Large-ring cyclodextrins. A molecular dynamics study of the conformational dynamics and energetics of CD10, CD14 and CD26

Martin G. Gotsev and Petko M. Ivanov*

Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, ul. Acad. G. Bonchev, bloc 9, 1113 Sofia, Bulgaria
E-mail: ivanov@bas.bg

Abstract

Molecular dynamics simulations in water were performed on the three largest large-ring cyclodextrins (LR-CD) for which X-ray data is available – CD10, CD14 and CD26. The Glycam-04 force field in AMBER and explicit water molecules (TIP3P) were used in the 20.0 ns simulations. Small variations about the starting conformation of CD10 were detected. Different structural motifs were monitored for CD14 that may represent chiral species of particular interest for exploring supramolecular and chiral molecular recognition effects: broadly opened macroring, that resembles the initial geometry of CD10 but in larger scale, a big circular loop with a small helical turn, deformed figure eight conformation, and a symmetrically squeezed open form ("a dumbbell"). The final geometry of CD14 is free of additional strain introduced by band-flips. The preferred conformation of CD26 contains a small helix and an extended helical portion that transforms also to an arc and a loop. Two modifications of the equilibration step of the simulation protocol were tested.

Keywords: Cyclodextrins, molecular dynamics, conformational analysis, macroring shapes, Glycam-04 force field, AMBER

Introduction

The large-ring cyclodextrins (LR-CDs)[1,2,3,4] attracted attention in recent years, and advances were marked in the study of their physicochemical properties,[5,6,7,8] in spite of existing difficulties in their synthesis,[9,10,11] isolation and purification.[8,12,13,14] The X-ray structures of the CDs composed by 9,[15] 10,[13,16,17] 14,[13,16] and 26[18,19] glucose units provided invaluable information on the conformations of these molecules in the crystal lattice.[20] New structural motifs were detected. The thermal and structural characterization of δ-CD, ε-CD, and ι-CD have been carried out,[21] and references to CDs with more than 60,[22] and several hundred[23] glucose units have been made.

CDs with more than 100 glucoses in the ring have been prepared by the action of disproportionating enzyme on amylase.[23]

Similarly to the case of native CDs (α-CD, β-CD, γ-CD), the study of the complexation properties of LR-CDs has also begun. Since CDs consist of optically active D-glucose units, they form with a racemic compound a pair of diastereoisomeric complexes, usually of different stability. Therefore, LR-CDs, as the native CDs,[24,25,26] are potential reagents for chiral resolution, that is of practical importance for separation science (molecular recognition and separation of closely related compounds, including geometrical and structural isomers) and for the pharmaceutical industry (different pharmacological activity of enantiomers of a chiral compound).[24] Besides, inclusion complex formation seems not to be a prerequisite for chiral recognition. CDs are able to form sufficiently strong external complexes enantioselectively with some chiral guest molecules.[26] LR-CDs may be good host molecules for relatively large guest compounds.[5] Alternatively, microanalytical separation techniques for enantiomers like capillary electrophoresis may serve as an useful source of experimental data (structure, dynamics, populations of the complexes) for evaluating reliability of molecular modeling protocols.[26] Ueda summarized in a mini-review the results about LR-CDs with regard to the potential for host-guest interactions and corresponding applications (commercially available CD-mixtures containing LR-CDs with a degree of polymerization from 9 to 21 were examined as food additives in Japan;[5] LR-CDs mixtures with a degree of polymerization from 22 to 45, and greater than 50 exhibited an efficient artificial chaperone effect for protein refolding (the first practical application of LR-CDs)[27]). The inclusion complexes of CDs of up to 17 glucose units with several benzoates, salicylate, ibuprofen anion and 1-adamantane carboxylate have been studied and their formation constants have been determined.[28] Association constants for those CDs having more than 10 glucoses are usually small (< 40 M^{-1}). However, the LR-CDs with 21-33 glucose units form stable inclusion complexes with iodine in aqueous solution as demonstrated by isothermal titration calorimetry (formation constants of about 1-7 10^3 M^{-1}).[29] δ-CD has demonstrated to form a stable complex with C_{70} that allows its solubilization in water.[30] It has been proved that η-CD (12 glucoses) is effective in the partial separation of carbon nanotubes.[31]

The structural features of the native and the LR-CDs, mainly X-ray structural analyses, have been surveyed and computational studies have also been quoted.[2,16,32,33,34] Reviews were especially devoted to the applications of molecular modeling techniques to the study of the static and the dynamical features of CDs, as well as their participation in host-guest complexation.[35,36] However, they are mainly centered in native CDs. Several studies used molecular dynamics (MD) simulations (with lengths of 100-400 ps) to examine the macroring conformations of LR-CDs containing 10, 14, 18, 21, 24 and 48 glucose units either in vacuum or in water solution.[37,38,39,40] The circularized three-turn single helical structure proposed for CD21 from small-angle X-ray scattering was shown, with rather short MD simulations (100 ps), to persist in water at 300 K.[34]

Recently we reported results from study on CDs with a degree of polymerization of 26, 30, 35, 40, 55, 70, 85 and 100, using much longer molecular dynamics simulations (5.0 ns for all

except for the one of 26 for which the MD was of 10.0 ns).[41,42] Although the results were later published,[43] the latter study was accompanied by examining the force field dependence of the modeling of LR-CDs by performing 5.0 ns simulations of CDn (n = 14, 21, 26, 28), as well as the native CDs, as isolated molecules (gas phase) and in aqueous solution, using four different force fields - MM3*,[44] (within Macromodel)[45,46] and the parm94,[47] parm99,[48] and glycam2000a[49] parameterizations of AMBER.[50] Two molecular dynamics studies focused on several LR-CDs flanking CD26, the largest CD for which X-ray data is available, using different starting geometries: *(i)* 5.0 ns simulations of CDn (n = 24, 25, 26, 27, 28, 29) with CD30-derived starting geometries and the AMBER parm99 force field as a molecular mechanics model,[51] and *(ii)* 10.0 ns simulations of CDn (n = 24, 25, 27, 28, 29, 30) with CD26-derived starting geometries[52] and the new Glycam-04 force field for carbohydrates.[53] Both, the monitoring of the structural variations during the simulations, as well as the analyses of energy balances were indicative for high flexibility of the macrorings. Different local structural motifs were detected. The results provided further support to the hypothesis for the existence of more than one cavity in large-ring cyclodextrins and suggested preferred conformations in water solution for the LR-CDs with degree of polymerization from 24 to 30.

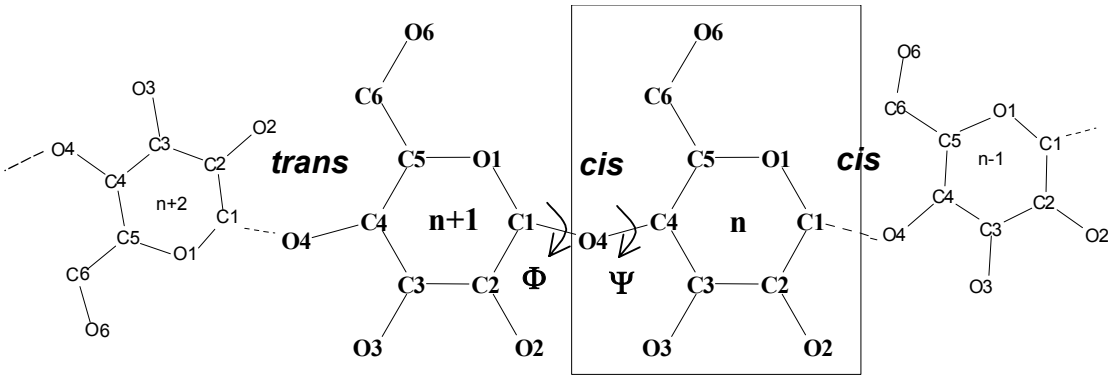


Figure 1. Schematic representation for a CD fragment showing the atomic numbering. Each individual glucose unit is designated by a number "n". Atoms are identified by a letter and a number that indicates its position in the glucose.

The latter simulations were not long enough to examine in full the conformational dynamics of these flexible molecules, but they provided useful information about the conformational dynamics and energetic of LR-CDs. We proceed now further by examining in more detail the conformations of the three largest LR-CDs for which X-ray data is available – CD10, CD14 and CD26, using X-ray starting geometries, much longer simulation times (20.0 ns), and the new Glycam-04 force field in AMBER for carbohydrates.[53] Minor variations in the previous simulation protocol[52] were also made, and they appeared to be of importance for the analyses of some properties, like hydrogen bonds distributions.

