

## Further studies on the conformations of large-ring cyclodextrins

P. Ivanov

<sup>1</sup>*Institute of Organic Chemistry with Centre of Phytochemistry, Bulgarian Academy of Sciences, Acad. G. Bonchev str., bl. 9, 1113 Sofia, Bulgaria*

Received May 07, 2014; Revised June 18, 2014

*Dedicated to Acad. Dimiter Ivanov on the occasion of his 120<sup>th</sup> birth anniversary*

The dependence of the results from the conformational search on the starting geometry of large-ring cyclodextrins (LR-CDs) was examined for the set of macrorings CD<sub>n</sub>, n = 16, 17, 18, 19, 20, 21. The structures obtained earlier for CD<sub>n</sub>, n=16, 17, 18, 19 were confirmed, but not those for CD<sub>20</sub> and CD<sub>21</sub>, although the tendency is evident for opening of the macrorings. Results from DFT studies on a model system, using different functionals, justify our choice to use the AMBER Glycam04 parameterization for modeling the conformations of large-ring cyclodextrins. The configurations of hydrogen bonds that determine the geometries of the preferred *syn* and *anti* conformations were also examined.

**Key words:** large-ring cyclodextrins, molecular dynamics, Glycam04 and Glycam06 AMBER parameterizations, DFT

### INTRODUCTION

As a natural consequence of the long-lasting interest and the numerous applications of the native cyclodextrins (small cyclodextrins;  $\alpha$ -CD,  $\beta$ -CD and  $\gamma$ -CD, respectively, cyclohexa-, cyclohepta-, and cyclooctaamylose), their larger analogues, the large-ring cyclodextrins (LR-CDs; CDs with a degree of polymerization (DP) higher than eight) [1-5], also attracted attention in recent years, and advances were marked in the study of their physicochemical properties [6,7], in spite of existing difficulties in their synthesis, isolation and purification. The existence of LR-CDs has been proven as products from the action of 4- $\alpha$ -glucanotransferases on amylose and amylopectin [2] almost a century later after the first evidences for the existence of cyclodextrins [8]. The expectations were that, in addition to the numerous well-known useful for the practice properties of the small cyclodextrins, namely to provide heterogeneous microenvironment to molecules, accommodated in their cavities, that acquire significantly modified properties compared to the free forms, some of the LR-CDs may have new, specific properties.

It took about twenty years after the first report for the existence of LR-CDs (DP = 9-14) [9] preparation method for LR-CDs mixtures to be worked out and  $\delta$ -CD (DP = 9) [10] to be isolated and its crystal structure to be characterized [11].

This opened the renewed interest on LR-CDs: within very short time the crystal determinations of CD<sub>10</sub> ( $\epsilon$ -CD) [12-14], CD<sub>14</sub> ( $\iota$ -CD) [12, 13], CD<sub>26</sub> ( $\nu$ -CD) [15,16] were reported. Important further steps related with the development of an effective purification method for LR-CDs [4] was the first chemical synthesis of a LR-CD, as well as the synthesis of the first chemically modified LR-CD (CD<sub>9</sub>). The experimental information about the conformations of LR-CDs is scarce. Crystal structure determinations were so far successful only for four of them. The difficulties in forming intramolecular and intermolecular hydrogen bonds, due to the high flexibility of their macrocyclic rings, as well as the low nucleation rates of sugars, that has to precede crystal growth [3], are considered to be the main cause for the low crystallinity of most LR-CDs.

An advantage of the LR-CDs in comparison with the small CDs, especially with  $\alpha$ -CD and  $\beta$ -CD, is that they are without any significant toxicity and that nutritionally they can be regarded as starch [2]. Commercially available CD-mixtures containing LR-CDs with a degree of polymerization from 9 to 21 were examined as additives to food (retrogradation retardant in breads, for freeze resistant jellies and for production of non-sticky rice) and drink products (high energy additive to soft drinks) for improving their texture, mouth feeling, flavor, taste, and palatability [6]. Others were studied as stabilization and solubilization for drugs. LR-CDs mixtures with a degree of polymerization from 22 to 45, and greater than 50 exhibited an efficient

\* To whom all correspondence should be sent:  
E-mail: ivanov@bas.bg

artificial chaperone effect for protein refolding (a protein refolding kit containing a mixture of LR-CDs as one of the active components is on the market [3] - the first practical application of LR-CDs in biotechnology). Such LR-CDs mixtures were able to strip detergent molecules of unfolded protein-detergent complexes, thus allowing the protein molecules to refold to their proper, folded, active state [2]. LR-CDs have been suggested for applications in the paper industry as an improved paper coating material and as starch substitutes in adhesives and biodegradable plastics.

