

*Molecular Dynamics Simulations of the Protein
Adsorption Process on Oxides*

Vom Fachbereich Produktionstechnik
der
UNIVERSITÄT BREMEN

zur Erlangung des Grades
Doktor-Ingenieur
genehmigte

Dissertation

von
Dipl.-Phys. Nils Alexander Hildebrand

Gutachter: Professor Dr.-Ing. Lucio Colombi Ciacchi
Zweitgutachter: Professor Dr. Dirk Zahn

Diese Arbeit ist meiner Tochter Josefine gewidmet,
die vor genau einem Monat das Licht der Welt erblickt hat.

Contents

1	Introduction	1
1.1	Proteins	4
1.2	Oxide Surfaces	5
1.3	Solvent over a Charged Surface	8
1.3.1	Electrical Double Layer Formation	9
1.3.2	DLVO Theory	11
1.3.3	Poisson-Boltzmann Equation	11
1.3.4	Generalized-Born Model	13
1.4	Protein Adsorption	14
2	Lysozyme and Chymotrypsin Binding Motifs on Silica	23
2.1	Results	25
2.1.1	Static Single-Point Calculations of the Protein Orientation	25
2.1.2	Molecular Dynamics Simulations	31
2.2	Discussion	40
3	Chymotrypsin Binding Motifs on Silica and Titania	45
3.1	Results	46
3.1.1	Long-range Interaction Force of Chymotrypsin over SiO ₂ and TiO ₂	46
3.1.2	Interaction Forces for Chymotrypsin Multilayers	49
3.1.3	Explicit-solvent MD Simulations of Chymotrypsin Adsorption on TiO ₂ and SiO ₂	51
3.2	Discussion	56

4	Surface Contact Stability	59
4.1	Methods	62
4.2	Results	63
4.3	Discussion	75
5	Surface-Induced Conformational Changes	79
5.1	Methods	83
5.1.1	Well-Tempered Metadynamics	83
5.1.2	Solute Tempering - Defining Parts of a Protein as “hot”	84
5.1.3	Replica Exchange with Solute Tempering in Combination with Well-Tempered Metadynamics (RESTmetaD)	85
5.1.4	Umbrella Simulation in Combination with Metadynamics (MetaUmbrella)	85
5.1.5	Umbrella Sampling with Multiple Walker Metadynamics and Exchanges (MulWalk-UmbrellaExchange)	87
5.2	Results	88
5.3	Discussion	105
5.3.1	Free energy MD simulations of proteins	105
5.3.2	Surface-induced changes of the helicity of chymotrypsin	109
6	Conclusion and Perspective	113
6.1	Conclusion	113
6.2	Perspective	118
6.2.1	Characterization of Alkaline Phosphatase	118
6.2.2	Mutation of Beta-Glucosidase	120
6.2.3	Amino Silane Functionalized Surfaces	123
7	APPENDIX: Molecular Dynamics Simulations	127
7.1	Standard Molecular Dynamics Simulations	127
7.1.1	Interactions Described by Force Field Equations	127
7.1.2	From the Force Field Equation to the Dynamics	129
7.1.3	The Simulated Solvent and Periodic Boundary Conditions	130
7.1.4	Thermostats	131
7.2	Explicit Solvent	134

7.3	Implicit Solvent Simulations	135
7.4	Collective Variables	136
7.4.1	HELIX Loops	136
7.5	Free Energy Sampling Methods	137
7.5.1	Arrhenius Equation	137
7.5.2	Umbrella Sampling	138
7.5.3	Steered Molecular Dynamics Simulations	139
7.5.4	Metadynamics	140
7.5.5	Replica Exchange Methods	142
7.6	Radial Distribution Function (RDF)	150
7.7	Dipole Moment	151
7.8	Cluster Analysis	151

Zusammenfassung

Die Protein- und Peptidadsorption zur Biofunktionalisierung von Oberflächen spielt eine wichtige Rolle in vielen biomedizinischen Anwendungen. Um den Adsorptionsprozess auf oxidbasierten Materialien auf mikroskopischer Ebene besser zu verstehen, werden in der vorliegenden Arbeit die einzelnen Phasen der Adsorption mithilfe Molekulardynamik - Simulationen (MD) untersucht. Insbesondere wird die Adsorption von Lysozym und Chymotrypsin auf Silica und Titania betrachtet.

Hierfür werden basierend auf der DLVO-Theorie zunächst die Fernfeld-Wechselwirkungen zwischen Protein und Oberfläche berechnet. Dies ermöglicht die Bestimmung einer bevorzugten Adsorptionsorientierung eines Proteins, die in weiteren dynamischen Simulationen mit explizitem Lösungsmittel bestätigt werden kann. Besonders das Dipolmoment des Proteins beeinflusst die resultierende Orientierung auf der Oberfläche maßgeblich. Dadurch bewirkt die stets saure Siliziumdioxidoberfläche eine einheitliche bevorzugte Orientierung der Proteine. Titandioxid hingegen hat einen amphoteren Charakter, was eine entgegengesetzte bevorzugte Proteinorientierung ober- und unterhalb des isoelektrischen Punktes erzeugt. Im Vergleich mit Adsorptionsexperimenten zeigt sich zudem, dass neben den DLVO-Kräften der einzelnen Moleküle auch Protein - Protein Wechselwirkungen und die genaue Wasserstrukturierung über der Oberfläche berücksichtigt werden müssen.

In einem weiteren Schritt wird daher die Kontaktstabilität der Proteine auf der Oberfläche durch simulierte und experimentell durchgeführte Rasterkraftmikroskopie näher betrachtet. Die Untersuchungen zeigen eine starke Stabilisierung des Proteins sowohl durch die umgebende Wasserstruktur als auch durch die proteininternen Disulfidbrücken. Dies bildet gemeinsam ein stabilisierendes Netzwerk und führt dazu, dass Lysozym ohne signifikante Konformationsänderungen desorbiert.

Um zudem eine mögliche spontane oberflächeninduzierte Konformationsänderung von Proteinen auf Oxidoberflächen zu untersuchen, werden mehrere Simulationsmethoden zur Analyse der freien Energie weiterentwickelt. Diese Methoden werden beispielhaft an auf Siliziumdioxid adsorbiertem Chymotrypsin angewendet. Das resultierende freie Energie-

profil des adsorbierten Proteins zeigt eine geringfügige Reduzierung der Helizität des Proteins im Sinne einer leichten Denaturierung, was durch Zirkulardichroismus - Experimente bestätigt werden kann. Diese leichte Denaturierung wird durch einen direkten Kontakt des nativen Helixanteils des Proteins mit der hydrophilen Oberfläche verursacht. Es wird somit nahegelegt, dass Chymotrypsin während des Adsorptionsvorganges nur kleine Konformationsänderungen erfährt und überwiegend seine native Struktur beibehält.

