



universität  
wien

# DISSERTATION

Titel der Dissertation

**„Molecular Modelling on  
Cyclodextrin Inclusion Complexes“**

Verfasser

Mag. Walter Snor

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr. rer.nat.)

Wien, 2009

Studienkennzahl lt. Studienblatt: A 091 419

Dissertationsgebiet lt. Studienblatt: 419 Chemie

Betreuer: Univ.-Prof. Dr. Peter Wolschann



## Table of Contents

I.	Danksagung.....	IV
II.	Abstract .....	V
III.	Zusammenfassung .....	VII
1	General Introduction.....	1
2	Methods and Programs .....	5
2.1	Molecular Modelling and Computational Chemistry.....	5
2.1.1	Overview of Computational Chemistry Methods .....	7
2.1.1.1	Molecular Mechanics (MM) .....	7
2.1.1.2	Ab Initio Electronic Structure Methods .....	10
2.1.1.3	Basic Quantum Mechanics (QM).....	11
2.1.1.4	The Hartree-Fock (HF) Self-Consistent Field Approximation.....	13
2.1.1.5	Electron Correlation .....	15
2.1.1.6	Density Functional Theory (DFT).....	16
2.1.1.7	Semi-empirical Methods.....	19
2.1.1.8	Basis Sets.....	21
3	Results and Discussion – Part 1 .....	26
	Consideration of the Basis Set Superposition Error Behaviour of a Hydrogen-Bonded Water-Methanol Complex at Several Levels of Theory .....	26
4	How to Conduct a Computational Molecular Modelling Study ....	29
4.1	Basic Considerations.....	29
4.2	Computer Graphics .....	33
4.2.1	Basic Considerations.....	33
4.2.2	Visualization of Molecular Models.....	33
4.2.3	Visualization of Molecular Properties.....	36
4.3	Geometry Optimization (Energy Minimization).....	38
4.3.1	The Molecular Potential Energy Surface (PES).....	38
4.3.2	Characterization of Stationary Points.....	40
4.3.3	Energy Minimization Methods .....	41
4.3.4	Conformational Searching.....	49
4.3.4.1	Systematic Search Procedure .....	49
4.3.4.2	Monte Carlo Methods (MC).....	50
4.3.4.3	Molecular Dynamics (MD).....	51
4.3.4.4	Simulated Annealing .....	52
4.4	Choosing the Best Computational Method.....	52
4.4.1	Model Chemistries.....	53
4.4.2	Selecting a Suitable Calculation Method.....	54
4.4.2.1	Semi-Empirical Methods .....	54
4.4.2.2	HF Theory.....	55
4.4.2.3	Electron Correlation and Post-SCF Methods .....	55
4.4.2.4	DFT Methods.....	56
4.4.2.5	QM/MM .....	56
4.4.3	Modelling Systems in Solution.....	56
5	Results and Discussion – Part 2 .....	59

---

Systematic Investigation of the Dependence of Geometric Parameters of Malonaldehyde on the Level of Theory Used.....	59	
6	Host-Guest-Chemistry.....	61
6.1	Classification of Host-Guest Compounds.....	63
6.2	The Importance of Preorganisation and Complementarity.....	64
6.3	Molecular Recognition.....	65
6.4	Chiral Recognition.....	67
7	Results and Discussion – Part 3.....	69
$\beta$ -CD as a Chiral Selector between D-Alanine and L-Alanine.....	69	
8	The Nature of Supramolecular Interactions.....	71
8.1	Basic Considerations.....	71
8.2	The Hydrogen Bond.....	73
8.2.1	Basic Considerations.....	73
8.2.2	$\sigma$ -Bond Cooperativity.....	74
8.2.3	$\pi$ -Bond Cooperativity.....	75
8.2.4	The Anticooperativity Effect; Homodromic and Antidromic Hydrogen Bonds.....	76
8.2.5	Flip-Flop Hydrogen Bonds.....	77
9	CDs as Supramolecular Hosts.....	80
9.1	Basic Considerations.....	80
9.2	Structural Features of CDs.....	81
9.3	Inclusion Chemistry of the CDs.....	87
10	Results and Discussion – Part 4.....	97
Quantum Mechanical Calculations on $\beta$ -CD Inclusion Complexes with Neutral and Anionic Conformations of Meloxicam.....	97	
11	Results and Discussion – Part 5.....	115
Papers I-V.....	115	
11.1	Paper I: “Homodromic Hydrogen Bonds in Low-Energy Conformations of Single Molecule CDs”.....	116
11.2	Paper II: “On the Structure of Anhydrous $\beta$ -CD”.....	121
11.3	Paper III: “Density Functional Calculations on CDs”.....	126
11.4	Paper IV: “Molecular Dynamics Simulations and Quantum Chemical Calculations on $\beta$ -CD-Spirolactone Complex”.....	136
11.5	(Paper V: “Density Functional Calculations on Meloxicam- $\beta$ -CD Inclusion Complexes”).....	142
12	Conclusion and Future Directions.....	174
13	Abbreviations.....	176
14	References.....	179
15	Appendix A.....	192
15.1	Molecule editors.....	192
15.1.1	ChemDraw.....	192
15.1.2	ISIS/Draw.....	193
15.2	Molecule Viewers.....	193
15.2.1	Discovery Studio Viewer Pro (formerly WebLabViewer).....	193
15.2.2	GaussView.....	194

---

15.2.3	HyperChem.....	194
15.2.4	Chem3D.....	195
15.3	Ab initio and DFT Software.....	197
15.3.1	Gaussian .....	197
15.3.2	HyperChem.....	197
16	List of publications and contribution to conferences.....	198
17	Curriculum Vitae.....	200

## I. Danksagung

Ich möchte vor allem meinem Betreuer, Herrn Univ.-Prof. Dr. Peter Wolschann, sehr herzlich danken für die interessante Aufgabenstellung sowie für die umfassende fachliche Betreuung und Unterstützung im Laufe dieser Arbeit. Seine ruhige und verständnisvolle Art, seine stete Diskussions- und Hilfsbereitschaft sowie sein großes Interesse am Fortgang meiner Arbeit waren für mich immer wieder Motivation und Ansporn zugleich.

Weiters ist es mir ein aufrichtiges Bedürfnis, Herrn Univ.-Prof. Dr. Helmut Viernstein für sein stets außerordentlich freundliches Entgegenkommen sowie seine Hilfe und Unterstützung bei all meinen Anliegen ganz herzlich zu danken.

Ein spezieller Dank ergeht auch an Herrn Univ.-Prof. Dr. Alfred Karpfen, der für meine Fragen und Probleme jederzeit ein offenes Ohr hatte und mir immer wieder mit Rat und Tat zur Seite stand.

Ein herzliches Dankeschön sage ich auch meinen Arbeitskollegen am Institut, allen voran Dr. Petra Weiss-Greiler, sowie Dr. Anton Beyer, Dr. Anna Stary, Mag. Marcus Seibold, Martina Ziehengraser, Victoria Stummer und Elisabeth Liedl für das angenehme Arbeitsklima und die Hilfe bei verschiedenen Problemen.

Ganz besonderen Dank aber schulde ich meiner Frau Regina und meiner Tochter Andrea für ihre Geduld und ihr Verständnis dafür, dass ihr Ehemann und Vater in dieser Zeit arbeitsbedingt sehr oft "in einer anderen Welt" weilte. Sehr herzlich möchte ich mich überdies bei meiner Frau für ihre Hilfe und Unterstützung bei Problemen im Hinblick auf die englische Grammatik bedanken.

*Meiner Frau Regina, meiner Tochter Andrea,  
meiner Enkelin Theresa und meinem  
Schwiegersohn Martin gewidmet*

\*\*\*

## II. Abstract

The main purpose of this thesis was to perform computational investigations on Cyclodextrins (CD) and CD complexes using ab initio and Density Functional Theory (DFT) methods at an appropriate high level of theory, in order to get better insight into the structures and energetic of these compounds.

After an introductory consideration of molecular modelling and computational chemistry (chapter 1), a short overview of the computational methods used in this thesis as well as a brief consideration of basis sets and its effects is given in chapter 2.

An investigation of the impact of the theory used on the magnitude of the BSSE correction energy on the calculated energies of the model system water-methanol-dimer was performed in chapter 3. It could be shown that as the basis set is improved the BSSE energy is reduced.

Chapter 4 provides guidelines on how to conduct properly a computational molecular modelling study.

In chapter 5, systematic investigations on the dependence of several geometric parameters of malonaldehyde on the level of theory used led to the outcome that optimizations with DFT B3LYP/6-31G(d) and MP2/6-311+G(d,p) model chemistries yield a resulting geometry which is in excellent agreement with the experimental findings.

Basic considerations of supramolecular chemistry including topics such as host-guest chemistry, molecular recognition, and chiral recognition are given in chapter 6.

The case study in chapter 7 with the aim to establish whether  $\beta$ -CD can act as a chiral selector between the two enantiomeric forms of the amino acid alanine revealed that the D-alanine complex is slightly more stable than the L-alanine complex. Thus, because of these findings it is confirmed that  $\beta$ -CD can act as a chiral selector between the two enantiomeric forms of alanine.

A review of the various molecular interactions in supramolecular systems with the focal point on hydrogen bonding is given in chapter 8.

Chapter 9 presents the CDs as supramolecular hosts, including structural features and inclusion chemistry of this very important group of compounds.

Because of its extraordinary importance as an effective COX-2 inhibitor in medicinal and pharmaceutical applications, inclusion complexes of several neutral and anionic conformations of meloxicam with  $\beta$ -CD were investigated using DFT B3LYP/6-31G(d,p) calculations (chapter 10). As a main result, it was found that hydrogen bonding interactions are involved in the complexation of both the neutral and anionic conformations and support considerably the stabilization of the respective complexes.

An overview of chemical structure drawing and viewing software as well as ab initio and DFT program packages that have been used in this thesis is given in Appendix A.

The results of all these investigations have been summarized in the following papers:

- I. **Homodromic Hydrogen Bonds in Low-Energy Conformations of Single Molecule CDs** (Karpfen A, Liedl E, Snor W, Wolschann P, 2007, *J Incl Phenom Macro* 57, 35)
- II. **On the Structure of Anhydrous  $\beta$ -CD** (Snor W, Liedl E, Weiss-Greiler P, Karpfen A, Viernstein H, Wolschann P, 2007, *Chem Phys Lett*, 441, 159)
- III. **Density Functional Calculations on CDs** (Karpfen A, Liedl E, Snor W, Viernstein H, Weiss-Greiler P, 2008, *Monatsh Chem* 139, 363)
- IV. **Molecular Dynamics Simulations and Quantum Chemical Calculations on  $\beta$ -CD-Spirolactone Complex** (Weinzinger P, Weiss-Greiler P, Snor W, Viernstein H, Wolschann P, 2007, *J Incl Phenom Macro*, 57, 29)
- V. **Density functional calculations on meloxicam- $\beta$ -cyclodextrin Inclusion Complexes** (Snor W, Liedl E, Weiss-Greiler P, Viernstein H, Wolschann P, 2008, doi:10.1016/j.i.jpharm.2009.05.012, *Int J Pharm*)

### III. Zusammenfassung

Hauptziel dieser Dissertation ist die Untersuchung von Cyclodextrinen (CD) und CD-Komplexen mit ab initio- und Dichtefunktionaltheorie (DFT)-Rechnungen auf einem entsprechend hohen theoretischen Niveau, um damit einen besseren Einblick in Struktur und Energetik dieser Verbindungen zu erhalten.

Nach einführenden Überlegungen zu Molecular Modelling und Computer-Chemie (Kapitel 1) wird in Kapitel 2 ein Überblick über die in dieser Dissertation verwendeten Rechenmethoden sowie eine kurze Einführung in Basissätze und ihre Effekte gegeben.

Die Fallstudie in Kapitel 3 untersucht den Einfluss der verwendeten Theorie auf die Größe der BSSE Korrektur der berechneten Interaktionsenergien im Modellsystem Wasser-Methanol. Es wird gezeigt, dass die BSSE-Korrektur-Energie mit zunehmender Verbesserung der Basissätze abnimmt.

Kapitel 4 enthält Richtlinien für eine zweckmäßige Durchführung von Molecular Modelling-Studien.

In Kapitel 5 führt eine systematische Untersuchung der Abhängigkeit verschiedener geometrischer Parameter des Malonaldehyds von der verwendeten Theorie zum Ergebnis, dass Optimierungen mit DFT B3LYP/6-31G(d) und MP2/6-311+G(d,p) Molekülgeometrien ergeben, die in ausgezeichneter Übereinstimmung mit den experimentellen Ergebnissen stehen.

Grundlegende Überlegungen zur Supramolekularen Chemie, insbesondere der Wirt-Gast-Beziehungen sowie der molekularen und chiralen Erkennung bilden den Inhalt von Kapitel 6.

Die Fallstudie in Kapitel 7 untersucht die Frage, ob  $\beta$ -CD als chiraler Selektor zwischen den beiden enantiomeren Formen der Aminosäure Alanin fungieren kann. Dabei zeigt sich, dass der D-Alanin Komplex etwas stabiler ist als der L-Alanin Komplex, womit die obige Annahme verifiziert wird.

Kapitel 8 gibt einen ausführlichen Überblick über die verschiedenen Wechselwirkungskräfte in supramolekularen Systemen mit dem Schwerpunkt auf der Wasserstoffbrückenbindung.

CDs zählen zu den wichtigsten Wirtsverbindungen in der supramolekularen Chemie. Kapitel 9 behandelt ihre strukturellen Aspekte und die Mechanismen bei der Komplexbildung.

Wegen seiner außerordentlichen Bedeutung als effektiver COX-2 Hemmstoff bei medizinischen und pharmazeutischen Anwendungen werden in Kapitel 10 Komplexe verschiedener neutraler und anionischer Konformationen von Meloxicam mit  $\beta$ -CD mittels DFT B3LYP/6-31G(d,p) Rechnungen untersucht. Als Hauptergebnis dieser Rechnungen wird gefunden, dass bei allen Komplexen Wasserstoffbrückenbindungen beteiligt sind und diese eine wichtige Rolle bei der Stabilisierung der jeweiligen Komplexe spielen.

Anhang A gibt einen Überblick über die in dieser Arbeit verwendeten Software Pakete zum Zeichnen und Betrachten von chemischen Strukturen sowie ab initio und DFT-Programme.

Die Ergebnisse der oben angeführten Untersuchungen wurden in den folgenden Veröffentlichungen zusammengefasst und angewendet:

- I. **Homodromic Hydrogen Bonds in Low-Energy Conformations of Single Molecule CDs** (Karpfen A, Liedl E, Snor W, Wolschann P, 2007, J Incl Phenom Macro 57, 35)
- II. **On the Structure of Anhydrous  $\beta$ -CD** (Snor W, Liedl E, Weiss-Greiler P, Karpfen A, Viernstein H, Wolschann P, 2007, Chem Phys Lett, 441, 159)
- III. **Density Functional Calculations on CDs** (Karpfen A, Liedl E, Snor W, Viernstein H, Weiss-Greiler P, 2008, Monatsh Chem 139, 363)
- IV. **Molecular Dynamics Simulations and Quantum Chemical Calculations on  $\beta$ -CD-Spironolactone Complex** (Weinzinger P, Weiss-Greiler P, Snor W, Viernstein H, Wolschann P, 2007, J Incl Phenom Macro, 57, 29)
- V. **Density functional calculations on meloxicam- $\beta$ -cyclodextrin Inclusion Complexes** (Snor W, Liedl E, Weiss-Greiler P, Viernstein H, Wolschann P, 2008, doi:10.1016/j.i.jpharm.2009.05.012, Int J Pharm)



# 1 General Introduction

**Molecular modelling** is centred on applying the fundamental laws of physics and chemistry to the study of molecules. Its ultimate aim is to create models and simulations, which can help by predicting, rationalizing, and estimating the properties of molecules and their interactions.

**Cyclodextrins** (CD) are macrocyclic oligomers of  $\alpha$ -D-glucose. They are shaped like truncated cones with primary and secondary hydroxyl groups at the narrower rim and wider rim, respectively. Three species of CDs are the most widely used: They have rings comprising six ( $\alpha$ -CD), seven ( $\beta$ -CD) and eight ( $\gamma$ -CD) glucose units. Because CDs have a hydrophilic exterior and a hydrophobic cavity of appropriate dimension, they can bind with various guest molecules to form **inclusion complexes**. This property has enabled CDs to be widely used in pharmaceutical science, catalysis, separation technology and other areas. Furthermore, the CD inclusion complexation has been considered an ideal model mimicking the enzyme-substrate interactions.

Because of the diverse applications of CDs, during the past several decades considerable efforts have been devoted to CD chemistry. Many experimental methods such as X-ray crystallography and NMR spectroscopy have been developed to study the complexation behaviours of native and functionalized CDs.

In order to get a better understanding of the binding events, a lot of theoretical methods including Molecular Mechanics (MM), Molecular Dynamics (MD), and more recently, Quantum Mechanical (QM) methods such as ab initio and Density Functional Theory (DFT), have also been used to study the CD complexes. All these experimental and theoretical methods, when properly utilized in combination with each other, have proven to be extremely powerful in solving the structural, energetic, and dynamic problems associated with CDs and CD complexes.

However, because the treatment of molecular systems as large as CDs and CD complexes by ab initio and DFT methods is extremely expensive computationally, comparatively few molecular calculations have been performed thus far based on more accurate ab initio and DFT methods on such systems.

Therefore, it was motivation and primary goal of this thesis to perform computational investigations on CDs and CD complexes using ab initio and DFT methods at an appropriate high level of theory, in order to get better insight into the structures and energetic of CDs and their complexes.

This thesis describes the determination of the geometries and energetic of CDs and CD inclusion complexes mainly by DFT calculations at the B3LYP/6-31 G(d,p) level of theory. This level of theory was chosen because it has at present been proven the best approach for investigations on such supramolecular systems.

This thesis is based on the following papers:

**I. Homodromic Hydrogen Bonds in Low-Energy Conformations of Single Molecule CDs** (Karpfen A, Liedl E, Snor W, Wolschann P, 2007, *J Incl Phenom Macro* 57, 35)

Low-energy conformations of  $\beta$ -CD under anhydrous conditions in the gas phase were investigated by ab initio HF/3-21G and HF/6-31G(d,p) as well as DFT B3LYP/6-31G(d,p) calculations, imposing  $C_7$  symmetry throughout. In these conformations, two homodromic hydrogen bond rings are formed with very short hydrogen bonds at the narrow side of the CD ring and a second one at the wider side. These hypothetical conformations are not comparable to experimentally studied open conformations, but their energy is significantly lower than those of the latter by an amount of  $\approx 10 \text{ kcal mol}^{-1}$

**II. On the Structure of Anhydrous  $\beta$ -CD** (Snor W, Liedl E, Weiss-Greiler P, Karpfen A, Viernstein H, Wolschann P, 2007, *Chem Phys Lett*, 441, 159)

Low-energy conformations of  $\beta$ -CD under anhydrous conditions in the gas phase were investigated by DFT B3LYP/6-31G(d,p) calculations, imposing  $C_7$  symmetry throughout. In one conformation, two homodromic hydrogen bond rings are formed with very short hydrogen bonds at the narrow side of the CD ring and a second one at the wider side. Four possibilities for the orientation of the homodromic rings have been taken into account. These “closed” conformations differ from conformations, which have been studied experimentally, forming inclusion complexes with small and medium-sized guest molecules. Their energy is significantly lower than those of the open conformations ( $\Delta E \approx 44 \text{ kJ mol}^{-1}$ ).

**III. Density Functional Calculations on CDs** (Karpfen A, Liedl E, Snor W, Viernstein H, Weiss-Greiler P, 2008, *Monatsh Chem* 139, 363)

Conformations of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs under anhydrous conditions in the gas phase were investigated by a DFT method, B3LYP/6-31G(d,p). These calculations resulted in several symmetric conformations with different energies. The lowest energy conformations contain two rings of homodromic hydrogen bonds and are referred to “one-gate-closed” conformations. Different orientations of hydrogen bonds lead to four minima. Other conformational minima were found for “open” conformations which correspond to some extent to experimentally determined structures.

**IV. Molecular Dynamics Simulations and Quantum Chemical Calculations on  $\beta$ -CD-Spirolactone Complex** (Weinzinger P, Weiss-Greiler P, Snor W, Viernstein H, Wolschann P, 2007, *J Incl Phenom Macro*, 57, 29)

MD simulations on  $\beta$ -CD in vacuo, with water and complexated with spironolactone were performed at a temperature of 300 K over a period of one ns. Two different orientations of SP in the cavity were considered. Along with conformational parameters, the formation of hydrogen bonds has been monitored during the whole simulation time. CDs have the capability to form hydrogen bonds with the surrounding water molecules or intramolecular ones. The incorporation of ligands into the hydrophobic interior of  $\beta$ -CD changes the preference of hydrogen bonds significantly and results in a contribution to the decrease of flexibility. In addition, QM calculations on the SP- $\beta$ -CD inclusion complexes with DFT method B3LYP/6-31G(d,p) were performed to determine the interaction energies and to prove the applicability of various methods.

**V. Density functional calculations on meloxicam- $\beta$ -cyclodextrin Inclusion Complexes** (Snor W, Liedl E, Weiss-Greiler P, Viernstein H, Wolschann P, 2008, doi:10.1016/j.i.jpharm.2009.05.012, *Int J Pharm*)

The geometries of the CD inclusion complexes with various tautomeric forms of meloxicam in vacuo were determined by DFT method B3LYP/6-31G(d,p). The interaction energies were estimated including BSSE correction. Two orientations of the meloxicam guest were considered: the benzene ring located near the narrow rim and at the wider rim of the  $\beta$ -CD, respectively. The calculations show that in all cases the molecules are located inside the CD-cavity. The preferred complexation orientation is that one, in which the benzene ring of meloxicam is located near the wider rim with the secondary hydroxyl groups of the CD. The stabilization energies for the encapsulation of the meloxicam guest molecules show an overall affinity ranking for the meloxicam guest molecule in the following order: anionic (deprotonated) form > zwitterionic form  $\sim$  enolic form > cationic (protonated) form. A comparison of the enolic and zwitterionic neutral forms shows, that the zwitterionic structure is better stabilized upon complexation due to the geometry of two extra hydrogen bonds between host and guest.

In addition to these papers presented above, the particular aims of this thesis comprise:

- An introductory consideration of molecular modelling and computational chemistry, including a short overview of the computational methods used in

this thesis as well as a brief consideration of basis sets and its effects (chapter 2.1)

- Investigation of the impact of the theory used on the magnitude of the BSSE correction energy on the calculated energies of the model system water-methanol dimer, where the water molecule acts as the hydrogen bond donor (chapter 3)
- Extensive considerations on how to conduct a computational molecular modelling study, including guidelines for the proper use of suitable coordinate systems and computational methods (chapter 4)
- Systematic investigation of the dependence of several geometric parameters of malonaldehyde on the level of theory used (chapter 5)
- Basic considerations of supramolecular chemistry including topics such as host-guest chemistry, molecular recognition and chiral recognition (chapter 6)
- Conduction of a case study with the aim to establish whether  $\beta$ -CD can act as a chiral selector between the two enantiomeric forms of the amino acid alanine (chapter 7)
- Review of the various molecular interactions in supramolecular systems with the focal point on hydrogen bonding (chapter 8)
- Review of CDs as supramolecular hosts, including structural features and inclusion chemistry of this very important group of compounds (chapter 9)
- Performance of a quantum mechanical calculation on  $\beta$ -CD inclusion complexes with neutral and anionic conformations of meloxicam; the latter was chosen because of its extraordinary importance as an effective COX-2 inhibitor in medicinal and pharmaceutical applications
- Overview of chemical structure drawing and viewing software as well as ab initio and DFT program packages that have been used in this thesis (Appendix A)

## 2 Methods and Programs

### 2.1 Molecular Modelling and Computational Chemistry

All chemists use **models**. Models are very useful because they often represent a simple way of describing and predicting scientific results without the work of performing the complex mathematical manipulations dictated by a rigorous theory. They can serve to simplify by limiting considerations to the essential phenomena that are believed to be the most important, and they can serve as didactical illustration of complicated problems, which are not easily accessible otherwise.

Besides models, **approximations** are another construct that is often encountered in chemistry. Even though a theory may give a rigorous mathematical description of chemical phenomena, the mathematical difficulties might be so great that it is just not feasible to solve a problem exactly. If a satisfactory result is desired, the best technique is often to do only part of the work, for example to completely leave out part of the calculation. Another approximation is to use an average rather than an exact mathematical description, to use perturbations, simplified functions, or to fit parameters to reproduce experimental results.

QM gives a mathematical description of the behaviour of electrons that has never been found to be wrong. However, the quantum mechanical equations have never been solved exactly for any chemical system other than the hydrogen atom. Thus, the entire field of computational chemistry is built around approximate solutions. Some of these solutions are very crude and others are expected to be more accurate than any experiment that has been conducted.

If an approximation is used, one must ask how accurate an answer should be. Computations of the energetic of molecules and reactions, for example, often attempt to attain what is called **chemical accuracy**, meaning an error of less than about one kcal/mol. This is sufficient to describe van der Waals interactions, the weakest interaction considered to affect most chemistry. Most chemists have no use for answers more accurate than this.

However, one must realize that models and approximations are powerful tools for understanding and achieving research goals. However, one also has to take into account that not all of them are perfect. Chemists are therefore advised to develop an understanding of the nature of computational chemistry approximations and what results can be trusted with any given degree of accuracy.

Molecular modelling is focused on applying the fundamental laws of physics and chemistry to the study of molecules. The ultimate aim is to create models and

simulations, which can help by predicting, rationalizing, and estimating the properties of molecules and their interactions. Today, computational techniques performed by powerful computers have revolutionised molecular modelling to the extent that most calculations could not be performed without the use of a computer. It allows chemists to study chemical phenomena by running calculations on computers rather than by examining reactions and compounds experimentally. Some methods can be used to model not only stable molecules, but also short-lived, unstable intermediates and even transition states. In this way, they can provide information about molecules and reactions, which is impossible to obtain through observations. Molecular modelling and computational chemistry is therefore both an independent research area and a vital adjunct to experimental studies.

Molecular modelling has undergone a dramatic change over the last decades mainly due to two factors:

1. Today's high level of computer technology has allowed an increase in the size of systems that can be studied, the degree of accuracy of the models and the number of interactions feasible to calculate on a reasonable time scale.
2. There has been tremendous progress in the experimental techniques that the different modelling tools rely on. X-ray crystallography and nuclear magnetic resonance (NMR) have been developed to a level where they are now applied routinely, which have an enormous impact on the number of experimentally determined molecular structures available.

Among the various properties most typically studied by computational chemists, the determination of the "best" structure of isolated molecules - as they are the fundamental units from which pure substances are constructed - is a very common undertaking. In this case, "best" is defined as having the lowest possible energy. This sounds relatively simple because it is about modelling of an isolated, single molecule. In the laboratory, however, one is much more typically dealing with an equilibrium mixture of a very large number of molecules at some non-zero temperature. In that case, measured properties reflect thermal averaging, possibly over multiple discrete stereoisomers, tautomers, etc., that are structurally quite different from the idealized model system, and great care must be taken in making comparisons between theory and experiment in such instances. To make a theory more closely mimic the experiment one has to consider not just one structure for a given chemical formula, but also all possible structures. That is, one characterizes the potential energy surface (PES) for a given chemical formula. Besides structural and energetic properties, several others that can be estimated by computational methods include spectral quantities, acidity, basicity (e.g.,  $pK_a$  values), hydrogen bond strengths, and so on.

### 2.1.1 Overview of Computational Chemistry Methods

All molecular calculation techniques can be classified under three general categories:

- ab initio and density functional electronic structure calculations,
- semi-empirical methods, and
- Molecular Mechanics

Table 1 summarizes some general characteristics for each of these methods.

*Tab. 1 General characteristic for computational methods*

Method	Method-description	Advantages	Disadvantages	Best for
Ab initio and DFT	Uses quantum physics; mathematically rigorous: no empirical parameters	Useful for a broad range of systems; does not depend on experimental data; calculates transition states and excited states	Computationally expensive	Small systems (tens of atoms); electronic transitions; systems without experimental data; systems requiring high accuracy.
Semi-empirical	Uses quantum physics, experimental parameters and extensive approximations	Less demanding computationally than ab initio methods; calculates transition states and excited states	Requires ab initio or experimental data for parameters; less rigorous than ab initio and DFT methods	Medium sized systems (hundreds of atoms)
Molecular Mechanics	Uses classical physics; relies on force fields with embedded empirical parameters	Computationally "cheap": fast and useful with limited computer resources.	Does not calculate electronic properties; requires ab initio or experimental data for parameters	Large systems (thousands of atoms); systems or processes that do not involve bond breaking

Several useful books covering fundamental and advanced principles of molecular modelling and computational chemistry are available (Foresman and Frisch 1993; Kunz 1997; Jensen 1999; Leach 2001; Young 2001; Schlick 2002; Hincliffe 2003; Höltje et al. 2003; Cramer 2004; Reinhold 2004; Atkins and Friedman 2005; Ramachandran et al. 2008)

#### 2.1.1.1 Molecular Mechanics (MM)

The term MM was coined in the 1970s to describe the application of classical mechanics to calculate the **static** properties of a molecule or a group of molecules, such as structure, energy, or electrostatics. If one is interested in **dynamic** properties like the time evolution of a molecular system, resulting in a trajectory of snapshots,

one has to use **molecular dynamics**. Finally, if one needs to know thermodynamic properties like enthalpies, or include entropy or free energy, an alternative to sampling the conformational space by molecular dynamics is to apply **Monte Carlo simulations**. The latter method does not concern time evolution at all, but is generally considered to generate statistically meaningful thermodynamic ensembles much more effectively.

MM is often the only feasible means with which to model very large and non-symmetric chemical systems such as proteins or polymers. MM is a purely empirical method that neglects explicit treatment of electrons, relying instead upon the laws of classical physics to predict the chemical properties of molecules. As a result, MM calculations cannot deal with problems such as bond breaking or formation, where electronic or quantum effects dominate. Furthermore, MM models are wholly system-dependent; MM energy predictions tend to be meaningless as absolute quantities, and are generally useful only for comparative studies. Despite these shortcomings, MM bridges the gap between quantum and continuum mechanics.

There are many different MM methods. Each one is characterised by its particular **force field**.

Generally, a standard modern MM force field can be written as

$$E = E_{stretch} + E_{bend} + E_{torsion} + E_{nonbonded} \quad (2.1)$$

Table 2 gives the mathematical forms of energy terms often used in popular force fields.

**Tab. 2** Common force field terms ( $l$ -bond length,  $\theta$ -bond angle,  $k$ ,  $\alpha$ ,  $A$ ,  $B$ -constants particular to the elements in a certain hybridization state,  $n$ -an integer,  $r$ -nonbonding distance,  $q$ -charge,  $D_e$  dissociation energy)

Name	Use	Energy Term
Harmonic	Bond stretch	$k(l-l_0)^2$
Harmonic	Angle bend	$k(\theta-\theta_0)^2$
Cosine	Torsion	$k[1+\cos(n\theta)]$
Lennard-Jones 6-12	Van der Waals	$4k\left(\frac{A}{r}\right)^{12}-\left(\frac{B}{r}\right)^6$
Lennard-Jones 10-12	Van der Waals	$4k\left(\frac{A}{r}\right)^{12}-\left(\frac{B}{r}\right)^{10}$
Coulomb	Electrostatic	$\frac{q_1q_2}{4\pi\epsilon_0r}$
Taylor	Stretch-bend	$k(\theta-\theta_0)\left[\left(l_1-l_{10}\right)\left(l_2-l_{20}\right)\right]$
Morse	Bond stretch	$D_e\left[1-e^{-\alpha(l-l_0)}\right]^2$

The constants may vary from one force field to another according to the designer's choice of unit system, zero of energy, and fitting procedure. All the constants in these equations must be obtained from experimental data or ab initio calculations. However, the database of compounds used to parameterize the method is crucial to its success. A MM method may be parameterized against a specific class of molecules, such as, for example, proteins. Such a force field would only be expected to have any relevance in describing other proteins. Other force fields are parameterized to give a reasonable description of a wide variety of organic compounds. A few force fields have even been parameterized for all the elements.

Many good force fields have already been developed. A comprehensive comparison of several force fields focusing on the calculation of conformational energies of organic molecules has been published by Pettersson and Liljefors (1996). Table 3 lists some commonly used MM force fields that have been designed for organic molecules and are implemented in more than one software package. There

tend to be minor differences in the implementation leading to small differences in results from one software package to another.

**Tab. 3** Commonly used MM force fields

Name	Description
MM <sub>2</sub> /MM <sub>3</sub> /MM <sub>4</sub>	The parameterizations provided by Norman Allinger and co workers are probably the best-known implementations of the MM concept and are widely used as a synonym for force field calculations in general. MM <sub>2</sub> is designed primarily for hydrocarbons. It was extended by many functional groups to cover almost all kinds of small organic molecules. The MM <sub>3</sub> method with significant improvements in the functional form is probably one of the most accurate ways of modelling hydrocarbons. It was also extended to handle amides, polypeptides, and proteins.
MM+	Extension of the MM <sub>2</sub> force field; developed by Allinger and co workers.
AMBER (Assisted model building with energy refinement)	One of the most popular force fields for modelling proteins and nucleic acids
GROMOS (Groningen molecular simulation)	It is a molecular dynamics computer simulation package for the study of biomolecular systems with the purpose of energy minimization and the simulation of molecules in solution or solid state by molecular dynamics, stochastic dynamics, or path-integral methods.
UFF (Universal Force Field)	It is a set of simple functional forms and parameters used to model the structure, movement, and interaction of molecules containing any combination of elements in the periodic table.
MMFF (Merck Molecular Force Field)	It is based primarily on quantum mechanical calculations of the energy surface and is able to handle all functional groups of interest in pharmaceutical design.

### 2.1.1.2 Ab Initio Electronic Structure Methods

The term *ab initio* is Latin for “from the beginning”. This name is given to computations that are derived from theoretical principles with no inclusion of experimental data. Over the past three decades, *ab initio* electronic structure methods have become an indispensable tool in the study of both atoms and molecules and in modelling complex systems, consisting of two or more components.

The underlying main technology is the computational solution of the electronic Schrödinger equation. In its exact form, the Schrödinger equation is a many-body problem, whose computational complexity grows exponentially with the number of electrons.

Two highly efficient approaches to solution of the electronic Schrödinger equation have arisen to date:

- **Wave function-based approaches** expand the electronic wavefunction as a sum of Slater determinants, the orbitals and coefficients of which are optimized by certain numerical procedures. **Hartree-Fock theory** is the simplest method of this type, involving the optimization of a single determinant only. However, its usefulness is limited because of complete neglect of electron correlation.
- The second class of theoretical approaches is based on **density functional theory**. The premise behind this theory is that the energy of a molecule can be determined from the electron density instead of a wave function. Kohn and Sham who formulated a method similar in structure to the Hartree-Fock method developed a practical application of this theory. The advantage of using electron density is that it depends on three coordinates instead of  $3N$  coordinates of  $N$  electrons only, thus scaling as  $N^3$ . Furthermore, at least some electron correlation can be included in the calculation. This results in faster calculations than HF calculations (which scale as  $N^4$ ) and computations are a bit more accurate as well. The better DFT functionals give results with accuracy similar to that of an MP2 calculation.

### 2.1.1.3 Basic Quantum Mechanics (QM)

QM is the correct mathematical description of the behaviour of electrons and thus of chemistry. In theory, QM can predict any property of an individual atom or molecule exactly. In practice, the QM equations have only been solved exactly for one-electron systems. A huge collection of methods has been developed for approximating the solution for multiple electron systems. These approximations can be very useful, but this requires an amount of sophistication on the part of the researcher to know when each approximation is valid and how accurate the results are likely to be.

Schrödinger and Heisenberg devised two equivalent formulations of QM. Here, only the Schrödinger form is presented since it is the basis for nearly all computational chemistry methods. The Schrödinger equation is

$$\hat{H}\psi = E\psi, \quad (2.2)$$

where  $\hat{H}$  is the Hamiltonian operator,  $\psi$  a wave function and  $E$  the energy.

In mathematics, an equation of this form is called an **Eigen equation**.  $\psi$  is then called the **eigenfunction** and  $E$  an **eigenvalue**. The operator and eigenfunction can be a matrix and vector, respectively, but this is not always the case.

The wave function  $\psi$  is a function of the electron and nuclear positions. As the name implies, this is the description of an electron as a wave. As such, it can describe the probability of electrons being in certain locations, but it cannot predict exactly where electrons are located. The wave function is also called probability amplitude because it is the square of the wave function that yields probabilities. This is the only rigorously correct meaning of a wave function. In order to obtain a physically relevant solution of the Schrödinger equation, the wave function must be continuous, single-valued, normalizable, and antisymmetric with respect to the interchange of electrons.

The Hamilton operator  $\hat{H}$  is, in general,

$$\hat{H} = - \sum_i^{\text{particles}} \frac{\nabla_i^2}{2m_i} + \sum_{i < j}^{\text{particles}} \sum \frac{q_i q_j}{r_{ij}}, \quad (2.3)$$

where

$$\nabla_i^2 = \frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2},$$

the Laplacian operator acting on particle  $i$ . Particles are both electrons and nuclei.

The symbols  $m_i$  and  $q_i$  are the mass and charge of particle  $i$ , and  $r_{ij}$  is the distance between particles. The first term gives the kinetic energy of the particle within a wave formulation. The second term is the energy due to Coulombic attraction or repulsion of particles.

This formulation is the time-independent, nonrelativistic Schrödinger equation.

Additional terms can appear in the Hamiltonian when relativity or interactions with electromagnetic radiation or fields are taken into account.

In currently available software, the Hamiltonian above is nearly never used. The problem, however, can be simplified by separating the nuclear and electron motions. This is called the **Born-Oppenheimer approximation**.

Using atomic units, the Hamiltonian for a molecule with stationary nuclei is

$$\hat{H} = - \sum_i^{\text{electrons}} \frac{\nabla_i^2}{2} - \sum_i^{\text{nuclei}} \sum_j^{\text{electrons}} \frac{Z_i}{r_{ij}} + \sum_{i < j}^{\text{electrons}} \sum \frac{1}{r_{ij}} \quad (2.4)$$

Here, the first term is the kinetic energy of the electrons only. The second term is the attraction of electrons to nuclei. The third term is the repulsion between electrons. The repulsion between nuclei is added onto the energy at the end of the calculation.

The motions of nuclei can then be described by considering this entire formulation to be a potential energy surface on which nuclei move.

Once a wave function has been determined, any property of the individual molecule can be determined. This is done by taking the expectation value of the respective quantum mechanical operator for that property, denoted with angled brackets  $\langle \rangle$ .

For example, the energy is the expectation value of the Hamiltonian operator:

$$\langle E \rangle = \int \psi^* \hat{H} \psi, \quad (2.5)$$

where  $\psi^*$  represents the complex conjugate wave function.

For an exact solution, this is the same as the energy predicted by the Schrödinger equation. For any approximate wave function, this gives an approximation of the energy, which is the basis for some of the techniques described below. It is called **variational energy** because it is always greater than or equal to the exact energy.

By substituting different operators, it is possible to obtain different observable properties, such as the dipole moment or electron density. Properties other than the energy are not variational, because only the Hamiltonian is used to obtain the wave function in the widely used computational chemistry methods.

#### 2.1.1.4 The Hartree-Fock (HF) Self-Consistent Field Approximation

The most common type of ab initio calculation is called a HF calculation, in which the primary approximation is the **central field approximation**. This means that the Coulombic electron-electron repulsion is taken into account by integrating the repulsion term. This gives the average effect of the repulsion, but not the explicit repulsion interaction. Instead, this is a variational calculation, meaning that the approximate energies calculated are all equal to or greater than the exact energy and tend to a limiting value called the **HF limit** as the basis set (see below) is improved.

One of the advantages of this method is that it breaks the many-electron Schrödinger equation into many simpler one-electron equations. Each one-electron equation is solved to yield a single-electron wave function, called an orbital, and energy, called an orbital energy. The orbital describes the behaviour of an electron in the net field of all the other electrons.

The second approximation in HF calculations is because the wave function must be described by some mathematical function, which is known exactly for only a few one-electron systems. The functions used most often are linear combinations of Gaussian-type orbitals (GTO). The wave function is formed from linear combinations of atomic orbitals or, stated more correctly, from linear combinations

of “basis functions”. Because of this approximation, most HF calculations give a computed energy greater than the HF-limit. The exact set of basis functions used is often specified by an abbreviation, such as STO-3G or 6-31G(d,p). Basis sets are discussed further in Chapter 2.1.1.8.

The Gaussian functions are multiplied by an angular function in order to give the orbital the symmetry of a  $s$ ,  $p$ ,  $d$ , and so on. A constant angular term e.g. yields  $s$  symmetry. Angular terms of  $x$ ,  $y$  and  $z$  give  $p$  symmetry. This pattern can be continued for the other orbitals.

These orbitals are then combined into a determinant, the so-called **Slater-Determinant**. This is done to satisfy two requirements of QM. One is that the electrons must be indistinguishable. By having a linear combination of orbitals in which each electron appears in each orbital, it is only possible to say that an electron was put in a particular orbital but not which electron it is. The second requirement is that the wave function for fermions (an electron is a fermion) must be antisymmetric with respect to interchanging two particles. Thus, if electron 1 and electron 2 are switched, the sign of the total wave function must change and only the sign can change. This is satisfied by a determinant because switching two electrons is equivalent to interchanging two columns of the determinant, which in turn changes its sign.

The functions put into the determinant do not need to be individual GTO functions, called Gaussian primitives. They can also be a weighted sum of basis functions on the same atom or on different atoms.

The steps in a Hartree-Fock calculation start with an initial guess for the orbital coefficients, usually using a semi-empirical method. This function is used to calculate energy and a new set of orbital coefficients, which can then be used to obtain a new set, and so on. This procedure continues iteratively until energies and orbital coefficients remain constant from iteration to the next one. This is called having the calculation converge. The iterative procedure itself is called a **self-consistent field procedure (SCF)**.

A variation on the HF procedure is the way that orbitals are constructed to reflect paired or unpaired electrons. If the molecule has a singlet spin, then the same orbital spatial function can be used for both the  $\alpha$  and  $\beta$  spin electrons in each pair. This is called the **restricted Hartree-Fock method (RHF)**. If otherwise two completely separate sets of orbitals for the  $\alpha$  and  $\beta$  electrons are used, this method is called the **unrestricted Hartree-Fock method (UHF)**. However, the latter introduces an error into the calculation, called **spin contamination**, which could be large enough to make the results unusable depending on the chemical system involved.

The RHF scheme results in forcing electrons to remain paired. This means that the calculation will fail to reflect cases where the electrons should uncouple. Therefore,

this limitation must be considered whenever processes involving pairing and impairing of electrons are modelled.

### 2.1.1.5 Electron Correlation

The HF method yields, even in favourable cases and if large basis sets are employed, only an approximation to the exact solution of the electronic Schrödinger equation. It does not consider the instantaneous Coulombic interaction between electrons, nor does it take into account the quantum mechanical effects on electron distributions. The effect of the N-1 electrons on the electron of interest is treated only in an average way. Therefore, in cases where accurate results are to be obtained, one has to go beyond the HF method. These methods are generally called post-SCF techniques. However, there are too many different methods for considering electron correlation to be described in detail here, so the discussion will be limited to the general principles of the most common techniques.

**Configuration interaction (CI)** solves the problem of electron correlation by considering more than a single occupation scheme for the MOs and by mixing the microstates obtained by permuting the electron occupancies over the available MOs. In its simplest form, a CI calculation consists of a preliminary SCF calculation, which gives the MOs that are used unchanged throughout the rest of the calculation. Microstates are then constructed by moving electrons from occupied orbitals to vacant ones according to preset schemes. However, the problem is that if you want to consider every possible arrangement of all the electrons in all the MOs (a full CI), the calculations would become far too large even for moderate-sized molecules with a large basis set. Thus, two types of restriction are usually used: only a limited number of MOs are included in the CI, and only certain types of rearrangement (excitation) of the electrons are used. The most economical form is that in which only one electron is promoted from the ground state to a virtual orbital (single excitations). This is abbreviated as CIS and has traditionally been used for calculating spectra. Adding all double excitations (in which two electrons are promoted) gives CISD, and so on.

A more practical way of considering electron correlation is to use perturbation theory to apply a correction to the SCF energy. Such an approach was first proposed by Møller and Plesset (1934) for atoms and was extended by Pople et al. (1976) to molecules. Because it is a perturbational treatment, **Møller-Plesset (MP) theory** can be applied considering the perturbation series to include different numbers of terms (i.e., to different orders). Second-order MP theory (MP2) is often used for geometry optimizations and fourth-order (MP4) for refining calculated energies. The reason, for instance, that MP3 theory is used less often is that the MP series tends to oscillate, so that using only the even-numbered orders gives results that are more consistent. MP techniques are size-consistent and computationally efficient, so that their use is very common.

The **coupled cluster (CC) methods** and **quadratic CI** form a further group of related techniques for considering electron correlation. These techniques represent the corrected wavefunction  $\psi$  as the result of applying a so-called cluster operator to the HF wavefunction. The cluster operator can be built up from a series of operators that consider excitations of one, two, three, n electrons, where n is the total number of electrons in the molecule. Thus, CC techniques can be truncated like MP methods, but are more accurate. However, they are also computationally more expensive. CC calculations using single and double excitations (CCSD) are common, but very often an additional perturbational term to take some triple excitations into account is used to give CCSD(T). CCSD(T) calculations (or the closely related QCISD(T) technique) represent about the best that is currently possible using an HF wavefunction as the starting point (reference wavefunction).

To sum up: ab initio calculations, in general, give very good qualitative results and can yield increasingly accurate quantitative results as the molecules in question become smaller. The advantage of ab initio methods is that they eventually converge to the exact solution once all the approximations are made sufficiently small in magnitude. In general, the relative accuracy of results is

$$HF \ll MP2 < CISD \cong MP4 \cong CCSD < CCSD(T) < CCSDT < Full CI$$

However, this convergence is not monotonic. Sometimes, the smallest calculation can give a very accurate result for a property under consideration.

In ab initio calculations, there are four sources of error:

- The Born-Oppenheimer approximation
- The use of an incomplete basis set
- Incomplete correlation
- The omission of relativistic effects

The disadvantage of ab initio methods is that they are computationally expensive. These methods often take enormous amounts of computer CPU time, memory, and disk space. The HF method scales as  $N^4$ , where N is the number of basis functions. This means that a calculation twice as big takes 16 times as long ( $2^4$ ) to complete. Correlated calculations often scale much worse than this. In practise, extremely accurate solutions are only obtainable when the molecule contains a dozen electrons or less. However, results with an accuracy rivalling that of many experimental techniques can be obtained for moderate sized organic molecules. The minimally correlated methods, such as MP2, are often used when correlation is important to the description of large molecules.

#### 2.1.1.6 Density Functional Theory (DFT)

DFT has become very popular in recent years, because it is less computationally intensive than other methods with similar accuracy. The premise behind DFT is that

the energy of a molecule can be determined from the electron density instead of a wave function. This theory originated with a theorem by Hohenberg and Kohn that stated this was possible (Hohenberg and Kohn 1964). Kohn and Sham who formulated a method similar in structure to the Hartree-Fock method (Kohn and Sham 1965) developed a practical application of this theory.

They suggested calculating the kinetic energy of the non-interacting electron density that corresponds to the real one exactly, and treating the correction from this energy to that of the real, interacting system approximately. The correction to the non-interacting kinetic energy is known as the **exchange correlation (XC) energy** and is calculated as a function of the electron density. As the electron density itself is a function, the XC energy is a function of a function, which is known as a **functional**; hence the name “density functional theory”. Its basic principles are described more fully by Koch and Holthausen (2001).

The advantage of using electron density is that the integrals for Coulomb repulsion need be done only over the electron density, which is a three-dimensional function, thus scaling as  $N^3$ . Furthermore, at least some electron correlation can be included in the calculation. This results in faster calculations than HF calculations (which scale as  $N^4$ ) and computations those are a bit more accurate as well. The better DFT functionals give results with accuracy similar to that of an MP2 calculation.

The problem is that one does not know the functional(s) that translate the electron density into the XC energy. There are now many alternative functionals available, but there is no way to say that functional A is better than functional B. Thus, the major advantage of ab-initio theory, the ability to improve it systematically, is lost in DFT.

There are, however, three basic types of functional.

The **local density approximation (LDA)** is the oldest and simplest of the functional types still in use. It is based on the idea of a uniform electron gas, a homogeneous arrangement of electrons moving against a positive background charge distribution that makes the total system neutral. This construct is abstract and not very realistic, but one does know the exact form of the exchange part of the XC functional for it and has accurate results to simulate for the correlation part. Importantly, the XC energy depends only on the electron density itself at a given position and so is easy to calculate. LDA calculations are thus very fast and often give good geometries. They tend, however, to give systematic errors in the energy and generally make bonds too strong. LDA calculations are therefore used less often for molecular applications than more sophisticated functionals.

The **generalized gradient approximation (GGA)** gives better results. GGA functionals are usually divided into **exchange** and **correlation functionals**, which are often derived separately and may be combined in different ways. The most

important practical feature of GGA functionals is that they depend not only on the value of the electron density itself, but also on its derivative (gradient) with respect to the position in space. The inclusion of the first derivative of the density allows GGA functionals to treat the inhomogeneities in the electron density better than LDA functionals. Koch and Holthausen (2001) give an up-to-date list of GGA exchange and correlation functionals.

The third class of density functional methods considered here, the **hybrid functionals**, are simply a combination of a GGA correlation functional with an exchange contribution that comes partly from an exchange functional and partly from HF theory, where the exchange energy is calculated exactly (Becke 1993a). The relative proportions of the HF exchange energy and those of the two GGA functionals vary between hybrid methods and are usually parameterized to fit a set of experimental data. Hybrid methods are generally the most accurate but suffer the disadvantage that calculating the HF exchange energy requires four-centre integrals. Hybrid DFT calculations are thus more expensive computationally than GGA.

Most DFT calculations today are being done with HF-optimized GTO basis sets. The accuracy of results tends to degrade significantly with the use of very small basis sets. For accuracy considerations, the smallest basis set used is generally 6-31G(d) or the equivalent. Interestingly, there is only a small increase in accuracy obtained by using very large basis sets. This is probably because the density functional is limiting accuracy more than the basis set limitations.

The accuracy of results from DFT calculations can be poor to good, depending on the choice of basis set and density functional. A variety of exchange-correlation functionals has been developed for use in DFT calculations; the names designate a particular pairing of an exchange functional and a correlation functional. For example, the popular BLYP functional is a combination of the gradient-corrected exchange functional developed by Becke (Becke 1986) and the gradient-corrected correlation functional developed by Lee, Yang, and Parr (Lee et al. 1988).

To date, the B3LYP hybrid functional (also called Becke3LYP) is the most widely used for molecular calculations with basis sets of 6-31G(d) or larger (Becke, 1993b). This is due to the often optimal accuracy versus CPU time, and therefore the B3LYP method is the method of choice for many organic molecule calculations.

Due to the newness of DFT, its performance is not completely known and continues to change with the development of new functionals. Cramer (2004) gives a broad overview of the applications and performance of DFT depending on the level of theory used. In addition, a detailed discussion on the advantages and disadvantages of DFT compared to MO theory is given. In a recent publication (Lynch and Truhlar 2003), a variety of DFT-based calculations were performed to compute barrier heights for six small-molecule reactions and atomization (complete dissociation) energies for six different molecules. The results for a variety of DFT

functionals and basis sets were compared with each other and with results of HF-based techniques.

Unfortunately, as mentioned above, there is no systematic way to improve DFT calculations, thus making them unusable for very-high-accuracy work. It is therefore prudent to look for relevant literature and run test calculations before using these methods.

#### 2.1.1.7 Semi-empirical Methods

Semi-empirical calculations are set up with the same general structure as a HF calculation in that they have a Hamiltonian and a wave function. Within this framework, certain pieces of information are approximated or completely omitted. Usually, the core electrons are not included in the calculation and only a minimal basis set is used. In addition, some of the two-electron integrals are omitted. In order to correct for the errors introduced by omitting part of the calculation, the method is parameterized. Parameters to estimate the omitted values are obtained by fitting the results to experimental data or ab initio calculations. Often, these parameters replace some of the integrals that are excluded.

The advantage of semi-empirical calculations is that they are much faster than ab initio calculations. The disadvantage of semi-empirical calculations is that the results can be erratic and fewer properties can be predicted reliably. If the molecule being computed is similar to molecules in the database used to parameterize the method, then the results may be good. If the molecule being computed is significantly different from anything in the parameterization set, the answers may be very poor.

Semi-empirical calculations have been very successful in the description of organic chemistry, where there are only a few elements used extensively and the molecules are of moderate size. Some semi-empirical methods have been devised specifically for the description of inorganic chemistry as well. Table 4 presents an overview of some of the most commonly used semi-empirical methods.

*Tab. 4 Overview of some commonly used semi-empirical methods*

Name	Description
HÜCKEL	The Hückel method is one of the earliest and simplest semi-empirical methods. A Hückel calculation models only the $n$ valence electrons in a planar conjugated hydrocarbon. A parameter is used to describe the interaction between bonded atoms. Hückel calculations do reflect orbital symmetry and qualitatively predict orbital coefficients. Hückel calculations can give crude quantitative information or qualitative insight into conjugated compounds, but are seldom used today
CNDO (Complete Neglect of Differential Overlap)	It is the simplest method of this type and models valence orbitals only by using a minimal basis set of Slater type orbitals. It is still sometimes used to generate the initial guess for ab initio calculations on hydrocarbons
MNDO (Modified Neglect of Diatomic Overlap)	This method has been found to give reasonable qualitative results for many organic systems and has been incorporated into several popular semi-empirical programs. It is still used, but the more accurate AM1 and PM3 methods have surpassed it in popularity
AM1 (Austin Model 1)	The Austin Model 1 method is still popular for modelling organic compounds. Hydrogen bonds are predicted to have the correct strength, but often the wrong orientation. Depending on the nature of the system and information desired, either AM1 or PM3 will often give the most accurate results obtainable for organic molecules with semi-empirical methods. On average, AM1 predicts energies and geometries better than MNDO, but not as well as PM3. Computed bond enthalpies are consistently low.
PM3	PM3 uses nearly the same equations as the AM1 method along with an improved set of parameters. The PM3 method is also currently extremely popular for organic systems. It is more accurate than AM1 for hydrogen bond angles, but AM1 is more accurate for hydrogen bond energies. The PM3 and AM1 methods are also more popular than other semi-empirical methods due to the availability of algorithms for including solvation effects in these calculations. There are also some known strengths and limitations of PM3. Overall heats of formation are more accurate than with MNDO or AM1. Hypervalent molecules are also treated more accurately. PM3 tends to predict that the barrier to rotation around the C-N bond in peptides is too low. Moreover, it tends to predict $sp^3$ nitrogen as always being pyramidal. Some spurious minima are predicted. Proton affinities are not accurate. Some polycyclic rings are not flat. The predicted charge on nitrogen is incorrect. No bonded distances are too short. Hydrogen bonds are too short by about 0.1 Å, but the orientation is usually correct. On average, PM3 predicts energies and bond lengths more accurately than AM1 or MNDO

To sum up: Semi-empirical methods can provide results accurate enough to be useful, particular for organic molecules with computation requirements low enough to make them convenient on PCs. These methods are generally good for predicting molecular geometry and energetic and can be used for predicting vibrational modes and transition structures but do so less reliably than ab initio methods. They generally give poor results for van der Waals and dispersion intermolecular forces, due to the lack of diffuse basis functions.

