Journal of Inclusion Phenomena and Macrocyclic Chemistry **44:** 9–14, 2002.

Theoretical Insights into the Formation, Structure, and Energetics of Some Cyclodextrin Complexes

NICHOLAS BODOR[1,2]* and PETER BUCHWALD[1]
[1]*IVAX Research, Inc., 4400 Biscayne Boulevard, Miami, FL 33137, USA;* [2]*Center for Drug Discovery, University of Florida, Gainesville, FL, USA*

(Received: 7 May 2002; in final form: 1 October 2002)

Key words: α-, β-, and γ-cyclodextrin complex, deformation index, estradiol, free energy, molecular size

Abstract

To investigate the physicochemical aspects relevant for the formation of various cyclodextrin inclusion complexes and to search for corresponding general structure–complex-stability relationships, stability data of 1 : 1 complexes for 179, 310, and 51 guest molecules with unsubstituted α-, β-, and γ-cyclodextrin were collected. Statistical analysis using structure-based parameters such as molecular size, hydrophobicity, rotatable bonds, electronic properties, and the presence or absence of more than 150 various functional or structural moieties were performed. The complexation thermodynamics could be well described within the framework of our recently introduced molecular size-based model for nonassociative liquids. With increasing guest size, 1 : 1 complex stability, as measured by ln K or ΔG^0, increases linearly up to a size limit characteristic for each CD, and the corresponding slopes and intercepts are in agreement with those predicted by the model. For larger structures, values level off and are scattered around an average value depending on shape, goodness of fit, and possibly lipophilicity and some specific effects (e.g. such as those caused by presence of phenol functionality). The complexation between β-cyclodextrin and certain large steroidal guest molecules, especially a brain-targeted estradiol chemical delivery systems (E_2-CDS) that is under clinical development, was investigated in details based on fully relaxed semiempirical AM1 quantum chemical calculations. A deformation index (DI) of the CD ring computed using these fully optimized host-guest geometries could be used to characterize the conformational change of the guest.

Introduction

Formation of cyclodextrin inclusion complexes is known to involve mainly nonspecific, weak interactions and no covalent bonding. Therefore, the transfer of organic guest molecules from the aqueous solution into the cyclodextrin cavity is in many aspects very similar to the transfer (partition) that takes place between an aqueous and a somewhat hydrophobic organic phase [1]. As our recently introduced general, molecular size-based model for nonassociative organic liquids and water [2–4] described the latter partition very well, we explored its ability to describe inclusion complexation with cyclodextrins. Therefore, we focused not as much on accurately predicting binding constants, but on placing the general physicochemistry of complex formation into the unified framework developed for organic liquids.

Experimental

Experimental thermodynamic data (log complex stability constants, log K, and standard free energies, ΔG^0) for 1 : 1 inclusion complexation of a large variety of guests with natural α-, β-, and γ-cyclodextrins were collected from the literature [1, 5–8]. For compounds with more experimental data available, average values were used. Three-dimensional molecular structures were built and the corresponding molecular volumes, used here as size descriptor, were calculated using computer models and softwares as described in our previous publications [2, 4]. Statistical analyses were performed using a standard spreadsheet program (Microsoft Excel 97). AM1 geometry-optimizations were performed using CAChe 5.0 software (Fujitsu, Ltd., Chiba, Japan).

* Author for correspondence: E-mail: Nicholas_Bodor@ivax.com

Results and discussion

Complex stability

Stability data of 1 : 1 complexes for 179, 310, and 51 guest molecules with unsubstituted α-, β-, and γ-CD, respectively were collected and used to investigate the physicochemical aspects relevant for the formation of various inclusion complexes and to search for corresponding general structure–complex-stability relationships. Analysis of the size (molecular volume)-dependence of these data (Figure 1) indicated, that as a first approximation, it is reasonable to assume that the free energy of complexation decreases linearly with size up to a size-limit, which has different values for α-, β-, and γ-CD, and more or less levels off after this.

