Journal of Biomolecular NMR, **20:** 351–363, 2001.
KLUWER/ESCOM

Quantitative conformational analysis of the core region of N-glycans using residual dipolar couplings, aqueous molecular dynamics, and steric alignment

Andrew Almond* & Jens Ø. Duus
Carlsberg Laboratory, Gamle Carlsberg Vej 10, 2500 Valby, Denmark

Received 4 April 2001; Accepted 18 May 2001

Key words: carbohydrate, conformation, molecular dynamics, N-glycan, oligosaccharide, residual dipolar coupling, water

Abstract

A method is described for quantitatively investigating the dynamic conformation of small oligosaccharides containing an α(1 → 6) linkage. It was applied to the oligosaccharide Man-α(1 → 3) {Man−α (1 → 6)} Man-α-O-Me, which is a core region frequently observed in N-linked glycans. The approach tests an aqueous molecular dynamics simulation, capable of predicting microscopic dynamics, against experimental residual dipolar couplings, by assuming that alignment is caused purely by steric hindrance. The experimental constraints were heteronuclear and homonuclear residual dipolar couplings, and in particular those within the α(1 → 6) linkage itself. Powerful spin-state-selective pulse sequences and editing schemes were used to obtain the most relevant couplings for testing the model. Molecular dynamics simulations in water over a period of 50 ns were not able to predict the correct rotamer population at the α(1 → 6) linkage to agree with the experimental data. However, this sampling problem could be corrected using a simple maximum likelihood optimisation, indicating that the simulation was modelling local dynamics correctly. The maximum likelihood prediction of the residual dipolar couplings was found to be an almost equal population of the *gg* and *gt* rotamer conformations at the α(1 → 6) linkage, and the *tg* conformation was predicted to be unstable and unpopulated in aqueous solution. In this case all twelve measured residual dipolar couplings could be satisfied. This conformer population could also be used to make predictions of scalar couplings with the use of a previously derived empirical equation, and is qualitatively in agreement with previous predictions based on NMR, X-ray crystallography and optical data.

Introduction

At the heart of biological systems the constituent molecules provide function by their dynamic interplay. Therefore, after entering an era where the chemical structure of the key players in the biological world are known, of which carbohydrates are a major group, a logical next step is to elucidate their self-dynamics (Homans, 1993). A premise underlying the work presented here is that these dynamics are resultant from water-solute interaction (Tvaroska et al., 1990). However, these interactions are potentially difficult to describe physically, in which case computer simulations can be used to make predictions in advance of the development of more simplistic models. It is proposed that when computer simulations can be tested against experimental data, then they can be used to provide reductive information about dynamics which is not evident from the experimental data alone. In effect it is proposed that the physics present in simulations can be used to constrain the interpretation of the experimental data. In our previous work, molecular dynamics simulations of carbohydrates have been tested against hydrodynamic, X-ray crystallography, and NMR data (Almond et al., 1998, 2000, 2001; Almond and Sheehan, 2000).

*To whom correspondence should be addressed. E-mail: aa@crc.dk

The $\alpha(1 \rightarrow 6)$ linkage plays an underlying role in the conformation and dynamics of N-linked oligosaccharides (Imberty et al., 1993). These molecules have been found to be involved in essential cellular processes (Dwek, 1996; Rudd and Dwek, 1997), particularly at the cell-surface, and their malfunction is implicated in the progression of diseases (Rudd et al., 1999). A detailed understanding of this type of linkage, and hence the molecules in which it is present, may be forthcoming from a study of dynamics and water. However, present molecular models require testing against experimental data before they can be considered useful.

Previous studies have used molecular modelling to gain theoretical insights into the dynamics of the $\alpha(1 \rightarrow 6)$ linkage in various environments (Dowd and French, 1995; Rutherford and Homans, 1994; Spronk et al., 1995; Quasba et al., 1994; Engelsen and Rasmussen, 1993). Complementary NMR techniques such as scalar couplings and relaxation effects have also been used to provide information about the linkage (Jansson et al., 1994; Sayers and Prestegard, 2000; Homans et al., 1987). Most recently residual dipolar couplings have been measured by NMR in oligosaccharides containing $1 \rightarrow 6$ linkages, using magnetically aligned co-solutes (Tian et al., 2001; Neubauer et al., 2001).

Since carbohydrate glycosidic linkages are potentially flexible, the interpretation of experimental data obtained from oligosaccharides must initially take dynamics into account. However, this has not precluded the use of residual dipolar couplings (Martin-Pastor and Bush, 2000a,b; Neubauer et al., 2001). Furthermore, the use of residual dipolar couplings in structure determination, and in particular when conformational exchange is present, has recently been reviewed (Prestegard et al., 2000). Compared with other techniques it has the ability to provide accurate global information, which is a stringent test for molecular models.

The use of a dilute crystalline aligning medium in NMR experiments to provide tuneable alignment of macromolecules, and hence observation of residual dipolar couplings is relatively new (Tjandra and Bax, 1997; Hansen et al., 1998). Fortunately, the theory describing such alignment is well established. If an internuclear vector subtends an angle Θ to a magnetic field then the dipolar contribution to coupling is given by Equation 1. Classically, the maximal value of the dipolar coupling is dependent on the gyromagnetic nuclear ratios γ_n and γ_m and the separation of the nuclei r, Equation 2.

$$D^{nm} = D^{nm}_{\max} \left\langle \frac{3\cos^2 \Theta - 1}{2} \right\rangle, \tag{1}$$

$$D^{nm}_{\max} = \frac{\mu_0 \gamma_n \gamma_m \hbar}{4\pi r^3}. \tag{2}$$

Under isotropic free tumbling in solution the dipolar spin-spin interactions average to zero due to their intrinsic symmetry. However, when an aligning material is present this is not quite true, and a net residual dipolar coupling is observed. The degree of alignment can be conveniently described by an order matrix, in this case a Saupe order matrix (Saupe and Englert, 1963), Equation 3. Here $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ otherwise. In this model, the angles α_i define the angle between the ith molecular axis and the director (the magnetic field) in the *laboratory frame*, as shown in Figure 1.

$$S_{ij} = \frac{1}{2} \left\langle 3 \cos \alpha_i \cos \alpha_j - \delta_{ij} \right\rangle \tag{3}$$

Using the order matrix, Equation 3, the expression for the residual dipolar couplings can be transformed to a molecular frame, Equation 4. Here β_i defines the angle between the internuclear vector and the ith molecular axis of the *molecular frame.*

$$D^{nm} = D^{nm}_{\max} S_{ij} \cos \beta^{nm}_i \cos \beta^{nm}_j \tag{4}$$

The aim of the present work is to test aqueous molecular dynamics simulations against measured residual dipolar couplings in a small oligosaccharide containing an $\alpha(1 \rightarrow 6)$ linkage. Assuming an agreement can be obtained, the simulation can then be used to understand how NMR in partially aligned systems can be used to study dynamics in such molecules, and consequently as a basis for understanding the dynamics of this linkage in solution.

