

The β -cyclodextrin/benzene complex and its hydrogen bonds – a theoretical study using molecular dynamics, quantum mechanics and COSMO-RS

Jutta Erika Helga Köhler* and Nicole Grczelschak-Mick

Full Research Paper

Open Access

Address:
Wacker Chemie AG, Consortium für elektrochemische Industrie,
Zielstattstrasse 20, D-81379 München, Germany

Email:
Jutta Erika Helga Köhler* - jutta.koehler@wacker.com

* Corresponding author

Keywords:
AM1; benzene; COSMO-RS; cyclodextrin; hydrogen bonds; inclusion
complex; molecular dynamics; quantum mechanics

Beilstein J. Org. Chem. **2013**, *9*, 118–134.
doi:10.3762/bjoc.9.15

Received: 18 September 2012
Accepted: 12 December 2012
Published: 18 January 2013

This article is part of the Thematic Series "Superstructures with cyclodextrins: Chemistry and applications".

Guest Editor: H. Ritter

© 2013 Köhler and Grczelschak-Mick; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

Four highly ordered hydrogen-bonded models of β -cyclodextrin (β -CD) and its inclusion complex with benzene were investigated by three different theoretical methods: classical quantum mechanics (QM) on AM1 and on the BP/TZVP-DISP3 level of approximation, and thirdly by classical molecular dynamics simulations (MD) at different temperatures (120 K and 273 to 300 K). The hydrogen bonds at the larger O2/O3 rim of empty β -CDs prefer the right-hand orientation, e.g., O3-H \cdots O2-H in the same glucose unit and bifurcated towards \cdots O4 and O3 of the next glucose unit on the right side. On AM1 level the complex energy was -2.75 kcal mol $^{-1}$ when the benzene molecule was located parallel inside the β -CD cavity and -2.46 kcal mol $^{-1}$ when it was positioned vertically. The AM1 HOMO/LUMO gap of the empty β -CD with about 12 eV is lowered to about 10 eV in the complex, in agreement with data from the literature. AM1 IR spectra displayed a splitting of the O–H frequencies of cyclodextrin upon complex formation. At the BP/TZVP-DISP3 level the parallel and vertical positions from the starting structures converged to a structure where benzene assumes a more oblique position (-20.16 kcal mol $^{-1}$ and -20.22 kcal mol $^{-1}$, resp.) as was reported in the literature. The character of the COSMO-RS σ -surface of β -CD was much more hydrophobic on its O6 rim than on its O2/O3 side when all hydrogen bonds were arranged in a concerted mode.

This static QM picture of the β -CD/benzene complex at 0 K was extended by MD simulations. At 120 K benzene was mobile but always stayed inside the cavity of β -CD. The trajectories at 273, 280, 290 and 300 K certainly no longer displayed the highly ordered hydrogen bonds of β -CD and benzene occupied many different positions inside the cavity, before it left the β -CD finally at its O2/O3 side.

Introduction

Cyclodextrins (CD) are a family of conical shaped cyclic oligosaccharides consisting of 6–8 (or up to 10) glucopyranose units linked by α -(1→4) glycosidic bonds. At the narrower rim of the truncated cone (O6 side) there is one primary hydroxy group per glucose unit whereas at the wider rim there are two secondary hydroxy groups (O2/O3 side) [1]. Cyclodextrins have many possibilities of forming hydrogen bonds [2]: firstly, cyclodextrin monomers form some intramolecular hydrogen bonds or closed rings of hydrogen bonds at both rims of the cone; secondly, they form intermolecular hydrogen bonds with water molecules in aqueous solution or with included guest molecules that fit into their cavity. This fit is extremely specific: the orientation of one and the same guest molecule inside the cyclodextrin cavity may be opposite in solution and in crystalline state, as was found for 4-fluorophenol in α -CD [3]. The reactivity of aromatic guest molecules, radicals or excited states, was found to be altered because of complex formation with cyclodextrins [4].

Cyclodextrins form three types of dimers, O2/O3 to O2/O3, O2/O3 to O6, and O6 to O6. They can also associate to extended stacks in crystalline state or solution, with guest and solvent

molecules located inside the cavities, between dimers, or in the channels next to the stacks [5]. The influence of cyclodextrin on a broad variety of molecules was reported: from changes of α -/ β -sheet ratio upon complex formation of cyclodextrins with side chains of proteins [6], to size control of the electrostatic self assembly of nanoparticles [7]. Also a pH-controlled release of many cyclodextrins in long stacks on polymers, such as polyrotaxanes, combined with mesoporous silica particles was observed [8].

Complex formation with solvent molecules can take place at the same time in different positions: the crystal structure of heptakis(6-*O*-triisopropylsilyl)- β -CD benzene pyrene solvate is a nice example for competing solvents that intercalate in the cavity of one β -CD (benzene), between β -CD dimers (pyrene), and in channels outside β -CDs (benzene). The β -CD dimers are connected with hydrogen bonds on their O2/O3 sides (Figure 1).

Investigations of such “flat energy hypersurfaces” due to hydrogen bonds and noncovalent interactions demand several theoretical methods to capture the entire network of forces on

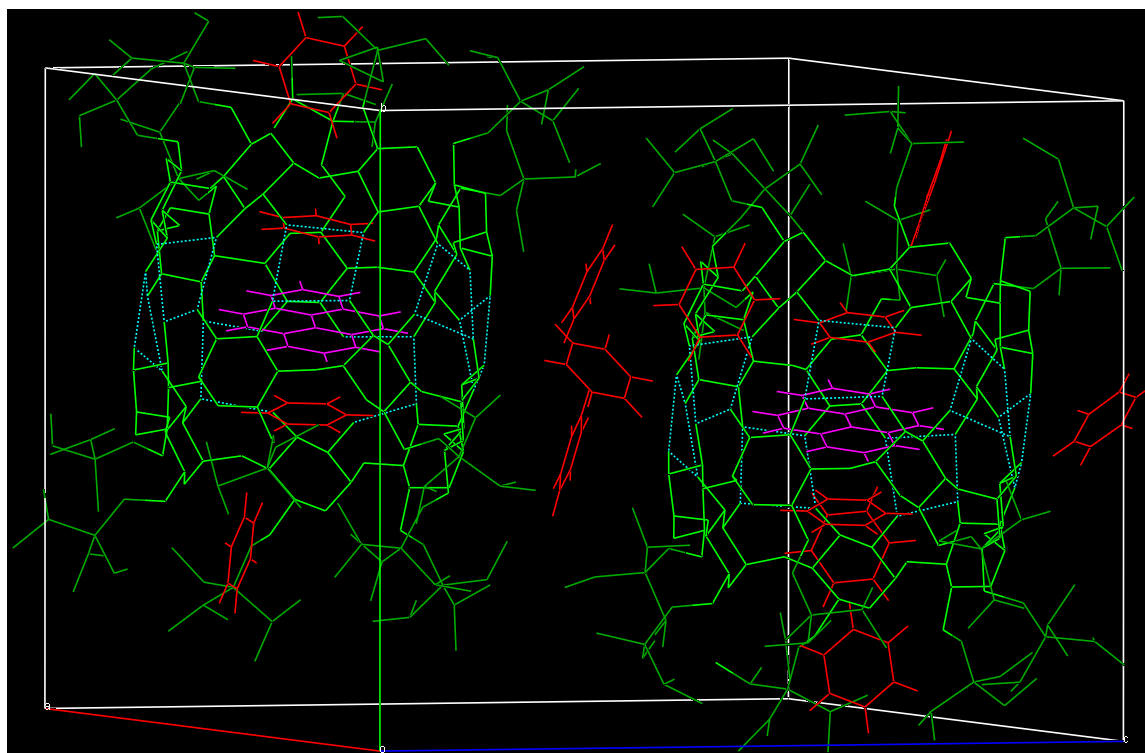


Figure 1: Crystal structure of heptakis(6-*O*-triisopropylsilyl)- β -cyclodextrin benzene pyrene solvate; $[\text{C}_{105}\text{H}_{210}\text{O}_{35}\text{Si}_{7,0,5}(\text{C}_{16}\text{H}_{10})_{3,5}(\text{C}_6\text{H}_6)]$, taken from The Cambridge Crystallographic Data Centre CCDC [9]. Colour code: benzene: red; pyrene: magenta; cyclodextrin corpus: light green; cyclodextrin side chains: dark green; hydrogen bonds: light blue, dotted lines.

which they rely. Molecular dynamics simulations displayed different hydrogen bond patterns of cyclodextrins in crystal and in solution and confirmed the experimental findings of soluble cyclodextrin complexes of cholesterol type versus precipitates of *cis/trans*-cyclohexadecenone//CDs [10–12]. MD with λ -dynamics and calculation of the solvation-free-energy differences was used to distinguish amylose helices from cellulose sheets by analysing the different reactivity of oxygen atoms O2, O3 and O6 of the sugar units with and without methylation, in line with experimental data [13]. Cyclodextrin-complex formation with substituted benzenes shows multiple topologies/configurations (guest up/down) in MD with λ -dynamics [14]. The association constant K_a for α - and β -CD inclusion complexes with several benzene derivatives was investigated by a genetic algorithm [15] and neuronal networks [16]. An overview of quantum mechanical methods to study cyclodextrin chemistry is given in [17]. The COSMO-RS solvation model [18] allows for calculations of several thermodynamic properties [19] once the polarisation charge surface of molecules has been determined by DFT calculations. With the current implementation of COSMO-RS, hydrogen-bond energies are predicted within 0.36 kcal mol⁻¹ relative to QM dimer calculations [20].

The necessity to understand the function of cyclodextrins in all details arises from their many applications in industry [21], agriculture, food [22] and health care [23]. Cyclodextrins and their complexes are produced in large industrial processes [24–26]. Our intention for this study was to find a “modular, robust molecular-modelling workflow” for how to analyse the formation of cyclodextrin complexes for many guest molecules in general, before extensive experimental work is carried out. Lemon grass oil [27] or retinol [23] can serve as typical molecular examples of the “practical industrial workbench”. To reach reasonable structures of these complexes we need to construct several models because of (a) isomers, (b) *n:m* guest/host stoichiometry, (c) preferred cyclodextrin cavity size, etc. We start here with the first step, the empty β -CDs, their subtle but strongly influential hydrogen bonds, and their complex formation with the simplest aromatic molecule of highest symmetry without substituents, i.e., benzene [28].

Methods

While calculating several cyclodextrin complexes with molecular dynamics [29], semi-empirical AM1, classical quantum mechanics [30] and COSMO-RS (Turbomole) [31], we observed that the orientation of all hydrogen bonds of the cyclodextrins already played a crucial role when the initial structures were constructed. It was not sufficient just to take the “normal route” of using the energetically lowest-lying structures from MD trajectories, and then proceed to AM1 and

further to QM optimizations; on the contrary, these structures turned out not to be the best ones possible. The best ones in QM finally were those that had hydrogen bonds oriented as symmetrically as possible. Therefore, we started with such manually constructed models and their AM1 optimisations.