LR-CDs may be good host molecules for relatively large guest compounds [6]. An increased flexibility of these molecules allows them either to present a more suitable cavity prior to complex formation or adapt to the guest molecules by an *induced fit* mechanism [2]. Since the large CDs are able to present a variety of cavity sizes, compared to the small CDs, they may be useful for special applications.  $\delta$ -CD, for example, has demonstrated to form a stable complex with C70 Buckminsterfullerene that allows its solubilization in water [17]. It has been proven that  $\eta$ -CD (12 glucoses) is effective in the partial separation of carbon nanotubes [18], i.e. not only fullerenes, but also single-wall carbon nanotubes were dissolved into water using LR-CD [3]. Increased interest is evident from the literature on developing techniques for isolation of LR-CDs [19,20] and for examining the properties of these macromolecules as new polysaccharide-based biomaterials [21].

In view of the difficulties with the experimental examination of the conformations of large-ring cyclodextrins, computational modeling and simulation methods provide useful tool to gain information about their conformational dynamics, the energetics, and the complex-forming ability. The conformations of some LR-CDs were systematically examined using molecular dynamics simulations as conformational search protocol [22,23-29]. Representative conformations of the macrocycles were obtained for different ranges of DP. There still exist however some problems that deserve further elucidation:

1. The range of DP from 16 to 21 was considered as appropriate test cases to examine alternative starting conformations for the conformational searches. This is the range of macroring sizes that differ significantly from the reference structures CD14 and CD26 for which experimental crystal structure determinations are available.

2. We used the AMBER (parm99) force field in our earlier studies [22]. The parameterization for carbohydrates, Glycam04 (AMBER v.8) [30] was

later adopted [23,24,28]. The most recent development of the Glycam series of parameterizations, Glycam06 (in AMBER v. 11) [31], is recommended to be used instead of Glycam04. It is natural now to adopt this new development. We tested however different methods on representative *syn* and *anti* conformations of  $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)- $\alpha$ -D-glucopyranoside (all they have six intramolecular hydrogen bonds). A *syn* conformation was estimated to have the lowest energy, whereas the following relative energies of the flipped (*anti*) conformation of the lowest energy were obtained (in kcal mol<sup>-1</sup>): DFT(6-311++G\*\*): 5.7 (B3LYP), 5.7 (LC-wPBE), 6.1 (CAM-B3LYP), 6.4 (wB97XD); SE-MO: 7.6 (AM1), 6.0 (PM3), 6.5 (PM6); UFF: 3.8; MMFF94: 5.7; Glycam04: 4.6; Glycam06: 2.9, i.e. Glycam06 estimated too low relative energy for the flipped conformation. How this result will reflect on the outcome from the Glycam06 modeling studies? Thus we consider important to justify our choice of using the Glycam04 parameterization for modeling the conformations of LR-CDs by comparing results for a model structure obtained with Glycam04 and DFT methods.

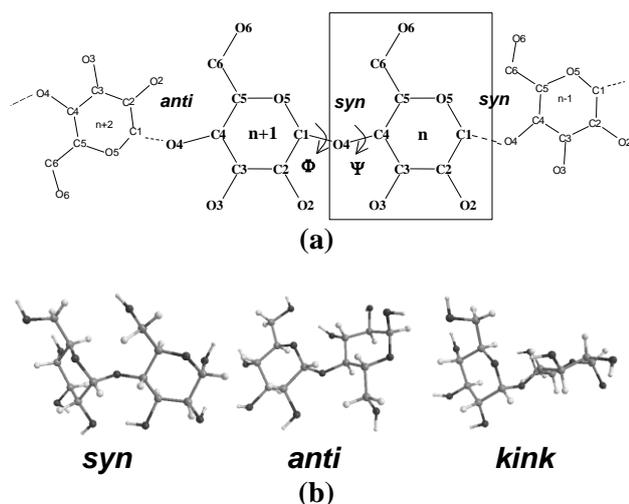
## COMPUTATIONAL DETAILS

The MD simulation trajectories were obtained with the AMBER program (version 11 [30]) using the parameterization for carbohydrates, Glycam04 [30]. The molecular dynamics simulations were run for aqueous solution (a box with TIP4P water molecules) using the particle mesh Ewald (PME) method for the treatment of the long-range electrostatics. A 9.0 Å distance cutoff was used for direct space nonbonded calculations and a  $1.0 \times 10^{-5}$  Ewald convergence tolerance for the inclusion of long-range electrostatic contributions. The 'solvateBox' command of LEaP was used to create cubic solvent box around the CD with buffer distances 15.0 Å between the walls of the box and the closest atoms of the solute. The SHAKE option (tolerance  $5.0 \times 10^{-5}$  Å) was used for constraining bonds involving hydrogen atoms. The starting geometries were derived following schemes based on the X-ray geometries of CD14 and CD26 as key reference structures and adding or removing residues. The same protocol from our previous studies was used for the preparation of the systems for the simulation [24,28,29]. Additional 500.0 ps simulation (NPT) was 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 an average

pressure 1.0 bar with isotropic position scaling. The simulation time was 100.0 ns. Samplings were taken every 2.0 ps.