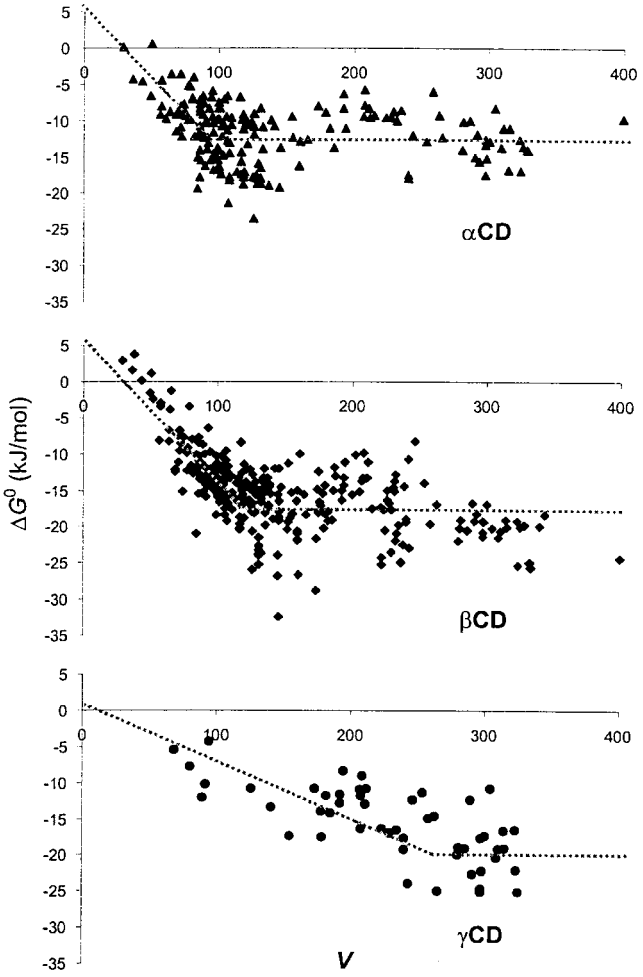

Figure 1. Size (molecular volume) dependence of standard free energies of 1 : 1 complexation with natural α-, β- and γ-CD for 179, 310 and 51 various guest molecules, respectively. The dashed line represents the value obtained from the unified liquid model and its extension for water. For sufficiently small guests, the slope (–0.2) and the intercept (6) are the same for α- and β-CD; however, limiting sizes are different: 95 Å^3 (–12.7 kJ/mol) and 120 Å^3 (–17.7 kJ/mol), respectively. Situation is somewhat different for γ-CD ($1 - 0.08v$ for $v < 260$ Å^3; –20.0 for $v \geq 260$ Å^3), see text for details.

Obviously, for every CD cavity there is an upper limit (Figure 2), and this is also shape-dependent. Molecular volume gives only an overall size-estimate, whereas the quality of the fit certainly has to be shape-dependent and can also be influenced by specific interactions of the functional moieties present. Flexible and rigid guests also behave differently, as flexible guest can more easily adopt a configuration better suited for inclusion. Nevertheless, this very simplified model (represented by the dashed line in Figure 1), which takes into account absolutely no shape- or specific interactions-related effects, already accounts for more than half of the variance for β-CD where most data were available. This confirms that, despite the large functional variety of the guest molecules, nonspecific interactions, which are to a great extent size-related, are dominant in determining CD-complex stability.