Step 1: Our four β -CD models are conformers; all the hydrogen bonds of each rim are oriented in the same direction. We named them BCDO231O6l, BCDO23rO6l, BCDO231O6r and BCDO23rO6r, referring to their manually constructed hydrogen-bond orientations when looking either at their O2/O3 or O6 rim (l = left hand, r = right hand orientation) by using GaussView [30]. Each model was optimised until convergence was reached “AM1 (opt=calcall, verytight)” by using the Gaussian9W program [30]. Molecular properties, such as energy, dipole moment, HOMO/LUMO molecular orbitals and IR spectral data, were analysed with GaussView.

Step 2: The four AM1 optimised β -CD models were used as starting structures in our QM calculations by using Turbomole v. 6.4 with COSMO-RS [31]. Two models were calculated for each structure, i.e., the structure in vacuo (.energy files) and the COSMO-RS structure (.cosmo files), the latter being optimised in a dielectric field with the dielectric constant of water (the value of the dielectric constant is a variable in the COSMO-RS program and can be chosen by the user). Up to now it was possible to calculate molecular structures with COSMO-RS on DMOL, BP/SVP or BP/TZVP level of approximation. Our trial to optimise the β -CD complexes with BP/TZVP failed: not always but in several cases, here, for example, the benzene guest was expelled for the BCDbenzeneO231O6r conformers. Now, however, the current version of Turbomole allows the calculation of more detailed energy hypersurfaces especially for hydrogen bonds (keyword BP-TZVP-DISP3) since Grimme’s dispersion corrections for nonbonded interactions were implemented recently [32]. The current release of COSMO-RS C30_1201 allows, for the first time, a more detailed σ -surface to be calculated (BP/TZVPD-FINE), but for the time being as single points (SP) only. COSMO-RS structures are saved in COSMO databanks to perform graphical and thermodynamic analysis with COSMOthermX [19]. In this study we used the BP/TZVP-DISP3 method to calculate the β -CD models, and with this method no benzene molecule was expelled from the β -CD cavity any more.

Step 3: Starting from the neutron diffraction study of β -CD [33] we constructed one model named “BCDcry” of one β -CD by adding hydrogen atoms using the Visualizer of Materials Studio v. 5.5 [29]. Additionally, we used the four empty β -CDs of the in vacuo runs (BP/TZVP optimised .energy files) and the two best structures from the AM1 optimisations, BCDO23rO6l/

benzene parallel and vertical. With Materials Studio molecular dynamics (Forcite plus program and Compass force field) we simulated their trajectories at 300 K. In detail: After force-field geometry optimisations the structures were heated up each from 0 to 3 K, to 100 K, to 200 K, and to 300 K. Finally at 300 K a trajectory of 6000 ps length was calculated using an integration step of 1 fs. From the last 2000 ps of the trajectories the averaged total energy $E_{\text{tot}} <4000\text{--}6000>$, averaged hydrogen-bond energy $E_{\text{HB}} <4000\text{--}6000>$, and the averaged hydrogen-bond distance and angle were calculated.

Next, the energetically two best optimised AM1 structures (BCDO23rO6l with benzene included in parallel and vertical positions) were used as starting structures and simulated at 300, 290, 280, and 273 K for 500 ps each, and their release of the guest was analysed from these trajectories.

Results and Discussion

AM1 Calculations

The four conformers of empty β -CD

The BCDO23rO6l structure turned out to be the lowest energy conformer, also having the lowest dipole moment of only 0.5618 Debye (dipole-moment components $x = 0.0$; $y = 0.0$; $-z = 0.5618$) pointing from the O6 rim towards the O23 rim of the cone. This structure is so symmetric that all its Mulliken partial atomic charges on all types of oxygen atoms were identical, e.g., equal to each average number (Table 1). Second, with about 6 kcal mol⁻¹ above, was conformer BCDO23rO6r. This oval structure has the highest dipole moment of 2.5639 D ($x =$

$+0.0268$; $y = +0.0236$; $z = -2.5637$) and its O6 atomic charges display a range from -0.340 to -0.349 (average -0.344), but all O2 atoms still have the same charge of 0.325, and O3 and O4 oxygen charges vary very little (-0.337 to -0.338 and -0.300 to -0.301 , respectively). The third conformer BCDO23lO6l, lying about 7 kcal mol⁻¹ above, was again total symmetric as shown by its dipole moment components of $x = 0.0$; $y = 0.0$; $z = -1.0687$ and also with respect to all its oxygen charge values. The fourth conformer, BCDO23lO6r, had the highest energy of about 12.7 kcal mol⁻¹ compared to the ground state BCDO23rO6l and the second largest dipole value of 2.4223 D ($x = +0.5806$; $y = -0.0872$; $z = -2.3501$). Its O6 oxygen atoms display a broad range of Mulliken partial charges from -0.338 to -0.360 , although their average hardly differs from the average values of the other O6 types (see Table 1); O2 and O3 have small ranges (-0.333 to -0.335 and -0.345 to -0.347) but some differences can be observed in the O4 values, i.e., -0.288 to -0.296 . These subtle differences were just given as a few examples here in the text (e.g., they are not given in Table 1) since they may be of importance during guest inclusion or reactivity. The Mulliken partial atomic charges of the benzene molecule in D_{6h} symmetry were C: -0.130 and H: $+0.130$.

The four conformers of the β -CD complexes with benzene in a parallel or vertical position

All four conformers of the complexes preferred the inclusion of benzene in a parallel position (Table 2). The energetic order of the benzene complexes and their sequence of dipole moments was qualitatively the same as for the empty ones: BCDO23rO6l,

Table 1: AM1 results for the four empty β -CD conformers. AM1 optimised energies [a.u.], relative energies [kcal mol⁻¹], dipole moments [D], LUMO/HOMO energies [a.u.] and gap [eV], and averaged Mulliken partial atomic charges of oxygen atoms. Benzene molecule with D_{6h} symmetry in vacuo. ΔE is the energy of each conformer with respect to the lowest conformer in this table, which is O23rO6l (bold).

AM1 β -CD empty ^a	abs. E [a.u.]	ΔE [kcal mol ⁻¹]	μ [D]	LUMO/HOMO [a.u.]; gap [eV]	O2 atomic charge	O3 atomic charge	O4 atomic charge	O6 atomic charge
O23lO6l				+0.06271 -0.38469 12.17	-0.335	-0.347	-0.291	-0.348
O23lO6r	-2.653658	7.06	1.0687	+0.06083 -0.38424 12.11	-0.334	-0.346	-0.293	-0.346
O23rO6l				+0.06072 -0.38650 12.17	-0.325	-0.337	-0.297	-0.348
O23rO6r	-2.644708	12.68	2.4223	+0.05835 -0.38786 12.14	-0.325	-0.337	-0.301	-0.344
benzene ^b	-2.655373	5.99	2.5639	+0.02039 -0.35474 10.21				
	+0.034953		0.0000					

^aAll four β -CD conformers reached a RMS gradient norm of 0.00000001 [a.u.] in the geometry optimisations.

^bBenzene reached a RMS gradient norm of 0.00000002 [a.u.].

Table 2: AM1 results for the four β -CD conformers with benzene placed in a parallel or vertical position. AM1 absolute energies, relative energies and dipole moments for the complexes: columns 1–3; for benzene alone: columns 4 and 5.

AM1 β -CD/benzene complex ^a	complex abs. E [a.u.]	complex ΔE [kcal mol ⁻¹]	complex μ [D]	benzene abs. E [a.u.]	benzene μ [D]
O23IO6l					
parallel	-2.62313280	7.0374744	0.9544	0.03501077	0.0446
vertical	-2.62266055	7.3338160	0.7084	0.03499769	0.0092
O23IO6r					
parallel	-2.61493920	12.1790404	2.9750	0.03499519	0.0174
vertical	-2.61452474	12.4391182	1.5541	0.03500111	0.0104
O23rO6l					
parallel	-2.63434772	0.00000000	0.4399	0.03502037	0.0543
vertical	-2.63388503	0.29034260	0.1579	0.03499812	0.0096
O23rO6r					
parallel	-2.62574636	5.39743941	2.5365	0.03499903	0.0256
vertical	-2.62459809	6.11799032	1.9127	0.03499288	0.0138

^aAll four complexes reached a RMS gradient norm $\leq 3 \times 10^{-8}$ [a.u.] in the geometry optimisations.

BCDO23rO6r, BCDO23IO6l, and BCDO23IO6r. Benzene was slightly distorted in each cavity, as can be seen from its energy and dipole moment inside the complex. Inside the β -CD/benzene complex the benzene neither in a parallel position nor in vertical position adopted an exact central position as displayed in Figure 2.

Our calculated AM1 HOMO/LUMO gaps of empty β -CD conformers, of benzene D_{6h} (Table 1) and the eight complexes (Table 3) are of the same order of magnitude as the values calculated with PM3 in the literature [34] where they found for β -CD 12.44 eV, for benzene 10.15 eV and for the β -CD/benzene complex 10.15 eV. These authors found β -CD/benzene