## RESULTS AND DISCUSSION

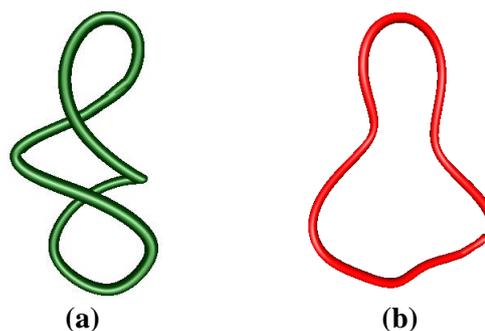
For the purpose of the representation we have to introduce some designations. Fig. 1 displays a schematic representation of a fragment of a cyclodextrin macroring with the numbering of the atoms and examples for *syn* and *anti* orientations of neighbouring glucose units.



**Fig. 1.** a) Schematic representation for 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. (b) Relative conformations of neighbor glucoses: *syn*, *anti* and *kink*.

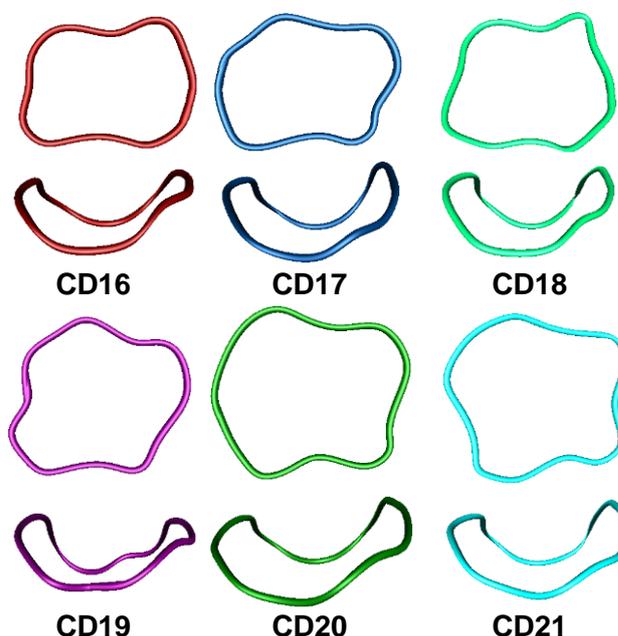
### Dependence of the results from the starting geometry

The three groups of LR-CDs examined by us so far [25-27] present structures from different ranges of values for the degree of polymerization. The starting geometries in two of these cases were derived from the experimental crystal structures of cyclodextrins which were closer in size, CD14 [26] and CD26 [25]. It should not be surprising that for the third group of cyclodextrins, CD<sub>n</sub> (n=18, 19, 20) [27], averaged geometries of CD20 were obtained that differ from the structures of the same macroring obtained earlier using different starting geometry (Fig. 2) [25]. This prompted us to examine in more detail the conformations of these LR-CDs using different starting geometries. The systems studied were CD<sub>n</sub>, n=16, 17, 18, 19, 20, 21. Indeed we found that such dependences exist.



**Fig. 2.** Averaged structures of CD20 derived after conformational searches with different starting geometries for the molecular dynamics simulations: a) Results from Ref. 27; b) Results from Ref. 25.

The scheme for generating the starting geometries differs considerably from the one we used before. The starting geometries in the earlier study were derived following a scheme based on the X-ray geometry of CD14 as a key reference structure and adding or removing residues: CD14→CD15→CD16→CD18→CD20→CD19. The leading idea in the present examination was to start from *flips*-free geometries (Fig. 3).



