Available online at www.sciencedirect.com
SCIENCE DIRECT®
Carbohydrate Research 339 (2004) 87–96
Carbohydrate RESEARCH

Use of ^{13}C chemical shift surfaces in the study of carbohydrate conformation. Application to cyclomaltooligosaccharides (cyclodextrins) in the solid state and in solution

Edward P. O'Brien and Guillermo Moyna*

Department of Chemistry and Biochemistry, University of the Sciences in Philadelphia, 600 South 43rd Street, Philadelphia, PA 19104-4495, USA

Received 16 July 2003; accepted 24 September 2003

Abstract—The anomeric carbon chemical shifts of free cyclomaltohexaose, -heptaose, -octaose, -decaose, and -tetradecaose (α-, β-, γ-, ε-, and ι-cyclodextrin, respectively), and of α-cyclodextrin inclusion complexes, both in the solid state and in solution, were computed using ab initio ^{13}C chemical shift surfaces for the D-Glc*p*-α-(1→4)-D-Glc*p* linkage as a function of the glycosidic bond $\langle \Phi, \Psi \rangle$ dihedral angles. Chemical shift calculations in the solid state used $\langle \Phi, \Psi \rangle$ angle pairs measured from cyclodextrin X-ray structures as input. For estimations in the liquid state two different approaches were employed to account for dynamic averaging. In one, the computed solid-state anomeric carbon chemical shifts for each cyclodextrin D-Glc*p* monomer were simply averaged to obtain an estimate of the ^{13}C shifts in solution. In the other, chemical shifts for the anomeric carbons were determined by averaging back-calculated ^{13}C shift trajectories derived from a series of 5 ns molecular dynamic simulations for the oligosaccharides with explicit representation of water. Good agreement between calculated and experimental ^{13}C shifts was found in all cases. Furthermore, our results show that the ab initio ^{13}C chemical shift surfaces are sufficiently sensitive to reproduce the small variations observed for the anomeric ^{13}C shifts of the different cyclodextrin D-Glc*p* units in the solid state with excellent accuracy. The use of chemical shift surfaces as tools in conformational studies of oligosaccharides is discussed.

Keywords: ^{13}C NMR; ab initio; Carbohydrate conformation; Chemical shift surfaces; Cyclodextrins; GIAO; Molecular dynamic simulations

1. Introduction

Oligosaccharides have arguably the most varied functions among the different classes of biological macromolecules. While their roles range from energy storage in living organisms to the control of cellular structure and shape, the most important biochemical process involving oligosaccharides is cellular signaling.[1] This is not surprising if one considers that carbohydrates, as opposed to amino acids and nucleic acids, have the possibility of forming highly branched polymers in which the units can be coupled through several types of glycosidic linkages. This rich structural diversity allows carbohydrate-containing macromolecules to present a virtually infinite number of signals to their surroundings, thus making them particularly well suited to participate in molecular recognition processes.[1] Regarding their three-dimensional (3D) structure and dynamics, oligosaccharides are also markedly different from other biopolymers. Despite there are almost 22,000 protein 3D structures available in the Protein Data Bank (PDB), only a handful of these correspond to glycoproteins.[2] Even if this small subgroup of entries is considered, the structure of the protein-bound oligosaccharide moiety is usually found to be ill-defined. Similarly, a survey of the Cambridge Structural Database (CSD) reveals that the number of linear or branched oligosaccharides for which high-resolution X-ray structures are available amounts

* Corresponding author. Tel.: +1-215-596-8526; fax: +1-215-596-8543; e-mail: g.moyna@usip.edu

0008-6215/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.carres.2003.09.034

to less than a hundred for disaccharides, trisaccharides, and tetrasaccharides combined.[3] These observations confirm the well-known ability of these molecules to adopt a wide range of conformations in dynamic equilibrium in solution.[1,4] Therefore, understanding the mechanism through which oligosaccharides participate in molecular recognition processes requires knowledge of their conformational behavior in solution.

From the variety of methods employed in the study of oligosaccharide conformation and dynamics, nuclear magnetic resonance (NMR) spectroscopy and molecular modeling simulations are perhaps the most valuable, and a number of protocols that employ a combination of the two techniques have been postulated.[5,6] In most of these approaches, estimations of internuclear distances between oligosaccharide proton pairs are obtained using 1D and/or 2D nuclear Overhausser enhancement (NOE) experiments, and information regarding the glycosidic bond Φ and Ψ dihedral angles is derived from homo- and heteronuclear vicinal (^3J) and geminal (^2J) coupling constants. Recently, the measurement of residual dipolar couplings in weakly aligned oligosaccharide samples has been employed to determine the relative orientation of C–H bond vectors.[7] Irrespective of the type of experiment, the information obtained from solution-state NMR spectroscopy is normally affected by conformational flexibility, and a single atomic model satisfying all the experimental data measured for an oligosaccharide cannot be obtained.[1,5–9] In order to account for these dynamical averaging effects, conformer ensembles, which can reproduce, on average, the experimental observations, are generated through the use of molecular modeling simulations such as Monte Carlo searches, molecular dynamics (MD), or simulated annealing (SA).[5–9]

Carbohydrate ^{13}C chemical shifts have also been recognized as valuable sources of structural information that could be used to complement or in some cases, as discussed in further detail below, replace NOE and J-coupling data. In particular, the anomeric carbon chemical shift is known to depend systematically with the glycosidic bond $\langle \Phi, \Psi \rangle$ dihedral angles,[10,11] and several experimental and theoretical studies have explained this relationship qualitatively.[12–15] Recently,[16] we revisited these original investigations and carried out an exhaustive ab initio study of the anomeric carbon chemical shift dependence versus Φ and Ψ for D-Glc*p*–D-Glc*p* disaccharides models with $(1\rightarrow1)$, $(1\rightarrow2)$, $(1\rightarrow3)$, and $(1\rightarrow4)$ linkages in both α- and β-configurations. Our work lead to the development of chemical shift surfaces for carbohydrates,[16] empirical functions relating anomeric carbon chemical shift with glycosidic bond conformation. These equations, of the form ^{13}C $\delta_{\text{anom}} = f(\Phi, \Psi)$, can be used to quantitatively incorporate isotropic ^{13}C shift measurements into the analysis of oligosaccharide 3D structure and dynamics.