CDs are usually named using Greek letters as prefixes. For a better understanding, hereinafter the nomenclature used will be based on that adopted for the cycloamyloses: a number added after CD designates the number of units in the macrocycle, e.g., CD24 is the cyclodextrin with 24 glucose residues. Figure 1 displays a schematic representation of a cyclodextrin fragment with the numbering of the atoms, as well as examples for *cis* and *trans* orientations of neighboring glucose units.

Computational Details

All computations were carried out with the AMBER program (version 7[50] and version 8[53]; the AMBER modules LEaP, SANDER and CARNAL were used, respectively, for the preparation of the input data, minimization and the MD simulation steps, and analysis of the MD trajectories) using the most recent parameterization for carbohydrates, Glycam-04.[53] The molecular dynamics simulations were run for water solution (a box with TIP3P[54] water molecules) using the particle mesh Ewald (PME) method[55,56,57,58] 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 cubic solvent box around the CD with buffer distances of 15.0 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: CD10 (51.9 Å; 4110); CD14 (56.8 Å; 5540); CD26 (60.1 Å; 6480). The SHAKE option (tolerance 0.00005 Å) was used for constraining bonds involving hydrogen atoms. Figure 2 presents the starting geometries from X-ray analyses.

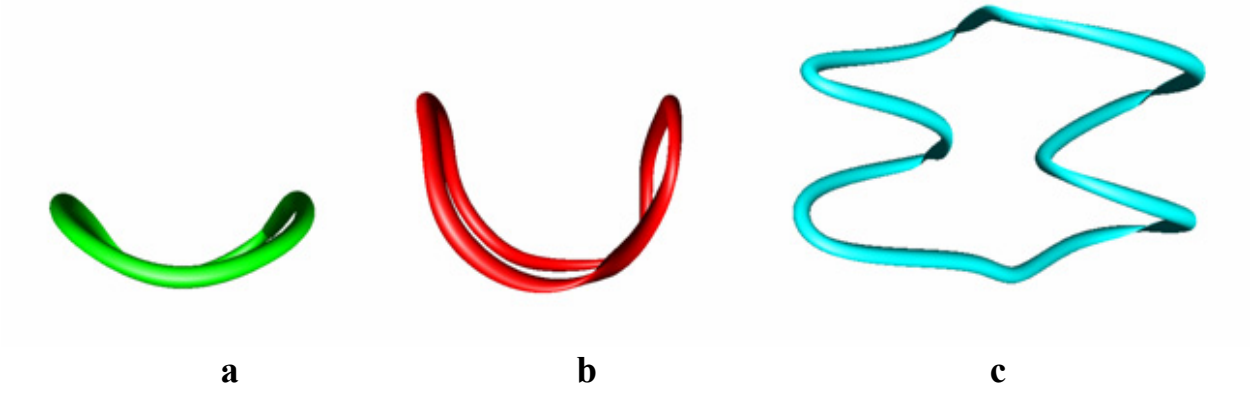


Figure 2. Starting conformations from X-ray. a) CD10 (green): [4s,a,4s,a]; b) CD14 (red): [6s,a,6s,a]; c) CD26 (cyan): [12s,a,12s,a].

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)* 50000 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 water relaxation takes place); *iv)* gradual release of the restraints on the LR-CD in a series of minimizations and MD steps: 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 100.0 ps MD simulation with time-step 1.0 fs after heating the system from 0 K to 300 K within the first 10000 steps, increasing the temperature in portions of 30 degrees; the remaining 90000 steps are equilibration at 300 K. Additional 500.0 ps simulations were executed before starting the productive runs. The productive runs were performed with the recommended maximum time-step 2.0 fs when SHAKE is used, at 300 K and a constant pressure of 1.0 bar with isotropic position scaling using the Berendsen algorithm for temperature coupling. The simulation time was 20.0 ns. Samplings were taken every 2.0 ps. Additional simulations were executed for CD26 that differ only in step "*v)*" of the present equilibration protocol. The approach adopted in the previous work was followed,[52] and we designated these data throughout the text by CD26*.

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. 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.

In addition, the variations with the 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[59]/Surface Area (LCPO)[60]) 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 contribution to these variations. 1000 structures from the whole MD trajectories were utilized to estimate the MM/GBSA energies in each case. Information about structural variations during the simulations was obtained from analysis of rms deviation from the last structure.

For the first time we estimated also the entropy contributions normalized per glucose unit with the purpose to examine its dependence on the macroring size. The NMODE module of AMBER was used.[53] Sampled structures were extracted from the MD trajectory files (200 structures from each 20.0 ns simulation; the water molecules were removed). The structures were

energy-minimized prior to the normal-mode analysis using distance-dependent dielectric constant of $4r_{ij}$ (where r_{ij} is the distance between atom *i* and atom *j*).

Results and Discussion

Relative conformation of glucose units and nomenclature

All glucose units in small CDs (from CD5 to CD9) present a *syn* relative arrangement between them. However, the X-ray structures for the CD10 and CD14 already show some *anti* arrangements and also a new structural motif, the *kink* arrangement. In this motif, neighbor glucoses are in a *clinal* (or gauche) disposition (Figure 3).

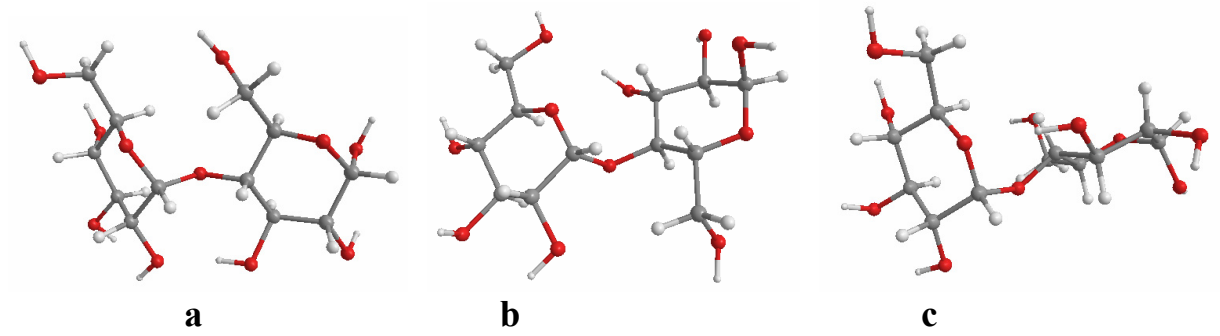


Figure 3. Relative conformations of neighbor glucoses: a) *syn*; b) *anti*; c) *kink*.

A nomenclature was proposed by Maestre et al.[43] for designating the conformation of the macroring in LR-CDs based on two nomenclatures used for cycloalkanes (using the torsional angles signs[61] and the nomenclature of Dale[62]). A character (*s*, *a*, + or -) is assigned to each relative position of a pair of glucoses depending on the value of a virtual dihedral angle (called *flip* and defined by O3(n)···C4(n)···C1(n+1)···O2(n+1), Figure 4). The ranges for the different portions were decided after the analyses of results from molecular dynamics simulations of several CDs, both in gas phase and in aqueous solution.[43] Thus, each inter-residue conformation can be assigned to *syn*, *anti* or *kink* form.

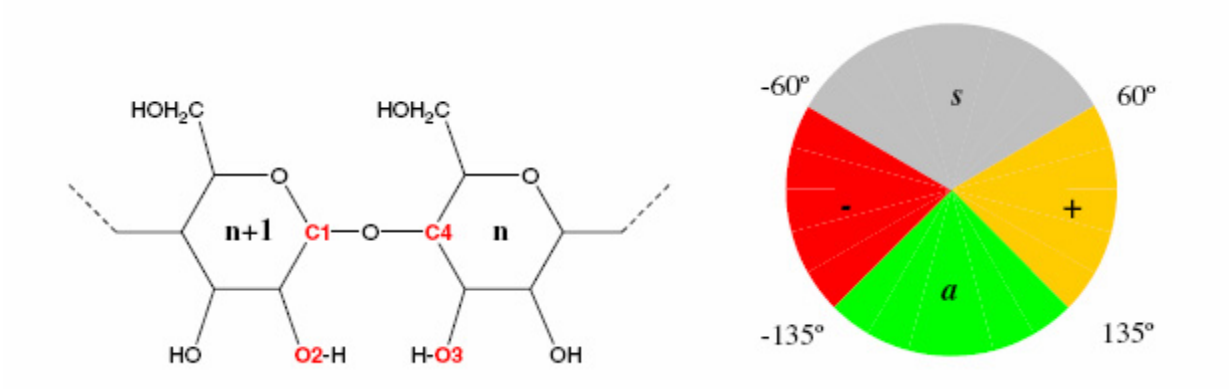


Figure 4. Definition of the relative orientation for a pair of glucose units as derived from the *flip* dihedral angle (O3(n)···C4(n)···C1(n+1)···O2(n+1)).

