Johannes Hoja, BSc

A Theoretical Study of Intermolecular Interactions in Hydrogen-Bonded Systems

MASTERARBEIT

zur Erlangung des akademischen Grades

Master of Science

Masterstudium Chemie

eingereicht an der

Technischen Universität Graz

Betreuer

Ao.Univ.-Prof. Dr.phil. Alexander F. Sax

Institut für Chemie
Karl-Franzens-Universität Graz

Graz, Juli 2014

EIDESSTATTLICHE ERKLÄRUNG

AFFIDAVIT

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Das in TUGRAZonline hochgeladene Textdokument ist mit der vorliegenden Masterarbeit identisch.

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis.

Datum / Date

Unterschrift / Signature

Publications

This master thesis is based on the following publications:

- J. Hoja, R. J. Maurer, A. F. Sax, Adsorption of Glucose, Cellobiose, and Cellotetraose onto Cellulose Model Surfaces, *J. Phys. Chem. B* **2014**, in press, DOI: 10.1021/jp5025685.

- J. Hoja, A. F. Sax, K. Szalewicz, Is Electrostatics Sufficient to Describe Hydrogen-Bonding Interactions?, *Chem. - Eur. J.* **2014**, *20*, 2292–2300.

Abstract

In this thesis, intermolecular interactions are studied for several hydrogen-bonded systems. There are four physical interaction types, which contribute to the total interaction energy between two interacting molecules: electrostatics, exchange, induction, and dispersion. At first the adsorption process of carbohydrates on cellulose model surfaces is studied. It is a popular belief that the interaction between the adsorbate and the cellulose surface is governed by hydrogen bonding, while ubiquitous dispersion interactions are regularly neglected. The adsorption of a glucose molecule onto a hydrophobic and a hydrophilic cellulose surface is studied with the GLYCAM06 force field and density-functional theory (DFT) with and without dispersion correction. It is shown that dispersion interactions are already important for the small glucose adsorbate and the use of the GLYCAM06 force field for larger adsorbates is justified by comparison with dispersion-corrected DFT. Next, the adsorption of cellobiose and cellotetraose is studied by using GLYCAM06. No direct correlation between the adsorption energy of the adsorbate/surface complex and the number of hydrogen bonds can be found. Therefore, the nature of the interaction between hydrogen-bonded systems is further studied by using symmetry-adapted perturbation theory. The interactions in the water dimer and several alcohol dimers are discussed. Against popular belief that hydrogen bonding is solely an electrostatic interaction it is shown that especially the exchange and dispersion interactions are also important. In large and bulky systems the dispersion plays a very crucial role in the stabilization of these dimers. For cellulose this means that dispersion interactions are of utmost importance for the adsorption process, even for polar molecules, and also for the interactions within cellulose bulk.

Kurzfassung

Es werden intermolekulare Wechselwirkungen in Systemen untersucht, die durch Wasserstoffbrücken (H-Brücken) gebunden sind. Es gibt vier physikalische Beiträge zur Wechselwirkungsenergie zwischen zwei Molekülen: Elektrostatik, Induktion, Austausch und Dispersion. Zuerst wird der Adsorptionsprozess von Kohlenhydraten auf Celluloseoberflächen untersucht. Es ist eine verbreitete Ansicht, dass die Wechselwirkung zwischen dem Adsorbat und der Celluloseoberfläche nur durch H-Brücken bestimmt wird, während die allgegenwärtige Dispersionswechselwirkung häufig vernachlässigt wird. Die Adsorption eines Glucose-Moleküls auf einer hydrophilen und hydrophoben Celluloseoberfläche wird mit dem Kraftfeld GLYCAM06 und Dichtefunktionaltheorie (DFT), sowohl mit als auch ohne Dispersionskorrektur, untersucht. Es wird gezeigt, dass bereits für das kleine Glucose-Adsorbat Dispersionswechselwirkungen wichtig sind. Weiters wird gezeigt, dass das GLYCAM06 Kraftfeld in der Lage ist, für diese Systeme sinnvolle Ergebnisse zu liefern und daher für die Untersuchung von größeren Adsorbaten verwendet werden kann. Bei der Untersuchung der Adsorbate Cellobiose und Cellotetraose kann keine direkte Korrelation zwischen der Adsorptionsenergie und der Anzahl an H-Brücken gefunden werden. Aufgrund dieser Ergebnisse wird in Folge die Art der Wechselwirkungen zwischen Systemen mit H-Brücken näher untersucht. Dazu wird symmetrieadaptierte Störungstheorie verwendet. Es werden die Wechselwirkungen im Wasserdimer und verschiedenen Alkoholdimeren besprochen. Entgegen der generellen Meinung, dass H-Brücken eine rein elektrostatische Wechselwirkung hervorrufen, wird gezeigt, dass die Austauschwechselwirkung und die Dispersion für diese Systeme auch sehr wichtig sind. Daher spielt die Dispersionswechselwirkung eine entscheidende Rolle in Adsorptionsprozessen auf Celluloseoberflächen, sogar wenn die Adsorbate polare Moleküle sind.

Acknowledgements

First, I want to thank my supervisor Prof. Alexander F. Sax for his excellent support and the pleasant working atmosphere in the computational chemistry research group.

Many thanks to Reinhard Maurer, who fortunately persuaded me to become a theoretical chemist. Thanks also for the many provocative discussions over the last years.

I also want to express my gratitude to Prof. Krzysztof Szalewicz for his support regarding SAPT.

Financial support by the EU Seventh Framework Program project SURFUNCELL is gratefully acknowledged.

Many thanks also to Prof. Volker Ribitsch, who gave me the opportunity to work in the SURFUNCELL project.

Finally, I want to express my deepest gratitude to my parents for their encouragement and their support.

Contents

Chapter 1

Introduction

In this thesis intermolecular interactions in hydrogen-bonded systems are discussed using computational chemistry methods. At first, the focus is on carbohydrate adsorption onto cellulose surfaces, in which hydrogen bonds are formed between the adsorbate and the surface. In the second part of this thesis, intermolecular interactions in the water dimer and several alcohol dimers are thoroughly investigated. It is shown that dispersion interactions can be very important and that electrostatic interactions play a much smaller role in the presented systems than it is generally claimed.

This chapter provides background information on a basic level regarding several topics essential for the understanding of the presented research. Therefore, the basics of intermolecular interactions are presented followed by an introduction to hydrogen bonding. Furthermore, the structure of cellulose is explained and recent work regarding the stability of cellulose and the adsorption onto cellulose surfaces is discussed. Finally, an outline of this thesis is given.

The intermolecular interactions section is based on the books of Kaplan [1], Maitland et al. [2], and Stone. [3] The section about hydrogen bonding follows the monograph of Jeffrey. [4]

1.1 Intermolecular Interactions

Intermolecular interactions are interactions between two molecules that are not covalently bound. Compared to a covalent bond, with bond strengths ranging from about 150 to 1000 $\mathrm{kJ\,mol^{-1}}$, intermolecular interactions are at least one order of magnitude smaller. This makes an accurate description challenging for various theoretical methods. Intermolecular interactions between two molecules arise because of a number of physical phenomena. In this thesis only interactions between closed-shell molecules in the ground state are discussed. For those molecules four components contribute to the total interaction energy: electrostatics, induction, dispersion, and exchange. These four components are illustrated in Figure 1.1 and will be discussed below. The first three terms are often referred to as *long-range interactions*, while the exchange is a *short-range interaction*. That means that the exchange interaction decays much faster with increasing distance between the molecules than any long-range interaction; the exchange decreases exponentially, while all long-range interactions are proportional to an inverse power of the intermolecular separation.

For distant molecules, where there is no overlap of the electron clouds, the total interaction between the molecules can be described by multipole-multipole interactions of the long-range components because the exchange is already negligible. Multipole-multipole interaction means that all possible interactions between the different multipoles of both monomers must be considered. A multipole expansion consists of the following multipole moments: monopole or charge (Q), dipole (μ), quadrupole (Θ), octopole, hexadecapole, etc. Therefore, the pair-potential V of a multipole-multipole interaction truncated at the quadrupole level can be written as

$$V = V_{QQ'} + V_{Q\mu} + V_{Q\Theta} + V_{\mu\mu'} + V_{\mu\Theta} + V_{\Theta\Theta'}. \tag{1.1}$$

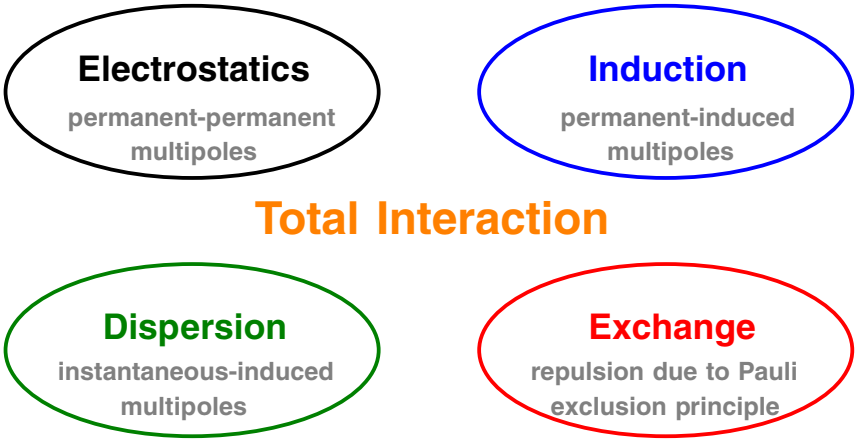


Figure 1.1: Illustration of the contributions to the total interaction energy between closed-shell molecules.

In this thesis only interactions between two uncharged molecules are discussed. Therefore, we can omit the first three terms in Equation 1.1, so the simplest term is now the dipole-dipole interaction $V_{\mu\mu'}$.

Now follows the discussion of the four previously mentioned components of the total interaction energy, starting with the three long-range components:

Electrostatics. It is the interaction between permanent multipoles of the two molecules or in other words, the interaction between static charge distributions. The electrostatic interaction strongly depends on the orientation of the molecules and can be either attractive or repulsive. In addition these interactions are strictly pairwise additive, so electrostatics can be perfectly described by using only pair-potentials. This will not be the case for all remaining components.

The interaction between two permanent dipole moments μ and μ' with fixed orientations at a separation r in vacuum can be expressed as

$$V_{\mu\mu'}^{\text{elst}}(r) = -\frac{1}{4\pi\epsilon_0}\frac{\mu\mu'}{r^3}\left(2\cos\theta\cos\theta' - \sin\theta\sin\theta'\cos\vartheta\right), \tag{1.2}$$

with ϵ_0 being the vacuum permittivity. The definitions of the angles θ, θ', and ϑ can be inferred from Figure 1.2. If these dipole moments are able to rotate freely, the interaction energy can be obtained by averaging over all orientations and weighting with a Boltzmann factor according to

$$V_{\mu\mu'}^{\text{elst}}(r) = -\frac{2}{3k_BT}\frac{1}{(4\pi\epsilon_0)^2}\frac{\mu^2\mu'^2}{r^6}, \tag{1.3}$$

where k_B is the Boltzmann constant and T the temperature. Now $V_{\mu\mu'}^{\text{elst}} \propto r^{-6}$ and always attractive, while for two permanent dipoles with fixed orientations $V_{\mu\mu'}^{\text{elst}} \propto r^{-3}$ and can be of either sign. For a detailed discussion including also dipole-quadrupole and quadrupole-quadrupole interactions the reader is referred to References 1–3.

Induction. This is the interaction between permanent and induced multipoles. In other words, induction arises because the electron distribution of one molecule is distorted by the

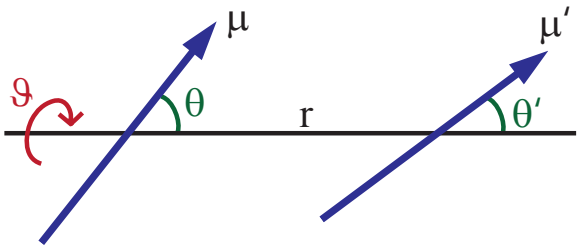


Figure 1.2: Definitions of the angles needed to describe the interaction between two permanent dipole moments with fixed orientations. The dipole moments drawn are located on a plane, the angle ϑ describes the rotation of the two dipole moments against each other.

electric fields of its neighboring molecules. The induction is always an attractive interaction but highly non-pairwise-additive due to the fact that additional electric fields can reinforce or also cancel an existing field. Here we also discuss the simplest term, the interaction of an uncharged molecule with a permanent dipole μ and an induced dipole μ_i in a nonpolar molecule with a static polarizability α. For a fixed orientation, the interaction in vacuum can be written as

$$V_{\mu\mu_i}^{\mathrm{ind}}(r) = -\frac{1}{2}\frac{1}{(4\pi\epsilon_0)^2}\frac{\mu^2\alpha}{r^6}\,(3\cos^2\theta + 1), \tag{1.4}$$

where the definition of θ is the same as shown in Figure 1.2. For a freely rotating system the interaction is

$$V_{\mu\mu_i}^{\mathrm{ind}}(r) = -\frac{1}{(4\pi\epsilon_0)^2}\frac{\mu^2\alpha}{r^6}. \tag{1.5}$$

The induction interaction between two freely rotating identical polar molecules with static polarizability α can be written as

$$V_{\mu\mu_i}^{\mathrm{ind}}(r) = -\frac{2}{(4\pi\epsilon_0)^2}\frac{\mu^2\alpha}{r^6}. \tag{1.6}$$

Equations 1.5 and 1.6 show that the induction interaction between a permanent dipole and an induced dipole is always proportional to r^{-6}, the interaction between a permanent quadrupole moment and an induced dipole moment would be proportional to r^{-8}.

Dispersion. This interaction is a phenomenon of quantum mechanics and originate from the correlation of electrons in the interacting molecules. As electrons move, instantaneous multipoles are created within the electron cloud of a molecule or atom. These instantaneous multipoles induce multipole moments in another molecule or atom. Therefore, dispersion can be seen as the interaction of instantaneously created multipoles with induced multipoles or in other words, the interaction between fluctuating multipoles. Dispersion interactions are the only interactions, which are always present between atoms or molecules because no permanent multipole moment is required. So uncharged non-polar molecules are held together mostly by dispersion interactions. The contributions of higher-order permanent multipole-multipole interactions are only significant at small separations because they are proportional to a high inverse power of r. Dispersion interactions are always attractive but not completely pairwise-additive. Sometimes these interactions are also called *van-der-Waals interactions* or even *hydrophobic interactions*. But the definition of these two terms is not unique, so in this thesis only the term *dispersion interactions* is used. Dispersion interactions can also be expressed by multipole-multipole interactions, while neglecting many-body effects:

$$V_{\mathrm{disp}}(r) = \frac{C_6}{r^6} + \frac{C_8}{r^8} + \frac{C_{10}}{r^{10}} + \dots \tag{1.7}$$

The terms C_X with $X = 6, 8, 10, \dots$ are coefficients. The C_6 coefficient corresponding to the interaction between fluctuating dipoles of identical molecules was first estimated by

London[5] in 1930 with

$$C_6 = -\frac{3}{4}\frac{\alpha^2 E_I}{(4\pi\epsilon_0)^2}, \tag{1.8}$$

where E_I is the ionization energy from the ground state of the molecules. Compared to recent calculations the error is around 30%.[2] In 1948 Casimir and Polder[6] suggested in atomic units

$$C_6^{\mathrm{AB}} = \frac{3}{\pi}\int_0^\infty \alpha(i\omega)^{\mathrm{A}}\alpha(i\omega)^{\mathrm{B}}d\omega, \tag{1.9}$$

where $\alpha(i\omega)^{\mathrm{X}}$ with X=A,B is the polarizability at imaginary frequencies for molecule A and B, respectively.

Exchange. The exchange is also a non-classical phenomenon and arises when the electron clouds of two interacting molecules start to overlap and the exchange of electrons becomes possible. It is a repulsion that can be related to the Pauli exclusion principle, which states that electrons with the same spin repel each other. It is sometimes also called Pauli repulsion. The exchange is often described by the r^{-12} term within the well known 12-6 Lennard-Jones potential.

The total interaction energy between two interacting molecules is the sum of the four previously discussed interactions. For a more detailed discussion of intermolecular interactions the reader is referred to Refs. 1–3.

1.2 Hydrogen Bonding

Hydrogen bonding is a central concept in chemistry introduced in the 1920s[7,8] for the description of the structure and stability of DNA molecules, polar liquids, molecular crystals with organic molecules, polysaccharides, and proteins.[4,9–14] The most important part of a hydrogen-bonded complex

R—D—H$\cdots$A—R

is the central moiety D—H$\cdots$A (shown in red), which is in most cases surrounded by some rests R. Atom D is more electronegative than hydrogen, so that there exists a polar bond between atoms D and H. The hydrogen of this D—H proton donor group is pointing towards a proton acceptor atom A, which has lone-pair electrons or at least polarizable π electrons. Hydrogen bonding is an attractive non-covalent interaction, which arises when the D—H group is *attached* to the acceptor atom A. Please note, that a hydrogen bond between two molecules does not involve the formation of a covalent bond, the two molecules retain their identity.

According to Jeffrey [4] hydrogen bonds can be classified as strong, moderate, and weak. Strong hydrogen bonds have bond energies ranging from around 60 to 160 kJ mol^{-1}. In this case the H$\cdots$A distance is comparable to the D—H bond length, eg. [F$\cdots$H$\cdots$F]$^-$. Therefore, one can argue if this interaction can still be considered hydrogen bonding, because the interacting molecules loose their identity.

In contrast, weak hydrogen bonds have bond energies of only around 4 to 15 kJ mol^{-1}. In this category D—H << H$\cdots$A, the distance between H and A varies between around 2.2 and 3.2 Å. Examples for this category include the central moieties C—H$\cdots$O and C—H$\cdots$N.

Between these two extremes there are the moderate hydrogen bonds. All hydrogen bonds discussed in this thesis belong to this category. Moderate hydrogen bonds have hydrogen bond energies ranging from around 15 to 60 kJ mol^{-1}, the D—H bond length is significantly smaller than the H$\cdots$A distance, with H$\cdots$A ranging from around 1.5 to 2.2 Å. The resulting distances between atoms D and A are between 2.5 and 3.2 Å. Another descriptor used for the classification of hydrogen bonds is the angle between the three atoms of the central moiety D—H$\cdots$A, the values for moderate hydrogen bonds are between 130 and 180 degrees. Here the acceptor atom must have at least one lone pair. Moderate hydrogen bonds are especially important because they can be found in nearly all biologically important molecules.

One very important central moiety belonging to moderate hydrogen bonds is O—H$\cdots$O. Hydrogen bonds containing this central moiety are responsible for the properties of liquid water or alcohols; they are also considered to be responsible for the structure and stability of biopolymers like cellulose.

The physical nature of hydrogen-bonding interactions is a quite controversial topic. The simplest approach was to assume that the attractive interaction stems from electrostatic interactions between the D—H group and the acceptor atom A, while neglecting all substituent effects. This interpretation was already questioned in 1952 by Coulson [15], who stated that also polarization effects (induction) and dispersion interactions must be considered. But still, it is very often claimed that for hydrogen bonding only electrostatic interactions are important. [4,14,16] Some authors even claim that hydrogen-bonding has a covalent character, [17,18] which is denied by Ghanty et al. [19] and Davidson. [20] Unfortunately the term *covalent* is often used in different contexts. For example sometimes charge transfer is considered as a sign for covalency. Charge transfer is a part of the induction interaction and we shall see later that induction plays only a minor role in the studied systems. So in that context there is no covalency in hydrogen bonding. Sometimes, the term *covalent* is used for interactions other than electrostatics. We shall see, that indeed non-electrostatic contributions are essential for hydrogen bonding, but one should certainly not use the term *covalent* for them. In this thesis hydrogen bonding will be discussed based on the four intermolecular interactions introduced in the previous section. We shall see, that the total interaction energy between two hydrogen-bonded systems is characterized by

an interplay of these four components, in which some components contribute more to the interaction than others.

1.3 Cellulose

The biopolymer cellulose is the most abundant material on earth, being the main constituent of plants and also present in animals, fungi, and algae.[21] It is a very useful component in fiber-based nanocomposites, which can be used in textiles,[22] flexible displays,[23] transparent films,[24] and even in barrier films.[25]

The structure of cellulose contains β-D-glucopyranose units, which are linked by 1→4 glycosidic bonds, making cellobiose the smallest repeating unit. Cellulose appears in several polymorphs, ranging from cellulose I to IV.[21] In this thesis only cellulose I will be discussed.

Native samples contain only cellulose I, which exists in two different crystalline forms, termed cellulose Iα and Iβ. The most recent crystal structures were published by Nishiyama et al.[26,27] In both crystalline forms sheets with parallel-aligned glucan chains were found. The crystal structure of cellulose Iα[26] is shown in Figure 1.3. It has a triclinic unit cell with space group P1. All sheets are equivalent and the glucan chains within a sheet are connected via hydrogen bonds, between two sheets no hydrogen bonds are observed. The crystal structure of cellulose Iβ[27] is shown in Figure 1.4, it has a monoclinic unit cell with space group $P2_1$. In this case, two alternating sheets were found with different conformations. The glucan chains within each sheet are also connected via hydrogen bonds, but both sheets show a different *hydrogen bonding network*. Like for cellulose Iα, no hydrogen bonds between sheets were found.

Traditionally the stability of crystalline cellulose within a sheet is only attributed to hydrogen bonding and it is claimed that the different sheets are held together by only weak van-der-Waals interactions,[28] whereas Qian et al.[29] and Nishiyama et al.[30] suggest that weak C—H$\cdots$O hydrogen bonds are responsible for the inter-sheet interaction. The role of dispersion interactions for the intra-sheet stabilization is commonly neglected and in the cellulose community hydrogen bonding is nearly always discussed only in terms of electrostatic interactions. Recently, Lindman and coworkers proposed that hydrophobic interactions (dispersion) are more important than hydrogen bonding for the stability of cellulose.[31–33]

Adsorption onto cellulose surfaces was studied for carbohydrates,[34–38] small proteins,[39] aromatic rings,[40,41] and also inorganic nanoparticles.[42] Experimental studies mostly focus on adsorption from solution but in many theoretical studies involving quantum theoretical calculations no explicit solvent is considered due to the computational expense. Mazeau and Charlier found that the structural features of adsorption can be calculated quite accurately without explicitly accounting for solvent effects, and they claim that such calculations lead to an overestimation of adsorption energies and therefore mimic "bad solvent conditions".[36]

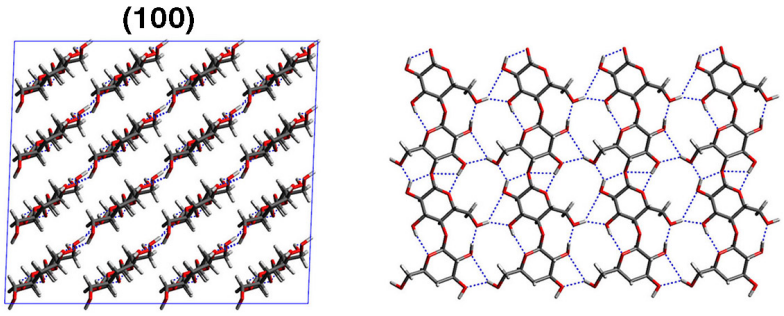


Figure 1.3: Supercell of cellulose Iα shown in the growth direction of the glucan chains (left) and glucan chains within a sheet (right).

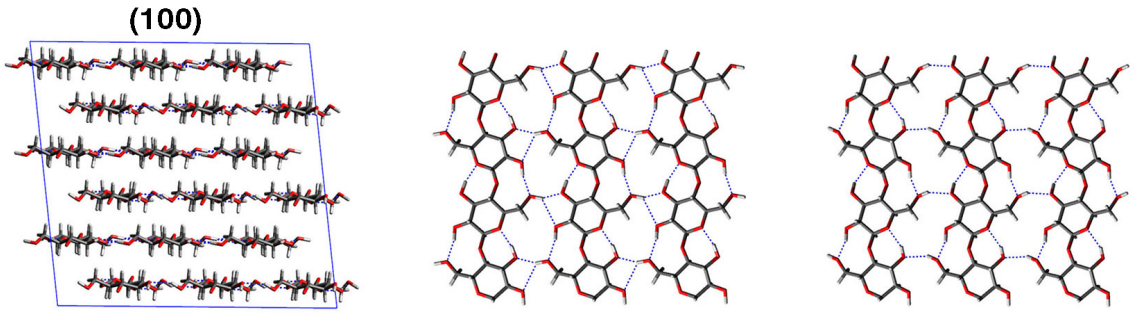


Figure 1.4: Supercell of cellulose Iβ shown in the growth direction of the glucan chains (left) and glucan chains within the two different sheets (middle and right).

Adsorption of molecules onto cellulose surfaces can take place on various surface types. In this thesis, only adsorption onto (100) surfaces are discussed. These surfaces can be seen in the supercells of Figure 1.3 and 1.4. The (100) surface of cellulose Iα has deep grooves and hydroxyl groups are accessible for adsorbates. Therefore, this surface is hydrophilic. In contrast, the (100) surface of cellulose Iβ is flat and the hydroxyl groups are less exposed. Therefore, this surface is hydrophobic.[43] Recently, Li et al.[44] stressed the importance of dispersion interactions for the crystal structure of cellulose Iα and Iβ as well as for the adsorption of water molecules.

1.4 Outline

The main goal of this master theses is to investigate the nature of the interaction between hydrogen-bonded systems, especially for the adsorption of carbohydrates onto cellulose surfaces. Additionally, the effect of the adsorbate size on the dominating interaction type should be discussed.

With the given computer resources, only the adsorption of a glucose molecule onto cellulose model surfaces could be studied by using Density-Functional Theory (DFT); larger adsorbates had to be treated with a force field method (GLYCAM06). Therefore, the quality of the GLYCAM06 force field was evaluated by comparing the results obtained with DFT and dispersion-corrected DFT for the adsorption of glucose onto a hydrophilic and

a hydrophobic cellulose model surface. The agreement between dispersion-corrected DFT and the GLYCAM06 force field was surprisingly good, so the adsorption of cellobiose and cellotetraose was studied with GLYCAM06 only. It was found, that dispersion interactions are essential for the adsorption of carbohydrates onto cellulose surfaces. A rigorous decomposition of the interaction energy between two molecules into electrostatics, induction, dispersion, and exchange interactions can be done by using symmetry-adapted perturbation theory (SAPT). Because of the huge computational costs of SAPT only small hydrogen-bonded systems (water dimer and progressively larger alcohol dimers) were studied with this method. It was found, that even in these small systems dispersion interaction are crucial for an accurate description of the interaction energy.

This thesis consists of five further chapters. In Chapter 2 all theoretical methods used for the calculation of interaction energies are discussed. Chapter 3 provides the results for the adsorption of a glucose molecule onto cellulose model surfaces including a method comparison. The results for the adsorption of cellobiose and cellotetraose are presented in Chapter 4. Furthermore, insight into the intermolecular interactions within the water dimer and several alcohol dimers is provided in Chapter 5. Finally, Chapter 6 contains the conclusions of this work and gives a short outlook for further research.

Chapter 2

Theoretical Methods

This chapter provides a short overview of the three theoretical methods, which are used in this thesis for the calculation of interaction energies. We start with a simple molecular-mechanics approach, the GLYCAM06 force field. Next, the basics of density-functional theory are presented including a description of several dispersion corrections. Finally, a short introduction into symmetry-adapted perturbation theory is given.

