Glycoconjugate Journal 15, 511–521 (1998)

EGC2

Conformational properties of a cyclic oligosaccharide: cyclotrikis-(1→6)-[*α*-D-glucopyranosyl-(1→4)-*β*-D-glucopyranosyl]

S. Spieser[1], K. Mazeau[1]*, M.C. Brochier[2], C. Gey[1], J.P. Utille[1] and F.R. Taravel[1]*

[1]*Centre de Recherches sur les Macromolécules Végétales (CERMAV), CNRS, BP 53, 38041 Grenoble cedex 9, France (affiliated with the Joseph Fourier University, Grenoble).*
[2]*Service RMN du LEDSS, Bâtiment Chimie Recherches, BP 53, 38041 Grenoble cedex 9, France*

The title compound is a cyclic oligosaccharide having six glucopyranose residues linked alternatively by *α*-(1→4) and *β*-(1→6) glycosidic linkages. Like cyclodextrin analogues it is expected to exhibit an internal cavity and to form inclusion complexes with other species. In order to investigate its conformational preferences, an extensive conformational search was carried out using a combination of Metropolis Monte-Carlo (MMC) procedure in the glycosidic torsion angle space and molecular mechanics procedures. To this end a specific program (METROCYCLIX) was developed. To reduce the MMC search, conformational maps of parent disaccharides were considered as starting entries. Fully minimized conformations were gathered into families using a clustering technique based on RMS fitting over the glycosidic torsion angle values. A wide range of local energy minima were identified in spite of ring closure conditions that constrained the structure of the oligosaccharide. Low energy conformers were stabilized by intramolecular interactions between distant residues. From the Bolzmann population of the best structures derived from the clustering results, various average properties were calculated and compared with experimental data obtained by high resolution NMR. Interpretation of these experimental values (heteronuclear coupling constants, rotating frame nuclear Overhauser effects, relaxation times) relies on the use of Karplus like equations (coupling constants) and analysis of the full relaxation rate matrix treatment (ROE). The quality of the molecular modelling strategy used is assessed by the agreement obtained between calculated and measured observables.

***Keywords*: Monte Carlo, NMR, ROESY, AMBER**

Introduction

The number of known inclusion complexes formed with cyclic oligosaccharides as potential hosts has rapidly grown in the last few years [1–3]. The wide range of shapes of the different guests able to accommodate the cavity of cyclodextrins is characteristic of a phenomenon, due to the fact that the two compounds associate with each other without any specific interaction.

The crystallographic data on inclusion compounds have further demonstrated the existence of an overall topological equivalence between the cavity of α, *β* and *γ* cyclodextrins, a hydrophobic property that concerns all the members of this family. The driving force for complex formation has been shown to be non-covalent [4–5]. Valuable insights concerning the hydrophobic effect [6–8] have also been considered. The results and the possibilities that this class of compounds offers to the research and industrial fields have initiated the search for new families of cyclic oligosaccharides with similar but also different properties, for example with a non-hydrophobic cavity able to complex polar molecule.

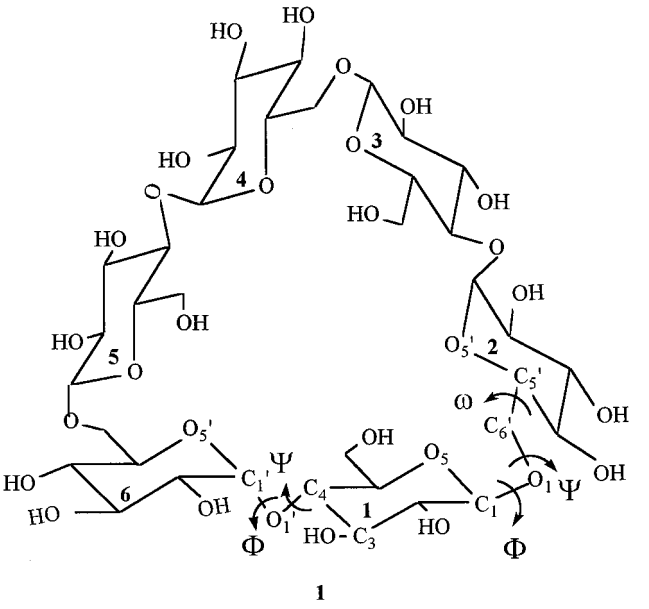
Analysis of the relationship between structure and molecular shape for carbohydrates is made difficult by their inherent flexibility, which allows saccharides to adopt multiple conformations. Knowledge of the conformational preferences of cyclic oligosaccharides through molecular modelling provides a structural basis for the design of analogues for which a controlled cavity size and character can enhance affinity towards potential hosts.

To achieve this aim, a study dealing with chemical synthesis [9], detailed shape characterization as well as conformational properties of cyclotrikis-(1→6)-[α-D-glucopyranosyl-(1→4)-*β*-D-glucopyranosyl], '**1**', has been initiated in

*To whom correspondence should be addressed. Tel.: 33(0)476 03 7603; E-mail: karim@cermav.cnrs.fr and taravel@cermav.cnrs.fr

our laboratory. An important issue in this study concerns the detailed shape and character of the cavity with respect to the macrocycle of '**1**', assessed through combined NMR and molecular modelling [10].

The title compound '**1**' possesses alternatively α-(1 → 4) and β-(1 → 6) glycosidic linkages [9] known to represent the most flexible part of carbohydrates. It was studied using Ramachandran-type maps [11] for both disaccharide models and it was expected that, compared to native cyclodextrins, the presence of β-(1 → 6) linkages should enhance the flexibility of the ring. It should be noticed also when dealing with flexible molecules of any size or complexity, a major difficulty arises which may be described as the local *vs.* global minimum problem [12]. To study the 'whole' conformational space of a molecule and to be able to converge to the global minimum, a maximum surface of the potential hyper-surface must be sampled. This is the classical multiple-minima problem.



1

Figure 1. Schematic representation of cyclotrikis-(1→4)-[α-D-glucopyranosyl-(1→6)-β-D-glucopyranosyl] with atom labelling and torsion angles of the inter-glycosidic linkages.

Materials and methods

Compound '**1**' has been prepared in our laboratory according to a procedure already described [9].

Molecular modelling

The initial structure of '**1**' was built by using coordinates of the global minimum optimized geometry of β-maltose (or α-D-glucopyranosyl-(1 → 4)-β-D-glucopyranose) and α-gentiobiose (or β-D-glucopyranosyl-(1 → 6)-α-D-glucopyranose) as determined by the study of the corresponding adiabatic (Φ, Ψ) potential energy maps. These maps were calculated by applying a systematic perturbation of the torsion angles defining the glycosidic linkage (*vide infra*) over the whole angular range. For each of the conformational microstates, a geometrical optimization was performed by allowing the cartesian coordinates of each atom to vary except those defining the torsional angles that were systematically varied. In this procedure a 10° stepping was applied on both Φ and Ψ torsion angles. For the ω torsion angle belonging to the (1 → 6) glycosidic linkage, only the three staggered orientations were considered. Slightly different results were obtained depending on the initial orientation of the pendent groups due to the multiple minima problem together with the inherent limitations of the minimization procedures. Therefore, 12 and 18 relaxed maps were calculated for maltose and gentiobiose respectively. Their corresponding adiabatic maps were then plotted by taking into account only the lowest energy conformer at each Φ, Ψ point.