In this study the molecule shown in Figure 2 was chosen as the model system. The molecule selected here is the core region of many N-linked glycans. In contrast to previous work (Tian et al., 2001), the residual dipolar couplings were measured using newly developed spin-state-selective pulse sequences to minimise overlap of the desired resonances. Our approach is also novel in that aqueous molecular dynamics simulations were then used to aid interpretation of the residual dipolar couplings, as opposed to static molecular models used in other studies of carbohydrates.

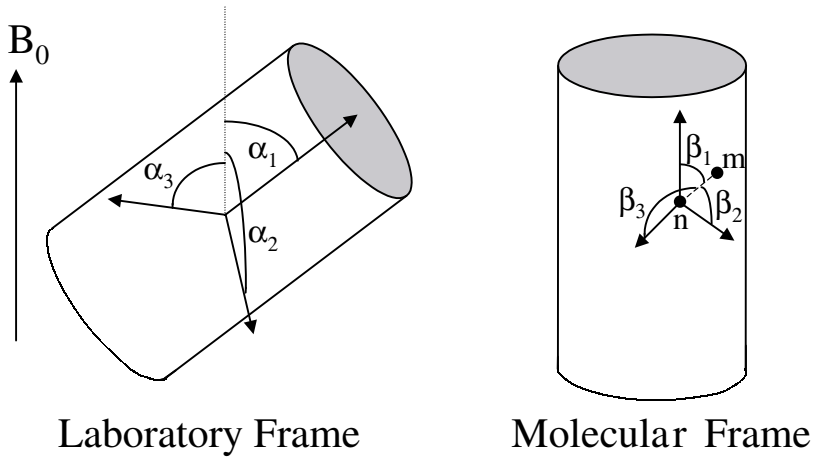

Figure 1. Angles used to express the molecular alignment. The α_i define the angles between the ith molecular axis and the director (the magnetic field) in the *laboratory frame*, and the β_i define the angles between the internuclear vector and the ith molecular axis of the *molecular frame.*

This was performed in the following way. Firstly, the overall performance of the simulation was tested by comparing theoretical predictions against, among others, heteronuclear residual dipolar couplings situated throughout the oligosaccharide. Secondly, using a maximum likelihood approach and including residual dipolar couplings in the $\alpha(1 \rightarrow 6)$ linkage region, the conformer preference of this flexible linkage is predicted. At the centre of our experimental design was that the $\alpha(1 \rightarrow 3)$ linkage, see Figure 2, only assumes a single dynamic conformation which does not significantly affect overall molecular conformation (Almond et al., 2001), and hence any conformational mobility of the $\alpha(1 \rightarrow 6)$ linkage can be more clearly identified. In contrast to previous approaches (Tian et al., 2001; Neubauer et al., 2001) the methodology presented here did not rely on back calculation of order matrices from the experimental data.

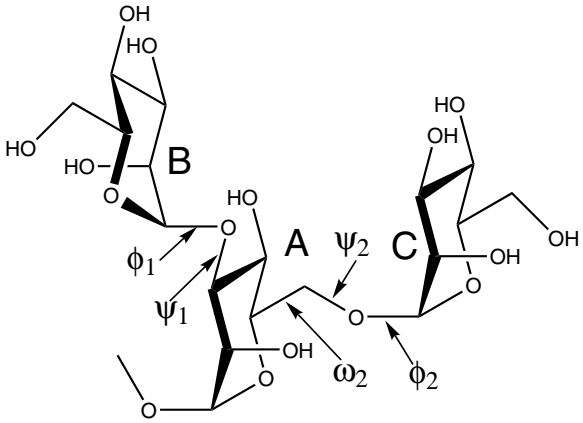

Figure 2. Structure of the mannose oligosaccharide. Also shown is the sugar ring labelling used throughout and the dihedral definitions. At the $\alpha(1 \rightarrow 3)$ linkage the definition was ϕ_1 = H1-C1-O3-C3 and ψ_1 = C1-O3-C3-H3. For the $\alpha(1 \rightarrow 6)$ linkage ϕ_2 = H1-C1-O6-C6, ψ_2 = C1-O6-C6-C5, and ω_2 = O6-C6-C5-O5.

Materials and methods

Sample preparation

The mannose oligosaccharide (Figure 2) was purchased from Glycorex AB (Sweden). The oligosaccharide was used without further purification. Two NMR samples were prepared at a concentration of 5 mg/ml oligosaccharide in D_2O, with 10 mM sodium phosphate buffer (pH 7.6) and 2 mM magnesium chloride. To one sample Pf1 phages were added to give a 14.0 ± 0.2 Hz splitting of the ^2H NMR signal of D_2O at 300 K. The phages were purchased from ASLA Ltd. (Latvia).

Nuclear magnetic resonance

All NMR spectra were recorded on a Varian Unity Inova 800 MHz spectrometer. The temperature was held constant at 300 K during experiments, which is the same temperature that all of the molecular dynamics simulations were performed at. Standard two-dimensional COSY and HSQC experiments were performed on the free molecule in solution to obtain both the proton and carbon chemical-shift assign-

ment. The assignments have been published elsewhere (Sayers and Prestegard, 2000).

One-bond couplings

One-bond couplings between carbon (S) and proton (I) were measured using a spin-state-selective (S^3) approach, and in particular the S^3CT element. The S^3CT element allows coherence transfer to either the I^-S^α or I^-S^β single-quantum coherence in IS systems, by a selective single-quantum S-spin π pulse (Sørensen et al., 1997). The pulse-sequence and editing methodology has been published previously (Lerche et al., 1999). However, it was modified slightly to enable the simultaneous measurement of IS couplings in both IS and I_2S systems. The editing scheme in the original S^3CT experiment did not allow for the extra coupling present in three-spin systems (Sørensen et al., 1997). To resolve this issue the editing scheme was changed to filter transverse S magnetisation. In this case there is no coupling to the third spin present after editing. This is analogous to the S^3P editing scheme (Meissner et al., 1998) and those proposed by others (Andersson et al., 1998; Ottiger et al., 1998). The observed effective coherence transfers (α/β) following editing are shown in Equation 5.

$$2\left(I_{1z}+I_{2z}\right)S^- \rightarrow \left(I_1^- + I_2^-\right)S^{\alpha/\beta}. \qquad (5)$$

Two two-dimensional S^3CT-HSQC spectra were recorded. They were edited such that one contained only the I^-S^α coherences and the other only the I^-S^β. Aquisition parameters: direct dimension was proton with a spectral width of 3 kHz, a digital resolution of 4096 complex points, and centred on the water resonance. Water presaturation was applied during the relaxation period. The indirect dimension was carbon with a spectral width of 12 kHz. For each t_1 16 scans were recorded, and a total of 512 t_1 increments were recorded. Analysis parameters: The data was transformed using a sine-bell squared apodisation function shifted by $2\pi/5$. Strip transforms were performed after zero-filling to an array of 16384×4096, to obtain a separation of 0.18 Hz between data points in the direct (measured) dimension.

