Translational Diffusion of Dilute Aqueous Solutions of Sugars as Probed by NMR and Hydrodynamic Theory

Céline Monteiro and Catherine Hervé du Penhoat*

Centre de Recherches sur les Macromolécules Végétales, CNRS (associated with University Joseph Fourier), BP 53, 38041 Grenoble Cedex 9, France

Received: April 12, 2001; In Final Form: August 3, 2001

The translational diffusion coefficients of dilute aqueous solutions of a series of carbohydrates that differ in size and branching pattern have been measured with pulsed field gradient (PFG) NMR experiments. Accurate data have been obtained by calibrating the PFG with respect to several carbohydrates of known translational diffusion. The most satisfactory results were observed with the stimulated echo experiments when the pulse sequence was kept as short as possible and transverse relaxation was accounted for. The molecular dimensions of several of these sugars were determined from optimized theoretical models previously reported in exhaustive conformational searches and compared to those evaluated from hydrodynamic theory. It was demonstrated that cylindrically shaped oligosaccharides with maximum extension about 10 times that of the water radius display classical Stokes behavior. In such cases, a combined approach including analysis of translational diffusion coefficients and molecular modeling can reveal average molecular shape. These measurements are very rapid, and such a strategy is complementary to studies based on much more time-consuming NMR probes of rotational diffusion (homo- and heteronuclear relaxation data).

Introduction

Understanding the biological roles and industrial properties of carbohydrates in aqueous solution requires a complete description not only of structure but also of molecular dynamics. Indeed, the hydrodynamic behavior of sugars is implicated in various phenomena such as hydration, gel formation, and binding to biological macromolecules. Pioneering work on the hydrodynamic properties of carbohydrates focused on the influence of concentration and temperature on the translational diffusion coefficients (D_t) of aqueous solutions of mono- and disaccharides measured by optical methods.[1−4] Deviations from Stokes−Einstein behavior were interpreted in terms of disruption of local water structure around the sugars, and an empirical relationship to the number equatorial hydroxyl groups was proposed.[5−7] In more recent years, considerable effort has been devoted to the development of tools for simulations of molecular dynamics trajectories[8−13] and for obtaining experimental probes of rotational diffusion that allow the quantitative interpretation of NMR relaxation data.[14] Only recently has there been a regain of interest in the translational diffusion of carbohydrates.

Drug design has motivated the development of diffusion-edited NMR spectroscopy (DOSY), which permits the screening of compound libraries based on the differences in observed translational diffusion coefficients, D_t, between binding and nonbinding components.[15−20] These experiments rely on the accurate measurement of the self-diffusion coefficients of small molecules with spin−echo sequences using pulsed field gradients (PFGSE).[21] Indeed, with the advent of modern NMR spectrometers equipped with units that produce well-defined pulsed-field gradients (PFG), D_t can be measured very rapidly, and in the future, widespread use of this dynamic information will prevail. To date, D_t coefficients of carbohydrates established with the PFGSE approach[13,17,22−26] have been undertaken (i) to validate the theoretical self-diffusion coefficients calculated from MD trajectories,[13,25] (ii) to demonstrate the complexation of lanthanide cations by sugars,[17] (iii) to probe the geometry of a molecular capsule formed by electrostatic interactions between oppositely charged β-cyclodextrins,[23] (iv) to study the influence of concentration and temperature dependence on the hydrodynamic properties of disaccharides,[24] and (v) to discriminate between extended and folded conformations of nucleotide−sugars.[26]

Hydrodynamic modeling of biological macromolecules has been a fruitful field of research, and formalisms are available for expressing both the rotational and translational diffusion of molecules with various distinctive molecular shapes.[27−29] Conversely, for rigid molecules of known shape, molecular dimensions can be established from experimental D_t or D_r coefficients. This approach has not been widely used in studies of medium-sized molecules because of the need of ambiguous empirical microviscosity correction factors, f.[30,31] Indeed, when the target molecule is approximately the same size as the solvent molecules (slip regime), f can vary over more than an order of magnitude ($0 < f < 1$) and depends on both molecular size and shape. However, recent conformational studies of medium-sized carbohydrates[26,32] pointed to classical Stokes behavior (stick regime in which the solvent can be treated as a continuum described by bulk viscosity), and a systematic study of this phenomenon was required.

Two major questions must be addressed to allow quantitative use of D_t in the study of carbohydrates: (1) the precision of D_t data must be established, for example, by comparing the results from PFG experiments with those from optical methods or tracer diffusion studies, and (2) the nature of carbohydrate hydrodynamic behavior (slip or stick regime) as a function of molecular size must be systematically evaluated. Molecules for which exhaustive conformational searching has been reported

* To whom correspondence should be addressed. Tel: 33-476-03-76-42. Fax: 33-476-54-72-03. E-mail: penhoat@cermav.cnrs.fr.

10.1021/jp0113836 CCC: $20.00
Published on Web 10/10/2001

1) α-Glc*p*(1→2)-β- D-Fru*f*

2) α- D-Glc*p*(1→4)-β-Glc

3) α- D-Glc*p*(1→2)-β- D-Fru*f*(1→2)-β- D-Fru*f*

4) Uridine diphosphoglucose

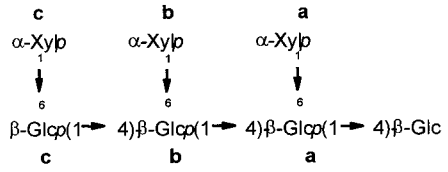
5) Guanosine diphosphomannose

6) α- D-GalA*p*(1→2)-α- D-Rha*p*(1 →4)-α- D-GalA*p*(1 →2)-α-L-Rha

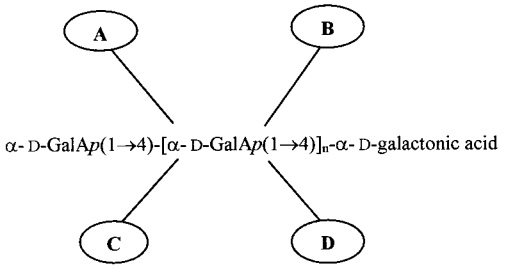
7) β- D-Glc*p*(1 →4)-β- D-Glc*p*(1→4)-β- D-Glc*p*(1→4)- D-Man

a b c d

8)