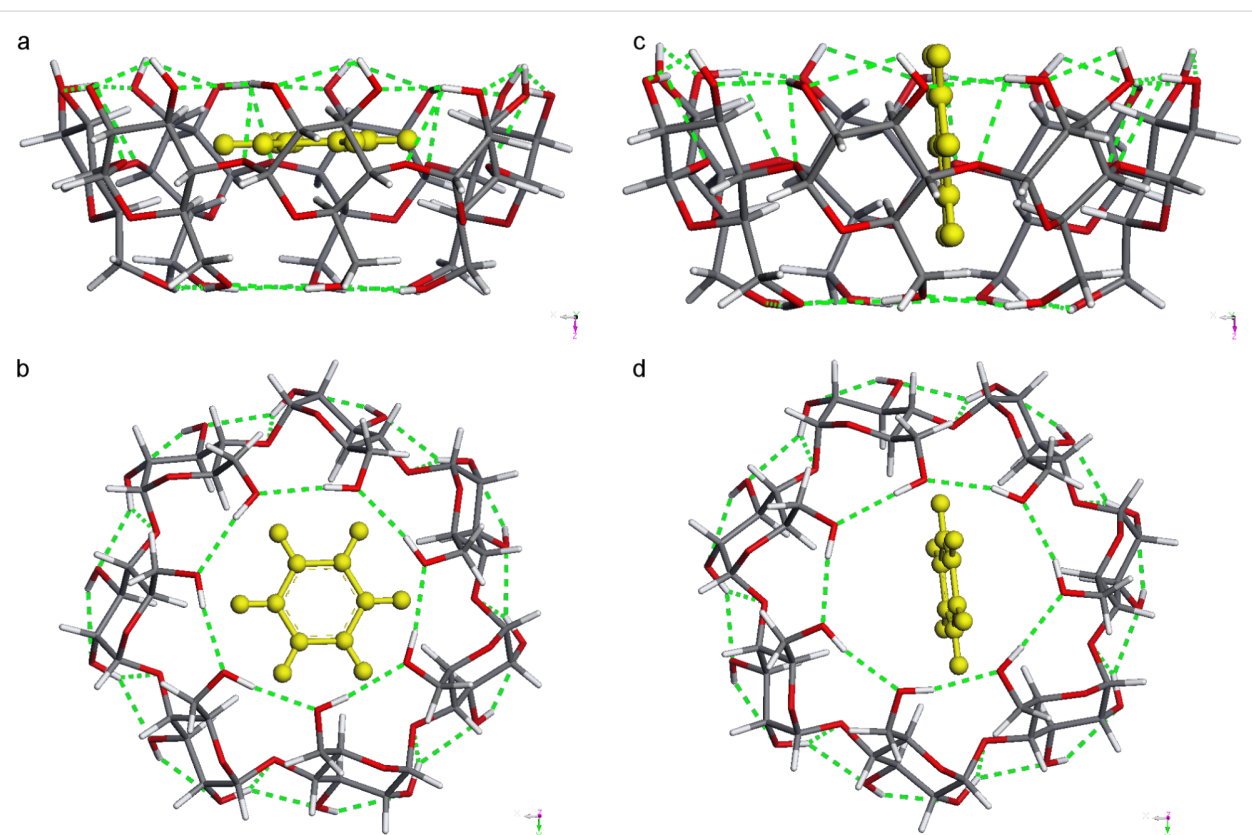
**Figure 2:** The two AM1-optimised stable conformers of BCDO23rO6l with benzene in a parallel (a and b) and vertical position (c and d). Top view from the O6 side.

Table 3: AM1 results for the four β -CD/benzene complex conformers with benzene placed in a parallel or vertical position. LUMO/HOMO energies and band gaps: columns 1–3. Absolute energies of complex formation $\Delta E_{\text{complexation}}$: columns 4,5. Data of column 5 are given in kilojoules per mole (kJ mol^{-1}) for easy comparison to the text above and [34].

AM1 β -CD/benzene complex	HOMO [a.u.]	LUMO [a.u.]	HOMO/LUMO gap [eV]	$\Delta E_{\text{complexation}}$ [kcal mol ⁻¹]	$\Delta E_{\text{complexation}}$ [kJ mol ⁻¹]
O23IO6l parallel	-0.34409	0.03063	10.196	4.29	17.93
vertical	-0.34163	0.03241	10.178	4.58	19.17
O23IO6r parallel	-0.35373	0.02092	10.195	9.43	39.44
vertical	-0.34592	0.02861	10.192	9.69	40.53
O23rO6l parallel	-0.34784	0.02688	10.197	-2.75	-11.52
vertical	-0.34530	0.02866	10.176	-2.46	-10.30
O23rO6r parallel	-0.35776	0.01689	10.195	2.65	11.07
vertical	-0.35353	0.02046	10.177	3.37	14.08

of 1:1 inclusion type in aqueous solution by experimental UV absorption spectra with various concentrations of β -CD. From their UV plots they calculated their stability constants K_s at three temperatures to be 116 L mol^{-1} at 291 K, 86 L mol^{-1} at 298 K and 68 L mol^{-1} at 303 K. According to the equation $\Delta G = -RT \ln K_s$ they obtained the corresponding three free energy values ΔG of $-11.5 \text{ kJ mol}^{-1}$, $-11.0 \text{ kJ mol}^{-1}$ and $-10.6 \text{ kJ mol}^{-1}$, indicating rather weak interactions between the guest and host molecule, but still a thermodynamically exothermic reaction of complex formation.

Here, our best conformer, BCDbenzeneparallelO23rO6l, displays an energy of complex formation

$$\Delta E_{\text{complexation}} = E_{\text{complex}} - E_{\text{BCDempty}} - E_{\text{benzD}_{6h}} \quad (1)$$

of $-2.75 \text{ kcal mol}^{-1}$ (Table 3), followed by the complex conformer BCDbenzeneverticalO23rO6l with $-2.46 \text{ kcal mol}^{-1}$. This is in very good agreement with the experimental value from [34]. All other complex conformers have positive $\Delta E_{\text{complexation}}$ values, up to BCDbenzeneverticalO23IO6r with $+9.69 \text{ kcal mol}^{-1}$, and thus are endothermic, Table 3.

IR spectra of empty β -CD conformers and their inclusion complexes

The most important lines of the calculated AM1 IR spectra are given in Table 4. The IR bands of benzene at 744, 1145, 1579 and 3194 cm^{-1} were characteristically shifted in the parallel versus the vertical complex. Especially the 3194 cm^{-1} H–C benzene stretch was split into a range from 3150 to 3199 cm^{-1} . Very interesting and easy to see was the concerted movement of all hydrogen bonds at the O23 and O6 rim: the empty β -CD's

Table 4: AM1 results for benzene and the two stable O23rO6l β -CD conformers with benzene placed in a parallel or vertical position. All IR bands [cm^{-1}] and relative intensities are listed for benzene; in the case of β -CD, single bands are included only if their relative intensity is >200 or for "very close frequency clusters" (bold)^a.

AM1 IR freq. [cm^{-1}]	rel. intensity benzene D_{6h} sym.	rel. intensity β -CD O23rO6l empty	rel. Intensity β -CD O23rO6l parallel	rel. Intensity β -CD O23rO6l vertical	comment
443					H(O6) up/d
444		729		713	H(O6) up/d
445			741		H(O6) up/d
511				276	H(O2) up/d
513		2 × 346	286		H(O2) up/d
514				240	H(O2) up/d
516			335	230	H(O2) up/d

Table 4: AM1 results for benzene and the two stable O23rO6l β -CD conformers with benzene placed in a parallel or vertical position. All IR bands [cm^{-1}] and relative intensities are listed for benzene; in the case of β -CD, single bands are included only if their relative intensity is >200 or for "very close frequency clusters" (bold)^a. (continued)

744	127				H–C up/d
757				122	H–C up/d
760			73		H–C up/d
1145	2 x 1.08				H–C in plane
1148				1.32	H–C in plane
1150			1.74		H–C in plane
1150			1.51		H–C in plane
1152				0.54	H–C in plane
1346				344	O4 and O5
1347		471	491		O4 and O5
1347		471	355		O4 and O5
1348				291	O4 and O5
1351		2 x 285			C-O6 frame
1356–1357			534(3)^a		C-frame
1357		2 x 312			C-frame
1357–1358				562(3)^a	C-frame
1418		346		338	H–C up/d
1419			339		H–C up/d
1577				14	C=C
1578			15		C=C
1579	2 x 12		15	10	C=C
3150				16	H–C benz.
3158				22	H–C benz.
3165				91	H–C benz.
3170				35	H–C benz.
3171			107		H–C benz.
3175	2 x 88		75		H–C benz.
3186			12		H–C benz.
3189				20	H–C benz.
3194					H–C benz.
3199				17	H–C benz.
3414–3421				879(7)^a	H–O2 stretch
3418–3421			861(7)^a		H–O2 stretch
3419		2 x 425			H–O2 stretch
3447–3455				733(5)^a	H–O6 stretch
3449–3451			1004(7)^a		H–O6 stretch
3453		2 x 536			H–O6 stretch
3455–3456			563(7)^a		H–O3 stretch
3457		532(3)^a			H–O3 stretch
3456–3461				832(9)^a	H–O6/H–O3 stretch

^aSum of (*n*) intensities within the given frequency range for a cluster of the same type (bold).

H–O2 stretch at 3419 cm^{-1} was spread out to a range from 3414 to 3421 cm^{-1} , the H–O6 stretch at 3453 cm^{-1} was spread out from 3447 to 3453 cm^{-1} , and the β -CD's H–O3 stretch at 3457 cm^{-1} from 3455 to 3456 cm^{-1} as a result of the interaction with the guest molecule. In general, the shifts were subtle; it seemed that the benzene itself was influenced very little if at all, the cyclodextrin on the contrary was substantially influenced in its frame vibrations over a wider range, its splitting most probably being caused by the “rigid benzene wheel” it had included,

see Figure 3. Interestingly, the concerted hydrogen-bond movements at the rims of β -CD rims were easily identified, for example in the BCDbenzeneparallelO23rO6l complex. Here, at 445 cm^{-1} , all hydrogen atoms bound to O6 atoms concertedly moved up and down the almost perfect plane of O6 oxygen atoms (marked H(O6) up/d in Table 4).

Since our two stable β -CD/benzene complex conformers (parallel and vertical) were energetically close neighbours with