Figure 1 shows a schematic representation of '**1**', with labelling of atoms and relevant torsion angles that are responsible for major conformational changes. According to the IUPAC nomenclature [13], the (1 → 4) glycosidic linkage was described by a set of two torsion angles Φ, Ψ that are defined as $\Phi(i) = O5'(i) - C1'(i) - O1'(i) - C4(i+1)$ and $\Psi(i) = C1'(i) - O1'(i) - C4(i+1) - C3(i+1)$. The (1 → 6) glycosidic linkage was described by $\Phi(i) = O5(i) - C1(i) - O1(i) - C6'(i+1)$, $\Psi(i) = C1(i) - O1(i) - C6'(i+1) - C5'(i+1)$ and an additional torsion angle: $\omega(i) = O1(i) - C6'(i+1) - C5'(i+1) - O5'(i+1)$. In our definition, i-odd corresponded to the (1 → 6) linkage and i-even corresponded to the (1 → 4) linkage. All the possible combinations of the three ω torsion angles were examined. They were constrained in the three staggered positions at values of 60°, −60° and 180°, corresponding to gt, gg and tg orientations respectively. In principle $3^3 = 27$ possible combinations should be studied, but because of symmetrical operations only 11 combinations were different. The potential hypersurface was explored by using a deductive, or stochastic [12], strategy. We believed that this procedure, in contrast to the inductive or systematic [12] one, allowed us to sample a larger conformational space in a given time. Therefore, wider energy barriers could be overcome and this offered a reasonable solution to the problem of finding the global minimum.

The starting cyclic structures were generated by using an in-house program, METROCYCLIX, which employs a Metropolis Monte Carlo approach [14]. At this step, the glucopyranose rings were treated as rigid units. The strategy could be commented as follows:

(1) The primary structure of the oligosaccharide was saved.

(2) The oligosaccharide was generated by addition of one residue at a time.

(2.1) The glycosidic torsion angle values for a new residue, n, were randomly assigned by applying a Metropolis

test [15], according to the equilibrium conformational states of the parent disaccharide as calculated from its corresponding adiabatic map.

(2.2) An excluded-volume calculation was then performed (by assuming a spherical shape for the glucopyranose residues with 4.5 Å diameter). If strong close contacts between the added residue, n, and all previous ones were detected, this residue was rejected.

(2.3) New glycosidic torsion angle values were chosen and step 2 was repeated until a successful structure was achieved (n = DP (degree of polymerization) for successful completion of a structure, or if a user-defined limit, m, was reached, in the case of unsolved contacts).

(2.4) If there was no possibility of addition, then previous steps were retraced *ie* residue n − 1 was removed and step 2 was repeated until a structure was obtained or a user-defined number, l, was reached.

(2.5) When the number l was reached, there was no possibility to build the conformer concerned without high energy interactions and this iteration step was terminated with no successful output.

(3) In the case of successful output, ring closure conditions were considered fulfilled when the distance between both ends of the molecule (C4, carbon atom of the first residue, and O1, oxygen atom of the sixth residue) was greater than 0.7 Å and smaller than 3.0 Å.

(4) When the conformer was accepted, an optimization was performed.

(5) If this was new, by comparison with all the previously generated structures, the conformer generated was accepted. This comparison was made on the basis of ordered dihedral angles, which should match within 10°. Because of the symmetry of conformers in some families, the atom-numbering system was permutated to give all possibilities.

(6) Steps 2 to 5 were repeated until a given number of iterations were reached. A typical simulation involved 5×10^5 iterations with m and l numbers set to 20 and 500 respectively (*ca* 1.5 CPU-days). The choice of m and l values was very critical in terms of calculation time. Increasing these values generated more condensed geometries.

This procedure led mainly to distorted conformers. In order to explore how relevant were the highly symmetrical structures, a method based on the search of helices having a pitch close to zero was developed. Stereoregular conformations were assessed by imposing the same Φ(i), Ψ(i), ω(i), Φ(i + 1), Ψ(i + 1) values over the whole molecule. The torsion angle range was restricted to the accessible space of each disaccharide model, and a 10° increment was used throughout the simulation except for the ω torsion angle which was contrained to the three staggered positions. This procedure replaced the preceding step 2.

The AMBER potential function [16] (AMBER 3.0 Rev A package) was used with the set of parameters and the partial charges defined by Homans [17] for molecular mechanics calculations. Energy minimization was achieved by using 600 steps of the steepest-descent algorithm [18]. Then the conjugated gradient algorithm [19] was used with a convergence criterion of $0.05\,\mathrm{kcal\,mol^{-1}\cdot Å^{-2}}$. For all calculations a non-variable dielectric constant $\varepsilon = 1$ was used. Both 1–4 nonbonded and 1–4 electrostatic scaling factors were set to 2. A nonbonded cutoff of 20.0 Å was imposed. The nonbonded-pair list was updated every five iterations. All these parameters have been optimized to reproduce correctly the *ab initio* calculated behaviour of 2-methoxy-tetrahydropyran [20]. However calculations refer to the isolated state of '**1**' and the procedure described here is unlikely to reproduce solvation effects.

The theoretical $^3J_{CH}$ values were obtained by the use of the appropriate Karplus relationship [21]. Full relaxation matrix calculations were performed by the CROSREL program [22], with incorporation of the HOHAHA effect in all theoretical spectra. The noise level was estimated at 0.5% of the initial magnetization. The M_0-scaling method [22] has been used in all cases presented in this paper.

Calculation of theoretical NMR parameters have been carried out through statistical mechanics from all conformers in each family. The relative population of each conformer is given by the Boltzmann distribution as:

$$P_i = \frac{\exp\left(\frac{-\Delta E_i}{RT}\right)}{\sum_i \exp\left(\frac{-\Delta E_i}{RT}\right)}$$

Where E_i is the total energy of the conformer after minimization, R is the gas constant and T the temperature (here a temperature of 500 K was used). Any parameter M_i may be calculated for the i-th conformer, and the overall average parameter $\langle M \rangle$ is given as:

$$\langle M \rangle = \sum_i P_i M_i$$

All calculations were carried out on Indigo 2 Silicon graphics™ and model 10 SUN™ SparcStation.