9)



A : 8-mer oligosaccharide side-chain
B : 7-mer oligosaccharide side-chain
C,D : disaccharides side chains

Figure 1. The formulas for the various carbohydrates studied in this work: (1) sucrose; (2) maltose; (3) kestose; (4) UDP-Glc; (5) GDP-Man; (6) rhamnogalacturonan I tetramer or RG I tetramer; (7) glucomannan tetramer or GGGM; (8) xyloglucan heptamer or XXXG; (9) rhamnogalacturonan II or RG II.

were chosen for this work, and overall molecular shape has been deduced from the dimensions of the corresponding low-energy conformers. Finally, molecular dimensions that have been established from D_t data and hydrodynamic theory have been compared to those of the theoretical models.

Methods

Nomenclature. A schematic representation of the carbohydrates used in this study is given in Figure 1 together with the labeling of the individual sugar residues of interest. The torsion angles have been defined as follows:

$$\Psi = \Theta(\mathrm{C1'-O1'-C}n\mathrm{-C}n{+}1)$$

$$\phi = \Theta(\mathrm{O5'-C1'-O1'-C}n)$$

$$\omega = \Theta(\mathrm{O5-C5-C6-O6})$$

The signs of the torsion angles are in agreement with the IUPAC−IUB conventions.[33]

Samples. The samples of glucose, sucrose, maltose, disodium uridine diphosphoglucose salt (UDP-Glc), and disodium guanosine diphosphomannose salt (GDP-Man) were commercially available (Sigma). 1-Kestose was a generous gift from industry. Isolation of the rhamnogalacturonan I (α-D-Gal*p*A(1→2)-α-L-Rha*p*(1→4)-α-D-Gal*p*A(1→2)-α-L-Rha[34]) and glucomannan (β-D-Glc*p*(1→4)-β-D-Glc*p*(1→4)-β-D-Glc*p*(1→4)-D-Man[35]) tetrasaccharides, the xyloglucan (XXXG[36]) heptasaccharide, and rhamnogalacturonan II[37] has been described previously.

All samples were freshly prepared by repeatedly (three times) dissolving 20 mg of carbohydrate in 0.7 mL of D_2O (99.8%, SDS) followed by evaporation of the solvent to remove exchangeable protons and then finally dissolving in 0.7 mL of D_2O (99.96%, SDS). Samples used for relaxation measurements were sealed in an NMR tube under argon after vacuum removal of dissolved oxygen. In the case of the nucleotide−sugars, 20 mg of sample was dissolved in 1 mL of a 10 mM aqueous solution of phosphate buffer containing 0.1% EDTA and 1% TSP. This mixture was lyophilized three times against D_2O (99.8%, SDS) and then dissolved in 0.5 mL of D_2O (99.96%, SDS) before vacuum removal of oxygen and sealing under argon.

NMR. ^{1}H NMR pulsed-gradient spin−echo experiments (PGSE) were conducted on either a Bruker DRX 400 or a Bruker DRX 500 (rhamnogalacturonan II) spectrometer at 298 K. Both the classical spin−echo sequence[38] and the stimulated spin−echo sequence[39] were used to measure the translational self-diffusion coefficients. The gradient duration (δ) was varied from 1 to 20 ms while keeping its strength fixed at 9.15 (9.9) G/cm for the 400 (500) MHz experiments, and the gradient recovery delay was 50 μs. The intergradient delay (Δ) was chosen to be as short as possible (1 ms for the SE experiment) while affording the maximum signal intensity for the shortest δ value. The translational self-diffusion coefficients have been obtained by fitting the intensities of a selected proton signal in spectra acquired with various lengths of the gradient pulses (8−15 data points) to the Stejskal−Tanner equation with inhouse software:

Pulse sequence A (90°-90°-90°)

$$M = {}^1/_2 M_0 \exp\{-(\tau_2 - \tau_1)/T_1 - 2\tau_1/T_2 - ((\tau_1 G\gamma)^2 D_t(\tau_2 - (\tau_1/3)))\}$$

Pulse sequence B (90°-180°)

$$M = M_0 \exp\{-2\tau/T_2 - ({}^2/_3(\delta G\gamma)^2 D_t \tau^3)\}$$

Several proton signals were monitored in separate experiments to estimate the experimental error in the estimation of D_t, and all experiments were run at least twice.

Spin−lattice relaxation times were measured with the inversion−recovery sequence. The recycle time was greater than $6T_1$, and data were collected for 10−20 τ values, which varied from 5 ms to $2T_1$. Spin−spin relaxation times were obtained with the Carr−Purcell−Meiboom−Gill sequence, and data were collected for roughly 20 spectra with various numbers of echos (total echo duration, $2n\tau$, varied from 2 ms to 1 s). The integrals of the peaks were fitted to a three- (T_1) or two-parameter (T_2) exponential function using spectrometer system software and in-house software, respectively, and all relaxation experiments were also run at least twice.

Molecular Models. Extensive molecular modeling studies have been reported for most of the molecules used in this work (sucrose,[10] maltose,[40] disodium uridine diphosphoglucose,[26] the XXXG heptasaccharide,[32] and rhamnogalacturonan II[41]). The

initial conformations for three of these molecules (sucrose, disodium uridine diphosphoglucose, the XXXG heptasaccharide) were built from the Cartesien coordinates of the global minimum. In the other cases (glucose, maltose, β-D-Glc*p*(1→4)-β-D-Glc*p*(1→4)-β-D-Glc*p*(1→4)-β-D-Man*p*, and α-D-Gal*p*A-(1→2)-α-L-Rha*p*(1→4)-α-D-Gal*p*A(1→2)-α-L-Rha*p*), starting models were constructed with the polysaccharide sequence builder of the commercially available Quanta97 package (MSI). The following protocol has been adopted to establish molecular dimensions from the molecular models. Conformers corresponding to the low-energy minima reported for each of the carbohydrates in Figure 1 were obtained by varying the Φ and Ψ torsion angles of the glycosidic linkages. In the case of the GGGM tetrasaccharide, the Φ and Ψ minima of the cellobiose disaccharide[42] were considered because molecular modeling of this tetrasaccharide has not been reported. Similarly, the Φ and Ψ values for the RG I tetrasaccharide were those described in a molecular mechanics study of a related pentasaccharide (α-L-Rha*p*(1→4)-α-D-Gal*p*A(1→2)-[α-D-Gal*p*(1→4)]-α-L-Rha*p*(1→4)-α-D-Gal*p*A[43]). The maximum atom-to-atom extension, L' (Å), and the corresponding atom-to-atom radius, r' (Å), of a hypothetical cyclinder or sphere encasing each structure were then measured for each of the geometries with the Quanta97 package.