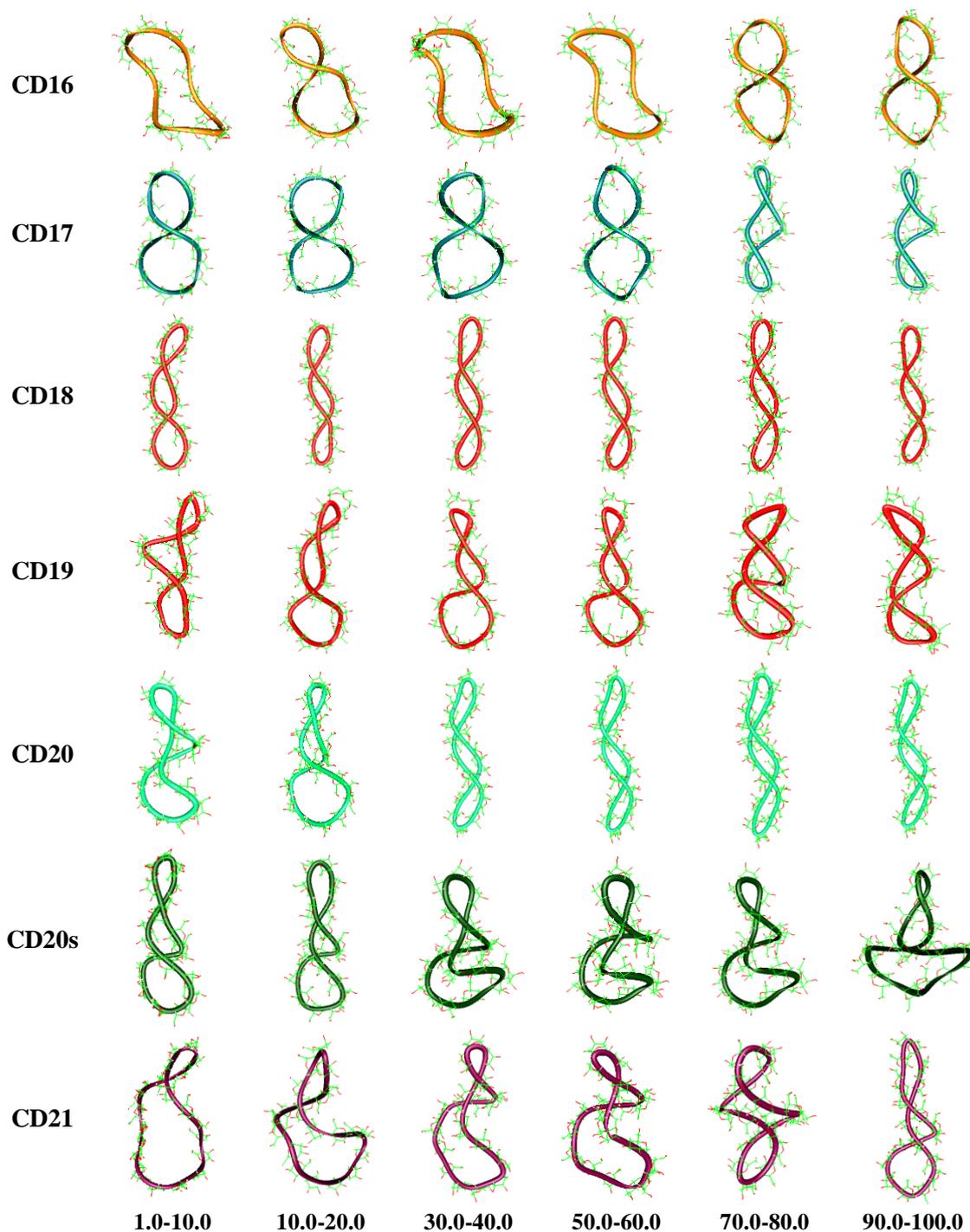
**Fig. 3.** Starting geometries for the conformational searches.

We prepared for the purpose starting geometry of CD16 using molecular graphics, whereas the other necessary structures were derived according to the sequence CD16→CD18→CD17→CD19→CD20, CD19→CD21. Fig. 4 contains average

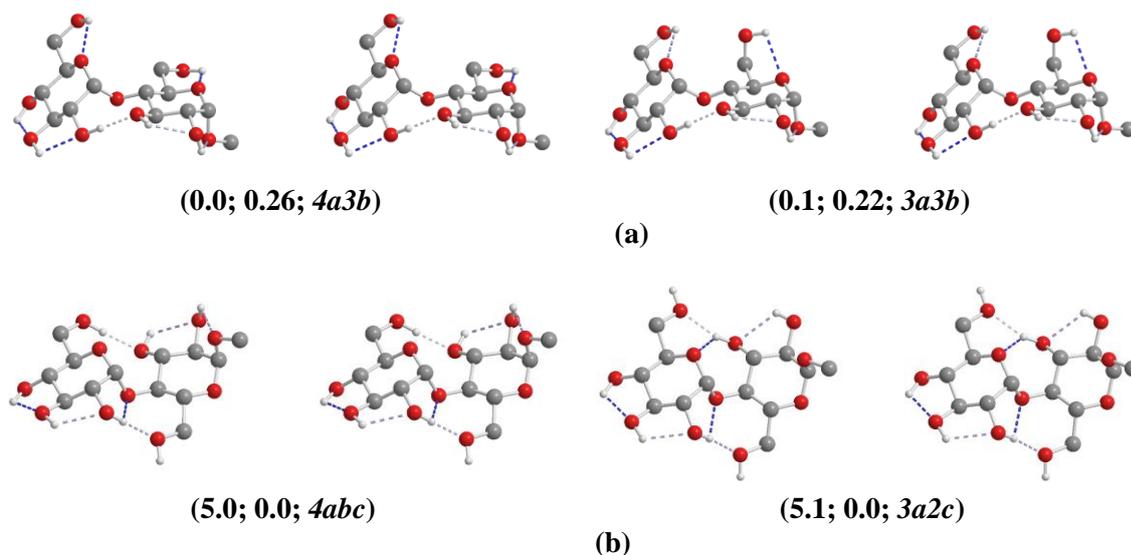
geometries for 10.0 ns simulation time intervals of the conformational searches. CD20 was used also to examine the dependence of the results from the value of the buffer parameter for the thickness of the water layer (Fig. 4). Using the same starting geometry, the value 15.0 Å for the solvent buffer

parameter was used in this study, whereas the results designated with CD20s were obtained with the smaller value 10.0 Å.