All glucoses are in principle equivalent by symmetry. Letter-code designations for the conformations of the macrorings were adopted that assume the name of the CD to start by that glucose residue having the largest number of *s* arrangement at the beginning of the name.[43] When two or more names fulfill this requirement, the one with the *a*, +, or – (in this order) as near as possible to the starting position, should be used. Since, this nomenclature produces long names (as many letters separated by comma as glucose units form the CD), whenever several identical characters are found, they are replaced by a number. As an example, the X-ray conformation of the CD10, (*s,s,s,s,a,s,s,s,s,a*), will be transformed into the (4*s,a*,4*s,a*). The names of the structures used for some of 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, CD24-2.0 represents a snapshot of the structure for CD24 after 2.0 ns MD simulation.

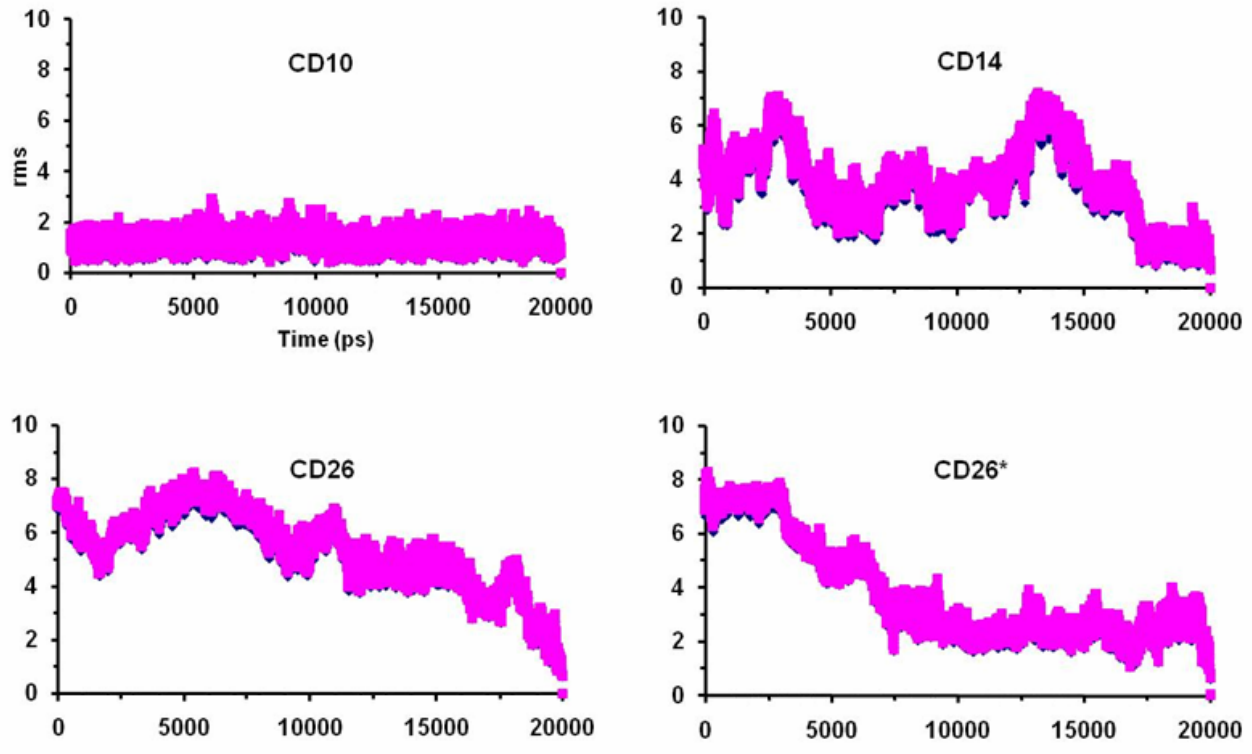


Figure 5. Rms deviations of atomic coordinates (in Å) relative to the final set of coordinates saved from the simulation. The C1, O4 and C4 atoms were used in the fit.

Structural variations of the large-ring cyclodextrins in solution

The analysis of the rms deviations from the last structure registered for each CD served to check and confirm the dynamism of the structural variations of the CDs. The outcome from the analysis indicates that, except for CD10, significant conformational changes indeed take place (Figure 5).

a **b** **c**

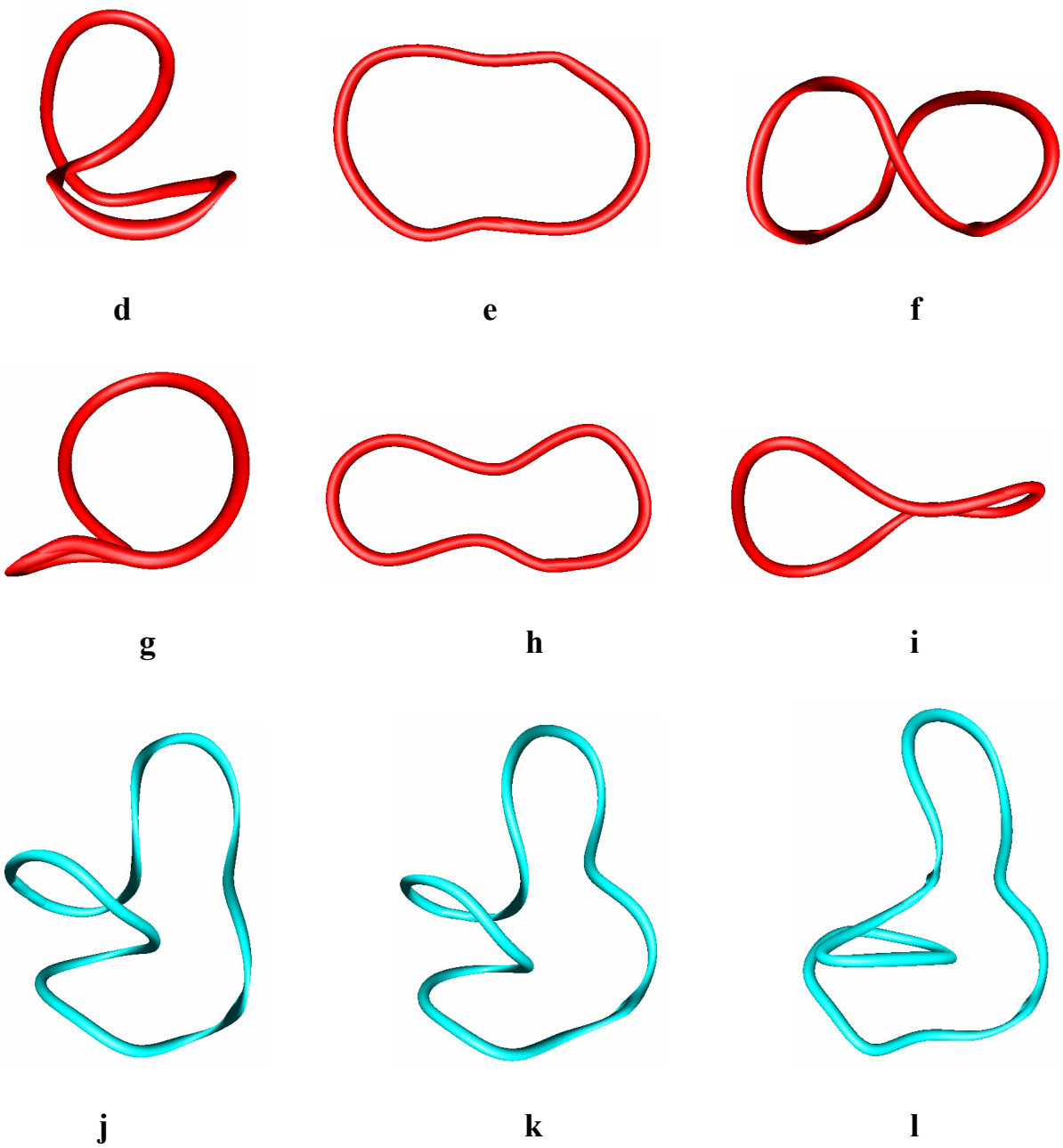


Figure 6. Typical appearances of some of the structural motifs: a) CD10-2.5, the most common form; b) CD10-5.5, the most opened form; c) CD10-19.0, the most closed form; d) CD14-1.5, two loops in perpendicular planes; e) CD14-5.0, the most open conformation; f) CD14-7.0, shaped as number eight; g) CD14-11.0, a big circular loop with a small helical turn; h) CD14-17.0, symmetrically squeezed open form ("a dumbbell"); i) CD14-20.0, two loops in perpendicular planes; j) CD26-3.5, a two turn helix with an extended helical portion; k) CD26-12.5, two turn helix with an arc and a loop of six glucose units; l) CD26-20.0, one turn helix with elongated loop and an arc.

