Doctoral Dissertation, 2014

Theoretical Study on the Intermolecular Interactions in Complexes of Cyclodextrins with Bile Acids and Bile Salts

OCHANOMIZU UNIVERSITY

Department of Chemistry and Biochemistry
Division of Advanced Sciences
Graduate School of Humanities and Sciences

YAO Lan

March, 2014

Table of Contents

Chapter 1

Introduction

1.1 Host-guest chemistry

Host-guest chemistry belongs to supramolecular chemistry, which involves molecular assemblies where no covalent bonds are established between the interacting species: i.e., molecules, ions, or radicals.[1-3] Host-guest chemistry underlies molecular recognition which is based on non-covalent interactions. Non-covalent bonding is crucial in maintaining three-dimensional structures of large molecules such as proteins, and it is involved in many biological processes in which large molecules bind specifically but transiently to one another. There are four commonly mentioned types of non-covalent interactions: hydrogen bonding, ionic bonding, van der Waals forces (dispersion forces), and hydrophobic interactions.[2] The host component is a molecule or ion whose binding sites converge in the complex, while the guest component is any molecule or ion whose binding sites diverge in the complex.[3] Host-guest chemistry involves the design, synthesis, and investigation of simpler model compounds that imitate working features of naturally occurring compounds, which is important for mimicking and understanding of biological systems. Application of host-guest chemistry includes catalysis, scavenging, sensors, pharmaceuticals, and so forth. Design of new hosts will open more fields for research.[4]

The subsequent two sections, 1.2 and 1.3, describe characteristic features of cyclodextrins and bile acids, which are the target host and guest molecules, respectively, in the present work.

1.2 Cyclodextrin (CD)

Cyclodextrins (CDs) are one of the most important host families. Utilization of CDs as host has the following advantages: (1) CDs are semi-natural products produced from a renewable natural material; thus industrial processes involving CDs are environmentally friendly technology; (2) The prices of CDs are economical and acceptable for most industrial purposes; (3) By formation of inclusion complexes, important properties of the complexed substances can be modified significantly. This unprecedented "molecular encapsulation" is widely utilized in a variety of industrial technologies and analytical methods; (4) Any of their toxic effect is of secondary character and can be eliminated by selecting the appropriate CD type or derivative or mode of application; and (5) As a result of point 4, CDs can be consumed by humans.

1.2.1 Discovery and development of CD

Cyclodextrin was first discovered by Villiers[5] in 1891. He isolated crystalline substance from digesting starch with Bacillus amylobacter, and named this product “cellulosine”. Villiers also observed that two distinct crystalline “cellulosines” were formed, probably α- and β-CDs. During 1903 to 1905, Schardinger,[6] who studied various isolated strains of bacteria and obtained the “cellulosines” from digesting starch with bacillus macerans. He developed an approach to distinguish between the α- and β-dextrins which is cellulosines-iodine complexes reaction. It can be said that the fundamentals of cyclodextrin chemistry were laid down by Schardinger. From the 1930s to the 1970s, Karrer[7] and Miekeley[8] concluded that the crystalline Schardinger's dextrins are built from maltose units and contain only α-1,4-glycosidic linkages. Freudenberg and co-workers postulated the cyclic structure of crystalline dextrins. [9] In late 1940s, the γ-CD had been discovered.[10] French et al.[11] and Cramer et al.[12] contributed to the work on enzymatic production of CDs. After 1970s, CDs were attracted more attention by scientists; over 15000 papers, patens, and conference abstracts were published until 1997. [13]

1.2.2 Structures and important properties of CD

Cyclodextrins (CDs) are cyclic oligosaccharides consisting of α-1,4-linked D-glucose units. The most common natural CDs are α-, β-, and γ-CDs, which are composed of six, seven, and eight glucose units, respectively. Topologically, CDs exhibit a truncated cone structure with a hydrophilic exterior and a hydrophobic interior, thereby rendering its cavity very suitable for inclusion of hydrophobic molecules (Figure 1.1). CDs are often described as a truncated cone with a polar outside and a non-polar cavity.[14] Owing to such topology and chemical properties, CDs can form inclusion complexes with various guest species.

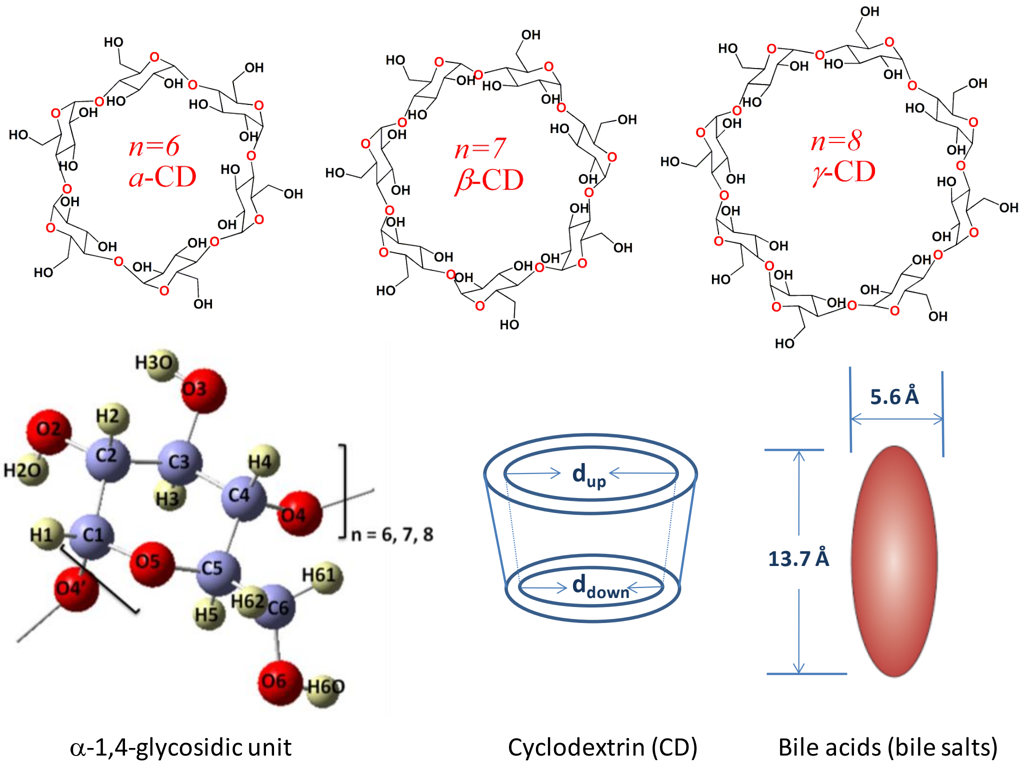


Figure 1.1 Sketch maps of α-, β-, γ-CD molecules, α-1,4 glycosidic unit, and the size of guest molecule.

Table 1.1 Inner diameters (Å) of α-, β-, and γ-CDs in open and closed conformations in the gas phase obtained from semi-empirical (PM3) calculations.

Host	**Conformer**	$\mathbf{d_{up}}$[a]	$\mathbf{d_{down}}$[b]
α-CD	Open[c]	7.0	6.7
	Closed[c]	8.1	5.0
β-CD[d]	Open	8.6	8.1
	Closed	9.8	5.8
γ-CD	Open	10.4	9.7
	Closed	11.8	6.6

[a] The inner diameter at the wide rim d_{up} is obtained from the distance between H3 in one α-1,4 glycosidic unit and that in the opposite unit for both the open and closed conformers. [b] For the open conformer, the inner diameter at the narrow rim d_{down} is obtained from the distance between H61 in one α-1,4 glycosidic unit and that in the opposite unit; for the closed conformer, it is obtained from the distance between H6O in one α-1,4 glycosidic unit and that in the opposite unit. [c] "Open" and "closed" denotes the open and closed conformations of CD which will be explained in Chapter 3. [d] The inner diameter of β-CD is shown in supporting information.

The structural features of CDs will be discussed in Chapter 3 in detail. Only a brief description is given here. There are two stable conformers of CD, open and closed conformations. The cavity sizes of α-, β-, and γ-CD molecules in open and closed conformations are listed in Table 1.1. The inner diameter of α-CD in the closed conformation (5.9 Å) is too small to introduce a bile acid inside. In the open conformation, the wide rim and the narrow rim are larger and smaller, respectively, than those in the closed conformation.

1.2.3 Application of CD

Cyclodextrins are useful molecular chelating agents. It is reported that CD can modify the chemical reactivity, improve the solubility, protect against degradation of guest molecules, etc. CDs have many applications; for instance, CDs are used as carriers for oral drug formulations because they can improve the stability and solubility of some drugs, enhancing their chemical and biological activities through the formation of inclusion complexes.[16] CD and their derivatives have been applied in different areas of drug delivery to reduce the toxicity of other enhancers without affecting their absorption-enhancing property.[17] β-CD or methylated β-CD reduced the serious nasal toxicity of sodium deoxycholate (DCA^-).[18]

1.3 Bile acid and bile salt

Bile salts play an important role in the digestive process, acting as biological surfactants allowing the dissolution and absorption of lipids by the body. Bile acids and bile salts have a characteristic steroid structure, with a side chain at C-17, methyl groups at C-10, C-13, and C-20, and a carboxyl group at C-23, see Chart 1.1. There are two kinds of primary bile acids in our human body, cholic acid (CA) and chenodeoxycholic acid (CDCA), which are produced in the liver and stored in gallbladder. When the primary bile acids are secreted into the lumen of the intestine, intestinal bacteria dehydroxylate a portion of each of them to form the secondary bile acids, deoxycholic acid (DCA) and lithocholic acid (LCA): CA becomes DCA, while CDCA becomes LCA as shown in Figure 1.2.

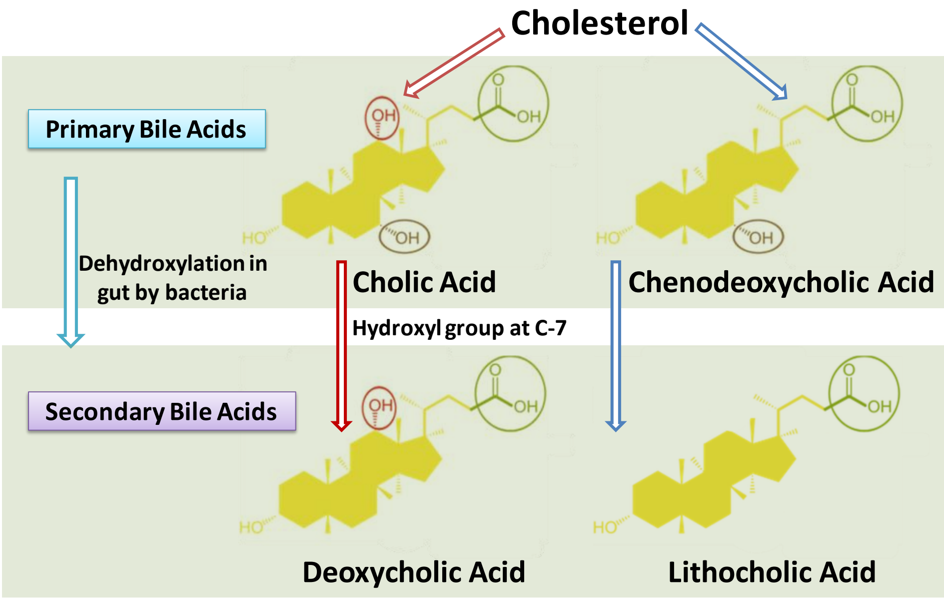


Figure 1.2 Synthetic routes of four bile acids.

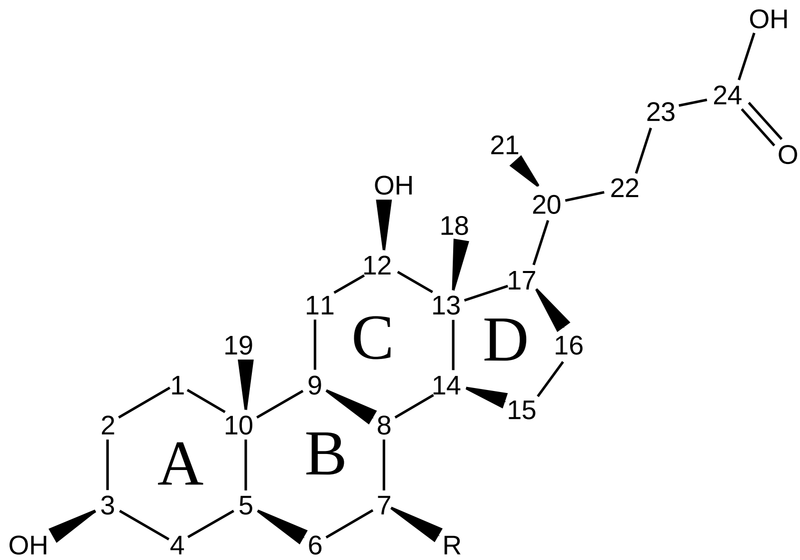


Chart 1.1 Atom numbering for bile acids. (R = OH (CA); R = H (DCA))

1.4 CD-based complex

α-, β-, or γ-CD and their derivatives form 1:1 or 2:1 (host:guest) inclusion complexes with bile salts such as cholate (CA^{-}) and deoxycholate (DCA^{-}) in aqueous solution.[19] Such CD/bile salt complexes have attracted increasing attention from a number of viewpoints. The metabolic pathways of bile salts may be altered by interactions with CDs, as suggested by bile-salt levels in animals fed with a cyclodextrin-enriched diet.[20] In some drug delivery systems, the interaction of the β-CD/drug complex with bile salts enhances the bioavailability of the drug.[21] CD/bile salt complexes, which are soluble in water, serve as models of CD/cholesterol complexes. Although cholesterol can bind to CD or its derivatives,[22] the low solubilities of cholesterol itself and the complexes in water prevent the structure determination of the complexes. Combinations of CDs and bile salts are also used in chromatographic and electrophoretic separation of enantiomers.

The stoichiometry and binding modes of CD/bile salt complexes have been determined mainly from NMR experiments in solution phase.[19] The association constants have also been determined for the equilibrium schematically shown in Figure 1.3. Although relative positions and orientations of the host and guest have been proposed, neither information on the detailed host–guest interactions nor crystal structures for CD/bile salt complexes are available. To understand the properties of CD/bile salt (or bile acid) complexes and to explore potential applications, it is essential to clarify the structures and intermolecular interactions of these complexes.

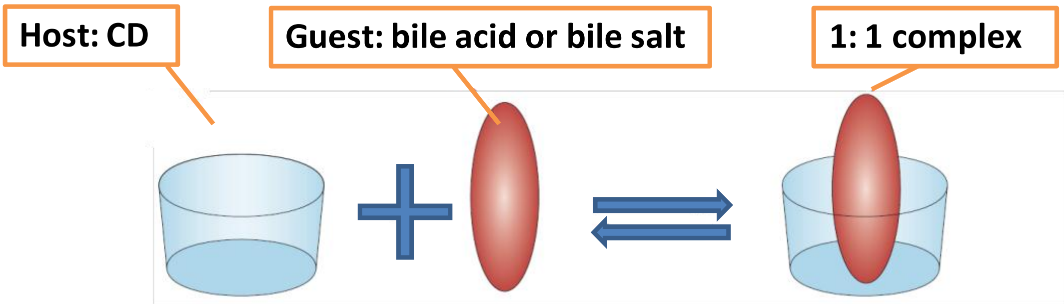


Figure 1.3 Schematic diagram for reversible association of CD and guest molecule.

1.5 Preceding theoretical studies on CDs and their complexes

A number of theoretical papers have been reported for CDs and their inclusion complexes with various guest molecules or ions such as water, ammonium, benzene derivatives, amino acids, cholesterol, and PCBs, typical persistent organic pollutants.[22-28] Those works use either molecular orbital (MO) calculations or molecular dynamics (MD) techniques to investigate the equilibrium structures, interaction energies, free energies of binding, and potential of mean force for the complexes.

In most of the previous MO calculations, semi-empirical (PM3, AM1,[23] or PM6[24]) or DFT (B3LYP)[23,25,26] levels are used, whereas MD simulations are based on parameterized force fields.[27-29] The computational results reproduce the experimentally determined structures of CDs rather well. However, validation on the structures of inclusion complexes is quite limited.[29] In some cases, experimental determination of the structures is difficult; for example, the low solubility of cholesterol−β-CD complex makes recrystallization or NMR study difficult. Comparisons of thermodynamic quantities between the calculations and experiments are also scanty. The accuracy of the computational results should be evaluated based on comparison of the results with data obtained by experiments or by calculations at higher levels of theory. In particular, accuracy of the DFT calculation depends on the functional used; it is reported that B3LYP cannot necessarily give proper description on dispersion interactions.[30]

In some of the MO-based studies, solvation effects are taken into consideration by using implicit solvation models such as COSMO.[25] In order to understand properties and intermolecular interactions of the CD-guest complexes, explicit solvation models may give better description on solvation effects of aqueous solvent. On the other hand, the MD simulations have been performed for models which include hundreds or more water molecules. The classical MD technique, however, cannot describe electronic polarization induced by intermolecular interactions. Although a few recent studies[26] utilize both MO and MD methods in order to survey plausible binding modes and to investigate driving forces in association processes, these two techniques are used independently without close linking of the available information obtained by each method. For understanding host-guest interactions which play important role in molecular recognition and modification of chemical properties, it is desirable to take advantages of these two methodologies via combination of computational results in sophisticated manner as well as use of models suitable for description of aqueous environment.

1.6 Objectives of this work

The target system of the present study is CD/bile acid (or bile salt) inclusion complexes. The significance and available experimental data for the system have been described in section 1.4. The main objectives of this work are: (1) to predict stable binding modes of the CD/bile acid complexes; (2) to analyze contributions to the stabilization of the complexes from different types of intermolecular interactions, hydrogen-bonding, dispersion forces, and others; (3) to compare host–guest interactions among different binding modes and among different bile acids; and (4) to clarify solvation effect of water on the structural features and the host−guest interactions of the complexes.

To this end, theoretical investigation is performed for 1:1 inclusion complexes of α-, β-, and γ-CDs with bile acids or bile salts. The methodologies used are density functional theory (DFT)-based or ab initio MO calculations and molecular dynamics (MD) simulation. The host–guest interactions are characterized at the ab initio MO calculations under influence of thermal motion and aqueous solvent by combining high-level MO methods and classical MD techniques. The obtained structures of the complexes are validated by comparing them with those proposed from the NMR experiments in aqueous solution. The differences in the host-guest interactions between in the gas phase and in aqueous phase are clarified. Details of the computational strategy will be described in the following chapter.

References and Notes

1 J. Szejtli, *Chem. Rev.* **1998**, *98*, 1743.

2 H. Lodish, A. Berk, C. Kaiser, *Molecular Cell Biology*, **2008**.

3 J. W. Steed, J. L. Atwood, *Supramolecular Chemistry*, **2000**, John Wiley&Sons.

4 E. M. Martin Del Valle, *Process Biochemistry*, **2004**. *39*. 1033.

5 A. Villiers, *Compt. Rend*. **1891**, *112*, 536.

6 (a) F. Schardinger, *Z. Unters. Nahr. u. Genussm*. **1903**, *6*, 865; (b) F. Schardinger, *Wien. Klin. Wochenschr*. **1904**, *17*, 207; (c) F. Schardinger, *Zentralbl. Bakteriol. Parasintenk. Abt. 2* **1905**, *14*, 772.

7 P. Karrer, C. Nageli, *Helv. Chim. Acta* **1921**, *4*, 169.

8 A. Miekeley, *Ber. Dtsch. Chem. Ges.* **1932**, *65*, 69.

9 (a) K. Freudenberg, W. Rapp, *Ber. Dtsch. Chem. Ges.* **1936**, *69*, 2041; (b) K. Freudenberg, H. Boppel, M. Meyer-Delius, *Naturwissenschaften* **1938**, *26*, 123; (c) K. Freudenberg, M. Meyes-Delius, *Ber. Dtsch. Chem. Ges.* **1938**, *71*, 1596; (d) K. Freudenberg, G. Blomquist, L. Ewald, K. Soff, *Ber. Dtsch. Chem. Ges.* **1936**, *69*, 1258.

10 K. Freudenberg, F. Cramer, *Z. Naturforsch*. **1948**, *3b*, 464.

11 D. French, *Adv. Carbohydr. Chem*. **1957**, *12*, 189.

12 F. Cramer, *Einschlussverbindungen (Inclusion Compounds)*, **1954**, Springer-Verlag: Berlin.

13 *Cyclodextrin News* (a monthly abstracting newsletter published since 1986, dedicated exclusively to the cyclodextrin literature, production, marketing, related topics, conferences, etc.), edited and published by Cyclolab Ltd., Budapest.

14 L. Peng, D. J. Zhang, J. H. Zhan, *J. Phys. Chem*. A, **2010**, *114*, 13123.

15 The width and length of the guest molecule are evaluated by the distances of 7-H…11β-H and 3-H…HOOC, respectively, based on the semi-empirical (PM3) calculations in the gas phase. The numbering scheme is shown in Chart 1.1.

16 (a) For example, K. Anzai, H. Kono, J. Mizoguchi, T. Yanagi, F. Hirayama, H.

Arima, K. Uekama, *Carbohydr. Res*. **2006**, *341*, 499; (b) C. R. Dass, W. Jessup, *J. Pharm. Pharmacol.* **2000**, *52*, 731.

17 R. Challa, A. Ahuja, J. Ali, R. K. Khar, *AAAPs Pharm. Sci. Tech.*, **2005**, *6(2)*, E329.

18 Y. Zhang, X. G. Jiang, J. Yao, *Acta Pharmacol. Sin.* **2001**, *22*, 1051.

19 R. Ramos, E. Alvarez-Parrilla, F. Meijide, J. A. Seijas, E. Rodriguez Nunez, J. Vazquez Tato, *Langmuir*, **1999**, *15*, 5489.

20 C. Abadie, M. Hug, C. Kubli, N. Gains, *Biochem. J*. **1994**, *299*, 725; N. Boehler, M. Riottot, J. Férézou, M. Souidi, F. Milliat, C. Sérougne, J. L. Smith, C. Lutton, *J. Lipid Res*. **1999**, *40*, 726.

21 K. Nakanishi, M. Masada, T. Nadai, K. Miyajima, *Chem. Pharm. Bull*. **1989**, *37*, 211.

22 (a) R. Breslow, B. Zhang, *J. Am. Chem. Soc.* **1996**, *118*, 8495; (b) H. Asanuma, M. Kakazu, M. Shibata, T. Hishiya, M. Komiyama, *Chem. Commun*. **1997**, 1971; (c) R. Ravichandran, S. Divakar, *J. Incl. Phenom. Mol. Recognit. Chem*. **1998**, *30*, 253.

23 M. Nagaraju, G. N. Sastry, *J. Phys. Chem. A*, **2009**, *113*, 9533.

24 Y. Xia, X.Y. Wang, Y. Zhang, B. Luo, *Comp. and Theor. Chem.*, **2011**, *967*, 213.

25 S. K. Xing, C. Zhang, H. Q. Ai, Q. Zhao, Q. Zhang, D. Z. Sun, *J. Molecular Liquids*, **2009**, *146*, 15.

26 P. Liu, D.J. Zhang, J. H. Zhan, *J. Phys. Chem. A*, **2010**, *114*, 13122.

27 Y. M. Yu, C. Chipot, W. S. Cai, X. G. Shao, *J. Phys. Chem. B*, **2006**, *110*, 6372.

28 M. Jana, S. Bandypadhyay, *J. Phys. Chem. B*, **2013**, *117*, 9280.

29 G. Raffaini, F. Ganazzoli, L. Malpezzzi, C. Fuganti, G. Fronza, W. Panzeri, A. Mele, *J. Phys. Chem. B*, **2009**, *113*, 9110.

30 J. Van der Mynsbrugge, K. Hemelsoet, M. Vandichel, M. Waroquier, V. Van Speybroeck, *J. Phys. Chem. C*, **2012**, *116*, 5499.

Chapter 2

Overview of Computational Details

2.1 Significance and originality of the present research

Theoretical research on CDs has been somewhat limited until recent years. The reason is that CDs are relatively large, flexible molecules. The size of CDs and their derivatives makes applications of quantum mechanics difficult even when symmetry conditions are imposed since they have many rotatable bonds; there exist an enormous number of conformational states. Moreover, CDs are usually studied experimentally in aqueous phase,[1] an environment that has been a barrier of computational chemistry at quantum level, creating a major hurdle for computational chemists.

In the present study, density functional theory (DFT) is used to obtain the equilibrium structures and the association energies for the CD/bile acid complexes. In general, current DFT methods exhibit relatively high performance at reasonable computational costs on prediction of equilibrium geometries, physical properties, and thermochemistry for a number of molecular systems. It is, however, still difficult for DFT methods to evaluate weak interactions such as van der Waals interaction with sufficiently high accuracy. Inclusion compounds are one of the difficult targets. The key point of success may be selection of the DFT functional; this issue will be discussed in the following section.

One of the aims of the present study is to evaluate the contribution from different types of interactions to the stabilization of the complexes as described in Chapter 1. To this aim, pair interaction energy decomposition analysis (PIEDA) method[2] is utilized at fragment molecular orbital (FMO)-MP2 level.[3] The FMO method is applicable for even larger systems such as proteins, and interactions between fragments can be analyzed. There were few attempts to calculate cyclic sugar molecules by the FMO level before; hence the validation of fragmentation scheme should be addressed, which will be described in section 5.2.

Another issue of the present study is solvation effects by aqueous environment. Structures of the CD/bile acid complexes are compared between in the gas phase and in aqueous phase, the latter being modeled with a polarizable continuum model (PCM) or explicit solvation model at molecular mechanics level. Evaluation of the solvation effects on the complexes will provide deeper insight on the behaviors of the CD/bile acid complexes in aqueous solution.

2.2 Selection of calculation methods

2.2.1 DFT functional

In the present study, two DFT functionals are mainly used; one is a dispersion-corrected density theory functional, B97-D,[4] and the other is M06-2X.[5] The specific features of these functionals are given below:

(1) B97-D belongs to density functionals of the generalized gradient approximation (GGA) type supplemented with an empirical correction term for dispersion interactions. The current version of B97-D is a slightly modified version for dispersion part of the approach described in Grimme's paper.[6] The total energy is given by

$$E_{DFT\text{-}D} = E_{KS\text{-}DFT} + E_{disp} \tag{2-1}$$

where $E_{KS\text{-}DFT}$ is the usual self-consistent Kohn-Sham energy, E_{disp} is an empirical dispersion correction term. It is reported[7] that B97-D exhibits high performance on 40 non-covalently bound complexes, including large stacked aromatic molecules and group II element clusters. The performance for non-covalent bound systems including many pure van der Waals complexes is exceptionally good, almost reaching on the average CCSD(T) accuracy. It is expected that the dispersion force between the host and guest is crucial for formation of the CD/bile acid complexes. Therefore, B97-D seems to be an appropriate choice.