NMR spectroscopy

NMR spectra were obtained at 298 K on a UNITY + 500 Varian™ spectrometer equipped with a Sun Sparc IPX computer. Samples were dissolved in high-quality D_2O. Reference shifts were taken from standard values. ^1H and ^{13}C assignments were based on homo- and heteronuclear correlation spectroscopy. Information about the conformational preferences of the hydroxymethyl groups can be obtained from NMR spectroscopy using the three-bond coupling constants between H5 and the two H6 protons. The two H6 protons were magnetically nonequivalent, and in the following procedure the *pro-R* proton was denoted H6R and the *pro-S* proton H6S. COSY spectra [23] were recorded using a 4096×512 acquisition matrix and processed using a 4096×4096 transformed matrix with

zero-filling in the F1 dimension. The spectral width was 2250 Hz in both dimensions. A sine-squared function was used prior to Fourier transformation. The ^{1}H-^{13}C shift correlation spectra [24] were recorded using a 1024×128 acquisition matrix and processed using a 4096×4096 transformed matrix with zero-filling in both dimensions. Spectral widths were 2250 Hz in F1 and 12 500 Hz in F2. Sine multiplication was performed prior to Fourier transformation.

2D ^{1}H-^{1}H NOESY [25] and ROESY [26] spectra were obtained in the phase-sensitive mode for different mixing times. For 2D-ROESY experiments, the spin-lock period involved a multi-pulse sequence [27] with a power level corresponding to a 90° pulse length of 128.5 μs. The spectral width was 1750 Hz in both dimensions. In each case, a 2048×512 acquisition matrix followed by a 2048×2048 transformed matrix were used. Peak intensities were evaluated by volume integration with Hartmann-Hahn and offset correction [28] in the case of ROESY spectra. One-bond and three-bond carbon–proton coupling constants were measured using a method proposed by Blechta *et al.* [29] and modified by Nishida *et al.* [30]. The method uses multiple-site selective excitation followed by polarization transfer and acquisition under band-selective proton decoupling conditions and was particularly designed to measure long-range carbon-proton couplings. Under modified conditions, *ie* without the BIRD module [31], one-bond carbon-proton couplings could also be measured with accuracy.

For calculation of the carbon-proton coupling constants, the torsion angles $\Phi_H = H1'–C1'–O1'–C4$ and $\Psi_H = C1'–O1'–C4–H4$ for $(1 \rightarrow 4)$ linkages and $\Phi_H = H1–C1–O1–C6'$ and $\Psi_{HR} = C1–O1–C6'–H6'R$ and $\Psi_{HS} = C1–O1–C6–H6'S$ for $(1 \rightarrow 6)$ linkages with reference to the hydrogen atoms, were used.

Results and discussion

Molecular modelling

As mentioned in the Material and methods section, the (Φ, Ψ) adiabatic energy maps of the parent disaccharides (maltose and gentiobiose) have been obtained by applying a systematic perturbation of the torsion angles defining the glycosidic linkages over the whole angular range. In both cases, the results were in good agreement with MM3 calculations previously reported [32, 33]. For oligosaccharides with $(1 \rightarrow 6)$ glycosidic linkage, the orientation around the C5–C6 bond constitutes an important factor in describing the overall shape of the molecules. In this study, the influence of the three ω torsion angles were particularly investigated. Their values were initially set to the three classical staggered positions *ie* gt, gg and tg corresponding to angles of 60°, $-60°$ and 180° respectively. Among $3^3 = 27$ possible combinations, eleven are non-equivalent (due to symmetrical reasons) and were studied. In order to investigate how the symmetry can affect the shape and the stabilization of '**1**', three symmetrical sub-spaces were added. Therefore, the conformational space of molecule '**1**' was divided into fourteen families noted F1–F14.

Non-symmetrical conformers

Table 1 shows the number of conformationally distinct cyclic structures found for each sub-space. Statisticaly, our program generated 580 different cyclic conformers in a run. On average, there was one successful trial for 850 iterations. The most represented families were F4 (gt–gt–gg), F7 (tg–gg–gt) and F9 (tg–gg–tg) with 850, 1064 and 1358 successful trials respectively. The poorly represented families were F3 (gg–gg–gt), F5 (tg–gg–gg), F6 (tg–gt–gt) and F11 (tg–tg–tg) with 64, 62, 52 and 21 distinct conformers respectively. All possible combinations of the three ω torsion angle values led to macrocycles. However, an attempt to generalize the dependence of cyclization as a function of the ω distribution was unsuccessful. Using this method to generate cyclic structures, the number of successful tries in a family reflected not only the ability to create cyclic geometries but also their flexibility. The low number of structures for some families might be explained by excluded volume effects, which strongly depend on the ω value used and/or on the combinations of torsion angles leading to extended conformations incompatible with the ring closure criterion.

The variation of the backbone torsion angles (Φ, Ψ) for the 850 conformers of the F4 family generated in the MMC sampling is summarized in Figure 2. This figure represents the adiabatic map for each disaccharide superimposed with the distribution of the Φ and Ψ torsion angle values calculated for the F4 family of '**1**' during the MMC sampling. In the figure, the contour lines denote iso-energy values with an upper limit of 8 kcal mol^{-1} above the global

Table 1. Number of different starting geometries (N_{ST}) and different optimized structures (N_{OP}) and relative energy (ΔE) in kcal mol^{-1} of the lowest energy conformer for each family.

	*Family**	N_{ST}	N_{OP}	ΔE
F1	gg–gg–gg	154	95	0.0
F2	gt–gt–gt	71	67	3.48
F3	gg–gg–gt	64	43	1.78
F4	gt–gt–gg	850	573	3.92
F5	tg–gg–gg	27	15	3.68
F6	tg–gt–gt	52	40	8.68
F7	tg–gg–gt	1064	463	6.18
F8	tg–gt–gg	584	389	7.58
F9	tg–gg–tg	1358	683	17.58
F10	tg–gt–tg	78	73	12.18
F11	tg–tg–tg	21	20	35.68
F12	Sgg–gg–gg	218	19	4.88
F13	Sgt–gt–gt	606	65	2.78
F14	Stg–tg–tg	55	23	7.08

*S defines symmetrical structures.