Hydrodynamic Theory and Overall Molecular Dimensions. The Stokes−Einstein equation has been used to determine the theoretical hydrodynamic radius, r, from D_t as follows (where k is Boltzmann's constant, T is the temperature in K (298 K in this work), and η_0 is the viscosity of D_2O at 298 K):

$$D_t = kT/(6\pi r\eta_0)$$

Expressions for the translational diffusion coefficients, $D_t^{\perp}$ and $D_t^{\parallel}$, have been reported for models of short cylinders with axial ratios of $p \geq 2$ ($p = L/2r$, where L and r are the length and the radius of the cylinder, respectively):[29]

$$D_t = (D_t^{\parallel} + 2D_t^{\perp})/3$$

$$4\pi\eta_0 L D_t^{\perp}/(kT) = \ln(p) + \gamma_{\perp}$$

$$2\pi\eta_0 L D_t^{\parallel}/(kT) = \ln(p) + \gamma_{\parallel}$$

where $\gamma_{\perp} = 0.839 + (0.185/p) + (0.233/p^2)$ and $\gamma_{\parallel} = -0.207 + (0.980/p) - (0.133/p^2)$. Values of the entire range of plausible average hydrodynamic molecular overall dimensions were estimated from D_t with the above equations.

Results and Discussion

Experimental Data. Among the numerous pulse sequences that have been recommended for measuring translational diffusion coefficients with the pulsed field gradient method (for reviews, see refs 19 and 21), the three-pulse stimulated spin−echo (90°-90°-90°, STE)[39] and the classical two-pulse spin−echo (90°-180°, SE)[38] sequences have been the most often used in studies of carbohydrates. Recently, the stimulated echo sequence using bipolar gradients with a longitudinal eddy delay (BPPLED)[44] has been applied to the study of the complexation of monosaccharides by lanthanides.[17]

Several factors inherent in PFG experiments on modern spectrometers, such as eddy currents and lack of a constant gradient, limit the accuracy of the translational diffusion coefficients measured by NMR. The requirement for a constant gradient over the sample volume is not generally satisfied with commercial NMR probes. A suggested way to work around this is to measure the gradient distribution on a sample of known diffusion coefficient and then include it in the diffusion equation.[45] The effects of macroscopic background gradients and radiation damping that increase with the static magnetic field strength, $\mathbf{B}_0$, have also been discussed.[46] The use of susceptibility matched NMR tubes and reduction of the periods in which the transverse magnetization is not spatially encoded, respectively, are the recommended remedies. Other macroscopic factors related to the experimental conditions, such as convection currents,[15,47] only become important in nonviscous liquids and at high temperature. In this work, the two basic sequences, STE and SE given in Figure 2a and b, respectively, have been systematically applied to dilute aqueous solutions of a series of carbohyrates to probe the relative merits of the two approaches. Some of the aforementioned deleterious effects are partially compensated for by the gradient calibration adopted in this study and the short values of the delays used in the pulse sequences.

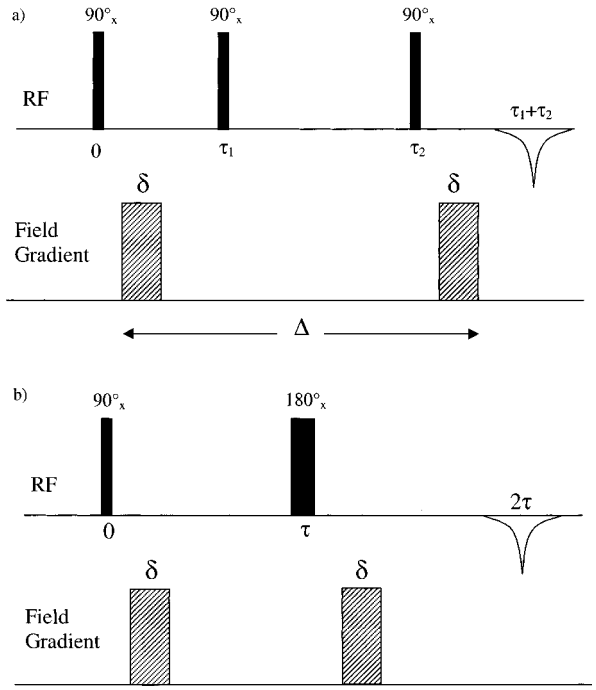


Figure 2. The pulsed field gradient sequences used to measure the translational diffusion of carbohydrates: (a) stimulated spin−echo or STE sequence; (b) classical spin−echo or SE sequence.