In an important development toward obtaining a unified description of organic liquids, this basic model of CD complex stability, could be very well fitted within the framework of our recently introduced general, molecular size-based model for nonassociative organic liquids and water [2–4]. Within this framework, as generally done for every property included in this model, the equation describing complex stability constants and/or corresponding standard free energies has to be obtained starting from the general form of the chemical potential $\mu_i^{\mathrm{sol},j}$ of solute i in solvent j [2–4]:

$$\mu_i^{\mathrm{sol},j} = kT\ln\left(\frac{\rho_i}{f}\Lambda_i^3\right) - \left(w_{ij} + \frac{kT}{v_j}\right)v_i - w_0 + kT. \quad (1)$$

Here k is the Boltzmann constant, T is the absolute temperature, Λ is the thermal de Broglie wavelength, $\rho_i = N_i/V$ represents particle density, v_i is (calculated) molecular volume, f denotes the fraction of the total volume of the liquid assumed to be available for translation ($f = 0.023$), and $w_0 = \omega_0 kT_0$ and $w = \omega kT_0$ ($T_0 = 298.15$ K) are the interaction-related constants that form the basis of the model. For all organic non-associative liquids described by the model, including haloalkanes, aromatics, haloaromatics, esters, and ketones, these interaction-related constants are assumed to have the same (or at least a very similar) value: $\omega_0 = 5.39$, $\omega = 0.082$. This chemical potential could be used in our unified liquid model to give good description of a variety of physicochemical properties, including boiling point (T_b), enthalpy of vaporization (H_{vap}), vapor pressure ($\log \rho^{\mathrm{gas}}$), Ostwald absorption coefficient ($\log \gamma$), surface tension (σ), and a number of partition and solubility data ($\log P$, $\log \rho_w$) for organic liquids that do not contain associative or strongly polar substituents [4]. Furthermore, despite the unusual properties of water, solvation in water can be described by the very same model, but a different, less favorable interaction constant ($\rho_w = -0.070$) has to be used because the unfavorable perturbation of hydrogen bonding in the solvent [3, 4]. The assumptions used for this extension fitted well with those of the modified hydration-shell hydrogen-bond model introduced by Muller [9, 10].

For CD complexation, the chemical potential at equilibrium has to have the same value for the guest molecule "freely" solvated in water and for that "bound" to the cyclodextrin host, which, in fact, is also solvated in water. Starting from here, and assuming that translational movement of "bound" guest molecules is restricted to a fraction ϕ of the part of the total liquid volume associated with the CD molecules, which, on its turn, for sufficiently dilute CD aqueous solutions, can be considered as $\rho_{\mathrm{CD}} v_{\mathrm{CD}}/\rho_w^0 v_w$, we could obtain an expression for the equilibrium constant, K = [CD·guest$_{\mathrm{bound}}$]/[guest$_{\mathrm{free}}$] [CD], of the guest-host complexation process

$$\mathrm{guest}_{\mathrm{free}} + \mathrm{CD} \overset{K}{\longleftrightarrow} \mathrm{CD}\cdot\mathrm{guest}_{\mathrm{bound}}.$$

After a few rearrangements and introducing all the numeric values (the ϕ fraction being the only undetermined one) the following expression is obtained:

$$\ln K = \ln\phi + \ln\frac{v_{\mathrm{CD}}}{v_w\rho_w^0} + (\omega - \omega_w - \frac{1}{v_w})v = \ln\phi + \ln\frac{1135.0}{14.6\times 55.2} + (0.082 + 0.070 - \frac{1}{14.6})v. \quad (2)$$

