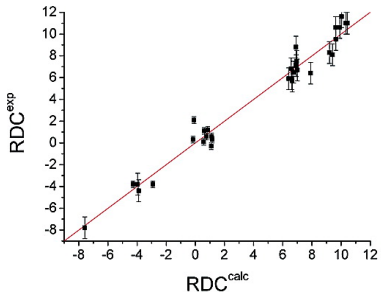
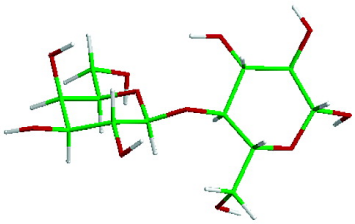
Article

Limited Flexibility of Lactose Detected from Residual Dipolar Couplings Using Molecular Dynamics Simulations and Steric Alignment Methods

Manuel Martn-Pastor, Angeles Canales, Francisco Corzana, Juan L. Asensio, and Jess Jimnez-Barbero

J. Am. Chem. Soc., **2005**, 127 (10), 3589-3595• DOI: 10.1021/ja043445m • Publication Date (Web): 19 February 2005





More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 6 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES
Published on Web 02/19/2005

Limited Flexibility of Lactose Detected from Residual Dipolar Couplings Using Molecular Dynamics Simulations and Steric Alignment Methods

Manuel Martín-Pastor,*,† Angeles Canales,‡ Francisco Corzana,‡ Juan L. Asensio,‡ and Jesús Jiménez-Barbero*,‡

Contribution from the Laboratorio Integral de Dinámica e Estructura de Biomoléculas José R. Carracido, Unidade de Resonancia Magnética, y Unidad Asociada al CSIC de RMN de Biomoléculas Santiago de Compostela, Spain, and Departamento de Estructura y Función de Proteínas, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

Received October 29, 2004; E-mail: jjbarbero@cib.csic.es; mmartin@usc.es

Abstract: The conformational flexibility of lactose in solution has been investigated by residual dipolar couplings (RDCs). One-bond carbon-proton and proton−proton coupling constants have been measured in two oriented media and interpreted in combination with molecular dynamics simulations (MD). Two different approaches, known as PALES (Zweckstetter et al., *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792) and TRAMITE (Azurmendi et al., *J. Am. Chem. Soc.* **2002**, *124*, 2426−2427), have been used to determine the alignment tensor from a shape-induced alignment model with the oriented medium. The steric alignment of the structures from several MD trajectories has provided ensemble averaged RDCs that have been compared with the experimental ones. The obtained results reveal the almost exclusive presence of a major low energy region defined as syn-Φ/syn-Ψ (> 97%), for which sampling occurs in a dynamic manner. This result satisfactorily agrees with that determined by standard NOE-based methods.

Introduction

The conformational flexibility of carbohydrates may play an important role for their molecular recognition by their biological targets, such as lectins, enzymes or antibodies.[1] Conformational analysis of saccharides has been usually based on inter-residual NOE data,[2] in some cases helped by interglycosidic long-range heteronuclear coupling constants measurements.[3,4] Frequently, these data are insufficient to determine plausible structural models of the carbohydrate and their relative weights; furthermore the direct interpretation of NOE data is hampered by certain difficulties that may lead to the generation of *virtual conformations*.[5] The $\langle r^{-6} \rangle$ dependence of cross-correlation rates tends to overestimate the weight of low populated conformers with short proton−proton distances. Additional problems for NOE quantification may arise from the uncertainty in the dynamics of exchange among conformations and/or when the rate of internal motion around the glycosidic linkages become similar to that of overall tumbling, a case particularly true for small sized oligosaccharides.[6] The assistance of molecular mechanics/dynamics calculations is an important step required for the proper interpretation of the experimental NMR data.[7] The use of residual dipolar couplings (RDC) has emerged as an alternative or complementary method to determine the conformational properties of biomolecules.[8,9] RDCs can be obtained by dissolving the solute, i.e., in this case the carbohydrate, with a special liquid crystal which is able to orient spontaneously in the NMR magnetic field and impose a certain degree of alignment to the solute as result of steric and electrostatic interactions with the liquid crystal.[8−11]

The structural interpretation based on the calculation of the expected RDCs is particularly simple when the orientation and magnitude of the three principal components of the so-called alignment tensor is known.[12] Attempts to predict the alignment tensor from physical models of the solute−liquid crystal system have been considered.[11,13−16] In particular, when neutral media

† Unidad Asociada al CSIC de RMN de Biomoléculas Santiago de Compostela.

‡ Centro de Investigaciones Biológicas.

