Dissertations, Master's Theses and Master's Reports - Open

Dissertations, Master's Theses and Master's Reports

2009

Molecular interaction between perthiolated [beta]-cyclodextrin (CD) and the guests molecules adamantaneacetic acid (AD) and ferroceneacetic acid (FC); and the effect of the interaction on the electron transition of CD anchored particles

Ming Ning
Michigan Technological University

Follow this and additional works at: https://digitalcommons.mtu.edu/etds

Part of the Chemistry Commons

Recommended Citation

Follow this and additional works at: https://digitalcommons.mtu.edu/etds

Part of the Chemistry Commons

Molecular Interaction between Perthiolated β-Cyclodextrin (CD) and the Guests Molecules Adamantaneacetic Acid (AD) and Ferroceneacetic acid (FC); and the Effect of the Interaction on the Electron Transition of CD Anchored Particles

By

Ming Ning

A DISSERTATION

Submitted in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

(Chemistry)

Michigan Technological University

2009

This dissertation, "Molecular Interaction between Perthiolated β-Cyclodextrin (CD) and the Guests Molecules Adamantaneacetic Acid (AD) and Ferroceneacetic acid (FC); and the Effect of the Interaction on the Electron Transition of CD Anchored Particles", is hereby approved in partial fulfillment of the requirements for the degree of DOCTOR OF PHILOSOPHY in the field of Chemistry.

DEPARTMENT:

Department of Chemistry

Signatures:

Dissertation Co-advisor _____

Typewritten Name: Richard E. Brown

Dissertation Co-advisor _____

Typewritten Name: Bahne C. Cornilsen

Department Chair _____

Typewritten Name: Sarah A. Green

Date _____

Abstract

The molecular interactions between the host molecule, perthiolated beta-cyclodextrin (CD), and the guest molecules, adamantaneacetic acid (AD) and ferroceneacetic acid (FC), have been investigated theoretically in both the gas and aqueous phases. The major computations have been carried out at the theoretical levels, RHF/6-31G and B3LYP/6-31G. MP2 electronic energies were also computed based at the geometries optimized by both the RHF and B3LYP methods in the gas phase to establish a better estimate of the correlation effect.

The solvent phase computations were completed at the RHF/6-31G and B3LYP/6-31G levels using the PCM model. The most stable structures optimized in gas phase by both the RHF and B3LYP methods were used for the computations in solution.

A method to systematically manipulate the relative position and orientation between the interacting molecules is proposed. In the gas phase, six trials with different host-guest relative positions and orientations were completed successfully with the B3LYP method for both the CD-AD and CD-FC complexes. Only four trials were completed with RHF method.

In the gas phase, the best results from the RHF method gives for the association Gibbs free energy (ΔG^0) values equal to -32.21kj/mol for CD-AD and -25.73kj/mol for CD-FC. And the best results from the B3LYP method have ΔG^0 equal to -47.57kj/mol for CD-AD and -41.09kj/mol for CD-FC. The MP2 correction significantly lowers ΔG^0 based on the geometries from both methods. For the RHF structure, the MP2 computations lowered ΔG^0 to -60.64kj/mol for CD-AD and -54.10 for CD-FC. For the structure from the B3LYP method, it was reduced to -59.87 kj/mol for CD-AD and -54.84 kj/mol for CD-FC. The RHF solvent phase calculations yielded following results: ΔG^0(aq) equals 107.2kj/mol for CD-AD and 111.4kj/mol for CD-FC. Compared with the results from the RHF method, the B3LYP method provided clearly better solvent phase results with ΔG^0(aq) equal to 38.64kj/mol for CD-AD and 39.61kj/mol for CD-FC. These results

qualitatively explain the experimental observations. However quantitatively they are in poor agreement with the experimental values available in the literature and those recently published by Liu et al.[1] And the reason is believed to be omission of hydrophobic contribution to the association. Determining the global geometrical minima for these very large systems was very difficult and computationally time consuming, but after a very thorough search, these were identified. A relevant result of this search is that when the complexes, CD-AD and CD-FC, are formed, the AD and FC molecules are only partially embedded inside the CD cavity. The totally embedded complexes were found to have significantly higher energies.

The semiempirical method, ZINDO, was employed to investigate the effect of complexation on the first electronic excitation of CD anchored to a metal nano-particle. The computational results revealed that after complexation to FC, the transition intensity declines to about 25% of the original value, and after complexation with AD, the intensity drops almost 50%. The tighter binding and transition intensity of CD-AD qualitatively agrees with the experimental result that the addition of AD to a solution of CD and FC restores the fluorescence of CD that was quenched by the addition of FC.

A method to evaluate the "hydrophobic force" effect is proposed for future work.

Acknowledgements

I would like to take this opportunity to express my deep gratitude to Dr. Richard Brown, who is my advisor, for his consistent patient support and help, and for so many wonderful conversations. Dr. Brown is one of two persons, who have had the most influence on my choice of theoretical research as my future career. The other person is Dr. Don Beck from the Physics Department, who taught me the course on quantum mechanics. Dr. Brown's continuous encouragement and help have been essential to completing this project. His advice and suggestions will certainly be what I will continue to follow in the future.

I am also in great debt to Dr. Bahne Cornilsen for his kindness and friendship. Dr. Cornilsen is the one who introduced me to the areas like the solid state and molecular spectroscopy. Without him, I would never have thought that I would have an interest in these areas. His class on group theory is really one of my favorite classes. As my co-advisor, he has been always available for help.

I also have great appreciation and admiration for Dr. Marshall Logue for his tremendous knowledge of chemistry, his humor and his kindness for agreeing to be a member of my committee.

Without any doubt the quantum mechanics class from Dr. Don Beck is one of the most interesting and rewarding classes in my life. Those illuminating discussions will benefit me for a long time to come. His readiness to help any student has been so impressive to me.

There are many thanks that I have for the many wonderful people in Chemistry Department for those memorable times that we had together. Kelley, Lorri and Aparna, it has been such a wonderful time working with all of you. And who can possibly forget the enormous help that Shane gave on computer issues. For my two group mates, Ping and Xin, thank you so much for your advice and help.

I would also like to thank the Chemistry Department for the generous financial support and facilities provided there. And this project would never have been possible without the computing power and consulting advice from the Pittsburgh Supercomputing Center provided through a generous NSF grant (grant number TG-CHE040036N).

Finally, I have to thank my whole family, my parent, my sister and my little adorable niece, for their consistent selfless support for the past many years. They are always my greatest source of power now matter how hard things get.

Table of Contents

1. Introduction

1.1. Origin of this project

This project originated from work in Dr. Jian Liu's group. This group had successfully synthesized water soluble perthiolated β-cyclodextrin modified Cadmium Sulfide Quantum dots (β-CD/CdS-QDs), whose fluorescence can be reversibly tuned by binding to different guest molecules.[1] Their intention was to prove that they could make fluorescence sensors, whose fluorescence emission can be controlled reversibly by complexation with neutral molecules in contrast to the approach of direct physical adsorption or chelation of metal ions.[1] The structure of β-CD used in their lab is shown in Figure 1 below.
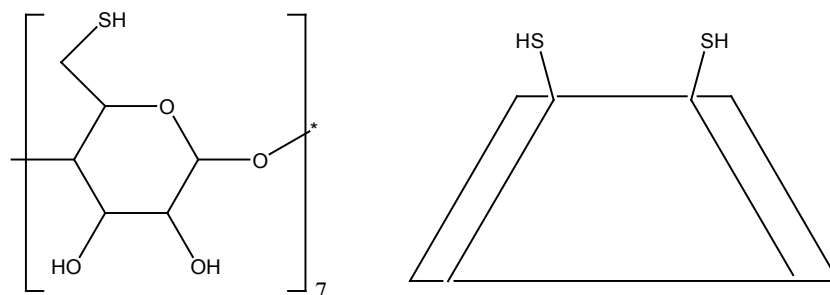
Fig 1.1. Structure of β-CD used for synthesizing quantum dots

The β-CD present above is a polycyclic molecule consisting of seven six-member monomer rings. The six-member monomer ring is monothiolated D-glucose. On the right hand side of Figure 1 is the conceptual diagram commonly used to describe the cone-shape of CD.

Dr. Liu's group believes that in the quantum dots, CD molecules are anchored to the CdS substrate by a bond between the sulfur atom in CD and the CdS molecule. ^{1}H NMR spectra for the free β-CD and the β-CD modified quantum

dots confirm that the CD molecules are immobilized on the surface of the quantum particles.[2,3] The TEM (transmission electron microscopy) measurements revealed that the average particle size is about 4 nm in diameter.[1] From an elemental analysis, approximately 80% of the surface of the quantum particles was estimated to be covered by β-CD.[1]

Two sets of molecules were employed to tune the fluorescence emission of these β-CD/CdS-QDs through the host-guest binding interaction. CD acts as the host molecule. One set is 1-Adamantaneacetic acid (AD), and another set consists of three Ferrocene derivatives (FC). Those molecules are shown in Figure 1.2.
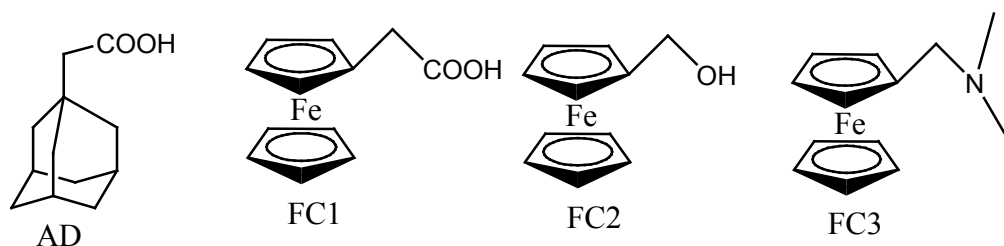
Fig 1.2. Structures of the guest molecules

From square wave voltammogram (SWV) experiments, it was concluded that CD exhibits its binding capability[1,2] to a variety of molecules including those of the two sets mentioned above.

According to the cited work[1], the β-CD/CdS-QDs have a typical two-band fluorescence emission. A sharp band-edge emission is centered around 410 nm followed by a broad trap-state emission ranging from 430 to 630 nm. However this fluorescence activity can be tuned by introducing different molecules from the two sets shown in Figure 2 to the aqueous solution containing β-CD/CdS-QDs. It has been demonstrated that by adding 1mM of FC3 to the solution, the intensity of the band-edge peak decreases significantly; however adding AD alone does not affect the fluorescence emission; but adding AD after FC3 can recover the emission intensity by as much as 90%. Clearly FC3 acts as a "fluorescence

2

quencher" and AD can offset this quenching effect from FC3. Also FC1 and FC2 have similar effects as a "fluorescence quencher". The authors here claimed that by doing such, they have successfully for the first time set up a conceptual model of a fluorescence sensor whose fluorescence emission can be reversibly manipulated using neutral molecules unlike other approaches which involve direct physical adsorption or chelation of a metal ion.

Since CD is an excellent receptor for all of these molecules, AD, FC1, FC2 and FC3, a model has been proposed to attribute this quenching and recovery effect of the fluorescence to CD binding the two different types of molecules. A straightforward explanation is that the fluorescence emission was first diminished because of the binding to FC1, FC2 or FC3. Then the AD molecules bind more strongly to CD and replaces FC1, FC2 or FC3 to recover the emission. If so, how much more strongly does AD bind to CD than FC1, FC2 and FC3 and how is the fluorescence emission quenched and recovered?

In order to answer the first question, we need to investigate the binding of CD to AD and the quenchers. Apparently, some thermodynamic quantities such as the binding enthalpy (ΔH), binding entropy (ΔS) and binding Gibbs free energy (ΔG) can provide helpful information, especially ΔG since the binding takes place under constant pressure and temperature. As for the mechanism of quenching and recovery of the fluorescence, electronic excited states must be studied for free β-CD/CdS-QDs as well as the complexes with AD, FC1, FC2 and FC3. First we need to determine the specific excitation in free QDS corresponding to the fluorescence (HOMO to LUMO), and then to locate the relevant excitation states in the complexes. By comparing the correlated excitations in free QDS and the complexes, we may be able to reveal what causes the fluorescence quenching and recovery.

3

1.2. The structural features of cyclodextrins

Naturally occurring α, β and γ cyclodextrins have the following general structure as illustrated in Figure 3 below.
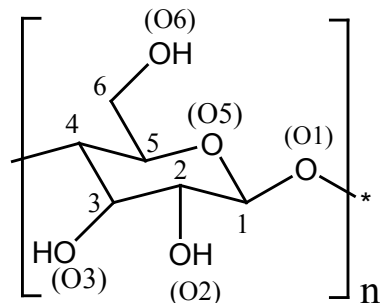
Figure 1.3. Structure of the natural cyclodextrins, n = 6, 7 and 8 for α, β and γ CDs, where the monomers are α-D-glucoses

The geometries of these macrocyclic molecules are usually described with the cone shape as illustrated in Figure 4 below.

Figure 1.4. The conceptual cone shape structure of the cyclodextrins

It is widely accepted that the narrower opening is formed by the primary hydroxyl groups, O_6H, bonded to C_6, and the wider opening is formed by the secondary hydroxyl groups, O_3H or O_2H, which are directly bonded to carbons on the monomer ring.

Apparently, the cyclodextrins are molecules with considerable size, which means that many geometry parameters are required to characterize the structure. In his review, Lipkowitz[4] depicted several important geometry descriptors. They are the intersaccharidic bond angle φ defined as the angle for C_1-O_1-C_4' (C_4 on adjacent monomer ring), the torsional angles Φ and Ψ about the glycosidic linkages

4

defined as C_5-C_1-O_1-C_4' and C_1-O_1-C_4'-C_3', the two torsional angles, Θ_1 and Θ_2 around C_4 for the C_1-C_2-C_3-C_4 and C_3-C_4-C_5-O_5 linkages, and the exocyclic torsional angle ω defined as O_5-C_5-C_6-O_6. These angles are listed in Table 1.0

Table 1.0 The definition of some important angle descriptors:[*]

Angle	Defining Atoms
φ	C_1-O_1-C_4'
Φ	C_5-C_1-O_1-C_4'
Ψ	C_1-O_1-C_4'-C_3'
Θ_1	C_1-C_2-C_3-C_4
Θ_2	C_3-C_4-C_5-O_5
ω	O_5-C_5-C_6-O_6

* Atoms with an apostrophe are located on an adjacent monomer ring

Besides X-ray structural data for crystalline cyclodextrin, an enormous amount of theoretical work has been carried out to study the free, gaseous CDs[4]. Most of them employed MM (Molecular Mechanics) or MD (Molecular Dynamics) simulations using various force fields. A few have tried semiempirical methods like the AM1 and PM3 methods. Most of those works aimed to find as many stable conformers as possible and determine the most stable conformation (exploring the potential energy surface (PES)) which has the lowest energy. This work focused on determining the geometry of those conformers, revealing the electrostatic potential and lipophilicity (hydrophilicity) outside and inside the cavity.

Although the structures of native CDs in most crystalline forms are symmetric possessing the C_n symmetry, asymmetric or distorted structures were obtained from some early theoretical studies for derivatized CDs[5] and for those with included guest molecules[6]. Kostense et al.[6] had carried out a conformation search for β-CDs to evaluate the best structure (cavity shape) for holding different guest

molecules. They started with 18 CD structures retrieved from the Cambridge Structural Database (CSD). They concluded their analysis by claiming that the O_4 atoms (the acetal linker oxygens) sometimes have to relocate to form an ellipse instead of circle to better fit some classes of guest molecules such as some benzene derivatives.

Two correlated works worth discussing are from Lipkowitz.[7, 8] Lipkowitz[7] started questioning the notion that CD molecules are inherently symmetric in their most stable state by arguing that the symmetric structures might be the consequence of crystal packing, water hydration and interaction from guest molecules. In order to prove his point, he carried out molecular mechanics optimizations on a large number of initial cyclodextrin conformations, some of which are related to those reported in the CSD. And some of his conclusions are:

1) Those conformers with C_n symmetry are not the most stable; and the symmetry must be broken to lower the energy.
2) Both gauche (-) and gauche (+) orientations for primary hydroxyl group are energetically accessible with the former one being more favorable.
3) The acetal linker oxygen atoms (O_4) do not have to be in one plane all the time.
4) Significant buckling can be achieved. The author concluded that the CDs as a class of compounds are remarkably more flexible than people had thought and the symmetric structures are likely due to time-averaged results of measurements.

In order to further support his argument, Lipkowitz, Green and Yang[8] carried out another investigation to evaluate the solid state structure of the cyclodextrins from the CSD. A total of 121 structures of cyclodextrins (43 α-, 70 β- and 8 γ-CD) were extracted from the 1990 Cambridge Structural Database (CSD). Those authors then employed MM using the MM2, CHARMM and AMBER force fields to energy minimize (geometry optimize) those 121 structures. The final results turned out to favor Lipkowitz's assumption. All 121 optimized structures are less

stable than those from Liptowitz's earlier work.[7] This further enhances Lipkowitz's point that the most stable CDs are not symmetric.[4,7] Those authors concluded that most CDs like those in the CSD are in high-energy states because of environmental effects like solid state packing, guest interaction, etc.

The work from Frieder Lichtenthaler and his coworkers had focused on analyzing cyclodextrin structures and properties, which led to a series of papers on the topic of *Molecular Modeling of Saccharides*. One highlight of one of these papers[9] is that they mapped the molecules' lipophilic surface and generated beautiful molecular graphics which are very easy to interpret and reveal those regions that are more or less hydrophilic than others. For native cyclodextrins the wide rim of the secondary hydroxyl groups is distinctly hydrophilic while the narrow rim of the primary hydroxyl groups is extensively hydrophobic. However a later work from Immel and Lichtenthaler[10] concerning the hydrophobic/hydrophilic nature of the peralkylated cyclodextrins has revealed that for permethylated cyclodextins, the area around both rims turn out to be hydrophobic since both are now covered by almost equally dense layers of alkyl groups. Actually the whole exterior region is enclosed by a hydrophobic band and one hydrophilic region remains in the cavity of the permethylated cyclodextrins. This helps to explain why some guest molecules would have to reverse their orientations when binding to native CDs vs peralkylated ones.

Besides these native cyclodextrins, a good number of investigations were carried out on the derivatized CDs. Reinhardt, Richter and Mager[11] have applied both MM and semiempirical methods to investigate the structural changes for α-, β- and γ-cyclodextrins before and after being permethylated. And some of their conclusions are:[4]

1) Alkylation always causes an increase in both the cavity openings and the ring formed by the acetal linker oxygens (O_4).

2) The increase in the opening of secondary hydroxyl groups is greater than that for the opening of the primary hydroxyl groups (except for α-Cyclodextrins).

3) However, the increase in the primary opening is often greater than that from the ring formed by the acetal linker oxygens.

We need to point out that for identifying the most important conformers, the most stable one and those near the global minimum, MD (Molecular Dynamics) methods have unique advantages over the MM, semiempirical and ab initio methods especially for large systems like the cyclodextrins. MD runs are dynamic simulations. The system can reach the many minima by meandering over a large part of potential energy surface (PES), given that the kinetic energy is high enough (the temperature is set high enough). On the other side, a geometry minimization using the MM, semiempirical and ab initio methods can only have the system fall into the minimum adjacent to the starting point. New initial structures are needed to seek new minima. And it is possible that the optimization from the new starting points may fall into the same minimum as the one already identified.

Many MD studies have been used to investigate the structure of the cyclodextrins in order to discover the global and near-global minima with the least number of constraints possible on the geometrical parameters. Please refer to Table 2 in Lipkowitz's review[4] for information about those investigations.

1.3. Inclusion activity of the cyclodextrins

The most noticeable nature of the cyclodextrins is the capability to bind guest molecules as the host molecule. In fact, β-cyclodextrin is a common organic solubilizing agent. The contributing driving forces have long been believed to be the following:[4]

1) Non-covalent Van der Waals interactions, which usually include dipole-dipole (permanent or induced) and dispersion forces. Van der Waals interactions by nature are electrostatic interactions.

2) Hydrogen bonding, which can be significant given the many hydroxyl groups in the cyclodextrin molecules and whose magnitude is considerable.

3) Energy from the relief of conformational strain, which means that either the cyclodextrin or the guest molecule adapts a conformationally preferred structure upon complexing.

4) Relief of hydrophobic forces where the guest molecules transfer to a more compatible environment inside the cavity.

5) Relief of high-energy water, which is displaced from the hydrophobic cavity to the hydrophilic bulk water.

The overall contribution from all of the driving forces listed above can be represented by the enthalpy change (ΔH), entropy change (ΔS) and Gibbs free energy change (ΔG) of the inclusion process. Apparently the ΔG is of special significance since the process takes place under constant pressure (P) and constant temperature (T). Many experiments have been carried out to measure those thermodynamics quantities. In their review, Rekharsky and Inoue[12] have reviewed most experimental methods for measuring the thermodynamic quantities of the inclusion process of the cyclodextrins and guest molecules. These include microcalorimetry (including batch microcalorimetry, flow microcalorimetry, titration "macro" calorimetry and titration microcalorimetry), electronic absorption (UV-vis), circular dichroism (CD), fluorescence (steady-state and time-resolve) spectroscopy, nuclear magnetic resonance (NMR), electron spin resonance (ESR), gas and liquid phase chromatography, capillary electrophoresis, pH potentiometry, ion selective electrodes (ISE), kinetic experiments, and solubility determinations. In rare cases, measurements of vapor pressure, conductivity and surface tension were applied to help to determine those quantities. In this review, the authors tabulated the experimental value of those thermodynamic quantities for a large number of complexes of different cyclodextrins and guest molecules. This data includes the solvents (mainly water) and measurement methods. In Table 1.1 are some data related to this project.

Much theoretical research has also been carried out to study the binding energetics. One early approach, Quantitative-Structure-Activity-Relations (QSAR), has been employed to investigate the binding interaction between cyclodextrins and guest molecules. For QSAR a mathematical model must be constructed to relate different interactions. Then experimental data are applied to parameterize the model. QSAR mathematical models do not contain information at the atomistic level. For a detailed description please refer to those references cited in Lipkowitz's review.[4]

Table 1.1: Experimental thermodynamic quantities for some β-CD inclusion complex at 298 K determined from calorimetry methods

Guest	solvent	ΔG^0 (kj/mol)	ΔH^0 (kj/mol)	$T\Delta S^0$ (kj/mol)
1-adamantanecarboxylic acid	H_2O	-32.4	-42.1	-9.7
Ferrocenecarboxylate	H_2O(pH8.6)	-19.0 ± 0.3	-15.3± 0.4	3.7± 0.7
Ferrocenecarboxylate	H_2O(pH8.6, 8Murea)	-16.2 ± 0.3	-11.1± 0.6	5.1± 1.0
Ferrocenylalkyldimethylammonium	H_2O(0.05 M N_aCl)	-19.7 ± 0.4	-28.5± 0.8	-8.8± 1.3
Ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 2M urea)	-18.3 ± 0.4	-25.5± 2.5	-7.1± 2.9
Ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 4M urea)	-17.9 ± 0.4	-24.3± 2.1	-6.3± 2.5
Ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 6M urea)	-17.1 ± 0.4	-21.3± 2.5	-4.2± 2.9
Ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 8M urea)	-16.3 ± 0.4	-20.5± 2.9	-4.2± 3.3
Ferrocenylalkyldimethylammonium	H_2O(pH 2.6-6.5)	-19.4 ± 0.4	-23.8± 0.4	-4.6± 0.8
Ferrocenylalkyldimethylammonium	H_2O(pH2.6)	-21.0 ± 0.4	-23.0± 0.4	-2.1± 0.8
Ferrocenylalkyldimethylammonium	H_2O(pH6.5)	-19.4 ± 0.4	-25.5± 0.4	-6.3± 0.8

There are many research groups who have intended to compute the inclusion free energy using methods at the atomistic level. Lukovits[13] had tried to investigate the energetics of alcohols binding to β-cyclodextrin. In this research, the geometry of both the host and guest were held fixed throughout the inclusion process (treated as rigid bodies). Only those geometry parameters, which dictate the relative position of the host and guest, were minimized using an empirical force field. The first derivatives of the computed intermolecular force (second derivative of energy) were used to guarantee the minimum. The result of this research shows that for the simplest alcohol, methanol, the van der Waals forces account for

about 50% of the total interaction energy and this percentage rises for heavier alcohols.

Because of the difficulty for evaluating the entropy, some groups have tried to study the inclusion energetics by only computing the enthalpy. One of interest is from Madrid et al.[14] In that study, a molecular mechanics computation using the Tripos force field was employed to investigate the energetics of methyl- 2-naphthoate binding to α- and β-Cyclodextrin with explicit water present. Periodic boundary conditions were applied to fully immerse the host and the guest both in the free and complexed states. The guest molecule was inserted into the cavity of the CD molecule incrementally and energy minimization was carried out at each step. These computations were carried out in both the gas and solution phases. For both cases the authors concluded that the van der Waals forces are the dominating stabilizing factor.

There are works from other groups such as Dong[15], Sato[16], Kuroda[17], Lipkowitz[18], Berg[19], Köhler[20], and Jursic[21], in which the guest molecules are gradually pushed through the cyclodextrin torus. The energy change associated with the relative position (distance) between the host and the guest was studied using MM methods. In order to define the distance, a mean plane was defined for those acetal linker oxygens (O_4) and the origin of the Cartesian coordinate system was place in the centroid of the mean plane, and the Z-axis was chosen to be perpendicular to the mean plane. Then for another reference point, a centroid was arbitrarily chosen on the guest molecule and it was moved along the Z-axis stepwise in small increments. In some cases the system was allowed to relax to different degrees. And the energy as the function of the distance between the two centroid points was computed using MM with the chosen empirical force field.

Besides the binding energy, a lot of effort was devoted to study the structure of included complex, especially the relative orientation between the host molecule and the guest molecule.

In his early computational study, Matsui[22] intended to relate the structures, stability and reactivity of some α-cyclodextrin inclusion complexes by computing the energy using MM. But only van der Waals energy was computed because the author assumed that the van der Waals forces and hydrophobic forces are the only types of significant interactions. As a result his calculation was not able to predict the correct structure for small guest molecules like Kr and MeOH. The author concluded that there was a need for considering more interactions besides the van der Waals and hydrophobic forces.

With the assumption in mind that the dipole-dipole interactions should be primarily important for stabilizing the complexes, Kitagawa et al.[23] applied the CNDO/2 semiempirical method to compute the dipole moments for both α-cyclodextrin and several other guest molecules and investigated the role of dipole-dipole interactions on stabilizing the complexes. Those guest molecules under study are some simple substituted benzenes: benzoic acid (BA), p-hydroxybenzoic acid (PHBA) and p-nitrophenol (PNP). These authors developed several angle parameters to relate the dipole moments vectors from both the host and guest molecules. At the stable conformation in the complexes, these two dipole moment vectors are almost antiparallel to each other, which enhances the assumption that the dipole-dipole interactions are the primary contributors for stabilizing the inclusion complex. The authors proposed that before inclusion, the host molecule changes its geometry to decrease the magnitude of its own dipole moment in order to reduce the energy barrier that would prevent the guest dipole moment from approaching the host. After the guest enters into the host cavity, the host varies its geometry again to raise the magnitude of its dipole moment so as to enhance the dipole-dipole interaction that stabilizes the complex.

The group of Rüdiger[24] has also used the CHARMM force field to deduce the favorable orientation of simple dipolar guest molecules in cyclodextrin cavities.

They compared their computed results from those from calorimetry and NMR experiments. Most interestingly, they have studied the inclusion between β-cyclodextrin and 1-adamantane carboxylic acid (which is quite close to one of our systems). And one interesting result they found for the complex of β-cyclodextrin and 1-adamantanecarboxylic acid is that the carboxyl group from 1-adamantanecarboxylic acid forms a hydrogen bond to the secondary hydroxyl group in β-cyclodextrin.

In the work from Pang and Whitehead[25], the binding of several highly symmetric guests (benzene and p-dihalobenzenes), to different cyclodextrins was studied. The author only considered two orientations. In one case, the cavity axis of the host (the axis perpendicular to the mean plane of the acetal linker oxygen atoms) coincides with the symmetry axis of the guest molecules (the C_2 axis lying in the benzene plane). In the other case these two axes are perpendicular to each other. The authors indicated that the optimal structure has the maximum contact surface between the host and guest molecules.

Tran, Delage and Buléon[26] have proposed a systematic way for docking the guest molecule inside the cyclodextrins. Unlike those docking methods used before, in which the guest molecules are placed in and around the host randomly or based on the researcher's intuition, these authors first invented six parameters (geometry descriptors) for dictating the relative position and orientation between the host and the guest. These six parameters are one distance, one angle, and four dihedral angles (it is different from the method proposed in this work). Then the authors could vary these six parameters systematically followed by minimization using the Tripos force field. By doing this one can search as much of the PES as possible by assigning values to as many sets of the six parameters as possible.

Another interesting subject on the inclusion complex structure is that cyclodextrins could form complexes with some guests with a stoichiometry other

13

than 1:1; in another words, the ratio of guest molecules to cyclodextrin molecules in the complex can be different from 1:1. Myles, Barlow, France and Lawrence[27] reported that the structure having a 2:1 ratio of β-CD to the drug, indomethacin, agrees with NMR data. Sato et al[28] reported in their paper that some cyanine dyes prefer a 2:1 ratio (dye to cyclodextrin) when binding to the cyclodextrins. They used MM to study the inclusion interaction between two types of cyclodextrins (β, and γ) and the monomers and dimers of three different dyes, 3,3'-diethyloxacarbocyanine iodide (DOC), 3,3'-diethyloxadicarbocyanine iodide (DODC), and 3,3'-diethylocatricarbocyanine iodide (DOTC). Through the calculations, they found that except for DOC and β-CD, the other two dyes bind to cyclodextrin with the dye to CD ratio of 2:1, D_2-CD. The authors attributed this mainly to van der Waals interactions.

Jaime and his coworkers made an extensive study into the structure of cyclodextrin inclusion complexes. In one of their studies[29] they have reported a 1:1 ratio complex of 1-bromoadamantane binding with β-cyclodextrin determined with the aid of NMR and NOE (Nuclear Overhauser Effect). MM computations using the MM2 force field were carried out to study the complex structure theoretically. Their calculations made them believe that no significant hydrogen bonding and electrostatic forces are present between the host and the guest and only in the solvent phase is the formation of the complex energetically favorable. This particular system differs from the one studied here in that those guests do not have any carboxylic acid groups and that might prevent a hydrogen bond from forming between the host and the guest.

In another work published a little later, Ivanov, Salvatiera and Jaime[30] employed NMR and MD simulations to study the inclusion complex of 1-bromoadmantane with α-, β-, and γ-cyclodextrin. The NMR results indicated that the host (CD)/guest ratios were 2:1, 1:1 and 1:1 for the α-, β-, and γ-cyclodextrin complexes respectively.

In their 2000 paper, Estrada, Perdomo-López and Torres-Labandeira[31] reported the results from their investigation on the inclusion complex of salbutamol (see figure 1.5) with β-cyclodextrin using both theoretical methods (MM2 and PM3) and experimental means (NMR and 2D ROESY). A 1:1 stoichiometry was determined by ^{1}H NMR studies. The theoretical computations indicated that two different orientations exist for the complex in both vacuo and aqueous solution. One orientation has the aromatic ring inside the CD cavity (I) and another one has the *tert*-butyl group in the cavity (II). The computational results show that the gas phase orientation (I) is energetically favored however, and in solution orientation (II) is more favorable. The reverse was confirmed by a 2D ROESY (rotating-frame Overhauser enhancement spectroscopy) experiment. The 2D ROESY results in solution showed no cross-peaks between the aromatic group protons and the protons of β-CD, but clear cross-peaks for the protons in the `tert`-butyl group and the protons inside β-CD.
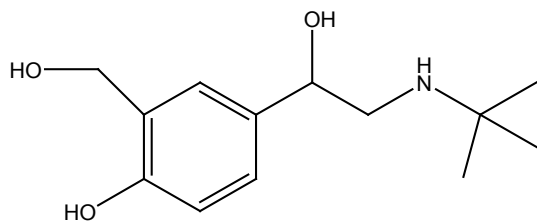
Figure 1.5. Structure of Salbutamol

Varady, Wu and Wang[32] have studied the binding interaction between benzyl alcohol and β-cyclodextrin in aqueous solution using molecular dynamics (MD) simulations. Simulations were achieved by using the CHARMM force field. Two types of simulations were tried; one is self-guided molecular dynamics (SGMD) and the other is conventional MD. Their results show an orientation with the phenyl group inside the cavity of β-CD but no hydrogen bond between the guest and the host was found. They concluded that the hydrophobic force is the main driving force.

Another MD simulation is from the work of Consonni and his coworkers[33]. The structure of the imazalil/β-cyclodextrin inclusion complex was investigated using

^{1}H and ^{13}C NMR and MD simulations. Their MD results indicate a complex structure in which the aromatic ring of the guest (imazalil) is deeply inserted into the cavity of the cyclodextrin withthe allyl and imidazole groups left outside. Hydrogen bond formation was observed between the hydroxyl groups of the cyclodextrin and the nitrogen atoms of the guest.
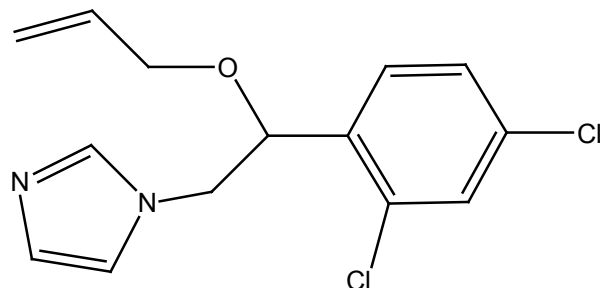
Figure 1.6. Structure of imazalil

It appears that very few investigations have been completed using ab initio methods. Most researchers are inclined to investigate the structure of cyclodextrin inclusion complexes using MM, MD and occasionally semiempirical methods. So inevitably the dispersion interactions are either completely ignored or poorly estimated. By employing both HF and those beyond like the DFT and MP2 methods, we might be able to reveal the role of dispersion interactions on stabilizing the complexes. And it seems that van der Waals forces (without dispersion) and hydrophobic forces are widely believed to be the primary contributions for binding. Relatively less attention has been paid to hydrogen bonding, which could be important.

2. Computational methods and theoretical background

2.1. Computational methods

2.1.1. Method outline

Consider the generic chemical reaction given in equation 1.1:

$$aA \; + \; bB \; \rightarrow \; cC \; + \; dD \qquad (1.1)$$

Experimentally we would use the following relations for calculating the thermodynamical quantities of the reaction such as the enthalpy ($\Delta_r H$), entropy ($\Delta_r S$) and gibbs free energy ($\Delta_r G$) changes under the condition of constant pressure (P) and temperature (T).

$$\Delta_r H \; = \; [c\Delta_f H(C) \; + \; d\Delta_f H(D)] \; - \; [a\Delta_f H(A) \; + \; b\Delta_f H(B)] \quad (1.2)$$

$$\Delta_r S \; = \; [c\Delta_f S(C) \; + \; d\Delta_f S(D)] \; - \; [a\Delta_f S(A) \; + \; b\Delta_f S(B)] \quad (1.3)$$

$$\Delta_r G \; = \; [c\Delta_f G(C) \; + \; d\Delta_f G(D)] \; - \; [a\Delta_f G(A) \; + \; b\Delta_f G(B)] \quad (1.4)$$

And for $\Delta_r H$, $\Delta_r S$ and $\Delta_r G$, we have following relation under constant temperature:

$$\Delta_r G \; = \; \Delta_r H \; - \; T\Delta_r S \qquad (1.5)$$

Those symbols like $\Delta_r H$, $\Delta_r S$ and $\Delta_r G$, with the subscript, r, refer to the change for the reaction. And those with the subscript, f, like $\Delta_f H$, $\Delta_f S$ and $\Delta_f G$ are the so called "molar quantity of formation". So $\Delta_f H(A)$ is the "molar enthalpy of formation for compound A". Please note that here we omit the temperature and the symbol indicating standard state. We imply that the temperature (T) is 298.15K and the standard state throughout this chapter unless we indicate otherwise.

The definition of these molar quantities of formation is the change associated with forming one mole of the compound of interest from all the pure elements in their most stable form. So here we set up a convention assigning a zero value to these thermodynamical quantities for elements in their most stable form under the standard state. For example, the most stable form for some elements are:[34] hydrogen gas (diatomic), helium gas (monatomic), solid lithium, solid beryllium, solid boron, solid carbon as graphite, nitrogen gas (diatomic), oxygen gas (diatomic), fluorine gas (diatomic), and neon gas (monatomic). According to this definition and convention, the $\Delta_f H^o$, $\Delta_f S^o$ and $\Delta_f G^o$ for Carbon Dioxide (CO_2) at 298.15K are the enthalpy change, entropy change and free energy change associated with following chemical reaction:

$$C(solid) \;+\; O_2(gas) \;\rightarrow\; CO_2(gas)$$

The superscript of o indicates the standard state. For the above reaction both C and O_2 are in their most stable form according the convention, in which C is in its graphite form, not diamond or others. For more information, please refer to any general Physical Chemistry textbook for a detailed discussion.

In a theoretical study, the direct computation of those thermodynamical quantities ($\Delta_r H$, $\Delta_r S$ and $\Delta_r G$) for a reaction like **1.1** can be achieved by evaluating the following equations:

$$\Delta_r H \;=\; [cH(C) \;+\; dH(D)] \;-\; [aH(A) \;+\; bH(B)] \quad (1.6)$$

$$\Delta_r S \;=\; [cS(C) \;+\; dS(D)] \;-\; [aS(A) \;+\; bS(B)] \quad (1.7)$$

$$\Delta_r G \;=\; [cG(C) \;+\; dG(D)] \;-\; [aG(A) \;+\; bG(B)] \quad (1.8)$$

In equations 1.6, 1.7 and 1.8, the quantities for H, S and G for compounds A, B, C and D are the value relative to the zero point of the state, in which all nuclei and electrons are infinitely far apart and at rest.

Before we move on to discuss how to evaluate H, S and G theoretically, we need to point out here that because the electronic correlation energies in different

molecules can be enormously different, the direct computation illustrated in equation 1.6-1.8 can yield very inaccurate results. However, identical bonds or lone pairs in different molecules have similar correlation energies and are calculated with similar error when using the same level of theory. If in chemical reaction 1.1, the number of various bonds and lone pairs on both sides are exactly identical, and then we would expect those errors in correlation energies to largely cancel out. So equation 1.6-1.8 can offer reasonably good results.[35] Such type of reaction is called "isodesmic". And this is the case for the system in our project; we expect that no chemical bonds are made or broken throughout the process of CD complexing to both AD and FC. So the complexation process that we will investigate is indeed isodesmic.

Actually, this technique has been often used for computing quantities such as $\Delta_f H$. For instance, if we know $\Delta_f H$ for all compounds except one, we can solve for it by using equation 1.2 and 1.6.

The internal energy (U) and entropy (S) can be directly computed by making use of statistical mechanical principles. Then, H and G can be obtained according to relations listed below:

$$H \quad = \quad U \quad + \quad PV \qquad (1.9)$$

$$G \quad = \quad H \quad - \quad TS \qquad (1.10)$$

P, V and T are the pressure, volume and temperature respectively. For an ideal gas PV equals nRT, in which R is the gas constant ($8.314 \text{J mol}^{-1} \text{ K}^{-1}$) and n is the number of moles of the gas. U and S can be evaluated according to the following equations based on statistical mechanics assuming a canonical ensemble:

$$U \quad = \quad k_b T^2 (\frac{\partial \ln Q}{\partial T})_{N,V} \qquad (1.11)$$

$$S \quad = \quad k_b \ln Q \quad + \quad k_b T^2 (\frac{\partial \ln Q}{\partial T})_{N,V} \qquad (1.12)$$

In equation 1.11 and 1.12, k_b is Boltzmann constant, which is the gas constant R divided by Avogadro's number, N_A. And Q is the partition function for the system containing N molecules. If these N molecules do not interact with one another (ideal gas assumption), we can then relate Q to the single molecular partition function q (V, T) as following:

$$Q(N,V,T) = \frac{[q(V,T)]^N}{N!} \qquad (1.13)$$

Here, we assume that all N molecules are indistinguishable. If we break the molecular energy down to the electronic, translational, rotational and vibrational components, we can further reduce q(V, T) to a product of partition functions for the electronic, translational, rotational and vibrational components. That is, if the energy is separable sum-wise, the partition function is separable product-wise due the exponential dependence of the partition function on the energy.

$$
\begin{aligned}
q(V,T) &= \sum_m^{levels} g_m EXP(-\varepsilon_m / k_b T) \\
&= \sum_m^{levels} g_m EXP[-(\varepsilon_{ele} + \varepsilon_{tran} + \varepsilon_{rot} + \varepsilon_{vib})_m / k_b T] \\
&= \sum_i^{ele} g_i EXP(-\varepsilon_i / k_b T) \ \sum_j^{tran} g_j EXP(-\varepsilon_j / k_b T) \ \sum_k^{rot} g_k EXP(-\varepsilon_k / k_b T) \ \sum_l^{vib} g_l EXP(-\varepsilon_l / k_b T) \\
&= q_{ele} q_{tran} q_{rot} q_{vib} \quad (1.14)
\end{aligned}
$$

In equation 1.14, g_m and ε_m are the degeneracy and energy for the mth level. Other g and ε (g_i and ε_i, g_j and ε_j, g_k and ε_k, g_l and ε_l) are the degeneracy and energy for the separable electronic, translational, rotational and vibrational levels respectively. From equation 1.11 and 1.12, it is clear that lnQ instead of Q plays an important role on computing U and S. So we need first to analyze lnQ. By combining equations 1.13 and 1.14, we have:

$$
\begin{aligned}
\ln Q &= \ln[\frac{(q_{ele} q_{tran} q_{rot} q_{vib})^N}{N!}] \\
&\approx N \ln q_{ele} \ + \ N \ln q_{rot} \ + \ Nq_{vib} \ + \ (N \ln q_{tran} \ - \ N \ln N \ + \ N) (1.15)
\end{aligned}
$$

Here, we make use of Stirling's approximation for ln(N!) that when N is sufficiently large, ln(N!) is very close to NlnN-N. It is apparently valid for any system of chemical interest since it always involves approximately Avogadro's number (10^{23}) of particles. And note that we collect the term of NlnN-N, which comes from N! in the denominator, with the translational part.

So we have reduced the system partition function, lnQ, into individual contributions from the electronic, translational, rotational and vibrational components, each of which can provide its own contribution for U and S by making use of equations 1.11 and 1.12. Now the problem left is how to evaluate those individual partition functions. Next we will briefly discuss how to compute those individual contributions, but for details, please refer to any standard Physical Chemistry textbook or Statistical Mechanics textbook.

2.1.2. Electronic contribution to U and S

Due to the significant energy gap between the electronic ground state and the lowest excited state for the vast majority of molecular systems, the occupancy of electronic excited states are negligible unless the temperature is extremely high. For a close-shell singlet molecule (g=1), the electronic partition function can be greatly simplified as following:

$$q_{ele} = EXP(-E_{ele}/k_b T) \qquad (1.16)$$

E_{ele} is the electronic energy for the ground state. Then by making use of equation 1.11 and 1.12, we have the following electronic internal energy and entropy contributions:

$$U_{ele} = E_{ele} \qquad (1.17)$$

$$S_{ele} = 0 \qquad (1.18)$$

Please note that the above equations are only valid for a singlet molecular state which the systems of interest in this work satisfy.

2.1.3. Translational contribution to U and S

For the model of "a particle in a box", the quantum mechanical energy levels for translation are given by:

$$\varepsilon_{tran} = \frac{h^2}{8m}(\frac{n_1^2}{a^2} + \frac{n_2^2}{b^2} + \frac{n_3^2}{c^2}) \quad (n_1, n_2, n_3 = 1,2,3,.......) \quad (1.19)$$

Integers n_1, n_2, n_3 are the quantum numbers for the energy level, h is the plank constant, m is the mass of a single molecule, and a, b and c are the dimensions of the box. Because of the close spacing between the energy levels, by approximating the sum with an integral, we get the following result:

$$q_{tran} = V[\frac{(2\pi m k_b T)^{3/2}}{h^3}] \quad (1.20)$$

V is the volume of the box (a×b×c). Then the molar translation contributions to U and S are:

$$U_{tran} = \frac{3}{2}RT \quad (1.21)$$

$$S_{tran} = R\ln[(\frac{2\pi m k_b T}{h^2})^{3/2} \frac{Ve^{5/2}}{N_A}] \quad (1.22)$$

N_A is Avogadro's number (6.02×10^{23}), and e is the base of the natural log. Here we assume one mole of molecules and included the NlnN-N part in the translational term (eq 1.15).

2.1.4. Rotational contributions to U and S

For the model of a rigid rotor, the rotational energy level for a linear molecule is:

$$\varepsilon_{rot} = \frac{\hbar^2}{2I} J(J+1) \qquad (J = 0,1,2,3,......)$$

$$g = 2J + 1 \qquad\qquad (1.23)$$

I is the principal moment of inertia, g is the degeneracy, $\hbar$ is Planck's constant h divided by 2π, and J is the quantum number. As with the approximation carried out for computing the partition function for translation, we can replace the summation with an integral because of the close spacing between the rotational energy levels. The rotational partition function for a linear molecule is given by:

$$q_{rot}^{linear} = \frac{2I k_b T}{\sigma \hbar^2} \qquad (1.24)$$

σ is the so called "symmetry number". But for a non-linear molecule with three principal moments of inertia I_A, I_B, and I_C, the rotational partition function becomes:[36]

$$q_{rot} = \frac{\sqrt{\pi I_A I_B I_C}}{\sigma} \left(\frac{2k_b T}{\hbar^2} \right)^{3/2} \qquad (1.25)$$

So, we have the following rotational contributions to U and S:

$$U_{rot} = \frac{3}{2}RT \qquad (1.26)$$

$$S_{rot} = R\left\{\ln\left[\frac{\sqrt{\pi I_A I_B I_C}}{\sigma}\left(\frac{2k_b T}{\hbar^2}\right)^{3/2}\right] + \frac{3}{2}\right\} \qquad (1.27)$$

For the systems that we study, the symmetry number σ equals 1.