### 2.1.1.8 Basis Sets

A basis set is a set of mathematical functions from which a wave function can be constructed. As considered in chapter 2.1.1.4., each MO in HF theory is expressed as a linear combination of basis functions, the coefficients for which are determined from the iterative solution of the HF SCF procedure. The full HF wave function is expressed as a Slater determinant formed from the individual occupied MOs.

In the abstract, the HF limit is achieved by use of an infinite basis set, which necessarily permits an optimal description of the electron probability density.

In practice, however, one cannot make use of an infinite basis set. Thus, much work has gone into identifying mathematical functions that allow wave functions to approach the HF limit arbitrarily closely in as efficient a manner as possible.

Efficiency in this case involves three considerations:

- Because the number of two-electron integrals increases as  $N^4$  where  $N$  is the number of basis functions, so keeping the total number of basis functions to a minimum is computationally attractive.
- In addition, however, it can be useful to choose basis set functional forms that permit the various integrals appearing in the HF equations to be evaluated in a computationally efficient fashion. Thus, a larger basis set can still represent a computational improvement over a smaller basis set if evaluation of the greater number of integrals for the former can be carried out faster than for the latter.
- Finally, the basis functions must be chosen to have a form that is useful in a chemical sense. That is, the functions should have large amplitude in regions of space where the electron probability density (the wave function) is also large, and small amplitudes where the probability density is small.

The simultaneous optimization of these three considerations is at the heart of basis set development.

Most semi-empirical methods use a predefined basis set. When ab initio or DFT calculations are done, a basis set must be specified. Although it is possible to create a basis set from scratch, most calculations are done using existing basis sets. The type of calculation performed and basis set chosen mainly determine the accuracy of results. What follows below is a discussion of standard basis sets and considerations on how to choose an appropriate one.

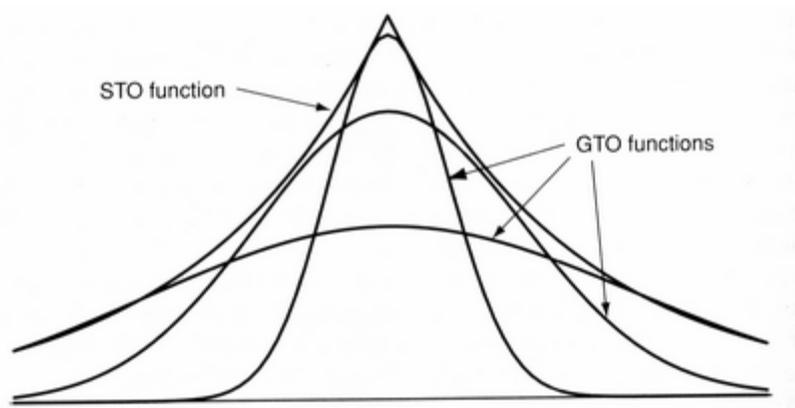
The orbitals used in ab initio calculations usually have the following functional form:

$$\varphi = Y_{lm} \sum_i C_i \sum_j C_{ij} e^{-\zeta_{ij} r^2} \quad (2.6)$$

The  $Y_{lm}$ -function gives the orbital the correct symmetry ( $s$ ,  $p$ ,  $d$ , etc.).  $e^{-r^2}$  is called a Gaussian primitive function. The contraction coefficients  $C_{ij}$  and exponents  $\zeta_{ij}$  are read from a database of standard functions and do not change over the course of the calculation. This predefined set of coefficients and exponents is called a **basis set**. By using such a predefined basis set, the program must only optimize the molecular orbital coefficients  $C_i$ . Each  $C_i$  may weigh a sum of typically one to nine primitive Gaussian functions, called a **contraction**. Basis sets of contracted functions are called **segmented basis sets**.

As mentioned above, the choice of basis set has a large effect on the amount of CPU time required to perform a calculation. In general, the amount of CPU time for HF calculations scales as  $N^4$ . This means that making the calculation twice as large will make the calculation take 16 times ( $2^4$ ) as long to run. Making the calculation twice as large can occur by switching to a molecule with twice as many electrons or by switching to a basis set with twice as many functions. Disk use for conventional calculations scales as  $N^4$  and the amount of RAM use scales as  $N^2$  for most algorithms. Some of the largest CI calculations scale as  $N^8$  or worse.

The orbitals in Equation 2.6 are referred to as **Gaussian type orbitals (GTOs)**, since they incorporate Gaussian functions,  $e^{-\zeta r^2}$ . The exact solution to the Schrödinger equation for the hydrogen atom is a **Slater type orbital (STO)** of the form  $e^{-\zeta r}$ . GTO basis sets require more primitives to describe the wave function than are needed for STO calculations, as shown in Figure 1.



*Fig. 1* Approximating a STO with several GTOs

However, the integrals over GTO primitives can be computed analytically, which is so much faster than the numeric integrals over STO functions that any given accuracy can be obtained most quickly using GTO functions. As such, STO basis sets are sometimes used for high-accuracy work, but most calculations are now done with GTO basis sets. Choosing a standard GTO basis set means that the wave function is being described by a finite number of functions. This introduces an

approximation into the calculation since an infinite number of GTO functions would be needed to describe the wave function exactly. Differences in results due to the quality of one basis set versus another are referred to as **basis set effects**. In order to avoid the problem of basis set effects, some high-accuracy work is done with **numeric basis sets**. These basis sets describe the electron distribution without using functions with a predefined shape. A typical example of such a basis set might be a cubic spline set in which a large number of third-order polynomials are used. Each polynomial would describe the wave function for just a small range of distances from the nucleus. The coefficients of these polynomials are then chosen so that the wave function and its derivatives will be continuous as well as describing the shape of the wave function.

Basis sets are identified by one of a number of notation schemes. These abbreviations are often used as the designator for the basis set in the input to ab initio computational chemistry programs. The following is a look at the notation for identifying some commonly available contracted GTO basis sets.

The smallest basis sets are called **minimal basis sets**. The most popular minimal basis set is the **STO-3G** set. This notation indicates that the basis set approximates the shape of a STO orbital by using a single contraction of three GTO orbitals. One such contraction would then be used for each orbital, which is the definition of a minimal basis. Minimal basis sets are used for very large molecules, qualitative results, and in certain cases quantitative results. There are **STO-nG** basis sets for  $n=2-6$ .

Another family of basis sets, commonly referred to as the **Pople basis sets**, are indicated by the notation **6-31G**. This notation means that each core orbital is described by a single contraction of six GTO primitives and each valence shell orbital is described by two contractions, one with three primitives, and the other with one primitive. These basis sets are very popular, particularly for organic molecules. Other Pople basis sets in this set are 3-21G, 4-31G, 4-22G, 6-21G, 6-311G, and 7-41 G.

The Pople basis set notation can be modified by adding one or two asterisks, such as **6-31G\*** or **6-31G\*\***. A single asterisk means that a set of *d*-primitives has been added to atoms other than hydrogen. Two asterisks mean that a set of *p*-primitives has been added to hydrogen as well. These are called **polarization functions** because they give the wave function more flexibility to change shape. Adding polarization functions usually decreases the variational total energy by about the same amount as adding another contraction. Polarization functions are used because they often result in more accurate computed geometries and vibrational frequencies. **The 3-21G\*** basis is an exception to the notation above. In this particular case, the *d*-functions are added only to 2nd row atoms, Al through Ar.

One or two plus signs can also be added, such as **6-31+G\*** or **6-31++G\***. A single plus sign indicates that **diffuse functions** have been added to atoms other than

hydrogen. The second plus sign indicates that diffuse functions are being used for all atoms. These diffuse functions are primitives with small exponents, thus describing the shape of the wave function far from the nucleus. Diffuse functions are used for anions, which have larger electron density distributions. They are also used for describing interactions at long distances, such as van der Waals interactions. The effect of adding diffuse functions is usually to change the relative energies of the various geometries associated with these systems. Basis sets with diffuse functions are also called **augmented basis sets**.

As the Pople basis sets have further expanded to include several sets of polarization functions,  $f$  functions and so on, there has been a need for a new notation. In recent years, the types of functions being added have been indicated in parentheses. An example of this notation is **6-31G(dp,p)** which means that extra sets of  $p$  and  $d$  functions have been added to non-hydrogen atoms and an extra set of  $p$  functions has been added to hydrogen. Thus, this example is synonymous with **6-31+G\*\***.

Some useful books discussing construction, notation schemes and performance of basis sets are (Hehre et al. 1986, 1995; Jensen 1999; Leach 2001; Young 2001; Atkins and Friedman 2005).

The **complete basis set (CBS)** scheme is a series of basis sets designed to extrapolate energies to the infinite basis set limit.

The **Basis Set Superposition Error (BSSE)** is a contribution to the inaccuracy of calculations that stems from the use of a finite basis set; this error may arise in the calculation of the interaction energy of two weakly bound systems such as van der Waals complexes or hydrogen bonded systems.

If, for example, the hydrogen bond between two water molecules is considered: the simplest approach consists in calculating the energy of the dimer and subtracting twice the energy of an isolated molecule. The electron distribution within each water molecule in the dimer is very close to that of the monomer. In the dimer, however, basis functions from one molecule can help compensate for the basis set incompleteness on the other molecule, and vice versa. The dimer will therefore be artificially lowered in energy, and the strength of the hydrogen bond overestimated. In the limit of a complete basis set, BSSE will be zero; adding additional basis functions does not give any improvement. The conceptually simplest approach for eliminating BSSE is therefore to add more and more basis functions, until the interaction energy no longer changes. Unfortunately, this requires very large basis sets and therefore enormous amounts of computer power.

An approximate way of assessing BSSE is the **Counterpoise (CP) correction** (Boys and Bernardi 1970). In this method, the BSSE is estimated as the difference between monomer energies with the regular basis and the energies calculated with the full set of basis functions for the whole complex.

Considering two molecules A and B, each having regular nuclear centred basis sets denoted by subscripts a and b, and the complex AB having the combined basis set  $a \cup b$ . The geometries of the two isolated molecules and of the complex are first optimized. The geometries of the A and B molecules in the complex will usually be slightly different from those of the isolated species, and the complex geometry will be denoted by an asterisk \*. The dimer energy minus the monomer energies is the directly calculated complexation energy.

$$\Delta E_{\text{complexation}} = E(AB)_{a \cup b}^* - E(A)_a - E(B)_b \quad (2.7)$$

To estimate how much of this complexation energy is due to BSSE, four additional energy calculations are needed. Using basis set a for A, and basis set b for B, the energies of each of the two fragments are calculated with the geometry they have in the complex. Two additional energy calculations of the fragments at the complex geometry are then carried out with the full  $a \cup b$  basis set. For example, the energy of A is calculated in the presence of both the normal a basis functions and with the b basis functions of fragment B located at the corresponding nuclear positions, but without the B nuclei present. Such basis functions located at fixed points in space are often referred to as **ghost orbitals**. The fragment energy for A will be lowered owing to these ghost functions, since the basis becomes more complete. The CP correction is defined as

$$\Delta E_{CP} = E(A)_{a \cup b}^* + E(B)_{a \cup b}^* - E(A)_a^* - E(B)_b^* \quad (2.8)$$

The counterpoise corrected complexation energy is given as

$$\Delta E_{BSSE} = \Delta E_{\text{complexation}} - \Delta E_{CP} \quad (2.9)$$

It should be noted that  $\Delta E_{CP}$  is an approximate correction, it gives an estimate of the BSSE effect, but it does not provide either an upper or lower limit.

It is usually observed that the CP correction for methods including electron correlation is larger and more sensitive to the size of the basis set, than that at the HF level. This is in line with the fact that the HF wave function converges much faster with respect to the size of the basis set than correlated wave functions.

There have been attempts to develop methods where the BSSE is excluded explicitly in the computational expressions; an example of this is the **chemical Hamiltonian approach (CHA)** (Mayer and Vibok 1997).

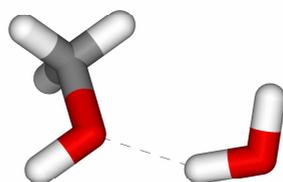
### 3 Results and Discussion – Part 1

#### Consideration of the Basis Set Superposition Error Behaviour of a Hydrogen-Bonded Water-Methanol Complex at Several Levels of Theory

The primary goal of this case study was the determination of the impact of the theory used on the magnitude of the BSSE correction energy on the calculated energies of the water-methanol dimer, where the water molecule acts as the hydrogen bond donor.

However, the quantitative description of non-bonded interactions between small polar molecules is a first step toward assessing the significance of these interactions in larger molecular systems. Thorough analysis of non-bonded interactions such as hydrogen bonding is particularly relevant in the case of CDs, where the molecular conformation depends in part on a balance between the weak forces among the hydroxyl groups within the molecule and those between the CD and the solvent. As water is the most common solvent, an accurate description of water-hydroxyl group interactions is essential for understanding the conformational properties of CDs and other carbohydrates. The smallest molecular system that can be used to model carbohydrate-water interactions is the water-methanol dimer.

The water-methanol dimer can adopt two possible configurations (WdM or MdW) depending on whether the water or the methanol acts as the hydrogen bond donor. For this study, the WdM form (Figure 2) was chosen, because it was found to be the energetically favoured configuration. (Jursic 1999).



**Fig. 2** Methanol-water dimer. The hydrogen bond is shown in black

All computational studies were performed with the GAUSSIAN 03 computational package (Frisch et al. 2004). Full geometry optimizations of the hydrogen-bonded WdM dimer were performed at the HF/3-21G, HF/6-31G(d), HF/6-31G(d,p), HF/6-31+G(d,p), HF/6-311+G(3d,2pd), DFT B3LYP/3-21G, DFT B3LYP/6-31G(d), DFT B3LYP/6-31G(d,p), DFT B3LYP/6-311+G(3d,2pd), MP2/3-21G, MP2/6-31G(d), MP2/6-31G(d,p), MP2/6-31+G(d,p), and MP2/6-311+G(3d,2pd) level of theory. The total energies were then corrected for basis set superposition error using the Boys-Bernardi counterpoise scheme, including full geometry optimization. (Boys and Bernardi 1970).

Table 5 summarizes the computed total energies of the water-methanol dimer (in Hartree) and the BSSE correction energy (in kcal/mol) in dependence of the level of theory used. From the results, it can be seen clearly that as the basis set is improved, the BSSE energy is reduced. It can also be seen that within a particular method (HF, B3LYP and MP2, respectively) the use of the 6-31G(d,p) basis set, particularly in connection with the B3LYP-DFT method, is supposed to be a reliable compromise between desired accuracy and computational expense. These data compare favourably with the findings of Del Bene and co workers who calculated binding energies for eight hydrogen bonded complexes and compared the DFT results with MP2 and experimental values. (Del Bene et al. 1995). They used Pople's 6-31G(d,p) and 6-31+G(d,p) basis sets and the B3LYP functional. Their calculated binding energies (uncorrected for BSSE) were generally close to the experimental values.

**Tab. 5** *Uncorrected, Counterpoise geometry optimized total energies, and Counterpoise BSSE-energy  $\Delta E$  of the water-methanol dimer in dependence of the level of theory used.*

Theory		Energy		
		Uncorrected (Hartree)	CP geometry optimized (Hartree)	$\Delta E$ BSSE (kcal/mol)
HF	3-21G	-190.0019759	-189.9940563	4.97
	6-31G(d)	-191.055019	-191.0534995	0.95
	6-31G(d,p)	-191.0791172	-191.0777431	0.86
	6-31+G(d,p)	-191.0918933	-191.0910227	0.55
	6-311+G(3d,2pd)	-191.1562687	-191.1558367	0.27
DFT B3LYP	3-21G	-191.07186	-191.0580488	8.67
	6-31G(d)	-192.135333	-192.1318876	2.16
	6-31G(d,p)	-192.155843	-192.1525999	2.04
	6-31+G(d,p)	-192.1788951	-192.1777012	0.75
	6-311+G(3d,2pd)	-192.2464784	-192.2459845	0.31
MP2	3-21G	-190.3376818	-190.3257763	7.47
	6-31G(d)	-191.5552944	-191.5513556	2.47
	6-31G(d,p)	-191.6132985	-191.6101104	2.00
	6-31+G(d,p)	-191.6378388	-191.6352119	1.65
	6-311+G(3d,2pd)	-191.8537978	-191.8525212	0.80

## 4 How to Conduct a Computational Molecular Modelling Study

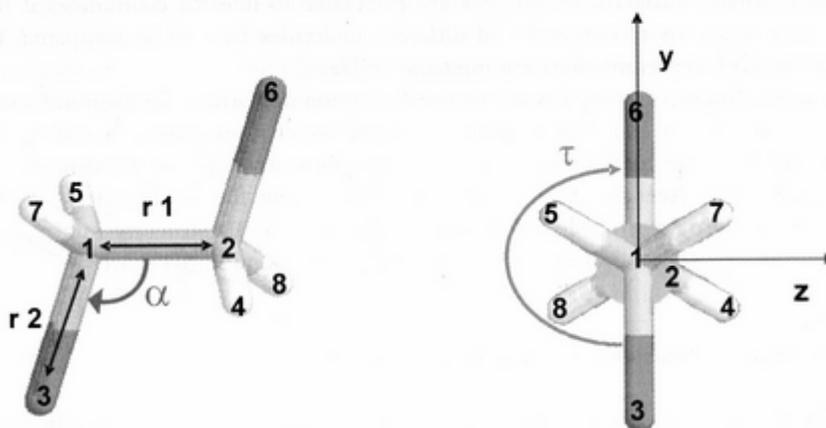
### 4.1 Basic Considerations

Molecules are real objects with a spatial extension in 3D space. In most cases, the 3D structure of a molecule is closely related to a large variety of physical, chemical, and biological properties. The actual 3D geometry of a molecule is called its conformation. Conformational isomers can be interconvert simply by rotation around rotatable bonds. Furthermore, most molecules can adopt more than one conformation of nearly equal energy content. Each of these geometries corresponds to one of the various minima of the potential energy function of the molecule, a high-dimensional mathematical expression that correlates the geometric parameters of a chemical structure with its energy content. Which of these conformations is the preferred one is heavily influenced by the interactions of the molecule with its current environment. Significantly different conformations can be observed for one and the same molecule if it is, e.g., isolated in the gas phase, influenced by solvent effects in solution, about to take part in a chemical reaction, or part of a crystal lattice in the solid state.

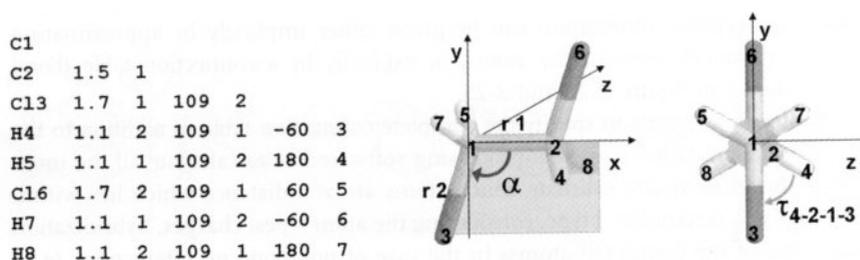
When conducting a computational molecular modelling study the first thing to do is to specify the atoms and/or molecules in the system to a modelling program. Then the program computes the energies and gradients of the energy to find the molecular geometry corresponding to the lowest energy (see chapter 4.3). However, a reasonable and reliable starting geometry is very important because it essentially determines the quality of the following investigations based on this geometry.

This can be done in two common ways. The most straightforward approach is to code the **CARTESIAN x-, y-, and z-coordinates** of each atom. Connectivity information should be given explicitly by a connection table (bond list).

The second method for representing a molecule in 3D space is to use **internal coordinates** such as bond lengths, bond angles, and torsion angles. Internal coordinates describe the spatial arrangement of the atoms relative to each other. Figure 3 illustrates this for 1, 2-dichloroethane. The most common way to describe a molecule by its internal coordinates is the so-called **Z-matrix**. Figure 4 shows the Z-matrix of 1, 2-dichloroethane.



**Fig. 3** Internal coordinates of 1, 2-dichloroethane: bond lengths  $r_1$  and  $r_2$ , bond angle  $\alpha$ , and torsion angle  $\tau$



**Fig. 4** Z-matrix of 1, 2-dichloroethane

A set of rules determines how to set up a Z-matrix properly. Each line in the Z-matrix represents one atom of the molecule. In the first line, atom 1 is defined as C1, which is a carbon atom and lies at the origin of the coordinate system. The second atom, C2, is at a distance of 1.5 Å (second column) from atom 1 (third column) and should always be placed on one of the main axes (the x-axis in Fig 4). The third atom, the chlorine atom Cl3, has to lie in the x-y-plane; it is at a distance of 1.7 Å from atom 1, and the angle between the atoms 3-1-2 is 109° (fourth and fifth columns). The third type of internal coordinate, the torsion angle or dihedral 4-2-1-3, is introduced in the fourth line of the Z-matrix in the sixth and seventh column. Except of the first three atoms, each atom is described by a set of three internal coordinates: a distance from a previously defined atom, the bond angle formed by the

atom with two previous atoms, and the torsion angle of the atom with three previous atoms. A total of  $3N-6$  internal coordinates, where  $N$  is the number of atoms in the molecule, is required to represent a chemical structure properly in 3D space. The number ( $3N-6$ ) of internal coordinates also corresponds to the number of degrees of freedom of the molecule.

Z-matrices are commonly used as input to QM (ab initio and semi-empirical) calculations as they properly describe the spatial arrangement of the atoms of a molecule. However, there is no explicit information on the connectivity present in the Z-matrix, as there is, e.g., in a connection table, but QM derives the bonding and no bonding intramolecular interactions from the molecular electronic wavefunction, starting from atomic wavefunctions and a crude 3D structure.

In contrast to that, most of the MM packages require the initial molecular geometry as 3D Cartesian coordinates plus the connection table, as they have to assign appropriate force constants and potentials to each atom and each bond in order to relax and optimize the molecular structure. Furthermore, Cartesian coordinates are preferable to internal coordinates if the spatial situations of ensembles of different molecules have to be compared. Of course, both representations are convertible into each other.

Another convenient way of obtaining suitable 3D structures is the use of X-ray crystallographic databases. However, the most important technique currently available for determining the three-dimensional structure of molecules is X-ray crystallography. The major database containing information obtained from X-ray structure analysis of small organic molecules is the **Cambridge Structural Database (CSD)**. 3D structures of macromolecules, especially proteins and nucleic acids, are found in the **Brookhaven Protein Data Bank (PDB)**.

The output of a database search in each case is a simple file containing the 3D-structural information about the molecule of interest. This data file can be read by most of the molecular modelling packages such as Gaussian. The atomic coordinates listed in the database are converted automatically to Cartesian coordinates when reading the file into the modelling program. Subsequently the structure can be displayed by molecular graphics and studied in its 3D shape.

As mentioned above, an obvious problem using computational chemistry to answer a chemical question is that one needs to know how to use the software. The difficulty sometimes overlooked is that one must estimate how accurate the answer will be in advance. In analytical chemistry, a number of identical measurements are taken and then an error is estimated by computing the standard deviation. With computational experiments, repeating the same step should always give exactly the same result (with the exception of Monte Carlo techniques). An error is estimated by comparing a number of similar computations to the experimental answers or much more rigorous computations.

The researcher may have to guess which method to use based on a fair amount of knowledge about available options: A careful researcher must know the merits and drawbacks of various methods and software packages in order to make an informed choice. It is then prudent to perform a short study to verify the method's accuracy before applying it to an unknown chemical system (see Results and Discussions Part 1 and 2 in this thesis). Appendix A provides short reviews of some useful software packages.

Tabulations of method accuracy or individual results can be found in literature (Bartlett 1985; Clark 1985; Hehre et al. 1986; Hehre 1995; Foresman and Frisch 1996; Scheiner et al. 1997; Curtiss et al. 1998; Feller and Peterson 1998; Irikura 1998; Jensen 1999).

Theoretical work is often reviewed in

- Advances in Chemical Physics
- Advances in Molecular Electronic Structure Theory
- Advances in Molecular Modelling
- Advances in Quantum Chemistry
- Annual Review of Physical Chemistry
- Recent trends in Computational Chemistry
- Reviews in Computational Chemistry

Reviews of computational work are sometimes found in

- Chemical Reviews
- Chemical Society Reviews
- Structure and Bonding

One issue of the Journal of Molecular Structure each year contains a tabulation referencing all theoretical computations published the previous year. Of this compilation, former years are available online at <http://qclodb.ims.ac.jp>

Some online databases allow the user to specify individual molecules, or select a set of molecules and then retrieve accuracy data. As an example, the computational chemistry list (CCL) consists of a list server and web site, which is at <http://server.ccl.net>. The web site contains a lot of information about computational chemistry. Subscribing to the list results in receiving about twenty messages per day. This is a good way to watch discussions of current issues. The etiquette on the list is that one should attempt to find an answer to one's question in the library and the web archives before asking that question. Once the question has been asked, one has to post a summary of the responses received.

<http://www.sciencecentral.com/> comprises a list of molecules for which various quantum chemistry methods give poor agreement with experimental results. It also recommends alternative computational methods that give better results. Moreover, it

compares experimental properties of small gas phase molecules with properties computed using a variety of ab initio methods.

In this context, some other useful sites on the internet include:

<http://srdata.nist.gov/cccbdb>

<http://www.emsl.pnl.gov/>

<http://www.rsc.org/index.asp>

## 4.2 Computer Graphics

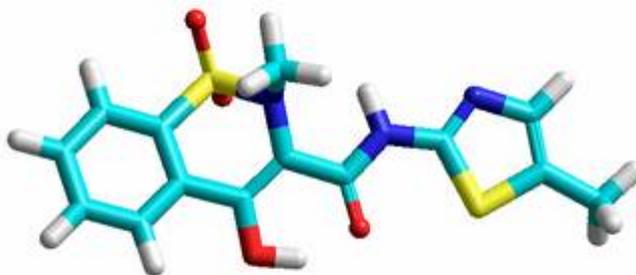
### 4.2.1 Basic Considerations

Since the early 20th century, chemists have represented molecular information by molecular models, because visualization makes complex information accessible to human understanding easily and directly with images. Diverse methods can be used for the display of a molecular model. The most widely used representations of molecules are 2D structure diagrams, which arrange the atoms in 2D space in such a way that bond lengths and bond angles are shown as undisturbed as possible and avoiding the overlap of atoms. In addition, stereo chemical information may be given through wedged and hashed bonds. However, no 2D structure representation can explain a 3D molecule in its entirety. The 3D structure information is especially important for an understanding of the chemical and biological properties of a compound and the relationships between structural features and molecular functions. Often, the physical and chemical properties of a molecule are more evident when one changes between different display styles.

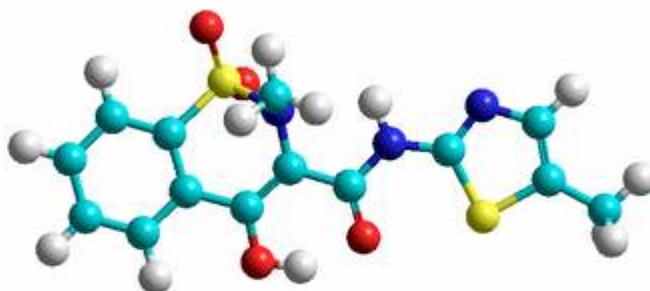
Computer graphics has had a dramatic impact upon molecular modelling. The interaction between molecular graphics and the underlying theoretical methods has enhanced the accessibility of molecular modelling methods and assisted the analysis and interpretation of such calculations.

### 4.2.2 Visualization of Molecular Models

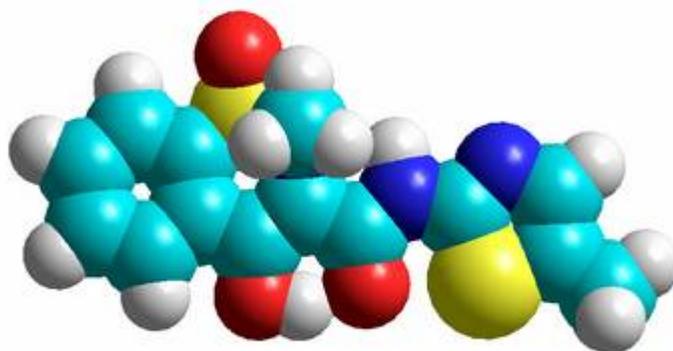
Molecules are most commonly represented on a computer graphics screen using “stick” or “space filling” representations, which are analogous to the Dreiding and Corey-Pauling-Koltun (CPK) mechanical models. Several variations on these two basic types have been developed, such as the ability to colour molecules by atomic number and the inclusion of shading and lighting effects, which give models a more realistic appearance. Some of commonly used molecular representations are shown in Figures 5-7.



*Fig. 5 Tube model of meloxicam*



*Fig. 6 Ball and stick model of meloxicam*

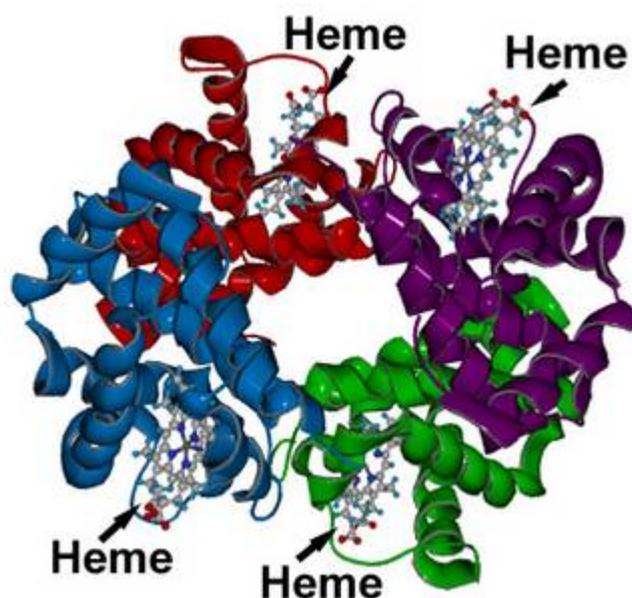


*Fig. 7 Corey-Pauling-Koltun (CPK) model of meloxicam*

Computer-generated models do have some advantages when compared with their mechanical counterparts. Of particular importance is the fact that a computer model can be very easily interrogated to provide quantitative information, from simple geometrical measures such as the distance between two atoms to more complex quantities such as the energy or surface area. Quantitative information such as this can be very difficult if not impossible to obtain from a mechanical model. Nevertheless, mechanical models may still be preferred in certain types of situation due to the ease with which they can be manipulated and viewed in three dimensions. A computer screen is inherently two-dimensional, whereas molecules are three-dimensional objects. Nevertheless, some impression of the three-dimensional nature of an object can be represented on a computer screen using techniques such as depth cueing (in which those parts of the object that are further away from the viewer are made less bright) and through the use of perspective. Specialised hardware enables more realistic three-dimensional stereo images to be viewed. In the future „virtual reality“- systems may enable a scientist to interact with a computer-generated molecular model in much the same way that a mechanical model can be manipulated.

Even the most basic computer graphics program provides some standard facilities for the manipulation of models, including the ability to translate, rotate and „zoom“ the model towards and away from the viewer. Packages that are more sophisticated can provide the scientist with quantitative feedback on the effect of altering the structure. For example, as a bond is rotated then the energy of each structure could be calculated and displayed interactively.

However, the visualization of hundreds or thousands of connected atoms, which are found in biological macromolecules such as proteins, is no longer reasonable with the molecular models described above. This problem can be solved with a simplified model, which serves primarily to represent the secondary structure of the protein or nucleic acid backbone: to produce a clearer picture the explicit representation of any atoms is dispensed and the molecule is presented using a „ribbon“(see Figure 8).



*Fig. 8 3-D structure of hemoglobin with the subunits displayed in the ribbon representation, which traces the backbone atoms of a protein. The four heme groups are displayed in the ball-and-stick representation. Regular structure elements (i.e.  $\alpha$ -helices and  $\beta$ -sheets) are highlighted in green and purple, respectively.*

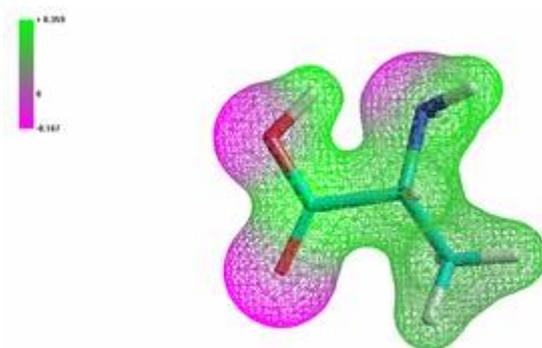
### 4.2.3 Visualization of Molecular Properties

Knowledge of the spatial dimensions of a molecule is insufficient to understand the details of complex molecular interactions. In fact, molecular properties such as electrostatic potential, hydrophilic/lipophilic properties, and hydrogen bonding ability should be taken into account.

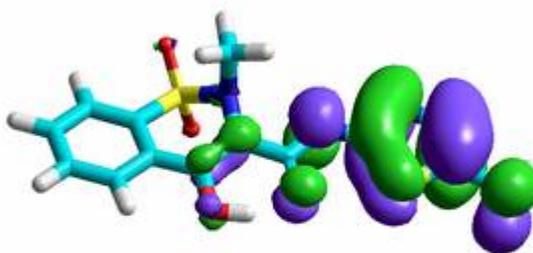
**Molecular orbitals** were one of the first molecular features that could be visualized with simple graphical hardware. The reason for this early representation is found in the complex theory of quantum chemistry. A structure is more attractive and easier to understand when orbitals are displayed, rather than numerical orbital coefficients. The molecular orbitals, calculated by semi-empirical or ab initio quantum mechanical methods, are represented by isosurfaces, corresponding to the electron density surfaces. Knowledge of molecular orbitals, particularly of the

HOMO (Highest Occupied Molecular Orbital) and the LUMO (Lowest Unoccupied Molecular Orbital), imparts a better understanding of reactions (Figure 10). Different colours (e.g. violet and green) are used to distinguish between the parts of the orbital that have opposite signs of the wavefunction.

Besides molecular orbitals, other molecular properties, such as the **molecular electrostatic potential (MEP)**, which is clearly the most important and most used property, can be represented by isovalue surfaces (Figure 9). This type of high-dimensional visualization permits fast and easy identification of the relevant molecular regions.



*Fig. 9* MEP of D-alanine; positive and negative areas are depicted in green and red, respectively

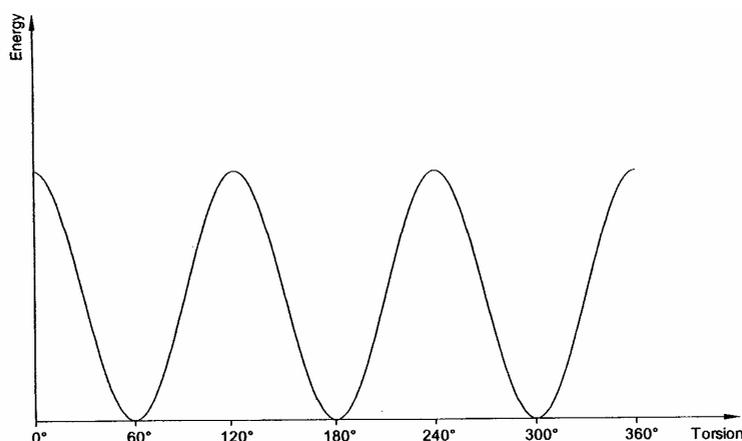


*Fig. 10* HOMO of meloxicam (the different colours represent opposite signs of the wave function).

## 4.3 Geometry Optimization (Energy Minimization)

### 4.3.1 The Molecular Potential Energy Surface (PES)

Almost certainly, the generated 3D model of a given molecule does not have ideal geometry; therefore, a geometry optimization must be performed subsequently. The way the energy of a molecular system varies with small changes in its structure is specified by its **potential energy surface (PES)**. A potential energy surface is a mathematical relationship linking molecular structure and the resultant energy. For a diatomic molecule, the PES is a two-dimensional plot. As an example, Figure 11 depicts the sine-like potential energy curve of ethane shown as function of the dihedral angle:



**Fig. 11** Sine-like potential energy curve of ethane shown as function of the dihedral angle (taken from Höltye et al. 2003).

For larger systems, the surface generally has as many dimensions as there are degrees of freedom within the molecule. Thus, for a system with  $N$  atoms the energy is a function of  $3N-6$  internal or  $3N$  Cartesian coordinates. It is therefore impossible to visualize the entire energy surface except for some simple cases where the energy is a function of just one or two coordinates. Such a PES is often represented by illustrations like that one in Figure 12.

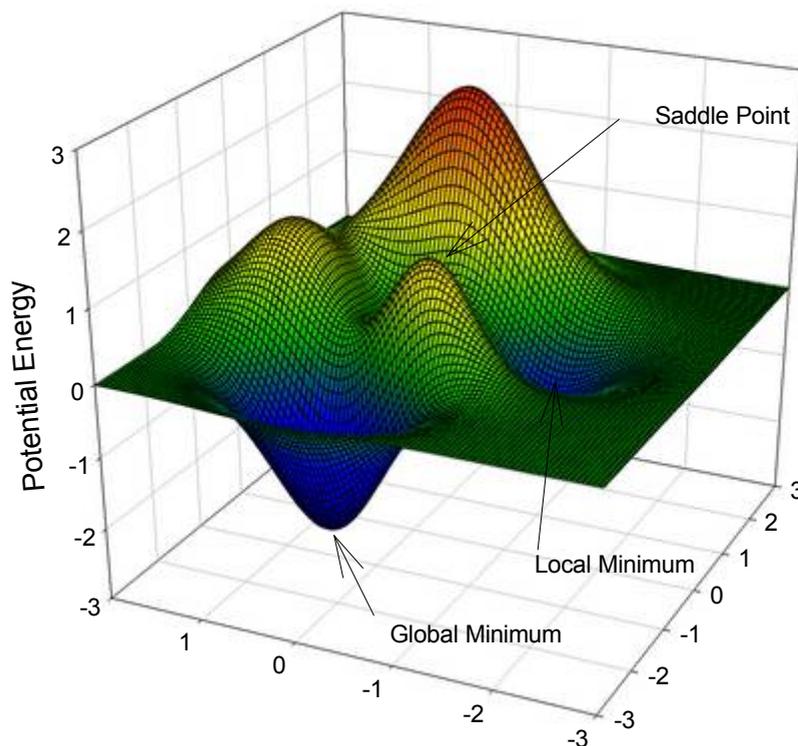


Fig. 12 Potential energy surface including global and local minimum and saddle point

This sort of drawing considers only two of the degrees of freedom within the molecule, and plots the energy above the plane defined by them, creating a literal surface. Each point corresponds to the specific values of the two structural variables and thus represents a particular molecular structure - with the height of the surface at that point corresponding to the energy of that structure.

In molecular modelling one is especially interested in minimum points on the energy surface. Minimum energy arrangements of the atoms correspond to stable states of the system; any movement away from a minimum gives a configuration with a higher energy. There may be a very large number of minima on the energy surface. The minimum with the very lowest energy is known as the **global energy minimum**.

The highest point on the pathway between two minima is of especial interest and is known as the **saddle point** (see Figures 12 and 13), with the arrangement of the atoms being the **transition structure**.

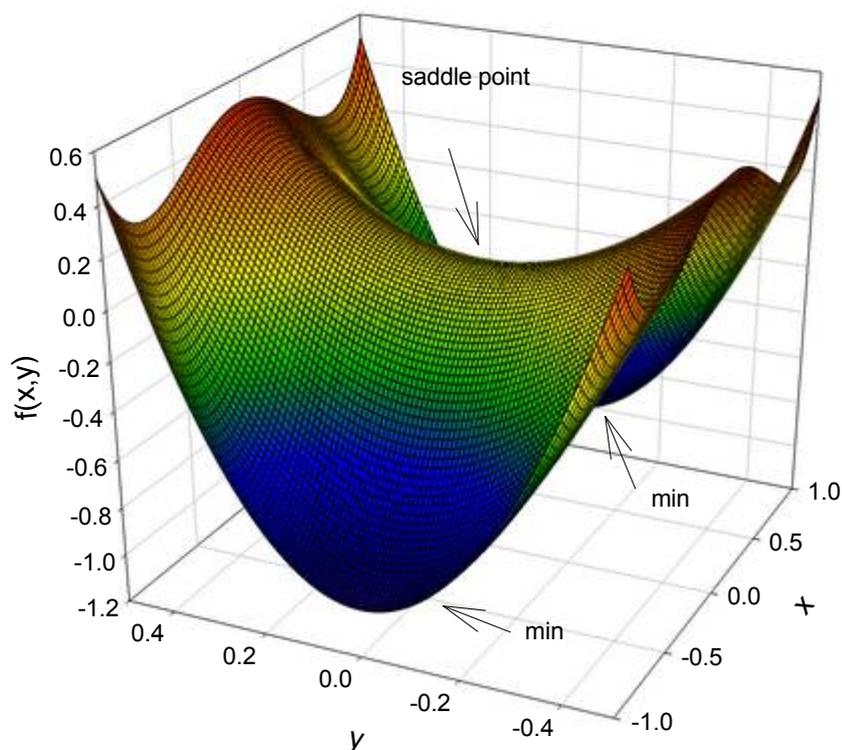


Fig.13 PES with a saddle point at  $(0, 0)$  and connecting two minima at  $(1, 0)$  and  $(-1, 0)$

Both minima and saddle points are **stationary points** on the energy surface, where the first derivative of the energy function is zero with respect to all the coordinates.

### 4.3.2 Characterization of Stationary Points

For a linear molecule with  $N$  Atoms, there are in total  $3N$  degrees of freedom. Three of these are translational degrees of freedom, three are rotational (two for a linear molecule) and the remaining  $p = 3N - 6$  ( $3N - 5$  if linear) are vibrational degrees of freedom. The molecular potential energy  $U$  will then depend on these  $3N - 6$  ( $3N - 5$  if linear) independent variables, which are known as the **internal coordinates**. It is traditional to denote such variables by the letter  $q(q_1, q_2, \dots, q_p)$ . The internal variables can be collected into the column vector  $\vec{q}$  (a matrix of dimension  $p \times 1$ ):

$$\vec{q} = \begin{pmatrix} q_1 \\ q_2 \\ \dots \\ q_p \end{pmatrix} \quad (4.1)$$

The **gradient**  $\vec{g}$  of a potential function  $U$ ,  $\vec{g} = \text{grad} U$  is defined as a column vector

$$\vec{g} = \begin{pmatrix} \frac{\partial U}{\partial q_1} \\ \frac{\partial U}{\partial q_2} \\ \dots \\ \frac{\partial U}{\partial q_p} \end{pmatrix}, \quad (4.2)$$

and it's **Hessian**  $\vec{H}$  as a symmetric square matrix

$$\vec{H} = \begin{pmatrix} \frac{\partial^2 U}{\partial q_1^2} & \dots & \frac{\partial^2 U}{\partial q_1 \partial q_p} \\ \dots & \dots & \dots \\ \frac{\partial^2 U}{\partial q_p \partial q_1} & \dots & \frac{\partial^2 U}{\partial q_p^2} \end{pmatrix} \quad (4.3)$$

As mentioned above, the molecular potential energy  $U$  depends on the variables  $q_p$ , that is,  $U(\vec{q})$ .

At a stationary point, the gradient is zero. In order to characterize the stationary point, one has to find the eigenvalues of the Hessian calculated at that point. If the eigenvalues are all positive, then the point is a **minimum**. If the eigenvalues are all negative, then the point is a **maximum**. Otherwise, the point is a **saddle point**. Saddle points of interest to chemists are those that are a minimum in all degrees of freedom except one, and the Hessian has just one negative eigenvalue at such a point. They are sometimes called **first-order saddle points**.

### 4.3.3 Energy Minimization Methods

To identify those geometries of the system that correspond to minimum points on the PES certain **minimization algorithms** are used.

The minimisation problem can be formally stated as follows: given a function  $f$  which depends on one or more independent variables  $x_1, x_2, \dots, x_i$ , find the values of those variables where  $f$  has a minimum value. At a minimum point the first derivative of the function with respect to each of the variables, i.e., the gradient is zero and the second derivatives are all positive:

$$\frac{\partial f}{\partial x_i} = 0 \quad \frac{\partial^2 f}{\partial x_i^2} > 0; \quad (4.4)$$

The functions of most interest in this context will be the QM or MM energy with the variables  $x_i$  being the Cartesian or the internal coordinates of the atoms.

MM minimisations are nearly always performed in Cartesian coordinates, where the energy is a function of  $3N$  variables.

In QM, it is more common to use internal coordinates (as defined in the  $Z$  - matrix).

For analytical functions, the minimum of a function can be found using standard calculus methods. However, this is not generally possible for molecular systems due to the complicated way in which the energy varies with the coordinates. Rather, minima are located using numerical methods, which gradually change the coordinates to produce configurations with lower and lower energies until the minimum is reached.

Minimization algorithms can be classified into two groups: those, which use derivatives of the energy with respect to the coordinates, and those, which do not. Derivatives can be useful because they provide information about the shape of the energy surface, and, if used properly, they can significantly enhance the efficiency with which the minimum is located.

There are many factors that must be taken into account when choosing the most appropriate algorithm (or combination of algorithms) for a given problem; the ideal minimisation algorithm is the one that provides the answer as quickly as possible, using the least amount of computer resources. A detailed discussion of an efficient use of computer resources is given by Young (2001).

No single minimisation method has yet proved to be the best for all molecular modelling problems and so most software packages offer a choice of methods.

In particular, a method that works well with QM may not be the most suitable for use with MM. This is partly because QM is usually used to model systems with fewer atoms than MM; some operations that are integral to certain minimisation procedures (such as matrix inversion) are trivial for small systems but formidable for systems containing thousands of atoms. QM and MM also require different amounts of computational effort to calculate the energies and the derivatives of the various

configurations. Thus, an algorithm that takes many steps may be appropriate for MM but inappropriate for QM.

Most minimisation algorithms can only go downhill on the energy surface and so they can only locate the minimum that is nearest (in a downhill sense) to the starting point. Thus, Figure 14 shows a schematic energy surface and the minima that would be obtained starting from three points A, B, and C.

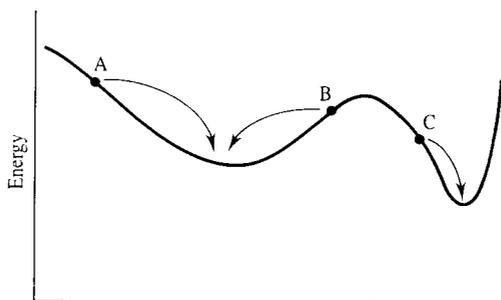


Fig. 14 Schematic energy surface with starting points A, B, and C, respectively

The minima can be considered to correspond to the locations where a ball rolling on the energy surface under the influence of gravity would come to rest. To locate more than one minimum or to locate the global energy minimum one therefore usually requires a means of generating different starting points, each of which is then minimised (see chapter 4.3.4.: Conformational Searching).

Some specialised minimisation methods can make uphill moves to seek out minima lower in energy than the nearest one, but no algorithm has yet proved capable of locating the global energy minimum from an arbitrary starting position.

The shape of the PES may be important if one wishes to calculate the relative populations of the various minimum energy structures. For example, a deep and narrow minimum may be less highly populated than a broad minimum that is higher in energy as the vibrational energy levels will be more widely spaced in the deeper minimum and so less accessible. For this reason, the global energy minimum may not be the most highly populated minimum. In any case, the “active” structure (e.g. the biologically active conformation of a drug molecule) may correspond to the global minimum or to the most highly populated conformation even to a minimum energy structure at all.

The input to a minimisation program consists of a set of initial coordinates for the system. The initial coordinates may come from a variety of sources. They may be obtained from an experimental technique, such as *X*-ray crystallography or NMR. In other cases, a theoretical method is employed, such as a conformational search algorithm. A combination of experimental and theoretical approaches may also be used. For example, to study the behaviour of a protein in water one may take a *X*-

ray structure of the protein and immerse it in a solvent “bath”, where the coordinates of the solvent molecules have been obtained from a Monte Carlo or molecular dynamics simulation.

In the following, some of the minimization methods most often used in molecular modelling are considered.

**The steepest descent minimizer** uses the numerically calculated first derivative of the energy function to approach the energy minimum. The energy is calculated for the initial geometry and then again, when one of the atoms has been moved in a small increment in one of the directions of the coordinate system. This process will be repeated for all atoms, which finally are moved to new positions downhill on the energy surface. The procedure will stop if the predetermined minimum condition is fulfilled. The optimization process is slow near the minimum, so the steepest descent method is often used for structures far from the minimum. It is the method most likely to generate low-energy structures of poorly refined crystallographic data or to relax graphically built molecules. In most cases, the steepest descent minimization is used as a first rough and introductory run followed by a subsequent minimization employing a more advanced algorithm like conjugate gradients.

The **conjugate gradients method** accumulates the information about the function from iteration to the next. With this proceeding, the reverse of the progress made in an earlier iteration can be avoided. For each minimization step, the gradient is calculated and used as additional information for computing the new direction vector of the minimization procedure. Thus, each successive step continually refines the direction towards the minimum. The computational effort and the storage requirements are greater than for steepest descent but conjugate gradients is the method of choice for larger systems. The greater total computational expense and the longer time per iteration is more than compensated by the more efficient convergence to the minimum achieved by conjugate gradients.

The **Powell method** is very similar to conjugate gradients. It is faster in finding convergence and is suitable for a variety of problems, but one should be careful when using the Powell algorithm because torsion angles may sometimes be modified to a dramatic extent. Therefore, the Powell method is not practicable for energy minimization after a conformational analysis because the located low-energy conformations will be altered in an undesired manner. It is advisable to perform a conjugate gradient minimization in this situation.

The **Newton-Raphson minimizer** as a second-derivative method uses, in addition to the gradient, the curvature of the function to identify the search direction. The second derivative is also applied to predict where the function passes through a minimum. The efficiency of the Newton-Raphson method increases as convergence is approached. The computational effort and the storage requirements for calculating larger systems are disadvantages of this method. For structures with high strain, the minimization process can become unstable, so the application of this algorithm is

mostly limited to problems where rapid convergence from a preoptimized geometry to an extremely precise minimum is required.

One of the most efficient quasi-Newton algorithms is the **Berny algorithm**, which internally builds up a second derivative Hessian matrix. In practise, it is usual to make an estimate and update these estimates at each of the iterations. A great strength of this method is that construction of the initial Hessian can be based on chemical ideas.

A number of factors, including the storage and computational requirements, the relative speeds with which the various parts of the calculation can be performed, the availability of analytical derivatives and the robustness of the method, dictate the choice of minimization algorithm. Thus, any method that requires the Hessian matrix to be stored may present memory problems when applied to systems containing thousands of atoms. Calculations on systems of this size are invariably performed using MM, and so the steepest descent and the conjugate gradients methods are very popular here. For MM calculations on small molecules, the Newton-Raphson method may be used, although this algorithm can have problems with structures that are far from a minimum. For this reason, it is usual to perform a few steps of minimisation using a more robust method such as the steepest descent method before applying the Newton-Raphson algorithm. Analytical expressions for both first and second derivatives are available for most of the terms found in common force fields.

Quantum mechanical calculations are restricted to systems with relatively small numbers of atoms, and so storing the Hessian matrix is not a problem. As the energy calculation is often the most time-consuming part of the calculation, it is desirable that the minimisation method chosen takes as few steps as possible to reach the minimum. For many levels of QM theory, analytical first derivatives are available. However, analytical second derivatives are only available for a few levels of theory and can be expensive to compute.

In any optimization problem, the choice of coordinates can have an important influence on the efficiency of the optimization.

One way of define the geometry of a molecule is by using a **Z-matrix**. (Constructing a Z-matrix is discussed in detail in chapter 4.1).

For many systems the Z-matrix can often be written in many different ways as there are many combinations of internal coordinates. There should be no strong coupling between the coordinates. **Dummy atoms** can often help in the construction of an appropriate Z-matrix. A dummy atom is used solely to define the geometry and has no nuclear charge and no basis functions.

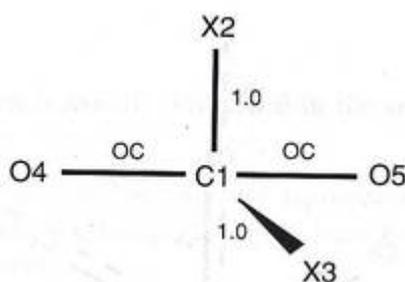
However, many optimization algorithms run into trouble when trying to optimize geometries having bond angles of  $180^\circ$ . If an angle of  $180^\circ$  is specified, then the dihedral angle referenced to that angle will be mathematically undefined.

To illustrate the use of dummy atoms, here is a carbon dioxide input with two dummy atoms, which is also shown in Figure 15:

```

C
X 1 1.0
X 1 1.0 2 90.0
O 1 OC 2 90.0 3 90.0
O 1 OC 2 90.0 3 -90.0
OC 1.2

```



*Fig. 15 Illustration of the geometry formed from the CO<sub>2</sub> Z -matrix example*

Strong coupling between coordinates can give long “valleys” in the energy surface, which may also present problems. Thus, care must be taken when defining the Z-matrix for cyclic systems in particular. The natural way to define a cyclic compound would be to number the atoms sequentially around the ring. However, this would then mean that the ring closure bond would be very strongly coupled to all of the other bonds, angles, and torsion angles. Therefore, molecules with rings should always be given a dummy atom in the centre of the ring. The atoms in the ring should then be referenced to the central dummy atom rather than each other.