The edited spectrum containing only I^-S^α coherences is shown in Figure 3. Equivalent one-dimensional slices in F2 were taken from both transformed spectra. The measured displacement between similar peaks in the two edited spectra corresponds to the scalar coupling. The above spectra were re-recorded with phages present, and the IS couplings

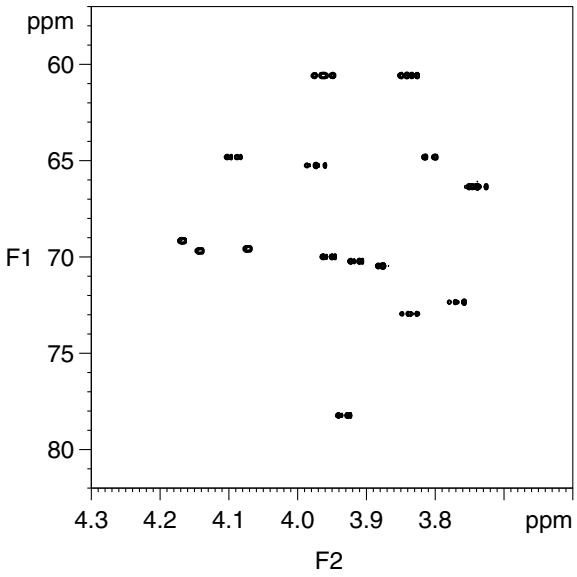

Figure 3. The α-edited version of the two-dimensional (S^3CT-HSQC) spectra recorded without phages present, at 800 MHz proton frequency. The F1 dimension is carbon and the F2 dimension is proton. Note: the H1 region of the spectra was omitted for clarity.

measured in an analogous way. The extra couplings correspond to the residual dipolar couplings induced by the phages.

Geminal couplings

Measurement of the geminal proton coupling at the $\alpha\,(1\rightarrow6)$ linkage was also performed using two-dimensional spectra, since significant overlap was observed in one-dimensional proton spectra when phages were present, even at 800 MHz proton resonance frequency. A coherence-order and spin-state-selective (COS^3) pulse sequence (Untidt et al., 1999) was used for this purpose. This pulse sequence and editing scheme achieved one (α/β) of the following specific transfers, Equation 6.

$$2\left(I_{1z}+I_{2z}\right)S^- \rightarrow I_1^- I_2^{\alpha/\beta} + I_1^{\alpha/\beta} I_2^-. \qquad (6)$$

Two two-dimensional spectra were recorded, one for the α and one for the β transfer in Equation 6, with and without phage. The scalar couplings and residual dipolar couplings were measured in an analogous fashion to that described above. Aquisition parameters: direct dimension was proton with a spectral width of 3 kHz, a digital resolution of 4096 complex points, and centred on the water resonance. Water presaturation was applied during the relaxation period. The

indirect dimension was carbon with a spectral width of 12 kHz. For each t_1 16 scans were recorded, and a total of 512 t_1 increments were recorded. Analysis parameters: The data was transformed using a sine-bell squared apodisation function shifted by $2\pi/5$. Strip transforms were performed after zero-filling to an array of 16384×4096, to obtain a separation of 0.18 Hz between data points in the direct (measured) dimension.

Molecular dynamics simulations

Simulations were performed using the general molecular modelling package CHARMM (Brooks et al., 1991). The most current carbohydrate force-field parameters suitable for CHARMM were used, with inclusion of terms to effectively model the exo-anomeric effect. A consistent set of partial charges for a mannose sugar in the $^4C_1$ configuration was obtained using a 6–31 g Hartree–Fock basis set computation. Each structure was initially relaxed in vacuum and then solvated in a rhomboidal dodecahedron of side 3.996 nm containing 1500 TIP3P water molecules (Jorgensen et al., 1983).

During simulation the covalent hydrogen bond lengths were kept constant with the SHAKE algorithm (van Gunsteren and Berendsen, 1977) and a 2 fs integration step was used. Long range electrostatic interactions were treated by using periodic-boundary conditions and a non-bonded interaction cut-off of 1.2 nm. The simulation was continued for a total of 50 ns, or 25 million steps, at a constant temperature of 300 K by weak coupling to a heat-bath. Co-ordinates were recorded every 0.2 ps for analysis purposes.

Simulating alignment

To simulate the alignment, it was assumed that alignment was caused by steric restriction at the phage surface, as described previously (Zweckstetter and Bax, 2000). Random rotation matrices R_i were created by a random rotation around the z-axis followed by transformation of the z-axis to a random position on the unit sphere. The molecular structure, after being translated to the centre of mass frame, $\left\{\mathbf{x}^{(1)}, \mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(n)}\right\}$, was rotated; Equation 7. Following rotation the atomic positions of the structure are $\left\{\mathbf{y}^{(1)}, \mathbf{y}^{(1)}, \ldots, \mathbf{y}^{(n)}\right\}$, for a particular rotation matrix R_i.

$$\mathbf{y}^{(r)} = R_i \mathbf{x}^{(r)}. \tag{7}$$

Rotation matrices were rejected if $\exists r$ such that $y_1^{(r)} > d$, where d is a random number $d \in [0, d_{\max}]$, which simulates steric hindrance, and d represents the distance of the molecule from the aligning surface. Here $d_{\max}$ is the maximum possible $y_1^{(r)}$ for any rotation matrix. This was continued for 250 000 selections of R and d for each frame in the simulation. If A is the set of the accepted rotation matrices during this process, then the expression for the Saupe order matrix is given by Equation 8, in anology to Equation 3. Here $\mathbf{e_i}$ are the cartesian unit vectors, and the magnetic field is taken to be parallel with $\mathbf{e_3}$. Thus S_{ij} could be calculated as a function of time during the simulation.

$$S_{ij} = \tfrac{1}{2}\left\langle 3\left(\mathbf{e_3} \cdot R_k \mathbf{e_i}\right)\left(\mathbf{e_3} \cdot R_k \mathbf{e_j}\right) - \delta_{ij}\right\rangle \quad \forall R_k \in A. \tag{8}$$

Predicting the scalar couplings

The scalar couplings were predicted from a previously developed equation (Haasnoot et al., 1980) which has been used previously to predict J-couplings in carbohydrates (Bock and Duus, 1994). The formulae are given in Equations 9 and 10, where ω_H is the dihedral angle between the relevant protons.

$$\begin{aligned} J_{56R} = {} & 13.22\cos^2\omega_H - 0.99\cos\omega_H \\ & -0.984\cos^2(0.139 + \omega_H) \\ & -6.396\cos^2(0.452 - \omega_H) + 2.61, \end{aligned} \tag{9}$$

$$\begin{aligned} J_{56S} = {} & 13.22\cos^2\omega_H - 0.99\cos\omega_H \\ & -0.984\cos^2(0.139 + \omega_H) \\ & -3.198\cos^2(0.452 + \omega_H) \\ & -3.198\cos^2(0.452 - \omega_H) + 2.61. \end{aligned} \tag{10}$$

These scalar couplings were calculated for every frame in the simulation and then averaged over relevant subsections of the trajectory.