The section focusing on density-functional theory is based on the books of Szabo and Ostlund,[45] Parr and Yang,[46], and Piela,[47] and the work of Becke[48] and Kohn et al.[49] Discussion of symmetry-adapted perturbation theory follows the work of Jeziorski and Szalewicz,[50] Piela,[47] and Szalewicz.[51]

2.1 The GLYCAM06 Force Field

The simplest method used for the calculation of interaction or adsorption energies is the GLYCAM06 force field.[52] This force field was chosen, because it was explicitly designed for carbohydrates. By using a force field method, the implicit quantum nature of the studied system is neglected. Atoms are reduced to mass points with bonds connecting these points. Each force field uses several atom types with corresponding parameters for all potentials. GLYCAM06 belongs to the AMBER force fields and its total energy E_{total} can be written as[53]

$$E_{\text{total}} = E_{\text{bonds}} + E_{\text{angles}} + E_{\text{dihedrals}} + E_{\text{qq}} + E_{\text{VDW}} \tag{2.1}$$

The bond stretch energy E_{bonds} sums over all present covalent bonds and describes the displacement of a bond length r away from its equilibrium bond length r_{eq} with a harmonic potential. It can be written as

$$E_{\text{bonds}} = \sum_{\text{bonds}} K_r(r - r_{eq})^2, \tag{2.2}$$

depending on the instantaneous bond length r with K_r being the bond stretching force constant. This harmonic potential is a good approximation if the actual bond length is close to the equilibrium bond length. The angle bend potential E_{angles} depends on the instantaneous angle θ of all angles between covalently bound atoms and is given by

$$E_{\text{angles}} = \sum_{\text{angles}} K_\theta(\theta - \theta_{eq})^2, \tag{2.3}$$

where K_θ is the angle-bending force constant and θ_{eq} the equilibrium angle. Next, the dihedral potential $E_{\text{dihedrals}}$ describes the potential energy depending on the instantaneous dihedral angle ϕ of all dihedrals for the rotation about a bond according to

$$E_{\text{dihedrals}} = \sum_{\text{dihedrals}} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)], \tag{2.4}$$

where V_n is the rotation barrier, γ a phase factor, and n the periodicity. The Coulomb potential E_{qq} describes the pair-interaction of partial charges q_i and q_j calculated with the restrained electrostatic potential[54] between all atoms, which are in different molecules or at least separated by three bonds. It is given by

$$E_{\text{qq}} = \sum_{i<j} \frac{q_i q_j}{\epsilon R_{ij}}, \tag{2.5}$$

where R_{ij} is the distance between atom i and j and ϵ is the dielectric constant. The van der Waals term E_{VDW} is calculated between the same atoms as E_{qq} and consists of a

Lennard-Jones potential, given by

$$E_{\mathrm{VDW}} = \sum_{i<j} \epsilon_{ij} \left[\left(\frac{R_{ij}^*}{R_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}^*}{R_{ij}} \right)^{6} \right], \tag{2.6}$$

where R_{ij}^* is the van der Waals radius and ϵ_{ij} the minimum energy. The interaction energy between two monomers can be calculated according to

$$E_{\mathrm{int}} = E_{\mathrm{dimer}} - E_{\mathrm{monomerA}} - E_{\mathrm{monomerB}}. \tag{2.7}$$

The adsorption energy of an adsorbate on a surface can be defined as

$$E_{\mathrm{ads}} = E_{\mathrm{surface}} + E_{\mathrm{adsorbate}} - E_{\mathrm{surface+adsorbate}}. \tag{2.8}$$

2.2 Density-Functional Theory

2.2.1 Basics

A molecular system with N electrons can be described in the Born-Oppenheimer approximation by the non-relativistic time-independent Schrödinger equation

$$\hat{H}\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = E\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N), \tag{2.9}$$

with E being the electronic energy and $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ being the wave function of that system. The coordinates $\mathbf{x}$ for each electron consist of the spatial coordinates $\mathbf{r}$ and the spin coordinates s. In the remaining chapter atomic units will be used throughout. The Hamiltonian $\hat{H}$ for the N-electron system is given by

$$\hat{H} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N} \nu(\mathbf{r}_i) + \sum_{i<j}^{N} \frac{1}{r_{ij}} = \hat{T} + \hat{V}_{en} + \hat{V}_{ee}, \tag{2.10}$$

where $\hat{T}$ describes the kinetic energy of the electrons, $\hat{V}_{en}$ the attraction between electrons and nuclei, and $\hat{V}_{ee}$ the repulsion between electrons. The expression $\nu(\mathbf{r}_i)$ is an external potential, which acts on electron i due to M nuclei with charges Z_A.

$$\nu(\mathbf{r}_i) = -\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} \tag{2.11}$$

In the Born-Oppenheimer approximation we assume fixed positions for the nuclei, so their kinetic energy is zero. Therefore, the total energy of a system is the sum of the electronic

energy and the nucleus-nucleus repulsion energy $\hat{V}_{nn}$, which is given by

$$\hat{V}_{nn} = \sum_{A<B} \frac{Z_A Z_A}{R_{AB}}. \tag{2.12}$$

One important requirement on the wave function $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ describing a Fermi system is the fulfillment of the antisymmetry principle. It states that $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ must be antisymmetric with respect to the exchange of two electrons,

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_N) = -\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N). \tag{2.13}$$

The wave function $\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$ for a single electron is a *spin orbital* $\psi(\mathbf{x})$. Let us assume that we have a system, where the electrons do not interact with each other. The Hamiltonian of such a noninteracting system would be just the sum of independent one-electron hamiltonians $\hat{h}(i)$,

$$\hat{H}_{\text{nonint}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 + \sum_{i=1}^{N} \nu(\mathbf{r}_i) = \sum_{i=1}^{N} \hat{h}(i), \tag{2.14}$$

and the wave function can simply be written as the product of a set of spin orbitals,

$$\Psi^{\text{HP}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \psi_i(\mathbf{x}_1)\psi_j(\mathbf{x}_2)\cdots\psi_k(\mathbf{x}_N) \tag{2.15}$$

Antisymmetrization of this product allows the description of a non-interacting fermionic system. This antisymmetrized product can be written as a determinant, called the Slater determinant $\Psi^{\text{SD}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N)$.

$$\Psi^{\text{SD}}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{x}_1) & \psi_2(\mathbf{x}_1) & \cdots & \psi_N(\mathbf{x}_1) \\ \psi_1(\mathbf{x}_2) & \psi_2(\mathbf{x}_2) & \cdots & \psi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \psi_1(\mathbf{x}_N) & \psi_2(\mathbf{x}_N) & \cdots & \psi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.16}$$

For large N these calculations can be quite time consuming. The expectation value of the Hamiltonian can be calculated by using the one-particle density $\rho(\mathbf{r})$ (in the remainder simply called electron density) and the two-particle density $\rho(\mathbf{r}, \mathbf{r}')$. The electron density $\rho(\mathbf{r})$ describes the number of electrons in a certain unit volume. The electron density at the position $\mathbf{r}_1$ can be calculated according to

$$\rho(\mathbf{r}_1) = N \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)|^2 \, ds_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N. \tag{2.17}$$

The two-particle density can be calculated according to

$$\rho(\mathbf{r}_1, \mathbf{r}_2) = N(N-1) \int \cdots \int |\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N)|^2 \, ds_1 \, ds_2 \, d\mathbf{x}_3 \cdots d\mathbf{x}_N. \tag{2.18}$$

Integration of $\rho(\mathbf{r})$ over the whole space $\mathbf{r}$ yields the total number of electrons

$$\int \rho(\mathbf{r}) d\mathbf{r} = N \tag{2.19}$$

2.2.2 The Two Hohenberg-Kohn Theorems

The foundations of density-functional theory (DFT) are the two Hohenberg-Kohn theorems.[55] The first theorem states, that it is possible to determine the external potential $\nu(\mathbf{r})$ of a system within an additive constant by only using the ground-state electron density $\rho(\mathbf{r})$. With the knowledge of $\rho(\mathbf{r})$ it is possible to determine the ground-state wave function. Therefore, one can write the total energy of a system as

$$E[\rho] = T[\rho] + V_{ne}[\rho] + V_{ee}[\rho] = \int \rho(\mathbf{r})\nu(\mathbf{r}) d\mathbf{r} + F[\rho], \tag{2.20}$$

where the first term is a system-dependent quantity and $F[\rho]$ is the universally valid Hohenberg-Kohn functional, given by

$$F[\rho] = T[\rho] + V_{ee}[\rho]. \tag{2.21}$$

The second Hohenberg-Kohn theorem states the so called Hohenberg-Kohn density variational principle,

$$F[\rho'] \int \rho'(\mathbf{r})\nu(\mathbf{r}) d\mathbf{r} \geq F[\rho] \int \rho(\mathbf{r})\nu(\mathbf{r}) d\mathbf{r} = E_0. \tag{2.22}$$

The energy obtained with any trial density ρ' is always larger or equal to the exact energy of the ground state E_0 corresponding to the ground state density $\rho(\mathbf{r})$.

So, it is theoretically possible to determine all ground state properties of a given system by using density functional theory, but unfortunately the exact form of the Hohenberg-Kohn functional $F[\rho]$ is not known.

2.2.3 Kohn-Sham DFT

Kohn and Sham proposed the use of a fictitious non-interacting reference system demanding that the ground state electron density of the non-interacting system is exactly the same as the ground state electron density of the interacting system.[56] The Hamiltonian $\hat{H}_s$ of that system can be written as

$$\hat{H}_s = -\frac{1}{2} \sum_i^N \nabla_i^2 + \sum_i^N \nu_s(\mathbf{r}_i). \tag{2.23}$$

This Hamiltonian does not consider the electron-electron repulsion. The advantage is, that the wave function Ψ_s of the reference system is simply a slater determinant (see Equation

2.16), and the kinetic energy $T_s[\rho]$ can be calculated rather easily with

$$T_s[\rho] = \langle \Psi_s | - \frac{1}{2} \sum_i^N \nabla_i^2 | \Psi_s \rangle = -\frac{1}{2} \sum_i^N \langle \psi_i | \nabla^2 | \psi_i \rangle. \tag{2.24}$$

So one can rewrite $F[\rho]$ as

$$F[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho], \tag{2.25}$$

where $J[\rho]$ is the so called *self-interaction* of our electron cloud given by

$$J[\rho] = \frac{1}{2} \int \int \frac{1}{|\mathbf{r} - \mathbf{r}'|} \rho(\mathbf{r}) \rho(\mathbf{r}') \, d\mathbf{r} \, d\mathbf{r}'. \tag{2.26}$$

All still missing contributions are simply subsumed into $E_{xc}[\rho]$, which is called *exchange-correlation* energy and can be expressed as

$$E_{xc}[\rho] = T[\rho] - T_s[\rho] + V_{ee}[\rho] - J[\rho] \tag{2.27}$$

Still, the exact form of $E_{xc}[\rho]$ is unknown. Many approximations of it have been proposed and a few of them are discussed below. $E_{xc}[\rho]$ is often referred to as the *holy grail of density functional theory*. One can now write the total energy $E[\rho]$ as

$$E[\rho] = T_s[\rho] + J[\rho] + E_{xc}[\rho] + \int \nu(\mathbf{r}) \rho(\mathbf{r}) d\mathbf{r}, \tag{2.28}$$

which leads to the Kohn-Sham (KS) equation

$$\left(-\frac{1}{2} \nabla^2 + \nu(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r}' + \nu_{xc}(\mathbf{r}) \right) \psi_i = \epsilon_i \psi_i \tag{2.29}$$

with

$$\rho(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2, \tag{2.30}$$

and

$$\nu_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \tag{2.31}$$

2.2.4 Approximations to the Exchange-Correlation Functional

The first attempt to describe the exchange-correlation functional $E_{xc}[\rho]$ was the *local density approximation* (LDA)[56]

$$E_{xc}^{\mathrm{LDA}}[\rho] = \int \rho(\mathbf{r}) \epsilon_{xc}(\rho(\mathbf{r})) \, d\mathbf{r}, \tag{2.32}$$

where $\epsilon_{xc}(\rho(\mathbf{r}))$ describes the exchange-correlation energy per electron in a homogeneous electron gas. Obviously, this approximation can only yield good results for systems with a nearly homogeneous electron distribution.

Next, the *generalized gradient approximation* (GGA) was introduced:

$$E_{xc}^{\mathrm{GGA}}[\rho] = \int f(\rho(\mathbf{r}), \nabla_\rho)\, d\mathbf{r}. \tag{2.33}$$

In this case the exchange-correlation energy depends not only on the density $\rho(\mathbf{r})$, but also on its first derivative ∇_ρ.

A further improvement are the so called *meta-GGA* functionals, where the exchange-correlation energy depends additionally to GGA also on the second derivative of the density $\rho(\mathbf{r})$. GGA and meta-GGA functionals perform quite well for a variety of systems but all of them and also LDA suffer from the so called *self-interaction error*. The exchange part is not described correctly and so an electron can interact with itself. In contrast to DFT, Hartree-Fock describes the exchange correctly. Therefore, to overcome this problem, *hybrid functionals* were introduced, which include a certain amount of exact exchange from Hartree-Fock. In this work the GGA functional BP86 [57,58] and the hybrid functionals B3LYP, [57,59,60] M06-2X, [61] ωB97X, [62,63] and PBE0 [64–66] are used.

2.2.5 Dispersion Corrections to DFT

Recently developed density functionals are able to describe covalent bonds in molecules quite accurately, but often fail for intermolecular interactions, because they do not account for long-range electron correlation. Therefore, simple dispersion corrections E_{disp} are used and just added to the DFT energy E_{DFT}

$$E_{\mathrm{DFT-D}} = E_{\mathrm{DFT}} + E_{\mathrm{disp}}, \tag{2.34}$$

where $E_{\mathrm{DFT-D}}$ is then the dispersion-corrected DFT energy.

In this thesis, two dispersion corrections developed by Grimme and coworkers [67,68] are used: In the D2 scheme [67] the dispersion correction is given by the atom-atom function

$$E_{\mathrm{disp}}^{\mathrm{D2}} = -s_6 \sum_{A=1}^{M-1} \sum_{B=A+1}^{M} \frac{C_6^{AB}}{R_{AB}^6} f_{\mathrm{dmp}}(R_{AB}), \tag{2.35}$$

where s_6 is a scaling factor that depends on the density functional used for the DFT calculation, M is the number of atoms, R_{AB} the distance between atom A and B, and C_6^{AB} is the dispersion coefficient between these two atoms. The range of this dispersion correction is determined by a damping function $f_{\mathrm{dmp}}(R_{AB})$ given by

$$f_{\mathrm{dmp}}(R_{AB}) = \frac{1}{1 + e^{-d(R_{AB}/R_r - 1)}}, \tag{2.36}$$

with R_r being the sum of the atomic van-der-Waals radii and d being a parameter. We see, that this equation has a form similar to the interaction between two fluctuating dipoles.

The coefficients C_6^{AB} are calculated from the atomic coefficients C_6^A and C_6^B by

$$C_6^{AB} = \sqrt{C_6^A C_6^B}. \tag{2.37}$$

C_6^i parameters for an atom i were determined by PBE0 calculations according to

$$C_6^i = 0.05\, K\, I_p^i\, \alpha^i, \tag{2.38}$$

where K is 2 for first-row elements and 10 for second-row elements of the periodic table, I_p^i the atomic ionization potential, and α^i the static dipole polarizability of atom i.

The D3 scheme[68,69] used in this work describes the energy of the dispersion correction by

$$E_{\text{disp}}^{\text{D3}} = -\frac{1}{2}\sum_{A\neq B} s_6 \frac{C_6^{AB}}{R_{AB}^6 + [f(R_{AB}^0)]^6} + s_8 \frac{C_8^{AB}}{R_{AB}^8 + [f(R_{AB}^0)]^8}, \tag{2.39}$$

where s_8 is an additional scaling factor and the damping function $f(R_{AB}^0)$ is given by[69]

$$f(R_{AB}^0) = a_1 R_{AB}^0 + a_2. \tag{2.40}$$

where a_1 and a_2 are fit parameters and R_{AB}^0 is given by

$$R_{AB}^0 = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}}. \tag{2.41}$$

In the D3 scheme dispersion coefficients were calculated with time-dependent DFT, for details see Grimme et al.[68]

Another approach was developed by Pernal et al.[70] They developed a so-called dispersionless density functional (dlDF) supplemented by an atom-atom dispersion function D_{as}, which accounts for the complete dispersion interaction. This dispersion function was refined by Podeszwa et al.[71] and has the following form

$$D_{\text{as}} = \sum_{a\in A}\sum_{b\in B} D_{\text{as}}^{ab}(R_{ab}), \tag{2.42}$$

with

$$D_{\text{as}}^{ab}(R_{ab}) = -\frac{\sqrt{C_6^a C_6^b}}{R_{ab}^6}\cdot f_6\left(\sqrt{\beta_a\beta_b}R_{ab}\right) - \frac{\sqrt{C_8^a C_8^b}}{R_{ab}^8}\cdot f_8\left(\sqrt{\beta_a\beta_b}R_{ab}\right). \tag{2.43}$$

This function calculates all possible pair interactions between an atom a in monomer A and an atom b in monomer B. The damping function $f_n\left(\sqrt{\beta_a\beta_b}R_{ab}\right)$ with $n = 6, 8$ depends on the distance R_{ab} between atom a and b, and also on two fit parameters β_a and β_b. It

can be expressed in the following way:

$$f_n(x) = 1 - \exp(-x) \sum_{i=0}^{n} \frac{x^i}{i!}. \tag{2.44}$$

Another popular dispersion correction using pair-potentials was developed by Tkatchenko and Scheffler,[72] and recently Tkatchenko et al. developed a many-body dispersion correction.[73]

2.2.6 Interaction Energies

Interaction energies and also adsorption energies can be calculated with DFT in basically the same way as with GLYCAM06 (see Equations 2.7 and 2.8). For practical calculations a finite basis set has to be used, which causes a basis set superposition error in the interaction or adsorption energy, respectively. This error can be corrected by using the so called counterpoise correction.[74] Therefore, *ghost* atoms are used for the two monomer calculations. This means that for every monomer calculation the basis functions of the other monomer are also included but the nuclear charges of all atoms of the second monomer are set to zero.

2.3 Symmetry-Adapted Perturbation Theory

With symmetry-adapted perturbation theory (SAPT) one can calculate very accurate interaction energies, but one also gets a rigorous decomposition of the interaction energy in all physical interaction components, i.e., electrostatics, exchange, induction, and dispersion. In this thesis SAPT(DFT) is used, which is symmetry-adapted perturbation theory based on DFT calculations of the two monomers of an interacting dimer.

At first the Schrödinger equation is solved for the isolated monomers

$$\hat{H}_X \Phi_X = E_X \Phi_X, \qquad X = A \text{ or } B, \tag{2.45}$$

and the wave function of monomer A (Φ_A) and monomer B (Φ_B) are obtained. In the dimer case the monomers interact with each other, this interaction can be described with the intermolecular interaction operator $\hat{V}$, which can be expressed as

$$\hat{V} = \sum_{a \in \mathrm{A}} \sum_{b \in \mathrm{B}} \frac{Z_a Z_b}{R_{ab}} - \sum_{a \in \mathrm{A}} \sum_{j \in \mathrm{B}} \frac{Z_a}{r_{ja}} - \sum_{i \in \mathrm{A}} \sum_{b \in \mathrm{B}} \frac{Z_b}{r_{ib}} + \sum_{i \in \mathrm{A}} \sum_{j \in \mathrm{B}} \frac{1}{r_{ij}}, \tag{2.46}$$

where a and b are used for nuclei of monomer A and B, respectively, and i and j are used for electrons of monomer A and B, respectively; Z_a and Z_b are nuclear charges, R_{ab} describes a nucleus-nucleus distance, r_{ja} and r_{ib} an electron-nucleus distance, and r_{ij} an

electron-electron distance. Therefore, the Hamiltonian for the dimer can be written as

$$\hat{H} = \hat{H}_A + \hat{H}_B + \hat{V} = \hat{H}_0 + \hat{V}, \tag{2.47}$$

and the corresponding Schrödinger equation is

$$\hat{H}\Psi_{AB} = E_{AB}\Psi_{AB}, \tag{2.48}$$

where Ψ_{AB} is the dimer wave function and E_{AB} the corresponding energy of the dimer. The interaction between the two monomers can be treated as perturbation of the isolated monomers by using $\hat{V}$ as perturbation operator. So the unperturbed Hamiltonian is now $\hat{H}_0$. The corresponding unperturbed wave function Φ_0 is simply the product of the two monomer wave functions:

$$\Phi_0 = \Phi_A\Phi_B, \tag{2.49}$$

and the energy E_0 of the unperturbed system is the sum of the monomer energies

$$E_0 = E_A + E_B. \tag{2.50}$$

For the perturbation of the isolated monomers, Rayleigh-Schrödinger (RS) perturbation theory is used. For a detailed description of that theory the reader is referred to the book of Szabo and Ostlund.[45] The wave function of the dimer Ψ_{AB} and the interaction energy E_{int} can be expressed with a polarization expansion, which describes the mutual polarization of the monomers

$$\Psi_{\text{AB}} \approx \Phi_{\text{pol}} = \Phi_0 + \Phi_{\text{pol}}^{(1)} + \Phi_{\text{pol}}^{(2)} + \cdots \tag{2.51}$$

$$E_{\text{int}} = E_0 + E_{\text{pol}}^{(1)} + E_{\text{pol}}^{(2)} + \cdots \tag{2.52}$$

The numbers in parenthesis mark the order of perturbation theory. According to RS perturbation theory $E_{\text{int}}^{\text{RS}}$ can be calculated with

$$E_{\text{int}}^{\text{RS}} = \frac{\langle\Phi_0|\hat{V}\Phi_{\text{pol}}\rangle}{\langle\Phi_0|\Phi_{\text{pol}}\rangle}. \tag{2.53}$$

The first order polarization energy $E_{\text{pol}}^{(1)}$ describes the electrostatic interaction $E_{\text{elst}}^{(1)}$ and can be calculated according to

$$E_{\text{pol}}^{(1)} = E_{\text{elst}}^{(1)} = \langle\Phi_0|\hat{V}\Phi_0\rangle \tag{2.54}$$

In order to calculate the second-order polarization energy, one has to calculate the first-order polarization wave function $\Phi_{\text{pol}}^{(1)}$ according to

$$\Phi_{\text{pol}}^{(1)} = \sum_{kl}{}' \frac{\langle\Phi_A^0\Phi_B^0|\hat{V}\Phi_A^k\Phi_B^l\rangle}{E_A^0 + E_B^0 - E_A^k - E_B^l}|\Phi_A^k\Phi_B^l\rangle, \tag{2.55}$$

where the superscript k and l denote the quantum states of monomer A and monomer B, respectively. The prime means that the term with $k = 0$ and $l = 0$ is excluded. Please note that $\Phi_A^0 = \Phi_A$ and $\Phi_B^0 = \Phi_B$. The second-order polarization energy can now be written as

$$E_{\text{pol}}^{(2)} = \sum_{kl}{}' \frac{|\langle \Phi_A^0 \Phi_B^0 | \hat{V} \Phi_A^k \Phi_B^l \rangle|^2}{E_A^0 + E_B^0 - E_A^k - E_B^l}. \tag{2.56}$$

This term contains induction and dispersion interactions. All singly excited states belong to the second-order induction energy $E_{\text{ind}}^{(2)}$, i.e. terms with $\Phi_A^0 \Phi_B^l$ and $\Phi_A^k \Phi_B^0$ $(k, l > 0)$. Doubly excited states belong to the second-order dispersion energy $E_{\text{disp}}^{(2)}$, i.e. terms with $\Phi_A^k \Phi_B^l$ $(k, l > 0)$.

With this approach one can get values for the electrostatic, the induction, and the dispersion interaction. However, there are no exchange interactions present. This is due to the fact that no term of Φ_{pol} fulfills the antisymmetry requirement. To overcome this problem, an antisymmetrization operator $\hat{A}$ was introduced, given by

$$\hat{A} = \frac{1}{N!} \sum_{\mathcal{P} \in S_n} \text{sgn}(\hat{P})\ \hat{P}, \tag{2.57}$$

where $\hat{P}$ denotes an intermonomer permutation operator. The wave function is now approximated by

$$\Psi_{AB} \backsimeq \hat{A} \Phi_{\text{pol}}, \tag{2.58}$$

which yields asymptotically exact results. So the interaction energy is now given by

$$E_{\text{int}}^{\text{SRS}} = \frac{\langle \Phi_0 | \hat{V} \hat{A} \Phi_{\text{pol}} \rangle}{\langle \Phi_0 | \hat{A} \Phi_{\text{pol}} \rangle}, \tag{2.59}$$

the superscript inducates that now Symmetrized Rayleigh-Schrödinger (SRS) perturbation theory is used. Now, one can express the energies up to second-order as

$$E_{\text{SRS}}^{(1)} = \frac{\langle \Phi_0 | \hat{V} \hat{A} \Phi_0 \rangle}{\langle \Phi_0 | \hat{A} \Phi_0 \rangle}, \tag{2.60}$$

$$E_{\text{SRS}}^{(2)} = \frac{\langle \Phi_0 | \hat{V} - E_{\text{SRS}}^{(1)} | \hat{A} \Phi_{\text{pol}}^{(1)} \rangle}{\langle \Phi_0 | \hat{A} \Phi_0 \rangle}. \tag{2.61}$$

The exchange energies are now simply the differences between SRS and RS perturbation theory:

$$E_{\text{exch}}^{(n)} = E_{\text{SRS}}^{(n)} - E_{\text{RS}}^{(n)} \tag{2.62}$$

From the first-order energy one gets the exchange interaction or Pauli repulsion $E_{\text{exch}}^{(1)}$, which was discussed in the introduction; from the second-order energy one gets two coupling terms, the exchange-induction energy $E_{\text{exch}-\text{ind}}^{(2)}$ and the exchange-dispersion energy

$E^{(2)}_{\mathrm{exch-disp}}$. These terms can be seen as modifications to the induction or dispersion energy, respectively, if exchange interactions are taken into account.[3] With SAPT(DFT) the perturbation theory is truncated after second order. However, higher-order induction contributions can also be important, especially for hydrogen-bonded systems. An estimate of these interactions can be calculated by a SAPT calculation based on Hartree-Fock. This estimate is abbreviated with δ^{HF}.

To sum up, interaction energies calculated with SAPT(DFT) can be decomposed according to

$$E_{\mathrm{SAPT}} = E^{(1)}_{\mathrm{elst}} + E^{(1)}_{\mathrm{exch}} + E^{(2)}_{\mathrm{disp}} + E^{(2)}_{\mathrm{exch-disp}} + E^{(2)}_{\mathrm{ind}} + E^{(2)}_{\mathrm{exch-ind}} + \delta^{\mathrm{HF}} \tag{2.63}$$

2.4 Summary

In this chapter three theoretical methods have been discussed, which are used in the further chapters to calculate interaction energies. Systems with several thousand atoms can be studied with the GLYCAM06 force field. With this method only the total interaction or adsorption energy can be obtained. For smaller model systems containing up to around 150 atoms (for our hardware) density-functional theory (DFT) can be applied. If a dispersion correction is used, one can roughly estimate the lower-bound of the dispersion interaction. The most accurate results and the most insight can be obtained by using symmetry-adapted perturbation theory (SAPT). With this method one gets accurate values for the electrostatic, the exchange, the induction, and the dispersion interaction. If DFT is used for the description of the monomers [SAPT(DFT)], systems up to around 40 atoms can be calculated given our current hardware.

Chapter 3

Adsorption of Glucose onto Cellulose Surfaces

In this chapter the adsorption of a glucose molecule onto two different cellulose model surfaces (i.e., the (100) surfaces of cellulose Iα and Iβ) is discussed. In order to validate the quality of the GLYCAM06 force field for describing adsorption processes involving carbohydrates, GLYCAM06 results are compared to DFT values with and without dispersion correction. It is also discussed if an evidence for the importance of dispersion interactions for this adsorption process can be found. These results were recently published in the *Journal of Physical Chemistry B.*[75]

3.1 Computational Details

In order to study the adsorption of a β-D-glucose molecule onto cellulose surfaces with quantum theoretical methods, small model systems must be used. To model the (100) surfaces of cellulose Iα and cellulose Iβ two adjacent cellobiose strands were used. The size of these model systems is sufficient to describe the major interactions of an adsorbed glucose molecule with the surface. The small contributions from neighboring surface strands and from lower-lying layers of the cellulose bulk are neglected. These model surfaces were constructed from surface slabs that were optimized and equilibrated with the GLYCAM06 force field by Maurer et al.[76] An illustration of the used model surfaces is shown in Figure 3.1. In the remaining chapter these model systems will be called *surface*.

To get a representative sample of structures for the adsorption of a glucose molecule, starting geometries for the geometry optimization were created for both surfaces as follows: The principal axis of inertia, which goes through the exocyclic hydroxymethyl group of the adsorbate (glucose), will be called *molecular axis*. At first, the adsorbate was positioned above the surface model with its molecular axis parallel to the growth direction of the glucan chains of the surface. The smallest distance between two carbon atoms were set to around 4 Å. The adsorbate was positioned on top of a surface glucose ring, on top of a glycosidic bond of the surface, and between the two surface strands. Next, the angle between the molecular axis of the adsorbate and the surface plane was set to 45 degrees and the positioning was done as described above. In one set of structures the exocyclic hydroxymethyl group of the adsorbate was pointing towards the surface and in another set it was pointing away from the surface. Finally, the angle between the molecular axis of the adsorbate and the surface plane was set to 90 degrees. Positioning was done as described above and also different orientations of the exocyclic hydroxymethyl group were considered.

