



## Refined structure of a flexible heptasaccharide using $^1\text{H}$ - $^{13}\text{C}$ and $^1\text{H}$ - $^1\text{H}$ NMR residual dipolar couplings in concert with NOE and long range scalar coupling constants

Manuel Martin-Pastor & C. Allen Bush\*

Department of Chemistry and Biochemistry, University of Maryland-Baltimore County, Baltimore, MD 21250, U.S.A.

Received 17 August 2000; Accepted 7 November 2000

**Key words:** carbohydrate structure, flexibility, NOE, residual dipolar couplings, scalar couplings, simulated annealing

### Abstract

The heptasaccharide isolated from the cell wall polysaccharide of *Streptococcus mitis* J22 serves as an important model for the dynamics and conformation of complex polysaccharides, illustrating the nature of flexibility with rigid epitopes joined by flexible hinges. One-bond C-H residual dipolar couplings ( $^1D_{\text{CH}}$ ) and long-range H-H residual dipolar couplings ( $^nD_{\text{HH}}$ ) were measured for the heptasaccharide in a cetylpyridinium chloride/hexanol/brine lamellar liquid crystal medium. A method is proposed to determine the  $^nD_{\text{HH}}$  in natural abundance based on a  $^{13}\text{C}$  resolved  $^1\text{H}$  TOCSY pulse sequence previously published to determine the homonuclear scalar couplings. Different methods for interpretation of the  $^1D_{\text{CH}}$  and the  $^nD_{\text{HH}}$  residual dipolar coupling data obtained were compared and combined with the NOE and long-range H,C and C,C scalar couplings available for this heptasaccharide. A flexible model of the heptasaccharide was determined in which two structurally well-defined regions involving four and two sugar residues, respectively are joined by a flexible hinge which involves two 1 $\rightarrow$ 6 glycosidic linkages.

**Abbreviations:** R-SA, restrained simulated annealing.

### Introduction

Most NMR studies of oligosaccharide conformation have used nuclear Overhauser effects (NOE) to establish proximity between protons, a method that has successfully guided molecular modeling of tightly folded and relatively rigid oligosaccharide epitopes (Bush et al., 1999). Complications in the interpretation of NOE data for more flexible oligosaccharides have stimulated a search for other experimental methods for obtaining data useful for the conformation of complex oligosaccharides, such as  $^{13}\text{C}$  scalar coupling (Martin-Pastor et al., 1999) and residual dipolar coupling in oligosaccharides partially oriented in liquid crystal so-

lutions (Ottiger et al., 1999). One-bond C-H residual dipolar couplings,  $^1D_{\text{CH}}$ , have recently been used in some studies of oligosaccharides either in the free state (Kiddle et al., 1998; Landersjo et al., 2000; Martin-Pastor et al., 2000a,b) or in the bound state (Bolon et al., 1999; Shimizu et al., 1999; Thompson et al., 2000), providing further refinement in the structures. Another residual dipolar coupling of interest in carbohydrates is the long-range  $^1\text{H}$ - $^1\text{H}$  residual dipolar coupling ( $^nD_{\text{HH}}$ ). The intra-residue  $^nD_{\text{HH}}$  can provide valuable information concerning the ring sugar puckering and orientation, while the inter-residue  $^nD_{\text{HH}}$  may provide information about the inter-glycosidic torsion angles in carbohydrates. Recently, a CT-COSY experiment has been used to determine  $^nD_{\text{HH}}$  in a trisaccharide (Tian et al., 1999). In this homonuclear method  $^nD_{\text{HH}}$  are determined quantitatively from

\*To whom correspondence should be addressed. E-mail: bush@umbc.edu

several experiments making use of the trigonometric dependence of the intensities of the signals with the constant time period used. Serious problems of signal overlap in  $^1\text{H}$  NMR spectra of oligosaccharides suggest that  $^{13}\text{C}$  chemical shift resolution could be valuable for  $^1\text{H}$ - $^1\text{H}$  dipolar coupling measurements in complex oligosaccharides. Some heteronuclear methods have been developed in the past for the measurement of  $^n\text{J}_{\text{HH}}$  scalar couplings. Since in many instances the  $^n\text{J}_{\text{HH}}$  and  $^n\text{D}_{\text{HH}}$  interactions evolve in an analogous way during the pulse sequence, some of these  $^n\text{J}_{\text{HH}}$  heteronuclear methods could be adapted for measurement of  $^n\text{D}_{\text{HH}}$ . In this paper we have tested a  $^1\text{H}$ ,  $^{13}\text{C}$   $^1\text{H}$ - $^1\text{H}$  TOCSY sequence previously published for the measurement of  $^n\text{J}_{\text{HH}}$  scalar couplings (Willker et al., 1992) in natural abundance samples. This experiment provides in-phase E.COSY type signals resolved in the  $^1\text{H}$  and  $^{13}\text{C}$  dimensions, from which the splittings in the proton dimension can be used to determine the sign and magnitude of the  $^n\text{D}_{\text{HH}}$  by its comparison between an oriented and a non-oriented sample.

The usual method of interpretation of the residual dipolar couplings requires an orientation tensor defining the direction and magnitude of alignment of the molecule in the liquid crystal medium (Tjandra et al., 1997). The alignment tensor can either be predicted based on the global shape of the molecule and its steric and electrostatic interactions with the liquid crystal (Zweckstetter et al., 2000), or it can be calculated for the complete structure or a substructure using a best fit approach within the order matrix calculation method (Fischer et al., 1999; Losonczi et al., 1999) or with a Powell minimization algorithm (Tjandra et al., 1996). If the structure is unknown and the number of residual dipolar couplings is large enough, the magnitude of the alignment can be estimated from the powder pattern of residual dipolar couplings (Clare et al., 1998a). An r-SA protocol has been proposed for the refinement of the structures using residual dipolar couplings in combination with other NMR structural restraints (Clare et al., 1998b). Briefly, in this r-SA protocol the residual dipolar couplings are treated by means of a floating tetra-atomic molecule, which is used to define as the axis of the alignment tensor in the molecular frame of the molecule of interest. During the initial stages of the r-SA refinement, the orientation of the tetra-atomic molecule is optimized for agreement with the observed experimental dipolar couplings of the molecule under study. Later in the r-SA simulation, when the NMR restraints have defined the main fea-

tures of the structure, the weight of the residual dipolar coupling restraints is gradually increased towards the desired value (Clare et al., 1998b). This r-SA protocol, initially implemented for the use of dipolar couplings with fixed distance, such as  $^1\text{D}_{\text{CH}}$  or  $^1\text{D}_{\text{NH}}$ , has been recently adapted to incorporate the residual dipolar couplings of variable distance such as  $^n\text{D}_{\text{HH}}$  (Tjandra et al., 2000).

The flexibility inherent to certain biomolecules in solution introduces an extra difficulty for the interpretation of the NMR structural data as the data reflect an average of the different conformations present. High amplitude flexibility in certain regions of the molecule may result in internal inconsistencies in the NMR structural data, and resolution of these inconsistencies may be required for the characterization of all the conformers involved and their respective population (Gorler et al., 2000). For the case of carbohydrates the structural inconsistencies may appear in some interglycosidic NOEs and/or in the long-range interglycosidic  $^3\text{J}_{\text{COCH}}$ ,  $^2\text{J}_{\text{COC}}$  or  $^3\text{J}_{\text{COC}}$  scalar couplings which reflect the flexibility among certain sugar linkages (Martin-Pastor et al., 1999). The interpretation of the residual dipolar couplings in flexible structures requires a determination of those structurally well defined regions or substructures in the molecule that could be considered as dynamically equivalent (Fischer et al., 1999), so that an independent orientation tensor can be defined for each of them. Ultimately, the features of the flexible hinges joining these more rigid substructures must also be determined.