(2) M06-2X functional is parameterized including both transition metals and nonmetals and a high-nonlocality functional with a high proportion (54%) of nonlocal exchange. M06-2X is recommended for applications involving main-group non-covalent interactions by the author.[5] It is also reported M06-2X performed best for the S22 database[8] of biologically importance interactions, which are bound by a mixture of electrostatic and dispersion-like force. The systems which have been studied with M06-2X are similar type as the present system; hence, M06-2X is considered to be another functional suitable to this work.

(3) B3LYP functional,[9] a widely used conventional functional, was also used for comparison with B97-D and M06-2X.

2.2.2 Basis set selection

All the DFT calculations are performed with the 6-31G(d) basis set.[10] This medium size basis set is a reasonable choice considering the balance between the size of basis set and the level of theory as well as the computational efficiency. In FMO calculations described in Chapter 5, the dependence of the interaction energies on the basis set is also examined by comparing the results between 6-31G(d) and cc-pVDZ.[11]

References

1 B. L. Kenny, *Chem. Rev.*, **1998,** *98*, 1829.

2 (a) D. G. Fedorov, K. Kitaura, *J. Comp. Chem.*, **2007**, *28*, 222; (b) D. G. Fedorov, K. Kitaura, *J. Chem. Phys.*, **2012,** *116*, 704.

3 (a) D. G. Fedorov, K. Kitaura, *J. Chem. Phys.*, **2004**, *120*, 6832; (b) D. G. Fedorov, K. Kitaura , *J. Chem. Phys.*, **2004**, *121*, 2483; (c) D. G. Fedorov, K. Kitaura, *J. Phys. Chem. A*, **2007**, *111*, 6904; (d) D. G. Fedorov, T. Nagata, K. Kitaura, *Phys. Chem. Chem. Phys.*, **2012**, *14*, 7562; (d) Y. Mochizuki, K. Yamashita, T. Murase, T. Nakano, K. Fukuzawa, K. Takematsu, H. Watanabe, S. Tanaka, *Chem. Phys. Lett.*, **2008**, *457*, 396.

4 S. Grimme, *J. Comput. Chem.*, **2006,** *27*, 1787.

5 Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.*, **2008**, *120*, 215.

6 S. Grimme, *J. Comput. Chem.*, **2004,** *25*, 1463.

7 (a) S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, *J. Chem. Phys.*, **2006**, *124*, 114304; (b) J. F. Olgivie, F. Y. H. Wang, *J. Mol. Struc.*, **1992**, *273*, 277; (c) J. F. Olgivie, F. Y. H. Wang, *J. Mol. Struc.*, **1993**, *291*, 313; (d) O. Hubner, A. Gloss, M. Fichtner, W. Klopper, *J. Phys. Chem. A*, **2004**, *108*, 3019; (e) J. Grant Hill, J. A. Platts, H.-J. Werner, *Phys. Chem. Chem. Phys.*, **2006**, *8*, 4072; (f) S. Tsuzuki, T. Uchimaru, M. Mikami, *J. Phys. Chem. A*, **2006**, *110*, 2027; (g) S. Tsuzuki, K. Honda, T. Uchimaru, M. Mikami, *J. Chem. Phys.*, **2005**, *122*, 144323.

8 S22 database is a large, diverse, non-covalent database.

9 A. D. Becke, *J. Chem. Phys.*, **1993**, *98*, 5648.

10 (a) R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.*, **1971**, *54* , 724; (b) W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.*, **1972**, *56*, 2257; (c) P. C. Hariharan, J. A. Pople, *Theor. Chem. Acc.*, **1973**, *28*, 213.

11 (a) T. H. Dunning Jr, *J. Chem. Phys.*, **1989,** *90,* 1007; (b) T. van Mourik, T. H. Dunning Jr, *J. Mol. Struct.*, **1996**, *388*, 339.

Chapter 3

DFT study on CD/bile salt complexes: Geometries, association energies, and hydrogen bonds

3.1 Introduction

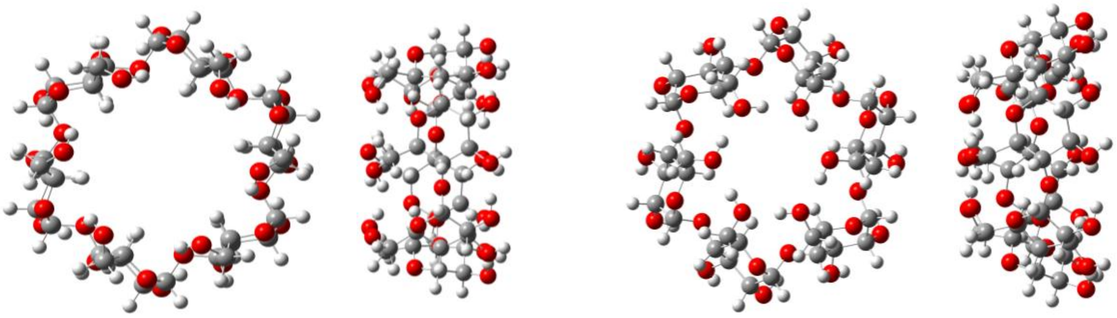
It is reported that the lowest energy structures of cyclodextrins (CDs) in the gas phase have two types of conformations, open and closed, which are shown in Figure 3.1.[1] In the first part of this chapter, the differences in energy and in hydrogen-bond structures between the open and closed conformations are discussed for free α-, β-, γ-CDs in the gas phase and in aqueous solution with DFT calculations.

As mentioned in Chapter 1, cyclodextrins (CDs) are able to encapsulate poorly water-soluble bile acids (e.g. CA and DCA) or bile salts (e.g. CA^- and DCA^-) due to the hydrophilic hydroxyl groups outside and the hydrophobic wall inside the cavity. The sizes of guest molecules and the binding modes should be suitable for the accommodation in the cavities of CDs. The second part of Chapter 3 describes the binding modes and the association energies for 1:1 inclusion complexes of CDs with CA and its related guest species (Chart 3.1) in the gas phase and in aqueous phase. In order to clarify which factor mainly contributes to stabilization of the complexes, intermolecular hydrogen bonds (H-bonds) between the host and guest molecules and geometrical changes of the individual components on complexation are analyzed.

In the last part of this chapter, the theoretical results are compared with experimental data in terms of the binding mode and the association enthalpy.

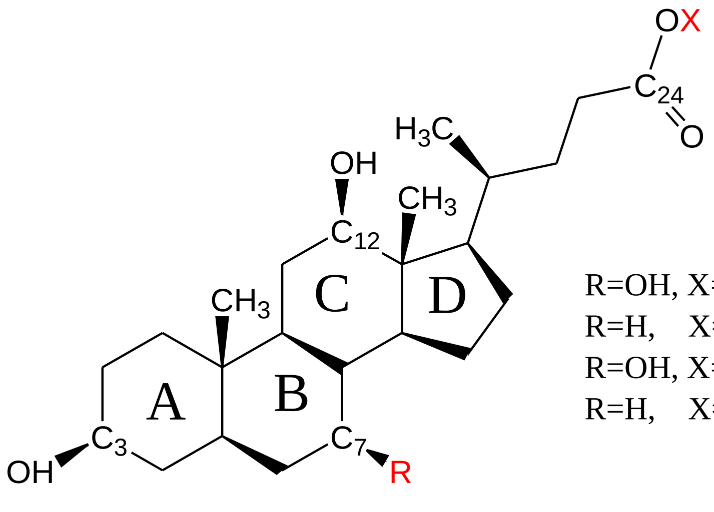
3.2 Computational Details

At first, geometry optimization and harmonic frequency analyses were carried out for free CDs in the open and closed conformations, the free guests in Chart 3.1, and the CD/guest complexes by the semi-empirical PM3 method[2] for rough explorations. Then the geometries were re-optimized by the DFT (B97-D[3], M06-2X[4] and B3LYP[5]) levels of theory with the standard 6-31G(d) basis set[6] both in the gas phase and in aqueous phase. A polarizable continuum model using integral equation formalism (IEF-PCM)[7] was employed for evaluation of solvation free energies. Although continuum solvation models such as PCM and COSMO have been utilized for evaluation of interaction energies in solution phase,[8] it is likely that such models are insufficient for the present system in terms of the accuracy in solvation free energy since hydrophobic hydration, for which the entropic term has major contribution, is difficult to be described with such models. Furthermore, the volume of the CD cavity is so small that solvent cannot be regarded as a continuum even though water is a small molecule. For the CD/guest complexes, however, the volume of the solute is large enough for solvent to be regarded as a continuum. Since the complexes are similar to each other in the hydrophobicity and shape, the errors may be cancelled out when the energies are compared between these complexes. In the present DFT calculations, therefore, PCM was mainly used to estimate relative energies among complexes with different conformations or configurations. A more realistic solvation model will be used in Chapter 4.



(a) Open conformation (b) Closed conformation

Figure 3.1. Top and side views of α-CD in the open (a) and closed (b) conformations. The hydrogen, oxygen, and carbon atoms are drawn in light gray, red, and gray, respectively.



R=OH, X=H: cholic acid (CA)
R=H, X=H: deoxycholic acid (DCA)
R=OH, X=$^-$: cholate (CA^-)
R=H, X=$^-$: deoxycholate (DCA^-)

Chart 3.1. The structural formulas of guest molecules.

The energy differences between the open and closed conformations of CDs were calculated by eq. 1:

$$\Delta E = E_{\text{open-host-opt}} - E_{\text{closed-host-opt}} \tag{1}$$

where "opt" means that the structure was optimized.

For the CD/guest complexes, the basis set superposition error (BSSE) was corrected for by the counterpoise method.[9] For calculations in aqueous phase, the BSSE correction calculated without PCM was applied. The BSSE-corrected association energy, E_{ass}, was evaluated according to eq. 2a or 2b for the open or closed conformation, respectively:

$$E_{\text{ass}} = E_{\text{open-complex-opt}} - (E_{\text{free-open-host-opt}} + E_{\text{free-guest-opt}}) + \text{BSSE} \tag{2a}$$

$$E_{\text{ass}} = E_{\text{closed-complex-opt}} - (E_{\text{free-closed-host-opt}} + E_{\text{free-guest-opt}}) + \text{BSSE} \tag{2b}$$

where $E_{\text{free-open-host-opt}}$, $E_{\text{free-closed-host-opt}}$, and $E_{\text{free-guest-opt}}$ are the energies of the host and the guest at the optimized geometry in free state.

Association enthalpies, H_{ass}, were calculated with the zero-point energy and thermal corrections at 298.15 K and 1 atm according to eq. 3.

$$H_{\text{ass}} = H_{\text{open-complex-opt}} - (H_{\text{free-open-host-opt}} + H_{\text{free-guest-opt}}) \tag{3a}$$

$$H = E_{\text{elec}} + \text{ZPE} + E_{\text{vib}} + E_{\text{rot}} + E_{\text{trans}} + RT \tag{3b}$$

where $H = E_{\text{elec}}$, ZPE, E_{vib}, E_{rot}, and E_{trans} represent the electronic energy (with BSSE-correction for the complex), zero-point energy, vibrational energy, rotational energy, and translational energy, respectively. The zero-point energies were calculated based on the harmonic frequency analyses.

Deformation energies, E_{deform}, which are energy changes accompanied with structural changes of the host and guest upon complexation, were evaluated by eq. 4.

$$E_{\text{deform}} = (E_{\text{bound-host}} - E_{\text{free-host-opt}}) + (E_{\text{bound-guest}} - E_{\text{free-guest-opt}}) \tag{4}$$

where $E_{\text{bound-host}}$ and $E_{\text{bound-guest}}$ are the energies of the host and the guest, respectively, at the geometries in the complex.

Hypothetical interaction energies, $E_{intrxn,}$ are defined as eq. 5:

$$E_{intrxn} = E_{ass} - E_{deform} \quad (5)$$

All the calculations in this chapter were performed by Gaussian 09 program package.[10]

3.3 Results and Discussion

3.3.1 Structures and relative energies of Free CDs

Relative energies of the closed and open conformations of free CDs

According to the previous calculations at the B3LYP/6-31G(d,p) level,[1] the lowest-energy conformation of α-, β-, and γ-CDs in the gas phase is a one-gate-closed shape as shown in Figure 3.1(b) and referred to as "closed conformation." Another local minimum, called "open conformation," is similar to the experimentally observed structures under hydrated conditions (Figure 3.1(a)).

The energy differences between the open and closed conformations (ΔE) for α-, β-, and γ-CDs were calculated with different levels of theory and listed in Table 3.1. All the calculation methods predict that the closed conformation is more stable than the open one in the gas phase for all the CDs investigated. The HF and DFT calculations show that the energy difference decreases with an increasing number of sugar units in agreement with the previous results.[1] In aqueous phase, the calculated ΔE values are smaller than those in the gas phase, indicating that the open conformation is more stabilized by solvation than the closed conformation.

Table 3.1 Energy differences between the open and closed conformations ($\Delta E = E_{open} - E_{closed}$ / kcal mol^{-1}) with different levels of theory with the standard 6-31G(d) basis set.

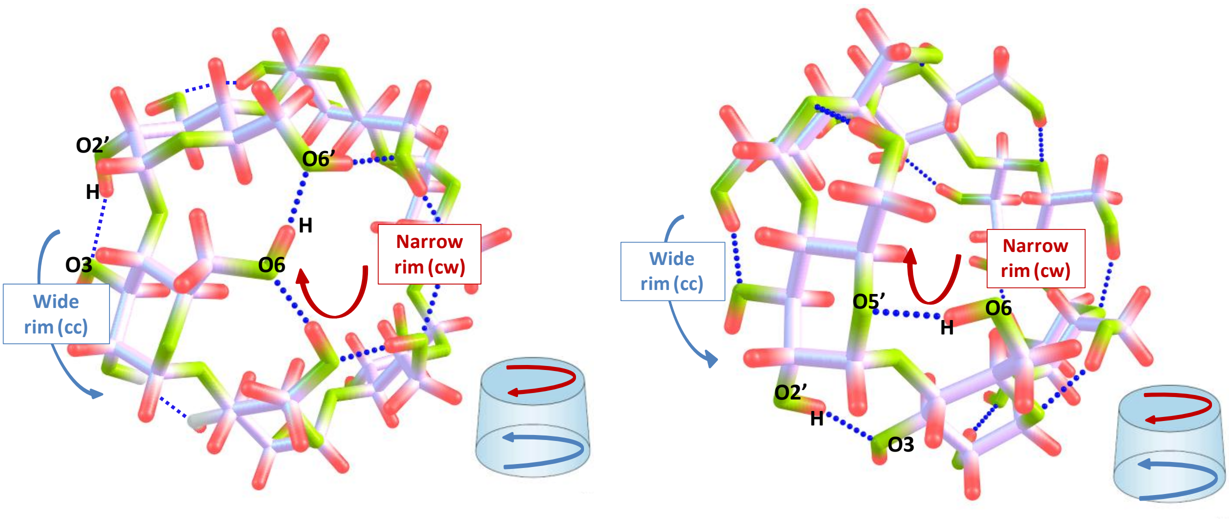
Molecule	**ΔE (HF)**	**ΔE (B97-D)**	**ΔE (M06-2X)**
α-CD	17.9	16.9	19.3 (10.3)[a]
β-CD	15.5	11.9	16.6 (5.1)
γ-CD	11.5	4.7	11.6 (–2.4)

[a] Data in the parentheses are the ΔE values calculated with PCM.

Hydrogen-bond rings in free CDs.

There are two rings of intramolecular hydrogen bond network for α-, β-, or γ-CD either in the open or closed conformation; one is located at the narrow rim and the other at the wide rim. When the orientation of an X…H-Y hydrogen bond is defined as the direction toward atom X, the orientations of the hydrogen bonds are described as clockwise (cw) and counter-clockwise (cc). As shown in Figure 3.2, the hydrogen bonds at the narrow rim (closed gate) are oriented clockwise (cw), whereas those at the wide rim are oriented counter-clockwise (cc) in both the closed and open conformations.

The distances ($r_{OH\ldots O}$) and angles (∠OH…O) of the H-bonds at the geometries optimized in the gas phase and with PCM are given in Table 3.2. In the closed conformation, the hydrogen-bond ring, O6-H…O6' (the primed oxygen atom belongs to the next sugar unit), which consists of the primary hydroxyl groups, involves a rather short O…H distance (1.79 ~ 1.88 Å) while the second one, O3…H-O2'**,** which consists of the secondary hydroxyl groups, involves a longer distance (2.09 ~ 2.15Å). The short hydrogen bonds at the narrow rim make a CD in the closed conformation a "one-gate-closed" shape. On the other hand, the open conformation contains one hydrogen-bond ring of O6-H…O5' type at the side of the primary hydroxyl groups and the second one of O3…H-O2' type, which is the same as observed in the closed conformation.



(a) α-CD in the closed conformation (b) α-CD in the open conformation

Figure 3.2 The orientations of the hydrogen bonds in α-CD in the closed (a) and open (b) conformations. The hydrogen, oxygen, and carbon atoms are drawn in red, violet, and yellow-green, respectively. The blue dotted lines denote H-bonds.

Table 3.2 Structural parameters of H-bonds for α-, β-, and γ-CDs in the gas phase and with PCM. The structures were optimized by the M06-2X[a]/6-31G(d) method.

Molecule	Confor-mation	Phase	Narrow rim (primary OH groups)			Wide rim (secondary OH groups)		
			n^b	$r_{OH...O}$ / Å	∠OH…O/°	n^b	$r_{OH...O}$ / Å	∠OH…O/°
α-CD	open	Gas	6 (cw)	1.88	153.7	6 (cc)	1.89	162.6
		PCM	6 (cw)	1.87	157.1	6 (cc)	1.89	158.3
	closed	Gas	6 (cw)	1.79	172.5	6 (cc)	2.15	160.1
		PCM	6 (cw)	1.80	171.1	6 (cc)	2.13	156.0
β-CD	open	Gas	7 (cw)	1.87	157.3	7 (cc)	1.87	161.1
		PCM	7 (cw)	1.86	161.0	7 (cc)	1.89	156.7
	closed	Gas	7 (cw)	1.82	170.2	7 (cc)	2.11	158.1
		PCM	7 (cw)	1.84	167.4	7 (cc)	2.10	153.9
γ-CD	open	Gas	8 (cw)	1.86	160.9	8 (cc)	1.85	159.9
		PCM	8 (cw)	1.85	163.0	8 (cc)	1.86	155.5
	closed	Gas	8 (cw)	1.86	167.2	8 (cc)	2.11	156.2
		PCM	8 (cw)	1.88	164.4	8 (cc)	2.09	152.3

[a] Not all structures were well converged by B97-D functional with PCM; however, the converged structures show almost the same tendency with that with M06-2X. [b] The number of hydrogen bonds. The orientation of the H-bond ring is shown in the parenthesis. cw and cc denote the clockwise and counter-clockwise, respectively.

The OH…O distance at the narrow rim in the closed conformation of α-CD (1.79/1.80 Å) is much shorter than that in the open conformation (1.88/1.8 Å), and the OH…O angle is close to an ideal value of 180º in the closed conformation. The stabilization due to the hydrogen-bonding is thus more significant in the closed conformation than in the open one, which likely causes a large difference in the energy ($\Delta E = E_{\text{open-}\alpha\text{-CD-opt}} - E_{\text{closed-}\alpha\text{-CD–opt}}$, see Table 3.1). The difference in the OH…O distance between the conformers decreases with an increasing number of sugar units, and the energy difference also decreases in the same order.

In the closed conformation, the OH…O distance (1.79/1.80 Å) at the narrow rim of α-CD is shorter than those at its wide rim (2.15/2.13 Å). β- and γ-CDs exhibit the same tendency, but the differences in the OH…O distance between the narrow and wide rims are smaller.

3.3.2 Structures and association energies of CD/CA complexes.

Survey of stable configurations by semi-empirical and DFT-D levels.

For construction of initial models for the CD/CA complexes, the structure of CD was fixed in the closed or open conformation, and the CA molecule was approached along its central axis toward the wider rim of CD. According to previous studies on a number of 1:1 CD complexes with oblate-shaped guests, [11] two or three energy minima were found when the guest, which was aligned so that its long axis was coincided with the axis of the CD cavity, was moved along this direction with step-by-step scan. In the present study, three different positions of CA (top, mid, and bottom) as well as two directions of insertion (1 and 2) were taken into consideration. "Top," "mid," and "bottom" refer to the position of the main part of CA in the cavity, and direction 1 represents such a direction that the A ring of CA points toward the wider rim of CD. This modeling procedure was applied to CD either in the open or the closed conformation to obtain 10 plausible configurations for each CD/CA complex (Figure 3.3). No models could be prepared for closed/mid2 or closed/bottom2 since the host and guest would be crashed upon docking. Optimization of each of the 10 initial structures with the PM3 method gave a stationary point, which corresponded to a local energy minimum.

The calculated association energies for all the 10 configurations with the PM3 and B97-D methods are presented in Table 3.3. The B97-D/6-31G(d) calculations were carried out by using the PM3-optimized structures. The association energies obtained by PM3 are quite different from those of B97-D. PM3 is a semi-empirical method and generally less accurate than typical DFT methods although it has often been used for large systems owing to lower computational cost.[12] On the basis of the B97-D calculations, eight configurations with most negative E_{ass} values (marked in bold in Table 3.3) were selected for further study. The schematic structures are shown in Figure 3.4.

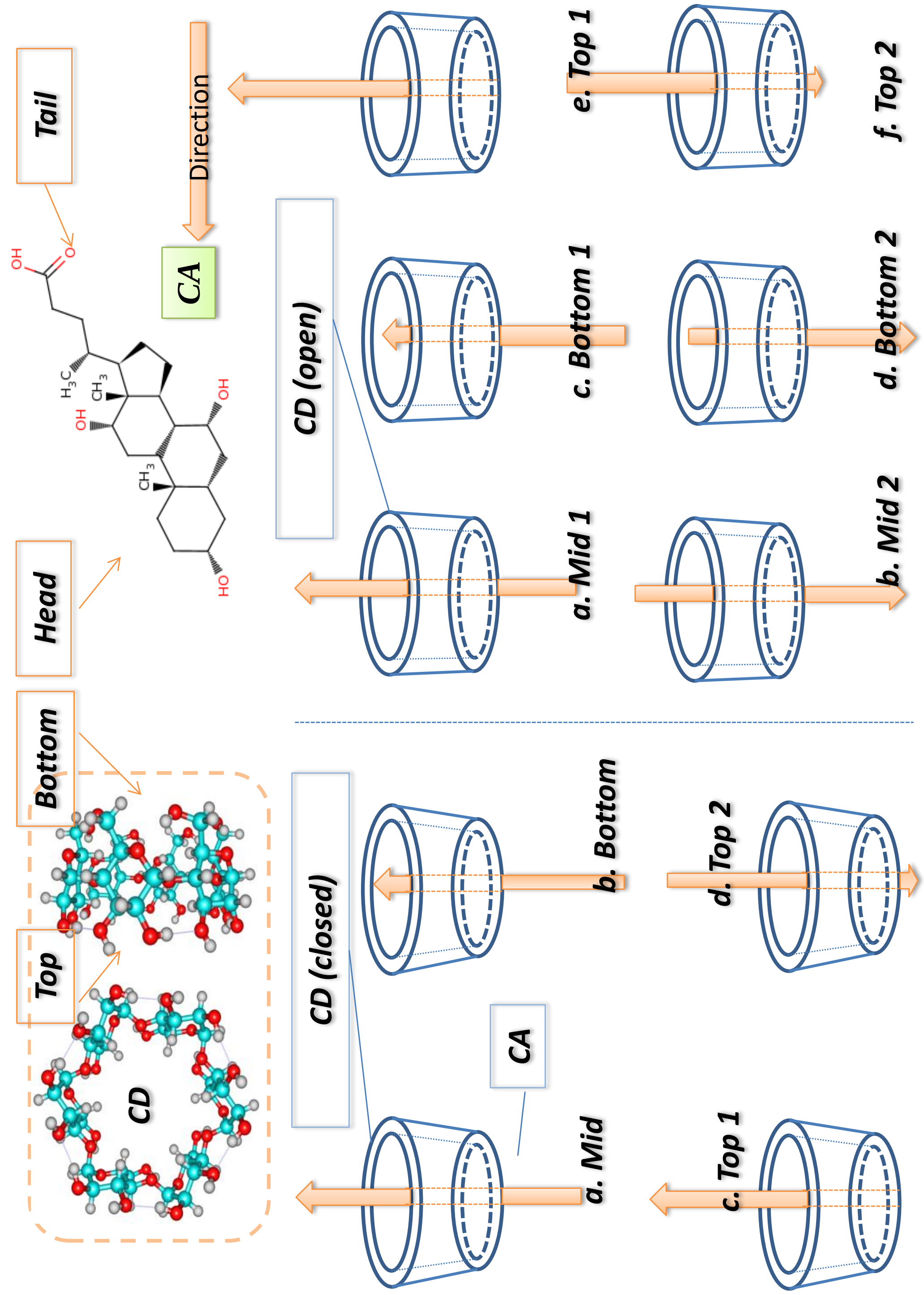


Figure 3.3 Schematic drawings of ten plausible structures of CD/CA complexes. The orange arrows indicate the direction pointing to the A-ring of CA. The hydrogen, oxygen, and carbon atoms are drawn in light gray, red, and blue, respectively.