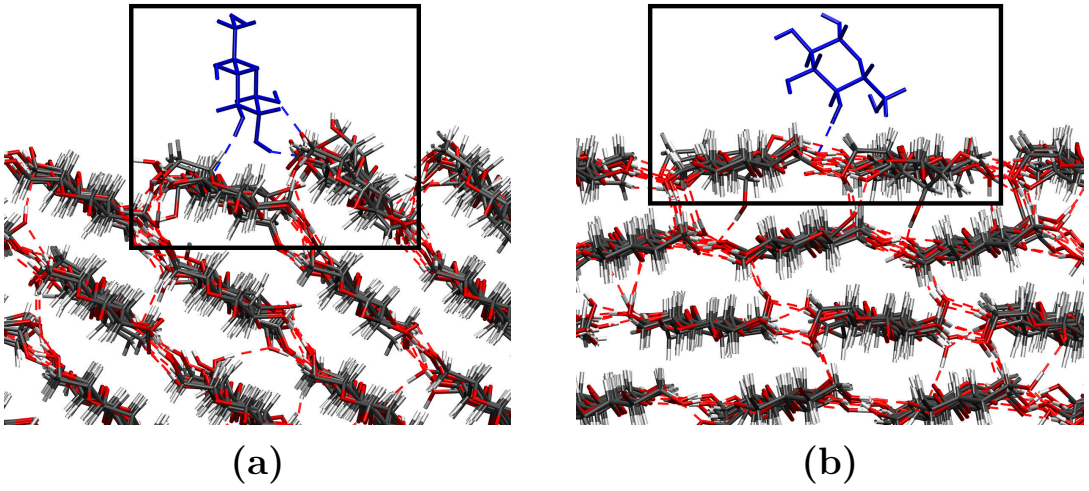


Figure 3.1: Illustration of the adsorption of a glucose molecule onto the (100) surfaces of cellulose Iα (a) and cellulose Iβ (b). The used surface models are shown in the black rectangles.[75]

These surface/glucose complexes were optimized at the BP86[57,58]/def2-SVP[77] level of theory by using the software package TURBOMOLE[78–81] with the resolution of the identity approximation.[82–87] The default convergence criteria were used for electronic structure and geometry. In the surface models only the exocyclic hydroxymethyl groups and the OH groups were optimized, the coordinates of all other atoms of the surface models were kept frozen.

The optimized geometries can be classified according to the orientation of the glucose on the surface. This can be done by using a parameter γ, which is the angle between the surface plane and the principal axis of inertia of the glucose molecule that corresponds to the largest moment of inertia (3-axis). An angle γ of zero degrees means that the glucose is oriented completely vertical to the surface and $|\gamma| = 90$ degrees means that the glucose is oriented completely parallel to the surface. The term *vertical* is used for structures with $|\gamma| \leqslant 45$ degrees and structures with $|\gamma| > 45$ degrees are called *parallel*. For each surface three representative structures (one parallel and two vertical structures) were chosen for our method comparison. In one vertical structure the adsorbate was located on top of a cellobiose molecule of the surface and in the other structure it was located between the two cellobiose molecules of the surface. These representative structures will be called BP86-optimized geometries. In order to estimate how explicit inclusion of dispersion interaction changes the geometries, they were re-optimized with BP86-D2,[57,58,67] yielding the BP86-D2-optimized geometries.

The main goal of this study is to evaluate the performance of the GLYCAM06[52] force field in describing the structure of the surface/adsorbate complexes. Therefore, the BP86-D2-optimized geometries were also re-optimized with GLYCAM06, these structures are termed GLYCAM06-optimized geometries. All GLYCAM06 calculations were done by using DL_POLY[88] with the ChemShell 3.3.1 interface.[89] For the optimizations the geometry optimizer DL-FIND[90] was used. For the generation of input files AmberTools 1.4[91] was used employing the GLYCAM06g parameters. Geometries were edited with Aten[92] and visualization was done with VMD.[93]

For all the selected representative structures the adsorption energy E_{ads} for the adsorption of the glucose molecule onto the cellulose surface was calculated according to

$$E_{\text{ads}} = E_{\text{surface}} + E_{\text{adsorbate}} - E_{\text{surface+adsorbate}} \tag{3.1}$$

The adsorption energy E_{ads} was calculated with BP86[57,58], B3LYP[57,59,60], BP86-D2[57,58,67], B3LYP-D2[57,59,60,67], B3LYP-D3[57,59,60,68,69], ωB97X-D[62,63], M06-2X[61], and GLYCAM06.[52]

The five methods that cover medium-range or long-range dispersion effects (BP86-D2, B3LYP-D2, B3LYP-D3, M06-2X, and ωB97X-D) will be called *DFT-D methods*; BP86 and B3LYP, will be called *DFT methods* throughout this chapter. The counterpoise correction was applied for all calculations of the adsorption energies with the DFT and DFT-D

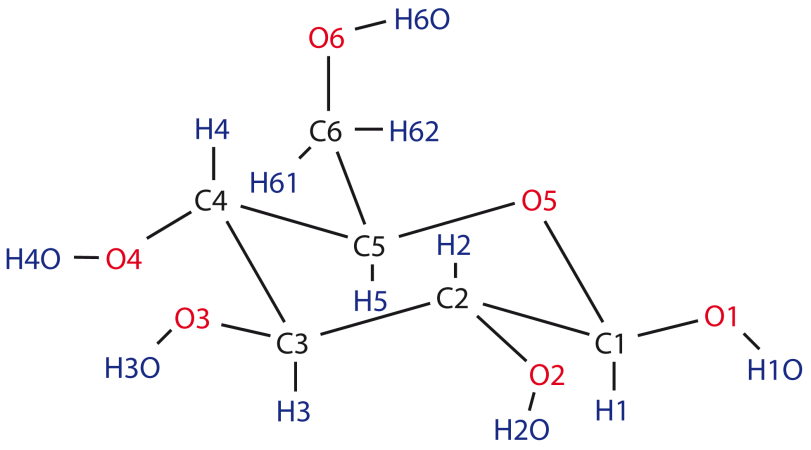


Figure 3.2: Naming convention employed for all atoms within a glucose molecule.[75]

methods.[74]

For all quantum theoretical calculations of adsorption energies the def2-TZVP basis set was used.[94] Gaussian 09[95] was used for all ωB97X-D and M06-2X calculations; all other quantum theoretical calculations were performed by using TURBOMOLE,[79–81] with the resolution of the identity (RI) approximation and the auxiliary basis set for def2-TZVP.[84–87] Default convergence criteria were used.

In order to discuss hydrogen bonds, a special naming convention will be used for all atoms within a glucose unit. This convention is shown in Figure 3.2. In this chapter, hydrogen bonds are named by the atoms involved in the O–H$\cdots$O moieties. The first mentioned oxygen atom will always belong to the adsorbate and the second one to the surface. Only hydrogen bonds between the adsorbate and the surface will be discussed.

3.2 Comparison of Optimized Geometries

For the adsorption of a β-D-glucose molecule onto the (100) surfaces of cellulose Iα and Iβ, six BP86-optimized geometries were selected as representatives for the method comparison. These geometries are shown in Figure 3.3 (blue) together with the corresponding structures that were re-optimized with BP86-D2 (red) and the GLYCAM06 force field (green).

In order to compare the results of the different optimization methods various parameters are used. It can be expected that optimization with BP86-D2 yields the most reliable results. Therefore, differences are only discussed with respect to BP86-D2. The first parameter is the previously defined angle γ (see Section 3.1). It measures the alignment of the molecular plane of the adsorbate to the surface plane, i.e. if the adsorbate is aligned parallel, vertical, or in between. All values of γ are shown in Table 3.1. We see, that for most model systems the agreement between the three methods is quite good. The ranges of the deviations of BP86 from BP86-D2 are -20.0 to 25.1 degrees and the corresponding mean of the absolute values amounts to 9.2 degrees. For GLYCAM06 the deviations from

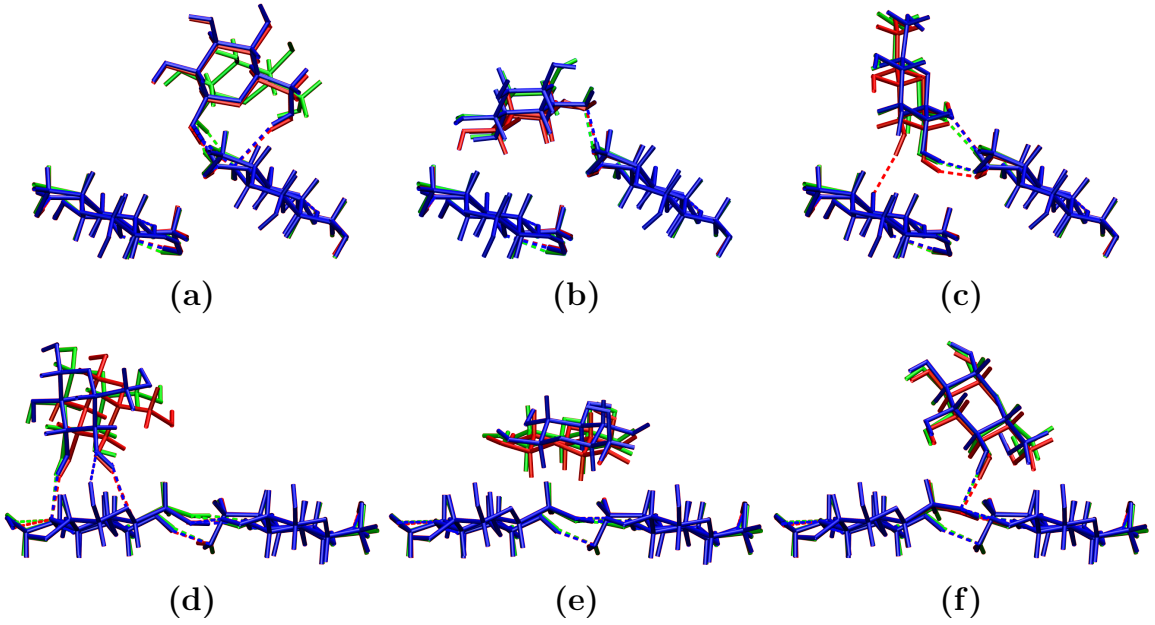


Figure 3.3: Optimized geometries of glucose onto the surfaces models for systems Iα-1 (a), Iα-2 (b), Iα-3 (c), Iβ-1 (d), Iβ-2 (e), and Iβ-3 (f) calculated with BP86 (blue), BP86-D2 (red), and GLYCAM06 (green).[75]

Table 3.1: Angle γ for all three optimizations in degrees.[75]

Struct.	BP86	BP86-D2	GLYCAM06
Iα-1	-21.6	-21.4	-56.4
Iα-2	67.2	87.2	71.6
Iα-3	5.1	4.6	8.5
Iβ-1	7.4	-17.7	-9.7
Iβ-2	69.1	76.5	76.2
Iβ-3	-32.0	-33.8	-33.4

BP86-D2 are between -35.0 and 8.0 degrees and the corresponding mean of the absolute values amounts to 10.5 degrees.

The next parameter is the normal distance d_n between the centers-of-mass of the glucose molecule and the surface model. It can be used to compare the relative distances between the surface and the adsorbate in the optimized geometry sets. All values of d_n are shown in Table 3.2. In general, BP86 and GLYCAM06 yield larger values for d_n than BP86-D2, the only exception is the geometry Iα-1 optimized with GLYCAM06. The ranges of the deviations of BP86 from BP86-D2 are 0.12 to 0.46 Å and the corresponding mean of the absolute values amounts to 0.29 Å. For GLYCAM06 the deviations from BP86-D2 are between -0.33 to 0.41 Å and the corresponding mean of the absolute values amounts to 0.26 Å.

Table 3.2: Distance d_n for all three optimizations in Å.[75]

Struct.	BP86	BP86-D2	GLYCAM06
Iα-1	3.06	2.94	2.61
Iα-2	2.39	1.93	2.34
Iα-3	3.48	3.20	3.49
Iβ-1	2.58	2.40	2.59
Iβ-2	2.31	1.88	2.07
Iβ-3	3.03	2.77	2.93

Table 3.3: RMSD values for the BP86-optimized geometries and the GLYCAM06-optimized geometries in comparison with the corresponding BP86-D2-optimized geometries in Å.[75]

Struct.	BP86	GLYCAM06
Iα-1	0.09	0.73
Iα-2	0.59	0.57
Iα-3	0.48	0.46
Iβ-1	0.77	0.45
Iβ-2	0.41	0.18
Iβ-3	0.20	0.21

Another parameter is the root mean square distance between two structures (RMSD). All RMSD values are shown in Table 3.3; no value is greater than 0.8 Å. The ranges of the RMSD values between BP86 and BP86-D2 are 0.09 to 0.77 Å and the corresponding mean value amounts to 0.42 Å. For GLYCAM06 the RMSD from BP86-D2 is between 0.18 to 0.73 Å and the corresponding mean value amounts to 0.43 Å. Based on the three discussed parameters, one can say that the agreement between the structures obtained with all three methods is remarkable, especially considering that only weak intermolecular interactions are holding the surface and the adsorbate together.

Finally, the occurrence of hydrogen bonds between the adsorbate and the surface can also be used to compare the optimized geometries. Figure 3.3 shows only hydrogen bonds, which have O—O distances below 3.5 Å and O—H$\cdots$O angles greater than 140 degrees. In the BP86 and BP86-D2-optimized geometry of Iα-1, there are three hydrogen bonds (O1$\cdots$H—O2, O1—H$\cdots$O3, and O6—H$\cdots$O3); whereas in the GLYCAM06-optimized geometry, the O6—H$\cdots$O3 hydrogen bond is missing. For the parallel structure Iα-2 all optimization methods find only one hydrogen bond (O6$\cdots$H—O2). For Iα-3, all methods yield two hydrogen bonds (O1$\cdots$H—O2 and O2—H$\cdots$O6); BP86-D2 yields also a O3—H$\cdots$O4 hydrogen bond.

For Iβ-1, there is no agreement between the three methods: BP86 finds three hydrogen bonds (O4—H$\cdots$O5, O3—H$\cdots$O3, and O4$\cdots$H—O1); BP86-D2 finds two hydrogen bonds (O4—H$\cdots$O5 and O3—H$\cdots$O3), and GLYCAM06 finds only one hydrogen bond (O3—H$\cdots$O3). For the parallel system Iβ-2 there is a good agreement among all methods.

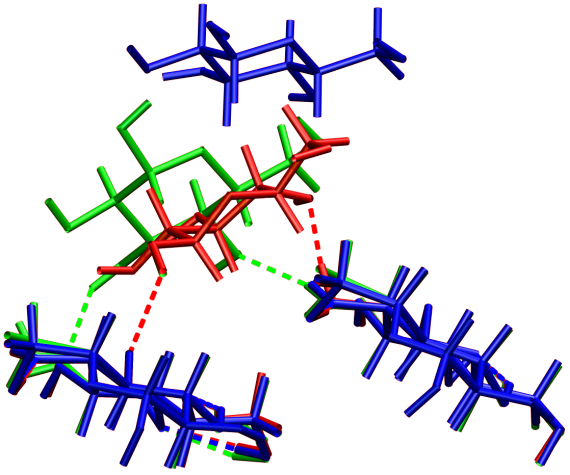


Figure 3.4: Illustration of the effect of dispersion interactions: Geometries were optimized using methods with and without account of dispersion interactions. BP86 (blue), BP86-D2 (red), and GLYCAM06 (green). For all optimizations the same starting geometry was used, which is practically identical to the BP86 optimized geometry.[75]

There, the parallel orientation of the glucose rings prevents the formation of any hydrogen bond with the surface. This is also observed between the layers of cellulose strands in crystalline cellulose Iβ[76,96]. For Iβ-3 all methods agree and predict only one hydrogen bond (O4—H$\cdots$O6).

For an accurate prediction of an adsorption structure dispersion interactions are very important, especially when weakly interacting complexes are optimized. Let us start an optimization with a structure where the adsorbate is aligned parallel to the surface plane and far away from the surface. With BP86 there are no attractive interactions between the adsorbate and the surface and the starting geometry remains unchanged and no hydrogen bonds are formed. In contrast to that, in geometry optimizations with GLYCAM06 and BP86-D2 the adsorbate is pulled towards the surface model and also hydrogen bonds are formed. An example for such a situation is shown in Figure 3.4.

3.3 Adsorption Energies for BP86-Optimized Geometries

For all previously shown geometries the adsorption energies were calculated with several methods. At first, the adsorption energies of the BP86-optimized geometries are discussed. All adsorption energies are listed in Table A.1 in the Appendix and plotted in Figure 3.5.

It can bee seen that adsorption energies calculated with BP86 and B3LYP (DFT methods) are nearly identical and only around half the size of the corresponding DFT-D values. For the system Iβ-2 the counterpoise-corrected DFT adsorption energies are even slightly negative. For the DFT-D methods M06-2X yields the lowest adsorption energies

and B3LYP-D2 the highest adsorption energies. The differences between them vary between 11 and 19 kJ/mol. The adsorption energies calculated with the GLYCAM06 force field are surprisingly close to the corresponding DFT-D results.

The trends in the adsorption energies for the different structures are very similar for all DFT-D methods, whereas the GLYCAM06 results are sometimes lower and sometimes higher than the DFT-D results. For the cellulose Iα structures there is a good agreement between the DFT-D and GLYCAM06 values, whereas for the cellulose Iβ structures larger differences are observed: for Iβ-1 the GLYCAM06 adsorption energy is smaller than any DFT-D value, but for Iβ-2 it is larger than any DFT-D value. In the BP86-optimized geometry of Iβ-1 there are three hydrogen bonds present, while the Iβ-2 geometry has no hydrogen bonds at all. So in the case of adsorption onto cellulose Iβ GLYCAM06 underestimates the adsorption energy for structures with many hydrogen bonds, while it overestimates the adsorption energy for structures without hydrogen bonds.

In order to estimate the quality of the GLYCAM06 adsorption energies more quantitatively, the following terms were calculated for every structure: The mean value of the DFT-D adsorption energies $\bar{E}_{\mathrm{DFT-D}}$ with the corresponding standard deviation σ, the difference between the GLYCAM06 adsorption energy and the DFT-D mean value of the adsorption energy $E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$, and the maximal difference between the five DFT-D values $\Delta_{\max} E_{\mathrm{DFT-D}}$. All values of these terms are listed in Table 3.4.

For the cellulose Iα structures the GLYCAM06 results agree quite well with the corresponding DFT-D mean values $\bar{E}_{\mathrm{DFT-D}}$. For all three structures the absolute value of the difference between the GLYCAM06 value and the DFT-D mean value, $|E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}|$, is even smaller than the spreading of the DFT-D results ($\Delta_{\max} E_{\mathrm{DFT-D}}$). For the cellulose I$\beta$ structures there are larger differences between GLYCAM06 adsorption energies and the

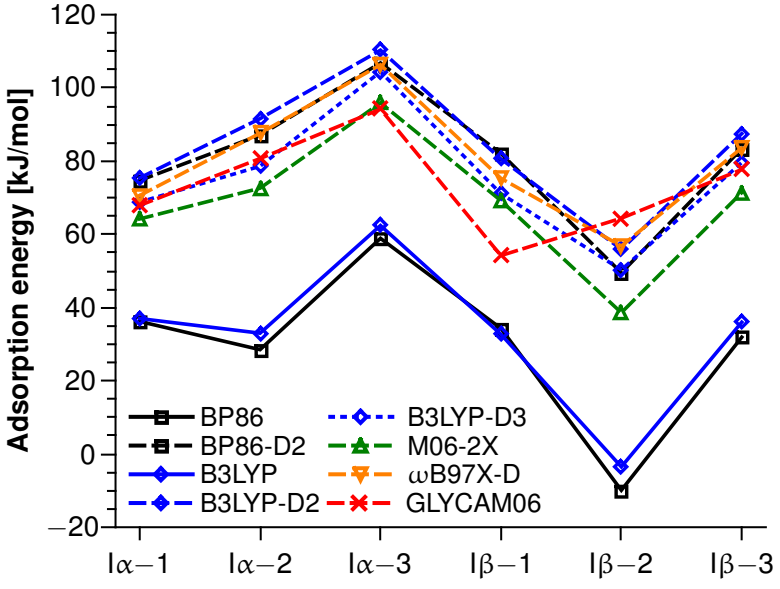


Figure 3.5: Adsorption energies for the BP86-optimized geometries.[75]

Table 3.4: Statistics of adsorption energies for BP86-optimized geometries. The mean values of DFT-D adsorption energies ($\bar{E}_{\mathrm{DFT-D}}$) with corresponding standard deviations (σ), the GLYCAM06 adsorption energies (E_{FF}), the differences between GLYCAM06 values and mean DFT-D values ($E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$) , and the maximal differences between DFT-D values ($\Delta_{\mathrm{max}}E_{\mathrm{DFT-D}}$) are shown. All energies are in kJ/mol.[75]

Structure	$\bar{E}_{\mathrm{DFT-D}} \pm \sigma$	E_{FF}	$E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$	$\Delta_{\mathrm{max}}E_{\mathrm{DFT-D}}$
Iα-1	70.8 ± 4.6	67.9	-2.9	11.2
Iα-2	83.5 ± 7.7	80.7	-2.8	19.0
Iα-3	104.7 ± 5.4	94.2	-10.5	14.4
Iβ-1	75.7 ± 5.7	54.2	-21.5	12.9
Iβ-2	50.2 ± 7.2	64.4	14.2	18.1
Iβ-3	81.1 ± 6.1	77.7	-3.4	16.1

DFT-D mean adsorption energies. In one case (Iβ-1) the absolute value of $E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$ is significantly larger than the spreading of the DFT-D results. For Iβ structures no systematic error can be observed for the GLYCAM06 results.

3.4 Adsorption Energies for BP86-D2-Optimized Geometries

All adsorption energies of the BP86-D2-optimized geometries are listed in Table A.2 in the Appendix and plotted in Figure 3.6. The GLYCAM06 values, the DFT-D mean values, and the previously defined deviations are presented in Table 3.5. Adsorption energies calculated with the DFT methods are for the BP86-D2-optimized geometries always significantly smaller than for the BP86-optimized geometries; in the parallel structure Iβ-2 the interaction between the adsorbate and the surface is now even repulsive by around 40 kJ/mol and in Iα-2 the interaction is nearly zero. It is not surprising that the BP86-D2 adsorption energies are for the BP86-D2-optimized geometries always larger than for the corresponding BP86-optimized geometries. However, for the remaining DFT-D methods this trend is not always observed. In fact, the spreading of the DFT-D adsorption energies ($\Delta_{\mathrm{max}}E_{\mathrm{DFT-D}}$) is for all BP86-D2-optimized geometries larger than for the corresponding BP86-optimized geometries.

Except for Iβ-2 all GLYCAM06 adsorption energies are significantly lower than the corresponding DFT-D mean values. In Iβ-2 there are no hydrogen bonds present and the GLYCAM06 adsorption energy is comparable with the largest DFT-D value. The lowest GLYCAM06 adsorption energy is found for the system Iβ-1, in which hydrogen bonds are present. For the BP86-D2-optimized geometries larger differences between the DFT-D mean values $\bar{E}_{\mathrm{DFT-D}}$ and the GLYCAM06 results are observed than for the BP86-optimized geometries; in two cases the absolute values of $E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$ are larger than the spreading of the DFT-D mean values $\Delta_{\mathrm{max}}E_{\mathrm{DFT-D}}$. Nevertheless, the quality of the GLYCAM06

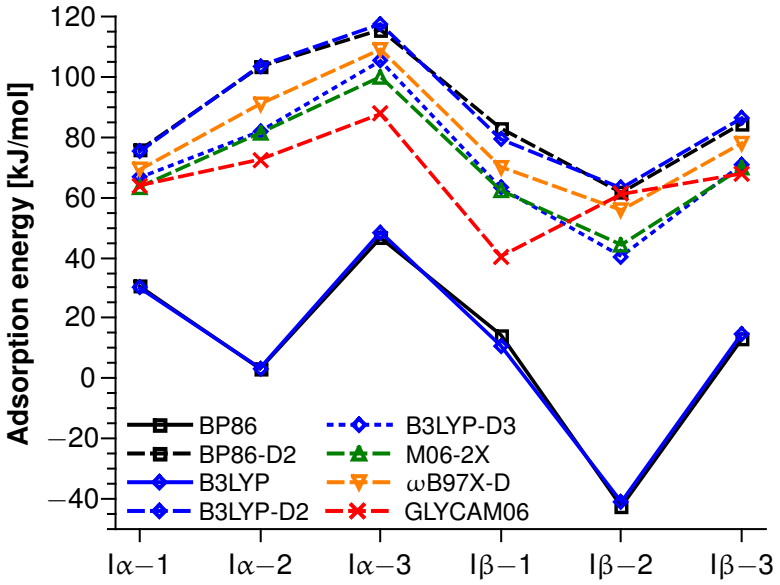


Figure 3.6: Adsorption energies for the BP86-D2-optimized geometries.[75]

Table 3.5: Statistics of adsorption energies for BP86-D2-optimized geometries. The mean values of DFT-D adsorption energies ($\bar{E}_{\mathrm{DFT-D}}$) with corresponding standard deviations (σ), the GLYCAM06 adsorption energies (E_{FF}), the differences between GLYCAM06 values and mean DFT-D values ($E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$) , and the maximal differences between DFT-D values ($\Delta_{\mathrm{max}} E_{\mathrm{DFT-D}}$) are shown. All energies are in kJ/mol.[75]

Structure	$\bar{E}_{\mathrm{DFT-D}} \pm \sigma$	E_{FF}	$E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$	$\Delta_{\mathrm{max}} E_{\mathrm{DFT-D}}$
Iα-1	70.0 ± 5.3	64.0	-6.0	12.2
Iα-2	92.3 ± 11.0	72.6	-19.7	22.3
Iα-3	109.5 ± 7.2	87.7	-21.8	17.5
Iβ-1	71.5 ± 9.3	40.5	-31.0	20.6
Iβ-2	53.1 ± 10.2	61.2	8.1	22.6
Iβ-3	77.9 ± 7.6	67.9	-10.0	16.5

adsorption energies is still quite remarkable.

3.5 Adsorption Energies for GLYCAM06-Optimized Geometries

All adsorption energies of the GLYCAM06-optimized geometries are listed in Table A.3 in the Appendix and plotted in Figure 3.7. The GLYCAM06 values, the DFT-D mean values, and the previously defined deviations are presented in Table 3.6.

As before, the DFT methods yield the lowest adsorption energies; The adsorbate/surface interaction in Iβ-2 is even repulsive by around 20 kJ/mol. For all vertical structures the GLYCAM06 adsorption energies agree very well with the DFT-D results. For all those structures the absolute values of $E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$ are below 5 kJ/mol and therefore also much

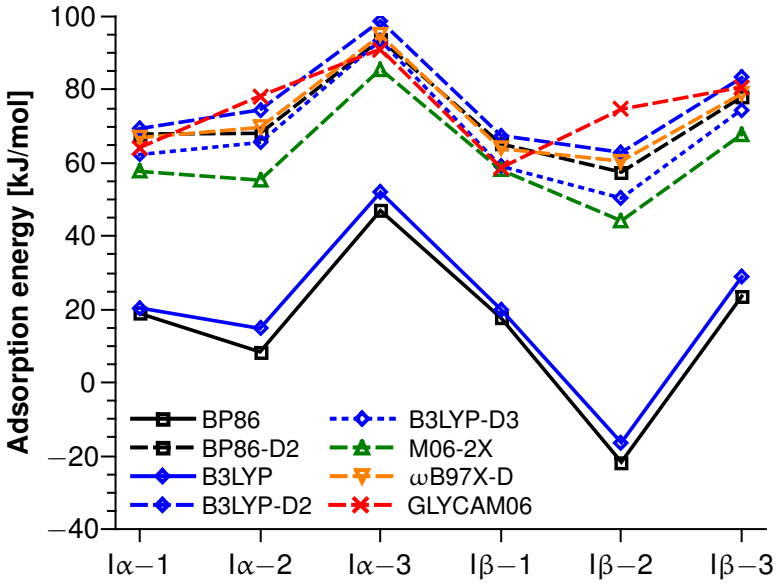


Figure 3.7: Adsorption energies for the GLYCAM06-optimized geometries.[75]

Table 3.6: Statistics of adsorption energies for GLYCAM06-optimized geometries. The mean values of DFT-D adsorption energies ($\bar{E}_{\mathrm{DFT-D}}$) with corresponding standard deviations (σ), the GLYCAM06 adsorption energies (E_{FF}), the differences between GLYCAM06 values and mean DFT-D values ($E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$) , and the maximal differences between DFT-D values ($\Delta_{\mathrm{max}} E_{\mathrm{DFT-D}}$) are shown. All energies are in kJ/mol.[75]

Structure	$\bar{E}_{\mathrm{DFT-D}} \pm \sigma$	E_{FF}	$E_{\mathrm{FF}} - \bar{E}_{\mathrm{DFT-D}}$	$\Delta_{\mathrm{max}} E_{\mathrm{DFT-D}}$
Iα-1	64.8 ± 4.8	64.2	-0.6	11.7
Iα-2	66.6 ± 7.1	78.3	11.7	19.2
Iα-3	93.3 ± 4.8	91.1	-2.2	13.2
Iβ-1	62.8 ± 4.0	58.7	-4.1	9.4
Iβ-2	55.1 ± 7.7	74.7	19.6	18.7
Iβ-3	76.6 ± 5.8	80.7	4.2	15.5

lower than the spreading between the various DFT-D results ($\Delta_{\mathrm{max}} E_{\mathrm{DFT-D}}$). However, for the parallel geometries GLYCAM06 overshoots all DFT-D adsorption energies. For the GLYCAM06-optimized geometries the GLYCAM06 adsorption energies of the cellulose Iα geometries are on average 3 kJ/mol larger than for the BP86-D2-optimized geometries; GLYCAM06 adsorption energies for the Iβ geometries are on average 15 kJ/mol larger.