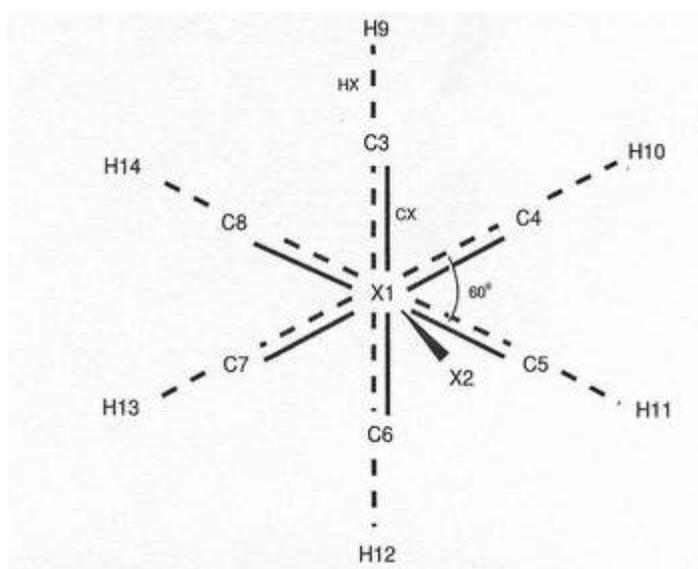
Moreover, it is often convenient to use two dummy atoms: one in the centre of the ring and one perpendicular to the ring. As an example, Figure 16 depicts a Z -matrix for a benzene molecule and the geometry formed from that Z-matrix:

```

X
X 1 1.0
C 1 CX 2 90.0
C 1 CX 2 90.0 3 60.0
C 1 CX 2 90.0 4 60.0
C 1 CX 2 90.0 5 60.0
C 1 CX 2 90.0 6 60.0
C 1 CX 2 90.0 7 60.0
H 1 HX 2 90.0 3 0.0
H 1 HX 2 90.0 4 0.0
H 1 HX 2 90.0 5 0.0
H 1 HX 2 90.0 6 0.0
H 1 HX 2 90.0 7 0.0
H 1 HX 2 90.0 8 0.0

CX 1.3
HX 2.3

```



*Fig. 16 Illustration of the geometry formed from the benzene Z-matrix example*

Another way to define the geometry of a molecule is a set of **Cartesian coordinates** for each atom. Graphic interface programs often generate Cartesian coordinates since this is the most convenient way to write those programs.

Though it is still possible to input a geometry manually in an ASCII input file into a computational chemistry program, it is becoming more common to use programs that have a graphical builder in which the user can essentially draw the molecule. There are several ways in which such programs work. Some programs allow the molecule to be built as a two-dimensional stick structure and then convert it into a three-dimensional structure. Some programs have the user draw the three-

dimensional backbone and then automatically add the hydrogen atoms. This works well for organic molecules. Many programs include a library of commonly used functional groups, which is convenient if it has the functional groups needed for a particular project. A number of programs have specialized building modes for certain classes of molecules, such as proteins, nucleotides, or carbohydrates (see Appendix A).

However, the way in which geometry was specified is not necessarily the coordinate system that will be used by the algorithm, which optimizes the geometry. For example, it is very simple for a program to convert a *Z*-matrix into Cartesian coordinates and then use that space for the geometry optimization.

Many *ab initio* and semi-empirical programs optimize the geometry of the molecule by changing the parameters in the *Z*-matrix. In general, this can be a very good way to change the geometry because these parameters correspond to molecular motions similar to those seen in the vibrational modes. However, if the geometry is specified in such a way that changing one of the parameters slightly could result in a large distortion to some portion of the molecule, then the geometry optimization is less efficient.

Thus, a poorly constructed *Z*-matrix can result in a very inefficient geometry optimization. Many computational chemistry programs will do the geometry optimization in **Cartesian coordinates**. This is often the only way to optimize geometry in MM programs and an optional method in orbital-based programs.

A Cartesian coordinate optimization may be more efficient than a poorly constructed *Z*-matrix. This is often seen in ring systems, where a badly constructed *Z*-matrix will perform very poorly (see above). Cartesian coordinates are often preferable when simulating more than one molecule since they allow complete freedom of motion between separate molecules.

In order to have the advantages of a well-constructed *Z*-matrix, regardless of how the geometry was defined, a system called **redundant internal coordinates** (Peng et al. 1996) was created. When redundant internal coordinates are used, the input geometry is first converted to a set of Cartesian coordinates. The algorithm then checks the distances between every pair of atoms to determine which are within a reasonable bonding distance. The program then generates a list of atom distances and angles for nearby atoms. This way, the algorithm does the job of constructing a sort of *Z*-matrix that has more coordinates than are necessary to specify the geometry completely. This is usually the most efficient way to optimize geometry.

Packages such as Gaussian offer a choice between *Z*-matrix, Cartesian coordinates, and redundant internal coordinate optimizations.

An optimization is complete when it has **converged**: essentially, when the forces are zero, the next step is very small, below some preset value defined by the algorithm, and some other conditions are met.

However, a geometry optimization alone cannot tell one about the nature of the stationary point that it finds. In order to characterize a stationary point, it is necessary to perform a **frequency calculation** on the optimized geometry. Electronic structure programs like Gaussian are able to carry out such calculations.

The completed frequency calculation will include a variety of results: frequencies, intensities, the associated normal modes, the structure's zero point energy (ZPE), and various thermo chemical properties. If any of the frequency values are less than zero, these frequencies are known as **imaginary frequencies**. The number of imaginary frequencies indicates the type of stationary point to which the given molecular structure corresponds. By definition, a structure, which has  $n$  imaginary frequencies, is referred to as  **$n^{\text{th}}$  order saddle point**. Thus, the minimum will have zero imaginary frequencies, and an ordinary transition structure will have one imaginary frequency since it is a first order saddle point. First order saddle points-which are a maximum in exactly one direction and a minimum in all other orthogonal directions-correspond to transition state structures linking two minima (see Figure 12).

In order to distinguish a local minimum from the global minimum, it is necessary to perform a **conformational search** (see chapter 4.3.4).

#### 4.3.4 Conformational Searching

Generally, the PES of a molecular cluster contains a large number of energy minima. These minima can be theoretically ascribed only if the whole PES is known. However, the number of energy minima increases very rapidly with the cluster size and the number of subsystems. The localization of every minimum is tedious, if not impossible, by standard methods based on experience and chemical intuition. Experimental techniques such as NMR only provide information on one or a few conformations of a molecule. A complete and exclusive overview of the conformational potential of molecules can be gained by theoretical techniques. Correspondingly a variety of theoretical methods for conformational analysis has been developed. Conformational energies can be calculated either using QM or MM methods. Because the quantum mechanical calculations are very time consuming, they cannot be applied to large or flexible molecules. For that reason, most of the conformational search programs use MM methods for the calculation of energies as a standard. In the following, some commonly used techniques for conformational analyses are described.

##### 4.3.4.1 Systematic Search Procedure

This method comprises perhaps the most natural of all different conformational analysis methods. It is performed by varying systematically each of the torsion angles in a molecule in order to generate all possible conformations. If the angle increment is appropriately small, the procedure yields a complete image of the PES of a molecule.

The number of generated conformations is

$$\left( \frac{360}{\text{step size}} \right)^n \quad (4.5)$$

where  $n$  is the number of rotatable bonds.

If for example a systematic conformational search is performed for a molecule with six rotatable bonds and a step size of  $30^\circ$  is employed, the number of generated conformers amounts to  $12^6$  or 2 985 984 structures. This huge amount of data cannot be handled properly; it therefore has to be reduced by appropriate methods.

#### 4.3.4.2 Monte Carlo Methods (MC)

MC is a random search technique and therefore of a statistical nature. At each stage of a MC search, the actual conformation is modified randomly in order to obtain a new one.

A random search starts with an optimized structure. At each of the iterations, new torsion angles or new Cartesian coordinates are assigned randomly. The resulting conformation is minimized using MM and the randomization process is repeated. The minimized conformation is then compared with the previously generated structures and is only stored if it is lower in energy than the current.

The random methods potentially cover all regions of conformational space, but this is only true if the process is allowed to run for a sufficiently long time. This may last extremely long because the probability to detect a new and unique conformation decreases dramatically depending on the growing number of conformers already discovered. However, even if the computation has been running very long, one cannot be certain that the conformational space has been completely covered. It is very important therefore to establish a means for testing the completeness of the analysis, for instance by performing several runs in a parallel mode, each one starting with a different initial conformation. If the results are identical or nearly identical, then completeness can be assumed.

The main advantage of random search methods is that, in principle, molecules of any size can be successfully treated. In practice, however, highly flexible molecules often do not give converging results, because the volume of the respective conformational space is too large.

An important variation of the MC method is the **Metropolis Approach**. While in MC, the new geometry after a randomly chosen trial move for the system is accepted as a starting point for the next step if it is lower in energy than the current, in the Metropolis procedure the energy of the new geometry is allowed to be higher, too. In this case, the Boltzmann factor  $e^{-\frac{\Delta E}{kT}}$  is calculated and compared to a random number

between zero and one. If  $e^{-\frac{\Delta E}{kT}}$  is less than this number the new geometry is accepted, otherwise the next step is taken from the old geometry.

Thus, the Metropolis method ensures that the configurations in the ensemble obey a Boltzmann distribution, and the possibility of accepting higher energy configurations allows climbing uphill on the PES and escaping from a local minimum.

#### 4.3.4.3 Molecular Dynamics (MD)

The aim of the MD approach is to reproduce the time-dependent motional behaviour of a molecule. MD is based on MM. It is assumed that the atoms in the molecule interact with each other according to the rules of the employed force field. At regular time intervals, the classical equation of motion represented by Newton's second law is solved:

$$F_i(t) = m_i a_i(t) \quad (4.6)$$

where  $F_i$  is the force on atom  $i$  at time  $t$ ,  $m_i$  is the mass of atom  $i$ , and  $a_i(t)$  is the acceleration of atom  $i$  at time  $t$ .

The gradient of the potential energy function is used to calculate the forces on the atoms while the initial velocities on the atoms are generated randomly at the beginning of the dynamics run. Based on the initial atom coordinates of the system, new positions and velocities on the atoms can be calculated at time  $t$  and the atoms will be moved to these new positions. Because of this, a new conformation is created. The cycle will then be repeated for a predefined number of time steps. The collection of energetically accessible conformations obtained by this procedure is called an **ensemble**.

Unlike the conservative geometry optimization procedures, MD is able to overcome energy barriers between different conformations. Therefore it should be possible to find local minima other than the nearest in the potential energy surface. However, if the energy barrier is high or the number of degrees of freedom in the molecule is very large, then some of the existing conformers of the investigated system possibly are not reached. Moreover, in view of the huge conformational space the completeness of the conformational search during the chosen simulation time is difficult to ensure.

To enhance conformational sampling an elevated temperature is added to the simulation. At high temperature, the molecule is able to overcome even large energy barriers that may exist between some conformations and therefore the chance for completeness of a conformation search increases.

#### 4.3.4.4 Simulated Annealing

It is a modification of the high-temperature MD. In this technique, the system is cooled down at regular time intervals by decreasing the simulation temperature. As the temperature approaches zero K, the molecule is trapped in the nearest local minimum conformation. The received geometry at the end of the annealing cycle is saved and subsequently used as starting point for further simulations at higher temperature. In order to obtain a set of low-energy conformations the cycle will be repeated several times.

To sum up: it may be stated that these methods considered above represent an additional and very valuable tool that can be used to sample the conformational space, especially when other conformational search methods have been unsuccessful. However, one should be careful when selecting the appropriate method and in setting the simulation conditions in order to ensure the completeness of the conformational search and the validity of the results. It should also be kept in mind that each approach has its strengths and its weaknesses and therefore, wherever possible, experimentally derived data should serve as verification.

### 4.4 Choosing the Best Computational Method

In general, real modelling research studies with computational methods involve multiple calculations to investigate thoroughly the system of interest. To predict the total energy of a system, the following procedure might be appropriate:

- Perform a low-level geometry optimization with a medium-sized basis set, for example, a HF calculation with the 6-31G (d,p) basis set. (For large systems, e.g., CDs, a smaller basis set might be necessary.)
- Run a frequency job at the optimized geometry, using the same method and basis set, to ensure the nature of the stationary points found and to calculate the zero point or thermal energy.
- If it is computationally feasible, improve on the structure by using it as the starting point for a more accurate optimization (using a larger basis set and/or run at a higher level of theory).
- Run a very high-level single point energy calculation - for example, MP2 - at the newly optimized structure, using a large basis set. Final energies should be computed at the most accurate model chemistry (see below) that is practical for the system under consideration.
- Once these steps are complete, the energy can then be computed as the sum of the final single point energy and the zero point or thermal energy, as appropriate.

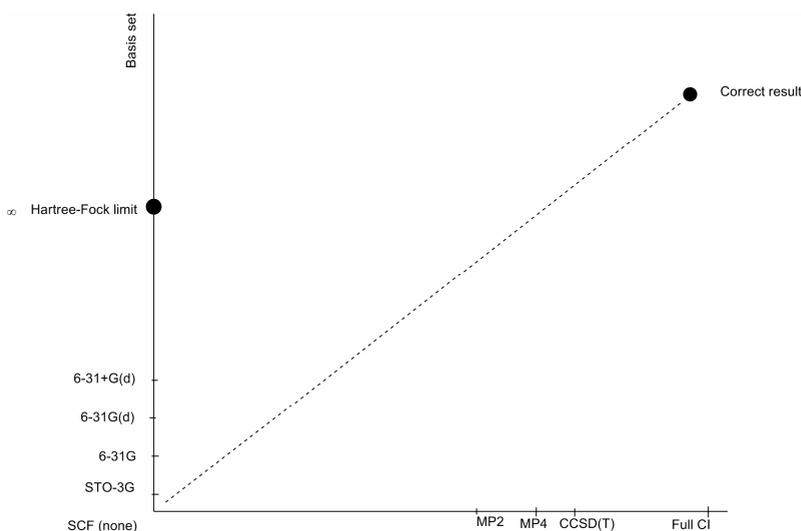
### 4.4.1 Model Chemistries

Model chemistry is an unbiased, uniquely defined, and uniformly applicable theoretical model for predicting the properties of chemical systems. It generally consists of the combination of a theoretical method with a basis set. Each such unique pairing of method with basis set represents a different approximation to the Schrödinger equation. Results for different systems generally may only be compared when they have been predicted via the same model chemistry. Different model chemistries may be tested by comparing their results on the same systems to one another and to the results of experiments.

The situation can be illustrated by a “Pople diagram” (Pople 1965) as shown in Figure 17. Different levels of ab-initio theory are represented on two axes. The vertical one gives the size of the basis set and the horizontal axis the correlation treatment. The level of correlation increases as one moves to the right, with the Hartree-Fock SCF method at the extreme left (including no correlation), and the Full Configuration Interaction method at the right (which fully accounts for electron correlation). In general, computational cost and accuracy increase as one moves to the right as well.

However, the diagram shows that one can never reach the correct result (top right-hand corner) simply by extending the basis set in SCF calculations. This would only take one to the top left-hand corner, which is marked “HF limit” where no further improvement in the energy can be obtained by basis set improvements.

Choosing specific model chemistry usually involves a trade-off between accuracy and computational cost. More accurate methods and larger basis sets make jobs run longer.



**Fig. 17** The „Pople diagram“. The vertical axis gives the size of the basis set and the horizontal axis the correlation treatment. Their positions on the axes (but not the order) are arbitrary.

Table 6 summarizes some of the most commonly used basis sets and provides some recommendations as to when each is appropriate.

**Tab. 6** Commonly used Standard Basis Sets

Basis Set	Description
STO-3G	Minimal basis set (stripped down in the interest of performance): use for more qualitative results on very large systems when one cannot afford even 3-21G
3-21G	Two sets of functions in the valence region provide a more accurate representation of orbitals. Use for very large molecules for which 6-31G(d) is too expensive
6-31G(d)	Adds polarization functions to heavy atoms: use for most jobs on up to medium/large sized systems
6-31G(d,p)	Adds polarization functions to the hydrogen atoms as well: use when the hydrogen atoms are the site of interest (for example, bond energies) and for final, accurate energy calculations
6-31+G(d)	Adds diffuse functions: important for systems with lone pairs, anions, excited states
6-31+G(d,p)	Adds extra valence functions (3 sizes of s and p functions) to 6-31+G(d). Diffuse functions can also be added to the hydrogen atoms via a second +

## 4.4.2 Selecting a Suitable Calculation Method

In this chapter, the strengths and weaknesses of various specific methods are considered.

### 4.4.2.1 Semi-Empirical Methods

Semi-empirical methods are characterized by their use of parameters derived from experimental data in order to simplify the approximation to the Schrödinger equation. As such, they are relatively inexpensive and can be practically applied to very large molecules. There is a variety of semi-empirical methods. Among the best-known are AM 1, PM3 and MNDO. Semi-empirical methods are appropriate for a variety of modelling tasks, including the following:

- For very large systems for which they are the only computationally practical quantum mechanical methods.
- As a first step for large systems. For example, one might run a semi-empirical optimization on a large system to obtain a starting structure for a subsequent HF or DFT optimization.
- For ground state molecular systems for which the semi-empirical method is well parameterized and well calibrated. In general, semi-empirical methods have been developed to focus on simple organic molecules.
- To obtain qualitative information about a molecule, such as its molecular orbitals, atomic charges or vibrational normal modes. In some cases, semi-empirical methods may also be successfully used to predict energy trends arising from alternate conformations or substituent effects in a qualitative or semi-quantitative way.

Semi-empirical methods may only be used for systems where parameters have been developed for all of their component atoms. In addition to this, semi-empirical models have a number of well-known limitations. Types of problems on which they do not perform well include:

- Hydrogen bonding
- Transition structures
- Molecules containing atoms for which they are poorly parameterized

Semi-empirical methods are discussed in chapter 2.1.1.7.

#### 4.4.2.2 HF Theory

HF theory is very useful for providing initial, first-level predictions for many systems. It is also reasonably good at computing the structures and vibrational frequencies of stable molecules and some transition states. As such, it is a good base-level theory. However, its neglect of electron correlation makes it unsuitable for some purposes. For example, it is insufficient for accurate modelling of the energetic of reactions and bond dissociation. The HF method is discussed in chapter 2.1.1.4.

#### 4.4.2.3 Electron Correlation and Post-SCF Methods

As mentioned above, the HF method provides a reasonable model for a wide range of problems and molecular systems. However, HF theory also has limitations. They arise principally from the fact that HF theory does not include a full treatment of the effects of electron correlation: the energy contributions arising from electrons interacting with one another. For systems and situations where such effects are important, HF results may not be satisfactory.

A variety of theoretical methods have been developed which include some effects of electron correlation. Traditionally, such methods are referred to as post-SCF methods because they add correlation corrections to the basic HF model. To date

there are many correlation methods available. Electron correlation is discussed in chapter 2.1.1.5.

#### 4.4.2.4 DFT Methods

The best DFT methods achieve significantly greater accuracy than HF theory at only a modest increase in cost (far less than MP2 for medium-size and larger molecular systems). They do so by including some of the effects of electron correlation much less expensively than traditional correlated methods. DFT methods are discussed in chapter 2.1.1.6.

#### 4.4.2.5 QM/MM

Various computational methods have strengths and weaknesses. QM can compute many properties and model chemical reactions. MM is able to model very large compounds quickly.

If the system under consideration is too large to be treated entirely by electronic structure methods, it is possible to combine these two methods into one calculation, which can model a very large compound using MM and one crucial section of the molecule with QM. This calculation is designed to give results that have very good performance when only one region needs to be modelled quantum mechanically. It can also be used to model a molecule surrounded by solvent molecules. This type of calculation is called a QM/MM calculation, and its underlying concept has been generalized in the ONIOM method (Vreven and Morokuma, 2000) to include several layers that are treated with different computational methods.

### 4.4.3 Modelling Systems in Solution

Most of the modelling methods discussed so far are performed in the gas phase in which it is reasonable to assume that there is no interaction with other molecules. However, most laboratory chemistry is done in solution where the interaction between the species of interest and the solvent is not negligible any more: The properties of molecules can differ considerably between the gas phase and solution. For example, electrostatic effects are often much less important for species placed in a solvent with a high dielectric constant than they are in the gas phase.

The most rigorously correct way of modelling chemistry in solution would be to insert all the solvent molecules explicitly and then run MD or MC calculations to give a time-averaged, ensemble average of the property of interest. This can be done using MM-style force fields, but even that is not a trivial amount of computational work. In general, however, this is a poor way to model solvation effects.

Another common approach is the **self-consistent reaction field (SCRF) method**. There are quite a number of variations on this method that model the solvent as a continuum of uniform dielectric constant  $\epsilon$ : the reaction field. The solute is placed

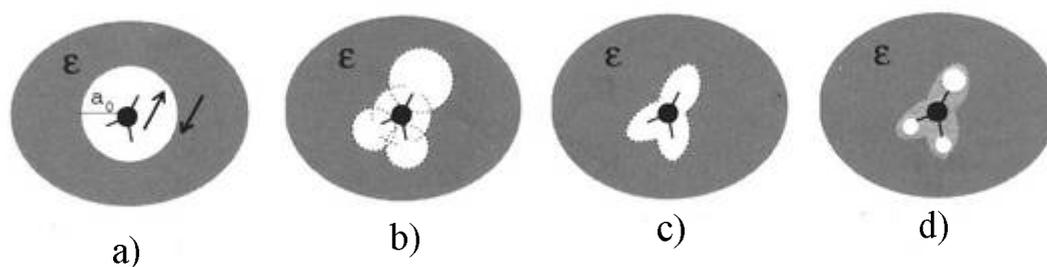
into a cavity within the solvent. SCRF approaches differ in how they define the cavity and the reaction field.

The simplest SCRF model is the **Onsager reaction field model** (Figure 18 a). In this method, the solute occupies a fixed spherical cavity of radius  $a_0$  within the solvent field. A dipole in the molecule will induce a dipole in the medium, and the electric field applied by the solvent dipole will in turn interact with the molecular dipole, leading to net stabilization. It should be noted that systems having a dipole moment of zero would not exhibit solvent effects for this method.

Tomasi's **Polarized Continuum Model (PCM)** (Figure 18 b) defines the cavity as the union of a series of interlocking atomic spheres. The effect of polarization of the solvent continuum is represented numerically: it is computed by numerical integration rather than by an approximation to the analytical form used in the Onsager model. The two isodensity surface-based SCRF models also use a numerical representation of the solvent field.

The **Isodensity PCM (IPCM) model** (Figure 18 c) defines the cavity as an isodensity surface of the molecule. This isodensity is determined by an iterative process in which a SCF cycle is performed and converged using the current isodensity cavity. The resultant wavefunction is then used to compute an updated isodensity surface, and the cycle is repeated until the cavity shape no longer changes upon completion of the SCF. An isodensity surface is a very natural, intuitive shape for the cavity since it corresponds to the reactive shape of the molecule to as great a degree as is possible (rather than being a simpler, pre-defined shape such as a sphere or a set of overlapping spheres). However, a cavity defined as an isosurface and the electron density are necessarily coupled.

The **Self-consistent Isodensity Polarized Continuum Model (SCI-PCM)** (Figure 18 d) was designed to consider this effect. It includes the effect of solvation in the solution of the SCF problem. This procedure solves for the electron density which minimizes the energy, including the solvation energy - which itself depends on the cavity which depends on the electron density. In other words, the effects of solvation are folded into the iterative SCF computation rather than comprising an extra step afterwards. SCI-PCM thus accounts for the full coupling between the cavity and the electron density and includes coupling terms that IPCM neglects.



*Fig. 18 Solvation Models (see text)*

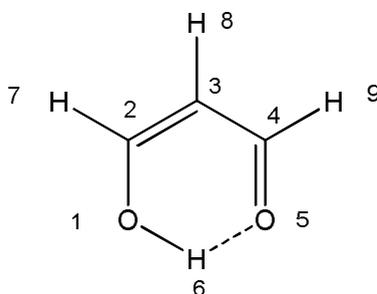
There is no one best method for describing solvent effects. The choice of method is dependent on the size of the molecule, type of solvent effects being examined and required accuracy of results. Many of the continuum solvation methods predict solvation energy more accurately for neutral molecules than for ions. The most accurate calculations are those that use a layer of explicit solvent molecules surrounded, in turn, by a continuum model. This adds the additional complexity of having to try various configurations of solvent molecules in order to obtain a statistical average.

## 5 Results and Discussion – Part 2

### Systematic Investigation of the Dependence of Geometric Parameters of Malonaldehyde on the Level of Theory Used

In this study geometry optimizations on malonaldehyde using several model chemistries in the gas phase were used to find out what level of theory is required to produce an accurate structure for this molecule. The results of this study are intended to contribute to the understanding of such events and to elect an appropriate method to obtain reliable energies and geometries for further computational modelling of larger molecular systems.

Malonaldehyde (Figure 19) provides one of the simplest examples of intramolecular hydrogen bonding. The quantitative description of small hydrogen bonded systems is a first step toward assessing the significance of such interactions in larger molecular systems. Thorough analysis of these no bonded interactions is particularly relevant in the case of CDs, where the molecular conformation depends in part on a balance between the weak forces among the hydroxyl groups within the molecule and those between the carbohydrate and the solvent, e.g., water. In other words, the calculations should prove the effectiveness of malonaldehyde as a reliable model system for the investigation of CDs and the like.



*Fig. 19 Structure of malonaldehyde with atom numbering for reference.*

All computational studies were performed with the GAUSSIAN 03 computational package (Frisch et al., 2004). Full geometry optimizations of malonaldehyde were performed at the semi-empirical PM3, HF/6-311+G(d,p), DFT B3LYP/6-31G(d), DFT B3LYP/6-31G(d,p) and MP2/6-311+G(d,p) level of theory. The relevant experimental values were taken from Foresman and Frisch (1996).

Table 7 depicts selected calculated bond lengths, and Table 8 depicts selected calculated bond angles of the optimized malonaldehyde structures.

**Tab. 7** Selected bond lengths (in Å) of malonaldehyde (for reference numbering see Figure 19).

Bond Length	PM3	HF/6-311+G(d,p)	B3LYP/6-31G(d)	B3LYP/6-31G(d,p)	MP2/6-311+G(d,p)	Experimental
O1-H6	0.97	0.95	1.01	1.00	0.9922	0.97
O1-C2	1.34	1.31	1.32	1.40	1.3241	1.32
C2-C3	1.36	1.34	1.37	1.34	1.3671	1.35
C3-C4	1.46	1.46	1.44	1.48	1.4451	1.45
C4-O5	1.23	1.20	1.24	1.24	1.2419	1.23
O5-H6	1.83	1.91	1.69	1.78	1.6864	1.68
C2-H7	1.10	1.08	1.09	1.00	1.0873	1.09
C3-H8	1.83	1.07	1.08	1.00	1.0823	1.09
C4-H9	1.10	1.09	1.05	1.00	1.1028	1.09

**Tab.8** Selected bond angles (in degrees) of malonaldehyde (for reference numbering see Figure 19).

Angle	PM3	HF/6-311+G(d,p)	B3LYP/6-31G(d)	B3LYP/6-31G(d,p)	MP2/6-311+G(d,p)	Experimental
H6-O1-C2	109.2	109.9	105.9	109.8	105.142	106.3
C2-C3-H8	120.2	119.3	120.2	118.5	119.92	128.1
C2-C3-C4	121.8	121.3	119.4	122.9	119.46	119.4
O1-C2-C3	124.8	126.3	124.4	122.6	124.183	124.5
C3-C4-O5	121.5	124.1	123.7	120.8	123.273	123.0
O1-H6-O5	140.3	138.5	146.9	139.8	148.43	-

The calculations show that the PM3 and HF geometries differ significantly from the experimental structure. This is most noticeable in the long hydrogen bonding distance (O5-H6) and the corresponding errors in the ring's internal bond angles. In contrast, the MP2 structure is in excellent agreement with experiment. The one major discrepancy comes with the C2-C3-H8 bond angle, but the reported experimental value is known to be quite uncertain. Most notably, the optimization with the DFT B3LYP/6-31G(d) model chemistry yields a resulting geometry which is in excellent agreement with the MP2 results with a larger basis set, while the B3LYP job takes less than 20% as long as the MP2 optimization.

## 6 Host-Guest-Chemistry

Modern chemistry is based on the understanding of the chemical bond. The chemical bond implies the distribution and delocalization of electrons over the entire molecule resulting in a strong, i.e., covalent interaction. Modern theoretical ab initio quantum chemistry methods have been extremely successful in describing the electronic structure of isolated molecules to a degree of precision that in some cases comes very close to high-resolution spectroscopic results.

Atoms and molecules can interact together leading to the formation of either a new molecule or a molecular cluster. The former is a **covalent interaction**; the latter one in which a covalent bond is neither formed nor broken is termed a **no covalent interaction**. The properties of the original subsystems in a molecular cluster are relatively unperturbed compared to the isolated molecules. Nevertheless, the stronger the no covalent interaction is, the larger are the changes in the properties of the subsystems. Most pronounced changes occur in H-bonded systems. No covalent interactions are considerably weaker (by one or two orders of magnitude) than covalent interactions.

The role of no covalent interactions in nature was fully recognized only in the last decades; they play an important role in chemistry and physics and are of key importance in the bio-disciplines. The structures of liquids, solvation phenomena, molecular crystals, physisorption, and the structures of bio-macromolecules such as DNA and proteins, and molecular recognition are only a few phenomena determined by no covalent interactions.

No covalent interactions play a special role in **supramolecular chemistry**, which has been defined by Lehn as “chemistry beyond the molecule” (Lehn 1988, 1990, 2002).

These no covalent interactions, as being most important for organic compounds, include:

- hydrogen bonds (H-bonds),
- stacking interactions ( $\pi$ - $\pi$  interactions),
- electrostatic interactions,
- hydrophobic interactions, and
- charge-transfer interactions.

No covalent interactions are discussed in detail in chapter 8

Supramolecular chemistry offers incredible applications in various fields such as medicinal chemistry (drug delivery systems), host-guest chemistry, catalysis, etc.

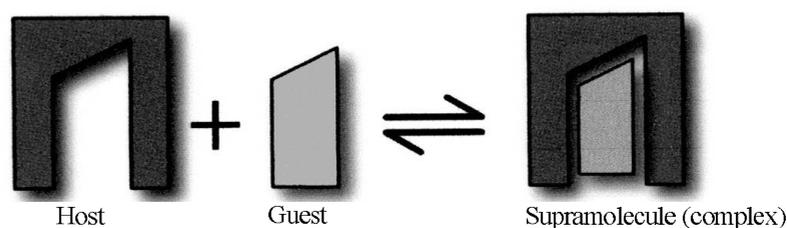
The existence of a condensed phase probably represents the most important example of no covalent interactions. The vast majority of chemical processes occur in solution, and the condensed phase affects the structure, properties, and reactivity of a system dramatically and plays an even more important role in biology.

Covalent and no covalent interactions differ considerably and have completely different origins. A covalent bond is formed when partially occupied orbitals of interacting atoms overlap and consists of a pair of electrons shared by these atoms. Covalent interactions are of short range and covalent bonds are generally shorter than two Å.

No covalent interactions are known to act at distances of several Ångstroms or even tens of Ångstroms and overlap is thus unnecessary (in fact overlap between occupied orbitals leads only to repulsion). The reason for the attraction between interacting subsystems must be sought elsewhere and it can lie only in the electrical properties of the subsystems. No covalent interactions originate from interaction between permanent multipoles, between a permanent multipole and an induced multipole, and finally, between an instantaneous time variable multipole and an induced multipole. The respective energy terms, called electrostatic, induction, and dispersion are attractive (only the electrostatic term, depending on the orientation of the subsystems, can be attractive or repulsive). The repulsive term, called exchange repulsion, connected with the above-mentioned overlap of occupied orbitals, prevents the subsystems from approaching too closely.

The total stabilization energy of a molecular cluster lies usually between 1 and 20 kcal mol<sup>-1</sup>, which is considerably smaller than the binding energy of a covalent bond of about 100 kcal mol<sup>-1</sup>. To describe and study no covalent interactions, it is essential to apply the most accurate methods of quantum chemistry, and experimental studies of these interactions belong to one of the most challenging tasks of contemporary science.

The construction of any supramolecular system involves selective molecular combination, i.e., molecules selectively catch other molecules in a process that is called “**molecular recognition**” (see below). The molecules that do the recognizing are called **host molecules**, and those that are recognized are known as **guest molecules**. Therefore, molecular recognition chemistry in its simplest form can be regarded as “host-guest-chemistry” (Cram, 1988). Figure 20 shows schematically the building of a supramolecular complex formed by a suitable host and guest molecule.

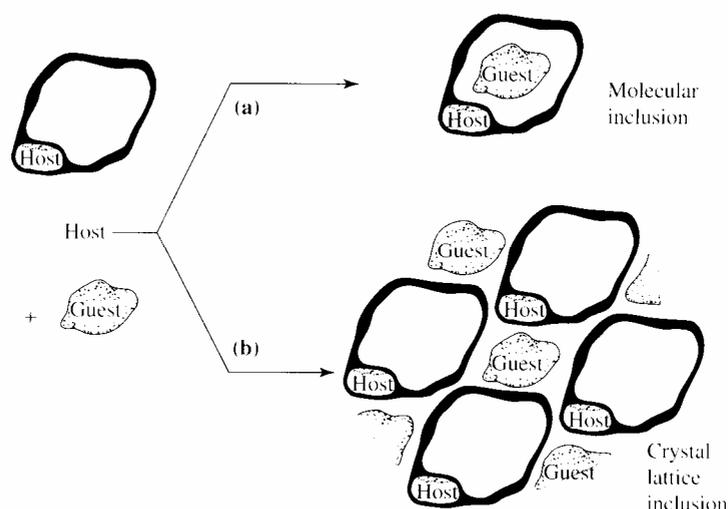


*Fig. 20 Building of a supramolecular complex by a suitable host and guest*

## 6.1 Classification of Host-Guest Compounds

In describing modern host-guest chemistry, it is useful to divide host compounds into two major classes according to the relative topological relationship between guest and host. **Cavitands** may be described as hosts possessing intramolecular cavities. This means that the cavity available for guest binding is an intrinsic molecular property of the host and exists both in solution and in the solid state.

Conversely, **clathrands** are hosts with extramolecular cavities (the cavity essentially represents a gap between two or more host molecules) and is of relevance only in the crystalline or solid state. The host-guest aggregate formed by a cavitand is termed a **cavitate**, while clathrands form **clathrates**. The distinction between the two host classes is illustrated schematically in Figure 21.



*Fig. 21 Scheme illustrating the difference between a cavitand and a clathrand (a) and the conversion of a cavitand into a cavitate by inclusion of a guest into the cavity of the host molecule (b) (taken from Steed and Atwood 2000)*

A further fundamental subdivision may be made based on the forces between host and guest. If primarily electrostatic forces weld the host-guest aggregate (including ion-dipole, dipole-dipole, hydrogen bonding etc.) the term “**complex**” is used. On the other hand, species held together by less specific (often weaker), no directional interactions, such as hydrophobic, van der Waals or crystal close-packing effects, then the terms “cavitate” and “clathrate” are more appropriate. It should be noted that there is a significant trend in the current literature to use the word “complex” to cover all of these phenomena.

## 6.2 The Importance of Preorganisation and Complementarity

In order to bind, a host must have binding sites that are of the correct electronic character (polarity, hydrogen bond donor/acceptor ability etc.) to complement those of the guest. Hydrogen bond donors must match acceptors; Lewis acids must match Lewis bases, and so on.

Furthermore, those binding sites must be spaced out on the host in such a way as to make it possible for them to interact with the guest in the binding conformation of the host molecule. If a host fulfils these criteria, it is said to be **complementary**.

If a host molecule does not undergo a significant conformational change upon guest binding, it is said to be **preorganised**. Host preorganisation is a key concept because it represents a major (in some cases decisive) enhancement to the overall free energy of guest complexation.

Neglecting the effects of solvation, the host guest binding process may be divided very loosely into two stages. First, the host undergoes conformational readjustment in order to arrange its binding sites in the fashion most complementary to the guest and at the same time minimising unfavourable interactions between one binding site and another on the host in an activation stage. This is energetically unfavourable, and because the host must remain in this binding conformation throughout the lifetime of the host-guest complex, this energy is never paid back.

Following rearrangement, binding occurs which is energetically favourable because of the enthalpically stabilising attraction between mutually complementary binding sites of host and guest.

The overall free energy of complexation represents the difference between the unfavourable reorganisation energy and the favourable binding energy. If the reorganisation energy is large, then the overall free energy is reduced, destabilising the complex. If the host is preorganised, this rearrangement energy is small.

The corollary of preorganisation is in the guest binding kinetics. Rigidly preorganised hosts may have significant difficulty in passing through a complexation transition state and so tend to exhibit slow guest binding kinetics. Conformationally

mobile hosts are able to adjust rapidly to changing conditions, and both complexation and decomplexation are rapid.

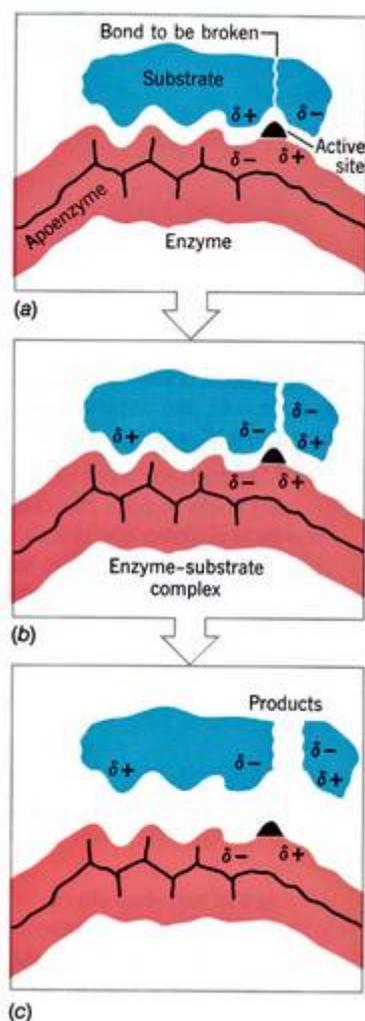
Solvation enhances the effects of preorganisation since the solvation stabilisation of the unbound host is often greater than the case when it is wrapped around the guest, effectively presenting less surface area to the surrounding medium.

### **6.3 Molecular Recognition**

Molecular recognition is the basis for supramolecular chemistry. All actions from the simple colour change in a reaction flask to highly sophisticated biological mechanisms are the result of chemical reactions and physicochemical interactions in various combinations.

However, these reactions and interactions occur seldom randomly, but most often between selected partners. This is especially true in highly organized biological systems, as the molecules recognize the best (or better) partners (Lamb and Jorgensen 2004). As mentioned above, this mechanism is called molecular recognition.

The importance of molecular recognition was realized around the middle of the nineteenth century. Pasteur noticed that there are two kinds of crystals of tartaric acid that are mirror images of each other, and these chiral isomers spontaneously self-recognize, resulting in the separate crystallization of each type. Living creatures such as mold and yeast recognize and utilize only one of these chiral isomers. Emil Fischer proposed that enzymes recognize substrates by a “lock and key” mechanism, where the structural fit between the recognizing molecule and the recognized molecule is important (Koshland 1994, Lichtenthaler 1994,). This important principle by Fischer is highlighted in Figure 22.



**Fig. 22.** Illustration of the lock-and-key model for enzyme action. (a) The enzyme and its substrate fit together to form an enzyme-substrate complex. (b) A reaction, such as the breaking of a chemical bond, occurs. (c) The product molecules separate from the enzyme.

The cyclic oligosaccharides CDs and the cyclic oligopeptide valinomycin were recognized as naturally occurring host molecules in the 1950s. Pedersen's discovery of crown ether in 1967 opened the door to research on artificial host molecules.

Cram applied the concept of artificial hosts to various kinds of molecules, and developed the research field of host-guest chemistry, referring to chemistry where a molecule (the host) accepts another particular molecule (the guest). Lehn combined the molecular assembly and host-guest chemistries into a unified concept, "supramolecular chemistry", reflecting the fact that this field deals with the complex entities - supermolecules - formed upon the association of two or more chemical species held together by intermolecular forces. The functionality of a supermolecule is expected to exceed a simple summation of its individual components. Lehn,

Pedersen, and Cram were jointly awarded the Nobel Prize in 1987 (Cram; 1988, Lehn 1988; Pedersen 1988).

## 6.4 Chiral Recognition

Especially in biochemical systems, enantiospecific receptor-substrate binding is of the utmost importance. As a result, there is a great deal of interest in the application of chiral supramolecular compounds such as CDs and their derivatives as enzyme mimics and models as well as in abiotic chiral catalysis. The latter might ultimately find application in the synthesis or separation of chiral pharmaceuticals owing to the usually different pharmacological activity of enantiomers of a chiral compound. The second enantiomer of a chiral drug, usually present as 50% impurity, can even be harmful. Therefore, the ability to predict which enantiomer forms a stronger complex with a particular CD and to estimate the energy difference between these diastereomeric complexes is of particular significance for the pharmaceutical industry as well as for separation science. However, these energy differences are usually very small. Differences of 10 cal are sufficient to achieve a satisfactory chromatographic resolution (Dodziuk 2006).

To get insight into the enantiospecific binding process, the inclusion complexes of D- and L-alanine in its zwitterionic form with  $\beta$ -CD were investigated in chapter 7.

Chiral recognition is based on the ability of a molecular receptor to form a complex preferentially with one of the enantiomers of a chiral molecule. The higher the difference of affinity of the receptor for these different enantiomers is, the better its discrimination efficiency is.

Most of receptors used for chiral recognition are based on macrocyclic structures such as CDs and its derivatives (Shahgaldian and Pieleles 2006). In a review dealing with enantiomeric recognition of chiral amines, Zhang et al. (1997) proposed general rules for effective chiral recognition using macrocyclic receptors.

CD molecules are chiral and are therefore able to form diastereoisomeric complexes, usually of different stability, with chiral guests. When this complex is of higher stability for one enantiomer of the guest molecule, it could be considered that enantioselective recognition occurs.

The difference of stability between the complexes formed with one enantiomer with regard to the other will define the efficiency of this chiral discrimination process. CDs are commonly used as selectors for chiral separation in high-performance liquid chromatography and capillary electrophoresis.

Mechanisms of chiral recognition by CDs have been elucidated and thoroughly reviewed elsewhere (Easton and Lincoln 1996; Kano 1997; Kano and Nishiyabu

2002; Dodziuk et al. 2004). These mechanisms could be mainly explained by “three point binding” (Davankov 1997) or “lock and key” rules (Behr 1994).

Chiral recognition is currently an active area of research because of its potential applications in areas as disparate as chemical synthesis, catalysis, enzyme mimetics, pharmaceuticals, geochemistry, and biotechnology.

## 7 Results and Discussion – Part 3

### $\beta$ -CD as a Chiral Selector between D-Alanine and L-Alanine

The purpose of this study was to establish whether  $\beta$ -CD could act as a chiral selector in the gas phase between the two enantiomeric forms of the amino acid alanine.

Because of the importance of chiral recognition, chiral molecules such as CDs continue to attract considerable attention (Ramirez et al. 1998). The formation of no covalently bound host-guest complexes in the gas phase have provided an useful method for discriminating enantiomers that takes advantage of the differences in the relative stability of the stereomeric complexes. The inclusion of the guest molecule inside the CD cavity is believed to provide the necessary requirements for chiral recognition and discrimination (Kano 1997; Kano and Nishiyabu 2002). Chiral discrimination requires that the host and guest form reasonably stable complexes. The cooperative interaction of several weak forces such as dipole - dipole, hydrophobic, electrostatic, van der Waals, and hydrogen bonding leads to molecular recognition and differentiation. The “three-point interaction” model describes more generally the nature of these interactions (Davankov 1997).

In  $\alpha$ -amino acids such as alanine two attractive interactions are present involving the protonated amine and the carboxylic acid group. The interaction of the side chain can be either attractive or repulsive depending on the respective functional group. The magnitude of the interaction can also be varied with the size and the type of the functional groups.

In this study, all calculations were performed with the GAUSSIAN 03 package (Frisch et al., 2004). Both the CD and the alanine structures were fully optimized with the B3LYP/6-31G(d,p) DFT method. The respective inclusion complexes were then constructed by placing the mass centre of the alanine enantiomer into the mass centre of the CD. Two orientations, with the carboxylic group of the alanine molecule near the narrow (=D-ALA-a and L-ALA-a), and the wide (= D-ALA-b and L-ALA-b) rim of the CD were considered. The complexes were then fully optimized with the B3LYP/6-31G(d,p) DFT method.

Table 9 summarizes the interaction energies and relative energies of the four alanine- $\beta$ -CD complexes.

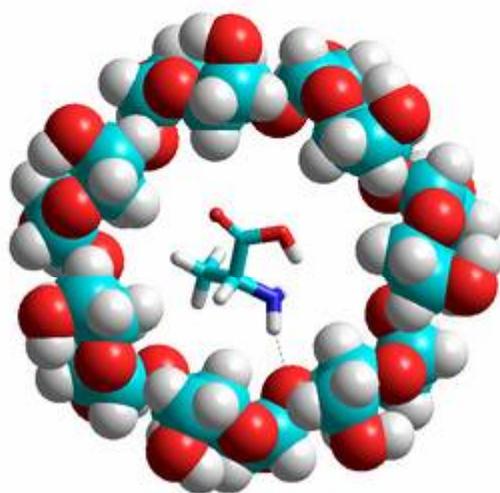
**Tab.9** Interaction energies and relative energies of several alanine- $\beta$ -CD complexes; the lowest energy is set to zero

Complex	Interaction Energy (kcal/mol)	Relative Energies in kcal/mol
D-ALA-a	-7.61	0.00
D-ALA-b	-6.37	1.24
L-ALA-a	-7.54	0.07
L-ALA-b	-6.30	1.30

The calculations revealed that in all cases the guest molecules are located inside the CD cavity, which remains almost undistorted on the complexation process.

The preferred complexation orientation is that one, in which the carboxylic group of the alanine is located near the narrow rim of the CD cavity. The D-alanine complex (Figure 23) is slightly more stable than the L-alanine complex. Moreover, in all forms an extra hydrogen bond between the ammonium group and a glycosidic oxygen atom of the CD was established.

Thus, because of the findings above it is recommended that  $\beta$ -CD can act as a chiral selector between the two enantiomeric forms of the amino acid alanine.



**Fig. 23** Lowest energy structure D-ALA-a; the extra hydrogen bond is depicted in black.

## 8 The Nature of Supramolecular Interactions

### 8.1 Basic Considerations

The motivation for the application of theoretical ab initio methods to molecular clusters comes from the need to determine the structure of the cluster, its stabilization energy, its (intermolecular) vibrational frequencies and the potential energy and free energy surfaces.

The primary property of an isolated (rigid) system is its structure, and a main goal is to determine the equilibrium structure of such a system. In the case of nonrigid (floppy) systems, the situation is different since these systems are dominated by large amplitude motions that make the concept of equilibrium structure meaningless. The majority of molecular clusters are nonrigid systems, and hence, the potential energy surface represents their primary property. Structures of global and local minima of the surface are found by optimizing the **stabilization energy** and not the total energy. Stabilization energy thus plays a central role in no covalent interactions.

The stabilization energy ( $\Delta E$ ) of a molecular cluster  $A \cdots B$  formed from subsystems  $A$  and  $B$  is evaluated following equation 8.1:

$$\Delta E = E(A \cdots B) - E(A) - E(B) \quad (8.1)$$

where  $E(A \cdots B)$ ,  $E(A)$ , and  $E(B)$  denote total energies of the molecular cluster ( $A \cdots B$ ), and subsystems  $A$  and  $B$ .

When concerning a supramolecular system it is important to consider the interplay of all of the molecular interactions and effects relating to the host and guest as well as their surroundings (e.g. solvation, crystal lattice, gas phase etc.).

In those systems, selective and effective recognition is achieved through various combinations of the molecular interactions discussed below. However, when several types of molecular interactions work together, a cooperative enhancement in molecular association is often observed. Finding an appropriate combination of molecular interactions is the key to designing efficient molecular recognition systems.

What follows is a brief overview of the various molecular interactions.

**Electrostatic interactions** occur between charged molecules. An attractive force is observed between oppositely charged molecules, and a repulsive force between molecules with the same type of charge (both either negative or positive).

The magnitude of this interaction is relatively large compared to other no covalent interactions, which means that the contributions from electrostatic interactions in molecular recognition systems cannot usually be ignored.

The strength of this interaction changes in inverse proportion to the dielectric constant of the surrounding medium. Therefore, in a more hydrophobic environment with a smaller dielectric constant, the electrostatic interaction becomes stronger.

If a functional group is in equilibrium between ionized and neutral forms, the population of the latter form decreases in a hydrophobic medium, resulting in a decreased contribution from the electrostatic interaction. **Dipole-dipole** and **dipole-ion interactions** play important roles in neutral species instead of electrostatic interactions.

**Hydrogen bonding** sometimes plays a crucial role during recognition, although a hydrogen bonding interaction is weaker than an electrostatic interaction. Hydrogen bonding only occurs when the functional groups that are interacting are properly oriented. Therefore, hydrogen bonding is the key interaction during recognition in many cases. The importance of hydrogen bonding to molecular recognition is illustrated for example by the base pairing that occurs in DNA strands, where nucleobases recognize their correct partners in a highly specific way.

Because of its great importance, hydrogen bonding is discussed in detail in chapter 8.2.

**Coordinate bonding** is another type of direction-specific interaction. This type of interaction occurs between metal ions and electron-rich atoms and is of moderate strength. Such interactions have also been utilized in the formation of supramolecular assemblies.

The **van der Waals interaction** is weaker and less specific than those described above, but it is undoubtedly important because this interaction generally applies to all kinds of molecules. It is driven by the interactions of dipoles created by instantaneous unbalanced electronic distributions in neutral substances.

Although individual interactions are negligible, the combined cooperative contributions from numerous van der Waals interactions make a significant contribution to molecular recognition. When the interacting molecules have surfaces with complementary shapes, as in the lock and key concept, the van der Waals interaction becomes more effective. This interaction is especially important when the host molecule recognizes the shape of the guest molecule.

Mainly in an aqueous medium, the **hydrophobic interaction** plays a very important role. It is the major driving force for hydrophobic molecules to aggregate as can be seen, for instance, in the formation of a cell membrane from lipid-based components.

The hydrophobic interaction is related to the hydration structure present around hydrophobic molecules. Water molecules form structured hydration layers that are not entropically advantageous. It is believed that hydrophobic substances aggregate to minimize the number of water molecules involved in hydration layers. However,

the mechanism and nature of the hydrophobic interaction is not that clear and is still under debate even today.

**$\pi$ - $\pi$  interactions** occur between aromatic rings, and these sometimes provide important contributions to molecular recognition. When the aromatic rings face each other, the overlap of  $\pi$ -electron orbitals results in an energetic gain. For example, the double-strand structure of DNA is partially stabilized through  $\pi$ - $\pi$  interactions between neighbouring base pairs.

## 8.2 The Hydrogen Bond

### 8.2.1 Basic Considerations

**The hydrogen bond** is the most important of all directional intermolecular interactions. It is operative in determining molecular conformation, molecular aggregation, and the function of a vast number of chemical systems ranging from inorganic to biological. In terms of modern concepts, the hydrogen bond is understood as a very broad phenomenon, and it is accepted that there are open borders to other effects. There are dozens of different types of  $X-H \cdots A$  hydrogen bonds that occur commonly in the condensed phases, and in addition, there are innumerable less common ones. Dissociation energies span more than two orders of magnitude (about  $0.2 - 40 \text{ kcal mol}^{-1}$ ). Within this range, the nature of the interaction is not constant, but its electrostatic, covalent, and dispersion contributions vary in their relative weights. The hydrogen bond has broad transition regions that merge continuously with the covalent bond, the van der Waals interaction, the ionic interaction, and the cation- $\pi$  interaction.

All hydrogen bonds can be considered as incipient proton transfer reactions, and for strong hydrogen bonds, this reaction can be in a very advanced state.

Most theoretical aspects of hydrogen bonding are covered by Jeffrey (1997) and Scheiner (1997).

In a hydrogen bond  $X-H \cdots A$ , the group  $X-H$  is called the **donor** and A is called the **acceptor** (short for “proton donator” and “proton acceptor”, respectively).

In a simple hydrogen bond, the donor interacts with one acceptor. Since the hydrogen bond has a long range, a donor can interact with two and three acceptors simultaneously. In such cases, the terms “bifurcated” and “trifurcated” are commonly used.

Hydrogen bonds exist with a continuum of strengths. Nevertheless, it can be useful for practical reasons to introduce a classification, such as “weak”, “strong”, and possibly also “in between”.

Following the system described by Jeffrey (1997), hydrogen bonds are called **moderate** if they resemble those between water molecules or in carbohydrates (one could also call them “**normal**”), and are associated with energies in the range 4-15  $\text{kcal mol}^{-1}$ . Hydrogen bonds with energies above and below this range are termed **strong** and **weak**, respectively. Some general properties of these categories are listed in Table 10. It must be stressed that there are no “natural” borderlines between these categories.

**Tab. 10** Strong, moderate, and weak hydrogen bonds following the classification of Jeffrey (1997)

	Strong	Moderate	Weak
Interaction type	Strongly covalent	Mostly electrostatic	Electrostat. /dispers.
Bond lengths [ $\text{\AA}$ ]	1.2-1.5	1.5-2.2	>2.2
$H \cdots A$			
lengthening of $X-H$ [ $\text{\AA}$ ]	0.08-0.25	0.02-0.08	<0.02
$X-H$ versus $H \cdots A$	$X-H \approx H \cdots A$	$X-H < H \cdots A$	$X-H \ll H \cdots A$
$X \cdots A$ [ $\text{\AA}$ ]	2.2-2.5	2.5-3.2	>3.2
directionality	strong	moderate	weak
$O-H \cdots O$ bond angles [ $^\circ$ ]	170-180	>130	>90
bond energy [ $\text{kcal mol}^{-1}$ ]	15-40	4-15	<4

However, many properties of  $n$  interconnected hydrogen bonds are not just the sum of those of  $n$  isolated bonds. Two principal mechanisms are responsible for this non-additivity, and both operate by mutual polarization of the involved groups:

- $\sigma$ -Bond Cooperativity and
- $\pi$ -Bond Cooperativity or Resonance-Assisted Hydrogen Bonding (RAHB; see below).

