Molecular Dynamics Simulation Study of Polyelectrolyte Adsorption on Cellulose Surfaces

Dissertation

dem
Fachbereich Chemie der Universität Dortmund
eingereicht

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften

von Dipl. Chem. Oliver Biermann
aus Unna, Westfalen

Diese Arbeit wurde in der Zeit vom 1. April 1998 bis zum 31. März 2001 am Max-Plack-Institut für Polymerforschung in Mainz erstellt.

Zusammenfassung

Es wurden verdünnte (≈ 2.5 Gewichtsprozent) Lösungen von Polyelektrolyten ((Carboxy methyl) Zellulose und Polyacrylat) in Wasser simuliert. Wasser und Natrium-Gegenionen waren Teil des Simulationsmodells.

Unsere Molekulardynamik Simulationen von CMC führen in Abhängigkeit vom Substitutionsmuster zu zwei unterschiedlichen CMC-Strukturen: Ein CMC-Oligomer behält seine gestreckte Konformation, ein zweites nimmt eine geknäulte Form an. Diese globulartige Struktur ist unabhängig von der Anfangskonformation (gestreckt oder schon geknäult). Polyacrylsäure (PAA) ist gestreckt.

CMC und PAA haben verschiedene inter- und intramolekulare Wasserstoffbrückenbindungen. Die COO^--Gruppe der Polyacrylsäure ist ein starker H-Akzeptor. Durch die negative Ladung ist an PAA mehr Wasser (pro Monomer, pro Masse) gebunden als an CMC. Die beiden simulierten CMC-Oligomere haben 0.029 H-Bindungen/amu (gestreckte Konformation) bzw. 0.019 H-Bindungen/amu (geknäulte). PAA hat 0.083 Bindungen/amu mit Wasser.

Beide Zelluloseoberflächen (110 und $1\bar{1}0$) eines monoklinen Kristalls haben gemeinsame Eigenschaften. Beide Oberflächen sind gegen Wasser stabil. Es findet keine Umstrukturierung in den Simulationen mit reinem Wasser statt. Die Kristallstruktur überträgt sich auf das Adsorptionsmuster von Argon und Wasser über der Grenzschicht. Trotz der OH-Gruppen auf der Oberflächen ist Cellulose lipophil und hydrophob. Wasser wird nicht von der Oberfläche angezogen. Jedoch stört die Gegenwart des Kristalls die Wasserstruktur und es bildet sich ein freies Volumen, in dem sich Argon-Atome aufhalten. Die breiteren Lücken zwischen den Cellbioseketten der 110-Oberfläche legen dort stärker lipophilen Bereiche frei: Auf der 110 Oberfläche adsorbiert mehr Argon. Vorstehenden OH-Gruppen auf der Oberfläche haben ähnliche Eigenschaften auf beiden – der "breiten" 110 und der "schmalen" $1\bar{1}0$– Oberfläche.

Es wurden vier verschiedene Kombinationen von Oberflächen und Polyelektrolyten simuliert. Der direkte Vergleich der Adsoption von CMC und PAA auf der Zelluloseoberfläche gibt das Experiment wieder: CMC adsorbiert, PAA nicht. CMC wird durch viele Wasserstoffbrückenbindungen von der Zelluloseoberfläche, besonders von der O2 und O6 Gruppe gebunden. Zwischen PAA und Zellulose gibt es kaum H-Brücken.

Die adsorbierten CMC-Oligomere stehen mit einem Ende auf der Oberfläche. Der Glucosering in Kontakt mit der Oberfläche richtet sich parallel zur 200 Kristallebene aus. Nur ein CMC-Oligomer liegt flach auf der Kristalloberfläche.

Der $1\bar{1}0$-Zellulosekristall ist in der Gegenwart von CMC oder PAA nicht stabil. Es bildet sich ein Defekt auf der Oberfläche in der Form eines Knicks.

Abstract

The adsorption of two polyelectrolyte ((carboxy methyl) cellulose and poly(acrylate) in water on crystalline cellulose is studied in this work. The multi-component problem has been splitted up into simulations of solutions of the polyelectrolyte (polyanions including sodium counterions) in water, into simulations of the interface of crystalline cellulose towards water. Finally polyelectrolyte-cellulose systems were studied.

Molecular dynamics simulations of diluted ($\approx$ 2.5 weight percent) aqueous solutions of two polyelectrolytes, namely sodium (carboxy methyl) cellulose (CMC) and sodium poly(acrylate) (PAA) have been performed. Water and counterions were taken into account explicitly. For CMC the substitution pattern is important. Simulations of CMC oligomers resulted in two different structures: One molecule takes a stretched conformation, while the other one takes a globule-like, collapsed state. In an additional simulation, which starts from a linear state, the second CMC molecules collapses possibly due to solute-solute hydrogen bonding. PAA is stretched during the whole simulation.

On a local atomistic scale, CMC and PAA have different hydrogen-bond properties. The COO^- groups of PAA can only act as hydrogen bond acceptors, but due to the high negative charge density, there are still more water molecules assembled around PAA than around CMC. There are 0.029 bonds/amu respectively 0.019 bonds/amu to water for the two CMC oligomers, but more than twice as many for PAA: 0.083 bonds/amu. Beside intermolecular hydrogen bonding, there is a significant amount of intramolecular H-bonding for CMC, which is influenced by the COO^- groups, which act as strong H-acceptor. In contrast to hydroxy- and carboxylic groups, ether oxygen are hardly involved in hydrogen bonding.

The water-cellulose interface shows common features for both simulations of the $1\bar{1}0$ and 110 surface of cellulose Iβ. We have simulated a cellulose-crystal with the $1\bar{1}0$ and 110 surface exposed to water. Both interfaces are stable with respect to surface reconstruction and water penetration, at least on the nanosecond time scale and in the absence of defects. Both show essentially the crystal structure of cellulose Iβ, also in the adsorption patterns for water and argon. In spite of the presence of OH groups, both surfaces are non-hydrophilic and lipophilic: Water molecules are not attracted to the surface, as they are equally happy in bulk water. On the other hand, the presence of the surface perturbs the water structure sufficiently to create free volume, in which an argon can dissolve more easily than in the denser bulk water. The wider spacing of cellulose chain allows the 110 surface to expose more of the hydrophobic grooves where lipophilic adsorption takes place. In contrast, the protruding OH groups are less affected by chain packing, so the behavior towards water is similar for both surfaces.

We did a direct comparison of the adsorption properties of poly(acrylate) and (carboxy methyl) cellulose on different monoclinic cellulose surfaces. Four systems have been simulated: 110-CMC, $1\bar{1}0$-CMC and 110-PAA and $1\bar{1}0$-PAA. Every systems consists of the water solvent, a cellulose-crystal and the polyelectrolyte.

As found in experiment, CMC does adsorb onto the cellulose surface, and PAA does not. This is due to a high number of hydrogen bonds from the cellulose surface to the CMC-oligomers. There are hardly any hydrogen bonds between cellulose and PAA. One (carboxy methyl) cellulose is flat on the surface, whereas other oligomers take a end-on-the-surface conformation. The anhydroglucose unit which is in contact with the cellulose crystal aligns parallel to the cellobiose-units of the crystal (200-plane).

Content

1 Introduction

The focus of this thesis is the adsorption properties of polymers on crystalline Iβ-cellulose. Cellulose I is our main interest, as it is the main constituent of cotton. Our adsorbates are sodium (carboxy methyl) cellulose (CMC) and sodium poly(acrylate) (PAA), both polymers belong to the class of polyelectrolytes. Experimentally, one finds CMC to adsorb on cellulose, whereas PAA does not [43]. The physics of the single components (cellulose, PAA, CMC) in aqueous solution was explored by molecular dynamics simulation of oligomers or small crystals in water. The different adsorption strengths of CMC and PAA on cellulose were tested by simulations of mixed systems, consisting of the polyelectrolyte, water and a cellulose-surface (either 110 or $1\bar{1}0$).

The scope of this thesis is therefore determined by the limits of atomistic molecular dynamics simulations, which are in the first place limited by system size and simulated time: Our systems are far to small to be representative for the "real-life". However, results from atomistic molecular dynamics benefit from the possibility to track down differences between CMC and PAA towards single sites or atoms.

Polymers have been in the focus of applied and material science for a long time. The term "*poly*mer", accounts for the repetition of some low-mass building unit (a "*mono*mer") into one long polymer molecule. They belong to the super-class of macromolecules, which do not necessitate a regular repeated monomer patter. The difference between macromolecules and other types of chemical compounds is the repetition of a building unit into one large molecule, with a molecular mass in the order of few thousand to several million atomic mass units. The term "macromolecule" or "giant molecule" was introduced by the German chemist Hermann Staudinger in 1922 [123]. A macromolecule may be classified by its repeat unit, and it's sequence.

While in the past polymers were directly used as they occur in nature (like natural rubber), there are is an increased effort to tailor polymers to special requirements. Industrial techniques include co-polymerization of different monomers or polymerization conditions to obtain different topologies and thereof polymers with distinct properties.

However, a precise and close-to-application development requires a good knowledge of the underlying polymer chemistry and physics. To this end, the influence of a certain monomer, as well as generic polymer properties are important. Polymer properties are determined on different length and time scales from the sub-atomistic to the macroscopic level. On the atomistic level most polymers are – depending on chemical constitution – entirely different. Not until the coarsened level the polymer "chemistry" diminishes and the polymer shows a more generic behavior [9]. However, the behavior (even on larger length scales) of neutral polymers is still determined on short ranges [8]. Excluded-volume interactions between molecules are typically much smaller than the scale which is responsible for physical properties [27]. The range of an induced dipole-dipole interaction is about one nanometer, whereas correlation lengths of polymer chains are tens of nanometers. [56, 75].

The physics of charged polymers is in sharp contrast to this. Polyelectrolytes bear ion-

izable groups, which dissociate into a charged polymer (macroion) and small counterions. Macroions and counterions are often separated, if polyelectrolyte chemistry takes place in solution, where water is the most important solvent. The high dielectric constant enables water to screen electrostatics. At a distance of 0.8 nm (Bjerrum length) the electrostatic energy of two elementary charges in water equals $1\,kT$. Examples of polyelectrolytes are DNA, proteins or synthetic materials like poly(acrylic acid) or sulfonated polystyrene. For polyelectrolytes the short-range excluded-volume interaction is superposed with the long-range Coulomb electrostatic interaction. This additional force changes the physics dramatically. The range of interaction is now in the same order of magnitude with the scale, which determines the mesoscopic and even macroscopic properties of polyelectrolytes. For neutral polymers, local chain structure disappears from a certain level of coarse graining. This does not longer hold true for polyelectrolytes, details like charge distribution on the polymer chain become now important and coupling of different length scales takes place. Scaling theories, which have proven valid for neutral polymers, become useless or at least very complex. Details, like counterion distributions or solvent properties are additional degrees of freedom. The situation can be further complicated by added salt or by different pH-values. For polyacids the degree of dissociation is dependent on the pH. In real-life polyelectrolyte systems – like blood, oil-slud or wash water – all of these effects take place.

In the past few years biopolymers received more consideration, but with the term "biopolymer" mostly referred to proteins or DNA. Carbohydrates (saccharides and polysaccharides) deserve the same attention because of their importance to living things, where they act as energy repository (glucogen, starch) or structural components (cellulose in wood/plants). Carbohydrates play an important role for both animal and herbal energy management. Glucogen is deposited in the liver and released to the blood on need. Plants transform carbon-dioxide and water to energy-rich carbohydrate during photosynthesis (assimilation). This process is the main source of carbohydrates on earth. Certain plants contain nutritious carbohydrates (like starch) and can be eaten to power life directly. In contrast cellulose, which is the main compound of wood, was utilized as a thermal energy source for cooking or later powering machines[1]. Carbohydrates gain industrial importance mainly as regenerating basic materials.

Carbohydrates are divided by size into mono-, di-, oligo- and polysaccharide. Examples for all four classes are glucose, sucrose (table sugar), raffinose and cellulose. As sugars bear polar groups (mostly OH) they are often soluble in water most of their chemistry takes place in aqueous solutions (at least for short carbohydrates). Solid sugars are crystalline, semicrystalline or less defined amorphous [59, 88]. In plants they occur as crystal (microcrystalline cellulose from algae) or as amorphous substance with impurities like wood-cellulose with lignin or hemicellulose [112]. Cellulose takes the leading role in industrially used carbohydrates. From the regrow by photosynthesis there is an annual harvest of 10^{11}–10^{12} tons [104]. Cellulose forms 33% of all plant matter. It is found in cotton and wood plant cells. Furthermore cellulose is found in fungi and cell walls of some algae and bacteria.

Cellulose is the isotactic β-1,4-polyacetal of cellobiose (4-*O*-β-D-glucopyranosyl-D'-glucose). The cellobiose unit consists of two glucose molecules, connected by a β-glucosidic linkage. But even with the molecular structure well defined, there is a broad range of cellulose material from different sources. Cellulose is polymorphic: depending on the origin

[1] This is valid for humans, many animals are able to digest cellulose

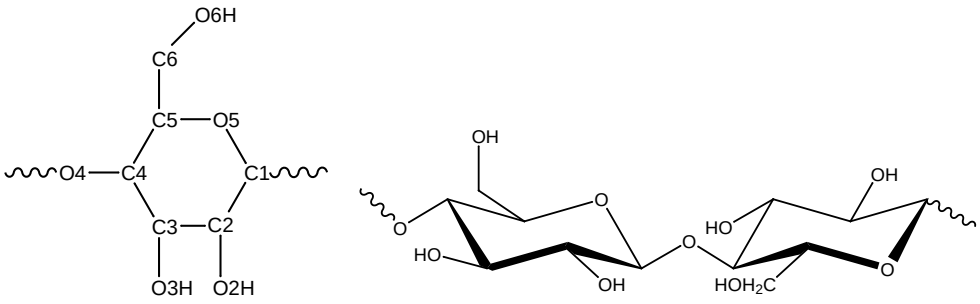


Figure 1.1: *Left:* Nomenclature of an D-glucose unit (AGU). *Right:* Cellobiose repeat unit with $\beta(1,4)$-glucosidic linkage.

or the conditions during isolation, it will adopt various crystal lattice structures as well as amorphous forms. Molecular size, size distribution or secondary structure (amount of crystalline and amorphous regions [82]) may vary as well. If cellulose is mixed with impurities like lignin or hemicellulose it is almost impossible to gain insight into the structure of the substrate without additional treatment or purification. Much research has been carried out on cellulose, but most of it from a technical point of view, with little understanding of the material characteristics on the molecular scale. Owing to this, the structure of many celluloses was unclear for a long time.

Industrial importance comes not only from the ease of cellulose extraction, but from the physical and chemical properties as well. The mechanical strength of cellulose is due to its fibrous structure and with additional good adsorbent properties cellulose is suitable for paper and textile mass production. Thus, a big extent of cellulose is used directly after purification for fiber spinning or paper making [18, 72].

There are three important crystal modifications of cellulose: Cellulose Iα, Iβ and cellulose II. The Iα and Iβ allomorphs are native, cellulose II ("regenerated cellulose") yields from the so called mercerization process. Native cellulose I from bacterial and algal sources is richer in the Iα phase, compared to material taken from higher plants and animals. Furthermore amorphous cellulose exists, mostly together with ordered parts in one microfibril entity.

1.1 Cellulose

1.1.1 Chemical and Physical Properties of Cellulose

Cellulose is the β-1,4-poly-acetal of glucose $(C_6H_{10}O_5)_n$. With respect to this unit, the degree of polymerization runs from 300 for bacterial cellulose up to 5000 for raw cotton. Each cellobiose consists of two units glucose, connected by a β-glucosidic linkage (see figure 1.1 for the structural formula). The conformation of the glucopyranosic six-ring is chair-like and energetically superior to a tray-conformation [45].

The properties of cellulose differ significantly between samples originating from miscellaneous sources or chemical treatments. Solid cellulose is polymorph and adopts various crystal lattice structures. Apart from the most common polymorphs, namely Iα, Iβ and cellulose II, there is evidence for more and exotic crystal allomorphs [74]. For cellulose I

	Iα (triclinic) one chain cell	Iβ (monoclinic) two chain cell
ρ/g cm^{-3}	1.582	1.599
V/nm^3	0.3395	0.6725
a/ nm	0.674	0.801
b/ nm	0.593	0.817
c/ nm	1.036	1.036
α/Deg	117	90
β/Deg	113	90
γ/Deg	81	97

Table 1.1: Selected parameters for cellulose Iα and Iβ. Originally the unit cell parameters have been published by Sugiyama [125], the density and volume was calculated in [50].

the chain alignment is parallel both for triclinic and monoclinic. The former is named Iα, the latter Iβ, some data taken from the literature [50] is summarized in table 1.1, see also [108] for a comprehensive discussion. The chain direction is still a question in dispute for cellulose II, but in [82] it is concluded that a parallel structure is most likely.

Most experiments are using green algal cellulose from *Valonia macrophysa* as a standard for native cellulose I, because of its high degree of crystallinity of 70%. Usually, ordered domains take up to 40–70% of the sample, which is notably higher than the values for regenerated material. The crystallization can be determined either with chemical [13] or physical methods like infrared spectroscopy [57]. Nevertheless the main difficulty for the experimentalist is the inability to separate amorphous and crystalline regions. Strongly related to the fact that cellulose I is a mixture of distinct phases is the problem of determining the crystalline structure. In 1984 Atalla and VanderHart [3, 136] analyzed CP/MAS ^{13}C spectra. They found a splitting patter from 81 ppm to 93 ppm which they assigned to glucose carbon-4 (C4, see figure 2.2) in two subphases in native cellulose. The intensities found are neither constant nor in the

> *"ratios of small whole numbers (as would be true if the same unit cell prevailed throughout all cellulose I crystallites)"* [136].

These results were later confirmed by two-dimensional C-H correlation NMR (HECTOR) [30]. Later, Sugiyama et. al. identified the unit cells of Iα and Iβ as triclinic and monoclinic with electron-diffraction experiments [125] and published relative stabilities: It was shown that the triclinic phase transformed into monoclinic during hydrothermal annealing. The analysis of the diffraction pattern showed similarities between Iα and Iβ molecular conformations. Differences mainly show up in the hydrogen-bond pattern. Sugiyama was able to explain the conversion Iα $\rightarrow$Iβ in terms of hydrogen-bond rearrangements using IR-spectroscopy [126]. The observed IR-bands at 3305 cm^{-1} and 3405 cm^{-1} can be assigned to *inter*molecular hydrogen-bonds. A 3240 cm^{-1} band coincides with expectations for triclinic Iα cellulose, whereas the band near 3270 cm^{-1} is proportional to the amount of Iβ: The former vanishes completely upon annealing of Iα cellulose.

There are no pure triclinic crystals, but in some kinds of cellulose like that taken from *Tunicate* only the monoclinic crystal form is present.

The cellulose surface texture consists of elementary fibrils on a length scale of μm and can be visualized by atomic force microscopy (AFM) [5–7, 113, 119] or transmission electron microscopy (TEM) [44]. TEM images of algal micro-fibrils show quadratic cross-sections of fibers with an edge length of 20 nm. Each microfibril is built of about 1000–1500 single cellulose chains. AFM images [5, 7] show the triclinic surface with a cellobiose repeating unit identical to the c unit cell length of 1.034 nm(figure 1.2. Further, the glucose interval (0.52 nm) and an approximate value for the intermolecular spacing (0.6 nm) were found. Baker identified the triclinic phase looking at the $c/4$ staggered surface pattern shown in figure 1.2. In a second paper on triclinic cellulose they calculated Connolly [23] surfaces of the triclinic Iα (110) surface from atomic positions taken from Sugiyama. For the triclinic crystal they found a good overlap of the real-space AFM-images with the X-ray data. The authors are convinced that this concurrence holds also for the monoclinic Iβ crystal.

1.1.2 Computer Simulations of Cellulose Prior to this Work

In the past, coordinates from electron-diffraction were widely used as starting configurations for computer simulations of cellulose. But one always has to bear in mind the problems connected with the interpretation of raw data. For example it is very difficult to determine the chain alignment in cellulose crystals (parallel or antiparallel). Meanwhile in the case of native cellulose the parallel orientation is established, but this question is left unsolved for regenerated cellulose [74, 82]. Cellulose Iα consists of parallel sheets of hydrogen-bonded chains. This was confirmed by several simulation studies, like the work of Aablo and Heiner [1, 50].

Previous simulation of various systems, including glucose [120], cellobiose and carbohydrates [1, 45, 75] have been undertaken. Some authors simulated crystal like cellulose, but under unrealistic conditions, like mini-crystals with only 24 to 36 monosaccharide residues and with vacuum boundary conditions [36, 121].

Aablo and French [1] calculated the energies for various packings of cellotetraose molecules. They found the *tg* conformation (the angle C4-C5-C6-O6 equals $-60°$) of the exocyclic O6 hydroxy group to be the most stable. Intrachain hydrogen bonding occurs between $\mathrm{O5}\cdots\mathrm{HO3'}$, $\mathrm{O2H}\cdots\mathrm{O6'}$ and intrasheet bonding between $\mathrm{O6H}\cdots\mathrm{O3}$ atoms in adjacent 110 planes (a prime indicating a second glucose ring). The conformational space of cellobiose and higher oligomers under vacuum conditions with respect to glucosidic torsions was explored by Hardy [45] using molecular mechanics, both for a charged and uncharged model. In the uncharged case the minima are in agreement with experimental crystal structures of cellobiose and cellulose. However, in the charged model the appearance of hydrogen bonding distorts this conformation, thereby leading to a new structure, with exocyclic O6 *gt* torsions.

1.1.3 Crystalline Systems and Surfaces

All simulations known from literature deal with the non-orthorhombic unit cells by using monoclinic/triclinic periodic boxes, as implemented in Molecular Dynamics simulation packages like AMBER [24] or GROMOS [135]. Pre-1984 work used the cellulose I crystal

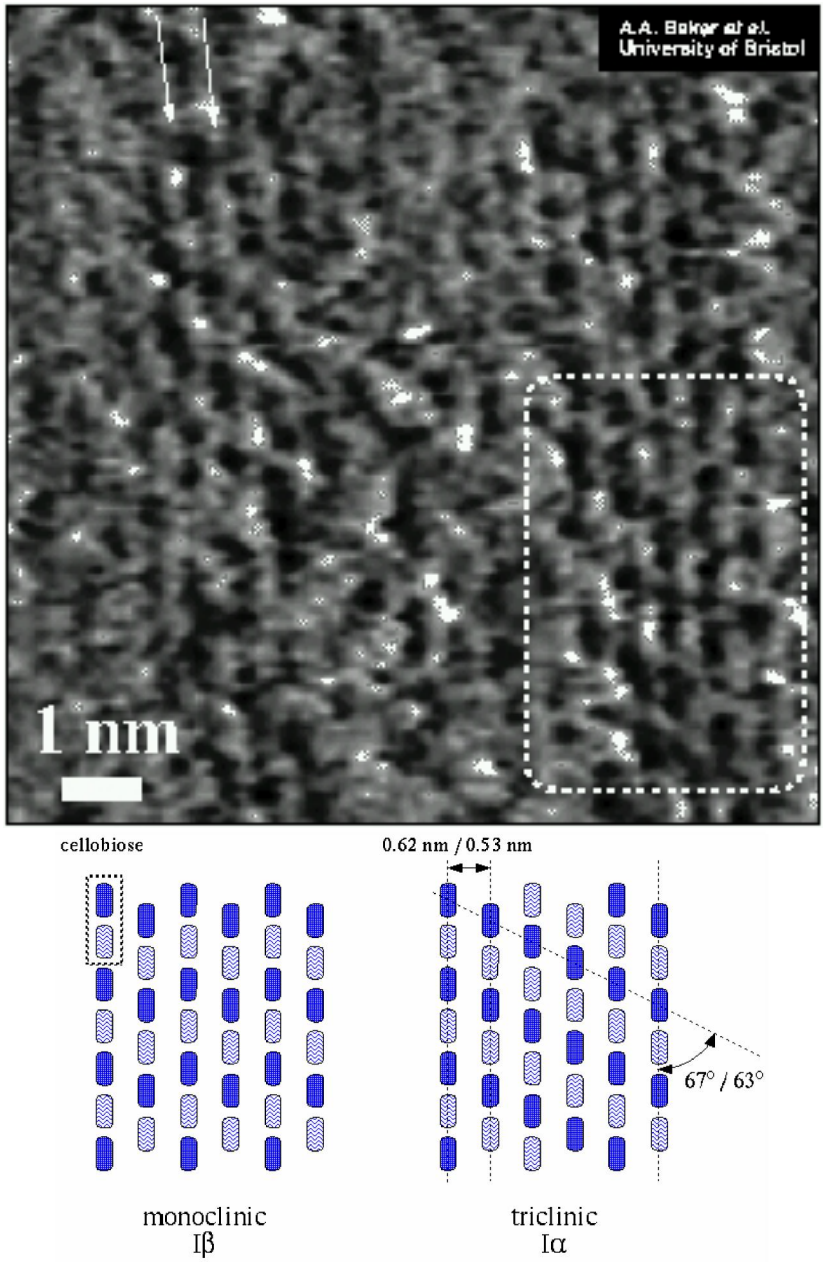


Figure 1.2: Top: A software zoom of an AFM image of microcrystalline *Valonia* surface. The arrows point along the cellulose chain direction, the dotted box highlights an area with spots in the length of the cellobiose repeating interval.
Bottom: The schematic diagram below shows the expected AFM-pattern for monoclinic and triclinic surfaces of cellulose I. Each rectangle represents a single glucose unit. Both pictures taken from Baker et. al. [5].

structure of Blackwell and Gardner [38], which is now considered to be wrong. Other simulations suffer from a restriction to small systems or from the chosen force field, which does not allow full atomistic details.

Recently Heiner [50] performed united-atom simulations of Iβ and Iα cellulose with the GROMOS87 force field starting from X-ray diffraction data of Sugiyama. The monoclinic system was built from an $3 \times 3 \times 3$ array of unit cells, the triclinic system from an $4 \times 6 \times 3$ array. Both runs extend over 1 ns of simulation time. The experimentally observed energy difference between the α and β polymorph was confirmed: the former one is stabilized by $-8.7\,\mathrm{kJ}\,(\mathrm{mol\ of\ cellobiose})^{-1}$. Most surprising is the small tilt angle of 11.5° observed between glucose ring planes in 200 crystal-planes of the monoclinic phase[2]. Alternating chains were termed even and odd, the different tilt angle with respect to the 200 surface was attributed to better interplane hydrogen-bonding [50] (see figure 1.3). More details about the hydrogen bonding pattern were determined using radial distribution functions and energy calculations.

The only existing simulation of an cellulose surface was performed by Heiner [49, 52]. In the first paper the 110 crystal face, in the second paper both the 110, $1\bar{1}0$ monoclinic (β) and the triclinic (α) 100 and 010 surfaces were exposed to water. Only the topmost cellulose layers (layers I in figure 1.3) are structurally affected by hydration. The cellulose properties of the interface layers (which are in contact with the solvent) differ only slightly from that of the crystal's bulk. The odd/even duplicity is absent in the interface layers towards water for Iα and Iβ cellulose and there are changes in the hydrogen-bond patterns, due to competion of cellulose-cellulose bonds with cellulose-water hydrogen bonds.

The cellulose-water interface for both the monoclinic and triclinic crystals was classified in te of "hydrophobic" and "hydrophilic". From a comparison of the surface-water pair distributions the monoclinic $1\bar{1}0$ and the triclinic 100 surface are found to be more hydrophobic than the monoclinic 110 and triclinic 010 surfaces. This becomes evident from the first hydration peak, which is is repelled from the surface and more unpronounced than for 110_β and 100_α.

In their second paper [49] the authors focus on similarities and differences between different cellulose surfaces. They found the monoclinic 110 and triclinic 010 surfaces to be very similar. Likewise, the monoclinic $1\bar{1}0$ surface is similar to the triclinic 100 surface. The latter surfaces are denser and more hydrophilic than the former two. As for 100_β the odd/even differences disappear on the cellulose-water interface.

1.1.4 Carboxy Methyl Cellulose

Cellulose is further substituted to cellulose-esters and -ethers, by either reaction with acid-anhydrides or halogen-carboxylate respectively. An example for cellulose-ether is (carboxy methyl) cellulose (CMC), which is gained by basic conversion of cellulose-slurry with sodium chloro-acetate or chloroacetic acid. CMC is mass-produced, because of its versatile properties. It is used as thickener, film former or protective colloid. Consumer care products take advantage of its non-toxicity, and it is employed particular for food-stuff, and as soil redeposition inhibitor. CMC has a high affinity to cellulose and it is therefore a good coating for textiles. CMC coatings improve paper properties, like ink and surface gloss.

[2] The 200 plane is perpendicular to the unit cell's **a**-vector, there are two 200-planes cutting the unit cell (see figure 1.3)

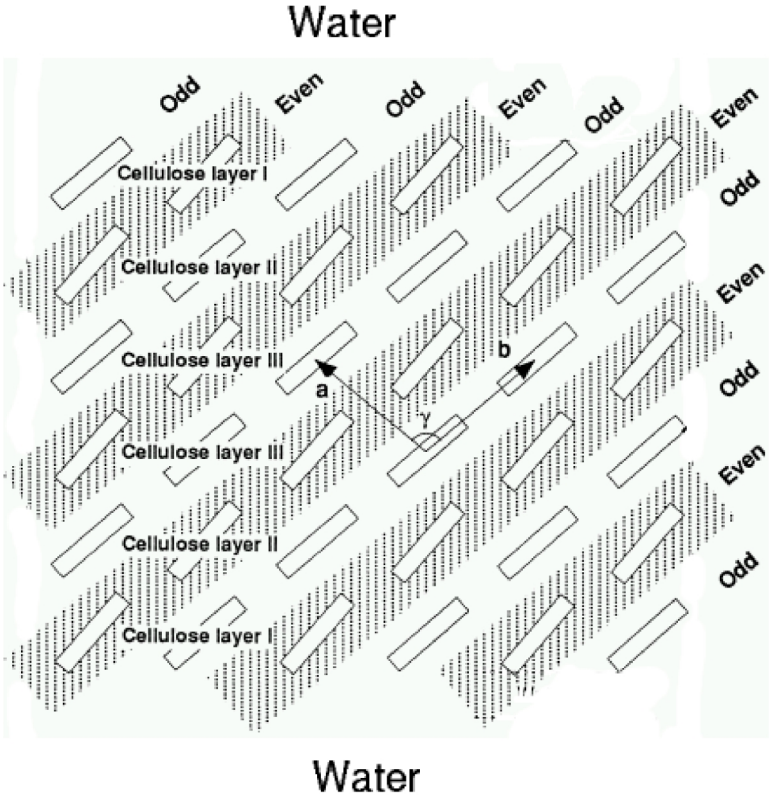


Figure 1.3: Schematic picture of the simulation cell of Heiner, taken from [49]. The scetch shows a monoclinic cellulose crystal, the 110-surfaces (top and bottom) exposed to water. Alternate even and odd 200 crystall planes are shaded white and gray. Glucose rings of even 200-planes are tilted with respect to the 200 planes and glucose rings of even planes.

Crude commercial grade CMC is produced for detergents, oil-drilling or for the paper industry. Product properties of CMC are mostly determined by the degree of substitution (DS) and the substituent distribution within one anhydroglucose unit (AGU) and along the CMC chain. The degree of substitution ranges from zero (no substitution) to three (all three hydroxy groups of a glucose unit substituted). In the case of heterogeneous cellulose derivatisation, statistical substitution patterns and polydisperse CMC are produced. The chemical characterization of CMC is almost restricted to the average degree of substitution, which is measured by titrimetric methods or chromatography [56, 101, 124, 143], where the persistence length of CMC is found by size exclusion chromatography to be 6 nm–16 nm. Owing to this, only little information is available on the interplay between the CMC structure (substitution pattern, degree of substitution) and its macroscopic properties.

The actual substitution pattern of industrial cellulose derivatives can be rationalized by both kinetic and energetic arguments if cellulose is produced as alkaline slurry. Hydroxyl groups can only be carboxy-methylated if the bulk cellulose is swollen and hydroxyl sites are free from hydrogen bonding and accessible to the solvent and substituting agent. The carboxy methylation reaction takes place at the solvent-cellulose phase boundary. From this, the dynamics of swelling and hydrogen-bond cleavage are responsible for the degree of substitution and the substitution pattern. By means of ^{13}C and ^{1}H NMR spectroscopy [4, 21] a substitution frequency O2 > O6 > O3 within the AGU is found, the average degree of substitution is 0.4–1.6. The glucosidic linkage of CMC was hydrolytical cleaved and the on this way information on the distribution along the backbone was lost. The same results can be found by high pressure liquid chromatography, which allows to separate and identify differently substituted units [53, 80]. There are eight different substituted monomer units possible.

CMC polyelectrolyte properties are only understood from a technical point of view. There is a good experience how to tailor a CMC through the production procedure to a certain property, like a high viscosity. The underlying molecular chemistry is still beyond our knowledge. The missing link between microscopic structure and (macroscopic) properties can be established by molecular dynamics simulations.

1.2 Polyelectrolytes

Polyelectrolytes play an important role in industrial chemistry. The fields of application range from tailor-made thickeners to paper finishing or ore preparation. Polysaccharide derivatives represent one interesting class of polyelectrolytes. In particular, cellulose products are important compounds. For our simulation study (carboxy methyl) cellulose (CMC) is chosen as an example for a polyelectrolyte derived from a natural polymer. Aqueous CMC solutions exhibit valuable properties, like a wide range of viscosity, non-toxicity and biodegradability. Particularly for the high-purity consumer-product market (cosmetics, food stuffs), CMC is used. However, pricing becomes more important in bulk applications (clay and ore treatment, oil-drilling). Hence it is desirable to replace some of the high-cost–high-selective chemicals with low-cost equivalents, like poly(acrylic acid) (PAA), which is the prototype of industrial synthetic polyelectrolytes. PAA is the other polymer studied in this work.

Most published work on aqueous CMC and PAA solutions was done experimentally

using chromatography [54], ^{13}C nuclear magnetic resonance [4, 21] and rheological techniques [65]. Theoretical approaches are scarce. We are aware of only one paper [26], which treats CMC by the worm-like-chain theory. This electrostatic theory successfully rationalizes some of the global properties of CMC, but as a rather generic approach it does not allow for detailed predictions on an atomistic time and length scale. Similar restrictions apply also to Monte Carlo simulations of polyelectrolyte chains in a cell model, where the solvent is treated as a dielectric continuum [131]. Especially local interactions such as hydrogen bonds (hydrogen bridges) are neglected in theories and non-atomistic simulations. With two or three hydrogen-bond donor groups per repeat unit and even more acceptors sites (including charged COO^- groups), this type of interaction is likely to be very important for the behavior of CMC in water. Experimental techniques, on the other hand, suffer from different problems: NMR provides averaged local properties. Rheology derives and verifies scaling laws, but different polyelectrolytes lose their chemical "identity" and are show the generic behavior of excluded-volume chains. Atomistic molecular dynamics (MD) simulations cannot overcome all these problems, but they can provide some more detailed information.

Poly(acrylic acid) is important for industrial applications because of good water soluability and as a (strong) model polyelectrolyte in science [70]. The polyanion poly(acrylate) (PAA) is known for only weak adsorption to cellulose, but binds with hydrophilic glass (SiO_2) surfaces [78]. At high concentration, poly(acrylic acid) forms networks, cross linked by hydrogen bonds and entanglements. The water structure around the polyelectrolyte was investigated by Tsukida et. al. [129] using Raman spectroscopy. They found a high perturbation of water-water hydrogen-bonds at a degree of neutralization below 0.3 (30% of all COOH-groups dissociated) and concluded that a certain amount of carboxylic groups enhances water-hydrogen bonds in the polymer vicinity. There is a minimal disruption of hydrogen-bonds near a degree of neutralization of 0.3. Even at higher ionization PAA is assumed not to be fully stretched [95, 137]. The persistence length of PAA as it is ionized to a degree of 93% changes from 1.46 Å to 1.7 Å in semidilute salt free aqueous solution. The local conformation of PAA is assumed not to be dependant of ionization or salt concentration.

Like in the case of CMC, no experimental method has been applied, which goes beyond a macroscopic view of PAA. Methods applied to PAA are rheology [102], viscometry, light scattering [14, 107, 117] or calorimetric methods [67]. The only simulation work was done by Ullner et. al. [131]. This is a Monte Carlo study of one PAA chain in solution. But even though counterions are accounted for explicitly the simulation is done in the generic cell-model for polyelectrolytes. However, there is some agreement, that PAA behaves like a *"flexible coil in a 'good solvent'"* [117] instead of having rodlike structure.

One aim of this thesis is to understand better the structural and dynamic aspects of the hydration of CMC and PAA and to compare the two polymers. To this end, we investigate both the chain properties and the interaction of chains with their immediate solvent environment. Atomistic simulation is confined to the study of small system sizes. However, in combination with coarse-graining methods [9, 127, 128] even some mesoscopic properties may be explored. Thus, a second goal of this study was to produce atomistic structural information, from which coarse grained models of e. g. PAA [110] and the cellulose surface can be generated. The coarse grained models can be used to study the adsorption of polyelectrolytes on cellulose beyond the size and time limitations of atomistic molecular dynamics simulation.

1.3 Techniques and Methods

Computer simulations came into fashion among scientist, as fast hardware became affordable. Early computer simulations were done on the "Maniac" computer in Los Alamos in 1952 by Metropolis and Rosenbluth [85]. From this milestone in scientific computing character and size of simulated systems changed. In the first years there was research to develop, validate and try new methods, even with simple systems like hard spheres, where often a theoretical solution was already present. With the Lennard-Jones potential it was possible to compare the outcome of simulations with experimental results. There was a need for a new method to simulate not only static properties (ensemble averages), but extend simulation to explore dynamical (transport) properties as well. Molecular dynamics (MD) is the new technique. If a given system follows Newton's equation of motion and we know one state of the system, than we can calculate every state of the system (both in the future and in the past). Classical particles are moved by integration of the system's equation of motion in time. By means of this, a molecular dynamics simulation is very simple; after initialization, the simulation cycles through successive molecular dynamics steps. For each step the force is calculated, particle velocities plus positions are updated and, finally, properties of interest are sampled.

This was first done for hard spheres by Alder and Wrainwright in 1957 and later in 1967 by Rahman for Lennard-Jones particles. Later, the method was extended to molecular systems (by the introduction of bonds), different algorithms devoted to handle different ensembles and conditions like flow (non-equilibrium molecular dynamics). In particular the Lennard-Jones potential has proven useful and it is the most common model for non-bonded interactions.

The work of Rahman was pioneering, because it showed the benefits of molecular dynamics simulation over theoretical approaches, which often fail to describe fluids. Moreover MD is valuable also for non-ideal, multi-particle systems.

One recent example to show the versatility of the method is the dissipative particle dynamics method. Even if it is still based on the simple molecular dynamics scheme, it can be used for mesoscopic simulations through a modified equation of motion. The second route to handle large systems is to use more elaborated programs on multi-processor computers. This enlarges the number of simulated particle from about 20 000 on a workstation to several million particles on a super-computer [144].

To mention but a few trends in atomistic molecular dynamics, there are

- the calculation of free energy related properties;
- the application to larger molecular systems, like polymers or biopolymers;
- the programming of user-friendly simulation programs to allow for easy standard calculations;
- or the development of new and better forcefields.

The last item – development of forcefields – is very crucial. Although there exists a variety of different forcefields, none of them describes all faces of a system completely. In consequence, there are many forcefields available and it is not always clear in advance, which one will give the best results to our questions. With other words, there is nothing like a

unique natural forcefield for a given molecular or system: Even approaches with a high number of terms [25] do not necessarily give good results.

The general layout is almost the same for all common forcefields. They divide interactions into non-bonded (Lennard-Jones dispersion repulsion interaction, Coulomb electrostatic interaction) and bonded (bonds, angles, torsions etc.). Some forcefields employ special terms to treat hydrogen bonds or similar phenomena, but this is not very common, as a good description may be obtained by other terms as well.

1.3.1 Forcefields

The Hamiltonian $\mathcal{H}$ of classical systems only depends on the particles' positions ($\mathbf{p}$) and velocities ($\mathbf{v}$) (respectively conjugate momenta $\mathbf{p}$) [77]. With the choice of doing our simulation in Cartesian space, the momenta $\mathbf{p}$ depend on the particle velocities $\mathbf{v}$; the classical Hamiltonian $\mathcal{H}$ is now the sum of a kinetic and potential energy part $V(\mathbf{r})$:

$$\mathcal{H} = \mathcal{H}(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}).$$

While the kinetic part of the Hamiltonian uses only particle masses as parameters, the potential energy part is dealt with by a forcefield. The forcefield is the major choice or input of a simulation. It gives an expression for the potential energy as a function of particle coordinates ($\mathbf{r}$). This expression consists out of different terms, which are usually chosen intuitively to mimic the nature of molecules. The splitting of the potential energy into a sum of bonds, angle and other terms is arbitrary and only reflects a human understanding of chemistry. The second approach to a forcefield is the pragmatic, technical one, where terms are not even meant to have a special physical meaning, but originate from some technical issue or procedure [32, 86]. To name but a few "technical" forcefield contributions, there are position restraint terms to keep atomic sites fixed in space or bond constraints to keep the distance between atoms constant. Nevertheless, the use of force fields instead of true electronic interactions has proven useful in lots of simulations from simple Lennard-Jones systems to much more complex molecular systems.

The negative of the derivative of the potential energy $V(\mathbf{r})$ with respect to the coordinates $\mathbf{r}_i$ equals the force $\mathbf{f}_i(\mathbf{r})$ acting on a particle i; the particle follows the gradient towards low-energy states:

$$\mathbf{f}_i(\mathbf{r}) = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}_i}. \tag{1.1}$$

With respect to Newton's classical equation of motion, the positions and time are connected:

$$\mathbf{v}_i = \frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t}, \quad \mathbf{v}_i' = \mathbf{a}_i = \frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t}, \tag{1.2}$$

where $\mathbf{a}_i$ is the acceleration acting on the particle. The connection to the potential energy and force is given by equation 1.1 and

$$\mathbf{a}_i = m_i \mathbf{f}_i. \tag{1.3}$$

Besides some technical issues, which are explained in section 1.3.5, the motion of a particle can be calculated from the forcefield and one starting conformation.

1.3.2 Non Bonded Interaction

The separation of a forcefield into distinct mathematical terms in molecular simulation is justified usually by computational convenience or reasons of transferability. There is no question, that all interaction would have to be calculated by quantum mechanical methods. Unfortunately, this way is by far too time-consuming, even if fast semi-empirical methods are employed [87]. For systems with more than several thousands or even hundreds of thousands of atoms it is inevitable to use a forcefield with pairwise additive terms. All electronic degrees of freedom are ignored, and every atom is taken as the position of it's nucleus. Methods which rely on this statement are commonly summarized under the term "Molecular Mechanics". Energy functions are called "effective potentials", as they try to incorporate many-body effects into a site-site potential. They do not resemble the potential as it would be correct for two interaction sites in vacuum, but are representative for say two argon atoms in liquid argon. On one hand this is an important breakthrough, as we do not bother about the explicit calculation of many-body terms, but on the other hand this may reduce transferability of a parameter set, as there is an influence of the environment onto a single site's forcefield parameters. This leads to a rough categorization into forcefields for anorganic (crystal), for organic (soft) matter and for solutions, the latter one with a special emphasis on water as solvent. Some forcefields are very biased towards aqueous solutions of (bio)organic compounds like DNA or carbohydrates [84, 135, 142]

To achieve transferability, which is often considered a key property of forcefields, the energy-function is divided into several contributions. To name but a few, there is a bond term, often modeled by a spring, or a bond angle modeled by a harmonic angle potential. Non-bonded interactions are divided into electrostatic and dispersion/repulsion (induced dipole, Pauli repulsion) contributions:

$$E_{\text{nonbonded}}(i,j) = E_{\text{Coulomb}}(i,j) + E_{\text{LJ}}(i,j) \tag{1.4}$$

$$E_{\text{Coulomb}}(i,j) = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \tag{1.5}$$

$$E_{\text{LJ}}(i,j) = 4\epsilon_{ij}\left(\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right), \tag{1.6}$$

where (partial) charges are q_i and q_j, ε_0 is the vacuum permittivity, the Lennard-Jones parameters are σ_{ij} (LJ-radius) and ε_{ij} (LJ-energy). As this process is somehow arbitrary, there exists a zoo of forcefields, which are often tailored to fit special needs. They differ sometimes only in parameter values, but often in the functional form of the interaction [22, 64, 106, 138]. One example is the dispersion energy, which is often expressed as a (computationally cheap) Lennard-Jones potential, but some approaches use the Buckingham-form, which models Pauli-repulsion using an exponential function [2].

Knowing the potential between two sites is only the first steps towards calculating the energy of a N-body system and all forces. Instead of using an order N^2 double loop over all site-site combinations ij, the most efficient way is the use of a cutoff r_c together with a neighbor list [2] to speed up the simulation by a factor of order N. This is a point where physical and technical issues meet and compete. From a physical point of view one wants to take the cutoff as large as possible, but with limited computer resources one should take it as short as possible.

First of all, short-range interactions are truncated at the cutoff and the short-range potential becomes

$$E_{\text{short-range}} = \begin{cases} E_{\text{LJ}}(r) - E_{\text{LJ}}(r_c) & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases} \tag{1.7}$$

The shift $E_{\text{LJ}}(r_c)$ ensures that the energy vanishes at r_c. If we precompute all site pairs, which are within the cutoff plus a safety "skin", the force calculation is only of order N. Getting the pair-pair list is a procedure of order N^2. This algorithm is accurate, until a particle travels across the neighbor list skin into the force cutoff of a site.

From both a physical and technical point of view there is a important difference between the dispersion (Lennard-Jones, see equation 1.6) and the electrostatic (Coulomb, see equation 1.6) energy. The former is a short-range potential, it falls of faster than r^{-d} where d is the spatial dimension of the system. Only the short-range interaction can be truncated at the cutoff, whereas long-range interactions can not, because they are to large to be negligible Rising the cutoff up to the distance where the long-range contributions are in order $\ll \mathrm{k}T$ is not feasible because of computational costs. The Coulomb potential goes with r^{-1}, a cutoff procedure can be ruled out, because the sum over particle-pairs and over the lattice of periodic repeated simulation boxes

$$E_{\text{Coulomb}} = \frac{1}{2} \sum_{i,j=1}^{N} {\sum_{\mathbf{n}}}' \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|} \tag{1.8}$$

is only conditionally convergent [29, 31]. Here L is the box size of the (cubic) box, and $\mathbf{n}$ is a index vector of simulation boxes in three dimensions. A cutoff-procedure either omits significant contributions of the Coulomb-interaction or the cutoff is in the order of some nanometers: The electrostatic interaction of two unit-charges in water equals $1\,\mathrm{k}T$ at 0.8 nm, hence the cutoff is much larger 0.8 nm.

The reaction field method [96, 97] assumes a dielectric continuum outside the cutoff-sphere. The dielectric constant (ε_{RF}) is taken a priori from experiment. The reaction field method comprises two contributions to the electrostatic energy: All interactions within the cutoff sphere are calculated directly and the surrounding environment contributes a polarization energy. The final interaction for the electrostatic potential is given by

$$V_q(r_{ij}) = \frac{q_i q_j}{4\pi\epsilon\epsilon_0} \left(\frac{1}{r_{ij}} + \frac{\epsilon_{\text{RF}} - 1}{2\epsilon_{\text{RF}} + 1} \frac{r_{ij}^2}{r_c^3} \right). \tag{1.9}$$

The implementation of this equations into a molecular dynamics program is straightforward and computationally efficient [90]. The effect of the reaction field is that the force is zero at the cutoff for infinite ε_{RF}, as can be shown by differentiation of equation 1.9. As ε_{RF} enters in the form, the value of $\varepsilon_{\text{RF}} = 78$ (water) is practically equivalent to infinity (see figure 1.4).

One advantage of the reaction field method is that it emphasizes the continuum character of a system and does not include any artifical periodicity. However, as we deal mostly with aqueous systems, where a strong dielectric screening takes place, the reaction field method should give good results, even for polyelectrolyte systems like (carboxy methyl) cellulose. The dielectric constant ε_{RF} for the continuum medium outside the cutoff-sphere is readily known for water or cellulose crystals.

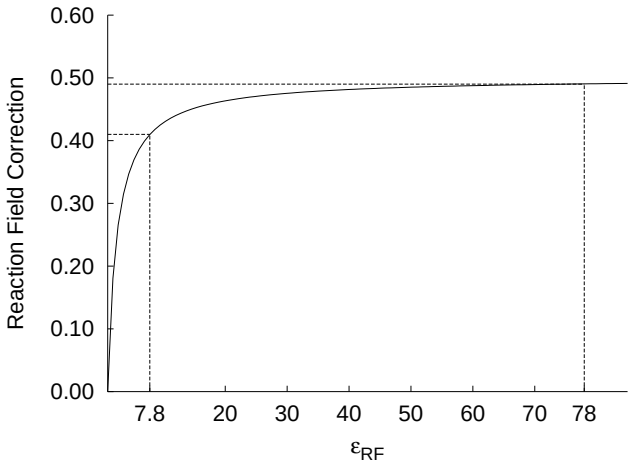


Figure 1.4: Reaction field correction $((\varepsilon_{\mathrm{RF}}-1)/(2\varepsilon_{\mathrm{RF}}+1))$ [96, 97] to the Coulomb energy as a function of the reaction field dielectric constant. For high values of the dielectric constant, the exact value is unimportant, as the curve is flat. Therefore the difference of the reaction field correction between water and cellulose is not as drastic (20%) as the dielectric constants suggests ($\varepsilon_{\mathrm{RF}}$(Cellulose) ~ 7 respectively $\varepsilon_{\mathrm{RF}}$(Water) ~ 78.)

For compound system (like interfaces) this method has one drawback, that is that we can only introduce one dielectricity constant $\varepsilon_{\mathrm{RF}}$ into a simulation. If two values are close, this is not a problem, because $\varepsilon_{\mathrm{RF}}$ entered equation (1.9) as

$$\frac{\varepsilon_{\mathrm{RF}}-1}{2\varepsilon_{\mathrm{RF}}+1}. \tag{1.10}$$

So in short the reaction field is the most pragmatic approach to deal with long range electrostatics at low computational costs and reasonable description of long-range forces in polar environments.

1.3.3 Bonding Interaction

Bonding interactions are somehow better to understand, because their concept is rather intuitive. Usually they define some minimum energy state in terms of a equilibrium angle or bond distance. Deviations from this value impose an energetic penalty. The only exception are bond lengths, which – if treated by a harmonic spring – would require very tiny time-steps and is thus not feasible. The bond vibrations of large molecules are of no interest. Therefore the harmonic bond potential is replaced with a rigid constrained bond, which on one other hand introduces additional calculations into the molecular dynamics simulation, but one the other hand allows to increase the integrator's time-step Δt by one order of magnitude [12].

1.3.3.1 Constraint Dynamics

The SHAKE procedure of Ryckaert, Ciccotti and Berendsen [12] is one of the most explained and cited[3] paper in molecular dynamics. The SHAKE method allows us to con-

[3] 2000 journal-citations found in the Web of Science Database on January 2001.

Formula		Comment
$V_{\text{bond angles}} =$	$\frac{k_\phi}{2}(\phi - \phi_0)^2$	Harmonic bond angle potential, parameters are k_ϕ and ϕ_0, ϕ is the angle between three atoms.
$V_{\cos \text{ bond angle}} =$	$\frac{K}{2}(\cos(\phi) - \cos(\phi_0))^2$	Bond angle potential, which is harmonic in the angle cosine instead of the angle itself, parameters ϕ_0 and K.
$V_{\text{HD}} =$	$\frac{k_\delta}{2}(\delta - \delta_0)^2$	Harmonic dihedral angle potential, parameters are k_δ and δ_0.
$V_{\text{torsions}} =$	$\frac{k_\tau}{2}\left(1 - \cos\left(n(\tau - \tau_0)\right)\right)$	Torsional potential with periodicity n, equilibrium dihedral angle τ_0 and force constant K_τ.

Table 1.2: List of bonding interactions as defined in Yasp [89]. Only contributions which are present in the GROMOS [135] forcefield are listed.

sider atomic connectivity without using harmonic bonds. Valence bonds vibrate at high frequency and impose a small integration time-step to a simulation. SHAKE now alleviates this shortcoming by fixing (constraining) the distance r between two sites to a parameter value $r = a$. This equality is usually written down in the form of a holonomic constraint:

$$r_k^2 - a_k^2 = 0. \tag{1.11}$$

First the unconstrained motion of all atoms are calculated, and after the equation of motion is expanded by the introduction of a constraining force (as a Lagrange multiplier).

The resulting equations of motions are now solved in an iterative fashion until all constraints k in equation 1.11 are within some tolerance. Our simulation package Yasp [90] uses a special flavor of the SHAKE algorithm, which performs well on vector machines like the Cray T90 or NEC SX5 [93]. As a final result of SHAKE one may raise the simulation time-step Δt into the order of magnitude of 0.001 ps to 0.002 ps.

A comprehensive introduction to SHAKE is provided in the textbooks of Allen [2] and Leach [76].

1.3.3.2 Other Forcefield Terms

Beside the non-bonded Lennard-Jones and Coulomb interaction and the constrained bonds there are several other force field terms. They are computational cheap and more generic as for example the non-bonded potential parameters. Examples include bond angles and bond torsions, all bonding interactions used in this study are listed in table 1.2.

1.3.4 Periodic Boundary Conditions

Periodic boundary conditions are the way to generate pseudo-infinite systems, thus simulations which do not suffer from boundary or edge effects. Periodic boundaries are achieved

by putting a grid of copies around the central simulation box. The algorithm ensures that no inter-atomic distance in one direction is larger than one box-length (this holds strictly only for orthorhombic boxes). To calculate the force on a site in the central box, neighbors from the central and surrounding boxes are used, if they are within the cutoff distance. If one atoms travels out of the central box, it re-enters it at the opposite side of the box. The internal coordinate format of `yaspmd` does not store these folded but unfolded coordinates, so that the folding is applied in the force loop. To avoid self-interaction the box has to be larger than two times the cutoff.

1.3.5 Equation of Motion and Integrator

The integration of Newton's equations of motion is done using the leap-frog algorithm, which is a descendent of the Verlet-Method [2, 76]. The leap-frog method calculates velocities and positions with a shift of one half of a time-step Δt:

$$\mathbf{v}_i\left(t_n + \frac{1}{2}\Delta t\right) = \mathbf{v}_i\left(t_n - \frac{1}{2}\Delta t\right) + \Delta t \frac{\mathbf{f}_i(t_n)}{m_i} + \mathcal{O}(\Delta t^3) \tag{1.12}$$

$$\mathbf{r}_i\left(t_n + \Delta t\right) = \mathbf{r}_i + \mathbf{v}_i\left(t_n + \frac{1}{2}\Delta t\right)\Delta t + \mathcal{O}(\Delta t^4). \tag{1.13}$$

Here $\mathbf{r}_i$, $\mathbf{v}_i$ and $\mathbf{f}_i$ are position, velocities and force of one particle. The positions are accurate to an order of $\mathcal{O}(\Delta t^3)$ of the simulation time-step.

1.3.6 Manostat and Thermostat Using Weak Coupling

The thermodynamic variables temperature T and pressure p are straightforward to access in a molecular dynamics simulation. For the temperature one has to sum up the kinetic energy of all particles divided by the number of degrees of freedoms (equipartition theorem [20]):

$$T = \left\langle \frac{1}{3N\mathrm{k}} \sum_N m v^2 \right\rangle, \tag{1.14}$$

where N is the number of sites, k_B the Boltzmann constant and the angle brackets $\langle\rangle$ denote ensemble averaging.

For the pressure, the situation is more complex: Two terms – A and B – contribute to the pressure:

$$p = \frac{1}{3V}\left(\underbrace{\left\langle \sum_N m v^2 \right\rangle}_{\mathsf{A}} + \underbrace{\left\langle \sum_N \mathbf{f}\,\mathbf{r} \right\rangle}_{\mathsf{B}}\right) \tag{1.15}$$

where $\mathbf{f}$ is the total force acting on a site and $\mathbf{r}$ is its position vector. Term A is the ideal gas contribution, and B the so called virial, which extends over all site-site interactions. The virial (respectively, the ensemble average $\langle W\rangle$) becomes an additional summand in the ideal-gas equation:

$$p = \frac{N k_B T}{V} + \frac{\langle W\rangle}{V} \tag{1.16}$$

Often molecular dynamics potentials are pair-potentials. In a simulation it is straightforward to express the virial in terms of site-site vectors $\mathbf{r}_{ij}$ instead of position vectors $\mathbf{r}$ [2, 90]. Term B now writes:

$$\sum_{i<j} \mathbf{r}_{ij}\mathbf{f}_{ij}. \tag{1.17}$$

This procedure has disadvantages for simulations with quasi infinite molecules. The position vector form requires every constrained bond to be between the minimum periodic image of two atoms. This is not true for cellulose crystals, where one end of the polymer chain connects to a chain in the adjacent periodic box.

The micro-canonical (*NVE*) ensemble is native to molecular dynamics simulation and is still common practice. However it is more realistic to perform simulations under pressure and temperature control. The Berendsen weak coupling method [11] rescales particle velocities by a factor s_T, which is computed from a first order decay of the actual temperature T towards the target value T_0:

$$s_T = \left(1 + \frac{\Delta t}{\tau_T}\left(\frac{T_0}{T} - 1\right)\right)^{\frac{1}{2}}. \tag{1.18}$$

The parameter τ_T (the relaxation time constant) is to be chosen significantly larger than the integration time-step Δt to be in the *weak* coupling regime, where ensemble averages of the Berendsen-ensemble are close to the exact *NVT*-ensemble.

The pressure is controlled using an equivalent of equation 1.18, but now scales particle positions and box sizes. We use isotropic scaling of the pressure p to the target p_0, except for the surface-simulations:

$$s_p = \left(1 + \frac{\Delta t \kappa_p}{\tau_p}(p - p_0)\right)^{\frac{1}{3}} \tag{1.19}$$

with τ_p as the pressure-coupling constant and κ_p the system's compressibility. In Yasp the atomic version of the virial is used and this scaling applies to all atoms. So if the coupling time τ_p is too short the manostat interferes with the physical trajectory of the system. Anyway, it has been pointed out by Schmitz et. al. that the scaling defines the origin of the simulated system and thus destroys translational invariance [116].

