Insights into the Structure of Large-Ring Cyclodextrins through Molecular Dynamics Simulations in Solution

Petko M. Ivanov* and Carlos Jaime

Departament de Química, Facultat de Ciències, Universitat Autònoma de Barcelona, E-08193 Bellaterra (Barcelona), Spain

Received: November 19, 2003; In Final Form: February 27, 2004

Molecular dynamics simulations were carried out using the AMBER parm99 force field and explicit water molecules (TIP3P) to gain insight into the structural deformations and energetics of several large-ring cyclodextrins (with a degree of polymerization 26, 30, 55, 70, 85, and 100) in solution. The structures displayed by CD26 during the MD simulation (10.0 ns) did not correspond to the conformation in the crystalline state. The two "flips" present in the macroring of CD26 in the crystal state disappeared after 1.5 and 3.0 ns simulations, respectively. The larger CDs bear a considerable degree of flexibility. They display different modes of folding and cavity-like regions of different sizes and shapes: circular and elongated loops of variable size, orientated in different fashion; portions of a double helical strand with the two single helices parallel to each other (CD30, CD70); a helix of three turns and a serpentiform portion containing six loops (CD55); a cone-shaped spiral region (CD70); a rounded dendritic fold with several arbitrarily oriented small loops on the surface of the clustering (CD85); three spiral portions and a tendency for bending into two (CD100). These results support the hypothesis for the existence of more than one cavity in large-ring cyclodextrins.

Introduction

The "native cyclodextrins" (CD6, CD7, and CD8) have been extensively studied both experimentally and computationally.[1] Their larger analogues, the large-ring cyxlodextrins (LR-CDs), have recently attracted significant research interest.[2,3] The structural features of the native and some of the LR-CDs have been surveyed, also quoting computational studies.[2,4−7] One review was devoted specifically to the application of molecular modeling techniques to the study of the static and the dynamical features of CDs, as well as their participation in the formation of host−guest complexes.[8] Macrocycles containing 14, 18, 24, and 48 glucose units have been studied by molecular dynamics simulations either in a vacuum or in water solution, but with rather short simulation times.[9] DFT/ab initio-derived empirical force field, AMB99C, was used in studies of the molecular properties of the cyclodextrins containing 10 and 21 glucoses.[10] The structures of the cyclodextrins with 14, 21, 26, and 28 glucose units were also simulated in the gas phase (6.0 ns) and in water solution (4.0 ns).[11]

The crystal structures of cyclodextrins with 9,[12] 10,[7,13,14] 14,[7,13] and 26[15] glucoses have been reported. Thus geometrical properties and characteristics of some LR-CD are available for the solid state. There is, however, limited information about either the structure of the macrocycles in solution, or their complex-forming properties.[2,3] The ^{13}C NMR spectra in D_2O at 50 °C of the series from CD6 to CD31 are indicative of all glucose residues being identical on the NMR time scale.[2,16]

In the present study we examine several LR-CDs with a degree of polymerization of 30, 55, 70, 85, and 100. CD26 was also considered as a reference structure, the largest one for which X-ray structural data are available.[15] We have addressed two important questions. First, does CD26 preserve its crystal geometry in solution, and if not, what would its most probable macroring conformation be in water solution? We arrived at an answer to this question from long (10.0 ns) molecular dynamics simulations in water. A preliminary SAXS study on CD26 by Kitamura and co-workers has been cited to show that the solution conformation of CD26 could not be simulated by its solid-state structure.[2] Second, what are the structural deformations and the modes of folding of the LR-CDs in water solution? Do they have cavities and/or channels that may eventually host small molecules, as well as characteristic folds and motifs of local structural deformations such as the ones we know already from crystal structural determinations (e.g., "flips")?[3] We conclude that the solution geometry of CD26 deviates significantly from the molecular geometry in the crystal lattice. The present results are in support of the expectations for the formation of multiple cavities in large cyclodextrins.[2]

Computational Details

All computations were carried out using the AMBER program (version 7; the AMBER modules LEaP, SANDER, and CARNAL were used for preparation of the input data, minimization, and the MD simulation steps, and analysis of the MD trajectories, respectively)[17] using the most recent parametrization of the AMBER-related additive force fields, parm99.[18] The molecular dynamics simulations were run for water solution (a box with TIP3P[19] water molecules) using the particle mesh Ewald (PME) method[20] for the treatment of the long-range electrostatics. A 9.0 Å distance cutoff was used for direct space nonbonded calculations and a 0.00001 Ewald convergence tolerance for the inclusion of long-range electrostatic contributions. The "solvateBox" command of LEaP was used to create a rectangular parallelopiped solvent box around the CD with buffer distances of 10.0 Å between the walls of the box and the closest atoms of the solute. The dimensions of the periodic TIP3P water boxes and the number of water molecules were as follows: CD26 (35.2 Å, 46.6 Å, 41.5 Å; 6309); CD30 (72.1 Å, 53.8 Å,

* Corresponding author. E-mail: petko.ivanov@uab.es. Permanent address: Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, Sofia, Bulgaria.

10.1021/jp037527t CCC: $27.50
Published on Web 04/23/2004

39.6 Å; 14730); CD35 (53.5 Å, 45.4 Å, 56.6 Å; 13014); CD55 (77.1 Å, 46.2 Å, 66.5 Å; 22488); CD70 (69.4 Å, 80.3 Å, 46.8 Å; 24669); CD85 (46.4 Å, 71.8 Å, 84.0 Å; 26184); CD100 (50.4 Å, 73.0 Å, 92.4 Å; 34199). The SHAKE option (tolerance 0.00005 Å) was used for constraining bonds involving hydrogen atoms. The MACROMODEL software package (version 5)[21] was also used for building the initial geometries (Figure 1S) and for preliminary optimizations with the AMBER* force field. Crystallographic coordinates were used for CD26 (CCDC11-5146).[15a] The generation of the initial structures and the execution of the simulations were conducted in the order CD30 → CD55 → CD100 → CD85 → CD70, using each preceding optimized structure as input for the derivation of the next one by computer graphics. Besides, the initial geometry of CD30 was obtained from AMBER* optimized coordinates of CD40, whereas the initial CD55 structure was derived from modeling and simulation computational experiments on CD35. In this way we tried to eliminate, as far as possible, the influence of the quality of the starting geometry on the computed results and use well relaxed starting geometries in each case.

The preparation of solvated macromolecules for the simulations comprised several stages: (i) 5000 steps steepest descent and 200 steps conjugate gradient minimization of the LR-CD in the gas phase, followed by 5000 steps steepest descent and 400 steps conjugate gradient minimization of the whole (LR-CD plus water molecules) fully unrestrained system; (ii) 50 000 steps steepest descent minimization with holding the solute fixed with positional restraints; (iii) 25.0 ps unrestrained MD run at 100 K on the water alone while the LR-CD was constrained (this is the stage of the equilibration process where the bulk of the water relaxation takes place); (iv) gradual release of the restraints on the LR-CD in a series of minimizations and MD steps (following the suggestions given in the decamer poly-(A)−poly(T) duplex DNA model structure example in the tutorial of AMBER7[17]): 1000 steps minimization and 3.0 ps MD with 25.0 kcal mol^{-1} Å position restraints, followed by five rounds of 600 steps minimization, reducing the positional restraints by 5.0 kcal mol^{-1} Å each run; (v) the equilibration process completed with 20.0 ps MD simulation after heating the system from 100 to 298 K in 2.0 ps. The productive runs were performed with the recommended maximum time step of 2.0 fs when SHAKE is used, at 298 K and a constant pressure of 1.0 bar with isotropic position scaling using the Berendsen algorithm for temperature coupling. The simulation time was 5.0 ns. An additional 5.0 ns simulation was executed for CD26. Samplings were taken every 2.0 ps. We monitored the variations with the simulation time of the temperature of the systems, the pressure on the walls of the boxes, the volumes of the boxes, the density of the solvent, and the relevant energy characteristics (total, kinetic, and potential energies), and they were all indicative of simulations of equilibrated systems (Figure 2S in Supporting Information displays total energies, and Figure 3S shows the rms deviations of atomic coordinates relative to the last structure registered for each CD). The plots of the total energy vs time after the equilibration show total energies oscillating around constant values for all CDs. Thus, all monitored conformational transitions occur in systems at thermal equilibrium.

Except for CD30 (an SGI Octane workstation was used), all other computations were performed using the Compaq AlphaServer HPC320 of the Load Share Facility cluster of CESCA. The jobs for CD70 and CD100 required about 1120 and 1440 h CPU, respectively. We attempted as far as possible to achieve the highest standards that are practically feasible under the computational resources currently available. Longer simulation times (with smaller time steps) and improved modeling are required for a more complete and precise analysis of the conformational flexibility and the complex-forming properties of these interesting systems (e.g., larger water boxes, with cubic or truncated octahedral forms, the consideration of several starting geometries for each CD, and further improvement of the molecular mechanics force fields for carbohydrate solution simulations).[10,22,23] Still at a preliminary stage of development are approaches that could provide the grounds for the development of molecular mechanics parametrizations that adequately balance the difference in strength between intramolecular and intermolecular hydrogen bonds for one and the same pair of functional groups.[23b] Such a force field could be used to study the effects of water molecules (inside and outside the cavities, as well as the bridging intermolecular water hydrogen bonds, formed with the hydroxyls at C2(n) and C3($n-1$) across the glycosidic linkages) on the degree of rotational flexibility of the glycosidic dihedral angles and the overall shape of the LR-CDs.