Results

Molecular dynamics simulation

The simulation was continued for a total of 50 ns at a constant temperature of 300 K. The $\alpha(1 \rightarrow 3)$ linkage was observed to be dynamic, but restricted in conformational space to a single region. This is represented in Figure 4, where the exploration of this linkage during the simulation is plotted. This result was anticipated and formed part of the experimental design, and allowed the conformational analysis to concentrate on the $\alpha(1 \rightarrow 6)$ linkage. Therefore, little

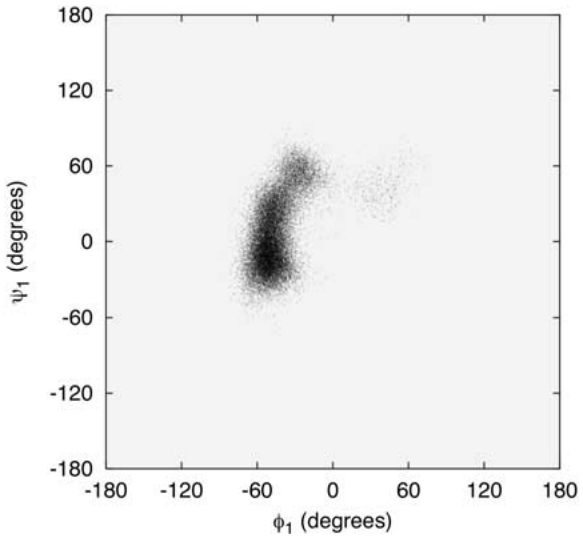

Figure 4. Exploration of the α (1 → 3) glycosidic linkage. The plot shows (ϕ_1, ψ_1), see Figure 2 for an explanation of these angular definitions.

Table 1. The measured scalar couplings J and residual dipolar couplings D in the mannose trisaccharide. The phage concentration was adjusted so that a deuterium splitting in water of 14.0 ± 0.2 Hz was observed

From	To	J (Hz)	$J + D$ (Hz)	D (Hz)
A C1	A H1	171.3 ± 0.2	179.2 ± 0.2	+7.9 ± 0.4
A C2	A H2	149.4 ± 0.2	143.3 ± 0.2	−6.1 ± 0.4
A C3	A H3	143.6 ± 0.2	147.7 ± 0.2	+4.1 ± 0.4
B C1	B H1	170.9 ± 0.2	175.0 ± 0.2	+4.1 ± 0.4
B C2	B H2	149.4 ± 0.2	149.0 ± 0.2	−0.4 ± 0.4
B C3	B H3	143.6 ± 0.2	143.2 ± 0.2	−0.4 ± 0.4
C C1	C H1	171.6 ± 0.2	175.0 ± 0.2	+3.4 ± 0.4
C C2	C H2	149.0 ± 0.2	152.8 ± 0.2	+3.8 ± 0.4
C C3	C H3	143.5 ± 0.2	139.2 ± 0.2	−4.3 ± 0.4
A C6	A H6R	145.7 ± 0.2	153.7 ± 0.2	+8.0 ± 0.4
A C6	A H6S	144.8 ± 0.2	146.7 ± 0.2	+1.9 ± 0.4
A H6R	A H6S	−11.6 ± 0.2	−5.9 ± 0.2	+5.7 ± 0.4
A H5	A H6R	4.3 ± 0.1	–	–
A H5	A H6S	1.9 ± 0.1	–	–

further reference is made to the α (1 → 3) linkage in this article.

At the α (1 → 6) linkage multiple distinct conformations were observed in the simulation. Previous work using NMR scalar couplings has indicated the presence of both *gg* and *gt* conformations at this linkage; the difference between them lies in the ω_2 angle (Figure 5). With the nomenclature used here (ω_2 = O6-C6-C5-O5) the *gg* conformation corresponds to $\omega_2 = -60°$ and *gt* to $\omega_2 = +60°$ (Jeffrey, 1990). No evidence was found during this simulation for a stable *tg* ($\omega_2 = \pm 180°$) conformation at the α(1 → 6) linkage. In a separate aqueous simulation the molecule was started in the *tg* conformation (data not shown here), and made the transition to the *gg* conformation after only 0.03 ns. It was therefore considered to be highly unstable.

In this simulation the α(1 → 6) linkage was started in the *gg* conformation. However, after a period of 8.5 ns a conformational change was noticed at the α(1 → 6) linkage. This was an exchange from the *gg* to *gt* conformation. Using the ω_2 angle this can be shown (Figure 6). During the following 41.5 ns no conformational exchange was observed back to the *gg* conformation. The dynamics of the α(1 → 3) linkage was found to be independent of the conformation at the α(1 → 6) linkage.

Nuclear magnetic resonance

One-bond carbon-proton couplings were measured using a spin-state-selective (S^3) pulse sequence and editing scheme (see Materials and methods). The couplings with and without phages are shown in Table 1. Using the S^3CT element and a specific editing scheme (see Materials and methods) allowed the measurement of C-H couplings in methylene groups. This was particularly important at the α(1 → 6) linkage.

Use of a COS^3 approach (see Materials and methods) allowed measurement of the geminal H-H coupling at the α(1 → 6) methylene even when phages were present, and hence the residual dipolar coupling between the two nuclei. In combination with the one-bond residual dipolar couplings this allows a good set of experimental data against which to compare the simulation (Table 1). The scalar couplings between H5 and H6R/S were measured from one-dimensional proton spectra in a sample without phages present.

Predicting the experimental observations

To predict the residual dipolar couplings it was necessary to consider the effect of alignment for all of the conformations present in the simulation. An alignment algorithm was constructed which could calculate the sterically induced Saupe order matrix for any conformation (see Materials and methods). This was used