In PFGSE experiments, the signal attenuation can be difficult to quantify because of phase distortion resulting from J-modulation and amplitude distortion due to relaxation. To understand these effects, it is worthwhile considering the following coherence pathway that produces the STE signal:

$$\mathbf{I}_z \xrightarrow{90_x} \mathbf{I}_y \xrightarrow[\tau_1]{\mathbf{I}_z,\ \mathbf{I}_z\mathbf{S}_z} \mathbf{I}_y \xrightarrow{90_x} -\mathbf{I}_z \xrightarrow[\tau_2-\tau_1]{} -\mathbf{I}_z \xrightarrow{90_x} -\mathbf{I}_y \xrightarrow[\tau_1]{\mathbf{I}_z,\ \mathbf{I}_z\mathbf{S}_z} \mathrm{Acq}$$

During the two τ_1 intervals, the magnetization undergoes J-modulation and transverse relaxation takes place, whereas during the $\tau_2 - \tau_1$ period, longitudinal decay of magnetization occurs without J-modulation. With both the classical two-pulse SE experiment (2τ) and the other coherence pathways in the STE experiment (corresponding to the $2\tau_1$, $2\tau_2 - 2\tau_1$, and $2\tau_2$ echos), both J-modulation and transverse relaxation are operative throughout the entire spin−echo period. It can been concluded that the STE echo signal is expected to have the least phase distortion due to J-modulation and the least amplitude distortion

TABLE 1: Experimental Translational Self-diffusion Coefficients, D_t, of Various Carbohydrates Measured by NMR Spectroscopy with STE (90°-90°-90°) or SE (90°-180°) Pulse Sequences[a]

carbohydrate	number of residues[b]	NMR-defined $D_t \times 10^6$ ($cm^2\ s^{-1}$) in D_2O STE	NMR-defined $D_t \times 10^6$ ($cm^2\ s^{-1}$) in D_2O SE	Optically defined $D_t \times 10^6$ ($cm^2\ s^{-1}$) in D_2O[c]	ref[d]
H_2O	na	18.9 ± 0.4[e]	17 ± 1[e]	18.7	
glucose	1	5.6 ± 0.4[e]	4.9 ± 0.1[e]	5.49	6
sucrose	2	4.35 ± 0.08	4.6 ± 0.1	4.25	6
maltose	2	4.09 ± 0.01	4.13 ± 0.02	4.23	5
raffinose	3			3.54	6
1-ketose	3	3.75 ± 0.08[e]			
α-L-Fuc*p*-(1→2)-β-D-Gal*p*-(1→3)-D-Glucitol	3		2.7		13
UDG-Glc M^+	na	3.15 ± 0.04	3.21 ± 0.07		
GDP-Man M^+	na	3.22 ± 0.08	2.69 ± 0.09		
RG I oligomer	4	2.74 ± 0.02	2.37 ± 0.1		
GGGM	4	2.49 ± 0.04	2.34 ± 0.03		
α-L-Fuc*p*-(1→2)-β-D-Gal*p*(1→3)-β-D-GlcNAc(1→3)-β-D-Gal*p*(1→4)-D-Glc*p*	5	1.8			25
β-cyclodextrin	7			2.69	6
XXXG	7	2.35 ± 0.07			
rhamnogalacturonan II	~30	1.7 ± 0.2			

[a] Values obtained with optical methods have been given for comparative purposes. [b] na = not appropriate. [c] Taking into account the η_{D_2O}/η_{H_2O} ratio of 1.23 at 298 K. [d] Reference refers to the optical data when both NMR and optical data are given and otherwise to the NMR data. [e] The relaxation contribution to signal attenuation has not been accounted for.

TABLE 2: 400 MHz Proton Longitudinal, T_1, and Transverse, T_2, Relaxation Times of Various Carbohydrates Measured at 298 K in D_2O

carbohydrate	proton	number of residues	T_1 (s)	T_2 (s)
sucrose	H1 (α-D-Glc*p*)	2	1.02 ± 0.03	0.50 ± 0.04
maltose	H1 (terminal α-D-Glc*p*)[a]	2	0.84 ± 0.003	0.48 ± 0.02
UDG-Glc M^+	H6	na	1.12 ± 0.05	0.47 ± 0.04
GDP-Man M^+	H1′ (β-D-Rib*f*)	na	2.50 ± 0.04	0.50 ± 0.03
RG I oligomer	H1 (internal α-L-Rha*p*)	4	0.70 ± 0.006	0.36 ± 0.02
GGGM	H1 (α,β-D-Man*p*)	4	0.58 ± 0.01	0.30 ± 0.03
XXXG	H1 (α-D-Xyl*p*)[b]	7	0.933 ± 0.004	0.29 ± 0.02
rhamnogalacturonan II	H1 (A5)	~30	1.20 ± 0.02	0.21 ± 0.02

[a] Anomeric signal for the terminal residue of both α- and β-maltose. [b] Signal of all three xylose residues.

due to transverse relaxation ($2\tau > 2\tau_1$), but with a very short intergradient delay (1 ms), the difference should be small. Limiting the intergradient delay ($\tau_1 - \tau_2$) in the STE experiment results in poor separation of the stimulated echo (occurring at $\tau_2 + \tau_1$) with respect to the other echos (appearing at $2\tau_1$, $2\tau_2 - 2\tau_1$, and $2\tau_2$), and this must be accounted for by considering M_0 as an adjustable parameter in the fitting routine.

The D_t coefficients of water and 14 carbohydrates along with the standard deviations from the fitting routine have been collected in Table 1 according to increasing molecular weight. It is to be noted that the variation in D_t from two independent experiments is less than the standard deviation from the data fitting procedure. In this study, the gradient field used in the NMR diffusion measurements was calibrated indirectly by analyzing the spin−echo signal of the first four entries in Table 1 (H_2O, glucose, sucrose, maltose). The precision and accuracy of the diffusion coefficients obtained from the STE and SE experiments have been assessed by comparison to data obtained with optical methods (Gouy or Raleigh interferometry).[5−7] The optically defined diffusion coefficients were determined in H_2O and conversion to the corresponding value in D_2O[48] was based on the viscosity ratio ($\eta_{D_2O}/\eta_{H_2O} \approx 1.23$). The data for the HOD signal (PFGSE experiments) were not corrected for isotope effects (<2%).[49] The average deviations between the optically and STE- or SE-defined D_t values for these four reference compounds are ±2% and ±10%, respectively. In the remaining discussion, we will only refer to the best experimental data set, namely, the D_t values obtained with the STE sequence.