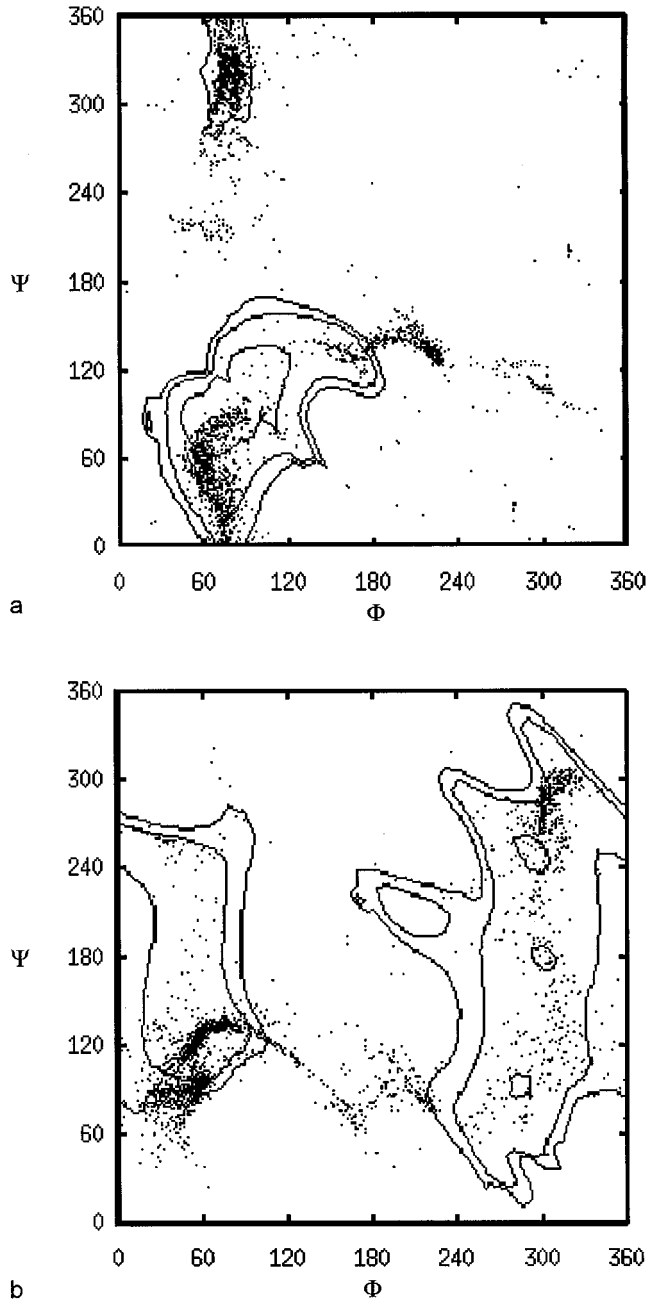


Figure 2. Distribution example of Φ, Ψ torsion angles (for F4) for the two glycosidic linkages (dots) over the Φ, Ψ map of the parent disaccharide for α-(1→4) linkage: a, and β-(1→6) linkage: b.

minimum.The wide range of values exhibited by each glycosidic bond was an indication of the flexibility of the molecule and revealed the sampling of different regions of the conformational (Φ, Ψ) map of the parent disaccharide. Most α-(1 → 4) glycosidic linkage (Figure 2a) showed values of Φ around 60°. This *+synclinal* position corresponds to the lowest minimum of 2-methoxytetrahydropyran with an axially-oriented methoxyl group. Thus, there was a contribution of the exo-anomeric effect which favours these conformers. In addition, conformational states with Φ in *− synclinal* and *antiperiplanar* orientations, less favourable energetically, were also sampled by some glycosidic linkages. The Ψ torsion angles describe the whole angular range. The (Φ, Ψ) values of the β-(1 → 6) linkage (Figure 2b) show an exo-anomeric stabilization of the Φ angle but a restricted Ψ torsional freedom. Most of the Φ values were distributed in the *+synclinal* and *−synclinal* positions that correspond to the minima of the equatorial form of 2-methoxytetrahydropyran. Some values were also found in the third minimum *ie antiperiplanar*. Although most of the points were found within the stable areas of the parent disaccharides, several points were found outside these areas. The latter might be the consequence of the Metropolis test criterion which accepts non-favoured conformations. Furthermore, it should be stressed that Figure 2 corresponds to the potential energy surfaces of both disaccharides on one hand and to the corresponding glycosidic bonds for '**1**' on the other hand. The differences observed are then not surprising because of the oligomeric nature of '**1**' and the cyclization effects that induce constraints on the torsion angles of each glycosidic bond which have to be satisfied.

The energy of the resulting conformers was in the range *ca* −108 to 1200 kcal mol^{-1}. Very high energy structures (above 100 kcal mol^{-1}) represented less than 3% and were discarded. The relative energy values reported in Table 1 identified the lowest energy minimum of each family, relative to the *minimum minimorum* of all MMC sampling. The lowest energy conformer belongs to family F1 (gg–gg–gg). The lowest energy conformation of families F3 (gg–gg–gt), F2 (gt–gt–gt), and F5 (tg–gg–gg) was 1.78, 3.48 and 3.68 kcal mol^{-1} higher respectively. Families with only gg and gt positions (except F5 (tg–gg–gg)) were the most stable (0 to 4 kcal mol^{-1}). Families with one tg position were in the range 4 to 7 kcal mol^{-1}. Families having two ω in tg orientation were in the range 10 to 16 kcal mol^{-1}. The worse energetical family had all three ω in tg orientation. This result was in good agreement with the (Φ, Ψ) map of the parent disaccharide, in which the tg orientation represented only 7% of the total surface of a synthetic map (drawn by using, at each Φ, Ψ state, the lowest energy conformer from the three ω combinations for the glycosidic linkage). Furthermore, the total accessible space was smaller for tg orientation than those for gg and gt. On the other hand, the gg orientation of ω led to the lowest energy families. This could explain the low energy of family F5 (tg–gg–gg) in spite of the presence of one ω in a tg orientation.

The backbone conformation of the lowest energy conformers from the five most populated non-symmetrical families together with the three symmetrical ones, is defined in Table 2 where the values of their backbone torsion angles are represented. A variety of different conformers was observed, revealing not only the flexibility of the molecule, but also the ability of the oligosaccharide to stabilize energetically unfavoured states (noted X in the Table) of individual

Table 2. Torsion angle values for the lowest energy conformers of selected families. X refer to conformers that are outside the low energy regions of the parent disaccharides.

Family	*(1→6)*			†	*(1→4)*		†
	Φ	*Ψ*	*ω*		*Φ*	*Ψ*	
	293.9	102.2	300.4		100.1	88.5	
F1 gg–gg–gg	46.4	115.1	299.9		60.1	63.3	
	54.2	132.3	306.0		107.5	138.5	X
	287.0	109.3	60.7	X	62.7	278.7	
F2 gt–gt–gt	10.6	150.7	55.2	X	75.1	306.4	
	48.3	97.1	51.2		75.8	15.6	X
	291.6	102.7	298.5		95.2	95.0	
F3 gg–gg–gt	52.1	121.5	298.5		70.2	296.4	
	59.0	139.2	55.6		91.0	124.9	X
	314.2	140.4	64.1	X	57.5	66.9	
F4 gt–gt–gg	38.4	114.9	294.9		64.2	88.7	
	299.3	107.6	61.2	X	64.7	281.2	
	319.8	320.4	176.6	X	72.2	309.6	
F5 tg–gg–gg	54.3	124.7	303.7		74.0	82.4	
	49.2	117.1	305.8		101.1	119.2	X
	50.8	129.0	299.8		62.5	65.3	
F12 Sgg–gg–gg	52.1	122.9	300.5		64.9	66.7	
	53.8	126.4	301.7		59.8	69.0	
	268.8	220.7	51.4	X	50.9	54.5	
F13 Sgt–gt–gt	269.1	220.2	51.5	X	51.2	53.8	
	268.8	219.9	51.5	X	51.3	53.9	
	289.4	303.2	178.0		100.1	120.5	X
F14 Stg–tg–tg	289.4	303.3	178.0		100.1	120.4	X
	289.4	303.3	178.0		100.1	120.5	X

glycosidic linkages. Ring closure constraints might force the torsion angles to deviate from the exact geometry of the minima and from the low-energy areas of the parent disaccharide. In fact, conformer stability could not be explained on the sole basis of torsion angle distribution. There must be a delicate balance between first-neighbour interactions that favour all Φ, Ψ couples corresponding to the lowest energy values of model compounds and long-range ones which, in conjunction with the cyclization conditions, might cause differences in the torsion angle values.