The structures obtained earlier for CDn, n=16, 17, 18, 19 were confirmed [26,27], but not those for



**Fig. 4.** Averaged geometries of CDn (n=16, 17, 18, 19, 20, 21) for different 10.0 ns intervals of the conformational searches.



**Fig. 5.** Stereo views of optimized geometries of the lowest-energy conformations of  $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)- $\alpha$ -D-glucopyranoside obtained with the MMFF94S force field: (a) *syn* conformations; (b) *anti* conformations. In parentheses are given the relative steric energies, the populations of the conformations and the letter-code describing the hydrogen bonding.

CD20 and CD21 [25], although the tendency is evident for opening of the macrorings. Longer duration of the simulations most probably may manifest also the geometries of CD20 and CD21 in Fig. 25 in Ref. 28. The observation of extended narrow antiparallel double helices for the cyclodextrins with DP in the range of 18 to 20 (Fig. 4), when using two different starting geometries, one of which has opened macroring, is additional evidence in support for the existence of such structures. Thicker water layer enhances the appearance of narrow extended geometries (Fig. 4, results for CD20). CD16 and CD17 display the figure-eight-like conformations with two  $\beta$ -CD pseudocavities. The cyclodextrins with even number of glucose residues, CD18 and CD20, present more symmetrical extended structures, probably due to more favorable torsional deformations about glycosidic bonds diametrically positioned along the perimeter of the macrocycle. Three segments, each of 6-7 glucoses, may produce more open cavities for CD21 (Fig. 4).

*Validation of the Glycam04 AMBER parameterization for conformational studies on large-ring cyclodextrins – a DFT study on a model system  $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)- $\alpha$ -D-glucopyranoside*

We will recall again the reasons hidden behind this examination. Tests with the Glycam06 parameterization (*vide supra*) gave too low relative

energy for the flipped conformation. A consequence of this result, when applying Glycam06 on the flexible LR-CD systems, could be unreliable modeling of local *syn/anti* transitions. In order to validate our preference for using the Glycam04 parameterization (adopted not only because of this reason), we examined with molecular mechanics (MMFF94S force field) the conformations of  $\alpha$ -D-glucopyranosyl-(1 $\rightarrow$ 4)- $\alpha$ -D-glucopyranoside. The total of 6697 conformations were obtained with relative energies less than 24.0 kcal mol<sup>-1</sup>. The first 130 lowest-energy conformations were further modeled also with four DFT functional [32]. The results for some of them are given in Table 1. We used two “standard” hybrid functionals not including dispersion (B3LYP, PBEh1PBE), one functional that has been constructed to account for dispersion (wB97XD) and the Zhao-Truhlar’s high-nonlocality functional M06 with double the amount of nonlocal exchange (2X) that has been recommended for noncovalent interactions. Fig. 5 presents MMFF94S optimized geometries of the lowest-energy *syn* and *anti* conformations. Table 1 contains also data for the dihedral angles of low-energy structures. For simplicity a letter-code was used to designate the values of the dihedral angles: *s* – dihedral angles in the range -30.0° – (+30.0°); *g*<sup>+</sup> – dihedral angles in the range 30.0° – 130.0°; *g*<sup>-</sup> – dihedral angles in the range -30.0° – (-130.0°); *a* – dihedral angles in the range 130.0° – (-130.0°). The pair of dihedrals ( $\Phi$ , $\Psi$ ) determines whether a particular conformation can be characterized as *syn*,

*kink* or *anti*. The four DFT methods estimate relative energies of ‘*band-flipped*’ conformations in accord with the Glycam04 results. The results justify our choice to use Glycam04 for modeling the conformations of large-ring cyclodextrins.

Let us examine the configurations of hydrogen bonds that determine the geometries of the preferred *syn* and *anti* conformations. For simplicity of the presentation we will introduce a letter-code for designation of the type of hydrogen bonding: *a* - a hydrogen bond between hydroxyl groups; *b* - a hydrogen bond between a hydroxyl group and a ring-oxygen atom or glycosidic oxygen; *c* - bifurcated hydrogen bond with proton acceptors hydroxyl oxygen and glycosidic or ring-oxygen atoms. Then, *4a2bc* will designate the case with four hydrogen bonds of type *a*, two hydrogen bonds of type *b* and one hydrogen bonding of type *c*.