3.6 Method Comparison

Geometry optimizations with BP86, BP86-D2, and GLYCAM06 yield quite similar geometries. The largest differences between the three methods were observed for geometries with a vertically adsorbed glucose molecule above a glucan chain of the surface (Iα-1 and Iβ-1).

For the remaining geometries the agreement between BP86, BP86-D2, and GLYCAM06 geometries is remarkable.

But the similarity between the BP86 and BP86-D2 optimized geometries can only be observed if the adsorbate and the surface model are in close-contact to each other already at the starting geometry for the BP86 optimization. If this is not the case, an optimization with BP86, or other DFT methods that do not fully account for dispersion interactions, would produce wrong geometries due to the missing attractive dispersion interaction between the two interacting systems. For our systems the shortest distance between two carbon atoms in the starting geometry has to be smaller than 5 Å for producing reasonable BP86 optimized geometries. In general, optimizations with BP86-D2 yield geometries with the shortest distances between the surface model and the adsorbate compared to BP86 and GLYCAM06 optimizations; only for the system Iα-1 GLYCAM06 predicts a shorter distance than BP86-D2.

For the adsorption energies there are larger differences between the used methods. Three sets of optimized geometries were compared (see Tables 3.4, 3.5, and 3.6). As it can be seen in Figures 3.5, 3.6, and 3.7, the DFT adsorption energies are always only a fraction of the corresponding GLYCAM06 and DFT-D adsorption energies, more than half of the adsorption energy originated because of dispersion interactions. The agreement between DFT-D and GLYCAM06 adsorption energies is for cellulose Iα geometries better than for cellulose Iβ geometries. In most cases there is a better agreement between the GLYCAM06 and the DFT-D mean adsorption energies than between the different DFT-D methods used in this work.

Considering the relative energetically ordering of the studied geometries, all DFT-D methods are in agreement but GLYCAM06 yields a different ordering. For DFT-D methods structure Iα-3 is most stable and Iβ-2 is least stable. In comparison with the DFT-D results, GLYCAM06 slightly underestimates the adsorption energy for geometries with a large number of intermolecular hydrogen bonds (e.g., Iβ-1 in all optimizations) and it overestimates adsorption energies in systems with no intermolecular hydrogen bonds, e.g. Iβ-2. The last effect is especially important for small and rigid adsorbates. Large and flexible adsorbates, such as medium-sized carbohydrates, can form hydrogen bonds with the surface rather easy and therefore this effect will be less pronounced.

3.7 Summary

The adsorption of a single glucose molecule onto the (100) surfaces of cellulose Iα and cellulose Iβ was investigated with the GLYCAM06 force field and DFT. It was found that GLYCAM06 is able to produce adsorption geometries, which are in good agreement with dispersion-corrected DFT geometries. For adsorption energies GLYCAM06 provides a generally good description compared to dispersion-corrected DFT. In fact, it does a much better job than DFT without dispersion correction. Therefore, it is justified to

use the GLYCAM06 force field to study the adsorption process of medium-sized flexible carbohydrate adsorbates such as cellobiose and cellotetraose onto cellulose surfaces. This topic will be discussed in Chapter 4.

The difference in adsorption energies between DFT and dispersion-corrected DFT can be used as a rough estimate of dispersion interaction if using the same density functional in both calculations. It was found that the contribution of dispersion interaction to the total interaction between the adsorbate and the surface model is huge. This is the case not only for adsorption geometries without hydrogen bonds, but also for structures involving multiple hydrogen bonds. Therefore, the nature of the interaction between hydrogen-bonded molecules should be further investigated. This will be done in Chapter 5.

Chapter 4

Adsorption of Cellobiose and Cellotetraose onto Cellulose Surfaces

We have learned in Chapter 3 that GLYCAM06 provides a reasonable description of the adsorption of medium-sized carbohydrates onto cellulose surfaces. Therefore, the adsorption of a cellobiose and a cellotetraose molecule onto the (100) surfaces of cellulose Iα and cellulose Iβ is studied in this chapter. By comparing the two adsorbates the effect of the adsorbate size on the interaction with the surface is discussed. These results were recently published in the *Journal of Physical Chemistry B.*[75]

4.1 Computational Details

The adsorption of a cellobiose molecule and a cellotetraose molecule was studied with the GLYCAM06 force field. For these calculations two cellooctaose strands are used as model systems for the (100) surfaces of cellulose Iα and cellulose Iβ. The principal axis of the adsorbate (cellobiose or cellotetraose) corresponding to the smallest principal moment of inertia (1-axis) defines the molecular axis of the adsorbate. The molecular plane of the adsorbate can be defined as a plane that is perpendicular to the principal axis corresponding to the largest principal moment of inertia (3-axis). To get a representative sample of surface/adsorbate complexes, starting geometries for optimizations were constructed in the following way: At first, the adsorbate was positioned on top of a glucan strand of the surface, i.e. the molecular axis of the adsorbate is parallel to the growth direction of the glucan chains of the surface and the molecular axis of the adsorbate is aligned parallel to the surface plane. The distance between the adsorbate and the surface was chosen in a way that the smallest distances between two carbon atoms were around 4 Å. Next, the adsorbate was shifted along the growth direction of the surface by half the size of a glucose ring and the adsorbate was also positioned at various positions between the glucan chains of the surface. After that, the adsorbate was translated as described but rotated in steps of 45 degrees around the 3-axis, which leads to a no longer parallel alignment between the growth direction of the surface and the molecular axis of the adsorbate. Finally, starting geometries were created by using the previous positions but also rotating the adsorbate molecule around its molecular axis, so that the molecular plane of the adsorbate is no longer parallel to the surface plane. All starting geometries were optimized and adsorption energies were calculated with the GLYCAM06 force field as described in Section 3.1.

4.2 Adsorption of Cellobiose

In order to describe the adsorption of a cellobiose molecule onto the (100) surfaces of cellulose Iα and Iβ various starting geometries were created and optimized with the GLYCAM06 force field. Several geometrical parameters are used to characterize the alignment of the cellobiose molecule on the surface.

At first, the absolute value of the angle between the projection of the principal axis of inertia of the cellobiose corresponding to the smallest moment of inertia (1-axis) on the surface plane and the growth direction of the surface constitutes the parameter α. The cellobiose molecule is oriented exactly parallel to the growth direction of the surface if $\alpha = 0$ degrees and it is oriented exactly orthogonal if $\alpha = 90$ degrees. The alignment is called *orthogonal* if $\alpha > 45$ degrees; in this study only one group of adsorbates is oriented orthogonal.

Second, the parameter β describes the absolute value of the angle between the surface plane and the 1-axis.

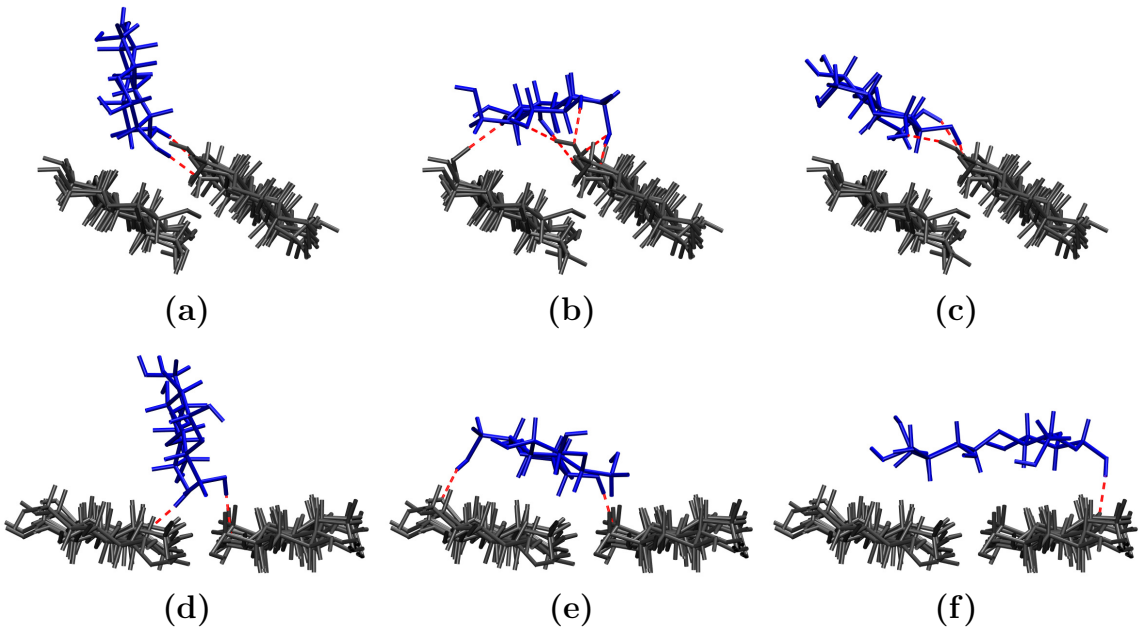


Figure 4.1: Representative structures for the cellobiose adsorption: Iα-vertical (a), Iα-parallel (b), Iα-diagonal (c), Iβ-vertical (d), Iβ-parallel (e), Iβ-parallel, orthogonal (f).[75]

Finally, the parameter γ describes the absolute value of the angle between the surface plane and the principal axis of inertia of the cellobiose molecule corresponding to the largest moment of inertia. The cellobiose molecule is oriented completely vertical to the surface plane if $\gamma = 0$ degrees and it is oriented completely parallel to the surface plane if $\gamma = 90$ degrees. The alignment is called *vertical* if γ is between 0 and 45 degrees, and *parallel* if $\gamma > 45$ degrees.

The geometry optimizations with GLYCAM06 yielded 43 stable surface/cellobiose structures, which are divided into 6 different groups based on the alignment of the cellobiose onto the surface. 17 structures were obtained for cellulose Iα, which are grouped into vertical, parallel, and diagonal structures. The diagonal structures look like an intermediate stage in crystal growth. For cellulose Iβ 26 structures were obtained and are grouped into vertical and two different parallel structures. For the parallel structures the growth direction of the cellobiose is either parallel or orthogonal to the growth direction of the surface. Figure 4.1 shows one representative geometry of each group. Pictures, adsorption energies, and structural parameters for all structures investigated are presented in Appendix A. In Table 4.1 the mean values, the standard deviations and the highest adsorption energy for each group are listed.

Parallel structures have on both surfaces higher adsorption energies than non-parallel structures, because for parallel structures more atom pairs between the surface and the adsorbate are in close-contact. Therefore, the loss of dispersion energy in non-parallel structures can not be compensated by forming hydrogen bonds.

The mean adsorption energy for the adsorption of cellobiose on the cellulose Iα surface

Table 4.1: GLYCAM06 mean adsorption energies of cellobiose (maximal values in parentheses) in kJ/mol.[75]

Adsorbate orientation	Iα	Iβ
vertical	155 ± 20 (178)	99 ± 17 (114)
parallel	209 ± 32 (238)	139 ± 32 (186)
parallel, orthogonal		120 ± 14 (141)
diagonal	180 ± 23 (206)	

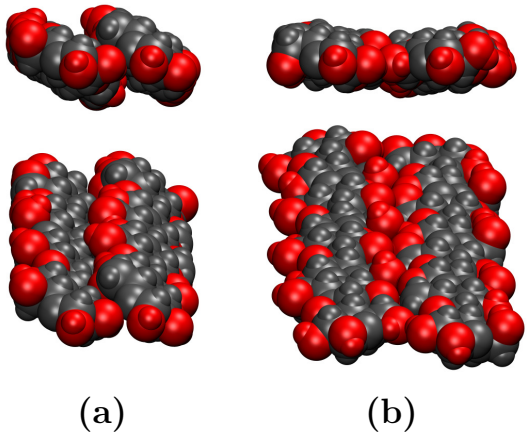


Figure 4.2: Parallel glucan molecules as model systems for the cellulose Iα(100) surface (a) and Iβ(100) surface (b).[75]

amounts to 185 ± 35 kJ/mol (calculated for 17 structures); for the adsorption on the cellulose Iβ surface it amounts to 120 ± 28 kJ/mol (calculated for 26 structures).

A reason for the strong adsorption to the cellulose Iα surface is the larger number of hydrogen bonds that can be formed, because on the cellulose Iα surface more oxygen atoms are accessible to the adsorbate than on the cellulose Iβ surface. The highest observed number of hydrogen bonds between the cellobiose and the surface model is 6 for cellulose Iα, but only 2 for cellulose Iβ. The cellulose Iα surface has a grooved shape and therefore small molecules can in general get closer to the surface compared to Iβ (see Figure 4.2).

4.3 Adsorption of Cellotetraose

For the description of the orientation of the cellotetraose molecule onto the surfaces the same parameters as for the cellobiose adsorption are used. Additionally, it is distinguished between flat and twisted structures by using

$$\sigma = \frac{I_3 - I_2}{I_3} \times 100, \tag{4.1}$$

where I_2 refers to the second largest principal moment of inertia and I_3 to the largest principal moment of inertia. The larger σ, the more flat is the cellotetraose molecule. Structures are termed as twisted structures, if an assignment to either the vertical or

Table 4.2: GLYCAM06 mean adsorption energies of cellotetraose (maximal values in parentheses) in kJ/mol.[75]

Adsorbate orientation	Iα	Iβ
vertical	249 ± 35 (289)	189 ± 31 (214)
parallel	325 ± 32 (374)	290 ± 20 (325)
twisted	231 ± 24 (251)	211 ± 27 (247)
diagonal	342	

parallel group is no longer reasonable. Structures with a σ value smaller than 6.0 are twisted structures.

For the adsorption of cellotetraose 14 structures adsorbed onto cellulose Iα and 18 onto cellulose Iβ were divided into vertical, parallel, and twisted structures. In addition, one diagonal structure was investigated for cellulose Iα. Pictures, adsorption energies, and structural parameters can be found in Appendix A.

Different numbers of stable structures of each type were found and sorted into the various groups. The number of stable parallel structures is much larger than the number of either vertical or twisted structures. In Figure 4.3, only representatives of those groups that were not found among the cellobiose structures are shown. Table 4.2 shows the mean values, the standard deviations, and the highest adsorption energies for all groups.

The highest adsorption energies for both cellulose surfaces were found for parallel adsorption; the differences between adsorption energies of vertical and twisted structures are rather small. The crystal-growth-like structure (diagonal) has one of the highest adsorption energies.

The mean energy for the adsorption of a cellotetraose molecule onto the cellulose Iα surface is 289 ± 53 kJ/mol (calculated for 14 structures); for adsorption onto the cellulose Iβ surface it is 255 ± 50 kJ/mol (calculated for 18 structures).

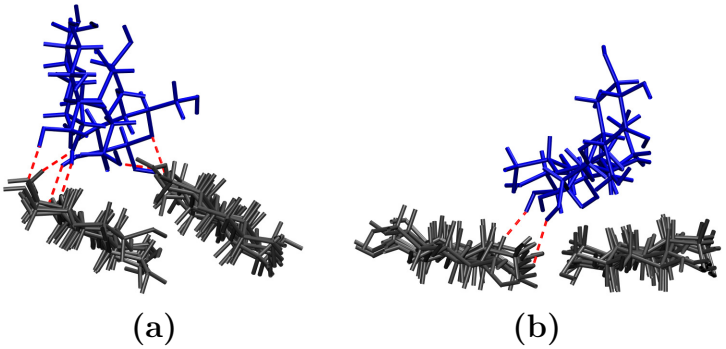


Figure 4.3: Representative geometries of the cellotetraose adsorption: Iα-twisted (a), Iβ-twisted (b). Only orientations that were not observed in the cellobiose adsorption are shown.[75]

4.4 Discussion of the Carbohydrate Adsorption

A Comparison of all adsorption energies shows, that adsorption onto the hydrophilic cellulose Iα surface is stronger than onto the hydrophobic cellulose Iβ surface. The difference becomes smaller with increasing size of the adsorbate. This can be clearly seen from the adsorption energy per glucose residue in the adsorbate (see below).

The highest adsorption energies are always observed when the glucose residues are parallel to the surface, because in this orientation the number of adsorbate atoms that are in close-contact with the surface atoms is maximized, and this leads to high adsorption energies due to large dispersion contributions. In contrast, for vertical structures, the number of hydrogen bonds is maximized but at the cost of a tremendous loss of dispersion energy, which cannot be compensated by the larger number of hydrogen bonds. With increasing size of the adsorbate, the energy difference between parallel and non-parallel alignment of the adsorbate with the glucan molecules vanishes. For small adsorbates, parallel alignment is preferred to non-parallel alignment with the glucan molecules.

Furthermore, no direct correlation between the adsorption energy and the number of intermolecular hydrogen bonds in the complex can be observed, as can be demonstrated by the Iβ/cellobiose structures: The structure shown in Figure 4.4 has no hydrogen bonds and an adsorption energy of 127 kJ/mol, the structure with the highest adsorption energy of 186 kJ/mol has two hydrogen bonds. The difference between the highest and lowest adsorption energy for structures with no hydrogen bond amounts to 38 kJ/mol; for structures with one hydrogen bond the corresponding difference is 100 kJ/mol, while for structures with two hydrogen bonds it is 109 kJ/mol. These differences are mainly due to the different dispersion interactions in these complexes.

That the dispersion contributions become more important for larger adsorbates can be seen from the mean adsorption energies per glucose residue: for cellobiose, it is 93 kJ/mol on the hydrophilic Iα surface and 60 kJ/mol on the hydrophobic Iβ surface; for cellotetraose, the values are 72 kJ/mol and 64 kJ/mol, respectively. We see that, with increasing size of the

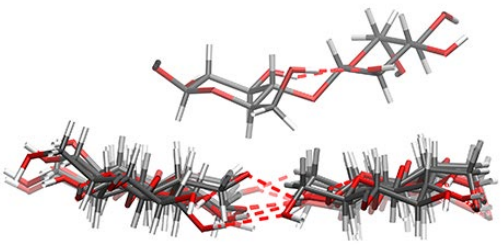


Figure 4.4: Structure Iβ+cb-po-7.

adsorbate, the adsorption energies become more similar, regardless of surface hydrophilicity. Averaging over the Iα and the Iβ values yields 77 kJ/mol for cellobiose and 68 kJ/mol for cellotetraose; this suggests that an adsorption energy between 70 and 75 kJ/mol could be a reasonable estimate for the adsorption energy per glucose residue for carbohydrates on any kind of cellulose I surface.

4.5 Summary

In this chapter the adsorption of cellobiose and cellotetraose onto two different cellulose model surfaces was discussed based on description with the GLYCAM06 force field. It was found that the adsorption energy increases with increasing adsorbate size. Adsorption onto the cellulose Iα surface is in general stronger than the adsorption onto the cellulose Iβ surface because the Iα surface is more hydrophilic. But with increasing adsorbate size the differences in adsorption energies between the hydrophilic and the hydrophobic surface are shrinking. This suggests that for much larger adsorbates the surface type is no longer an important factor and dispersion interactions dominate. An estimate of an averaged adsorption energy per glucose residue on any cellulose I surface was given, which amounts to 70-75 kJ/mol. It was also found that adsorption parallel to the surface is stronger than vertical adsorption.

Also no direct correlation between the adsorption energy and the number of hydrogen bonds between the adsorbate and the surface can be found. This also suggests that hydrogen bonding is not solely responsible for the adsorbate-surface interaction. But still, the amount of dispersion interactions can not be quantified because the force field approach allows no reasonable decomposition of interaction energies.

Chapter 5

Water and Alcohol Dimers

As already discussed in the previous chapters, the interaction between two hydrogen-bonded molecules can not be directly related to the number of hydrogen bonds between them, ubiquitous dispersion interactions must also be important. Generally, hydrogen bonding is often attributed to only electrostatic interactions. To determine the *nature* of the interaction between two hydrogen-bonded molecules, small model systems are studied with symmetry-adapted perturbation theory (SAPT). The used model systems consist of the water dimer and several progressively larger alcohol dimers. The interaction between the two monomers is discussed in terms of electrostatic, exchange, induction, and especially dispersion interactions. These results were recently published in *Chemistry - A European Journal.*[97]

5.1 Computational Details

Various small hydrogen-bonded dimers ranging from the water dimer to the *tert*-butanol dimer were optimized using the second-order approximate coupled cluster method (CC2);[98] calculations were performed with the cc-pVDZ basis set[99] using the program package TURBOMOLE.[78–81]

For every dimer the interaction energy was calculated with symmetry-adapted perturbation theory based on DFT description of the monomers [SAPT(DFT)] for several O–O distances of the oxygen atoms involved in the hydrogen bond. Based on the CC2 optimized dimer geometries rigid displacements of the monomers were performed in 0.1 Å steps along the line spanned by the oxygen atoms of the hydrogen bond. Any relaxation of the monomers was prohibited. The coordinates of each monomer were transformed in a way that one oxygen atom was placed at the coordinate origin and the other one on the x-axis. Therefore, this displacements consisted of simply changing the x-coordinates of every atom in one monomer.

All SAPT(DFT) calculations of the interaction energies were performed by using the SAPT2012 suite of programs[100] with the density fitting approach.[101–104] The software packages DALTON[105] and GAMESS[106] were used to calculate DFT orbitals of monomers and integrals containing auxiliary functions, respectively. The PBE0 functional[64–66] was applied together with the Fermi-Amaldi-Tozer-Handy asymptotic correction.[107] All ionization potentials, which were needed for the asymptotic correction were taken from Ref. 108. For all calculations the aug-cc-pVTZ basis[109] was used together with the corresponding auxiliary basis from Ref. 110. The δ_{HF} terms were computed using the aug-cc-pVDZ basis set due to the fact, that calculations in aug-cc-pVTZ were becoming very time consuming for the biggest systems investigated, because the SAPT(HF) codes within SAPT2012 do not utilize density fitting. Hartree-Fock energies and HF-level SAPT components converge fast with the basis set size. Therefore, the use of the smaller basis set for δ_{HF} had only negligible effect on the calculated interaction energies, which are discussed in this chapter.

SAPT(DFT) calculations yield a variety of interaction components, which were already discussed in Section 2.3. In order to interpret the calculated interaction energies, the different contributions to the total interaction energy E_{SAPT} were collected into four components: electrostatics (E_{elst}), induction (E_{IND}), dispersion (E_{D}) and exchange interactions (E_{exch}). E_{IND} includes the induction, the exchange-induction and the δ_{HF} terms. E_{D} consists of the dispersion and the exchange-dispersion component. The sum of E_{elst}, E_{IND}, and E_{exch} is referred to as the dispersionless interaction energy E_{DL}.

In addition, also the atom-atom dispersion function D_{as} was used.[70,71] With D_{as} it is possible to determine approximate dispersion contributions from groups of atoms.

5.2 Total Interaction Energies

The following systems are investigated in this chapter: water dimer $(H_2O)_2$, methanol dimer $(MeOH)_2$, ethanol dimer $(EtOH)_2$, *n*-propanol dimer $(n\text{PrOH})_2$, *n*-butanol dimer $(n\text{BuOH})_2$, isopropanol dimer $(i\text{PrOH})_2$, and the *t*-butanol dimer $(t\text{BuOH})_2$. Minimum configurations of these dimers were calculated by using the CC2 method. The optimized structures are shown in Figure 5.1. Note, that some of them might not be global minima, because the topologies of the minima were kept as similar as possible so that a reasonable comparison between the structures is possible. The coordinates of all optimized geometries are listed in Appendix B.

Figure 5.2 shows the total interaction energies E_{SAPT} of all investigated systems as functions of the O–O distance. These potential curves were obtained by a rigid variation of the O–O distances. It can be seen that the minima of the E_{SAPT} curves are all located very close to 2.9 Å. In Table 5.1 the total interaction energies and the O–O distances at the minima are shown together with the dispersionless interaction energies E_{DL} and the O–O distances at the corresponding minima (see below). The largest difference in the O–O distances at the minimum of E_{SAPT} was observed between the two smallest systems and it amounts to 0.11 Å, whereas for the remaining dimers the values differ by only 0.01 Å. The closeness of these minima will simplify the analysis of the results. The SAPT(DFT) minima have slightly larger O–O distances than the CC2 optimized dimers, the former are on average 0.08 Å larger than the latter. This is due to the fact that the optimization with CC2 did not include the counterpoise correction, whereas SAPT(DFT) interaction energies do not contain the basis set superposition error at all. For the methanol dimer the difference between the CP-corrected CC2 minimum and the SAPT(DFT) minimum is smaller than 0.01 Å.

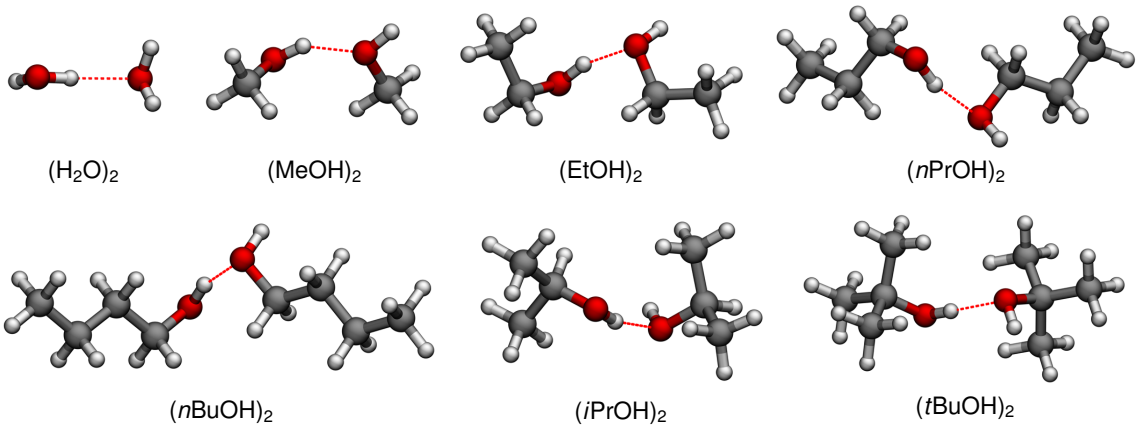


Figure 5.1: Optimized geometries of the studied dimers.[97]

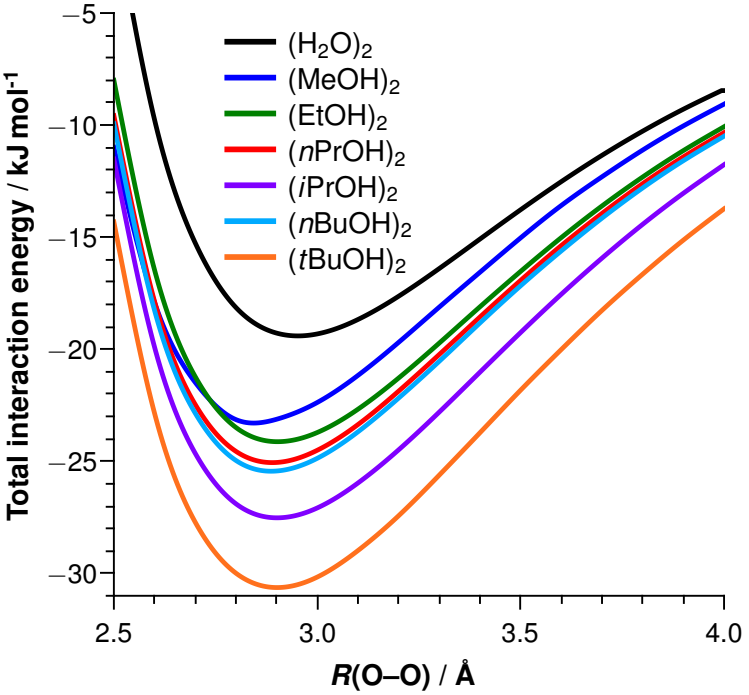


Figure 5.2: Total interaction energies.[97]

Table 5.1: O–O distances in Å and interaction energies in kJ mol^{-1} at the local minima of the respective E_{SAPT} and E_{DL} curves.[97]

System	E_{SAPT}	O–O$_{\mathrm{SAPT}}$	E_{DL}	O–O$_{\mathrm{DL}}$
$(H_2O)_2$	-19.4	2.95	-11.9	3.19
$(MeOH)_2$	-23.3	2.84	-10.9	3.19
$(EtOH)_2$	-24.1	2.90	-10.0	3.28
$(n\mathrm{PrOH})_2$	-25.1	2.89	-10.2	3.27
$(n\mathrm{BuOH})_2$	-25.4	2.89	-10.0	3.27
$(i\mathrm{PrOH})_2$	-27.5	2.90	-8.3	3.37
$(t\mathrm{BuOH})_2$	-30.6	2.90	-8.7	3.40

The binding energy is the interaction energy at the minimum of E_{SAPT} with sign reversed. It increases with dimer size from 19.4 to 30.6 kJ mol^{-1}, which corresponds to an increase by 58%. This is a significant variation given the fact that the range of medium-strength hydrogen bonds is often assumed to be between 20 and 50 kJ mol^{-1}. This shows immediately that the simplest physical interpretation of hydrogen bonding, which utilizes only the properties of the central moiety O–H$\cdots$O, fails for the presented sequence of dimers. This moiety is almost the same in all the studied dimers, so no significant variation of the binding energy should be observed.