### 8.2.2 $\sigma$ -Bond Cooperativity

If a  $X^{\delta-} - H^{\delta+}$  group forms a hydrogen bond  $X^{\delta-} - H^{\delta+} \cdots A^{\delta-}$ , it becomes more polar. The same is true if it accepts a hydrogen bond,  $Y^{\delta-} - H^{\delta+} \cdots X^{\delta-} - H^{\delta+}$ . Thus, in a chain with two hydrogen bonds,  $Y - H \cdots X - H \cdots A$ , **both** become stronger. This effect is often called “ **$\sigma$ -bond cooperativity**”, (Jeffrey 1997, Schuster and Wolschann 1999; Karpfen 2002), since the charges flow through the  $X-H$   $\sigma$  bonds. Model calculations on moderate strength hydrogen bonds yield typical energy gains of around 20 % relative to isolated interactions (Scheiner 1997).

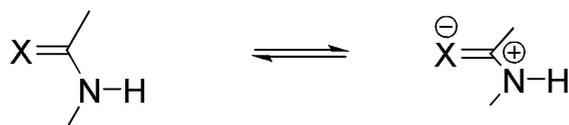
$\sigma$ -Bond cooperativity drives the clustering of polar groups. In the condensed phases, this leads to formation of  $X-H \cdots X-H \cdots X-H$  chains and rings, in

particular for  $X=O$ , but also for  $X=N$  or  $S$ . If double donors (such as  $H_2O$ ) and/or double acceptors are involved, they can interconnect chains and rings to form complex networks.

The topology of such networks has been documented in detail for the O-H-rich carbohydrates (Jeffrey and Saenger 1991, Steiner et al. 1991).

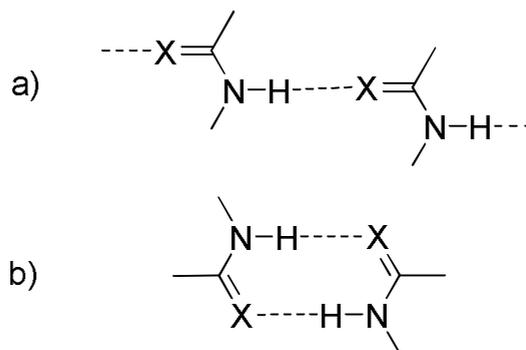
### 8.2.3 $\pi$ -Bond Cooperativity.

$X-H$  groups may also be polarized by charge flow through  $\pi$  bonds. For example, an amide  $N-H$  group becomes a stronger donor if the amide  $O$  atom accepts a hydrogen bond,  $X-H \cdots O=C-N-H$ . This results because the zwitterionic resonance form is stabilized (Figure 24).



**Fig. 24.** Resonance forms of amide and thioamide groups. The neutral form is always dominating, but the weight of the zwitterionic form is increased by accepted as well as by donated hydrogen bonds;  $X=O$  or  $S$ .

The same effect occurs in thio-amides. Amide units, because of their dual donor and acceptor capacity, often form hydrogen-bonded chains or rings (such as in protein secondary structure; Figure 25).

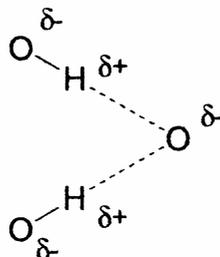


**Fig. 25.** Chains (a) and rings (b) as formed by amides and thioamides through the  $\pi$ -bond cooperativity  $X=O$  or  $S$ .

Since the polarization occurs through  $\pi$  bonds, the effect is often called  **$\pi$ -bond cooperativity**.

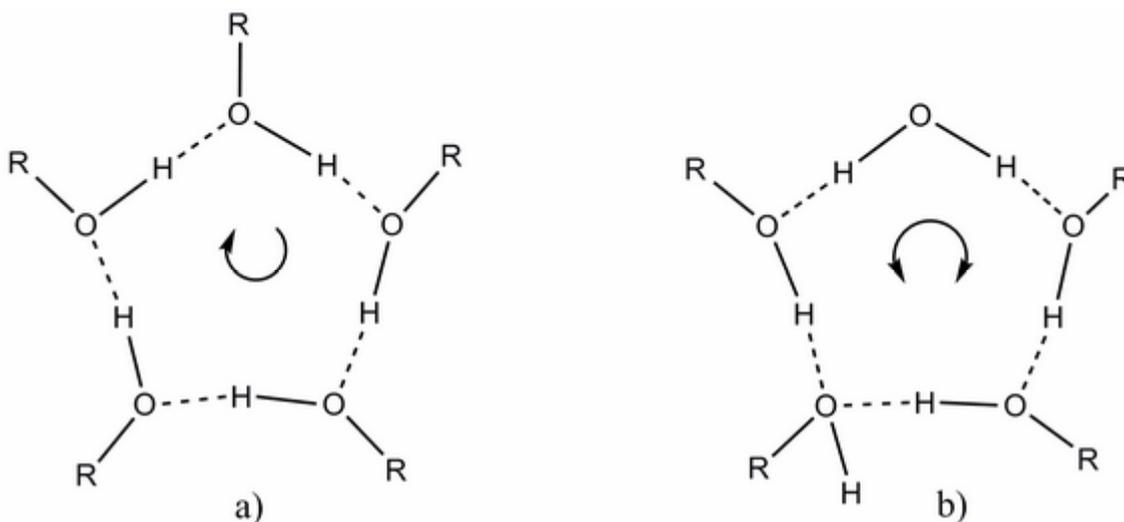
### 8.2.4 The Anticooperativity Effect; Homodromic and Antidromic Hydrogen Bonds

Hydrogen bonds may not only enhance, but also reduce the strengths of each other. This occurs, for instance, at double acceptors where two roughly parallel donor dipoles repel each other (Figure 26).



**Fig. 26.** Anticooperative hydrogen bonds. The two donors represent roughly parallel dipoles that repel each other.

This effect is called “**anticooperativity**” and is probably responsible for the preferences of “homodromic” over “antidromic” cycles of hydrogen bonds (Figure 27).



**Fig. 27.** Cycles of five hydrogen bonds (Saenger 1979). In the preferred “homodromic” arrangement (a), all hydrogen bonds run in the same direction. In the less common “antidromic” arrangement (b), a change of orientation leads to local anticooperativity.

CD-hydrate crystal structures present a valuable source for the investigation of complex  $OH \cdots O$  hydrogen bonding patterns. In  $\alpha$ -,  $\beta$ - and  $\gamma$ - CD crystal structures, it was found that “infinite” chains of  $O-H \cdots O-H \cdots OH$  hydrogen bonds are formed in which all  $OH \cdots O$  bonds are homodromic (see above). These infinite chains can also close up to produce cycles with four, five, six, or more  $O-H$  groups (Saenger 1979; Saenger and Lindner 1980).

Besides the homodromic arrangement, in some cycles an anti-dromic system was observed. Here, one water molecule donates two hydrogen bonds, which give rise to two chains running in opposite directions and colliding at one oxygen atom acting as double acceptor.

As considered above, the reason for homodromic chain-like or cyclic arrangements is found in the cooperativity, which results if extended to  $O-H \cdots O$  hydrogen bonding networks. This cooperativity can only occur if the hydrogen bond donor acts simultaneously as acceptor, as in the case of the  $O-H$  group. If an  $O-H$  is engaged in a hydrogen bond, it is polarized such that the oxygen becomes more negative and the hydrogen becomes more positive, i.e. it is a stronger donor and acceptor compared with an isolated  $O-H$  group. For this reason, strings of  $O-H \cdots O-H \cdots O-H$  hydrogen bonds are preferred over single  $O-H$  interactions.

In many cases, the cyclic structures do not occur isolated but they are fused to form more extended patterns. Shunts or connections are provided in most cases by water molecules, which engage their two donor sites in two different cycles. As an example, homo- and antidromic cycles can be fused and infinite chains can be incorporated.

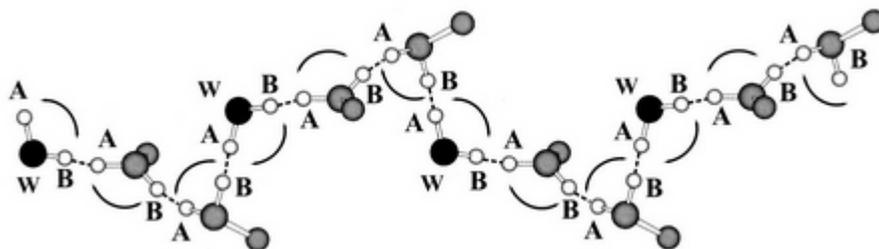
### 8.2.5 Flip-Flop Hydrogen Bonds

In the  $\beta$ -CD $\cdot$ 11H<sub>2</sub>O crystal structure, the eleven water molecules are distributed statistically over 16 sites (Betz et al. 1984). The distribution is not even, some sites being fully occupied and some others to only 20 %.

In this crystal structure, besides normal  $OH \cdots O$  hydrogen bonds, a large number of  $O-H \cdots H-O$  interactions are found.

In these, the  $O \cdots O$  distances are in the usual range, 2.7-2.9 Å, and hydrogen atoms are only  $\sim 1$  Å apart, i.e. shorter than the van der Waals  $H \cdots H$  distance of 2.4 Å. Consequently, they cannot be present simultaneously; their positions are occupied only  $\sim 0.5$  and in each  $O-H \cdots H-O$ , hydrogen atom occupancies add up to  $\sim 1.0$ .

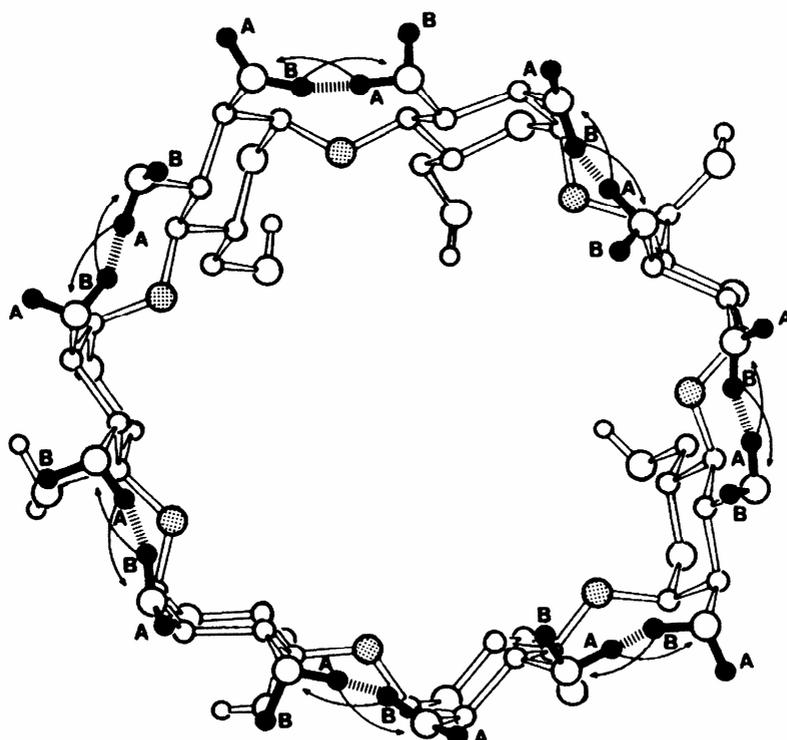
Because these hydrogen bonds represent a statistical average over two states they were termed **flip-flop hydrogen bonds** (Saenger et al. 1982; Figure 28).



**Fig. 28** A “Flip-Flop” hydrogen chain containing water molecules shown by filled circles. Hydrogen atoms cannot simultaneously occupy positions A and B (taken from Dodziuk 2006).

In the  $\beta$ -CD crystal structure, several of these flip-flop hydrogen bonds form more extended systems, all of which can occur in two states. If transition from one to the other state takes place, all hydrogen bonds in a system have to rotate from one into the other position in a concerted, cooperative motion. Zabel and co workers (1986) clearly demonstrated that most of the flip-flop hydrogen bonds represent a dynamic equilibrium in the solid state.

In the CDs, all O (2) and O (3) hydroxyls are on the same side of the macrocyclic molecule and are able to form intramolecular, interglucose  $O(2)\cdots O(3)$  hydrogen bonds. In the case of the  $\beta$ -CD $\cdot$ 11H<sub>2</sub>O crystal structure, all of the seven intramolecular hydrogen bonds are of the flip-flop type, (Figure 29).



*Fig. 29. In the  $\beta$ -CD-11H<sub>2</sub>O crystal structure, all intramolecular, interglucose hydrogen bonds are of the flip-flop type. Hydroxyl hydrogen atoms are drawn black. If positions A are filled, positions B have to be empty and vice versa (Betzel et al. 1984).*

They are probably responsible for the rigidity of the molecule in aqueous solution (Saenger, 1980).

## 9 CDs as Supramolecular Hosts

### 9.1 Basic Considerations

CDs are indispensable excipients not only in pharmacy and pharmaceutical technology, but also in many other scientific disciplines, like environmental, technical, and analytical chemistry, for stereo-specific separations of diastereomers and optical isomers, extraction of natural products, protection, and stabilization of light-, temperature-, or oxidation-sensitive compounds.

The reason for this broad field of applications of CDs is their ability to form inclusion complexes with small or even medium-sized organic or inorganic compounds. Such an inclusion influences the physico-chemical behaviour of the guest molecules, like the reactivity or the solubility significantly. Emulsification of highly apolar compounds, change of the catalytic activities, support in organic syntheses, masking of odour or taste, increase of bioavailability and subsequently higher efficiency of the active substance because of solubility enhancement, and the permission of controlled release are topics of actual CD research.

Steric as well as electronic parameters of both the CDs and the guest molecules determine the driving forces of the complexation and the geometries of the inclusion complexes (see below). In addition, the use of various CDs and CD derivatives enhances the variability of applications tremendously.

Many review articles have been published, which give excellent overviews about a large number of applications and detailed descriptions of molecular properties of CDs and CD complexes (French 1957; Bender and Momiyama 1978; Szejtli, 1982, 1988, 1996, 1998, 2004; Hedges 1998; Schneider et al. 1998; Del Valle, 2004; Dodziuk 2006). Particularly, as a consequence of the high importance of CDs in pharmaceutical applications many reviews and articles have been published (Frömming and Szejtli 1994; Loftsson and Brewster 1996; Stella and Rajewski 1997; Hirayama and Uekama 1999; Szente and Szejtli 1999; Loftsson and Masson 2001; Singh et al. 2002; Davis and Brewster 2004; Uekama 2004; Challa et al. 2005; Uekama et al. 2006; Brewster and Loftsson 2007; Vyas et al. 2008).

Finally, a review emphasizing historical development perspectives of pharmaceutical applications has been presented quite recently (Loftsson and Duchene 2007).

Native CDs are obtained by the degradation of starch [ $\alpha$  (1 $\rightarrow$ 4) linked polyglucose] by  $\alpha$ -1,4-glucan-glycosyltransferases. Depending on the respective transferase, different types of CDs result, consisting of six ( $\alpha$ -CD), seven ( $\beta$ -CD), or eight ( $\gamma$ -CD  $\alpha$  (1 $\rightarrow$ 4) linked glucose units. There also exist larger CDs, as  $\delta$ -

$\epsilon$ -, and  $\iota$ -CDs with nine, ten, and fourteen glucose residues. CDs with higher degrees of polymerization up to several hundred glycosyl units produced by a number of  $\alpha$ -4-glucotransferase have been also described briefly (Endo et al. 2002; Larsen 2002; Zheng et al. 2002; Taira et al. 2006). Because of the high ring flexibility and the consequently distorted interior, the importance of these CDs is rather low.

Modified CDs have been synthesized to change their inclusion properties and to induce biomimetic functions. Random mono- and dimethylation as well as permethylation have been applied widely, leading to an increase of the solubility and to changes of the thermodynamic parameters of host guest association (Khan et al. 1998). Other modified CD derivatives result from the hydroxypropyl substitution at the positions O6 and O2 (Pitha and Trinadha 1990).

Monosubstitution of CDs offers the entrance into a broad field of applications as supramolecular capsules. Entities grafting on one of the CD faces are so called “caps”, whereas caps attached to only one glucose unit are called “flexible caps” and in the case of “rigid caps” an organic fragment forms a bridge by linking two glucose units at one end of the CD. These caps induce significant cavity distortions and therefore have been believed to provide better complementarity between host and guests leading to higher binding constants or higher catalytic reaction rates (Engeldinger et al. 2003). The use of CDs in nanoparticles for controlled drug release is also of high interest (Duchene et al. 1999a, b; Pariot et al. 2000)

## 9.2 Structural Features of CDs

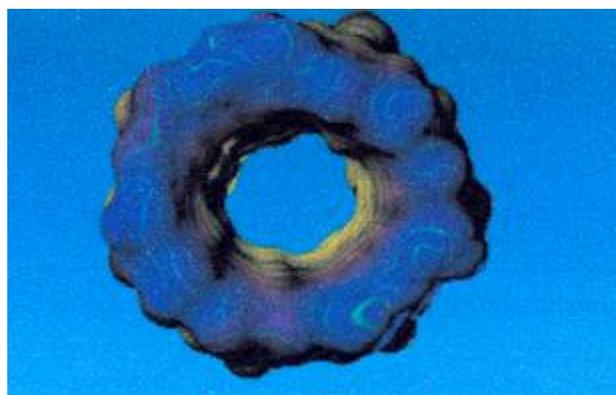
The special structures of CDs have been investigated widely, mainly by X-ray crystallography (Manor and Saenger 1974; Chacko and Saenger 1981; Lindner and Saenger 1982a, b; Hingerty et al. 1984; Harata 1987, 1998; Lipkowitz 1991; Steiner et al. 1991; Lipkowitz et al. 1992; Steiner and Koellner 1994; Steiner and Saenger 1994; Saenger et al., 1998; Saenger and Steiner 1998) and by neutron diffraction (Klar et al. 1980; Betzel et al. 1984; Zabel et al. 1986; Imamura et al 2001, Dobado et al. 2004). In most of these experimentally determined structures, water molecules are included, because of strong hydrogen bonds, which are formed between the hydroxyl groups of the CDs rims and the solvent. A few experimental data are available about the structure of CDs obtained from more or less anhydrous conditions (Diot et al. 1998; De Brauer et al. 2000; Steiner and Saenger 1996).

In 1996 Lichtenthaler and Immel conducted a statistical analysis of the solid-state structures available for the CDs and their inclusion compounds - 42 for  $\alpha$ -CD, 48 for  $\beta$ -CD, and 8 for  $\gamma$ -CD (Lichtenthaler and Immel 1996a) and found their mean molecular geometry parameters to be within normal ranges (see below).

The mean 2-O $\cdots$ O-3' distances between adjacent glucose portions decrease in the Order  $\alpha$ -CD >  $\beta$ -CD >  $\gamma$ -CD from 3.05 to 2.84 Å, allowing more intense 2-O $\cdots$ HO-3' hydrogen bonding interactions.

This reduces the overall flexibility of the macro cycles correspondingly. The intersaccharidic oxygen atoms that without exception point toward the inside of the macro cycle essentially lie in one plane, deviations from planarity being in the range of only 0.02-0.12 Å. The global molecular shape of  $\alpha$ -,  $\beta$ - and  $\gamma$ -CD in their various hydrates and inclusion complexes is thus uniformly characterized by essentially unstrained, torus-shaped cones with a nearly unpuckered mean plane. These results justify considering the solid-state structures of  $\alpha$ -,  $\beta$ - and  $\gamma$ -CD hydrates, crystallizing with 8-14 water molecules, as relevant “frozen molecular images” of their solution conformations.

Moreover, Lichtenthaler and Immel used the solid-state data to compute the contact surfaces, cavity dimensions, and molecular lipophilicity patterns (MLPs) of  $\alpha$ -,  $\beta$ - and  $\gamma$ -CD. They found that the MLPs, presented in colour-coded form, provide a lucid picture of how these CDs are balanced with respect to their hydrophilic (blue) and hydrophobic (yellow) areas (Figure 30). The larger opening of the cone-shaped macro cycles carrying the 2-OH and 3-OH groups is intensely hydrophilic; the opposite, narrower opening, ringed by the CH<sub>2</sub>OH groups, is considerably less hydrophilic, and is partially permeated by hydrophobic areas, whereas the bulk of the intensely hydrophobic regions are concentrated on the inner region of the cavities.



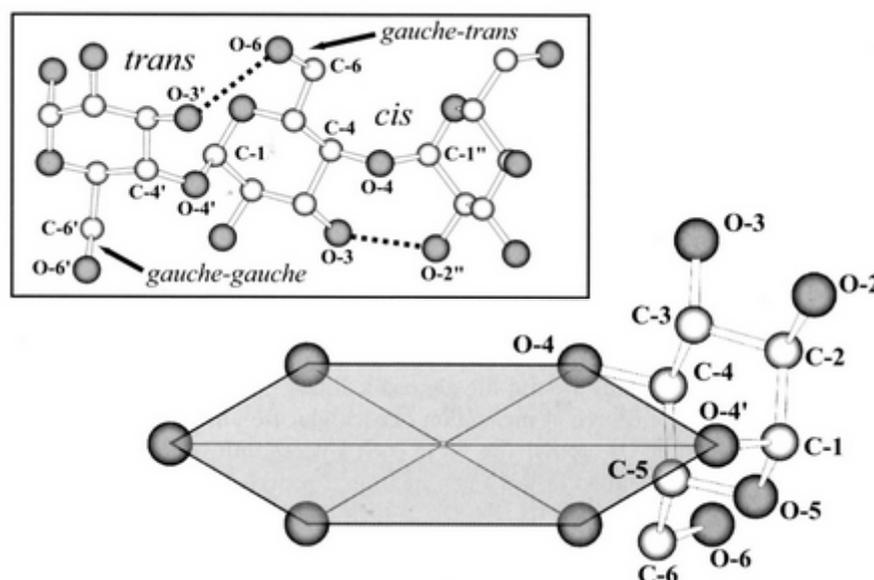
*Fig. 30. Molecular lipophilicity (MLP) projected onto the contact surface of  $\beta$ -CD. (Taken from Lichtenthaler and Immel 1996a).*

Thus, the complexation of suitable guest molecules by  $\alpha$ -,  $\beta$ - and  $\gamma$ -CD, which is governed by a variety of factors, can be rationalized with respect to the hydrophobic interactions based on their MLP profiles.



Crystallographic investigations revealed that the structures of CDs seem to be rather rigid, which can be also assumed from the fact that relatively rigid glucose units are joined in a cyclic arrangement, taking the  ${}^4C_1$  chair conformation. Nevertheless, some flexibility has been postulated for CDs as well as for their complexes (Köhler et al. 1995; Naidoo et al. 2004).

Some structural characteristics of CDs are illustrated in Figure 32



**Fig. 32** Illustrations showing characteristics of the CD structure. The plane of glycosidic O4 atoms is shaded. Intramolecular hydrogen bonds are denoted by dotted lines (taken from Dodziuk 2006)

Primary hydroxyl groups have rotational flexibility around the C5-C6 bond. Their conformation is confined to two types, *gauche-gauche* and *gauche-trans*. In the former type, the C6-O6 bond is *gauche* to both the C5-O5 and C4-C5 bonds while in the other type the C6-O6 bond is *gauche* to the C5-O5 bond and *trans* to the C4-C5 bond (Lichtenthaler and Immel 1996a).

Secondary hydroxyl groups are circularly aligned like the rim of a torus and hydrogen bonds are formed between an O2H hydroxyl group and the O3H hydroxyl group of an adjacent glucose unit.

These intramolecular hydrogen bonds make the molecules more rigid, and in solution there is a competition between inter- and intramolecular hydrogen bond networks, which are reflected in extraordinary physicochemical properties.

For example, the solubilities of the parent compounds in water are influenced so far that  $\beta$ -CD is less soluble than  $\alpha$ - and  $\gamma$ -CD (Buvári-Barcza and Barcza 2000; Plazanet et al. 2004), and surprisingly methyl substituted derivatives. Negative

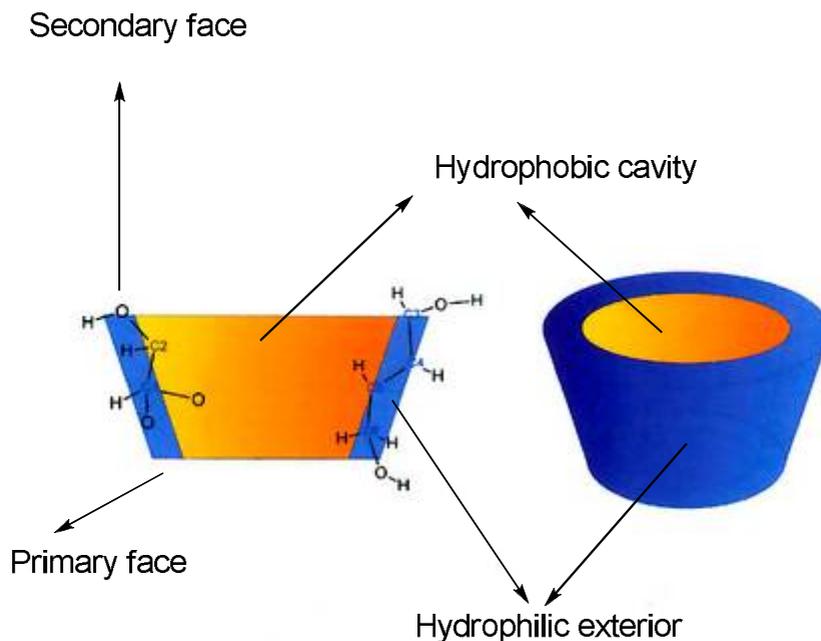
solvation enthalpies result from the competition of intramolecular hydrogen bonds with hydrogen bonds to the solvent (Kozar and Venanzi 1997; Dodziuk 2002).

In the crystalline state, CDs form both intra- and intermolecular hydrogen bonds, which stabilize not only the molecular conformation but also the crystal packing. In CDs consisting of six to nine glucose units, two adjacent glucose units are in a *cis* arrangement and secondary hydroxyl groups, either  $O2-H$  or  $O3-H$ , form hydrogen bonds. On the other hand, the *trans* - arrangement observed in larger CDs allows the formation of a hydrogen bond between an  $O3-H$  hydroxyl group and an  $O6-H$  hydroxyl group of an adjacent glucose unit.

These hydrogen bonds are incorporated in a hydrogen-bond network that stabilizes the crystal structure by a cooperative effect. In crystals of  $\alpha$ -CD hydrate, the  $O-H$  direction in the hydrogen bond network is not random but the bonds point in the same direction. On the other hand, in  $\beta$ -CD crystals, the direction of the O-H-bond can change cooperatively through the network, shown as the flip-flop hydrogen bond disorder (see chapter 8.2.5).

However, all the hydroxyl groups of the glucopyranose building blocks of a CD molecule are orientated to the exterior of the molecule, with the primary hydroxyl groups located at the narrow primary face of the torus and the secondary hydroxyl groups on the wide secondary face (Figures 32 and 33).

The CD exterior is therefore hydrophilic, whereas the central cavity, lined with skeletal carbon and ether oxygen atoms of the glucopyranose units, is postulated to be hydrophobic to some extent. For instance, the inclusion of some compounds with proton transfer equilibrium by  $\beta$ -CD shifts the equilibrium in the same direction as the addition of dioxane to 24% (v/v) (Colman et al. 1992).



*Fig. 33. Functional scheme of CD anatomy*

CDs consisting of six to nine glucose units exhibit a doughnut-shaped structure (Figure 33). These CDs are approximately in the shape of a truncated polygonal cone. The macro cycle of CDs is distorted from its ideal structure with  $n$ -fold symmetry to relieve the steric hindrance between glucose units and distortion of the glycosidic linkage.

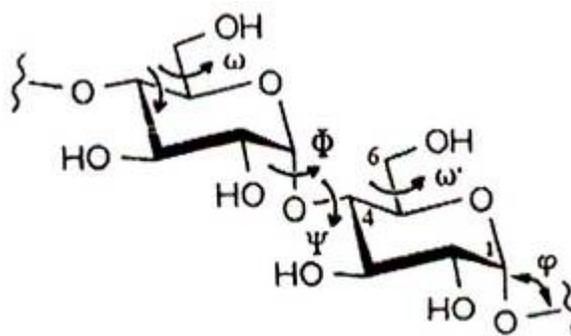
Upon complex formation, CDs change their macrocyclic structure and adjust the structure of the cavity to accommodate the guest molecule. The shape of the CD ring is well characterized by using a polygon consisting of glycosidic O4 atoms, which are nearly coplanar (Figure 32). The radius of the polygon, that is the average of the distance from the centre of the polygon to each O4 atom, is a good measure to estimate the size of the macro cycle. The radius varies in the range from 4.2 Å in  $\alpha$ -CD to 6.4 Å in  $\delta$ -CD (Dodziuk 2006).

The  $\alpha$ -1,4 linkage is responsible for the conformational flexibility of the macro cycle. The pyranose ring is almost perpendicular to the O4 plane but the average over all the glucose units indicates that the pyranose ring is generally tilted with its primary hydroxyl side towards the inside of the macrocyclic ring. The rotational movement of glucose units around the glycosidic linkage distorts the macrocyclic ring by increasing the degree of inclination of each glucose unit against the plane through the glycosidic O4 atoms. The tilt of each glucose unit is evaluated by the "tilt-angle"  $\tau$ , which is defined as the angle measured between the mean plane of the O4 polygon and the least-squares best-fit mean plane through the six pyranoid ring atoms C1-C2-C3-C4-O5 (Lichtenthaler and Immel 1996a).

Other important descriptors of the molecular shape of CDs include:

- the intersaccharidic angle  $\varphi$  (C1-O1-C4')
- the torsion angles  $\Phi$  (O5-C1-O1-C4') and  $\Psi$  (C1-O1-C4'-C3') which describe the conformations relative to the glycosidic linkages and
- the exocyclic torsion angle  $\omega$  (O5-C5-C6-O6), which describes the orientation of the primary O6H relative to the pyranoid ring.

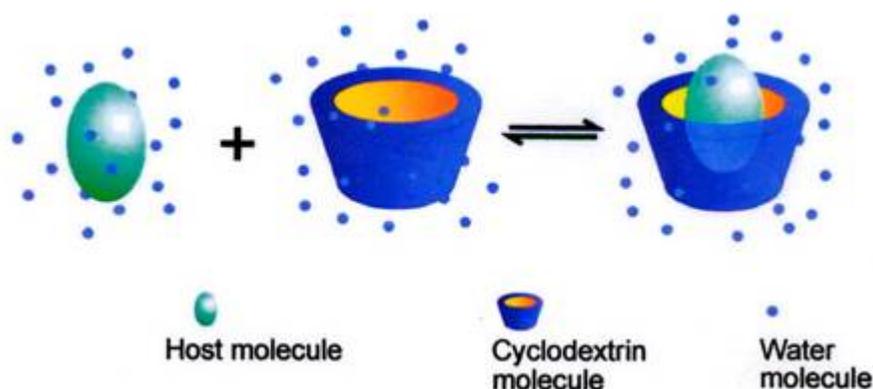
The relevant numbering for these definitions is depicted in Figures 32 and 34.



*Fig. 34. Reference numbering for the definition of geometric descriptors of CDs.*

### 9.3 Inclusion Chemistry of the CDs

As mentioned above, the most notable feature of CDs is their ability to form inclusion complexes (host guest complexes) with a very wide range of solid, liquid, and gaseous compounds. (Cramer and Hettler 1967; Saenger 1980, 1984; Atwood et al. 1984; Jursic et al. 1996; Ross and Rekharsky 1996; Mayer et al. 1997; Liu et al. 1999; Lebrilla 2001; Bodor and Buchwald 2002; Lawtrakul et al. 2003; Viernstein et al. 2002, 2003; Charumanee et al. 2006). In these complexes (Figure 35), a guest molecule is held within the cavity of the CD host molecule.



*Fig. 35 CD inclusion complex formation in an aqueous solution*

Complex formation is a dimensional fit between host cavity and guest molecule. The lipophilic cavity of CD molecules provides a microenvironment into which appropriately sized non-polar moieties can enter to form inclusion complexes. No covalent bonds are broken or formed during formation of the inclusion complex. In aqueous solutions, the main driving force of complex formation is the release of enthalpy-rich water molecules from the cavity. Water molecules are displaced by more hydrophobic guest molecules present in the solution to attain an apolar-apolar association and decrease of CD ring strain resulting in a more stable lower energy state.

The binding of guest molecules within the host CD is not fixed or permanent but rather constitutes a dynamic equilibrium. Binding strength depends on how well the host-guest complex fits together and on specific local interactions between surface atoms. Complexes can be formed either in solution or in the crystalline state and water is typically the solvent of choice. Inclusion complexation can be accomplished in a co-solvent system and in the presence of any non-aqueous solvent.

Inclusion in CDs exerts a profound effect on the physicochemical properties of guest molecules as they are temporarily locked or caged within the host cavity giving rise to beneficial modifications of guest molecules, which are not achievable otherwise. These properties include:

- solubility enhancement of highly insoluble guests,
- stabilisation of labile guests against the degradative effects of oxidation, visible or UV light and heat,
- control of volatility and sublimation,
- physical isolation of incompatible compounds,

- chromatographic separations,
- taste modification by masking off flavours, unpleasant odours and controlled release of drugs and flavours.

Therefore, CDs are used in food, pharmaceuticals, cosmetics, environment protection, bioconversion, packing, and the textile industry.

The potential guests for molecular encapsulation in CDs include straight or branched chain aliphatics, aldehydes, ketones, alcohols, organic acids, aromatics, gases, and polar compounds such as halogens, oxyacids, etc.

Due to the availability of multiple reactive hydroxyl groups, the functionality of CDs is greatly increased by chemical modification, and their applications are expanded largely. CDs are modified through substituting various functional compounds on the primary and/or secondary face of the molecule. Modified CDs are useful as enzyme mimics because the substituted functional groups act in molecular recognition. The same property is used for targeted drug delivery and analytical chemistry as modified CDs show increased enantioselectivity over native CDs.

The ability of a CD to form an inclusion complex with a guest molecule is a function of two key factors. The first is steric and depends on the relative size of the CD to the size of the guest molecule or certain key functional groups within the guest. If the guest has wrong size, it will not fit properly into the CD cavity. The second critical factor is the thermodynamic interactions between the different components of the system (CD, guest, solvent). To form the complex, a favourable net energetic driving force that pulls the guest into the CD is necessary.

While the height of the CD cavity is the same for all three types, the number of glucose units determines the internal diameter of the cavity and its volume. Based on these dimensions,  $\alpha$ -CD can typically complex low molecular weight molecules or compounds with aliphatic side chains,  $\beta$ -CD will complex aromatics and heterocycles and  $\gamma$ -CD can accommodate larger molecules such as macro cycles and steroids.

The driving force for guest inclusion involves a number of contributions, the importance of which is still a matter of some debate (Connors 1997, Liu and Guo 2002).

Chiefly the factors of importance are:

- steric fit;
- release of high-energy water;
- hydrophobic effects;
- van der Waals interactions;
- dispersive forces;
- dipole-dipole interactions;
- charge-transfer interactions;
- electrostatic interactions; and
- hydrogen bonding.

While the initial equilibrium to form the complex is very rapid (often within  $\mu\text{sec}$ ), the final equilibrium can take much longer to reach. Once inside the CD cavity, the guest molecule makes conformational adjustments to take maximum advantage of the weak van der Waals forces that exist.

Complexes can be formed by a variety of techniques that depend mainly on the properties of the active material and the equilibrium kinetics. (Del Valle 2004). However, each of these processes depends on a small amount of water to aid the different thermodynamic processes.

Water is by far the most commonly used solvent in which complexation reactions are performed. The more soluble the CD in the solvent, the more molecules become available for complexation. The guest must be able to displace the solvent from the CD cavity if the solvent forms a complex with the CD. Water, for example is very easily displaced. The solvent must be easily removed if solvent-free complexes are desired. In the case of multi-component guests, one of the components may act as a solvent and be included as a guest.

Not all guests are readily solubilized in water, making complexation either very slow or impossible. In such cases, the use of an organic solvent to dissolve the guest is desirable. The solvent should not complex well with CD and be easily removed by evaporation. Ethanol and diethyl ether are good examples of such solvents.

Dissociation of the inclusion complex is a relatively rapid process usually driven by a large increase in the number of water molecules in the surrounding environment. The resulting concentration gradient shifts the equilibrium in Figure 35 to the left. In highly dilute and dynamic systems like the body, the guest has difficulty finding another CD to reform the complex and is left free in solution.

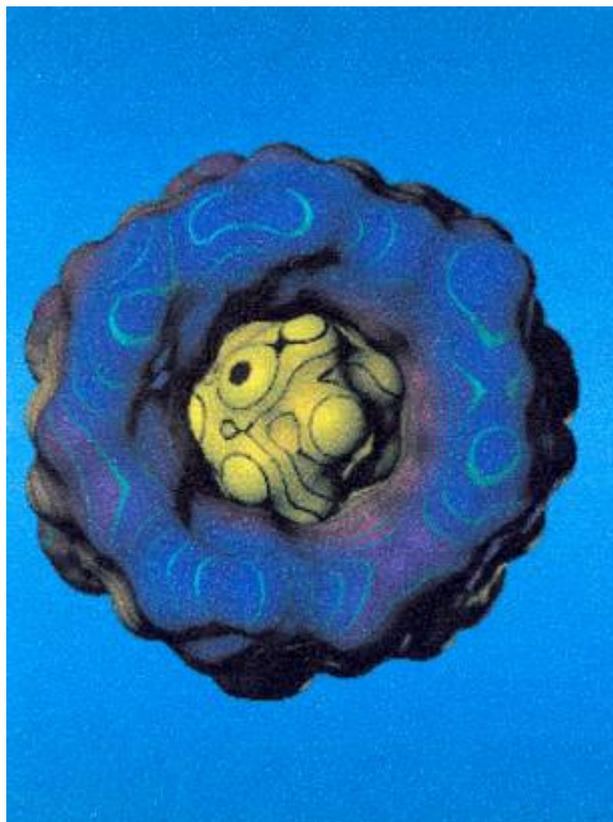
As mentioned above, the central cavity of the CD molecule is lined with skeletal carbon atoms and ethereal oxygen atoms of the glucose residues and thus provides a lipophilic microenvironment into which suitably sized guest molecules may enter and include.

In 1996, Lichtenthaler and Immel subjected five CD inclusion complexes - typical examples for their solid state as well as exemplary “frozen molecular images” of their status in solution - to a molecular modelling study comprising the computer-assisted generation of the molecular lipophilicity patterns (MLP's). (Lichtenthaler and Immel 1996b).

Based on their solid state structures, the “solvent-accessible” contact surfaces were generated by the MOLCAD program (Brickmann et al. 2000). Calculation and colour-coded projection of the MLP's onto these surfaces easily allows locating the hydrophobic and hydrophilic surface regions of CD-host and guest alike.

Their detailed analysis revealed a far-reaching, mostly complete conformity between the hydrophobic surface areas of the guest and the hydrophobic domains in the CD cavity.

This striking tendency to optimize the reciprocal concurrence of lipophilic as well as hydrophilic domains at the guest-host-interface may accordingly be concluded to be an important, if not the decisive element in orienting the guest into the cavity and in determining the stability of the complex. This is particularly true in cases where the guest is devoid of polar groups. If these are present, dipole-dipole alignments and the need for their solvation may diminish the importance of hydrophobic attractions for orienting and stabilizing the guest in the CD cavity. Figure 36 shows the molecular lipophilicity pattern (MLP) of the inclusion complex  $\beta$ -CD adamantanecarboxylic acid.



**Fig. 36.** Molecular lipophilicity pattern (MLP) of the inclusion complex  $\beta$ -CD-adamantanecarboxylic acid. View from the secondary face of the CD. Blue=hydrophilic, yellow=lipophilic (taken from Lichtenthaler and Immel 1996b).

On complexation, generally, one guest molecule is included in one CD molecule, although in the case of some low molecular weight molecules, more than one guest molecule may fit into the cavity, and in the case of some high molecular weight molecules, more than one CD molecule may bind to the guest. In principle, only a portion of the molecule must fit into the cavity to form a complex. As a result, one-to-one molar ratios are not always achieved, especially with high or low molecular weight guests.

The evaluation of association free energies and the respective enthalpic and entropic contributions is of central significance for the functional description of host-guest systems. A concise introduction to experimental methods that cover the energetic of supramolecular complexes is given, for example, by Schneider and Yatsimirsky (2000).

Measurements of stability or equilibrium constants ( $K_c$ ) or the dissociation constants ( $K_d$ ) of the guest-CD complexes are important since this is an index of changes in physicochemical properties of a compound upon inclusion.

The equilibrium constants are the stoichiometric constants defined by application of the mass action law to the concentrations of all components in equilibrium under given conditions. This can be done rigorously only if the activity coefficients of all components remain constant.

When an experimental method exists that allows one to determine the concentrations of all components under equilibrium, the calculation of the equilibrium constant is straightforward. As a rule, this is impossible, however. Fortunately, if the concentration of at least one component can be determined, or if merely a solution property proportional to the concentration of at least one component can be measured, the equilibrium constant can be calculated from the changes of this property upon variation of component concentrations. One of the most useful and widely applied analytical approaches in this context is the Phase-solubility method described by Higuchi and Connors (1965).

Hence, ignoring activity effects, the equilibrium constant  $K_c$  for the reaction



is given by

$$K_c = \frac{[CD \cdot \text{Guest}]}{[CD] \times [\text{Guest}]} \quad (9.2)$$

(units:  $dm^3 mol^{-1}$ , or  $M^{-1}$ )

Thus, a large binding constant corresponds to a high equilibrium concentration of bound guest molecules, and hence a more stable CD-guest complex.

Higher equilibria involving the formation of 1:2 complexes, or multiple aggregates involving more than one CD molecule, are common, and often exist simultaneously.

In thermodynamic terms, **selectivity** is simply the ratio of the binding constant for one guest over another:

$$\text{Selectivity} = \frac{K_{\text{Guest1}}}{K_{\text{Guest2}}} \quad (9.3)$$

This kind of selectivity tends to be the most easy to achieve because it is highly susceptible to manipulation by intelligent application of concepts such as the lock and key analogy, preorganisation and complementarity, coupled with a detailed knowledge of the CD guest interactions.

Because binding constants are thermodynamic parameters, they are related to the free energy of the complexation process according to the Gibbs equation:

$$\Delta G = -RT \ln K_c \quad (9.4)$$

The enthalpies of reactions can likewise be determined from  $K$  obtained at various temperatures using the van't Hoff equation. If two sets of data are available (i.e. two  $K$  values determined at two different temperatures in  $K$ ) then

$$\log \left( \frac{K_2}{K_1} \right) = \frac{\Delta H}{2.303R} \left( \frac{T_2 - T_1}{T_1 T_2} \right) \quad (9.5)$$

On the other hand, if a range of values is available, the  $\Delta H$  values can be obtained from a plot of  $\log K$  versus  $1/T$  using the following relationship:

$$\log K = -\frac{\Delta H}{2.303R} \frac{1}{T} + const. \quad (9.6)$$

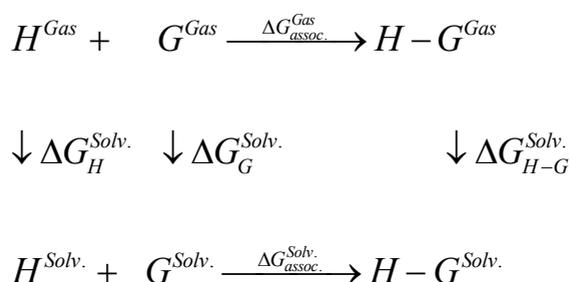
where the slope will provide the enthalpy data.

The entropy for the complexation can then be calculated using the expression:

$$\Delta S = \frac{\Delta H - \Delta G}{T} \quad (9.7)$$

Thus, the general affinity of CD for a guest under specific conditions (temperature, solvent etc.) either may be given in terms of  $K$  or  $-\Delta G$  values.

If the host-guest interaction is done in solution, the free energy of host-guest binding in solution,  $\Delta G_{H-G}^{Solv}$  can be analyzed according to the following thermodynamic cycle involving binding energies in the gas phase  $\Delta G^{Gas}$  and the solvation energies  $\Delta G^{Solv}$  of host (H), guest (G), and the host-guest complex (H-G):



This cycle allows analysis of direct host-guest interactions and of the role of solvation on binding. Quantum mechanical methods can be used for the quantitative evaluation of the binding of relatively small molecules in the gas phase. Modern quantum mechanical methods are able to compute interaction energies very accurately in the gas phase to precisions of a few tenths of a kcal mol<sup>-1</sup>, which is well within common experimental errors. However, the sizes of the systems discussed in this thesis preclude the use of such high accuracy methods; rarely have systems of more than 10 heavy atoms been treated by such methods.

The calculation of the binding free energies of larger flexible hosts and guests requires not just computation of the minimum energy of a complex but averaging over all accessible conformations of the host and of the host-guest complex. This statistical problem becomes even more severe for the proper computation of solvation; accurate results require averaging over a huge number of possible solvent-solute configurations.

Although direct quantum dynamics techniques, such as the Car-Parrinello method are applied more and more frequently to phenomena occurring during solvation (Payne et al. 1990), highly accurate quantum mechanical treatments of host-guest complexes in solution are currently impossible.

For this reason a variety of approximate methods have been developed, such as the Polarizable continuum models (PCM) for solvation (see chapter 4.4.3).

The affinities of hosts - ranging from small synthetic cavitands to large proteins – for organic molecules are well documented. In a concise review Houk and co workers (Houk et al. 2003), for example has surveyed the average binding energies and range of complex stabilities for many types of hosts, such as CDs and synthetic non-CD hosts.

CD complexes are by far the most widely used supramolecular systems; at the same time their binding mechanisms are the most difficult to rationalize and to predict.

For example, prediction models for the free energy of complexation of CDs reveal that for  $\beta$ -CD mainly hydrophobic interactions contribute to the driving forces for the complexation; hydrogen bond dependent descriptors are suggested to play a minor role. For the other CDs ( $\alpha$ - and  $\gamma$ -CD) hydrogen bond donor as well as acceptor properties determine significantly the prediction correlation (Klein et al. 2000).

As already mentioned, hydrogen bonding in CDs has been investigated intensively by various experimental and theoretical methods. Most of the experimentally available structures of CDs contain solvent molecules, mainly water; they are therefore unsymmetrical and show intermolecular hydrogen bonds. Not many examples are known for structures of CDs under anhydrous conditions.

Comparatively few molecular calculations have been performed based on semi-empirical (Lipkowitz 1998; Li et al. 2000; Liu and Guo 2004) and more accurate ab initio and DFT methods (Avakyan et al. 1999, 2001, 2005; Pinjari et al., 2006, 2007; Rudyak et al. 2006; Anconi et al. 2007; Karpfen et al. 2007, 2008; Weinzinger et al. 2007; Snor et al. 2007; Jimenez and Alderete 2008).

Complex formation is usually associated with a relative large negative  $\Delta H$  and  $\Delta S$ , which can either be negative, but also depends on the properties of the guest molecules.

Hydrophobic interactions are associated with a slightly positive  $\Delta H$  and a large positive  $\Delta S$  and, therefore, classical hydrophobic interactions are entropy driven suggesting that they are not involved with CD complexation since, as indicated, these are enthalpically driven processes.

Furthermore, for a series of guests there tends to be a linear relationship between enthalpy and entropy, with increasing enthalpy related with less negative entropy values. This effect, termed **enthalpy-entropy-compensation**, is often correlated with water acting as a driving force in complex formation.

However, Connors has pointed out that, in general, the most non-polar portions of guest molecules are enclosed in the CD cavity and, thus, hydrophobic interactions must be important in many CD complexes (Connors 1997). The main driving force for complex formation is considered by many investigators to be the release of enthalpy-rich water from the CD cavity. The water molecules located inside the cavity cannot satisfy their hydrogen bonding potentials and therefore are of higher enthalpy. The energy of the system is lowered when these enthalpy-rich water molecules are replaced by suitable guest molecules, which are less polar than water.

Other mechanisms that are thought to be involved with complex formation have been identified in the case of  $\alpha$ -CD where release of ring strain is thought to be involved with the driving force for compound-CD interaction. Hydrated  $\alpha$ -CD is associated with an internal hydrogen bond to an included water molecule, which perturbs the cyclic structure of the macro cycle. Elimination of the included water and the associated hydrogen bond is related with a significant release of steric strain decreasing the system enthalpy.

In addition, “non-classical hydrophobic effects” have been invoked to explain complexation (Connors 1997). These non-classical hydrophobic effects are a composite force in which the classic hydrophobic effects (characterised by large positive  $\Delta S$ ) and van der Waal’s effects (characterised by negative  $\Delta H$  and negative  $\Delta S$ ) are operating in the same system.

In 1998, Rekharsky and Inoue examined some thermodynamic data of  $\alpha$ - and  $\beta$ -CDs upon complexation with adamantanecarboxylates and other guest molecules as probes (Rekharsky and Inoue 1998):

In the case of  $\alpha$ -CD, the experimental data indicated small changes in  $\Delta H$  and  $\Delta S$  consistent with little interaction between the bulky probe and the small cavity. In the case of  $\beta$ -CD, a deep and snug-fitting complex was formed leading to a large negative  $\Delta H$  and a small  $\Delta S$ . Analogous results were found in the case of p-nitro phenolate and p-nitro phenol (see Table 12).

In this context, another example of significant differences is the inclusion reaction of triflumizole to  $\beta$ -CD and dimethyl- $\beta$ -CD. Although the complexation constants are similar for both cases, completely different reaction enthalpies are observed (Viernstein et al. 2002). Enthalpy-entropy compensation occurs for  $\beta$ -CD, whereas the reaction is mainly entropy-controlled for dimethyl- $\beta$ -CD.

These examples clearly demonstrate that the thermodynamical parameters of the complexation reaction depend, among other parameters, to some extent on the type of the CD involved in the complexation process.

**Tab. 12.** Thermodynamic data [ $\text{kJ mol}^{-1}$ ] for selected CD complexes.

Host		Guest		
		p-nitro phenolate	p-nitro phenol	Adamantane carboxylate
$\alpha$ -CD	$-\Delta G$	18.7	11.5	11.6
	$-\Delta H$	42.8	23.0	14.3
	T $\Delta S$	-24.1	-11.5	-2.7
$\beta$ -CD	$-\Delta G$	15.0	14.2	24.5
	$-\Delta H$	16.1	10.2	21.6
	T $\Delta S$	-1.1	3.9	2.9

To sum up: as can be seen from these considerations, there is no simple construct to describe the driving force for the complexation process between CD hosts and several guest molecules. Although release of enthalpy-rich water molecules from the CD cavity is probably an important driving force for the guest-CD complex formation several other forces may also be important (see above).

## 10 Results and Discussion – Part 4

### Quantum Mechanical Calculations on $\beta$ -CD Inclusion Complexes with Neutral and Anionic Conformations of Meloxicam

In this case study quantum mechanical DFT B3LYP/6-31G(d,p) calculations were performed in the gas phase to investigate the complexation process between  $\beta$ -CD and meloxicam considering the most stable neutral and anionic tautomers of this compound.

The motivation for this study was because so far no comparable DFT investigations on anionic and neutral meloxicam- $\beta$ -CD complexes at this level of theory have been conducted.

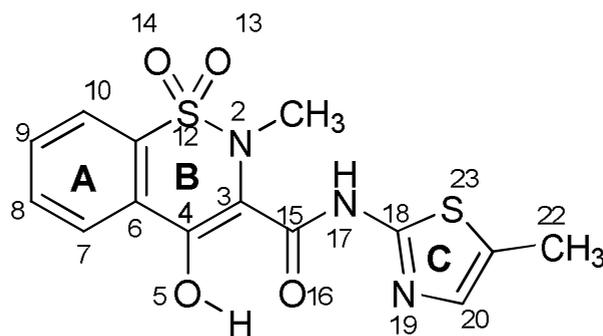
Meloxicam was chosen as guest because of its extraordinary importance as an effective COX-2 inhibitor (see below) in medicinal and pharmaceutical applications.

$\beta$ -CD is the most known and widely employed CD and can form host-guest complexes with a variety of organic and inorganic compounds (Alcaro et al. 2004). The study of these inclusion complexes has become a major goal in both basic and applied research due to a wide range of applications, such as the study of mimetic systems for enzyme reactivity, the enhancement of solubility and stability of drugs, and their use as novel carriers for drugs. Both the energetic and structural characterization of inclusion complexes in solution are central issues in CD research.

However, because the applicability of some experimental techniques such as NMR in certain molecular systems is limited, computational calculations have become very powerful tools for the structural analysis of inclusion complexes. Due to their relatively large size, most theoretical studies on CDs have been performed by MM or molecular dynamic methods and only few reports have been focused on the use of quantum chemical calculations at higher levels on the study of CD complexes.

**Meloxicam** (4-hydroxy-2-methyl-N-(5-methyl-2-thiazolyl)-2H-1,2-benzothiazine-3-carboxamide-1,1-dioxide), is a NSAID (see below) of the enolic acid class of compounds that was introduced in the 1980s and has been demonstrated to be - in comparison to all NSAIDs - the first one to inhibit COX-2 more effectively than COX-1. Its acidity is comparable to various carboxylic groups, and the presence of two aryl-moieties provides the lipophilic characteristics for protein binding and tissue distribution. It is currently marketed in more than 30 countries worldwide and has a wider spectrum of anti-inflammatory activity, combined with less gastric and local tissue irritation than NSAIDs available prior to its discovery (Luger et al., 1996; Davies and Skjodt, 1999; Banerjee et al., 2004; Naidu et al, 2004; Baboota and Agarwal, 2005; Abdoh et al., 2007).

Figure 37 depicts the enolic conformation of meloxicam with numbering of the atoms (not systematic, for reference only).



*Fig. 37. Enolic conformation of meloxicam (atom numbering for reference)*

Previous studies reported the characterization of the stoichiometry and some insights on the 3D geometry of these inclusion complexes (Tsai et al. 1993; Luger et al. 1996; Banerjee and Sarkar 2002; Banerjee et al. 2003, 2004; Naidu et al. 2004; Abdoh 2007). However, there are no clear conclusions in the literature on the mode of inclusion and the conformation of the drug inside the CD cavity. Luger et al. (1996) estimated two  $pK_a$  values for meloxicam in aqueous solution, at 1.09 and 4.18, corresponding to ionization of the thiazole ring nitrogen and the enolic OH group, respectively. In contrast to other oxicams, meloxicam exists as an anion at neutral pH and in basic solutions. In slightly acidic aqueous solution, neutral meloxicam exhibits two tautomeric conformations: zwitterion and enol. Banerjee et al. (2004) investigated the meloxicam-  $\beta$ -CD inclusion complexation in aqueous solution at pH=5.5 and they found that meloxicam was incorporated into the  $\beta$ -CD cavity with a 1:1 binding stoichiometry and a complex formation constant of  $114M^{-1}$ .