Several well-known parameters were used to analyze the structures of the LR-CDs (see Supporting Information). Their computed values were compared with experimental determinations from X-ray analyses of CD26 and other smaller CDs. In addition, the variations with simulation time were monitored for the moments of inertia. Time-averaged values of the O4 to the center of mass distances were evaluated for each residue of all CDs. The energy for each LR-CD was computed also by the MM/GBSA (generalized Born[24]/surface area (LCPO)[25]) methodology implemented in AMBER. Values normalized per glucose unit were obtained to access potential variations of these quantities between different LR-CDs, and to determine which energy terms have the most pronounced contributions to these variations. About 250 structures from the whole MD trajectories were utilized to estimate the MM/GBSA energies in each case.

Results and Discussion

Figure 1 displays a schematic representation of a fragment of a cyclodextrin macroring with the numbering of the atoms and examples for cis and trans orientations of neighboring glucose units. The names of the structures used for the figures and in the text were composed by (i) the cyclodextrin abbreviation and (ii) a number indicating the simulation time (in nanoseconds) and referring approximately to the moment of the simulation at which this structure has been picked up from the MD simulation trajectory file (a snapshot). In this way, CD100-2.0 represents a snapshot of the structure for CD100 after 2.0 ns MD simulation.

Figures 2−7 display stereogeometries from different stages of the simulations. The experimental X-ray geometry of CD26 (CD26(X-ray)) is also given and compared with snapshots of geometries from the first 5.0 ns simulation (CD26-5.0) and the subsequent restarted simulation of additional 5.0 ns (CD26-10.0). Some geometrical and energy details, as well as estimates of other properties of the simulated systems are included in the Supporting Information.

Structural Variations of the Large Cyclodextrins in Solution. We examined, in stereo, 10−15 snapshots for each CD in order to trace the tendencies in the deformations of the macrorings, focusing on the formation and disappearance of cavities and/or channels that may eventually host small molecules, as well as characteristic folds and motifs of local structural deformations.

The MD simulation of CD26 was performed starting with the experimental X-ray geometry (CD26(X-ray) in Figure 2).

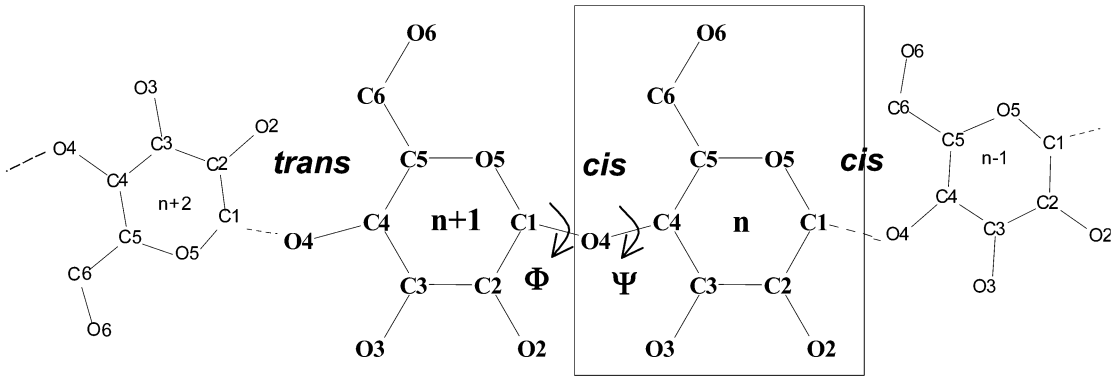


Figure 1. Schematic representation of a cyclodextrin showing the atomic numbering. Each individual glucose unit is designated by a number "*n*". Atoms are identified by a letter and a number that indicates its position in the glucose.

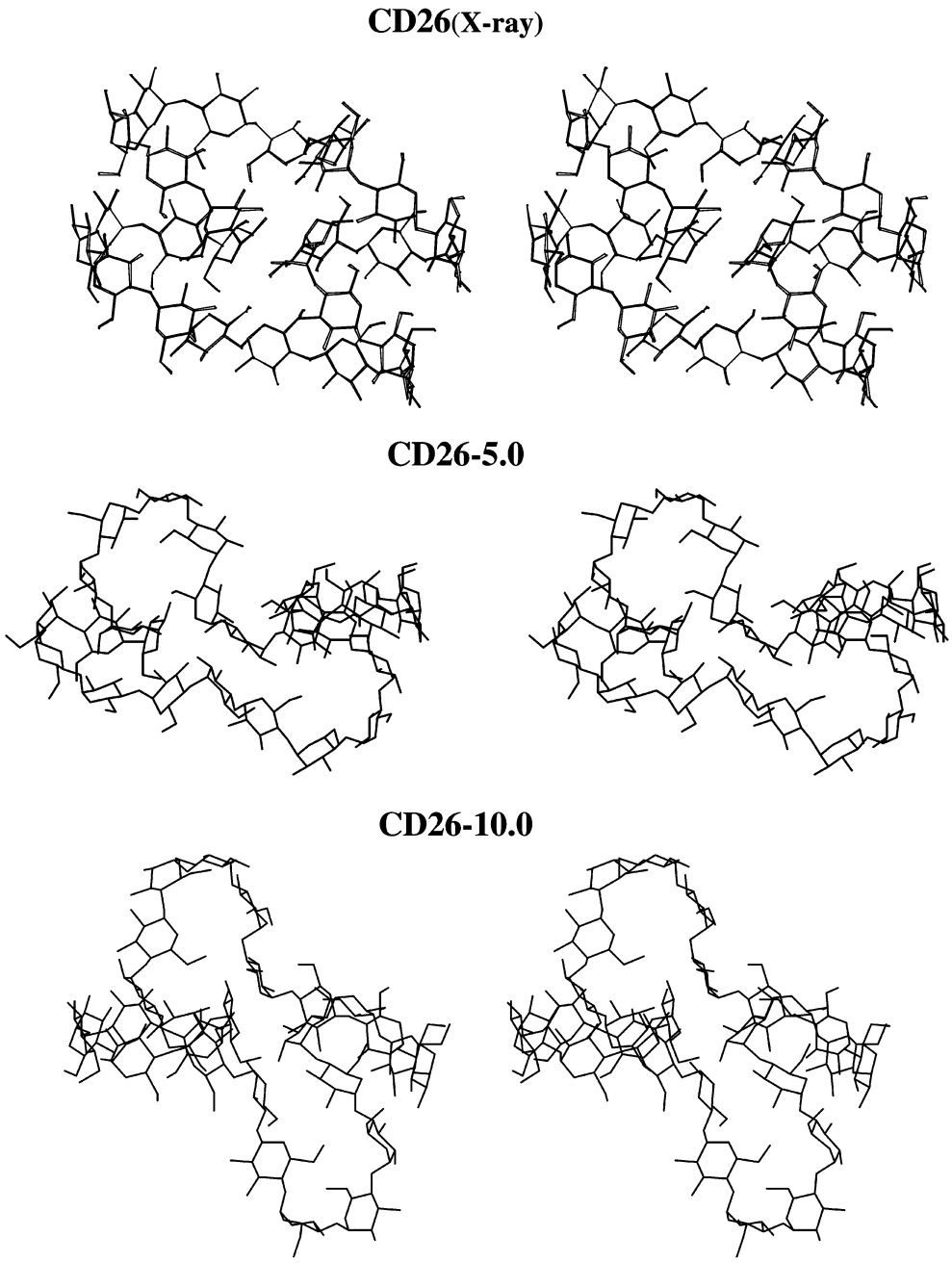
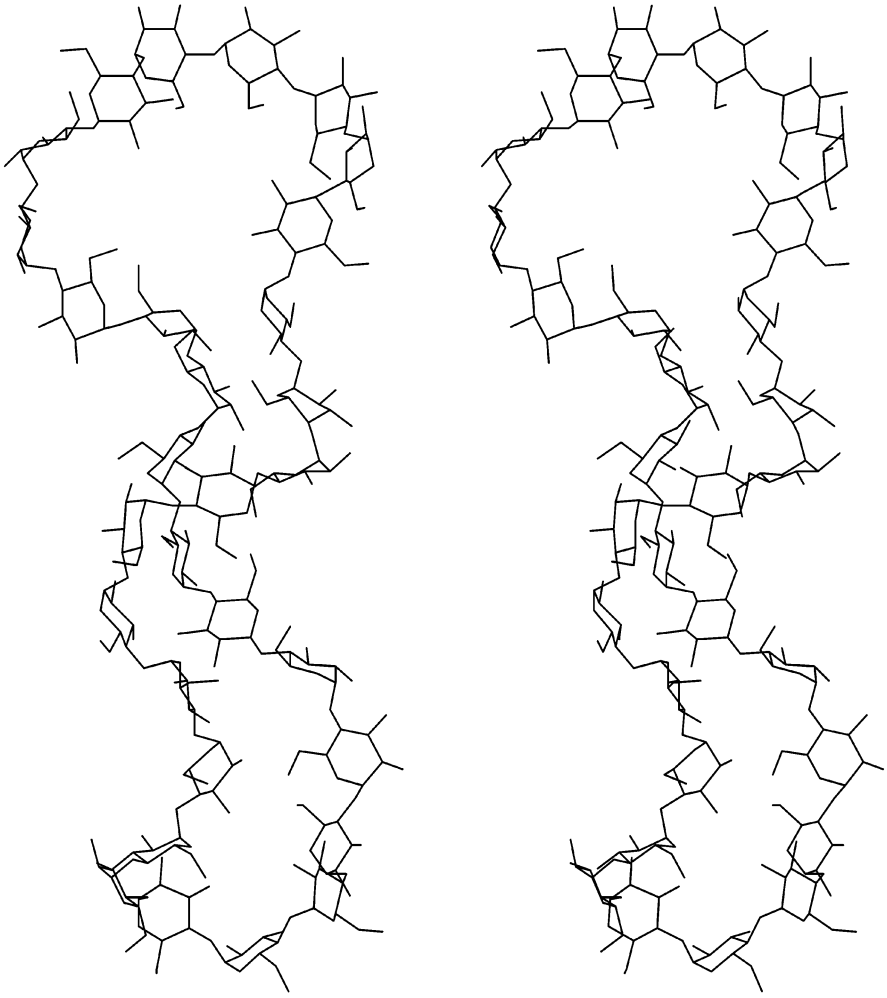