In this report we study the applicability of ^{13}C chemical shift surfaces to the study of carbohydrate conformation, using the surface corresponding to the D-Glc*p* α-$(1\rightarrow4)$ linkage and α-, β-, γ-, ε-, and ι-cyclodextrin (CD) as model systems (Figs. 1 and 2). The results presented below show that ^{13}C estimations obtained from theoretical chemical shift surfaces reproduce experimental observations well, both in the solid state as well as in solution if conformational averaging is taken into account, and serve to demonstrate the usefulness of the methodology.

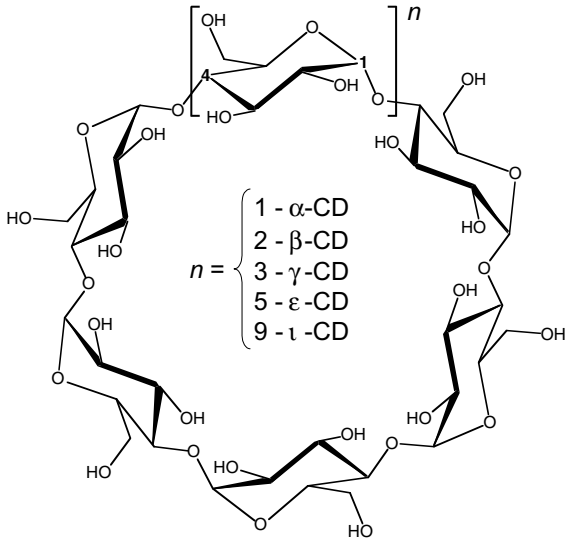

Figure 1. Structures of the cyclomaltooligosaccharides α-, β-, γ-, ε-, and ι-CD.

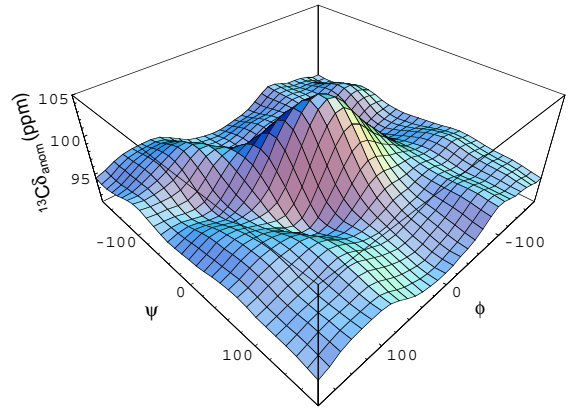

Figure 2. ^{13}C chemical shift surface for the anomeric carbon of the D-Glc*p* α-$(1\rightarrow4)$ glycosidic linkage as a function of the Φ and Ψ dihedral angles, defined as ⟨H-1–C-1–O-1–C-4′⟩ and ⟨H-1–C-1–O-1–C-4′⟩, respectively.

2. Computational methodology

2.1. General considerations

Molecular models for free α-, β-, γ-, ε-, and ι-CD, and for α-CD inclusion complexes were built from their corresponding X-ray coordinates found in the CSD (Tables 1, 2, and 5).[17–22] Sybyl 6.8 (Tripos Associates, Inc.) and AMBER 6.0 were used to convert CSD coordinate files into PDB files and to add hydrogens in idealized positions to structures bearing only heavy-atom coordinates, respectively. Solution ^{13}C NMR data for the CD systems considered in this study, as well as solid-state cross polarization magic angle spinning (CP/MAS) ^{13}C NMR data for free α-, β-, and γ-CD, and for α-CD inclusion complexes were obtained from the literature.[22–27]

2.2. Generation of the D-Glc*p*-α-(1 → 4)-D-Glc*p* chemical shift surface

The derivation of the chemical shift surface for the anomeric carbon of the D-Glc*p* α-(1 → 4) linkage has been presented in detail elsewhere,[16] and is briefly described here. An 18 × 18 matrix in the (−180° ↔ 180°) range for Φ and Ψ at 20° intervals for the D-Glc*p*-α-(1 → 4)-D-Glc*p* disaccharide was constructed, resulting in grid with 324 conformers. For each structure in the grid, the Φ and Ψ angles were kept constant with dihedral constraints on heavy atoms and the monosaccharide held in the $^4C_1$ conformation, while the geometry of the rest of the molecule was optimized using the AM1 semiempirical Hamiltonian as provided in Spartan 5.0.1 (Wavefunction, Inc.). These structures were used as input for ab initio ^{13}C chemical shielding calculations at the 3-21G level, utilizing Hartee–Fock (HF) theory and the Gauge-Including Atomic Orbital (GIAO) method as implemented in Gaussian 98.[28] The results obtained with the 3-21G basis set were then converted to chemical shifts by subtracting the chemical shielding obtained for the methyl carbons of the reference standard tetramethylsilane (TMS) using the same basis set, and scaled to results from a series of reference 6-311G** level calculations. The raw ab initio chemical shift surface data were finally fitted to a trigonometric series expansion of the form (Eq. 1):

$$
\begin{aligned}
{}^{13}\mathrm{C}\delta_{\mathrm{anom}}(\Phi,\Psi) = A_0 + \sum_i [A_i \cdot \sin(i\cdot\Phi) + B_i \cdot (i\cdot\Phi) \\
+ C_i \cdot \sin(i\cdot\Psi) + D_i \cdot \sin(i\cdot\Psi)] + \sum_{i,j,\alpha,\beta} [A_{i,j,\alpha,\beta} \cdot \sin(i\cdot\alpha)\cdot\cos(j\cdot\beta) \\
+ B_{i,j,\alpha,\beta} \cdot \sin(i\cdot\alpha)\cdot\sin(j\cdot\beta) + C_{i,j,\alpha,\beta} \cdot \cos(i\cdot\alpha)\cdot\cos(j\cdot\beta)],
\end{aligned}
\tag{1}
$$