As can be seen in Figure 5.2, the increase of the binding energy with dimer size is not uniform and can therefore not depend only on the size of the monomers; the shape of the monomers is also important. The average increase of the binding energy amounts

to 1.9 kJ mol^{-1}, but the increase between $(H_2O)_2$ and $(MeOH)_2$ is with 3.9 kJ mol^{-1} much larger than the average. When moving to $(EtOH)_2$, the increase amounts to 0.8 kJ mol^{-1}, which is about twice smaller than the average. Going to $(n\text{PrOH})_2$, the increase is virtually the same with 1.0 kJ mol^{-1}. Finally, moving from $(n\text{PrOH})_2$ to $(n\text{BuOH})_2$ yields an increase of only 0.3 kJ mol^{-1}. In contrast to that, the increases in the sequence $(EtOH)_2 \rightarrow (i\text{PrOH})_2 \rightarrow (t\text{BuOH})_2$ are 3.4 kJ mol^{-1} and 3.1 kJ mol^{-1}, respectively, which are similar in size to the first step. Obviously, there is no chance that these observed variations can be explained by only considering the central moiety O–H$\cdots$O. In fact, we need to consider all components of the interaction energy.

5.3 Dispersionless Components of the Total Interaction Energy

5.3.1 Electrostatic Interactions

All components of the total interaction energy E_{SAPT} at the minima of the E_{SAPT} curves are listed in Table 5.2. The potential curves of the dispersionless components as functions of the O–O distance are plotted in Figure 5.3. We start the discussion with the electrostatic interactions, because a lot of interpretations of hydrogen bonding only use these interactions. Some models only consider the central moiety and in biochemical applications purely electrostatic models are often used. Therefore, we will discuss the question whether (complete) electrostatic interactions alone can describe the observed variation of the hydrogen bond stabilization in the presented sequence of dimers. The electrostatic interaction energy is the energy of the Coulomb interactions of permanent charge distributions on monomer A with permanent charge distributions on monomer B. Since the electrostatic interactions of uncharged molecules with a nonzero dipole moment decay with the third inverse power of the intermolecular separation, there is no minimum observed on the electrostatic energy curves. Therefore an electrostatic model is always tacitly supplemented by a simple model of the repulsive exchange forces. The simplest model is just a hard-sphere repulsion at atomic van der Waals radii.

Table 5.2: Energy components in kJ mol^{-1} at the minima of the E_{SAPT} curves.[97]

System	E_{elst}	E_{exch}	E_{IND}	E_{DL}	E_{D}	E_{Das}	E_{SAPT}
$(H_2O)_2$	-32.4	31.3	-8.8	-10.0	-9.4	-9.4	-19.4
$(MeOH)_2$	-35.1	38.6	-10.2	-6.7	-16.5	-17.2	-23.2
$(EtOH)_2$	-36.4	42.0	-11.2	-5.5	-18.6	-19.4	-24.1
$(n\text{PrOH})_2$	-36.9	42.8	-11.6	-5.7	-19.4	-20.2	-25.1
$(n\text{BuOH})_2$	-36.8	43.0	-11.7	-5.4	-20.1	-20.9	-25.4
$(i\text{PrOH})_2$	-43.0	54.1	-13.3	-2.1	-25.4	-26.1	-27.5
$(t\text{BuOH})_2$	-46.6	59.3	-14.5	-1.8	-28.8	-29.1	-30.6

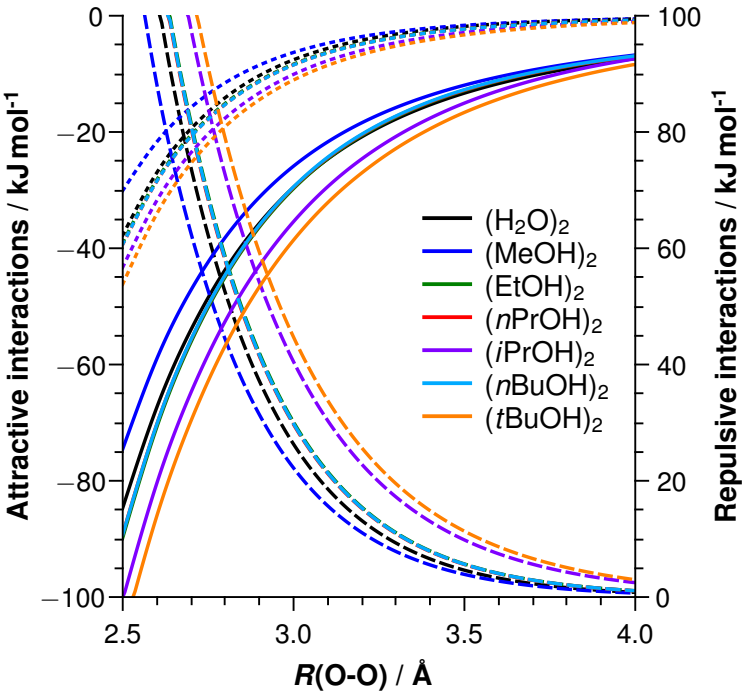


Figure 5.3: Dispersionless components of the total interaction energy as a function of the O–O distance.[97]

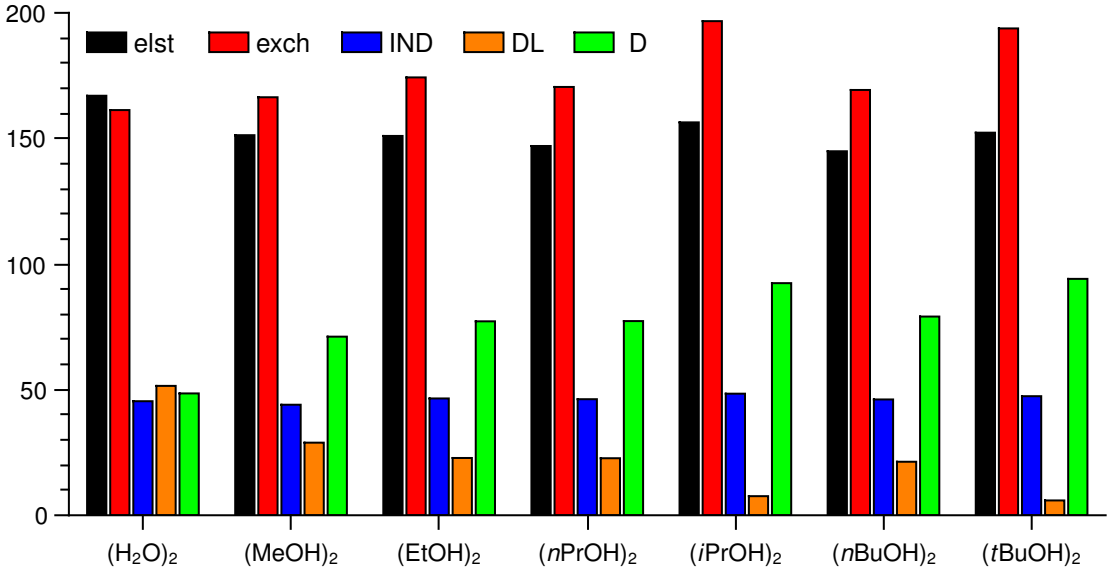


Figure 5.4: Absolute values of percentage contributions to total interaction energies at the minima of the E_{SAPT} curves.[97]

Absolute values of percentage contributions of all interaction components to the total interaction energy are shown for all studied dimers at the minima of the E_{SAPT} curves in Figure 5.4. It can be seen that the electrostatic interaction is surprisingly not the largest contribution to E_{SAPT}, but only the second largest. Except for the water dimer, the largest contribution is always the exchange-repulsion. Nevertheless, the contribution of the electrostatic interaction is between 145% and 167% of the total interaction energy. Note that this percentage contribution is remarkably constant. Nevertheless, the electrostatic contribution increases in the presented sequence of dimers (from $(H_2O)_2$ to $(t\mathrm{BuOH})_2$) by $14.2\,\mathrm{kJ\,mol^{-1}}$ in absolute terms or by a factor of 1.4 in relative terms. This absolute increase is fairly close to the corresponding change of E_{SAPT}, which amounts to $11.2\,\mathrm{kJ\,mol^{-1}}$.

In order to find out how well the electrostatic interactions can actually predict the stabilization of hydrogen-bonded systems, Figure 5.5 shows the correlation between all possible differences of electrostatic energies and all possible differences of the total interaction energies at the minima of the E_{SAPT} curves. One can see, that there is only modest correlation between differences of the electrostatic energy ΔE_{elst} and differences of the total interaction energy ΔE_{SAPT}. The observed ratios $\Delta E_{\mathrm{elst}}/\Delta E_{\mathrm{SAPT}}$ vary between -0.3 and 3.0. Only one negative ratio was observed, which belongs to the difference between $(n\mathrm{PrOH})_2$ and $(n\mathrm{BuOH})_2$. The ratios are highest for differences between systems, in which at least one is either $(i\mathrm{PrOH})_2$ or $(t\mathrm{BuOH})_2$ (blue points); for these systems the correlation is worse than for the others. So only for structural homogeneous systems, like the linear hydrogen

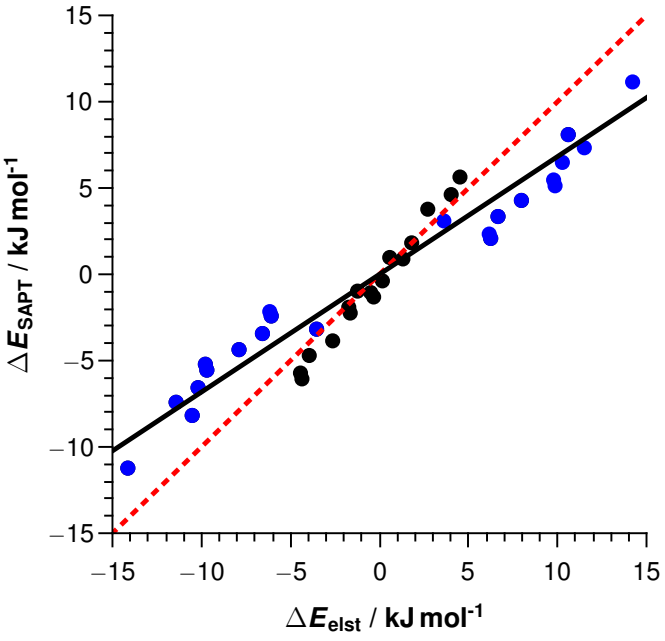


Figure 5.5: Correlation of all possible differences of total interaction energies ΔE_{SAPT} versus those of electrostatic energies ΔE_{elst}. The linear regression is shown by the solid line. Ideal correlation is indicated with the red dashed line. Differences containing $(i\mathrm{PrOH})_2$ or $(t\mathrm{BuOH})_2$ are shown in blue, all others in black. [97]

bonded systems (black points), the electrostatic model might be able to describe changes in stabilization of hydrogen bonds, for all others this is not the case.

5.3.2 Induction Interactions

The next component to discuss is the induction interaction, which is the interaction between permanent and induced multipoles. For neutral systems induction energies decay as sixth inverse powers of the intermonomer separation and are always negative. Therefore, an electrostatic plus induction model needs additionally an implicit hard-sphere repulsion. It can be seen in Figure 5.4 that the induction interactions in the presented sequence of dimers are always a quite constant fraction of the total interaction energy; the percentage contributions vary between 44% and 48%. Between $(H_2O)_2$ and $(t\mathrm{BuOH})_2$ the induction energy increases by 5.7 kJ mol^{-1} in magnitude, which corresponds to a factor of 1.6. Adding the induction contribution to the electrostatic one makes the predictions of the variation of the total interaction energies even worse. The overall change of the sum of electrostatics and induction amounts to -19.9 kJ mol^{-1}, which is by a factor of 1.8 larger than the overall change of the total interaction energy. The ratios $\Delta E_{\mathrm{elst+IND}}/\Delta E_{\mathrm{SAPT}}$ calculated for all possible differences are between 0.2 and 3.7. Therefore, the electrostatics plus induction model is a little worse than the pure electrostatics one. It can be seen in Figure 5.3 that the induction interactions are decaying much faster than the electrostatic interactions. So at larger separations than those discussed here, the induction will be even less important.

5.3.3 First-Order Exchange Interactions

Exchange interactions are a quantum mechanical phenomenon related to the Pauli exclusion principle. These interactions decay exponentially with increasing distance. Since electron densities also show this exponential decay, the magnitude of the exchange interactions can be related to an overlap of electron densities.

Table 5.2 and Figure 5.4 show that the exchange contributions are always destabilizing and have percentage contributions between 161% and 197%. Except for the water dimer, the exchange has the largest percentage contributions. In the water dimer the percentage contribution of electrostatics is slightly higher than that of the exchange, for the remaining systems, the ratios of magnitudes of exchange and electrostatic interactions are between 1.1 and 1.3. Between $(H_2O)_2$ and $(t\mathrm{BuOH})_2$ the exchange energy increases in magnitude by 28 kJ mol^{-1}, which corresponds to a factor of 1.9. Note, that this is the largest absolute increase of all components. The exchange repulsion plays an important role in the studied systems and it is obvious that it is not possible to model this large exchange effect accurately by just using a hard-sphere approximation.

The importance of the exchange interaction is increasing relative to electrostatic interactions with increasing dimer size. This can be rationalized as follows: The latter interactions originate to a large extent from charge distributions on the central moiety and its vicinity

since alkyl groups are fairly nonpolar. Thus, the increase of the dimer size leads to an increase of the magnitude of the electrostatic energy by a factor of 1.4 only. In contrast, significant exchange interactions take place whenever two atoms come close to each other. As seen in Figure 5.1, there are several such close contacts in larger dimers, in particular in those with branched monomers. The exchange interactions are the largest for $(i\mathrm{PrOH})_2$ and $(t\mathrm{BuOH})_2$. This explains the relative increase by a factor of 1.9 for the exchange energies, which is much larger than that for the electrostatic energies.

5.3.4 Dispersionless Interaction Energy

One may now ask how well the sum of the electrostatic, induction, and exchange energies predicts the stabilization of hydrogen bonds. This is again a very important question since numerous papers have been published investigating hydrogen bonds at the supermolecular Hartree-Fock and DFT level. Both methods are unable to reproduce dispersion interactions (exact DFT would be able to but such a method is unknown), so the predictions of these methods should be similar to those given by the dispersionless energies.

Table 5.2 and Figure 5.4 show that the predictions of the dispersionless approach are even worse than those of the purely electrostatic model. In fact, E_{DL} generally decreases in magnitude when the total interaction increases, the contribution to the total interaction energy drops from 52% in $(H_2O)_2$ to 6% in $(t\mathrm{BuOH})_2$, or put in other words, the dispersionless method predicts for this dimer a binding energy that is 17 times too small in magnitude.

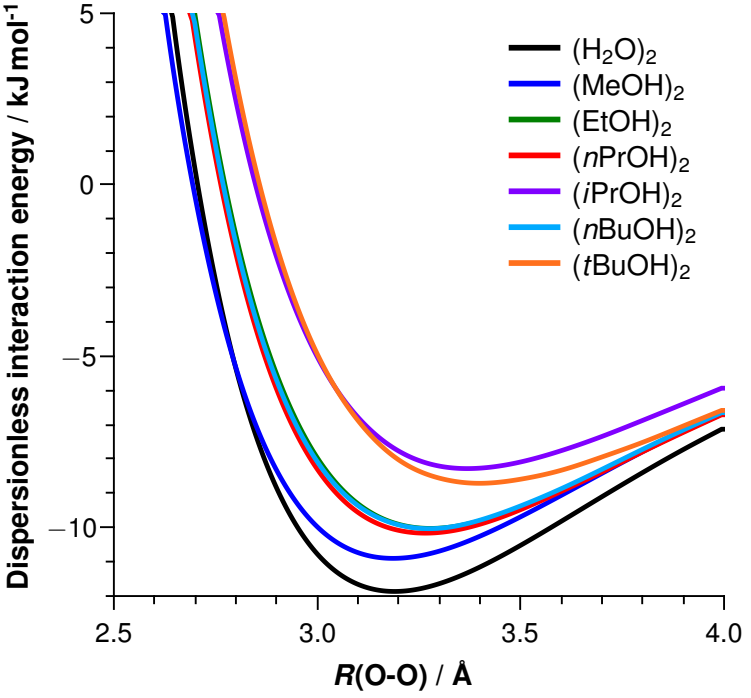


Figure 5.6: Dispersionless interaction energies.[97]

One could argue that the poor performance of the dispersionless method is due to the fact that we performed the comparisons at the minima of the E_{SAPT} curves rather than at the minima of the E_{DL} curves. However, it is shown in Table 5.1 that using the minima of the E_{DL} curves does not help. The values of E_{DL} at these minima are almost constant, so that this method cannot predict the observed trends. The E_{DL} minimum O–O distances are also significantly larger than those of the real minima, and the minima of the branched dimers are always larger than those of the linear ones (see Table 5.1). The dependence of E_{DL} on the O–O distance is presented in Figure 5.6. The Comparison with Figure 5.2 shows again a complete failure of the dispersionless approach for the investigated sequence of dimers. It can be clearly seen that one cannot describe the stabilization of the investigated dimers without considering dispersion interactions.

The components of E_{DL} are plotted in Figure 5.3. This figure shows that for distances smaller than 2.9 Å, E_{DL} becomes quickly dominated by the first-order exchange interaction. For distances larger than 2.9 Å, the electrostatic component gradually becomes more and more important, and at O–O distances of 4.0 Å and larger, it dominates E_{DL}. Since at such distances the dispersion interaction is also small in magnitude, electrostatic interactions dominate the whole interaction energy. This is an expected behavior since electrostatics is the slowest decaying component of the total interaction energy in the investigated dimers.

5.4 Dispersion Interactions

Since we were not able to rationalize the dependence of the stabilization energy of the investigated dimers by using the electrostatic, induction, and exchange interactions, it is clear that one has to take into account also the dispersion interactions, which are very rarely discussed in the context of the physical interpretation of hydrogen bonds. The dispersion interactions, similarly as the exchange interactions, are a pure quantum phenomenon and result from correlations of motions of electrons in monomer A with those in monomer B. The interaction between fluctuating dipoles decays as the sixth inverse power of the intermonomer separation. Dispersion energies are always negative. Table 5.2 and Figure 5.4 show that the dispersion energy gives the third largest in magnitude contribution to the total interaction energy, varying from 48% to 94%. Whereas the magnitude of the dispersion energy is only 30% of the value of the first-order exchange energy for $(H_2O)_2$, the same ratio is 49% for $(t\mathrm{BuOH})_2$. Despite the modest size, the dispersion energy is the component that increases fastest with increasing dimer size: by 19.4 kJ mol^{-1} or a factor of 3.1 between $(H_2O)_2$ and $(t\mathrm{BuOH})_2$. Therefore, the absolute change is the second largest and the ratio is the largest, significantly larger than the ratio of 1.9 for the exchange contributions. The differences of the total interaction energy are reasonably well correlated with the differences of the dispersion energy: the ratios $\Delta E_{\mathrm{D}}/\Delta E_{\mathrm{SAPT}}$ over all possible differences vary in a small interval between 0.8 and 2.5. Thus, it is the dispersion energy which best predicts the increases of the total interaction energies with increasing dimer size. Nevertheless, it is

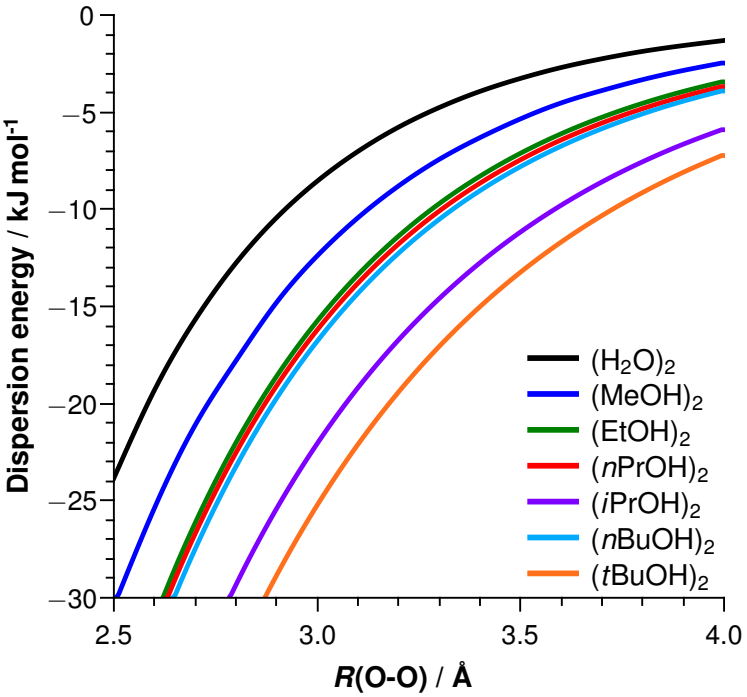


Figure 5.7: Dependence of the dispersion energy on the O–O distance.[97]

also clear that one has to consider all interaction energy components in order to get really reliable predictions.

Since for $(t\text{BuOH})_2$ the dispersion energy is -28.8 kJ mol^{-1} and the total interaction energy is -30.6 kJ mol^{-1}, one may be tempted to say that this dimer is bound by dispersion. However, we believe that this is not the right point of view since the electrostatic and first-order exchange components are actually larger in magnitude than the dispersion component. Nevertheless, one can definitely say that the dispersion interaction plays a crucial role in stabilization of all discussed hydrogen-bonded dimers.

Figure 5.7 shows the dependence of the dispersion energy on the O–O distance. As one can see, the curves are virtually parallel. Note that the curves for $(\text{EtOH})_2$, $(n\text{PrOH})_2$, and $(n\text{BuOH})_2$ almost overlap. Up to $(\text{EtOH})_2$ the increase of the substituent sizes increases the dispersion interaction, but the terminal methyl groups in $(n\text{PrOH})_2$ and the terminal ethyl groups in $(n\text{BuOH})_2$ contribute only very little to the dispersion interaction. In other words, the dispersion interactions in $(n\text{PrOH})_2$ and $(n\text{BuOH})_2$ take place almost exclusively within a sphere corresponding to the size of $(\text{EtOH})_2$, while the contributions of atoms outside this sphere can be neglected.

To analyze in more detail how various regions of monomers contribute to the dispersion energy, we used the atom-atom dispersion function D_{as}, which allows calculations of approximate dispersion contributions between a pair of atoms or between groups of atoms. As can be inferred from Table 5.2, this function predicts dispersion energies in excellent agreement with the E_{D} values from SAPT(DFT). All values of the dispersion interaction

within the O–H$\cdots$O moiety at the minima of the E_{SAPT} curves, i.e., the dispersion interaction between the hydroxyl group of the donor molecules and the O atom of the acceptor molecule, lie in the narrow interval from -6.7 to -7.7 kJ mol^{-1}, giving an average value of -7.4 kJ mol^{-1}. All other dispersion contributions amount to 28% of the total dispersion energy for $(H_2O)_2$ and vary between 55% and 74% for the other systems. Thus, except for $(H_2O)_2$ in which this effect obviously cannot be large because there are only a few electrons beyond the central moiety, the non-central dispersion interactions dominate, as long as the atoms are within the sphere discussed above. In $(t\text{BuOH})_2$ the non-central dispersion interactions are especially important: they amount to 74% of the total dispersion energy, which itself recovers 94% of the total interaction energy.

5.5 Comparison of Linear and Branched Systems

As already discussed for Table 5.2 and Figure 5.4, our sequence of dimers can be better analyzed if the dimers are grouped into one sequence consisting of linear monomers and another sequence of branched ones. Also, it is interesting to compare interactions of linear and branched monomers. Figure 5.8 shows such a comparison on the example of $(n\text{BuOH})_2$ and $(t\text{BuOH})_2$. As one can see, all contributions are for all O–O separations larger in magnitude in $(t\text{BuOH})_2$ than in $(n\text{BuOH})_2$. The relative increases at the minima of the E_{SAPT} curves are 43%, 38%, 27%, and 24% for dispersion, exchange, electrostatic, and induction energies, respectively. The reason for the stronger increase of the two former components is that there are several close-contact atoms in $(t\text{BuOH})_2$, not present in $(n\text{BuOH})_2$, which contribute to the dispersion and the exchange interactions. Since these fragments are quite nonpolar, changes of the electrostatic and induction energies are less pronounced.

Table 5.3 shows the contributions of various groups to the dispersion energy of $(n\text{BuOH})_2$ and $(t\text{BuOH})_2$ at the minima of the E_{SAPT} curves calculated using D_{as}. Note that the OH donor group is located in monomer A and monomer B contains the acceptor atom O. The

Table 5.3: Contributions to the dispersion energy in kJ mol^{-1} calculated using D_{as} for different groups of the two butanol structures at the minima of the E_{SAPT} curves. The two monomers are denoted with A and B. [a] without OH group. [b] without O. [c] CH_3 group that is most remote to the other monomer. [97]

Group	$(n\text{BuOH})_2$	$(t\text{BuOH})_2$
E_{Das}	-20.9	-29.1
OH(A)$\cdots$O(B)	-7.5	-7.6
A$^{[a]}\cdots$B$^{[b]}$	-4.5	-6.7
OH(A)$\cdots$B$^{[b]}$	-6.4	-12.0
A$^{[a]}\cdots$O(B)	-2.5	-2.9
CH_3(A)$^{[c]}\cdots$B	-0.2	-0.8
CH_3(B)$^{[c]}\cdots$A	-0.1	-0.6

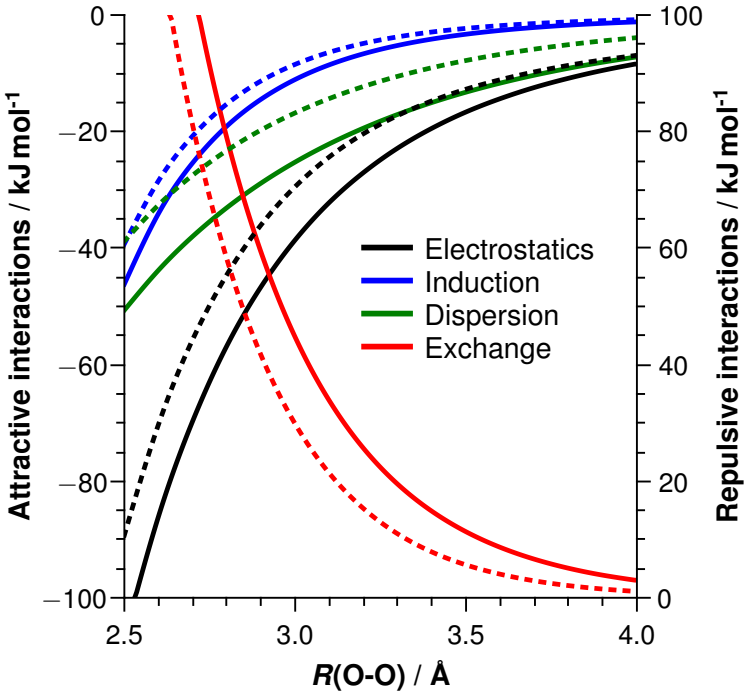


Figure 5.8: Dependence of all contributions to the interaction energy of $(n\text{BuOH})_2$ (dashed lines) and $(t\text{BuOH})_2$ (solid lines) on the O-O separation.[97]

total dispersion interaction for $(t\text{BuOH})_2$ is by 8.2 kJ mol^{-1} stronger than for $(n\text{BuOH})_2$. If one considers only the central moiety O–H$\cdots$O, the corresponding dispersion energies for $(n\text{BuOH})_2$ and $(t\text{BuOH})_2$ are almost the same: -7.5 and -7.6 kJ mol^{-1}, respectively. Omitting the whole O–H$\cdots$O group in the calculations yields for $(n\text{BuOH})_2$ a dispersion energy of only -4.5 kJ mol^{-1} and for $(t\text{BuOH})_2$ it amounts to -6.7 kJ mol^{-1}. This means that in $(t\text{BuOH})_2$ the contribution of the dispersion interaction between the alkyl groups is of comparable magnitude to the O–H$\cdots$O contribution. For $(t\text{BuOH})_2$, the largest contribution to the dispersion energy, -12.0 kJ mol^{-1} or 41%, originates from the interaction of the alkyl group of monomer B with the OH group of monomer A, so that this term clearly dominates. The corresponding term in $(n\text{BuOH})_2$ is only half as large (31%), so that it is comparable to the contribution of the central moiety. Note that the sum of the second to fifth numerical rows in the table is equal E_{Das} (modulo rounding errors). The dispersion energy contributions between the most remote CH_3 group of monomer A with all atoms of monomer B and vice versa are quite small in magnitude due to the large distances, but for $(t\text{BuOH})_2$ these interactions are on average still 5 times stronger than for $(n\text{BuOH})_2$.