1.3.7 Free Energies

The driving force for chemical reactions and physical processes is the Helmholtz free energy F for canonical (*NVT*)-systems or the Gibbs free energy G for the *NpT*-ensemble. From the definition equation

$$F(N,V,T) = U - TS = -\,\mathrm{k}T \ln Q(N,V,T) \tag{1.20}$$

$$\text{with} \quad Q(N,V,T) = \sum_i \exp(-\beta E_i); \quad \beta = 1/\,\mathrm{k}T. \tag{1.21}$$

F and G depend on the Entropy S of a system, which itself is a function of the accessible phase space. The calculation requires the knowledge of the partition function Q which is the sum over all states E_i (equation (1.21)) of the system. In the case of classical

mechanics the sum over distinct states turns into a continuous integral over all states with the Hamiltonian $\mathcal{H}$:

$$Z(N,V,T) = \frac{1}{N!}\int \exp\left(-\beta\mathcal{H}(\mathbf{q}^N,\mathbf{p}^N)\right)\mathrm{d}\mathbf{q}^N\mathrm{d}\mathbf{p}^N \tag{1.22}$$

In classical mechanics we can separate $\mathcal{H}$ into a kinetic part $\mathcal{K}$ and a potential part $\mathcal{U}$:

$$\exp\left(-\beta\mathcal{H}\right) \approx \exp\left(-\beta\mathcal{K}\right) + \exp\left(-\beta\mathcal{U}\right) \tag{1.23}$$

$$\tag{1.24}$$

We can now integrate equation 1.22 over Cartesian coordinates and momentum in two steps and split the partition function in a kinetic and potential energy part. The former is also known as "ideal gas" contribution, because it describes a system without any interactions. All interactions between particles contribute exclusively to the potential part, which is named excess part, the partition function now becomes:

$$\begin{aligned} Z(N,V,T) &= \left(\frac{2\pi m\,\mathrm{k}T}{h^2}\right)^{3N/2}\frac{1}{N!}\int_V\cdots\int_V \exp\left(-\beta U(\mathbf{q}^N)\right)\mathrm{d}\mathbf{q}^N && (1.25)\\ &= Z(N,V,T)_{\text{ideal}}\,Z(N,V,T)_{\text{excess}}. && (1.26)\end{aligned}$$

The second part of equation (1.26) equals $V^N/N!$ for the ideal gas case. From this equation we can derive the expression of the free energy. With $F(N,V,T) = -kT\ln Z(N,V,T)$, F is the sum of an ideal and interaction-dependent (excess) term:

$$F = F(N,V,T)_{\text{ideal}} + F(N,V,T)_{\text{excess}}. \tag{1.27}$$

A method to access free energy differences is the particle insertion method, which is explained in more detail in the next section.

1.3.8 Particle Insertion

In a molecular dynamics simulation, the complete phase space or even parts thereof can not be completely surveyed, because the simulated system stays in a low energy sub-space. Molecular dynamics are thus not representative for a system's whole phase space. Therefore Q and F cannot be calculated directly. There are methods to circumvent this drawback [66, 73, 114, 115, 134] like umbrella sampling or particle insertion They allow the calculation of derivatives of the free energy with respect to some other system variable. It is now possible to gain the difference of the free energy between two states of the system. For example if the reference system is chosen to be the ideal (gas) state, excess thermodynamic variables comprise the difference to the ideal state and incorporate the (non-ideal) interactions of a system. From this free energy differences, one can calculate the outcome of a given physical process like the solvation of apolar particles in water (coefficients of solubility) or other equilibrium constants. The only prerequisite for calculating the free energy is sampling all states of a system, including high-energy, low-probability (low Boltzmann factor) ones.

The concept of particle insertion was independently introduced by Widom [139] and Jackson & Klein [60]. In molecular dynamics one determines the interaction of a so called, randomly placed "ghost" particle within the ensemble of other simulated particles. For every insertion, the Boltzmann factor $\exp(-\beta U)$ is evaluated, where U equals the interaction

energy of the ghost with all simulated sites:

$$U = \sum_i \phi(r, i).$$

There are many low-probability insertions possible, and they contribute mostly to Q.

There are different technical methods of the particle insertion method, which are all similar in spirit. Either the position of the "ghost" particle is fixed in space an the simulation is advanced without any influence of the ghost, but the interaction energy U and Boltzmann factor are recorded. The result is μ_{ex}, which is the excess chemical potential $\mu = \mu_{\mathrm{id}} + \mu_{\mathrm{ex}}$:

$$\mu_{\mathrm{ex}} = -kT \ln \langle \exp(-U/\,\mathrm{k}T) \rangle \tag{1.28}$$

Other methods access μ_{ex} while throwing in the ghost particle at random positions and evaluating U for each insertion position. More elaborated schemes use some kind of excluded-volume map to find favorable insertion coordinates. However the excluded volume technique has to ensure a uniform sampling, which is often guaranteed for by a weighting correction for probing only limited parts of phase space. The insertion energy can be estimated efficiently by using a grid-method [103]. Every frame of a simulation is partitioned into small grid points (usually with the distance of 1/100 of an atom diameter). Grid points are set to allow or forbid an test insertion, depending on how far the next simulated site is apart (the term "forbidden" is used in the way that no energy calculation has to be performed and the insertion is done, but with infinitely high insertion energy). Every site occupies grid points according to a core excluded volume. We have chosen the Lennard-Jones radius σ_{LJ} to determine the volume of an atom. To allow for high-energy spaces to sampled as well, not the original value of the Lennard-Jones radius, but a scaled-down value $\sigma_{\mathrm{ex.\ vol}}$ is used:

$$\sigma_{\mathrm{ex.\ vol.}} = s\sigma_{\mathrm{LJ}} \tag{1.29}$$

with $s < 1$.

Even if it is not the fastest choice, we have taken s equal to 0.6, which still gives some speedup and is on the safe side regarding the outcome of an insertion.

If a particle is thrown in the simulation box at random position, there is a first check, if it falls into a forbidden grid point, in this case the insertion energy is set to infinity. If a "free" grid point is found, the interaction energy is evaluated with a Lennard-Jones cutoff of 1.0 nm. Other parameters are given in table 2.7 on page 43.

1.3.9 Radial Distribution Functions

Radial distribution functions $g(r)$ are an important tool to describe the interaction of particles or simulation sites. They are useful to determine *local* structure. They are easily connected to X-ray or neutron scattering experiments. As they are reasonably accessible through atomistic molecular dynamics simulation [2, 76] they are used extensively in this work. The radial distribution of two type of atoms a and b is given through

$$g_{ab}(r) = \left\langle \frac{1}{N_a} \sum_{i=1}^{N_a} \frac{1}{4\pi r^2 \rho_b} \sum_{j=1}^{N_b} \delta(r - r_{ij} \right\rangle \tag{1.30}$$

where ρ_b is the particle density (N_b/V) of particle b. The delta function $\delta(r - r_{ij})$ equals 1 if $r - r_{ij} = 0$, otherwise $\delta(r - r_{ij})$ equals 0.

The integral n of $g(r)$ over r defines the number of neighboring particles b with respect to an integration cutoff r' and particle type a:

$$n_{ab}(r') = \rho_b \int_0^{r'} g(r) \mathrm{d}V \tag{1.31}$$

where $\mathrm{d}V$ is the integration over a spherical volume with radius r around a particle.

2 Computational Details

2.1 Interactions and Forcefields – Choice of the GROMOS Forcefield

The actual parameterization of the GROMOS forcefield for carbohydrates [135] has evolved since 1987 [68, 69] and was mainly tested on cyclodextrines [83, 132, 133, 141] but as well on other sugars [71]. Kroon-Batenburg, Bouma and Kroon [74, 75] made use of the GROMOS parameter set for simulations of cellulose in solution and compiled an overview of different parameter sets in conjuncture with Ewald sums by Kouwijzer et. al. [71].

One application of the GROMOS forcefield to crystalline cellulose was reported by Heiner, Teleman and coworkers [49–52]. Their simulations covered both the crystalline phase of cellulose and the interface with water.

The successful simulations and the compatibility of the GROMOS forcefield terms – in particular the treatment of electrostatics *without* an Ewald sum – with our program suite Yasp (version of January 2001) [90] were decisive factors to choose the GROMOS forcefield. A second point is our interest in multi component, heterogeneous systems, with both a cellulose-surface, solvent and a polyelectrolyte solute molecule. Our forcefield of choice should be able to give a good description for every component, not only of the sugar. So the use of a building block based and thus flexible forcefield is sensible.

However, there are plenty alternatives for carbohydrate force fields. Beside generic ones, like AMBER [24], CHARMM [19, 47, 71] and the OPLS [63, 64] parameter sets some authors developed special approaches for carbohydrates. Most of these expert-models have some special application in mind [108, 109], like the exploration of anomeric equilibrium of sugar rings [118, 120, 122]. An overview is given by French [34] and in various articles published in a special issue "Carbohydrate Modeling" of the "Journal of Molecular Structure" [35]. A recent approach was done by Neyertz et. al. [99] to develop a cellulose all-atom potential model from various origins (mainly from other sugar parameters and quantum chemistry). Despite the mixed sources the Neyertz approach reproduces unit cell parameters, thermodynamic stability and moduli in close agreement with experiment. Tests or applications for this forcefield in solution are not known jet. The authors develop the forcefield with PEO-cellulose interface simulations in mind [98, 100].

In contrast to the Neyertz model, the GROMOS approach is a so called "united-atom" forcefield. Aliphatic oxygens are not modeled by an explicit interaction site. Only polar/OH hydrogens are explicitly treated. Aliphatic hydrogen atoms are accounted for by a change of the parameters for the parent carbons, which grow in size σ_{LJ} (+7% compared to all atom forcefield (like Amber [24, 39]) and get a higher minimal energy ϵ_{LJ} (+10%. For organic materials this has proven feasible if the stereochemistry at chiral centers is preserved using additional terms in the parameterization. The GROMOS forcefield tackles this by harmonic dihedral angles, which fixes four atoms in a given tetrahedral geometry. The major benefit is the reduced amount of computer time. However, for our cellulose-

water systems the savings are less pronounced. Even if the united atom model removes one third of all atoms for a glucose ring (7 alphatic hydrogens), there is still the large amount of polar hydrogen sites left.

Torsions in the GROMOS forcefield are considered by (a) the torsion potential (see table 1.2) and (b) modified 1-4 interactions. United atoms separated by exactly three bonds interact through a reduced Lennard-Jones potential. For atoms other than united atoms the 1-4 interaction is not modified. The sum of both terms results in a physical torsional potential. There is usually only one torsional term for each bond i-**j**-**k**-l, but this rule is changed for sugar rings, where additional torsions guarantee for the correct ring puckering. All 1-2 and 1-3 interactions are dealt with by some forcefield term (bonds, angles) all atom pairs with a topological distance greater than four bonds interact by unmodified non-bonded interactions.

We make use of an effective model potential. This is because the parameter set is chosen to incorporate many-body effects by the physico-chemical environment. For example charges are taken to reproduce average polarization effects by the solvent. Because of this, we avoid mixing of different forcefields and used SPC-water (Simple Point Charge) throughout [10], where parameters harmonize well with the GROMOS cellulose forcefield. Furthermore, it is a rigid model and it has the minimal number of sites, which makes it computational efficient. The CMC and PAA force field terms were taken to be compatible with the cellulose parameters. Because of this CMC and PAA are modeled with the GROMOS forcefield as well. The CMC-parameterization is based on cellulose, with extra parameters for the CH_2 -COO^- -group. The CH_2 parameters were taken from an aliphatic sub-chain in the GROMOS forcefield and the carboxylic group is an generic parameter set which is used for all kinds of carboxylic acids in the GROMOS-handbook.

2.2 Computational Details: Polyelectrolyte in Dilute Solution

For the polyelectrolyte simulations in dilute solution (CMC and PAA with counterions, but without a surface), a single oligomer was solvated in about 4000 (see below for exact numbers) water molecules. The water configuration has been prepared from a cubic centered lattice by an equilibration run of 200 ps until density converged to $\rho = 0.955\,\mathrm{kg\,m^{-3}}$ at 333.15 K). The initial polyelectrolytes' configurations were generated from a Z-matrix [61] for a linear molecule. The sodium counterions were placed into the simulation box at random, but not closer than 0.4 nm to any atoms of the solute. All coordinate sets were joined together and overlap was removed by either

- removing water molecules, closer than 0.3 nm to any of the solute atoms;
- pushing overlapping water molecules away from the solute: All water molecules in the vicinity of the solute are moved away from the polyelectrolyte along a solute-water vector $\mathbf{r}$, defined for every water molecule. The displacement vector starts at the polyelectrolyte site, which is closest to the water oxygen and ends at this oxygen. The length of the vector is scaled by an exponential decay function $\exp(|\mathbf{r}|/c)$, c was choosen by trial an error to be 0.6 nm. Water molecules far apart from the solute are hardly displaced at all.

In spite of these precautions all, systems were in a state of very high energy and as a consequence did not run at usual timesteps of $\Delta t = 0.002$ ps. To quench the system into

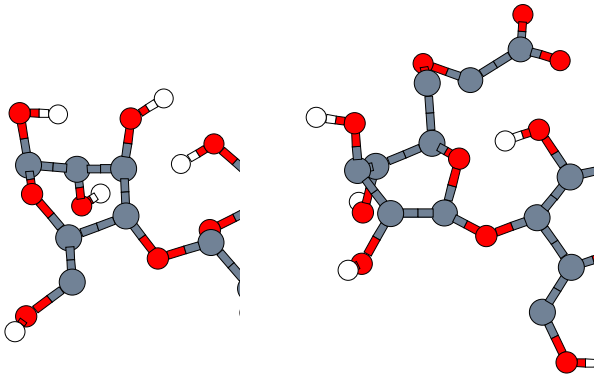


Figure 2.1: Two terminal anhydroglucose units during the initial phase of a CMC I equilibration. The ${}^4C_1$ conformation is lost and the OH-oxygens take an axial position. Both rings are twisted neither in a proper chair nor boat form. The reasons for this behavior are not know, as the present conformation is energetically unfavorable due to the axial oxygen-oxygen repulsion.

a more physical low-energy state we used our MD-program as a primitive minimizer: The equilibration was started at very small timesteps ($\Delta t = 10^{-8}$ ps) and only run for three MD-steps. After this, the velocities were discarded and re-initialized according to a Maxwell-Boltzmann distribution at a temperature of 333.15 K. The timestep was raised by a factor of 1.5 and the procedure was repeated until an integration timestep of 0.001 ps was reached at which a first extended simulation was performed for 65 ps. During this time two artefacts turned out for (carboxy methyl) cellulose:

- Some glucose rings undergo a ring inversion which makes the exocyclic OH-groups change from an equatorial to the axial position (figure 2.1). As the oxygens are very close now, this state should not be favored and is probably only entered through collisions with fast water molecules. However, if this state is present and the simulation is settled, the energy barrier for the transition back is quite high and does not happen. We avoided this artefact by raising the torsional transition barrier in the forcefield (prior to any transitions) to some arbitrary high value of 100 kJ/mol and switched it back to the true forcefield values for later simulation after the initial equilibration of 65 ps.

- Due to the lack of a repulsive core at the hydroxy-hydrogen sites there is no minimum-energy distance in the hydrogen-oxygen potential. This artefact leads to an non-physical strong attraction between these sites, if not compensated by adjacent atoms. We observed a hydrogen-bond between the terminal OH-group of the first glucose unit and ether oxygen O5 of the same glucose ring (see figure 2.2) for atom labeling in cellulose). The oxygen-oxygen distance O4-O5 was as close as 0.22 nm, which caused some significant ring deformation (see figure 2.1). To avoid further problems by this atom-pair, the electrostatic-interaction between the terminal O4H hydrogen and O5 was set to zero.

For the production runs, the water-polyelectrolyte systems were coupled to a heat reservoir of 333.15 K and to a pressure reservoir of 1.013 bar using the weak coupling thermo- and manostat of Berendsen et. al. [11]. The temperature coupling constant was $\tau_T = 0.5\,\text{ps}$ and the pressure coupling constant to $\tau_p = 2.5\,\text{ps}$ at a compressibility of $4.5 \cdot 10^{-6}\,\text{kPa}^{-1}$. The integration timestep of the leap-frog [2] algorithm was $\Delta t = 0.002\,\text{ps}$. Bond lengths were held constant using the SHAKE procedure [12] with a maximum of 500 iterations. The iterative process was stopped "converged" if bond length deviations were less $0.15 \cdot 10^{-8}\,\text{nm}$. The neighbor list was updated every 12 MD steps using a Verlet-scheme. All neighbor sites of a selected site within a cutoff of 1.2 nm are included into the neighbor-list of a site. The non-bonded-force cutoff was set to 1.0 nm with a reaction-field correction for electrostatics [96, 97] (see figure 1.4). The reaction field dielectric constant was 78, and was taken from [37]. Coordinates, velocities and some energies are written every 300 MD steps (0.6 ps) for analysis. The center of mass-motion of the simulation box was removed every 1200 molecular dynamic steps.

Both the simulations and the analysis were done with the molecular dynamics package Yasp, where the molecular-dynamics program `yaspmd` is described in reference [90].

2.2.1 Carboxy-Methyl Cellulose (CMC)

The first polyelectrolyte of interest is (carboxy methyl) cellulose (CMC). The exact molecular composition of commercial CMC is a consequence of the molecule's production history. Usually, OH-groups are randomly substituted by carboxy-methyl units [4, 53] and the degree of substitution (DS) for most commercial CMC's lies in the range DS = 0.6–1.2 substitutions per anhydroglucose unit (AGU) [53, 55]. There is evidence for a preferential substitution during heterogeneous, wet conversion of cellulose to CMC: Heinze and coworkers used liquid chromatography to analyze the substitution pattern of Sodium-CMC. Their analysis was preceded by hydrolysis of the glucosidic link, by which information on the ring-substitution sequence is lost. For one AGU, the following substitution statistics was found [80]: the O2-oxygen was substituted most often (23% of all O2 are substituted), followed by the O6 (15%) and O3 (11%) atom.[1] The glucose nomenclature is shown in figure 2.2. The diversity of molecular substitution patterns of CMC makes it difficult to model such a compound by an oligomer, because there is no "typical" molecule or structure.

To have some variety, two different substituted CMC oligomers were simulated (CMC I and CMC II). Substitution patterns were generated at random, with weights for the three exocyclic oxygens ($P(\text{O2}) = 0.23$, $P(\text{O3}) = 0.12$, $P(\text{O6}) = 0.15$). The molecular composite is summarized in table 2.1. The first molecule (CMC I) consists of seven AGUs and is substituted five times, the second one (CMC II) consist of eight AGUs with six OH groups substituted. The degrees of substitution are DS = 0.70 and DS = 0.75, respectively. Due to three sites, at which substitution can take place (neglecting the chain ends) there are eight different possible substitution patterns on a single AGU. Four of these (O2 O3, O3, O2, O6) are present in the first simulated CMC oligomer and one more (O3 O6) in the second CMC oligomer. CMC I is charged $-5e$, CMC II $-6e$.

All carboxylic acid groups were assumed to be dissociated, neutrality of the systems was maintained by five, respectively six sodium counterions. Molecular weights without counterions are 1440 amu (seven AGUs, CMC I) and 1660 amu (eight AGUs, CMC II), respec-

[1] The percent points are relative to the total number of the specified atom types: e.g. 11%of all O3-atoms are substituted.

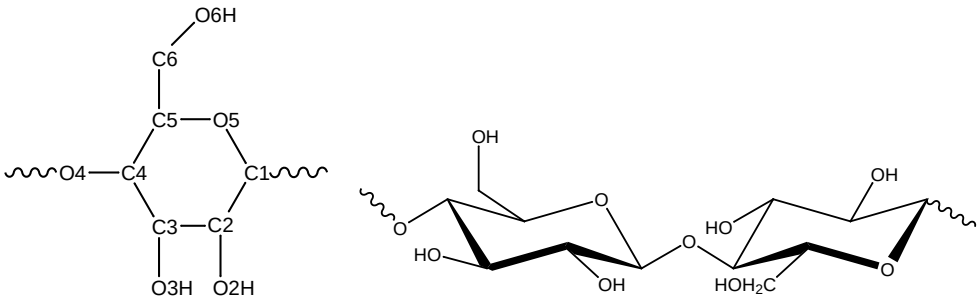


Figure 2.2: *Left:* Nomenclature of an D-glucose unit (AGU). *Right:* Cellobiose repeat unit with $\beta(1,4)$-glucosidic linkage.

	Anhydroglucose unit number							
	1	2	3	4	5	6	7	(8)
O2			I		II	I, II		
O3		II	I		I		II	
O6							I, II	II

Table 2.1: Substitution pattern of both simulated (carboxy methyl) cellulose (CMC) molecules. An entry in the table for one of the CMC molecules (I or II) indicates a substitution at the respective OH-Group (O2, O3 and O6). Anhydroglucose units are numbered from 1 to 7 (CMC I) or 1 to 8 (CMC II).

tively. Molecule CMC I was solvated in 3789, molecule CMC II in 3917 water molecules. The concentrations are about 2.1 and 2.4 weight percent (excluding the sodium counterions). Therefore, both simulations were performed in the dilute regime (no added salt), where polyelectrolytes are expected to be stretched [65]. Other simulation parameters are as given before in section 2.2.

2.2.2 Poly(Acrylate) (PAA)

Poly(acrylate) (PAA) was modeled as one atactic oligomer strand, with 23 repeat units (1636 amu) solvated in 3684 water molecules, the concentration excluding counterions was 2.4 weight percent. All COO groups are negatively charged and there are 23 sodium counterions. Runs were performed under similar conditions as for CMC, including the GROMOS force field [135] as detailed in table 2.2 and 2.3.

Some settings were different from CMC-parameters as the neighbor list, which was updated every 15 steps with a list cutoff of 1 nm. The nonbonded interaction cutoff was 0.9 nm. Weak coupling to a temperature ($T = 333.15$ K) and pressure ($p = 1$ atm) bath [11] with coupling times of $\tau_T = 0.5$ ps and $\tau_p = 2.5$ ps (water compressibility $4.5\,\mathrm{kPa}^{-1}$) was used. The center of mass movement of the simulation box was removed every 1200 MD steps.

Nonbonded Parameters CMC					
Name	Type	σ [nm]	ε [kJ mol^{-1}]	q [e]	m [amu]
C	COO^-	0.33611	0.405870	0.2700	12.0100
CH_1	O-C1-O	0.38004	0.313940	0.4000	13.0190
	C2-OH, C3-OH	0.38004	0.313940	0.1500	13.0190
	C4-OR_2, C5-C-OH	0.38004	0.313940	0.1600	13.0190
CH_2	R-CH_2-OH and				
	R-CH_2-O-CH_2-COO^-	0.39199	0.489590	0.1500	14.0270
	R-CH_2-OCH_2-COO^- and				
	R-O-CH_2-COO^-	0.39199	0.489590	0.2080	14.0270
O	-O-H	0.28706	1.010650	−0.5480	15.9994
	-O-H (chain end)	0.28706	1.010650	−0.5730	15.9994
	R-O-CH_2 and				
	R-CH_2-O-CH_2-COO^-	0.28706	1.010650	−0.3580	15.9994
	R-O-R (ether)	0.28706	1.010650	−0.3600	15.9994
	-COO^-	0.26259	1.725440	−0.6350	15.9994
H	-O-H	0.00000	0.000000	0.3980	1.0080
H	-O-H (chain end)	0.00000	0.000000	0.3730	1.0080
OW	water	0.31650	0.650300	−0.8200	15.9994
HW	water	0.00000	0.000000	0.4100	1.0080
Na^+		0.27300	0.358000	1.0000	22.9898
Additional Nonbonded Parameters PAA					
C	CH_1	0.38004	0.31394	0.000	13.0190
	CH_2	0.39199	0.48959	0.000	14.0270
	CH_3	0.38750	0.73227	0.000	15.0350

Table 2.2: Overview of all non-bonded force field parameter of our simulations. A united-atom model is used, so no explicit aliphatic hydrogens are present. Lennard Jones and electrostatic interactions between atoms less than 3 bonds apart are switched off. Lennard-Jones interactions between unlike atoms are treated using the Lorentz-Berthelot mixing rules [2]. R refers to an aliphatic site. Underlining is used to make the assignment unambiguous.

Bond constraints CMC and PAA

Name	r_0 [nm]
CH_n-$(COO)^-$	0.1530
CH_n-$CH_{1\|2}$	0.1520
CH_n-O	0.1435
C-(OO)	0.1250
O-H	0.1000

Bond angles CMC and PAA

Name	ϕ_0 [degree]	K_ϕ [kJ mol^{-1} rad^{-2}]
CH_1-O-H	109.5	450.0
CH_n-CH_n-CH_n	109.5	285.0
CH_n-CH_n-O	109.5	320.0
CH_n-O-CH_n	109.5	380.0
O-CH_n-O	109.5	285.0
CH_n-(COO)-(OO)	117.0	635.0
(COO)-CH_n-O	109.5	520.0
(OO)-(COO)-(OO)	126.0	770.0
CH_2-CH_1-CH_2	111.0	530.0

Harmonic dihedrals CMC and PAA

Name	δ_0 [degree]	K_δ [kJmol^{-1}rad^{-2}]
C1-O5-O4-O2	35.3	334.9
C5-O5-C6-C4	35.3	334.9
C4-C3-O4-C5	35.3	334.9
C3-O3-C2-C4	35.3	334.9
C2-O2-C3-C1	35.3	334.9
CH_2-(OO)-(OO)-(COO)	0.0	167.5

Torsional angles

Name	τ_0 [degree]	n	K_τ [kJ mol^{-1}]
CMC			
H4-O4-C1-C2	60.0	3	2.520
O4-C1-C2-C3	60.0	3	11.720
O4-C1-C2-C3	90.0	2	0.836
O5-C1-C2-C3	90.0	2	0.836
O5-C1-C2-O2	90.0	2	4.180
O4-C1-C2-O2	90.0	2	4.180
C1-C2-O2-H2	60.0	3	2.520
C1-C2-C3-C4	60.0	3	11.720
C1-C2-C3-O3	90.0	2	0.836
O2-C2-C3-C4	90.0	2	0.836
O2-C3-C3-O3	90.0	2	4.180
C2-C3-O3-H3	60.0	3	2.520
C2-C3-C4-C5	60.0	3	11.720
C2-C3-C4-O4	90.0	2	0.836
O3-C3-C4-C5	90.0	2	0.836
O3-C3-C4-O4	90.0	2	4.180
C2-C1-O5-C5	60.0	3	7.540
C1-O5-C5-C4	60.0	3	7.540
C4-C5-C6-O6	60.0	3	11.720
C4-C3-C6-O6	90.0	2	0.836
O5-C5-C6-O6	90.0	2	4.180
C5-C6-O6-H6	60.0	3	2.520
C6-C5-C4-C3	60.0	3	11.720
O5-C5-C4-C3	90.0	2	0.836
C6-C5-C4-O4	90.0	2	0.836
O5-C5-C4-O4	90.0	2	4.180
C3-C4-O4-C1	60.0	3	7.540
C2-C3-O3-CH_2	60.0	3	7.540
C3-O3-CH_2-(COO)	60.0	3	7.540
O2-CH_2-(COO)-(OO)	0.0	6	2.000
O3-CH_2-(COO)-(OO)	0.0	6	2.000
PAA			
CH_3-CH_1-CH_2-CH_1	60.0	3	11.720
CH_1-CH_2-CH_1-CH2	60.0	3	11.720
CH_2-CH_1-CH_2-CH1	60.0	3	11.720
$CH_{2\|3}$-CH_1-(COO)-(OO)	30.0	6	2.000

Table 2.3: Overview of bonding parameters of our simulations. "(COO)" is carboxylic carbon, "(OO)" carboxylic oxygen. All other atom type names are self-explanatory (see figure 2.2). Analytical forms of force field term are as explained in ref. [90].

Orthorhombic simulation cells are computational more efficient than monoclinic geometries, because they avoid expensive trigonometric functions in the force-loop. We now used a method which allows for an efficient orthorhombic super-cell instead of the native, but unfavorable monoclinic cellulose cell.
Every non-orthorhombic unit-cell $\mathbf{A}$ (matrix representation) can be packed into a larger orthorhombic super-cell ($\mathbf{A}'$). This is simply a linear combination by a coefficient matrix $\mathbf{T}$:

$$\mathbf{A}' = \mathbf{T} \cdot \mathbf{A} \quad \text{with}$$

$$\mathbf{A} = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3)$$

where $\mathbf{u}_{\{1,2,3\}}$ are the cell vectors spanning the unit cell. However, large means usually *very* large and we have to fallback onto a box which is the smallest feasible with all three angles of the new super-cell close to 90°. The deviation α from a strict orthogonal cell is measured by $\alpha^2 = (1 - \det(\mathbf{A}'))^2$ where $\det(\mathbf{A}')$ is the determinate of $\mathbf{A}'$ calculated from the unit vectors.
In the first step, we test all feasible transformation matrices and evaluate the deviation α and sort the results below an angle-threshold according to the super-cell volumes. The smallest volume which has there unit-cell vectors of approximately equal length is taken. In the last step, the almost-orthonormal basis vectors of the super-cell are sheared to gain an orthorhombic cell and scaled to preserve the correct volume of the cell, which was changed by the shearing.

Table 2.4: Generation of an orthorhombic super-cell from a monoclinic geometry.

2.3 Computational Details: Cellulose Surfaces

We examined two different cellulose-water systems. The first one with an interface between the monoclinic 110 surface and water, the second one with the monoclinic $1\bar{1}0$ surface exposed to water. The simulation setup is almost identical with that used by Heiner [49, 52]. A slab of six cellulose layers has two interfaces towards water. The z-axis of the periodic box is parallel to the normal of the respective surface. The cellulose crystal's c-axis (chain-direction) runs along the cartesian x-axis. Eight cellulose chains of each three cellobiose units are staggered with a shift of $c/4$ along the periodic box's y-axis. However, the simulation of the monoclinic cellulose surface employing an orthorhombic simulation box, leads to some distortions of the molecular coordinates. The deviations from the native monoclinic structure are minor and were neglected. The exact procedure, to change the unit cell's geometry is described in figure 2.4.

The angle angle between the 110 and $1\bar{1}0$ surface of 88.858° [126] is close to 90 degree (this angle is not to be confused with the unit cell angle $\gamma = 97°$). From a small change of the monoclinic cell we arrive at an unit cell which packs into an orthorhombic lattice.

Our transformation of the unit cell basis vectors preserved the overall volume of the cell and only changes the angle between the crystallographic a and b axis. As our simulations took place at 333.15 K the thermal expansion of cellulose was accounted for by an isotropic scaling of the unit cell at 300 K by an factor of 1.0113, the thermal expansion coefficient ($13 \cdot 10^{-5}\,\mathrm{K}^{-1}$) of cellulose was taken from reference [37].

This procedure avoids the computationally expensive trigonometric functions in the force loop, which, for a system size of $12 \cdot 10^3$, atoms gains a factor of 1.5 in simulation time. The larger part of the simulated system is liquid, and thus is not affected by the modification of the cellulose crystal structure.

The inter-molecular cellulose-cellulose hydrogen-bonds can still form despite the changed, orthorhombic crystal symmetry: The inter-chain distance in the 110 plane enlarges by 0.005 nm during the monoclinic→orthorhombic transformation, this is about 2% of the length of a hydrogen-bond. By these arguments we are confident that the modification of the crystal structure is permitted.

The simulation periodic box had an elongated shape, with the cellulose at the bottom and free space filled with 2197 SPC (Simple Point Charge) water molecules. The total water volume was about 70 nm^3, filling 40% of the simulation box. The total monoclinic 110 surface was of size

$$2 \times L_x \times L_y = 2 \times 4.91 \times 3.14 = 2 \times 15.4\,\mathrm{nm}^2,$$

and the $1\bar{1}0$ surface of size

$$2 \times L_x \times L_y = 2 \times 4.32 \times 3.14 = 2 \times 13.6\,\mathrm{nm}^2.$$

As the cellulose crystal structure from X-ray diffraction does not contain positions for hydrogen atoms, they were placed 0.1 nm away from the oxygen, with an C-O-H angle of 109°. The hydrogen positions are identical for every cellobiose repeat unit. However during the equilibration every bias towards the initial position vanishes. Example conformations of the 110 and $1\bar{1}0$ system are shown in picture 2.3.

Equilibrium was achieved by fixing the positions of the heavy backbone atoms (five carbons and two oxygen per glucose building unit) in space using position restraints: Every atom of the glucosidic ring, including the linking O4 were tied to their initial (crystal) position with a harmonic spring of strength 3000 kJ/mol nm^2. Only exocyclic atoms (O2, O3, C6, O6 and corresponding hydrogens) were allowed to move unrestrained. The system was filled with 2197 water molecules on a disturbed lattice (fcc with rotational disorder). From this unphysical starting point the system was brought into a geometry suitable for starting the simulation by an initial equilibration phase of 5 ps. During this phase the timestep increases every 3 simulation steps, starting from 10^{-7} ps until a maximum of 0.002 ps is reached. During this phase the system was quenched by removing all velocities every three steps and reassignment from a Gaussian distribution [90] at 333 K. Next a second equilibration (50 ps) with the final simulation timestep of 0.002 ps was performed.

The position restraints were removed for all, but the glucosidic O4 atoms. They, too, were removed after additional 20 ps of equilibration. Trajectory writing was turned on and sampling was begun after another 1000 ps run. The output frequency for trajectory writing was 0.6 ps, the linked-cell [2] neighbor-list was updated every 12 MD-steps. Temperature control was achieved by weak coupling to a heat bath of 333 K [11] with a coupling constant of 0.3 ps. We were restricted to a constant volume simulation, because the endless

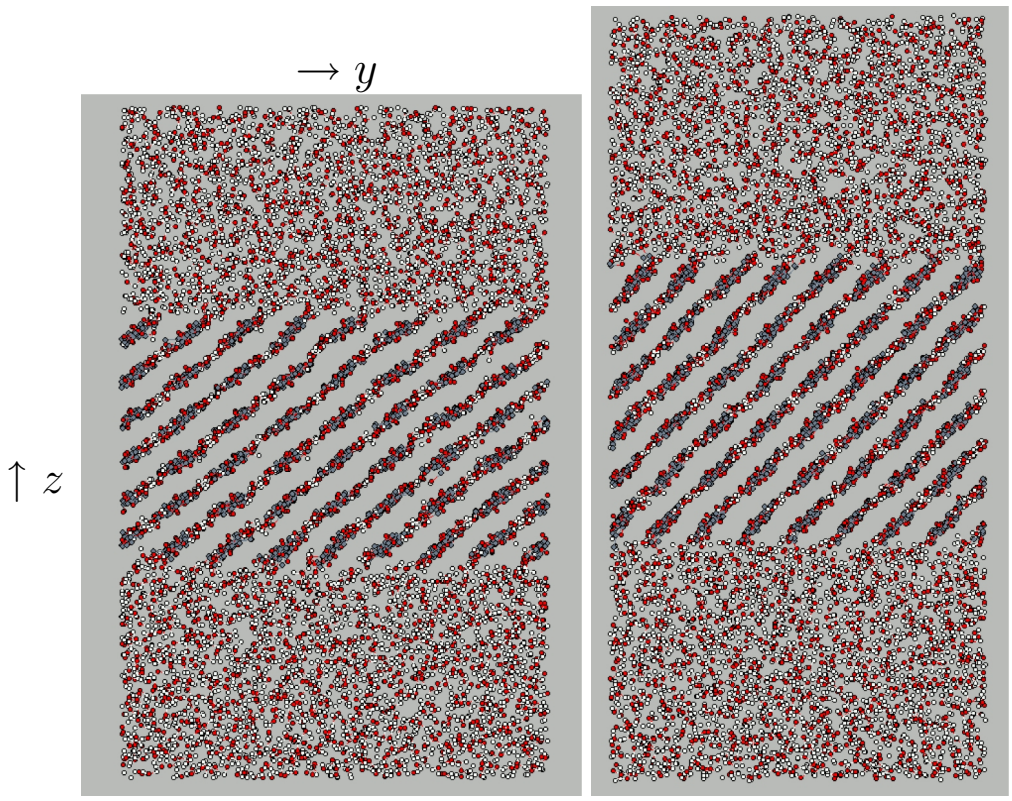


Figure 2.3: Pictures of surface-water simulations after equilibration of about 400 ps. *Left:* 110-cellulose (wide), *Right:* $1\bar{1}0$-cellulose (narrow). The cellulose chains run into the drawing plane (x-direction).

(periodically repeated) cellulose chains are incompatible with correct pressure calculation in the presence of constraints [90]. The cutoff for Lennard-Jones and Coulomb interactions is 1.0 nm, the neighbor list cutoff 1.2 nm.

For our simulations we used the GROMOS forcefield together with SPC-water. The water model was chosen, because it is purported to give better results with the cellulose forcefield [51], even if other water parameters like SPC/E are known for a better description of bulk water. To have an overall simulation setup as consistent as possible, we used a cutoff as used in the GROMOS forcefield: Lennard-Jones interactions were truncated and shifted at a cutoff of 1.0 nm applying an isotropic correction to the energy. Electrostatics were dealt with using the Coulomb-Potential with a reaction-field correction [96, 97], the reaction field dielectric constant equals 78, which is the value for water at 333 K. We are aware that we overemphasize the electrostatics in the cellulose bulk by this procedure, but the reaction field approach does not allow the correct treatment of dielectric heterogeneities.

2.4 Computational Details: Aqueous Polyelectrolyte-Cellulose Systems

The third kind of systems examined were the combination of the two systems above: A cellulose-water interface simulation with CMC or PAA oligomers dissolved in the aqueous phase. As the oligomers of CMC and PAA described above are to large to fit into a reasonable sized cellulose-water simulation box, we used smaller oligomers: For CMC the first molecules CMC I was cut into two pieces, a trimer and a tetramer (see table 2.5) with molecular weights of 619 and 838 amu (excluding counterions). Two trimers and two tetramers including 10 sodium counterions are simulated. PAA was modeled as a 357 amu (excluding counterions) pentamer.

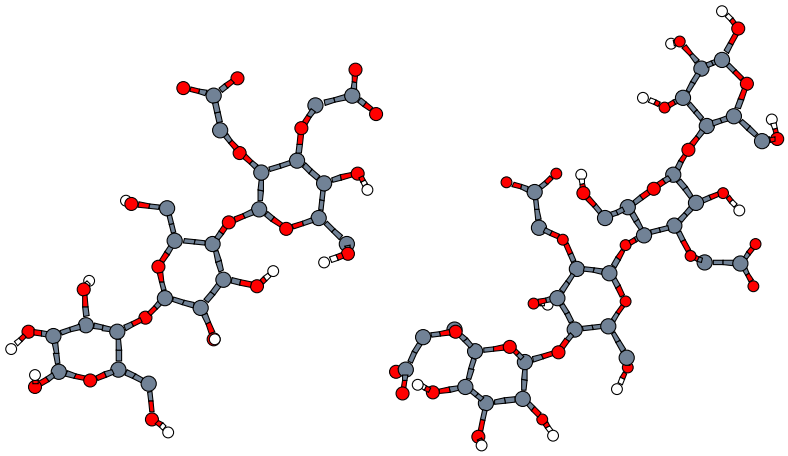
With two polymers and two different surfaces we have four possible solute-surface combinations. The polymers are irregularly placed into the dense water system, by shifting water molecules away from the solute atoms radially as described on page 32. By this procedure we obtain suitable starting coordinates without overlapping atoms.

The cellulose sites are fixed in space using position restraints and the system is quenched into a low energy state (see section 2.2) and equilibration is started afterwards without position restraints. All simulation parameters (temperature, timestep, weak coupling etc.) are as for the cellulose-water system. Cellulose and water coordinates are taken from the respective cellulose-water simulations. An overview over all four solute-surface simulations is given in table 2.6.

2.5 Computational Details: Particle Insertion

The results gained by the particle insertion method [17, 28, 103, 140, 140] are strongly dependent on the sampling uniformity at which the test particle is inserted. We have inserted the test particle at random positions, chosen by the pseudorandom number generator supplied by the Fortran math library of the Fortran77-compiler [33]. In a dense fluid the acceptance ratio of insertions is comparatively low. Most attempted insertion coordinates hit a host atom and thus produce unfavorable high energies, which do not contribute to $\langle \exp(-\beta U) \rangle$ (see section 1.3.8).

The method can be improve by sampling only a certain region, like the cellulose-water interface. All properties by particle insertion are further recorded as a function of the z-coordinate, which is the distance of the point of insertion from the interface. Furthermore, we have also used spatially resolved particle insertion with small sub-volumes. Properties are thus resolved on the surface on an $(x, y$-grid), giving a map of lipophilicity and lipophobicity.



	Trimer			Tetramer			
	1	2	3	1	2	3	4
O2			×			×	
O3			×		×		
O6							×

Table 2.5: Pictures and substitution pattern for CMC trimer and tetramer in the polyelectrolyte-cellulose simulations.

	Simulation Length	
	time [ns]	saved conformations
110 PAA	4.1	6788
$1\bar{1}0$ PAA	5.0	8341
110 CMC	5.3	8750
$1\bar{1}0$ CMC	4.4	7344

Table 2.6: Overview of polyelectrolyte-cellulose simulations.

Inserted argon LJ-parameter [135]	
ε	1.03925 kJ/mol
σ	0.34150 nm
s	0.60
LJ-force cutoff	1.0 nm
Minimum number of accepted insertions	5000
Average Grid Dimensions	
n_x	386
d_x	0.02 nm
n_y	245
d_y	0.02 nm
n_z	157
d_z	0.02

Table 2.7: Input parameters for the particle insertion method.

3 Hydration of Polyelectrolytes Studied by Molecular Dynamics Simulation

3.1 Introduction

In this section compare the sodium salt of (carboxy methyl) cellulose and sodium polyacrylate in dilute aqueous solution. Carboxy-methyl cellulose (CMC) was modeled by two different molecules, a 7-mer and a 8-mer (see section 2.2). In the following they are named CMC I (7-mer) and CMC II (8-mer). For CMC II we did two simulations starting from different geometries. They are named CMC IIa (extended intitial conformation) and CMC IIb (curved initial conformation). For PAA there is only one 23-mer simulated, starting from an elongated geometry.

3.2 Results

3.2.1 Global Chain Properties

Carboxy-methyl cellulose I (CMC I) and CMC II have very different structures: CMC I is an extended, but slightly bent chain, whereas CMC II assumes a cyclic conformation (figure. 3.1 and 3.2). Poly(acrylic acid) (PAA) behaves more like CMC I, having mostly a straight conformation with only some bending (figure 3.3). These structural features cause the end-to-end distance distributions (figure 3.4) of CMC I and CMC II to be very different. The end-to-end distance R_{ete} is calculated between the center of mass of the first and last AGU in the chain. For PAA it is calculated from the center of mass of the terminal repeat units. For CMC I there are two conformers, a major one at $R_{\mathrm{ete}} = 2.6\,\mathrm{nm}$ and a second state at 1.8 nm. This second state corresponds to a bent conformation, which shows up several times in the simulation. Its probability is about 30% of the stretched structure and on average this state is kept for about 50.8 ps until the oligomer's end-to-end distance lengthens again. For CMC II the situation is different: The end-to-end distribution is very narrow if the simulation run is started from an already compact structure. If started from a stretched state, the CMC II molecules collapses at 1 ns (figure 3.5). This sudden collapse results in two peaks for CMC IIb in figure 3.4, an initial one at 3 nm and a final one at 1 nm.

Poly(acrylic acid) shows a wider distribution than either CMC, resulting from the larger number of torsional degrees of freedom. The shoulder at 2.8 nm is due to the bent conformation shown in figure 3.3.

The shapes of the molecules are also reflected in the eigenvalues of the gyration tensor (table 3.1). The stretched conformations of CMC I and PAA, lead to one large and two small eigenvalues, whereas CMC II has more isotropic eigenvalues, with only a factor of two between the largest and second largest value. The correlation between the eigenvalues of the gyration tensor follows opposite trends (table 3.1): Whereas CMC I and PAA exhibit

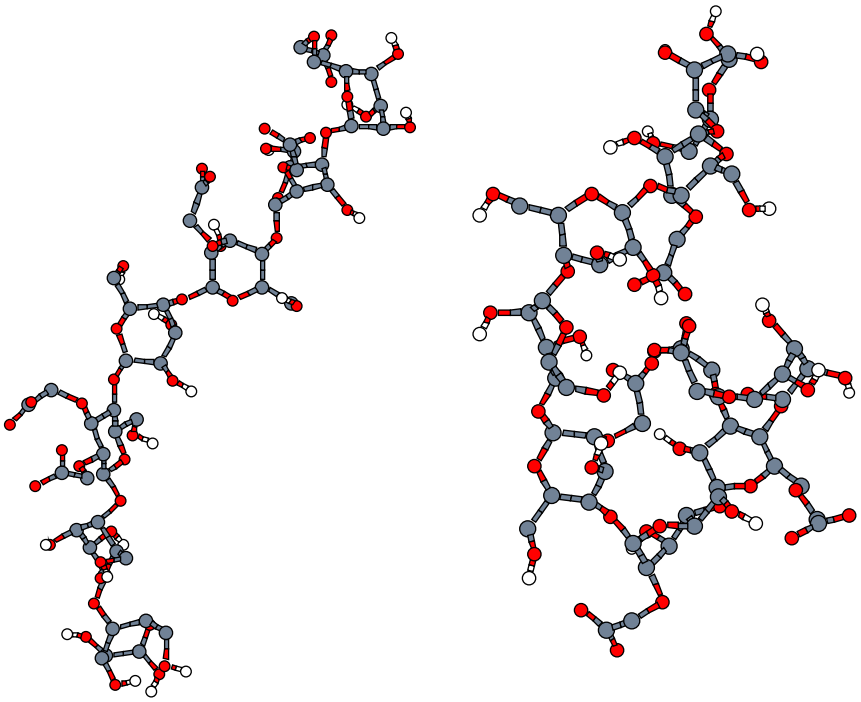


Figure 3.1: Snapshots of CMC I (left at 2.0 ns simulation time) and CMC IIa (right at 3 ns).

	CMC I	CMC IIa	CMC IIb		PAA
			0–0.9 ns	1.2–2.5 ns	
$\langle R_{\mathrm{ete}} \rangle$	2.3 nm	1.01 nm	2.8	0.87	3.6 nm
$\langle R_{\mathrm{gyr}} \rangle$	0.90 nm	0.69 nm	1.00	0.68	1.27 nm
$\langle R_{11} \rangle$	0.69 nm	0.27 nm	1.02	0.27	1.35 nm
$\langle R_{22} \rangle$	0.09 nm	0.15 nm	0.11	0.18	0.21 nm
$\langle R_{33} \rangle$	0.03 nm	0.04 nm	0.04	0.04	0.05 nm
$c_{R_{11}R_{22}}$	−0.86	0.12	−0.78	0.30	−0.76
$c_{R_{11}R_{33}}$	−0.28	−0.43	0.25	−0.13	−0.30
$c_{R_{22}R_{33}}$	0.06	0.07	−0.58	−0.51	−0.15
C_∞	6.7	2.1	7.0	1.8	8.3

Table 3.1: Averages of the end-to-end distance R_{ete}, the radius of gyration R_{gyr} and the eigenvalues R_{nn} of the gyration tensor ($\langle R_{11} \rangle \geq \langle R_{22} \rangle \geq \langle R_{33} \rangle$) of CMC I and II and PAA. The second CMC II simulation is separated into the stretched (0–0.9 ns) and the folded state (1.2–3.0 ns). Additionally, correlation coefficients between the eigenvalues ($c_{R_{11}R_{22}}$ etc.) and characteristic ratios C_∞ are given.

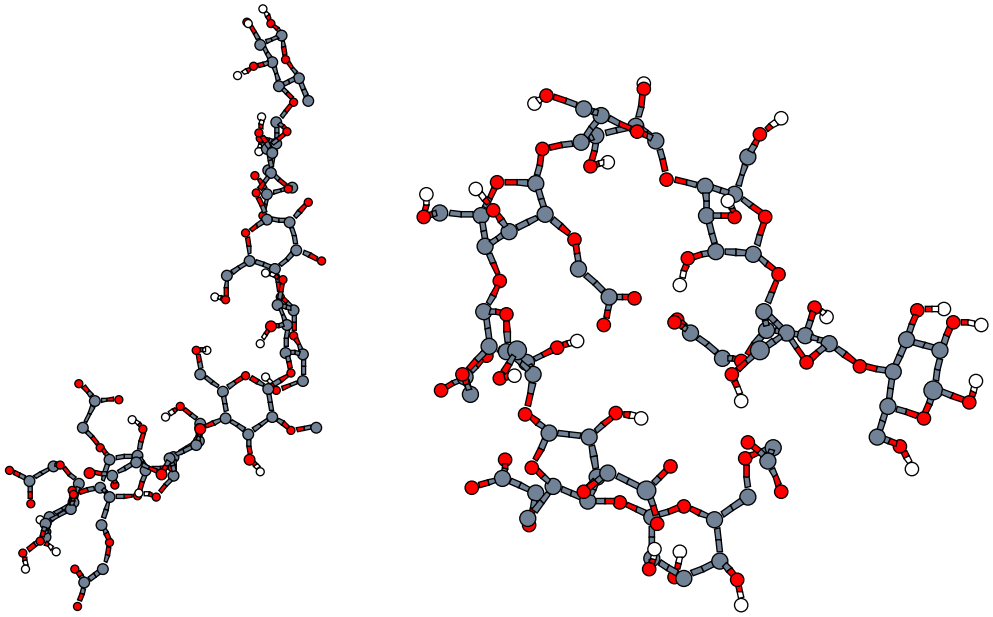


Figure 3.2: Snapshots of CMC IIb during folding (left at 0.3 ns, right at 2.1 ns).

a negative correlation (−0.86 and −0.76) between the two largest eigenvalues, there is a small positive correlation (0.12, 0.30) for CMC II, which is found throughout for CMC IIa and for CMC IIb after folding. This shows how structural fluctuations happen: CMC I and PAA behave like bending rods – they shrink in one dimension and necessarily grow in another. In contrast the ring pulsation of CMC II, leads to the two largest eigenvalues growing and shrinking simultaneously.

3.2.2 Local Chain Properties

Two selected CMC I glucosidic torsional angle's ϕ (C4-O4-C1'-C2') and ψ (C3-C4-O4-C1') trajectories are shown in figure 3.6. The positions of the maxima differ (for all angles) from results found for unsubstituted cellobiose in vacuum [45] and in water [46]: For aqueous solutions there are maxima for (ϕ, ψ) at (284, 63), (274, 193), (294, 242) and (50, 237) degrees (in our definition). We found most of the linking dihedral angles to be unimodally distributed with large fluctuations around a dominant conformation on a time scale of 3–4 ps (terminal links: 2 ps). These fluctuations probe only briefly alternative conformations and rarely lead to lasting transitions. Deviations from unimodality depend on the position along the chain and the substitution pattern: Torsion ψ, for instance, is unimodal for chain-terminating glucosidic links. A second ψ-state is populated only for links connecting

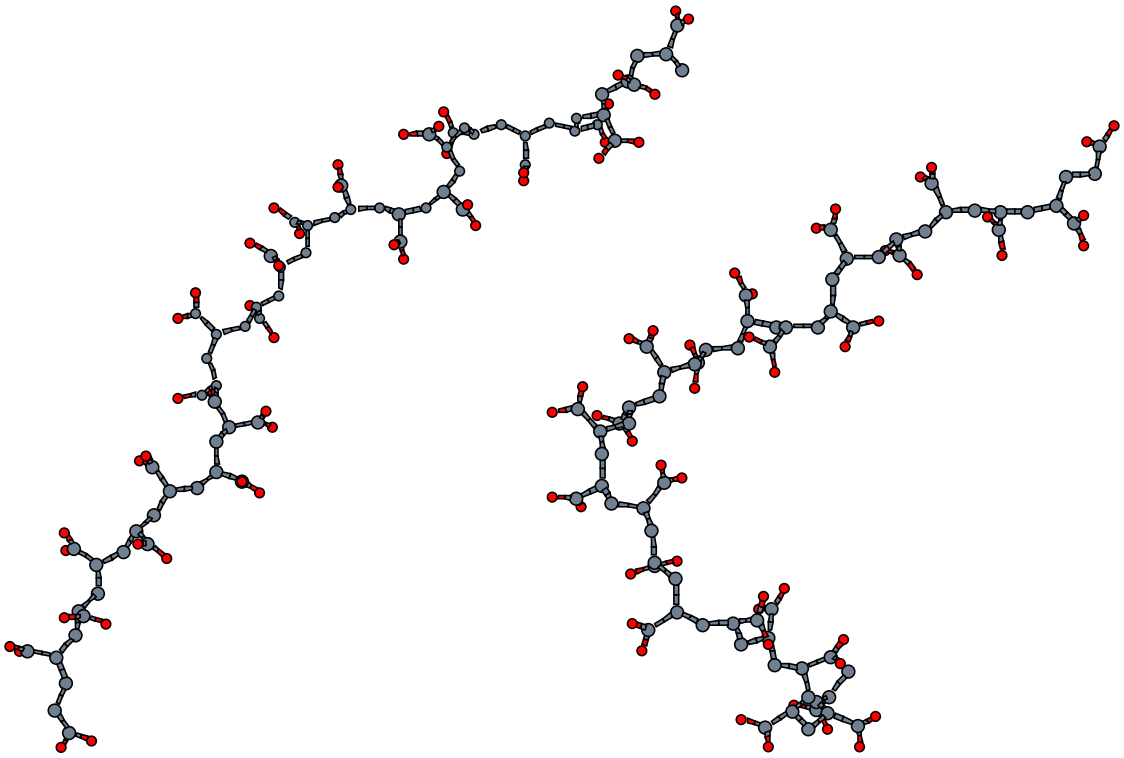


Figure 3.3: Snapshots of the PAA oligomer after 3.5 ns and 4.5 ns.

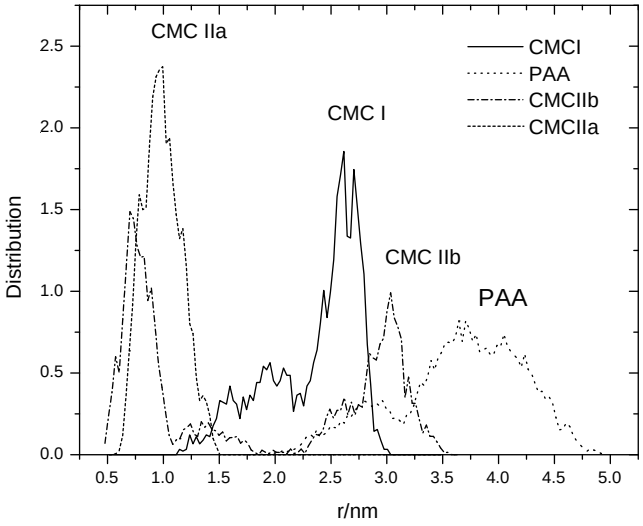


Figure 3.4: Distribution of the end-to-end distance for the polymers simulated.

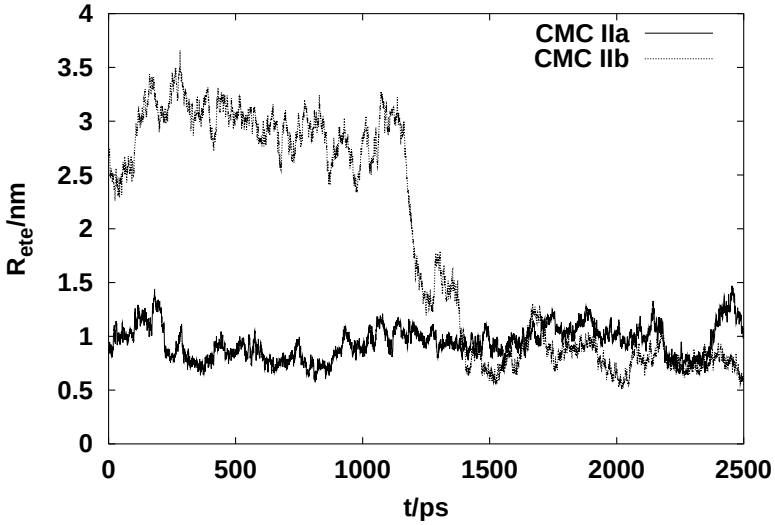


Figure 3.5: End-to-end distance for CMC II (simulations IIa and IIb). The radius of gyration (not plotted) has the same general tendency as the end-to-end distance.

repeat unit 3 with its neighbors. Transitions between two states are scarce and are observed on average every 0.9 ns. Other than this, no significant trends can be found.

There is hardly any torsional dynamics for CMC IIa, which is a consequence of the static, folded state, with torsional links less mobile than those of CMC I. Typical examples are shown in figure 3.7. These distributions are sharp and unimodal.

Torsions are more complex for PAA (figure 3.8), with up to three states for each torsion. The time scale on which the torsions fluctuate (2–5 ps) is of the same order of magnitude as for CMC I.

3.2.3 Hydrogen Bonding

For biopolymers and modified biopolymers hydrogen bonding (H-bonds) is one of the most significant interaction on atomistic scales, both intramolecular (conformation) and intermolecular (solvation, aggregation). Both CMC and PAA undergo intensive hydrogen bonding. In CMC hydroxy groups, carboxylic oxygens, and to a small extent, ether oxygens are involved. Poly(acrylic acid) can only form H-acceptor bonds via carboxylic groups.

Our definition of a hydrogen bond is geometric: If the two oxygens are within a distance of less than 0.29 nm with the hydrogen between them, so that the angle O-H-O is greater 130°, a hydrogen bond is assumed. In our polyelectrolyte-water systems, we find both solute-solute, solute-solvent and solvent-solvent hydrogen bonding.

In addition to the distribution of hydrogen bonds, their lifetimes τ were calculated from bond-correlation functions [92]: For each possible hydrogen-bond we calculate an array, which contains 'true=1' in case of a bond between two oxygen sites, otherwise 'false=0'. The time-correlation function is calculated on this array, integration leads to the lifetime τ. The integration was done by fitting a stretched exponential (Kohlrausch-Williams-Watt) $\exp[-(t/\alpha)^{\beta}]$ in the range 0 to 80 ps, followed by analytical integration (Gamma function).