Table 1. Glycosidic bond $\langle\Phi,\Psi\rangle$ dihedral angles, calculated, and experimental CP/MAS anomeric carbon chemical shifts for α-, β-, and γ-CD in the solid state

D-Glc*p* residue	α-CD (CHXAMH)[17,26]				β-CD (BUVSEQOI) [18,26]				γ-CD (CIWMIEIO)[19,26]			
	Φ	Ψ	δ_a Calcd	δ_a Exp[a]	Φ	Ψ	δ_a Calcd	δ_a Exp[a]	Φ	Ψ	δ_a Calcd	δ_a Exp[a]
1	−33.3	−5.7	102.70	102.1	−2.5	9.6	104.10	*102.8*	0.0	17.8	103.07	*101.1*
2	−32.1	53.5	98.86	98.1	−7.6	9.0	104.34	*103.7*	−18.4	−0.9	104.32	*104.2*
3	−20.0	2.8	104.25	*103.2*	−13.9	8.9	104.35	*104.5*	−16.8	8.4	104.29	*103.0*
4	−8.4	18.2	103.53	*102.8*	−18.0	−0.6	104.35	*103.8*	−42.1	4.1	101.93	*101.1*
5	−18.3	11.7	104.06	*102.8*	−11.5	15.7	103.86	*102.7*	−14.9	−2.9	104.47	*105.1*
6	−15.7	5.3	104.43	*103.8*	−2.5	21.8	102.76	*101.9*	−12.1	7.2	104.46	*104.7*
7	—	—	—	—	−14.8	−6.8	104.32	*103.1*	−7.1	16.7	103.64	*101.5*
8	—	—	—	—	—	—	—	—	−10.9	−11.2	104.22	*102.6*
Mean δ_a	—	—	102.97	102.1	—	—	104.01	103.2	—	—	103.80	103.2

[a] Experimental CP/MAS ^{13}C shifts denoted in italics are tentatively assigned (see text for details).

Table 2. Glycosidic bond $\langle\Phi,\Psi\rangle$ dihedral angles, calculated, and experimental CP/MAS anomeric carbon chemical shifts for α-CD inclusion complexes with *p*-hydroxybenzoic acid (PHBA) and *p*-nitrophenol (PNP) in the solid state

D-Glc*p* residue	α-CD-PHBA (ACDHBA)[20,27]				α-CD-PNP (ACDPNP)[20,27]			
	Φ	Ψ	δ_a Calcd	δ_a Exp[a]	Φ	Ψ	δ_a Calcd	δ_aExp[a]
1	−7.9	6.0	104.51		−11.1	12.0	103.07	
2	−6.2	8.0	104.37		−3.5	11.4	104.01	
3	−13.3	15.8	103.85		−16.7	18.5	103.57	104.3 (3)
4	−6.4	12.3	104.05	103.5 (b)	−5.8	14.1	103.86	and
5	10.1	5.1	103.53		5.7	9.2	103.56	102.7 (3)
6	−5.6	0.1	104.67		−6.8	6.1	104.49	
Mean δ_a	—	—	104.16	103.5	—	—	103.94	103.5

[a] For the experimental CP/MAS ^{13}C shifts, (b) denotes a broad signal, and the numeral in parenthesis the number of overlapping resonances.

where A_0, A_i, B_i, C_i, D_i, $A_{i,j,\alpha,\beta}$, $B_{i,j,\alpha,\beta}$, and $C_{i,j,\alpha,\beta}$ are the fit parameters, and α and β can be either Φ or Ψ. The data were fitted using Mathematica 4.1 (Wolfram Research, Inc.), using a 91-term series (i, $j = 1$–3). The resulting function, shown graphically in Figure 1, was incorporated into Perl scripts for use in data analysis, which are available from the authors upon request.

2.3. Molecular dynamic simulations

MD simulations for α-, β-, and γ-CD were carried out and analyzed with the AMBER 6.0 molecular mechanics software suite. The Cornell et al. force field potential energy function was employed for all simulations,[29] using the GLYCAM_2000a parameter set and ensemble-averaged charges for oligosaccharides developed by Woods and coworkers.[30,31] Starting from their X-ray coordinates, the CD models were positioned in the center of a truncated octahedron large enough to provide a 10 Å buffer around them, and solvated with TIP3P waters.[32] A total of 1028, 1088, and 1258 water molecules were used for α-, β-, and γ-CD, respectively, resulting in a density approaching 1.0 g/mL for all three systems. Prior to MD simulations, the solvated models were subjected to 500 energy minimization steps to eliminate bad van der Waals and electrostatic interactions. The relaxed systems were then allowed to equilibrate to 298 K for 500 ps under NTP conditions, using a time constant of 0.2 ps for coupling to the external thermal bath and an integrator step of 1 fs. A cutoff of 8.0 Å was employed for of all nonbonded interactions, the scaling factors for 1–4 nonbonded interactions were set to 1.0, and the particle-mesh Ewald approach was used in the treatment of electrostatics.[33] The SHAKE algorithm was employed to constrain the lengths of all bonds involving hydrogen atoms.[34] After equilibration, the MD simulations were continued under the same conditions for an additional 5-ns period in which data analysis structures were written to disk every 100 fs. All MD simulations were performed on a LINUX Beowulf cluster consisting of 20 Pentium-III 450 MHz compute nodes. Analysis of the data from the MD trajectories was done with the JMP statistical package, version 3.1.5 (SAS Institute, Inc.).