The typical chain of four hydrogen bonds (*4a*) between the secondary hydroxyl groups determine the geometry of the lowest energy conformation optimized with molecular mechanics. The two primary hydroxyls form hydrogen bond with the ring-oxygen atoms. This is the typical case encountered in cyclodextrins. The second low-energy conformation has the same hydrogen bonds, the only difference being the orientation of the primary hydroxyl group with respect to the C5(n)-O5(n) bond. This hydrogen bonding is not present in the third conformation. The same type of hydrogen bonding of conformations one and two characterizes also conformations four (*kink* conformation) and five, but some dihedral angles have different values. Five hydrogen bonds of type *a*, constituting a ‘proton-relay’-like fragment, determine the conformation of the sixth *syn* conformation. One and

**Table 1.** Relative energies (kcal mol<sup>-1</sup>) of some low-energy conformations of α-D-glucopyranosyl-(1→4)-α-D-glucopyranoside determined with molecular mechanics and DFT methods.

No.	ΔE <sup>a</sup>					Dihedral angles <sup>d,e,f</sup>											
	B3LYP	wB97XD	PBEh1PBE	M062X	MMFF94S	Φ	Ψ	θ <sub>1</sub> <sup>1</sup>	θ <sub>1</sub> <sup>2</sup>	θ <sub>1</sub> <sup>3</sup>	θ <sub>1</sub> <sup>4</sup>	θ <sub>1</sub> <sup>2</sup>	θ <sub>2</sub> <sup>2</sup>	θ <sub>2</sub> <sup>3</sup>	θ <sub>2</sub> <sup>4</sup>	Φ <sub>OH</sub>	Φ <sub>OG+</sub>
1	1.6	1.8	1.7	2.6	0.0	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
2	0.0	0.0	0.0	0.0	0.1	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
3	1.1	0.5	0.7	0.6	0.1	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	a	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
4 <sup>b</sup>	0.0	0.0	0.0	0.0	0.8	a	a	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
5	1.6	1.6	1.7	2.7	1.0	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
6	1.7	1.7	1.7	2.8	1.1	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
7	2.0	2.3	2.2	2.9	1.6	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
8	0.2	0.6	0.2	0.8	1.7	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
9	1.1	0.5	0.9	1.1	1.7	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	a	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
10 <sup>b</sup>	1.4	1.8	1.5	2.6	1.8	a	a	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
11	1.7	1.4	1.5	1.8	1.9	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
12	0.0	0.6	0.0	0.9	1.9	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	a	g <sup>+</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
13	2.2	3.2	2.4	3.4	2.1	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	a	a	a	a	a	a	a	g <sup>+</sup>
14	1.1	0.7	0.8	0.7	2.1	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	a	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
15 <sup>b</sup>	0.2	0.6	0.2	0.8	2.1	a	a	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
16 <sup>b</sup>	0.0	0.6	0.0	0.9	2.3	a	a	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	a	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
17	3.0	3.4	3.2	4.2	2.4	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
18 <sup>b</sup>	2.2	2.5	2.4	2.6	2.4	a	a	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
19	2.2	2.5	2.4	2.6	2.5	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	a	a	a	a	a	a	a	g <sup>+</sup>
20	2.9	3.6	3.1	4.1	2.6	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	a	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
21 <sup>b</sup>	1.0	1.6	1.1	2.2	2.8	a	a	g <sup>+</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
71 <sup>c</sup>	<b>6.7</b>	<b>6.2</b>	<b>6.3</b>	<b>6.2</b>	<b>5.0</b>	g <sup>+</sup>	g <sup>-</sup>	a	a	g <sup>-</sup>	g <sup>-</sup>	g <sup>+</sup>	s	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>
72 <sup>c</sup>	<b>4.7</b>	<b>4.1</b>	<b>4.4</b>	<b>3.8</b>	<b>5.1</b>	g <sup>+</sup>	g <sup>-</sup>	a	a	a	a	g <sup>+</sup>	a	s	g <sup>-</sup>	g <sup>+</sup>	g <sup>+</sup>