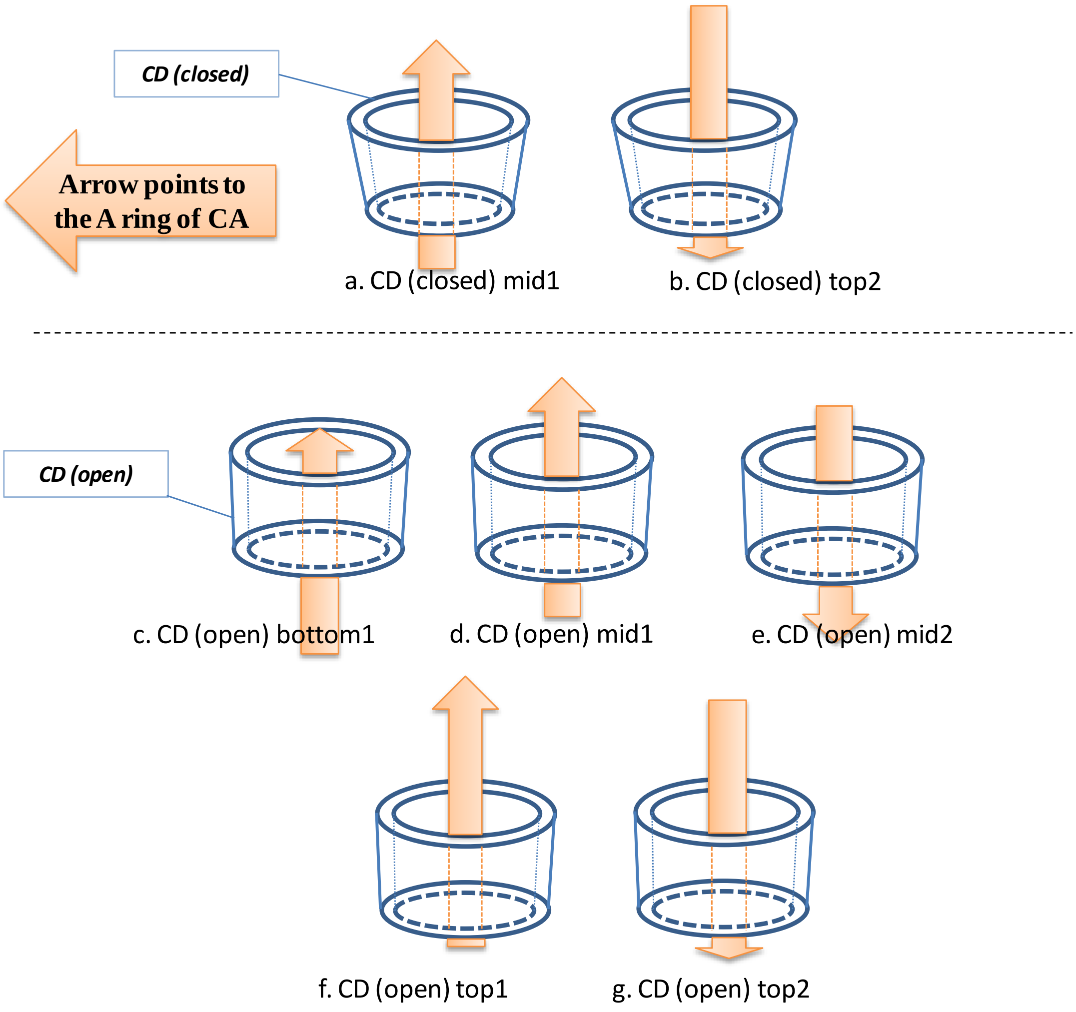


Figure 3.4 Schematic drawings of eight stable structures of CD/CA complexes selected based on the energies calculated by the B97-D functional with the 6-31G(d) basis set.

Table 3.3 Association energies (E_{ass} / kcal mol^{-1}) calculated with the PM3 and B97D/6-31G(d) methods in the gas phase.

Host molecules	Closed conformation			Open conformation		
	CA position	E_{ass}		CA position	E_{ass}	
		PM3	B97-D[a]		PM3	B97-D[a]
α-CD	Bottom1	-3.4	-38.5	Bottom1	-16.4	-52.0
	Mid 1	-12.1	-51.6	Bottom2	-8.1	-51.9
	Top1	-8.1	-39.5	Mid 1	-20.1	**-60.3**
	Top2	-14.9	-42.4	Mid2	-15.3	-51.0
				Top1	-18.3	-57.3
				Top2	-2.1	**-65.9**
β-CD	Bottom1	-2.3	-42.9	Bottom1	-25.3	-56.4
	Mid1	-26.1	**-68.5**	Bottom2	-14.2	-57.6
	Top1	0.3	-42.8	Mid 1	-29.6	-70.6
	Top2	-11.9	-51.8	Mid2	-27.2	**-77.7**
				Top1	-17.5	**-81.8**
				Top2	-19.7	-53.1
γ-CD	Bottom1	-19.1	-64.1	Bottom1	-15.4	**-74.4**
	Mid1	-22.2	-59.7	Bottom2	-17.1	-61.2
	Top1	0.2	-66.6	Mid 1	-22.9	-64.6
	Top2	-9.6	**-72.4**	Mid2	-17.6	-71.3
				Top1	-16.7	**-75.8**
				Top2	-16.9	-64.4

[a] BSSE was not corrected. In this subsection, only B97-D functional was employed. In the next subsection, M06-2X functional is also used. The two functionals give same tendency of association energies.

BSSE-corrected association energies of the CD/CA complexes in the gas phase

The geometries of the configurations selected above were re-optimized by the B97-D and M06-2X functionals with the 6-31G(d) basis set in the gas phase. The association energies (E_{ass}) of the CD/CA complexes are listed in Table 3.4 (The absolute electronic energies and BSSE values are given in Table S1 and Table S2 in Supporting Information). The complex with CD in the open conformation shows more negative association energy than the corresponding complex with CD in the closed conformation. The association energies of the complexes of β-CD and γ-CD are more negative than those of α-CD. While β-CD and γ-CD can include a part of the steroid skeleton of CA within their cavities, α-CD can accommodate only the side chain of CA owing to its smaller cavity diameter. For the β-CD/CA complex, the open/top1 and open/mid2 configurations are energetically favored. The open/top1 configuration agrees with the biding mode proposed for the β-CD/CA^{-} complex from NMR experiments in aqueous solution,[13,14] while open/mid2 resembles the structure observed for a complex of CA with a fluorophore-modified β-CD.[15] For the γ-CD/CA complex, the open/bottom1 and open/top1 configurations are energetically favored. The open/bottom1 structure is close to the structure proposed for the γ-CD complex with glycocholate in aqueous solution.[16] Thus, the present calculation for the β-CD/CA and γ-CD/CA complexes, although in the gas phase, reproduced the binding modes of the β- or γ-CD complexes with bile salts in aqueous solution rather well.

Table 3.4 also shows deformation energies, E_{deform}, and hypothetical interaction energies, E_{intrxn}, which are defined by eqs. 3 and 5, respectively. No distinctive correlation is observed between the association energy and the interaction energy. Instead weak correlation be seen between the deformation energy and the interaction energy; complexes with higher E_{deform} values tend to show more negative E_{intrxn} values. The association energy is thus determined by subtle balance between E_{deform} and E_{intrxn}.

Table 3.4 BSSE-corrected association energies (E_{ass} /kcal mol^{-1}), deformation energies (E_{deform} /kcal mol^{-1}), and interaction energies (E_{intrxn} /kcal mol^{-1}) of CD/CA complexes in the gas phase optimized with the B97-D and M06-2X functionals with the 6-31G(d) basis set.

CD	**CA position**[a]	$\boldsymbol{E}_{\mathbf{ass}}$		$\boldsymbol{E}_{\mathbf{deform}}$		$\boldsymbol{E}_{\mathbf{intrxn}}$	
		B97-D	**M06-2X**	**B97-D (H/G)**[b]	**M06-2X (H/G)**[b]	**B97-D**	**M06-2X**
α-CD (open)[c]	top2 (g)	–39.4	–30.5	26.1 (19.4/6.7)	15.2 (11.2/3.9)	–65.5	–45.7
	mid1 (d)	–36.7	–24.9	16.9 (10.2/6.7)	19.7 (10.5/9.1)	–53.6	–44.7
β-CD (open)	top1 (f)	–52.2	–34.1	14.5 (9.9/4.6)	26.6 (19.9/6.7)	–66.7	–60.7
	mid2 (e)	–48.7	–31.1	11.5 (9.0/2.5)	10.0 (5.8/4.2)	–60.2	–41.1
β-CD (closed)	mid1 (a)	–37.3	–25.6	36.7 (30.3/6.4)	24.0 (17.3/6.6)	–74.0	–49.6
γ-CD (open)	bottom1 (c)	–47.5	–40.4	13.5 (9.3/4.3)	7.5 (5.4/2.1)	–61.0	–47.9
	top1 (f)	–47.6	–37.5	21.1 (15.1/6.1)	11.8 (9.1/2.7)	–68.7	–49.3
γ-CD (closed)	top2 (b)	–41.1	–25.0	17.5 (12.8/4.7)	26.8 (23.2/3.6)	–58.6	–51.8

[a] The letter in the parentheses indicates the structure in Figure 3.4. [b] The deformation energies for the host and guest. [c] The conformation of CD is listed in the parenthesis. parts are shown in the parentheses.

3.3.3 Effects of DFT functional

In the preceding subsection, the calculations were performed with the B97-D and M06-2X functionals, which are considered to be suitable for evaluation of dispersion interactions. In order to examine the effects of the functional on the structure and the association energy of the β-CD/CA complex, the geometry optimization and evaluation of the association energy were performed with B3LYP, which is one of the most popular functionals. Figure 3.5 shows the geometries of the β-CD/CA complex in the open/top1 configuration optimized with the B97-D, M06-2X, and B3LYP methods with the 6-31G(d) basis set. The root mean square deviation (RMSD) values were 0.58, 1.06, and 1.46 Å for the B97-D/M06-2X, M06-2X/B3LYP, and B97-D/B3LYP pairs, respectively. The position of CA relative to β-CD in the structure optimized with B3LYP significantly differs from that predicted with B97-D or M06-2X; the steroid skeleton of CA is extruded from the cavity of β-CD in the former structure, whereas it resides inside the cavity in the latter one.

As shown in Table 3.5, the calculated association energies are quite different among the functionals used, although all the methods predict that open/top1 has the lowest energy. The E_{ass} value obtained with B97-D is more negative than that with M06-2X for any of the configuration investigated. The E_{ass} values obtained with B3LYP are positive, which is inconsistent with the experimental results that the association enthalpy is negative for the β-CD/CA complex. The difference among the functionals likely arises from the different treatments of intermolecular interactions. Grimme's B97-D functional,[3] is based on B97,[17] includes an additional correction term for dispersion interactions, and performs exceptionally well for supramolecular systems including van der Waals complexes.[18] On the other hand, M06-2X is a hybrid functional with parameters optimized for systems consisting of nonmetal elements.[4] The M06-2X functional has been employed to investigate thermochemistry, kinetics, and non-covalent interactions.[19] M06-2X is reported to perform best on non-covalent interactions for benchmarks in the S22 database.[20] On the other hand, B3LYP is a typical hybrid functional and has been conventionally used for organic molecules and 22supramolecular systems including CD inclusion complexes.[21, 22] The performance of B3LYP is, however, considered to be poorer than that of B97-D or M06-2X on evaluation of dispersion interactions. Thus, the dependence of the association energy on the functional suggests that dispersion interactions play an important role in the formation of CD/CA complexes. This point will be further examined with energy decomposition analyses with FMO-MP2 method in Chapter 5.

Table 3.5 BSSE-corrected association energies (E_{ass}) of β-CD/CA complex in aqueous phase calculated with the B97-D, M06-2X and B3LYP methods with the 6-31G(d) basis set.

CD	**CA**	**E_{ass} /kcal mol^{-1} [a]**		
(conformation)	**position**	**B97-D**	**M06-2X**	**B3LYP**
β-CD (open)	top1	–49.4 (–52.2)	–28.8 (–34.1)	7.3 (–7.2)
	mid2	–44.1 (–48.7)	–24.4 (–31.1)	19.3 (5.5)
β-CD (closed)	mid1	–41.5 (–37.3)	–24.9 (–25.6)	16.8 (–2.2)

[a] The values obtained in the gas phase are shown in the parentheses.

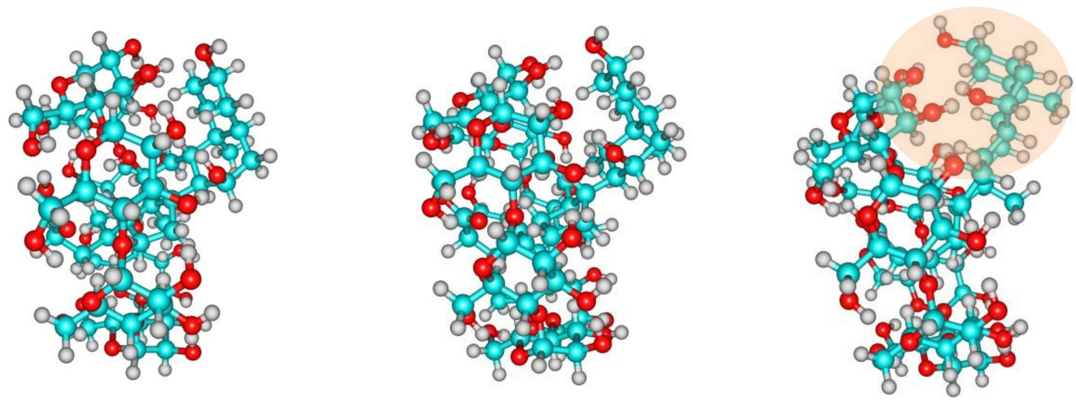


Figure 3.5 Structures of β-CD/CA in the open/top1 configuration optimized by the B97-D (left), M06-2X (middle), and B3LYP (right) functionals with the 6-31G(d) basis set in aqueous phase. The hydrogen, oxygen, and carbon atoms are drawn in light gray, red, and blue, respectively.

3.3.4 Hydrogen bonds between the host and the guest

Formation of intermolecular hydrogen bonds (H-bonds) is a significant driving force for complexation of a guest by CD. The numbers and geometries of H-bonds for the β-CD complexes are listed in Table 3.6 together with the association energies (H-bonds with H..O distances less than 2 Å are listed in Table 3.6). The views of the H-bonds are illustrated in Figure 3.6.

The number of H-bonds depends on the direction of CA; the open/top1 and closed/mid1 configurations involve six intermolecular H-bonds whereas the open/mid2 involves only three. One intramolecular H-bond in the narrow rim of β-CD is broken when it forms a complex in the open/top1 configuration. The lack of 7-OH in DCA or DCA^- decreases the number of H-bonds by two compared with CA or CA^- complexes, respectively. Despite the smaller number of H-bonds, the association energy for the β-CD/DCA (–28.1 kcal mol^{-1}) or β-CD/DCA^- (–30.7 kcal mol^{-1}) is similar to that for β-CD/CA (–28.8 kcal mol^{-1}) or β-CD/CA^- (–31.9 kcal mol^{-1}), respectively. Thus, the difference in the association energy among the complexes cannot be explained by the difference in the number of H-bonds.

Table 3.6 Intermolecular H-bonds in the β-CD/guest complexes together with BSSE-corrected association energies (E_{ass} /kcal mol-1) in aqueous phase calculated with the M06-2X functional with the 6-31G(d) basis set.

Guest	E_{ass}	**H-bond geometries ($r_{OH\cdots O}$/Å; ∠O-H···O/°)** [a]		**Number of H-bonds** [b]
		β-CD to guest	**Guest to β-CD**	
CA open/top1	–28.8	3-OH (#3)[c]···O7 (1.71; 163) 3-OH (#2)···O3 (1.84; 155) 3-OH (#1)···O12 (1.95; 148) 6-OH (#3)[d]···O=C (2.04; 132)[e]	7-OH···2-O (#3) (1.89; 138) COOH···6-O (#2) (1.80; 146)	6/13
CA open/mid2	–24.4	3-OH (#1)···O=C (1.81; 164)	7-OH···6-O (#7) (1.87; 146) 3-OH···6-O (#1) (1.88; 164)	3/14
CA closed/mid1	–24.9	3-OH (#7)···O7 (1.77; 159) 3-OH (#6)···O3 (1.79; 162) 6-OH (#7)[d]···O=C (1.92; 165) 6-OH (#6)[d]···O=C (1.94; 154)	7-OH···2-O (#7) (1.89; 137) 12-OH···4-O (#5) (1.97; 146)	6/12
CA^- open/top1	–31.9	3-OH (#3)···O7 (1.73; 157) 3-OH (#2)···O3 (1.83; 154) 3-OH (#1)···O12 (1.81; 153) 6-OH (#3)[d]···O=C (1.75; 165)	7-OH···2-O (#3) (1.99; 134)	5/13
DCA open/top1	–28.1	3-OH (#2)···O3 (1.82; 159) 3-OH (#1)···O12 (1.91; 150) 6-OH (#3)[d]···O=C (1.93; 154)	COOH···6-O (#2) (1.64; 169)	4/13
DCA^- open/top1	–30.7	3-OH (#2)···O3 (1.84; 156) 3-OH (#1)···O12 (1.76; 157) 6-OH (#3)[d]···O=C (1.72; 166)		3/13

[a] Labels for sugar units in β-CD are shown in the parentheses. [b] Number of inter- and intra-molecular H-bonds of the β-CD/guest complexes. [c] # denotes fragment number which is defined in Chapter 5. [d] The intramolecular H-bond in β-CD is broken on complexation with the guest. [e] Bond 6-OH (#6)···O=C is also considered here because of its relatively strong interaction energy (3.45 kcal mol^{-1}; see Chapter 5).

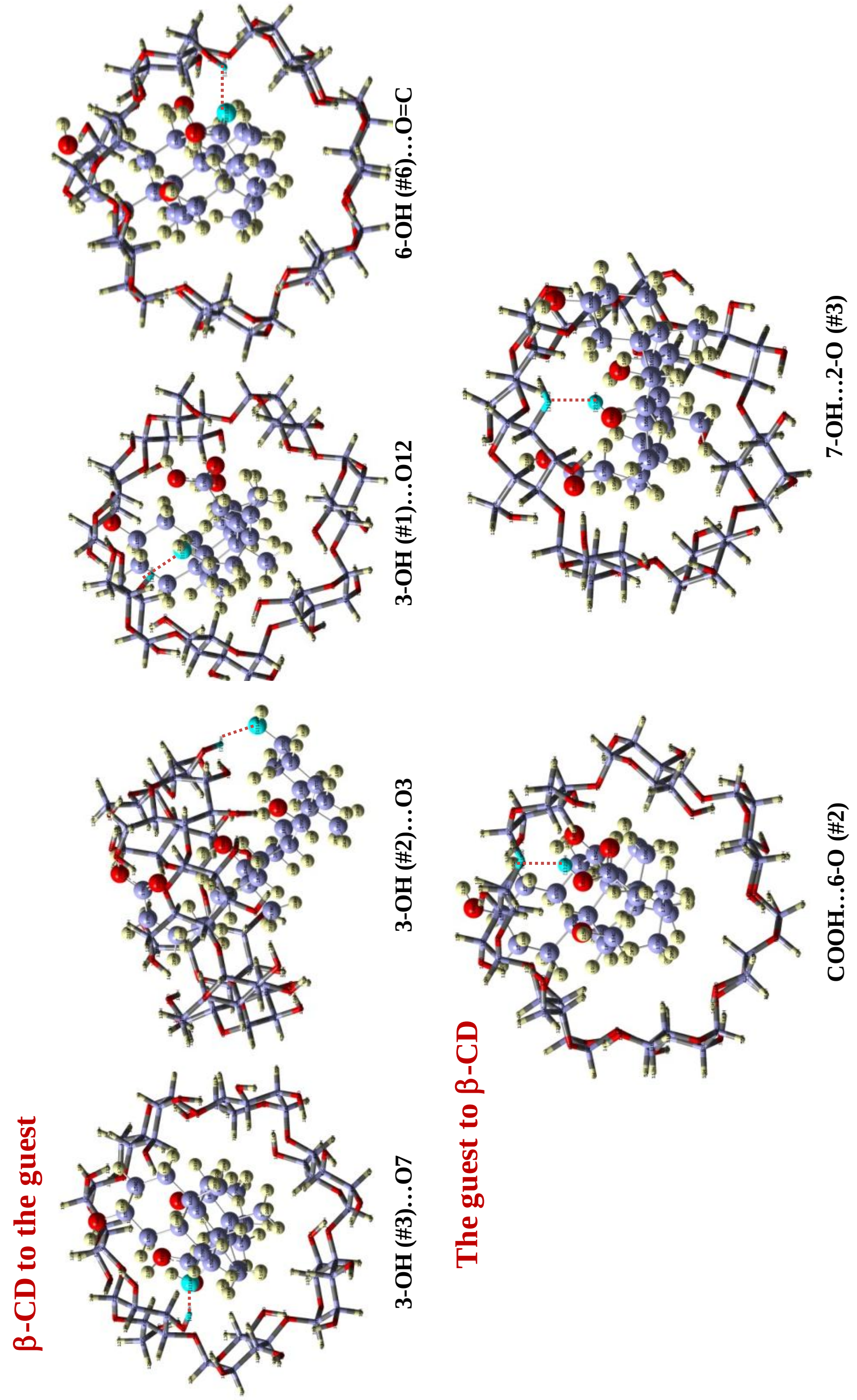


Figure 3.6 Intermolecular hydrogen bonds for β-CD/CA complex in the open/top1 configuration optimized by the M06-2X/6-31G(d) method with PCM. The hydrogen bonds are shown by dotted lines. Labels for sugar units in β-CD are shown in the parentheses. The atoms involving H-bonds are drawn in turquoise. The other hydrogen, oxygen, and carbon atoms are drawn in yellow, red, and blue, respectively.

3.3.5 Deformation of host and guest molecules on complexation.

Formation of an inclusion complex generally accompanies changes in the geometry of the host and guest, which raises the energy. Sizable deformation energies were indeed predicted for the CD/CA complexes as shown in Table 3.4. The geometrical changes are discussed in detail in this section. Figures 3.7 and 3.8 show the structures of guest molecules (CA and DCA) and β-CD in free state and in the complex. For either CA or DCA, the curvature of the molecule is enhanced on complexation, which decreases the O3-C24 distance. The root-mean-square deviation (RMSD) and the deformation energies (E_{deform}) were calculated for the β-CD/CA and β-CD/DCA complexes both in the gas phase and with PCM; the results are listed in Table 3.7. In the calculation of E_{deform} for the geometries optimized with PCM, the energies calculated without PCM at the same geometries were used since the deformation energy is considered to be an intrinsic quantity irrespective of the nature of the surrounding.

The deformation energy (E_{deform}) of β-CD is larger than that of the guest although the RMSD value is smaller for the former than for the latter. Since the alkyl chain in the guest is flexible, the geometrical changes in this part causes no severe destabilization. A distinctive change in β-CD is deviation from the seven-fold symmetry (See panels (a-3) and (a-4) in Figure 3.7). The RMSD value for DCA is smaller than that for CA, and the reduction of the O3-C24 distance on complexation is also less prominent for DCA. As a result, the E_{deform} value for β-CD/DCA is smaller than for β-CD/CA. In spite of the higher E_{deform}, β-CD/CA shows E_{ass} almost the same as that for β-CD/DCA (−28.8 and −28.1 kcal mol^{-1} for β-CD/CA and β-CD/DCA, respectively; see Table 3.6). Formation of two intermolecular H-bonds involving the 7-OH group in CA likely contributes to compensation for the increase in the deformation energy.

The reduction of the O3-C24 distance partially results from the conformational changes in the flexible tail part. The dihedral angles for the tail of the guest molecules are listed in Table 3.8. The atom numbering is shown in Chart 1.1. The dihedral angle of C16-C17-C20-C22 exhibits slight change on complexation whereas the dihedral angle of C20-C22-C23-C24 largely varies from 177.8° to −87.9° in the gas phase, indicating that the COOH group is rotated from trans to gauche with respect to the C20 atom (See panels (a-1) and (a-2) in Figure 3.7). Such rotation of the COOH group enables H-bonds to be formed with β-CD. Similarly, the dihedral angle of C20-C22-C23-C24 in DCA changed from 176.7° to −93.8° on complexation in the gas phase. The dihedral angles calculated with PCM are quite similar to those in the gas phase for either CA or DCA as shown in Table 3.8.

Table 3.7 O3-C24 Distances, root-mean-square-deviation (RMSD), and deformation energies (E_{deform}) of the guest and host for the β-CD/CA and β-CD/DCA complexes optimized by the M06-2X/6-31G(d) method in the gas phase and with PCM.

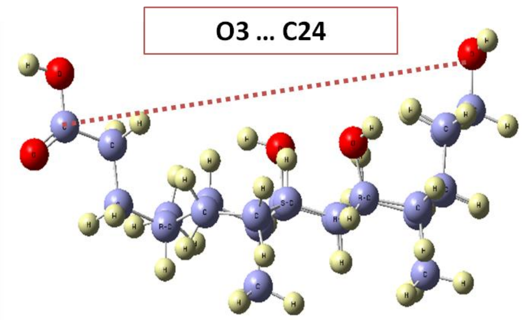
Guest	**Phase**	**O3-C24 / Å**		**RMSD / Å**		**E_{deform} / kcal mol^{-1} [a]**		
		free	**bound [b]**	**guest**	**host**	**guest**	**host**	**total**
CA	Gas	11.6	9.2 (2.4)	0.78	0.54	6.7	19.9	26.6
DCA	Gas	11.4	9.4 (2.0)	0.69	0.50	3.8	20.1	23.9
CA	PCM	11.7	9.2 (2.5)	0.79	0.47	6.1 (7.0)	20.7 (16.4)	26.6 (23.4)
DCA	PCM	11.4	9.6 (1.8)	0.74	0.50	3.5 (3.8)	13.2 (11.1)	16.8 (14.8)

[a] E_{deform} values evaluated from the energies with PCM are shown in the parentheses. Only the electrostatic solute-solvent interaction is included in the solvation free energy.
[b] Difference form the value in free state is shown in the parentheses.

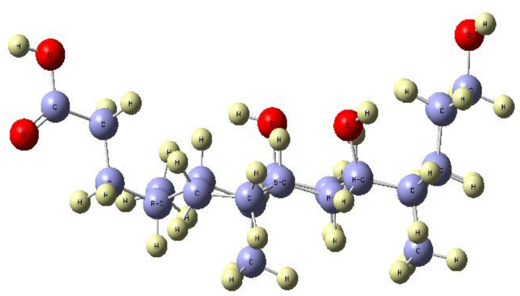
Table 3.8 The dihedral angles (°) for the guest molecules in the gas phase and with PCM obtained with the M06-2X/6-31G(d) method.

Guest	**C16-C17-C20-C22**		**C17-C20-C22-C23**		**C20-C22-C23-C24**	
	Gas phase	**PCM**	**Gas phase**	**PCM**	**Gas phase**	**PCM**
CA	53.6 (59.2) [a]	54.6 (58.9)	62.2 (75.4)	62.9 (75.3)	177.8 (–87.9)	177.4 (–87.6)
DCA	54.3 (59.5)	54.6 (52.9)	62.1 (71.1)	62.3 (79.4)	176.7 (–93.8)	176.7 (–80.8)

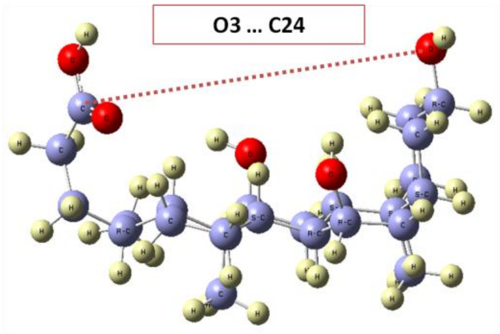
[a] The dihedral angles of the guest molecules in β-CD complexes (open/top1) are listed in the parentheses.