2.1.5. Vibrational contributions to U and S

Under the model for the harmonic oscillator, a non-linear molecule has 3N-6 vibrational modes which means that there are 3N-6 harmonic oscillators. N here is the number of atoms in the molecule. For linear molecules, the number of modes reduces to 3N-5. Here we only discuss non-linear molecules since none of the systems studied here are linear. The total vibrational energy is the sum of energies from these 3N-6 modes. And the overall vibrational partition function is the product of the partition functions from those individual modes. For the ith mode, we have:

$$\varepsilon_n^i = \left(n+\frac{1}{2}\right)h\omega_i \qquad (n=0,1,2,....) \qquad (1.28)$$

$$q_{vib}^i = \sum_n^{\infty} EXP\left[\frac{-(n+\frac{1}{2})h\omega_i}{k_b T}\right] \qquad (1.29)$$

ε_n^i is the nth energy level for the ith vibration mode, ω_i is the frequency for the ith mode and q_{vib}^i is the partition function for the ith mode. Equation 1.29 is a convergent series. It can be expressed as a closed form:

$$q_{vib}^{i} = \frac{EXP\left(-h\omega_i/2k_bT\right)}{1 - EXP\left(-h\omega_i/k_bT\right)} \qquad (1.30)$$

The overall partition function for vibration is given as follows:

$$q_{vib} = \prod_{i=1}^{3N-6}\left[\frac{EXP\left(-h\omega_i/2k_bT\right)}{1 - EXP\left(-h\omega_i/k_bT\right)}\right] \qquad (1.31)$$

The molar vibrational contributions to U and S are:

$$U_{vib} = R\sum_{i=1}^{3N-6}\left[\frac{h\omega_i}{k_b\left(e^{h\omega_i/k_bT} - 1\right)} + \frac{h\omega_i}{2k_b}\right](1.32)$$

$$S_{vib} = R\sum_{i=1}^{3N-6}\left[\frac{h\omega_i}{k_bT\left(e^{h\omega_i/k_bT} - 1\right)} - \ln\left(1 - e^{-h\omega_i/k_bT}\right)\right](1.33)$$

Now, we can get the total internal energy and entropy by summing up all individual contributions:

$$U_{total} = U_{ele} + U_{tran} + U_{rot} + U_{vib} \qquad (1.34)$$

$$S_{total} = S_{ele} + S_{tran} + S_{rot} + S_{vib} \qquad (1.35)$$

By making use of equations 1.9 and 1.10, we can get the value of H and G for the molecule of interest.

For the translational and rotational parts, no more information about the structure of the molecule is needed to compute the U and S (hence H and G), but for the electronic and vibrational parts, quantum mechanical calculations are needed in

order to obtain the electronic energy levels and vibrational frequencies. The vibrational frequencies can only be computed after the geometry of the molecule is optimized to the global minimum, which means that the electronic energy is the lowest at that specific nuclear arrangement (geometry).

The free energy change in solution is evaluated under the aid of following cycle:

$$A(g) + B(g) \quad \xrightarrow{\Delta G(g)} \quad C(g) + D(g)$$
$$\downarrow \Delta G_1 \qquad\qquad\qquad \uparrow \Delta G_2$$
$$A(aq) + B(aq) \quad \xrightarrow{\Delta G(sol)} \quad C(aq) + D(aq)$$
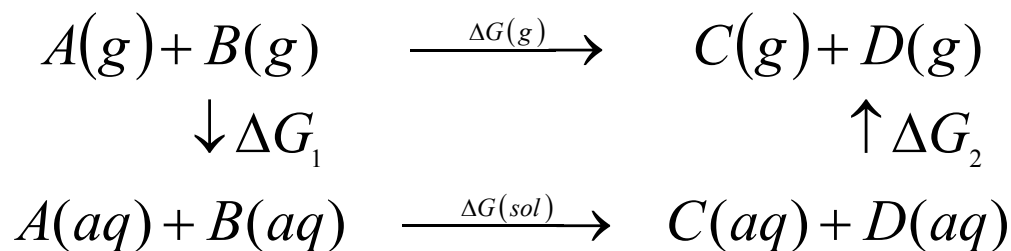
Figure 2.1: cycle for computing free energy change in solution

In the above diagram, $\Delta G(g)$ and $\Delta G(sol)$ are the gibbs free energy change for the reaction in gas phase and solvent phase respectively. ΔG_1 is the solvation free energy for the reactants and ΔG_2 is the negative of the solvation free energy for the products. From figure 1, we have:

$$\Delta \boldsymbol{G}(\boldsymbol{sol}) = \Delta \boldsymbol{G}(\boldsymbol{g}) - \Delta \boldsymbol{G}_1 - \Delta \boldsymbol{G}_2 \qquad (1.36)$$

$\Delta G(g)$ can be computed using the methods discussed above. We need to carry out solvent phase calculations to obtain ΔG_1 and ΔG_2 in order to compute $\Delta G(sol)$ using the relation shown in equation 1.36.

There are many quantum mechanical methods for computing the electronic wave function and energy such as the HF, MPn (n=1, 2, 3,….) and DFT methods. Since the HF self-consistent-field method is the basis for other higher level methods, we present it in a fairly detailed way here. For the details of the other methods, please refer to the references. While reading these theories, please refer to section 2.3 for a discussion of the variation theorem.

2.2. Quantum mechanics background.

2.2.1. Hartree-Fock SCF method for computing electronic wave functions

Here we adopt the methods developed by Roothaan.[37] A molecule with N nuclei and m electrons would have a Hamiltonian, the energy operator, like the one below:

$$H = -\sum_{\mu}^{m} \frac{\hbar^2}{2m_e} \nabla_{\mu}^2 - \sum_{K}^{N} \frac{\hbar^2}{2M_K} \nabla_K^2 - \sum_{\mu}^{m}\sum_{K}^{N} \frac{e^2 Z_K}{R_{iK}}$$

$$+ \sum_{\mu<v} \frac{e^2}{R_{\mu v}} + \sum_{K<L} \frac{e^2 Z_K Z_L}{R_{KL}} \qquad (2.1)$$

On the right hand side of equation 2.1, we have the terms in the order given for the kinetic energy for the m electrons, kinetic energy for N nuclei, nuclear-electron attraction, electron-electron repulsion and nuclear-nuclear repulsion. M_K is the mass of kth nucleus, m_e is the electronic mass, $R_{\mu K}$ is the distance between the μth electron and Kth nucleus, $R_{\mu v}$ is the distance between the μth and vth electrons, and R_{KL} is the distance between the Kth and Lth nuclei. And e is the electronic charge and Z_K and Z_L are the atomic numbers for Kth and Lth nuclei.

The first thing we need to do is to separate the electronic and nuclear motions. This is achieved by invoking the so called "Born-Oppenheimer Approximation". Given the fact that protons and neutrons are about 1800 times heavier than electrons[38], the nuclei move much more slowly than electrons, we can assume at any moment, the electrons move in an environment where all nuclear positions are fixed. In other words, each time the nuclei adapt a new configuration, the electrons have plenty of time to adjust to reach an equilibrium state before the nuclei adopt another configuration. The electrons appear to move in a field of

frozen or fixed nuclei. This makes it possible to separate the motion of the electrons from that of the nuclei. The Hamiltonian for the electronic part after separation is given below:

$$H_e = \sum_{\mu}^{m} h^{\mu} + \frac{1}{2} e^2 \sum_{\mu \neq \nu} \frac{1}{r^{\mu\nu}}$$

$$h^{\mu} = -\frac{\hbar^2}{2m_e} \nabla^2_{\mu} - \sum_{K}^{N} \frac{e^2 Z_K}{r^{\mu K}} \qquad (2.2)$$

In equation 2.2, we move most subscripts to superscripts in order to be consistent with those in the Roothaan paper. The subscripts e in H_e indicates that it is the electronic Hamiltonian. On the first line of equation 2.2, we multiply the electron repulsion term by ½ to avoid double-counting, which has the same effect as using μ<ν as specified in equation 2.1. In equation 2.2, we introduce a new operator, h^{μ}, which is the Hamiltonian for the μth electron moving in the field of the frozen nuclei; and this operator is linear and Hermitian.[37]

Here, we are trying to solve for the electronic wave function for a closed-shell system with an even number of electrons, 2n, all of which are paired (S=0). For the HF method, we assume that the solution for the N=2n electron wave function has the structure of a product of 2n one electron wave functions. This simple product allows the wavefunction to satisfy some necessary boundary equations. These one electron wave functions are defined as *Molecular Orbitals (MO),* where each only depends on the space coordinates of the electron occupying it:

$$\varphi_i^{\mu} = \varphi_i(x^{\mu}, y^{\mu}, z^{\mu}) \qquad (2.3)$$

Above x^{μ}, y^{μ}, z^{μ} are the Cartesian coordinates for electron μ. The subscript i indicates that it is the ith *MO*. The Pauli exclusion principle dictates that one *MO* can only hold two electrons with different spin, α and β. So for one *MO* we can define two so called *Molecular Spinorbitals (MSO):*

28

$$\psi_k^\mu = \psi_k(x^\mu, y^\mu, z^\mu, s^\mu)$$
$$= \varphi_i(x^\mu, y^\mu, z^\mu)\eta_k(s^\mu) = \varphi_i^\mu \eta_k^\mu$$

$$\eta_k^\mu = \alpha^\mu, or \beta^\mu \qquad (2.4)$$

So for a N (N=2n) electron wave function, we need N *MSO's* constructed from n *MO's*. As we mentioned above, the N electron wave function has the structure of the product of the N *MSO's*. Because electronic wave functions have to be antisymmetric, the N electron wave function is constructed as an *antisymmetrized product (AP) of MSO* in the form of a Slater determinant:

$$\Phi = (N!)^{-\frac{1}{2}} \begin{vmatrix} \psi_1^1 & \psi_2^1 & \cdots & \psi_N^1 \\ \psi_1^2 & \psi_2^2 & \cdots & \psi_N^2 \\ \cdots & \cdots & \cdots & \cdots \\ \psi_1^N & \psi_2^N & \cdots & \psi_N^N \end{vmatrix} \qquad (2.5)$$

We require that all *MO's* be orthonomalized. Thus all *MSO's* are also orthnormalized since the α spin orbital is orthogonal to the β spin orbital:

$$\int \overline{\varphi_i}\varphi_j d\tau = \delta_{ij}$$
$$\int \overline{\psi_i}\psi_j d\tau = \delta_{ij} \qquad (2.6)$$

With the traditional delta notation, these integrals equal zero when $i \neq j$ and equal unity when $i = j$. With all of the MSO orthonormalized, the N electron wavefunction as defined above in equation 2.5 is also normalized.

And the energy of this N electron wave function is:

$$E = \int \overline{\Phi} H_e \Phi d\tau \qquad (2.7)$$

The operator H_e is the electronic Hamiltonian defined in equation 2.2. The overbar on Φ implies that it is the complex conjugate of Φ. And this notation will be used for all equations that follow. By substituting Φ as defined in equation 2.5 into equation 2.7, we get:

$$E = 2\sum_i h_i + \sum_{ij} 2J_{ij} - K_{ij} \qquad (2.8)$$

Where the nuclear-field energy h_i, the coulomb integral J_{ij}, and the exchange integral K_{ij} are defined as follows:

$$h_i = \overline{h_i} = \int \overline{\varphi}_i h^i \varphi_i d\tau \qquad (2.9)$$

$$J_{ij} = J_{ji} = \overline{J}_{ij} = \overline{J}_{ji} = e^2 \int \frac{\overline{\varphi}_i^\mu \overline{\varphi}_j^\nu \varphi_i^\mu \varphi_j^\nu}{R^{\mu\nu}} d\tau^{\mu\nu} \qquad (2.10)$$

$$K_{ij} = K_{ji} = \overline{K}_{ij} = \overline{K}_{ji} = e^2 \int \frac{\overline{\varphi}_i^\mu \overline{\varphi}_j^\nu \varphi_j^\mu \varphi_i^\nu}{R^{\mu\nu}} d\tau^{\mu\nu} \qquad (2.11)$$

And apparently we have:

$$J_{ii} = K_{ii} \qquad (2.12)$$

It is useful to define two extra operators, the *coulomb operator J_i* and the *exchange operator K_i*:

$$J_i^\mu \varphi^\mu = e^2 \left(\int \frac{\overline{\varphi}_i^\nu \varphi_i^\nu}{R^{\mu\nu}} d\tau^\nu \right) \varphi^\mu$$

$$K_i^\mu \varphi^\mu = e^2 \left(\int \frac{\overline{\varphi}_i^\nu \varphi^\nu}{R^{\mu\nu}} d\tau^\nu \right) \varphi_i^\mu \qquad (2.13)$$

Therefore, by making use of the *coulomb operator J_i* and *exchange operator K_i*, we can rewrite J_{ij} and K_{ij},

$$J_{ij} = \int \overline{\varphi_i} J_j \varphi_i d\tau = \int \overline{\varphi_j} J_i \varphi_j d\tau$$

$$K_{ij} = \int \overline{\varphi_i} K_j \varphi_i d\tau = \int \overline{\varphi_j} K_i \varphi_j d\tau \qquad (2.14)$$

The goal of the HF-SCF procedure is to determine the best *MO's* or *MSO's* that give the lowest possible energy. Mathematically we need to locate the minimum of E defined in equation 2.8 by varying the φ_i's. By taking the total differential of E with respect to φ_i while applying the constraints of orthonormality (the first line of 2.6), Roothaan[37] demonstrated that the best *MO's* have to satisfy following equations in order to minimize E:

$$\left\{ h + \sum_j \left(2J_j - K_j \right) \right\} \varphi_i = \sum_j \varphi_j \varepsilon_{ji}$$

$$\left\{ \overline{h} + \sum_j \left(2\overline{J}_j - \overline{K}_j \right) \right\} \overline{\varphi_i} = \sum_j \overline{\varphi_j} \varepsilon_{ij} \qquad (2.15)$$

In equation 2.15, operator h is the operator h'' defined in equation 2.2. The coefficients, ε_{ij}, were introduced to facilitate the constraints defined by equation 2.6. It was proven that the two equations in 2.15 are actually equivalent and:

$$\varepsilon_{ji} = \overline{\varepsilon_{ij}} \qquad (2.16)$$

So the matrix of $\boldsymbol{\varepsilon}$ which has ε_{ij} as elements is Hermitian, which can be diagonalized by a unitary transformation. We define now two additional operators, the *total electron interaction operator G* and the *Hartree-Fock Hamiltonian operator F* to make the expressions in 2.15 more concise.

$$G = \sum_i \left(2J_i - K_i \right) \qquad (2.17)$$

31

$$F = h + G \qquad (2.18)$$

Again the operator h in equation 2.18 is the h^μ defined in equation 2.2. Then we can rewrite equation of 2.15 as:

$$F\varphi_i = \sum_j \varphi_j \varepsilon_{ji} \qquad (2.19)$$

Or in matrix notation:

$$F\phi = \phi\varepsilon \qquad (2.20)$$

$\emptyset$ is the matrix having the *MO's* φ_i as column vectors and ε is the matrix with ε_{ij} as an element. As we mentioned above, matrix ε is Hermitian, therefore it can be transformed to a diagonal matrix ε' by a unitary matrix. When transforming equation 2.20 to diagonal form, the *MO's*, φ_i also undergo a transformation by the same unitary matrix to form a new set of *MO's*, φ_i'. We need point out here that the operator of F in equation 2.20 depends on these *MO's*. It can be seen in equation 2.13, which defines the operators, J and K are needed to construct F. But as demonstrated by Roothaan in his paper[37], the operator F remains unchanged under such a unitary transformation. Hence the equation of 2.20 is still valid after the diagonalization of ε. So without losing generality, we can conclude that the set of best *MO's* satisfies the following simple equation:

$$F\varphi_i' = \varepsilon_i' \varphi_i' \qquad (2.21)$$

Where we can drop the new notation and just use the equation 2.20, but the matrix, ε, would be a diagonal one. The ε_i eigenvalues are interpreted as the energy levels of the *MO's*, φ_i.

So those best *MO's* are eigenfunctions of operator F. We need n such functions with lowest eigenvalues to construct the ground state close-shell wave function for the system with 2n electrons as the form of a Slater determinant. All eigenvalues are real. Any two eigenfunctions having different eigenvalues are orthogonal to each other automatically.

But for those functions having the same eigenvalue (degenerate orbitals), orthonormalization is still needed to make them orthonormal.

But solving equation 2.21 is quite tricky. First we want to solve for the φ_i's from it, however, the operator $\boldsymbol{F}$ depends on the φ_i's. The general procedure is one of trial and error. One first makes a guess at a set of φ_i's, then constructs the Fock operator, $\boldsymbol{F}$ based on the guessed functions, solve the equation 2.21 to get a new set of functions, compare them to the guessed set. The new set is then used to construct the new Fock operator. This procedure is repeated until the new set and energy equal to previous values within certain tolerance limits. That is why it is called the iterative "self-consistent" method.

2.2.2.Constructing *MO's* using basis sets

For ab inito computations, we construct the *MO's* using so called **basis sets** (**basis functions**). Basis sets are a set of mathematical functions (theoretically an infinite set), of which any arbitrary mathematical function can be constructed as a linear combination. In practice, computers are not able to deal with sets with an infinite number of functions. We always use finite sets. The underlying idea is that we can represent the *MO's* as a linear combination of some known, conventional and effective set of functions. Then by solving for the coefficients in the linear combination through a specified algorithm, we can construct all the *MO's* needed. For *MO's*, a natural choice for basis sets would be atomic orbitals from atoms in the molecule. And it is called a Linear Combination of Atomic Orbitals (LCAO). In his paper, Roothaan[37] derived a mathematically rigorous way to solve for those coefficients through the variation method, which is quite similar to the method we described above for deriving the HF SCF solution. And this is also applicable for solving the problem with other general basis sets. So here we try to present the general procedures to construct the *MO's* from general basis sets.

Suppose, we try to construct the set of MO's, $\left\{ \phi_j \right\}_{j=1}^{n}$, from the set of basis functions,

$\left\{ \chi_p \right\}_{p=1}^{m}$.

$$\varphi_i = \sum_p \chi_p C_{pi} \qquad (2.22)$$

In matrix notation:

$$\varphi_i = \chi C_i \quad and \quad \phi = \chi C$$
$$\phi = \begin{bmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_n \end{bmatrix}$$
$$\chi = \begin{bmatrix} \chi_1 & \chi_2 & \cdots & \chi_m \end{bmatrix}$$

$$C_i = \begin{bmatrix} C_{1i} \\ C_{2i} \\ \vdots \\ C_{mi} \end{bmatrix} \qquad C = \begin{bmatrix} C_{11} & C_{12} & \cdots & C_{1n} \\ C_{21} & C_{22} & \cdots & C_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ C_{m1} & C_{m2} & \cdots & C_{mn} \end{bmatrix} \qquad (2.23)$$

The goal of the LCAO treatment is to determine the best coefficient matrix C, with which the energy E reaches the minimum. By applying the variational method in a manner similar to that used in deriving HF SCF *MO's* mentioned above, Roothaan[37] illustrated that the best matrix, C, satisfies following equation:

$$FC_i = \varepsilon_i SC_i \qquad (2.24)$$

F is the Fock operator, C_i is the column vector in the matrix C, ε_i is a real number, S is the *overlap matrix*, whose element S_{pq} is defined as below:

$$S_{pq} = \int \overline{\chi_p} \chi_q d\tau \qquad (2.25)$$

The eigenvalue, ε_i, is called the *LCAO orbital energy* for MO φ_i. We rewrite equation 2.24 to get C_i:

$$(F - \varepsilon S)C_i = 0 \qquad (2.26)$$

Please note ε in above equation is a number not a matrix. And we need to solve the secular equation below for ε:

$$\begin{vmatrix} F_{11} - \varepsilon S_{11} & F_{12} - \varepsilon S_{12} & \cdots & F_{1m} - \varepsilon S_{1m} \\ F_{21} - \varepsilon S_{21} & F_{22} - \varepsilon S_{22} & \cdots & F_{2m} - \varepsilon S_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ F_{m1} - \varepsilon S_{m1} & F_{m2} - \varepsilon S_{m2} & \cdots & F_{mm} - \varepsilon S_{mm} \end{vmatrix} = 0 \qquad (2.27)$$

For the secular equation and the following discussions, we adopt some notations from Cramer's book.[39] The elements, $S_{\mu\nu}$, are from the overlap matrix S; ε is the eigenvalue that we need to solve from the secular equation above; $F_{\mu\nu}$ is defined as below as:

$$F_{\mu\nu} = \int \overline{\mu} F \nu \, d\tau \qquad (2.28)$$

F is the Fock operator defined in 2.18. But here it is defined based on the basis sets since the *MO's* are constructed through the basis sets. Here we use the subscripts μ and ν to refer to basis functions χ_μ and χ_ν and we will use this notation below.

In detail, $F_{\mu\nu}$ includes following terms:

$$F_{\mu\nu} = \left\langle \mu \left| -\frac{1}{2}\nabla^2 \right| \nu \right\rangle - \sum_K^{nuclei} Z_K \left\langle \mu \left| \frac{1}{r_k} \right| \nu \right\rangle$$

$$+ \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma) \right] \qquad (2.29)$$

The first two terms on the right side of equation 2.29 are straightforward, but we need to explain a little more for the third term.

$$(\mu\nu|\lambda\sigma) = \iint \overline{\chi_\mu}(1)\overline{\chi_\nu}(1)\frac{1}{r_{12}}\chi_\lambda(2)\chi_\sigma(2)\,dr(1)dr(2) \qquad (2.30)$$

The number 1 and 2 in 2.30 refer to the position coordinates for electron 1 and electron 2, and r_{12} is the distance between the two electrons (the relative distance between the two

positions indicated by the two coordinates). The integral in 2.30 is often referred to as the "two electron integral" or "four-index integral", which is probably the most time-consuming part for evaluating the Fock matrix. $P_{\lambda\sigma}$ is the element from the so called "density matrix" **P**. $P_{\lambda\sigma}$ is given as:

$$P_{\lambda\sigma} = 2 \sum_{i}^{occupied} C_{\lambda i} C_{\sigma i} \qquad (2.31)$$

The sum in 2.31 is over all occupied *MO's* (the n MO's with lowest energy or eigenvalue). $C_{\lambda i}$ is the coefficient for the λth basis function (χ_λ) in the ith MO (φ_i).

In summary, we need to solve the secular equation 2.27 to obtain the eigenvalues ε and eigenvector C_i (surely quite many). By using matrix methods, all m ε_i and C_i can be obtained in one matrix transformation and diagonalization. Then the n *MO's* with lowest energy can be constructed from the n coefficient column vectors with the lowest eigenvalues.

2.2.3. Variation principle

The variation theorem is the principle behind most quantum mechanical approximation methods. It simply states: the expectation value of a system Hamiltonian, **H,** for any trial wavefunction (approximate solution) is higher than the exact system ground-state energy level. The expectation value of the approximate wavefunction sometimes is also called "average energy" of the wavefunction. A brief proof is presented below.

Suppose, we have an arbitrary normalized function, ψ, which could be obtained from a guess, or any other approximation method. Let the set of functions, {u_i },(i = 0,1, 2, 3….) be the set of exact eigenfunctions of **H** which is a complete orthonormal set. Then the function of ψ can be expanded in terms of these eigenfunctions as shown below:

$$\psi = \sum_{i} A_i u_i \qquad where \qquad H u_i = E_i u_i \qquad (2.32)$$

In equation 2.32, u_i is the eigenfunction of $\boldsymbol{H}$ with eigenvalue E_i (energy level i). And u_0 and E_0 are the ground state function and ground state energy respectively.

The expectation value of $\boldsymbol{H}$ for ψ is given:

$$\langle H \rangle = \int \overline{\psi} H \psi \, d\tau = \sum_i E_i |A_i|^2 \qquad (2.33)$$

Since E_0 is the ground state energy, which is the lowest, we can easily get the following inequality:

$$\langle H \rangle = \sum_i E_i |A_i|^2 \geq \sum_i E_0 |A_i|^2 \qquad (2.34)$$

Since ψ is normalized, we have:

$$\sum_i |A_i|^2 = 1 \qquad (2.35)$$

So we end up with following relation:

$$E_0 \leq \int \overline{\psi} H \psi \, d\tau \qquad (2.36)$$

That is what the variation theorem states.

The variation principle provides a method to solve for the wavefunction approximately. If the trial function ψ has a number of unknown parameters, the best wavefunction from it can be obtained by varying those parameters in order to minimize the expectation value. For details, please refer to any quantum mechanics textbook such as the one by Schiff[40].

2.2.4. A brief introduction to the STO and GTO basis sets

We gave a very brief definition for a basis set in Section 2.2 without mentioning any details. Here we will be discussing more about basis sets by describing two specific types, the STO and GTO sets.

STO is the abbreviation for Slater-type of orbital, which is a set of functions with the following form:

$$\varphi(r,\theta,\phi,\zeta,n,l,m) = \frac{(2\zeta)^{n+\frac{1}{2}}}{[(2n)!]^{\frac{1}{2}}} r^{n-1} e^{-\zeta r} Y_l^m(\theta,\phi) \qquad (2.37)$$

In equation 2.37, n is the principle quantum number; $Y_l^m(\theta, \emptyset)$ is the spherical harmonic functions with dependence on angular momentum quantum numbers l and m, which are identical to those from the Schrodinger equation solution for Hydrogen atom. The exponent, ζ, can be decided according to a simple set of rules developed by Slater[41] or by optimization and can be interpreted as a screened nuclear charge.

Since the STOs bear a lot of resemblance with and originate from the hydrogen atomic orbitals, they are a natural choice as an atomic centered basis set in molecular calculations.[42] They possess the correct exponential decay with increasing r, have a well known and separable angular part, and have the correct cusp at the nucleus for the 1s orbital.[43,44]

Despites all these attractive merits, STOs do have shortcomings. The exponential term makes it very difficult to calculate the "four-index integral" shown in equation 2.30. This difficulty would be solved if the exponential term in the STOs is changed from EXP (-r) to EXP (-r^2), which has the form of a Gaussian function. So Gaussian-type orbitals (GTO) were proposed as an alternative. GTOs have following general function form:

$$\phi(x,y,z,\alpha,i,j,k) = \left(\frac{2\alpha}{\pi}\right)^{\frac{3}{4}} \left[\frac{(8\alpha)^{i+j+k} i! j! k!}{(2i)!(2j)!(2k)!}\right]^{\frac{1}{2}} x^i y^j z^k e^{-\alpha(x^2+y^2+z^2)} \qquad (2.38)$$

Where α is the factor controlling the width of the GTO, and i, j, and k are positive or zero integers that dictate the nature of the orbitals. And these functions are also called "primitive GTOs"

When i, j, and k are all zero, we have a s type orbital; when one of i, j, and k equals 1 and other two are zero, we have a p type orbital, and apparently we have three types of p orbitals, p_x, p_y, and p_z, depending on which one equals 1. When the sum of i, j, and k

equals 2, we have a d type orbital and there are six of these, x^2, y^2, z^2, xy, xz, and yz, which is different from the d orbitals resulting from the Schrodinger equation solution for the hydrogen atom.

Although GTOs have the advantage with the "two-electron integral problem", they are not as atomic-orbital like as STOs.[45] For instance GTOs do not have the cusp at the nucleus and the radial decay in a GTO is exponential in r^2 instead of r. In this aspect STOs have the advantage over the GTOs. So in order to keep the best features of both the STO and GTO sets, a new strategy for developing basis sets was proposed. Instead of using individual GTOs, linear combinations of GTOs are built up to best fit the STOs. And these new linear combination of GTOs serves as the basis functions (basis set)[45]:

$$\varphi(x, y, z; \{\alpha\}, i, j, k) = \sum_{a=1}^{M} c_a \phi(x, y, z; \alpha_a, i, j, k) \quad (2.39)$$

Where φ is the basis function (STO like) and Ø is the GTO, the building block. These new linear combinations of GTOs are called "contracted GTOs". And basis functions constructed in this way are often referred to as "contracted basis functions". The basis function is characterized by the exponents, α_a, the coefficients, c_a and the number M, i.e., how many GTOs are used in the combination.

Hehre, Stewart, and Pople[46] were the first to form a library of the STO-MG basis sets, "Slater-Type Orbital approximated by M Gaussians" by systematically determining the exponents and coefficients in each of these basis sets.

For example, a STO-3G basis set for a carbon atom will have each basis function for carbon: 1s, 2s, $2p_x$, $2p_y$ and $2p_z$ atomic orbitals. Each basis function is a linear combination of three GTOs of corresponding type, like a 1s from three GTOs of s type with fixed coefficients and exponents from the library.

The basis set 6-31G used in this work is a "split valence shell" basis. For the inner shell, each atomic orbital (simulated STO) is represented as a linear combination of six GTOs. But for the valence shell, each atomic orbital is represented by two independent basis

functions, one constructed with three GTOs and another consisting of one GTO. Many more complete and complex libraries of basis sets have been developed, and the reader is referred to the literature for the details on these.[39]

2.2.5. Methods for including electron Correlation

The fundamental assumption of the HF theory is that each electron moves in the average static electric field generated by nuclei and other electrons. And Fock operators are very much one electron operators, so electron correlation is completely ignored. From mathematical point of view, HF solutions have the structure of a single product and use a single determinant. This kind of constraint contributes to the failing to account for the correlation effect mathematically. One apparent way to improve it is to incorporate more determinants in the solution, constructing it as linear combination of multiple determinants, which is what the MP2 method does. DFT methods try to address this issue differently by changing the Hamiltonian with correlation functionals to correct the non-interacting wavefunction. Next we will briefly discuss how the MP2 and DFT methods work, but for the details please refer to the references.

2.2.5.1. The MP2 method

The MPn (n=1, 2, 3,……) method is the perturbation method proposed by Møller and Plesset[47], where n is the term after which the perturbation series is terminated. So for the MP2 method, the series is terminated after the term with a second power.

The perturbation method is an approximation method for solving the Schrodinger equation when the system Hamiltonian, H, can be broken down into two parts, H_0 and H', where the Schrodinger equation constructed from H_0 can be solved exactly, and the perturbation term, H', is small compared to H_0. So for H_0, the eigenfunctions, u_k with eigenvalues E_k, are available. And for the system we have[48]:

$$H\psi = W\psi, \quad H = H_0 + H', \quad H_0 u_k = E_k u_k \qquad (2.40)$$

In the above, H is the system Hamiltonian with eigenfunction ψ and eigenvalue W. Because $H^{'}$ is small, we can expand ψ and W as power series in terms of the eigenfunctions and eigenvalues of H_o. This would be more conveniently accomplished by introducing a new parameter λ and expanding the power series in λ. For that we rewrite the second equation in 2.40 as follows:[48]

$$H = H_0 + \lambda H^{'} \qquad (2.41)$$

We can then expand the system wavefunction ψ and eigenvalue W as follows:[48]

$$\psi = \psi_0 + \lambda \psi_1 + \lambda^2 \psi_2 + \lambda^3 \psi_3 + \cdots \qquad (2.42)$$

$$W = W_0 + \lambda W_1 + \lambda^2 W_2 + \lambda^3 W_3 + \cdots$$

After substituting 2.41 and 2.42 into the system Schrodinger equation, we have:

$$(H_0 + \lambda H^{'})(\psi_0 + \lambda \psi_1 + \lambda^2 \psi_2 + \cdots)$$
$$= (W_0 + \lambda W_1 + \lambda^2 W_2 + \cdots)(\psi_0 + \lambda \psi_1 + \lambda^2 \psi_2 + \cdots) \qquad (2.43)$$

Equation 2.43 is supposed to be valid for any λ ranging from 0 to 1, so the coefficients for the collected terms for the identical powers of λ on both sides of 2.43 must be equal. We get upon collecting the terms,

$$(H_0 - W_0)\psi_0 = 0$$
$$(H_0 - W_0)\psi_1 = (W_1 - H')\psi_0$$
$$(H_0 - W_0)\psi_2 = (W_1 - H')\psi_1 + W_2\psi_0$$
$$(H_0 - W_0)\psi_3 = (W_1 - H')\psi_2 + W_2\psi_1 + W_3\psi_0 \qquad etc \qquad (2.44)$$

From the first equation in 2.44, we can conclude that the ψ_0 can be any of the eigenfunctions of the unperturbed Hamiltonian H_0, and W_0 is the corresponding eigenvalue (i.e., ψ_0 and W_o correspond to any u_m and E_m).

$$\psi_0 = u_m, \quad W_0 = E_m \quad m = 0,1,2,....... \qquad (2.45)$$

Before we move on, we need to point out two facts. First, from those equations in 2.44, it is clear that any ψ_s ($s \neq 0$) can have an arbitrary multiple of ψ_0 added to it and still obey the relations shown in equation 2.44. In order to simplify the situation, we choose ψ_s in such a way that all ψ_s's are orthogonal to ψ_0:

$$(\psi_0, \psi_s) = \int \overline{\psi_0} \psi_s d\tau = 0 \quad s > 0 \qquad (2.46)$$

Second, because of the Hermitian nature of H_0, if we multiply the left side of any equation in 2.44 by ψ_0 and take the integral over the whole space (inner product), the integral vanishes. By making use of the orthogonality shown in equation 2.46, we can immediately get following relation:

$$W_s = \frac{\int \overline{\psi_0} H' \psi_{s-1} d\tau}{\int \overline{\psi_0} \psi_0} = \int \overline{\psi_0} H' \psi_{s-1} d\tau = \langle \psi_0 | H' | \psi_{s-1} \rangle \qquad (2.47)$$

In equation 2.47, we assumed that ψ_0 (u_m) is normalized.

From the first equation in 2.44, we can solve for ψ_0 and W_0, which are u_m and E_m, substitute ψ_0 and W_0 to the second equation we can solve for ψ_1 and W_1 and so on. Because the set of u_m is a complete orthonormal set, ψ_s can be expanded as linear combination of u_m.

$$\psi_s = \sum_i a_i^{(s)} u_i \qquad (2.48)$$

Here, we will only demonstrate how to solve for ψ_1, ψ_2, W_1 and W_2. We designate u_m as the reference wavefunction, which means $\psi_0 = u_m$. By substituting equation 2.48 for ψ_1 and ψ_2 into equation 2.44 we can get:

$$W_1 = \int \overline{u_m} H' u_m d\tau = \langle m | H' | m \rangle \qquad (2.49)$$

$$a_k^{(1)} = \frac{\int \overline{u_k} H' u_m d\tau}{E_m - E_k} = \frac{\langle k | H' | m \rangle}{E_m - E_k} \quad k \neq m$$

$$a_m^1 = 0$$

$$W_2 = \sum_{n \neq m} \frac{\left| \int \overline{u_m} H' u_n \right|^2}{E_m - E_n} = \sum_{n \neq m} \frac{\left| \langle m | H' | n \rangle \right|}{E_m - E_n}$$

$$a_k^{(2)} = \sum_{n \neq m} \frac{\langle k | H' | n \rangle \langle n | H' | m \rangle}{(E_m - E_k)(E_m - E_n)} - \frac{\langle k | H' | m \rangle \langle m | H' | m \rangle}{(E_m - E_k)^2} \quad k \neq m$$

$$a_m^2 = 0$$

42

We have adopted the following notation for the equations above:

$$\langle k | H' | n \rangle = \int \overline{u_k} H' u_n d\tau \qquad (2.50)$$

The MP2 method is a many-body perturbation method. The reference functions are the set of Slater determinants, which are the solutions for the HF SCF method discussed in Section 2.1 in this chapter. The unperturbed Hamiltonian H_0 is the sum of the one-electron Fock operators:[47,49]

$$H_0 = \sum_i f_i \qquad (2.51)$$

The perturbation operator is defined as follows:[49]

$$V = \sum_i^{occ} \sum_{j>i}^{occ} \frac{1}{r_{ij}} - \sum_i^{occ} \sum_j^{occ} (J_{ij} - \frac{1}{2} K_{ij}) \qquad (2.52)$$

J_{ij} and K_{ij} are the coulomb and exchange integral respectively, and these summations cover all occupied MO's. In other words, the perturbation is just the difference between the exact and Hartree Fock Hamiltonians. For the MP2 method, we need to substitute the unperturbed operator H_0 defined in equation 2.51, the perturbation operator defined in equation 2.52 and the Slater determinants from HF solution into those equations from 2.40 to 2.49. So the solutions for the MP2 method are linear combination of multiple determinants.

2.2.5.2. The DFT method

Density Functional Theory (DFT) is based the fact that the system Hamiltonian can be sufficiently determined by the electron density, ρ, and thus so are the wavefunction and energy.[50, 51] In order to facilitate the calculations, Kohn and Sham[52] proposed a new procedure to construct the electronic energy as a function of the density. They first started with a fictitious system with only non-interacting electrons. For the fictitious system, the energy is broken down to non-interacting electronic kinetic energy, nuclear-electron attraction energy and electron-electron repulsion energy. Then an exchange and correlation correction term is added to those non-interacting energy components to

correct for the correlation from electron interaction and the non-classical electron repulsion energy. With the basis set, χ_i, the electronic energy consisting of those components can be described as follows:[53]

$$E[\rho(r)] = \sum_i^N \left(\left\langle \chi_i \left| -\frac{1}{2}\nabla_i^2 \right| \chi_i \right\rangle - \left\langle \chi_i \left| \sum_k^{nuclei} \frac{Z_k}{|r_i - r_k|} \right| \chi_i \right\rangle \right)$$
$$+ \sum_i^N \left\langle \chi_i \left| \frac{1}{2}\int \frac{\rho(r')}{|r_i - r'|} dr' \right| \chi_i \right\rangle + E_{XC}[\rho(r)] \qquad (2.53)$$

The four terms on the right side of equation 2.53 are the kinetic energy for non-interacting electrons, the nuclear-electron attraction energy, and the electronic repulsion energy and exchange-correlation correction, respectively. N is the number of electrons, Z_k is the nuclear charge of the kth atom and the electronic density, $\rho(r)$ is constructed as follows:[53]

$$\rho(r) = \sum_i^N \left\langle \chi_i \left| \chi_i \right\rangle \right. \qquad (2.54)$$

A great number of research efforts has been carried out to find the function to approximate E_{XC} such as the Local Density Approximation (LDA) and Generalized Gradient Approximation (GGA). A widely used one is the B3LYP hybrid functional, which is defined as:[54]

$$E_{XC}^{B3LYP} = (1-a)E_X^{LSDA} + aE_X^{HF} + b\Delta E_X^B + (1-c)E_C^{LSDA} + cE_C^{LYP} \quad (2.55)$$

where a=0.1161, b=0.9262, and c=0.8133. The 5 terms on the right side of 2.55 are the LSDA exchange, the HF exchange, the B exchange, the LSDA correlation and the LYP correlation functionals. Please refer to the references for a detailed discussion.

2.2.6. Solvent phase computations and the PCM model

The computational methods used in this work for the solvent phase is the Polarized Continuum Medium model (PCM) developed by Tomasi and coworkers.[55] It is an

implicit solvent model because no explicit solvent molecules are present. In the PCM model, the statistically averaged solvent effect is directly addressed by describing the solvent macroscopically as continuum medium with suitable properties (dielectric constant, thermal expansion coefficient, etc) instead of a microscopic description.

The solute molecule (or ion) is placed in a cavity embedded in the infinite polarizable dielectric medium of the solvent. In the PCM method the cavity is represented as a system of interlocking spheres. Those individual spheres are centered on appropriate nuclei in the solute molecule with radii R_1, R_2, R_3…….. The number of spheres and the values of the radii are introduced into the computational procedures as empirical parameters.

The solute is represented by a charge distribution $\rho(r)$ inside the cavity. The initial distribution, $\rho^0(r)$, is usually taken as the one from gas phase calculations (with the geometry optimization, population analysis etc in gas phase). Through electrostatic interaction, net charges will be induced on the cavity surface. The induced surface charges will act back and change the charge distribution in the solute. This mutual interaction goes on until equilibrium is reached. Then we have a stable electrical field in the cavity with contributions from both solute charge distribution and induced surface charges. The potential built by the induced surface charges, $V_\sigma(r)$, will act on the solute as an additional potential. In the computation this potential is treated as a perturbation term.

$$H = H^0 + V_\sigma(r) \qquad (2.56)$$

where H is the Hamiltonian including the solvent effect, H^0 is Hamiltonian without the solvent effect (gas phase Hamiltonian), and $V_\sigma(r)$ is the potential generated by the surface charges. The computation is carried out using perturbation theory.

3. Computations and results

3.1. Preparation of individual molecules and complexes

3.1.1. Construction of perthiolated beta-cyclodextrin

The perthiolated beta-cyclodextrin (CD) molecule was constructed using the Hyperchem graphical facility as shown in Figure 5.1 in Appendix.

3.1.1.1. Conformational search for CD

With seven inter-linked glucose rings (see Figure 3.1 and 3.2 for the monomer structure), the cyclodextrin molecule is a very large molecule and thus has many conformers, especially those with different torsional angles. The conformational search was carried out using the PM3 method available in **HyperChem 7.5.**
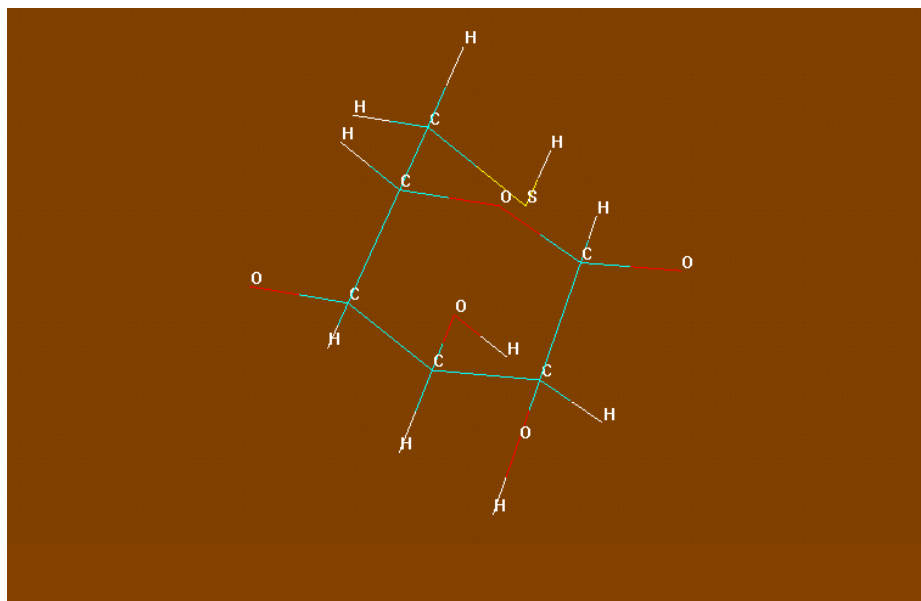

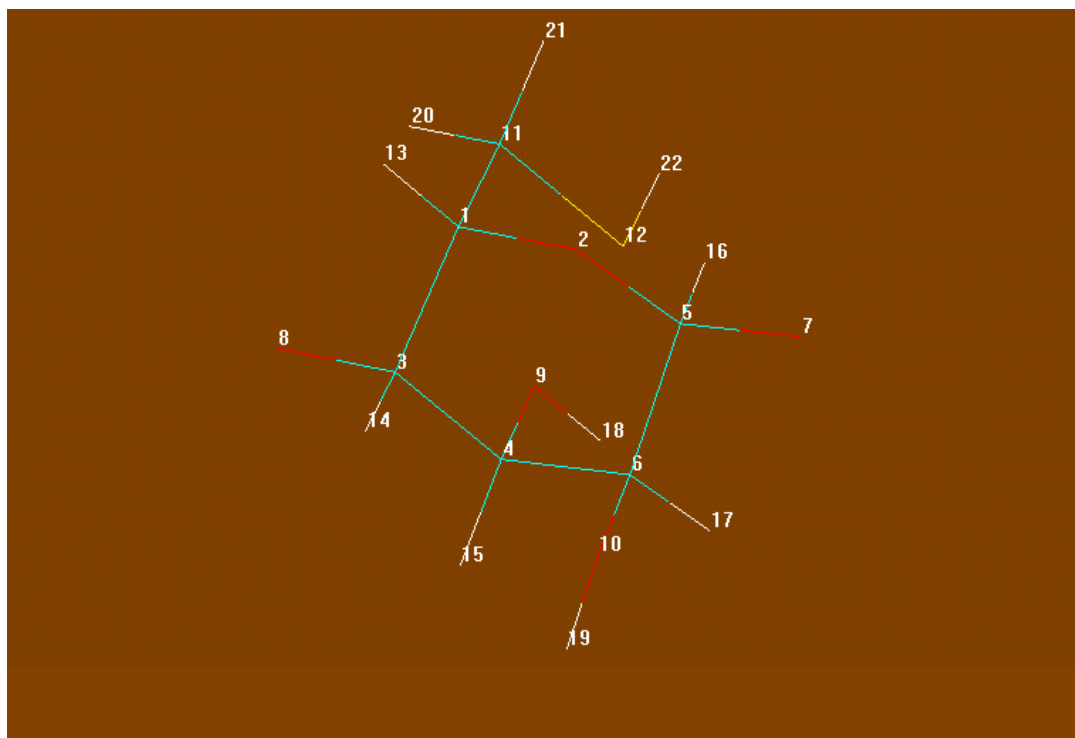Figure 3.1. The monomer ring in cyclodextrin shown with atom symbols

46

Figure 3.2. The monomer ring in cyclodextrin shown with atom numbers.

I. Choosing torsional angles:

Before the conformational search was started, 56 torsional angles were selected, eight of which were located in each unit ring. The way, by which those eight torsion angles in each unit ring were chosen, is described in Table 1 with the atoms and numbering given in Figures 3.1 and 3.2.

Terms like "torXY" are used to identify the torsional angles. The X (1-7) specifies on which ring the torsional angle resides. The Y (1-8) specifies one of eight angles in the ring specified by the X. Since those seven rings are identical to each other, the way to choose the torsional angle is the same for all rings. Table 3.1 lists all eight torsional angles for the ring shown above.

Table 3.1. The seven torsional angles from the ring shown in Figures 3.2 and 3.3

Torsion name	Illustrated in Fig 2	Illustrated in Fig 3
Tor11	H-C-O-H	15-4-9-18
Tor12	H-C-O-H	17-6-10-19
Tor13	H-C-C-H	13-1-11-20
Tor14	H-C-C-H	13-1-11-21
Tor15	H-C-C-S	13-1-11-12
Tor16	C-C-S-H	1-11-12-22
Tor17	H-C-S-H	20-11-12-22
Tor18	H-C-S-H	21-11-12-22

II. The results of the conformational search:

The conformational search was completed using the PM3 method built into HyperChem 7.5. The first ten conformers with the lowest energy were chosen for further computations at a higher level of theory. Table 3.2 lists the energy levels for these 10 conformers. The geometries are illustrated in Figures 5.2 through 5.11 in the Appendix.