Table 5.3 demonstrates clearly that: first, the dispersion energy is proportional to the number of electrons that can be dynamically polarized, thus, dispersion increases with increasing size of the molecules; second, the dispersion energy decreases with increasing distance between the interacting electrons. It appears that a sphere of the size of $(\text{EtOH})_2$ provides an approximate boundary for the balance of the two trends: if with the increasing size of the substituents all atoms stay within this sphere, their contribution to the dispersion

energy is significant, if atoms are positioned outside the sphere, their contribution could be neglected.

5.6 Interactions at Distances Larger than the Equilibrium

So far mainly the interactions at the minima of the E_{SAPT} curves have been discussed. But what happens at distances, which are larger than the equilibrium distance? To answer that question, the absolute values of the percentage contributions of all interaction components to the total interaction energies are shown for $(H_2O)_2$ and $(t\mathrm{BuOH})_2$ at three O–O distances (R) in Figure 5.9.

The O–O distance $R = 2.9$ Å is very close to the equilibrium distances of both systems. It can be seen that in both cases the exchange has the highest and the induction the lowest percentage contribution. For $(H_2O)_2$ the contribution of the electrostatic interaction amounts to 185%, while for $(t\mathrm{BuOH})_2$ it is 152%. The dispersion interaction contributes for $(H_2O)_2$ only 54% but for $(t\mathrm{BuOH})_2$ already 94%.

When moving to $R = 3.5$ Å all percentage contributions are decreasing. It can be seen that the exchange contribution decreases much faster than the contribution of the electrostatic interaction. For $(H_2O)_2$ the contribution of electrostatics is now 97%, while the exchange amounts to only 34%. For $(t\mathrm{BuOH})_2$ the difference between electrostatics and exchange is much smaller, the electrostatic interaction contributes 76% and the exchange interaction still 52%. As can be seen, the induction interaction decreases much faster than the dispersion interaction. For $(H_2O)_2$ the contributions of the induction and dispersion interactions amount to 13% and 24%, respectively. For $(t\mathrm{BuOH})_2$ the induction contribution is quite the same as for $(H_2O)_2$ (15%), while the dispersion contribution is still large with 61%. So at $R = 3.5$ Å one can say that for $(H_2O)_2$ the electrostatic interaction is most important. In contrast, for $(t\mathrm{BuOH})_2$ also the exchange and dispersion interactions are still important.

At $R = 4.0$ Å the electrostatics contributes for $(H_2O)_2$ 88% to the total interaction energy, while the second largest contribution is only 16% (dispersion). Therefore, one can say that the electrostatic interaction clearly dominates the interaction between the two water molecules at an O–O distance of 4.0 Å. In contrast, this is not the case for $(t\mathrm{BuOH})_2$. There the electrostatic interaction contributes 61% and the dispersion interaction contributes still 52%. So at this distance the dispersion is still of comparable strength to the electrostatic interaction. Therefore, for large and bulky dimers dispersion interactions can play a crucial role even at distances significantly larger than the equilibrium distance.

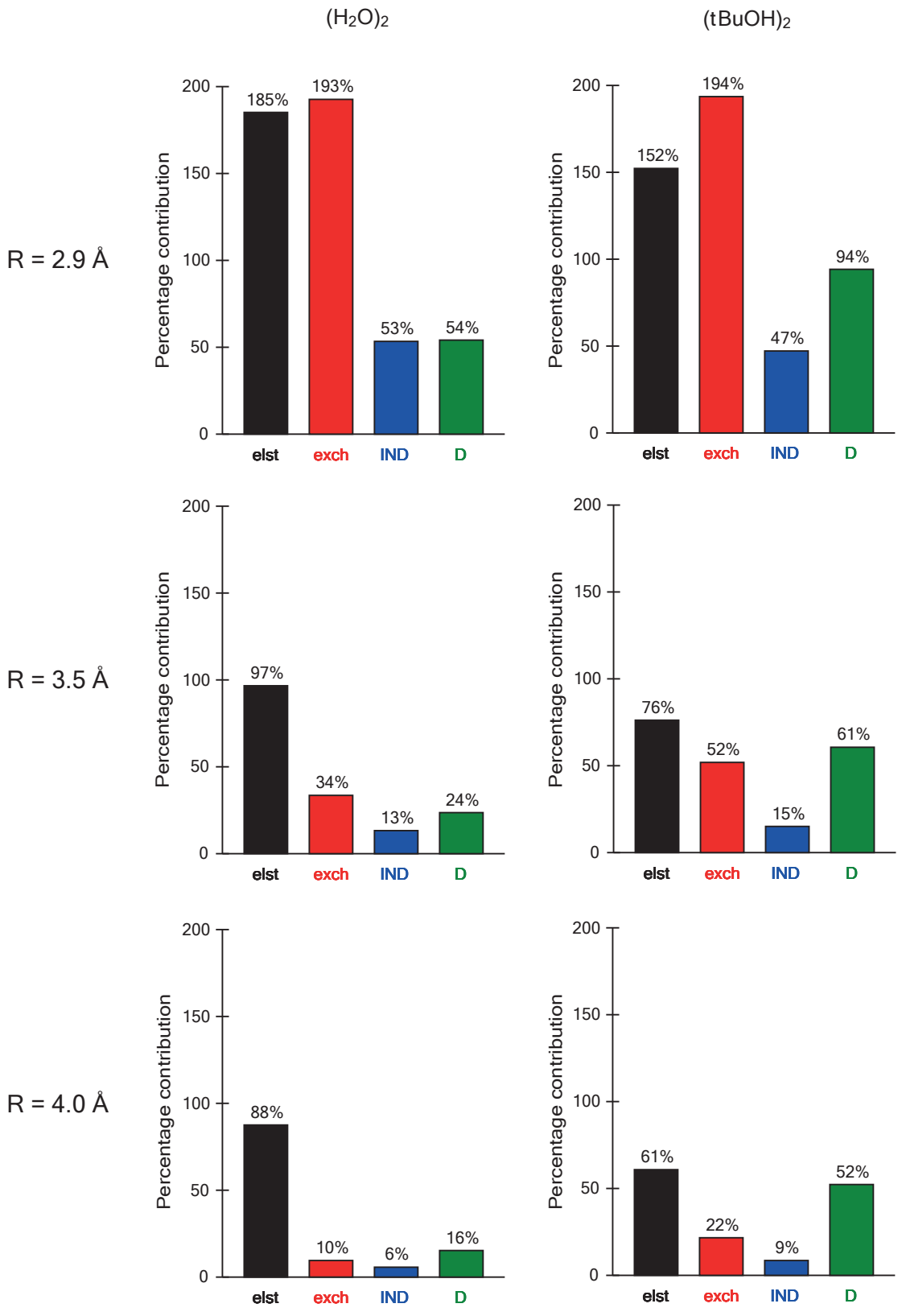


Figure 5.9: Absolute values of the percentage contributions to the total interaction energies for $(H_2O)_2$ and $(t\text{BuOH})_2$ at O–O distances (R) 2.9 Å, 3.5 Å, and 4.0 Å.

5.7 Summary

The intermolecular interactions between small hydrogen-bonded dimers were studied by using symmetry-adapted perturbation theory. Interaction energies were decomposed into electrostatic, exchange, induction, and dispersion interactions. At first the water dimer was studied, followed by consecutive alcohol dimers. In the sequence of systems studied in this chapter, the interaction energy increases by 58% when moving from the smallest to the largest and bulkiest system. Except for the water dimer, exchange interactions are always largest in magnitude and can be up to 27% larger than corresponding electrostatic interactions at equilibrium geometries. It was found that dispersion contributions exhibit the fastest variation with dimer size.

It was discussed whether it is possible to describe the observed trends of the total interaction energy with increasing system size at the equilibrium geometries by only considering the electrostatics or other subsets of components. The general trends of the electrostatic interactions are the same as for the total interaction energy, but electrostatics alone is not capable to describe energy differences between linear and branched dimers correctly. Inclusion of the induction made the description even worse. A dispersionless energy was defined, which is the sum of electrostatics, induction, and exchange interactions. It turned out that the dispersionless interaction is an even worse predictor than electrostatics alone, the stability trends on the investigated dimers are somehow reversed. So it is clear that these systems can not be described without considering the dispersion interaction. In fact, the dispersion interaction alone was the best descriptor for the observed trends of the total interaction energy. But for an accurate description all components have to be considered, unfortunately there is no simple single descriptor for describing intermolecular interactions in hydrogen-bonded dimers.

Also all interaction components were studied as a function of the oxygen-oxygen distance. Surprisingly, all minima of the total interaction energy are located around an O–O distance of 2.9 Å. The electrostatic interaction has no minimum at all for the considered variation of the O–O distance. Therefore, electrostatics alone can never be solely responsible for the observed equilibrium distances. Minima are only observed when the exchange interaction is considered, because it is the only repulsive component. The minima of the dispersionless method are located at distances larger than 2.9 Å and their location also depends on the shape of the monomers. It turned out that the dispersionless method predicts somehow reversed stability trends over the whole potential curves. Therefore, inclusion of dispersion interaction is important for getting correct interaction energies and geometries.

In addition, significant differences were observed between linear and branched dimers. In the sequence of linear dimers the total interaction energy reaches already a plateau at $(EtOH)_2$. But for the sequence of branched dimers, the total interaction energy decreases by -6.5 kJ mol^{-1} when moving from $(EtOH)_2$ to $(t\text{BuOH})_2$. These differences arise because in branched dimers there is a larger number of close-contact atoms; this leads to larger

dispersion energies. For branched dimers dispersion is not only very important at the equilibrium geometry but plays also an important role at larger distances.

Chapter 6

Conclusion

In this thesis intermolecular interactions were studied for several hydrogen-bonded systems by using different computational methods. First, the adsorption of a glucose molecule onto cellulose surfaces was investigated. Geometries and adsorption energies were calculated with the GLYCAM06 force field and density-functional theory (DFT) with and without dispersion correction. GLYCAM06 is able to produce structure in good agreement with dispersion-corrected DFT results. Concerning adsorption energies, the GLYCAM06 energies are in most cases surprisingly close to dispersion-corrected DFT results. However, for structures with very few hydrogen bonds, GLYCAM06 tends to slightly overestimate the adsorption energies. We believe that this effect is much smaller for larger adsorbates like cellobiose and cellotetraose because they are quite flexible around the glycosidic bond and can therefore form more hydrogen bonds. These results suggested that GLYCAM06 can be safely used to study the adsorption of medium-sized carbohydrates onto cellulose surfaces. It was found that DFT without dispersion correction produces completely wrong adsorption energies, they are far too small. This suggests that dispersion interactions are important for this adsorption process, although the adsorbate is small and hydrogen bonds are formed. The exact amount of the dispersion interaction can not be calculated with this method, because the dispersion correction accounts only for dispersion interactions missing in the used density functional.

Next, the adsorption of a cellobiose and a cellotetraose molecule onto the hydrophilic (100) surface of cellulose Iα and the hydrophobic (100) surface of cellulose Iβ was investigated by using the GLYCAM06 force field. It was found that the adsorption onto the Iα surface is in general stronger than onto the Iβ surface, but with increasing adsorbate size the differences are getting smaller. These findings indicate that for very large adsorbates the hydrophilicity of the surface is probably not important any longer. In addition, it was found that adsorption parallel to the surface is stronger than vertical adsorption. Also no pronounced correlation between the adsorption energy and the number of hydrogen bonds between the surface and the adsorbate can be found. Therefore, the observed adsorption

energies can certainly not be attributed to hydrogen bonding alone. This also suggests that dispersion interactions are important.

Finally, we wanted to quantify the contribution of dispersion interactions to the interaction energy between two hydrogen-bonded molecules. By using symmetry-adapted perturbation theory (SAPT) interaction energies can be rigorously decomposed into four contributions: electrostatics, exchange, induction, and dispersion. Because of the computational costs of this method, the water dimer and progressively larger alcohol dimers were investigated up to the *tert*-butanol dimer. The interaction in hydrogen-bonded systems is often only attributed to electrostatic interactions. The presented results strongly contradict that. Although electrostatic interactions are very important, one has to consider all four interaction components for an accurate description of the interaction energy. It was found that the dispersion exhibits the largest variation with dimer size. For the *tert*-butanol dimer it contributes 94% to the total interaction energy, while the sum of electrostatics, exchange, and induction amount to only 6%. That means that without considering dispersion interactions one can not describe the interaction in these systems. But this does not mean that only dispersion is important, the interaction between two hydrogen-bonded molecules consists of an interplay of all four interaction components.

The dispersion interaction depends on the number of atoms in one monomer, which are in close contact to the other monomer. So whenever two molecules get closer together, the dispersion interaction gets stronger. However, at small distances the repulsive exchange prevents the molecules from getting any closer.

The results obtained from the SAPT study can now be related to the adsorption process of carbohydrates onto cellulose surfaces. For very small adsorbates the formation of hydrogen bonds may be important but for larger adsorbates the dispersion interaction clearly dominates. Large adsorbates will try to minimize the distance between every atom of themselves and the surface. We have already seen this in the adsorption study, in which the preferred orientation of the adsorbates was parallel to the surface. We now also know that dispersion interactions are even important for the hydrogen bond itself. So the formation of hydrogen bonds also increases the dispersion energy. Therefore, one can safely say that the adsorption of carbohydrates is dominated by dispersion interactions. Based on these findings, one can argue that dispersion interactions might even be of utmost importance between hydrogen-bonded glucan chains within the cellulose bulk.

But there is still room for further studies related to that topic. For example free energies of the adsorption process of larger carbohydrates onto various cellulose surfaces could be determined by molecular dynamics simulations using the GLYCAM06 force field. Next, the importance of dispersion interactions for the stabilization of bulk cellulose could be studied by using the dispersion function D_{as} together with the dispersionless density functional. It would also be interesting to study cooperative effects of multiple hydrogen bonds by using SAPT(DFT) for systems like the propanetriol dimer. In the presented SAPT(DFT) study

only the distance between the two monomers was varied. Therefore, one could also study the angular dependence of the interaction energy by using SAPT(DFT).

Appendix A

Adsorbate/Surface Complexes

A.1 Adsorption Energies for the Glucose Adsorption

Table A.1: Adsorption energies in kJ/mol for the BP86-optimized structures.

Method	Iα-1	Iα-2	Iα-3	Iβ-1	Iβ-2	Iβ-3
BP86	36.2	28.5	58.9	34.3	-9.9	31.9
B3LYP	37.0	33.0	62.4	32.7	-3.5	36.0
BP86-D2	74.7	87.1	106.8	82.1	49.5	83.4
B3LYP-D2	75.4	91.6	110.3	80.6	55.9	87.5
B3LYP-D3	68.8	78.6	104.4	71.1	50.3	79.6
M06-2X	64.2	72.6	95.9	69.2	38.6	71.4
ωB97X-D	70.6	87.6	106.3	75.3	56.8	83.6
GLYCAM06	67.9	80.7	94.2	54.2	64.4	77.7

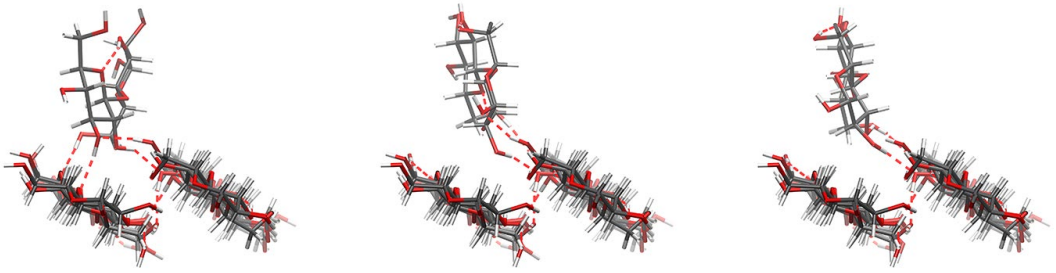
Table A.2: Adsorption energies in kJ/mol for the BP86-D2-optimized structures.

Method	Iα-1	Iα-2	Iα-3	Iβ-1	Iβ-2	Iβ-3
BP86	30.5	3.2	46.6	14.4	-42.5	13.0
B3LYP	30.0	3.4	48.5	10.8	-41.0	14.8
BP86-D2	75.7	103.4	115.6	82.9	61.6	84.5
B3LYP-D2	75.2	103.7	117.5	79.3	63.1	86.3
B3LYP-D3	66.6	81.9	105.3	63.1	40.5	70.9
M06-2X	63.5	81.4	100.0	62.3	44.3	69.8
ωB97X-D	69.2	91.0	109.1	70.1	55.8	78.0
GLYCAM06	64.0	72.6	87.7	40.5	61.2	67.9

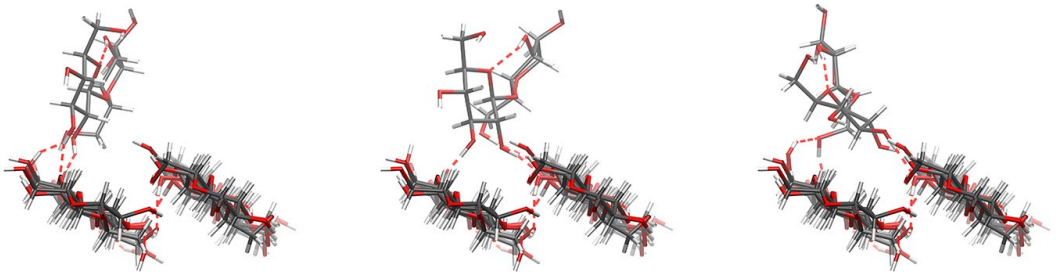
Table A.3: Adsorption energies in kJ/mol for the GLYCAM06-optimized structures.

Method	Iα-1	Iα-2	Iα-3	Iβ-1	Iβ-2	Iβ-3
BP86	18.9	8.4	47.0	17.7	-21.7	23.6
B3LYP	20.4	14.8	52.0	20.0	-16.4	28.9
BP86-D2	67.9	68.1	93.7	65.2	57.5	78.0
B3LYP-D2	69.4	74.5	98.8	67.5	62.9	83.3
B3LYP-D3	62.3	65.6	93.5	59.1	50.4	74.6
M06-2X	57.7	55.3	85.6	58.1	44.2	67.8
ωB97X-D	66.9	69.7	94.9	63.9	60.5	78.9
GLYCAM06	64.2	78.3	91.1	58.7	74.7	80.7

A.2 Geometries for the Adsorption of Cellobiose and Cellotetraose

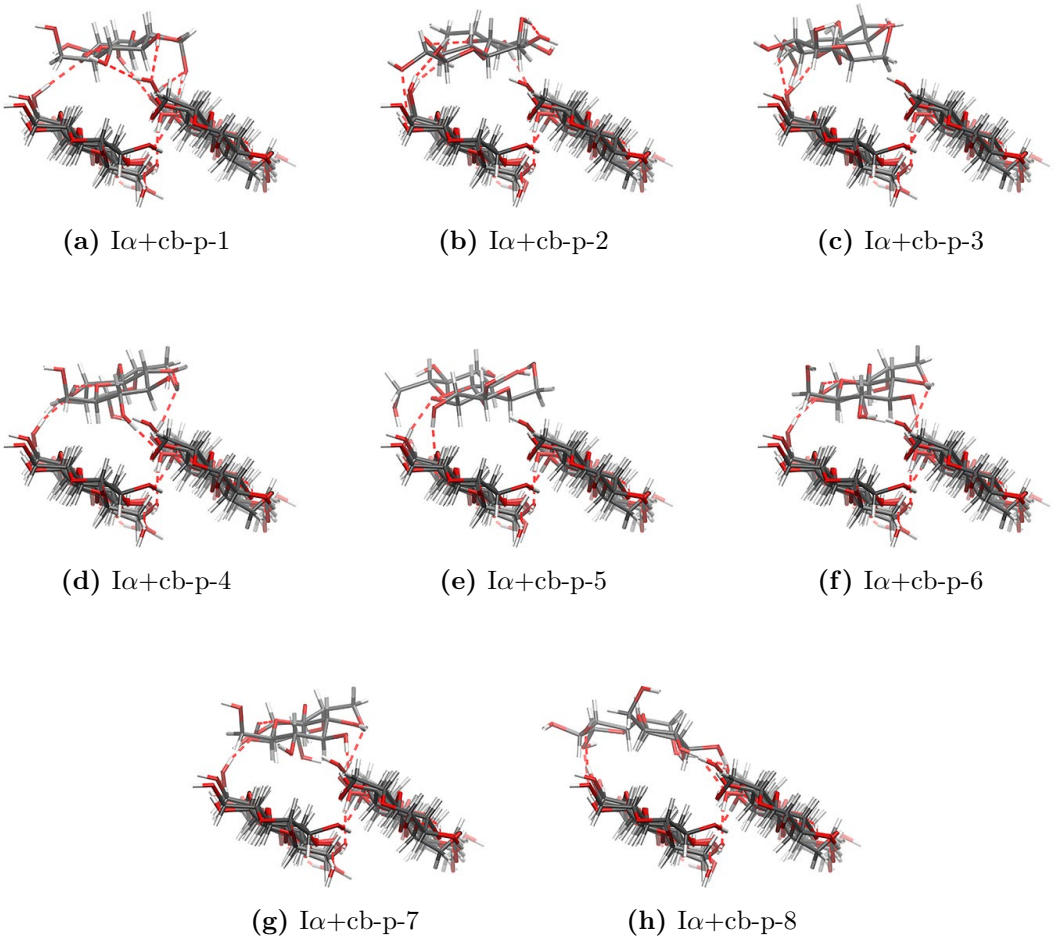


(a) Iα+cb-v-1 **(b)** Iα+cb-v-2 **(c)** Iα+cb-v-3



(d) Iα+cb-v-4 **(e)** Iα+cb-v-5 **(f)** Iα+cb-v-6

Figure A.1: Geometries of the vertical adsorption of cellobiose onto the cellulose Iα model surface.

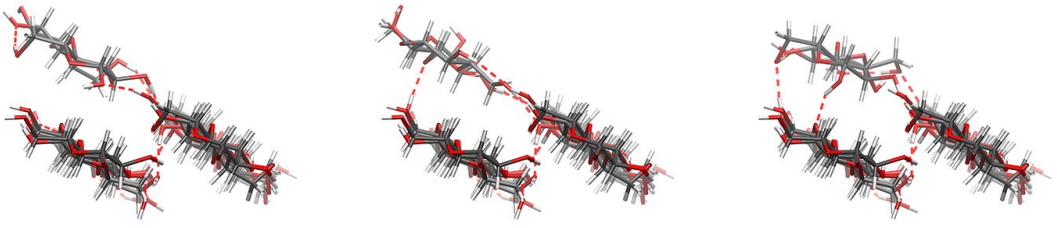


(a) Iα+cb-p-1 **(b)** Iα+cb-p-2 **(c)** Iα+cb-p-3

(d) Iα+cb-p-4 **(e)** Iα+cb-p-5 **(f)** Iα+cb-p-6

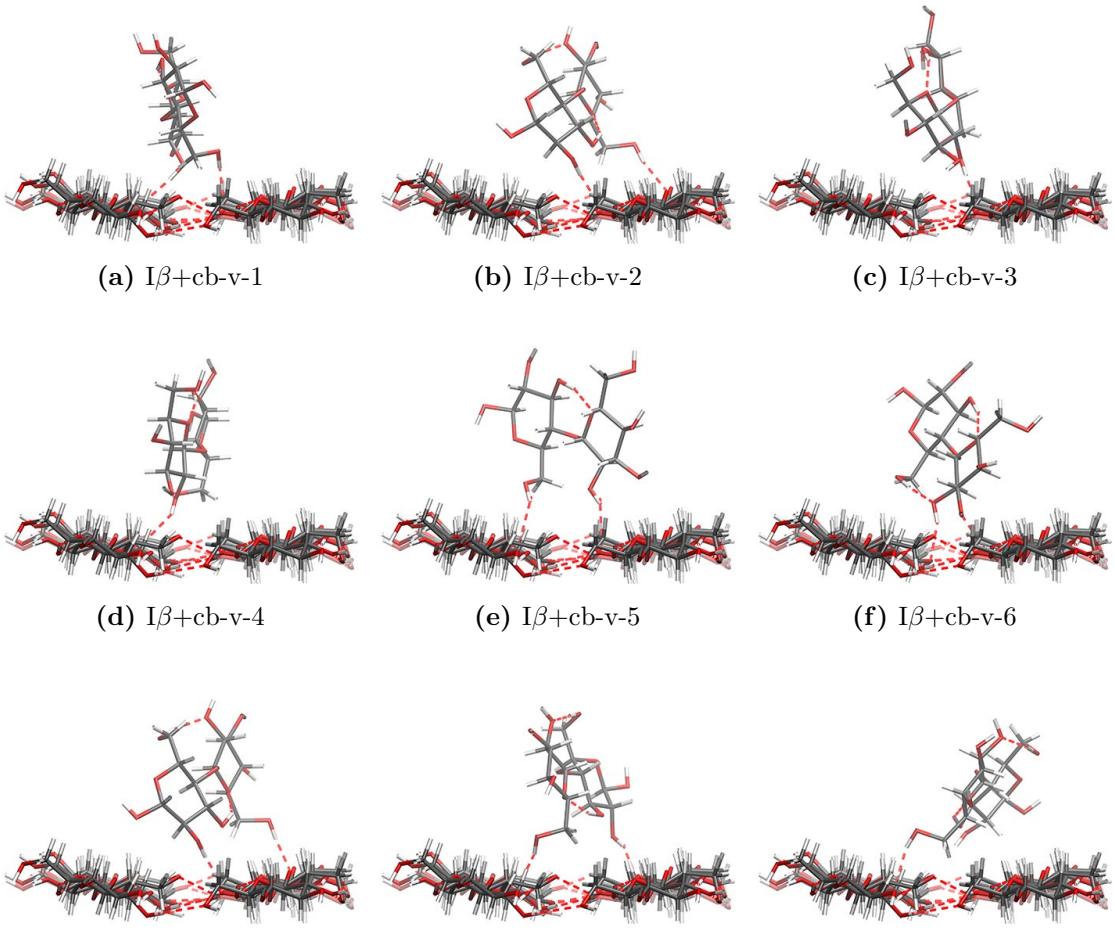
(g) Iα+cb-p-7 **(h)** Iα+cb-p-8

Figure A.2: Geometries of the parallel adsorption of cellobiose onto the cellulose Iα model surface.



(a) Iα+cb-d-1 **(b)** Iα+cb-d-2 **(c)** Iα+cb-d-3

Figure A.3: Geometries of the diagonal adsorption of cellobiose onto the cellulose Iα model surface.