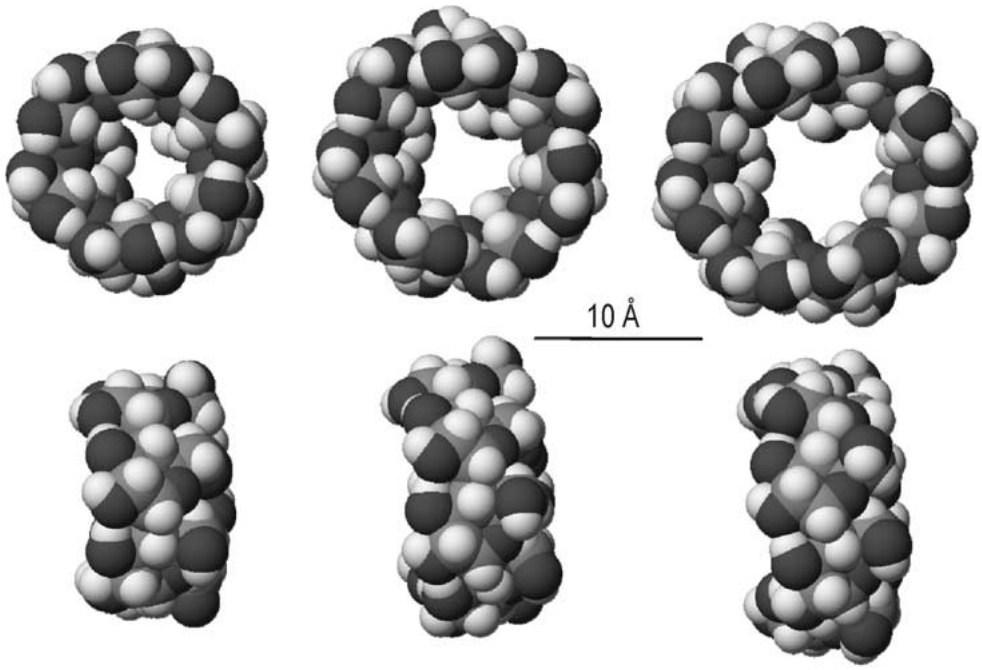

Figure 2. CPK structures showing fully AM1-optimized geometries for natural α-, β- and γ-cyclodextrins containing α-1,4-linked 6, 7 and 8 glucose units, respectively. These structures tend to have a truncated cone shape with the secondary hydroxyls along the wider edge (front and left side, respectively) and the primary hydroxyls along the narrower edge.

This can be easily converted to standard free energies ($\Delta G^0 = -RT_0 \ln K$) by multiplying with $-RT_0 (= -2.479$ kJ/mol):

$$\Delta G^0(\text{kJ/mol}) = -2.48 \ln\phi - 0.849 - 0.207 v(\text{Å}^3). \quad (3)$$

This entirely model-derived slope of –0.20 is in excellent agreement with experimental results. For β-CD, where most data are available, the average slope for all sufficiently small molecules ($v < 120$ Å^3) is –0.192 (Figure 1). For n-alcohols, which because of their additional flexibility and linearity tend to give a larger slope, the slopes are around –0.22 both for α- (up to n-hexanol) and β-CD. This slope gives a general size-dependence, but for comparison with another, often used way of presenting size-dependence, which works only for congener series, it can be converted to contribution per methylene units. As a methylene (CH_2) unit increases v almost exactly with 14 Å^3, eq. (3) corresponds to a decrease of about 2.8 kJ/mol per methylene unit in ΔG^0, the very same average value suggested by Rekharsky and Inoue in their revue [1].

The value of the intercept is less certain, but a value of around 5–6 seems to agree with the experimental data. If a slope of –0.2 is used, best fit (i.e. minimum sum of squared errors) for the basic model assumed here (ΔG^0 decreases linearly for $v \leq v_{\text{lim}}$ and is constant for $v > v_{\text{lim}}$, as shown in Figure 1) is obtained with an intercept of 6.48 and a limiting v of 121 Å^3 (corresponding to ΔG^0 of –17.7). If the restriction on slope is also relaxed, best fit is obtained with a slightly smaller slope (–0.176) and an intercept of 4.40 and a limiting v of 127 Å^3 (corresponding to ΔG^0 of –17.8). It is reassuring that the value of the slope obtained by fitting to experimental data (–0.176) is very close to that obtained from the model (–0.207, eq. (6)). Based on eq. (3), an intercept of 5–6 corresponds to a ϕ value of around 0.06–0.08, indicating that CD-bound guest molecules are restricted to about 6–8% of the translational space associated with the corresponding CD-phase, a very reasonable assumption.