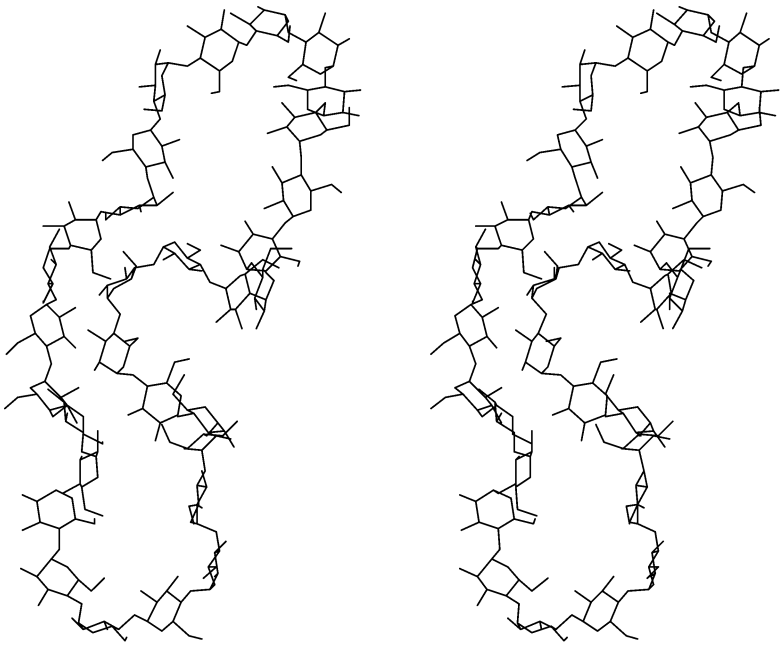
Figure 2. Stereodisplays of the initial geometry of CD26 and snapshots of geometries from different stages of the simulation.

An earlier 4.0 ns simulation with the AMBER parm94 force field[11] rendered only one average structure that closely resembles the starting geometry, although less compact than CD26(X-ray). The present study revealed more significant structural variations

CD30-1.5



CD30-3.0



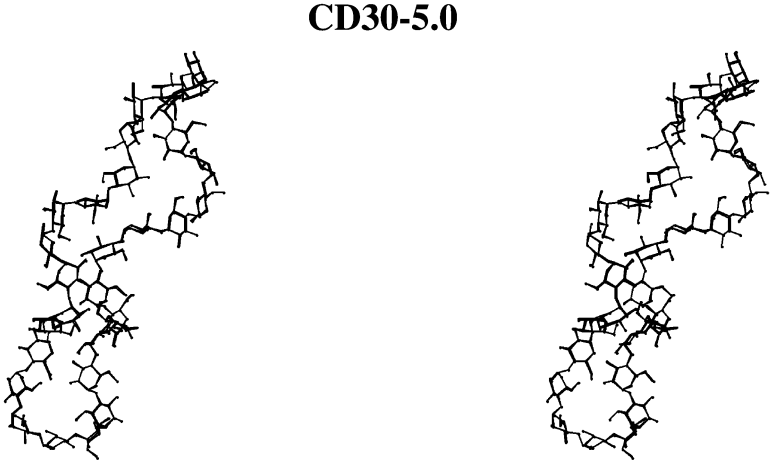


Figure 3. Stereodisplays of snapshots of geometries of CD30 from different stages of the simulation.

during even shorter simulation times. The end turns of the two antiparallel left-handed single helices (each one having ca. two turns; the structure of each helix resembles V-amylose with six glycosyl units per repeating turn[2]) were the first to broaden symmetrically, one from the "upper" and the other from the "lower" sides of the stretches that connect the two short helices (Figure 4S, CD26-0.75). One of the two expanding loops tilted its average plane gradually, at about 1.5 ns, with respect to the plane of the other neighbor turn of the single helix until it became almost perpendicular to it (CD26-2.0). Accordingly, one of the helices transformed into two mutually perpendicular loops. The tilt at 1.5 ns was accompanied by the disappearance of one of the "flips". Meanwhile, the other end loop that was also enlarging at the earlier stages of the simulation shrank in response to the tilt in the first helix. The second helix acquired the same geometry as the first one at about 3.0 ns, and accordingly, the second "flip" also disappeared. The overall appearance of CD26 at 3.0 ns was with the shape of two crossing each other figures of eight (CD26-3.0). One of them had circular loops of ca. 6 glucoses. The other two loops were bigger and more elongated. The next mode of the dynamic reorganization of the macromolecule was a close approach and a tendency for self-inclusion of portions of the larger loops into the cavities of the neighbor smaller turns (CD26-5.0 in Figure 2 and Figure 4S).

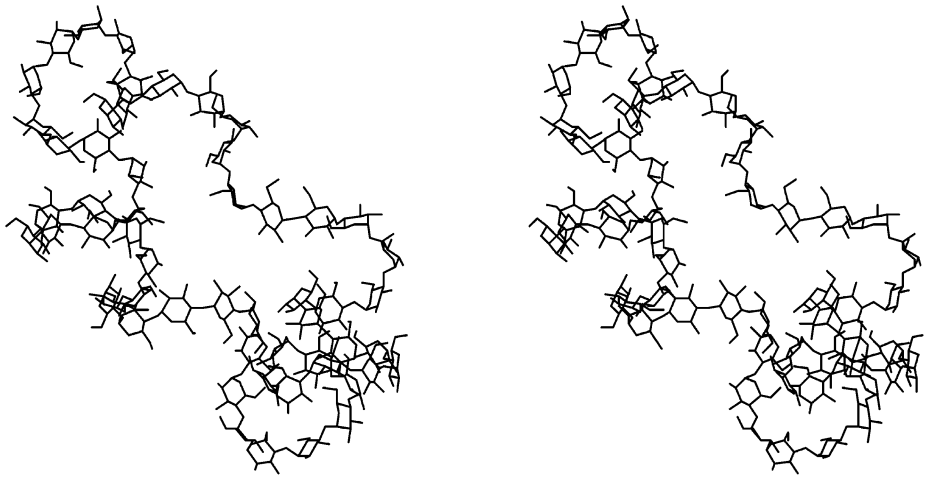
All monitored structural variations of CD26 took place in a manner that maintained a periodicity of about 6−7 glucose units for the average distances of the O4 atoms to the center of mass of the solute (Figure 8 and Figure 9). An examination of the average coordinates of CD26 derived from the trajectory file for the last 5.0 ns simulation revealed that half of the macroring retained practically the same conformation (the right parts of CD26-5.0 and CD26-10.0 in Figure 2), whereas the geometries of the other 13 glucose units underwent more significant deformations about this "reference core" (see also Figure 9).

The present simulations are too short and insufficient to simulate the NMR time scale (acquisition time about 1 s). Nevertheless, we have sufficient grounds to conclude that the LR-CDs do not adopt a single geometry but rather change the conformation of the macroring and thus all glucoses should average in time. Because all glucose units are identical, each one of them can occupy any position in the macroring; thus they will all finally appear equivalent in the NMR time scale. The observation of a single signal for each carbon atom type (Cn, n = 1−6)[16] does not contradict the results from our simulation studies. There is a fast exchange among all glucoses that averages the chemical shifts for each carbon atom.

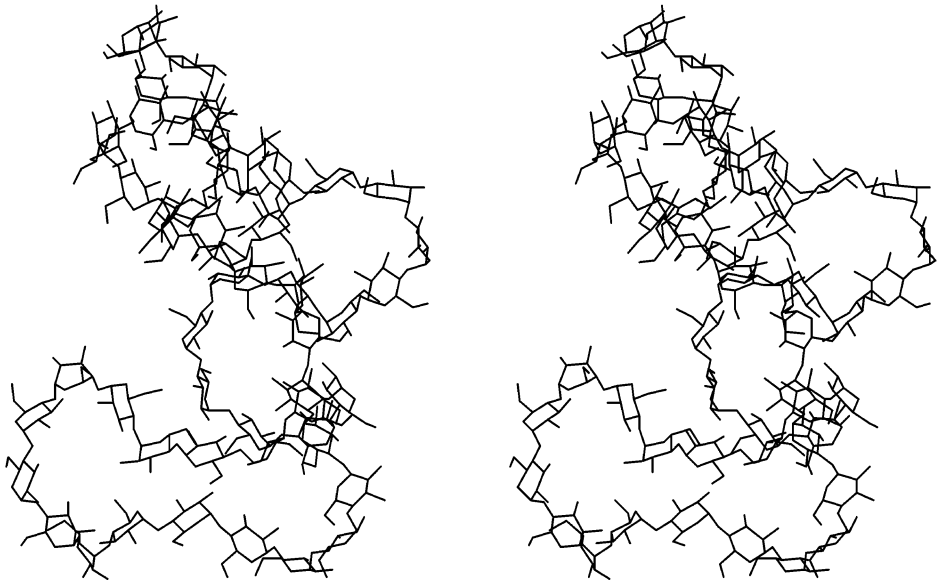
Having answered the first question, the geometry of CD26 in water solution deviates significantly from its solid-state structure, we proceeded to investigate the structural deformations of LR-CDs in water solution, and the tendencies to form cavities and/or channels and other characteristic folds and motifs of local structural deformations. Due to the dimensionality of the problem, we cannot examine in full the dynamics of these systems. Executing several longer runs for each CD, starting from different initial structures may have been the preferred computational protocol. In view of the enormous computational resources needed for such a study, however, we attempted an alternative approach by selecting several LR-CDs and executing MD simulations for each one of them for a more practically feasible length of time. Thus, in effect we acquired structural information about the LR-CDs in water solution. Moreover, it is unlikely that a useful result from the study, i.e., the presence of multiple cavities, would be invalidated if several runs with different initial structures are carried out, or results from several molecular mechanics force fields are compared. We expect that other groups will also report results from simulation studies on the LR-CDs resulting in a more complete picture of the dynamics of the LR-CDs. Such studies necessitate enormous resources and, at present, only by contributions from different groups are we likely to attain the necessary information.