From a qualitative point of view, perusal of Table 1 shows that the D_t values decrease fairly regularly with molecular size (i.e., for linear carbohydrates containing one to five sugar residues). Two entries from the literature (a trisaccharide[13] and a linear pentasaccharide[25]) have been included to demonstrate how robust the correlation of D_t data of dilute aqueous solutions of linear carbohydrates with the number of sugar residues really is. However, closer examination of the data for the larger nonlinear carbohydrates suggests that molecular volume is as important as the number of constituent residues. The D_t values of both the branched xyloglucan heptasaccharide with a tetrasaccharide mainchain (XXXG, 2.35×10^{-6} $cm^2\ s^{-1}$) and β-cyclodextrin (a D_t of 2.69×10^{-6} $cm^2\ s^{-1}$)[6] are analogous to those of the two tetrasaccharides (2.74×10^{-6} and 2.49×10^{-6} $cm^2\ s^{-1}$). Thus, at first glance, the molecular volume of the carbohydrates would seem to offer the best correlation with the D_t values, and to test this hypothesis, optimized theoretical model structures were considered (vide infra).

The nonselective longitudinal, T_1, and transverse, T_2, relaxation times of the majority of the carbohydrates investigated in this study have been given in Table 2. In many cases, the contribution of relaxation to the attenuation of the spin−echo signal in diffusion measurements of carbohydrates has been ignored. As regards the present work, the error introduced by neglect of the T_1 and T_2 contributions to attenuation is small for disaccharides (4−5%), moderate for nucleotide sugars (8%) but very substantial for larger carbohydrates (i.e., 50% for the RG II 30-mer, Table 1). It is known that, in the case of a good signal-to-noise ratio, accurate D_t values can still be obtained when the diffusion effect represents as little as 20% of the signal attenuation as long as relaxation is accounted for.[21] However, for high molecular-weight polysaccharides, T_2 would obviously

TABLE 3: Comparison of the Overall Molecular Dimensions[a] of the Carbohydrates Derived from the STE D_t Measurements and Hydrodynamic Theory with Those Established from Theoretical Models[b]

carbohydrate	molecular dimensions from PFG STE NMR and hydrodynamic theory; r or L/r (p) (Å)	average molecular dimensions from molecular modeling, r or L/r (p) (Å)	variations in molecular dimensions from molecular modeling, Δr or $\Delta L/\Delta r$ (Δp) (Å)	number of minima from modeling studies (ref)
glucose	3.7	4.6	0.1	2
sucrose	4.6	5.9 (1.1)	0.8	4 (10)
maltose	5.0	6.4 (1.4)	1.3	6 (40)
UDP-Glc	15.5/4.5 (1.7)	16.5/4.6[c](1.8)	3/0 (0.5)	7 (26)
model RG1 tetrasaccharide	18/5 (1.8)	18.5/5.0 (1.9)	1.6/0.2 (0.1)	5 (43)
glucomannan tetrasaccharide, GGGM	23/4.5 (2.6)	23.6/4.6 (2.6)[d]	0/0 (0)	4 (42)
xyloglucan heptasaccharide, XXXG	19.3/6.5 (1.5)	22.4/6.5 (1.7)	3.7/0.3 (0.2)	5 (32)

[a] r = radius in Å; L = cylinder length in Å; p = axial ratio $L/2r$. [b] Defined by the corresponding Φ,Ψ values. The atom-to-atom dimensions (L', r') have been adjusted ($\Delta L = +2.3$ Å; $\Delta r = +1.2$ Å) for comparison with the hydrodynamic ones as described in the text. [c] Average value of L for 3-ns molecular dynamics trajectories in explicit water (the r value is that of the glucose monomer). [d] Identical values were obtained for low-energy conformers with a cylindrical shape, while the least favorable minima presented an almost spherical C shape with a radius of 7.4 Å.

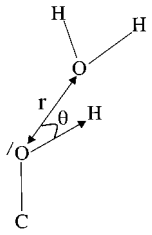


Figure 3. Geometric parameters (θ,d_{O_s/O_w}) that describe the minimal distance between sugar OH groups and the first layer of water.

be the limiting factor in the measurement of translational diffusion.

Molecular Dimensions from Theoretical Models. As described in the Methods Section, the ranges of plausible molecular dimensions for the various carbohydrates in Table 3 have been established from data published for theoretical models, and the number of low-energy conformers that have been considered has been indicated in the last column. The largest atom-to-atom distance, L' (Å), was determined for the low-energy conformers followed by the largest atom-to-atom distance in the direction perpendicular to L', which is considered to be the atom-to-atom radius, r' (Å), of a hypothetical cylinder encasing the molecule. These molecular descriptors were chosen to allow comparison of D_t-defined molecular dimensions with the corresponding theoretical values for cylindrical or spherical models of overall molecular shape.[29] Much more sophisticated hydrodynamic models might be more appropriate for some of the carbohydrates in Table 3, but these two suffice for our purposes. Average values for these molecular dimensions (all low-energy conformers were considered to have the same population) and the maximum variations in the atom-to-atom geometric parameters ($\Delta r'$, $\Delta L'$, $\Delta p'$) for the energy minima structures were systematically evaluated.

Rigorous treatment of hydration is beyond the scope of this study, but an estimate of the distance between the peripheral hydroxyl hydrogens of the carbohydrates and the solvent is necessary for comparison of the hydrodynamic (r, L) and atom-to-atom (r', L') dimensions. State-of-the-art MD trajectories[8−10,12] include pair distribution functions, which generally place the oxygens of first layer of water (O_w) at roughly 2.7−2.8 Å from the sugar oxygens (O_s). The definitions of the variables that determine the O_w-O_s distance are illustrated in Figure 3. The value of θ is zero (i.e., $O_s-H_s-O_w$ are colinear) when the sugar hydroxy group acts as a proton donor in forming a hydrogen bond with the water molecule. This situation corresponds to the shortest H_s-O_w distance possible, and as the O_s-H_s bond length is close to 0.95 Å, it is approximately 1.8 Å. For the sake of comparison, the upper limit for the O_w-H_s distance will be assumed to be that of the average O_w-O_s distance for the first water layer in the MD simulations. In the following discussion, r' and L' are increased by, respectively, 50% and 100% of the average O_w-H_s distance (i.e., 1.2 and 2.3 Å, respectively) for comparison with the hydrodynamic r and L parameters, and only these latter values will be referred to.