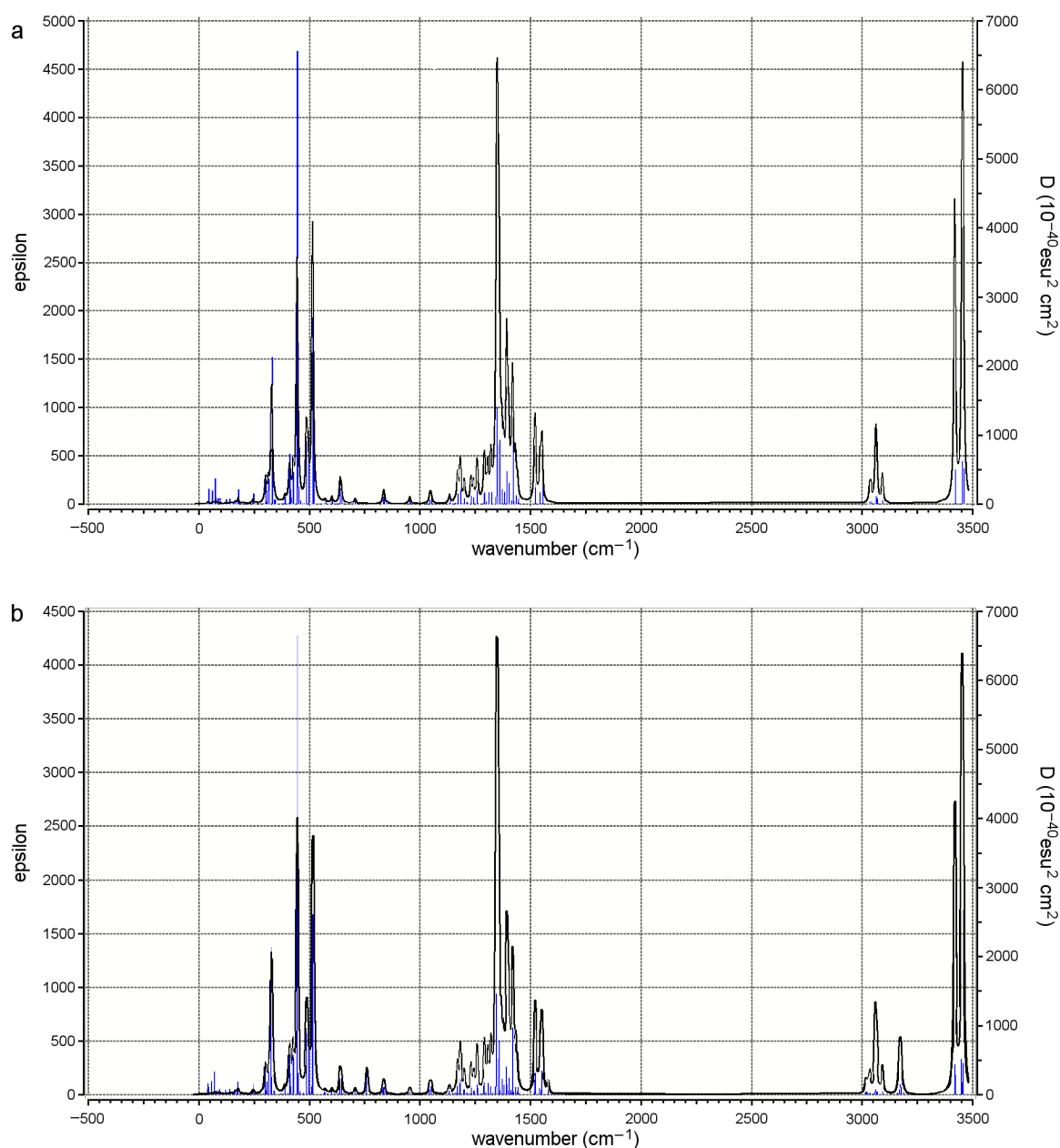


Figure 3: (a) AM1 calculated IR spectra of BCDO23rO6l empty and (b) the BCDO23rO6l/benzene inclusion complex with benzene in a parallel orientation.

only 0.29 kcal mol⁻¹ energy difference calculated at 0 K, the experimental spectrum at room temperature should display all cyclodextrin absorptions, i.e., both our conformers exist at the same time.

Summarised result of the AM1 calculations

AM1 energy/geometry optimisations in vacuo at 0 K showed that the β -CD/benzene complex with all hydrogen bonds at the O2/O3 rim in a right-hand orientation and at the narrower O6 rim in a left-hand orientation can include benzene in a parallel (best energy) or vertical position (higher in energy by only 0.29 kcal mol⁻¹); both conformers were thermodynamically allowed. The HOMO/LUMO gap of the empty β -CD with about 12 eV was lowered to about 10 eV in the complex. The β -CD/benzene inclusion complex was formed only by weak noncovalent interactions, which influence the IR lines of the β -CD most at its frame frequencies and at highly ordered concerted movements of its hydrogen bonds. The IR frequencies of benzene were influenced only marginally upon inclusion.

COSMO-RS calculations with Turbomole

Quantum mechanical geometry/energy optimisations

For the four empty β -CD models, for benzene in D_{6h} geometry, and for all eight BCD/benzene complexes, quantum mechanical geometry/energy optimisations with Turbomole and the implemented COSMO-RS method were performed at 0 K. For all starting structures the AM1 optimised geometries were used, and the results are given in Table 5 for the calculations in vacuo, and in Table 6 for the calculations in the COSMO-RS dielectric field of water (no explicit water molecules, but instead dipoles that model the dielectric field of the solvent are used by this method [18]).

The empty BCDO23rO6l model is still the lowest energy conformer among the four in vacuo (Table 5), but now the BCDO23lO6l is second and energetically close (0.26 kcal mol⁻¹), followed by BCDO23rO6r with 2.17 kcal mol⁻¹ and BCDO23lO6r with 2.3 kcal mol⁻¹ higher. In aquo (Table 6) the empty hydrogen bond model BCDO23rO6l was again the most preferred (Figure 4).

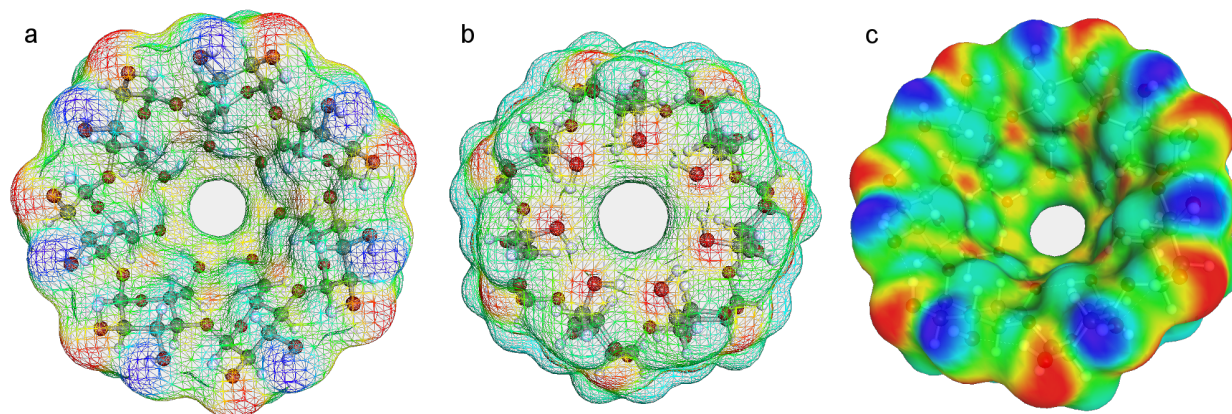
Although all optimisations were started from the AM1-optimised parallel and vertical geometries (from which the naming stems), the BP/TZVP-DISP3 method with its enhanced description of dispersion forces led in general to structures in which the benzene molecule adopts a more oblique position [14,20,32]. The whole energy hypersurface of β -CD complexes is rather flat, and therefore it is difficult to identify the many local minima [17]. Therefore, geometry optimisations of β -CD complexes still remain a multiple-minimum problem in general. The COSMO-RS method combines these identified multiple-minima structures as a set of conformers in a databank (which is updated principally when new minima are identified by more sensitive theoretical methods). From these sets the thermodynamic properties of the material at various temperatures can be calculated (by analytical formulas with COSMOthermX) more realistically than for one structure only [18–20]. Molecules adopt many conformers at temperatures higher than 0 K, which will be shown for the β -CD/benzene complex by MD simulations in the next section. The energetically best structure of the BCD/benzene complex in vacuo due to the BP/TZVP-DISP3 method resulted from the BCDO23rO6l/benzenevertical starting structure. The AM1 favourite BCDO23rO6l/benzeneparallel model followed second with 0.06 kcal mol⁻¹ above, and the others were higher up to a maximum of 2.34 kcal mol⁻¹ (BCDO23lO6r/benzeneparallel), Table 5. All optimised BP/

Table 5: Optimised BP/TZVP-DISP3 energies in vacuo. Lowest energy conformer of each group in bold.

BP/TZVP-DISP3 in vacuo	$E_{\text{abs.}}$ [a.u.]	$E_{\text{rel.}}$ [a.u.]	$E_{\text{rel.}}$ [kcal mol ⁻¹]
BCDO23lO6l	-4277.4641711793	0.0004226810	0.2652365729
BCDO23lO6r	-4277.4609298546	0.0036640057	2.2992002354
BCDO23rO6l	-4277.4645938603	0.0000000000	0.0000000000
BCDO23rO6r	-4277.4611238596	0.0034700007	2.1774601643
Benzene	-232.3358156576	0.0071740456	4.5017853670
parallelO23lO6l	-4509.8320598074	0.0005758687	0.3613633618
parallelO23lO6r	-4509.8289075499	0.0037281263	2.3394365033
parallelO23rO6l	-4509.8325394613	0.0000962148	0.0603757489
parallelO23rO6r	-4509.8291543743	0.0034813018	2.1845516926
verticalO23lO6l	-4509.8321043279	0.0005313482	0.3334263276
verticalO23lO6r	-4509.8289158257	0.0037198504	2.3342433180
verticalO23rO6l	-4509.8326356761	0.0000000000	0.0000000000
verticalO23rO6r	-4509.8291521426	0.0034835335	2.1859520815

Table 6: Optimised BP/TZVP-DISP3 COSMO-RS energies in aquo. Lowest energy conformer of each group in bold.

BP/TZVP-DISP3 in aquo	$E_{\text{abs.}}$ [a.u.]	$E_{\text{rel.}}$ [a.u.]	$E_{\text{rel.}}$ [kcal mol ⁻¹]
BCDO23IO6I	-4277.5397933200	0.0044818408	2.8123999202
BCDO23IO6r	-4277.5390925770	0.0051825838	3.2521231603
BCDO23rO6I	-4277.5442751608	0.0000000000	0.0000000000
BCDO23rO6r	-4277.5432219210	0.0010532398	0.6609185069
Benzene	-232.3408624099	0.0021272933	1.3348978187
parallelO23IO6I	-4509.9064609135	0.0041674010	2.6150858014
parallelO23IO6r	-4509.9056675976	0.0049607169	3.1128994621
parallelO23rO6I	-4509.9106283145	0.0000000000	0.0000000000
parallelO23rO6r	-4509.9096646704	0.0009636441	0.6046963092
verticalO23IO6I	-4509.9004752425	0.0101530720	6.3711542109
verticalO23IO6r	-4509.9055329919	0.0050953226	3.1973658847
verticalO23rO6I	-4509.9106053475	0.0000229670	0.0144120224
verticalO23rO6r	-4509.9095882112	0.0010401033	0.6526752221

**Figure 4:** The lowest energy empty conformer BCDO23rO6I, COSMO-RS structure optimised with BP/TZVP-DISP3 (solvent = water), σ -surface with charge code: oxygen red, hydrogen blue, hydrophobic area yellow and green. View a: from O2/O3 side, view b: from O6 side, view c: from O2/O3 side.

TZVP-DISP3 COSMO-RS energies in aquo were lower than the corresponding ones in vacuo, which is reasonable (Table 6). Here the BCDO23rO6I hydrogen-bond starting models remained the best ones after energy/geometry minimisations, for the empty β -CD and for the complex as well.

The BP/TZVP-DISP3 energies of complex formation were calculated from the data of Table 5 and Table 6 by Equation 1 and are shown in Table 7. All energies of complex formation $\Delta E_{\text{Complexation}}$ in vacuo and in aquo were negative, indicating that the complexes will be formed. The complex structures of the reactions with the lowest energies were BCDO23rO6I/benzeneparallel in aquo and BCDO23rO6I/benzenevertical in vacuo.