(1) Dwek, R. A. *Chem. Rev.* **1996**, 96, 683.
(2) Kroon-Batenburg, L. M. J.; Kroon, J.; Leeflang, B. R.; Vliegenthart, J. F. G. *Carbohydr. Res.* **1993**, *245*, 21−42.
(3) Tvaroska, I.; Taravel, F. *Adv. Carbohydr. Chem. Biochem.* **1995**, *51*, 15−61.
(4) Bose, B.; Zhao, S.; Stenutz, R.; Cloran, F.; Bondo, F.; Bondo, G.; Hertz, B.; Carmichael, I.; Serianni, A. *J. Am. Chem. Soc.* **1998**, *120*, 11158−73.
(5) Cumming, D. A.; Carver, J. P. *Biochemistry* **1987**, *26*, 6664−6676.
(6) Catoire, L.; Braccini, I.; Bouchemal-Chibani, N.; Jullien, L.; Herve du Penhoat, C.; Perez, S. *Glycoconjugate J.* **1997**, *14*, 935−943.
(7) Weimar, T.; Woods, R. In *NMR spectroscopy of glycoconjugates*; Jiménez-Barbero, J., Peters, T., Eds.; Wiley-VCH: Weinheim, 2002; pp 111−144.
(8) Prestegard, J. H. *Nat. Struct. Biol.* **1998**, *5*, 517−522.
(9) Tjandra N.; Bax A. *Science*,**1997**, 278, 1111−1114.
(10) Bax, A. *Protein Sci.* **2003**, *12*, 1−16.
(11) Annila, A.; Permi, P. *Concepts Magn. Reson.* **2004**, *23A*, 22−37.
(12) Kramer, F.; Deshmukh, M. V.; Kessler, H.; Glaser, S. J. *Concepts NMR* **2004**, *21A*, 10−21.
(13) Zweckstetter, M.; Bax A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.
(14) Fernandez, M. X.; Bernado, P.; Pons, M.; de la Torre, J. G. *J. Am. Chem. Soc.* **2001**, *123*, 12037−12047.
(15) (a) Almond, A.; Duus, J. Ø. *J. Biomol. NMR* **2001**, *20*, 351−363. (b) Almond, A.; Axelsen, J. B. *J. Am. Chem. Soc.* **2002**, *124*, 9986−9987. (c) Almond, A.; Petersen, B. O.; Duus, J. Ø. *Biochemistry* **2004**, *43*, 5853−5863.

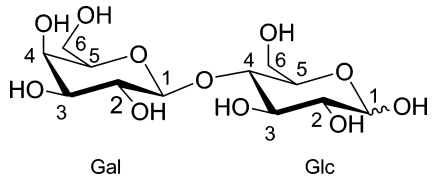


Figure 1. Molecular structure of lactose disaccharide β-D-Gal*p*-(1 → 4)-D-Glc*p* **1**, showing the atomic labels used throughout the text. Both anomers (Glc-α and Glc-β) are present in solution.

or solutes are used, the electrostatic contribution can be neglected, and the calculations of RDCs can be made from a steric obstruction alignment model, a method known as PALES (Prediction of ALignmEnt from Structure).[13] Since the alignment in neutral media just reflects the asymmetries in the shape of the molecule under study, the key structural information is already encoded in its inertia tensor, and using this inertia tensor to predict the alignment tensor as been proposed, except for a scaling factor which depends on experimental conditions. This method is known as TRAMITE (TRacking Alignment from the Moment of Inertia TEnsor).[16] Both TRAMITE and PALES methods allow checking the correctness of a derived structure by agreement with the experimental RDCs, or alternatively, TRAMITE can be used to evaluate the agreement of multiple-conformer models in flexible molecules.[11,16] The key point of this approach is that it provides an attractive way to overcome some of the aforementioned difficulties in the application of NOE methods to carbohydrate conformational analysis, and provided that enough RDCs are measured, they can be computed by a weighted average over multiple conformations.[17,15]

In this article, we have boarded the structural study of the flexible disaccharide lactose (Figure 1) using only RDCs, which have been interpreted with the assistance of molecular dynamic simulations in explicit solvent (MD). These results have been compared with our previous results using NOE information.[18−20] To obtain a significant number of nonambiguous RDCs avoiding signal overlapping problems, we have used a ^{13}C-labeled disaccharide. Different applications of RDCs to the conformational and dynamics analysis of carbohydrate molecules have been reported.[21−23] Also, several authors have previously used basic steric models to perform a fit of the RDC data in oligosaccharides.[15−17] In this work, we explicitly consider the use of the steric alignment models TRAMITE and PALES for the simulation of RDCs.

Experimental Section

Sample Preparation. The synthesis of uniformly ^{13}C labeled (U-^{13}C) lactose (Figure 1) will be described elsewhere, although it was based on the method of Field and co-workers.[24] For the measurement of RDC, three samples of U-^{13}C lactose were prepared. A lactose isotropic sample was prepared by dissolving 5 mg of U-^{13}C lactose in 0.3 mL of D_2O. A first oriented sample was prepared by dissolving 5 mg of U-^{13}C lactose in 0.3 mL in the alignment medium composed of cetylpyridinium chloride (CPCl)/hexanol 1:1 and brine (0.2 M NaCl)[25−27] at a CPCl/hexanol/brine concentration of 5% (w/w). A second oriented sample was prepared by dissolving 5 mg of U-^{13}C lactose in 0.3 mL of the alignment medium formed by pentaethylene glycol mono-dodecyl ether (C12E5)/hexanol[28] at a molar ratio of 1:0.96 and at a concentration of 5% (w/w) in D_2O. The two oriented samples are referred in the text as CPCl and C12E5, respectively. The presence of alignment in the magnet was monitored by observing the symmetric quadrupolar splitting on the deuterium D_2O lock signal, which is proportional to the degree of achieved alignment.

NMR Spectroscopy. Experiments were acquired at 25 °C on a Varian INOVA NMR instrument operating at 750 MHz, and the data were processed with MestRe-C v3.0 software.[29]

Measurement of One-Bond Carbon-Proton RDC. The one bond C−H dipolar couplings, $^1D_{CH}$, were determined for the nonoriented D_2O sample and the two oriented CPCL and C12E5 samples using 2D constant-time coupled-enhanced HSQC experiments, 2D CT-CE-HSQC.[30]

The constant time period was set to a nominal value of 22 ms to remove the one bond carbon−carbon couplings. To maximize the resolution in the indirect dimension, t_1-folded experiments were acquired, the ^{13}C carrier position was set at 70 ppm and the spectral width was set to 30 ppm (∼5600 Hz) in this dimension. For each experiment, a 2D matrix of 1024 × 300 complex FID data points was acquired. The 2D FIDs were apodized in both dimensions with a 90°-shifted sine-bell function and zero-filling to give, after Fourier transformation, a 2D spectrum of 2048 × 1024 real points. The column corresponding to each C−H signal was carefully phased and stored. After an inverse Hilbert transformation and a zero-filling to 16 K, the splitting of the doublet was determined by the line fitting routine in the Mestre-C software.[29] The measured C−H splitting in the indirect F1 dimension of this experiment corresponds to twice the scalar $^1J_{CH}$ (or scalar + RDC) coupling. The sign and magnitude of the $^1D_{CH}$ couplings were determined from the difference between the real coupling in an oriented sample and the nonoriented sample. The $^1D_{CH}$ results obtained for the CPCL and C12E5 samples are given in Table 1. The estimated experimental error is ±1 Hz.