Die entwickelten und vorgestellten Simulationen können leicht auf weitere Umgebungseinflüsse wie spezielle Oberflächen oder Lösungsmittel angewendet werden, welche im Experiment teilweise nur schwer zu untersuchen sind. Darüber hinaus kann zum Einen das Adsorptionsverhalten bestimmter Oberflächenfunktionalisierungen und des Weiteren die bevorzugte strukturelle Konformation eines Bereiches eines Makromoleküls berechnet werden. Somit können die Simulationen experimentelle Arbeiten sinnvoll ergänzen und deren Effizienz durch aussagekräftige Vorhersagen deutlich steigern.

Introduction

Protein and peptide adsorption on oxide surfaces is particularly relevant when an artificial material interacts with the biological environment. Since many related technical applications stem from the medical sector, this environment is often a living organism. Therefore, the adsorption behavior of endogenous proteins on exogenous materials such as ceramics or metals brought into contact with the organism needs to be considered. Next to the surface functionalization of implant materials, innovative applications including protein coatings for nanomaterials are used in a variety of organic-inorganic hybrid materials for medical and biotechnological applications. For example, nanoparticles can be used as carrier systems in nanomedicine [1, 2] as well as in imaging and sensing approaches [3, 4].

When the artificial material is in contact with an organism, it becomes covered with various proteins within seconds to hours [5–8]. These proteins include mainly plasma proteins such as albumin, fibrinogen, IgG and fibronectin [8–10]. Dependent on the type of application the adsorption of these proteins can enhance or prevent the biocompatibility of the materials. Fibrinogen and thrombin are part of the blood coagulation whereas the adsorption of the adhesion protein fibronectin mediates the cell adhesion, which is advantageous for implants [5, 11]. Völker et al. [12] generated carboxyl groups on a silicon corneal prosthesis to bind fibronectin. The thus achieved fibronectin enrichment increases the surface biocompatibility and stimulates the integration into the endogenous tissue. However, in most of the cases, surfaces are engineered to reduce the protein-surface interactions, because the amount and the subsequent conformational changes of the adsorbed proteins might induce an immune reaction or blood coagulation [8, 13]. Furthermore, in many biomedical devices such as the bearing area of artificial joints,

tubes or nanodevices, no tissue integration is wanted and thus the adsorption of neither fibronectin nor other plasma proteins is desired. Therefore, the implant materials need to be engineered with care in order to prevent these body reactions otherwise caused by the artificial surfaces.

It is known that proteins bind to hydrophobic surfaces in a higher amount and more tightly compared to hydrophilic surfaces [6, 14, 15]. Furthermore, it is assumed that hydrophobic surfaces induce conformational changes which could result in a reduced biocompatibility [6, 16]. The water above the hydrophobic surface forms structured layers which are rich in hydrogen bonds whereas no bonds are formed with the surface itself. The adsorbing proteins break the water structures with increasing entropy. At the surface, the proteins undergo surface-induced conformational changes – possibly mediated by hydrophobic interactions – and expose hydrophobic domains which tightly bind to the surface [17, 18]. The hydrophilic polar domains of the protein are turned towards the solvent offering possible hydrogen bonding sites, thus acting as a surfactant [5]. Therefore, to minimize the protein-surface interactions, hydrophilic surfaces are more promising candidates for implant materials.

In addition to the surface hydrophilicity, protein adsorption is also influenced by the surface charge, which has been the focus of several recent investigations. Especially the effect of the four most common surface functionalizations with methyl (-CH₃), amino (-NH₂), carboxyl (-COOH) and hydroxyl (-OH) groups on proteins has been studied [6]. A neutrally charged, hydrophobic methyl (-CH₃) functionalization, for example, has shown a strong fibrinogen binding and thrombocyte accumulation as part of the blood coagulation [19]. A positively charged, hydrophilic amine (-NH₂) functionalization, on the other hand, has promoted the adsorption of fibronectin and subsequent focal adhesion components [16] but also triggered an acute inflammatory response [20]. A negatively charged, hydrophilic carboxyl (-COOH) functionalization has shown an increased adsorption of fibronectin and albumin compared to (-OH) [21]. The surface functionalized with hydrophilic, neutrally charged hydroxyl (-OH) groups, promoted the exposure of cell adhesive domains of fibronectin [16] but also has shown a reduced plasma protein adsorption [19, 22]. In summary, neutrally and negatively charged hydrophilic surfaces demonstrated favorable biocompatible properties as also suggested by Ostuni et al. [23].

These empirically measured trends reveal first insights of suitable materials for biomedical applications. However, in order to engineer suitable surface functionalizations, a pro-

found understanding of the protein adsorption process on the microscopic level is needed. The presented work contributes to this knowledge gap by the introduction of simulation tools to study the individual phases of the adsorption with an atomic resolution.

Under physiological conditions, the huge variety of micro- and macromolecules creates a rather complex adsorption environment. The atomistic simulation of this interplay of several solutes, the solvent and the adsorbent is too much of a computational challenge in many cases of interest. However, the simulation of an isolated part of the adsorption process is easier to realize and provides broad opportunities for further investigations, as shown in this work. The presented investigations can be seen as an artificial atomic microscope with the ability to change different material properties of the system, even to nonphysical conditions. These changes can include for example the scaling of specific interactions within the system, the usage of extremely steep temperature ramps or the replacement of the explicit water molecules with a continuum solvent model as done in this work. The detailed description of the isolated part of the adsorption process contributes to the general image of the complete adsorption process.

This thesis investigates exemplarily the adsorption of two proteins on two oxide surfaces in different adsorption phases. The used proteins chymotrypsin and lysozyme are particularly interesting for biomedical applications due to the generic proteolytic properties of chymotrypsin [24] and the antibacterial properties of lysozyme [25–27]. The proteolytic properties of chymotrypsin are particularly suitable for nanotechnology applications. Nanoparticles coated with chymotrypsin can easily be centrifuged to disable e.g. the proteolysis in a suspension. The antibacterial properties of lysozyme could be used as a coating for implants since post-surgical inflammation and bacterial films on artificial implants are still a major complication in implantology [28, 29]. Both used oxides, silica and titania, are hydrophilic surfaces with different surface charge depending on the pH and are used in numerous nanotechnology applications such as bone-repairing devices, drug delivery systems [30] and orthopedic implants [31]. In the following, the used proteins and surfaces, as well as the liquid-solid interfaces, are introduced in detail followed by a description of the different phases of adsorption.