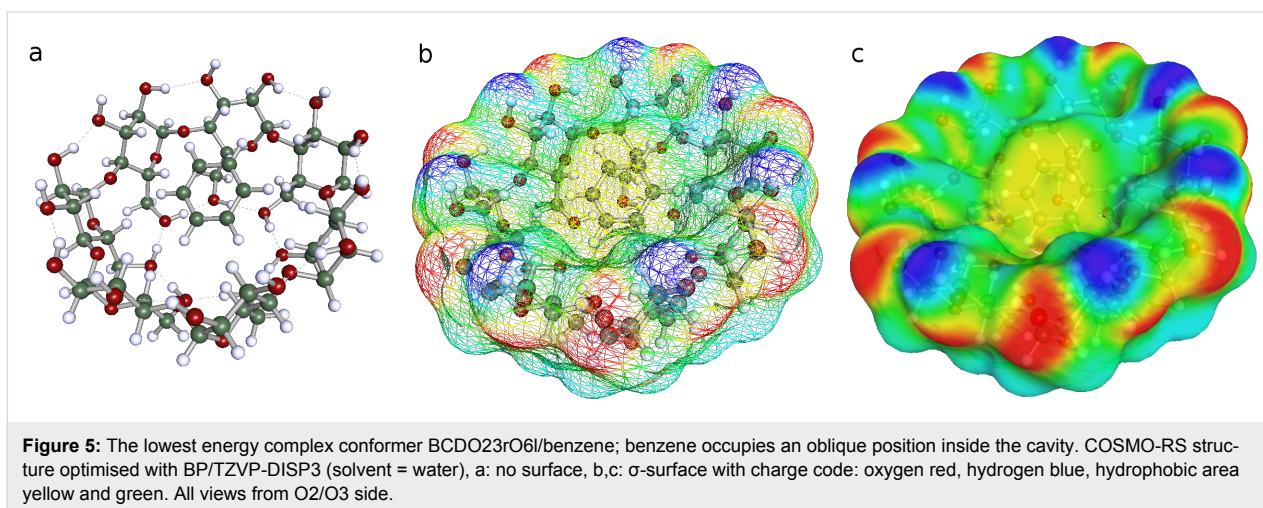
Figure 4a–c and Figure 5a–c display different views of the best BCDO23rO6I empty model and the BCDO23rO6I/benzene complex with their COSMO-RS σ -surfaces. Hydrophilic areas are coloured in red (negative charge of oxygen atoms) or blue (positive charge of hydrogen atoms), which were mostly located at the O23 rim. Hydrophobic areas (coloured green/yellow) can be observed easily inside the BCDO23rO6I empty model (caused by its O6 highly ordered hydrogen-bond rim) and the whole surface of benzene while sitting inside the cyclodextrin cavity.

Consequences for thermodynamic analysis

The basis of thermodynamic analysis with COSMOthermX [18–20] is the so called σ -profiles and σ -potentials of each molecule,

Table 7: BP/TZVP-DISP3 energies of complex formation in vacuo and in aquo, calculated with the absolute data of BCDO23rO6I for E_{BCDempty} . Relative $\Delta E_{\text{Complexation}}$ with respect to the energetically lowest complex conformer, column 2 and column 4.

BP/TZVP-DISP3	abs. $\Delta E_{\text{Complexation}}$ in vacuo [kcal mol ⁻¹]	rel. $\Delta E_{\text{Complexation}}$ in vacuo [kcal mol ⁻¹]	abs. $\Delta E_{\text{Complexation}}$ in aquo [kcal mol ⁻¹]	rel. $\Delta E_{\text{Complexation}}$ in aquo [kcal mol ⁻¹]
parallelO23IO6I	-19.861	0.361	-13.381	2.615
parallelO23IO6r	-17.883	2.339	-12.883	3.113
parallelO23rO6I	-20.162	0.060	-15.996	0.000
parallelO23rO6r	-18.038	2.184	-15.391	0.605
verticalO23IO6I	-19.889	0.333	-9.625	6.371
verticalO23IO6r	-17.888	2.334	-12.798	3.198
verticalO23rO6I	-20.222	0.000	-15.981	0.015
verticalO23rO6r	-18.036	2.186	-15.343	0.653



which show the characteristic differences in the electrostatic molecular surfaces that are calculated quantum mechanically at 0 K for molecular conformers of up to 10 kcal mol⁻¹ in the standard COSMO databases. These σ -profiles and σ -potentials were calculated with COSMOthermX for the empty β -CD models and are displayed in Figure 6 and Figure 7. Especially the σ -potentials (Figure 7) showed that the empty β -CD models were split into two groups. One group had the two hydrogen-bond models with all hydrogen bonds in the left-hand orientation (O23l), the other one with the two models with right-hand (O23r) orientation. This mirrored quantitatively the result of how much the two energetically preferred O23r models were closer to each other and were most different from the two O23l models, regardless of which orientation the hydrogen bonds had at the O6 rim. By the time these four empty β -CD models have been combined into one set of β -CD conformers and stored like this in the COSMO database, thermodynamic analysis with COSMOthermX will take into account these differences and give more accurate results for many material properties over a wider temperature range than with simple “one-molecule sets” [18–20].

Summarised result of the COSMO-RS calculations

In aquo the hydrogen-bond model BCDO23rO6I was again the most preferred, not only among the empty β -CD conformers but also as the BCDO23rO6I/benzene parallel complex model. According to the quantum mechanical COSMO-RS calculations, all empty β -CD and all complex conformers of β -CD/benzene were thermodynamically allowed, i.e., the complexes should be formed. The highest relative energy of all conformers was less than 2.4 kcal/mol⁻¹ in vacuo and less than 6.4 kcal mol⁻¹ in aquo. The benzene molecule adopted an oblique position in the inclusion complex. The σ -surface of the empty β -CD “looked” more hydrophilic from the outside at its O2/O3 side (red and blue) than from the O6 side (only red and yellow/green). From the inside, the O6 rim of β -CD “looked” hydrophobic.

Molecular-dynamics simulations

The empty cyclodextrin at 300 K

The “highly ordered hydrogen bonds” of the BCDO23IO6I, BCDO23rO6I, BCDO23IO6r and BCDO23rO6r models cannot be observed at 300 K where the temperature is high enough to

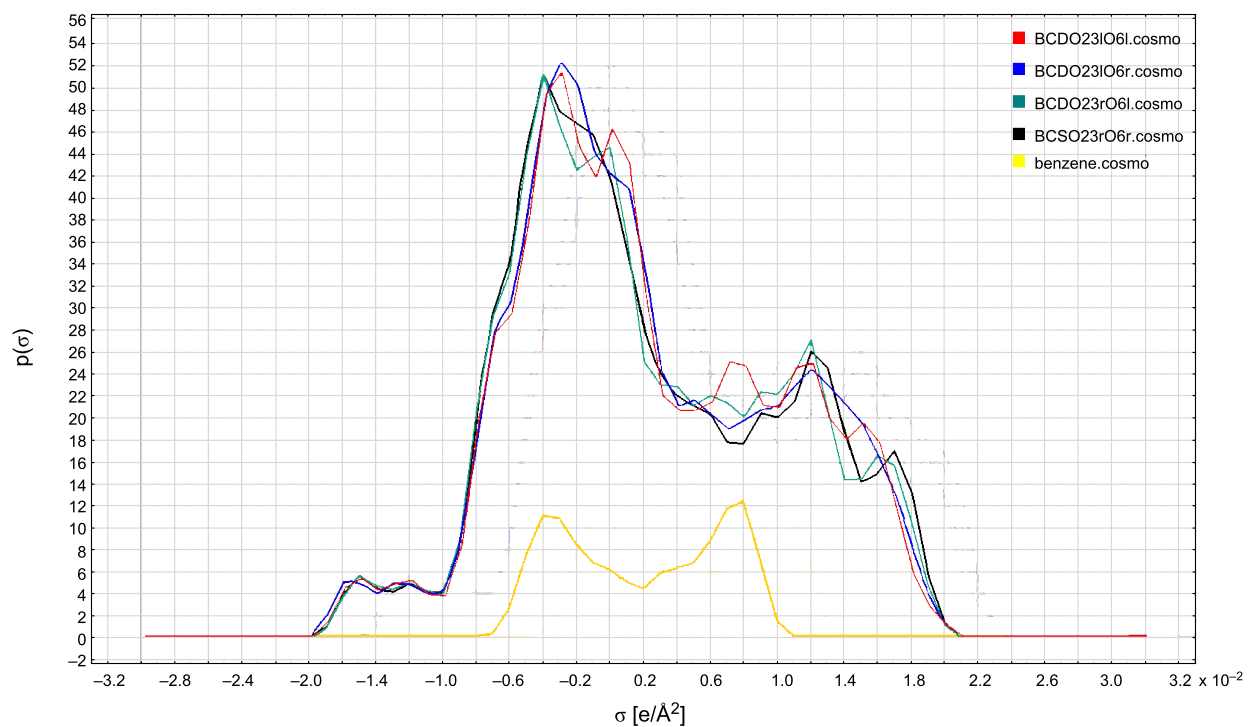


Figure 6: σ -Profiles of the COSMO-RS method for the four empty β -CD models and benzene (BP/TZVP-DISP3 method).