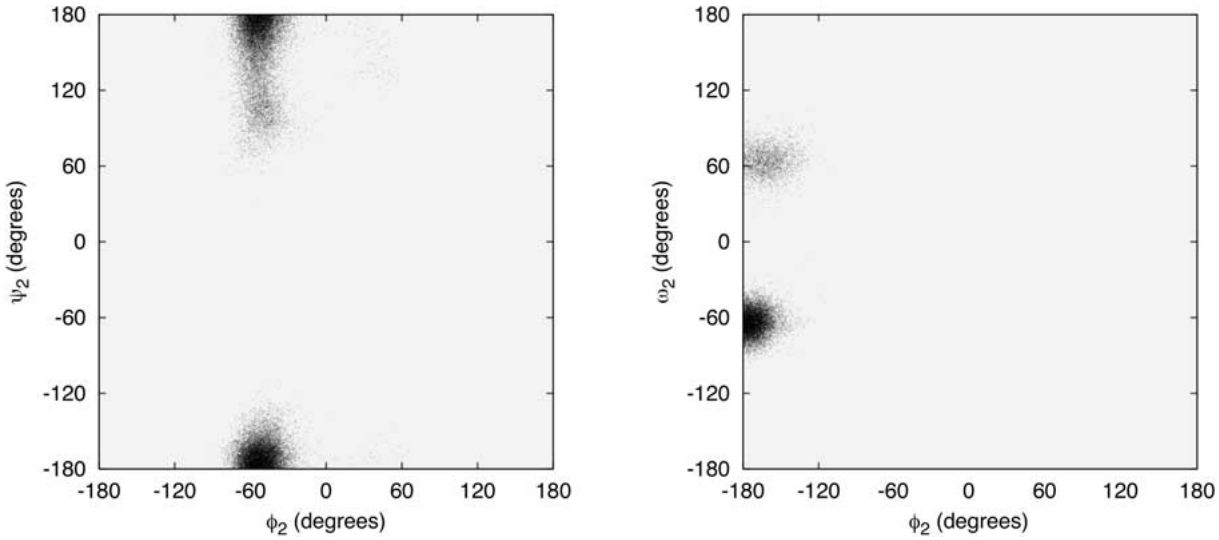

Figure 5. Exploration of the α (1 → 6) glycosidic linkage. On the left the angles (ϕ_2, ψ_2) are plotted and on the right (ϕ_2, ω_2) are plotted. See Figure 2 for the angular definitions.

Table 2. Predictions of the experimental data using different dynamic models. All values are expressed in Hz. The mixed model uses 55% *gg* and 45% *gt*. $\chi^2_{\min} = 83.6$ (10 d.f.)

From	To	*gg*	*gt*	*gg* + *gt*	Experimental
A C1	A H1	+5.9	+5.8	+6.9	+7.9 ± 0.4
A C2	A H2	−3.5	−7.5	−6.1	−6.1 ± 0.4
A C3	A H3	+6.1	+3.2	+5.7	+4.1 ± 0.4
B C1	B H1	+3.7	+3.9	+4.4	+4.1 ± 0.4
B C2	B H2	+1.0	−3.9	−1.3	−0.4 ± 0.4
B C3	B H3	0.0	0.0	0.0	−0.4 ± 0.4
C C1	C H1	+5.0	+0.5	+3.6	+3.4 ± 0.4
C C2	C H2	+2.9	+1.0	+2.5	+3.8 ± 0.4
C C3	C H3	−3.3	−6.5	−5.4	−4.3 ± 0.4
A C6	A H6R	+6.1	+5.3	+6.8	+8.0 ± 0.4
A C6	A H6S	+5.7	−1.0	+3.4	+1.9 ± 0.4
A H6R	A H6S	+5.0	+2.0	+4.3	+5.7 ± 0.4

to calculate the order matrix as a function of time during the simulation. It was found that in the diagonal frame of the inertia tensor the calculated order matrix was almost diagonal. The calculated diagonal components of the order matrix (S_{11},S_{22},S_{33}) in the diagonal frame of the inertia tensor as a function of time are presented in Figure 7. Using Equation 3 and S_{ij} the residual dipolar couplings for each frame in the simulation were calculated. It was then asked whether *gg*, *gt* or a mixture of the two gave the best fit to the residual dipolar couplings. So the average value of the D-values was calculated from the simulation for the *gg* and the *gt* portions. Let these be D_i^{gg} and D_i^{gt}, and let the proportion of *gt* be p. In the alignment algorithm the concentration of phages was not taken into account and our final predictions need to be scaled by a factor k to take into account the concentration of phages (c), and $k \propto c$. In that way, Equation 11 shows how the predicted value is dependent on p and k.

$$D_i^{pred}(k, p) = k\left\{pD_i^{gt} + (1-p)\,D_i^{gg}\right\}. \quad (11)$$

To estimate the most probable value of p a maximum likelihood function, Equation 12, was constructed based on the square distance of the meaured values from the predicted values. In this expression $D_i^{\exp}$ are the experimental data values and $\sigma^{\exp}$ are their respective errors.

$$\chi^2(k, p) = \sum_{i=1}^{N} \frac{\left\{D_i^{\text{pred}}(k, p) - D_i^{\exp}\right\}^2}{(\sigma^{\exp})^2}. \quad (12)$$

In Table 2 the best predictions for the pure dynamic *gg* ($p = 0$) and *gt* ($p = 1$) conformations are shown. The fitness function χ^2 was calculated as a function of p, and optimised against k, to see if a mixture of the two dynamic conformations allowed better prediction of the residual dipolar couplings. The minimum value of χ^2 as a function of p is shown in Figure 8. The optimal fit (minimum χ^2) to the experimental data occurs when $p = 0.45$ and this is significantly better than the

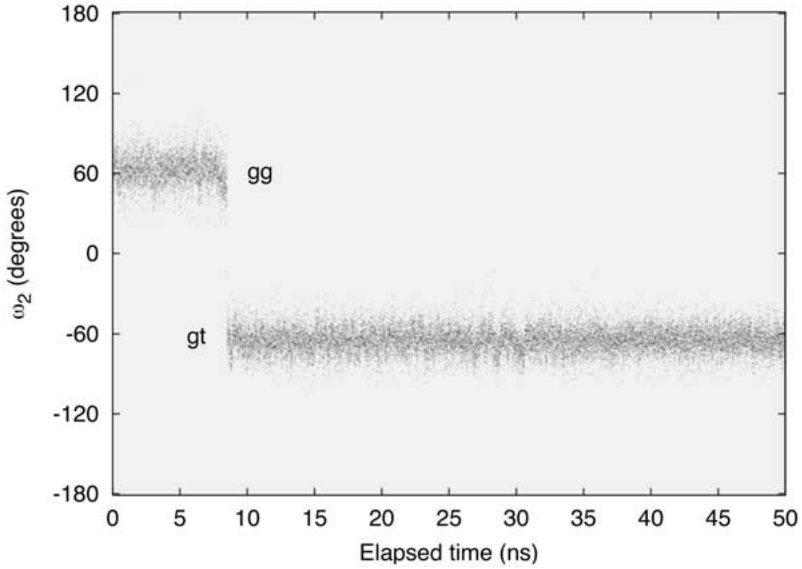

Figure 6. Conformational exchange from *gg* to *gt* conformation at the α(1 → 6) linkage of the oligosaccharide. In this plot the ω_2 angle (ω_2 = O6-C6-C5-O5) is shown as a function of elapsed time in the simulation.

Table 3. Predictions of scalar couplings at the α(1 → 6) linkage, using the simulation and an empirical relationship between *J* and relative proton dihedral angles (see Materials and methods for details). All values are expressed in Hz

Coupling	*gg*	*gt*	$gg + gt^a$	$gg + gt^b$	Experimental
J_{H5-H6R}	9.9	1.3	4.4	6.0	4.4 ± 0.1
J_{H5-H6S}	2.8	2.8	2.8	2.8	1.7 ± 0.1

[a]Best fit prediction to the scalar couplings using the mixed population 36% *gg* and 64% *gt*.
[b]Best fit prediction to the residual dipolar couplings using the mixed population 55% *gg* and 45% *gt*.

predictions using only one dynamical conformation. Thus from the residual dipolar coupling data a 55% *gg* and 45% *gt* conformational mix is predicted.