Table 3.2. The energy levels for the first ten conformers with lowest energy

conformer	Energy (kcal/mol)
1	109.3676
2	109.5789
3	109.734
4	109.8334
5	109.9717
6	109.9761
7	109.9802
8	110.0331
9	110.0749
10	110.1093

III. Re-optimization of all 10 conformers at the RHF/6-31G(d) level

From the conformational search results shown above, it is apparent that the energy gaps among those conformers are not significant; the difference between the conformer with highest energy and the one with lowest energy is 0.7417 kcal/mol (about 3.103 kj/mol). Hence re-optimization with a higher level method is necessary. All ten above conformers were re-optimized at the RHF/6-31G(d) level of theory. The energies of these optimized conformers are listed in Table 3.3. Please refer to Figures 5.12 to 5.21 in the Appendix for the geometries optimized at RHF/6-31G(d).

In order to consider the correlation effect on the energy level of these conformers, single point calculations at the B3LYP/ 6-31G level were carried out on four conformers with lowest energy from the RHF optimization, namely the 9th, 6th, 8th and 5th, based on the geometries minimized at the RHF/6-31G level. The energies obtained from the B3LYP calculations are listed in Table 3.4. Both calculations have indicated that the 9th conformer is the most stable one with energies equal to -6509.70422au and -6535.8694au with the RHF and B3LYP methods, respectively, as shown in Tables 3.3 and 3.4.

Table 3.3. Energy levels of the first ten conformers with the RHF/6-31G(d) method

Conformer	Energy (Hartrees/molecule)
1	-6509.69414
2	-6509.68942
3	-6509.68747
4	-6509.69469
5	-6509.69579
6	-6509.70246
7	-6509.69317
8	-6509.69835
9	-6509.704224
10	-6509.69096

Table 3.4. B3LYP/6-31g(d) energy for four conformers with lowest RHF/6-31g(d) energy

Conformer	9th	6th	8th	5th
Energy (Hartree)	-6535.86941	-6535.86789	-6535.86718	-6535.86023

49

3.1.2. Construction of the two Guest molecules

No conformational searches were carried out for the molecules of adamantaneacetic acid (AD) and ferroceneacetic acid (FC), partially because they are much smaller and have fixed or rigid conformations compared to the CD molecule. The structures for them are illustrated in Figures 5.22 and 5.23 in the Appendix respectively.

3.1.3. Construction of the complexes

There are many ways to construct the complexes formed by cyclodextrin (CD) and admantaneacetic acid (AD), and cyclodextrin and ferroceneaceticacid (FC). We first did it by carefully inserting the molecules of AD and FC into the center of the optimized CD molecule (referred as the "direct method"). In 3.2.3, we illustrate how to combine the Z-matrices of two individual molecules together by generating six more geometry parameters (controlling parameters). By varying these parameters, we can manipulate the relative orientation between the host molecule and the guest molecule systematically. The structures of the complexes constructed by the direct method are illustrated in Figures 5.24 and 5.25 in the Appendix.

3.2. Computations of the association free energy and binding constant

3.2.1. Background

For a more detailed discussion about theoretically computing the Gibbs free energy change for chemical reactions, please refer to chapter 2 (Computational methods and Theoretical Background).

The binding constant K for forming a binary complex, AB, from molecules A and B is defined as below:

$$A + B \Leftrightarrow AB \qquad (3.1)$$

$$K = \frac{[AB]}{[A][B]} \qquad (3.2)$$

$$K = EXP\left(-\frac{\Delta_r G^0}{RT}\right) \qquad (3.3)$$

$\Delta_r G^0$ is the Gibbs free energy change for the complexation reaction, which by most definitions, is the Gibbs free energy of formation of product, AB minus the sum of Gibbs free energy of formation of the two reactants, A and B. But here we replace the Gibbs free energy of formation $\Delta_f G^0$ with the Gibbs free energy G at 298.15K relative to that of the completely separated nuclei and electron at rest at 0K. Please refer to the Computational Methods and Theoretical Background section for more discussions.

$$\Delta_r G^0 = G(AB) - G(A) - G(B) \qquad (3.4)$$

In Equation 3.4 G(AB), G(A) and G(B) are the Gibbs free energy at 298.15K relative to that of the state of nuclei and electrons infinitely far apart, which can be computed through the vibrational analysis for AB, A and B respectively, as discussed in chapter 2. For the systems we study, AB represents the complexes, CD-AD and CD-FC. For complex CD-AD, A and B are CD and AD. For complex CD-FC A and B are CD and FC. So for the systems under study, we have:

$$K(CD - AD) = \frac{[CD - AD]}{[CD][AD]} = EXP\left(-\frac{\Delta_r G^0(CD - AD)}{RT}\right) \qquad (3.5)$$

$$K(CD - FC) = \frac{[CD - FC]}{[CD][FC]} = EXP\left(-\frac{\Delta_r G(CD - FC)^0}{RT}\right) \qquad (3.6)$$

From the two equations above, if we assume at equilibrium (the system with all three species CD, AD and FC), that the concentration of AD and FC are equal, which is

experimentally achievable, then we can derive that the concentrations of the CD-AD and CD-FC complexes are related according to following equation:

$$\frac{[CD-AD]}{[CD-FC]} = \frac{K(CD-AD)}{K(CD-FC)} \quad (3.7)$$

And also we have the following relationships:

$$\Delta_r G^0(CD-AD) = G(CD-AD) - G(CD) - G(AD) \quad (3.8)$$

$$\Delta_r G^0(CD-FC) = G(CD-FC) - G(CD) - G(FC) \quad (3.9)$$

For the above species, CD, AD, FC, CD-AD and CD-FC, we optimized all the geometries followed by vibrational analyses. By doing so we can obtain electronic energies plus those thermodynamic quantities such as zero point energy, internal energy (U), enthalpy (H), entropy (S) and gibbs free energy (G) for all species. So in all tables listing those absolute quantities in this chapter, we include values both with and without electronic contribution. For instance for those entries only labeled "zero point", "internal", "enthalpy", etc we refer to those without electronic contributions; but for those presented as "elec + zero point", "elec + enthalpy", "elec + internal" etc indicate that those are the quantities including electronic energy. With all this data, we can then calculate $\Delta_r G^0$ (at 298.15K) for both complexes CD-AD and CD-FC. We intend to determine the energetically most favorable structures for the complexes of CD-AD and CD-FC in the gas phase, which thermodynamically are those yielding minimum $\Delta_r G^0$ (most negative). And we use the geometries for solvent phase calculations to determine the Gibbs free energy change in solution.

From the paper of Rekharsky and Inoue[12], we have the following thermodynamic data for complexes formed between β cyclodextrin and some guest molecules:

52

Table 3.5. Thermodynamic data for complexes of CD and various guest molecules

Host	Guest	solvent	ΔG^0 (kj/mol)
β-CD	1-adamantanecarboxylic acid	H_2O	-32.4
β-CD	ferrocenecarboxylate	H_2O(pH8.6)	-19.0 ± 0.3
β-CD	ferrocenecarboxylate	H_2O(pH8.6, 8M urea)	-16.2 ± 0.3
β-CD	ferrocenylalkyldimethylammonium	H_2O(0.05 M N_aCl)	-19.7 ± 0.4
β-CD	ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 2M urea)	-18.3 ± 0.4
β-CD	ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 4M urea)	-17.9 ± 0.4
β-CD	ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 6M urea)	-17.1 ± 0.4
β-CD	ferrocenylalkyldimethylammonium	H_2O(0.1 M N_aCl; 8M urea)	-16.3 ± 0.4
β-CD	ferrocenylalkyldimethylammonium	H_2O(pH 2.6-6.5)	-19.4 ± 0.4
β-CD	ferrocenylalkyldimethylammonium	H_2O(pH2.6)	-21.0 ± 0.4
β-CD	ferrocenylalkyldimethylammonium	H_2O(pH6.5)	-19.4 ± 0.4

We note that the β-cyclodextrin used by Rekharsky and Inoue is the natural one[12] not the thiolated one that we are using in this work and that the guest molecules are very similar, but not exactly the same as ours. However, since our systems are structurally very similar to those listed in Table 3.5, we expect that our calculated values to be comparable to those on the table. But we must keep in mind that all the data shown in Table 3.5 are those in aqueous solution, so we need to be careful when comparing those values to the results for the gas phase.

We have carried out calculations for the individual molecules and complexes using different theoretical models in both the vacuum and aqueous phases. The MP2 and DFT (B3LYP) methods can be applied to provide better electronic energies based on the HF optimized geometries. Both methods account for the effects of dispersion interactions and electronic correlation which the HF method omits. Experimental work implies that the binding between CD and AD is stronger than that between CD and FC, which means the ΔG^0 for the formation of the CD-AD complex would be more negative than that for the CD-FC complex. In section 3.2.3.1, we describe how we systematically control the relative position and orientation between the host molecule (CD) and guest molecules

(AD or FC) by making use of the z-matrix. The z-matrix or internal coordinates is a way to describe the molecular geometry. It specifies the position of an atom in the molecule relative to other atoms through parameters like bond distance, bond angle and torsional angle. These parameters determine the relative position of atoms in the molecule and they are independent with the choice of the origin (reference)

A ZINDO[56-58] semiempirical method was used along with a rough population analysis in order to study the origin (or lack thereof) the fluorescence in these species. This allowed us to decipher the relationship between the excited states of free cyclodextrin and the two complexes that it forms with adamantaneacetic acid and ferroceneacetic acid. The detailed results are presented in next section.

3.2.2. Computational results for complexes constructed by direct insertion

3.2.2.1. Results from molecular mechanics

Due to the size of our system, we started our computing task with the molecular mechanics method (MM) to get some rough idea about how resource-consuming this project is. We also intended to get an initial estimate of the energies to obtain some reasonably good starting structures for later computations at a higher theoretical level. We used two different MM models with two different force fields (the Dreiding and UFF parameterizations). The results from the MM methods are shown in Tables 3.6 through 3.8.

Table 3.6. Dreiding force field energies for all species

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.1982	0.2976	0.2300	1.5006	1.4322
Internal	1.2771	0.3085	0.2408	1.5922	1.5235
Enthalpy	1.2781	0.3095	0.2417	1.5932	1.5244
Free energy	1.0784	0.2605	0.1920	1.3675	1.2999
Entropy(cal/mol*K^{-1})	420.188	103.126	104.770	474.944	472.583

Table 3.7. UFF force field energies for all species

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.3251	0.3065	0.2552	1.6378	1.5668
Internal	1.3930	0.3158	0.2643	1.7162	1.6456
Enthalpy	1.3939	0.3167	0.2653	1.7171	1.6465
Free energy	1.2198	0.2708	0.2181	1.5211	1.4496
Entropy(cal/mol*K^{-1})	366.496	96.599	99.289	412.620	414.525

.

Table 3.8. Association energies for CD-AD and CD-FC from MM calculations

	CD-AD (dreiding)	CD-FC (dreiding)	CD-AD (uff)	CD-FC (uff)
ΔU^0(kj/mol)	17.22	14.64	19.60	-30.74
ΔH^0(kj/mol)	14.74	12.16	17.12	-33.22
ΔG^0(kj/mol)	75.09	77.51	80.10	30.74
ΔS^0(kj/mol*K^{-1})	-0.202	-0.219	-0.211	-0.214
K_{eq}	6.99E-14	2.63E-14	2.48E-31	1.79E-12

Note:

For the MM calculations with the UFF force field, the individual ferrocene molecule could only be completed by using the Newton Raphson optimization algorithm instead of the default RFO (Rational Optimization step) algorithm.

Because the electronic energies given by MM calculations are suspicious (force fields do not have terms for the electronic energy), we exclude the electronic contribution from all energetic terms.

These force fields are obviously too crude for predicting accurate thermodynamical properties. However, they can be used to quickly generate a set of initial or test conformations that can be studied at higher levels of theory.

3.2.2.2. Results from ab initio methods

I. Results at the RHF/3-21G level

A. Gas phase calculation

Table 3.9. RHF/3-21G energies for all species

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.2609	0.3076	0.2270	1.5709	1.4904
Internal	1.3367	0.3181	0.2398	1.6590	1.5807
Enthalpy	1.3377	0.3190	0.2408	1.6599	1.5817
Free energy	1.1464	0.2709	0.1841	1.4408	1.3581
Elec+Zero point	-6474.2410	-610.9681	-1863.7494	-7085.2287	-8338.0179
Elec+Internal	-6474.1651	-610.9577	-1863.7366	-7085.1406	-8337.9276
Elec+Enthalpy	-6474.1642	-610.9567	-1863.7356	-7085.1397	-8337.9266
Elec+Free energy	-6474.3554	-611.0049	-1863.7923	-7085.3588	-8338.1501
Eelectronic	-6475.5018	-611.2757	-1863.9764	-7086.7996	-8339.5083
Entropy(cal/mol*K^{-1})	402.605	101.345	119.305	461.266	470.509

Table 3.10. Association energies for CD-AD and CD-FC from RHF/3-21G calculations

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-57.87	-78.97
ΔU^0(kj/mol)	-46.87	-68.00
ΔH^0(kj/mol)	-49.36	-70.48
ΔS^0(kj/mol*K^{-1})	-0.179	-0.215
ΔG^0(kj/mol)	3.90	-6.34
K_{eq}	0.207	12.93

B. Solvent phase calculation at the RHF/3-21G level

In order to get a rough idea how the value of ΔG would change when the system is moved from the gas phase to the solvent phase (water as solvent), we carried out solvent phase calculations at the RHF/3-21G level. The calculations were completed at the RHF/3-21G level using the SCRF(PCM) solvation model. For all species, only single point calculations were completed based on the geometries optimized in gas phase at the RHF/3-21G level, and thus the gas phase vibrational analyses were used since no solvent phase optimizations have taken place.

Table 3.11. Solvation energy of each individual species at the RHF/3-21G level

Species	CD	AD	FC	CD-AD	CD-FC
ΔG^{sol}(kj/mol)	-36.58	-7.19	-15.38	-14.62	-34.47

Table 3.12. Solvent phase association Gibbs free energy and equilibrium constant of CD-AD and CD-FC at the RHF/6-31G level

	CD-AD	CD-FC
ΔG^0(aq) (kj/mol)	33.05	11.15
Keq(aq)	1.62E-6	0.0111

The results to indicate that as we move from the gas phase to the solvent phase, the ΔG tends to turn more positive because the combined ΔG of two individual molecules (CD, AD or CD, FC) switch to more negative values than that of the complexes (CD-AD or CD-FC).

Those results shown in Table 3.10 appear to be much better than those from the MM methods. However the $\Delta_r G$ values still compare poorly to the experimental values shown in Table 3.5 and CD-FC is predicted to have a higher binding constant than that for CD-AD, which does not agree with experimental observations.

To improve the results, we may change the binding sites for CD-AD and CD-FC and use larger basis sets and higher levels of theory. Before we find a way to systematically maneuver the relative orientation of the monomer molecules in the complexes, we decided to try some higher level models. We first tried MP2 single point calculations on

the geometries optimized with the RHF/3-21G method. Then the MP2 and B3LYP (and the RHF method) methods were employed with larger basis sets such as the 6-31G and 6-31G(d) sets.

II. MP2 correction for the RHF/3-21G calculation

The MP2 single point calculations were completed at the RHF/3-21g optimized geometries, and the thermal energy corrections are from the RHF/3-21g vibrational analysis.

Table 3.13. MP2 energies for all species after RHF/3-21G optimization

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.2609	0.3076	0.2270	1.5709	1.4904
Internal	1.3367	0.3181	0.2398	1.6590	1.5807
Enthalpy	1.3377	0.3190	0.2408	1.6599	1.5817
Free energy	1.1464	0.2709	0.1841	1.4408	1.3581
Mp2elec+Zero point	-6481.6861	-612.2900	-1865.2051	-7094.0142	-8346.9421
Mp2elec+Internal	-6481.6102	-612.2796	-1865.1922	-7093.9262	-8346.8518
Mp2elec+Enthalpy	-6481.6093	-612.2786	-1865.1913	-7093.9252	-8346.8509
Mp2elec+Free energy	-6481.8006	-612.3268	-1865.2480	-7094.1444	-8347.0744
Eelectronic(mp2)	-6482.9469	-612.5976	-1865.4321	-7095.5852	-8348.4325
Entropy(cal/mol*K^{-1})	402.605	101.345	119.305	461.266	470.509

Table 3.14. Association energies for CD-AD and CD-FC with MP2 corrections

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-106.53	-140.55
ΔU^0(kj/mol)	-95.53	-129.57
ΔH^0(kj/mol)	-98.01	-132.05
ΔS^0(kj/mol*K^{-1})	-0.179	-0.215
ΔG^0(kj/mol)	-44.76	-67.91
K_{eq}	6.95E+7	7.92E+11

These results imply tighter binding, but the value for ΔG^0(CD-FC) is still more negative than ΔG^0(CD-AD). But those values are greatly improved from those from the MM

58

methods, so we can carry those geometries of CD-AD and CD-FC to the higher level computations. We first tried using a larger basis set (6-31G) followed by computations using the B3LYP method.

III. Results at the RHF/6-31G level

Table 3.15. RHF/6-31G energies for all species

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.2657	0.3082	0.2282	1.5774	1.4972
Internal	1.3430	0.3186	0.2411	1.6660	1.5886
Enthalpy	1.3440	0.3195	0.2421	1.6669	1.5896
Free energy	1.1467	0.2714	0.1849	1.4422	1.3569
Elec+Zero point	-6506.5634	-614.0970	-1872.7908	-7120.6603	-8379.3826
Elec+Internal	-6506.4861	-614.0866	-1872.7779	-7120.5717	-8379.2911
Elec+Enthalpy	-6506.4852	-614.0857	-1872.7769	-7120.5708	-8379.2902
Elec+Free energy	-6506.6824	-614.1338	-1872.8341	-7120.7954	-8379.5228
Eelectronic	-6507.8291	-614.4052	-1873.0190	-7122.2377	-8380.8798
Entropy(cal/mol*K^{-1})	415.161	101.332	120.345	472.880	489.641

Table 3.16. Association energies for CD-AD and CD-FC from RHF/6-31G

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-8.69	-83.05
ΔU^0(kj/mol)	2.80	-71.22
ΔH^0(kj/mol)	0.32	-73.70
ΔS^0(kj/mol*K^{-1})	-0.182	-0.192
ΔG^0(kj/mol)	54.74	-16.47
K_{eq}	2.54E-10	771.86

These results show no improvement over those from the RHF/3-21G method in that they show even less bonding for the CD-AD complex which does not agree with experimental observations.

Because the practical counter-parts of our systems are those with cyclodextrin attached to a nano-particle through sulfur atoms on cyclodextrin, we use the redundant coordinate

option in Gaussian to fix the coordinates of these hydrogen atoms. Thus these chosen hydrogen atoms were not allowed to move throughout the optimization process to simulate the situation that the cyclodextrin is anchored to the nano-particle substrate.

IV.   Results from the RHF/6-31G with redundant coordinates

Table 3.17. Energies for all species at the RHF/6-31G level with redundant coordinates

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.2656	0.3082	0.2282	1.5774	1.4972
Internal	1.3433	0.3186	0.2411	1.6660	1.5886
Enthalpy	1.3443	0.3195	0.2421	1.6669	1.5896
Free energy	1.1426	0.2714	0.1849	1.4423	1.3569
Elec+Zero point	-6506.5556	-614.0970	-1872.7908	-7120.6603	-8379.3826
Elec+Internal	-6506.4779	-614.0866	-1872.7779	-7120.5717	-8379.2911
Elec+Enthalpy	-6506.4770	-614.0857	-1872.7769	-7120.5708	-8379.2902
Elec+Free energy	-6506.6786	-614.1338	-1872.8341	-7120.7954	-8379.5228
Eelectronic	-6507.8212	-614.4052	-1873.0190	-7122.2377	-8380.8798
Entropy(cal/mol*K^{-1})	424.356	101.332	120.345	472.840	489.674

Table 3.18. Association energies for CD-AD and CD-FC from RHF/6-31G calculations with redundant coordinates

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-29.46	-103.81
ΔU^0(kj/mol)	-18.80	-92.81
ΔH^0(kj/mol)	-21.28	-95.29
ΔS^0(kj/mol*K^{-1})	-0.221	-0.230
ΔG^0(kj/mol)	44.66	-26.64
K_{eq}	1.50E-8	4.64E+4

It is clear that this gives little improvement, since with a $\Delta_r G$ of +44.66 kj/mol for CD-AD, we have a much weaker interaction between CD and AD compared to that between CD and FC with $\Delta_r G$ equal to -26.64kj/mol. And since we know that the RHF method does not include dispersion and correlation effects, we decided to try the DFT method using the B3LYP method.

V.   Results at the B3LYP/6-31G level

Table 3.19. B3LYP energies for all species

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.1677	0.2884	0.2133	1.4591	1.3830
Internal	1.2493	0.2996	0.2255	1.5529	1.4779
Enthalpy	1.2503	0.3005	0.2265	1.5539	1.4789
Free energy	1.0444	0.2509	0.1728	1.3213	1.2452
Elec+Zero point	-6533.3808	-618.1412	-1878.1713	-7151.5361	-8411.5697
Elec+Internal	-6533.2991	-618.1301	-1878.1591	-7151.4423	-8411.4748
Elec+Enthalpy	-6533.2982	-618.1291	-1878.1581	-7151.4414	-8411.4738
Elec+Free energy	-6533.5041	-618.1788	-1878.2118	-7151.6740	-8411.7075
Eelectronic	-6534.5485	-618.4297	-1878.3846	-7152.9952	-8412.9527
Entropy(cal/mol*K^{-1})	433.274	104.438	113.024	489.549	491.815

Table 3.20. Energy changes for CD-AD and CD-FC from B3LYP/6-31G calculations

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-44.97	-51.58
ΔU^0(kj/mol)	-34.41	-43.49
ΔH^0(kj/mol)	-36.88	-45.97
ΔS^0(kj/mol*K^{-1})	-0.202	-0.228
ΔG^0(kj/mol)	23.21	22.01
K_{eq}	8.55E-5	1.39E-4

And once again this method predicts no binding. It was becoming clear that cyclodextrin may be interacting with AD and FC at different sites and that a more thorough conformational search will be necessary. Each conformation could be trapped in a local minimum, but not the global minimum. It would be ideal if we can systematically change the binding site at which the host molecule (CD) and guest molecules (AD and FC) interact and also adjust the relative orientation of the host molecule and guest molecule. Then we can figure out the optimum site and orientation which combine to produce the best results. The detailed procedures are discussed in next section.

3.2.3. Results for the systematic construction of the complexes

3.2.3.1. Systematic procedure for constructing the binary complexes

A new approach was used to build the two complexes between cyclodextrin and AD and between cyclodextrin and FC under which we can systematically adjust the binding site and the relative orientation between the host molecule (cyclodextrin) and the guest molecules (AD and FC). The method to achieve this follows:

First generate the z-matrices for the optimized CD, AD, and FC respectively; then attach the optimized z-matrices of AD and FC to the optimized z-matrix of CD respectively; in this process, we need to create six new internal coordinates, namely one bond length, two bond angles and three torsional angles in order to combine the two z-matrices into a single whole one. These six parameters dictate the binding site and the relative orientation of cyclodextrin to AD and FC. So we refer them as the six controlling parameters. Special care is needed for re-numbering the atoms and re-naming the symbols in the z-matrices of AD and FC before they are attached to the z-matrix of cyclodextrin. All of the six new internal coordinates involve at least on atom from each of the two molecules in the binary complex.

In order to illustrate the whole process better, we demonstrate here with graphics how we built up the single z-matrix for the complexes by putting together two z-matrices, one from the host molecule (CD) and another from one of the two guest molecules (AD and FC).

First, we optimized all individual molecules with the method of choice (RHF, MP2 and B3LYP). Then for the final optimized geometries, we need to generate the z-matrix for the complex for which the program, MOLDEN, can greatly facilitate the process. We

change the order of how atoms are arranged in the z-matrices for molecules of AD and FC such that those atoms involved in the new controlling parameters occur first. Next, we illustrate the whole process using an example from the RHF calculations. Only a part of each z-matix is shown below with part of the connectivity section and all of the variable-value section excluded.

The partial z-matrix of CD:

```
h   52 hc144     45 hcc144     44 dih144
h   53 hs145     52 hsc145     45 dih145
h   50 ho146     47 hoc146     48 dih146
h   49 ho147     48 hoc147     43 dih147
```

The partial z-matrix of AD (Represented by the top):

```
c
c    1 cc2
c    1 cc3       2 ccc3
h    1 hc4       2 hcc4        3 dih4
```

The partial z-matrix of FC (Represented by the top):

```
c
c    1 cc2
c    1 cc3       2 ccc3
o    2 oc4       1 occ4        3 dih4
```

Now, if we attach the z-matrix from AD and FC to the end of the z-matrix of CD respectively in order to form the two z-matrices for the two complexes, they would appear as follows:

CD-AD:

```
h   52 hc144     45 hcc144     44 dih144
h   53 hs145     52 hsc145     45 dih145
h   50 ho146     47 hoc146     48 dih146
h   49 ho147     48 hoc147     43 dih147
c
c    1 cc2
c    1 cc3       2 ccc3
h    1 hc4       2 hcc4        3 dih4
```

CD-FC:

```
h   52 hc144      45 hcc144      44 dih144
h   53 hs145      52 hsc145      45 dih145
h   50 ho146      47 hoc146      48 dih146
h   49 ho147      48 hoc147      43 dih147
c
c    1 cc2
c    1 cc3         2 ccc3
o    2 oc4         1 occ4         3 dih4
```

Apparently, in order to make them a whole z-matrix, we need to use six more internal coordinates (referred to as **controlling parameters**) for each new z-matrix, namely one bond length, two bond angles and three torsional angles. The completed matrices are shown below.

CD-AD:

```
h   52 hc144      45 hcc144      44 dih144
h   53 hs145      52 hsc145      45 dih145
h   50 ho146      47 hoc146      48 dih146
h   49 ho147      48 hoc147      43 dih147
c    7 oc148      14 ooc148      21 dih148
c  148 adcc2       7 cco149      14 dih149
c  148 adcc3     149 adccc3       7 dih150
h  148 adhc4     149 adhcc4     150 addih4
```

CD-FC:

```
h   52 hc144      45 hcc144      44 dih144
h   53 hs145      52 hsc145      45 dih145
h   50 ho146      47 hoc146      48 dih146
h   49 ho147      48 hoc147      43 dih147
c    7 oc148      14 ooc148      21 dih148
c  148 fecc2       7 cco149      14 dih149
c  148 fecc3     149 feccc3       7 dih150
o  149 feoc4     148 feocc4     150 fedih4
```

64

Those highlighted six parameters are those constructed for completing the large z-matrix for the complexes. These are oc148 (bond length), ooc148 and cco149 (bond angles) and dih148, dih149 and dih150 (torsional angles). As for these six controlling parameters, one worth special mentioning here is the dih148. It turns out that the magnitude of dih148 governs the distance between the overall guest molecule and the mean plane of cyclodextrin such that the smaller the magnitude, the closer the distance; and the sign of dih148 dictates which side of cyclodextrin the guest molecule is on with a positive sign for one side and a negative sign for the other side.

The following precautions must be taken when generating the new larger z-matrix:
1. Because AD and FC have far less atoms than CD does (32 and 27 atoms compared to 147 atoms), it is better attaching the z-matrices from AD and FC to the one for CD. By doing so, we reduce the changes needed in the renumbering.
2. Due to the 147 atoms in CD, the atom sequence number in the z-matrices of AD and FC start from 148. We add 147 to the sequence number in the original individual matrices of AD and FC in order to get the right sequence number for the corresponding atoms in the large z-matrix for the complex. But those in the z-matrix of CD are kept unchanged.
3. In order to avoid having the same acronym used for more than one coordinate, all the variable names for the z-matrices of AD and FC have an extra "AD" and "FC" added as a prefix.
4. Three oxygen atoms (7, 14, 21) from CD, three Carbon atoms from both AD and FC are chosen to construct the six new internal coordinates.   All these atoms are circled in yellow in the two figures below. The three carbon atoms chosen from AD and FC are the first three atoms in the original AD and FC z-matrices but they are numbered 148, 149 and 150 in the final large z-matrices for the complexes.

By accident, the three carbon atoms chosen from AD for the B3LYP calculations are different from those of AD for the RHF calculations. They are also circled in yellow in the figure for AD for the B3LYP calculations. Certainly it does not affect the

results. The choices for the three carbon atoms in FC are the same for both the RHF and B3LYP calculations.
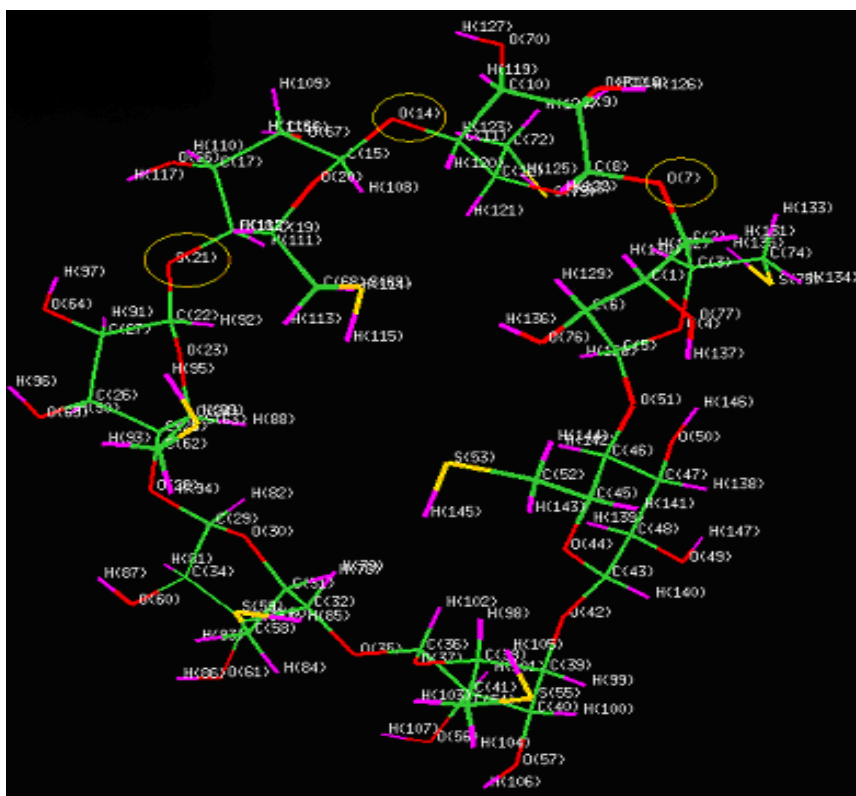
Figure 3.3. CD unit for constructing complexes for all calculations
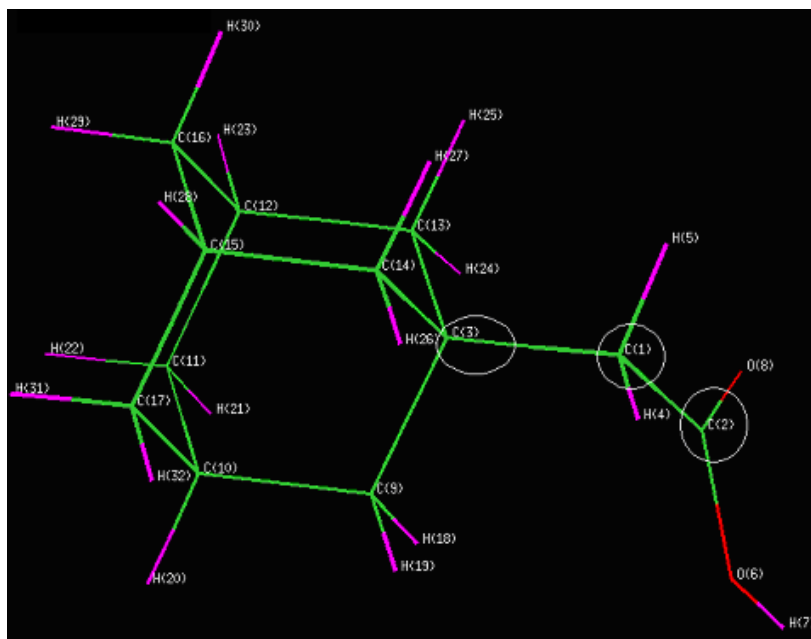
Figure 3.4. AD unit for constructing CD-AD complex for the RHF calculation
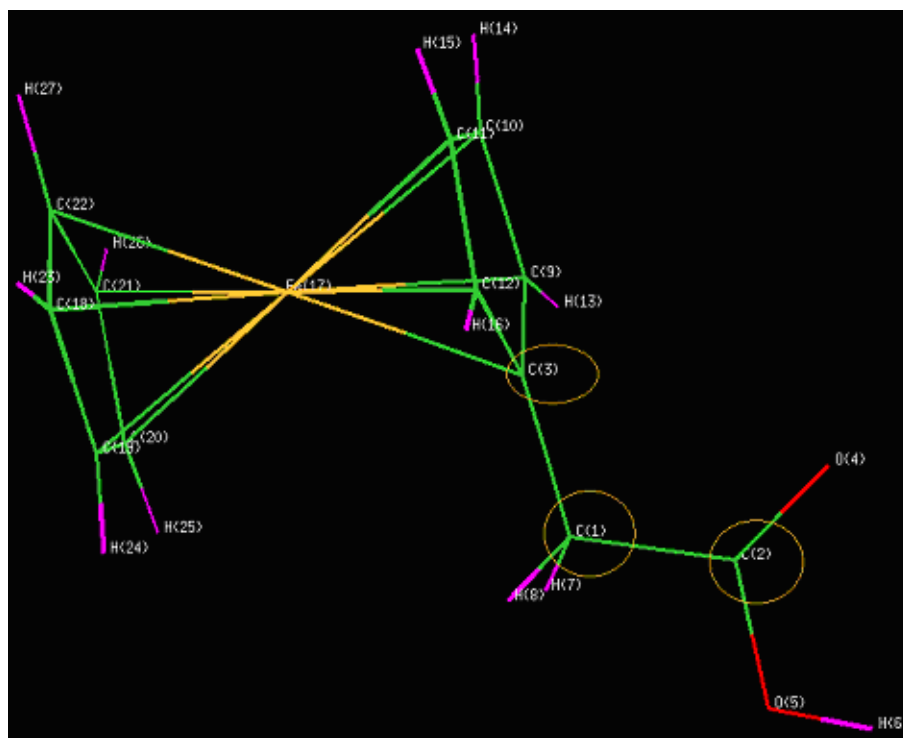
Figure 3.5. FC unit for constructing CD-FC complex for all calculations
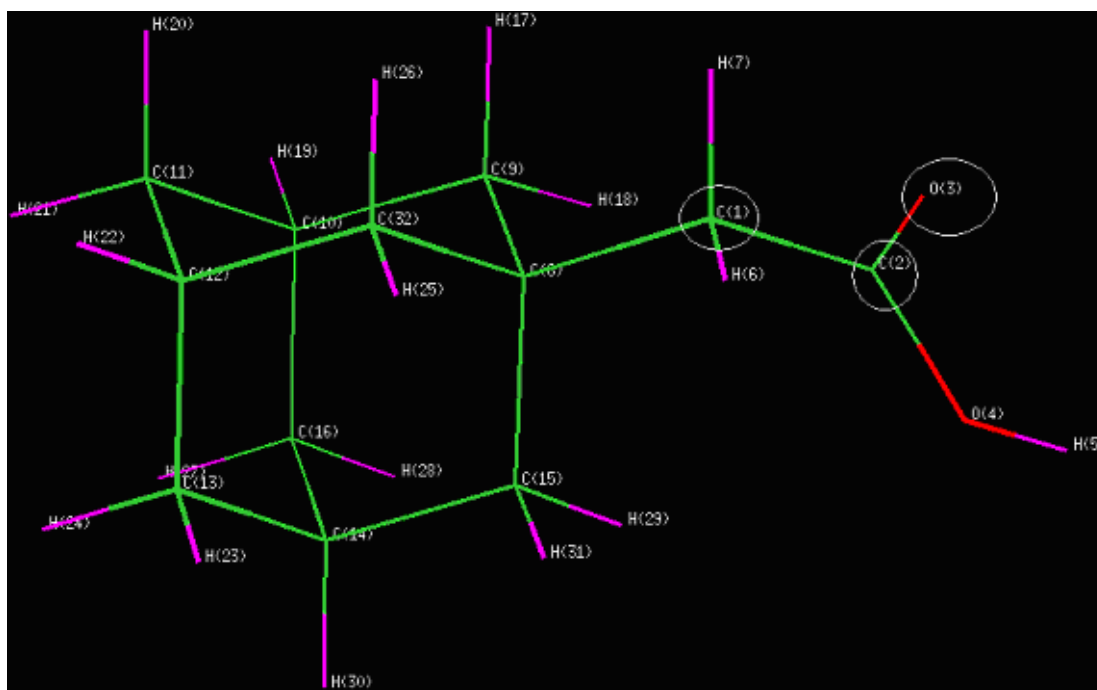
Figure 3.6. AD unit for constructing CD-AD complex for the B3LYP calculation

After constructing the z-matrices for the complexes, we are then able to adjust the binding site and the relative orientation between the host molecule (CD) and the guest molecules (AD or FC) by varying these six controlling parameters. Theoretically we can obtain all possible conformations, and we would need a mechanism to explore all possible combinations of these six quantities. However, from a practical point of view due to the size of these complexes, we are only able to investigate a smaller set of initial geometries. Details of computational results are discussed in next section.

3.2.3.2. Computational results in vacuum

I. Results from partial optimizations

We first carried out a partial optimization using both the RHF/6-31g and B3LYP/6-31g methods, in which only the six controlling parameters are optimized with the remainder left unchanged. The initial values are assigned as reasonable guesses and given below. And we used the same set of initial values for CD-AD and CD-FC for both the RHF and B3LYP calculations. Those initial values were chosen intuitively.

Table 3.21. Initial values of the controlling parameters for CD-AD

oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
7.30	75.00	-45.00	35.00	135.00	95.00

Table 3.22. Initial values of the controlling parameters for CD-FC

oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
7.30	75.00	-45.00	35.00	135.00	95.00

Table 3.23. Final values of the controlling parameters for CD-AD from RHF partial optimization

oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
8.17	105.95	-24.11	18.58	91.12	58.22

Table 3.24. Final values of the controlling parameters for CD-FC from RHF partial optimization

oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
7.40	93.35	-37.39	24.82	95.12	86.97

Table 3.25. Energies for all species from RHF partial optimization

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.2657	0.3082	0.2282	1.5770	1.4969
Internal	1.3430	0.3186	0.2411	1.6658	1.5885
Enthalpy	1.3440	0.3195	0.2421	1.6668	1.5894
Free energy	1.1467	0.2714	0.1849	1.4398	1.3537
Elec+Zero point	-6506.5634	-614.0970	-1872.7908	-7120.6914	-8379.3832
Elec+Internal	-6506.4861	-614.0866	-1872.7779	-7120.6025	-8379.2915
Elec+Enthalpy	-6506.4852	-614.0857	-1872.7769	-7120.6016	-8379.2906
Elec+Free energy	-6506.6824	-614.1338	-1872.8341	-7120.8286	-8379.5264
Eelectronic	-6507.8291	-614.4052	-1873.0190	-7122.2683	-8380.8800
Entropy(cal/mol*K^{-1})	415.161	101.332	120.345	477.752	496.206

Table 3.26. Association energies for CD-AD and CD-FC from RHF partial optimization

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-89.17	-83.67
ΔU^0(kj/mol)	-78.07	-72.26
ΔH^0(kj/mol)	-80.54	-74.74
ΔS^0(kj/mol*K^{-1})	-0.162	-0.164
ΔG^0(kj/mol)	-32.21	-25.71
K_{eq}	4.42E+5	3.21E+4

Table 3.27. Final values of controlling parameters for CD-AD from B3LYP partial optimization

oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
7.29	76.11	-22.93	56.30	103.45	133.72

Table 3.28. Final values of the controlling parameters for CD-FC from B3LYP partial optimization

oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
7.74	88.96	-37.94	38.51	131.23	82.19

Table 3.29. Energies for all species from B3LYP partial optimization

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.1677	0.2884	0.2134	1.4576	1.3843
Internal	1.2493	0.2996	0.2266	1.5520	1.4803
Enthalpy	1.2503	0.3005	0.2275	1.5530	1.4813
Free energy	1.0444	0.2509	0.1711	1.3187	1.2438
Elec+Zero point	-6533.3808	-618.1412	-1878.1715	-7151.5431	-8411.5848
Elec+Internal	-6533.2991	-618.1301	-1878.1583	-7151.4487	-8411.4887
Elec+Enthalpy	-6533.2982	-618.1291	-1878.1574	-7151.4477	-8411.4878
Elec+Free energy	-6533.5041	-618.1788	-1878.2138	-7151.6820	-8411.7252
Eelectronic	-6534.5485	-618.4297	-1878.3849	-7153.0007	-8412.9690
Entropy(cal/mol*K^{-1})	433.274	104.438	118.777	493.135	499.787

Table 3.30. Association energies from B3LYP partial optimization

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-59.32	-93.71
ΔU^0(kj/mol)	-51.10	-82.00
ΔH^0(kj/mol)	-53.58	-84.48
ΔS^0(kj/mol)	-0.187	-0.219
ΔG^0(kj/mol)	2.04	-19.28
K_{eq}	0.439	2.39E+3

The RHF/6-31G results look much better in that it gives rise to more negative ΔG^0. However we had expected that B3LYP would yield more favorable results since it takes into account the correlation effect which is not included in the RHF method, which should result in stronger binding (ΔG^0 of more negative value).

In order to reveal the cause of this unexpected result, we examine the final structures from the RHF and B3LYP partial optimizations, which are shown in Figures 3.7 through 3.10 below.
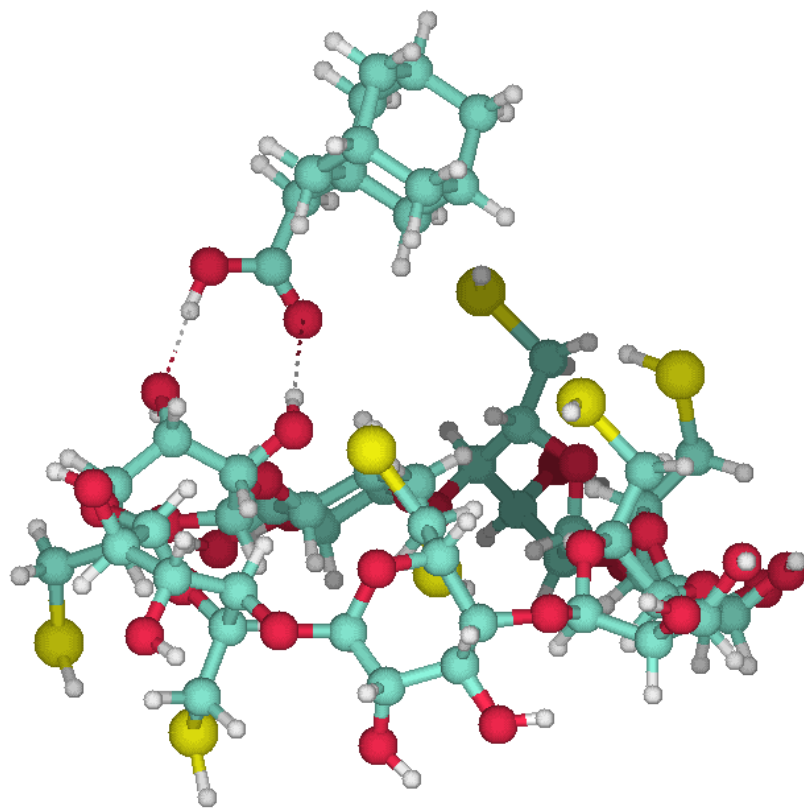
Figure 3.7. Final geometry of the RHF partially optimized CD-AD
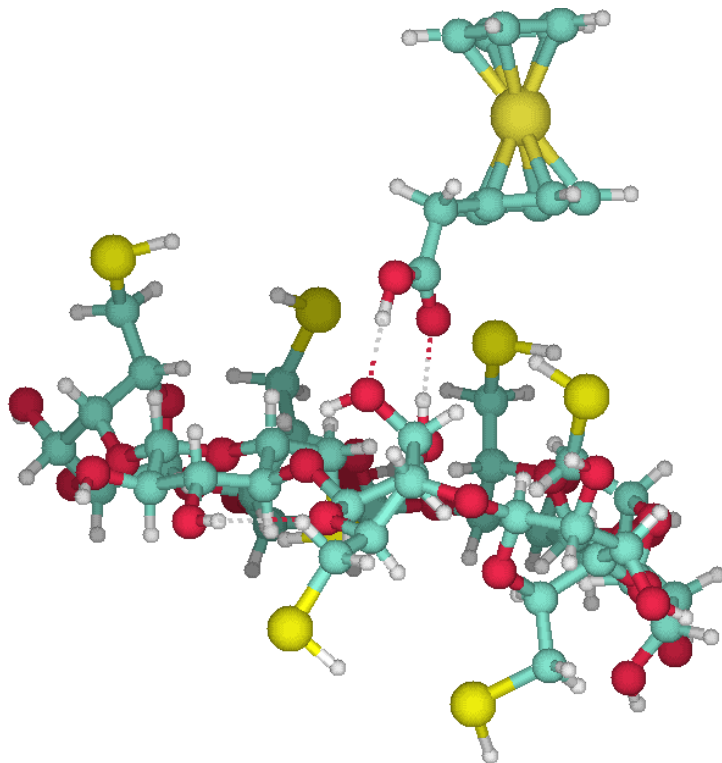
Figure 3.8. Final geometry of the RHF partially optimized CD-FC
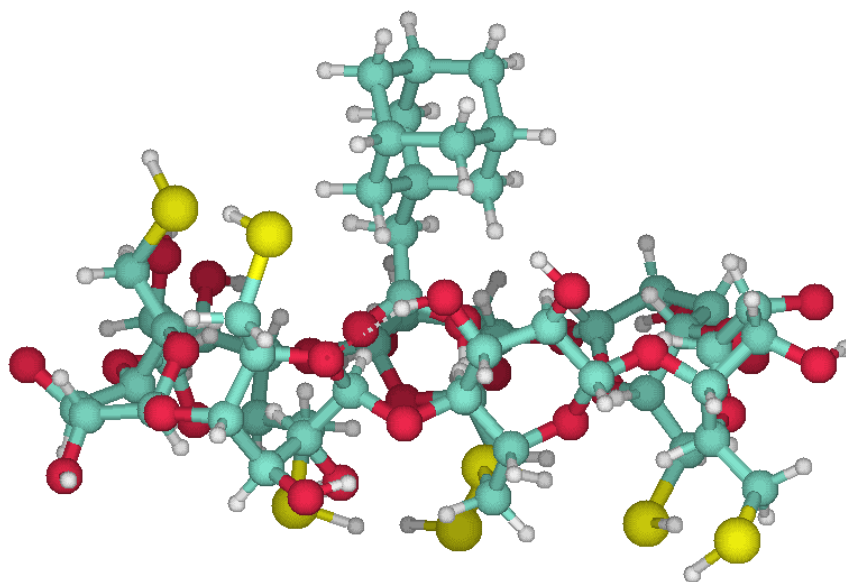
Figure 3.9. Final geometry of the B3LYP partially optimized CD-AD
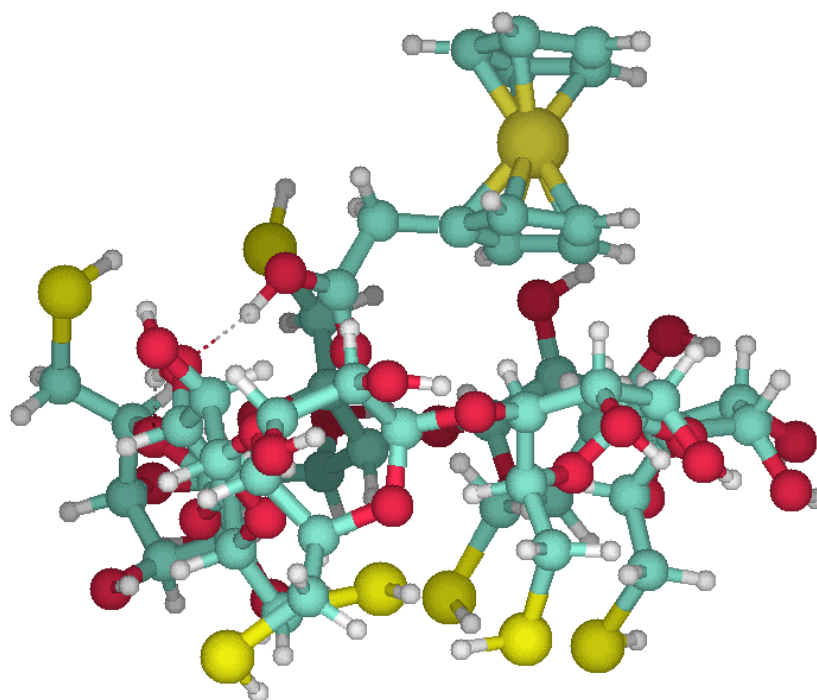
Figure 3.10. Final geometry of the B3LYP partially optimized CD-FC

From these figures, the most notable fact is that each of these two structures partially optimized by the RHF method has two hydrogen bonds present, but those partially

optimized by the B3LYP method have none for CD-AD or only one for CD-FC. This finding implies that the formation of multiple hydrogen bonds may play a critical role here. We must next complete the full optimization of all internal coordinates. We can take the final RHF partially optimized geometries as initial structures or we can intentionally generate two hydrogen bonds by varying the six controlling parameters in the partially optimized B3LYP z-matrices. In order to be more cautious, we tried both. Totally we have six different initial structures for the B3LYP optimizations and four for the RHF optimizations. And clearly all these different initial structures are dictated by different sets of the six controlling parameters.