## Experimental

### *Sample preparation*

A CPCI/hexanol stock solution was prepared by mixing equal amounts of cetylpyridinium chloride (CPCI) and hexanol in  $\text{D}_2\text{O}$  ~5% w/w. The resulting solution was submitted to several cycles of vortexing and equilibration, first at 70 °C and then at 30 °C for 1 h.

Fourty mg of the J22 heptasaccharide repeating unit (Figure 1) was prepared by mild acid hydrolysis of the J22 polysaccharide (provided by J. Cisar) as previously described (Cisar et al., 1995) and was subsequently exchanged three times in  $\text{D}_2\text{O}$ . A sample of J22 heptasaccharide (~37 mM) was prepared by dissolving 20 mg of the heptasaccharide in  $\text{D}_2\text{O}$ . A CPCI/hexanol 5% liquid crystal sample of J22 heptasaccharide ~37 mM was prepared by dissolving 20 mg of the heptasaccharide in the corresponding



nal was carefully phased and stored. After an inverse Hilbert transformation, extensive zero filling to 16K was used to obtain a digital resolution of 0.2 Hz/pt. This way of increasing the digital resolution does not improve the resolution of the data and was chosen simply to facilitate the subsequent comparison of the vectors. Vectors from the liquid crystal sample were compared with the corresponding vector from the D<sub>2</sub>O spectrum and the offset required for superposition of the multiplet components was used to calculate the sign and magnitude of <sup>1</sup>D<sub>CH</sub>. The experiments were repeated twice and <sup>1</sup>D<sub>CH</sub> couplings averaged; the error from the linewidths obtained was estimated to be ±1.0 Hz.

Residual long-range H-H dipolar coupling values, <sup>n</sup>D<sub>HH</sub>, were measured at 25 °C for the J22 heptasaccharide by measuring the differences in the splittings in the F<sub>2</sub> dimension obtained in a <sup>1</sup>H-<sup>13</sup>C, <sup>1</sup>H-<sup>1</sup>H TOCSY experiment (Wilker et al., 1992) between the heptasaccharide partially oriented in the CPCL/hexanol 5% or 3% liquid crystal sample and a sample of the heptasaccharide in D<sub>2</sub>O isotropic solution. All the experiments were acquired with the proton carrier placed at the center of the proton spectrum at 3.5 ppm with a spectral width of 2500 Hz. The carbon carrier was placed at 70 ppm and the spectral width used was 12 500 Hz. The INEPT delays corresponding to a nominal value of <sup>1</sup>J<sub>CH</sub> were set to 155 Hz. For the H-H TOCSY transfer, an MLEV-17 spin locking sequence was used in which the seventeenth pulse was set to 60°. A TOCSY mixing time of 40 ms with two trim pulses of 1 ms was used and the spin locking field strength was 5102 Hz. Carbon decoupling during acquisition was carried out with WALTZ-16 with a field strength of 1612.9 Hz. For each experiment a 2D matrix of 1024×256 complex FID data points was acquired for the t<sub>1</sub> and t<sub>2</sub> dimensions. The 2D FIDs were apodized in both dimensions with a 90° shifted sine-bell function and zero filled to give, after Fourier transformation, a 2D spectrum of 2048×512 real points. Several C-H correlations through one or several bonds were observed in the spectra, which were assigned by comparison with the <sup>1</sup>H and <sup>13</sup>C chemical shifts previously reported (Abeygunawardana et al., 1990; Martin-Pastor et al., 2000c). The measurement of the splitting was accomplished by addition of all the rows of each E. COSY doublet component of the signal. The resulting 1D vector was then carefully phased and stored. After an inverse Hilbert transformation and zero filling to 16K, the sign and magnitude of the H-H couplings were

estimated from the offset required for superposition of the two E. COSY components. <sup>n</sup>D<sub>HH</sub> residual dipolar couplings were estimated from the difference between the <sup>1</sup>H-<sup>1</sup>H couplings measured with the J22 heptasaccharide in the oriented CPCL/hexanol 3% or 5% liquid crystals and those measured in D<sub>2</sub>O; the estimated error is ±1.0 Hz.

#### *Molecular calculations*

Molecular modeling calculations were used to determine the three-dimensional structure of the J22 heptasaccharide consistent with the measured residual dipolar couplings and also with the NMR structural data previously reported (Martin-Pastor et al., 2000c). Individual sugar residues of the J22 heptasaccharide are indicated by the letters given in Figure 1 and glycosidic dihedral angles are defined with respect to the glycosidic hydrogen φ<sub>H</sub>: H1-C1-O1-C<sub>x</sub> and ψ<sub>H</sub>: C1-O1-C<sub>x</sub>-H<sub>x</sub>, except for the 1→6 linkages where ψ is defined as ψ<sub>O</sub>: C1-O1-C6<sub>x</sub>-C5<sub>x</sub>, where x is an atom in the contiguous residue. The hydroxymethyl dihedral angles are defined as ω<sub>d</sub>: O1c-C6-C5-C4 and ω<sub>e</sub>: O1d-C6-C5-O5. The additional torsion angle of the Galf residue d is defined as γ: O4-C4-C5-C6.

#### *Conformational grid search*

Energy-relaxed φ/ψ conformational grid searches were calculated independently for the glycosidic torsions of those residues of the J22 heptasaccharide which are not linked by 1→6 glycosidic linkages. Two different grid searches were performed, one for the antigenic tetrasaccharide component a-b-(g)-c and another for the lectin-binding disaccharide e-f, using InsightII/Discover software (Biosym Technologies, San Diego, CA, U.S.A.). For each grid search all the glycosidic dihedral angles were restrained by a cosine type potential with a force constant of 100 kcal mol<sup>-1</sup> and the remainder of the sub-structure was minimized using the CVFF (Hagler et al., 1979) force field with a distance dependent dielectric of 80\*/r. The grid step used for scanning the glycosidic torsions φ and ψ was 2.5°. In order to reduce the computational time of the a-b-(g)-c multidimensional grid search, any available interresidual NOE data were used to restrict the grid search. For this purpose upper distance limits of 3.5 Å and 4.0 Å for strong and medium NOE, respectively, were used as a criterion to select or reject a conformation prior to minimization. Similarly, the inter-glycosidic long-range scalar couplings <sup>3</sup>J<sub>CH</sub> and <sup>3</sup>J<sub>CC</sub> (Martin-Pastor et al., 2000c) were also used to reject conformations prior to the minimization by means

of their relationship with the glycosidic dihedral angle employing a Karplus-type correlation curve (Tvaroska et al., 1989; Bose et al., 1998). Conformations with calculated  $^3J_{CH}$  and  $^3J_{CC}$  values over an error limit of  $\pm 1$  Hz were discarded prior to minimization. All the accepted conformers were then minimized and those conformers within an energy limit of  $3 \text{ kcal mol}^{-1}$  over the global minimum were finally considered and tested for agreement with the  $^1D_{CH}$  dipolar couplings. A best fit approach of the  $^1D_{CH}$  data was used to calculate the alignment tensor by using a Powell optimization algorithm (Tjandra et al., 1996) modified as previously used in our laboratory (Martin-Pastor et al., 2000a,b). Only the part of the  $^1D_{CH}$  data corresponding to the residues a-b-(g)-c or e-f was used for these calculations. A  $\chi^2$  error function defined as in Martin-Pastor et al. (2000a) was used to measure the goodness of the  $^1D_{CH}$  fitting. Those conformations with a  $\chi^2$  fitting error over 1.0 were discarded. The conformers finally accepted after the grid search protocol were represented in  $\phi/\psi$  maps using Origin software (Microcal Software Inc).