The same methodology of the analyses was followed as for CD26. We will not go into details here but instead mark only the most important observations. The starting geometry for CD30 is completely different from the initial structure of CD26: it resembles an elongated open macroring with longer left-oriented helical sides of 11−12 glucoses. The central portions of the longer sides approached closer at the initial stages of the simulation and wound around each other in a cross fashion (an interwound fragment). A deformed loop of 10 glucoses with a "flip" was formed at the one side of a double helical strand at 1.5 ns (resembling the X-ray structure of CD10). Two smaller loops of 6−7 glucoses flanked the extended structure after 2.0 ns simulation. A larger deformed loop of ca. 11−12 glucoses was formed again at one of the ends at 3.0 ns (Figure 3 and Figure 4S, CD30-3.0). The macromolecule was "breathing" about these patterns gradually approaching the geometry of a symmetrical double helical strand with the two single helices parallel to each other (Figure 3, CD30-5.0). The periodicity of 2 (15 glucose units per a period) is manifested in the curve presenting the average O4 to the center of mass distances (Figure

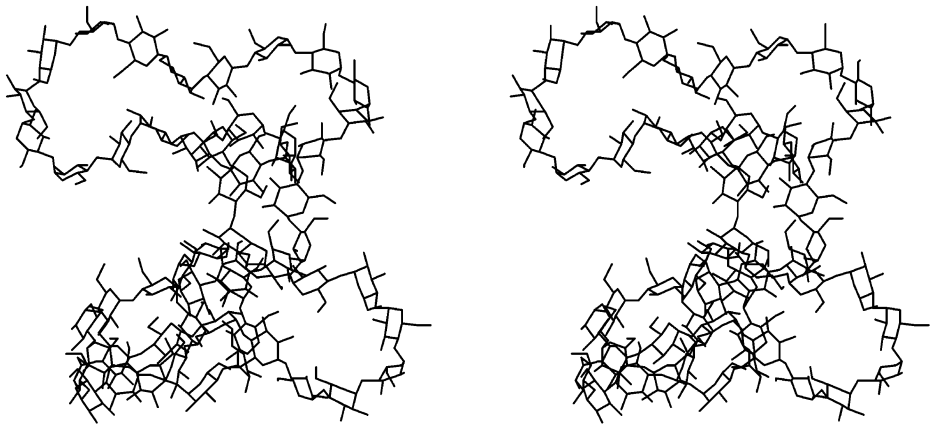
CD55-1.5



CD55-2.0



CD55-3.0



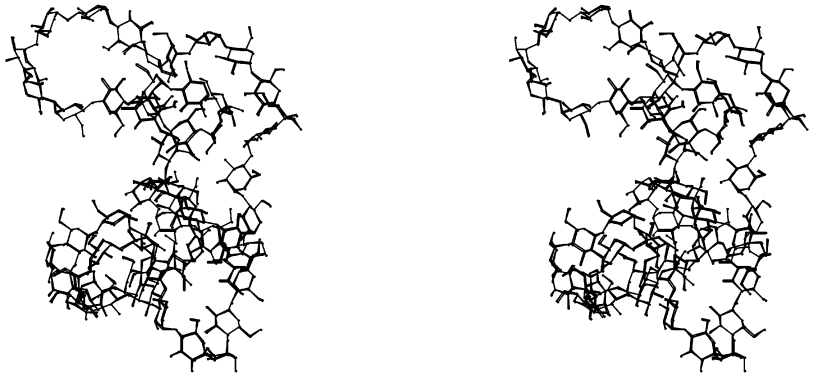


Figure 4. Stereodisplays of snapshots of geometries of CD55 from different stages of the simulation.

8). The moments of inertia of CD26 and CD30 had the least significant changes throughout the entire simulation times of all CDs studied. Three "flips" at residues 4, 15, and 22 were present in the starting structure and they persisted until the end of the simulation. None of them disappeared nor were new "flips" formed. A glance at Figure 8 reveals that this triple of glucoses presents alternating largest−smallest−largest average O4 to the center of mass distances. At the same time, sharp changes in the average populations of intramolecular hydrogen bonds involving the same residues were evaluated (Figure 12S).

The starting geometry of CD55 presented an open structure with a small helix at the broader portion of the macroring. A loop of six glucoses was formed at the beginning (Figure 4S, CD55-0.25) that opposed the helix along the perimeter of the macroring. A second helical fragment started to develop in the neighborhood of the loop (1.0 ns), and the structure evolved into two deformed helical portions that were connected by stretches with lengths of five (residues 46−50) and nine residues (from 22 to 30), respectively (1.5 ns). It seems that once a smaller loop of 6−7 glucoses is formed, the next natural mode for deformation of the macromolecule is toward the creation of a short single helix that further enhances the stability of the structure. The macroring contained at 2.0 ns one spiral of three turns (six glucoses per turn, with a longer but deformed channel compared with CD26) and a serpentine-form portion of six loops of varying sizes oriented differently in space (residues 19−55, plus 1). The snapshot at 3.0 ns displayed significant folding of the helical portion and four well shaped loops (CD55-3.0). A helix, two helical stretches, and three loops were present at the final stages of the simulation (Figure 4, CD55-4.7). The only "flip" introduced in the starting structure was also present at the end of the simulation. Flips were not present in the starting geometries of CD70, CD85, and CD100, and the simulations did not produce that structural motif.

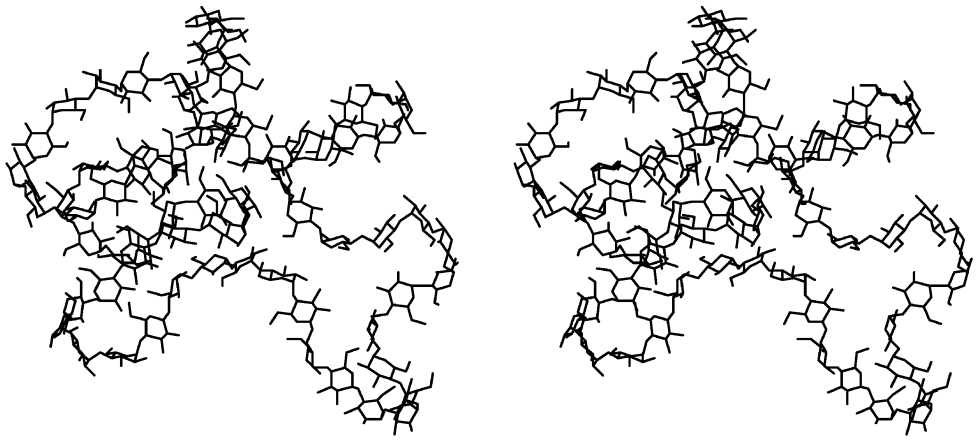
The analyses of the other three larger CDs also presented interesting structural motifs. Two deformed loops of fifteen and nine residues connected by deformed spiral fragments were contained in the starting structure of CD70. These motifs transformed at the initial stages of the simulation (Figure 4S, CD70-0.25) into an expanded three-turn single helix with eight glucoses per turn and a parallel cone-shaped spiral region. The two modules were linked by a small loop of seven residues from the one side and a twisted chain of nine glucoses from the other. The overall mode of dynamical deformation with the time was toward the creation of a more compact, folded central core that contains several overlapping loops. The macromolecule at the end of the simulation (CD70-5.0) acquired a geometry that could be partitioned into a helix + two loops (six and nine units) + a double helical portion. These were formed by the sequences of residues 15−36, 52−57, 41−49, and (61−70 plus 1−5), respectively (Figure 8). We can also conclude from the analyses of the variations with time of the dihedral angles Φ-(O5(n)−C1(n)−O4(n−1)−C4(n−1)), Ψ(C1(n)−O4(n−1)−C4-(n−1)−C3(n−1)), and "flip" (O3(n)···C4(n)···C1(n+1)···O2-(n+1)) that the local deformation "flip" introduces higher energy in the system than the single helical pattern but appears necessary for the formation of the favorable architecture of two antiparallel left-handed single helices in CD26 in the crystal. The presence of "flips" in LR-CDs, larger than CD26, does not significantly increase the energy of the solvated system. We did not encounter a single case of a "flip" appearing during the simulation; i.e., it also seems that the presence of "flips" is not always necessary in LR-CDs in solution.

The starting geometry of CD85 contained a seven-turn spiral portion with ends connected through a loop of thirteen glucoses. The overall appearance was of a more open structure and resembled to some extent the shape of the starting geometry of CD70. CD100 started from completely spiral geometry that closes a narrow elongated cavity. Local deformations prevailed at the beginning of the simulation, followed by more folded configurations of the macrocycles at later times. The big open cavity of CD85 disappeared at about 1 ns, and the structure at 2.0 ns resembled a dendritic fold with several arbitrarily oriented small loops on the surface of the clustering. Several six, seven, and eight residues loops were facing the exterior during the final stages of the simulation (Figure 6, CD85-4.5). These results further support the presence of more than one cavity in LR-CDs.