Therefore, with the present model not only the size-dependence (slope) of complex stability, but also the corresponding intercept can be obtained based on very reasonable assumptions. It is a remarkable, never before achieved result that the very same ω interaction constants derived from boiling point and enthalpy of vaporization data can describe partition, and now complex stability data as well. While it was long obvious that all such intermolecular interaction-related physicochemical properties (e.g. T_b, H_{vap}, log γ, log ρ^{gas}, σ, log P, log ρ_w, log K_{CD}) have to be connected, it is for the first time that they could be placed within the framework of a unified and quantitative model.

Experimental complexation data for β-CD (Figure 1), where by far the most data are available ($n = 310$), confirm this general idea. The free energy decreases linearly with size up to about 120–130 Å^3, and more or less levels off after this. The average value for molecules with $v > 120$ Å^3 is –17.7 kJ/mol. This very simplified model that is represented by the dashed line in Figure 1 accounts here for more than 50% of the variance ($r^2 = 0.52$ between experimental and "predicted" values).

Only about half as much experimental complex-stability data are available for α-CD ($n = 179$), and they also agree to a good extent with these prediction (Figure 1). Existence of a limiting v value is less obvious here, it could be somewhere around 90–100 Å^3. Sufficiently flexible molecules can still

give a good fit even for relatively large sizes, whereas rigid molecules might not give a good fit at considerably smaller sizes. The average for molecules larger than 90–110 Å^3 is quite stable and is about –12.7 kJ/mol. The slope of the linear section of all data is somewhat less certain than for β-CD, but it seems to have a very similar, maybe slightly larger value. For *n*-alcohols up to hexanol ($v = 99.5$ Å^3), the slope is the same as for β-CD. As Figure 1 shows, the same equation used for β-CD complexation ($6 - 0.2v$) also gives a good fit here for the α-CD data if it is assumed to reach a limiting value at a smaller size ($v = 95$ Å^3). However, the overall fit is considerably poorer than was for β-CD: the model here accounts for only about 25% of the variability.

Much less ($n = 51$) data are available for γ-CD, and as most published values are for drugs, the paucity of data is especially evident on simple structures. Nevertheless, the limiting size seems much larger; it may be even around 250–300 Å^3, but it is difficult to tell based on available data (Figure 1). It is also obvious that for molecules smaller than this limit, the slope is smaller, probably somewhere around one-third of the value of –0.20 used for β-CD. This can be easily rationalized, because γ-CD has a much larger inside cavity, and contrary to α- and β-CD, for most molecules, here water is not entirely excluded from the cavity during complexation. Hence, the guest will not be surrounded entirely by a CD "environment", and the average favorable interaction will be less pronounced here ($\omega_{avg} < 0.082$). Also, as the fit is much less specific (especially for smaller molecules), the position of γ-CD-bound molecules is much less restricted, and the intercept might be considerably less positive. For example, the equation $1 - 0.08v$ for $v < 250$ corresponds to reasonable values of $\omega = 0.5$ and $\omega_{avg} = 0.03$ and is in good agreement with the experimental data (Figure 1).

The limiting size values obtained here by fitting the model (95, 120, and 260 Å^3 for α-, β-, and γ-CD, respectively) certainly represent only approximate values, and actual values also depend on the atomic van der Waals radii used. Nevertheless, they are not unreasonable. For example, the number of water molecules per CD-cavity estimated based on crystallization data are 6, 11, and 17 for α-, β-, and γ-CD, respectively [11]. With the v_w value used here for water (14.6 Å^3), this suggest cavity sizes of around 88, 161, and 248 Å^3 for α-, β-, and γ-CD, respectively, which are in reasonable agreement with the values obtained here (except of a somewhat larger value for β-CD).

Other descriptors

Size alone cannot account for a considerable amount of variance in the complex stability data, and this is especially true for structures larger than the limiting size requirement of the corresponding CD cavity. Shape or goodness of fit within the cavity is also a major determining factor. For example, the rigid and symmetric adamantyl moiety seems to give an excellent fit within the β-CD cavity and, hence, result in very stable complexes. Di- or tri-substitution of benzene moieties in different positions results in different stabilities, even if size is not affected. The presence of certain functional moieties (e.g. phenol groups) may also result in specific interactions.