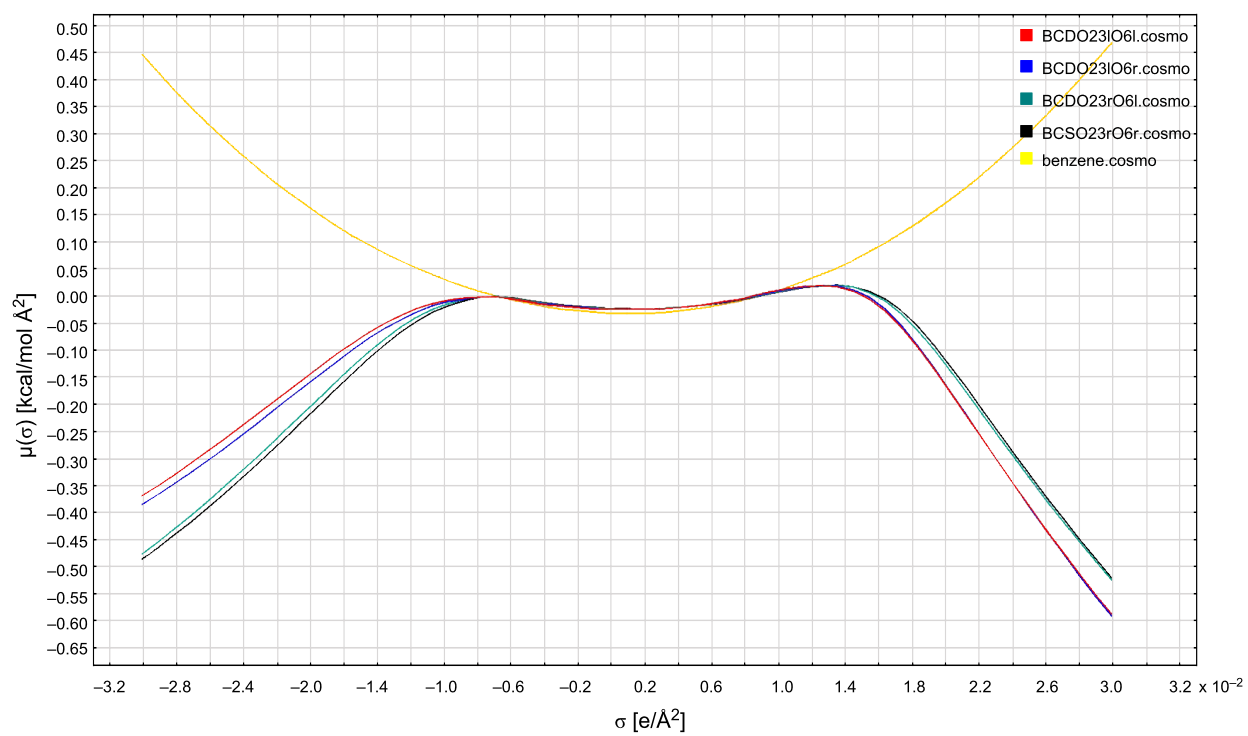
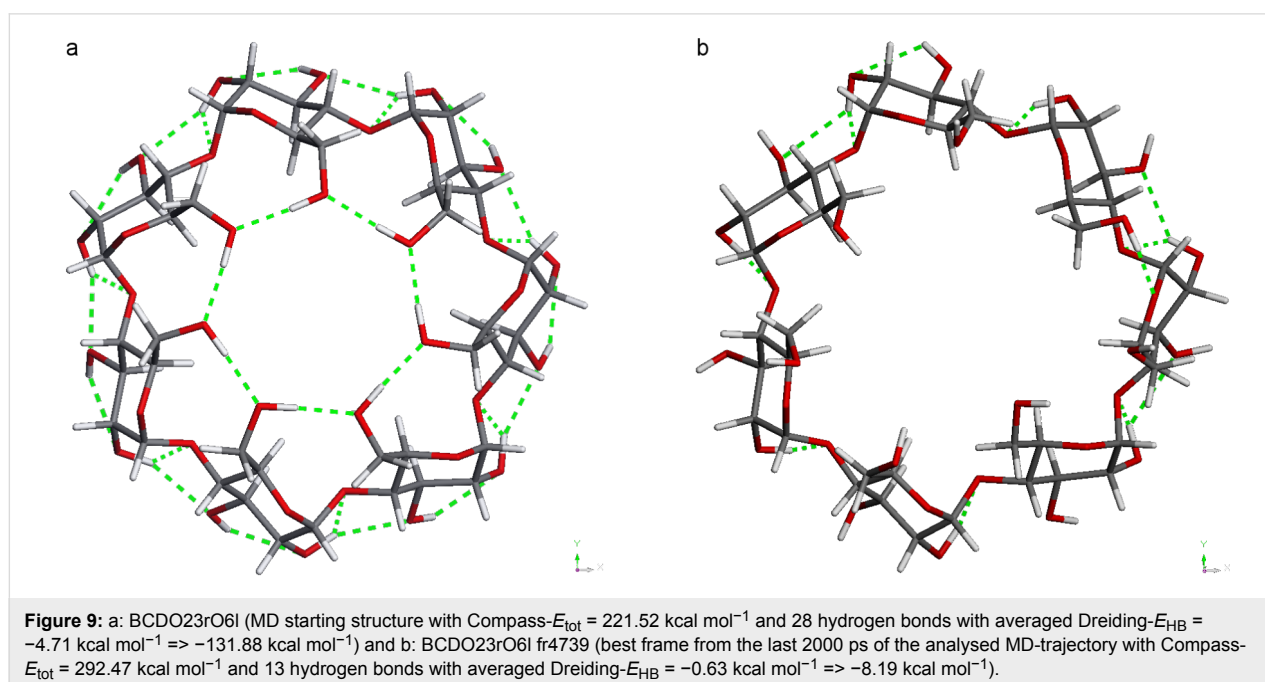
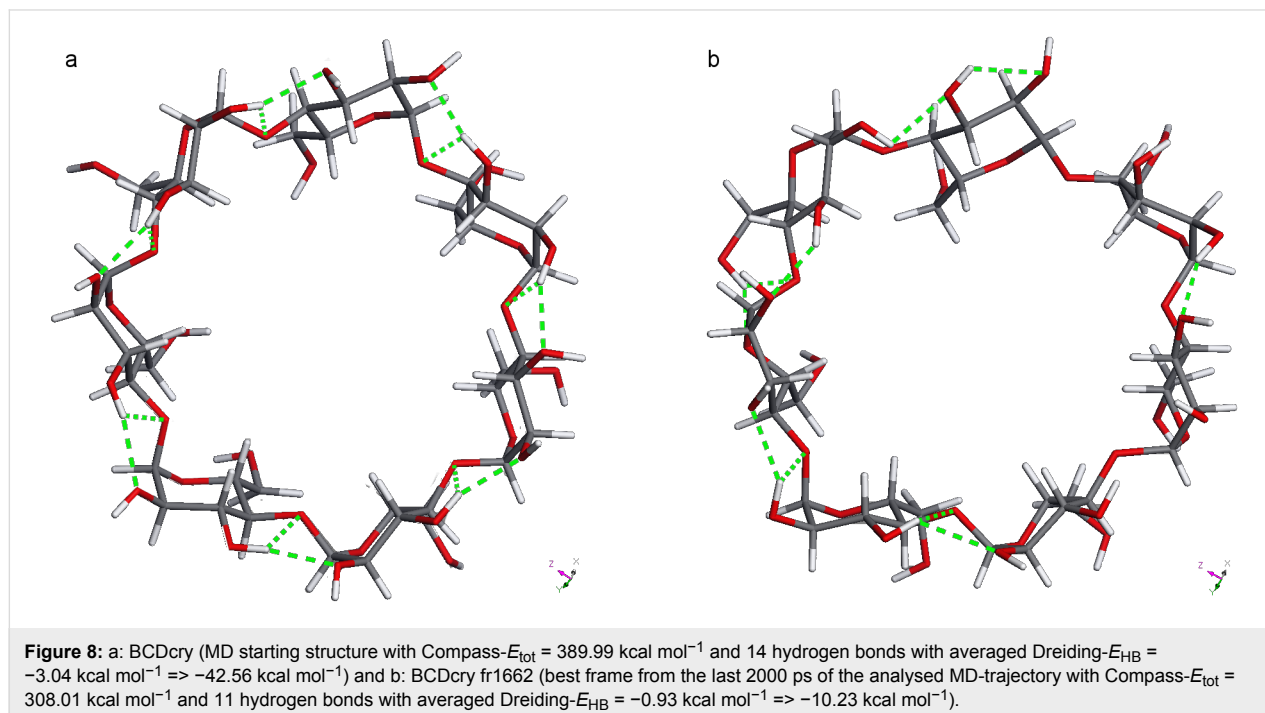
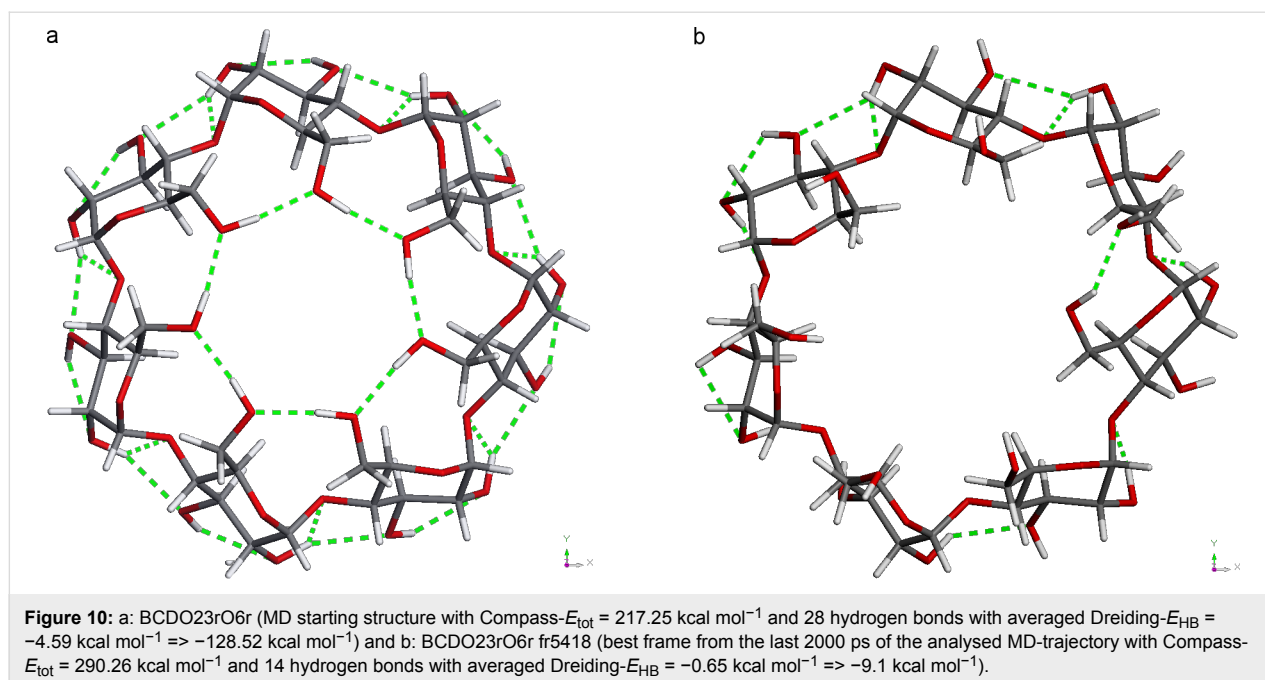


Figure 7: σ -Potentials of the COSMO-RS method for the four empty β -CD models and benzene (BP/TZVP-DISP3 method).

make the β -CDs flexible and the energy barriers of hydrogen bonds are overcome. As was observed in the neutron diffraction studies of β -CDs at 300 K [33] and 120 K [35] the temperature determined the hydrogen-bond distribution. In the MD simulations started from the “highly ordered” 0 K AM1 models the hydrogen bonds switched to more disordered arrangements. For comparison the crystal structure model of 300 K was added. The starting structures and the energetically best frame from the

last 2000 ps of the 6000 ps trajectories are displayed in Figure 8, Figure 9 and Figure 10. Only the two energetically best models are shown: BCDO23rO6l and BCDO23rO6r. The Compass- E_{tot} energy of BCDcry was higher in the starting structure than in the best frame of the trajectory at 300 K, which accounts for the number of packing forces in the crystal that deformed the cyclodextrin monomer in vacuo. For the two other models BCDO23rO6l and BCDO23rO6r this was reversed;





their “highly ordered hydrogen-bond structures” in vacuo from the start were energetically lower than their structures from their best frames of the trajectories at 300 K, because they “lost their hydrogen-bond stabilisation” and became flexible instead. All starting structures had much higher hydrogen-bond energies (analysis using Dreiding force field) than the best structures from their trajectories.