Another way to obtain information about the linkage is to use scalar couplings, and an empirical relation between vicinal scalar coupling and angle (Bock and Duus, 1994). Using Equations 9 and 10 in the Materials and methods section allowed predictions of the vicinal scalar couplings J_{H5-H6R}/J_{H5-H6S}. Both the pure *gg* and *gt* dynamic conformations were predicted by averaging relevant parts of the simulation. These averaged values are given in Table 3, alongside the experimental data. The best agreement to the experimental scalar couplings was obtained with a 36% *gg* and 64% *gt* conformer mix at the α(1 → 6) linkage. This could not be improved by including a *tg* conformer, in agreement with the simulation which suggested that the *tg* conformer is unstable in aqueous solution.

It can be seen that the conformer mix predicted by the residual dipolar couplings also provides predictions of scalar couplings in qualitative agreement with the experimental data, but the trend is for the predictions to overestimate the actual values. However, the predictions of scalar couplings are notoriously inaccurate due to errors in the empirical equations used to predict them. Further, it has been commented previously that these empirical equations can overestimate the experimental data when dynamics is taken into account (Bock and Duus, 1994).

Discussion

Microscopic interpretation of dynamics

In accordance with our previous work (Almond et al., 2001) molecular dynamics simulations in water have proven to be a useful technique for aiding the interpretation of experiments. Also, it has been found that predicting residual dipolar coupling induced by phages, using a theory based on steric alignment (Zweckstetter and Bax, 2000), is useful even for dynamic and anisotropic uncharged oligosaccharides. Underlying this is the fact that comparison with experimental data allows the prediction capacity of the simulations to be assessed.

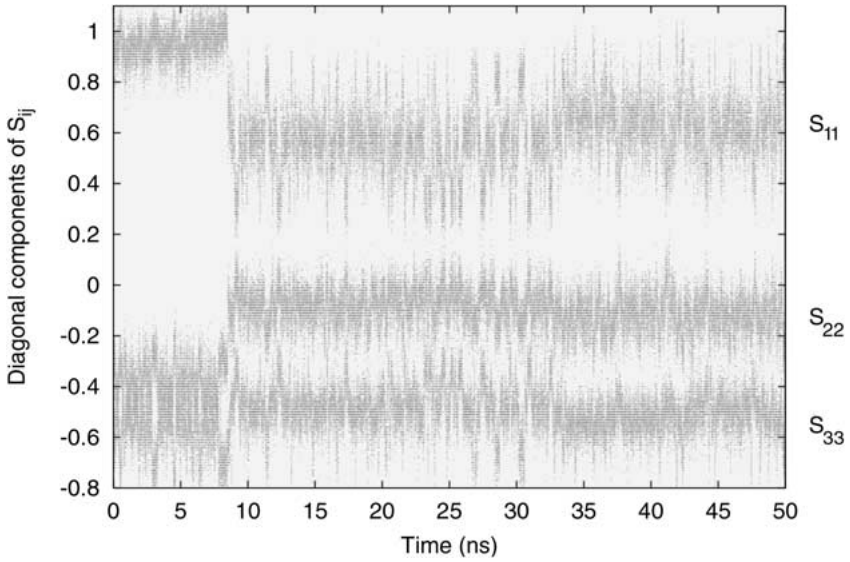

Figure 7. The calculated principal components (S_{11}, S_{22}, S_{33}) of the Saupe order tensor as a function of time in the simulation (see text for details).

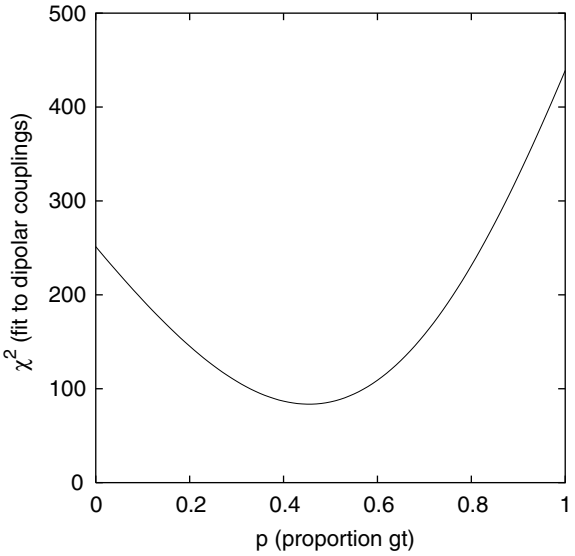

Figure 8. How fit to experimental data changes with proportion of *gg* to *gt*. The p value is the proportion of *gt* conformation, and hence the proportion of *gg* is $(1 - p)$. The predictions are made using Equation 11, and χ^2 is the maximum likelihood fit of the experimental data to the predictions, Equation 12. The minimum value of χ^2 occurs at $p = 0.45$, and $\chi^2_{\text{min}} = 83.6$ (10 d.f.).

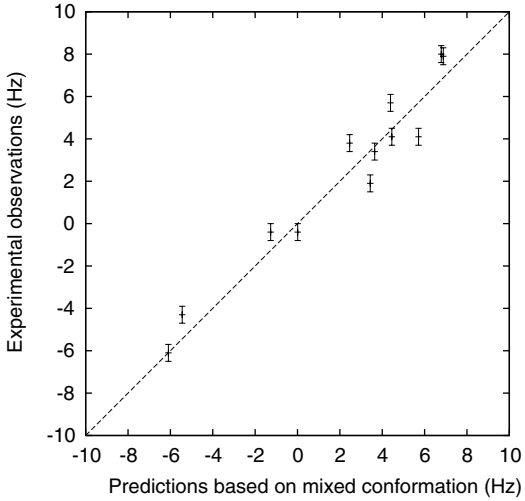

Figure 9. The fit of the experimental data to the computational predictions using a mix of 55% *gg* and 45% *gt*. Twelve experimental data points and predictions are shown in this plot. The straight line indicates where experimental data exactly agrees with theoretical predictions.

Analysis of a 50 ns trajectory in water and comparison with experimental data showed that the sampling of major molecular conformers was not correct. This sampling problem could be simply recorrected and then a satisfactory fit to the experimental data was obtained. Therefore, it is concluded that the local dynamics within particular major conformations is well represented by the simulation, but the population of major conformers is not.

The incorrect sampling of major conformers could be explained in a number of ways. One is that the force-field does not well represent the energy barriers for transition between major conformations, and a

second is that the simulation is not long enough to statistically sample these conformations. To correctly conclude one or the other would require more simulations and experiments, for example to investigate the rate of transition between conformers (Mäler et al., 1996). However, the preliminary data here is suggestive of problems with the molecular force-field, and this is not unexpected since dynamic molecular models have been largely untested by experimental data. Furthermore, it may be possible to use experimental data of this kind to provide feedback for systematic improvement of the force-fields (molecule and solvent) and their predictive capacity.