3. Results and discussion

3.1. Chemical shift calculations on CD X-ray structures

In order to investigate the applicability of ^{13}C chemical shift surfaces to the study of oligosaccharide conformation and dynamics, their predictive power and accuracy need to be determined first. As we have stated above and discussed in more detail in our previous report,[16] this is hampered by the fact that most of the NMR data available for carbohydrates is from studies carried out in solution, and as a result the measurements are affected by conformational averaging.[5–7] Although we have shown that these effects can be taken into account through the use of high-quality ab initio potential energy diagrams and Boltzmann statistics,[35] the methodology is only accurate for small molecules with few degrees of conformational freedom. Therefore, ^{13}C NMR data for single oligosaccharide conformers, such as those obtained from CP/MAS studies on crystalline samples, are required for our purposes. Our attention was directed toward the α-(1→4)-linked cyclomaltooligosaccharides, α-, β-, and γ-CD, as these molecules and their inclusion complexes with small ligands have been the subject of several solid-state NMR, X-ray crystallographic, and molecular modeling studies.[17–27,36–39] These systems give us the opportunity to compute anomeric carbon shifts using the D-Glc*p* α-(1→4) chemical shift surface and $\langle \Phi, \Psi \rangle$ dihedral angles obtained from X-ray structures as input, and compare the theoretical estimations directly against experimental ^{13}C NMR CP/MAS data. The results from these comparisons are presented in Table 1. With the exception of two resonances for α-CD, the experimental chemical shifts are not assigned to specific anomeric carbons in the literature,[26] and these resonances are italicized in Table 1 to reflect this fact. The assignments we present are tentative and based solely on the trends found for the computed chemical shifts. Despite this, and that the calculations tend to overestimate the observed CP/MAS data by an average of 0.8 ppm, the computed ^{13}C shifts correlate favorably with experiment, particularly if the mean values obtained for the three molecules are compared. The results obtained for α-CD are more informative. The torsional strain caused by its smaller ring size forces D-Glc*p* residues 1 and 2 in this cyclomaltooligosaccharides to deviate from a regular toroidal spatial arrangement, resulting in glycosidic linkage angles for the 1–2 and 2–3 monomer pairs of $\langle \Phi = -33.3°, \Psi = -5.7° \rangle$ and $\langle \Phi = -32.1°, \Psi = -53.5° \rangle$, respectively.[17] The calculated ^{13}C shifts for the anomeric carbons of these two residues are 102.70 and 98.86 ppm, respectively, which are in very good agreement with the observed CP/MAS ^{13}C shifts of 102.1 and 98.1 ppm.

To further test our methodology, we decided to investigate inclusion complexes of α-CD with 1,4-substituted benzenes.[20] In order create a larger cavity capable of accommodating the aromatic ligands, all the D-Glc*p* α-(1→4) linkages in these systems adopt similar conformations, with Φ and Ψ angles in the (−16.7° ↔ 10.1°) and (0.1° ↔ 18.5°) range, respectively (Table 2). Formation of the inclusion complexes manifests itself in the CP/MAS spectra by the absence of the upfield anomeric ^{13}C resonance at 98.1 ppm observed for free α-CD.[23,27] While ring-current effects from the aromatic ligands can

shield some of the cyclomaltooligosaccharide protons by up to 0.5 ppm, they cause relatively minor perturbations to ^{13}C resonances.[25] Therefore, the variations in anomeric carbon shifts that occur upon complexation are governed mainly by changes in the glycosidic bond conformations. Accordingly, the shifts for the α-CD inclusion complexes with *p*-hydroxybenzoic acid (PHBA) and *p*-nitrophenol (PNP) calculated using the D-Glc*p* α-(1→4) ^{13}C chemical shift surface reproduce the changes observed experimentally, predicting no anomeric carbon resonances below 103.07 ppm (Table 2). Although individual ^{13}C estimations cannot be correlated to experimental CP/MAS data due to the lack of assignments and signal overlap, the change in the average shifts for α-CD before and after complexation with PHBA and PNP can be compared. Analysis of the observed CP/MAS data indicates average downfield shifts of 1.1 ppm upon formation of either PHBA- or PNP-α-CD inclusion complexes.[27] Our calculations predict changes in the average anomeric carbon shift of +1.19 and +0.97 ppm for PHBA and PNP, respectively, which are virtually identical to the average deviations determined from experimental data.

3.2. Effects of conformational averaging

The results presented above confirmed that the ^{13}C chemical shift surface methodology is well suited to predict experimental anomeric carbon shifts in the solid state, and showed that the approach is sufficiently sensitive to accurately reproduce chemical shift changes observed with variations of the glycosidic linkage conformation. The next step in the study was to incorporate conformational averaging effects taking place in solution over the NMR time-scale into the chemical shift estimations, and two different methods were employed with this purpose.

The first approach takes advantage of the fact that all monomers and glycosidic linkages in the cyclomaltooligosaccharides are equivalent due to symmetry.[36] In solution, the $\langle \Phi, \Psi \rangle$ dihedral angles of all the residues in a CD sample the same regions of conformational space in the NMR time-scale, and, as evidenced in the solution-state ^{1}H and ^{13}C spectra of these molecules, all the observables dependent on the glycosidic bond conformation for different D-Glc*p* monomers adopt identical values.[25] Since the D-Glc*p* α-(1→4) linkage conformations found in the CD X-ray structure can be expected to contribute considerably to the solution conformational ensemble, the anomeric carbon chemical shifts for individual residues measured in the solid state should have a significant weight on the shift observed in solution. Indeed, the average ^{13}C shifts derived from CP/MAS data deviate from experimental solution-state anomeric carbon shifts by only 0.48 ± 0.30 ppm (Table 3). Thus, the average of the calculated solid-state ^{13}C shifts should compare well with the solution-state ^{13}C shifts. These averages and the corresponding experimental ^{13}C chemical shifts in solution for α-, β-, and γ-CD are presented in Table 3. Although this approach relies on a number of simplifying assumptions (vide infra), the estimations compare reasonably well with experimental measurements in solution. As was the case for solid-state ^{13}C shift predictions, the calculations based on results from chemical shift surfaces overestimate the observed solution-state ^{13}C shifts by an average of 0.94 ± 0.49 ppm. Despite it slightly exceeds the experimental error, this deviation is relatively small if we consider that there is a spread of up to 0.5 ppm among the reported CD anomeric carbon shifts in aqueous solution found in the literature.[23–25]

