Conformational Studies of Human Milk Oligosaccharides Using $^{1}H-^{13}C$ One-Bond NMR Residual Dipolar Couplings†

Manuel Martin-Pastor and C. Allen Bush*

Department of Chemistry and Biochemistry, University of Maryland−Baltimore County, Baltimore, Maryland 21250

Received September 1, 1999; Revised Manuscript Received January 18, 2000

ABSTRACT: $^{1}H-^{13}C$ one-bond dipolar coupling values were measured for natural abundance samples of the human milk oligosaccharides "lacto-*N*-fucopentaose" (LNF-1 LNF-2, and LNF-3), "lacto-*N*-difucohexaose" (LND-1), "lacto-*N*-tetraose" (LNT), and "lacto-*N*-*neo*-tetraose" (LNnT), four of which have Lewis blood group epitopes. Each oligosaccharide was dissolved in a 7.5% solution of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine/1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DMPC/DHPC) bicelle liquid crystals oriented in the NMR magnetic field. The dipolar coupling data and NOE were fitted to conformational models with calculations of an optimum orientation tensor which best represents the dipolar coupling values for a fragment hypothesized to adopt a single conformation. In the case of LNF-1, LNF-2, LNF-3, and LND-1, the models confirm previous conformational models for the Lewis epitopes based on NOE and molecular dynamics simulations. Extensions of the model provided new structural data for the remaining residues. In all cases, upper limits for the errors in the glycosidic angles of the models were estimated. Since residual dipolar coupling provides information on long-range order, it is a valuable complement to other types of NMR data such as NOE and scalar coupling for exploring conformations of complex oligosaccharides.

The solution conformation of a complex biologically active oligosaccharide is an essential reference state for calculation of the binding free energy to antibodies, lectins, glycosidases, and glycosyl transferases. Proper evaluation of the conformation requires a scheme for understanding the influence of the conformational flexibility. We have proposed that internal motions of oligosaccharides be classified as two kinds, the first of which may be envisioned as a single conformation with modest excursions of the glycosidic dihedral angles of the order of ±15° and of an approximately symmetric type (*1*). The second kind of internal motion features larger excursions of glycosidic dihedral angles between distinct energy minima which may be separated by energy barriers. For oligosaccharides exhibiting internal motion of the first kind, NOE[1] data can generally be interpreted by a single conformational model whereas for more flexible oligosaccharides exhibiting internal motion of the second kind, NOE cross-peaks may be found which cannot be simultaneously reconciled with any single conformation.

Suitable mixtures of lipids form bicelles which can induce weak orientation of proteins following an approach introduced by Tjandra and Bax (*2*). Several research groups have shown that residual dipolar couplings can be measured in oligosaccharides, which are weakly oriented in these liquid-crystalline media. Bolon and Prestegard (*3*) have demonstrated H−H dipolar coupling, and Rundlöf et al. (*4*) and Kiddle and Homans (*5*) have measured C−H dipolar coupling in oligosaccharides. The method appears to be extremely promising both for the solution conformation and for conformations of protein-bound oligosaccharides (*6*).

While interpretation of residual dipolar coupling data in highly flexible oligosaccharides exhibiting internal motion of the second kind could be difficult and uncertain, interpretation should be especially straightforward in the relatively less flexible oligosaccharides. The Lewis blood group epitopes, which are held by a broad consensus of research groups to be well represented by single well-defined conformations, should be a good test for interpretation of dipolar coupling data. If they can be represented by single conformation models and given an adequate number and quality of dipolar coupling values, it should be possible to calculate an orientation tensor for the Lewis epitope which explains the dipolar coupling data (*7*). It has been shown that multiple solutions for the orientation tensor may be found which are consistent within experimental error (*8*). In this paper, we present evidence that at least for these single conformations a limited region of conformational space can be identified and ambiguities can be resolved with the addition of other sources of structural data such as NOE and molecular mechanics calculations. Models deduced are consistent with widely accepted models for Lewis epitopes, and the models were extended to include additional sugar residues of the human milk oligosaccharides.

† Research supported by NSF Grant MCB9724133 and by a Ministerio de Educacion y Cultura of Spain postdoctoral fellowship to M.M.-P.

* Corresponding author. Phone: (410) 455-2506, fax: (410) 455-2608, e-mail: bush@umbc.edu.

[1] Abbreviations: NOE, nuclear Overhauser effect; DMPC, 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine; DHPC, 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine; "lacto-*N*-fucopentaose", LNF-1, LNF-2, LNF-3; "lacto-*N*-difucohexaose", LND-1; "lacto-*N*-tetraose", LNT; "lacto-*N*-*neo*-tetraose", LNnT; g-HSQC, gradient heteronuclear single quantum coherence experiment; NOESY, nuclear Overhauser experiment; MD, molecular dynamics simulation; CVFF, covalent valence force field.

10.1021/bi992050q CCC: $19.00
Published on Web 03/30/2000

EXPERIMENTAL SECTION

Sample Preparation. A ~15% bicelle stock solution in molar ratio 3:1 of 1,2-dimyristoyl-*sn*-glycero-3-phosphocholine (DMPC) and 1,2-dihexanoyl-*sn*-glycero-3-phosphocholine (DHPC) in D_2O was prepared. A 30 mM phosphate buffer at pH 7 was added, and the solution was mixed by vortexing in several cycles of heating to 40 °C and cooling on ice over a period of 20 h.

Of the human milk oligosaccharides, whose structures are shown in Figure 1, LNT, LNnT, LNF-1, and LND-1 were purchased from V-Labs (Covington, LA). LNF-3 was purchased from Oxford Glycosciences (Oxford, U.K.), and LNF-2 was isolated as previously described (*9*). Each oligosaccharide was exchanged in ^2H_2O and lyophilized for three cycles. A sample, ~10 mM, of each oligosaccharide in ~7.5% of the bicelle medium was prepared by mixing a 0.25 mL solution of the oligosaccharide in D_2O with 0.25 mL of the stock bicelle solution. A sample of LNT, ~10 mM, and ~15% of the bicelle medium was also prepared by dissolving the oligosaccharide in 0.25 mL of the bicelle medium. Each sample was homogenized by vortexing in several cycles of heating to 40 °C and cooling on ice.

NMR Spectroscopy. NMR data were acquired on a GE-Omega PSG 500 NMR instrument and processed on a Silicon Graphics workstation using Felix 2.3 software (Biosym Technologies, San Diego, CA). 2D-NOESY spectra were acquired for the oligosaccharides LNT, LNF-1, and LND-1 in D_2O at 5 °C using a mixing time of 250 ms. A total of 512 × 1024 complex data points were acquired for the t_1 and t_2 dimensions, respectively. Both dimensions were apodized with a 90°-shifted sine-bell function and zero-filled before being Fourier-transformed. The final 2D matrix size for each spectrum was 1024 × 1024 real points. 2D-NOESY data for LNF-2 and LNF-3 were taken from data sets previously recorded (*10*, *11*). After assignment of NOE cross-peaks, the intensities were classified as strong (s), medium (m), or weak (w).