**Nonsteroidal anti-inflammatory drugs (NSAIDs)** are the most widely used drugs for the treatment of inflammatory diseases. They also exhibit various other therapeutic properties (Wenz 2000; Celotti and Laufer 2001; Banerjee et al. 2004). These drugs are not only important for their great therapeutic potential, but also for their interesting chemical properties by virtue of their dynamic structural features, which make them extremely sensitive to their microenvironment (Banerjee and Sarkar 2002).

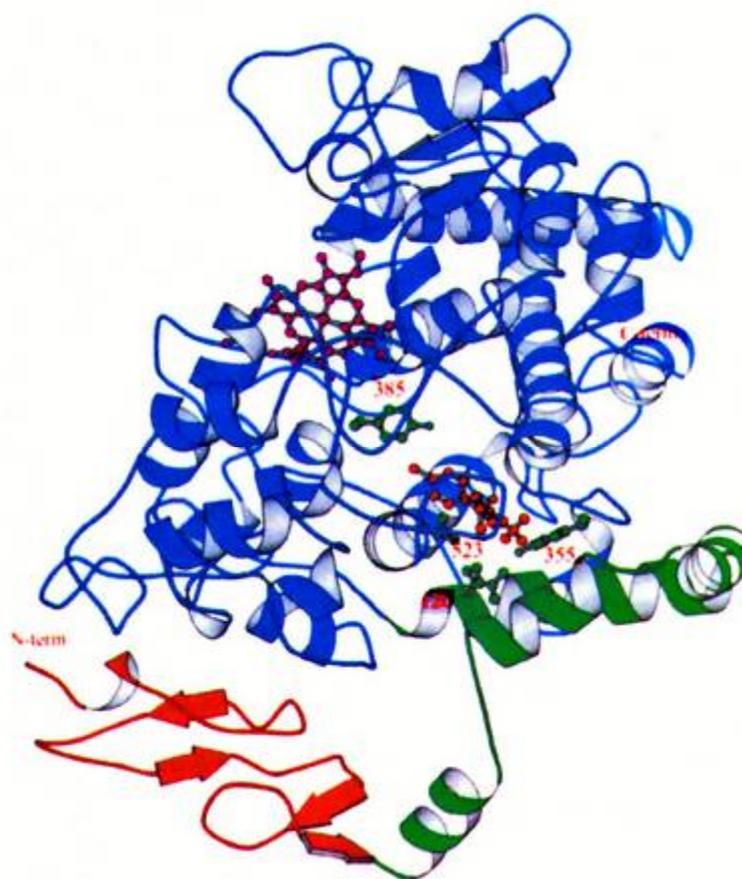
NSAIDs act as anti-inflammatory and analgesic agents by inhibiting the two isoforms of cyclooxygenase, viz., COX-1 and COX-2, the enzymes responsible for the formation of prostaglandins from dietary arachidonic acid (Vane 1971; Griswold and Adams 1996; Vane and Botting, 1998).

Prostaglandins are lipid mediators made by most cells in the body except for red blood cells and released by almost any type of chemical or mechanical stimulation; they produce a broad spectrum of effects, comprising practically every biological function as well as being of great importance as mediators of negative effects in

inflammation. Inhibition of their biosynthesis by the NSAIDs causes major changes of the pathophysiological functions of the organism.

**COX-1 and COX-2** are integral membrane proteins, which interact via three short  $\alpha$ -helices with the hydrocarbon chains of the lipid bilayer of the membrane. The active site of the enzymes is localized in a 25 Å long narrow channel that runs from the membrane-binding site up through the centre of the protein and ends at the cytoplasmic heme-binding site adjacent to the peroxidase active site. At the upper end of the mainly hydrophobic channel, the amino acids Tyr 385 and Ser 530 can be found. Arg 120 is located in the middle of the channel, offering the possibility to interact either with the carboxylate of the arachidonate or with the carboxylate of an NSAID.

The three-dimensional structure of COX-2 (Figure 38) closely resembles the structure of COX-1, except that the COX-2 active site is slightly larger and can accommodate bigger structures than those fitting into the active site of COX-1. (Kurumbail et al. 1996). Additionally, a second internal side pocket of COX-2 together with a wider central channel contributes to a larger volume of this enzyme.



**Fig. 38** Ribbon drawing of the structure of COX-2 complexed with the selective inhibitor SC-558. The structure consists of three distinct domains: an N-terminal epidermal growth factor (EGF) domain (red), followed by a membrane-binding motif (green), and the C-terminal catalytic domain (blue) which contains the cyclooxygenase and peroxidase active sites. The inhibitor is bound in the cyclooxygenase active site. Arg 120, Tyr 355, Tyr 385, and Val 523 are shown, as is the heme cofactor. (Taken from Kurumbail et al. 1996)

COX-1 can be found in most tissues and is responsible for the synthesis of prostaglandins, which regulate normal cell activities (e.g. the production of prostacyclin, which prevents gastric erosions and ulceration.). The concentration of COX-1 remains constant largely during inflammatory processes. COX-2, which is induced by cytokines, mitogens and endotoxins in inflammatory cells, is responsible for the enhanced production of prostaglandins during inflammation. Accordingly, many undesirable properties of the NSAIDs can be attributed to their insufficient COX-2 selectivity.

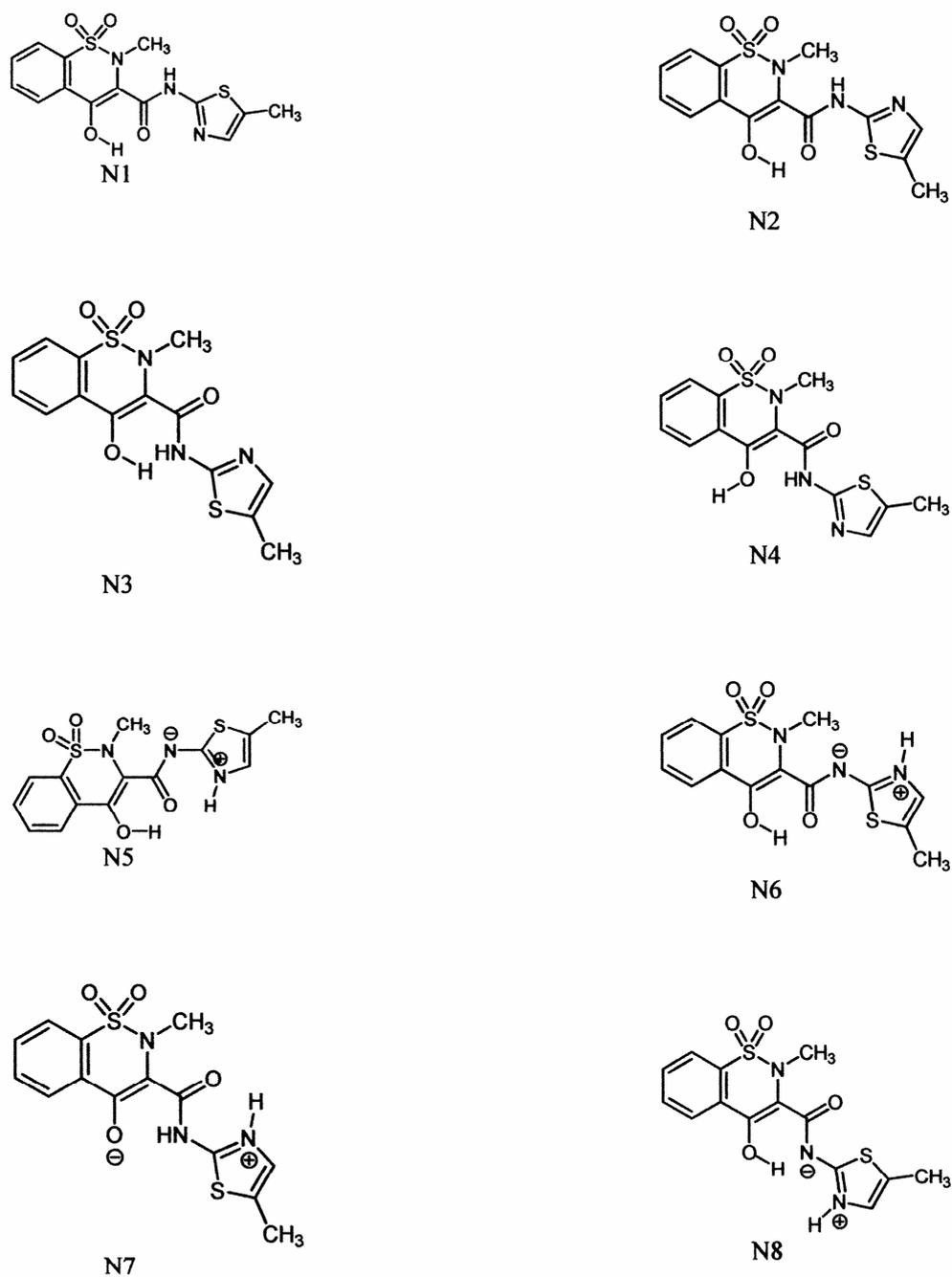
The principal targets of meloxicam are cyclooxygenases, particularly COX-2 (Hawkey 1999), which are membrane-associated enzymes (see above). To bind with the target, meloxicam has to pass through the membrane and hence its interaction

with biomembranes should play a major role in guiding its interaction with cyclooxygenases.

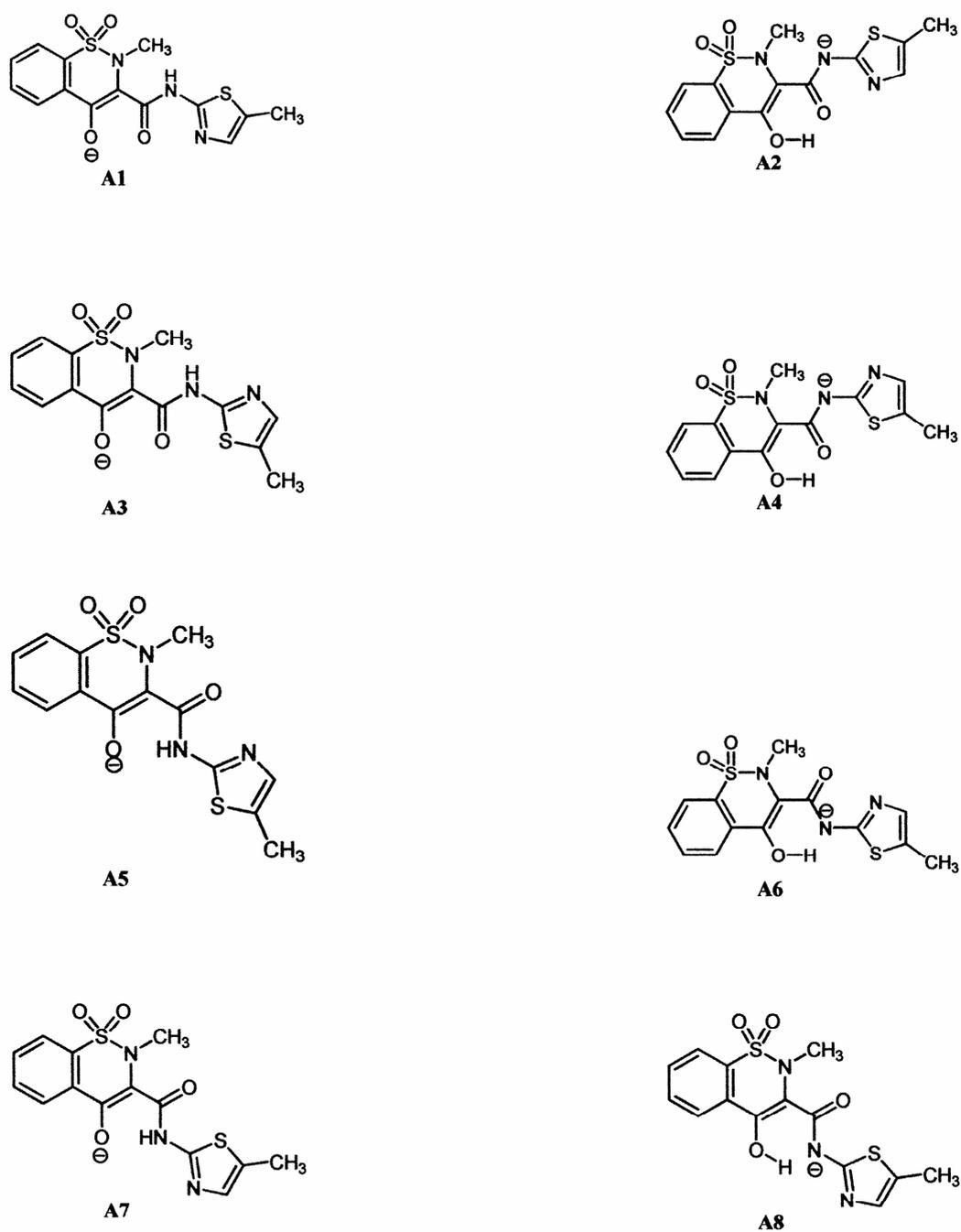
To elucidate this mechanism, Chakraborty et al. (2005) conducted an extensive investigation with small unilamellar vesicles formed by the dimyristoylphosphatidylcholine (DMPC), serving as a membrane mimetic system to study the effect of absence of net surface charges on oxicam NSAID-membrane interaction. This study clearly demonstrated that the different prototropic forms of two oxicams, namely piroxicam and meloxicam could be incorporated into DMPC vesicles, which do not have any net surface charge. In both cases, the neutral form partitions into the vesicles more than the corresponding anionic form. In addition, these results imply that for bio-membranes having no net surface charges, hydrophobic effects would play a principal role in guiding these drugs to their targets.

In the present study DFT B3LYP/6-31G(d,p) calculations have been performed on the inclusion complexes of  $\beta$ -CD with 8 neutral and 8 anionic conformations of meloxicam in two orientations: with its benzene ring near the narrow rim and the wide rim of the CD cavity, respectively.

All calculations were performed in the gas phase using the GAUSSIAN 03 package (Frisch et al., 2004). Structures of eight neutral and eight anionic tautomers of meloxicam were fully optimized at DFT B3LYP/6-31G(d,p) level, in order to identify the most stable species in each case. Figure 39 depicts the most stable neutral conformations of meloxicam (N1-N8), while Figure 40 depicts the most stable anionic conformations of meloxicam (A1-A8).



*Fig. 39* The most stable neutral conformations of meloxicam, N1-N8



*Fig. 40* The most stable anionic conformations of meloxicam, A1-A8

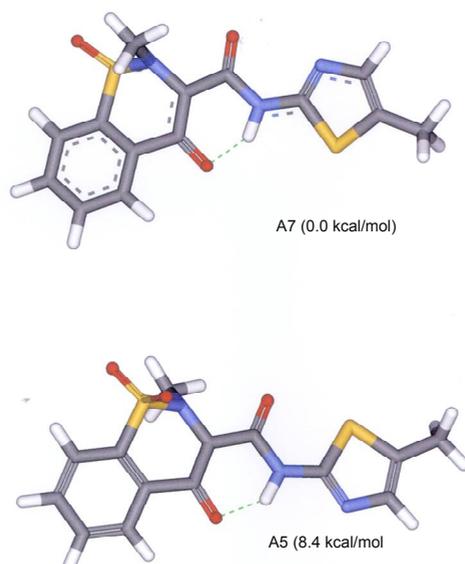
Table 13 shows the energy ranking of the eight neutral and eight anionic meloxicam conformations, respectively.

**Tab.13** Energy- ranking of the most stable neutral and anionic meloxicam conformations. The lowest energy conformation is set to zero in each case.

Meloxicam Conformation	Relative Energies in kcal/mol
N1	6.7
N2	0.0
N3	15.8
N4	14.7
N5	5.3
N6	5.7
N7	5.5
N8	7.8
A1	20.6
A2	14.1
A3	10.6
A4	5.4
A5	8.4
A6	13.2
A7	0.0
A8	8.0

The most stable neutral and anionic species found herein are N2 and A7, respectively.

To get better insight into the intramolecular forces in meloxicam molecules, a closer examination of two meloxicam conformations, viz. A5 and A7 (Figure 41), has been performed.

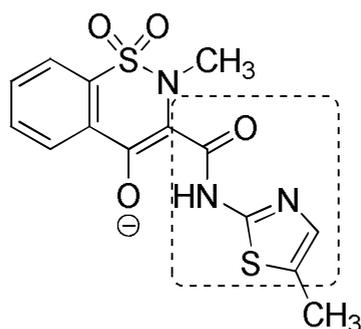


**Fig. 41** Structures of A5 and A7 with its relative energies. Hydrogen bonds are depicted in green.

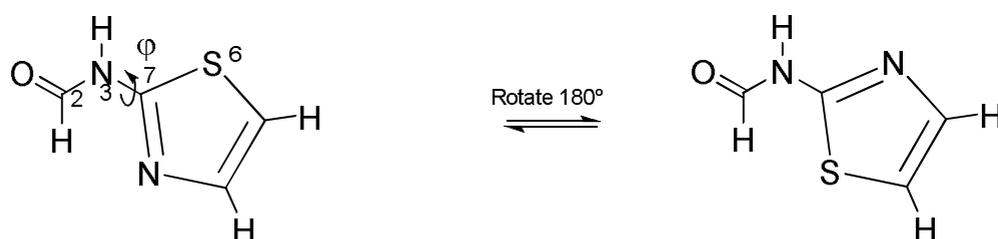
Though these two conformations are quite similar in geometry, their energies differ by an amount of 8.4 kcal/mol. This could be attributed to an extra hydrogen bond. However, as can be seen in Figure 40, both in A5 and A7 only one hydrogen bond is formed between O16 and N17 in each case (For reference numbering see Figure 37).

Therefore, it has been suggested that other forces might be responsible for the calculated energy difference between these two structures.

Structure A5 is converted into structure A7 by rotating dihedral C15-N17-C18-S23 by 180°. For this reason, a model structure MOD was constructed from A5 for further investigations (See Figures 42 and 43).



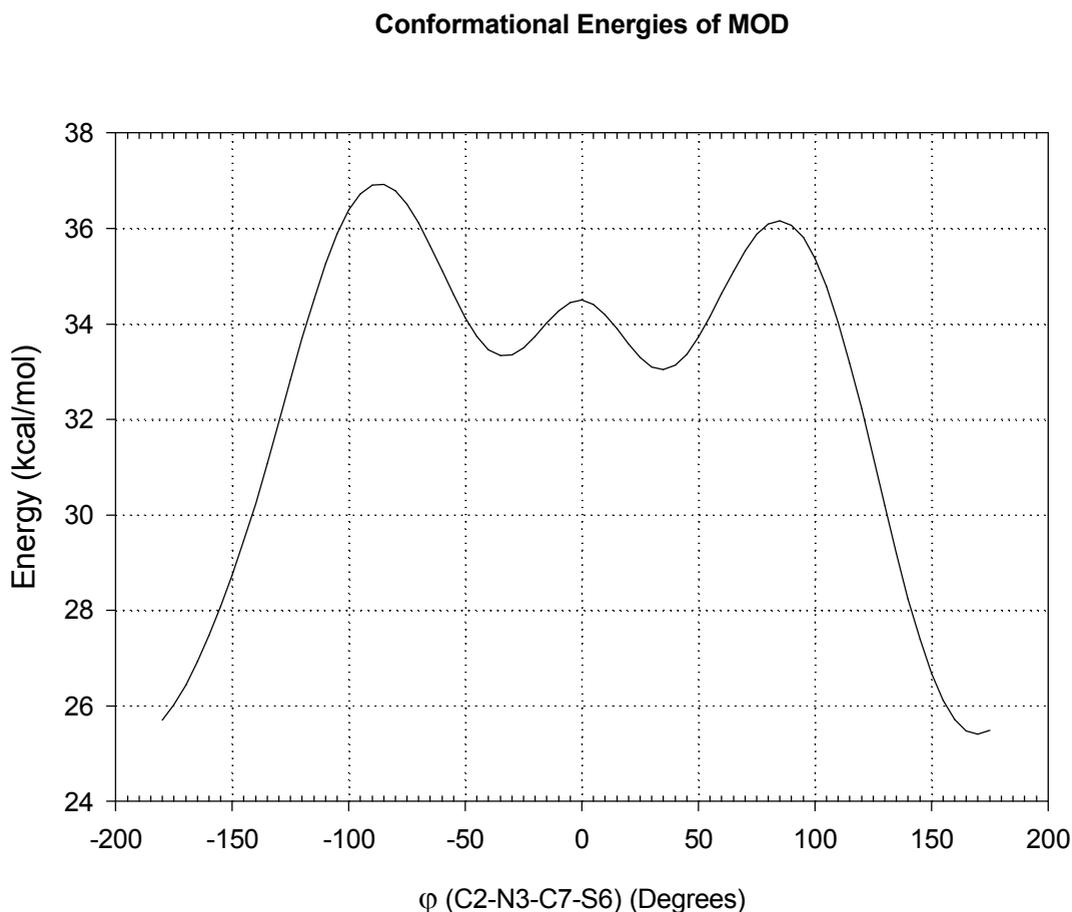
**Fig. 42** Structure of A5; the partial structure (marked by a dashed line) was used to construct the model structure MOD (Figure 43)



**Fig. 43** Structure of MOD. The unrotated and rotated forms are shown

In this structure, dihedral  $\phi$  (C2-N3-C7-S6) was rotated by 180°, using the DIHEDRAL DRIVER of the CHEM 3D program package (see Appendix A).

A plot of the resulting conformational MM2-energies is depicted in Figure 44.



**Fig. 44** Conformational Energies of the model structure MOD. The dependence of the energy on the dihedral  $\phi$  is shown

In comparison, the conformational energies of A5 on rotating the dihedral C15-N17-C18-S23 (For reference numbering see Figure 37) by  $180^\circ$  were assessed in the same way.

Moreover, single point energy calculations of the unrotated and rotated structures of MOD and A5 at different levels of theory have been performed.

The results of these calculations as considered above are presented in Table 14.

**Tab.14** Energy differences between the unrotated and rotated structures of MOD and A5, using different theoretical methods of calculation (Rotation by 180° in each case)

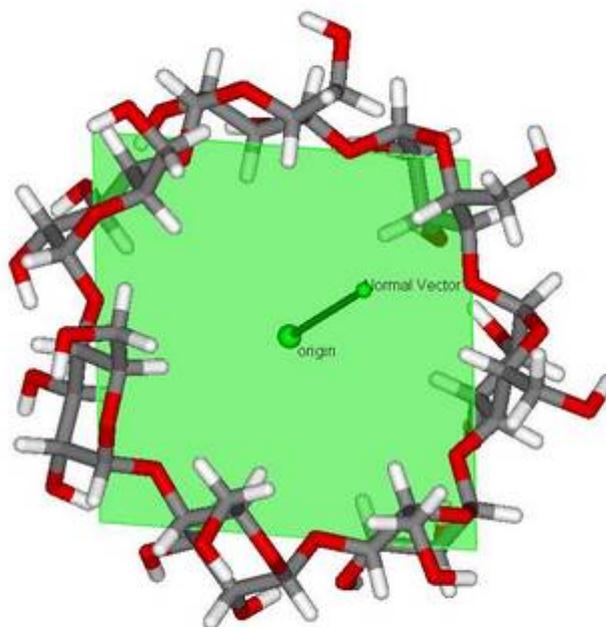
Structure	Calculation Method	$\Delta E$ (kcal/mol)
MOD	MM2	8.8
	B3LYP/6-31G(d,p)	9.6
	MP2/6-311+g(d,p)	9.1
A5	MM2	6.0
	B3LYP/6-31G(d,p)	8.4

The comparison of the results for MOD and A5 supports the suggestion that the calculated energy difference between A5 and A7 might not be caused by hydrogen bonding. For this reason, other forces have to be taken into account:

With respect to the anionic nature of A5 and A7 and the presence of several bond dipoles, it is very likely that electrostatic forces and possibly steric effects instead of hydrogen bonding effects are responsible for this energy difference.

All of the 16 structures listed in Table 13 were considered in the further investigation of the inclusion complexation process between meloxicam and  $\beta$ -CD. The geometry of  $\beta$ -CD was fully optimized at DFT B3LYP/6-31G(d,p) level, without any symmetry constraint.

For the construction of the inclusion complexes, the mass centre of the CD was put in the origin of the reference coordinate system (Figure 45). The glycosidic oxygen atoms were then located in the x-y-plane. In each case, the guest molecule was located perpendicularly to the x-y-plane with its main axis coincident with the normal vector onto the x-y-plane and with its mass centre lying in the mass centre of the CD, considering two possible orientations. The orientation in which the benzene ring of the meloxicam guest points toward the narrow rim of the CD cavity is named “a”, while the other where the benzene ring points to the wide rim is named “b”.



*Fig. 45 Plane of the glycosidic oxygen atoms with the origin of the coordinate system*

The resulting 16 complexes were fully geometry optimized with DFT B3LYP/6-31G(d,p) calculations. Tables 15 and 16 show the energy ranking of the neutral and anionic meloxicam- $\beta$ -CD complexes, respectively.

**Tab. 15** Energy ranking of the neutral meloxicam- $\beta$ -CD complexes. The letters “a” and “b” correspond to the different orientations of the benzene ring of the meloxicam in the CD-cavity (see the text). The lowest energy complex is set to zero.

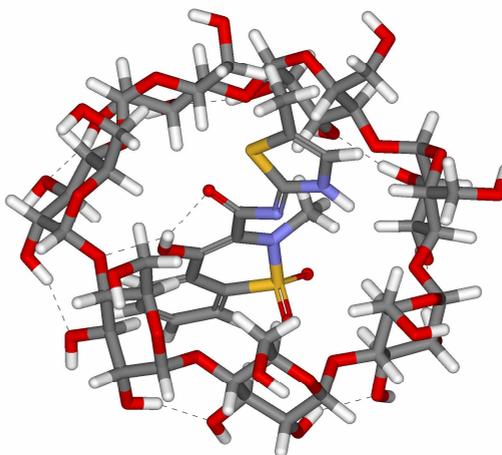
Meloxicam- $\beta$ -CD complex	Relative energies in kcal/mol
N1a	7.8
N1b	11.6
N2a	11.3
N2b	4.7
N3a	8.0
N3b	3.6
N4a	5.5
N4b	3.1
N5a	5.7
N5b	4.1
N6a	3.8
N6b	0
N7a	4.1
N7b	4.7
N8a	5.7
N8b	0.2

**Tab. 16** Energy ranking of the anionic meloxicam- $\beta$ -CD complexes. The letters “a” and “b” correspond to the different orientations of the benzene ring of the meloxicam in the CD-cavity (see the text). The lowest energy complex is set to zero

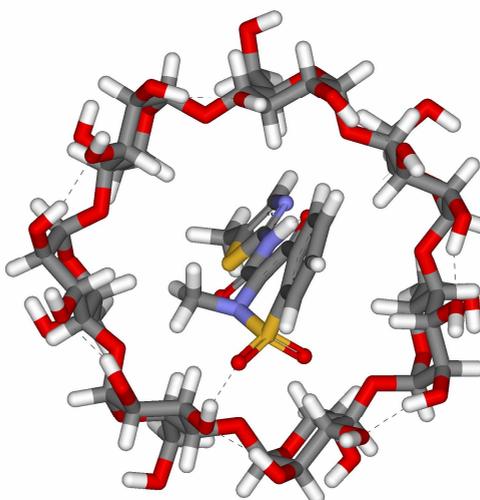
Meloxicam- $\beta$ -CD complex	Relative energies in kcal/mol
A1a	8.7
A1b	2.7
A2a	9.6
A2b	3.1
A3a	12.6
A3b	9.8
A4a	17.8
A4b	16.5
A5a	13
A5b	8.8
A6a	12.9
A6b	4.9
A7a	12.7
A7b	9.7
A8a	4.4
A8b	0

The computational calculations revealed that all host-guest systems are stabilized because of guest penetration into the  $\beta$ -CD cavity. To examine the stability of the complexes, the binding energy was obtained from the difference between the complex energy and the energy sum of the isolated host and guest.

Figures 46 and 47 show the structures corresponding to the most stable neutral and anionic inclusion complexes, respectively.



**Fig. 46** Structure of the most stable neutral meloxicam- $\beta$ -CD inclusion complex N6b, calculated at DFT B3LYP/6-31G(d,p) level. Hydrogen bonds are depicted in black.



**Fig. 47** Structure of the most stable anionic meloxicam- $\beta$ -CD inclusion complex A8b, calculated at DFT B3LYP/6-31G(d,p) level. Hydrogen bonds are depicted in black.

In both cases, the guest is considerably included in the CD cavity, which remains almost undistorted. As can be seen in Figure 46, two extra hydrogen bonds are formed between the neutral meloxicam guest and the CD. Figure 47 shows the formation of one extra hydrogen bond between the anionic conformation of meloxicam and the CD host.

According to the parameters calculated for the hydrogen bonds in the neutral and anionic complexes, these interactions could be classified as moderate hydrogen bonds (Scheiner 1997). It is suggested that these extra hydrogen bonds contribute significantly to the stability of the complexes. However, even though the formation of hydrogen bonds can explain the additional stabilization of the neutral and anionic inclusion complexes between meloxicam and  $\beta$ -CD, it must be kept in mind that inclusion takes place in aqueous solution and other interactions should be involved in order to overcome the energy cost of desolvation of the guest molecule.

To conclude, the complexation of neutral and anionic conformations of meloxicam with  $\beta$ -CD was studied by quantum mechanical B3LYP/6-31G(d,p) density functional theory calculations in the gas phase. The calculations reveal that hydrogen bonding interactions are involved in the complexation of both the neutral and anionic conformations and support considerably the stabilization of the respective complexes.

## 11 Results and Discussion – Part 5

### Papers I-V

The main goal of the theoretical considerations and case studies that have been described above was the selection and evaluation of proper methods to perform computational calculations on CDs and CD inclusion complexes. As the choice of computational methods must be based on a clear understanding of both the chemical system and the information to be computed, a careful researcher must be aware of the merits and drawbacks of various methods and software packages in order to make an informed choice.

In particular, the following points have been highlighted:

- As the quantitative description of non bonded interactions between small polar molecules is a first step toward assessing the significance of these interactions in larger molecular systems such as the CDs, the BSSE behaviour of a hydrogen bonded water-methanol complex at several levels of theory has been investigated in chapter 3.
- Guidelines on how to choose both the best computational method and model chemistry to conduct properly a computational molecular modelling study were mentioned in chapter 4.
- Geometry optimizations on malonaldehyde using several model chemistries in the gas phase were performed in chapter 5 to find out what level of theory is required to produce an accurate structure for this molecule. The results of this study are mainly intended to elect an appropriate method to obtain reliable energies and geometries for further computational modelling of larger molecular systems.
- Because of the importance of chiral recognition in chiral molecules such as CDs, the behaviour of  $\beta$ -CD as chiral selector between the two enantiomeric forms of the amino acid alanine has been investigated in chapter 7.
- Chapter 8 gives an exhaustive description of the various molecular interactions in supramolecular systems with the focal point on hydrogen bonding.
- The structural features and inclusion chemistry of the CDs as supramolecular hosts are presented in chapter 9.
- To gain a better insight into the complexation behaviour of  $\beta$ -CD and due to the extraordinary importance of meloxicam in medicinal and pharmaceutical

applications, a closer examination of inclusion complexes of several neutral and anionic conformations of meloxicam with  $\beta$ -CD was performed in chapter 10.

The results of all these investigations as mentioned above have been summarized in the following papers I-V

### **11.1 Paper I: “Homodromic Hydrogen Bonds in Low-Energy Conformations of Single Molecule CDs”**

As the application of ab initio and DFT calculations was limited in the past by the size of the molecular system, only few theoretical studies on CDs and CD complexes were performed using DFT methods.

The primary goal of Paper I was to get some insight into structure and intramolecular hydrogen bond network of the  $\beta$ -CD molecule under anhydrous conditions in the gas phase using higher level ab initio as well as DFT calculations.

## Homodromic hydrogen bonds in low-energy conformations of single molecule cyclodextrins

Alfred Karpfen · Elisabeth Liedl · Walter Snor ·  
Petra Weiss-Greiler · Helmut Viernstein ·  
Peter Wolschann

Received: 15 May 2006 / Accepted: 20 October 2006  
© Springer Science+Business Media B.V. 2007

**Abstract** Low-energy conformations of  $\beta$ -cyclodextrin under anhydrous conditions in the gas phase were investigated by DFT calculations. In these conformations, two homodromic hydrogen bond rings are formed with very short hydrogen bonds at the narrow side of the cyclodextrin ring and a second one at the wider side. These hypothetical conformations are not comparable to those conformations, which have been studied experimentally, forming inclusion complexes with small and medium-sized guest molecules, but their energy is significantly lower than the open conformations ( $\Delta E = 10$  kcal/mol).

**Keywords** Cyclodextrin · ab initio · Density functional calculations · Homodromic hydrogen bonds

### Introduction

Hydrogen bonding is very important for the structure of cyclodextrin and cyclodextrin complexes. Particularly, crystal structures of  $\beta$ -CD exhibit a number of O-H...O hydrogen bonds between the three different hydroxyl groups and water molecules, which were investigated by X-ray crystallography and by neutron diffraction [1–6].

In addition to intermolecular hydrogen bonds, intramolecular hydrogen bonds are formed between the secondary hydroxyl groups of the glucose units to the adjacent units. These intramolecular hydrogen bonds stabilize the macrocyclic conformation of  $\beta$ -CD, increase the rigidity of the molecule and lead in this special case to a reduced solubility of the compound in comparison to other CDs, with a smaller or larger number of glucose units, and also to substituted derivatives.

The application of ab initio and DFT calculations was limited in the past by the size of the molecular system. Numerous theoretical studies were performed on the structure of CDs and CD complexes, based on semiempirical quantum chemical methods [7]. Accurate molecular calculations were performed only for smaller systems like glucose dimers [8–11]. For the structure of CDs only a few papers were published, comparing empirical, semiempirical and low-level ab initio methods (HF/3-21G), concluding that Molecular Mechanics is the most convenient method to describe structural features of CDs, followed by the ab initio method [12]. Some recent investigations were performed using DFT methods [13].

In the present paper, we describe the gas phase structures of  $\beta$ -CD as obtained from ab initio (HF/3-21G and HF/6-31G(d,p)) as well as DFT (B3LYP/6-31G(d,p)) calculations, in order to get some insight into the intramolecular hydrogen bond network of this molecule.

### Methods of calculation

C7 symmetry was forced by building up the Z-matrix starting from the oxygen-oxygen distances at the more

A. Karpfen · E. Liedl · W. Snor ·  
PetraWeiss-Greiler · PeterWolschann (✉)  
Institute of Theoretical Chemistry, University of Vienna,  
Währinger Straße 17, 1090 Wien, Austria  
e-mail: Karl.Peter.Wolschann@univie.ac.at

H. Viernstein  
Department of Pharmaceutical Technology and  
Biopharmaceutics, University of Vienna, Althanstraße 14,  
1090 Wien, Austria

narrow rim of the  $\beta$ -CD moiety. HF/3-21G and HF/6-31G(d,p), as well as B3LYP/6-31G(d,p) were used for full geometry optimisation. The calculations were performed, using the program package GAUSSIAN03 [14] on the Schrödinger III cluster of the University of Vienna. Scanning of the oxygen-oxygen distance was done in order to obtain the structures of minimum energies.

### Results and discussion

X-ray crystals of CDs and CD complexes are grown from aqueous solution and also most of the equilibrium constants and the related thermodynamic parameters were obtained from solution of pure water or water co-solvent mixtures. Only a very few examples are known for CD-guest complexation in non-aqueous conditions [15, 16].

In order to investigate the conformational behaviour of CDs, we have analysed the gas phase conformations of  $\beta$ -CD under symmetric conditions (C7) by a systematic scan of the oxygen-oxygen distance of the O6 hydroxyl groups. Three conformational minima were found.

The energies of the minimum conformations are given in Table 1, together with selected structural parameters.

The lowest energy minimum (HF/3-21G and B3LYP/6-31G(d,p)) was found for a rather short O6-O6 distance, a second less pronounced at a slightly elongated distance, a third minimum at a distance around 6.5 Å, which is similar to the experimentally (X-ray crystallographic) determined structure.

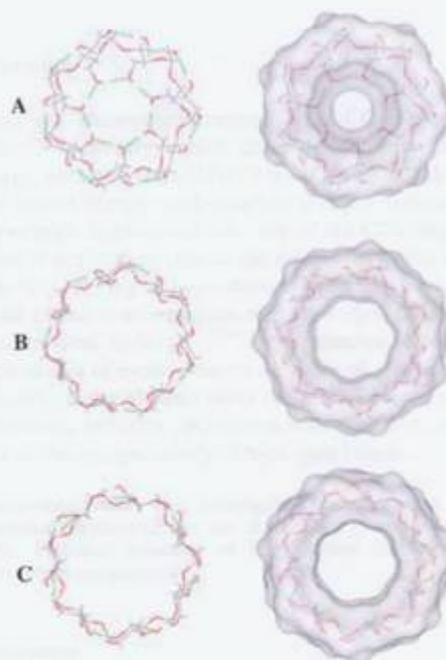
Although all methods used describe a very similar geometry, the energy differences between the obtained minima vary drastically. Even the energy ranking is described differently by HF/6-31G(d,p), whereas HF/3-21G and B3LYP/6-31G(d,p) calculations lead to the same ranking. The reason for these discrepancies is that hydrogen bonding superimposed by quite subtle steric interactions determines the geometry of the molecule. A higher-level DFT method is therefore necessary to describe the geometries sufficiently well, and this method will be used for further considerations.

**Table 1** Calculated energies of three symmetric  $\beta$ -CD conformations, relative to the energy of conformation A

	Energy values (kcal mol <sup>-1</sup> )		
	HF/3-21G	HF/6-31G(d,p)	B3LYP/6-31G(d,p)
A	0.0	0.0	0.0
B	43.4	10.6	18.6
C	24.1	-0.6	10.5

The geometries of the three symmetric conformations are depicted in Fig. 1. The top view shows that the lowest energy minimum (A), which is 10 kcal/mol more stable than the second one (C), possesses a very narrow rim at the upper side of the  $\beta$ -CD ring. Both other conformations are somewhat more open, and differ in the arrangement of the glucose units. Some selected structural parameters are given in Table 2 including the numbering of the atoms for one glucose unit in the adjacent scheme. Scheme 1

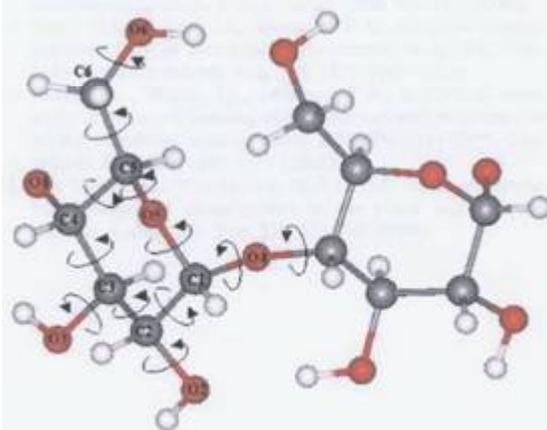
Table 2 is divided into four groups of structural parameters. In the first three rows, the distances between oxygen atoms are reported, which are involved in hydrogen bonding contacts at least in one conformation. Strong hydrogen bonds exist in conformation A within the ring built up from oxygens 6 (O6). This distance is very close to those found for hydrogen bond networks in open chains and cyclic polymers of water and alcohols [17]. Also the valence angle of the hydrogen bond is rather close to the optimal value of 180°. A second hydrogen bond system occurs between oxygens O2 and O3. These interactions are somewhat weaker, as the heavy atom distance is longer and also the valence angle is less favourable. The shortest distance exhibits conformation C. The third distance



**Fig. 1** Top-view of the symmetric conformations of  $\beta$ -CD (A, B, and C). On the left side stick presentation, including hydrogen bonds, on the right side the molecular surfaces

**Table 2** Structural parameters of three symmetrical conformational minima of  $\beta$ -CD as obtained from B3LYP/6-31G(d,p) calculations. Distances in Å, torsional angles in degrees

Structural parameters			
	A	B	C
Distance O6-O6'	2.80 (173.7)	4.24	6.11
Distance O3-O2'	3.08 (163.5)	2.84 (164.0)	2.84 (164.4)
Distance O2-O3'	2.86 (108.8)	2.78 (111.7)	2.77 (112.0)
Orientation of hydroxyl groups			
H6-O6-C6-C5	-61.9	-56.7	56.6
C4-C3-O3-H3	52.3	52.4	50.1
C3-C2-O2-H2	-46.7	-43.6	-43.6
Outer-ring connections			
O6-C6-C5-C4	-122.5	174.5	61.2
C2-C1-O4-C4'	-138.7	-124.1	-124.1
C1-O4-C4'-C3'	131.1	122.9	124.1
Connections in the glucose units			
C6-C5-C4-C3	180.0	-172.7	-169.0
C5-C4-C3-C2	54.6	52.4	52.4
C4-C3-C2-C1	-50.5	-52.9	-54.9
C3-C2-C1-O5	53.4	56.0	56.8
C2-C1-O5-C5	-62.8	-62.4	-59.2
C1-O5-C5-C4	65.5	62.1	57.2

**Scheme 1**

(O2-O3) describes a very weak hydrogen bond contact for all conformations. The orientation of the hydroxyl groups is identical for all conformations except for the dihedral angle O6-C6-C5-C4 (third group in the table).

Surprisingly small changes occur at the bonds connecting the glucose rings together, only the above-mentioned orientation of the  $-\text{CH}_2\text{OH}$  group changes completely. Remarkably, some distortions of the glucose rings can be observed for the different conformations.

Because of the co-operativity of the homodromic hydrogen bonds at the upper and the lower side of the

**Table 3** Calculated energies of the conformers derived from conformation A. Conformation A1 is taken as reference

Conformations	Energy values (kcal mol <sup>-1</sup> )
A1	0.00
A2	2.15
A3	0.04
A4	2.30

$\beta$ -CD ring, four conformations of A with different energies are possible as a consequence of the chirality of the glucose rings. For conformation A these energies are given in Table 3.

The top view of the molecule (Fig. 1) indicates, that the strong hydrogen bonds at the smaller rim of  $\beta$ -CD are oriented counter clockwise, whereas the more extended ring is oriented clockwise. This conformation shows the lowest conformational energy (A1, reference in Table 3). Flipping of the orientation of the hydrogen bonds at the larger ring has almost no influence on the energy and enhances the energy value only slightly (conformation A3). Reorientation of the ring of the strong hydrogen bonds leads to a significant increase of the energy (2.15 kcal mol<sup>-1</sup> for the clockwise orientation of the larger ring (A2) and 2.30 kcal mol<sup>-1</sup> for the counterclockwise one (A4).

## Discussion

Three gas phase conformations of  $\beta$ -CD were found with  $C_7$  symmetry, which differ significantly in their energy, obtained from B3LYP/6-31G(d,p) calculations. The lowest-energy conformation is characterized by a surprisingly tight ring at one side of the CDs ring. This shape is not comparable to the structures found experimentally. A ring of homodromic hydrogen bonds was found in this conformation, with hydrogen bonds very close to ideal hydrogen bridge geometries obtained in open chains of cyclic clusters of water or alcohols. The chirality of the glucose units of  $\beta$ -CD induces energy differences between geometries with different orientation of the co-operativity of hydrogen bonds.

**Acknowledgements** This investigation was supported by the Hochschuljubiläumstiftung der Stadt Wien (Project P H-778/2005). Technical assistance of Ms. Martina Ziehegraser is gratefully acknowledged.

## References

1. Chacko, K.K., Saenger, W.: Topography of cyclodextrin inclusion complexes. 15. crystal and molecular structure of the cyclohexaamylose-7.57 water complex. form III. four-

- and six-membered circular hydrogen bonds. *J. Am. Chem. Soc.* **103**, 1708–1715 (1981).
- Betzel, C., Saenger, W., Hingerty, B.E., Brown, G.M.: Circular and flip-flop hydrogen bonding in  $\beta$ -cyclodextrin undecahydrate: a neutron diffraction study? *J. Am. Chem. Soc.* **106**, 7545–7557 (1984).
- Zabel, V., Saenger, W., Mason, S.A.: Neutron diffraction study of the hydrogen bonding in  $\beta$ -cyclodextrin undecahydrate at 120 K: from dynamic flip-flops to static homodromic chains. *J. Am. Chem. Soc.* **108**, 3664–3673 (1986).
- Steiner, T., Mason, S.A., Saenger, W.: Disordered guest and water molecules. Three-center and flip-flop O–H...O hydrogen bonds in crystalline,  $\beta$ -cyclodextrin ethanol octahydrate at  $T = 295$  K: a neutron and X-ray diffraction study? *J. Am. Chem. Soc.* **113**, 5676–5687 (1991).
- Steiner, T., Koellner, G.: Crystalline  $\beta$ -cyclodextrin hydrate at various humidities: fast, continuous, and reversible dehydration studied by X-ray diffraction. *J. Am. Chem. Soc.* **116**, 5122–5128 (1994).
- Saenger, W., Jacob, J., Gessler, K., Steiner, T., Hoffmann, D., Sanbe, H., Koizumi, K., Smith, S.M., Takaha, T.: Structures of the common cyclodextrins and their larger analogues-beyond the doughnut. *Chem. Rev.* **98**, 1787–1802 (1998).
- Liu, L., Guo, Q.-X.: Use of quantum chemical methods to study cyclodextrin chemistry. *J. Incl. Phenom. Macrocycl. Chem.* **50**, 95–103 (2004).
- French, A.D., Kelterer A.M., Johnson, Dowd M.K.: B3LYP/6-31G\*, RHF/6-31G\* and MM3 heats of formation of disaccharide analogs. *J. Mol. Struct.* **556**, 303–313 (2000).
- Strati, G.L., Willett, J.L., Momany, F.A.: Ab initio computational study of  $\beta$ -cellobiose conformers using B3LYP/6-311++G\*\*. *Carbohydr. Res.* **337**, 1833–1849 (2002).
- Strati, G.L., Willett, J.L., Momany, F.A.: A DFT/ab initio study of hydrogen bonding and conformational preference in model cellobiose analogs using B3LYP/6-311++G\*\*. *Carbohydr. Res.* **337**, 1851–1859 (2002).
- Da Silva, C.O., Nascimento, M.A.C.: Ab initio conformational maps for disaccharides in gas phase and aqueous
12. Britto, M.A.F.O., Nascimento, C.S., dos Santos, H.F.: Structural analysis of cyclodextrins: a comparative study of classical and quantum mechanical methods. *Quim. Nova.* **27**, 882–888 (2004).
13. Avakyan, V.G., Nazarov, V.B., Voronezhcheva, N.I.: DFT and PM3 calculations of the formation enthalpies and intramolecular H-bond energies in alpha-, beta-, and gamma-cyclodextrins. *Russ. J. Phys. Chem.* **79**, S18–S27 (2005).
14. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery, J.A., Jr., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar, S.S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G.A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J.E., Hratchian, H.P., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Ayala, P. Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels, A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nanayakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B., Chen, W., Wong, M.W., Gonzales, C., Pople, J.A.: Gaussian Inc., Wallingford CT (2004).
15. Shannigrahi, M., Bagchi, S.: Time resolved fluorescence study of ketocyanine dye- $\beta$  cyclodextrin interactions in aqueous and non-aqueous media. *Chem. Phys. Letters.* **403**, 55–61 (2005).
16. Panja, S., Chowdhury, P., Chakravorti, S.: Modulation of complexation of 4(1H-pyrrole 1-yl) benzoic acid with  $\beta$ -cyclodextrin in aqueous and non-aqueous environments. *Chem. Phys. Letters.* **393**, 409–415 (2004).
17. Karpfen, A.: Cooperative effects in hydrogen bonding. *Adv. Chem. Phys.* **123**, 469–510 (2002).

## **11.2 Paper II: "On the Structure of Anhydrous $\beta$ -CD"**

In continuation of Paper I, main emphasis was laid on the orientation of the hydrogen bonds in the  $\beta$ -CD molecule that can be clockwise or counterclockwise. Four possibilities for the orientation of the homodromic rings of hydrogen bonds have been taken into account.



## On the structure of anhydrous $\beta$ -cyclodextrin

Walter Snor<sup>a</sup>, Elisabeth Liedl<sup>a</sup>, Petra Weiss-Greiler<sup>a</sup>, Alfred Karpfen<sup>a</sup>,  
Helmut Viernstein<sup>b</sup>, Peter Wolschann<sup>a,\*</sup>

<sup>a</sup> University of Vienna, Institute of Theoretical Chemistry, Währinger Straße 17, A-1090 Wien, Austria

<sup>b</sup> University of Vienna, Department of Pharmaceutical Technology and Biopharmaceutics, Althanstraße 14, A-1090 Wien, Austria

Received 9 March 2007; in final form 30 April 2007

Available online 6 May 2007

### Abstract

Low-energy conformations of  $\beta$ -cyclodextrin under anhydrous conditions in the gas phase were investigated by DFT calculations. In one conformation, two homodromic hydrogen bond rings are formed with very short hydrogen bonds at the narrow side of the cyclodextrin ring and a second one at the wider side. Four possibilities for the orientation of the homodromic rings have been taken into account. These 'closed' conformations differ from conformations which have been studied experimentally, forming inclusion complexes with small and medium-sized guest molecules. Their energy is significantly lower than those of the open conformations ( $\Delta E \approx 44$  kJ mol<sup>-1</sup>).

© 2007 Elsevier B.V. All rights reserved.

### 1. Introduction

Natural cyclodextrins (CDs) consist of several linked glucose subunits forming a pore, where small or medium-sized molecules can be included. There exists a broad variety of applications of such inclusion complexes, particularly in pharmacy and pharmaceutical technology [1,2]. The interior of CDs is to some extent hydrophobic, nevertheless, hydrogen bonding plays an essential role for the structure of CDs and their inclusion complexes. The shape of CDs resembles that of a broad ring. A rim of primary hydroxyl groups is found on one side of the molecule and another of secondary hydroxyl groups exists on the opposite side [3]. In aqueous solution, the intramolecular hydrogen bonds compete with intermolecular associations to water molecules, influencing the structure and the solubility of various CDs, e.g. the inverse freezing of  $\alpha$ -CD, reported recently is a consequence of such a competition [4–7].

Crystal structures of  $\beta$ -CD with varying numbers of water molecules are available from X-ray crystallography and from neutron diffraction [8–13]. However, not much is known about the geometry of  $\beta$ -CD under anhydrous conditions. Thermodynamic investigations were performed on anhydrous  $\beta$ -CD [14,15], and hydrogen bonding connected with flip-flop motion in the secondary hydroxyl groups' rim was considered. Investigations on various conformations of  $\beta$ -CD, including open and closed conformations, were performed using empirical force fields [16]. Conformations of anhydrous  $\beta$ -CD in the gas phase were investigated by DFT calculations just recently [17–19], where hydrogen bonding in one rim was studied and the related energies were estimated. A preliminary report of this work has been given at a recent conference [20].

### 2. Methods of calculation

We have applied systematic B3LYP calculations with 6–31G(d,p) basis set on anhydrous  $\beta$ -CD. Imposing  $C_7$  symmetry throughout, the oxygen–oxygen distances were scanned and full geometry optimisations of all remaining geometry parameters were then performed, using the program package GAUSSIAN03 [21].

\* Corresponding author. Fax: +1 4277 9527.

E-mail address: Karl.Peter.Wolschann@univie.ac.at (P. Wolschann).

### 3. Results and discussion

The dependence of the minimized energy on the oxygen–oxygen distance is depicted in Fig. 1. Three conformational minima were found. By subsequent frequency calculations, it was verified that these geometries are real minima on the potential energy surface. Their energies together with the distances of the oxygen–oxygen atoms of the primary hydroxyl groups and the oxygen–oxygen distances at the secondary hydroxyl group rim are given in Table 1. For the latter, two alternating distances occur, one with hydrogen bonds, the other without hydrogen atoms in a proper geometry between the oxygen atoms. In the lowest energy conformation (A), two homodromic hydrogen bond rings are formed, one with very short hydrogen bonds at the side of the primary hydroxyl groups of the cyclodextrin ring and the second with the secondary hydroxyl groups. The geometry of this conformation of  $\beta$ -CD is shown in Fig. 2, where both hydrogen bond rings are given as dotted lines in green<sup>1</sup> (primary hydroxyl groups) and violet (secondary hydroxyl groups). Such a conformation has not been found experimentally up to now and corresponds to a *one-gate-closed* geometry of  $\beta$ -CD.

Two further minimum conformations (B and C) were detected with larger oxygen–oxygen distances of the primary hydroxyl groups, and no hydrogen bonds between these groups. Such open conformations exhibit significantly higher energies than the low-energy conformation A, as a consequence of the missing hydrogen bonds. In both these conformations, only one hydrogen bond rim built from the secondary hydroxyl groups exists. The open conformation C is similar to non-symmetric conformations found in crystal structures and, moreover, by semiempirical and lower level *ab initio* calculations [22–24].

The orientation of the hydrogen bonds can be clockwise (cw) or counterclockwise (cc) as presented in Scheme 1. As a consequence of the asymmetry of the glucose units, four different conformations can be constructed. The top view of the molecule (smaller ring in Fig. 2) indicates that the strong hydrogen bonds at the closed gate of  $\beta$ -CD are oriented counter clockwise (cc), whereas at the more extended ring they are oriented clockwise (cw) or counter clockwise (cw). Flipping of the orientation of the hydrogen bonds in the larger ring has almost no influence on the energy. Reorientation of the ring of the strong hydrogen bonds leads to a significant increase of the energy ( $9.04 \text{ kJ mol}^{-1}$ ) for the cw orientation of the larger ring and almost the same energy ( $9.63 \text{ kJ mol}^{-1}$ ) for the cc one. The oxygen–oxygen distances of the primary hydroxyl groups vary between 2.80 and 2.82 Å only. Surprisingly small changes occur also at the distances of the secondary hydroxyl groups which are between 2.86 and 2.89 Å for the hydrogen bonds, and 3.08 to 3.12 Å for the non-hydrogen-bonded

<sup>1</sup> For interpretation of color in Fig. 2, the reader is referred to the web version of this article.

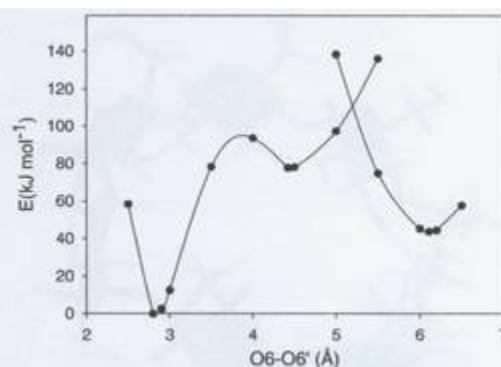


Fig. 1. Energy of  $\beta$ -CD obtained from B3LYP/6-31G(d,p) calculations in dependence on the oxygen–oxygen distance of the primary hydroxyl groups (O6–O6').