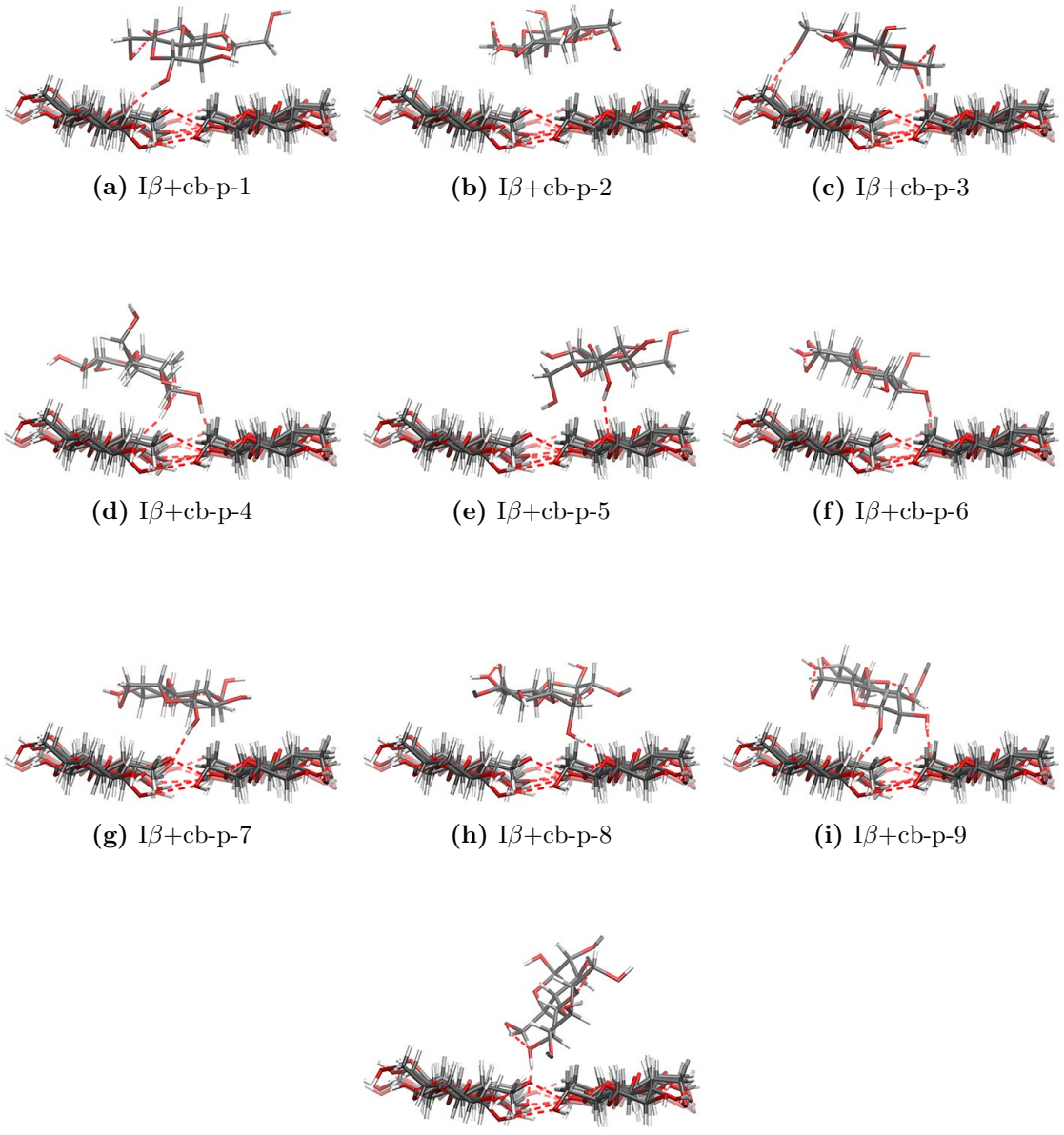


(a) Iβ+cb-v-1 **(b)** Iβ+cb-v-2 **(c)** Iβ+cb-v-3

(d) Iβ+cb-v-4 **(e)** Iβ+cb-v-5 **(f)** Iβ+cb-v-6

(g) Iβ+cb-v-7 **(h)** Iβ+cb-v-8 **(i)** Iβ+cb-v-9

Figure A.4: Geometries of the vertical adsorption of cellobiose onto the cellulose Iβ model surface.



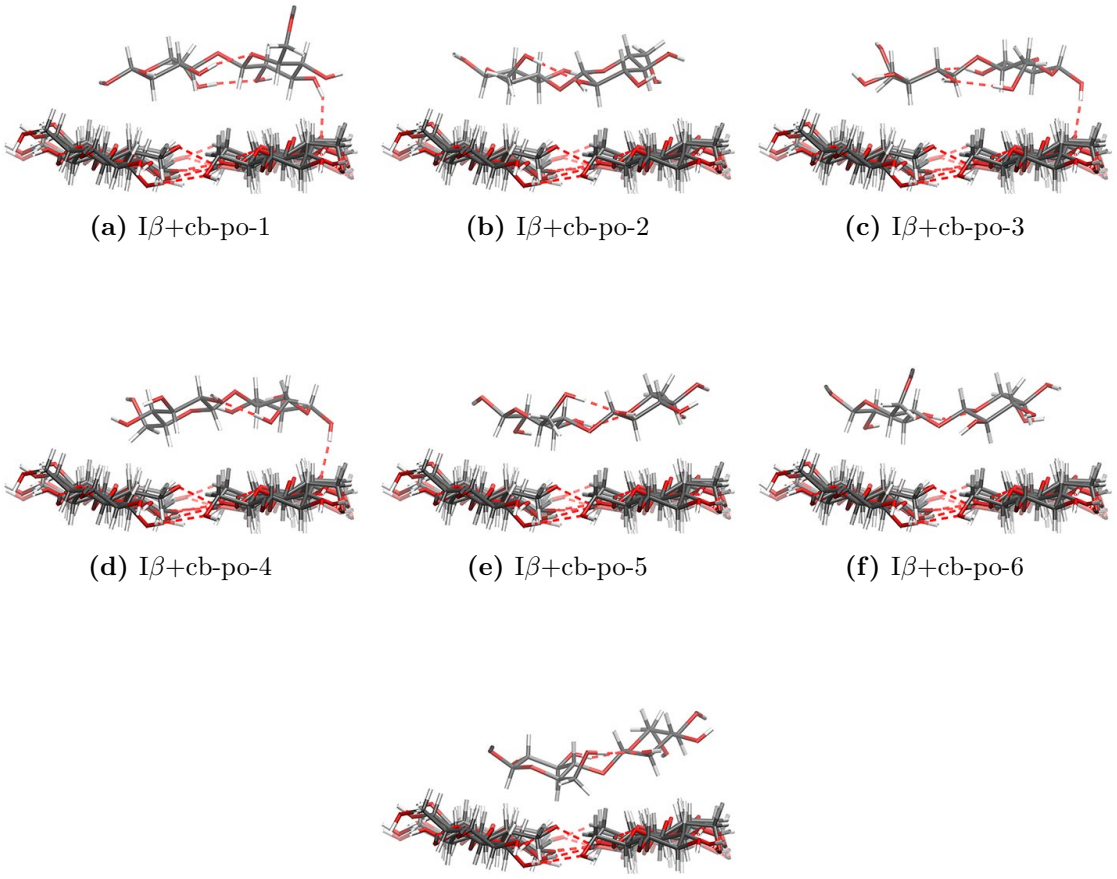
(a) Iβ+cb-p-1 **(b)** Iβ+cb-p-2 **(c)** Iβ+cb-p-3

(d) Iβ+cb-p-4 **(e)** Iβ+cb-p-5 **(f)** Iβ+cb-p-6

(g) Iβ+cb-p-7 **(h)** Iβ+cb-p-8 **(i)** Iβ+cb-p-9

(j) Iβ+cb-p-10

Figure A.5: Geometries of the parallel adsorption of cellobiose onto the cellulose Iβ model surface.

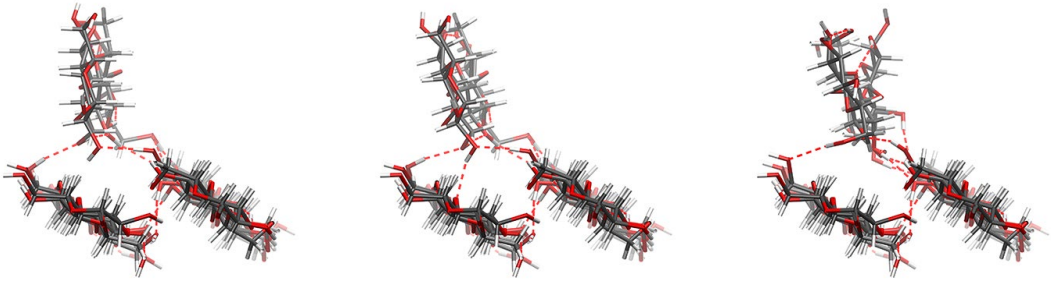


(a) Iβ+cb-po-1 **(b)** Iβ+cb-po-2 **(c)** Iβ+cb-po-3

(d) Iβ+cb-po-4 **(e)** Iβ+cb-po-5 **(f)** Iβ+cb-po-6

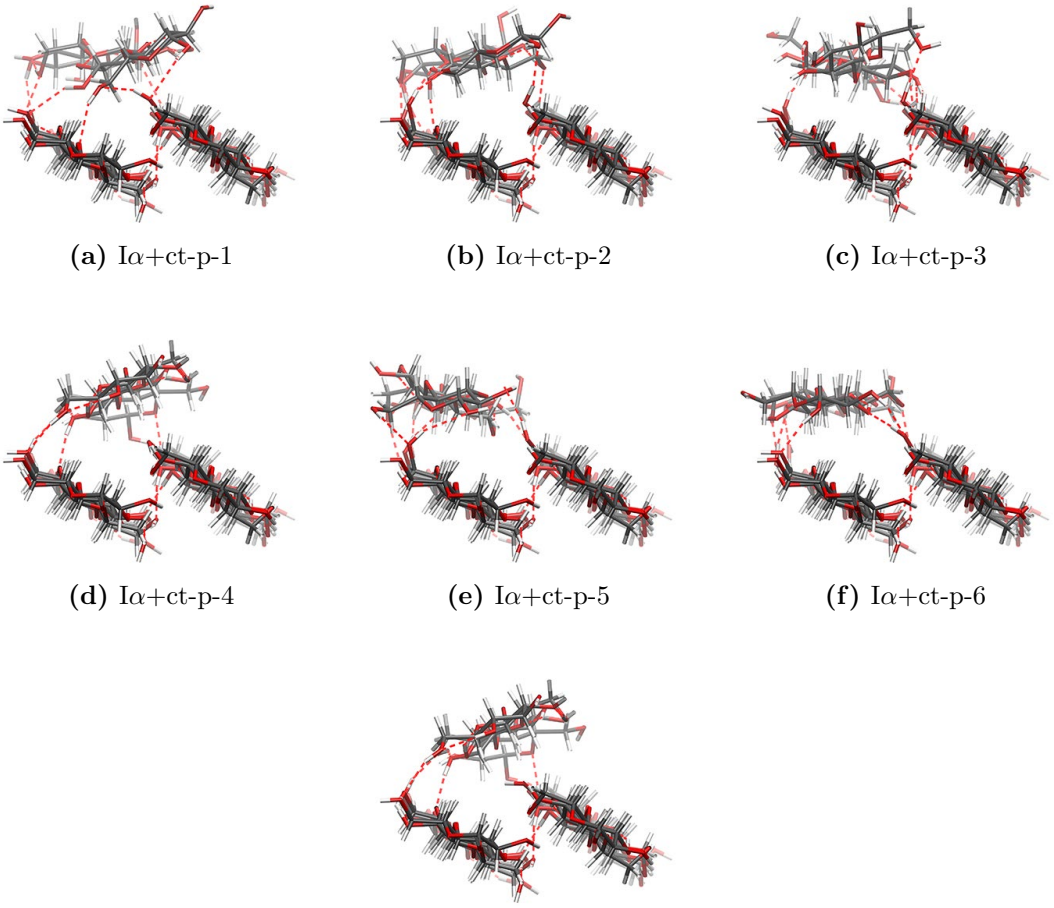
(g) Iβ+cb-po-7

Figure A.6: Geometries of the parallel, orthogonal adsorption of cellobiose onto the cellulose Iβ model surface.



(a) Iα+ct-v-1 **(b)** Iα+ct-v-2 **(c)** Iα+ct-v-3

Figure A.7: Geometries of the vertical adsorption of cellotetraose onto the cellulose Iα model surface.

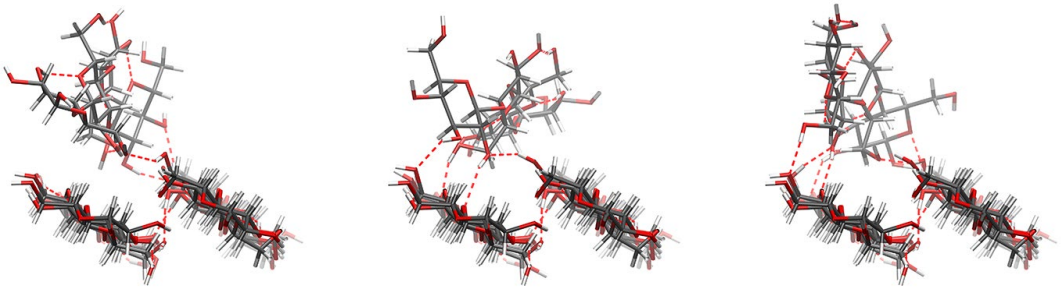


(a) Iα+ct-p-1 **(b)** Iα+ct-p-2 **(c)** Iα+ct-p-3

(d) Iα+ct-p-4 **(e)** Iα+ct-p-5 **(f)** Iα+ct-p-6

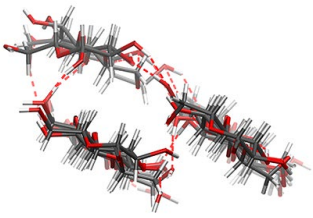
(g) Iα+ct-p-7

Figure A.8: Geometries of the parallel adsorption of cellotetraose onto the cellulose Iα model surface.



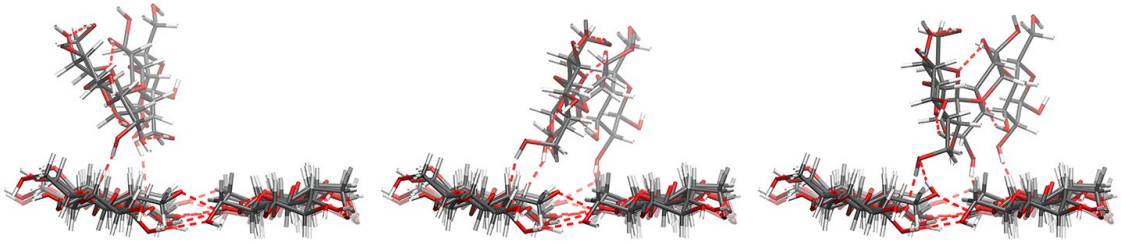
(a) Iα+ct-t-1 **(b)** Iα+ct-t-2 **(c)** Iα+ct-t-3

Figure A.9: Geometries of the twisted adsorption of cellotetraose onto the cellulose Iα model surface.



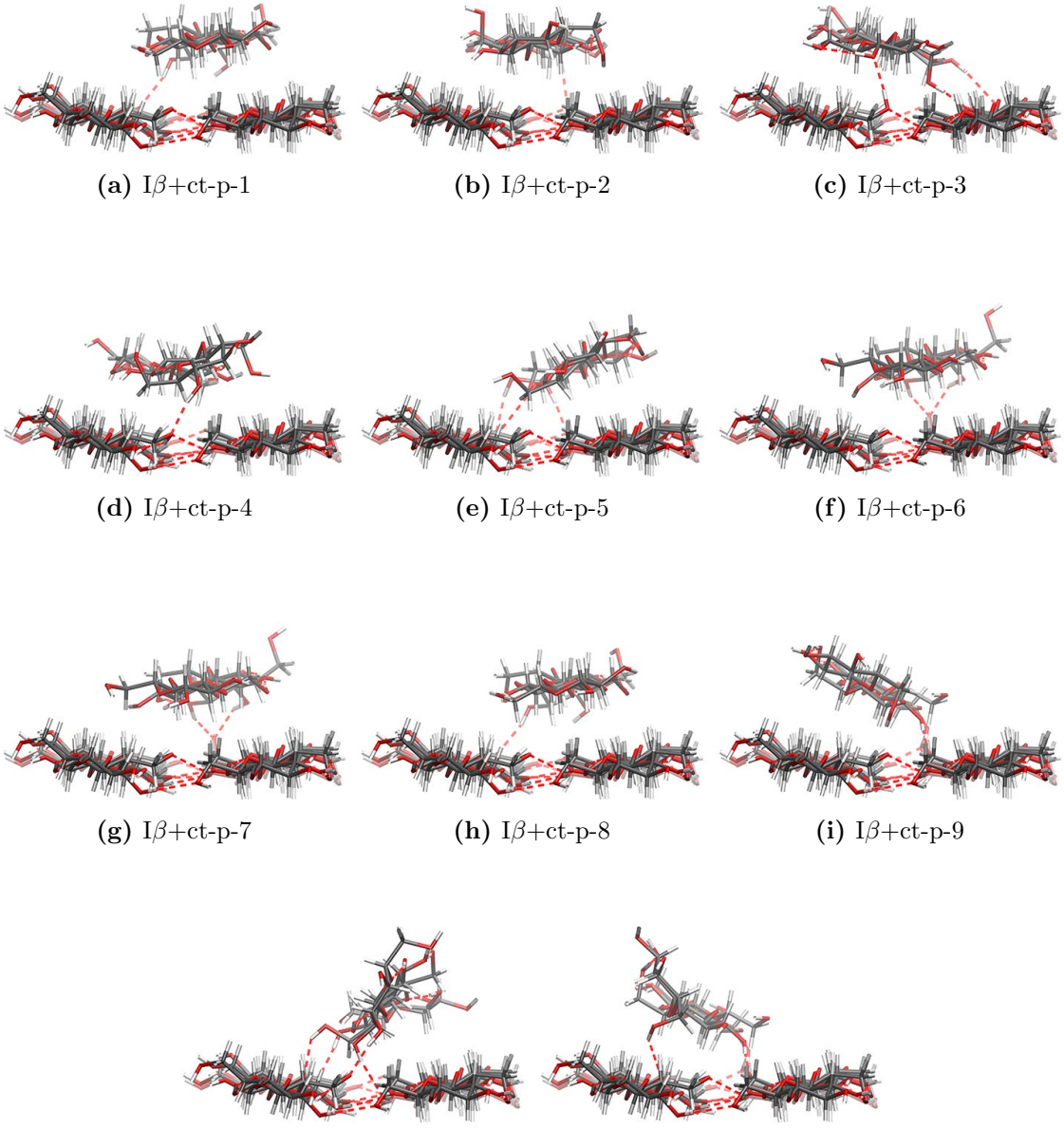
(a) Iα+ct-d-1

Figure A.10: Geometry of the diagonal adsorption of cellotetraose onto the cellulose Iα model surface.



(a) Iβ+ct-v-1 **(b)** Iβ+ct-v-2 **(c)** Iβ+ct-v-3

Figure A.11: Geometries of the vertical adsorption of cellotetraose onto the cellulose Iβ model surface.



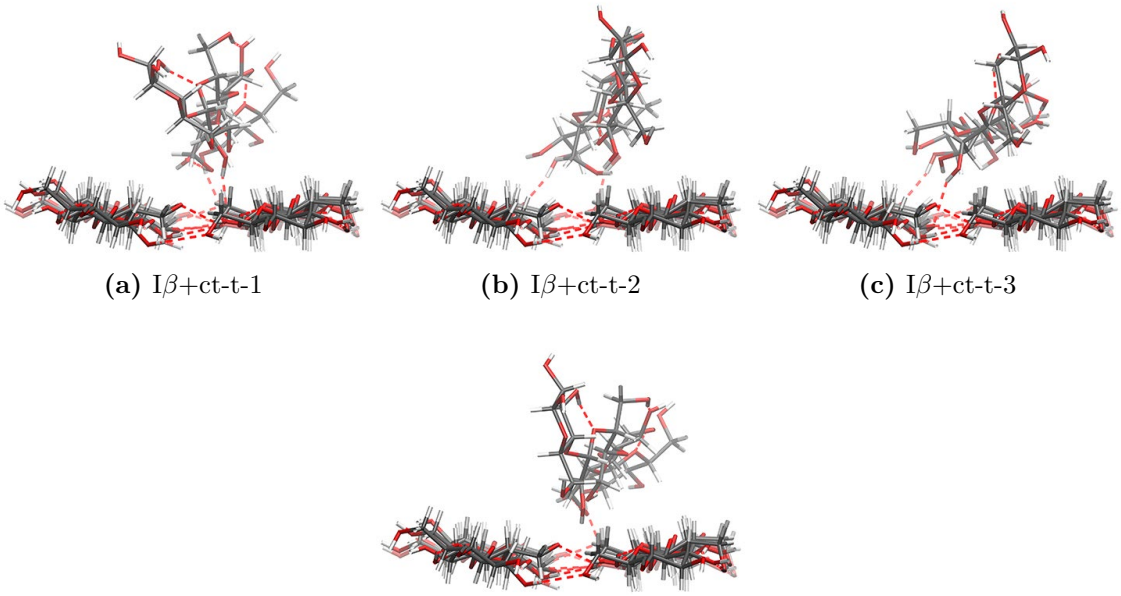
(a) Iβ+ct-p-1 **(b)** Iβ+ct-p-2 **(c)** Iβ+ct-p-3

(d) Iβ+ct-p-4 **(e)** Iβ+ct-p-5 **(f)** Iβ+ct-p-6

(g) Iβ+ct-p-7 **(h)** Iβ+ct-p-8 **(i)** Iβ+ct-p-9

(j) Iβ+ct-p-10 **(k)** Iβ+ct-p-11

Figure A.12: Geometries of the parallel adsorption of cellotetraose onto the cellulose Iβ model surface.



(a) Iβ+ct-t-1 **(b)** Iβ+ct-t-2 **(c)** Iβ+ct-t-3

(d) Iβ+ct-t-4

Figure A.13: Geometries of the twisted adsorption of cellotetraose on the cellulose Iβ model surface.

A.3 Structural Parameters for the Adsorption of Cellobiose and Cellotetraose

In Tables A.4 and A.5 the geometrical parameters for every structure of the cellobiose and cellotetraose adsorption (Chapter 4) are shown. The parameters α, β, and γ are given in degrees. The principal moments of inertia I_1, I_2, and I_3 are given in amu $\times$ Å^2, where the indexes 1 and 3 mark the smallest and the largest principal moment of inertia, respectively. Adsorption energies are listed in kJ/mol.

Table A.4: Structural parameters for the adsorption of cellobiose.

Structure	E_{ads}	α	β	γ	I_1	I_2	I_3	σ
Iα+cb-v-1	178	13.6	8.8	5.1	1259	3510	4409	20.4
Iα+cb-v-2	161	7.8	12.4	11.3	1205	3611	4614	21.7
Iα+cb-v-3	146	0.4	10.1	14.8	1188	3705	4647	20.3
Iα+cb-v-4	128	10.7	2.9	20.3	1269	3567	4638	23.1
Iα+cb-v-5	142	25.8	14.4	14.5	1242	3570	4376	18.4
Iα+cb-v-6	175	5.1	18.6	27.4	1261	3617	4419	18.1
Iα+cb-p-1	235	4.3	0.2	88.6	1250	3582	4599	22.1
Iα+cb-p-2	205	18.9	3.5	83.6	1207	3615	4618	21.7
Iα+cb-p-3	149	5.6	2.1	81.1	1188	3731	4664	20.0
Iα+cb-p-4	238	3.3	1.2	80.5	1254	3477	4396	20.9
Iα+cb-p-5	175	14.2	4.9	76.8	1215	3585	4620	22.4
Iα+cb-p-6	232	2.2	0.5	85.7	1261	3487	4384	20.5
Iα+cb-p-7	215	3.5	0.0	86.2	1262	3488	4413	21.0
Iα+cb-p-8	224	25.7	0.2	72.7	1169	3789	4653	18.6
Iα+cb-d-1	172	13.4	10.0	63.5	1219	3590	4603	22.0
Iα+cb-d-2	162	0.1	3.0	60.5	1243	3561	4637	23.2
Iα+cb-d-3	206	2.2	2.7	75.8	1181	3759	4747	20.8
Iβ+cb-v-1	111	6.2	4.7	8.7	1244	3508	4493	21.9
Iβ+cb-v-2	110	22.2	6.9	16.8	1229	3582	4505	20.5
Iβ+cb-v-3	86	7.0	21.6	12.1	1245	3483	4472	22.1
Iβ+cb-v-4	68	11.2	5.0	14.8	1240	3511	4553	22.9
Iβ+cb-v-5	77	42.1	10.3	6.3	1247	3530	4563	22.6
Iβ+cb-v-6	108	19.1	23.1	34.2	1229	3528	4540	22.3
Iβ+cb-v-7	113	28.1	9.5	11.1	1228	3596	4495	20.0
Iβ+cb-v-8	105	19.5	5.0	8.3	1224	3545	4556	22.2
Iβ+cb-v-9	114	13.0	5.5	42.9	1254	3453	4452	22.4
Iβ+cb-p-1	137	19.5	6.6	78.7	1246	3478	4496	22.6
Iβ+cb-p-2	89	5.1	4.4	85.1	1205	3599	4606	21.9
Iβ+cb-p-3	165	12.2	1.3	68.3	1231	3547	4587	22.7
Iβ+cb-p-4	186	19.2	2.5	57.3	1156	3844	4599	16.4
Iβ+cb-p-5	116	4.7	3.2	84.4	1265	3501	4453	21.4
Iβ+cb-p-6	139	9.0	7.8	66.4	1221	3537	4585	22.9
Iβ+cb-p-7	168	7.8	3.1	78.8	1237	3516	4559	22.9
Iβ+cb-p-8	135	18.1	3.5	70.2	1207	3595	4603	21.9
Iβ+cb-p-9	158	4.3	11.3	66.3	1241	3469	4441	21.9
Iβ+cb-p-10	94	5.2	14.7	46.4	1233	3508	4527	22.5
Iβ+cb-po-1	110	84.5	3.5	76.9	1238	3571	4507	20.8
Iβ+cb-po-2	110	68.1	6.0	84.0	1236	3458	4480	22.8
Iβ+cb-po-3	141	76.9	6.4	80.7	1223	3541	4582	22.7
Iβ+cb-po-4	135	65.8	6.2	83.8	1218	3537	4545	22.2
Iβ+cb-po-5	111	85.0	3.0	85.9	1243	3542	4538	21.9
Iβ+cb-po-6	104	87.3	6.7	81.9	1246	3567	4502	20.8
Iβ+cb-po-7	127	76.9	3.1	83.5	1253	3494	4510	22.5

Table A.5: Structural parameters for the adsorption of cellotetraose.

Structure	E_{ads}	α	β	γ	I_1	I_2	I_3	σ
Iα+ct-v-1	227	4.4	0.0	6.0	2554	27012	29104	7.2
Iα+ct-v-2	230	5.4	0.2	17.7	2572	27023	29147	7.3
Iα+ct-v-3	289	3.4	1.0	12.2	2735	26219	28067	6.6
Iα+ct-p-1	282	11.8	0.9	74.5	2562	26961	28983	7.0
Iα+ct-p-2	374	8.5	1.1	74.3	2580	27089	29223	7.3
Iα+ct-p-3	358	3.2	0.9	85.8	2625	27028	28818	6.2
Iα+ct-p-4	311	2.8	3.1	70.3	2552	27346	29483	7.2
Iα+ct-p-5	324	3.0	0.4	84.3	2569	27387	29376	6.8
Iα+ct-p-6	324	2.9	0.1	87.3	2511	27833	30037	7.3
Iα+ct-p-7	300	3.0	3.2	70.4	2557	27306	29458	7.3
Iα+ct-t-1	205	11.8	0.1	(7.3)	2749	26555	28075	5.4
Iα+ct-t-2	236	12.9	2.6	(43.7)	2801	26475	27655	4.3
Iα+ct-t-3	251	10.6	8.5	(13.4)	2651	26932	28282	4.8
Iα+ct-d-1	342	2.6	0.9	75.6	2476	28108	30102	6.6
Iβ+ct-v-1	155	8.5	1.1	12.8	2538	27254	29095	6.3
Iβ+ct-v-2	198	7.8	3.2	29.9	2583	26853	28932	7.2
Iβ+ct-v-3	214	14.2	1.6	13.6	2613	26992	29090	7.2
Iβ+ct-p-1	282	0.3	0.7	89.2	2569	27064	29209	7.3
Iβ+ct-p-2	287	1.8	1.2	81.7	2567	26821	28989	7.5
Iβ+ct-p-3	303	0.3	0.2	70.1	2580	27291	29516	7.5
Iβ+ct-p-4	288	8.5	0.2	83.1	2576	26927	28929	6.9
Iβ+ct-p-5	280	0.8	2.0	77.3	2597	26992	29272	7.8
Iβ+ct-p-6	292	8.6	0.6	88.2	2536	27078	29088	6.9
Iβ+ct-p-7	284	8.5	0.6	89.0	2533	27116	29144	7.0
Iβ+ct-p-8	288	0.2	0.7	89.2	2570	27045	29170	7.3
Iβ+ct-p-9	311	4.3	4.0	58.2	2583	26857	29031	7.5
Iβ+ct-p-10	246	2.6	3.9	60.2	2612	27054	28805	6.1
Iβ+ct-p-11	325	5.8	6.0	51.4	2613	26798	28608	6.3
Iβ+ct-t-1	195	9.4	6.7	(15.2)	2712	26771	27612	3.0
Iβ+ct-t-2	213	7.4	9.1	(22.6)	2599	26870	28244	4.9
Iβ+ct-t-3	247	7.5	7.9	(57.7)	2852	26609	27855	4.5
Iβ+ct-t-4	187	8.1	8.3	(34.8)	2666	27006	28393	4.9

Appendix B

Coordinates of the Water and Alcohol Dimers

Tables B.1 to B.7 show the cartesian coordinates in Å of all CC2/cc-pVDZ optimized geometries discussed in Chapter 5. In every table the total interaction energy E_{SAPT} as well as the total CC2/cc-pVDZ energy E_{CC2} are listed.