Considering the fact that the simulations provide a reasonable representation of the local dynamics allows the simulation to be used as a source of microscopic information about linkage dynamics which would otherwise not be available. This information is invaluable in generating preliminary hypotheses, and hence in experimental design. It is also of use in the drive to understand how these molecules interact with proteins and other macromolecules. Figure 10 shows an overlay of the two dynamic conformations, *gg* and *gt*. The figures show the amount of dynamic flexibility predicted at the linkages, as would be observed by an interacting receptor, for example. Since this oligosaccharide represents the core region of many N-linked glycans, it can be anticipated that much of their overall dynamics emerges from flexibility in the core, and particularly in biantennary structures. Using residual dipolar couplings allows an estimate to be placed on the populations of *gg* and *gt* rotamers at the $\alpha(1 \rightarrow 6)$ linkage. The experimental data suggests an approximate equal mix of the two, and aqueous molecular dynamics simulations allow the local flexibility of major conformers to be predicted. The good agreement with the experimental data indicates that no further major conformers are necessary to explain the observations. Hence, even though this linkage is flexible, it is likely that simulations in the range 100 ns can sample $\alpha(1 \rightarrow 6)$ linkage conformations correctly, with a refined potential energy model and that may include an improved model for water.

Comparison with previous work

The approach presented here extends previous similar work (Tian et al., 2001; Neubauer et al., 2001) by using spin-state-selective elements to avoid NMR resonance overlap, and aqueous molecular dynamics to aid interpretation of the experimental data. Also, the residual dipolar couplings measured here are in agreement with the primary data reported elsewhere on the same molecule (Tian et al., 2001), within experimental error.

In a previous article, also using residual dipolar couplings, it was concluded that the $\alpha(1 \rightarrow 3)$ linkage was "a good representation of a single preferred structure". Using the simulation and the overall agreement with experimental data allows a more specific conclusion to be made, namely that the $\alpha(1 \rightarrow 3)$ linkage is dynamic but restricted to a single conformational region, as described previously (Almond et al., 2001; Rutherford and Homans, 1994; Dowd and French, 1995; Homans et al., 1987). Based on this experimental data there is no reason to suggest otherwise. Hence the dynamic average, as observed by experiments, appears equivalent to, and can be interpreted as, a single structure. Other studies have interpreted two states in fast exchange at the $\alpha(1 \rightarrow 3)$ linkage (Brisson and Carver, 1983a; Al-Hashimi et al., 2000). More detailed analysis shows this to be the case, but because the exchange is fast and does not alter the overall molecular shape, the linkage was considered as a single dynamic conformation, for convenience. Hence, the residual dipolar couplings are directly sensitive to large molecular motion at the $\alpha(1 \rightarrow 6)$ linkage.

In a previous article (Tian et al., 2001) it was reported that the "motional amplitudes are large" at the $\alpha(1 \rightarrow 6)$ linkage. The simulation presented here predicts that this linkage has significant local dynamics and two major conformations. However, the interpretation of the residual dipolar couplings in Tian et al. (2001) required knowledge of molecular geometry and assumed that the fragment is rigid or exhibits small local librations. The simulation here shows that the residual dipolar couplings can be interpreted by a highly flexible molecule.

In contrast to previous approaches (Tian et al., 2001; Neubauer et al., 2001) the methodology presented here does not rely on back calculation of order matrices from the experimental data using a static molecular model. As can be seen in Figure 10 exchange between conformers can dramatically affect the overall molecular shape, and hence the calculated order matrix (Figure 7). In the *gt* conformation there is an average reduction in the principal component of the order matrix and an increase in the asymmetry, compared with *gg*. This is apparent in Figure 10, and local dynamics in the glycosidic linkages were seen to contribute to fluctuations in the order tensor, which

Figure 10. The two dynamic conformations of the oligosaccharide. On the left is the *gg* conformation of the α(1 → 6) linkage and on the right the *gt*. In this diagram only the ring carbons, oxygens and the glycosidic oxygens have been drawn. They have been overlayed by best-fitting to the ring of sugar A.

take place on the sub-nanosecond timescale. Although an average order matrix can always be calculated, it would appear important to retain the relative sizes of the order matrix in different conformations to ensure the correct conformational weighting is effected.

For normal *D*-hexoses it has been found that the relative orientation of the substituent at C4, adjacent to the α(1 → 6) linkage, influences its conformer preference. In sugars with *trans* hydroxyl at C4, for example glucose and mannose, approximate populations of *gg*:*gt*:*tg* ≈ 6:4:0 have been reported. This is based on previous NMR studies (Nishida et al., 1988), statistical analysis of X-ray crystal structures (Marchessault and Peréz, 1979) and fluorescence energy transfer (Rice et al., 1993). Other approaches have used maximum entropy along with scalar couplings and NMR relaxation to obtain the conformer preference in a β(1 → 6) linked disaccharide (Poppe, 1993), and again a combination of *gg* and *gt* conformations was predicted at the linkage. All of this data supports the claim here that the α(1 → 6) linkage in this molecule is a mixture of conformations.

More specifically, in studies of the oligosaccharides containing the core region of N-linked glycans the α(1 → 6) linkage has been proposed to take on a mixture of *gg* and *gt* conformations. However, dynamic analysis of such large molecules by current experimental techniques is prone to errors. Subsequently, the results were found to depend on the particular technique employed. There are examples of studies showing preference for *gg* (Berman, 1987), *gt* (Cumming and Carver, 1987; Paulsen et al., 1987) or an equal mixture (Brisson and Carver, 1983b; Homans et al., 1986). Within the limitations of the particular experiments used the agreement with the conclusions presented here is satisfactory. It appears that residual dipolar couplings combined with a dynamic model promises to be a more quantitative way to obtain this information. Further, although a trisaccharide was the subject of the present study there is no reason why this could not be extended to larger oligosaccharides if more experimental data is forthcoming from them.

Conclusions

Residual dipolar couplings have been measured in a trimannose structure containing an α(1 → 6) linkage using a spin-state-selective approach, which offers the advantage of reducing resonance overlap and increasing precision. The oligosaccharide represents the carbohydrate core region in many N-linked glycans.

The approach presented here tests a 50 ns molecular dynamics simulation in water against the experimental observations. By assuming that alignment by phages is induced by sterics it was possible to calculate the order matrix for all conformations in the simulation. It was found that the simulation could not correctly predict the conformational mixture at the α(1 → 6) linkage. Using a maximum likelihood approach it was possible to re-correct for sampling of major conformers, since local flexibility appears to be well represented. With this approach an almost equal population of *gg* and *gt* rotamers was predicted at this

linkage, while the $\alpha(1 \rightarrow 3)$ linkage was restricted to a single conformational region. The conformer predictions are in agreement with those previously made on the basis of NMR in dilute solution, optical techniques and X-ray crystallography.

This approach offers promise for correcting the output from molecular dynamics simulations and therefore provide feedback for systematic improvement of force-fields and their future reliability, since it has the ability to provide quantitative rotamer preference at such flexible linkages. Although this approach was restricted to a trisaccharide, there is no reason why it cannot be applied to larger uncharged complex oligosaccharides.