Measurement of Multiple-Bond Proton−Proton RDC. The proton−proton dipolar couplings, $^nD_{HH}$, were determined using the 2D HSQC-(selC,selH)-CT-COSY experiment.[31] This selective experiment provides a 2D ^{1}H/^{13}C HSQC-type correlation spectrum with the selected one-bond ^{1}H/^{13}C HSQC peaks (autopeaks) and the ^{1}H−^{1}H CT-COSY correlations (cross-peaks) with each autopeak. The later correlations that appear at the same carbon chemical shift as the autopeak are generated by scalar and/or dipolar coupling with the autopeak proton.

(16) Azurmendi, H.; Bush, C. A. *J. Am. Chem. Soc*. **2002**, *124*, 2426−2427.
(17) Staaf, M.; Hoog, C.; Stevensson, B.; Maliniak, A.; Widmalm, G. *Biochemistry* **2001**, *40*, 3623−3628.
(18) (a) Asensio, J. L.; Cañada, F. J.; Jiménez-Barbero, J. *Eur. J. Biochem*. **1995**, *233*, 618−630. (b) Asensio, J. L.; Jiménez-Barbero, J. *Biopolymers* **1995**, *35*, 55−73.
(19) (a) Rivera, A.; Solis, D.; Diaz-Mauriño, T.; Jimenez-Barbero, J.; Martin-Lomas, M. *Eur. J. Biochem*. **1991**, *197*, 217−228. (b) Espinosa, J. F.; Cañada, F. J.; Asensio, J. L.; Martín-Pastor, M.; Dietrich, H.; Martín-Lomas, M.; Schmidt, R.; Jiménez-Barbero, J. *J. Am. Chem. Soc*. **1996**, *118*, 10862−10871.
(20) Martín-Pastor, M.; Espinosa, J. F.; Asensio, J. L.; Jiménez-Barbero, J. *Carbohydr. Res*. **1997**, *298*, 15−49.
(21) Martin-Pastor, M.; Bush, C. A. In *NMR spectroscopy of glycoconjugates*; Jiménez-Barbero, J., Peters, T., Eds.; Wiley-VCH: Weinheim, 2002; pp 23−28 and references therein.
(22) (a) Prestegard, J. H.; Gluskha, J. In *NMR spectroscopy of glycoconjugates*; Jiménez-Barbero, J., Peters, T., Eds.; Wiley-VCH: Weinheim, 2002; pp 231−245 and references therein. (b) Tolman, J. R.; Al-Hashimi, H. M.; Kay, L. E.; Prestegard, J. H. *J. Am. Chem. Soc*. **2001**, *123*, 1416−1424.
(23) Pham, T. N.; Hinchley, S. L.; Rankin, D. W.; Liptaj, T.; Uhrin, D. *J. Am. Chem. Soc*. **2004**, *126*, 13100−13110.
(24) Shimizu, H.; Brown, J. M.; Homans, S. W.; Field, R. A. *Tetrahedron* **1998**, *54*, 9489−9497.
(25) Porte, G.; Gomati, R.; El Haitamy, O.; Appell, J.; Marignan, J. *J. Phys. Chem*. **1986**, *90*, 5746−5751.
(26) Gomati, R.; Appell, J.; Bassereau, P.; Marignan, J.; Porte, G. *J. Phys. Chem.* **1987**, *91*, 6203−6210.
(27) Prosser, R. S.; Losonczi, J. A.; Shiyanovskaya, I. V. *J. Am. Chem. Soc.* **1998**, *120*, 11010−11011.
(28) Ruckert, M.; Otting, G. *J. Am. Chem. Soc.* **2000**, *122*, 7793−7797.
(29) Cobas, J. C.; Sardina, F. J. *Concepts Magn. Reson*. **2003**, *19A*, 80−96.
(30) Tian, F.; Al-Hashimi, H. M.; Craighead, J. L.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 485−492.
(31) Martín-Pastor, M.; Canales-Mayordomo, A.; Jiménez-Barbero, J. *J. Biomol. NMR* **2003**, *26*, 345−353.

Table 1. Experimental and Calculated RDCs for β-Lactose 1 (Glc-β Anomer) in CPCL and C12E5 Alignment Media[a]

	CPCL EXP	CPCL TRAMITE	CPCL PALES	C12E5 EXP	C12E5 TRAMITE	C12E5 PALES
H1-C1Gal	9.5 ± 1	9.64	9.86	8.8 ± 1	6.90	7.01
H2-C2Gal	11.0 ± 1	10.43	10.77	6.7 ± 1	7.00	7.18
H3-C3Gal	11.0 ± 1	10.33	10.65	7.5 ± 1	6.95	7.12
H4-C4Gal	−7.8 ± 1	−7.59	−7.04	−3.8 ± 1	−3.96	−3.70
H5-C5Gal	11.6 ± 1	10.03	10.29	5.7 ± 1	6.65	6.73
H1-C1Glc	6.4 ± 1	7.91	8.40	6.8 ± 1	6.57	6.54
H2-C2Glc	8.3 ± 1	9.20	8.90	6.5 ± 1	6.74	6.61
H3-C3Glc	10.6 ± 1	9.91	9.66	7.1 ± 1	6.92	6.84
H4-C4Glc	10.6 ± 1	9.63	9.31	5.9 ± 1	6.40	6.39
H5-C5Glc	8.1 ± 1	9.41	9.12	5.9 ± 1	6.61	6.45
H1-H2Gal	0.6 ± 0.3	1.12	1.16	0.6 ± 0.3	0.77	0.78
H1-H3Gal	−3.8 ± 0.3	−4.28	−3.82			
H2-H3Gal	0.3 ± 0.3	1.15	1.25	1.2 ± 0.3	0.85	0.90
H3-H4Gal	0.3 ± 0.3	−0.15	−0.33	2.1 ± 0.3	−0.09	−0.12
H1-H2Glc	0.1 ± 0.3	0.57	0.62	1.1 ± 0.3	0.62	0.59
H2-H3Glc	−0.3 ± 0.3	1.09	0.80			
H1Gal-H4Glc	−4.4 ± 0.3	−3.88	−3.62	−3.8 ± 0.15	−2.92	−2.77