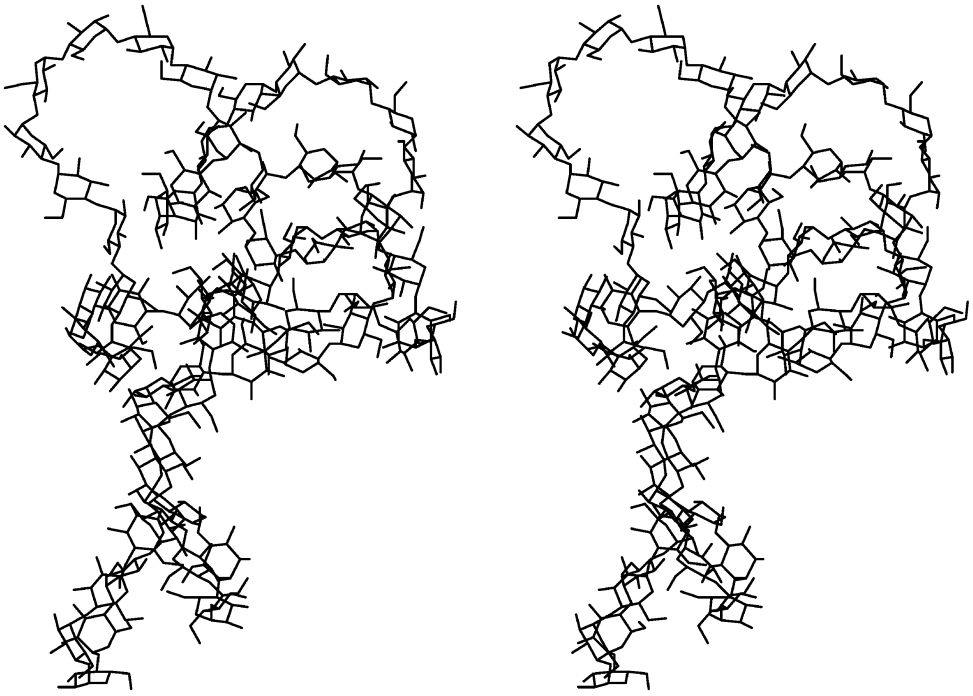
The initial mode of deformation of the CD100 structure was toward a closer approach of the two longer helical sections along the perimeter of the macroring. Three spiral portions were interconnected with a double helical strand after 1.0 ns MD simulation. A deep channel (Figure 8, residues 37−62; the highest intramolecular hydrogen bond population is for a residue of this portion of the macroring, Figure 12S) and several deformed loops were portrayed at 3.0 ns. A tendency for bending into two was observed at this stage that clearly manifested itself at the end of the simulation (Figure 7, CD100-5.0). After more pronounced conformational transitions at the initial stages of the simulations the moments of inertia of the larger CDs (CD55 to CD100) approached values about which they oscillate.

Outline of the Important Structural Parameters Describing the Conformations of the Large-Ring Cyclodextrins. More complete data of important geometrical parameters are contained in Table 2S in the Supporting Information. In agreement with the crystallographic data, the distances O4(n)·

CD70-0.25



CD70-2.0



CD70-3.0

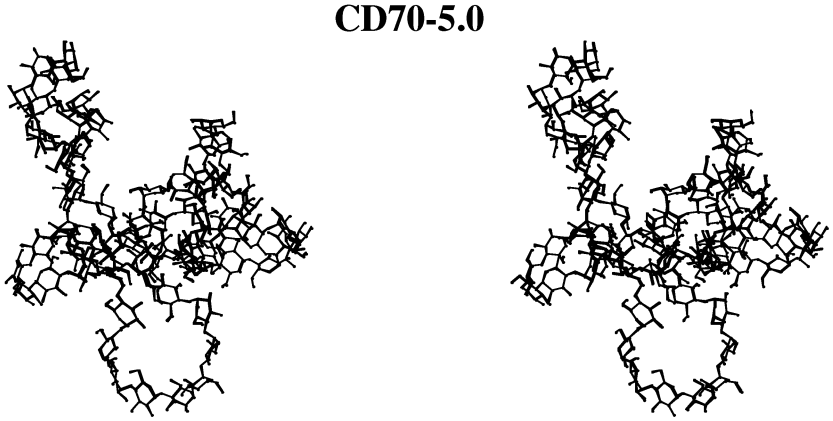


Figure 5. Stereodisplays of snapshots of geometries of CD70 from different stages of the simulation.

$\cdot\cdot$O4($n-1$) present almost the same values as for the helical parts of CD26, 4.4 Å. The distances O2(n)$\cdots$O3($n-1$) are rather small in the crystal, about 2.8−2.9 Å (Table 1S), as a consequence of the formation of intramolecular interglucose hydrogen bonds.[3] The computed average O2(n)$\cdots$O3($n-1$) distances (3.5−4.1 Å) are somewhat longer than the experimental ones. The same quantities in an earlier study[11] with the parm94 force field[17] were 4.2−4.8 Å, which signifies an improved performance of the new parm99 parametrization[18] in carbohydrate modeling. The larger computed values for the distances O2(n)$\cdots$O3($n-1$) are due partly also to the loss of the alignment (partial flipping) between segments of the macrocycles such that clusters of several glucoses form hydrogen bonds with other units further along the ring, breaking the continuity of the belt of hydrogen bonds between neighboring glucoses. It may also indicate still unsatisfactory modeling of the effect of the bridging intermolecular water hydrogen bonds, formed with the hydroxyls at C2(n) and C3($n-1$) across the glycosidic linkages.[23b] In fact, distances of 5.5 Å between neighbor glucoses are described for CD10 and CD14.[14] The shortest computed O2(n)$\cdots$O3($n-1$) contacts are 2.4 Å and indicate the presence of strong hydrogen bonds.

The total number of hydrogen bonds increases with increasing ring size (Figure 10). The large CDs have folded and unfolded portions and hydrogen bonds are thus also formed between glucose residues that are more distant in the macrocycle and the overall expected effect is a slight increase in the number of hydrogen bonds per monomer with increasing the size of the macroring. The higher flexibility of the larger macromolecules facilitates the formation of hydrogen bonds between distant atoms on the macroring but closely spaced pairs of glucose residues. Only CD30 deviates from the general tendency in the series regarding the participation of the hydroxyl groups in hydrogen bonds—the smallest value for "hb1" per residue is estimated in this case. This result reflects the presence of a comparatively more open structure for CD30 during the simulation (vide supra). CD26 possesses helices and, accordingly, the numbers of hydrogen bonds per residue are larger than in CD30. The number of hydrogen bonds with proton donors from the water molecules (hb2) increases per residue with the ring size—the larger macrorings have much more exposed surfaces for hydrogen bonding with the water molecules.

Variations of 3.5° were obtained for the computed bond angles at the glycosidic oxygens (C1(n)−O4($n-1$)−C4($n-1$)), and the average values were about 116°. The experimental determinatons display significant differences in the C1(n)−O4($n-1$)−C4($n-1$) values in the sequence CD6 → CD8. This bond angle is 119.0° in CD6, and diminishes to 112.6° in CD8. According to the crystal structural data, the angle O4(n)$\cdots$O4($n-1$)$\cdots$O4(n−2) increases from 119.9° in CD6 to 134.9° in CD8. The computed data for CD26 are in agreement with the experimental determinations. The computed average values for the other CDs more closely resemble the crystallographic data for CD8, CD9, CD10, and CD14.

The dihedral angles defined by the glycosidic oxygens are the measure for their coplanarity. The average values computed for the series vary in the range −19.0° to −40.0°. The O3(n)$\cdots$C4(n)$\cdots$C1($n+1$)$\cdots$O2($n+1$) dihedrals determine the relative orientation of the secondary hydroxyl groups of adjacent glucose residues. These dihedral angles have zero values for the CDs with the geometry of a truncated cone. A value 180.0° for O3(n)$\cdots$C4(n)$\cdots$C1($n+1$)$\cdots$O2($n+1$) indicates a rotation by this angle of one glucose monomer, which in effect results in an exchange of the positions of the primary and the secondary hydroxyls, a "flip". The average computed values were determined in the range −30.0° to −42.0°.

The angles Φ(O5(n)−C1(n)−O4($n-1$)−C4($n-1$)) and Ψ-(C1(n)−O4($n-1$)−C4($n-1$)−C3($n-1$)) describe the orientation of the residues about the $\alpha(1\rightarrow4)$ glycosidic linkage (adopting the definitions of Φ and Ψ utilized by Saenger and coworkers[3,7,15]). The glucose residues have the same orientations in the crystal phase if the two dihedrals acquire values in the ranges 94.0° to 110.0° and 97.0° to 135.0°, respectively.[7] When a glucose unit is rotated at about 180°, then these angles have values 82.0° to 84.0° (Φ) and −65.0° to −69.0° (Ψ).[7] Φ and Ψ have computed average values from 87.3° (CD30) to 104° (CD100), and from 75.0° (CD30) to 103.2° (CD70), respectively. The rms deviations are ca. 20° in all cases.