<sup>a</sup> Total energies of the lowest-energy conformations:  $(E_o)_g^{B3LYP} = -1337.662286$  a.u.,  $(E_o)_g^{wB97XD} = -1337.269492$  a.u.,  $(E_o)_g^{PBEh1PBE} = -1336.333340$  a.u.,  $(E_o)_g^{M062X} = -1337.1721753$  a.u.,  $(E_o)_g^{MMFF94S} = 176.1$  kcal mol<sup>-1</sup>. <sup>b</sup> *kink*-conformations. <sup>c</sup> *band-flipped* conformations. <sup>d</sup> Definition of the dihedral angles according to the numbering of the atoms in Fig. 1: Φ = C4(n)-O4(n)-C1(n+1)-O5(n+1); Ψ = C3(n)-C4(n)-O4(n)-C1(n+1); θ<sub>1</sub><sup>1</sup> = O5(n)-C5(n)-C6(n)-O6(n); θ<sub>1</sub><sup>2</sup> = C5(n)-C6(n)-O6(n)-H; θ<sub>1</sub><sup>3</sup> = C1(n)-C2(n)-O2(n)-H; θ<sub>1</sub><sup>4</sup> = C2(n)-C3(n)-O3(n)-H; θ<sub>2</sub><sup>1</sup> = O5(n+1)-C5(n+1)-C6(n+1)-O6(n+1); θ<sub>2</sub><sup>2</sup> = C5(n+1)-C6(n+1)-O6(n+1)-H; θ<sub>2</sub><sup>3</sup> = C1(n+1)-C2(n+1)-O2(n+1)-H; θ<sub>2</sub><sup>4</sup> = C2(n+1)-C3(n+1)-O3(n+1)-H; φ<sub>OH</sub> = C3(n+1)-C4(n+1)-O4(n+1)-H; φ<sub>OG+</sub> = C4(n-1)-O4(n-1)-C1(n)-H. <sup>e</sup> Results with MMFF94S. <sup>f</sup> A letter-code: *s* - dihedral angles in the range -30.0° - (+30.0°); *g*<sup>+</sup> - dihedral angles in the range 30.0° - 130.0°; *g*<sup>-</sup> - dihedral angles in the range -30.0° - (-130.0°); *a* - dihedral angles in the range 130.0° - (-130.0°).

the same **4abc**-type *anti* conformation was found by all methods used. In addition to the two sets of three hydrogen bonds, a bifurcated hydrogen bond is also formed with the participation of the glycosidic oxygen atom. With the same energy, determined with MMFF94S, and with lower energy, determined by all DFT methods, is another *anti* conformation characterized by **4ac** or **3a2c** types of hydrogen bonding. The lowest energy *anti* conformation has one or two hydrogen bonds less than the corresponding *syn* counterpart. What is more interesting however is that the lowest energy *anti* form has one hydrogen bond less than the other *anti* conformation.

## CONCLUSION

Evidences were acquired on the validation of the modeling and simulation results on large-ring cyclodextrins from the starting geometry for the conformational search. The dependence of the results from the starting geometries was examined for the set of macrorings CD<sub>n</sub>, n= 16, 17, 18, 19, 20, 21. The structures obtained earlier for CD<sub>n</sub>, n=16, 17, 18, 19 were confirmed, but not those for CD20 and CD21, although the tendency is evident for opening of the macrorings. Results from DFT studies on a model system using different functionals (B3LYP, wB97XD, PBEh1PBE, M06-2X) justify our choice to use Glycam04 for modeling the conformations of large-ring cyclodextrins. The configurations of hydrogen bonds that determine the geometries of the preferred *syn* and *anti* conformations were also examined.

**Acknowledgements:** The financial support is from the National Science Fund within the project TK01/0167. CPU-hours grants from HPC-EUROPA2 (RII3-CT-2003-506079 – Research Infrastructure Action under the FP6 and the FP7 “Structuring the European Research Area” Programme) and RNF01/0110 (NSF-Bulgaria) projects, as well as stimulating discussions with partners in the COST action CM1002 are also acknowledged.

