the chemical shift scale is ppm (parts per million). The zero ppm in ^1H and ^{13}C NMR spectra is defined using the signal of a reference compound, TMS (tetramethylsilane) and the chemical shift is defined as:

$$\delta(\text{ppm}) = 10^6 \times (\upsilon - \upsilon_{\text{ref}})/\upsilon_{\text{ref}} \qquad \textbf{(Eq.I.6.6)}$$

NMR signals are usually not singlets but also show a fine structure. This peak splitting comes from the spin-spin coupling (scalar coupling) transferred through chemical bonds between individual nuclei. The coupling interaction causes splitting of energy levels introducing several additional transitions, which results in the splitting of peaks. The scalar coupling is independent of the magnetic field strength and enables establishing the bonds between atoms. For coupled spin-half nuclei, each spin can be in one of the two spin states, α or β. The first spin experiences the two spin states of the second spin and its transitions are split equally into a doublet with a frequency difference of the scalar coupling constant, J in Hz. The doublet is centered at the frequency of the first spin with individual lines shifted by $\pm J/2$. The same doublet with the splitting of J, exists at the frequency of the second spin (Fig.I.6.2a). If more than two spins are coupled to each other, the shape of the multiplet can be predicted by a tree diagram considering one coupling at a time (Fig.I.6.2b). This rule applies to both homo- and heteronuclear couplings.

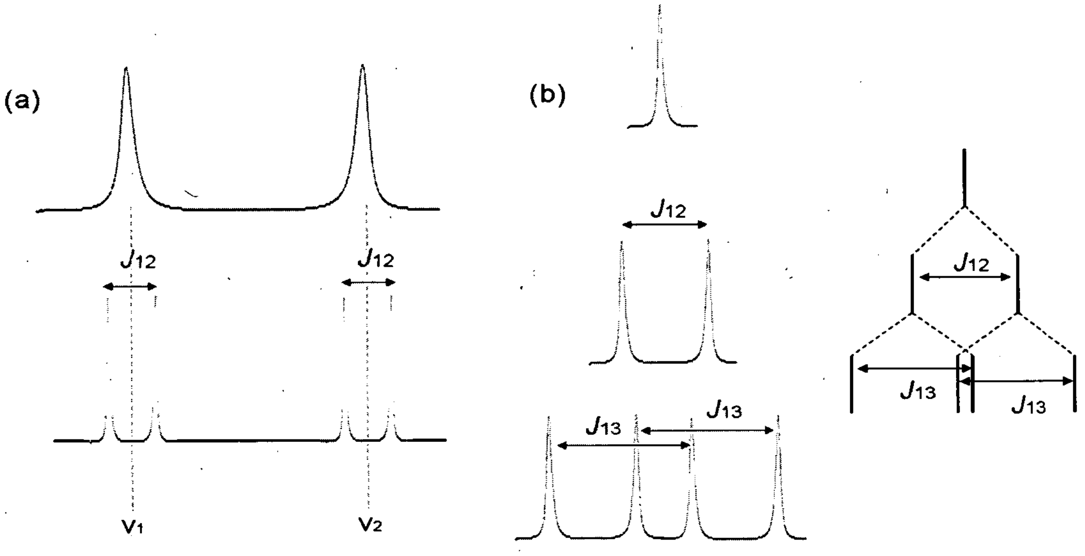


Fig.I.6.2 (a) Doublet formed by the scalar coupling. (b) Tree diagram to predict simple couplings among multiple spins

The spin-spin interactions are classified by the number of bonds between the two coupled nuclei. The couplings between two nuclei that are separated by two, three or four bonds are defined as geminal (^2J), vicinal (^3J) or long-range (^4J) couplings, respectively.

The geminal proton-proton (^2J) couplings are sensitive to α- and β-substituents. The α-substitution by an electronegative group leads to a positive change in the coupling constant. In contrast, the β-substitution gives a negative change in the coupling constant. Similarly to the geminal couplings, an electronegative substituent results in a negative change in vicinal (^3J) coupling constants. At the same time, vicinal coupling constants are dependent on the dihedral angles, Φ, between the coupled nuclei. The theoretical prediction of the coupling constant from the dihedral angle can be made by using the Karplus equation. In pyranose chairs of monosaccharides, the coupling constants between two axial protons are always larger than those between two equatorial protons or one axial and one equatorial proton. Karplus equations, appropriately parameterized for various molecular fragments, are a valuable tool in the conformational analysis of carbohydrates. The Karplus curve has the following general form:

$$^{3}J = \mathrm{A} + \mathrm{B}\cos\Phi + \mathrm{C}\cos^{2}\Phi \qquad \textbf{(Eq.I.6.7)}$$

Long-range (^{4}J) couplings are very small in magnitude but still can be observed especially when in the form of the W shape.

In addition to through bond spin-spin couplings, through space interactions exist between nuclei. These dipole-dipole (dipolar) couplings are the direct interactions between the nuclear moments of spins. The magnitude of the dipolar interaction depends on the distance between the two coupled nuclei and the angle formed by their internuclear vector and the external magnetic field (Eq.I.5.8). In solution, the angle θ varies with time because of the molecular tumbling. Therefore the integral of $(3\cos^{2}\theta\text{-}1)$ over time results in zero and the dipolar couplings vanish. No splitting of the signals is observed from the dipolar interactions.

$$\Delta \mathrm{B} = 3\mu(3\cos^{2}\theta\text{-}1)\mathrm{r}^{-3}(\mu_{0}/4\pi) \qquad \textbf{(Eq.I.6.8)}$$

Pulses and relaxation

In FT NMR spectroscopy the spectra are obtained by a Fourier transformation of the FID (Free Induced Decay). The FID is generated by rotating magnetic moments and is measured in the transverse plane. This precession in the xy plane starts when the magnetization vector is rotated away from the equilibrium position along the z axis. This rotation is achieved by applying a radiofrequency (RF) pulse along the x-axis producing a linearly oscillating magnetic field, B_1, with a transmitter frequency, ω_{tx}. The RF field becomes static when a nuclear spin system is rotating about the z-axis at the same frequency as the RF field. This frame of references is called rotating frame in which the B_1 field is stationary along the x-axis. The apparent frequency of the Larmor precession of the nucleus becomes $(\omega_0 - \omega_{tx})$ in the rotating frame. As a consequence, the apparent magnetic field along the z-axis changes to a reduced field, ΔB. The effective field, $\mathrm{B}_{\mathrm{eff}}$, in the rotating frame, is a vector sum of B_1 and ΔB. B_{eff} forms an angle, θ, with the z-axis

(Fig.I.6.3). The magnetization performs precession about the effective field, while the RF pulse is on.

$$\Delta B = -(\omega_0 - \omega_{tx})/\gamma \qquad \textbf{(Eq.I.6.9)}$$

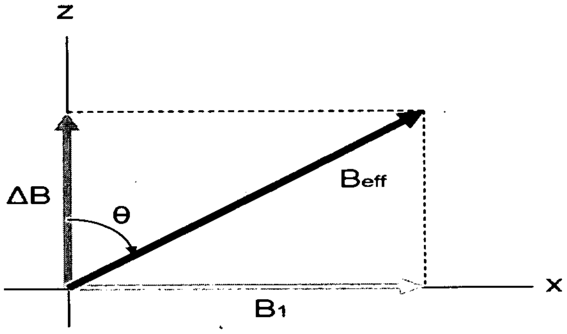


Fig.I.6.3 The effective field in the rotating frame.

When the transmitter frequency is the same as the nucleus spin Larmor frequency ($\omega_{tx} = \omega_0$), as well as when $B_1 >> \Delta B$, the angle $\theta = 90°$ and the effective field lies along the x-axis. The magnetization vector is rotated from the z-axis to the –y-axis. Similarly, an 180° x-pulse (inversion pulse) will rotate the magnetization vector from z-axis to the –z-axis. Both 90° and 180° x-pulses are commonly used in the NMR experiments.

During the RF pulse the equilibrium magnetization is rotated away from the z-axis and generates a non-zero transverse magnetization. After the RF pulse the transverse magnetization precesses around the external magnetic field inducing oscillating electric currents in the detection coils of NMR spectrometer. Relaxation precesses drive the z-magnetization back to the steady state along the z-axis and the x-, and y-magnetizations decay to zero (Fig.I.6.4). This process is characterized by two rate constants, $1/T_1$ and $1/T_2$, and referred to as longitudinal and transverse relaxation. T_1 and T_2 are the spin-lattice and spin-spin relaxation times, respectively.

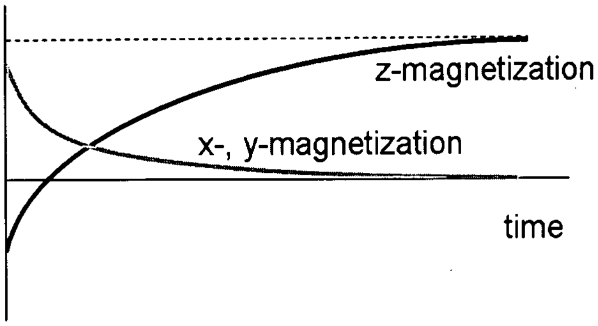


Fig.I.6.4 Magnetization changes during the relaxation

Product operators

A convenient way of describing the fate of coupled spin systems during a multi-pulse sequence is to follow the evolution of the density operator, $\rho(t) = |\Psi(t)\rangle\langle\Psi(t)|$, where $\Psi(t)$ is the wavefunction for an ensemble of spin systems. For an ensemble of spin-half nuclei, the matrix representing the density operator can be expressed as a linear combination of the orthogonal basis operators in the form of matrices:

$$\rho(t) = a_x(t)\hat{I}_x + a_y(t)\hat{I}_y + a_z(t)\hat{I}_z + a_E\hat{E} \qquad \textbf{(Eq.I.6.10)}$$

where $\hat{I}_x$, $\hat{I}_y$ and $\hat{I}_z$ are the spin operators representing the x-, y- and z-components of the spin angular momentum, $\hat{E}$ is the unit operator and the amount of M_i magnetization is given by the coefficient a_i. Neglecting relaxation effects the time evolution of the density operators can be given by the Hamiltonian, $\hat{H}$:

$$\rho(t) = \exp(-i\hat{H}t)\rho(0)\exp(i\hat{H}t) \qquad \textbf{(Eq.I.6.11)}$$

The Hamiltonian determines how the spin evolves in time, which is different during pulses (**Eq.I.5.12**) and in free precession (**Eq.I.5.13**).

$$\hat{H}_{\text{free}} = \Omega\hat{I}_z \qquad \textbf{(Eq.I.6.12)}$$

$$\hat{H}_{\text{x-pulse}} = \Omega\hat{I}_z + \omega_{tx}\hat{I}_x \qquad \textbf{(Eq.I.6.13)}$$

When two spins coupled to each other with a coupling constant J_{12}, the coupling Hamiltonian becomes:

$$\hat{H}_J = 2\pi J_{12}\tau\hat{I}_{1z}\hat{I}_{2z} \qquad \textbf{(Eq.I.6.14)}$$

During the period of free precession, the rotation of a spin operator can be represented by the diagrams shown in Fig.I.6.5.

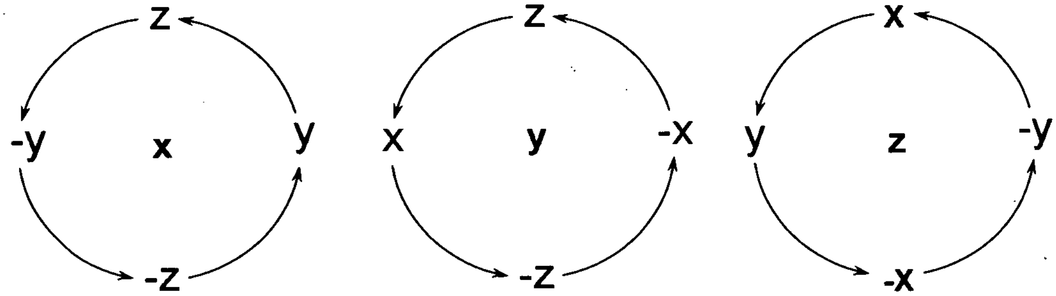


Fig.I.6.5 Determination of a spin operator rotation about x, y and z. The 'new operator' is obtained by rotating the 'initial operator' according to the arrow direction. The result of the rotation is [cosθ × (initial operator) + sinθ × (new operator)].

When two spins are coupled to each other, two set of operators ($\hat{I}_x$, $\hat{I}_y$, $\hat{I}_z$ and $\hat{E}$) are used to describe each of the spins. The possible combinations of the operators in a two-spin system are listed in Table.I.6.1. The scalar coupling between the two spins causes the interconversion between in- and antiphase magnetization. This coupling evolution can be used to transfer the coherence between the two spins through the antiphase terms. A diagram predicting the result of the scalar coupling evolution is shown in Fig.I.6.6. During acquisition antiphase operators evolve into observable in-phase operators (single-quantum coherence) yielding splitting because of the coupling.

Table I.6.1 Possible products of operators for a *J* coupled two-spin system.

	z-magnetization	In-phase x-, y-magnetizations (SQ)	Antiphase x-, y-magnetizations	Multiple-quantum coherences	Non-equilibrium population
Spin 1	$\hat{I}_{1z}$	$\hat{I}_{1x}$, $\hat{I}_{1y}$	$2\hat{I}_{1x}\hat{I}_{2z}$ $2\hat{I}_{1y}\hat{I}_{2z}$	$2\hat{I}_{1x}\hat{I}_{2x}$ $2\hat{I}_{1x}\hat{I}_{2y}$ $2\hat{I}_{1y}\hat{I}_{2x}$ $2\hat{I}_{1y}\hat{I}_{2y}$	$2\hat{I}_{1z}\hat{I}_{2z}$
Spin 2	$\hat{I}_{2z}$	$\hat{I}_{2x}$, $\hat{I}_{2y}$	$2\hat{I}_{2x}\hat{I}_{1z}$ $2\hat{I}_{2y}\hat{I}_{1z}$		

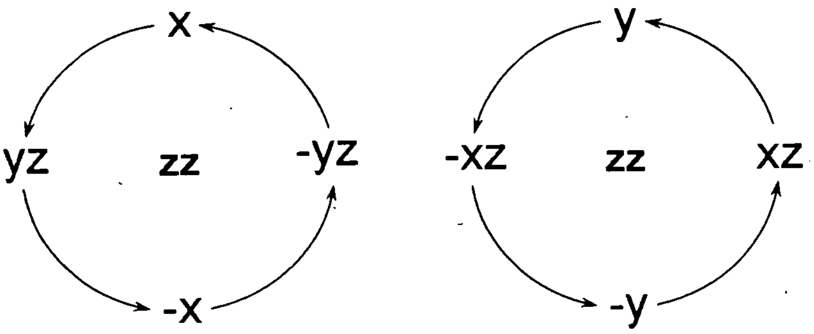


Fig.I.6.6 Interconversion of the scalar coupling during a time period, t, between in-phase (x or y) and antiphase (xz or yz). The 'new phase' is obtained by rotating the 'initial phase' according to the arrow direction. The evolution result is [cos(πJt) × (initial phase) + sin(πJt) × (new phase)].

The antiphase terms are essential in the multi-pulse NMR sequences because they can be used in both coherence transfers and generation of multiple-quantum coherences. The antiphase term is only generated during a scalar coupling evolution period. When a pulse is applied to the antiphase term, multiple-quantum (a mixture of zero- and double-quantum) coherences will be generated. A specific coherence can be selected by phase cycling or pulsed field gradients to generate required signals while a mixture of the unwanted coherences is eliminated.

Building blocks

The most commonly used component of multi-pulse NMR experiments is a spin-echo (Fig.I.6.7). The spin-echo sequence starts with a $90°_x$ pulse followed by two equal time evolution periods with an $180°_x$ pulse in the middle. Both offsets (*Ω*, chemical shift) and scalar couplings (*J*) evolve independently during the whole pulse sequence. The special feature of spin-echo is that at the end of the sequence the offset is refocused but the coupling evolves during the whole time period, 2τ.

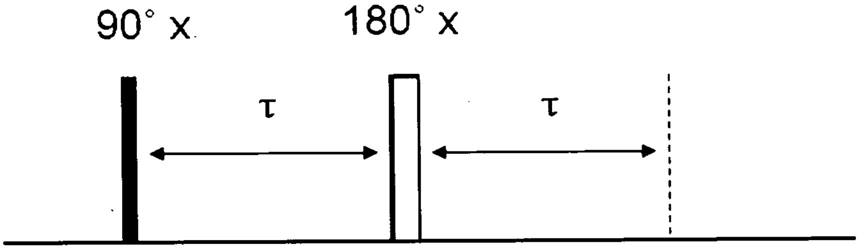


Fig.I.6.7 Spin-echo pulse sequence. The sequence ends after the second delay τ, at the point of the dashed line. The delay τ can be optimized for the scalar coupling evolution

In the following the evolution of homonuclear two-spin is followed by using the product of operators. The magnetization is rotated from z-axis [$\hat{I}_z$] to -y-axis [$-\hat{I}_y$] by the first $90°_x$ pulse. The magnetization evolves during the first delay τ. Offsets of different spins make the precession rates different which results in different effective evolution [$-\cos(\Omega\tau)\hat{I}_y + \sin(\Omega\tau)\hat{I}_x$]. The $180°_x$ pulse rotates the magnetizations around the x-axis to the mirror side of the xz-plane. This is reflected in change of the sign of the $\hat{I}_y$ term [$\cos(\Omega\tau)\hat{I}_y + \sin(\Omega\tau)\hat{I}_x$]. The magnetization keeps evolving during the second delay τ. Because the two delays are equal, the magnetizations evolve back to the y-axis [$\hat{I}_y$] irrespective of the offset. In other words, the chemical shift of the evolution is refocused during a spin-echo.

Considering the coupling evolution for spin1 during the first delay τ, the in-phase term [$-\hat{I}_{1y}$] evolves into the antiphase term [$-\cos(\pi J_{12}\tau)\hat{I}_{1y} + \sin(\pi J_{12}\tau)2\hat{I}_{1x}\hat{I}_{2z}$] as predicted from the diagram shown in Fig.I.6.6. The $180°_x$ pulse changes the signs of both terms [$\cos(\pi J_{12}\tau)\hat{I}_{1y} - \sin(\pi J_{12}\tau)2\hat{I}_{1x}\hat{I}_{2z}$], yielding [$\cos(2\pi J_{12}\tau)\hat{I}_{1y} - \sin(2\pi J_{12}\tau)2\hat{I}_{1x}\hat{I}_{2z}$] at the end of the second delay τ. The same evolution applies to spin2. The final result of the scalar coupling evolution during the spin-echo is effectively an evolution of the coupling for the time 2τ combined with spin inversion. The in- and antiphase terms interconvert between each other independently of the chemical shift evolution. This is referred as the *J*-modulated spin-echo. By optimizing the delay τ, exclusive in- or antiphase doublets can be obtained.

Heteronuclear two-spin system (I, S) can be easily manipulated to yield different spin states depending on the combination of I and S 180° pulses (Fig.I.6.8). The application of 180° pulses on both nuclei (Fig.I.6.8a), has the same effect as the application of a single 180° pulse to a homonuclear spin system (Fig.I.6.7). In addition, in heteronuclear spin-systems the 180° pulse can be applied only on one of the two nuclei, I or S (Fig.I.6.8.b,c).

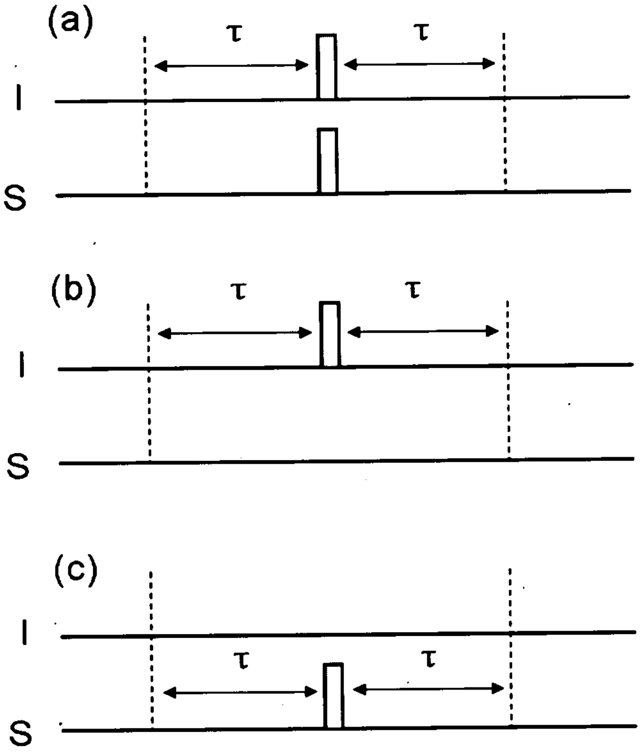


Fig.I.6.8 Spin-echo in a heteronuclear spin system. The 90° pulse is omitted compared with the pulse sequence in Fig.I.6.7. The spin-echo is represented by the blocks in between the two dashed lines. (a) 180° pulse on both spins; (b) $180°_x$ pulse only on I spin; (c) $180°_x$ pulse only on S spin.

Taking the sequence in Fig.I.6.8b as an example, I spin experiences the complete spin-echo and its chemical shift is refocused. However, in the absence of the 180° pulse, the chemical shift of S spin will evolve during the whole period of delay 2τ. Considering the coupling evolution the term $-\hat{I}_y$ will have the form of $[-\cos(\pi J_{IS}\tau)\hat{I}_y + \sin(\pi J_{IS}\tau)2\hat{I}_x\hat{S}_z]$ before the 180° pulse. The 180° pulse only affects the I spin, changing the sign of the first term and yielding $[\cos(\pi J_{IS}\tau)\hat{I}_y$ +

$\sin(\pi J_{IS}\tau)2\hat{I}_x\hat{S}_z]$. After the second τ delay, only the $\hat{I}_y$ operator is present meaning that the heteronuclear coupling is refocused. This is also true for the S spin.

The INEPT (Insensitive Nuclei Enhanced by Polarization Transfer) sequence (Fig.I.6.9) is usually used for the coherence transfer from the sensitive nucleus (I) to the insensitive one (S) through the antiphase terms. This polarization transfer increases the intensity of the insensitive nucleus by a factor of γ_I/γ_S, which is 33% more than that obtained from the heteronuclear NOE $(1 + \gamma_I/2\gamma_S)$ for $^{13}C/^{1}H$ only.

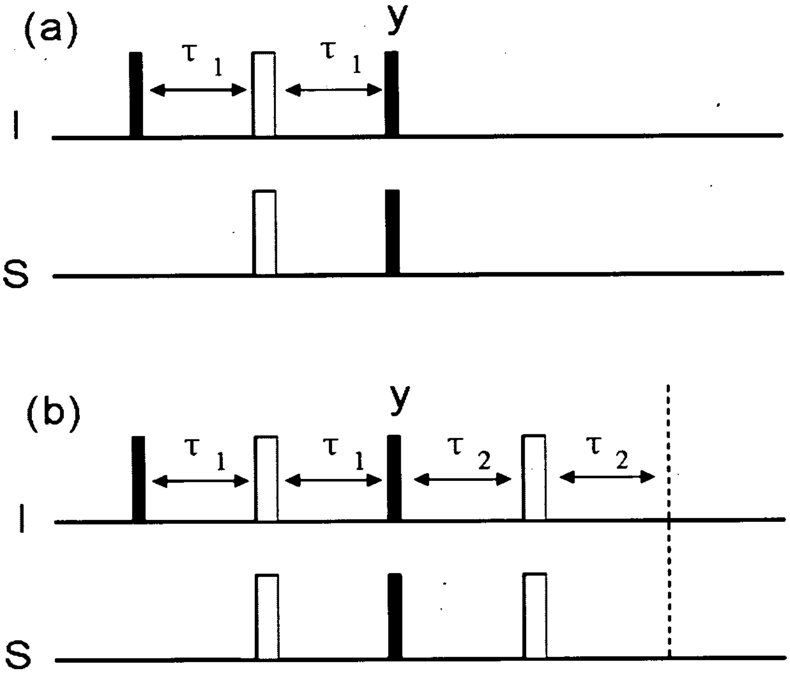


Fig.I.6.9 INEPT pulse sequences. All pulses are from the x-axis unless otherwise indicated; (a) the coherence ends up with an antiphase term; (b) the coherence evolves into an in-phase term

Before the second 90° pulse (a spin-echo sequence), an antiphase coherence $[-\sin(2\pi J_{IS}\tau_1)2\hat{I}_x\hat{S}_z]$ is generated through scalar coupling evolution while the chemical shift is refocused. The 90°_y pulse on the I spin and 90°_x pulse on the S spin transfer the antiphase term from I to S $[-\sin(2\pi J_{IS}\tau_1)2\hat{I}_z\hat{S}_y]$. At this stage, as shown by the sequence in Fig.I.6.9a, antiphase doublet can be observed on S spin directly. By adding another spin-echo after the 90° pulses (Fig.I.6.9b), the antiphase magnetization evolves into in-phase during the delay $2\tau_2$. Both τ_1 and τ_2 can be optimized to obtain optimal sensitivity for spin S.

INADEQUATE (Incredible Natural Abundance DoublE-QUAntum Transfer Experiment) uses double-quantum (DQ) coherences to correlate homonuclear chemical shifts and to measure the ^{13}C-^{13}C coupling constants (J_{CC}). The spectrum only contains signals from adjacent ^{13}C-pairs, which is a very good way to establish the carbon connectivity of the molecule. The generation of the multiple-quantum coherence can be achieved by the pulse sequence shown in Fig.I.6.10.

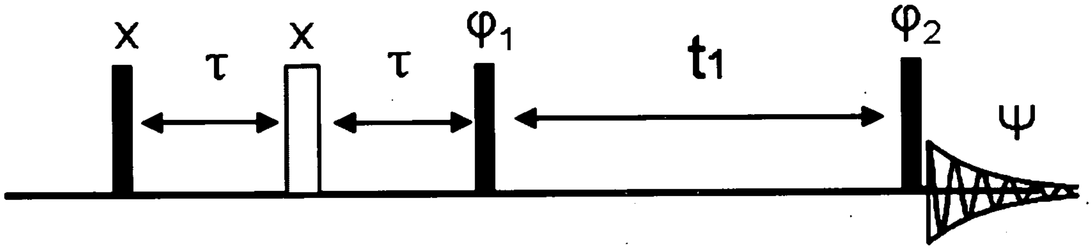


Fig.I.6.10 2D pulse sequence for DQ correlation; $\varphi_1 = 2(x, -x)$, $\varphi_2 = x, y, -x, -y$ and $\psi = x, -y, -x, y$.

The antiphase magnetization is obtained by the initial spin-echo during which the chemical shifts are refocused and J_{CC} couplings evolve during the time of 2τ. The antiphase term is then converted into multiple-quantum coherence by the second 90° pulse. The DQ coherence is selected either by phase cycling or pulsed field gradients and evolves for time t_1, before being converted back to single-quantum coherence by the last 90° pulse. The unwanted signals from isolated ^{13}C spin are discarded by a coherence selection technique. Antiphase doublets in the directly detected dimension can be recorded immediately. Alternatively, a second spin-echo can be appended to evolve the antiphase into in-phase terms to obtain the in-phase doublets. Large $^1J_{CC}$ coupling constants can be extracted from the frequency splitting. The same pulse sequence can be optimized also for $^nJ_{CC}$ coupling constants yielding long-range carbon-carbon correlations. In this case the determination of coupling constants is not straightforward due to poor separations of the two lines of the doublet. Solutions to this problem are discussed in Chapter III.3.

As mentioned above coherence selection can be achieved by using phase cycling or the pulsed field gradients (PFG) technique. Each method has strengths and limitations. Phase cycling selects the required coherence by repeating the experiments several times with different phases of pulses and receiver. In each transient, both required and unwanted signals are recorded with a particular phase. When adding up transients, only the required coherences are added up coherently while the unwanted ones cancel out. Phase cycling thus relies on the cancellation of the unwanted signals. This might be not complete, especially when the unwanted signals are much stronger than the required ones. The other drawback of the phase cycling is the prolonged experimental time as each t_1-increment of a 2D experiment must be recorded using the complete phase cycle.

The PFG method relies on the creation of a magnetic field during a short period, which changes the main magnetic field B_0 in a linear way along an arbitrary axis, usually the z-axis (Fig.I.6.11). Both the strength and polarity of PFGs can be controlled to select required coherences. This variation of the homogeneity of B_0 causes different phase shifts to be experienced by individual spins located in different parts of the sample. As a result, the transverse magnetization is averaged to zero (Fig.I.6.11).

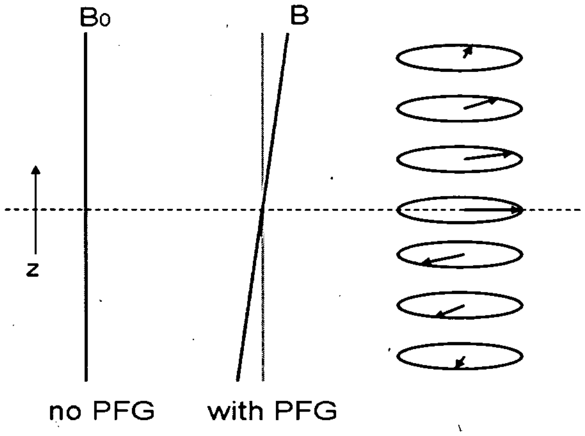


Fig.I.6.11 Effect of a pulsed field gradient on the main magnetic field B_0 and the transverse magnetization. The result field during the PFG varies linearly along the z-axis. Transverse magnetizations cancel out across the whole sample because of the dephasing.

When using a pair of PFGs, the coherences are dephased by the first and rephased by the second pulsed field gradient. The point is to control the ratio of the two PFGs to select certain coherence (**Eq.I.6.15**):

$$\Phi_z = \Sigma(-\gamma P_i G_i z \tau_i) = 0 \qquad \textbf{(Eq.I.6.15)}$$

where P_i is the coherence order and other symbols have the usual meaning.

PFGs can only de/rephase transverse magnetizations. Therefore the z-magnetization always survives during PFGs. The coherence selection by PFGs in a single scan is the strength of this method compared with the phase cycling method. However, usually only half of the signal can be selected by coherence selecting PFGs, which leads to 50% reduction of signal intensities. This loss can occasionally be eliminated or reduced. The first instance applies to sensitivity-enhanced experiments [51]. The signal-to-noise loss of only 29% is sustained when PFGs are used for the selection of echo-antiecho coherences during t_1 as is the case for INADEQUATE sequences [52].

Some routine NMR experiments for the study of molecular structures include COSY (TOCSY), NOESY (ROESY), HSQC and HMBC. COSY and NOESY experiments are commonly used to establish the connectivity of the molecular frame through chemical bonds and distances in space, respectively. HSQC and HMBC experiments provide one-bond and long-range heteronuclear chemical shift correlations.

COSY (Correlation Spectroscopy) and TOCSY (Total Correlation Spectroscopy) are homonuclear experiments (Fig.I.6.12) mostly used to trace the proton-proton couplings and to correlate chemical shifts of coupled protons. The difference between the two experiments is in the distance which they transfer the magnetizations. COSY spectrum only records signals from coupled protons. TOCSY spectrum can trace all protons in the same spin system with unbroken chain of couplings. The mixing period of COSY experiment consists of a single 90° pulse which transfers the magnetization between spins.

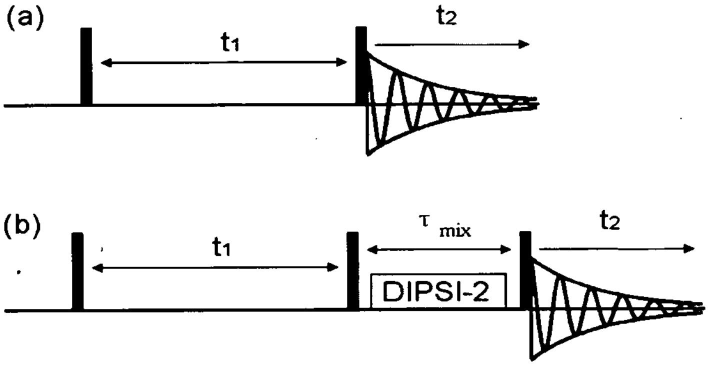


Fig.I.6.12 Pulse sequences of COSY (a) and TOCSY (b) experiments.

In both experiments, the transverse magnetization evolves during the t_1 period coding both chemical shift and scalar couplings evolution. The COSY cross peaks show the active couplings in antiphase, while the passive spins are in-phase. The cross peaks and the diagonal peaks are shifted by 90° in-phase in the basic experiment. Double-quantum filtration, typically using PFGs for coherence selection, produces pure phase cross peaks and diagonal peaks and is typically used. In TOCSY experiment, the second 90° pulse is followed by a spin-echo mixing period, τ_{mix}. The magnetization is transferred between all protons of one spin system during the mixing achieved by a mixing sequence, e.g. DIPSI-2 [53].

NOESY (Nuclear Overhauser Effect Spectroscopy) and ROESY (Rotating frame Nuclear Overhauser Effect Spectroscopy) experiments (Fig.I.6.13), respectively, generate cross peaks from the cross relaxations of z-magnetizations and spin-locked tilted magnetizations between spins close in space.

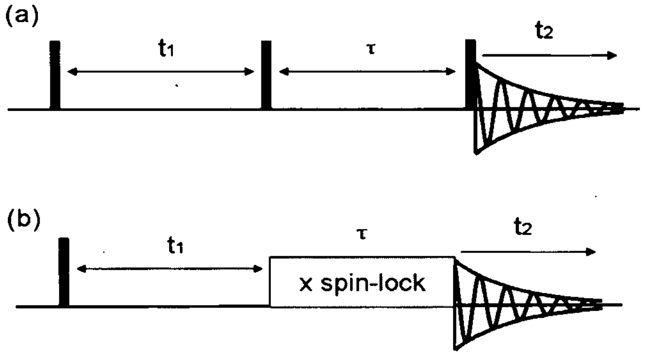


Fig.I.6.13 Pulse sequences of NOESY (a) and ROESY (b) experiments.

In both experiments, the frequency-labeled magnetization during the t_1 period is subject to a mixing period that transfers the magnetization between spins through dipolar relaxation. The 90° pulse rotates the magnetization back to the z-axis in NOESY, where it undergoes cross relaxation during the τ mixing period transferring the z-magnetization to other spins close in space. The problem of NOESY transfer is that its efficiency depends on the molecular tumbling and the magnetic field strength and that it can change the sign. For certain regimes no NOE is observed. This can be avoided in ROESY experiment by using the cross-relaxation of the transverse magnetizations. After the frequency labeling period t_1, the x-magnetization is locked along the effective axis using a weak spin-lock, where it undergoes the cross relaxation. The cross peaks in ROESY spectra are always positive, while the diagonal peaks are negative.

HSQC (Heteronuclear Single-Quantum Correlation) experiment (Fig.I.6.14) is widely used to correlate the proton and carbon chemical shifts. The magnetization is generated from the I spin

(^1H) and transferred to the S spin (^{13}C) using the previously described INEPT sequence. During the t_1 period, the spin-echo on I spins refocuses the heteronuclear couplings but the offsets of S spins evolve during the whole t_1. The following two 90° pulses transfer the magnetization back to I spin and antiphase signals can be detected during t_2. In practice, unless the $^1J_{CH}$ coupling constants are to be measured, an additional spin-echo follows and the signal is detected with ^{13}C decoupling.

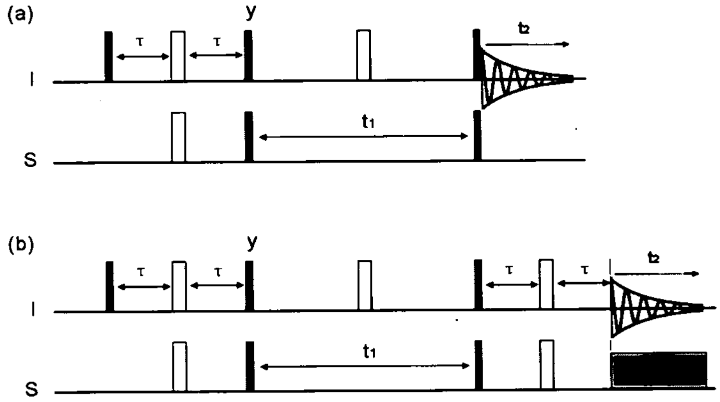


Fig.I.6.14 Pulse sequences for heteronuclear correlation HSQC experiments (b) with refocusing period and decoupling during acquisition

The HMBC (Heteronuclear Multiple-Bond Correlation) experiment (Fig.I.6.15) is used to establish the long-range correlation and to measure the long-range coupling constants. This experiment starts by rotating the $\hat{I}_z$ magnetization to the transverse planes, where couplings evolve for the period of τ generating antiphase terms. The 90° pulse on S spin converts the antiphase magnetization into multiple-quantum coherences, which evolve for t_1 time. During the t_1 period, the I spin offset is refocused and the multiple-quantum coherence is labeled by the effect of the S spin. The final 90° pulse on the S spin converts the multiple-quantum coherences back to single-quantum proton coherences, which are recorded during t_2 on the I spin yielding antiphase correlation peaks. Cross peaks in HMBC spectra have mixed phase because of proton chemical shift and proton-proton coupling evolution during τ delay.

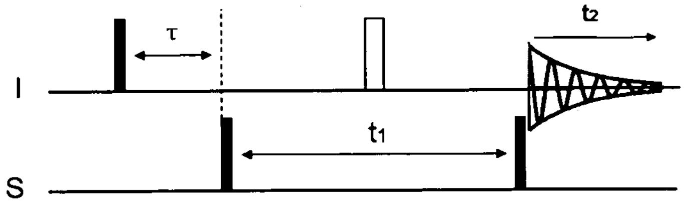


Fig.I.6.15 Pulse sequence of HMBC experiment

I. 7 Alignment system

In addition to scalar couplings, spins are also coupled via dipolar couplings. The dipolar interaction comes from direct, through space, interaction of magnetic moments of nuclei. Similarly to *J* couplings, two different orientations of spin-half nuclei cause the coupled spin to experience different local magnetic fields. This time, however, the strength of the interaction depends on the distance between the two spins and the immediate orientation of this internuclear vector with regard to the external magnetic field (**Eq.I.7.1**, **Eq.I.7.2** and Fig.I.7.1). The dipolar coupling between the two spins is only derived from the z-component of the local magnetic field of the two spins. In isotropic solutions the direct dipole-dipole couplings cancel out because of thermal motion and random molecular tumbing during the acquisition time. The integral given in **Eq.I.6.2** is zero and no splitting of signals can be observed. The dipolar interactions only contribute to the relaxation in isotropic solutions (Fig.I.7.1). The dipolar interactions reflect the mutual relative positions of spins in space, providing useful long-range structural information about the molecules. If dipolar coupling constants could be measured in solution, they can be used to complement the short range structural information obtained from NOE.

$$D_{AX} = (\mu_0/8\pi^2)(\hbar\gamma_A\gamma_X/r^3)(3\cos^2\theta - 1)$$ **(Eq.I.7.1)**

$$\int_0^{\pi} 3(\cos^2\theta - 1)\sin\theta d\theta = 0$$ **(Eq.I.7.2)**

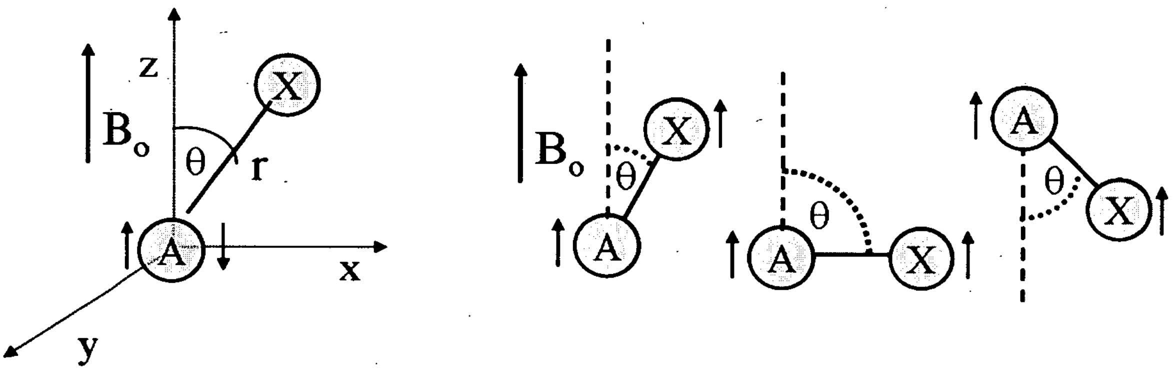


Fig.I.7.1 Dipolar interaction between two spins A and X in isotropic solution

The cancellation of the dipolar interactions in solution comes from the random Brownian motion. If this motion can be restricted slightly, the dipolar interaction will not vanish completely. Minor restriction of the Brownian motion can be achieved by using liquid crystals as co-solvents to dissolve the solute. Through the interactions with the liquid crystal the solute molecule is forced to adopt preferably certain orientations, while still undergoing relatively free tumbling in the nematic phases formed by the liquid crystals (Fig.I.7.2).

This interaction is either electrostatic or steric, and generates preferred angles between the internuclear vectors and the external magnetic field. The small resulting dipole-dipole coupling reflects this time averaged angle. Relating this angular information for many bonds of the molecule leads to structural information. By controlling the strength of the alignment media, the dipolar interactions are recovered only slightly without producing complicated spectra. The residual dipolar coupling (RDC) constants can be obtained from the difference between the splitting observed in the aligned media ($J + D$) and the scalar coupling constant (J) observed in isotropic samples.

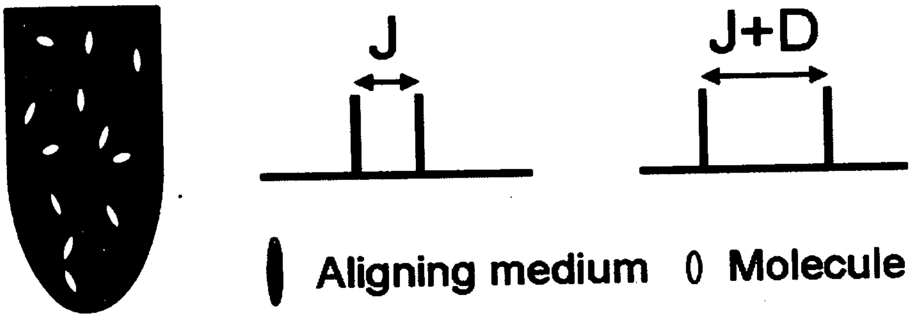


Fig.I.7.2 Schematic representation of the aligned sample and splitting obtained in isotropic and anisotropic solutions

Some frequently used liquid crystalline media include phospholipid bicelles (DMPC/DHPC or DIODPC/CHAPSO) [54], filamentous phages [55], cetylpyridinium chloride/bromide and hexanol with NaCl/NaBr [56], purple membrane fragments [57] and PEG (n-alkyl-poly(ethylene glycol))/n-alkyl alcohol mixtures [58, 59]. Each medium has strengths and limitations in the application conditions, which are compared in Table I.7.1. Almost all media are dependent on temperature, salt concentration and pH. The strength of the media can be conveniently monitored via the quadrupole splitting of the deuterium signal of D_2O. For the same medium the splitting of the quadrupole signals is very sensitive to the temperature.

Table I.7.1 Comparison of different liquid crystals applied as alignment media.

	Phospholipid bicelle	Filamentous phage	CPBr/hexanol	Purple membrane	PEG ($C_{12}E_5$)
Charge	none	negative	none	negative	none
pH region	6-7	6.5-8	2.2-8.1	2.5-10	insensitive
Salt condition (mM)	≤ 300	~ 100	5-200 CPCl: 200-500	~ 5000	little sensitive
Concentration	3-15% (w/v)	~ 60 (mg/ml)	1-10 % (w/v)	7-33 (mg/ml)	3-25% (wt%)
Temperature (°C)	22-50	5-45	15-60	-269-69	10-40
Splitting (Hz)	5-27	~ 50	0-21	2-14	13-138[i]
Limitation	DMPC/DHPC: Irreversible degradation at low pH [60] DIODPC/CHAPSO: nonstable at neutral pH	Aggregate below pH 6	High salt concentration of NaCl with CPCl/hexanol	Binding tightly with solute	Solutes are difficult to recover from the lamellar phases

The PEG/hexanol medium was used as the dilute liquid crystalline solvent to form the bilayer surface, which is oriented parallel to the external magnetic field [61]. The PEG molecule, $C_{12}E_5$ (3,6,9,12,15-pentaoxaheptacosan-1-ol), has 12 carbons in the n-alkyl group and 5 glycol units. When mixed with n-hexanol and different amounts of water, different strengths of alignment media can be prepared. The strength of the medium can be characterized by measuring the residual quadrupole splitting of the deuterium signal of D_2O using 1D ^2H spectrum. The composition of the alignment medium is characterized by the weight percentage ratio, *wt%*, of the $C_{12}E_5$ to the D_2O and the molar ratio, *r*, of $C_{12}E_5$ to n-hexanol. A medium with smaller r is stable over a smaller range of temperatures. The aligning strength of the same medium is weaker at higher temperatures and reflected by smaller splitting of the deuterium signal.

The splitting caused by the residual dipolar coupling of a pair of spin-half nuclei A and X can be expressed by the following equation:

---

[i] 138 Hz was used for a very strong alignment required for the ^{13}C-detected INADEQUATE experiment development in this project, 15-40 Hz is the normal range according to [59]

$$D_{AX} = (\mu_0/4\pi^2)\ (\hbar\gamma_A\gamma_X/r^3) < (3\cos^2\theta-1)/2> \qquad \textbf{(Eq.I.7.3)}$$

where γ_A and γ_x are the gyromagnetic ratios, r_{AX} is the internuclear distance and θ is the angle between the internuclear vector and the external magnetic field, brackets indicate average due to motion. This anisotropic averaging can be simplified and expressed by the Saupe order matrix [62], which is a symmetric and traceless 3 × 3 matrix consisting of S_{ij} elements corresponding the molecular coordinates defined in an arbitrary Cartesian coordinate system.

$$S_{ij} = 3< (\cos\theta_i\cos\theta_j>/2-k_{ij}/2 \qquad \textbf{(Eq.I.7.4)}$$

$$S_{ij} = S_{ji} \qquad \textbf{(Eq.I.7.5)}$$

$$S_{xx} + S_{yy} + S_{zz} = 0 \qquad \textbf{(Eq.I.7.6)}$$

$$<P_2(\cos\theta)> = \Sigma\ (S_{ij}\cos\Phi_i\cos\Phi_j) \qquad \textbf{(Eq.I.7.7)}$$

i, j = {x, y, z}, k_{ij} is the Kronecker delta function; Φ and θ are defined as follows. The internuclear vector within the molecule is defined by the direction cosines (Fig.I.7.3), $\cos\Phi_i$ and $\cos\Phi_j$, which describe the internal geometries of the molecule. Therefore, the solution of the matrix parameters gives the molecular structure information.

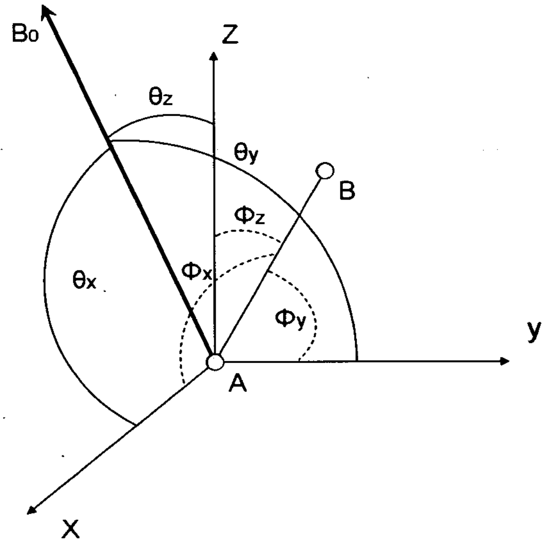


Fig.I.7.3 Definition of the direction cosines of the internuclear vector AB within arbitrary molecular axis (angles Φ) and the magnetic field axis (angles θ) in the arbitrary molecular frame

The definition of the Saupe order matrix requires only five independent parameters, which can be solved by measuring five independent RDCs of the molecule. The solved matrix can be further diagonalized to obtain the axial and rhombic parameters and a transformation matrix. The determination of this Saupe's order tensor relates the principle frame of the alignment tensor to the arbitrary molecular system. This procedure is carried out with Singular Value Decomposition [63] (SVD) method by solving a series of linear equations.

$$\mathrm{Ax} = \mathrm{b} \qquad \textbf{(Eq.I.7.8)}$$

$$\begin{pmatrix} \cos^2\phi_y^1 - \cos^2\phi_x^1 & \cos^2\phi_z^1 - \cos^2\phi_x^1 & 2\cos\phi_x^1\cos\phi_y^1 & 2\cos\phi_x^1\cos\phi_z^1 & 2\cos\phi_y^1\cos\phi_z^1 \\ \cos^2\phi_y^2 - \cos^2\phi_x^2 & \cos^2\phi_z^2 - \cos^2\phi_x^2 & 2\cos\phi_x^2\cos\phi_y^2 & 2\cos\phi_x^2\cos\phi_z^2 & 2\cos\phi_y^2\cos\phi_z^2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cos^2\phi_y^n - \cos^2\phi_x^n & \cos^2\phi_z^n - \cos^2\phi_x^n & 2\cos\phi_x^n\cos\phi_y^n & 2\cos\phi_x^n\cos\phi_z^n & 2\cos\phi_y^n\cos\phi_z^n \end{pmatrix} \begin{pmatrix} S_{yy} \\ S_{zz} \\ S_{xy} \\ S_{xz} \\ S_{yz} \end{pmatrix} = \begin{pmatrix} RDC_1 \\ RDC_2 \\ \cdot \\ \cdot \\ \cdot \\ RDC_n \end{pmatrix}$$

The tensor parameters (S_{xx}, S_{yy} and S_{zz}) can be plotted in the manner of the Sanson-Flamsteed projection (Fig.I.7.4) to represent the orientation of the principle axis system.

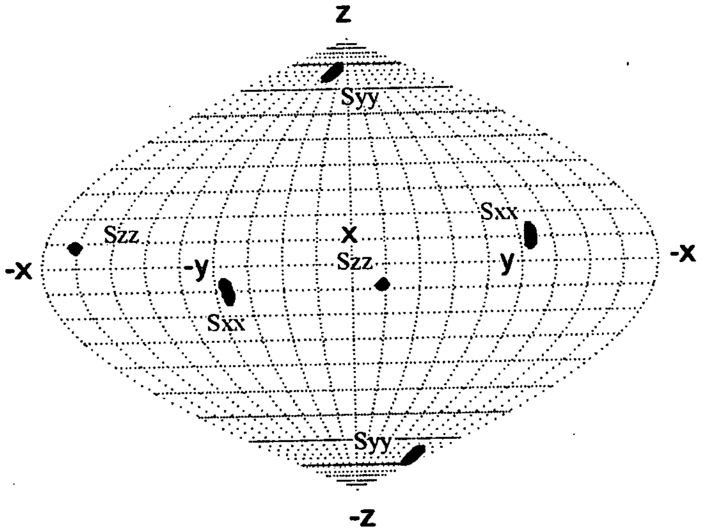


Fig.I.7.4 Sanson-Flamsteed projection represents the orientation of the principle axis system

Chapter II

Materials and Methods

II. 1 Heparin-derived di- and tetrasaccharides

One disaccharide and three tetrasaccharides (Fig.II.1.1) were prepared by enzymatic digestion of low molecular weight heparin (LMWH) (Innohep; Leo Laboratories, Princes Risborough, U.K.) by our collaborators in the University of Manchester.

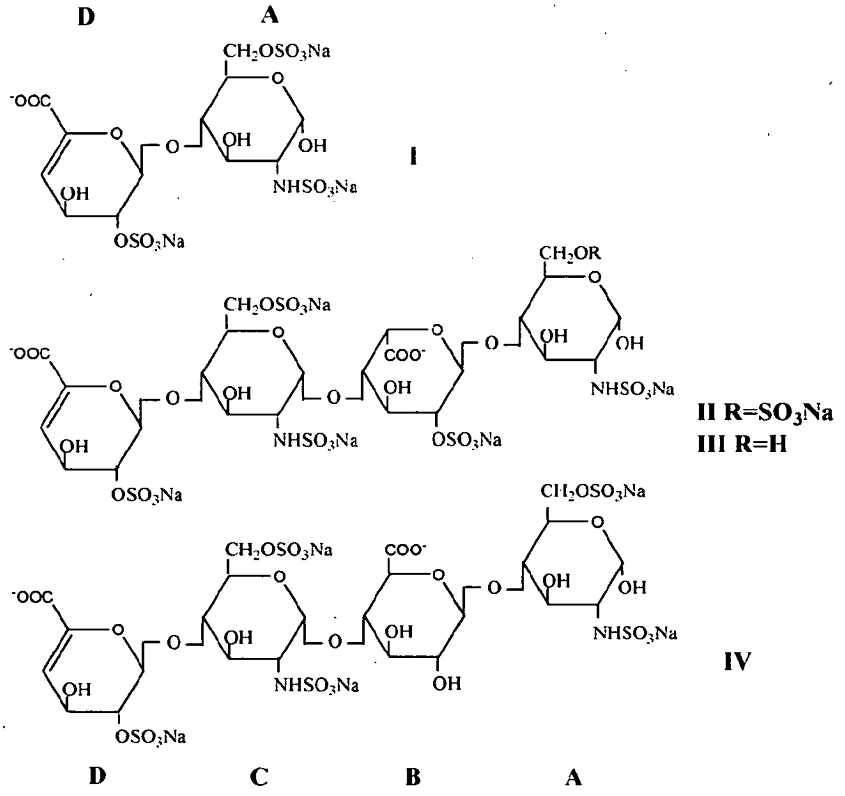


Fig.II.1.1 Structures of the heparin-derived disaccharide (**I**) and three tetrasaccharides (**II-IV**) used in this study. ΔUA(2S)-GlcNS(6S)-Na$_3$ (**I**), ΔUA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S)-Na$_6$ (**II**), ΔUA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS-Na$_6$ (**III**) and ΔUA(2S)-GlcNS(6S)-GlcA-GlcNS(6S)-Na$_5$ (**IV**). Letters **A-D** are used to label individual rings starting from the reducing end (**A**) and progressing towards the unsaturated non-reducing end (**D**) of the oligosaccharide.

The final digest was resolved into its constituent oligosaccharide size fractions by gel filtration chromatography on a Bio-Gel P10 (Bio-Rad Laboratories, Hemel Hempstead, UK) column (2.5 x 120 cm) eluted with 0.25 M NH_4HCO_3 at a low rate of 10 ml.hr with collection of 4 ml fractions. (Fig.II.1.2a)

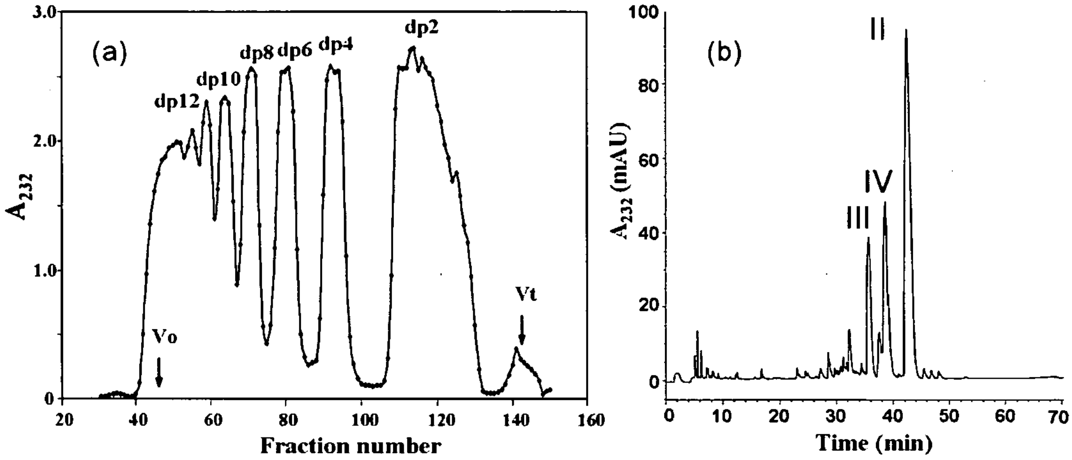


Fig.II.1.2 (a) The initial Bio-Gel P10 size fractionation of the heparin oligosaccharides; (b) the strong anion-exchange HPLC separation of the tetrasaccharides.

The elution profile was monitored by UV absorption at 232 nm, specific to the unsaturated nonreducing ends of the oligosaccharides introduced by the lyase action of the heparinase enzyme. The disaccharide and tetrasaccharide peaks were separately pooled, freeze-dried twice, desalted on PD-10 columns (Amersham Biosciences Ltd, Chalfont St. Giles, UK) and then dried. The di-and tetrasaccharide pools were resolved into their constituent species by strong anion-exchange HPLC chromatography on an IonPac AS17 column (0.4 x 25 cm; Dionex, Camberley, UK) (Fig.II.1.2b). After a brief wash with pH3.5 water, the column was eluted with a 60 ml linear gradient of 0 – 1.0 M NaCl, pH3.5, at a flow rate of 1 ml/min with collection of 1 ml fractions. The disaccharide compositions of specific oligosaccharides were determined by complete digestion to disaccharides with a mixture of heparinases I, II (*F. heparinum*; no EC number assigned) and III (*F. heparinum*; heparitin-sulfate lyase EC 4.2.2.8) (Grampian Enzymes, Orkney, UK) in 50 mM sodium acetate, 0.5 mM calcium acetate, pH 7.2 at 37oC. Digests were then applied to a ProPac PA1 strong-anion exchange HPLC column (0.4 x 25cm; Dionex). After a brief wash with pH3.5 water, they were eluted on a 45 ml gradient of 0 – 1.0 M NaCl, pH3.5 at a flow rate of 1ml/min. Disaccharides were monitored by on-line absorption at 232 nm and were identified by comparison with the elution positions of known disaccharide standards.

For NMR experiments, each sample was dissolved in 0.5 ml deuterium oxide (D_2O >99.96%) and lyophilized to remove exchangeable protons. The powder was then redissolved in 320 µl D_2O to be used as an isotropic sample for NMR experiments. Each sample was transferred into a 5 mm Shigemi tube in order to maximize the sensitivity. Approximately 25 µg for each sample was used for the primary structure determination.

Heparin-derived fully sulfated tetrasaccharide (compound **II**) was selected for the measurement of scalar and dipolar coupling constants to characterize its conformation in solution. 550 µg were dissolved in 320 µl D_2O. In the initial sample, the proton peaks of B ring were too broad to obtain reliable coupling constants. This was the result of the equilibrium of the two exchangeable conformations. In order to slow down this exchange [64] $CaCl_2$ was added into the isotropic sample in concentration 4 times larger than that of the tetrasaccharide **II**. The pH value was adjusted to 7.5 by adding NaOH dissolved in D_2O. The sample was transferred back into a 5 mm Shigemi tube yielding sharp peaks suitable for the coupling constants measurements. Identical conditions were applied to the aligned sample using an open Shigemi tube without the plunger. The composition of the solvent was 18.6 µl $C_{12}E_5$, 6.2 µl hexanol and 430 µl D_2O.

II. 2 Compounds used for the development of NMR methods

Deuterium oxide (D_2O > 99.96%) was purchased from Sigma-Aldrich. All isotropic samples were prepared with 100% D_2O as solvent.

Pentaethylene glycol monododecyl ether ($C_{12}E_5 \geq 98.0\%$, GC, purchased from Biochemika) and hexanol (> 99%, anhydrous, purchased from Sigma-Aldrich) were used as the alignment medium (PEG) [59] for the preparation of aligned samples.

Pf1 filamentous phage (50 mg/ml, 10 mM kp_i, pH = 7.6, 2 nM $MgCl_2$, 0.02% NaN_3) was purchased from Profos, Germany. The solvent was further diluted with D_2O to 1.7 mg/ml as the aligned medium to dissolve the cholic acid sodium salt.

Cellobiose (Fig.II.2.1a) used for development of *J*-modulated methods for determination of scalar and proton-proton dipolar coupling constants was purchased from Sigma-Aldrich. 20 mg were dissolved in 550 µl of D_2O. The aligned sample was prepared using the same amount of the compound in 550 µl D_2O containing 25.6 µl $C_{12}E_5$ and 8.5 µl hexanol yielding 4% (w/w) of the medium. Splitting of the ^2H signal was 23.6 Hz at 25 °C.

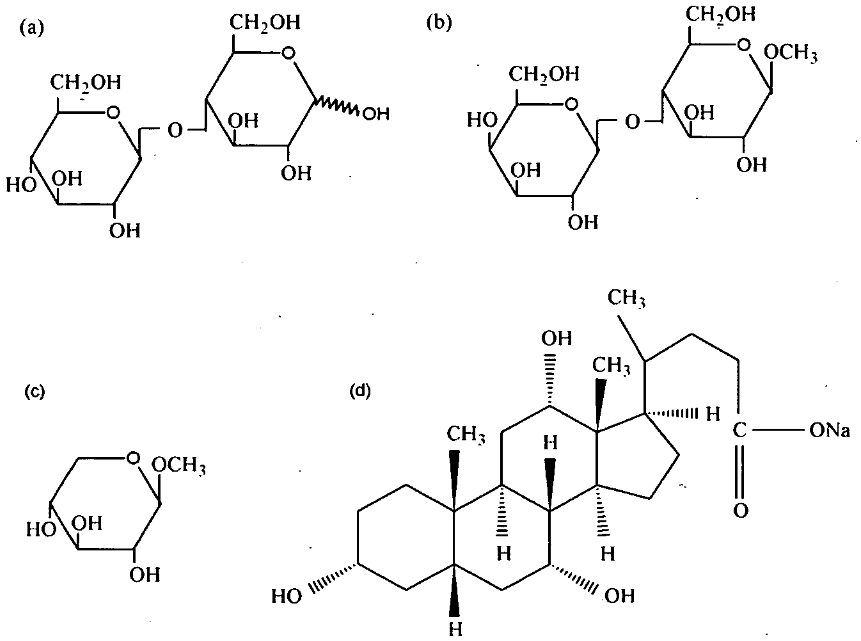


Fig.II.2.1 Structures of (a) cellobiose (β-D-glucopyranose disaccharide), (b) Me-β-D-lactoside, (c) Me-β-D- xylopyranoside and (d) cholic acid sodium salt

Me-β-D-lactoside and Me-β-D-xylopyranoside (Fig.II.2.1b, c), used for the development of INADEQUATE methods, were purchased from Sigma. 25 mg of disaccharide and 71 mg of monosaccharide were dissolved in 350 μl of D_2O and transferred into 5 mm Shigemi NMR tubes. The Me-β-D-xylopyranoside was aligned using a mixture of $C_{12}E_5$/hexanol medium for the measurement of ^{13}C-^{13}C RDCs. 128 μl of $C_{12}E_5$, 38.8 μl of hexanol and 600 μl of D_2O were used to dissolve 140 mg Me-β-D-xylopyranoside. The mass fraction of the medium was 22% with the molar ratio $r = 0.99$ of $C_{12}E_5$/hexanol. The splitting of the deuterium signal at 25 °C was 138.8 Hz.

Cholic acid sodium salt ($C_{24}H_{39}NaO_5 \geq 96\%$, Fig.II.2.1d), used to illustrate the application of *J*-modulated methods for the measurement of ^1H-^1H RDCs, was purchased from Biochemika. Approximately 1 mg was dissolved in 550 μl D_2O for the isotropic sample. The aligned sample was prepared by dissolving 2 mg of cholic acid sodium salt in 625 μl of 1.68 mg/ml phage medium (Pf1 filamentons phage). The splitting of the deuterium signal was 3.3 Hz at 25 °C.

$^{13}C_1$-glucopyranose and ^{13}C-CH_3COONa used for the development of INADEQUATE methods were purchased from Sigma-Aldrich. 20 mg ^{13}C-CH_3COONa was dissolved in 550 μl D_2O. 150 mg $^{13}C_1$-glucopyranose was dissolved in 550 μl D_2O. A mixture of 50 mg ^{13}C-CH_3COONa and 50 mg $^{13}C_1$-glucopyranose was also dissolved in 550 μl D_2O.

II. 3 Spin-labelled disaccharide

Spin-labelled heparin-derived fully sulfated disaccharide (compound V Fig.II.3.1) was synthesized by Astrid Jahnke [65]. 4-amino-2,2,6,6-tetramethyl-piperidine-1-oxyl, 4-amino-TEMPO (a TEMPO derivative), was used as the spin label, linked with the carboxyl group of the non-reducing ring of the disaccharide. An NMR sample was prepared by dissolving 200 μg of the molecule in 500 μl D_2O.

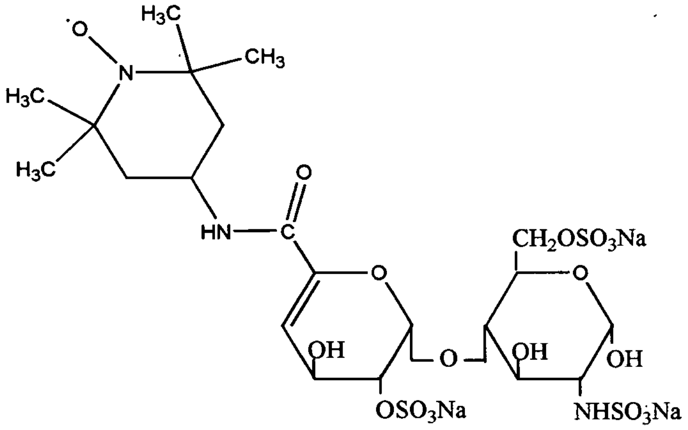


Fig.II.3.1 Spin-labelled heparin-derived fully sulfated disaccharide (compound V)

II. 4 Ion Mobility Mass Spectrometry [49]

Gas phase collision cross-sections of heparin oligosaccharides **I-IV** were determined using the Ion Mobility Mass Spectrometer in the Chemistry Department at the University of California at Santa Barbara[ii].

Samples were sprayed at a concentration of 200 μM from a solution of 50% methanol 50% water and 2% acetic acid using the nano-electrospray method (negative electrospray ionization). The prepared samples contained 50 μg of oligosaccharides and we estimate that around 20 μg were actually used to collect the data. Ions were first selected by mass-to-charge ratio before injected into the drift cell.

The drift cell was temperature regulated (298 K), 4 cm long and pressurized with helium to 5 Torr. The ions drifted under the influence of a 5-30 V cm^{-1} electrostatic field while retarded by collisions with the buffer gas. On the exit of the cell, the ions were analyzed using a

---

[ii] All experiments were carried out by Dr. Perdita E. Barran

quadrupole mass spectrometer (Extrel) operating in scanning mode. Mass spectra were obtained for each of the oligosaccharides, which were presented in singly or doubly charged forms as partially sodiated species. The experimental and predicted masses for the major species observed for each oligosaccharide are listed in Table II.4.1.

Table II.4.1 Major peaks observed in mass spectra of oligosaccharides **I–IV**

Compound	Observed (m/z)	Calculated (m/z)	Assigned species
I	642.7	642.4	$[C_{12}H_{15}NO_{19}S_3Na_3]^-$
	631.3	630.9	$[C_{24}H_{31}N_2O_{38}S_6Na_5]^{2-}$
	620.6	619.9	$[C_{12}H_{16}NO_{19}S_3Na_2]^-$
II	642.8	642.4	$[C_{24}H_{30}N_2O_{38}S_6Na_6]^{2-}$
	631.6	630.9	$[C_{24}H_{31}N_2O_{38}S_6Na_5]^{2-}$
	621.1	620.4	$[C_{24}H_{32}N_2O_{38}S_6Na_4]^{2-}$
III	592.3	591.4	$[C_{24}H_{31}N_2O_{35}S_5Na_5]^{2-}$
	581.2	580.4	$[C_{24}H_{32}N_2O_{35}S_5Na_4]^{2-}$
	570.1	569.4	$[C_{24}H_{33}N_2O_{35}S_5Na_3]^{2-}$
IV	592.4	591.4	$[C_{24}H_{31}N_2O_{35}S_5Na_5]^{2-}$
	581.1	580.4	$[C_{24}H_{32}N_2O_{35}S_5Na_4]^{2-}$
	570.1	569.4	$[C_{24}H_{33}N_2O_{35}S_5Na_3]^{2-}$

When an ion was pulsed into the cell, the quadrupole was then 'parked' on the ion of interest and its arrival time at the detector, relative to the start time of the pulse, was recorded. Arrival time distributions (ATDs) were collected at several drift voltages (90, 60, 45, 30 and 20 V).

II. 5 Nuclear Magnetic Resonance spectroscopy

II. 5.1 Parameters of NMR experiments on compound II

Unless stated otherwise, all spectra were recorded at 298 K on an 800 MHz Bruker Avance spectrometer equipped with a 5 mm triple-resonance probe and xyz-gradients. All data were processed using XWINNMR program (Linux, version 3.6).

The 1D ^{1}H NMR spectrum (Fig.III.1.2.1) of an isotropic sample was recorded using water presaturation during the relaxation delay. The 1D ^{1}H NMR spectrum (Fig.III.4.4.1.1a) of the aligned sample was acquired using a CPMG pulse sequence with an interpulse delay of 1.6 ms and a total duration of the spin-echoes of 32 ms.

1D TOCSY spectra (Fig.III.4.4.1.1b, c) were acquired using selective excitation of anomeric protons by a 30 ms Gaussian pulse followed by a DIPSI-2 spin-lock. The duration of the spin-lock was 100 ms for A and C rings, 80 ms for D ring and 160 ms for B ring.

The 2D NOESY spectrum (Fig.III.4.2.2) of the isotropic sample was recorded using water presaturation during the relaxation delay and also during the mixing time of 300 ms. 50 ms Crp40.1000 pulse with a 2 Gauss/cm PFG was used to suppress the zero-quantum (ZQ) coherence. Acquisition times in both t_1 and t_2 were 256 ms. Sixteen scans were acquired per each of 1536 increments and the total experimental time was 18 hours.

The 2D ROESY spectrum (Fig.III.1.2.3) of the isotropic sample was recorded with water presaturation during the relaxation delay. The duration of the spin-lock was 300 ms using a CW pulse with γ_{B1} = 25000 Hz. Acquisition times in t_1 and t_2 were 80 and 256 ms, respectively. Number of scans was 48 and 512 increments were acquired yielding the total experimental time of 14 hours.

The proton-detected ^{1}H-^{13}C gradient selected phase-sensitive HMBC [68] spectrum (Fig.III.4.3.1) was acquired on a 600 MHz Bruker Avance spectrometer equipped with a 5 mm cryoprobe with inverse geometry and z-gradients. Acquisition times in t_1 and t_2 were 21.2 and 852 ms, respectively. The number of scans was 256 with 256 increments and the total experimental time was 48 hours. Long-range C-H scalar coupling constants were obtained by a fitting procedure described by Keeler et al. [68] in the frequency domain. This procedure requires pure phase ^{1}H multiplets that are used to reconstruct the HMBC multiplets. These were extracted from 1D ^{1}H or 1D TOCSY spectra. The C-H coupling constants across the three glycosidic linkages were of particular interest and were obtained by a home written program (Dr. Tran N. Pham).

Non-refocused 2D ^{1}H-^{13}C HSQC spectra of both isotropic and aligned samples were recorded on a 600 MHz Bruker Avance spectrometer equipped with a 5 mm cryoprobe with inverse geometry and z-gradients. The spectra were acquired without decoupling in the directly detected dimension. Acquisition times of both isotropic and aligned samples in t_1 and t_2 were 63.6 and 341 ms, respectively. Numbers of increments of both samples were 768. Numbers of scans in experiments using of isotropic and aligned samples were 160 and 116, respectively. The total experimental times were 62.5 and 45.5 hours, respectively. One-bond heteronuclear C-H scalar and dipolar coupling constants were extracted based on the frequencies of the antiphase doublets from the spectra processed without using any window function.

Typically SPFGSE-COSY and CSSF-COSY experiments were used for the measurement of ^1H-^1H scalar and dipolar coupling constants using isotropic and aligned samples. For severely overlapped protons especially in the aligned sample, HOHAHA-COSY spectra were acquired.

The spectra of the isotropic sample were normally acquired using 80 scans, while those of the aligned sample were recorded using 200 scans. The acquisition and relaxation times were 0.96 s and 1.5 s, respectively, for both isotropic and aligned samples. 20 ms Gaussian pulses were used for the selective pulses and also during the COSY part in the SPFGSE-COSY experiments using isotropic sample, while 30 ms Gaussian pulses were used in the aligned sample. For ^1H-^1H coupling constants smaller than 1 Hz, 0.8 s constant-time SPFGSE-COSY spectra were acquired with 512 scans. More selective Gaussian pulses, e.g. 45 ms, were used for crowded protons.

CSSF-COSY spectra were typically acquired using eight 5 ms or 10 ms CSSF increments for small or large ^1H-^1H coupling constants in both isotropic and aligned sample, respectively. 30 ms Gaussian pulses were used for the selective pulses and also during the COSY part. The acquisition and relaxation times were 0.64 s and 1.5 s, respectively. The spectra were acquired using 4 scans. More selective Gaussian pulses, e.g. 45 ms, were used for crowded protons.

1D double-selective HOHAHA spectra (Fig.III.4.4.1.1e) were acquired using the aligned sample. A 30 ms Gaussian pulse applied at δ_{H2}, 80 ms double-selective rectangular pulse centered at δ_{H2} and δ_{H3} with RF field strength of 50 Hz [$\gamma B_1/2\pi = (J_{3\alpha,4\alpha}/4)(4n^2 - 1)^{1/2}$, where J_{23} = 10 Hz and n = 10], acquisition and relaxation times of 1.0 and 1.5 s, respectively, and 400 scans were used. A 50 ms 180° adiabatic smoothed (10%) CHIRP pulse with a frequency sweep of 40 kHz was used to purge dispersive components [69]. A 30 ms Gaussian pulse was used to select magnetization of proton H3. The selective pulses applied to H3 and H4 protons were 30 ms Gaussian pulses.

II. 5.2 Alignment media for the measurement of residual dipolar coupling constants

The $C_{12}E_5$/hexanol medium was initially tested on hyaluronan tetrasaccharide [70] β-D-glucuronic acid-β-(1→3)-N-acetylglucosamine-β-1,4 (AB_2) (Fig.II.5.2.1). A series of different strength media were prepared and their effects tested.

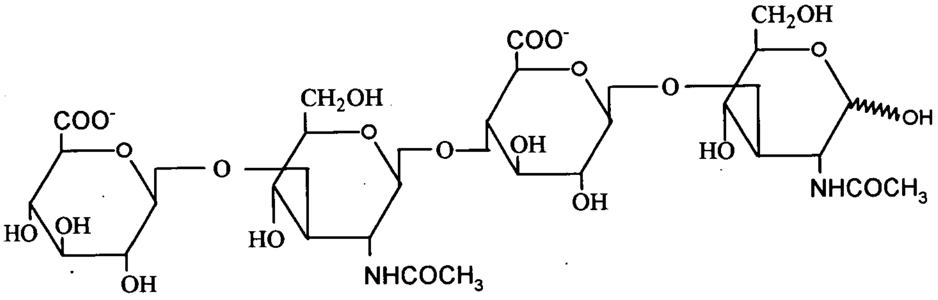


Fig.II.5.2.1 Structure of AB_2 molecule

The aligned media with r = 0.9 were used, which are suitable for the temperature range around 25 °C [59]. A series of media with *wt%*'values of 6%, 5.2%, 4.5%, 4% and 3.3% were prepared to dissolve the AB_2 molecule using the same salt concentration and pH as was intended to be used for the heparin-derived fully sulfated tetrasaccharide **II**. Samples were kept in the magnet until the deuterium splitting was stable and symmetrical. All spectra were acquired on a 600 MHz Bruker spectrometer equipped with a 5 mm triple-resonance probe and xyz-gradients. The splittings of the deuterium signals at 25 °C are shown in Fig.II.5.2.2.

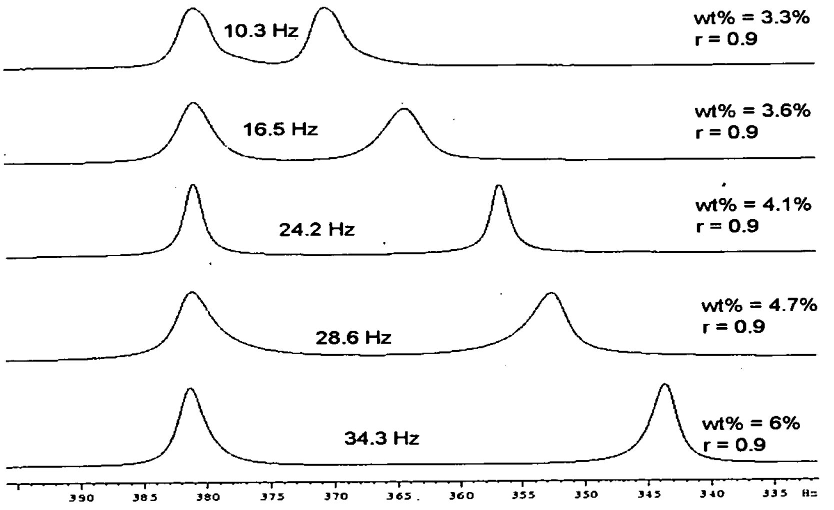


Fig.II.5.2.2 Splittings of deuterium signals on different strength of PEG media

HSQC spectra were acquired and the corresponding one-bond (J + D) splittings of C1H1 cross peaks were extracted by peak picking. The $^1D_{CH}$ coupling constants were calculated using $^1J_{CH}$ values from isotropic sample. Representative $^1D_{CH}$ coupling constants are summarized in Table II.5.2.1.

Table II.5.2.1 Dipolar coupling constants of D_{C1H1} in different alignment media

r = 0.9 / wt%	6%	4.7%	4.1%	3. 6%	3.3%
D_{C1H1} (Hz)	19	15	13	8	5

The medium with r = 0.9 and $wt\%$ = 3.7% was chosen for the aligned sample of the tetrasaccharide **II**. The composition of the mixture was 18.6 μl $C_{12}E_5$, 6.2 μl hexanol and 430 μl D_2O. The splitting of the deuterium signal was 22.5 Hz at 25 °C.

II. 5.3 Parameters of new INADEQUATE experiments

All ^{13}C-^{13}C ^1H-detected INADEQUATE spectra were acquired on a 600 MHz Bruker Avance spectrometer equipped with a 5 mm cryoprobe with inverse geometry and z-gradients. The ^{13}C-detected INADEQUATE experiments on Me-β-D-xylopyranose were acquired on a 600 MHz cryoprobe with a cold ^{13}C preamplifier (University of Glasgow). The experiment on Me-β-D-lactoside was acquired on an 800 MHz cryoprobe with a cold ^{13}C preamplifier (University of Edinburgh). Unless stated otherwise, the ^1H-detected INADEQUATE spectra were processed with no window function in F_1 and a Gaussian function in F_2.

DJM-INEPT-INADEQUATE:

The 2D DJM-INEPT-INADEQUATE spectrum (Fig.III.3.1.2) of Me-β-D-xylopyranoside was recorded using 32 scans in each of 800 complex t_1 increments. Acquisition times in T_1 and T_2 were 94.7 and 106 ms, respectively. The relaxation time was 1s and the total acquisition time was 23 hours. Adiabatic 180° pulses described in [71] were used as 500 μs (inversion) and 2 ms (refocusing) pulses. This experiment did not contain any fixed ^{13}C-^{13}C evolution delays. A scaling factor of 4 was used for J-modulation. An acquisition time t_1 of 95 ms was used to measure coupling constants as small as 2 Hz. The maximum J-modulation time was 380 ms. A compromised delay τ_3 = 2 ms was used in the experiment.