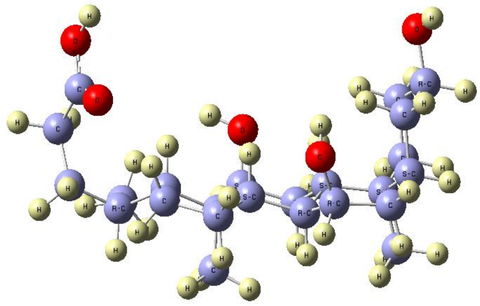
(a-1) Free CA in the gas phase



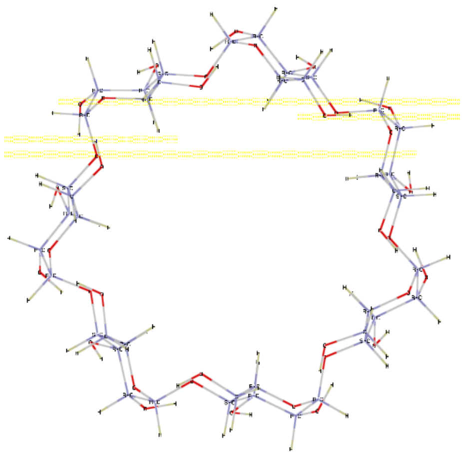
(b-1) Free CA in aqueous phase (PCM)



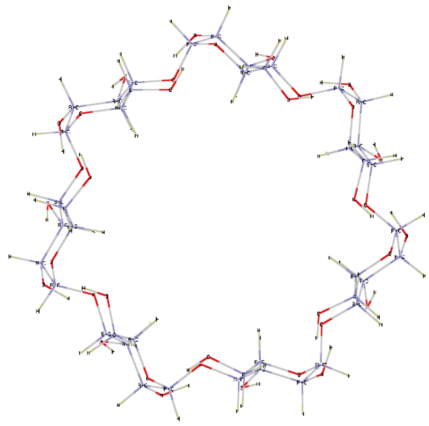
(a-2) CA in β-CD open/top1 in the gas phase



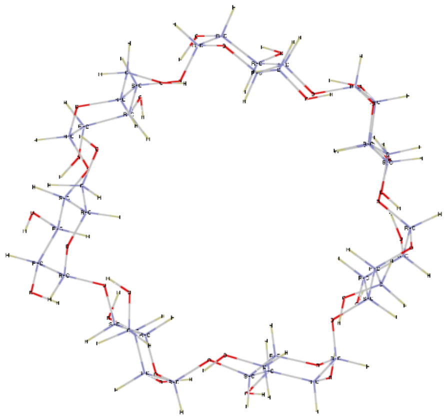
(b-2) CA in β-CD open/top1 in aqueous phase (PCM)



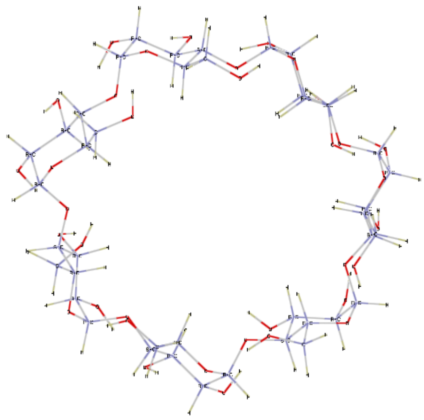
(a-3) Free CD in the gas phase



(b-3) Free CD in aqueous phase (PCM)

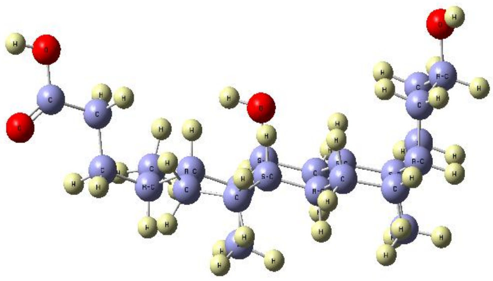


(a-4) CD in β-CD open/top1 in the gas phase

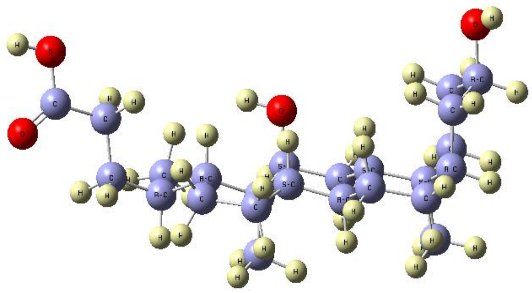


(b-4) CD in β-CD open/top1 in aqueous phase (PCM)

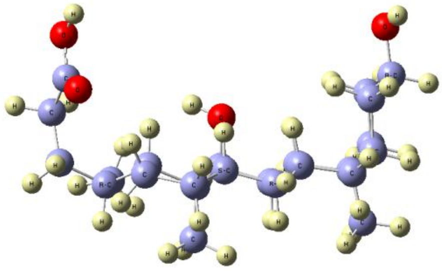
Figure 3.7 Structures of CA and β-CD in free state and in the complex optimized by the M06-2X/6-31G(d) method in the gas phase and with PCM. The hydrogen, oxygen, and carbon atoms are drawn in yellow, red, and blue, respectively.



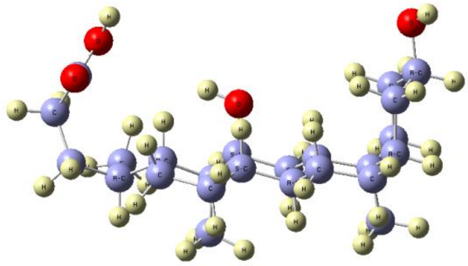
(c-1) Free DCA in the gas phase



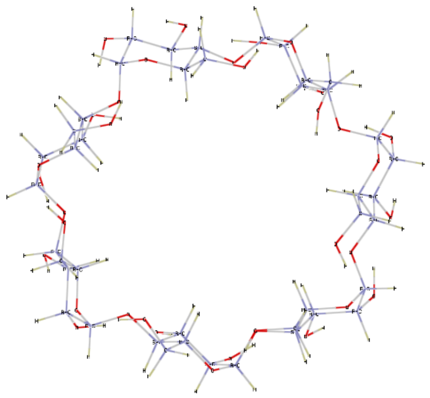
(d-1) Free DCA in aqueous phase (PCM)



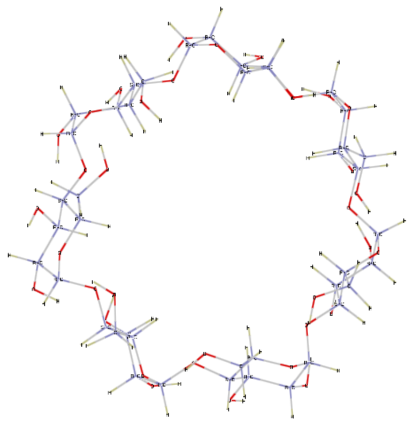
(c-2) DCA in β-CD open/top1 in the gas phase



(d-2) DCA in β-CD open/top1 in aqueous phase (PCM)



(c-3) CD in β-CD open/top1 in the gas phase



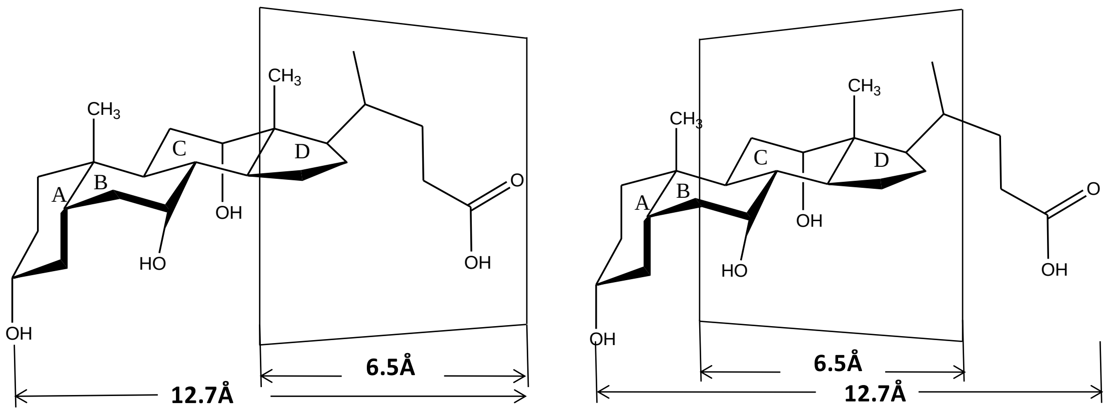
(d-3) CD in β-CD open/top1 in aqueous phase (PCM)

Figure 3.8 Structures of DCA and β-CD in free state and in the complex optimized by the M06-2X/6-31G(d) method in the gas phase and with PCM. The hydrogen, oxygen, and carbon atoms are drawn in yellow, red, and blue, respectively.

3.3.6 Comparison with experimental study

As mentioned in 3.3.2, the open/top1 and open/mid2 configurations are energetically favored for the β-CD/CA complex. As shown in Chart 3.3, a small part of the C-ring, the D-ring, and the tail of CA are inside the β-CD cavity in the most stable open/top1 configuration, whereas a half of the B-ring, the C-ring, and the D-ring in the open/mid2 configuration. The optimized structure of the open/top1 configuration is consistent with the binding mode proposed based on NMR experiments.[11,12]

It is reported that the experimental association enthalpies are −5.5 and −6.2 kcal mol^{-1} for β-CD/ CA^{-} and β-CD/DCA^{-}, respectively, in aqueous solution at 298 K.[12] The association enthalpies calculated by the M06-2X/6-31G(d) method with PCM are −26.0 and −24.1 kcal mol^{-1} for β-CD/CA^{-} and β-CD/DCA^{-}, respectively. The discrepancy with the experimental data may arise from the use of an implicit solvation model. In aqueous solution of CD, water molecules are included in the cavity of CD when other guests are absent,[20] and they are released when another guest enters into the cavity. Thus, the energy difference between water in the bulk and that confined in the CD cavity should be taken into consideration for evaluating thermodynamic quantities for the association equilibrium.[23] In addition, the complexes, free CDs, and guests can form H-bonds with water. Effects of such specific solute–solvent interactions, which are neglected in the PCM, will be considered by using MD simulation techniques in the next chapter.



β-CD/CA (open/top1) β-CD/CA (open/mid2)

Chart 3.3 Schematic drawings of the open/top1 (left) and open/mid2 (right) configurations of the β-CD/CA complex. β-CD is represented by cavity and approximate lengths of the host and the guest are shown.

Table 3.8 BSSE-corrected E_{ass} and H_{ass} (kcal mol^{-1}) in aqueous phase for the complexes of β-CD with bile acids or their anions in the open/top1 configuration optimized by the M06-2X/6-31G(d) method with PCM.

Guest	$\boldsymbol{E_{ass}}$	$\boldsymbol{H_{ass}}$	**Experimental** $\boldsymbol{H_{ass}}$
CA	–28.8	–22.5	———
CA^-	–31.9	–26.0	–5.5
DCA	–28.1	–23.3	———
DCA^-	–30.7	–24.1	–6.2

3.4 Conclusions

In Chapter 3, the stable structures, association energies (E_{ass}), and deformation energies (E_{deform}) have been studied for the CD/bile acid complexes by the DFT calculations both in the gas phase and aqueous phase. The relative stabilities of the open and closed conformations for free α-, β-, and γ-CDs are also discussed. The calculation results show that:

(1) The closed conformation is more stable than the open conformation for free α-, β-, and γ-CDs in the gas phase. This is likely because the hydrogen bonds at the narrow rim are shorter in the closed conformation than in the open one. On the other hand, the open conformation is more favored in the CD/bile acid complexes since it provides suitable space for a guest molecule to be logged in.

(2) The β-CD/CA and γ-CD/CA complexes are more stable than the α-CD/CA complex. The association energy depends on the contact area between the host and the guest which is governed by the cavity size of CD as well as on the position and direction of the guest (top, mid, or bottom; direction 1 or 2). For the β-CD/CA complex, the open/top1 and open/mid2 configurations are energetically favored, which is in agreement with experimental observations in aqueous solution.

(3) The shapes of the guest (CA or DCA) and host (β-CD) molecules are changed when they combine together either in the gas phase or in aqueous phase. The COOH group of CA or DCA is rotated from trans to gauche with respect to the C20 atom, resulting in formation of intermolecular hydrogen bonds with β-CD.

(4) The results for the CD/bile acid complexes obtained with the B97-D and M06-2X

functionals have similar tendencies. When the B3LYP functional is used, however, the optimized structure of β-CD/CA is quite different from that optimized with B97-D or M06-2X, and the association energy is predicted to be positive. The poor performance of B3LYP is likely due to the underestimation of dispersion interactions.

References and Notes

1 (a) W. Snor, E. Liedl, P. Weiss-Greiler, A. Karpfen, H. Viernstein, P. Wolschann, *Chem. Phys. Lett*. **2007**, *441*, 159. (b) A. Karpfen, E. Liedl, W. Snor, H. Viernstein, P. Weiss-Greiler, P. Wolschann, *Monatsh. Chem*. **2008**, *139,* 363.
2 (a) J. J. P. Stewart. *J. Comp. Chem.* **1989**, *10*, 209. (b) J. J. P. Stewart, *J. Comp. Chem.* **1989**, *10,* 221.
3 S. Grimme, *J. Comput. Chem.* **2006**, *27*, 1787.
4 Y. Zhao, D. G. Truhlar, *Theor. Chem. Acc.* **2008**, *120*, 215.
5 A. D. Becke, *J. Chem. Phys*. **1993**, *98*, 5648.
6 (a) R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54* , 724. (b) W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys*. **1972**, *56*, 2257. (c) P. C. Hariharan, J. A. Pople, *Theor. Chem. Acc*. **1973**, *28,* 213.
7 (a) M. T. Cances, B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *106*, 5151. (b) B. Mennucci, J. Tomasi, *J. Chem. Phys.* **1997**, *107*, 3032. (c) M. Cossi, V. Barone, B. Mennucci, J. Tomasi, *Chem. Phys. Lett*. **1998**, *286*, 253.
8 J. Tomasi, B. Mennucci, R. Cammi, *Chem. Rev*., **2005**, *105*, 2999.
9 S. F. Boys, F. Bernardi, *Mol. Phys.* **1970**, *19*, 553.
10 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 09 (Revision C.01), Gaussian, Inc.,

Wallingford CT, 2009.
11 Y. Yu, C. Chipot, W. Cai, X. Shao, *J. Phys. Chem. B*, **2006**, *110*, 6372.
12 M. Nagaraju, G. N. Sastry, *J. Phys. Chem. A.* **2009**, *113*, 9533.
13 P. R. Cabrer, E. Álvarez-Parrilla, F. Meijide, J. A. Seijas, E. R. Nuñez, J. V. Tato, *Langmuir* **1999**, *15*, 5489.
14 Y. Liu, Y.-W. Yang, E.-C. Yang, X.-D. Guan, *J. Org. Chem*. **2004**, *69*, 6590.
15 Y. Liu, J. Shi, D.-S. Guo, *J. Org. Chem.* **2007**, *72*, 8227.
16 S. Holm, C. Schönbeck, S. Askjar, P. Westh, *J. Incl. Phenom. Macrocycl. Chem.* **2013**, *75*, 223.
17 A. D. Becke, *J. Chem. Phys*. **1997**, *107*, 8554.
18 (a) F. F. Contreras-Torres, E. V. Basiuk, V. A. Basiuk, V. Meza-Laguna, T. Yu. Gromovoy, *J. Phys. Chem. A* **2012**, *116*, 1663. (b) S. Grimme, J. Anthony, S. Ehrlich, H. Krieg, *J. Chem. Phys*. **2010**, *132*, 154104. (c) P. Jurečka, J. Ćerny, P. Hobza, D. R. Salahub, *J. Comput. Chem*. **2007**, *28*, 555.
19 Y. Zhao, D. G. Truhlar, *J. Chem. Theory Comput.* **2007**, *3*, 289.
20 P. Jurečka, J. Sponer, J. Ćerny, P. Hobza, *Phys. Chem.* **2006**, *8*, 1985.
21 For example, N. Leila, H. Sakina, A. Bouhadiba, F. Madi. *J. Mol. Liq.* **2011,** *160*, 8; Y. Xia, X. Wang, Y. Zhang, B. Luo, *Comput. Theoret. Chem.* **2011**, *967*, 213.
22 D. Akase, M. Aida, S. S. Xantheas, *The VIth Annal Meeting of Japan Society for Molecular Science*, **2013**, 2E11.
23 Since the energy of water confined in the CD cavity is higher than that of water in the bulk, the predicted association enthalpy becomes more negative for the equilibrium described by CD/nH$_2$O(aq) + guest(aq) → CD/guest(aq) + nH$_2$O(aq).

Chapter 4

Molecular Dynamics (MD) simulations of CD/bile acid complexes

4.1 Introduction

In Chapter 3, the equilibrium structures and the association energies were investigated for the CD/bile acid complexes in the most stable configurations by the MO calculations. The solvent effects of water were taken into consideration only with an implicit solvation model, PCM. In real solution, the conformations of the host and guest molecules as well as the configurations of the complexes fluctuate in various timescales. Furthermore, water molecules can form hydrogen bonds with the complexes; such specific solute-solvent interactions cannot be adequately described with PCM. Molecular dynamics (MD) simulation is a useful method that provides valuable information on behaviors of chemical systems at the molecular level.[1,2] Classical MD methods can treat more than thousands of atoms at relatively low computational cost by using empirical force fields. Thus, solvation models are constructed with explicit solvent molecules in many cases, which may be able to describe specific solute-solvent interactions such as hydrogen-bonding better than simple continuum models do.

In Chapter 4, aqueous solutions of the β-CD/CA and γ-CD/CA complexes are investigated by MD simulations in order to clarify effects of thermal motion and solvent on the interactions between CA and CD. Simulations are also carried out for free β-CD and CA. Time evolutions and distributions of characteristic geometrical parameters are analyzed for trajectories of equilibrated systems. The results of analyses clarify the structural flexibility of CD and CA, the hydrogen-bonding interaction, and the hydration structures of the complexes. The simulation time is relatively short (~5 ns); the dissociation dynamics of the complexes are beyond the scope of the present study.

4.2 Computational Details

4.2.1 Molecular models

The initial coordinates of β-CD/CA and γ-CD/CA were taken from the geometries of the open/top1 configurations which were optimized with M06-2X/6-31G(d) in aqueous phase (described in Chapter 3). The initial coordinate of free CA was taken from the geometry optimized with M06-2X/6-31G(d) in aqueous phase while that of β-CD was from the crystal structure (PDB code: 1DMB),[3] which may be regarded as an exemplary structure of β-CD under hydrated conditions. For hydration of the model, a rectangular water box confined by a periodic boundary condition was added to the simulation system, see Figure 4.1. The sizes of the simulation boxes and the numbers of the water molecules included are listed in Table 4.1.

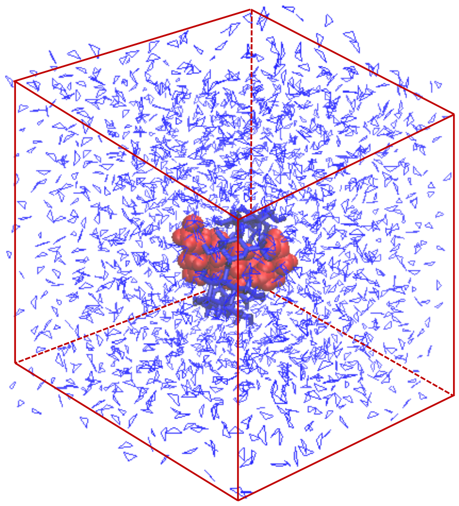


Figure 4.1 TIP3P water box model for β-CD/CA inclusion complex.

Table 4.1 Dimensions and contents of simulation cells for CA, β-CD, β-CD/CA, and γ-CD/CA

Solute	**Size of cell (Å^3)**[a]	**Number of water molecules**
CA (69)[b]	38.2 × 32.0 × 31.5	1252
β-CD (147)	39.6 × 39.6 × 34.6	1764
β-CD/CA (216)	40.2 × 39.8 × 37.3	1906
γ-CD/CA (237)	42.3 × 40.1 × 36.6	1985

[a] The box size for an NVT ensemble is shown. [b] Numbers of atoms in solute are listed in the parentheses.

4.2.2 Molecular dynamics (MD) simulations

All the MD simulations were performed by AMBER11[4] with GLYCAM06 force field[5] for β-CD or γ-CD, GAFF force field[6] for CA, and TIP3P water model.[7] GAFF is a general AMBER force field which was developed for rational drug design, and it has parameters for almost all the organic molecules made of C, N, O, H, S, P, F, Cl, Br, and

I. As a complete force field, GAFF is suitable to study a great amount of molecules.

The atomic charges used for CD are supplied in the GLYCAM06 parameter sets, which were derived from the HF/6-31G(d)//HF/6-31G(d) molecular electrostatic potential employing the restrained electrostatic potential (RESP) charge fitting methodology[8a] with a hyperbolic charge restraint weight of 0.01.[5] The Merz-Kollman electrostatic potential (ESP) charges[8b,c] were used for the atomic charges for CA, which were calculated by M06-2X/6-31G(d) with PCM.[8d,e] The same values for the atomic charges were used for the free species and the complexes.

The default factors of 1/2 and 1/1.2 were applied to scale 1−4 Lenard-Jones and electrostatic interactions for the GAFF parameters, whereas these factors were set to unity for the GLYCAM06 parameters. Such a combination of the force field parameters were used for MD simulations of β-CD/isoflavon complexes in aqueous alcohol solutions.[9] The cutoff for short-range interactions was set to 10.0 Å. Long-range electrostatic interactions were treated with the particle mesh Ewald method.[10]

The simulation conditions for equilibration and production runs are summarized in Table 4.2. At first, an MD run of 100 ps was carried out with the water molecules moving and the atoms in the solute frozen at their starting positions at constant temperature (T = 298 K) and pressure at 1 bar (NPT) with a time step of 1.0 fs. The Andersen thermostat was applied while the pressure was regulated by a weak-coupling scheme with a pressure relaxation time of 5.0 ps and a compressibility of 44.6×10^{-6} bar^{-1}. The density became 1.0 g cm^{-3} during the run. Next, both the water molecules and the solute were allowed to move for equilibration of whole the systems for 550 ps at a constant temperature (T = 298 K) and volume (NVT). SHAKE[11] was applied to constrain the C-H and O-H bond lengths at the ideal values. The temperature was regulated by Langevin dynamics with a collision frequency of 2.0 ps^{-1}. Lastly, a production run was performed under the same conditions to collect a trajectory for 5 ns. The trajectories for γ-CD/CA and free β-CD were obtained only for 4.4 ns and 2.5 ns, respectively. The coordinates were saved every 1.0 ps.

PTRAJ program in the AMBER11 package was used to analyze radial distribution functions, geometrical parameters of interest, and the root-mean-square deviation (RMSD)[12] which is defined as eq. 4-1.

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{N_{atoms}}(\mathrm{r_i(t)}-\mathrm{r_i(ref)})^2}{N_{atoms}}} \qquad (4\text{-}1)$$

where N_{atoms} is the number of atoms whose positions are being compared, $r_i(t)$ is the position of atom i at time t and r_i(ref) is that in the reference coordinate. The reference coordinates for free CA, β-CD/CA, and γ-CD/CA are the structures optimized at the M06-2X/6-31G(d) level with PCM while that for free β-CD was taken from the crystal structure (PDB code: 1DMB).

Table 4.2 MD simulation procedures for equilibration and production runs.

Simulation Procedure	# of Steps	Time step/ps	Time	Ensemble
(1) Equilibration (only solvent)	100,000	0.001	100 ps	NPT
(2) Equilibration (whole system)	550,000	0.001	550 ps	NVT
(3) Production run	5,000,000	0.001	5 ns	NVT

4.3 Results and Discussion

4.3.1 Structural features of the hosts

Figure 4.2(a) shows the time evolutions of root-mean-square deviations (RMSDs) for β-CD/CA (blue), γ-CD/CA (red), free β-CD (green), and free CA (violet) in the production runs. Although the RMSD values of the complexes fluctuated with time, no abrupt change was observed, indicating that the complex did not dissociate within the simulation time (equilibration for 0.55 ns followed by a production run for 5 ns). In average, the position of the CA in the cavity of β-CD is consistent with the results of NMR experiments reported for β-CD/CA^- complex in D_2O solution; distinctive peaks are seen at 2.3–2.5 Å in the radial distribution functions of the proton-proton pairs for which nuclear Overhauser effect (nOe) was observed (See Figure S2 in Supporting Information and reference 13). Inspection of several snapshots for β-CD/CA and γ-CD/CA reveals that the configurations in the most snapshots examined are classified as the open/top1, which is similar to the structure proposed for β-CD/CA^- in aqueous solution from the experimental data, as described in Chapter 3. The RMSD values of the β-CD/CA and γ-CD/CA complexes are similar, and the fluctuation is somewhat larger for β-CD/CA (see Table 4.3).