[a] Calculated RDCs are the average of the MD-I (syn-Φ/syn-Ψ) trajectory aligned either with the TRAMITE or PALES method. See Supporing Information for some data on the α-anomer.

The relationship between the peak amplitudes of a cross-peak, I_{cross}, and its autopeak, I_{auto}, is given by eq 1:

$$J + D \propto \arctan(I_{cross}/I_{auto})/\pi\Delta \qquad (1)$$

where J and D represent the proton−proton scalar and the RDC coupling and Δ is the constant time period. A nonlinear fit of I_{cross}/I_{auto} for different values of Δ was made using Origin v6.1 software (OriginLab Corporation, http://www.OriginLab.com) that provided the proton−proton scalar (or scalar + RDC) coupling. For the C12E5 sample the ${}^nD_{HH}$ values were determined by subtraction of the values in the oriented and nonoriented sample[8,9] (Table 1). The ${}^nD_{HH}$ results for the CPCL sample have been previously reported by us[31] and are also given in Table 1. The experimental error is ±0.3 Hz.

Molecular Dynamics Simulations with Explicit Solvent (MD). Previous conformational studies of lactose have determined the existence of three low energy regions,[18−20] centered around the following Φ/Ψ glycosidic torsion angles: region I, syn-Φ /syn-Ψ (60°/0°); region II, syn-Φ /anti-Ψ (60°/180°); and region III, anti-Φ/syn-Ψ 180°/0°). The Φ/Ψ dihedral angles are defined as the following: (Φ) H_1 Gal-C_1 Gal-O_1 Gal-C_4 Glc and (Ψ) C_1 Gal-O_1 Gal-C_4 Glc−H_4 Glc.

The conformational space available of regions I, II, and III was independently investigated by MD simulations in explicit water, which were performed using the AMBER 8.0 program package[32] together with the GLYCAM (Glycam_2000a and Glycam_04c) parameters designed for carbohydrates.[33] These MD simulations are referred in the text as MD-I, MD-II, and MD-III, respectively. The starting conformation for every MD was the molecular mechanics minimized conformation of lactose with the anomeric carbon of the glucose residue in the beta configuration and the glycosidic dihedral angles at the corresponding region of the Φ/Ψ map. It is important to note that similar results were obtained when the simulations were starting from the alpha configuration of the anomeric carbon of glucose (see Supporting Information). The molecules were hydrated in the Xleap module of AMBER by a periodic box of TIP3P waters.[34] All the simulations were run with the SANDER module of AMBER with SHAKE algorithm[35] (tolerance = 0.0005 Å) to constrain covalent bonds to hydrogens, using periodic boundary conditions, a 2 fs time step, a temperature of 300 K with Berendsen temperature coupling,[36] a 9 Å cutoff applied to the Lennard−Jones interactions, and a constant pressure of 1 atm using isotropic position scaling. All the simulations were performed using the particle-mesh Ewald approach to introduce long-range electrostatic effects. The nonbonded list was updated every 10 steps. Both the default scaling factors (1.2 and 2.0) and the optimized parameters (1.0 and 1.0), as given by Kirschner and Woods,[37] were used for 1−4 electrostatic and van der Waals interactions, respectively, obtaining in both cases comparable results. In all the simulations, a periodic box of TIP3P waters was extended by 10 Å in each direction from the carbohydrate atoms. Equilibration of the system was carried out as follows; as a first step, a short minimization with positional restraints on the solute atoms was run to remove any potentially bad contact. The force constant for the positional restraints was 500 kcal mol^{-1} Å. Then, a 25 ps molecular dynamics calculation was run at 300 K maintaining positional restraints on the sugar in order to equilibrate the water box. For these two steps, a 9 Å cutoff was used for the treatment of the electrostatic interactions. As a next step, the system was equilibrated (25 ps) using the particle-mesh Ewald method. Then the system was subjected to several minimizations cycles (each using 1000 steepest descent iterations) gradually reducing positional restraints on the sugar from 500 kcal mol^{-1} Å to 0. Finally, trajectory coordinates were sampled for 10 ns with a spacing of 500 fs and analyzed using the CARNAL module.