DJM-REVINEPT-INADEQUATE:

The 2D spectrum (Fig.III.3.2.5) of Me-β-D-xylopyranoside was acquired using 8 scans in each of 2688 complex t_1 increments. A scaling factor of 4 was used for *J*-modulation. Acquisition times in t_1 and t_2 were 159.1 and 106.8 ms. A compromised delay $\tau_3 = 2$ ms was used in the experiment. The relaxation time was 1.5 s and the total acquisition time was 14 hours.

The 2D spectrum (Fig.III.3.2.6) of Me-β-D-lactoside was acquired using 100 scans in each of 1280 complex t_1 increments. A scaling factor of 4 was used for *J*-modulation. Acquisition times in t_1 and t_2 were 75.8 and 106.8 ms. The delay of $\tau_3 = 3.3$ ms was optimized for the CH couplings in the experiment. The relaxation time was 1.3 s and the total acquisition time was 63 hours.

^{13}C spin-echo experiments (Fig.III.3.2.2) with or without ^{1}H decoupling were acquired using the following parameters. The acquisition times were 0.45 s for mono-, di- and tetrasaccharides. The relaxation times were 1 s for mono- and tetrasaccharides and 2 s for the disaccharide. Variable delays of 0–800 ms, 0–720 ms and 0–250 ms were used for mono-, di- and tetrasaccharides, respectively. The number of scans of 4, 512 and 2048 were used for mono-, di- and tetrasaccharides for the spectra acquisition, respectively.

JM-DEPT-INADEQUATE:

The in- and antiphase multiplets (Fig.III.3.3.3) of the Me-β-D-xylopyranoside were acquired in an interleaved manner using 8 scans in each of 3072 complex t_1 increments. Acquisition times in t_1 and t_2 were 181.8 and 106.8 ms. The $^{1}J_{CH}$, $^{1}J_{CC}$ and $^{n}J_{CC}$ were optimized for 150, 42 and 3 Hz couplings, respectively. A scaling factor of 4 was used for *J*-modulation. The relaxation time was 1.0 s and the total acquisition time was 37.5 hours.

The interleaved spectrum (Fig.III.3.3.4) including in- and antiphase slices of the Me-β-D-lactoside were acquired using 24 scans in each of 1750 complex t_1 increments. Acquisition times in t_1 and t_2 were 103.6 and 102.6 ms. The J_{CH}, $^{1}J_{CC}$ and $^{n}J_{CC}$ are optimized for 150, 42 and 3 Hz couplings, respectively. A scaling factor of 4 was used for *J*-modulation. The relaxation time was 1.0 s and the total acquisition time was 54 hours.

INEPT-INADEQUATE and REVINEPT-INADEQUATE:

Both spectra (Fig.III.3.4.2) of Me-β-D-xylopyranoside were acquired using identical parameters. Each spectrum was acquired using 4 scans in each of 512 complex t_1 increments.

Acquisition times in t_1 and t_2 were 30.3 and 106.8 ms. The ${}^{1}J_{CH}$ and ${}^{n}J_{CC}$ were optimized for 150 and 3 Hz couplings, respectively. The relaxation time was 1.3 s and the total acquisition time was 1 hour.

REVINEPT-INADEQUATE spectra (Fig.III.3.4.4) optimized for long-range or one-bond J_{CC} coupling constants were acquired on the sample of Me-β-D-lactoside. Acquisition times in t_1 and t_2 were 30.3 and 106.8 ms, and the relaxation time was 1 s in both cases. The ${}^{1}J_{CH}$ was optimized for 150 Hz couplings. The ${}^{n}J_{CC}$ and ${}^{1}J_{CC}$ were optimized for 3 and 50 Hz, respectively. The long-range spectrum was acquired using 40 scans and 9.5 hours of spectrometer time. The one-bond correlation spectrum was acquired using 8 scans and 1.5 hours of experimental time.

*J*M-REVINEPT-INADEQUATE:

The in- and antiphase spectra (Fig.III.3.5.2) of Me-β-D-xylopyranoside were acquired in an interleaved manner. Experiment with a refocusing delay yielded the in-phase spectrum; the experiment without refocusing period resulted in the antiphase spectrum. Each spectrum was acquired using 4 scans in each of 3072 complex t_1 increments. Acquisition times in t_1 and t_2 were 181.8 and 106.8 ms. The ${}^{1}J_{CH}$ and ${}^{n}J_{CC}$ were optimized for 150 and 3 Hz couplings, respectively. A scaling factor of 4 was used for *J*-modulation. The relaxation time was 1.3 s and the total acquisition time was 19.5 hours.

The interleaved spectrum (Fig.III.3.5.3) of the Me-β-D-lactoside was acquired using 24 scans in each of 2176 complex t_1 increments. Acquisition times in t_1 and t_2 were 128.8 and 106.8 ms. The ${}^{1}J_{CH}$ and ${}^{n}J_{CC}$ were optimized for 150 and 3 Hz couplings, respectively. A scaling factor of 4 was used for *J*-modulation. The relaxation time was 1.3 s and the total acquisition time was 74 hours.

^{13}C-detected INADEQUATE:

Two interleaved refocused and non-refocused 2D spectra (Fig.III.3.6.3) of Me-β-D-xylopyranoside (isotropic and aligned samples) were acquired using 64 scans in each of 364 complex t_1 increments. Acquisition times in t_1 and t_2 were 46.4 and 452 ms. The ${}^{n}J_{CC}$ was optimized for 3.07 Hz coupling. The relaxation time was 1.0 s and the total acquisition time was 26 hours. The spectra were processed using Gaussian and Exponential window functions in the F_1 and F_2 dimensions, respectively. Line-broadening of 1.0 Hz was used in the F_2 dimension. Linear prediction was used in both dimensions.

Two interleaved refocused and non-refocused 2D spectrum (Fig.III.3.6.7) of Me-β-D-lactoside were acquired using 256 scans in each of 112 complex t_1 increments. Acquisition times in t_1 and t_2 were 18.6 and 511 ms. The $^{n}J_{CC}$ was optimized for 3.0 Hz couplings. The relaxation time was 0.5 s and the total acquisition time was 48 hours. The spectra were processed with Gaussian and Exponential window functions in the F_1 and F_2 dimension, respectively. Line-broadening of 1.0 Hz was used in the F_2 dimension. Linear prediction in the F_1 dimension was used.

II. 5.4 Parameters of NMR experiments for the measurement of J_{HH} and D_{HH} coupling constants

All spectra were recorded on an 800 MHz Bruker Avance spectrometer equipped with a 5 mm triple resonance probe and xyz-gradients.

Cellobiose:

A 1D spectrum of the aligned cellobiose (Fig.III.2.2.4a) was recorded using a CPMG pulse sequence with an interpulse delay of 1.6 ms and the total duration of the spin-echoes of 32 ms. 1D TOCSY spectrum (Fig.III.2.2.4b) was acquired using selective excitation [72] of the anomeric proton $H1_\alpha$ via a 30ms Gaussian pulse followed by a 120 ms DIPSI-2 spin-lock.

A 1D double-selective HOHAHA spectrum (Fig.III.2.2.4c) was acquired using the pulse sequence of Fig.III.2.2.2a. Acquisition started at point **a**. The following parameters were used: 30 ms Gaussian pulse applied at $\delta_{H3\alpha}$, 80 ms double selective rectangular pulse centered at $\delta_{H3\alpha}$ and $\delta_{H4\alpha}$ with RF field strength of 44 Hz [$\gamma B_1/2\pi = (J_{3\alpha,4\alpha}/4)(4n^2 - 1)^{1/2}$, where $J_{3\alpha,4\alpha} = 8.9$ Hz and n = 10], acquisition and relaxation times of 1 and 2s, respectively, and 4 scans. The spectrum of Fig.III.2.2.4d was acquired using identical parameters but acquisition was started at point **b**. 50 ms 180° adiabatic smoothed (10%) CHIRP pulse with a frequency sweep of 40 kHz was used to purge dispersive components. Practically all $H3_\alpha$ magnetization was transferred to $H4_\alpha$. The spectrum of Fig.III.2.2.4e was obtained by starting the acquisition at point **c**. A 30 ms Gaussian pulse was used to select magnetization of proton $H4_\alpha$. The J-modulated 1D directed HOHAHA-COSY spectrum (Fig.III.2.2.4f) was acquired using the pulse sequence of Fig.III.2.2.2b. The parameters for the doubly selective HOHAHA step were identical to those used to acquire the spectrum of Fig.III.2.2.4e. T and τ_r were set to 200 ms and 250 ms, respectively, and refocusing was used. The selective pulses applied to $H4_\alpha$ and $H2_\alpha$ protons were 50 ms Gaussian pulses; 128 scans were accumulated.

The 1D ge-VT-CSSF spectrum [73] (Fig.III.2.3.2c) was acquired using the pulse sequence of Fig. III.3.2.3.1a and twelve 4 ms CSSF increments with 2 scans each. A 30 ms Gaussian pulse was applied to $H3_{\alpha}$. The *J*-modulated 1D directed CSSF-COSY spectrum (Fig.III.2.3.2d) was acquired using the pulse sequence of Fig.III.2.3.1a using 32 increments with 16 scans each. Other parameters of the CSSF were identical to those used to acquire the spectrum of Fig.III.2.3.2c. 45 ms Gaussian pulses were applied during the COSY part of the sequence; T and τ_r were set to 100 ms and a refocusing period was used.

Sodium cholate:

A ^{1}H spectrum of the aligned sample (Fig.III.2.5.1b) was recorded using a CPMG pulse sequence with an interpulse delay of 1.6 ms and total duration of the spin-echoes of 32 ms.

1D TOCSY spectra (Fig.III.2.5.1c-f) were acquired using selective excitation of the proton H3, H7, H12 and C18-Me via a 30ms Gaussian pulse followed by a 140 ms DIPSI-2 spin-lock. .

The non-refocused 2D ^{1}H-^{13}C HSQC spectra of isotropic and aligned samples were acquired without decoupling in the directly detected dimension. Identical acquisition parameters were used for the isotropic and aligned samples. Acquisition times in t_1 and t_2 were 15.9 and 604.8 ms, respectively. Numbers of increments were 512. Numbers of scans in experiments were 48 and the total experimental times were 6.5 hours. One-bond heteronuclear C-H scalar and dipolar coupling constants were extracted based on the frequencies of the antiphase doublets from the spectra processed without using any window function.

A series of typical 1D SPFGSE-COSY spectra were acquired using the variable delays from 30 to 360 ms for both isotropic and aligned samples. 30 ms Gaussian pulses were used for the selective pulse and also during the COSY part. A more selective Gaussian pulse, e.g. 70 ms, was used for seriously overlapping protons such as $H4_{ax}$. The acquisition and relaxation times were 0.8 s and 1.5 s, respectively. The spectra of the isotropic sample were acquired using 32 scans, while those of the aligned sample were recorded using 64 scans.

A series of typical 1D non-refocused CSSF-COSY spectra were acquired using sixteen 5 ms CSSF increments for both isotropic and aligned samples. 30 ms Gaussian pulses were used for the selective pulse and also during the COSY part. The variable delays were set to be 30 to 360 ms. The acquisition and relaxation times were 0.8 s and 2 s, respectively. The spectra of the isotropic sample were acquired using 4 scans, while those of the aligned sample were recorded using 8 scans.

II. 6 Parameters used for docking and modeling

II. 6.1 Docking of heparin-derived fully sulfated tetrasaccharide to fH modules

The C, N, O, S and H atoms of the tetrasaccharide were used for the calculation of the affinity grid maps (Appendix I) as well as the electrostatic map before docking. The dielectric constant was set to be a constant of 1.0. The interactions of hydrogen-bonds between N-H, O-H and S-H were also calculated. The protein was kept static during the whole process. The grid maps of the atomic affinity and electrostatic potentials were calculated for each element in the ligand. The potential energies of the interaction between the atoms of the ligand and the macromolecule were stored in points within the grid maps. These maps were used during the docking process (Appendix I) to evaluate the energy during each run at 298.15 K. The ligand changed the orientation along the rotatable bonds when moving around the surface of the protein. The energy was calculated and compared with the previous step. The configuration with the lowest energy was accepted in each cycle. The Darwinian genetic algorithm (GA) was used for the docking calculation in 250 runs, each including 50 cycles and producing 250 possible conformations around the binding sites.

II. 6.2 Molecular modeling of the heparin-derived oligosaccharides

II. 6.2.1 Parameterization of heparin-oligosaccharides for AMBER calculations

Heparin-derived fully sulfated di- and tetrasaccharides were built using the atom types contained in the newly parameterized force field (Appendix II). Partial atom charges were calculated with the RESP (Restrained ElectroStatic Potential fit) procedure using Gaussian98 with the 6-31G* basis set. The single point charge calculation was carried out with the convergence criteria of 1.00D-02 and 1.00D-04 for the maximal and RMSDs change of the density matrix elements, respectively. The net charges for the fully sulfated di- and tetrasaccharide were set to -4 and -8, respectively (PDB files with partial atom charges are shown in Appendix III.) Three and six sodium ions were added randomly as counter ions to the di- (**I**) and tetrasaccharide (**II**), respectively, to generate sodiated oligosaccharide models with total charges of -1 and -2, respectively. The topology and coordinate files were used for subsequent simulations. The same procedure was applied to build up the other two tetrasaccharides **III** and **IV**. Five sodium ions were presented in the final systems resulting in a total charge of -2.

II. 6.2.2 Protocols for gas/solution modeling of GAGs in AMBER

The complexity of oligosaccharide conformations is mainly associated with the flexibility of the glycosidic linkages. In order to sample the conformational space sufficiently, a large number of conformers needs to be generated. Therefore, a simulated annealing procedure was used to generate candidate structures for each oligosaccharide ions in the gas phase.

Simulated annealing (Appendix IV) is a cyclic calculation protocol containing initial energy minimization, high temperature molecular dynamics followed by a rapid cooling step to a very low temperature to freeze out individual conformers, and a final energy minimization. The calculation was run in vacuum without any solvent or periodic box. In the initial energy minimization period, the starting conformation was optimised with a 3 ps steepest descent step followed by a gradient method of a duration 0.5 ns. The optimized structure was subjected to a 30 ps of molecular dynamics at 800 K with constraint on the bonds involving H-atoms (SHAKE). The system was then cooled to 0 K during a period of 10 ps using ten cooling steps of 80 K in order to freeze out one conformation. This geometry was then optimized by minimizing the energy in the same way as the initial energy minimization was carried out. The first candidate conformation was obtained in this way and used as the input for the next round of the simulated annealing. One hundred candidate structures were generated for di- and tetrasaccharides. Conformations within 5 kcal mol^{-1} of the lowest potential energy were considered to be the correct models. The 10% lowest potential energy conformations out of the 100 structures were used as the final candidate structures for the theoretical calculation of collision cross-sections. These included 10 and 5 conformations for sodiated fully sulfated di- **I** and tetrasaccharide **II** models, respectively, 5 and 4 for the tetrasaccharides **III** and **IV**, respectively. Detail analysis and discussion for these gas phase conformations are in Chapter III.1.

Fully sulfated tetrasaccharide, **II,** was also studied in solution. The structure with the lowest potential energy from gas phase simulation was dissolved in a water box for explicit solution periodical calculation (Appendix IV). Before adding the water molecules, eight sodium ions were added randomly to neutralize the whole unit. A cubic 40.675 Å water box with a spacing of 8 Å containing 1938 triangulated 3-point water molecules was generated. The whole water box was first energy minimized with a tight restraint on the tetrasaccharide atoms to relax the water molecules. Then the whole system was energy minimized without any restraint. The temperature of the system was then increased from 0 K to 300 K in 0.002 ns and stabilized at 300 K for 0.02 ns at a constant volume state with constraint on the bonds involving H-atoms while the tetrasaccharide molecule was restrained weakly. Another 0.1 ns equilibration molecular dynamic at 300 K and a constant pressure (1 atm) was simulated without any restraint before the molecular dynamic calculations.

Restrained and unrestrained molecular dynamics (Appendix IV) were performed at constant pressure (1 atm) and temperature of 300 K for 4 and 2 ns, respectively. RDCs of both glucosamine rings (A and C rings in compound **II**) were used as experimental restraints. Comparisons of these two simulations are discussed in detail in Chapter III.4.

II. 6.2.3 Calculation of Φ/Ψ energy maps of disaccharide fragments

The relative orientation of individual monosaccharide rings is described by the dihedral angles Φ ($H_{1'}$-$C_{1'}$-O-C_4) and Ψ ($C_{1'}$-O-C_4-H_4). These angles were restrained in 10-degree steps between -180 and 180 degrees and the structures were energy minimized during a 3 ps steepest descent step followed by a gradient method of 0.5 ns, using the parameterized force field. The calculations (Appendix IV) were performed in vacuum and in the absence of Na^+ ions. The non-bonded cut-off value was set to 999 Å and a distance-dependant dielectric constant was used. For ΔUA2S and IdoA2S residues (rings D and B) two ($^1H_2$ and $^2H_1$) and three ($^1C_4$, $^4C_1$ and $^2S_0$) starting conformations were used, respectively. These conformations were not fixed during the calculations, although they did not change. Maps that produced only the lowest energy minima are shown in Fig.III.1.6.1. The resulting (Φ, Ψ) maps were drawn using ten isoenergy contours in 1 kcal mol^{-1} steps above the global minimum.

II. 6.3 Molecular modeling of the spin-labeled disaccharide

Partial atom charges were calculated with the RESP procedure using Gaussian98 and the 6-31G* basis set. The single point charge calculation was carried out with the convergence criteria of 1.00D-02 and 1.00D-04 for the maximal and RMSD change of the density matrix elements, respectively. The net charge of the whole molecule was set to be -3. Both $^1H_2$ and $^2H_1$ conformations of the uronic acid ring (D ring) were considered as starting structures.

The TEMPO-disaccharide molecule (compound **V**) was dissolved in a cubic 32.683 Å water box with a spacing of 8 Å containing 981 triangulated 3-point water molecules. The water box was first energy minimized using tight restrains on the solute atoms, followed by energy minimization of the complete system without restraining the solute. The temperature of the system was then increased from 0 K to 300 K in 0.002 ns and stabilized at 300 K for 0.02 ns at a constant volume using SHAKE for bonds involving H-atoms. The position of the solute was weakly restrained. This step was followed by a 0.1 ns equilibration molecular dynamics at 300 K and constant pressure, 1 atm, without any restraints on the position of the solute. Finally, a constant temperature free molecular dynamics with SHAKE at 300 K was performed for 1 ns. Both $^1H_2$ and $^2H_1$ starting conformations converged to the $^2H_1$ form during the simulation.

Chapter III

Results and Discussion

The heparin-derived oligosaccharides were studied in the gas phase and in solution experimentally. As the interpretation of the results requires theoretical structures, the description of the protocols used in molecular modeling is given first.

III. 1 Gas phase conformations of heparin-derived oligosaccharides [77]

III. 1.1 Parameterization of the AMBER force field for GAGs

We have further parameterized the Parm99 force field in AMBER7 [74] by introducing the GLYCAM_2000a [75] parameter set for carbohydrates. This extension included creation of new atom types such as α-, β-anomeric carbons, definition of parameters for glycosidic linkages and carbohydrate specific bonds, angles, torsion angles and improper torsion angles. However, GLYCAM_2000a does not contain parameters for sulfate groups, which are common in GAGs. Therefore, parameters for sulfates and sulfamates were applied according to Huige and Altona [76]. A new atom type, SO, for the sulfate group was created. Some parameters for sulfate groups that were not available from the work of Huige and Altona were approximated by using those for phosphates from Parm99. The compatibility of parameters arising from the presence of the double bond contained in the non-reducing terminal monosaccharide ring was also resolved (Appendix II). All of these modifications were written into a new file, which was used as a force field modified file loaded together with the original parm99 force field into XLEAP sessions.

β-D-glucopyranosyl-(1→2)-β-D-manopyranose was used as a pilot study to test the performance of the modified force field parameters for carbohydrates. This compound was chosen because its glycosidic linkage was known to exist in two conformations. MM3 force field structures of two energetically favored conformations were provided by S. Perez. The dihedral angles (Φ, Ψ) across the glycosidic linkages were (60, 120) (conformer I) and (280, 140) (conformer II), respectively. The starting structures were used as input files to calculate the atom charges with the AM1-BCC basis set using ANTECHAMBER. The output file was loaded into an XLEAP session to generate topology and coordinate files, which contain proper atom types and geometries, and used for further calculation.

Free molecular dynamics was simulated in vacuum at 300 K for 6 ns with the two models. Bonds involving H-atoms were constrained using SHAKE during dynamic .calculations eliminating stretching motions. From the trajectory files, Φ and Ψ torsion angles across the glycosidic linkage were extracted and plotted as a function of time (Fig.III.1.1.1). The Φ angle of conformer I changed from an average value of 60 degrees to 280 degrees, resulting in the changing from one minimum (conformer I) to the other (conformer II). The torsion angle distribution is shown in Fig.III.1.1.2.

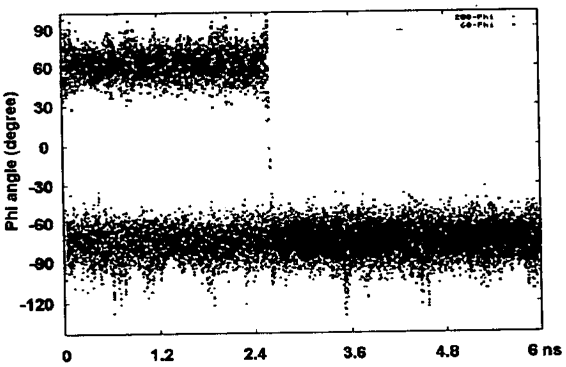


Fig.III.1.1.1 Phi angle (H-C-O-C') across the glycosidic linkage of β-D-glucopyranosyl-(1→2)-β-D-manopyranose during the simulation. Red dots show the phi angle of conformer II. Green dots show the phi angle of conformer I. The phi angle jumped from 60 degrees to -80 degrees during the molecular dynamics of conformer I, which indicates the conformation change from one minimum to the other.

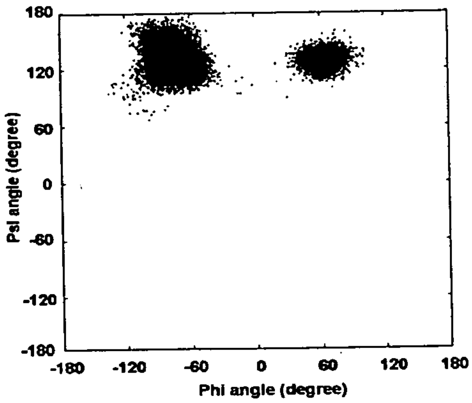


Fig.III.1.1.2 The distribution of (Φ, Ψ) angles during the simulation.

These calculations suggest that the parameterized force field can be applied to oligosaccharides and that the flexibility across the glycosidic linkages can be represented properly. The parameterization of heparin-derived oligosaccharides and protocols used for gas and solution modeling of GAG in AMBER are described in Chapter II.6.2.

III. 1.2 The studied species

Heparin-derived disaccharide (**I**) and three tetrasaccharides (**II-IV**) were studied in the gas phase. Disaccharide **I** and tetrasaccharide **II** are fully sulfated species possessing three and six sulfate/sulphamate groups, respectively. Their structures were therefore obvious from disaccharide analysis (i.e. $[\Delta UA(2S)\text{-}GlcNS(6S)]_{1\text{-}2}$), but were confirmed by NMR and MS. Two lesser sulfated tetrasaccharides (**III** and **IV**) were also obtained that possessed asymmetric disaccharide compositions.

The ^{1}H NMR spectra of oligosaccharides **I-IV**, obtained using water presaturation during the relaxation delay, are shown in Fig.III.1.2.1. In these spectra, signals with chemical shifts between 5.0 and 5.5 ppm are those of the anomeric protons. The signal at around 6.0 ppm seen in all spectra belongs to H4 of the terminal uronic acid ring (D ring). In the spectrum of sample **IV**, there is an additional signal at 4.542 ppm. This can be assigned to the H1 proton of a glucuronic acid. In heparin-derived tetrasaccharides glucuronic ring can only exist as a B ring. All other 19 carbon-bonded protons of tetrasaccharides resonate within the 2.9-4.9 ppm region. Inspection of the 1D spectrum also revealed that sample **III** is heterogeneous, which complicated the analysis of the ^{1}H NMR spectra of this compound.

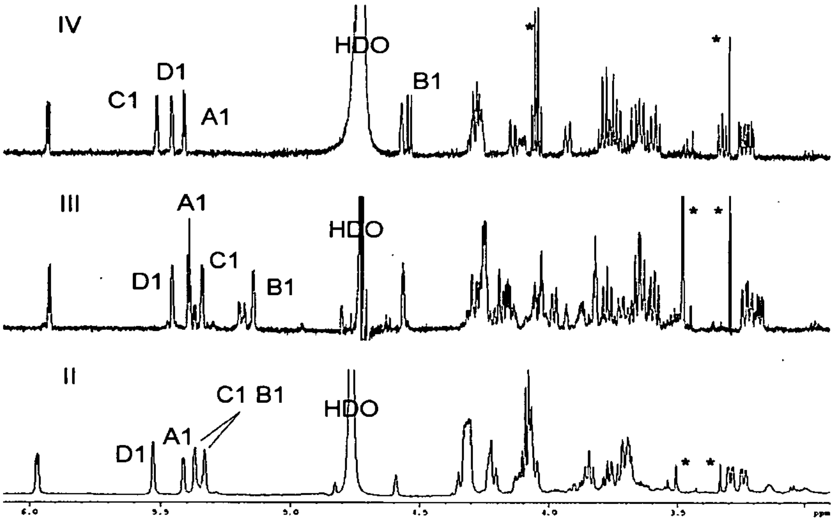
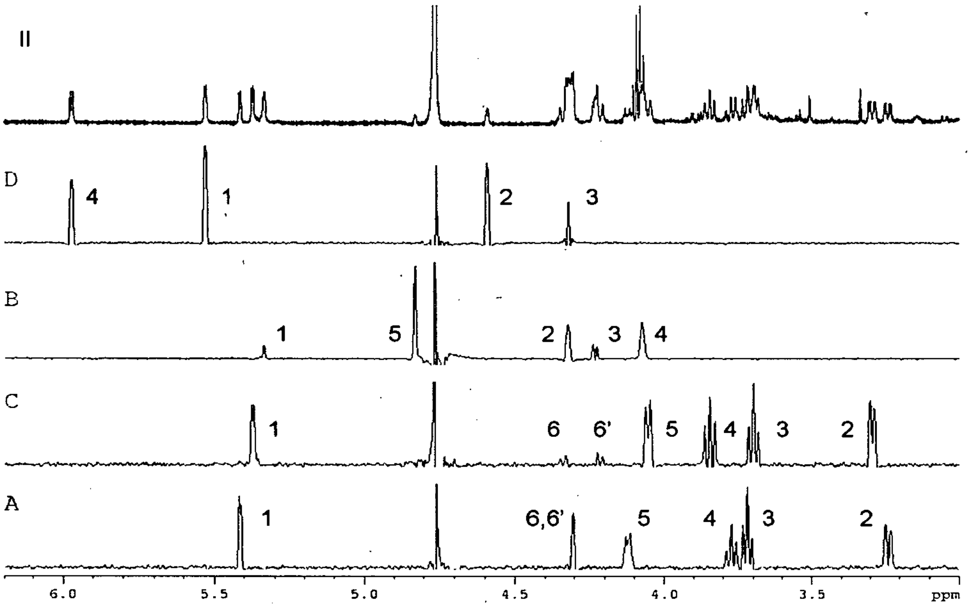
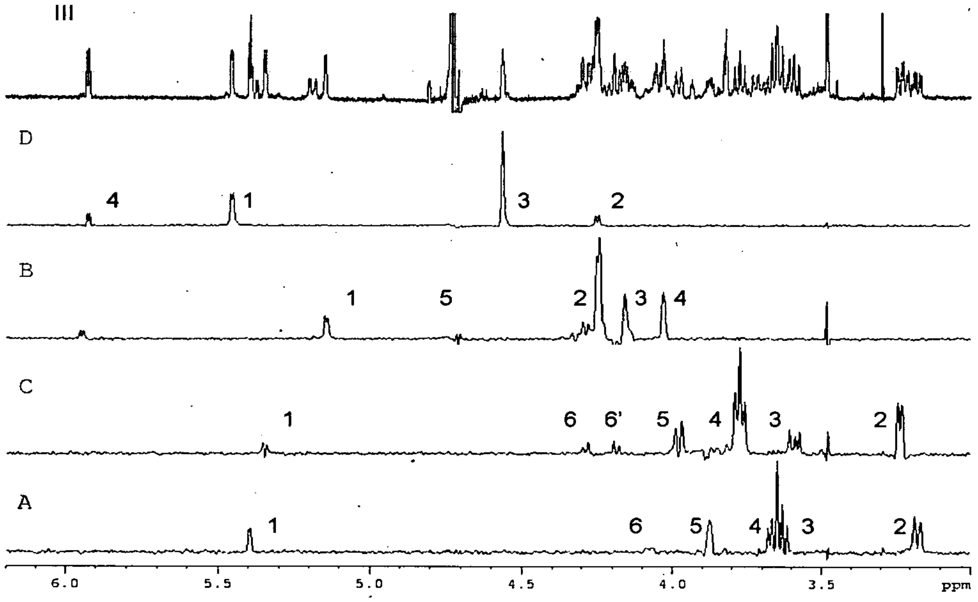


Fig.III.1.2.1 ^{1}H NMR spectra of heparin-derived tetrasaccharides (**II-IV**). Signals labeled with * are non-carbohydrate impurities.

Using 1D traces of 2D TOCSY spectra of samples **II-IV** (Fig.III.1.2.2), four different spin systems associated with four carbohydrate rings were assigned. The signals of two iduronic rings in samples **II** and **III** can be assigned to rings B and D by inspection of TOCSY cross-peaks alone, but this is not the case for the two glucosamine rings. The signals of rings B (glucuronic acid) and D (unsaturated uronic acid) can be unambiguously identified from the TOCSY spectrum of compound **IV**.





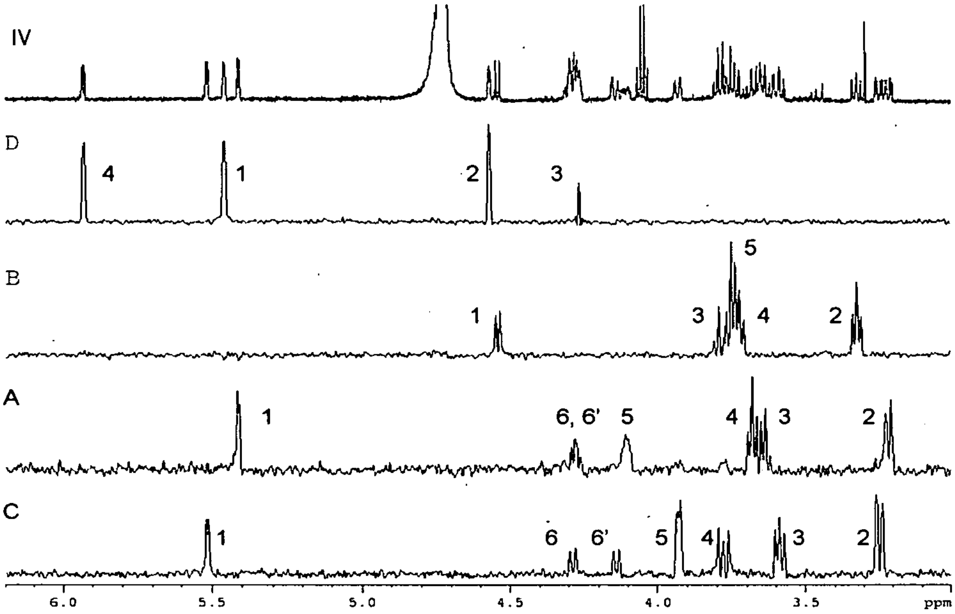


Fig.III.1.2.2 1D traces through anomeric protons taken from 2D TOCSY spectra for samples **II-IV**.

A 2D ROESY experiment (Fig.III.1.2.3) of sample **II** was therefore used to confirm the sequence of monosaccharides as D-C-B-A and assign the glucosamine signals of the two rings A and C. Several ROE cross peaks between protons on both sides of the glycosidic linkage were identified. This allowed the assignment of the chemical shifts for each of the four rings of the three samples (Table III.1.2.1).

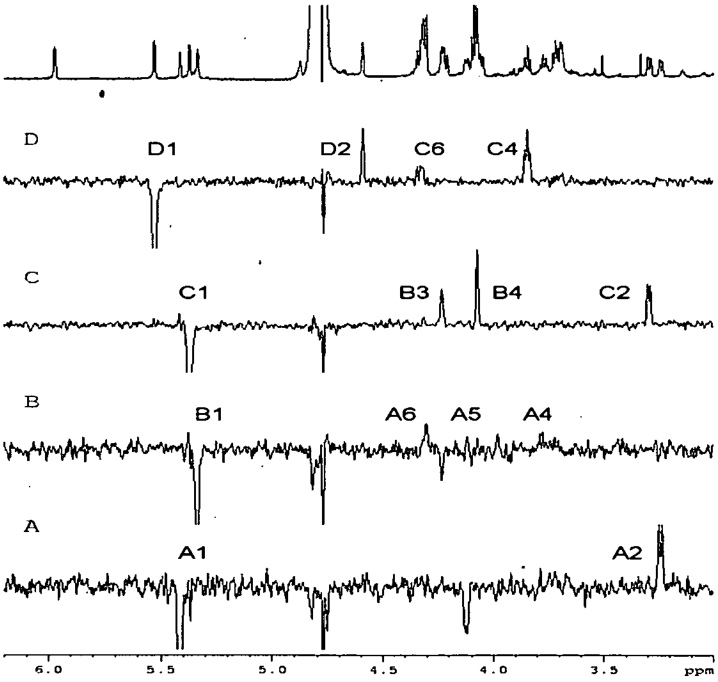


Fig.III.1.2.3 1D traces through anomeric protons taken from 2D ROESY spectrum of sample **II**.

Table III.1.2.1 ^{1}H chemical shifts of the three tetrasaccharides **II-IV**. Literature [33] values are shown in bold; the differences are shown in parentheses

Compound/Rings		H_1	H_2	H_3	H_4	H_5	H_6	$H_{6'}$
II	A	5.379 / **5.430** (-0.051)	3.208 / **3.245** (-0.037)	3.682 / **3.682** (0)	3.739 / **3.732** (0.007)	4.089 / **4.117** (-0.028)	4.271 / **4.290** (-0.019)	4.271 / **4.350** (-0.079)
	B	5.301 / **5.203** (0.098)	4.285 / **4.303** (-0.018)	4.197 / **4.198** (-0.001)	4.035 / **4.097** (-0.062)	4.794 / **4.765** (0.029)	-	-
	C	5.335 / **5.416** (-0.081)	3.260 / **3.285** (-0.025)	3.660 / **3.631** (0.029)	3.809 / **3.823** (-0.014)	4.019 / **4.024** (-0.005)	4.179 / **4.237** (-0.058)	4.301 / **4.341** (-0.040)
	D	5.497 / **5.495** (0.002)	4.556 / **4.612** (-0.056)	4.285 / **4.308** (-0.023)	5.938 / **5.984** (-0.046)	-	-	-
III	A	5.395 / **5.433** (-0.038)	3.174 / **3.220** (-0.046)	3.623 / **3.671** (-0.048)	3.661 / **3.698** (-0.037)	- -	3.823 / **3.860** (-0.037)	3.873 / **3.860** (0.013)
	B	5.145 / **5.196** (-0.051)	4.247 / **4.293** (-0.046)	4.159 / **4.219** (-0.060)	4.031 / **4.065** (-0.034)	4.704 / **4.796** (-0.092)	-	-
	C	5.345 / **5.357** (-0.012)	3.236 / **3.286** (-0.050)	3.591 / **3.645** (-0.054)	3.773 / **3.824** (-0.051)	3.981 / **4.008** (-0.027)	4.185 / **4.231** (-0.046)	4.284 / **4.342** (-0.058)
	D	5.453 / **5.495** (-0.042)	4.562 / **4.612** (-0.050)	4.247 / **4.308** (-0.061)	5.926 / **5.984** (-0.058)	-	-	-
IV	A	5.409 / **5.448** (-0.039)	3.216 / **3.260** (-0.044)	3.637 / **3.694** (-0.057)	3.677 / **3.721** (-0.044)	4.101 / **4.142** (-0.041)	4.290 / **4.330** (-0.040)	4.290 / **4.330** (-0.040)
	B	4.542 / **4.589** (-0.047)	3.324 / **3.372** (-0.048)	3.793 / **3.839** (-0.046)	3.713 / **3.771** (-0.058)	3.749 / **3.812** (-0.063)	-	-
	C	5.517 / **5.555** (-0.038)	3.248 / **3.292** (-0.044)	3.587 / **3.631** (-0.044)	3.775 / **3.823** (-0.048)	3.927 / **3.976** (-0.049)	4.137 / **4.188** (-0.051)	4.286 / **4.333** (-0.047)
	D	5.459 / **5.495** (-0.036)	4.572 / **4.612** (-0.040)	4.264 / **4.308** (-0.044)	5.934 / **5.984** (-0.050)	-	-	-

Overall, there is a good agreement between the measured ^{1}H chemical shifts and the literature data for each tetrasaccharide. Nevertheless there are some differences. The largest difference (0.098 ppm) was found for H1 (B) of compound **II**. The reason for this is likely to be the different sample conditions. Sample **II** used in this work was at pH 7.5 and 4:1 excess of Ca^{2+} was used.

In tetrasaccharide **III,** position C6 of ring A is not sulfated (i.e. GlcNS). In the case of tetrasaccharide **IV**, ring B is a nonsulfated GlcA rather than a 2-O-sulfated IdoA as in **II** and **III**. All oligosaccharides contain a C4-C5 double-bond in the non-reducing uronate ring D resulting from the enzymatic lyase cleavage that was used to excise the oligosaccharides from the heparin polysaccharide. At the reducing end the α-anomers are the prevailing species in solution (α:β = 9:1), as determined by NMR. There is no significant difference in the theoretical collision cross-sections between these two forms. This is in agreement with the experimental ATDs that did not show any indications of separate peaks attributable to different anomeric forms. Only α-anomeric configurations of the reducing end monosaccharides are therefore considered below.

III. 1.3 Measurement of the experimental collision cross-sections[ii]

The ions generated from disaccharide **I** showed three distinct arrival distribution times (Fig.III.1.3.1a) at an apparent mass-to-charge ratio of 642.4 m/z. Early in the data analysis it became clear that the species with longer ATDs could not be reproduced by considering only a monomeric ion ($[C_{12}H_{15}NO_{19}S_3Na_3]^-$) and these corresponded to a dimer and trimer of **I** ($[C_{12}H_{15}NO_{19}S_3Na_3]_2^{2-}$ and $[C_{12}H_{15}NO_{19}S_3Na_3]_3^{3-}$, respectively) that were formed in the collision cell. The ion mobility, K, was obtained via a plot of arrival times versus the pressure of helium divided by the drift voltage (Fig.III.1.3.1b). The experimental cross-sections were calculated using those values from **Eq.I.5.2** (Table III.1.3.1).

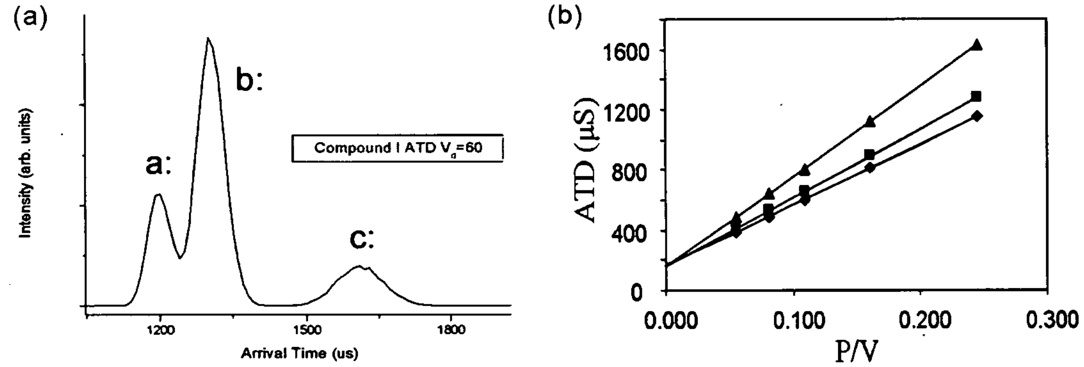


Fig.III.1.3.1 (a) Arrival time distributions for the 642.4 m/z ion of disaccharide **I**. An electric gradient of 60 V was used. The peaks labelled a-c correspond to a monomer, dimer and trimer of **I**, respectively. (b) Determination of experimental ion mobilities, K, for three arrival time distributions observed for disaccharide monomer (◆), dimer (■) and trimer (▲) of **I**. (Data from Table III.1.3.1)

Table III.1.3.1 Experimental data of ion mobilities of the disaccharide ion **I**.

Fully sulfated disaccharide (I)						
Pressure(bar)	4.965	4.939	4.957	4.943	4.953	
Voltage (V)	91.39	61.20	45.69	30.70	20.13	
P/V	0.054	0.081	0.108	0.161	0.246	
ATD (μs)	385	490	600	810	1150	Monomer
	410	530	653	890	1280	Dimer
	485	645	805	1120	1625	Trimer
$K_0(cm^2V^{-1}s^{-1})$	4.024		5.274		5.992	
σ (Å^2)	132.1		201.6		266.8	
	Monomer		Dimer		Trimer	

[ii] All experiments were carried out by Dr. Perdita E. Barran at UCSB (USA)

The ions generated using tetrasaccharides **II-IV** were principally monomeric and corresponded to the following species: $[C_{24}H_{30}N_2O_{38}S_6Na_6]^{2-}$ (**II**, 642.4 m/z) and $[C_{24}H_{31}N_2O_{35}S_5Na_5]^{2-}$ (**III**, **IV**, 591.4 m/z). The arrival time distributions of the oligosaccharides **I-IV** were analysed as explained above for the disaccharide, and the resulting experimental collision cross-sectional areas are summarized in Table III.1.3.2. The number of Na atoms lost during the ionization process (one for the disaccharide and two for tetrasaccharides) corresponds to the number of carboxyl groups, implying a stronger affinity of sulfate and sulfamate groups for sodium compared with carboxyl groups.

Table III.1.3.2 Experimental and theoretical cross sectional areas of oligosaccharides **I-IV.**

	Compound I			II	III	IV
σ (Å^2)	Monomer	Dimer	Trimer			
Experimental	132.1	201.6	266.8	198.4	190.4	191.6
Calculated	134.8	191.7	250.2	204.1	190.3	191.8

III. 1.4 Modelling of the gas phase structures and calculation of the theoretical collision cross-sections

Candidate structures of oligosaccharides and their oligomers were obtained by molecular modelling described in Chapter II.6.2. For each of these structures the theoretical collision cross-sections were calculated (Table III.1.3.2).

The ions were modelled using the projection approximation method as a collection of overlapping hard spheres with radii equal to hard sphere collision distances, corrected for collision temperature [78]. The orientationally averaged geometric cross-section was determined by averaging the geometric cross-section over all possible collision geometries. The program Mobcal [79] was used to calculate the theoretical collision cross-sections for the candidate oligosaccharides. Fig.III.1.4.1 shows plots of theoretical collision cross-sections of 100 conformations of the sodiated heparin-derived oligosaccharide ions.

The average theoretical collision cross-sections reported in the Table III.1.3.2 for each of the oligosaccharide models were obtained by averaging the collision cross-sections of 10% lowest potential energy structures. The agreement between the experimental and theoretical cross-sections is very good. For monomers the differences were < 2.1%, while for oligomers they were < 6.2%. Representative structures of the oligosaccharides are shown in Fig.III.1.4.2.

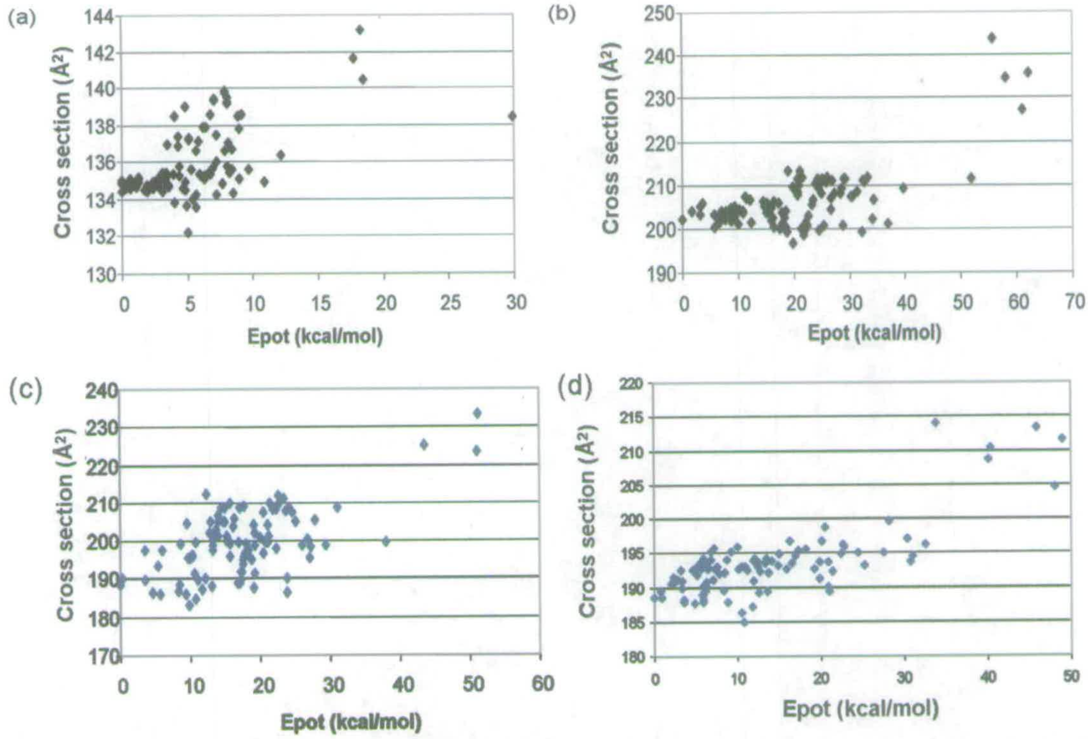


Fig.III.1.4.1 Theoretical cross sections of 100 candidate structures of (a) fully sulfated disaccharide **I**, (b) fully sulfated tetrasaccharide **II**, (c) tetrasaccharide **III** and (d) tetrasaccharide **IV** as a function of the potential energy.

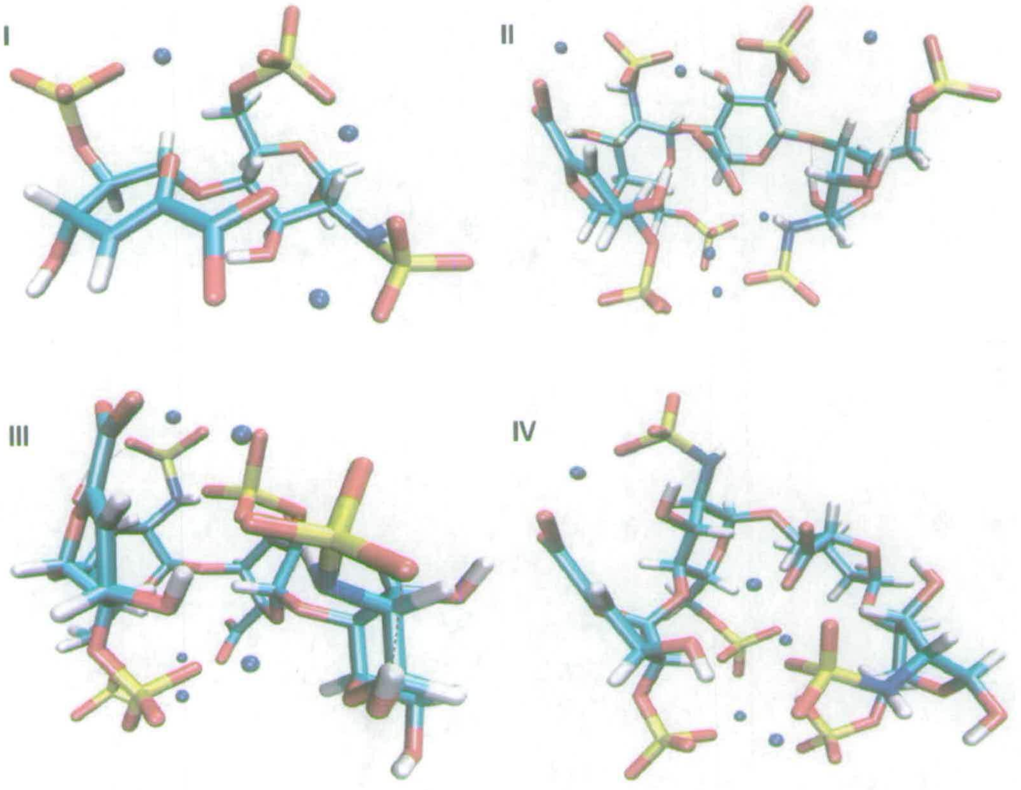


Fig.III.1.4.2 Typical structures of sodiated oligosaccharides, blue dots are sodium ions. **I** ΔUA(2S)-GlcNS(6S)-Na_3; **II** ΔUA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S)-Na_6; **III** ΔUA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS-Na_5 and **IV** ΔUA(2S)-GlcNS(6S)-GlcA-GlcNS(6S)-Na_5

III. 1.5 Analysis of the modelled gas phase structures

Ten lowest energy structures of the monomeric disaccharide **I** ion, $[C_{12}H_{15}NO_{19}S_3Na_3]^-$, differed only slightly in the orientation of their functional groups (Fig.III.1.5.1).

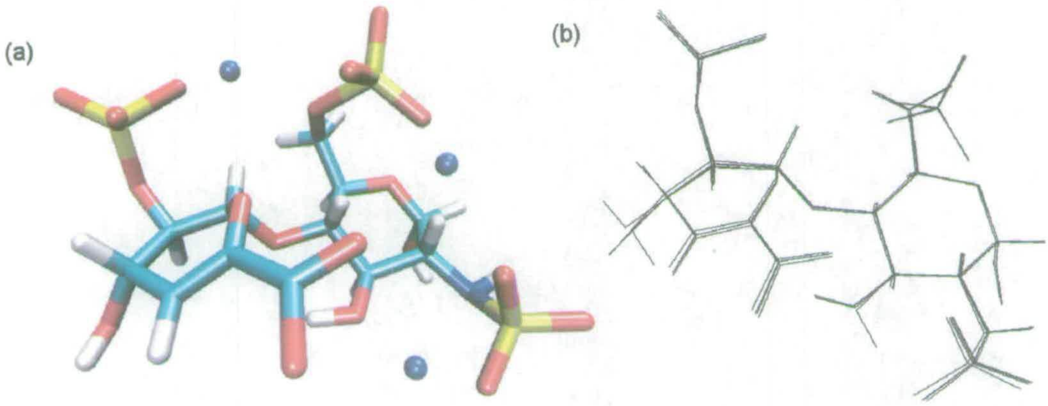


Fig.III.1.5.1 Low energy structures of the ΔUA(2S)-GlcNS(6S)-Na$_3$ ($[C_{12}H_{15}NO_{19}S_3Na_3]^-$) disaccharide **I** ion. (a) A typical structure of **I**. Na$^+$ is shown as spheres. (b) Overlay of ten lowest energy structures. The structures were generated using MolMol [80].

The ΔUA ring (D) was always found in the $^1H_2$ conformation while the glucosamine residue (ring A) invariably occupied the $^4C_1$ conformation. Two Na$^+$ ions were coordinated by two sulfate groups, while the third was positioned between the sulfamate and carboxyl groups. As a general feature of all structures simulated, the Na$^+$ ions were found between 2.2 and 2.9 Å from oxygen atoms of at least two sulfate, sulfamate or carboxyl groups.

Although the prevailing conformations of rings D and A in the dimer and the trimer (Fig.III.1.5.2) were the same as observed for the monomer disaccharide ($^1H_2$ and $^4C_1$), other forms ($B_{3,O}$ for ring D and $^1S_5$, $^5S_3$ for ring A) were also found. The existence of these forms was attributed to the presence of the Columbic interactions between the Na$^+$ ions and negatively charged groups causing distortions of the ring conformations. In all dimers the Na$^+$ cations were simultaneously coordinated to both monomer units. In the trimers all Na$^+$ ions were coordinated by at least two monomers, while some showed simultaneous interactions with all three monomers. In addition, several hydrogen bonds were observed between some monomers using the criterion for the distance between the atom, carrying hydrogen donor (D) and oxygen acceptor (A) < 3 Å and the angle |D-H-A| < 20°.

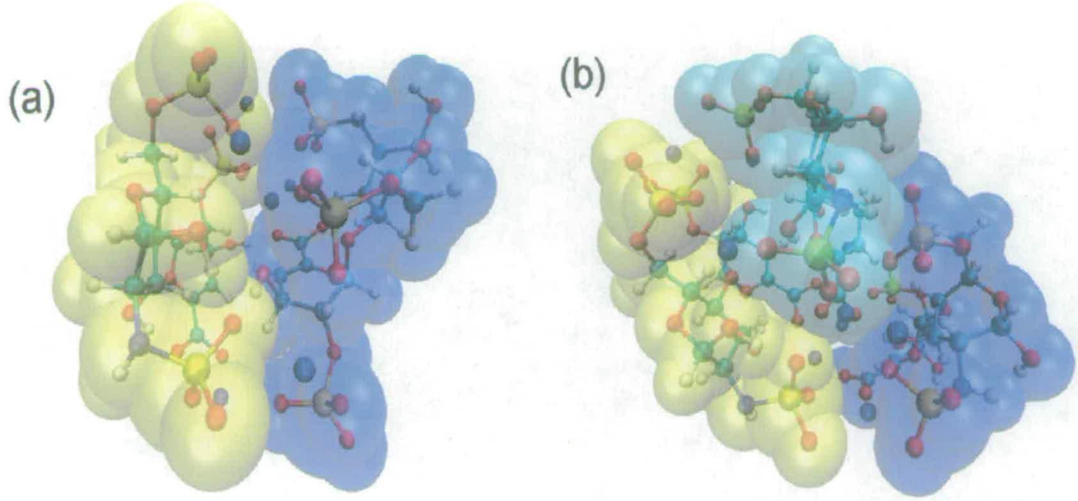


Fig.III.1.5.2. Representative structures of (a) $[C_{12}H_{15}NO_{19}S_3Na_3]_2^{2-}$ and (b) $C_{12}H_{15}NO_{19}S_3Na_3]_3^{3-}$ ions corresponding to the dimer and trimer of **I**. The structures were generated with VMD [81]. Individual monomers are highlighted using van der Waals radii of their atoms.

The overlay (Fig.III.1.5.3) of five of the lowest energy structures of ions tetrasaccharide **II** showed some heterogeneity in the conformation of individual monosaccharide rings. In particular, both $^1H_2$ and $^2H_1$ forms were present in ring D. Ring A also showed, in addition to a $^4C_1$ conformation, both boat and skew boat forms. The IdoA2S (ring B) consistently occupied a $^4C_1$ conformation, although the solution structure of **II** indicates the prevalence of the ^2S_o form [31a].

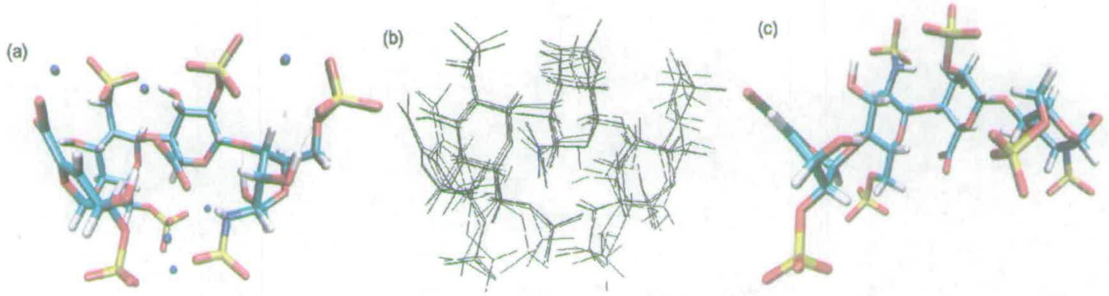


Fig.III.1.5.3 Fully sulfated tetrasaccharide **II**. (a) Representative structures of $C_{24}H_{30}N_2O_{38}S_6Na_6^{2-}$ ion. (b) Overlay of the five lowest energy structures (c) Crystal structure [86] of **II** in complex with bFGF.

Calculations performed on this tetrasaccharide with the Born Solvent model [31b] resulted in interconversion between the usual $^1C_4$ and ^2S_o forms for the B ring. This indicates that the $^4C_1$ is specific to the gas phase rather than the AMBER force-field we have used. In order to see whether different conformations of the iduronic acid could be distinguished by IMMS, theoretical collision cross-sections of 89.35 Å^2 ($^1C_4$), 87.80 Å^2 ($^4C_1$) and 89.40 Å^2 ($^2S_0$) were calculated. As can be seen, the difference is < 2%, which would not be detected by an IMMS instrument. Nevertheless, in the context of a larger molecule different conformations of IdoA2S monosaccharide may influence the overall shape of the molecule and therefore be indirectly traceable via molecular modelling.

A notable feature of the lowest energy structures of the tetrasaccharide ions is their 'compact' shape (Fig.III.1.5.3b), which differs from the more extended conformation observed in the X-ray structure [86] of **II** complexed with basic fibroblast growth factor (bFGF) (Fig.III.1.5.3c). The theoretical cross section of the X-ray structure (231.8 Å^2) is substantially larger than that of the corresponding low energy structures in the gas phase (204.1 Å^2). Clearly, the value obtained from the calculated structures is much closer to the experimental value of 198 Å^2. Even better agreement between the theoretical and experimental cross sections was observed for tetrasaccharide ions **III** and **IV** (Table III.1.3.2), indicating that the calculated model structures are a true representation of their gas phase conformations (Fig.III.1.5.4 and Fig.III.1.5.5).

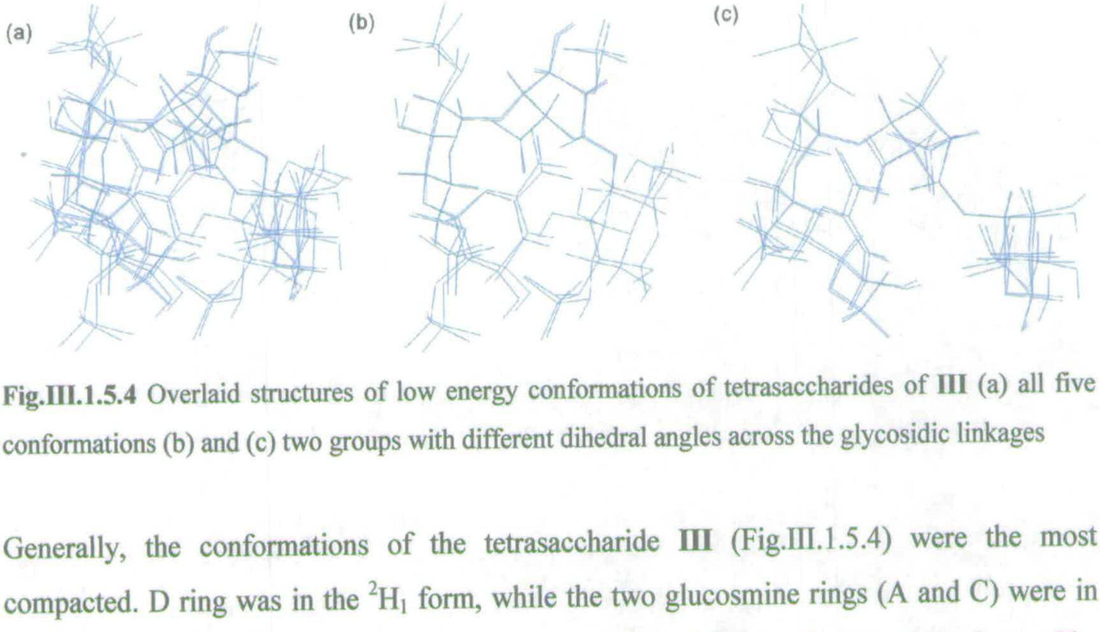


Fig.III.1.5.4 Overlaid structures of low energy conformations of tetrasaccharides of **III** (a) all five conformations (b) and (c) two groups with different dihedral angles across the glycosidic linkages

Generally, the conformations of the tetrasaccharide **III** (Fig.III.1.5.4) were the most compacted. D ring was in the $^2H_1$ form, while the two glucosmine rings (A and C) were in the $^4C_1$ conformation. However, the IdoA2S (ring B) were in the $B_{1,4}$ form. The conformational changes across the glycosdic linkages of the tetrasaccharide **III** (Fig.III.1.5.4b, c) did not affect the overall collision cross-sections of the whole molecule.

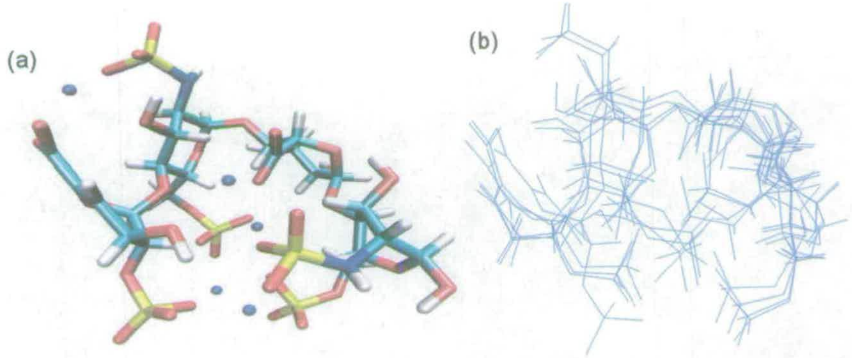


Fig.III.1.5.5 Low energy structures of the tetrasaccharide **IV** ion. (a) A typical structure of **IV**. Na^+ is shown as spheres. (b) Overlay of four lowest energy structures.

Conformation of low energy structures of tetrasaccharide **IV** was heterogeneous (Fig.III.1.5.5b) but different conformers had similar cross-sections. Both $^1H_2$ and $^2H_1$ forms were obtained for the D ring, while both $^4C_1$ and $^1C_4$ forms existed in the C ring. The GlcA (B ring) was in the $^{1,4}B$ form and the A ring showed $^1S_5$ form. These special conformations are very likely caused by distortions from the interactions between the negatively charged sulfate and carboxyl groups of **IV** and the sodium ions. The following tables summarize the dihedral angles of the lowest energy conformers of gas phase oligosaccharide ion structures.

Table III.1.5.1. Structural parameters of the lowest energy conformers of the disaccharide **I** and its oligomers.

	Ring D	Ring A	Φ	Ψ
Monomer	$^1H_2$	$^4C_1$	64	28
X-ray [86]	$^1H_2$	$^4C_1$	42.3	18.3
Solution [85]	$^1H_2$	$^4C_1$	45.7	13.2
Dimer	$^1H_2$	$^1S_5$	41	-30
	$^1H_2/B_{3,O}$	$^4C_1$	45	-16
Trimer	$^1H_2$	$^4C_1$	26	-18
	$^1H_2/B_{3,O}$	$^4C_1$	27	-4
	$^1H_2/B_{3,O}$	$^5S_3$	45	-47

Table III.1.5.2 Structural parameters of the lowest energy conformers of the tetrasaccharide ions **II**, **III** and **IV**

Dihedral angle	Crystal structure [86]	Solution structure [85]	**II**	**III**	**IV**
Φ_{DC}	42.3	45.7	175.8	172.0	61.0
Ψ_{DC}	18.3	13.2	-149.4	-159.7	20.0
Φ_{CB}	-18.5	-43.3	-48.8	-3.9	-129.8
Ψ_{CB}	-3.5	-42.2	-10.8	5.0	-11.2
Φ_{BA}	57.0	45.2	25.5	176.6	34.7
Ψ_{BA}	25.1	15.0	-78.5	0.8	-30.0

It can be seen that a small change in the primary structure can lead to very different conformations for tetrasaccharides **II-IV** in the gas phase. For example, tetrasaccharides **II** and **III**, which differ only in one sulfate group at position A6, show markedly different conformations along the CB and BA glycosidic linkages. Similarly, GlcA/IdoA2S isomerism (compounds **II** and **IV**) resulted in very different conformations. This is an interesting result showing that, at least in the gas phase, differences in the primary structures of HS/heparin sequences can lead to different conformations. It is also possible that the conformations revealed by IMMS exist transiently in solution generating short-lived "kinks and turns" in otherwise regular, helical structures of HS/ heparin [82]. It can also be seen that large differences exist between these structures and the conformations in solution or solid state. In order to investigate the source of these differences next we have performed valuation in vacuum in the absence of sodium cations.

III. 1.6 Molecular modelling of heparin-derived tetrasaccharide ions in the gas phase and in the absence of Na^+ ions

Compact structures of gas phase ions have been observed in an IMMS study of neutral, sodiated carbohydrates [83], and recent work by Hill et al. [84] reports the use of atmospheric pressure ion mobility to separate sodiated di- and trisaccharides. In our work, interactions of Na^+ with negatively charged sulfates, sulfamates and carboxyl groups resulted in the formation of higher oligomers of disaccharide ions; aggregation was not reported by Hill. Are these interactions also responsible for the "compact" shape of the tetrasaccharide ions that reproduced so well our IMMS data? Or is it because of the newly parameterized force field favours certain conformations around the glycosidic linkage leading to compact structures? In order to answer these questions, structures of three disaccharide segments, contained in each of the tetrasaccharides **II-IV**, in the absence of Na^+, were calculated. For these calculations conformations of individual rings corresponding to the most populated species found in the modelled gas phase structures in the presence of Na^+ were used. The results of these calculations are presented in the form of (Φ, Ψ) energy maps (Fig.III.1.6.1), that were calculated as described in Chapter II.6.2.3. On these maps, the dihedral angles observed in the model structures of the gas phase ions (Tables III.1.5.1 and III.1.5.2), and also those found in the solution [85] and solid state structures [86] are indicated.

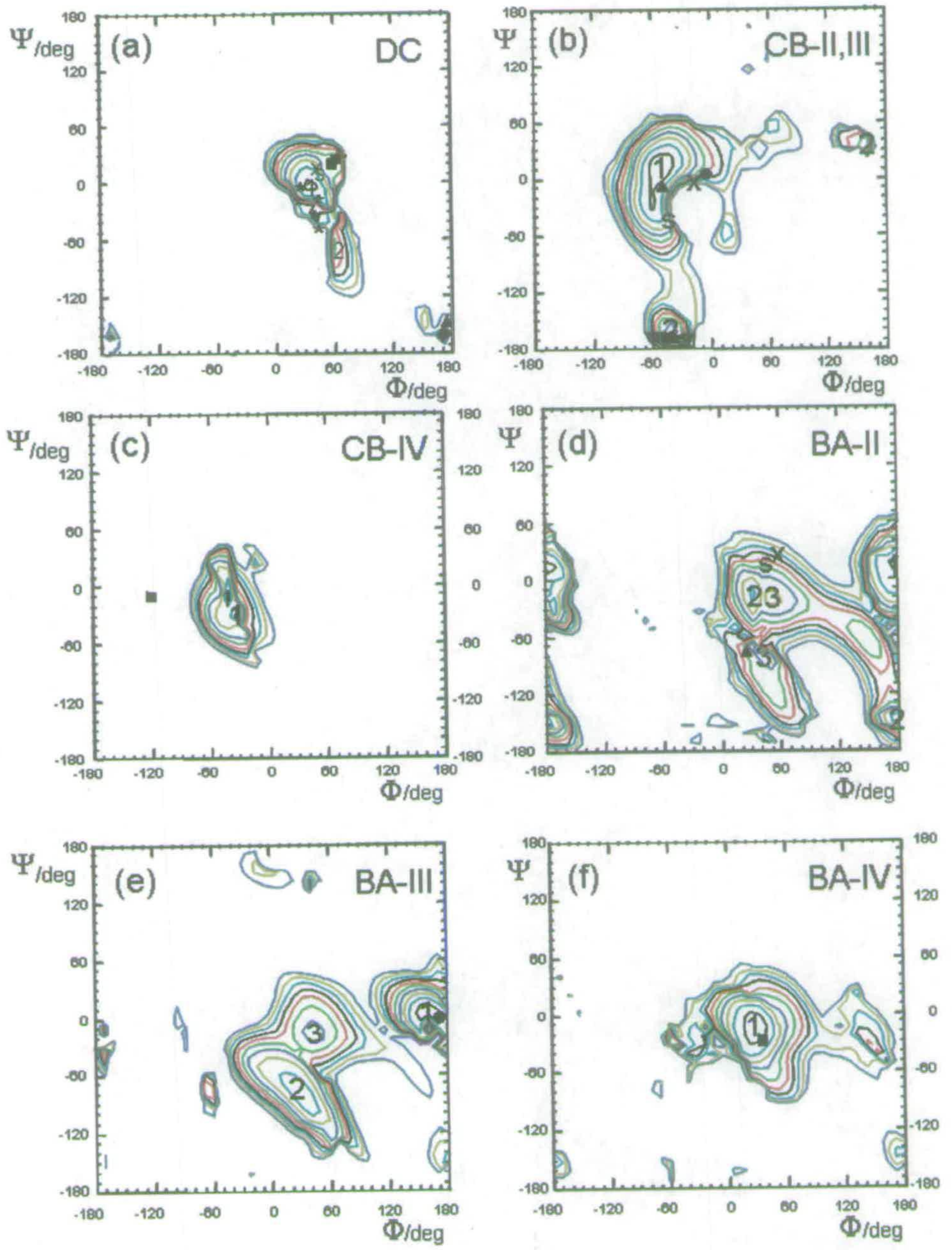


Fig.III.1.6.1 (Φ, Ψ) energy maps calculated for disaccharide fragments of oligosaccharides **I-IV**. The following conformations of individual rings were considered: DC: $^{1}H_{2}$-$^{4}C_{1}$; CB and BA: $^{4}C_{1}$-$^{4}C_{1}$. Ten isoenergy contours were drawn by interpolation of 1 kcal/mol above the global minimum. The conformation found for disaccharide ions are marked by + (monomer) and * (dimer and trimer). Conformations found for tetrasaccharide ions **II**, **III** and **IV** are labelled using ▲, • and ■, respectively. The dihedral angles found in the solution structures of the tetrasaccharide **II** are marked by **x** and **s**, respectively. (a) CD disaccharide fragments. The energy minima labelled as **1** and **2** are separated by 4 kcal mol^{-1}. (b) Rings C and B in tetrasaccharide ions **II**, **III**. The energy minima **1** and **2** are separated by 3 kcal mol^{-1}. (c) Rings C and B in the tetrasaccharide ion **IV**. (d) Rings B and A in tetrasaccharide ion **II**. The energy minima marked as **1**, **2** and **3** are separated by 1 kcal mol^{-1}. (e) Rings B and A in tetrasaccharide ion **III**. The energy minima **1** and **2**, and **1** and **3** are separated by 1 and 3 kcal mol^{-1}, respectively. (f) Rings B and A in tetrasaccharide ion **IV**.

Starting from the non-reducing end of the molecule, the DC disaccharide fragment was first examined. The dihedral angles observed for the monomeric disaccharide ion $[C_{12}H_{15}NO_{19}S_3Na_3]^-$ deviate up to +30° on both angles from the deepest minimum of the (Φ, Ψ) map. Deviations in the opposite direction, and of up to -45° in the Ψ_{DC} angle, were observed for dimeric and trimeric disaccharide ions. The largest differences were however observed for the DC linkages in the tetrasaccharide ions **II** and **III**, corresponding to the existence of anti conformers (symbols ▲ and • in Fig.III.1.6.1a). This region of the (Φ, Ψ) map was not particularly energetically favoured in the absence of the sodium, but presumably became stabilized by the Na^+ cations interacting with both D and A rings. The corresponding dihedral angles observed in the X-ray and solution structures were close to the minimum of the (Φ, Ψ) map.

The next two disaccharide fragments can only be examined in the context of the tetrasaccharides. Considering tetrasaccharide ions **II** and **III** first, the CB fragments are structurally identical in these two compounds and the dihedral angles of the tetrasaccharide ion **II** are close to the minimum of the (Φ, Ψ) map (Fig.III.1.6.1b). However, in compound **III** the Φ_{CB} angle differs by 45°. It is possible that this difference is a consequence of a different conformation found in the BA disaccharide fragment of ions **II** and **III** (see Table III.1.5.1 and Fig.III.1.6.1d,e), and it is therefore not possible to attribute this difference solely to effects of the Na^+ ions. On the other hand, a large difference in the dihedral angle Φ_{CB} (80°) was observed between the sodium-free disaccharide fragment and the gas phase conformation of tetrasaccharide ion **IV** (Fig.III.1.6.1c). The latter clearly occupies a high energy region of the (Φ, Ψ) map.

Because of the differences in the primary sequence of tetrasaccharides **II-IV**, three separate cases must be considered when analysing the BA disaccharide fragment. In compound **II** the gas phase ion structure is not found in the global minimum (Φ, Ψ) = (180°, 10°), but occupies a third minimum (+2 kcal mol^{-1}) with coordinates (26, -79). This conformation also differs by ~105° in the Ψ_{AB} angle from the X-ray and solution structures. Ions **III** and **IV** on the other hand show conformations coinciding with the global minima of the sodium-free disaccharide segments (Fig.III.1.6.1e, f).

The analysis presented above showed that considerable differences exist between the structures calculated in the presence or absence of Na^+. It appears that sodium atoms play an important role in stabilising the gas phase conformation of ions, leading to the appearance of "compact" structures of tetrasaccharide ions. In these structures rings A and D are brought close to each other and bridged by sodium cations. In order to accommodate such conformations, some glycosidic linkages adopt conformations that differ from those calculated for the sodium free disaccharide segments, or those observed in the solution or X-ray structures of tetrasaccharide **II**.

Na^+ ions interact differently with the negatively charged groups of GAGs in solution and in the gas phase. In solution the hydrated sodium counter-ions are delocalized in the volume around the polyanion [87], while in the gas phase the Na^+ cations are localized near several negatively charged groups as seen in this study. This effect will lessen the Columbic repulsion that isolated sulfate groups would impart on each other, and has the effect of 'tightening' the structures. This explains the lower cross section obtained in the gas-phase compared with that from X-ray data.

In summary, the conformational analysis of ions of heparin-derived oligosaccharides using ion-mobility mass spectrometry and molecular modelling was performed in this project. Negative mode electrospray ionisation has produced singly (disaccharide) and doubly charged (tetrasaccharide) ions containing three and six Na^+ ions, respectively. A good agreement was obtained between the experimental and theoretical cross sections. The latter were obtained using modelled structures generated by the AMBER-based force field. Analysis of the data shows that the sodium cations play a major role in stabilizing the ions in the gas phase. This was seen in the formation of oligomers of the disaccharide and "compact" structures of tetrasaccharide ions. Interestingly, the conformations of the three tetrasaccharides, differing by relatively small changes in primary structure, were significantly different. The Columbic interaction between sodium ions and negatively charged carboxyl and sulfate groups of oligosaccharides distorted their conformations. This made the whole structure compact.

By using the ionization conditions as described by Zaia and Costello [88] it may be possible to obtain Na^+-free ions of heparin-derived oligosaccharides for ion-mobility studies that would bear a closer resemblance to conditions seen in solution. In particular, this may be possible for low sulfated HS species which so far have received much less attention compared with highly sulfated domains. Although the low sulfated species are less likely to be involved in protein recognition, they may well be the source of flexibility in HS/heparin polymers and this could be probed by IMMS. A significant advantage of IMMS over other biophysical methods applicable to conformational studies of biomolecules is that the ion-mobility experiments can be performed at near liquid nitrogen temperatures. Under these conditions different conformers can potentially be separated providing they exhibit distinct cross sectional areas. Other methods, such as NMR, can only report on the time-averaged structures.

III. 2 Development of methods for the measurement of J_{HH} and D_{HH} coupling constants

III. 2.1 *J*-modulated 1D directed COSY

We have developed a set of robust intensity based 1D NMR methods for accurate and precise measurement of the scalar and dipolar proton-proton coupling constants for small and medium size molecules. These techniques work despite the presence of overlapping and broad ^{1}H multiplets that are typically observed for molecules in the weakly aligned media. The new methods were developed based on the *J*-modulated 1D directed COSY experiment which, in the original form [89], can only be used for compounds with resolved ^{1}H resonances. Overlapping ^{1}H resonances in spectra of small and medium size molecules cannot be directly used for the selective magnetization transfer. Therefore a more selective technique is required that is capable of selecting one proton despite extensive overlap with other protons.

The original technique was a single-pulse field-gradient spin-echo (SPFGSE, Fig.III.2.1.1) first to select the magnetization of proton *k*, which is, amongst other protons, coupled to proton *l*. This is followed by a variable evolution interval, *T*, during which only protons *k* and *l* are selectively inverted. The signal intensity of proton *l* generated by the subsequent polarization transfer step depends on the length of this evolution interval. The experiment is repeated several times while varying *T* and the coupling constants J_{kl} are obtained by fitting the signal intensities of proton *l* using a simple transfer function [89]. An optional refocusing interval, which partially restores the in-phase magnetization, is advantageous and recommended for aligned samples and/or small coupling constants.

The high precision of coupling constant determination achieved by this experiment is due to the fact that the fitted signal is modulated by no more than one coupling constant regardless of how many other protons may have a scalar or a dipolar coupling with proton *k*. This basic method, SPFGSE-COSY, is limited to instances when proton *k* is sufficiently isolated to allow for its selection by a SPFGSE. The COSY evolution interval can be either variable or constant time (Fig.III.2.1.1). The constant-time version is more suitable for the measurement of very small splittings (< 1 Hz), since the relaxation time can be omitted from the fitting, which increases the accuracy of the minimization procedure.