II. Results of the full optimizations

A. RHF/6-31G

Four full RHF optimizations were completed for the complexes of CD-AD and two for CD-FC using RHF different starting geometries, which can be controlled by varying the six parameters. The values assigned for those six controlling parameters were decided by varying them in order to get the guest molecule as deep as possible in to the CD cavity but avoid any chemical bond formation between the host and the guest molecules. And those successful sets were chosen for the computations.

The first RHF full optimization used the final geometry from the RHF partial optimization as the initial structure; values for those six parameters are listed in Table 3.31. The results are shown in Tables below.

Table 3.31. Initial values of the controlling parameters for CD-AD

	oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	8.17	105.95	-24.11	18.58	91.12	58.22
II	7.70	60.00	14.00	130.00	-15.00	-109.00
III	9.50	75.00	-17.00	85.00	35.00	95.00
IV	6.50	125.00	-51.00	35.00	25.00	90.00

Table 3.32. Initial values of the controlling parameters for CD-FC

	oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	7.40	93.35	-37.39	24.82	95.12	86.97
II	6.00	60.00	14.00	20.00	-2.00	-90.00

Table 3.33. Final RHF/6-31G values of the controlling parameters for CD-AD

	oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	8.17	105.96	-24.12	18.58	91.11	58.22
II	7.46	60.92	37.49	114.92	-45.53	-176.64
III	9.97	62.19	-15.36	88.57	63.17	136.66
IV	6.27	61.30	-77.39	32.90	92.08	126.41

Table 3.34. Final RHF/6-31G values of the controlling parameters for CD-FC

	oc148(angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	7.40	93.35	-37.38	24.82	95.12	86.97
II	7.70	70.75	36.39	54.49	-34.61	-89.88

Table 3.35. RHF/6-31G energies for CD, AD and FC

(Hartrees/particle)	CD	AD	FC
Zero point	1.2657	0.3082	0.2282
Internal	1.3430	0.3186	0.2411
Enthalpy	1.3440	0.3195	0.2421
Free energy	1.1467	0.2714	0.1849
Elec+Zero point	-6506.5634	-614.0970	-1872.7908
Elec+Internal	-6506.4861	-614.0866	-1872.7779
Elec+Enthalpy	-6506.4852	-614.0857	-1872.7769
Elec+Free energy	-6506.6824	-614.1338	-1872.8341
Eelectronic	-6507.8291	-614.4052	-1873.0190
Entropy(cal/mol*K^{-1})	415.161	101.332	120.345

Table 3.36. RHF/6-31G energies for CD-AD complexes at various orientations

(Hartrees/particle)	CD-AD I	CD-AD II	Cyclo-Ad III	Cyclo-Ad IV
Zero point	1.5770	1.5759	1.5763	1.5765
Internal	1.6658	1.6655	1.6657	1.6657
Enthalpy	1.6668	1.6665	1.6666	1.6666
Free energy	1.4398	1.4409	1.4409	1.4416
Elec+Zero point	-7120.6914	-7120.6787	-7120.6719	-7120.6811
Elec+Internal	-7120.6025	-7120.5890	-7120.5826	-7120.5919
Elec+Enthalpy	-7120.6016	-7120.5881	-7120.5816	-7120.5909
Elec+Free energy	-7120.8286	-7120.8137	-7120.8074	-7120.8159
Eelectronic	-7122.2683	-7122.2545	-7122.2482	-7122.2576
Entropy(cal/mol*K^{-1})	477.752	474.824	475.117	473.526

Table 3.37. RHF/6-31G energies for CD-FC complexes at various orientations

(Hartrees/particle)	CD-FC I	CD-FC II
Zero point	1.4969	1.4965
Internal	1.5885	1.5880
Enthalpy	1.5894	1.5890
Free energy	1.3537	1.3565
Elec+Zero point	-8379.3832	-8379.3691
Elec+Internal	-8379.2915	-8379.2775
Elec+Enthalpy	-8379.2906	-8379.27656
Elec+Free energy	-8379.5264	-8379.5090
Eelectronic	-8380.8800	-8380.8656
Entropy(cal/mol*K^{-1})	496.224	489.228

Table 3.38. RHF/6-31G association energies for CD-AD at various orientations

CD-AD	I	II	III	IV
ΔE_{ele}(kj/mol)	-89.17	-52.94	-36.34	-60.89
ΔU^0(kj/mol)	-78.07	-42.68	-25.70	-50.18
ΔH^0(kj/mol)	-80.54	-45.15	-28.18	-52.66
ΔS^0(kj/mol*K^{-1})	-0.162	-0.174	-0.173	-0.180
ΔG^0(kj/mol)	-32.21	6.84	23.45	0.953
K_{eq}	4.42E+5	0.0633	7.76E-5	0.681

Table 3.39. RHF/6-31G association energies for CD-FC at various orientations.

CD-FC	I	II
ΔE_{ele}(kj/mol)	-83.67	-45.77
ΔU^0(kj/mol)	-72.26	-35.51
ΔH^0(kj/mol)	-74.74	-37.99
ΔS^0(kj/mol*K^{-1})	-0.164	-0.194
ΔG^0(kj/mol)	-25.73	19.75
K_{eq}	3.24E+4	3.46E-4

These results from the first full optimization are almost identical to those from the RHF partial optimization, which indicates that throughout the first full optimization, the structures of the individual molecules in the complexes (CD, AD and FC) change very little, which can be easily noted from the similarity between geometries before and after optimization as shown from Figures 5.26 through 5.29 in the Appendix.

The results from the second full optimization are not as good as those from the first one in that the second one results in much more positive value for ΔG^0 for both CD-AD and CD-FC compared to those values obtained from the first optimization. So the best results from the RHF full optimizations are those from the first.

Figures 5.30 through 5.33 in the Appendix give the geometries for CD-AD and CD-FC before and after the second full optimization. And Figures 5.34 to 5.37 in the Appendix illustrate the geometries of CD-AD before and after the third and fourth optimizations. Clearly the co-occurrence of double hydrogen bonds and very negative ΔG from the first full optimization for both of the CD-AD and CD-FC complexes makes us further believe that the two hydrogen bonds are the main contribution for stabilizing the complexes.

B. B3LYP/6-31G full optimization

Six trial B3LYP calculations have been carried out on different initial structures for the complexes of CD-AD and CD-FC with different values for the controlling parameters. Like those for the RHF full optimization calculations, values used for the B3LYP method

were chosen by carefully varying all six parameters to move the guest molecules into the CD cavity but to avoid the formation of chemical bonds between the host molecule and the guest molecule. The fifth one used the corresponding final geometries from the RHF partial optimizations as the initial structures for both CD-AD and CD-FC. But because of the different way used to set up complexes for the B3LYP computations from those for the RHF computations, those values for the fifth CD-AD and CD-FC are not identical from those final values from the corresponding RHF computations. These values are listed below.

Table 3.40. Initial values assigned for the controlling parameters for CD-AD

	oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	7.60	100.00	55.44	70.00	-107.97	23.00
II	8.10	78.00	29.00	35.00	-50.00	65.00
III	8.10	78.00	-29.00	15.00	35.00	65.00
IV	6.00	45.00	-50.00	15.00	35.00	65.00
V	8.30	109.41	46.21	62.33	-107.97	31.87
VI	13.50	70.00	-34.00	85.00	100.00	120.00

Table 3.41. Initial values assigned for the controlling parameters for CD-FC

	oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	6.00	80.00	54.00	65.00	220.00	80.00
II	7.50	80.00	30.00	50.00	20.00	-90.00
III	6.80	57.50	-30.00	50.00	20.00	-80.00
IV	7.74	88.96	-37.94	38.51	131.23	82.19
V	9.63	99.79	35.86	51.27	-112.64	-92.79
VI	13.00	60.00	-30.00	95.00	135.00	100.00

Table 3.42. Final B3LYP/6-31G values of the controlling parameters for CD-AD

	oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	7.96	99.44	58.93	59.55	-114.53	-0.28
II	9.80	91.47	37.05	13.58	-17.64	48.14
III	4.96	75.81	-62.11	19.54	27.50	164.25
IV	6.11	53.94	-89.00	42.67	17.76	82.13
V	8.28	102.89	45.61	62.85	-112.10	42.53
VI	16.75	74.38	-22.00	28.24	85.32	63.79

Table 3.43. Final B3LYP/6-31G values of the controlling parameters for CD-FC

	oc148 (angs)	ooc148(deg)	dih148(deg)	cco149(deg)	dih149(deg)	dih150(deg)
I	7.12	81.03	41.64	48.73	13.92	-19.14
II	7.11	81.24	41.56	49.20	14.00	-21.16
III	7.47	64.36	-40.47	21.37	-29.72	16.62
IV	7.70	92.74	-31.40	39.44	126.39	83.17
V	9.55	96.25	35.09	52.51	-113.00	-105.16
VI	16.45	55.34	-10.84	38.78	174.70	-36.41

Table 3.44. B3LYP/6-31G energies for CD, AD and FC

(Hartrees/particle)	CD	AD	FC
Zero point	1.1677	0.2884	0.2134
Internal	1.2493	0.2996	0.2266
Enthalpy	1.2503	0.3005	0.2275
Free energy	1.0444	0.2509	0.1711
Elec+Zero point	-6533.3808	-618.1412	-1878.1715
Elec+Internal	-6533.2991	-618.1301	-1878.1583
Elec+Enthalpy	-6533.2982	-618.1291	-1878.1574
Elec+Free energy	-6533.5041	-618.1788	-1878.2138
Eelectronic	-6534.5485	-618.4297	-1878.3849
Entropy(cal/mol*K^{-1})	433.274	104.438	118.777

Table 3.45. B3LYP/6-31G energies for CD-AD at various orientations

(Hartrees/particle)	CD-AD I	CD-AD II	CD-AD III	CD-AD IV	CD-AD V	CD-AD VI
Zero point	1.4581	1.4583	1.4584	1.4583	1.4592	1.4584
Internal	1.5522	1.5524	1.5525	1.5527	1.5528	1.5524
Enthalpy	1.5532	1.5534	1.5535	1.5537	1.5537	1.5533
Free energy	1.3191	1.3199	1.3200	1.3191	1.3199	1.3185
Elec+Zero point	-7151.5468	-7151.5393	-7151.5466	-7151.5378	-7151.5616	-7151.5482
Elec+Internal	-7151.4527	-7151.4452	-7151.4525	-7151.4434	-7151.4680	-7151.4543
Elec+Enthalpy	-7151.4517	-7151.4443	-7151.4515	-7151.4424	-7151.4671	-7151.4533
Elec+Free energy	-7151.6858	-7151.6778	-7151.6850	-7151.6770	-7151.7009	-7151.6882
Eelectronic	-7153.0049	-7152.9977	-7153.0050	-7152.9961	-7153.0208	-7153.0067
Entropy(cal/mol*K^{-1})	492.694	491.442	491.495	493.738	492.215	494.227

Table 3.46. B3LYP/6-31G association energies for CD-AD at various orientations

CD-AD	I	II	III	IV	V	VI
ΔE_{ele}(kj/mol)	-70.29	-51.31	-70.60	-47.15	-112.07	-74.96
ΔU^0(kj/mol)	-61.56	-42.04	-61.04	-37.14	-101.85	-65.80
ΔH^0(kj/mol)	-64.04	-44.52	-63.52	-39.62	-104.33	-68.28
ΔS^0(kj/mol*K^{-1})	-0.188	-0.194	-0.193	-0.184	-0.190	-0.182
ΔG^0(kj/mol)	-7.873	13.211	-5.859	15.241	-47.567	-14.025
K_{eq}	23.951	0.00484	10.628	0.00214	2.1573E+8	286.610

Table 3.47. B3LYP/6-31G energies for CD-FC

(Hartrees/particle)	CD-FC I	CD-FC II	CD-FC III	CD-FC IV	CD-FC V	CD-FC VI
Zero point	1.3840	1.3838	1.3832	1.3836	1.3840	1.3845
Internal	1.4799	1.4798	1.4789	1.4790	1.4800	1.4800
Enthalpy	1.4809	1.4807	1.4798	1.4799	1.4810	1.4810
Free energy	1.2422	1.2420	1.2421	1.2438	1.2394	1.2406
Elec+Zero point	-8411.5759	-8411.5760	-8411.5652	-8411.5839	-8411.5889	-8411.5825
Elec+Internal	-8411.4800	-8411.4800	-8411.4695	-8411.4886	-8411.4929	-8411.4869
Elec+Enthalpy	-8411.4790	-8411.4791	-8411.4686	-8411.4876	-8411.4920	-8411.4860
Elec+Free energy	-8411.7177	-8411.7178	-8411.7063	-8411.7237	-8411.7335	-8411.7263
Eelectronic	-8412.9599	-8412.9598	-8412.9484	-8412.9675	-8412.9729	-8412.9669
Entropy(cal/mol*K^{-1})	502.317	502.399	500.367	496.855	508.448	505.845

Table 3.48. B3LYP/6-31G association energies for CD-FC at various orientations

CD-FC	I	II	III	IV	V	VI
ΔE_{ele}(kj/mol)	-69.66	-69.46	-39.47	-89.76	-103.92	-88.21
ΔU^0(kj/mol)	-59.03	-59.18	-31.62	-81.67	-93.01	-77.31
ΔH^0(kj/mol)	-61.51	-61.66	-34.10	-84.15	-95.49	-79.79
ΔS^0(kj/mol*K^{-1})	-0.208	-0.208	-0.216	-0.231	-0.182	-0.193
ΔG^0(kj/mol)	0.541	0.294	30.383	-15.281	-41.086	-22.145
K_{eq}	0.804	0.888	4.75E-6	475.566	1.5792E+7	7583.369

From the results above, it is apparent that for both CD-AD and CD-FC, the fifth and sixth complexes have the best structure given that they have the most negative ΔG^0 values. From Tables of 3.42 and 3.43, and Figures 5.38 through 5.61 in the Appendix, we note that for both CD-AD and CD-FC, the first, second and fifth complexes have the guest molecule (AD or FC) on the same side of CD, and the third, fourth and sixth have the

guest molecule on the other side of CD. The 1st, 2nd and 5th complexes have positive dih148 and the 3rd,4th and 6th complexes have negative dih148.

III. MP2 corrections

MP2 single point calculations were carried out to obtain a better description of the electronic and correlation energies for those structures from the first RHF full optimization as well as for the fifth and sixth B3LYP full optimizations since these give the lowest ΔG values.

A. MP2 correction to the RHF results

Table 3.49. MP2/6-31G energies for the RHF optimized species

(Hartrees/particle)	CD	AD	FC	CD-AD	CD-FC
Zero point	1.2657	0.3082	0.2282	1.5770	1.4969
Internal	1.3430	0.3186	0.2411	1.6658	1.5885
Enthalpy	1.3440	0.3195	0.2421	1.6668	1.5894
Free energy	1.1467	0.2714	0.1849	1.4398	1.3537
Elec+Zero point	-6514.0647	-615.4247	-1874.2646	-7129.5312	-8388.3690
Elec+Internal	-6513.9874	-615.4144	-1874.2516	-7129.4424	-8388.2774
Elec+Enthalpy	-6513.9865	-615.4134	-1874.2507	-7129.4414	-8388.2764
Elec+Free energy	-6514.1838	-615.4616	-1874.3078	-7129.6684	-8388.5122
Eelectronic	-6515.3305	-615.7330	-1874.4927	-7131.1082	-8389.8658
Entropy(cal/mol*K^{-1})	415.161	101.332	120.345	477.752	496.224

Table 3.50. Association energies for CD-AD and CD-FC from the MP2 corrections to the RHF results

	CD-AD	CD-FC
ΔE_{ele}(kj/mol)	-117.61	-112.04
ΔU^0(kj/mol)	-106.50	-100.63
ΔH^0(kj/mol)	-108.98	-103.11
ΔS^0(kj/mol*K^{-1})	-0.162	-0.164
ΔG^0(kj/mol)	-60.64	-54.10
K_{eq}	4.26E+10	3.04E+9

B. MP2 correction to the B3LYP results

Table 3.51. MP2/6-31G energies for the B3LYP optimized CD, AD and FC

(Hartrees/particle)	CD	AD	FC
Zero point	1.1677	0.2884	0.2134
Internal	1.2493	0.2996	0.2266
Enthalpy	1.2503	0.3005	0.2275
Free energy	1.0444	0.2509	0.1711
Elec+Zero point	-6514.2389	-615.4534	-1874.3353
Elec+Internal	-6514.1573	-615.4423	-1874.3221
Elec+Enthalpy	-6514.1564	-615.4413	-1874.3212
Elec+Free energy	-6514.3622	-615.4909	-1874.3776
Eelectronic	-6515.4066	-615.7418	-1874.5487
Entropy(cal/mol*K^{-1})	433.274	104.438	118.777

Table 3.52. MP2/6-31G energies for B3LYP optimized CD-AD and CD-FC

(Hartrees/particle)	CD-AD (V)	CD-AD (VI)	CD-FC (V)	CD-FC (VI)
Zero point	1.4592	1.4584	1.3840	1.3845
Internal	1.5528	1.5524	1.4800	1.4800
Enthalpy	1.5537	1.5533	1.4810	1.4810
Free energy	1.3199	1.3185	1.2394	1.2406
Elec+Zero point	-7129.7366	-7129.7154	-8388.6161	-8388.6014
Elec+Internal	-7129.6430	-7129.6215	-8388.5201	-8388.5058
Elec+Enthalpy	-7129.6421	-7129.6205	-8388.5192	-8388.5049
Elec+Free energy	-7129.8760	-7129.8553	-8388.7607	-8388.7452
Eelectronic	-7131.1958	-7131.1738	-8390.0001	-8389.9859
Entropy(cal/mol*K^{-1})	492.694	494.227	508.448	505.845

Table 3.53. Association energies for CD-AD and CD-FC from the MP2 corrections to the B3LYP
results

	CD-AD V	CD-AD VI	CD-FC V	CD-FC VI
ΔE_{ele}(kj/mol)	-124.36	-66.64	-117.67	-80.22
ΔU^0(kj/mol)	-114.15	-57.48	-106.76	-69.32
ΔH^0(kj/mol)	-116.63	-59.96	-109.24	-71.81
ΔS^0(kj/mol*K^{-1})	-0.190	-0.182	-0.182	-0.193
ΔG^0(kj/mol)	-59.87	-5.70	-54.84	-14.16
K_{eq}	3.08E+10	9.98	4.05E+9	302.0

3.2.3.3. Solvation energy

The following diagram represents the cycle we use to determine the solvation effects for the reaction that forms a binary complex, $A + B \rightarrow$ A-B.

$$\Delta_r G^0(g)$$

$$A(g) + B(g) \rightleftharpoons \text{A-B}(g)$$

$$-\Delta G^{sol} \quad -\Delta G^{sol} \qquad \qquad \Delta G^{sol}$$

$$\Delta_r G^0(aq)$$

$$A(aq) + B(aq) \rightleftharpoons \text{A-B}(aq)$$
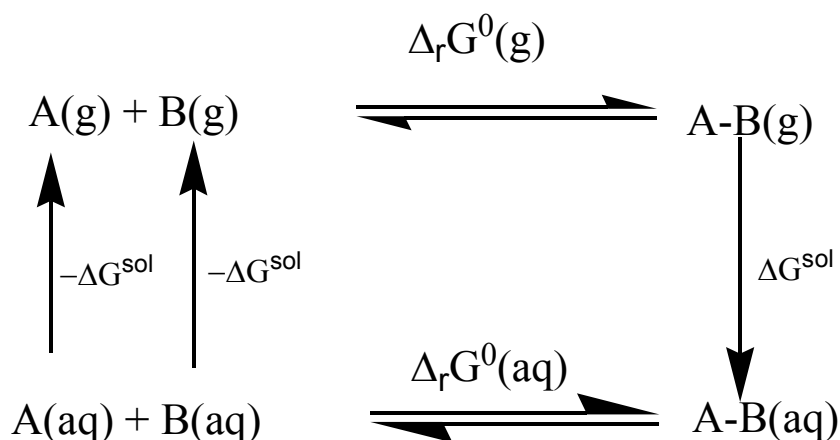
From the above diagram, we get:

$\Delta_r G^0$(CD-AD)(aq) = -ΔG^{sol}(CD) + [-ΔG^{sol}(AD)] + $\Delta_r G^0$(CD-AD)(g) + ΔG^{sol}(CD-AD) (3.10)

$\Delta_r G^0$(CD-FC)(aq) = -ΔG^{sol}(CD) + [-ΔG^{sol}(FC)] + $\Delta_r G^0$(CD-FC)(g) + ΔG^{sol}(CD-FC) (3.11)

In the above, $\Delta_r G^0$(CD-AD)(aq) and $\Delta_r G^0$(CD-FC)(aq) are the Gibbs free energies of association for CD-AD and CD-FC respectively in water solution. In turn, ΔG^{sol}(CD), ΔG^{sol}(AD), ΔG^{sol}(FC), ΔG^{sol}(CD-AD) and ΔG^{sol}(CD-FC) are the solvation Gibbs free energies for CD, AD, FC, CD-AD and CD-FC. $\Delta_\mathbf{r} G^0$(CD-AD)(g) and $\Delta_\mathbf{r} G^0$(CD-FC)(g) are the Gibbs free energies of association for CD-AD and CD-FC in the gas phase.

As with the MP2 calculations, the solvent phase calculations have only been carried out on the first RHF optimized complexes and the fifth and sixth B3LYP optimized complexes because those ΔG values were the lowest. The model used is the PCM/SCRF

method. Only single point calculations were applied. The following results were obtained.

Table 3.54. The RHF/6-31G solvation energies for the 1st fully optimized set

Species	CD	AD	FC	CD-AD	CD-FC
ΔG^{sol}(kal/mol)	-58.81	-8.260	-14.86	-33.77	-40.89
ΔG^{sol}(kj/mol)	-246.1	-34.56	-62.17	-141.3	-171.1

Table 3.55. The RHF/6-31G solvent phase association Gibbs free energies and equilibrium constants of CD-AD and CD-FC

	CD-AD	CD-FC
ΔG^0(aq) (kj/mol)	107.2	111.4
Keq(aq)	1.691E-19	2.997E-20

Table 3.56. B3LYP/6-31G solvation energies for CD-AD

Species	CD	AD	FC	CD-AD(V)	CD-AD(VI)
ΔG^{sol} (kal/mol)	-38.25	-6.660	-11.89	-22.62	-32.31
ΔG^{sol} (kj/mol)	-160.0	-27.86	-49.75	-94.64	-135.2

Table 3.57. B3LYP/6-31G solvation energies for CD-FC

Species	CD	AD	FC	CD-FC(V)	CD-FC(VI)
ΔG^{sol} (kal/mol)	-38.25	-6.66	-11.89	-25.42	-35.37
ΔG^{sol} (kj/mol)	-160.0	-27.86	-49.75	-106.4	-148.0

Table 3.58. B3LYP/6-31G solvent phase association Gibbs free energies and equilibrium constants for CD-AD and CD-FC

	CD-AD V	CD-AD VI	CD-FC V	CD-FC VI
ΔG^0(aq) (kj/mol)	45.65	38.64	62.26	39.61
Keq(aq)	1.005E-8	1.704E-7	1.236E-11	1.152E-7

So, the best results in solvent phase are 38.64 kj/mol for CD-AD and 39.61 kj/mol for CD-FC complexes respectively. Interestingly, for the complex of CD-FC, the best result is from the sixth complex, different from the gas phase results, in which the fifth one has the best association Gibbs free energy. Furthermore, those solvent phase results are in

very poor agreement with experimental data in Table 3.3. We believe the reason behind this is that we have completely ignored the hydrophobic (hydrophilic) contribution to the association process. See discussion part for detail.

3.3. The excited state calculations

3.3.1. Method

In order to study the electron absorption and fluorescence behavior of these metal nano-particles anchored to individual cyclodextrin molecules and the complexes of CD-AD and CD-FC, the ZINDO method was applied to study the excited states. The nano-particles of the metal (M) are linked to the CD component through one or more sulfur linkages, -M-S-C- where the carbon atom resides in one of the CD rings. This is modeled here by inserting one metal atom into the thiol H-S- bond to give a   H-M-S-C- motif. For these calculations, Zn is used as the metal since the ZINDO method is well parameterized for this transition metal element. These structures of Zn-CD, Zn-(CD-AD) and Zn-(CD-FC) are presented in Figure 5.62, 5.63 and 5.64 respectively in the Appendix.

These three structures were then optimized using the semiempirical PM3 method before a ZINDO single point calculation was completed on each of the three optimized structures. The ZINDO method calculates for the excited states the transition energies, the molecular orbitals involved in each transition and the major contributing atomic orbitals making up each molecular orbital through a Mulliken Population analysis.

Our goal is to decipher what kind of change is brought to the first excitation transition (HOMO to LUMO) in the nano-particle-CD system after AD and FC are complexed to the CD moiety. First we determine the first excitation transition in the metal-CD unit, and then locate the corresponding excitation in the metal-CD-AD and the metal-CD-FC

complexes by comparing the excitation energy, the energy level of the associated molecular orbitals and the major atomic orbitals making up these molecular orbitals.

We obtained the following results. Only those atomic orbitals with a MO coefficient greater than 0.1, are listed.

3.3.2.   The ZINDO Results

3.3.2.1. Results for a nano-particle with only CD

The first excitation transition occurs as an electron is excited from orbital 225 (HOMO) to orbital 226 (LUMO). The excitation energy is 7.0151 eV (about 176.74 nm) with an oscillator strength f equal to 0.0103. The major components for orbital 225 and 226 are listed in Table 3.59 and 3.60 respectively.

Table 3.59. The atomic orbital composition for molecular orbital 225 (Eigenvalue = -0.36480 eV)

O #13	C #72	S #73			Zn #74	H #124	H #125
$2P_x$	$2P_y$	$3P_x$	$3P_y$	$3P_z$	$3P_y$	1S	1S
0.12	-0.17	-0.18	0.89	0.22	0.14	-0.17	0.12

Table 3.60. The atomic orbital composition for molecular orbital 226 (Eigenvlue = 0.06449 eV)

C #11	C#12	C#72	S#73		Zn #74			H#121	H#122	H#148
2S	2S	2S	3S	$3P_x$	4S	$3P_x$	$3P_z$	1S	1S	1S
0.21	-0.30	0.26	-0.14	0.18	0.41	0.23	0.58	-0.10	0.12	-0.25

The first orbital (225) is mainly centered on the sulfur atom (#73) and the second orbital (226) is very much concentrated on the Zn atom (#74), which is bonded to the sulfur atom. So for particles with CD-AD and CD-FC, the corresponding excitation has to involve two molecular orbitals mostly centered on the bonded S and Zn atoms.

3.3.2.2.   Results for a nano-particle with the CD-AD complex

The corresponding excitation in the nano-particle with CD-AD complex is the one from MO 264 to MO 266 at an excitation energy of 6.9410 eV (178.63 nm) with an oscillator strength f equal to 0.0074. Orbital 264 is mainly localized on the sulfur atom (#85) and MO 266 mostly involves Zn (#86) which is bonded to the sulfur atom. The major components for orbital 264 and 266 are listed in Table 3.61 and 3.62 respectively.

Table 3.61. The atomic composition for molecular orbital 264 (Eigenvalue = -0.36361 eV)

C#84		S#85			S#88	H#139
$2P_x$	$2P_y$	$3P_x$	$3P_y$	$3P_z$	$3P_x$	1S
0.10	-0.12	-0.56	0.59	-0.45	0.10	0.16

Table 3.62. The atomic composition for molecular orbital 266 (Eigenvalue = 0.06338 eV)

S#85			Zn #86		
3S	$3P_x$	$3P_y$	4S	$3P_x$	$3P_z$
0.15	0.16	0.12	-0.41	0.37	0.46

Table 3.62 (continued)

H#135	H#136	H#180	C #32	C#33	C#34	C#35	C#84
1S	1S	1S	1S	1S	1S	1S	1S
0.11	-0.11	-0.25	-0.11	0.13	-0.22	0.28	-0.27

3.3.2.3.   Results for a nano-particle with the CD-FC complex

The excitation of the nano-particle with the CD-FC complex of that corresponds to the first excitation for the nano-particle with CD only is the one from MO 259 to orbital MO 270. The excitation energy is 6.9164 eV (179.26 nm) with an oscillator strength of 0.0052. And orbital 259 is mainly centered on the sulfur atom (#78) and MO 270 mainly involves the Zn atom (#132), which is bonded to the sulfur atom. Table 3.63 and 3.64 lists the major components for orbital 259 and 270 respectively.

Table 3.63. The atomic orbital composition for molecular orbital 259 (Eigenvalue = -0.35871 eV).

C#77	S#78		H#130	H#131	Zn#132
$2P_x$	$3P_x$	$3P_y$	$1S$	$1S$	$3P_x$
-0.15	0.80	-0.48	-0.16	0.13	0.12

Table 3.64. The atomic orbital composition for molecular orbital 270 (Eigenvalue = 0.06286 eV)

S#78		Zn#132				C#166		C#170	
$3S$	$3P_y$	$4S$	$3P_x$	$3P_y$	$3P_z$	$2P_x$	$2P_y$	$2P_x$	$2P_y$
0.12	-0.15	-0.37	-0.11	-0.12	0.51	0.17	-0.12	-0.13	0.11

Table 3.64 (continued)

C#7	C#8	C#9	C#10	C#11	C#77	H#129	H#148	C#167
$2S$	$2S$	$3S$	$2S$	$2S$	$2S$	$1S$	$1S$	$2P_x$
0.28	-0.24	0.21	-0.16	0.13	-0.21	-0.12	0.22	-0.10

Figures 3.11 through 3.16 give the density plots for the oribtals discussed above. The MO illustrated in Figures 3.11, 3.13 and 3.15 are very similar, and those illustrated in Figures 3.12, 3.14 and 3.16 are also very similar. These similarities in density, energy level and orbital composition assure us that these orbitals are related to one another, as are the three corresponding transitions.
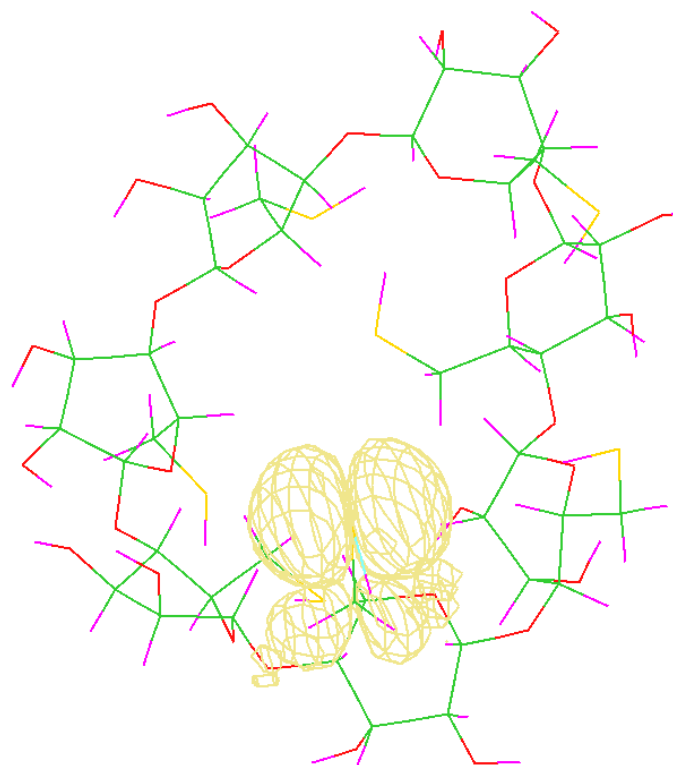
Figure 3.11. The HOMO (225) of CD-ZnS
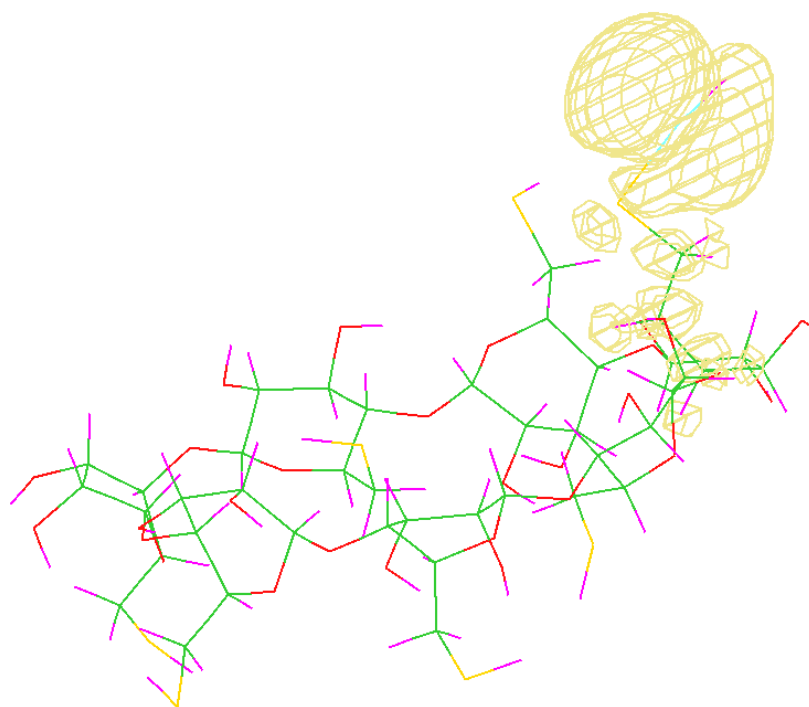
Figure 3.12. LUMO (226) for CD-ZnS
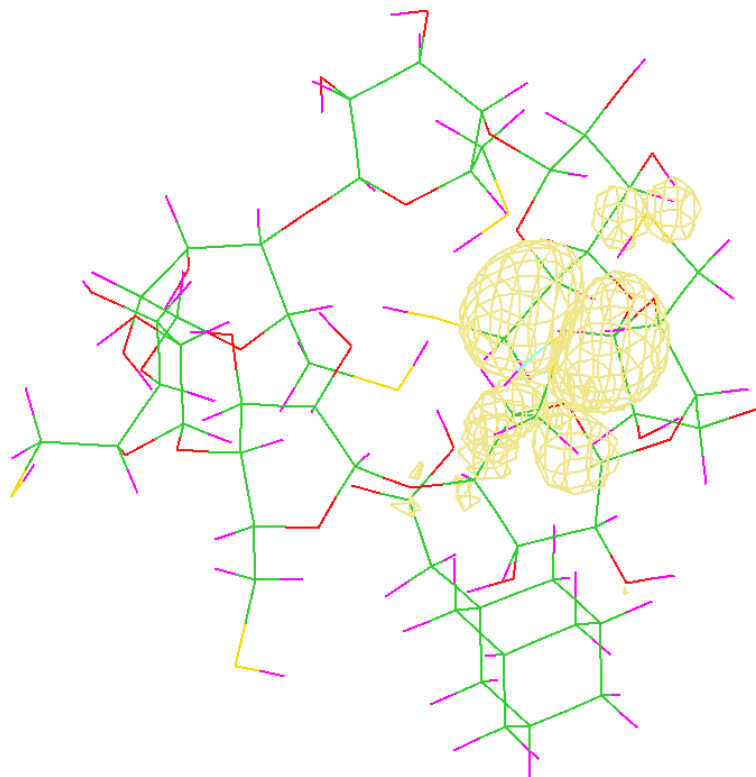
Figure 3.13. Orbital 264 for CD-AD-ZnS.
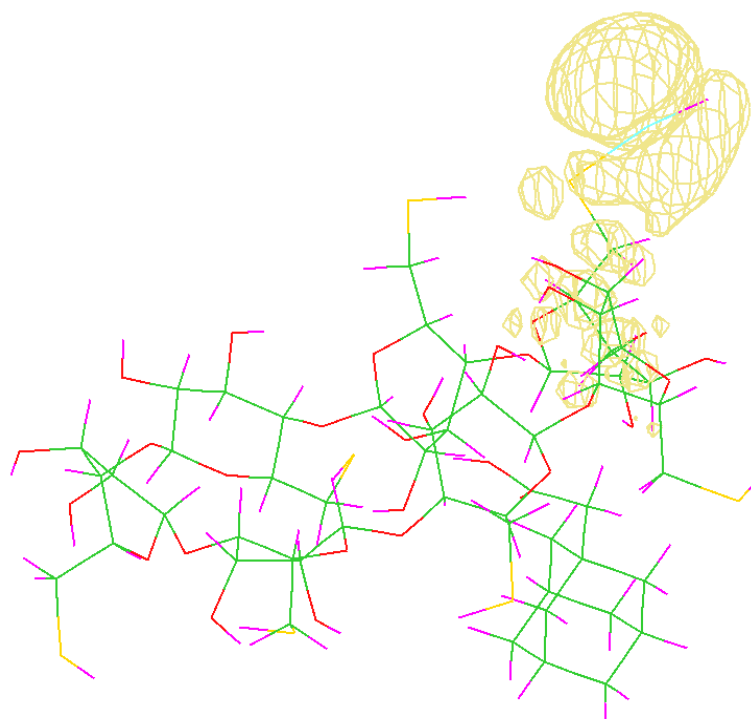
Figure 3.14. Orbital 266 for CD-AD-ZnS
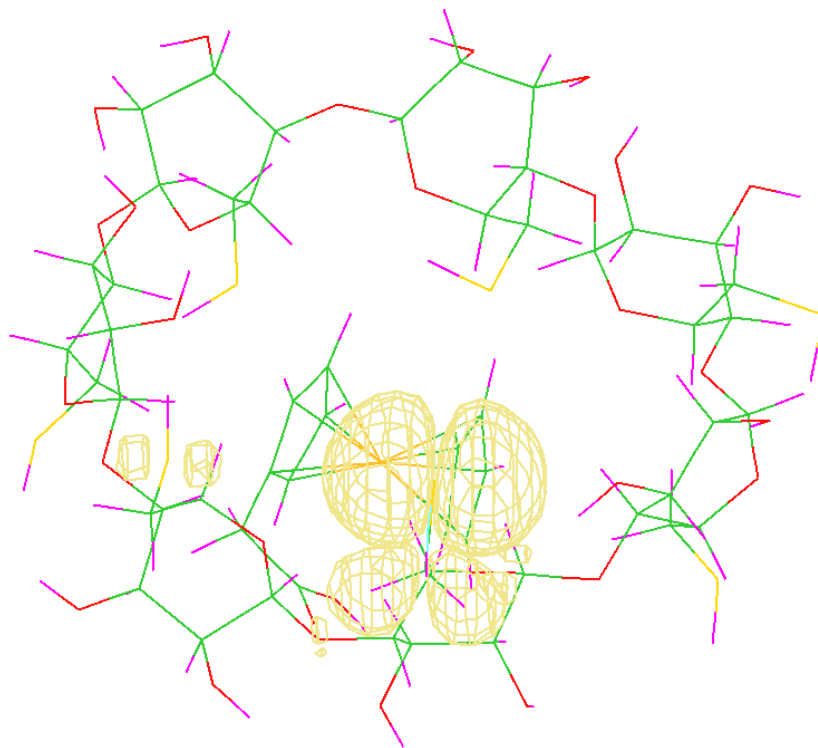
Figure 3.15. Orbital 259 of CD-FC-ZnS.
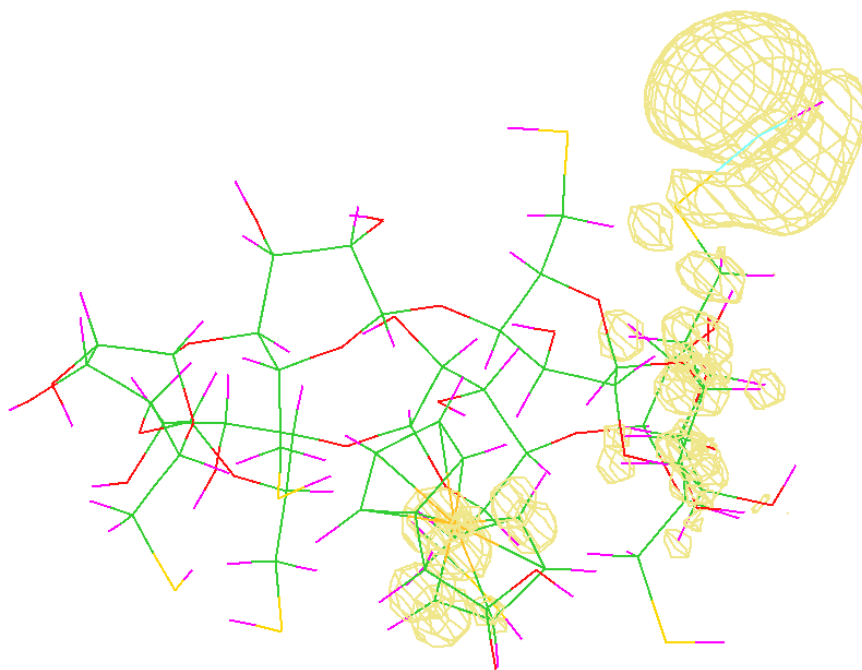
Figure 3.16. Orbital 270 of CD-FC-ZnS

4. Discussion

4.1. The structure and energy levels of CD

The CD structure with the lowest computed energy from the RHF/6-31G(d) optimizations and confirmed by B3LYP/6-31G(d) single point calculations is the 9th conformer. It is illustrated below.
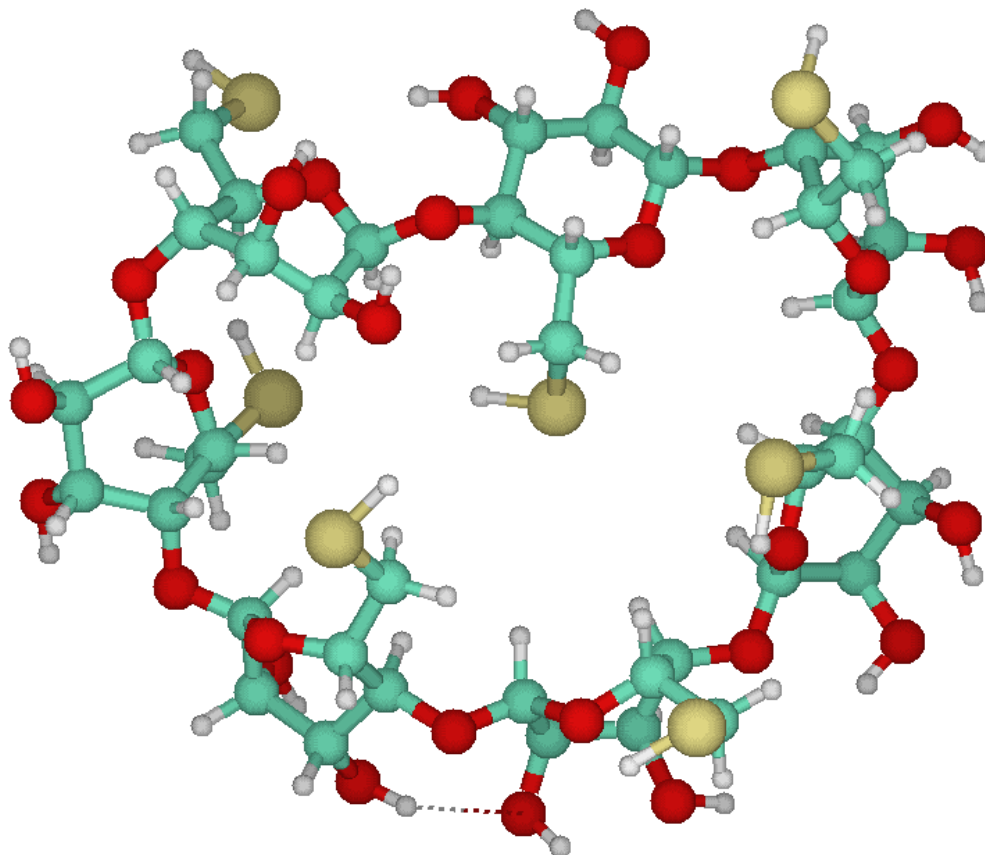
Figure 4.1. Structure of the ninth CD conformer

In Table 4.1, the ten conformers with the lowest energy from RHF/6-31G(d) optimizations are listed. The B3LYP/6-31G(d) method was employed to calculate the energy on the RHF-optimized geometries of the four conformers with lowest energy, and these results are listed in Table 4.1b. These calculations have revealed that the ninth conformer is the most stable one, and it will be used for all inclusion complexes.