Acknowledgements

We wish to thank the Danish Instrument Centre for NMR Spectroscopy of Biological Macromolecules for providing time on the 800 MHz Varian Unity Inova NMR spectrometer. Bent O. Petersen is thanked for confirming the measurement of the residual dipolar couplings. Jakob Bunkenborg is also thanked for help in implementing the NMR pulse sequences and editing schemes. A.A. is funded by the Wellcome Trust (U.K.) on a 'Prize Travelling Research Fellowship', grant reference number 058154.

References

Al-Hashimi, H.M., Valafar, H., Terrell, M., Zartler, E.R., Eidsness, M.K. and Prestegard, J.H. (2000) *J. Magn. Reson.*, **143**, 402–406.

Almond, A., Brass, A. and Sheehan, J.K. (1998) *J. Mol. Biol.*, **284**, 1425–1437.

Almond, A., Brass, A. and Sheehan, J.K. (2000) *J. Phys. Chem.*, **B104**, 5634–5640.

Almond, A., Bunkenborg, J., Franch, T., Gotfredsen, C.H. and Duus, J.Ø. (2001) *J. Am. Chem. Soc.*, **123**, 4792–4802.

Almond, A. and Sheehan, J.K. (2000) *Glycobiology*, **10**, 329–338.

Andersson, P., Weigelt, J. and Otting, G. (1998) *J. Biomol. NMR*, **12**, 435–441.

Berman, E. (1987) *Eur. J. Biochem.*, **165**, 385–391.

Bock, K. and Duus, J.Ø. (1994) *J. Carbohydrate Chem.*, **13**, 513–543.

Brisson, J.R. and Carver, J.P. (1983a) *Biochemistry*, **22**, 1362–1368.

Brisson, J.R. and Carver, J.P. (1983b) *Biochemistry*, **22**, 3680–3686.

Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M. (1991) *J. Comput. Chem.*, **12**, 292–300.

Cumming, D.A. and Carver, J.P. (1987) *Biochemistry*, **26**, 6676–6683.

Dowd, M.K. and French, A.D. (1995) *J. Carbohydrate Chem.*, **14**, 589–600.

Dwek, R.A. (1996) *Chem. Rev.*, **96**, 683–720.

Engelsen, S.B. and Rasmussen, K. (1993) *Int. J. Biol. Macromol.*, **15**, 56–62.

Haasnoot, C.A.G., de Leeuw, F.A.A.M. and Altona, C. (1980) *Tetrahedron*, **36**, 2783–2792.

Hansen, M.R., Mueller, L. and Pardi, A. (1998) *Nat. Struct. Biol.*, **5**, 1065–1074.

Homans, S.W. (1993) *Glycobiology*, **3**, 551–555.

Homans, S.W., Dwek, R.A., Boyd, J., Mahmoudian, J., Richards, W.G. and Rademacher, T.W. (1986) *Biochemistry*, **25**, 6342–6350.

Homans, S.W., Pastore, A., Dwek, R.A. and Rademacher, T.W. (1987) *Biochemistry*, **26**, 6649–6655.

Imberty, A., Peréz, S., Hricovíni, M., Shah, R.N. and Carver, J.P. (1993) *Int. J. Biol. Macromol.*, **15**, 17–23.

Jansson, P., Kenne, L. and Kolare, I. (1994) *Carbohydrate Res.*, **257**, 163–174.

Jeffrey, G.A. (1990) *Acta Crystallogr.*, **B46**, 89–103.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1983) *J. Chem. Phys.*, **79**, 926–935.

Lerche, M.H., Meissner, A., Poulsen, F.M. and Sørensen, O.W. (1999) *J. Magn. Reson.*, **140**, 259–263.

Mäler, L., Widmailm, G. and Kowalewshi, J. (1996) *J. Biomol. NMR*, **7**, 1–7.

Marchessault, R.H. and Peréz, S. (1979) *Biopolymers*, **18**, 2369–2374.

Martin-Pastor, M. and Bush, C.A. (2000a) *Carbohydrate Res.*, **323**, 147–155.

Martin-Pastor, M. and Bush, C.A. (2000b) *Biochemistry*, **39**, 4674–4683.

Meissner, A., Schulte-Herbrüggen, T. and Sørensen, O.W. (1998) *J. Am. Chem. Soc.*, **120**, 3803–3804.

Neubauer, H., Meiler, J., Peti, W. and Griesinger, C. (2001) *Helv. Chim. Acta*, **84**, 243–258.

Nishida, Y., Hori, H., Ohrui, H. and Meguro, H. (1988) *J. Carbohydr. Chem.*, **7**, 239–250.

Ottiger, M., Delaglio, F. and Bax, A. (1998) *J. Magn. Reson.*, **131**, 373–378.

Paulsen, H., Peters, T. and Sinnwell, V. (1987) *Carbohydr. Res.*, **165**, 251–266.

Poppe, L. (1993) *J. Am. Chem. Soc.*, **115**, 8421–8426.

Prestegard, J.H., Al-Hashimi, H.M. and Tolman, J.R. (2000) *Quart. Rev. Biophys.*, **33**, 371–424.

Quasba, P.K., Balaji, P.V. and Rao, V.S.R. (1994) *Glycobiology*, **4**, 805–815.

Rice, K.G., Wu, P., Brand, L. and Lee, Y.C. (1993) *Biochemistry*, **32**, 7264–7270.

Rudd, P.M. and Dwek, R.A. (1997) *Crit. Rev. Biochem. Mol. Biol.*, **32**, 1–100.

Rudd, P.M., Endo, T., Colominas, C., Groth, D., Wheeler, S.F., Harvey, D.J., Wormald, M.R., Serban, H., Prusiner, S.B., Kobata, A. and Dwek, R.A. (1999) *Proc. Natl. Acad. Sci. USA*, **96**, 13044–13049.

Rutherford, T.J. and Homans, S.W. (1994) *Biochemistry*, **33**, 9606–9614.

Saupe, A. and Englert, G. (1963) *Phys. Rev. Lett.*, **11**, 462–465.

Sayers, E.W. and Prestegard, J.H. (2000) *Biophys. J.*, **79**, 3313–3329.

Sørensen, M.D., Meissner, A. and Sørensen, O.W. (1997) *J. Biomol. NMR*, **10**, 181–186.

Spronk, B.A., Rivera-Sagredo, A., Kamerling, J.P. and Vliegenthart, J.F.G. (1995) *Carbohydrate Res.*, **273**, 11–26.

Tian, F., Al-Hashimi, H.M., Craighead, J.L. and Prestegard, J.H. (2001) *J. Am. Chem. Soc.*, **123**, 485–492.

Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1113.
Tvaroska, I., Imberty, A. and Peréz, S. (1990) *Biopolymers*, **30**, 369–379.
Untidt, T.S., Schulte-Herbrüggen, T., Sørensen, O.W. and Nielsen, N. (1999) *J. Phys. Chem.*, **A103**, 8921–8926.
van Gunsteren, W.F. and Berendsen, H.J.C. (1977) *Mol. Phys.*, **34**, 1311–1327.
Zweckstetter, M. and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 3791–3792.