In addition to the moments of inertia, estimated with the CARNAL module of AMBER,[17] three more descriptors of the molecular size and anisometry have been also evaluated to assess differences in size and shape between the CDs (Table 1):[26] radius of gyration (also provides an absolute measure of compactness), span (defined as the radius of the smallest sphere, centered at the center of mass, which encloses the distribution of points), and asphericity (defined in terms of all moments of inertia and measures the deviation from a spherical form). The DRAGON program[27] was used. Due to size limitations, only the C1−O4−C4 fragment (CD26 and CD30) or only the O4 atom (CD55 to CD100) of each glucose unit were used when evaluating the descriptors. Tests of the "one-atom" (O4) and the "three-atoms" (C1−O4−C4) presentations on estimates of the radius of gyration, the span, and the asphericity for one and the same geometry of CD26 yielded identical values. The values

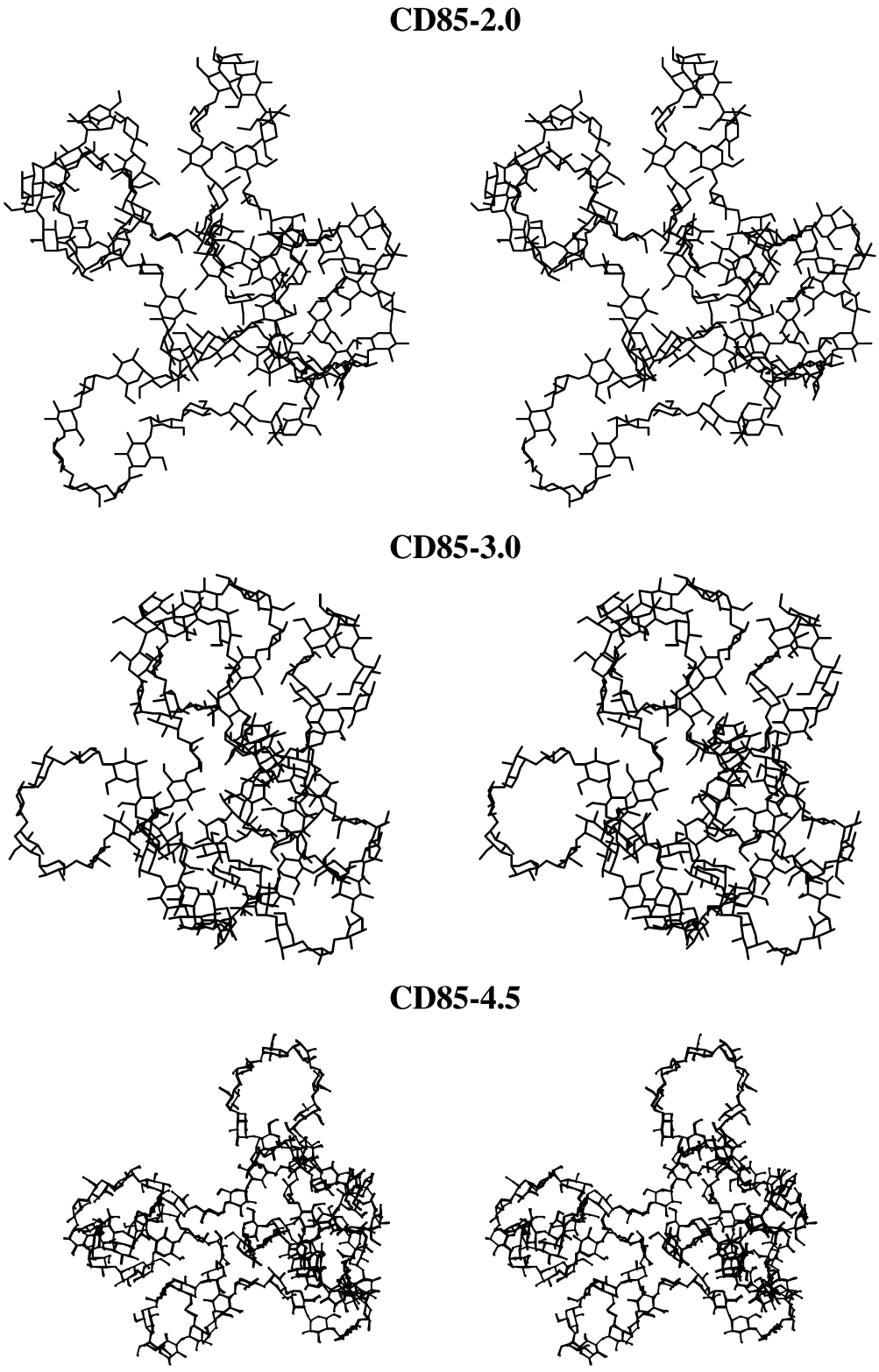


Figure 6. Stereodisplays of snapshots of geometries of CD85 from different stages of the simulation.

reported were obtained by averaging over all snapshots picked up from the trajectories.

The radius of gyration and the span increase with the CD size. The asphericity values are smaller than 0.3, indicating similarity with a cylindrical shape. Only CD30 deviates from these trends. CD30 has the less compact configuration of the macroring (only CD85 and CD100 have values for the radius of gyration higher than CD30). The comparable values for the span of CD70 and CD85 and the difference in their radii of gyration correlate with the tendency we noticed for CD70 toward the creation of a more compact, central core of residues, and with the rounded shape of the dendritic fold of CD85 with small loops on the surface of the clustering (CD85 has the smallest value for asphericity).

Energies of the Large-Ring Cyclodextrins Estimated with the MM/GBSA Method. Small variations in the total energies per residue were obtained ($E_{(\text{total,GB})}$, Table 2). The polarization contribution to the energy of solvation, E(GB), lowers the total

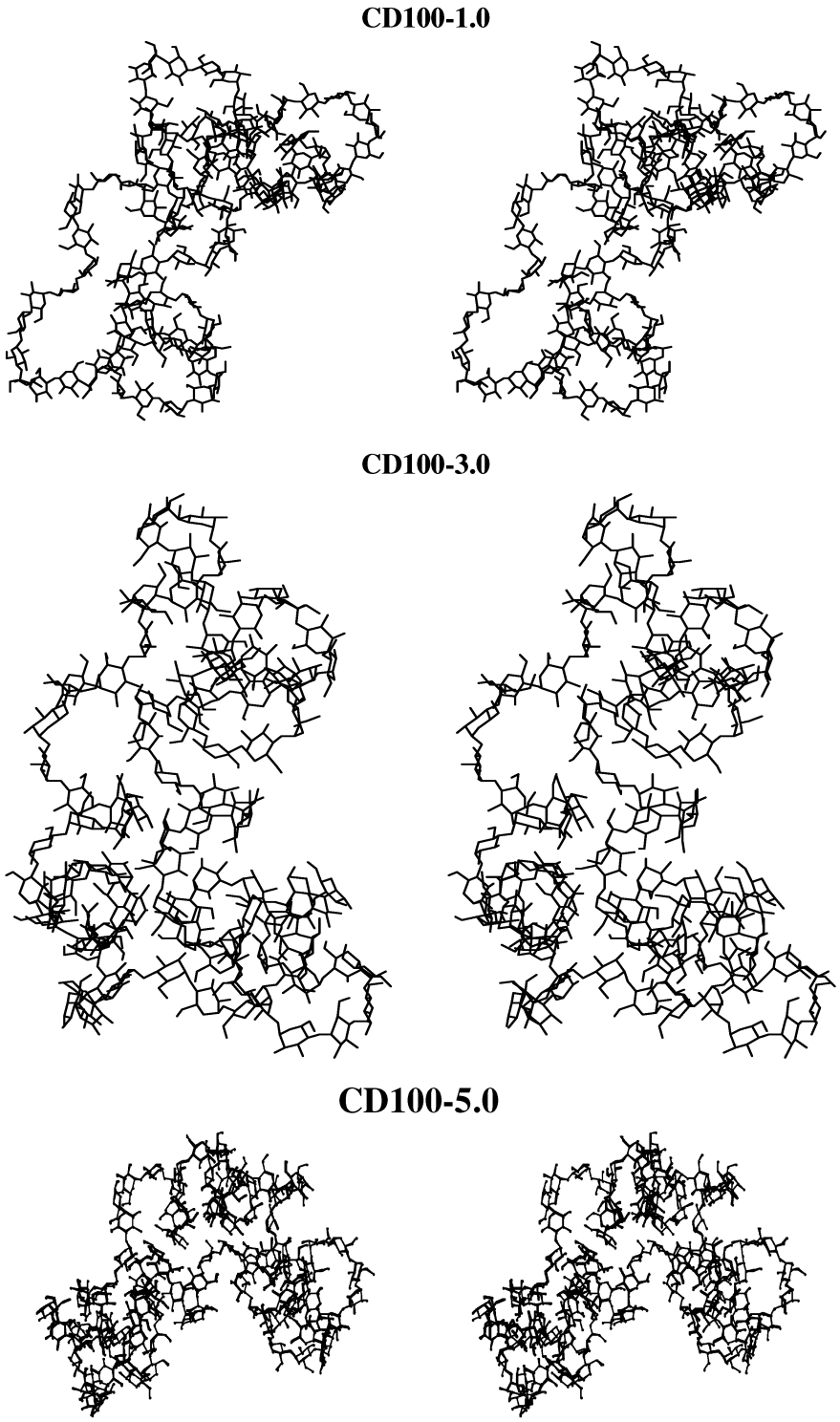


Figure 7. Stereodisplays of snapshots of geometries of CD100 from different stages of the simulation.

energy per monomer by about 11−15 kcal mol^{-1}, whereas the electrostatic and the internal energy terms contribute mostly for the resulting averaged per residue MM/GBSA energy values. CD30 has the most disadvantageous energy contribution from the nonbonded interactions (van der Waals + electrostatic), and this result corroborates the smaller population of hydrogen bonds (Figure 10) and the appearance of a more open structure in this case. CD30 has at the same time the most favorable contribution from the polarization component of the solvation energy. A mutual compensation between the contributions of the nonbonded and the solvation energy terms was monitored in the simulations (Table 2, Figure 11S, Figure 13S). We noticed analogous behavior for a balance between electrostatic and solvation energy components in earlier studies on inclusion properties of native CDs.[28] Such a compensating balance of energy contributions modeling different interactions is a prerequisite for high flexibility of the structure.