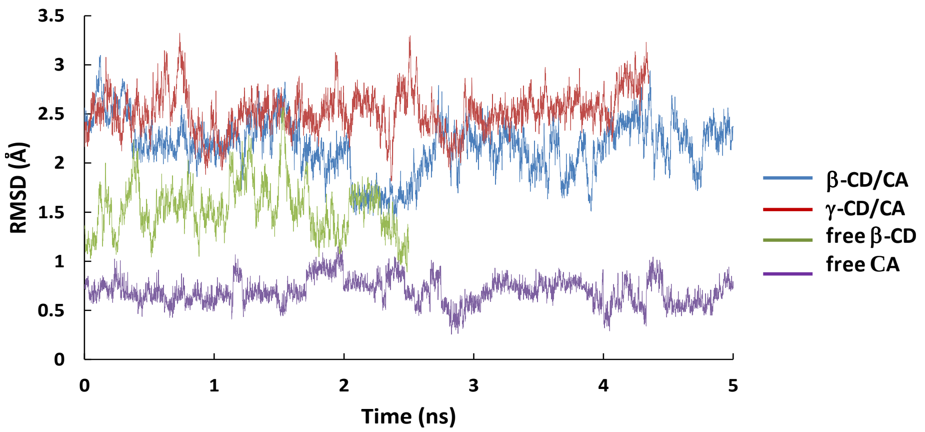
Figure 4.2(b) indicates the RMSDs of CDs in the free state and in the complexes. Free β-CD is more flexible than β-CD in β-CD/CA complex. γ-CD in the γ-CD/CA complex is more flexible than β-CD in β-CD/CA, suggesting that the increase in sugar units makes the host molecule more flexible. Figure 4.2(c) shows the RMSDs of CA in the free state and in the complexes. Since the side chain of CA is rather flexible as discussed in Chapter 3, the geometrical change of this moiety may have large contribution to the RMSD. Free CA varies its conformation more frequently than CA in the complexes. The fluctuation of CA in γ-CD/CA is smaller than that of CA in β-CD/CA or in the free state. It should be noted that the conformational fluctuation of the host part is larger than that of the guest part in the γ-CD/CA complex.

Table 4.3 The averages and standard deviation values (shown in the parentheses) of RMSDs (Å) for β-CD/CA, γ-CD/CA, free β-CD, and free CA.

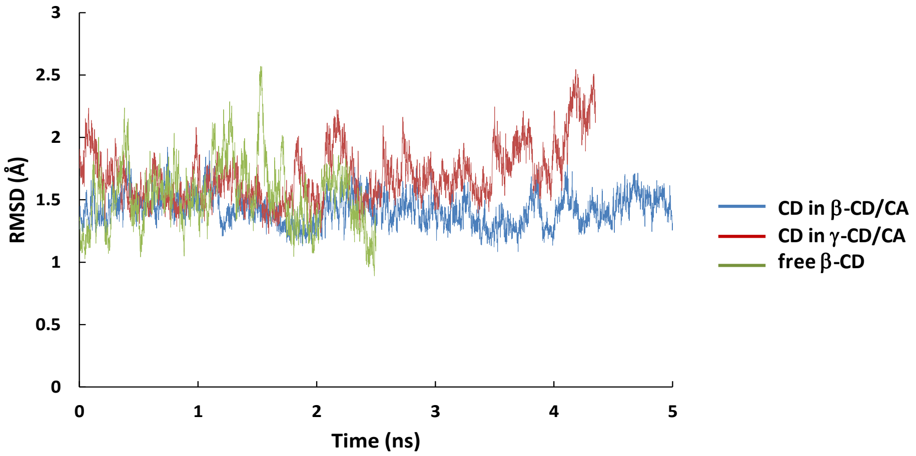
Structures	**Complex**	**Host (CD)**	**Guest (CA)**
β-CD/CA	2.17 (0.28)	1.42 (0.13)[a]	1.63 (0.14)[b]
γ-CD/CA	2.53 (0.21)	1.70 (0.23)[a]	1.13 (0.09)[b]
free β-CD	-------------	1.54 (0.26)	-------------
free CA	-------------	-------------	0.69 (0.13)

[a] The atoms belonging to CD were superimposed with the corresponding atoms in the optimized structure of the CD/CA. [b] The atoms belonging to CA were superimposed with the corresponding atoms in the optimized structure of the CD/CA.

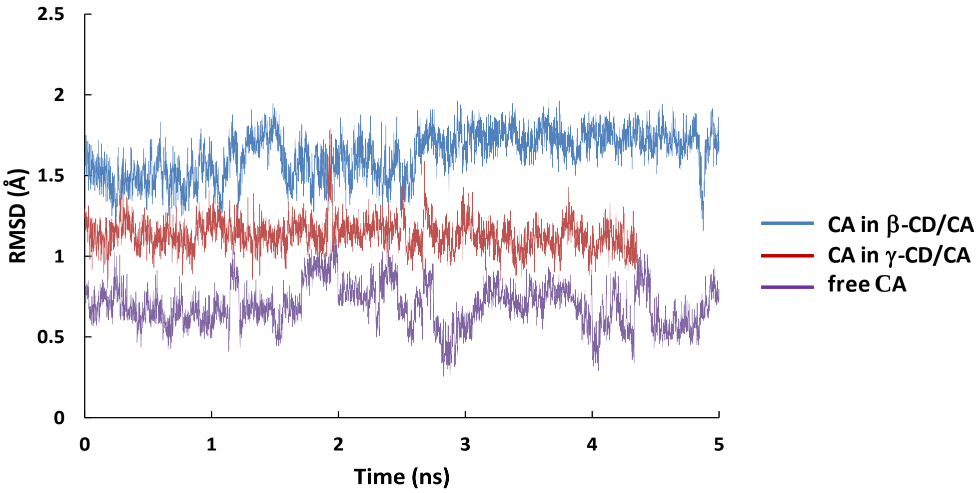
The conformational flexibility of the CD was further analyzed in terms of the diagonal distance, $D(i)$ for β-CD or $D'(i)$ for γ-CD, and the inter-residue angle $\theta(i)$, which were defined as eq. 4-2 and eq. 4-3, respectively; see also Figure 4.3. The results are summarized in Table 4.4. The averages of $D(i)$, $D'(i)$, and $\theta(i)$ are also shown as D_{ave}, D'_{ave}, and θ_{ave}, respectively.



(a) RMSD for β-CD/CA (blue), γ-CD/CA (red), free β-CD (green), and free CA (violet).



(b) RMSD for β-CD in β-CD/CA (blue), γ-CD in γ-CD/CA (red) and free β-CD (green).



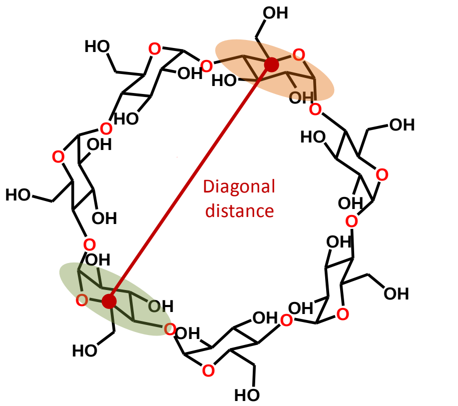
(c) RMSD for CA in β-CD/CA (blue), CA in γ-CD/CA (red) and free CA (violet).

Figure 4.2 Time evolutions of the root-mean-square deviations (RMSD/Å) for β-CD/CA (blue), γ-CD/CA (red), free β-CD (green), and free CA (violet).

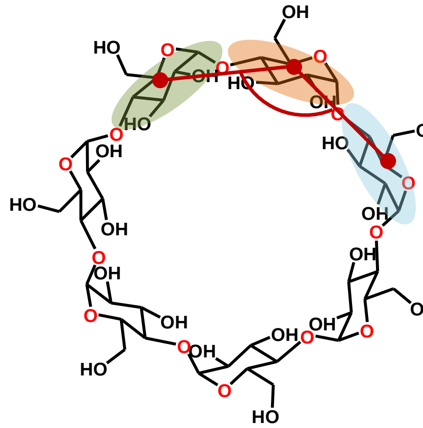
Table 4.4 Diagonal distances ($D(i)$ or $D'(i)$ / Å) and inter-residue angles $\theta(i)$ /° of β-CD and γ-CD in the free state and in the CD/CA complexes.

Structures	***D*(1)**	***D*(2)**	***D*(3)**	***D*(4)**	***D*(5)**	***D*(6)**	***D*(7)**		D_{ave} [a]
β-CD-crystal[b]	11.2	10.5	10.7	10.9	11.0	10.5	10.8		10.8
Free β-CD-MD[c]	11.4	11.1	11.2	11.5	11.2	11.1	11.3		11.3
	(0.59)	(0.71)	(0.61)	(0.65)	(0.69)	(0.61)	(0.66)		(0.25)
β-CD/CA-opt[d]	10.7	10.6	10.4	11.0	10.6	10.5	10.7		10.6
β-CD/CA-MD[c]	11.4	11.4	11.3	11.4	11.4	11.3	11.3		11.3
	(0.43)	(0.35)	(0.43)	(0.41)	(0.37)	(0.44)	(0.38)		(0.11)
	***D'*(1)**	***D'*(2)**	***D'*(3)**	***D'*(4)**					D'_{ave} [a]
γ-CD/CA-opt[d]	11.6	12.5	13.2	12.2					12.4
γ-CD/CA-MD[c]	12.9	12.8	13.1	13.3					13.0
	(0.53)	(0.56)	(0.62)	(0.59)					(0.16)
	$\theta(1)$	$\theta(2)$	$\theta(3)$	$\theta(4)$	$\theta(5)$	$\theta(6)$	$\theta(7)$	$\theta(8)$	θ_{ave} [e]
β-CD-crystal[b]	135.0	126.9	123.7	129.4	127.2	132.0	121.2		127.9
Free β-CD-MD[c]	124.9	127.8	126.6	123.5	127.8	126.9	125.2		126.1
	(8.18)	(7.64)	(8.87)	(8.33)	(8.55)	(7.61)	(9.36)		(1.45)
β-CD/CA-opt[d]	125.1	131.3	124.6	127.1	134.2	121.2	133.4		127.1
β-CD/CA-MD[c]	126.5	127.2	128.2	127.1	126.4	127.6	127.9		127.3
	(5.46)	(5.06)	(6.57)	(6.83)	(5.08)	(5.92)	(6.38)		(0.80)
γ-CD/CA-opt[d]	134.3	128.3	131.7	141.7	135.1	121.2	143.5	136.1	133.8
γ-CD/CA -MD[b]	133.7	130.8	134.9	129.7	133.8	136.3	127.1	131.0	132.2
	(6.90)	(6.61)	(5.82)	(6.58)	(6.73)	(8.68)	(7.59)	(6.35)	(1.20)

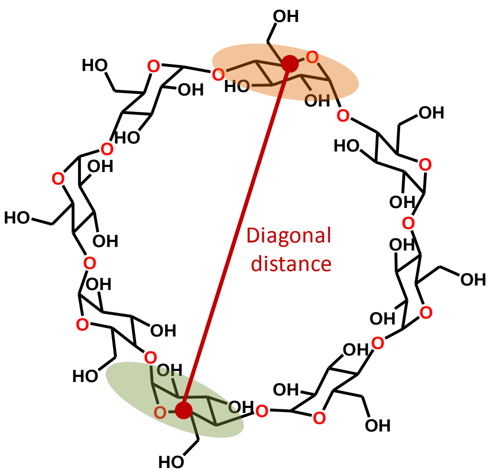
[a] Average of seven $D(i)$ or four $D'(i)$ values for each snapshot was analyzed for the ensemble. [b] Taken from PDB code: 1DMB.[3] [c] Ensemble-average over the MD run. The standard deviation values are listed in the parentheses. [d] The open/top1 configuration optimized by M06-2X/6-31G(d) in aqueous phase with PCM. [e] Average of seven or eight $\theta(i)$ values for each snapshot was analyzed for the ensemble.



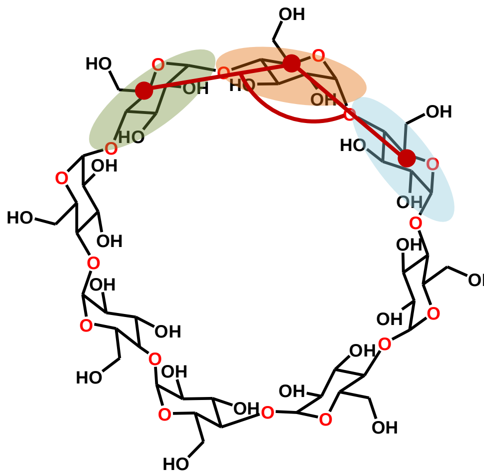
(a) Diagonal distance $D(i)$



(b) Inter-residue angle $\theta(i)$



(c) Diagonal distance $D'(i)$



(d) Inter-residue angle $\theta(i)$

Figure 4.3 Definitions of the diagonal distance $D(i)$ or $D'(i)$ and the inter-residue angle $\theta(i)$ for β-CD (panels a and b) and γ-CD (panels c and d).

$$D(i) = |\,\mathbf{r}_{\mathrm{COM}}(i) - \mathbf{r}_{\mathrm{COM}}(i+3)\,| \quad \text{for } i = 1,2,3,4$$

$$\text{or } = |\,\mathbf{r}_{\mathrm{COM}}(i) - \mathbf{r}_{\mathrm{COM}}(i-4)\,| \quad \text{for } i = 5,6,7 \tag{4-2a}$$

$$D'(i) = |\,\mathbf{r}_{\mathrm{COM}}(i) - \mathbf{r}_{\mathrm{COM}}(i+4)\,| \quad \text{for } i = 1,2,3,4 \tag{4-2b}$$

$$\theta(i) = \angle\, \mathbf{r}_{\mathrm{COM}}(i-1) - \mathbf{r}_{\mathrm{COM}}(i) - \mathbf{r}_{\mathrm{COM}}(i+1) \tag{4-3}$$

where $\mathbf{r}_{\mathrm{COM}}(i)$ represents the positional vector of the center-of-mass for the i-th residue.

The averaged diagonal distance, D_{ave}, of free β-CD or β-CD/CA in the MD run (11.3 Å) is larger than the value for β-CD in the crystal (10.8 Å)[3] or in the optimized structure of β-CD/CA (10.6 Å) (optimized by M06-2X/6-31G(d) in aqueous phase, see Table 4.4). Similarly, the D'_{ave} value for γ-CD/CA in the MD run (13.0 Å) is larger than that in the optimized structure (12.1 Å). The expansion of the CD ring is likely due to thermal motion in the MD simulation at 298 K.

The averaged inter-residue angle, θ_{ave}, of β-CD in the crystal (127.9°) or in the optimized β-CD/CA complex (127.1°) is close to the angle of a regular heptagon (128.6°), which suggests that the center-of-mass of the sugar units are not significantly deviated from the mean plane. The θ_{ave} value of free β-CD in the MD run (126.1°) is slightly smaller, reflecting the deformation of flexible β-CD. A similar tendency is also seen for γ-CD when the θ_{ave} value is compared with the angle of a regular octagon (135.0°).

A comparison of the standard deviations of these structural parameters provides clear insight on the flexibility of the CDs. The standard deviation increases in the order of β-CD/CA < γ-CD/CA < free β-CD for any parameters examined. Figure 4.4 shows the distribution of the diagonal distances, D(i) or $D'(i)$, and the inter-residue angles, $\theta(i)$. The distribution of β-CD/CA is narrower than that of free β-CD or γ-CD/CA for both the parameters. These results clearly reveal that β-CD in β-CD/CA is less flexible than free β-CD or γ-CD in γ-CD/CA. Thus, complexation with CA makes β-CD more rigid.

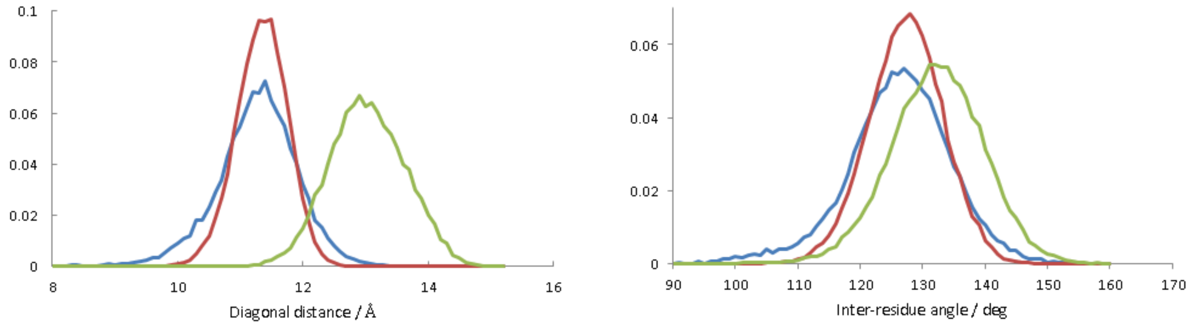


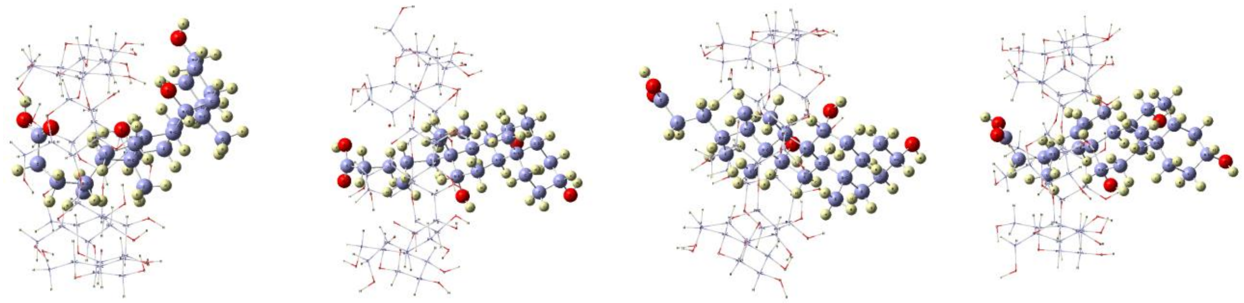
Figure 4.4 Distribution of diagonal distances $D(i)$ or $D'(i)$/Å (left panel) and inter-residue angles $\theta(i)$ / ° for free β-CD (blue), β-CD/CA (red), and γ-CD/CA (green).

4.3.2 Conformation and position of CA in the CD/CA complexes

As was mentioned above, the guest molecule did not escape from the CD cavity but resided in its inside for the simulation time. However, the position of CA relative to the host varied with time. As shown in Figure 4.5, the β-CD/CA complex adopted in the open/top1 configuration at 1 ns while it has changed to open/mid1 at 3 ns and again to open/top1 at 5 ns. The configuration of γ-CD/CA is kept to be open/top1 at 1 ns, 3 ns, and 4.4 ns. The comparison of snapshots between β-CD/CA and γ-CD/CA indicates that the structures of the two complexes differ in the conformation of CA and the relative orientation of the host and guest. In the cavity of β-CD, CA adopts an extended conformation and lies with its long axis almost parallel to the short axis of β-CD. On the other hand, CA in γ-CD/CA is in a more curly conformation and inclined with respect to the short axis of γ-CD so that better contact can be attained between the host and guest.

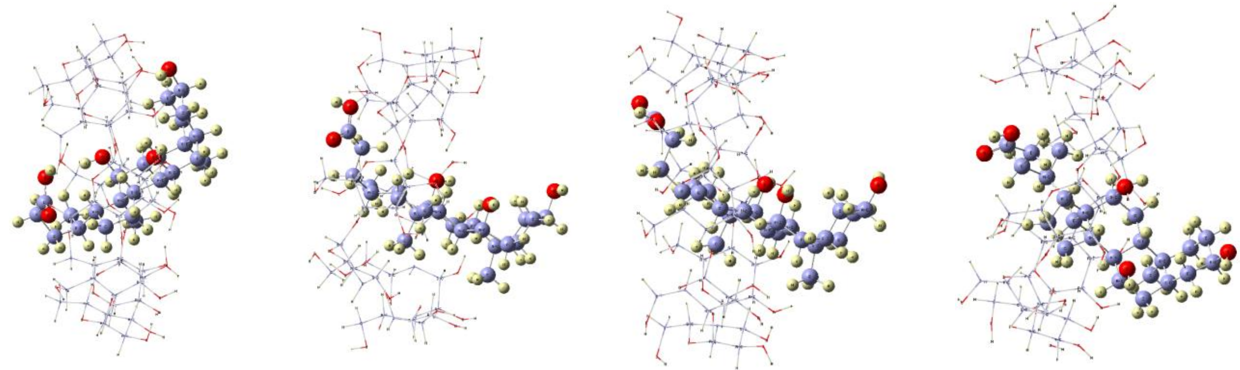
The O3–C24 distances of free CA and CA in β-CD or γ-CD cavity were discussed in Chapter 3 (Figures 3.7 and 3.8 and Table 3.7). Here, O3 denotes the oxygen in 3-OH and C24 denotes the carbon atom in the COOH group. Table 4.4 summarizes the O3–C24 distances obtained by the MD simulation together with those obtained by the optimization. The O3-C24 distance of free CA in the optimized structure (11.7 Å) is close to that in the MD simulation (12.0 Å), suggesting that the equilibrium geometry of CA in aqueous solution could be obtained rather accurately either by the MO calculation with PCM or by the classical MD simulation with explicit water molecules.

β-CD/CA



(a) Reference structure (b) Snapshot at 1 ns (c) Snapshot at 3 ns (d) Snapshot at 5 ns

γ-CD/CA



(a’) Reference structure (b’) Snapshot at 1 ns (c’) Snapshot at 3 ns (d’) Snapshot at 4.4 ns

Figure 4.5 The reference structures optimized with M06-2X/6-31G(d) in PCM (panels a and a’) and snapshots of β-CD/CA (panels b-d) and γ-CD/CA (panels b’-d’).

The O3–C24 distance of CA in β-CD/CA obtained by the MD simulation (13.3 Å) is longer than that of free CA (12.0 Å), indicating that CA in the cavity of β-CD adopts a more extended conformation than that in the free state. On the other hand, the O3–C24 distance of CA in γ-CD/CA obtained by the MD simulation (12.1 Å) is close to that of free CA, suggesting that the structural change of CA upon complexation with γ-CD may be small. The standard deviation of the O3–C24 distance for free CA in the MD run (0.59 Å) is larger than that for CA in the complex (0.38 or 0.43 Å), which is consistent with the larger fluctuation of the RMSD for the former than the latter, as shown in Figure 4.2(c). The O3–C24 distance of CA in β-CD/CA (9.2 Å) or in γ-CD/CA (9.6 Å) obtained by the optimization is shorter than that obtained by the MD simulation. The discrepancy likely results from the difference in hydrogen-bonding pattern predicted by the MO and MD methods, as will be discussed in the next subsection.

Table 4.4 O3-C24 distance of free CA and CA in β-CD and γ-CD cavities.

Structures	**O3-C24 distance / Å**
Free CA-opt[a]	11.7
Free CA-MD[b]	12.0 (0.59)
CA in β-CD/CA[c]-opt[a]	9.2
CA in β-CD/CA-MD[b]	13.3 (0.38)
CA in γ-CD/CA[c]-opt[a]	9.6
CA in γ-CD/CA-MD[b]	12.1 (0.43)

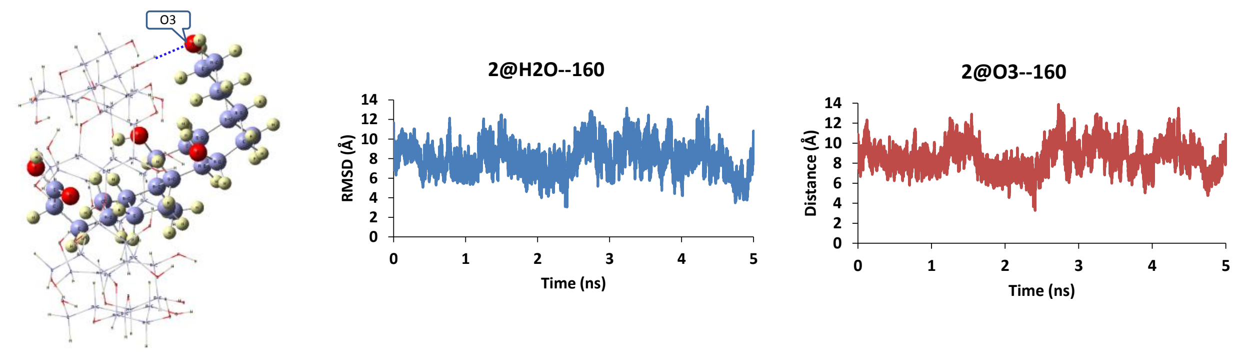
[a] Structure was optimized by M06-2X/6-31G(d) in aqueous phase with PCM. [b] Average values for the MD trajectory. The standard deviation values are listed in the parentheses. [c] CD/CA complex in the open/top1 configuration.

4.3.3 Hydrogen bonds

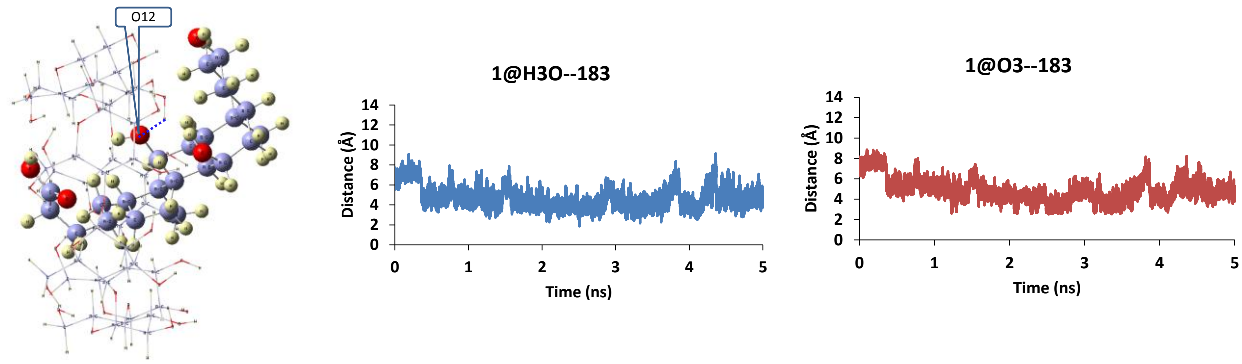
In the optimized structure of β-CD/CA complex in the open/top1 configuration, the five oxygen atoms in CA form six hydrogen bonds (H-bonds) with hydroxyl groups in β-CD (see Table 3.4). In order to examine whether the H-bonds are retained or broken in aqueous solution, the interatomic distances involving the H-bonds were traced over the MD trajectory. As shown in Figure 4.6, H-bonds are not always formed between β-CD and CA during the 5-ns simulation; the original H-bond donor (OH) and accepter (O) are far from each other for most of the time. Some of the OH and COOH groups in CA form H-bonds with water; for example, the 3-OH group in CA forms two H-bonds with water molecules at 1 ns as shown in Figure 4.7.

Hydrogen-bonding schemes can be characterized by radial distribution functions (RDFs). Figure 4.8(a) shows the RDFs of oxygen atoms in β-CD around 3-O, 7-O, 12-O, and C=O in CA for the β-CD/CA complex. The RDF for 12-O exhibits a broad peak at ~2.9 Å, suggesting that the 12-OH in CA sometimes forms an H-bond with an oxygen atom in β-CD. On the other hand, the 3-O atom in CA is far from any oxygen atoms in β-CD. The difference in the H-bond partner between 3-OH and 12-OH in CA is also seen in RDFs of oxygen atoms in water around the 3-O and 12-O atoms. As shown in Figure 4.8(b), the hydration structure of the 3-OH group (broken lines) is little affected by complexation with β-CD, whereas the first peak of RDF for 12-O (solid lines) is diminished upon the complexation.