Symmetrical conformers

The most represented families were F12 (gt–gt–gt) and F13 (gg–gg–gg) with respectively 606 and 218 cyclic structures. A poorly represented family was F14 (tg–tg–tg) with 55 cyclic structures. The average number of acceptable symmetrical structures, 290, found before minimization was, as expected, less important than for the non-symmetrical ones. This was due to conformational space limitations (see Materials and methods). Here, also, all combinations of the three ω torsion angle values led to macrocycles.

The energy of the resulting conformers was in the range *ca* −106 to 30 kcal mol^{-1}. The relative energy values are reported in Table 1. The lowest energy conformer belonged to family F13 (gt–gt–gt) that was only 2.78 kcal mol^{-1} higher than the global minimum. The lowest energy conformation of families F12 (gg–gg–gg) and F14 (tg–tg–tg) were 4.88 and 7.08 kcal mol^{-1} relative energy. Again the family with all three ω in tg orientation was the less favourable one. It was nevertheless strongly stabilized as compared to non-symmetrical analogues.

NMR spectroscopy

The ^{1}H and ^{13}C signals were assigned by 2D COSY and 2D heteronuclear correlated proton–carbon experiments respectively. ^{1}H and ^{13}C chemical shifts and proton–proton coupling constants are presented in Table 3. Identification of the prochiral protons H6R and H6S have been performed according to results already published [34]. The ^{1}H and ^{13}C spectra show in both cases only two different glucopyranose units corresponding to the two types of glycosidic linkages. There are two possible explanations for these features. In the first, each residue involved in the same glycosidic linkage are magnetically equivalent about a C_3 symmetry axis. In the second, a fast conformational equilibrium occurs between different conformers compared to the NMR time scale. Hence values obtained from NMR experiments will be averages of averages. Firstly, averages because NMR gives only average values. Secondly, averages because one signal is in fact the superposition of three signals. The magnitude of the following proton coupling constants $J_{1,2}$ ($\sim$7 Hz), and $J_{2,3}$, $J_{3,4}$, $J_{4,5}$ ($\sim$9 Hz), are characteristic of β-D-glucopyranose ring having the $^4C_1$ chair conformation. Similarly the magnitude of the following coupling constants

Table 3. Experimental ^{1}H and ^{13}C chemical shifts (ppm) and ^{1}H vicinal coupling constants (Hz) of **1** in D_2O solution at 298 K.

	1	*2*	*3*	*4*	*5*	*6R*	*6S*
β-linked residue							
H	5.41	3.44	3.53	3.09	3.82	3.66	3.83
C	103.81	74.67	76.75	75.06	75.34	62.07	
	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$	$J_{5,6R}$	$J_{5,6S}$	$^2J_{6R,6S}$
$^3J_{H,H}$	7.90	9.41	8.51	9.82	7.62	2.06	−11.50
α-linked residue							
H	4.42	3.21	3.70	3.57	3.54	3.83	3.99
C	98.15	72.28	73.64	71.34	72.82	70.97	
	$J_{1,2}$	$J_{2,3}$	$J_{3,4}$	$J_{4,5}$	$J_{5,6R}$	$J_{5,6S}$	$^2J_{6R,6S}$
$^3J_{H,H}$	3.98	9.96	9.20	9.61	5.39	2.12	−12.10

$J_{1,2}$ (~4 Hz), and $J_{2,3}$, $J_{3,4}$, $J_{4,5}$ (~9 Hz), are characteristic of α-D-glucopyranose ring having the $^4C_1$ chair conformation.

Conformation around the C5–C6 bonds: ω torsion angles

Measurement of the $^3J_{5,6S}$, and $^3J_{5,6R}$ proton coupling constants provides a way to estimate the rotamer distribution about the C5–C6 bond in solution. However, for the interpretation of these coupling constants, it is necessary to choose a set of values for the three staggered torsion angle conformations and for their corresponding coupling constant values. These values can be derived from different sources [35]. For our purpose the set of equations established by Manor *et al.* [36], and modified by Nishida *et al.* [37] was used. In the present study a straightforward solution of the equations yields the respective percentage of (gg : gt : tg) conformers: (58.4 : 40.1 : 1.5) for the β units (those with the free hydroxymethyl groups) and (35.7 : 65.3 : −1.0) for the α units having the hydroxymethyl groups involved in a glycosidic linkage. The negative population of the tg conformer has no physical meaning. In general, a common procedure was used in which the populations of the gg and gt conformers were calculated assuming a tg population of 0. It has been recently reviewed [35] that several factors might complicate the accurate interpretation of the experimental homonuclear couplings. Among them, the choice of the limiting values of the coupling constants that derive from an empirical-based formula, and the exact position of the three staggered rotamers and/or the flexibility around each minimum, are of importance. From our modelling results, we observe that the minima of the ω torsion angles were not trapped into discrete canonical values. On the contrary, one could observe an angular distribution that reached 30°. Interpretation of our experimental data showed qualitatively that the gt rotamer of the (1 → 6) glycosidic bonds was more populated than the gg one (in the ratio $\frac{2}{3}$ of gt and $\frac{1}{3}$ of gg) while the tg one was not represented in the equilibrium mixture. Analysis of the calculated data showed that within 4 kcal mol^{-1}, the represented families were F1, F2, F3, F4, F5 and F13. Therefore, the conformations around the C5–C6 bonds in the glycosidic linkages of '**1**' were mostly gt (50%) and gg (44%), while tg conformers were a minor species (6%). In our study, the F1 family with all three ω torsion angles in the gg conformation were predicted to be the lowest-energy one. It is well known that theoretical investigations of the hydroxymethyl group orientation have not been successful in reproducing experimental results, as empirical force-fields overestimate the influence of intra-unit hydrogen bonding. Also, solvation effects may change the estimation of the relative population of the different conformers.