Therefore, a number of structure-derived decriptors were analyzed in a linear regression model using the unexplained residual of the present model. In addition to size and calculated log octanol-water partition coefficient (QLogP), variables included the number of rotatable bonds, the number of donor, acceptor, and total hydrogen bonds using a classical additive scale, indicator variables (I, 1 or 0) for the presence or absence of a total of more than 120 functional moieties (e.g. –OH, Ar–OH, –NH_2, –NH–, Ar–SO_2NH_2, etc.), indicator variables for the presence or absence of structural elements, such as –CH_3 groups or three- to eight-membered aliphatic, aromatic-, or heterocyclic-rings.

For β-CD, the model-derived prediction (Figure 1)

$$\begin{aligned}\Delta G^0_{\text{model},\beta} &= 6 - 0.2v \text{ for } v < 120\text{Å}^3 \text{ and} \\ &\quad -17.7 \text{ for } v \geq 120\text{Å}^3\end{aligned} \tag{4}$$

results in:

$$\Delta G^0_{\text{exp},\beta} = 0.440 + 1.014\Delta G^0_{\text{model},\beta} \tag{5}$$

$$n = 310;\ r^2 = 0.515;\ \sigma = 3.58;\ F = 327.1.$$

Out of the additional structure-derived descriptors, QLogP, was by far the most relevant ($r^2 = 0.22$ with the residual for β-CD) indicating that overall lipophilicity/hydrophobicity further enhances complex stability.

$$\Delta G^0_{\text{exp},\beta} = -0.740 + 0.702\Delta G^0_{\text{model},\beta} - 1.517\text{QLogP} \tag{6}$$

$$n = 310;\ r^2 = 0.653;\ \sigma = 3.032;\ F = 289.1.$$

Even if inclusion of a general lipophilicity/hydrophobicity term seems justifiable and improves the correlation, inclusion of this additional descriptor already worsens the F-statistics. A further warning sign against inclusion of a general lipophilicity term is that, contrary to β-CD, inclusion of QLogP provides essentially no improvement in the correlation for α-CD. A number of other investigated descriptors have statistically significant coefficients, and their inclusion improves somewhat the descriptive power of the model (r^2, σ), but the F-statistics is further worsened. Because our main focus here was to include cyclodextrin complexation within the general framework of the size-based model and not to obtain accurate prediction of stabilities, these regressions will not be further explored here. Nevertheless, for illustrative purposes one more equation with indicator variables for the presence of –CH< groups, eight-membered rings, and phenol functionalities is included:

$$\begin{aligned}\Delta G^0_{\text{exp},\beta} &= -0.203 + 0.670\Delta G^0_{\text{model},\beta} - 1.554\text{QLogP} \\ &\quad -2.545 I_{-\text{CH}<} - 6.159 I_{\text{Ring8}} \\ &\quad -1.916 I_{\text{Ar}-\text{OH}}\end{aligned} \tag{7}$$