#### *Restrained simulated annealing simulations*

r-SA simulations, including both  $^1D_{CH}$  and  $^nD_{HH}$  restraints, were performed independently for the tetrasaccharide formed by residues a-b-(g)-c and the disaccharide e-f of the J22 heptasaccharide using a version of the X-plor program (Brünger, 1992) modified to incorporate these restraints (Tjandra et al., 2000). The  $^1D_{CH}$  data sets obtained at 5% and 3% liquid crystal concentrations were introduced as long order harmonic restraints in the r-SA calculation using a floating tetra-atomic axis molecule to describe the alignment tensor (Clare et al., 1998b; Tjandra et al., 2000). As the alignment tensor was gradually optimized during the r-SA, the force constant used for the residual dipolar restraints was gradually increased from a starting value of 0.01 up to  $2 \text{ kcal Hz}^{-2}$ . The magnitude  $A_a$  of the alignment tensor was taken from the structure with the lowest  $\chi^2$  obtained during the previous grid search for the tetrasaccharide a-b-(g)-c. The  $A_a$  values used were 10.94 and 5.37 for the data at 5% and 3% liquid crystal concentration, respectively. An identical value for the rhombicity was used to treat data sets at different liquid crystal concentration. Several estimates of the rhombicity values, 0.2, 0.3, 0.4, 0.5 and 0.6, were tested in different r-SA simulations and the best results were obtained for values of 0.3 and 0.4. In all the r-SA calculations NOE and long-range scalar coupling restraints were also ap-

plied. A flat-bottomed harmonic potential was used for the interglycosidic  $^1H$ - $^1H$  NOE distance restraints of the a-b-(g)-c or e-f sub-structures. No NOE energy penalty was applied for NOE distances in the range 2.5 to 3.0 Å or 2.8 to 3.5 Å for strong and medium NOEs, respectively. Beyond the distance limits a harmonic potential is applied with a force constant of  $150 \text{ kcal Å}^{-2}$ . Scalar coupling  $^3J_{CH}$  and  $^3J_{CC}$  harmonic restraints were applied for the sub-structures a-b-(g)-c or e-f using the corresponding Karplus-type correlation curves (Tvaroska et al., 1989; Bose et al., 1998) with a force constant of  $1 \text{ kcal Hz}^{-2}$ . The temperature of the first high-temperature r-SA stage was set at 8000 K during 100 ps. This temperature is somewhat lower than commonly used for r-SA calculations in peptides. This period was followed by a slow cooling annealing stage to 300 K during 60 ps. Each structure was then minimized by 2000 conjugate gradient steps of minimization. These parameters, especially the lower temperature for the first stage, were necessary to avoid a high proportion of distorted pyranoside chair conformations in these carbohydrates.

A final r-SA calculation was performed for the complete J22 heptasaccharide using the X-plor program. A starting model of the heptasaccharide was built using the previously r-SA refined structures of a-b-(g)-c and e-f. NOE,  $^3J_{CH}$  and  $^3J_{CC}$  restraints previously used for the individual r-SA simulations of residues a-b-(g)-c and e-f were similarly included for this refinement. The  $^1D_{CH}$  and  $^nD_{HH}$  residual dipolar couplings of residues a-b-(g)-c and e-f were included as restraints in this calculation using a different tetra-atomic axis molecule to define the alignment tensor in each sub-structure. The r-SA proceeded until a total of 3000 trial structures had been calculated and provided 89 low-energy conformers, whose similarity was analyzed by cluster analysis. A variable RMS cut-off window was used in the cluster analysis, which was performed with the program Nmcluster (Oxford Molecular Ltd) (Kelley et al., 1996). A total of nine clusters were found, from which the structure of the most representative conformation was selected and further analyzed.

## **Results and discussion**

Residual dipolar coupling  $^1D_{CH}$  and  $^nD_{HH}$  values were measured for a natural abundance sample of J22 heptasaccharide (Figure 1) in suitable mixtures of CPCI/hexanol/brine liquid crystals.  $^1D_{CH}$  were mea-





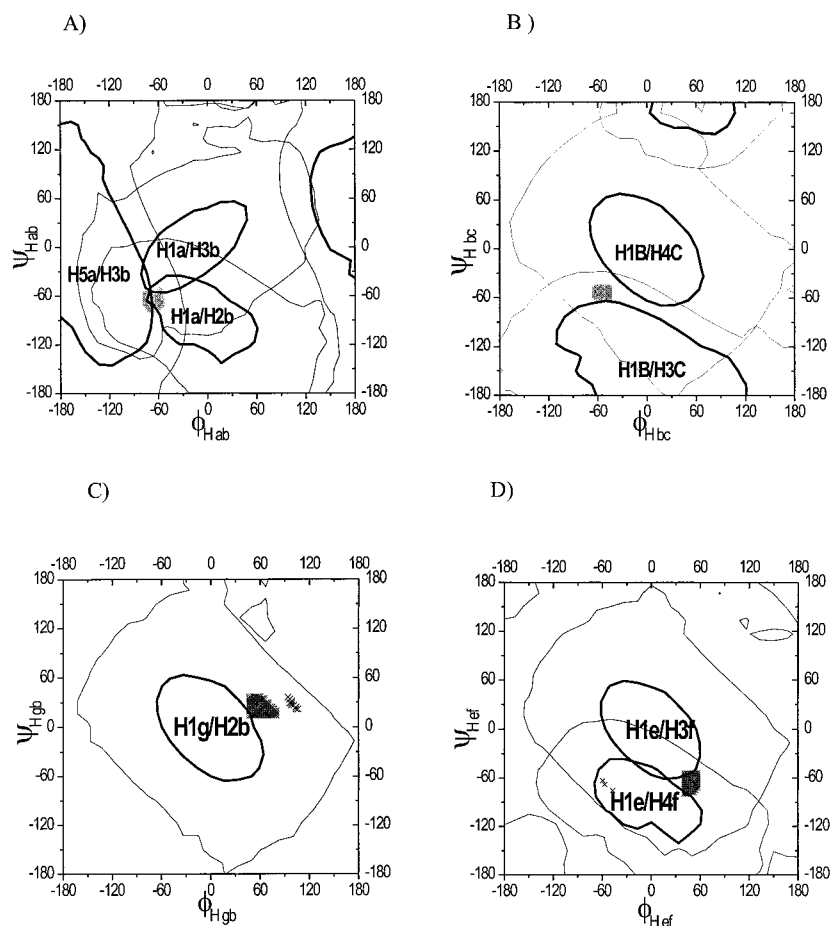


Figure 4.  $\phi/\psi$  plot of the conformations found in agreement with the interglycosidic NOE,  $^3J_{CH}$  and  $^3J_{CC}$  and  $^1D_{CH}$  during the grid search performed for residues a-b-(g)-c and e-f of J22 heptasaccharide. The experimental interglycosidic NOE considered are represented as contours corresponding to distances of 2.5 and 3.5 Å.

liquid crystal sample of J22 heptasaccharide are shown in Figure 3. The visual comparison of these spectra shows that for the case of the heptasaccharide in the 5% liquid crystal there is a loss in sensitivity, a larger line width and vanishing of some signals, which were present in the  $D_2O$  sample. Several effects may contribute to this loss of sensitivity such as the shortening in  $T_2$  relaxation in the more viscous liquid crystal media, the additional  $^1H$ - $^1H$  dipolar coupling interactions with passive nuclei during the proton acquisition, and a certain mismatch in the INEPT transfer periods caused by the additional  $^1D_{CH}$  contribution.