The previous method provides a rapid route for the estimation of isotropic ^{13}C shifts in solution, but is limited to CDs for which high-resolution X-ray structures are available. Even in these cases, some of the glycosidic bond conformers observed in certain CD polymorphs may be absent in solution. For example, ^{13}C CP/MAS measurements on frozen aqueous solutions of α-CD indicate that the conformation found in crystalline samples of α-CD hexahydrate is not significantly populated in solution.[37] In addition, the calculations neglect the effects that crystal packing forces and solvation have on the conformation of the cyclomaltooligosaccharides. The polar hydroxyl groups in these

Table 3. Comparison of calculated and experimental anomeric carbon chemical shifts for α-, β-, and γ-CD in solution

Molecule	Calculated shifts (ppm)		Experimental shifts (ppm)	
	X-ray av[a]	MD av[b]	CP/MAS av[c]	Solution
α-CD	102.97	103.08	102.1	102.55
β-CD	104.01	102.75	103.2	103.00
γ-CD	103.80	102.14	103.2	102.40
Average deviation[d]	0.94 ± 0.49	0.34 ± 0.16	0.48 ± 0.30	—

[a] ^{13}C shifts computed by taking the average for equivalent anomeric carbons obtained from calculations on X-ray structures.
[b] ^{13}C shifts back-calculated from the 5-ns MD simulations.
[c] Average of the experimental CP/MAS ^{13}C data for equivalent anomeric carbons.
[d] The average deviations were obtained as the mean of the individual deviations between ^{13}C shifts calculated as described in (a), (b), and (c), and the experimental ^{13}C shifts in solution.

molecules interact strongly with water through intricate and extensive H-bonding networks, and these interactions can alter the conformational space accessible to the CDs in solution considerably.[36] In fact, early studies by Prabhakaran have shown that the average $\langle \Phi, \Psi \rangle$ dihedral angles computed from a 100 ps MD simulation of β-CD in water deviated substantially from the average $\langle \Phi, \Psi \rangle$ angles obtained from the X-ray structure.[38] Therefore, we decided to sample the conformational space accessible to fully solvated CDs extensively through a series of 5 ns MD simulations for α-, β-, and γ-CD with explicit representation of water. The general statistics from these simulations are summarized in Table 4. Our results compare favorably with earlier, albeit shorter, MD simulations of CDs in aqueous solution found in the literature.[38–40] The glycosidic bonds in the three cyclomaltooligosaccharides are highly mobile, and sample a vast region of $\langle \Phi, \Psi \rangle$-space in the 5 ns analyzed. As expected for cyclic molecules composed of identical repeating fragments, the Φ and Ψ dihedral angles from different α-(1→4) linkages in each CD cover similar ranges, which are, on average, ($61.5° \leftrightarrow -65.7°$, $109.4° \leftrightarrow -51.4°$), ($61.5° \leftrightarrow -63.9°$, $76.6° \leftrightarrow -54.4°$), and ($77.2° \leftrightarrow -69.6°$, $74.4° \leftrightarrow -57.0°$) for α-, β-, and γ-CD, respectively. Additionally, the average $\langle \Phi, \Psi \rangle$ angles computed from the MD trajectories of the three molecules deviate from the corresponding X-ray averages (Table 4), corroborating previous findings.[37,38] The simulations also reveal an interesting conformational behavior for the largest CD considered. While the $\langle \Phi, \Psi \rangle$ angles of all the α-(1→4) linkages in α- and β-CD oscillate around their averages values following a nearly normal distribution, the glycosidic bond dihedral angles in γ-CD cluster around two distinct regions of $\langle \Phi, \Psi \rangle$-space. The α-(1→4) linkages in this molecule spend nearly half of the simulation time centered at $\langle \Phi = 14.4°, \Psi = 15.0° \rangle$, and the remaining half of the time around $\langle \Phi = -22.1°, \Psi = -10.5° \rangle$. This bimodal conformational behavior indicates that the general shape of free γ-CD deviates slightly from the regular toroidal geometry adopted by α- and β-CD.

Table 4. Summary of the most relevant parameters obtained from the 5-ns MD simulations in water for α-, β-, and γ-CD

Parameter	Molecule		
	α-CD	β-CD	γ-CD[a]
Φ_{min}	−65.7	−63.9	−69.6
Φ_{max}	61.5	61.5	77.2
Φ_{avg}	−10.2 ± 15.9	−7.3 ± 20.8	14.4 ± 17.7 −22.1 ± 11.1
Φ_{avg} (X-ray)	−21.3 ± 9.7	−10.1 ± 6.1	−15.3 ± 12.3
Ψ_{min}	−51.4	−54.4	−57.0
Ψ_{max}	109.4	76.8	74.4
Ψ_{avg}	7.3 ± 15.3	4.7 ± 15.2	15.0 ± 9.1 −10.5 ± 12.2
Ψ_{avg} (X-ray)	14.3 ± 20.8	8.2 ± 9.6	4.9 ± 9.8

[a] As described in the text, Φ and Ψ angles cluster around two values in the MD simulation of γ-CD, and the averages for each cluster are shown.