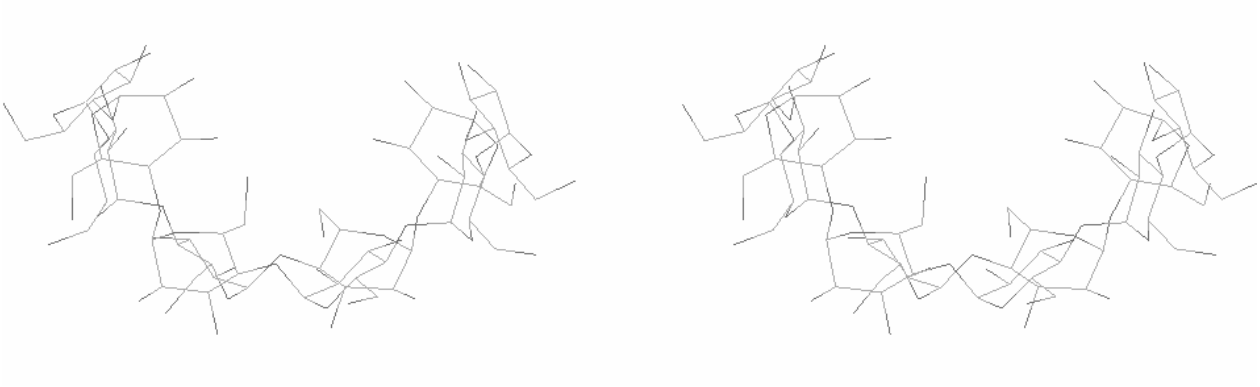
We examined, in stereo, 40 snapshots from each simulation (in intervals of 0.5 ns) in order to trace the tendencies in the deformations of the macrorings, focusing on the formation and disappearance of cavities that may eventually host small molecules, as well as characteristic folds and motifs of local structural deformations. Typical appearances of some of the structural motifs are presented in Figure 6. Figure 7 displays geometries from the last stages of the simulations. They are not representative or averaged geometries, but rather illustrate the variations of the macrorings conformations during the simulations, compared with the initial geometries (Figure 2).

For CD14 and CD26 different structural motifs were monitored (Figure 6), that may represent chiral species of particular interest for exploring supramolecular and chiral molecular recognition effects. Up to now, researchers in the field of cyclodextrins have been focusing their attention at the cavities formed by these macrorings, especially the native CDs. However, the results from our previous studies,[51,52] as well as the present one, with large variety of chiral structural shapes, open the possibility of using not the cavities but the whole macroring as a suitable ligand for some large chiral molecules with chirality originating from the symmetry group to which the molecule belongs, not from the existence of chiral atoms.

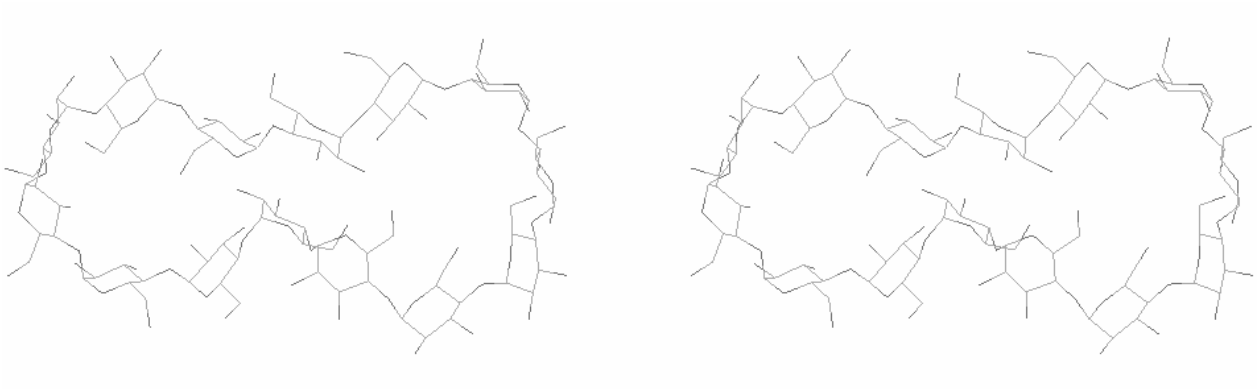
Due to the flipping of about 180º in *anti* orientation of two diametrically opposed glucoses along the perimeters of the macrorings, the molecular shapes of CD10 and CD14, determined from X-ray analyses, resemble the shape of a butterfly in which the wings are formed by cyclodextrin-like fragments and the conformational band-flips ("flips") are located at the body.[16] The macroring of CD10 underwent small deformations throughout the whole simulation and the overall shape of the molecule oscillated around the starting geometry. The flip angles at the two band-flips vary between 90º and 165º, whereas all other eight flip angles have negative values. The dihedral angles defined by the glycosidic oxygens (Table 1) are the measure for their coplanarity. The average value computed for CD10 is practically zero, with rms deviations of about 20º.

One of the "flips" in CD14 disappeared after only 1.0 ns simulation, whereas the second "flip" remained for about 15.0 ns. Thus the final geometry of CD14 is free of additional strain introduced by band-flips. Variety of macroring deformations were monitored during the simulation. The starting saddle-like geometry underwent twisting and two mutually perpendicular loops were formed after 1.5 ns, accompanied with the disappearance of one of the "flips". This conformation dominated for about 3.0 ns simulation. The overall shape resembles twisted number eight (Figure 6(d)). The first "flip" appeared again at 5.0 ns and survived until 10.0 ns simulation. The macro ring at 5.0 ns is opened and resembles the initial geometry of CD10 but in larger scale (Figure 6(e)). The macroring conformation passed again through the twisted-eight form and the two loops of figure eight are almost coplanar at 7.0 ns. The geometry oscillated about the latter two forms during the next 4.0-5.0 ns simulation. The macroring deformations produced a small spiral portion with a big circular loop at 11.0 ns (Figure 6(g)). The appearance of the shape of twisted figure eight dominates during the last 5.0 ns simulation, except for about 17.0 ns when the two loops are in the same plane and not crossed (Figure 6(h);

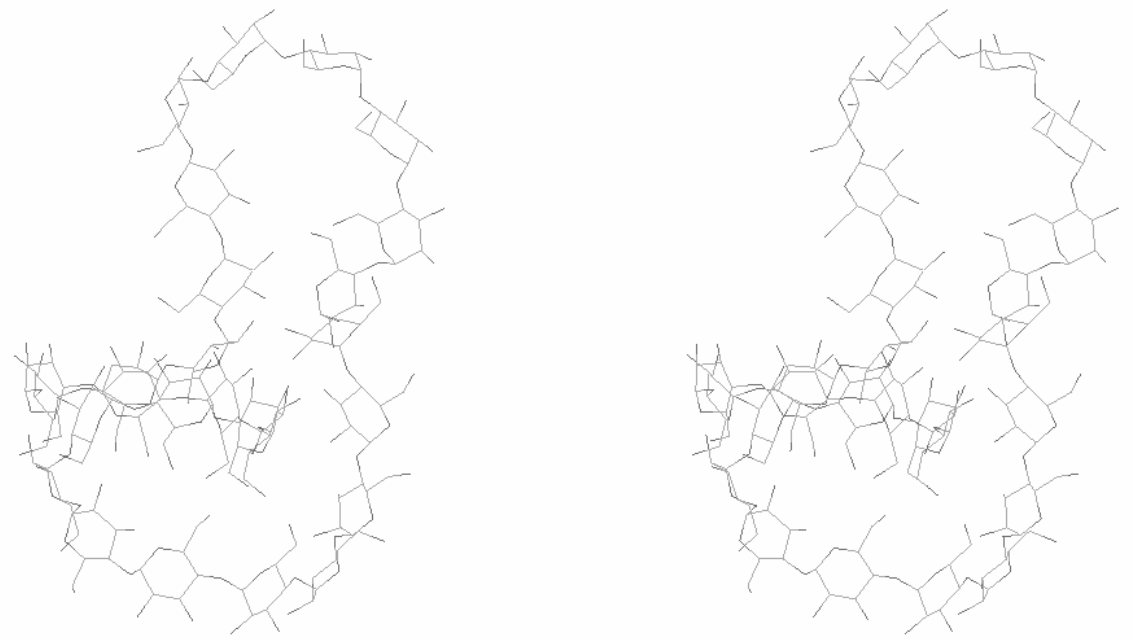
"dumbbell"). Thus, the overall analysis of the shape suggests preferred conformation of CD14 in aqueous solution in the form of deformed figure eight (this conformation is present about 60% of the simulation time).