$^1D_{HH}$  values were measured for those signals having reasonable quality in both spectra and which could be unambiguously assigned. The  $^1D_{HH}$  obtained for J22 heptasaccharide in the 3% liquid crystal solution were practically zero, as the splittings observed in the

experiment were very similar to those in the  $D_2O$  sample. The  $^1D_{HH}$  values in the 5% liquid crystal were of a measurable magnitude and they were used for the structure calculations (Table 2).

A conformational study of J22 heptasaccharide was done using the residual dipolar couplings reported in Tables 1 and 2 combined with NMR structural data previously published for this heptasaccharide (Martin-Pastor et al., 2000c) including  $^1H$ - $^1H$  NOEs of Figure 4 and long-range inter-glycosidic scalar couplings  $^3J_{CH}$  and  $^3J_{CC}$  of Tables 3 and 4. As a first stage in the conformational study of J22 heptasaccharide, relaxed  $\phi/\psi$  conformational grid searches were performed independently for the glycosidic linkages of residues a-b-(g)-c and e-f components. These systematic searches provided information about all the possible single conformations of these residues compatible



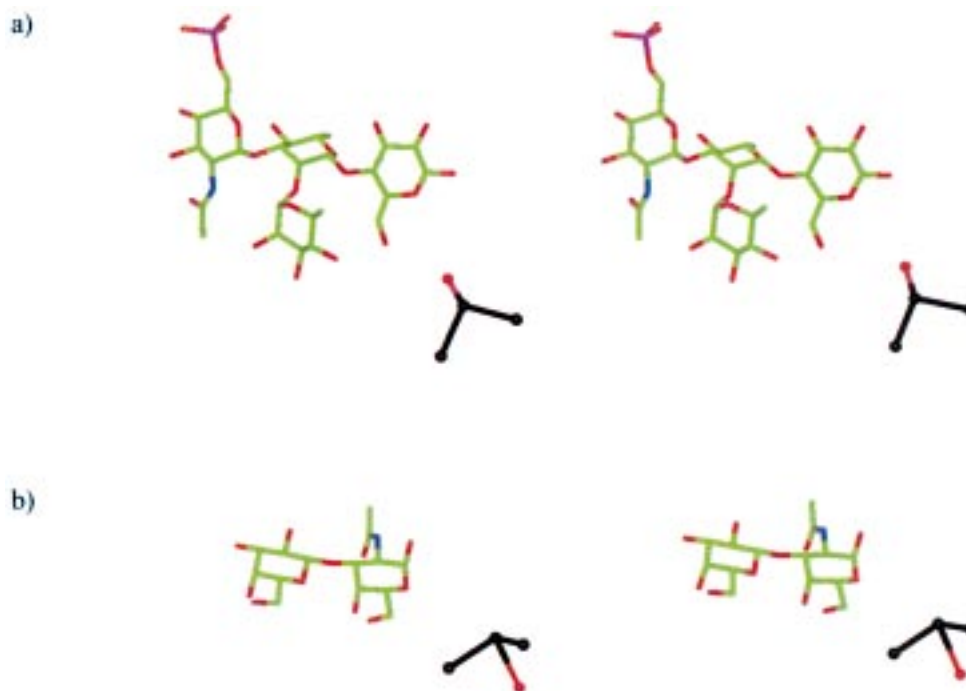


Figure 5. Stereo-view of the best conformer found in the r-SA calculations and the alignment tensor obtained for (a) r-SA of residues a-b-g-c of J22 heptasaccharide; (b) r-SA of residues e-f of J22 heptasaccharide.

Table 2. Experimental  ${}^nD_{HH}$  obtained for J22 heptasaccharide at 5% bicelle concentration and values calculated with the r-SA refined models of residues a-b-(g)-c and e-f

Proton pair	${}^nD_{HH}^{\text{exp}}$	${}^nD_{HH}^{\text{calc}}$ r-SA
H2-H3a	2.0	1.4
H3-H4g	0.0	1.3
H3-H4b	4.0	3.0
H3-H5b	2.5	1.4
H3-H4g	2.0	1.3
H1-H2c	-2.3	-0.9
H4-H5c	-1.0	-1.0
H4-H5e	1.0	0.9
H3-H2e	1.2	0.9
H2-H3f	1.8	1.5

with the available interglycosidic  ${}^1\text{H}$ - ${}^1\text{H}$  NOE,  ${}^3J_{\text{CH}}$ , and  ${}^3J_{\text{CC}}$  and the  ${}^1D_{\text{CH}}$  experimental data for these substructures. The intra-residual  ${}^nD_{\text{HH}}$  experimental data were not considered for the grid search protocol, because these restraints depend on both the internuclear distance and the orientation. Their interpretation requires further optimization of the pyranoside chair

conformation, as will be described below. The  $\phi/\psi$  maps obtained for the accepted conformers in the grid searches are given in Figure 4. These maps show that for all the glycosidic torsions considered, there is a single region where  ${}^1\text{H}$ - ${}^1\text{H}$  NOE distances,  ${}^3J_{\text{CH}}$ , and  ${}^3J_{\text{CC}}$  and the  ${}^1D_{\text{CH}}$  experimental data can be satisfactorily explained. Although the extents of the regions found are governed by the upper limits considered for the experimental error ( $\pm 1$  Hz for these figures), these results suggest that residues a-b-(g)-c and e-f of J22 heptasaccharide adopt a single well-defined conformation around the angles found in these maps with rather limited flexibility. Although the possibility that this single structure could correspond to a virtual conformation cannot be totally ruled out, its consistency with experimental data of different average time scales in solution, i.e.  ${}^1\text{H}$ - ${}^1\text{H}$  NOE,  ${}^3J_{\text{CH}}$  and  ${}^3J_{\text{CC}}$  scalar couplings, and  ${}^1D_{\text{CH}}$ , considerably reduces this possibility.

A complete model of the J22 heptasaccharide requires the conformation adopted by residues c-d-e. A simple calculation of the glycosidic torsion angles  $\phi$  and  $\psi$  of these two 1 $\rightarrow$ 6 glycosidic linkages using the available interglycosidic scalar coupling data  ${}^3J_{\text{CH}}$  and  ${}^3J_{\text{CC}}$  of Tables 4 and 5 and the correspond-

ing Karplus-type correlation curves (Tvaroska et al., 1989; Bose et al., 1998) showed that some of the scalar coupling constants are not compatible with a single dihedral angle within an error limit of  $\pm 1$  Hz. This situation indicates that there must be flexibility around these linkages, a situation similar to that observed in the study of the analogue J22 polysaccharide (Martin-Pastor et al., 1999). In addition, the linkages of residues c-d-e present three additional backbone torsions,  $\omega_d$ ,  $\gamma_d$  and  $\omega_e$ , for which there are no experimental data available and that could also provide additional flexibility. This flexibility around the linkages connecting residues c-d-e precludes the interpretation of all the residual couplings obtained for the heptasaccharide using a single alignment tensor.