The most important step in our study was the back-calculation of the anomeric carbon chemical shift trajectories and averages utilizing the MD conformational ensembles. For this purpose, the $\langle \Phi, \Psi \rangle$ angle pairs from the trajectories of each glycosidic linkage were used as input for the D-Glc*p* α-(1→4) chemical shift surface equation. The mean values from the resulting 5 ns ^{13}C chemical shift trajectories were then computed, and the results for equivalent carbons were further averaged to obtain the back-calculated anomeric carbon chemical shift in solution for each molecule. This process is summarized mathematically in Eq. 2:

$$\langle {}^{13}C\delta_{anom} \rangle = \frac{1}{n_{anom}} \sum_{i=1}^{i=n_{anom}} \left(\frac{1}{n_{steps}} \sum_{j=1}^{j=n_{steps}} {}^{13}C\delta_{anom,i,j} \right), \quad (2)$$

where n_{anom} is the number of anomeric carbons in the CD, n_{steps} is the number of steps in the MD simulation, and $^{13}C\delta_{anom,i,j}$ is the computed shift for the ith anomeric carbon in the jth step of the trajectory. The estimations, presented in Table 3, are in excellent agreement with experiment. In this case, the differences between calculated and observed ^{13}C shifts are well within the experimental error. Although indirectly, our results confirm that CD conformations not found in the X-ray structures are populated in solution. More importantly, they clearly demonstrate that through the use of chemical shift surfaces, chemical shift measurements can be employed in the validation of oligosaccharide MD simulations in a manner analogous to *J*-couplings and NOE correlations.

To further asses the validity of the MD simulations, we decided to investigate if they reflected the symmetric structure of the cyclomaltooligosaccharides. As mentioned above, symmetry causes the chemical shifts of equivalent anomeric carbons to adopt identical time-averaged values in solution. In theory, this behavior should be mimicked by sufficiently long MD simulations that properly reproduce the mobility of the solvated CDs. If this is the case, the running averages of the chemical shift trajectories for equivalent anomeric carbons should converge asymptotically toward a common value, which at the limit should correspond to the anomeric ^{13}C shift in solution. Plots of these running averages for α-, β- and γ-CD are presented in Figure 3. It is clear from the plots that the chemical shifts for equivalent anomeric carbons approach a common value in all cases. The dispersion of the back-calculated chemical shifts for the anomeric carbons in α-CD is 3.05 ppm during the initial stages of the MD simulation, dropping to 0.36 ppm after 5 ns. For β- and γ-CD, the

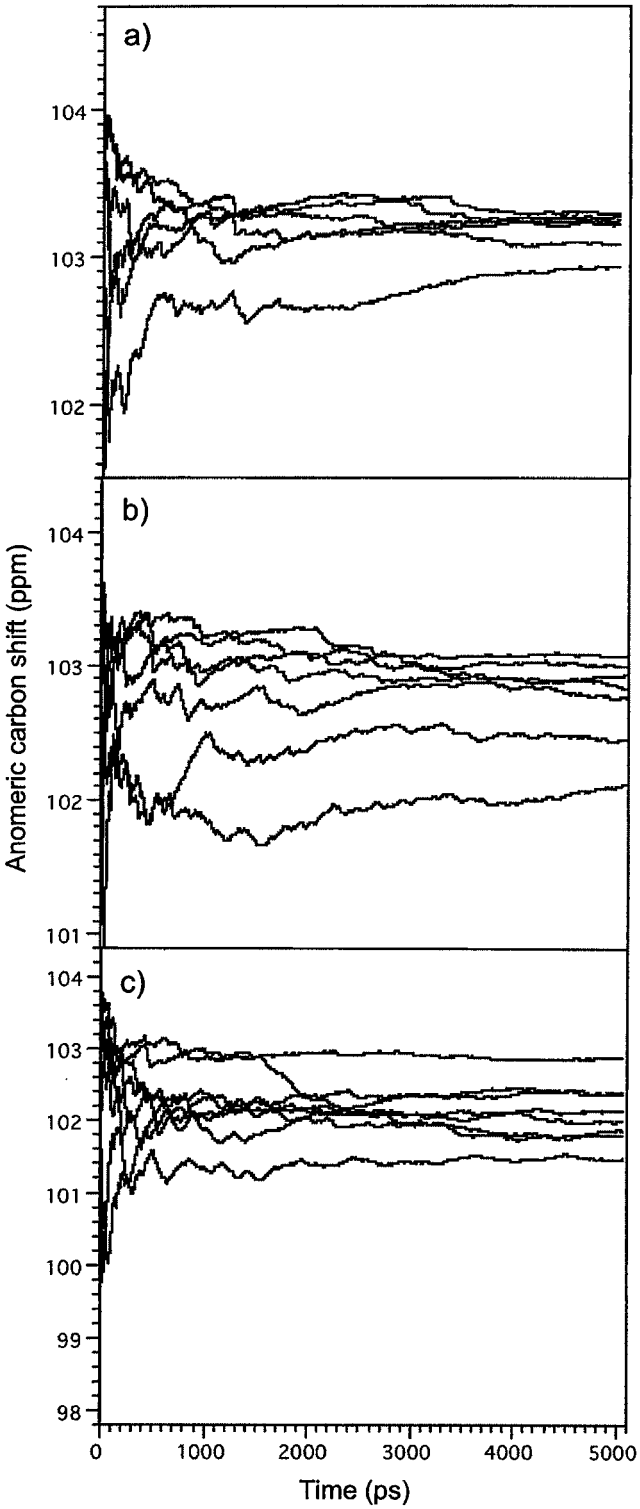

Figure 3. Running averages for equivalent anomeric carbons in α- (a), β- (b), and γ-CD (c) computed from the corresponding 5-ns ^{13}C chemical shift trajectories (see text for details).

initial and final dispersions are 3.47 and 0.95 ppm, and 6.54 and 1.40 ppm, respectively. These results thus indicate that the MD simulations follow the theoretically expected behavior, confirming that they properly reproduce molecular symmetry effects.