The five MD trajectories of the empty β -CD models had the following averaged energies (Compass force field): BCDery $311.15 \text{ kcal mol}^{-1}$, BCDO23IO6I $304.26 \text{ kcal mol}^{-1}$, BCDO23IO6r $304.02 \text{ kcal mol}^{-1}$, BCDO23rO6I $295.14 \text{ kcal mol}^{-1}$ and BCDO23rO6r $291.22 \text{ kcal mol}^{-1}$. According to this the BCDO23rO6r model was the lowest energy conformer, closely followed by the BCDO23rO6I model, which was the best conformer in the AM1 method. This means, that the hydrogen bonds at the larger O2/O3 rim of empty β -CDs preferred the right-hand orientation, i.e., O3–H \cdots O2–H in the same glucose unit and bifurcated towards O4 and O3 of the next glucose unit on the right side.

The β -CD/benzene complex at different temperatures from 300 K to 120 K

The MD trajectories of the parallel and vertical BCDO23rO6I model were practically the same since the benzene guest changed its position already in the first few pico seconds, and later a significant difference could not be observed any more. It was observed that the benzene guest left the cyclodextrin host at slightly different times, depending on the temperature, (Figure 11). At 120 K benzene was always inside the cyclodex-

trin cavity. “Inside the cavity” was defined by a distance of up to 2 Å between the centre of mass of benzene and the seven O4 oxygen atoms of the β -CD that define a plane in the middle of the torus. At 273 K this distance extended up to 3 to 4 Å starting at about 140 ps, and at about 320 ps it became longer than 5 Å indicating that now benzene had left the cavity without return. At 280 K and 290 K these events appeared earlier, at 140 ps/160 ps and 50 ps/125 ps, respectively. This was in good qualitative agreement with the complex formation constants that were experimentally determined from UV spectra in [34] and showed that the complex was more stable at lower temperatures. At 300 K benzene tried to leave the cavity early at the O6 side, but always returned. It spent most of the time at a greater distance of about 4 to 5 Å to the O4-plane of β -CD and finally left the cavity on the O2/O3 side as all the others did, but later at 2400 ps. Benzene had a high mobility in many geometrical positions in these trajectories, and the most characteristic ones from the trajectory at 300 K are shown in Figure 12. Figure 13 displays how different the energies of the β -CD/benzene complex were during the MD trajectory at 290 K.

Summarised result of the molecular dynamics simulations

Molecular dynamics simulations at different temperatures (120 K to 300 K) displayed qualitatively correctly the dynamics of the β -CD/benzene complex. Benzene accepted many different positions inside the β -CD cavity before it finally left at the O2/O3 side. The highly ordered hydrogen bonds at the O6 or O2/O3 rim were not existent at these temperatures and thus were a phenomenon of very low temperatures.

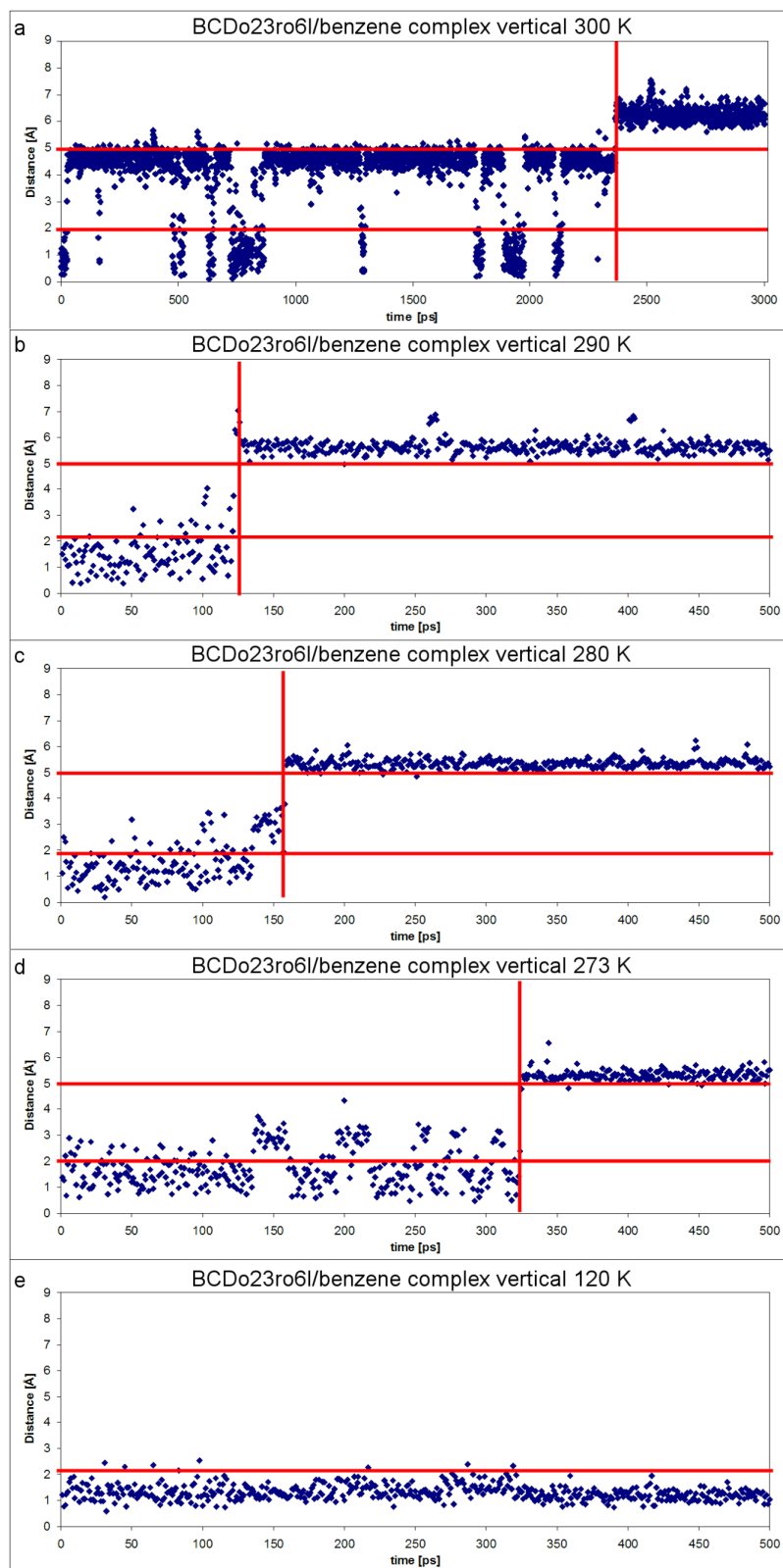
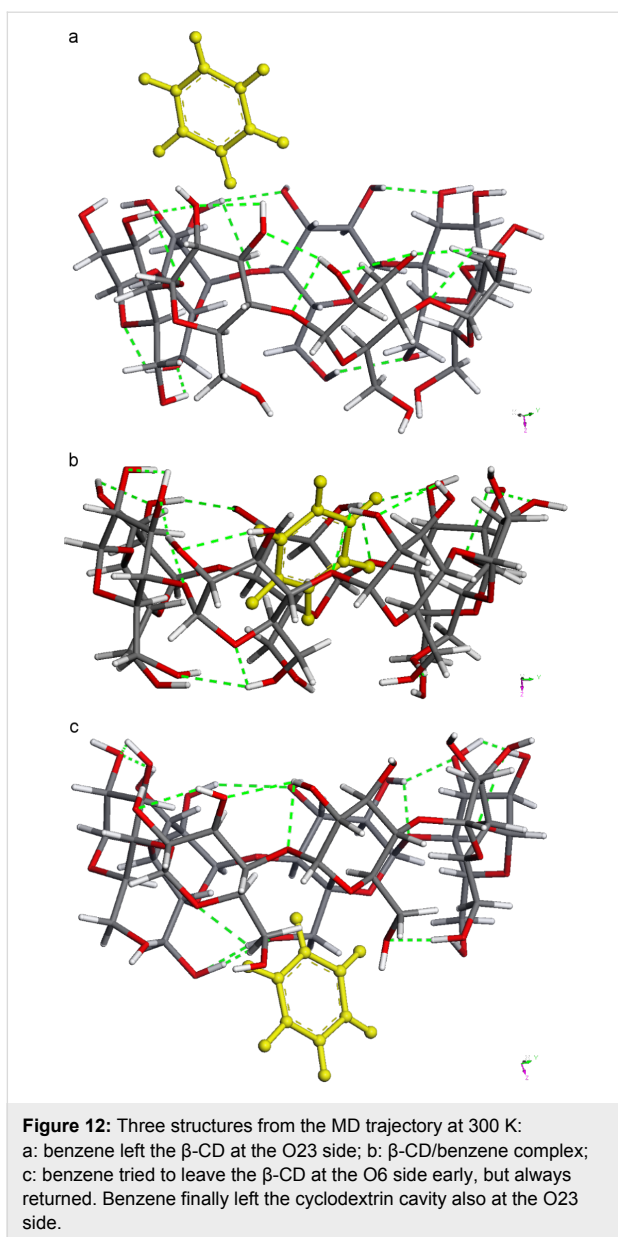


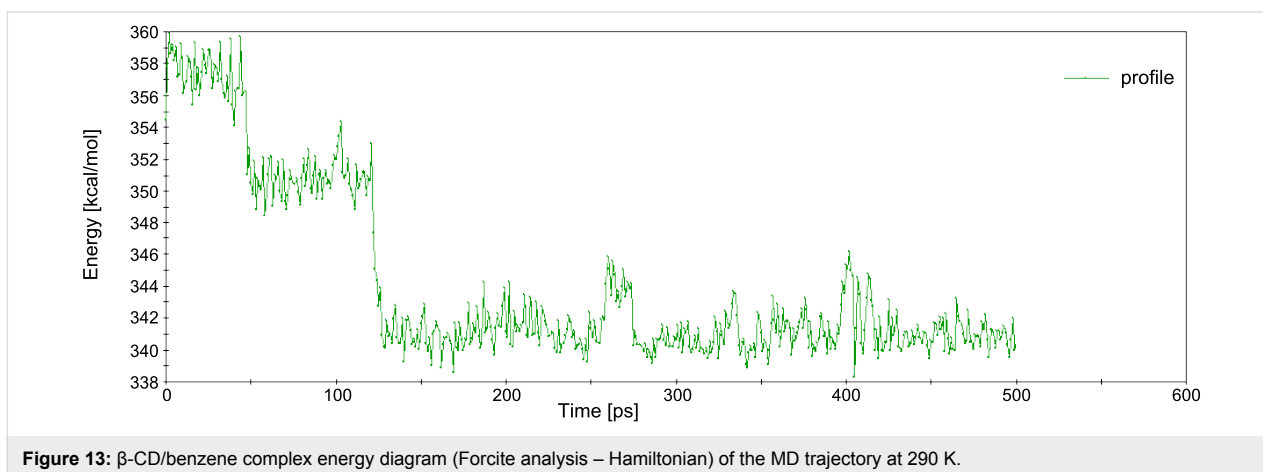
Figure 11: Distance plots from the MD trajectories of 500 ps each (x axis) for β -CD/benzene complex at different temperatures: a: 300 K, b: 290 K, c: 280 K, d: 273 K and e: 120 K. Distance [Å] (y axis) between the centres of mass of benzene and the seven O4 atoms of β -CD; horizontal red lines at distances of 2 Å and 5 Å, see text; vertical red lines indicate the time when the benzene left the β -CD cavity without return.