Therefore three dynamically independent regions can be defined in the heptasaccharide corresponding to the substructures a-b-(g)-c, d and e-f, and a different alignment tensor is required for the interpretation of the residual dipolar coupling in the heptasaccharide in each of them. At this point the number of vectors for the refinement of the substructure of residue d (Table 1) is insufficient to determine an alignment tensor and because of that no further effort was made to interpret these data.

The substructures a-b-(g)-c and e-f of the J22 heptasaccharide obtained in the grid search were further refined in two independent r-SA calculations. Each simulation included all the available experimental data, i.e.  $^1D_{CH}$  and  $^nD_{HH}$  residual dipolar couplings of Tables 1 and 2, the interglycosidic NOE restraints (Figure 4), and long-range scalar  $^3J_{CH}$  and  $^3J_{CC}$  interglycosidic restraints of Table 3. The conformations obtained for the a-b-(g)-c and e-f substructures were again consistent with a single model within the regions previously found in the grid searches. The best conformation obtained in each r-SA of substructures a-b-(g)-c and e-f and its corresponding alignment tensor is shown in Figure 5a and 5b, respectively. The agreement of these structures with  $^1D_{CH}$  and  $^nD_{HH}$  data is shown in Tables 1 and 2. In Table 3 are shown the experimental and calculated long-range scalar couplings for these models and Figure 6 summarizes the agreement obtained with these data. Most of the scalar couplings and residual dipolar couplings are within the  $\pm 1$  Hz error limit considered. The glycosidic torsion angles of these models, given in Table 3, agree well with the strong and medium interglycosidic NOE distances of Figure 4. The starting structures for the r-SA calculation are taken from the best conformers of the grid calculation illustrated in Figure 4. Although the

r-SA results for  $^1D_{CH}^{calc}$  and  $^nD_{HH}^{calc}$  agree better with the experiment, the r-SA introduces only small changes of about 5 degrees in the glycosidic dihedral angles. More significant to the agreement with the experimental dipolar coupling values, especially of the  $^nD_{HH}^{calc}$ , are the changes in the puckering of the pyranoside rings which can be seen in the data of Table 6, which compares the sugar puckering angles resulting from the r-SA calculation with those of the grid calculation. The dihedral angles in the latter case are determined only by the force field, CVFF in this case, while in the r-SA calculation, the ring puckering is adjusted by a few degrees by the pseudopotential of the dipolar coupling restraint. These subtle changes could correspond to real effects in the sugar puckering or be caused by artifacts introduced by the residual dipolar couplings in order to find good agreement without reorienting the complete sugar chair as reported for the case in some proteins (Fischer et al., 1999). Although values of pyranoside ring puckering can be extracted from crystallographic data on monosaccharides, the values vary and are not known with precision. Also included in Table 6 are ring puckering values for the r-SA calculation without dipolar restraints, which shows that the experimental restraints introduce changes of a size comparable to differences between the X-plor (Charmm) and CVFF force fields.

A final r-SA simulation of the complete J22 heptasaccharide was performed with the aim of determining the flexibility around the linkages of residues c-d-e. The substructures of residues a-b-(g)-c and e-f were restrained in this simulation with the NOE,  $^3J_{CH}$ ,  $^3J_{CC}$ ,  $^1D_{CH}$  and  $^nD_{HH}$  experimental data and a different alignment tensor was used for each substructure. The linkages connecting residues c-d-e and the GalF ring d were modeled only with the molecular mechanics force field. This calculation provided 89 different low energy conformers in which different possibilities were found for the conformation around residues c-d-e while the structure of residues a-b-(g)-c and e-f in all of them was similar to the single structure previously found and represented in Figure 5. Cluster analysis (Kelley et al., 1996) of these 89 conformers grouped them into 9 significantly different clusters. According to these results the major points of flexibility in J22 heptasaccharide are located in the torsions  $\omega_d$ ,  $\gamma_d$  and  $\omega_e$ , which in general may adopt a combination of any of the three staggered conformations around angles  $-60^\circ$ ,  $60^\circ$  or  $180^\circ$ . Steric and electrostatic interactions imposed by the a-b-(g)-c and e-f substructures select these 9 combinations found in the clusters from

Table 3. Experimental interglycosidic  ${}^3J_{CH}$ ,  ${}^3J_{CC}$  and  ${}^2J_{CC}$  (Hz) and those calculated with the best conformer obtained in r-SA of substructures a-b-(g)-c and e-f of J22 heptasaccharide

Linkage	$\phi_H/\psi_H^b$	$\phi$		$\psi$		
		${}^3J_{H1C1O1Cx}$	${}^3J_{C2C1O1Cx}$	${}^3J_{C1O1CxCxHx}$	${}^3J_{C1O1CxCx-1}$	${}^3J_{C1O1CxCx+1}$
a – b	Exp.	2.0	2.0	1.7	<0.8	2.8
	r-SA	-54.5/-53.2	2.0	<sup>a</sup>	2.2	0.3
b – c	Exp.	0.9	2.5	2.4	<0.8	2.2
	r-SA	-73.3/-50.5	0.8	3.5	2.4	0.5
g – b	Exp.	0.7		4.3	<0.8	
	r-SA	68.3/26.9	1.0		4.5	
e – f	Exp.		2.4	1.3		<0.8
	r-SA	50.6/-60.6	2.4	1.6		0.4

<sup>a</sup>In-plane oxygen precludes calculation of this scalar coupling (Bose et al., 1998).

<sup>b</sup>These angles also correspond to the final structures found with MC/EM.

Table 4. Torsion angles for the c-d linkages and the experimental and calculated  ${}^3J_{CH}$  and  ${}^3J_{CC}$  for the more representative conformers found in the r-SA and cluster analysis of J22 heptasaccharide

	$\omega_d$	$\gamma_d$	$\phi_{H\ cd}/\psi_{O\ cd}$	$\phi$		$\psi$	
				${}^3J_{H1C1O1C6}$	${}^3J_{C2C1O1C6}$	${}^3J_{H6C6C1O1}$	${}^3J_{C1O1C6C5}$
Exp.				1.4	<0.8	1.9	1.8
Clust. 1	168.8	46.3	111.9/80.9	1.5	1.4	3.0 <sup>a</sup>	0.2
Clust. 2	-62.0	-72.6	110.7/-177.4	1.4	1.4	1.8 <sup>a,b</sup>	3.6
Clust. 3	-137.8	60.9	110.7/-60.8	1.4	1.5	1.7 <sup>a</sup>	0.9
Clust. 4	-162.8	50.9	110.0/-73.7	1.4	1.5	2.2 <sup>a</sup>	0.4
Clust. 5	-179.8	-165.0	110.3/-156.4	1.4	1.5	3.2 <sup>a</sup>	3.1
Clust. 6	-164.4	47.1	110.1/-155.1	1.4	1.5	3.4 <sup>a</sup>	3.0
Clust. 7	57.8	-66.4	110.4/-156.4	1.4	1.5	3.3 <sup>a</sup>	3.1
Clust. 8	-167.2	-168.2	110.4/-161.0	1.4	1.5	3.1 <sup>a</sup>	3.2
Clust. 9	-66.6	-80.0	110.6/-108.0	1.4	1.5	4.0 <sup>b</sup>	0.5

<sup>a</sup>Calculated for H6<sub>pro-R</sub>.

<sup>b</sup>Calculated for H6<sub>pro-S</sub>.