Heteronuclear long-range coupling constants: Φ and Ψ torsion angles

Comparison between experimental and computed observable NMR parameters is an usual probe for molecular modelling strategy. From the Φ and Ψ torsion angles of all predicted conformations within each family, it is possible to calculate Boltzmann average values for the three bond carbon-proton coupling constants across the glycosidic linkages, by using the established angular dependence of $^3J_{CH}$ for a C–O–C–H pathway (Table 4). Because of the impossibility to distinguish the six different units of '**1**' (*vide supra*), the values reported in the table are also averaged over the three equivalent (1 → 6) and (1 → 4) glycosidic linkages. The corresponding experimental couplings are also reported. In principle the theoretical error of the experimental measurement corresponds to the digital resolution (± 0.2 Hz). However, in practice there are always some uncertainties when measuring long-range couplings and hence the real error is difficult to quantify. An error of ± 0.5 Hz can be considered. As may be seen from Table 4, only two families (F1 and F8) are in fair agreement with the observed values, with however a 0.97 Hz and 0.88 Hz difference for the coupling corresponding to the Φ torsion angle of the three (1 → 6) glycosidic linkages. In Table 4, the average coupling constant values computed for the six lowest energy families, are also given. In this case, the agreement between calculated and experimental data is better and supports the above-mentioned conclusions drawn on the basis of homonuclear couplings *ie* an equilibrium exists between conformers that are, within 4 kcal mol^{-1}, in fast exchange as compared to the NMR time scale.

CROSREL analysis

Due to the fact that the product of the Larmor frequency ω, and the rotational correlation time τ_c is close to $(5/4)^{1/2}$, for many oligosaccharides, no NOE cross peaks could be detected in NOESY spectra. For this reason ROESY spectra were recorded. The spectra were analysed following the full relaxation matrix calculations allowed by the CROSREL method [22]. The first step was the determination of the interproton distance matrices, the estimation of the leakage relaxation rate (R_L) and the rotational correlation time τ_c. Here the distance matrices for each non-exchangeable proton have been obtained after conformational averaging based on the respective Boltzmann population of conformers within a family. In the rest of the study, the assumption of an isotropic reorientation of the molecule was made. CROSREL permitted a grid search in which R_L and τ_c were incremented and compared to experimental intensities [22]. The optimal fit, using intra-residue cross peaks with defined distances such as H1–H2 or H2–H4 for β anomers and H1–H3 or H1–H5 for α anomers, was obtained at $\tau_c = 340$ ps and $R_L = 1.4$ s^{-1} (R_w factor = 0.34) for both F1, the lowest family, and F4, a relevant family when considering the conformation around the C5–C6 bond ($\frac{2}{3}$ of gt and $\frac{1}{3}$ of gg). The fact that the same values were found for both families was in good agreement with the assumption of isotropic reorientation. Some typical, observed and calculated ROEs for F1 and F4 are plotted in Figure 3,

Table 4. Experimental and calculated ^{3}J proton–carbon coupling constants (Hz) at 298 K for all families. Error values were equal to the digital experimental resolution. However, as noted in the text, the actual error may be significantly greater.

	Family	*(1 → 6) linkage*			*(1 → 4) linkage*	
		J_{Φ_H}	$J_{\Psi_{HR}}$	$J_{\Psi_{HS}}$	J_{Φ_H}	J_{Ψ_H}
F1	gg–gg–gg	5.07	2.67	5.37	3.73	3.78
F2	gt–gt–gt	4.10	3.34	4.53	2.75	4.38
F3	gg–gg–gt	5.29	2.26	5.25	3.23	5.51
F4	gt–gt–gg	2.82	2.34	5.03	1.77	4.09
F5	tg–gg–gg	4.69	1.57	5.82	3.31	5.27
F6	tg–gt–gt	3.33	3.61	3.73	1.92	4.16
F7	tg–gg–gt	3.62	3.66	4.72	2.69	3.93
F8	tg–gt–gg	4.98	2.59	4.67	2.98	4.20
F9	tg–gg–tg	4.63	2.29	5.62	1.21	1.75
F10	tg–gt–tg	3.98	3.01	3.71	3.35	4.54
F11	tg–tg–tg	5.19	3.01	2.68	1.52	3.09
F12	Sgg–gg–gg	6.69	1.76	5.54	1.25	2.16
F13	Sgt–gt–gt	4.63	4.83	1.02	1.68	2.21
F14	Stg–tg–tg	2.42	1.26	6.72	3.93	5.49
	Fmean*	4.43	2.83	4.50	2.75	4.21
	Experimental	4.1 ± 0.2	2.5 ± 0.2	5.2 ± 0.2	3.2 ± 0.2	3.9 ± 0.2

* For the six lowest energy families.

taking now into account all cross peaks. The curves were plotted for different dipolar interactions (*eg* H1′–H2′). When an overlap of signals was observed as in the case of H4 and H5, the resulting cross peaks could not be analysed independently. Therefore, they were treated as originating from one virtual proton and are indicated as H1′–H4, H5. The shape and intensity for most of the ROE build-up curves from F1 and F4 compared reasonably well with the experimental curves. Again the worst fit took place between protons of the (1 → 6) glycosidic linkages, in accordance with results derived from long-range heteronuclear coupling constants (*vide supra*). The agreement, however, was sometimes better for family F1 and sometimes for family F4. This could indicate that a conformational equilibrium between the best families might occur, that these MMC calculations failed to reproduce.

Geometrical description of the molecule

The overall shape of the main conformers may be described as a triangle more or less distorted. Figure 4 shows typical examples. The residues having free hydroxymethyl groups that are involved in (1 → 6) linkages create the sides whereas the other residues involved in (1 → 4) linkages correspond to the corners of the triangle. This arrangement indicates that the (1 → 6) linkages, because of its greater flexibility as compared to the (1 → 4) ones, causes kinks in the structure, thus favouring the cyclization phenomenon. This general shape displays a cavity whose sizes are reported in Table 5. The three free hydroxymethyl group positions are of importance because they determine the accessibility of the cavity. Table 5 reports these positions relative to the macro-ring. They are shown as: IN or OUT for an hydroxymethyl directed inward or outward the cavity, UP or DOWN for hydroxymethyl groups that point upside or downside the mean plane of the molecule. Also, the term CAP indicates an hydroxymethyl group that covers the cavity. The largest cavity is created when all hydroxymethyl groups adopt an UP orientation. This is confirmed by conformer 12. Strangely, conformer 13, which adopts the same arrangement has no cavity. This can be explained by the tilt of glucopyranose residues carrying the three hydroxymethyl groups (*exo-skew* position). This leads the secondary hydroxyls on the other side of the molecule to close the cavity. Thus, conformer 13 shows a cup-like shape. The same result appears with conformer 14. In this case, the three primary hydroxyls covers the cavity (CAP orientation). For conformers 1, 3, 4, 5, 7, 8, IN or OUT orientations lead to a substantial reduction of the cavity size by cavity fill-in of one or more hydroxymethyl groups. Also cavity fill-in by one or more glucose units can be seen, until a complete disappearance is reached (conformer 2). This is in fact a consequence of the force-field parametrization of MM programs: compact structures are favoured by maximization of Van der Waals and electrostatic interactions.