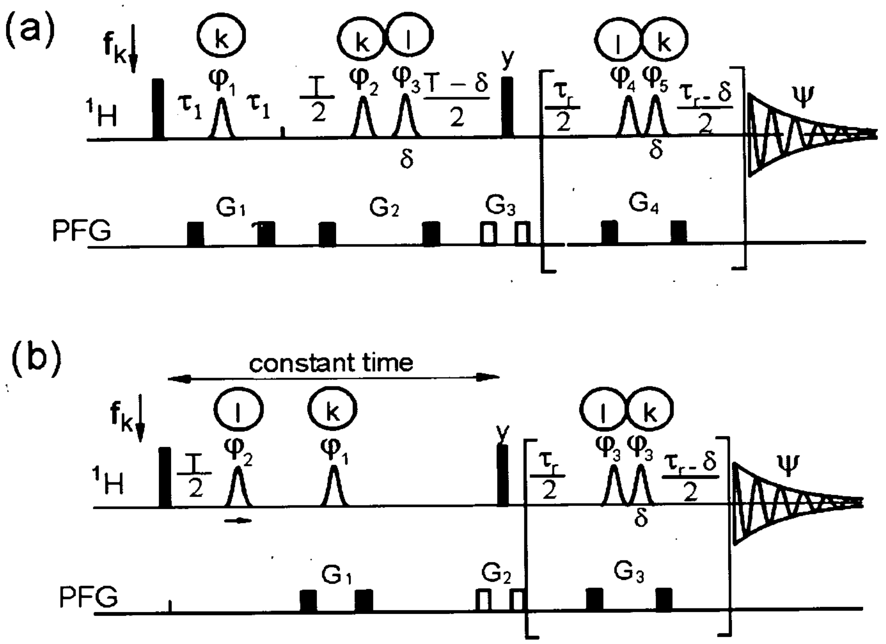
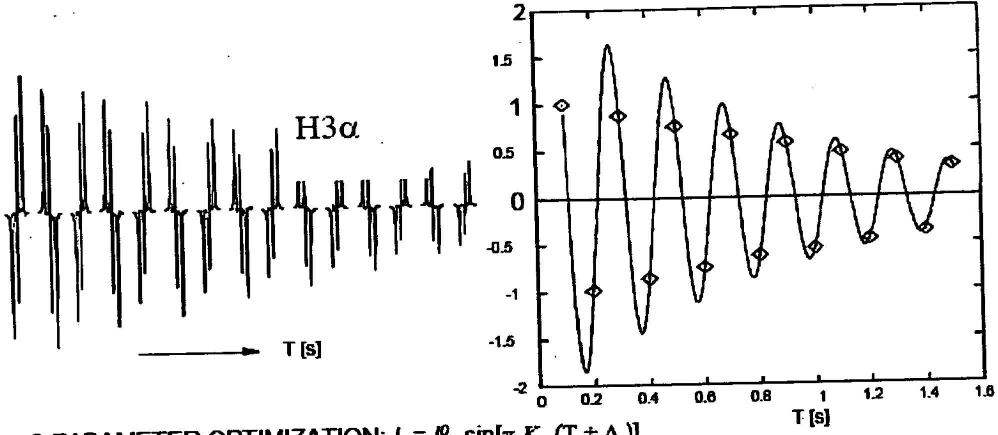


Fig.III.2.1.1 Filled and open rectangles represent 90° and 180° nonselective pulses, respectively, applied from the x axis unless specified otherwise. Selective 180° Gaussian pulses are indicated by open Gaussian envelopes and applied to spins j, k or l, as indicated. The refocusing interval enclosed in square brackets is optional; τ_1 = 1.2 ms and T and τ_r are the variable and refocusing delays, respectively. Sine shaped gradient pulses were applied along the z (open squares) or x, y directions (filled squares). All pulse field gradients had duration δ = 1 ms unless specified otherwise. All phase cycling given for pulse sequences, derived using the Bruker simulation program, ensure that distortion-free multiplets are acquired when 30-70 ms selective Gaussian pulses are used during the J-modulated COSY step. (a) A variable-time 1D directed-COSY. Gradient strengths were G_1 = 20 G/cm, G_2 = 15 G/cm, G_3 = 8 G/cm, G_4 = 11 G/cm and the following phase cycling was applied: φ_1 = x, y; φ_2 = 2x, 2y, 2(-x), 2(-y); φ_3 = 8x, 8y, 8(-x), 8(-y); φ_4 = x; φ_5 = 32x, 32(-x); ψ = x, 2(-x), x. (b) Constant-time J-modulated 1D directed COSY. τ_1 = 1.1 Gradient strengths were G_1 = 15 G/cm, G_2 = 8 G/cm, G_3 = 11 G/cm and the following phase cycling were applied: φ_1 = x, y; φ_2 = 2x, 2y, 2(-x), 2(-y); φ_3 = 4x, 4(-x); ψ = x, 2(-x), x.

As an example of the fitting process and a comparison of the data between the variable and constant-time methods, the determination of the $J_{H2\alpha H3\alpha}$ coupling constant of Me-β-D-cellobiose is shown in Fig.III.2.1.2. Although these data were obtained using the newly developed methods for the measurement of ^{1}H-^{1}H coupling constants, the principles behind the fitting are the same as for the basic method.

a) 4-PARAMETER OPTIMIZATION: $I_l = I^0_k \sin[\pi K_{kl}(T + \Delta)] \exp[-(T + \Delta)/T_{2k}^{eff}]$
Iterations: 11
Minimum found at:
K_{kl} = 9.86 Hz I^0_k = 2.29 T_{2k}^{eff} = 0.80s Δ = -15.34 ms
Minimum function value = 0.005130



b) 3-PARAMETER OPTIMIZATION: $I_l = I^0_k \sin[\pi K_{kl}(T + \Delta)]$
Iterations: 11
Minimum found at:
K_{kl} = 9.84 Hz I^0_k = 1.05 Δ = -57.22 ms

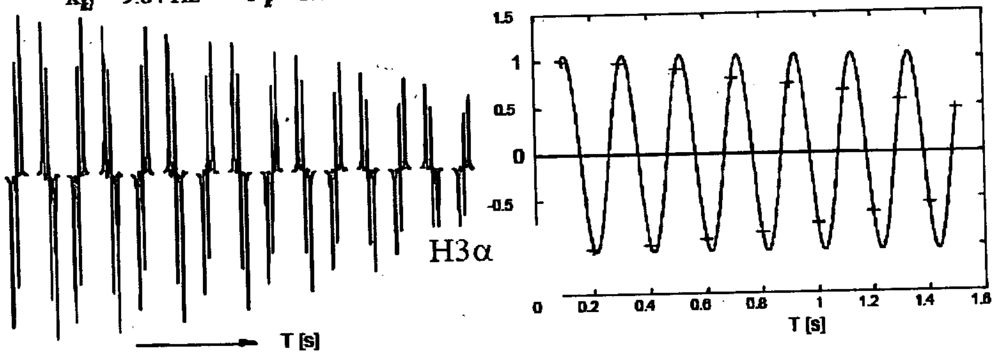


Fig.III.2.1.2 Coupling constant determination [$J_{H2\alpha H3\alpha}$ of Me-β-D-cellobiose] from signal intensities. I_l is the signal intensity of proton l, I^0_k is the scaling factor; $K_{kl} = J_{kl}$ or $J_{kl} + D_{kl}$; T_{2k}^{eff} is the effective relaxation time of proton k; Δ is the effective K evolution during the selective pulses.

As can be seen from this example, the coupling constants are determined by fitting the peak intensities to the COSY delays. In order to obtain accurate intensities the peaks need to have pure phase, especially in the case of refocused spectra. Phase cycling of the *J*-modulated 1D directed COSY was therefore optimized during both selection of the magnetization and the refocusing period to achieve this. Multiplets obtained using the Bruker simulation program and different phase cyclings are shown in Fig.III.2.1.3. As multiple evolution points are required for the intensity based method, a concern is to keep the phase cycling to a minimum in order to reduce the experimental time.

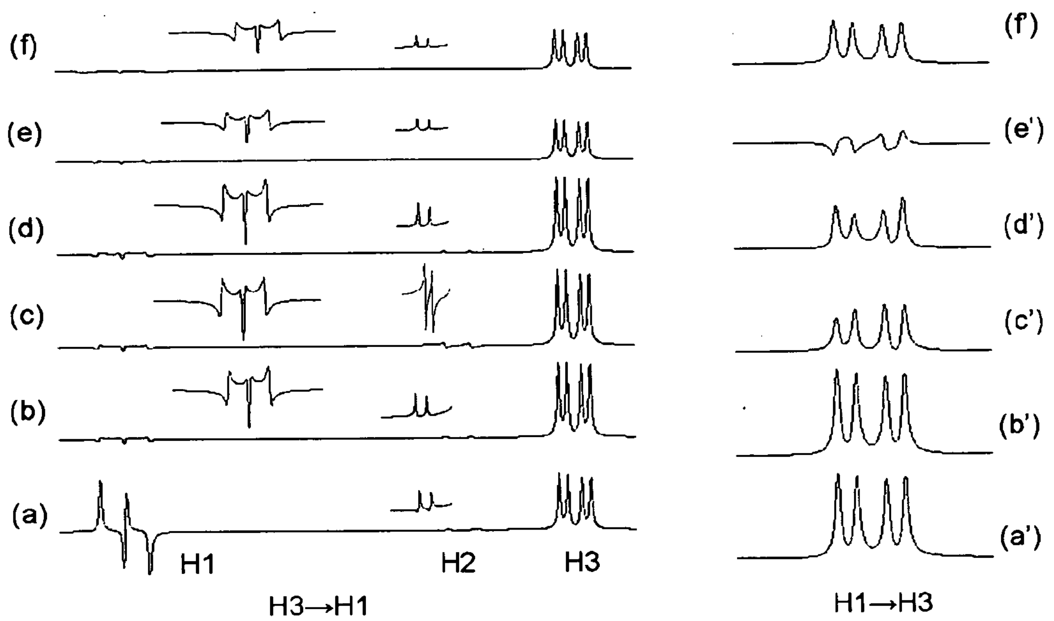


Fig.III.2.1.3 Multiplets obtained from the *J*-modulated COSY experiments using different combinations of phase cycling in the pulse sequences of Fig.III.2.1.1a, simulated using NMRSIM within XWINNMR. (a) to (f) were from the non-refocused pulse sequence to obtain magnetization transfer from H3 to H1 (Fig.III.2.1.1a) using phases φ_1, φ_2, φ_3 and ψ; (a') to (f') were from the refocused pulse sequence to transfer magnetization from H1 to H3 with phases φ_1, φ_2, φ_3, φ_4, φ_5 and ψ.

Phase cycling optimization was first performed on the non-refocused pulse sequence (Fig.III.2.1.3a-f) by focusing on the 180° pulses (φ_1, φ_2 and φ_3). Pure phase multiplets were obtained when all these pulses were phase cycled using full EXORCYCLE [90] independently. The spectrum simulated with such phase cycling (φ_1 = x, y, -x, -y; φ_2 = 4x, 4y, 4(-x), 4(-y); φ_3 = 16x, 16y, 16(-x), 16(-y) and ψ = 4(x, -x), 4(-x, x)) was used as a reference (Fig.III.2.1.3a). In this spectrum the evolution delay was set to yield sizeable transfer of magnetization from H3 to H1. At the same time, the signal of proton H2, which should not appear in the spectrum, is small. In order to see the distortion on proton H1 or appearance of signal of H2, an evolution delay that resulted in a poor transfer of magnetization from H3 to H1 was used. Phase cycling φ_1 was then reduced to (x, y) in spectrum (b), which also gave a H1 multiplet with a good phase. Phase cycling φ_2 was then reduced to (8x, 8y) or (8x, 8(-x)) in spectra (c) or (d). Spectrum (c) gave a large signal of proton H2. When phase cycles φ_2 = (8x, 8(-x)) and ψ = (x, -x) were used in spectrum (d) a good result was obtained. In spectrum (e), phase cycling φ_1 = x, y, -x, -y; φ_2 = 4x, 4(-x); φ_3 = 8x, 8y, 8(-x), 8(-y) and ψ = x, -x were used, which also gave a good result but with half the number of scans compared to previous spectra (a) to (d). The spectrum (f) showing a distortion free mutliplet was simulated using phase cycling of φ_1 = x, y; φ_2 = 2x, 2y, 2(-x), 2(-y); φ_3 = 8x, 8y, 8(-x), 8(-y) and ψ = x, 2(-x), x with the same number of scans as that in the spectrum (e). As phase cycling in both (e) and (f) have the same member of items, the phase cycling used to obtain spectrum (f) was adopted arbitrarily for future experiments.

The next step was to optimize the phases of φ_4 and φ_5, which were used during the refocusing period. The simulated results of the magnetization transfer from protons H1 to H3 are shown in Fig.III.2.1.3a'-f'. Spectra (a') and (b') were acquired using full EXORCYCLE phase cycling on either φ_4 or φ_5 and their mutiplets were in pure phase. Phase cycling φ_4 = 32x, 32y, 32(-x), 32(-y) and φ_5 = x were used for spectrum (a') and φ_4 = x and φ_5 = 32x, 32y, 32(-x), 32(-y) for spectrum (b'). Multiplets in spectra (c'), (d') and (e') were acquired by reducing the phase cycling of either φ_4 or φ_5. These multiplets were phase distorted and their intensities were not reliable for intensity-based coupling constant determination. Spectrum (f') was simulated using phases of φ_4 = x, φ_5 = 32x, 32(-x) and ψ = x, 2(-x), x. The multiplet has a pure phase and the number of scans was reduced to half of that used in the spectrum (a') or (b').

The above simplest combination of the phase cycling gave distortion-free multiplets and was implemented in all *J*-modulated experiments using 30-70 ms selective Gaussian pulses. Spectra acquired using the phase cycle given in Fig.III.2.1.1 lead to increased accuracy of the coupling constant determination.

Next we explore two approaches in order to make the *J*-modulated 1D directed COSY a more versatile method and to overcome the limitations of the initial SPFGSE excitation. The first one relies on multiple magnetization transfers, while the second explores a more selective signal selection than afforded by a SPFGSE.

III. 2.2 *J*-modulated 1D directed HOHAHA-COSY

In the first approach, a double-selective Hartman-Hahn transfer [91] is used to generate the magnetization of the proton *k*. In this experiment, the Hartman-Hahn condition is satisfied only at the resonance frequencies of the two *J*-coupled protons *j* and *k*. This results in a highly efficient magnetization transfer because of the absence of the leakage of the magnetization to other *J*-coupled protons and the lack of cross-relaxation during the double-selective spin-lock [91]. The sensitivity of this method was compared with that of a regular

1D TOCSY experiment (Fig.III.2.2.1). The signals in the double-selective HOHAHA spectra (Fig.III.2.2.2.1d,e) are 1.8 times more intense than those in the TOCSY spectra .(Fig.III.2.2.1b,c).

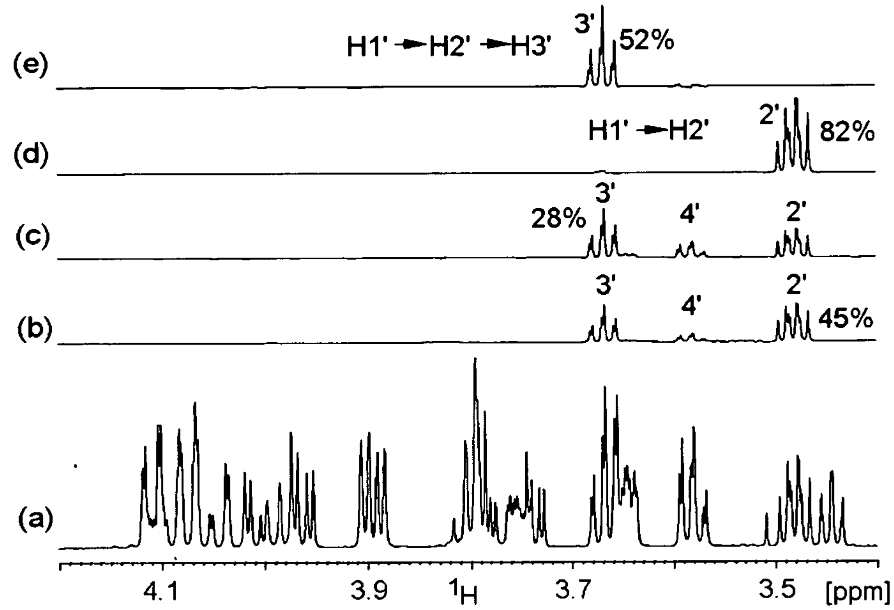


Fig.III.2.2.1 A comparison of 1D TOCSY and 1D double-selective HOHAHA spectra of isotropic cellobiose. All spectra were acquired using identical parameters and are plotted using the same vertical scale. The percentages given in spectra (b)-(e) show the relative signal intensity in comparison with the 1D spectrum (a). In spectra (b)-(e) the magnetization transfer pathway started on proton H1' (outside the spectral range shown) and the purging scheme of Thrippleton and Keeler [92] was applied. 1D TOCSY spectra (b) and (c) were acquired using a DIPSI-2 spin-lock of 47 and 67 ms which yielded maximum magnetization for protons H2' and H3', respectively. Spectra (d) and (e) were acquired using one- and two-step double-selective HOHAHA (pulse sequences of Fig.III.2.2.2a, acquisition point **c** and pulse sequences of Fig.III.2.2.3, acquisition point **d**), respectively. Spin-lock times of 125 and 108 ms were used for the H1' → H2' and H2' → H3' transfer steps, respectively. The strengths of the corresponding B_1 fields were 46 and 38 Hz, respectively.

The selectivity of the double-selective HOHAHA experiment exceeds that of the SPFGSE. Even if the multiplet of proton *j* is not isolated, providing none of the overlapping protons has a coupled partner exactly on resonance with proton *k*, a double-selective Hartman-Hahn transfer is capable of generating selectively the magnetization of proton *k*.

Any antiphase magnetization of proton *k* arising from the mismatch between J_{jk} and the length of the spin-lock period is removed by a purge element proposed by Thrippleton and Keeler [92] (Fig.III.2.2.2a, acquisition point **b**). Finally, the magnetization of proton *k* is selected by a SPFGSE (Fig.III.2.2.2a, acquisition point **c**). When this building block (Fig.III.2.2.2a) is combined with the pulse sequence of Fig.III.2.1.1a, a *J*-modulated 1D directed HOHAHA-COSY experiment is obtained (Fig.III.2.2.2b).

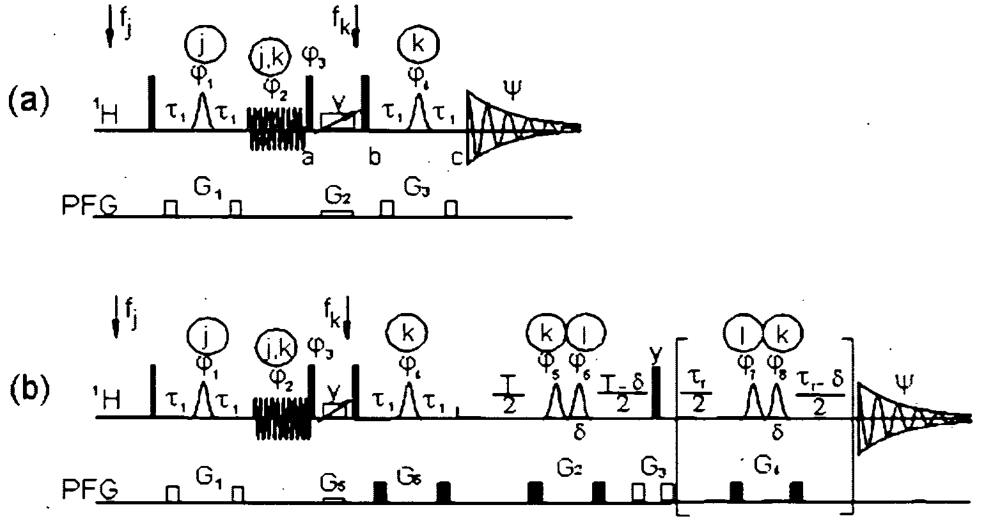


Fig.III.2.2.2 Filled and open rectangles represent 90° and 180° nonselective pulses, respectively, applied from the *x* axis unless specified otherwise. Selective 180° Gaussian pulses are indicated by open Gaussian envelopes and applied to spins *j*, *k* or *l*, as indicated. The refocusing interval enclosed in square brackets is optional; τ_1 = 1.2 ms and *T* and τ_r are the variable and refocusing delays, respectively. Gradient sine shaped pulses were applied along the *z* (open squares) or *x*, *y* directions (filled squares). All pulse field gradients had duration δ = 1 ms unless specified otherwise. (a) 1D double-selective HOHAHA. Gradient strengths were G_1 = 13 G/cm, G_2 = 4 G/cm and G_3 = 17 G/cm. The length of G_2 gradient was 49.8 ms. The following phase cycling was applied: $\varphi_1 = x, y$; $\varphi_2 = y + \alpha$; $\varphi_3 = x + \beta$; $\varphi_4 = 2x, 2y$; When the acquisition is started at points **a** and **b**, the receiver is phase cycled as $\psi = x, -x$; while for the acquisition starting at point **c** it is $\psi = x, 2(-x), x$. Phases φ_2 and φ_3 are optimized (α and β) in increments of 10° for the maximum signal in two experiments with acquisition starting at points **a** and **b**, respectively. The ^{1}H frequency is changed at times specified by vertical arrows. The selective HOHAHA spin-lock is generated by a frequency-modulated rectangular pulse at resonances of *j* and *k* protons. The RF field strength is calculated as $(J_{j,k}/4)\,(4n^2-1)^{1/2}$, where $J_{j,k}$ is the proton-proton coupling constant; n=10 is typically used. The 180° adiabatic pulse is indicated as a rectangle with an inclined arrow. (b) 1D *J*-modulated 1D directed HOHAHA-COSY; The following phase cycling was applied: $\varphi_1 = x, y$; $\varphi_2 = y$; $\varphi_3 = x$; $\varphi_4 = 2x, 2y$; $\varphi_5 = 4x, 4y, 4(-x), 4(-y)$; $\varphi_6 = 16x, 16y, 16(-x), 16(-y)$; $\varphi_7 = x$; $\varphi_8 = 64x, 64(-x)$; $\psi_1 = x, 2(-x), x, (-x), 2x, (-x)$. Gradient strengths were G_1 = 13 G/cm, G_2 = 15 G/cm, G_3 = 8 G/cm, G_4 = 11 G/cm, G_5 = 4 G/cm and G_6 = 17 G/cm. When the stability of a spectrometer is not an issue, pulse sequences (b) can be modified to use phase cycling rather than the gradient selection for the COSY transfer step. This entails setting G_3 = 0 and imposing a two-step phase cycle (y,-y) on the 90° polarization transfer pulse. This results in a two-fold increase of signal intensities.

The applicability of this approach can be further broadened by combining the COSY step with multiple selective HOHAHA transfers utilizing extended polarization transfer pathways (Fig.III.2.2.3).

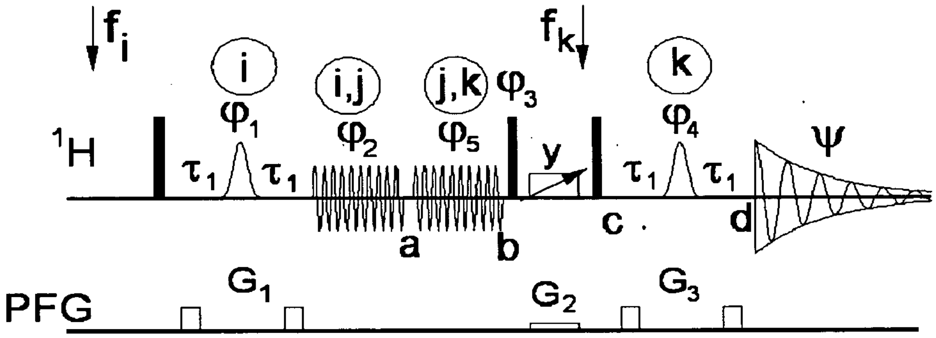


Fig.III.2.2.3 1D two-step double-selective HOHAHA experiment. Filled rectangles represent 90° nonselective pulses applied from the x axis unless specified otherwise. Selective 180° Gaussian pulses are indicated by open Gaussian envelopes and applied to spins I and k, as indicated. Sine shaped gradient pulses were applied along the z axis. All pulse field gradients had duration $\delta = 1$ ms ($\tau_1 = 1.1$ ms) unless specified otherwise. Gradient strengths were $G_1 = 13$ G/cm, $G_2 = 4$ G/cm and $G_3 = 17$ G/cm. The length of G_2 gradient was 49.8 ms. The following phase cycling was applied: $\varphi_1 = x, y$; $\varphi_2 = y + \alpha$; $\varphi_3 = x + \beta$; $\varphi_4 = 2x, 2y$; $\varphi_5 = y + \gamma$; When the acquisition is started at points **a**, **b**, or **c** the receiver is phase cycled as $\psi = x, -x$; while for the acquisition starting at point **d** it is $\psi = x, 2(-x), x$. Phases φ_2 , φ_5 and φ_3 are independently optimized (i.e. angles α, γ and β determined) in increments of 10° for the maximum signal in three separate experiments with the acquisition starting at points **a**, **b** and **c**, respectively. The ^{1}H frequency is changed at positions specified by vertical arrows. The selective HOHAHA spin lock is generated by a frequency modulated rectangular pulse at frequencies of i,j and j,k protons, respectively. The RF field strength is calculated as $(J_{HH}/4)\ (4n^2 - 1)^{-1/2}$, where J_{HH} is the proton-proton coupling constant. The 180° adiabatic pulse is indicated as a rectangle with an inclined arrow.

A two-step double-selective HOHAHA transfer (Fig.III.2.2.4e) via coupling constants > 7 Hz is still 1.8 times more sensitive than a regular 1D TOCSY experiment (Fig.III.2.2.4c). The use of two double-selective steps makes highly selective experiments possible.

Application of the double-selective HOHAHA experiment to the measurement of ^{1}H-^{1}H RDCs of an aligned sample of cellobiose is shown in Fig.III.2.2.4. Selective excitation of H4$_\alpha$, which is buried in a crowded region, is required to initiate the transfer of the magnetization to H2$_\alpha$ for the measurement of the mutual splitting. This is achieved by a double-selective HOHAHA transfer from H3$_\alpha$ to H4$_\alpha$ (Fig.III.2.2.4c-e). Once the pure multiplet of H4$_\alpha$ is obtained, a variable time COSY step is appended (Fig.III.2.2.4f) and magnetization is transferred to H2$_\alpha$.

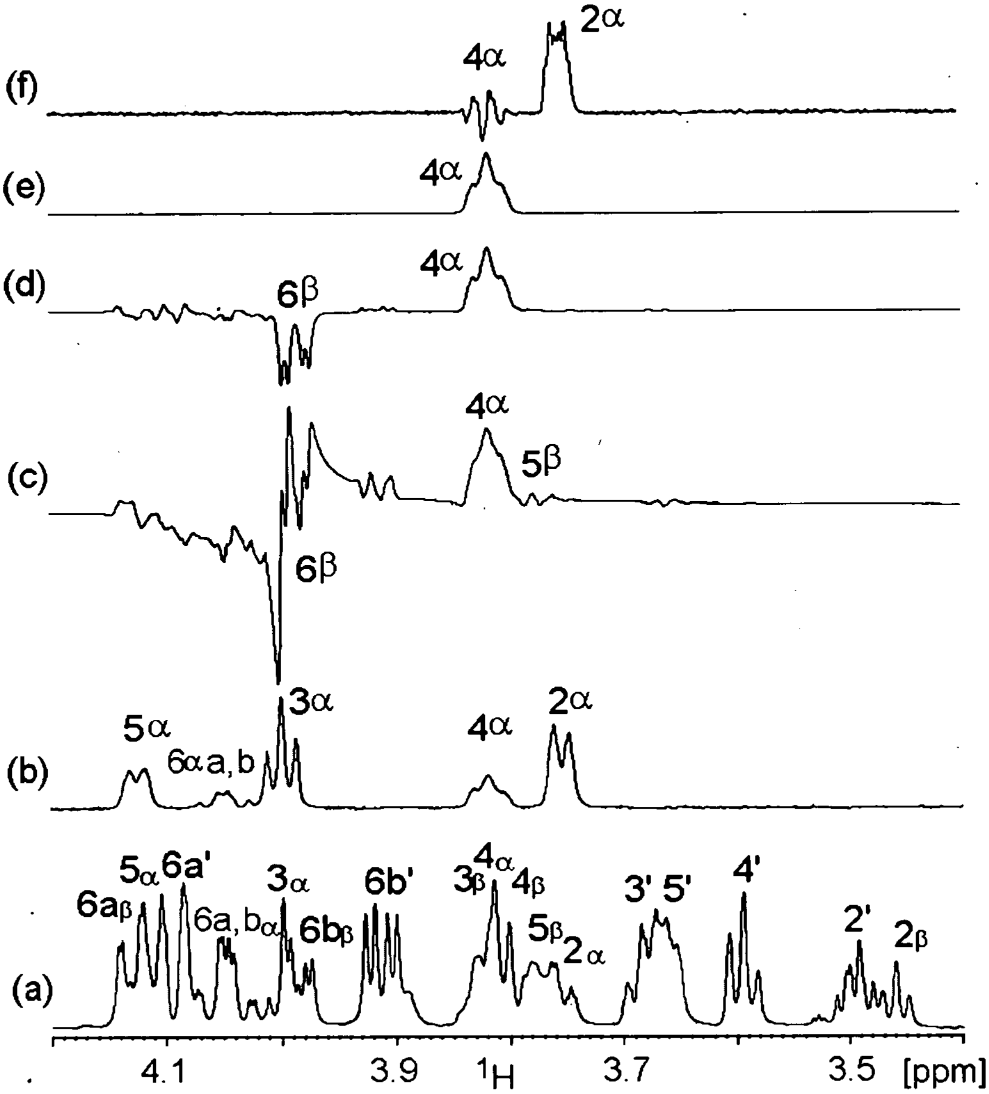


Fig.III.2.2.4 (a) Partial 1D CPMG spectrum of aligned cellobiose. (b) Partial 1D TOCSY spectrum with selective excitation of the anomeric proton H1$_\alpha$. (c) 1D double-selective HOHAHA spectrum with a selective transfer of magnetization from H3$_\alpha$ to H4$_\alpha$ acquired using the pulse sequence of Fig.III.2.2.2a and the acquisition started at point **a**. Note a very small amount of H5$_\beta$ that was generated in spite of close proximity of resonances H3$_\alpha$/H6$_\beta$ and also H4$_\alpha$/H5$_\beta$. (d) The same as (c), but the acquisition started after purging of the antiphase magnetization, at point **b**. (e) The same as (d), but the acquisition started after the selection of H4$_\alpha$ by a SPFGSE, at point **c**. Spectra (c)-(e) are plotted using the same vertical scale. (f) 1D *J*-modulated directed HOHAHA-COSY acquired with the pulse sequence of Fig.III.2.2.2b and the polarization transfer pathway H3$_\alpha$ → H4$_\alpha$ → H2$_\alpha$.

The intensities of $H2_\alpha$ signals as a function of the COSY delay are fitted (Fig.III.2.2.5) yielding an accurate $D_{4\alpha 2\alpha} + J_{4\alpha 2\alpha}$ splitting.

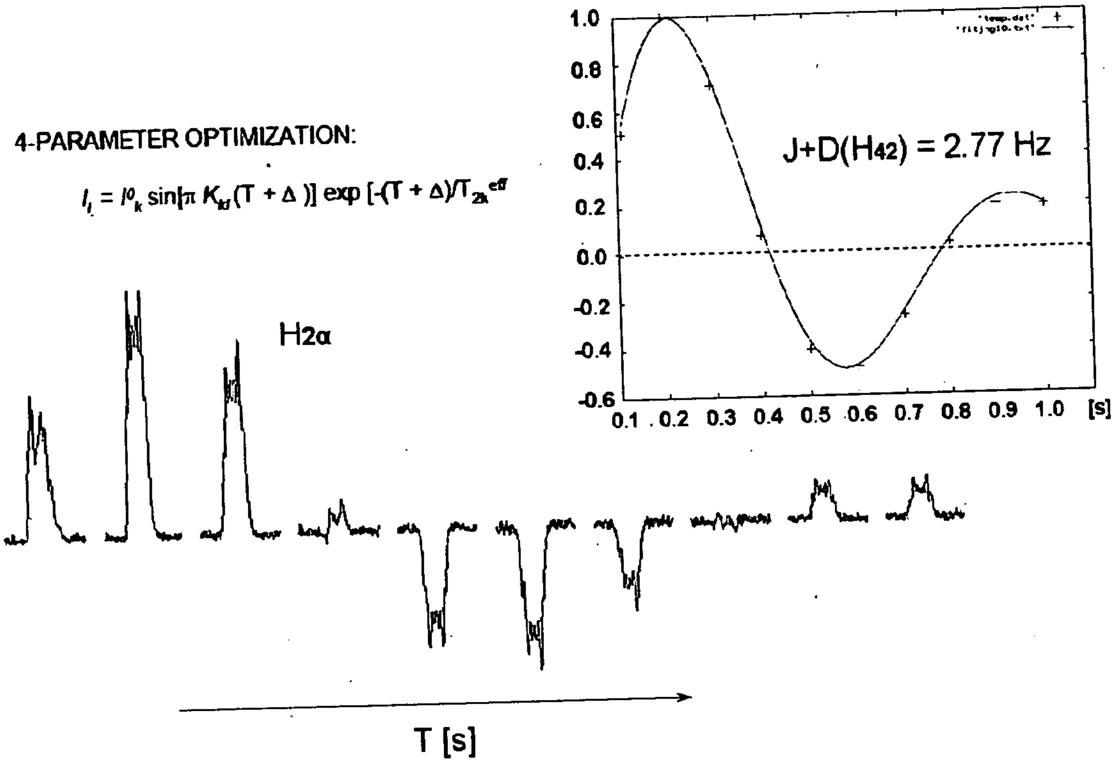


Fig.III.2.2.5 Coupling constant determination, $|J+D|_{H4\alpha H2\alpha}$ of cellobiose using the variable-time *J*-modulated 1D HOHAHA-COSY experiment. Signal intensities of proton $H2_\alpha$ were fitted using the transfer function given in the inset.

III. 2.3 *J*-modulated 1D directed CSSF-COSY

The second approach relies on a direct selection of a proton of interest without the need for an initial transfer of magnetization. The overlapping ^1H multiplets can be selectively excited by gradient-enhanced chemical shift selective filters [73] (CSSFs, Fig.III.2.3.1a). Chemical shift differences as small as 1-2 Hz are sufficient to achieve this. Replacing the first SPFGSE of the *J*-modulated 1D directed COSY by a ge-CSSF yields a pulse sequence of a *J*-modulated 1D directed CSSF-COSY (Fig.III.2.3.1b).

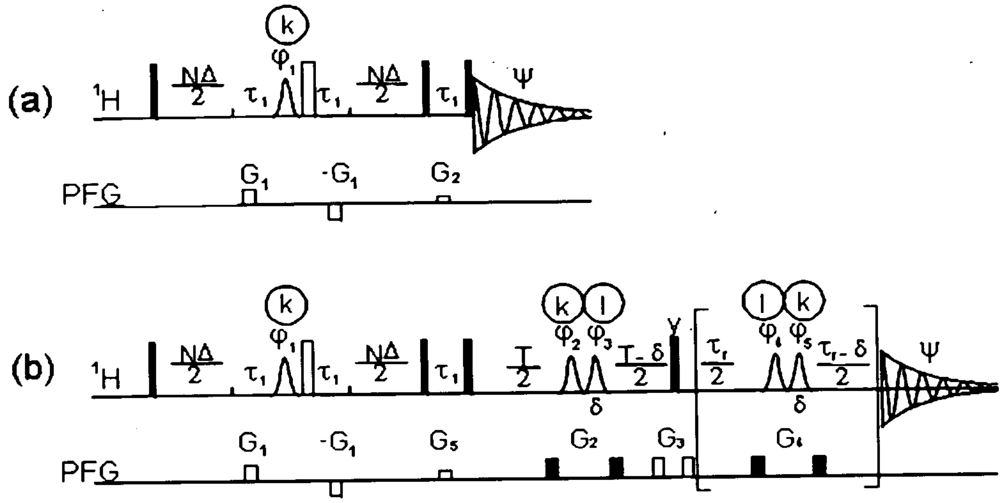


Fig.III.2.3.1 Filled and open rectangles represent 90° and 180° nonselective pulses, respectively, applied from the *x* axis unless specified otherwise. Selective 180° Gaussian pulses are indicated by open Gaussian envelopes and applied to spins *j*, *k* or *l*, as indicated. The refocusing interval enclosed in square brackets is optional; τ_1 = 1.2 ms and *T* and τ_r are the variable and refocusing delays, respectively. Gradient sine shaped pulses were applied along the *z* (open squares) or *x*, *y* directions (filled squares). All pulse field gradients had duration δ = 1 ms unless specified otherwise. (a) 1D ge-VT-CSSF; Δ is the increment of the CSSF, N=0, 1, 2 ...n, and t_{max}=(N+1)Δ. The following phase cycling was applied: $\varphi_1 = x, y$; $\psi_1 = x, -x$; Gradient strengths were G_1= 20 G/cm and G_2=13 G/cm (b) 1D *J*-modulated 1D directed CSSF-COSY; The following phase cycling was applied: $\varphi_1 = x, y$; $\varphi_2 = 2x, 2y, 2(-x), 2(-y)$; $\varphi_3 = 8x, 8y, 8(-x), 8(-y)$; $\varphi_4 = x$; $\varphi_5 = 32x, 32(-x)$ $\psi_1 = x, 2(-x), x$. Gradient strengths were G_1 = 13 G/cm, G_2 = 15 G/cm, G_3 = 8 G/cm, G_4 = 11 G/cm and G_5 = 4 G/cm. When the stability of a spectrometer is not an issue, pulse sequences (b) can be modified to use phase cycling rather than the gradient selection for the COSY transfer step. This entails setting G_3 = 0 and imposing a two-step phase cycle (y,-y) on the 90° polarization transfer pulse. This results in a two-fold increase of signal intensities.

Fig.III.2.3.2 shows the application of this experiment to the measurement of the $D_{3\alpha4\alpha} + J_{3\alpha4\alpha}$ splitting of the aligned cellobiose. A CSSF filter, utilizing a 10 Hz separation between resonance frequencies of overlapping protons H3$_\alpha$ and H6$_\beta$, is used to select the signal of H3$_\alpha$. The subsequent *J*-modulated 1D directed COSY step transfers the magnetization to proton H4$_\alpha$. A series of experiments, with a variable COSY delays, yield the value of the $D_{3\alpha4\alpha} + J_{3\alpha4\alpha}$ splitting.

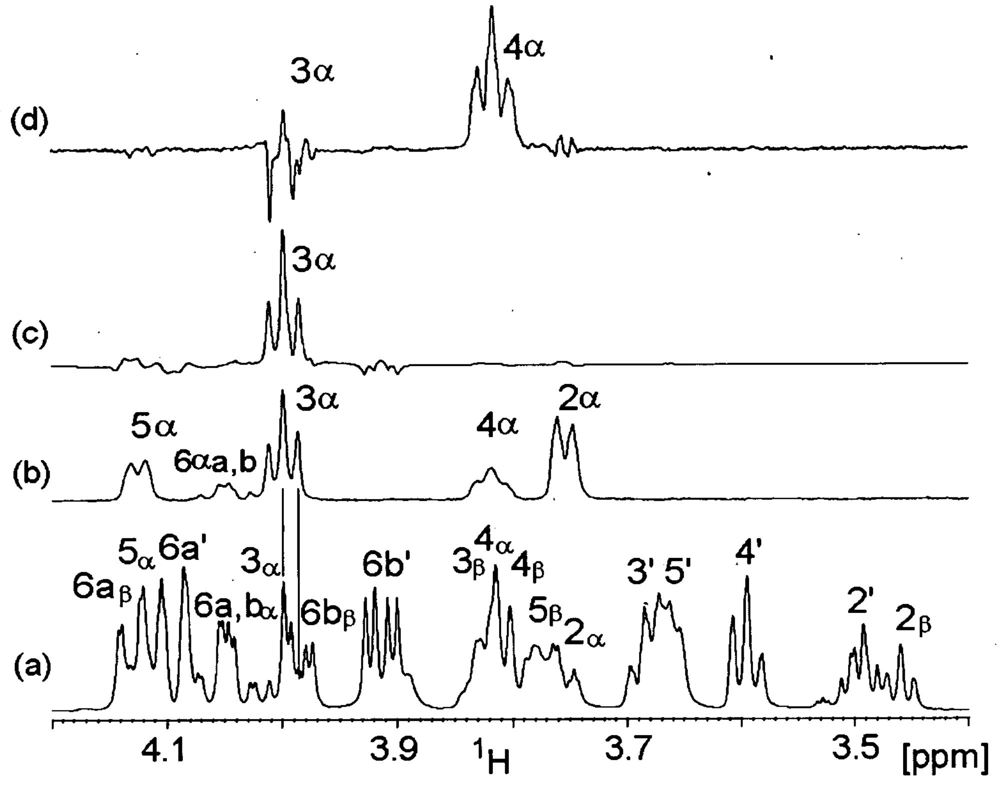


Fig.III.2.3.2 (a) Partial 1D CPMG spectrum of aligned cellobiose. Lines indicate positions of chemical shifts of protons $H3_\alpha$ and $H6_\beta$; (b) Partial 1D TOCSY spectrum with selective excitation of the anomeric proton $H1_\alpha$. (c) 1D ge-VT-CSSF spectrum with a selection of proton $H3_\alpha$ acquired with the pulse sequence of Fig.III.2.3.1a. (d) 1D *J*-modulated 1D directed CSSF-COSY spectrum acquired with the pulse sequence of Fig.III.2.3.1b and a selective transfer of magnetization from $H3_\alpha$ to $H4_\alpha$.

Both variable-time and constant-time COSY evolution intervals were tested in combination with the double-selective HOHAHA and CSSF experiments. This is because the constant-time experiments do not require fitting of the relaxation time. For small coupling constants, where the build-up of the magnetization competes with the relaxation in variable time experiments, the fit is not reliable. On the negative side, the constant-time experiments are less sensitive.

Double-selective HOHAHA and CSSF based methods are especially useful for accurate and precise determination of the ${}^{n}D_{HH}$ coupling constants of medium size molecules that show complex, overlapping ${}^{1}H$ NMR spectra. This greatly extends the applicability of the basic double-selective COSY technique which relies on the presence of resolved ${}^{1}H$ multiplets. The choice between the CSSF and double-selective HOHAHA based methods depends on the severity of spectral overlap and sizes of coupling constants involved. Determination of an extensive set of proton-proton residual dipolar coupling constants is essential for the calculation of the alignment tensor of many molecules. $^{1}D_{CH}$ that are most easily accessible are often degenerate due to the limited variability of the orientations of CH vectors in small molecules.

III. 2.4 Determination of the signs of coupling constants

Only absolute values of J and $(J + D)$ splittings are obtained from the J-modulated COSY experiments. Their signs can be determined either experimentally (e.g. E.COSY [93]) or iteratively during the calculation of the alignment tensor. Experimental determination is preferred, but can be difficult for very small splittings. The iterative procedure outlined below relies on the tensor being reasonably well defined by a basic set of RDCs. This definition is firmed up by including additional splittings for which the signs are not known. This concerns the four- and five-bond ^{1}H-^{1}H RDCs. The signs of two-bond and three-bond J coupling constants are negative and positive, respectively, and therefore the determination of the signs of corresponding RDCs is trivial.

Each pair of $^{4}|J|$ and $^{4}|J+D|$ satisfies four possible combinations of signs of J and D, consequently four different $^{4}D_{HH}$ coupling constants can be obtained. Here an example of the calculation of the alignment tensor of α-D-glucopyranose in aligned sample of cellobiose is given. The alignment tensor is initially calculated using unambiguous $^{1}D_{CH}$ and $^{3}D_{HH}$ coupling constants and the neutron-diffraction structure [94] of α-D-glucopyranose. $^{4}D_{H2H4}$ is considered first. As the $^{4}J_{H2H4}$ is almost zero, possible values of $^{4}D_{H2H4}$ are ± 2.77 Hz. The negative $^{4}D_{H2H4}$ gives smaller RMSDs between the experimental and back-calculated RDCs and is therefore taken forward for further analysis. Similarly $^{4}J_{H3H5}$ is also close to zero meaning that only two rather than four values of $^{4}D_{H3H5}$ must be tested. Therefore 32 combinations of $^{4}D_{H1H3}$, $^{4}D_{H1H5}$ and $^{4}D_{H3H5}$ are considered next (Fig.III.2.4.1). For each combination the alignment tensor is optimized and the RDCs are back-calculated. The sum of the square of the differences (SSRs) between the experimental and calculated RDCs are shown in Fig.III.3.2.10. The lowest SSR (0.19 Hz^{2}) obtained for the combination of $^{4}J_{24} = 0$ Hz, $^{4}J_{13} = -0.52$ Hz, $^{4}J_{15} = -0.61$ Hz and $^{4}J_{35} = 0$ Hz and $^{4}D_{24} = -2.77$ Hz, $^{4}D_{13} = -0.19$ Hz, $^{4}D_{15} = -0.6$ Hz and $^{4}D_{35} = 0.7$ Hz represents the correct solution.

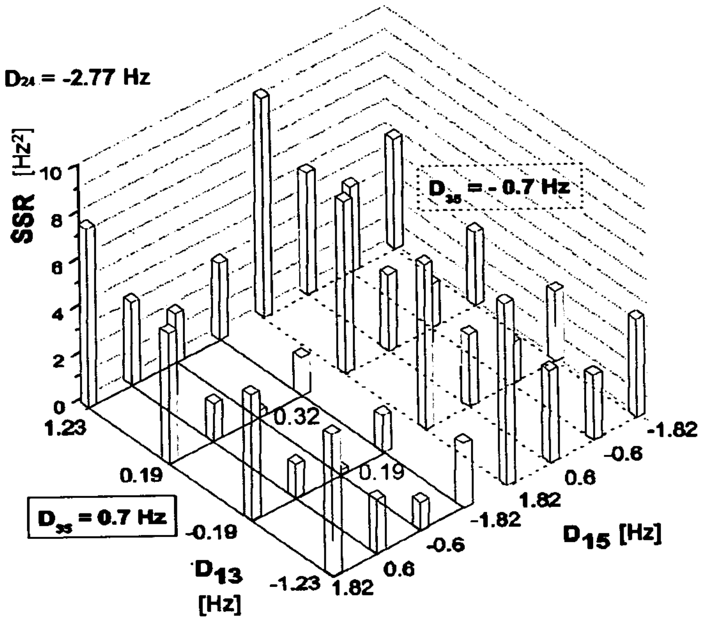


Fig.III.2.4.1 The parameters of the alignment tensor were first determined using four ${}^{1}D_{CH}$ and four ${}^{3}D_{\mathrm{HH}}$ RDCs, and the largest ${}^{4}D_{\mathrm{HH}}$ (${}^{4}D_{24}$ = ± 2.77 Hz). The value -2.77 yielded lower SSR (sum of the square of the differences) between the experimental and back-calculated RDCs. Only this value was taken forward for further analysis. Possible values for the remaining three ${}^{4}D_{\mathrm{HH}}$ RDCs (${}^{4}D_{13}$, ${}^{4}D_{15}$ and ${}^{4}D_{35}$) were calculated as $\pm |J + D| \pm |J|$ and SSR were obtained using all 32 permutations.

Values of ^{1}H-^{1}H J and D coupling constants given in Table III.2.4.1 were obtained as averages of four measurements (RMSDs < 0.04 Hz) using forward and backward direction and both HOHAHA and CSSF based methods. The heteronuclear one-bond ^{1}H-^{13}C splittings were determined from non-refocused 2D HSQC spectra acquired without decoupling in the directly detected dimension.

Table III.2.4.1 Scalar and dipolar coupling constants of the reducing α-D-glucopyranose of cellobiose.

	H_1C_1[a]	H_2C_2[a]	H_3C_3[a]	H_4C_4[a,b]	H_1H_2	H_2H_3	H_3H_4	H_4H_5	H_2H_4	H_1H_3	H_1H_5	H_3H_5
J	169.9	144.5	146.5	145.6	3.83	9.8	8.9	10.0	0[c]	-0.5	-0.6	0[c]
J+D	165.4	151.1	152.9	151.5	3.52	10.7	9.6	11.1	-2.8	-0.7	-1.2	0.7
D_{exp}	-4.5	6.6	6.4	5.9	-0.31	0.9	0.7	1.1	-2.8	-0.2	-0.6	0.7
D_{cal}[d]	-4.44	6.54	6.27	6.19	-0.26	0.69	0.69	0.91	-2.77	-0.21	-0.56	0.91

[a] Determined from F_2 proton-coupled ^{1}H-^{13}C HSQC spectra. [b] H5C5 was excluded because of higher order effects in the ^{13}C satellite spectra. [c] No transfer was observed. [d] Calculated using REDCAT [66]. Parameters of the alignment tensor are: $S_{x'x'}$ = (-7.35 ± 0.45)e-05, $S_{y'y'}$ = (-14.68 ± 0.18)e-05, $S_{z'z'}$ = (22.03 ± 0.50) e-05. The Euler angles in the coordinate system of neutron diffraction structure of cellobiose: α = - 27.0 ± 1.1, β = 74.6 ± 2.3 and γ = 92.5 ± 6.4 degrees.

Five independent RDCs are required as a minimum in order to define the alignment tensor. In a pyranose chair five C-H bonds provide four different vectors at the most, therefore ${}^{n}D_{HH}$ are essential for the accurate determination of the alignment tensor. In carbohydrates ${}^{1}D_{CH}$ and ${}^{3}D_{HH}$ are typically used to define the tensor while inclusion of ${}^{4}D_{HH}$ improves the accuracy of the definition (Fig.III.2.4.2). In α-D-glucopyranose, only two CH vectors point in independent directions (equatorial and axial). Therefore the alignment tensor calculated relying mostly on the heteronuclear coupling constants is expected to show large ambiguity. When only five dipolar coupling constants (four ${}^{1}D_{CH}$ and ${}^{3}D_{H1H2}$) were used in the calculation of the tensor parameters, the largest component, S_{zz}, was defined but the other two components S_{xx} and S_{yy} were not (Fig.III.2.4.2a). In this calculation, ${}^{3}D_{H1H2}$ coupling constant was used. This coupling constant could be measured using the basic *J*-modulated 1D directed COSY. Next, the coupling constants measured by the newly developed methods were added to the calculations in two steps. Initially, the other three three-bond coupling constants were added, followed by four four-bond coupling constants. As can be seen in Fig.III.2.4.2 the ambiguity in the definition of the tensor is gradually reduced when all ${}^{3}D_{HH}$ and then also ${}^{4}D_{HH}$ coupling constants are included. The S_{xx} and S_{yy} components are now well defined.

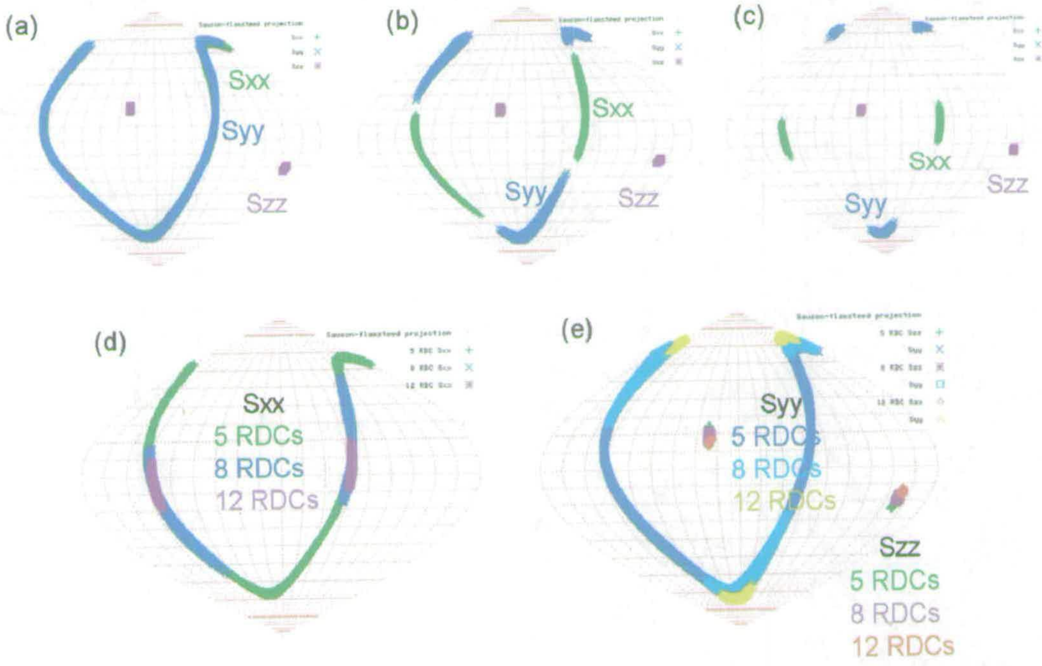


Fig.III.2.4.2 Sanson-Flamsteed projections showing orientations of the alignment tensor principal axis system in the molecular frame of the neutron-diffraction structure of cellobiose calculated using REDCAT [66] (a) results based on five coupling constants (four ${}^{1}D_{CH}$ and ${}^{3}D_{H1H2}$); (b) results based on four ${}^{1}D_{CH}$ and four ${}^{3}D_{HH}$ coupling constants; (c) results based on all twelve RDCs given in Table III.2.4.1, (d) is an overlay of the S_{xx} vectors of the alignment system for the three cases; (e) is an overlay of the S_{yy} and S_{zz} vectors of the alignment system for the three cases.

III. 2.5 Configuration of sodium cholate from ^{1}H-^{1}H RDCs

In the next section, the application of the proposed methods for the measurement of proton-proton RDCs in the determination of the relative configuration of small molecules is illustrated. The problem arises when in six-member rings the number of independent orientations of CH vectors is limited. Considering a chair conformation, only 4 CH vectors are independent and this number can be reduced further when rings are fused, eliminating some CH vectors due to substitutions. On the other hand, the numbers of independent orientations considering $^{2}D_{HH}$, $^{3}D_{HH}$, $^{4}D_{HH}$ and $^{5}D_{HH}$ are 3, 12, 10 and 9, respectively. Determining the alignment tensor using only the $^{1}D_{CH}$ coupling constants is rarely possible as more than five independent CH vectors are not available. Approximations need to be introduced, e.g. assuming the alignment tensors to be axially symmetric reduces the minimum number of required RDCs from 5 to 3. This approximation has been used in confirming the absolute configuration of three stereocenters at C3, C7 and C12 in sodium cholate using only $^{1}D_{CH}$ coupling constants [95]. This, however, is not a general approach. Supplying several proton-proton RDCs can provide an answer without using any approximations.

Fig.III.2.5.1 shows a comparison of 1D spectra of isotropic and aligned samples of sodium cholate together with 1D TOCSY spectra of the aligned sample. These spectra illustrate that despite considerable overlap in the 1D spectrum, protons of individual rings of sodium cholate (labeled as in Fig.II.2.1d) are resolved sufficiently for the measurement of numerous proton-proton RDCs.

Fig.III.2.5.2 shows a particularly crowded region of the spectra where six protons resonate between 1.5 and 1.7 ppm. Despite this overlap, proton H8 can be selected by CSSF (Fig.III.2.5.2d) and the splitting H8H9 measured in a 1D J-modulated 1D directed CSSF-COSY experiment (Fig.III.2.5.2e).

Using a combination of CSSF and double-selective HOHAHA methods, four two-bond and ten three-bond proton-proton RDCs were measured together with fourteen $^{1}D_{CH}$ coupling constants (Table III.2.5.1).

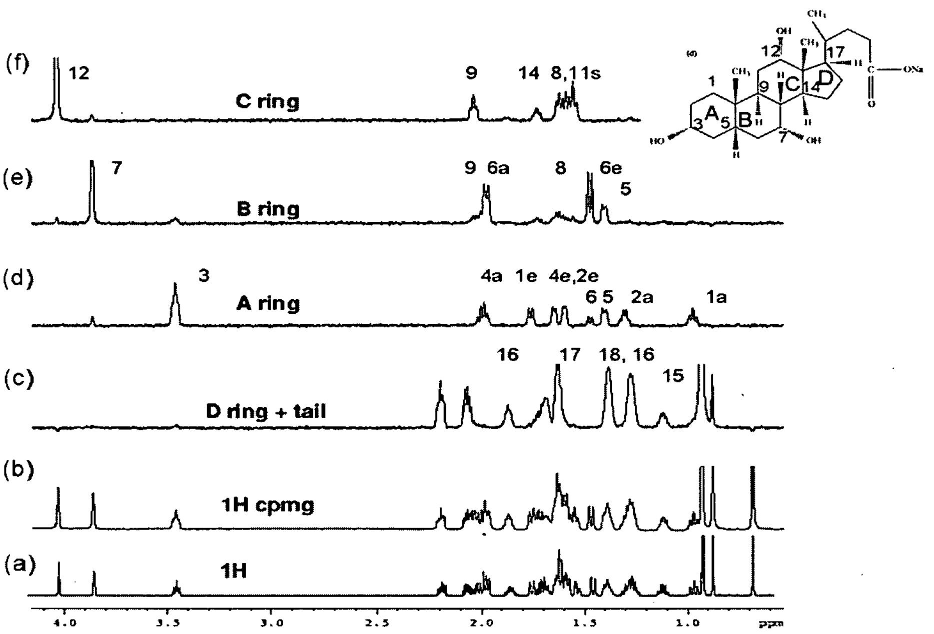


Fig.III.2.5.1 (a) and (b) ^{1}H spectra of isotropic and aligned samples of sodium cholate; (c-f) 1D TOCSY spectra of different spin systems from the aligned sample of sodium cholate. A 50ms CPMG pulse is used to reduce the signals from the alignment medium.

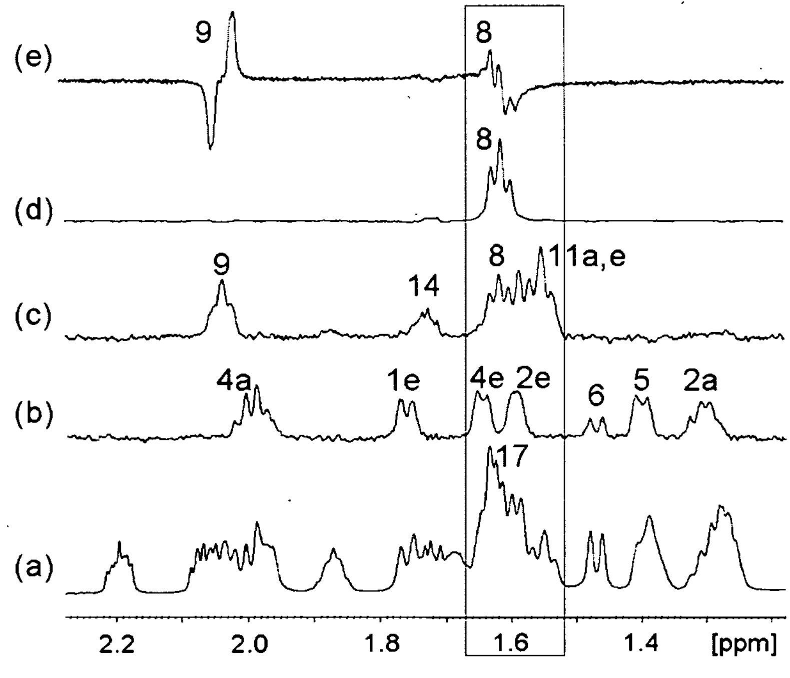


Fig.III.2.5.2 (a) Partial 1D CPMG spectrum of aligned sodium cholate. The crowded region of overlapped protons is enclosed in the box. (b) Partial 1D TOCSY spectrum of A ring with selective excitation from H3. (c) Partial 1D TOCSY spectrum of C ring with selective excitation from H7. (d) 1D ge-VT-CSSF spectrum with a selection of proton H8 acquired with the pulse sequence of Fig.III.2.3.1a.

Table III.2.5.1 Scalar and dipolar coupling constants of sodium cholate. RDCs in bold italic related to the atoms of these stereocenters were not used in the determination of the alignment tensor

	J (Hz)	$J + D$ (Hz)	RDC$_{\text{Expt}}$ (Hz)	RDC$_{\text{cal}}$ (Hz)
C_1H_{1ax}	123.7	126.0	2.3 ^{a}	2.56
C_1H_{1eq}	128.6	126.0	-2.6 ^{b}	-2.52
C_2H_{2ax}	125.3	128.0	2.7 ^{a}	2.49
C_2H_{2eq}	129.4	128.2	-1.2 ^{c}	-0.95
C_3H_3	***141.7***	***144.4***	***2.7*** ^{a}	***2.56***
C_4H_{4eq}	126.0	123.6	-2.4 ^{b}	-2.55
C_4H_{4ax}	126.2	128.9	2.7 ^{a}	2.54
C_5H_5	126.1	128.3	2.2 ^{a}	2.57
C_6H_{6eq}	127.3	124.8	-2.5 ^{b}	-2.44
C_6H_{6ax}	122.2	124.2	2.0 ^{a}	1.88
C_7H_7	***144.9***	***143.8***	***-1.1*** ^{c}	***-0.99***
C_9H_9	122.4	122.8	1.5 ^{d}	1.66
$C_{12}H_{12}$	145.3	142.8	-2.5 ^{b}	-2.53
$H_{1a}H_{1e}$	-14.41	-14.65	-0.24 ^{e}	-0.16
$H_{2a}H_{2e}$	-12.21	-10.92	1.29 ^{f}	0.98
$H_{1e}H_{2e}$	3.72	3.86	0.14 ^{g}	-0.14
$H_{2a}H_{1e}$	3.24	3.82	0.58 ^{h}	0.39
$H_{2e}H_3$	***4.40***	***3.66***	***-0.74*** ^{i}	***-0.59***
H_3H_{4e}	***4.50***	***4.50***	***0.00*** ^{j}	***-0.21***
H_3H_{4a}	***11.45***	***11.75***	***0.30*** ^{k}	***0.22***
$H_{4e}H_5$	3.87	4.30	0.43 ^{h}	0.39
$H_{4a}H_{4e}$	-12.50	-12.66	-0.16 ^{l}	-0.20
$H_{4a}H_5$	13.37	14.08	0.71 ^{m}	0.53
$H_{6a}H_{6e}$	-15.00	-15.31	-0.31 ^{n}	-0.19
$H_{6a}H_7$	***3.59***	***4.49***	***0.90*** ^{f}	***0.77***
H_7H_8	***3.14***	***2.77***	***-0.37*** ^{o}	***-0.53***
H_8H_9	11.85	12.07	0.22 ^{g}	0.10

Parameters of the alignment tensor are: $S_{x'x'} = (8.19 \pm 3.17)\text{e-}06$, $S_{y'y'} = (6.04 \pm 0.80)\text{e-}05$, $S_{z'z'} = (-6.86 \pm 0.85)\text{ e-}05$. The Euler angles in the coordinate system of neutron diffraction structure of sodium cholate: $\alpha = 12.6 \pm 5.6$, $\beta = 33.5 \pm 3.7$ and $\gamma = 6.2 \pm 4.0$ degrees; a to o in superscripts on RDC$_{\text{Expt}}$ are groups of same orientation for CH and HH vectors

The alignment tensor was initially calculated using unambiguous $^1D_{\text{CH}}$ and $^3D_{\text{HH}}$ (excluding those containing H3, H7 and H12) coupling constants and the neutron-diffraction structure [96] of sodium cholate. Because of the lack of proton-proton RDCs involving proton H12, only $^1D_{\text{CH}}$ is considered for this stereocenter. Table III.2.5.1 contains RDCs calculated based on the correct stereochemistry, which is reflected in a good agreement between the measured and calculated values. Back-calculated RDCs considering equatorial and axial positions of protons H3, H7 and H12 are shown in Table III.2.5.2. These results indicate that protons H3, H7 and H12 are indeed equatorial.

Table III.2.5.2 Back calculated dipolar coupling constants considering different orientations of protons of H3, H7 and H12 in sodium cholate. Bold values indicate a match between the experimental and theoretical RDCs.

	Experiment	H_{7a}	**H_{7e}**
C_3H_{3e}	2.7	**2.56**	**2.56**
C_7H_7	-1.1	2.11	**-0.99**
$C_{12}H_{12e}$ [$C_{12}H_{12a}^*$]	-2.5	**-2.53** [2.0]	**-2.53** [2.0]
$H_{3e}H_{2e}$	-0.7	**-0.59**	**-0.59**
$H_{3e}H_{4a}$	0.3	**0.22**	**0.22**
$H_{3e}H_{4e}$	0.0	**-0.40**	**-0.40**
H_7H_{6a}	0.9	0.55	**0.77**
H_7H_8	-0.4	0.02	**-0.53**
C_3H_{3a}	2.7	1.41	1.41
C_7H_7	-1.1	2.11	**-0.99**
$C_{12}H_{12e}$ [$C_{12}H_{12a}^*$]	-2.5	**-2.53** [2.0]	**-2.53** [2.0]
$H_{3a}H_{2e}$	-0.7	**-0.56**	**-0.56**
$H_{3a}H_{4a}$	0.3	-0.17	-0.17
$H_{3a}H_{4e}$	0.0	**-0.38**	**-0.38**
H_7H_{6a}	0.9	0.55	**0.77**
H_7H_8	-0.4	0.02	**-0.53**

* Proton H12 at the axial position does not fit the experimental data

In summary, the results presented in this chapter [97] demonstrate that we are now able to measure accurately ^1H-^1H RDCs in spectra with significant overlap of resonances. The techniques outlined here were used to measure RDCs of the aligned sample of fully sulfated tetrasaccharide **II**. These were used in the conformational analysis of this molecule presented in Chapter III.4.

III. 3 Methods for the measurement of J_{CC} coupling constants

The ultimate NMR experiments for establishing the structures of small organic molecules are based on the tracing of their carbon-carbon connectivity. This can be achieved by using INADEQUATE [98] (Incredible Natural-Abundance DoublE-QUAntum Transfer Experiment) experiments. At the heart of this experiment is a double-quantum filter that extracts ^{13}C-^{13}C satellites either by phase cycling or pulsed field gradients, while suppressing single-quantum carbon coherences. Besides providing the connectivity information, the INADEQUATE experiment opens a route towards the measurement of carbon-carbon coupling constants. In principle, both $^{1}J_{CC}$ and $^{n}J_{CC}$ coupling constants can be measured, although the latter are not easily obtained.

The glycosidic linkage of the disaccharide unit is described by two dihedral angles Φ and Ψ (Fig.III.3.1a). Both angles can be defined in two ways either as COCH or COCC angles across the linkage. Two proton-carbon coupling constants ($^{3}J_{HC1OCx}$ and $^{3}J_{C1OCxHx}$) across the glycosidic linkage report the Φ and Ψ angles, respectively. These coupling constants are easily obtained via HSQC or HMBC based experiments. However, as can be seen by way of example in Fig.III.3.1b four solutions are possible for a $^{3}J_{HCOC}$ coupling constant of 3.8 Hz. Addition of a $^{3}J_{COCC}$ of coupling constant of 3.1 Hz leaves only one solution. Care must be taken when interpreting coupling constants to account for possible flexibility, nevertheless availability of an extended set of coupling constants is clearly beneficial to this analysis. This is the motivation behind our efforts in developing INADEQUATE methods with an emphasis on the measurement of long-range coupling constants presented below.

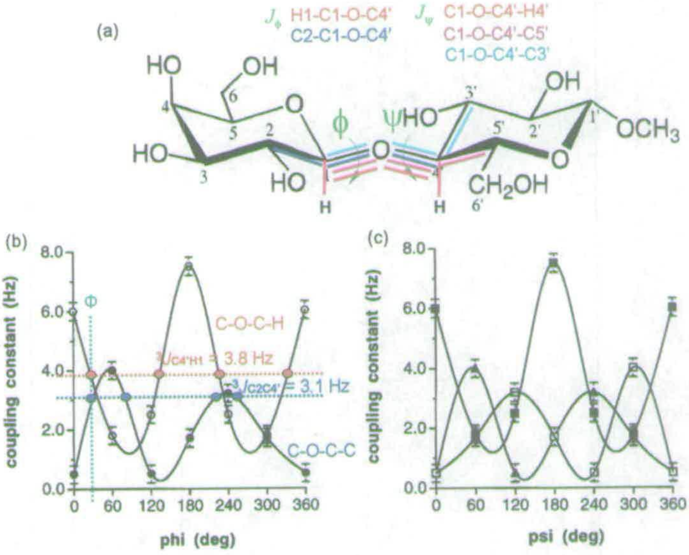


Fig.III.3.1 (a) Coupling constants reporting on dihedral angles, Φ and Ψ, across the glycosidic linkage in a disaccharide unit. (b) Correlation between the Φ angle and $^{3}J_{COCH}$ and $^{3}J_{COCC}$ coupling constants. [99] (c) Correlation between Ψ angle and one $^{3}J_{COCH}$, two $^{3}J_{COCC}$ coupling constants [99].

The original INADEQUATE experiment [98] utilizes ^{13}C polarization and ^{13}C detection. Introduction of inverse probes and pulsed field gradients made proton-detected INADEQUATE experiments possible. The sensitivity of NMR experiments is determined by the product of the gyromagnetic ratio in the form of $\gamma_{Start}\gamma^{3/2}_{Detected}$, where γ_{Start} and $\gamma_{Detected}$ correspond respectively to the nucleus used to generate the initial polarization and to the nucleus used for detection. Four possible combinations have relative sensitivities given in Table III.3.1. However, depending on individual experiments, an additional factor comes into play, and as will be seen later, the relative sensitivities of various implementations of the INADEQUATE experiment are much closer than given in Table III.3.1.

Table III.3.1 Relative sensitivities of four experiments with different combinations of γ_{Start} and $\gamma_{Detected}$

Experiment	1D ^{13}C	INEPT	Reverse INEPT	HSQC
Relative sensitivity	$\gamma_C\gamma_C^{3/2} \sim 1$	$\gamma_H\gamma_C^{3/2} \sim 4$	$\gamma_C\gamma_H^{3/2} \sim 8$	$\gamma_H\gamma_H^{3/2} \sim 32$

Several INADEQUATE-based experiments developed for the measurement of the carbon-carbon coupling constants and tested using mono- and disaccharides are presented next. These experiments provide valuable information about the geometry of the glycosidic linkage in a disaccharide unit.

III. 3.1 DJM-INEPT-INADEQUATE

One of the factors contributing to the low sensitivity of the INADEQUATE experiments aimed at the measurement of long-range carbon-carbon coupling constants is the large spread of their values (0 - 6 Hz). Some magnetization is lost because of a mismatch between the coupling constants and the evolution intervals which are optimized for an average coupling constant. The first INADEQUATE experiment tested in our laboratory removed this limitation. It is referred to as ^{1}H-detected double-*J*-modulated (DJM) INEPT-INADEQUATE [100] and contains no fixed delays. Because of this, it is capable of the measurement of both one-bond and long-range carbon-carbon coupling constants in a single experiment. When I joined this project, it was well advanced. My contribution was in assessing the performance of this method for fast relaxing compounds and small amounts of sample. The method is first briefly outlined before describing my analysis.

In DJM-INEPT-INADEQUATE both fixed C-C coupling evolution intervals of the original INEPT-INADEQUATE experiment [101] are replaced by variable time intervals as shown in the pulse sequence of Fig.III.3.1.1. These variable time intervals are incremented simultaneously with the double-quantum (DQ) evolution period (t_1) causing evolution of both ${}^1J_{CC}$ and ${}^nJ_{CC}$ coupling constants. The signal is modulated by $\sin^2(\pi J_{CC}\kappa t_1) = 0.5\ [1 - \cos(\pi J_{CC}\kappa t_1)]$, where κ is the scaling factor of the J-evolution in t_1. Therefore the cross peaks consist of three absorption lines in F_1. The two outer lines are separated by $2\kappa J_{CC}$ Hz and have opposite phases with respect to the central line that is twice as intense.

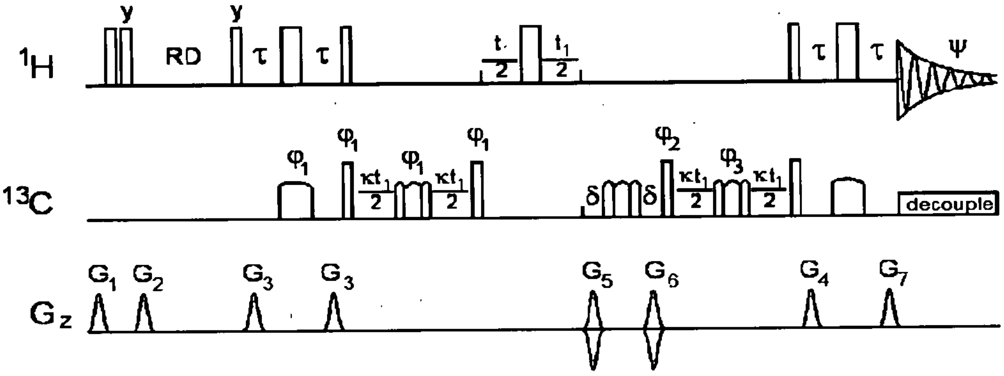


Fig.III.3.1.1 Pulse sequence of DJM-INEPT-INADEQUATE. Thin and thick open bars represent 90° and 180° rectangular pulses. The ${}^{13}C$ pulse applied after G_4 was a 120° pulse. The first and the last ${}^{13}C$ 180° pulses were regular smoothed chirp pulses (500 μs), while the other five ${}^{13}C$ 180° pulses were composite smoothed chirp pulses (2 ms). Unless stated otherwise, all pulses were applied from the x direction. The following delays were used: RD = 1 s, $\tau = 0.25/{}^1J_{CH}$, $\delta = 1.2$ ms, $\delta 1 = 3$ μs, $\tau_1 = 1/(k^*{}^1J_{CH}) + 0.5$ ms, where k=4 for CH only and k=8 for all multiplicities. Addition of 0.5 ms compensates for partial refocusing during the composite adiabatic pulse. The phases of pulses were cycled as follows: φ_1 = x, -x; φ_2 = 8x, 8(-x); φ_3 = 4x, 4(-x); φ_4 = 2x, 2(-x); φ_5 = 4y, 4(-y); Ψ = (x, -x, x, -x), 2(-x, x, -x, x), (x, -x, x, -x). PFGs were applied during 1 ms using the following strengths (100% = 75 Gauss/cm): G_1 = 33%, G_2 = 22%; while G_3 = 48%, G_4 = -48%, and G_5 = 48.19% were used for N-type signal selection, G_3 = -48%, G_4 = 48% and G_5 = 48.19% were used for P-type signal selection together with incrementing phase φ_5 by 180°. The spectra were processed using the echo-antiecho protocol.

The element PFG-90_x90_y-PFG at the end of the t_2 acquisition time was used to remove artifacts originating in the coherent proton magnetization of ${}^{12}C$-attached protons that survives between individual scans. The artifacts arising from the imperfect 180° ${}^{13}C$ pulses and single-quantum ${}^{13}C$ coherences were removed by the phase cycling, PFGs (G_3 and G_4) and the use of the adiabatic 180° ${}^{13}C$ pulse [102]. The use of long (2 ms) adiabatic pulses resulted in the same initial J-evolution for $t_1 = 0$. Consequently, particularly for ${}^1J_{CC}$ cross peaks, mixed phase multiplets were obtained. As will be shown later, this did not compromise the coupling constant analysis. A notable improvement of S/N was observed when adiabatic pulses were used.

Common to all INEPT-INADEQUATE methods the cross peaks between quaternary carbons are missing in the DJM-INEPT-INADEQUATE spectra. Nevertheless, the presence of long-range cross peaks in the spectrum reduces the ambiguities arising from the absence of the one-bond cross-peaks between quaternary carbons in the INEPT-INADEQUATE optimized for ${}^{1}J_{CC}$ coupling constants. The DJM-INEPT-INADEQUATE method is illustrated by tracing the carbon skeleton of the Me-β-D-xylopyranoside and measuring all its carbon-carbon coupling constants (Fig.III.3.1.2). All cross-peaks in the spectra show fine structure (Fig.III.3.1.2 and Fig.III.3.1.3), from which the values of J_{CC} coupling constants can be extracted. A more complicated shape of multiplets reduces the sensitivity of the method. The sensitivity of the DJM-INEPT-INADEQUATE is 50% of a JM-INEPT-INADEQUATE [103] experiment optimized for ${}^{1}J_{CC}$ coupling constants (Fig.III.3.1.3 e,f). The intensity of the long-range cross peaks in the DJM-INEPT-INADEQUATE is, however, enhanced by the fact that no fixed intervals are contained in the pulse sequence. If only the correlation information is of interest, this can be obtained from the presence of the central peak of the multiplets which is half as intense as the cross peaks of a regular INEPT-INADEQUATE experiment.

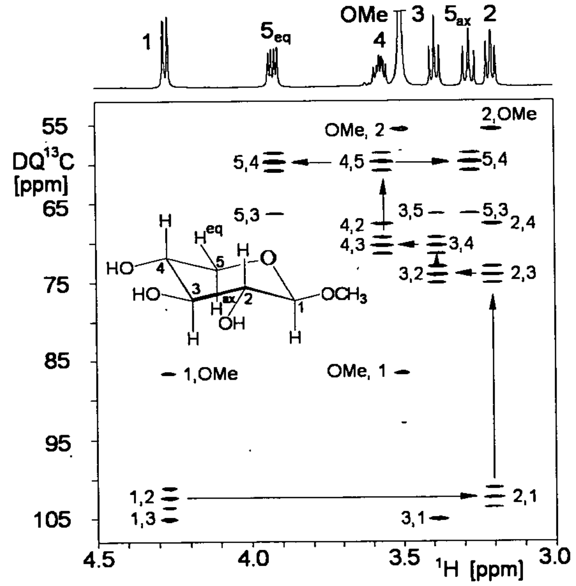


Fig.III.3.1.2 A 2D DJM-INEPT-INADEQUATE spectrum of Me-β-D-xylopyranoside. Cross peaks are numbered according to the structure shown in the inset. The first number in the labels of ${}^{13}C$ DQ coherences identifies the proton on which the magnetization originated is detected. Arrows indicate the carbon–carbon connectivity mediated by ${}^{1}J_{CC}$ coupling constants.

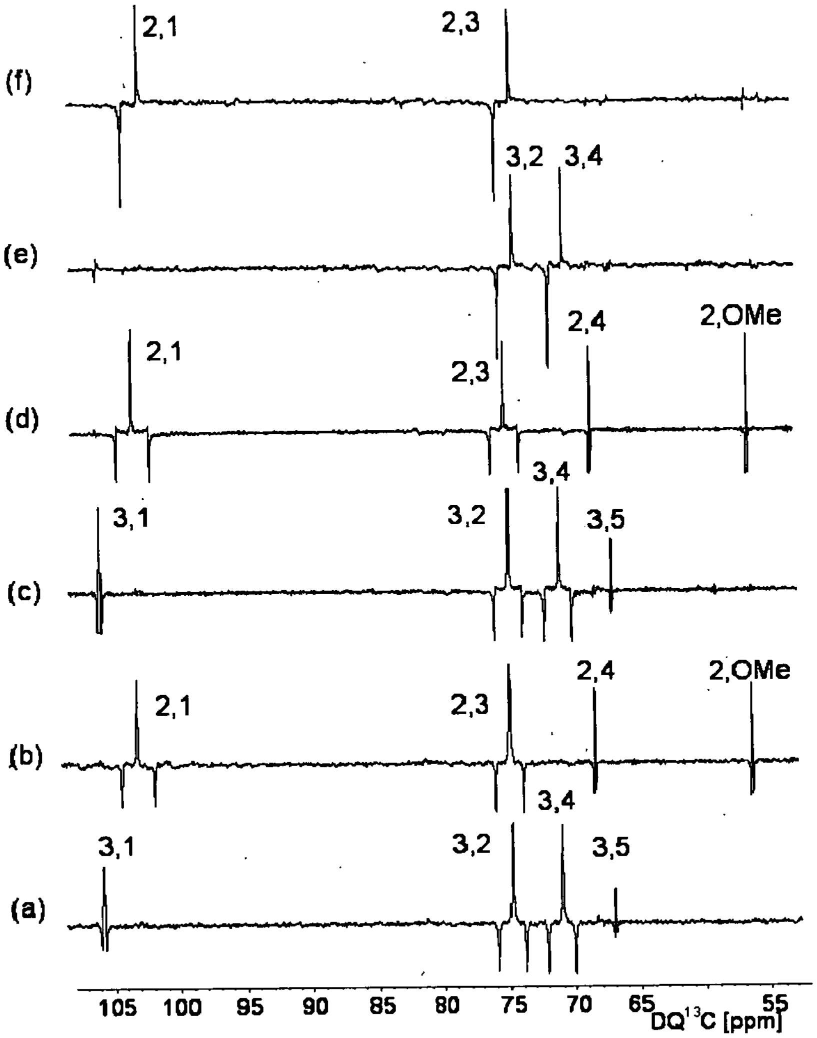


Fig.III.3.1.3 F_1 traces from 2D DJM-INEPT-INADEQUATE spectra at ^1H chemical shifts of H3 (a) and (c) and H2 (b) and (d), respectively. The pulse sequence in Fig.III.3.1.1 with rectangular (a) and (b) and adiabatic (c) and (d) 180 ° ^{13}C pulses were used to acquire the 2D spectra. (e) and (f) Equivalent traces from a 2D JM-INEPT-INADEQUATE spectrum optimized for $^1J_{CC}$. Cross peaks are labeled using the numbering shown in Fig.III.3.1.2.