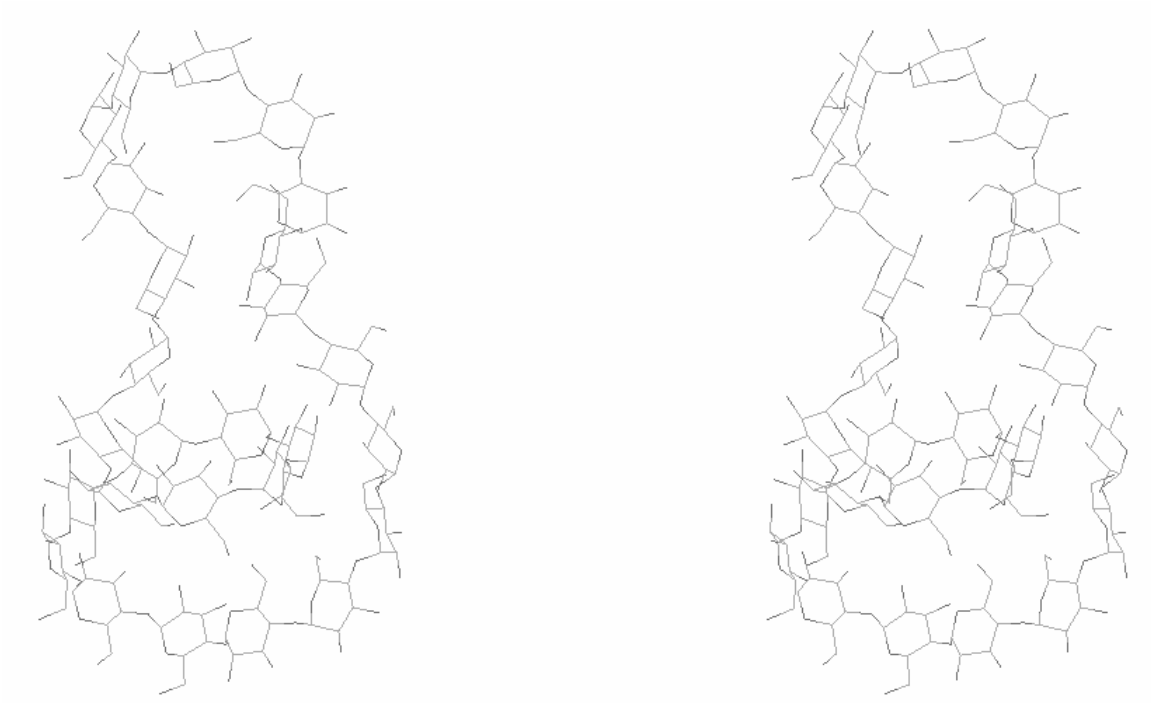
CD10-20.0 [3s,+,2-,2s,+,-]



CD14-20.0 [4s,3-,4s,-,s,-]



CD26-20.0 [12s,-,a,5s,2-,4s,-]



CD26*-20.0 [10s,a,-,5s,-,5s,3-]

Figure 7. Snapshots of geometries from the final stages (20.0 ns) of the simulations. These present an idea about the variations of the macrorings conformations during the simulations, starting from the initial geometries shown on Figure 2.

At the end of the equilibration step, CD26 acquired geometry that is still very close to the starting X-ray geometry. The crystallographic coordinates display 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]) with two stretches that connect the short helices from the "upper" and from the "lower" sides (Figure 2(c)). One "flip" is present in each bridging portion. One of the "flips" disappeared during the preliminary 500 ps simulation, whereas the other "flip" survived until the end of the simulation (except for short time intervals about 8.5 ns and 15.0 ns). The general appearances of conformations (Figure 6(j-l)) are analogous to the trends monitored in our previous simulations.[41,51] The most representative structural motifs during the simulation are a helical portion, one or two loops and an arc. The only difference in the results from the previous simulation protocol[41,51,52] (data designated with CD26*) is that the first "flip" did not disappear from the very beginning of the simulation, but survived during the first 3.0 ns.

Outline of the important structural parameters describing the conformations of the large-ring cyclodextrins

The following geometrical parameters are usually used for the structural analysis of large-ring cyclodextrins (Table 1):

O4(n)···O4(n-1): the distance between the glycosidic oxygen atoms;

O2(n)···O3(n-1): the distance between secondary hydroxyl groups of adjacent glucoses;

C1(n)-O4(n-1)-C4(n-1): the angle at the glycosidic oxygen connecting two glucoses;

O4(n)···O4(n-1)···O4(n-2): the angle formed by three neighbor glycosidic oxygen atoms;

O5(n)-C1(n)-O4(n-1)-C4(n-1): the dihedral angle ϕ;

C1(n)-O4(n-1)-C4(n-1)-C3(n-1): the dihedral angle ψ (ϕ and ψ describe the orientation of the residues about the $\alpha(1\rightarrow4)$ glycosidic linkage);

O4(n)···O4(n-1)···O4(n-2)···O4(n-3): the dihedral angle of four consecutive glycosidic oxygen atoms (a measure for the coplanarity of the macrorings);

O3(n)···C4(n)···C1(n+1)···O2(n+1): the dihedral "flip" between secondary hydroxyls of adjacent glucoses.

All computed structural parameters (Table 1) have values in the range of previously obtained data that were already discussed in detail and correlation was also made with experimental structural determinations.[41,42] In agreement with the crystallographic data, the distances O4(n)···O4(n-1) present almost the same values as for the helical parts of CD26, 4.5 Å. The distances O2(n)···O3(n-1) are rather small in the crystal, about 2.8-2.9 Å,[1,18] as a consequence of the formation of intramolecular inter-glucose hydrogen bonds.[1] The computed average O2(n)···O3(n-1) distances (3.4-3.9 Å) are somewhat longer than the experimental ones. This computed result is due partly to the loss of the alignment (partial flipping) between segments of the macrocycles such as a portion 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. In fact, distances of 5.5 Å between neighbor glucoses are described for CD10 and CD14.[17]

Table 1. Computed distances (Å), bond angles and dihedral angles (in degrees) of neighboring glucose units obtained from the MD simulations in water solution

	O4(n) …O4(n-1)				O2(n)…O3(n-1)			
	av.	rms	max.	min.	av.	rms	max.	min.
CD10	4.5	0.2	5.2	3.7	3.9	0.4	5.5	3.0
CD14	4.5	0.2	5.3	3.5	3.9	0.6	6.0	2.8
CD26	4.5	0.2	5.2	3.6	3.4	0.5	5.2	2.6
CD26*	4.5	0.2	5.2	3.7	3.4	0.5	5.2	2.6
	C1(n)—O4(n-1)—C4(n-1)				O4(n)…O4(n-1)…O4(n-2)			
	av.	rms	max.	min.	av.	rms	max.	Min.
CD10	116.9	3.4	130.8	104.2	135.3	7.3	162.8	99.5
CD14	116.6	3.5	130.6	104.1	135.8	9.9	167.7	93.8
CD26	116.4	3.4	130.0	103.9	135.1	8.1	162.6	99.4
CD26*	116.4	3.4	130.6	104.1	135.0	8.3	161.8	100.2
	O4(n)···O4(n-1)··· O4(n-2)···O4(n-3)				O3(n)···C4(n)··· C1(n+1)···O2(n+1)			
	av.		rms		av.		rms	
CD10	-0.2		18.4		-13.3		19.0	
CD14	-13.9		36.9		-26.4		30.8	
CD26	-18.6		22.2		-28.3		22.1	
CD26*	-20.1		22.4		-30.0		21.2	
	Φ				Ψ			
	O5(n)-C1(n)-O4(n-1)-C4(n-1)				C1(n)-O4(n-1)-C4(n-1)-C3(n-1)			
CD10	89.1		12.4		63.3		11.7	
CD14	95.2		17.6		93.5		21.0	
CD26	98.4		14.1		99.9		14.9	
CD26*	97.9		13.3		98.7		15.5	

Variations of 3.4° 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 determinations 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)…O4(n-1)…O4(n-2) increases from

119.9° in CD6 to 134.9° in CD8. Our computed data are in agreement with the experimental determinations. The average planarity of the macroring of CD10 (vide supra) is manifested in the almost zero value of the dihedral angle determined by the neighbor glycosidic oxygen atoms. The O3(n)···C4(n)···C1(n+1)···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)···C4(n)···C1(n+1)···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'. Except for CD10, the average computed values were determined in the range -26.0° to -30.0°.