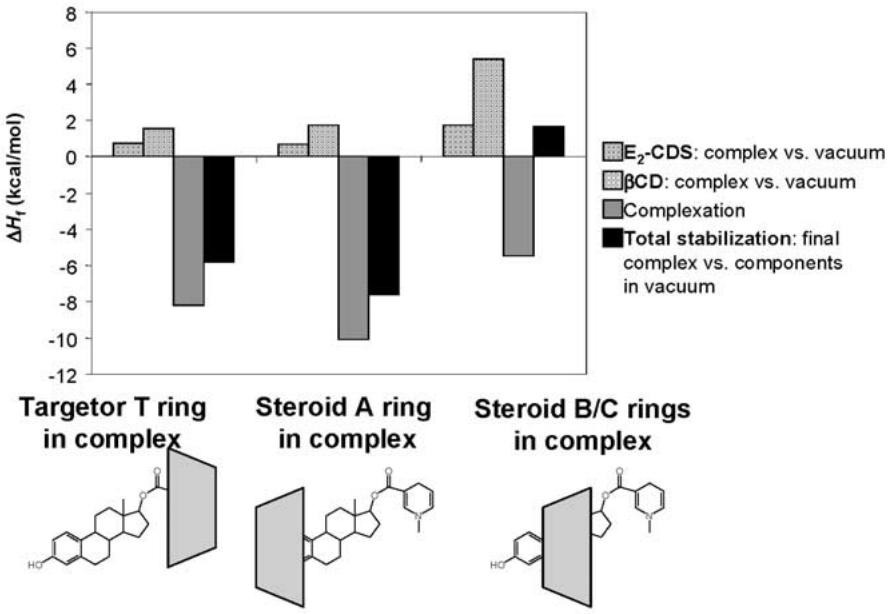

Figure 3. AM1-calculated energies for possible E_2-CDS–β-CD complexes.

$$n = 310;\ r^2 = 0.731;\ \sigma = 2.686;\ F = 164.8.$$

This brief exploration only confirms that complexation is mainly determined by nonspecific interactions, which are to a good extent well accounted for by the size-based liquid model, and to some degree by overall lipophilicity/hydrophobicity and by the presence of specific interactions such as those by a phenol group. The shape of the included (sub)moiety is also relevant as it determines the goodness of fit within the CD-cavity.

Estradiol – βCD complexation

Because of our continuous interest in developing soft drugs and chemical delivery systems [12–14] as well as in the theoretical aspects of CD complex formation [15–17], the complexation between β-cyclodextrin and certain large steroidal guest molecules, especially a brain-targeted estradiol chemical delivery systems (E_2-CDS) that is under clinical development, was investigated in details. For such a large guest, complexation is possible with various subunits. It is now possible to perform fully relaxed semiempirical AM1 quantum chemical calculations [18] on such systems, and Figure 3 presents the results of our calculations. Inclusion of the steroid A ring or of the N-methyl substituted dihydropyridine targetor (T) moiety puts relatively little structural strain on the CD ring or on the guest structure, and this is more than compensated by the gain from the complexation itself. On the other hand, full insertion (B+C rings) is energetically unfavorable. All our AM1-based calculations indicated that inclusion of the steroid ring is more favorable than inclusion of the N-substituted targetor ring. Based on these calculations and on earlier experimental data [19], it is therefore suggested that 1 : 1 complexes are first formed with inclusion of the A ring, and at higher CD-concentrations, 1 : 2 complexes with inclusion at both the A and T ring are formed.

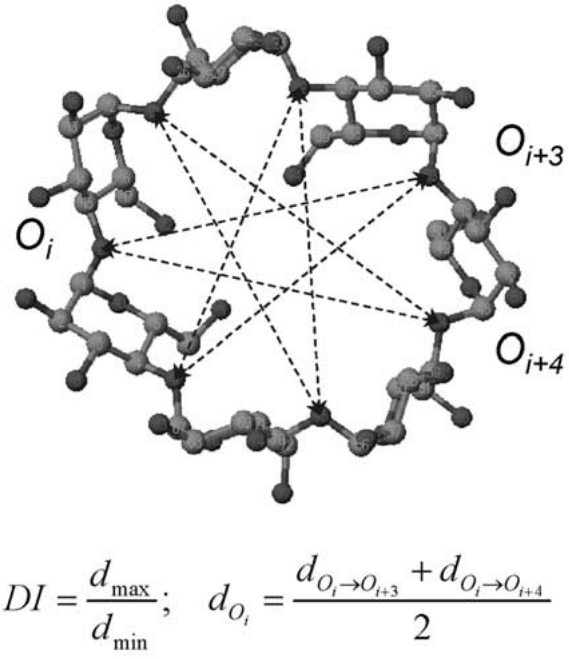

Figure 4. Definition of the deformation index (DI) for a β-CD ring based on diagonal distances between the glucose-linking oxygens (O_i).