Residual dipolar coupling values, $^1D_{CH}$, were measured for each sample of oligosaccharide dissolved in the bicelles from the difference in the splitting of two t_1-coupled, gradient HSQC spectra (t_1-coupled g-HSQC) (*12*), acquired at 35 and 19 °C, corresponding to the oriented and isotropic solution, respectively. The ^1H and ^{13}C chemical shift assignments were taken from references (*10*) and (*13*−*15*) and were very similar to those obtained for the oligosaccharides in the bicellar medium. The only discrepancy in chemical shifts is an apparent typographical error in the reported value for proton H5a of LN-neo-T (*15*) which is 3.72 ppm instead of 3.94 ppm. To maximize the resolution in the indirect dimension, two sets of t_1-folded experiments were acquired, one with ^{13}C carrier position in the center of the anomeric region (~98 ppm) and the other in the center of the ring region (~75 ppm). For the former set of experiments, the spectral width was 3200 Hz, and the INEPT delays corresponding to a nominal value of $^1J_{CH}$ were set to 165 Hz. For the other set of experiments, the spectral width was 4300 Hz, and the INEPT delays were set to a nominal value of 155 Hz. Carbon decoupling during acquisition was carried out with WALTZ-16 with a field strength of 1612.9 Hz. Two sine-shaped gradients were applied during 2.5 and 1 ms which act as a zz-filter, purging the HSQC spectrum (*12*). For each experiment, a 2D matrix of 1024 × 256 complex FID data points was acquired for the t_1 and t_2 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 1024 × 1024 real points. The column corresponding to each C−H signal was carefully phased and stored. After an inverse Hilbert transformation and zero-filling to 16K, the retransformed vector showed a digital resolution of 0.2 Hz/point. Vectors from the 35 °C spectrum were compared with the corresponding vector from the 19 °C spectrum, and the offset required for superposition of the multiplet components was used to calculate the sign and magnitude of $^1D_{CH}$. The experiments were repeated twice and $^1D_{CH}$ couplings averaged; the estimated error is ±0.8 Hz.

Dipolar coupling values were interpreted using the formula given by Tjandra and Bax (*2*). The orientation tensor was determined by iterative fitting of the measured dipolar coupling values for a fixed structural model (*7*). In the algorithm used for the fitting (*16*), the direction cosines of the bond vectors are calculated from the Cartesian coordinates of the model molecular structure. A rigid model with at least five nonparallel C−H bond vectors is required to uniquely determine the orientation (three Euler angles) as well as the magnitude (A_a) and rhombicity (A_r) of the alignment tensor. The Lewis epitopes of the milk oligosaccharides were used for that purpose in our analysis. Initial trial values of the Euler angles (θ, ϕ, ψ) were selected, and the goodness of the parameters was evaluated by the merit function χ^2 defined in reference (*7*).

Molecular Modeling. Molecular modeling was used both to refine the initial model of the Lewis epitope and to extend the model to include additional sugar residues. In the molecular modeling, the glycosidic dihedral angles are defined according to the IUPAC heavy-atom convention in which ϕ is defined by $O_5-C_1-O_1-C_x$ and ψ by $C_{x-1}-C_x-O_1-C_1$. Energy-relaxed conformational maps were generated for the glycosidic linkages of the oligosaccharides studied using InsightII/Discover software (Biosym Technologies, San Diego, CA). For the relaxed energy maps, the glycosidic dihedral angles were restrained by a cosine type potential with a force constant of 100 kcal mol^{-1} on the glycosidic dihedral angle, and the remainder of the molecule was minimized using the CVFF (*17*) force field and a distance-dependent dielectric of 80**r*. Relaxed grid maps were generated by scanning the glycosidic torsions ϕ and ψ using a grid step of 2.5° for each linkage of each oligosaccharide.

For grid mapping of the Lewis epitope portion of the oligosaccharide, two pairs of dihedral angles were scanned in a four-dimensional search specifying the geometry of three sugar residues. To save computational time, any available interresidual NOE data were used to restrict the four-dimensional grid search. For this purpose, upper distance limits of 3.5 and 4.0 Å for strong and medium NOEs, respectively, were used for considering a given conformation prior to minimization. The energy-minimized structures from this search were then submitted to the dipolar coupling calculation in which the alignment tensor optimization protocol included only the dipolar coupling values for the three residues. For LND-1, whose Lewis^b epitope contains four residues, the three residues of the Lewis^a epitope were selected for the four-dimensional grid mapping described above.

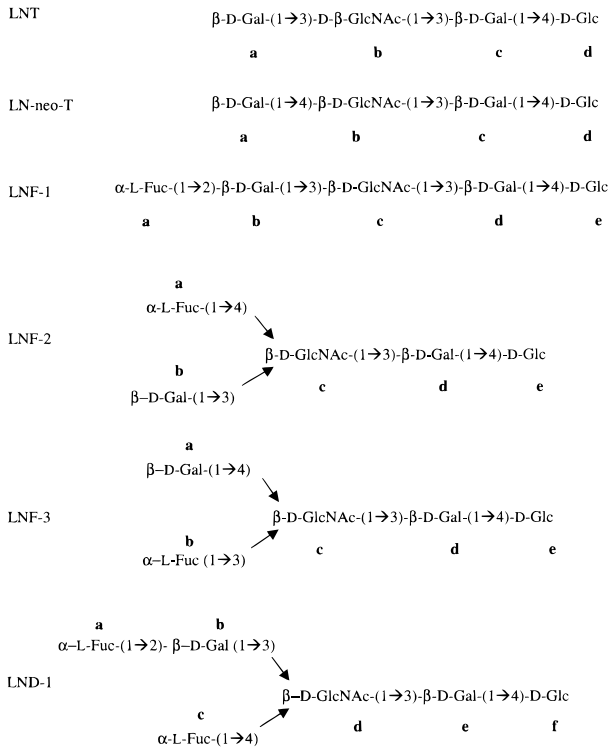

FIGURE 1: Schematic representation of the oligosaccharides studied and the corresponding label used to name each residue.

Once an energy-minimized model consistent with the experimental data for the Lewis epitope was determined, then additional residues were added to the model for independent grid searches from −180° to 180° over the ϕ and ψ dihedral angles joining the newly added residue to the initial core whose conformation is held fixed in the previously determined conformation. To the Lewis[a] core of LND-1 is first added the α-Fuc-(1→2) residue to form the Lewis[b] epitope.