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 α(1→4) glycosidic linkage. The glucose residues have the same orientations in the crystal phase if the two dihedrals acquire values in the range of 94.0° to 110.0°, and 97.0° to 135.0°, respectively.[16] 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).[16] Φ and Ψ have computed average values from 89.1° (CD10) to 98.4° (CD26), and from 63.3° (CD10) to 99.9° (CD26), respectively. The rms deviations are ca. 15° in most of the cases.

Analysis of hydrogen bonds distribution

Differences were obtained here in comparison with all earlier analyses of hydrogen bond distributions,[41,42,51,52] and these discrepancies resulted from the minor variations introduced in the equilibration stage of the simulation protocol (heating of the system from 0 K to 300 K within the first 10000 time-steps by increasing in ten steps the temperature in portions of 30 degrees, plus 90000 time-steps equilibration at 300 K). It appears that some properties, like intermolecular hydrogen bonds distributions are very sensitive to parameters of the simulation protocol. More reasonable numbers of intermolecular hydrogen bonds were estimated for the interactions of the water molecules as proton donors to the oxygen atoms of the CD (Figure 8). The population of the intramolecular OH…OH hydrogen bonds is preserved during the simulations, while very complicated and delicate dynamical distribution characterizes the intermolecular hydrogen bonding with the water molecules (Figure 9 and Table 2). Figure 9 displays nicely the disruption of the intramolecular C2(n)OH…OH(C(n-1)) hydrogen bonds of CD10 at the residues with "flips" which are present during the whole simulation.

Table 2. Number and population of hydrogen bonds obtained for each CD from the MD simulations in water solution (See Figure 8 for the meaning of hb1, hb2 and hb3).

		Total number of hydrogen bonds[a]	Number of hydrogen bonds with different population		
Population (%)			> 10.0	5.0 – 10.0	1.0 – 5.0
hb1	CD10	17237	32	31	98
	CD14	23927	53	34	191
	CD26	32860	131	90	437
	CD26*	693	126	66	148
hb2	CD10	43377	0	0	14
	CD14	60129	0	0	68
	CD26	84952	0	0	127
	CD26*	454	0	0	0
hb3	CD10	124	32	31	35
	CD14	245	53	34	72
	CD26	544	131	87	140
	CD26*	513	126	66	146

[a] This quantity is not identical with the population of hydrogen bonds presented in Figure 8, e.g. the number 32860 of hb1 for CD26 means that 32860 different hydrogen bond distances, shorter than the assumed O…O cut-off distance 3.5 Å, were registered during the whole simulation. Only 131 of them are with population higher than 10%, 90 hydrogen bonds have populations in the range of 5%-10%, etc.

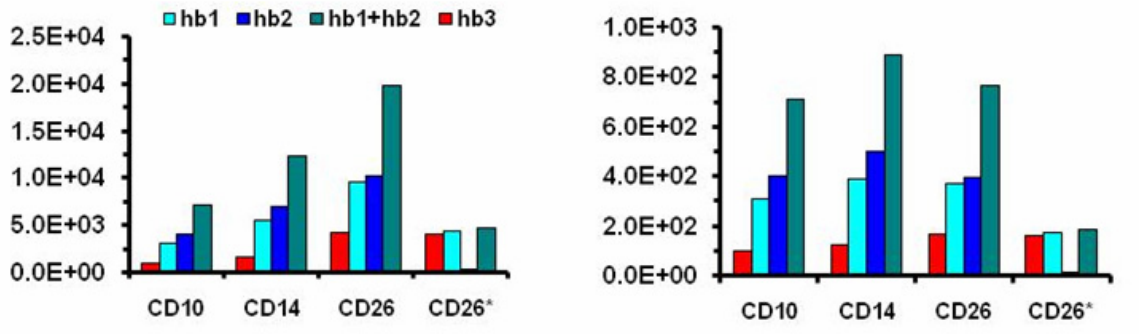


Figure 8. Graphical representation of the population (%) of hydrogen bonds obtained for each CD from the MD simulations in aqueous solution [hb1 – intramolecular + hydrogen bonds to oxygen atoms of water molecules (i.e., this quantity is a measure for the occupancy of the hydroxyl groups of the macromolecule to participate in hydrogen bonding as proton donors); hb2 – hydrogen bonds of water molecules as proton donors to oxygen atoms of the macro-ring; hb3 –

intramolecular hydrogen bonds]. (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.

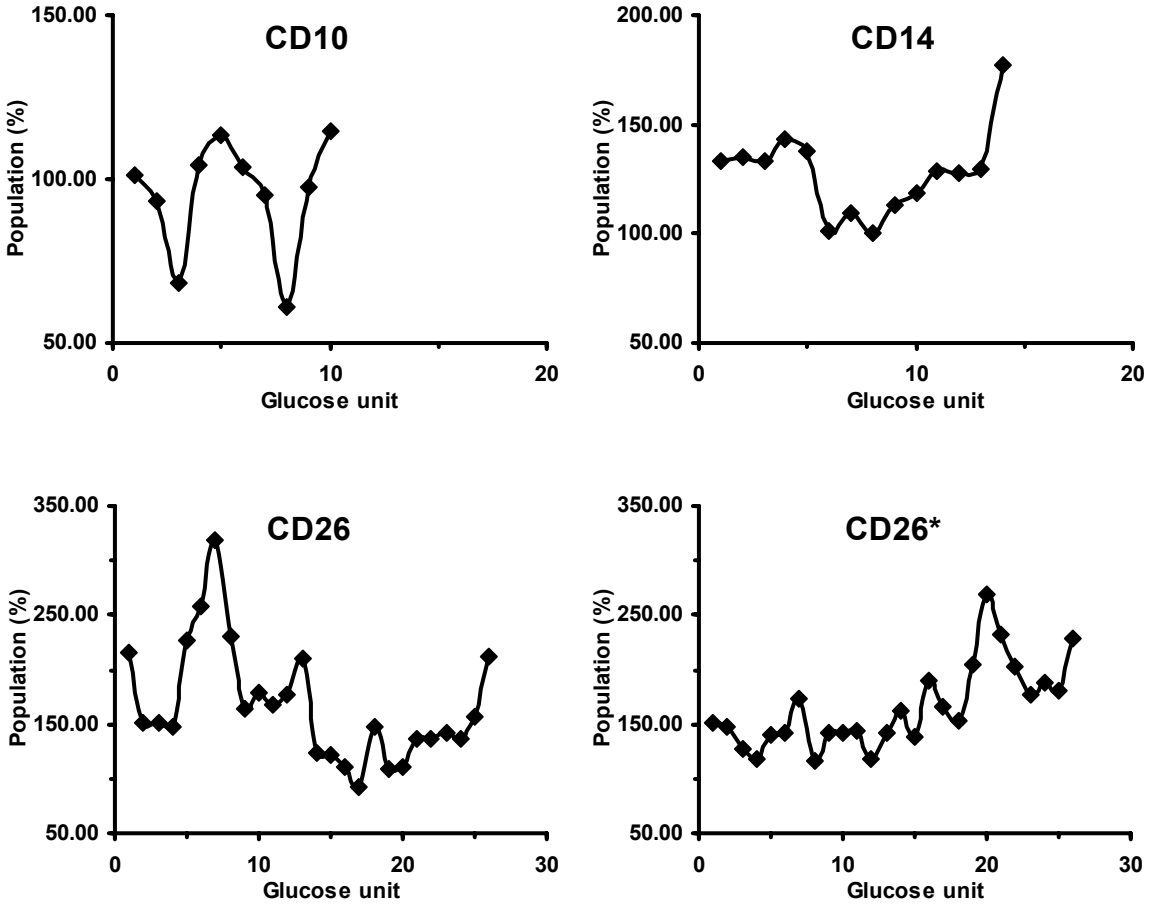


Figure 9. Time averaged populations of intramolecular hydrogen bonds estimated for each glucose residue. The OH groups participate as proton donors in hydrogen bonding.