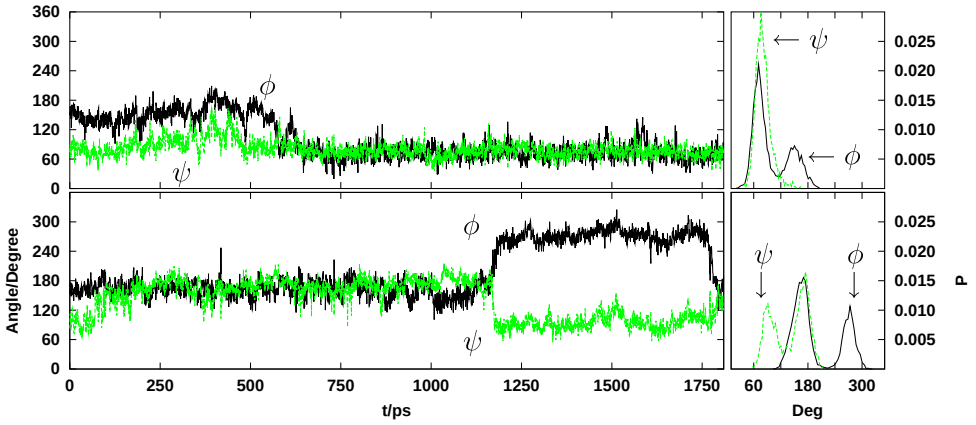


Figure 3.6: Two selected trajectories of ring-linking dihedral angles for CMC I (ϕ : C4-O4-C1'-C2' (black) and ψ: C3-C4-O4-C1' (grey)). The link between rings 1 and 2 is shown at the top and the link from ring 5 to 6 at bottom. Distribution histograms are shown on the right. Both links undergo one clearly visible transition. They, as all other links, also show large-amplitude fluctuations in between, during which alternative conformations are only briefly probed.

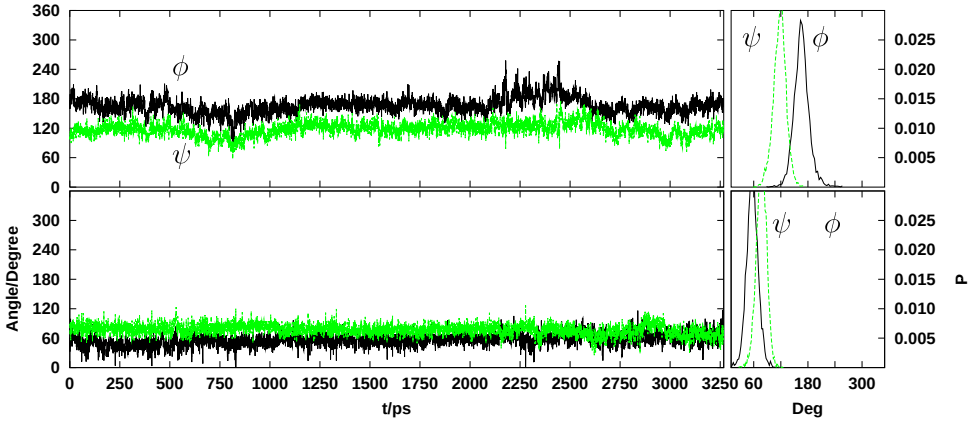


Figure 3.7: Selected trajectories of ring-linking dihedral angles for CMC IIa (ϕ : C4-O4-C1'-C2' (black) and ψ: C3-C4-O4-C1' (grey)), link 6-7 (top), link 3-4 (bottom). Distributions are shown on the right. The sharp one-state peak for the 3-4 link distribution is also characteristic for CMC II glucosidic links in the chain center such as links 4-5 or 5-6. Distribution histograms are shown on the right.

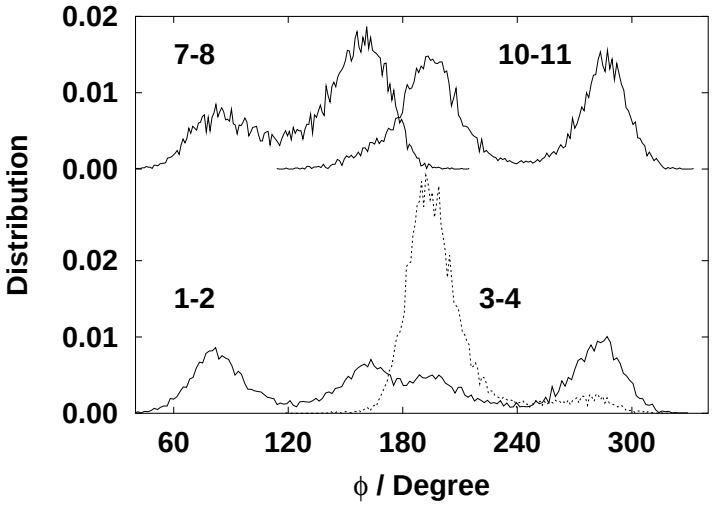


Figure 3.8: Distributions of four representative PAA backbone dihedral angles. The numbers indicate the repeat units between which the dihedral angle is located. The above two torsions have two maxima, with clear minima beween them. The lower panel shows one almost unimodal torsion and a second one with four maxima. PAA has two backbone torsions per repeat unit.

3.2.3.1 Solute-Solute Hydrogen Bonding

The average absolute numbers of intramolecular hydrogen bonds are 3.4 for CMC I and 7.4 for CMC II respectively (see figure 3.9 and table 3.2). The large number of intramolecular hydrogen bonds stabilizes the more compact structure of CMC II compared to CMC I.

For CMC, the hydrogen bond lifetime τ (see page 49) differs between the CMC I and CMC II molecule: CMC I has a bond lifetime of 23 ps, CMC IIa of 86 ps (table 3.2), indicating that the cyclic compact structure of CMC II is stabilized by long-lived hydrogen bonds. This trend is observed for CMC IIb as well, with an apparent average bond lifetime of 50 ps. However, τ is smaller if we consider the two phases of this simulation separately (figure 3.4). This indicates that some hydrogen bonds persist during the process of folding. Cutting one trajectory into two interrupts the long-living bonds, which would otherwise contribute to a tail in the H-bond correlation function (defined on page 49). Namely there are 4 bonds, persisting during the collapse, but two of them are intra-ring bonds, and two bridge only one glucosidic bond employing a carboxylic group. So the lifetime of 50 ps calculated for the trajectory as a whole is more likely to be a statistical artifact, as correlation functions are only meaningful in equilibrium. It is interesting to note that the total number of H-bonds of CMC IIb does not change during the folding process.

Intramolecular hydrogen bonds can be further subdivided into two sets: *(a)* intra-ring and *(b)* inter-ring bonds, which connect two distinct glucose moieties and which are presumed to influence the overall chain structure of CMC: The ratio of inter- and intra-ring bonds for CMC I is 2.7 : 1. Lifetimes are shorter for intra-ring bonds ($\tau = 13$ ps), which demonstrates a lower stability, compared to inter-ring bond with $\tau = 24$ ps. Further differences are found in the hydrogen bonding atom types: intra-ring bonds involve O3-

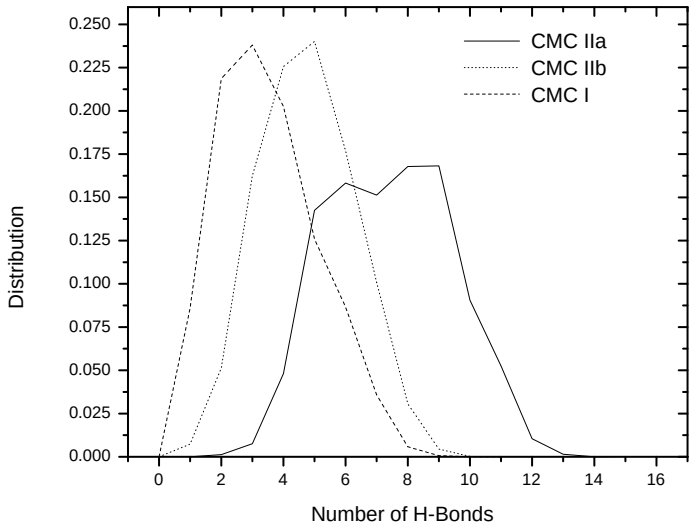


Figure 3.9: Distribution of the number of intramolecular hydrogen bonds for CMC I, CMC IIa and IIb.

		max	min	average	τ
CMC	I	9	0	3.4	23 ps
CMC	IIa	15	2	7.4	86 ps
CMC	IIb 0–0.9 ps	9	1	4.7	25 ps
CMC	IIb 1.2-2.5 ps	10	1	4.8	37 ps

Table 3.2: Number of solute-solute hydrogen bonds in one conformation: maximum (max), minimum (min), average and mean lifetime $\langle\tau\rangle$ (fit to H-bond correlation function from 0 to 80 ps).

	$n =$	0	1	2	3	4	5	6	7
CMC I		27	63	10	0	0	0	0	0
CMC IIa		17	40	14	0	5	4	17	2
CMC IIb		21	56	10	0	0	6	7	0
CMC IIb 0–0.9 ps		45	55	0	0	0	0	0	0
CMC IIb 1.2–2.5 ps		9	57	15	0	0	9	10	0

Table 3.3: Solute-solute hydrogen bonds for CMC. First row (n) contains the distance along the chain between two rings involved in a hydrogen bond, the next rows give the percentage of each type of hydrogen bond. $n = 0$ denotes intra-ring hydrogen bonds.

O6 hydrogen bonds (71%, all bonds in both directions O3H→O6 and O6H→O3) and to a smaller amount O3-O4 bonds (17%). The remaining 12% are contributed by miscellaneous bonds involving chain terminating OH-Groups or O6-O4-hydrogen bonds. Inter-ring involve mainly O6-COO^- (58%) and O2-COO^- (23%) and O3-COO^- (10% each) oxygen atoms. Hydrogen bonds from O3 to O5, which are known for crystalline cellulose [15] are not present.

Thus the picture of intramolecular hydrogen bonds in CMC I is as follows: There is a high fraction of long-lived, ring-connecting hydrogen bonds, intra-ring bonds are of less importance. Inter-ring hydrogen bonding happens mostly via COO^- groups, which makes *(a)* the interaction energy favourable and *(b)* involves a flexible, exocyclic group. For this reason, changes in the global chain conformation may be compensated by carboxy-methyl torsional changes to keep the OH-O alignment in an energetically low geometry. Intra-ring H-bonds, on the other hand, are either formed by neutral OH groups, which have a lower energy, or they necessitate sharp turns of the carboxy methyl group, which leads to distortions.

In CMC I, inter-ring hydrogen bonds nearly always connect neighboring rings (table 3.3). This is different for CMC IIa, where also more distant rings are connected by hydrogen bonds: In contrast to CMC I, 51% of the inter-ring hydrogen bonds connect rings which are not nearest neighbors: most probable are bonds which span 2 or 6 rings. These bonds are found to be very stable, some of them are intermittently present for an overall time of 0.3 ns. For comparison, CMC IIa intra-ring bonds never live longer than 0.15 ns. The same kind of persistent H-bonds is present in the folded structure of CMC IIb with a maximum lifetime of 0.2 ns.

Furthermore, in CMC IIa there are pairs of hydrogen bonds (about 10), which bind almost exclusively (see figure 3.10 for examples; the same kind of plot for CMC I is shown in figure 3.12). Out of these pairs, only one bond is present at any given time, but alternating with the second one. The presence of one bonds prevents the second bond. Even, if the mean lifetime of an individual bond is short, both bonds together form a steady connection between two glucose rings. One example is the bond pair O3(ring 6)-COO(ring 6) and O6(ring 5)-COO(ring 6). Each particular bond lives 750 ps, respectively 950 ps and together they span 1800 ps. This is evident for H-bonds formed via a COO^- group, which account for half of all solute-solute bonds. A distribution of all bond-bond cross correlation coefficients for CMC I and IIa is shown in figure 3.11. They are calculated as

$$c_{ij} = \frac{\langle (h_i(t) - \langle h_i \rangle)(h_j(t) - \langle h_j \rangle) \rangle}{(\langle h_i(t) - \langle h_i \rangle \rangle^2 \langle h_j(t) - \langle h_j \rangle \rangle^2)^{1/2}},$$

where $h_i(t)$ is 1 when hydrogen bond i is present and 0 otherwise. Most hydrogen bonds are uncorrelated: Only 15 hydrogen-bond-pairs have a strong negative ($c_{ij} < -0.3$), 79 bond-pairs a positive correlation ($c_{ij} > 0.3$).

High positve correlations ($c_{ij} > 0.6$) occur particularly if the compact configuration of both the CMC IIa and IIb oligomer allows for some bonds to form simultaneously, like H-bonds involving O2H and O6H donor groups of one, and a single carboxylic group of a second ring. The COO^- group is embedded between the two H-donors (figure 3.13). The effect is very strong for the flexible COO^- group at rings 7 and 8, but is still noticeable for other acceptors.

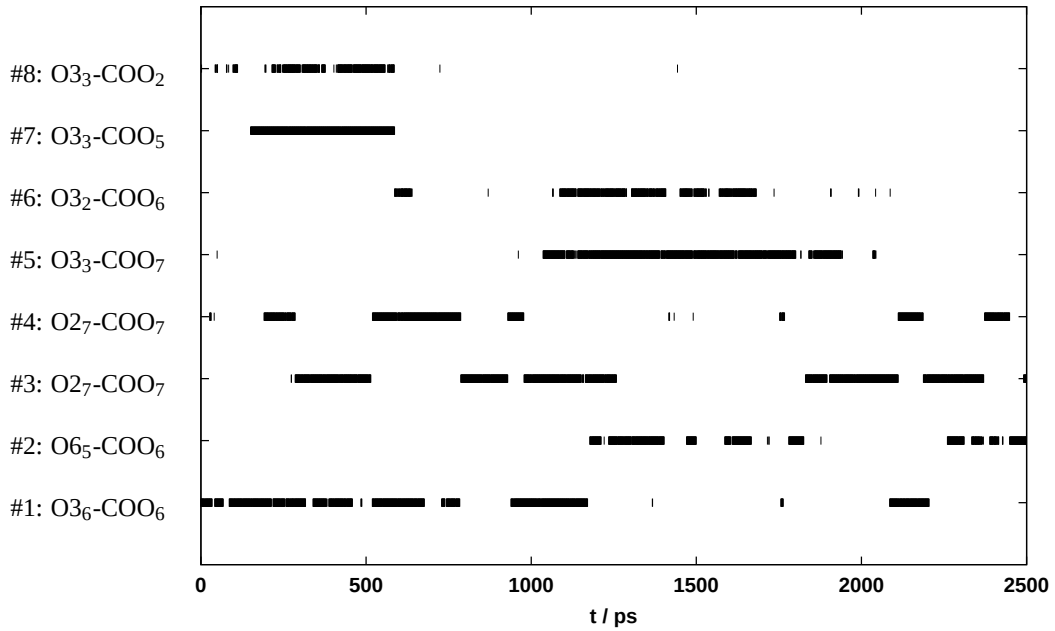


Figure 3.10: Time evolution of eight selected hydrogen bonds for the compact CMC oligomer IIa. For bond pairs (like #1/#2 or #3/#4) most of the time there is only one bond out of a pair intact (indicated by a bar). Bond #1, #3 and #4 are intra-ring H-bonds, the remaining are inter-ring bonds: (subscripts indicate the anhydro glucose unit number.)

The cross correlation coefficient c

$$c = \frac{\langle (n - \langle n \rangle)(r - \langle R_{\mathrm{ete}} \rangle) \rangle}{(\langle (n - \langle n \rangle)^2 \rangle \langle (R_{\mathrm{ete}} - \langle R_{\mathrm{ete}} \rangle)^2 \rangle)^{1/2}}. \tag{3.1}$$

shows some coincidence between the number of solute-solute hydrogen-bonds n and the polyelectrolyte's end-to-end distance R_{ete} for CMC I. Whereas there is a correlation of extension with the total number of solute-solute hydrogen-bonds ($c = 0.24$), it is slightly more pronounced ($c = 0.28$) for the number of intra-ring hydrogen bonds versus the end-to-end distance: If the molecular extension shrinks, hydrogen bonds are broken. This is not in line with the view, that a compact structure favours H-bonding. Instead, this is a hint that the straight conformation of CMC I is stabilised by hydrogen bonds between neighbouring rings. For CMC IIb we find the opposite: There is a negative correlation ($c = -0.20$) between the end-to-end distance and the total number of solute-solute hydrogen-bonds. The correlation is much stronger for inter-ring bonds: $c = -0.68$. The compact polymer structure favours more contacts.

3.2.3.2 Hydrogen Bonds to Solvent

A general picture of the solvent distribution around CMC is given by radial distribution functions (RDF, [2]) between the center of mass of the glucose unit and water oxygens (figure 3.14). The center of mass of a repeat unit is calculated using all its atoms, including

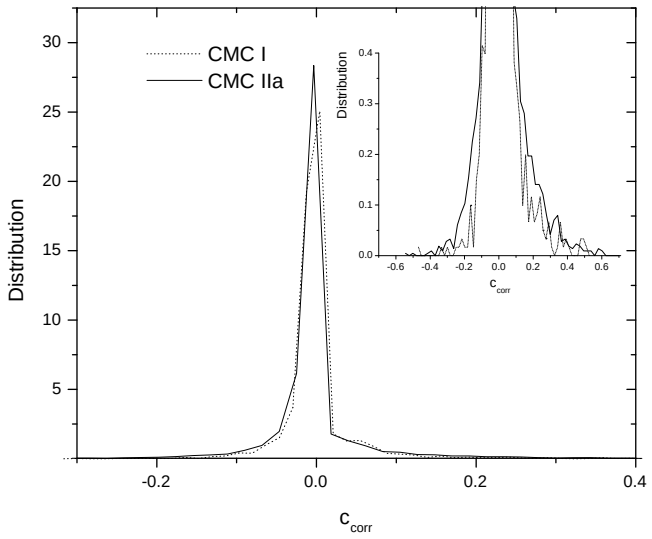


Figure 3.11: Distribution of correlation coefficients (see section 3.2.3.1) between all intramolecular hydrogen-bonds occurring during the CMC I and CMC IIa simulation. The majority of all hydrogen bonds does not show correlations with other bonds for neither simulation. Only a small part shows negative and positive correlations (inset).

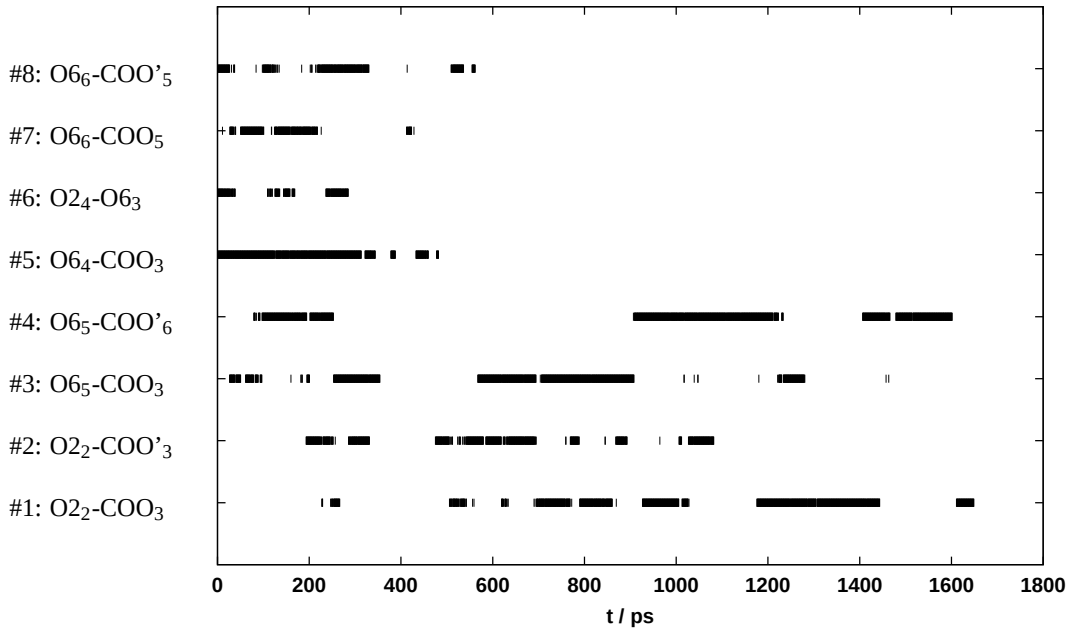


Figure 3.12: Time evolution of selected hydrogen bonds for CMC I. Hydrogen-bonds #1 and #2 are anticorrelated, there is only one bond present at any given time. Bonds #5–#8 are not found beyond 0.6 ns. Description of the H-bonds, subscripts indicate the number of the anhydro glucose unit.

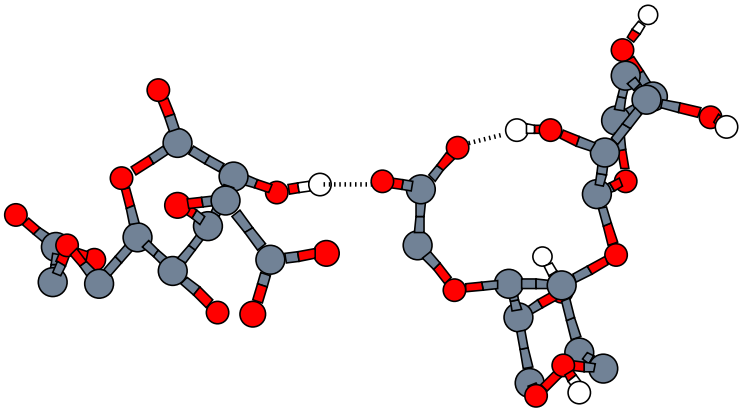


Figure 3.13: Example of a solute-solute hydrogen bond for CMC IIb. There are two inter-ring H-bonds, which occur only simultaneously.

any side groups if present. The CMC repeat unit begins at C1 and ends at O4 according to the numbering in figure 2.2. The acrylate repeat unit coincides with the propylate unit -CH_2-CH-COO^-.

The differences in the peak heights of different rings are explained as follows: The outermost glucose rings have the best defined hydration shell. Environments of inner rings are more perturbed, either by neighboring rings or by attached carboxylic groups. This is evident for the doubly substituted ring 3 in CMC I, whose first solvation peak disappears entirely. From the appearance of the RDFs in figure 3.14, the solvation shell of all three CMC oligomers are similar. The radial distribution function of the AGU-center of mass with water is not sensitive to the global chain conformation, but more to the local substitution pattern. The solvation shell of PAA (fig 3.14 d) is uniform: chain ends attract more water than inner segments, but, apart from that, the center of mass-water RDF is the same with the first peak at 0.35 nm.

On average, the CMC I oligomer has 12 donor and 29 acceptor hydrogen bonds with water, CMC IIa 9 donor and 22 acceptor bonds (table 3.4). The difference is in line with the result for intramolecular bonds: The CMC II molecule, which bonds with itself, has fewer free sites for hydrophilic interaction with the solvent. Poly(acrylic acid), with no hydrogen bond donors of its own, has on average no less than 136 bonds with water. As an estimate, this corresponds to water molecules with a total mass of 2300 amu, which is 1.5 times the polymer's own mass. The mass of associated water is smaller for CMC: About 60% of the polymer's mass for CMC I and only 50% for CMC II. The detailed distributions of H-bonds are shown in figure 3.15. They are approximately Gaussian in all cases and they reflect the trends already seen in the averages.

In order to compare the hydration of the different molecules on an equal basis, one may normalize the number of H-bonds by the mass of the polymer. This yields the H-bonds per weight, which should in some way be correlated with the specific enthalpy of

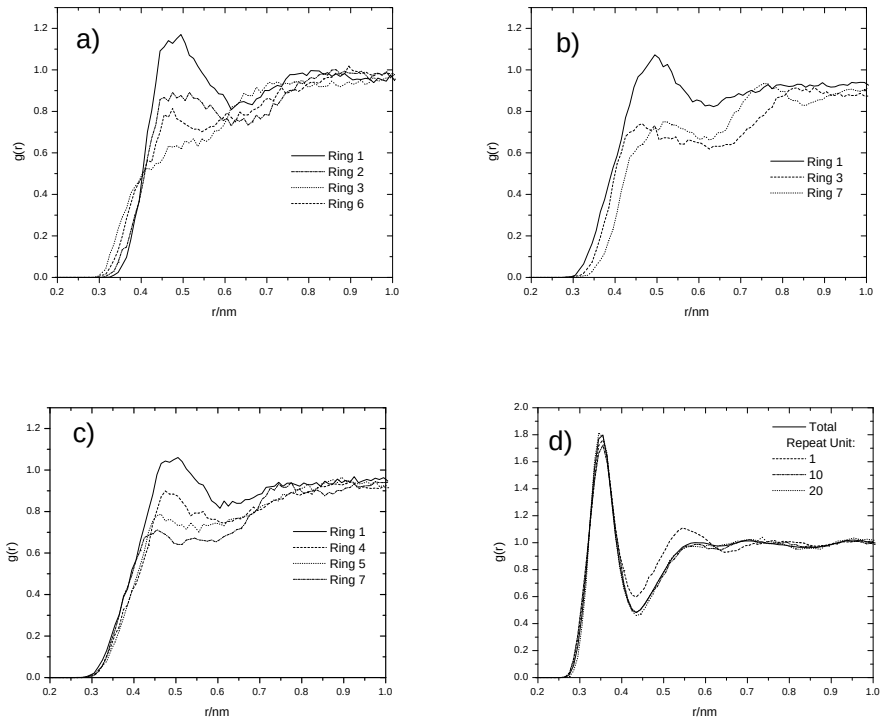


Figure 3.14: Radial distribution functions $g(r)$ of selected CMC glucose rings and PAA monomer units (center of mass) with water oxygens. a) CMC I, b) CMC IIa, c) CMC IIb (total simulation) and d) PAA. For CMC I the first hydration shell($< 0.65\,\mathrm{nm}$) of each monomer contains, on average, 15 water molecules. The first shell of CMC II monomers is more irregular, with hydration numbers ranging from 10 to 18. PAA has the most regular and smooth hydration shell ($< 0.45\,\mathrm{nm}$) with 8 neighbors per repeat unit.

	CMC I		CMC IIa		PAA	
Donor-OH	11.8	(0.66)	8.9	(0.44)		
Acceptor COO^-	17.4	(1.75)	21.6	(1.80)	136	(2.9)
Acceptor OH	10.0	(0.55)	11.6	(0.56)		
Acceptor ether-O	1.6	(0.12)	0.82	(0.05)		
Sum per AGU	5.8		5.4			

Table 3.4: Average number of hydrogen bonds to solvent, grouped according to solute binding sites. In parentheses: numbers with respect to hydroxy, ether and carboxylic oxygen sites of the polyelectrolytes.

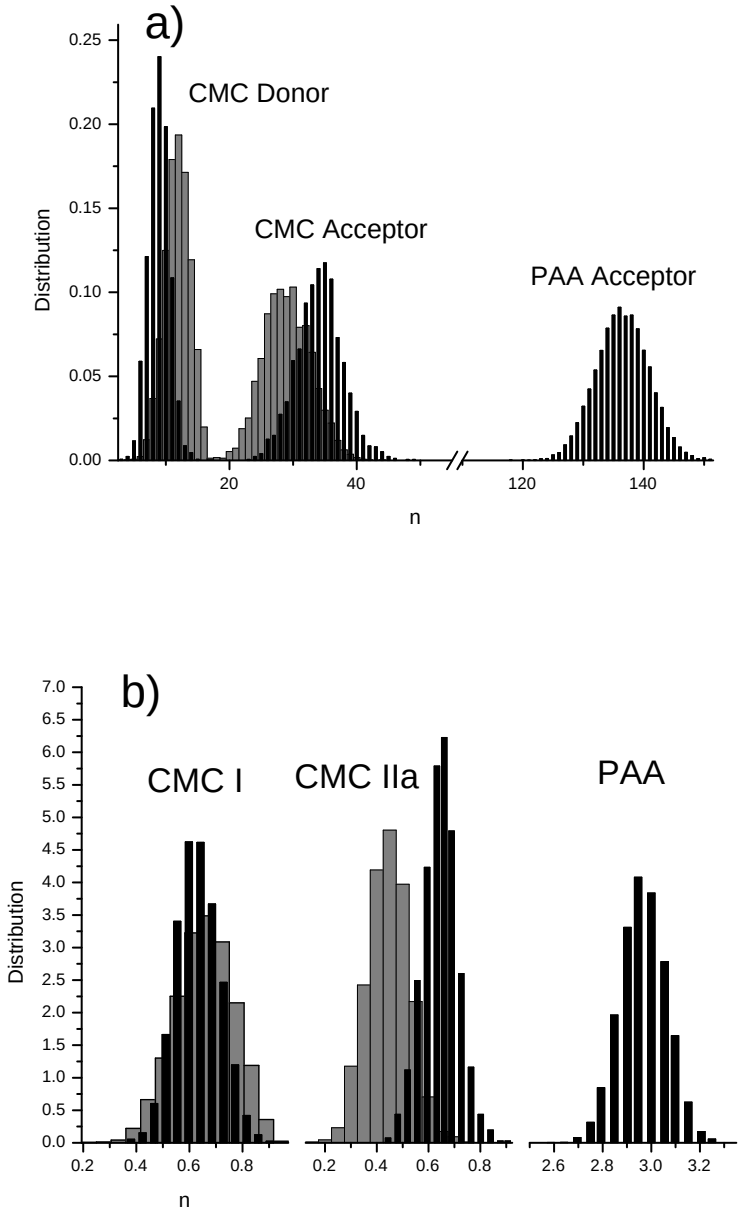


Figure 3.15: a) Total number of donor and acceptor hydrogen bonds between water and CMC I (grey) and CMC IIa (black) and PAA. b) Number of solute-solvent hydrogen bonds for CMC I, CMC IIa and PAA (from left to right) normalized with respect to the number of possible binding sites of the polyelectrolytes (counting all polymer oxygens as acceptors, all hydroxy-groups as donors to water). Acceptor bonds are grey, donor bonds are black.

solvation. One finds 0.029 bonds/amu for CMC I, 0.019 bonds/amu for CMC II and more than twice as many for PAA: 0.083 bonds/amu. Thus, there seems to be a competition between inter- and intra-molecular hydrogen bonds for (carboxy methyl) cellulose, which is dominated by the intra-molecular bonds in case of CMC IIa and to a lesser extent for CMC I. Counting intramolecular bonds (table 3.2) twice as they involve two sites, one finds a very similar total number of hydrogen bonds for CMC I (47.6) and CMC IIa (50.3): The larger number of intramolecular bonds in CMC IIa leads to a decrease of the number of bonds to solvent. In PAA, there is no competition from intramolecular hydrogen-bonds. In addition, one can reasonably expect that hydrogen-bonds to COO^- with its negative charge are energetically more favorable than to –OH or –O–. Hence, PAA should have a more negative enthalpy of solvation than CMC. We are, aware of experimental to compare with data only for PAA: Klein et al. [67] used solution calorimetric measurements to determine the solution enthalphy of about 25 kJ/(mol repeat unit) (0.27 kJ/g) for sodium-PAA, this value corresponds to 575 kJ/mol for the PAA 23-mer.

3.2.4 Counterions

The distribution of counterions around the solute molecules has also been investigated. The solutes are fragments of polyelectrolytes, and since the work of Manning [81] the distribution of counterions in the field of polyelectrolytes has been of considerable theoretical interest (see e.g., ref. [8] and references therein). Because of the small size of the charged solutes investigated here and the small number of counterions, this study is necessarily limited to the immediate neighborhood of the solutes.

Radial distribution functions (RDFs) between the centers of mass of the repeat units (for definition, see section 3.2.3.2) and the Na^+-ions are summarised in figure 3.16.

All CMCs show multiple peaks, which are very high. Even though, the number of close counterions (integral of $g(r)$ [2]) is small. For CMC I, the integrals including the first, second and third peak, are 0.14, 0.45 and 0.80, respectively. Unity is reached at a distance of 1.9 nm: The counterion concentration in the vicinity of the stretched CMC I is small and there is not much of a difference to CMC IIa. The positions of the seconds peaks of CMC IIa are different but the integral equals one at a distance of 0.64 nm. The difference with CMC I and IIa is explained by the higher charge density of the compact structure which attracts more counterions. The large second peak of the pair distribution of CMC IIa falls off directly to unity. Both pair distribution have a first peak at the same position (0.34 nm) with the same number of neighbors. Both the (dynamic) flexibility of the carboxylate and the (static) different substitution patterns on different AGUs account for the large width of the second peak. CMC II has a lower intensity at short distance. This arises from its more compact structure which prevents counterions from approaching it freely from all sides. This behaviour is similar to what was already seen for the solute-solvent H-bonds. The corresponding RDF of poly(acrylate) (fig 3.16 b) is better defined due to the larger number of counterion-monomer pairs. It shows a relatively sharp first maximum at about 0.51 nm caused by Na^+ ions near the carboxylate and a broader second maximum ($\approx$ 0.8 nm) possibly due to solvent-separated ion pairs. Up to the first minimum a PAA repeat unit has 1.4 Na^+ neighbors. On average the CMC oligomer keeps fewer counterions in its electrostatic vicinity than the PAA oligomer. This is caused by the difference in charge.

Pair distribution functions for different oxygen atoms with sodium-ions reveal that the

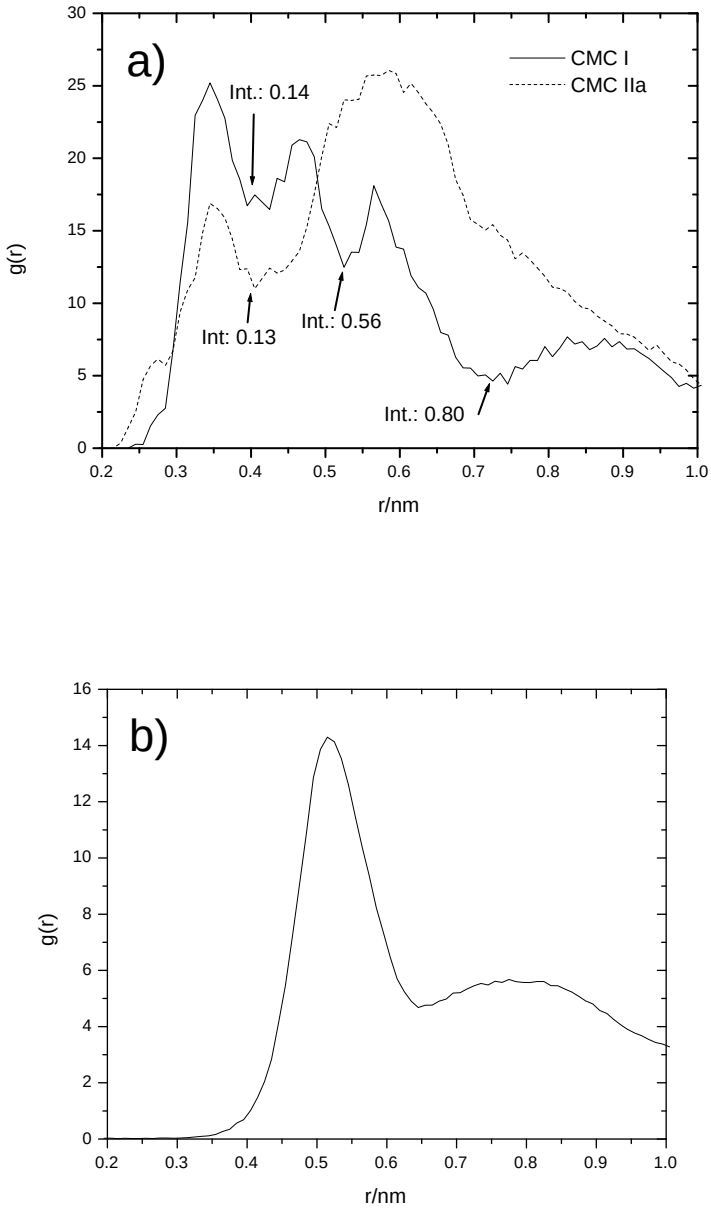


Figure 3.16: Monomer(center of mass)-sodium radial distribution functions for a) CMC I, CMC IIa and b) PAA. Both CMC radial distributions reach unity at about $r = 1.9\,\mathrm{nm}$. The PAA radial distribution owns a well defined first peak (1.4 Na^+ neighbors) and a broad second one with 4.6 counterions. The PAA-rdf falls down to $g(r) = 1$ at $r = 1.9\,\mathrm{nm}$.

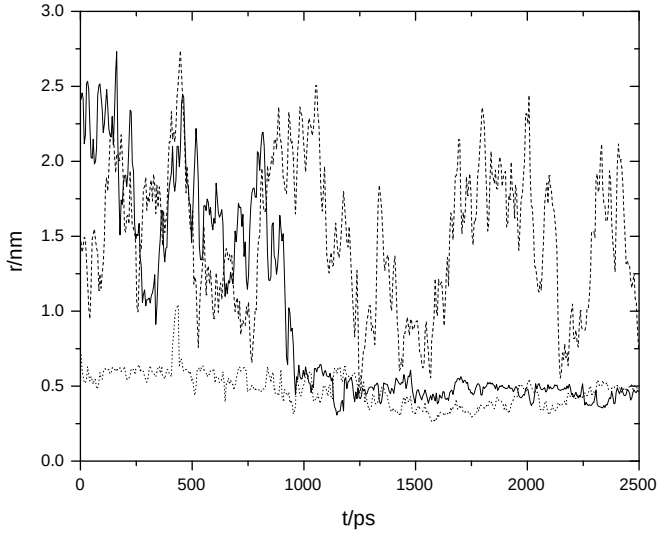


Figure 3.17: Distance between three selected sodium-ions and the CMC IIb-backbone. The shortest distance between a given Na^+ and any AGU-center of mass is plotted against the time. There are three curves: a) The counterion moves freely and is not trapped at the polymer. b) the counterion *is* trapped by the polymer (short polymer-counterion distance) c) the counterion *gets* trapped by the polymer (combination of a) and b)). These types of counterions are typical for CMC I and CMC IIa as well.

solute-counterion distribution is determined mainly by the carboxylate groups. Here our focus is only on the first peak of the radial distribution function and its integral, which corresponds to the number of direct neighbors. For all simulated polyelectrolytes most of the sodium ions near the solute bind to COO^-: For CMC I (table 3.5) there are 0.67 sodium neighbors near a carboxylate oxygen, whereas there are only 0.02 neighbors near O2, but 0.12 near O3. This is not observed for the collapsed CMC IIa, which on average has most (0.89) counterions near the COO^--group. This averaging is resolved in time for CMC IIb, where we divide again into the elongated and collapsed structure. During the first 0.9 ns of simulation, there are no counterions near O2 and O3, but with the chain folding two peaks with 0.32 and 0.36 sodium atoms arise near O2 and O3 respectively. The number of ions near CMC IIa and CMC IIb is almost the same. For CMC different types of counterion dynamics are observed (figure 3.17 for CMC IIb examples): Ions stay close to the polymer or move freely. For CMC IIb there is one sodium-ion, approaching the polymer from 2.5 nm to 0.5 nm prior to the collapse. The number of permanently (on the scale of teh simulation) trapped counterions plays a crucial role for the collapse of CMC II. There are 4 trapped ions for CMC IIa and CMC IIb after the collapse. For CMC I we find only 2 ions trapped closely to the sugar chain.

		1st Peak		2nd Peak	
		r_1/nm	n_1	r_2/nm	n_2
CMC I	O2	0.31	0.02	0.57	0.32
	O3	0.37	0.12	0.58	0.37
	O4/5	0.38	0.02		
	O6	0.34	0.04	0.56	0.32
	C(OO)	0.32	0.67	0.47	0.88
CMC IIa	O2	0.33	0.08	0.67	0.92
	O3	0.34	0.09	0.65	0.90
	O4/5	0.36	0.04	0.60	0.46
	O6	0.32	0.06	0.47	0.36
	C(OO)	0.31	0.89	0.47	1.26
CMC IIb	O2			0.34–0.61 (very broad)	0.38
0–0.9 ns	O3			0.31–0.61 (very broad)	0.34
	O4/5	0.37	0.03	0.89	1.09
	O6	0.33	0.04	0.67	0.40
	C(OO)	0.34	0.46	0.55	0.58
CMC IIb	O2	0.32	0.18	0.47	0.70
1.2–	O3	0.36	0.29	0.56	1.15
2.5 ns	O4	0.40	0.13		
	O5	0.36	0.11		
	O6	0.36	0.23	0.55	0.81
	C(OO)	0.32	1.11	0.48	1.61
PAA	C(OO)	0.36	0.93	0.74	1.88

Table 3.5: Summary of oxygen-sodium contacts, r_1, r_2 is the distance of the first or second minimum of $g(r)$, n_1, n_2 is the number of neighbours up to the minimum distance. Fields are empty, if the shell cannot be well defined. (The number given as the integral for the second peak is the cumulative value for both the first and second one: $n_2 = \int_0^{r_2} g(r)\mathrm{d}\, r^3$).

3.3 Conclusions

Even though poly(acrylate) and (carboxy methyl) cellulose both are water-soluble polyelectrolytes, their behaviour in water and towards water differs markedly. This is due to the different charge density as well as to the different type and quality of hydrogen bonds that either forms with water. In PAA, there is one strong hydrogen bond with the deprotonated carboxylate acting as an acceptor. In CMC, the smaller density of carboxylates is only partly offset by the possibility of alcoholic OH groups participating both as donors and as acceptors in hydrogen bonds. Hydrogen bonds to the ether oxygens are irrelevant. Taken per molecular weight of the polymer, it seems safe to say that PAA forms at least twice as many hydrogen bonds to water as CMC and that they are of larger binding energy (charge-dipole, rather than, dipole-dipole). Based on this argument, the solvation of PAA in water should be more exothermic than that of CMC. Unfortunately, no measurements appear to be available for comparison.

The comparison of the two CMC oligomers shows that the particular carboxy-methylation pattern has an immense influence on the local structure in solution. The two assume entirely different conformations: CMC I is stretched and flexible, whereas CMC II favors a rigid cyclic conformation. We are therefore left to conclude that industrial CMC with its statistical substitution of OH groups, behaves locally very diversely. As a consequence of its globular structure, CMC II shows more intramolecular hydrogen bonds than CMC I, fewer hydrogen bonds to water, slower hydrogen bond dynamics, and more contacts with the counterions.

4 The Monoclinic Cellulose-Water Interface

Cellulose is one of the most abundant biopolymers and a basis for many industrial derivatives and products. Its surface physics and chemistry are important for the understanding of adhesion to other materials (e.g. nanocomposite of cellulose-fiber thermoplastics) [58], adsorption processes from solution (e.g. in paper production) [111], and of industrial processes that start from a cellulose slurry (e.g. in the preparation of cellulose derivatives, where often counter-intuitive substitution preferences are found) [53]. Applications cover many areas of chemistry, biology and technology. Still, the cellulose surface is not well understood. The situation is complicated by the simultaneous presence of several phases in natural and processed cellulose fibers: There are three principal crystalline phases, triclinic and monoclinic and type II with varying contents of amorphous material. Not even the behavior of the crystalline surfaces towards water and hydrophobic organics is fully characterized on a molecular level. Therefore, we make an attempt to characterize the cellulose in its behavior towards water. (hydrophilicity/hydrophobicity) and towards apolar parts of organic molecules (lipophilicity/lipophobicity).

Experimentally, the surface of crystalline cellulose has been visualized by atomic force microscopy. One important results is that it preserves the underlying crystal structure, as it can be seen from unit cell parameters obtained by AFM which compare well with those from X-ray diffraction [5–7].

We have simulated the 110 (wide) and $1\bar{1}0$ (narrow) surfaces of monoclinic (Iβ) cellulose in contact with water. Except for the in-plane distance between parallel chains (110: 0.614 nm, $1\bar{1}0$: 0.540 nm) they are very similar, as the chains run parallel to the surface and the cellobiose units are tilted out of the surface plane by about 45°.

4.1 Radial Distribution Functions

We computed the radial distribution functions $g(r)$ (see equation 1.30 and [2, pages 55 ff]) between cellulose atoms of the interface layers one and water molecules. All atoms of the cellulose surface layer are selected regardless of whether they are exposed to water or if they point inwards toward the cellulose bulk. As the radial distribution functions are calculated only in a half-sphere over the surface instead of the full radial space around the reference atom they should tend towards one half. However the radial distribution functions in figure 4.1 converge to 0.6–0.7. This peculiarity is caused by the $g(r)$-normalization with respect to the total simulation box instead of the two correct sub-volumes of the crystal and bulk water. The value for large distances now becomes $g(\infty) = 0.75$. Plots of the radial distribution functions for cellulose oxygens and water are shown in figure 4.1. The plots 4.1 a and 4.1 c compare radial distribution funcitions for selected cellulose sites (C1, O2 and O4) on the lower (points drawn) and upper (lines drawn) surface towards water. These two plots show that both surfaces are hydrated in exactly the same way.

Oxygen 2 (O2) of the 110 surface (figure 4.1 b) has a sharp first peak at 0.27 nm

	110			$1\bar{1}0$		
	$r_{\min}$ [nm]	$n(r_{\min})$	$n(0.29\,\text{nm})$	$r_{\min}$ [nm]	$n(r_{\min})$	$n(0.29\,\text{nm})$
C1	0.54	3.49	(0.00)	0.41	0.63	(0.00)
C2	0.48	2.57	(0.00)	0.52	2.77	(0.00)
C3	0.57	4.41	(0.00)	0.48	2.21	(0.00)
C4	0.61	6.02	(0.00)	0.61	5.17	(0.00)
C5	0.55	3.90	(0.00)	0.56	3.70	(0.00)
C6	0.44	1.76	(0.00)	0.49	3.00	(0.00)
O2	0.33	0.67	(0.39)	0.36	1.10	(0.49)
O3	0.38	1.08	(0.32)	0.40	1.24	(0.22)
O4	0.34	0.30	(0.07)	0.34	0.16	(0.03)
O5	0.66	8.25	(0.01)	0.38	0.51	(0.07)
O6	0.34	0.98	(0.52)	0.34	0.94	(0.43)

Table 4.1: First minima ($r_{\min}$) and integrals (see equation 1.31) of oxygen-oxygen and carbon-oxygen radial distribution functions (partly shown in figure 4.1) between the interface (oxygens O2, O3, O4, O5, O6 and carbons C1, C2, C4, C5 and C6) and water solvent. For better orientation the integrals at 0.29 nm are given as well in parentheses "()". This is the oxygen-oxygen distance threshold in a hydrogen-bond. The integrals at $r = 0.29$ nm reflect the hydrogen-bond strength of the different oxygens, which is discussed later in table 4.9.

followed by a minimum at 0.33 nm and with an integral of 0.67 atoms (see equation 1.30 and 1.31). The number of neighbors is higher for both other atom types: Oxygen 3 (O3) has a broad first peak with a minimum at 0.38 nm and an integral of 1.08, while O6's radial distribution function has a narrow peak (minimum at 0.34) with 0.98 atoms below.

The situation is different for the $1\bar{1}0$ surface: The radial distribution function's peaks are shifted to larger distances with more water atoms in the vicinity of the surface. O2 and O6 have almost identical first peaks of the radial distribution functions with water (fig 4.1 c and d). However, the positions of the first minima are different: O2 has a minimum at 0.36 nm with 1.10 water neighbors and O6 has a minimum at 0.34 nm with 0.94 atoms under the first peak. Oxygen 3 has only a low but broad peak with the minimum at 0.40 nm (1.24 neighbors).

O6 has the most exposed position of all oxygens (figure 4.1 b and d). As it is detached from the surface, O6 has almost identical rdfs with water on both surfaces. The differences are more pronounced for the O2- and O3-water distributions. The account in equal parts for the difference, which about one quarter water atoms difference in the integrals for the wide and narrow surface. This corresponds to the higher surface density of hydrophilic oxygen atoms the $1\bar{1}0$ surface. Even if the shapes of the radial distributions here are not absolutely identical with the distributions given by Heiner [49] for 298 K, the tendency is equal, where the narrow cellulose surface has more water molecules up to the first peak of the radial distribution functions.

As the integrals up to the threshold distance for H-bonds (see section 3.2.3) are almost equal, both the narrow $1\bar{1}0$ and the wide 110 surface are similar on a very local scale.

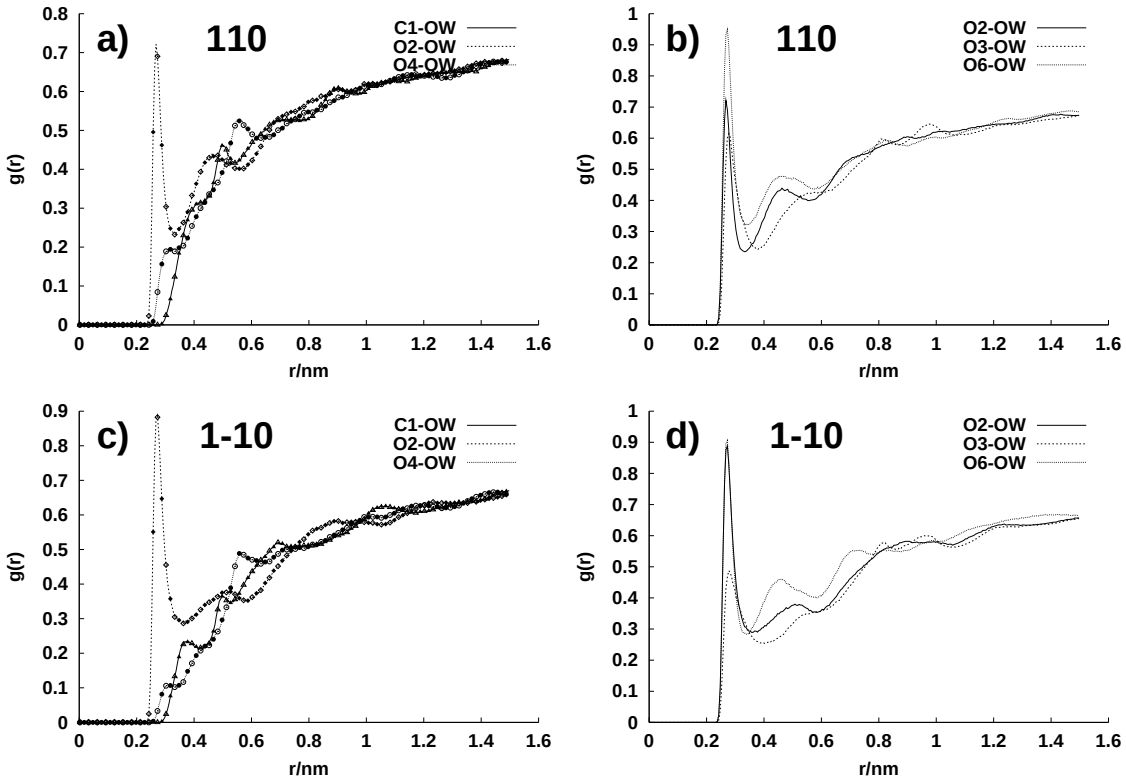


Figure 4.1: Radial distribution functions for oxygen O2, O3 and O6 of the cellulose surface with water oxygens. a) and b) are for the monoclinic 110; c) and d) for the $1\bar{1}0$ surface.

a) *Example* radial distribution functions between cellulose atoms (C1, O2, O4) on both cellulose interfaces (up- and downside, layer 1 and layer 6 of the cellulose crystal) to water. Lines are curves for layer one, points (only every fifth shown) for layer six. There is no difference in the rdfs between the two layers. The curves do not reach unity ($g(r) = 1$ as demanded for large distances) because of the reasons explained in the text. b) Radial distributions for exocyclic oxygens O2, O3 and O6. All three have the first maximum at roughly 0.273 nm, but differ in peak height, position of the first minimum and integrals up to this point. The respective radial distribution functions for the $1\bar{1}0$-surface are shown in c) and d).

All radial distribution functions (O2, O3, O4, O5, O6) and C1, C2, C3, C4, C5 and C6 are given in the appendix on page 145.

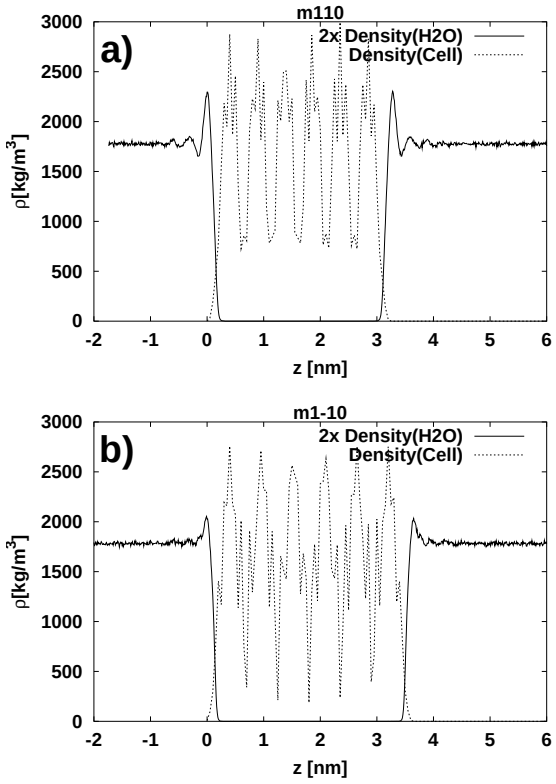


Figure 4.2: Density of the cellulose crystal and water along the surface normal. The water density is multiplied by a factor of two. a) 110 system b) $1\bar{1}0$ system. For the crystal (dashed) both sub-figures show six layers: every main peak is flanked by two side-maxima. The water density (solid) has one first maximum about 0.18 nm away from the interface. We define the position z_0 of the interface between cellulose and water as the point where densities are equal.

4.2 Local Densities perpendicular to the Surface

The crossing-point of the water and cellulose densities (see figure 4.2) is used to define the position of the crystal-water interface (z_0). For 110 the water and cellulose density are equal at $z_0 = 0.14\,\text{nm}$ and 3.13 nm; for $1\bar{1}0$ the cross at $z_0 = 0.13\,\text{nm}$ and 3.5 nm. The maximum heights of both cellulose-surface density profiles are equal, not so the water density, where the peak for the wide 110 surface is better defined and has more structure up to 1 nm away from the interface (see also section 4.5.2.1.

The distribution of atom densities in the direction of the surface normal is shown in figure 4.3. The plots are averaged over the x and y direction, which "span" the surface. There are single plots for oxygen and carbon atoms and for the wide (110) and narrow ($1\bar{1}0$) surface. For clarity, a sketch of the simulation boxes including a z-axis is given in figure 4.4 (the axes are chosen so that the surface normal of both surfaces is parallel to the z-axis).

Even if the peak intensities are lower towards the interface, the distance between the layers (for example between the C4 peaks in figure 4.3 b)) is constant. This shows, that

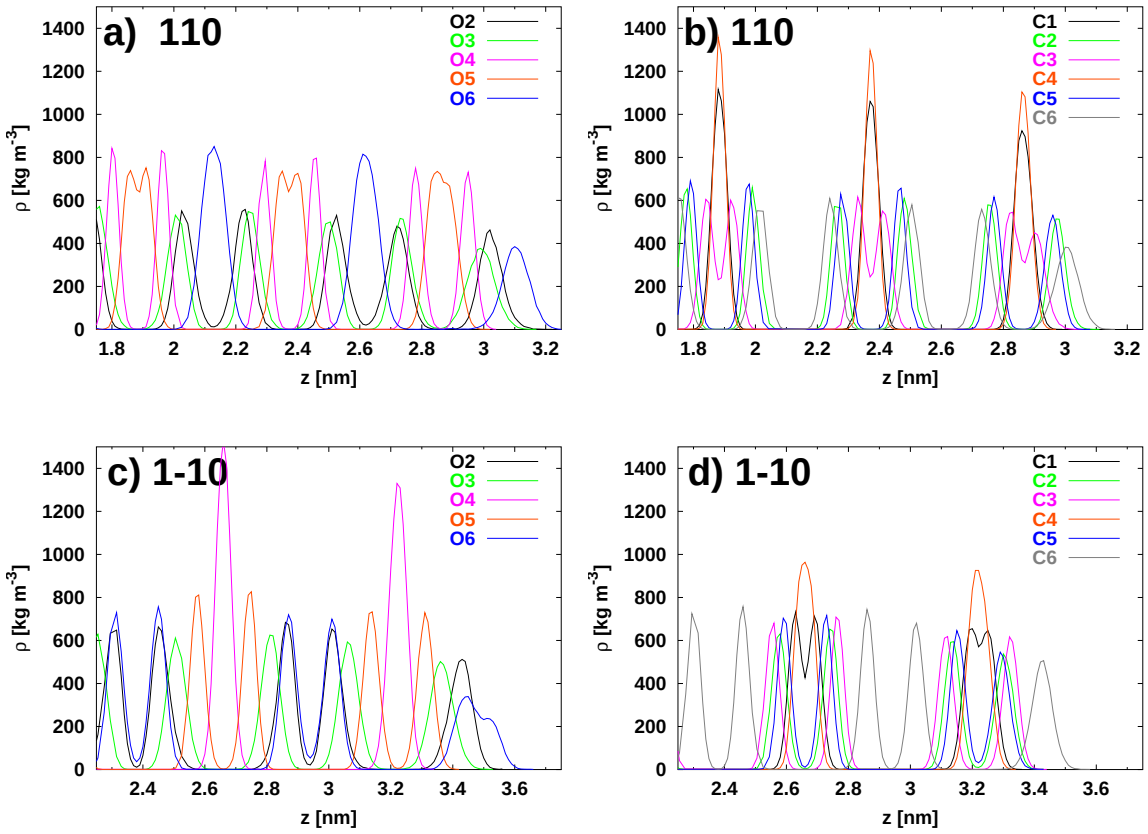


Figure 4.3: Density of different atom-types for 110 and $1\bar{1}0$-simulations along the surface normal vector (z-axis). a) 110, oxygen and b) 110, carbon. Subfigure c) and d) show the carbon- and oxygen distribution for $1\bar{1}0$.