Conclusions

The elucidation of the dynamics of the structural deformations of the large CDs is important for the correct interpretation of the complex-forming properties of these molecules. The present

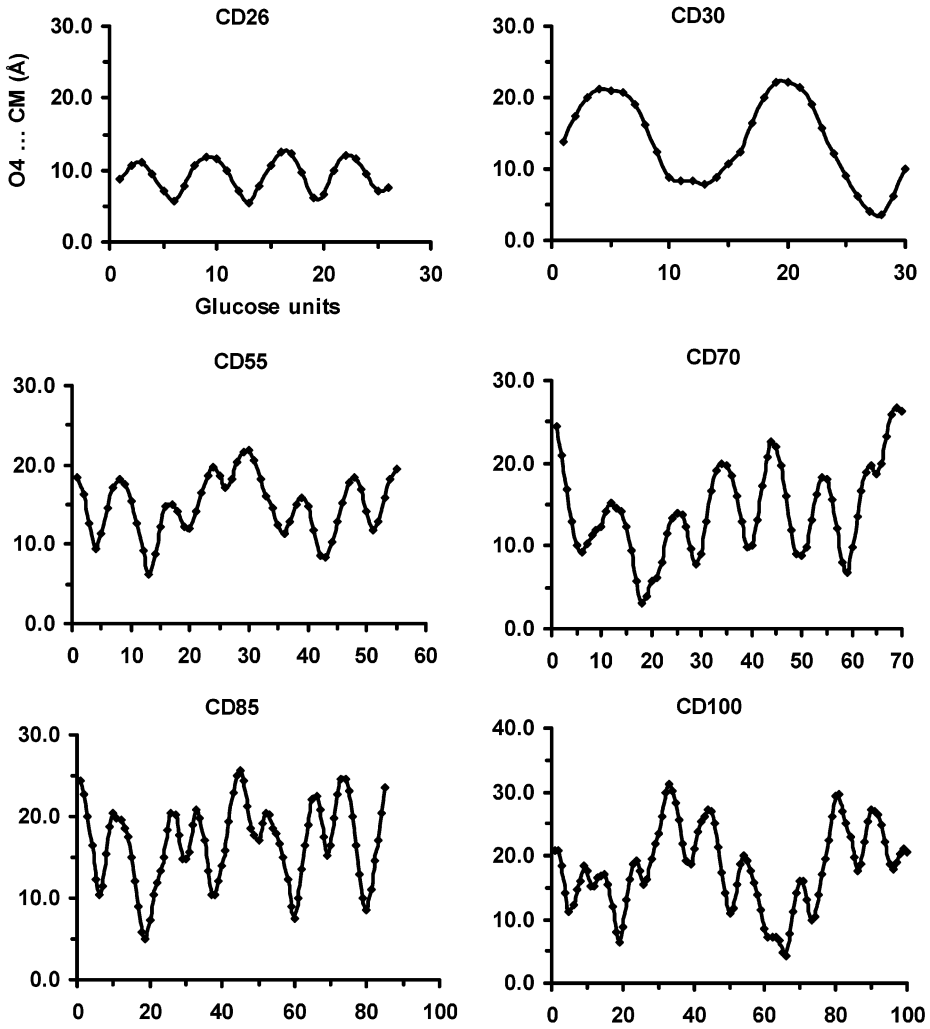


Figure 8. Variation of the time average values (5.0 ns simulations) of the O4 to the center of mass distances (O4···CM) evaluated for each residue.

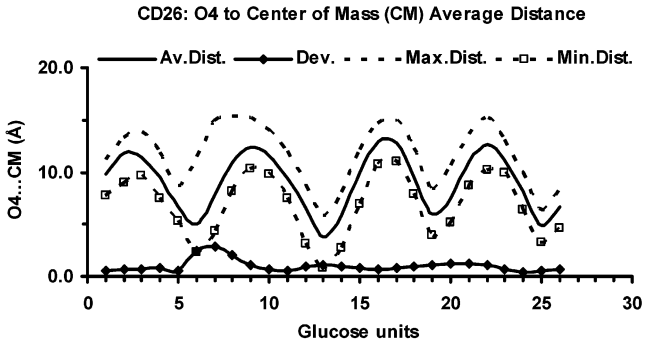


Figure 9. Variation of the time average values of the O4 to the center of mass distance evaluated for each residue of CD26 for the simulation interval from the fifth to the tenth nanosecond. Rms deviations and maximum and minimum distances are also given. The asymmetry of the curves reflects the difference in flexibility of two 13-residue portions of CD26.

results provide at least an overall view of the deformations that these structures could exhibit, as well as the potential of portions of the macrorings to form cavities of different sizes and shapes that may host small molecules.

The folding and bending of the residues in the large CDs vary. Tendencies for the creation of spiral folded fragments, as well as overall folding of the structures in water, were also observed. The crystal structure of CD26 is symmetrical and consists of two antiparallel, left-handed, single helices of almost two turns, connected by two band-flipped glycosidic bonds. The structure of each helix resembles V-amylose with six glycosyl units per repeating turn. The structures acquired by CD26 during the MD simulation in water do not correspond to the conformation in the crystalline state. The two "flips" present in the macroring of CD26 in the crystal state disappeared after 1.5 and 3.0 ns simulations, respectively.

The band-flip motif has only been observed in the crystal structures of CD10, CD14, and CD26. So far, no evidence, besides the crystal structures, has been obtained to support its existence in solution. It has been suggested that band-flips are induced by conformational strain.[7,14,29] The band-flips could be considered also as an allowed structure, due to the higher conformational freedom of the macrocycle and thus not necessarily a strain-induced conformation.[2] The present simulations in water solution give some degree of support to both views. The disappearance of the two "flips" in CD26 during the simulation in water could be interpreted as an indication of an additional strain in the macroring when "flips" are present. In none of the cases studied did we witness the formation of a new "flip". However, "flips" introduced in the starting structures of the larger CDs also remained present at the end of the simulations; i.e., the high conformational freedom of the large macrorings is not significantly perturbed by the presence of a "flip". The preferred formation of helical fragments that was noticed during the MD simulations of the large CDs may eventually require a "flip" when such two portions are very close

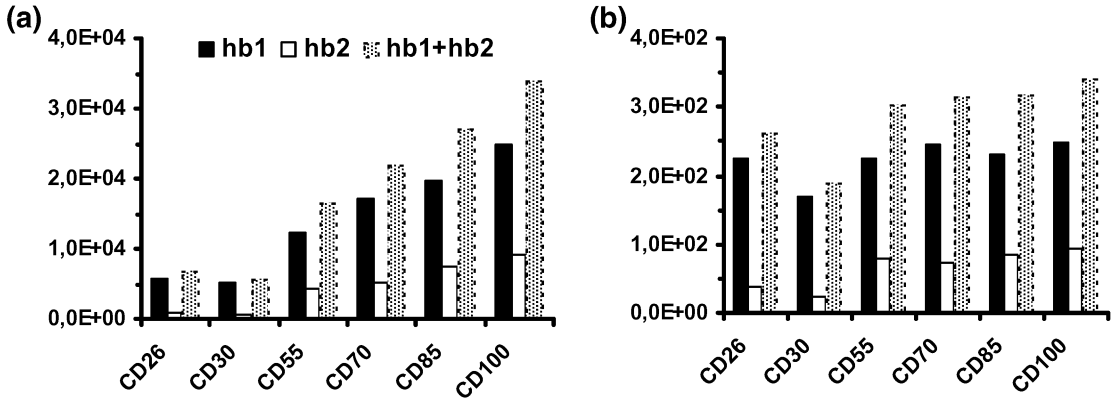


Figure 10. Graphical representation of the population (%) of hydrogen bonds obtained for each CD from the MD simulations in water solution. hb1 = intramolecular + hydrogen bonds to oxygen atoms of water molecules (i.e., this quantity is a measure of the occupancy of the hydroxyl groups of the macromolecule to participate in hydrogen bonding as proton donors); hb2 = hydrogen bonds of water molecules as proton donors to oxygen atoms of the macroring. (a) Sum of all hydrogen bond populations. (b) Ratio between the sum of all hydrogen bond populations and the number of glucose residues for each CD. A quantity of 100 for the hydrogen bond population corresponds to the existence on average of one hydrogen bond for the corresponding residue.

TABLE 1: Computed Average Values of Radius of Gyration (Å), Span (Å), and Asphericity

	CD26	CD30	CD55	CD70	CD85	CD100
radius of gyration	25.6[a]	41.0	35.1	36.6	48.5	55.8
span	14.2	23.6	24.6	26.8	28.1	32.5
asphericity	0.25	0.55	0.24	0.20	0.16	0.29

[a] An experimental estimate 19.6 Å is cited in ref 2. Values 31.9 and 17.7 Å were also determined from molecular dynamics simulations in the gas phase and in water solution, respectively.[11b]

TABLE 2: Contributions per a Monomer Unit of Different Energy Terms to the Total MM/GBSA Energies of the CDs (kcal mol^{-1})[a]

	CD26	CD30	CD55	CD70	CD85	CD100
$E_{\text{electrostat}}$	16.1	17.9	16.4	15.9	16.2	15.9
E_{vdW}	−0.8	1.0	−0.1	−1.0	−0.3	−0.7
E_{internal}	38.3	37.5	37.9	37.9	37.8	38.0
E_{gas}	53.6	56.3	54.2	52.8	53.8	53.2
$E_{(\text{nonpolar})}$	0.6	0.7	0.6	0.5	0.6	0.5
$E_{(\text{GB})}$	−11.7	−14.6	−12.4	−11.0	−12.0	−11.1
$E_{(\text{solvation})}$	−11.2	−13.9	−11.8	−10.5	−11.4	−10.5
$E_{(\text{GB+elect})}$	4.4	3.3	4.0	4.9	4.8	4.8
$E_{(\text{total,GB})}$	42.4	42.4	42.4	42.3	42.3	42.6

[a] $E_{\text{electrostat}}$ = electrostatic energy; E_{vdW} = van der Waals energy; E_{internal} = internal energy (stretching + bending + torsional); E_{gas} = gas-phase energy (electrostatic + van der Waals + internal); $E_{(\text{nonpolar})}$ = the nonpolar component of the solvation energy (surface area energy); $E_{(\text{GB})}$ = the polarization component of the solvation energy in the generalized Born solvation model; $E_{(\text{solvation})} = E_{(\text{nonpolar})} + E_{(\text{GB})}$; $E_{(\text{GB+elect})} = E_{(\text{GB})} + E_{\text{electrostat}}$; $E_{(\text{total,GB})} = E_{\text{gas}} + E_{(\text{solvation})}$.

along the perimeter of the macroring. Most probably, under the conditions of the crystal packing forces the barrier for a "flip" to occur can be overcome, and nonbonded interactions between such neighbor helices could counterbalance the strain introduced by the band-flip.