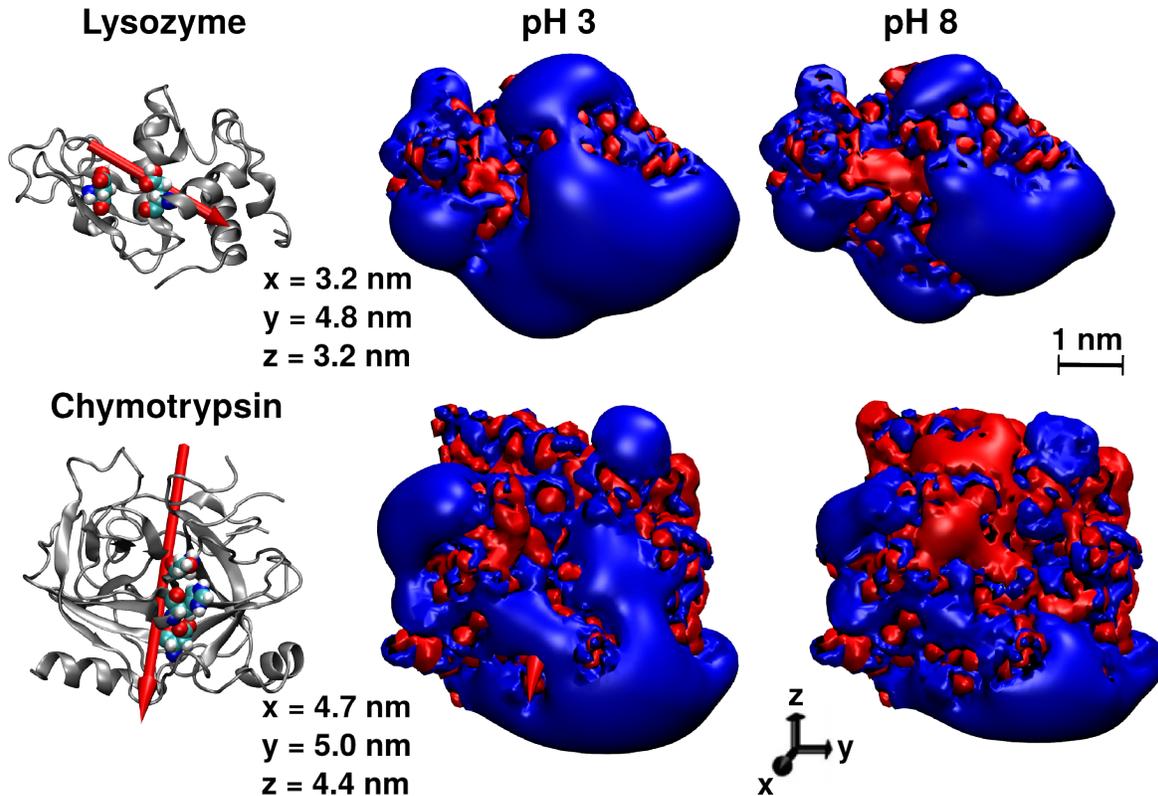


Figure 1.1: *Hen-egg-white lysozyme and α -chymotrypsin* is taken from the protein data bank. The dipole moment is shown as red arrow and the active sides as VDW-style colored atoms. The isoelectric surfaces at $\pm 1.4 k_B T/e$ is shown in blue and red for the pH of 3 and 8. The protein dimensions are indicated in all axes.

1.1 Proteins

Hen-egg-white lysozyme is a glycoside hydrolases protein. It is a rather small protein with 129 amino acids, four disulfide bonds and exposes an ellipsoidal shape of about $3.2 \times 4.8 \times 3.2 \text{ nm}^3$ with a positive net charge at physiological pH (Fig. 1.1 top). Its isoelectric point is located at a pH of approximately 11 [32]. In the human body, lysozyme is present in e.g. tears, saliva and human milk as well as in large amounts in animal egg white. It is the first enzyme structure ever resolved via X-ray diffraction methods in 1963 [33]. In combination with many investigations on further protein properties, the enzyme is an exemplary protein for many conceptual research topics in various fields. The enzymatic activity of lysozyme is the hydrolysis of the 1,4-beta linkage between the N-acetylmuramic acid (NAM) and the N-acetyl-D-glucosamine residues (NAG). This catalytic process is mediated by the protein residues GLU 35 and ASP 52 inside the active site of the protein

(Fig. 1.1 top). The bacterial cell wall of gram-positive bacteria is rich of an NAM-NAG peptidoglycan network which can be attacked by lysozyme – making the protein antibacterial.

The MD simulation of lysozyme is based on the protein database structure 1GXV.pdb. Missing residues in the crystal structure are added and the pH is adjusted by the protonation of sensitive residues. By using an AMBER force field and solving the Poisson-Boltzmann equation numerically, the electrostatic field around the protein can be shown by its isosurface as presented in Fig. 1.1. Further details of this representation can be found in [34] and in Chapter 2 and 3.

The proteolysis protein α -chymotrypsin is a slightly larger protein with 247 amino acids, five disulfide bonds and exposing a spherical shape of about $5.0 \times 4.7 \times 4.4 \text{ nm}^3$. It has a slightly positive net charge at physiological pH (Fig. 1.1 bottom). The simulation model of this protein is based on the database structure 4CHA.pdb and is similarly prepared as in the case of lysozyme. The proteins isoelectric point is located at a pH of approximately 8.8 [35]. Chymotrypsin is a digestive enzyme of the pancreatic juice of many organisms which perform proteolysis of proteins or peptides. The amino acids involved in the catalytic activity are part of its catalytic triad. This catalytic triad of chymotrypsin consists of the uncharged SER195, the protonated neutrally charged HIS57 with a proton at the delta position and the anionic ASP102. The protein activity is maximized at pH 8 which is similar to the pancreatic milieu [36]. The used PDB structure contains a dimer configuration of the protein, based on a crystallization at pH 4.2 [37]. Further studies [36, 38] confirm a maximum in the dimerization ability at about pH 4. However, mass spectrometry measurements revealed clearly a monomer configuration at pH 8 [39, 40].

1.2 Oxide Surfaces

The two oxide surfaces used in this work, amorphous silica (SiO_2) and titania (TiO_2) are introduced in the following.