A qualitative analysis of the geometries leads us to believe that the repulsive interactions among the sulfur atoms plays an important role on the relative stability of the conformers, perhaps even more than the intra-molecular hydrogen bonds. The 2nd, 3rd and 4th conformers have two intramolecular hydrogen bonds formed at their optimized geometries and the rest have only one (see Figures 5.12-5.21 in the Appendix). But those with two hydrogen bonds are not those with the lowest energy; instead the 2nd and 3rd are the two with highest energy. In Table 4.2 we list these distances between all seven sulfur atoms (21 pairs) in these ten conformers.

Table 4.1. Energy level of the first ten conformers at the RHF/6-31G(d) level

Conformer	Energy (Hartree/molecule)
1	-6509.6941
2	-6509.6894
3	-6509.6875
4	-6509.6947
5	-6509.6958
6	-6509.7025
7	-6509.6932
8	-6509.6984
9	-6509.7042
10	-6509.6910

(Note: the order from the lowest to highest is 9, 6, 8, 5, 4, 1, 7, 10, 2, 3.)

Table 4.1b. B3LYP/6-31G(d) energies for the four conformers with lowest RHF/6-31G(d) energy

Conformer	9th	6th	8th	5th
Energy (Hartree)	-6535.8694	-6535.8679	-6535.8672	-6535.8602

Actually based on our qualitative analysis we have concluded that not all sulfur-sulfur interactions have a significant effect on the energy. The further the sulfur pair is apart, the

weaker the interaction would be. It seems that only interactions from those pairs within 5 Å are significant. For instance the sum of all pair distances reaches the maximum for the 3rd conformer, which is 185.136 Å, but the 3rd conformer has the highest energy. However, both the 3rd and 2nd conformers have four pair distances shorter than 5 Å, the highest among all ten conformers. So the 3rd and 2nd conformers are the two with the highest energy. On the other hand, the 9th and 6th conformers have only 1, the lowest number of pair distance shorter than 5 Å, which makes them the two with the lowest energy. But this is only a qualitative analysis. The 7th conformer is an exception. It only has two pairs of sulfur atoms falling within 5 Å, but it has an energy level higher than conformers 4 and 1, both of which have 3 pairs of sulfur atoms with distances less than 5 Å. This can be explained partially by counting the number of distances shorter than 6 Å but higher than 5 Å. For that, the 7th conformer has 4 pairs but both the 1st and the 4th have only 3 pairs. And also it may be explained by the strength of the intramolecular hydrogen bonds. The 4th one has two intramolecular hydrogen bonds. Both the 1st and the 7th only have one, but the one from the 1st conformer has a shorter O-H distance and a more linear O-H-O angle (closer 180 degrees) than those in the hydrogen bond for 7th conformer (see Table 4.3).

But a more precise quantitative analysis is needed to reveal the subtle relation between the structure and energy. For example, the 9th and 6th conformers both have only one sulfur atom pair with a distance shorter than 5Å and comparable hydrogen bonds. So why is the 9th one more stable than the 6th one energetically?

In Table 4.4 are listed the nonbonded interaction energies from MM calculations for the ten conformers. They are ordered from the lowest to the highest.

From the data in Table 4.4, the nonbonded interactions (sometimes it is termed the van der Waals interactions including the nuclear repulsion) play an important role in determining the relative energy levels of the conformers, especially for making the 9th and the 6th the two most stable ones. But for the rest of the conformers, other effects such as the electronic contributions including the electron-electron repulsion and nuclear-electron attraction surpass the inter-nuclear interactions. For example, the 2nd

93

and the 3rd conformers both have double intramolecular hydrogen bonds and very low nonbonded energy terms (ranked the third and the fourth), but after the electronic energy is taken into account, they become the two with the highest energy levels.

Table 4.2. Distance between the sulfur atoms in the first ten conformers (Angstroms)

1	2	3	4	5	6	7	8	9	10
8.45	6.40	8.44	5.80	7.32	6.67	6.29	6.22	6.93	6.36
4.64	4.62	4.65	4.37	5.12	4.62	4.78	4.41	4.58	4.67
4.62	6.67	6.79	5.27	6.27	6.19	4.79	4.58	6.26	4.84
7.84	7.31	7.27	6.54	7.56	7.62	7.67	7.57	7.63	7.75
4.68	4.59	4.58	4.80	4.68	5.91	5.20	5.68	5.61	5.02
6.98	6.89	7.08	5.68	7.03	7.55	6.90	7.94	7.56	6.64
9.92	10.73	10.75	9.83	9.92	10.84	10.20	10.82	10.87	10.16
8.98	11.24	11.28	9.50	11.33	10.72	9.53	8.90	10.78	9.46
7.81	6.64	7.33	8.13	7.58	6.99	7.25	6.57	6.94	6.97
5.65	4.60	4.67	5.22	5.47	5.34	5.37	5.52	5.20	5.12
5.02	4.46	4.45	4.70	4.76	5.19	5.26	7.01	5.12	6.06
7.10	8.10	7.94	7.58	7.78	7.55	7.92	7.44	6.96	6.83
12.32	12.53	13.41	12.58	12.44	11.50	12.11	11.47	11.22	10.92
11.24	12.67	12.74	12.39	12.23	11.21	11.23	11.84	10.74	9.87
13.30	14.64	14.65	15.51	13.80	12.60	12.93	13.63	12.28	11.91
5.68	8.31	8.35	10.31	6.94	6.00	5.88	6.43	5.65	4.79
9.03	8.39	8.22	7.20	8.59	9.78	10.40	9.29	9.32	9.78
14.67	12.69	13.32	12.04	13.62	14.22	14.85	13.86	14.17	14.68
12.02	11.20	11.30	10.27	11.50	12.31	12.34	12.55	12.14	12.02
12.26	11.24	11.26	12.04	10.84	11.02	12.04	12.14	10.95	11.83
6.87	6.65	6.68	7.85	6.25	7.64	7.78	8.32	7.58	8.18

Table 4.3. The O-H distance and O-H-O angle in the intramolecular bonds in the RHF/6-31G(d) optimized conformers (Note: in the second and third rows, the distance and angle are listed for the first and the second hydrogen bond (if it exists).

1	2	3	4	5	6	7	8	9	10
2.00Å	2.122Å	2.00 Å	1.98 Å	2.00Å	1.97Å	2.122Å	2.00 Å	1.99 Å	2.14 Å
160.4^0	158.9^0	162.8^0	153.8^0	165.7^0	150.5^0	153.9^0	152.9^0	150.7^0	152.7^0
NA	2.18Å	2.17Å	2.07Å	NA	NA	NA	NA	NA	NA
	146.6^0	146.6^0	145.6^0						

Table 4.4. MM Nonbonded energy terms

Conformer	Nonbonded energy (Hartrees)
9th	0.3506
6th	0.3593
2nd	0.3679
3rd	0.3679
8th	0.3681
5th	0.3691
10th	0.3704
1st	0.3752
7th	0.3800
4th	0.3878

4.2. The structures and association energies of the complexes

The most stable structures for the CD-AD and CD-FC complexes in the gas phase are illustrated below for the RHF/6-31G and B3LYP/6-31G levels of theory respectively.
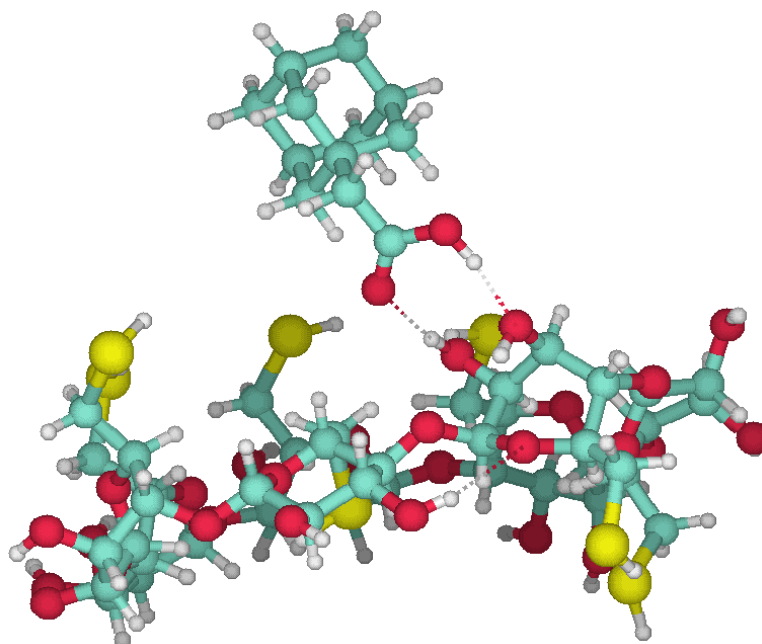
Figure 4.2. The most stable RHF/6-31G structure of CD-AD
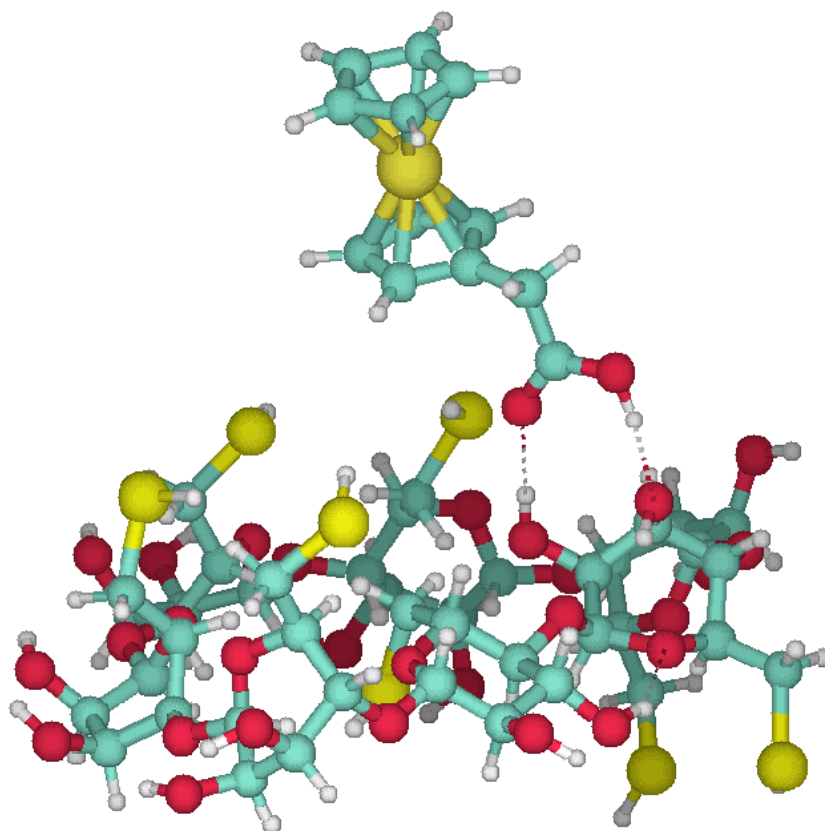
Figure 4.3. The most stable RHF/6-31G structure of CD-FC
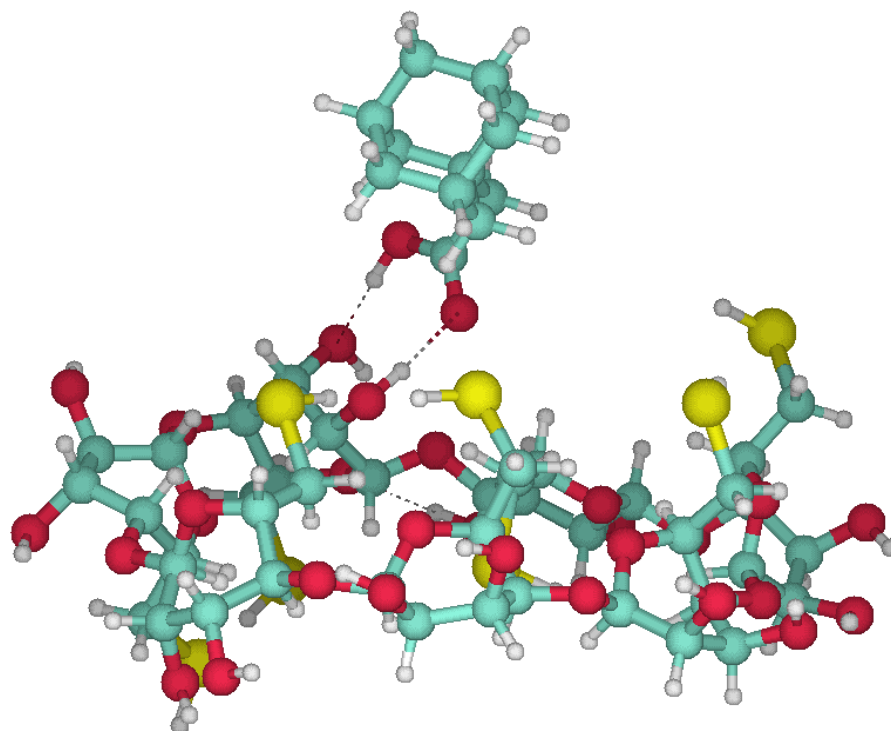
Figure 4.4. The most stable B3LYP/6-31G structure of CD-AD
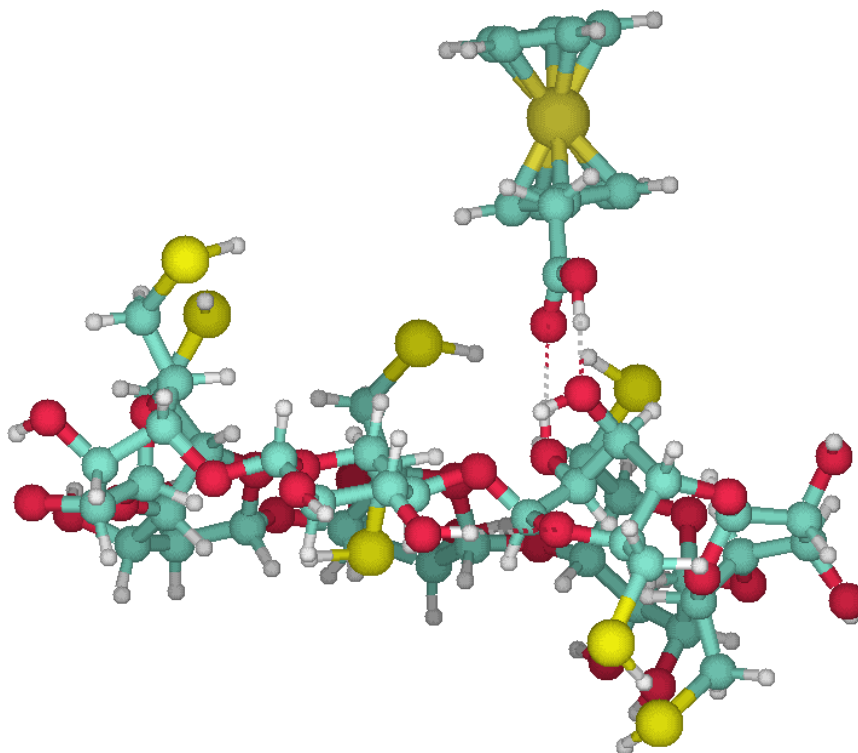
Figure 4.5. The most stable B3LYP/6-31G structure of CD-FC

A noticeable feature of the above structures is the double intermolecular hydrogen bonds.
In Tables 4.3 and 4.4 are listed the O-H length and O-H-O angle of the double hydrogen
bonds in CD-AD and CD-FC complexes at the RHF and B3LYP levels respectively.

Table 4.5. The RHF/6-31G O-H length and O-H-O angle of the double hydrogen bonds in CD-AD and CD-FC

	CD-AD		CD-FC	
	O-H	O-H-O	O-H	O-H-O
Bond #1	1.73769Å	171.995^0	1.70188 Å	170.146^0
Bond #2	1.86231 Å	163.278^0	1.82117 Å	169.843^0

Table 4.6. The B3LYP/6-31G O-H length and O-H-O angle of the double hydrogen bonds in CD-AD and CD-FC

	CD-AD		CD-FC	
	O-H	O-H-O	O-H	O-H-O
Bond #1	1.61070Å	175.687^0	1.56055Å	175.705^0
Bond #2	1.69535Å	164.542^0	1.74543Å	168.263^0

The data in the above tables indicates strong hydrogen bond interactions. The lengths are quite short (1.56 – 1.86 Å) and the O-H-O angles are quite linear (163.3° – 175.7°). The gas phase association Gibbs free energy (ΔG), enthalpy (ΔH), entropy (ΔS), and electronic energy ($\Delta E_{electron}$) for the most stable CD-AD and CD-FC complexes are listed in Tables 4.7 and 4.8

Table 4.7. The RHF/6-31G association energies for the most stable complexes

kj/mol	ΔG^0	ΔH^0	ΔS^0	$\Delta E_{electron}$
CD-AD	-32.205	-80.542	-0.162	-89.175
CD-FC	-25.730	-74.741	-0.164	-83.673

Table 4.8. The B3LYP/6-31G association energies for the most stable complexes

kj/mol	ΔG	ΔH	ΔS	$\Delta E_{electron}$
CD-AD	-47.567	-104.334	-0.190	-112.066
CD-FC	-41.086	-95.487	-0.182	-103.923

The solvation Gibbs free energy for all species were calculated with the PCM model at both the RHF/6-31G and B3LYP/6-31G levels. These gas phase optimized geometries were used for the solvent phase calculations. For calculations with RHF methods, only CD-AD, CD-FC complexes with the most stable gas structures were used in solvent phase calculations; but for B3LYP methods, the second most stable structures from gas phase were also tried in solvent phase along with the most stable ones.

Table 4.9. The RHF/6-31G solvation Gibbs free energies

Species	CD	AD	FC	CD-AD	CD-FC
ΔG^{sol} (kj/mol)	-58.81	-8.26	-14.86	-33.77	-40.89

Table 4.10. The B3LYP/6-31G solvation Gibbs free energies

Species	CD	AD	FC	CD-AD	CD-FC
ΔG^{sol} (kj/mol)	-38.25	-6.66	-11.89	-22.62	-25.42

From equations 3.10 and 3.11, we calculated the solvent phase association Gibbs free energy [ΔG^0(aq)] for CD-AD and CD-FC as follows (also see tables 3.55 and 3.58):

Table 4.11a. Solvent phase association Gibbs free energy ΔG^0(aq) for the CD-AD and CD-FC complexes

	CD-AD	CD-FC
ΔG^0(aq)(RHF/6-31G)(kj/mol)	107.2	111.4
ΔG^0(aq)(B3LYP/6-31G)(kj/mol)	38.64	39.61

Table 4.11b. Solvent phase association equilibrium constant for the CD-AD and CD-FC complexes

	CD-AD	CD-FC
K_{eq}(aq)(RHF/6-31G)	1.691E-19	2.997E-20
K_{eq}(aq)(B3LYP/6-31G)	1.704E-7	1.152E-7

In the gas phase, the most stable structures of the CD-AD and CD-FC complexes are not those having the guest molecule deep inside the cavity of the host molecule. As we discussed in Chapter III, the magnitude of the parameter, dih148, among the six controlling parameters, indicates how close the guest molecule is to the mean plane of the CD molecule (the sign indicates which side of CD the guest is on). The smaller the magnitude is, the closer the guest is to CD. As shown in Table 3.37 and 3.38, for the RHF method, we have dih148 equal to -24.12 for the most stable CD-AD complex and -37.38 for the most stable CD-FC. These two values are not the smallest in magnitude compared to other available ones. We have one with dih148 equal -15.36 for CD-AD and 36.39 for CD-FC. For the B3LYP method, the values of dih148 for the most stable complexes are 45.61 for CD-AD and 35.09 for CD-FC according to Tables 3.46 and 3.47. Other conformations have smaller values for this angle. For example at the B3LYP/6-31G level, we have -22.00 and 37.05 for CD-AD and -31.40 and -10.84 for CD-FC, all of which have smaller magnitude than those of the two most stable complexes. For both the RHF and B3LYP models, all of the most stable complexes for CD-AD and CD-FC have the double intermolecular hydrogen bonds present. So it seems that in the gas phase the hydrogen bond interaction is the main driving force for inclusion.

From the data in Tables 4.11a and 4.11b, it is clear that the B3LYP method provides better results in terms of giving the lowest ΔG^0(aq) values which are also closer to those experimental measurements shown in Table 3.5. If we accept the B3LYP solvent phase results as the most reliable. According to equation 3.7 the ratio of the CD-AD complex to the CD-FC complex in the solution can be estimated as follows:

$$\frac{[CD-AD]}{[CD-FC]}=\frac{K_{eq}^{CD-AD}(aq)}{K_{eq}^{CD-FC}(aq)}=\frac{1.704E-7}{1.152E-7}=1.479 \quad (4.1)$$

According to the result from equation 5.1, given that we have equal amounts of free AD and FC in the solution mixture at equilibrium, we have about 40% of the complex as CD-FC and 60% as CD-AD. Of course here we assume we have excess AD and FC in solution for complexing with CD. This results indicates that after enough AD is added to the solution composed of the CD-FC complex and excess FC, we have most of the CD-FC complex convert to the CD-AD complex. This explains why most of the fluorescence emission was recovered by adding AD to a solution mixture of CD and FC (With a large K_{eq} and FC in excess, most of the CD is complexed with FC before the addition of AD).[1] The electronic contribution plays a critical role in stabilizing the complexes. In Table 4.12 through 4.15, are listed the energy changes without the electronic contribution. For comparison we still keep the electronic energy changes in these tables.

Table4.12. The RHF/6-31G energy changes for CD-AD without electronic contribution

CD-AD	I	II	III	IV
ΔE_{ele}(kj/mol)	-89.17	-52.94	-36.34	-60.89
ΔU(kj/mol)	11.11	10.27	10.64	10.71
ΔH(kj/mol)	8.63	7.79	8.16	8.23
ΔS(kj/mol)	-0.162	-0.174	-0.173	-0.180
ΔG^0(kj/mol)	56.97	59.78	59.79	61.84

Table 4.13. The RHF/6-31G energy changes for CD-FC without electronic contribution

CD-FC	I	II
ΔE_{ele}(kj/mol)	-83.67	-45.77
ΔU(kj/mol)	11.41	10.26
ΔH(kj/mol)	8.93	7.78
ΔS(kj/mol)	-0.164	-0.194
ΔG^0(kj/mol)	57.94	65.52

Table 4.14. The B3LYP/6-31G energy changes for CD-AD without electronic contribution

CD-AD	I	II	III	IV	V	VI
ΔE_{ele}(kj/mol)	-70.29	-51.31	-70.60	-47.15	-112.07	-74.96
ΔU(kj/mol)	8.73	9.27	9.56	10.01	10.21	9.16
ΔH(kj/mol)	6.25	6.79	7.08	7.53	7.73	6.68
ΔS(kj/mol)	-0.188	-0.194	-0.193	-0.184	-0.190	-0.182
ΔG^0(kj/mol)	62.42	64.52	64.74	62.39	64.50	60.94

Table 4.15. The B3LYP/6-31G energy changes for CD-FC without electronic contribution

CD-FC	I	II	III	IV	V	VI
ΔE_{ele}(kj/mol)	-69.66	-69.46	-39.47	-89.76	-103.92	-88.21
ΔU(kj/mol)	10.63	10.29	7.85	8.09	10.91	10.90
ΔH(kj/mol)	8.15	7.81	5.37	5.61	8.43	8.41
ΔS(kj/mol)	-0.208	-0.208	-0.216	-0.231	-0.182	-0.193
ΔG^0(kj/mol)	70.20	69.76	69.86	74.48	62.83	66.06

The entropy change, ΔS, for all the complexes always has a negative value, which is what we expect since we have two individual molecules combine into one single molecule. And this is an unfavorable factor thermodynamically. All the enthalpy changes, ΔH, are positive when the electronic component is omitted which also thermodynamically inhibits the complexing process. And the combined the effect of a negative ΔS and positive ΔH leads to a positive Gibbs free energy change, ΔG, with a quite significant magnitude around 60~70 kj/mol when the electronic contributions are omitted.

The ΔS values are not sensitive to the complex structure and the computational method. It remains relatively stable around -0.2 kj/mol. But the electronic energy change greatly depends on the complex structure and the computational method. We have observed as much as 64.92 kj/mol electronic energy difference for the CD-AD complex at B3LYP/6-31G level arising from different conformations (the fifth vs the fourth). A similar difference, 64.45 kj/mol, has been obtained for the CD-FC complex (the fifth vs the third) at the B3LYP/6-31G level for different conformations. At the RHF level, the largest electronic energy changes were 52.83 kj/mol for CD-AD and 37.90 kj/mol for CD-FC. .

In order to assess the dependence of the electronic energy change on the method of computation, we need to compare the results from different methods on a similar structure for the complex. So it is only meaningful to compare the results from the first RHF and the fifth B3LYP calculations, since the final geometry from these calculations are very similar. Both methods started the optimization from the structure partially optimized at the RHF/6-31G level, and very little geometrical changes were observed during these optimizations. The MP2 calculations help to give more reliable estimates of the correlation effect.

The differences in the electronic energy changes from the RHF and B3LYP methods are basically due to the correlation effect. In the gas phase, the B3LYP method estimates the association electronic energy for CD-AD to be -112.066 kj/mol , which is about 22.891 kj/mol lower (more negative) than the one estimated from the RHF method, which is -89.175 kj/mol. Similar trends can be found for CD-FC with the results equal to -103.923 kj/mol and -83.673 kj/mol for the B3LYP and RHF methods respectively. The results from the MP2 method also exhibit the importance of the correlation effect. The MP2 electronic energy for the complexes from the first RHF optimization and the fifth B3LYP optimization are listed below.

Table 4.16. MP2 association electronic energies for CD-AD and CD-FC

	CD-AD	CD-FC
ΔE_{ele} (RHF/6-31G geometry) (kj/mol)	-117.610	-112.043
ΔE_{ele} (B3LYP/6-31G geometry)(kj/mol)	-124.363	-117.672

It is clear that for the same geometries, compared to results from the RHF method, the MP2 result is 28.435 kj/mol less for CD-AD and 28.370 kj/mol less for CD-FC. And the MP2 electronic energy is much closer to those from the B3LYP method: -124.363 kj/mol(MP2) vs -112.066 kj/mol (B3LYP) for CD-AD and -117.672 kj/mol (MP2) vs -103.923kj/mol (B3LYP) for CD-FC. It is clear that the electron correlation effect and hence the dispersion interactions play an important role on stabilizing these complexes. The association Gibbs free energy for the most stable CD-AD and CD-FC complexes in aqueous solution, are -25.28kj/mol for CD-AD and -16.37kj/mol for CD-FC as shown in

Table 3.62. These results were obtained from the computation with the PCM model at the B3LYP level and the complex geometries were taken from the fifth B3LYP optimization in the gas phase. Compared to the gas phase association Gibbs free energy, the solvent phase values are more positive (or less negative). The reason for that is that the sum of the solvation Gibbs free energies from the two monomers is always more negative than the one from the corresponding complex formed by them. As indicated in Table 3.58, 3.60 and 3.61, the combined solvation Gibbs free energy of CD and AD is always more negative than the solvation Gibbs free energy of the complex of CD-AD. And that is also true for CD, FC and the complex of CD-FC. There appears to be a loss of total polarity upon complexation.

Due to the implicit model for the solvent phase computations, there are no solvent molecules (water) present. So the so called "hydrophobic forces" are ignored. That is what we believe contribute to the poor solvent phase results compared to the experimental data. The "hydrophobic force" contribution includes guest molecules being released from the hydrated state to the free state, and guest molecules pushing out the water molecules residing in the cavity of the host molecule into the bulk water. A better model may be one, in which the CD molecules have several water molecules inside the cavity before complexation. For future study, we propose the following algorithm to re-calculate the association Gibbs free energy in solution.

$$Guest(g) + CD \bullet nH_2O(g) \xrightarrow{\Delta G(g)} CD \bullet Guest(g) + nH_2O(g)$$
$$\downarrow \Delta G_1 \qquad\qquad\qquad\qquad\qquad \uparrow \Delta G_2$$
$$Guest(l) + CD \bullet nH_2O(l) \xrightarrow{\Delta G(aq)} CD \bullet Guest(l) + nH_2O(l)$$
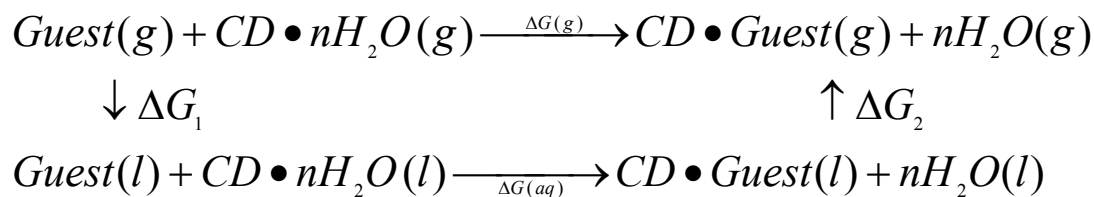
Figure 4.6. Cycle for calculating association Gibbs free energy

From the above cycle, the association Gibbs free energy in solution $\Delta G(aq)$ can be calculated according to the equation below.

$$\Delta G(aq) = \Delta G(g) - \Delta G_1 - \Delta G_2 \qquad (4.2)$$

ΔG_1 is the sum of the solvation Gibbs free energies for the Guest and CD·nH$_2$O. The way to compute ΔG_1 has already been discussed before. ΔG_2 can also be calculated in a similar way, but it is the Gibbs free energy change for a process opposite to solvation. So ΔG_2 is

the negative of the sum of the solvation Gibbs free energy for CD·Guest and nH_2O. ΔG_2 includes two parts: the solvation Gibbs free energy for CD·Guest and the Gibbs free energy change for turning liquid water to gaseous water at 298.15K and 1atm (ΔG^{vap}). The experimental data for ΔG^{vap} are available for a wide range of temperatures and pressures. We can also evaluate it in the way demonstrated below.

$$nH_2O(l) \xrightarrow{\Delta H^{vap},\Delta S^{vap},\Delta G^{vap}} nH_2O(g) \quad (298.15K, 1atm)$$
$$\downarrow \Delta H_a, \Delta S_a \qquad\qquad \uparrow \Delta H_c, \Delta S_c$$
$$nH_2O(l) \xrightarrow{\Delta H_b,\Delta S_b} nH_2O(g) \quad (398.15K, 1atm)$$
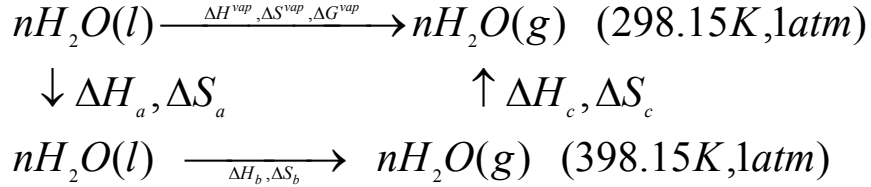
Figure 4.7. Cycle for calculating ΔG^{vap} for water

With the help from the above cycle, the vaporization Gibbs free energy, ΔG^{vap}, at 298.15K and one atmosphere pressure can be evaluated as follows:

$$\Delta G^{vap} = \Delta H^{vap} - 298.15K \times \Delta S^{vap} (at \quad 298.15K) \quad (4.2)$$

$$\Delta H^{vap} = \Delta H_a + \Delta H_b + \Delta H_c \quad (4.3)$$

$$\Delta S^{vap} = \Delta S_a + \Delta S_b + \Delta S_c \quad (4.4)$$

$$\Delta H_a = \int_{298.15K}^{398.15K} C_p^l dT, \Delta S_a = \int_{298.15K}^{398.15K} \frac{C_p^l dT}{T}, \Delta S_b = \frac{\Delta H_b}{398.15K} \quad (4.5)$$

$$\Delta H_c = \int_{398.15K}^{298.15K} C_p^g dT, \Delta S_c = \int_{398.15K}^{298.15K} \frac{C_p^g dT}{T} \quad (4.6)$$

ΔH_b is the heat of vaporization for water at 398.15K and 1atm, C_p^l, and C_p^g are the constant pressure heat capacity for liquid and gaseous water respectively. Experimental values for ΔH_b, C_p^l, and C_p^g are readily available in the literature. So we have several ways involving theoretical and experimental data to calculate a reliable estimate of ΔG_2 in equation 4.1.

In order to estimate the contribution from the "hydrophobic force", we propose the following procedures.

$$CD \bullet nH_2O(g) \xrightarrow{\Delta G(g)} CD(g) + nH_2O(g)$$
$$\downarrow \Delta G_3 \qquad\qquad\qquad \uparrow \Delta G_4$$
$$CD \bullet nH_2O(l) \xrightarrow{\Delta G(aq)} CD(l) + nH_2O(l)$$
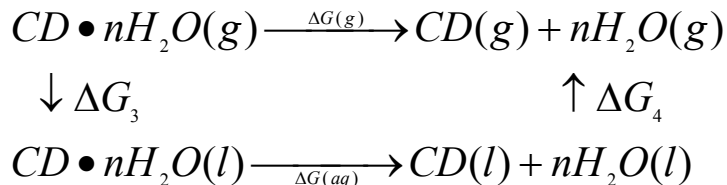
Figure 4.8. Cycle for estimating the hydrophobic contribution to the association energy

The Gibbs free energy contribution for the "hydrophobic force", $\Delta G(aq)$ in Figure 4.8, is evaluated according to the equation below.

$$\Delta G(aq) = \Delta G(g) - \Delta G_3 - \Delta G_4 \qquad (4.7)$$

$\Delta G(g)$, ΔG_3 and ΔG_4 can be calculated in the way discussed before. As with ΔG_2 in Figure 4.6, we also have several ways to estimate ΔG_4.

Investigating the "hydrophobic forces" will probably be a difficult problem since the number of included waters is not known and the complexation enforced expulsion of the water molecules may not be complete. Since the most stable complexes found here appear to be predominantly surface complexes rather than internal or encapsulated complexes, it is not clear how many embedded water molecules, if any, will be expelled.

4.3. Excited state calculations

Based on the orbital composition and the orbital shapes, we conclude that the following three excitations are correlated: transition from MO's 225 to 226 for the isolated CD, the transition from MO's 264 to 266 for the CD-AD complex and the one from MO's 259 to 270 for the CD-FC complex. From the data in Tables 3.63, 3.65 and 3.67, it is clear that the largest contributions to the three lower states in these three transitions, (MO 225 for CD, MO 264 for CD-AD and MO 259 for CD-FC), are from the $3P_x$, $3P_y$ and $3P_z$ orbitals of the sulfur atom bonded to the Zn metal atom. And from the data in Tables 3.64, 3.66 and 3.68, we can conclude that the major contributions to the higher states in these three transitions, (MO 226 for CD, MO 266 for CD-AD and MO 270 for CD-FC), are from the 4s and 3p orbitals of the Zn atom. There are two differences in the results which are

worth discussion. One is that the lower state for the transition from CD-AD, MO 264, does not have the minor contribution from the Zn atom, as its counterparts, CD and CD-FC. Another is that the higher state for the transition for CD-FC includes a minor contribution from a carbon atom (#170) in the FC molecule, the guest molecule. But for CD and CD-AD, this does not happen.

The calculated oscillator strengths for the three transitions are 0.0103, 0.0074 and 0.0052 respectively. Since the probability of an excitation transition is proportional to the square of the oscillator strength and the transition energy, and here all three transition energies are very close to one another (7.02ev, 6.94ev and 6.92ev), the transition for the CD-FC complex is about 25% as intense as the transition for CD; and the transition for the CD-AD complex is about 52% as intense as that for CD. Although these results give the correct qualitative ordering of the effect of the guest molecules on the fluorescence spectrum of CD, they are actually a little disappointing. We expect to see higher intensity from the particle with CD-AD since according to the work of Liu[1], adding AD to the solution mixture of CD and FC can bring back more than 90% of fluorescence emission. However we note here that the transition process that is under study here is in some ways the opposite process of the fluorescence emission. In fluorescence emission, the electron transits from the singular excited state with a higher energy level to the singular lower state with less energy. Although here we could make an argument that a lower transition probability caused by a weaker oscillator strength indicates that fewer electrons were able to be excited to the higher state and hence we have a smaller excited population to feed the fluorescence. But the fluorescence emission is a very complicated process as evidenced by the fact that there has been little predictive theoretical work devoted to this area. After the electron is excited to the higher state, the molecule can undergo various decay and kinetic processes that can dramatically affect the fluorescence emission. Furthermore, in the experimental work from Liu[1], the metal in the nano-particle is Cadmium instead of Zinc that we used in this work. And in order to make the task quantum mechanically treatable, we simply have the one single metal atom attached to one arbitrarily chosen sulfur atom in CD (in either free CD, CD-AD or CD-FC). Apparently this simple model is not able to represent all of the interactions between the nano-particles and the anchored complexes (either CD, CD-AD or CD-FC). So in order

106

to have a better understanding of the mechanism for fluorescence quenching and recovery[1], we need:

1. A proper method to predict the probability or intensity of electrons transiting from the excited state to lower state that includes the effect of the various decay and rearrangement processes that occur between the absorption and emission.

2. A better model to simulate the presence of the metal nano-particles (or the metal atoms) and a better method to describe the interaction between the nano-particle and the species CD, CD-AD and CD-FC.

These will be topics for future work.