To each of the Lewis epitopes was added the Gal residue which is linked to GlcNAc by a β-(1→3) linkage. A full grid search over the latter linkage was performed, and each

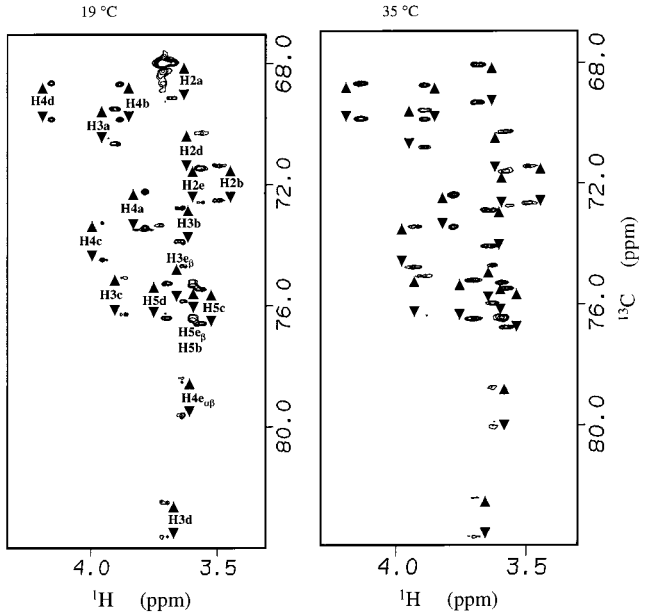

FIGURE 2: t_1-coupled g-HSQC of LNF-3 acquired isotropic at 19 °C and oriented in the liquid-crystal phase at 35 °C.

energy-minimized structure was submitted to the alignment tensor optimization protocol. Those structures which can be aligned with low values of χ^2 are judged to be consistent with the dipolar coupling data, and the lowest energy conformers up to 4 kcal/mol above the lowest energy conformer consistent with NOE data can be considered as candidate structures. The same procedure was repeated for mapping the terminal β-Gal-(1→4)-Glc linkage.

An estimate of the uncertainties in each glycosidic dihedral angle was done by considering all the different low-energy structures obtained during the grid searches which are consistent with the experimental data within a χ^2 window which gives mean deviations of the simulated dipolar couplings ≤0.8 Hz. The corresponding allowed conformations were represented as points in a ϕ/ψ dipolar map.

Table 1: Experimental Dipolar Couplings and Those Calculated with the Oriented Models for Tetrasaccharides LNT and LN-neo-T

	LNT[a] $^1D_{CH}$calc	LNT[a] $^1D_{CH}$exp	LNT[b] $^1D_{CH}$calc	LNT[b] $^1D_{CH}$exp	LN-neo-T $^1D_{CH}$calc	LN-neo-T $^1D_{CH}$exp
H1−C1a	8.4	8.6				
H2−C2a			28.1	27.6	23.1	22.8
H3−C3a	9.7	10.0	28.0	30.9	18.3	18.6
H4−C4a	−2.9	−2.7	−9.5	−11.2	−21.8	−21.9
H5−C5a	9.4	9.1	24.2	24.2	21.1	21.0
H1−C1b	18.7	19.2				
H2−C2b	18.9	19.0				
H3−C3b	19.2	18.5			29.4	29.4
H4−C4b	18.8	17.7	48.0	47.0	26.4	26.6
H5−C5b	18.8	19.8	48.8	50.0	26.6	26.6
H1−C1c	15.0	16.6				
H3−C3c	16.5	16.4				
H4−C4c	2.0	2.2	3.0	1.2	−9.2	−9.2
H5−C5c	15.2	14.5	43.5	42.2	23.4	23.3
H3−C3d	9.7	9.7	21.2	24.3	18.4	18.6
H4−C4d					17.0	16.7
H5−C5d	11.3	11.6	25.2	23.0	15.9	15.8
av dev (Hz)	0.3		1.4		0.14	
χ^2	0.291		2.292		0.021	
R^2	0.998		0.996		0.999	

[a] Bicelles concentrated 7.5%. [b] Bicelles concentrated 15%.

Table 2: Experimental Dipolar Couplings and Those Calculated with the Oriented Model for Pentasaccharides LNF-1, LNF-2, and LNF-3

atom pairs	LNF-1 $^1D_{CH}$calc	LNF-1 $^1D_{CH}$exp	LNF-2 $^1D_{CH}$calc	LNF-2 $^1D_{CH}$exp	LNF-3 $^1D_{CH}$calc	LNF-3 $^1D_{CH}$exp
H1−C1a	−17.5	−18.4	2.0	1.8		
H2−C2a	−6.4	−6.7	−2.1	−1.3	10.4	10.8
H3−C3a	−4.1	−5.5	−0.5	−0.8	9.4	9.3
H4−C4a	−16.9	−16.2	0.2	0.3	−17.4	−17.2
H5−C5a	−3.7	−3.2				
H1−C1b			−5.6	−5.5		
H2−C2b	5.5	5.9	−3.6	−3.2	10.3	11.1
H3−C3b	6.4	5.5	−5.0	−4.8	10.4	10.0
H4−C4b	9.4	9.6	6.0	5.8	−6.2	−5.6
H5−C5b	4.5	6.0	−5.1	−5.8		
H2−C2c	9.9	12.1	11.7	12.1	17.3	17.2
H3−C3c	11.5	10.6	12.3	11.9	16.9	17.2
H4−C4c	9.1	8.7	11.1	10.3	16.8	15.6
H5−C5c	8.9	9.2	11.3	11.5	17.0	17.4
H1−C1d	9.6	10.2	13.1	12.5		
H2−C2d	10.8	10.5	13.1	13.7	17.7	18.5
H3−C3d			13.5	15.3		
H4−C4d	7.3	7.5	−2.6	−2.6	−4.0	−4.2
H5−C5d	9.8	9.9	13.3	12.4	17.1	17.4
H3−C3e	5.1	4.9	6.3	6.3	9.8	10.8
H4−C4e	5.6	6.7	8.3	8.2	10.1	9.2
H5−C5e	5.6	5.0			10.0	10.0
av dev (Hz)	0.7		0.4		0.5	
χ^2	0.514		0.318		0.251	
R^2	0.995		0.996		0.998	

Table 3: Experimental Dipolar Couplings and Those Calculated with the Oriented Model for Hexasaccharide LND-1

atom pairs	LND-1 $^1D_{CH}$calc	LND-1 $^1D_{CH}$exp
H1−C1a	−26.6	−26.4
H2−C2a	−3.6	−2.3
H3−C3a	−1.3	−1.4
H4−C4a	−26.3	−26.5
H5−C5a	−1.1	−0.9
H2−C2b	−3.0	−2.8
H3−C3b	−2.8	−3.1
H4−C4b	17.7	19.0
H5−C5b	−4.5	−4.5
H1−C1c	10.6	9.1
H2−C2c	6.5	7.0
H4−C4c	6.4	6.7
H3−C3d	19.1	19.2
H4−C4d	18.4	16.3
H5−C5d	18.6	18.8
H1−C1e	16.7	15.2
H2−C2e	17.9	20.0
H4−C4e	0.1	0.9
H5−C5e	16.7	16.3
H3−C3f	13.2	13.4
av dev (Hz)	0.7	
χ^2	0.649	
R^2	0.997	

RESULTS AND DISCUSSION

^{1}H−^{13}C one-bond residual dipolar couplings were measured in dilute bicelle media composed of DMPC/DHPC 3:1 (*18*) for the six oligosaccharides of Figure 1. Dipolar couplings were determined from the splittings in the ^{13}C dimension of a t_1-coupled g-HSQC experiment (*12*, *19*) as shown in Figure 2 for the ring region of LNF-3. The experimental dipolar couplings obtained for the different oligosaccharides are given in Tables 1−3.