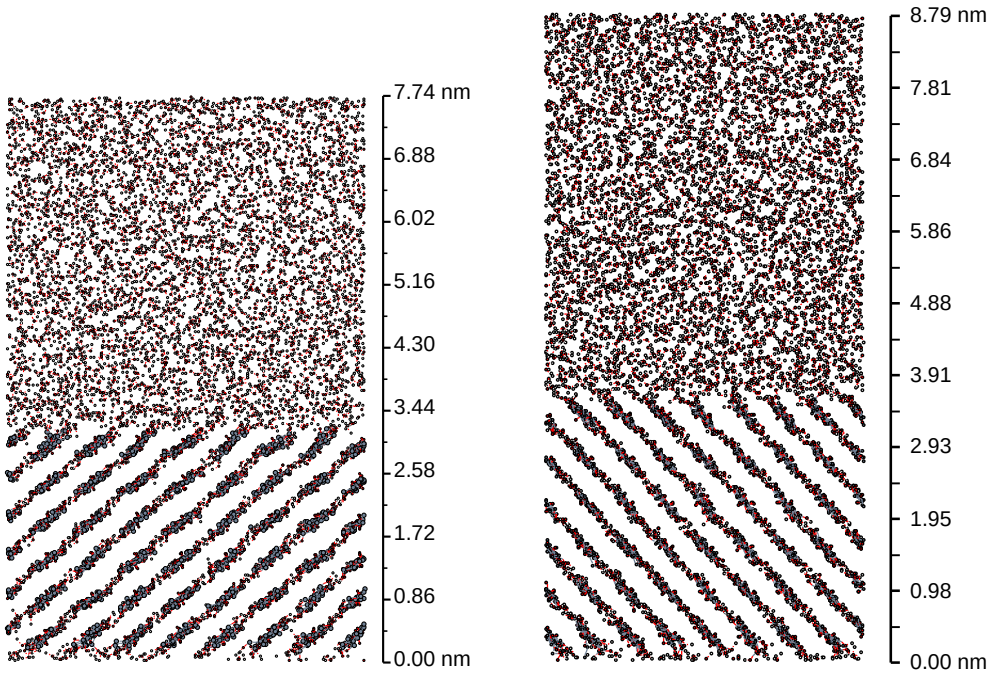


Figure 4.4: Sketch of the simulation box including the z-axis which runs of the left side of every sketch with a nanometers-scale. Left: wide 110 surface. Right: narrow $1\bar{1}0$ surface.

the cellulose crystal is stable and does not dissolve. This finding is in accordance with the simulations of Heiner and experiment, as cellulose does not dissolve in water.

Furthermore, the density profiles explain some of the radial distribution functions. The surface or interface between the cellulose crystal and water is defined as the z-coordinate where the densities meet: For the wide surface, the maximum of the O6 density is shifted by 0.08 nm away from the surface and 0.12 nm with respect to O3. So O6 is well exposed to the solvent, while O2 and O3 are both closer to the crystal bulk. For example, O3 is only 0.05 nm in z-direction away from O4.

The peak structure is different for the narrow surface: If in contact with water, O6 owns a double shoulder structure and O2 is located in the middle between O2 and O6, pointing more into the solvent. If we take the position of the outmost O3 to be zero, O2 is shifted by 0.20 nm and O6 by 0.25 nm (main peak) or 0.47 nm (shoulder). In the cellulose bulk there is only one main peak (O6), which coincides with the O2-peak. In short the different activities of the OH-groups [105] as they are found in experiment and the radial distribution of cellulose OH-groups with water are reflected and explained by the OH-groups distribution along the surface normale.

There are only a few more atoms protruding into the water for the $1\bar{1}0$ system (figure 4.3). For both systems only the C6 and hydroxy oxygens are found beyond the interface at z_0.

If the cellulose's mass (sum of all sites) is integrated from z_0 into the solvent, the 110-surface has a mass of $13.25 \cdot 10^{-9}\,\mathrm{kg\,m^{-2}}$ and $1\bar{1}0$ of $14.37 \cdot 10^{-9}\,\mathrm{kg\,m^{-2}}$, hence 8% higher mass pointing into the solvent. This is mainly due to O6, which is more exposed on the narrow $1\bar{1}0$-surface. Details are given and discussed in table 4.2.

Both surfaces seem to be quite "smooth", with some "roughness" introduced by the exocyclic CH_2OH-groups. This is in accordance with the atomic force microscopic images

Atom	110 $z > 3.125\,\mathrm{nm}$ $[10^{-9}\,\mathrm{kg\,m^{-2}}]$	$1\bar{1}0$ $z > 3.5\,\mathrm{nm}$ $[10^{-9}\,\mathrm{kg\,m^{-2}}]$
C6	0.06	0.41
O2	0.41	0.48
O3	0.07	0.03
O6	12.71	13.45
sum	*13.25*	*14.37*

Table 4.2: Cellulose single atom densities. Only atoms protruding into the solvent (z-value larger z_0) are summed. Oxygen 6 is the most protruding atom type. This is in close agreement with the AFM images of Baker [5–7], where a "flat" cellulose surface with CH_2O6 "bumps" was found.

by Baker [6] (see AFM-image in figure 1.2), where every second C6-O6 is found pointing away from the interface into the solvent.

4.3 Torsions of Side-Groups

Dihedral angle distributions for exocyclic atom-groups were calculated for different layers of the cellulose crystal. As expected, the distributions are almost identical and monomodal for the cellulose bulk in the 110 and $1\bar{1}0$ simulation. Differences were found only for solvent-exposed layers.

In the bulk, exocyclic C6O6 groups (C4-C5-C6-O6, τ) are directed backwards in chain direction and point towards O2 of the previous AGU (figure 4.5). This conformation allows for a short oxygen-oxygen distance and therefore enables hydrogen bonds (see next section, the definition is given in section 3.2.3). The conformation is compatible with a C1-C2-O2-H2 angle χ of 50° (figure 4.7), which implies a O2H2-O6 H-bond. From the appearance of solvent-exposed torsions, this hydrogen bond is no longer dominant at the surface, even if there is some intensity left. For both surfaces there is one additional peak at −175° and for the $1\bar{1}0$ surface as well at 60°.

In the cellulose bulk, the C5-C6-O6-H6 η equilibrium angle directs O6H6 to adjacent cellulose chains. Therefore the maximum at −150° is no longer present on the surface. The hydrogen at O6 has a comparative flat distribution on the surface, with only minor differences between 110 and $1\bar{1}0$ surfaces.

Only C3-O3H has similar torsion statistics for both the bulk and the surface (figure 4.8): The C2-C3-O3-H3 torsion's γ main peak is centered around 175°, its intensity is lower for solvent-exposed groups with some distributed probability in the range from −150° to 0°. This bond is found (page 75) to exist on the surface and in the cellulose bulk. There is no difference between 110 and $1\bar{1}0$ surfaces.

From the dihedral distributions it becomes obvious, that the cellulose bulk is not affected by the presence of the water solvent. Torsions change conditionally direct on the surface, but atom groups of the surface layers have bulk-like dihedral distributions, if they point towards the cellulose crystal.

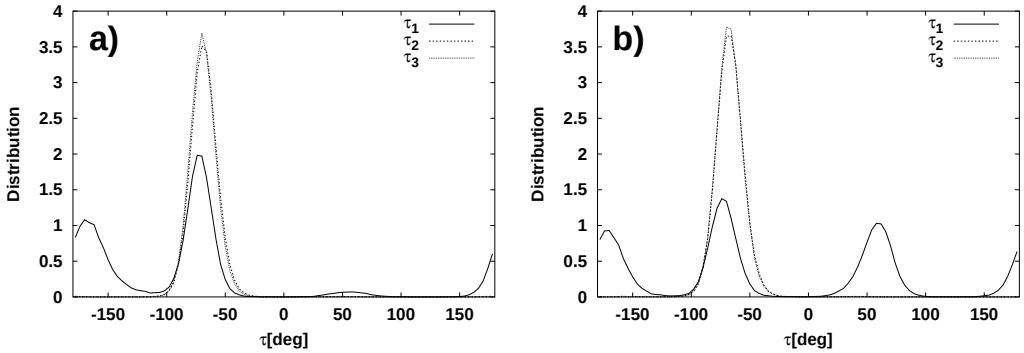


Figure 4.5: Cellulose dihedral angle distribution of C4-C5-C6-O6 (τ). Distributions are given for C6O6H groups of the top and bottom layer, pointing towards the solvent (τ_1, solid), and of the same layers, but pointing to the crystal bulk (τ_2, dotted). The Bulk values τ_3 are dashed. a) 110 surface b) $1\bar{1}0$ surface. The bulk values of $\tau_1 \approx -90°$ are consistent with an intramolecular hydrogen bond between O6 and O2 of the previous AGU in the same chain. However the situation is somewhat different on the surface, where $1\bar{1}0$ has one state $\tau_1 \approx 60°$, which is hardly populated for 110.

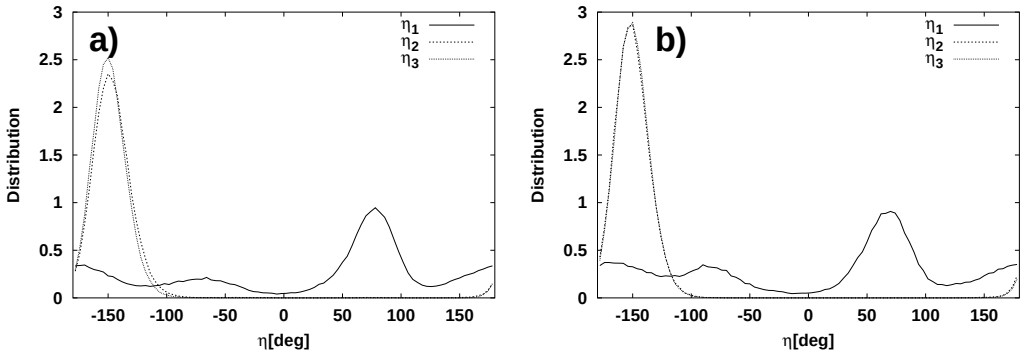


Figure 4.6: Cellulose dihedral angle distribution of C5-C6-O6-H6 (η). Distributions are given for torsions of the top and bottom layer, pointing towards the solvent (η_1, solid), and of the same layers, but pointing to the crystal bulk (η_2, dotted). The Bulk values η_3 are dashed. a) 110 b) $1\bar{1}0$ surface.

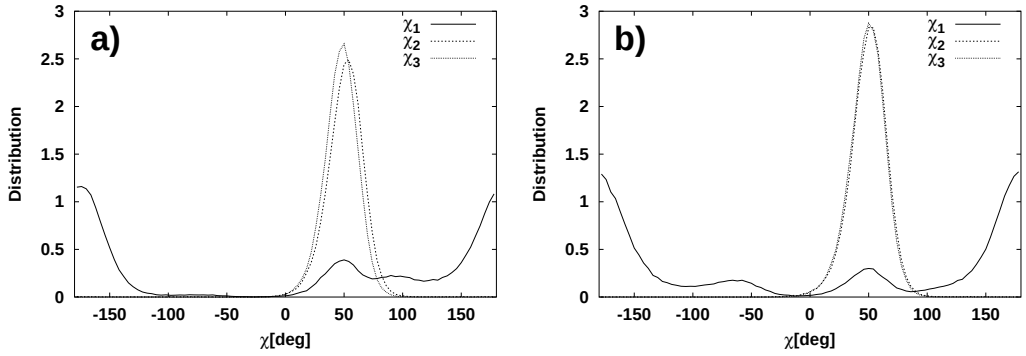


Figure 4.7: Cellulose dihedral angle distribution of C1-C2-O2-H2 (χ). Distributions are given for torsions of the top and bottom layer, pointing towards the solvent (χ_1, solid), and of the same layers, but pointing to the crystal bulk (χ_2, dotted). Bulk values χ_3 are dashed. a) 110 b) $1\bar{1}0$ surface.

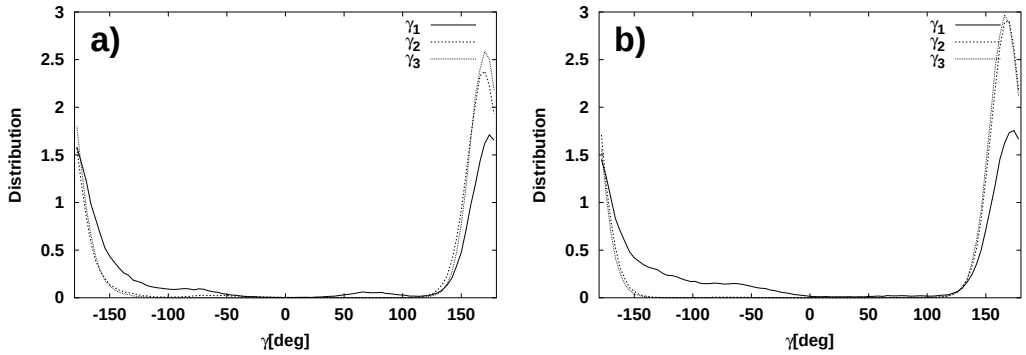


Figure 4.8: Cellulose dihedral angle distribution of C2-C3-O3-H3 (γ). Distributions are given for torsions of the top and bottom layer, pointing towards the solvent (γ_1, solid), and of the same layers, but pointing to the crystal bulk (γ_2, dotted). The Bulk values γ_3 are dashed. a) 110 b) $1\bar{1}0$ surface.

110				
Torsion	div. angle			f_{trans} [ps^{-1}]
C5-C6-O6-H6	−125	0	125	0.218
C4-C5-C6-O6	−120	0	125	0.054
C1-C2-O2-H2	−100	0	125	0.145
C2-C3-O3-H3	−100	0	125	0.100
$1\bar{1}0$				
C5-C6-O6-H6	−125	0	125	0.260
C4-C5-C6-O6	−110	0	120	0.018
C1-C2-O2-H2	−100	0	100	0.124
C2-C3-O3-H3	−100	0	100	0.146

Table 4.3: Transition frequencies f_{trans} of exocyclic dihedral angles (normalised to transitions per dihedral per time). First column states the torsion, following columns are the dividing angles (minima) between dihedral states. Last column (f_{trans}) is the transition frequency. Only transitions in the cellulose surface layers are viewed. See text (page 74) for details.

The dihedral states are classified by their minima, transition frequencies are calculated. The frequencies are determined by classifying dihedral angles into three intervals and counting the number of transitions. Results are given in table 4.3. On the wide surface 110 there is a factor of 2–4 in the frequencies between the C4-C5-C6-O6 torsion and torsions with a terminal hydrogen. The heavier O6 rotates slower than the light hydrogen atoms. On the $1\bar{1}0$ surface this rotation is slowed by a factor of 3, whereas the hydrogen torsions are unaffected. This is explained, as the 110 surface is the wider one with less steric hindrance for the free rotation.

4.4 Hydrogen Bonding

Crystalline cellulose is composed of distinct poly-glucose chains forming layers, different layers forming the crystal. The definition of a layer depends on the surface examined. We take a layer as one sheet of cellulose-chains which is parallel to the interface. By this we divide hydrogen-bonding into intra-chain, inter-chain, intra-layer, inter-layer and into crystal-water interactions. A focus is on the inter-molecular (inter-layer, inter-chain and crystal-solvent) hydrogen bonding, as the interplay determines, if the cellulose crystal is stable or is solvated in water. The cellulose layers are numbered from 1 to 6 (L1 to L6), where the water-exposed cellulose atoms belong to layer L1 and L6. Other layers (L2 to L5) form the crystal's bulk.

4.4.1 Intra-Layer Hydrogen Bonding

As the 110 and $1\bar{1}0$ simulation have both the same underlying crystal structure, intra-layer hydrogen bonds are similar, except for water exposed surface atoms. In general

		O2	O3	O4	O5	O6	sum
L1	O2H	0.00	0.15	0.00	0.00	25.30	
	O3H	1.63	0.01	0.01	30.43	0.12	
	O6H	7.80	0.04	0.29	0.03	0.00	65.8
L2	O2H	0.00	0.01	0.00	0.00	44.30	
	O3H	0.50	0.00	0.01	38.24	0.00	
	O6H	0.24	0.00	0.01	0.00	0.00	83.3
L3	O2H	0.00	0.00	0.00	0.00	44.66	
	O3H	0.44	0.00	0.01	38.45	0.00	
	O6H	0.00	0.00	0.00	0.00	0.00	83.3
L4	O2H	0.00	0.00	0.00	0.00	44.71	
	O3H	0.46	0.00	0.01	38.40	0.00	
	O6H	0.00	0.00	0.00	0.00	0.00	83.6
L5	O2H	0.00	0.01	0.00	0.00	44.49	
	O3H	0.45	0.00	0.01	38.34	0.00	
	O6H	0.08	0.00	0.00	0.00	0.00	83.4
L6	O2H	0.00	0.10	0.00	0.00	25.39	
	O3H	1.64	0.00	0.01	30.66	0.10	
	O6H	7.74	0.02	0.27	0.03	0.00	66.0

Table 4.4: Total number of intra-layer hydrogen bonds for the 110 simulation. There is one entry for each layer, with fields for every possible hydrogen bond.

their number is reduced by 20% on the surface. Every bulk layer (L2 to L5) contains 48 anhydroglucose units and has about 83 hydrogen bonds within the layer. There are on average 0.58 bonds per donor group, but only O3H-O5 and O2H-O6 bonds are found in noticeable amounts, with 0.63 bonds per O3H group and 0.92 bonds per O2H group. Surface layers (L1 and L6) of the wide 110 surface have 66 intra-layer bonds but only 63 bonds in the narrow $1\bar{1}0$ surface layers. The detailed hydrogen-bond statistics in tables 4.3 and 4.4 shows that both surfaces have a common hydrogen-bond pattern on the surface. There are 0.50 bonds per O2H donor and 0.60 bonds per O3H donor. In comparison with the bulk, the O6H-O2 bond is enforced, but still plays a minor role with 0.16 bonds per group on the wide 110 surface and 0.12 for the narrow one. This is a first argument for the water-stability of cellulose: There is still a significant amount of hydrogen bonds a) in the bulk, fixing surface layers and b) even within the surface layer itself. Solvent molecules are not able to cleave cellulose-cellulose hydrogen bonds.

There is a systematical deviation between the 110 and $1\bar{1}0$ system of about one intra-layer bond, which is due to a higher populated O3H-O5 bond in the bulk with 38.4 and 40.0 bonds for the 110 and $1\bar{1}0$ system respectively.

The average (geometric) conformations of O2H-O6 and O3H-O6 intra-layer hydrogen bonds are shown in figure 4.9 for both surfaces. Figures 4.9 a) and c) show the distribution of oxygen-oxygen distances and b) and d) show the distribution of the O-H-O angle ϕ. The maxima of the distributions are $r_{\max} = 0.263\,\mathrm{nm}$, $\phi_{\max} = 163°$ for O2H-O6 and $r_{\max} = 0.275\,\mathrm{nm}$, and $\phi_{\max} = 165°$ for O3H-O5 bonds. However, even if the maximum values for $\phi_{\max}$ are similar between both bonds, O3H-O5 has a broader distribution, with more

		O2	O3	O4	O5	O6	sum
L1	O2H	0.00	0.03	0.36	0.00	24.15	
	O3H	0.05	0.00	0.02	30.46	0.05	
	O6H	5.63	0.05	1.54	0.06	0.01	62.4
L2	O2H	0.00	0.00	0.01	0.00	44.41	
	O3H	0.00	0.00	0.02	40.09	0.00	
	O6H	0.01	0.00	0.03	0.01	0.00	84.6
L3	O2H	0.00	0.00	0.00	0.00	44.42	
	O3H	0.00	0.00	0.01	40.35	0.00	
	O6H	0.00	0.00	0.01	0.00	0.00	84.8
L4	O2H	0.00	0.00	0.00	0.00	44.43	
	O3H	0.00	0.00	0.02	40.38	0.00	
	O6H	0.00	0.00	0.01	0.00	0.00	84.8
L5	O2H	0.00	0.00	0.01	0.00	44.38	
	O3H	0.00	0.00	0.02	40.06	0.00	
	O6H	0.02	0.00	0.03	0.01	0.00	84.5
L6	O2H	0.00	0.03	0.34	0.00	24.94	
	O3H	0.05	0.01	0.02	30.56	0.07	
	O6H	5.41	0.04	1.26	0.08	0.01	62.8

Table 4.5: Total number of intra-layer hydrogen bonds for the $1\bar{1}0$ simulation. There is one entry for each layer, with fields for every possible hydrogen bond. L1 to L6 are the names of the cellulose sheets/layers of the crystal. L1 and L6 are the upper and lower layer, which are in contact with water.

intensity at lower and unfavorable bond angles. Obviously the oxygen-oxygen distance is determined by the rigid cellulose-backbone, with the torsion of the hydrogen atoms as the only degree of freedom for the O3H-O5 hydrogen bond. This is in accordance with the intra-chain nature of the O3H-O5 bond and explains the unexpected O3H-O5 bond with an ether-oxygen as acceptor.

The O6H-group features the $C6CH_2$-spacer group, which allows for a larger "action-radius" and therefore better defined hydrogen bonds are shown in figure 4.9 a) and b). So from an energetic and geometric point of view the O2H-O6 hydrogen bond is more stable than the O3H-O5. There is no difference found between the two surfaces.

4.4.2 Inter-Layer Hydrogen Bonding

Hydrogen bonding between different layers (inter-layer H-bonding) is similar for both surfaces and changes only little at the interface. This is again an evidence that the properties of the bulk crystal are hardly affected if exposed to water (see table 4.6 a) and b)). The number of hydrogen bonds between two bulk layers does not depend on the choice of the donor and acceptor layer, for example there are 22.1 hydrogen bonds from either layer L2→L3 and the 22.1 in the direction L3→L2. The layer pair L4/L5 has the same amount

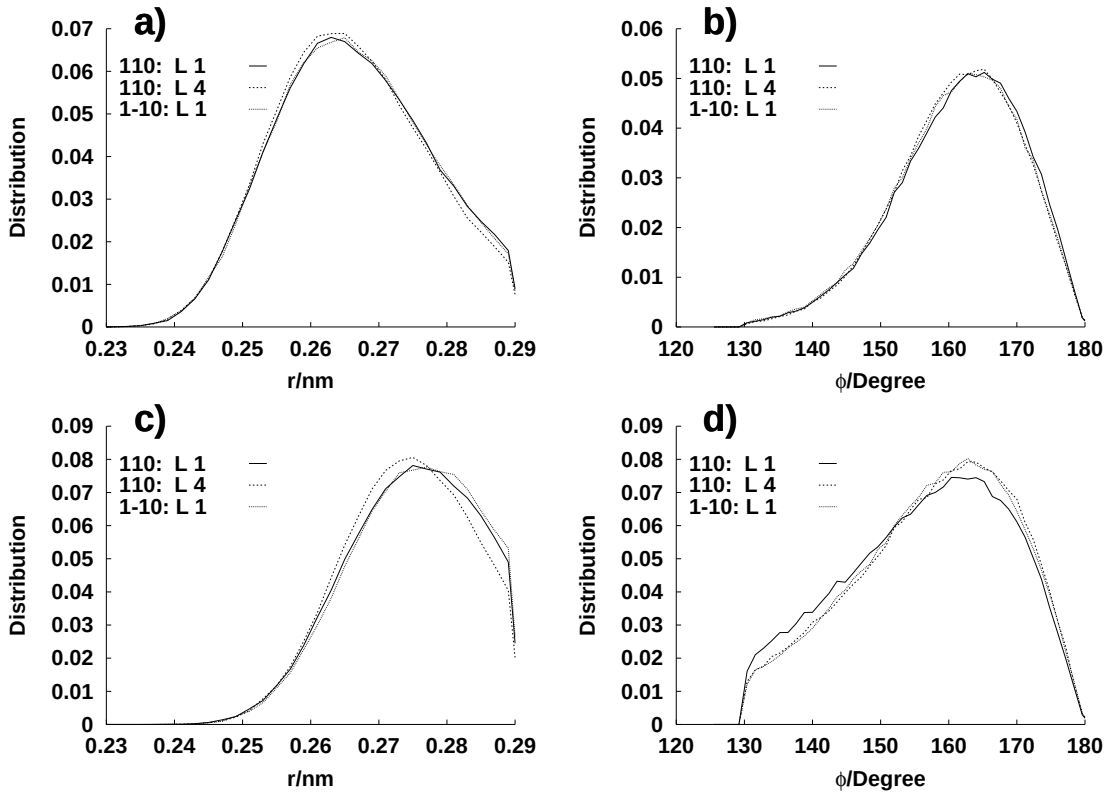


Figure 4.9: Sample intra-layer hydrogen-bond geometries for 110 and $1\bar{1}0$ surfaces for different layers (L1,L4). Distributions of oxygen-oxygen distances and O-H-O angles are shown. a) distance and b) angle distribution of O2H-O6 bonds. c) distance d) angle distribution of the O3H-O5 bond. The O2H-O6 bonds are more favorable from geometric considerations: The bonds distance is shorter and the bond angle larger, deviating less from an linear conformation. The distance distributions have an upper limit (0.29 nm) and the bond angle distributions an lower border of 130° because of our definition of a hydrogen-bond.

a)

Layer D	Layer A	mean	sum D+A
L1	L2	21.9	
L2	L1	21.4	43.3
L2	L3	22.1	
L3	L2	22.1	44.2
L3	L4	22.2	
L4	L3	22.2	44.4
L4	L5	22.1	
L5	L4	22.0	44.1
L5	L6	21.5	
L6	L5	21.8	43.3

b)

Layer D	Layer A	mean	sum D+A
L1	L2	21.9	
L2	L1	21.7	43.6
L2	L3	22.1	
L3	L2	22.1	44.2
L3	L4	22.1	
L4	L3	22.1	44.2
L4	L5	22.0	
L5	L4	22.0	44.0
L5	L6	21.6	
L6	L5	22.1	43.7

Table 4.6: Absolute number of hydrogen-bonds between different layers of a) 110 and b) $1\bar{1}0$ cellulose. Layer 1 (L1) and Layer 6 (L6) are exposed to water. The first column gives the OH donor (D) layer, second column the hydrogen-acceptor (A) one. The change at the interface is diminishing small. There is only one hydrogen-bond less than in between bulk layers. The number of bonds is identical for both surfaces. Every layer has 48 possible hydrogen-donor groups of which one half points at the upper and the other half points at the lower layer.

of inter-layer H-bonds (22.1 and 22.0). The hydrogen bond-pattern are very regular in the crystall bulk, which is in good accordance to the regular spaced density peaks of cellulose in figure 4.2. There is only the O6H-O3 bond populated. This bond is present between adjacent surface layer pairs (L1, L2 and L6, L5).

The absolute number of inter-layer hydrogen-bonds between L1-L2 and L6-L5 does change only by one bond towards the surface. There are on average 22 donor hydrogen-bonds between adjacent layers, almost all of type O6H-O3, where 90% of all O6H groups bind (with respect to 24 groups pointing to an adjacent layer).

Tables 4.6 a) and b) are further detailed in table 4.7 and 4.8 for inter-layer hydrogen bonding. Almost all inter-layer bonds go from the exocyclic O6H donor to an O3 oxygen. Other contributions are hardly found and below 4% for both surface simulations.

For the cellulose crystal both simulations have almost identical hydrogen bond pattern.

Geometric data for the inter-layer hydrogen bond O6H-O3 is given in figure 4.10 a) and b). Compared to the intramolecular bonds above, the O6H-O3 bonds exhibits on average a smaller bond angle (158°), but with approximately the same distance-distribution as the O2H-O6 bond. There are only small changes at the surface and between the two surfaces.

4.4.3 Summary Cellulose-Cellulose Hydrogen Bonds

Using all the detailed data from table 4.4, 4.5, 4.7 and 4.8 one can gains a picture of hydrogen bonding for monoclinic cellulose:

Layer D→A		Acceptor Group					
	Donor Group	O2	O3	O4	O5	O6	
L1 to L2	O2H	0.00	0.00	0.01	0.00	0.01	
	O3H	0.00	0.00	0.01	0.00	0.30	
	O6H	0.17	21.32	0.00	0.02	0.03	21.9
L2 to L1	O2H	0.00	0.00	0.01	0.00	0.09	
	O3H	0.00	0.00	0.00	0.00	0.02	
	O6H	0.20	20.92	0.01	0.14	0.00	21.4
L2 to L3	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.00	0.00	0.00	0.00	0.01	
	O6H	0.14	21.90	0.00	0.01	0.00	22.1
L3 to L2	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.00	0.00	0.00	0.00	0.00	
	O6H	0.13	21.99	0.00	0.02	0.00	22.1
L3 to L4	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.00	0.00	0.00	0.00	0.00	
	O6H	0.12	22.04	0.00	0.01	0.00	22.2
L4 to L3	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.00	0.00	0.00	0.00	0.00	
	O6H	0.12	22.06	0.00	0.01	0.00	22.2
L4 to L5	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.00	0.00	0.00	0.00	0.00	
	O6H	0.13	22.00	0.00	0.01	0.00	22.1
L5 to L4	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.00	0.00	0.00	0.00	0.00	
	O6H	0.13	21.89	0.00	0.01	0.00	22.0
L5 to L6	O2H	0.00	0.00	0.00	0.00	0.02	
	O3H	0.00	0.00	0.00	0.00	0.00	
	O6H	0.20	21.17	0.01	0.12	0.00	21.5
L6 to L5	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.00	0.00	0.01	0.00	0.15	
	O6H	0.19	21.40	0.00	0.02	0.02	21.8

Table 4.7: Inter-layer hydrogen bonds for the 110 cellulose surface. Numbers are absolute and only normalized to the number of simulated frames but not to the number of sites. Data is organized with respect to layers (L1 to L6), Donor (D) and Acceptor (A) layers and groups (donor: O2H, O3H and OH6, acceptor: O2, O3, O4, O5 and O6). Every pair of adjacent layers has two datasets with donor and acceptor role reversed. The O6H-O3 accounts for almost all of all bonds with a very small contribution from bonds originating from O2. The ether oxygens O4 and O5 do not contribute at all.

Layer D→A		Acceptor Group					
	Donor Group	O2	O3	O4	O5	O6	
L1 to L2	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.05	0.00	0.00	0.00	0.05	
	O6H	0.17	21.66	0.00	0.00	0.00	21.9
L2 to L1	O2H	0.00	0.00	0.00	0.00	0.01	
	O3H	0.07	0.00	0.00	0.00	0.00	
	O6H	0.15	21.46	0.00	0.00	0.00	21.7
L2 to L3	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.01	0.00	0.00	0.00	0.00	
	O6H	0.10	21.93	0.00	0.00	0.00	22.0
L3 to L2	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.03	0.00	0.00	0.00	0.00	
	O6H	0.10	21.94	0.00	0.00	0.00	22.1
L3 to L4	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.02	0.00	0.00	0.00	0.00	
	O6H	0.09	21.96	0.00	0.00	0.00	22.1
L4 to L3	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.02	0.00	0.00	0.00	0.00	
	O6H	0.09	21.99	0.00	0.00	0.00	22.1
L4 to L5	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.03	0.00	0.00	0.00	0.00	
	O6H	0.10	21.90	0.00	0.00	0.00	22.0
L5 to L4	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.02	0.00	0.00	0.00	0.00	
	O6H	0.10	21.89	0.00	0.00	0.00	22.0
L5 to L6	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.07	0.00	0.00	0.00	0.00	
	O6H	0.15	21.41	0.00	0.00	0.00	21.6
L6 to L5	O2H	0.00	0.00	0.00	0.00	0.00	
	O3H	0.15	0.00	0.00	0.00	0.07	
	O6H	0.18	21.67	0.00	0.00	0.00	22.1

Table 4.8: Inter-layer hydrogen bonds for $1\bar{1}0$ cellulose. Numbers are are only normalized to the number of simulated frames. No normalization to the number of atomic sites happened. Data is organized with respect to layers (L1 to L6), Donor (D) and Acceptor (A) layers and groups (donor: O2H, O3H and OH6, acceptor: O2, O3, O4, O5 and O6). Every pair of adjacent layers has two datasets with donor and acceptor role reversed. There is no qualitative difference to the 110-cellulose simulation: O6-O3 bonds account for the majority of all bonds with a vanishing contribution from bonds originating from O2. The ether oxygens O4 and O5 do not contribute at all.

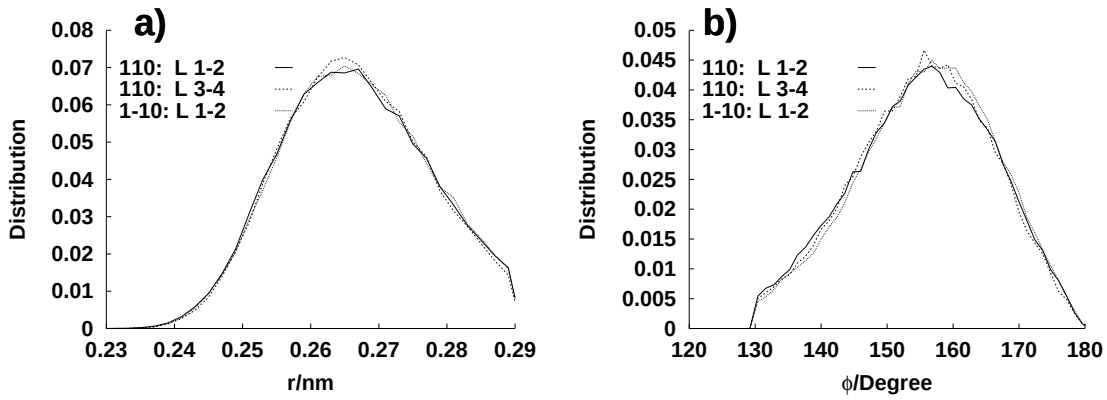


Figure 4.10: Distributions of geometric values for the inter-layer O6H-O3 hydrogen bond of cellulose for selected layer-pairs (1-2 and 3-4) a) Distribution of the oxygen-oxygen distance (r) b) oxygen-hydrogen-oxygen angle ϕ distribution. Sample distributions are shown for both surfaces and the cellulose bulk and water-interface.

O6H→O3 "diagonal" (see figure 4.4) inter-layer bond, as it shows up as inter-layer for both 110 and $1\bar{1}0$ (0.90 bonds/donor-group);

O3H→O5 intra-chain (intra-layer for both simulations), (0.81 bonds/donor-group); there is a much smaller population of diagonal O6H→O2 bonds;

O2H→O6 intra-chain for 110 and $1\bar{1}0$-cellulose (0.92 bond/donor-group);

O6H→O2 only present on the surface, intra-chain (≈ 0.25 bonds/donor-group).

For every donor group in cellulose there is one specific hydrogen bond, which uses 90% of all donor groups in the bulk. However, only one (O6H-O3) connects different layers in the crystal, the other two bonds are intra-chain respectively intra-molecular, hence they do not stabilize the crystal.

Differences in the geometry of the O6H-O3, O3H-O5 and O2H-O6 hydrogen bonds are minute. The O3H-O5 bond is less favorable in terms of oxygen-oxygen distances and bond angles compared to the O2H-O6 bond. This is not in line with the stability on the surface against water. The reduced bond stability on the surface is thus not an effect of hydrogen bond energy, but depends more on steric factors. The O6H group is stronger exposed to water and hence has more bonds with the solvent. This is in agreement with oxygen-oxygen radial distribution functions (figure 4.1) and the density distributions along the z-axis as discussed in section 4.2.

There some other hydrogen bonds, like intra-layer O6H-O4 on the surface, but they all fall back behind the first three bonds mentioned above. Together with unbound donor groups, they contribute to the missing 10% of possible bonds, which are not covered by the "major"-bonds.

4.4.4 Cellulose-Water Hydrogen Bonding

It is clear from the water distribution along the surface normal (see 4.5.2.1) that cellulose is somehow hydrophobic, and this phenomenon has to be explained on the basis of cellulose-cellulose hydrogen bonding as the dominant interaction.

110				$1\bar{1}0$			
	D	A	sum		D	A	sum
O2	0.29	0.07	0.36	O2	0.28	0.16	0.44
O3	0.04	0.23	0.27	O3	0.09	0.11	0.20
O4	-	0.04	0.04	O4	-	0.02	0.02
O5	-	0.00	0.00	O5	-	0.03	0.03
O6	0.24	0.25	0.49	O6	0.24	0.22	0.46
sum	0.57	0.59	1.16	sum	0.61	0.54	1.15

Table 4.9: Average number or hydrogen bonds per AGU between cellulose and water on the cellulose-water interface (averaged over both surface-layers L1 and L6). Column D(onor) are cellulose-water bonds, column A(cceptor) vice versa.

If the crystal gets into contact with water, it forms hydrogen bonds with the solvent and there is a competition between cellulose-cellulose and cellulose-water bonds.

Hydrogen bonds between water and cellulose are summarized in table 4.9. There are on average 1.17 bonds/AGU for the 110 and almost the same amount (1.15) on the $1\bar{1}0$ surface, adding up donor and acceptor bonds. From the tables we can deduce a hydrogen-bonds propensity:

$$\mathrm{O6} \approx \mathrm{O2} > \mathrm{O3},$$

with small variations between the two surfaces, even if they have in total the same number of H-bonds.

The hydroxy group at O6 has the same number of donor- and acceptor bonds for both surfaces. It is the most exposed oxygen, as it is not directly connected to the glucose ring, but through a methylene spacer. As we see from the density profile along the z-axis in section 4.2, O6 reaches into water furthest. The very local environment of O6 is thus not influenced by the surface structure.

The hydroxy group O3H on the $1\bar{1}0$ surface does only participate in hydrogen bonding with water as an acceptor. There are almost no donor-bonds and only few acceptor ones. Oxygens O2 and O6 have in total twice as many hydrogen-bonds. This is in line with the stable O3H-O5 intra-chain hydrogen bond (see section 4.4.1), which is obviously favored over bonds with water. As O3 is exposed to water it is a moderate acceptor for water-hydrogen.

The two ether oxygen types O4 and O5 are hardly involved in water hydrogen bonds: about 2% act as acceptors. This can be attributed to a) their sheltered position *in* the framework of the backbone and b) their small negative charge compared to hydroxy oxygens, which makes hydrogen bonding less attractive c) in the cue of O5, the competing O3H→O5 bond.

The cellulose crystal is stabilized by cellulose-cellulose hydrogen bonds, which are not cleaved by water. This is in accordance with experiment (cellulose is not water solvable) and the results of Heiner [49]. Even if the density profile of water close to the cellulose surface (see figure 4.2 on page 4.2) is characteristic for a hydrophobic surface [40, 49], cellulose appears to be well hydrated in terms of cellulose-water hydrogen bonds.

4.5 Lateral Structure of the Cellulose-Water Interface

The adsorption of a polyelectrolyte on a surface can be caused not only by polar hydrogen-bonds, but also by dispersion interactions (lipophilic, hydrophobic interactions) as they are typical for aliphatic polymers. Even if there are many polar sites (charged and partially charged oxygens, OH-hydrogens) present in (carboxy methyl) cellulose (CMC), and poly-acrylate (PAA), cellulose and water, the polyelectrolytes are built by apolar sites (aliphatic carbons and ether oxygens) as well. Now the cellulose-water interface is analyzed to identify hydrophobic/hydrophilic and lipophobic/lipophilic domains by means of the excess chemical potential of water (hydrophilic molecule) and argon (lipophilic particle). Surface maps of the chemical potential are completed by maps of atom densities.

The chemical potential of water and argon (Ar) and other observables are resolved spatially with respect to a volume around point $\mathbf{r}$ on the surface. The surface is defined as the z-coordinate where the water and cellulose densities are equal (figure 4.2). For 110 we take $z_0 = 3.125\,\mathrm{nm}$ and for $1\bar{1}0$ $z_0 = 3.5\,\mathrm{nm}$, all analysis here is confined to the upper interface L1.

All surface maps are a 33×48 matrix of small orthorhombic volume elements with dimension and volume

$$0.095 \times 0.102 \times 0.2\,\mathrm{nm} = 1.94 \cdot 10^{-3}\,\mathrm{nm}^3 \quad \text{for the 110 simulation}$$

and

$$0.095 \times 0.090 \times 0.2\,\mathrm{nm} = 1.71 \cdot 10^{-3}\,\mathrm{nm}^3 \quad \text{for } 1\bar{1}0 \text{ simulation.}$$

All observables are calculated and averaged for these sub-volumes, which are taken for different z-heights to calculate a three dimensional representation of the crystal-interface-water structure The color-coding of the maps is given in figure 4.11 and ranges from *red* for high values over neutral *green* to *blue* negative values.

During the equilibration phase the positions of the glucose ring atoms within the anhydro glucose unit were fixed to the initial crystal positions (see section 2.4). However, after the restraints had been removed, the entire crystal started to diffuse in xy direction (figure 4.12). This process was balanced by the diffusion of the center of mass of the solvent, so that the total momentum of the simulation box was zero (which is guaranteed by the simulation program). There is no component along the surface normal, because this cannot be balanced by the solvent without the solvent penetrating through the crystal. All maps in the following sections are therefore corrected to the initial center of mass. They show the upper surface, viewed from the water bulk along the z-axis. Cellulose chains run from the left to the right, eight chains are staggered from the top to the bottom.

4.5.1 Atom Density Maps of the Interface

A simple picture of the surface is given by single-atoms densities. The color-range in legend (figure 4.11) comprises values from blue ($\rho < 200\,\mathrm{kg\,m^{-3}}$) to red ($\rho > 1200\,\mathrm{kg\,m^{-3}}$). Green is approximately equal to a density of $600\,\mathrm{kg\,m^{-3}}$. The maps in figures 4.13 to 4.17 account only for atoms, which contribute to the outermost peak (of a given atom type) in the z-distributions (see figures 4.3 a) to d)). On going from the bulk water to the interfacial cellulose, the atom-maps comprise only for the z-coordinates of the interfacial, outermost peak of the given atom type. Therefore the maps of different atom-types extend over

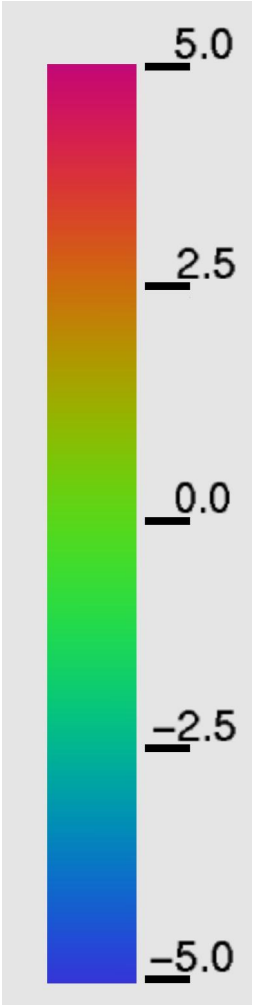


Figure 4.11: Legend for maps of chemical potential difference $\Delta\mu$ for water and argon. Red color stands for a high (+5 kJ/mol), blue for a negative (−5 kJ/mol) difference to the bulk value. This color scheme is applied to other maps as well, but with different reference points.

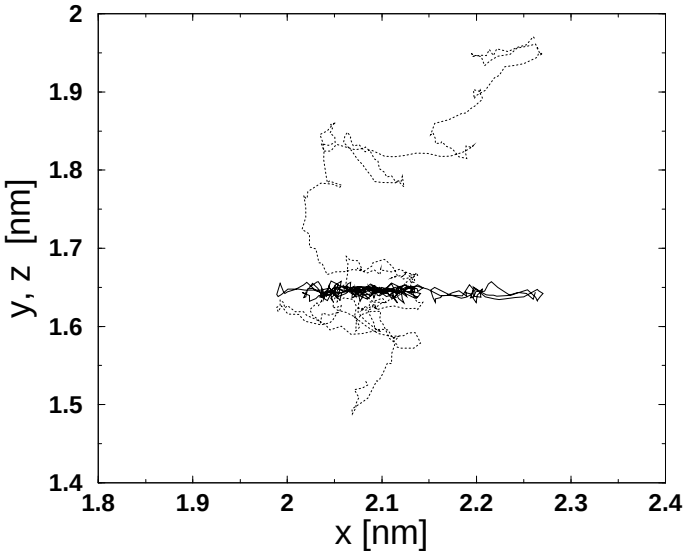


Figure 4.12: Coordinates of the crystal center-of-mass (com) during the 110 simulation. The solid line is the xz-component of the com, there is no movement in z-direction. The dotted line is the displacement along the xy-plane. The total amount is in the order of 0.5 nm during the whole simulation. This corresponds to a diffusion constant which is about 1500 times smaller than that of a single bulk water molecule. The drift was corrected for in all surface maps as they rely on absolute positions. No correction was necessary if an analysis depends only on particle distances, like radial distribution functions or hydrogen-bonds.

different intervals in z-direction. Figure 4.13 and 4.14 show the locations of two atoms types incorporated in the backbone (C1 and O4) for both surface simulations. The atoms are located on the surface in a region of about $0.2 \times 0.2\,\mathrm{nm}^2$ and the maps show the regular pattern of the underlying crystal. One clearly identifies eight parallel cellulose chains running from the left to the right in a zig-zag fashion. Every second chain is shifted by $c/4$ to their neighbor chains. There are six C1-spots in a chain, one for every AGU.

Cellulose glycosidic bonds are directed up- or downwards in an alternating fashion. Because of this only three O4-atoms per show up for the 110 surface in figure 4.14 a), whereas all six atoms of a chain are visible for the narrow $1\bar{1}0$ surface. This corresponds to the double peak found for O4-density in z-direction in figure 4.3 a and single-peak structure in 4.3 b.

In addition to the two surfaces, there are two crystallographically different types of chains, denoted here as *even* and *odd*. The terms even and odd refer to the sequence number of the cellulose chains counted from the top of the map to the bottom. Depending on the observable and surface both chains do have different characteristics [49, 52] which is suggested already by the atom densities. The differences between the even and odd density patterns are more obvious for the narrow $1\bar{1}0$ surface, but are present for the wide 110-surface one as well.

The density pattern for O2 and O3 are similar (figures 4.15 and 4.16) with no remarkable features. According to figure 4.17, the high density patches for O6 are more diffuse,

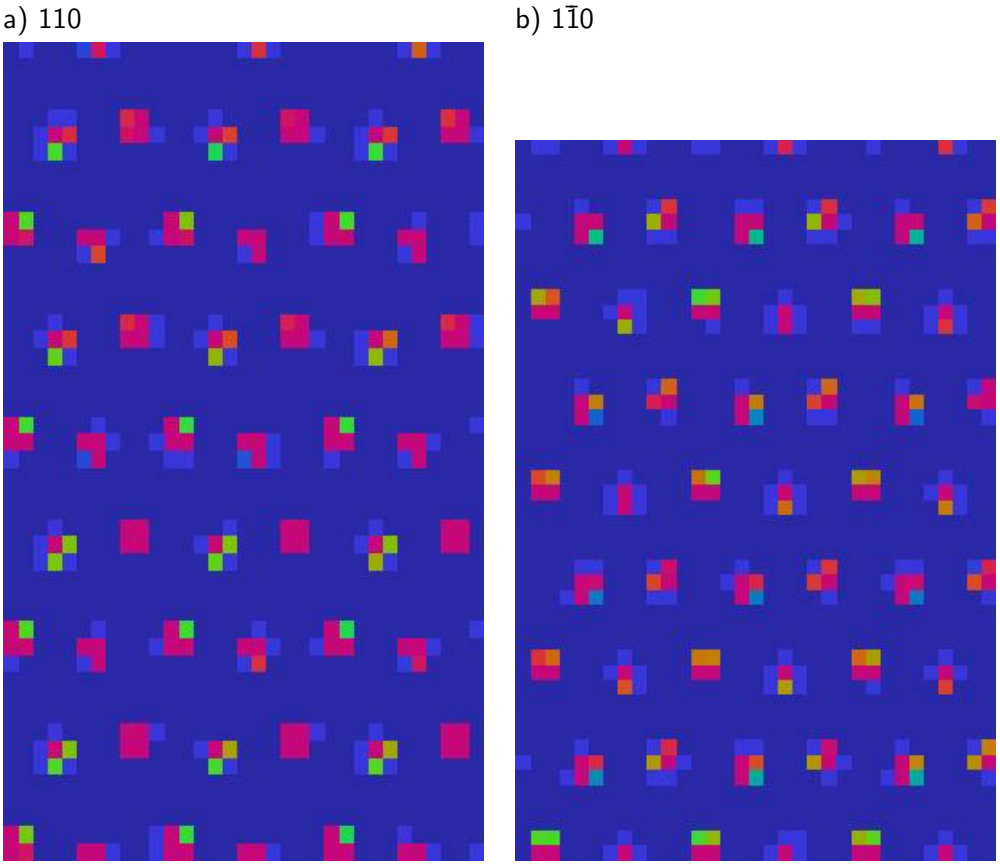


Figure 4.13: Density map of C1-atoms on the a) 110 and b) 1$\bar{1}$0 surface. There are eight cellulose chains running from the left to the right, one identifies the $c/4$ shift of adjacent chains, which are packed vertically (along the y-axis). Every cellobiose unit is visible through two C1-atoms, so there are in total 6×8 C1-spots visible. The map gives a impression on the size differences between wide (110) and narrow (1$\bar{1}$0) surfaces. Apart from this both surfaces look very similar. Every atom resides in a region of about 2×2 grid points, which is an area of about $0.2 \times 0.2\,\mathrm{nm}^2$. The color code ranges from blue ($\rho < 200\,\mathrm{kg\,m^{-3}}$) to red ($\rho > 1200\,\mathrm{kg\,m^{-3}}$). There is almost no difference between even and odd chains for 110, but for 1$\bar{1}$0 every second chain has a little different structure.

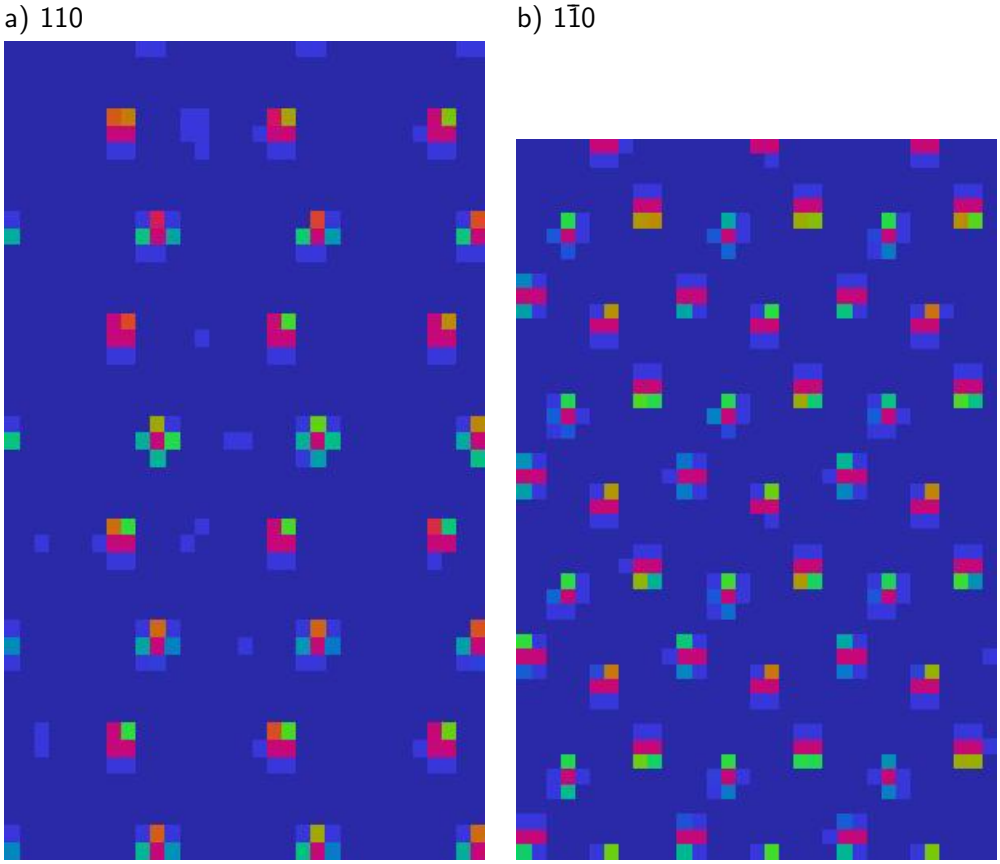


Figure 4.14: Density map of O4-atoms on the a) 110 and b) 1$\bar{1}$0 surface. As the maps are centered in z-direction around the outermost density peaks of O4 (see section 4.2) there are three spots/chain for the wide surface and six spots/chain for the narrow surface. This is due the cellulose crystal packing, where there is no axis possible through all O4-atoms which is parallel to the c-axis. The color code ranges from blue ($\rho < 200\,\mathrm{kg\,m^{-3}}$) to red ($\rho > 1200\,\mathrm{kg\,m^{-3}}$).

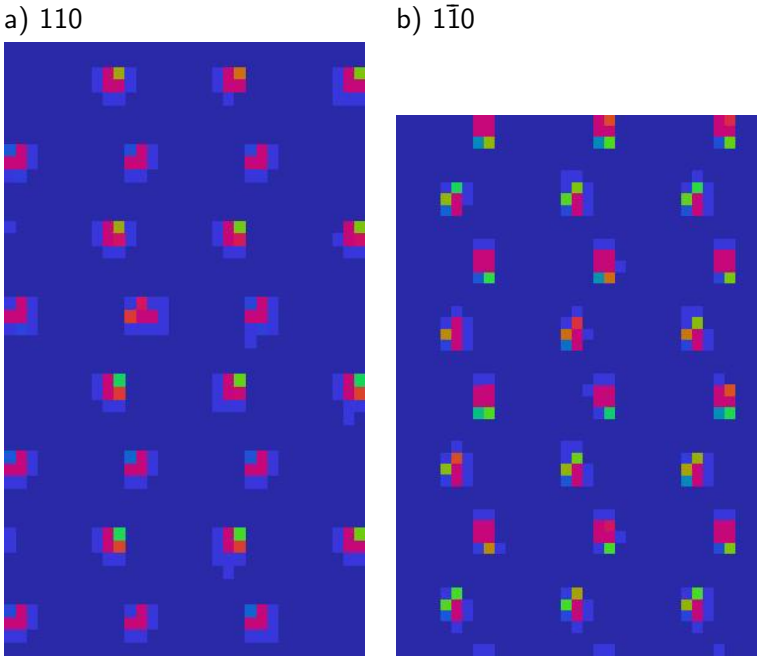


Figure 4.15: Density map of O2-atoms on the a) 110 and b) 1$\bar{1}$0 surface. Every second O2 is protruding into the water bulk, the pattern for the wide and narrow surface are much alike as proposed in the schematic diagram in figure 1.2. The color code ranges from blue ($\rho < 200\,\mathrm{kg\,m^{-3}}$) to red ($\rho > 1200\,\mathrm{kg\,m^{-3}}$).

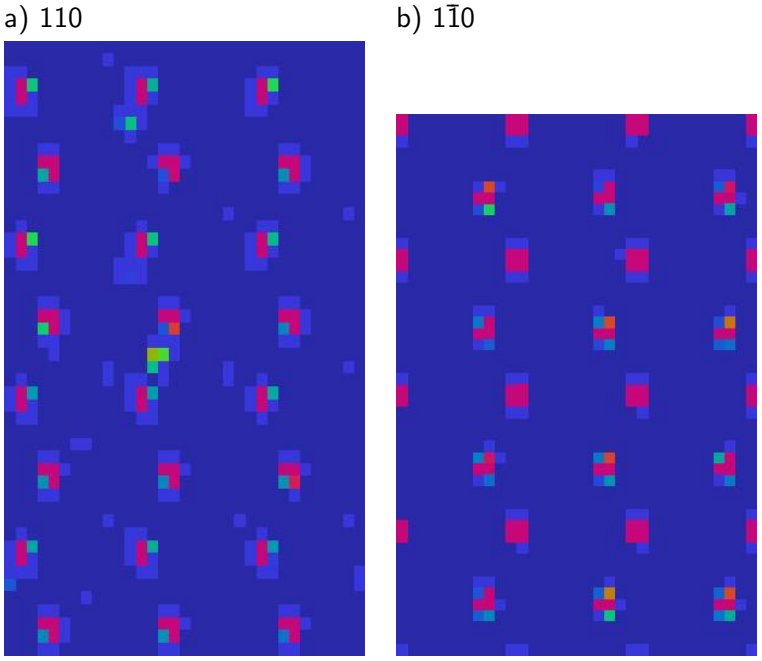


Figure 4.16: Density map of O3-atoms on the a) 110 and b) 1$\bar{1}$0 surface. The color code ranges from blue ($\rho < 200\,\mathrm{kg\,m^{-3}}$) to red ($\rho > 1200\,\mathrm{kg\,m^{-3}}$).

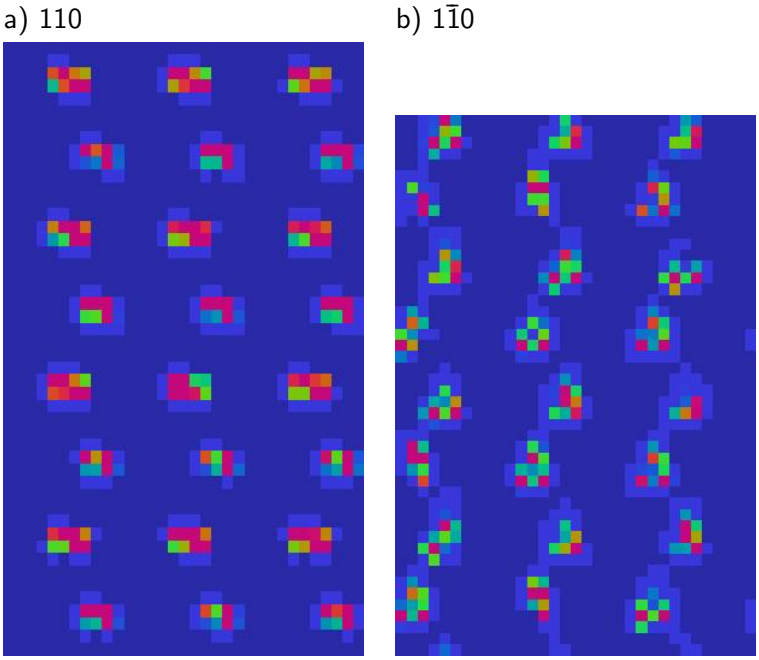


Figure 4.17: Density map of O6-atoms on the a) 110 and b) 1$\bar{1}$0 surface. The color code ranges from blue ($\rho < 200\,\mathrm{kg\,m^{-3}}$) to red ($\rho > 1200\,\mathrm{kg\,m^{-3}}$).

covering $0.4 \times 0.2\,\mathrm{nm}^2$ (110) and $0.4 \times 0.4\,\mathrm{nm}^2$ (1$\bar{1}$0) but with increased disorder for single atoms on the narrow 1$\bar{1}$0 surface.

Atoms on the surface are connected tightly to the cellulose backbone. Only in the case of O6, which is connected to the chain by a spacer, there is some wider distribution of atom positions so it is more diffuse (figure 4.17). Oxygen atoms are recognized as density patches with an area of about $0.04\,\mathrm{nm}^2$ (O2 and O3), respectively $0.16\,\mathrm{nm}^2$ for O6. If grid points with significant oxygen density $\rho > 400\,\mathrm{kg\,m^{-3}}$ are summed, about 58% of the surface are covered with oxygen density/atoms.

The density pattern of O6 fits the height profile from AFM [5], and shows a zig-zag line of O6 intensities perpendicular to the cellulose chain direction (from top to bottom in figure 4.17). Density spots along the cellulose chain direction are shifted by $c/4$, as proposed by Baker (see [5, figure 6]). Other atom-types, which do not extent into the solvent are not visible AFM-images.