Reference structure **O...H distance** **O...O distance**



(a-1) 3-OH(CD)…O3 (1.84 Å) (a-2) (a-3)



(b-1) 3-OH(CD)…O12 (1.95 Å) (b-2) (b-3)

Figure 4.6 Time evolutions of the distances between the atoms which formed H-bonds in the optimized structure of β-CD/CA complex. The left panels show the views of the optimized structure. The middle and right panels display the H…O and O…O distances, respectively. The upper and lower panels show the distances involving the 3-O and 12-O atoms, respectively.

In the β-CD/CA of the open/top1 configuration, the 12-OH group in CA (see Chart 3.1) is located in the vicinity of the wider rim of β-CD, along which the secondary hydroxyl groups are aligned, as seen in panels (b) and (d) of Figure 4.5. When the structure of β-CD/CA is classified as open/mid1 (see panel (c) of Figure 4.5), 12-OH is berried in the cavity of β-CD, and the wall of the cavity keeps this OH group out from solvent molecules. The characteristic features of RDFs confirm that the snapshots in Figures 4.5 and 4.7 are representative structures in the MD trajectory.

Effects of the complexation with CA on hydration structures of the CDs are also examined. Figure 4.9 displays RDFs of oxygen atoms in water around the oxygen atoms of the 2-OH and of the 6-OH groups in free β-CD, β-CD/CA, and γ-CD/CA. The RDFs around 3-OH (not shown) are similar to the corresponding curves around 2-OH. The first peak of RDF for the primary hydroxyl groups, 6-OH, in β- or γ-CD is little affected by complexation with CA whereas that for the secondary groups is diminished when the CD accommodates CA in its cavity. The results of RDFs for the CDs are consistent with occasional formation of H-bonds between CD and CA as described above.

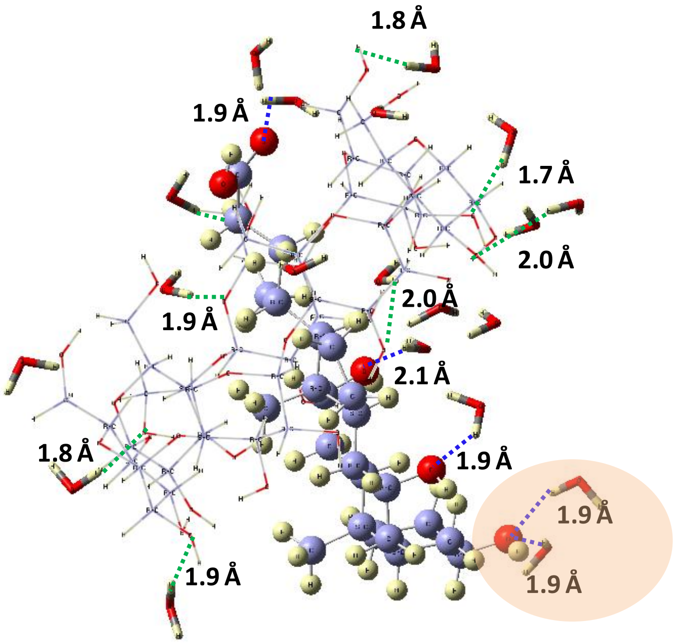
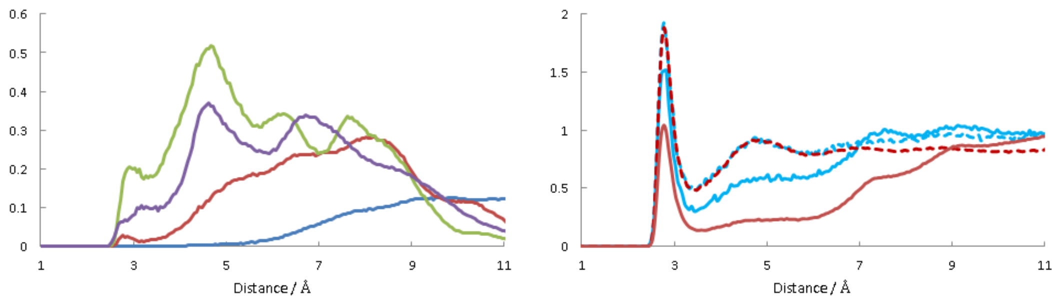
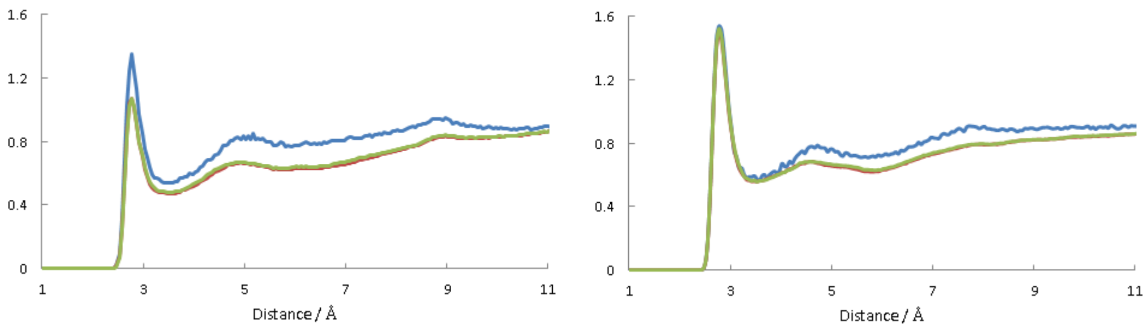


Figure 4.7 Snapshot of β-CD/CA with 20 closest water molecules at 1 ns. The green dotted lines denote the hydrogen bonds between β-CD and water molecules, while the blue dotted lines denote the hydrogen bonds between CA and water molecules. The H…O distances are also given.



(a) (b)

Figure 4.8 (a) Radial distribution functions of O atoms in β-CD around the oxygen atoms at 3-OH (blue), 7-OH (red), 12-OH (green), and C=O (violet) in CA for the β-CD/CA complex. (b) Radial distribution functions of O atoms in water around the oxygen atoms at 12-OH (solid lines) and at 3-OH (broken lines) in free CA (turquoise) and CA in β-CD/CA (red).



(a) RDF of O_{H_2O} from 2-O in CD (b) RDF of O_{H_2O} from 6-O in CD

Figure 4.9 Radial distribution functions of the O atoms in water around the 2-O atoms (a) and 6-O atoms (b) in free β-CD (blue), β-CD/CA (red), and γ-CD/CA (green).

4.4 Conclusions

In this chapter, aqueous solutions of the β-CD/CA and γ-CD/CA complexes with explicit water model have been investigated by MD simulations. The structural features of the host molecules, the positions of the guest molecule, and the hydrogen bonds between the host and guest are focused on. It is concluded that:

(1) The guest molecule, CA, does not escape from the CD cavity but resides in its inside for the simulation time of 5 ns. The RMSD values of the β-CD/CA and

γ-CD/CA complexes from their reference structures are similar, and the fluctuation is somewhat larger for β-CD/CA. The distributions of the diagonal distance and the inter-residue angle indicate that β-CD in β-CD/CA is more rigid than free β-CD or γ-CD in γ-CD/CA.

(2) The O3-C24 distance of CA in γ-CD/CA is similar to that in free CA whereas CA in β-CD/CA adopts a more extended conformation. The position of CA in the CD cavity fluctuates with time; the configuration of the β-CD/CA complex is open/top1 in some periods and open/mid1 in other occasion.

(3) Hydrogen bonds (H-bonds) are not always formed between β-CD and CA during the 5-ns simulation. Instead either β-CD or CA forms H-bonds with water molecules. The 12-OH group, which is located at the middle of CA molecule, forms an H-bond with a β-CD with a higher probability than the other OH and COOH groups in CA. On the other hand, the 3-OH, which is located at the top of the A-ring, is hydrated similarly to that in free CA. The OH groups in β-CD form H-bonds with water as well as intramolecular H-bonds. The hydration structure of the secondary OH groups is affected by complexation with CA.

The thermal fluctuation of the structure and the alteration of the H-bond partners will affect the host-guest interactions in the CD/CA complexes. This issue will be addressed by using ab initio MO calculations in the next chapter.

References and Notes

1 K. B. Lipkowitz, *Chem. Rev.*, **1998**, *98*, 1829.

2 (a) C. Jullian, S. Miranda, G. Zapata-Torres, F. Mendizabal, C. Olea-Azar, *Bioorg. Med. Chem*,. **2007**, *15*, 3217; (b) G. Raffaini, F. Ganazzoli, L. Malpezzi, C. Fuganti, G. Fronza, W. Panzeri, A. Mele, *J. Phys. Chem. B*, **2009**, *113*, 9110; (c) G. Zifferer, B. Sellner, A. Kornherr, D. Krois, U. H. Brinker, *J. Phys. Chem. B*, **2008**, *112*, 710; (d) I. Bea, M. G. Gotsev, P. M. Ivanov, C. Jaime, P. A. Kollman, *J. Org. Chem.* **2006**, *71*, 2056; (e) P. M. Ivanov, C. J. Jaime, *J. Phys. Chem. B*, **2004**, *108*, 6261; (f) M. Jana, S. Bandypadhyay, *J. Phys. Chem. B*, **2013**, *117*, 9280; (g) Y. Yu, C. Chipot, W. Cai, X. Shao, *J. Phys. Chem. B*, **2006**, *110*, 6372.

3 A. J. Scharff, L. E. Rodseth, F. A. Quiocho, *Biochemistry*, **1993**, *32*, 10553.

4 D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, J. Liu, X. Wu, S. R. 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, *AMBER 11*, University of California, San Francisco (**2010**).

5 K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. Gonzalez-Outeirino, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.*, **2008**, *29*, 622.

6 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.*, **2004,** *25*, 1157.

7 W. L. Jorgensen; J. Chandrsekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.*, **1983**, *79*, 926.

8 (a) C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, J. Phys. Chem., **1993**, *97*, 10269. (b) B. H. Besler, K. M. Merz, Jr., P. A. Kollman, *J. Comp. Chem.* **1990**, *11*, 431. (c) U. C. Singh, P. A. Kollman, *J. Comp. Chem.*, **1984**, *5*, 129. (d) Although ESP charges are suitable for description of intermolecular interactions between rigid molecules, usage of RESP charges8a with an appropriate restraint weight is a better choice for CA since the conformational fluctuation of the flexible tail part alters the molecular electrostatic potential. The data in Table 4.4 suggest that the ensemble of β-CD/CA contains a high fraction of conformations more extended than the optimized structure while the ensemble-averaged structure in free CA or γ-CD/CA is similar to the optimized geometry. For simplicity, the ESP charges at the optimized geometry were used instead of determining an optimal restraint scheme to use for evaluation of RESP charges. (e) In the DFT calculation of the electrostatic potential of CA in aqueous media, PCM was employed. Such a treatment is more rational for free CA in aqueous phase than a calculation in the gas phase, even though the PCM is not a high-quality model for description of solvation effects in aqueous solution. Although the effective polarity of space confined in the CD cavity is much lower than that of bulk water, the same atomic charges were used for free CD, β-CD/CA, and γ-CD/CA for consistency because an identical set of atomic charges for the CD part (given in the GLYCAM06 parameters) were used for free β-CD, β-CD/CA, and γ-CD/CA.

9 H. Zhang, C. Ge, D. van der Spoel, W. Feng, T. Tan, *J. Phys. Chem. B,* **2012**, *116*, 3880.

10 U. Essmann, L. Perera, and Max L. Berkowitz, *J. Chem. Phys.*, **1995**, *103(19)*, 8577.

11 J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys*., **1977**, *23*, 327.

12 E. A. Coutsias, C. Seok, K. A. Dill, *J. Comput. Chem.*, **2004**, *25*, 1849.

13 R. Ramos, E. Alvarez-Parrilla, F. Meijide, J. A. Seijas, E. Rodriguez Nunez, J. Vazquez Tato, *Langmuir*, **1999**, *15*, 5489.

Chapter 5

Energy decomposition analyses with FMO-MP2 for CD/CA complexes

5.1 Introduction

Computational investigations on inclusion complexes such as CD/bile acid systems, in which van der Waals interactions play an important role in complexation, remain challenging. The DFT and dispersion-corrected DFT (DFT-D) methods were used in Chapter 3 to investigate the plausible binding modes and the association energies of the CD/bile acid complexes. Post-Hartree-Fock methods[1] such as second-order Mϕller-Plesset perturbation theory (MP2)[2] generally evaluate dispersion interactions with higher accuracy than DFT methods. The fragment molecular orbital (FMO)-MP2 method[3] was applied in the present study to deal with rather large systems.

The focus of Chapter 5 is analyses of contributions to the stabilization of the complexes from different types of intermolecular interactions, particularly the electrostatic and dispersion interactions. The electrostatic and van der Waals interaction energies are also evaluated with molecular mechanics (MM)-based pair potentials for the models of aqueous solution of β-CD/CA. A comparison of the MM-based interaction energies with the results of the FMO-MP2 calculations will reveal the importance of quantum chemical approaches, particularly, consideration of electronic polarization for characterization of intermolecular interactions in the present systems whereas the temporal variation due to thermal fluctuation will be described well with the MM-based approach owing to a merit that a quite large number of conformers can be fed to the energy calculations.

5.2 Computational Details

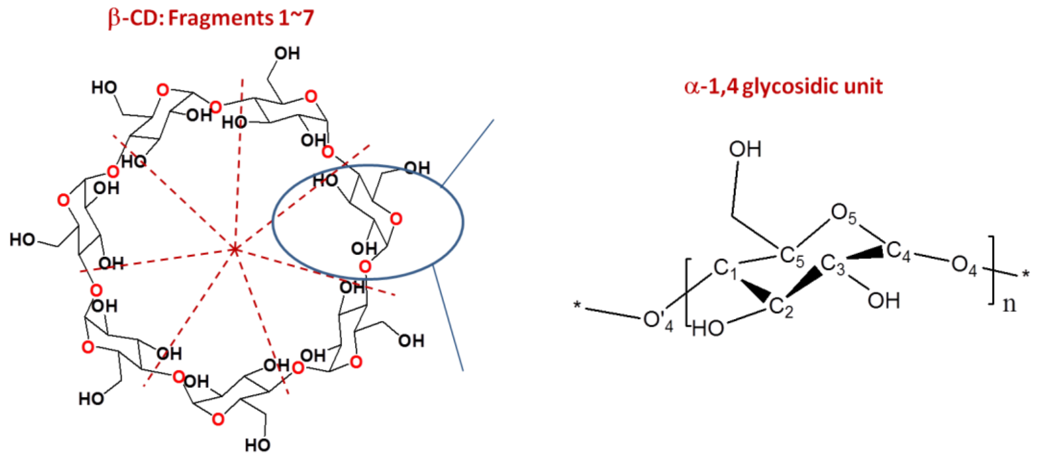
Prior to FMO-MP2 calculations, the energy of β-CD/CA complex was calculated with the normal HF/6-31G(d) and FMO-HF/6-31G(d) methods. The difference was quite small, which confirms that the FMO method can be applied to the present system. A pair interaction energy decomposition analysis (PIEDA)[4] was performed at the FMO-MP2 level of theory in the gas phase, for clusters with explicit H_2O molecules, and in aqueous solution, which was modeled with a conductor-like polarizable continuum (CPCM).[5] The cc-pVDZ[6] and 6-31G(d)[7] basis sets were used. The geometries of the complex in the gas phase and for the PCM-based solution models were optimized at the DFT level with 6-31G(d) as described in Chapter 3. The geometries of the hydrated clusters were generated from snapshots in the classical MD simulations described in Chapter 4 by extracting the coordinates of the complex and the closest water molecules. The FMO-MP2 computations with CPCM[8] were performed with GAMESS 2013 (R1)[9] whereas the gas phase calculations were with GAMESS 2012 (R2)[9]. In both cases, the FMO program version 4.3 was employed.

In an FMO calculation, a target system is segmented into several fragments. On the segmentation of the complex, each α-1,4-glycosidic unit in β- or γ-CD was treated as one fragment, whereas CA was divided into two fragments at the C17–C20 bond, as shown in Chart 5.1(a) – (c). Each H_2O molecule in the hydrated clusters was treated as one fragment. Since two segmentation methods (Method1 and Method2) were possible for CD, the results of PIEDA with these methods were compared.

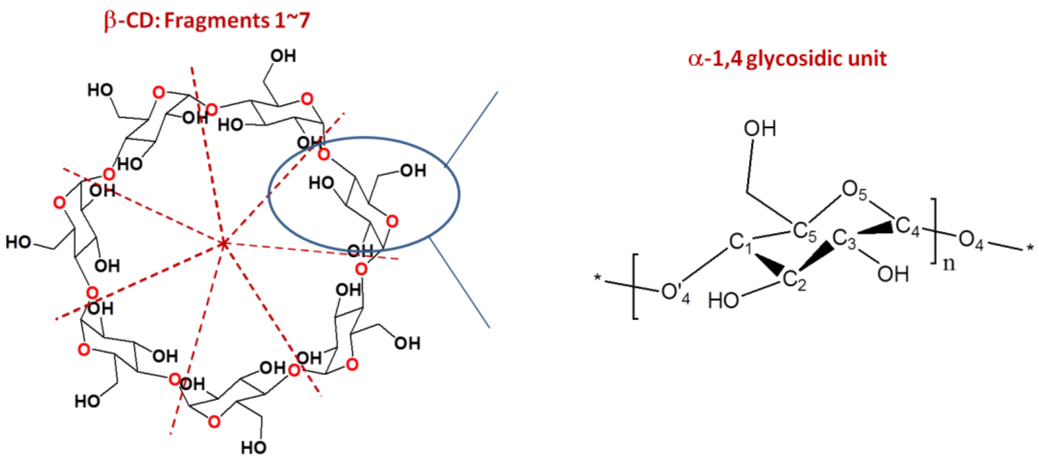
The interaction energies with Method1 and Method2 are shown in Table 5.1. E_{total}, E_{es}, E_{ex}, $E_{\mathrm{ct+mix}}$, and E_{disp} denote the total interaction energy, contributions from electrostatic interaction term, exchange repulsion term, charge-transfer (CT) interaction together with the mixed term, and dispersion interaction term, respectively. E_{ex} ($\Delta E_{\mathrm{ex}} = \Delta E_{\mathrm{ex\,(method1)}} - \Delta E_{\mathrm{ex\,(method2)}} = 1.6$ kcal mol^{-1}) and $E_{\mathrm{ct+mix}}$ ($\Delta E_{\mathrm{ct+mix}} = -3.1$ kcal mol^{-1}) terms are more sensitive to the segmentation method than E_{es} ($\Delta E_{\mathrm{es}} = 0.6$ kcal mol^{-1}) and E_{disp} ($\Delta E_{\mathrm{disp}} = -1.0$ kcal mol^{-1}). Since the main aim of PIEDA in the present study is to analyze the contributions from the electrostatic and dispersion terms, either Method1 or Method2 may be used. In the following part of Chapter 5, all the calculations were performed with Method1.

Table 5.1 Comparison of interaction energies (kcal mol^{-1}) with segmentation Method1 and Method2 at the FMO-MP2/6-31G(d)//B97-D/6-31G(d) level for β-CD/CA in the open/top1 configuration in the gas phase.

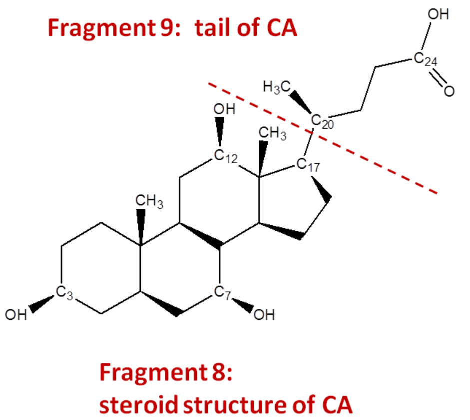
Segmentation method	$\boldsymbol{E_{\mathrm{total}}}$	$\boldsymbol{E_{\mathrm{es}}}$	$\boldsymbol{E_{\mathrm{ex}}}$	$\boldsymbol{E_{\mathrm{ct+mix}}}$	$\boldsymbol{E_{\mathrm{disp}}}$
Method1	–92.4	–80.8	99.3	–41.0	–69.9
Method2	–90.4	–81.4	97.7	–37.9	–68.9



(a) Segmentation Method1of host molecule (β-CD)



(b) Segmentation Method2 of host molecule (β-CD)



(c) Segmentation method of guest molecule (CA)

Chart 5.1 Segmentation schemes for CD and CA.

5.3 Results and Discussion

5.3.1 PIEDA for β-CD/CA and γ-CD/CA complexes in the gas phase

Figure 5.1 shows the interaction energies between CD and CA for the β-CD/CA and γ-CD/CA complexes obtained at the FMO-MP2/cc-pVDZ level of theory. The total interaction energies, E_{total}, for β-CD/CA in open/top1 and γ-CD/CA in open/top1 are more negative than those for β-CD/CA in open/mid2 and γ-CD/CA in open/bottom1. This tendency is similar to that observed in the hypothetical interaction energy calculated by the DFT method as shown in Table 3.4.[10] As is seen in Figure 5.1, both the electrostatic (E_{es}) and dispersion (E_{disp}) terms significantly contribute to the total interaction energies (E_{total}). The dispersion term contributes to almost the same extent as the electrostatic term although dispersion forces are generally considered to be weak. The energy of the CT interaction and mixed term ($E_{\mathrm{ct+mix}}$) also has a significant contribution to E_{total}. The physical meaning of $E_{\mathrm{ct+mix}}$ is, however, not so clear because $E_{\mathrm{ct+mix}}$ obtained with PIEDA, which is calculated by eq. (5-1), contains not only the charge-transfer interaction energy but also various coupling terms.[11] Thus, only the E_{total}, E_{es}, and E_{disp} values are discussed here since the physical meanings of these terms are easily understood.

$$E_{\mathrm{ct+mix}} = E_{\mathrm{total}} - E_{\mathrm{es}} - E_{\mathrm{ex}} - E_{\mathrm{disp}} \tag{5-1}$$

The most stable configuration, open/top1, of β-CD/CA exhibits the most negative total and electrostatic interaction energies. The dispersion term in β-CD/CA is more negative than that in γ-CD/CA, which is attributed to the difference in the contact between CD and CA; some of the sugar units in γ-CD have no van der Waals contact with CA, as shown in Figure 5.2.

Figure 5.3 shows the interactions for each pair of fragments for β-CD/CA in the open/top1 configuration obtained by the FMO-MP2/cc-pVDZ calculation. The warm and cold colors represent destabilizing and stabilizing pair interactions, respectively. The total pair interactions are attractive for any pair between β-CD and CA. The pairs #8–#2, #8–#3, and #9–#2 have the largest contribution to the electrostatic term (E_{es} = −15.4, −30.5, and −16.7 kcal mol^{-1}, respectively) and the total interaction energy (E_{total} = −20.3, −28.8, and −15.0 kcal mol^{-1}, respectively). These pairs involve one, two, and one intermolecular hydrogen bond(s), respectively, as shown in Table 3.6. Significant dispersion interactions are also seen for pairs #8–#1 (E_{disp} = −10.3 kcal mol^{-1}), #8–#2

(E_{disp} = −10.4 kcal mol^{-1}), and #8–#3 (E_{disp} = −15.1 kcal mol^{-1}), which involve a number of van der Waals contacts between the medial curved surface of CA and the inside wall of the β-CD cavity, as seen in Figure 5.2.

Figure 5.1 also indicated that the direction of CA has a significant effect on the electrostatic term (E_{es}): Direction1 gives a more negative value of E_{es} than direction2 since the former direction forms more hydrogen bonds with CD than the latter one (see Table 3.6).

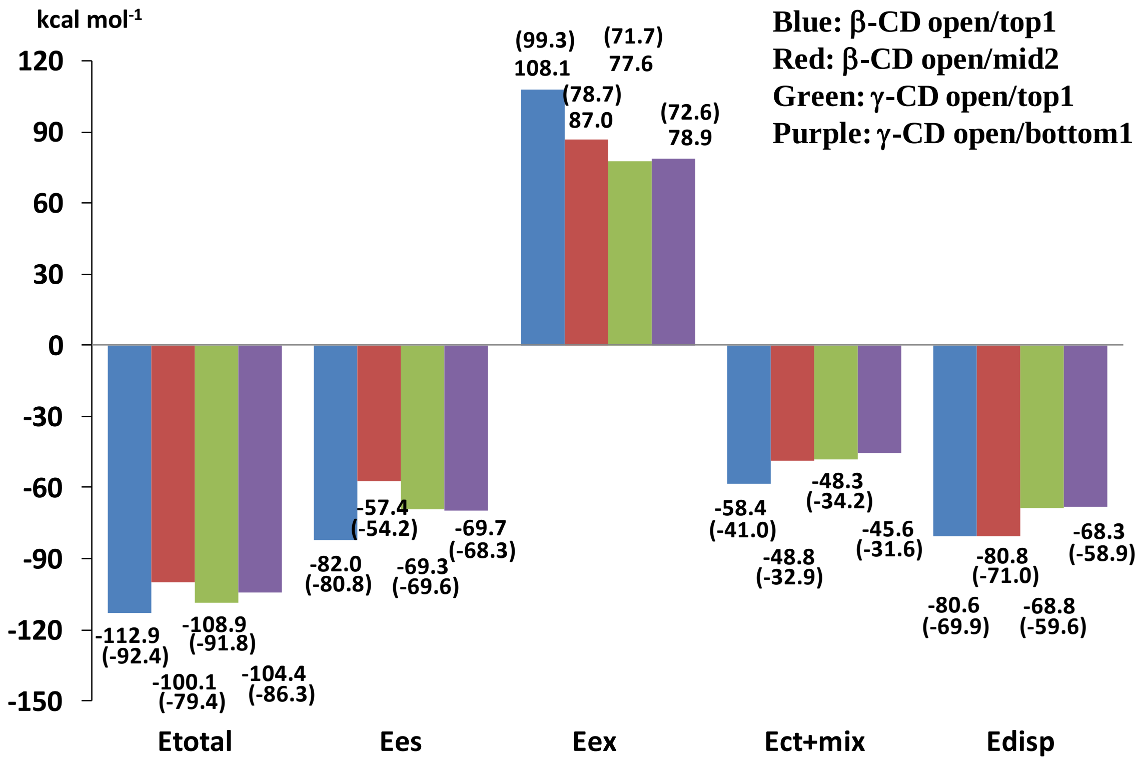
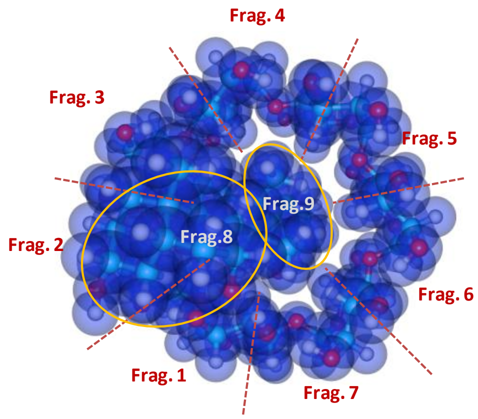
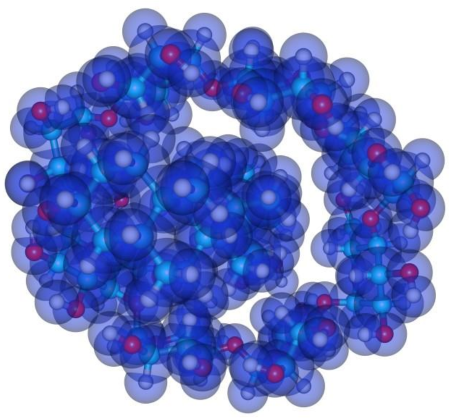


Figure 5.1 Comparison of interaction energies between CD and CA for β-CD/CA in open/top1 (blue), β-CD/CA in open/mid2 (red), γ-CD/CA in open/top1 (green), and γ–CD/CA in open/bottom1 (violet) obtained by the FMO-MP2/cc-pVDZ //B97-D/6-31G(d) (FMO-MP2/6-31G(d)//B97-D/6-31G(d) in the parentheses) method in the gas phase. The interaction energies were evaluated as sums of pair interaction energies, $\Sigma_{i \in \mathrm{host}} \Sigma_{j \in \mathrm{guest}} E_{ij}$, where E_{ij} is the pair interaction energy between the i-th fragment belonging to the host and the j-th fragment belonging to the guest.