Analysis of the shape

The moments of inertia were estimated with the CARNAL module of AMBER[50] and they correlate with the differences in size between the three CDs. The variation of the time average values of the O4 to the center of mass distances (O4…CM) evaluated for each residue reproduce correctly the variations with the time of the macroring conformations as determined from the analyses of snapshots extracted from the simulation trajectories, namely, small variations about one conformation for CD10, preferentially populated figure eight conformation for CD14, and more distant or closer positioning of some residues from the center of mass in CD26 (Figure10). Three more descriptors of the molecular size and anisometry have been also evaluated to assess

differences in size and shape between the CDs (Table 3):[63] radius of gyration (also provides an absolute measure of compactness), span (defined as the radius of the smallest sphere, centred at the centre 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[64] was used. Due to size limitations of the program, only the C1-O4-C4 fragment of each glucose unit was used when evaluating the descriptors. The values reported were obtained by averaging over all snapshots picked up from the trajectories. All computed values are in accord with our previous estimates.[41,43,51,52]

Table 3. Computed average values of radius of gyration (Å), span (Å), and asphericity

	CD10	CD14	CD26	CD26*
Radius of gyration	11.4	16.0	25.6[a]	26.6[a]
Span	8.5	12.0	15.9	17.2
Asphericity	0.2	0.3	0.2	0.3

[a] An experimental estimate 19.6 Å is cited in Ref. 2.

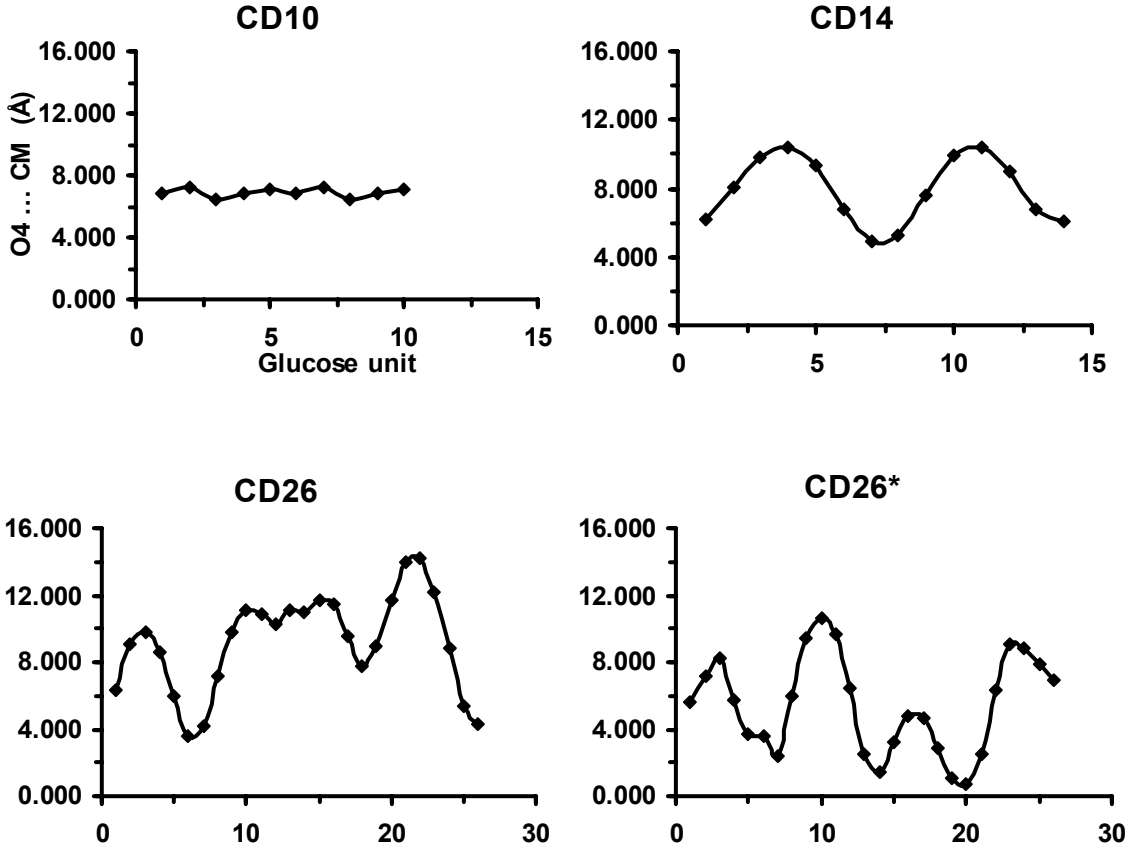


Figure 10. Variation of the time average values of the O4 to the center of mass distances (O4…CM) evaluated for each residue.

Energies of the large-ring cyclodextrins estimated with the MM/GBSA method

Exactly the same total energies per residue were obtained for CD26 ($E_{(total,GB)}$, Table 4) as were previously reported with the same force field for the CDs in the range of CD24 to CD30.[52] These quantities for the smaller CDs, CD10 and CD14, are about 1.0 kcal mol^{-1} larger. CD26 has more favorable van der Waals and electrostatic contributions, whereas better solvation lowers the energies of CD10 and CD14. The electrostatic and the internal energy terms contribute mostly to the resulting averaged per residue MM/GBSA energy values. In comparison with CD26, the total entropy contributions (TΔS) for CD10 and CD14, normalized per residue, are 1.6 kcal mol^{-1} and 1.1 kcal mol^{-1} more favorable, respectively. The vibrational entropy contributions per residue are the same for the three CDs. The total rotational entropy contributions correlate with the differences in the moments of inertia.

Table 4. Contributions per a monomer unit of different energy terms to the total MM/GBSA energies of the CDs and the entropy (in kcal/mol)[a]

	CD10		CD14		CD26		CD26*	
	av.	std	av.	std	av.	std	av.	std
$E_{electrostat}$	71.8	1.2	70.1	1.2	68.3	0.8	68.2	0.9
E_{vdW}	-0.8	0.5	-1.1	0.5	-2.4	0.4	-2.6	0.4
$E_{internal}$	23.3	1.1	23.6	1.0	23.3	0.7	23.3	0.7
E_{gas}	94.3	1.5	92.6	1.4	89.2	1.0	88.9	1.1
$E_{(nonpolar)}$	0.7	0.01	0.7	0.02	0.5	0.02	0.5	0.02
$E_{(GB)}$	-20.8	0.9	-19.1	0.9	-16.5	0.7	-16.5	0.7
$E_{(solvation)}$	-20.1	0.8	-18.5	0.9	-16.0	0.6	-15.9	0.7
$E_{(GB+elect)}$	51.0	0.6	51.0	0.6	51.7	0.4	51.7	0.4
$E_{(total,GB)}$	74.2	1.1	74.1	0.9	73.2	0.7	73.0	0.7
$TS_{(total)}$	15.5	1.1	15.0	0.8	13.9	0.5	13.9	0.9

[a] $E_{electrostat}$ – 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); $E_{(nonpolar)}$ – the nonpolar component of the solvation energy (surface area energy); $E_{(GB)}$ – the polarization component of the solvation energy in the generalized Born solvation model; $E_{(solvation)} = E_{(nonpolar)} + E_{(GB)}$; $E_{(GB+elect)} = E_{(GB)} + E_{electrostat}$; $E_{(total,GB)} = E_{gas} + E_{(solvation)}$.

Conclusions

The present long 20.0 ns molecular dynamics simulations on CD10, CD14, and CD26 in water, starting from X-ray geometries, gave small variations about the starting conformation for CD10 and preferentially populated deformed figure eight conformation for CD14. Different structural motifs were monitored for CD14 that may represent chiral species of particular interest for

exploring supramolecular and chiral molecular recognition effects: broadly opened macroring, that resembles the initial geometry of CD10 but in larger scale (CD14-5.0), a big circular loop with a small helical turn (CD14-11.0), symmetrically squeezed open form, "a dumbbell" (CD14-17.0). The final geometry of CD14 is free of additional strain introduced by band-flips. The preferred conformation of CD26 contains a small helix and an extended helical portion that transforms also to an arc and a loop. The general appearances of conformations of CD26 are analogous to the trends monitored in our previous simulations.[41,51] Differences were obtained in the present study, in comparison with all our earlier analyses of hydrogen bond distributions,[41,42,51,52] and these differences resulted from modification in the equilibration step of the simulation protocol.

Acknowledgements

Generous software support from Professor David Case and the AMBER team is acknowledged. The financial support is from the National Research Fund of Bulgaria (contract grant number X-1406).