At ∼7.5% bicelle concentration, the minimum and maximum values for the experimental dipolar couplings vary for the different oligosaccharides ranging from those for LNT (−2.7 and 19.8 Hz), for LN-Neo-T (−21.9 and 29.4 Hz), for LNF-1 (−18.4 and 12.1 Hz), for LNF-2 (−5.5 and 15.3 Hz), for LNF-3 (−17.2 and 18.5 Hz), and for LND-1 (−26.5 and 20.0 Hz). These values may reflect either differences in the orientation of the alignment tensor, the degree of alignment, or the fact that the orientation of the vectors sampled is not isotropically distributed in space. The number of C−H vectors considered is small, and many of them have nearly parallel orientations. This situation differs from that of the N−H vectors in a protein which are both more numerous and more randomly oriented, a fact which has been exploited to estimate the components A_a and A_r of the alignment tensor from the maximum, minimum, and most frequent value of the set of dipolar couplings (*8*, *20*).

For LNT, the dipolar couplings were measured both at ∼7.5% and at ∼15% bicelle concentration (Table 1). It can be seen that at the higher concentration the dipolar couplings are approximately 2 or 3 times larger than those at the low concentration with the ratio of the values remaining similar. This observation is consistent with a higher degree of alignment at ∼15% bicelle concentration with a similar orientation of the alignment tensor.

For the interpretation of the experimental dipolar couplings in terms of the structure of the oligosaccharide, a relaxed grid search strategy, as described under Experimental Section, was used to investigate the conformations of the Lewis epitopes of LNF-1, LNF-2, LNF-3, and LND-1. The glycosidic torsion angles ϕ and ψ of the Lewis epitope of each oligosaccharide were simultaneously scanned over their full range in a four-dimensional grid search using a step of 2.5°. Each conformation was tested for agreement with NOE data, and those in agreement were energy-minimized and then aligned to reproduce the experimental dipolar couplings of the Lewis epitope substructure (*16*). Since each of these Lewis epitopes has five nonparallel C−H vectors, it is anticipated that a small set of orientations of the epitope

Table 4: Experimental Interresidual NOE for Oligosaccharides Studied and the Corresponding Distances in the Structures Obtained

	proton pair	exp NOE	distance (Å)		proton pair	exp NOE	distance (Å)
LNT	H1a−H3b[c]	M	2.56	LNF-1[a]	H1a−H2b	S	2.47
	H1b−H3c	S	2.48		H5a−H2c	S	2.38
	H1c−H4d[c]	S	2.33		H1b−H3c	S	2.61
					NHc−H1b[a]	S	2.38
LNF-2[d]	H1a−H4c	S	2.57		H1c−H3d	S	2.50
	H5a−H2b	M	2.25		H1d−H4e	S	2.85
	H1b−H3c	M	2.64				
	H1c−H3d	S	2.19	LN-neo-T	H1a−H4b	−	2.62
	H1c−H4d	M	3.10		H1b−H3c	−	2.46
	H1d−H4e[f]	S	2.30		H1c−H4d	−	2.27
LNF-3[g]	H1a−H3c	M	3.08	LND-I[d]	H1a−H2b	S	2.36
	H5a−H2b	M	2.81		H5a−H2d	S	2.97
	H1b−H4c	S	2.43		H1b−H3d	S	2.53
	H1b−H6c[h]	S	2.89		H1c−H4d	S	2.53
	H1c−H3d	S	2.22		H5c−H2b	S	2.40
	H1d−H4e[b]	S	2.69		H1d−H3e	S	2.44
					H1d−H4e[i]	S	2.38
					H1e−H4f[e]	S	2.65

[a] NOEs reported in (*32*). [b] Possible overlap with NOE H1c−H3d. [c] Tentative assignment due to overlapping. [d] NOEs reported in (*11*). [e] Possible overlap with NOE H1e−H3f$_\beta$. [f] Possible overlap with NOE H1d−H3e$_\beta$. [g] NOEs reported in (*10*). [h] Distance for H1b−H6c$_{proR}$. [i] This NOE overlap with NOE H1d−H3d.

Table 5: Glycosidic Dihedral Angles Grouped by Linkage Type for the Best Conformer Found in Agreement with Dipolar Couplings and NOE Data of the Oligosaccharides Studied

	Fuc α(1→2) Gal	Fuc α(1→3) GlcNAc	Fuc α(1→4) GlcNAc	Gal β(1→3) GlcNAc	Gal β(1→4) GlcNAc	GlcNAc β(1→3) Gal	GlcNAc β(1→4) Glc
LNF-1	ϕab/ψab −80 ± 10°/ 140 ± 10°			ϕbc/ψbc −65 ± 10°/ −103 ± 10°		ϕcd/ψcd −137°/ −160° [a]	ϕde/ψde −35°/ 108° [a]
LNF-2			ϕac/ψac −63 ± 10°/ 131 ± 10°	ϕbc/ψbc −61 ± 10°/ −101 ± 10°		ϕcd/ψcd −105°/ −143° [a]	ϕde/ψde −135°/ 95° [a]
LNF-3		ϕac/ψac −40 ± 10°/ −90 ± 10°			ϕbc/ψbc −65 ± 10°/ 114 ± 10°	ϕcd/ψcd −95°/−114° [a]	ϕde/ψde −69°/146° [a]
LND-1	ϕab/ψab −109.3°/ 152.4° [b]		ϕcd/ψcd −74.9 ± 5°/ 139.8 ± 5°	ϕbd/ψbd −72.5 ± 5°/ −100.7 ± 5°		ϕde/ψde −139.1°/ −153.9° [a]	ϕef/ψef −72.8°/ 146.5° [a]
LNT				ϕab/ψab −59.7 ± 10°/ −118.8 ± 15°		ϕbc/ψbc −60.5 ± 10°/ −157.0 ± 10°	ϕcd/ψcd −137.7°/ 94.1° [b]
LN-neo-T					ϕab/ψab −67.7°/ 139.2° [b]	ϕbc/ψbc −75.5°/ −172.0° [b]	ϕcd/ψcd −91.4°/ 124.5° [b]

[a] Uncertainties in angles are shown in the dipolar map of Figure 3. [b] Uncertainties in angles are shown in the dipolar map of Figure 4.

should be found to be consistent with the dipolar data if it can be adequately represented by a single rigid model. This latter assumption is supported by a broad consensus in the literature (*10, 21*−*29*). This single structure model for the oligosaccharide is consistent with either a single and rigid structure in solution or an average structure which represents average dihedral angles if fast internal motions are present (*18*). This case corresponds to an oligosaccharide having internal motions of the first kind (*1*). It proved to be possible to calculate conformations and orientation tensors which allowed an excellent fit of calculated dipolar coupling values within experimental error and in agreement with the interresidual NOE data of Table 4 (see Tables 1−3).