Before ending this subsection, effects of the basis set on the interaction energies will be considered. The results with the 6-31G(d) basis set, which are shown in parentheses in Figure 5.1, showed the same tendencies as those with cc-pVDZ; however, the values of E_{total}, E_{es}, E_{ex}, $E_{\mathrm{ct+mix}}$, and E_{disp} for β-CD/CA open/top1 with 6-31G(d) amount 82%, 99%, 92%, 70%, and 87%, respectively, of the corresponding values with cc-pVDZ. The difference in the basis has set little effect on the electrostatic term (E_{es}). These results suggest that a comparison of the interaction energies obtained with 6-31G(d) among these four complexes provides sufficiently accurate information for the present study, in which the electrostatic and dispersion interactions are focused on. In the viewpoint of computational cost, the CPU time consumed with cc-pVDZ was 26.5 h by GAMESS 2012 version at Linux computer (Intel Xeon5500, 1.3GHz) in our laboratory for the γ-CD/CA complex while 6-31G(d) took only 7.2 h for the same complex. Thus, the following calculations were performed with the 6-31G(d) basis set.



β-CD/CA



γ-CD/CA

Figure 5.2 Space-filling models drawn with the van der Waals radii for β-CD/CA (left) and γ-CD/CA (right) complexes in the open/top1 configuration whose geometries were optimized by the B97-D/ 6-31G(d) calculation in the gas phase.

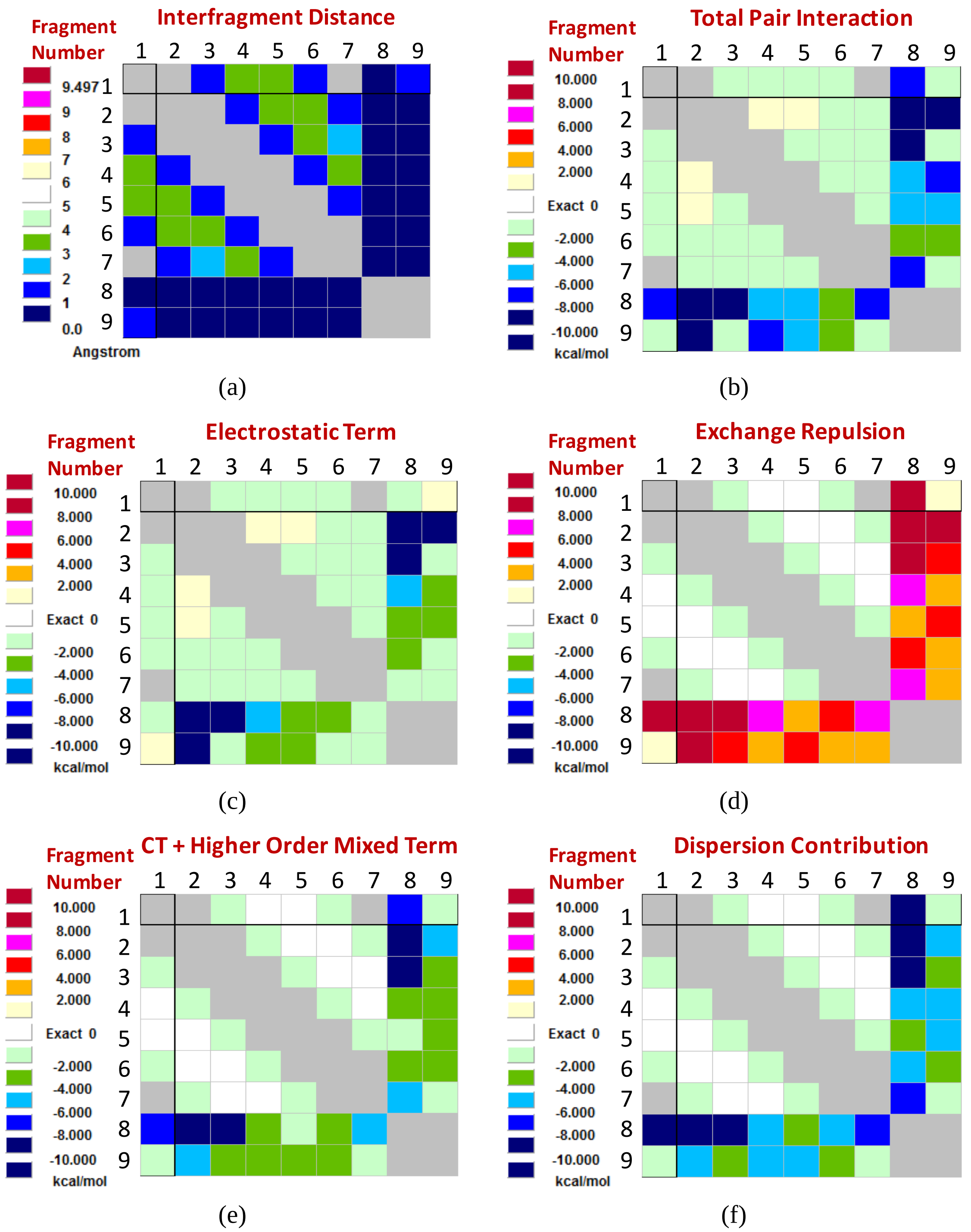


Figure 5.3 Inter-fragment distance (a) and pair interaction energies (b-f) indicated by its color for β-CD/CA in the open/top1 configuration obtained by the FMO-MP2/cc-pVDZ method at the B97-D/6-31G(d)-optimized geometry in the gas phase.

5.3.2 Effects of the DFT functionals used for the geometry optimization

As mentioned in section 5.2, the FMO-MP2 single-point calculations were performed at the geometries optimized by the DFT methods. In order to examine the sensitivity of the interaction energies on the geometry, the results of PIEDA were compared among the geometries optimized with the different functionals used in Chapter 3.

Table 5.2 shows the results of PIEDA obtained with the FMO-MP2/6-31G(d) method at the geometries optimized by the B97-D and M06-2X functionals in the gas phase. Although the absolute values are somewhat different from each other, the tendencies in the relative magnitudes of contributions and in the order of magnitudes among the three complexes are quite similar. Furthermore the E_{disp} terms are almost the same for the two geometries.

As was discussed in Chapter 3, the geometry optimization with the B3LYP functional gave an energy minimum which was rather different from that with the B97-D or M06-2X functional (See Figure 3.5). The results of PIEDA at the geometries optimized with the B97-D, M06-2X, and B3LYP functionals with PCM are shown in Figure 5.4. Obviously, E_{disp} for the B3LYP-optimized geometry is less negative than that for the B97-D- or M06-2X-optimized one. This result can be explained by the difference in the position of the head part in CA as shown in Figure 3.5.

Table 5.2 Interaction energies (kcal mol^{-1}) for CD/CA complexes obtained with FMO-MP2/ 6-31G(d) in the gas phase at the geometries optimized by the B97-D and M06-2X functionals with 6-31G(d) in the gas phase.

Host	**Configuration**	**Method**	$\boldsymbol{E}_{\mathbf{total}}$	$\boldsymbol{E}_{\mathbf{es}}$	$\boldsymbol{E}_{\mathbf{ex}}$	$\boldsymbol{E}_{\mathbf{ct+mix}}$	$\boldsymbol{E}_{\mathbf{disp}}$
β-CD	open/top1	B97-D	–92.4	–80.8	99.3	–41.0	–69.9
		M06-2X	–101.8	–97.1	104.9	–40.5	–69.2
β-CD	open/mid2	B97-D	–79.4	–54.2	78.7	–32.9	–71.0
		M06-2X	–73.6	–59.7	88.4	–31.9	–70.3
γ-CD	open/top1	B97-D	–91.8	–69.6	71.7	–34.2	–59.6
		M06-2X	–89.5	–68.2	73.5	–34.0	–60.8

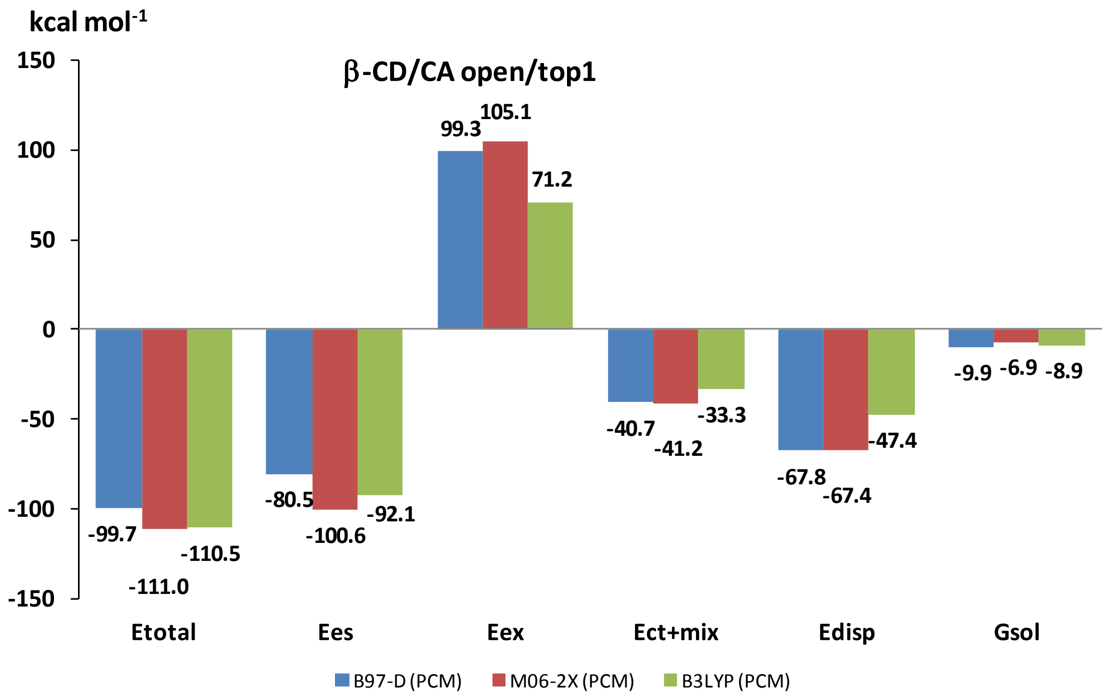


Figure 5.4 Interaction energies for β-CD/CA in the open/top1 configuration obtained by the FMO-MP2/6-31G(d) with PCM for the geometries optimized with the B97-D (blue), M06-2X (red), and B3LYP (green) functionals. The geometry optimization was performed using the 6-31G(d) basis set with PCM.

5.3.3 Effects of solvation with an implicit solvation model

Table 5.3 shows the interaction energies with the FMO-MP2/6-31G(d) method both in the gas phase and with PCM. The solvation term (G_{sol}) for the β-CD/CA complex is more negative than that for the γ-CD/CA complex. The E'_{total} values in the parentheses are the summation without G_{sol} term. The difference in the total interaction energy between the gas phase and PCM-based aqueous phase almost comes from G_{sol} (E_{total} = –92.4, E_{total}' = –92.2 for β-CD/CA in open/top1; E_{total} = –79.4, E_{total}' = –79.8 for β-CD/CA in open/mid2). The contributions from the other terms are quite similar. The solvation effect on the host–guest interactions is not clear with the PCM-based consideration.

Table 5.3 Interaction energies (kcal mol^{-1}) obtained with FMO-MP2/6-31G(d) with or without PCM for CD/CA complexes. The geometries were optimized at the B97-D/6-31G(d) level in the gas phase or with PCM.

Configuration	**Phase (opt.)**	**Phase (FMO)**	E_{total} (E'_{total}) [a]	E_{es}	E_{ex}	E_{ct+mix}	E_{disp}	G_{sol} [b]
β-CD/CA	Gas	Gas	–92.4	–80.8	99.3	–41.0	–69.9	0.0
open/top1	Gas	PCM	–101.2 (–92.2)	–81.1	99.9	–41.0	–70.0	–9.0
	PCM	PCM	–99.7 (–89.8)	–80.5	99.3	–40.7	–67.8	–9.9
β-CD/CA	Gas	Gas	–79.4	–54.2	78.7	–32.9	–71.0	0.0
open/mid2	Gas	PCM	–89.9 (–79.8)	–54.9	78.8	–32.5	–71.3	–10.1
	PCM	PCM	–90.8 (–79.6)	–57.2	80.7	–33.1	–70.0	–11.2
γ-CD/CA	Gas	Gas	–91.8	–69.6	71.7	–34.2	–59.6	0.0
open/top1	Gas	PCM	–97.3 (–90.3)	–69.0	71.8	–33.4	–59.7	–7.0
γ-CD/CA	Gas	Gas	–86.3	–68.3	72.6	–31.6	–58.9	0.0
open/bottom1	Gas	PCM	–91.3 (–86.0)	–68.4	72.6	–31.2	–59.1	–5.3

[a] $E'_{total} = E_{es} + E_{ex} + E_{ct+mix} + E_{disp.}$ [b] Solvation term evaluated as a free energy change based on PCM.

5.3.4 Effects of thermal motion and solvation with an explicit solvation model

In this subsection, attention will be paid to effects of thermal motion and solvation with an explicit solvation model. The CD/CA complex in an aqueous solution varies the relative position of CA with respect to CD as well as the conformation of each component with time as demonstrated in Chapter 4. Such structural fluctuation modulates the host-guest interactions.

Figure 5.5 shows time evolutions of the electrostatic, Lenard-Jones, and total interaction energies between CA and β-CD in the 5-ns MD run. The interaction energies show large random fluctuation. The variation in the electrostatic term is related to the occasional formation and breaking of hydrogen bonds between β-CD and CA as shown in Figure 5.6. In this analysis, the predetermined atomic charges were used for calculation of the Coulomb potentials, no polarization effects being taken into consideration; the stabilization effect is hence underestimated. In order to analyze the host-guest interactions with higher accuracy, the FMO-MP2/6-31G(d) method was used, and the results were compared with those obtained for the DFT-optimized geometries in the previous subsection.

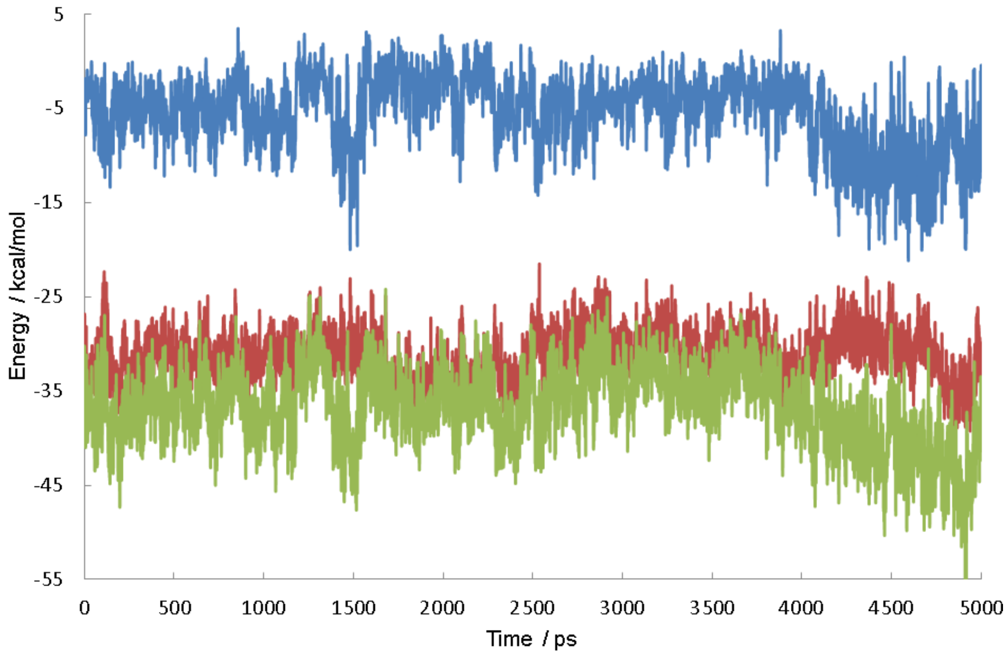


Figure 5.5 Time evolution of the electrostatic (blue), Lenard-Jones (red), and total interaction (green) energies between β-CD and CA in the MD simulation of β-CD/CA complex for 5 ns in Chapter 4. The energies were calculated every 1 ps with the same fractional charges and van der Waals parameters as used in the MD run.

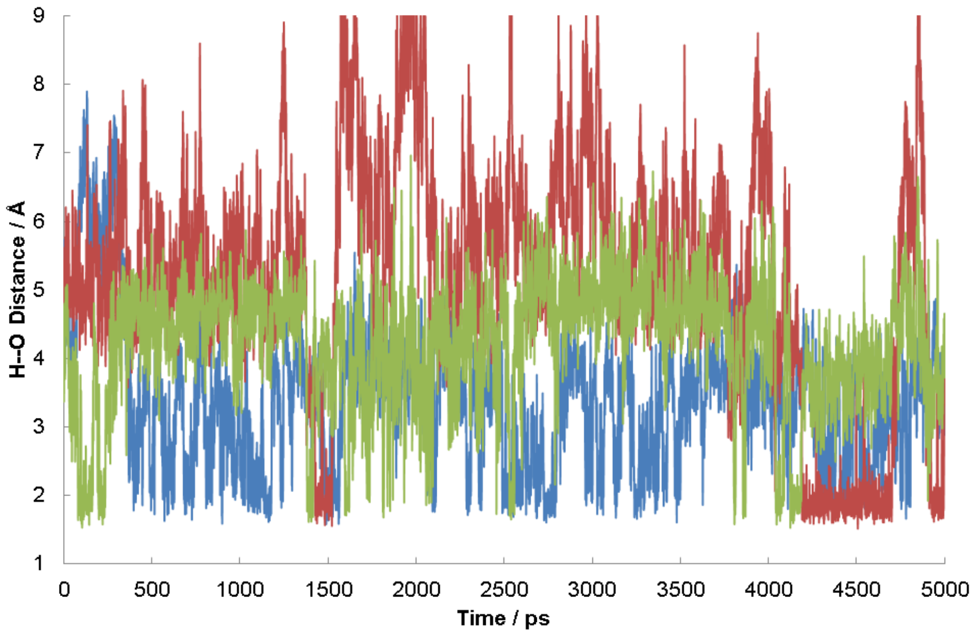


Figure 5.6 Time evolution of the shortest 12-OH···O* distances (blue; O* denotes any oxygen atom of the secondary OH groups in β-CD), the COOH···O6(#4) (red), and the shortest C=O···H* (green; H* denotes any hydrogen atom of the primary OH groups in β-CD) during the 5-ns MD simulation for β-CD/CA. The H-bond partner was switched several times for the 12-OH and C=O groups in CA.

The effects of aqueous environment on the host-guest interactions should also be considered. As an explicit solvation model, clusters consisting of the complex and surrounding water molecules were used here. In the snapshot at 5 ns in the MD simulation of β-CD/CA, the first solvation shell contains 106 water molecules and exhibits an almost spherical shape with a diameter of about 21 Å as illustrated in Figure 5.7. Here the first solvation shell was defined as a region within 3.4 Å from any of the solute atoms, which was the default criterion of the ptraj program in the AMBER11 package.[12] Three models were generated from the snapshot at 5 ns: the bare complex without any water molecules (5ns_noWAT), a small cluster with 20 water molecules closest from any oxygen atoms in the complex (5ns_20WAT), and a larger cluster with 106 water molecules (5ns_106WAT). Most of the water molecules in the 5ns_20WAT model form H-bonds with β-CD and/or CA.

FMO-MP2/6-31G(d) calculations were carried out for the three models. The results of PIEDA are listed in Table 5.4. In the FMO method, the SCF calculation of each monomer or dimer is performed in an external electric field generated by all the other fragments, that is, the electron density of the target monomer or dimer is altered by the presence of other fragments. Nevertheless, the incorporation of water molecules little affected the CD–CA interaction energies in a direct manner.[13] In the 5ns_20WAT model, most of the water molecules are in the peripheral region of the complex as shown in Figure 5.8; hence the effect on the CD–CA contact area is likely small. The contact area is also shielded from water in the first solvation shell by the outer wall of the cavity. Although models without water could be used, the models with the 20 closest water molecules were used for subsequent calculations.

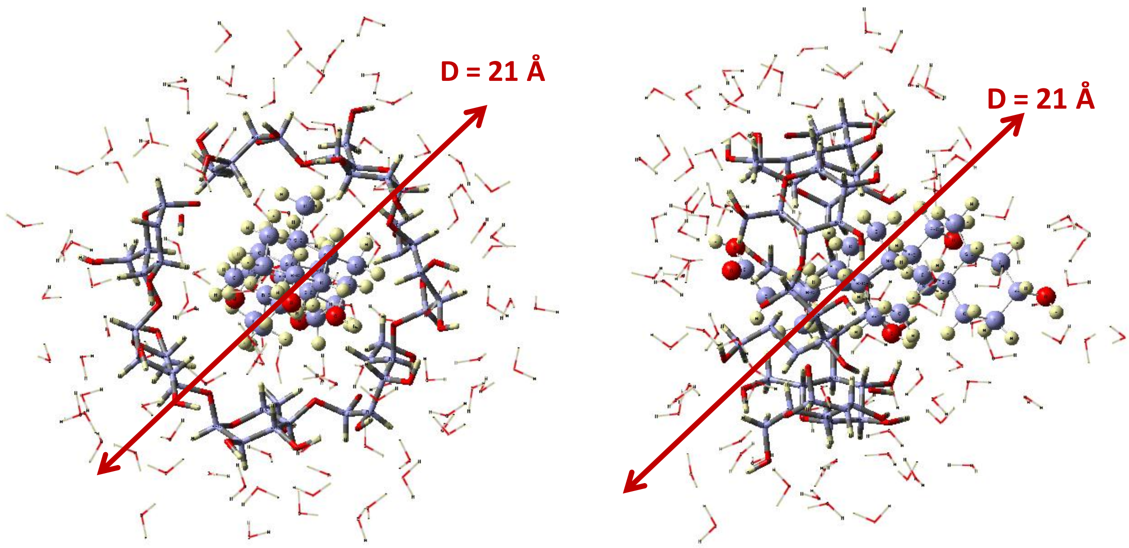


Figure 5.7 Two different views for the snapshot of β-CD/CA complex with 106 water molecules at 5 ns. The hydrogen, oxygen, and carbon atoms are drawn in light gray, red, and bluish gray, respectively. The red arrow denotes the region of water sphere.

Table 5.4 Interaction energies (kcal mol^{-1}) between β-CD and CA for the three models generated from the snapshot at 5 ns in the MD simulation for the β-CD/CA complex.