To characterize the deformation of the host CD molecule, we introduced a deformation index (DI) computed on these fully optimized host-guest geometries. To better accommodate larger, rigid guest molecules, the CD ring has to undergo deformation (has to deviate from circular symmetry). A possible measure for this is a deformation index defined as the ratio between the largest and smallest diagonal, as illustrated in Figure 4 for a seven-glucose containing β-CD.

The larger the DI value, the larger the deformation (the flatter the ring); for a perfect circle (cylinder), DI = 1. Diagonal distances measured between glucose-linking oxygens were used, as these are more likely to represent the conformation of the CD ring itself and are less dependent on the orientation of the glucose subunits within the CD ring. Also, because in the seven-subunit β-CD ring there are two such opposite oxygens, the average of the corresponding two values was used (Figure 4). Variations in DI are not very large, and even if they indicate deviation from circularity, they are not necessarily related to structural strain. Our fully AM1-optimized structure is the most symmetric having DI of 1.054, a crystal structure (β-CD·11H_2O) determined by neutron diffraction [20] has DI of 1.060. When complexed with E_2-CDS, DI is 1.098 if the steroid A ring and 1.132 if the targetor T ring is included in the cavity, which, in agreement with the energy calculations, also indicates that inclusion of the A ring causes less deformation than inclusion of the N-methyl-substituted T ring.

References

1. M. V. Rekharsky and Y. Inoue: *Chem. Rev.* **98**, 1875 (1998).
2. P. Buchwald and N. Bodor: *J. Phys. Chem. B* **102**, 5715 (1998).
3. P. Buchwald: *Perspect. Drug Disc. Des.* **19**, 19 (2000).
4. P. Buchwald and N. Bodor: *J. Am. Chem. Soc.* **122**, 10671 (2000).
5. G. Castronuovo, V. Elia, A. Iannone, M. Niccoli and F. Velleca: *Carbohydr. Res.* **325**, 278 (2000).
6. C. T. Klein, D. Polheim, H. Viernstein and P. Wolschann: *J. Inclusion Phenom. Macro. Chem.* **36**, 409 (2000).
7. C. T. Klein, D. Polheim, H. Viernstein and P. Wolschann: *Pharm. Res.* **17**, 358 (2000).
8. T. Suzuki: *J. Chem. Inf. Comput. Sci.* **41**, 1266 (2001).
9. N. Muller: *J. Solution Chem.* **17**, 661 (1988).
10. N. Muller: *Acc. Chem. Res.* **23**, 23 (1990).
11. M. E. Brewster and T. Loftsson: *Pharmazie* **57**, 94 (2002).
12. N. Bodor and P. Buchwald: . *Pharmacol. Ther.* **76**, 1 (1997).
13. N. Bodor and P. Buchwald: *Adv. Drug Deliv. Rev.* **36**, 229 (1999).
14. N. Bodor and P. Buchwald: *Med. Res. Rev.* **20**, 58 (2000).
15. N. Bodor: in: *Proceedings of the 7th International Cyclodextrins Symposium*, April 25–28, 1994, Tokyo, Japan, p. 3 (1994).
16. N. Bodor, P. Buchwald and M.-J. Huang: in J. Leszczynski (ed), *Computational Molecular Biology, Theoretical and Computational Chemistry*, Vol. 8, Elsevier, Amsterdam, p. 569 (1999).
17. P. Buchwald: *J. Phys. Chem. B.* **106**, 6864 (2002).
18. M.J.S. Dewar, E.G. Zoebisch, E.F. Healy and J.J.P. Stewart: *J. Am. Chem. Soc.* **107**, 3902 (1985).
19. M.E. Brewster, *K*. E. Estes, T. Loftsson, R. Perchalski, H. Derendorf, G. Mullersman and N. Bodor: *J. Pharm. Sci.* **77**, 981 (1988).
20. C. Betzel, W. Saenger, B.E. Hingerty and G.M. Brown: *J. Am. Chem. Soc.* **106**, 7545 (1984).