A continuum of molecular shapes from perfectly spherical ($p = 1$, where p is the axial ratio and $p = L/2r$) to cylindrical ($p > 1$) is found by fitting the experimental D_t coefficients to theoretical ones. In Table 3, the results for spherical models have been given when the p values of the most favorable conformers were less than 1.5, and data for cylindrical models have been proposed when the p values of the optimized geometries were greater than 1.5. An independent experimental estimation of the anisotropy of molecular shape can be obtained from NMR relaxation data. In the case of carbohydrates, significant variations in methine carbon T_1 values and heteronuclear Overhauser effects, η_{C-H}, (>10%) often indicate anisotropic rotational diffusion.[50] Such experimental data can be fitted with appropriate spectral densities while considering the rotational correlation times as adjustable parameters. This leads to an NMR-defined motional model for rotational diffusion that can also be expressed in terms of molecular dimensions and compared to the D_t-defined dimensions.

It has been demonstrated by many NMR investigations that the conformation of glucose, the only monomer in Table 3, is adequately described by the classical $^4C_1$ chair form of the glucopyranose ring[51] and gauche−gauche and gauche−trans orientations of the exocyclic group.[52] The estimated hydrodynamic radius (4.6 Å) of the molecular models is larger (124%) than the one established from the translational diffusion coefficient.

As regards the disaccharides, extensive carbon NMR relaxation data have shown that the overall shape of sucrose is almost spherical as only very small variations (<5%) in the carbon T_1 values and the heteronuclear Overhauser effects of the methine carbons have been observed.[10,53] An exhaustive molecular mechanics study[54] and molecular dynamics simulations of sucrose in explicit solvent[10] have revealed four stable low-energy structures that were all examined to establish the average 5.9-Å hydrodynamic radius. Only small variations in molecular dimensions (<1 Å) are predicted for the various stable conform-

ers. In the case of maltose, carbon T_1 values[55] point to a less spherical overall shape ($\Delta T_1 = 18\%$ and 26%, respectively, for α- and β-maltose). The average hydrodynamic radius of 6.4 Å was evaluated by considering the six major conformational families that have been described in a molecular mechanics investigation[40] of this disaccharide. Estimates of the axial ratios for hypothetical cylinders encasing these sugars (1.1 and 1.4 for sucrose and maltose, respectively) corroborate a spherical shape for sucrose, whereas this shape is expected to be less satisfactory for maltose. The dimensions calculated from the D_t coefficients are again larger (128% for both sugars) than the experimental ones.

Extensive carbon relaxation data has been reported for the UDP-Glc nucleotide sugar,[56] and here, the large variations in the multi-field methine carbon relaxation parameters indicate anisotropic overall shape. Simulations of these data in terms of rotational diffusion have afforded hydrodynamic dimensions of 15.2 and 4.0 Å, respectively, for L and r. At least nine stable conformational families have been revealed through exhaustive molecular mechanics and molecular dynamics simulations in the presence of explicit solvent and counterions.[26] The average extensions (L) for seven 3-ns trajectories were also reported, and the corresponding radius (r) measured for the dominant conformer of each trajectory was very close to that of the glucose monomer, and the corresponding hydrodynamic dimensions are indicated in Table 3. The agreement between the D_t-defined (L/r 15.5/4.5) and D_r-defined dimensions is excellent, and these dimensions are only slightly smaller ($\Delta L \approx 1$ Å, $\Delta r \approx 0.1$ Å) than the average ones determined from the MD trajectories. Because of the considerable flexiblity of the pyrophosphate linkage (four more torsion angles than a six-linked disaccharide), larger variations in overall extension ($\Delta L > 3$ Å) are predicted for UDP-Glc models.

In the molecular mechanics study of a pentasaccharide (α-L-Rha*p*(1→4)-α-D-Gal*p*A(1→2)-[α-D-Gal*p*(1→4)]-α-L-Rha*p*(1→4)-α-D-Gal*p*A)[43] related to the RG I tetrasaccharide, the potential energy surfaces (Φ,Ψ maps) of all the corresponding disaccharide fragments were reported. The Φ,Ψ values of the energy minima were used to build five geometries of the RG I tetrasaccharide that served to establish its molecular dimensions. All conformers displayed axial ratios close to 2 indicating that a cylindrical overall shape would be appropriate for this tetrasaccharide, but experimental data corroborating this shape are lacking. The fit between the D_t-defined dimensions (18/5 L/r) and the average hydrodynamic dimensions of the minima structures (18.5/5) is very good.

The glucomannan tetrasaccharide has not been investigated with theoretical methods, but the cellobiose disaccharide has been extensively studied.[42] Its Φ,Ψ map contains a large low-energy region (the region containing the Φ,Ψ values of most crystal structures; two Φ,Ψ pairs in this region at 40,−40 and 50,−10 were considered) and two less favorable narrow wells (180,0 and 30,180). When the Φ,Ψ values of the global minimum region were used to construct GGGM conformers, extended cylindrical geometries resulted. The D_t-defined hydrodynamic dimensions (L/r 23/4.5) are almost identical to the ones calculated for the average structure of the molecular models in the case of the tetrasaccharides ($\Delta L = 0.6$ Å, $\Delta r = 0.1$ Å). Some much more folded conformers were reported, and the overall shape (that of the letter C) was best represented by a sphere ($r = 7.4$ Å for the molecular models and $r = 8.0$ Å for a spherical model according to hydrodynamic theory). However, these compact spherical GGGM conformers are located in the least-favorable energy well arguing against a significant population.