Average RDC Calculations from MD Trajectories with TRAMITE. In the TRAMITE method, the three principal axes of the alignment tensor are chosen as to correspond to the three axes of the inertia tensor (i.e., their eigenvectors are parallel).[16] The three eigenvalues of the inertia tensor, I_x, I_y, I_z, were then normalized to $I_x + I_y + I_z = 1$ to represent the probability ellipsoid for the alignment.[12] The three alignment magnitudes A_x, A_y, and A_z were obtained by subtraction of $^1/_3$ to the normalized I_x, I_y, and I_z value, to satisfy the null-trace condition $A_x + A_y + A_z = 0$. [12]

The calculation of the average RDC from an MD trajectory proceeds as follows: for each conformation in an MD trajectory, the inertia tensor is calculated and diagonalized. The three inertia tensor eigenvectors were ordered by the size of the corresponding eigenvalue $I_z \geq I_x \geq I_y$. The molecule is then rotated so as to make the principal axis system of the alignment parallel to the laboratory frame. For a conformer in this orientation, the calculation of the instantaneous RDC for any spin vector, r (r_x, r_y, r_z), is given by eq 2:[12]

$$D = \frac{k' \cdot k_{max}}{R^3}\left(\cos^2\theta \frac{3}{3}\right) = \frac{k' \cdot k_{max}}{R^3} A_x r_x^2 + A_y r_y^2 + A_z r_z^2 \qquad (2)$$

where D is the expected RDC, $\cos^2\theta$ is the angle of the vector with

(32) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(33) Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
(34) Jorgensen, W. L. *J. Am. Chem. Soc.* **1981**, *103*, 335.
(35) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

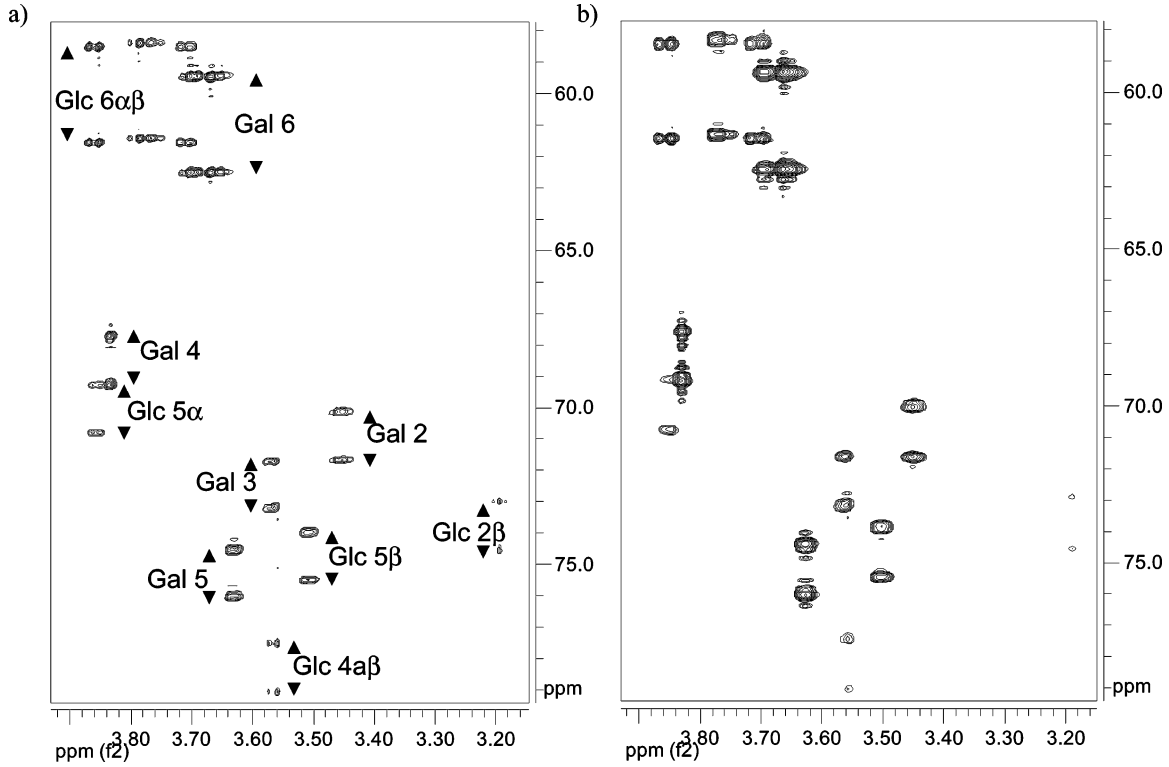


Figure 2. 2D CT-CE-HSQC used for the determination of one-bond proton-carbon RDCs in **1**. (a) D_2O nonoriented sample, (b) C12E5 oriented sample

the alignment frame, k_{max} is a constant that depends on the type of RDC, R is the C−H distance, and k' is a scaling factor, which depends on the particular system and the experimental liquid crystal concentration; this parameter was optimized for each conformation in the MD trajectory. This scaling factor was optimized for maximum agreement with the experimental RDCs. Average RDCs were calculated from the linear average of the instantaneous RDCs calculated for every single conformer in a MD trajectory. The average RDCs were calculated independently for the MD-I, MD-II, and MD-III trajectories using homemade written software available upon request.

Average RDC Calculations from MD Trajectories with PALES. The alignment tensor was also predicted from the shape of each conformation in a MD trajectory using the PALES program[13] (Phttp://www.mpibpc.gwdg.de/abteilungen/030/zweckstetter/_links/software_pales.htmP). The PALES program provided the three alignment magnitudes A_x, A_y, and A_z and performed a rotation of each conformer in such a way as to make the axes of the alignment tensor parallel to the laboratory frame. For a conformer in this orientation, the expected RDC for any spin vector, r (r_x, r_y, r_z) was calculated using eq 2.[12] Average RDCs were obtained by calculating the linear average of the RDCs predicted along the conformations of an MD trajectory. The average RDCs were calculated independently for the MD-I, MD-II, and MD-III trajectories.