The analysis of the coupling constants was performed in the time domain in order to eliminate the effects of limited digital resolution in the t_1 dimension of 2D experiment. The initial 2D Fourier transformation was followed by extraction of columns containing signals (Fig.III.3.1.4). Each multiplet excised from these F_1 traces was then inverse-Fourier transformed. The fast ^{13}C-^{13}C DQ frequency component of the resulting FID was removed by applying a procedure described by Stonehouse and Keeler [104], which has been further modified to include phase optimization by using the Powell minimization algorithm (Dr. Tran N. Pham). The simplified FID is described by the following equation:

$$I_0 \times \sin^2(\pi J_{CC}\kappa t_1 + \varphi) \times \exp(-t_1/T_{2\mathrm{eff}}) \qquad \textbf{(Eq.III.3.1.1)}$$

A four-parameter fit (I_0, φ, T_{2eff} and κJ_{CC}) produces the values of J_{CC}. The whole process is illustrated in Fig.III.3.1.4 for one-bond and long-range cross peaks and the extracted coupling constants of Me-β-D-xylopyranoside are summarized in Table III.3.1.1. The obtained $^1J_{CC}$ coupling constants (Table III.3.1.1) are identical to literature values [105].

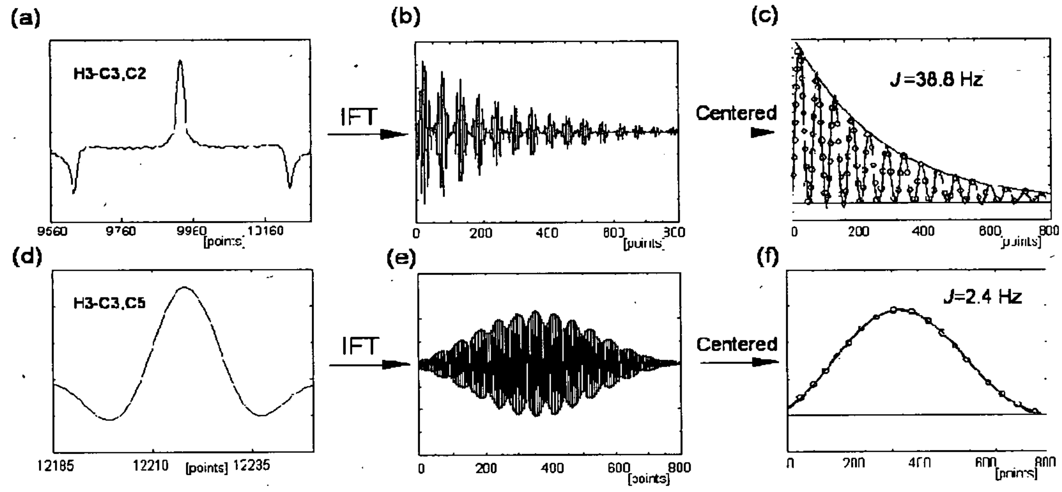


Fig.III.3.1.4 Two examples of the coupling constant determination from 2D DJM-INEPT-INADEQUATE cross peaks. (a) and (d) C3, C2 and C3, C5 cross peaks, respectively, excised from the H3 F_1 trace of the spectrum shown in Fig.III.3.1.3. (b) and (e) 800 real points of FIDs obtained after inverse Fourier transformation of the multiplets (a) and (d), respectively. The centered FIDs obtained by deconvolution with DQ frequencies are shown in (c) and (f). Also shown are the fitted centered FIDs (–o–) obtained using **Eq.III.3.1.1**. The exponential decay function in (c) was drawn using the value $T_{2\mathrm{eff}} = 297$ ms obtained by the fitting.

The effective relaxation time is used rather than the ^{13}C relaxation time for several reasons. Firstly, the proton spin flips contribute towards the relaxation of ^{13}C coherences. Secondly, the t_1 interferograms reflect the relaxation of a mixture of single- and double-quantum coherences. Thirdly, the fitted relaxation rates absorb the losses of the signal due to the diffusion taking place between the de- and refocusing PFGs. Finally and most significantly, the use of J-scaling accelerates the apparent relaxation of these coherences.

Table III.3.1.1 Carbon-carbon coupling constants of Me-β-D-xylopyranoside, in Hz, determined from the spectrum shown in Fig.III.3.1.2

	C_1C_2	C_2C_3	C_3C_4	C_4C_5	C_1C_3	C_2C_4	C_3C_5	C_1OMe	C_2OMe
J_{CxCy}	46.8	38.8	39.1	40.0	4.1	2.9	2.4	2.3	3.2
	46.8	38.8	39.0	39.9[b]	4.0	2.9	2.2[b]	2.4	3.2
				39.8[c]			2.3[c]		
	46.8[d]	38.8[d]	39.1[d]	39.9[d]					

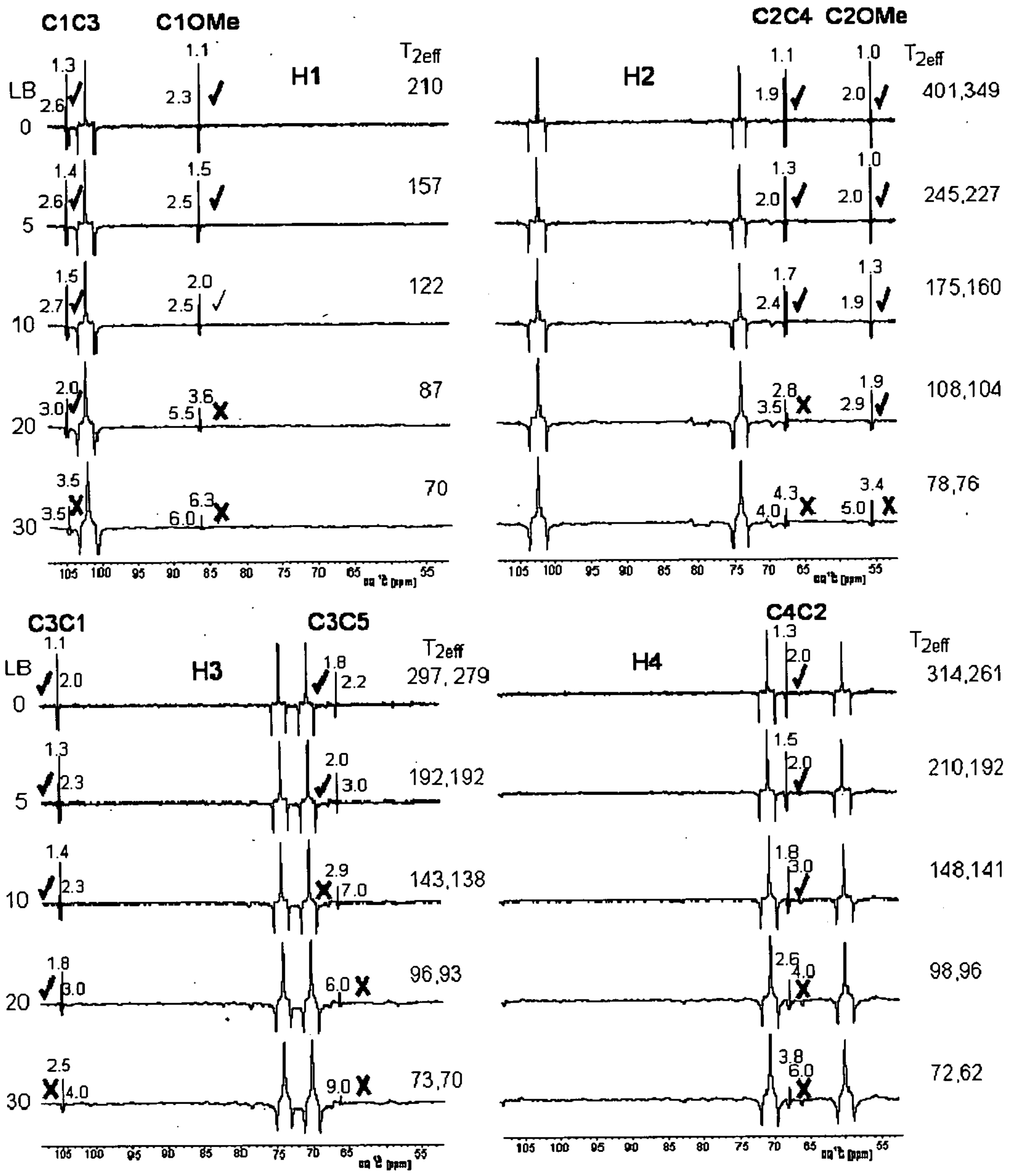
[a] The first row gives values determined from HxCxCy cross-peaks, while the second corresponds to the HyCyCx cross-peaks

[b] Determined from H5$_{eq}$ cross-peak

[c] Determined from H5$_{ax}$ cross-peak

[d] Values taken from [105]

This method worked very well for a monosaccharide which has long relaxation times. We have decided to simulate the effects of the relaxation factor using the acquired spectrum of Me-β-D-xylopyranoside in order to assess the applicability of this method to larger molecules. The effect of relaxation was simulated by applying exponential line broadening prior to the Fourier transformation in the t_1 dimension. A series of F_1-traces obtained by using exponential line-broadening values of 0, 5 10, 20 and 30 Hz is shown in Fig.III.3.1.5. This simulates approximately four-fold reduction of T_{2eff} relaxation times (from about 300 to 70 ms). It can be seen from Fig.III.3.1.5 that, with the exception of the cross peaks of the OMe carbon, the long-range cross peaks show a lower intensity compared to one-bond cross peaks. This is because of partial cancellation of opposite phase lines. It is well known that the cancellation of lines in antiphase multiplets leads to inaccurate coupling constants. Such cancellations are more pronounced for broad lined associated with faster relaxation. Individual multiplets were analyzed as described above and the obtained coupling constants are summarized in Table III.3.1.2 and graphically presented in Fig.III.3.1.6. The deviations of the determined $^1J_{CC}$ coupling constants increased steadily from their $J_{LB} = 0$ values with increasing line broadening to a maximum of ±0.4 Hz for $J_{LB} = 30$ values. This is a relatively small deviation and we can conclude that for $^1J_{CC} \sim 40$ Hz and a scaling factor $\kappa = 4$, DJM-INEPT-INADEQUATE provides $^1J_{CC}$ coupling constants accurate within 1% for T_{2eff} relaxation times as short as 70 ms.


C1C3
C1OMe
H1
T2eff
LB
0
5
10
20
30
210
157
122
87
70
C2C4
C2OMe
H2
401,349
245,227
175,160
108,104
78,76
C3C1
C3C5
H3
297, 279
192,192
143,138
96,93
73,70
C4C2
H4
314,261
210,192
148,141
98,96
72,62
105
100
95
90
85
80
75
70
65
55
δ¹³C [ppm]

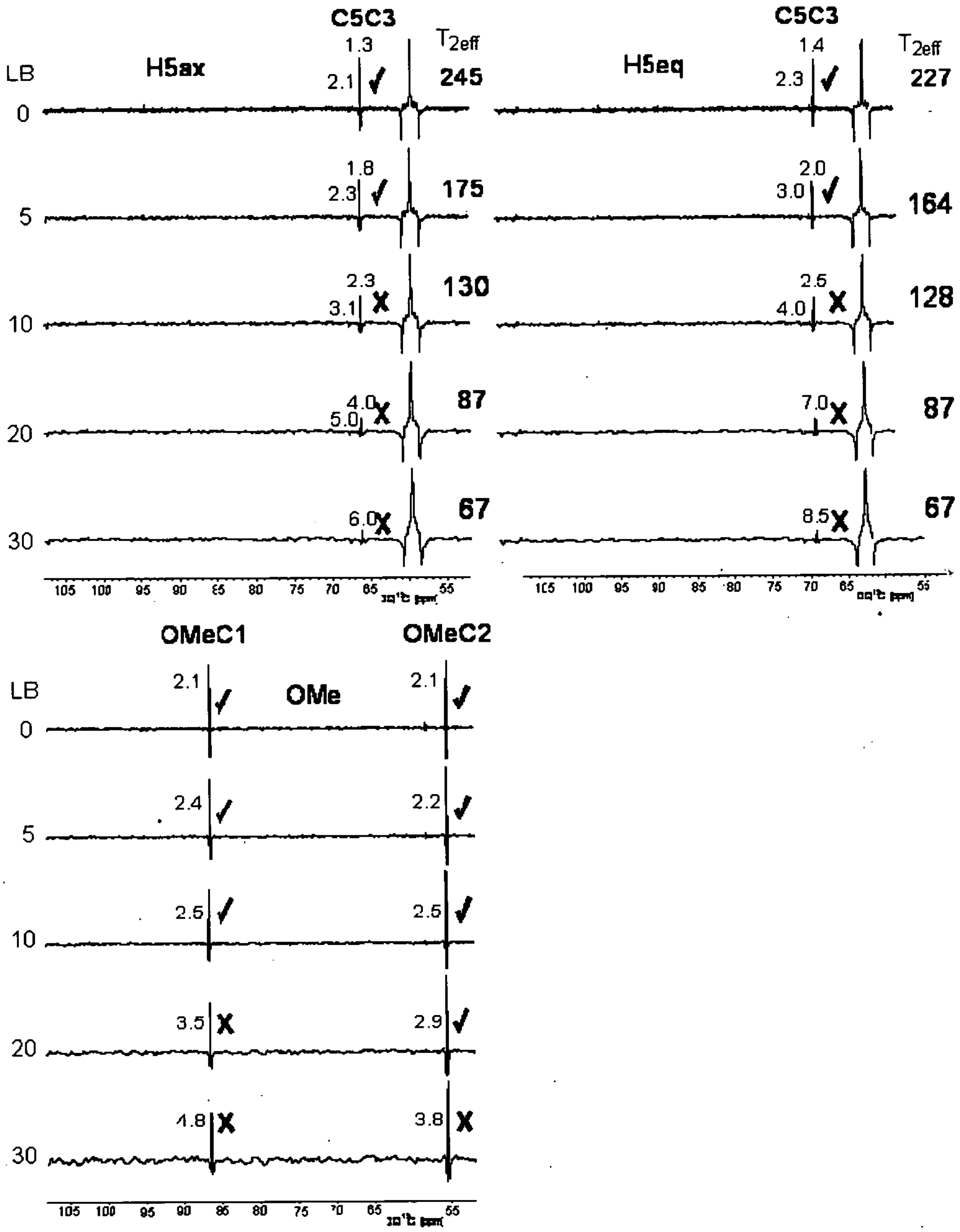


Fig.III.3.1.5 Effects of fast T_{2eff} relaxation on the appearance of DJM-INEPT-INADEQUATE spectra. Shown are F_1 traces through individual protons extracted from spectra processed using increasing line broadening (LB). Numbers beside and above the long-range correlation cross peaks give the intensity ratio of the positive and negative peaks *within* the long-range peaks (R_1) and the intensity ratio between the positive peaks of the one-bond and long-range peaks (R_2), respectively. Symbol √ is used to indicate that $R_1 \leq 3$ and $R_2 \leq 2$ and thus the cross peak is suitable for coupling constant determination; otherwise symbol × is used. The effective spin-spin relaxation times, T_{2eff}, as determined by fitting the one-bond cross peaks are given on the right edges of the spectra.

Table III.3.1.2 $^{n}J_{CC}$ coupling constants in Hz obtained from spectra processed with increasing line broadening

$^{n}J_{CC}$ (Hz)	LB = 0	LB = 5	LB = 10	LB = 20	LB = 30
H_1C_1-C_3	4.2	3.9	3.7	3.5	3.9
H_1C_1-C_2	46.8	46.8	46.8	46.9	46.9
H_1C_1-OMe	2.3	2.3	2.5	2.7	3.0
H_2C_2-C_1	46.8	46.8	46.7	46.7	46.6
H_2C_2-C_3	38.8	38.8	38.8	38.8	38.8
H_2C_2-C_4	2.9	3.0	3.0	3.1	3.8
H_2C_2-OMe	3.2	3.2	3.2	3.4	1.9
H_3C_3-C_1	4.0	4.0	4.1	4.1	4.4
H_3C_3-C_2	38.8	37.7	38.7	38.6	38.6
H_3C_3-C_4	39.1	39.1	39.2	39.2	39.3
H_3C_3-C_5	2.4	2.5	2.4	1.0	0.7
H_4C_4-C_2	2.9	3.0	3.0	1.6	1.3
H_4C_4-C_3	39.0	39.0	39.0	39.0	38.9
H_4C_4-C_5	40.0	40.0	40.0	39.9	39.9
$H_{5eq}C_5$-C_3	2.2	2.2	2.3	0.4	0.6
$H_{5eq}C_5$-C_4	39.9	39.8	39.8	39.6	39.4
$H_{5ax}C_5$-C_3	2.3	2.3	2.3	0.8	0.8
$H_{5ax}C_5$-C_4	39.7	39.6	39.5	39.4	39.2
OMe-C_1	2.4	2.4	2.6	2.8	2.1
OMe-C_2	3.2	3.2	3.3	3.4	3.6

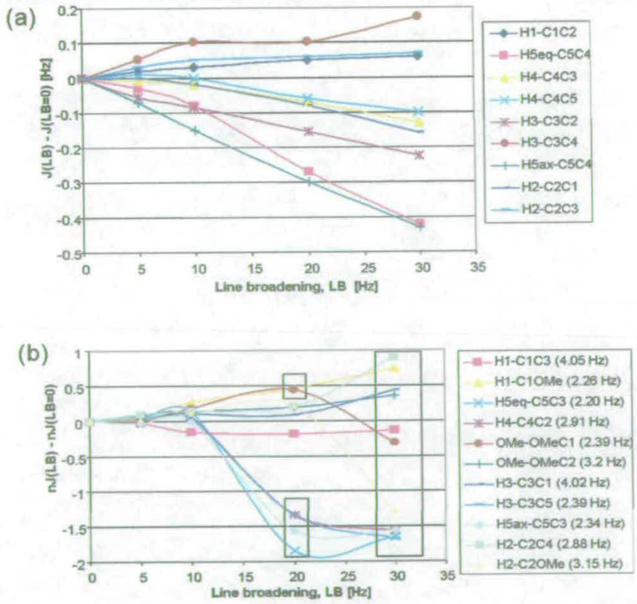


Fig.III.3.1.6 Effects of fast T_{2eff} relaxation on the determination of carbon–carbon coupling constants. (a) The deviations from $^{1}J_{CC}$ values determined using LB = 0 increase steadily with increasing line broadening (shorter T_{2eff} relaxation times). This is most prominent for the coupling constants determined using H5 protons due to the fastest relaxation of coherences involving this CH_2 group. (b) The same as in (a) but for $^{n}J_{CC}$ coupling constants. The values in rectangular boxes are from the analysis of cross peaks that do not comply with the criteria for reliable determination of coupling constants as defined in the text.

For small ${}^{n}J_{CC}$ values, the results of the analysis were not so positive. The use of large line broadening resulted in unreliable results because of the increasing cancellation of negative and positive lines in the long-range correlation cross-peaks. Two criteria, independent of T_{2eff} or ${}^{n}J_{CC}$, were adopted to assess the reliability of the obtained long-range coupling constants. The first criterion is based on the observation that the partial line cancellation increases the ratio, R_1, between the absolute intensities of the central and the two outer lines of a multiplet. Ideally, when there is no cancellation $R_1 = 2$. Based on our analysis, reliable coupling constants are obtained when $R_1 \leq 3$ (Fig.III.3.1.5). The second criterion is the ratio of the intensities of the central lines of ${}^{1}J_{CC}$ and ${}^{n}J_{CC}$ cross-peak, $R_2 \leq 2$ (Fig.III.3.1.5). Even when these criteria are fulfilled the values of ${}^{n}J_{CC}$ obtained at larger line broadening differ < ±0.3 Hz from ${}^{n}J_{LB} = 0$ values. Considering the small values of long-range coupling constants this is a large relative error. It is therefore advisable that all precautions are taken in order to prolong the effective relaxation times. One such approach is presented in the next chapter.

For the moment it can be concluded that similarly to simple antiphase multiplets, the coupling constants determined from $\sin^2$ modulated multiplets can yield erroneous coupling constants due to partial cancellation of lines. This problem is aggravated by the poor signal-to-noise ratios (SNRs). Therefore it was decided to determine the effect of poor SNRs on the reliability of coupling constants determined from DJM-INEPT-INADEQUATE spectra. This analysis was combined with increased effective relaxation simulated by using increasing LB values during the spectra processing. Only the long-range peaks that fulfilled the criteria outlined above were taken forward for this analysis.

Random noise was extracted from the blank regions of the spectrum processed using ${}^{n}J_{LB} = 0$. Using increasing scaling factors, the noise traces were added to the traces containing peaks in order to construct new traces with increasing SNR values. SNR was adjusted separately for individual cross peaks to achieve values of 10, 20, 40 and 60:1 as illustrated in Fig.III.3.1.7. Five different noise spectra were used and the average coupling constant together with RMSD was calculated. For ${}^{1}J_{CC}$ coupling constants the RMSDs increased exponentially with decreasing SNR (Fig.III.3.1.8), but even for SNR = 10 and LB = 30 Hz these were < 0.25 Hz. In other words, the deviation was always smaller than the systematic errors caused by fast relaxation. The RMSDs for ${}^{n}J_{CC}$ coupling constants followed a similar trend but increased faster with increasing line broadening values and decreasing SNR (Fig.III.3.1.8). For SNR ≥ 20 the RMSD were < 0.25 Hz.

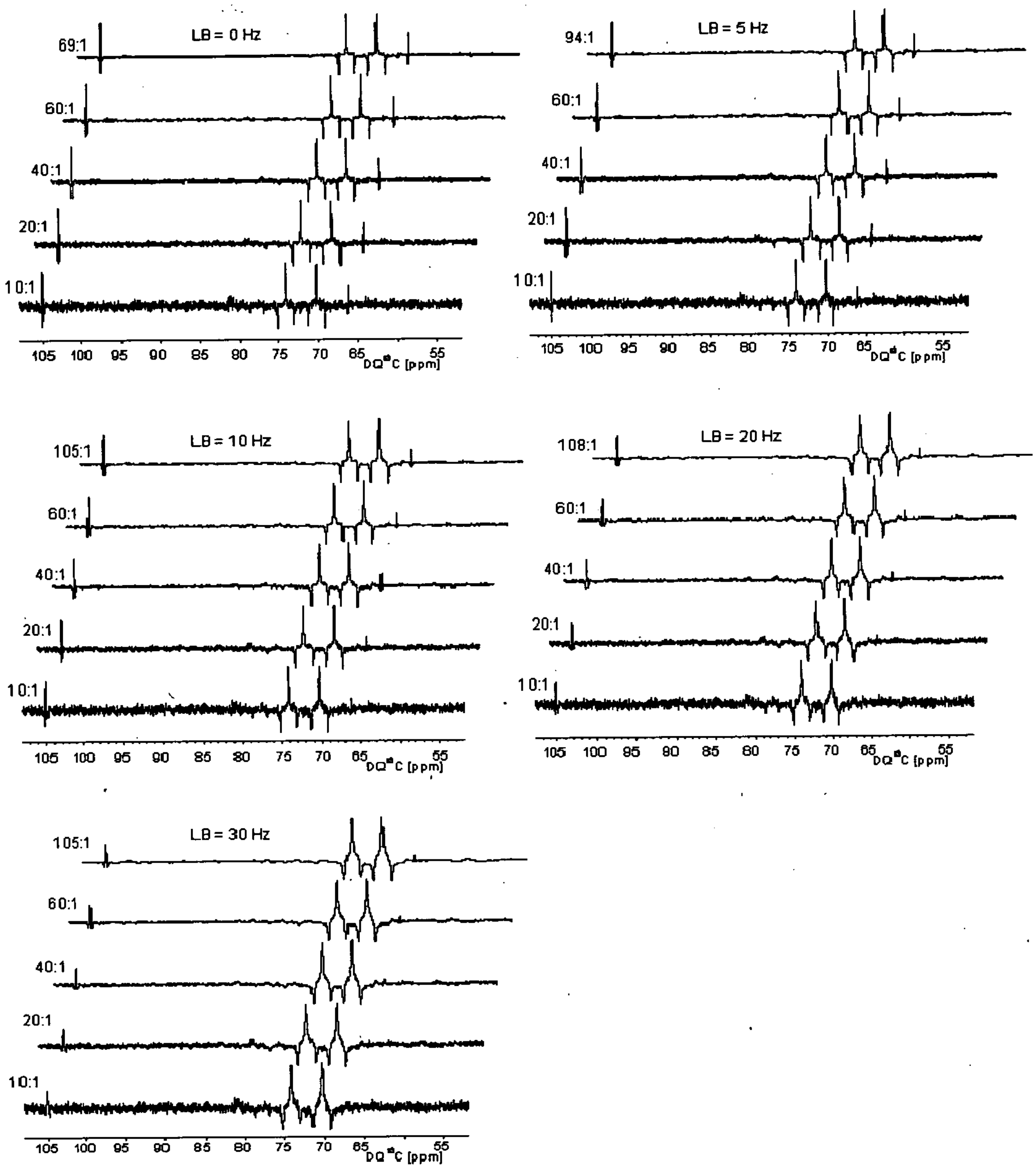


Fig.III.3.1.7 Effects of decreasing signal-to-noise ratios on the DJM-INEPT-INADEQUATE spectra. F_1 traces through H3 taken from the spectrum shown in Fig.III.3.3.1.2 processed with different line broadenings in F_1 and increasing levels of noise as described in the main text. This is one of the five sets of data used in the analysis.

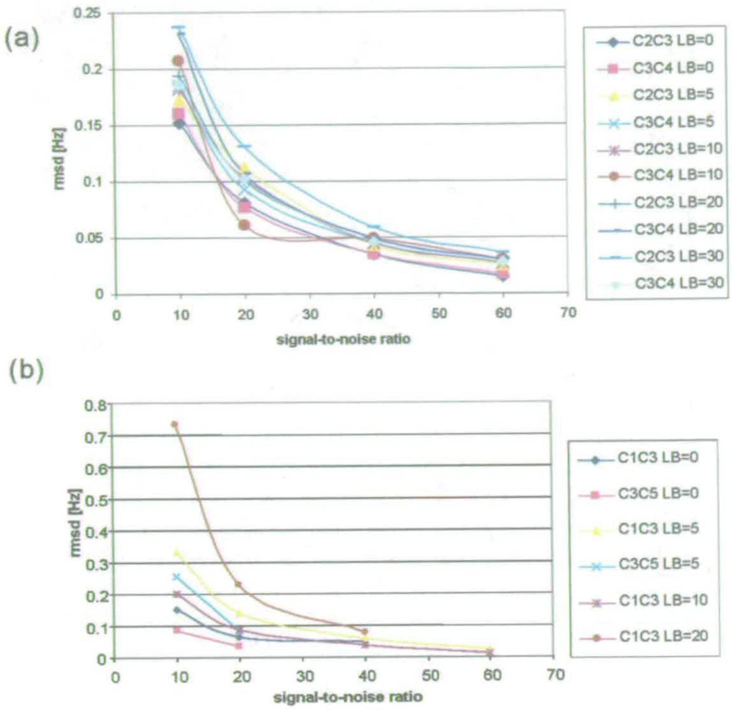


Fig.III.3.1.8 Effects of decreasing signal-to-noise ratios on values of carbon–carbon coupling constants. (a) $^1J_{CC}$ coupling constants. Standard deviations of $^1J_{C3C2}$ (38.8 Hz) and $^1J_{C3C4}$ (39.1 Hz) coupling constants obtained by the analysis of spectra processed with different line broadenings as a function of the increasing SNRs. (b) the same as (a) but for $^2J_{C3C1}$ (4.0 Hz) and $^2J_{C3C5}$ (2.4 Hz) coupling constants. Only cases which met the criteria for the analysis of $^nJ_{CC}$ coupling constants, as defined in the main text, were considered. Higher SNRs were not available for some of the cross-peaks.

In summary, one-bond $^1J_{CC}$ coupling constants can be determined with high accuracy (± 0.25 Hz) when $T_{2eff} \sim 70$ ms and SNR ≥ 10, while the determination of long-range $^nJ_{CC}$ coupling constants requires $T_{2eff} \geq 100$ ms and SNR ≥ 20 for the same accuracy. As the long-range coupling constants are of particular interest to this work, improvements of the current DJM-INEPT-INADEQUATE were explored aiming at increasing the effective relaxation times of involved coherences.

III. 3.2 DJM-REVINEPT-INADEQUATE:

Some sources of the increased relaxation of coherences discussed in the previous section can be removed. A modification described here prolongs the relaxation of coherences by eliminating the contributions of proton spin flips. During both the *J* evolution and DQ labeling intervals, the coherences in the DJM-INEPT-INADEQUATE contain the proton terms e.g. $\hat{I}_{1z}\hat{S}_{1x}\hat{S}_{2z}$ or $\hat{I}_{1z}\hat{S}_{1x}\hat{S}_{2y}\hat{I}_{2z}$. If these could be reduced to $\hat{S}_{1x}\hat{S}_{2z}$ and $\hat{S}_{1x}\hat{S}_{2y}$, longer effective relaxation times would result. The central part of the pulse sequence of DJM-INEPT-INADEQUATE (Fig.III.3.2.1a) was modified by adding the decoupling during the *J*-modulated and DQ evolution intervals (Fig.III.3.2.1b). Samples of mono-, di- and tetrasaccharides have been used to verify this experimentally. The effective relaxation times of ^{13}C single-quantum coherences of these carbohydrates were measured by a simple ^{13}C-detected spin-echo experiment with or without ^{1}H decoupling.

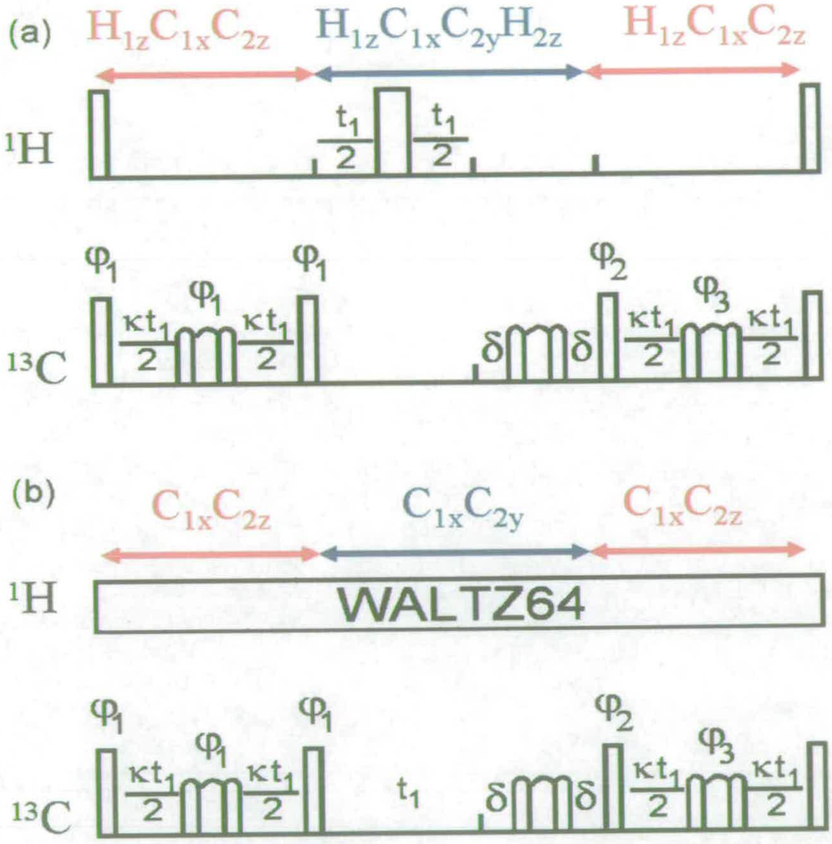


Fig.III.3.2.1 Partial DJM-INEPT-INADEQUATE pulse sequences without (a) and with (b) decoupling. Relevant coherences are shown above.

Fig.III.3.2.2 shows a comparison of the effective T_2 relaxation times for mono-, di- and tetrasaccharides applicable to the above experiments. These were acquired using a ^{13}C spin-echo experiment with or without the ^{1}H decoupling. When using ^{1}H decoupling the effective relaxation time was prolonged although this effect was progressively smaller for large molecules.

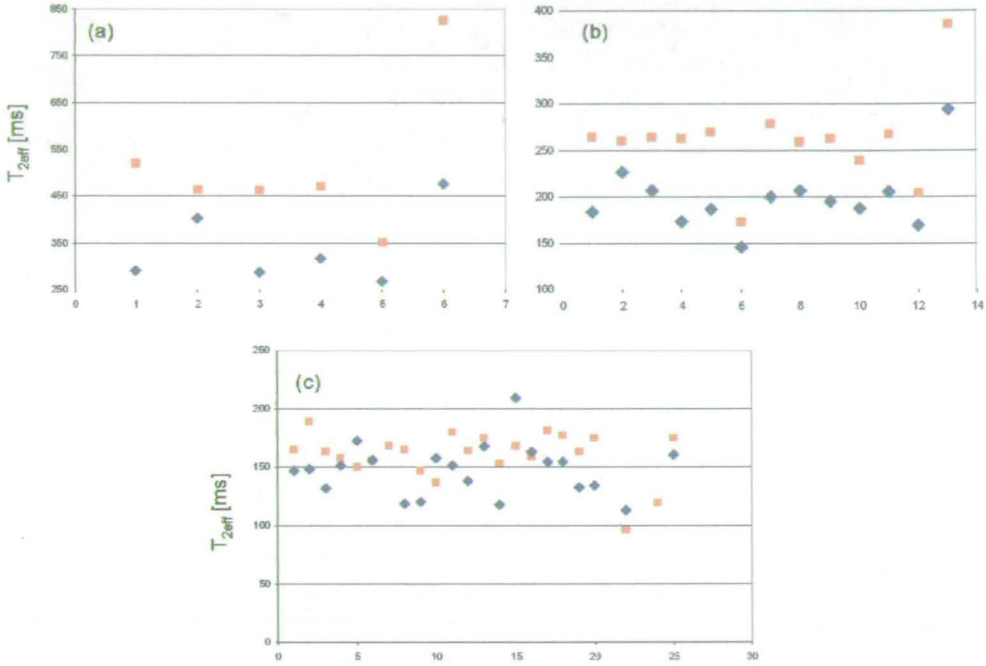


Fig.III.3.2.2 Effective relaxation time results, showing the effects of using decoupling during the coupling evolution intervals for (a) mono-, (b) di- and (c) tetrasaccharides. Red squares are results from decoupling and blue diamonds are without decoupling.

A practical implementation of this idea into D*J*M-INEPT-INADEQUATE requires several changes to be made to the original pulse sequence. To start with, the refocusing of $^1J_{CH}$ coupling constants is required before the ^{1}H decoupling can be switched on. During this refocusing (~3.3 ms) the carbon-carbon coupling constants would be evolving. Similar evolution would also take place during the defocusing period that is required for the reversed INEPT step. Even more detrimental, once the magnetization is on carbon, refocusing for CH_2 and CH_3 carbons requires a compromised setting of evolution intervals that leads to signal losses as illustrated in Fig.III.3.2.3.

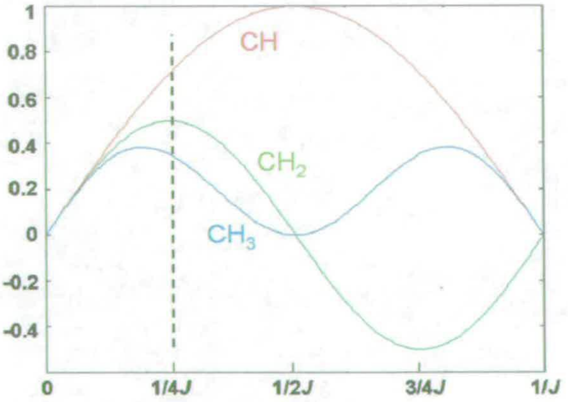


Fig.III.3.2.3 Creation of the in-phase ^{13}C magnetization for CH, CH_2 and CH_3 carbons during the refocusing interval. The dashed line represents the optimal setting for all multiplicities.

In order to alleviate these problems it is advantageous to start the experiment by utilizing the carbon polarization. This eliminates the first INEPT step and the associated refocusing interval. Such an implementation of the DJM-INEPT-INADEQUATE experiment shown in Fig.III.3.2.4 is referred to as DJM-REVINEPT-INADEQUATE. This experiment is only slightly less sensitive than its double INEPT version. The factor of 4 gained by starting the experiment with the proton polarization is countered by the heteronuclear NOE (max $1 + \gamma_H/2\gamma_C = 3$) in a carbon initiated experiment. In addition, the use of decoupling during the DQ evolution interval recovers one-half of the signals that are otherwise lost for correlations between two protonated carbons [106] as is often the case in carbohydrates. Theoretical relative intensities of the two experiments are given in the following Table III.3.2.1.

Table III.3.2.1 Sensitivity comparison between the INEPT and REVINEPT-INADEQUATE

Method	Sensitivity
INEPT-INADEQUATE (CH_x)	32 × 0.5 (ZQC) = 16
REVINEPT-INADEQUATE (CH only)	8 × 3 (NOE) = 24
REVINEPT-INADEQUATE (CH)	8 × 3 (NOE) × 0.71 = 17.0
REVINEPT-INADEQUATE (CH_2)	8 × 3 (NOE) × 0.5 = 12
REVINEPT-INADEQUATE (CH_3)	8 × 3 (NOE) × 0.35 = 8.4

Returning to the pulse sequence of Fig.III.3.2.4, the DJM-REVINEPT-INADEQUATE experiment starts from ^{13}C magnetization, and utilizing the SQ(^{13}C)-DQ-SQ(^1H) pathway, the magnetization is transferred to protons for detection. The modulation of cross peaks by J_{CC} coupling constants is produced during two J-modulated variable-time intervals. The second interval is followed by a defocusing interval that can be optimized either for CH only or for all carbon multiplicities. A compromised setting of this interval leads to some losses of signal as shown in Table III.3.2.1.

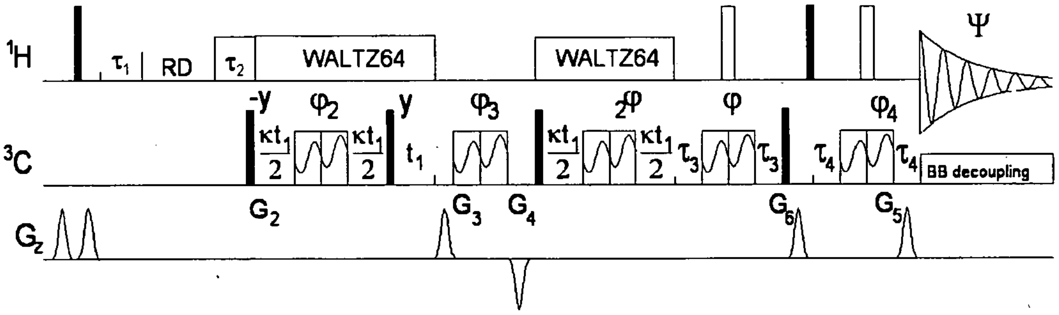


Fig.III.3.2.4 Pulse sequence of DJM-REVINEPT-INADEQUATE. Thin closed and thick open bars represent 90° and 180° rectangular pulses.

A 2D DJM-REVINEPT-INADEQUATE spectrum of Me-β-D-xylopyranoside optimized for all carbon multiplicities is shown in Fig.III.3.2.5. All cross peaks show fine structures (Fig.III.3.2.5), from which J_{CC} coupling constants were extracted.

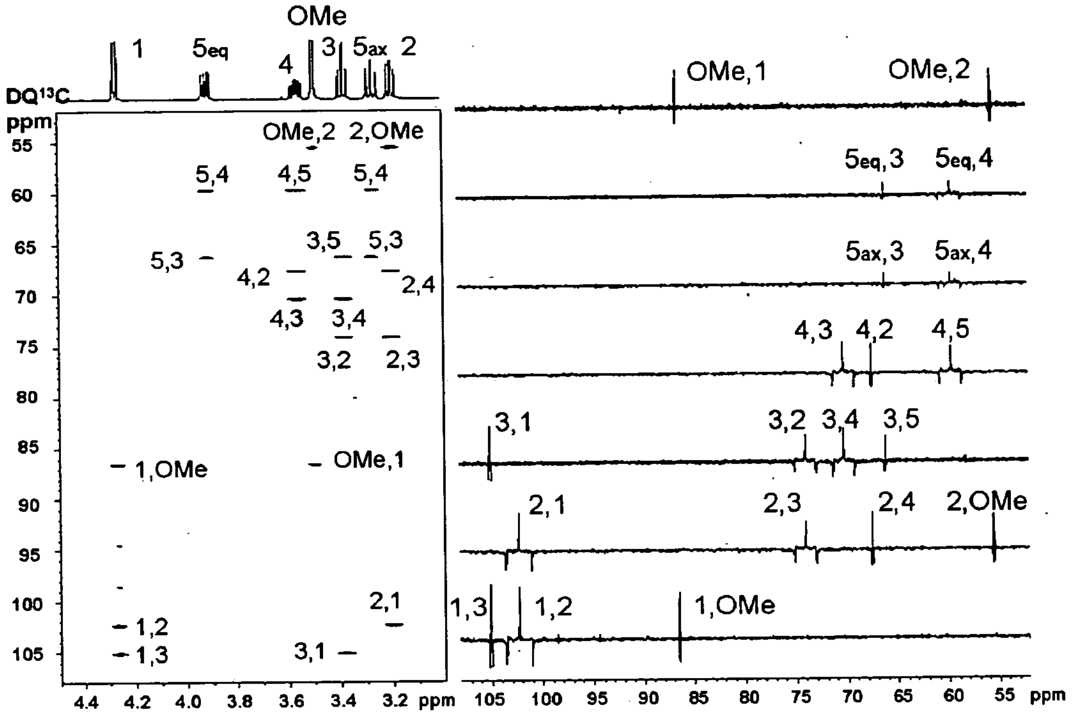


Fig.III.3.2.5 A 2D DJM-REVINEPT-INADEQUATE spectrum of Me-β-D-xylopyranoside and the F_1 traces at ^{1}H chemical shifts. Cross peaks are numbered the same as in Fig.III.3.1.2. The first number in the labels of ^{13}C DQ coherences identifies the proton on which the magnetization originated/is detected.

The intensity ratios (Table III.3.1.2) of the central positive lines in long-range and one-bond cross peaks is on average 20% higher compared to the DJM-INEPT-INADEQUATE spectrum. This reflects longer effective relaxation times in the REVINEPT experiment.

Table III.3.2.2 Ratio of the central positive lines of long-range to one-bond cross peaks in columns taken from DJM-INEPT-INADQUATE and DJM-REVINEPT-INADQUATE spectra.

Cross peaks	C_1C_3	C_1OMe	C_2C_4	C_2OMe	C_3C_1	C_3C_5	C_4C_2	$C_{5ax}C_3$	$C_{5eq}C_3$
DJM	0.80	0.97	0.80	0.92	0.84	0.51	0.80	0.74	0.72
DJM-REV	1.07	0.88	0.92	0.83	1.06	0.75	0.99	0.97	0.85

The J_{CC} coupling constants were obtained from the DJM-REVINEPT-INADQUATE spectrum (Table III.3.2.3) using the same procedure developed for the analysis of the DJM-INEPT-INADQUATE spectra. The long-range coupling constants (< 4.0 Hz) of the former spectrum are slightly smaller, which perhaps is an indication that even modest cancellation of lines can cause an overestimation of coupling constants.

Table III.3.2.3 J_{CC} coupling constants from the DJM-REVINEPT-INADEQUATE method

	C_1C_2	C_2C_3	C_3C_4	C_4C_5	C_1C_3	C_2C_4	C_3C_5	C_1OMe	C_2OMe
J_{CxCy}	46.7	38.8	39.1	39.8	4.1	2.8	2.1	2.3	3.0
	46.8	38.8	39.1	39.8[b]	4.2	2.8	2.3[b]	2.1	3.0
				39.7[c]			1.8[c]		

[a] The first row gives values determined from HxCxCy cross-peaks, while the second corresponds to the HyCyCx cross-peaks

[b] Determined from the $H5_{eq}$ cross-peak

[c] Determined from the $H5_{ax}$ cross-peak. These two last values of J_{C3C5} are clearly affected by a poor SNR.

In order to show the usefulness of this method for larger molecules than a monosaccharide and smaller amounts of sample, 25 mg of Me-β-D-lactoside were used to acquire a DJM-REVINEPT-INADEQUATE spectrum (Fig.III.3.2.6) in 63 hours. The J_{CC} coupling constants (Table III.3.2.4) were obtained from the F_1 proton traces (Fig.III.3.2.7).

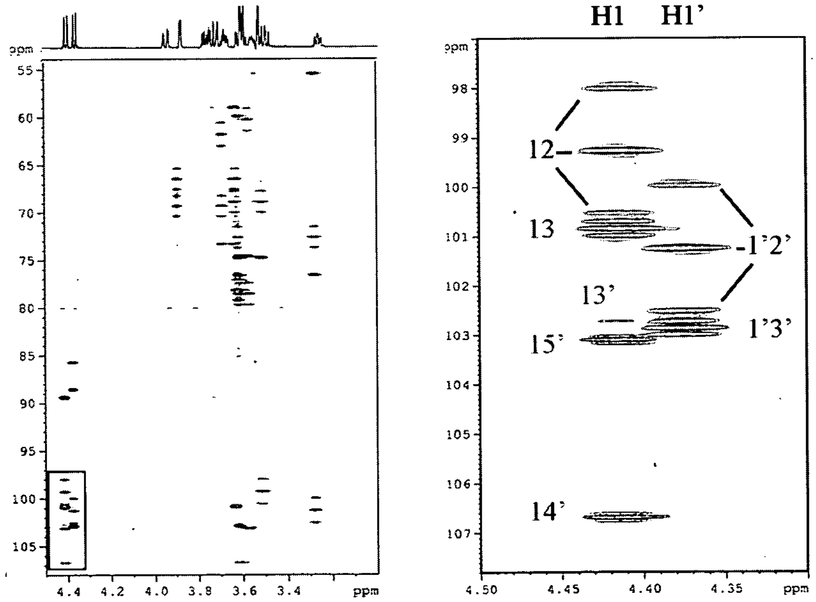


Fig.III.3.2.6 A 2D DJM-REVINEPT-INADEQUATE spectrum of Me-β-D-lactoside. On the right an expansion of the correlations of two anomeric protons is shown. Cross peaks 13', 14' and 15' correlate carbons of two rings.

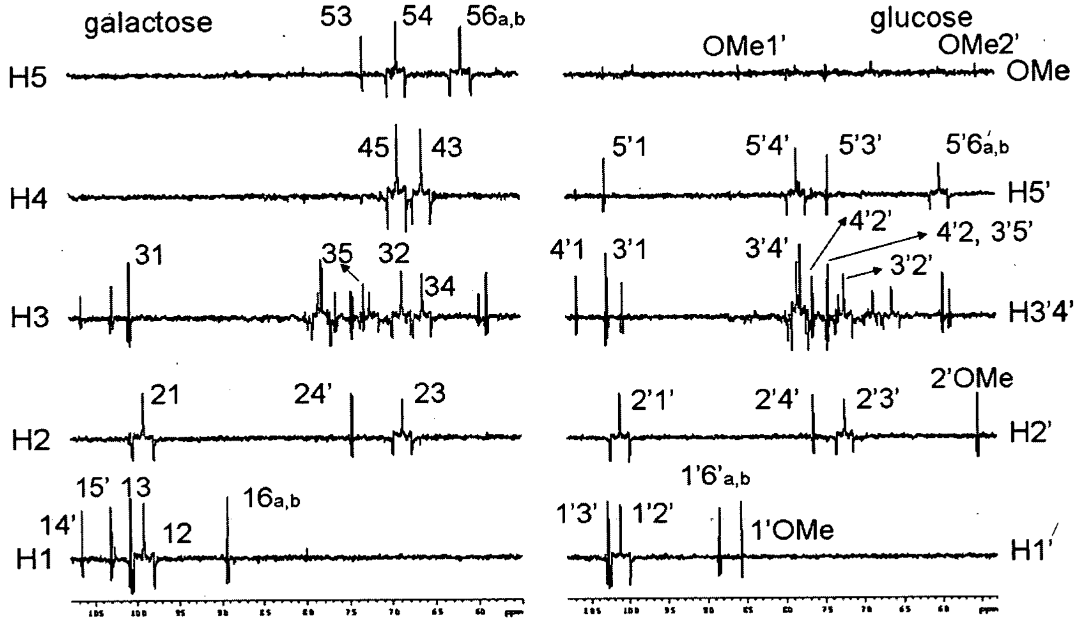


Fig.III.3.2.7 F_1 traces from the 2D DJM-REVINEPT-INADEQUATE spectrum of Me-β-D-lactoside

Table III.3.2.4 $^{1,n}J_{CC}$ coupling constants of Me-β-D-lactoside obtained from a 2D DJM-REVINEPT-INADEQUATE spectrum

C_1C_2	C_2C_3	C_3C_4	C_4C_5	$C_{1'}C_{2'}$	$C_{2'}C_{3'}$	$C_{3'}C_{4'}$	$C_{4'}C_{5'}$	$C_{1'}OMe$	$C_{2'}OMe$
46.8 46.8	40.0 40.2*	39.0 38.9	38.5 38.5	47.2 47.2	39.7 39.8	38.9	 41.6	2.1 2.9^+	3.0 2.8^+
C_1C_3	C_3C_5	$C_{1'}C_{3'}$	$C_{2'}C_{4'}$	$C_{3'}C_{5'}$	$C_1C_{4'}$	$C_1C_{3'}$	$C_1C_{5'}$	$C_2C_{4'}$	
5.2* 5.0	2.2 2.2	4.8* 4.4	2.5 2.1*	3.2* 3.0	2.2 2.6	~0 ~0	2.4 2.4	3.0 3.3*	

The first row gives values determined from HxCxCy cross-peaks, while the second corresponds to the HyCyCx cross-peaks

* extracted from overlapped cross peaks, likely in error

$^+$ represents values from cross peaks with very poor SNR

The $^nJ_{CC}$ coupling constants across the glycosidic linkage of the disaccharide are the most interesting. Traces through protons containing cross peaks of carbons from the two rings are shown in Fig.III.3.2.8.

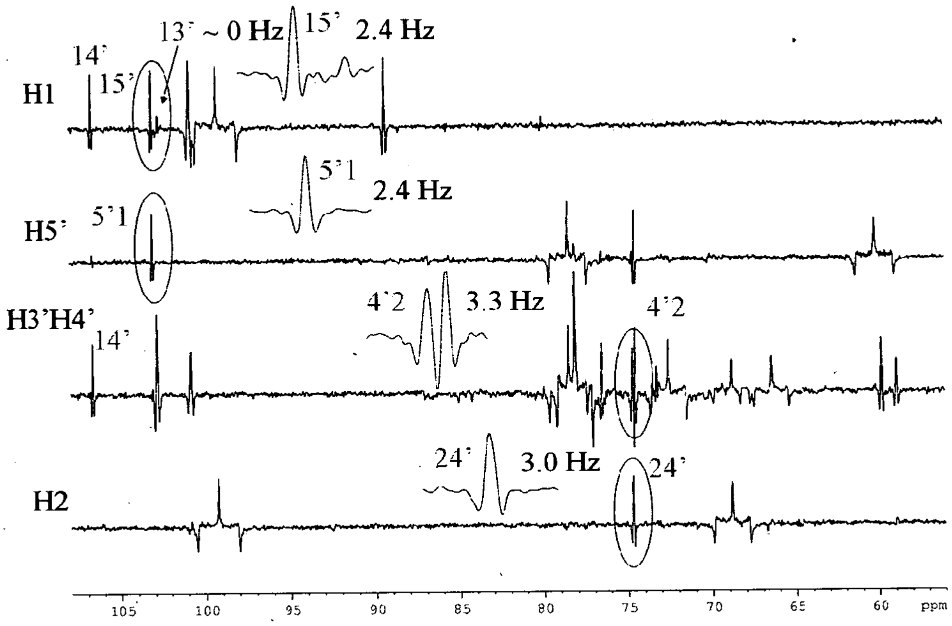


Fig.III.3.2.8 F_1 traces containing cross peaks of carbons across the glycosidic linkage of Me-β-D-lactoside

III. 3.3 *J*M-DEPT-INADEQUATE:

We have shown that the performance of the basic D*J*M-INEPT-INADEQUATE experiment can be improved by using ^{1}H decoupling during the long evolution delays. Longer effective relaxation times allow more accurate determination of small long-range carbon-carbon coupling constants. However, the improvements for large molecules are smaller due to shorter relaxation times of ^{13}C.

It seems that partial cancellation of opposite phase lines is a fundamental problem when it comes to the extraction of coupling constants from antiphase multiplets [107]. The original *J*-modulated INEPT-INADEQUATE experiment [108] suffers from this problem, as antiphase $^nJ_{CC}$ doublets are recorded in this experiment. In antiphase multiplets, the peak-to-peak distances are larger than the true coupling constants. In an attempt to remove this limitation, a modification has been proposed [109] that records in-phase doublets in F_1. However, if these doublets are not resolved, the coupling constants cannot be determined by this experiment and poorly resolved doublets lead to underestimation of the coupling constants. In the following a solution to this problem is proposed, which provides a way for the determination of very small $^nJ_{CC}$ coupling constants. It is based on acquisition of both in- and antiphase $^nJ_{CC}$ doublets. It has been shown [110] that addition/subtraction of in- and antiphase multiplets leads to simplified multiplets in which the active (antiphase) coupling constant is missing. Its value can still be determined as a displacement of the two simplified multiplets as illustrated in Fig.III.3.3.1 for resolved and nonresolved multiplets. It can be seen that in both cases the coupling constant can be determined. Crucial to the success of this method is that the scaling factor between the in- and antiphase multiplets is known [110]. This is not always the case, as the two kinds of multiplets are obtained with different efficiency. In this section, assisted by an additional experiment, a solution to this problem for the INADEQUATE experiments is presented. A scaling factor is obtained by comparing one resolved doublet and non-resolved doublets. Firstly, the ratio of the integrals of the two doublets in the in-phase spectra is used to calculate an initial J_{CC} coupling constant. Secondly, this J_{CC} coupling constant is used to calculate a scaling factor of the antiphase doublet. In this calculation the relaxation during the additional evolution interval was taken into account. Thirdly, the antiphase doublet is scaled down using the calculated scaling factor and added to or subtracted from the in-phase doublet to construct two singlets. Finally, the non-resolved J_{CC} coupling constant is read out from the splitting chemical shift difference of the two singlets. The coupling constant is therefore determined by the manipulation of in- and antiphase multiplets.

Different implementations of this basic principle are presented below to proton-detected INADEQUATE and later also to a carbon-detected INADEQUATE pulse sequences.

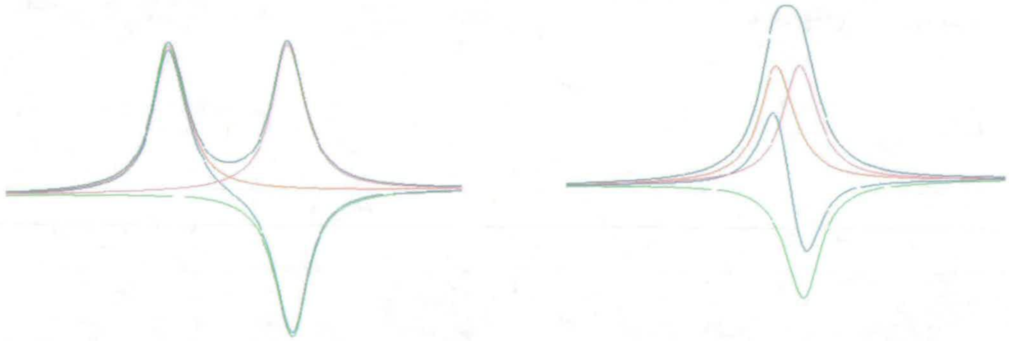


Fig.III.3.3.1 Simulated doublets for $J = 4$ Hz ($T_2 = 160$ ms) and $J = 2$ Hz ($T_2 = 111$ ms). Red, pink and green curves are singlets obtained by addition/subtraction of in- and antiphase multiplets, blue curves are antiphase multiplets and dark-blue curves are in-phase multiplets. The 4 Hz doublet is well resolved and the splitting represents the coupling constant but the 2 Hz in-phase doublet is not resolved and the value obtained from the antiphase doublets is not accurate.

The *J*M-DEPT-INADEQUATE experiment starts and ends up on protons. In order to maximize its sensitivity, ^1H decoupling is used during the long evolution intervals. As discussed previously, this doubles the signal intensities compared with INEPT-

INADEQUATE without ^{1}H decoupling as no leakage of magnetization to ZQ coherences occurs. The decoupling also prolongs the relaxation times of ^{13}C coherences as illustrated in the previous chapter. DEPT rather than INEPT was used for polarization transfer, which reduced the numbers of pulses. This experiment contains two CH refocusing/defocusing intervals which would reduce its sensitivity, if all CH_x multiplets were to be recorded. Therefore this experiment was only used for the measurement of $^{n}J_{CC}$ coupling constants between two CH carbons. Fig.III.3.3.2 shows the pulse sequence of *J*M-DEPT-INADEQUATE. Refocusing of ^{13}C-^{1}H coupling constants required before the ^{1}H decoupling means that some evolution of the ^{13}C-^{13}C coupling constant cannot be avoided. It was therefore decided to purge this magnetization in order to select only the in-phase ^{13}C magnetization. For this, a gradient enhanced z-filter proposed by Thrippleton and Keeler [92] was used. At the same time, purging of one-bond correlation was achieved by setting the $2\Delta_1$ delay to $1/2(^{1}J_{CC})$. This reduces the possibility of overlap in the final spectrum. The following incrementable *J* evolution interval therefore starts with pure in-phase carbon magnetization of long-range coupled ^{13}C pairs. Although some magnetization is lost, the long-range correlations are barely affected. Selection of DQ coherences and their frequency labeling is followed by a fixed refocusing interval of $2\Delta_2$ optimized for $^{n}J_{CC}$ coupling constants. In such an experiment antiphase $^{n}J_{CC}$ are recorded in F_1. The in-phase doublets are acquired by inserting another block into the pulse sequence prior to the 90° ^{13}C pulse that generates MQ coherences. This block selects the cosine modulated magnetization generated during the κt_1 interval by purging the antiphase component. An additional defocusing interval of $2\Delta_2$ generates the required antiphase magnetization. As mentioned above, this experiment is optimized for CH carbons in order to maximize its sensitivity. This is achieved by setting the flip angle of the DEPT to 90°.

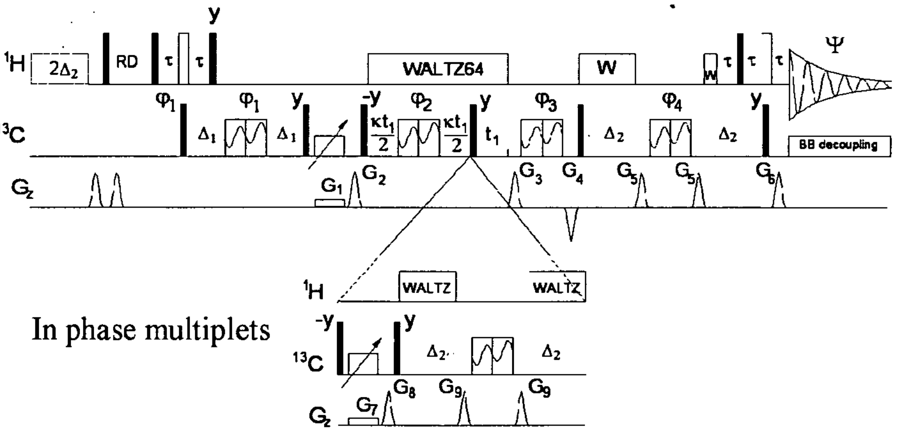


Fig.III.3.3.2 Pulse sequence of *J*M-DEPT-INADEQUATE. Thin closed and thick open bars represent 90° and 180° rectangular pulses.

The application of this pulse sequence to Me-β-D-xylopyranoside and Me-β-D-lactoside is presented next. The spectra (Fig.III.3.3.3) were processed with no window function in the F_1 dimension. Long relaxation times of this small molecule, together with scaling up ($\kappa = 4$) of the *J*-evolution meant that the coupling constants could be extracted by a peak picking and not by manipulation of in- and antiphase doublets. The only exception was the C3C5 cross peak which is mediated by a small coupling constant. This is a CH-CH_2 correlation observed on the CH proton, which together with a faster relaxation of the CH_2 carbon contributed to the lower intensity of this cross peak.

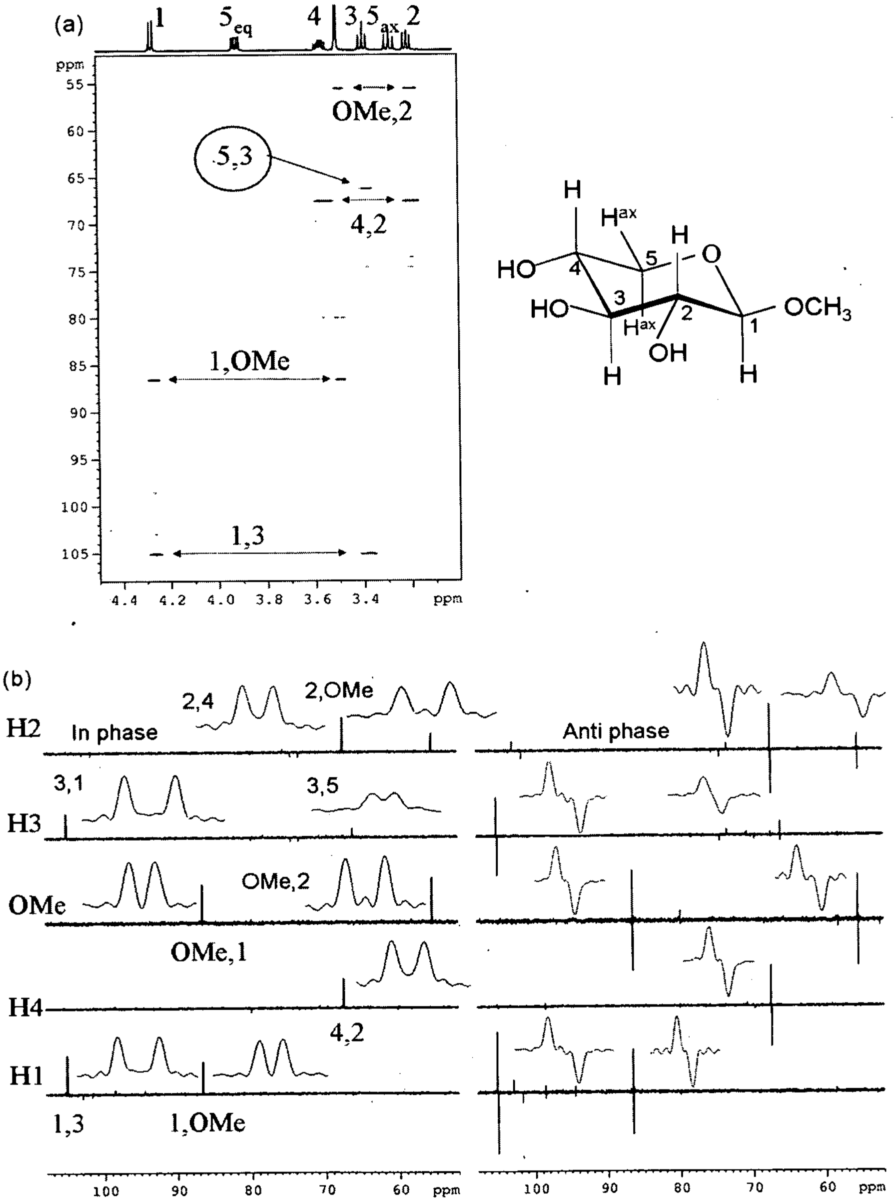


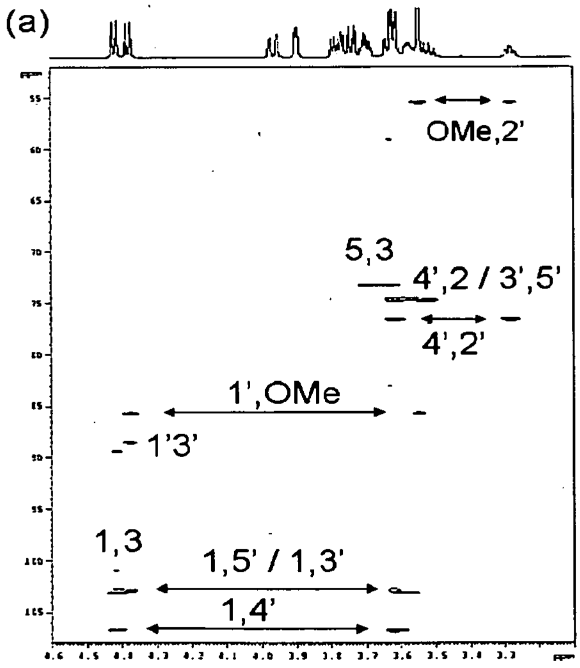
Fig.III.3.3.3 (a) 2D *J*M-DEPT-INADEQUATE spectrum of Me-β-D-xylopyranoside and (b) F_1 traces of long-range doublets from both in- and antiphase spectra

Table III.3.3.1 shows the long-range ${}^{n}J_{CC}$ coupling constants directly extracted from the in- and antiphase traces (excluding the middle number of the ${}^{n}J_{C3C5}$ coupling constant (1.9 Hz), which was from the manipulation of in- and antiphase doublets) in Fig.III.3.3.3. These ${}^{n}J_{CC}$ coupling constants are very similar to those determined by the two previously discussed methods (Table III.3.1.1 and Table III.3.2.3).

Table III.3.3.1 ${}^{n}J_{CC}$ coupling constants, in Hz, of Me-β-D-xylopyranoside from 2D *J*M-DEPT-INADEQUATE in- and antiphase spectra

J_{CxCy}	C_1C_3	C_2C_4	C_3C_5	C_1OMe	C_2OMe
In-phase	4.2 4.1	2.8 2.7	- - 1.9	2.1 2.3	3.2 3.2
Antiphase	4.2 4.2	2.8 2.7	1.8 2.1	2.1 2.3	3.2 3.2

The method was next tested on a more realistic sample of Me-β-D-lactoside. The antiphase *J*M-DEPT-INADEQUATE spectrum is shown in Fig.III.3.3.4 together with the F_1 traces extracted from both in- and antiphase spectra for inter-ring correlations. It can be seen that the signal-to-noise ratios are much worse (two times on average) in the in-phase spectra. This is due to the extra refocusing interval used in the in-phase experiment. Shorter relaxation times of Me-β-D-lactoside also meant that the coupling constants cannot be determined any more by a simple peak picking.



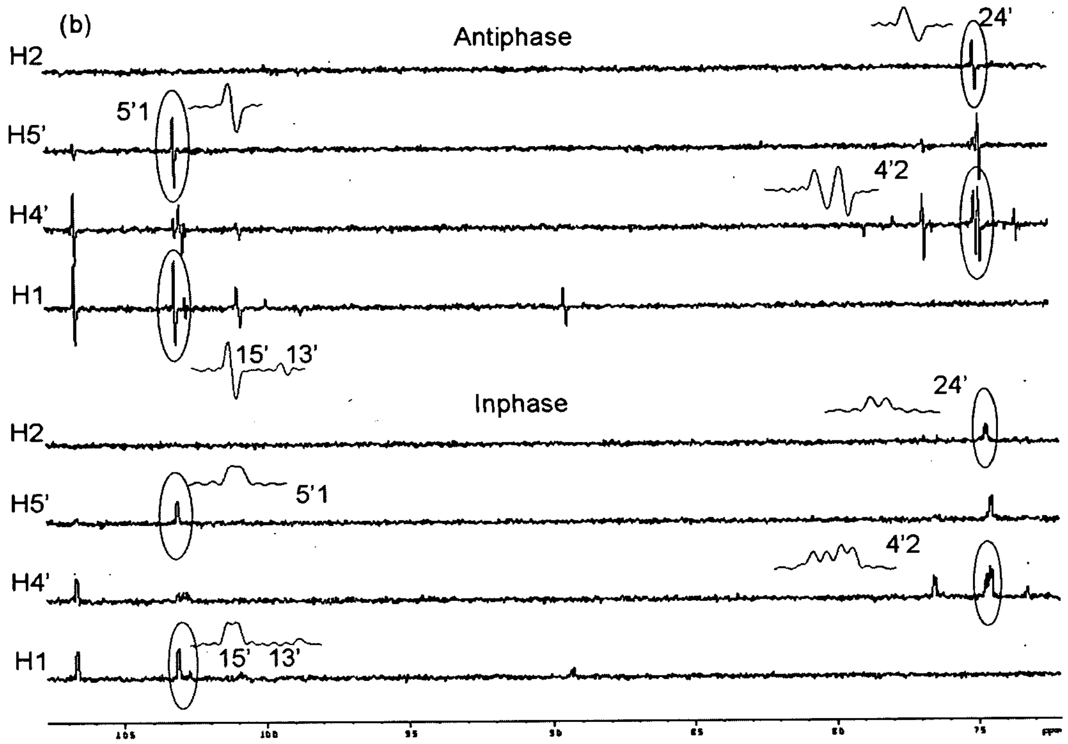


Fig.III.3.3.4 (a) 2D JM-DEPT-INADEQUATE spectrum of Me-β-D-lactoside and (b) F_1 traces of long-range doublets from both in- and antiphase spectra showing inter-ring correlations.

The long-range $^nJ_{CC}$ coupling constants (Table III.3.3.2) obtained by the manipulation of in- and antiphase doublets from the spectra in Fig.III.3.3.4 are similar to those determined by the DJM-REVINEPT-INADEQUATE method (Table III.3.2.4). Small coupling constants < 2.5 Hz gave slightly smaller values in the JM-DEPT-INADEQUATE spectrum. This indicates overestimation of small coupling constants by the DJM-REVINEPT-INADEQUATE method.

Table III.3.3.2 $^nJ_{CC}$ coupling constants, in Hz, of Me-β-D-lactoside from 2D JM-DEPT-INADEQUATE in- and antiphase spectra

C_1C_2	C_2C_3	C_3C_4	C_4C_5	$C_{1'}C_{2'}$	$C_{2'}C_{3'}$	$C_{3'}C_{4'}$	$C_{4'}C_{5'}$	$C_{1'}OMe$	$C_{2'}OMe$
-	-	-	-	-	-	-	-	2.1 2.1	3.4 3.1
C_1C_3	C_3C_5	$C_{1'}C_{3'}$	$C_{2'}C_{4'}$	$C_{3'}C_{5'}$	$C_1C_{4'}$	$C_1C_{3'}$	$C_1C_{5'}$	$C_2C_{4'}$	
5.1 4.9	1.7 1.6	4.6 -	- 2.4	2.5	2.0 2.0	~0	2.0 1.9	3.1 3.1	

The first row gives values determined from HxCxCy cross-peaks, while the second corresponds to the HyCyCx cross-peaks

III. 3.4 REVINEPT-INADEQUATE:

The *J*M-DEPT-INADEQUATE technique presented in the previous chapter performs well only for CH carbons. Taking into account our experience with the conversion of D*J*M-INEPT-INADEQUATE into the D*J*M-REVINEPT-INADEQUATE pulse sequence, a reversed INEPT-based method for acquisition of in- and antiphase J_{CC} doublets is explored next. Before proceeding with this, the sensitivity of the basic INEPT-INADEQUATE is compared with that of REVINEPT-INADEQUATE.

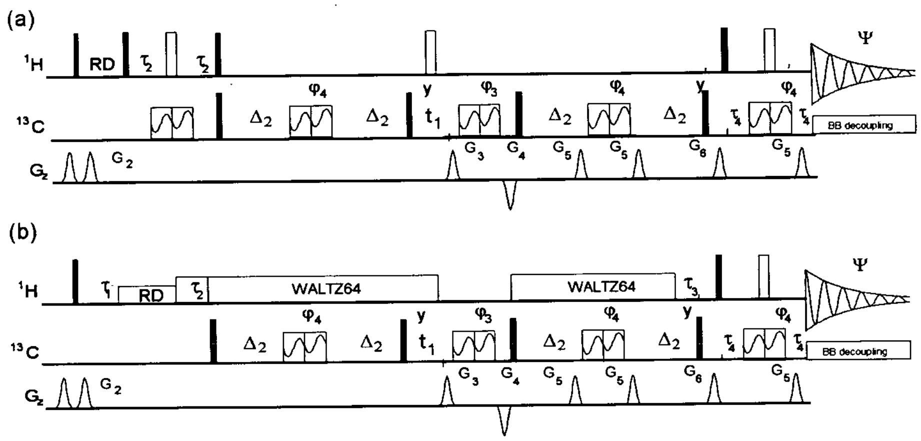


Fig.III.3.4.1 Pulse sequences of INEPT-INADEQUATE (a) and REVINEPT-INADEQUATE (b). Thin closed and thick open bars represent 90° and 180° rectangular pulses.

In the ^1H-detected INEPT-INADEQUATE experiment [103] (Fig.III.3.4.1), the magnetization starts on ^1H and is transferred to ^{13}C where after the creation of the antiphase carbon-carbon magnetization, the desired double-quantum coherences are selected. Subsequently, the J_{CC} is refocused and the magnetization is transferred to protons for detection. In the reversed INEPT method, the magnetization starts directly from ^{13}C. The ^{13}C magnetization is enhanced by heteronuclear NOE, which can yield up to a three-fold gain in sensitivity. ^1H decoupling applied throughout the pulse sequence ensures that DQ coherences are efficiently converted into observable magnetization. On the other hand, the acquisition of all carbon multiplets in the same experiment requires a compromise setting of a defocusing delay, τ_3, which leads to some loss of signal. A theoretical comparison, neglecting relaxation, given in Table III.3.2.1, suggests that the sensitivity of both methods is comparable.

The INEPT-INADEQUATE and REVINEPT-INADEQUATE spectra of Me-β-D-xylopyranoside are shown in Fig.III.3.4.2. These experiments were acquired using a $2\Delta_2$ delay of 167 ms. This value was set to be an even multiple of $1/2(^1J_{CC})$ and close to the optimum delay for $^nJ_{CC}$ of 3 Hz. As a consequence, both one-bond and long-range correlations appear in the spectra. Both spectra were acquired using the same overall experimental time of 1 hour. A compromise delay τ_3 = 2 ms was used in the REVINEPT-INADEQUATE experiment in order to see also the correlations of CH_2 and CH_3 carbons. The F_2 traces plotted with the same absolute scale for both experiments are shown in Fig.III.3.4.3 for direct comparison of the sensitivity of the two methods. Some F_2 traces through CH carbons from the REVINEPT-INADEQUATE are even more intense than those from the INEPT-INADEQUATE. This can be attributed to slower relaxation in the REVINEPT-INADEQUATE experiment during the long (334 ms) evolution delays.

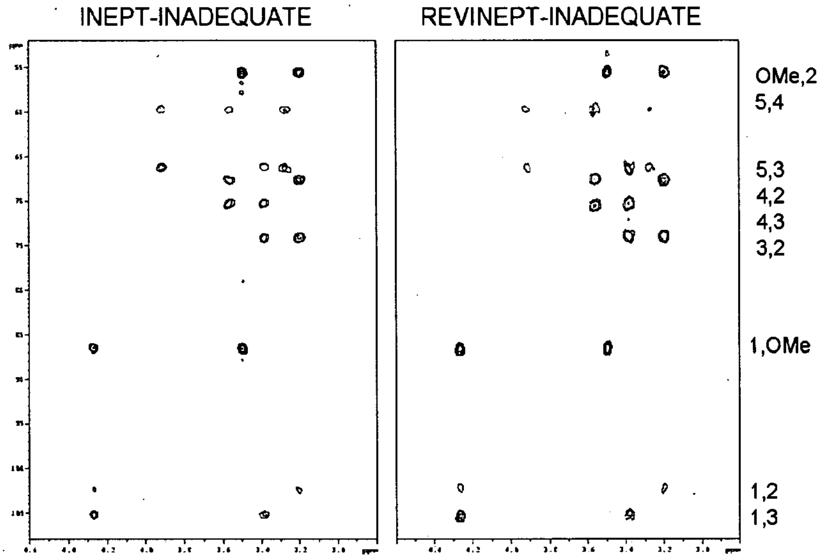


Fig.III.3.4.2 2D INEPT-INADEQUATE and REVINEPT-INADEQUATE spectra of Me-β-D-xylopyranoside

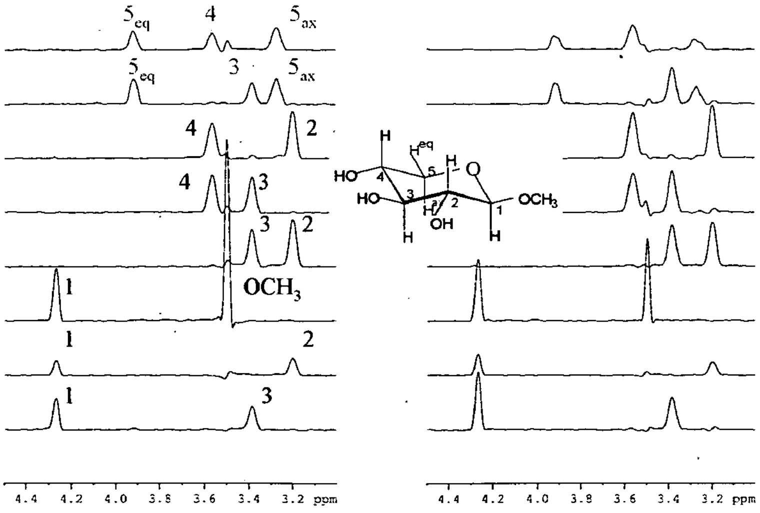


Fig.III.3.4.3 (a) F_2 dimension proton traces from the 2D INEPT-INADEQUATE and (b) REVINEPT-INADEQUATE spectra of Me-β-D-xylopyranoside.

The 2D REVINEPT-INADEQUATE experiment can be optimized to provide simultaneously both one-bond and long-range correlations (Fig.III.3.4.4a). Both types of correlations can be distinguished as the fast evolving one-bond couplings are very sensitive to the evolution delays, while the long-range couplings are not. Acquisition of two spectra optimized for long-range couplings but differing by $0.25/{}^1J_{CC}$ in setting of the carbon-carbon evolution interval will avoid a zero-point for one-bond correlations at least in one of the spectra. Subtraction of the two spectra yields a spectrum with only one-bond cross peaks, albeit with reduced intensities. These can be used in the first instance to assign the carbon skeleton of the molecule supplemented by long-range correlations. This procedure is illustrated in Fig.III.3.4.4 using the REVINEPT-INADEQUATE spectra of Me-β-D-lactoside.