Finally, as regards the xyloglucan heptasaccharide, XXXG, 25 low-energy conformers have been described in an exhaustive search of its conformational space.[32] The five most favored geometries were examined to establish the average molecular dimensions that correspond to a fairly anisotropic shape ($1.7 < p < 1.9$). The xyloglucan STE-defined molecular dimensions are too small (L/r 19.3/6.5) when compared to those of the average cylindrical model (L/r 22.4/6.5), whereas they are in good agreement with the dimensions of the most folded low-energy structure (L/r 19.6/6.3), which turns out to be the global minimum of the conformational search.[32]

Stokes−Einstein Behavior of Carbohydrates. In the case of mono- (glucose) and disaccharides (sucrose and maltose), the STE-defined radii are only about 75% of those established for theoretical models indicating the necessity for a microviscosity correction factor (slip regime) in accordance with earlier reports.[5−7] However, as regards the two tetrasaccharides, STE-defined translational diffusion obeys the Stokes−Einstein relation, and very good agreement ($\Delta L \leq 1$ Å when Δr is set to ∼0) is obtained between the experimentally defined molecular dimensions for cylindrical shape and those of the theoretical models.

In the case of the majority of the carbohydrates in Table 3, small variations in molecular dimensions are observed for the various low-energy conformers ($\Delta L < 1.6$ Å), whereas both UDP-Glc and the XXXG heptasaccharide are much more flexible ($\Delta L > 3$ Å). The experimental error associated with the D_t measurement of these latter compounds corresponds to a ΔL variation of 0.7 Å so that an indication of the more populated conformational families could be obtained from the D_t coefficient. For UDP-Glc, the average structure from nine MD trajectories (27 ns) in explicit water reproduces the translational diffusion coefficient, whereas the xyloglucan heptasaccharide is best described by the global minimum.

Conclusions

Hydrodynamic modeling has not been fully exploited in the case of carbohydrates because of the need for ambiguous correction factors in the diffusion equations. Generally speaking, for spherical models, microviscosity correction factors only become negligeable when the ratio of the solvent radius to that of the solute is about 0.01.[30] Because the radius of the water molecule is about 1.7 Å,[31] none of the carbohydrates in Table 3 would be expected to display classical Stokes behavior. However, the results in Table 3 show that the stick regime applies for cylindrically shaped oligosaccharides when the maximum extension of the carbohydrate is about 10 times larger than the water radius. The definition of the hydrodynamic dimensions of carbohydrates as outlined in Figure 3 may appear somewhat arbitrary. However, this model is corroborated by a recent study[55] of the hydration of sugars based on MD trajectories in the presence of explicit solvent and D_t measurements of water for dilute solutions of disaccharides. It was demonstrated that the experimental water translational diffusion coefficient (which includes the 30 molecules in the first hydration shell around the sugar) is identical to that of bulk water.

The quantitative analysis of hydrodynamic parameters in terms of molecular dimensions and, conversely, the prediction of D_r or D_t coefficients from molecular models are of considerable interest. Indeed, D_t measurements represent another type

of experimental probe of molecular volume and, as such, are complementary to the arsenal of NMR methods routinely used in the conformational analysis of carbohydrates. It is not possible to distinguish between spherical and cylindrical models on the basis of the D_t coefficient alone. Average molecular shape must be determined from experimental probes of rotational diffusion (such as a spread in the NMR relaxation parameters of the methine carbons) or from an exhaustive molecular modeling study. In the case of very flexible molecules, the translational diffusion coefficient can help discriminate between conformers with distinctly different volumes. To this end, a 2-fold reduction in the experimental error associated with the D_t measurements (a 1% error has been achieved with optimized equipment and pulse sequences[46]) would afford a precision in average molecular extension of better than 0.5 Å for moderately sized oligosaccharides.

Quantitative analysis of NMR relaxation data requires a motional model for rotational diffusion, and in the case of carbohydrates of sufficient size (degree of polymerization ≥ 4), this dynamic information could be rapidly established from molecular dimensions of model structures validated by D_t measurements. Finally, secure prediction of D_t coefficients would facilitate the design of DOSY experiments in studies of substrate/receptor interactions with carbohydrate ligand libraries, and this will be a very active field of research in the coming years.

Acknowledgment. C.M. was supported by a grant from the French Ministère de l′Enseignement Supérieur et de la Recherche.