SiO_2 is mostly found in the form of the crystalline quartz with a network of tetrahedral bonds. The two most frequent crystalline structures are the α and β phases. In both cases, the bonding tetrahedra form a hexagonal network. However, in many technical and especially nanotechnological applications of SiO_2 , the interface properties are mostly

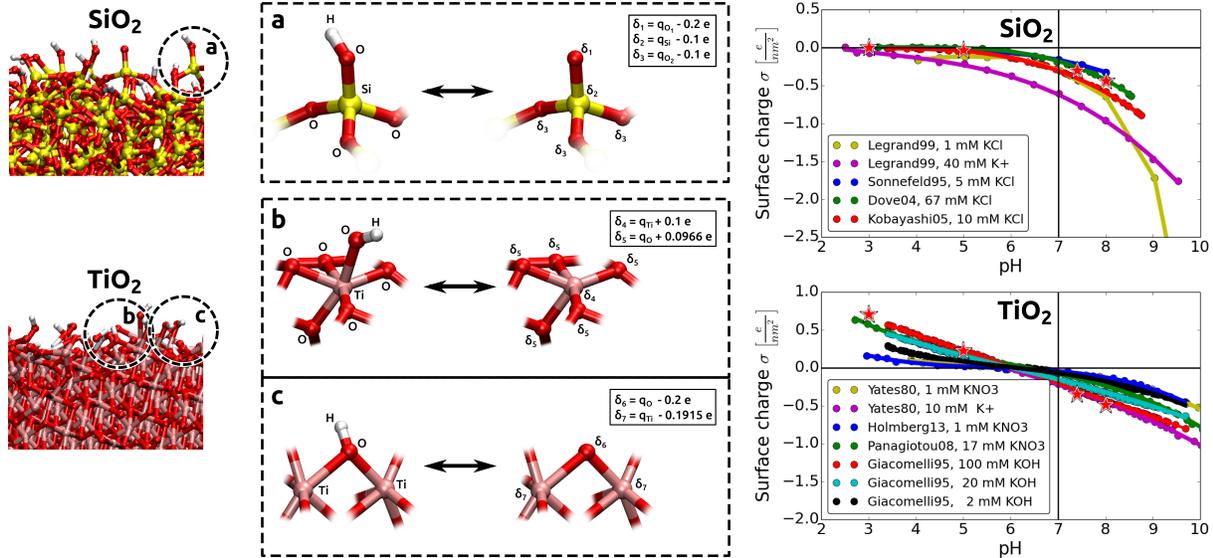


Figure 1.2: Surface termination and adjustment of the pH value of the amorphous SiO_2 and TiO_2 surfaces. In the case of the acidic SiO_2 surface, the terminal silanol groups are protonated to adjust the surface charge to a value evaluated from several potentiometric titration experiments by the named authors [42–45]. In the case of the amphoteric TiO_2 surface, the surface terminal groups are either hydroxylated or protonated on bases of the MUSIC model described for TiO_2 by Köppen and Langel [46]. The resulting surface charge is in relatively good agreement with the named authors [47–50].

determined by the amorphous phase. In the following, some of these application systems are presented. The exposition of a silicon wafer to air forms a thin amorphous oxide layer with a thickness of a few nm at room temperature [41]. Prominent techniques to produce amorphous SiO_2 nanoparticles are based on gas-phase (fumed SiO_2) or solvent (Stöber SiO_2) methods. Fumed SiO_2 particles are produced by a flame spray method whereas the production of Stöber SiO_2 is a wet chemistry approach. Another technique to produce amorphous SiO_2 is the fast quenching of the oxide from the liquid phase. This quenching technique leads to an amorphous phase which is very similar to a “frozen liquid”. The chemical bonding is identical to the crystal phase. Small variations of the angles and lengths of the bonds create a short-range order of the amorphous material but not a long-range order as in a crystal.

The amorphous SiO_2 surface is modeled in this work with the slab structure obtained in a previous work by Cole et al., [51] (Fig. 1.2) using the SiO_2 -water interaction potential and parameters of Butenuth et al. [52]. Silica forms a strong acidic surface and has an isoelectric point at 1.7 - 3.5 [53]. The surface reactive groups are silanol groups which in

the case of deprotonation cause the negative charge of the surface. The surface charge is adjusted by the removal of one proton and the additional reduction of the charge of the neighboring Si and O as shown in Fig. 1.2 a). Hence, the total charge of the slab is reduced by $-1 e$, representing the deprotonation of one silanol group. By this approach, the surface charge density is adjusted to the pH value so as to reproduce average values of the results of several potentiometric titration experiments [42–45]. This results in final surface charge densities of 0.0, -0.03 , -0.31 and $-0.44 e \text{ nm}^{-2}$ at pH 3.0, 5.0, 7.4 and 8.0, respectively (Fig. 1.2 right column, marked with stars).

The charged silanol groups and the charge distribution of the surface atoms generate hydrophilic surface properties [52, 54, 55].

TiO₂ exists in nature in three different crystalline forms: Rutile, Anatase and Brookite. Tons of nanoparticles are produced by the pigment industry worldwide. TiO₂ is mainly produced by either the sulfate or the chloride process [56, 57]. Another technique to produce TiO₂ nanoparticles is the flame spray reactor method introduced by Mädler et al. [58, 59].

It is known from experiments that a thin oxide layer on Ti exhibits a mainly amorphous character [60–62]. In the same way, it is expected, that the first layers of TiO₂ have an amorphous or at least disordered character, especially for small nanoparticles.

In contrast to SiO₂, TiO₂ forms an amphoteric surface with an IEP at about 5 - 7 [53]. The reactive groups are the hydroxylation of a five-fold (or smaller coordinated) titanium atom and the protonation of a bridging oxygen. These groups form the basic and acidic properties of the surface, respectively. The reaction is described by the MUSIC model of Hiemstra et al. [63]. The construction of the surface and its distribution of the reactive groups have been developed by Simon Kunze [64, 65]. After removing the hydroxyl or the proton from the slab, the charges of the neighboring and next neighboring atoms are reduced as shown in Fig. 1.2 b) - c). The charge reduction follows the approach of Butenuth et al. [52] introduced for SiO₂, namely to reduce the charge of these neighboring atoms by 0.1 - 0.2 e . The ratio of both groups as a function of pH is taken from Köppen et al. [46] resulting in a surface charge density of 0.71, 0.23, -0.34 , and $-0.49 e \text{ nm}^{-2}$ at pH 3.0, 5.0, 7.4, and 8.0. It can be seen in Fig. 1.2 (right/bottom) that these values marked with red stars are in good agreement with several electrical titration experiments [47–50]. The force field of the interface interactions has been developed by Schneider et al. [55].

The simulated interface contains the described proteins and surfaces plus an intermediate solvent. This solvent crucially influences the surface properties, as presented in the following.

1.3 Solvent over a Charged Surface

The specific distribution and protonation state of the surface reactive groups do strongly influence the solvent and need to be chosen with care (Fig. 1.2).

In the case of an ionic solvent, the surface charge influences the position of the mobile solvent ions to form an electrical double layer. An aqueous solvent creates a structured configuration in the very proximity to the surface. This structure is often present in layers and is caused by the polarity of the water together with entropic effects.