The energy-minimized models for the Lewis epitopes are very similar to those previously proposed, which is consistent with our assumption of internal motion of the first kind (*11, 30*−*34*). Thus, the rigid models given in Table 5 may also represent models having internal motion of the first kind with symmetric fluctuations in the dihedral angles. Molecular dynamics simulations suggest fluctuations of about ±15° for the Lewis oligosaccharide epitopes.

Since single conformation models very similar to those previously proposed for these Lewis blood group epitopes can be oriented to give excellent agreement with experimental dipolar coupling data, we can conclude that the calculated orientation tensor is valid for the Lewis epitope and that it can be used to deduce structural information for the remaining residues. Although there is little evidence in the literature about the flexibility of the β-GlcNAc-(1→3)-Gal linkage, we propose as a hypothesis that is has flexibility of the first kind and could therefore fit a single conformation model. Accordingly, the glycosidic dihedral angles for the β-GlcNAc-(1→3)-Gal linkage were varied in 2.5° intervals with the Lewis epitope held in a fixed conformation. After energy minimization, each structure was submitted to the protocol for calculating the orientation tensor. The dipolar maps of

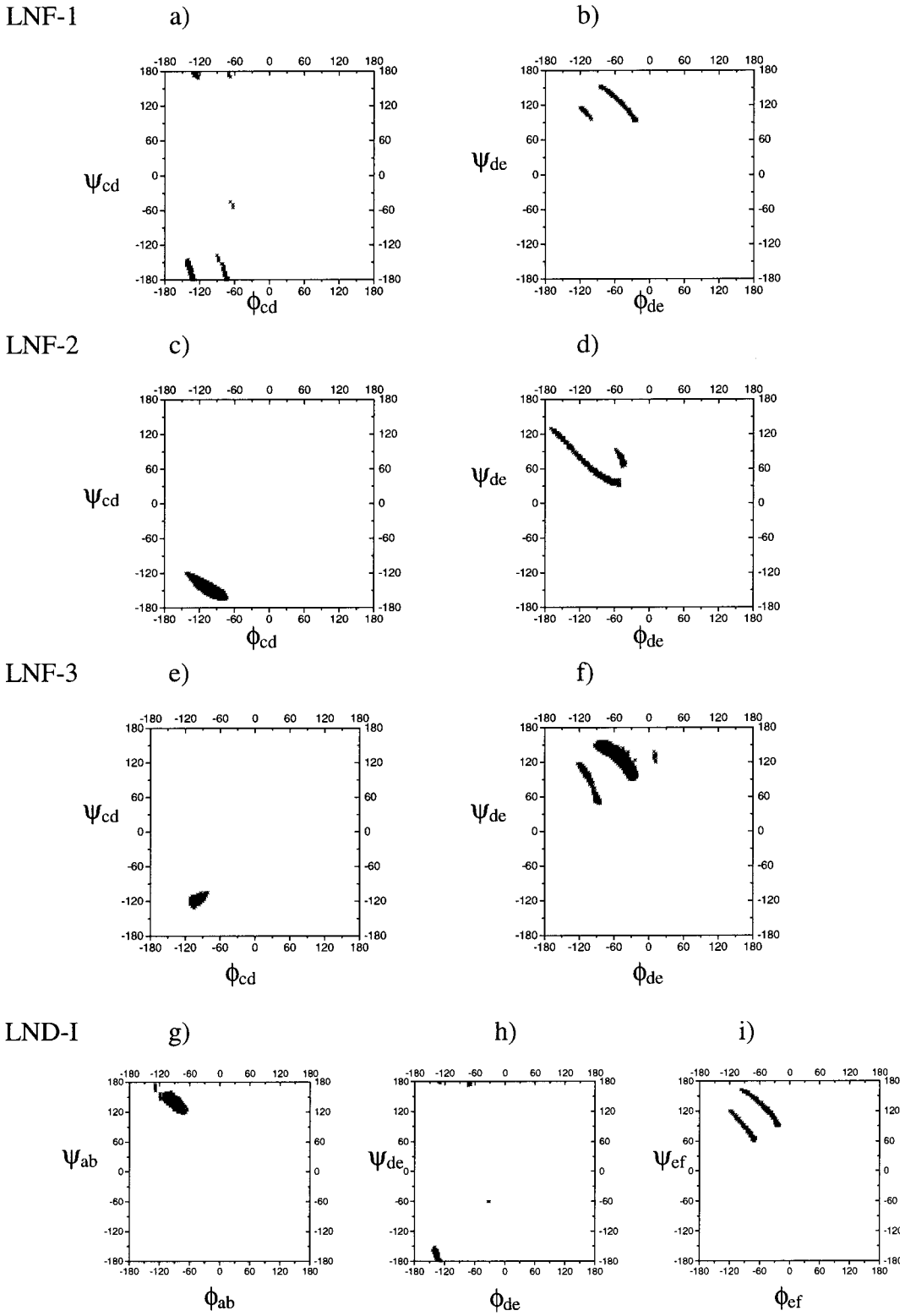

FIGURE 3: ϕ/ψ dipolar maps for LNF-1, LNF-2, LNF-3, and LND-1 in which each point is a low-energy conformer consistent with the experimental dipolar experimental data. In all cases, points represented are conformations with mean deviations ≤0.8 Hz.

Figure 3 show dots for each resulting low-energy structure within 4 kcal/mol of the global minimum for which the χ^2 optimization gives a mean deviation ≤0.8 Hz from the experimental dipolar couplings. For LND-1 (Figure 3h), a small region around ϕ/ψ −139°/−154° for the β-D-GlcNAc-(1→3)-D-Gal linkage gave good agreement. These values are in a low-energy region and are also consistent with the NOE data of Table 4. For LNF-2 and LNF-3, a slightly larger area (Figure 3c,e) of ϕ and ψ gives satisfactory agreement with the dipolar coupling data, and the best-fit/low-energy conformer in those areas is indicated in Table 5. For LNF-1 (Figure 3a), several regions in ϕ/ψ space for the β-D-Glc-NAc-(1→3)-D-Gal linkage can be reconciled with the dipolar coupling data, but only the region around angles ϕ/ψ −137°/−160° is consistent with the NOE and energy calculations.