Table 1

Energy values and structural parameters of three symmetric  $\beta$ -CD conformations (cccw), relative to the energy of conformation A (*one-gate-closed* conformation) as obtained from B3LYP/6-31G(d,p) calculations

Conformation	Energy in kJ mol <sup>-1</sup>	Primary H-bonds (O6–O6')	Secondary H-bond (O2–O3)	Secondary H-bond (O2–O3')
A	0.0	2.80	2.86	3.08
B	77.8 (68.8) <sup>a</sup>	4.42	2.78	2.84
C	43.8 (35.6)	6.11	2.77	2.84

Distances in Å.

<sup>a</sup> Zero point energy corrected.

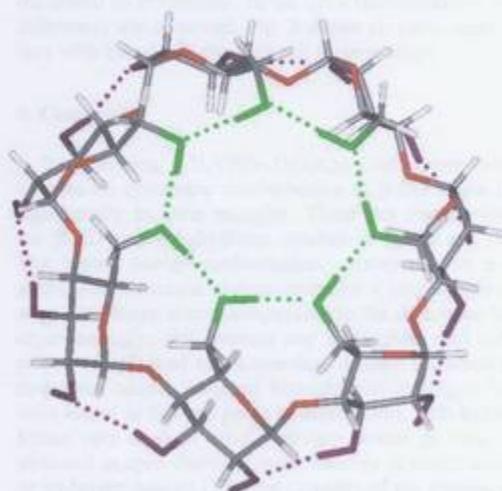
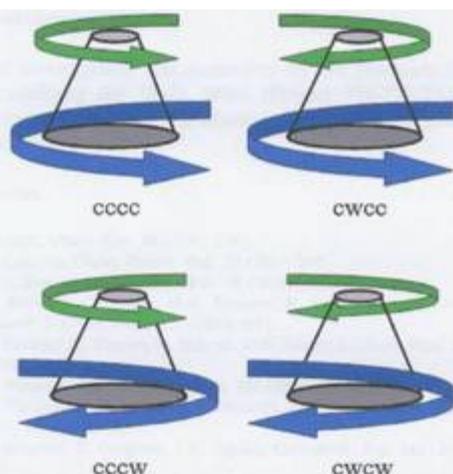


Fig. 2. Top view of the symmetric conformation A (*one gate closed* conformation) of  $\beta$ -CD. Stick presentation, including hydrogen bonds oriented ccw.



Scheme 1. Presentation of the possible orientations of the hydrogen bonds in conformation A.

oxygen–oxygen contacts. An overview of the distances for the four isomers is given in Table 2.

The geometries of the three symmetric conformations were calculated taking into account the orientation of the

Table 2  
Energy values and structural parameters of the four possible orientations of the hydrogen bonds of the symmetrical conformational minima A of  $\beta$ -CD (*one-gate-closed* conformation) as obtained from B3LYP/6-31G(d,p) calculations

Orientation of the H-bonds	Energy in $\text{kJ mol}^{-1}$	Primary H-bonds (O6–O6')	Secondary H-bond (O2–O3)	Secondary H-bond (O2–O3')
cccw	0.0	2.80	2.86	3.08
cccc	0.0	2.80	2.89	3.09
cwcw	8.8	2.82	2.86	3.11
cwcc	9.4	2.82	2.89	3.12

Distances in Å.

Table 3  
Geometrical data of the cones of different conformations of  $\beta$ -CD under consideration of the orientation of the homodromic hydrogen bond rings (measured between O6 and O3)

Conformations of $\beta$ -CD		$r_1$ (O6) Å	$r_2$ (O3) Å	$\tau(^{\circ})$	$\sigma(^{\circ})$
cccw	A	3.23	6.45	57.3	70.4
	B	5.09	6.12	80.3	82.3
	C	7.04	6.05	101.1	83.5
cccc	A	3.23	6.45	57.2	70.4
	B	5.09	6.02	81.2	82.4
	C	7.04	5.96	102.2	84.2
cwcw	A	3.25	6.55	57.3	69.7
	C	6.14	6.32	88.0	76.2
cwcc	A	3.25	6.49	57.7	69.7
	C	6.15	6.23	89.1	76.2

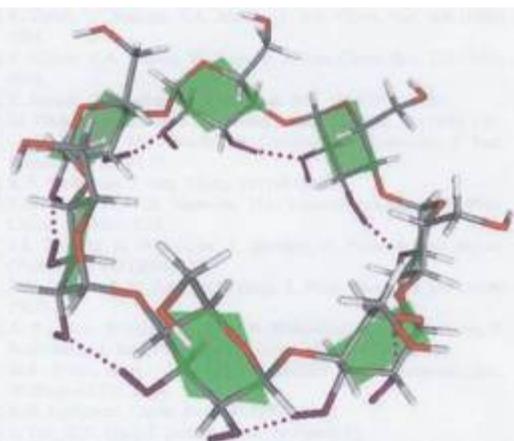


Fig. 3. Open conformation (C orientation of the H-bonds ccw) of  $\beta$ -CD with average planes through the glucose rings.

hydrogen bonds. For the orientations cwcc, no conformational minima corresponding to B were found. To compare the different structures of the three conformations, the angle  $\tau$  was calculated from the difference of the cone radii of O6 and of O3 and the height of the cones. The results are presented in Table 3. For the closed conformations A, the difference of  $\tau$  is negligible, whereas for the open conformations B and C,  $\tau$  varies between  $80.3^{\circ}$  and  $101.07^{\circ}$ . The tilt angle  $\sigma$  of the average planes through the glucose rings is also depicted in Table 3. Again, there is no influence of the orientation of the hydrogen bonds on these angles for the closed conformation. In the open conformations larger differences are observed. Fig. 3 shows an open conformation with the planes through the glucose rings.

#### 4. Conclusion

Summarizing, B3LYP/6-31G(d,p) calculations resulted in three  $C_7$  symmetric conformation of  $\beta$ -CD which differ significantly in their energies. These are representations for  $\beta$ -CD under anhydrous conditions in the gas phase. The lowest energy conformation corresponds to a *one-gate-closed* structure, characterized by a surprisingly tight ring. This shape is not comparable to the structures found experimentally, also because any interaction with solvent molecules will lead to an opening of some intramolecular hydrogen bonds. Rings of homodromic hydrogen bonds were found in this gas phase conformation, with hydrogen bonds very close to ideal hydrogen bridge geometries as obtained in open chains of cyclic clusters of water, alcohols or hydrogen halides [25]. The chirality of the glucose units of  $\beta$ -CD induces energy differences between geometries with different orientation of the cooperativity of hydrogen bonds.

**Acknowledgements**

This investigation was supported by the Hochschuljubiläumstiftung der Stadt Wien (Project PH-778/2005). Technical assistance of Ms. Martina Ziehengraser is gratefully acknowledged.

**References**

- [1] J. Szejtli, *Chem. Rev.* 98 (1998) 1743.
- [2] K. Uekama, *Chem. Pharm. Bull.* 52 (2004) 900.
- [3] W.R. Saenger et al., *Chem. Rev.* 98 (1998) 1787.
- [4] M. Plazanet, C. Floare, M.R. Johnson, R. Schweins, H.P. Trommsdorff, *J. Chem. Phys.* 121 (2004) 5031.
- [5] E. Tombari, C. Ferrari, G. Salvetti, G.P. Johari, *J. Chem. Phys.* 123 (2005) 051104.
- [6] M. Plazanet et al., *J. Chem. Phys.* 125 (2006) 154504.
- [7] M. Plazanet, M.R. Johnson, R. Schweins, H.P. Trommsdorff, *Chem. Phys.* 331 (2006) 35.
- [8] E. Sabadini, T. Cosgrove, F.C. Egidio, *Carbohydr. Res.* 341 (2006) 270.
- [9] K.K. Chacko, W. Saenger, *J. Am. Chem. Soc.* 103 (1981) 1708.
- [10] C. Betzel, W. Saenger, B.E. Hingerty, G.M. Brown, *J. Am. Chem. Soc.* 106 (1984) 7545.
- [11] V. Zabel, W. Saenger, S.A. Mason, *J. Am. Chem. Soc.* 108 (1986) 3664.
- [12] T. Steiner, S.A. Mason, W. Saenger, *J. Am. Chem. Soc.* 113 (1991) 5676.
- [13] T. Steiner, G. Koellner, *J. Am. Chem. Soc.* 116 (1994) 5122.
- [14] M. Diot, C. de Braver, P. Germain, *J. Incl. Phenom.* 30 (1998) 143.
- [15] C. De Brauer, M.P. Merlin, P. Germain, T. Guerlandel, *J. Incl. Phenom.* 37 (2000) 75.
- [16] K.B. Lipkowitz, *J. Org. Chem.* 56 (1991) 6357.
- [17] V.G. Avakyan, V.B. Nazarov, N.I. Voronezhcheva, *Russ. J. Phys. Chem.* 79 (2005) S.18.
- [18] J.A. Dobado, N. Benkadour, S. Melchor, D. Portal, *J. Mol. Struct. (Theochem)* 672 (2004) 127.
- [19] R.V. Pinjari, K.A. Joshi, S.P. Gejji, *J. Phys. Chem. A* 110 (2006) 13073.
- [20] A. Karpfen, E. Liedl, W. Snot, P. Weiss-Greiler, H. Viernstein, P. Wolschann, *J. Incl. Phenom. Macrocycl. Chem.* 57 (2007) 35.
- [21] M.J. Frisch et al., *GAUSSIAN 03*, Revision C.02. Gaussian, Inc., Wallingford CT, 2004.
- [22] K.B. Lipkowitz, *Chem. Rev.* 98 (1998) 1829.
- [23] L. Liu, Q.X. Guo, *J. Incl. Phenom.* 50 (2004) 95.
- [24] X.S. Li, L. Liu, T.W. Mu, Q.X. Guo, *Monatsh. Chem.* 131 (2000) 849.
- [25] A. Karpfen, *Adv. Chem. Phys.* 123 (2002) 469.

### **11.3 Paper III: “Density Functional Calculations on CDs”**

In contrast to Papers I+II, where only conformations of  $\beta$ -CD have been taken into account, Paper III takes a closer look at various conformations of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs under anhydrous conditions in the gas phase using DFT methods. Moreover, the calculated structural parameters were compared with related structural parameters that have been obtained experimentally.

Monatsh Chem (2008)  
DOI 10.1007/s00706-007-0811-2  
Printed in The Netherlands

Monatshefte für Chemie  
Chemical Monthly

## Density functional calculations on cyclodextrins

Alfred Karpfen<sup>1</sup>, Elisabeth Liedl<sup>1</sup>, Walter Snor<sup>1</sup>, Helmut Viernstein<sup>2</sup>, Petra Weiss-Greiler<sup>1</sup>, Peter Wolschann<sup>1</sup>

<sup>1</sup> Institute of Theoretical Chemistry, University of Vienna, Wien, Austria

<sup>2</sup> Department of Pharmaceutical Technology and Biopharmaceutics, University of Vienna, Wien, Austria

Received 30 August 2007; Accepted 18 September 2007; Published online 14 March 2008

© Springer-Verlag 2008

**Abstract** Conformations of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs under anhydrous conditions in the gas phase were investigated by a density functional method, *B3LYP/6-31G(d,p)*. These calculations resulted in several symmetric conformations with different energies. The lowest energy conformations contain two rings of homodromic hydrogen bonds and are referred to "one-gate-closed" conformations. Different orientations of hydrogen bonds lead to four minima. Other conformational minima were found for "open" conformations which correspond to some extent to experimentally determined structures.

**Keywords** Hydrogen bonding; Conformation; Molecular modeling; Homodromic.

### Introduction

Cyclodextrins (CDs) are indispensable recipients not only in pharmacy and pharmaceutical technology, but also in many other scientific disciplines, like environmental, technical and analytical chemistry, for stereo-specific separations of diastereomers and optical isomers, extraction of natural products, protection and stabilization of light-, temperature-, or oxidation-sensitive compounds. The reason for this broad field of applications of CDs is their ability to

form inclusion complexes with small or even medium-sized organic or inorganic compounds. Such an inclusion influences the physico-chemical behavior of the guest molecules, like the reactivity or the solubility significantly. Emulsification of highly apolar compounds, change of the catalytic activities, support in organic syntheses, masking of odor or taste, increase of bioavailability and subsequently higher efficiency of the active substance as a consequence of solubility enhancement, and the permission of controlled release are topics of actual CD research.

Steric as well as electronic parameters of both the CDs and the guest molecules determine the driving forces of the complexation and the geometries of the inclusion complexes. Also the use of various CDs and CD derivatives enhances the variability of applications tremendously. Many review articles have been published, which give excellent overviews about a large number of applications and detailed descriptions of molecular properties of CDs and CD complexes [1–4]. Particularly, as a consequence of the high importance of CDs in pharmaceutical applications many extensive reviews have been published [5–13]. Finally, a review emphasizing historical development perspectives of pharmaceutical applications has been presented quite recently [14].

Native CDs are obtained by the degradation of starch [ $\alpha(1\rightarrow4)$  linked polyglucose] by  $\alpha$ -1,4-glucan-glycosyltransferases. Depending on the respective transferase, different types of CDs result, consisting of 6 ( $\alpha$ -CD), 7 ( $\beta$ -CD), or 8 ( $\gamma$ -CD)

Correspondence: Peter Wolschann, Institute of Theoretical Chemistry, University of Vienna, Währinger Straße 17, A-1090 Wien, Austria. E-mail: karl.peter.wolschann@univie.ac.at

$\alpha(1\rightarrow4)$  linked glucose units. There also exist larger CDs, as  $\delta$ -,  $\epsilon$ -, and  $\iota$ -CDs with 9, 10, and 14 glucose residues. CDs with higher degrees of polymerization up to several hundred glycosyl units produced by a number of  $\alpha$ -4-glucotransferase have been also described briefly [15–18]. Because of the high ring flexibility and the consequently distorted interior, the importance of these CDs is rather low. Modified CDs have been synthesized to change their inclusion properties and also to induce biomimetic functions. Random mono- and dimethylation as well as permethylation have been applied widely, leading to an increase of the solubility and to changes of the thermodynamic parameters of host guest association [19]. Other modified CD derivatives result from the hydroxypropyl substitution at the positions O6 and O2 [20]. Monosubstitution of CDs offers the entrance into a broad field of applications as supramolecular capsules. Entities grafting on one of the CD faces are so called “caps”, whereas caps attached to only one glucose unit are called “flexible caps” and in the case of “rigid caps” an organic fragment forms a bridge by linking two glucose units at one end of the CD. These caps induce significant cavity distortions and therefore have been believed to provide better complementarity between host and guests leading to higher binding constants or higher catalytic reaction rates [21]. The use of CDs in nanoparticles for controlled drug release is also of high interest [22–24].

The special structures of CDs have been investigated widely, mainly by X-ray crystallography [25–30] and by neutron diffraction [31–34]. In most of these experimentally determined structures water molecules are included, as a consequence of strong hydrogen bonds, which are formed between the hydroxyl groups of the CDs rims and the solvent. A few experimental data are available about the structure of CDs obtained from more or less anhydrous conditions [35–38].

Crystallographic investigations reveal that the structures of CDs seem to be rather rigid, which can be also assumed from the fact that relatively rigid glucose units are joined in a cyclic arrangement. Nevertheless, some flexibility has been postulated for CDs as well as for their complexes [39, 40].

The interior of CDs is postulated to be hydrophobic to some extent. The inclusion of some compounds with a proton transfer equilibrium by  $\beta$ -CD shifts the equilibrium in the same direction as the

addition of dioxane to 24% (v/v) [41]. The possibility to form hydrogen bonds is another important driving force for complexation of CDs. Hydrogen bonding determines the geometries of CDs and their physico-chemical properties. Intramolecular hydrogen bonds make the molecules more rigid, and in solution there is a competition between inter- and intramolecular hydrogen bond networks, which is reflected in extraordinary physico-chemical properties. For example, the solubilities of the parent compounds in water are influenced so far that  $\beta$ -CD is less soluble than  $\alpha$ - and  $\gamma$ -CD [42–44], and surprisingly methyl-substituted derivatives. Negative solvation enthalpies result from the competition of intramolecular hydrogen bonds with hydrogen bonds to the solvent [45, 46].

Also the thermodynamical parameters of the complexation reaction depend on the type of CD. An example of significant differences is the inclusion reaction of triflumizole to  $\beta$ -CD and dimethyl- $\beta$ -CD. Although the complexation constants are similar for both cases, completely different reaction enthalpies are observed [47]. Enthalpy-entropy compensation occurs for  $\beta$ -CD, whereas the reaction is mainly entropy-controlled for dimethyl- $\beta$ -CD.

Prediction models for the free energy of complexation of CDs reveal that for  $\beta$ -CD mainly hydrophobic interactions contribute to the driving forces for the complexation; hydrogen bond dependent descriptors are suggested to play a minor role. For the other CDs ( $\alpha$ - and  $\gamma$ -CD) hydrogen bond donor as well as acceptor properties determine significantly the prediction correlation [48].

As already mentioned, hydrogen bonding in CDs has been investigated intensively by various experimental and theoretical methods. Most of the experimentally available structures of CDs contain solvent molecules, mainly water; they are therefore unsymmetric and show also intermolecular hydrogen bonds. Not many examples are known for structures of CDs under anhydrous conditions. Comparatively few molecular calculations have been performed based on semiempirical [49–51] and more accurate ab initio and density functional theory (DFT) methods [52–54]. In this paper, we describe in continuation of our studies on intramolecular hydrogen bonds in CDs [55, 56] a systematic investigation on the structures of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD by a DFT method, *B3LYP*, using 6-31G(d,p) basis set, in order to get more detailed information about the conformational minima and the energy differences between them.

Density functional calculations on cyclodextrins

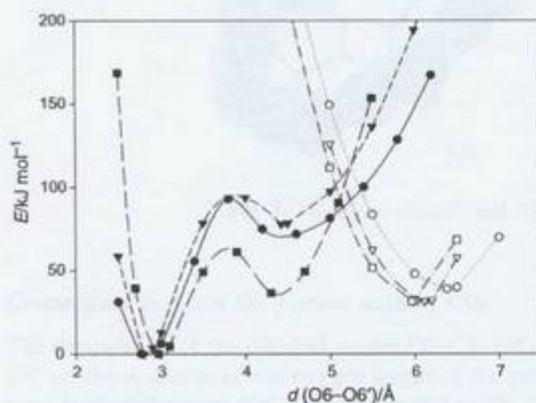
## Results and discussion

### Homodromic hydrogen bonds

Under symmetric conditions several conformational minima can be found for  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD by varying the distances of the oxygen atoms of the primary hydroxyl groups. The resulting energy profiles are given in Fig. 1.

The global minima, the lowest energy conformations (A) are detected at rather close O6–O6' distances. Two homodromic hydrogen bond rings are formed in these conformations, one very short at the primary hydroxyl groups and another intramolecular hydrogen bond ring include the secondary hydroxyl groups of CDs. This "one-gate-closed" conformation (A) shows a basket-like shape. The geometry of these conformations of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs is shown in Fig. 2, where both hydrogen bond rings are given as *van der Waals* balls in gray (primary hydroxyl groups) and black (secondary hydroxyl groups).

The second important conformational minima occur at larger O6–O6' distances and describe open conformations, which correspond to some extent to structures, which are found experimentally. These open conformations exhibit significantly higher energies than the low-energy conformations A, as a consequence of the missing hydrogen bonds at the primary hydroxyl groups. In these conformations on-



**Fig. 1.** Energy profiles of the various conformations of  $\alpha$  (circle)-,  $\beta$  (triangle)-, and  $\gamma$  (rectangle)-CD by varying the distances of the oxygen atoms of the primary hydroxyl groups. Filled and unfilled symbols represent structures according to A and C, respectively

ly one hydrogen bond rim built from the secondary hydroxyl groups exists. The conformational minima B (Fig. 2) are rather flat and will not be considered in detail here.

Four orientations of the homodromic hydrogen bond rings are possible, which differ in energy: both hydrogen bond rims orientated counterclockwise (cccc) or clockwise (cwcw) and the primary rim clockwise and the secondary rim counterclockwise (cwcc) and *vice versa* (cccw). For the orientations cccw and cccc the conformational minima B do not exist.

In Table 1, the distances of the heavy atoms of the hydrogen bridges are given for the minima A and C and also possible orientations of the homodromic hydrogen bonds together with some characteristic parameters describing the shape of the CDs. Moreover, the relative energies calculated by DFT/6-31G(d,p) are included in the table. The energy of the lower energy conformation is set to zero. For all CDs, the conformations A exhibit short distances of the oxygen atoms of the primary hydroxyl groups indicating short and strong hydrogen bonds. A slight increase only of these distances can be observed going from  $\alpha$ - to  $\gamma$ -CD. No significant differences are found for the various orientations of the hydrogen bonds. In the open conformations C no hydrogen bonds are observable at the primary hydroxyl groups, the distances O6–O6' depend on the orientation of the hydroxyl groups at these positions. cw orientation is connected with slightly smaller distances throughout. The hydrogen bonds at the second rim, built by the secondary hydroxyl groups, do not show any remarkable changes for all conformations and CDs. Somewhat smaller distances can be found for O2–O3' between conformations A and C.

### Comparison with experimental parameters

The comparison to related structural parameters obtained experimentally is given in Table 2a and b.

Three randomly selected crystallographic geometries were considered [34, 57, 58]. As water molecules are included in the crystallographic data and intermolecular hydrogen bonds are formed, the resulting structures are unsymmetric with quite large deviations for the oxygen-oxygen distances. Nevertheless, the mean values of these distances are in relatively good agreement with the calculated structural data of the conformations C (open conformations).

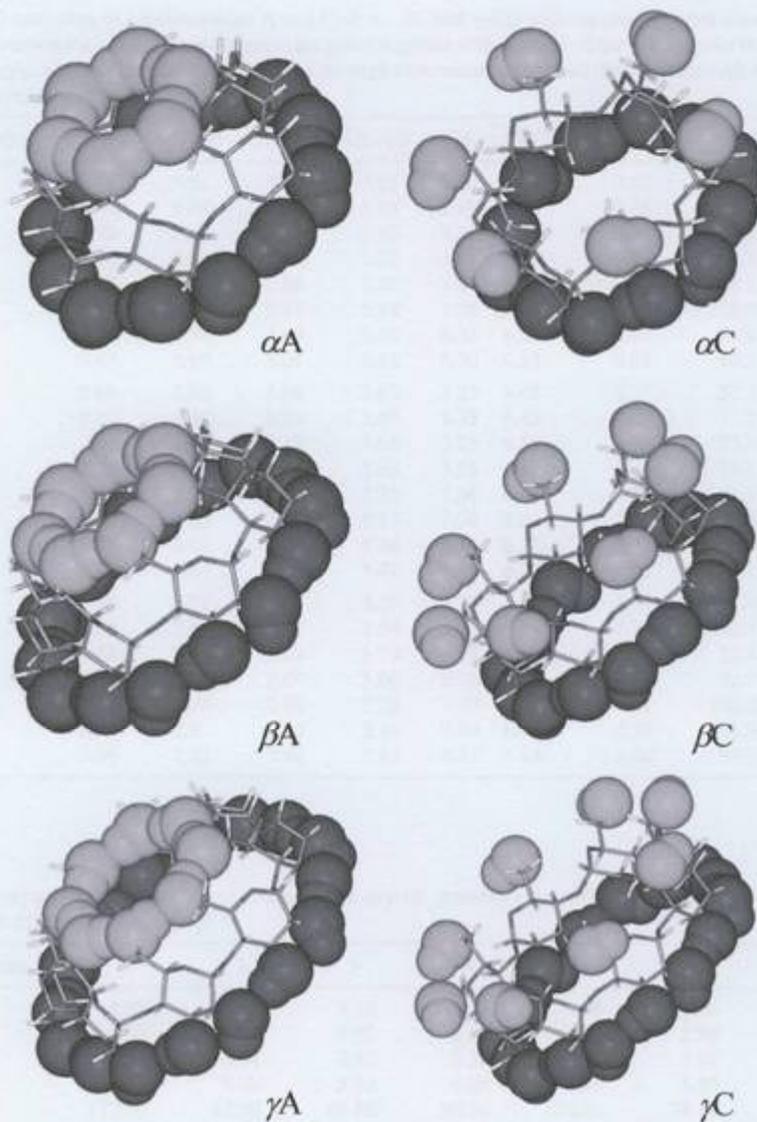


Fig. 2. "One-gate-closed" and "open" conformations of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD

#### Connections between the glucose units of CDs

The dependence of the dihedral angles O5–C1–O4'–C4' on the distances of the oxygen atoms of the primary hydroxyl groups (O6–O6') is depicted in Fig. 3.

Moreover, the dihedral angles describing the connections between the glucose units of CDs (O5–C1–O4'–C4', O6–C6–C5–C4) are given in Table 3a, b and c. There are evidently differences of these angles

between conformations A and C. The torsional angles O5–C1–O4'–C4' are somewhat larger in conformations C, the torsional angles C1–O4'–C4'–C3' are here slightly decreased. There is almost no significant influence of the orientation of the homodromic hydrogen bonds on both dihedral angles in conformations A, some influence of the orientation at the rim built from the secondary hydroxyl groups on O5–C1–O4'–C4' can be observed.

**Table 1.** Geometry parameters of conformations A and C of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD with various orientations of the hydrogen bond rings. The hydrogen containing heavy atoms distances are given together with the radii of the primary and the secondary hydroxyl groups.  $\tau$  is the tilt angle,  $\sigma$  is the angle between the plane through equivalent oxygen and the plane through the heavy atoms of the glucose unit (distances/Å, angles/°)

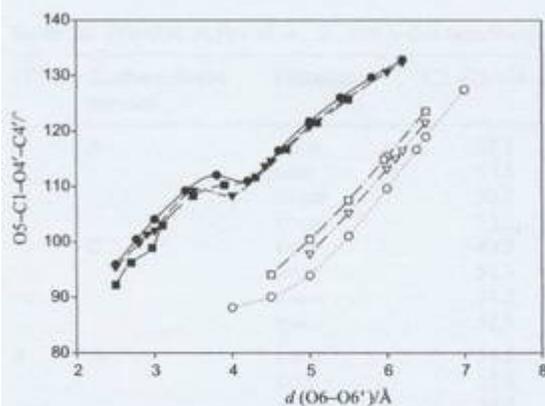
CD	E <sub>min</sub>	Orientation	O6–O6'	O2–O3	O2–O3'	O3–O3'	rO6	rO3	$\Delta$ (rO6–rO3)	$\tau$	$\sigma$	E/kJ mol <sup>-1</sup>
$\alpha$	A	cccw	2.77	2.86	3.12	5.61	3.19	6.47	0.87	58.5	73.8	0.6
		cccc	2.76	2.89	3.15	5.59	3.19	6.45	0.86	58.4	76.8	0.1
		ewew	2.78	2.86	3.16	5.65	3.21	6.52	0.87	58.0	72.8	0.1
		ewcc	2.78	2.89	3.18	5.62	3.21	6.49	0.87	58.7	72.5	0.0
	C	cccw	6.38	2.78	2.86	5.22	7.37	6.03	0.81	105.1	85.3	48.4
		cccc	6.37	2.83	2.87	5.14	7.35	5.94	0.80	106.0	85.0	51.0
		ewcw	5.50	2.84	2.99	5.45	6.35	6.29	0.84	90.7	78.5	106.0
$\beta$	A	cccw	2.80	2.86	3.08	5.65	3.23	6.45	3.23	57.3	70.7	0.0
		cccc	2.80	2.89	3.09	5.60	3.23	6.45	3.23	57.2	70.4	0.0
		ewcw	2.82	2.86	3.11	5.68	3.25	6.55	3.30	57.3	69.7	8.8
		ewcc	2.82	2.89	3.12	5.63	3.25	6.49	3.24	57.7	69.7	9.4
	C	cccw	6.11	2.77	2.84	5.25	7.04	6.05	-0.99	101.1	83.5	43.8
		cccc	6.11	2.82	2.85	5.17	7.04	5.96	-1.08	105.8	83.5	44.1
		ewcw	5.33	2.83	2.96	5.48	6.14	6.32	0.17	88.0	76.2	88.7
$\gamma$	A	cccw	2.85	2.87	3.06	5.70	3.73	7.45	1.75	51.4	66.6	0.0
		cccc	2.85	2.9	3.07	5.64	3.73	7.37	1.73	51.9	66.7	0.7
		ewcw	2.88	2.87	3.08	5.72	3.76	7.47	1.75	52.4	66.2	18.4
		ewcc	2.88	2.89	3.09	5.66	3.76	7.40	1.74	52.9	66.4	20.0
	C	cccw	5.96	2.76	2.82	5.26	7.79	6.87	1.61	100.2	82.7	31.3
		cccc	5.97	2.8	2.83	5.16	7.80	6.74	1.58	101.9	82.6	28.4
		ewcw	5.26	2.82	2.91	5.47	6.87	7.14	1.68	87.0	75.2	64.8

**Table 2a.** Oxygen-oxygen distances of three examples of crystal structures for  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD. The related calculated parameters are given (distances/Å, angles/°)

CD	Struct. param.	1	2	3	4	5	6	7	8
$\alpha$	O6–O6'	4.67	6.01	5.35	5.74	4.96	6.06	–	–
	O2–O3	2.94	2.88	2.92	2.88	2.91	2.86	–	–
	O2–O3'	2.95	3.01	2.82	3.37	4.21	3.02	–	–
	O3–O3'	5.50	5.60	5.20	6.02	6.16	5.59	–	–
	$\tau$	122.00	92.90	89.80	89.50	125.10	74.10	–	–
	$\sigma$	112.12	98.59	99.14	104.53	132.52	99.35	–	–
$\beta$	O6–O6'	3.82	5.20	6.26	4.67	5.80	5.85	6.09	–
	O2–O3	2.92	2.91	2.90	2.87	2.86	2.92	2.89	–
	O2–O3'	2.96	2.96	2.86	2.77	2.78	2.90	2.88	–
	O3–O3'	5.42	5.30	5.58	5.62	5.69	5.75	5.70	–
	$\tau$	113.97	109.04	77.41	114.88	92.26	75.63	78.14	–
	$\sigma$	111.00	106.20	85.00	102.00	103.00	97.50	95.00	–
	$\sigma$ lit	112.30	106.50	85.30	103.60	103.50	92.20	94.90	–
$\gamma$	O6–O6'	5.11	5.55	5.89	6.01	4.06	5.31	6.09	3.93
	O2–O3	2.92	2.91	2.85	2.92	2.84	2.89	2.90	2.86
	O2–O3'	2.91	2.78	2.77	2.83	2.81	2.84	2.77	2.88
	O3–O3'	5.50	5.57	5.04	5.58	5.31	5.47	5.45	5.47
	$\tau$	116.50	74.90	91.00	86.10	115.50	90.80	80.20	113.90
	$\sigma$	108.31	90.69	111.62	95.56	107.56	103.94	101.13	112.53

**Table 2b.** Mean values of three examples of crystal structures for  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD of oxygen-oxygen distances have been determined and compared to the values of the symmetric conformations (distances/Å, angles/°)

CD	Structural parameters	Mean value	sym A (cccw)	sym C (cccw)
$\alpha$	O6-O6'	5.47	2.77	6.38
	O2-O3	2.90	2.86	2.78
	O2-O3'	3.23	2.86	2.86
	O3-O3'	5.68	5.61	5.22
	$\tau$	98.90	121.50	74.94
	$\sigma$	107.72	106.19	94.75
$\beta$	O6-O6'	5.38	2.80	6.11
	O2-O3	2.90	2.86	2.77
	O2-O3'	2.87	3.08	2.84
	O3-O3'	5.58	5.60	5.25
	$\tau$	86.98	122.81	78.83
	$\sigma$	99.67	109.28	96.53
$\gamma$	$\sigma$ lit	99.76		
	O6-O6'	5.24	2.85	5.96
	O2-O3	2.89	2.87	2.76
	O2-O3'	2.82	3.06	2.82
	O3-O3'	5.42	3.73	5.26
	$\tau$	96.11	128.57	79.79
$\sigma$	103.92	113.45	97.30	



**Fig. 3.** Dependence of the dihedral angle O5-C1-O4'-C4' on the O6-O6' distance for various conformations of  $\alpha$  (circle)-,  $\beta$  (triangle)-, and  $\gamma$  (rectangle)-CD. Filled and unfilled symbols represent minima A and C

A graphical presentation of the dihedral angles of the bonds connecting the glucose units is given in Fig. 4, together with averaged values from some selected crystal structures.

All dihedral angles can be found in a rather limited range, as larger deviations lead to an increase of

**Table 3a.** Dihedral angles of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD describing the connection between the glucose units (angles/°)

CD	Conformational minima	orientation	O5-C1-O4'-C4'	C1-O4'-C4'-C3'
$\alpha$	A	cccw	100.4	132.7
		cccc	99.8	135.6
		cwcc	99.0	133.8
	C	cwcc	99.0	136.4
		cccw	116.7	123.1
		cccc	115.7	126.0
$\beta$	A	cwcc	107.2	127.5
		cccw	105.8	130.6
		cccc	99.8	131.1
	C	cccc	100	133.5
		cwcc	98.9	132.0
		cwcc	99.3	134.3
$\gamma$	A	cccw	115.1	124.1
		cccc	114.5	126.7
		cwcc	106.7	127.0
	C	cwcc	106.5	129.3
		cccw	98.9	130.1
		cccc	99.3	132.4
$\gamma$	A	cwcc	98.5	130.7
		cccw	99.0	133.0
		cccc	99.0	133.0
	C	cccw	114.9	124.1
		cccc	114.7	126.7
		cwcc	107.6	125.6
		cwcc	107.2	128.1

energy caused by steric repulsion. The “one-gate-closed” conformations occur somewhat separated at lower values of O5-C1-O4'-C4'. The values for the “open” conformations vary according to the orientation of the hydrogen bonds, and they are close to the averaged experimental values throughout.

#### Comparison of the glycosidic angles

Due to the structure of the cyclic systems (six-membered glucose units) no pronounced changes of the dihedral angles inside the rings are possible as no inversion of the rings can be performed within the frame of the macrocyclic systems. In Table 3a, the dihedral angles describing the conformations of the glucose units are given.

Surprisingly, there are some significant differences between the values for conformations A and C, particularly for the torsional angles C3-C4-C5-O5 and C4-C5-O5-C1. Moreover, some limited distortions of the glucose units can be found for the different CDs. Somewhat smaller changes of these

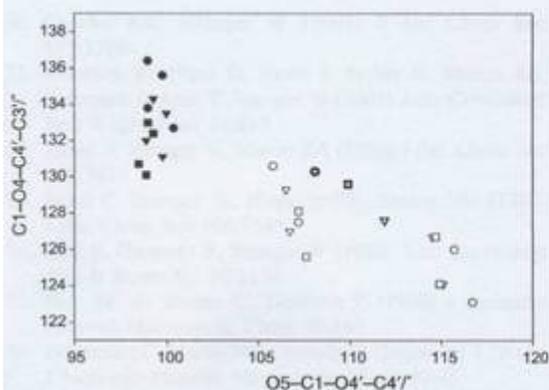
## Density functional calculations on cyclodextrins

**Table 3b.** Dihedral angles of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD describing the orientation of the primary hydroxyl groups substituent (angles/°)

CD	Conformational minima	Orientation	H6-O6-C6-C5	O6-C6-C5-C4	H3-O3-C3-C2	H2-O2-C2-C1
$\alpha$	A	cccw	-57.0	-127.4	176.9	-167.7
		cccc	-56.7	-128.6	-48.0	-35.9
		cwcw	107.9	-142.6	177.0	-168.0
		ewcc	108.4	-143.2	-47.8	-35.5
	C	cccw	56.4	61.7	174.7	-165.6
		cccc	56.7	61.4	-44.7	-39.1
		cwcw	-93.0	23.8	176.3	-167.6
		ewcc	-95.9	26.1	-47.1	-36.7
$\beta$	A	cccw	-61.9	-122.5	174.0	-169.3
		cccc	-61.7	-123	-48.1	-34.4
		ewew	111.6	-139.2	174.3	-169.6
		ewcc	112.3	-139.6	-48.1	-34.0
	C	cccw	56.6	61.2	170.6	-166.5
		cccc	56.9	60.9	-44.0	-38.3
		cwcw	-95.6	31.0	172.9	-169.0
		ewcc	-97.6	32.0	-46.5	-35.8
$\gamma$	A	cccw	-65.2	-119.2	171.8	-170.7
		cccc	-65.0	-119.4	-48.7	-33.0
		ewew	114.8	-136.6	172.2	-171.1
		ewcc	115.6	-136.9	-48.5	-32.6
	C	cccw	56.7	60.8	167.4	-166.9
		cccc	56.8	60.5	-42.9	-37.9
		cwcw	-96.1	34.0	170.3	-170.0
		ewcc	-98.7	35.6	-46.1	-35.0

**Table 3c.** Dihedral angles of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD describing the conformation of the glucose subunits (angles/°)

CD	Conformational minima	Orientation	C2-C3-C4-C5	C3-C4-C5-O5	C4-C5-O5-C1	C5-O5-C1-C2
$\alpha$	A	cccw	52.1	-55.6	63.4	-64.0
		cccc	53.1	-55.6	62.2	-62.3
		cwcw	52.3	-55.5	63.2	-63.6
		ewcc	53.1	-55.6	62.2	-62.1
	C	cccw	49.2	-48.7	56.8	-61.7
		cccc	51.3	-48.9	55.3	-60.6
		cwcw	51.2	-54.3	62.0	-64.0
		ewcc	52.8	-54.8	61.1	-62.9
$\beta$	A	cccw	54.5	-59.6	65.5	-62.8
		cccc	55.9	-59.8	64.4	-61.4
		cwcw	54.7	-59.4	65.2	-62.2
		ewcc	56.0	-59.6	64.1	-60.8
	C	cccw	52.4	-52.0	57.2	-59.2
		cccc	54.6	-52.0	55.5	-58.1
		cwcw	53.8	-57.5	63.0	-62.0
		ewcc	55.6	-57.8	61.8	-61.0
$\gamma$	A	cccw	56.3	-62.6	67.0	-61.7
		cccc	57.8	-62.9	65.8	-60.4
		cwcw	56.3	-62.2	66.4	-60.9
		ewcc	57.9	-62.5	65.3	-59.7
	C	cccw	54.7	-53.4	57.0	-57.0
		cccc	57.0	-53.8	55.0	-55.9
		cwcw	55.5	-59.5	63.2	-60.4
		ewcc	57.5	-59.7	61.9	-59.4



**Fig. 4.** Plot of the dihedral angle  $O5-C1-O4'-C4'$  against  $C1-O4'-C4'-C3'$  for  $\alpha$  (circle)-,  $\beta$  (triangle)-, and  $\gamma$  (rectangle)-CD. Filled and unfilled symbols represent minima A and C. Specially marked symbols for the averaged values of the experimental structures

dihedral angles occur for the different orientations for the hydrogen bonds.

Summarizing, *B3LYP/6-31G(d,p)* calculations have been performed on natural  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CD in the gas phase applying  $C_n$  symmetry. Generally, two pronounced energy minimum conformations were detected by varying the oxygen-oxygen distances of the primary hydroxyl groups ( $O6-O6'$ ). The lowest energy conformations correspond to “one-gate-closed” structures, characterized by two intramolecular hydrogen bond rings. The hydrogen bond lengths are very close to that found for open chains of water and alcohols. There is no experimental evidence for such conformations, in contrary to the second minima (“open” conformations) of higher energy, which show similarities to geometries of experimentally obtained structures. According to the orientation of the homodromic hydrogen bonds, four different conformations are found for each geometry (see also Table 3b) with different energies as a consequence of the hydrogen bonds and the chirality of the glucose units.

#### Methods of calculation

Systematic *B3LYP* calculations with 6-31G(d,p) basis set on anhydrous  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs have been applied. Imposing  $C_n$  symmetry throughout, the oxygen-oxygen distances were scanned and full geometry optimizations of all remaining geometry parameters were then performed, using the program package Gaussian03 [59]. By subsequent frequency calcula-

tions, it was verified that the calculated geometries are indeed minima on the potential energy surface.

#### Acknowledgements

Technical assistance of Ms. M. Ziehrgraser, Ms. A. and Ms. V. Stummer is gratefully acknowledged.

#### References

- Szejtli J (1998) *Chem Rev* 98:1743
- Szejtli J (2004) *Pure Appl Chem* 76:1825
- Dodziuk H (2006) *Cyclodextrins and Their Complexes*. Wiley-Vch, Weinheim
- Del Valle EMM (2004) *Process Biochem* 39:1033
- Challa R, Ahuja A, Ali J, Khar RK (2005) *AAPS Pharm Sci Tech* 6: E 329; [www.aapspharmsci.org](http://www.aapspharmsci.org)
- Loftsson T, Masson M (2001) *Int J Pharm* 225:15
- Uekama K (2004) *Chem Pharm Bull* 52:900
- Loftsson T, Brewster ME (1996) *J Pharm Sci* 85(10):1017
- Szente L, Szejtli J (1999) *Adv Drug Deliver Rev* 36:17
- Hirayama F, Uekama K (1999) *Adv Drug Deliver Rev* 36:125
- Devis ME, Brewster ME (2004) *Nat Rev Drug Discovery* 3:1023
- Uekama K, Hirayama F, Arima H (2006) *J Inclusion Phenom Macrocyclic Chem* 56:3
- Singh M, Sharma R, Banerjee UC (2002) *Biotech Advances* 20:341
- Loftsson T, Duchene D (2007) *Int J Pharm* 329:1
- Endo T, Zheng M, Zimmermann W (2002) *Aust J Chem* 55:39
- Zheng M, Endo T, Zimmermann W (2002) *J Inclusion Phenom Macrocyclic Chem* 44:387
- Larsen KL (2002) *J Inclusion Phenom Macrocyclic Chem* 43:1
- Taira H, Nagase H, Endo T, Ueda H (2006) *J Inclusion Phenom Macrocyclic Chem* 56:23
- Khan AR, Forgo P, Stine KJ, D'Souza VT (1998) *Chem Rev* 98:1977
- Pitha J, Trinadha RC (1990) *Carbohydr Res* 200:429
- Engeldinger E, Armspach D, Matt D (2003) *Chem Rev* 103:4147
- Duchene D, Ponchel G, Wouessidjewe D (1999) *Adv Drug Deliver Rev* 36:29
- Duchene D, Wouessidjewe D, Ponchel G (1999) *J Controlled Release* 62:263
- Patriot N, Edwards-Levy F, Andry MC, Levy MC (2000) *Int J Pharm* 211:19
- Saenger W, Jakob J, Gessler K, Steiner T, Hoffmann D, Sambe H, Koizumi K, Smith SM, Takaha T (1998) *Chem Rev* 98:1787
- Steiner T, Mason SA, Saenger W (1991) *J Am Chem Soc* 113:5676
- Steiner T, Koellner G (1994) *J Am Chem Soc* 116:5122
- Steiner T, Saenger W (1994) *Carbohydr Res* 259:1
- Saenger W, Steiner T (1998) *Acta Crystallogr Sect A Found Crystallogr* 54:798

## density functional calculations on cyclodextrins

- Chacko KK, Saenger W (1981) *J Am Chem Soc* 103:1708
- Imamura K, Nimz O, Jacob J, Myles D, Mason SA, Kitamura S, Aree T, Saenger W (2001) *Acta Crystallogr Sect B Struct Sci* 57:833
- Zabel V, Saenger W, Mason SA (1986) *J Am Chem Soc* 108:3664
- Betzl C, Saenger W, Hingerty BE, Brown GM (1984) *J Am Chem Soc* 106:7545
- Klar B, Hingerty B, Saenger W (1980) *Acta Crystallogr Sect B Struct Sci* 36:1154
- Diot M, de Brauer C, Germain P (1998) *J Inclusion Phenom Macrocyclic Chem* 30:143
- De Brauer C, Merlin MP, Germain P, Guerlandel T (2000) *J Inclusion Phenom Macrocyclic Chem* 37:75
- Steiner T, Saenger W (1996) *Carbohydr Res* 282:53
- Steiner T, Saenger W (1996) *Carbohydr Res* 275:73
- Köhler G, Martinek H, Parasuk W, Rechthaler K, Wolschann P (1995) *Monatsh Chem* 126:299
- Naidoo KJ, Yu-Jen CJ, Jansson JLM, Widmalm G, Maliniak A (2004) *J Phys Chem B* 108:4236
- Colman AW, Nicolis I, Keller N, Dalbiez JP (1992) *J Inclusion Phenom Macrocyclic Chem* 13:139
- Buvari-Barcza A, Barcza L (2000) *J Inclusion Phenom Macrocyclic Chem* 36:355
- Plazanet M, Floare C, Johnson MR, Schweins R, Trommsdorff HP (2004) *J Chem Phys* 121:5031
- Plazanet M, Johnson MR, Schweins R, Trommsdorff HP (2006) *Chem Phys* 331:35
- Kozar T, Venanzi CA (1997) *J Mol Struct* 395-396:451
- Doziuk H (2002) *J Mol Struct* 614:33
- Viernstein H, Weiss-Greiler P, Wolschann P (2002) *J Inclusion Phenom Macrocyclic Chem* 44:235
48. Klein CT, Polheim D, Viernstein H, Wolschann P (2000) *J Inclusion Phenom Macrocyclic Chem* 36:409
49. Li XS, Liu L, Mu TW, Guo QX (2000) *Monatsh Chem* 131:849
50. Liu L, Guo QX (2004) *J Inclusion Phenom Macrocyclic Chem* 50:95
51. Lipkowitz KB (1998) *Chem Rev* 98:1829
52. Avakyan VG, Nazarov VB, Voronezhcheva NI (2005) *Russ J Phys Chem* 97:18
53. Avakyan VG, Nazarov VB, Alfimov MV, Bagaturyants AA, Voronezhcheva NI (2001) *Russ Chem Bull* 50:206
54. Pinjari RV, Joshi KA, Gejji SP (2006) *J Phys Chem A* 110:13073
55. Karpfen A, Liedl E, Weiss-Greiler P, Snor W, Viernstein H, Wolschann P (2007) *J Inclusion Phenom Macrocyclic Chem* 57:35
56. Snor W, Liedl E, Weiss-Greiler P, Karpfen A, Viernstein H, Wolschann P (2007) *Chem Phys Lett* 441:159
57. Lindner K, Saenger W (1982) *Carbohydr Res* 99:103
58. Harata K (1987) *Bull Chem Soc Jpn* 60:2763
59. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski JA, Montgomery JJA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (2004) GAUSS- IAN 03, Revision C.02, Gaussian Inc., Wallingford CT

#### **11.4 Paper IV: “Molecular Dynamics Simulations and Quantum Chemical Calculations on $\beta$ -CD-Spirolactone Complex”**

The purpose of this study was to investigate the complexation behaviour of  $\beta$ -CD with respect to spironolactone. For this reason MD simulations on  $\beta$ -CD in vacuo, with water and complexated with spironolactone were performed and two different orientations of SP in the  $\beta$ -CD cavity were considered. Moreover, DFT calculations were used in order to get more reliable structures and interaction energies of the  $\beta$ -CD-SP complexes.

## ORIGINAL ARTICLE

## Molecular dynamics simulations and quantum chemical calculations on $\beta$ -cyclodextrin–spironolactone complex

Philipp Weinzinger · Petra Weiss-Greiler ·  
Walter Snor · Helmut Viernstein · Peter Wolschann

Received: 15 May 2006 / Accepted: 20 October 2006 / Published online: 24 January 2007  
© Springer Science+Business Media B.V. 2007

**Abstract** Molecular dynamics simulations on  $\beta$ -cyclodextrin in vacuo, with water and complexed with spironolactone (SP) were performed at a temperature of 300 K over a period of 1 ns. Two different orientations of SP in the cavity were considered. Along with conformational parameters, the formation of hydrogen bonds has been monitored during the whole simulation time. Cyclodextrins have the capability to form hydrogen bonds with the surrounding water molecules or intramolecular ones. The incorporation of ligands into the hydrophobic interior of  $\beta$ -cyclodextrin changes the preference of hydrogen bonds significantly and results in a contribution to the decrease of flexibility. Quantum chemical calculations on SP– $\beta$ -CD inclusion complex were performed to determine the interaction energy and to prove the applicability of various methods. Although all applied methods describe reasonable geometries for the association complex, higher level methods (e.g., B3LYP/6-31G(d,p)) seem to be necessary to determine reliable interaction energies.

**Keywords**  $\beta$ -Cyclodextrin · Molecular dynamics · DFT calculations · Inclusion complex · Hydrogen bonding · Spironolactone · Solubility enhancement

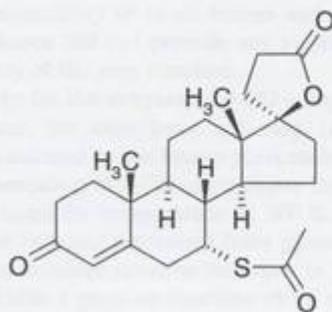
### Introduction

The affinity of host molecules to the cavity of cyclodextrins (CDs) depends on the agreement of the molecular surface of the guest and the complementary surface of the host. Moreover, the complementarity of electrostatic potentials, hydrogen bond donor and acceptor properties are features which influence the stability of the inclusion complexes significantly. Therefore, investigations on the geometry of the inclusion complexes are very helpful to get some insight into the detailed reaction mechanism, the driving forces of complexation and the related geometrical conditions. Molecular calculations are useful methods, which contribute to answer these questions. Empirical, ab initio and DFT methods can be used to obtain energetically minimized structures. Molecular dynamics (MD) and Monte-Carlo (MC) simulations allow to explore the energy hypersurfaces and to find local and global conformational minima distinct from a certain starting geometry. Particularly, MD simulations are widely used to explore the structural features of CDs and cyclodextrin complexes, as exemplified by some recent publications [1–8].

In continuation of our research concerning the reaction mechanisms of CD inclusion complexation [1, 9, 10] we have performed molecular calculations on spironolactone (SP). This compound (Scheme 1) was selected because of its exceptionally high affinity to  $\beta$ -CD, which results in an extensive enhancement of its solubility. SP is a partial synthetic steroid-analogue of aldosterone, which works as competitive aldosterone-antagonist. The solubility of SP in water is rather low (2.8 mg/100 mL at 25°). For that reason the solubility

P. Weinzinger · P. Weiss-Greiler · W. Snor ·  
P. Wolschann  
Institute of Theoretical Chemistry, University of Vienna,  
Währinger Straße 17, 1090 Vienna, Austria

H. Viernstein (✉)  
Department of Pharmaceutical Technology and  
Biopharmaceutics, University of Vienna, Althanstraße 14,  
1090 Vienna, Austria  
e-mail: Helmut.Viernstein@univie.ac.at



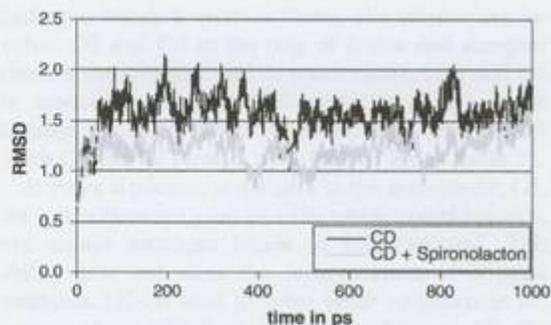
Scheme 1

enhancement by complexation is of high importance for its practical application.

#### Methods of calculations

Starting geometries of  $\beta$ -CD and SP were obtained from the Cambridge database. The ligand was placed in two orientations (in complex A the lactone ring of SP is located close to the more narrow rim of  $\beta$ -CD, in complex B SP is turned 180° along the  $\beta$ -CD-entry axis) and energy-minimized with Gaussian03 [11] using the HF/6-31G(d) basis set. The AMBER 7 program package [12] was used to generate the MD trajectories using the general amber force field (gaff) and a step size of 1fs. The gaff is a practical force field for organic molecules such as the ligands and was used uniformly for all molecules, including water and  $\beta$ -CD. AMBER tools were used in part for the subsequent analysis. The simulations in aqueous environment were performed with approximately 2000 water molecules of the TIP3P type in a periodic cell. The system was heated to a temperature of 300 K and afterwards held at constant pressure and temperature using a weak coupling algorithm [13]. A 12 Å cutoff was used for the Lennard-Jones interactions. In order to evaluate the quality of the simulations the root mean square deviation (RMSD) of the atom positions have been estimated for the CDs (Fig. 1) in reference to the starting structure.

The trajectory was recorded at 0.5 ps intervals thus resulting in 2000 frames for the 1ns simulation. In order to estimate the hydrogen-bonding properties of the  $\beta$ -CD molecule all frames were searched for hydrogen bonds with the following characteristics: a maximum heavy atom distance of 3.2 Å and a maximum angle deviation of 45° from linearity. Although the simulations could be considered equilibrated after 50 ps, data for further analysis were collected after 100 ps.



**Fig. 1** RMSD values of cyclodextrin atom positions. In the simulation without a ligand an average of 1.57 was estimated and 1.22 in the simulation with spironolactone as ligand. Both set-ups included a periodic water box

To calculate the binding energy different energy contributions have been determined (e.g., the change in entropy for the binding process). These calculations were enabled by the nmode module of AMBER. Since this is a time consuming procedure only 100 snapshots at regular intervals were picked up for this calculation.

The interaction energy between  $\beta$ -CD and SP was also calculated by quantum chemical methods using Gaussian03 [11]. A semiempirical (PM3) and two ab initio methods (Hartree Fock with basis sets 3-21G and 6-31G(d,p)) and a DFT method (B3LYP with the basis set 6-31G(d,p)) were used for the full geometry optimization of  $\beta$ -CD, SP and the complexes. The starting geometries for the complexes were taken from MD calculations using the two different orientations of SP. The binding energies were calculated from the energy differences between both isolated compounds and the complexes.

#### Results and discussion

One of the most obvious observations during the MD simulations was the bend in the C1 backbone of the  $\beta$ -CD molecule. In the minimized structure  $\beta$ -CD rim is not completely planar, but shows a slight bent to the side of the O2/O3 groups. The RMS average of all seven C1-C1-C1-C1 torsion angles is 14°. During the simulation and the effects of kinetic energy this angle increases as the  $\beta$ -CD is allowed to occupy conformations farther away from the minimum. The root-mean-square of the angle was chosen in order to get a general parameter for the flexion of the  $\beta$ -CD, which is not modified by the algebraic sign of each individual torsion angle. In the summation of all seven torsion angles, without RMS, the torsions cancelled each other

out to approximately  $0^\circ$  in all frames and all simulations and hence did not provide any information for the flexibility of the ring structure.