The adsorbate mobility prior to adsorption depends on its diffusion in the solvent. Similar to the solvent structure formation over the surface, the solvent also forms structures around the adsorbate as well, referred to as a solvation shell. The positions of the ions and the polar water molecules around the adsorbate are determined by its charge distribution. If the adsorbate approaches the oxide, its behavior is influenced by direct adsorbate-surface interactions as well as the overlap of the solvent structures of the surface and the adsorbate. The free energy gain of a constructive interference of these water layers supports the adsorption whereas a destructive interference obstructs it.

Therefore, the dynamics involved in the adsorption process include a constant rearrangement of the solvent molecules. These dynamics are very well reproduced by molecular dynamics simulations, which motivates the approach of this thesis (Appendix 7).

An analytical approximation to describe the liquid-solid interface theoretically is based on the Poisson-Boltzmann (PB) equation and Derjaguin-Landau-Verwey-Overbeek (DLVO) theory [66,67]. These theories do not include the just described solvation shell. They focus on the ion distribution within the solvent. These approaches are historically based on particle interactions. While the application of these theories on protein adsorption will be investigated later in this thesis, their derivation is described in the following section. The equations of these theories describe the far-field interactions between protein and surface as a function of surface charges and ionic strength.

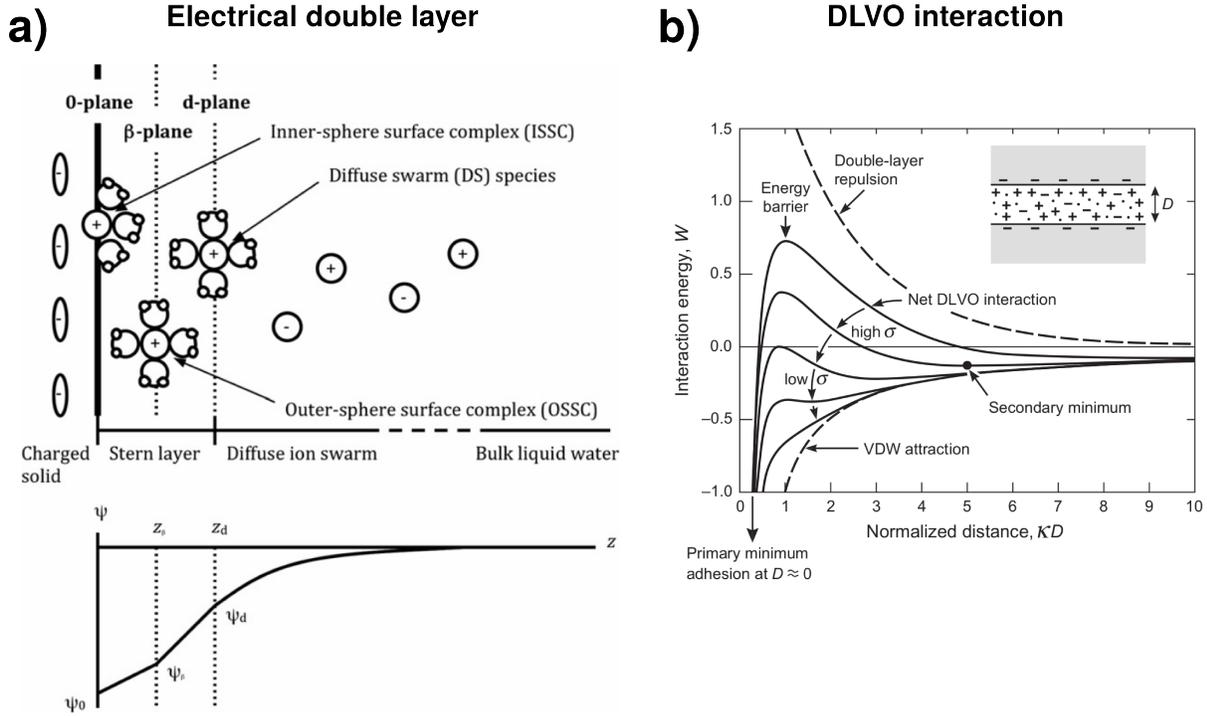


Figure 1.3: a) Accommodation of ions over a negatively charged surface using the triple-layer model of Bourg et al. [68, 69]. (top) Ions form the 0-, β - and d-plane in the proximity of the surface. These planes form the Stern layer as the opposite pole to the surface. Further ions form the diffuse ion layer. The present ion complexes are called ISSC, OSSC and DS as declared in the image. (bottom) The electrostatic potential decreases linearly first between the 0- and β -planes and then exponentially between the β - and d-plane following the Poisson-Boltzmann equation. b) DLVO theory of equally charged surfaces or particles. The electrostatic double-layer repulsion is in competition with the VdW attraction. At medium salt concentrations and surface charge σ , two attraction minima are formed, separated by an energy barrier. At high σ and low salt concentrations the double-layer repulsion dominates and the secondary minimum disappears. Similarly, for low σ and high salt concentrations with dominating VdW interactions. κ is the inverse Debye length and D the particle distance. The figures are reprinted from Bourg et al. [68] and Israelachvili [70], ©2011, with permission from Elsevier.

1.3.1 Electrical Double Layer Formation

If a charged surface is in contact with an aqueous electrolyte solution, an electrical double layer is formed by primary counterions in the proximity of the surface. This double layer is influenced by entropic properties of the solvent as well as Coulomb and VdW interactions. Among various forms of this layer model the triple-layer model proposed by Bourg et al. [68, 69] is described in the following (Fig. 1.3 a)).

Due to the polarity of the water, each ion contains a rather rigid (depending on the valency and size) hydration shell forming an ion-water complex (ISSC, OSSC and DC

as declared in Fig. 1.3 a)). The Pauli exclusion principle generates a gap between the surface and the 0-plane of ions and partly truncates the hydration shell of ISSC complexes to the cost of solvation energy. The 0-plane formed by ions and water atoms is also known as the inner Helmholtz plane. Depending on the surface hydrophilicity, this plane often contains the most prominent water layer on top of the surface with increased water density compared to the bulk water. The layer between 0-plane and d-plane is called the Stern layer and contains most of the physisorbed ions.

The electrostatic potential of the interface has been first described by the Gouy-Chapman model [71, 72]. It states that the ions above the charged surface exhibit a Boltzmann distribution following the potential calculated by the Poisson equation. This Poisson-Boltzmann equation (described in detail in Section 1.3.3) implies that the electrostatic potential decreases exponentially with increasing distance from the surface. It was found by Grahame [73] that this equation is only valid for the diffuse ion layer. Instead, he defined the inner and outer Helmholtz planes (see 0- and β - plane in Fig. 1.3 a)). These planes of high ion concentrations are explained by the overlap of the hydration shells of the ions and the surface and the resulting entropic effects. This theory assumes a linear decrease of the electrical potential in the Stern layer based on the description of a plate capacitor as shown in Fig. 1.3 a) (bottom).