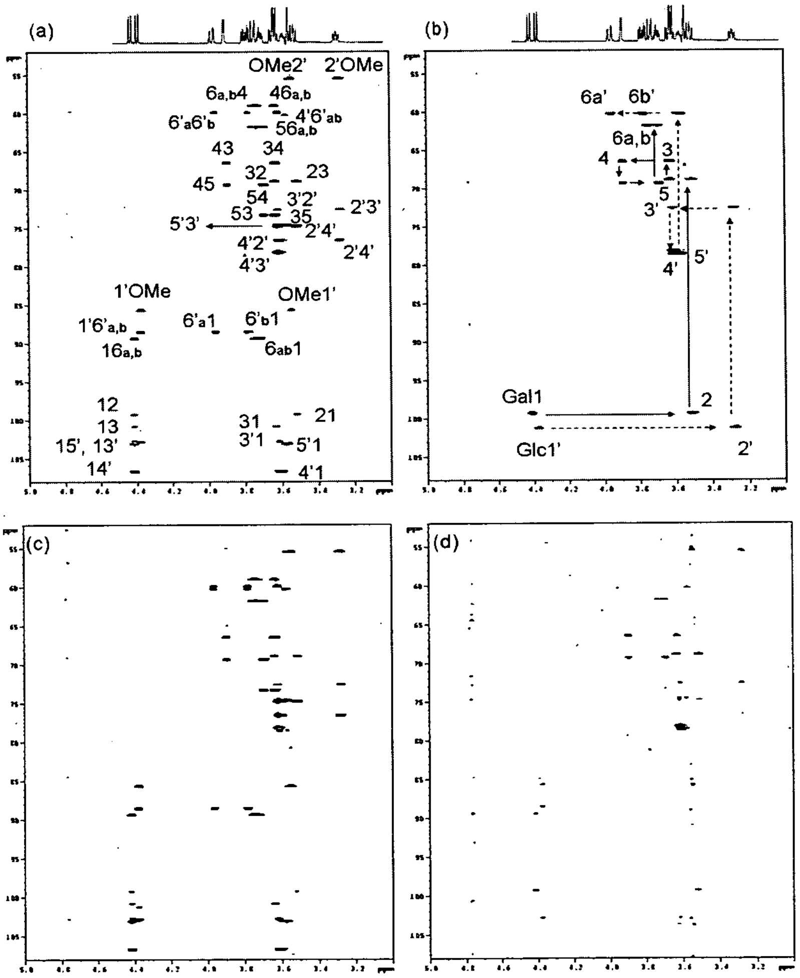


Fig.III.3.4.4 2D REVINEPT-INADEQUATE spectra of Me-β-D-lactoside (a) and (b) 2D spectra optimized for long-range and one-bond J_{CC} couplings, respectively. (c) Addition of the two long-range correlation spectra acquired with different delays ($2\Delta_2$, $2\Delta_2 + 0.25/{}^1J_{CC}$). (d) Subtraction of the two spectra yielding one-bond multiplets only.

In summary, REVINEPT-INADEQUATE is a sensitive experiment suitable for conversion into a *J*-modulated experiment for the measurement of $^{n}J_{CC}$ coupling constants.

III. 3.5 *J*M-REVINEPT-INADEQUATE:

The REVINEPT-INADEQUATE method can be easily adapted into a *J*-modulated experiment using the building blocks introduced in the discussion of *J*M-DEPT-INADEQUATE. Compared with the original *J*M-INEPT-INADEQUATE method [101], *J*M-REVINEPT-INADEQUATE (Fig.III.3.5.1) utilizes the carbon polarization as in all previous reversed experiments. The first fixed carbon-carbon coupling evolution interval ($0.5/^{n}J_{CC}$) is converted into a *J*-modulated interval.

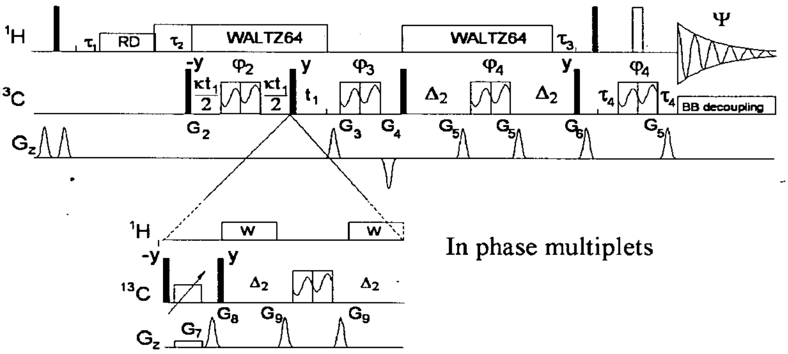


Fig.III.3.5.1 Pulse sequence of *J*M-REVINEPT-INADEQUATE. Thin closed and thick open bars represent 90° and 180° rectangular pulses.

The in- and antiphase spectra are acquired in the interleaved manner as was the case for *J*M-DEPT-INADEQUATE. The in-phase and antiphase 2D *J*M-REVINEPT-INADEQUATE spectra of Me-β-D-xylopyranoside are shown in Fig.III.3.5.2. All doublets, except for C_3C_5, were sufficiently resolved in both spectra to allow the coupling constants (Table III.3.5.1) to be determined by simply peak picking.

Table III.3.5.1 $^{1,n}J_{CC}$ coupling constants, in Hz, of Me-β-D-xylopyranoside from spectrum of 2D *J*M-REVINEPT-INADEQUATE in Fig. III.3.5.2

J_{CxCy}	C_1C_2	C_2C_3	C_3C_4	C_4C_5	C_1C_3	C_2C_4	C_3C_5	C_1OMe	C_2OMe
In-phase	46.8 46.7	38.9 38.9	39.0 39.0	40.0 -	4.2 4.3	2.8 2.8	1.7 - 1.9* 1.9*$_{eq}$	2.1 2.1	3.1 3.1
Antiphase	46.8 46.8	38.8 38.9	39.1 39.2	40.0 40.1^{eq}	4.1 4.2	2.7 2.7	2.1 2.0$_{eq}$	2.3 2.3	3.2 3.2

* determined by the manipulation of the in- and antiphase multiplets.

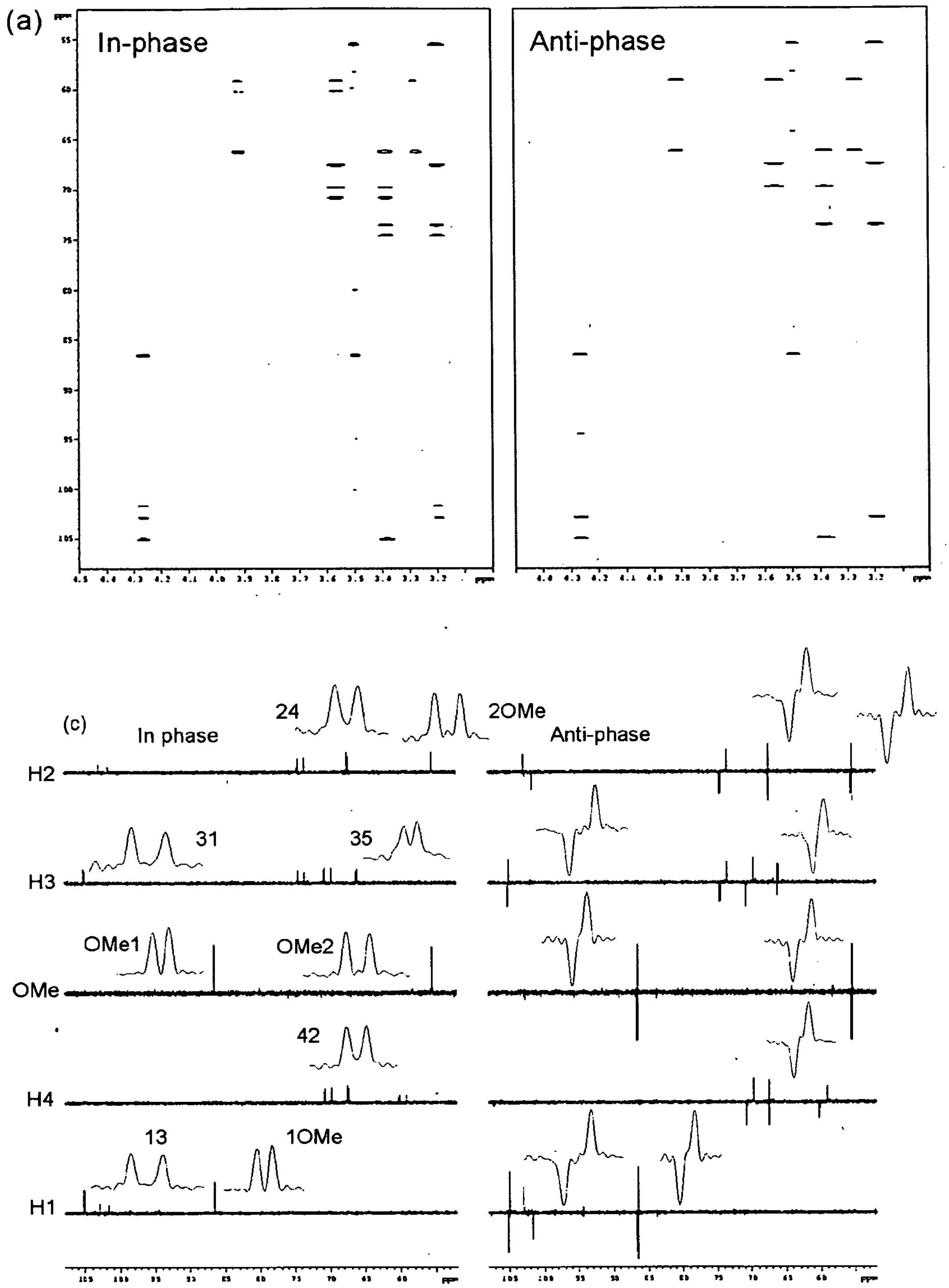


Fig.III.3.5.2 (a) 2D *J*M-REVINEPT-INADEQUATE in- and antiphase spectra of Me-β-D-xylopyranoside. (b) In/antiphase F_1 traces from the 2D spectra.

The C3C5 multiplet is not well resolved. The J_{C3C5} coupling constant determined from the in-phase spectrum is much smaller compared with the result obtained from the antiphase spectrum. The coupling constants determined in this experiment are practically identical with the coupling constants extracted from JM-DEPT-INADEQUATE spectra of Me-β-D-xylopyranoside.

The JM-REVINEPT-INADEQUATE experiment was further tested using the sample of Me-β-D-lactoside. Both in- and antiphase spectra were acquired in an interleaved manner. The in-phase spectrum is shown in Fig.III.3.5.3 and the traces from both spectra through carbons around the glycosidic linkage are shown in Fig.III.3.5.4. Table III.3.5.2 shows the coupling constants determined by the manipulation of in- and antiphase multiplets.

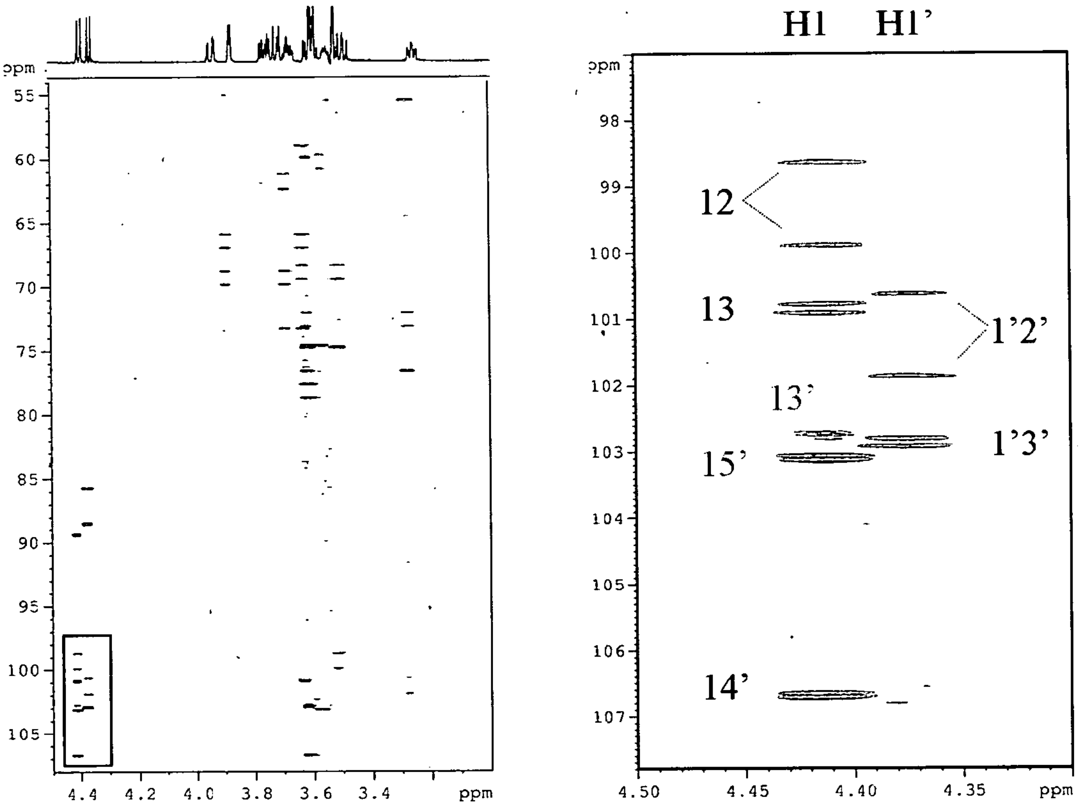


Fig.III.3.5.3 2D JM-REVINEPT-INADEQUATE in-phase spectrum of Me-β-D-lactoside

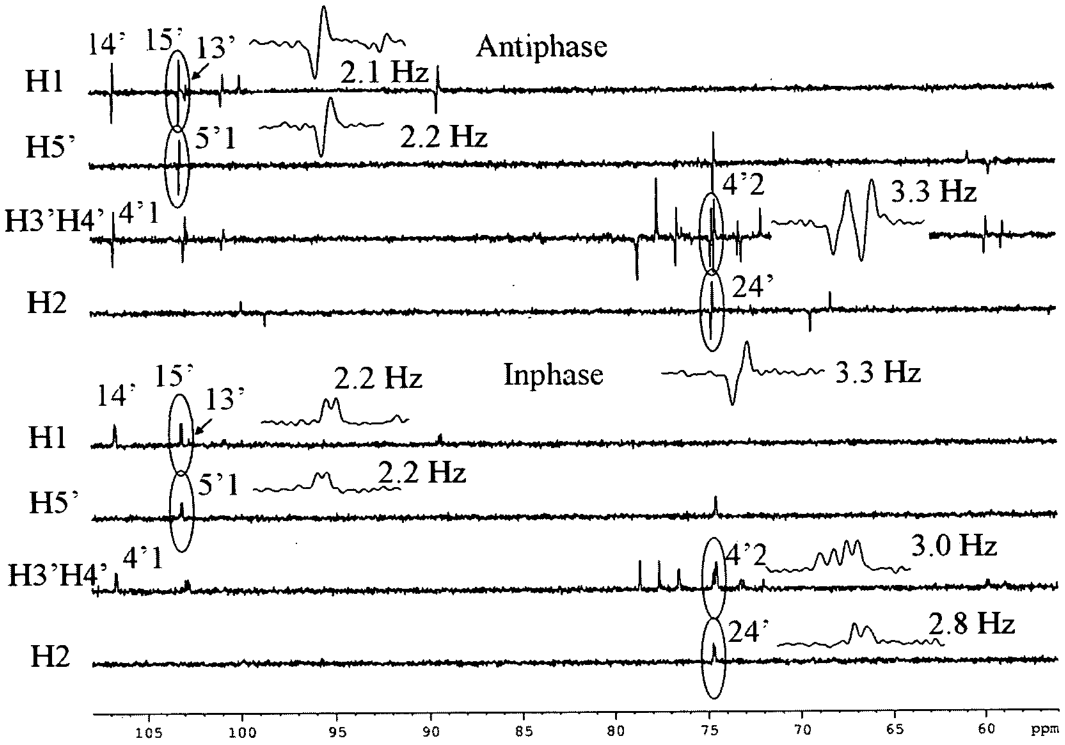


Fig.III.3.5.4 F_1 traces across the glycosidic linkage of Me-β-D-lactoside from 2D spectrum of *J*M-REVINEPT-INADEQUATE

Table III.3.5.2 $^{1,n}J_{CC}$ coupling constants, in Hz, of Me-β-D-lactoside from 2D *J*M-REVINEPT-INADEQUATE spectrum

C_1C_2	C_2C_3	C_3C_4	C_4C_5	$C_{1'}C_{2'}$	$C_{2'}C_{3'}$	$C_{3'}C_{4'}$	$C_{4'}C_{5'}$	$C_{1'}OMe$	$C_{2'}OMe$
47.0	40.2	39.2	38.6	47.1	39.9	39.2	-	2.1	3.2
47.0	40.4	38.9	38.7	-	39.8	39.2		-	-
C_1C_3	C_3C_5	$C_{1'}C_{3'}$	$C_{2'}C_{4'}$	$C_{3'}C_{5'}$	$C_1C_{4'}$	$C_1C_{3'}$	$C_1C_{5'}$	$C_2C_{4'}$	
5.2	1.5	4.5	2.4	2.5	2.0	~0	2.1	3.1	
5.2	1.7	4.6	2.4	2.5	2.0		2.1	3.2	

The first row gives values determined from HxCxCy cross-peaks, while the second corresponds to the HyCyCx cross-peaks

The coupling constants summarized in Table III.3.5.2 are practically identical to those determined by the *J*M-DEPT-INADEQUATE (Table III.3.3.2).

III. 3.6 ^{13}C-detected INADEQUATE:

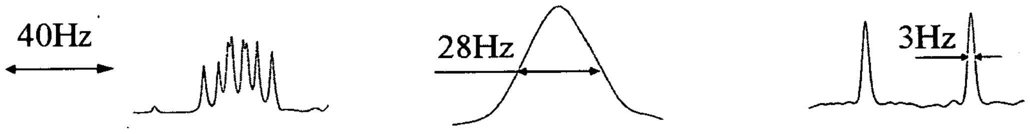
Next we turn our attention to ^{13}C-detected INADEQUATE. At a first glance it may seem that the gap between the sensitivity of the INEPT-INADEQUATE polarization transfer pathway (proton-carbon-proton) and that of ^{13}C-detected INADEQUATE (carbon-carbon polarization transfer pathway) is too big for the latter experiment to be worth contemplating. However, a more detailed analysis reveals that this is not the case and in practice, the ^{13}C-detected INADEQUATE is a very competitive alternative to the ^{1}H-detected INADEQUATE. Despite the low initial relative sensitivity of the ^{13}C-detected INADEQUATE of 1/32, when other factors are taken into account a more realistic picture emerges (Table III.3.6.1).

Table III.3.6.1 Sensitivity of ^{1}H- and ^{13}C-detected INADEQUATE experiments

INEPT-INADEQUATE	$32 \times 0.5 \times 0.71 = 11.3$
^{13}C-detected INADEQUATE	$1 \times 3 = 3$

Here the loss of half of the magnetization, due to the leakage to other protonated carbons (0.5) and the loss associated with gradient selection (0.7), is considered for INEPT-INADEQUATE. The stability of today's spectrometers does not require gradients for the signal selection in the ^{13}C-detected INADEQUATE and therefore phase cycling [111] is the preferred option. This also eliminates diffusion-related losses associated with gradient selection. Considering heteronuclear NOE for the ^{13}C-detected method, the relative sensitivity of the two experiments is ~ 1/4.

In ^{1}H-detected INADEQUATE the signal is detected on nuclei that show extensive splitting of their signals due to *J* couplings. Acquisition is performed in the presence of ^{13}C decoupling which limits the acquisition times considerably and a value of ~100 ms is typical. When using appropriate window functions, the signals in the ^{1}H dimension are broad. On the other hand, acquisition in ^{13}C-detected INADEQUATE is performed with ^{1}H decoupling, which allows longer acquisition times. The signal is only a doublet and can be properly sampled. A comparison of typical multiplets obtained in both experiments (Fig.III.3.6.1) clearly illustrates this difference.



1D ^{1}H　INEPT-INADEQUATE　^{13}C detection

Fig.III.3.6.1 Signals from ^{1}H- and ^{13}C-detected INADEQUATE spectra

The above analysis shows that when the width of signals is taken into account the sensitivity of both experiments is comparable. The other factor that has to be taken into account when deciding which experiment to use is the available equipment. The availability of cryoprobes means that the range of NMR sensitivities has widened as illustrated in Table III.3.6.2.

Table III.3.6.2 Sensitivity of probes at 600 MHz (0.1%v EtBz)

Probes	^{1}H Sensitivity	^{13}C
Triple resonance inverse probe	1000:1	
Triple resonance inverse cryoprobe	5000:1	
^{13}C direct detection probe		200:1
^{13}C Triple resonance inverse cryoprobe		400:1
Inverse cryoprobe with a cold ^{13}C preamplifier		800:1
^{13}C direct detection cryoprobe		1600:1

Capitalizing on the know-how acquired during the development of the ^{1}H-detected INADEQUATE experiments, the ^{13}C-detected experiment was designed as to provide (i) simultaneous one-bond and long-range correlations, and (ii) accurate values of all J_{CC} coupling constants including the small $^{n}J_{CC}$ coupling constants, in a single measurement.

The original non-refocused INADEQUATE experiment [98] with a modified phase-cycling [111] and the preparation period, τ, optimized for long-range correlations (Fig.III.3.6.2), is used for the acquisition of antiphase ^{13}C doublets in the F_2 dimension.

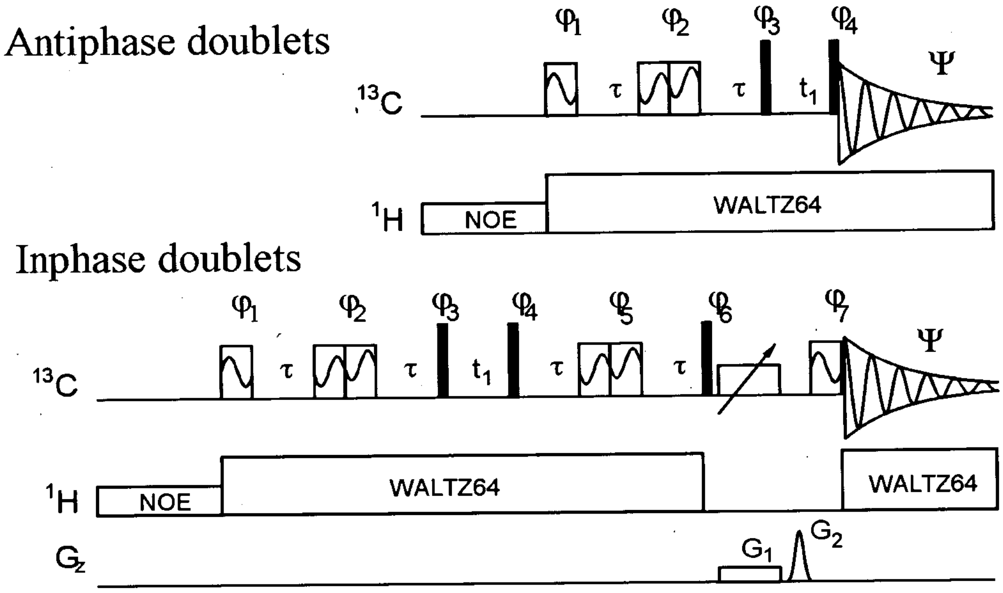


Fig.III.3.6.2 ^{13}C-detected, phase cycled (a) non-refocused and (b) refocused INADEQUATE. $\tau = 0.25/nJ_{CC}$ or $0.25/nJ_{CC} + 0.5/{}^{1}J_{CC}$. The phase cycling according to [111] was applied; 90° and 180° BEBOP pulses [112] were applied where appropriate. A 20 ms adiabatic pulse [92] was applied in (b) simultaneously with a 2.1 Gauss/cm pulsed field gradient (G_1) followed by a purging gradient G_2=23.8 G/cm.

Two experiments with $\tau = 0.5/^{n}J_{CC}$ and $\tau = 0.5/^{n}J_{CC} + 0.5/^{1}J_{CC}$ are acquired in an interleaved manner. Neglecting the relaxation effects, data obtained by addition or subtraction of the two spectra will always contain one-bond cross-peaks, albeit with a reduced intensity (at most by a factor 0.5-0.71 compared with an experiment optimized for $^{1}J_{CC}$). Variations of the evolution interval will only marginally affect the intensities of long-range cross peaks. The readout of $^{1}J_{CC}$ values from these spectra is trivial due to the large separation of the two components of the one-bond antiphase doublets. This is, however, not the case for the long-range cross peaks, as partial cancellation of closely spaced antiphase lines produces an apparent splitting discussed previously. Similarly to ^{1}H-detected *J*M-REVINEPT-INADEQUATE and *J*M-DEPT-INADEQUATE, in order to determine the $^{n}J_{CC}$ coupling constants the in-phase multiplets must also be acquired. Editing of the in- and antiphase doublets by using the correct intensity ratio of these multiplets leads to accurate coupling constants.

Two additional INADEQUATE experiments ($\tau = 0.5/\,^{n}J_{CC}$ and $\tau = 0.5/\,^{n}J_{CC} + 0.5/\,^{1}J_{CC}$), which include a refocusing period followed by an efficient purging of the residual antiphase component (Fig.III.3.6.2b) are therefore acquired. The purging is achieved by a z-filter with simultaneous suppression of DQ/ZQ coherences [92] placed after the refocusing interval. Two non-refocused and two refocused INADEQUATE spectra that contain both one-bond and long-range cross peaks are acquired in an interleaved manner. The scaling factor is calculated using the relevant transfer functions and the effective ^{13}C spin-spin relaxation times are determined in a separate experiment. Spectral editing of these four spectra, which may also utilize the symmetry of INADEQUATE spectra in the directly detected dimension, leads to accurate determination of $^{n}J_{CC}$ coupling constants.

The new experiment is illustrated for accurate measurement of scalar and dipolar carbon-carbon coupling constants of carbohydrates. ^{13}C-detected INADEQUATE spectra of Me-β-D-xylopyranoside were acquired at 600 MHz using an inverse detected cryoprobe with a cold ^{13}C preamplifier (University of Glasgow). The long-range evolution delay was optimized to an odd multiple of $0.5/^{1}J_{CC}$, which was known from the previous experiments. Therefore in this instance only two spectra were acquired rather than four. The 2D INADEQUATE spectrum is shown in Fig.III.3.6.3. Also shown are F_2 traces of carbons C2 to C5 taken from the refocused and non-refocused spectra. The determined coupling constants are listed in Table III.3.6.3.

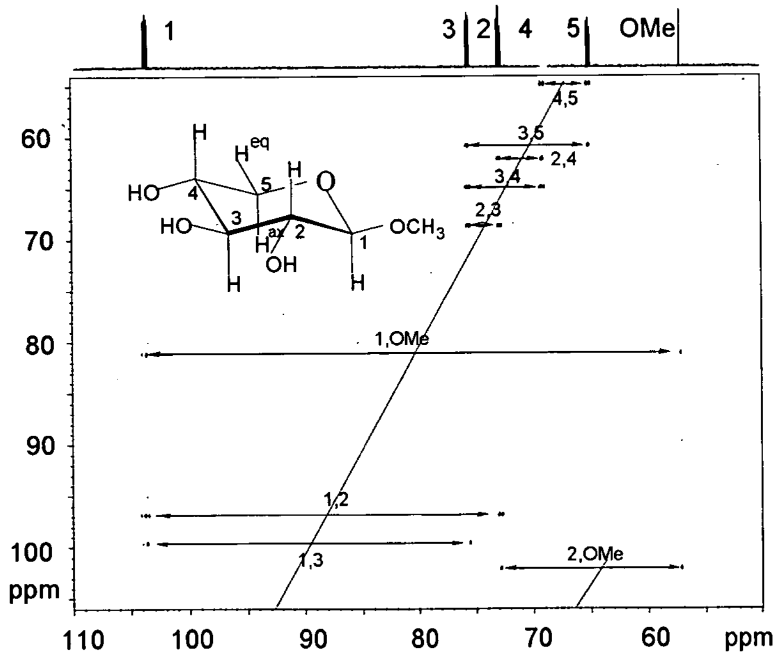
Table III.3.6.3 Coupling constants of Me-β-D-xylopyranoside, in Hz, determined from in- and antiphase spectra of isotropic (Fig.III.3.6.3) and aligned (not shown) samples

	C_1C_2	C_2C_3	C_3C_4	C_4C_5	C_1C_3	C_2C_4	C_3C_5	C_1OMe	C_2OMe
J	46.8	38.7	39.0	39.8	4.2	2.7	2.0	2.2	3.1
$J+D$	50.6	36.7	38.5	43.5	4.6	1.8	2.8	1.8	3.3
D_{expt}	3.8	-2.0	-0.5	3.6	0.4	-0.9	0.8	-0.3	0.2
$D_{cal}{}^{a}$	3.8	-2.0	-0.8	3.5	0.7	-0.7	1.0	0.3	0.1
$D_{cal}{}^{b}$	3.8	-2.0	-0.4	3.6	0.3	-0.8	0.6	0.3	0.2

	C_1H_1	C_2H_2	C_3H_3	C_4H_4	C_5H_{5a}	C_5H_{5b}
J	161.6	144.7	141.6	145.0	141.6	151.2
$J+D$	151.1	134.8	133.5	134.6	130.8	149.9
D_{expt}	-10.5	-9.9	-8.1	-10.4	-10.8	-1.4
$D_{cal}{}^{a}$	-10.6	-10.2	-8.4	-10.5	-10.1	-1.5
$D_{cal}{}^{b}$	-10.6	-9.9	-8.0	-10.6	-10.4	-1.4

[a] from ab initio structure; Parameters of the alignment tensor are: $S_{x'x'}$ = (2.53± 0.06)e-04, $S_{y'y'}$ = (9.67 ± 0.16)e-04, $S_{z'z'}$ = (-1.22 ± 0.02)e-03. The Euler angles in the coordinate system of the ab initio structure of Me-β-D-xylopyranoside: α = -48.6 ± 0.2, β = -280.1 ± 0.5 and γ = -5.4 ± 1.0 degrees.

[b] from solution structure [89a] Parameters of the alignment tensor are: $S_{x'x'}$ = (2.43± 0.02)e-04, $S_{y'y'}$ (9.09 ± 0.13)e-04, $S_{z'z'}$ = (-1.15 ± 0.01)e-03. The Euler angles in the coordinate system of the solution structure of Me-β-D-xylopyranoside: α = 260.0 ± 0.3, β = 128.3 ± 0.2 and γ = 85.9 ± 1.1 degrees. Please note that the two structures are in different starting molecular coordinate systems. Because of that the Euler angles are very different between the two.



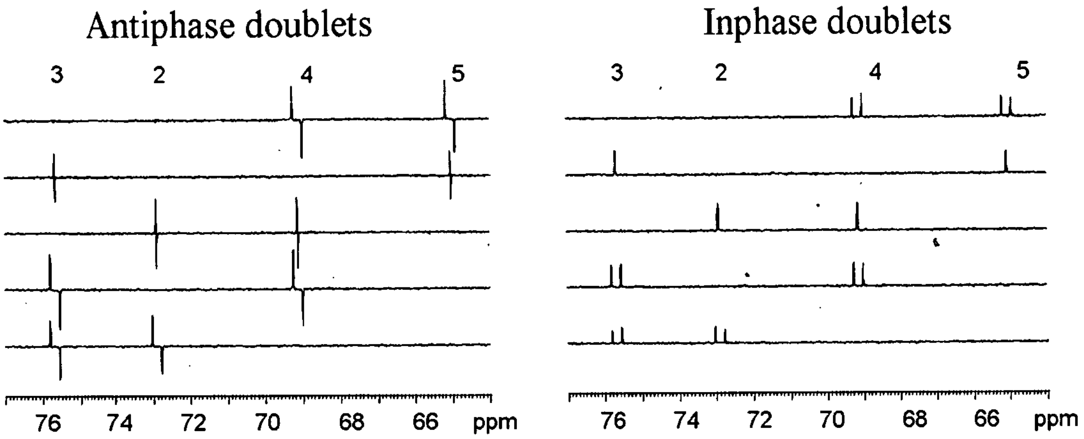


Fig.III.3.6.3 2D ^{13}C-detected INADEQUATE spectrum and the traces from the non-refocused/ refocused 2D INADEQUATE spectra.

The second example illustrates the potential of the refocused/non-refocused INADEQUATE in studying compounds in aligned media. Providing a sufficient amount of compound is available, the use of D_{CC} coupling constants for structural investigation of small organic molecules in anisotropic media is a very attractive proposition. Many media proposed for organic solvents thus far tend to align molecules rather strongly. Consequently, higher order effects prevent accurate measurement of D_{HH} and $^{1}D_{CH}$ coupling constants. Since only two ^{13}C spins in each molecule are selected by the INADEQUATE pulse sequence, the extraction of numerous carbon-carbon RDCs is no different from the measurement of $^{1,n}J_{CC}$ coupling constants. The ^{1}H-detected INADEQUATE is not particularly suited for the measurement of carbon-carbon RDCs because of the broadening of ^{1}H multiplets by numerous proton-proton RDCs in strongly aligned media.

An aligned sample of Me-β-D-xylopyranoside was prepared yielding the residual quadrupolar splitting of D_2O of 138.8 Hz. A CPMG spectrum of the aligned sample is compared with the ^{1}H spectrum of the isotropic sample in Fig.III.3.6.4, where a considerable broadening of ^{1}H resonance is noticeable. ^{13}C-detected INADEQUATE spectra of the aligned sample were of similar quality to the spectra acquired on the isotropic sample. The spectra were acquired using exactly the same conditions as for the non-aligned sample. The one-bond antiphase cross peaks from both spectra were used for the determination of $^{1}J_{CC}$ RDCs including their signs (Fig.III.3.6.5a). The analysis of three long-range cross peaks shown in Fig.III.3.6.5b illustrates that coupling constant determination is as straightforward as for the isotropic sample.

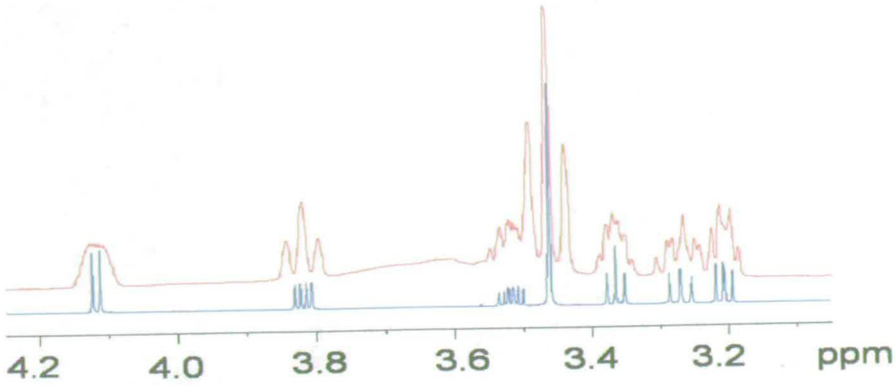


Fig.III.3.6.4 Comparison of 600 MHz ^{1}H spectra of isotropic (blue) and aligned (red) samples

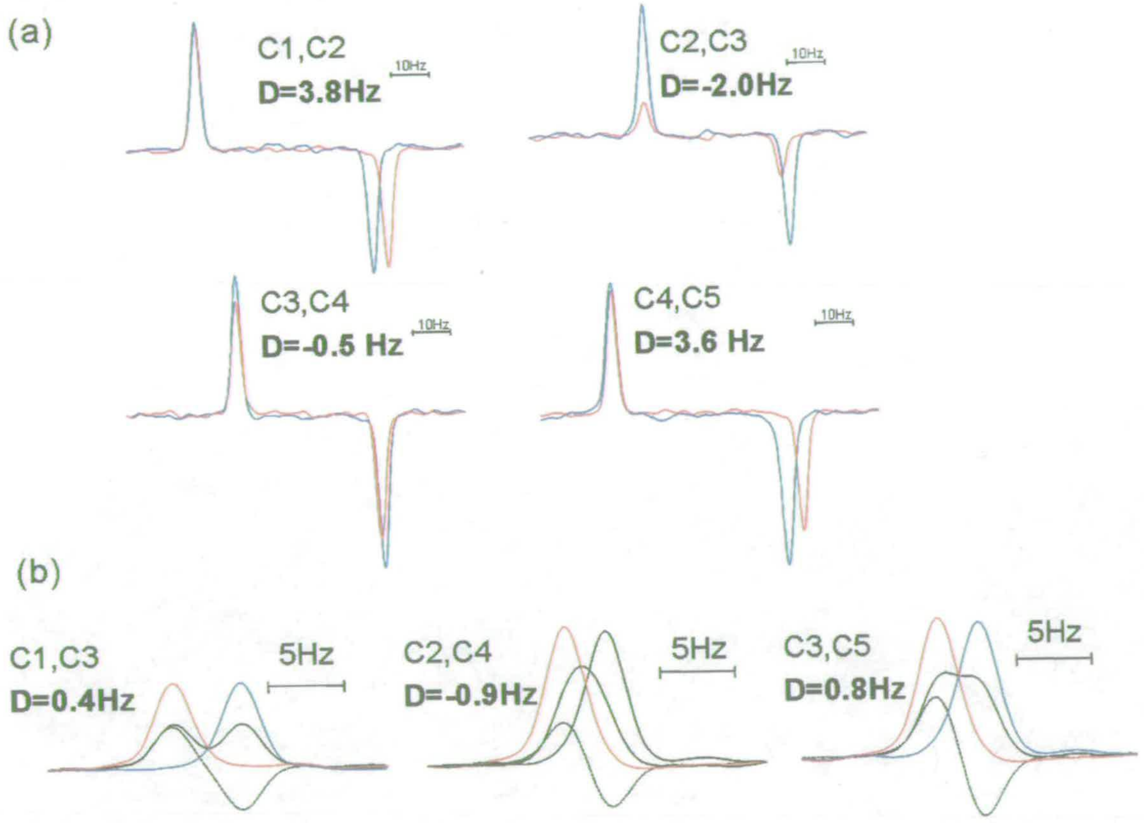


Fig.III.3.6.5 Determination of $^{1}J_{CC}$ and $^{n}J_{CC}$ coupling constants from in- and antiphase spectra of ^{13}C-detected INADEQUATE F_2 traces of Me-β-D-xylopyranoside sample

In addition to ^{13}C-^{13}C RDCs one-bond ^{1}H-^{13}C RDCs were also obtained using 1D intensity based experiments (Fig.III.3.6.6). In this experiment a BIRD[d,x] is used to remove the evolution of long-range couplings and the signal is only modulated by $^{1}J_{CH}$ coupling constants.

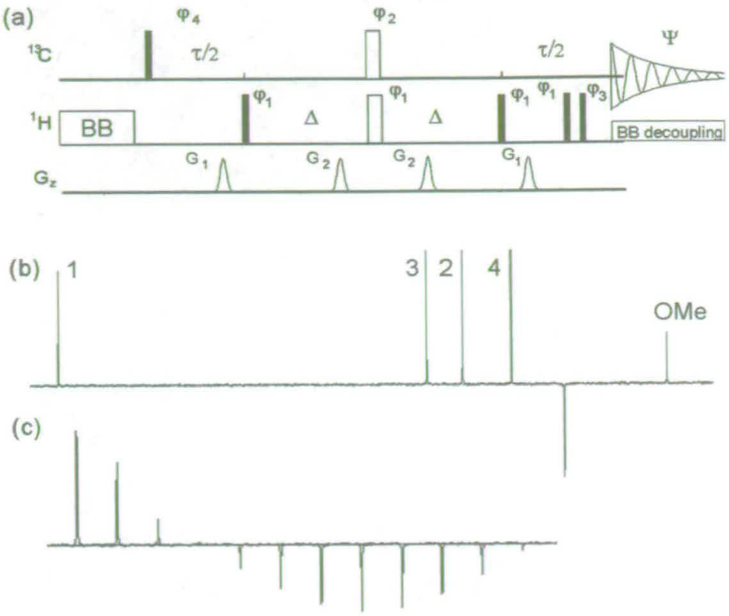


Fig.III.3.6.6 (a) Pulse sequence of the 1D intensity based experiment for the measurement of $^1J_{CH}$ coupling constant. (b) Spectrum of Me-β-D-xylopyranoside carbons at one delay. (c) Signal of carbon C1 acquired using 12 delays (1.5 – 207.2 ms). The acquisition and relaxation times of 0.45 and 3 s and 512 scans were used.

The experimental and back-calculated RDCs, using ab initio and solution [89a] structures of Me-β-D-xylopyranoside, refined previously by an extensive set of RDCs, are listed in Table III.3.6.3. The alignment tensor parameters are shown in Fig.III.3.6.7. The back-calculated RDCs from the refined solution structure give better agreement with the experimental RDCs compared to those from the ab initio structure (RMSD of 0.16 versus 0.26 Hz).

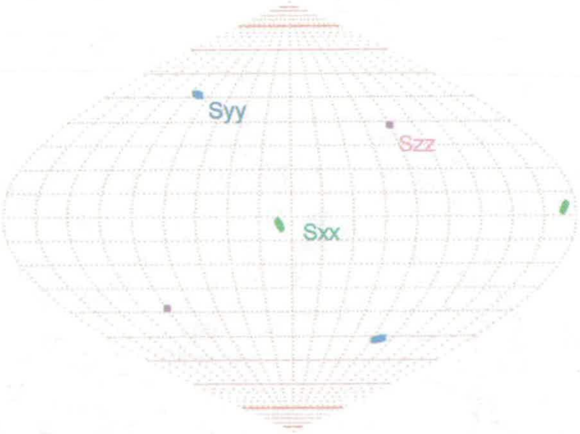


Fig.III.3.6.7 Sanson-Flamsteed projections showing the orientation of the alignment tensor principal axis system in the molecular frame of the solution structure of Me-β-D-xylopyranoside calculated using REDCAT. RDCs used for the calculation were taken from in Table III.3.6.3 (excluding C_1OMe and C_2OMe).

The third example of the ^{13}C-detected INADEQUATE comes from the measurement of $^{n}J_{CC}$ across the glycosidic linkage of a disaccharide. 25 mg of a Me-β-D-lactoside was used to record the non-refocused/refocused INADEQUATE spectra in 25 hours each (Fig.III.3.6.8). The S/N ratio of the antiphase long-range cross peaks of 7.1-14.6: 1 was obtained.

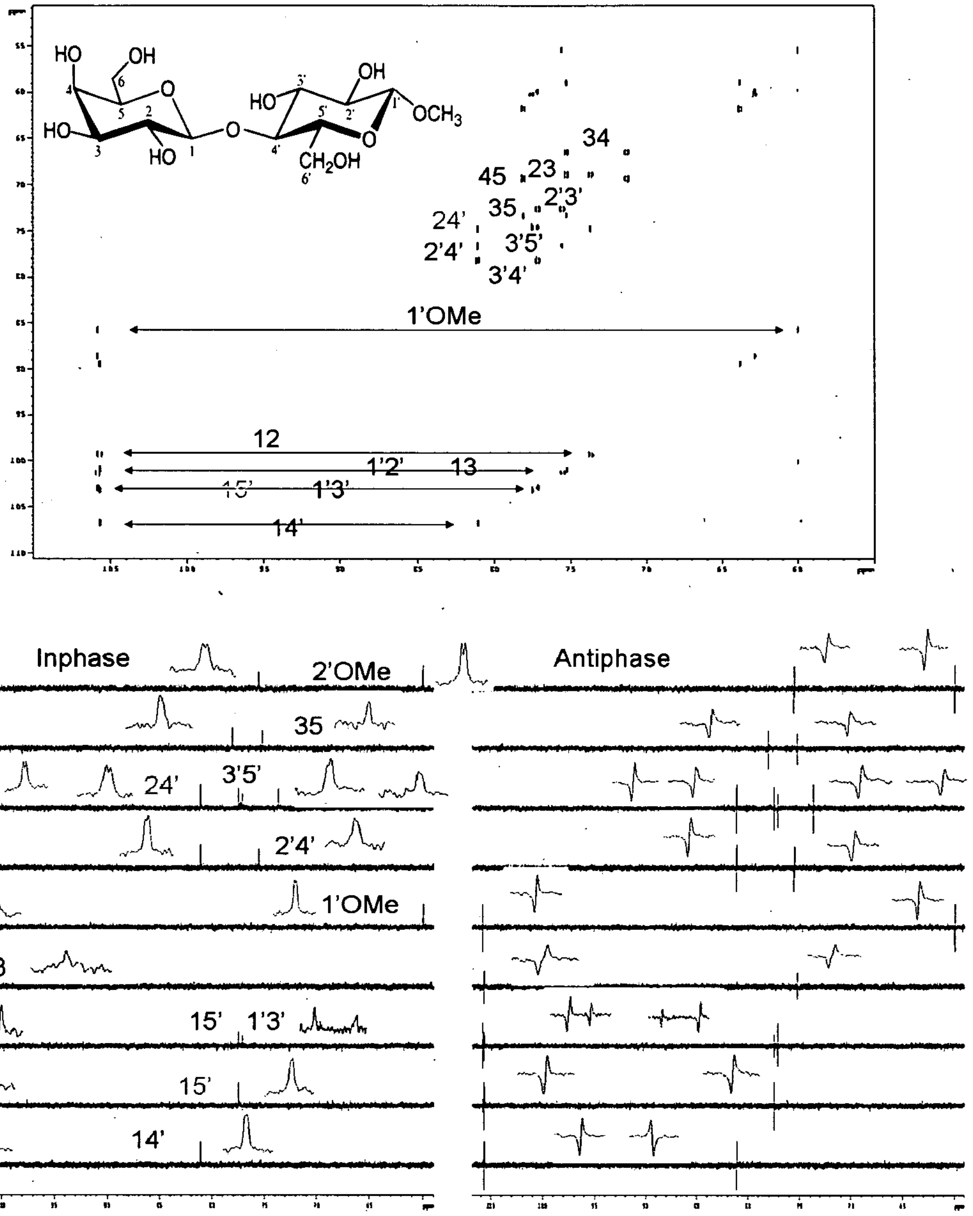


Fig.III.3.6.8 2D ^{13}C-detected INADEQUATE spectrum of Me-β-D-lactoside and traces of long-range ^{13}C-^{13}C doublets from the refocused (in-phase)/non-refocused (antiphase) 2D INADEQUATE spectra.

The carbon-carbon coupling constants determined by the manipulation of the in- and antiphase doublets are summarized in Table III.3.6.4. These coupling constants are practically identical to those determined by the ^{1}H-detected methods.

Table III.3.6.4 $^{1,n}J_{CC}$ coupling constants, in Hz, of Me-β-D-lactoside determined from the 2D ^{13}C-detected INADEQUATE spectra

C_1C_2	C_2C_3	C_3C_4	C_4C_5	$C_{1'}C_{2'}$	$C_{2'}C_{3'}$	$C_{3'}C_{4'}$	$C_{4'}C_{5'}$	$C_{1'}OMe$	$C_{2'}OMe$
46.9	39.8	38.8	38.5	44.5	47.6	39.6	-	2.0	3.0
46.6	40.1	38.9	38.5	44.5	47.1	39.8		2.0	2.9
C_1C_3	C_3C_5	$C_{1'}C_{3'}$	$C_{2'}C_{4'}$	$C_{3'}C_{5'}$	$C_1C_{4'}$	$C_1C_{3'}$	$C_1C_{5'}$	$C_2C_{4'}$	
5.1	1.4	4.8	2.4	2.4	2.0	~ 0	2.0	3.0	
5.1	1.6	4.8	2.3	2.4	2.0		1.9	3.1	

The first row gives values determined from HxCxCy cross-peaks, while the second corresponds to the HyCyCx cross-peaks

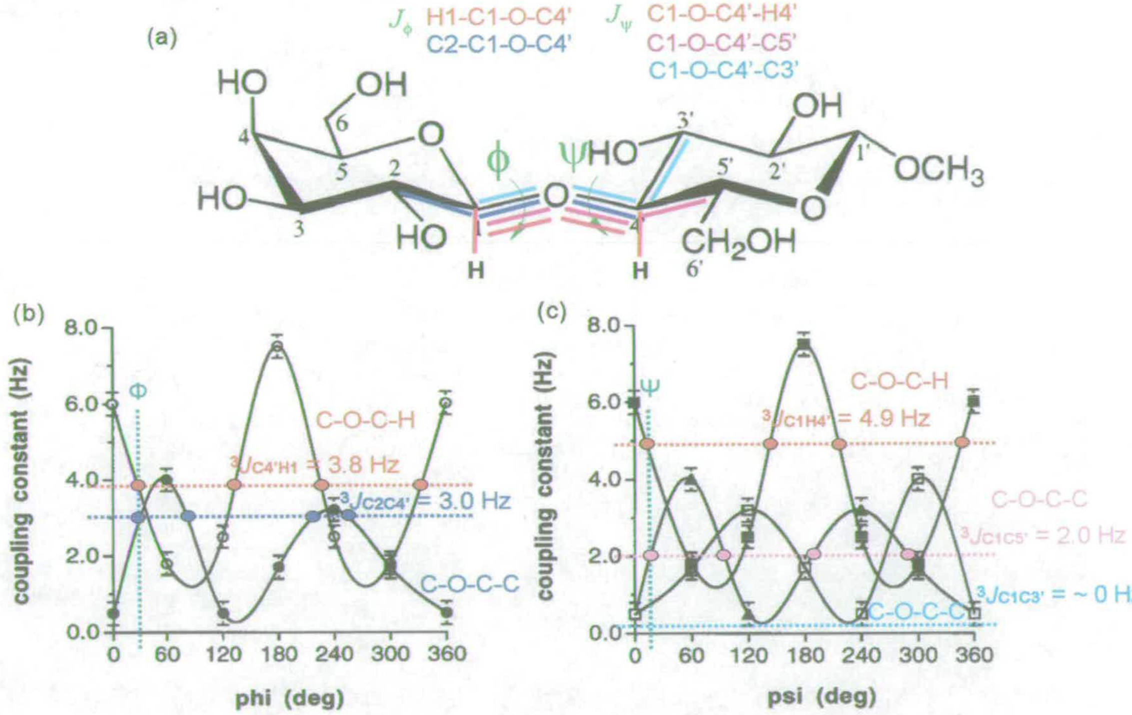


Fig.III.3.6.9 (a) Definition of the dihedral angles across the glycosidic linkage of the Me-β-D-lactoside. (b) and (c) Determination of the Φ and Ψ angles using the $^3J_{COCC}$ (results from the 2D ^{13}C-detected INADEQUATE experiment) and $^3J_{COCH}$ [99] coupling constants.

The measurement of $^3J_{COCC}$ coupling constants across glycosidic linkages reduces the ambiguity in the corresponding dihedral angles which are determined from $^3J_{COCH}$ coupling constants as shown in Fig.III.3.6.9. The $^3J_{C2C4'}$ coupling constant (Table III.3.6.4) is dependent on the Φ angle, while $^3J_{C1C3'}$ and $^3J_{C1C5'}$ reflect the Ψ angle [99]. The $^3J_{C2C4'} = 3.0$ Hz coupling constant was measured in our experiments. Using this value, the possible dihedral angles are 30, 90, 210, and 260 degrees. When considering the $^3J_{C4'H1}$ [99] coupling constant (Fig.III.3.6.9b), possible values of the Φ angle are 30, 130, 230 and 330 degrees. A common value of 30 degrees therefore represents the experimental conformation.

The intensity of the C1C3' cross peak (Fig.III.3.6.8) is very small, which suggests a close to zero coupling constant ${}^{3}J_{C1C3'}$. The C1C5' multiplets are isolated in both C1 and C5' traces. A value of 2.0 Hz for the ${}^{3}J_{C1C5'}$ coupling constant was used for the analysis. When combined with the ${}^{3}J_{C1H4'}$ [99] coupling constant, the Ψ angle across the glycosidic linkage of the Me-β-D-lactoside (Fig.III.3.6.9c) should be ~20°. This combination of (Φ, Ψ) = (30°, 20°) is not the same as found in the crystal structure of β-lactoside, (40°, -15°). Our results also differ from those of Bose [99], who determined the Ψ angle at -15°. This discrepancy is likely caused by the difference in the ${}^{3}J_{C1C5'}$ coupling constant (2.0 vs. 1.6 Hz) between us and Bose.

III. 3.7 Comparison of the sensitivity of developed INADEQUATE experiments

The sensitivity of the new INADEQUATE experiments (DJM-INEPT-INADEQUATE, DJM-REVINEPT-INADEQUATE, JM-DEPT-INADEQUATE, JM-REVINEPT-INADEQUATE and ^{13}C-detected INADEQUATE) is discussed next. The comparison is based on the isotropic samples of both Me-β-D-xylopyranoside and Me-β-D-lactoside. The acquisition parameters of DJM-INEPT-INADEQUATE were used as the reference for the Me-β-D-xylopyranoside sample. All ^{1}H-detected spectra were processed using a Gaussian window function in the F_2 dimension but no window function was used in the F_1 dimension. ^{13}C-detected spectra were processed using no window function in the F_2 dimension and an exponential window function with LB = 1 Hz in the F_1 dimension. Signal-to-noise ratios were calculated for the C2C4 multiplets in the traces extracted from 2D spectra. This coupling constant (J_{C2C4} = 2.7 Hz) is close to 3 Hz, which was used to set the evolution delay. In this way, when multiplet delays were used, only relaxation effects contributed to the loss of signal. Table III.3.7.1 summarizes the results.

Table III.3.7.1 Sensitivities of INADEQUATE methods using Me-β-D-xylopyranoside *

Parameters	DJM-INEPT	DJM-REVINEPT	JM-DEPT	JM-REVINEPT	^{13}C-detected INADEQUATE
Scaling factor	4	4	4	4	none
Maximum t_1 (ms)	94	91	180	180	23
Experimental time (hours)	23	11	15	8.5	13
Corrected SNR	70:1	43:1	50:1	38:1	58:1
Percentage	100%	61%	71%	54%	83%

* All parameters are for a single 2D spectrum; two experiments were acquired for the last three experiments. SNR was the average value calculated from both C2C4 and C4C2 multiplets. In the J-modulated and ^{13}C-detected INADEQUATE experiments, the SNR is an average of SNRs measured

in in- and antiphase spectra. Because of different total experimental times the SNR was normalized to the 23 hours DJM-INEPT-INADEQUATE experiment.

It would appear that the most sensitive experiment is DJM-INEPT-INADEQUATE. This could be rationalized by the absence of any long fixed delays in this pulse sequence. It should be noted that the SNR given is for the central peak of the multiplet which does not carry the information about the size of the coupling constant. The opposite phase satellites which are separated by the *J* coupling are only half as intense. The relatively high sensitivity of the DJM-INEPT-INADEQUATE composed with the DJM-REVINEPT-INADEQUATE is likely caused by less than 100% heteronuclear NOE build up in the latter experiment. In this experiment the defocusing interval was optimized for all CH_x multiplicities, which also reduced its sensitivity. Similarly, JM-DEPT-INADEQUATE was optimized for CH only, while JM-REVINEPT-INADEQUATE was optimized for all multiplicities. The latter experiment does not contain purging of one-bond cross peaks that is likely responsible for some loss of signal in the DEPT version. As anticipated, the ^{13}C-detected INADEQUATE shows good SNR for the reasons discussed in Chapter III.3.6.

Table III.3.7.2 contains the comparison of SNRs of INADEQUATE spectra of Me-β-D-lactoside. The SNRs in these spectra are comparable. The best performing experiment was DJM-REVINEPT-INADEQUATE. This can be attributed to the lack of fixed intervals in this experiment. In the remaining three experiments one or two 167 ms delays (non-refocused or refocused spectra) are comparable to effective T_2 relaxation times (~ 260 ms). This decreases their sensitivities substantially. Also for this compound, the ^{13}C-detection compares well with ^{1}H-detected experiments.

Table III.3.7.2 Sensitivities of INADEQUATE methods using Me-β-D-lactoside*

Parameters	DJM-REVINEPT	JM-DEPT	JM-REVINEPT	^{13}C-detected INADEQUATE
Scaling factor	4	4	4	none
Maximum t_1 (ms)	75.8	103.6	128.8	9.3
Experimental time (hours)	63.5	23.8	33	25
Corrected SNR	22.3	17.6	13.0	17.2
Percentage	100%	79%	58%	77%

* All parameters are for a single 2D spectrum; two experiments were acquired for the last three experiments. SNR was the average value calculated from both C2C4' and C4'C2 multiplets. In the *J*-modulated and ^{13}C-detected INADEQUATE experiments, the SNR is an average of SNRs measured in in- and antiphase spectra. Because of different total experimental times the SNR was normalized to the 63.5 hours DJM-REVINEPT-INADEQUATE experiment.

III. 3.8 DNP-INADEQUATE:

The Dynamic Nuclear Polarization (DNP) technique has been shown to increase the sensitivity of NMR experiments significantly. In collaboration with Oxford Instruments, we have developed an INADEQUATE pulse sequence that yields the suppression of single-quantum coherences in a single scan.

The polarization of ^{13}C nuclei is close to zero at 1 K, while that of electrons is ~98% (Fig.III.3.8.1). If the polarization of electrons can be transferred to ^{13}C nuclei, this will dramatically increase the ^{13}C-detection. The polarization transfer can be achieved by using microwave irradiation through a combination of solid effect and thermal mixing mechanisms [113], which provides ^{13}C hyperpolarization of ~20% in the solution state and the resulting magnetization enhancement of ~10^4 times. The typical ^{13}C DNP sample is prepared by mixing the ~100 μg solution sample with an equal quantity of trityl free radical in glassing solvent (e.g. DMSO/ethylene glycol), which forms a glass after being frozen with a volume of ~10 to 200 μl. After hyperpolarization by DNP, the solid sample is dissolved and flushed into an NMR tube. A rapid transfer of the liquid sample into the NMR magnet is required for a one scan acquisition of a spectrum.

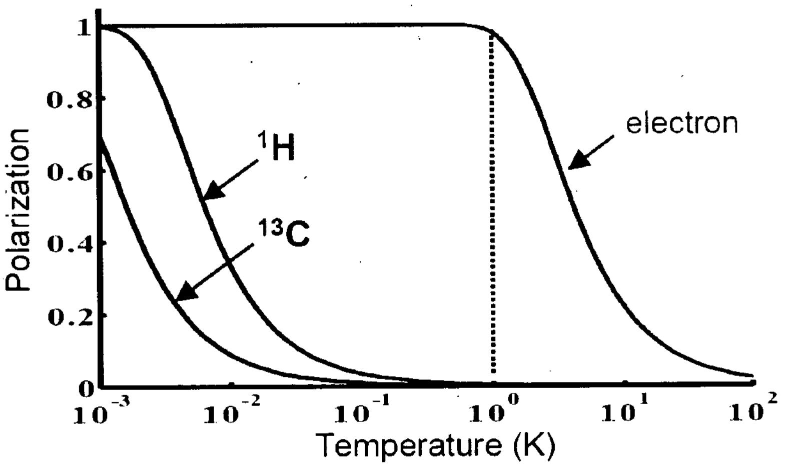


Fig.III.3.8.1 Polarizations of ^1H, ^{13}C nuclei and electrons at low temperature

The demand of a single scan dictates that pulse field gradients must be used for coherence selection. Fig.III.3.8.2a shows the pulse sequence of the basic nonrefocused gradient selected INADEQUATE pulse sequence, while Fig.III.3.8.2b shows the refocused version. The latter gave much better suppression of SQ ^{13}C coherences and therefore only this pulse sequence is discussed next.

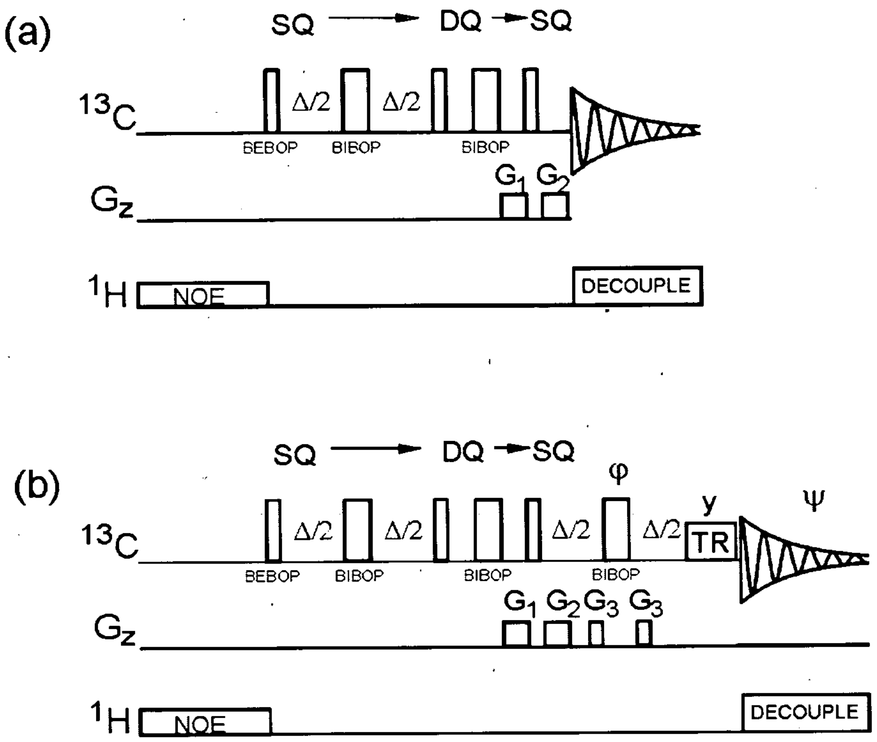


Fig.III.3.8.2 Nonrefocused (a) and refocused (b) pulse sequences of one-scan DNP-INADEQUATE method. Narrow and wide rectangles indicate 90 and 180 deg pulses, respectively. Unless stated otherwise, pulses were applied from the x-axis. For 4 scans experiments φ= x,y,-x,-y and Ψ= x,-x. BEBOP (excitation) [112a] or BIBOP (inversion) [112b] can be used instead of rectangular pulses. Gradients were shaped according to CHIRP pulses and were applied only from z. G_1=40%, G_2=80% and G_3=55%. The first two were 3.5ms long, the last two 0.5 ms.

Over the wide range of ^{13}C chemical shifts, rectangular pulses have poor amplitude and phase profiles as illustrated in Fig.III.3.8.3 on an inversion profile of a 26 μs 180° rectangular pulse. In the context of the INADEQUATE pulse sequence this can have two consequences: (i) strong off-resonance effects can hamper suppression of SQ ^{13}C coherences and (ii) loss of signal for off-resonance DQ coherences. In addition, poor B_1 homogeneity can also cause the two above effects. Therefore, more sophisticated pulses that minimize the problems associated with rectangular pulses were used in the pulse sequence of DNP-INADEQUATE. These pulses were also used in all INADEQUATE experiments discussed before with the exception of DJM-INEPT-INADEQUTE.

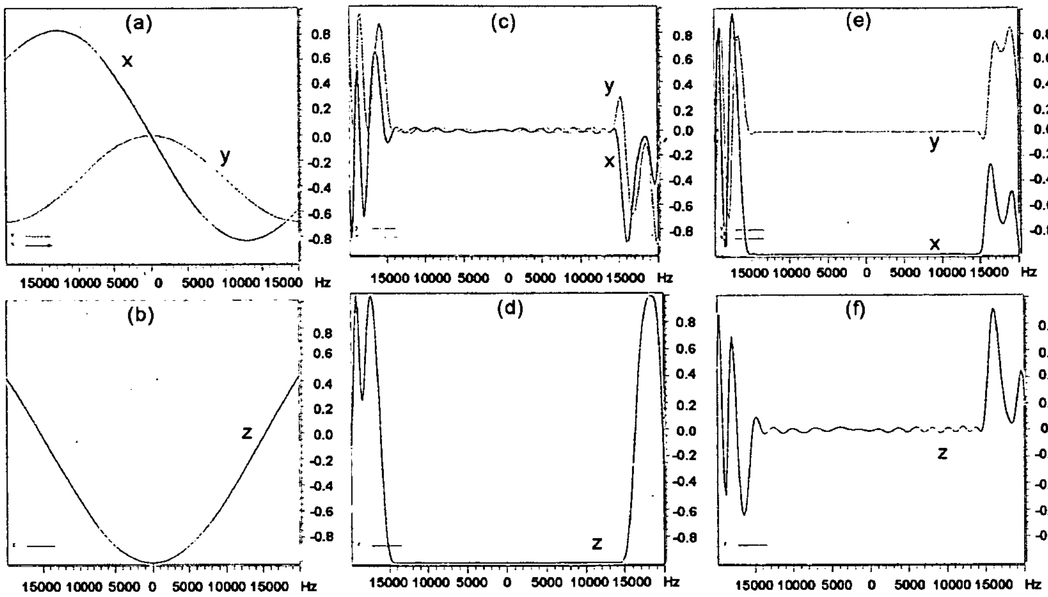


Fig.III.3.8.3 Simulated inversion profiles of 180° pulses (a) and (b) 26 μs hard rectangular pulses and 1400 μs BIBOP pulses (c) to (f). Starting magnetization was in z in (a-d) and x in (e) and (f).

Recently published BEBOP (Broadband Excitation By Optimized Pulses) [112a] and BIBOP (Broadband Inversion By Optimized Pulses) pulses [112b] are resistant to miscalibration and B_1 inhomogenieties, cover larger offsets with stable phases and amplitudes compared to the hard rectangular pulses as illustrated in Fig.III.3.8.4. The BIBOP pulse is a universal rotator, constructed from two BEBOP pulses, meaning that it can be applied to all coherences irrespective of their state. This can be seen in Fig.III.3.8.3c, d, where the starting z magnetization is inverted within ±15 KHz completely by the BIBOP pulse. Similarly, starting from x magnetization, –x magnetization is created across the ±15 KHz range (Fig.III.3.8.3e, f) by the inversion pulse.

As BEBOP and BIBOP pulses are both amplitude and phase modulated, they put high demands on the spectrometer hardware. These pulses were implemented on our spectrometers and their performances were tested. The experimental excitation (Fig.III.3.8.4) profiles show the expected performance of these pulses. The BEBOP pulse excites the magnetizations within ±13.5 KHz uniformly. The deviations of the amplitudes and phases are within 0.7% and 12°, respectively.

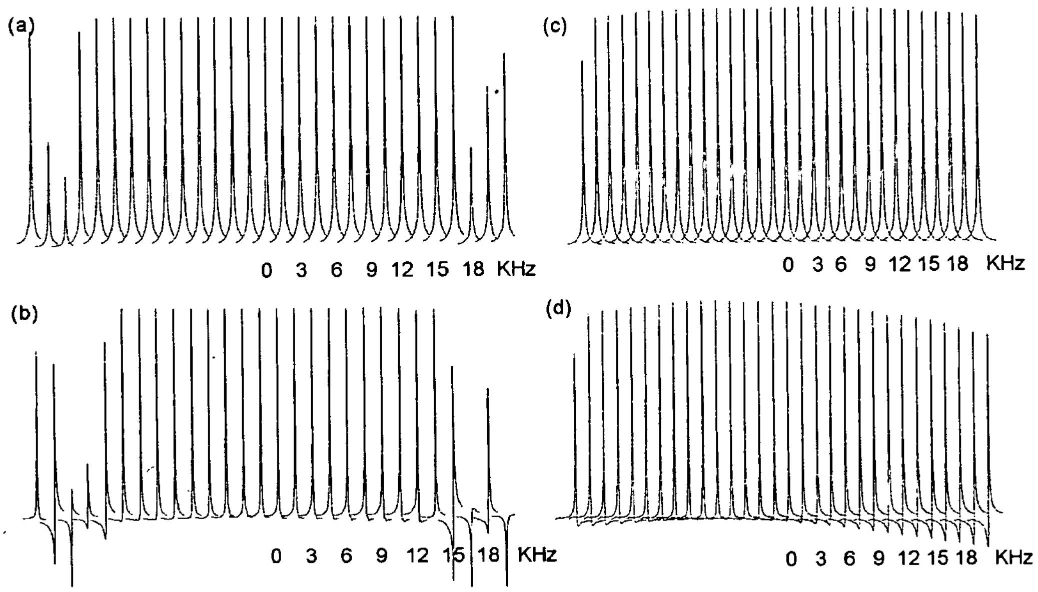


Fig.III.3.8.4 Amplitude (a, c) and phase (b, d) profiles of C1$_\beta$ signals of $^{13}C_1$-glucopyranose using 700 μs BEBOP (a, b) and 13.2 μs hard rectangular (c, d) excitation pulses. The offset was changed ± 21 KHz from the resonance of C1$_\beta$ on a 600 MHz NMR spectrometer.

Both the simulation and experimental results indicate that BEBOP and BIBOP pulses are suitable for use in the DNP-INADEQUATE experiment. As BEBOP is not a universal rotator, only the first 90° pulse of the INADEQUATE pulse sequence could be replaced by this pulse. This is not the case for the BIBOP pulse, where all three 180° rectangular pulses can be replaced. Fig.III.3.8.5 shows a dramatic effect of this replacement by following the signal of $^{13}C_1$ of glucose in a DNP-INADEQUATE pulse sequence as a function of the carrier frequency.

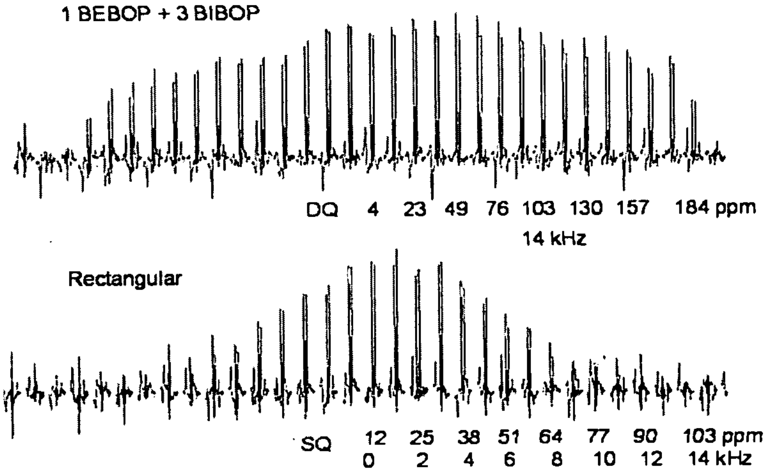


Fig.III.3.8.5 INADEQUATE signal intensities as a function of the offset from the C1$_\beta$ doublet using BEBOP/BIBOP and hard rectangular pulses

1D INADEQUATE spectra of $^{13}C_1$-glucopyranose are shown in Fig.III.3.8.6 using the refocused pulse sequence of Fig.III.3.8.2. One scan, designed for DNP-INADEQUATE, and four scans, using the full phase cycling, spectra are compared in combination with BEBOP/BIBOP and hard rectangular pulses. As the range of ^{13}C frequencies in glucopyranose is small, no difference was seen in the signal intensities in spectra acquired with rectangular or broad-band pulses. As can be seen, the phase cycling removes the artefacts seen around the $C1_{\alpha,\beta}$ signals, while for the $C2_{\alpha,\beta}$ signals the quality of the spectra is identical for one and four scans. As will be shown later, the natural abundance ^{13}C samples gave good suppression of unwanted signals also with one scan. In addition to one-bond correlations, also long-range correlations were observed, but only in the one-scan spectra.

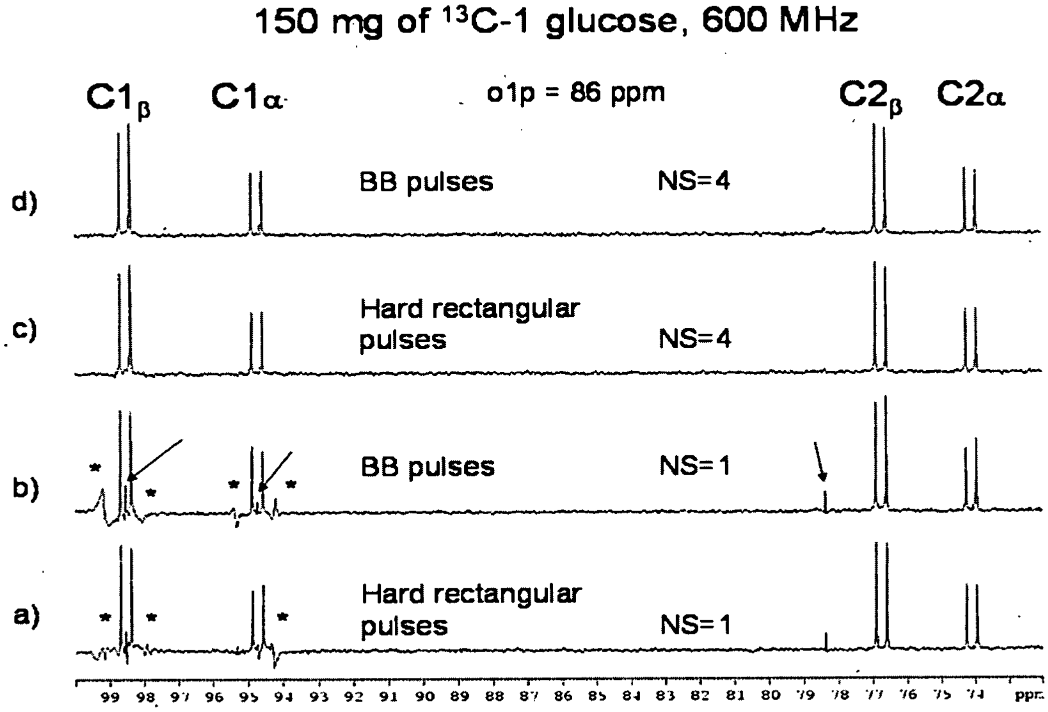


Fig.III.3.8.6 1D INADEQUATE spectra of $^{13}C_1$-glucopyranose acquired with the pulse sequence of Fig.III.3.8.2a. Arrows indicate long-range correlations; asterisks indicate artefacts originating from the main ^{13}C-1 signal. The intensity of the artefacts varies depending on the offset. In (c, d) a trim pulse was used to suppress the long-range correlation together with the phase cycling.

Encouraged by these results, we have tested the refocused INADEQUATE pulse sequence using a natural abundance ^{13}C sample of pyridine (500 μl). The results are presented in Fig.III.3.8.7 and show a good level of suppression of SQ coherences. The pulse sequence was also tested by Oxford Instruments using DNP. For this initial test 30 mg of ^{13}C labeled glucopyranose were used. The spectrum in Fig.III.3.8.8 shows that the pulse sequence worked well, but the signal intensity is low. Considering the T_1 relaxation time of 0.7 s of C1 of glucopyranose and at least 3 s for the transfer of the sample from the polarizer, it can be calculated that only 1.4% of the initial polarization has survived. This points to a major limitation of this technique. Further experiments using natural abundance ^{13}C samples with a more favorable relaxation time are planned.

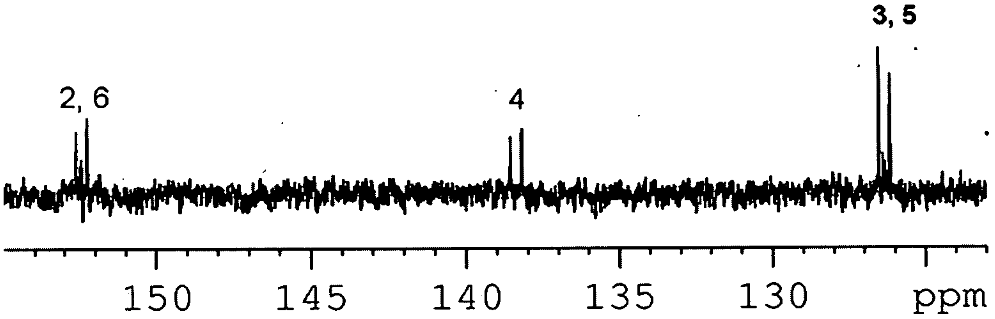


Fig.III.3.8.7 600 MHz natural abundance ^{13}C 1D INADEQUATE spectrum of pyridine without DNP. 500 μL of pyridine + 50 μL of CD_3OD. Eight scans with no phase cycling were acquired ($^1J_{23}$ = 54.1 Hz, $^1J_{34}$ = 54.1 Hz).

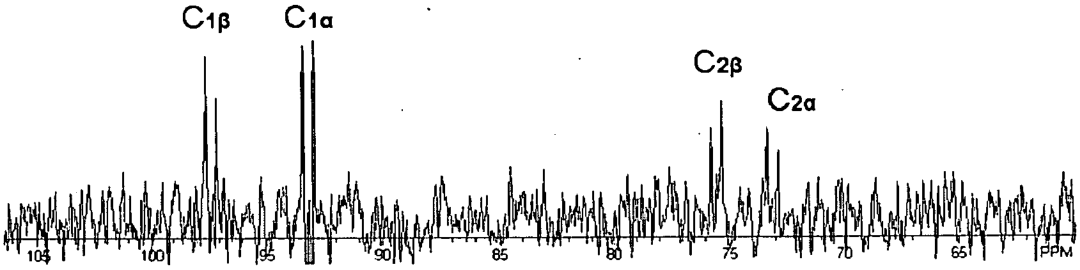


Fig.III.3.8.8 300 MHz hyperpolarized DNP-INADEQUATE spectrum of C-1 ^{13}C-labelled glucopyranose (30 mg in 4 ml CH_3OH)

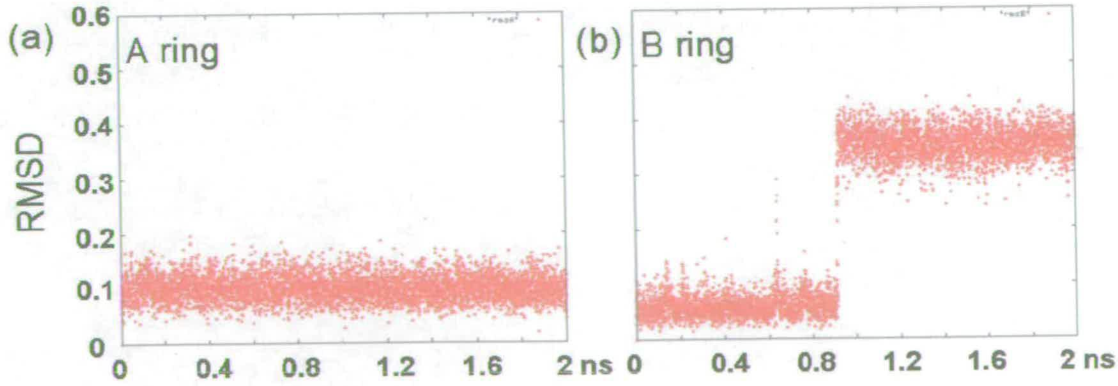
In conclusion, this one-scan INADEQUATE pulse sequence is capable of yielding sufficient suppression of SQ coherences in the DNP setup. In principle, such an experiment can give 1D ^{13}C INADEQUATE spectra from which the $^1J_{CC}$ coupling constants could be determined and in favorable circumstances used to imply the structure of a molecule. If the technology of DNP improves, we would be interested to measure selected long-range ^{13}C-^{13}C coupling constants that could be used in conformational analysis of carbohydrates. This type of experiment would require the use of selective ^{13}C pulses, but we do not foresee any problems with designing such pulse sequences.

III. 4 Solution conformation of the heparin-derived fully sulfated tetrasaccharide II

III. 4.1 AMBER based solution structures without experimental restraints

Force field based solution structures of the fully sulfated tetrasaccharide **II** were generated by free molecular dynamics using the protocol described in Chapter III.1.3. The starting conformations of ${}^4C_1$, ${}^1C_4$, ${}^4C_1$ and ${}^1H_2$ for rings A to D, respectively, (Fig.II.1.1) were considered.

During a 2 ns free molecular dynamics simulation, the A and C rings (glucosamines) stayed in the ${}^4C_1$ conformation. The average RMSDs of the A and C monosaccharide ring atoms relative to the starting structures were 0.1 Å and 0.08 Å, respectively (Fig.III.4.1.1). The B ring, the internal iduronic acid, changed from the starting ${}^1C_4$ conformation to ${}^2S_o$ and stayed in this form until the end of the 2 ns period. This results in a significant change of the positions of the monosaccharide ring atoms (Fig.III.4.1.1). The RMSD of this ring's atoms increases from 0.07 Å to 0.35 Å with reference to the starting positions. The inter-atom distance between H2 and H5 decreases from 3.8 Å to 2.5 Å. When the B ring is in the ${}^1C_4$ form, the two $NHSO_3^-$ groups of the A and C rings are placed on the opposite sides of the molecule ('side-side' geometry). This geometry changes to a 'side-top' when ring B flips to the ${}^2S_o$ form (Fig.III.4.1.2). The nonreducing terminal unsaturated uronic acid, D ring, changed from the starting ${}^1H_2$ conformation to ${}^2H_1$, which is then stable until the end of the 2 ns period. The RMSD of this ring's atoms changes from 0.06 Å to 0.3 Å (Fig.III.4.1.1) with reference to the starting positions. The distance between H1 and H3 decreases from 4.2 Å to 2.7 Å. The D ring tends to be in an 'open' form when it is in the ${}^2H_1$ conformation with respect to the rest of the tetrasaccharide, compared to a 'closed' form when it occupies the ${}^1H_2$ conformation (Fig.III.4.1.2). In other words, the change of the D ring conformation is associated with a significant change of the conformation of the glycosidic linkages between rings CB and DC.