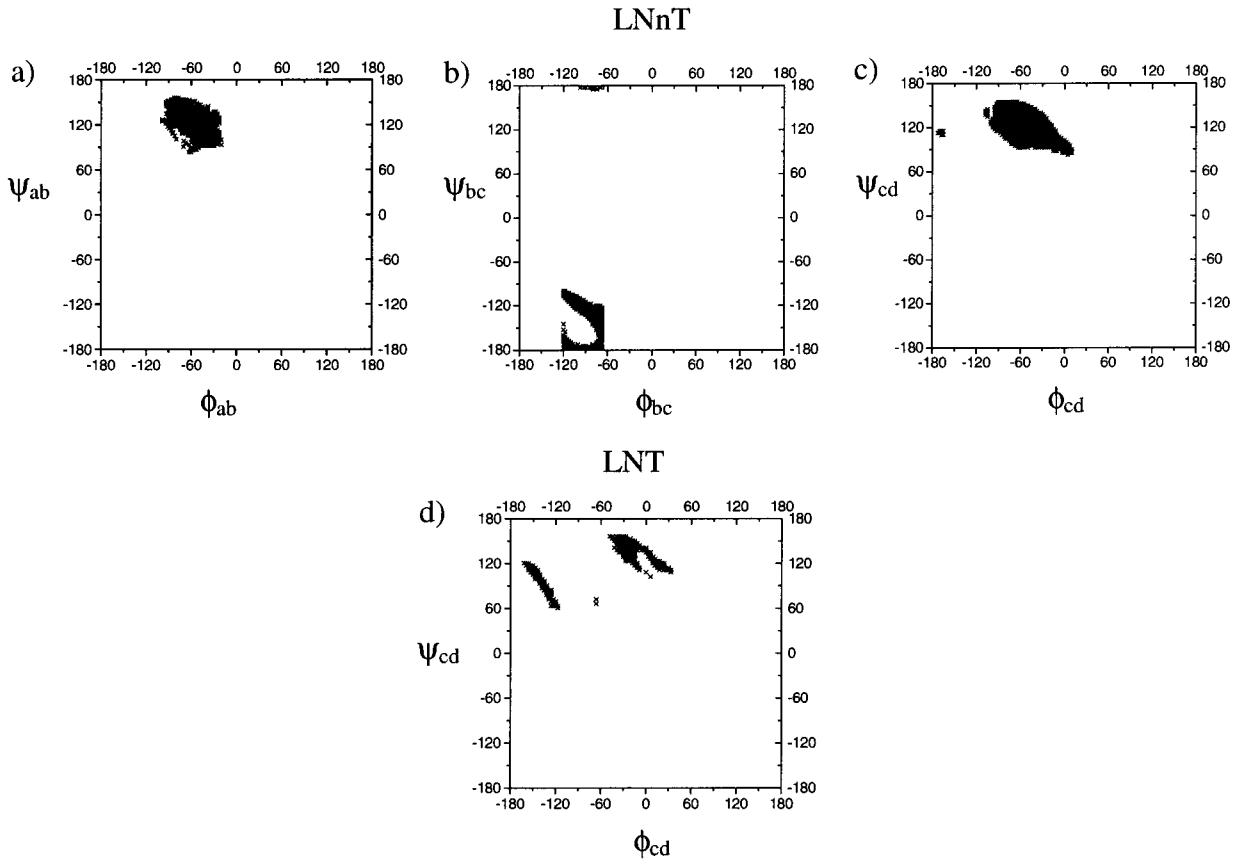

FIGURE 4: ϕ/ψ dipolar maps for LNnT and LNT in which each point is a low-energy conformer consistent with the experimental dipolar experimental data. In all cases, points represented are conformations with mean deviations ≤0.8 Hz.

These results support the hypothesis that the β-D-GlcNAc-(1→3)-D-Gal linkage is relatively rigid, exhibiting internal motion of the first kind, and suggest that the calculated orientation tensor could be used to search for conformations of the β-D-Gal-(1→4)-D-Glc linkage which agree with the dipolar coupling data. Therefore, another complete grid search was performed for the terminal β-D-Gal-(1→4)-D-Glc linkage by varying the dihedral angles in 2.5° intervals while maintaining fixed the rest of the structure. The structures generated were analyzed in a similar way, and the corresponding dipolar maps are shown in Figure 3b,d,f,i. It can be seen that although the energy calculations and dipolar couplings limit considerably the conformational space available for this linkage, there are several possible regions compatible with the dipolar coupling experimental data. The conformations of the β-D-Gal-(1→4)-D-Glc linkage region in agreement with dipolar coupling data are all in the syn region, corresponding to the major conformer of lactose (*35*). The dipolar coupling method is less sensitive to minor conformers than is NOE since it involves a simple linear average of coupling values of multiple conformers. Therefore, our data do not rule out a contribution by a few percent of the anti conformer. The low-energy conformation with the best fit to the dipolar coupling data is indicated in Table 5. The uncertainty in dihedral angles is indicated in Figure 3, which shows all conformations within 4 kcal of the global minimum whose calculated dipolar coupling values agree within 0.8 Hz of experiment. All these conformations are consistent with the NOE data of Table 4.

Applications of this sequential method for model building, when applied to ordinary molecular mechanics calculations, are subject to the criticism that the results can depend strongly on the choice of the initial structure. But in this application to the interpretation of dipolar coupling data, since both the conformation of the Lewis epitope and the orientation tensor are likely to be correct, little new information is required to correctly orient the 3-linked galactose residue. The requirements of the molecular geometry combined with the two independent coupling values measured for the galactose residue should be adequate to provide an unambiguous orientation. The orientation of the 4-linked glucose residue is somewhat more speculative since it depends on the assumption of rigidity for the conformation of the linkage to galactose and on dipolar coupling values for glucose which correspond to only a single direction.

The results for the β-D-GlcNAc-(1→3)-D-Gal linkage in the larger milk oligosaccharides suggested that this linkage could be represented by a single conformer in the tetrasaccharides LNT and LNnT (residues b−c). Under the hypothesis that motions about the linkages of the nonreducing terminal residues might be adequately described as belonging to the first kind, four-dimensional grid searches were carried out on the two glycosidic linkages of the nonreducing terminal trisaccharide fragments of these tetrasaccharides. The results of this search for the a−b−c fragment of LNT show a very small region of conformational space in which the energy is within 4 kcal/mol of the global minimum and the calculated dipolar coupling values are within 0.8 Hz of experiment. The estimated uncertainty of the dihedral angles is given in Table 5 along with the conformation having the best fit to the dipolar coupling data. For LNnT, the error limits indicated in Figure 4 show a larger but still limited