Molecular dynamics simulations of an NaCl - CaCl₂ solution on a smectite surface by Bourg et al. [68] showed that the most prominent concentration of ions could be found in the β -plane. In this plane the positively charged Na⁺ and Ca²⁺ are embedded with their complete hydration shell directly between the first and second layer of dense water above the negatively charged smectite surface. The hydration shells of the ions and the surface overlap. Ions in the 0-plane which have to contain a truncated hydration shell are much less likely due to their loss of solvation energy of the hydration shell. In the study by Bourg et al. [68] all three planes 0, β and d are equal to the position of the first three prominent peaks of the ion concentrations above the charged surface. The Stern layer in this study is located at about 1.1 nm above the surface. In comparison, the Debye length of this system is about $\lambda_D = 0.3$ nm (0.14 M = 0.8 nm, 3 mM = 5.6 nm). λ_D is defined as the distance where the potential is $\Psi(x) = \Psi_0/e$, with Ψ_0 being the potential at the surface. It is an indicator for the electrical screening of the solvent and depends highly on the salt type and concentration. The Debye length is given for monovalent ions at room temperature by $\lambda_D = \kappa^{-1} = \frac{0.304}{\sqrt{c_s}}$ [70], in nm for a molar salt concentration c_s .

1.3.2 DLVO Theory

The theory by Derjaguin and Landau [66] and Verwey and Overbeek [67] (DLVO) combines the electrostatic repulsion (in the case of two equally charged surfaces) and the VdW attraction. It is developed and mainly applied for particle suspensions in an aqueous electrolyte solution. If the particles are all equally charged, the electrostatic interactions are repulsive, which is termed double-layer repulsion (Fig. 1.3). The interactions in the Stern layer as described in Section 1.3.1 are neglected and only the interactions in the diffuse layer with the exponential decrease of the electrical potential are included. The attractive VdW interactions are not salt-dependent, but the electrostatic interactions are highly influenced by the type and concentrations of the ions (see Section 1.3.3 and 1.3.4). The VdW attraction is usually described by a power-law interaction ($\propto -1/r^n$) [70]. Due to the difference of the exponential decrease of the electrostatics and the $\propto -1/r^n$ decrease of the VdW attraction two energy minima can appear as shown in Fig. 1.3 b). The first minimum caused by the VdW attraction is separated by an energy barrier from the second minimum, which is located typically at distances of about 3 nm [70]. In comparison to the typical VdW radius of 0.4 nm [74], this minimum is far away from the surface. The second minimum disappears for very low salt concentrations or high surface charges due to dominating double-layer repulsions. Similarly, it disappears for very high salt concentrations or low surface charge based on a high electrostatic screening and a dominating VdW attraction, as shown in Fig. 1.3 b). Therefore, an aqueous particle suspension exhibiting the presented interactions is stable at low salt concentrations and high surface charges σ , whereas the suspension often precipitates for contrary conditions.

In systems of protein adsorption, the proteins and surfaces are not necessarily of the same charge. Furthermore, the charge distribution, especially of the proteins, can be very inhomogeneous. This complex electrostatic distribution is shown by the isosurface and fieldlines of chymotrypsin over a SiO_2 surface in Fig. 1.4 a). The computation of this system is performed in this work on the bases of the Poisson-Boltzmann equation and its approximation within the Generalized-Born approach described in the following.

1.3.3 Poisson-Boltzmann Equation

The exponential decrease of the electrical potential in the diffuse layer as shown in Fig. 1.3 a) (*bottom*) is part of the Gouy-Chapman model [71, 72]. The model includes the Boltzmann distribution of the ions based on the potential calculated by the Poisson equation.

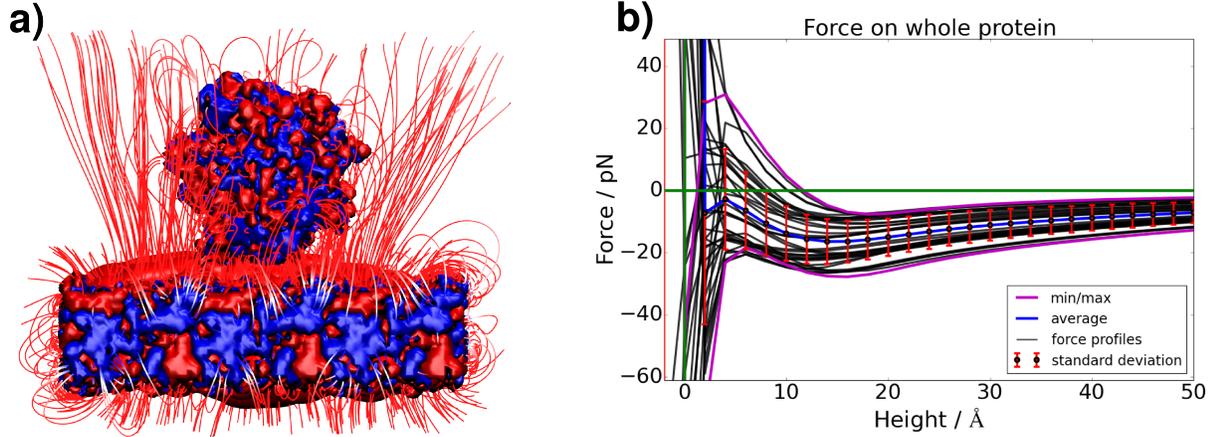


Figure 1.4: a) *Electrostatic isosurface of chymotrypsin over a negatively charged SiO₂ surface including fieldlines. The numerical calculation of the Poisson-Boltzmann equation is performed at a monovalent salt concentration of 0.1 M. The isosurfaces are shown at $\pm 5 k_B T / e$ (blue/red).* b) *Interaction profiles calculated by the Generalized-Born model of chymotrypsin over SiO₂ using a 0.1 M monovalent salt concentration. The profile is calculated for 36 different protein orientations shown as gray curves. Negative force values are defined as attractive interactions between surface and protein.*

In the following, the Poisson-Boltzmann equation is deduced for a system of two equally charged surfaces at a distance D with a solvent and counterions in between [70]. The Boltzmann distribution of the ions is given by:

$$\rho = \rho_0 e^{-\frac{ze\Psi}{k_B T}}$$

with the density of the ions ρ , the density in the middle of the two surfaces ρ_0 , the valency z , the charge e , the electrostatic potential Ψ , the Boltzmann constant k_B and temperature T . The term $ze\Psi$ is equal to the electrical potential energy of the ions. The electrostatic potential Ψ can be calculated by the integration of the Poisson equation, which is given by [70]:

$$ze\rho = -\epsilon\epsilon_0 \frac{d^2\psi}{dx^2}$$

with the permittivity of vacuum ϵ_0 and the dielectric constant of the solvent ϵ . Both equations yield the Poisson-Boltzmann equation:

$$\frac{d^2\Psi}{dx^2} = -\frac{ze\rho_0}{\epsilon\epsilon_0} \cdot e^{-\frac{ze\Psi}{k_B T}} \quad (1.1)$$

When this nonlinear second order differential equation is solved, it gives the potential Ψ , the electrical field $E = \partial\Psi/\partial x$ and the counterion density ρ . For long distances away from the surface and if $\frac{ze\Psi}{k_B T} \ll 1$ the Debye-Hückel approximation [75] with: $e^{-\frac{ze\Psi}{k_B T}} \approx 1 - \frac{ze\Psi}{k_B T}$ can be used to transform the nonlinear in a linear Poisson-Boltzmann equation.