5. Appendix

Figures
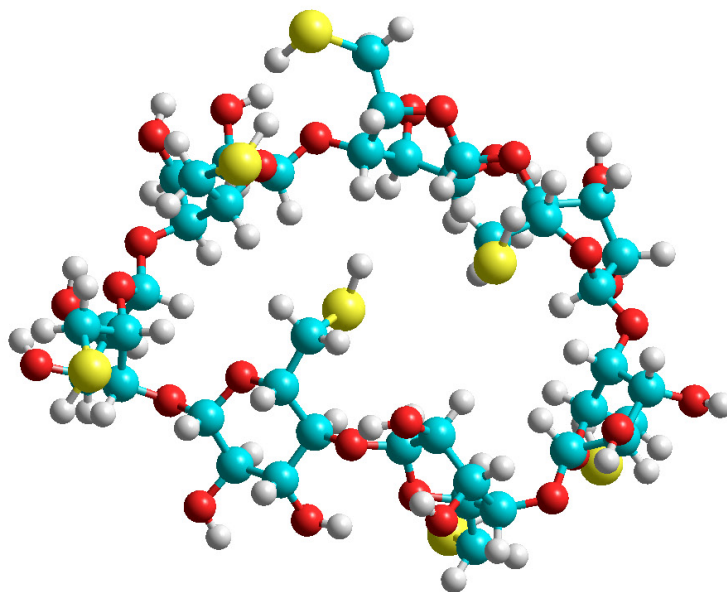
Figure 5.1. The structure of perthiolated beta-cyclodextrin (CD)
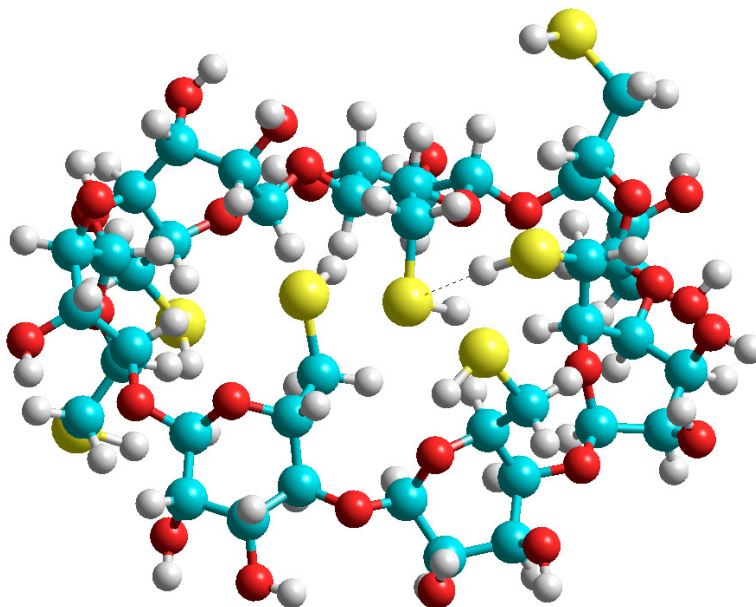
Figure 5.2. 1st conformer with energy of 109.3676 kcal/mol
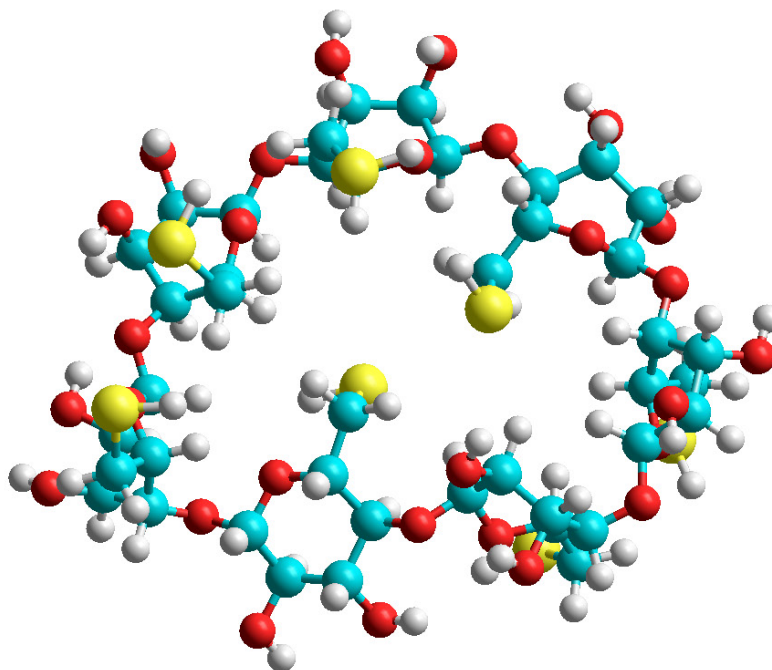
Figure 5.3. 2nd conformer with energy of 109.5789 kcal/mol
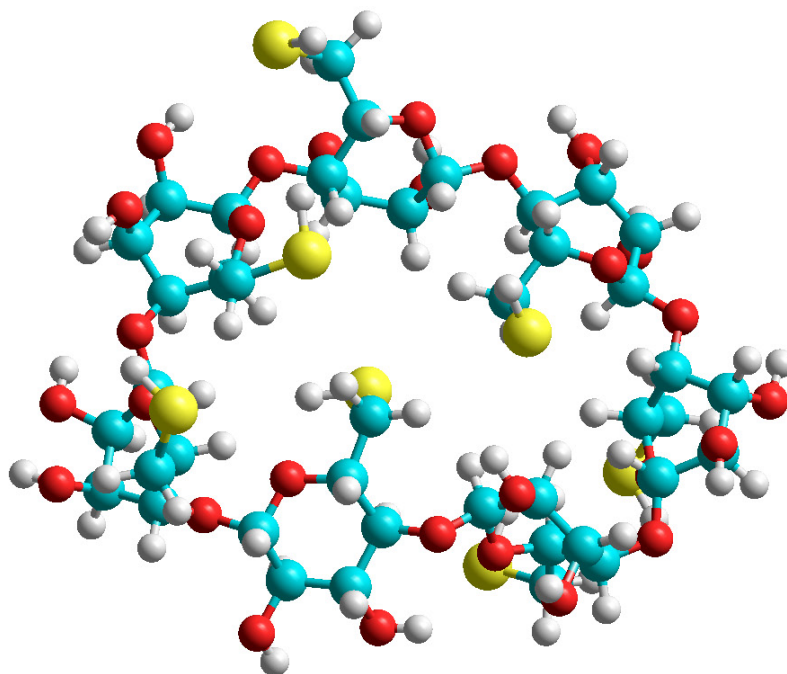
Figure 5.4. 3rd conformer with energy of 109.734 kcal/mol
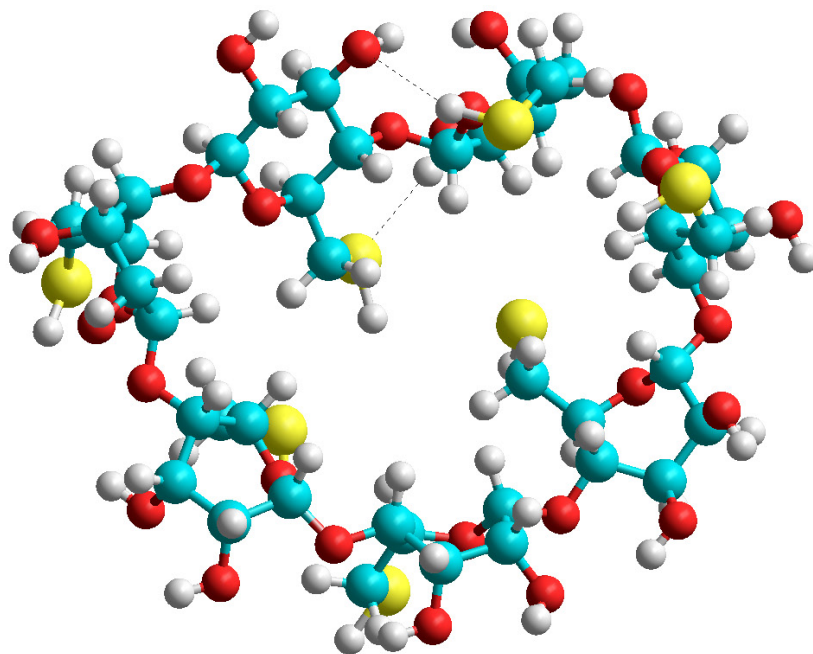
Figure 5.5. 4th conformer with energy of 109.8334 kcal/mol
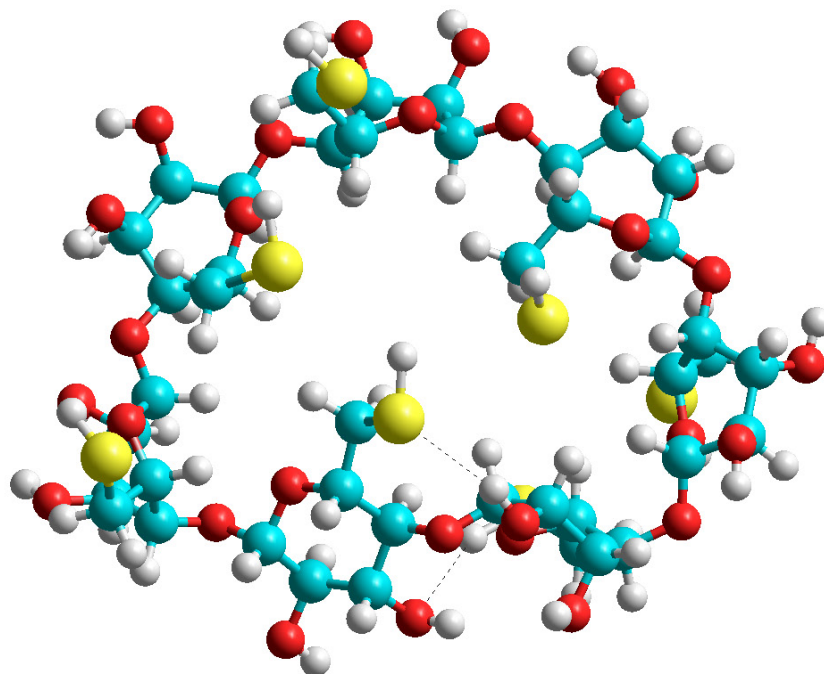
Figure 5.6. 5th conformer with energy of 109.9717 kcal/mol
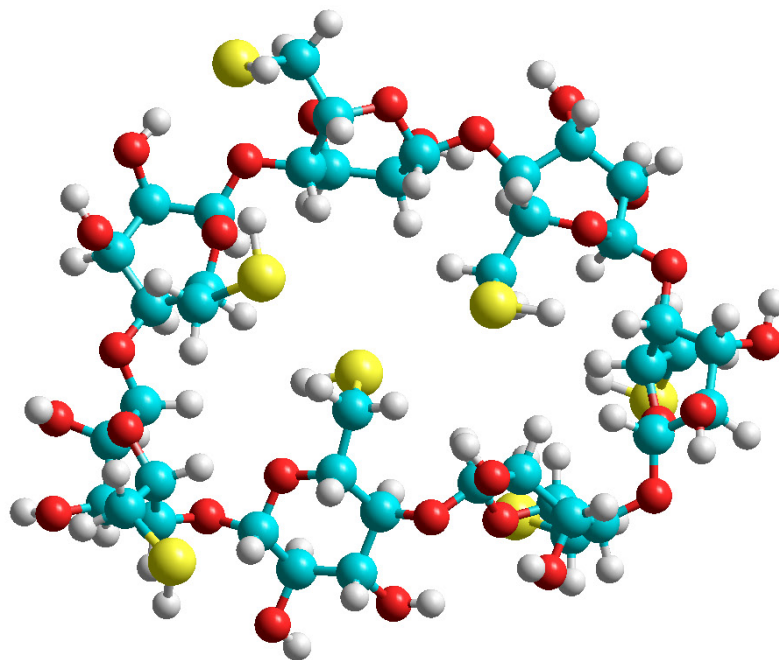
Figure 5.7. 6th conformer with energy of 109.9761 kcal/mol
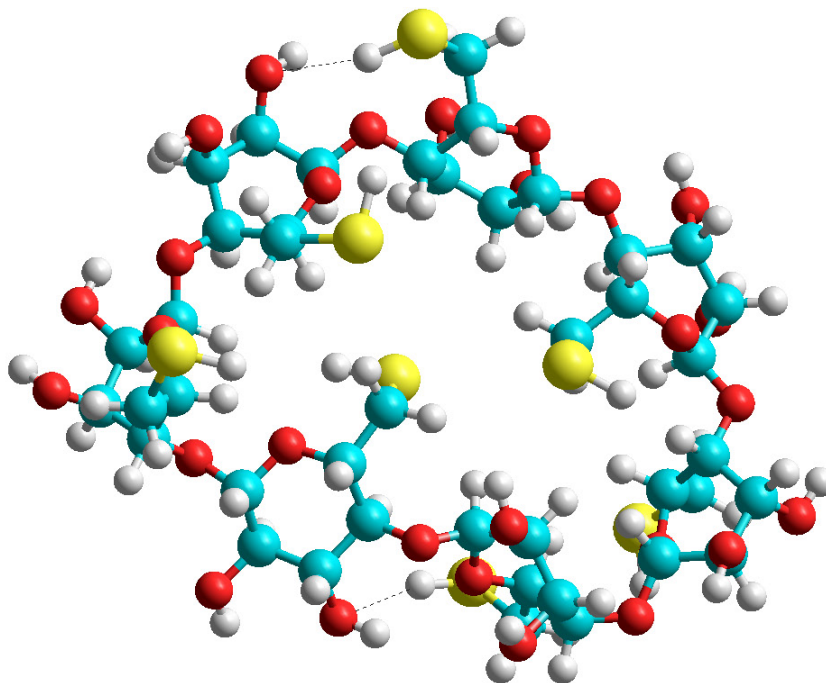
Figure 5.8. 7th conformer with energy of 109.9802 kcal/mol
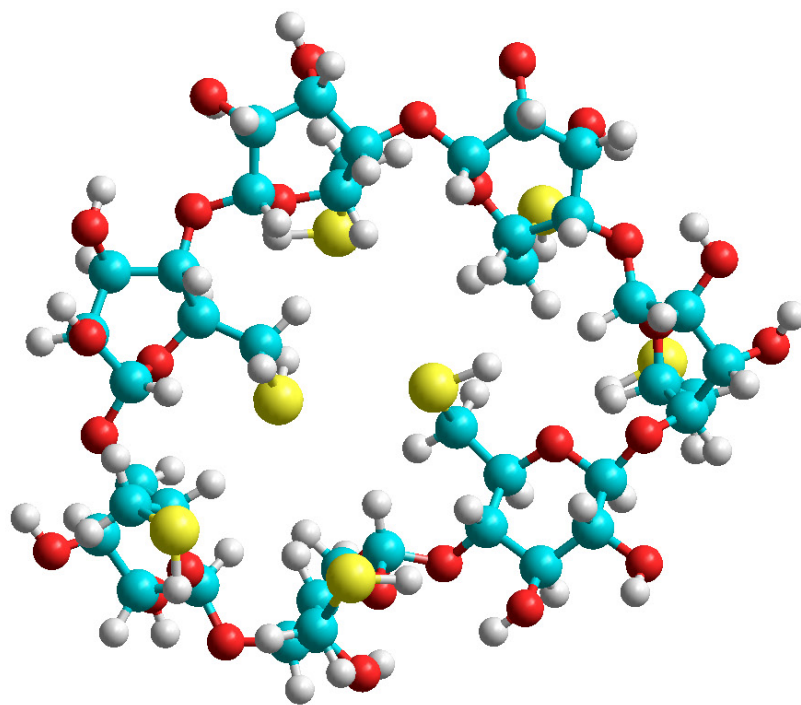
Figure 5.9. 8th conformer with energy of 110.0331 kcal/mol
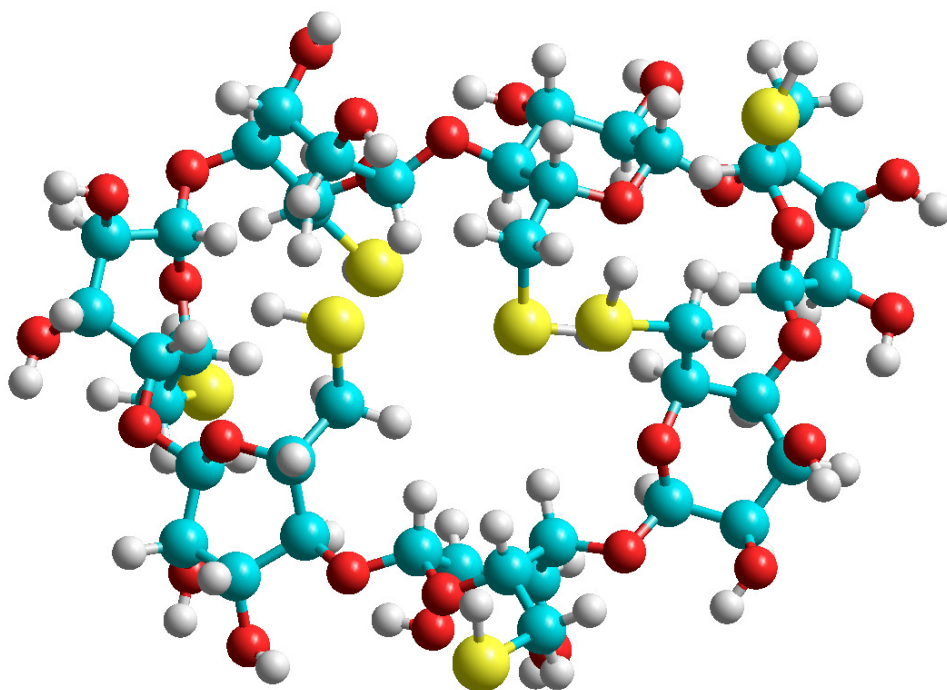
Figure 5.10. 9th conformer with energy of 110.0749 kcal/mol

112
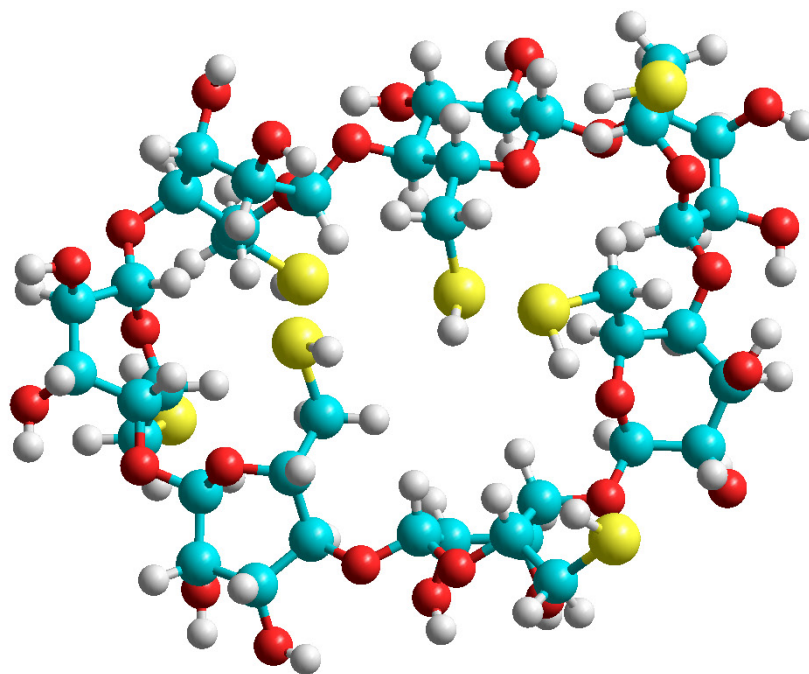
Figure 5.11. 10th conformer with energy of 110.1093 kcal /mol
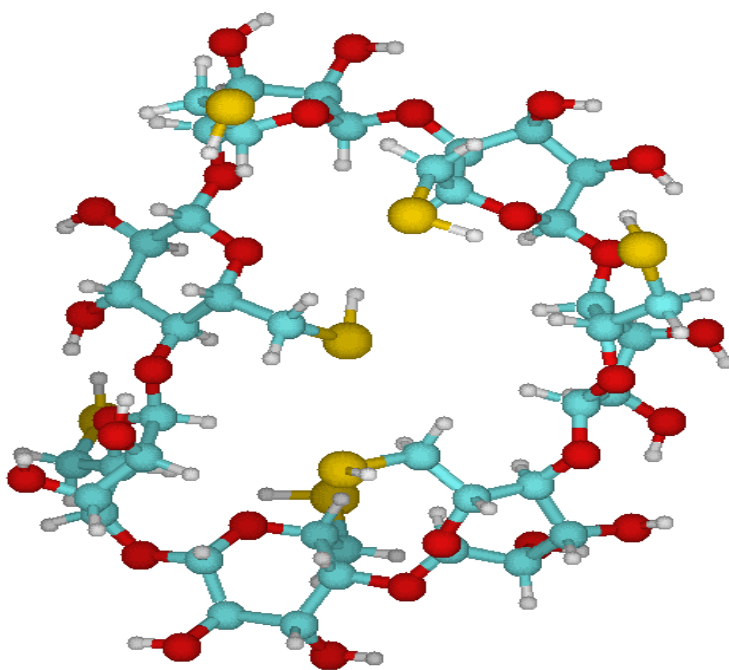
Figure 5.12. The 1st conformer optimized with the RHF/6-31G(d) method

113
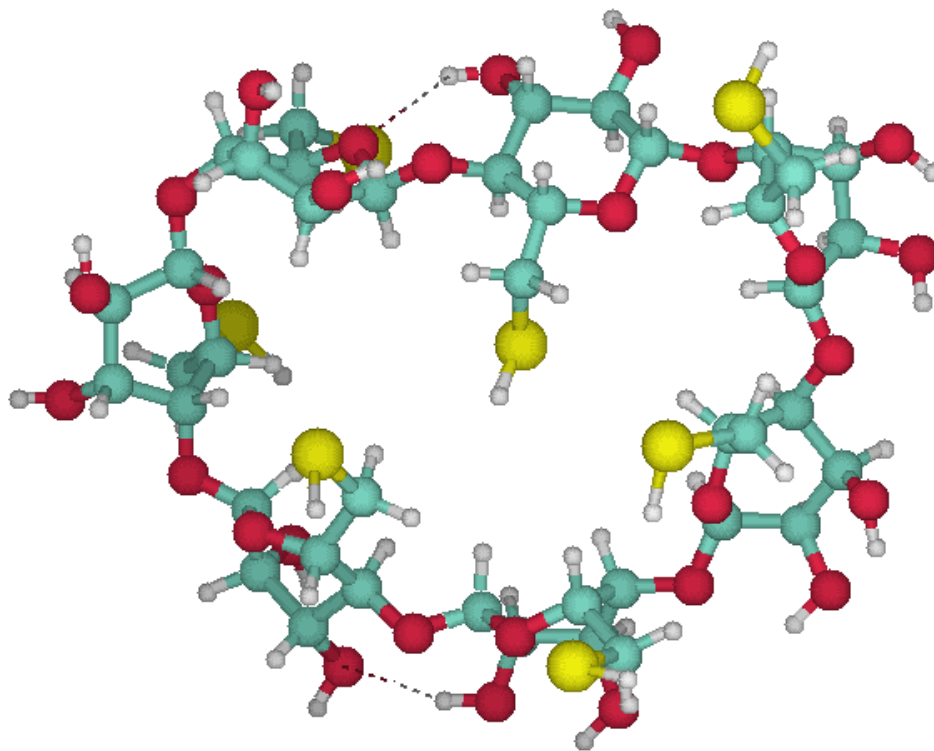
Figure 5.13. The 2nd conformer optimized with the RHF/6-31G(d) method
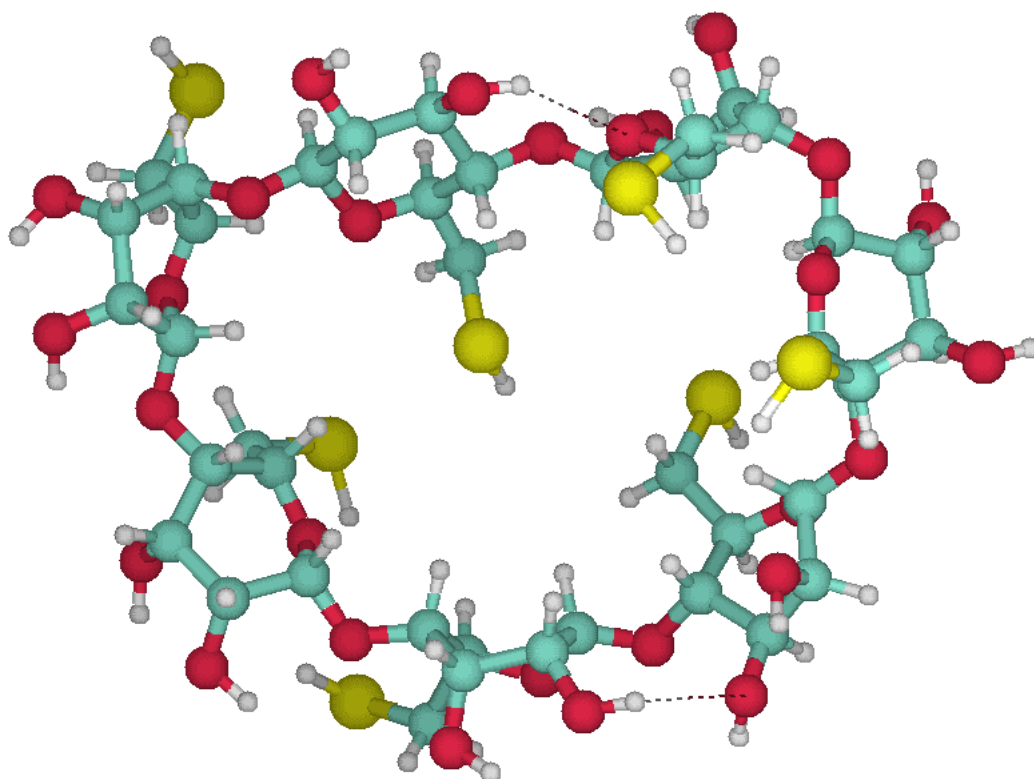
Figure 5.14. The 3rd conformer optimized with the RHF/6-31G(d) method

114
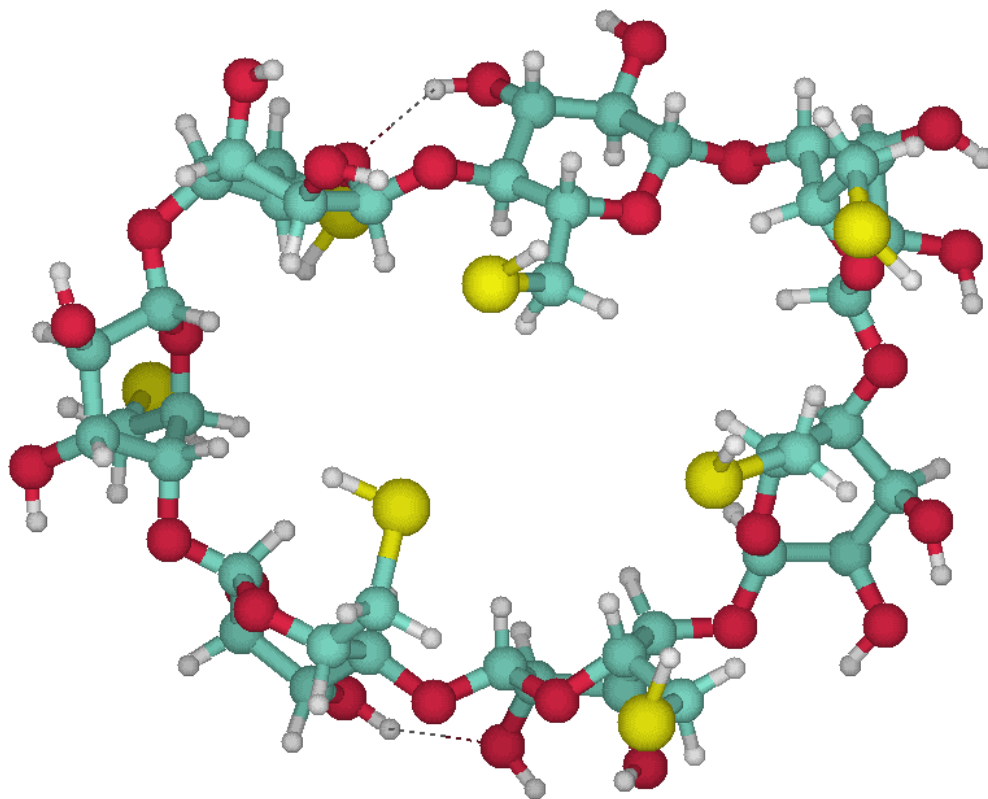
Figure 5.15. The 4th conformer optimized with the RHF/6-31G(d) method
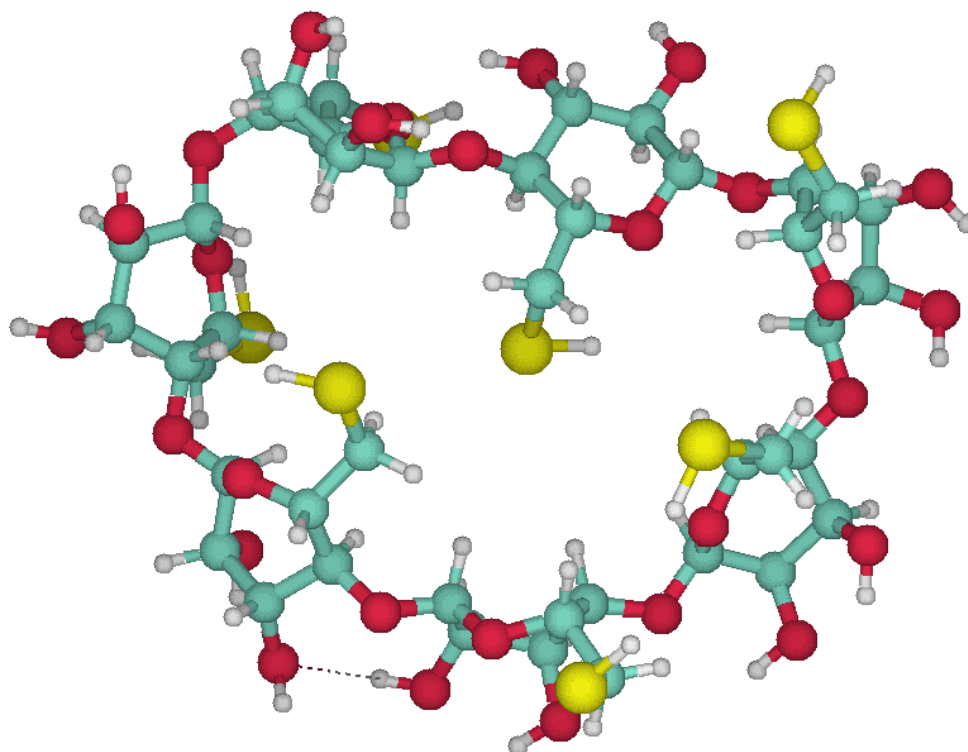
Figure 5.16. The 5th conformer optimized by with the RHF/6-31G(d) method

115
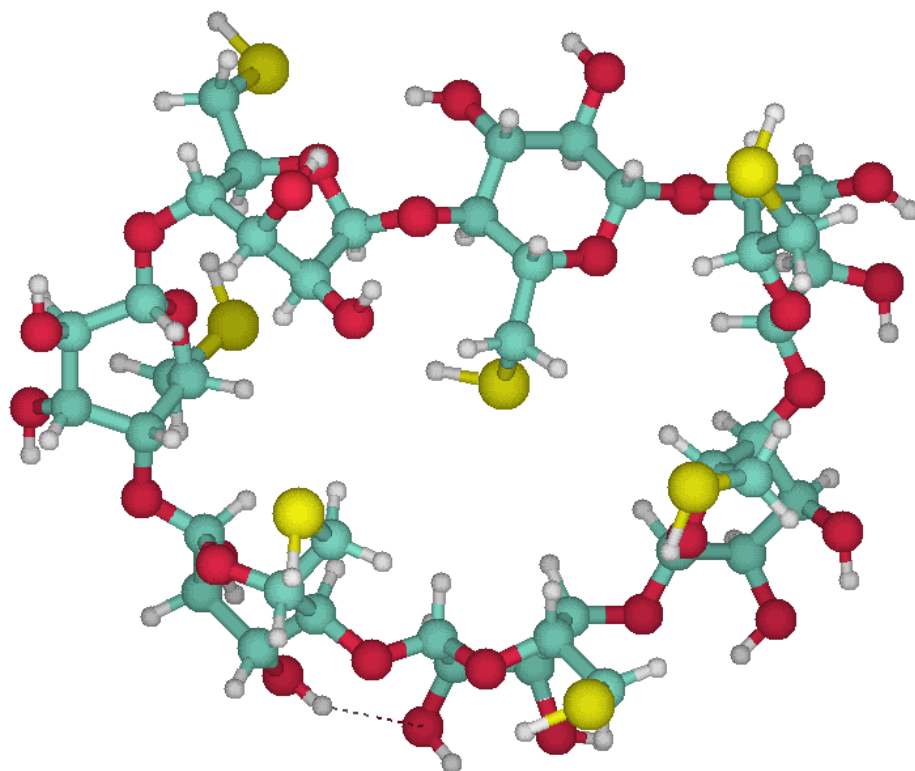
Figure 5.17. The 6th conformer optimized with the RHF/6-31G(d) method
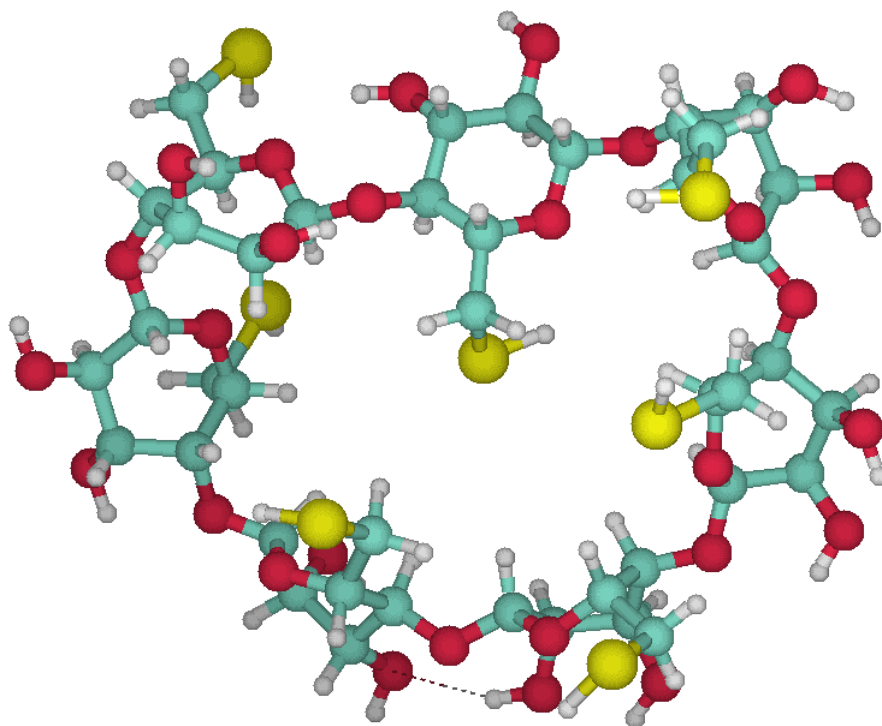
Figure 5.18. The 7th conformer optimized with the RHF/6-31G(d) method

116
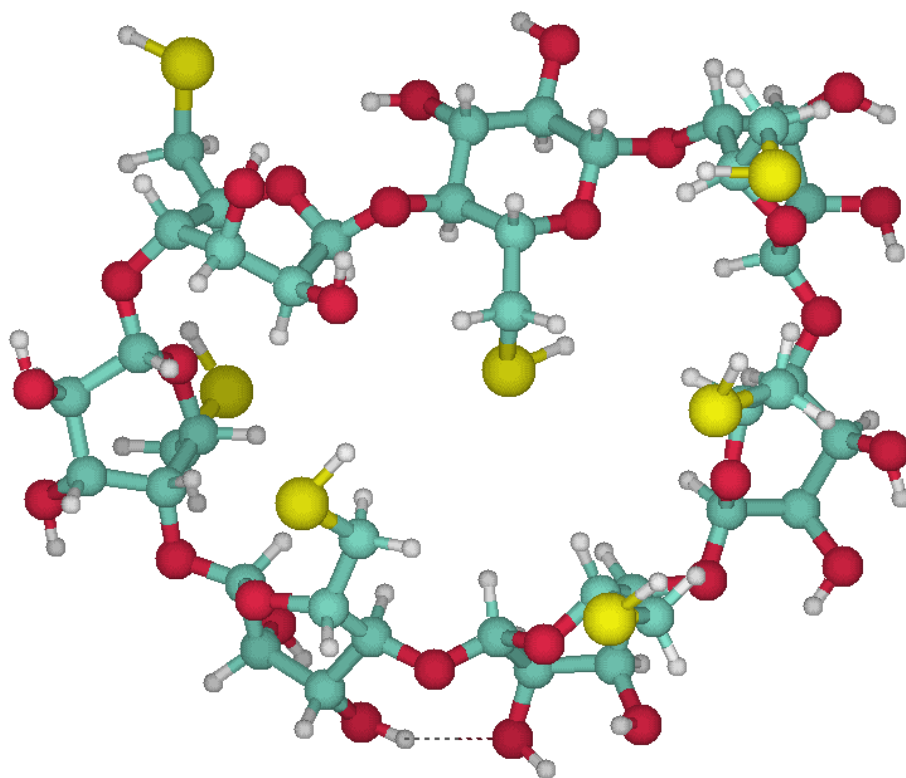
Figure 5.19. The 8th conformer optimized with the RHF/6-31G(d) method
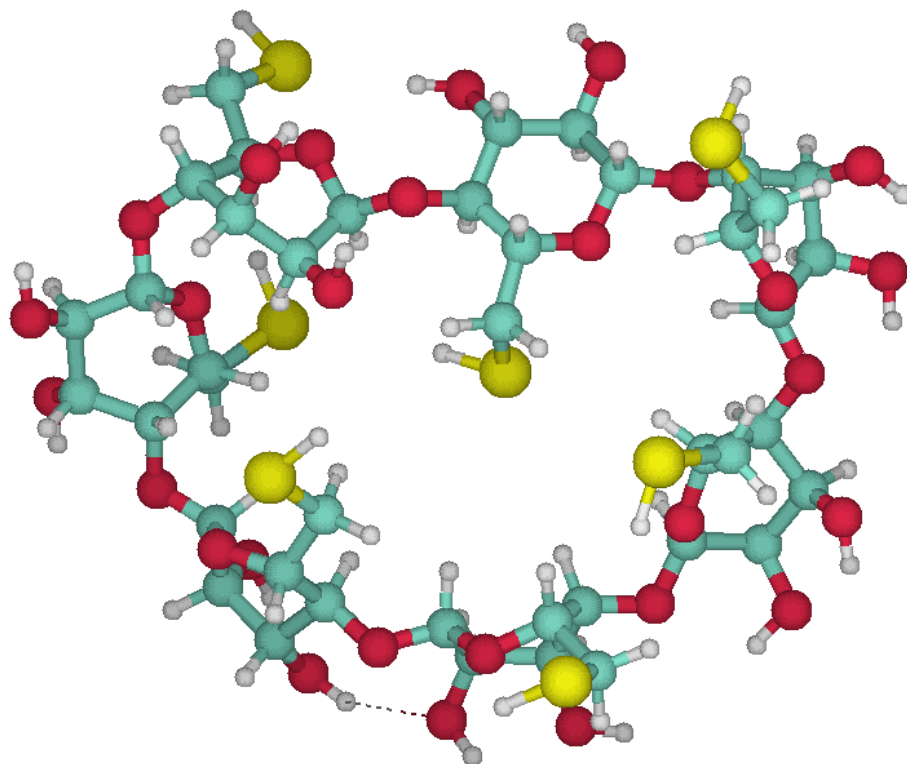
Figure 5.20. The 9th conformer optimized with the RHF/6-31G(d) method
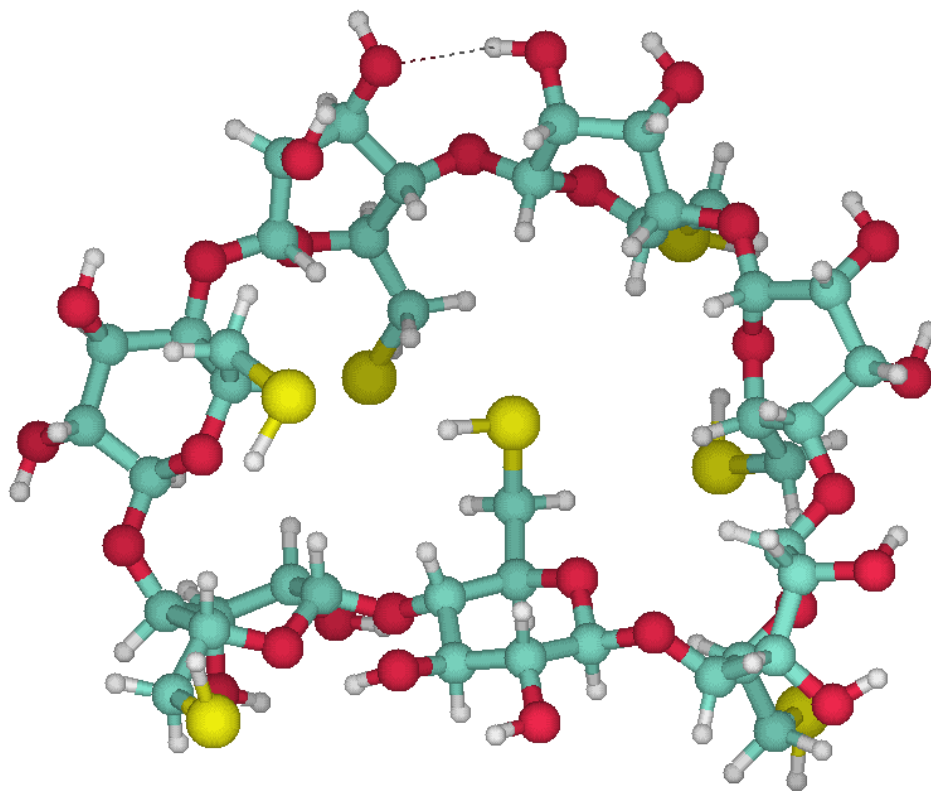
Figure 5.21. The 10th conformer optimized with the RHF/6-31G(d) method
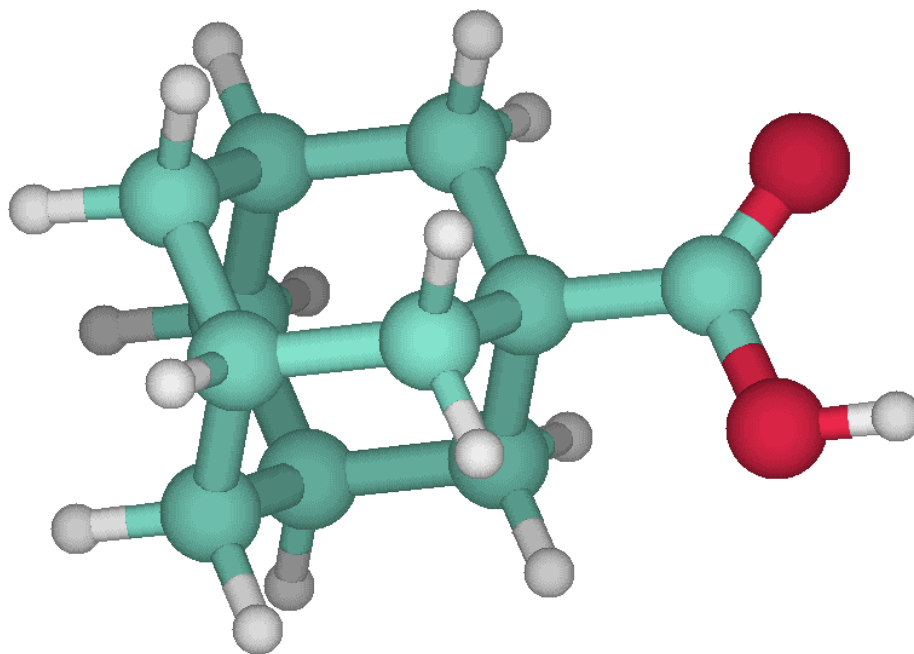
Figure 5.22. Structure of admantane carboxylic acid (AD)

118
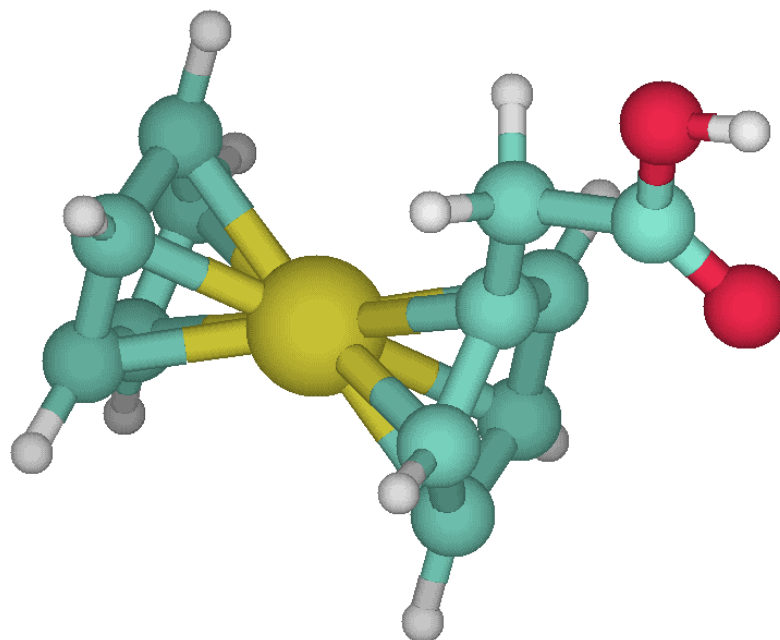
Figure 5.23. Structure of ferrocene carboxylic acid (FC)
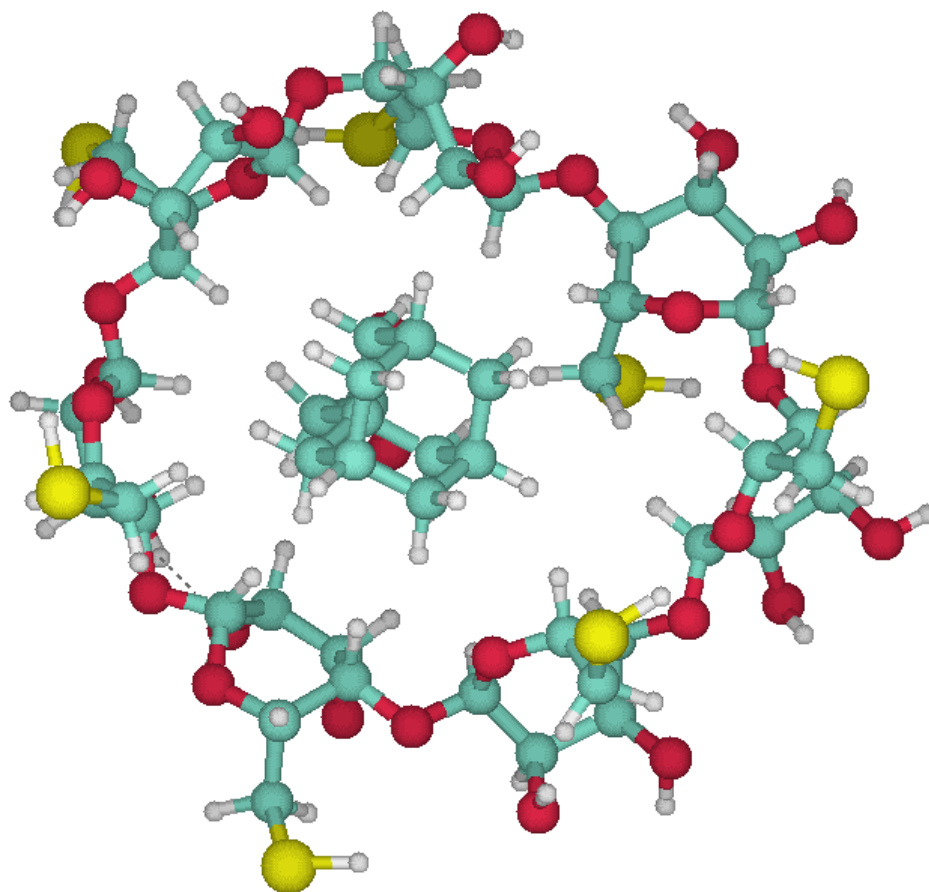
Figure 5.24. Structure of the complex of CD-AD
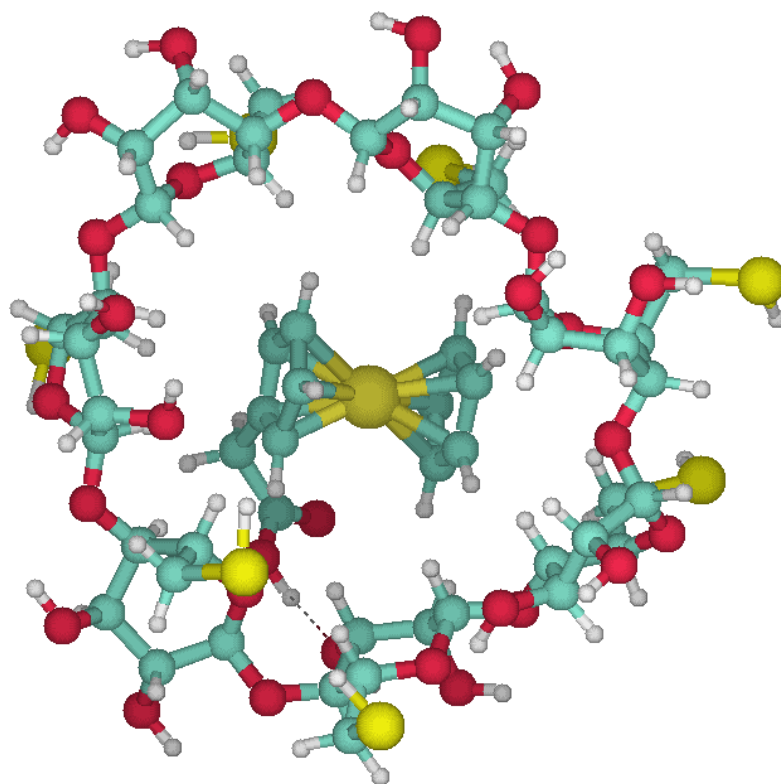
Figure 5.25. Structure of the complex of CD-FC
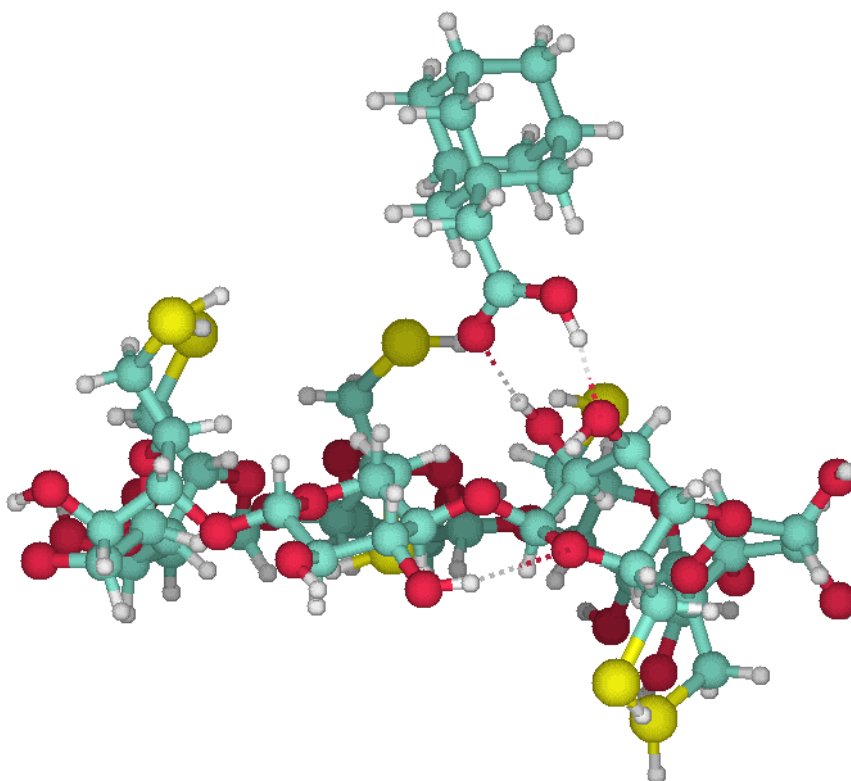
Figure 5.26. The initial geometry of CD-AD for the 1st RHF full optimization