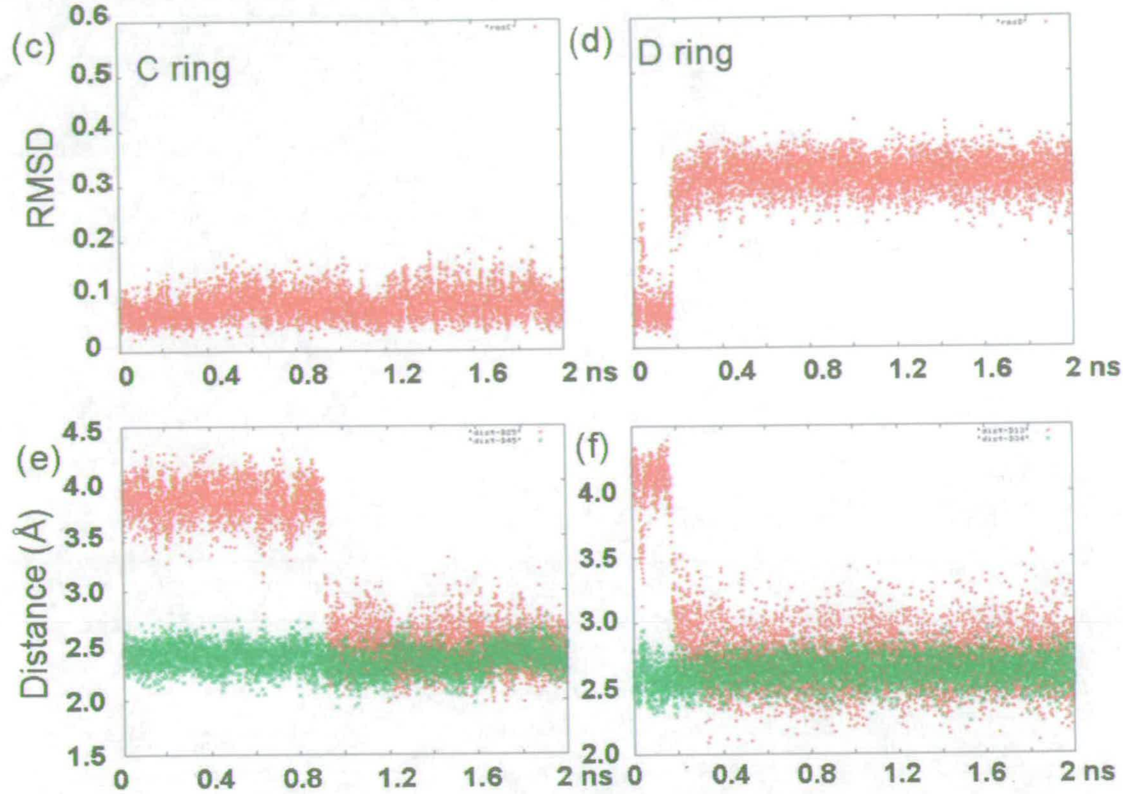


Fig.III.4.1.1 (a)-(d): RMSDs of individual monosaccharide ring atoms (5 carbons and 1 oxygen) with regard to the starting structures. (e) Change of the H2-H5 distance in ring B associated with the transition between the ^2S_O and $^1C_4$ forms; (f) Change of the H1-H3 distance in ring D indicates the change of the conformation of this ring from $^1H_2$ to $^2H_1$.

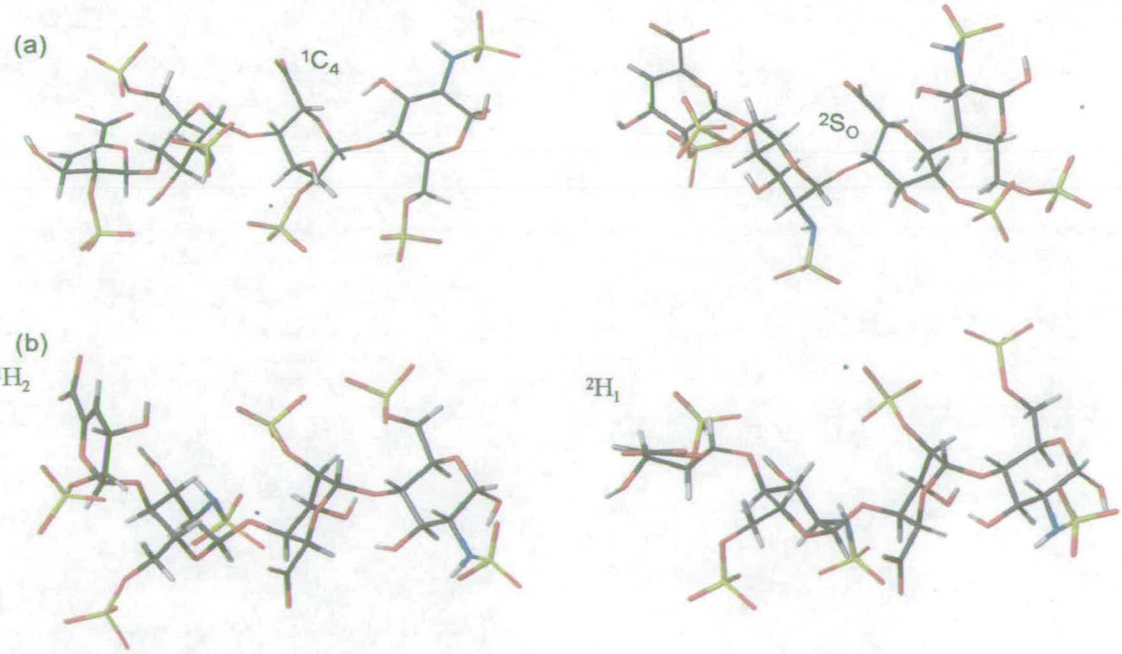
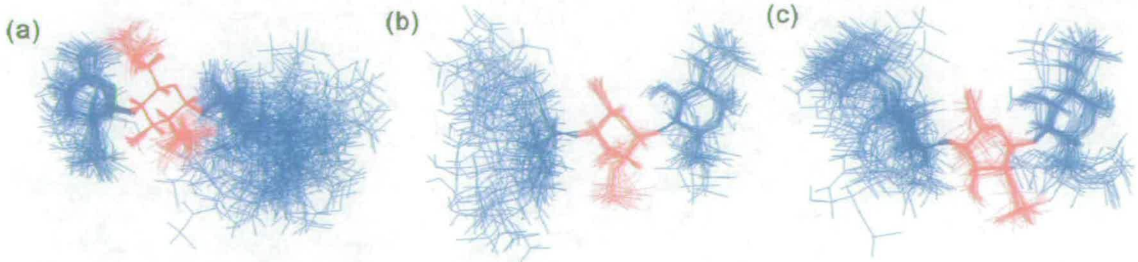
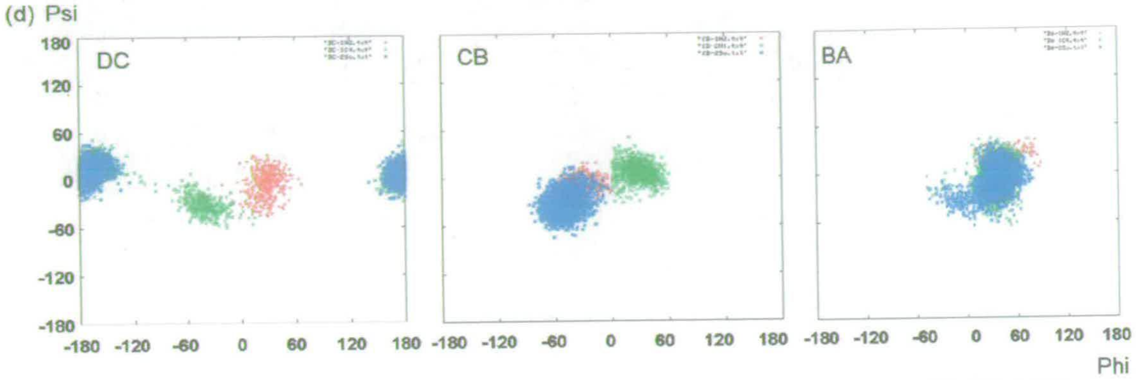


Fig.III.4.1.2 Closest to mean structures showing the differences caused by changed conformations of (a) B ring and (b) D ring.

The conformations of individual monosaccharide rings at the end of the simulation converged to a stable geometry with RMSD values for the ring atoms of individual monosaccharide rings of less than 0.13 Å. The remaining conformational heterogeneity of the tetrasaccharide comes from the varying conformation across the glycosidic linkages. As thousands of structures were calculated during free MD, a subset was taken for further analysis. Fifty structures with the best agreement between the back-calculated and experimental RDCs of the A and C rings are overlaid on the C ring as shown in red in Fig.III.4.1.3a. For this selection the RDCs were calculated separately for rings A and C in order not to bias the subset, but only to choose molecules with good geometries of individual rings.

The glycosidic dihedral angles across the AB ring linkage (Fig.III.4.1.3.d) cluster around (Φ, Ψ) = (38°, 0°) irrespective of the conformation of the B ring. The CB linkage (Fig.III.4.1.3.d) is affected by the conformational change of the B ring and shows a difference of 80 and 20 degrees in Φ and Ψ angles between the $^1C_4$ and ^2S_o conformations. This linkage is also affected by the conformation of the D ring ($^1H_2$ and $^2H_1$) with the Φ angle changes ~70 degrees, while the Ψ angle changes 16 degrees. The DC linkage (Fig.III.4.1.3.d) shows the heterogeneity, which originates in conformational changes of both B and D rings. When the B ring is in the $^1C_4$ form, the Ψ_{DC} is similar for both $^1H_2$ and $^2H_1$ conformations of the D ring. However, for the $^1H_2$ form the Φ_{DC} angle is 28°, but for the $^2H_1$ form it oscillates between -30° and 180°. When the B ring is in the ^2S_o form Φ_{DC} is 180° only and the Ψ_{DC} is similar to the other conformer. The geometry of glycosidic linkages will be further characterized by using RDCs restrained molecular dynamics (Chapter III.4.4.3). Before that, NMR data are analyzed in order to characterize the internal dynamics of individual rings.





$^1C_4$ & $^1H_2$	Phi: 27.7 +/- 11.5 Psi: -4.1 +/- 16.5	Phi: -35.0 +/- 12.9 Psi: -11.2 +/- 11.2	Phi: 42.4 +/- 12.7 Psi: 10.5 +/- 11.1
$^1C_4$ & $^2H_1$	Phi: -41.5 +/- 15.2 or -165.0 +/- 10.3 Psi: 3.9 +/- 22.0	Phi: 29.3 +/- 12.9 Psi: 5.6 +/- 11.2	Phi: 36.7 +/- 13.1 Psi: -3.2 +/- 18.1
^2S_O & $^2H_1$	Phi: -168.4 +/- 9.0 Psi: 9.4 +/- 11.0	Phi: -50.8 +/- 12.3 Psi: -32.9 +/- 13.3	Phi: 36.0 +/- 16.2 Psi: -7.4 +/- 15.8

Fig.III.4.1.3 (a) 50 structures overlaid on C ring; (b) 20 out of 50 structures overlaid on B ring in the form of $^1C_4$; (c) 30 out of 50 structures overlaid on B ring in the form of ^2S_O; (d) dihedral angles across the glycosidic linkages of 5,000 structures and statistics for dihedral angles.

III. 4.2 Conformations of individual monosaccharide rings of compound II

Three-bond proton-proton coupling constants ($^3J_{HH}$) of the four monosaccharide rings (Table III.4.2.1) were measured using *J*-modulated SPFGSE-COSY, CSSF-COSY and HOHAHA-COSY methods described in Chapter III.2. As evidenced by large values of $^3J_{23}$, $^3J_{34}$ and $^3J_{45}$ coupling constants, the two glucosamine rings (A and C) are stabilized in the $^4C_1$ conformation. The internal iduronic acid ring (B) and the nonreducing terminal of the unsaturated uronic acid ring (D) exist in a conformational equilibrium. Fig.III.4.2.1 shows the possible conformations for these rings.

Table III.4.2.1 $^3J_{HH}$ coupling constants, in Hz, of heparin-derived fully sulfated tetrasaccharide **II**

Rings / $^3J_{HH}$	H_1-H_2	H_2-H_3	H_3-H_4	H_4-H_5
A	3.5	10.2	8.8	9.8
C	3.7	10.6	9.0	10.1
B	2.4	4.8	3.4	2.5
D	3.4	2.6	4.7	-

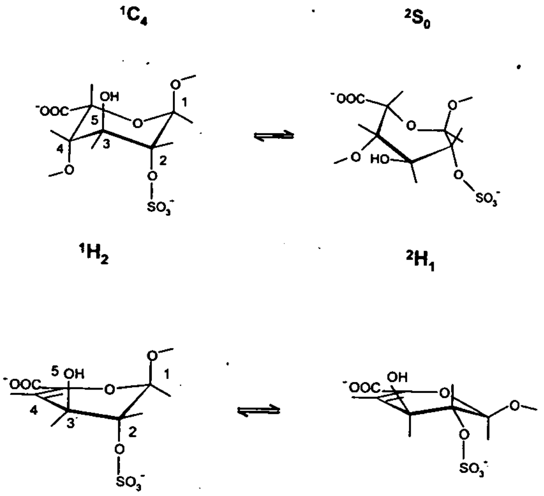


Fig.III.4.2.1 Possible conformations of the internal iduronic acid ring (B) and the nonreducing end unsaturated uronic acid ring (D)

The three-bond proton-proton coupling constants ($^{3}J_{HH}$) depend on the dihedral angle across the C-C bond and the electronegativities of the substituents. Considering the effects of substituents and incorporating them into a generic Karplus equation, an empirically parameterized (P_1 to P_6) equation has been obtained from a set of 315 coupling constants [114] (**Eq.IV.2.2.1**):

$$^{3}J_{HH} = P_1\cos^2\Phi + P_2\cos\Phi + P_3 + \Sigma\Delta\chi_i\{P_4 + P_5\cos^2(\xi_i\Phi + P_6\,|\Delta\chi_i|)\}$$

$$\Delta\chi_i^{\text{group}} = \Delta\chi_i^{\alpha\text{-sub}} - P_7\,\Sigma\Delta\chi_i^{\beta\text{-sub}}$$

$\Delta\chi_i$ is the difference of electronegativities between the substituent and proton attached to the carbon atom.

ξ_i is either +1 or -1 depending on the orientation of the substituent.

Using the Pauling electronegativity scale, the above equation for individual coupling constants of the iduronic ring (B ring) was parameterized as:

$$\begin{aligned}^{3}J_{H1H2} = {} & 13.24\cos^2\Phi - 0.91\cos\Phi + 1.1735\{0.53 - 2.41\cos^2(\Phi + 15.5\times 1.1735)\} \\ & + 1.1735\{0.53 - 2.41\cos^2(-\Phi + 15.5\times 1.1735)\} \\ & + 0.0479\{0.53 - 2.41\cos^2(-\Phi + 15.5\times 0.0479)\} \\ & + 1.1678\{0.53 - 2.41\cos^2(\Phi + 15.5\times 1.1678)\}\end{aligned}$$

$$
\begin{aligned}
{}^3J_{H2H3} = 13.24\cos^2\Phi - 0.91\cos\Phi &- 0.1212\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.1212)\} \\
&+ 0.0479\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.0479)\} \\
&+ 1.24\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.24)\} \\
&+ 1.1678\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.1678)\}
\end{aligned}
$$

$$
\begin{aligned}
{}^3J_{H3H4} = 13.24\cos^2\Phi - 0.91\cos\Phi &+ 1.24\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.24)\} \\
&+ 0.0479\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.0479)\} \\
&+ 0.0479\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.0479)\} \\
&+ 1.1735\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.1735)\}
\end{aligned}
$$

$$
\begin{aligned}
{}^3J_{H4H5} = 13.24\cos^2\Phi - 0.91\cos\Phi &+ 1.1735\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 1.1735)\} \\
&+ 1.1735\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 1.1735)\} \\
&+ 0.0479\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 0.0479)\} \\
&- 0.3568\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 0.3568)\}
\end{aligned}
$$

The iduronic acid ring can adopt three different conformations, ${}^4C_1$, ${}^1C_4$ and ${}^2S_o$, in solution but only ${}^1C_4$ and ${}^2S_o$ were found for an internal iduronic acid (B ring) in fully sulfated tetrasaccharide [1, 2]. The dihedral angles (Table III.4.2.2a) across the C-C bonds were measured for standard ${}^1C_4$ and ${}^2S_o$ conformations and used for the calculation of the theoretical ${}^3J_{HH}$ coupling constants using the above parameterized equations. The population of each conformer was calculated by fitting the calculated coupling constants to the experimental data. The result is shown in Table III.4.2.2b. The large deviation of ${}^3J_{12}$ might be caused by not taking properly into account the geometry of the glycosidic linkage.

Table III.4.2.2 (a) Dihedral angles of H-C-C-H bonds of the B ring in the conformations of ${}^1C_4$ and ${}^2S_o$ (b) Least-square analysis of the populations of individual conformers

(a)

Dihedral angles (degree)	$H_1C_1C_2H_2$	$H_2C_2C_3H_3$	$H_3C_3C_4H_4$	$H_4C_4C_5H_5$
${}^1C_4$	70.8	-79.1	89.8	44.1
${}^2S_O$	148.2	-167.2	145.2	33.5

(b)

$^3J_{HH}$(Hz)	H_1-H_2	H_2-H_3	H_3-H_4	H_4-H_5	wt%
$^1C_4$	2.6	3.2	2.5	1.4	76%
^2S_o	5.6	9.2	5.5	4.7	24%
Expt.	2.4	4.8	3.4	2.5	RMSD = 0.47 Hz
Cal.	3.3	4.6	3.2	2.2	

The population of each conformer can be further inspected by analyzing the 2D NOESY spectrum. The distance between H5 and H4 is around 2.40 Å in both $^1C_4$ and ^2S_o conformers. But there is a significant change of the distance between atoms H5 and H2, when the B ring transits from $^1C_4$ (3.90 Å) to ^2S_o (2.51 Å) form. The distance of 3.90 Å is too large to be detected in the NOESY experiment for a molecule of the size of a tetrasaccharide. Therefore, the H2-H5 cross peaks in the NOESY spectrum (Fig.III.4.2.2a) can be assumed to originate from the ^2S_o conformer. Based on the identical distance between H4 and H5 in the two conformers, the comparison of the intensities [115] of the cross peaks of H2 and H4 in the H5 trace (Fig.III.4.2.2b) can be used to calculate the population of the ^2S_o form. The ^2S_O form is present in 22% based on this analysis, which is close to 24% determined based on the analysis of $^3J_{HH}$ coupling constants. Further analysis of the population of the two conformers of this internal iduronic ring based on RDCs and molecular modeling is presented in Chapter III.4.4.4 from the RDC-restrained MD results.

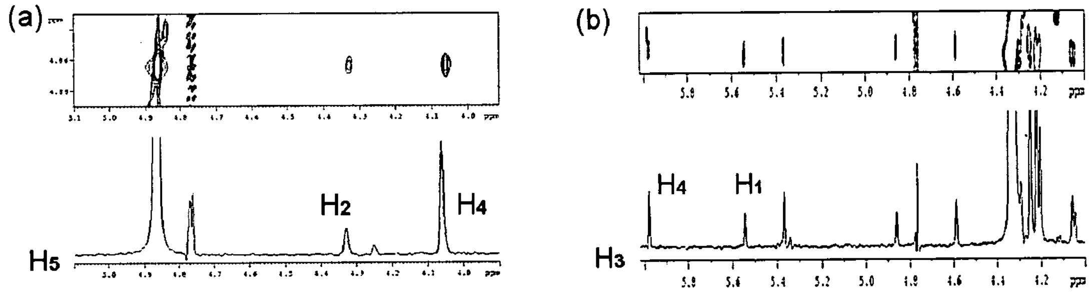


Fig.III.4.2.2 Partial 2D NOESY spectra of compound **II** and proton traces of (a) H5 on B ring and (b) H3 on D ring.

In the case of the D ring, the parameterized Karplus equations of three-bond H-H coupling constants are shown below:

$$\begin{aligned}{}^3J_{H1H2} = 13.24\cos^2\Phi - 0.91\cos\Phi &+ 1.1735\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.1735)\} \\ &+ 1.1735\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 1.1735)\} \\ &+ 0.0479\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.0479)\} \\ &+ 1.1678\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.1678)\}\end{aligned}$$

$$\begin{aligned}{}^3J_{H2H3} = 13.24\cos^2\Phi - 0.91\cos\Phi &+ 1.1678\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.1678)\} \\ &+ 1.24\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.24)\} \\ &- 0.1212\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.1212)\} \\ &+ 0.217\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.217)\}\end{aligned}$$

$$\begin{aligned}{}^3J_{H3H4} = 13.24\cos^2\Phi - 0.91\cos\Phi &+ 1.24\{0.53 - 2.41\cos^2(\Phi + 15.5 \times 1.24)\} \\ &+ 0.0958\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.0958)\} \\ &+ 0.0479\{0.53 - 2.41\cos^2(-\Phi + 15.5 \times 0.0479)\}\end{aligned}$$

This unsaturated uronic acid ring exists in the equilibrium of ${}^1H_2$ and ${}^2H_1$ conformations. The dihedral angles (Table III.4.2.3a) across the C-C bonds were measured for standard ${}^1H_2$ and ${}^2H_1$ conformations and used for the calculation of the theoretical ${}^3J_{HH}$ coupling constants using the above parameterized equations. The population of each conformer was calculated by fitting the calculated coupling constants to the experimental data (Table III.4.2.3b).

Table III.4.2.3 (a) Dihedral angles of H-C-C-H bonds of ${}^1H_2$ and ${}^2H_1$ conformers of the D ring. (b) Least-square analysis of the populations of individual conformers

(a)

Dihedral angles (degree)	$H_1C_1C_2H_2$	$H_2C_2C_3H_3$	$H_3C_3C_4H_4$
${}^1H_2$	57.8	-75.5	44.9
${}^2H_1$	-178.9	-161.6	72.1

(b)

${}^3J_{HH}$ (Hz)	H_1-H_2	H_2-H_3	H_3-H_4	wt%
${}^1H_2$	3.10	0.18	5.98	78%
${}^2H_1$	8.30	10.20	1.67	22%
Expt.	3.36	2.65	4.66	RMSD = 0.57 Hz
Cal.	4.24	2.38	5.03	

The large deviation of $^3J_{H1H2}$ might be because of the effects of the glycosidic linkage and/or not very accurate parameterization of the Karplus equation for this particular coupling constant. The unsaturated double bond of C4-C5 affects the accuracy of the analysis based on the $^3J_{H3H4}$ coupling constant. All of these factors contribute to the relative ambiguity of this population analysis. Therefore, only a qualitative indication of the domination of the $^1H_2$ conformation of ring D can be reached.

Similar to the analysis of ring B, the NOESY spectrum can be used to analyze the conformation of ring D. The distance between H3 and H4 is around 2.50 Å in both $^1H_2$ and $^2H_1$ conformers. But there is a significant change of the distance between atoms H3 and H1 from $^1H_2$ (4.20 Å) to $^2H_1$ (2.50 Å). Only the distance of 2.50 Å can be detected in the NOESY experiment. Therefore, the H3-H1 cross peaks in the NOESY spectrum (Fig.III.4.2.2b) must come from the $^2H_1$ conformer. By comparing the intensities of the H3-H1 and H3-H4 cross peaks in the H3 trace, the population of the $^2H_1$ form can be calculated as 56%. This result is very different from the scalar coupling constants analysis, which gives the population of the form $^2H_1$ to be 22%.

III. 4.3 Geometry of glycosidic linkages

Besides the flexibility of individual rings, the conformation of a carbohydrate is usually associated with the flexibility across the glycosidic linkages. The conformation of the glycosidic linkage is characterized by two dihedral angles Φ ($H_{1'}$-$C_{1'}$-O-C_4) and Ψ ($C_{1'}$-O-C_4-H_4). Via the Karplus equation, Φ and Ψ angles can be related to the three-bond coupling constants. Three-bond heteronuclear C-H coupling constants ($^3J_{COCH}$) across the glycosidic linkages of tetrasaccharide **II** were determined from an HMBC spectrum (Fig.III.4.3.1a and Table III.4.3.1). Two identical multiplets extracted from 1D ^1H or 1D TOCSY spectra are manipulated first in order to simulate the chemical shift and coupling evolution of multiplets in the HMBC [68] experiment. Such modified multiplets are then shifted in opposite directions and subtracted yielding a theoretical HMBC multiplet (Fig.III.4.3.1b). The mutual shift of the two multiplets that reproduce best the HMBC multiplet represents the $^nJ_{CH}$ coupling constants (Fig.III.4.3.1c).

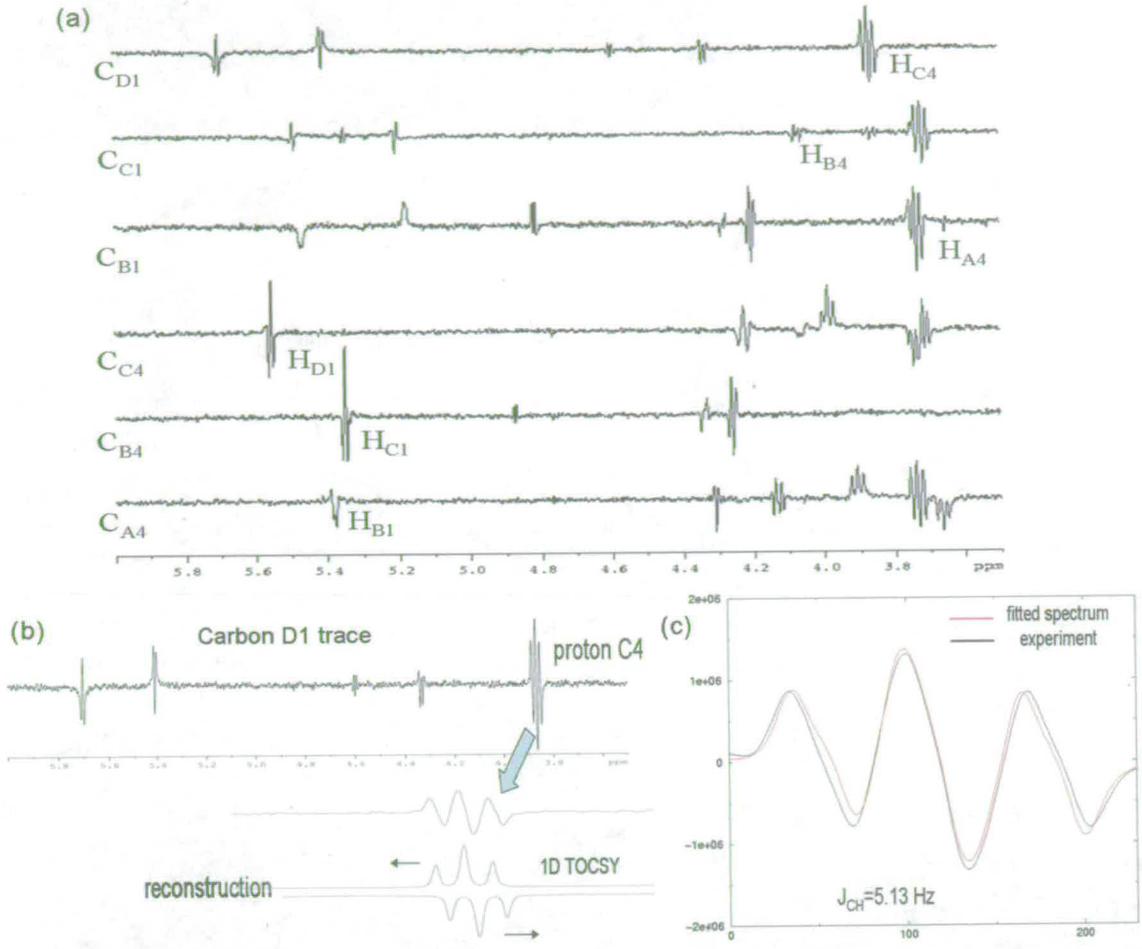


Fig.III.4.3.1 (a) Traces of carbons of the glycosidic linkages extracted from the 2D HMBC spectrum of compound **II**; (b) Reconstruction of the HMBC multiplets from 1D TOCSY spectra; (c) Best fit between the experimental and reconstructed mutliplets.

Table III.4.3.1 $^3J_{COCH}$ across the glycosidic linkages from the HMBC spectrum of compound **II**. The first row: 1D ^1H spectrum was used as a determined reference. The second row: 1D TOCSY spectrum was used as a reference

$C_{A4}H_{B1}$	$C_{B4}H_{C1}$	$C_{C4}H_{D1}$	$C_{D1}H_{C4}$	$C_{C1}H_{B4}$	$C_{B1}H_{A4}$
3.8	4.4	4.1	5.0	-	4.8
4.6	4.4	4.1	5.1	5.0	4.8

The Karplus equation parameterized for the glycosidic linkage of carbohydrates was used to convert the measured coupling constants into dihedral angles [116]. The Karplus equations are not unambiguous and for each measured coupling constant there are four possible dihedral angles (Fig.III.4.3.2). $^3J_{COCC}$ would need to be measured to complement the $^3J_{COCH}$ coupling constants in order to decide if a single angle satisfies the experimental data or whether there is flexibility around the linkages. Several methods have been developed in this project to achieve this aim (Chapter III.3). Unfortunately, we did not have compound **II** in sufficient quantity to obtain $^3J_{COCC}$ coupling constants and such analysis was left for a later day.

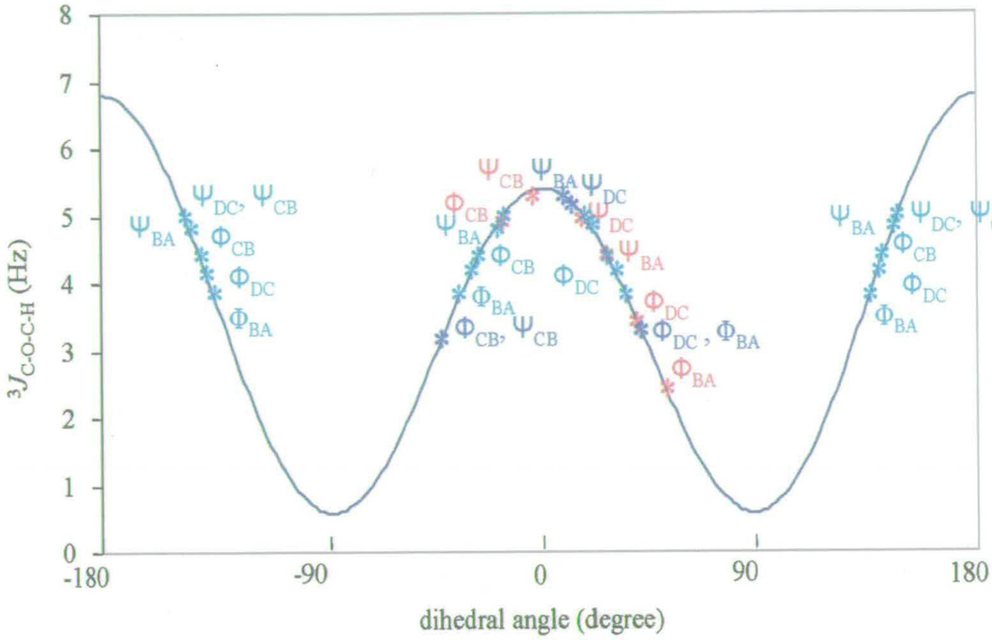


Fig.III.4.3.2 Dihedral angles across the glycosidic linkages of tetrasaccharide **II** on the Karplus curve [116] $^3J_{COCH} = 5.7\cos^2\theta - 0.6\cos\theta + 0.5$; red, blue and green asterisks represent the values obtained from the crystal structure [86], solution structure [85] of **II**, and our measurements, respectively.

Table III.4.3.2 Dihedral angles (in degrees) across the glycosidic linkages of the fully sulfated tetrasaccharide **II** from the measured $^3J_{CH}$ coupling constants, solution [85] and crystal structure [86] in the complex with bFGF; values based on $^3J_{CH}$ and NOESY data of compound **II** are shown in bold.

	Φ_{DC}	Ψ_{DC}	Φ_{CB}	Ψ_{CB}	Φ_{BA}	Ψ_{BA}
Tetrasaccharide **II**	±145 ±30 **30**	±150 ±20 **±20**	±147 ±25 **-25**	±150 ±20 **-20**	±145 ±35 **35**	±150 ±20 **±20**
Solution structure [85]	45.7	13.2	-43.3	-42.2	45.2	15.0
Crystal structure [86]	42.3	18.3	-18.5	-3.5	57.0	25.1

The dihedral angles are compared to the literature values in Table III.4.3.2. Analysis of the NOESY spectrum allowed reducing the ambiguity of the six dihedral angles based on the interpretation of $^3J_{CH}$ coupling constants. The absence of cross peaks of the anti conformer (dihedral angles around ± 180°) in the NOESY spectra means that only conformations with Φ and Ψ angles within ± 90° are possible. In some instances a preference for a positive or negative dihedral angle could also be established. For example, in the NOESY spectrum, the cross peak D1C3 (H1 of D ring and H3 of C ring) is weaker than that of D1C5, which indicates the distance between D1C3 is larger than D1C5. This selects the Φ_{DC} angle to be 30°. The Ψ_{DC} angle can be either 20° or -20°. The distances C1B4 and C1B3 are much smaller compared to C1B5, which restricts Φ_{CB}, Ψ_{CB} angles to -25° and -20°, respectively. The distance of A3B5 is larger compared to that of A3B1, which indicates the Φ_{BA} angle to be 35°. The Ψ_{BA} angle can be either 20° or -20°. The ambiguity of Ψ_{DC} or Ψ_{BA} angles can be explained as the result of the conformational equilibria between forms of $^1H_2$ and $^2H_1$ for the D ring and $^1C_4$ and ^2S_O for the B ring.

The solution [85] and crystal [86] structures of tetrasaccharide **II** are both in the $^1H_2$-$^4C_1$-$^1C_4$-$^4C_1$ form, while the ^{1}H-^{1}H coupling constant analysis of tetrasaccharide **II** in solution indicated the presence of an equilibrium process. Despite this, there is a qualitative agreement between our results and the crystal structure with differences of less than 20°. Comparing the experimental dihedral angles with the non-restrained molecular dynamics shows significant differences only for ring D, where the presence of the anti conformer in simulations is not confirmed by the experiment.

RDC-restrained molecular dynamics calculations were used to further characterize the orientations across the glycosidic linkages. A detailed analysis is presented in the following chapter.

III. 4.4 Conformation from residual dipolar coupling constants

III. 4.4.1 Measurement of $^3D_{HH}$ and $^1D_{CH}$ residual dipolar coupling constants

Scalar coupling constants can be measured from the isotropic sample. The aligned sample yields the sum of the scalar and dipolar coupling constants. Therefore, RDCs can be obtained from NMR experiments acquired under the same conditions using the two samples. The three-bond ^{1}H-^{1}H coupling constants were obtained from *J*-modulated SPFGSE-COSY, CSSF-COSY and HOHAHA-COSY experiments developed in this work (Chapter III.2). An example of the coupling constant determination of tetrasaccharide **II** is shown in Fig.III.4.4.1.1.

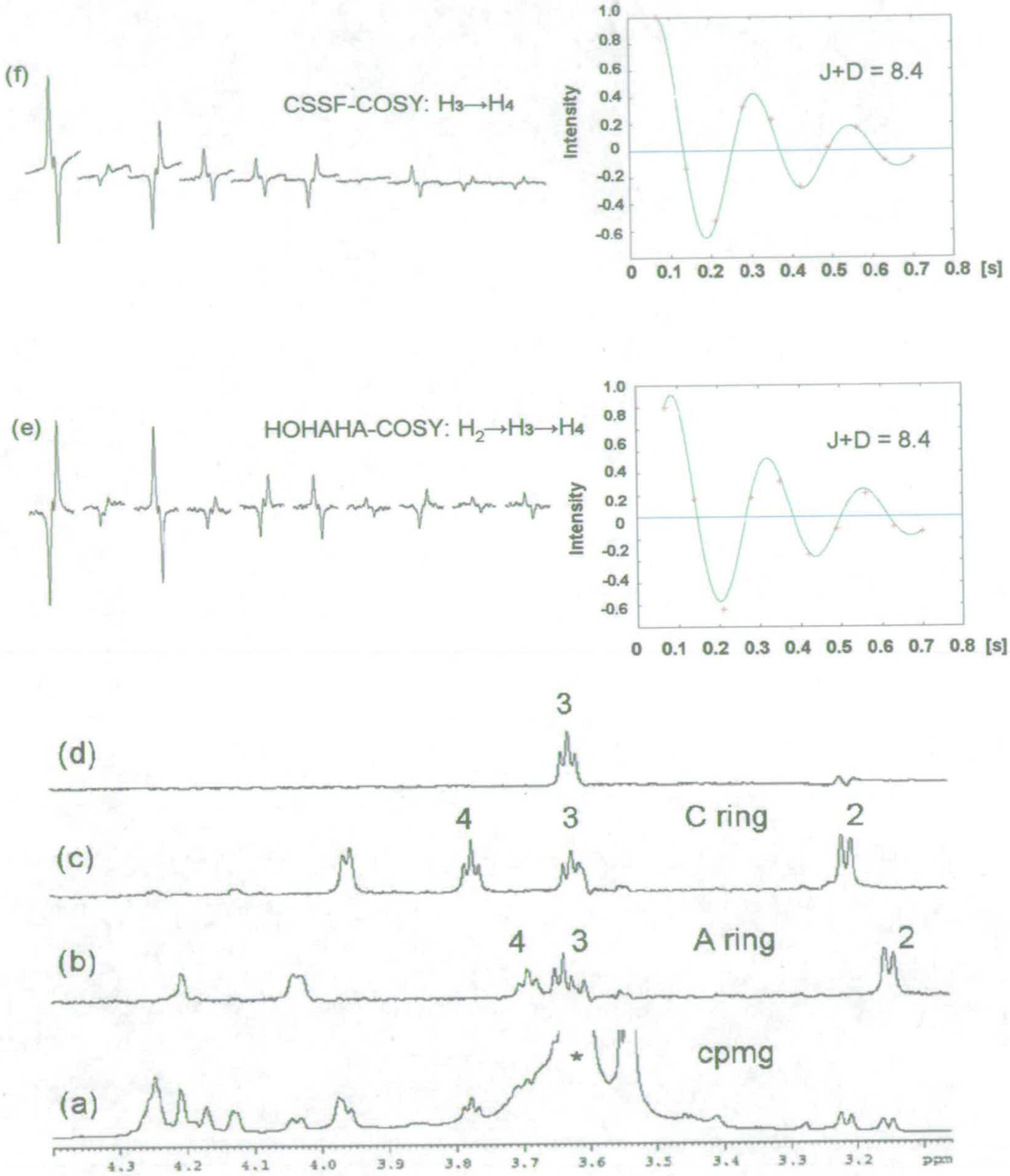


Fig.III.4.4.1.1 (a) 1D CPMG spectrum of the aligned sample of tetrasaccharide **II**; the asterisk indicates the residual ^{1}H signal of the medium. (b) and (c) are 1D TOCSY spectra of the A and C rings; (d) 1D double selective HOHAHA spectrum of **II** with a selective transfer of magnetization from H2 to H3 of the C ring; (e) and (f) coupling constant determinations of $^3|J+D|_{H3H4}$ from the variable-time *J*-modulated 1D HOHAHA-COSY and CSSF-COSY experiments, respectively. Signal intensities of proton H4 were fitted to the COSY delays.

The one-bond C-H coupling constants were extracted from the F_2 dimension of the nonrefocused ^{1}H-^{13}C HSQC spectrum. The coupling constants (Table III.4.4.1.1) of the two glucosamine rings (A and C) reflect the $^4C_1$ chair form, but those of the uronic acid rings (B and D) are average values of the exchanging conformations.

Table III.4.4.1.1 Scalar and dipolar coupling constants of tetrasaccharide **II**

Tetrasaccharide **II**		*J* (Hz)	*J* + *D* (Hz)	RDC (Hz)
C-H coupling constants	C1-H1	172.2	168.8	-3.4
	C2-H2	139.1	144.9	5.8
A ring	C3-H3	148.0	154.1	6.1
	C4-H4	147.2	152.6	5.4
	C5-H5	146.9	153.1	6.2
B ring	C1-H1	174.1	178.9	4.8
	C2-H2	151.4	152.5	1.1
	C3-H3	151.5	148.9	-2.6
	C4-H4	148.8	152.8	4.0
	C5-H5	146.1	145.7	-0.4
C ring	C1-H1	172.6	177.7	5.1
	C2-H2	138.7	135.6	-3.1
	C3-H3	147.4	144.4	-3.0
	C4-H4	147.2	142.4	-4.8
	C5-H5	147.0	143.4	-3.6
D ring	C1-H1	175.9	177.0	1.1
	C2-H2	154.4	153.0	-1.4
	C3-H3	151.6	155.7	4.1
	C4-H4	169.6	169.5	-0.1
H-H coupling constants	H1-H2	3.53	2.71	-0.82
	H2-H3	10.25	11.13	0.88
A ring	H3-H4	8.79	9.75	0.96
	H4-H5	9.85	10.75	0.90
B ring	H1-H2	2.42	3.97	1.55
	H2-H3	4.76	3.53	-1.22
	H3-H4	3.41	3.30	-0.11
C ring	H1-H2	3.67	2.46	-1.21
	H2-H3	10.59	11.14	0.55
	H3-H4	9.04	8.43	-0.61
	H4-H5	10.16	8.80	-1.36
D ring	H1-H2	3.36	3.83	0.47
	H2-H3	2.65	3.50	0.85
	H3-H4	4.66	3.90	-0.76

III. 4.4.2 Analysis of the RDCs of A and C rings using disaccharides

We have noticed in the past that RDCs are very sensitive to the actual geometry of a carbohydrate ring [89a]. Differences of a few degrees in dihedral and bond angles can provide measurable differences in RDCs. It is also known that the neighboring rings can affect the geometry of pyranose rings in oligosaccharides. In collaboration with Dr. Miloš Hricovini we have therefore decided to investigate whether we can see any proof for this by analyzing the geometry of rings A and C in the context of the tetrasaccharide **II**. By investigating these rings separately, we are not making any assumptions about the alignment of the whole tetrasaccharide. The flexibility of the molecule is absorbed by the alignment allowing accurate calculation of the alignment tensor. Dr. Hricovini has calculated ab initio structures of two disaccharides (private communication) GlcNS6S-1->4-IdoA2SMe and IdoA2S-1->4-GlcNS6SMe, in explicit water. These structures correspond to the C-B and B-

A fragments of tetrasaccharide **II**, respectively. Both $^1C_4$ and ^2S_o conformations were considered for the B ring (Fig.III.4.4.2.1).

Fig.III.4.4.2.1 CB and BA disaccharides overlaid on C (CB) and A (BA) rings.

Selected structural parameters of the four disaccharides are given in Table III.4.4.2.1. Differences in dihedral angles >10° and >3° in bond angles are highlighted in bold.

Table III.4.4.2.1 Dihedral and bond angles (in degree) of rings C and A in disaccharides CB and BA, respectively.

	CB ($^1C_4$-$^4C_1$)	**CB** (^2S_O-$^4C_1$)	**Diff.**	**BA** ($^4C_1$-$^1C_4$)	**BA** ($^4C_1$-^2S_O)	**Diff.**
H_1-C_1-O_1-$C_{4'}$ (**Φ**)	-66.2	-50.3	**-15.9**	-42.3	67.7	**-110.0**
C_1-O_1-$C_{4'}$-$H_{4'}$ (**Ψ**)	-37.3	-44.2	6.9	-45.6	50.8	**-96.4**
H_1-C_1-C_2-H_2	55.7	68.4	**-12.7**	55.7	56.9	-1.2
H_2-C_2-C_3-H_3	179.2	166.0	**13.2**	166.6	-173.7	**19.7**
H_3-C_3-C_4-H_4	170.7	172.5	-1.8	-167.0	171.1	**21.9**
H_4-C_4-C_5-H_5	-166.4	-155.3	**-11.1**	175.4	-176.2	8.4
C_5-O-C_1	116.9	116.9	0	117.5	110.5	**7.0**
O-C_1-H_1	105.3	106.1	-0.8	103.8	109.8	**-6.0**
H1-C1-C2	111.6	112.6	-1.0	110.8	109.7	1.1
C1-C2-C3	106.3	105.8	0.5	105.3	110.2	**-4.8**
H2-C2-C3	109.0	109.4	-0.4	107.3	110.2	-2.9
C2-C3-H3	109.2	104.3	**4.9**	109.8	109.9	-0.1
H3-C3-C4	108.3	107.2	1.1	109.8	106.9	2.9
C3-C4-H4	109.2	109.2	0	108.5	111.1	-2.6
H4-C4-C5	109.3	110.7	-1.4	108.0	110.3	-2.3
C4-C5-H5	108.1	106.6	1.5	108.9	107.0	1.9
H5-C5-C6	109.4	110.5	-0.9	109.3	107.6	1.7

There are differences >10° in several dihedral angles of the C ring in the CB disaccharide between the two conformations of the B ring (Table III.4.4.2.1). As the Φ and Ψ dihedral angles across the glycosidic linkage are very close in the two forms, it is possible that the primary reason for these differences is the changed orientation of the sulfate group at C2 of the iduronic acid which then affects the orientation of the sulfamate group of the glucosamine ring. The back-calculated RDCs agree very well with the experimental data using the $^1C_4$ form (Table III.4.4.2.2), while larger differences were observed for the ^2S_O conformation. This is in agreement with the dominance of the $^1C_4$ conformer (76%) in the tetrasaccharide **II** as implied from the analysis of the $^3J_{HH}$ coupling constants.

Table III.4.4.2.2 Experimental and back-calculated RDCs of the C ring using CB disaccharide

RDC [Hz]	Expt.	$^4C_1$-$^1C_4$ (CB)		$^4C_1$-^2S_O (CB)	
		Cal.	Diff.	Cal.	Diff.
C_1-H_1	5.1	5.11	-0.01	5.14	-0.04
C_2-H_2	-3.1	-2.98	-0.12	-2.65	-0.45
C_3-H_3	-3.0	-3.09	0.09	-4.10	1.10
C_4-H_4	-4.8	-4.88	0.08	-4.39	-0.41
C_5-H_5	-3.6	-3.57	-0.03	-3.22	-0.38
H_1-H_2	-1.21	-1.22	0.01	-1.10	-0.11
H_2-H_3	0.55	0.46	0.09	0.03	0.52
H_3-H_4	-0.61	-0.48	-0.13	-0.32	-0.29
H_4-H_5	-1.36	-1.23	-0.13	-1.06	-0.30
RMSD		0.09		0.49	

In the case of the BA disaccharide, both conformers of the B ring gave very good agreement between the back-calculated and experimental RDCs of the A ring (Table III.4.4.2.3).

Table III.4.4.2.3 Experimental and back-calculated RDCs of the A ring using BA disaccharide

RDC (Hz)	Expt.	$^4C_1$ - $^1C_4$		$^4C_1$ - ^2S_O	
		Cal.	Diff.	Cal.	Diff.
C_1-H_1	-3.4	-3.40	0.00	-3.41	0.01
C_2-H_2	5.8	5.89	-0.09	5.89	-0.09
C_3-H_3	6.1	6.07	0.03	6.07	0.03
C_4-H_4	5.4	5.65	-0.25	5.59	-0.19
C_5-H_5	6.2	5.94	0.26	6.04	0.16
H_1-H_2	-0.82	-0.82	0.00	-0.77	-0.05
H_2-H_3	0.88	0.70	0.18	0.74	0.14
H_3-H_4	0.96	0.86	0.10	0.76	0.20
H_4-H_5	0.90	0.91	-0.01	0.43	0.47
RMSD		0.14		0.20	

This is despite the fact that large differences (up to 22°) were observed in the dihedral angles of the A ring in the BA disaccharide between the $^1C_4$ and ^2S_O conformations of the B ring, which presumably are a consequence of different conformations of the glycosidic linkage in these disaccharides.

There are several factors which can affect our analysis. We are using data acquired on a tetrasaccharide, but are using the disaccharide structures to interpret the RDCs. Therefore, for example, the conformation of ring C in the CB disaccharide might be affected by the ring D which is absent in the ab initio structure. Nevertheless, we have obtained good agreement for one form of the CB disaccharide. Perhaps surprisingly, it is the disaccharide BA which gave us an inconclusive result; despite the fact that ring A is at the reducing end in both di- and tetrasaccharides. Also the dihedral angles across the glycosidic linkage of the ab initio structures are not compatible with the measured coupling constants of the tetrasaccharide **II** and in the case of the BA disaccharide show large changes between the two conformations of the B ring. This observation puts the validity of our analysis into question and the ab initio tetrasaccharide structure (which is not available at present) needs to be analyzed in order to draw any conclusion. It remains to be seen if a detailed analysis of RDCs on the level of individual monosaccharides can be used to verify the ab initio structures of oligosaccharides.

III. 4.4.3 Restrained molecular dynamics with RDCs

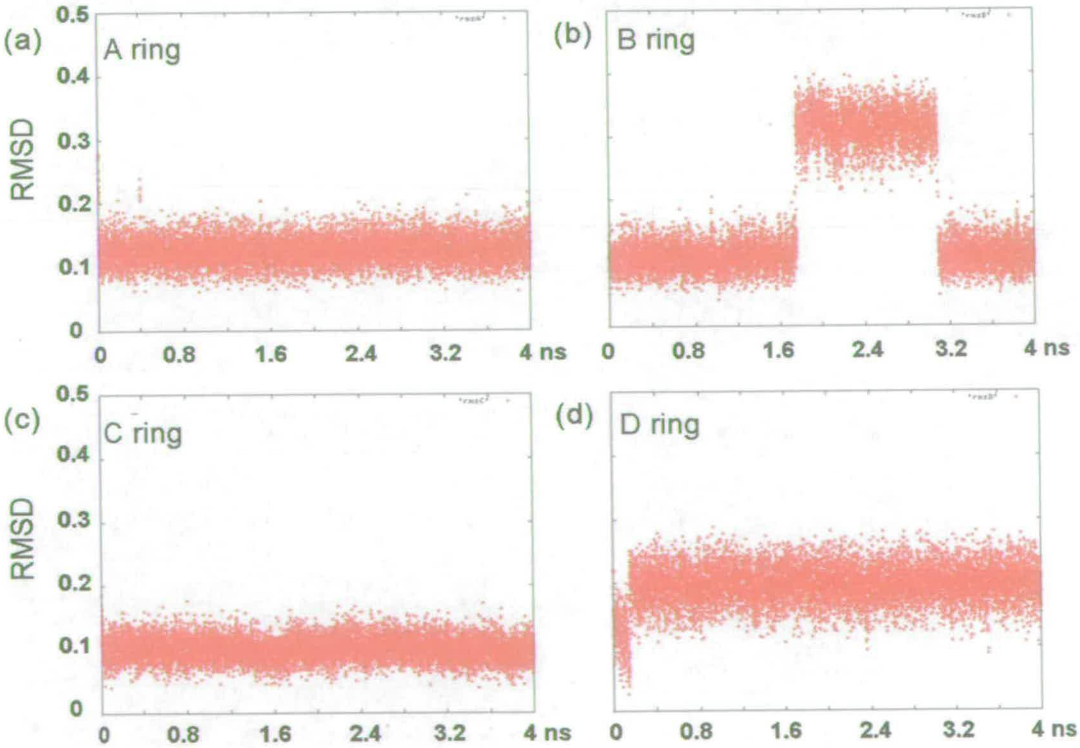
Selected $^1D_{CH}$ and $^3D_{HH}$ coupling constants were used to refine the conformation of the fully sulfated tetrasaccharide **II**. The glucosamine rings are stable in the $^4C_1$ chair form. In this conformation, all but the anomeric proton are in the axial position and therefore the CH RDCs sample only two independent orientations. However, at least five independent vectors are needed to define the alignment tensor. Therefore, $^3D_{HH}$ coupling constants played an essential role as they sample more directions. In the following, the analysis of RDCs is presented in an attempt to describe the relative orientations of individual monosaccharide rings of tetrasaccharide **II**.

This analysis is complicated as in the tetrasaccharide **II** only rings A and C are rigid chairs. They are surrounded by flexible B and D rings and connected by possibly flexible glycosidic linkages. Our analysis of RDCs is based on assumptions that (i) despite this flexibility the overall shape of the molecule does not change implying that it can be described by one alignment tensor, and (ii) the averaged RDCs can be interpreted by a single alignment tensor.

For this analysis we chose RDCs of rings A and C, which have rigid conformations on the monosaccharide level.

In our model, we thus assume that possible motions across the glycosidic linkages and the conformational exchanges of the B and D rings will only contribute to minor changes of the overall shape of the molecule. At the same time, we assume that the average measured RDCs can be attributed to an average alignment tensor reflecting the average molecule shape. The mutual orientation of the A and C rings does change only slightly, e.g. as a consequence of conformational transitions of ring B, allowing us to interpret RDCs using a single alignment tensor.

Eighteen RDCs from the A and C rings sampling twelve independent orientations were used as restraints during a 4 ns molecular dynamics using the protocol described in Chapter II.6.2.2. The alignment tensor was optimized during the RDC-restrained MD. These calculations require initial tensor parameters. These were calculated based on the free molecular dynamics structure that showed the best agreement between the experimental and back-calculated RDCs for individual rings. The qualitative analysis about the populations of each conformation can only be calculated from two ensembles of $^1C_4$ and ^2S_o forms generated during the dynamics. Results are shown in the next section.



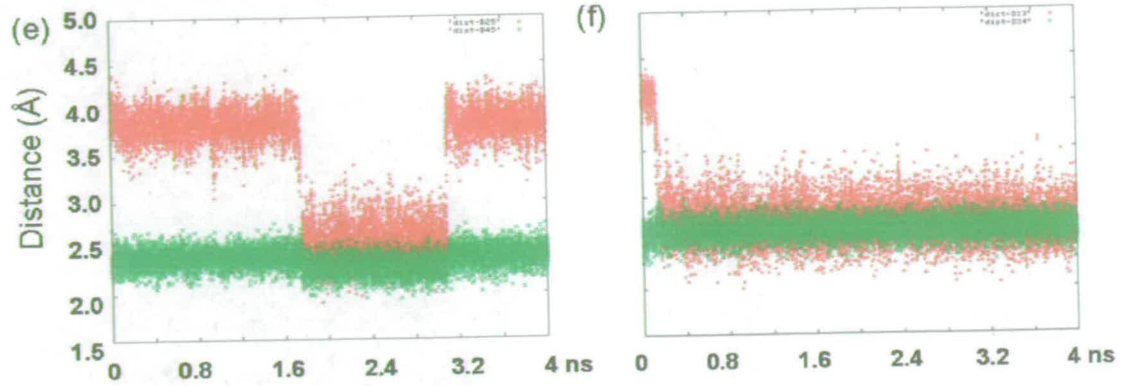
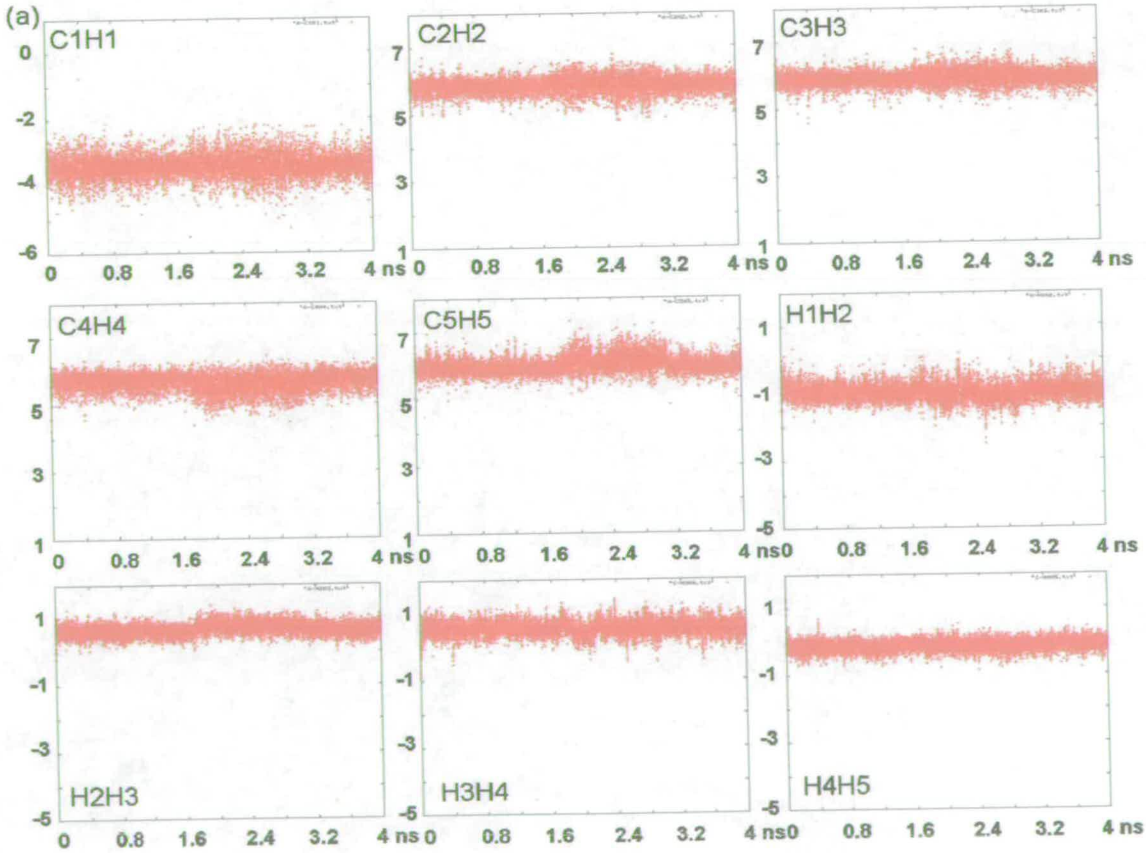


Fig.III.4.4.3.1 (a) to (d) RMSDs of the ring atoms for each monosaccharide; (e) H2-H5 and (f) H1-H3 distance change

All monosaccharide rings converged to good conformations. The two NHSO$_3^-$ groups of the A and C rings were placed on the opposite sides of the molecule ('side-side' geometry). During the 4 ns simulation, the B ring showed transitions between $^1C_4$ and ^2S_o forms, which is in agreement with the existence of the conformational equilibrium of the B ring in solution (Fig.III.4.4.3.1). Ring D showed only one transition from the initial $^1H_2$ to the $^2H_1$ form. The glucosamine rings stayed in $^4C_1$ conformations. The trajectories of the RDCs of the A and C rings during the restrained MD are shown in Fig.III.4.4.3.2.



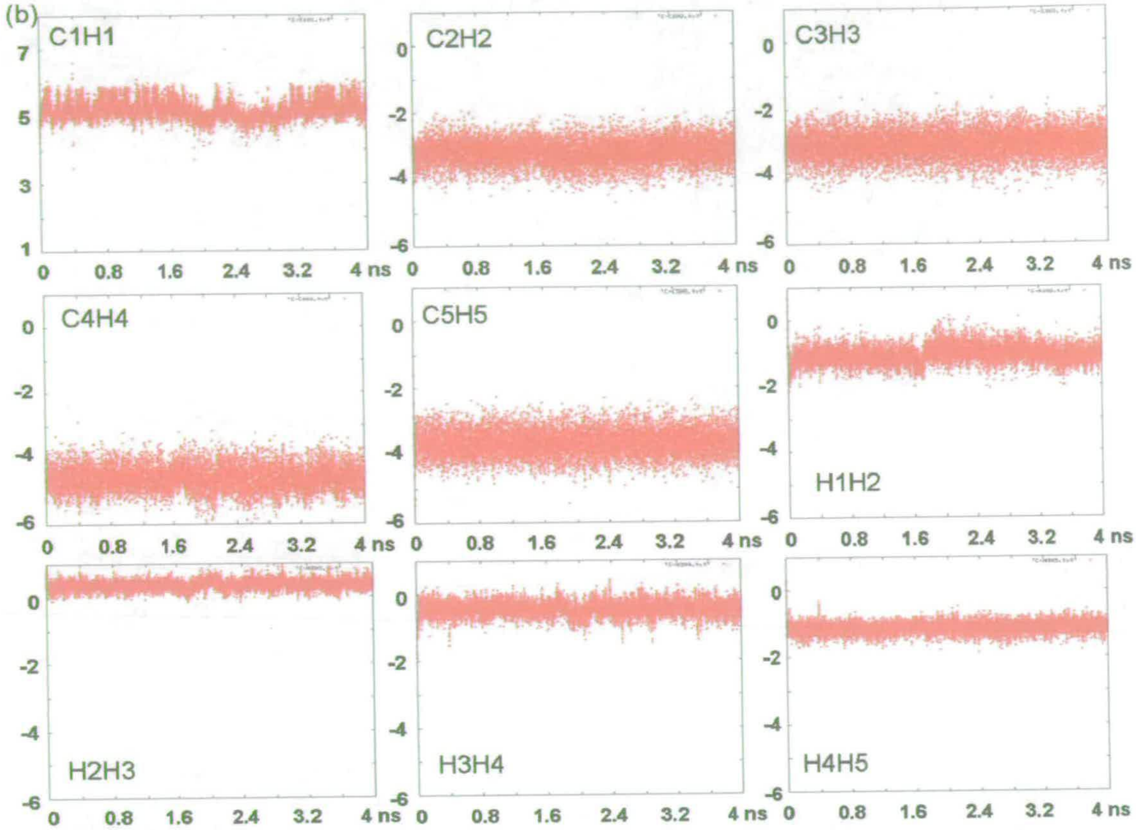
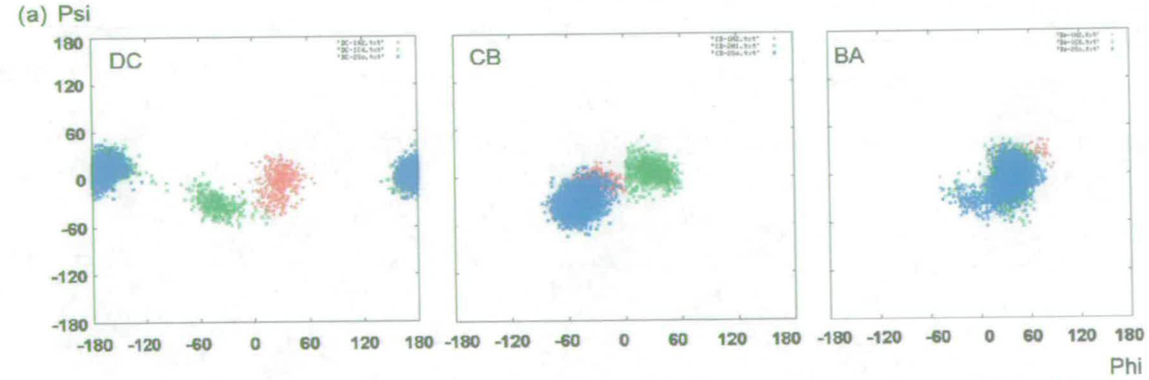


Fig.III.4.4.3.2 Back-calculated RDCs from restrained MD structures of (a) A and (b) C rings

The (Φ, Ψ) plots of all structures show that the dihedral angles converged to certain regions for all three glycosidic linkages under the constraints of RDCs. The orientations across the DC and CB linkages settled to one of the regions (Fig.III.4.4.3.3) found in the free MD.



$^1C_4$ & $^1H_2$	Phi: 27.7 +/- 11.5 Psi: -4.1 +/- 16.5	Phi: -35.0 +/- 12.9 Psi: -11.2 +/- 11.2	Phi: 42.4 +/- 12.7 Psi: 10.5 +/- 11.1
$^1C_4$ & $^2H_1$	Phi: -41.5 +/- 15.2 or -165.0 +/- 10.3 Psi: 3.9 +/- 22.0	Phi: 29.3 +/- 12.9 Psi: 5.6 +/- 11.2	Phi: 36.7 +/- 13.1 Psi: -3.2 +/- 18.1
$^2S_0$ & $^2H_1$	Phi: -168.4 +/- 9.0 Psi: 9.4 +/- 11.0	Phi: -50.8 +/- 12.3 Psi: -32.9 +/- 13.3	Phi: 36.0 +/- 16.2 Psi: -7.4 +/- 15.8

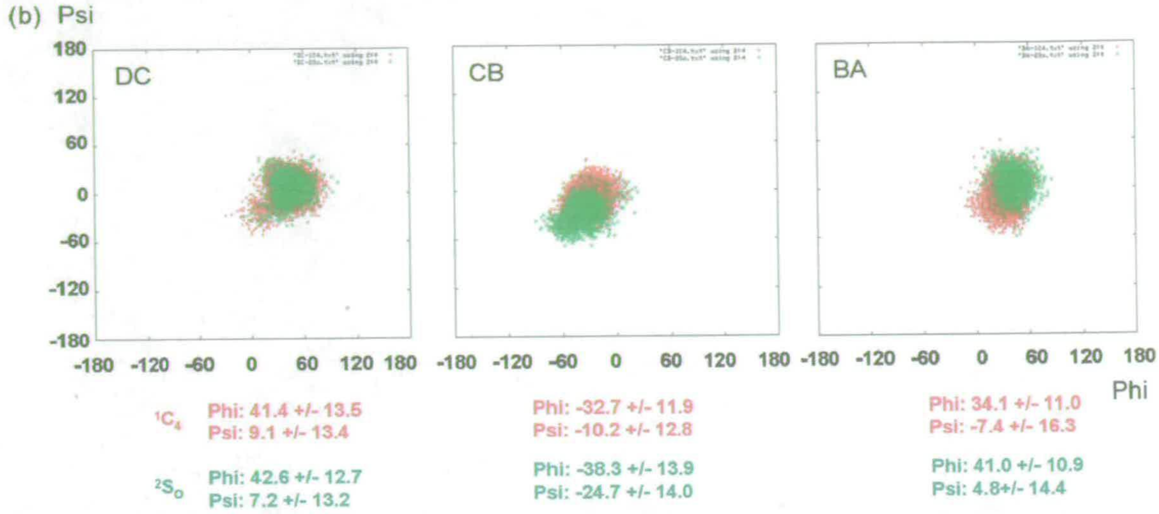


Fig.III.4.4.3.3 Comparison of the (Φ, Ψ) maps obtained in nonrestrained and RDC-restrained molecular dynamics. (a) shows (Φ, Ψ) dihedral angles across the glycosidic linkages from nonrestrained MD; (b) are results of RDC-restrained simulations.

This is reflected in reduced conformational heterogeneity of the tetrasaccharide as illustrated in the overlay of selected structures in Fig.III.4.4.3.4. 28 structures were selected based on the best fitted back-calculated RDCs of both A and C rings at the same time. Out of these, 21 structures showed the B ring with the $^1C_4$ form, while the other 7 of the B ring were in the ^2S_O form.

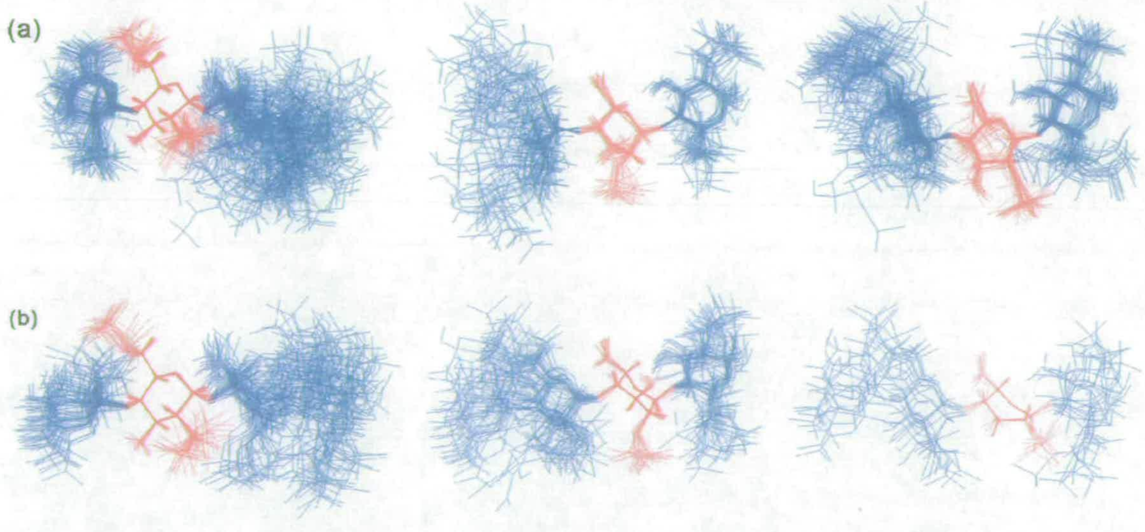


Fig.III.4.4.3.4 Comparison of structures obtained in nonrestrained and RDC-restrained molecular dynamics. (a) 50 best overlaid structures on the C ring from nonrestrained MD; 20 structures overlaid on the $^1C_4$ form of the B ring and 30 structures overlaid on the ^2S_O form of the B ring. (b) 28 structures overlaid on the C ring from RDC-restrained MD; 21 structures overlaid on the $^1C_4$ form of the B ring and 7 structures overlaid on the ^2S_O form of the B ring

The average dihedral angles across the glycosidic linkages are given in Table III.4.4.3.1. These are practically identical between the two forms with different conformations of B and D rings. This is in contrast with the unrestrained MD which showed heterogeneity of Φ angles. These agree well with the analysis of torsion angles based on the analysis of the $^{3}J_{CH}$ coupling constants (Table III.4.4.3.1).

Table III.4.4.3.1 Dihedral angles (in degrees) across the glycosidic linkages of the tetrasaccharides **II** in RDCs-restrained MD structures.

Tetrasaccharide	Φ_{DC}	Ψ_{DC}	Φ_{CB}	Ψ_{CB}	Φ_{BA}	Ψ_{BA}
RDC-restrained MD[a]	42.6 ± 12.7	7.2 ± 13.2	-38.3 ± 3.9	-24.7 ± 14.0	41.0 ± 10.9	4.8 ± 14.4
	41.4 ± 13.5	9.1 ± 13.4	-32.7 ± 11.9	-10.2 ± 12.8	34.1 ± 11.0	-7.4 ± 16.3
NMR analysis	30	± 20	-25	-20	35	± 20

[a] dihedral angles are the values from RDC-refined solution structures including 3200 $^{2}S_{O}$ (first row) and 5700 $^{1}C_{4}$ (second row) for the internal iduronic ring (B ring); all structures of the D ring are in the $^{2}H_{1}$ form.

III. 4.4.4 Flexibility of the iduronic acid ring as described by RDCs

During the 4 ns RDC-restrained MD dynamics, the A and C rings converged to a certain orientation reflecting their common alignment tensor. The RMSDs of the five carbon atoms on both rings are smaller than 0.2 Å (Fig.III.4.4.3.1). In the frame work of mutually restricted orientation of rings A and C, the B ring changes between the two forms $^{1}C_{4}$ and $^{2}S_{o}$. During this transition the distances between certain pairs of protons, e.g. H2 and H5, changed significantly (Fig.III.4.4.3.1), as was also seen in the unrestrained calculations. The dihedral angles across the CB and BA linkages, however, changed less than 14° (Table III.4.4.3.1) between the $^{1}C_{4}$ and $^{2}S_{O}$ forms. The structures were separated into two ensembles with the $^{1}C_{4}$ and $^{2}S_{o}$ conformations of the B ring. 3200 structures were in the $^{2}S_{o}$ form and 5700 in $^{1}C_{4}$ form. These two ensembles were used to back-calculate RDCs for each form individually and are shown in Fig.III.4.4.4.1. These RDCs were calculated using individual structures and corresponding alignment tensors generated during the simulation.

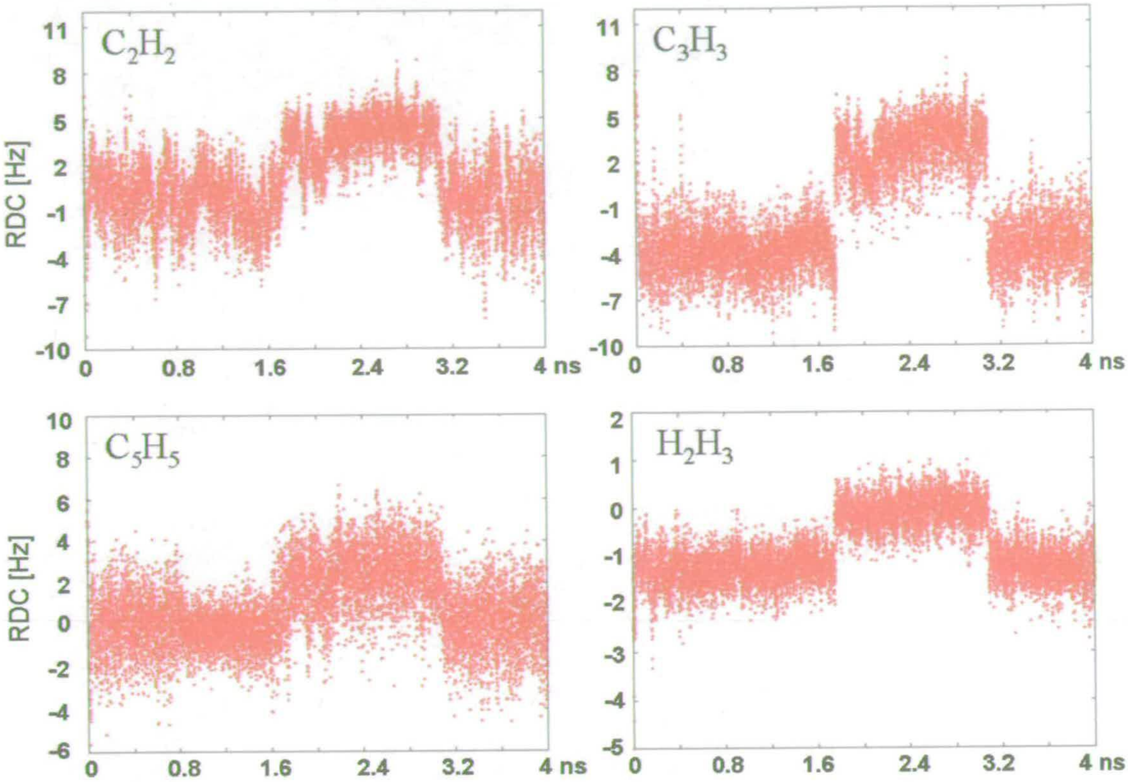


Fig.III.4.4.4.1 Back-calculated RDCs from restrained structures of the iduronic acid ring (B ring) in both forms of $^1C_4$ and ^2S_o

These plots show a large spread of values with average RMSDs of 1.6 and 0.35 Hz for $^1D_{CH}$ and $^3D_{HH}$ coupling constants, respectively. The average back-calculated RDCs of both conformers are shown in Table III.4.4.4.1. Here only C_2H_2, C_3H_3, C_5H_5 and H_2H_3 RDCs are presented as only these were significantly different between the $^1C_4$ and ^2S_O forms (Fig.III.4.4.4.1). These RDCs were used to determine the relative populations of the two forms by finding the best fit between the experimental and theoretical RDCs. The population of the ^2S_o form was found to be 18%. This qualitatively agrees with the analysis of the $^3J_{HH}$ coupling constants, where the population of the ^2S_o form was found to be 24%.

Table III.4.4.4.1 Population analysis of the $^1C_4$ and ^2S_o forms of the iduronic acid ring (B ring) from the back-calculated average RDCs

RDC (Hz)	C_2H_2	C_3H_3	C_5H_5	H_2H_3	wt%
$^1C_4$	-0.20	-3.72	-0.01	-1.27	82%
^2S_O	3.80	2.78	2.38	-0.07	18%
Expt.	1.1	-2.6	-0.4	-1.22	RMSD =0.51
Theory	0.52	-2.55	0.42	-1.05	

For the D ring (the terminal unsaturated non-reducing uronic acid) the RDC-restrained molecular dynamics shows that the ${}^{2}H_{1}$ form is more stable than the ${}^{1}H_{2}$ form. This contradicts the experimental analysis based on the ${}^{3}J_{HH}$ coupling constants, which shows that the ${}^{1}H_{2}$ form is the dominant conformer (78%). Meanwhile the NOE analysis shows 56% of the ${}^{2}H_{1}$ form. However, a similar analysis as for the B ring using the average C_1H_1, C_2H_2, H_1H_2 and H_2H_3 RDCs from the D ring showed 50% of each conformer, ${}^{1}H_{2}$ and ${}^{2}H_{1}$ (Table III.4.4.4.2) despite the fact that the average RMSDs of 1.5 and 0.35 Hz for ${}^{1}D_{CH}$ and ${}^{3}D_{HH}$ coupling constants, respectively, are comparable to these in the B ring. The quality of the fit is poor.

Table III.4.4.4.2 Population analysis of the ${}^{1}H_{2}$ and ${}^{2}H_{1}$ forms of D ring from the back-calculated average RDCs

RDC (Hz)	C_1H_1	C_2H_2	H_1H_2	H_2H_3	wt%
${}^{1}H_{2}$	-2.39	-2.08	0.10	0.59	50%
${}^{2}H_{1}$	2.09	1.86	0.68	-0.29	50%
Expt.	1.1	-1.4	0.47	0.85	RMSD =0.96
Theory	-0.15	-0.11	0.39	0.15	

In the free molecular dynamics distinct dihedral angles were found for different conformations of B and D rings. Fig.III.4.4.4.2 shows that almost the same conformational space is occupied by both forms of the B ring from RDC-restrained MD.

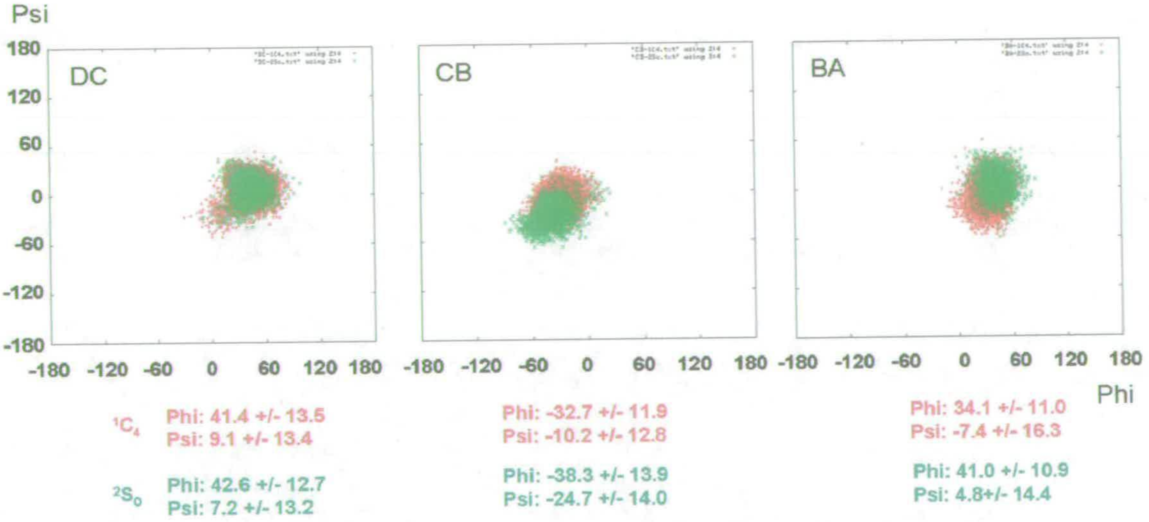


Fig.III.4.4.4.2 (Φ, Ψ) dihedral angles across the glycosidic linkages of BA, CB and DC rings of the RDC-restrained MD structures. Red dots are in the ${}^{1}C_{4}$ form (5700 structures) of the B ring, while the green ones are in the ${}^{2}S_{O}$ form (3200 structures).