Table 5. Torsion angles for the d-e linkages and the experimental and calculated  ${}^3J_{CH}$  and  ${}^3J_{CC}$  for the more representative conformers found in the r-SA and cluster analysis of J22 heptasaccharide

	$\omega_e$	$\phi_{H\ de}/\psi_{O\ de}$	$\phi$		$\psi$	
			${}^3J_{H1C1O1C6}$	${}^3J_{C2C1O1C6}$	${}^3J_{H6C6C1O1}$	${}^3J_{C1O1C6C5}$
Exp.			1.5	2.1	1.3	2.6
Clust. 1	-148.9	61.8/167.8	1.5	3.6	0.8 <sup>a</sup>	3.5
Clust. 2	84.9	62.3/-161.7	1.4	3.6	0.6 <sup>b</sup>	3.3
Clust. 3	177.3	62.3/-163.3	1.4	3.6	0.7 <sup>b</sup>	3.4
Clust. 4	68.9	62.1/-178.3	1.5	3.6	1.5 <sup>b,a</sup>	3.6
Clust. 5	78.2	110.5/-171.2	1.4	1.6	1.0 <sup>b</sup>	3.6
Clust. 6	-169.0	62.3/-154.0	1.4	3.6	0.5 <sup>b</sup>	3.0
Clust. 7	75.2	61.7/155.7	1.5	3.6	0.5 <sup>a</sup>	3.1
Clust. 8	-62.7	61.7/-152.2	1.5	3.6	0.5 <sup>b</sup>	2.9
Clust. 9	-46.3	62.0/-132.1	1.5	3.6	1.3 <sup>b</sup>	1.7

<sup>a</sup>Calculated for H6<sub>pro-R</sub>.

<sup>b</sup>Calculated for H6<sub>pro-S</sub>.

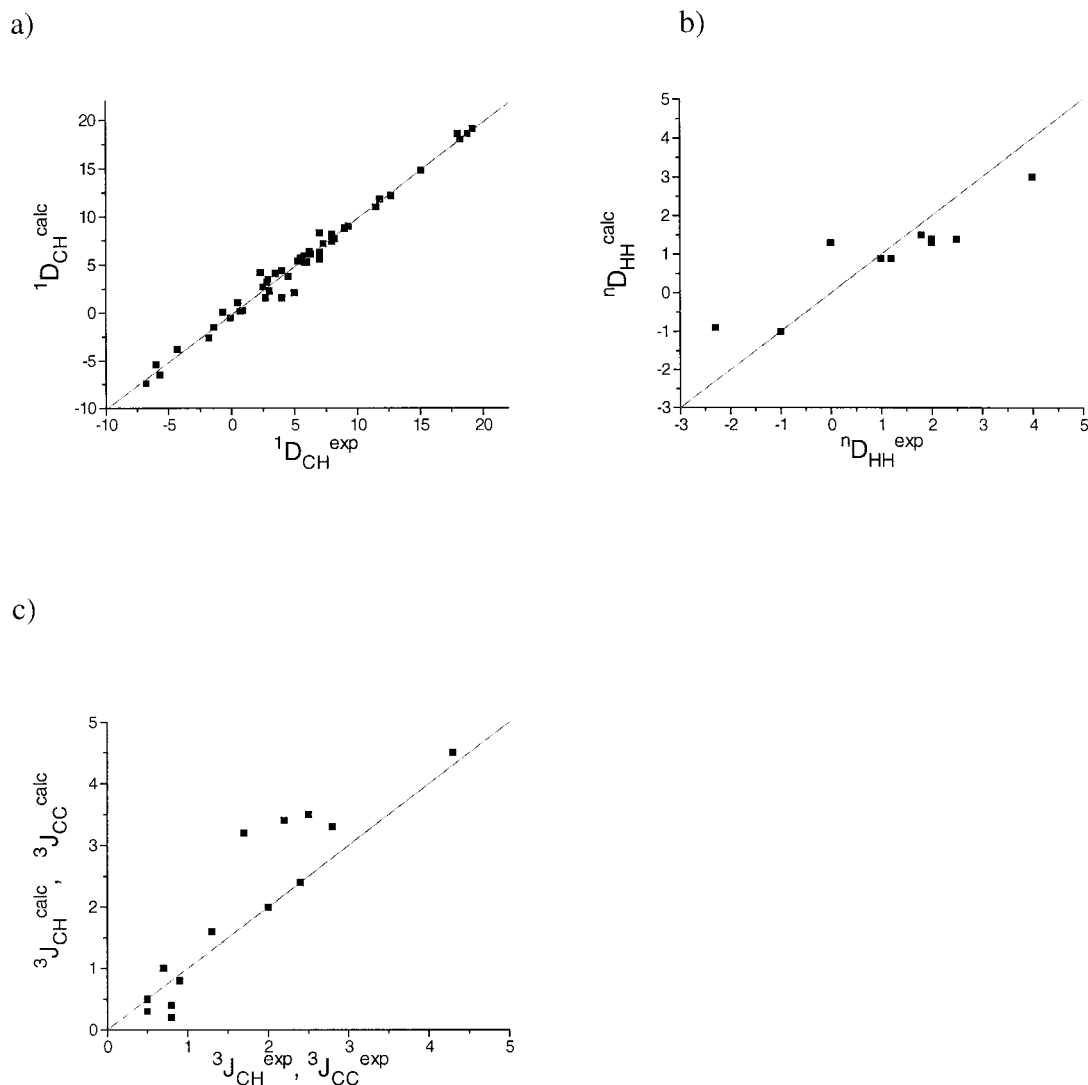


Figure 6. Plot of the experimental versus calculated data with the r-SA oriented models of substructures a-b-(g)-c and e-f. (a)  $^1D_{CH}$ , (b)  $^nD_{HH}$ , (c)  $^3J_{CH}$  and  $^3J_{CC}$ .

the 27 possible combinations. The glycosidic torsions  $\phi_{H cd}/\psi_{O cd}$  adopt either  $110^\circ / -60^\circ$  or  $110^\circ / -160^\circ$ ,  $110^\circ / -110^\circ$  and  $\phi_{H de}/\psi_{O de}$  angles of  $60^\circ / 180^\circ$  or  $110^\circ / 180^\circ$ . The 9 representative structures found were tested for agreement with the experimental  $^3J_{CH}$  and  $^3J_{CC}$  of the 1→6 glycosidic torsions which were not included in the r-SA calculation. These results are given in Tables 4 and 5 with the respective glycosidic dihedral angles. It can be seen in these tables that the structures from clusters 2, 3 and 4 agree with all the  $^3J_{CH}$  of residues c-d and d-e data within  $\pm 1$  Hz. From these clusters, only 3 and 4 agree with all the  $^3J_{CC}$  data within  $\pm 1$  Hz if the coupling  $^3J_{C2C1O1C6}$  of linkage

d-e is excepted. This last coupling is overestimated ( $>1$  Hz) by all clusters except by cluster 5. A linear combination involving at least cluster 5 is required to explain the data. A combination of clusters 3, 4 and 5 can be proposed to satisfy both the  $^3J_{CH}$  and  $^3J_{CC}$  data of residues c-d-e within  $\pm 1$  Hz by adjusting the relative populations, for example to 0.10:0.15:0.75, respectively. A view of the more representative structure of these three clusters can be seen in Figure 7a-c. However, this solution is not unique and several other combinations may be proposed with different populations for some of the other clusters. At this point more experimental restraints would be required to bet-