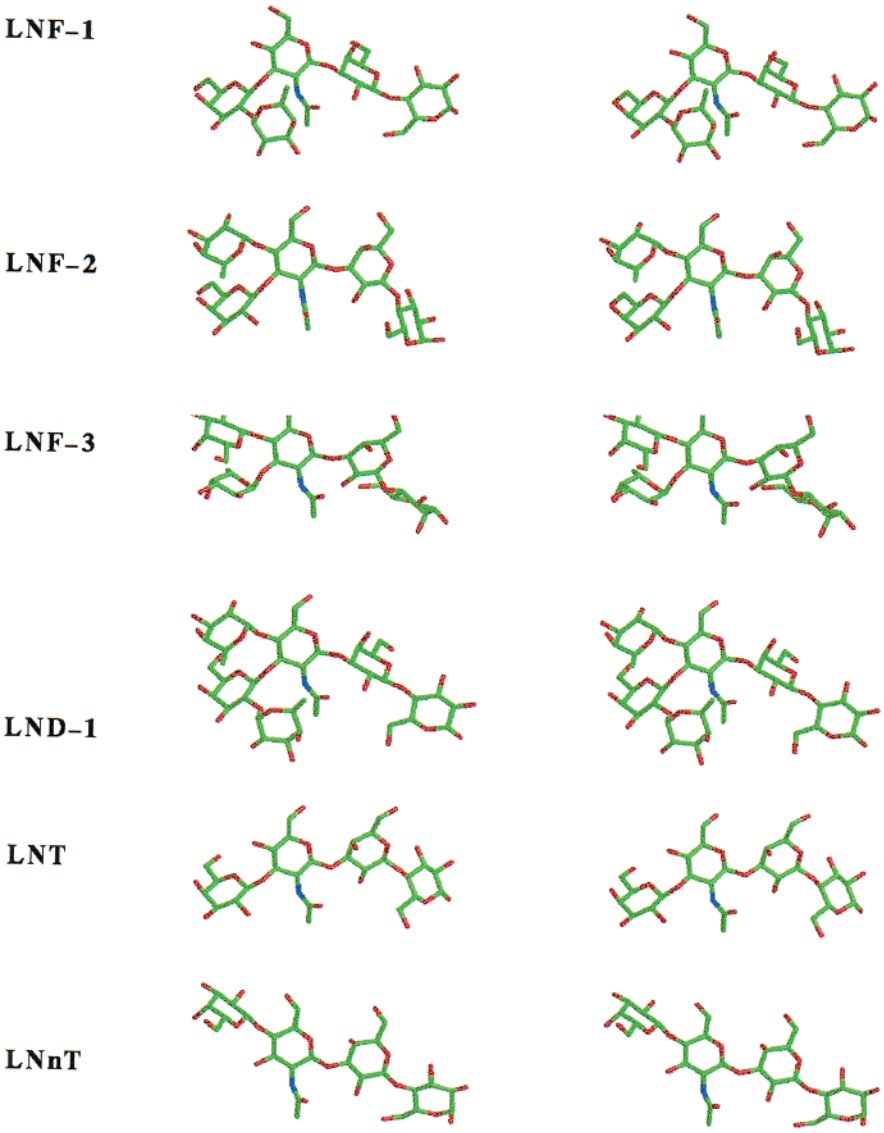

FIGURE 5: Stereoview of the best conformer found in agreement with dipolar couplings and NOE data of the oligosaccharides studied.

region of agreement with experiment, and the model with the best fit is indicated in Table 5.

The best structure obtained at the trisaccharide stage was fixed and used for the orientation of the terminal β-D-Gal-(1→4)-D-Glc linkage of LNT and LNnT by performing a full grid search with this linkage. The results showed that only a limited region of the conformational space gave dipolar coupling values consistent with experiment. The dipolar map of Figure 4a for the terminal linkage β-Gal-(1→4)-Glc in LNT shows two possible low-energy regions consistent with the dipolar coupling data. Both regions of the map are compatible with a NOE distance for H-1Gal−H-4Glc of <3.5 Å. Nevertheless, NOE distances in the region on the left of the map range from 2.5 to 3.0 Å, which is more consistent with a strong NOE (Table 4) than the central region which ranges from 3.0 to 3.5 Å. For this reason, the conformer in the region on the left with best agreement with the dipolar couplings was chosen (Table 5). For LNnT, the dipolar map of Figure 4d shows a single region compatible with dipolar couplings. The models giving the best agreement with the dipolar coupling data are indicated in Table 5. Without carbon scalar coupling data and definitive NOE data, the evidence for single conformations of LNT and LNnT is not as strong as that for the Lewis oligosaccharides. Since no previous models have been proposed for these tetrasaccharides, the models of Table 5 should be regarded as hypotheses.

The model for LNT given in Table 5 is based on calculations of the dipolar coupling data recorded in the 7.5% bicelle solution. This model was also oriented using the Euler angles obtained in the fit of the 7.5% bicelle solution and optimizing exclusively the magnitude components of the asymmetric tensors A_a and A_r, to fit the dipolar coupling data in the 15% bicelle solution. The agreement between this simulation and the experimental data in Table 1 shows an adequate fit suggesting that there is little difference between the conformation of the oligosaccharide in the two bicelle solutions (Table 1).

In summary, this protocol has made it possible to orient all the residues of each oligosaccharide to reproduce the corresponding dipolar coupling data. Structures which best reproduce the experimental dipolar couplings were selected with χ^2 from the lowest value up to those corresponding to deviations of the experimental $^1D_{CH}$ values of 0.8 Hz. The proposed models are within 4 kcal/mol of the global minimum and are consistent with the NOE data of Table 4.

The conformations of the reducing terminal disaccharide are extended and do not predict any NOE which is not observed. A stereoview of the best conformation obtained for each oligosaccharide is given in Figure 5 and the corresponding glycosidic dihedral angle values are given in Table 5. The calculated dipolar couplings obtained for these structures are given in Tables 1−3. The model for LNF-2 is identical to that recently published (*7*).

It can be seen in Tables 1−3 that the linear regressions between experimental and calculated $^1D_{CH}$ values show excellent correlation with coefficients R^2 ranging from 0.995 to 0.999. The χ^2 parameter is more sensitive to the total deviation between experimental and calculated data. For the oligosaccharides at 7.5% bicelle concentration, χ^2 ranges from 0.291 to 0.649, and the mean deviations between experimental and calculated $^1D_{CH}$ are in a range from 0.3 to 0.7 Hz. For the case of LNT at 15% bicelle concentration, the deviations are larger with $\chi^2 = 2.292$, and the mean deviation is 1.4 Hz. The higher deviation is not unexpected because of the larger dipolar couplings at higher bicelle concentration.

CONCLUSIONS

For the structural interpretation of the $^1D_{CH}$ data, we have used the approach proposed by Tjandra and Bax (*2*) in which the orientation tensor defining the alignment of the carbohydrate in the bicelle liquid crystal is obtained by fitting the $^1D_{CH}$ data to a model structure. The method is effective for a model substructure in which at least five C−H bond vectors are directed differently and which exhibits internal motion of the first kind. While the requirement for five nonparallel vectors is restrictive in pyranosides for which all axial C−H bonds are approximately parallel, the comparatively rigid Lewis epitopes provide an especially favorable test case.

In this paper exhaustive conformational grid searches are used as an alternative to the simulated annealing method (*5*) in the interpretation of $^1D_{CH}$ data. An exhaustive grid search is generally practical for only a limited number of degrees of freedom. For interpretation of dipolar coupling data, a limited substructure containing a few degrees of freedom can be selected for the initial search. Once its orientation is known, the other degrees of freedom can be searched exhaustively. NOE and other NMR experimental data such as scalar couplings can also be included and effectively reduce the computational time for the exhaustive searches.