The orientation of the alignment tensor can be represented by the Sanson-Flamsteed projection plot (Fig.III.4.4.4.3) showing the eigenvectors of the alignment tensor parameters (S_{xx}, S_{yy} and S_{zz}), which represent the directions of the principal axis in the molecular coordinate system. Using the minimal difference between the experimental and back-calculated RDCs of the A and C rings a small set of structures was chosen to calculate the alignment tensor. These structures were divided into two groups depending on the conformation of B ring. Both of these tensors show large rhombicity. Their two largest components (S_{zz} and S_{yy}) differ by less than 10%. In the $^{1}C_{4}$ conformer the two largest components swapped occasionally as seen in Fig.III.4.4.4.3d circled in colors pink or blue. As shown in Fig.III.4.4.4.4 one of the two large axes is nearly parallel to the long tetrasaccharide axis. The other large component reflects the presence of bulky sulfate groups on two sides of the molecule. The average alignment tensor parameters in the $^{2}S_{O}$ forms (S_{xx} = -1.44e-05, S_{yy} = 1.43e-04 and S_{zz} = 1.58e-04) are slightly larger than those of the $^{1}C_{4}$ forms (S_{xx} = -0.70e-05, S_{yy} = 1.36e-04 and S_{zz} = 1.44e-04).

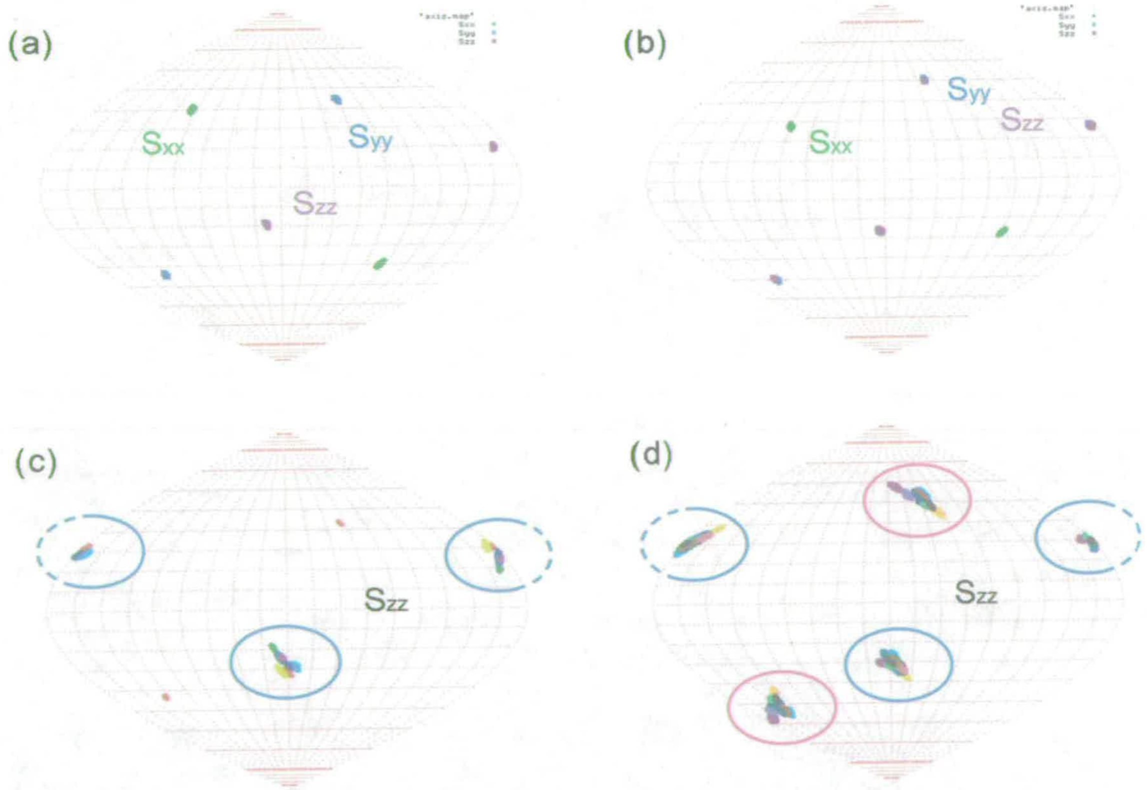


Fig.III.4.4.4.3 Sanson-Flamsteed projection plots of the eigenvectors constituting the principal axis (a) best fitted structure with the B ring in the form of $^{2}S_{O}$; (b) best fitted structure with the B ring in the form of $^{1}C_{4}$; (c) the largest components (S_{zz}) for 7 structures with the B ring in the form of $^{2}S_{O}$; (d) the largest components (S_{zz}) for 21 structures with the B ring in the form of $^{1}C_{4}$

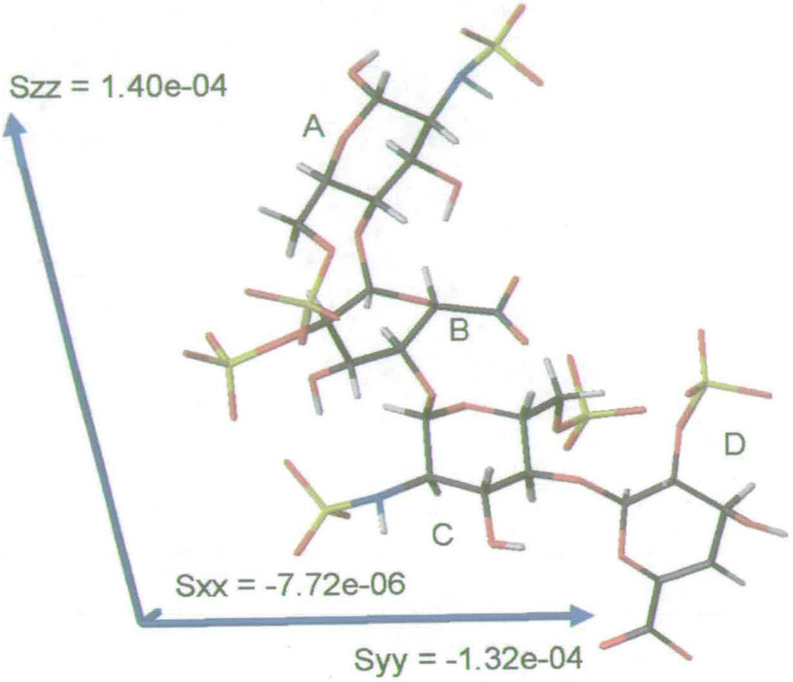


Fig.III.4.4.4.4 Eigenvectors constituting the principal axes for the best fitted structure (a) in Fig.III.4.4.4.3

III. 4.5 Conclusions and future directions

The heparin-derived fully sulfated tetrasaccharide **II** was refined using RDCs. This refinement yielded structures that are compatible with ${}^{3}J_{CH}$ coupling constants across the glycosidic linkages. They are also broadly compatible (less than 20° difference) with the CB and BA glycosidic dihedral angles of the heparin solution structure from [117]. As the refinement was performed using certain assumptions, it would be desirable to verify the conformations of glycosidic linkages by measuring the J_{CC} coupling constants. The effect of different substituents (e.g. compound **III** and **IV**) on the conformations of monosaccharide rings and the glycosidic linkages can be further studied. Any effects of the degree of polymerization on the orientation of the whole molecule through establishing the dihedral angles across the glycosidic linkages can be further studied using hexa- or even polysaccharides.

Chapter IV

The protein-heparin interactions

Heparan sulfate (HS) and heparin play important roles in a variety of biological functions. These include protein folding, activation of cell-surface receptors, participation in cell-cell recognition and immunological recognitions [2]. These processes are mediated by HS-protein interactions. Therefore understanding of protein-heparin interactions at the molecular level is important for studying physiological and pathophysiological processes and the design of new therapeutics. Factor H, constituted of 20 complement control protein modules (CCPS), is responsible for the host recognition via controlling the alternative pathway of complement activation [118], which defends organisms against invaders. This regulation occurs via factor H binding to cell surfaces which contain polyanions. Factor H interacts with polyanions, such as HS and related oligosaccharides, through CCP modules 7, 13 and 20 [119]. Structures of modules 7 and 19, 20 of fH were solved by Dr. A. Herbert in Prof. Barlow's group in Edinburgh. These were used to model the protein-GAG interaction. Docking of heparin-derived fully sulfated tetrasaccharide, **II**, to factor H modules provided initial identification of binding sites prior to experimental studies.

IV. 1 Docking of heparin-derived fully sulfated tetrasaccharide to fH modules

All dockings were performed using AUTODOCK 2.4 [120] (autodock tools or ADT). Autodock calculates the binding energy using grid maps, which show the potential energy of individual atoms or functional groups of the ligand as well as the electrostatic potentials of the protein. The configurations of the ligand on the surface of the macromolecule are explored and their binding energies are calculated. In this way, the ligand with a random initial conformation searches a defined region of the macromolecule. By adopting different geometries through rotations around the flexible bonds numerous conformations at the binding sites are considered.

Solution structures of heparin-derived fully sulfated tetrasaccharide **II** with different starting monosaccharide conformations, $^1H_2$-$^4C_1$-$^1C_4$-$^4C_1$, $^2H_1$-$^4C_1$-$^1C_4$-$^4C_1$ and $^2H_1$-$^4C_1$-^2S_O-$^4C_1$, were used as ligands for docking to fH modules. The above three conformations were chosen due to their presence in the free molecular dynamics calculation in explicit water. The NMR-refined solution structures of fH~19, 20 and fH~7 (both Y402 and H402 isoforms) of factor H were provided by Dr. Herbert.

The ligand was first prepared by transferring the AMBER pdb format to the Sybyl Mol2 format in order to keep the exact partial atomic charges for each atom before loading it into the ADT program. ADT automatically merged non-polar hydrogens to heavy atoms. The root fragment was set to be the six-member C ring since its conformation is identical in all considered structures. During the whole docking process, all monosaccharide rings were fixed at their starting conformations, while the ring substituents were defined as flexible (although they could be made rigid during the docking). Altogether 27 torsion angles were set to be active and could rotate freely. The rotatable parts included all of the substituents of the ring atoms, i.e. 5 sulfate groups, 2 carboxyl groups, all 3 glycosidic linkages, 2 C5-C6 bonds and 2 C2-N bonds of the glucosamine rings, and 5 hydroxyl groups (Fig.IV.1.1). Allowing a significant degree of freedom to the carbohydrate resulted in a variety of conformations of the bound sugar. Therefore, the docking results reflect the specific conformations of the individual monosaccharide rings.

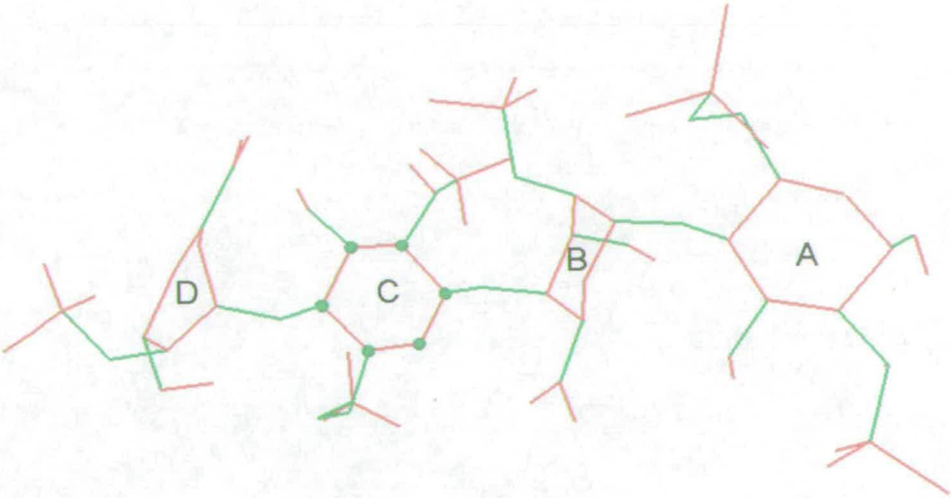


Fig.IV.1.1 Definition of the rigid (red) and rotatable (green) bonds of the tetrasaccharide **II** ligand ΔUA(2S)-GlcNS(6S)-IdoA(2S)-GlcNS(6S) (D-A rings).

IV.1.1 Docking of the tetrasaccharide II to fH~19, 20

fH~19, 20 was prepared by adding Kollman charges and solvation parameters automatically using ADT. A volume of (120, 120, 120) grid points with 0.7 Å spacing was used. The protein was put in the center of a cube large enough to cover the whole surface of the protein. The tetrasaccharide ligand was placed randomly into the box.

Two binding sites were found, which contained tetrasaccharides with both $^1C_4$ and ^2S_O conformations of the iduronic acid B ring. Fig.IV.1.1.1 shows the result of 250 docking calculations. 85% of the tetrasaccharides bound to the C-terminus of module 20 interacting mainly with residues Arg1206 and Arg1231 and 15% appeared to bind close to the linker region between the modules 19 and 20 interacting primarily with residues Lys1186 and Lys1188. These four amino acids interacted with the sulfate and carboxyl groups of the tetrasaccharides. Binding conformations with lowest docking energies representing typical binding modes are shown in Fig.IV.1.1.2.

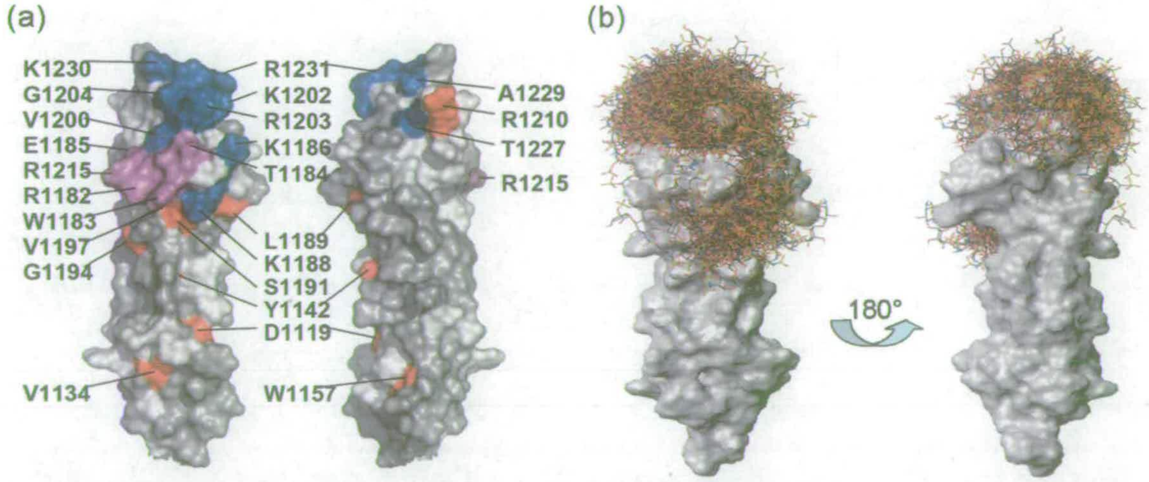


Fig.IV.1.1.1 Surface map of fH~19, 20. (a) two views of surface representation [121], red = disease-associated mutations where the substituted residue does not undergo chemical shift perturbations upon heparin-binding; purple = both chemical shift perturbation occurs and these residues have been linked to aHUS; blue = significant chemical shift perturbations, where substitutions of these residues have not so far been linked to aHUS; pink = amino acid residues for which no chemical shift perturbation information could be obtained. (b) Two binding sites of tetrasaccharide on fH~19, 20 obtained by AUTODOCK.

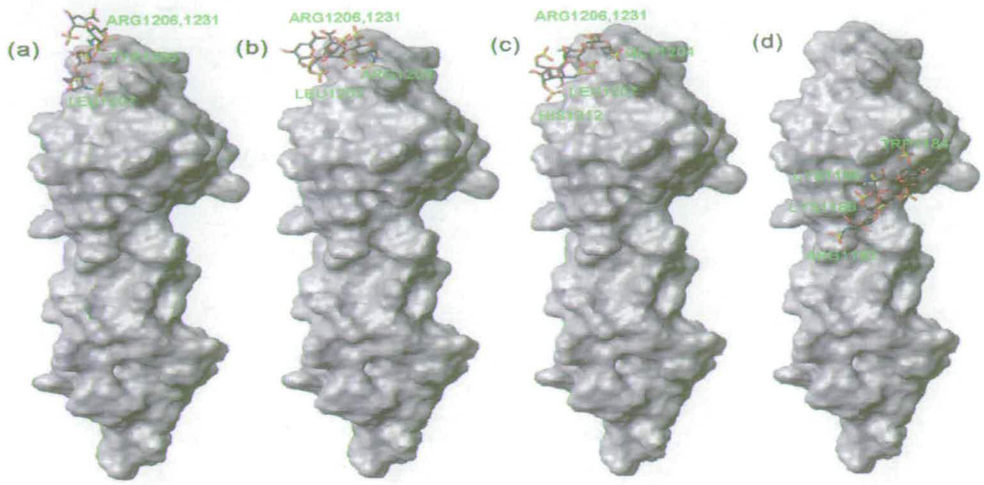


Fig.IV.1.1.2 Docking of heparin-derived fully sulfated tetrasaccharide to fH~19, 20. (a-c) different binding conformations of the tetrasaccharide at the main binding site at the top of module 20; (d) the second binding site close to the linker between modules 19 and 20.

The binding site in module 20 was centered on Arg1206 and Arg1231 (Fig.IV.1.1.2a,b,c). This could be further extended to Arg1203 (Fig.IV.1.1.2b) or His1212 (Fig.IV.1.1.2c). This suggests that if the ligand is long enough, the oligosaccharide should be able to cover a larger surface. In all of the three binding conformations at least three sulfate or carboxyl groups interacted with the basic residue and A-6S (sulfate group on carbon 6 of ring A) and C-NS (sulfamate on carbon 2 of ring C) were always actively interacting with Arg1206 or 1231. At the binding site close to the linker C-NS, A-6S, B-2S and the two carboxyl groups (B-COO^- and D-COO^-) interacted with Lys1186 and Lys1188, while A-NS interacted with Arg1192 and D-2S with Typ1183 (Fig.IV.1.1.2d). These docking results are similar to those of Gordon et al., which concluded that at least four positively charged residues interact with the negatively charged sulfate and carboxyl groups of fH [122].

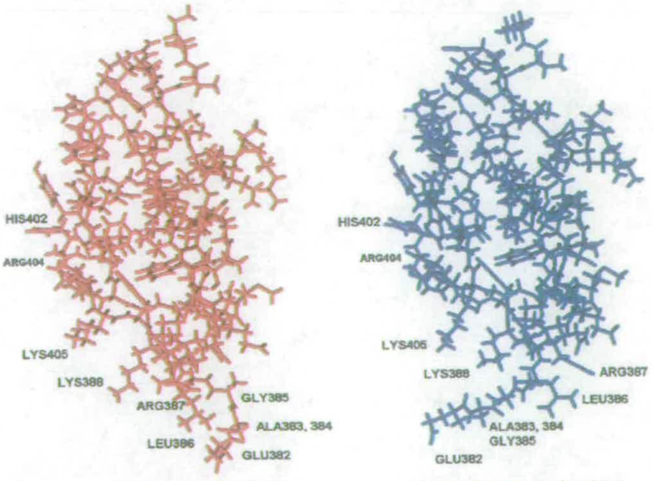
The docking results agreed well with the NMR chemical shift mapping of the bound fH~19, 20 by Dr. Herbert [121]. These studies found that the binding sites contained residues Arg1182, Lys1186 and Lys1188 in the hypervariable loop, the surrounding basic residues of Lys1202, Arg1203 and Arg1215, and the C-terminal residues of Lys1230 and Arg1231. Their chemical shifts changed more than 0.025 ppm between the free and bound forms. The chemical shift changes of Tyr1205 and Arg1206 were 0.015 and 0.022 ppm, respectively, while unambiguous assignment of Gly1204 was not available. The docking results found two binding sites containing residues Arg1231 and (Lys1186, Lys1188) were found. However, some extra surrounding residues, Gly1204, Tyr1205 and Arg1206 other than Lys1202 and

Arg1203, showed contacts with the ligand. This suggests that these residues should be involved in the binding. The two binding sites obtained from docking mainly cover the blue region (Fig.IV.1.1.1) but not the purple one, even though residue Trp1184 does interact with the sulfate group.

IV.1.2 Module 7 of protein factor H

Two NMR structures of fH~7 were used for docking; the only difference was in the residue 402, which was either tyrosine (fH~7Y) or histidine (fH~7H). This isomerism has been recently linked with age-related macular degeneration (AMD), a leading cause of irreversible visual impairment in the elderly population of western countries [123]. People who are homozygous for an at-risk allele in the fH gene are 7 fold more likely to develop AMD, and heterozygous people are 2.6 times more susceptible [124]. The at-risk allele encodes a His rather than Tyr at position 402. Experimental evidence is emerging showing that the two forms differ in their binding to GAGs. Experimental studies of the bindings of both forms to GAGs are carried out in our lab. In this project we studied the binding using docking.

The lowest energy and closest to mean structures were the same for fH~7Y. But for fH~7H they were different as shown in Fig.IV.1.2.1. The differences were mostly confined to the N-terminal region of the protein. It was therefore decided to use three structures for docking, one for fH~7Y and two for fH~7H. A volume of (100, 100, 100) grid points with 0.6 Å spacing was used for the whole surface exploration of fH~7. Otherwise, identical docking procedures and starting tetrasaccharide conformations were used as for fH~19, 20. The modeled structures of complexes for all three proteins are shown in Fig.IV.1.2.2.



lowest energy structure

closest to mean structure

Fig.IV.1.2.1 NMR-refined fH~7H [125] with labeled residues in different geometries mainly on the N-terminus

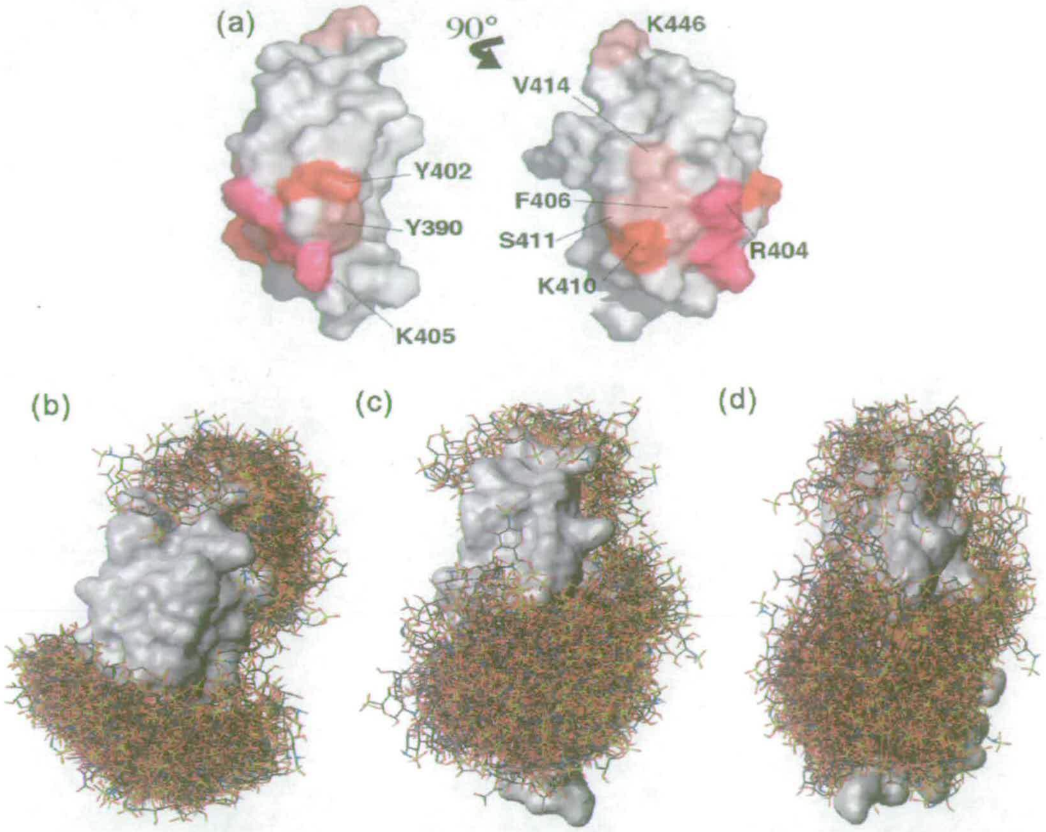


Fig.IV.1.2.2 Surfaces of fH~7s. (a) Two views [125] of surface representation of fH~7Y; the side-chains of 10 residues that undergo the largest (combined) chemical shift perturbation are highlighted in pink or purple. Arg404 and Lys405 (in purple) exhibit large chemical shift perturbations. Tyr402 (exhibits 11th biggest CSP) and Lys410 (important according to mutagenesis but exhibiting no significant chemical shift perturbation) are colored red; binding of the tetrasaccharide to (b) fH~7Y (c) fH~7H (lowest energy structure) (d) fH~7H (closest to mean structure)

Different monosaccharide conformations gave similar binding results for fH~7Y. Two binding sites were found with 50% of tetrasaccharides binding to each site. Fig.IV.1.2.3 shows the conformations of the tetrasaccharide in the binding sites. In the first binding site, Lys388 and Lys405 were the most active residues interacting with at least four negatively charged sulfate or carboxyl groups (Fig.IV.1.2.3a, b, c) always including A-6S, D-COO$^-$ and B-2S. The rest of the negatively charged groups occasionally interact with the protein. These included interactions with extra residues, e.g. Leu386, Arg387, Arg404 and Phe406. In the second binding site located at the C-terminal (Fig.IV.1.2.3), Arg441, Arg444 and Lys446 were important for binding with the tetrasaccharide.

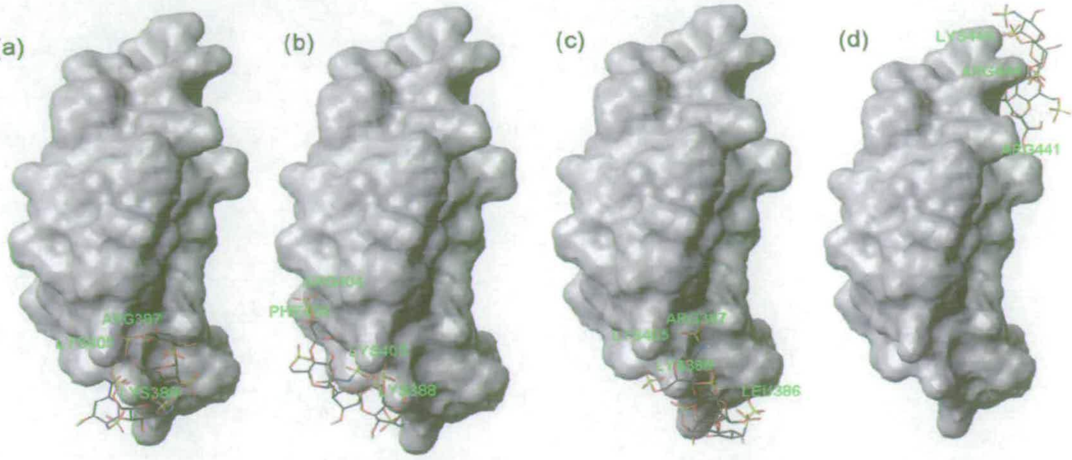


Fig.IV.1.2.3 Docking of heparin-derived fully sulfated tetrasaccharide to fH~7Y. (a-c) different binding conformations of the tetrasaccharide at the main binding site around Arg388 and Lys405; (d) the second binding site close to the Arg444

The results of the docking to fH~7H were partially different when compared to the results of the binding to fH~7Y. Firstly, the C-terminal binding site was less well defined as indicated by a long spread of tetrasaccharide structures in this region. This was the case for both the closest to mean and the lowest energy structures of fH~7H. Secondly, differences were seen between the closest to mean and the lowest energy structures in the main binding site in fH~7H. In the lowest energy structure, extra conformations were found (Fig.IV.1.2.4 and Fig.IV.1.2.5) characterized by the involvements of extra residues Tyr390 and Arg441 in the binding. These may be caused by slight differences in the protein surface due to the replacement of Tyr by His at residues 402.

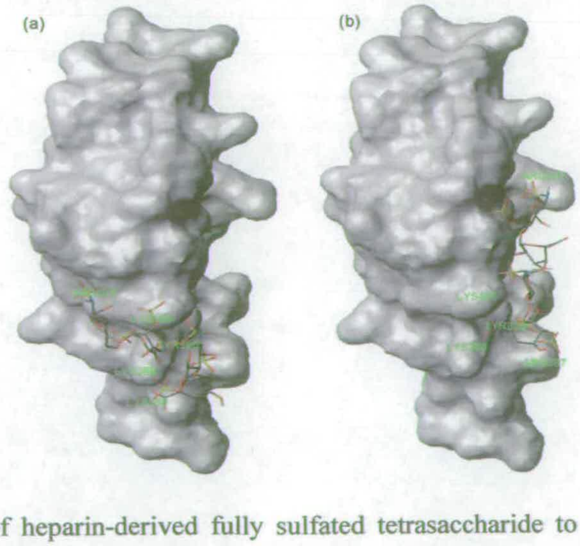


Fig.IV.1.2.4 Docking of heparin-derived fully sulfated tetrasaccharide to the NMR-refined lowest energy structure of fH~7H. (a) tetrasaccharide with $^{1}H_{2}$ of the terminal unsaturated uronic acid ring; (b) tetrasaccharide with $^{2}H_{1}$ conformation, showing extra interaction with residue Arg441

Different conformations of the N-terminus in the closest to mean structure of fH~7H (Fig.IV.1.2.1) gave extra binding possibilities as shown in Fig.IV.1.2.5 in addition to the above binding conformations. This additional mode of binding was mainly associated with the Glu382 residue. Their residues provide a binding surface and interact with the negatively charged sulfate groups together with Lys388 and Lys405.

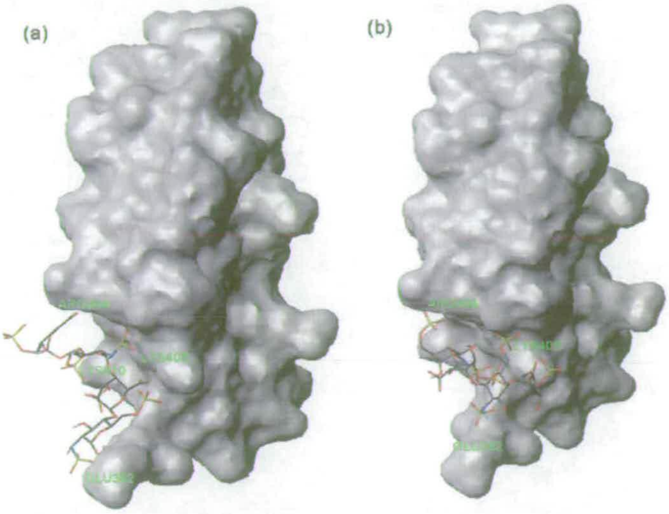


Fig.IV.1.2.5 Docking of heparin-derived fully sulfated tetrasaccharide to the NMR-refined closest to mean structure of fH~7H. These extra binding conformations are associated with the different geometry of Glu382 as shown in Fig.IV.1.2.1

Our docking results indicate that at least three positively charged amino acids of either fH~7 or fH~19, 20 are essential for the binding with the tetrasaccharide. Changes of the protein conformation will affect the binding mode as was seen in fH~7H structures differing in the orientation of the N-terminus. In accord with the observation of others (B. Mulloy, personal communication), the results of AUTODOCK simulations do not give an accurate picture of the conformation of the bound carbohydrate. They can only locate potential binding sites. As anticipated, the tetrasaccharides adopted different conformations; this was the case for both binding sites. They showed changing geometries of the glycosidic linkages and the assignment of rotatable sulfate groups, but also different directionality of the binding. This study also showed that allowing too much freedom in carbohydrate conformation will produce multiple binding modes. Perhaps freezing the geometry of the glycosidic linkages at values found in crystal structures of the GAG-protein complex will provide more accurate identification of crucial binding interactions.

IV. 2 Spin-labeled heparin-derived fully sulfated disaccharide

Structure determination of the protein-GAG complex is a challenging task. The interactions are often weak and the bulky sulfate groups keep the GAG relatively far from the protein surface. Because of this, the intermolecular NOEs between the protein and GAG protons are difficult to detect. In the lack of such NOEs, the three dimensional coordinates of the protein-GAG complex cannot be determined by standard NMR methods. Chemical shift mapping, the main experimental technique used for locating the binding site, or theoretical docking studies can characterize the binding sites only qualitatively. For example, the directionality of the binding as shown in Fig.IV.2.1a, b is unknown. If a paramagnetic tag can be attached to the ligand then its exact binding position can be determined (Fig.IV.2.1c, d). The tag could be a spin-labeled molecule, which causes enhanced relaxation of nuclei in its vicinity. In a complex, effects of a spin-label attached to a ligand can be felt by the protein protons, e.g. NH protons. Ideally, two tags attached to the reducing and nonreducing ends of the molecule should be investigated in two separate experiments, triangulating the ligand accurately on the protein surface.

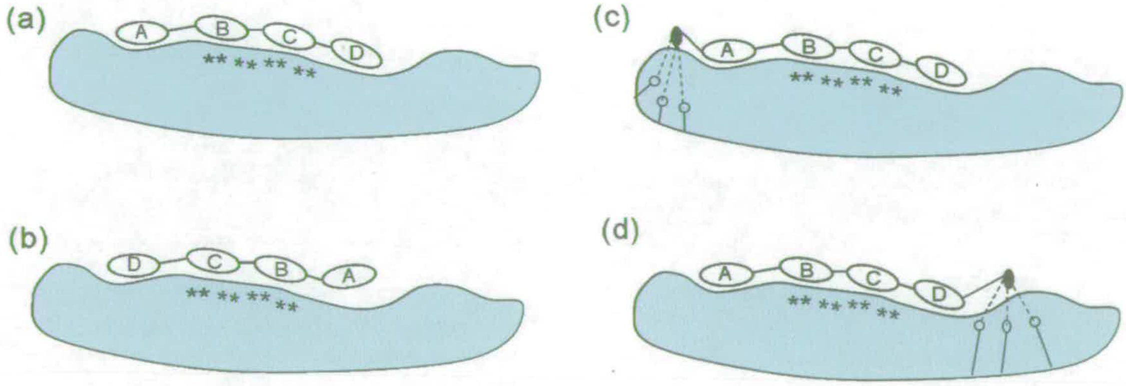


Fig.IV.2.1 Protein-oligosaccharide binding using paramagnetic tags. Asterisks represent the active residues. (a) and (b) show two possible binding modes. (c) and (d) show the NH protons (○) affected by the paramagnetic tag.

As an initial step in the development of this methodology, a 4-amino-TEMPO labeled disaccharide (compound **V**, Fig.IV.2.2) was prepared [65] in collaboration with Alison Hulme (Edinburgh). Below the results of an initial study of a spin-labeled disaccharide are presented.

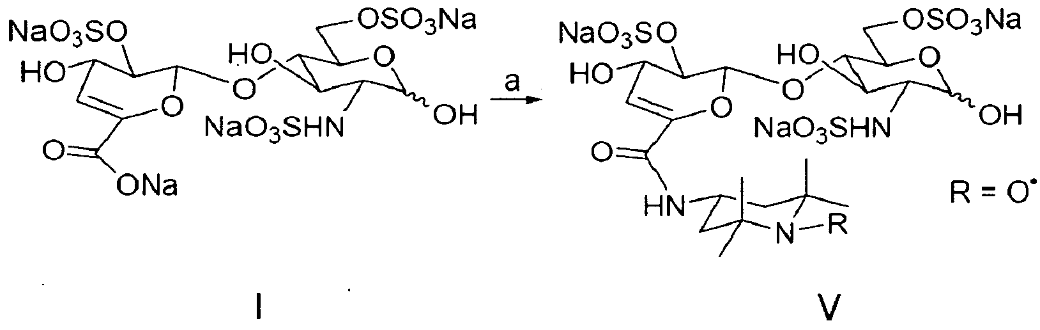


Fig.IV.2.2 Preparation of spin-labeled disaccharide using 4-amino-TEMPO

IV. 2.1 Primary structure determination of compound V by MS and NMR

Mass spectra of compound V were obtained using a ZMD mass spectrometer. Experimental masses for the major species observed are listed in Table IV.2.1.1.

Table IV.2.1.1 Major peaks observed in MS of the spin-labeled disaccharide V

Observed (m/z)	Calculated (m/z)	Assigned species
773.8	773.8	$[C_{21}H_{33}N_3O_{19}S_3Na_2^-]^\bullet$
729.2	729.8	$[C_{21}H_{35}N_3O_{19}S_3^-]^\bullet$
375.2	375.4	$[C_{21}H_{33}N_3O_{19}S_3Na^{2-}]^\bullet$
364.2	364.4	$[C_{21}H_{34}N_3O_{19}S_3^{2-}]^\bullet$
242.8	242.6	$[C_{21}H_{33}N_3O_{19}S_3^{3-}]^\bullet$

An NMR sample was prepared by dissolving 200 μg of compound V in 500 μl D_2O. The 600 MHz ^1H NMR spectrum of compound V is shown in Fig.IV.2.1.1 and the resonance assignment is summarized in Table IV.2.1.2.

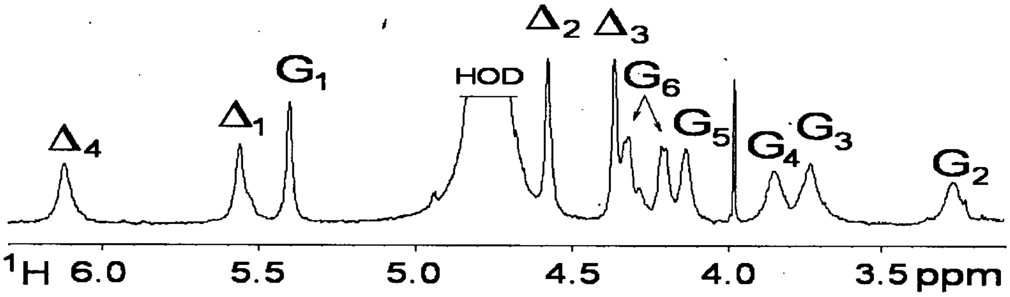


Fig.IV.2.1.1 600 MHz ^1H spectrum of compound V (4-amino-TEMPO-ΔUA2S-GlcNS6S). A significant broadening of resonances indicates the presence of the spin-label.

Table IV.2.1.2 (a) ^{1}H Chemical shifts of compound **V**. Literature [89c] values of disaccharide are shown in bold and the differences are shown in parentheses

	H_1	H_2	H_3	H_4	H_5	H_6	$H_{6'}$
GlcNS6S	5.41	3.27	3.74	3.85	4.14	4.21	4.33
	5.45	**3.28**	**3.76**	**3.83**	**4.16**	**4.21**	**4.35**
	(0.04)	(0.01)	(0.02)	(-0.02)	(0.02)	(0.00)	(0.02)
ΔUA2S	5.56	4.58	4.37	6.12	-	-	-
	5.50	**4.57**	**4.34**	**5.97**			
	(-0.06)	(-0.01)	(-0.03)	(-0.15)			

The MS and NMR data confirm the primary structure of compound **V**. The ^{1}H chemical shifts of compound **V** are practically identical to those of the free disaccharide with the exception of ΔUA2S proton H4, whose chemical shift is affected by the TEMPO substitution. ^{1}H signals of compound **V** are considerably broadened by the effects of the radical. The non-selective ^{1}H T_1 relaxation times of the compound were measured and the results are listed in Table IV.2.1.3.

Table IV.2.1.3 The non-selective ^{1}H T_1 relaxation times of compound **V** in ms.

	H_1	H_2	H_3	H_4	H_5	H_6	$H_{6'}$
GlcNS6S	84	27	42	32	59	77	84
ΔUA2S	43	140	126	35	-	-	-

In order to interpret these results, the structure of compound **V** was modeled using AMBER.

IV. 2.2 Molecular modeling

The model of compound **V** was built within XLEAP by adding 1-oxyl-4-carboxamido-2,2,6,6-tetra-methyl-piperidine (4-amino-TEMPO) to the carboxyl group of the uronic acid of the previously refined structure of the fully sulfated heparin-derived disaccharide (**I**). A representative modeled structure of compound **V** is shown in Fig.IV.2.2.1. Dihedral angles, (Φ, Ψ), across the glycosidic linkage of compound **V** are compared with those observed for the disaccharide fragment in the context of a tetrasaccharide in Table IV.2.2.1.

Table IV.2.2.1 Comparison of the dihedral angles across the glycosidic linkage of compound **V** and the disaccharide fragment of the tetrasaccharide **II**

	D Ring	A Ring	Φ (degree)	ψ (degree)
Compound V	$^2H_1$	$^4C_1$	51.0	4.8
X-ray [86] tetrasaccharide	$^1H_2$	$^4C_1$	42.3	18.3
Solution [85] tetrasaccharide	$^1H_2$	$^4C_1$	45.7	13.2

Electron-proton distances between the spin-label and protons of the disaccharide obtained from the modeled structure are listed in Table IV.2.2.2. The closer the proton of the disaccharide is to the radical the shorter will be its T_1 relaxation time. It is expected that similar effects will be observed for protein protons if they are part of the binding site in the GAG-protein complex, which will provide a better decision of the complex than chemical shift mapping.

Table IV.2.2.2 Electron-proton distances between the 4-amino-TEMPO and protons of the disaccharide in Å

	H_1	H_2	H_3	H_4	H_5	H_6	$H_{6'}$
GlcNS6S	8.1	6.3	8.4	7.7	10.1	11.11	11.14
ΔUA2S	9.3	10.2	10.8	9.3	-	-	-

The modeled structure shows a good qualitative agreement between the proton-electron distances and the measured relaxation times. Two CH protons with the fastest relaxation times (H2' and H4' of the glucosamine ring) show the shortest distances, while the most distant protons (H2 and H3 of the uronic acid ring) relax most slowly (Fig.IV.2.2.1).

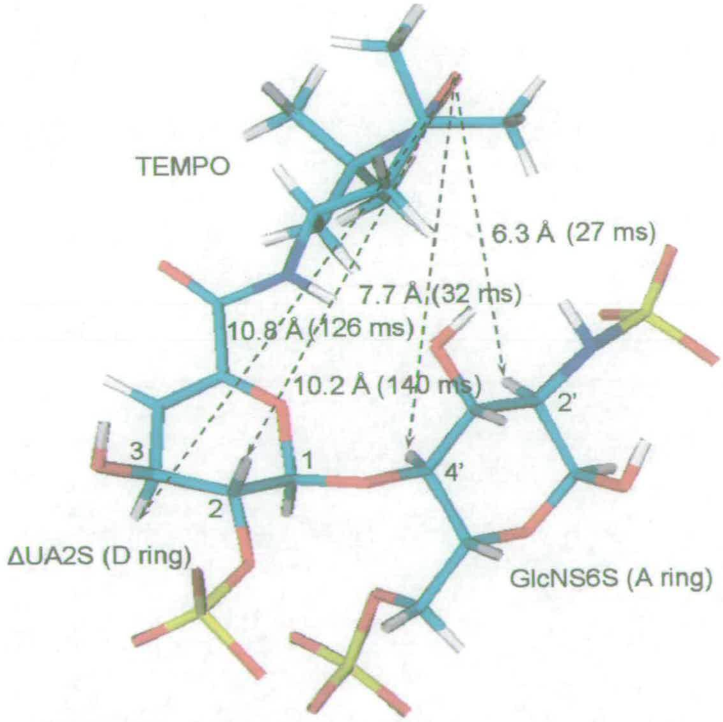


Fig.IV.2.2.1 The lowest energy structure of compound **V** (4-amino-TEMPO-ΔUA2S-GlcNS6S, ${}^2H_1$-${}^4C_1$). The dihedral angles across the glycosidic linkage are within 10° of those observed for this disaccharide fragment in solution or X-ray structures of a tetrasaccharide.

As the next step, a restrained molecular dynamics study using the relaxation data will be performed so that such a model of compound **V** could be used to study the complex. Our collaborators are currently working on the preparation of a tetra- or hexasaccharide labeled separately at both nonreducing and reducing ends. Once the labeled oligosaccharides are available, the protein-heparin complex will be studied. We will start with protein NK1 (part of hepatocyte growth factor), which binds GAG strongly ($K_D \sim 10^{-9}$ M) and later extend the methodology to more weakly binding proteins such as factor H ($K_D \sim 10^{-6}$ M).

In principle, the radical-proton distances can also be implied from the relaxation measurements of carbohydrate protons in the complex and thus provide the bound conformation of carbohydrates. This is, however, not straightforward, and would require the use of double labeled protein and the X-filtered technique. Further complications would come from averaging between the free and bound forms of the ligand. Therefore, at present the main interest lies in using the labeled GAGs to characterize the binding site.

Conclusion

New methodologies for the conformational study of free and bound GAGs in the gas phase and solution were developed and tested in this project.

Gas phase conformations of heparin-derived oligosaccharides were obtained through interpretation of the IMMS experimental data using molecular modelling. Very good agreement was shown between the experimental and theoretical structures. The more compact conformations of the heparin oligosaccharides in the gas phase were attributed to the effects of sodium cations interacting with the negatively charged sulfate and carboxyl groups in the oligosaccharides. Formation of higher oligomers of the disaccharide, observed in the experiment, was confirmed by the modelling. Adiabatic maps of dihedral angles vs. potential energy of disaccharide fragments of the tetrasaccharides were calculated in the absence of the sodium cations. These results confirmed the reliability of the parameterized AMBER force field as the compact structures were not observed. Further studies using different ionization conditions are planned, in order to generate less sodiated species.

Two new *J*-modulated COSY experiments, HOHAHA-COSY and CSSF-COSY, were developed for the measurement of scalar and dipolar ^{1}H-^{1}H coupling constants. These methods allow accurate and precise measurement of coupling constants from spectra with severely overlapping resonances. Good performance of these methods was obtained using model compounds. The developed methods were used to measure dipolar coupling constants of heparin-derived fully sulfated tetrasaccharide.

A series of INADEQUATE experiments (DJM-INEPT-INADEUATE, DJM-REVINEPT-INADEQUATE, JM-REVINEPT-INADEQUATE, JM-DEPT-INADEQUATE, REVINEPT-INADEQUATE, ^{13}C-detected INADEQUATE and DNP-INADEQUATE) was developed for the measurement of long-range ^{13}C-^{13}C coupling constants in samples with natural abundance of ^{13}C. These experiments were developed primarily for the measurement of ^{13}C-^{13}C coupling constants across the glycosidic linkage of carbohydrates. Current sensitivity of NMR spectrometers requires ca. 20 mg of a disaccharide for reliable determination of small $^nJ_{CC}$ coupling constants. For samples where the quantity of material is not a limiting factor, ^{13}C-detected INADEQUATE is particularly suited for the measurement of ^{13}C-^{13}C residual dipolar coupling constants. DNP-INADEQUATE was developed for one scan acquisition of ^{13}C-^{13}C DQ filtered spectra.

Solution conformation of heparin-derived fully sulfated tetrasaccharide was studied using NMR spectroscopy. The conformations of the two flexible monosaccharide rings were investigated by ^{1}H-^{1}H scalar coupling constants and the population of each conformer contributing to the dynamic equilibrium was calculated. RDC-restrained molecular dynamics was carried out to study the conformations of the glycosidic linkages of the tetrasaccharide in solution. Using an assumption of a single alignment tensor, the analysis yielded a subset of structures seen in unrestrained MD. Insufficient material prevented us from measuring the $^3J_{CC}$ coupling constants across the glycosidic linkages. These experiments would be very valuable as they would validate or not the assumption we made during the interpretation of RDCs.

The study of the interaction between the heparin-derived fully sulfated tetrasaccharide and factor H modules, fH~19, 20 and fH~7, using docking identified possible binding sites on the protein surface. Some residues identified by the modelling were also implicated in the binding based on the chemical shift perturbation studies (Dr. A. Herbert). However, the binding conformations of the tetrasaccharide could not be determined from the docking. Various conformations and orientations of the oligosaccharide were observed. The spin-labelled heparin-derived fully sulfated disaccharide showed significant broadening of spectral lines caused by the spin label (4-amino-TEMPO). The conformation of the 4-amino-TEMPO-disaccharide generated with free molecular modelling was compatible with the measured T_1 relaxation times. Further studies of protein-GAG interactions using larger spin-labelled oligosaccharide are planned.

Appendix I

Docking Parameters

Grid maps calculation:

```
receptor fh7H40.pdbqs             # macromolecule
gridfld fh7H40.maps.fld           # grid_data_file
npts 100 100 100                # num.grid points in xyz
spacing 0.6                   # spacing(A)
gridcenter 0.299 0.1 0.226        # xyz-coordinates or auto
types CNOSH                     # atom type names
smooth 0.5                     # store minimum energy w/in rad(A)

map fh7H40.C.map                  # atom-specific affinity map
nbp_r_eps  4.00 0.0222750 12  6  # C-C lj
nbp_r_eps  3.75 0.0230026 12  6  # C-N lj
nbp_r_eps  3.60 0.0257202 12  6  # C-O lj
nbp_r_eps  4.00 0.0257202 12  6  # C-S lj
nbp_r_eps  3.00 0.0081378 12  6  # C-H lj
nbp_r_eps  3.00 0.0081378 12  6  # C-H lj
nbp_r_eps  3.00 0.0081378 12  6  # C-H lj
sol_par  12.77 0.6844   # C atomic fragmental volume, solvation parameters
constant  0.000     # C grid map constant energy

map fh7H40.N.map                  # atom-specific affinity map
nbp_r_eps  3.75 0.0230026 12  6  # N-C lj
nbp_r_eps  3.50 0.0237600 12  6  # N-N lj
nbp_r_eps  3.35 0.0265667 12  6  # N-O lj
nbp_r_eps  3.75 0.0265667 12  6  # N-S lj
nbp_r_eps  1.90 0.3280000 12 10  # N-H hb
nbp_r_eps  1.90 0.3280000 12 10  # N-H hb
nbp_r_eps  1.90 0.3280000 12 10  # N-H hb
sol_par   0.00 0.0000   # N atomic fragmental volume, solvation parameters
constant  0.000     # N grid map constant energy

map fh7H40.O.map                  # atom-specific affinity map
nbp_r_eps  3.60 0.0257202 12  6  # O-C lj
nbp_r_eps  3.35 0.0265667 12  6  # O-N lj
nbp_r_eps  3.20 0.0297000 12  6  # O-O lj
nbp_r_eps  3.60 0.0297000 12  6  # O-S lj
nbp_r_eps  1.90 0.3280000 12 10  # O-H hb
```

```
nbp_r_eps  1.90 0.3280000 12 10   # O-H hb
nbp_r_eps  1.90 0.3280000 12 10   # O-H hb
sol_par   0.00 0.0000    # O atomic fragmental volume, solvation parameters
constant  0.236      # O grid map constant energy

map fh7H40.S.map                # atom-specific affinity map
nbp_r_eps  4.00 0.0257202 12  6   # S-C lj
nbp_r_eps  3.75 0.0265667 12  6   # S-N lj
nbp_r_eps  3.60 0.0297000 12  6   # S-O lj
nbp_r_eps  4.00 0.0297000 12  6   # S-S lj
nbp_r_eps  2.50 0.0656000 12 10   # S-H hb
nbp_r_eps  2.50 0.0656000 12 10   # S-H hb
nbp_r_eps  2.50 0.0656000 12 10   # S-H hb
sol_par  0.000 0.000    #S atomic fragmental volume, solvation parameters
constant  0.000        #S grid map constant energy

map fh7H40.H.map                # atom-specific affinity map
nbp_r_eps  3.00 0.0081378 12  6   # H-C lj
nbp_r_eps  1.90 0.3280000 12 10   # H-N hb
nbp_r_eps  1.90 0.3280000 12 10   # H-O hb
nbp_r_eps  2.50 0.0656000 12 10   # H-S hb
nbp_r_eps  2.00 0.0029700 12  6   # H-H lj
nbp_r_eps  2.00 0.0029700 12  6   # H-H lj
nbp_r_eps  2.00 0.0029700 12  6   # H-H lj
sol_par   0.00 0.0000    # H atomic fragmental volume, solvation parameters
constant  0.118     # H grid map constant energy

elecmap fh7H40.e.map             # electrostatic potential map
dielectric 1.0000             # <0, distance-dep.diel;>0, constant
#dielectric -0.1146             # <0, distance-dep.diel;>0, constant
#
```

Docking:

```
seed pid time                # seeds for random generator
types CNOSH                    # atom type names
fld fh7H40.maps.fld             # grid_data_file
map fh7H40.C.map                 # atom-specific affinity map
map fh7H40.N.map                 # atom-specific affinity map
map fh7H40.O.map                 # atom-specific affinity map
map fh7H40.S.map                 # atom-specific affinity map
map fh7H40.H.map                 # atom-specific affinity map
map fh7H40.e.map                # electrostatics map
move 161.pdbq                  # small molecule
about 19.7909 14.3745 17.4982       # small molecule center
tran0 random                  # initial coordinates/A or random
quat0 random                   # initial quaternion
ndihe 27                     # number of active torsions
dihe0 random                   # initial dihedrals (relative) or random
tstep 2.0                   # translation step/A
qstep 50.0                   # quaternion step/deg
dstep 50.0                   # torsion step/deg
```

```
torsdof 22 0.3113              # torsional degrees of freedom and coefficients

intnbp_r_eps  4.00 0.0222750 12 6   # C-C lj
intnbp_r_eps  3.75 0.0230026 12 6   # C-N lj
intnbp_r_eps  3.60 0.0257202 12 6   # C-O lj
intnbp_r_eps  4.00 0.0257202 12 6   # C-S lj
intnbp_r_eps  3.00 0.0081378 12 6   # C-H lj
intnbp_r_eps  3.50 0.0237600 12 6   # N-N lj
intnbp_r_eps  3.35 0.0265667 12 6   # N-O lj
intnbp_r_eps  3.75 0.0265667 12 6   # N-S lj
intnbp_r_eps  2.75 0.0084051 12 6   # N-H lj
intnbp_r_eps  3.20 0.0297000 12 6   # O-O lj
intnbp_r_eps  3.60 0.0297000 12 6   # O-S lj
intnbp_r_eps  2.60 0.0093852 12 6   # O-H lj
intnbp_r_eps  4.00 0.0297000 12 6   # S-S lj
intnbp_r_eps  3.00 0.0093852 12 6   # S-H lj
intnbp_r_eps  2.00 0.0029700 12 6   # H-H lj
#
outlev 1                    # diagnostic output level
rmstol 0.5                   # cluster_tolerance/A
extnrg 1000.0                  # external grid energy
e0max 0.0 10000                  # max initial energy; max number of retriesga_pop_size 50

# number of individuals in population
ga_num_evals 250000              # maximum number of energy evaluations
ga_num_generations 27000           # maximum number of generations
ga_elitism 1                  # number of top individuals to survive to next generation
ga_mutation_rate 0.02             # rate of gene mutation
ga_crossover_rate 0.8             # rate of crossover
ga_window_size 10                 #
ga_cauchy_alpha 0.0               # Alpha parameter of Cauchy distribution
ga_cauchy_beta 1.0               # Beta parameter Cauchy distribution
set_ga                      # set the above parameters for GA or LGA
do_global_only 250                # do this many GA runs
analysis                    # perform a ranked cluster analysis
```

Appendix II

Parameters of Sulfate/Sulfamate groups added to AMBER7

MASS

SO	32.06	0.00	Sulfate/Sulfamate group

BOND

CT-OS	320	1.41	
CT-NT	367	1.471	
SO-OS	230	1.62	
SO-O2	525	1.44	
SO-NT	230	1.685	
OH-SO	230.0	1.610	JCC, 7, (1986), 230; NA Phosphate

ANGL

SO-NT-CT	100	110.4		
SO-NT-H2	45	109.5		
SO-OS-CT	100	120.5		
H2-NT-CT	35	109.5		
NT-CT-CT	80	109.7		
NT-SO-O2	100	109.6		
O2-SO-O2	140	119.9		
HC-CT-NT	35	109.5		
OS-SO-O2	100	108.23		
OS-SO-OS	45	102.6		
O2-SO-OH	45.0	108.23	Parm99 Phosphate	
NT-SO-OH	100	109.6	Appr. NT-SO-O2	
HO-OH-SO	45.0	108.50	Parm99 Phosphate	
SO-NT-H	45	109.5	Appr. SO-NT-H2	
NT-CT-AC	80	109.7	Appr. NT-CT-CT	
NT-CT-EC	80	109.7	Appr. NT-CT-CT	
OH-SO-OS	45.0	102.60	Parm99 Phosphate	
HO-OH-SO	45.0	108.50	Parm99 Phosphate	
CM-OS-AC	60.0	117.00	Junmei et al, 1999	Appr. CM-OS-CT
CM-OS-EC	60.0	117.00	Junmei et al, 1999	Appr. CM-OS-CT
C -CM-OS	80.0	125.00	Junmei et al, 1999	Appr. CM-CM-OS
CM-CT-OH	50.0	109.50	Junmei et al, 1999	Appr. CM-CT-OS
CM-C -OH	80.0	125.30	Appr. CM-C -O	
CM-C -O2	80.0	125.30	Appr. CM-C -O	

DIHE

```
X -CT-NT-X     3     0.00     0.00     6
HC-CT-NT-SO    3     1.05     0.00     3
CT-CT-NT-SO    3     1.05     0.00     3
CT-CT-NT-SO    1     -6.4     0.00     3
CT-CT-OS-SO    1     -1.16    0.00     1
CT-CT-OS-SO    2     0.25     0.00     1
CT-CT-OS-SO    3     -0.15    0.00     1
CT-CT-OS-SO    4     -0.30    0.00     1
HC-CT-OS-SO    3     0.15     0.00     1
X -CT-OS-X     3     1.15     0.00     3
X -OS-SO-X     3     0.75     0.00     3
X -NT-SO-X     3     0.75     0.00     6
CT-NT-SO-O2    3     1.60     0.00     6
X -OH-SO-X     3     0.75     0.00     3.  JCC, 7, (1986), 230  Phosphate
```

IMPR

HBON

NONB

Appendix III

Pdb files of heparin-derived oligosaccharides with partial charges

Disaccharide I

```
ATOM    1 O1  DAP  1    2.287  1.412  0.719 -0.6260 1.6612
ATOM    2 S1  DAP  1    1.973  1.307 -0.711  1.0710 2.0000
ATOM    3 O2  DAP  1    2.771  1.913 -1.725 -0.6260 1.6612
ATOM    4 O3  DAP  1    0.500  1.864 -0.924 -0.6260 1.7210
ATOM    5 O4  DAP  1    1.889 -0.285 -1.141 -0.3950 1.6837
ATOM    6 C1  DAP  1    1.569 -0.783 -2.414  0.0290 1.9080
ATOM    7 H1  DAP  1    1.801 -0.075 -3.210  0.0840 1.3870
ATOM    8 H2  DAP  1    0.492 -0.942 -2.454  0.0840 1.3870
ATOM    9 C2  DAP  1    2.420 -2.080 -2.788  0.1200 1.9080
ATOM   10 H3  DAP  1    2.199 -2.383 -3.811  0.0690 1.3870
ATOM   11 O5  DAP  1    3.822 -1.793 -2.655 -0.2960 1.6837
ATOM   12 C3  DAP  1    4.716 -2.915 -2.726  0.0250 1.9080
ATOM   13 H4  DAP  1    5.738 -2.569 -2.569  0.1750 1.2870
ATOM   14 O6  DAP  1    4.613 -3.470 -4.011 -0.6040 1.7210
ATOM   15 H5  DAP  1    5.005 -4.345 -3.963  0.4350 0.0000
ATOM   16 C4  DAP  1    4.411 -3.948 -1.631  0.0950 1.9080
ATOM   17 H6  DAP  1    4.432 -3.305 -0.751  0.1140 1.3870
ATOM   18 N1  DAP  1    5.411 -5.084 -1.319 -0.6640 1.8240
ATOM   19 H7  DAP  1    4.809 -5.853 -1.062  0.3570 0.6000
ATOM   20 S2  DAP  1    6.353 -5.705 -2.473  1.0460 2.0000
ATOM   21 O7  DAP  1    6.689 -7.287 -2.182 -0.6330 1.7210
ATOM   22 O8  DAP  1    5.586 -5.729 -3.668 -0.6330 1.6612
ATOM   23 O9  DAP  1    7.662 -5.124 -2.532 -0.6330 1.6612
ATOM   24 C5  DAP  1    2.944 -4.436 -1.796  0.2170 1.9080
ATOM   25 H8  DAP  1    2.962 -4.996 -2.731  0.1100 1.3870
ATOM   26 O10 DAP  1    2.678 -5.261 -0.665 -0.6760 1.7210
ATOM   27 H9  DAP  1    1.728 -5.397 -0.634  0.4400 0.0000
ATOM   28 C6  DAP  1    1.943 -3.301 -1.900 -0.0020 1.9080
ATOM   29 H10 DAP  1    1.770 -2.945 -0.884  0.1360 1.3870
ATOM   30 O11 DAP  1    0.648 -3.859 -2.386 -0.3560 1.8310
ATOM   31 C7  DAP  1   -0.456 -3.971 -1.511  0.1450 1.9080
ATOM   32 H11 DAP  1   -0.527 -3.091 -0.872  0.1480 1.2870
ATOM   33 O12 DAP  1   -0.352 -5.089 -0.602 -0.2880 1.6837
ATOM   34 C8  DAP  1   -0.771 -6.259 -0.790  0.0900 1.9080
ATOM   35 C9  DAP  1   -0.618 -7.326  0.184  0.8380 1.9080
```

```
ATOM   36  O13 DAP   1    0.122  -7.193   1.195 -0.8020  1.6612
ATOM   37  O14 DAP   1   -1.283  -8.520   0.080 -0.8020  1.7210
ATOM   38  C10 DAP   1   -1.536  -6.543  -1.908 -0.4520  1.9080
ATOM   39  H12 DAP   1   -1.857  -7.568  -2.020  0.1490  1.4590
ATOM   40  C11 DAP   1   -1.922  -5.502  -2.929  0.3520  1.9080
ATOM   41  H13 DAP   1   -2.971  -5.609  -3.206  0.0910  1.3870
ATOM   42  O15 DAP   1   -1.123  -5.815  -3.998 -0.7110  1.7210
ATOM   43  H14 DAP   1   -0.246  -5.872  -3.611  0.3940  0.0000
ATOM   44  C12 DAP   1   -1.766  -4.029  -2.378  0.1320  1.9080
ATOM   45  H15 DAP   1   -1.563  -3.366  -3.220  0.0810  1.3870
ATOM   46  O16 DAP   1   -2.910  -3.488  -1.709 -0.4220  1.6837
ATOM   47  S3  DAP   1  -4.275  -3.008  -2.450  1.1410  2.0000
ATOM   48  O18 DAP   1   -5.080  -4.337  -2.743 -0.6410  1.7210
ATOM   49  O19 DAP   1   -3.920  -2.471  -3.752 -0.6410  1.6612
ATOM   50  O17 DAP   1   -5.094  -2.191  -1.614 -0.6410  1.6612
```

Tetrasaccharide II

```
ATOM    1  O6  TAP   1   13.482  10.732  20.442 -0.6600  1.6612
ATOM    2  S2  TAP   1   13.925  11.479  19.247  1.1300  2.0000
ATOM    3  O7  TAP   1   13.100  11.604  18.083 -0.6600  1.6612
ATOM    4  O8  TAP   1   14.046  13.073  19.610 -0.6600  1.7210
ATOM    5  O5  TAP   1   15.437  11.070  18.774 -0.4140  1.6837
ATOM    6  C4  TAP   1   15.755   9.731  18.257  0.0770  1.9080
ATOM    7  H2  TAP   1   16.768   9.526  18.605  0.0940  1.3870
ATOM    8  H3  TAP   1   15.021   9.026  18.647  0.0940  1.3870
ATOM    9  C3  TAP   1   15.825   9.589  16.730  0.0370  1.9080
ATOM   10  H4  TAP   1   16.491   8.749  16.530  0.1640  1.3870
ATOM   11  O4  TAP   1   14.532   9.434  16.132 -0.3580  1.6837
ATOM   12  C2  TAP   1   14.400   9.112  14.784  0.0650  1.9080
ATOM   13  H6  TAP   1   13.341   9.119  14.526  0.2070  1.2870
ATOM   14  C1  TAP   1   15.048  10.339  14.093  0.1560  1.9080
ATOM   15  H7  TAP   1   14.476  11.240  14.315  0.1440  1.3870
ATOM   16  C5  TAP   1   16.495  10.676  14.526  0.0370  1.9080
ATOM   17  H9  TAP   1   17.127   9.824  14.279  0.1070  1.3870
ATOM   18  O10 TAP   1   16.859  11.861  13.880 -0.6020  1.7210
ATOM   19  H8  TAP   1   17.425  11.675  13.128  0.4030  0.0000
ATOM   20  N1  TAP   1   14.823  10.060  12.645 -0.7110  1.8240
ATOM   21  H1  TAP   1   15.597   9.554  12.237  0.3780  0.6000
ATOM   22  S1  TAP   1   14.635  11.362  11.721  1.0700  2.0000
ATOM   23  O1  TAP   1   14.101  12.485  12.406 -0.6600  1.6612
ATOM   24  O2  TAP   1   13.603  11.072  10.523 -0.6600  1.7210
ATOM   25  O3  TAP   1   15.869  11.668  11.103 -0.6600  1.6612
ATOM   26  O9  TAP   1   15.015   7.936  14.413 -0.5860  1.7210
ATOM   27  H5  TAP   1   15.169   7.972  13.466  0.3920  0.0000
ATOM   28  C6  TAP   1   16.561  10.793  16.073  0.0220  1.9080
ATOM   29  H10 TAP   1   16.085  11.743  16.316  0.1260  1.3870
ATOM   30  O11 TAP   1   17.940  10.720  16.479 -0.3450  1.8310
ATOM   31  C7  TAP   1   18.313  11.890  17.261  0.1480  1.9080
ATOM   32  H12 TAP   1   17.539  12.143  17.985  0.1840  1.2870
ATOM   33  O12 TAP   1   18.398  13.133  16.503 -0.3790  1.6837
ATOM   34  C8  TAP   1   19.436  13.120  15.473  0.0380  1.9080
ATOM   35  H11 TAP   1   19.256  12.285  14.796  0.0760  1.3870
```

```
ATOM   36 C9  TAP  1   19.330 14.394 14.609 0.8320 1.9080
ATOM   37 O13 TAP  1   18.694 15.443 14.811 -0.7990 1.6612
ATOM   38 O14 TAP  1   19.940 14.301 13.372 -0.7990 1.7210
ATOM   39 C10 TAP  1   19.703 11.681 17.933 0.0470 1.9080
ATOM   40 O15 TAP  1   19.825 12.675 18.953 -0.3780 1.6837
ATOM   41 S3  TAP  1   19.413 12.315 20.445 1.0860 2.0000
ATOM   42 O16 TAP  1   20.283 11.249 20.890 -0.6370 1.6612
ATOM   43 O17 TAP  1   18.001 12.109 20.501 -0.6370 1.6612
ATOM   44 O18 TAP  1   19.754 13.559 21.324 -0.6370 1.7210
ATOM   45 H13 TAP  1   19.680 10.699 18.404 0.1450 1.3870
ATOM   46 C11 TAP  1   20.886 11.643 16.919 0.0440 1.9080
ATOM   47 O19 TAP  1   20.816 10.511 15.981 -0.5940 1.7210
ATOM   48 H14 TAP  1   19.979 10.542 15.513 0.4140 0.0000
ATOM   49 H15 TAP  1   21.815 11.532 17.478 0.1750 1.3870
ATOM   50 C12 TAP  1   20.909 12.953 16.043 0.0080 1.9080
ATOM   51 H16 TAP  1   21.611 12.863 15.214 0.0830 1.3870
ATOM   52 O20 TAP  1   21.270 14.109 16.935 -0.2380 1.8310
ATOM   53 C13 TAP  1   22.603 14.516 17.016 -0.0250 1.9080
ATOM   54 H17 TAP  1   23.234 13.765 16.540 0.1480 1.2870
ATOM   55 O25 TAP  1   22.775 15.676 16.264 -0.2480 1.6837
ATOM   56 C16 TAP  1   22.297 16.915 16.888 -0.1010 1.9080
ATOM   57 H24 TAP  1   21.217 16.788 16.953 0.1400 1.3870
ATOM   58 C17 TAP  1   22.548 18.038 15.822 -0.0330 1.9080
ATOM   59 H22 TAP  1   22.119 17.670 14.891 0.1220 1.3870
ATOM   60 H23 TAP  1   21.919 18.901 16.041 0.1220 1.3870
ATOM   61 O26 TAP  1   23.986 18.480 15.750 -0.3600 1.6837
ATOM   62 S5  TAP  1   25.068 17.433 15.361 1.1710 2.0000
ATOM   63 O27 TAP  1   26.449 18.120 15.029 -0.6640 1.7210
ATOM   64 O28 TAP  1   25.452 16.714 16.524 -0.6640 1.6612
ATOM   65 O29 TAP  1   24.807 16.736 14.138 -0.6640 1.6612
ATOM   66 C14 TAP  1   23.082 14.677 18.530 0.1110 1.9080
ATOM   67 H18 TAP  1   24.166 14.685 18.646 0.2090 1.3870
ATOM   68 N2  TAP  1   22.728 13.412 19.244 -0.5960 1.8240
ATOM   69 H21 TAP  1   22.743 13.592 20.237 0.3580 0.6000
ATOM   70 S4  TAP  1   23.949 12.245 19.151 1.0630 2.0000
ATOM   71 O22 TAP  1   24.701 12.420 20.318 -0.6610 1.6612
ATOM   72 O23 TAP  1   23.480 10.760 18.986 -0.6610 1.7210
ATOM   73 O24 TAP  1   24.656 12.431 17.910 -0.6610 1.6612
ATOM   74 C15 TAP  1   22.540 15.917 19.169 0.1530 1.9080
ATOM   75 H19 TAP  1   21.461 15.811 19.280 0.0970 1.3870
ATOM   76 O21 TAP  1   23.146 16.148 20.401 -0.6280 1.7210
ATOM   77 H20 TAP  1   22.405 16.285 20.996 0.4280 0.0000
ATOM   78 C18 TAP  1   22.815 17.158 18.288 -0.0850 1.9080
ATOM   79 H25 TAP  1   23.886 17.363 18.282 0.1480 1.3870
ATOM   80 O30 TAP  1   21.942 18.276 18.606 -0.2210 1.8310
ATOM   81 C19 TAP  1   22.441 19.212 19.548 0.0670 1.9080
ATOM   82 H26 TAP  1   23.525 19.217 19.437 0.1780 1.2870
ATOM   83 O31 TAP  1   22.063 18.858 20.904 -0.2730 1.6837
ATOM   84 C20 TAP  1   20.886 18.776 21.210 0.1050 1.9080
ATOM   85 C21 TAP  1   20.456 17.626 22.044 0.8600 1.9080
ATOM   86 O32 TAP  1   19.161 17.471 22.371 -0.8310 1.7210
ATOM   87 O33 TAP  1   21.313 16.816 22.482 -0.8310 1.6612
ATOM   88 C22 TAP  1   20.029 19.765 20.747 -0.4130 1.9080
```

```
ATOM   89 H27 TAP  1   18.973 19.781 20.974  0.1290 1.4590
ATOM   90 C23 TAP  1   20.555 20.895 19.880  0.2880 1.9080
ATOM   91 H28 TAP  1   20.612 21.784 20.509  0.1150 1.3870
ATOM   92 O34 TAP  1   19.556 21.196 18.905 -0.6980 1.7210
ATOM   93 H29 TAP  1   19.220 20.370 18.550  0.4080 0.0000
ATOM   94 C24 TAP  1   21.894 20.591 19.255  0.0730 1.9080
ATOM   95 H30 TAP  1   21.863 20.628 18.166  0.1180 1.3870
ATOM   96 O35 TAP  1   22.775 21.636 19.636 -0.3680 1.6837
ATOM   97 S6  TAP  1   23.506 21.888 21.039  1.1010 2.0000
ATOM   98 O37 TAP  1   24.697 21.116 20.950 -0.6460 1.6612
ATOM   99 O38 TAP  1   22.582 21.657 22.133 -0.6460 1.6612
ATOM  100 O36 TAP  1   24.002 23.391 20.860 -0.6460 1.7210
```

Tetrasaccharide III

```
ATOM    1 O1  TAD  1    3.999  0.901  1.091 -0.6120 1.7210
ATOM    2 H31 TAD  1    3.399  0.542  1.748  0.3590 0.0000
ATOM    3 C1  TAD  1    3.777  2.308  0.922  0.0740 1.9080
ATOM    4 H1  TAD  1    3.961  2.584 -0.116  0.0710 1.3870
ATOM    5 H2  TAD  1    2.752  2.605  1.145  0.0710 1.3870
ATOM    6 C2  TAD  1    4.617  3.162  1.909  0.0420 1.9080
ATOM    7 H3  TAD  1    4.569  2.610  2.848  0.1770 1.3870
ATOM    8 O2  TAD  1    5.945  3.261  1.368 -0.3680 1.6837
ATOM    9 C3  TAD  1    6.916  4.052  2.060  0.0330 1.9080
ATOM   10 H4  TAD  1    7.826  4.101  1.463  0.2070 1.2870
ATOM   11 O7  TAD  1    7.207  3.527  3.343 -0.5550 1.7210
ATOM   12 H9  TAD  1    6.434  3.628  3.902  0.3730 0.0000
ATOM   13 C4  TAD  1    6.381  5.508  2.196  0.0760 1.9080
ATOM   14 H5  TAD  1    6.161  5.871  1.192  0.1690 1.3870
ATOM   15 N1  TAD  1    7.399  6.322  2.876 -0.6890 1.8240
ATOM   16 H8  TAD  1    6.934  7.187  3.114  0.3670 0.6000
ATOM   17 S1  TAD  1    8.623  6.797  1.892  1.1770 2.0000
ATOM   18 O4  TAD  1    8.943  8.216  2.238 -0.6820 1.7210
ATOM   19 O5  TAD  1    8.228  6.934  0.542 -0.6820 1.6612
ATOM   20 O6  TAD  1    9.830  6.094  2.203 -0.6820 1.6612
ATOM   21 C5  TAD  1    5.025  5.428  2.949  0.0460 1.9080
ATOM   22 H6  TAD  1    5.149  5.077  3.973  0.1430 1.3870
ATOM   23 O3  TAD  1    4.516  6.755  3.024 -0.5900 1.7210
ATOM   24 H7  TAD  1    3.570  6.652  3.154  0.4020 0.0000
ATOM   25 C6  TAD  1    3.961  4.489  2.191 -0.0350 1.9080
ATOM   26 H10 TAD  1    3.727  4.996  1.255  0.1490 1.3870
ATOM   27 O8  TAD  1    2.777  4.404  3.063 -0.3810 1.8310
ATOM   28 C7  TAD  1    1.442  4.476  2.521  0.0990 1.9080
ATOM   29 H11 TAD  1    1.489  4.314  1.444  0.1860 1.2870
ATOM   30 C10 TAD  1    0.572  3.390  3.213  0.0890 1.9080
ATOM   31 O12 TAD  1    1.035  2.168  2.672 -0.3610 1.6837
ATOM   32 S2  TAD  1    1.014  0.753  3.496  1.0730 2.0000
ATOM   33 O13 TAD  1   -0.281  0.470  3.984 -0.6170 1.6612
ATOM   34 O14 TAD  1    1.940  0.960  4.758 -0.6170 1.7210
ATOM   35 O15 TAD  1    1.745 -0.225  2.757 -0.6170 1.6612
ATOM   36 H13 TAD  1    0.803  3.354  4.278  0.1500 1.3870
ATOM   37 C11 TAD  1   -0.972  3.570  3.054  0.0290 1.9080
ATOM   38 O16 TAD  1   -1.606  2.880  4.151 -0.6730 1.7210
```

```
ATOM   39  H14 TAD  1    -1.257   1.986   4.125  0.4690  0.0000
ATOM   40  H15 TAD  1    -1.240   3.302   2.032  0.1360  1.3870
ATOM   41  O9  TAD  1     0.892   5.754   2.706 -0.2970  1.6837
ATOM   42  C8  TAD  1    -0.482   5.969   2.340 -0.0320  1.9080
ATOM   43  H12 TAD  1    -0.732   7.000   2.592  0.0580  1.3870
ATOM   44  C9  TAD  1    -0.672   5.814   0.795  0.8430  1.9080
ATOM   45  O10 TAD  1    -0.848   7.003   0.144 -0.8100  1.7210
ATOM   46  O11 TAD  1    -0.697   4.727   0.221 -0.8100  1.6612
ATOM   47  C12 TAD  1    -1.418   5.069   3.183  0.0550  1.9080
ATOM   48  H16 TAD  1    -1.245   5.446   4.191  0.1050  1.3870
ATOM   49  O17 TAD  1    -2.869   5.292   2.878 -0.2230  1.8310
ATOM   50  C13 TAD  1    -3.643   6.271   3.661 -0.0460  1.9080
ATOM   51  H17 TAD  1    -2.964   7.022   4.065  0.1760  1.2870
ATOM   52  C16 TAD  1    -4.381   5.686   4.911  0.0480  1.9080
ATOM   53  H21 TAD  1    -4.888   6.531   5.379  0.1890  1.3870
ATOM   54  N2  TAD  1    -3.347   5.092   5.812 -0.5910  1.8240
ATOM   55  H22 TAD  1    -3.299   4.089   5.703  0.3600  0.6000
ATOM   56  S4  TAD  1    -3.645   5.291   7.455  1.0440  2.0000
ATOM   57  O23 TAD  1    -5.057   5.454   7.681 -0.6520  1.6612
ATOM   58  O24 TAD  1    -2.978   4.233   8.150 -0.6520  1.6612
ATOM   59  O25 TAD  1    -3.000   6.639   8.000 -0.6520  1.7210
ATOM   60  C17 TAD  1    -5.447   4.740   4.408  0.1390  1.9080
ATOM   61  H23 TAD  1    -4.946   3.959   3.836  0.1600  1.3870
ATOM   62  O26 TAD  1    -6.058   4.136   5.593 -0.6130  1.7210
ATOM   63  H24 TAD  1    -5.814   4.678   6.347  0.4280  0.0000
ATOM   64  O18 TAD  1    -4.669   6.969   2.942 -0.2540  1.6837
ATOM   65  C14 TAD  1    -5.795   6.213   2.390 -0.0230  1.9080
ATOM   66  H18 TAD  1    -5.383   5.510   1.666  0.0980  1.3870
ATOM   67  C15 TAD  1    -6.596   7.279   1.573 -0.0080  1.9080
ATOM   68  H19 TAD  1    -6.038   7.748   0.762  0.1080  1.3870
ATOM   69  H20 TAD  1    -7.492   6.828   1.146  0.1080  1.3870
ATOM   70  O19 TAD  1    -7.171   8.390   2.327 -0.3690  1.6837
ATOM   71  S3  TAD  1    -8.015   9.570   1.558  1.1390  2.0000
ATOM   72  O20 TAD  1    -8.488  10.722   2.548 -0.6540  1.7210
ATOM   73  O21 TAD  1    -7.178  10.300   0.606 -0.6540  1.6612
ATOM   74  O22 TAD  1    -9.264   9.019   1.099 -0.6540  1.6612
ATOM   75  C18 TAD  1    -6.534   5.394   3.498 -0.0610  1.9080
ATOM   76  H25 TAD  1    -7.134   6.070   4.108  0.1360  1.3870
ATOM   77  O27 TAD  1    -7.323   4.336   2.852 -0.3060  1.8310
ATOM   78  C19 TAD  1    -8.695   4.577   2.541  0.2460  1.9080
ATOM   79  H26 TAD  1    -8.822   5.544   2.054  0.1240  1.2870
ATOM   80  O28 TAD  1    -9.574   4.586   3.663 -0.2940  1.6837
ATOM   81  C20 TAD  1   -10.174   3.547   4.121  0.1070  1.9080
ATOM   82  C21 TAD  1   -11.052   3.548   5.321  0.8720  1.9080
ATOM   83  O29 TAD  1   -11.555   2.484   5.705 -0.8360  1.6612
ATOM   84  O30 TAD  1   -11.394   4.634   6.001 -0.8360  1.7210
ATOM   85  C22 TAD  1   -10.092   2.352   3.421 -0.4490  1.9080
ATOM   86  H27 TAD  1   -10.587   1.477   3.815  0.1280  1.4590
ATOM   87  C23 TAD  1    -9.401   2.176   2.136  0.3710  1.9080
ATOM   88  H28 TAD  1    -9.976   1.519   1.483  0.0850  1.3870
ATOM   89  O31 TAD  1    -8.203   1.559   2.517 -0.7440  1.7210
ATOM   90  H29 TAD  1    -7.877   2.147   3.203  0.4340  0.0000
ATOM   91  C24 TAD  1    -9.177   3.530   1.434  0.1010  1.9080
```

```
ATOM   92 H30 TAD  1   -8.405  3.474  0.667  0.0830 1.3870
ATOM   93 O32 TAD  1  -10.393  4.015  0.799 -0.3990 1.6837
ATOM   94 S5  TAD  1  -11.291  3.199 -0.319  1.0860 2.0000
ATOM   95 O34 TAD  1  -12.074  4.170 -1.067 -0.6410 1.6612
ATOM   96 O35 TAD  1  -11.952  2.014  0.171 -0.6410 1.6612
ATOM   97 O33 TAD  1  -10.273  2.664 -1.451 -0.6410 1.7210
```