Results

Previous molecular mechanics conformational studies of lactose have determined three low energy regions in the Φ/Ψ map separated by medium size energy barriers.[18−20] Interconversion is fast in the NMR chemical shift time scale. These regions are syn-Φ/syn-Ψ (region I), syn-Φ/anti-Ψ (region II), and anti-Φ/syn-Ψ (region III) (see details in the Experimental Section). In these studies, the observation of certain region specific interresidue NOEs and a full-relaxation matrix quantitative NOE analysis permitted us to determine the population of these regions in solution.[18−20] The obtained molecular mechanics and NMR data for the α- and β-methyl derivatives of lactose have shown a very similar conformational behavior for both analogues.[18−20]

Region I, characterized by the strong interresidue NOE H-1 Gal-H-4 Glc, has the largest population (>90%) in solution and is the region found in the X-ray structure of β-lactose.[38] Two weak NOEs exclusive of region II are also observed, namely H-1 Gal-H-3 Glc and H-1 Gal-H-5 Glc, that suggest the existence of a certain population in the anti-Ψ (region II) (ca. <10%). Although a third minimum (region III, anti-Φ) is predicted by molecular mechanics, there is no direct evidence of its existence in solution since the H-2 Gal-H-4 Glc NOE, exclusive of region III is not observed.[18−20] However, it has been recently reported that this is the major conformer in the gas phase.[39] Thus, and in order to obtain a different perspective, independent of that provided on the extent of flexibility of this key disaccharide, we have boarded the structural and dynamic study of lactose using residual dipolar couplings (RDCs). In particular, one-bond carbon-proton RDCs, $^1D_{CH}$, have been determined for samples of lactose oriented in cetylpyridinium chloride/hexanol/brine (CPCL) and in pentaethylene glycol mono-dodecyl ether (C12E5) oriented media, using the CT-CE-HSQC experiment (see Experimental Section).[30] These results are given in Table 1, and a relevant region of these spectra is shown in Figure 2. The Supporting Information shows the similarity of the data obtained for some of the CH vectors of the Glc units in the α- and β-anomers of lactose. In addition,

(36) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(37) Kirschner, K. N.; Woods, R. J. *Proc. Natl. Acad. Sci. U.S.A* **2001**, *98*, 10541−10545.
(38) Hirotsu, K.; Shamada, A. *Bull. Chem. Soc. Jpn.* **1974**, *47*, 1872−1879.
(39) Jockusch, R. A.; Kroemer, R. T.; Talbot, F. O.; Snoek, L. C.; Carcabal, P.; Simons, J. P.; Havenith, M.; Bakker, J. M.; Compagnon, I.; Meijer, G.; von Helden, G. *J. Am. Chem. Soc.* **2004**, *126*, 5709−5714.

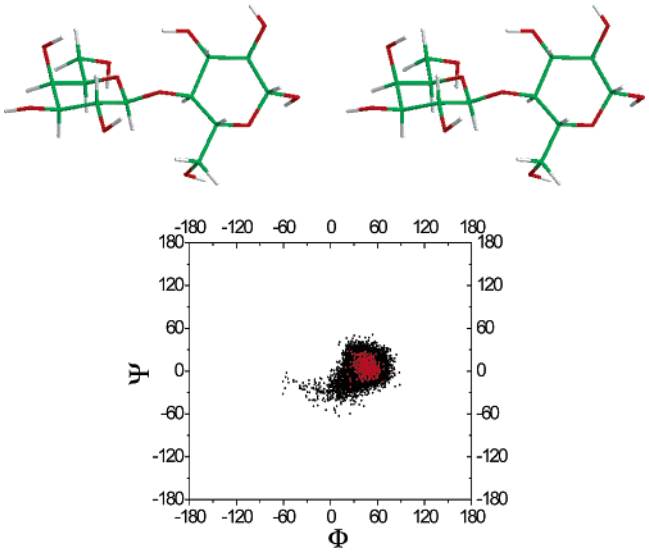


Figure 3. Scatter plot of Φ versus Ψ obtained for the 20 000 conformers obtained with simulation MD-I for compound **1** (Glc-β anomer). Data points of the 4000 with the lowest RMSD deviation from the experimental RDCs are represented in red. A stereoview of the syn-Φ/syn-Ψ conformation with the lowest RMSD deviation from the RDCs is also shown (Φ/Ψ: 49°/12.8°).

proton−proton RDCs, $^nD_{HH}$, have also been determined in both media, using a HSQC(selC,selH)-CT COSY experiment.[31] The $^nD_{HH}$ data for CPCL and C12E5 are given in Table 1.

While the direct interpretation of the experimental RDCs to determine the structural and dynamics properties can be a cumbersome task for flexible molecules with no satisfactory outcome guarantied, molecular dynamics simulations in explicit solvent (MD) can provide plausible conformational and dynamical models, which can be tested for consistency with available NMR data, such as RDCs.[15,17] Although the affordable lengths of MD simulations, currently in the nanoseconds time scale, may be not long enough to sample the slower motions such as those that may cause transitions toward other minima separated by relatively large energy barriers, MD can still be useful to sample the conformational space accessible around the proximity of a local minimum.[40] This limitation implies that the relative population of the three different local minima previously described for lactose cannot be accurately predicted from a MD trajectory. Instead, the relative population was treated in our study as an adjustable parameter in order to reproduce the experimental RDCs.

For the lactose case, 10 ns MD simulations in explicit solvent MD-I, MD-II, and MD-III were performed using as starting conformation a minimized structure within the I, II, and III regions, respectively. The analysis of the glycosidic torsions during the MD-I trajectory is represented in the scattered plot of Figure 3 for MD-I. This simulation remains around the starting conformation in region I during the complete simulation. For the other two simulations, MD-II and MD-III (Figure S1), after the first 2100 and 900 picoseconds, respectively, there was a transition toward the more stable region I, where the molecule remained during the rest of the MD trajectory.