Obviously, for the unliganded  $\beta$ -CD a larger bend is allowed than for complexes suffering from steric restrictions induced by the bound guest molecules. The in vacuo simulation of  $\beta$ -CD alone shows the effects of motion induced by temperature at 300 K. Although temperature inflicted deviations from planarity occur, the average structure remains the same as after minimization. Table 1 gives an overview of all simulations and the RMS torsion averaged over the full trajectory. The investigated simulations include  $\beta$ -CD without a ligand in water and in vacuo,  $\beta$ -CD complexed with SP in two different orientations in water and one of them in vacuo.

In the presence of water  $\beta$ -CD undergoes conformational changes. After the initial stages of the simulation two opposing glucose units turn their O2/O3 side inwards, resulting in higher torsional deviations from planarity. The flexions of  $\beta$ -CD are a result of an effort to minimize the volume of the hydrophobic interior and the formation of a hydrogen bond network, which includes more or less seven water molecules. All O2/O3 of both bent glucose units are involved in this network.

Steric hindrances and a modified hydrogen bond network provided by a large ligand like SP let the  $\beta$ -CD assume a vacuum-like conformation. After complexation the water network in the interior is removed and  $\beta$ -CD can adopt its planar vacuum structure again. The hydrogen bonds are not simply broken, but shifted from the inside to the area on the rim of the ring and the exterior (esp. O6). So no energetic penalties occur after ligand binding. This supports the claim that desolvation is the driving force of complexation.  $\beta$ -CDs hydroxyl groups are encouraged to form an intramolecular hydrogen-bonding network, which involves the whole O2/O3 side of the ring.

Since each glucose unit contains one hydroxyl group (O6-H) on one side and two (O2-H and O3-H) on the other side of the donut shaped structure of  $\beta$ -CD the

**Table 1** Average RMS C1-C1-C1-C1 backbone torsion of neighboring glucose units of  $\beta$ -CD

	CD in water	Complex A in water	Complex B in water	CD in vacuo	SP in vacuo
Average RMS torsion	62.02	39.90	35.48	39.19	38.01

The standard deviation for these values is 6.90

hydrogen bonds formation during the simulations involves O2 and O3 in the role of donor and acceptor (both intramolecular and to water molecules) and O6 in association mostly with water, although some hydrogen bonds to O6 of the neighboring glucose unit or an O5 can be formed from time to time (Table 2).

If water is present in the core of the macrocycle, O1, the linker between glucose units which points inwards, can accept hydrogen bonds to some amount. This ability gets lost after the incorporation of a guest molecule. O2-Hs tend to favor other cyclodextrin acceptors above water especially after ligation. On the other hand O3s prefer hydrogen bonds to water molecules, a preposition which is reduced by complexation. The acceptor properties of O2/O3 act vice versa to their donor activities. In this case O3 prefers more hydrogen donors from other cyclodextrin hydroxyl groups compared to O2 as a result of O3's donor preference to water. But again, the occurrence of intramolecular hydrogen bonds is increased after the interior water is removed by a ligand. The glucosidal O5 can act only as hydrogen acceptor and since it is fairly accessible by water because of its position on the rim of the tube-like structure it accepts preferably water hydrogens in comparison to CD hydrogens.

**Table 2** Occurrence of hydrogen bonds

	CD	Complex A	Complex B
<b>Donors</b>			
O2 to water	0.439	0.407	0.323
O3 to water	0.823	0.600	0.650
O6 to water	0.913	0.867	0.863
O2 to CD	0.563	0.605	0.687
O3 to CD	0.185	0.389	0.341
O6 to CD	0.082	0.121	0.131
All O2	1.001	1.012	1.010
All O3	1.008	0.989	0.990
All O6	0.995	0.989	0.994
<b>Acceptors</b>			
O1...H-OH	0.228	0.035	0.042
O2...H-OH	1.230	0.908	0.961
O3...H-OH	0.679	0.649	0.570
O5...H-OH	0.912	0.747	0.744
O6...H-OH	1.290	1.282	1.314
O1...H-CH	0.006	0.025	0.031
O2...H-CD	0.182	0.382	0.331
O3...H-CD	0.560	0.589	0.669
O5...H-CD	0.077	0.114	0.126
O6...H-CD	0.004	0.005	0.001
Lig from water	-	1.090	1.128
Lig from CD	-	0.101	0.085

The numbers represent the total count of hydrogen bonds divided by the number of frames and glucose residues (seven). In the case of SP the average of its four oxygen acceptors is given. All of them are sufficiently apart from the hydrophobic core of the  $\beta$ -CD to accept hydrogen bonds from the solvent

Throughout the simulations the capabilities of hydrogen bond formation are always covered even though the preferences for donors and acceptors shifts. Thus no negative influence of the formation could be found in this matter. The entropic effects of the loss of the hydrogen bond network inside the ring core are highly favorable since  $\beta$ -CD would otherwise put up with unfavorable bending measures to minimize this volume and increase the hydrophilic inside by turning two glucose units.

The calculated binding energies show a high affinity for both complexes (A and B). The nmode module of AMBER allows the calculation of the entropic energy contribution using a normal mode method. Its drawback is that it cannot take water molecules explicitly into account whose hydrogen bond network has a substantial contribution to this energy. The energies in Table 3 represent the  $\beta$ -CD–ligand complex formation alone without respect to this phenomenon.

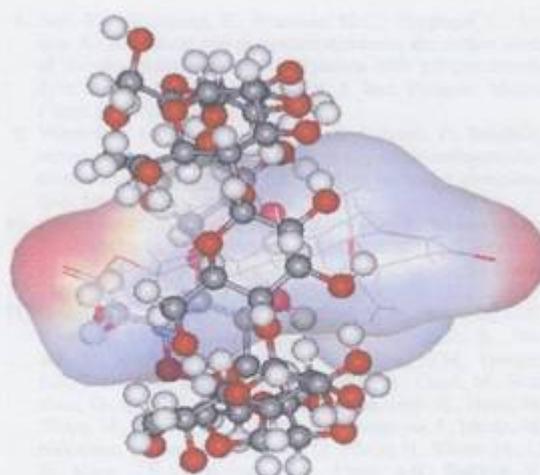
There are some differences of the energies for both orientations of SP. First, the size of the entered side of complex A is larger than in complex B and it establishes more VdW contacts to  $\beta$ -CD's hydrophobic interior core. Although in A more hydrogen bonds to  $\beta$ -CD can be formed, its tight packing does not allow an orientation for optimal electrostatic interaction, which results in the slightly worse overall binding energy (Table 3, electrostatic energy contribution).

The results of quantum chemical calculations on the interaction energy show rather different results depending on the methods and basis set used. Complex B possesses a lower energy throughout. Although in all cases a more or less complete inclusion of SP is observed (Fig. 2), the interaction energies differ significantly.

Applying the semiempirical method PM3 low, but quite diverse energies are obtained (Table 4). The low-level HF methods overemphasizes the association affinity, whereas HF/6-31G(d,p) results in low, but somewhat more similar energy values. The probably most reliable method (B3LYP/6-31G(d,p) delivers interaction energies, which are rather similar for both orientation of SP.

**Table 3** Binding energies and their contributions of both SP complexes with  $\beta$ -CD simulated in water environment

	Complex A	Complex B
Electrostatic	-11.48	-18.71
VdW	-33.66	-31.78
-TAS	15.33	15.38
$\Delta G_{\text{formation}}$	-29.81	-35.11



**Fig. 2** Inclusion complex between SP and  $\beta$ -CD in lowest-energy orientation (complex B) fully minimized by B3LYP/6-31G(d,p)

## Conclusion

Molecular dynamics simulations on  $\beta$ -CD and SP/ $\beta$ -CD complexes show that the formation of a rigid hydrogen bond network in the hydrophobic interior is responsible for large deformations of the  $\beta$ -CD ring. This flexion is energetically unfavorable for the ring itself and is a contributing factor for complex formation. After a ligand is placed hydrogen bonds are formed only on the rim and outside the ring structure. Furthermore, this study supports the idea that solvation is responsible for the excellent ligand binding properties of  $\beta$ -CD. The  $\beta$ -CD hydrate MD simulations are interesting model systems to investigate how conformational flexibility and dynamic disordering can influence inclusion complexes of a variety of compounds.

Molecular calculations on the geometry of SP/ $\beta$ -CD complexes show that the molecule is inserted completely into the cavity. For both possible orientations of SP, there is good agreement between the molecular surface of the guest molecule and the complementary  $\beta$ -CD surface, which explains to some extent the high

**Table 4** Binding energies of complexes A and B in gas phase, obtained from various quantum chemical calculations

Complex	PM3	HF/3-21G	HF/6-31G(d,p)	B3LYP/6-31G(d,p)
A	-2.86	-17.06	-4.80	-9.51
B	-7.94	-20.66	-5.37	-10.86

equilibrium constant of the inclusion complex in aqueous solution ( $K = 29,000$  L/mol). The relative large dependency of the calculated interaction energy shows that higher level ab initio or DFT methods have to be used to obtain reliable results.

**Acknowledgements** This investigation was supported by the Hochschuljubiläumsstiftung der Stadt Wien (Project P H-778/2005). Technical assistance of Ms. Martina Ziehengraser is gratefully acknowledged.

## References

1. Lawtrakul, L., Viernstein, H., Wolschann, P.: Molecular dynamics simulations of  $\beta$ -cyclodextrin in aqueous solution. *Int. J. Pharm.* **256**, 33–41 (2003).
2. Bea, I., Jaime, C., Kollman, P.: Molecular recognition by  $\beta$ -cyclodextrin derivatives: molecular dynamics, free-energy perturbation and molecular mechanics/Poisson-Boltzmann surface area goals and problems. *Theor. Chem. Acc.* **108**, 286–292 (2002).
3. Choi, Y.H., Yang, C.H., Kim, H.W., Jung, S.: Molecular modeling studies of the  $\beta$ -cyclodextrin in monomer and dimer form as hosts for the complexation of cholesterol. *J. Incl. Phenom. Macro. Chem.* **39**, 71–76 (2001).
4. Naidoo, K.J., Chen, J.Y.J., Jansson, J.L.M., Widmalm, G., Maliniak, A.: Molecular properties related to the anomalous solubility of  $\beta$ -cyclodextrin. *J. Phys. Chem. B* **108**, 4236–4238 (2004).
5. Choi, Y., Jung, S.: Simulations for the prediction of chiral discrimination of *N*-acetylphenylalanine enantiomers by cyclomaltoheptaose ( $\beta$ -cyclodextrin,  $\beta$ -CD) based on the MM-PBSA (molecular mechanics-Poisson-Boltzmann surface area) approach. *Carbohydr. Res.* **339**, 1961–1966 (2004).
6. Winkler, R.G., Fioravanti, S., Ciccotti, G., Margheritis, C., Villa, M.: Hydration of  $\beta$ -cyclodextrin: a molecular dynamics simulation study. *J. Comput.-Aid. Mol. Des.* **14**, 659–667 (2000).
7. Yu, Y., Chipot, C., Cai, W., Shao, X.: Molecular dynamics study of the inclusion of cholesterol into cyclodextrins. *J. Incl. Phenom. Macro. Chem.* **39**, 71–76 (2001).
8. Jara, F., Mascayano, C., Rezende, M.C., Tirapegui, C., Urzua, A.: A spectral and molecular dynamics simulation study of  $\beta$ -cyclodextrin inclusion complexes with solvatochromic dyes derived from barbituric acid. *J. Incl. Phenom. Macro. Chem.* **54**, 95–99 (2006).
9. Viernstein, H., Weiss-Greiler, P., Wolschann, P.: Solubility enhancement of low soluble biologically active compounds—temperature and cosolvent dependent inclusion complexation. *Int. J. Pharm.* **256**, 85–94 (2003).
10. Charumane, S., Titwan, A., Sirithunyalug, J., Weiss-Greiler P., Wolschann, P., Viernstein, H., Okonogi, S.: Thermodynamics of the encapsulation by cyclodextrins. *J. Chem Technol. Biotechnol.* **81**, 523–529 (2006).
11. Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Montgomery, J.A. Jr., Vreven, T., Kudin, K.N., Burant, J.C., Millam, J.M., Iyengar S.S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G.A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J.E., Hratchian, H.P., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski J.W., Ayala, P.Y., Morokuma, K., Voth, G.A., Salvador, P., Dannenberg, J.J., Zakrzewski, V.G., Dapprich, S., Daniels A.D., Strain, M.C., Farkas, O., Malick, D.K., Rabuck, A.D., Raghavachari, K., Foresman, J.B., Ortiz, J.V., Cui, Q., Baboul, A.G., Clifford, S., Cioslowski, J., Stefanov, B.B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R.L., Fox, D.J., Keith, T., Al-Laham, M.A., Peng, C.Y., Nana yakkara, A., Challacombe, M., Gill, P.M.W., Johnson, B. Chen, W., Wong, M.W., Gonzalez, C., Pople, J.A., Gaussian Inc., Wallingford CT, (2004).
12. Pearlman, D.A., Case, D.A., Caldwell, J.W., Ross, W.R., Cheatham, T.E., DeBolt, S., Ferguson, D., Seibel, G., Kollman, P.: AMBER, a package of computer program for applying molecular mechanics, normal mode analysis molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Comp Phys. Commun.* **91**, 1–41 (1995).
13. Berendsen, J.C.J., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., Haak, J.R.: Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690 (1984).

### **11.5 (Paper V: “Density Functional Calculations on Meloxicam- $\beta$ -CD Inclusion Complexes”)**

*This paper has been accepted on May 6<sup>th</sup>, 2009 for publication in International Journal of Pharmaceutics by Sasa Baumgartner, Special Issue Guest Editor, International Journal of Pharmaceutics.*

Owing to its extraordinary importance as an effective COX-2 inhibitor in medicinal and pharmaceutical applications the geometries and interaction energies of the  $\beta$ -CD inclusion complexes with various neutral, anionic, and cationic forms of meloxicam in gas phase were determined by DFT calculations. Moreover, the interaction energies were estimated including BSSE correction. In the complexes, two orientations of the meloxicam guest were considered in each case.

## Accepted Manuscript

Title: Density functional calculations on meloxicam- $\beta$ -cyclodextrin Inclusion Complexes

Authors: Walter Snor, Elisabeth Liedl, Petra Weiss-Greiler, Helmut Viernstein, Peter Wolschann

PII: S0378-5173(09)00292-0  
DOI: doi:10.1016/j.ijpharm.2009.05.012  
Reference: IJP 10645

To appear in: *International Journal of Pharmaceutics*

Received date: 27-11-2008  
Revised date: 22-4-2009  
Accepted date: 6-5-2009

Please cite this article as: Snor, W., Liedl, E., Weiss-Greiler, P., Viernstein, H., Wolschann, P., Density functional calculations on meloxicam- $\beta$ -cyclodextrin Inclusion Complexes, *International Journal of Pharmaceutics* (2008), doi:10.1016/j.ijpharm.2009.05.012

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



## Density functional calculations on meloxicam- $\beta$ -cyclodextrin

### Inclusion Complexes

Walter Snor<sup>a</sup>, Elisabeth Liedl<sup>a</sup>, Petra Weiss-Greiler<sup>a</sup>, Helmut Viernstein<sup>b\*</sup>,  
Peter Wolschann<sup>a</sup>

<sup>a</sup>Institute of Theoretical Chemistry, University of Vienna, Währinger Straße 17, A-1090 Wien, Austria

<sup>b</sup>Department of Pharmaceutical Technology and Biopharmaceutics, University of Vienna, Althanstraße 14, A-1090 Wien, Austria

\* Corresponding author: Tel.: +43 1 4277 55400; fax: +43 1 4277 9544.  
E-mail address: Helmut.Viernstein@univie.ac.at

#### ABSTRACT

The geometries of the cyclodextrin (CD) inclusion complexes with various tautomeric forms of meloxicam in gas phase were determined by DFT calculation (B3LYP/6-31G(d,p)). The interaction energies were estimated including basis set superposition error (BSSE) correction. Two orientations of the meloxicam guest were considered: the benzene ring located near the narrow rim and at the wider rim of the  $\beta$ -cyclodextrin, respectively. The calculations show that in all cases the molecules are located inside the CD-cavity. The preferred complexation orientation is that one, in which the benzene ring of meloxicam is located near the wider rim with the secondary hydroxyl groups of the CD. The stabilization energies for the encapsulation of the meloxicam guest molecules show an overall affinity ranking for the meloxicam guest molecule in the following order: anionic (deprotonated) form > zwitterionic form ~ enolic form > cationic (protonated) form. A comparison of the enolic and zwitterionic neutral forms

## ACCEPTED MANUSCRIPT

shows, that the zwitterionic structure is better stabilized upon complexation due to the geometry of two extra hydrogen bonds between host and guest.

*Keywords:* Meloxicam;  $\beta$ -Cyclodextrin; Inclusion complexation; Molecular Calculation; Density functional theory

Accepted Manuscript

## 1. Introduction

Cyclodextrins (CDs) are indispensable recipients not only in pharmacy and pharmaceutical technology, but also in many other scientific disciplines, like environmental, technical and analytical chemistry, for stereo-specific separations of diastereomers and optical isomers, for extraction of natural products, for protection and stabilization of light-, temperature- or oxidation-sensitive compounds. The reason for this broad field of various applications of CDs is their ability to form inclusion complexes with small or even medium-sized organic or inorganic compounds. Such an inclusion influences the physico-chemical behaviour of the guest molecules, like the reactivity or the solubility significantly. Emulsification of highly apolar compounds, change of the catalytic activities, support in organic syntheses, masking of odour or taste, increase of bioavailability and efficiency of the active substance as a consequence of solubility enhancement and the permission of controlled release are topics of actual CD research. (Szejtli 1982, 1996, 1998; Szejtli and Osa 1996; Bender and Momiyama (1978; Duchene and Wouessidjewe 1992).

Native CDs are obtained by the degradation of starch ( $\alpha(1 \rightarrow 4)$  linked polyglucose) by  $\alpha$ -1,4-glucan-glycosyltransferases. Depending on the respective transferase, different types of CDs result, consisting of 6 ( $\alpha$ -CD), 7 ( $\beta$ -CD), 8 ( $\gamma$ -CD) or more  $\alpha(1 \rightarrow 4)$  linked glucose units.  $\beta$ -cyclodextrin is one of the most widely used compound owing a cavity with an internal diameter of 6.5 Å and a depth of 8 Å.

Steric as well as electronic parameters of both the CDs and the guest molecules determine the driving forces of the complexation and the geometries of the inclusion

## ACCEPTED MANUSCRIPT

complexes. Many review articles have been published, which give excellent overviews about detailed descriptions of structural properties of CDs and CD complexes (Del Valle 2004; Dodziuk 2006; Szejtli 2004). Particularly, as a consequence of the high importance of CDs in pharmaceutical applications many reviews about pharmaceutical applications have been published (Vyas et al. 2008; Brewster and Loftsson 2007; Challa et al. 2005; Davis and Brewster 2004; Frömring and Szejtli 1994; Loftsson and Duchene 2007; Rajewski and Stella 1996).

In CDs all hydroxyl groups of the glucopyranose subunits of the CD molecule are orientated to the exterior of the molecule, with the primary hydroxyl groups located at the narrow rim of the torus and the secondary hydroxyl groups on the wider rim. The CD exterior is therefore hydrophilic, whereas the central cavity, lined with skeletal carbon and ether oxygen atoms of the glucopyranose units, is relatively hydrophobic and comparable to the lipophilicity of an aqueous ethanolic solution (Frömring and Szejtli 1994). The lipophilic cavity of a cyclodextrin molecule provides a microenvironment into which an appropriately sized nonpolar drug molecule, or more often nonpolar parts of the drug molecule, can enter to form an inclusion complex (Alcaro et al. 2004; Bodor and Buchwald 2002; Connors 1997; Kozar and Venanzi 1997; Lawtrakul et al. 2003; Lebrilla 2001; Liu and Guo 2002; Rekharsky and Inoue 1998; Viernstein et al. 2002).

The use of molecular modeling techniques for the study of cyclodextrins was somewhat limited in the past due to the size and flexibility of such molecules (Lipkowitz 1998). Some studies using molecular calculations based on empirical force field methods or on low-level semiempirical quantum chemical methods have been performed (Liu and Guo 2004; Liu et al 1999), but only a limited amount of calculations based on ab initio and DFT methods with suitably high levels of theory and

## ACCEPTED MANUSCRIPT

large basis sets, have been reported in the last years (Anconi et al 2007; Avakyan et al. 2005; Karpfen et al. 2007, 2008; Pinjari et al 2006, 2007; Snor et al. 2007, Weininger et al. 2007). However, the combination of experimental and computational studies has been recognized as a powerful tool for the study of the geometry of complexation.

In continuation of the work of Charumaneet et al. (2006) the current study examines the inclusion complexation of meloxicam and  $\beta$ -CD by means of density functional calculations. Meloxicam [4-hydroxy-2-methyl-N-(5-methyl-2-thiazolyl)-2H-1,2-benzothiazine-3-carboxamide-1,1-dioxide], (Fig. 1) belongs to the class of Nonsteroidal anti-inflammatory drugs (NSAID) of the enolic acid type of compounds. Previous studies reported the characterization of the stoichiometry and some insights on the 3D geometry of this inclusion complex (Abdoh 2007, Banerjee et al. 2003, 2004; Luger et al. 1996; Naidu et al. 2004; Tsai et al. 1993). However, there are no unequivocal conclusions in the literature on the mode of inclusion and the conformation of the drug inside the CD cavity.

Insert FIGURE 1

Tsai et al. (1993) reported that meloxicam exhibits only one  $pK_a$  (4.08) corresponding to the enolic OH. In contrast, Luger et al. (1996) estimated two  $pK_a$  values for meloxicam in aqueous solution, at 1.09 and 4.18, corresponding to ionization of the thiazole ring nitrogen and the enolic OH group, respectively. Thus, meloxicam exists as an anion at neutral pH and in weakly basic solutions, and is converted to cationic species at very low pH.

## ACCEPTED MANUSCRIPT

## 2. Materials and methods

We have analyzed the possible tautomers of all protonation states of meloxicam by DFT methods and have determined the geometries of the inclusion complexes with  $\beta$ -CD.

The initial structure of  $\beta$ -CD was obtained by manually building up the Z-matrix starting from the oxygen-oxygen distances of the primary hydroxyl groups at the more narrow rim of the  $\beta$ -CD molecule. The glycosidic oxygens of the CD were placed onto the XY plane and their center was defined as center of the coordinate system. Only a single combination of the OH-group orientations was considered (all hydroxymethyl and hydroxyl groups were oriented counterclockwise). This structure with  $C_7$  symmetry was then fully geometry optimized without symmetry restriction with a DFT B3LYP/6-31G(d,p) calculation using the program package GAUSSIAN 03 (Frisch et al., 2004). The obtained geometry was taken for further modeling of the inclusion complexes with meloxicam.

### 2.1. Selection of the tautomers of meloxicam.

Four low-energy basic conformations of meloxicam, given in Figure 2, were used to derive all theoretical possible tautomeric forms of the molecule.

Insert FIGURE 2

Each of these possible tautomeric structures was subjected to a 5 picosecond simulated annealing run, using the MM+ force field and monitoring the three dihedral angles, viz.  $\theta_1$ , defined by C(4)-C(3)-C(15)-N(17),  $\theta_2$ , defined by O(16)-C(15)-N(17)-C(18), and  $\theta_3$ , defined by C(15)-N(17)-C(18)-N(19), respectively. Upon placing the dihedrals  $\theta_1$ ,  $\theta_2$  and  $\theta_3$  at the positions with the energetically most favoured values, a new structure was created in either case and was fully energy optimized with B3LYP/6-31G (d,p). In this way a total amount of sixteen structures was reached.

## ACCEPTED MANUSCRIPT

Out of these 16 structures the energetically most favourable enolic, zwitterionic, anionic (deprotonated) and cationic (protonated) form of meloxicam, viz. Bn, Az, Dd and Ap respectively (Fig 3.) were used to build up the inclusion complexes with  $\beta$ -CD.

Insert FIGURE 3

### 2.2. Force Field Calculations on the complex geometry

The four energetically most favourable forms of meloxicam were used to create CD inclusion complexes. These meloxicam- $\beta$ -CD complexes were constructed by manually introducing the meloxicam molecule into the  $\beta$ -CD cavity through the narrow and wide rim of the latter, respectively, centering it on a vector perpendicular to the mean plane through the glycosidic oxygen atoms.

For the complexation process, the CD was kept in this position while the meloxicam approached along the Z-axis toward the narrow and wide rim of the CD, respectively. Two possible orientations of the guest molecule were considered. The relative position between meloxicam and CD was measured by the distance C3 of the meloxicam guest and the origin. The molecule was initially located at a distance of 10 Å from the origin and was moved in decreasing increments of 2 Å towards the center of the CD cavity. At each step the entire structure was optimized without any restriction, using the United Force Field (UFF) implemented in GAUSSIAN 03. Then the increments were increased and the meloxicam moved away from the CD. Once the meloxicam had been translated to 10 Å beyond the origin, the procedure was terminated. By this procedure local minima can be found and avoided.

### 2.3. Density functional theory calculations.

To construct these inclusion complexes for the optimization with the DFT B3LYP/6-31G (d,p) method the results of the preliminary force field calculations (see chapter 2.2) were used. For this purpose the C3 of the meloxicam molecule was placed in each case on the central axis of the CD molecule at a distance from the origin that corresponds to the energy minimum of the force field calculations (Fig 4.).

For the DFT calculation, no constraints were imposed on the whole system, especially no parameters were fixed. So the meloxicam molecule was free to move in the cavity of

## ACCEPTED MANUSCRIPT

the  $\beta$ -CD during the whole optimization process. Therefore, conformational changes of the host as well as of the guest molecule were explicitly allowed.

In this way eight complex structures of CD/meloxicam complexes were fully optimized with B3LYP/6-31G (d,p). No restrictions were imposed on the complexes during the geometry minimization process, and no close contacts have been established.

Table 1 lists particular free input parameters before and their respective output values after the optimization process.

Insert TABLE 1

The stabilization energy ( $\Delta E$ ) between meloxicam and  $\beta$ -CD was calculated for the minimum energy structure according to

$$\Delta E = E_{\text{complex}} - (E_{\text{meloxicam}} + E_{\beta\text{-CD}}).$$

In order to take into account the basis set superposition error (BSSE) correction on the energy determination, the counterpoise (CP) correction (Boys and Bernardi 1970) was estimated, and, additionally, a full geometry optimization was performed including BSSE correction.

### 3. Results and Discussion

#### 3.1. Force Field Calculations

As preliminary studies based on empirical force fields systematic conformational searches on the geometry of CD/meloxicam complexes were performed by varying the position of meloxicam inside the cavity. In Fig. 4 the calculated force field energies are plotted as a function of the internal distance between the CD ring and the meloxicam guest molecules.

## ACCEPTED MANUSCRIPT

Insert FIGURE 4

Only one pronounced energy minimum can be observed for each orientation together with some smaller local minima, which means that the inclusion complexes should have well-defined geometries depending only on the orientation of the inserted molecule. Also the rotation of the molecule inside the cavity leads to a single geometry only.

### 3.2. Density Functional Theory Calculations

From the 16 tautomeric forms of meloxicam of lowest energy 4 energetically most favourable enolic, zwitterionic, anionic (deprotonated) and cationic (protonated) forms of meloxicam, viz. Bn, Az, Dd and Ap, respectively (Fig 3.) were used to investigate the inclusion complexes with  $\beta$ -CD. Full geometry optimizations without any constraint with the DFT B3LYP/6-31G(d,p) method were employed in the study of the complexation process of these tautomeric forms of meloxicam with  $\beta$ -CD.

Figures 5-8 show the top view and side view of the most stable deprotonated (anionic) and neutral (zwitterionic) inclusion complexes, respectively.

Insert FIGURE 5

Insert FIGURE 6

Insert FIGURE 7

Insert FIGURE 8

### 3.3. Structural behaviour of the meloxicam-CD complexes

In all complex geometries, a considerable part of the guest molecule is accommodated in the  $\beta$ -CD cavity. The long axis of the meloxicam guests is oriented along the axis of the  $\beta$ -CD, with the B and C rings (Fig. 1) lying within the cavity. Both orientations, form 1 as well as form 2, are energetically possible and lead to conformational minima.

## ACCEPTED MANUSCRIPT

The preferred complexation orientation is that one, in which the benzene ring of meloxicam is located near the wider rim with the secondary hydroxyl groups of the CD (form 2). In all cases, the ring of seven hydrogen bonds of the CD, comprising the secondary OH groups at the wider rim, remains unchanged upon complexation.

Moreover, extra hydrogen bonds are formed between the meloxicam guest and secondary OH-groups of  $\beta$ -CD in the case of the anionic and zwitterionic complexes, respectively.

While in form 1 of the anionic complex one extra hydrogen bond has been established between N19 of the meloxicam guest and a secondary OH group of the CD (not shown in the Figures), in form 2 of the anionic complex one extra hydrogen bond between O13 of meloxicam and a secondary OH group of CD was observed (see Fig. 5 and 6).

In the case of form 1 of the zwitterionic complex, one extra hydrogen bond was established between N19 of the meloxicam and a secondary OH group of the CD (not shown in the Figures). In form 2 of the zwitterionic complex, both sulfur oxygens of the meloxicam guest molecule, viz. O13 and O14, were found to be involved in an extra hydrogen bond to a secondary OH group of the CD (see Fig. 7 and 8), respectively.

Upon complexation, the CD remains basically undistorted. To confirm this, the mean distance of the CD's mass center (centroid) to the seven glycosidic oxygens was calculated in each case. It was found that the extent of variation of this distance was marginal ( $5.09 \pm 0.15$  Å). The corresponding values of the uncomplexed CD were  $5.11 \pm 0.002$  Å.

Table 1 comprises a comparison of certain parameters of meloxicam and meloxicam in the cavity of CD-complex. The calculated parameters show, that the geometry of the meloxicam guest remains to a great extent undistorted, too.

According to the parameters calculated for the extra hydrogen bonds considered above between the meloxicam guest and secondary OH groups of the CD host in the anionic (deprotonated) and neutral (zwitterionic) inclusion complexes, respectively, these interactions can be classified as moderate hydrogen bonds (Scheiner 1979). The criteria employed in this classification state that moderate hydrogen bonds are characterized by A...H-D distances between 1.5-2.2 Å and D-H... angles between  $\approx$

## ACCEPTED MANUSCRIPT

130-180 degrees, where A represents a hydrogen acceptor and D a hydrogen donor atom, respectively.

As an example, the partial charges on the different atoms of the meloxicam conformation Az have been calculated and are depicted in Figure 9.

Insert Figure 9

### 3.4. Energetic behaviour of the meloxicam-CD complexes.

To examine the stability of the respective complexes, the stabilization energy was obtained from the difference between the complex energy and the sum of the energy of the isolated host and guest molecules. The calculated stabilization energies for the encapsulation of the meloxicam guest molecules show an overall affinity ranking for the meloxicam guest molecule in the following order: Deprotonated form > enolic form ~ zwitter ionic form > protonated form.

Table 2 presents the DFT B3LYP/6-31G (d,p) interaction energies ( $\Delta E$ ) of the minimum energy forms of various 1:1 meloxicam  $\beta$ -CD complexes. The uncorrected and BSSE-corrected energy values are depicted

Insert TABLE 2

## 4. Conclusions

The complexation of neutral, zwitterionic, anionic (deprotonated) and cationic (protonated) meloxicam with  $\beta$ -CD was studied by force field and density functional B3LYP/6-31G(dp) calculations, in the latter case including a fully optimized

## ACCEPTED MANUSCRIPT

Counterpoise correction of the BSSE effect. The application of the DFT method leads to more reliable geometries than the widely used molecular mechanics and semiempirical methods. More possibilities of interactions as Van der Waals forces and H-bonds are taken into account by these accurate calculations, leading to well defined complexes and interaction energies with higher accuracy compared to the other methods.

The tautomeric forms of meloxicam studied here show that they are capable of forming stable inclusion complexes with the host  $\beta$ -cyclodextrin.

The optimized structures of the inclusion complexes reveal an overall affinity ranking for the meloxicam guest molecule in the following order: Deprotonated form > enolic form ~ zwitter ionic form > protonated form.

In the case of the deprotonated (anionic) complex, in both forms one extra hydrogen bond has been established between the meloxicam guest and the CD host which may account for the stabilization of these complexes. In the case of the neutral forms of meloxicam, the better stabilization of the zwitterionic form over the enolic form may be contributed to the establishment of two extra hydrogen bonds between host and guest.

Upon complexation, no remarkable distortion of the  $\beta$ -CD host and the meloxicam guest molecules were detected. For the  $\beta$ -CD this was confirmed by comparing the variation of the mean distances between the centroid of the CD and the seven glycosidic oxygen atoms. In the case of the meloxicam guest molecules, certain parameters such as dihedral angles before and after the complexation process have been compared (see Table 1).

Nevertheless, the calculations were performed on molecules in the gas phase. There is, however, a discussion about the binding forces. In water, cyclodextrins are

## ACCEPTED MANUSCRIPT

believed to form inclusion complexes that are mostly stabilized by hydrophobic interactions between the unpolar surface inside the cyclodextrin cavity and the surface of the guest. In the gas phase, the energetics that benefit from this arrangement do not exist because there are no water molecules that surround the complex. Consequently, there are no benefits from hydrophobic interactions in the gas phase. Rather, van der Waals forces and, if possible, hydrogen bonding remain.

It should also be mentioned, that the absence of explicit solvent did not allow us to consider the hydrophobic effect in a proper way. Despite this drawback, we still get a favorable interaction energy between the two components of the respective host-guest complexes. The solvophobic effect will further improve the stabilization of the complex. The aryl moieties in the meloxicam molecule, viz., ring A and B, being predominantly hydrophobic in nature should be partially stabilized by the solvophobic effect, while the interaction of the third ring C with  $\beta$ -CD is expected to be guided by the electrostatic effect.

Moreover, stronger interactions can be expected for a host with so many hydroxy dipoles that may arrange in a favorable fashion around a ionic guest. For a optimization of the number of such interactions, it is likely that the guest is located inside the cavity to provide a geometrically reasonable fit between guest and cavity. The conclusion from these considerations is that the complexes exist in solution as well as in the gas phase, but for different reasons.

**Acknowledgements**

The authors thank Ms. M. Ziehengraser for technical assistance

## ACCEPTED MANUSCRIPT

**References**

- Abdoh AA, El-Barghouthi MI, Zughul MB, Davies JE, Badwan AA, 2007. Changes in the conformational structure, microscopic and macroscopic  $pK_{aS}$  of meloxicam on complexation with natural and modified cyclodextrins. *Pharmazie*, 62, 55
- Alcaro S, Battaglia D, Ortuso F, 2004. Molecular modeling of  $\beta$ -cyclodextrin inclusion complexes with pharmaceutical compounds. *ARKIVOC*, V,107
- Anconi CPA, Nascimento CS jr., Fedoce-Lopes J, Dos Santos HF, De Almeida WB, 2007. Ab Initio Calculations on Low-Energy Conformers of  $\alpha$ -Cyclodextrin. *J Phys Chem A*, 111, 12127
- Avakyan VG, Nazarov VB, Voronezhcheva NI, 2005. DFT and PM3 Calculations of the Formation Enthalpies and Intramolecular H-Bond Energies in  $\alpha$ -,  $\beta$ -, and  $\gamma$ -Cyclodextrins. *Russ J Phys Ch*, 79 (1), S18
- Banerjee R, Chakraborty H, Sarkar M, 2004. Host-Guest Complexation of Oxicam NSAIDs with  $\beta$ -Cyclodextrin. *Biopolymers*, 75, 355
- Bender ML, Momiyama M, 1978. *Cyclodextrin Chemistry*. Springer, Berlin
- Bodor N, Buchwald P, 2002. Theoretical Insights into the Formation, Structure, and Energetics of Some Cyclodextrin Complexes. *J Incl Phenom Macro*, 44, 9
- Boys SF, Bernardi F, 1970. Calculation of small molecular interactions by differences of separate total energies – some procedures with reduced errors. *Mol Phys*, 19, 553
- Brewster ME, Loftsson T, 2007. Cyclodextrins as pharmaceutical solubilizers. *Adv Drug Deliver Rev*, 59, 645

## ACCEPTED MANUSCRIPT

- Challa R., Ahuja A, Ali J., Khar RK, 2005. Cyclodextrins In Drug Delivery: An Updated Review. *AAPS PharmSciTech*, 6, 329
- Charumane S, Titwan A, Sirithunyalug J, Weiss-Greiler P, Wolschann P, Viernstein H, Okonogi S, 2006. Thermodynamics of the encapsulation by cyclodextrins. *J Chem Technol Biot*, 81, 523
- Connors, KA, 1997. The Stability of Cyclodextrin Complexes in Solution. *Chem. Rev*, 97, 1325
- Davis ME, Brewster ME, 2004. Cyclodextrin-based Pharmaceuticals: Past, Present and Future. *Nat Rev Drug Discov*, 3, 1023
- Del Valle EMM, 2004. Cyclodextrins and their uses: a review. *Process Biochem*, 39, 1033
- Dozduik H, 2006. Cyclodextrins and their Complexes. Wiley-Vch, Weinheim
- Duchene D, Wouessidjewe D, 1992. Industrial Uses of Cyclodextrins and their Derivatives. *J Coord Chem*, 27 (1-3), 223
- Frömming KH, Szejtli J, Ed, 1994. Cyclodextrins in Pharmacy. Dordrecht: Kluwer Academic Publishers.
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski JA, Montgomery JJA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL,

## ACCEPTED MANUSCRIPT

Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (2004) GAUSSIAN 03, Revision C.02, Gaussian Inc., Wallingford CT

Karpen A, Liedl E, Snor W, Wolschann P, 2007. Homodromic hydrogen bonds in low-energy conformations of single molecule cyclodextrins. *J Incl Phenom Macro* 57, 35

Karpen A, Liedl L, Snor W, Viernstein H, Weiss-Greiler P, 2008. Density functional calculations on cyclodextrins. *Monatsh Chem* 139, 363

Kozar T, Venanzi CA, 1997. Reconsidering the conformational flexibility of  $\beta$ -cyclodextrin. *J Mol Struct*, 395, 451

Lawtrakul L, Viernstein H, Wolschann P, 2003. Molecular dynamics simulations of  $\beta$ -cyclodextrin in aqueous solution. *Int. J. Pharm*, 256, 33

Lebrilla CB, 2001. The Gas-Phase Chemistry of Cyclodextrin Inclusion Complexes. *Accounts Chem Res*, 34, 653

Lipkowitz KB, 1998. Applications of Computational Chemistry to the Study of Cyclodextrins. *Chem Rev*, 98, 1829

Liu L, Guo QX, 2002. The Driving Forces in the Inclusion Complexation of Cyclodextrins. *J Inc Phenom. Macro*, 42, 1

Liu L, Guo QX, 2004. Use of Quantum Chemical Methods to Study Cyclodextrin Chemistry. *J Incl Phenom Macro*, 50, 95

Liu L, Li XS, Guo QX, Liu YC, 1999. Hartree-Fock and Density Functional Theory Studies on the Molecular Recognition of the Cyclodextrin. *Chinese Chem Lett*, 10 (12), 1053

Lofsson T, Duchene D, 2007. Cyclodextrins and their pharmaceutical applications. *Int J Pharm*, 329, 1

## ACCEPTED MANUSCRIPT

- Luger P, Daneck K, Engel W, Trummlitz G, Wagner K, 1996. Structure and physicochemical properties of meloxicam, a new NSAID. *Eur J Pharm Sci*, 4, 175
- Naidu NB, Chowdary KPR, Murthy KVR, Satyanarayana V, Hayman AR, Becket G, 2004. Physicochemical characterization and dissolution properties of meloxicam-cyclodextrin binary systems. *J Pharmaceut Biomed*, 35, 75
- Pinjari RV, Joshi KA, Gejji SP, 2006. Molecular Electrostatic Potentials and Hydrogen Bonding in  $\alpha$ -,  $\beta$ -, and  $\gamma$ -Cyclodextrins. *J PhysChem A*, 110 (48), 13037
- Pinjari RV, Joshi KA, Gejji SP, 2007. Theoretical Studies on Hydrogen Bonding, NMR Chemical Shifts and Electron Density Topography in  $\alpha$ ,  $\beta$  and  $\gamma$ -Cyclodextrin Conformers. *J Phys Chem A*, 111,13583
- Rajewski RA, Stella VJ, 1996. Pharmaceutical applications of cyclodextrins.2. In vivo drug delivery. *J Pharm Sci*, 85, 1142
- Rekharsky MV, Inoue Y, 1998. Complexation Thermodynamics of Cyclodextrins. *Chem. Rev*, 98, 1875
- Scheiner S, 1979. *Hydrogen Bonding. A Theoretical Perspective*. Oxford: University Press
- Snor W, Liedl E, Weiss-Greiler P, Karpfen A, Viernstein H, Wolschann P, 2007. On the structure of anhydrous  $\beta$ -cyclodextrin. *Chem Phys Lett*, 441, 159
- Szejtli J, 1982. *Cyclodextrins and their Inclusion Complexes*. Budapest: Akademiai Kiado.
- Szejtli J, 1998. Introduction and General Overview of Cyclodextrin Chemistry. *Chem Rev*, 98, 1743

## ACCEPTED MANUSCRIPT

Szejtli J, 2004. Past, present, and future of cyclodextrin research. *Pure Appl Chem*, 76, 1825

Szejtli J, Osa T (eds) 1996. *Comprehensive Supramolecular Chemistry: Cyclodextrins*. Pergamon

Tsai RS, Carrupt PA, El Tayar N, Giroud Y, Anrade P, Testa B, Bree F, Tillement JP, 1993. Physicochemical and Structural Properties of Nonsteroidal Anti-inflammatory Oxicams. *Helv Chim Acta*, 76, 842

Viernstein H, Weiss-Greiler P, Wolschann P, 2002. Solubility Enhancement of Low Soluble Biologically Active Compounds by  $\beta$ -Cyclodextrin and Dimethyl- $\beta$ -cyclodextrin. *J Incl Phenom Macro*, 44, 235

Vyas A, Saraf S, Saraf S, 2008. Cyclodextrin based novel drug delivery systems. *J Incl Phenom Macro*, 62, 23

Weinzinger P, Weiss-Greiler P, Snor W, Viernstein H, Wolschann P, 2007. Molecular dynamics simulations and quantum chemical calculations on  $\beta$ -cyclodextrin-spironolactone complex. *J Incl Phenom Macro*, 57, 29

## ACCEPTED MANUSCRIPT

**Figure Captions**

Fig.1. Enolic conformation of meloxicam (atom numbering for reference only, not systematic)

Fig. 2. Basic molecular structures of meloxicam

Fig.3. Structures of the energetically most favourable enolic (Bn), zwitterionic (Az), anionic (Dd) and cationic (Ap) conformation of meloxicam. Hydrogen bonds are indicated by dotted lines

Fig. 4 calculated force field energies as a function of the internal distance between C3 of the meloxicam guest molecule and the origin. The dotted line corresponds to form 2 of the anionic (deprotonated) complex, and the full line to form 1 of the anionic complex, respectively.

Fig.5 B3LYP/6-31G (d,p) energy minimized structure of the meloxicam Dd- $\beta$ -cyclodextrin complex (form 2). View along the z-axis. Hydrogen bonds are indicated by dotted lines

Fig.6 B3LYP/6-31G (d,p) energy minimized structure of the meloxicam Dd- $\beta$ -cyclodextrin complex (form 2). View perpendicular to the z-axis. Hydrogen bonds are indicated by dotted lines

Fig.7 B3LYP/6-31G (d,p) energy minimized structure of the meloxicam Az- $\beta$ -cyclodextrin complex (form 2). View along the z-axis. Hydrogen bonds are indicated by dotted lines

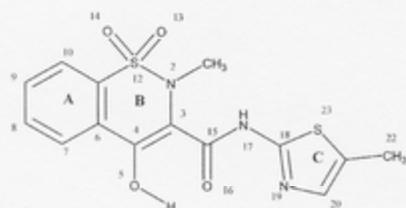
Fig.8 B3LYP/6-31G (d,p) energy minimized structure of the meloxicam Az- $\beta$ -cyclodextrin complex (form 2). View perpendicular to the z-axis. Hydrogen bonds are indicated by dotted lines

Fig.9 B3LYP/6-31G(d,p) energy minimized structure of the meloxicam Az tautomer; the partial charges on each atom are depicted

ACCEPTED MANUSCRIPT

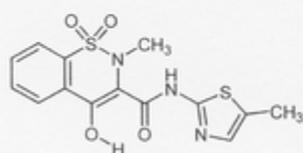
## FIGURES

Figure 1

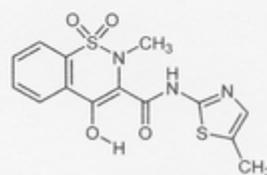


ACCEPTED MANUSCRIPT

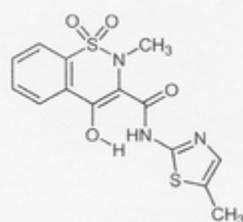
Figure 2



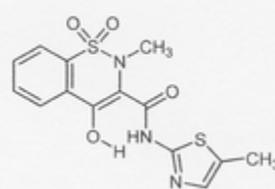
Type A



Type B



Type C

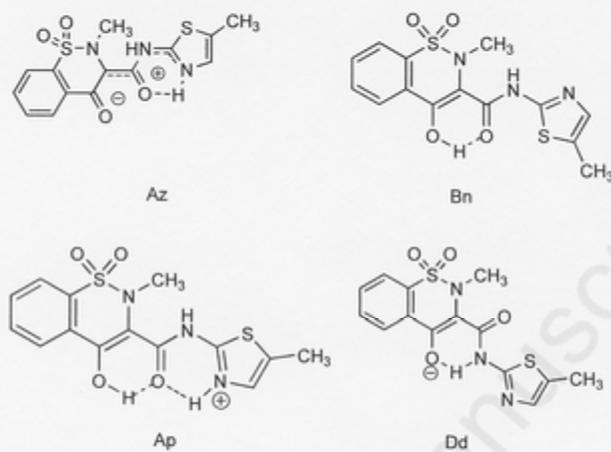


Type D

Accepted

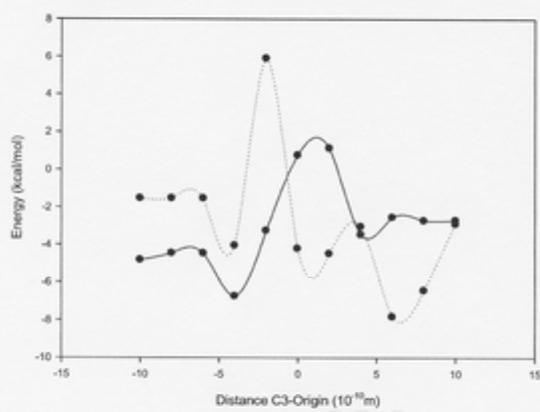
ACCEPTED MANUSCRIPT

Figure 3



ACCEPTED MANUSCRIPT

Figure 4



ACCEPTED MANUSCRIPT

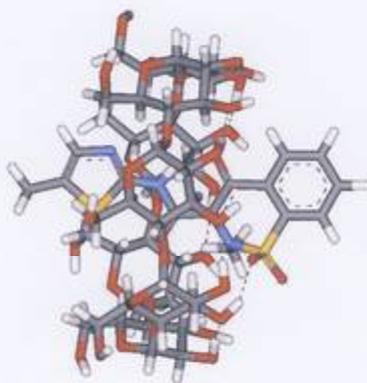
Figure 5



Accepted Manuscript

ACCEPTED MANUSCRIPT

Figure 6



Accepted M,

ACCEPTED MANUSCRIPT

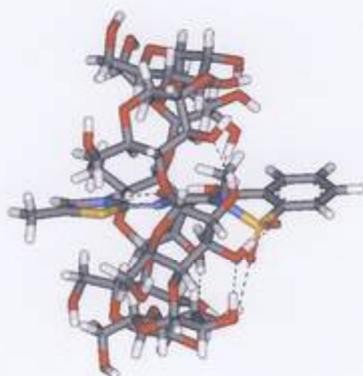
Figure 7



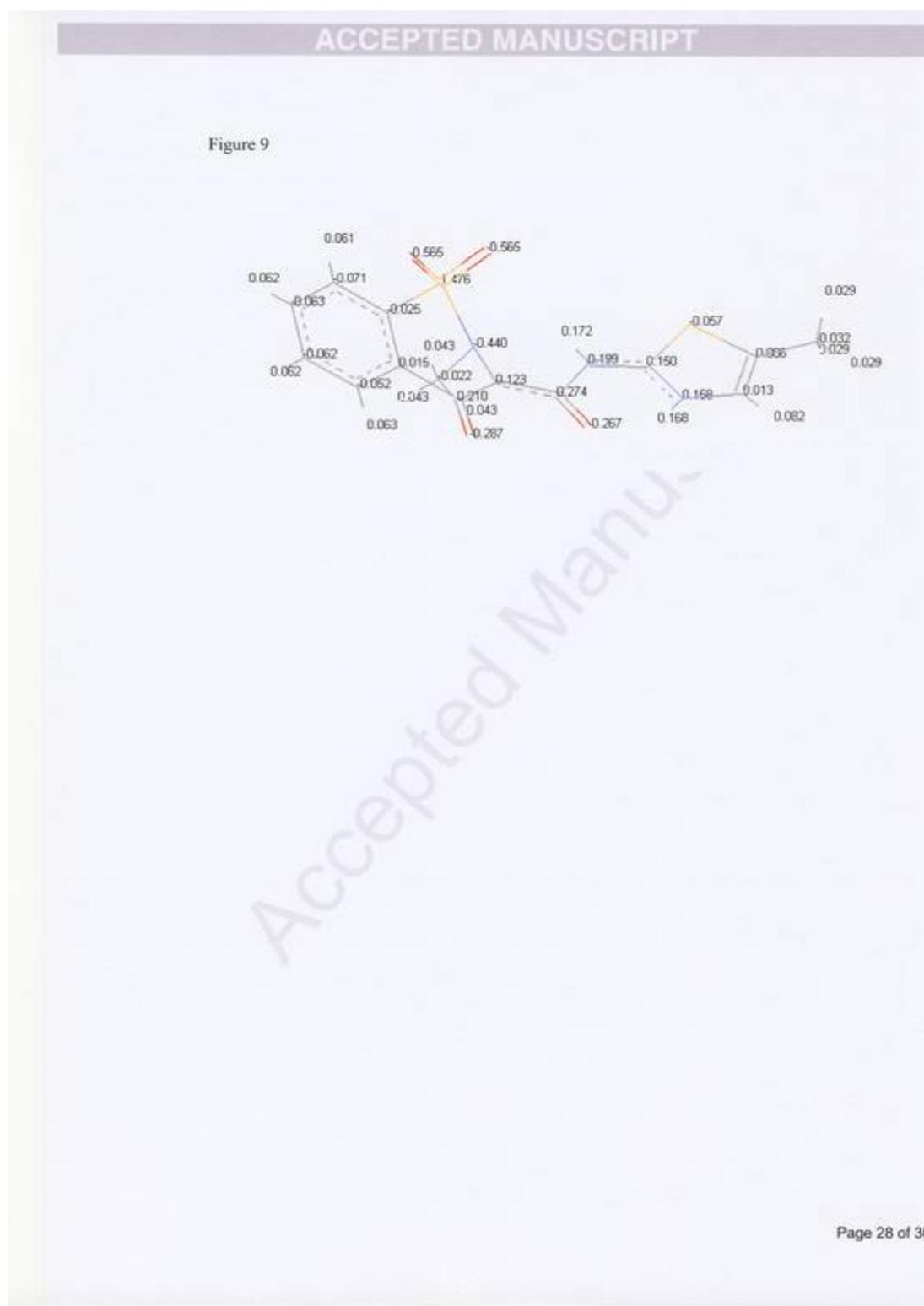
Accepted M.

ACCEPTED MANUSCRIPT

Figure 8



Accepted M,



## ACCEPTED MANUSCRIPT

## TABLES

Table 1

Tab.1. Comparison of particular parameters of various meloxicam guest molecules in their free and complexed form, respectively.  $\theta_1$ ,  $\theta_2$  and  $\theta_3$  are defined by C(4)-C(3)-C(15)-N(17), O(16)-C(15)-N(17)-C(18) and C(15)-N(17)-C(18)-N(19), respectively. A represents a hydrogen acceptor atom and D a hydrogen donor atom, respectively.

Guest	Complex-form	$\theta_1(^{\circ})$	$\theta_2(^{\circ})$	$\theta_3(^{\circ})$	H-Bond D→A	Length(Å)	Angle(D-H...A; $^{\circ}$ )
Bn	-	-170.5	178.7	-177.3	O5-O16	1.64	149.0
	1	-169.0	177.6	-177.4	O5-O16	1.71	144.8
	2	-172	-179.3	-178.8	O5-O16	1.67	146.4
Az	-	179.6	-177.3	-179.7	O5-16	1.58	150.6
					N19-O16	1.87	125.8
	1	174.9	-172.5	-179.2	O5-O16	1.57	151.1
					N19-O16	1.98	120.6
	2	178.3	-177.6	179.4	O5-O16	1.57	150.1
					N19-O16	1.86	126.0
Dd	-	177.9	179.5	-179.2	N17-O5	1.73	140.6
	1	-177.8	174.2	-169.4	N17-O5	1.67	143.0
	2	178.1	178.5	-177.9	N17-O5	1.7	141.5
Ap	-	165.6	-177.6	176.4	N19-O16	1.9	126.7
					O5-O16	1.76	144.2
	1	166.3	-175.5	-168.9	N19-O16	2.16	112.9
					O5-O16	1.81	140.6
	2	166.3	179.5	177.5	N19-O16	1.86	127.6
					O5-O16	1.81	140.0

## ACCEPTED MANUSCRIPT

Table 2

Tab.2 DFT B3LYP/6-31G (d,p) interaction energies ( $\Delta E$ ; kcal/mol) of the minimum energy forms of various 1:1 meloxicam  $\beta$ -cyclodextrin complexes. The uncorrected and BSSE-corrected energy values are depicted.