The numerical calculation of Equation 1.1 is implemented in the AMBER simulation package [76]. This package is used to calculate the electrical potential of chymotrypsin over SiO_2 as shown in Fig. 1.4 a).

1.3.4 Generalized-Born Model

The Generalized-Born model is an analytical approximation to the computational expensive solution of the Poisson-Boltzmann differential equation (Eq. 1.1). The Generalized-Born approach is also included in the AMBER simulation package [76]. In this implementation, each atom is represented as a sphere of radius R_i (effective Born radius) and a charge q_i in its center. For an isolated ion the Born radius is equal to its VdW radius. For an embedded atom the Born radius depends on the molecule conformation and needs to be calculated for each molecular dynamics step. Various methods exist to calculate the Born radii. All methods are based on a different approach to compute the solute surface. Two commonly used methods are based on the approach of Hawkins, Cramer, Truhlar [77, 78] or Onufriev, Bashford, Case [79, 80]. The interior of the atoms below the Born radii is filled with dielectric constant $\epsilon = 1$ and the solvent with $\epsilon = 80$ at 300 K [81]. The partial charge q_i of the atom i is located in the center of the sphere with this Born radii R_i . The Generalized Born approximation of the electrical part of the solvation free energy (see Appendix Section 7.3) is given by [81–83]:

$$\Delta G_{el} = -\frac{1}{2} \sum_{ij} \frac{q_i q_j}{f_{GB}(r_{ij}, R_i, R_j)} \left(1 - \frac{e^{-\kappa f_{GB}}}{\epsilon} \right) \quad (1.2)$$

A common choice of the function f_{GB} is [82]:

$$f_{GB} = \sqrt{r_{ij}^2 + R_i R_j} e^{-\frac{r_{ij}^2}{4R_i R_j}}$$

with the distance r_{ij} between atoms i and j , the effective Born radii $R_{i,j}$ and the Debye-Hückel electrostatic screening parameter κ [81]. The Debye length is given as κ^{-1} . There-

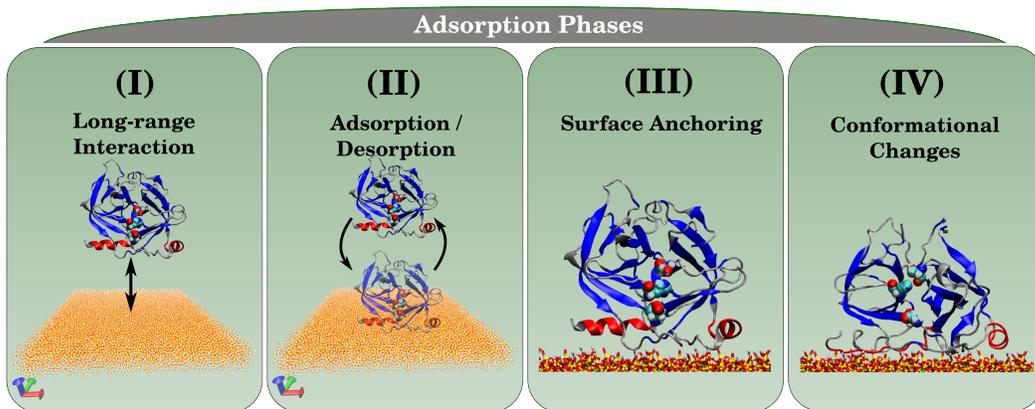


Figure 1.5: Separation of the adsorption process in four phases as investigated in this thesis. Far-field interactions (I) in Chapter 2 and 3, adsorption and desorption processes in Chapter 3, surface anchoring in Chapter 2, 3 and 4 and surface-induced conformational changes in Chapter 5.

fore the screening effect of the salt is included by the temperature depended Debye length and the screening of the solvent by the dielectric constant ϵ . In case of the PB approach, the ions are directly included by considering their effect on the electrical potential. Due to the exponential equation in f_{GB} it can be seen that for small distances of the interacting atoms r_{ij} (in the order of magnitude of $R_{i,j}$ or smaller), the effect of the used Born radii can be significant. Therefore, whenever the nearfield interactions at these distances are of interest, the Born radii $R_{i,j}$ have to be chosen with care.

The Generalized-Born approach is used to calculate the force profiles of chymotrypsin over SiO_2 as shown in Fig. 1.4 b). This method is extensively used in Chapter 2 and 3.

1.4 Protein Adsorption

The protein adsorption process can be divided into four phases, as shown in Fig. 1.5. These phases are a guideline for the chapter sequence in this work. The long-range interactions (I) attract or repel the molecules to the surface. In competition with other solutes and the solvent, the molecules approach the near-field of the surface. The protein adsorbs and can reversibly desorb again forming an adsorption/desorption equilibrium (II). By strong physical or chemical interaction the protein stably anchors to the surface (III), followed by possible surface-induced conformational changes (IV).

(I) Long-range Interactions

The long-range interactions between protein and surface are influenced by Coulomb, van der Waals and solvent interactions. The screening ability of the solvent depends strongly on the valency and concentration of the solvent salt and influences the electrostatic field significantly. The DLVO forces are calculated by the Poisson-Boltzmann equation in combination with the van der Waals attraction. Here, it is used to investigate the long-range interactions of proteins over the charged oxide surfaces. Attraction/repulsion and even preferred protein orientations can be derived from the calculations (Chapter 2 and 3).

Recent experiments [84, 85] have shown that protein adsorption on oxides can be mediated by the electrostatic distribution of protein and surface. Arguments based on the DLVO theory are used to explain the amount of adsorption under different pH and ionic strength conditions. Indeed, the experimentally measured fraction of adsorbed proteins in the presence of amorphous SiO₂ nanoparticles increase steadily with increasing pH, as shown in Fig. 1.6 [34]. A high pH value is associated with an increasing surface charge, which results in strong DLVO forces between surface and proteins. The data presented in Fig. 1.6 display the amounts of adsorbed enzymes measured by the Bicinchoninic acid assay (BCA depletion assay) [40]. In particular, the experiment reveals that only chymotrypsin reaches an adsorption density close to the theoretical jamming coverage (at pH 8), while lysozyme adsorbs in a much lower amount. The theoretical jamming coverage is calculated by two monolayer models [84], corresponding to the proteins occupying a rectangular or a dense hexagonal lattice (Fig. 1.6 b), inset).