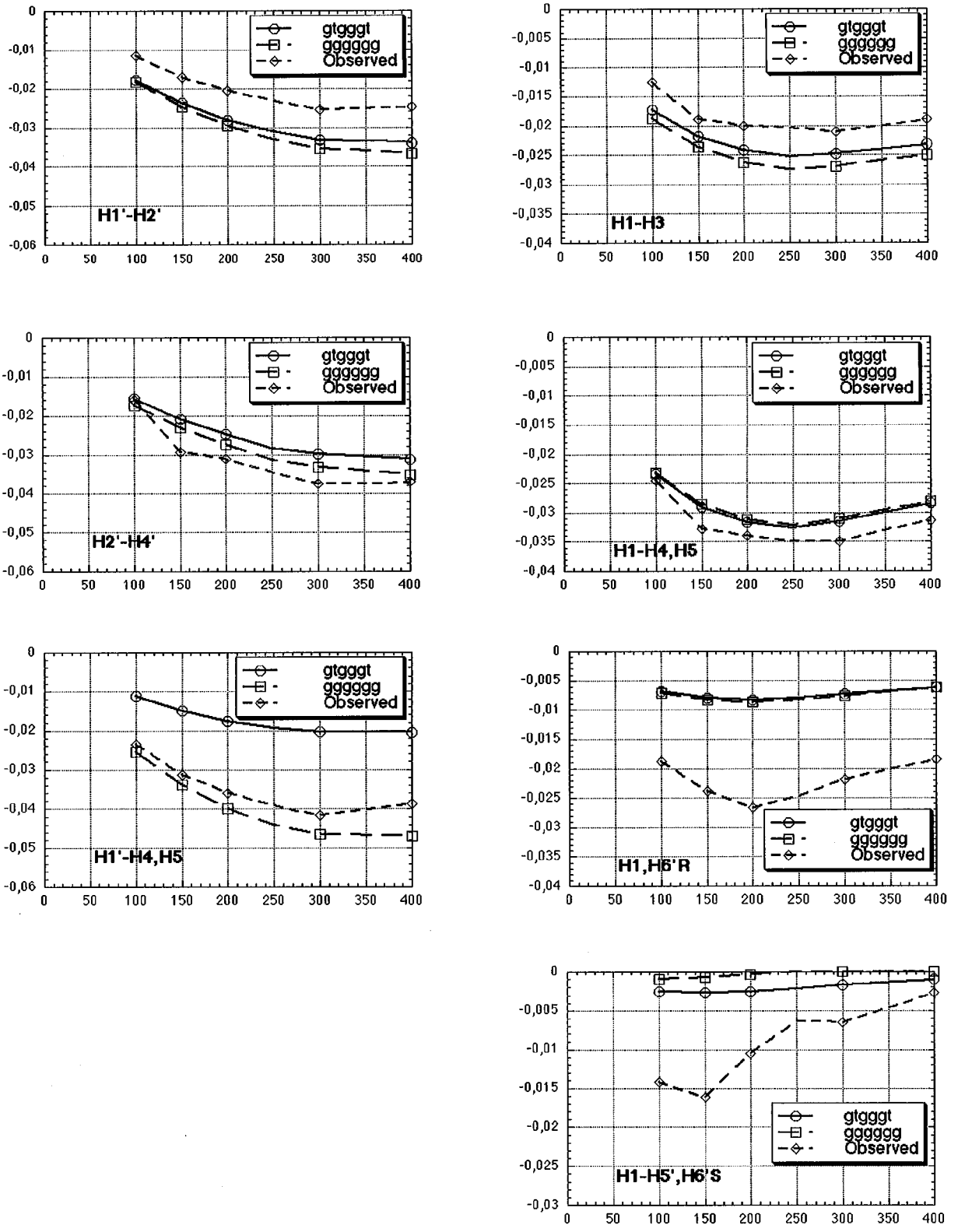


Figure 3. Selected observed and calculated ROE build-up curves relative to M_{0-} in function of the mixing time (ms); for *α*-linked residue (left) and *β*-linked residue (right) (see text for details).

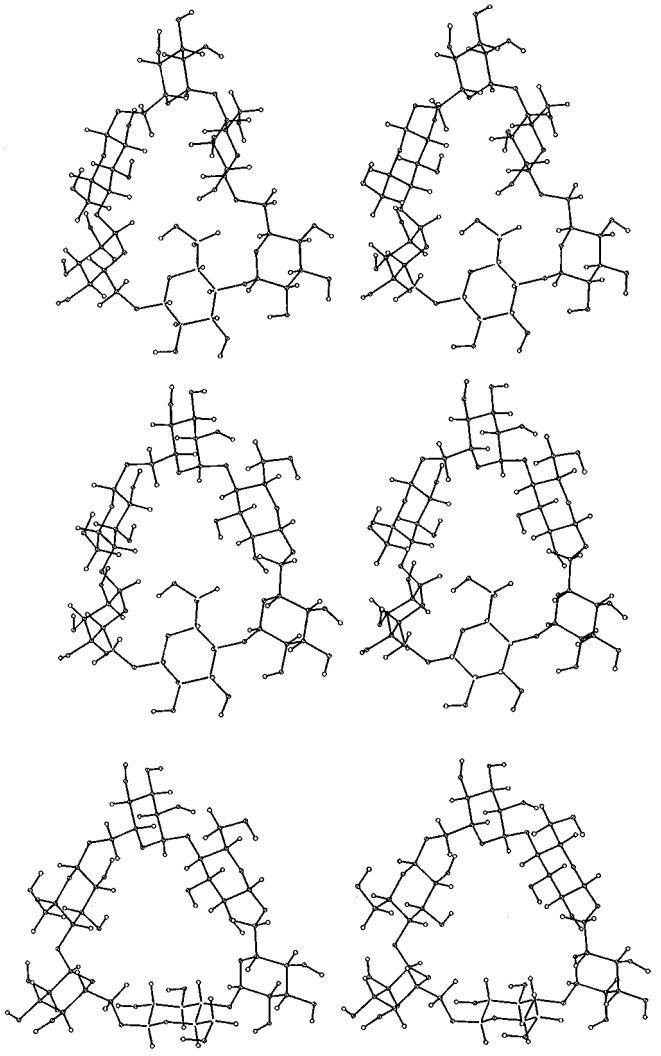


Figure 4. Stereo top-view plots of lowest energy conformers from F3, F1 and F12 from top to bottom respectively.

Concerning arrangement of atoms that are in the α-residue, it can be seen that the intracyclic oxygen atom O5 always points towards the centre of the cavity while the remaining atoms point outside. In most cases, the hydrogen atoms H4 and H2, and oxygen atom O5, of α-residue, are oriented towards the interior of the cavity. On the contrary, the glycosidic O1 atoms are oriented outside. This particular arrangement confers to the cavity of '**1**', a polar character, and corresponds also to the exact opposite found for native cyclodextrins [8]. The polar cavity of the models proposed here, would be more likely to chelate or coordinate ions or form complexes with polar guest molecules. Moreover the greater flexibility of the structures brought by β-(1 → 6) linkages should also facilitate the association of guest and host molecules. This is currently under study in our laboratory.

Table 5. Estimated cavity diameter and depth in Å and relative position of the hydroxymethyl groups (x represents one hydroxymethyl group) of the lowest energy conformers for each family.