Figure 5.26. The initial geometry of CD-AD for the 1st RHF full optimization

120
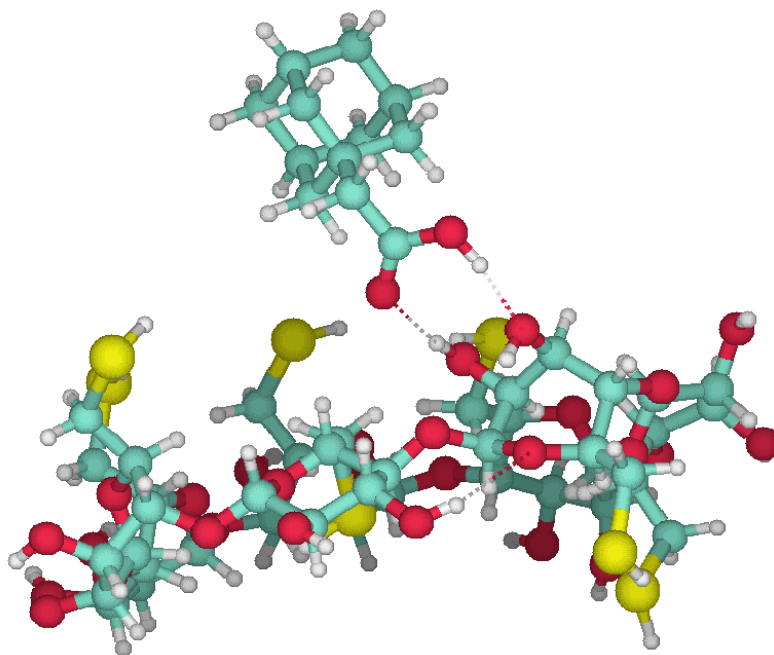
Figure 5.27. The 1st fully RHF optimized geometry of the CD-AD
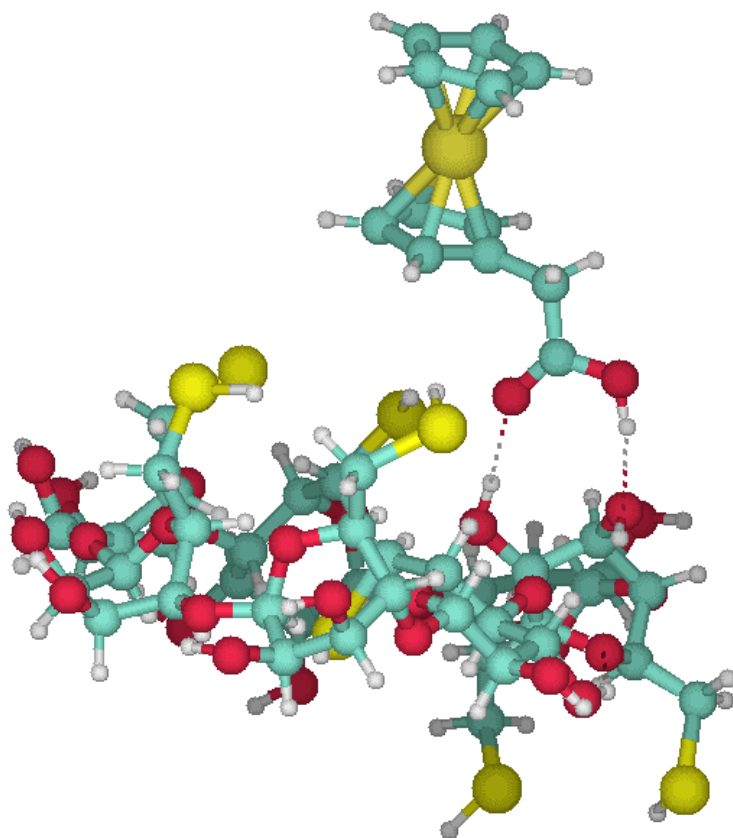
Figure 5.28. The initial geometry of the CD-FC for the 1st RHF full optimization

121
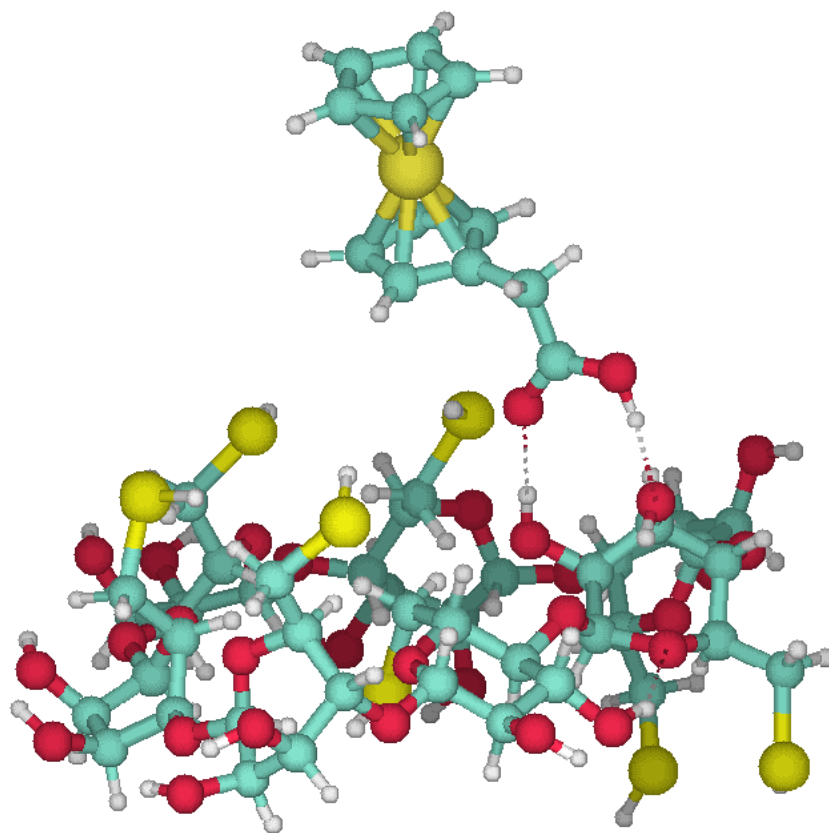
Figure 5.29. The 1st fully RHF optimized geometry of the CD-FC
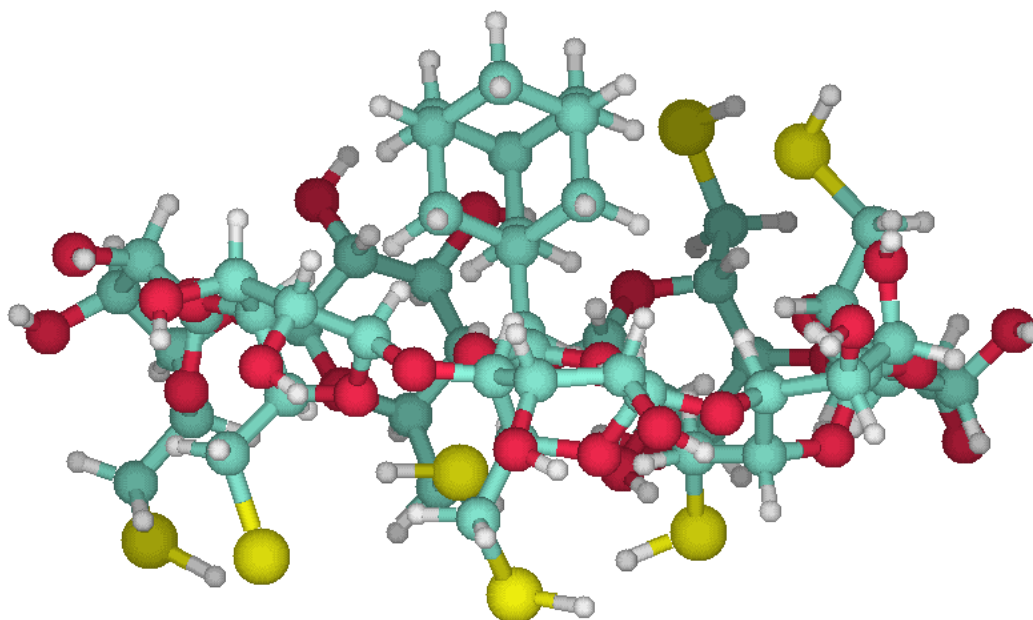
Figure 5.30. The initial geometry of CD-AD for the 2nd RHF full optimization
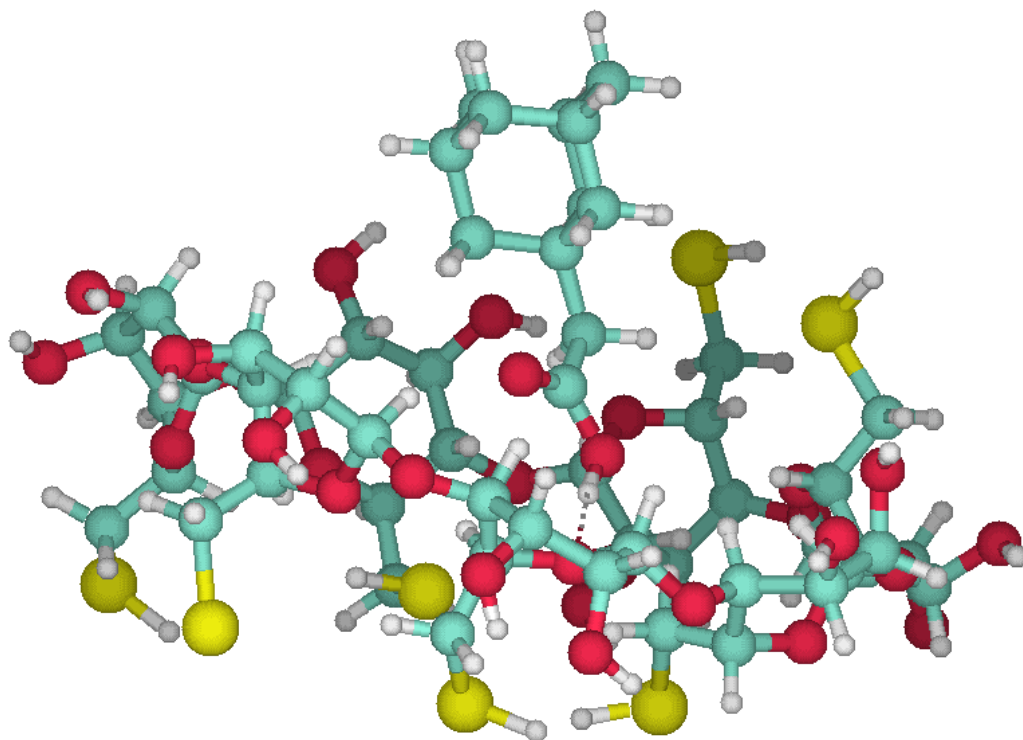
Figure 5.31. The 2nd fully RHF optimized geometry of the CD-AD
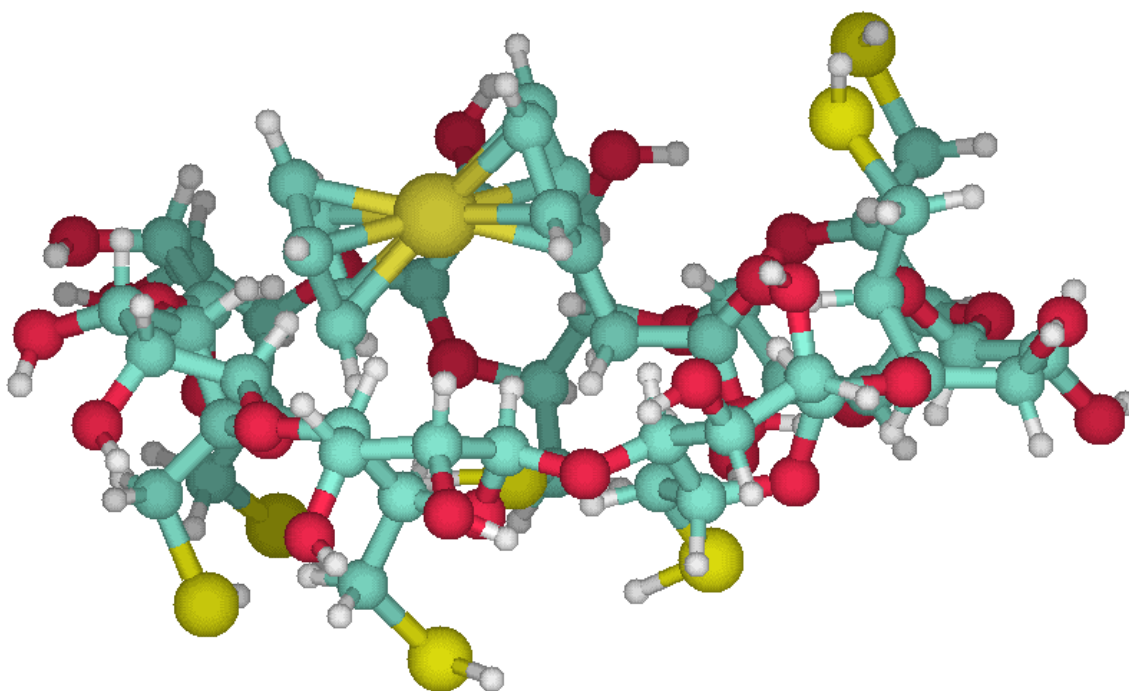
Figure 5.32. The initial geometry of the CD-FC for the 2nd RHF full optimization
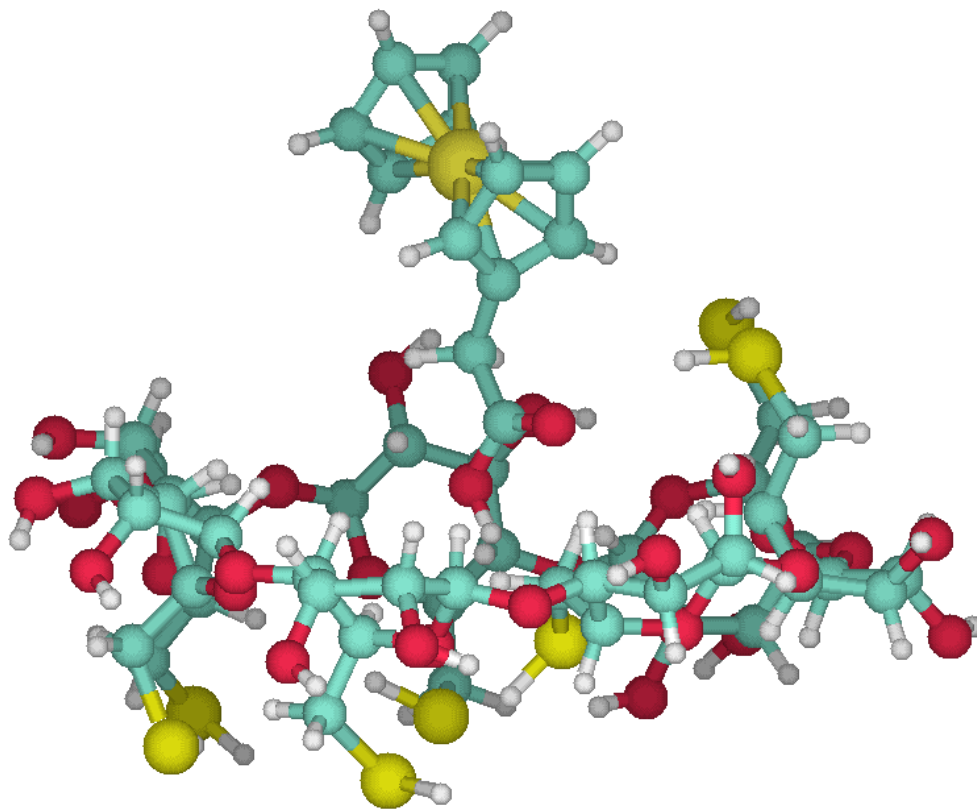
Figure 5.33. The 2nd fully RHF optimized geometry of the CD-FC
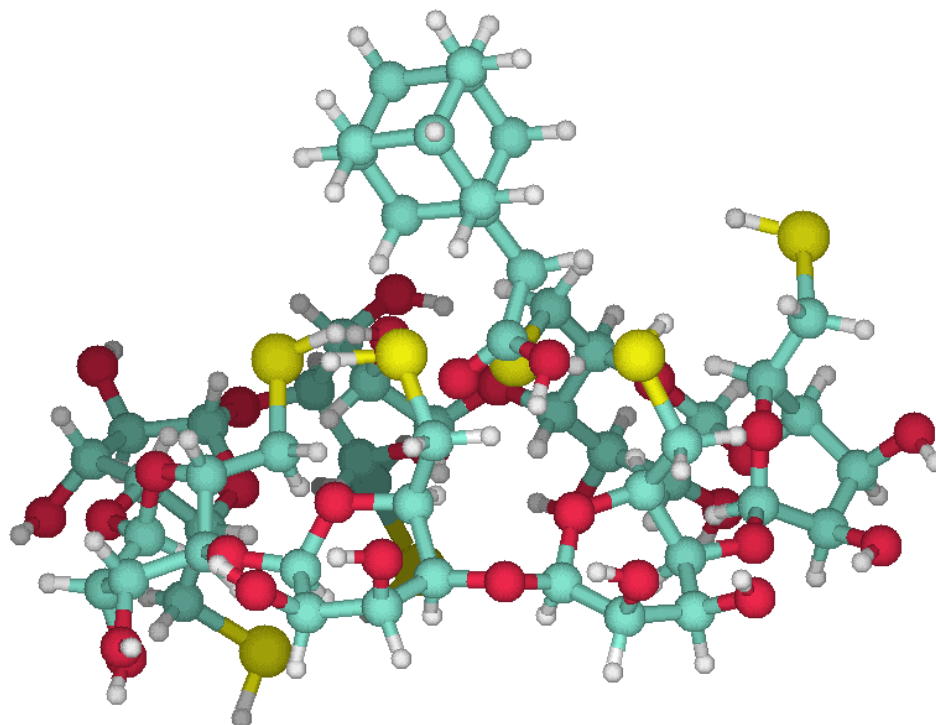
Figure 5.34. The initial geometry of CD-AD for the 3rd RHF full optimization
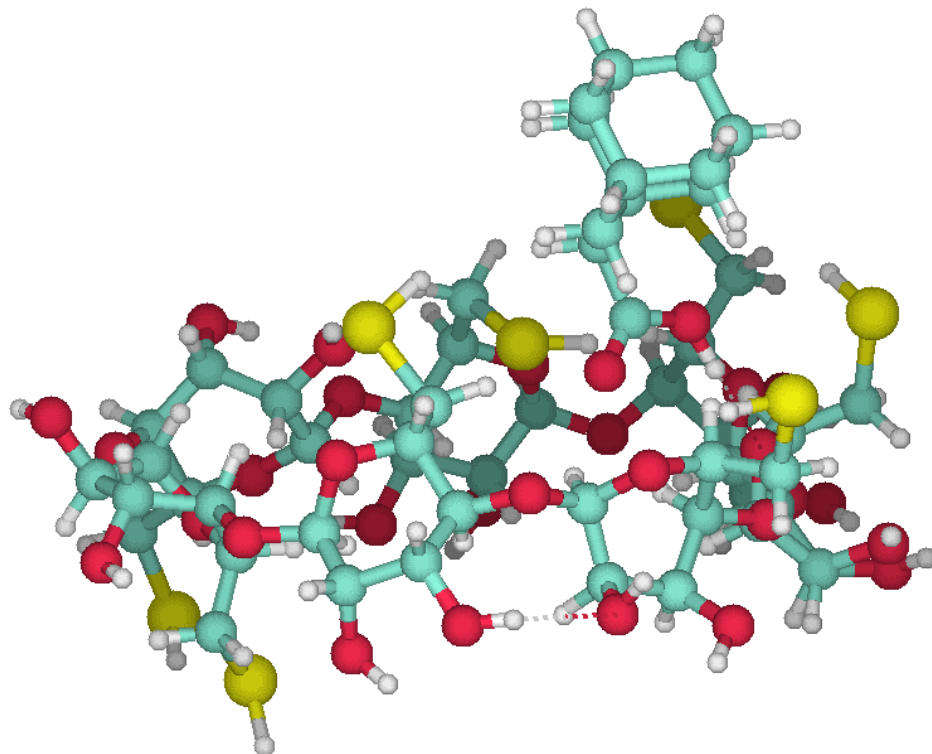
Figure 5.35. The 3rd fully RHF optimized geometry the 3rd CD-AD
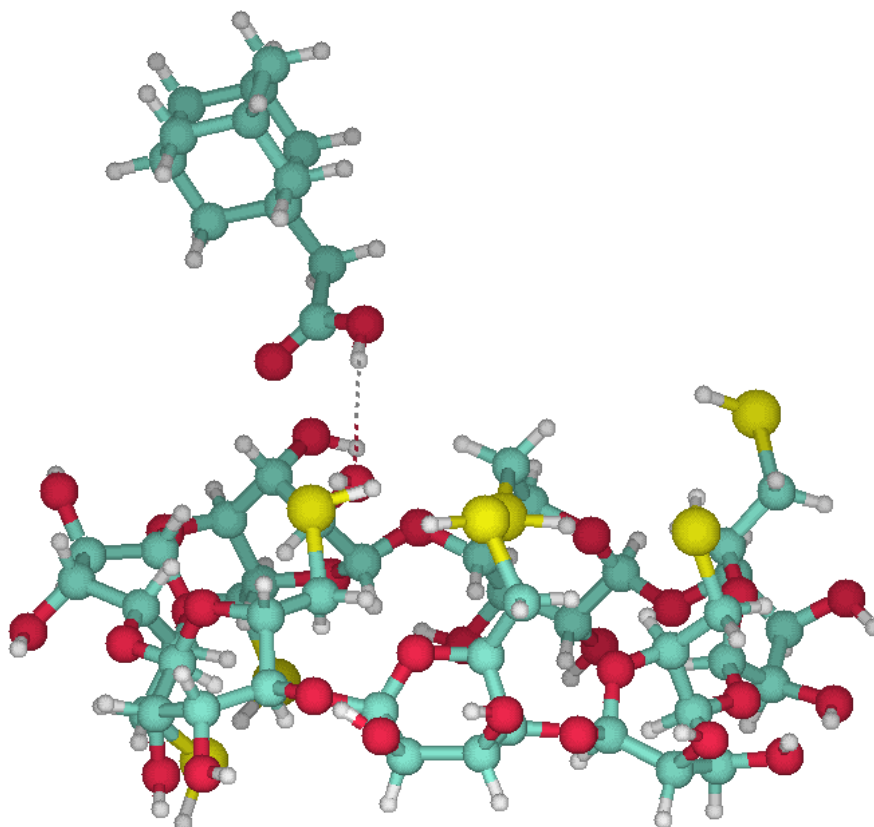
Figure 5.36. The initial geometry of CD-AD for the 4th RHF optimization

125
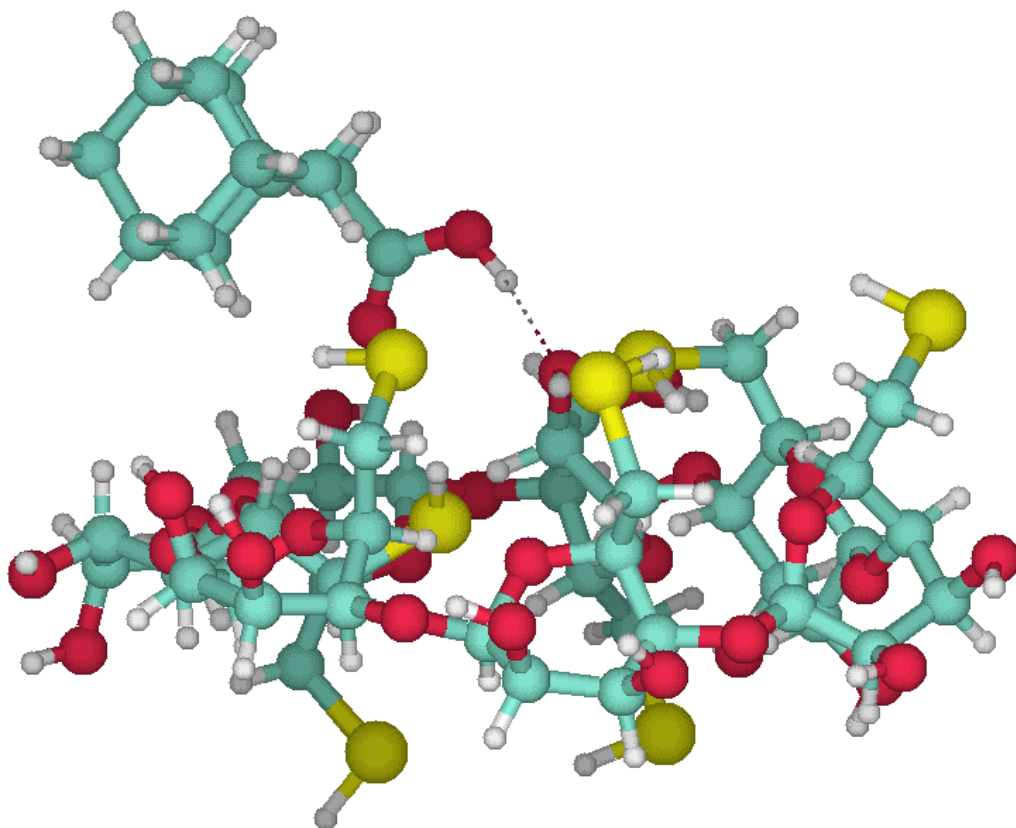
Figure 5.37. The 4th fully RHF optimized geometry of the CD-AD
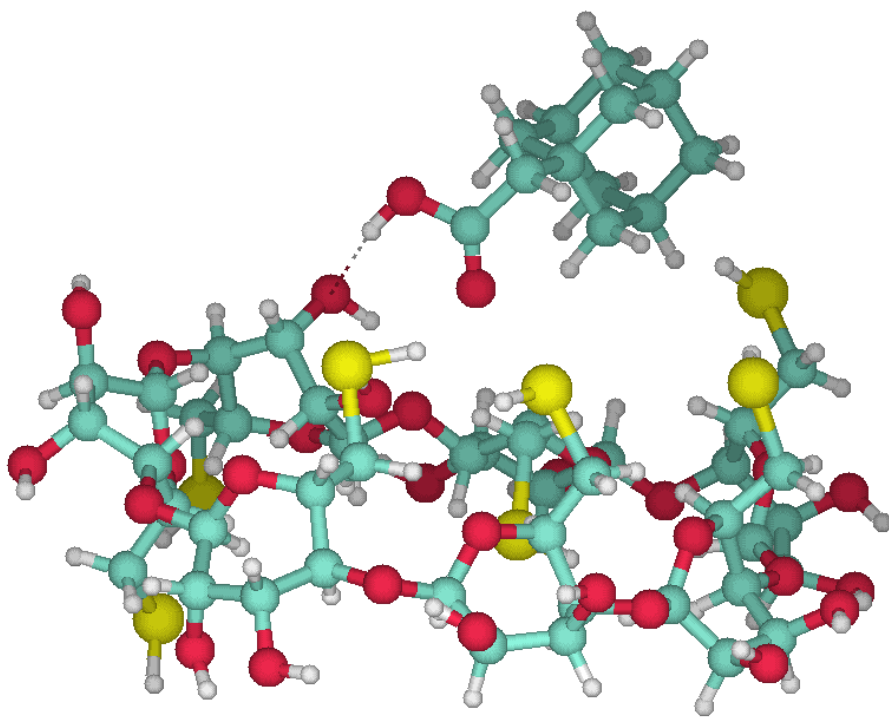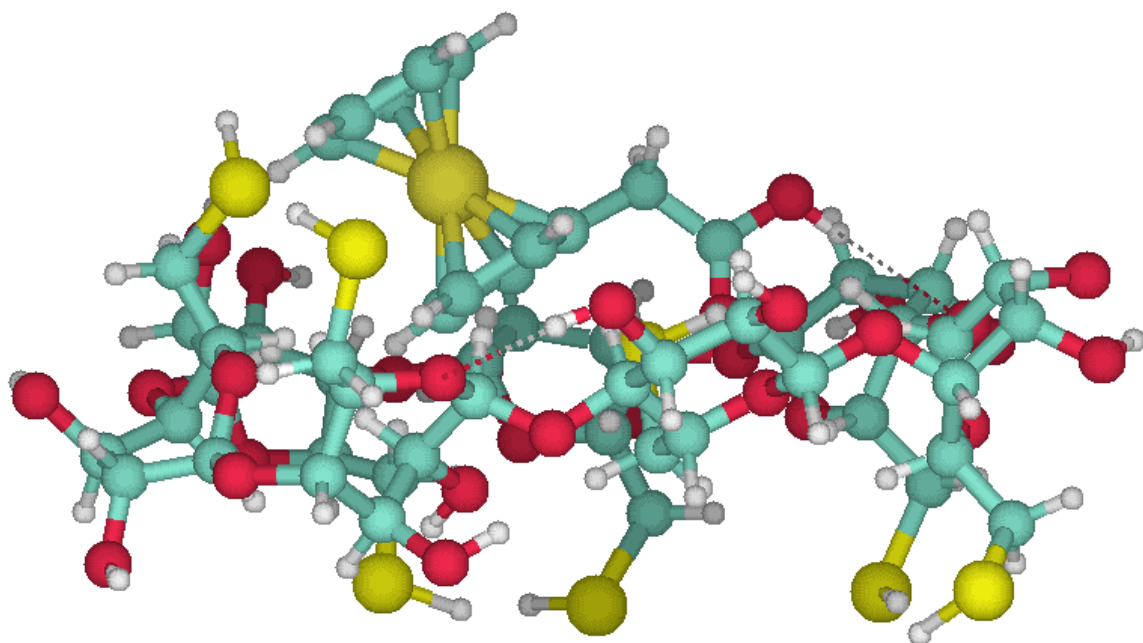

Figure 5.38. Initial CD-AD geometry for the 1st B3LYP full optimization
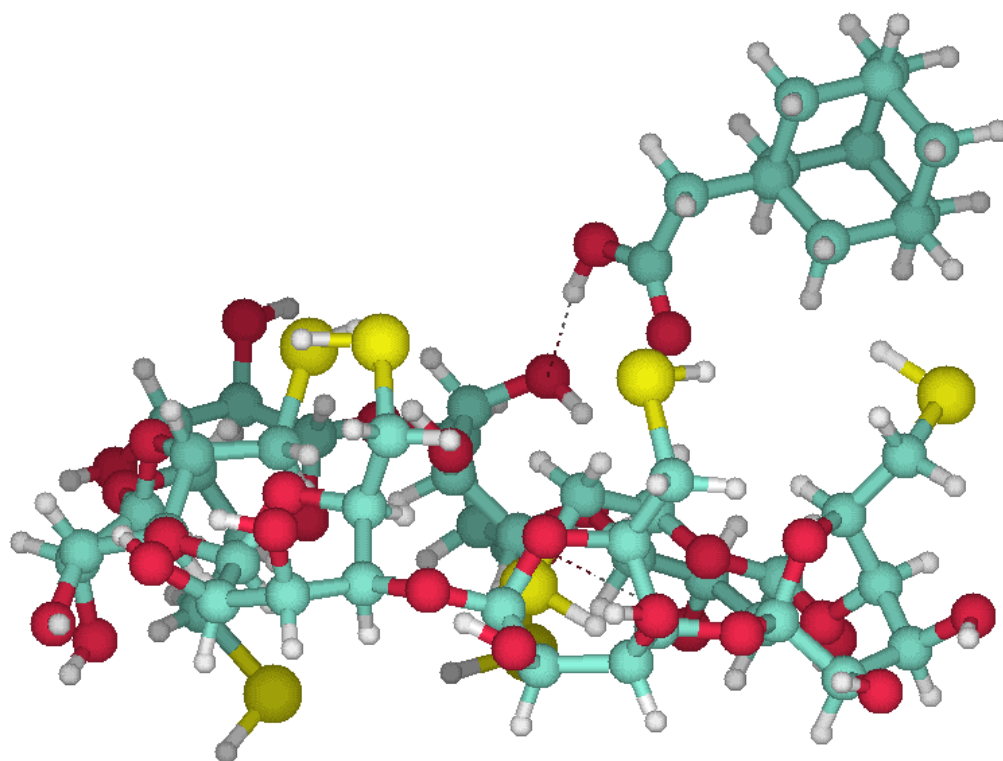
Figure 5.39. The optimized CD-AD geometry from the 1st B3LYP full optimization
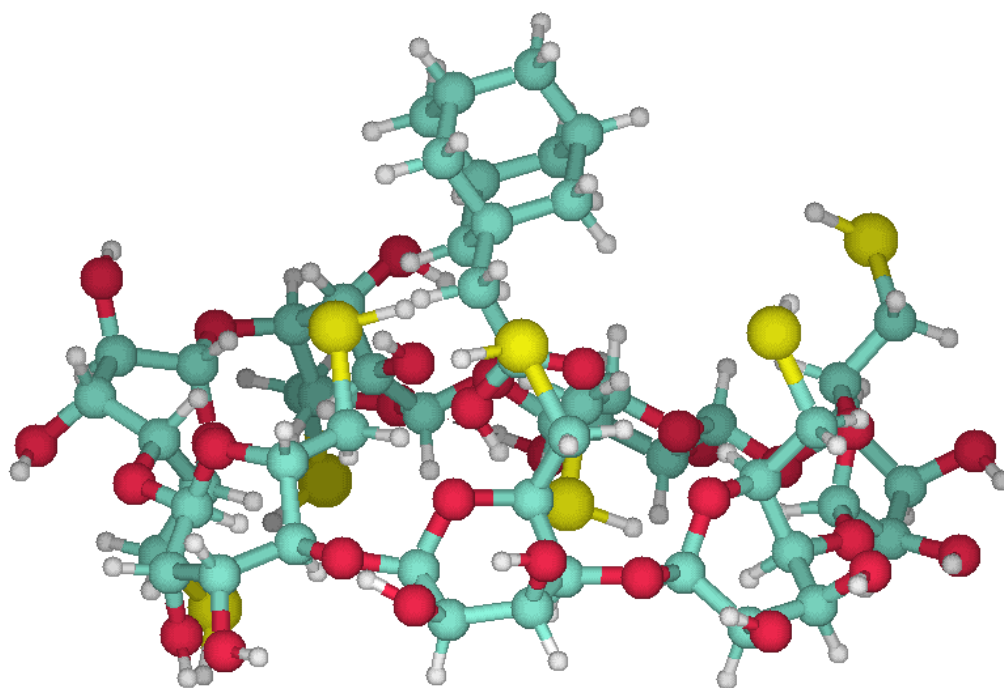
Figure 5.40. Initial CD-AD geometry for the 2nd B3LYP full optimization
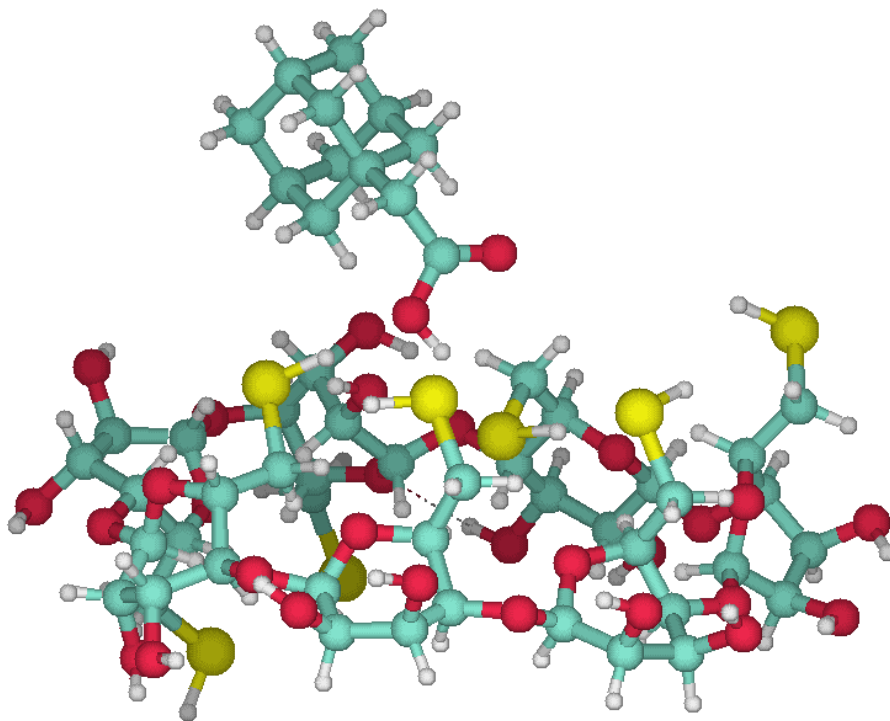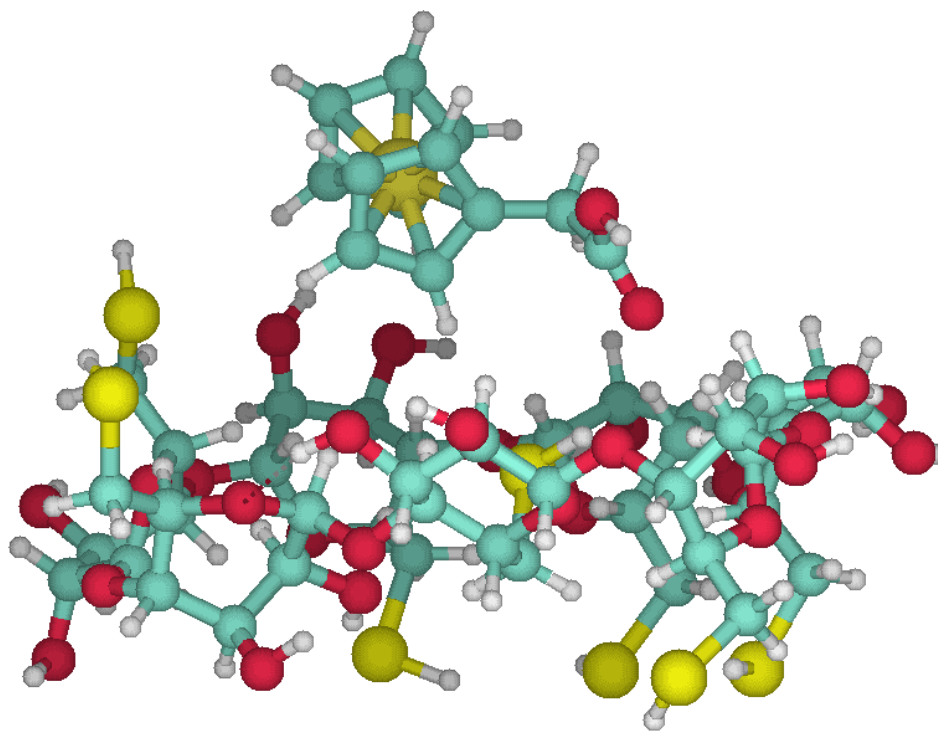
Figure 5.41. The optimized CD-AD geometry from the 2nd B3LYP optimization
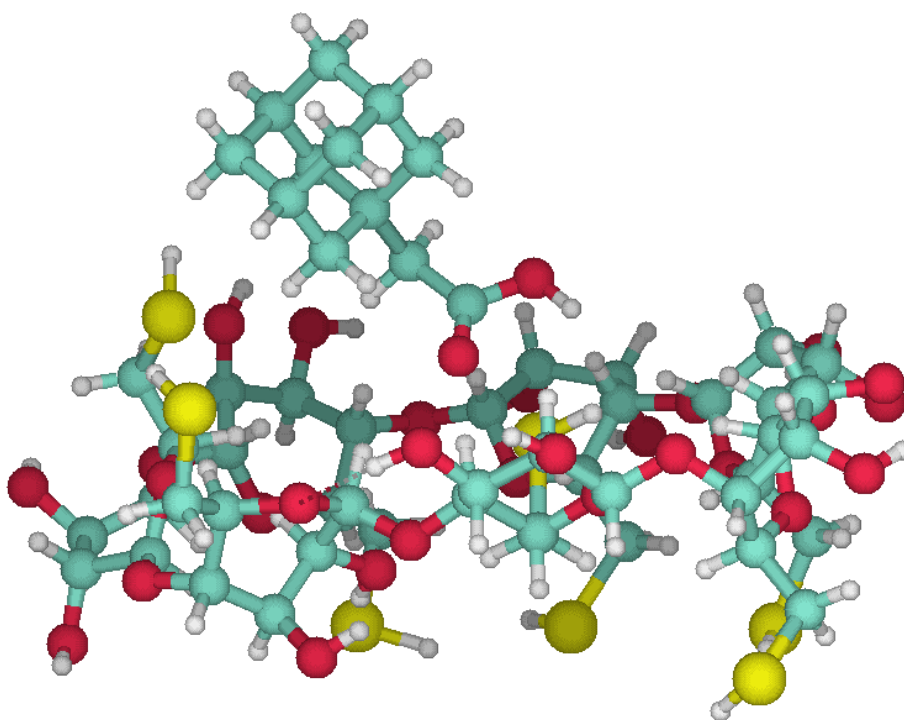
Figure 5.42. Initial CD-AD geometry for the 3rd B3LYP full optimization
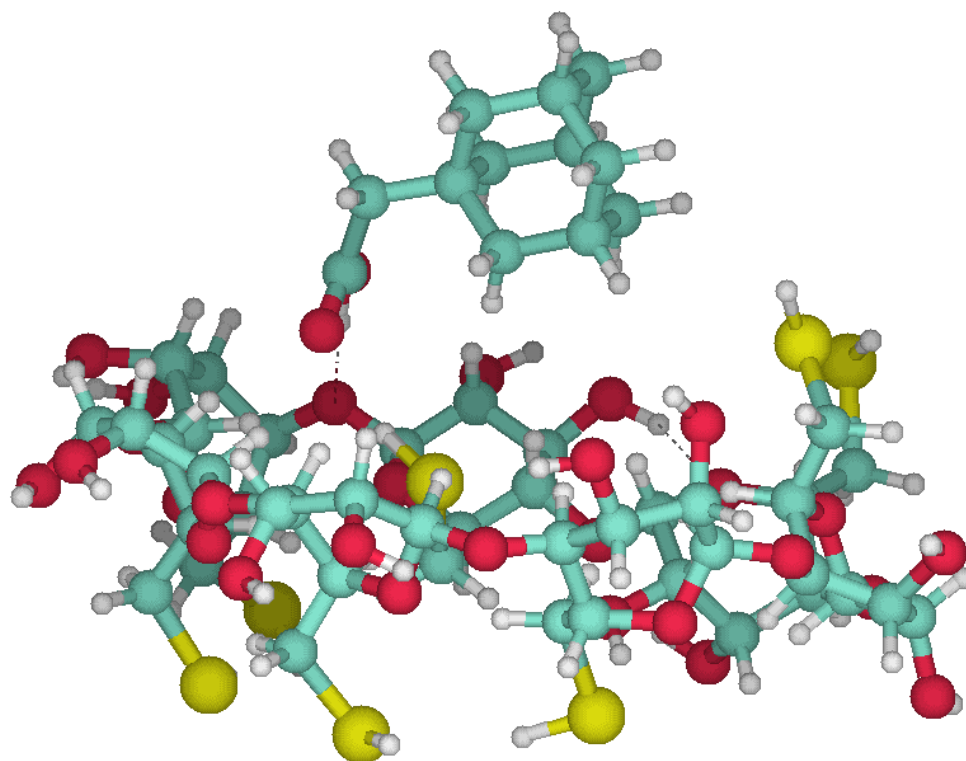
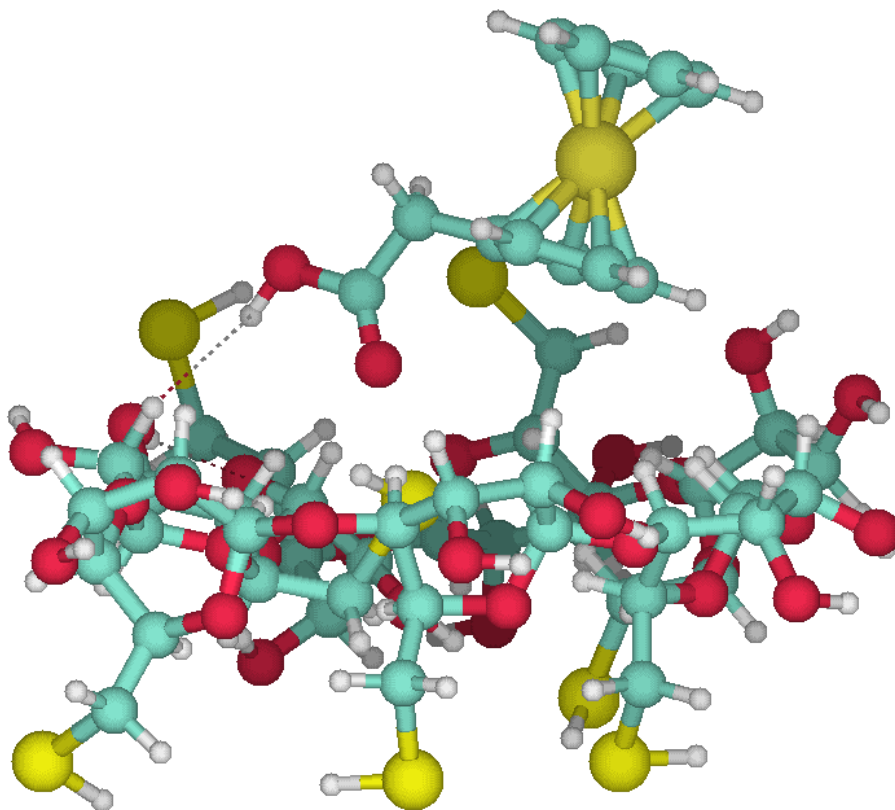Figure 5.43. The optimized CD-AD geometry from the 3rd B3LYP optimization
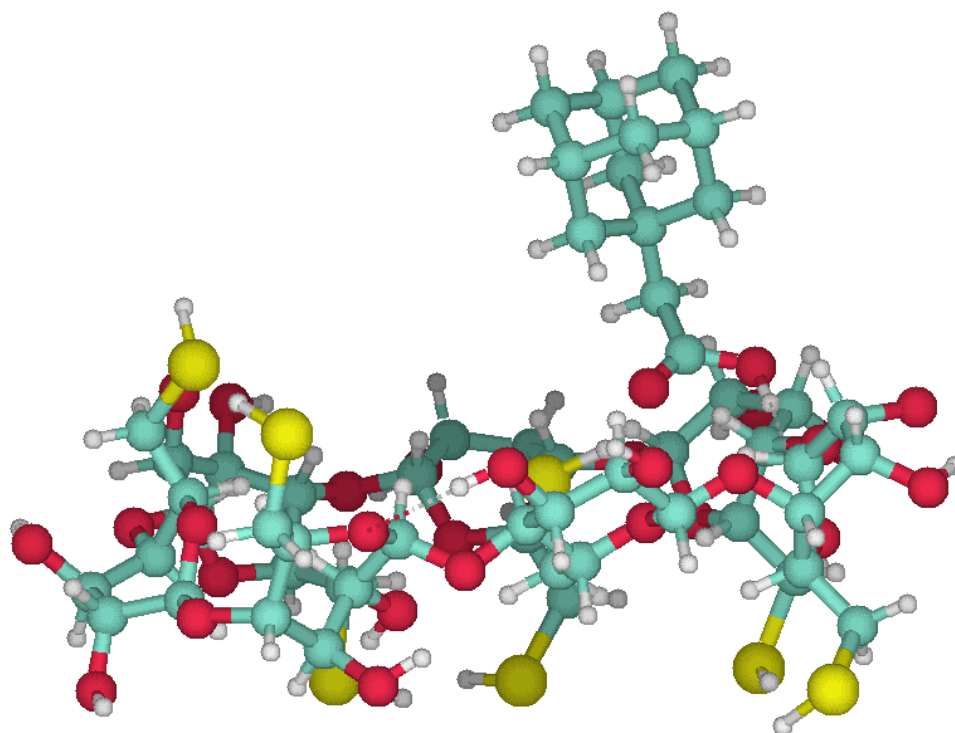
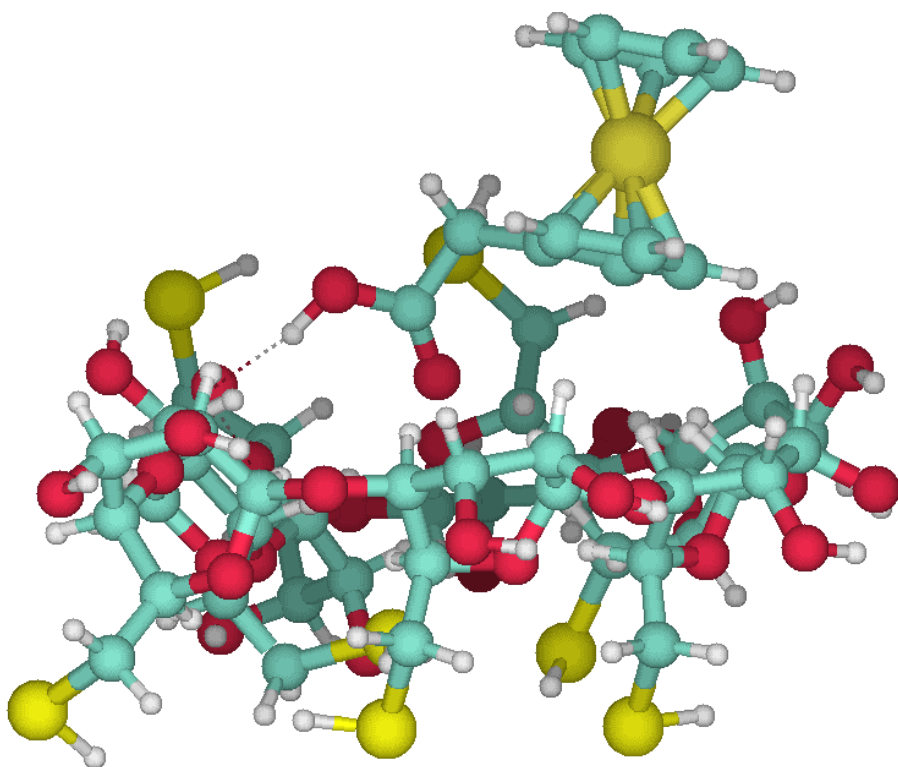Figure 5.44. Initial CD-AD geometry for the 4th B3LYP full optimization
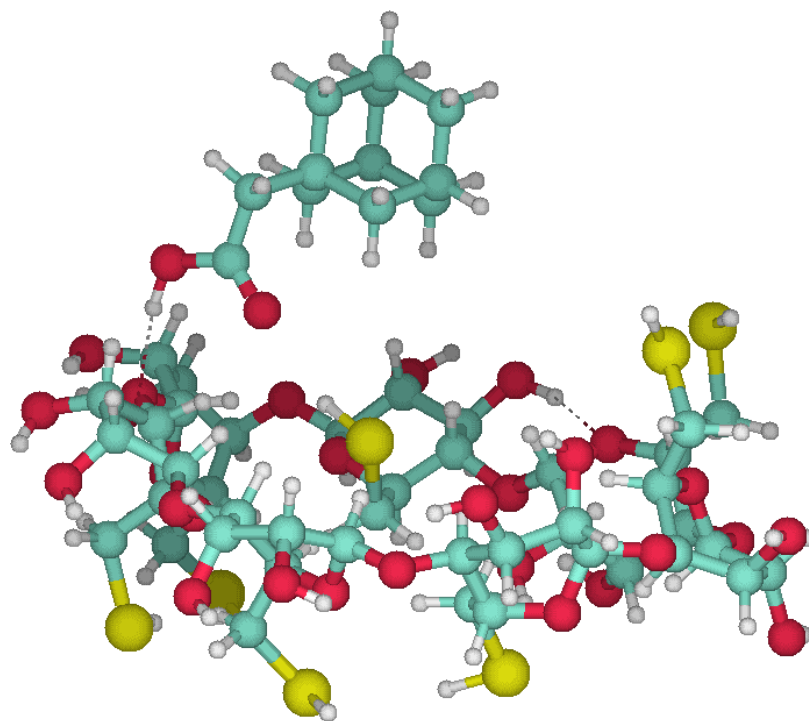
Figure 5.45. The optimized CD-AD geometry from the 4th B3LYP optimization
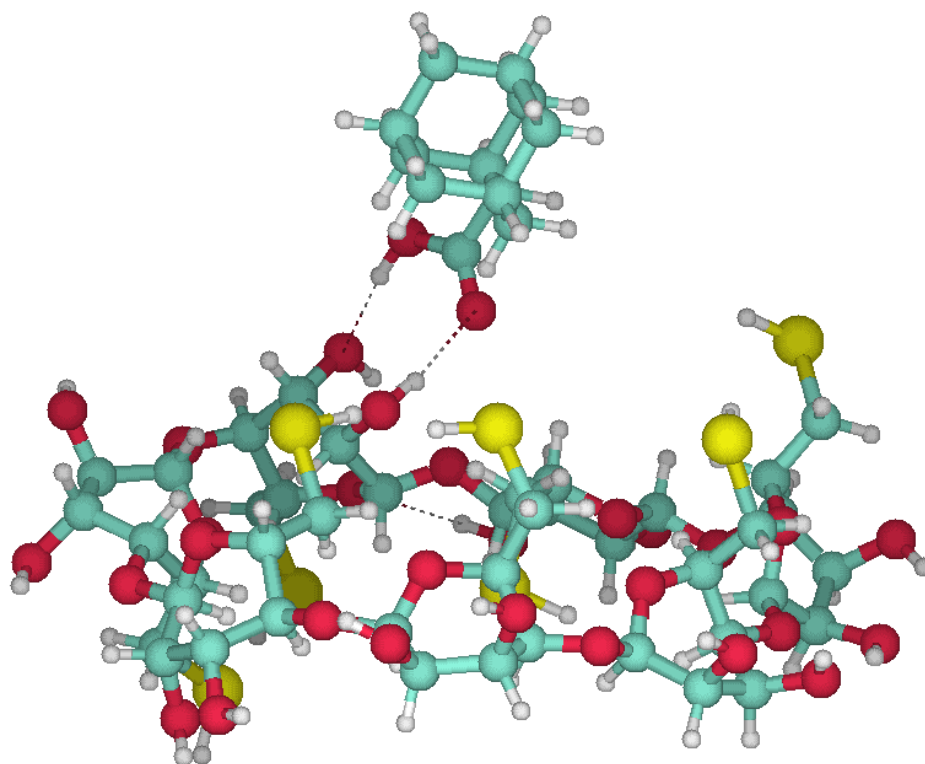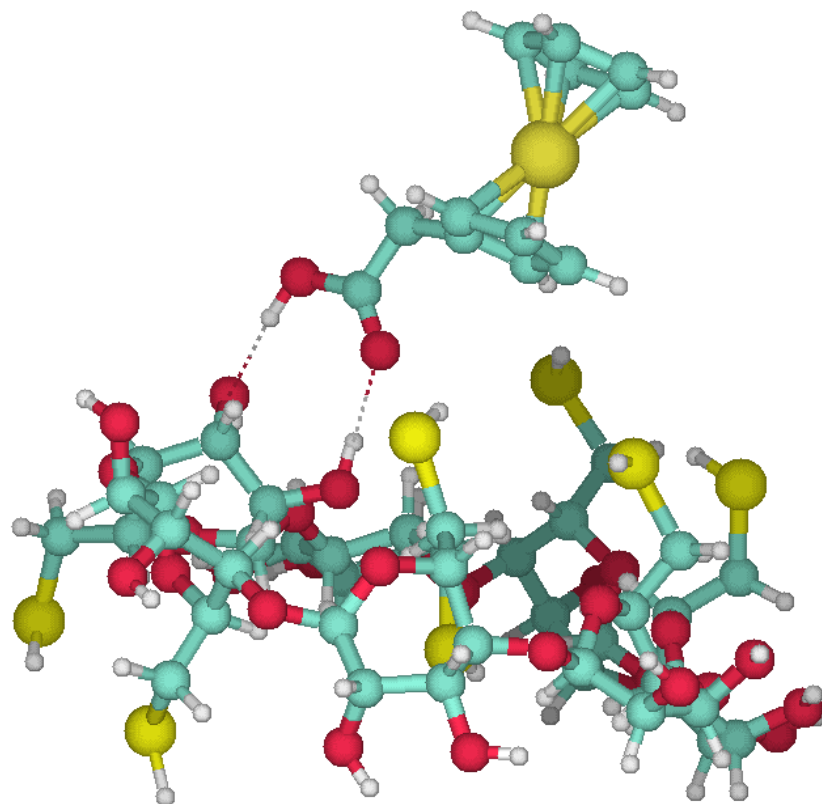