4.5.2 The Chemical Potential of Water and Argon

The adsorption behavior of different polyelectrolytes onto cellulose involves desolvatisation of the surface as well as hydrophobic polymer-cellulose interactions. Hence, we make an attempt to characterize the cellulose in its behavior towards water (hydrophilicity/lipophobicity) and towards apolar parts of organic molecules (lipophilicity/hydrophobicity).

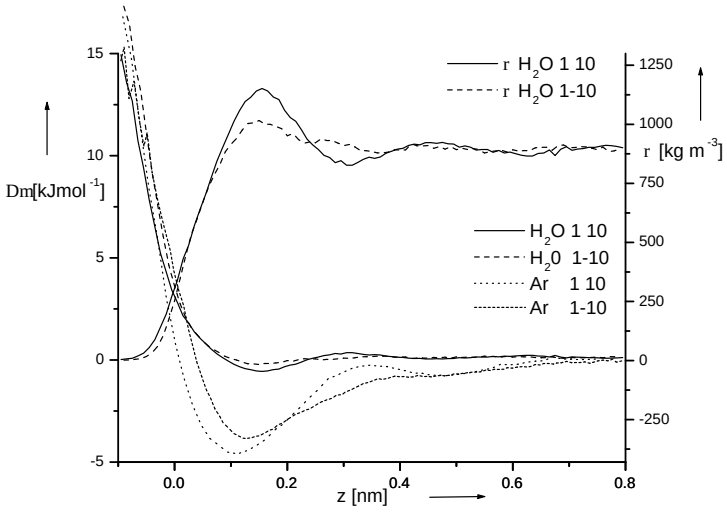


Figure 4.18: Local chemical potential in kJ/mol of water $\Delta\mu_{\mathrm{H_2O}}(z)$ and argon $\Delta\mu_{\mathrm{Ar}}(z)$ as a function of the z-distance from the cellulose Iβ 110 and $1\bar{1}0$ surfaces, respectively. The water density is plotted on the right y-axis.

4.5.2.1 Chemical Potential on the Surface

Lipophilicity is not necessarily the opposite of hydrophilicity, but must be quantified by the local chemical potential of a "lipophile", which has to be chosen. Since our interest is in the adsorption of small lipophilic moieties, rather than entire molecules, we chose argon as the model lipophile [42]. The argon atom is similar in size and cohesion to a methylene group. Moreover, its size is comparable to that of a water molecule, so probes of similar diameter are used for both hydrophilicity and lipophilicity. Its local excess chemical potential $\Delta\mu^{\mathrm{ex}}_{\mathrm{Ar}}(\mathbf{r})$ is obtained by spatially resolved test-particle insertion [91] (section 1.3.8):

$$\mu^{\mathrm{ex}}_{\mathrm{Ar}}(r) = -\mathrm{R}T \ln \langle \exp(-\beta U(\mathbf{r})) \rangle \tag{4.1}$$

where $U(\mathbf{r})$ is the would-be change in potential energy when an argon atom is inserted into the cellulose-water system at position $\mathbf{r}$. The excess chemical potential of Ar in bulk water $\Delta\mu^{\mathrm{ex}}_{\mathrm{Ar}}(\mathrm{bulk}) = 6.0\,\mathrm{kJ/mol}$ (experiment: $\approx 10\,\mathrm{kJ/mol}$ [79]) is calculated in a similar way. Figure 4.18 shows

$$\Delta\mu_{\mathrm{Ar}}(\mathbf{r}) = \mu^{\mathrm{ex}}_{\mathrm{Ar}}(\mathbf{r}) - \mu^{\mathrm{ex}}_{\mathrm{Ar}}(\mathrm{bulk}).$$

Both surfaces exhibit clear minima of $-4.5\,\mathrm{kJ/mol}$ (110) and $-3.8\,\mathrm{kJ/mol}$ ($1\bar{1}0$). As kT at 333 K corresponds to 2.8 kJ/mol, both would attract small lipophiles from aqueous solution. More obvious than in the corresponding $\Delta\mu$ curves (see below) for water, the argon curve of 110 has two minima, the one for $1\bar{1}0$ has one. The intervening maximum at $z \approx 0.32\,\mathrm{nm}$ of the 110 curve is a shoulder for $1\bar{1}0$.

The hydrophilicity/hydrophobicity at a point $\mathbf{r}$ is quantified by the local value of the chemical potential of water $\mu_{\mathrm{H_2O}}(\mathbf{r})$ relative to that of bulk water $\mu_{\mathrm{H_2O}}(\mathrm{bulk})$. It is given by the ratio of the local water density $\rho_{\mathrm{H_2O}}(\mathbf{r})$ and the bulk density $\rho_{\mathrm{H_2O}}(\mathrm{bulk})$;

$$\Delta\mu_{\mathrm{H_2O}}(\mathbf{r}) \quad = \mu_{\mathrm{H_2O}}(\mathbf{r}) - \mu_{\mathrm{H_2O}}(\mathrm{bulk}) \tag{4.2}$$

$$= -\mathrm{k}T \ln \frac{\rho_{\mathrm{H_2O}}(\mathbf{r})}{\rho_{\mathrm{bulk}}(\mathrm{H_2O})}. \tag{4.3}$$

A value of $\Delta\mu_{\mathrm{H_2O}}(\mathbf{r}) < 0$ indicates hydrophilicity and $\Delta\mu_{\mathrm{H_2O}}(\mathbf{r}) > 0$ hydrophobicity at $\mathbf{r}$. Figure 4.18 shows $\Delta\mu_{\mathrm{H_2O}}$ as a function of the distance z from the cellulose surface, i.e. averaged over x and y. The enrichment of water near both surfaces is minute and the basins of negative chemical potential at $z \approx 0.15\,\mathrm{nm}$ are shallow (110: $-0.56\,\mathrm{kJ/mol}$, $1\bar{1}0$: $-0.21\,\mathrm{kJ/mol}$). Hence, both surfaces are in essence non-hydrophilic. This is also evident in the water density profile normal to the surface (shown on the second axis in figure 4.18), which lacks the sharp first peak characteristic for hydrophilic surfaces [40]. The 110 surface has a second, even shallower, minimum (at $z \approx 0.47\,\mathrm{nm}$), which we attribute to its wider lateral spacing between the cellulose chains.

Hydrophilicity and lipophilicity can also be resolved laterally like the atom densities to study the distribution of hydrophilic/phobic and lipophilic/phobic spots on the surface. We use the same grid setup as in 4.5.1, but with a fixed slab thickness (in z-direction) of 0.2 nm. The effect of the crystal structure on hydrophobicity and lipophilicity patterns is clearly visible (figures 4.19 and 4.20), with only some variation due to a small degree of disorder. In most cases, the lipophobic and hydrophobic areas (both dark red) coincide for both surfaces, the lipophobic areas being somewhat larger. Surprising at first, this is easily explained, as the exclusion is steric for both water and argon. There is, however, a qualitative difference between the hydrophilic and lipophilic regions: Lipophilic patches (dark blue) are nearly equal in size to the lipophobic ones. In contrast, there are only small hydrophilic patches, most of the non-hydrophobic surface is "neutral" green ($\Delta\mu_{\mathrm{H_2O}} \approx 0\,\mathrm{kJ/mol}$), leading to the same water density as bulk water. The differences between the patterns of the two surfaces are small for both hydrophilicity and lipophilicity.

This is further stressed by the two dimensional radial distribution function $g_{2\mathrm{D}}(r)$ of lipohilic ($\Delta\mu_{\mathrm{Ar}} < 0\,\mathrm{kJ/mol}$, "A"ttractive) and lipophobic regions ($\Delta\mu_{\mathrm{Ar}} > 0\,\mathrm{kJ/mol}$, "R"epulsive) patches on the surface: The distance r is defined as $r = (x^2 + y^2)^{1/2}$ in two dimensions parallel to the surface. Where x and y are the coordinates of an argon insertion point (the ghost argon is inserted into a slab between $z_0 = 0\,\mathrm{nm}$ and $z = 0.2\,\mathrm{nm}$). The z-coordinate of the ghost particle is not taken into account for the calculation of $g_{2\mathrm{D}}(r)$. Depending on the insertion energy or chemical potential at the insertion location, this point is termed attractive (A) or repulsive (R) to argon atoms. After 5000 insertion attempts for one simulation frame the radial distribution function is calculated between "A" and "R" particles. The distributions are given in figure 4.21. They show the extension of the attractive and repulsive patches.

The attractive-attractive and repulsive-repulsive rdf functions exhibit a first maximum for $r = 0\,\mathrm{nm}$, as it is expected: There are always patches of one kind close together. As a consequence, "A" and "R" patches repel each other and the corresponding A-R distribution in figure 4.21 starts at $g(0\,\mathrm{nm}) = 0.5$. It does not start at 0, because the z-coordinate is not taken into account, and attractive and repulsive sites may overlap ($r = 0$ but $\Delta z \neq 0$). Attractive "A" locations are clustered together in a stronger way, as the repulsive ones. The "A"-patches are larger. However it is surprising to find the curves for the 110 and $1\bar{1}0$ surface to be almost equal, which is not evident from the maps in section 4.5.

Local hydrophilicity and lipophilicity are related to the functional groups being presented to the water phase. Figures 4.22 and 4.23 show hydrophilicity and lipophilicity

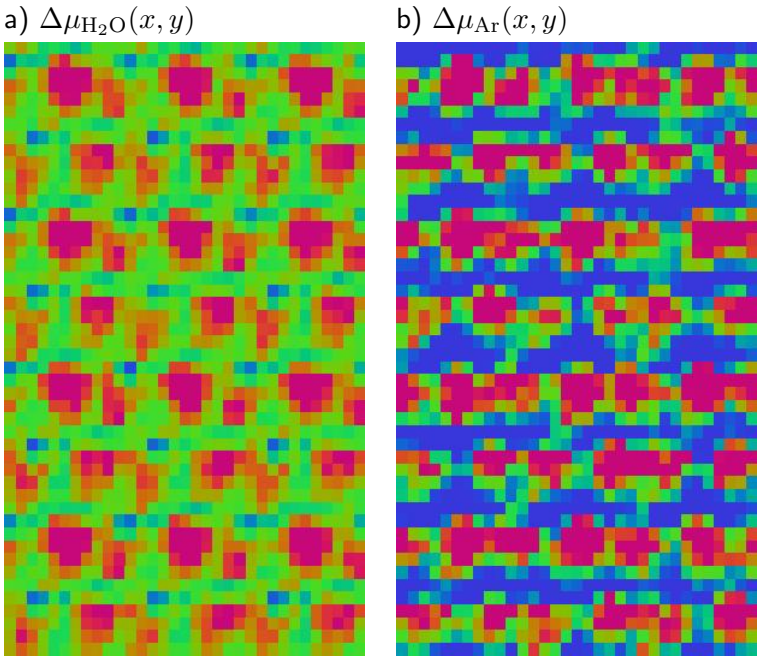


Figure 4.19: Hydrophilicity and lipophilicity charts of the monoclinic cellulose 110 surface in a layer between $z = 0\,\text{nm}$ and $z = 0.2\,\text{nm}$. Eight parallel cellulose chains run from left to right. The lateral grid spacing is approximately 0.1 nm, see text. The color code ranges from $-5\,\text{kJ/mol}$ (blue) to $+5\,\text{kJ/mol}$ (red) a) $\Delta\mu_{H_2O}(x, y)$, b) $\Delta\mu_{Ar}(x, y)$. The figures are given in figure A.3 in the appendix with drawn in cellulose strucure formulae.

above a single cellobiose unit on the surface. They have been obtained by averaging over all corresponding repeat units and, thus, include static and dynamic disorder. The charts are given for even and odd cellulose chains (see section 4.5.1) separately.

For both surfaces and both types of cellulose chain (figure 4.22 a/b and 4.23 a/b), the protruding C6/O6 units and, to a lesser extent, the C2/O2 moieties create hydrophobic areas (high water chemical potential). This is due to their space requirements. The effect is slightly more pronounced for the narrow $1\bar{1}0$ surface (figure 4.23) due to the stronger tilt of the glucose rings. The chain packing in the two surfaces also has an effect on the type of contacts a water molecule has with the surface groups (table 4.10). This can be extracted by integrating the appropriate partial radial distribution function [94]. For both surfaces, half of the contacts of a nearby water molecule are, on average, hydrophilic OH groups. On the wider 110 surface the remaining surface contacts are equally divided between hydrophobic groups and "intermediate" groups. This distribution is shifted from the hydrophobic to the intermediate groups in case of the narrow $1\bar{1}0$ surface.

The lipophobicity distribution (high argon chemical potential = red) follows qualitatively the hydrophobicity, since it is also mainly due to excluded volume (figure 4.22 c/d and 4.23 c/d). However, the spatial variation of $\Delta\mu_{Ar}(x, y)$ is larger than that of $\Delta\mu_{H_2O}(x, y)$, and there are attractive (blue) sites as well. For the 110 surface, they form continues grooves (figure 4.22 c/d) , whereas on the narrow $1\bar{1}0$ surface the attractive

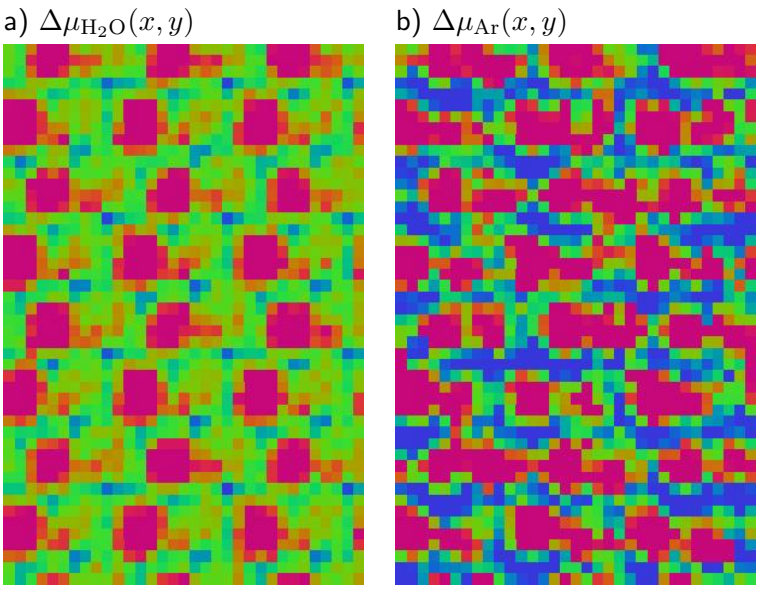


Figure 4.20: Hydrophilicity and lipophilicity charts of the monoclinic cellulose $1\bar{1}0$ (narrow) surface in a layer between $z = 0\,\text{nm}$ and $z = 0.2\,\text{nm}$. Details as in figure 4.19. a) $\Delta\mu_{H_2O}(x, y)$, b) $\Delta\mu_{Ar}(x, y)$.

		H_2O		Ar	
		110	$1\bar{1}0$	110	$1\bar{1}0$
Hydropphilic :	O2, O3, O6	49	52	20	31
Intermediate:	C2, C3, C6	26	29	22	27
Hydrophobic :	C1, C4, C5, O4, O5	25	19	58	42

Table 4.10: Types of surface contacts (in percent of all surface contacts) for a water molecule and an argon as the prototype lipophile. The absolute number of surface contacts of argon/water to a surface's site is defined as the integral of the respective radial distribution function to the first minimum (number or nearest neighbors).

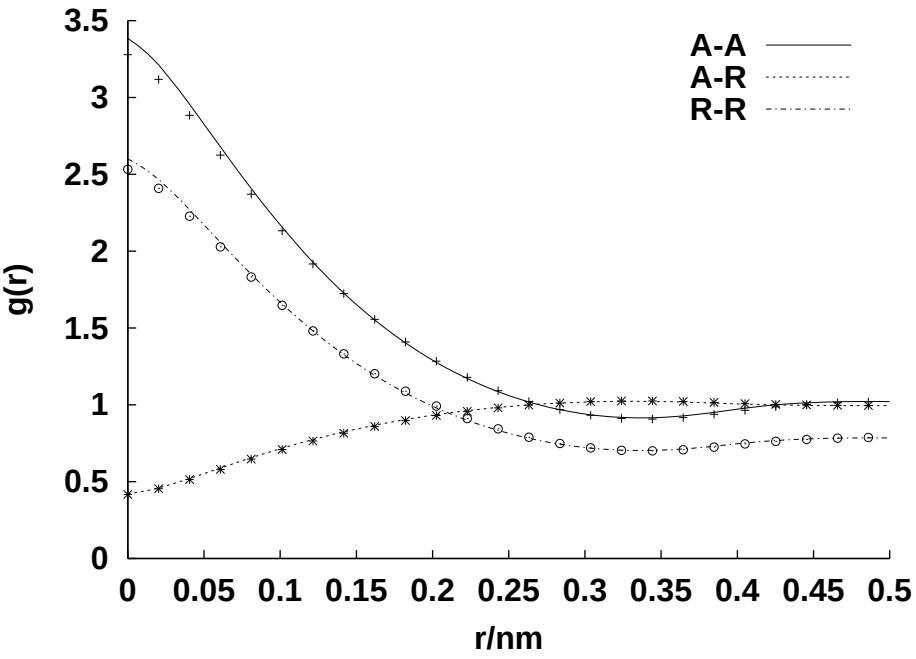


Figure 4.21: Radial distribution functions ($g_{2D}(r)$, no units) in two dimensions between lipophilic (A) and -phobic (R) insertion points on the 110 (lines) and $1\bar{1}0$ (points) surface. Insertion points with an instant chemical potential lower 0 kJ/mol are named "A"ttractive, otherwise they are "R"epulsive. Radial distribution functions are calculated in the surface slabs (0.0–0.2 nm). The distance r is only calculated in two dimensions, irrespective of the z-value of the insertion: $r = \sqrt{x^2 + y^2}$.

patches are not connected (figure 4.23 c and d), probably due to the lack of space. The lipophilic areas are generally located in depressions on the surface and are lined by hydrophobic atoms such as ether oxygen (O4 and O5) and apolar carbons (C1, C4 and C5). This is also reflected in the surface contacts of argon (table 4.10): On the $1\bar{1}0$ surface, the distribution of Ar contacts closely resembles that of water contacts, steric accessibility being the criterion for both interactions. In contrast, on the wider 110 surface, the hydrophobic groups are more exposed, and more interactions with them are found.

The radial distribution function between the inserted argon ghost particle and cellulose (see section 1.3.8) atoms in the surface layer are given for the 110-system in figure 4.24 and for $1\bar{1}0$ in 4.25. For the hydrophilic sites O3 and O6 there is no significant difference between the surface, but for O2 there is. With respect to the $1\bar{1}0$ surface, the 110 one lacks the first peak at (0.32 nm), but has only a small hump. This is in accordance to the trend of the O2-water radial distribution functions, even if the difference is rather pronounced for the Ar-H_2O distributions.

One difference between the wide 110 and narrow $1\bar{1}0$ surface is observed for oxygen O5, which is accessible on the wide surface, whereas it does not exhibt any clear peak on the narrow surface, which is plausible by simple steric arguments: the core glucose ring's atoms are more shielded by the hydroxy-groups on the narrow $1\bar{1}0$ surface. The same holds true for O4 and other ring-atoms (figure 4.24 and 4.25 b), c) and d)). The carbon atoms

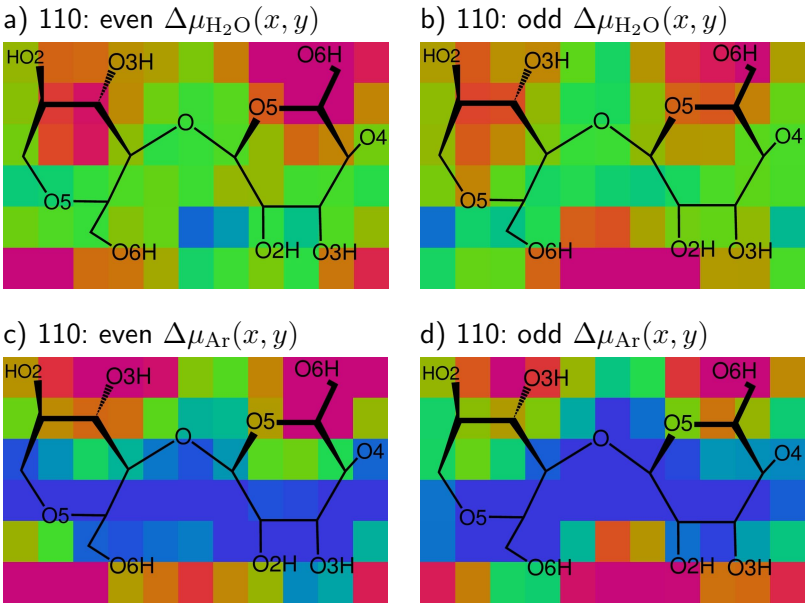


Figure 4.22: Averaged hydrophilicity $\Delta\mu_{H_2O}(x,y)$ and lipophilicity $\Delta\mu_{Ar}(x,y)$ above a single cellobiose unit on the monoclinic cellulose 110 (wide) surface. a) $\Delta\mu_{H_2O}(x,y)$ for even-numbered cellulose chains, b) $\Delta\mu_{H_2O}(x,y)$ for odd-numbered chains. c) $\Delta\mu_{Ar}(x,y)$ for even-numbered chains. d) $\Delta\mu_{Ar}(x,y)$ for odd-numbered chains, respectively, in a layer above the surface between $z = 0\,\text{nm}$ and $z = 0.2\,\text{nm}$. The color code is as in figure 4.19. The drawn-in structural formulae reflect average atomic positions and are meant as orientation aid. See figures 4.13 to 4.17 for calculated atom positions.

C1, C4, C5 have first peaks intensity from the wide 110 surface at about 0.33 nm, whereas on the narrow surface the Ar-carbon distributions raise slower to at $r_{max} = 0.44\,\text{nm}$.

4.5.2.2 Chemical Potential in the Bulk

The lipophilicity/hydrophilicity maps do not only provide information in the lateral (x,y) direction, but are also resolved in z-direction. Thus can be used to reveal a three-dimensional picture.

Figure 4.26 shows the chemical potential $\Delta\mu_{H_2O}(x,y)$ of water in overlapping layers of different z-height from the cellulose 110 bulk to the water solvent. As it is already shown in figure 4.18, the raise of the water density towards the bulk value is abrupt. Below the surface ($z = -0.2\,\text{nm}$ to $0.0\,\text{nm}$) there is hardly any water. Apart from tiny green or yellow spots the entire map is red. This indicates a $\Delta\mu_{H_2O}(x,y)$ larger 5 kJ/mol. The few spots in figure 4.26 exhibit a pattern along the cellulose chains with three repeat units, resembling the cellobiose units. They are located in a groove close to O4 and C1. The slab from $z = -0.1$ to $0.1\,\text{nm}$ shows a regular pattern, which is repeated three times along the chain direction. Even and odd cellulose chains are joined by vertical regions (along y-axis) of water density. These patches run from the glycosidic links (O4) of one chain to that one of the next chain (the bottom to the top of the chart). In the next higher slab (the interface-slab from 0 nm to 0.2 nm) the water density becomes close to the water

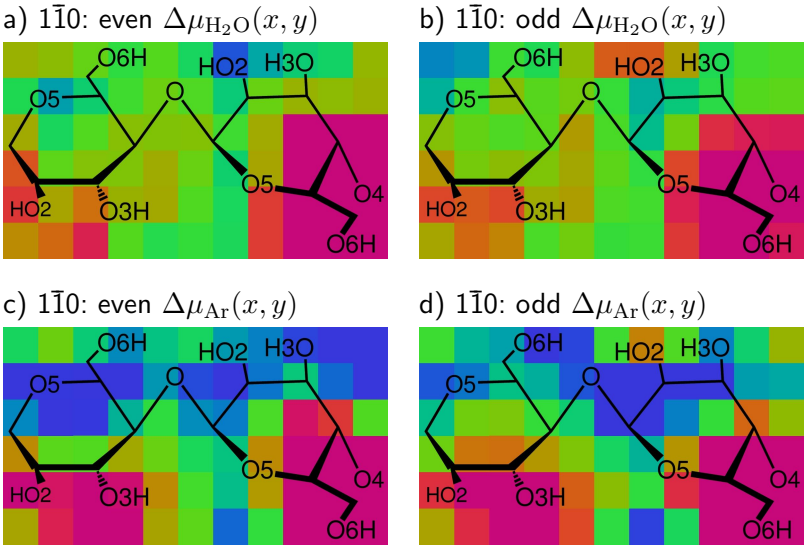


Figure 4.23: Averaged hydrophilicity $\Delta\mu_{H_2O}(x,y)$ and lipophilicity $\Delta\mu_{Ar}(x,y)$ above a single cellobiose unit on the cellulose 1$\bar{1}$0 (narrow) surface. a) $\Delta\mu_{H_2O}(x,y)$ for even-numbered cellulose chains, b) $\Delta\mu_{H_2O}(x,y)$ for odd-numbered chains. c) $\Delta\mu_{Ar}(x,y)$ for even-numbered chains. d) $\Delta\mu_{Ar}(x,y)$ for odd-numbered chains. Details as in figure 4.22.

bulk value (green areas) with some blue spots which are between an oxygen O3 of the one cellulose chain and O2 of the next one one in y-direction. Single cellobiose units are resolved in figure 4.26 c, even with a bigger red patch for the first (from left to right) AGU visible (O6 pointing along z-axis) and a smaller one for the AGU with O2 and O3 pointing upwards. The next map (0.1 nm–0.3 nm) of $\Delta\mu_{H_2O}(x,y)$ is almost uniformly green. Only a few tenths of a nanometer above the surface, the water properties are close to the bulk values, even if the cellulose pattern is still present.

The narrow 1$\bar{1}$0 surface maps of water density are shown in figure 4.27. The general trend is as in figure 4.26 for the 110 surface: 0.2 nm below the surface ($z = -0.2$ to 0 nm) there is hardly any evidence of water. In the slab above ($z = -0.1$ to 0.1 nm) there are only some localized regions with noticeable water density. Along the cellulose chain direction there is a three-patch pattern, resembling the number of cellobiose units. The chains are connected by vertical (along the y-axis) density regions, like for the wide 110 surface. Different from the 110 surface these regions of high water density zig-zag diagonal across the surface. The tilt of the cellulose chains is opposite to that shown for the wide 110 surface (see figure 4.22 or 4.23 for comparison). Patches of high water density connect two glycosidic links O4 with each other.

On the interface ($z = 0$ nm to 0.2 nm), a high percentage of the surface is covered with water at bulk density. Only few excluded-volume patches (discussion above) are visible. Stepping further into the bulk (0.1 nm–0.3 nm), the water profile is not as flat as for the other surface with still some cellobiose-structure. This is due to the close packing of the narrow surface, where cellulose atoms are forced to extend into the water bulk (see atoms-densities along the z-axis in section 4.2)

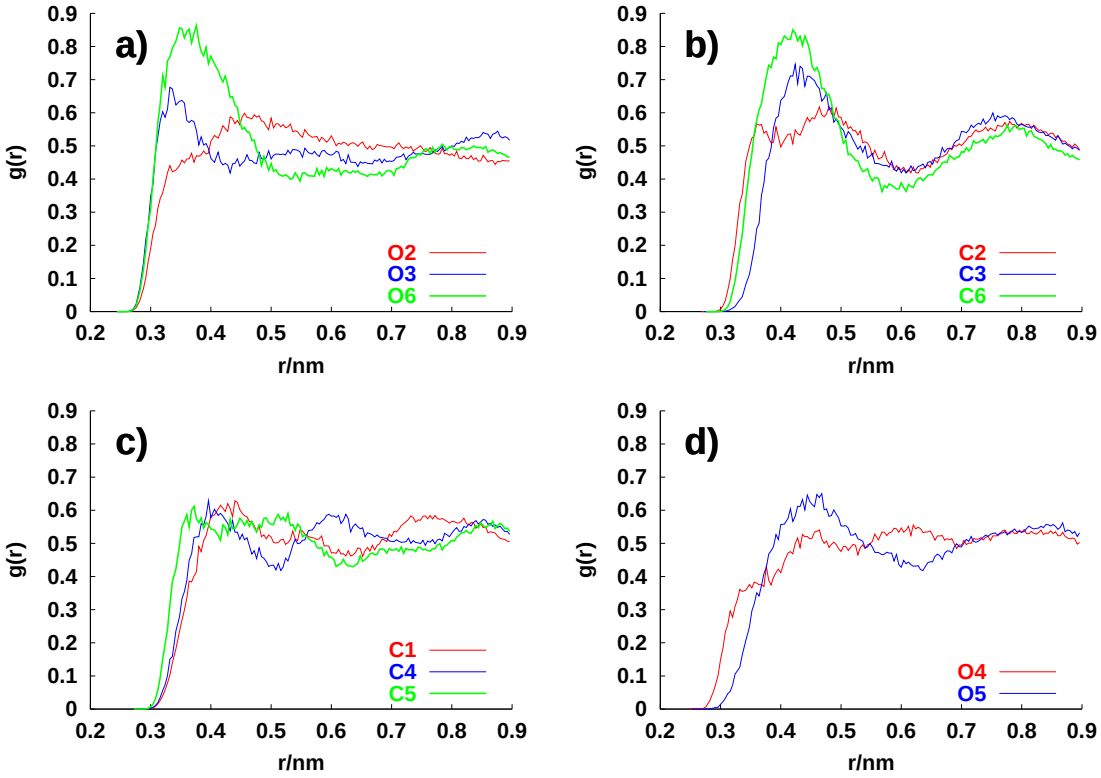


Figure 4.24: Radial distribution functions between a ghost argon atom and different atom sites on the 110-surface. The functions were calculated by inserting the ghost particle and calculating a Boltzman-weighted distribution to the reference sites which is converted to a radial distribution through a Boltzman-inversion [103]. The functions do not approach unity, as they are only taken in a half-space (z from 2.9 to 2.8 nm) above the surface. Differing from the input in table 2.7 a minimum of 15 000 successful insertions per frame was used. Reference-sites of the surface's interface layers: a) O2, O3 and O6 b) C2, C3, C6 c) C1, C4, C5 and d) O4, O5.

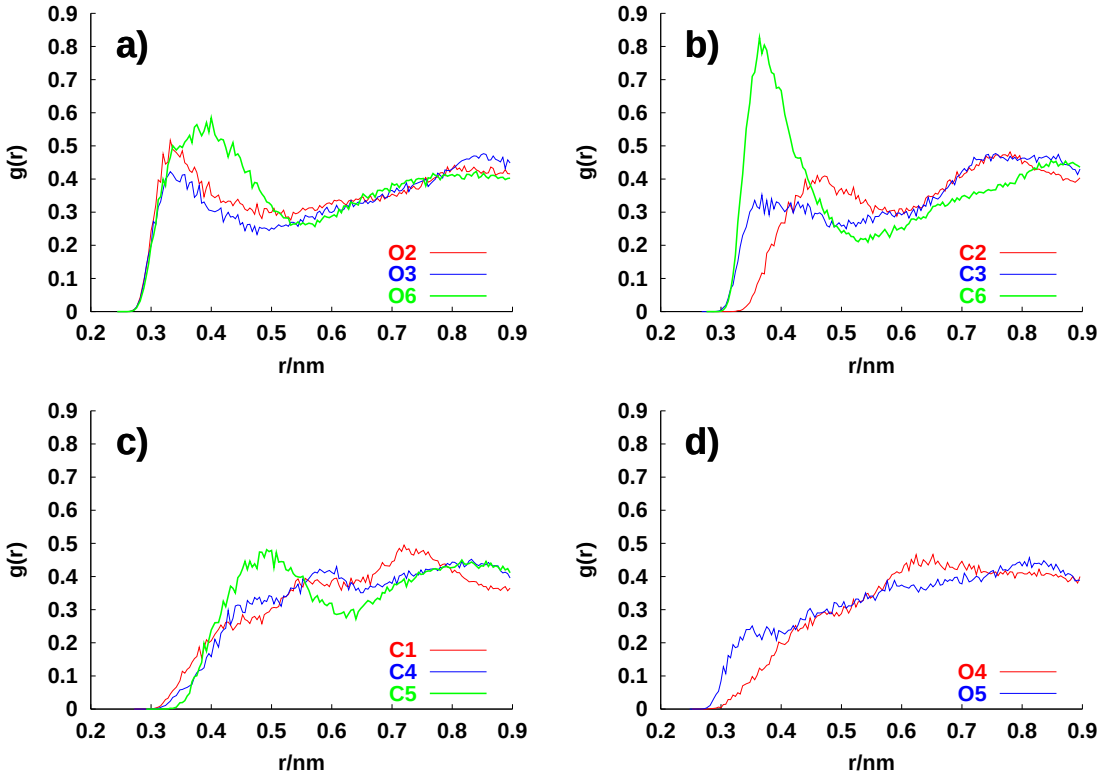


Figure 4.25: Radial distribution functions between a ghost argon atom and different atom sites on the $1\bar{1}0$-surface. The functions were calculated by inserting the ghost particle and calculating a Boltzman-weighted distribution to the reference sites which is converted to a radial distribution through a Boltzman-inversion [103].

The functions do not approach unity, as they are only calculated in a half-space (z from 3.5–4.4 nm) above the surface.

Reference-sites of the surface's interface layers: a) O2, O3 and O6 b) C2, C3, C6 c) C1, C4, C5 and d) O4, O5.

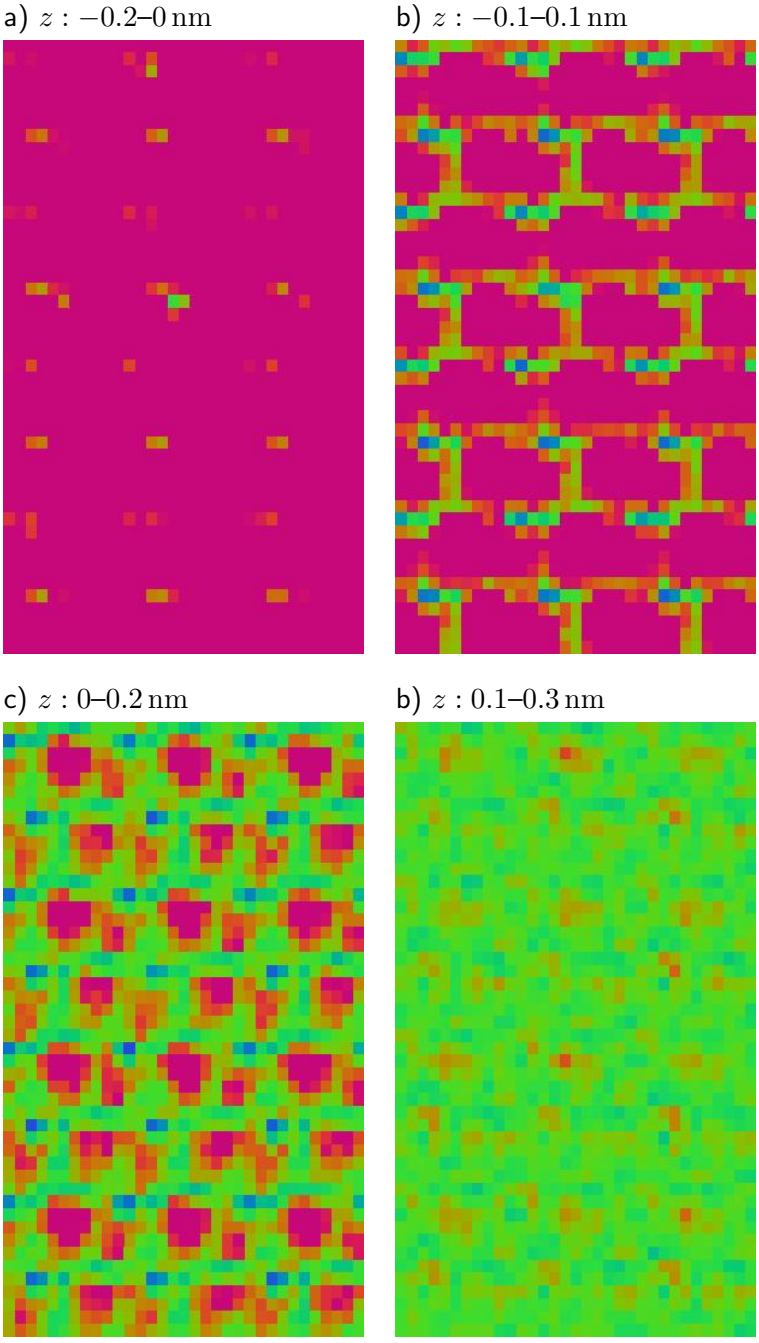


Figure 4.26: Hydrophilicity/hydrophobicity maps ($\Delta\mu_{H_2O}(x, y)$) for the cellulose Iβ wide 110 surface. The maps are calculated for different overlapping z-regions above or below the surface, z-values are given in the figure heads. The color code ranges from −5 kJ/mol (blue) to 5 kJ/mol (red). See legend on page 84 for the color-coding.

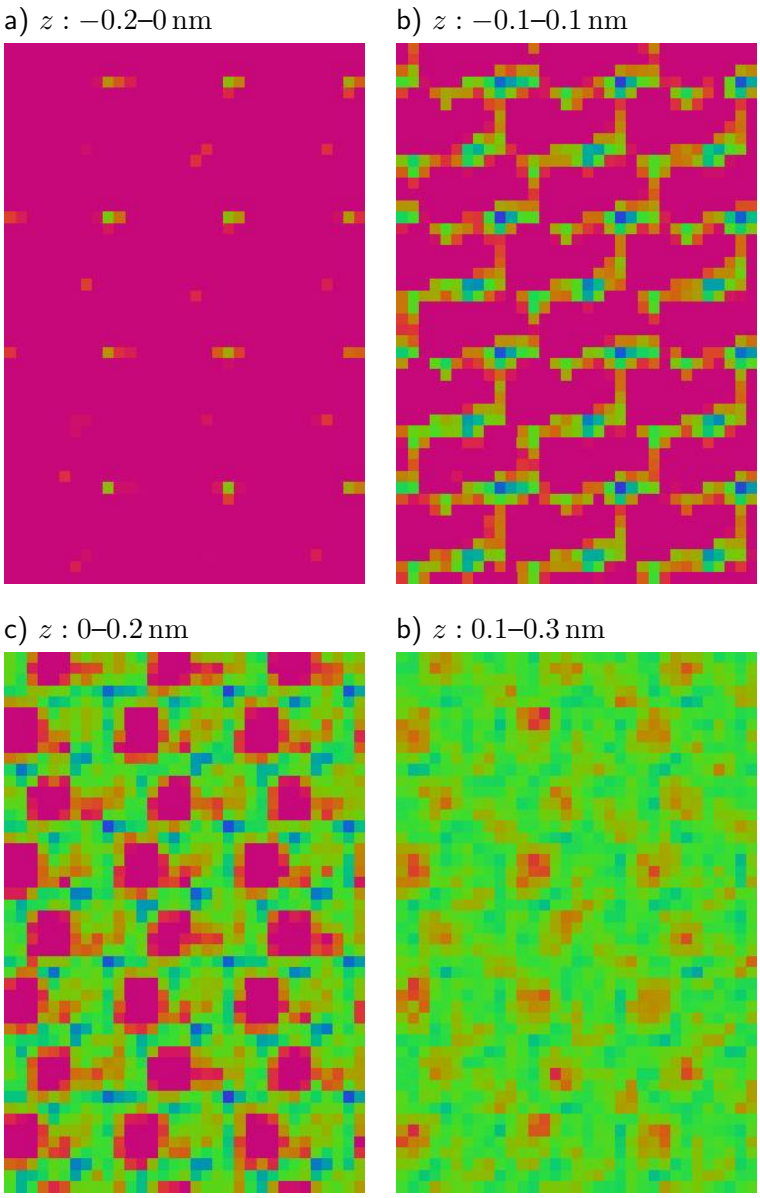


Figure 4.27: Hydrophilicity/hydrophobicity maps ($\Delta\mu_{H_2O}(x, y)$) for the cellulose Iβ narrow $1\bar{1}0$ surface. The maps are calculated for different overlapping z-heights above or below the surface, z-values are given in the figure heads.

The structure of the crystal is responsible for the argon chemical potential $\Delta\mu_{\mathrm{Ar}}(x, y)$ in figures 4.28 a and b. Below the surface (−0.2 nm–0 nm) there is hardly any structure, but only some local cavities, which are favorable for argon. In the layer above (−0.1 nm–0.1 nm), there is a strong excluded-volume structure, leading to alternating horizontal (along the x-axis) grooves with negative and positive chemical potential differences. In the surface layer (z from 0 nm to 0.2 nm), the blue patches grow larger and there are disconnected red and blue islands. In the solvent layer (0.1 nm–0.3 nm), there are only few isolated small red patches left, large regions are favorable for argon in this layer. Going further (z from 0.3 nm to 0.5 nm) into the water solvent (figure not shown), the pattern is washed out. The blue "negative" patches shrink and $\Delta\mu_{\mathrm{Ar}}(x, y)$ takes the value for bulk water (figure not shown).

The situation on the narrow $1\bar{1}0$ surface (figure 4.29) is the same as described for argon $\Delta\mu_{\mathrm{Ar}}(x, y)$ on the 110 surface. In the crystal, below the surface the chemical potential $\Delta\mu_{\mathrm{Ar}}(x, y)$ is greater than 5 kJ/mol, except on some single spots (figure 4.29 a). Pronounced argon-densities (blue) run as horizontal bars (along the x-axis) from left to right, about one third of all grid points have a negative argon chemical potential. On the surface, the pattern looks much alike that one for the wide surface, but with fewer negative regions, which is reflected by the z-dependence of $\Delta\mu_(z)$ as shown in figure 4.18. However, structure is lost if we move away from the surface: at a distance of 0.1–0.3 nm the pattern of the $\Delta\mu_{\mathrm{Ar}}(x, y)$ map has only some red spots left, while green ($\Delta\mu_{\mathrm{Ar}}(x, y) = 0$) gets the dominant color for larger z-values (not shown).

On both surfaces, the chemical potential of argon μ_{Ar} is more structured than for water $\mu_{\mathrm{H_2O}}$. Additionally the argon structure extend more into the water bulk than the water structure does. This is explained by a water-structure imposed from the surface, which does not alter the water density from the bulk value, but creates free volume to accommodate argon. The existence of a surface-induced water order, which is not visible in the density profile or chemical potential is validated in the next section (4.5.3) by the analysis of the water-orientation. However, as water itself is able to accommodate small lipophiles [42, 103], there is always some argon structure, even if the interfacial water is much like bulk water.

In summary, both interfaces of cellulose Iβ share common features. Both are stable with respect to surface reconstruction and water penetration, at least on the nanosecond timescale and in the absence of defects. Both show essentially the crystal structure of cellulose Iβ, also in the adsorption patterns for water and argon. In spite of the presence of OH groups, both surfaces are non-hydrophilic and lipophilic: Water molecules are not attracted by the surface, as they are equally happy in bulk water. On the other hand, the presence of the surface perturbs the water structure sufficiently to create free volume, in which an argon atom can dissolve more easily than in denser bulk water. The wider spacing of cellulose chains allows the 110 surface to expose more of the hydrophobic grooves where lipophilic adsorption takes place. In contrast, the protruding OH groups are less affected by chain packing, so the behavior towards water is similar for both surfaces.

4.5.3 Structure of Water close to the Surface

Up to this section, the water-cellulose interplay has been characterized only by the chemical potential, deduced from the density, which does not tell us anything about the structure of the solvent near the surface.

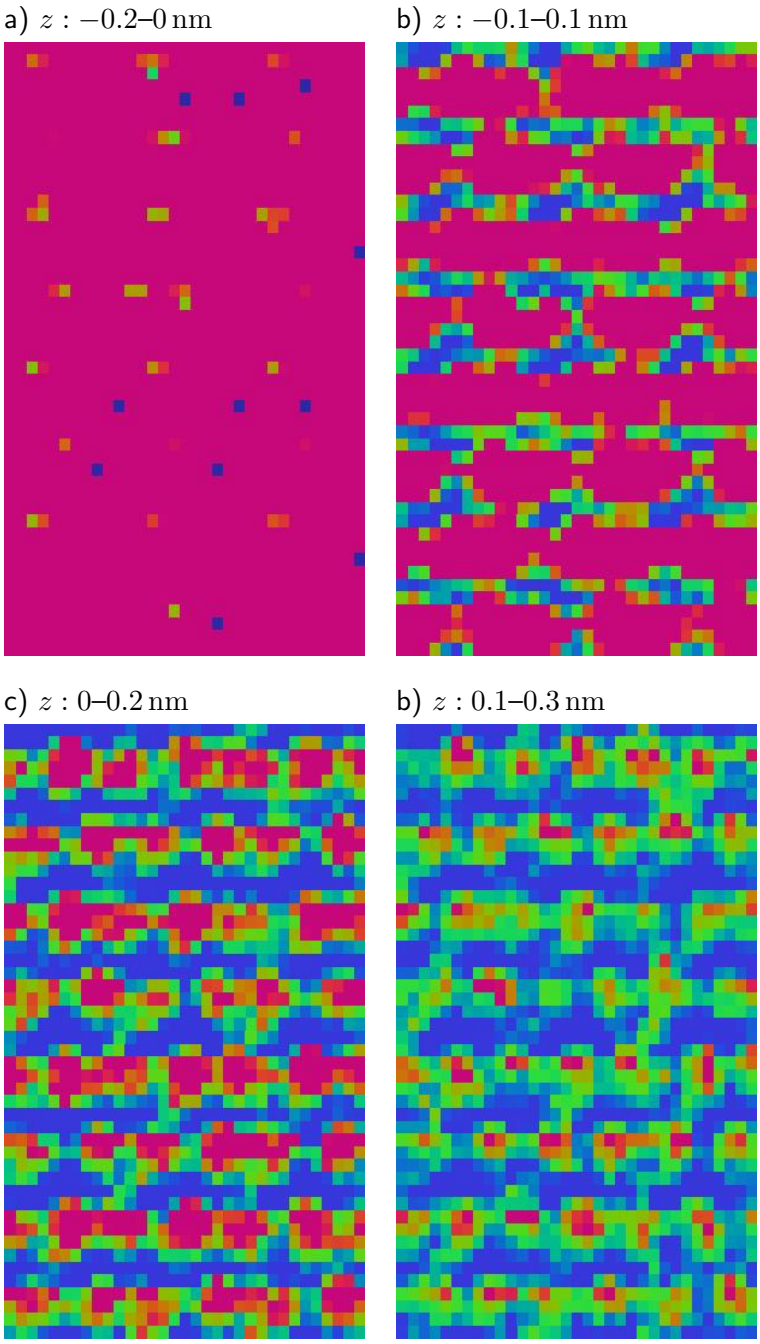


Figure 4.28: Lipophilicity/lipophobicity maps $\Delta\mu_{\mathrm{Ar}}(x, y))$ for the cellulose Iβ wide 110 surface. The maps are calculated for different overlapping z-regions above or below the surface, z-values are given in the figure heads.

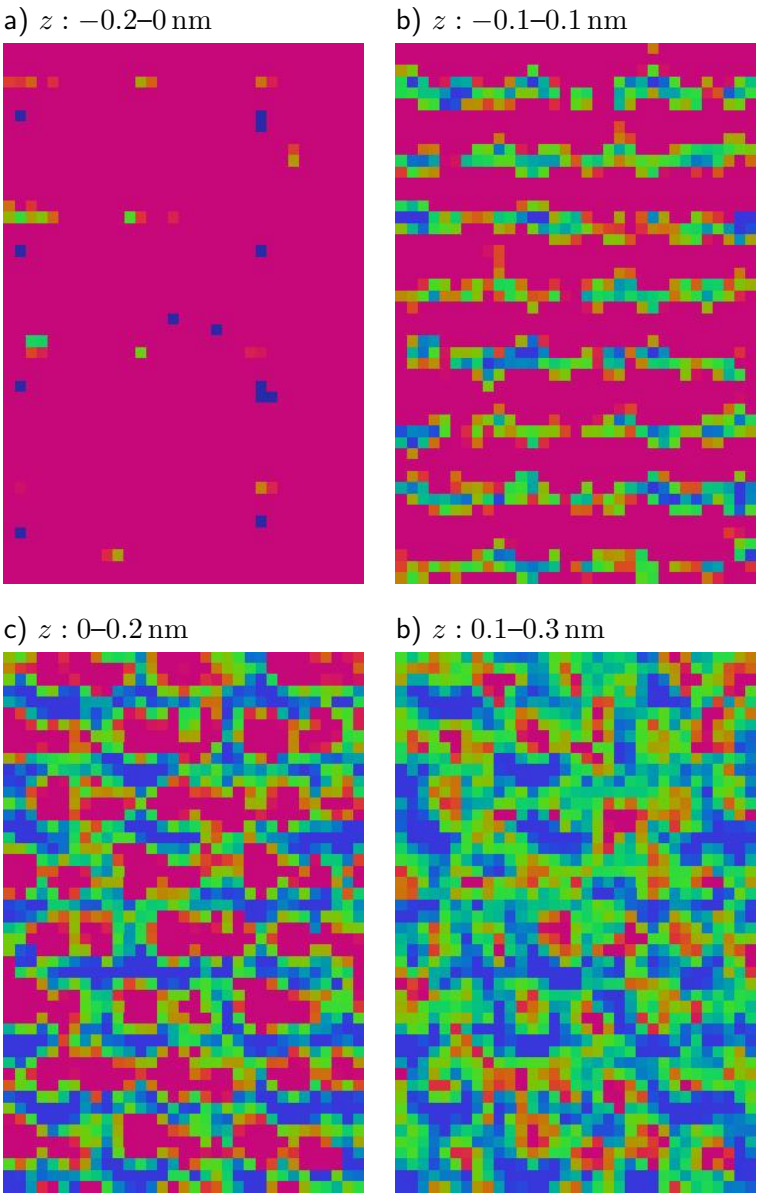


Figure 4.29: Lipophilicity/lipophobicity maps ($\Delta\mu_{\mathrm{Ar}}(x, y)$) for the cellulose I$\beta$ narrow $1\bar{1}0$ surface. The maps are calculated for different overlapping z-regions above or below the surface, z-values are given in the figure heads.

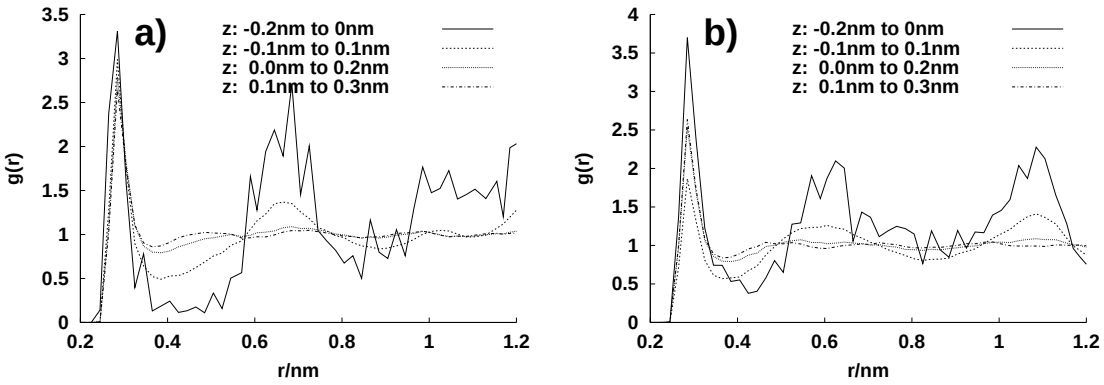


Figure 4.30: Radial distribution of water oxygens in different slabs over above the a) wide 110 surface and b) the narrow $1\bar{1}0$ surface. Only water atoms in the respective slab are accounted for.

4.5.3.1 Water-Water Radial Distribution Functions

Radial distribution functions unveil the intermolecular relationship between water molecules in different slabs above the surface (figure 4.30). Below the surface ($z = -0.2\,\text{nm}$ to $0.0\,\text{nm}$) the water structure surface is equal up to the first minimum. Differences are only found in the second and third peak's structure. On the wider 110 surface the water has more structure with deeper minima and sharper maxima. There is one additional peak centered around 1.25 nm which is absent for the $1\bar{1}0$ surface. The structure is transferred from the lower to to the upper water layer (z from -0.1 to 0.1) as well, but is no longer present in the surface layer (z from 0.0 nm to 0.2) and above. From the size of the water-patches (low chemical potential of water) in figure 4.26 the peak at 0.65 nm for the 110-surface (figure 4.30 a) is due to water molecules within one patch: the distance between regions with low chemical potential in figure 4.26 is about 1 nm. Therefore, the two water-peaks on the $1\bar{1}0$ surface at 0.6 nm and 1.1 nm correspond to water molecules within one patch and between two patches.

The structure close to the surface is probably due to the distance between polar sites on the surface, where water can interact with. The structure enhancement due to hydrophobic surfaces [40] is found in the water layer above the surface, where for example the second peak is of the OW-OW radial distribution function is stronger than for bulk water (not shown). This holds true for both – wide and narrow – surfaces.

The number of water-water hydrogen-bonds is given in table 4.11. "Below" the water-cellulose interface ($z < 0$), there are only few water molecules at all, and so there are hardly any water-water hydrogen-bonds inside a layer, but still some bonds to connect water molecules to the bulk. There are no "isolated" water molecules, which are totally saturated by cellulose H-bonds. However, these water molecules have about 70% of the bonds of a bulk water molecule. Differences between the two surface are minute for negative and small positive z-values, but the wide 110 surface has about 5% more bonds than the $1\bar{1}0$ surface. As the separation to the surface becomes larger, the differences between the two surfaces flattens, for both there are on average 3.1 water-water hydrogen bonds.

a) 110					b) $1\bar{1}0$				
z_l [nm]	z_u [nm]	total	intra	inter	z_l [nm]	z_u [nm]	total	intra	inter
-0.2	-0.0	1.7	0.0	1.7	-0.20	0.00	1.6	0.0	1.5
-0.1	0.1	2.1	0.2	1.8	-0.10	0.10	1.9	0.2	1.7
-0.0	0.2	2.1	0.6	1.4	0.00	0.20	2.0	0.6	1.4
0.1	0.3	2.3	0.8	1.5	0.10	0.30	2.5	0.8	1.7
0.2	0.4	2.9	0.7	2.2	0.20	0.40	3.0	0.7	2.3
0.3	0.5	3.1	0.6	2.5	0.30	0.50	3.1	0.7	2.4

Table 4.11: Number of water-water hydrogen bonds per water oxygen in different slabs. a) 110 surface, b) $1\bar{1}0$ surface. The height above the surface is given in the first two columns (lower z_l and upper value z_u), columns three to five state the absolute number of water-water hydrogen bonds per oxygen (total) which is the sum of two terms. First the bonds inside (intra) a slab and second bonds from one slab to another one (inter).

4.5.3.2 Orientational Order of Water Near the Surface

Distribution functions of the HH-, dipole- and normal-vector Legendre polynome $\cos(\theta)$ show the orientation of water molecules near the surface in figure 4.31 and 4.32. The distribution functions are defined as the probability of the cosine of the angle θ between the corresponding molecular vector and the surface normal. A random distribution of water molecules would show up as a flat horizontal (along the x-axis) line. Again, different layers are defined (see legends).

Water molecules close to both (narrow and wide) surfaces own one broad peak, with a preferential orientation where the HH- and dipole vector align parallel to the surface, the normal-vector of water ($\overline{\mathrm{OH}}_1 \times \overline{\mathrm{OH}}_2$) enclosing an angle of 0° or 180° (figure 4.31) with the surface normale. Many water molecules lie "flat" on the surface, avoiding water donor hydrogen-bonds with cellulose (while others direct one hydrogen towards cellulose, building a hydrogen bond). The distribution-maximum for the HH- and dipole-vector is not exactly zero, but approximately 0.1, which corresponds to 84°, this is very close to the angle of 87° found by Grigera [40] for a model-hydrophobic surface.

At larger separation from the surface the distributions are almost constant, but with a negative slope from the left to the right (z: 0.4 nm–0.6 nm) in (fig 4.31 a), the HH- and dipole-vector exhibit "long"-range surface-induced structure, which is not found for the normal-vector (figure 4.31 b). The distance of 0.6 nm corresponds to about two water layers, so the ordering is not a surprise; one can expect the ordering to be preserved up to the second or even third hydration shell of the surface [40].

Even if the $\cos(\theta)$ distributions are very similar for both surfaces, some details are different. The dipole distribution for the wide surface (figure 4.31 a) tend to zero for $\cos(\theta) = -1$, whereas they own a sub-maximum figure 4.32 a/b for the other surface. There are some water molecules, which direct hydrogen atoms to the surface, forming probably H-bonds.

In summary the water structure near the surface is perturbed by the presence of the hydrophobic surface, repelling the solvent. This leads to some structure-enhancement for water molecules in the vicinity of cellulose and the effect is washed out at larger

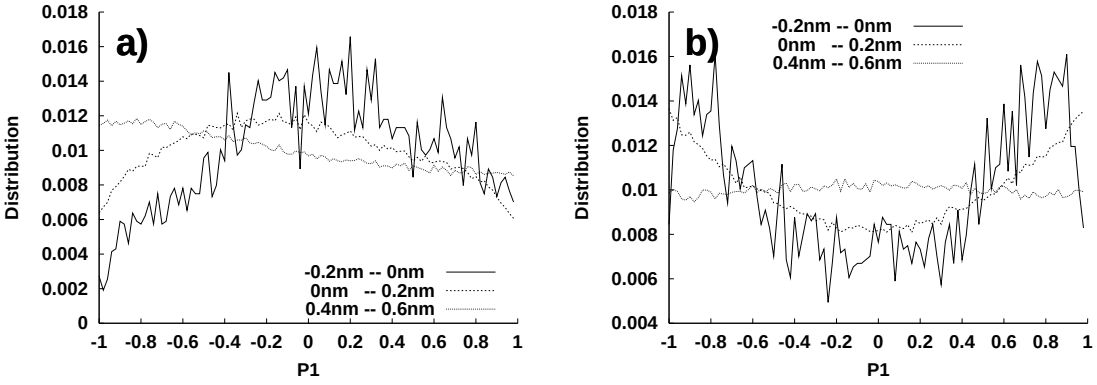


Figure 4.31: Distribution functions of the the a) dipole- and b) normal vector Legendre polynome $\cos(\theta)$ of water molecules on the 110 (wide) surface. The HH-vector has a distribution almost identical to the dipole vector and is not shown. (The dipole vector is defined that it points along the bisecting line of the HOH angle.)