Family	*Diameter*	*Depth*	*IN**	*OUT**	*UP**	*DOWN**	*CAP**
F1	4.9	4.6	x		xx		
F2	–			xx	x		
F3	3.4	4.0	x		x	x	
F4	3.4	6.2		x	x		x
F5	4.4	5.8	x		xx		
F6	4.6	6.4			xx		x
F7	4.5	6.4	x		x	x	
F8	4.0	5.8	x		x	x	
F12	6.0	5.5			xxx		
F13	–				xxx		
F14	–						xxx

* See text for definition

Conclusion

In this paper, we have reported NMR data and results of a theoretical investigation of cyclotrikis-(1 → 6)-[α-D-glucopyranosyl-(1 → 4)-β-D-glucopyranosyl], '**1**', using MMC search and molecular mechanics calculations. The cyclization possibility, as a function of the glycosidic torsion angle ω, has been explored. All optimized cluster families give rise to stable cyclic geometries. The best ones present the following distribution (50% gt, 44% gg and 6% tg) which is in qualitative agreement with results extracted from homonuclear coupling constants. These structures are also in good agreement with results from heteronuclear coupling constants and related to the Φ and Ψ torsion angles. They fit correctly the NMR data using CROSREL analysis as well. The calculated low-energy conformers have a non-symmetrical (and sometimes symmetrical) triangular overall shape and a cavity whose diameter is in the range 4–6 Å. These structures are flexible, in particular hydroxymethyl group rotation and residue tilting are allowed which may cause fill-in of the cavity. The size and the polar character of the cavity should permit a favourable accommodation of polar guest molecules or ions. Nevertheless, the role of the solvent on the shape of '**1**' is still unclear and further experiments and molecular dynamic simulations with explicit water molecules are necessary for a better understanding of the actual conformational equilibrium. These topics, as well as, theoretical docking and experimental evidence of complexes of '**1**' with various species are actually under investigation in our laboratory.

References

1 Saenger W (1980) *Angew Chem Int Ed Engl* **19**: 344–62.

2 Kuroda Y, Suzuki Y, He J, Kawabata T, Shibukawa A, Wada H, Fujima H, Go-oh Y, Imai E, Nakagawa T (1995) *J Chem Soc Perkin Trans 2* 1749–59.
3 Botsi A, Yannakopoulou K, Hadjoudis E, Waite J (1996) *Carbohydr Res* **283**: 1–16.
4 Cramer F, Saenger W, Spatz HC (1967) *J Am Chem Soc* **89**: 14–20.
5 Bender ML, Komiyama M (1978) *Cyclodextrin Chemistry*. New York: Springer-Verlag; D'souza V, Bender ML (1987) *Acc Chem Res* **20**: 146–52.
6 Komiyama M, Bender ML (1978) *J Am Chem Soc* **100**: 2259–60; Harrison JC, Eftink MR (1982) *Biopolymers* **21**: 1153–66.
7 Gelb RI, Schwartz LM, Cardelino B, Fuhrman HS, Johnson RF, Laufer DA (1981) *J Am Chem Soc* **103**: 1750–57.
8 Geld RI, Schwartz LM, Laufer DA (1984) *J Chem Soc Perkin Trans 2* 15–21; Lichtenthaler FW, Immel S (1994) *Tetrahedron: Asymmetry* **5**: 2045–60.
9 Driguez H, Utille JP (1994) *Carbohydr Lett* **1**: 125–8.
10 Meyer B (1990) *Top Curr Chem* **154**: 143–208.
11 Rao VSR, Sundararajan PR, Ramakrishnan C, Ramachandran GN (1967) *Conformation in Biopolymers*. Vol 2 Ramachandran, GN London: Academic Press.
12 Goto H, Osawa E (1993) *J Chem Soc Perkin Trans 2* 187–98.
13 IUPAC-IUB Joint Commission on Biochemical Nomenclature (1983) *Eur J Biochem* **131**: 5–7.
14 Binder K, Heermann DW (1988) *Springer Series in Solid-State Sciences*: *Monte Carlo Simulation in Statistical Physics* Vol 80 Berlin: Springer-Verlag.
15 Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E (1953) *J Chem Phys* **21**: 1087–92.
16 Weiner S, Kollman PA, Case DA, Chandra Singh U, Ghio C, Alagona G, Profeta SP, Weiner P (1984) *J Am Chem Soc* **106**: 765–784; Singh UC, Weiner P, Caldwell J, Kollman PA (1986) AMBER 3.0, University of California, San Francisco.
17 Homans SW (1990) *Biochemistry* **29**: 9110–18.
18 Fletcher R, Reeves CM (1964) *Comput J* **7**: 149–54.
19 Powell MJD (1964) *Comput J* **7**: 155–62.
20 Tvaroska I, Carver JP (1994) J *Phys Chem* **98**: 9477–85.
21 Tvaroska I, Hricovini M, Petrakova E (1989) *Carbohydr Res* **189**: 359–62.
22 Leeflang BR, Kroon-Batenburg LMJ (1992) *J Biomol NMR* **2**: 495–518.
23 Wokaun A, Ernst RR (1977) *Chem Phys Lett* **52**: 407–12.
24 Bax A (1983) *J Magn Reson* **53**: 517–20.
25 Bodenhausen G, Kogler H, Ernst RR (1984) *J Magn Reson* **58**: 370–88.
26 Bothner-By AA, Stephens RL, Lee J, Warren CD, Jeanloz RW (1984) *J Am Chem Soc* **106**: 811–13; Bax A, Davis DG (1985) *J Magn Reson* **63**: 207–13.
27 Hwang TL, Shaka AJ (1992) *J Am Chem Soc* **114**: 3157–9.
28 Bax A (1988) *J Magn Reson* **77**: 134–47.
29 Blechta V, del Rio-Portilla F, Freeman R (1994) *Magn Res Chem* **32**: 134–7.
30 Nishida T, Widmalm G, Sandor P (1995) *Magn Res Chem* **33**: 596–9.
31 Garbow JR, Weitekamp DE, Pines A (1982) *Chem Phys Lett* **93**: 504–9.
32 French AD, Dowd MK (1993) *J Mol Struct* (*Theochem*) **286**: 183–201.
33 Dowd MK, Reilly PJ, French AD (1994) *Biopolymers* **34**: 625–38.
34 Gagnaire D, Horton D, Taravel FR (1973) *Carbohydr Res* **27**: 363–372; Gagnaire D, Taravel FR (1975) *FEBS Letters* **60**: 317–21.
35 Bock K, Duus JO (1994) *J Carbohydr Chem* **13**: 513–43.
36 Manor PC, Saenger W, Davies DB, Jankowski K, Rabczenko A (1974) *Biochim Biophys Acta* **340**: 472–83.
37 Nishida Y, Ohrui H, Meguro H (1984) *Tetrahedron Lett* **25**: 1575–78.

Received 9 September 1996, revised 18 March 1997, accepted 30 April 1997