3.3. Chemical shift predictions on ε- and ι-CD

As a final test for the ^{13}C chemical shift surfaces, we applied them to the study of ε- and ι-CD, cyclomaltooligosaccharides with 10 and 14 D-Glc*p* monomers,

Table 5. Glycosidic bond $\langle \Phi, \Psi \rangle$ dihedral angles, calculated anomeric ^{13}C shifts in the solid state and in solution, and observed anomeric ^{13}C shifts in solution for ε- and ι-CD

D-Glc*p* residue	ε-CD (NOBBOB)[21,22]			D-Glc*p* residue	ι-CD (NOBBUB)[21,22]		
	Φ	Ψ	δ_a Calcd		Φ	Ψ	δ_a Calcd
1 and 6	−16.1	−22.9	102.60	1 and 8	−21.4	−14.5	103.22
2 and 7	−24.1	−20.4	102.28	2 and 9	−28.8	−31.2	100.28
3 and 8	−15.2	0.7	104.51	3 and 10	−22.8	−14.1	103.14
4 and 9	−39.7	−32.2	99.19	4 and 11	−10.9	16.5	103.77
5 and 10	−34.0	177.5	93.77	5 and 12	−15.4	−3.0	104.41
—	—	—	—	6 and 13	−16.0	−10.4	104.01
—	—	—	—	7 and 14	−38.7	173.4	93.92
Mean δ_a	—	—	100.45		—	—	101.82
Solution δ_a	—	—	99.7		—	—	100.5

respectively. Due to their scarcity, only a handful of studies have been published for these molecules. For example, the only NMR data available for these CDs is in solution,[22,41] and no information regarding their dynamics, neither theoretical nor experimental, has been reported. However, capillary electrophoresis studies have shown that their complex-forming capacities are lower than those of α-, β-, and γ-CD.[42] This has been attributed to the cyclic oligosaccharide rings becoming saddle- or butterfly-shaped to relieve steric and torsional strain, which results in a collapsed internal cavity.[42] Recent X-ray crystallographic studies of ε- and ι-CD have confirmed this structural feature, and show that in both molecules two diametrically opposed D-Glc*p* units are 'band-flipped' nearly 180°, with Φ and Ψ angles of approximately −35° and 175° (Table 5).[21,22]

Since our ^{13}C chemical shift surfaces are particularly sensitive to changes in glycosidic bond conformation, they are especially well suited to detect these band-flipped D-Glc*p* units. Using the $\langle \Phi, \Psi \rangle$ angle pairs from the corresponding X-ray structures, we computed the chemical shift of the anomeric carbons of ε- and ι-CD, and the results are presented in Table 5. Our estimations indicate that the chemical shift of the anomeric carbons in the band-flipped D-Glc*p* residues are almost 10 ppm upfield from the others. Although there are no CP/MAS data to corroborate our predictions, the ^{13}C shifts calculated from the crystal structures can be averaged to obtain good approximations of the chemical shifts in solution following the method described earlier. The resulting estimates are 100.45 and 101.82 ppm for ε- and ι-CD, respectively, both of which are in good agreement with the observed anomeric carbon shifts in solution (Table 5).[22,40] A detailed MD simulation would be required to fully understand the mobility of these larger CDs in an aqueous environment, but this is beyond the scope of the present work. However, our results reveal that the glycosidic linkage geometries found in the X-ray structure have a substantial weight in the solution conformational ensemble average, indicating that the overall shape of these molecules in solution is similar to the one observed in the solid state.

4. Conclusions

In this report we presented the application of chemical shift surfaces derived from GIAO ab initio calculations to the study of a number of CDs and CD inclusion complexes. Our results show that the methodology is capable of predicting the anomeric carbon chemical shifts of cyclomaltooligosaccharides and their complexes reliably, both in the solid state as well as in solution. Results from calculations of solid-state ^{13}C shifts indicate that the approach is particularly sensitive to variations in the glycosidic linkage conformation. In principle, the method could be used to guide the building of 3D molecular models of CDs from high-quality CP/MAS NMR data when X-ray data is not available. The estimates obtained for chemical shifts in solution were also comparable to experimental measurements, particularly when MD simulations were employed to describe conformational averaging on the solvated CDs. Furthermore, these latter results show that ^{13}C shift measurements can be as useful as *J*-couplings and NOE correlations in the validation MD simulations of oligosaccharides. It is also worth noting that our MD simulations detected an interesting bimodal conformational behavior for γ-CD in solution which, to the best of our knowledge, has not been reported previously.

We believe that the work presented here serves to demonstrate the utility of the chemical shift surface method in the study of oligosaccharide conformation. Despite the investigations were limited to the study of anomeric carbon chemical shifts, the approach could be applied to other nuclei whose shielding varies systematically with conformation, particularly the nonanomeric carbon of the glycosidic linkage.[10,11] In fact, preliminary results from our laboratory indicate that the experimental CP/MAS shifts for the C-4′ carbons in α-,

β, and γ-CD can be predicted using the corresponding chemical shift surface and $\langle \Phi, \Psi \rangle$ angles from the X-ray structures. Although the approach requires the computation of new surfaces for glycosidic linkages not yet considered in our earlier studies, the increasing availability of powerful and inexpensive computer hardware will make this task trivial in the near future. While in this respect *J*-couplings and NOE correlations can be considered simpler to use in conformational studies, the measurement of ^{13}C shifts is undoubtedly easier. Furthermore, ^{13}C resonances are not affected by signal overlap as seriously as *J*-couplings, NOE correlations, and other NMR parameters dependent on the ^{1}H chemical shift dispersion are. Thus, experimental ^{13}C chemical shifts, together with the appropriate ^{13}C shift surfaces, could provide crucial information regarding the structure of oligosaccharides in cases where ^{1}H spectral overlap makes the measurement of other NMR parameters impractical.