Table B.1: Coordinates of $(H_2O)_2$ [97]

H	0.96684	0.09795	-0.00277
O	0.00000	0.00000	0.00000
H	-0.28968	0.92066	0.00446
H	3.03798	-0.59166	0.75480
O	2.90430	0.00000	0.00000
H	3.04610	-0.59380	-0.75161

$E_{\mathrm{CC2}} = -152.471346$ a.u.
$E_{\mathrm{SAPT}} = -19.3$ kJ/mol

Table B.2: Coordinates of $(MeOH)_2$ [97]

H	0.50885	2.03588	0.21850
C	-0.21666	1.36957	-0.28961
H	-1.22592	1.63165	0.06638
H	-0.17435	1.59331	-1.37500
O	0.00000	0.00000	0.00000
H	0.92357	-0.16800	-0.25827
H	1.87253	0.02191	1.82187
C	2.87082	-0.18632	1.41329
H	3.60473	0.50712	1.86316
H	3.14719	-1.22338	1.67778
O	2.75773	0.00000	0.00000
H	3.57754	-0.31955	-0.40152

$E_{\mathrm{CC2}} = -230.795063$ a.u.
$E_{\mathrm{SAPT}} = -22.7$ kJ/mol

Table B.3: Coordinates of $(EtOH)_2$ [97]

H	-0.24917	3.19817	1.13332
C	0.04295	2.41209	0.41512
C	-0.32018	1.02289	0.93168
H	1.13266	2.47994	0.25163
H	-0.46221	2.60219	-0.54629
H	0.17761	0.85362	1.91151
O	0.00000	0.00000	0.00000
H	-1.40703	0.95400	1.11042
H	0.95889	0.09817	-0.15660
H	4.28395	-2.19916	2.13603
C	4.25650	-1.51186	1.27378
C	2.85354	-0.96329	1.06758
H	4.59286	-2.07479	0.38495
H	4.96685	-0.68887	1.45228
H	2.12768	-1.77846	0.88691
O	2.80863	0.00000	0.00000
H	2.51615	-0.41481	1.96107
H	3.14067	-0.45636	-0.79010

$E_{\mathrm{CC2}} = -309.169774$ a.u.
$E_{\mathrm{SAPT}} = -23.6$ kJ/mol

Table B.4: Coordinates of $(n\mathrm{PrOH})_2$ [97]

H	-0.00930	4.50774	1.08717
C	-0.30114	3.50901	1.45187
C	0.07183	2.41461	0.45036
H	-1.38977	3.52325	1.63636
H	0.19978	3.34689	2.42267
C	-0.30990	1.01800	0.93921
H	1.16186	2.42745	0.26029
H	-0.42433	2.58964	-0.52143
H	0.18038	0.83545	1.92223
O	0.00000	0.00000	0.00000
H	-1.39971	0.96645	1.11349
H	0.96004	0.08994	-0.15471
H	5.40505	-2.81294	2.61330
C	4.38127	-2.42635	2.48446
C	4.29258	-1.47980	1.28555
H	4.09567	-1.90907	3.41659
H	3.70656	-3.29075	2.35889
C	2.88137	-0.94585	1.08047
H	4.61230	-2.00988	0.36699
H	4.97614	-0.62186	1.41354
H	2.16609	-1.77734	0.92386
O	2.80926	0.00000	0.00000
H	2.54898	-0.39097	1.97339
H	3.14615	-0.46162	-0.78506

$E_{\mathrm{CC2}} = -387.533253$ a.u.
$E_{\mathrm{SAPT}} = -24.7$ kJ/mol

Table B.5: Coordinates of $(i\text{PrOH})_2$ [97]

H	-2.04641	-1.33316	-0.91463
C	-1.87622	-1.43216	0.16959
H	-2.44638	-0.63851	0.68212
H	-2.25487	-2.41344	0.50171
C	-0.39593	-1.28073	0.48810
H	0.16497	-2.08407	-0.04207
O	0.00000	0.00000	0.00000
H	0.96513	0.05494	0.15617
C	-0.11705	-1.40095	1.98603
H	-0.66229	-0.60959	2.52823
H	-0.43089	-2.38479	2.37723
H	0.96039	-1.27546	2.19244
H	2.94125	2.45620	-0.78861
C	2.47582	1.78293	-1.52603
H	1.38527	1.77345	-1.36219
H	2.67857	2.17102	-2.53777
C	3.03489	0.37586	-1.38147
H	4.12727	0.39277	-1.57177
O	2.81531	0.00000	0.00000
H	3.14619	-0.90877	0.09038
C	2.35994	-0.61760	-2.32169
H	1.27882	-0.63861	-2.10820
H	2.51778	-0.32929	-3.37486
H	2.77224	-1.63390	-2.18582

$E_{\text{CC2}} = -387.548390$ a.u.
$E_{\text{SAPT}} = -27.1$ kJ/mol

Table B.6: Coordinates of $(n\text{BuOH})_2$ [97]

C	0.14100	4.90713	0.96384
C	-0.27051	3.51733	1.45523
H	1.23076	4.95519	0.79361
H	-0.12386	5.69208	1.69226
H	-0.35729	5.14835	0.00876
C	0.08973	2.41094	0.46203
H	-1.36107	3.49864	1.64400
H	0.21603	3.30983	2.42777
C	-0.30297	1.01561	0.94527
H	1.18149	2.41954	0.27382
H	-0.40351	2.59521	-0.51098
H	0.18650	0.82304	1.92636
O	0.00000	0.00000	0.00000
H	-1.39294	0.97036	1.11910
H	0.95886	0.09135	-0.16147
C	5.73033	-3.00018	2.74071
C	4.32668	-2.42986	2.52684
H	5.76499	-3.66895	3.61663
H	6.46253	-2.19031	2.90251
H	6.05823	-3.57849	1.85950
C	4.25322	-1.49280	1.31852
H	4.00651	-1.88008	3.43210
H	3.60439	-3.25726	2.39295
C	2.85481	-0.93505	1.09246
H	4.57557	-2.04197	0.41061
H	4.95447	-0.64764	1.44617
H	2.12480	-1.75329	0.93700
O	2.81050	0.00000	0.00000
H	2.52399	-0.36293	1.97484
H	3.14654	-0.47688	-0.77634

$E_{\text{CC2}} = -465.896549$ a.u.
$E_{\text{SAPT}} = -25.1$ kJ/mol

Table B.7: Coordinates of $(t\mathrm{BuOH})_2$ [97]

H	-2.46941	-0.76090	0.29196
C	-2.02101	-0.28840	1.18196
H	-2.20378	0.79761	1.12550
H	-2.51520	-0.68713	2.08424
C	-0.52003	-0.56053	1.21530
C	-0.24237	-2.06814	1.23837
O	0.00000	0.00000	0.00000
H	0.97270	-0.11439	0.05249
C	0.13492	0.12151	2.42166
H	-0.07647	1.20434	2.40612
H	-0.23954	-0.29708	3.37266
H	1.22990	-0.01722	2.38511
H	-0.66534	-2.53974	0.33551
H	-0.68064	-2.54893	2.13085
H	0.84687	-2.25390	1.24980
H	2.84453	2.44033	0.90543
C	2.38931	2.32373	-0.09239
H	1.31354	2.12080	0.02800
H	2.52147	3.26333	-0.65422
C	3.05001	1.16843	-0.83810
C	4.55891	1.37307	-0.96609
O	2.81013	0.00000	0.00000
H	3.20545	-0.75301	-0.47230
C	2.39607	0.94894	-2.20385
H	1.31543	0.76632	-2.08464
H	2.53980	1.83013	-2.85258
H	2.84811	0.07824	-2.71391
H	4.78479	2.28111	-1.55095
H	5.02699	0.51419	-1.48092
H	5.01387	1.47069	0.03318

$E_{\mathrm{CC2}} = -465.927977$ a.u.
$E_{\mathrm{SAPT}} = -30.1$ kJ/mol

List of Figures

List of Tables

Bibliography

[1] I. G. Kaplan, Intermolecular Interactions: Physical Picture, Computational Methods and Model Potentials, John Wiley & Sons, Ltd, **2006**.

[2] G. C. Maitland, M. Rigby, E. B. Smith, W. A. Wakeham, Intermolecular Forces: Their Origin and Determination, of *International Series of Monographs on Chemistry*, Clarendon Press, Oxford, **1981**.

[3] A. Stone, The Theory of Intermolecular Forces, 2nd Edition, Oxford University Press, **2013**.

[4] G. A. Jeffrey, An Introduction to Hydrogen Bonding, Oxford University Press, New York, **1997**.

[5] F. London, Über einige Eigenschaften und Anwendungen der Molekularkräfte, *Z. Phys. Chem.* **1930**, *B11*, 222–251.

[6] H. B. G. Casimir, D. Polder, The Influence of Retardation on the London-Van der Waals Forces, *Phys. Rev.* **1948**, *73*, 360–372.

[7] W. M. Latimer, W. H. Rodebush, Polarity and Ionization from the Standpoint of the Lewis Theory of Valence, *J. Am. Chem. Soc.* **1920**, *42*, 1419–1433.

[8] G. N. Lewis, Valence and the Structure of Atoms and Molecules, Chemical Catalogue Co., New York, **1923**.

[9] R. Rohs, S. West, A. Sosinsky, P. Liu, R. Mann, B. Honig, The Role of DNA Shape in Protein-DNA Recognition, *Nature* **2009**, *461*, 1248–1253.

[10] M. Ahmad, W. Gu, T. Geyer, V. Helms, Adhesive Water Networks Facilitate Binding of Protein Interfaces, *Nat. Commun.* **2011**, *2*, 261.

[11] J. Guo, X. Meng, J. Chen, J. Peng, J. Sheng, X.-Z. Li, L. Xu, J.-R. Shi, E. Wang, Y. Jiang, Real-Space Imaging of Interfacial Water With Submolecular Resolution, *Nat. Mater.* **2014**, *13*, 184–189.

[12] G. R. Desiraju, T. Steiner, The Weak Hydrogen Bond, Oxford University Press, New York, **2001**.

[13] S. J. Grabowski, Hydrogen Bonding - New Insights, Springer, Dortrecht, **2006**.

[14] G. Gilli, P. Gilli, The Nature of the Hydrogen Bond, Oxford University Press, New York, **2009**.

[15] C. A. Coulson, Valence, Clarendon Press, Oxford, **1952**.

[16] J. N. Israelachvili, Intermolecular and Surface Forces, Third Edition, Academic Press, **2010**.

[17] E. D. Isaacs, A. Shukla, P. M. Platzman, D. R. Hamann, B. Barbiellini, C. A. Tulk, Covalency of the Hydrogen Bond in Ice: A Direct X-Ray Measurement, *Phys. Rev. Lett.* **1999**, *82*, 600–603.

[18] S. J. Grabowski, What Is the Covalency of Hydrogen Bonding?, *Chem. Rev.* **2011**, *111*, 2597–2625.

[19] T. K. Ghanty, V. N. Staroverov, P. R. Koren, E. R. Davidson, Is the Hydrogen Bond in Water Dimer and Ice Covalent?, *J. Am. Chem. Soc.* **2000**, *122*, 1210–1214.

[20] E. R. Davidson, Bonding in FHF^-, $(HF)_2$, and FHF, *Int. J. Quantum Chem.* **2004**, *98*, 317–324.

[21] A. O'Sullivan, Cellulose: The Structure Slowly Unravels, *Cellulose* **1997**, *4*, 173–207.

[22] G. M. Raghavendra, T. Jayaramudu, K. Varaprasad, R. Sadiku, S. S. Ray, K. M. Raju, Cellulose-Polymer-Ag Nanocomposite Fibers for Antibacterial Fabrics/Skin Scaffolds, *Carbohydr. Polym.* **2013**, *93*, 553–560.

[23] Y. Okahisa, A. Yoshida, S. Miyaguchi, H. Yano, Optically Transparent Wood-Cellulose Nanocomposite as a Base Substrate for Flexible Organic Light-Emitting Diode Displays, *Compos. Sci. Technol.* **2009**, *69*, 1958–1961.

[24] M. Nogi, S. Iwamoto, A. N. Nakagaito, H. Yano, Optically Transparent Nanofiber Paper, *Adv. Mater.* **2009**, *21*, 1595–1598.

[25] G. Findenig, S. Leimgruber, R. Kargl, S. Spirk, K. Stana-Kleinschek, V. Ribitsch, Creating Water Vapor Barrier Coatings from Hydrophilic Components, *ACS Appl. Mater. Interfaces* **2012**, *4*, 3199–3206.

[26] Y. Nishiyama, J. Sugiyama, H. Chanzy, P. Langan, Crystal Structure and Hydrogen Bonding System in Cellulose Iα from Synchrotron X-Ray and Neutron Fiber Diffraction, *J. Am. Chem. Soc.* **2003**, *125*, 14300–14306.

[27] Y. Nishiyama, P. Langan, H. Chanzy, Crystal Structure and Hydrogen-Bonding System in Cellulose Iβ from Synchrotron X-Ray and Neutron Fiber Diffraction, *J. Am. Chem. Soc.* **2002**, *124*, 9074–9082.

[28] T. Shen, S. Gnanakaran, The Stability of Cellulose: A Statistical Perspective from a Coarse-Grained Model of Hydrogen-Bond Networks, *Biophys. J.* **2009**, *96*, 3032–3040.

[29] X. Qian, S.-y. Ding, M. R. Nimlos, D. K. Johnson, M. E. Himmel, Atomic and Electronic Structures of Molecular Crystalline Cellulose Iβ: A First-Principles Investigation, *Macromolecules* **2005**, *38*, 10580–10589.

[30] Y. Nishiyama, G. P. Johnson, A. D. French, V. T. Forsyth, P. Langan, Neutron Crystallography, Molecular Dynamics, and Quantum Mechanics Studies of the Nature of Hydrogen Bonding in Cellulose Iβ, *Biomacromolecules* **2008**, *9*, 3133–3140.

[31] B. Medronho, A. Romano, M. G. Miguel, L. Stigsson, B. Lindman, Rationalizing Cellulose (In)solubility: Reviewing Basic Physicochemical Aspects and Role of Hydrophobic Interactions, *Cellulose* **2012**, *19*, 581–587.

[32] B. Medronho, B. Lindman, Competing Forces During Cellulose Dissolution: From Solvents to Mechanisms, *Curr. Opin. Colloid Interface Sci.* **2014**, *19*, 32–40.

[33] B. Lindman, G. Karlström, L. Stigsson, On the Mechanism of Dissolution of Cellulose, *J. Mol. Liq.* **2010**, *156*, 76–81.

[34] J. Hanus, K. Mazeau, The Xyloglucan-Cellulose Assembly at the Atomic Scale, *Biopolymers* **2006**, *82*, 59–73.

[35] M. Bergenstråhle, J. Wohlert, M. E. Himmel, J. W. Brady, Simulation Studies of the Insolubility of Cellulose, *Carbohydr. Res.* **2010**, *345*, 2060–2066.

[36] K. Mazeau, L. Charlier, The Molecular Basis of the Adsorption of Xylans on Cellulose Surface, *Cellulose* **2012**, *19*, 337–349.

[37] S. Peri, L. Muthukumar, M. Nazmul Karim, R. Khare, Dynamics of Cello-Oligosaccharides on a Cellulose Crystal Surface, *Cellulose* **2012**, *19*, 1791–1806.

[38] Z. Zhao, V. H. Crespi, J. D. Kubicki, D. J. Cosgrove, L. Zhong, Molecular Dynamics Simulation Study of Xyloglucan Adsorption on Cellulose Surfaces: Effects of Surface Hydrophobicity and Side-Chain Variation, *Cellulose* **2014**, *21*, 1025–1039.

[39] M. R. Nimlos, G. T. Beckham, J. F. Matthews, L. Bu, M. E. Himmel, M. F. Crowley, Binding Preferences, Surface Attachment, Diffusivity, and Orientation of a Family 1 Carbohydrate-Binding Module on Cellulose, *J. Biol. Chem.* **2012**, *287*, 20603–20612.

[40] K. Mazeau, C. Vergelati, Atomistic Modeling of the Adsorption of Benzophenone onto Cellulosic Surfaces, *Langmuir* **2002**, *18*, 1919–1927.

[41] K. Mazeau, M. Wyszomirski, Modelling of Congo Red Adsorption on the Hydrophobic Surface of Cellulose Using Molecular Dynamics, *Cellulose* **2012**, *19*, 1495–1506.

[42] T. Nypelö, H. Pynnönen, M. Österberg, J. Paltakari, J. Laine, Interactions Between Inorganic Nanoparticles and Cellulose Nanofibrils, *Cellulose* **2012**, *19*, 779–792.

[43] K. Mazeau, A. Rivet, Wetting the (110) and (100) Surfaces of Iβ Cellulose Studied by Molecular Dynamics, *Biomacromolecules* **2008**, *9*, 1352–1354.

[44] Y. Li, M. Lin, J. W. Davenport, Ab Initio Studies of Cellulose I: Crystal Structure, Intermolecular Forces, and Interactions with Water, *J. Phys. Chem. C* **2011**, *115*, 11533–11539.

[45] A. Szabo, N. S. Ostlund, Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory , Dover Publications, Mineola, **1996**.

[46] R. Parr, W. Yang, Density-Functional Theory of Atoms and Molecules, of *International Series of Monographs on Chemistry*, Oxford University Press, USA, **1989**.

[47] L. Piela, Ideas of Quantum Chemistry *2nd ed.*, Elsevier, Oxford, **2014**.

[48] A. D. Becke, Perspective: Fifty Years of Density-Functional Theory in Chemical Physics, *J. Chem. Phys.* **2014**, *140*, 18A301.

[49] W. Kohn, A. D. Becke, R. G. Parr, Density Functional Theory of Electronic Structure, *J. Phys. Chem.* **1996**, *100*, 12974–12980.

[50] B. Jeziorski, K. Szalewicz, Handbook of Molecular Physics and Quantum Chemistry, *Vol. 3*, S. Wilson (Ed.), John Wiley & Sons, Inc., **2003**, chapter Symmetry-Adapted Perturbation Theory, pp. 232–279.

[51] K. Szalewicz, Symmetry-Adapted Perturbation Theory of Intermolecular Forces, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 254–272.

[52] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeiriño, C. R. Daniels, B. L. Foley, R. J. Woods, GLYCAM06: A Generalizable Biomolecular Force Field. Carbohydrates, *J. Comput. Chem.* **2008**, *29*, 622–655.

[53] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.

[54] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: the RESP Model, *J. Phys. Chem.* **1993**, *97*, 10269–10280.

[55] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev.* **1964**, *136*, B864–B871.

[56] W. Kohn, L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.* **1965**, *140*, A1133–A1138.

[57] A. D. Becke, Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior, *Phys. Rev. A* **1988**, *38*, 3098–3100.

[58] J. Perdew, Density-Functional Approximation for the Correlation Energy of the Inhomogeneous Electron Gas, *Phys. Rev. B* **1986**, *33*, 8822–8824.

[59] C. Lee, W. Yang, R. G. Parr, Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density, *Phys. Rev. B* **1988**, *37*, 785–789.

[60] A. D. Becke, Density-Functional Thermochemistry. III. The Role of Exact Exchange, *J. Chem. Phys.* **1993**, *98*, 5648 – 5652.

[61] Y. Zhao, D. G. Truhlar, The M06 Suite of Density Functionals for Main Group Thermochemistry, Thermochemical Kinetics, Noncovalent Interactions, Excited States, and Transition Elements: Two New Functionals and Systematic Testing of Four M06-Class Functionals and 12 Other Functionals, *Theor. Chem. Acc.* **2008**, *120*, 215–241.

[62] J.-D. Chai, M. Head-Gordon, Systematic Optimization of Long-Range Corrected Hybrid Density Functionals, *J. Chem. Phys.* **2008**, *128*, 084106.

[63] J.-D. Chai, M. Head-Gordon, Long-Range Corrected Hybrid Density Functionals with Damped Atom-Atom Dispersion Corrections, *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615–6620.

[64] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[65] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)], *Phys. Rev. Lett.* **1997**, *78*, 1396–1396.

[66] C. Adamo, M. Cossi, V. Barone, An Accurate Density Functional Method for the Study of Magnetic Properties: The PBE0 Model, *J. Mol. Struct.: THEOCHEM* **1999**, *493*, 145–157.

[67] S. Grimme, Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction, *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[68] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu, *J. Chem. Phys.* **2010**, *132*, 154104.

[69] S. Grimme, S. Ehrlich, L. Goerigk, Effect of the Damping Function in Dispersion Corrected Density Functional Theory, *J. Comput. Chem.* **2011**, *32*, 1456–1465.

[70] K. Pernal, R. Podeszwa, K. Patkowski, K. Szalewicz, Dispersionless Density Functional Theory, *Phys. Rev. Lett.* **2009**, *103*, 263201.

[71] R. Podeszwa, K. Pernal, K. Patkowski, K. Szalewicz, Extension of the Hartree-Fock Plus Dispersion Method by First-Order Correlation Effects, *J. Phys. Chem. Lett.* **2010**, *1*, 550–555.

[72] A. Tkatchenko, M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.* **2009**, *102*, 073005.

[73] A. Tkatchenko, R. A. DiStasio, R. Car, M. Scheffler, Accurate and Efficient Method for Many-Body van der Waals Interactions, *Phys. Rev. Lett.* **2012**, *108*, 236402.

[74] S. Boys, F. Bernardi, The Calculation of Small Molecular Interactions by the Differences of Separate Total Energies. Some Procedures with Reduced Errors, *Mol. Phys.* **1970**, *19*, 553–566.

[75] J. Hoja, R. J. Maurer, A. F. Sax, Adsorption of Glucose, Cellobiose, and Cellotetraose onto Cellulose Model Surfaces, *J. Phys. Chem. B* **2014**, in press, DOI: 10.1021/jp5025685.

[76] R. J. Maurer, A. F. Sax, V. Ribitsch, Molecular Simulation of Surface Reorganization and Wetting in Crystalline Cellulose I and II, *Cellulose* **2013**, *20*, 25–42.

[77] A. Schäfer, H. Horn, R. Ahlrichs, Fully Optimized Contracted Gaussian Basis Sets for Atoms Li to Kr, *J. Chem. Phys.* **1992**, *97*, 2571.

[78] TURBOMOLE V6.3 2011, A Development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

[79] R. Ahlrichs, M. Bär, M. Häser, H. Horn, C. Kölmel, Electronic Structure Calculations on Workstation Computers: The Program System Turbomole, *Chem. Phys. Lett.* **1989**, *162*, 165–169.

[80] O. Treutler, R. Ahlrichs, Efficient Molecular Numerical Integration Schemes, *J. Chem. Phys.* **1995**, *102*, 346–354.

[81] M. Von Arnim, R. Ahlrichs, Performance of Parallel TURBOMOLE for Density Functional Calculations, *J. Comput. Chem.* **1998**, *19*, 1746–1757.

[82] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, Auxiliary Basis Sets to Approximate Coulomb Potentials, *Chem. Phys. Lett.* **1995**, *240*, 283–290.

[83] K. Eichkorn, O. Treutler, H. Öhm, M. Häser, R. Ahlrichs, Auxiliary Basis Sets to Approximate Coulomb Potentials (Chem. Phys. Letters 240 (1995) 283-290), *Chem. Phys. Lett.* **1995**, *242*, 652–660.

[84] K. Eichkorn, F. Weigend, O. Treutler, R. Ahlrichs, Auxiliary Basis Sets for Main Row Atoms and Transition Metals and Their Use to Approximate Coulomb Potentials, *Theor. Chem. Acc.* **1997**, *97*, 119–124.

[85] F. Weigend, Accurate Coulomb-Fitting Basis Sets for H to Rn, *Phys. Chem. Chem. Phys.* **2006**, *8*, 1057–1065.

[86] F. Weigend, A Fully Direct RI-HF Algorithm: Implementation, Optimised Auxiliary Basis Sets, Demonstration of Accuracy and Efficiency, *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

[87] M. K. Armbruster, F. Weigend, C. van Wüllen, W. Klopper, Self-Consistent Treatment of Spin-Orbit Interactions with Efficient Hartree-Fock and Density Functional Methods, *Phys. Chem. Chem. Phys.* **2008**, *10*, 1748–1756.

[88] W. Smith, The DL_POLY Molecular Simulation Package, CSE Department, STFC Daresbury Laboratory.

[89] ChemShell, a Computational Chemistry Shell, see `www.chemshell.org`.

[90] J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander, P. Sherwood, DL-FIND: An Open-Source Geometry Optimizer for Atomistic Simulations, *J. Phys. Chem. A* **2009**, *113*, 11856–11865.

[91] D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin, P. A. Kollman, Amber 11, University of California, San Francisco, **2008**.

[92] T. G. A. Youngs, Aten-An Application for the Creation, Editing, and Visualization of Coordinates for Glasses, Liquids, Crystals, and Molecules, *J. Comput. Chem.* **2010**, *31*, 639–648.

[93] W. Humphrey, A. Dalke, K. Schulten, VMD: Visual Molecular Dynamics, *J. Molec. Graph.* **1996**, *14*, 33–38.

[94] F. Weigend, M. Häser, H. Patzelt, R. Ahlrichs, RI-MP2: Optimized Auxiliary Basis Sets and Demonstration of Efficiency, *Chem. Phys. Lett.* **1998**, *294*, 143 – 152.

[95] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N.

Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09 Revision A.02, Gaussian Inc. Wallingford CT 2009.

[96] J. F. Matthews, C. E. Skopec, P. E. Mason, P. Zuccato, R. W. Torget, J. Sugiyama, M. E. Himmel, J. W. Brady, Computer Simulation Studies of Microcrystalline Cellulose Iβ, *Carbohydr. Res.* **2006**, *341*, 138–152.

[97] J. Hoja, A. F. Sax, K. Szalewicz, Is Electrostatics Sufficient to Describe Hydrogen-Bonding Interactions?, *Chem. - Eur. J.* **2014**, *20*, 2292–2300.

[98] O. Christiansen, H. Koch, P. Jørgensen, The Second-Order Approximate Coupled Cluster Singles and Doubles Model CC2, *Chem. Phys. Lett.* **1995**, *243*, 409–418.

[99] T. H. Dunning Jr., Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen, *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[100] R. Bukowski, W. Cencek, P. Jankowski, M. Jeziorska, B. Jeziorski, S. A. Kucharski, V. F. Lotrich, A. J. Misquitta, R. Moszyński, K. Patkowski, R. Podeszwa, S. Rybak, K. Szalewicz, H. L. Williams, R. J. Wheatley, P. E. S. Wormer, P. S. Żuchowski, SAPT2012: An *Ab Initio* Program for Many-Body Symmetry-Adapted Perturbation Theory Calculations of Intermolecular Interaction Energies, University of Delaware and University of Warsaw, **2012**. `http://www.physics.udel.edu/~szalewic/SAPT/SAPT.html`.

[101] A. J. Misquitta, R. Podeszwa, B. Jeziorski, K. Szalewicz, Intermolecular Potentials Based on Symmetry-Adapted Perturbation Theory with Dispersion Energies from Time-Dependent Density-Functional Calculations., *J. Chem. Phys.* **2005**, *123*, 214103.

[102] A. Hesselmann, G. Jansen, M. Schütz, Density-Functional Theory-Symmetry-Adapted Intermolecular Perturbation Theory with Density Fitting: A New Efficient Method to Study Intermolecular Interaction Energies, *J. Chem. Phys.* **2005**, *122*, 014103.

[103] R. Bukowski, R. Podeszwa, K. Szalewicz, Efficient Calculations of Coupled Kohn-Sham Dynamic Susceptibility Functions and Dispersion Energies with Density Fitting, *Chem. Phys. Lett.* **2005**, *414*, 111–116.

[104] R. Podeszwa, R. Bukowski, K. Szalewicz, Density-Fitting Method in Symmetry-Adapted Perturbation Theory Based on Kohn-Sham Description of Monomers, *J. Chem. Theory Comput.* **2006**, *2*, 400–412.

[105] DALTON, A Molecular Electronic Structure Program, Release 2.0 (2005), see `http://www.kjemi.uio.no/software/dalton/dalton.html`.

[106] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, General Atomic and Molecular Electronic Structure System, *J. Comput. Chem.* **1993**, *14*, 1347–1363.

[107] A. J. Misquitta, K. Szalewicz, Symmetry-Adapted Perturbation-Theory Calculations of Intermolecular Forces Employing Density-Functional Description of Monomers., *J. Chem. Phys.* **2005**, *122*, 214109.

[108] D. R. Lide, CRC Handbook of Chemistry and Physics *82nd ed.*, CRC Press, Boca Raton, London, New York, Washington D.C., **2001**.

[109] R. A. Kendall, T. H. Dunning, Jr., R. J. Harrison, Electron-Affinities of the 1st-Row Atoms Revisited - Systematic Basis-Sets and Wave-Functions, *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[110] F. Weigend, A. Köhn, C. Hättig, Efficient Use of the Correlation Consistent Basis Sets in Resolution of the Identity MP2 Calculations, *J. Chem. Phys.* **2002**, *116*, 3175–3183.