Figure 5.46. Initial CD-AD geometry for the 5th B3LYP optimization
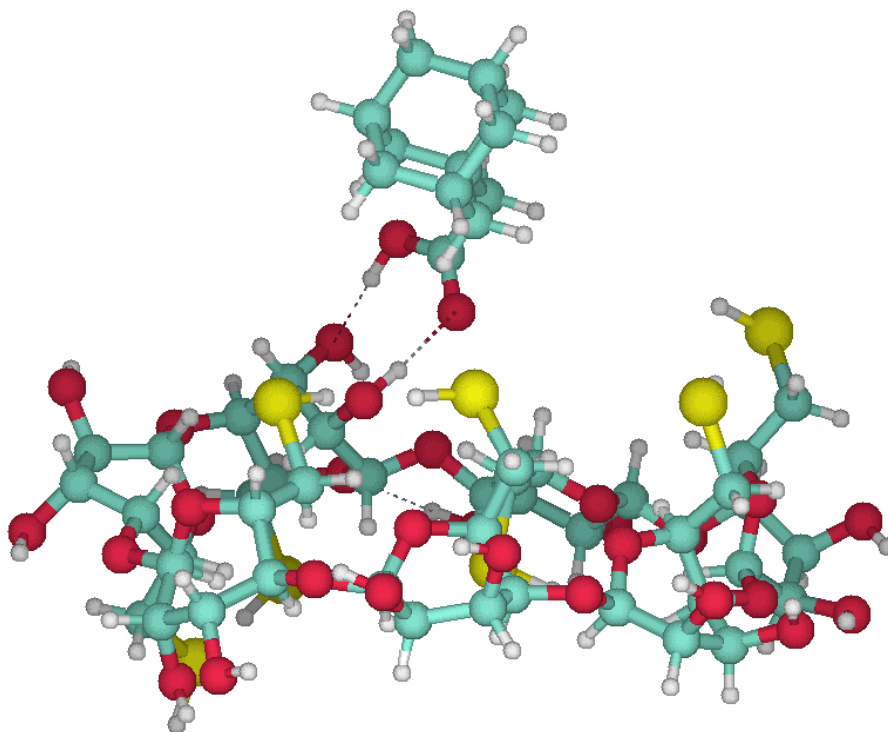
Figure 5.47. The optimized CD-AD geometry from the 5th B3LYP full optimization
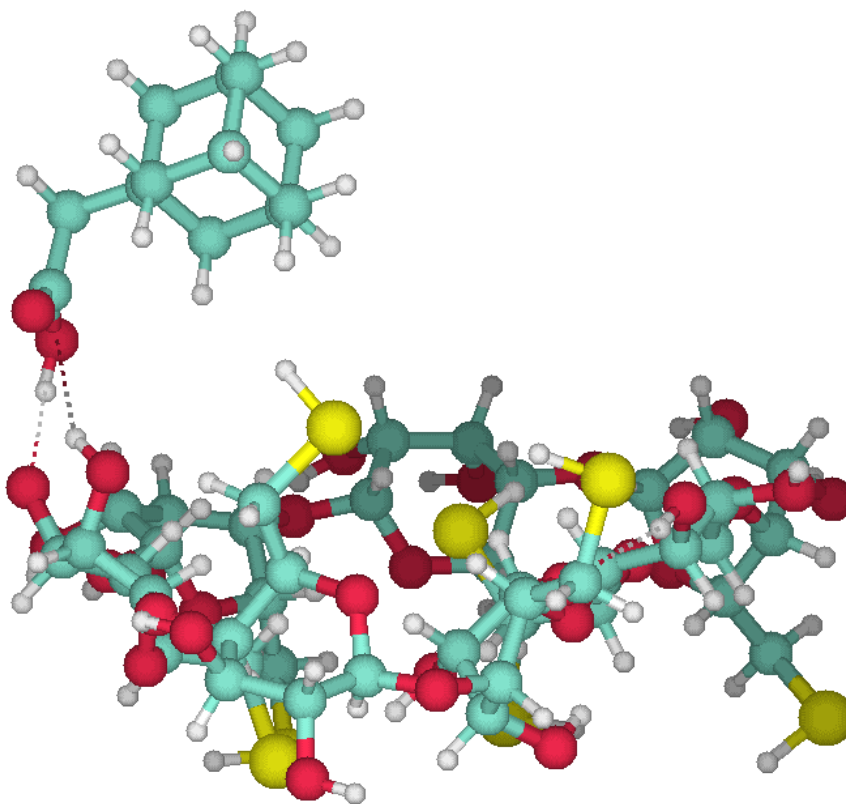
Figure 5.48. Initial CD-AD geometry for the 6th B3LYP full optimization
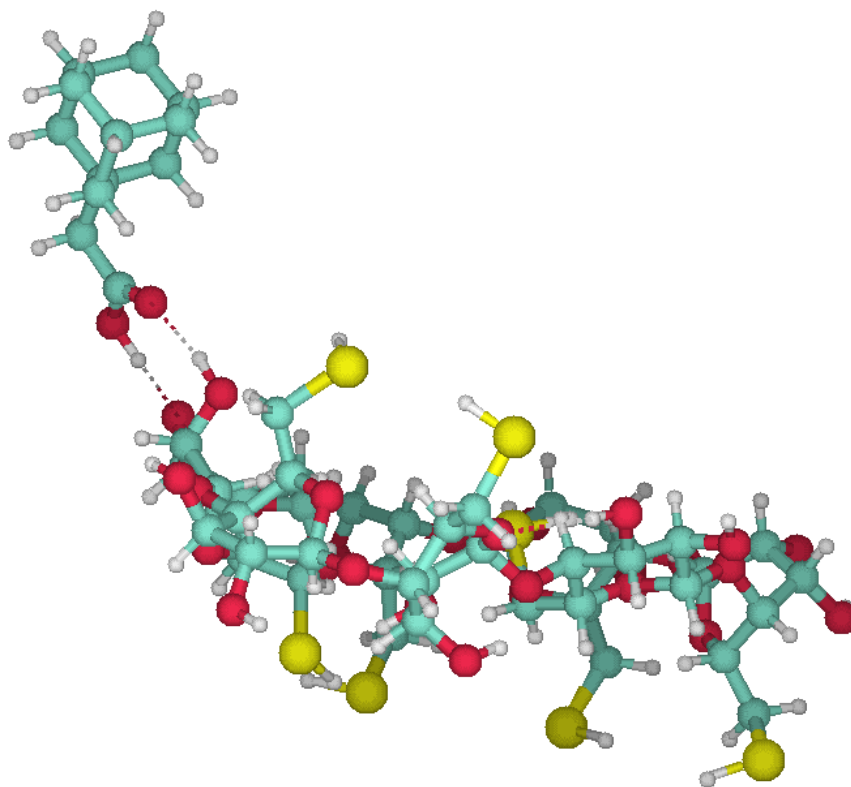
Figure 5.49. The optimized CD-AD geometry from the 6th B3LYP full optimization
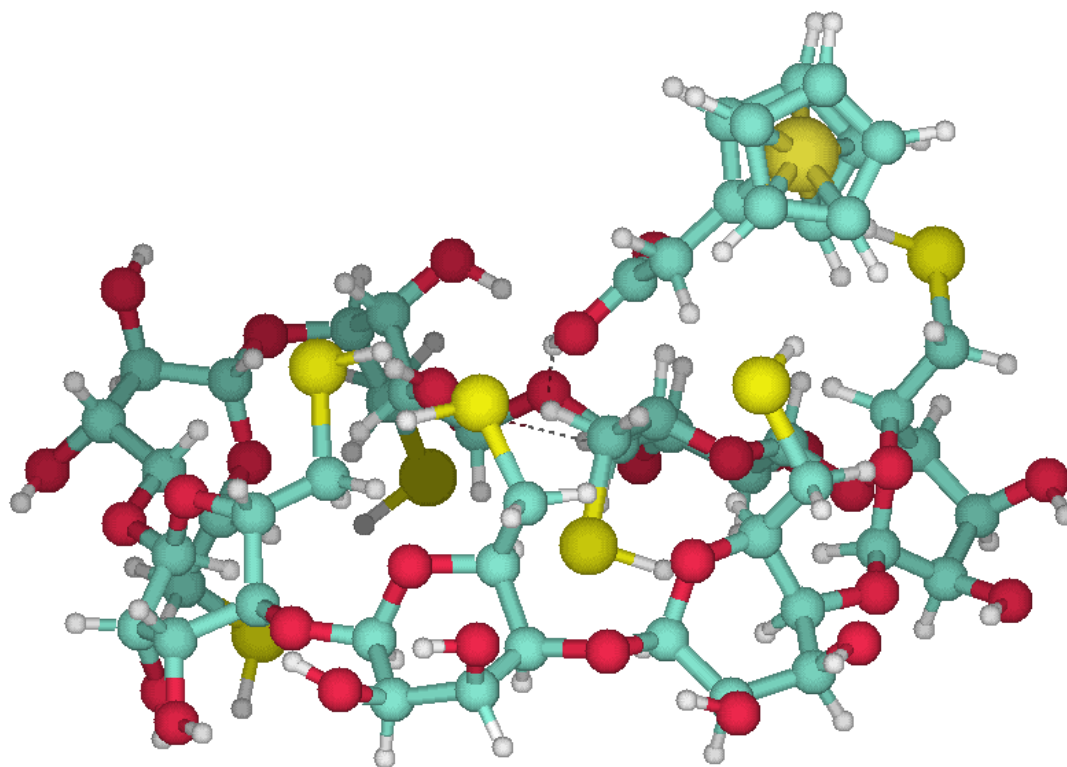
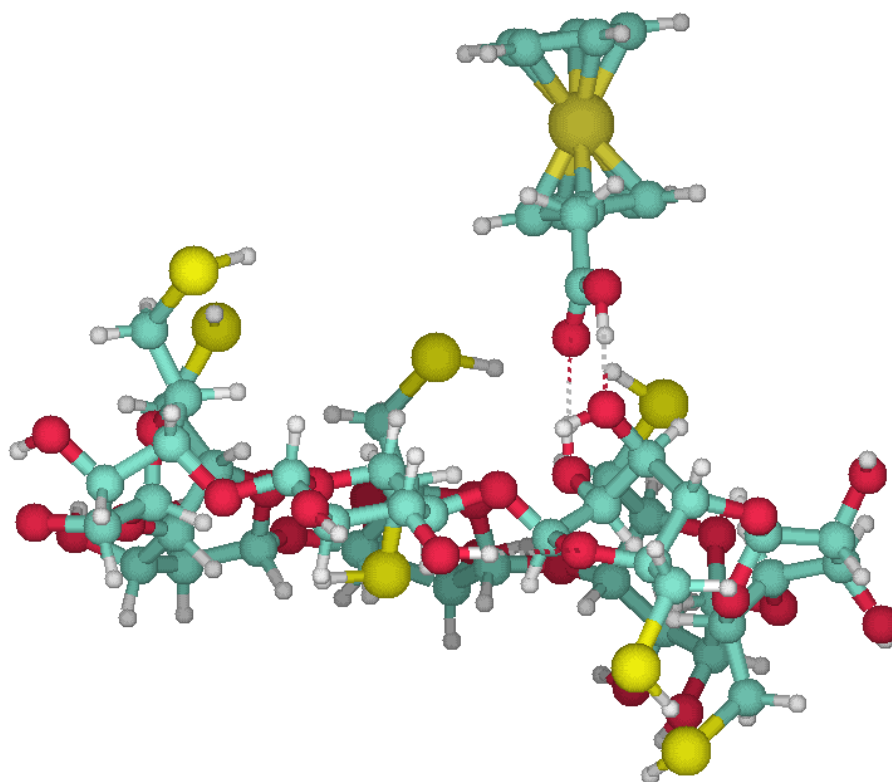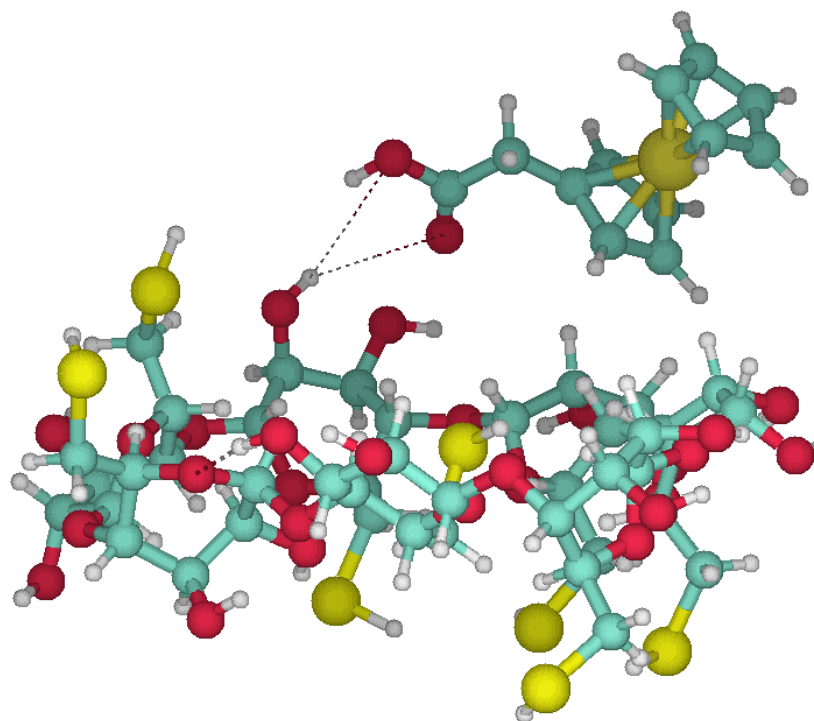Figure 5.50. Initial CD-FC geometry for the 1st B3LYP full optimization
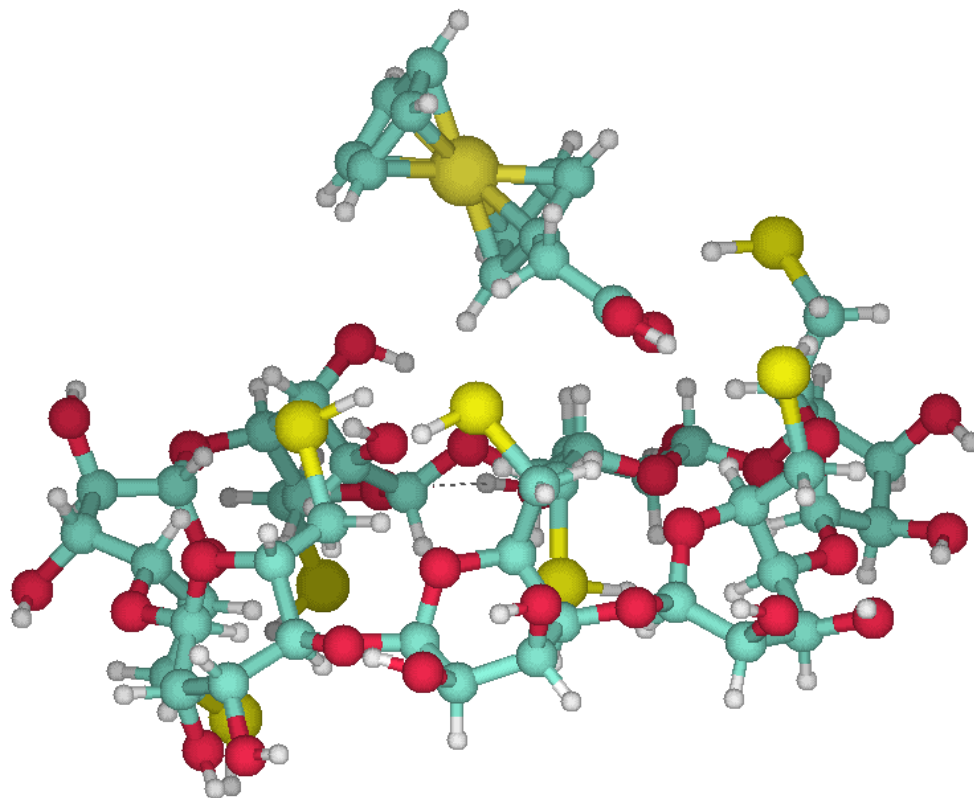
Figure 5.51. The optimized CD-FC geometry from the 1st B3LYP optimization
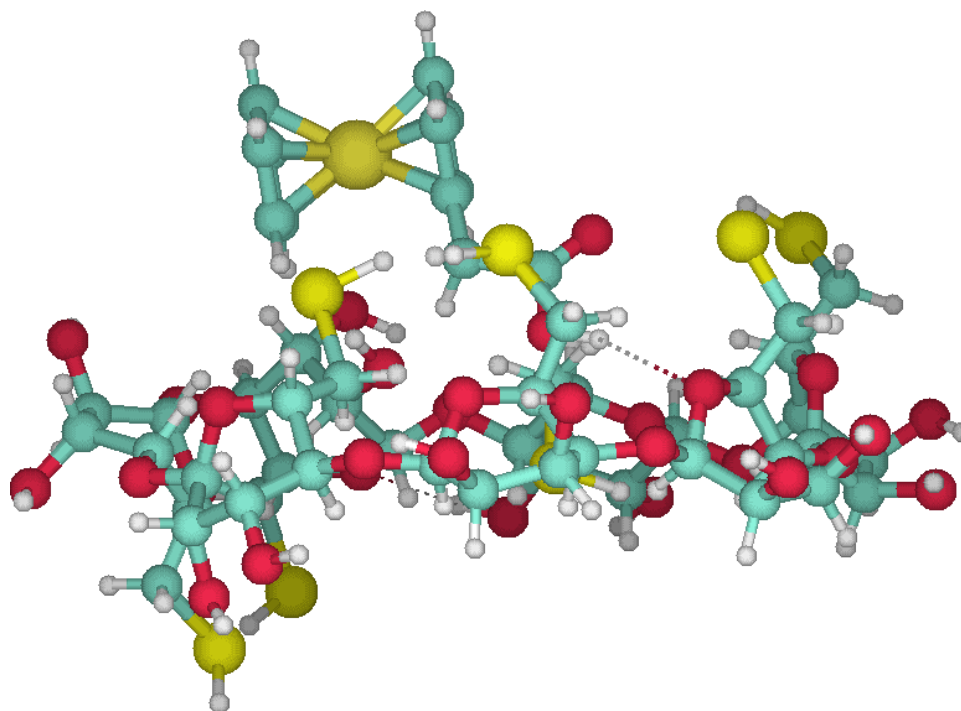
Figure 5.52. Initial CD-FC geometry for the 2nd B3LYP optimization
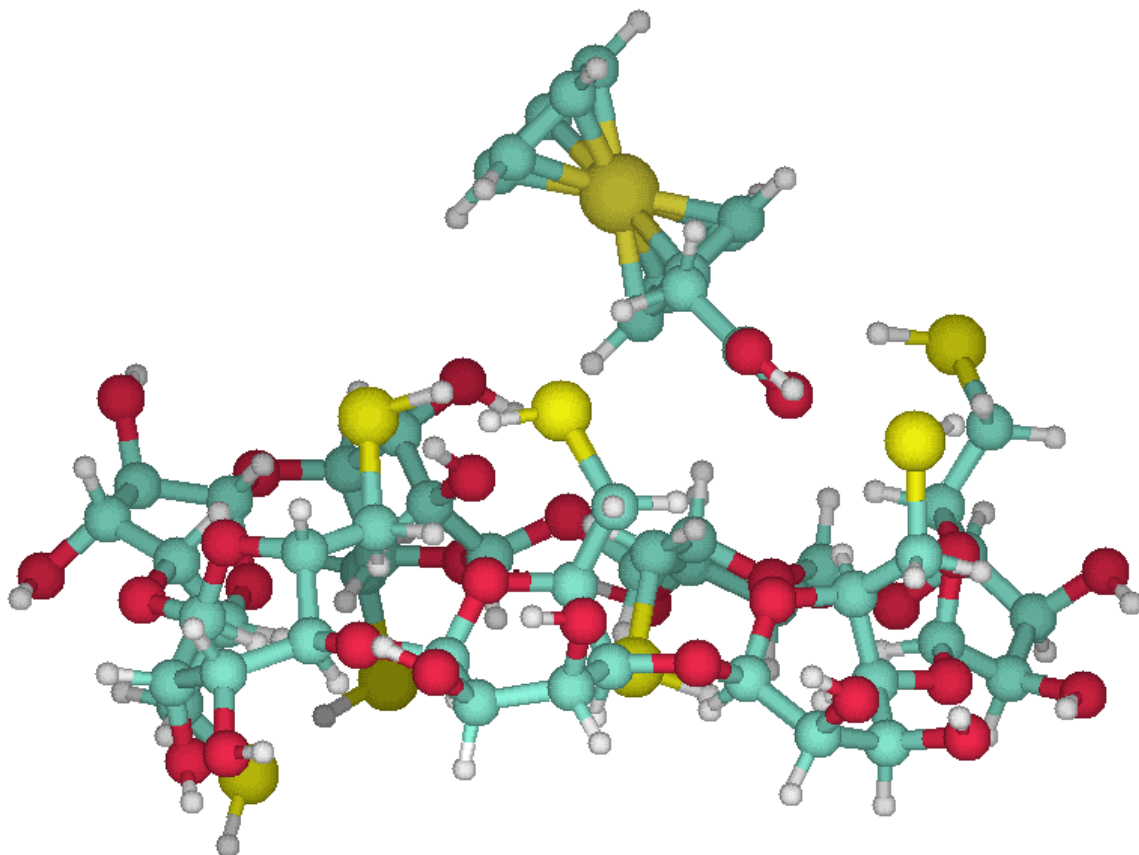
Figure 5.53. The optimized CD-FC geometry from the 2nd B3LYP full optimization

Figure 5.54. Initial CD-FC geometry for the 3rd B3LYP full optimization

Figure 5.55. The optimized CD-FC geometry from the 3rd B3LYP full optimization

Figure 5.56. Initial CD-FC geometry for the 4th B3LYP full optimization
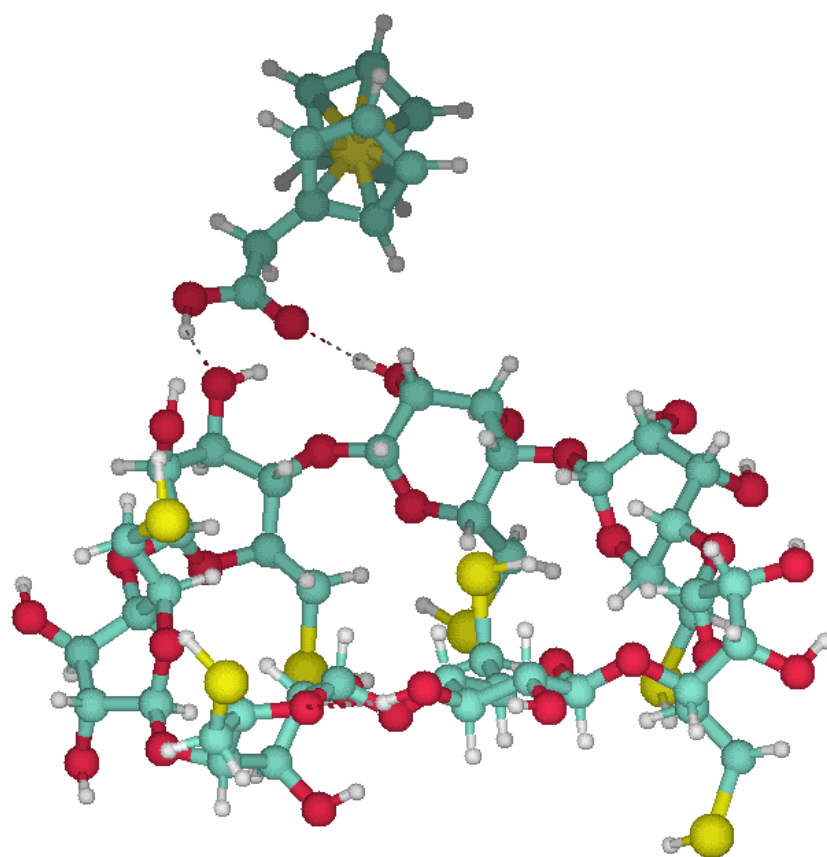
Figure 5.57. The optimized CD-FC geometry from the 4th B3LYP full optimization

Figure 5.58. Initial CD-FC geometry for the 5th B3LYP full optimization

136

Figure 5.59. The optimized CD-FC geometry from 5th B3LYP full optimization

Figure 5.60. Initial CD-FC geometry for the 6th B3LYP full optimization

137

Figure 5.61. The optimized CD-FC geometry from the 6th B3LYP full optimization
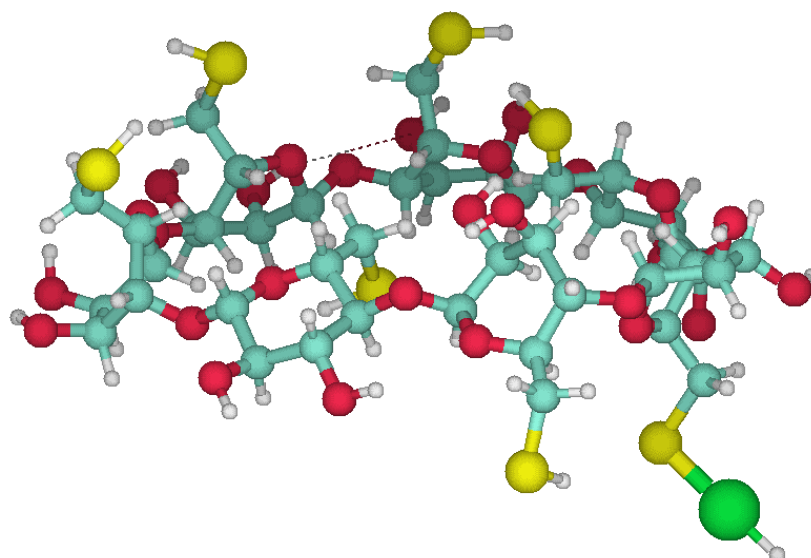

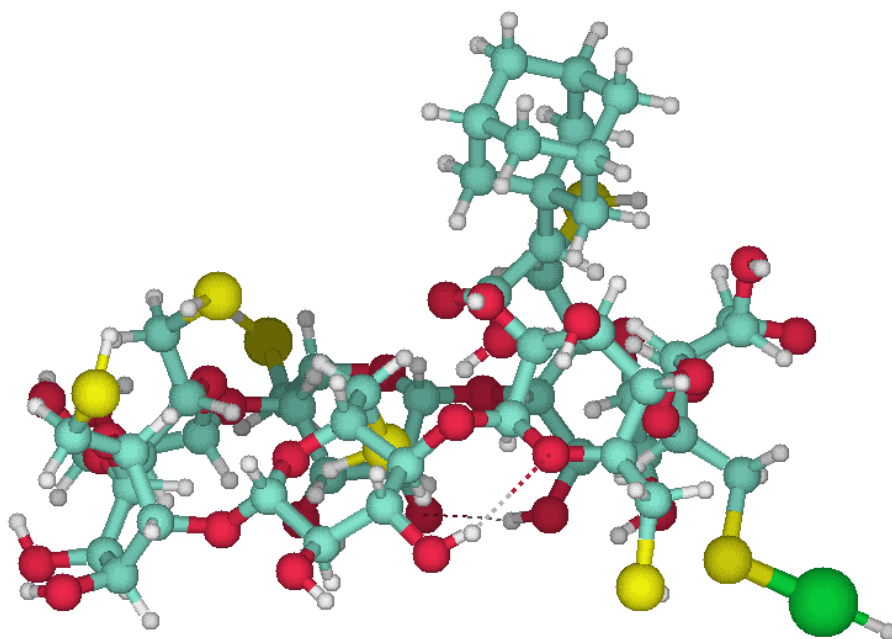Figure 5.62. Structure of the model of a Zn nano-particle attached to CD

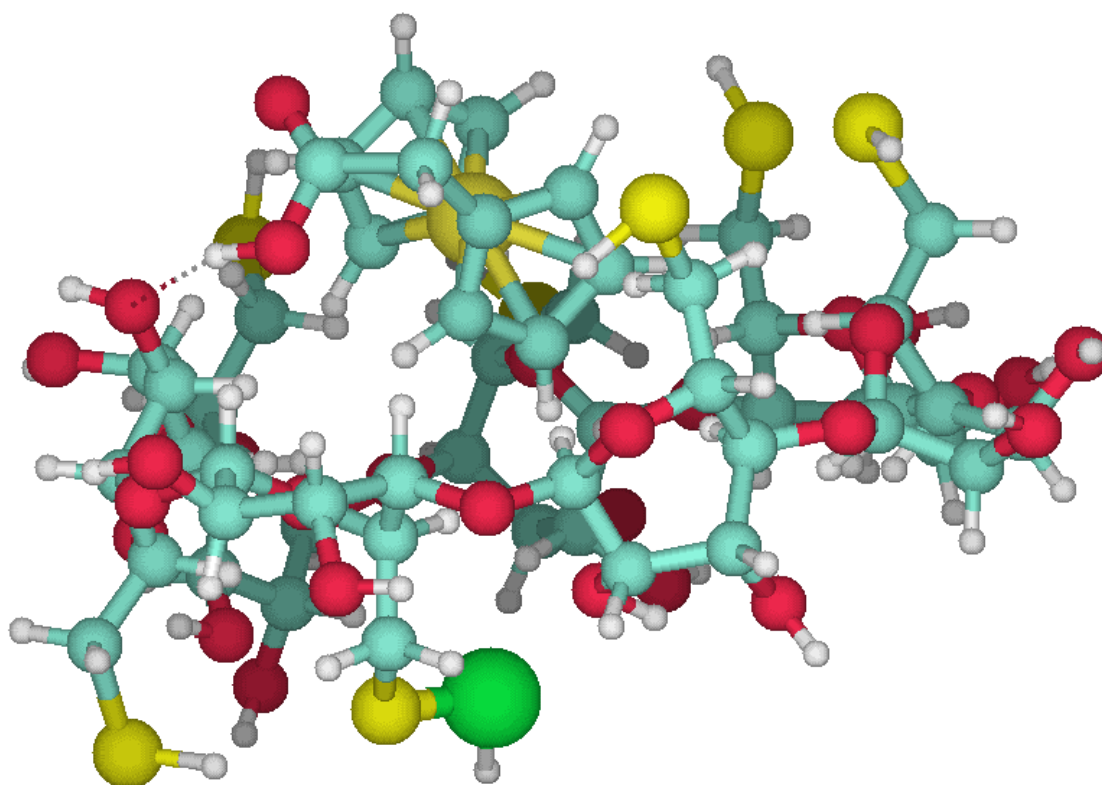Figure 5.63. Structure of the model of a Zn nanoparticle attached to CD-AD

Figure 5.64. Structure of the model of a Zn nanoparticle attached to CD-FC

References

1. Palaniappan, K.; Hackney, S. and Liu, J. *Chem.commun.* 2004, 2704-2705.

2. Liu, J.; Ong, W.; Román, E.; Lynn, M.; Kafer, A., *Langmuir*, 2000, **16**, 3000-3002.

3. Liu, J.; Alvarez, J.; Ong, W.; Román, E.; Kafer, A., *J. Am. Chem. Soc.*, 2001, **123**, 11148-11154.

4. Lipkowitz, K., *Chem. Rev.* 1998, **98**, 1829-1873.

5. Venanzi, C. A.; Canzius, P. M.; Zhang, Z.; Bunce, J. D., *J. Comput. Chem.* 1989, **10,** 1038.

6. Kostense, A. S.; van Helden, S. P.; Janssen, L. H. M., *J. Computer-Aided Mol. Design.* 1991, **5,** 525.

7. Lipkowitz, K. B., *J. Org. Chem.* 1991, **56**, 6357.

8. Lipkowitz, K. B.; Green, K.; Yang, J. A., *Chirality.* 1992, **4,** 205.

9. Lichtenthaler, F. W.; Immel, S., *Tetrahedron: Asymmetry.* 1994, **5**, 2045.

10. Immel, S.; Lichtenthaler, F. W., *Starch/Staerke.* 1996, **48**, 225.

11. Reinhardt, R.; Richter, M.; Mager, P. P., *Carbohydr. Res.* 1996, **291**, 1.

12. Rekharsky, M.; Inoue, Y., *Chem. Rev.* 1998, **98**, 1875-1917.

13. Lukovits, I., *J. Mol. Struct. (THEOCHEM)* 1988, **170**, 249.

14. Madrid, J.M.; Pozuelo, J.; Mendicuti, F.; Mattice, W. L., *J. Colloid Interface Sci.* 1997, **193**, 112.

15. Lü, T.-X.; Zhang, D.-B.; Dong, S-J., *J. Chem. Soc., Faraday Trans. 2* 1989, **85**, 1439.

16. Ohashi, M.; Kasatani, K.; Shinohara, H.; Sato, H., *J. Am. Chem. Soc.* 1990, **112,** 5824.

17. Kuroda, Y.; Ito, M.; Sera, T.; Ogoshi, H., *J. Am. Chem. Soc.* 1993, **115**, 7003.

18. Lipkowitz, K. B.; Raghothama, S.; Yang, J.-A., *J. Am. Chem. Soc.* 1992, **114,** 1554.

19. Berg, U.; Gustavsson, M.; Åström, N, *J. Am. Chem. Soc*. 1995, **117**, 2114.

20. Marconi, G.; Monti, S.; Mayer, B.; Köhler, G., *J. Phys. Chem*. 1995, **99**, 3943.

21. Jursic, B.; Zdravkovski, Z.; French, A. D., *J. Mol. Struct. (THEOCHEM)* 1996, **366**, 113.

22. Matsui, Y., *Bull. Chem. Soc. Jpn*. 1982, **55,** 1246.

23. Kitagawa, M.; Hoshi, H.; Sakurai, M.; Inoue, Y.; Chûjô, R., *Bull. Chem. Soc. Jpn*. 1988, **61**, 4225.

24. Rüdiger, V.; Eliseev, A.; Simova, S.; Schneider, H.-J.; Blandamer, M. J.; Cullis, P. M.; Meyer, A. J., *J. Chem. Soc., Perkin Trans. 2* 1996, 2119.

25. Pang, L.; Whitehead, M. A., *Supramol. Chem*. 1992, **1**, 81.

26. Tran, V.; Delage, M. M.; Buléon, A., *J. Inclusion Phenom. Mol. Recognit. Chem*. 1992, **14,** 271.

27. Myles, A. M. C.; Barlow, D. J.; France, G.; Lawrence, M. J., *Biochim. Biophys. Acta* 1994, 1199, 27.

28. Ohashi, M.; Kasatani, K.; Shinohara, H.; Sato, H., *J. Am. Chem. Soc.* 1990, 112, 5824.

29. Jaime, C.; Redondo, J.; Sánchez-Ferrando, F.; Virgili, A., *J. Org. Chem*. 1990, **55**, 4772.

30. Ivanov, P. M.; Salvatierra, D.; Jaime, C., *J. Org. Chem*. 1996, **61**, 7012.

31. Estrada, E.; Perdomo-López.; Torres-Labandeira, J., *J. Org. Chem*. 2000, **65**, 8510.

32. Varady, J.; Wu, X.; Wang, S., *J. Phys. Chem. B* 2002, **106**, 4863.

33. Consonni, R.; Recca, R.; Dettori, M. A.; Fabbri, D.; Delogu, G., *J. Agric. Food Chem*. 2004, **52**, 1590.

34. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 366.

35. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 372-375.

36. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 363.

37. Roothaan, C. C. J., *Rev. Mod. Phys.* 1951, **23**, 69..

38. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 110.

39. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 127-128.

40. Schiff, L. I., *Quantum Mechanics*, Third Edition, McGraw-Hill: USA, 1968 p 255.

41. Slater, J. C., *Phys. Rev.*, 1930, **36**, 57.

42. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004 p 134.

43. Mulliken, R. S., Rieke, C. A., and Orloff, H., *J. Chem. Phys.*, 1949, **17**, 1248.

44. Bishop, D. M., *J. Chem. Phys.*, 1966, **45**, 1880.

45. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 168.

46. Heher, W. J., Stewart, R. F., and Pople, J. A., *J. Chem. Phys.*, 1969, **51**, 2657.

47. Møller, C. and Plesset, M. S., *Phys. Rev.*, 1934, **46**, 359.

48. Schiff, L. I., *Quantum Mechanics*, Third Edition, McGraw-Hill: USA, 1968, p 245.

49. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 219-222.

50. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p249-254.

51. Hohenberg, P., Hohn, W., *Phys. Rev.*, 1964, **136**, B864.

52. Kohn, W., Sham, L. J., *Phys. Rev.*, 1965, **140**, A1133.

53. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 255.

54. Cramer, C. J., *Essentials of Computational Chemistry: Theories and Models*, Second Edition, John Wiley & Sons Ltd: Chichester, England, 2004, p 267.

55. Miertuš, S., Scrocco, E., Tomasi, J., *Chem. Phys*., 1981, **55**, 117.

56. Ridley, J., Zerner, M., *Theoret. Chim. Acta.,* 1973, **32**, 111.

57. Bacon, A., Zerner, M., *Theoret. Chim. Acta.,* 1979, **53**, 21.

58. Ridley, J., Zerner, M., *Theoret. Chim. Acta.,* 1976, **42**, 223.