Tetrasaccharide IV

```
ATOM    1 O1  HSA  1   -0.117 11.975 -2.879 -0.6520 1.6612
ATOM    2 S1  HSA  1    0.347 13.095 -2.102  1.1210 2.0000
ATOM    3 O2  HSA  1    1.472 13.874 -2.559 -0.6520 1.6612
ATOM    4 O3  HSA  1   -0.867 14.153 -2.023 -0.6520 1.7210
ATOM    5 O4  HSA  1    0.589 12.664 -0.548 -0.4060 1.6837
ATOM    6 C1  HSA  1    0.001 11.511  0.065  0.0810 1.9080
ATOM    7 H1  HSA  1    0.051 11.640  1.146  0.0800 1.3870
ATOM    8 H2  HSA  1   -1.051 11.444 -0.217  0.0800 1.3870
ATOM    9 C2  HSA  1    0.713 10.189 -0.291  0.0160 1.9080
ATOM   10 H3  HSA  1    0.491  9.966 -1.336  0.1430 1.3870
ATOM   11 O5  HSA  1    2.123 10.371 -0.143 -0.2960 1.6837
ATOM   12 C3  HSA  1    2.973  9.313 -0.580  0.0480 1.9080
ATOM   13 H4  HSA  1    3.999  9.573 -0.305  0.1520 1.2870
ATOM   14 O10 HSA  1    2.935  9.141 -1.969 -0.5190 1.7210
ATOM   15 H9  HSA  1    3.516  8.387 -2.133  0.3570 0.0000
ATOM   16 C4  HSA  1    2.621  7.993  0.130  0.0520 1.9080
ATOM   17 H5  HSA  1    2.919  8.094  1.176  0.1600 1.3870
ATOM   18 N1  HSA  1    3.319  6.824 -0.490 -0.6320 1.8240
ATOM   19 H8  HSA  1    3.088  6.008  0.047  0.3590 0.6000
ATOM   20 S2  HSA  1    4.966  6.839 -0.385  1.1020 2.0000
ATOM   21 O7  HSA  1    5.469  7.545 -1.530 -0.6610 1.6612
ATOM   22 O8  HSA  1    5.368  5.475 -0.188 -0.6610 1.6612
ATOM   23 O9  HSA  1    5.432  7.635  0.921 -0.6610 1.7210
ATOM   24 C5  HSA  1    1.095  7.756  0.108  0.0200 1.9080
ATOM   25 H6  HSA  1    0.808  7.578 -0.931  0.1590 1.3870
ATOM   26 O6  HSA  1    0.835  6.584  0.872 -0.5650 1.7210
ATOM   27 H7  HSA  1   -0.119  6.535  1.064  0.3670 0.0000
ATOM   28 C6  HSA  1    0.282  8.986  0.603  0.0130 1.9080
ATOM   29 H10 HSA  1    0.547  9.208  1.638  0.1260 1.3870
ATOM   30 O11 HSA  1   -1.152  8.711  0.495 -0.3270 1.8310
ATOM   31 C7  HSA  1   -1.993  8.556  1.655  0.0660 1.9080
ATOM   32 H11 HSA  1   -2.996  8.336  1.286  0.1550 1.2870
ATOM   33 C10 HSA  1   -1.580  7.394  2.585  0.1670 1.9080
ATOM   34 O15 HSA  1   -1.800  6.157  1.916 -0.6710 1.7210
ATOM   35 H31 HSA  1   -2.719  5.917  2.078  0.4160 0.0000
ATOM   36 H13 HSA  1   -0.514  7.499  2.786  0.0970 1.3870
ATOM   37 C11 HSA  1   -2.282  7.376  3.968  0.0520 1.9080
ATOM   38 O16 HSA  1   -1.315  7.600  4.993 -0.6620 1.7210
ATOM   39 H14 HSA  1   -1.271  8.574  5.168  0.4280 0.0000
ATOM   40 H15 HSA  1   -2.679  6.375  4.142  0.0390 1.3870
ATOM   41 O12 HSA  1   -2.106  9.746  2.424 -0.2600 1.6837
ATOM   42 C8  HSA  1   -3.056  9.765  3.516 -0.0490 1.9080
ATOM   43 H12 HSA  1   -3.965 10.255  3.166  0.0280 1.3870
ATOM   44 C9  HSA  1   -2.464 10.620  4.652  0.8200 1.9080
```

```
ATOM   45  O13 HSA  1    -3.091  11.822   4.837 -0.8010  1.7210
ATOM   46  O14 HSA  1    -1.484  10.188   5.308 -0.8010  1.6612
ATOM   47  C12 HSA  1    -3.460   8.373   4.089  0.2300  1.9080
ATOM   48  H16 HSA  1    -3.676   8.487   5.155  0.1110  1.3870
ATOM   49  O17 HSA  1    -4.663   7.792   3.467 -0.3610  1.8310
ATOM   50  C13 HSA  1    -5.901   7.900   4.191 -0.0040  1.9080
ATOM   51  H17 HSA  1    -5.774   8.654   4.973  0.1380  1.2870
ATOM   52  C16 HSA  1    -6.422   6.616   4.884  0.1150  1.9080
ATOM   53  H21 HSA  1    -7.329   6.903   5.424  0.1550  1.3870
ATOM   54  N2  HSA  1    -5.464   6.058   5.875 -0.3850  1.8240
ATOM   55  H22 HSA  1    -5.731   5.097   6.009  0.2700  0.6000
ATOM   56  S4  HSA  1    -5.721   6.649   7.417  1.0380  2.0000
ATOM   57  O23 HSA  1    -6.056   8.045   7.360 -0.6430  1.6612
ATOM   58  O24 HSA  1    -6.628   5.755   8.081 -0.6430  1.6612
ATOM   59  O25 HSA  1    -4.372   6.521   8.278 -0.6430  1.7210
ATOM   60  C17 HSA  1    -6.886   5.539   3.875  0.0880  1.9080
ATOM   61  H23 HSA  1    -6.032   5.234   3.267  0.1300  1.3870
ATOM   62  O26 HSA  1    -7.373   4.444   4.661 -0.7020  1.7210
ATOM   63  H24 HSA  1    -7.196   3.563   4.243  0.5070  0.0000
ATOM   64  O18 HSA  1    -6.972   8.375   3.371 -0.2720  1.6837
ATOM   65  C14 HSA  1    -7.554   7.505   2.395  0.0700  1.9080
ATOM   66  H18 HSA  1    -6.791   7.317   1.639  0.1070  1.3870
ATOM   67  C15 HSA  1    -8.711   8.282   1.723 -0.0690  1.9080
ATOM   68  H19 HSA  1    -8.319   9.238   1.373  0.1120  1.3870
ATOM   69  H20 HSA  1    -9.066   7.722   0.857  0.1120  1.3870
ATOM   70  O19 HSA  1    -9.822   8.534   2.584 -0.3590  1.6837
ATOM   71  S3  HSA  1   -11.027   9.562   2.166  1.1200  2.0000
ATOM   72  O20 HSA  1   -11.606   8.990   0.776 -0.6450  1.7210
ATOM   73  O21 HSA  1   -12.093   9.468   3.133 -0.6450  1.6612
ATOM   74  O22 HSA  1   -10.471  10.853   1.853 -0.6450  1.6612
ATOM   75  C18 HSA  1    -8.000   6.126   2.956  0.0370  1.9080
ATOM   76  H25 HSA  1    -8.885   6.292   3.574  0.1150  1.3870
ATOM   77  O27 HSA  1    -8.326   5.249   1.819 -0.5180  1.8310
ATOM   78  C19 HSA  1    -9.362   4.271   1.993  0.5260  1.9080
ATOM   79  H26 HSA  1   -10.114   4.697   2.658  0.0760  1.2870
ATOM   80  O28 HSA  1    -8.842   3.106   2.619 -0.4010  1.6837
ATOM   81  C20 HSA  1    -8.080   2.307   2.018  0.0560  1.9080
ATOM   82  C21 HSA  1    -7.074   1.530   2.771  0.8640  1.9080
ATOM   83  O29 HSA  1    -6.399   0.457   2.232 -0.8200  1.7210
ATOM   84  O30 HSA  1    -6.801   1.870   3.950 -0.8200  1.6612
ATOM   85  C22 HSA  1    -8.157   2.172   0.662 -0.3970  1.9080
ATOM   86  H27 HSA  1    -7.495   1.485   0.157  0.1270  1.4590
ATOM   87  C23 HSA  1    -9.085   3.033  -0.163  0.3710  1.9080
ATOM   88  H28 HSA  1    -9.613   2.422  -0.893  0.0850  1.3870
ATOM   89  O31 HSA  1    -8.204   3.891  -0.849 -0.7460  1.7210
ATOM   90  H29 HSA  1    -7.721   4.338  -0.140  0.4390  0.0000
ATOM   91  C24 HSA  1   -10.081   3.851   0.685  0.0780  1.9080
ATOM   92  H30 HSA  1   -10.346   4.752   0.134  0.0940  1.3870
ATOM   93  O32 HSA  1   -11.301   3.132   0.967 -0.4240  1.6837
ATOM   94  S5  HSA  1   -12.277   2.534  -0.194  1.1110  2.0000
ATOM   95  O34 HSA  1   -11.765   1.268  -0.649 -0.6410  1.6612
ATOM   96  O35 HSA  1   -12.558   3.567  -1.154 -0.6410  1.6612
ATOM   97  O33 HSA  1   -13.662   2.199   0.563 -0.6410  1.7210
```

Appendix IV

Molecular Modelling Protocols

Gas phase simulated annealing (modified from file given by Dr. Barran):

```
#!/bin/csh

#   Simulated annealing using Amber7        ann7.job
# ------------------------------------------------
# does an initial minimization, runs dynamics at a
# specified temperature (hightemp) for a specified
# time (mdtime), cools the molecule from "hightemp"
# to "lowtemp" in "coolstep" steps in "cooltime" time,
# then minimizes again.  The final minimized structure
# is then used as input for a repeat of the above steps.
#
# time is in femtoseconds (generally, set by "dt" in the
# input file).
#
#   Edit this file as well as min7.in and md_x.in
#      before running or submitting to a queue.
#
#   Must have initial.crd and prm.top in the directory
#     given just below:
#


# -----------------------
# * Initial minimization *
# -----------------------

cp initial.crd prelim.crd
rm -f initial_min.out initial.crd initial_min.crd \
     mdinfo mden mdcrd mincor

/usr/progs/amber/exe/sander -i min7.in \
            -o initial_min.out \
            -p prm.top \
            -c prelim.crd \
            -r initial.crd          || goto error

cp initial.crd initial_min.crd
```

```
# --------------
# * Set limits *
# --------------

set index = 1
set j = 0
set limit = 100
set hightemp = 800
set lowtemp = 000
set numcoolsteps = 10
set mdtime = 30000
set cooltime = 10000
set ntx = 1

@ coolstep = ( ( $hightemp - $lowtemp ) / $numcoolsteps )
@ coolsteptime= ( $cooltime / $numcoolsteps )

sed "s/temp0_xxx/$hightemp/g" md_x.in > scr1.lst
sed "s/ntx_xxx/$ntx/g" scr1.lst > scr2.lst
sed "s/nstlim_xxx/$mdtime/g" scr2.lst > hotmd.in

# ---------------------
# * Beginning of loop *
# ---------------------

while ($index <= $limit)
  /bin/rm -f mdinfo mdcrd dyn.$index.out dyn.crd

#  --------------------------------
#  * dynamics at high temperature *
#  --------------------------------

   /usr/progs/amber/exe/sander -i hotmd.in \
            -o dyn.$index.out \
            -p prm.top \
            -c initial.crd \
            -r dyn.crd               || goto error

   set incrd = dyn.crd

#  --------------------
#  * stepwise cooling *
#  --------------------

   set ntx = 5
   set coolcyc = 1
   set temp = $hightemp

   sed "s/nstlim_xxx/$coolsteptime/g" md_x.in > scr1.lst
   sed "s/ntx_xxx/$ntx/g" scr1.lst > scr2.lst

#  ----------------
#  * cooling loop *
```

```
# ----------------

  while ($coolcyc <= $numcoolsteps)

     @ temp = ( $temp - $coolstep )
     sed "s/temp0_xxx/$temp/g" scr2.lst > coolmd.in
     set outcrd = crd.$index.$coolcyc.crd
     rm -f cooldown.$index.$coolcyc.out \
         $outcrd mdinfo mden mdcrd mincor

     /usr/progs/amber/exe/sander -i coolmd.in \
           -o cooldown.$index.$coolcyc.out \
           -p prm.top \
           -c $incrd \
           -r $outcrd                        || goto error

     set incrd = $outcrd

     @ coolcyc++
  end

# -----------------------
# * energy minimization *
# -----------------------

  /bin/rm -f  mdinfo mdcrd min.$index.out min.$index.crd
  /bin/rm -f  initial.crd mden mincor
   /usr/progs/amber/exe/sander -i min8.in \
            -o min.$index.out \
            -p prm.top \
            -c $incrd \
            -r min.$index.crd          || goto error

  cp min.$index.crd initial.crd

  @ index++

# -----------------------
# * clean up some files *
# -----------------------

if $j > 0 then
 rm -f dyn.$j.*
 rm -f cooldown.$j.*.out
 rm -f crd.$j.*.crd
endif

  @ j++
# ---------------
# * End of loop *
# ---------------

end
```

```
#-----------------
# * Final cleanup *
#-----------------

cp prelim.crd initial.crd
/bin/rm -f dyn.$j.* cooldown.$j.*.out crd.$j.*.crd
/bin/rm -f scr1.lst scr2.lst scr3.lst 0 dyn.crd prelim.crd \
        mdinfo dyn.$index.out

echo " "
echo No errors detected
exit(0)

error:
echo " --Failure: look at *.out, fix problem and run this script again"
echo " "
exit(1)
```

md_x.in

```
Amber7 sander dynamics
 &cntrl
  nstlim = nstlim_xxx, temp0 = temp0_xxx, ntx = ntx_xxx,
  ntc = 2, ntf = 2, ntt = 1, ntb = 0, ntpr = 1000,
 &end
```

min7.in

```
minimization with sander (Amber7)
 &cntrl
  imin = 1, maxcyc = 500000, drms = 0.005, ntb = 0,
  ntpr = 1000, cut = 999., ncyc = 3000,
 &end
```

min8.in

```
minimization with sander (Amber7)
 &cntrl
  imin = 1, maxcyc = 500000, drms = 0.0005, ntb = 0,
  ntpr = 1000, cut = 999., ncyc = 3000,
 &end
```

Gas phase (Φ, Ψ) map calculation:

min-res.in

```
Minimization with restraints on the glycosidic angles
 &cntrl
  imin = 1, maxcyc = 50000, drms = 0.0005, ntb = 0,
  ntpr = 100, cut = 999., ncyc = 1000,
  nmropt = 1,
 &end

 &wt
   type = 'END',
 &end
DISANG=restraints.dat
LISTIN=POUT
LISTOUT=POUT
```

Solution molecular dynamics:

min1.in

```
Initial minimization with Cartesian restraints on solute to reduce bad VDW on solvent
 &cntrl
  imin = 1, maxcyc = 1000, ncyc = 500,
  drms = 0.0005, ntb = 1, cut = 10, ntr = 1,
 &end
Group input for restrained atoms
 500
RES 1
END
END
```

min2.in

```
Minimization on the whole system
 &cntrl
  imin = 1, maxcyc = 3000, ncyc = 1000,
  drms = 0.0005, ntb = 1, cut = 10, ntr = 0,
 &end
```

md1.in

```
Group input for restraint atoms
 10.0
RES 1 1
END
END
MD heating from 0 to 300K with restraints on the solute
 &cntrl
   imin = 0, ntpr = 100, ntwx = 100, ntwe = 100, ntwr = 1000,
   nstlim =10000, dt = 0.002,
   ntb = 1,cut = 10, ntr = 1, ntc = 2, ntf = 2, nmropt = 1,
   tempi = 0.0, temp0 = 300.0, ntt = 1, tautp = 0.5,
```

```
&end
&wt
 type='TEMP0', istep1=0,     istep2=1000,
          value1=0.0,  value2=300.0,
 &end
 &wt
 type='TEMP0', istep1=1000,   istep2=20000,
          value1=300.0, value2=300.0,
 &end
 &wt
 type='END',
 &end
```

md2.in

MD equilibration on the whole system at 300K

```
 &cntrl
  imin = 0, ntpr = 100, ntwx = 100, ntwe = 100, ntwr = 1000,
  nstlim = 50000, dt = 0.002, irest = 1, ntx = 5,
  ntb = 2, pres0 = 1.0, ntp = 1, taup = 2.0,
  cut = 10, ntr = 0, ntc = 2, ntf = 2,
  tempi = 300.0, temp0 = 300.0, ntt = 1,
 &end
```

md-production.in

Free MD production run

```
 &cntrl
  imin = 0, ntpr = 200, ntwx = 200, ntwe = 200, ntwr = 10000, ntwv = 200,
  nstlim = 200000, dt = 0.002, irest = 1, ntx = 5,
  ntb = 2, pres0 = 1.0, ntp = 1, taup = 2.0,
  cut = 8, ntr = 0, ntc = 2, ntf = 2,
  tempi = 300.0, temp0 = 300.0, ntt = 1,
 &end
```

md-RDC-refinement.in

MD production run with RDCs restraints on A and C rings

```
 &cntrl
  imin = 0, ntpr = 200, ntwx = 200, ntwe = 200, ntwr = 20000, ntwv = 200,
  nstlim = 200000, dt = 0.002, irest = 0, ntx = 5,
  ntb = 2, pres0 = 1.0, ntp = 1, taup = 2.0,
  cut = 8, ntr = 0, ntc = 2, ntf = 2,
  tempi = 300.0, temp0 = 300.0, ntt = 1,
  iscale = 5, nmropt = 1,
 &end

 &wt type = 'END',  &end

LISTOUT = POUT
DIPOLE = RST-AC.f
```

RST-AC.f

RDCs restraints of A and C rings

```
 &align
   ndip = 18, dcut = -1.0, gigj = 10*7.847, 8*31.199, dwt = 18*1.0,
   s11 = -5.357, s22 = 8.746, s12 = -1.867, s13 = -1.617, s23 = -13.253,
   id(1) = 12, jd(1) = 13, dobs(1) = -3.4,
   id(2) = 14, jd(2) = 15, dobs(2) = 5.8,
   id(3) = 16, jd(3) = 17, dobs(3) = 6.1,
   id(4) = 28, jd(4) = 29, dobs(4) = 5.4,
   id(5) = 9, jd(5) = 10, dobs(5) = 6.2,
   id(6) = 53, jd(6) = 54, dobs(6) = 5.1,
   id(7) = 66, jd(7) = 67, dobs(7) = -3.1,
   id(8) = 74, jd(8) = 75, dobs(8) = -3.0,
   id(9) = 78, jd(9) = 79, dobs(9) = -4.8,
   id(10) = 56, jd(10) = 57, dobs(10) = -3.6,
   id(11) = 13, jd(11) = 15, dobs(11) = -0.82,
   id(12) = 15, jd(12) = 17, dobs(12) = 0.88,
   id(13) = 17, jd(13) = 29, dobs(13) = 0.96,
   id(14) = 29, jd(14) = 10, dobs(14) = 0.90,
   id(15) = 54, jd(15) = 67, dobs(15) = -1.21,
   id(16) = 67, jd(16) = 75, dobs(16) = 0.55,
   id(17) = 75, jd(17) = 79, dobs(17) = -0.61,
   id(18) = 79, jd(18) = 57, dobs(18) = -1.36,
 &end
```

Appendix V

Pulse sequences of new NMR methods

HOHAHA-COSY:

```
; 1D double-selective COSY with an initial one-step selective HOHAHA step
; R. Konrat, I. Burghardt and G. Bodenhausen J. Am. Chem. Soc. 1991, 113, 9135-9140
; The magnetization is transferred from the on-resonance proton and the proton to which
; the magnetization is transferred is inverted.
; Selective spin-lock is used

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

"d12=20u"
"d13=4u"
"p18=p13-200u"
"TAU=2*d16+p16"
"DELTA1=0.5*d2-TAU"
"DELTA2=0.5*d2-2*TAU-p12"
"DELTA3=0.5*d3-2*TAU"
"DELTA4=0.5*d3-p11-TAU-4u"

1 ze
2 20u BLKGRAD
 d1 pl1:f1
  d12 pl1:f1
  d12  fq=cnst1(bf ppm):f1

/* spfgse */
 p1 ph1
 d12  UNBLKGRAD
 p16:gp6
 d16  pl0:f1
  p11:sp3 ph6
 d12
 p16:gp6
 d16

/* sel. spin-lock */
```

```
  4u
    p21:sp21:f1 ph21
  4u

/* purging pulse */
# ifdef PURGE
  d12 pl1:f1
    p1 ph7:r
    d12  fq=cnst2(bf ppm):f1
    d12 pl0:f1
      (p13:sp13 ph30) (100u p18:gp5 100u)
    d13
    d12 pl1:f1
    p1 ph1
# endif

# ifdef SPFGSE

  d13          ;SEL 180 pulse
  p16:gp3
  d16   pl0:f1
    p11:sp1 ph3
  d13
  p16:gp3
  d16
# endif

 DELTA1
    d16
    p16:gp1
    d16
      (p11:sp1 ph4)
      4u
      (p12:sp2 ph8)
    d16
    p16:gp1
    d16
  DELTA2
  d16                pl1:f1
  p16:gp2
  d16
    p1 ph2
# ifdef REFOCUS
    d16
    p16:gp2
    d16
      DELTA3
      d16
      p16:gp4
      d16
        (p12:sp2 ph5)
        4u
        (p11:sp1 ph5)
```

```
    d16
    p16:gp4
    d16
    DELTA4
# else
 d16
 p16:gp2
 d16
# endif

# ifdef SPFGSE
 go=2 ph31
# else
 go=2 ph29
# endif
 20u BLKGRAD
 wr #0
exit

ph1=0 1
ph2=1
ph3=0
ph4=0 0 1 1
ph5=0 0 0 0 1 1 1 1
      2 2 2 2 3 3 3 3
ph6={0}*16 {1}*16 {2}*16 {3}*16
ph7=0
ph8={0}*64 {2}*64
ph21=1
ph30=1
ph31=0 2 2 0 2 0 0 2
ph29=0 2 0 2 2 0 2 0

;pl1 : power level for pulse (default = -5dB)
;p1 : 90 degree high power pulse (7.5us)
;pl0 : 120dB
;sp1: power for p11 (180 Gauss) H-2
;sp3: power for p11 (180 Gauss) H-1
;sp2: power for p12 (180 Gauss) H-3

;cnst1 - ppm of H-1
;cnst2 - ppm of H-2 (=o1p)

;spoff1=0
;spoff2=H3-H2
;spoff3=0
;p21 : spin-lock time (1/J)
;sp21 : Rectangular shaped pulse (J/4)*sqrt(4*n*n -1), n=10, J=7.936Hz => sp21 = 39.6Hz
;p1 : 90 degree high power pulse
;d1 : relaxation delay; 1-5 * T1
;NS : 4*n
;DS : 2
;phcor21 : phase correction (0 degree)
```

CSSF-COSY:

```
; 1D double-selective COSY (variable time) with a chemical shift selective filter
; P. T. Robinson, T. N. Pham and D. Uhrín J. Magn. Reson. 2001, 151, 97-103

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

"d12=20u"
"d13=4u"
"p2=p1*2"

"TAU=2*d16+p16"
"DELTA1=0.5*d2-TAU-d12"
"DELTA2=0.5*d2-2*TAU"
"DELTA3=0.5*d3-2*TAU"
"DELTA4=0.5*d3-p12-TAU-4u"

1 ze
2 30m
3 50u BLKGRAD

#ifdef PRESAT
  d12  fq=cnst1(bf ppm):f1
  d12 pl9:f1
  d1 cw:f1 ph29
  d13 do:f1
  d12 pl1:f1
  d12  fq=cnst2(bf ppm):f1
  d12 UNBLKGRAD
# else
  d1 pl1:f1
  d12 UNBLKGRAD
#endif

;[cssf]
  p1 ph1
  d23*0.5
  p16:gp1
  d16 pl0:f1
  24u
  (p11:sp3 ph2:r):f1
  4u
  d12 pl1:f1
  p2 ph1
  p16:gp2
  d16
  d23*0.5
  p1 ph5
  p16:gp3
  d16
  p1 ph6
```

```
;[cosy]
 d12 pl0:f1
    (p12:sp2 ph8)
 DELTA1
   d16
   p16:gp4
   d16
    (p12:sp1 ph3)
     4u
    (p12:sp2 ph8)
   d16
   p16:gp4
   d16
  DELTA2
  d16              pl1:f1
  p16:gp5
  d16
   p1 ph7

# ifdef REFOCUS
   d16
   p16:gp5
   d16
     DELTA3 pl0:f1
     d16
     p16:gp6
     d16
       (p12:sp2 ph4)
        4u
       (p12:sp1 ph4)
     d16
     p16:gp6
     d16
     DELTA4
# else
  d16
  p16:gp5
  d16
# endif

  go=2 ph31
  30m mc #0 to 2 F0(id23 & zd)

  50u rd23
  lo to 3 times l3

  wr #0
  4u BLKGRAD.
exit

ph1=0
ph2=0 1
ph3=0 0 1 1 2 2 3 3
```

```
ph8={0}*8 {1}*8 {2}*8 {3}*8
ph4={0}*32 {2}*32
ph5=0
ph6=0
ph7=1
ph29=0
ph31=0 2 2 0

;pl0 : 120dB
;pl1 : f1 channel - power level for pulse (default)
;sp1 : f1 channel - shaped pulse 180 degree
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel -  180 degree high power pulse
;p11:  180 CSSF shaped pulse
;sp3: for p11
;p12: 180 selective COSY shaped pulses
;sp1: on-reson p12
;sp2: off-reson p12
;p16: homospoil/gradient pulse

;d1 : relaxation delay; 1-5 * T1
;d2 : defocusing delay
;d3 : refocusing delay
;d12: delay for power switching
;d16: delay for homospoil/gradient recovery
;d23: CSSF delay
;id23: increment of CSSF
;NS: 8 * TD0 * l3
;DS: 4
;l3 long term accum. Total NS=NS*Td0*l3
;use gradient ratio:    gp 1 : gp 2 =  40 :  -40

;for z-only gradients:
;gpz1: 40%
;gpz2: -40%
;gpz3: 5%
;gpz4: 11%
;gpz5: 4%
;gpz6: 17%

;use gradient files:
;gpnam1: CRP.100
;gpnam2: CRP.100
;gpnam3: CRP.100
;gpnam4: CRP.100
;gpnam1: CRP.100
;gpnam5: CRP.100
;gpnam6: CRP.100
```

DJM-REVINEPT-INADEQUATE:

```
; 2D doublet J-modulated reversed-INEPT-INADEQUATE starts on carbon
; both one-bond and long-range proton-carbon correlation and coupling constants obtained
; phase sensitive and phase cycling according to:
; M. Bourdonneau & B. Ancian, J. Magn. Reson. 132, 316-327 (1998)
; BEBOP and BIBOP pulses are used to cover the whole 13C chemical shift range
; (a) T. E. Skinner, T. O. Reiss, B. Luy, N. Khaneja and S. J. Glaser J. Magn. Reson. 2004,
; 167, 68-74; (b) B. Luy, K. Kobzar, T. E. Skinner, N. Khaneja and S. J. Glaser
; J. Magn. Reson. 2005, 176, 179-186

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

"p2=p1*2"
"p4=p3*2"
"d2=1s/(cnst2*4)"  ;JCH
;"d3=1s/(cnst2*4)"  ;kJCH
"d11=30m"
"d12=20u"
"DELTA2=d2-p16-d16"
"DELTA3=d3-p19-d16"
"DELTA4=2*d0"
"p18=2ms"

# ifdef F1180
"d0=in0/2 + 1u"
# else
"d0=3u"
#endif

"d20=3u"
"in20=cnst1*in0"
"in21=cnst1*in0"
"l3=(td1/2)"
"d21=(in21+0.5*in0)*l3+4u"
"TAU=d1-2*d11-4*d21"

;---------------------- heating compensation
1 ze
 d11 pl12:f2
2 d11 do:f2
 d12 pl1:f1   ;removing the effects of 13C decoupling
 50u UNBLKGRAD
 p16:gp1
 d16
 (p1 ph1):f1
 4u
 (p1 ph1):f1
 p16:gp2
 d16 BLKGRAD
 d0*2.0
```

```
 d20*4.0
 d12 pl2:f2 pl19:f1
 d11 cpds1:f1 ph1
 TAU
 4u do:f1
 d12 pl18:f1
 d12 cpds1:f1 ph1
 d21*4
;--------------------- heating compensation

50u UNBLKGRAD

# ifdef BB90   ; real start
3  (p11:sp1 ph3):f2          ; shifted because of shape by 90deg
  4u
# else
3 (p3 ph3):f2
  4u
# endif

  d20

# ifdef BB180
  4u
  (p12:sp2 ph3):f2
  4u
# else
 (p4 ph3):f2
# endif

  4u
  d20

 (p3 ph3):f2  ;MQ generated

 d0*2.0
 d12 do:f1
 p16:gp5*EA

# ifdef BB180
 d16 pl0:f2
 (p12:sp2 ph1):f2
# else
 d16
 (p4 ph1):f2
# endif

# ifdef F1180
 1u
# else
 DELTA4
#endif
 p16:gp6*EA
```

```
 d16 pl2:f2
 d12 cpds1:f1 ph1

 (p3 ph1):f2

 d20  ;refocusing starts here

# ifdef BB180
   4u
   (p12:sp2 ph4):f2
   4u
# else
 (p4 ph4):f2
# endif

 d20
 4u do:f1
 DELTA3 pl1:f1
 p19:gp8
 d16

# ifdef BB180
   (center (p12:sp2 ph5):f2 (p2 ph1):f1)
# else
 (center (p4 ph5):f2 (p2 ph1):f1)
# endif

 p19:gp8
 d16
 DELTA3
 (p3 ph2):f2
 p16:gp9
 d16

 p1 ph1
 d2

# ifdef BB180
   (center (p12:sp2 ph1):f2 (p2 ph1):f1)
# else
 (center (p4 ph1):f2 (p2 ph1):f1)
# endif

 p16:gp7
 d16
 DELTA2 pl12:f2 BLKGRAD

 go=2 ph31 cpd2:f2
 d11 do:f2 mc #0 to 2 F1EA(igrad EA, id0 & id20 & dd21)
 d11 do:f2
exit

ph0=0
```

```
ph1=0
ph2=1
ph3=0 1 2 3
ph4=0
ph5=2
ph31=0 2

;pl1 : f1 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d3 : 1/(4*1JCH) for CH, 1/(8*1JCH) for CHx
;d4 : 1/(4J(CC))
;d11: delay for disk I/O                    [30 msec]

;cnst1: scaling factor
;cnst2: = J(CH)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 1
;d21 : compensation for heating during increasing t1
;NS: 32 * n
;DS: 16
;td1: number of experiments
;l1=td1/2: number of real points
;FnMODE: Echo:antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;gpz1: 20
;gpz2: 35
;gpz5: 40.0
;gpz6: -40.0
;gpz7: 40.16
;gpz8: 7
;gpz9: 90

;Processing

;PHC0(F1): 90
;PHC1(F1): -180
;FCOR(F1): 1

;$Id: inadph,v 1.3 2002/06/12 09:04:58 ber Exp $
```

JM-DEPT-INADEQUATE:

```
; 2D DEPT-INADEQUATE experiment yields interleaved in-phase/antiphase spectra in F1
; two fixed evolution intervals are optimized for long-range couplings and one-bond
; correlations are suppressed
```

```
; gradient selections are used
; adiabatic purge pulse/gradient block used according to
; M. J. Thrippleton and J. Keeler Angew. Chem. Int. Ed. Engl..2003, 42, 3938-3941

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

prosol relations=<triple>

"p2=p1*2"
"p4=p3*2"

# ifdef F1180
"d0=0.5*in0"
"DELTA1=1u"
# else
"d0=3u"
"DELTA1=2*d0"
#endif

"d20=4u"

"d3=1s/(cnst2*2)"
"d5=1s/(cnst4*4)"       ; nJCC
"d6=1s/(cnst3*4)"       ; 1JCC
"TAU=d5-p19-d16"
"TAU1=d6-p16-d16"

"in20=cnst1*in0"
"in21=(cnst1)*in0"
"l3=(td1/4)"
"d21=(in21+in0)*l3+4u"
"d11=30m"
"d13=4u"
"d12=20u"

"DELTA=d3-p16-d16"
"DELTA3=d5-p19-d16-20u" ;nJCC
"DELTA4=d5-p19-d16-2.0*d3-20u"  ;nJCC
"DELTA6=d6-d3-p1-p16-d16" ;1JCC

"l0=1"

1 ze
  d11 pl12:f2
2 d11 do:f2 BLKGRAD

  d12 pl19:f1
  20u cpds1:f1 ph1      ;1H dec on
  d21
  d21
```

```
if "l0 %2 == 0"        ; heat compensation for antiphase multiplets
 {
 DELTA3
 DELTA3
 }

 d12 do:f1             ; 1H dec off
 d12 pl1:f1
  50u UNBLKGRAD

  p16:gp6
  d16
  (p1 ph1):f1
  4u
  (p1 ph2):f1
  p16:gp7
  d16 BLKGRAD

  d1 pl1:f1

3 (p1 ph1)      ; this is the real start
  d3 pl2:f2
  (ralign (p2 ph1) (p3 ph3):f2) ;90 deg 13C pulse
  d3 UNBLKGRAD
  (p1 ph2)
  DELTA6
  p16:gp4
  d16

# ifdef COMP2
  d12 pl0:f2
  (p24:sp7 ph3):f2
  d12
# else
  (p4 ph3):f2
# endif
  p16:gp4
  d16 pl2:f2
  TAU1
;----------------

  (p3 ph2):f2  ;purge one-bond

  d12 pl0:f2
  10u gron0
  (p30:sp30 ph11):f2
  20u groff
  p19:gp8
  d16 pl19:f1
  d12 pl2:f2
  20u cpds1:f1 ph1      ;1H dec on

  (p3 ph8):f2
```

```
# ifdef COMP2
 d20 pl0:f2
 4u
 (p24:sp7 ph10):f2
 4u
 d20 pl2:f2
# else
 d20
 (p4 ph10):f2
 d20
# endif
;------------------

if "l0 %2 == 1"        ; acquistion of inphase multiplets
 {
 (p3 ph8):f2   ;purge antiphase
 20u do:f1     ;1H dec off
 d12 pl0:f2
 10u gron0
 (p30:sp30 ph11):f2
 20u groff
 p19:gp9
 d16 pl2:f2
 20u cpds1:f1 ph1      ;1H dec on

 (p3 ph9):f2

# ifdef COMP2   ;evolve nJCC
 DELTA3
 20u do:f1     ;1H dec off
 p19:gp10
 d16 pl0:f2
 4u
 (p24:sp7 ph1):f2
 4u
 p19:gp10
 d16
 20u cpds1:f1 ph1      ;1H dec on

 DELTA3 pl2:f2
# else
 DELTA3
 p19:gp10
 d16 pl0:f2
 4u
 (p4 ph3):f2
 4u
 p19:gp10
 d16
 DELTA3 pl2:f2
# endif
}
```

```
;----------------------------
 (p3 ph2):f2   ;MQ 13C13C created
 d0
 d0
 20u do:f1    ;1H dec off
 p16:gp1*EA
 d16

# ifdef COMP2
 d12 pl0:f2
 (p24:sp7 ph7):f2
 d12
# else
 (p4 ph7):f2
# endif

 DELTA1
 p16:gp2*EA
 d16 pl2:f2
 20u cpds1:f1 ph1     ;1H dec on
 (p3 ph5):f2 ; MQ -- SQ transfer
 TAU
 20u do:f1    ;1H dec off
 p19:gp5
 d16
 d13

# ifdef COMP2
 d12 pl0:f2
 (p24:sp7 ph12):f2
 d12
# else
 (p4 ph1):f2
# endif

 d13
 p19:gp5
 d16 pl2:f2
 20u cpds1:f1 ph1     ;1H dec on

 DELTA4

20u do:f1    ;1H dec off
 d3 pl1:f1
 (p1 ph1):f1
 d3
 (lalign (p2 ph1) (p3 ph6):f2)
 d13
 DELTA

 p16:gp3
 d16 BLKGRAD pl12:f2
```

```
 go=2 ph31  cpd2:f2
 d11 do:f2 mc #0 to 2
    F1I(iu0, 2)
   F1EA(igrad EA, id0 & id20 & dd21)
exit

ph1=0
ph2=1
ph3=0 2
ph4=1
ph5=0
ph6=1
;ph6=1 1 1 1 3 3 3 3
ph7=0 0 1 1
ph8=3
ph9=1
ph10=0 0 0 0 2 2 2 2
ph11=0
ph12=0 0 0 0 0 0 0 0
       1 1 1 1 1 1 1 1
ph31=0 2 0 2 0 2 0 2
       2 0 2 0 2 0 2 0
;ph31=0 2 0 2 2 0 2 0
       2 0 2 0 0 2 0 2

;pl1 : f1 channel - power level for pulse (default)
;pl2 : f2 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;sp7: composite adiabatic 180: Crp60comp.4
;sp13:180: Crp60,0.5,20.1
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse
;p3 : f2 channel -  90 degree high power pulse
;p4 : f2 channel - 180 degree high power pulse
;p24: composite adiabatic 180: 2ms
;p8: Crp60,0.5,20.1 500us
;p16: homospoil/gradient pulse

;d0 : incremented delay (2D)              [3 usec]
;d1 : relaxation delay; 1-5 * T1
;d11: delay for disk I/O                    [30 msec]
;d13: short delay                         [4 usec]
;d16: delay for homospoil/gradient recovery

;cnst1: = scaling factor
;cnst2: = 1J(CH)
;cnst3: = 1J(CC)
;cnst4: = nJ(CC)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 2
;NS: 1 * n
;DS: >= 16
;td1: number of experiments
```

```
;FnMODE: echo-antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;use gradient ratio:   gp 1 : gp 2 : gp 3  =  40 : -40  : 40.16   for C-13

;for z-only gradients:
;gpz0: 4%
;gpz1: 40%
;gpz2: -40%
;gpz3: 40.16%
;gpz4: 57%
;gpz5: 77%
;gpz6: 60%
;gpz7: 70%
;gpz8: 8%
;gpz9: 15%
;gpz10: 90%

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

;$Id: invietgpsi,v 1.6 2000/05/08 11:40:23 eng Exp $
```

REVINEPT-INADEQUATE:

```
; 2D reversed-INEPT-INADEQUATE experiment yields interleaved in-phase/antiphase
; correlations only with gradient selection
; phase sensitive and phase cycling according to:
; M. Bourdonneau & B. Ancian, J. Magn. Reson. 132, 316-327 (1998)

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

"p2=p1*2"
"p4=p3*2"

"d2=1s/(cnst2*4)"  ;JCH
;"d3=1s/(cnst2*4)"  ;kJCH
"d4=1s/(cnst3*4)"  ;nJCC
"d5=1s/(cnst3*4)+1s/(cnst1*8)"  ;delay for alternative exp.
"d11=30m"
"d12=20u"

"DELTA1=d4-d3-d12"
"DELTA3=d5-d3-d12"     ; for alternative experiment
"DELTA2=d2-p16-d16"

# ifdef F1180
```

```
"d0=in0/2 + 1u"
# else
"d0=3u"
#endif

"DELTA4=2*d0"
"p18=2ms"
"l0=1"

"in21=in0"
"l3=(td1/4)"
"d21=in0*l3+4u"
"TAU=d1-2*d11-2*d21"

1 ze
 d11 pl12:f2
2 d11 do:f2 pl1:f1

 50u UNBLKGRAD
 p16:gp1      ;purge
 d16
 (p1 ph1):f1
 4u
 (p1 ph2):f1
 p16:gp2
 d16 BLKGRAD

;--------------------- heating compensation
 d0*2.0
 d12 pl2:f2 pl19:f1
 d11 cpds1:f1 ph1
 TAU
 4u do:f1
 d12 pl18:f1
 d12 cpds1:f1 ph2
 d21*2.0
;--------------------- heating compensation

 50u UNBLKGRAD

# ifdef BB90   ; real start
3  (p11:sp1 ph3):f2          ; shifted because of shape by 90deg
   4u
# else
3 (p3 ph3):f2
   4u
# endif

 if "l0 %2 == 1"
    {
  d4
    }
  else
```

```
    {
   d5
    }

# ifdef BB180
    4u
    (p12:sp2 ph3):f2
    4u
# else
  (p4 ph3):f2
# endif

 if "l0 %2 == 1"
    {
   d4
    }
  else
    {
   d5
    }
   4u

  (p3 ph3):f2   ;MQ generated

  d0*2.0
  d12 do:f1
  p16:gp5*EA

# ifdef BB180
  d16 pl0:f2
  (p12:sp2 ph1):f2
# else
  d16
  (p4 ph1):f2
# endif

# ifdef F1180
  1u
# else
  DELTA4
#endif

  p16:gp6*EA
  d16 pl2:f2
  d12 cpds1:f1 ph1

  (p3 ph4):f2

;refocusing starts here
 if "l0 %2 == 1"
    {
   d4
    }
```

```
  else
    {
   d5
    }

# ifdef BB180
   4u
   (p12:sp2 ph5):f2
   4u
# else
  (p4 ph5):f2
# endif

 if "l0 %2 == 1"
    {
  DELTA1
    }
  else
    {
  DELTA3
    }

  d12 do:f1
  d3 pl1:f1  ; optimize d3 for a compromise between CH and CH2
  (p3 ph2):f2
  p18:gp3
  d16
  p1 ph1
  d2

# ifdef BB180
   (center (p12:sp2 ph1):f2 (p2 ph1):f1)
# else
  (center (p4 ph1):f2 (p2 ph1):f1)
# endif

   p16:gp7 ;40.16
  d16
  DELTA2 pl12:f2 BLKGRAD

  go=2 ph31 cpd2:f2
  d11 do:f2 mc #0 to 2
  F1I(iu0, 2)
  F1EA(igrad EA, id0 & dd21)
  d11 do:f2
exit

ph0=0
ph1=0
ph2=1
ph3=0 1 2 3
ph4=0 0 0 0 2 2 2 2
ph5=0 0 0 0 0 0 0 0
```

```
      1 1 1 1 1 1 1 1
ph31=0 2 0 2 2 0 2 0
      2 0 2 0 0 2 0 2
;pl1 : f1 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse

;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d3 : 1/(2*1JCH) for CH, 1/(4*1JCH) for CHx
;d4 : 1/(4J(CC))
;d11: delay for disk I/O                    [30 msec]

;cnst1: = 1J(CC)
;cnst2: = J(CH)
;cnst3: = nJ(CC)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 1
;NS: 32 * n
;DS: 16
;td1: number of experiments
;l1=td1/2: number of real points
;FnMODE: States-TPPI, TPPI, States or QSEC
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;;gpz1: 20%
;;gpz2: 35%
;gpz3: 90%
;gpz5: 40%
;gpz6: -40%
;gpz7: 40.16

;Processing

;PHC0(F1): 90
;PHC1(F1): -180
;FCOR(F1): 1

;$Id: inadph,v 1.3 2002/06/12 09:04:58 ber Exp $
```

*J*M-REVINEPT-INADEQUATE:

```
; 2D reversed-INEPT-INADEQUATE J-modulated experiment yields interleaved in-phase/
; antiphase spectra of one-bond and long-range proton-carbon correlation and coupling
; constants in F1 with gradient selection
; phase sensitive and phase cycling according to:
; M. Bourdonneau & B. Ancian, J. Magn. Reson. 132, 316-327 (1998)

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

```
#include <Grad.incl>

"p2=p1*2"
"p4=p3*2"

"d2=1s/(cnst2*4)"  ;JCH
;"d3=1s/(cnst2*4)"  ;kJCH
"d4=1s/(cnst3*4)"  ;nJCC
"d11=30m"
"d12=20u"

"DELTA1=d4-d3-d12"
"DELTA2=d2-p16-d16"
"DELTA3=2*d0"
"DELTA4=d4-p16-d16-20u" ;nJCC

"p18=2ms"

# ifdef F1180
"d0=in0/2 + 1u"
# else
"d0=3u"
#endif

"d20=3u"
"in20=cnst1*in0"
"in21=cnst1*in0"
"l3=(td1/4)"
"d21=(in21+in0)*l3+4u"
"TAU=d1-2*d11-2*d21"

"l0=1"

1 ze
  d11 pl12:f2
2 d11 do:f2 pl1:f1

  50u UNBLKGRAD
  p16:gp1      ;purge
  d16
  (p1 ph1):f1
  4u
  (p1 ph2):f1
  p16:gp2
  d16 BLKGRAD

;---------------------- heating compensation
  d0*2.0
  d20*2.0
  d12 pl2:f2 pl19:f1
  d11 cpds1:f1 ph1
  TAU
  4u do:f1
```

```
 d12 pl18:f1
 d12 cpds1:f1 ph1
 d21*2

if "l0 %2 == 0"       ; heat compensation for antiphase multiplets
 {
 DELTA4
 DELTA4
 }
;---------------------- heating compensation

 50u UNBLKGRAD

if "l0 %2 == 1"       ; acquistion of inphase multiplets
 {
# ifdef BB90   ; real start
3  (p11:sp1 ph1):f2          ; shifted because of shape by 90deg
   4u
# else
3 (p3 ph1):f2
   4u
# endif

  d20

# ifdef BB180
   4u
   (p12:sp2 ph1):f2
   4u
# else
 (p4 ph1):f2
# endif

  4u
  d20

 (p3 ph1):f2  ;purge antiphase
 20u do:f1    ;1H dec off
 d12 pl0:f2
 10u gron0
 (p30:sp30 ph1):f2
 20u groff
 p16:gp9
 d16 pl2:f2
 20u cpds1:f1 ph1     ;1H dec on
 (p3 ph3):f2

# ifdef BB180  ;evolve nJCC
 DELTA4
 20u do:f1    ;1H dec off
 p16:gp10
 d16 pl0:f2
 4u
```

```
 (p12:sp2 ph3):f2
 4u
 p16:gp10
 d16
 20u cpds1:f1 ph1     ;1H dec on
 DELTA4 pl2:f2
# else
 20u do:f1    ;1H dec off
 DELTA4
 p16:gp10
 d16 pl0:f2
 4u
 (p4 ph3):f2
 4u
 p16:gp10
 d16
 20u cpds1:f1 ph1     ;1H dec on
 DELTA4 pl2:f2
# endif
}
 else
 {

# ifdef BB90   ; real start
4  (p11:sp1 ph3):f2          ; shifted because of shape by 90deg
   4u
# else
4 (p3 ph3):f2
   4u
# endif

  d20

# ifdef BB180
   4u
   (p12:sp2 ph3):f2
   4u
# else
 (p4 ph3):f2
# endif

  4u
  d20
 }

 (p3 ph3):f2  ;MQ generated

 d0*2.0
 d12 do:f1
 p16:gp5*EA

# ifdef BB180
 d16 pl0:f2
```

```
 (p12:sp2 ph1):f2
# else
 d16
 (p4 ph1):f2
# endif

# ifdef F1180
 1u
# else
 DELTA3
#endif

 p16:gp6*EA
 d16 pl2:f2
 d12 cpds1:f1 ph1

 (p3 ph4):f2

 d4   ;refocusing starts here

# ifdef BB180
   4u
   (p12:sp2 ph5):f2
   4u
# else
 (p4 ph5):f2
# endif

 DELTA1
 d12 do:f1
 d3 pl1:f1  ; optimize d3 for a compromise between CH and CH2
 (p3 ph2):f2
 p18:gp3
 d16
 p1 ph1 ;back to proton
 d2

# ifdef BB180
   (center (p12:sp2 ph1):f2 (p2 ph1):f1)
# else
 (center (p4 ph1):f2 (p2 ph1):f1)
# endif

 p16:gp7
 d16
 DELTA2 pl12:f2 BLKGRAD

 go=2 ph31 cpd2:f2
 d11 do:f2 mc #0 to 2
 F1I(iu0, 2)
 F1EA(igrad EA, id0 & id20 & dd21)
 d11 do:f2
exit
```

```
ph1=0
ph2=1
ph3=0 1 2 3
ph4=0 0 0 0 2 2 2 2
ph5=0 0 0 0 0 0 0 0
    1 1 1 1 1 1 1 1
ph31=0 2 0 2 2 0 2 0
      2 0 2 0 0 2 0 2

;pl1 : f1 channel - power level for pulse (default)
;pl12: f2 channel - power level for CPD/BB decoupling
;p1 : f1 channel -  90 degree high power pulse
;p2 : f1 channel - 180 degree high power pulse

;d0 : incremented delay (2D)
;d1 : relaxation delay; 1-5 * T1
;d3 : 1/(2*1JCH) for CH, 1/(4*1JCH) for CHx
;d4 : 1/(4J(CC))
;d11: delay for disk I/O                    [30 msec]

;cnst1: scaling factor
;cnst2: = J(CH)
;cnst3: = J(CC)
;in0: 1/(2 * SW(X)) = DW(X)
;nd0: 1
;d21 : compensation for heating during increasing t1
;NS: 32 * n
;DS: 16
;td1: number of experiments
;l1=td1/2: number of real points
;FnMODE: Echo:antiecho
;cpd2: decoupling according to sequence defined by cpdprg2
;pcpd2: f2 channel - 90 degree pulse for decoupling sequence

;gpz0: 4%
;gpz1: 67%
;gpz2: 77%
;gpz3: 90%
;gpz5: 40%
;gpz6: -40%
;gpz7: 40.16%
;gpz9: 13%
;gpz10: 7%

;Processing
;PHC0(F1): 90
;PHC1(F1): -180
;FCOR(F1): 1

;$Id: inadph,v 1.3 2002/06/12 09:04:58 ber Exp $
```

Bibliography

[1] B. Mulloy and M. J. Forster *Glycobiology* 2000, **10**, 1147-1156

[2] (a) I. Capilar and R. J. Linhardt *Angew. Chem. Int. Ed* 2002, **41**, 390-412; (b) J. Angulo, P. M. Nieto and M. Martin-Lomas *Chem. Commum (Camb)* 2003, **13**, 1512-1513

[3] (a) J. R. S. Day, R. C. Landis and K. M. Taylor *J. Cardiothorac. Vasc. Anesth.* 2004, **18**, 93-100; (b) B. Mulloy and R. J. Linhardt *Curr. Opin. Struct. Biol.* 2001, **11**, 623-628

[4] (a) S. Faham, R. E. Hileman, J. R. Fromm, R. J. Linhardt and D. C. Rees *Science* 1996, **271**, 1116-1120; (b) B. Casu and U. Lindahl *Adv. Carbohydro. Chem. Biochem.* 2001, **57**, 159-208; (c) L. Pettegrini, D. F. Burke, F. V. Delft, B. Mulloy and T. L. Blundell *Nature* 2000, **407**, 1029-1034

[5] (a) M. Guerrini, T. Agulles, A. Bisio, M. Hricovini, L. Lay, A. Naggi, L. Poletti, L. Sturiale, G. Torri and B. Casu *BBRC* 2002, **292**, 222-230; (b) A. D. DiGabriele, I. Lax, D. I. Chen, C. M. Svahn, M. Jaye, J. Schlessinger and W. A. Hendrickson *Nature* 1998, **395**, 812-817; (c) A. Canales, R. Lozano, B. Lopez-Mendez, J. Angulo, R. Ojeda, P. M. Nieto, M. Martin-Lomas, G. Gimenez-Gallego and J. Jiménez-Barbero *FEBS J.* 2006, **273(20)**, 4716-4727

[6] (a) R. E. Hileman, J. R. Fromm, J. M. Weiler and R. J. Lihardt *BioEssays* 1998, **20**, 156-167; (b) M. A. Palladino, R. E. Morris, H. F. Starnes and A. D. Levinson *Ann. N. Y. Acad. Sci.* 1990, **593**, 181-187; (c) T. A. McCaffrey, D. J. Falcone and B. Du *J. Cell Physiol.* 1992, **152**, 430-440

[7] M. Lyon, J. A. Deakin, K. Mizuno, T. Nakamura and J. T. Gallagher *J. Biol. Chem.* 1994, **269**, 11216-11223

[8] S. Sue, J. Brisson, S. Chang, W. Huang, S. Lee, H. C. Jarrell, W. Wu *Biochemistry* 2001, **40**, 10436-10446

[9] R. D. Rosenberg and P. S. Damus *J. Biol. Chem.* 1973, **248**, 6490-6505

[10] J. R. S. Day, R. C. Landis and K. M. Taylor *J. Cardiothorac. Vasc. Anesth.* 2004, **18**, 93-100

[11] K. Sugahara, T. Mikami, T. Uyama, S. Mizuguchi, K. Nomura and H. Kitagawa *Curr. Opin. Struct. Biol.* 2003, **13**, 612-620

[12] K. E. Kuettner, R. Schleyerbach, J. G. Peyron and V. C. Hascall (Eds) *Articular cartilage and osteoarthritis New York, Raven Press,* 1992

[13] B. M. Cooke, S. J. Rogerson, G. V. Brown and R. L. Coppel *Blood* 1996, **88**, 4040-4044

[14] (a) M. Fried and P. E. Duffy *J. Mol. Med.* 1998, **76**, 162-171; (b) M. Fried, R. M. Lauder and P. E. Duffy *Exp. Parasitol* 2000, **95**, 75-78

[15] M. C. Bourin, A. Lundgren, E. Kerlund and U. Lindahl *J. Biol. Chem.* 1990, **265**, 15424-5431

[16] M. Lyon, J. A. Deakin, H. Rahmoune, D. G. Fernig, T. Nakamura and J. T. Gallagher *J. Biol. Chem.* 1998, **273**, 271-278

[17] M. M. Maimone and D. M. Tollefsen *J. Biol. Chem.* 1990, **265**, 18263-18271

[18] (a) A. R. Poole *Biochem. J.* 1986, 236, 1-14; (b) L. Å. Fransson *Trends Biochem. Sci.* 1987, **12**,406-411

[19] (a) E. J. Thonar, M. E. Lenz, G. K. Klintworth, B. Caterson, L. M. Pachman, P. Glickman, R. Katz, J. Huff and K. E. Kuettner *Arthritis Rheum.* 1985, **28**, 1367-1376; (b) M. B. Sweet, A. Coelho, C. M. Schnitzler, T. J. Schnitzer, M. E. Lenz, I. Jakim, K. E. Kuettner and E. J. Thonar *Arthritis Rheum.* 1988, **31**, 648-652

[20] N. Papageorgakopoulou, A. D. Theocharis, S. S. Skandalis, D. H. Vynios, D. A. Theocharis and C. P. Tsiganos *Biochimie* 2001, **83**, 973-978

[21] G. Tai, T. N. Huckerby and I. A. Nieduszynski *Carbohydr. Res.* 1994, **255**, 303-309

[22] M. Pena, C. Williams and E. Pfeiler *Carbohydr. Res.* 1998, **309**, 117-124

[23] J. Fraser, T. Laurent and U. Laurent *J. Intern. Med.* 1997, **242**, 27-33

[24] (a) M. Goebeler, D. Kaufmann, E. B. Brocker and C. E. Klein *J. Cell Sci.* 1996, **109**, 1957-1964; (b) J. Catterall and M. Gardner *Cancer J.* 1995, **8**, 320-324; (c) P. Rooney, S. Kumar, J. Ponting and M. Wang *Int. J. Cancer* 1995, **60**, 632-636; (d) L. Zhang, C. B. Underhill and L. Chen *Cancer Res.* 1995, **55**, 428-433

[25] P. Teriete, S. Banerji, M. Noble, C. D. Blundell, A. J. Wright, A. R. Pickford, E. Lowe, D. J. Mahoney, M. I. Tammi, J. D. Kahmann, I. D. Campbell, A. J. Day and D. G. Jackson *Mol. Cell* 2004, **13**, 483-496

[26] A. J. Day and G. D. Prestwich *J. Biol. Chem.* 2002, **277**, 4585

[27] B. Mulloy, P. A. S. Mourão and E. Gray *J. Biotechno.* 2000, **77**, 123-135

[28] D. R. Ferro, A. Provasolli, M. Ragazzi, B. Casu, G. Torri, V. Bossennec, B. Perty, P. Sinay, M. Petitou and J. Choay *Carbohydr. Res.* 1990, **195**, 157-167

[29] H. Verli and J. A. Guimarães *Carbohydr. Res.* 2004, **339**, 281-290

[30] D. R. Ferro, A. Provasolli, M. Ragazzi, G. Torri, B. Casu, G. Gatti, J. C. Jacquinet, P. Sinay, M. Petitou and J. Choay *J. Am. Chem. Soc.*1986, **108**, 6773-6778

[31] (a) D. Mikhailov, K. H. Mayo, I. R. Vlahov, T. Toida, A. Pervin and R. J. Linhardt *Biochem. J.* 1996, **318**, 93-102; (b) D. Bashford and D. A. Case *Ann. Rev. Phys. Chem.* 2000, **51**, 129-152; (c) A. Canales, J. Angulo, R. Ojeda, M. Bruix, R. Fayos, R. Lozano, G. Gimenez-Gallego, M. Martin-Lomas, P. M. Nieto and J. Jiménez-Barbero *J. Am. Chem. Soc.* 2005, **127**, 5778-5779

[32] S. E. B. Gould, R. O. Gould, D. A. Rees and A. W. Wight *J. Chem. Soc. Perkin Trans. II* 1975, **392**

[33] J. Jiménez-Barbero and T. Peters (Eds.) *NMR Spectroscopy of Glycoconjugates* 2003, P74

[34] (a) D. Mikhailov, K. H. Mayo, I. R. Vlahov, T. Toida, A. Pervin and R. J. Linhardt *Biochem. J.* 1996, **318**, 93-102; (b) J. Angulo, M. Hricovini, M. Gairi, M. Guerrini, J. L. de Paz, R. Ojeda, M. Martin-Lomas and P. M. Nieto *Glycobiology* 2005, **15(10)**, 1008-1015

[35] M. Lyon and J. T. Gallagher *Matrix Biol.* 1998, **17**, 485-493

[36] (a) H. Tsuda, S. Yamada, Y. Yamane, k. Yoshida, J. J. Hopwood and K. Sugahara *JBC* 1996, **271**, 10495-10502; (b) B. Mulloy, M. J. Forster, C. Jones, A. F. Drake, E. A. Johnson and D. B. Davies *Carbohydr. Res.* 1994, **255**, 1-26; (c) E. A. Yates, F. Santini, M. Guerrini, A. Naggi, G. Torri and B. Casu *Carbohydr. Res.*1996, **294**, 15-27

[37] H. G. Garg, L. Yu, C. A. hales, T. Toida, T. Islam and R. J. Linhardt *Biochim.Biophys. Acta* 2003, **1639**, 225-231

[38] I. Capila, M. J. Hernáiz, Y. D. Mo, T. R. Mealy, B. Campos, J. R. Dedman, R. J. Linhardt and B. A. Seaton *Structure* 2001, **9**, 57-64

[39] B. Mulloy, M. J. Forster, C. Jones and D. B. Davies *Biochem. J.* 1993, **293**, 849-857

[40] B. Mulloy, M. J. Forster, C. Jones, A. F. Drake, E. A. Johnson and D. B. Davies *Carbohydr. Res.* 1994, **255**, 1-26

[41] V. S. R. Rao, P. K. Qasba, P. V. Balaji and R. Chandrasekaran *Conformation of Carbohydrates (Harwood Academic Publishers)* 1998

[42] C. J. Cramer and D. G. Truhlar *J. Am. Chem. Soc.* 1993, **115**, 5745-5753

[43] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham III, S. Debolt, D. Ferguson, G. Seibel and P. Kollman *Comput. Phys. Commun.* 1995, **91**, 1-41

[44] R. J. Woods, R. A. Dwek and C. J. Edge *J. Phys. Chem.* 1995, **99**, 3832-3846

[45] R. R. Hudgins, J. Woenckhaus and M. F. Jarrold *Int. J. Mass Spectrom. Ion Process.* 1997, **165/166**, 497-507

[46] B. S. Kinnear, M. R. Hartings and M. F. Jarrold *J. Am. Chem. Soc.* 2001, **123**, 5660

[47] (a) G. Von Helden, N. G. Gotts, P. R. Kemper and M. T. Bowers *Chem. Phys. Lett.* 1993, **204**, 15; (b) G. Von Helden, M. T. Hsu, P. R. Kemper and M. T. Bowers *MRS Symp. Series* 1992, **270**, 117

[48] (a) S. Lee, T. Wyttenbach and M. T. Bowers *Int. J. Mass Spectrom. Ion Process.* 1997, **167/168**, 605-614; (b) D. Lee, C. Wu and H. H. Hill Jr. *J. Chromatogr. A* 1998, **822**, 1-9

[49] (a) http://bowers.chem.ucsb.edu/theory_analysis/index.shtml Bowers Group (UCSB); (b) T. Wyttenbach and M. T. Bowers *Top. Curr. Chem.* 2003, **225**, 207-232;

[50] (a) *Understanding NMR Spectroscopy* James Keeler, 2005, Wiley (b) *NMR Spectroscopy: Basic principles, concepts, and applications in chemistry* (Second Edition) Harald Günther 1996, John Wiley & Sons

[51] L. E. Kay, P. Keifer and T. Saarinen *J. Am. Chem. Soc.* 1992, **114**, 10663-10665

[52] (a) J. Weigelt and G. Otting *J. Magn. Reson. A* 1995, **113**, 128-130; (b) N. C. Nielsen, H. Thogersen and O. W. Sørensen *J. Am. Chem. Soc.* 1995, **117**, 11365-11366

[53] A. J. Shaka, C. J. Lee and A. Pines *J. Magn. Reson.* 1988, **77**, 274-293

[54] (a) N. Tjandra and A. Bax *Science* 1997, **278**, 1111-1114; (b) S. Cavagnero, H. J. Dyson and P. E. Wright *J. Biomol. NMR* 1999, **13**, 387-391; (c) M. Ottiger and A. Bax *J. Biomol. NMR* 1998, **12**, 361-372

[55] (a) M. R. Hansen, L. Mueller and A. Pardi *Nat. Struct. Biol.* 1998, **5**, 1065-1074; (b) G. M. Clore, M. R. Starich and A. M. Gronenborn *J. Am. Chem. Soc.* 1998, **120**, 10571-10572

[56] (a) R. S. Prosser, J. A. Losonczi and I. V. Shiyanovskaya *J. Am. Chem. Soc.* 1998, **120**, 11010-11011; (b) L. G. Barrientos, C. Dolan and A. M. Gronenborn *J. Biomol. NMR* 2000, **16**, 329-337

[57] (a) B. W. Koenig, J. Hu, M. Ottiger, S. Bose, R. W. Hendler and A. Bax *J. Am. Chem. Soc.* 1999, **121**, 1385-1386; (b) J. Sass, F. Cordier, A. Hoffmann, M. Rogowski, A. Cousin, J. G. Omichinske, H. Lőwen and S. Grzesiek *J. Am. Chem. Soc.* 1999, **121**, 2047-2055

[58] (a) M. Jonströmer and R. Strey *J. Phys. Chem* 1992,**96**, 5993-6000; (b) M. H. G. M. Penders and R. Strey *J. Phys. Chem.* 1995, **99**, 6091-6095; (c) E. Freyssingeas, F. Nallet and D. Roux Langmuir 1996, **12**, 6028-6035

[59] M. Rückert and G. Otting *J. Am. Chem. Soc.* 2000, **122**, 7793-7797

[60] M. Ottiger and A. Bax *J. Biomol. NMR* 1998, **12**, 361-372

[61] C. R. II Sanders, J. E. Schaff and J. H. Prestegard *Biophys. J.* 1993, **64**, 1069-1080

[62] A. Saupe *Angew. Chem., int. Ed. Engl.* 1968, **7**, 97

[63] J. A. Losonczi, M. Andrec, M. W. F. Fischer and J. H. Prestegard *J. Mag. Reson.* 1999, **138**, 334-342

[64] F. Chevalier, R. Lucas, J. Angulo, M. Martin-Lomas and P. M. Nieto *Carbohydr. Res.* 2003, **339**, 975-983

[65] E. Gemma, A. N. Hulme, A. Jahnke, L. Jin, M. Lyon, R. M. Müller and D. Uhrín *Chem. Comm.* In press

[66] H. Valafar and J. H. Prestegard *J. Magn. Reson.* 2004, **167**, 228-241

[67] M.J. Thrippleton and J. Keeler *Angewandte Chemie* 2003, **115**, 4068-4071

[68] R.A.E. Edden and J. Keeler *J. Magn. Reson.* 2004, **166**, 53-68

[69] M. J. Thrippleton and J. Keeler *Angew. Chem. Int. Ed. Engl.* 2003, **42**, 3938-3941

[70] The Chemistry, Biology and Medical Applications of Hyalüronan and its Derivatives (Laurent TC, ed) *Wenner-Gren International Series*, 1998, 72, Portland Press, London

[71] M. Köck, R. Kerssebaum and W. Bermel *Magn. Reson. Chem.* 2003, **41**, 65-69

[72] K. Stott, J. Stonehouse, J. Keeler, T.L. Hwang and A.J. Shaka *J. Am. Chem. Soc.* 1995, **117**, 4199-4200

[73] P. T. Robinson, T. N. Pham and D. Uhrín *J. Magn. Reson.* 2004, **170**, 97-103

[74] AMBER 7, D.A. Case, D.A. Pearlman, J.W. Caldwell, T.E. Cheatham III, J. Wang, W.S. Ross, C.L. Simmerling, T.A. Darden, K.M. Merz, R.V. Stanton, A.L. Cheng, J.J. Vincent, M. Crow-ley, V. Tsui, H. Gohlke, R.J. Radmer, Y. Duan, J. Pitera, I. Massova, G.L. Seibel, U.C. Singh, P.K. Weiner and P.A. Kollman (2002), University of California, San Francisco.

[75] Dr. Woods website, http://glycam.ccrc.uga.edu/glycam/glycam_parameters.html

[76] C. J. M. Huige and C. Altona *J. Comp. Chem.* 1995, **16**, 56-79

[77] L. Jin P. E. Barran, J. A. Deakin, M. Lyon and D. Uhrín *Phys. Chem. Chem. Phys.*, 2005, **7**, 3464-3471

[78] T. Wyttenback and M. T. Bowers, *Top. Curr. Chem.* 2003, **255**, 207-232

[79] http://nano.chem.indiana.edu/software.html Dr. Jarrold website

[80] R. Koradi, M. Billeter and K. Wüthrich *J. Mol. Graphics* 1996, **14**, 51-55

[81] W. Humphrey, A. Dalke and K. Schulten *J. Mol. Graphics* 1996, **14**, 33-38. VMD -Visual Molecular Dynamics, http://www.ks.uiuc.edu/Research/vmd/

[82] B. Mulloy, M. J. Forster and C. Jones *Biochem. J.*, 1993, **293**, 849

[83] S. Lee, T. Wyttenbach and M. T. Bowers *Int. J. Mass Spectrum*, 1997, **167**, 605-614

[84] B. H. Clowers, P. Dwivedi, W. E. Steiner and H. H. Hill, Jr *J. Am. Soc. Mass Spectrom*, 2005, **16**, 660-669

[85] D. Mikhailov, K. H. Mayo, I. R. Vlahov, T. Toida, A. Pervin and R. J. Linhardt *Biochem. J.* 1996, **318**, 93-102.

[86] S. Faham, R. E. Hileman, J. R. Fromm, R. J. Linhardt and D. C. Rees *Science*, 1996, **271**, 1116-1120

[87] D. L. Rabenstein, J. M. Robert and J. Peng *Carbohyd Res.*, 1995, **278**, 239-256

[88] J. Zaia and C.E. Costello *Anal. Chem.*, 2001, **73**, 233-239

[89] (a) T. N. Pham, S. L. Hinchley, D. W. H. Rankin, T. Liptaj and D. Uhrín *J. Am. Chem. Soc.* 2004, **126**, 13100-13110; (b)T. N. Pham, T. Liptaj, P. N. Barlow and D. Uhrín *Magn. Res. Chem.* 2002, **40**, 729-732; (c) J. R. Brisson, S.-C. Sue, W.-G. Wu, G. McManus, T. N. Pham and D. Uhrín *NMR and Glyco-Conjugates* (Eds.: J. Jimenez-Barbero and T. Peters), WILEY-VCH Verlag, Weinheim, 2003, P59-94

[90] G. Brodenhausen, R. Freeman and D. L. Turner *J. Magn. Reson.* 1977, **27**, 511

[91] R. Konrat, I. Burghardt and G. Bodenhausen *J. Am. Chem. Soc.* 1991, **113**, 9135-9140

[92] M. J. Thrippleton and J. Keeler *Angew. Chem. Int. Ed. Engl.* 2003, **42**, 3938-3941

[93] W. Willker, D. Leibfritz, R. Kerssebaum and J. Lohman *J. Magn. Reson. A,* 1993, **102**, 348-350

[94] G. M. Brown and H. A. Levy *Acta. Cryst.* 1979, **B35**, 656-659

[95] A. Mangoni, V. Esposito and A. Randazzo *Chem. Commun.* 2003, 154-155

[96] M. C. Wahle and S. R. Bryn *Acta. Cryst.* 1996, **C52**, 2495-2499

[97] L. Jin, T. N. Pham, and D. Uhrín *ChemPhysChem* In press

[98] A. Bax, R. Freeman and S. P. Kempsell *J. Am. Chem. Soc.* 1980, **102**, 4849-4851

[99] B. Bose, S. Zhao, R. Stenutz, F. Cloran, P. B. Bonodo, G. Bondo, B. Hertz, I. Carmichael and A. S. Serianni *J. Am. Chem. Soc.* 1998, **120**, 11158-11173

[100] T. N. Pham, K. E. Kövér, L. Jin and D. Uhrín *J. Magn. Reson.* 2005, **176**, 199-206

[101] J. Weigelt and G. Otting *J. Magn. Reson. A* 1995, **113**, 128-130

[102] M. Kővér and P. Forgo *J. Magn. Reson.* 2004, **166**, 47-52

[103] W. Kozminski, D. Sperandio and D. Nanz *Magn. Reson. Chem.* 1996, **34**, 311-315

[104] J. Stonehouse and J. Keeler *J. Magn. Reson. A.* 1995, **112**, 43-47

[105] A. Neszmelyi and G. Lukacs *J. Am. Chem. Soc.* 1982, **104**, 5342-5346

[106] A. Meissner, D. Moskau, N. C. Nielsen and O. W. Sørensen *J. Magn. Reson.* 1997, **124**, 245-249

[107] F. Delaglio, Z. R. Wu and A. Bax *J. Magn. Reson.* 2001, **149**, 276-281

[108] W. Kozminske and D. Nanz *J. Magn. Reson. (Series A)* 1996, **122**, 245-247

[109] K. E. Kover and P. Forgo *J. Magn. Reson.* 2004, **166**, 47-52

[110] H. Kessler, A. Muller and H. Oschkinat *Magn. Reson. Chem.* 1985, **23**, 844-852

[111] M. Bourdonneau and B. Ancian *J. Magn. Reson.* 1998, **132**, 316-327

[112] (a) T. E. Skinner, T. O. Reiss, B. Luy, N. Khaneja and S. J. Glaser *J. Magn. Reson.* 2004, **167**, 68-74; (b) B. Luy, K. Kobzar, T. E. Skinner, N. Khaneja and S. J. Glaser *J. Magn. Reson.* 2005, **176**, 179-186

[113] M. Goldman Clarendon Press 1970 DNP-polarization

[114] C. A. G. Haasnoot, F. A. A. M. de Leeuw and C. Altona *Tetrahedron* 1980, **36**, 2783-2792

[115] D. R. Ferro, A. Provasoli and M. Ragazzi *Carbohydr. Res.* 1990, **195**, 157-167

[116] I. Tuarosha, M. Hricouini and E. Petrarova *Carbohydr. Res.* 1989, **189**, 359-362

[117] B. Mulloy, M. J. Forster, C. Jones and D. B. Davies *Biochem. J.* 1993, **293**, 849-858

[118] K. Whaley and S. Ruddy *Science* 1976, **193**, 1011-1013

[119] (a) T. K. Blackmore, J. Hellwage, T. A. Sadlon, N. Higgs, P. F. Zipfel, H. M. Ward and D. L. Gordon *J. Immunol.* 1998, **160**, 3342-3348; (b) T. K. Blackmore, T. A. Sadlon, H. M. Ward, D. M. Lublin and D. L. Gordon *J. Immunol.* 1996, **157**, 5422-5427; (c) M. K. Pangburn, M. A. Atkinson and S. Meri *J. Biol. Chem.* 1991, **266**, 16847-16853

[120] G. M. Morris, D. S. Goodsell, R. Huey and A. J. Olson *J. Comput. Aided Mol. Des.* 1996, **10**, 293-304

[121] A. P. Herbert, D. Uhrín, M. Lyon, M. K. Pangburn and P. N. Barlow *J. Biol. Chem.* 2006, **281**, 16512-16520

[122] T. S. Jokiranta, J. Hellwage, D. A. Male, E. Giannakis, P. F. Zipfel, S. Meri and D. L. Gordon *Immunopharmacology* 2000, **49**, 54

[123] Bok, D. Evidence for an inflammatory process in age-related macular degeneration gains new support. *Proc Natl Acad Sci U S A* 2005, **102**, 7053-4 Epub

[124] Hageman, G. S. et al. A common haplotype in the complement regulatory gene factor H (HF1/CFH) predisposes individuals to age-related macular degeneration. *Proc. Natl. Acad. Sci. U S A* 2005, **102**, 7227-32. Epub

[125] A. P. Herbert, J. A. Deakin, C. Q. Schmidt, B. S. Blaum, C. Egan, V. Ferreira, M. K. Pangburn, M. Lyon, D. Uhrín and P. N. Barlow; submitted to *Biol. Sci. Immun.*

PhD Records

Courses attended: Annual fire safety lectures, Fire extinguisher course, Assessment and control of risk (chemical lab) course, Introduction course for laboratory demonstrator;
PG courses: Advanced experimental and computational methods, Mass spectrometry, Carbohydrates, Topspin NMR course, PG course in Nuclear Magnetic Resonance (2005), Resonance Techniques in NMR and EPR;
Optional courses: Effective writing, Poster production, Preparing a presentation using Powerpoint, Internet literature research for science, First year report writing, PhD thesis workshop (general session), EaSTChem thesis workshop; Graphing packages, Unix 2, Time management, Interviewing skills, Alternative career for scientific students.

School seminars and presentations: Biophysical chemistry section seminar and annual oral presentations, BioNMR group weekly meeting and oral presentations, Monthly NMR technical meeting and presentations, Poster presentation on PG open days, Annual Firbush presentations;
Colloquia: Getting into shape-biological molecules in the gas phase (Prof. John Simons); Mechanisms of cooperativity and desolvation in molecular recognition: implications for rational drug design (Prof. Paul H. Axelsen); Templated molecular recognition materials: theory and computer simulation (Dr. Lev Sarkisov); Diversity oriented organic synthesis and the small molecule approach to chemical biology (Prof. Steve Caddick); Statistical mechanics of liquids: from disorder to complexity (Prof. Jean-Pierre Hansen); Surveying protein folding energy landscapes: from dynamics to disease (Prof. Sheena Radford); Recent advances in the total synthesis of antibiotic natural products (Prof. Anthony Barrett); Single molecular magnet: a molecule approach to nanoscale magnetic materials (Prof. George Chrisou); The chemistry of oxygen sensing in human (Chris Schofield); Force generation by self-assembling protein polymers (Dr. Marileen Dogteron).

Demonstrating: Biological chemistry (including foundation course) laboratory supervision and reports marking (three years).

Conferences attended and work presentations: British Society of Matrix Biology (BSMB) spring anniversary meeting 2004, 15 minutes oral presentation, 1.5 days poster exhibition; Collaborative Computing Project for NMR (CCPN) Conferences (4th to 6th, 2004-2006); Biophysical Chemistry 2004: Ligand Binding and Drug Discovery Sep. 6-8th; EUROMAR/EENC 2005, Veldhoven-Netherlands Jul. 3-8th, 2005, poster exhibition; Symposium: the Edinburgh Protein Interaction Centre (EPIC): New Insight into Protein Interactions 2005; 2006 AMBER Workshop, London.

Publications:

1. '^{1}H-Detected Double-*J*-Modulated INEPT-INADEQUATE for simultaneous determination of one-bond and long-range carbon-carbon connectivities and the measurement of all carbon-carbon coupling constants', T. N. Pham, K. E. Kövér, L. Jin and D. Uhrín, *J. Magn. Reson.* **2005**, 176, 199-206
2. 'Conformation of Glycosaminoglycans by Ion Mobility Mass Spectrometry and Molecular Modelling', L. Jin, P.E. Barran, J. A. Deakin, M. Lyon and D. Uhrín, *Phys. Chem. Chem. Phys.* **2005**, 7, 3464-3471; Hot-paper and selected for the Chemical biology Virtual Journal of Molecular Biosystems
3. 'DMT-MM Mediated Functionalisation of the Non-Reducing End of Glycosaminoglycans', E. Gemma, A. N. Hulme, A. Jahnke, L. Jin, M. Lyon, R. M. Müller and D. Uhrín, *Chem. Comm.,* In press
4. 'Measurement of ^{1}H-^{1}H Residual Dipolar Coupling Constants for Structural Studies of Medium Size Molecules', L. Jin, T. N. Pham and D. Uhrín, *ChemPhysChem,* In press
5. 'The ^{13}C-detected INADEQUATE for accurate measurement of scalar and dipolar carbon-carbon coupling constants The INADEQUATE experiment has come of age' L. Jin and D. Uhrín submitted to *Magn. Reson. Chem.*