Our interest in the large CA's arose from the expectation that they could have the potential to form inclusion complexes. The small CDs (the native CDs) have clearly shaped cavities and can form inclusion complexes. The intermediate CD14 possesses more deformed structures with cavities not enough clearly seen.[7,13] Narrow, but deeper, cavities could eventually manifest amplification effects for binding small molecules.[30] The X-ray structures of two cyclomaltohexaicosaose (CD26) triiodide inclusion complexes (with NH_4I_3 and $Ba(I_3)_2$) were reported.[31] It has been already observed for CD50 (unpublished results)[3] that in the presence of alcohols or acids, the long chain makes a turn that probably arises from the formation of an inclusion complex. Different aspects of the complex-forming properties of the large-ring cyclodextrins were also recently addressed.[32]

The large-ring CDs we studied display cavity-like portions of different sizes and shapes: circular and elongated loops of variable size orientated in different fashions; portions of a double helical strand with the two single helices parallel to each other (CD30, CD70); a helix of three turns and a serpentine form portion of six loops (CD55); a cone-shaped spiral region (CD70); a rounded dendritic fold with several arbitrarily oriented small loops on the surface of the clustering (CD85); a deep channel and three spiral portions with the tendency for bending onto two (CD100). The present results support the hypothesis for the existence of more than one cavity in large-ring cyclodextrins.[2] In view of the difficulties in the purification of individual large-ring cyclodextrins, molecular dynamics simulation techniques provide a useful tool to gain insight into the complex-forming ability of individual CDs as components of mixtures of LR-CDs that possess such properties.[2] It has been reported that LR-CD mixtures with a degree of polymerization from 22 to 45, and greater than 50 exhibit an efficient artificial chaperone for protein refolding.[33] It is expected also that LR-CDs may provide useful means for molecular manipulation of carbon nanotubes.[34]

Acknowledgment. This work was performed mainly using the CESCA facilities (Catalonia, Spain). We thank the "Ministerio de Ciencia y Tecnología" (Spain) for financial support through grant no. PPQ2000-0369. We also thank CIRIT (Generalitat de Catalunya, Catalonia, Spain) and the "Ministerio de Educación, Cultura y Deporte" (Spain) for providing grants to P.M.I. as a visiting professor.

Supporting Information Available: Figures of initial geometries, the variations with time of total energies, rms deviation curves of atomic coordinates relative to a reference structure, O4 to center of mass distances, and average values of individual energy terms (MM/GBSA) and tables (the main computed and experimental structural parameters), as well as some other details. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) (a) *Chem. Rev.* **1998**, *98*(5), all pages. (b) Connors, K. A. *Chem. Rev.* **1997**, *97*, 1325−1357.

(2) Larsen, K. L. *J. Inclusion Phenom. Macrocycl. Chem.* **2002**, *43*, 1−13.

(3) Saenger, W.; Jacob, J.; Gessler, K.; Steiner, T.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T. *Chem. Rev.* **1998**, *98*, 1787−1802.

(4) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1987**, *15*, 211−224.

(5) Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1988**, *16*, 153−168.

(6) Kitamura, S.; Isuda, H.; Shimada, J.; Takada, T.; Takaha, T.; Okada, S.; Mimura, M.; Kajiwara, K. *Carbohydr. Res.* **1997**, *304*, 303−314.

(7) Jacob, J.; Gessler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 606−609.

(8) Lipkowitz, K. B. *Chem. Rev.* **1998**, *98*, 1829−1873.

(9) (a) Shimada, J.; Handa, S.; Kaneko, H.; Takada, T. *Macromolecules* **1996**, *29*, 6408−6421. (b) Shimada, J.; Kaneko, H.; Takada, T.; Kitamura, S.; Kajiwara, K. *J. Phys. Chem. B* **2000**, *104*, 2136−2147.

(10) Momany, F. A.; Willett, J. L. *Carbohydr. Res.* **2000**, *326*, 210−226.

(11) (a) Beà, I. Ph.D. Thesis, Universitat Autònoma de Barcelona, Barcelona, 2001. (b) Maestre, I.; Beà, I.; Ivanov, P. M.; Jaime, C. *J. Mol. Struct*., submitted for publication.

(12) Fujiwara, T.; Tanaka, N.; Kobayashi, S. *Chem. Lett*. **1990**, 739−742.

(13) Ueda, H.; Endo, T.; Nagase, H.; Kobayashi, S.; Nagai, T. *J. Inclusion Phenom. Mol. Recognit. Chem.* **1996**, *25*, 17−20.

(14) Jacob, J.; Gessler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W. *Carbohydr. Res.* **1999**, *322*, 228−246.

(15) (a) Gessler, K.; Usón, I.; Takaha, T.; Krauss, N.; Smith, S. M.; Okada, S.; Sheldrick, G. M.; Saenger, W. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 4246−4251. (b) Nimz, O.; Gebler, K.; Usón, I.; Saenger, W. *Carbohydr. Res.* **2001**, *336*, 141−153.

(16) Koizumi, K.; Sanbe, H.; Kubota, Y.; Terada, Y.; Takaha, T. *J. Chromatogr. A* **1999**, *852*, 407−416.

(17) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheathan, T. E., III; Wang, J.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K. M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Gohlke, H.; Radmer, R. J.; Duan, Y.; Pitera, J.; Massova, I.; Seibel, G. L.; Singh, U. C.; Weiner, P. K.; Kollman, P. A. *AMBER 7*, University of California: San Francisco, 2002.

(18) Wang, J.; Cieplack, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049−1074.

(19) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(20) (a) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092. (b) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577−8593. (c) Sagui, C.; Darden, T. A. In *Simulation and Theory of Electrostatic Interactions in Solution*; Pratt, L. R., Hummer, G., Eds.; American Institute of Physics: Melville, NY, 1999; pp 104−113. (d) Toukmaji, A.; Sagui, C.; Board, J.; Darden, T. *J. Chem. Phys.* **2000**, *113*, 10913−10927.

(21) (a) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440−467. (b) *MacroModel/BatchMin*, v. 5; Department of Chemistry, Columbia University, New York, 1995. The AMBER* force field is a well-documented version of AMBER force field, implemented in MacroModel/BatchMin, v.5 (see pp 81−82 of the BatchMin Reference Manual).

(22) Melani, F.; Mura, P.; Adamo, M.; Maestrelli, F.; Gratteri, P.; Bonaccini, C. *Chem. Phys. Lett.* **2003**, *370*, 280−292.

(23) (a) Kuttel, M.; Brady, J. W.; Naidoo, K. J. *J. Comput. Chem.* **2003**, *23*, 1236−1243. (b) Chen, J. Y.; Naidoo, K. J. *J. Phys. Chem. B* **2003**, *107*, 9558−9566.

(24) Tsui, V.; Case, D. A. *Biopolymers (Nucl. Acid. Sci.)* **2001**, *56*, 275−291.

(25) Weiser, J.; Shenkin, P. S.; Still, W. C. *J. Comput. Chem*. **1999**, *20*, 217−230.

(26) Arteca, G. A. in *Reviews in Computational Chemistry*; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1996; Vol. 9, Chapter 5, pp 191−253.

(27) Todeschini, R.; Consonni, V.; Mauri, A.; Pavan, M. DRAGON, version 3.0, 2003.

(28) (a) Ivanov, P. M.; Jaime, C. *Anal. Quím. Int. Ed.* **1996**, *92*, 13−16. (b) Ivanov, P. M.; Jaime, C. *J. Mol. Struct.* **1997**, *377*, 137−147. (c) Ivanov, P. M.; Salvatierra, D.; Jaime, C. *J. Org. Chem.* **1996**, *61*, 7012−7017.

(29) Harata, K.; Endo, T.; Ueda, H.; Nagai, T. *Supramol. Chem.* **1998**, *9*, 143−150.

(30) Whitlock, B. J.; Whitlock, H. W. *J. Am. Chem. Soc.* **1994**, *116*, 2301−2311.

(31) Nimz, O.; Gessler, K.; Usón, I.; Laettig, S.; Welfle, H.; Sheldrick, G. M.; Saenger, W. *Carbohydr. Res.* **2003**, *338*, 977−986.

(32) *J. Inclusion Phenom. Macromol. Chem.* **2002**, *44*, 1−4, all pages.

(33) Machida, S.; Ogawa, S.; Xiaohua, S.; Takaha, T.; Fijii, K.; Hayashi, K. *FEBS Lett.* **2000**, *486*, 131−135.

(34) Dodziuk, H.; Ejchart, A.; Anczewski, W.; Ueda, H.; Krinichnaya, E.; Dolgonos, G.; Kutner, W. *Chem. Commun*. **2003**, 985−987.