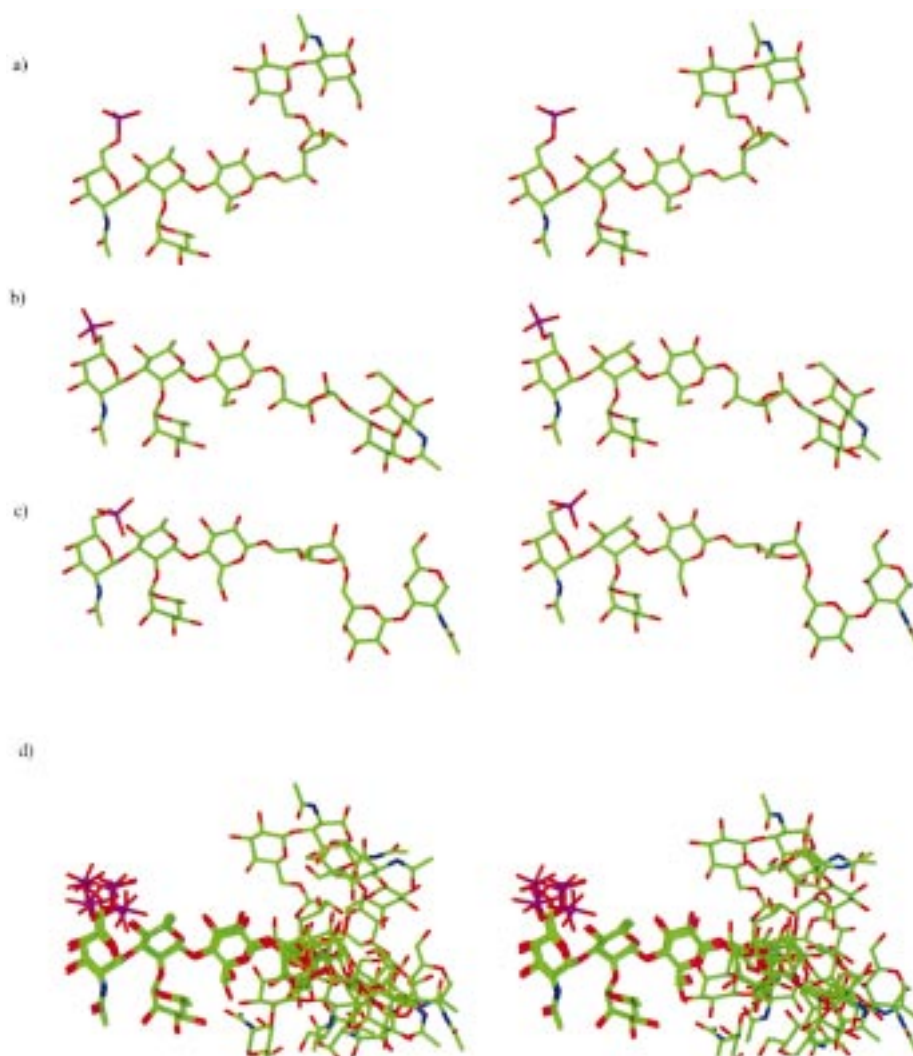


Figure 7. Stereo-view of the four most populated cluster families found during the r-SA of J22 heptasaccharide.

ter define the populations and structure of the flexible linkages connecting residues c-d-e. A superposition of the 9 most representative cluster structures found is given in Figure 7d.

### Conclusions

The use of residual dipolar couplings  $^1D_{CH}$  and  $^nD_{HH}$  provides a new and valuable set of structural restraints that can be used in concert with other NMR restraints for conformational studies with carbohydrates.  $^1D_{CH}$  were measured for a heptasaccharide in a  $t_1$ -coupled HSQC experiment. This paper has shown that the  $^nD_{HH}$  can be measured in a natural abundance carbo-

hydrate sample using a  $^1H$ - $^{13}C$   $^1H$ - $^1H$  TOCSY experiment (Willker et al., 1992), providing  $^{13}C$  chemical shift resolution critical for complex oligosaccharides for which overlap of  $^1H$  spectral lines is a serious problem.

We have compared different complementary protocols for the interpretation of the residual dipolar coupling data in terms of oligosaccharide conformation. The r-SA method allows for small adjustments in the molecular conformation to represent the dipolar coupling data. For sugar pyranosides, this allows for small corrections in the ring puckering which may not be described with sufficient accuracy by the currently available molecular modeling force fields. But on the other hand, the relaxed energy grid search

Table 6. Pyranose ring puckering torsion angles for the best conformers of residues a-b-(g)-c and e-f obtained for J22 heptasaccharide in different calculations. The differences in the glycosidic dihedral angles in the conformers selected are below 5°

Dihedral	Force field	Dipolar restraints	Residue					
			a	b	g	c	e	f
C1-C2-C3-C4	X-plor <sup>a</sup>	Yes	-55.5	50.6	46.8	-51.6	-51.6	-52.1
	X-plor <sup>b</sup>	No	-53.6	55.1	54.0	-50.5	-52.8	-52.2
	CVFF <sup>c</sup>	No	-51.8	56.6	51.6	-50.1	-51.5	-51.5
C2-C3-C4-C5	X-plor <sup>a</sup>	Yes	52.2	-53.2	-52.0	47.8	48.9	54.5
	X-plor <sup>b</sup>	No	53.7	-53.3	-55.2	50.4	52.6	50.8
	CVFF <sup>c</sup>	No	53.1	-52.9	-51.2	48.4	52.3	52.4
C3-C4-C5-O5	X-plor <sup>a</sup>	Yes	-52.0	58.4	59.0	-51.1	-53.1	-58.0
	X-plor <sup>b</sup>	No	-55.6	55.0	57.0	-55.9	-56.6	-53.5
	CVFF <sup>c</sup>	No	-54.0	50.3	51.1	-50.7	-54.0	-53.9
C4-C5-O5-C1	X-plor <sup>a</sup>	Yes	59.0	-63.6	-62.7	60.5	62.7	60.9
	X-plor <sup>b</sup>	No	60.5	-60.8	-59.0	63.2	62.7	60.6
	CVFF <sup>c</sup>	No	58.3	-56.5	-56.0	59.3	59.7	57.7
C5-O5-C1-C2	X-plor <sup>a</sup>	Yes	-62.8	61.9	56.5	-65.2	-65.7	-58.2
	X-plor <sup>b</sup>	No	-60.5	63.9	57.3	-62.9	-62.8	-62.3
	CVFF <sup>c</sup>	No	-57.0	60.7	56.2	-60.5	-58.4	-56.7
O5-C1-C2-C3	X-plor <sup>a</sup>	Yes	59.5	-54.2	-47.9	59.9	59.0	53.1
	X-plor <sup>b</sup>	No	55.9	-59.9	-53.6	55.8	57.1	56.8
	CVFF <sup>c</sup>	No	51.3	-57.7	-51.5	53.6	51.8	52.1

<sup>a</sup>r-SA with NOE, <sup>n</sup>J<sub>CH</sub>, <sup>n</sup>J<sub>CC</sub>, <sup>n</sup>D<sub>HH</sub>, and <sup>1</sup>D<sub>CH</sub> restraints.

<sup>b</sup>r-SA with NOE, <sup>n</sup>J<sub>CH</sub>, <sup>n</sup>J<sub>CC</sub> restraints.

<sup>c</sup>Grid search with glycosidic dihedral angle restraints.

method gives reasonably good conformations for rigid oligosaccharide substructures. The conformations are very similar to the final r-SA results and the protocol does not require a prior model or estimate of the orientation tensor and thus can be used to obtain starting values for r-SA procedures.