As stated in the introduction, the CDs were used as test cases for our current methodology. We are presently developing the chemical shift surfaces required to study biologically relevant noncyclic carbohydrates, particularly sialyl Lewis^{x} and other blood group oligosaccharides. The findings from these ongoing investigations will be reported in due course.

Acknowledgements

The authors wish to thank Chet W. Swalina and Prof. Randy J. Zauhar for their invaluable comments and suggestions regarding the manuscript. Financial support provided by the Schering-Plough Research Institute (EPO, GM), and by the Office of the Vice-President of Academic Affairs, USP (GM), is also acknowledged.

References

1. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683–720.
2. Berman, M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235–242.
3. Allen, F. H. *Acta Crystallogr. B* **2002**, *58*, 380–388.
4. Imberty, A.; Pérez, S. *Chem. Rev.* **2000**, *100*, 4567–4588.
5. Bush, C. A.; Matrin-Pastor, M.; Imberty, A. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 269–293.
6. Wormald, M. R.; Petrescu, A. J.; Pao, Y.-.L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371–386.
7. Martin-Pastor, M.; Bush, C. A. *Biochemistry* **2000**, *39*, 4674–4683.
8. Adeyeye, J.; Azurmendi, H. F.; Stroop, C. J. M.; Sozhamannan, S.; Williams, A. L.; Adetumbi, A. M.; Johnson, J. A.; Bush, C. A. *Biochemistry* **2003**, *42*, 3979–3988.
9. Martin-Pastor, M.; Bush, C. A. *Biochemistry* **1999**, *38*, 8045–8055.
10. Saitô, H. *Magn. Reson. Chem.* **1986**, *24*, 835–852.
11. Jarvis, M. C. *Carbohydr. Res.* **1994**, *259*, 311–318.
12. Durran, D. M.; Howlin, B. J.; Webb, G. A.; Gidley, M. J. *Carbohydr. Res.* **1995**, *277*, C1–C5.
13. Wilson, P. J.; Howlin, B. J.; Webb, G. A. *J. Mol. Struct. (Theochem.)* **1996**, *385*, 185–193.
14. Hricovíni, M.; Malkina, O. L.; Bízik, F.; Nagy, T.; Malkin, V. G. *J. Phys. Chem. A* **1997**, *101*, 9756–9762.
15. Zhang, P.; Klymachyov, A. N.; Brown, S.; Ellington, J. G.; Grandinetti, P. J. *Solid State Nucl. Magn. Reson.* **1998**, *12*, 221–225.
16. Swalina, C. W.; Zauhar, R. J.; DeGrazia, M. J.; Moyna, G. *J. Biomolec. NMR* **2001**, *21*, 49–61.
17. Manor, P. C.; Saenger, W. *J. Am. Chem. Soc.* **1974**, *96*, 3630–3639.
18. Steiner, T.; Koellner, G. *J. Am. Chem. Soc.* **1994**, *776*, 5122–5128.
19. Harata, K. *Bull. Chem. Soc. Jpn.* **1987**, *60*, 2763–2767.
20. Harata, K. *Bull. Chem. Soc. Jpn.* **1977**, *50*, 1416–1424.
21. Jacob, J.; Geßler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W. *Angew. Chem., Int. Ed.* **1998**, *110*, 626–629.
22. Jacob, J.; Geßler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W. *Carbohydr. Res.* **1999**, *322*, 228–246.
23. Takeo, K.; Hirare, K.; Kuge, T. *Chem. Lett.* **1973**, 1233–1236.
24. Colson, P.; Jennings, H. J.; Smith, I. C. P. *J. Am. Chem. Soc.* **1974**, *96*, 8081–8087.
25. Schneider, H.-J.; Hacket, F.; Rüdiger, V. *Chem. Rev.* **1998**, *98*, 1755–1785.
26. Heyes, S. J.; Clayden, N. J.; Dobson, C. M. *Carbohydr. Res.* **1992**, *233*, 1–14.
27. Inoue, Y.; Okuda, T.; Chûjô, R. *Carbohydr. Res.* **1985**, *141*, 179–190.
28. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T. A.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B. G.; Chen, W.; Wong, M. W.; Andres, J. L.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98 (Revision A.7)*. Gaussian: Pittsburgh, PA, 1998.
29. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
30. Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832–3846.
31. Basma, M.; Calgan, D.; Varnali, T.; Woods, R. J. *J. Comput. Chem.* **2001**, *22*, 1125–1137.
32. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
33. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Hsing Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577–8593.
34. Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comp. Phys.* **1977**, *23*, 327–341.

35. Swalina, C. W.; O'Brien, E. P.; Moyna, G. *Magn. Reson. Chem.* **2002**, *40*, 195–201.
36. Lipkowitz, K. B. *Chem. Rev.* **1998**, *98*, 1829–1873.
37. Gidley, M. J.; Bociek, S. M. *Carbohydr. Res.* **1998**, *183*, 126–130.
38. Prabhakaran, M. *Biochem. Biophys. Res. Commun.* **1991**, *178*, 192–197.
39. Manunza, B.; Deiana, S.; Pintore, M.; Gessa, C. *J. Mol. Struct. (Theochem.)* **1997**, *419*, 133–137.
40. Lawtrakul, L.; Viernstein, H.; Wolschann, P. *Int. J. Pharm.* **2003**, *256*, 33–41.
41. Endo, T.; Nagase, H.; Ueda, H.; Shigihara, A.; Kobayashi, S.; Nagai, T. *Chem. Pharm. Bull.* **1997**, *45*, 1856–1859.
42. Szejtli, J. *Chem. Rev.* **1998**, *98*, 1743–1753.