Since quantum-mechanical calculations lead to these highly ordered optimised structures, they represent the lowest energy conformers in vacuo at 0 K. Nevertheless, they will be the basis for the assumptions of the COSMO-RS method and its thermodynamic formulas and calculations/extrapolations based on the energetically lowest single-molecular states, plus conformers on top of that, up to 10 kcal mol^{-1} . These conformer ensembles will make it possible to calculate several thermodynamic properties at room temperature and above, not only for the gas phase but also for the liquid and solid states of matter made from these molecules.

Conclusion

The β -CD/benzene AM1 energies of complex formation after geometry optimisations for both positions of the guest, parallel and vertical were $-2.75 \text{ kcal mol}^{-1}$ and $-2.46 \text{ kcal mol}^{-1}$, respectively. The concerted highly ordered hydrogen bonds at the O2/O3 and O6 rims of cyclodextrin had a strong influence on the structures. This was confirmed by the calculated AM1 IR spectra, which showed that the β -CDs O–H frequencies were mostly split up upon complex formation with benzene. The HOMO/LUMO gap of the empty cyclodextrin with about 12 eV was lowered to about 10 eV in the complex, in agreement with other calculations (PM3) from the literature. All four hydrogen-bond models studied here (BCDO23IO6I, BCDO23rO6I, BCDO23IO6r and BCDO23rO6r) were energetically advantageous after geometry optimisations with BP86/TZVP-DISP3 in vacuo and in aquo, their relative energies were, all but one, less than $3.2 \text{ kcal mol}^{-1}$. The character of the COSMO-RS σ -surface of β -CD was much more hydrophobic on its O6 rim than on its O23 side when all hydrogen bonds were arranged in a concerted mode. COSMO-RS BP/TZVP-DISP3 calculations in vacuo and in aquo (dielectric field approximation) preferred an oblique position (over parallel or vertical) of benzene inside the β -CD cavity and suggested energies of complex formation up to $20.2 \text{ kcal mol}^{-1}$ in vacuo and of about 16 kcal/mol^{-1} in aquo.



The static picture of these energetically lowest structures of the β -CD/benzene complex at 0 K was extended by the molecular dynamics simulations at different temperatures (120 K to 300 K). They displayed qualitatively correctly the dynamics of the β -CD/benzene complex, which was more stable at lower temperatures, especially at 120 K. Benzene accepted many different positions inside the β -CD cavity before it finally left at the O2/O3 side (273 to 300 K). The highly ordered hydrogen bonds at the O6 or O2/O3 rim were a subtle phenomenon at very low temperatures. Still, we assume that hydrogen bonds of cyclodextrins nevertheless play a crucial role in all their special properties, i.e., their many inclusion complexes and remarkable influence on reactions, catalysis and supramolecular structures.

References

- Szejtli, J. *Chem. Rev.* **1998**, *98*, 1743–1754. doi:10.1021/cr970022c
- Jeffrey, G. A.; Saenger, W. Chapter 4. Theoretical calculations of hydrogen bond geometries. *Hydrogen Bonding in Biological Structures*; Springer Verlag: Berlin Heidelberg, 1991. See also Chapter 18, Cyclodextrins.
- Alderfer, J. L.; Eliseev, A. V. *J. Org. Chem.* **1997**, *62*, 8225–8226. doi:10.1021/jo970772m
- Schmid, B. V. K. J.; Hetzer, M.; Ritter, H.; Barner-Kowollik, C. *Macromolecules* **2011**, *44*, 7220–7232. doi:10.1021/ma2011969
- Eliadou, K.; Yannakopoulou, K.; Rontoyianni, A.; Mavridis, I. M. *J. Org. Chem.* **1999**, *64*, 6217–6226. doi:10.1021/jo990021f
- Aachmann, F. L.; Otzen, D. E.; Larsen, K. L.; Wimmer, R. *Protein Eng., Des. Sel.* **2003**, *16*, 905–912. doi:10.1093/protein/gzg137
- Willerich, I.; Schindler, T.; Ritter, H.; Gröhn, F. *Soft Matter* **2011**, *7*, 5444–5450. doi:10.1039/c1sm05405h
- Park, C.; Oh, K.; Lee, S. C.; Kim, C. *Angew. Chem., Int. Ed.* **2007**, *46*, 1455–1457. doi:10.1002/anie.200603404
- Kida, T.; Iwamoto, T.; Fujino, Y.; Tohnai, N.; Miyata, M.; Akashi, M. *Org. Lett.* **2011**, *13*, 4570–4573. doi:10.1021/ol2017627
- Koehler, J. E. H.; Saenger, W.; van Gunsteren, W. F. *J. Biomol. Struct. Dyn.* **1988**, *6*, 181–198. doi:10.1080/07391102.1988.10506490
- Köhler, J.; Hohla, M.; Söllner, R.; Eberle, H.-J. *Supramol. Sci.* **1998**, *5*, 101–116. doi:10.1016/S0968-5677(97)00065-5
- Köhler, J.; Hohla, M.; Söllner, R.; Amann, M. *Supramol. Sci.* **1998**, *5*, 117–137. doi:10.1016/S0968-5677(97)00066-7
- Yu, H.; Amann, M.; Hansson, T.; Köhler, J.; Wich, G.; van Gunsteren, W. F. *Carbohydr. Res.* **2004**, *339*, 1697–1709. doi:10.1016/j.carres.2004.05.003
- Damodaran, K. V.; Banba, S.; Brooks, C. L., III. *J. Phys. Chem. B* **2001**, *105*, 9316–9322. doi:10.1021/jp010361g
- Liu, L.; Guo, Q.-X. *J. Phys. Chem. B* **1999**, *103*, 3461–3467. doi:10.1021/jp984545f
- Liu, L.; Guo, Q.-X. *J. Chem. Inf. Comput. Sci.* **1999**, *39*, 133–138. doi:10.1021/ci980097x
- Liu, L.; Guo, Q.-X. *J. Inclusion Phenom. Macrocyclic Chem.* **2004**, *50*, 95–103. doi:10.1007/s10847-003-8847-3
- Klamt, A. *COSMO-RS From Quantum Chemistry to Fluid Phase Thermodynamics and Drug Design*, 1st ed.; Elsevier: The Netherlands, 2005.
- Klamt, A.; Eckert, F.; Arit, W. *Annu. Rev. Chem. Biomol. Eng.* **2010**, *1*, 101–122. doi:10.1146/annurev-chembioeng-073009-100903
- Klamt, A.; Reinisch, J.; Eckert, F.; Hellweg, A.; Diedenhofen, M. *Phys. Chem. Chem. Phys.* **2012**, *14*, 955–963. doi:10.1039/c1cp22640a
- Saenger, W. *Angew. Chem.* **1980**, *92*, 343–361. doi:10.1002/ange.19800920505
- Ho, B. T.; Joyce, D. C.; Bhandari, B. R. *Food Chem.* **2011**, *129*, 259–266. doi:10.1016/j.foodchem.2011.04.035
- Regiert, M. *SOFW J.* **2009**, *8*, 135.
- Schmid, G. Preparation and industrial production of cyclodextrins. In *Comprehensive Supramolecular Chemistry*; Atwood, J.; Davies, E. D.; MacNicol, D. D.; Vogtle, F., Eds.; Pergamon Press: Oxford, U.K., 1996; pp 41–56.
- van der Veen, B. A.; Uitdehaag, J. C. M.; Dijkstra, B. W.; Dijkhuizen, L. *Biochim. Biophys. Acta* **2000**, *1543*, 336–360. doi:10.1016/S0167-4838(00)00233-8
- Li, Z.; Wang, M.; Wang, F.; Gu, Z.; Du, G.; Wu, J.; Chen, J. *Appl. Microbiol. Biotechnol.* **2007**, *77*, 245–255. doi:10.1007/s00253-007-1166-7
- Rungsardthong Ruktanonchai, U.; Srinuanchai, W.; Saesoo, S.; Sramala, I.; Puttipipatkachorn, S.; Soottitantawat, A. *Biosci., Biotechnol., Biochem.* **2011**, *75*, 2340–2345. doi:10.1271/bbb.110523
- Hobza, P.; Selzle, H. L.; Schlag, E. W. *J. Am. Chem. Soc.* **1994**, *116*, 3500–3506. doi:10.1021/ja00087a041
- Materials Studio*, Version 5.5; Accelrys: San Diego, USA.
- Gaussian 09*, Revision A.1; Gaussian, Inc.: Wallingford, CT, 2009. GaussView v 5.0.9 Visualizer and Builder.
- Turbomole*, Version 6.4; COSMOlogic GmbH: Leverkusen, Germany.
- Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104. doi:10.1063/1.3382344
- Betzler, C.; Saenger, W.; Hingerty, B. E.; Brown, G. M. *J. Am. Chem. Soc.* **1984**, *106*, 7545–7557. doi:10.1021/ja00336a039
- Trofymchuk, I. M.; Belyakova, L. A.; Grebenyuk, A. G. *J. Inclusion Phenom. Macrocyclic Chem.* **2011**, *69*, 371–375. doi:10.1007/s10847-010-9757-9
- Zabel, V.; Saenger, W.; Mason, S. A. *J. Am. Chem. Soc.* **1986**, *108*, 3664–3673. doi:10.1021/ja00273a020

License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The license is subject to the *Beilstein Journal of Organic Chemistry* terms and conditions: (<http://www.beilstein-journals.org/bjoc>)

The definitive version of this article is the electronic one which can be found at: doi:10.3762/bjoc.9.15