References

1. Saenger, W.; Jacob, J.; Gessler, K.; Steiner, T.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T. *Chem. Rev.* **1998**, *98*, 1787.
2. Larsen, K. L. *J. Incl. Phenom. Mol. Recognit. Chem.* **2002**, *43*, 1.
3. Endo, T.; Ueda, H. *J. Pharm. Sci.* **2004**, *29*, 27.
4. Ueda, H.; Endo, T. In *Cyclodextrins and their complexes. Chemistry, analytical methods, applications;* Dodziuk, H. Eds.; Weinheim: Wiley-VCH, 2006; pp 370-380.
5. Ueda, H. *J. Incl. Phenom. Macrocyclic Chem.* **2002**, *44*, 53.
6. Motohama, S.; Ishii, E.; Endo, T.; Nagase, H.; Ueda, H.; Takaha, T.; Okada, S. *Biologicheskii Zhurnal Armenii* **2001**, *53*, 27.
7. Ueda, H.; Wakisaka, M.; Nagase, H.; Takaha, T.; Okada, S. *J. Incl. Phenom. Macrocyclic Chem.* **2002**, *44*, 403.
8. Taira, H.; Nagase, H.; Endo, T.; Ueda, H. *J. Incl. Phenom. Macrocyclic Chem.* **2006**, *56*, 23.
9. Zheng, M.; Endo, T.; Zimmermann, W. *J. Incl. Phenom. Macrocyclic Chem.* **2002**, *44*, 387.
10. Endo, T.; Zheng, M.; Zimmermann, W. *Aust. J. Chem.* **2002**, *55*, 39.
11. Qi, Q.; She, X.; Endo, T.; Zimmermann, W. *Tetrahedron* **2004**, *60*, 799.
12. Endo, T.; Ueda, H.; Kobayashi, S.; Nagai, T. *Carbohydr. Res.* **1995**, *269*, 369.
13. Ueda, H.; Endo, T.; Nagase, H.; Kobayashi, S.; Nagai, T. *J. Inclusion Phenom. Mol. Recognit. Chem.* **1996**, *25*, 17.

14. Larsen, K. L.; Mathiesen, F.; Zimmermann, W. *Carbohydr. Res.* **1997**, *298*, 59.
15. Fujiwara, T.; Tanaka, N.; Kobayashi, S. *Chem. Lett.* **1990**, 739.
16. Jacob, J.; Gessler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W. *Angew. Chem. Int.Ed.* **1998**, *37*, 606.
17. Jacob, J.; Gessler, K.; Hoffmann, D.; Sanbe, H.; Koizumi, K.; Smith, S. M.; Takaha, T.; Saenger, W. *Carbohydr. Res.* **1999**, *322*, 228.
18. Gessler, K.; Uson, I.; Takaha, T.; Krauss, N.; Smith, S. M.; Okada, S.; Sheldrick, G. M.; Saenger, W. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4246.
19. Nimz, O.; Gessler, K.; Uson, I.; Saenger, W. *Carbohydr. Res.* **2001**, *336*, 141.
20. Harata, K. In *Cyclodextrins and their complexes. Chemistry, analytical methods, applications;* Dodziuk, H. Ed.; Weinheim: Wiley-VCH, 2006; pp 147-198.
21. Bettinetti, G.; Sorrenti, M. *Thermochimica Acta.* **2002**, *385*, 63.
22. Terada, Y.; Yanase, M.; Takata, H.; Takaha, T.; Okada, S. *J. Biol. Chem.* **1997**, *272*, 15729.
23. Takaha, T.; Yanase, M.; Takata, H.; Okada, S.; Smith, S. M. *J. Biol. Chem.* **1996**, 271, 2902.
24. Dodziuk, H. In *Cyclodextrins and their complexes. Chemistry, analytical methods, applications;* Dodziuk, H. Ed.; Weinheim: Wiley-VCH, 2006; pp 1-30.
25. Asztemborska, M.; Bielejewska, A. In *Cyclodextrins and their complexes. Chemistry, analytical methods, applications;* Dodziuk, H. Ed.; Wiley-VCH: Weinheim, 2006; pp 106-118.
26. Chankvetadze, B. In *Cyclodextrins and their complexes. Chemistry, analytical methods, applications;* Dodziuk, H. Ed.; Weinheim: Wiley-VCH, 2006; pp 119-146.
27. Machida, S.; Ogawa, S.; Xiaohua, S.; Takaha, T.; Fujii, K.; Hayashi, K. *FEBS Letters* **2000**, *486*, 131.
28. Larsen, K. L.; Endo, T.; Ueda, H.; Zimmermann, W. *Carbohydr. Res.* **1998**, *309*, 153.
29. Kitamura, S.; Nakatani, K.; Takaha, T.; Okada, S. *Macromol. Rapid Commun.* **1999**, *20*, 612.
30. Furuishi, T.; Endo, T.; Nagase, H.; Ueda, H.; Nagai, T. *Chem. Pharm. Bull.* **1998**, *46*, 1658.
31. Dodziuk, H.; Ejchart, A.; Anczewski, W.; Ueda, H.; Krinichnaya, E.; Dolgonos, G.; Kutner, W. *Chem. Commun.* **2003**, 986.
32. Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1987**, *15*, 197.
33. Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *Eur. Biophys. J.* **1987**, *15*, 211.
34. Kitamura, S.; Isuda, H.; Shimada, J.; Takada, T.; Takaha, T.; Okada, S.; Mimura, M.; Kajiwara, K. *Carbohydr. Res.* **1997**, *304*, 303.
35. Lipkowitz, K. B. *Chem. Rev.* **1998**, *98*, 1829.
36. Dodziuk, H. In *Cyclodextrins and their complexes. Chemistry, analytical methods, applications;* Dodziuk, H. Ed.; Wiley-VCH: Weinheim, 2006; pp 333-355.
37. Shimada, J.; Handa, S.; Kaneko, H.; Takada, T. *Macromolecules* **1996**, *29*, 6408.
38. Shimada, J.; Kaneko, H.; Takada, T.; Kitamura, S.; Kajiwara, K. *J. Phys. Chem. B* **2000**, *104*, 2136.
39. Momany, F. A.; Willett, J. L. *Carbohydr. Res.* **2000**, *326*, 194.

40. Momany, F. A.; Willett, J. L. *Carbohydr. Res.* **2000**, *326*, 210.
41. Ivanov, P. M.; Jaime, C. *J. Phys. Chem. B* **2004**, *108*, 6261.
42. Ivanov, P. M.; Gotsev, M. G.; Jaime, C. *Bulg. Chem. Commun.* **2005**, *37*, 380.
43. Maestre, I.; Bea, I.; Ivanov, P. M.; Jaime, C. *Theoret. Chem. Acc.* **2007**, *117*, 85.
44. MM3* is a modified and adapted version of Allinger's MM3 force field (Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.)
45. Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.
46. 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).
47. Cornell, W. D.; Cieplack, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
48. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
49. Woods, R. J.; Dwek, R. A.; Edge, C. J.; Fraser-Reid, B. *J. Phys. Chem.* **1995**, *99*, 3832.
50. Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; 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.
51. Gotsev, M. G.; Ivanov, P. M.; Jaime C. *Chirality* **2007**, *19*, 203.
52. Gotsev, M. G.; Ivanov, P. M. *Int. J. Quantum Chem.* **2007**, *10*, 1657.
53. Case, D. A.; Darden, T. A.; Cheatham III, T. E.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Merz, K. M.; Wang, B.; Pearlman, D. A.; Crowley, M.; Brozell, S.; Tsui, V.; Gohlke, H.; Mongan, J.; Hornak, V.; Cui, G.; Beroza, P.; Schafmeister, C.; Caldwell, J. W.; Ross, W. S.; Kollman, P. A. AMBER 8. University of California: San Francisco 2004.
54. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
55. Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
56. Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
57. 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; p. 104.
58. Toukmaji, A.; Sagui, C.; Board, J.; Darden, T. *J. Chem. Phys.* **2000**, *113*, 10913.
59. Tsui, V.; Case, D. A. *Biopolymers* **2000**, *56*, 275.
60. Weiser, J.; Shenkin, P. S.; Still, W. C. *J. Comput. Chem.* **1999**, *20*, 217.

61. Burkert, U.; Allinger, N. L. In *Molecular Mechanics;* ACS Monograph 177 Washington ACS, 1982; pp 107–108.
62. Dale, J. *Acta. Chem. Scand.* **1973**, *27*, 1115.
63. Arteca, G. A. *Reviews in Comput. Chem.* **1996**, *9*, 191.
64. Todeschini, R.; Consonni, V.; Mauri, A.; Pavan, M. *http://www.disat.unimib.it/chm/*