Two different methods available to predict the alignment tensor from the shape of a neutral molecule were applied,

(40) Larsson, E. A.; Staaf, M.; Soderman, P.; Hoog, C.; Widmalm, G. *J. Phys. Chem. A* **2004**, *108*, 3932−3937

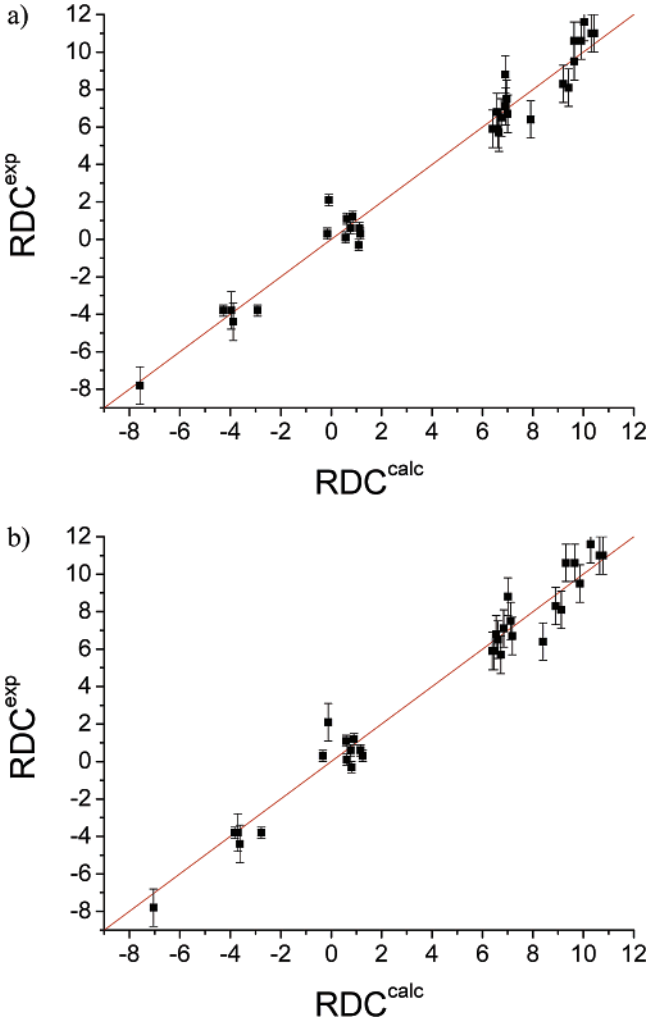


Figure 4. Plot of the experimental RDCs determined for compound **1** (Glc-β anomer) with the CPCL and C12E5 liquid crystals vs the calculated average RDCs with MD-I. (a) TRAMITE method, linear fit (R^2: 0.986). (b) PALES method, linear fit (R^2: 0.986).

PALES[13] and TRAMITE.[16] Each of these methods was used to determine the instantaneous RDCs associated with every snapshot in the MD simulations, and then the averaged RDCs were calculated by using a linear average over the MD trajectory. The calculations were independently performed for MD-I, MD-II, and MD-III. For the MD-II and MD-III trajectories, the trajectory frames belonging to region I of the Φ/Ψ map were discarded for the final averaging.

The average RDCs calculated from MD-I with either the TRAMITE or PALES methods show a good agreement with the experimental RDCs determined in both CPCL and C12E5 liquid crystals (Table 1). Indeed, both TRAMITE and PALES methods, based in the steric model, gave very similar predicted RDCs. In fact, when a given conformer of lactose was aligned using either the TRAMITE or PALES method, the corresponding structures almost superimpose with each other.

The linear fits between the calculated and experimental RDCs have a regression coefficient of 0.986 for both the TRAMITE and PALES methods (Figure 4). MD-I predicts the intrasugar $^1D_{CH}$ with a mean deviation of ±0.71 Hz for both TRAMITE and PALES methods. The intrasugar $^nD_{HH}$ are predicted with a mean deviation of ±0.73 Hz for both TRAMITE and PALES methods. The key calculated interresidue $^nD_{HH}$ between H1 Gal

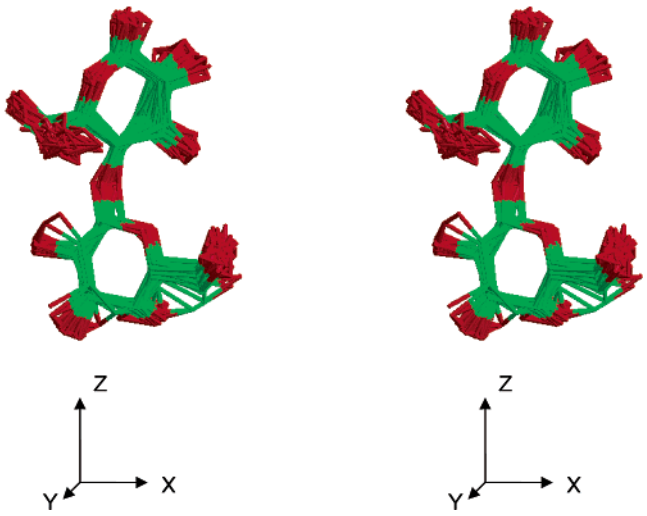


Figure 5. Trajectory frames obtained from MD-I (syn-Φ/syn-Ψ) aligned with the alignment tensor parallel to the laboratory frame using the TRAMITE method (*X*−*Z* plane).

and H4 Glc, which is extremely sensitive to the relative orientation of the sugars, is also fairly similar to the experimental one. Indeed mean deviations of ±0.70 Hz and ±0.90 Hz were obtained for the TRAMITE and PALES methods, respectively. Although this is a favorable comparison of the MD-I results with the experimental RDCs, it was found that the agreement can be improved when only a subset of 4000 conformers with the lowest RMSD is considered for the RDCs calculation (Table S1 in the Supporting Information). These conformers which are represented with red points in the map of Figure 3 are distributed around a compact rounded area of the syn-Φ/syn-Ψ map. Close to the center of this area, each single conformer has an agreement with the experimental RDC data significantly better than that in any of the averages of the MD trajectory previously considered. An example of this type of conformer has the glycosidic dihedral angles Φ/Ψ, 49°/12.8° and is represented in Figure 3. Provided that the liquid crystal media do not bias the conformational distribution, these observations could be an indication that the flexibility of the lactose is somewhat overestimated by the MD-I trajectory (and from NOE data); a more compact dynamics around the central conformer Φ/Ψ, 49°/13° would be even more consistent with the RDC data.