This straightforward interpretation of dipolar coupling data is possible only for molecules which can be represented by a single conformation. While NOE data and MD simulations support this assumption for the Lewis epitopes, it may be questioned for LNT, LNnT, and the reducing terminal disaccharide of the other milk oligosaccharides which are extended rather than folded back. It is possible that the conformations listed in Table 5, which represent the best agreement with the dipolar coupling data, could be virtual conformations. While the conformations of Table 5 all lie in low-energy regions, more stringent tests for internal motion of the second kind would require additional data such as scalar coupling. While few such data are available, $^3J_{H-C-O-C}$ data for LNF-2 have been reported which are in reasonable agreement with the Lewis[a] model (*26*), and $^3J_{H1-C1-O1-Cx}$ data have been reported to be ∼3.4 Hz for the β-D-GlcNAc-(1→3)-β-D-Gal and β-Gal-(1→4)-Glc linkages (*14*). Thus, better agreement with these scalar coupling data could be obtained by changing ϕ for these two linkages slightly from the values given in Table 5 which represent the best fit to dipolar coupling data. While such a conformation would still be in reasonable agreement with the dipolar maps represented in Figure 3c,d, the fit is slightly poorer.

Discrepancies between the calculated and experimental dipolar coupling values could arise from our assumption that the models contain the correct geometries for the monosaccharides. In our treatment, this is determined by the CVFF force field which might not reproduce the pyranose ring bond angles and dihedral angles with sufficient accuracy. Comparison of experimental and calculated $^1D_{CH}$ values in Tables 1−3 shows that coupling values for axial C−H bonds on a given residue are similar but not identical as they should be if the bond vectors are exactly parallel. While the pyranose ring chairs calculated with the force field CVFF gave reasonable agreement with the average $^1D_{CH}$ experimental value for a set of C−H parallel vectors in a chair, the small distortions of the pyranose chairs required to perfectly fit these differences are not reproduced.

REFERENCES

1. Martin-Pastor, M., and Bush, C. A. (1999) *Biochemistry 38*, 8045−8055.
2. Tjandra, N., and Bax, A. (1997) *Science 278*, 1111−1114.
3. Bolon, P. J., and Prestegard, J. H. (1998) *J. Am. Chem. Soc. 120*, 9366−9367.
4. Rundlöf, T., Landersjö, C., Lycknert, K., Malinak, A., and Widmalm, G. (1998) *Magn. Reson. Chem. 36*, 773−776.
5. Kiddle, G. R., and Homans, S. W. (1998) *FEBS Lett. 436*, 128−130.
6. Shimizu, H., Donohue-Rolfe, A., and Homans, S. W. (1999) *J. Am. Chem. Soc. 121*, 5815−5816.
7. Martin-Pastor, M., and Bush, C. A. (2000) *Carbohydr. Res. 323*, 147−155.
8. Fischer, M. W. F., Losonczi, J. A., Weaver, J. L., and Prestegard, J. H. (1999) *Biochemistry 38*, 9013−9022.
9. Bush, C. A., Panitch, M. M., Dua, V. K., and Rohr, T. E. (1985) *Anal. Biochem. 145*, 124−136.
10. Miller, K. E., Mukhopadhyay, C., Cagas, P., and Bush, C. A. (1992) *Biochemistry 31*, 6703−6709.
11. Cagas, P., and Bush, C. A. (1990) *Biopolymers 30*, 1123−1138.
12. John, B. H. (1992) *J. Magn. Reson., Ser. A 101*, 113.
13. Breg, J., Romijn, D., Vliegenthart, J. F. G., Strecker, G., and Montreuil, J. (1988) *Carbohydr. Res. 183*, 19−34.
14. Xu, Q., Gitti, R., and Bush, C. A. (1996) *Glycobiology 6*, 281−288.
15. Strecker, G., Wieruszeski, J. M., Michalski, J. C., and Montreuil, J. (1989) *Glycoconjugate J. 6*, 67−83.
16. Tjandra, N., Grzesiek, S., and Bax, A. (1996) *J. Am. Chem. Soc. 118*, 6264−6272.
17. Hagler, A. T., Lifson, P., and Dauber, P. (1979) *J. Am. Chem. Soc. 101*, 5122−5130.
18. Sanders, C. R., II, Hare, J. H., Howard, K. P., and Prestegard, J. H. (1994) *Prog. NMR Spectrosc. 26*, 421−444.
19. Bax, A., and Tjandra, N. (1997) *J. Biomol. NMR 10*, 289−292.
20. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) *J. Magn. Reson. 131*, 159−162.
21. Lemieux, R. U., Bock, K., Delbaere, T. J., Koto, S., and Rao, V. S. (1980) *Can. J. Chem. 58*, 631−653.
22. Duus, J. O., Nifant'ev, N., Shashkov, A. S., Khatuntseva, E. A., and Bock, K. (1996) *Carbohydr. Res. 288*, 25−44.
23. Homans, S. W., and Forster, M. (1992) *Glycobiology 2*, 143−151.

24. Delbaere, L. T. J., Vandonselaar, M., Prasad, L., Quail, J. W., Pearlstone, J. R., Carpenter, M. R., Spohr, U., and Lemieux, R. U. (1990) *Can. J. Chem. 68*, 1116−1121.
25. Rutherford, T. J., Spackman, D. G., Simpson, P. J., and Homans, S. W. (1994) *Glycobiology 4*, 59−68.
26. Kogelberg, H., and Rutherford, T. J. (1994) *Glycobiology 4*, 49−57.
27. Perez, S., Mouhous-Riou, N., Nifantev, N. E., Tsevetkov, Y. E., Bachet, B., and Imberty, A. (1996) *Glycobiology 6*, 537−542.
28. Berthault, P., Birlirakis, N., Rubenstein, G., Sinay, P., and Desvaux, H. (1996) *J. Biomol. NMR 8*, 23−35.
29. Kurutz, J. M., and Kiessling, L. (1997) *Glycobiology 7*, 337−347.
30. Mukhopadhyay, C., and Bush, C. A. (1991) *Biopolymers 31*, 1737−1746.
31. Yan, Z.-Y., and Bush, C. A. (1990) *Biopolymers 29*, 799−811.
32. Cagas, P., Kaluarachchi, K., and Bush, C. A. (1991) *J. Am. Chem. Soc. 113*, 6815−6821.
33. Rao, B. N. N., Dua, V. K., and Bush, C. A. (1985) *Biopolymers 24*, 2207−2229.
34. Bagley, S., Hovacs, H., Kowaleski, J., and Widmalm, G. (1992) *Magn. Reson. Chem. 30*, 733−739.
35. Espinosa, J. F., Canada, J., Asensio, J. L., Martin-Pastor, M., Dietrich, H., Martin-Lomas, M., and Jimenez-Barbero, J. (1996) *J. Am. Chem. Soc. 118*, 10862−10871.

BI992050Q