## REFERENCES

1. W. Saenger, J. Jacob, K. Gessler, T. Steiner, D. Hoffmann, H. Sanbe, K. Koizumi, S. Smith, T. Takaha, *Chem. Rev.*, **98**, 1787 (1998).
2. K. L. Larsen, *J. Incl. Phenom. Mol. Recognit. Chem.*, **43**, 1 (2002).
3. T. Endo, H. Ueda, *J. Pharm. Sci.*, **29**, 27 (2004).
4. H. Ueda, T. Endo, in: Cyclodextrins and their complexes. Chemistry, analytical methods, applications; H. Dodziuk, Ed.; Weinheim: Wiley-VCH, 2006; pp. 370-380.
5. T. Endo, *Trends Glycosci. Glycotechnol.*, **23**, 79 (2011).
6. H. Ueda, *J. Incl. Phenom. Macrocyclic Chem.*, **44**, 53 (2002).
7. K. L. Larsen, F. Mathiesen, W. Zimmermann, *Carbohydr. Res.*, **298**, 59 (1997).
8. M. A. Villier, *Comptes. Rendus. Acad. Sci.*, **112**, 536 (1891).
9. D. French, A. O. Pulley, J. A. Effenberger, M. A. Rougvie, M. Abdullah, *Arch. Biochem. Biophys.*, **111**, 153 (1965).
10. S. Kobayashi, M. Fukuda, M. Monma, T. Harumi, M. Kubo, Abstracts of Papers, the Annual Meeting of the Agricultural Chemical Society of Japan, Kyoto, April 1986, p. 649 (in Japanese).
11. T. Fujiwara, N. Tanaka, S. Kobayashi, *Chem. Lett.*, 739 (1990).
12. J. Jacob, K. Gessler, D. Hoffmann, H. Sanbe, K. Koizumi, S. M. Smith, T. Takaha, W. Saenger, *Angew. Chem. Int. Ed. Eng.*, **37**, 606 (1998).
13. J. Jacob, K. Gessler, D. Hoffmann, H. Sanbe, K. Koizumi, S. M. Smith, T. Takaha, W. Saenger, *Carbohydr. Res.*, **322**, 228 (1999).
14. K. Imamura, O. Nimz, J. Jacob, D. Myles, S. A. Mason, S. Kitamura, T. Aree, W. Saenger, *Acta Cryst. B*, **B57**, 833 (2001).
15. K. Gessler, I. Uson, T. Takaha, N. Krauss, S. M. Smith, S. Okada, G. M. Sheldrick, W. Saenger, *Proc. Natl. Acad. Sci. USA*, **96**, 4246 (1999).
16. O. Nimz, K. Gessler, I. Uson, S. Laettig, H. Welfle, G. M. Sheldrick, W. Saenger, *Carbohydr. Res.*, **338**, 977 (2003).
17. T. Furuishi, T. Endo, H. Nagase, H. Ueda, T. Nagai, *Chem. Pharm. Bull.*, **46**, 1658 (1998).
18. H. Dodziuk, A. Ejchart, W. Anczewski, H. Ueda, E. Krinichnaya, G. Dolgonos, W. Kutner, *Chem. Commun.*, 986 (2003).
19. F. Ellouze, N. Ben Amar, A. Deratani, W. Zimmermann, Results presented at the 15th International Cyclodextrin Symposium, O12, May 9-12, 2010, Vienna, Austria.
20. J. Wang, A. Jiao, Y. Tian, X. Xu, Z. Jin, Y. Li, *J. Chromatography A*, **1218**, 863 (2011).
21. S. Toita, N. Morimoto, K. Akiyoshim, *Biomacromolecules*, **11**, 397 (2010).
22. P. M. Ivanov, C. Jaime, *J. Phys. Chem. B*, **108**, 6261 (2004).
23. M. G. Gotsev, P. M. Ivanov, *J. Phys. Chem. B*, **113**, 5752 (2009).
24. P. M. Ivanov, *J. Phys. Chem. B*, **114**, 2650 (2010).
25. P. M. Ivanov, *Chirality*, **23**, 628 (2011).
26. P. M. Ivanov, *J. Mol. Struct.*, **1009**, 3 (2012).

27. P. M. Ivanov, in: *Nanoscience & Nanotechnology - Nanostructured materials - Application and Innovation Transfer*, E. Balabanova, E.; I. Dragieva, I. Eds., 2011, issue 11, 14.
28. P. Ivanov, *Curr. Phys. Chem.*, **2**, 413 (2012), and references cited therein.
29. P. Ivanov, E. Atanassov, C. Jaime, *J. Mol. Struct.*, **1056-1057**, 238 (2014).
30. D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossvary, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko, P. A. Kollman, (2010), AMBER 11, University of California, San Francisco.
31. K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.*, **29**, 622 (2007).
32. Gaussian 09, Revision A.1, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian, Inc., Wallingford CT, 2009.

## ИЗСЛЕДВАНЕ ВЪРХУ КОНФОРМАЦИИТЕ НА ГОЛЕМИ ЦИКЛОДЕКСТРИНИ

П. Иванов

*Институт по Органична химия с Център по Фитохимия, Българска Академия на Науките, ул. Акад. Г. Бончев, бл. 9, 1113 София, България*

Постъпила на 07 май 2014 г.; Коригирана на 18 юни 2014 г.

(Резюме)

Изследвана е зависимостта на резултатите от изходните геометрии при конформационно търсене за набор от големи циклодекстрини, CD<sub>n</sub>, n = 16, 17, 18, 19, 20, 21, Потвърдени бяха получени по-рано структури за CD<sub>n</sub>, n=16, 17, 18, 19, но не и такива за CD<sub>20</sub> и CD<sub>21</sub>, но се наблюдава тенденция към разтваряне на макропръстените, установено по-рано при тях. Резултати от DFT пресмятания с различни функционали на моделна система са в подкрепа на направения от нас избор на силово поле, Amber Glucam04, за моделиране на конформациите на големи циклодекстрини. Проучени са и конфигурациите от вътрешномолекулни водородни връзки, определящи геометриите на предпочетените *syn* и *anti* конформации.