Models	E_{total}	E_{es}	E_{ex}	E_{ct+mix}	E_{disp}
5 ns_noWAT	–34.9	–12.6	32.1	–15.1	–39.4
5 ns_20WAT	–35.2	–12.8	32.1	–15.2	–39.3
5 ns_106WAT	–34.1	–11.6	32.1	–15.2	–39.4

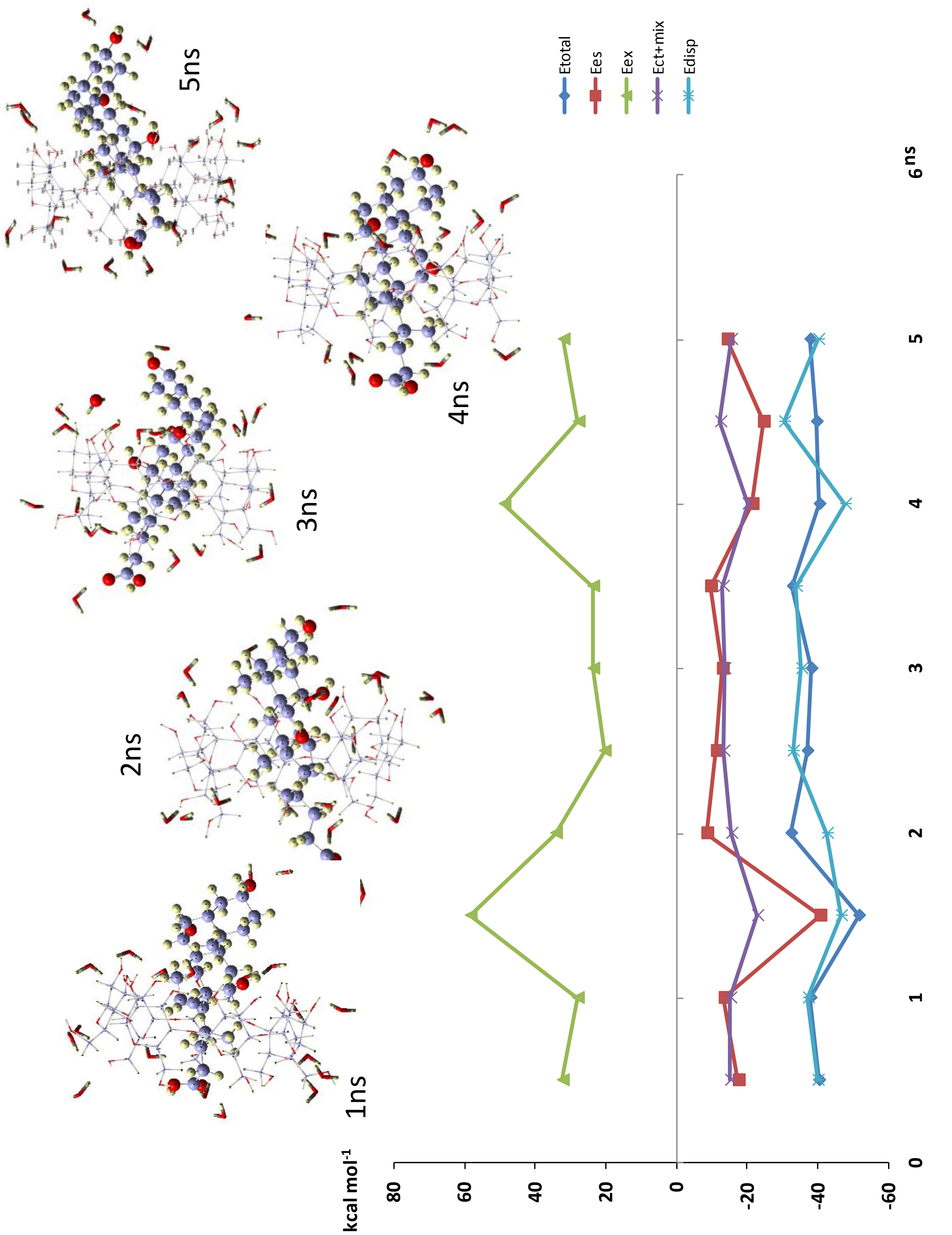


Figure 5.8 Variation in the interaction energies for the snapshots of β-CD/CA complex with 20 closest water molecules from 0.5 ns to 5 ns. The calculations were performed at the FMO-MP2 /6-31G(d) level. Straight lines are drawn merely for the aid of eyes.

Since the interaction energies between CA and β-CD significantly fluctuate at a relatively short timescale due to thermal motions as seen in Figure 5.5, it is desirable to perform calculations for a large number of snapshots in order to obtain statistical averages of the total and individual interaction energies. However, only ten snapshots were arbitrarily selected with an interval of 0.5 ns from the MD trajectory[14] because of high computational costs of the FMO-MP2 calculations. The variations of the interaction energies obtained with the FMO-MP2/6-31G(d) method are shown in Figure 5.8 together with some of the snapshots.

The total interaction energy and the contributions from the individual terms are fluctuated as was predicted by the preliminary analysis shown in Figure 5.5. The electrostatic interaction (E_{es}) for each snapshot (–8.5 to –40.6 kcal mol^{-1}) is more negative than the corresponding value (–5.5 ± 3.8 kcal mol^{-1}) predicted as the sum of Coulombic pair potentials with the predetermined atomic charges. This result clearly indicates that it is essential to consider the electronic polarization effects for accurate evaluation of the electrostatic interaction energy. The E_{es} value for the snapshot at 1.5 ns is more negative than that for the other snapshots. As shown in the right panel of Figure 5.9, strong electrostatic interactions are seen in the #8-#7 (E_{es} = –10.5 kcal mol^{-1}) and #9-#4 (E_{es} = –17.1 kcal mol^{-1}) pairs. Hydrogen bonds are formed between the fragments in these pairs as illustrated in Figure 5.10. It was confirmed that the electrostatic term becomes more negative when hydrogen bonds are formed between β-CD and CA.

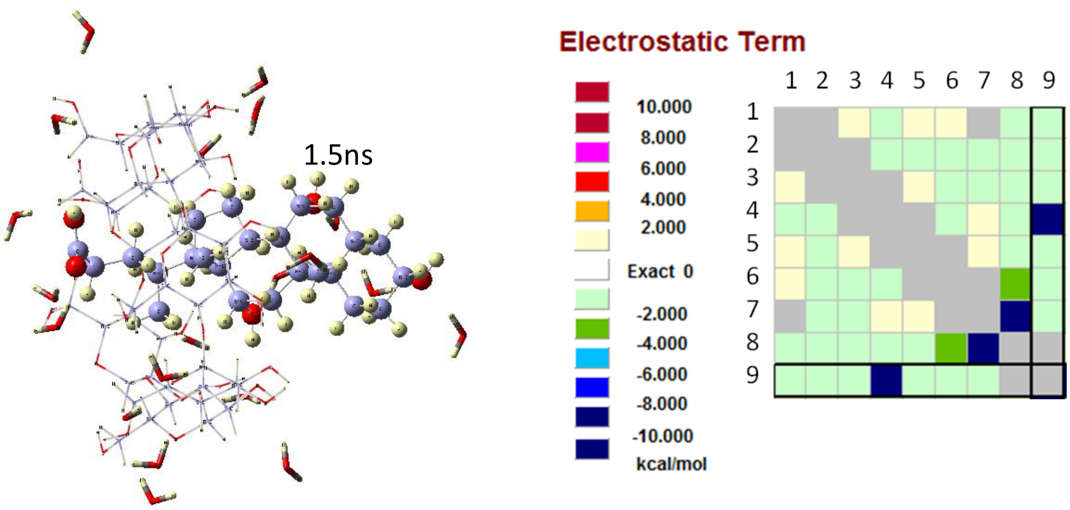
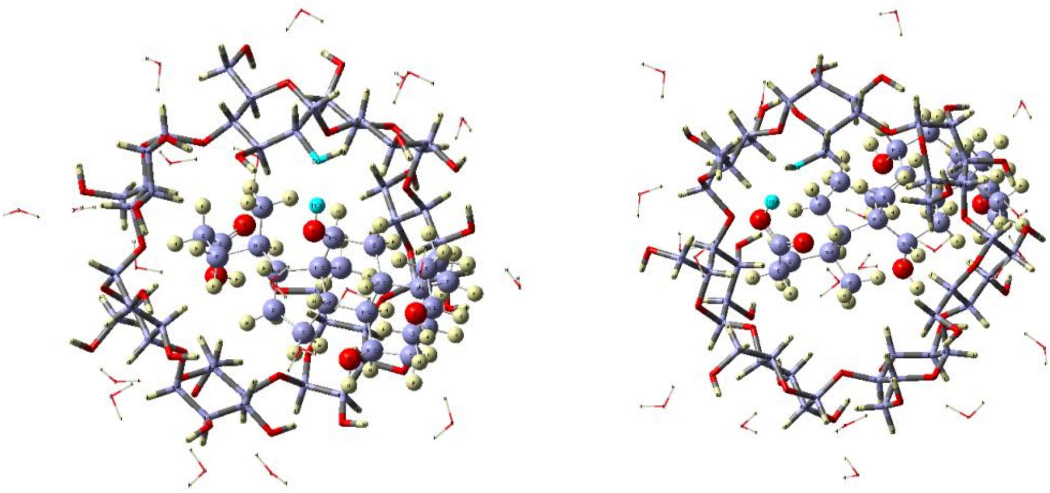


Figure 5.9 (Left) The structure of the complex in the snapshot at 1.5 ns. The hydrogen, oxygen, and carbon atoms are drawn in light gray, red, and bluish gray, respectively. (Right) Color chart of the electrostatic interaction energy for each pair of fragments.



#8–#4 $r_{H\ldots O}$ =1.80Å   #9–#4 $r_{H\ldots O}$ =1.72Å

Figure 5.10 Two different views of β-CD/CA in the snapshot at 1.5 ns. The atoms involving hydrogen bonds between β-CD and CA are drawn in turquoise. The other hydrogen, oxygen, and carbon atoms are drawn in light gray, red, and bluish gray, respectively.

5.4 Conclusions

The results of PIEDA have clarified the contributions from different types of intermolecular interactions to the stabilization of the CD/CA complexes as follows:

(1) Dispersion forces (E_{disp}) play an important role in the formation of CD/CA complexes, almost as important as the electrostatic term (E_{es}). The β-CD/CA complex shows more negative values of E_{disp} than the γ-CD/CA complex because the van der Waals contacts between the host and guest molecules in β-CD/CA are larger than those in γ-CD/CA.

(2) The direction of CA has a significant effect on the electrostatic term (E_{es}): Direction1 gives a more negative value of E_{es} than direction2 since the former direction forms more hydrogen bonds with CD than the latter one.

(3) The total interaction energy (E_{total}) is the most negative in the open/top1 configuration of β-CD/CA, which was predicted to be the most stable by the DFT calculations in Chapter 3. This point validates the applicability of the DFT methods, particularly the B97-D and M06-2X functionals, to study on inclusion complexes such as the present CD/bile acid systems.

(4) The interaction energies calculated for the hydrated cluster model are rather

different from those for the complex in the gas phase. The dispersion term (E_{disp}) shows a larger contribution to the total interaction energy (E_{total}) than the electrostatic term (E_{es}) for the hydrated clusters, which was arbitrarily sampled from an MD trajectory of β-CD/CA. The hydrogen bonds between CD and CA are frequently broken and formed in aqueous solution, and the number of H-bonds is varied from 0 to 2. The decrease in the number of H-bonds is a main reason for the reduction of the contribution from E_{es} for the hydrated cluster as compared with that for the complex in the gas phase.

(5) The comparison of the electrostatic energies evaluated with the MM-based pair potential and those with the FMO-MP2 calculations revealed that it is essential to consider electronic polarization accompanied with the complexation for evaluation of accurate interaction energies.

References and Notes

1 (a) J. Cramer, Christopher, *Essentials of Computational Chemistry*, **2002**, John Wiley & Sons. Ltd. (b) Jensen F. *Introduction to Computational Chemistry 2nd edition.* **2006**. John Wiley & Sons. Ltd.

2 C. Mϕller, M. S. Plesset, *Phys. Rev.*, **1934**, *46*, 618.

3 (a) D. G. Fedorov, K. Kitaura, *J. Chem. Phys.*, **2004**, *120*, 6832. (b) D. G. Fedorov, K. Kitaura , *J. Chem. Phys.*, **2004**, *121*, 2483. (c) D. G. Fedorov, K. Kitaura, *J. Phys. Chem. A*, **2007**, *111*, 6904. (d) D. G. Fedorov, T. Nagata, K. Kitaura, *Phys. Chem. Chem. Phys.*, **2012**, *14*, 7562.

4 (a) D. G. Fedorov, K. Kitaura, *J. Comp. Chem.*, **2007**, *28*, 222. (b) D. G. Fedorov, K. Kitaura, *J. Chem. Phys.*, **2012,** *116*, 704.

5 (a) V. Barone, M. Cossi, *J. Phys. Chem. A*, **1998**, *102*, 1995. (b) M. Cossi, N. Rega, G. Scalmani, V. Barone, *J. Comput. Chem.*, **2003**, *24*, 669.

6 (a) T. H. Dunning Jr., *J. Chem. Phys.*, **1989,** *90*, 1007. (b) T. van Mourik, T. H. Dunning Jr., *J. Mol. Struct. (THEOCHEM)*, **1996**, *388*, 339.

7 (a) R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.*, **1971**, *54* , 724. (b) W. J. Hehre, R. Ditchfield, J. A. Pople, *J. Chem. Phys.*, **1972**, *56*, 2257. (c) P. C. Hariharan, J. A. Pople, *Theor. Chem. Acc.*, **1973**, *28*, 213.

8 D. G. Fedorov, K. Kitaura, H. Li, J. H. Jensen, M. S. Gordon. *J. Comp. Chem.*, **2006**, *27*, 976.

9 M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. J.

Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem.*, **1993**, *14*, 1347.

10 The meaning of the total interaction energy (E_{total}) obtained by the FMO-MP2 method is different from that of the hypothetical interaction energy (E_{intrxn}) evaluated according to eq. 3.4 in Chapter 3 since the energy of each fragment is calculated as the polarized state in the external field generated by the other fragments in the FMO method; such polarization destabilizes the fragment, and the stabilization accompanied with pair formation is exaggerated.

11 A quantity of charge transferred from a fragment to another is evaluated in the PIEDA method. The partial charge transferred from CA to CD were evaluated to be 0.14, 0.20, 0.11, and 0.14 for β-CD/CA (open/top1), β-CD/CA (open/mid2), γ-CD/CA (open/bottom1), and γ-CD/CA (open/top1), respectively, with the FMO-MP2/cc-pVDZ method. The positive value means that the CA part exhibits a negative partial charge while the CD part a positive one in the complex. The transferred charge in the cluster models are smaller, ranging from 0.03 to 0.12 with the FMO-MP2/6-31G(d) calculations. These results suggest that the charge-transfer interaction between CA and CD is not so strong but non-negligible.

12 D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, *AMBER*, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp. Phys. Commun.*, **1995**, *91*, 1.

13 There are a few water molecules that exhibit non-negligible electrostatic attraction both with β-CD and with CA. Such an indirect interaction via a single water molecule would be evaluated with the three-body FMO method, so-called FMO3-method, which is much more computationally demanding than the two-body FMO method.

14 Although it is better use the structure averaged over ~0.1 ps rather than a single snapshot, the coordinates were damped every 0.1 ps in the present MD simulation since a shorter damping interval (e.g. 1 fs) would require a huge volume of disk.

Chapter 6

General overview

This study mainly describes the interactions between the host molecules (α-, β-, and γ-CDs) and the guest molecules (CA, DCA and their deprotonated forms) by using computational methods. Specifically, (1) the stable structures and the association energies have been studied for the CD/bile acid inclusion complexes by the density functional theory calculations mainly in the gas phase; (2) the structural features of the host molecules, the positions of the guest molecule, and the hydrogen bonds between the host and guest of β-CD/CA and γ-CD/CA complexes in aqueous solution have been investigated by molecular dynamics simulations with explicit water model; and (3) the contributions from different types of intermolecular interactions to the stabilization of the CD/bile acid complexes have been studied by fragment molecular orbital calculations at the MP2 level. The following conclusions have been drawn from the current study.

DFT level of study

According to the DFT calculations, the closed conformation is more stable than the open conformation for free α-, β-, and γ-CDs in the gas phase whereas the open conformation is more favored in the CD/bile acid complexes. β-CD and γ-CD are more favored than α-CD as a host molecule in CD/bile acid complexes. The association energy depends on the cavity size of CD, as well as on the position and direction of the guest. Intermolecular hydrogen bonds are formed between the host and guest molecules in the gas phase. For the β-CD/CA complex, the open/top1 and open/mid2 configurations are energetically favored, which is in agreement with experimental observations in aqueous solution. The comparison of the optimized structure and the association energy for the β-CD/CA complex obtained with B97-D, M06-2X, and B3LYP indicates that the accurate evaluation of dispersion forces is important in DFT calculations.

MD simulations

MD simulations were performed for aqueous solution of β-CD/CA or γ-CD/CA at 298 K by staring from the most stable open/top1 configuration for 5 ns, during which the guest molecule, CA, resided in the CD cavity. The analyses of the diagonal distance and the inter-residue angle indicate that β-CD in β-CD/CA is more rigid than free β-CD or γ-CD in γ-CD/CA. The length of CA in γ-CD/CA measured by the O3–C24 distance is similar to that in free CA whereas CA in β-CD/CA adopts a more extended conformation. The position of CA in the CD cavity fluctuates resulting in the change in the configuration of the β-CD/CA complex with time. Intermolecular hydrogen bonds are not always formed between CD and CA during the 5-ns simulation. Instead, most of the OH groups in CD and the OH and COOH groups in CA form hydrogen bonds with water similarly to the individual host and guest species in the free state.

FMO-MP2 level of study

The results of PIEDA indicate that dispersion force have a large contribution to the stabilization of the CA/bile acid complex. The contributions from the electrostatic and dispersion terms are almost the same in the gas phase. The β-CD/CA complex shows more negative values of dispersion force than the γ-CD/CA complex because the van der Waals contacts between the host and guest molecules in β-CD/CA are larger than those in γ-CD/CA. The direction of CA has a significant effect on the electrostatic term. The total interaction energy is the most negative in the open/top1 configuration of β-CD/CA, which was predicted to be the most stable by the DFT calculations. The dispersion term in hydrated cluster model shows a larger contribution to the total interaction energy than the electrostatic term. Since most of the hydrogen bonds between CD and CA are replaced by the hydrogen bonds with water in aqueous solution, the electrostatic interactions are reduced as compared with those in the gas phase. The comparison of the electrostatic energies evaluated with the molecular mechanics (MM)-based pair potential and those with the FMO-MP2 calculations revealed that it is essential to consider electronic polarization accompanied with the complexation for evaluation of accurate interaction energies.

Supporting Information

Table S1. Absolute electronic energies (a.u.) for complexes, free hosts, and free guest in the gas phase calculated with the B97-D/6-31G(d) method.

CD Config.	CA position	E_{complex}	$E_{\text{free-host}}$	$E_{\text{free-CA}}$	BSSE
α-CD (open)	top2	-4976.191723	-3662.181788	-1313.905002	0.042216
	mid1	-4976.182836	-3662.181788	-1313.905002	0.037555
β-CD (open)	top1	-5586.584620	-4272.548827	-1313.905002	0.047258
	mid2	-5586.577702	-4272.548827	-1313.905002	0.046295
β-CD (closed)	mid1	-5586.583385	-4272.567737	-1313.905002	0.051272
γ-CD (open)	bottom1	-6196.934510	-4882.911014	-1313.905002	0.042864
	top1	-6196.936760	-4882.911014	-1313.905002	0.044916
γ-CD (closed)	top2	-6196.938834	-4882.918512	-1313.905002	0.049785

Table S2. Absolute electronic energies (a.u.) for complexes, free hosts, and free guest in the gas phase calculated with the M06-2X/6-31G(d) method.

CD Config.	CA position	$E_{complex}$	$E_{free\text{-}host}$	$E_{free\text{-}CA}$	BSSE
α-CD (open)	top2	-4977.443538	-3663.139012	-1314.225003	0.030959
	mid1	-4977.435334	-3663.139012	-1314.225003	0.031716
β-CD (open)	top1	-5587.984920	-4273.666495	-1314.225003	0.0390136
	mid2	-5587.976651	-4273.666495	-1314.225003	0.035594
β-CD (closed)	mid1	-5588.001273	-4273.692920	-1314.225003	0.042611
γ-CD (open)	bottom1	-6198.514717	-4884.188287	-1314.225003	0.0370173
	top1	-6198.508949	-4884.188287	-1314.225003	0.035864
γ-CD (closed)	top2	-6198.519317	-4884.206826	-1314.225003	0.047592

The inner diameter of β-CD is evaluated by eq. S1:

$$d_{\beta\text{-}CD} = \frac{1}{2}\, r_{BC}\,(\tan\Theta + \cot\Theta) \qquad \text{(S1)}$$

where r_{BC} is the H3…H3' for d_{up}, H6…H6' for d_{down} in the open conformation, H6O…H6O' for d_{down} in the closed conformation, $\Theta = 12.85^{\circ}$ for both the open and closed conformations, in which the primed atoms belong to the neighboring sugar unit. Detailed information is given in Figure S1.

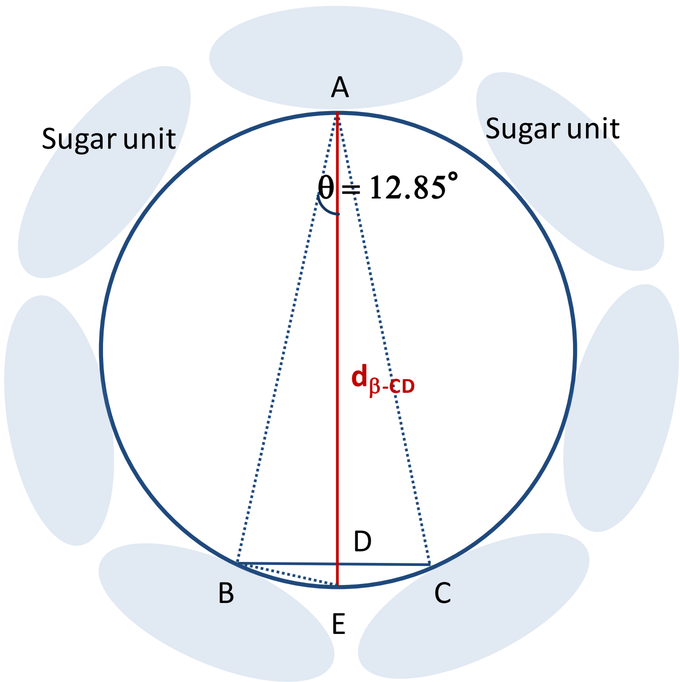


Figure S1. Schematic diagram of the definition of inner diameter of β-CD for both the open and closed conformers.

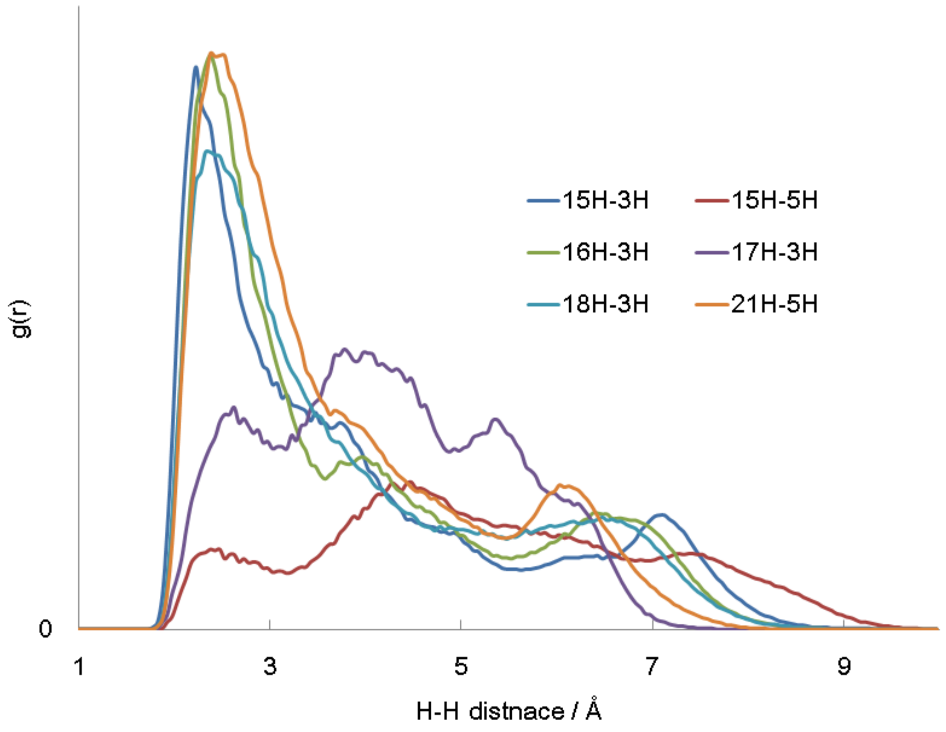


Figure S2. Radial distribution functions for 3-H and 5-H in β-CD from 15-H, 16-H, 17-H, 18-H, and 21-H in CA in the MD trajectory of β-CD/CA complex. Cross peaks were observed in the NOESY spectrum of β-CD/CA^- measured in a buffered D_2O solution at room temperature.[1]

References

1 R. Ramos, E. Alvarez-Parrilla, F. Meijide, J. A. Seijas, E. Rodriguez Nunez, J. Vazquez Tato, *Langmuir*, **1999**, *15*, 5489.

List of publications

Original paper

Lan Yao, Yukie Mori, Keiko Takano. Theoretical Study on Intermolecular Interactions in Complexes of Cyclodextrins with Bile Acids: DFT and ab initio Fragment Molecular Orbital Calculations.

Bull. Chem. Soc. Jpn., **2014**, *87*, 258. (Selected Papers)

Domestic conferences

[1] ○Lan Yao, Yukie Mori, Keiko Takano. Theoretical Study on Intermolecular Interactions in Complexes of Cyclodextrins with Bile Acids: DFT and ab initio Fragment Molecular Orbital Calculations.

The VIth Annual Meeting of Japanese Society for Molecular Science, **2013**, 1P131 (poster)

[2] ○Lan Yao, Yukie Mori, Keiko Takano. Theoretical Study on Complexes of Cyclodextrins (CDs) with bile acids and bile salts.

The VIIth Annual Meeting of Japanese Society for Molecular Science, **2012**, 4P101 (poster)

[3] ○Lan Yao, Yukie Mori, Keiko Takano. Theoretical Study on Complexes of Cyclodextrins (CDs) with Cholic Acid.

Annual Meeting of Chemical Society of Japan (Spring) **2012**, 1PC-184 (poster)

Acknowledgement

First and foremost, I am most grateful to my supervisor, Professor Keiko Takano, whose useful suggestions, incisive comments, and constructive criticism have contributed greatly to the completion of this thesis. She devotes a considerable portion of her time to reading my manuscripts and making suggestions for further revisions. I would like to express my thanks to her tremendous assistance in developing the framework for analysis and having gone through the draft versions of this thesis several times. Without her patient instruction, insightful criticism, and expert guidance, the completion of this thesis would not been possible. Meanwhile, her great cares in my job hunting as well as my daily life deserve more thanks than I can find words to express.

It gives me great pleasure in acknowledging the support and help of Professor Hirotoshi Mori, who has offered me valuable advices in the academic studies. In the preparation of the thesis, he has spent much time to provide me with inspiring advice.

I owe a special debt of gratitude to Emeritus Professor Michiko Konno, who always encourages me to overcome the adversity.

I am also greatly indebted to Dr. Yukie Mori who has helped me directly in my study. She makes me understand the interesting of quantum chemistry and theoretical chemistry more than before.

I consider it an honor to work with the members in Professor Takano's laboratory and Professor Mori's laboratory, whose names are Noriko Tsuchida, Aya Mastuda, Miho Otsuka, Sayuri Fujiki, Yuka Koyama, Kanako Sato, and Natsumi Hirayama. I could not adapt to the life in Japan without their kind help.

In addition, I would like to thank to the financial support for graduate student research from Ochanomizu University. The computations were carried out at the Research Center of Computer Science in Okazaki, Japan.

At last, I owe my deepest gratitude to my family. Special thanks to the unconditional support from my husband, Yoshio Son.