In any event, the favorable comparison of the MD-I results with the experimental RDCs suggests a model of the lactose with the almost exclusive presence of the energy minimum syn-Φ /syn-Ψ (>97%) around which sampling occurs in a dynamic manner (Figure 5). Attempts to improve the agreement by considering different populations of the conformers represented by MD-II and MD-III did not improve the agreement (data not shown); on the contrary, populations higher than 3% of the MD-II or higher than 1% of MD-III worsen the agreement beyond acceptable limits.

Conclusions

Previous analyses of NOE data have allowed postulating a major conformation of lactose in solution, corresponding to region I. A minor additional population in region II has also been proposed. The results presented herein complement those studies probably providing a more accurate description of the flexibility and conformational space available. The RDCs data analyzed in this work are consistent with the major presence (>97%) of minimum syn-Φ /syn-Ψ around which sampling occurs in a dynamic manner. In contrast with the NOE-based conclusions, the presence of other minor conformers did not improve the agreement between the experimental and predicted RDC datasets. Moreover, the best fit between expected and measured RDCs is provided by considering a reduced set of conformers around Φ/Ψ, 49°/13° values.

Although, from an energy difference perspective, the conformer population differences between region I and II based on NOEs (ca. 90:10) and RDCs (ca. >97:<3) are basically negligible, they provide distinct views of oligosaccharide flexibility. It is nowadays well accepted that fast motions around low energy minima geometries do indeed exist, as exemplified by any of the MD simulations discussed above, and both NOEs and RDCs are in agreement with the presence of these motions. However, the actual existence of additional and distinct energy minima (predicted by molecular mechanics calculations) seems to depend on the experimental protocol. The RDC-based conformational behavior basically provides a single populated region I, while the experimental NOEs also require the minor presence of conformers within region II. The different type of averaging could be responsible for these subtle differences. The $\langle r^{-6} \rangle$ averaging make NOEs more sensitive to the presence of minor conformers, provided that short interproton distances do exist for the key proton pairs. Other authors have claimed that the oriented medium can bias the conformational distribution in solution.[41] However, for these neutral molecules, in principle, the type of liquid crystal used should not affect the steric alignment mode. Indeed, the experimental RDCs obtained for lactose both in C12E5 and CPCL liquid are notoriously similar except for the scaling factor (Table 1); the linear regression coefficient of both experimental data sets was 0.95. Subtle differences in the alignment and/or internal dynamic averaging in both alignment media are more likely reflected in the deviations observed in the multiple bond H−H RDCs than in the one-bond H−C RDCs due to the distance and angular dependence of the former. Thus, this observation indicates that the structural information in both data sets is somewhat redundant and robust. A third explanation for the subtle differences could be related to the coupling of internal and global motions in small saccharides that make NOEs basically impossible to interpret in a quantitative manner.

In any case, the combined RDC/molecular dynamics approach used herein avoids the generation of virtual conformations that could explain the experimental data, being physically nonexistent.

The approach presented does not completely rely on the number of nondegenerate RDCs, since it is not aimed to deduce the complete structure of dynamic of the sugar from the experimental data. It implicitly assumes that one or several MDs provide the relevant geometries and dynamic of the sugar in solution except for the possible exchange between local minima separated by energy barriers. It also assumes that the sugar aligns due to a steric alignment model. The RDCs are used in this approach as a way to check for the consistency of the theoretical and experimental results in order to probe (or discard) flexible conformations. There are still certain limitations in the approach

(41) Landersjo, C.; Hoog, C.; Maliniak, A.; Widmalm, G. *J. Phys. Chem. B* **2000**, *104*, 5618−5624.

described here to deduce flexible structures using RDCs. The first one is that, for charged carbohydrates, such as the glycosylaminoglycans or sialic acid containing molecules, both the steric and electrostatic interactions with the liquid crystal modulate the alignment and the simple steric obstruction model implicit in TRAMITE and PALES for prediction of the alignment failing.[10,11,42] On this regard, an extension of the PALES method has been recently developed to consider charged solute−liquid crystal systems[41] that could be safely used. The second situation that can make the prediction of the alignment failing is when the carbohydrate incorporates a pendant lipid chain (as a *n*-butyl group), as has been recently observed.[43] In this situation, the carbohydrate may be dynamically attached to the liquid crystal and modify its average orientation.

Acknowledgment. Financial support by the DGICYT of Spain (Grant BQU2003-03550) is gratefully acknowledged. M.A.C. and F.C. thank the Ministry of Science of Spain for an FPI fellowship and Comunidad de la Rioja for a postdoctoral fellowship, respectively.

Supporting Information Available: Scatter plots of Φ versus Ψ obtained with simulations MD-II and -III for **1**. Experimental and calculated RDCs (average of 4000 structures with the lowest RMSD of MD-I aligned either with the TRAMITE or PALES method) for **1**. This material is available free of charge via the Internet at http://pubs.acs.org.

JA043445M

(42) Zweckstetter, M.; Hummerm, G.; Bax, A. *Biophys. J.* **2004**, *86*, 3444−3460.

(43) Yi, X.; Venot, A.; Glushka, J.; Prestegard, J. H. *J. Am. Chem. Soc.* **2004**, *126*, 13636−13638.