References and Notes

(1) Creeth, J. M. *J. Am. Chem. Soc.* **1955**, *77*, 6428−6440.
(2) English, A. C.; Dole, M. *J. Am. Chem. Soc.* **1950**, *72*, 3261−3267.
(3) Gladden, J. K.; Dole, M. *J. Am. Chem. Soc.* **1952**, *75*, 3900−3904.
(4) Schliephake, D.; Schneider, F. *Zucker* **1965**, 138−142.
(5) Uedaira, H.; Uedaira, H. *Bull. Chem. Soc. Jpn.* **1969**, *42*, 2140−2142.
(6) Uedaira, H.; Uedaira, H. *J. Phys. Chem.* **1970**, *74*, 2211−2214.
(7) Uedaira, H.; Uedaira, H. *J. Solution Chem.* **1985**, *14*, 27−34.
(8) Brady, J. W. *J. Am. Chem. Soc.* **1989**, *111*, 5155−5165.
(9) Brady, J. W.; Schmidt, R. K. *J. Phys. Chem.* **1993**, *97*, 958−966.
(10) Engelsen, S. B.; Hervé du Penhoat, C.; Pérez, S. *J. Phys. Chem.* **1995**, *99*, 13334−13350.
(11) Kaufmann, J.; Möhle, K.; Hofmann, H.-J.; Arnold, K. *Carbohydr. Res.* **1999**, *318*, 1−9.
(12) Sakurai, M.; Murata, M.; Inoue, Y.; Hino, A.; Kobayashi, S. *Bull. Chem. Soc. Jpn.* **1997**, *70*, 847−858.
(13) Widmalm, G.; Venable, R. M. *Biopolymers* **1994**, *34*, 1079−1088.
(14) Dais, P. *Adv. Carbohydr. Chem. Biochem.* **1995**, *51*, 63−131.
(15) Augé, S.; Amblard-Blondel, B.; Delsuc, M.-A. *J. Chim. Phys.* **1999**, *96*, 1559−1565.
(16) Chen, A.; Shapiro, M. J. *J. Am. Chem. Soc.* **1998**, *120*, 10258−10259.
(17) Diaz, M. D.; Berger, S. *Carbohydr. Res.* **2000**, *329*, 1−5.
(18) Hajduk, P. J.; Olejniczak, E. T.; Fesik, S. W. *J. Am. Chem. Soc.* **1997**, *119*, 12257−12261.
(19) Johnson, C. S. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 203−256.
(20) Waldeck, A. R.; Kuchel, P. W.; Lennon, A. J.; Chapman, B. E. *Prog. Nucl. Magn. Reson. Spectrosc.* **1997**, *30*, 39−68.
(21) Stilbs, P. *Prog. Nucl. Magn. Reson. Spectrosc.* **1987**, *19*, 1−45.
(22) Hall, L. D.; Luck, S. D. *Carbohydr. Res.* **1984**, *134*, C1−C3.
(23) Hamelin, B.; Jullien, L.; Derouet, C.; Hervé du Penhoat, C.; Berthault, P. *J. Am. Chem. Soc.* **1998**, *120*, 8438−8447.
(24) Rampp, M.; Buttersack, C.; Lüdemann, H.-D. *Carbohydr. Res.* **2000**, *328*, 561−572.
(25) Rundlöf, T.; Venable, R. M.; Pastor, R. W.; Kowalewski, J.; Widmalm, G. *J. Am. Chem. Soc.* **1999**, *121*, 11847−11854.
(26) Petrova, P.; Monteiro, C.; Hervé du Penhoat, C.; Koca, J.; Imberty, A. *Biopolymers* **2001**, *58*, 617−635.
(27) de la Torre, J. G.; Bloomfield, V. A. *Q. Rev. Biophys.* **1981**, *14*, 81−139.
(28) Tirado, M. M.; de la Torre, J. G. *J. Chem. Phys.* **1979**, *71*, 2581−2586.
(29) Tirado, M. M.; Martinez, C. L.; de la Torre, J. G. *J. Chem. Phys.* **1984**, *81*, 2047−2052.
(30) Boeré, R. T.; Kidd, R. G. Rotational Correlation Times in Nuclear Magnetic Resonance. In *Annual Reports on NMR Spectroscopy*; Webb, G. A., Ed.; Academic Press: London, 1982; Vol. 13, pp 319−385.
(31) Easteal, A. J. *J. Chem. Eng. Data* **1996**, *41*, 741−744.
(32) Picard, C.; Gruza, J.; Derouet, C.; Renard, C. M. G. C.; Mazeau, K.; Koca, J.; Imberty, A.; Hervé du Penhoat, C. *Biopolymers* **2000**, *54*, 11−26.
(33) IUPAC−IUB *Arch. Biochem. Biophys.* **1971**, *145*, 405−421.
(34) Renard, C. M. G. C.; Lahaye, M.; Mutter, M.; Voragen, A. F. G.; Thibault, J. F. *Carbohydr. Res.* **1998**, *305*, 271−280.
(35) Goldberg, R.; Gillou, L.; Prat, R.; Hervé du Penhoat, C.; Michon, V. *Carbohydr. Res.* **1991**, *210*, 263−276.
(36) Renard, C. M. G. C.; Lomax, J. A.; Boon, J. J. *Carbohydr. Res.* **1992**, *232*, 303−320.
(37) Pellerin, P.; Doco, T.; Vidal, S.; Williams, P.; Brillouet, J.-M.; O'Neill, M. A. *Carbohydr. Res.* **1996**, *290*, 183−197.
(38) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288−292.
(39) Tanner, J. E. *J. Chem. Phys.* **1970**, *52*, 2523−2526.
(40) Tran, V.; Buleon, A.; Imberty, A.; Pérez, S. *Biopolymers* **1989**, *28*, 679−690.
(41) Mazeau, K.; Pérez, S. *Carbohydr. Res.* **1998**, *311*, 207−217.
(42) French, A. D.; Dowd, M. K. *J. Mol. Struct.: THEOCHEM* **1993**, *286*, 183−201.
(43) Broadhurst, M.; Cros, S.; Hoffmann, R.; Mackie, W.; Pérez, S. Modelling a pentasaccharide fragment of rhamnogalacturonan I. In *Pectins and Pectinases*; Visser, J., Voragen, A. G. J., Eds.; Elsevier Science: New York, 1996.
(44) Wu, D.; Chen, A.; Johnson, C. S. J., Jr. *J. Magn. Reson. A* **1995**, *115*, 123−126.
(45) Damberg, P.; Jarvet, J.; Gräslund, A. *J. Magn. Reson.* **2001**, *148*, 343−348.
(46) Price, W. S.; Stilbs, P.; Jönsson, B.; Söderman, O. *J. Magn. Reson.* **2001**, *150*, 49−56.
(47) Sorland, G. H.; Seland, J. G.; Krane, J.; Anthonsen, H. W. *J. Magn. Reson.* **2000**, *142*, 323−325.
(48) Matsunga, N.; Nagashima, A. *J. Phys. Chem. Ref. Data* **1983**, *12*, 933−966.
(49) Mills, R. *J. Phys. Chem.* **1973**, *77*, 685−688.
(50) Dais, P. *Adv. Carbohydr. Chem. Biochem.* **1995**, *51*, 63−131.
(51) Altona, C.; Haasnoot, C. A. G. *Org. Magn. Reson.* **1980**, *13*, 417.
(52) Bock, K.; Duus, J. O. *J. Carbohydr. Chem.* **1994**, *13*, 513−543.
(53) McCain, D. C.; Markley, J. L. *J. Am. Chem. Soc.* **1986**, *108*, 4259−4264.
(54) Casset, F.; Imberty, A.; Hervé du Penhoat, C.; Koca, J.; Pérez, S. *J. Mol.Struct.: THEOCHEM* **1997**, *395*−*396*, 211−224.
(55) Engelsen, S. B.; Monteiro, C.; Hervé du Penhoat, C.; Pérez, S. *J. Biophys. Chem.* **2001**, in press.
(56) Monteiro, C.; Neyret, S.; Leforestier, J.; Hervé du Penhoat, C. *Carbohydr. Res.* **2000**, *329*, 141−155.