Components/form	B3LYP uncorr.	BSSE-Energy	B3LYP corr.
$\beta$ -CD+Bn/1	-3.3	4.79	1.49
$\beta$ -CD+Bn/2	-9.9	7.53	-2.37
$\beta$ -CD+Dd/1	-25.8	10.76	-15.04
$\beta$ -CD+Dd/2	-28.9	8.91	-19.99
$\beta$ -CD+Ap/1	-3.6	8.71	5.11
$\beta$ -CD+Ap/2	-5.6	7.45	1.85
$\beta$ -CD+Az/1	-8.9	7.37	-1.53
$\beta$ -CD+Az/2	-10.5	11.43	0.93

## 12 Conclusion and Future Directions

It was the aim of this thesis to perform computational investigations on CDs and CD complexes by using quantum mechanical ab initio and DFT methods at an appropriate high level of theory.

The molecular systems studied in this context include:

- The hydrogen-bonded water-methanol complex (chapter 3)
- Malonaldehyde (chapter 5)
- $\beta$ -CD Inclusion complexes with D- and L Alanine (chapter 7)
- $\beta$ -CD Inclusion complexes with neutral and anionic conformations of meloxicam (chapter 10)
- $\alpha$ -,  $\beta$ - and  $\gamma$ -CD single molecules (Paper I, II and III)
- $\beta$ -CD Spironolactone Inclusion complex (Paper IV)
- $\beta$ -CD Inclusion complexes with neutral, anionic and cationic conformations of meloxicam (Paper V)

All calculations were performed in gas phase. Geometries and energies of CDs and CD inclusion complexes as listed above were obtained mainly by DFT calculations at the B3LYP/6-31G(d,p) level of theory.

So far, no comparable investigations have been performed at this accurate level of theory.

As to the further development of molecular modelling and computational chemistry: Over the next decade, it is expected to see significant increase in accuracy in all principal quantum chemical directions, such as the development of better DFT functionals etc. Continued reductions in the cost/performance of computing and improvements in algorithmic details should continue to yield shorter time to solution for increasingly larger systems.

Similarly, improvement can be expected in treatment of the solvent effect. Optimization of the accuracy of continuum solvation methods is far from a solved problem to date.

Hence, it is hoped that a large number of important new applications are to be carried out over the next decade; with advances in theory, software and computational hardware, larger data sets and systems of increasing size should be amenable to study.

However, the most exciting possibility is that the parallel advances in theory and experiment will enable fully explanatory and predictive models to be constructed for the complex, condensed-phase processes that govern most of the natural world.

## 13 Abbreviations

AMBER	Assisted Model Building with Energy Refinement
AM1	Austin Model 1
AO	Atomic Orbital
BLYP	Functional developed by Becke, Lee, Yang, and Parr
BSSE	Basis Set Superposition Error
B3LYP	Hybrid Functional, also called Becke3LYP
CBS	Complete Basis Set
CCDC	Cambridge Crystallographic Data Centre
CC	Coupled Cluster
CD	Cyclodextrin
CCL	Computational Chemistry List
CCSD	CC Including Single and Double Excitation
CCSD(T)	CC Including Single, Double and Triple Excitation
CHA	Chemical Hamiltonian Approach
CI	Configuration Interaction
CIS	CI including Single Excitation
CISD	CI including Single and Double Excitation
CNDO	Complete Neglect of Differential Overlap
COX-2	Cyclooxygenase-2
CP	Counterpoise
CPK	Corey-Pauling-Koltun
CPU	Central Processing Unit
CSD	Cambridge Structural Database
DFT	Density Functional Theory
DMPC	Dimyristoylphosphatitylcholine
DNA	Desoxyribonucleic acid
GGA	Generalized Gradient Approximation

GROMOS	Groningen Molecular Simulation
GTO	Gaussian-Type Orbital
GUI	Graphical User Interface
HF	Hartree-Fock
HOMO	Highest Occupied Molecular Orbital
IPCM	Isodensity PCM Model
LDA	Local Density Approximation
LUMO	Lowest Unoccupied Molecular Orbital
MC	Monte Carlo
MD	Molecular Dynamics
MEP	Molecular Electrostatic Potential
MLP	Molecular Lipophilicity Pattern
MMFF	Merck Molecular Force Field
MM	Molecular Mechanics
MNDO	Modified Neglect of Diatomic Overlap
MO	Molecular Orbital
MPn	Møller-Plesset n <sup>th</sup> Order
NMR	Nuclear Magnetic Resonance
NSAID	Nonsteroidal Anti-inflammatory Drug
ONIOM	Our Own N-Layered integrated MO and MM Method
PC	Personal Computer
PCM	Polarized Continuum Model
PDB	Brookhaven Protein Data Bank
PES	Potential Energy Surface
PM3	Parameterization Method 3
QCISD(T)	Quadratic CI Including Single, Double and Triple Excitation
QM	Quantum Mechanics
RAHB	Resonance-Assisted Hydrogen Bonding
RHF	Restricted Hartree Fock
ROHF	Restricted open shell Hartree-Fock
SCF	Self consistent field

---

SCI-PCM	Self-Consistent Isodensity Polarized Continuum Model
SCRf	Self-Consistent Reaction Field
SP	Spironolactone
STO	Slater-Type Orbital
UFF	United Force Field
UHF	Unrestricted Hartree-Fock
VSEPR	Valence-Shell Electron Pair Repulsion
XC	Exchange Correlation
ZPE	Zero Point Energy

## 14 References

Abdoh AA, El-Barghouthi MI, Zughul MB, Davies JE, Badwan AA, 2007. Changes in the conformational structure, microscopic and macroscopic  $pK_a$ s of meloxicam on complexation with natural and modified CDs. *Pharmazie*, 62, 55

Alcaro S, Battaglia D, Ortuso F, 2004. Molecular modelling of  $\beta$ -CD inclusion complexes with pharmaceutical compounds. *ARKIVOC*, V, 107

Anconi CPA, Nascimento CS jr., Fedoce-Lopes J, Dos Santos HF, De Almeida WB, 2007. Ab Initio Calculations on Low-Energy Conformers of  $\alpha$ -CD. *J Phys Chem A*, 111, 12127

Atkins P, Friedman R, 2005. *Molecular QM*. Oxford: University Press

Atwood JL, Davis JE, MacNicol DD, Eds, 1984. *Inclusion Compounds*. London: Academic Press

Avakyan VG, Nazarov VB, Alfimov MV, Bagaturyants AA, Voronezhcheva NI, 2001. The role of intra- and intermolecular hydrogen bonds in the formation of  $\beta$ -CD head-to-head and head-to-tail dimers. The results of ab initio and semi-empirical quantum chemical calculations. *Russ Chem Bull*, 50 (2), 206

Avakyan VG, Nazarov VB, Voronezhcheva NI, 2005. DFT and PM3 Calculations of the Formation Enthalpies and Intramolecular H-Bond Energies in  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs. *Russ J Phys Ch*, 79 (1), S18

Avakyan VG, Nazarov, VB, Alfimov MV, Bagatur'yants AA, 1999. Structure of guest-host complexes of  $\beta$ -CD with arenes: a quantum-chemical study. *Russ Chem B*, 48 (10), 1833

Baboota S, Agarwal SP, 2005. Preparation and Characterisation of Meloxicam Hydroxy Propyl  $\beta$ -CD Inclusion Complex. *J Incl Phenom Macro*, 51, 219

Banerjee R, Chakraborty H, Sarkar M, 2003. Photophysical studies of oxamic group of NSAIDs: piroxicam, meloxicam and tenoxicam. *Spectrochim acta A*, 59, 1213

Banerjee R, Chakraborty H, Sarkar M, 2004. Host-Guest Complexation of Oxamic NSAIDs with  $\beta$ -CD. *Biopolymers*, 75, 355

Banerjee R, Sarkar M, 2002. Spectroscopic studies of microenvironment dictated structural forms of piroxicam and meloxicam. *J Lumin*, 99, 255

Bartlett RJ Ed., 1985. *Comparison of Ab Initio Quantum Chemistry with Experiment for Small Molecules*. Dordrecht: D. Reidel.

Becke AD, 1986. Density Functional Calculations of Molecular Bond Energies. *J Chem Phys*, 84, 4524

Becke AD, 1993a. A New Mixing of Hartree-Fock and Local Density-Functional Theories. *J Chem Phys*, 98, 1372

Becke AD, 1993b. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J Chem Phys*, 98, 5648

Behr JP, Ed, 1994. The Lock and Key Principle. The State of the Art-100 Years on. Chichester, New York: Wiley

Bender ML, Momiyama M, 1978. CD Chemistry. Springer, Berlin

Betzel C, Saenger W, Hingerty BE, Brown GM, 1984. Circular and Flip-Flop Hydrogen Bonding in  $\beta$ -CD Undecahydrate: A Neutron Diffraction Study. *J Am Chem Soc*, 106, 7545

Bodor N, Buchwald P, 2002. Theoretical Insights into the Formation, Structure, and Energetics of Some CD Complexes. *J Incl Phenom Macro*, 44, 9

Boys SF, Bernardi F, 1970. Calculation of small molecular interactions by differences of separate total energies – some procedures with reduced errors. *Mol Phys*, 19, 553

Brewster ME, Loftsson T, 2007. CDs as pharmaceutical solubilizers. *Adv Drug Deliver Rev*, 59, 645

Brickmann J, Exner TE, Keil M, Marhöfer RJ, 2000. Molecular Graphics – Trends and Perspectives. *J Mol Mod*, 6, 328

Buvari-Barcza A, Barcza L, 2000. Changes in the Solubility of  $\beta$ -CD on Complex Formation: Guest Enforced Solubility of  $\beta$ -CD Inclusion Complexes. *J Incl Phenom Macro*, 36, 355

Celotti F, Laufer S, 2001. Anti-Inflammatory Drugs: New Multitarget Compounds to face an old Problem. The Dual Inhibition Concept. *Pharmacol Res*, 43 (5), 429

Chacko KK, Saenger W, 1981. Topography of CD Inclusion Complexes. 15. Crystal and Molecular Structure of the Cyclohexaamylose-7.57 Water Complex, Form III. Four- and Six-Membered Circular Hydrogen Bonds. *J Am Chem Soc*, 103, 1708

Chakraborty H, Roy S, Sarkar M, 2005. Interaction of oxicam NSAIDs with DMPC vesicles: differential partitioning of drugs. *Chem Phys Lipids*, 138, 20

Challa R., Ahuja A, Ali J., Khar RK, 2005. CDs In Drug Delivery: An Updated Review. *AAPS PharmSciTech*, 6, 329

Charumanee S, Titwan A, Sirithunyalug J, Weiss-Greiler P, Wolschann P, Viernstein H, Okonogi S, 2006. Thermodynamics of the encapsulation by CDs. *J Chem Technol Biot*, 81, 523

Clark T, 1985. A Handbook of Computational Chemistry. New York: John Wiley & Sons.

Colman AW, Nicolis I, Keller N, Dalbiez JP, 1992. Aggregation of CDs-An Explanation of the Abnormal Solubility of  $\beta$ -CD. *J Incl Phenom Macro*, 13, 139

Connors, KA, 1997. The Stability of CD Complexes in Solution. *Chem. Rev*, 97, 1325

Cram DJ, 1988. The Design of Molecular Hosts, Guests, and Their Complexes (Nobel Lecture). *Angew Chem Int Edit*, 27, 1009

Cramer CJ, 2004. Essentials of Computational Chemistry. Theories and Models. 2<sup>nd</sup> ed. Chichester: John Wiley & Sons Ltd

Cramer F, Hettler H, 1967. Inclusion Compounds of CDs. *Die Naturwissenschaften*, 24, 625

Curtiss LA, Redfern PC, Raghavachari K, Pople, JA, 1998. Assessment of Gaussian-2 and Density Functional Theories for the Computation of Ionization Potentials and Electron Affinities. *J Chem Phys*, 109, 42

Davankov V, 1997. The Nature of Chiral Recognition: Is it a Three Point Interaction?. *Chirality*, 9, 99-102

Davies NM, Skjodt NM, 1999. Clinical Pharmacokinetics of Meloxicam. A Cyclo-Oxygenase-2 Preferential Nonsteroidal Anti-Inflammatory Drug. *Clin Pharmacokinet*, 36 (2), 115

Davis ME, Brewster ME, 2004. CD-based Pharmaceutics: Past, Present and Future. *Nat Rev Drug Discov*, 3, 1023

De Brauer C, Germain P, Merlin MP, Guerandel T, 2000. Thermal Behaviour of Anhydrous  $\alpha$ -,  $\beta$ - and  $\gamma$ -CD at Low Temperature. *J Incl Phenom Macro* 37, 75

Del Bene JE, Person WB, Szczepaniak K, 1995. Ab initio theoretical and matrix isolation experimental studies of hydrogen bonding: evidence of a dramatic effect of the matrix on the structure and vibrational spectrum of HBr:3,5-dichloropyridine. *J Phys Chem*, 99, 1172

Del Valle EMM, 2004. CDs and their uses: a review. *Process Biochem*, 39, 1033

Diot M, deBraver C, Germain P, 1998. Calorimetric Behaviour of Anhydrous  $\beta$ -CD at Very Low Temperature. *J Incl Phenom Macro*, 30, 143

Dobado JA, Benkadour N, Melchor S, Portal, D, 2004. Structure and theoretical NMR chemical shifts of modified CDs. *J Mol Struct*, 672, 127

Dodziuk H, 2002. Rigidity versus flexibility. A review of experimental and theoretical studies pertaining to the CD nonrigidity. *J Mol Struct*, 614, 33

Dodziuk H, 2006. CDs and their Complexes. Wiley-Vch, Weinheim

Dodziuk H, Kozminski W, Ejchart A, 2004. NMR Studies of Chiral Recognition by CDs. *Chirality*, 16, 90

- Duchene D, Ed, 1987. CDs and their Industrial Uses. Editions de Sante, Paris
- Duchene D., Ponchel G., Wouessidjewe D, 1999a. CDs in targeting Application to nanoparticles. *Adv Drug Deliver Rev*, 36, 29
- Duchene D., Wouessidjewe D., Ponchel G, 1999b. CDs and carrier systems. *J Control Release*, 62, 263
- Easton CJ, Lincoln SF, 1996. Chiral discrimination by modified CDs. *Chem Soc Rev*, 25, 163-170
- Easton CJ, Lincoln SF, 1999. Modified CDs: Scaffolds and Templates for Supramolecular Chemistry. London: Imperial College Press, p 43-100
- Endo T, Zheng M., Zimmermann W, 2002. Enzymatic Synthesis and Analysis of Large-Ring CDs. *Aust J Chem*, 55, 39
- Engeldinger E, Armspach D, Matt D, 2003. Capped CDs. *Chem Rev*, 103 (11), 4147
- Feller D, Peterson KA, 1998. An examination of intrinsic errors in electronic structure methods using the Environmental Molecular Sciences Laboratory computational results database and the Gaussian-2 set. *J Chem Phys*, 108 (1), 154
- Foresman JB, Frisch A, 1996. Exploring Chemistry with Electronic Structure Methods. 2<sup>nd</sup> ed. Pittsburgh: Gaussian, Inc.
- French D, 1957. The Schardinger dextrans. *Adv Carbohyd Chem*, 12, 189
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski JA, Montgomery JJA, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (2004) GAUSSIAN 03, Revision C.02, Gaussian Inc., Wallingford CT
- Frömring KH, Szejtli J, Ed, 1994. CDs in Pharmacy. Dordrecht: Kluwer Academic Publishers.
- Griswold DE, Adams JL, 1996. Constitutive Cyclooxygenase (COX-1) and Inducible Cyclooxygenase (COX-2): Rationale for Selective Inhibition and Progress to Date. *Med Res Rev*, 16 (2), 181
- Harata K, 1987. The Structure of the CD Complex. XX. Crystal Structure of Uncomplexed Hydrated  $\gamma$ -CD. *B Chem Soc Jpn*, 60, 2763

Harata K, 1998 Structural Aspects of Stereodifferentiation in the Solid State. Chem. Rev. 98, 1803

Hawkey CJ, 1999. COX-2 Inhibitors. Lancet, 353, 307

Hedges AR, 1998. Industrial Applications of CDs. Chem Rev, 98, 2035

Hebre WJ, 1995. Practical Strategies for Electronic Structure Calculations. Irvine: Wavefunction, Inc.

Hebre WJ, Radom L, Schleyer PV, Pople, JA, 1986. Ab Initio Molecular Orbital Theory. New York: John Wiley & Sons.

Higuchi T, Connors KA, 1965. Phase-solubility techniques. Adv Anal Chem Instr, 4, 212

Hinchliffe A, 2003. Molecular Modelling for Beginners. Chichester: John Wiley & Sons Ltd

Hingerty B, Klar B, Hardgrove GL, Betzel C, Saenger W, 1984. Neutron diffraction of  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs : hydrogen bonding patterns. J Biomol Struct Dyn, 2, 249

Hirayama F, Uekama K, 1999. CD-based controlled drug release system. Adv Drug Deliver Rev, 36, 125

Höltje H-D, Sippl W, Rognan D, Folkers G, 2003. Molecular Modelling. Basic Principles and Applications. 2<sup>nd</sup> ed. Weinheim: Wiley-VCH

Hohenberg P, Kohn W, 1964. Inhomogenous Electron Gas. Phys Rev, 136, B864

Houk KN, Leach AG, Kim SP, Zhang X, 2003. Binding Affinities of Host-Guest, Protein-Ligand, and Protein-Transition-State Complexes. Angew Chem Int Edit, 42, 4872

Imamura K, Nimz O, Jacob J, Myles D, Mason SA, Kitamura S, Aree T, Saenger W, 2001. Hydrogen-bond Network in Cyclodecaamylose Hydrate at 20 K; Neutron Diffraction Study of Novel Structural Motifs Band-Flip and Kink in  $\alpha$ -(1  $\rightarrow$  4)-D-Glucoside Oligosaccharides. Acta Crystallogr B, 57, 833

Irikura KK, 1998. Systematic Errors in Ab Initio Bond Dissociation Energies. J Phys Chem A, 102 (45), 9031

Jeffrey GA, 1997. An Introduction to Hydrogen Bonding. Oxford: University Press

Jeffrey GA, Saenger W, 1991. Hydrogen Bonding in Biological Structures. Berlin: Springer

Jensen F, 1999. Introduction to Computational Chemistry. New York: John Wiley & Sons.

Jimenez V, Alderete JB, 2008. Hartree-Fock and Density Functional Theory Study of  $\alpha$ -CD Conformers. *J Phys Chem A*, 112, 678

Jursic BS, Zdravkovski Z, French AD, 1996. Molecular modelling methodology of  $\beta$ -CD inclusion complexes. *J Mol Struct*, 366, 113

Jursic BS, 1999. Study of the water-methanol dimer with gaussian and complete basis set ab initio, and density functional theory methods. *J Mol Struct*, 466, 203

Kano K, 1997. Mechanisms for chiral recognition by CDs. *J Phys Org Chem*, 10, 286-291

Kano K, Nishiyabu R, 2002. General mechanism for chiral recognition by native and modified CDs. *J Incl Phenom Macro*, 44, 355-359

Karpfen A, 2002. Cooperative effects in hydrogen bonding. *Adv Chem Phys*, 123, 469

Karpfen A, Liedl E, Snor W, Wolschann P, 2007. Homodromic hydrogen bonds in low-energy conformations of single molecule CDs. *J Incl Phenom Macro* 57, 35

Karpfen A, Liedl L, Snor W, Viernstein H, Weiss-Greiler P, 2008. Density functional calculations on CDs. *Monatsh Chem* 139, 363

Khan AR., Forgo P, Stine KJ, Souza VT, 1998. Methods for Selective Modifications of CDs *Chem Rev*, 98, 1977

Klar B, Hingerty B, Saenger W, 1980. Topography of CD Inclusion Complexes. XII. Hydrogen Bonding in the Crystal Structure of  $\alpha$ -CD Hexahydrate: The Use of a Multicounter Detector in Neutron Diffraction. *Acta Crystallogr B*, 36, 1154

Klein CT, Polheim D, Viernstein H, Wolschann P, 2000. A Method for Predicting the Free Energies of Complexation between  $\beta$ -CD and guest molecules. *J Incl Phenom Macro*, 36, 409

Koch W, Holthausen MC, 2001. *A Chemist's Guide to Density Functional Theory*. 2<sup>nd</sup> edition. Weinheim: Wiley-VCH

Koehler G, Martinek H, Parasuk W, Rechthaler K, Wolschann P, 1995. The Influence of CD Complexation on Proton-Transfer in Piperidinomethyl-2-Naphthol. *Monatsh Chem*, 126, 299

Kohn W, Sham LJ, 1965. Self Consistent Equations Including Exchange and Correlation Effects. *Phys Rev*, 140, A1133

Koshland DE, 1994. The Key and Lock Theory and Induced-Fit Theory. *Angew Chem Int Ed*, 33, 2375

Kozar T, Venanzi CA, 1997. Reconsidering the conformational flexibility of  $\beta$ -CD. *J Mol Struct*, 395, 451

Kunz RW, 1997. *Molecular Modelling für Anwender*. 2<sup>nd</sup> ed. Stuttgart: Teubner

Kurumbail RG, Stevens AM, Gierse JK, McDonald JJ, Stegeman RA, Pak JY, Gildehaus D, Miyasiro JM, Penning TD, Seibert K, Isakson PC, Stallings WC, 1996. Structural basis for selective inhibition of cyclooxygenase-2 by anti-inflammatory agents. *Nature*, 348,644

Lamb ML, Jorgensen WL, 2004. Computational approaches to molecular recognition. *Curr Opin Chem Biol*, 1, 449

Larsen KL, 2002. Large CDs. *J Incl Phenom Macro*, 43, 1

Lawtrakul L, Viernstein H, Wolschann P, 2003. Molecular dynamics simulations of  $\beta$ -CD in aqueous solution. *Int. J. Pharm*, 256, 33

Leach AR, 2001. *Molecular Modelling. Principles and Applications*. 2<sup>nd</sup> ed. Dorchester: Dorset Press

Lebrilla CB, 2001. The Gas-Phase Chemistry of CD Inclusion Complexes. *Accounts Chem Res*, 34, 653

Lee C, Yang W, Parr RG, 1988. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys Rev B*, 37, 785

Lehn JM, 1988. Supramolecular Chemistry-Scope and Perspectives: Molecules, Supermolecules, and Molecular Devices (Nobel Lecture). *Angew Chem Int Ed*, 27, 89

Lehn JM, 1990. Perspectives in Supramolecular Chemistry-From Molecular Recognition towards Molecular Information-Processing and Self-Organization. *Angew Chem Int Ed*, 29 (11), 1304

Lehn JM, 2002. Toward Complex Matter: Supramolecular Chemistry and Self-Organization. *Proc Natl Acad Sci USA*, 99, 4736

Li XS, Liu L, Mu TW, Guo Q.X, 2000. A Systematic Quantum Chemistry Study on CDs. *Monatsh Chem*, 131, 849

Lichtenthaler FW, 1994. "100 Years Schlüssel-Schloss-Prinzip- What made Emil Fischer use this Analogy?". *Angew Chem Int Ed*, 33, 2364

Lichtenthaler F, Immel S, 1996a. On the Hydrophobic Characteristics of CDs: Computer-Aided Visualization of Molecular Lipophilicity Patterns. *Liebigs Ann*, 27

Lichtenthaler FW, Immel S, 1996b Towards Understanding Formation and Stability of CD Inclusion Complexes: Computation and Visualization of their Molecular Lipophilicity Patterns. *Starch*, 4, 145

Lindner K, Saenger W, 1982a. Crystal and Molecular Structure of Cyclohepta-amylose Dodecahydrate. *Carbohyd Res*, 99, 103

Lindner K, Saenger W, 1982b. Topography of CD Inclusion Complexes. XVI. Cyclic System of Hydrogen Bonds: Structure of  $\alpha$ -CD Hexahydrate, Form (II): Comparison with Form (I). *Acta Crystallogr*, 38, 203

- Lipkowitz KB, 1991. Symmetry Breaking in CDs: A MM Investigation. *J Org Chem*, 56, 6357
- Lipkowitz KB, 1998. Applications of Computational Chemistry to the Study of CDs. *Chem Rev*, 98, 1829
- Lipkowitz KB, Green K, Yang JA, 1992 Structural Characteristics of CDs in the Solid State. *Chirality*, 4, 205
- Liu L, Guo QX, 2002. The Driving Forces in the Inclusion Complexation of CDs. *J Inc Phenom. Macro*, 42, 1
- Liu L, Guo QX, 2004. Use of Quantum Chemical Methods to Study CD Chemistry. *J Incl Phenom Macro*, 50, 95
- Liu L, Li XS, Guo QX, Liu YC, 1999. Hartree-Fock and Density Functional Theory Studies on the Molecular Recognition of the CD. *Chinese Chem Lett*, 10 (12), 1053
- Loftsson T, Brewster ME, 1996. Pharmaceutical Applications of CDs. 1. Drug Solubilization and Stabilization. *J Pharm Sci*, 85(10), 1017
- Loftsson T, Duchene D, 2007. CDs and their pharmaceutical applications. *Int J Pharm*, 329, 1
- Loftsson T, Masson M, 2001. CDs in topical drug formulations: theory and practice. *Int J Pharm*, 225, 15
- Luger P, Daneck K, Engel W, Trummlitz G, Wagner K, 1996. Structure and physicochemical properties of meloxicam, a new NSAID. *Eur J Pharm Sci*, 4, 175
- Lynch BJ, Truhlar DG, 2003. Small representative benchmarks for thermochemical calculations. *J Phys Chem A*, 107 (42), 8996
- Manor PC, Saenger W, 1974. Topography of CD Inclusion Complexes. III. Crystal and Molecular Structure of Cyclohexaamylose Hexahydrate, the (H<sub>2</sub>O)<sub>2</sub> Inclusion Complex. *J Am Chem Soc*, 96 (11), 3630
- Mayer B, Marconi G, Klein C, Köhler G, Wolschann P, 1997. Structural Analysis of Host-Guest Systems. Methyl-substituted Phenols in  $\beta$ -CD. *J Incl Phenom Macro*, 29, 79
- Mayer I, Vibok A, 1997. BSSE-free second-order intermolecular perturbation theory. *Mol Phys*, 92, 503
- Møller C, Plesset MS, 1934. Note on an Approximation Treatment for Many Electron Systems. *Phys Rev*, 46, 618
- Naidoo KJ, Yu-Jen CJ, Jansson JLM, Widmalm G., Maliniak A, 2004. Molecular Properties Related to the Anomalous Solubility of  $\beta$ -CD. *J Phys Chem B*, 108, 4236

Naidu NB, Chowdary KPR, Murthy KVR, Satyanarayana V, Hayman AR, Becket G, 2004. Physicochemical characterization and dissolution properties of meloxicam-CD binary systems. *J Pharmaceut Biomed*, 35, 75

Pariot N, Edwards-Levy F, Andry MC, Levy MC, 2000. Cross-linked  $\beta$ -CD microcapsules: preparation and properties. *Int J Pharm*, 211, 19

Payne MC, Teter MP, Allan DC, 1990. Car-Parrinello methods. *J Chem Soc Faraday Trans*, 86, 1221

Pedersen CJ, 1988. The Discovery of Crown Ethers (Nobel Lecture). *Angew Chem Int Ed*, 27, 1021

Peng C, Ayala PY, Schlegel B, 1996. Using Redundant Internal Coordinates to Optimize Equilibrium Geometries and Transition States. *J Comput Chem*, 17 (1), 49

Pettersson I, Liljefors T, 1996. MM calculated conformational energies of organic molecules: a comparison of force fields. In: Lipkowitz KB, Boyd DB, eds. *Reviews in Computational Chemistry*, Vol. 9, New York: VCH Publishers, 167-189

Pinjari RV, Joshi KA, Gejji SP, 2006. Molecular Electrostatic Potentials and Hydrogen Bonding in  $\alpha$ -,  $\beta$ -, and  $\gamma$ -CDs. *J PhysChem A*, 110 (48), 13037

Pinjari RV, Joshi KA, Gejji SP, 2007. Theoretical Studies on Hydrogen Bonding, NMR Chemical Shifts and Electron Density Topography in  $\alpha$ ,  $\beta$  and  $\gamma$ -CD Conformers. *J Phys Chem A*, 111, 13583

Pitha J, Trinadha RC, 1990. Distribution of Substituents in 2-Hydroxypropyl Ethers of Cyclomaltoheptaose. *Carbohydr Res*, 200, 429

Plazanet M, Floare C, Johnson MR, Schweins R, Tromsdorff HP, 2004. Freezing on heating of liquid solutions. *J Chem Phys*, 121, 5031

Pople JA, 1965. Two Dimensional Chart of Quantum Chemistry. *J Chem Phys*, 43 (10), 229

Pople JA, Binkley JS, Seeger R, 1976. Theoretical Models Incorporating Electron Correlation. *Int J Quant Chem Symp*, 10, 1

Rajewski RA, Stella VJ, 1996. Pharmaceutical applications of CDs.2. In vivo drug delivery. *J Pharm Sci*, 85, 1142

Ramachandran KI, Deepa G, Namboori K, 2008. *Computational Chemistry and Molecular Modelling*. Berlin Heidelberg: Springer

Ramirez J, He F, Lebrilla CB, 1998. Gas-Phase Chiral Differentiation of Amino Acid Guests in CD Hosts. *J Am Chem Soc*, 120, 7387

Redenti E, Szente L, Szejtli J, 2001. CD Complexes of Salts of Acidic Drugs. Thermodynamic Properties, Structural Features, and Pharmaceutical Applications. *J Pharm Sci*, 90, 979

- Reinhold J, 2004. *Quantentheorie der Moleküle*. 2<sup>nd</sup> ed. Stuttgart: B.G. Teubner
- Rekharsky MV, Inoue Y, 1998. Complexation Thermodynamics of CDs. *Chem. Rev*, 98, 1875
- Ross PD, Rekharsky, MV, 1996 Thermodynamics of Hydrogen Bond and Hydrophobic Interactions in CD Complexes. *Biophys J*, 71, 2144
- Rudyak VY, Avakyan VG, Nazarov VB, Voronezheva NI, 2006. DFT calculation of  $\alpha$ -CD dimers. Contribution of hydrogen bonds to the energy of formation. *Russ Chem B+*, 55 (8), 1337
- Saenger W, 1979. Circular Hydrogen Bonds. *Nature*, 289, 343
- Saenger W, 1980. CD Inclusion Compounds in Research and Industry. *AngewChem Int Edit*, 19 (5), 344
- Saenger W, 1980. CD Inclusion Compounds in Research and Industry. *Angew Chem Int Ed*, 19 (5), 344
- Saenger W, 1984. Structural aspects of CDs and their inclusion complexes. In: Atwood JL, Davies JED, MacNicol DD, eds. *Inclusion compounds*. London: Academic Press, 231
- Saenger W, Betzel C, Hingerty B, Brown GM, 1982. Flip-flop hydrogen bonding in a partially disordered system. *Nature*, 296, 581
- Saenger W, Jacob J, Gessler K, Steiner T, Hoffmann D, Sanbe H, Koizumi K, Smith SM., Takaha T, 1998. Structures of the Common CDs and Their Larger Analogues Beyond the Doughnut. *Chem Rev*, 98, 1787
- Saenger W, Lindner K, 1980. OH Clusters with Homodromic Circular Arrangement of Hydrogen Bonds. *Angew Chem Int Ed*, 19 (5), 398
- Saenger W, Steiner T, 1998. CD Inclusion Complexes: Host-Guest Interactions and Hydrogen-Bonding Networks. *Acta Crystallogr A*, 798
- Scheiner AC, Baker J, Andzelm JW, 1997. Molecular Energies and Properties from Density Functional Theory: Exploring Basis Set Dependence of Kohn-Sham Equation using several Density Functionals. *J Comput Chem*, 18 (6), 775
- Scheiner S, 1997. *Hydrogen Bonding. A Theoretical Perspective*. Oxford: University Press
- Schlick T, 2002. *Molecular Modelling and Simulation*. New York: Springer
- Schneider HJ, Hacket F, Rüdiger V, Ikeda H, 1998. NMR Studies of CDs and CD Complexes. *Chem Rev*, 98, 1755
- Schneider HJ, Yatsimirsky AK, 2000. *Principles and Methods in Supramolecular Chemistry*. Chichester: John Wiley & Sons

Schuster P, Wolschann P, 1999. Hydrogen Bonding: From Small Clusters to Biopolymers. *Monatsh Chem*, 130, 947

Shahgaldian P, Pieles U, 2006. CD Derivatives as Chiral Supramolecular Receptors for Enantioselective Sensing. *Sensors*, 6, 593-615

Singh M, Sharma R, Banerjee UC, 2002. Biotechnological applications of CDs. *Biotechnol Adv*, 20, 341

Snor W, Liedl E, Weiss-Greiler P, Karpfen A, Viernstein H, Wolschann P, 2007. On the structure of anhydrous  $\beta$ -CD. *Chem Phys Lett*, 441, 159

Steed JW, Atwood JL, 2000. *Supramolecular Chemistry*. Chichester: John Wiley & Sons

Steiner T, Koellner G, 1994. Crystalline  $\beta$ -CD Hydrate at Various Humidities: Fast, Continuous and Reversible Dehydration Studied by X-Ray Diffraction. *J Am Chem Soc*, 116, 5122

Steiner T, Mason SA, Saenger W, 1991. Disordered Guest and Water Molecules. Three-Center and Flip-Flop  $O-H\cdots O$  Hydrogen Bonds in Crystalline  $\beta$ -CD Ethanol Octahydrate at T=295K: A Neutron and X-ray Diffraction Study. *J Am Chem Soc* 113, 5676

Steiner T, Saenger W, 1994. Reliability of assigning H...O hydrogen bonds to short intermolecular O...O separations in CD and oligosaccharide crystal structures. *Carbohydr Res*, 259, 1

Steiner T, Saenger W, 1996. Crystal structure of anhydrous hexakis-(2,3,6-tri-O-methyl)cyclomaltohexaose (permethyl- $\alpha$ -CD) grown from hot water and from cold NaCl solutions. *Carbohydr Res*, 282, 53

Stella VJ, Rajewski RA, 1997. CDs: their future in drug formulation and delivery. *Pharm Res*, 14, 556

Stella VJ, Rao VM, Zannou EA, Zia V, 1999. Mechanism of drug release from CD complexes. *Adv. Drug Deliver. Rev*, 36, 3

Szejtli J, 1982. *CDs and their Inclusion Complexes*. Budapest: Akademiai Kiado.

Szejtli J, 1998. Introduction and General Overview of CD Chemistry. *Chem Rev*, 98, 1743

Szejtli J, 2004. Past, present, and future of CD research. *Pure Appl Chem*, 76, 1825

Szejtli J, Ed, 1988. *CD Technology*. Dordrecht: Kluwer Academic Publishers.

Szejtli J, Ed, 1996. *Comprehensive Supramolecular Chemistry, Vol.3: CDs*. Oxford: Pergamon.

Szente L, Szejtli J, 1999. Highly soluble CD derivatives: chemistry, properties, and trends in development. *Adv Drug Deliver Rev*, 36, 17

Taira H, Nagase H, Endo T, Ueda H, 2006. Isolation, Purification and Characterization of Large-Ring CDs (CD<sub>36</sub>-CD<sub>39</sub>). *J Incl Phenom Macro*, 56, 23

Tsai RS, Carrupt PA, El Tayar N, Giroud Y, Anrade P, Testa B, Bree F, Tillement JP, 1993. Physicochemical and Structural Properties of Nonsteroidal Anti-inflammatory Oxicams. *Helv Chim Acta*, 76, 842

Uekama K, 2004. Design and Evaluation of CD-Based Drug Formulation. *ChemPharmBull*, 52, 900

Uekama K, Hirayama F, Arima H, 2006. Recent aspect of CD-based drug delivery system. *J Incl Phenom Macro*, 56, 3

Vane JR, 1971. Inhibition of prostaglandin synthesis as a mechanism of action for aspirin-like drugs. *Nature-New Biol*, 231, 232

Vane JR, Botting RM, 1998. Anti-inflammatory drugs and their mechanism of action. *Inflamm Res*, 47 (Supplement2 ), S78

Viernstein H, Weiss-Greiler P, Wolschann P, 2002. Solubility Enhancement of Low Soluble Biologically Active Compounds by  $\beta$ -CD and Dimethyl- $\beta$ -CD. *J Incl Phenom Macro*, 44, 235

Viernstein H, Weiss-Greiler P, Wolschann P, 2003. Solubility enhancement of low soluble biologically active compounds-temperature and cosolvent dependent inclusion complexation. *Int J Pharm*, 256, 85

Vreven T, Morokuma K, 2000. The ONIOM (our own N-layered integrated molecular orbital+MM) method for the first singlet excited ( $S_1$ ) state photoisomerization path of a retinal protonated Schiff base. *J Chem Phys*, 113 (8), 2969

Vyas A, Saraf S, Saraf S, 2008. CD based novel drug delivery systems. *J Incl Phenom Macro*, 62, 23

Weinzinger P, Weiss-Greiler P, Snor W, Viernstein H, Wolschann P, 2007. Molecular dynamics simulations and quantum chemical calculations on  $\beta$ -CD-spirolactone complex. *J Incl Phenom Macro*, 57, 29

Wenz G, 2000. An Overview of Host-Guest Chemistry and its Application to Nonsteroidal Anti-Inflammatory Drugs. *Clin Drug Invest*, 19 (2), 21

Young DC, 2001. Computational Chemistry. A Practical Guide for Applying Techniques to Real-World Problems. New York: Wiley Interscience

Zabel V, Saenger W, Mason SA, 1986. Neutron Diffraction Study of the Hydrogen Bonding in  $\beta$ -CD Undecahydrate at 129 K: From Dynamic Flip-Flops to Static Homodromic Chains. *J Am Chem Soc*, 108, 3664

Zhang XX, Bradshaw JS, Izatt RM, 1997. Enantiomeric recognition of amine compounds by chiral macrocyclic receptors. *Chem Rev*, 97, 3313-3361

Zheng M, Endo T, Zimmermann W, 2002. Synthesis of Large-Ring CDs by CD Glucanotransferases from Bacterial Isolates. *J Incl Phenom Macro*, 44, 387

## 15 Appendix A

### Chemical Structure Drawing Software - Molecule Editors and Viewers

A chemical structure-drawing package is more than a simple, conventional drawing package, such as CorelDraw or MS Paint. In fact, molecule editors are often used to create professional-looking structure diagrams or even chemical reaction equations for publication. Furthermore, a structure-drawing program should incorporate the possibility of deducing additional information on the compound from the drawing in order to make the molecule amenable to processing in computational chemistry. Therefore, the electronic structure input and the further use of the structure or structure data play a dominant role in the field of molecular modelling.

#### 15.1 Molecule editors

Generally, molecule editors display the structure diagrams as 2D images. Stereochemical information such as R/S identity can be visualized in the software by using wedged or hashed bonds, which are well known to chemists. Nevertheless, the structure still is only two-dimensional. If the 3D geometry of the molecule (the 3D coordinates of the atoms) is required, it is necessary to send the 2D structure from the editor to a 3D structure generator. Some editors already provide this 3D feature within the software. Of course simple, appropriate drawing tools, such as different bond types (simple, double, etc.) and element symbols are available in chemistry and text settings in the editors discussed here.

The most important feature of editing software is the option to save the structure in standard file formats, which contain information about the structure (e.g., Mol-file, PDB-file). Most of these file formats are ASCII text files (which can be viewed in simple text editors) and cover international standardized and normalized specifications of the molecule, such as atom and bond types or connectivities. Thus, with these files, the structure can be exchanged between different programs. Furthermore, they can serve as input files to other chemical software, e.g., to calculate 3D structures or molecular properties.

##### 15.1.1 ChemDraw

ChemDraw Ultra is among the most popular commercial chemical drawing software. It is available as a separate program or integrated into the commercial software suite ChemOffice from CambridgeSoft Inc. with Chem3D (3D molecule viewer, modelling software), ChemFinder (database manager), ChemInfo (chemical databases and catalogs), and E-Lab Notebook (organizer). The drawing software comprises a comprehensive collection of standard tools to sketch 2D chemical structures, but there are also some other features that are very useful for chemists. One of these features enables the prediction of  $^1H$  and  $^{13}C$  NMR shifts from

structures and the correlation of atoms with NMR peaks. IUPAC standard names can be generated from chemical structures and on the other hand, for the most substances a structure can be created by typing in a systematic chemical name. Stereocenters can be identified using Cahn-Ingold-Prelog rules. Additional physical properties such as boiling point, melting point, etc., of chemical compounds can also be computed.

### 15.1.2 ISIS/Draw

ISIS/Draw from MDL Information Systems, Inc. is free of charge for non-profit use. The powerful structure editor has drawing tools for 2D molecular structures, reaction equations, and publishing graphics (journal settings). Structures can easily be rotated with the “3D rotate tool”. The graphical user interface provides the ability to create chemical structures intuitively. Besides many templates (pre-defined structures), the program incorporates additional features that automatically display hydrogen atoms at free valences or check properties (bond order, etc.) of the chemical structure with the “Chem Inspector”. Some additional molecular parameters can be calculated, e.g., molecular weight. Structure and reaction databases can be generated easily with the help of this structure editor.

## 15.2 Molecule Viewers

Molecule editors represent only two-dimensional chemical structures (thus also could be considered as 2D viewers); the third dimension is visualized by 3D viewers, mainly user-interactive. In order to represent 3D molecular models it is necessary to supply structure files with 3D information (e.g., pdb, xyz, mol, etc.). If structures from a structure editor are used directly, the files do not normally include 3D data. Inclusion of such data can be achieved only via 3D structure generators, force field calculations, etc. 3D structures can then be represented in various display modes, e.g., wire frame, balls and sticks, space-filling spheres, etc. Proteins are visualized by various representations of helices,  $\beta$ -strains, or tertiary structures. An additional feature is the ability to colour the atoms according to subunits, chain types, etc. During all such operations the molecule can be interactively moved, rotated, or zoomed by the user.

### 15.2.1 Discovery Studio Viewer Pro (formerly WebLabViewer)

The DS Viewer Pro is an innovative software tool for examining the 3D structure of molecular models, and for communicating the resulting information with other scientists. It uses a native file format that is capable of describing molecular or crystal structures along with surfaces and labels. It also reads a dozen different common molecular file formats. Two-dimensional structures can be automatically

converted to three-dimensions when the file is read. Work can be saved in common molecule file formats, VRML, SMILES, and bit-mapped graphic files.

DS Viewer Pro gives a very-high-quality display suitable for publication and presentation. Molecules can be displayed as lines, sticks, ball and stick, CPK, and polyhedrons. In addition, different atoms within the same structure may be displayed in different ways. Text can be added to the display as well as labelling parts of the structure in a variety of ways. The user has control over colours, radii, and display quality. The program can also replicate a unit cell to display a crystal structure. Several types of molecular surfaces can be displayed.

DS Viewer Pro can also be used to build molecular structures. It has building modes for adding individual atoms, creating chains, and creating rings. All these create structures of carbon atoms. Once the backbone has been created, atoms can be changed to other elements and hydrogen atoms added. There is a function to clean up the shape of the molecule, which does a basic MM minimization using a simple Dreiding-type force field. DS Viewer Pro can be used to create animations and has a scripting language to automate tasks.

### 15.2.2 GaussView

GaussView is a graphic interface for use with the Gaussian ab initio program. It can be used to build molecules, set up the options in the input file, run a calculation, and display results. The program has several building modes. Compounds can be built one atom at a time by selecting the element and hybridization. There are also libraries of ring systems, amino acids, nucleosides, and common organic functional groups. The user can manually set bond lengths, angles, and dihedral angles. A clean function gives an initial optimization of the structure using a rule-based VSEPR algorithm. A Z-matrix editor gives some control over how the Z-matrix is constructed, but does not go as far as giving the user the ability to enforce symmetry constraints.

GaussView can also be used to set up ONIOM QM/MM calculations. The graphic molecule-building functions are very easy to use. The molecular structures are rendered with good-quality shading on a blue background. Isosurfaces produced from cube files or checkpoint files can also be displayed. Molecular vibrations can be animated on screen and vibrational displacement vectors displayed. The vibrational line spectrum may be displayed too, but the user has no control over the axes. There is no way to set the background colour. The display can be saved using several image file formats.

### 15.2.3 HyperChem

HyperChem is an integrated graphic interface, computational, and visualization package. It has seen the most use on PCs. HyperChem incorporates ab initio, semi-empirical, and MM programs. These can be used for computing vibrational

frequencies, transition states, electronic excited states, and QM/MM, MD, and MC simulations. The program has a drawing mode in which the backbone can be sketched out and then hydrogen atoms added automatically. This sketcher does not set the bond lengths or angles, so the use of a MM optimization before doing more time-consuming calculations is highly advised. Building biomolecules is made easier with a sugar builder and amino-acid sequence editor. Periodic systems can be constructed with a crystal builder and a polymer builder, which are very easy to use.

The graphic interface incorporates a variety of rendering modes. It is possible to visualize molecular surfaces and animations of vibrational modes. Both electronic and vibrational spectra can be displayed with intensities. The program can produce good-quality graphics, including ray-traced renderings, suitable for publication. The GUI is integrated tightly with the computational modules, thus changing settings in one menu and the options available in other menus.

A number of common-structure file formats can be read and written. By default, the calculation results are displayed on screen, but they are not saved to disk. The user can specify that all results for a given session be written to a log file. While a calculation is running, no actions can be taken other than changing the molecular orientation on screen. By default, calculations are run on the PC that the GUI is running on. The program has a scripting ability that can be used to automate tasks. The built-in scripting allows the automation of menu selections and execution of jobs.

The MM force fields available include MM+, OPLS, BIO+, and AMBER. Parameters missing from the force field will be automatically estimated. The user has some control over cut-off distances for various terms in the energy expression. Solvent molecules can be included along with periodic boundary conditions. Biomolecule computational abilities are aided by functions for superimposing molecules, conformation searching, and QSAR descriptor calculation.

The semi-empirical techniques available include, among others, CNDO, INDO, MINDO/3, ZINDO, MNDO, AM1, and PM3. The ab initio module can run HF, MP2 (single point), and CIS calculations. A number of common basis sets are included. Some results, such as population analysis, are only written to the log file.

#### 15.2.4 Chem3D

Chem3D is a molecular modelling package for the PC and Macintosh. It can perform calculations using MM2 and extended Hückel as well as acting as a graphic interface, e.g., for Gaussian. There are also browser plug-ins available for viewing structures and surfaces. Chem3D can read a wide variety of popular chemical structure files, including Gaussian, MDL, MOPAC, PDB and SYBYL. Two-dimensional structures imported from ChemDraw or ISIS/Draw are automatically converted to three-dimensional structures. The Chem3D native file format contains

both the molecular structure and results of computations. Data can be exported in a variety of chemical-structure formats and graphics files.

Chem3D has both graphic and text-based structure-building modes. Structures can be generated graphically by sketching out the molecule. The builder creates carbon atoms, which can be edited by typing text to substitute other elements or functional groups. As the structure is built, the valence is filled with hydrogen atoms and typical bond lengths and angles are set. Several hundred predefined functional groups are available and users can define additional ones. The text-based mode allows the user to input a simple text string (similar to SMILES, but not identical). This text mode can be used to build structures entirely or to add functional groups.

A number of mechanisms are available for manually defining aspects of the molecular geometry. These include defining dummy atoms as well as setting bond lengths, angles, and dihedral angles. It is also possible to set distances between nonbonded atoms. The molecular structure is maintained internally in both Cartesian coordinates and a Z-matrix. A number of functions for defining how the Z-matrix is constructed make this one of the best GUIs available for setting up calculations that must be done by Z-matrix.

Chem3D uses a MM2 force field that has been extended to cover the full periodic table with the exception of the *f* block elements. The program will estimate unknown parameters and a message generated to inform the user of this. MM2 can be used for both energy minimization and MD calculations. The user can add custom atom types or alter the parameters used for one specific atom in the calculation. Extended Hückel may be used for the calculation of charges and molecular surfaces.

A number of properties can be computed from various chemical descriptors. These include physical properties, such as surface area, volume, molecular weight, and moments of inertia. Chemical properties available include boiling point, melting point, critical variables, Henry's law constant, heat capacity, log P, refractivity, and solubility. Several display modes are available. Molecules can be displayed as wire frames (lines), sticks (wider lines), ball and stick models (with line or cylindrical bonds), and as space-filling models. Protein structures can be displayed as ribbons. Dot surfaces of van den Waals radii or extended Hückel charges may be added to any of these. When molecular surfaces from extended Hückel, MOPAC, or Gaussian calculations are displayed, a different set of rendering algorithms with improved three-dimensional shading is used. These surfaces can be displayed as solid, mesh, dots, or translucent surfaces. The graphics quality in this display mode is very good. Movies can be created from operations generating multiple structures, such as MD simulations. These movies can be viewed within Chem3D, but cannot be saved in a common movie file format.

## 15.3 Ab initio and DFT Software

Some of these software packages also have semi-empirical or MM functionality. There are also ab initio programs bundled with products discussed previously in this appendix, e.g., Hyperchem.

### 15.3.1 Gaussian

Gaussian is a monolithic ab initio program. Gaussian probably incorporates the widest range of functionality of any ab initio program. It does include a few semi-empirical and MM methods that can be used alone or as part of QM/MM calculations. It uses one of the simplest ASCII input file formats. There are also many graphic interfaces available for creating Gaussian files and viewing results, such as GaussView or Chem3D.

Gaussian contains a wide range of ab initio functionality, such as HF, ROHF, MPn, CI, CC, QCL MCSCF, CBS, and G2. A number of basis sets and pseudo potentials are available. It also supports a large number of DFT functionals. Semi-empirical methods available include AM1, PM3, and ZINDO (single point only). MM methods are Amber, Dreiding, and UFF.

There are a wide range of molecular properties that can be computed, such as NMR chemical shifts, nonlinear optical properties, several population analysis schemes, vibrational frequencies, and intensities and data for use in visualization programs. QM/MM calculations can be performed using the ONIOM method. Transition structures and intrinsic reaction coordinates may also be computed. It is additionally possible to manually specify which sections of code are to be called and in what order.

Gaussian has one of the ASCII input formats most convenient to use without a graphic interface. Even though graphic interfaces are available, many researchers still construct input files manually due to the amount of control this gives them over the choice of computation method and molecular geometry constraints. There are a large number of options for controlling how the algorithms are executed. There are also a variety of options that allow the user to make efficient use of the hardware configuration, such as in core, direct, and semi direct integral evaluation. In addition, Gaussian can take advantage of parallel architectures. The program output contains a large amount of information. Gaussian has seen the widest use mainly in modelling organic molecules. The program is designed to execute as a batch job. It can readily be used with common batch-queuing systems.

### 15.3.2 HyperChem

See Chapter 15.2.3.

## 16 List of publications and contribution to conferences

### Publications

Karpfen A, Liedl E, Snor W, Wolschann P, 2007. Homodromic hydrogen bonds in low-energy conformations of single molecule CDs. *J Incl Phenom Macro* 57, 35

Snor W, Liedl E, Weiss-Greiler P, Karpfen A, Viernstein H, Wolschann P, 2007. On the structure of anhydrous  $\beta$ -CD. *Chem Phys Lett*, 441, 159

Weinzinger P, Weiss-Greiler P, Snor W, Viernstein H, Wolschann P, 2007. Molecular dynamics simulations and quantum chemical calculations on  $\beta$ -CD-spirolactone complex. *J Incl Phenom Macro*, 57, 29

Karpfen A, Liedl L, Snor W, Viernstein H, Weiss-Greiler P, 2008. Density functional calculations on CDs. *Monatsh Chem* 139, 363

Snor W, Liedl E, Weiss-Greiler P, Viernstein H, Wolschann P, 2009. Density functional calculations on meloxicam- $\beta$ -cyclodextrin inclusion complexes, *International Journal of Pharmaceutics* (2008), doi:10.1016/j.ijpharm.2009.05.012.

### Poster presentations and contributions to proceeding papers at scientific conferences

05/2006 **Poster presentation: “Quantum chemical calculations of symmetric conformations of anhydrous  $\beta$ -cyclodextrin”**

13<sup>th</sup> International Cyclodextrins Symposium, Turin, Italy

**05/2006 Poster presentation: “Spironolactone-Physicochemical Aspects of the Complexation with Cyclodextrin”**

13<sup>th</sup> International Cyclodextrins Symposium, Turin, Italy

**04/2006 Poster presentation:”Inclusion complexation of various tautomers and protonation states of Meloxicam”**

19<sup>th</sup> Scientific Meeting of the Austrian Pharmaceutical Organisation, Innsbruck, Austria

**04/2006 Poster presentation:”Molecular Calculations on Cyclodextrin and Cyclodextrin Inclusion Complexes”**

19<sup>th</sup> Scientific Meeting of the Austrian Pharmaceutical Organisation, Innsbruck, Austria

**09/2007 Poster presentation: “Molecular Calculations on Hydrogen Bonds in Cyclodextrins”**

15<sup>th</sup> International Workshop on Bioencapsulation, Vienna, Austria

**09/2007 Poster presentation: “Density Functional Calculations on Meloxicam- $\beta$ -Cyclodextrin Inclusion Complexes”**

15<sup>th</sup> International Workshop on Bioencapsulation, Vienna, Austria

**09/2007 Contribution to proceeding paper: “Molecular Calculations on Hydrogen Bonds in Cyclodextrins”**

15<sup>th</sup> International Workshop on Bioencapsulation, Vienna, Austria

**09/2007 Contribution to proceeding paper: “Density Functional Calculations on Meloxicam- $\beta$ -Cyclodextrin Inclusion Complexes”**

15<sup>th</sup> International Workshop on Bioencapsulation, Vienna, Austria

**05/2008 Poster presentation: “DFT Calculations on Meloxicam- $\beta$ -Cyclodextrin Inclusion Complexes”**

14<sup>th</sup> International Cyclodextrins Symposium, Kyoto, Japan

## 17 Curriculum Vitae

### WALTERSNOR

Institute for Theoretical Chemistry,  
University of Vienna,  
Währinger Straße 17, 1090 Vienna, Austria  
Tel: 06766319080  
walter.snor@kstp.at

#### Personal data

---

Date of Birth: 02.04.1945, Dörfl/Kasten

Citizenship: Austrian

#### Education and professional experience

---

Since 10/1964:

Study of chemistry and physics at University of Vienna

Since 09/1972:

Teaching at Bundesgymnasium und Bundesrealgymnasium St.Pölten

1977-1978:

Diploma thesis: Institute for Analytical Chemistry, University of Vienna

“Fortschritte auf dem Gebiete ionenselektiver Elektroden”

10/1978:

Diploma “Magister der Naturwissenschaften in der Studienrichtung Chemie,  
Nebenfach Physik; Studienzweig für das Lehramt an höheren Schulen”

09/2004 retired as teacher

10.04 – current

Doctoral thesis: Institute for Theoretical Chemistry, University of Vienna

“Molecular Modelling on Cyclodextrin Inclusion Complexes”