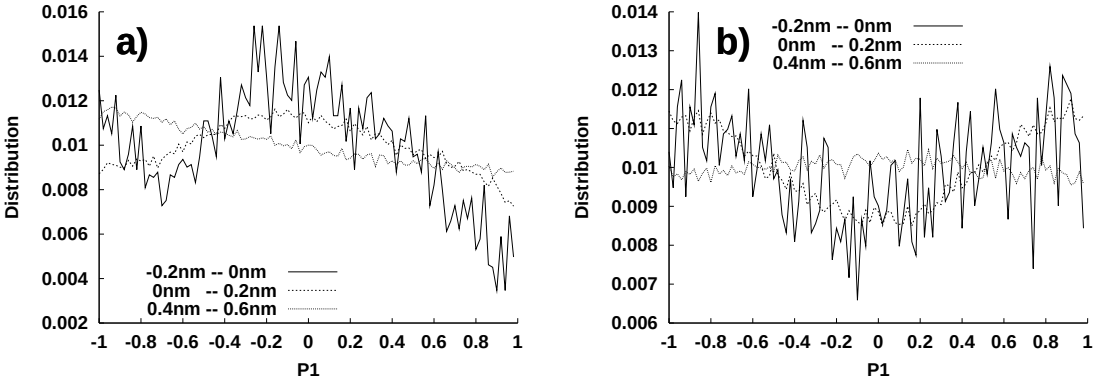


Figure 4.32: Distribution functions of the the a) dipole- and b) normal vector Legendre polynome $\cos(\theta)$ of water molecules on the $1\bar{1}0$ (narrow) surface. The HH-vector has a distribution almost identical to the dipole vector and is not shown. (The dipole vector is defined to point along the bisecting line of the HOH angle.)

distances ($z > 0.6\,\mathrm{nm}$). The water solvent close to the surface takes a typical structure for water in the presence of hydrophobic walls (see [40]). This structure is characterized by at strong hydrophobic ordering of the water molecules, which allows to accomodate small hydrophilic molecules into cavities. The water structure (orientational distributions, water-water radial distribution functions) are typical for hydrophobic ordered water (ref. [41, 42, 103] and references therein). From the orientational distributions $\cos(\theta)$ both, the wide and narrow, surface are similar with only minor differences. The narrow surface has not as much structure from the OW-OW rdf and the orientational distribution gives some evidence for water-molecules directed towards the surface.

5 Polyelectrolyte-Cellulose-Surface Systems

In the last two chapters we have dealt with solutions of polyelectrolytes and cellulose in water, which are now put together to a three component-system: cellulose, polyelectrolyte and solvent.

In experiment, sodium (carboxy methyl) cellulose (CMC) is found to adsorb on cellulose, whereas sodium poly(acrylic acid) (PAA) does not [43]. This difference in adsorption behavior now becomes evident from the direct comparison by molecular dynamics simulations of four systems:

- wide (110) cellulose Iβ surface with CMC-polyelectrolyte,
- wide cellulose surface with PAA,
- narrow (1$\bar{1}$0) surface with CMC and
- narrow cellulose surface with PAA-polyelectrolyte.

As a comparative study of both cellulose surfaces with both polyelectrolytes is desirable, we have simulated all four systems. Unfortunately in both the 1$\bar{1}$0-CMC and 1$\bar{1}$0-PAA simulations the 1$\bar{1}$0-cellulose crystal is not stable: After 1 ns of equilibration the cellulose crystal reconstructs. For this reason, only the 110 cellulose surface with PAA and CMC adsorbate are discussed here. The two 1$\bar{1}$0 systems are addressed in section 5.5, which describes and explains the rupture of the crystal. The affinity of a given CMC- or PAA-oligomer can now directly be monitored in the simulation.

Simulation details for both 110-simulations are given in section 2.4 in chapter "Computational Details". Both simulations have run-times from 4 ns and 5 ns, after 1 ns of equilibration each. The CMC-oligomers are named trimer #1, tetramer #1 and trimer #2 and tetramer #2. PAA pentamers are numbered from 1 to 6.

The stability of the cellulose crystal against solvation or penetrating water is checked by the cellulose and water densities in z-direction in figure 5.1. Whereas there is good agreement of the density profiles along the z-axis between the cellulose-water and 110-cellulose-water-CMC simulations (figure 5.1 a), this is not exactly true for the 110-PAA system. The 110 crystal simulations have the same density profile which is found for the cellulose-water systems in chapter 4. The GROMOS-modeled cellulose-110-surface is stable on a multi-nanosecond timescale. The presence of an additional polyelectrolyte-component does not influence the stability of the crystal and the transformation of the unit-cell from monoclinic to orthorhombic (see section 2.3) does not change this as well on a multi-nanosecond timescale.

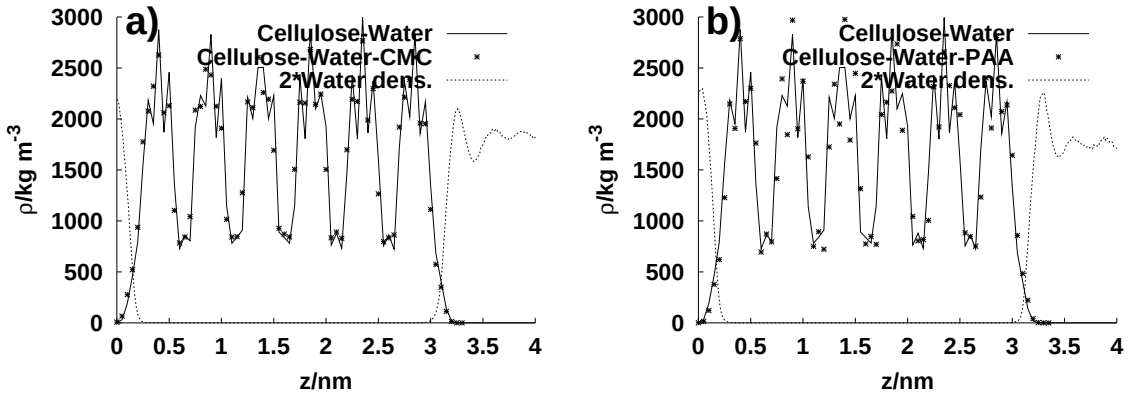


Figure 5.1: Comparison of two cellulose density profiles for a) the 110-CMC and b) the 110-PAA system. There is hardly any deviation between the system with and without the CMC-polyelectrolyte. In subfigure b) 110-PAA: there is a difference, but somehow located in the interior of the crystal, whereas the edges towards water are equal to that of the water-crystal simulations. The difference is most likely due to different statistics in single histogram bins. The water densities are taken from the polyelectrolyte-110-surface simulations. The surface is not penetrated by water.

5.1 Adsorption Behavior of CMC and PAA on Cellulose Surfaces

A first impression of the adsorption of a polyelectrolyte is gained from the time evolution run of the z-coordinate of the oligomer-center of mass (figure 5.2).

For the wide 110 surface, all four CMC-oligomers adsorb on either one of the surfaces, but with different averages of the z-value of the center of mass. One oligomer (trimer #1) sticks to its initial state close to the surface, which is taken during the equilibration. All other CMC oligomers have first a diffusive trail of 1.5 ns, where they slowly approach to the surface.

Center of mass z-trajectories for PAA-oligomers are shown in figure 5.2 b. Two oligomers (#2 and #4) move towards the surface and keep their z-positions at about 0.45 nm and 0.75 nm. All others PAA-pentamers (including #3 and #6, which are not shown) stay away from the surface. Both CMC and PAA molecules adsorb, but PAA with fewer oligomers than CMC.

Complete xyz-trajectories of two selected CMC-oligomers' centers of mass are shown in figure 5.3. Trimer #1 is fixed to the surface, there is almost no center of mass translation in z-direction. The center of mass diffusion coefficients are given in table 5.1 and plotted in figure 5.4. The diffusion coefficient is calculated on the basis of coordinates, which are relative to the center of mass of the aqueous phase including polyelectrolyte and ions. From the CMC-oligomers trimer #1 is the slowest, while other molecules are faster. In general the larger tetramers are slower than the trimer. If a CMC molecule approaches the surface it slows down, because of interaction with the surface and the loss of one degree of freedom along the surface normal, as it shown in figure 5.4 a. While for trimer #1 the diffusion coefficients in x- and y-direction are almost equal, the diffusion-coefficient in

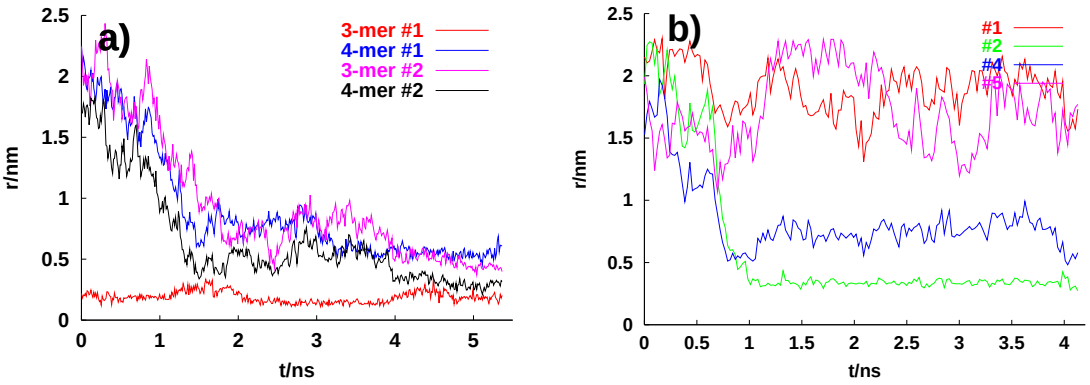


Figure 5.2: a) Distance between the "centers" of mass of all CMC oligomers and the closest surface of the 110-crystal: trimer #1 adsorbs onto the upper crystal-surface (at 3.125 nm), all other three molecules adsorb to the lower surface (at $z = 0.1$ nm).

b) PAA center of mass to surface distance in z-direction for selected oligomers (#1, #2, #4 and #5). Oligomers #3 and #6 (not shown) are like #1 or #5. One oligomer adsorbs on the surface, but all other solute stay in the bulk. Oligomer #4 (blue line) approaches the surface up to 0.6 nm, but never has any hydrogen bond with the surface. Every 40th data point is shown.

z-direction is usually smaller by a factor of 40:

$$D_{xx} = 7.0 \cdot 10^{-7}, D_{yy} = 6.8 \cdot 10^{-7}, D_{zz} = 0.2 \cdot 10^{-7} \,/\mathrm{cm}^2\,\mathrm{s}^{-1}$$

This ratio is not as strong for other CMC oligomers or PAA, which is shown for trimer #2 (after adsorption at 2.0 ns) and PAA-pentamer #1 (total simulation time), where the difference between lateral and the diffusion along the z-axis is a factor of two for CMC trimer #2 :

$$\begin{aligned} \text{CMC trimer \#2:} \quad & D_{xx} = 9.3 \cdot 10^{-7}, D_{yy} = 9.6 \cdot 10^{-7}, D_{zz} = 3.1 \cdot 10^{-7} \,/\mathrm{cm}^2\,\mathrm{s}^{-1} \\ \text{PAA pentamer \#1:} \quad & D_{xx} = 3.5 \cdot 10^{-7}, D_{yy} = 3.6 \cdot 10^{-7}, D_{zz} = 4.0 \cdot 10^{-7} \,/\mathrm{cm}^2\,\mathrm{s}^{-1} \end{aligned}$$

As it becomes evident, all CMC-oligomers, except trimer #1 are fixed to the surface in a loose fashion which still allows for some fluctuation of the oligomer center of mass. The diffusion coefficients decrease in time (see table 5.1), the longer a CMC-molecule is close to the surface, the slower it gets. This indicates that the adsorption is a slow process and CMC oligomers are not equilibrated on the surface after some nanoseconds of simulation.

The PAA diffusion coefficients are of the same order of magnitude as for CMC, but they do not become smaller with simulation time, so there is no hint that PAA adsorbs on the crystal surface in a very slow process, except for pentamer #2, which does not exhibit a significant diffusion along the z-axis (figure 5.4 c).

In short, CMC molecules are attracted by the cellulose surface. Oligomers move from the water bulk to the surface, where they adsorb. There is no case of desorption. In contrast most PAA molecules stay solved in the bulk. However there is one PAA pentamer, which adsorbs, and a second one which is in close distance to the surface. Once adsorbed, the PAA molecule keeps the position close to surface.

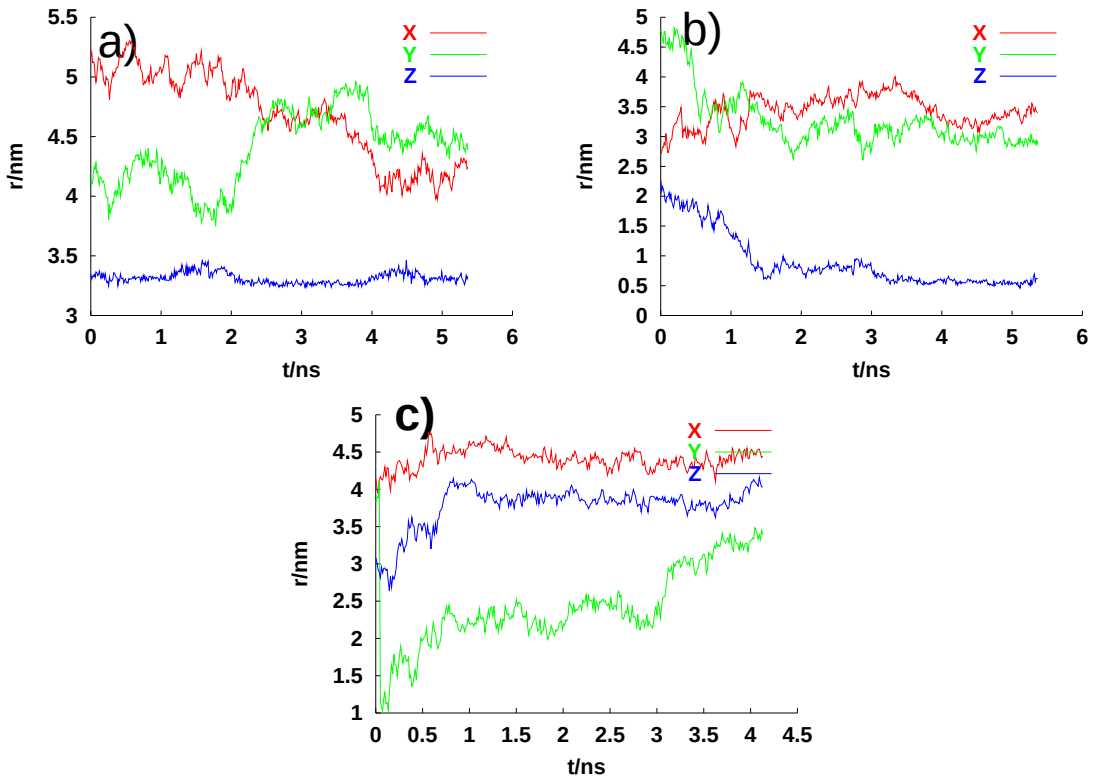


Figure 5.3: Components (x, y and z) of the center of mass of a) trimer #1, b) tetramer #1 of the CMC-110 surface-simulation and c) pentamer #4 of the PAA-110 simulation. The tetramer #2 plot is representative for other oligomers. (x is the the coordinate perpendicular to the cellulose chain direction, y is parallel, z is directed along the surface normalen vector). Only every 20th data point is shown.

5.1.1 Distribution of Monomers along the surface normal (z-coordinate)

A distribution for the z-distance of anhydroglucose unit's center of mass is given in figure 5.5. The center of mass is calculated using the six atoms of the glucose ring. The density of glucose-repeat units is much higher near the cellulose-surface than in the water-bulk. This is most obvious for trimer #1, which adsorbs almost flat on the surface. The CMC-density is zero for distances from the surface larger 0.4 nm. All three repeat units of trimer #1 own the same small distance to the surface and overlap in z-direction.

The other three oligomers (tetramer #1, trimer #2, tetramer #2) are more or less perpendicular or inclined with respect to the surface: there is a moderate overlap of the anhydroglucose repeat units. The distance between the peak-maxima is about 0.35 nm along the system's z-axis. So this is a "end on the surface" configuration, where probably only one AGU is in contact with cellulose. The shape of all AGU-density peaks is characteristic: The peak closest to the surface is sharp with a high intensity, as the respective AGU is (hydrogen-)bonded to some sites on the crystal surface, whereas the other repeat units protrude into the water bulk. If the peak of the closest AGU is below 0.15 nm it is very likely to find the glucose ring adsorbed flat on the surface, the ring-plane parallel to the surface.

CMC-oligomers	simulation time [ns]					
	0-1	1-2	2-3	3-4	4-5	0-5
Trimer #1	2.20	1.60	1.22	3.21	1.18	1.88
Tetramer #1	2.62	2.58	1.96	2.52	2.19	2.37
Trimer #2	3.72	3.59	2.66	2.96	2.97	3.18
Tetramer #2	2.58	3.21	2.32	2.80	1.82	2.55
PAA-pentamers	**simulation time [ns]**					
	0-1	1-2	2-3	3-4		0-4
Pentamer #1	6.28	3.27	2.32	2.93		3.70
Pentamer #2	4.60	3.73	1.39	1.51		2.81
Pentamer #3	5.22	3.31	2.70	2.35		3.40
Pentamer #4	3.27	2.07	1.60	1.20		2.04
Pentamer #5	4.73	2.76	3.29	3.59		3.59
Pentamer #6	3.30	2.05	1.98	1.40		2.18

Table 5.1: Diffusion coefficients (in $10^{-6}\mathrm{cm}^2\,\mathrm{s}^{-1}$) of CMC and PAA oligomers. The values are listed for different time-regimes of each 1 ns length. The last column states the average value for the whole simulation.

Only CMC-trimer #2 (figure 5.5 c) has its closest peak's intensity at a larger separation from the crystal beyond 0.3 nm, which is still close enough to form hydrogen bonds with the cellulose surface.

An example plot of the z-distribution is given for PAA-pentamer #1 in figure 5.6 a). The center-of-mass distribution for almost all of the repeat units is centered at a distance of about 1.8 nm away from the surface, which is in the middle of the water bulk. The half-width of the distribution is twice as large as for CMC-repeat units. PAA pentamers stay solvated in the water bulk and in general do not adsorb on the surface. There is one exception: PAA pentamer #2 is close to the surface, all repeat units are less than 0.5 nm away from the surface.

5.1.2 Orientation of Anhydroglucose Units on the Surface

The orientation of glucose rings of CMC can be evaluated by the vector $\mathbf{u}$, which is defined as the mean of the normals of the two planes defined by C1-C3-C5 and C2-C4-O5 (see figure 5.7 and ref. [48]):

$$\mathbf{u} = \mathrm{norm}\left(\mathrm{norm}((\mathbf{C5}-\mathbf{C1})\times(\mathbf{C3}-\mathbf{C1})) + \mathrm{norm}((\mathbf{C4}-\mathbf{O5})\times(\mathbf{C2}-\mathbf{O5}))\right).$$

Where $\mathbf{C1}$,... represent the position vectors of the respective atoms. The "norm"-function returns a unit-vector of its argument. The angle-distributions of vector $\mathbf{u}$ with the surface normale $\mathbf{n}$ are shown in figure 5.8. From this definition, $\cos(\phi) = \mathbf{n}\cdot\mathbf{u} = 1$ means that **C1**, **C3**, and **C5** point toward the surface (chair conformation of AGU), $\cos(\phi) = -1$ means that **C2**, **C5** and **O5** point toward the surface. Trimer #1 (figure 5.8 a), is flat on the surface, $\mathbf{u}$ and surface normal $\mathbf{n}$ enclose an angle of either 0° or 180° ($\cos(\phi) = \pm 1$). There is a small local maximum for the middle anhydroglucose unit at $\cos(\phi) = 0$ (AGU

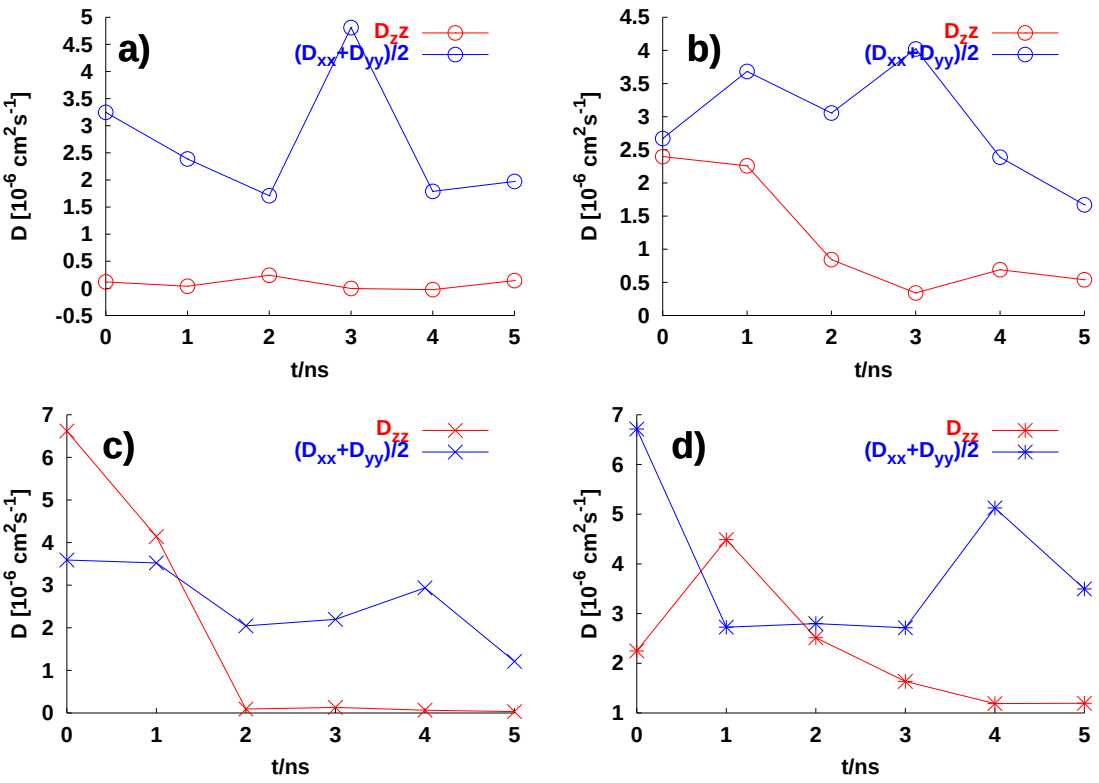


Figure 5.4: Plot of the diffusion-coefficients along the xy-plane ($(D_{xx} + D_{yy})/2$) and along the surface normalen z (D_{zz}) of a) CMC trimer #1 b) tetramer #2 and c) PAA pentamer #1 and d) PAA pentamer #2 against the simulation time in ps. The diffusion coefficients are averages over one nanosecond each.

#2). Because the cosine-function is symmetric, there is only one maximum for alternating rings of a cellobiose unit: Glucose rings of the adsorbed CMC oligomer are rotated by 180° in an alternating fashion and subsequent AGUs (1 and 2 or 3 and 4 of an oligomer) have distributions symmetric to $\cos(\phi) = 0$. As an approximation, this holds true for both tetramers as well. For trimer #2 two glucose rings exhibit a maximum at $\cos(\phi) = \pm 0.7$ ($\phi \approx 135°$). For tetramer #1 the repeat unit #1 is closest to the surface and exhibits a broad $\cos(\phi)$-distribution (figure 5.8 b) with a maximum at 0.75. AGU #2 is parallel to the surface, but its center of mass is about 0.5 nm away from the cellulose surface. Other oligomers like tetramer #2 have the maximum for the surface-closest AGU at $\cos(\phi) \approx= 1$, too.

On adsorption some but not all anhydroglucose unit's orientation changes. While in solution angle distributions are almost featureless, there is some preferential AGU orientation on the surface. While trimer #1 (see figure 5.9) repeat units are always flat on the cellulose surface, other glucose-rings are tilted to the surface-normal. They are aligned parallel to the 200-planes of the crystal, the glucose-ring **C1**-**C4**-vector perpendicular to the crystal chain-direction (see figure 5.10).

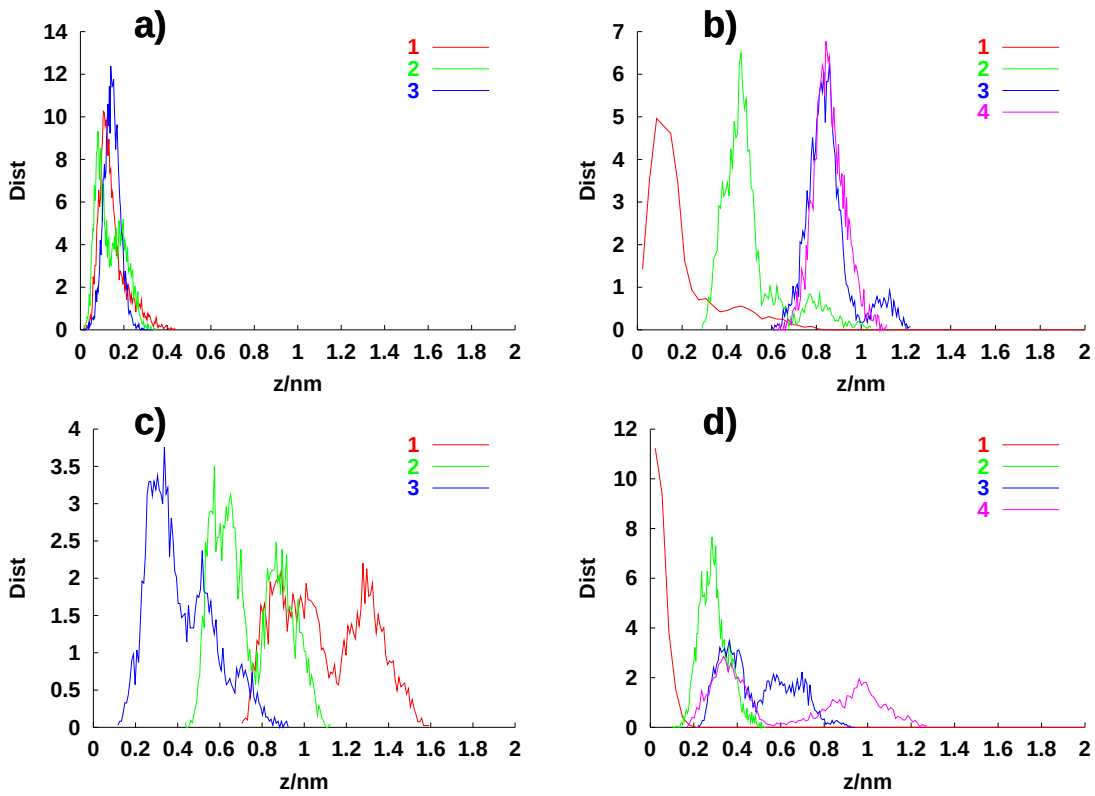


Figure 5.5: Distribution of the single distance anhydroglucose units' center of mass to the 110 surface. a) trimer #1, b) tetramer #1, c) trimer, #2 d) tetramer #2. The numbers in the legend indicate AGU numbers within the oligomers. AGUs are numbered along the oligomer-strand starting at the non-reductive end. The histograms are calculated for adsorbed states of CMC only.

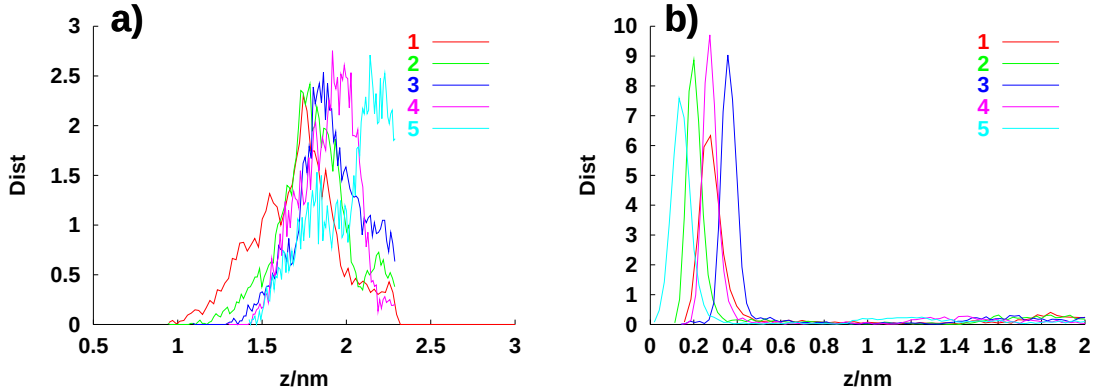


Figure 5.6: a) Example repeat unit z-distribution for PAA-pentamer #1. No one of the five repeat units approach the surface and there is no special orientation of the PAA pentamer is visible. b) distribution of pentamer #2 which is adsorbed on the surface. The molecule is almost flat on the surface. The numbers in the legends refer to the indices of the acrylate repeat units.

Figure 5.7: Sketch of a glucose ring with the C1-C3-C5 plane drawn in. The second plane C2-C2-O5 is not shown, but the respective vector points along the shown one. The orientation of the glucose ring is determined by the mean of both vectors.

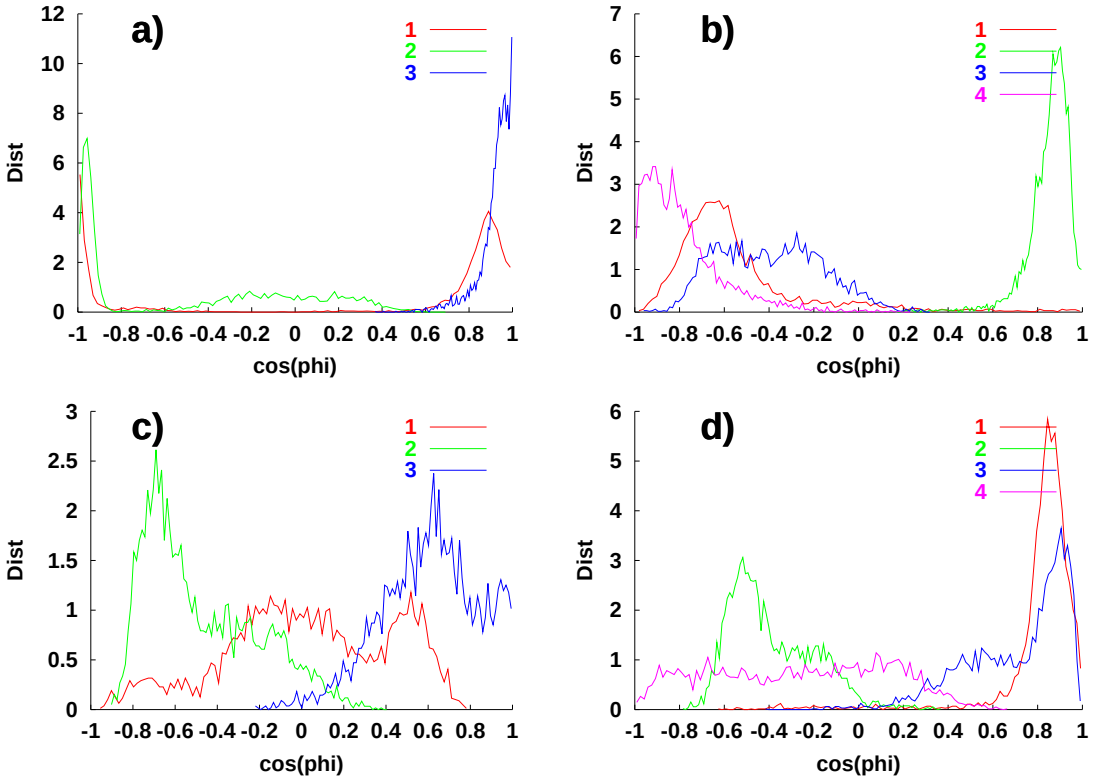


Figure 5.8: Distribution of angle between the surface normal and the vector **u** for all anhydroglucose units of CMC. **u** is defined by the atoms of a glucose ring in such a way, that it is perpendicular to an "imaginary" ring-plane (see text on page 113. a) trimer #1 b) tetramer #1 c) trimer #2 and d) tetramer #2. The histograms are calculated for adsorbed states of CMC (when there is at least one hydrogen bond between the oligomer and the surface).

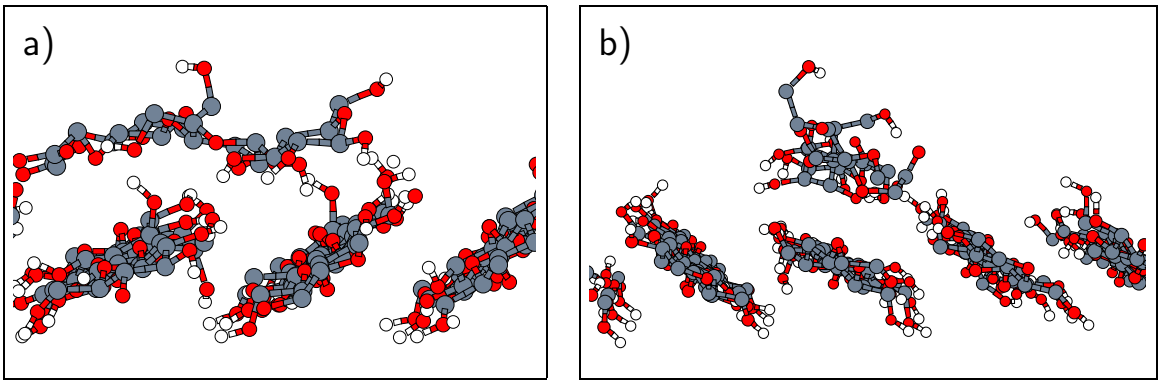


Figure 5.9: Snapshots of trimer #1 adsorbed on the cellulose surface. a) Trimer #1 on the surface at 3 ns. One looks along the cellulose chains of the surface (bottom), the trimer is almost perpendicular to the chain-direction. b) as subfigure a), but $t = 4.5$ ns. The trimer is now parallel to the chain direction, the point of adsorption is located over one cellulose-chain of the crystal.

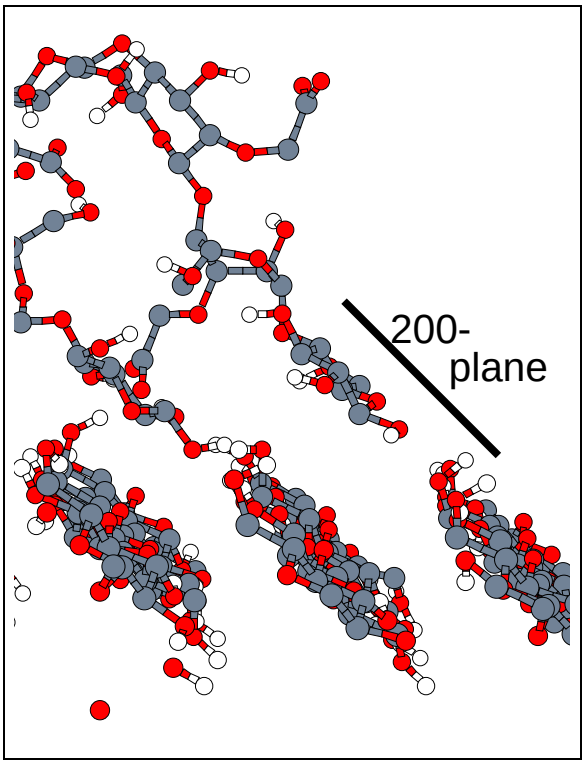


Figure 5.10: Snapshot of tetramer #1 and tetramer #2 on the surface at $t = 4.5$ ns. There is only one glucose ring of each CMC-molecule close to the surface. The contacting AGU takes a position, which is aligned parallel to the 200-plane (parallel to the cellulose-rings of the crystal; AGU 1 tetramer #1 in the front, AGU 1 of tetramer #2 in the back).

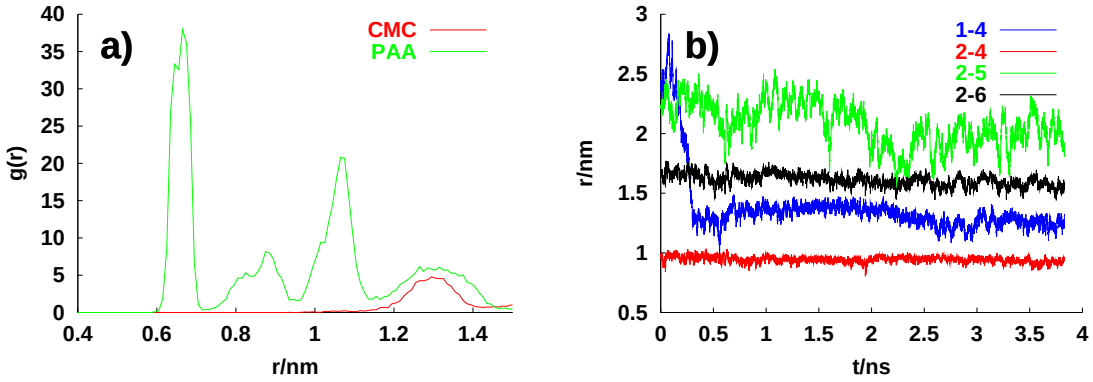


Figure 5.11: a) Center of mass-center of mass radial distribution functions for CMC and PAA oligomers (mean over all oligomers) of the polyelectrolyte-110-cellulose simulations, average using whole simulation.
b) Distance between selected poly(acrylic acid) pentamers from the 110-PAA simulation as a function of time. The plot shows the center of mass-center of mass distance of two PAA-pentamers. Examples for pentamer-pairs are shown, which attract or keep some distance.

5.2 Aggregation of CMC and PAA oligomers

Contrary to the expectation for polyelectrolyte molecules, the oligomers in the water bulk of the cellulose-water-polyelectrolyte simulations attract each other. This is easy to see from the center of mass-center radial distributions for CMC/PAA molecules or the center of mass-distance of poly(acrylate) oligomers in figure 5.11 a) and b). In figure 5.11 b) there are three cases for either a dynamic process of aggregation (pentamer pair 1-4), a persistent aggregation (pair 2-4) and a pair with constant center of mass-separation (2-5 or 2-6).

The PAA pentamers come much closer than CMC oligomers, which keep a distance larger one nanometer.

This behavior is in contradiction to the stretched conformation found for the PAA 23-mer in chapter 3: The short poly(acrylic acid) pentamers attract in aqueous solution. In figure 5.11 the center of mass distance is shown as a function of time. While there are pairs with a continuous short distance like 2-4 and (not shown in figure 5.11) 3-5, 3-6, 4-6, other pairs have a larger separation after the equilibration phase which becomes smaller during the first 0.5 ns of the simulation (like 1-4 and (not shown) 1-5, 1-3). Pentamer #2 does not move close to other oligomers and keeps separated from them. The electrostatic repulsion of the carboxylic groups is screened by the counterions: Almost all sodium ions are close the aggregated polyelectrolytes (rdf not shown): Every carboxylic oxygen has on average 1.0 counterions within a distance of 0.25 nm. In chapter 3.2.4 we found 1.5 sodium counterions in the vicinity < 0.65 nm of the PAA 23mer. The distance between the PAA center of mass and the carboxylic oxygens is approximately 0.25 nm. Counterions are much closer to the aggregated PAA pentamers, because of end-effects which lead to a higher local charge density (there are 10 polymer ends instead of only two of the 23mer). High surface charge density is produces a high contact density of counterions [62, 130]. The second reason is the high concentration of counterions in the simulation box, just

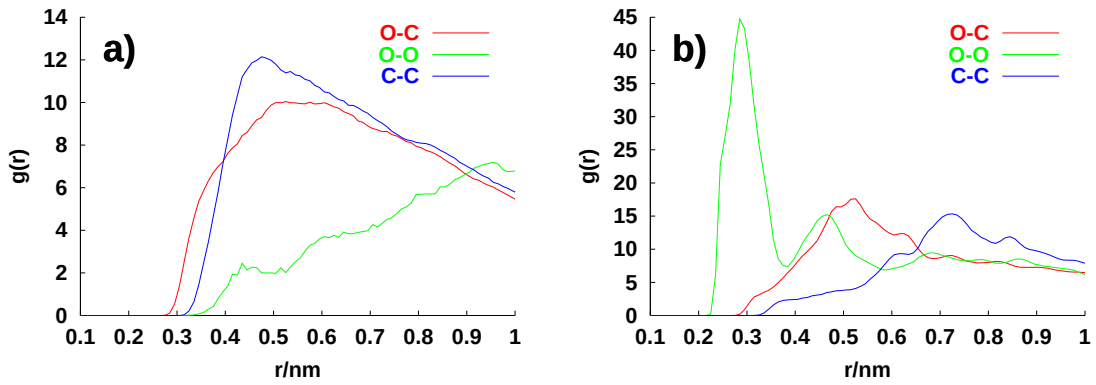


Figure 5.12: Intermolecular radial distribution functions between CH_n-sites (C) and COO^--sites (O) of a) CMC-oligomers (mean over trimer #2, tetramer #1 and #2), and b) PAA-pentamers (mean of all pentamers) in the 110-simulations.

favoring more contacts for the simulation of 110-cellulose with CMC-pentamers.

Intermolecular PAA-PAA radial distribution functions (intermolecular CH_n with CH_n or $C(OO)^-$ with $C(OO)^-$, figure 5.12 b) are almost featureless. At about 0.7 nm the CH_n-CH_n distribution has one broad peak, with is found with a similar structure for the $C(OO)^-$-CH_n distribution with a first maximum at 0.51 nm. The COO^--COO^- distribution has a remarkable first maximum at (0.3 nm), which can only be rationalized by the presence of a large amount of sodium-counterions close to the PAA-pentamers, which shield the electrostatic repulsion of the negative charged carboxylic groups.

The aggregation is of importance to explain later results, like the hydration sphere around the poly(acrylic acid) and the the number of hydrogen bonds with water.

The aggregation phenomen is present for CMC as well. From the results of chapter 3 this is understandable if not expected, because there is one molecule (CMC II) which is folded, thus aggregates with itself. The intermolecular radial distribution functions for carbon and oxygen sites are shown in figure 5.12 a). While the oxygen-oxygen radial distribution has some structure, the oxygen-carbon or carbon-carbon distributions do only have one featureless main peak, showing that there is neither a strong repulsion nor attraction between these atoms nor a very distinctive radial distribution function between CMC's COO^--sites and sodium (not shown). Compared to the CMC distributions, the radial distribution functions of PAA are much more structured. This is a sign of a stronger interaction between PAA pentamers, which is consistent with strong sodium-bridges between the carboxylic groups.

5.2.1 Interaction of surface and polyelectrolyte atoms

5.2.1.1 Carboxy-methyl cellulose with the cellulose 110 surface

The interaction of different atom types of the cellulose surface and the polyelectrolyte oligomers is straightforward evaluated using radial distribution function of the respective atom pairs. However with about twelve distinct types of atoms on both the surface's and polyelectrolyte's side there are more than 150 different radial distribution functions. They are grouped according to the surface's atom and shown in the appendix on page 149 to 152.

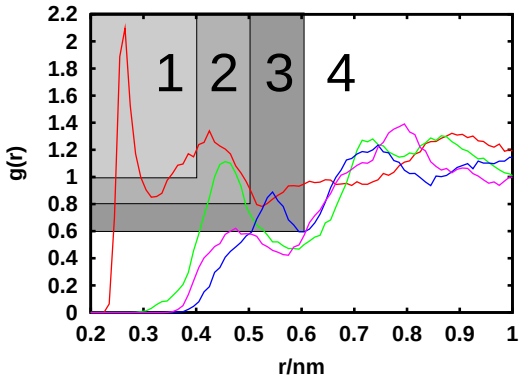


Figure 5.13: Four example radial distribution functions between the cellulose surface and polyelectrolytes in the aqueous phase. The distribution functions are categorized from *strong* (1) to not attraction (4). If the first main peak of a distribution falls into the region denoted 1, the respective atom-pairs attract strongly, in region 2, they attract modestly; in region 3 a weak interaction is assumed and if there is no peak in either one of the first three patches, the surface and polyelectrolyte atoms repel.

The red rdf-function is strong attractive (1) (CMC O2- with surface O6), the green one is modestly attractive (2) (C4-O2), blue is weak (C5-C4) and the magenta colored function is repulsive with not attraction (4) (CMC C5 with C3 of the surface).

Integrals are supplied in the appendix in tables B.2 and B.3. To shed some light onto all the functions they have been sorted into four categories.

1. Strong attraction,

2. modest attraction,

3. weak attraction and

4. no attraction

The categories are assigned according to the position of the first main peak of the rdf function (figure 5.13). The classification is done with the help of figure 5.13.

For the CMC-110 system, all 154 functions are classified according to figure 5.13. The population of the different categories is as follows:

$$1 : 18\%; \quad 2 : 42\%; \quad 3 : 37\%; \quad \text{and } 4 : 3\%$$

So most of all atom-atom interactions are either modest or weak, leaving about 21% together for the outer categories 1 and 4, but with a strong emphasis for strong interactions (category 1). The statistics is further detailed in an on-atom-basis for the different categories (table 5.2).

	CMC-oligomers categories			
Sites	Cat 1	Cat 2	Cat 3	Cat 4
C1	1	7	3	0
C2	1	**9**[a]	1	0
C3	1	4	5	1
C4	0	3	8	0
C5	1	2	7	1
C6	1	2	6	**2**[b]
CE	4	6	1	0
CH2	3	7	1	0
O2	4	5	2	0
O3	2	4	5	0
O4	2	7	2	0
O5	0	2	**9**[c]	0
O6	2	2	6	1
OE	**5**[d]	5	1	0
sum:	27	65	57	5

[a]CMC-C2 with C1, C2, C3, C4, C5, C6, O2, O4, O5 of the cell. surface.
[b]CMC-C6 with C5, O3 of the cell. surface.
[c]CMC-O5 with C1, C2, C3, C4, C5, C6, O2, O3, O4 of the cell. surface.
[d]CMC-OE with C2, C6, O2, O3, sO6 of the cell. surface.

	Cellulose-surface categories			
Sites	Cat 1	Cat 2	Cat 3	Cat 4
C1	1	6	7	0
C2	3	7	4	0
C3	0	**8**[a]	5	1
C4	0	5	**9**[b]	0
C5	1	7	5	1
C6	3	7	3	1
O2	6	4	4	0
O3	4	5	4	1
O4	0	6	7	1
O5	1	6	7	0
O6	**8**[c]	4	2	0
sum:	27	65	57	5

[a]Cell.-C3 with C1, C2, CE, CH2, O2, O4, O6, OE of the CMC monomers.
[b]Cell.-C4 with C1, C3, C4, C5, C6, CH2, O4, O5, O6 of the CMC monomers.
[c]Cell.-O6 with C1, C2, CE, CH2, O2, O3, O4, OE of the CMC monomers.

Table 5.2: Statistics for atom types of either the cellulose surface and the CMC-oligomers. The atom-site's radial distribution functions are grouped according to their category as described in figure 5.13. The highest populated values in a column are typed boldface. For example there are five OE (of the CMC monomers) radial distribution functions in category 1. Atom types OE is the carboxylic oxygen site, CE the carboxylic carbon. CH2 is the α-carbon with respect to the carboxylic functional group.

However on the side of the CMC-oligomer the carboxylic oxygen[1] (OE) dominates the "attractive" radial distribution functions, followed by O2[2] and CE[3]. Other sites like O6[4] or O3[5] do not exhibit more than two radial distribution functions in category 1. So the strongest interaction involves atom types which are known as good hydrogen-bond acceptors like OE and O2. Oxygen O3, which is involved into a stable bond with O5 is involved less distinctive. On the other hand, CMC-oxygen 5 (O5)[6] has many weak interactions of class 3. Radial distribution functions, which are more "repulsive" are scarce, only two of them are found for C6[7]. In general, some of the hydroxy-oxygens have distinct radial distribution functions, whereas O5 has mostly weak interactions with the surface sites. The apparently high population (7 hits) of O4 in class 2 is due to hydroxy O4-oxygens, which terminate a CMC-oligomer. Carbon atoms like C5[8] and C6[9] undergo only class 3 interactions with less polar surface atoms like carbons or O4 and O5. These trends are in accordance with the observed lipophilicity/phobicity or hydrophilicity/phobicity as observed before, where carbon atoms and oxygens like O4 and O5 are (section 4.5.2.1) hydrophobic/lipophilic.

5.2.1.2 Poly(acrylic acid) with the cellulose 110 surface

Because there exist only three different atom types for PAA (CH_n, CE and OE); only 33 different radial distributions exist between the PAA pentamers and the cellulose surface. The radial distributions can therefore be analyzed without any classification, as like for CMC. The radial distribution functions are supplied in figure B.5 in the appendix, integrals are given in table B.4 and B.5.

The radial distribution functions between carboxylic sites (OE and CE) of PAA with surface carbon atoms have first peaks around 0.4 nm for CE with C5/C6 of the 110-cellulose surface and around 0.33 nm for the rdf of OE with C5/C6 of the 110-surface. All other distributions involving surface carbons do not exhibit significant intensity below 0.5 nm and hence would contribute to class 4. The group of radial distributions between the CH_n sites of PAA and carbon atoms of cellulose has one member with a significant peak for the mCH_n-sC6 rdf and some low-intensity peaks for mCH_n-sC1, mCH_n-sC4 and mCH_n-sC5. All of them would be likely to be in class 2, all other distribution functions would be class 4. So of the cellulose surface's carbons, only C6 and C5 do approach PAA-sites.

The corresponding cellulose-oxygen do have similar distribution functions. There is

[1] The following five rdf involve CMC-OE atoms and fall into class 1: mOE-sc1, mOE-sC3, mOE-sC4, mOE-sC5, mOE-sO4. The coding convention for radial distributions is *m* for oligo*m*er and *s* for *s*urface atoms.

[2] The following four rdfs, involve a CMC-O2 atoms and fall into class 1: mO2-sC1, mO2-sC5, mO2-sO5, mO2-sO6.

[3] The following four rdfs involve CMC-CE atoms and fall into class 1: mCE-sC6, mCE-sO2, mCE-sO3, mCE-sO6,

[4] The following two rdfs involve CMC-O6 atoms and fall into class 1: mO6-sC2, mO6-sO2.

[5] The following two rdfs involve CMC-O3 atoms and fall into class 1: mO3-sO2, mO3-sO6.

[6] The following nine rdfs involve CMC-O5 atoms and fall into class 3: mO5-sC1, mO5-sC2, mO5-sC3, mO5-sC4, mO5-sC5, mO5-sC6, mO5-sO2, mO5-sO3, mO5-sO4.

[7] The following two rdfs involve CMC-C6 atoms and fall into class 4: mC6-sO3, mO6-sC6)

[8] The following seven rdfs involve CMC-C5 atoms and fall into class 3: mC5-sC4, mC5-sC5, mC5-sO3, mC5-sO4, mC5-sO5, mC5-sO6.

[9] The following six rdfs involve CMC-C6 atoms and fall into class 3: mC6-sC1, mC6-sC3, mC6-sC4, mC6-sC6, mC6-sO4, mC6-sO5.

always a strong peak at very short distances for mCH_n-sO6, mCE-sO6 and mOE-sO6. None of the other's sites radial distribution functions have some intensity, only mCH_n-sO2 and mCH_n-sO4 exhibit some intensity.

In summary, only the outermost sites of the surface (C6, O6) come close to PAA-sites, which are mainly the carboxylic sites OE and CE.

5.3 Sodium-Counterions

The role of the sodium counterions is likely to be important, as we are aware from the polyelectrolyte simulations in solution (section 3.2.4). There is a similarity between hydrogen bonds and interactions with and through counterions. Both sodium and hydrogen are atoms with either a whole or partial positive charge, respectively, while all other sites are otherwise apolar (carbon, CH_n) or charged negative (oxygen). It is likely that sodium-cations are able to act like a hydrogen in a H-bond. We used the same method, which was previously used to locate hydrogen-bonds. A sodium-bridge between two oxygens is defined "intact", if the oxygen-oxygen distance is lower than 0.31 nm and the oxygen-sodium-oxygen angle is larger 160°. This restriction is necessary to avoid a too high number of sodium-bridges. Because sodium ions are not connected by a chemical bond to a particular donor-oxygen, the angle-limit becomes now much more important than the distance. One can think of a situation, with the distance limit satisfied for a high number of sodium-bonds: For an oxygen on the surface, there are about 10 neighbor atoms in a 0.30 nm-radius. If the angle-criterion is chosen rather weak, every sodium in the vicinity of both oxygens raises a sodium-bridge. Even with this precaution, the number of sodium-bridges on the surface appears to be rather high. There are on average *total* 12.5 sodium-bridges for the CMC-cellulose simulations (1.2 per ion and simulation step) and a *total* of 23.2 contacts for the PAA-cellulose simulations (0.8 per ion and step). Only a small percentage of the *total* numbers given is contributed by cellulose-polyelectrolyte (2.5% in the CMC-simulation, and 0.0% for PAA) or polyelectrolyte-polyelectrolyte (10% for CMC, 5% for PAA) interactions. However, the majority of all CMC-CMC sodium bridges is intramolecular: there are 0.818 intramolecular sodium-bridges and only 0.121 intermolecular. PAA oligomers have even an increased activity for intermolecular sodium-bridge, which is attributed to the higher charge density and higher sodium concentration in the solvent. There are 0.975 (see table 5.4) intermolecular solute-solute sodium-bridges for PAA. Trimer #1 is the only oligomer, which has a significant amount of sodium-bonds with the surface (0.2 bonds per frame on average), which is in most cases an O2-COO or O6-COO bond. Sodium-bridges are due to sodium-ions close to the surface. An overview of polyelectrolyte-polyelectrolyte sodium-bonds is given in tables 5.3 and 5.4. The high amount of cellulose-cellulose bridges is still caused by single ions on the surface, which are involved into several sodium-bonds due to the high density of donor- and acceptor oxygens on the surface.

The radial distribution function between CMC-AGUs and sodium counterions is given in figure 5.14. The plot in figure 5.14 is similar to that for CMC-Na in chapter 3, figure 3.16 a). The respective PAA-sodium distributions (not shown), has an obfuscated multi-peak structure, which is not amenable to an interpretation.

The repeat-unit(center of mass)-sodium radial distribution function (figure 5.14) shows three peaks, with maxima at about 0.36 nm, 0.46 nm and 0.76 nm. The first two are sub-

		$\langle n \rangle$
Trimer #1	Trimer #1	0.221
Trimer #1	Tetramer #1	0.000
Tetramer #1	Tetramer #1	0.118
Tetramer #1	Trimer #2	0.028
Tetramer #1	Tetramer #2	0.026
Trimer #2	Trimer #2	0.196
Trimer #2	Tetramer #2	0.067
Tetramer #2	Tetramer #2	0.283
sum:		0.939

Table 5.3: Average number of CMC-CMC sodium-bridges per frame (no normalization to the number of sodium-ions) for 110-CMC simulation. Oligomer pairs without any bonds are not displayed.

		$\langle n \rangle$			$\langle n \rangle$
#1	#1	0.120	#4	#4	0.025
#1	#2	0.000	#4	#5	0.000
#1	#3	0.012	#4	#6	0.005
#2	#2	0.181	#5	#5	0.472
#2	#3	0.000	#5	#6	0.020
#2	#4	0.051	#6	#6	0.016
#3	#3	0.161			
#3	#5	0.023			
#3	#6	0.017			
sum:				1.103	

Table 5.4: Average number of PAA-PAA sodium-bridges per frame (not normalized to the number of sodium-ions). Pentamer pairs without bonds are not shown.

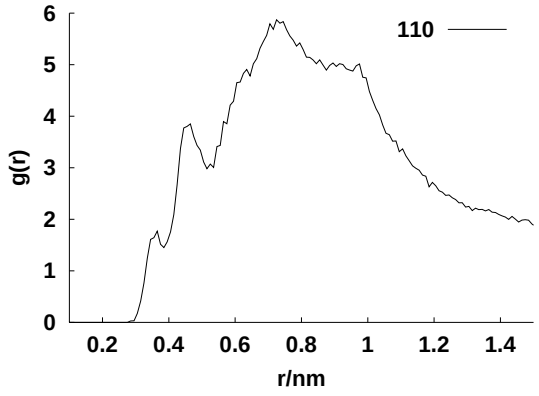


Figure 5.14: Average repeat unit (center of mass)-sodium radial distribution functions for CMC-oligomers (average of trimers and tetramers) of the 110 simulation. The respective plot for PAA is not shown, as it is obfuscated by a multi-peak structure without sensible physical interpretation. This is due to the higher concentration PAA and Na^+-ions. Furthermore there are several small PAA strands instead of one long molecule.

peaks of the main one, which tends towards $g(r) = 2$ at 1.40 nm.

5.4 Hydrogen Bonds

Hydrogen bonding was found to be one, if not the dominant interaction for both the polyelectrolyte in water (see chapter 3 and reference [16]) and for the cellulose-water-interface (section 4 and reference [15]). By the combination of three components into one system, there are four kinds of hydrogen-bonds possible:

- cellulose-polyelectrolyte,
- water-polyelectrolyte,
- water-surface and
- polyelectrolyte-polyelectrolyte.

5.4.1 Solute-Surface Hydrogen Bonds

Hydrogen bonds between the surface and CMC are detailed in tables 5.5 with every oligomer having its own entry. PAA-cellulose bonds are discussed in the text.