A challenging problem for the interpretation of the residual dipolar couplings and other NMR restraints in carbohydrates is provided by the inherent flexibility among certain sugar linkages, which may occasionally result in incompatible NMR restraints. Incompatible NMR structural restraints were detected for <sup>3</sup>J<sub>CH</sub> and <sup>3</sup>J<sub>CC</sub> scalar coupling data involving the 1→6 glycosidic linkages of J22 heptasaccharide, which corresponded to regions of higher flexibility. The detection of such a flexible region in the heptasaccharide precludes a simple interpretation of the residual dipolar couplings of the complete heptasaccharide in terms of a single alignment tensor. Although the observed dipolar coupling values for a flexible structure represent a simple average over contributing conformers, calculation of the average values requires not only the statistical weight of the contributing conformers but

also their correct orientation tensors. The many unknowns involved preclude use of the dipolar coupling data to refine the flexible part of the J22 oligosaccharide structure. However, the residual dipolar coupling data were still useful to refine the substructures formed by residues a-b-(g)-c and e-f for which such flexibility was not present. For each substructure the different sugar units were considered as being dynamically equivalent (Fischer et al., 1999) and the residual dipolar coupling data were interpreted with a unique alignment tensor.

A previous study (Martin-Pastor et al., 2000) showed striking similarities between the NOE, <sup>n</sup>J<sub>CH</sub> and <sup>n</sup>J<sub>CC</sub> across the glycosidic linkages, suggesting that the structured regions of the heptasaccharide and polysaccharide adopt the same conformation. On the other hand, we have no experimental evidence concerning the similarity of the conformation adopted by the other, more flexible glycosidic linkages joining residues c-d-e. In addition there is a phosphate linkage connecting the repeating units in the polysaccharide which could adopt different conformations affecting the overall polysaccharide structure.

The influence of the interactions with the liquid crystalline medium on the molecular structure and dynamics is uncertain at present. Zweckstetter and Bax (2000) have recently proposed a way of predicting the alignment tensor for certain proteins based on the overall shape of the molecule and the steric interactions with the external surface of a neutral bilayer medium such as DMPC/DHPC or with phages. For liquid crystals such as the positively charged CPCI/hexanol/brine system used here and the negatively charged heptasaccharide, attractive electrostatic interactions may also be present. At this point, there is no plausible model to treat attractive interactions with the liquid crystal and the flexibility found in the heptasaccharide could severely complicate this kind of study. In any case, the strength of the interaction responsible for the orientation is very small. It seems unlikely that these interactions could have any influence on the oligosaccharide conformation.

### Acknowledgements

The authors would like to thank Ad Bax and Markus Zweckstetter for helpful discussions. We thank Nico Tjandra for providing his version of the X-plor program used in the calculations, and John Cisar for providing a J22 polysaccharide sample. This research was supported by NSF Grant MCB9724133 and Ministerio de Educacion y Cultura of Spain for a postdoctoral fellowship to M.M.-P.

### References

- Abeygunawardana, C., Bush, C.A. and Cisar, J.O. (1990) *Biochemistry*, **29**, 234–248.
- Bolon, P.J., Al-Hashimi, H.M. and Prestegard, J.H. (1999) *J. Mol. Biol.*, **293**, 107–115.
- Bose, B., Zhao, S., Stenutz, R., Cloran, F., Bondo, F., Bondo, G., Hertz, B., Carmichael, I. and Serianni, A. (1998) *J. Am. Chem. Soc.*, **120**, 11158–11173.
- Brünger, A.T. (1992) *X-Plor, version 3.840. A system for X-ray crystallography and NMR*, Yale University, New Haven, CT.
- Bush, C.A., Martin-Pastor, M. and Imberty, A. (1999) *Annu. Rev. Biophys. Biomol. Struct.*, **28**, 269–293.
- Cisar, J.O., Sandberg, A.L., Abeygunawardana, C., Reddy, G.P. and Bush, C.A. (1995) *Glycobiology*, **5**, 655–662.
- Clore, G.M., Gronenborn, A.M. and Bax, A. (1998a) *J. Magn. Reson.*, **133**, 216–221.
- Clore, G.M., Gronenborn, A.M. and Tjandra, N. (1998b) *J. Magn. Reson.*, **131**, 159–162.
- Fischer, M.W.F., Losonczi, J.A., Weaver, J.L. and Prestegard, J.H. (1999) *Biochemistry*, **38**, 9013–9022.
- Gomati, R., Appell, J., Bassereau, P., Marignan, J. and Porte, G. (1987) *J. Phys. Chem.*, **91**, 6203–6210.
- Gorler, A., Ulyanov, N.B. and James, T.L. (2000) *J. Biomol. NMR*, **16**, 147–164.
- Hagler, A.T., Lifson, P. and Dauber, P. (1979) *J. Am. Chem. Soc.*, **101**, 5122–5130.
- John, B.H. (1992) *J. Magn. Reson.*, **A101**, 113.
- Kelley, L.A., Gardner, S.P. and Sutcliffe, M.J. (1996) *Protein Eng.*, **9**, 1063–1065.
- Kiddle, G.R. and Homans, S.W. (1998) *FEBS Lett.*, **436**, 128–130.
- Landersjö, C., Hoog, C., Maliniak, A. and Widmalm, G. (2000) *J. Phys. Chem. B*, **104**, 5618–5624.
- Losonczi, J.A., Andrec, M., Fischer, M.W.F. and Prestegard, J.H. (1999) *J. Magn. Reson.*, **138**, 334–342.
- Martin-Pastor, M. and Bush, C.A. (1999) *Biochemistry*, **38**, 8045–8055.
- Martin-Pastor, M. and Bush, C.A. (2000a) *Carbohydr. Res.*, **323**, 147–155.
- Martin-Pastor, M. and Bush, C.A. (2000b) *Biochemistry*, **39**, 4674–4683.
- Martin-Pastor, M. and Bush, C.A. (2000c) *Biopolymers*, **54**, 235–248.
- Ottiger, M. and Bax, A. (1999) *J. Biomol. NMR*, **13**, 187–191.
- Ottiger, M., Delaglio, F. and Bax, A. (1998) *J. Magn. Reson.*, **131**, 373–378.
- Porte, G., Gomati, R., El Haitamy, O., Appell, J. and Marignan, J. (1986) *J. Phys. Chem.*, **90**, 5746–5751.
- Shimizu, H., Donohue-Rolfe, A. and Homans, S.W. (1999) *J. Am. Chem. Soc.*, **121**, 5815–5816.
- Thompson, G.S., Shimizu, H., Homans, S.W. and Donohue-Rolfe, A. (2000) *Biochemistry*, **39**, 13153–13156.
- Tian, F., Bolon, J. and Prestegard, J.H. (1999) *J. Am. Chem. Soc.*, **121**, 7712–7713.
- Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.
- Tjandra, N., Grzesiek, S. and Bax, A. (1996) *J. Am. Chem. Soc.*, **118**, 6264–6272.
- Tjandra, N., Marquardt, J. and Clore, G.M. (2000) *J. Magn. Reson.*, **142**, 393–396.
- Tvaroska, I., Hricovini, H. and Perakova, E. (1989) *Carbohydr. Res.*, **189**, 359–362.
- Willker, W. and Leibritz, D. (1992) *J. Magn. Reson.*, **99**, 421–425.
- Zweckstetter, M. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 3791–3792.