Normalization of the number of hydrogen bonds is now more delicate, as compared to the polyelectrolyte-water systems in chapter 3 and 4: Because there is no longer an immobile component (the crystal or polyelectrolyte) surrounded by a mobile solvent, the number of hydrogen bonds is not to be normalized to the number of simulation frames. The polyelectrolyte adsorption-time is a more suitable normalization basis. It is measured by the number of time steps when an oligomer is adsorbed on the cellulose-surface. For analysis of the hydrogen bonds, a polyelectrolyte is defined to be adsorbed, if there is at least one hydrogen bond with the cellulose surface. With respect to this basis, the average

CMC → 110-surface	O2	O3	O4	O5	O6
Trimer #1: 0.92 bonds in total (5.3 ns)					
O2H	0.05	0.14	0.16	0.00	0.14
O3H	0.16	0.06	0.00	0.00	0.14
O6H	0.01	0.02	0.00	0.00	0.04
Tetramer #1: 0.08 bonds in total (2.0 ns)					
O2H	0.00	0.01	0.00	0.00	0.00
O3H	0.00	0.00	0.00	0.00	0.00
O6H	0.00	0.01	0.00	0.00	0.06
Trimer #2: 0.01 bonds in total (2.5 ns)					
O2H	0.00	0.00	0.00	0.00	0.00
O3H	0.00	0.00	0.00	0.00	0.00
O6H	0.00	0.00	0.00	0.00	0.01
Tetramer #2: 0.44 bonds in total (3.3 ns)					
O2H	0.00	0.03	0.00	0.00	0.05
O3H	0.00	0.02	0.00	0.00	0.04
O6H	0.02	0.16	0.00	0.00	0.12

110-surface → CMC	O2	O3	O4	O5	O6	OE
Trimer #1: 3.78 bonds in total (5.3 ns)						
O2H	0.02	0.19	0.05	0.03	0.26	1.21
O3H	0.00	0.01	0.01	0.00	0.00	0.58
O6H	0.08	0.05	0.08	0.01	0.08	1.12
Tetramer #1: 1.26 bonds in total (2.0 ns)						
O2H	0.00	0.00	0.03	0.00	0.02	0.03
O3H	0.00	0.00	0.00	0.00	0.00	0.00
O6H	0.01	0.00	0.11	0.05	0.08	0.93
Trimer #2: 2.29 bonds in total (2.5 ns)						
O2H	0.00	0.00	0.02	0.00	0.00	0.62
O3H	0.00	0.00	0.00	0.00	0.00	0.22
O6H	0.00	0.02	0.02	0.00	0.00	1.39
Tetramer #2: 1.41 bonds in total (3.3 ns)						
O2H	0.03	0.02	0.30	0.00	0.22	0.00
O3H	0.00	0.00	0.00	0.00	0.01	0.00
O6H	0.48	0.17	0.05	0.02	0.08	0.03

Table 5.5: Hydrogen bonds of the CMC-oligomers (trimer #1 and #2, tetramer #1 and #2) with the 110 surface. First colum of each tabular shows donor groups, acceptor oxygens are arranged in lines. Values with respect to the number of states where the polyelectrolyte is adsorbed (see text page 125). The normalization base is different for all oligomers and given in nanoseconds in the oligomer sub-header lines. There is *no* normalization with respect to the number of interaction sites.

number of hydrogen bonds between the polyelectrolyte and the cellulose surface is given in table 5.5.

This procedure still does not deal with the fact that the number of polyelectrolyte-surface hydrogen-bonds depends on the local arrangement of the polyelectrolyte. The orientation of the polyelectrolyte with respect to the surface becomes decisive. A polymer lying flat on the surface has more possibilities for multiple hydrogen-bonds than a polymer with a tip-only surface contact. Trimer #1 is the only CMC oligomer where all repeat units have hydrogen-bonds with the surface but with a bias towards AGU 3 (80%). AGU 1 and 2 contribute with 10% to all trimer #1-cellulose hydrogen-bonds. The other oligomers have only one or two repeat units in contact with the surface: Tetramer #1 has 80% of all hydrogen-bonds with the surface via AGU #1. Trimer #2 has all bonds to the surface via AGU #3 and tetramer #2 90% via AGU #1. The remaining percentage is covered by hydrogen-bonds from adjacent repeat units. With respect to the solute molecular weight there are in total $3.5 \cdot 10^{-3}$ bonds/amu CMC between the 110 surface and CMC.

The cellulose crystal donates more hydrogen bonds to the CMC-polyelectrolyte than there are bonds in the reverse direction polyelectrolyte→surface. For example 0.92 bonds reach from the CMC trimer #1 to the cellulose surface, but 3.78 bonds from the surface to trimer #1. The difference is more pronounced for trimer #2, where the number of hydrogen bonds CMC→surface is minute, but on average 2.29 hydrogen-bonds span from the surface to the trimer. In general, we find more hydrogen-bonds with trimers than with tetramers.

This is due to the lower number of intermolecular bonds involving carboxylic donor oxygens (OE) on side of the tetramers: There are 2.91 and 2.23 surface→OE bonds with trimers and either no or only some (max: 0.96) hydrogen-bonds with tetramers (table 5.5).

In section 3.2.3.1 ("Solute-Solute Hydrogen Bonding") intramolecular hydrogen-bonds via COO^- groups have been observed for CMC-oligomers and similar bonds are found here as well: on average 0.54 intramolecular hydrogen-bonds via a carboxylic-group (OH-OE) are found for trimer #1 and 0.38 H-bonds for tetramer #1 involve OH-OE sites as well. Trimer #2 and tetramer #2 have 0.13 and 0.35 hydrogen-bonds with an COO^- acceptor group. The detailed statistics of solute-solute hydrogen bonds is given in the appendix in table B.1 on page 147.

Other trends for hydrogen bonding are as observed before: O4 and O5 are rather inactive and do hardly undergo any H-bonding, the only exception is an occasional O6H→O5 bond (0.05 bonds/step) for tetramer #1 and the surface→CMC hydrogen bond from O2H to O4 in tetramer #2.

In contrast to the donor activity of the polyelectrolyte, the surface's O2H- and O6H hydroxy groups are moderately active donor-sites and even O3H has donor bonds with trimer #1 and #2. The total number of hydrogen-bonds from the surface to CMC for trimer #1 and tetramer #2 are:

trimer #1: O2H: 1.76, O3H: 0.60, O6H: 1.42 (sum: 3.78),
and tetramer #2: O2H: 0.57, O3H: 0.01, O6H: 0.83 (sum: 1.41).

Oxygen 6 (O6) and the carboxylic oxygen play a marked role during the process of adsorption: They form the first contact between the polyelectrolyte and the surface. For example during 190 ps one single O6H-OE bonds is the only bond between the surface and trimer #2. During this time the O6H bond changes four times from one oxygen of the COO^- group to the other one. The closeness of two good (and accessible) acceptor oxygens is inevitable for the adsorption process.

For oligomers other than trimer #2 the first contact is established from a cellulose O6H group, but occasionally with other acceptor groups on side of the polyelectrolyte. Times where exclusively one bond via a cellulose O6H group is present range from 10 ps to 20 ps.

For the PAA-110 system there is one pentamer hydrogen-bonded to the surface. Pentamer #2 has two bonds, from an O2H and an O6H group of the surface to COO oxygens, both bonds live for 3.2 ns; in total this averages to 0.026 bonds per adsorbed state for O2H and 1.19 O6H→C(OO) bonds per state. There are three types of hydrogen-bonds found, for about 4.1 ns there is at least one bond present. The total number of hydrogen bonds per adsorbed simulation step for PAA is

O2H: 0.71, O3H: 0.05 and O6H: 1.04.

If a PAA molecule is readily solvated with water, it is kept in the bulk and does not form hydrogen bonds with cellulose. In chapter 3.3 it was concluded, that PAA has a large solvation shell, which shields the polyelectrolyte from adsorption and inhibits any contact.

5.4.2 CMC-Solute-Solute Hydrogen-Bonds

In section 3.2.3, CMC in infinite dilution was found to exhibit strong intramolecular hydrogen bonds. These bonds connect distant glucose-units and are responsible for the com-

		CMC m110
		$\langle n \rangle$
Trim.#1	Trim.#1	1.712
Tetram.#1	Trim.#1	0.000
Tetram.#1	Tetram.#1	1.602
Tetram.#1	Trim.#2	1.278
Tetram.#1	Tetram.#2	3.714
Trim.#2	Trim.#2	0.644
Tetram.#2	Trim.#1	0.000
Tetram.#2	Trim.#2	3.139
Tetram.#2	Tetram.#2	1.522

Table 5.6: Average total number ($\langle n \rangle$) of CMC-CMC hydrogen-bonds per frame. Oligomer-pairs without any H-bonds are not shown. More details are given in the appendix on page 147.

pact, folded structure of CMC. As the oligomers here are about half in length (trimer- and tetramer vs. hepta- and octamer in chapter 3) we do not expect folded or collapsed configurations, which were found concurrently with strong intra-molecular hydrogen bonding. Indeed, there are no hydrogen-bonds, which span multiple rings for the trimers and tetramers: For trimer #1 and tetramer #1, 10%, respectively 13%, of intramolecular hydrogen bonds are within one glucose unit; for trimer #2 and tetramer #2 we find 34% and 20%. The others connect neighboring glucose units. The values are in accordance with results for CMC I (in chapter 3 there are on average 0.78 intramolecular hydrogen bonds per anhydroglucose and simulation step).

Because of the CMC aggregation (section 5.2), intermolecular hydrogen bonding is expected to play a important role. There are some oligomers found to have excessive intermolecular hydrogen-bonding like trimer #2 with tetramer #2. Even if trimer #2 and tetramer #2 have not less than 31 different hydrogen bonds, almost all (92%) happen via a COO acceptor. The O2H-group of trimer #2 is the strongest donor-group and O3H of tetramer #2 as the second strongest one. The CMC-CMC H-bonding is tabulated on page 147 in the appendix, and summarized in table 5.6.

5.4.3 Hydrogen Bonds with Water

For the CMC-cellulose simulation the number of hydrogen-bonds between the surface and water decreases with ongoing simulation time. The process is continuous without abrupt change when a polyelectrolyte adsorbs. The distribution of the number of hydrogen-bonds (not shown) is unimodal.

The average total number of hydrogen bonds between water and cellulose during the first 0.6 ns is 110, whereas it is only 104 during the last 3 ns. It is likely that CMC is capable to displace water from the surface. The overall number of hydrogen bonds is in the same order of magnitude as for the surface-water systems in section 4.4, 4.9. The details for different atoms-types are given in table 5.7.

	110 CMC		
	Donor	Acceptor	Sum
O2	0.259	0.080	0.339
O3	0.047	0.233	0.280
O4	-	0.040	0.040
O5	-	0.004	0.004
O6	0.213	0.234	0.447
sum	0.519	0.591	1.110

Table 5.7: Average total number of hydrogen bonds between the surface and water per anhydroglucose unit. Average over both surface layers (top and bottom). "Donor" column are surface→water hydrogen-bonds, column "Acceptor" water→surface bonds.

CMC m110				CMC m110			
	Acc.	Don.	sum		Acc.	Don.	sum
Trimer #1				Trimer #2			
O2	0.63	0.93	1.56	O2	0.89	0.68	1.57
O3	0.89	0.65	1.54	O3	1.17	0.52	1.69
O4	1.16	-	1.16	O4	1.14	-	1.14
O5	0.32	-	0.32	O5	0.34	-	0.34
O6	1.41	1.38	2.78	O6	1.75	1.57	3.32
COO	6.25	-	6.25	COO	7.33	-	7.33
Quadromer #1				Quadromer #2			
O2	1.41	1.83	3.24	O2	1.12	1.08	2.20
O3	1.47	1.04	2.51	O3	1.49	0.66	2.15
O4	1.04	-	1.04	O4	1.17	-	1.17
O5	0.54	-	0.54	O5	0.32	-	0.32
O6	1.66	1.31	2.97	O6	1.43	0.96	2.39
COO	11.98	-	11.98	COO	12.23	-	12.23

Table 5.8: Average total number of hydrogen-bonds per simulation step between CMC oligomers and water. In figure 5.15 the number of H-bonds are plotted as a function of time.

The number of hydrogen bonds between water and all four CMC-oligomers is given in table 5.8. The number per AGU is higher than for the heptamer in chapter 3, because of the higher H-bond activity of terminal O4-oxygens. For PAA there is only one type of solvent-solute hydrogen-bond with 11 bonds/pentamer on average, which corresponds to 2.2 bond/COO group. Due to the aggregated PAA-pentamers, there are less bonds found here than for the 23-mer in table 3.4, where 5.8 bonds/COO^- group were encountered. A plot of the number of hydrogen bonds between polyelectrolyte and water is shown

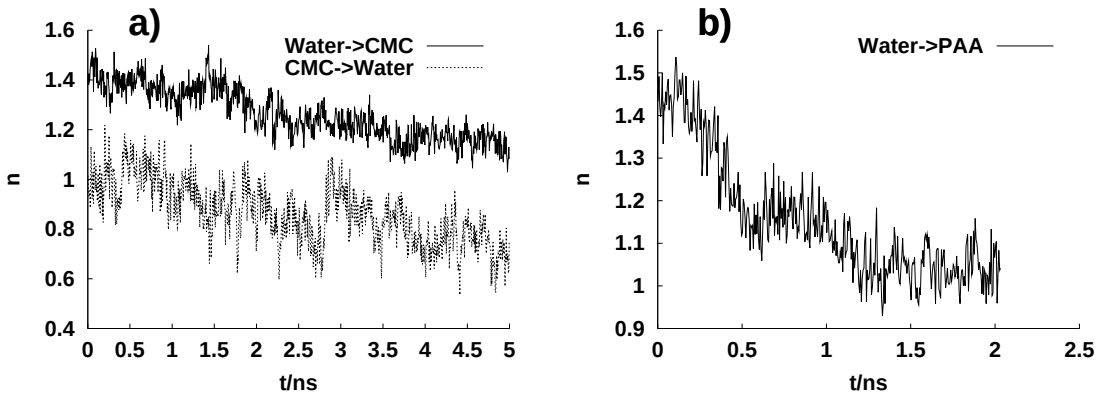


Figure 5.15: Average number of polyelectrolyte-water and water-polyelectrolyte hydrogen-bonds as a function of simulation time (t). Normalization to the number of CMC donor- or acceptor sites (resp. atoms). a) Hydrogen bonds between CMC and water, b) between PAA and water, the plot is truncated at 2.0 ns, but the number of H-bonds does not change after in the later course of the simulation.

in figure 5.15. In both the CMC and PAA oligomer-water systems, bonds are broken with ongoing simulation time. However, there is a plateau for PAA after about 1.5 ns of equilibration. CMC has a loss of polyelectrolyte→water bonds due to O2H(tetramer #2), O3H(trimer #2) and O3H(tetramer #2) bonds, where the number of hydrogen bonds goes down almost to 50% of the initial ($t = 0$ ns) value.

Averages of *all* (normalisation with respect to the number of simulation steps, whereas normalisation in table 5.9 happens to a different basis) hydrogen-bonds for CMC-polyelectrolyte are given in table 5.9. There are

Trimer #1	Tetramer #1	Tetramer #2	Trimer #2
20.9 (1.0)	31.1 (1.2)	23.2 (1.2)	32.4 (1.2)

hydrogen-bonds for CMC-oligomers. Numbers in parentheses "()" are with respect to the number of oxygen-sites. Trimer #1 has fewest bonds, which is partly due to its reduced hydrogen-bond activity towards water, but to a significant extent also to the lack of intermolecular CMC-CMC-bonds. Trimer #1 has substituted intermolecular hydrogen-bonds from one like molecule (CMC) by another (cellulose). Because trimer #1 is close to the surface, it cannot be well accessed by solvent molecules. All other molecules have 1.2 hydrogen-bonds per oxygen-atoms, but the flat-adsorbed trimer #1 only 1.0.

5.4.3.1 Water-Repeat Unit Distributions

Radial distribution functions of the polyelectrolyte repeat units are given for some example oligomers in figure 5.16. They look different from the distributions in chapter 3, figure 3.14. First of all, the distribution of the adsorbed CMC oligomers (figure 5.16 a) do not approach unity, because water is not distributed symmetrically around the oligomers. The radial distribution of trimer #1, which has the smallest center of mass-surface distance, goes to 0.40 The oligomer is almost flat on the surface, each AGU has only a hydration half-sphere. Other oligomers exhibit more hydration water, they can access more than a half-sphere full of water.

a) Cellulose↔CMC

	CMC-sites							
	O2	O3	O4	O5	O6	OE	sum	
Trimer #1	0.6	0.6	0.1	0.0	0.4	2.9	4.6	(0.23)
Tetramer #1	0.0	0.0	0.1	0.0	0.3	1.1	1.5	(0.05)
Trimer #2	0.0	0.0	0.0	0.0	0.0	2.3	2.3	(0.15)
Tetramer #2	0.6	0.3	0.4	0.0	0.6	0.0	1.9	(0.07)

b) Water↔CMC

	CMC-sites							
	O2	O3	O4	O5	O6	OE	sum	
Trimer #1	1.6	1.5	1.2	0.3	2.8	6.3	13.7	(0.69)
Tetramer #1	3.2	2.5	1.0	0.5	3.0	12.0	22.2	(0.82)
Trimer #2	1.6	1.7	1.1	0.3	3.3	7.3	15.3	(0.77)
Tetramer #2	2.2	2.2	1.2	0.3	2.4	12.2	20.5	(0.76)

c) CMC↔CMC (intermolecular)

	CMC-sites							
	O2	O3	O4	O5	O6	OE	sum	
Trimer #1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	(0.00)
Tetramer #1	0.5	1.1	0.1	0.0	0.4	2.8	4.9	(0.11)
Trimer #2	1.1	1.1	0.0	0.0	0.5	1.7	4.4	(0.22)
Tetramer #2	1.3	1.6	0.0	0.1	0.8	3.1	6.9	(0.26)

d) CMC↔CMC (intramolecular)

	CMC-sites							
	O2	O3	O4	O5	O6	OE	sum	
Trimer #1	0.1	0.3	0.1	0.1	0.6	0.6	1.8	(0.09)
Tetramer #1	0.1	0.3	0.0	0.0	0.6	0.6	1.5	(0.06)
Trimer #2	0.1	0.1	0.1	0.1	0.2	0.1	0.7	(0.04)
Tetramer #2	0.1	0.3	0.0	0.1	0.5	0.5	1.5	(0.06)

Table 5.9: Overview of hydrogen-bonds in the CMC-110 simulation. For every site-type and oligomer the table gives the total number (acceptor and donor) of hydrogen bonds.
a) Cellulose-Surface with CMC b) water and CMC and c) intermolecular d) intramolecular CMC-CMC hydrogen-bonds. Normalization is with respect to the total number of frames for b) , c) and d). The cellulose-CMC hydrogen-bonds (a) are divided by the number of adsorbed states (see text). Numbers in parentheses "()" in the last column are normalized with respect to the number of oxygen sites of the respective oligomer.

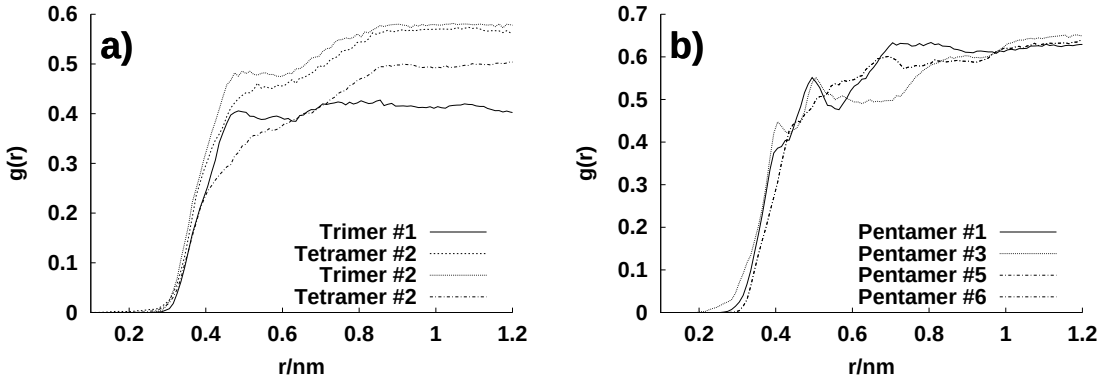


Figure 5.16: Radial distribution functions $g(r)$ for a) CMC and b) selected PAA oligomers (average over building units's center of mass) with water solvent oxygens for the 110 simulations only. The distributions are calculated for the whole simulation, disregarding the adsorption of a polyelectrolyte. However the change of the water-sphere on adsorption is significant, even if the overall shape of the rdf is preserved (figure 5.17 below). Radial distribution functions are normalized with respect to the density of water in the aqueous phase (instead of using the overall water denisity in the simulation box).

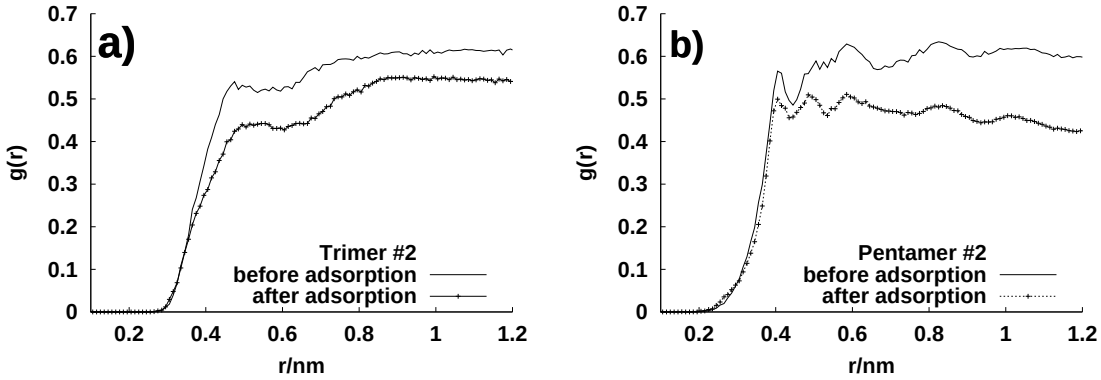


Figure 5.17: Radial distribution functions as above in figure 5.16, but before and after adsorption of the respective oligomer for a) CMC trimer #2 and b) PAA pentamer #2. The adsorption of a given polyelectrolyte is accompanied by a loos of water from the hydration shell. Radial distribution normalized with respect to the water density in the aqueous phase.

PAA distributions reach almost 0.6, as they do not approach the surface but nevertheless do not have an isotropic solvation shell, because of aggregation. Pentamer #2 is an exception, as it is close to the surface. The amount of hydration water of PAA is not changed by the presence of the surface. However the distribution for PAA look different than for the 23-mer in chapter 3, which is explained by the PAA-aggregation.

5.5 Reconstruction of the 1$\bar{1}$0 Crystal Surface

In both 1$\bar{1}$0-simulations (with CMC and PAA-solute) the cellulose-crystal deforms. After about 1 ns of equilibration the crystal is stable for 150 ps and then undergoes a deformation which is finished after 60 ps. The deformed structure is stable for the rest of the simulation.

The simulation parameters for the two 1$\bar{1}$0 simulations

- 1$\bar{1}$0-cellulose surface with CMC oligomers (trimer #1, tetramer #1, trimer #2 and tetramer #2) and
- 1$\bar{1}$0-cellulose surface with PAA (six pentamers)

are as given for the 110-simulations. Simulation times are 4 ns in both cases after 1 ns of equilibration.

A series of snapshots of the deformation phase is given in figures 5.18 a–f. In subfigure 5.18 a), the crystal is intact, in b) there is a small dip between two cellulose chains, which is absent in c). After 55 ps the dent is deeper, and the left part of the crystal starts sliding upwards, which results in the deformation shown in figure 5.18 f). A picture of the total crystal is given in figure 5.19. The single cellulose chains keep parallel to the y-axis. The same process is found for the PAA-1$\bar{1}$0 simulation runs at about the same time.

The density profiles (after deformation) in z-direction are shown in figure 5.20. There is a lot of structure present: six layers can be identified as for the unruptured CMC-water system. The cellulose now extends stronger into the aqueous phase as before. The crystal-density is shifted outwards by 0.8 nm.

Both the nonbonded energy and the pressure perpendicular to the chain-direction (c-axis) are plotted in figure 5.21. Both curves show a strong change, the pressure increases from the negative and both the total and non-bonded energy drop down (figure 5.21).

Because simulations were performed under constant-volume condition this might be a consequence of the negative pressure perpendicular the chain direction (y-axis) (about 150 kPa for both 1$\bar{1}$0 simulations). Probably the narrow surface does not have sufficient space perpendicular to the cellulose-chain, to accommodate pressure changes by the polyelectrolyte, whereas the wide surface has more room left to balance negative pressure. This view is supported by the fact that the pressure's y-component after the rupture equals that one of the pure water-cellulose system.

However, there are undisturbed patches on the surface, where CMC adsorbs. The CMC molecules do not adsorb at the kink of the surface, but instead are ordered perpendicular to the cellulose chains as it was found for trimer #1 in section 5.1.2. Two trimers are flat on the surface (see figure 5.22). Hydrogen bonding seems to be again the prevailing interaction between CMC and the surface: There are 5.4 hydrogen bonds between trimer #1 and the surface and 5.4 bonds for trimer #2. As found for the 110 surface, there are many more bonds from the surface to CMC (85%, respectively 75% of all bonds for trimer #1 and trimer #2) than in the opposite direction. Bonds from the O2H and O6H surface groups to C(OO)-atoms of the polyelectrolyte account for almost all cellulose→trimer bonds. The hydrogen bond averages for trimer #1 and trimer #2 are given in table 5.10. The atom-type statistics is as for the 110-system: O2H and O6H of the surface are better donor-groups than O3H, the ether-sites are of no importance for the overall H-bond statistics. Of the tetramers, only tetramers #1 adsorbs: After $\sim$ 3.0 ns there are two surface-CMC

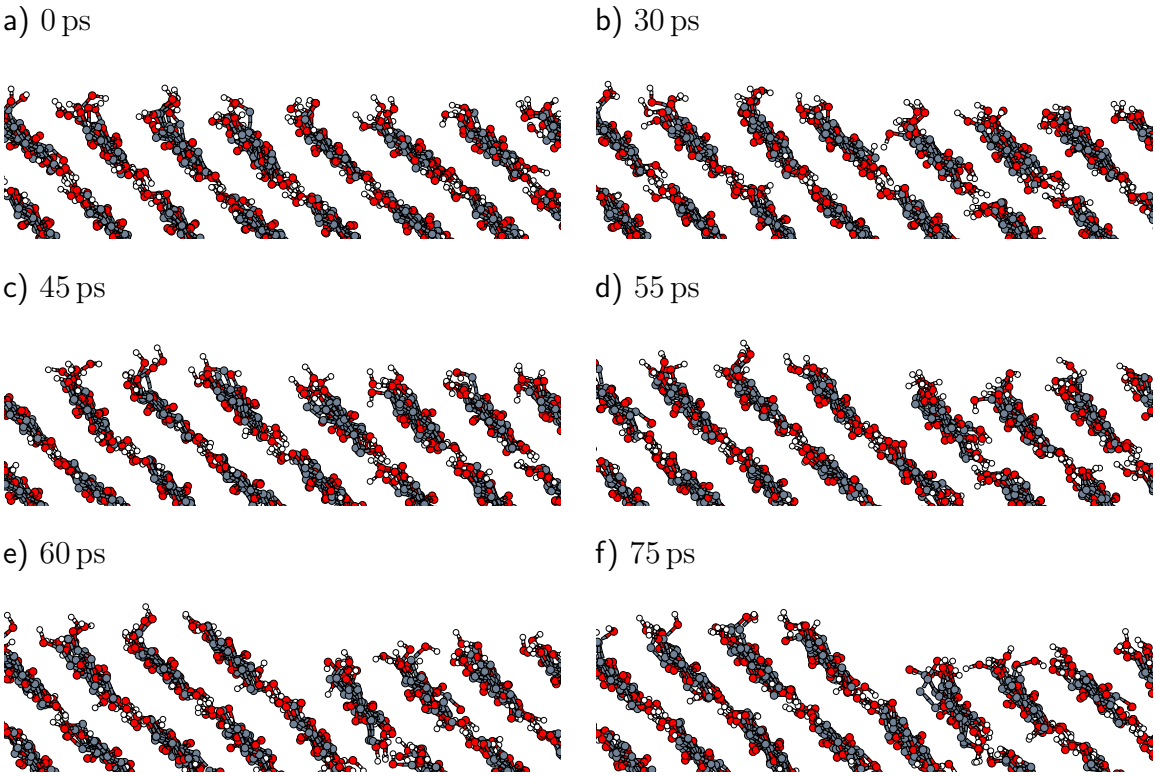


Figure 5.18: Timeseries of intermediate states of the crystal rupture in the CMC-1$\bar{1}$0-simulation. The origin in time (subfigure a)) is $t = 150\,\text{ps}$ of simulation time after the equilibration. The pictures show the top cellulose interface towards water and parts of the crystal. The crystal slides along the 200-plane. The process is similar for the PAA-1$\bar{1}$0-system.

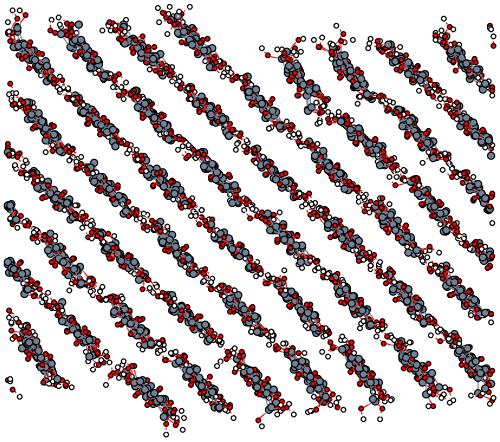


Figure 5.19: Picture of the distorted 1$\bar{1}$0-cellulose-crystal.

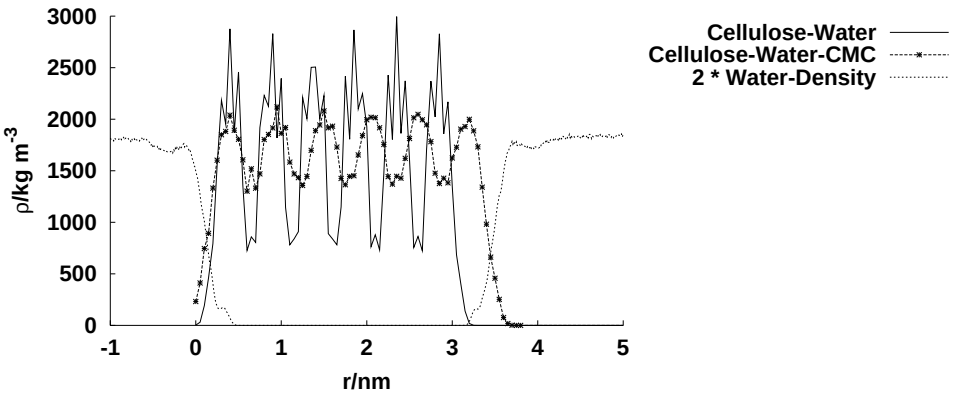


Figure 5.20: Cellulose- and water-density profiles in z-direction of the CMC-1$\bar{1}$0 (narrow) system. In the broken system, there are six layers present, but in comparison with the 1$\bar{1}$0-water system their structure is washed out and the crystal extents in z-direction by about 0.8 nm into the aqueous phase.

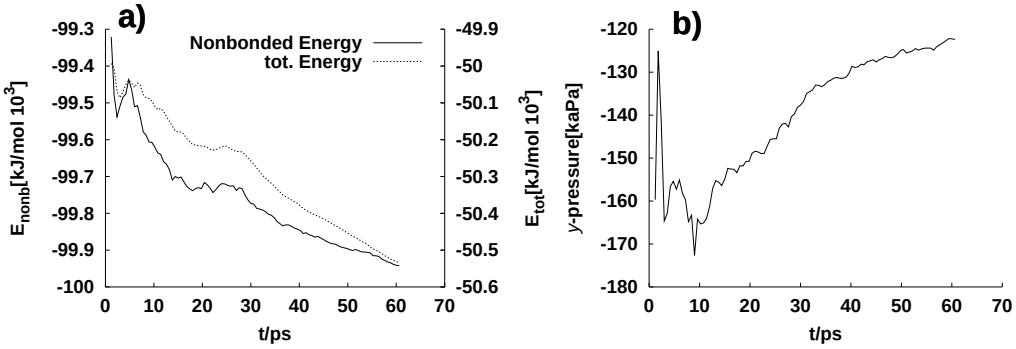


Figure 5.21: Energy (total and nonbonded) and pressure (component perpendicular to the chain directionand to the surface normal) as a function of time during the cellulose crystal's transition. The pressure tends toward the pressure-value of the pure 1$\bar{1}$0-water simulation. The origin in time equals that one in figure 5.18.

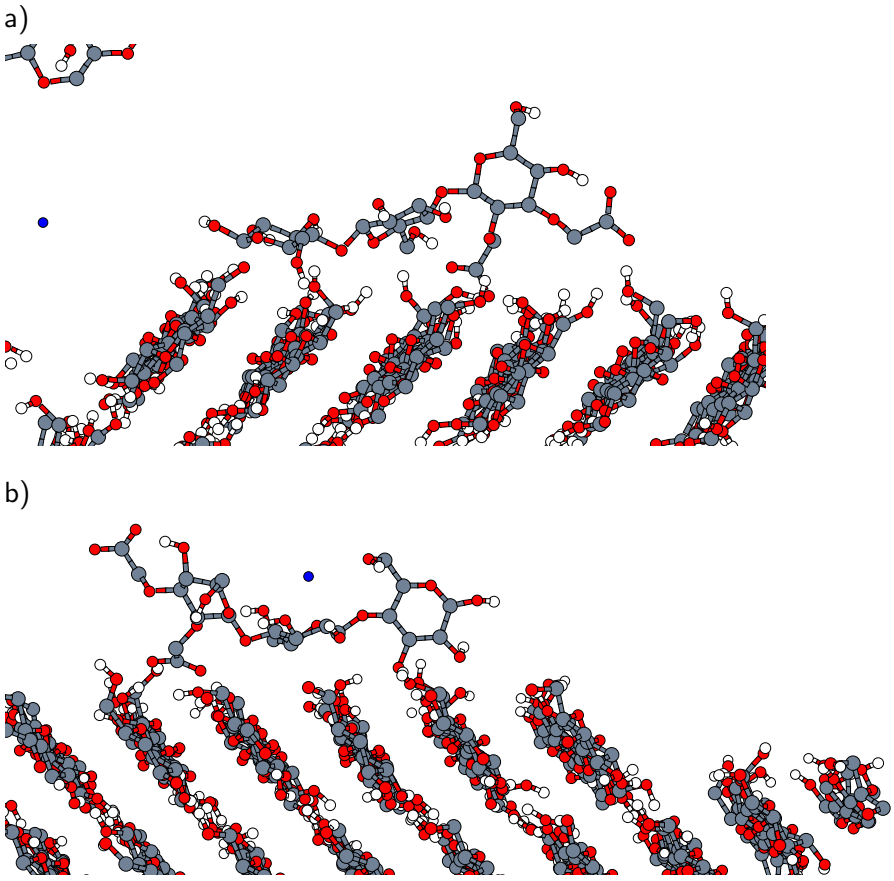


Figure 5.22: Snapshots of two CMC trimers on the $1\bar{1}0$-surface. Both of them are flat on the surface and adsorb on the undisturbed parts of the cellulose-crystal.

1$\bar{1}$0-surface → CMC						
	O2	O3	O4	O5	O6	OE
Trimer #1 (4.3 ns) sum: 4.5 bonds						
O2H	0.102	0.027	0.029	0.073	0.255	1.673
O3H	0.127	0.007	0.004	0.001	0.013	0.805
O6H	0.058	0.003	0.127	0.142	0.102	0.984
Trimer #2 (4.3 ns) sum: 4.1 bonds						
O2H	0.362	0.062	0.044	0.024	0.036	1.670
O3H	0.022	0.007	0.186	0.003	0.100	0.767
O6H	0.033	0.043	0.081	0.016	0.107	0.535

CMC → 1$\bar{1}$0-surface					
	O2	O3	O4	O5	O6
Trimer #1 (4.3 ns) sum: 0.84 bonds					
O2H	0.147	0.017	0.000	0.001	0.121
O3H	0.081	0.032	0.000	0.000	0.061
O6H	0.105	0.026	0.019	0.131	0.095
Trimer #2 (4.3 ns) sum: 1.3 bonds					
O2H	0.090	0.029	0.002	0.015	0.218
O3H	0.401	0.011	0.000	0.000	0.064
O6H	0.013	0.003	0.005	0.013	0.394

Table 5.10: Hydrogen bonds per adsorbed state between the 1$\bar{1}$0-surface and CMC-trimers. The normalization is done using the number of states a given trimer has hydrogen-bonds with the surface, instead of using the simulation length as a basis.

hydrogen bonds, but the oligomer takes up a end on the surface conformation with no more than two bonds to the surface simultaneously.

PAA does not adsorb on the 1$\bar{1}$0 surface with the exception of one pentamer (pentamer #4, see figure 5.23. This pentamer is tied by sodium bonds (see section 5.3) to the surface. There are on average 0.1 sodium-bonds between the pentamer's COO-groups and the surface. Even if this value is small, there are three ions located in the space between the polyelectrolyte and the surface, so it is likely, that they play a role, even if there are not many sodium bridges with an exact geometric alignment. Additionally there are about 2–3 hydrogen bonds with the surface, all of them are either via a O2H or O6H donor group of the surface. However, later the PAA-molecules are located close to the kink (figure 5.24).

The reasons for the rupture of the 1$\bar{1}$0-crystal are not understood completely. As the process happens in one concerted step and no solutes are close to the position, where the crystal slides, it is most likely that there is some external driving force like the total energy or the pressure. Even if system escapes by changing the crystal-structure, the stability of the intact crystal is demonstrated by the regular crystal conformation beside the defect, where almost intact cellulose-crystal patches exist. It is also not clear why the cellulose crystal deforms for both polyelectrolytes in a very similar way, while it appears

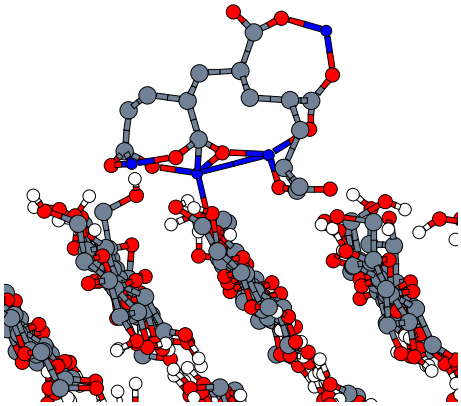


Figure 5.23: One PAA pentamer adsorbs at an undisturbed part of the $1\bar{1}0$-cellulose surface. There are three sodium-ions near the surface, which form sodium-bridges (blue lines) (picture after a simulation time of about 1 ns).

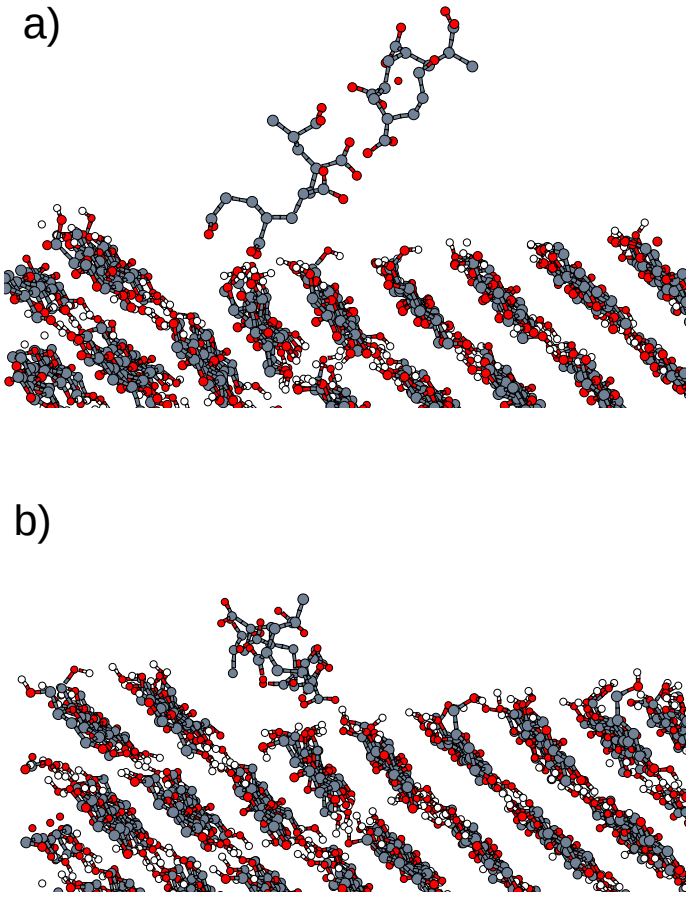


Figure 5.24: Two PAA pentamers close to the surface's kink at simulation times of a) 3.2 ns and b) 4.0 ns.

to be completely stable for the cellulose-water simulations (with a constant pressure of $p_{yy} - 110\,\mathrm{kPa}$, $p_{zz} = -60\,\mathrm{kPa}$.

However, the adsorption properties are the following: CMC adsorbs well and trimers keep a flat on the surface conformation which multiple hydrogen-bonds. PAA pentamers do not adsorb, except for one molecule, but the molecules are located close to the surface-kink without any H-bonds or sodium-bridges (except pentamer #4).

5.6 Conclusion

In chapter 5 we have studied four different combinations of a polyelectrolyte (CMC, PAA) in water together with the 110 and $1\bar{1}0$ surface of the monoclinic cellulose crystal. These simulations allow for an direct comparison of the different adsorption properties of CMC and PAA on crystalline cellulose.

The experimental observation [43] – CMC adsorbs, whereas PAA does not – is reproduced by the simulations. The CMC oligomers (trimers and tetramers) do approach the cellulose 110 surface and undergo hydrogen bonding, which involves sterically exposed functional groups of both CMC and cellulose, mainly O6H, O2H and COO^- groups.

All CMC oligomers adsorb on the surface while most of them form hydrogen-bonds with the surface. The z-coordinate of four CMC oligomers is close to either one of the celluloses surfaces, the CMC molecules keep their z-position and no desorption takes place. In contrast, PAA molecules do not approach the surface (with one exception).

During the whole simulation time one CMC-oligomer is flat on the surface and some others take a end on the surface conformation with only one AGU in contact with the cellulose crystal. Every adsorption starts from a surface-O6H hydrogen-bond to a polyelectrolyte. There is a large degree of inter- and intramolecular hydrogen bonding for the CMC oligomers. Linking of the polyelectrolyte by counterions does not take place, most of counterion-bridges are intramolecular. On adsorption, the diffusion along the surface normal is significantly slowed down. Once in contact with the cellulose surface, some CMC's anhydroglucose units try to align parallel to the 200-cellulose plane (figure 5.10). While CMC oligomers keep some intermolecular separation, PAA pentamers aggregate in solution with counterions and other PAA molecules. Only one PAA pentamer has few hydrogen bonds with the surface, as all but two pentamers stay in the solvent bulk.

By the measure of radial distribution functions, most of site-site interactions between the polyelectrolyte solute and the cellulose surface are rather indifferent and featureless. Only some rdfs resemble a strong attraction: For example CMC-OE has pronounced "attractive" rdfs with C2/O2 and C6/O6 sites of the surface; there are radial distribution functions of O6 with OE, O2, O3 of CMC oligomers. PAA has few "attractive" rdfs with the cellulose surface, like PAA-OE with O6 or O2.

A second – unexpected – behavior is shown by the simulations of CMC and PAA with the narrow $1\bar{1}0$ cellulose surfaces: The crystal is not stable, but undergoes some deformation forming a kink on the water-cellulose interface. The transition from the intact to the disturbed crystal is abrupt and occurs in less than 100 ps. However, once the crystal geometry is changed it is again stable for several nanoseconds, with the undisturbed patches of the crystal much like intact ones. The distortion is accompanied by a change in pressure (p_{yy}, p_{zz}, which raises from strong negative but keeps below zero after the transition, while the system escapes into a state of lower nonbonded energy.

6 Discussion

In this work, aqueous polyelectrolyte-solutions of (carboxy methyl) cellulose (CMC) and poly(acrylate) (PAA) have been investigated.

With respect to the size and CH_2COO^--distribution pattern, two aqueous solutions of different CMC oligomers (one heptamer and octamer) result in two chain structures: We observe one stretched structure – which is for a polyanion is rationalized by repulsion of negative charges – and a globule like, collapsed structure. The compact structure is held together by intramolecular hydrogen bonds, which bridge multiple anhydroglucose units and often involves COO^--groups. The carboxy methyl side groups are strong H-acceptors, because of their flexibility from the carbohydrate backbone and because of their negative charge. The globular CMC conformation is both stable through a multi-nanoseconds simulation and builds dynamically from a stretched starting geometry. On a local scale, the globular conformation undergoes less hydrogen bonding with the solvent, as more intramolecular H-bonds are present and some H-donor and acceptor sites are buried inside the globule and are not accessible to water. This is also visible in the CMC-water radial distribution functions, where the globule state has a more distorted and irregular hydration shell.

The PAA oligomer in aqueous solution is stretched ($C_\infty = 8.3$) and is readily solvated by water molecules. Because PAA has a high charge density on COO^--groups. The mass of solvation water exceeds the polymer's own mass by a factor of 1.5.

The hydrogen bonding for CMC is more complex. There is inter- and intramolecular hydrogen bonding, latter one can be sub-divided into inter- and intra glucose-ring bonding. For the collapsed, globule CMC molecules all kinds of H-bonds are found in a significantly amount. Most remarkable are H-bonds spanning six or seven glucose-rings, thus closing the CMC-chain to a ring.

There are only few counterions close to the CMC-strand (about 0.8 Na^+-ions/AGU). The PAA-sodium radial distribution function is better defined due to the higher number of counterions. There are about 1.4 sodium ions under the first peak of the PAA-Na^+-rdf. The difference between CMC and PAA is understood, as the main interaction of sodium takes place with the carboxylic side groups of the polymers, which favors PAA.

Counterions play a role during the (dynamic) collapse of the second CMC molecule. Simultaneously with the collapse the number of sodium-ions close to the backbone raises. The increased concentration of positive charges screens the repulsion of COO^--groups and initiates the collapse.

Two crystal planes (110 and $1\bar{1}0$) of the monoclinic cellulose crystal were simulated for several nanoseconds with an interface to water. Both surfaces are representative for other surfaces like the triclinic ones [49]. The 110 surface has a wider inter-chain spacing than the $1\bar{1}0$ surface. Both surfaces are stable against water and they are not penetrated by the solvent. This is due to the mainly hydrophobic and lipophilic character of both surfaces. This property has been accessed either through water densities on the surface

and particle insertion with a test lipophile (argon). Even if there are hydrogen bonds with the surface (involving atoms O2, O6 and to lesser degree O3), the cellulose surface does not attract water. Water close to the surface is structured through the surface and there is a large amount of water molecules being "flat" on the surface (thus all three sites have the same distance from the surface). The water density has only a small maximum close to the surface, which is observed for hydrophobic hydration by other authors. [40].

The void space between cellulose and water is able to accommodate some small lipophile like argon. The chemical potential of argon thus resembles much of the underlying structure of the cellulose crystal – the single cellobiose units and even glucose rings are identified and the resulting patterns are in accordance to AFM images [5–7]. Most of all the protruding CH_2O6H groups are visible, as they create excluded space for both lipophile test particles but also for water molecules close to the surface. The second dominant feature on the cellulose-water interface are trenches between the cellulose-chains on the surfaces. These trenches are lipophilic and allow for the insertion of argon test lipophile. They are larger on the wide 110 surface than on the narrow $1\bar{1}0$ surface allowing for a slightly larger argon concentration. Nevertheless, the differences are minor.

Simulations with both smaller oligomers of (carboxy methyl) cellulose and poly(acrylate) in aqueous solution in the presence of either the wide 110 or the narrow $1\bar{1}0$ surface were performed to investigate the adsorption properties directly. While the 110-crystal was stable upon the addition of an polyelectrolyte, the $1\bar{1}0$ crystal underwent a transition, leaving behind the intact crystal structure and forming a defect on the surface. This defect can best be described as a kink, running parallel to the chain direction (x-axis). During the deformation, both the pressure raises and the nonbonded energy of the whole system drops.

(Carboxy methyl) cellulose adsorbs onto both surfaces. On the 110 surface, all CMC molecules deposit. One trimer is flat on the surface, and takes a conformation above one cellobiose chain of the crystal. All other oligomers have a "tip-to-the-surface" conformation with one or two glucose units in contact (H-bonds) with the crystal, whereas the other ends point towards the bulk water. The main interaction is hydrogen bonding between cellulose and CMC.

PAA molecules stay solvated in the bulk, but aggregate together with sodium ions. Only one ion stays away from the PAA-aggregate and adsorbs onto the surface, while forming few hydrogen bonds with cellulose.

Finally there are three main results from this thesis:

- (Carboxy methyl) cellulose has a high affinity to cellulose or cellulose like molecules like CMC itself. The main interaction – may it be intermolecular or intramolecular – is hydrogen bonding; PAA behaves like a classical polyelectrolyte with a strong hydration shell and a elongated chain conformation.

- Monoclinic cellulose crystal surfaces (110 and $1\bar{1}0$) are mainly hydrophobic, but undergo hydrogen bonding with water. The H-bonding is dominated by oxygen O2 and O6 of the anhydro glucose repeat unit;

- (Carboxy methyl) cellulose adsorbs onto monoclinic cellulose, and PAA does not. The adsorbtion happens via hydrogen bonds, where mainly O2 and O6 groups are involved from the surface's side and COO^- on the polyelectrolyte's side.

	CMC	PAA
Solvation properties in water	CMC is well hydrated in water. COO^-- and OH-groups are responsible for keeping the cellobiose backbone into solution (cellulose itself is water insoluble) by forming H-bonds with the solvent. The modest hydration sphere does not interfere with the adsorption of CMC on cellulose.	PAA is well hydrated by the COO^--water H-bonds. The hydration is very strong and cannot waived easily. The water hydration sphere shields PAA against interaction with cellulose.
Lipophilicity	CMC has some lipophilic regions and patches and therefore has one additional interaction with cellulose, which is absent in PAA.	PAA has hydrophilic properties throughout, which is governed by the strong COO^--water interaction.
Hydrogen-bonding	CMC undergoes both donor- and acceptor hydrogen bonds with the cellulose surface. As a) the strength of single bonds and b) the absolute number of bonds is important for a molecules adsorption strength this favors adsorption of CMC. The CMC-CMC hydrogen bonds are likely to be transferable to CMC-cellulose bonds.	PAA (poly acrylate) does not exhibit hydrogen-donor groups. In the unlike event of a hydrogen bond between PAA and the cellulose surface this bond is very stable and lasting. Nevertheless bonds are very rare.
Aggregation	CMC is able to form inter- and intramolecular hydrogen bonds with compounds derived from cellobiose-structures. This leads to some aggregation and folding and makes CMC as well a good adsorbat onto cellulose surfaces.	At higher concentration short PAA oligomers do aggregate in solution. This aggregation prevents adsorption with the cellulose surface, as the COO^--contacts are saturated within the aggregate. PAA does not exhibit any cellulose-like patterns and thus does not absorb well.

Table 6.1: Summary of CMC and PAA properties and their implication on the polyelectrolyte adsorption on cellulose.

An overview of properties of PAA and CMC and implications on the polyelectrolyte's adsorption properties is given in table 6.1.

The hydration sphere of PAA is very strong and cannot waived without an high enthalpic penalty. Water in the vicinity of CMC is not bound that strong, even if the COO^--groups assist to solve oligomers in water. However, the amount of water, which bounds to OH-groups is only modest, there is no "water-shield" around CMC, prohibiting adsorption as for PAA.

The molecular shape of CMC is commensurable with cellulose (cellobiose backbone), this is obvious from the adsorption geometry of some oligomers. PAA does not exhibit any pattern, commensurable with cellulose and it is not able to form donor hydrogen bonds with cellulose. Hydrogen bonds from CMC to cellulose were found to be important for the adsorption as well as lipophilic contacts between CMC and cellulose.

Appendix A: Cellulose-Water Interface Simulations

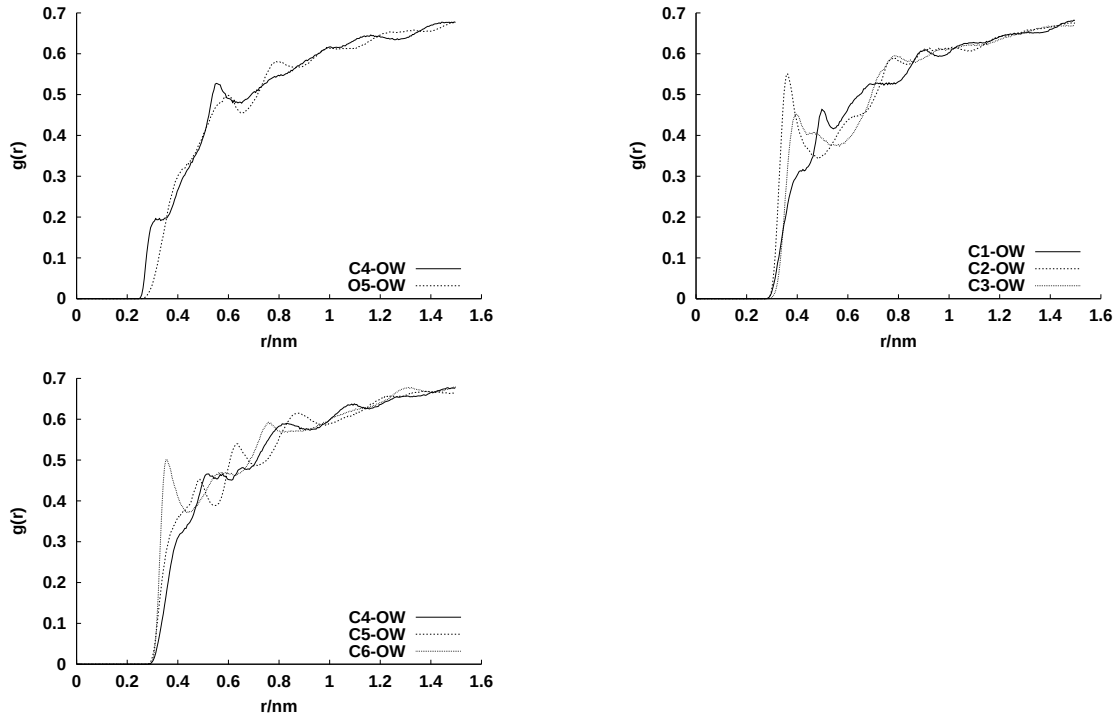


Figure A.1: Radial distribution functions for oxygen O2, O3, O4, O5 and O6 of the cellulose surface with water oxygens for the 110 surface.

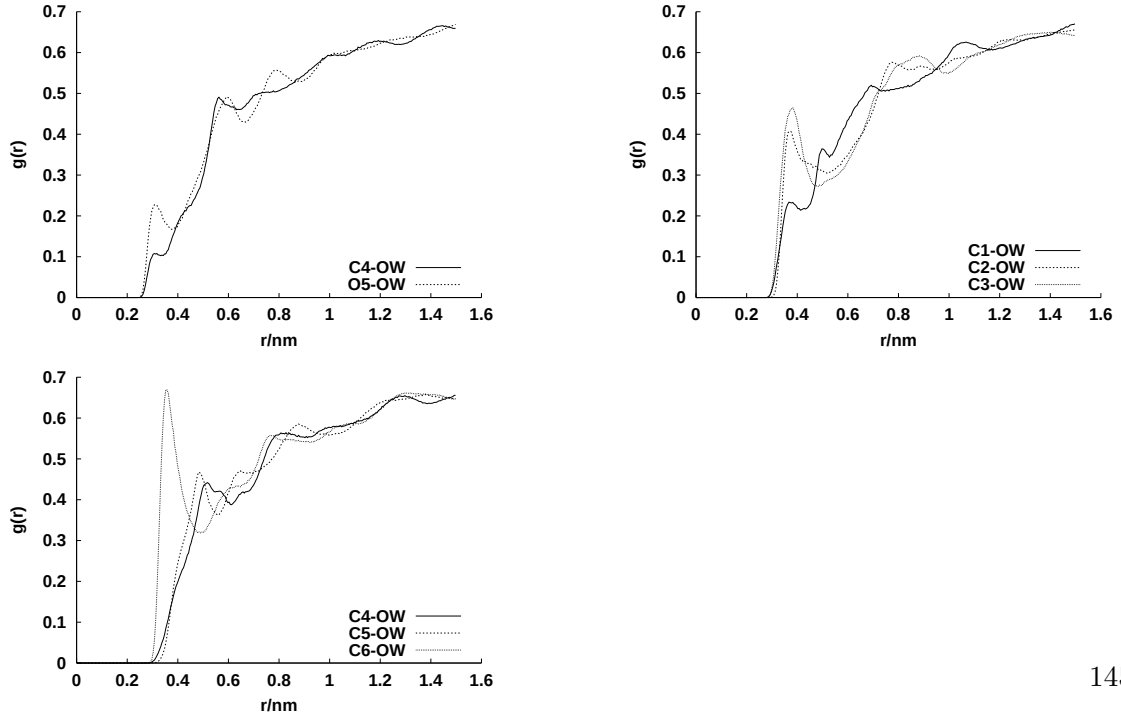


Figure A.2: Radial distribution functions for oxygen O2, O3, O4, O5 and O6 of the cellulose surface with water oxygens for the $1\bar{1}0$ surface.

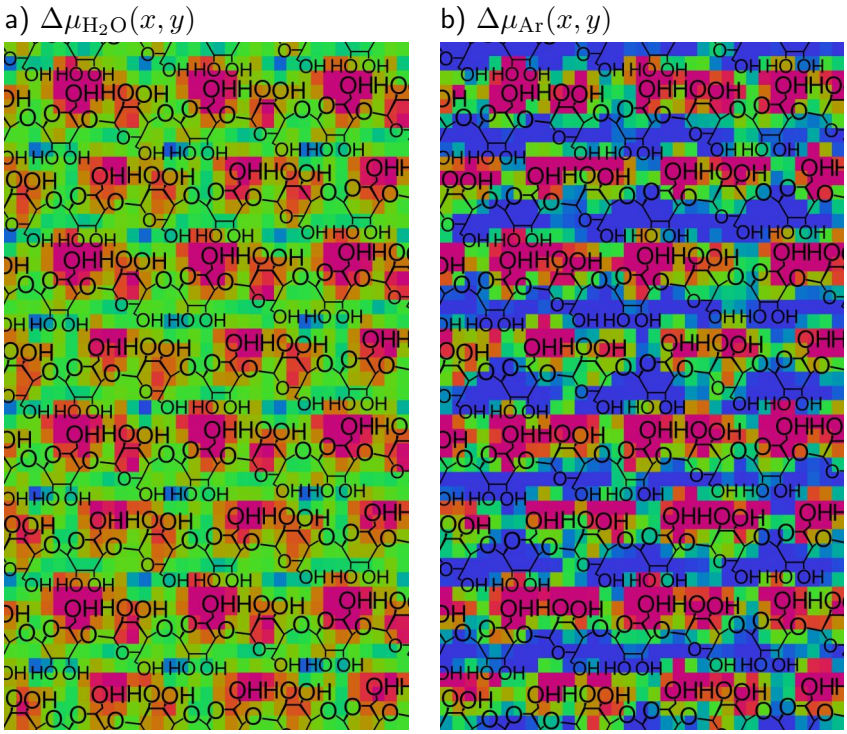


Figure A.3: Hydrophilicity and lipophilicity charts of the monoclinic cellulose 110 surface in a layer between $z = 0\,\text{nm}$ and $z = 0.2\,\text{nm}$. The structure formulae show the underlying structure of the respective cellulose atoms.

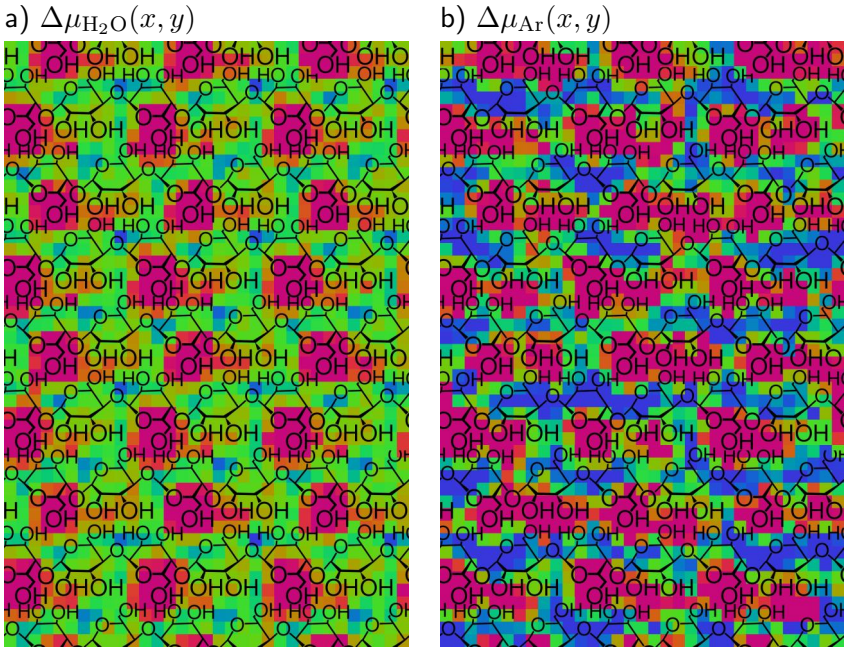


Figure A.4: Hydrophilicity and lipophilicity charts of the monoclinic cellulose $1\bar{1}0$ (narrow) surface in a layer between $z = 0\,\text{nm}$ and $z = 0.2\,\text{nm}$. The structure formulae show the cellulose structure. a) $\Delta\mu_{H_2O}(x, y)$, b) $\Delta\mu_{Ar}(x, y)$.

Appendix B: Polyelectrolyte-Cellulose-Water Interface Simulations

	O2	O3	O4	O5	O6	OE	O2	O3	O4	O5	O6	OE
	Trim.#1 to Trim.#1											
O2H	0.000	0.000	0.001	0.000	0.058	0.000						
O3H	0.000	0.000	0.001	0.277	0.028	0.254						
O6H	0.067	0.115	0.077	0.000	0.000	0.833						
	Trim.#1 to Tetram.#1						Tetram.#1 to Trim.#1					
	No hydrogen bonds at all											
	Trim.#1 to Trim.#2						Trim.#2 to Trim.#1					
	No hydrogen bonds at all											
	Trim.#1 to Tetram.#2						Tetram.#2 to Trim.#1					
	No hydrogen bonds at all											
	Tetram.#1 to Tetram.#1											
O2H	0.005	0.009	0.002	0.002	0.033	0.081						
O3H	0.003	0.000	0.000	0.048	0.230	0.237						
O6H	0.055	0.073	0.012	0.000	0.005	0.808						
	Tetram.#1 to Trim.#2						Trim.#2 to Tetram.#1					
O2H	0.000	0.003	0.000	0.000	0.000	0.220	0.000	0.007	0.002	0.000	0.004	0.328
O3H	0.003	0.000	0.000	0.000	0.001	0.138	0.000	0.002	0.069	0.000	0.008	0.347
O6H	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.024	0.021	0.000	0.003	0.099
	Tetram.#1 to Tetram.#2						Tetram.#2 to Tetram.#1					
O2H	0.000	0.008	0.001	0.000	0.062	0.235	0.000	0.002	0.003	0.000	0.000	0.668
O3H	0.012	0.002	0.000	0.001	0.013	0.921	0.000	0.001	0.000	0.000	0.003	0.826
O6H	0.001	0.001	0.003	0.000	0.023	0.318	0.006	0.006	0.007	0.000	0.035	0.557
	Trim.#2 to Trim.#2											
O2H	0.000	0.000	0.003	0.004	0.040	0.000						
O3H	0.000	0.000	0.000	0.101	0.014	0.000						
O6H	0.051	0.107	0.063	0.000	0.000	0.259						
	Trim.#2 to Tetram.#2						Tetram.#2 to Trim.#2					
O2H	0.000	0.000	0.000	0.007	0.004	0.719	0.000	0.000	0.025	0.001	0.000	0.556
O3H	0.000	0.019	0.002	0.073	0.015	0.580	0.001	0.000	0.016	0.001	0.006	0.694
O6H	0.003	0.007	0.000	0.000	0.001	0.305	0.029	0.004	0.008	0.003	0.000	0.057
	Tetram.#2 to Tetram.#2											
O2H	0.000	0.007	0.003	0.040	0.160	0.029						
O3H	0.000	0.000	0.000	0.044	0.033	0.475						
O6H	0.092	0.072	0.022	0.000	0.000	0.543						

Table B.1: Table with all hydrogen bonds between CMC molecules in the CMC-110 system.

	CMC atom types													
	C1	C2	C3	C4	C5	C6	CE	CH2	O2	O3	O4	O5	O6	OE
C1	0.0008	0.0057	0.0043	0.0000	0.0001	0.0006	0.0009	0.0016	0.0171	0.0043	0.0072	0.0012	0.0067	0.0062
C2	0.0071	0.0069	0.0053	0.0032	0.0049	0.0044	0.0057	0.0043	0.0135	0.0055	0.0225	0.0080	0.0132	0.0230
C3	0.0009	0.0015	0.0023	0.0001	0.0010	0.0003	0.0035	0.0019	0.0083	0.0044	0.0125	0.0032	0.0054	0.0128
C4	0.0016	0.0025	0.0052	0.0005	0.0002	0.0002	0.0010	0.0046	0.0093	0.0028	0.0126	0.0014	0.0064	0.0054
C5	0.0079	0.0081	0.0021	0.0036	0.0017	0.0023	0.0027	0.0052	0.0219	0.0112	0.0103	0.0032	0.0090	0.0124
C6	0.0054	0.0052	0.0022	0.0064	0.0022	0.0018	0.0110	0.0109	0.0185	0.0150	0.0140	0.0054	0.0085	0.0302
O2	0.0077	0.0044	0.0063	0.0031	0.0099	0.0129	0.0179	0.0055	0.0109	0.0095	0.0161	0.0135	0.0166	0.0319
O3	0.0130	0.0122	0.0147	0.0076	0.0107	0.0055	0.0086	0.0118	0.0142	0.0121	0.0197	0.0077	0.0081	0.0128
O4	0.0033	0.0080	0.0040	0.0015	0.0023	0.0024	0.0063	0.0057	0.0126	0.0051	0.0090	0.0048	0.0093	0.0206
O5	0.0033	0.0080	0.0022	0.0001	0.0002	0.0007	0.0026	0.0054	0.0150	0.0052	0.0100	0.0012	0.0026	0.0049
O6	0.0211	0.0158	0.0091	0.0113	0.0123	0.0117	0.0264	0.0192	0.0275	0.0199	0.0248	0.0146	0.0141	0.0397

Table B.2: Integrals of radial distribution functions $g(r)$ of cellulose (110) surface atoms (data for a given cellulose-atom organized in rows) with CMC-oligomer atoms (data given in columns). Integrals (see equation (1.31) recorded at $r = 0.395\,\text{nm}$.

	cellulose atom types										
	C1	C2	C3	C4	C5	C6	O2	O3	O4	O5	O6
C1	0.0052	0.0486	0.0061	0.0111	0.0544	0.0372	0.0531	0.0891	0.0228	0.0228	0.1447
C2	0.0390	0.0475	0.0105	0.0175	0.0555	0.0353	0.0301	0.0836	0.0546	0.0552	0.1082
C3	0.0295	0.0361	0.0161	0.0356	0.0141	0.0149	0.0434	0.1010	0.0276	0.0148	0.0624
C4	0.0000	0.0219	0.0006	0.0034	0.0244	0.0437	0.0213	0.0521	0.0102	0.0005	0.0773
C5	0.0006	0.0333	0.0067	0.0012	0.0117	0.0150	0.0680	0.0736	0.0157	0.0012	0.0843
C6	0.0043	0.0304	0.0021	0.0017	0.0155	0.0120	0.0887	0.0379	0.0167	0.0048	0.0804
CE	0.0087	0.0549	0.0337	0.0093	0.0263	0.1058	0.1719	0.0830	0.0609	0.0251	0.2538
CH2	0.0151	0.0417	0.0178	0.0443	0.0501	0.1049	0.0524	0.1135	0.0543	0.0519	0.1840
O2	0.1174	0.0928	0.0569	0.0636	0.1504	0.1269	0.0745	0.0971	0.0863	0.1026	0.1885
O3	0.0293	0.0374	0.0301	0.0193	0.0765	0.1026	0.0649	0.0829	0.0352	0.0354	0.1368
O4	0.0383	0.1202	0.0667	0.0671	0.0547	0.0749	0.0858	0.1050	0.0479	0.0532	0.1324
O5	0.0080	0.0547	0.0222	0.0093	0.0221	0.0371	0.0928	0.0530	0.0328	0.0082	0.1000
O6	0.0457	0.0902	0.0368	0.0441	0.0619	0.0583	0.1139	0.0557	0.0636	0.0175	0.0967
OE	0.0299	0.1102	0.0615	0.0259	0.0594	0.1452	0.1531	0.0615	0.0988	0.0237	0.1905

Table B.3: Integrals of radial distributions $g(r)$ of CMC atoms (data organized in rows) with monoclinic 110-cellulose (data in columns). Integrals taken in a distance of $r = 0.395\,\text{nm}$.

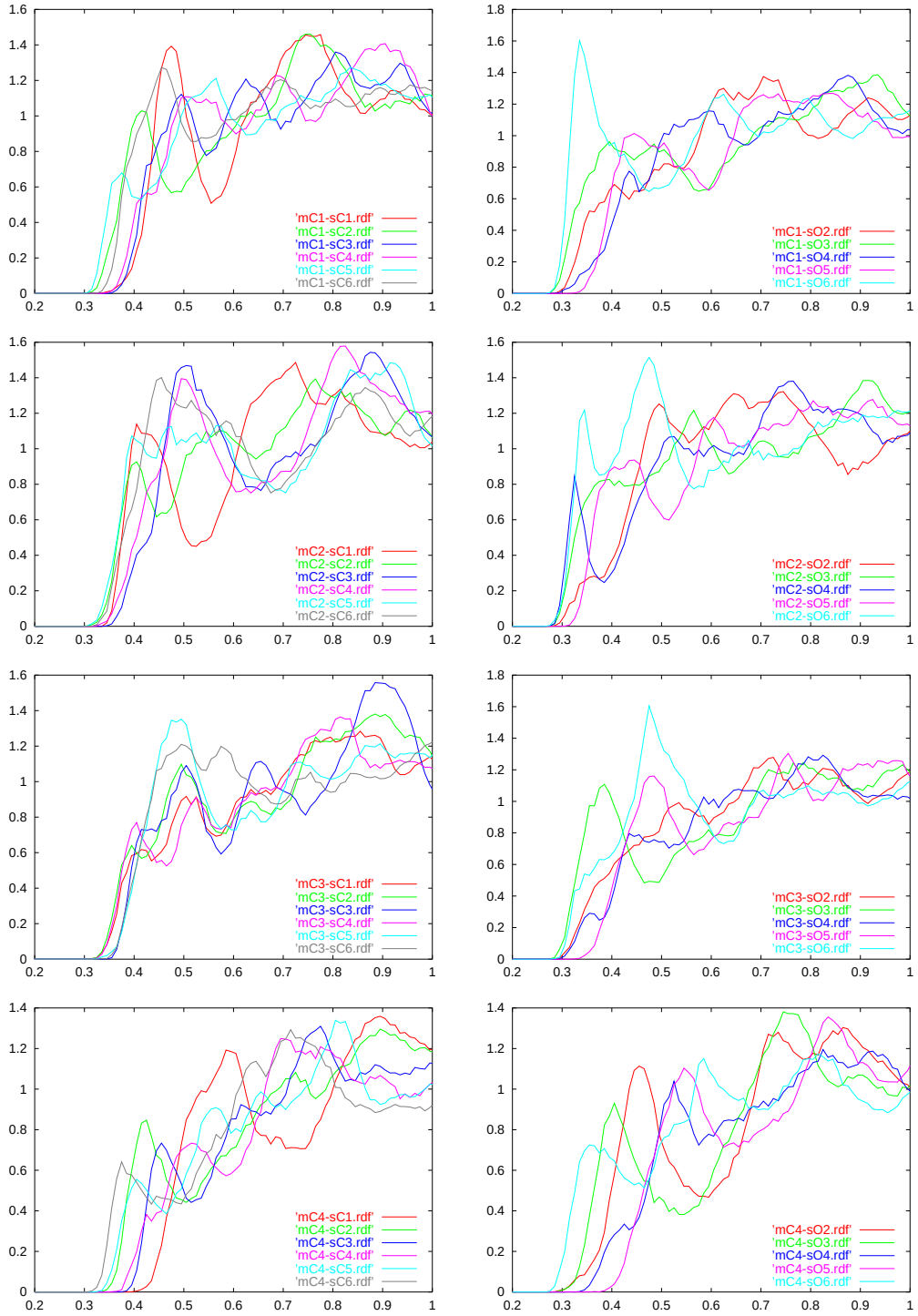


Figure B.1: Part I. Radial distribution functions (x-axis: r, y-axix: $g(r)$) between oligomer and surface atom sites for the (carboxy methyl)cellulose 110-surface simulation. The key shows the respective atoms: the prefix 'm' denotes the oligomer atom type and 's' the surface atom type. Due to the normalization of all rdf to the total number of sites for all oligomers, the limit of $g(r)$ at $r \to \infty$ depends on the number of adsorbed polyelectrolyte oligomers.

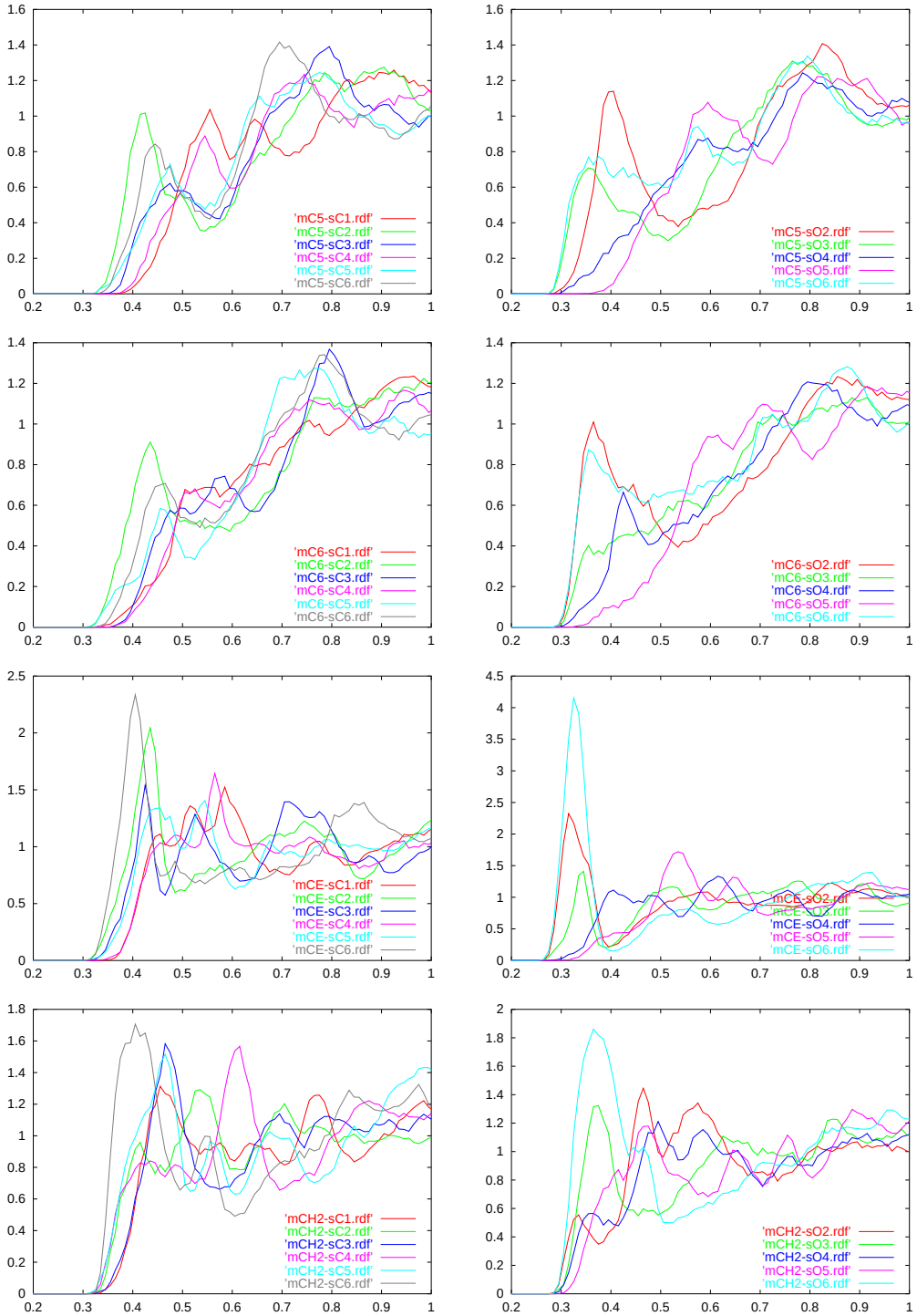


Figure B.2: Part II. Radial distribution functions (x-axis: r, y-axix: $g(r)$) between oligomer and surface atom sites for the (carboxy methyl)cellulose 110-surface simulation. The key shows the respective atoms: the prefix 'm' denotes the oligomer atom type and 's' the surface atom type. Due to the normalization of all rdf to the total number of sites for all oligomers, the limit of $g(r)$ at $r \to \infty$ depends on the number of adsorbed polyelectrolyte oligomers. (CE: carboxylic carbon (C)OO , CH_2 : CH_2 -oxygen of the $-CH_2COO^-$ -group.

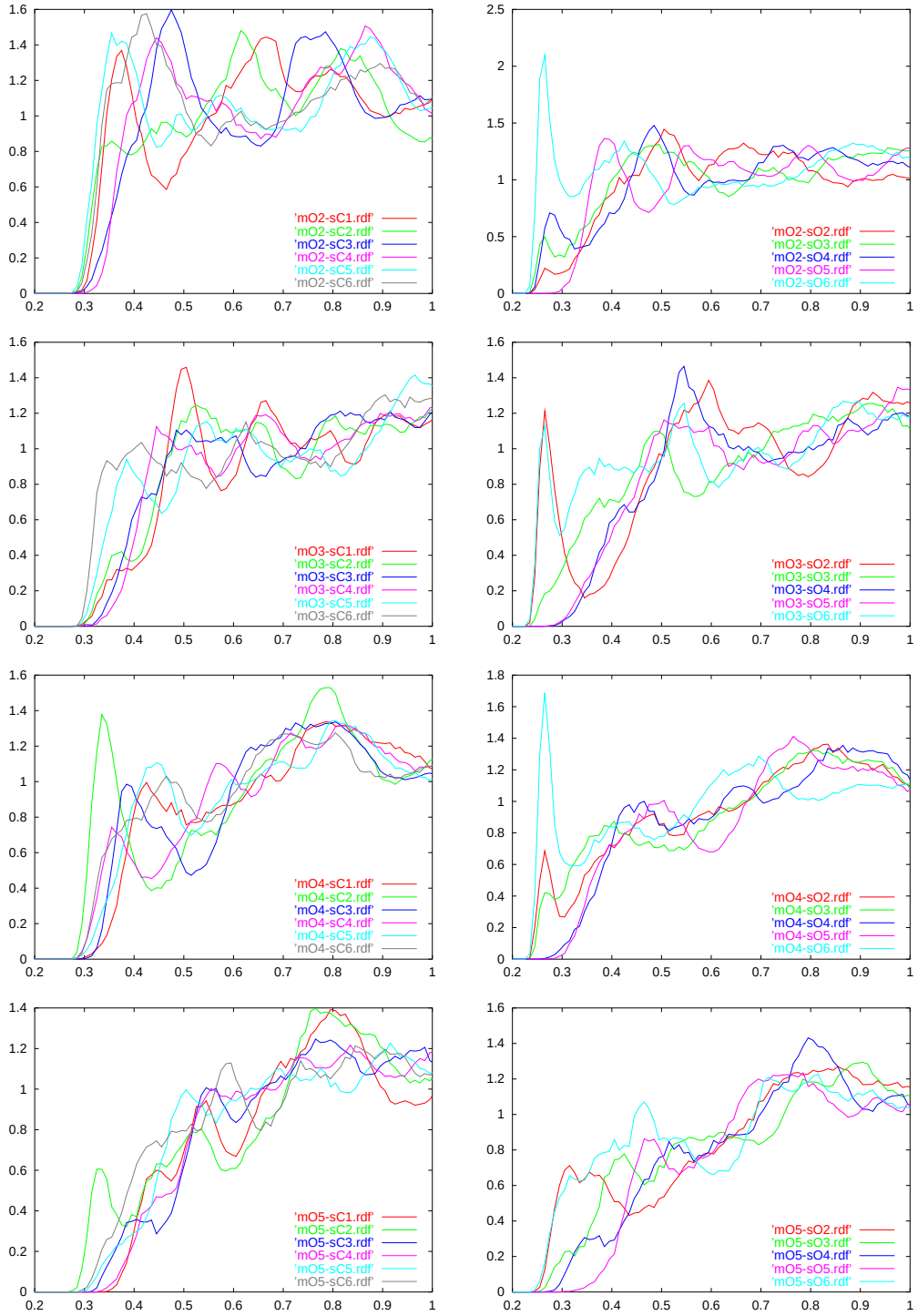


Figure B.3: Part III. Radial distribution functions (x-axis: r, y-axix: $g(r)$) between oligomer and surface atom sites for the (carboxy methyl)cellulose 110-surface simulation. The key shows the respective atoms: the prefix 'm' denotes the oligomer atom type and 's' the surface atom type. Due to the normalization of all rdf to the total number of sites for all oligomers, the limit of $g(r)$ at $r \to \infty$ depends on the number of adsorbed polyelectrolyte oligomers.

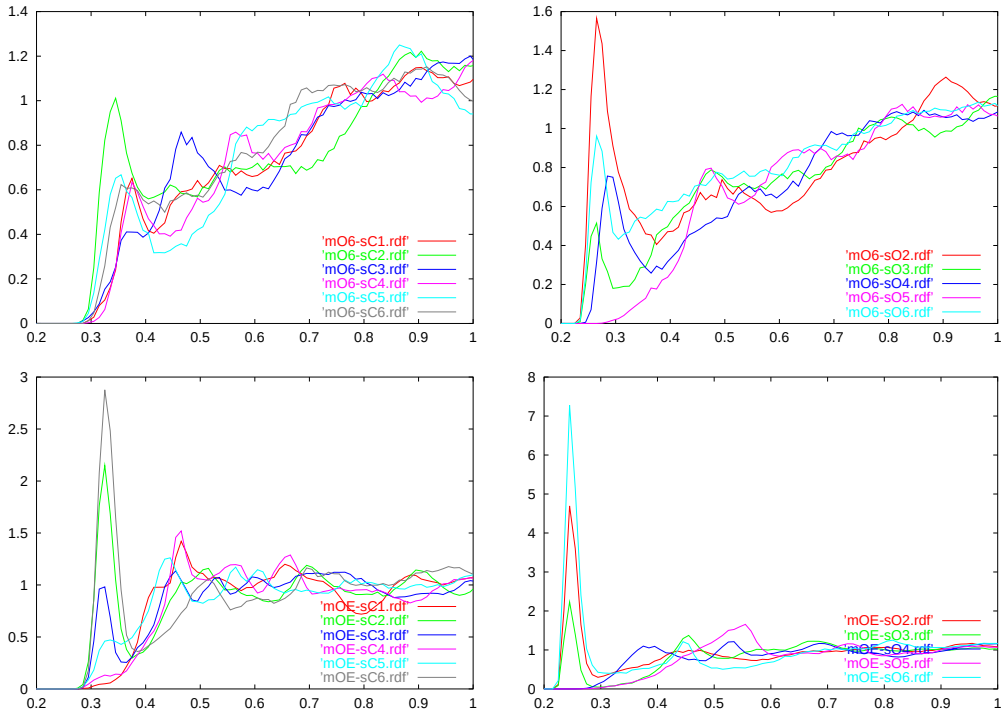


Figure B.4: Part IV. Radial distribution functions (x-axis: r, y-axix: $g(r)$) between oligomer and surface atom sites for the (carboxy methyl)cellulose 110-surface simulation. The key shows the respective atoms: the prefix 'm' denotes the oligomer atom type and 's' the surface atom type. Due to the normalization of all rdf to the total number of sites for all oligomers, the limit of $g(r)$ at $r \to \infty$ depends on the number of adsorbed polyelectrolyte oligomers. (OE: carboxylic oxygen C(OO)

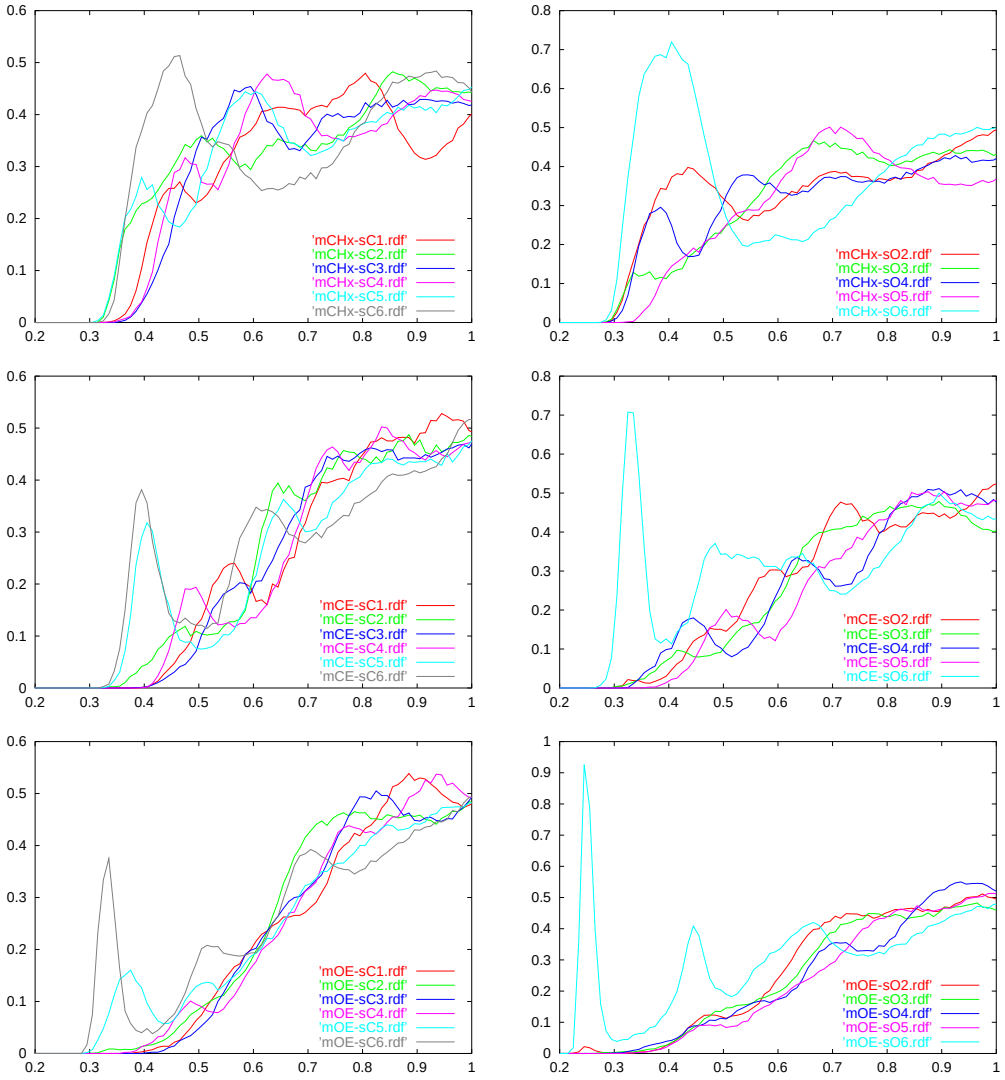


Figure B.5: Part V. Radial distribution functions (x-axis: r, y-axix: $g(r)$) between poly(acrylic acid) pentamers and cellulose 110-surface atoms. The key shows the respective atoms or atom groups: the prefix 'm' denotes the oligomer atom type and 's' the surface atom type. The atom type CHx denotes CH , CH_2 and CH_3 sites. OE is the carboxylic oxygen C(OO) and CE the carboxylic sp_2 carbon atom. Due to the normalization of all rdf to the total number of sites for all oligomers, the limit of $g(r)$ at $r \to \infty$ depends on the number of adsorbed polyelectrolyte oligomers.

	PAA atom types		
	CHx	CE	OE
C1	0.0021	0.0000	0.0001
C2	0.0096	0.0005	0.0006
C3	0.0007	0.0000	0.0000
C4	0.0009	0.0000	0.0002
C5	0.0110	0.0036	0.0083
C6	0.0122	0.0058	0.0122
O2	0.0176	0.0007	0.0009
O3	0.0084	0.0013	0.0010
O4	0.0155	0.0016	0.0016
O5	0.0036	0.0001	0.0005
O6	0.0424	0.0135	0.0172

Table B.4: Integrals of radial distribution functions $g(r)$ of cellulose (110) surface atoms (data for a given cellulose-atom organized in rows) with CMC-oligomer atoms (data given in columns). Integrals (see equation (1.31) recorded at $r = 0.395\,\text{nm}$.

	cellulose atom types										
	C1	C2	C3	C4	C5	C6	O2	O3	O4	O5	O6
CHx	0.0034	0.0153	0.0011	0.0014	0.0175	0.0196	0.0281	0.0134	0.0247	0.0058	0.0678
CE	0.0000	0.0016	0.0000	0.0000	0.0117	0.0185	0.0021	0.0040	0.0053	0.0004	0.0431
OE	0.0001	0.0009	0.0000	0.0004	0.0133	0.0195	0.0015	0.0017	0.0026	0.0009	0.0275

Table B.5: Integrals of radial distributions $g(r)$ of PAA atoms (data organized in rows) with monoclinic 110-cellulose (data in columns). Integrals taken in a distance of $r = 0.395\,\text{nm}$.

Bibliography

[1] A. Aabloo and A. D. French. *Preliminary potential energy calculations of cellulose Iα crystall structure. Macromol Theory Simul*, 3, 185–91, 1994.

[2] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids.* Clarendon, Oxford, U.K., 1997.

[3] R. H. Atalla and D. L. VanderHart. *Native cellulose: A composit of two distinct crystalline forms. Science*, 223, 283–4, 1984.

[4] A. Baar, W. M. Kulicke, K. Szablikowski and R. Kiesewetter. *Nuclear-magnetic-resonance spectroscopic characterization of carboxymethylcellulose. Macromol Chem Phys*, 195(5), 1483–92, 1994.

[5] A. A. Baker, W. Helbert, J. Sugiyama and M. J. Miles. *High resolution atomic force microscopy of native valonia cellulose I microcrystals. J Structural Biology*, 119, 129–38, 1997.

[6] A. A. Baker, W. Helbert, J. Sugiyama and M. J. Miles. *Surface structure of native cellulose microcrystals by AFM. Appl Phys A*, 66, S559–S563, 1998.

[7] A. A. Baker, W. Helbert, J. Sugiyama and M. J. Miles. *New insight into cellulose structure by atomic force microscopy shows the Iα crystal phase at near-atomic resolution. Biophys J*, 79, 1139–45, 2000.

[8] J.-L. Barrat and J.-F. Joanny. *Theory of polyelectrolyte solutions. Adv Chem Phys*, 94, 1–66, 1996.

[9] J. Baschnagel, K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Müller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter and V. Tries. *Bridging the gap between atomistic and coarse-grained models of polymers: Status and perspectives. Adv Polym Sci*, 152, 41–156, 2000.

[10] H. J. C. Berendsen, J. R. Grigera and T. P. Straatsma. *The missing term in effective pair potentials. J Phys Chem*, 91, 6269–71, 1987.

[11] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak. *Molecuar dynamics with coupling to an external bath. J Chem Phys*, 81, 3684–90, 1984.

[12] H. J. C. Berendsen, J. P. Ryckaert and G. Ciccotti. *Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-Alkanes. J Comput Phys*, 23, 327–41, 1977.

[13] N. R. Bertoniere and S. H. Zeronian. *Chemical characterization of cellulose.* In R. H. Atalla (ed.), *The structures of cellulose – Characterization of the solid state*, ACS Symposium Series, pp. 254–71. American Chemical Society, 1985.

[14] P. Beyer and E. Nordmeier. *Ultracentrifugation, viscometry, pH, and dynamic light scattering studies of the complexation of ionene with poly(acrylic acid) and poly(methacrylic acid). European Polymer Journal*, 35(7), 1351–65, 1999.

[15] O. Biermann, E. Hädicke, S. Koltzenburg and F. Müller-Plathe. *Hydrophilicity and lipophilicity of cellulose crystal surfaces. accepted by Angew Chem*, 2001.

[16] O. Biermann, E. Hädicke, S. Koltzenburg, M. Seufer and F. Müller-Plathe. *Hydration of polyelectrolyte studied by molecular dynamics simulation. submitted to Macromolecules*, 2001.

[17] R. Bieshaar. *Free energies of aqueous solutions from molecular dynamics simulations.* Ph.D. thesis, University of Dortmund, Dep. of Chemistry., 1996.

[18] K. Bredereck, M. Gruber, A. Otterbach and F. Schulz. *Die Hydrogelstruktur von Cellulosefasern und ihre Bedeutung für Fasereigenschaften und Textilveredlung. Textilveredlung*, 31, 194–200, 1996.

[19] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus. *CHARMM: A program for macromolecular energy, minimization and dynamics calculations. J Comput Chem*, 4(187-217), 1983.

[20] H. B. Callen. *Thermodynamics and an Introduction to Thermostatistics.* John Wiley & Sons, 1985.

[21] S. N. K. Chaudhari, K. C. Gounden, G. Srinivasan and V. S. Ekkundi. *High resolution ^{13}C-NMR spectroscopy of sodium carboxy methyl cellulose. J Polym Sci Pol Chem*, 25, 337–42, 1987.

[22] M. Clark, R. D. Cramer III and N. Van Opdenbosch. *Validation of the general purpose tripos force field. J Comput Chem*, 10(8), 982–1012, 1989.

[23] M. L. Connolly. *Solven-accessible surfaces of proteins and nucleic acids. Science*, 221, 709–13, 1983.

[24] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. J. Merz, F. D. M, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman. *A second generation forcefield for the simulation of proteins nucleic acids and organic molecules. J Am Chem Soc*, 117, 1995.

[25] S. K. Cousins and R. M. Brown Jr. *Cellulose I micorfibril assembly: computational molecular mechanics energy analysis favours bonding by van der Waals forces as the initial step in crystallization. Polymer*, 36, 3885–8, 1995.

[26] R. M. Davis. *Analysis of dilute-solutions of (carboxymethyl)cellulose with the electrostatic wormlike chain theory. Macromolecules*, 24(5), 1149–55, 1991.

[27] P. G. de Gennes. *Scaling Concepts in Polymer Physics.* Cornel University Press, Ithaca, 1979.

[28] G. L. Deitrick, L. E. Scriven and H. T. Davis. *Efficient molecular simulation of chemical potentials. J Chem Phys*, 90, 2370–85, 198.

[29] M. Deserno and C. Holm. *How to mesh up ewald sums: A theoretical and numerical comparison of various paricle mesh routines. J Chem Phys*, 108(18), 7678–93, 1998.

[30] T. Erata, T. Shikano, Y. Shimizu, M. Takai and J. Hayashi. *NMR studies on the structure of cellulose 2-dimensional solid state NMR approach. Macromol Symp*, 99, 25–9, 1995.

[31] U. Essmann, L. Perera, M. Berkowitz, T. Darden, H. Lee and L. Pedersen. *A smooth particle mesh ewald method. J Chem Phys*, 103(19), 8577–93, 1995.

[32] R. Faller, H. Schmitz, O. Biermann and F. Müller-Plathe. *Automatic parameterization of force fields for liquids by simplex optimization. J Comput Chem*, 20, 1009–17, 1999.

[33] *DEC Fortran user's manual*, 1998.

[34] A. D. French and J. W. Brady (eds.). *Computer Modeling of Carbohydrate Molecules*, vol. 430 of *ACS Symposium Series*. American Chemical Society, Washington DC, 1990.

[35] A. D. French and I. G. Csizmadia (eds.). *Carbohydrate Modeling*, vol. 395-396 of *J Mol Struc (Theochem)*, pp. 1–515. Elsevier Science, 1997.

[36] A. D. French, D. P. Miller and A. Aabloo. *Miniature crystal models of cellulose polymorphs and other carbohydrates. Int J Biol Macromol*, 15, 30–36, 1993.

[37] J. Ganster and H.-P. Fink. *Polymer Handbook*, chap. "Physical Constants of Cellulose", p. V/135. John Wiley & Sons, 4th edn., 1999.

[38] K. Gardner and J. Blackwell. *The structure of native cellulose. Biopolymers*, 13, 1975–2001, 1974.

[39] T. M. Glennon and K. M. Merz Jr. *A carbohydrate force field for AMBER and its application to the study of saccharide to surface adsorption. J Mol Struc Theochem*, 395-396, 157–71, 1997.

[40] J. R. Grigera, S. G. Kalko and J. Fischbarg. *Wall-water interface. A molecular dynamics study. Langmuir*, 12, 154–8, 1996.

[41] B. Guillot and Y. Guissani. *A computer simulation study of the liquid-vapor coexistence curve of water. J Chem Phys*, 98, 8221–35, 1993.

[42] B. Guillot and Y. Guissani. *A computer simulation study of the temperature dependence of the hydrophobic hydration. J Chem Phys*, 99, 8075–94, 1993.

[43] E. Haedicke. private communication. BASF Ludwigshafen.

[44] S. J. Hanley, J. Giasson, J.-F. Revol and D. G. Gray. *Atomic force microscopy of cellulose microfibrils: comparison with transmission electron microscopy. Polymer*, 33, 4639–42, 1992.

[45] B. J. Hardy and A. Sarko. *Conformational analysis and molecular dynamics simulation of cellobiose and larger cellooligomers. J Comput Chem*, 14, 831–47, 1993.

[46] B. J. Hardy and A. Sarko. *Molecular-dynamics simulation of cellobiose in water. J Comput Chem*, 14, 848–57, 1993.

[47] B. J. Hardy and A. Sarko. *Molecular dynamics simulations and diffraction-based analysis of the native cellulose fibre: Structural modeling of the I-α and I-β phases and their interconversion. Polymer*, 37, 1833–39, 1996.

[48] F. M.-P. Heiko Schmitz, Roland Faller. *Yasp documentation homepage*, 2000. URL `http://www.mpip-mainz.mpg.de/~yasp/`.

[49] A. P. Heiner, L. Kuutti and O. Teleman. *Comparison of the interface between water and four surfaces of native crystalline cellulose by molecular dynamics simulations. Carbohyd Res*, 306, 205–20, 1998.

[50] A. P. Heiner, J. Sugiyama and O. Teleman. *Crystalline cellulose Iα and Iβ studied by molecular dynamics simulation. Carbohyd Res*, 273, 207–23, 1995.

[51] A. P. Heiner and O. Teleman. *Structural reporter parameter for the characterisation of crystalline cellulose. Pure Appl Chem*, 68(11), 2187–92, 1996.

[52] A. P. Heiner and O. Teleman. *Interface between monoclinic crystalline cellulose and water: Breakdown of the odd/even duplicity. Langmuir*, 13, 511–8, 1997.

[53] T. Heinze, U. Erler, I. Nehls and D. Klemm. *Determination of the substituent pattern of heterogeneously and homogeneously synthesized carboxymethyl cellulose by using high-performance liquid chromatography. Angew Makromol Chem*, 215, 93–106, 1994.

[54] T. Heinze and U. Heinze. *The first approach to non-aqueous solutions of carboxymethylcellulose. Macromol Rapid Commun*, 18, 1033–40, 1997.

[55] T. Heinze and K. Rahn. *Cellulose-p-toluenesulfonates: a valuable intermediate in cellulose chemistry. Macromol Symp*, 120, 103–13, 1997.

[56] C. Hoogendam, A. Dekeizer, M. Stuart, B. Bijsterbosch, J. Smit, J. Vandijk, P. Vanderhorst and J. Batelaan. *Persistence length of carboxymethyl cellulose as evaluated from size exclusion chromatography and potentiometric titrations. Macromolecules*, 31(18), 6297–6309, 9 1998.

[57] L. M. Ilharco, A. R. Garcia, J. L. da Silva and L. V. Ferreira. *Infrared approach to the study of adsorption on cellulose: Influence of cellulose crystallinity on the adsorption of benzophenone. Langmuir*, 13, 4126–32, 1997.

[58] S. Imam, R. V. Green and B. R. Zaidi (eds.). *Biopolymers – Utilizing Nature's Advanced Materials*. American Chemical Society, Washingto, 1998.

[59] M. Ioelovitch and M. Gordeev. *Crystallinity of cellulose and its accessibility during deuteration. Acta Polym*, 45, 121–3, 1994.

[60] J. L. Jackson and L. S. Klein. *Phys Fluids*, 7, 228, 1964.

[61] F. Jensen. *Introduction to Computational Chemistry*. Wiley VCH, 1998.

[62] B. Jönsson, C. Peterson and B. Söderberg. *A variational approach to the structure and thermodynamics of linear polyelectrolytes with coulomb and screend coulomb interactions. J Chem Phys*, 99, 1251, 1995.

[63] W. J. Jorgensen, J. D. Madura and C. J. Swenson. *Optimized intermolecular potential functions for liquid hydrocarbons. J Am Chem Soc*, 106(6638-46), 1984.

[64] W. L. Jorgensen. *Optimized intermolecular potential functions for liquid alcohols. J Phys Chem*, 90, 1276–84, 1986.

[65] U. Kästner, H. Hoffmann, R. Dönges and J. Hilbig. *Structure and solution properties of sodium carboxymethyl cellulose. Colloid Surface A*, 123, 307–28, 1997.

[66] P. M. King. *Free energy via molecular simulations: A primer*. In van Gunsteren *et al.* [134], p. 315.

[67] J. Klein and W. Scholz. *Kalorimetrische Untersuchungen zum Lösungsverhalten von Polyacrylsäure und Poly(natriumacrylat in Wasser, Formamid und Ethylenglycol). Makromol Chem*, 180, 1477–86, 1979.

[68] J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren. *The flip-flop hydrogen bonding phenomenon. Eur Biophys J*, 16, 153–68, 1988.

[69] J. E. H. Koehler, W. Saenger and W. F. van Gunsteren. *A molecular dynamics simulation of crystalline α-cyclodextrin hexahydrate. Eur Biophys J*, 15, 197–210, 1987.

[70] A. Konop and R. Colby. *Polyelectrolyte charge effects on solution viscosity of poly(acrylic acid). Macromolecules*, 32(8), 2803–5, 1999.

[71] M. L. C. E. Kouwijzer, B. P. van Eijck, S. J. Kroes and J. Kroon. *Comparison of two force fields by molecular dynamics simulations of glucose crystals: Effect of using ewald sums. J Comput Chem*, 14, 1281–89, 1993.

[72] H. Krässig, R. G. Steadman, K. Schliefer and W. A. Albrecht. *Cellulose.* In *Ullmann's encyclopedia of industrial chemistry*, vol. A5, pp. 375–418. 1986.

[73] S. Krishnaswami, D. Ramkrishna and J. M. Caruthers. *Statistical-mechanically exact simulation of polymer conformation in an external field. J Chem Phys*, 107, 5929–44, 1997.

[74] L. M. J. Kroon-Batenburg, B. Bouma and J. Kroon. *Stability of cellulose structures studied by MD simulations. Could mercerized cellulose II be parallel? Macromolecules*, 29, 5695–99, 1996.

[75] L. M. J. Kroon-Batenburg, P. H. Kruiskamp, V. J. F. G. and J. Kroon. *Estimation of the peristence length of polymers by MD simulation on small fragments in solution. Application to cellulose. J Phys Chem B*, 101, 8454–59, 1997.

[76] A. R. Leach. *Molecular Modelling.* Longman Ltd., 1996.

[77] I. N. Levine. *Quantum Chemistry.* Prentice Hall, 4th edn., 1991.

[78] H. Li, B. Liu, X. Zhang, C. Gao, J. Shen and G. Zou. *Single-molecule force spectroscopy on poly(acrylic acid) by AFM. Langmuir*, 15(6), 2120–4, 1999.

[79] D. R. Lide (ed.). *Handbook of Chemistry and Physics.* CRC Press, Boca Raton, 74th edn., 1993.

[80] T. Liebert and T. Heinze. *Induced phase preparation: A new synthesis concept in cellulose chemistry.* In T. J. Heinze and W. G. Glasser (eds.), *Cellulose derivatives, Modification, Characterization and Nanostructures*, ACS Symposium Series No. 688, pp. 61–72. American Chemical Society, Orlando, Florida, 1998.

[81] S. G. Manning. *J Chem Phys*, 51, 924, 934, 3249, 1969.

[82] R. J. Marhöfer, S. Reiling and J. Brickmann. *Computer simulations of crystal structures and elastic properties of cellulose. Ber Bunsenges Phys Chem*, 100, 1350–4, 1996.

[83] A. E. Mark, S. P. van Helden, P. E. Smith, L. H. M. Janssen and W. F. van Gunsteren. *Convergence properties of free energy calculations: α-cyclodextrin complexes as a case study. J Am Chem Soc*, 116, 6293–302, 1994.

[84] S. L. Mayo, B. D. Olafson and W. A. Goddard III. *DREIDING: a generic force field for molecular simulations. J Phys Chem*, 94, 8897–909, 1990.

[85] N. Metropolis, R. A. W, M. N. Rosenbluth *et al. Equation of state calculations by fast computing machines. J Chem Phys*, 21(6), 1087–92, 1953.

[86] H. Meyer, O. Biermann, R. Faller, D. Reith and F. Müller-Plathe;. *Coarse graining of nonbonded inter-particle potentials. J Chem Phys*, 2000.

[87] A. Mokrane, P. Friant-Michel, A. Catier and R. J. L. *Scaled semiempirical method for the calculation of vibrational spectra: Molecular vibrational frequencies of monosaccharides and disaccharides. J Mol Struc Theochem*, 395-396, 71–80, 1997.

[88] M. Müller, C. Czihak, H. Schober, Y. Nishiyama and G. Vogl. *All disordered regions of native cellulose show common low-frequency dynamics. Macromolecules*, 33, 1834–40, 2000.

[89] F. Müller-Plathe. *Parallelising a molecular dynamics algorithm on a multi-processor workstation. Comput Phys Commun*, 61, 285–93, 1990.

[90] F. Müller-Plathe. *Yasp: A molecular simulation package. Comput Phys Commun*, 78, 77–94, 1993.

[91] F. Müller-Plathe. *Unexpected diffusion behaviour of gases in crystalline poly(4-methyl-1-pentene). J Chem Phys*, 103, 4346–51, 1995.

[92] F. Müller-Plathe. *Different states of water in hydrogels. Macromolecules*, 31, 6721–3, 1998.

[93] F. Müller-Plathe and R. Brown. *Multi-colour algorithms in molecular simulation: Vectorisation and parallelisation of internal forces and constraints. Comput Phys Commun*, 64, 7–14, 1991.

[94] F. Müller-Plathe and W. van Gunsteren. *Solvation of poly(vinyl alcohol) in water, ethanol and an equimolar water-ethanol mixture: Structure and dynamics studied by molecular dynamics simulation. Polymer*, 38, 2259–68, 1997.

[95] Y. Muroga, I. Noda and M. Nagasawa. *Investitagion of local conformations of polyelectrolytes in aqueous solution by small-angle X-ray scattering 1. Local conformations of poly(sodium acrylate). Macromolecules*, 18, 1576–9, 1985.

[96] M. Neumann. *Dipole moment fluctuation formulas in computer simulations of polar systems. Mol Phys*, 50(4), 841–58, 1983.

[97] M. Neumann and O. Steinhauser. *The influence of boundary conditions used in machine simulations on the structure of polare systems. Mol Phys*, 39(2), 437–54, 1980.

[98] S. Neyertz, D. Brown and J. O. Thomas. *Molecular-dynamics simulation of the crystalline phase of poly(ethylene oxide)-sodium iodide, PEO(3)NAI. Electrochim Acta*, 40(13-14), 2063–9, 1995.

[99] S. Neyertz, A. Pizzi, A. Merlin, B. Maigret, D. Brown and X. Deglise. *A new all-atom force field for crystalline cellulose I. J Applied Polymer Science*, 78, 1939–46, 2000.

[100] S. Neyertz, J. O. Thomas and D. Brown. *Molecular-dynamics simulations of the amorphous polymer electrolyte PEO(X)NAL. Comput Polym Sci*, 5(3), 107–20, 1995.

[101] K. Niemela and E. Sjostrom. *Studies of carboxymethylcellulose. 2. Characterization of hardwood-derived carboxymethylcellulose by gas-liquid chromatography and mass-spectrometry. Polymer Com*, 30(8), 254–6, 1989.

[102] O. Nikolaeva, T. Budtova, Y. Brestkin, Z. Zoolshoev and S. Frenkel. *Rheological properties of an interpolymer complex formed between poly(acrylic acid) and methyl cellulose. J Applied Polymer Science*, 72(12), 1523–8, 1999.

[103] D. H. Paschek. *Molekulardynamic Simulation der hydrophoben Hydratation nichtionischer Tenside.* Ph.D. thesis, University of Dortmund, Dep. of Chemistry., 1998.

[104] B. Philipp. *Cellulose–new prospects for an old polymer. Polymer News*, 22, 205–8.

[105] B. Philipp, W. Wagenknecht, I. Nehls, D. Klemm, A. Stein and T. Heinze. *Regioselective derivatization of cellulose–routes of synthesis, effects on properties and areas of application. Polymer News*, 21, 155–61, 1996.

[106] A. K. Rappe, C. J. Casewit, C. W. A, W. A. Goddard III and W. M. Skiff. *UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. J Am Chem Soc*, 114, 10024–35, 1992.

[107] W. F. Reed, S. Ghosh, G. Medjahdi and J. Francois. *Dependence of polyelectrolyte apparent persistence lengths, viscosity, and diffusion on ionic strength and linear chain density. Macromolecules*, 24, 6189–98, 1991.

[108] S. Reiling and J. Brickmann. *Theoretical investigations on the structure and physical properties of cellulose. Macromol Theory Simul*, pp. 725–41, 1995.

[109] S. Reiling, M. Schlenkrich and J. Brickmann. *Force field parameters for carbohydrates. J Comput Chem*, 17, 450–68, 1994.

[110] D. Reith. *Neue Methoden zur Computersimulation von Polymersystemen auf verschiedenen Lngenskalen und ihre Anwendung.* Ph.D. thesis, University of Mainz, 2001.

[111] J. C. Roberts. *The Chemisty of Paper.* Royal Society of Chemistry, Cambridge, 1996.

[112] *Römpp Lexikon Chemie*, 1999.

[113] M. W. Rutland, A. Carambassis, G. A. Willing and R. D. Neumann. *Surface force measurements between cellulose surfaces using scanning probe microscopy. Colloid Surface A*, 123-124, 369–74, 1997.

[114] S. O. Samuelson and M. G. J. *Two dimensional umbrella sampling techniques for the computer simulation study of helical peptides at thermal equilibrium: The 3K(I) peptide in vacuo and solution. J Chem Phys*, 109, 11061–73, 1998.

[115] R. Schmidt, B. Teo and J. W. Brady. *Use of umbrella sampling in the calculation of the potential of mean force for maltose in vacuum from molecular dynamcis simulations. J Phys Chem*, 99, 11339–43, 1995.

[116] H. Schmitz, R. Faller and F. Müller-Plathe. *Molecular mobility in cyclic hydrocarbons – A simulation study. J Phys Chem B*, 1999.

[117] K. S. Schmitz and J.-W. Yu. *On the electrostatic contribution to the persistence length of flexible polyelectrolytes. Macromolecules*, 21, 484–93, 1988.

[118] H. Senderowitz and W. C. Still. *Anomeric free energy of d-mannose in water: A comparison of free energy perturbation, potential of mean force, and mc(jbw)/sd simulations. J Phys Chem B*, 101(8), 1409–12, 1997.

[119] J. Shaune, J. Hanley and D. G. Gray. *Atomic force microscope images of black spruce wood sections and pulp fibres. Holzforschung*, 48, 29–34, 1994.

[120] C. Simmerling, T. Fox and P. A. Kollman. *Use of a locally enhanced sampling in free energy calculations: testing and application to the $\alpha \rightarrow \beta$ anomerization of glucose. J Am Chem Soc*, 120, 5771–82, 1998.

[121] I. Simon, H. A. Scheraga and M. R. S. John. *Structure of cellulose. 1. Low-energy conformations of single chains. Macromolecules*, 21, 983–90, 1988.

[122] S. A. H. Spieser, J. A. van Kuik, L. M. J. Kroon-Batenburg and J. Kroon. *Improved carbohydrate force field for GROMOS: ring and hydroxymethyl group conformations and exo-anomeric effect. Carbohyd Res*, 322(3-4), 264–74, 1999.

[123] H. Staudinger. *Das Wissenschaftliche Werk von Hermann Staudinger*. Hüthig & Wepf Verlag, 1972.

[124] P. Stenius and J. Laine. *Studies of cellulose surfaces by titration of esca. Appl Surf Sci*, 75, 213–9, 1994.

[125] J. Sugiyama, T. Okano, H. Yamamoto and F. Horij. *Transformation of valonia cellulose crystalls by an alkaline hydrothermal treatment. Macromolecules*, 23, 3196–9, 1990.

[126] J. Sugiyama, R. Vuong and H. Chanzy. *Electron diffraction study on the two crystalline phases occurring in native cellulose from algal cell wall. Macromolecules*, 24, 4168–75, 1991.

[127] W. Tschöp, K. Kremer, J. Batoulis, T. Bürger and O. Hahn. *Simulation of polymer melts. I. Coarse-graining procedure for polycarbonates. Acta Polym*, 49, 61–74, 1998.

[128] W. Tschöp, K. Kremer, O. Hahn, J. Batoulis and T. Bürger. *Simulation of polymer melts. II. From coarse-grained models back to atomistic description. Acta Polym*, 49, 75–79, 1998.

[129] N. Tsukida, H. Muranaka, M. Ide, Y. Maeda and H. Kitano. *Effect of neutralization of poly(acrylic acid) on the structure of water examined by raman spectroscopy. J Phys Chem B*, 101(34), 6676–6679, 1997.

[130] M. Ullner, B. Jönsson, B. Söderberg and C. Peterson. *A monte carlo study of tritrating polyelectrolytes. J Phys Chem*, 104, 3048–57, 1996.

[131] M. Ullner, G. Staikos and D. Theodorou. *Monte Carlo simulations of a single polyelectrolyte in solution: Activity coefficients of the simple ions and application to viscosity measurements. Macromolecules*, 31(22), 7921–33, 1998.

[132] A. R. van Buuren and H. J. C. Berendsen. *Molecular dynamics simulations of carbohydrate-based surfactants in surfactant/water/oil systems. Langmuir*, 10, 1703–13, 1994.

[133] A. R. van Buuren, S.-J. Marrink and H. J. C. Berendsen. *Characterisation of aqueous interfaces with different hydrophobicities by molecular dynamcis. Colloid Surface A*, 102, 143–57, 1995.

[134] W. F. van Gunsteren, T. C. Beutler, P. M. King, A. E. Mark and P. Smith. *Computation of free energy in practice: Choice of approximations and accuray limiting factors.* In W. F. van Gunsteren, P. K. Weiner and A. J. Wilkinson (eds.), *Computer Simulation of Biomolecular Systems*, vol. 2, p. 315. Escom, Leiden, 1993.

[135] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. G. Tironi. *Biomolecular simulation: The GROMOS96 manual and user guide*. VDF Hochschulverlag, Zürich, 1996.

[136] D. L. VanderHart and R. H. Atalla. *Studies of microstructure in native celluloses using solid-state ^{13}C NMR. Macromolecules*, 17, 1465–72, 1984.

[137] W. J. Walczak, D. A. Hoagland and S. L. Hsu. *Analysis of polyelectrolyte chain conformation by polarized raman spectroscopy. Macromolecules*, 25, 7317–23, 1992.

[138] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. J. Profeta and P. Weiner. *A new force field for molecular mechanical simulation of nucleic acids an proteins. J Am Chem Soc*, 106(3), 765–84, 1984.

[139] B. Widom. *J Chem Phys*, 39, 2808, 1963.

[140] B. Widom. *Potential-distribution theory and the statistical mechanics of fluids. J Phys Chem*, 86, 869–72, 1982.

[141] R. G. Winkler, S. Fioravanti, G. Ciccotti, C. Margheritis and M. Villa. *Hydration of β-cyclodextrin: A molecular dynamics simulation study. J Comp Aid Mol Design*, 14, 659–67, 2000.

[142] R. J. Woods. *Carbohydrate force fields.* In *Encyclopedia of Computational Chemistry.* John Wiley and Sons, 1998.

[143] S. G. Zeller, G. W. Griesgraber and G. R. Gray. *Analysis of positions of substitution of O-carboxymethyl groups in partially O-carboxymethylated cellulose by the reductive-cleavage method. Carbohyd Res*, 211(1), 41–5, 1991.

[144] S. J. Zhou, D. L. Preston, P. S. Lomdahl and D. M. Beazley. *Large-scale molecular dynamics simulatios of dislocation intersection in copper. Science*, 279, 1525–7, 1998.