Increasing the pH from 3.0 to 8.0 leads to an increase of the negative surface charge density from 0.0 to $-0.44 e \text{ nm}^{-2}$ (Fig. 1.2). The positive net protein charge, instead, decreases from +7 to +3 e for chymotrypsin, and from +10 to +9 e for lysozyme. The fact that chymotrypsin adsorbs in a larger amount than lysozyme (up to 73% vs. 13.7% at pH 8) despite its lower net charge suggests that trivial electrostatic arguments cannot be used to explain the observed adsorption behavior. These apparently contradictory observations motivate further analysis of the adsorption isotherms, described in the following, and the atomistic simulations presented in Chapter 2 and 3.

(II) Adsorption - Desorption

In the case of a long-range attraction, proteins approach the surface, possibly bind and desorb again. The amount of potential surface binding sites, the concentration in solution

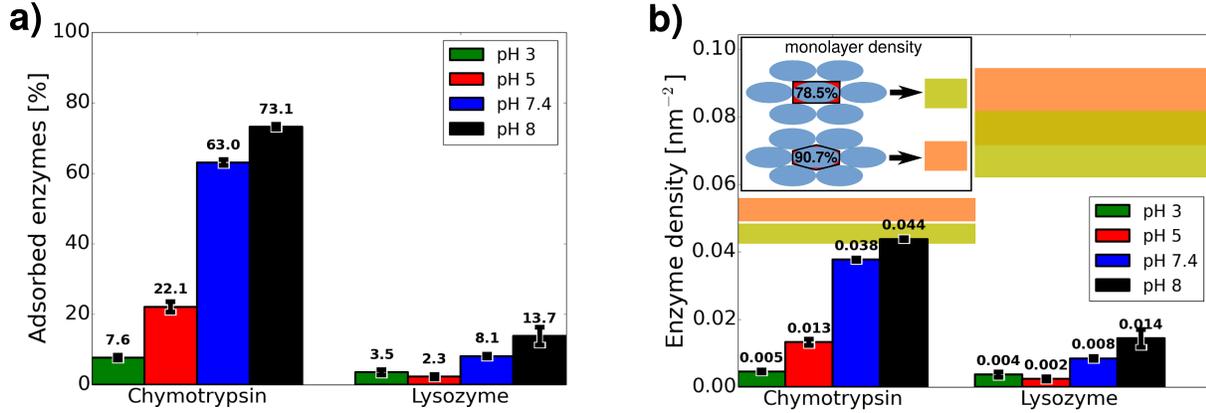


Figure 1.6: a) Percentage of adsorbed chymotrypsin and lysozyme on amorphous SiO_2 nanoparticles at different pH values, obtained by means of BCA depletion assay after 20 hours of incubation. b) Corresponding surface protein densities with respect to the specific surface area (BET) of the particles. The inset shows the assumed jamming coverages of a protein monolayer for a rectangular (light green) and hexagonal (orange) packing, using an ellipsoidal protein model. The bar width results from the difference between a side-on and an end-on orientation [86] of the adsorbed proteins [34].

and the surface affinity of the proteins influence the equilibrium reaction between adsorbed and desorbed proteins. Adsorption isotherms describe the equilibrium as a function of the protein concentration in solution and allow thereby the calculation of the adsorption free energy.

A simple adsorption isotherm is predicted by the Langmuir model. It assumes an adsorption up to a monolayer, a homogeneous distribution of equal adsorption sites at the surface and no interactions between the adsorbate molecules in solution or on the surface [87]. With a fixed number of surface sites, the Langmuir isotherm is derived in the following [88] using the number of the adsorbates in solution A , the number of surface sites S and the number of adsorbed species SA . Thus, the reaction equation is given by: $A + S \rightleftharpoons SA$.

The relative surface coverage Θ is given as:

$$\Theta = \frac{\Gamma}{\Gamma_{max}} \quad (1.3)$$

with the occupied adsorption sites Γ and the maximal available adsorption sites Γ_{max} . The rate of change of the surface coverage is proportional to the concentration c of the adsorbate, and the number of free adsorption sites is $\Gamma_{max}(1 - \Theta)$. The desorption rate is

proportional to the number of occupied adsorption sites $\Gamma_{max}\Theta$. In case of an equilibrium, the sum of adsorption and desorption rate is zero:

$$\frac{d\Theta}{dt} = k_a c \Gamma_{max} (1 - \Theta) - k_d \Gamma_{max} \Theta = 0$$

with the adsorption and desorption rate constant k_a and k_d . This equals the Langmuir isotherm:

$$\Theta = \frac{Kc}{1 + Kc}$$

with the equilibrium constant $K = \frac{k_a}{k_d}$. With equation 1.3 it follows the Langmuir isotherm of concentrations with the Langmuir constant $K_L = K^{-1}$:

$$\Gamma = \frac{c\Gamma_{max}}{K_L + c}. \quad (1.4)$$

Equation 1.4 can be used in a nonlinear approximation to evaluate K_L . Furthermore, K_L can be used to calculate the Gibbs energy of adsorption ΔG_{ads}^0 [89–91]:

$$\Delta G_{ads}^0 = -RT \cdot \ln \left(\frac{c_{sol}}{K_L} \right) \quad (1.5)$$

where R is the gas constant, T is the temperature, and c_{sol} is the molar concentration of pure water at room temperature (55.5 mol l^{-1}).

An extension of the Langmuir model for multilayer concentrations is the Brunauer-Emmett-Teller (BET) model [92,93]. It is frequently used experimentally in combination with N_2 gas adsorption to calculate the surface of porous materials such as ceramics.

The adsorption isotherms of chymotrypsin on SiO_2 and TiO_2 nanoparticles are measured by BCA protein adsorption experiments [39,40,65] as shown in Fig. 1.7. The adsorbed proteins are normalized to the nanoparticle BET surface and to the protein accessible nanoparticle surface (PAA), directly determined by the radius found in TEM images. Both normalization methods result in different maximal adsorption quantities as shown by Γ_{max} in Fig. 1.7. The difference between BET and PAA surface stems from the porosity of the oxides, allowing an increased adsorption of the small N_2 gas molecules used to calculate the BET surface. Due to the relative difference between Γ_{max} of the BET and PAA surface, SiO_2 has an increased porosity compared to TiO_2 (Fig. 1.7).

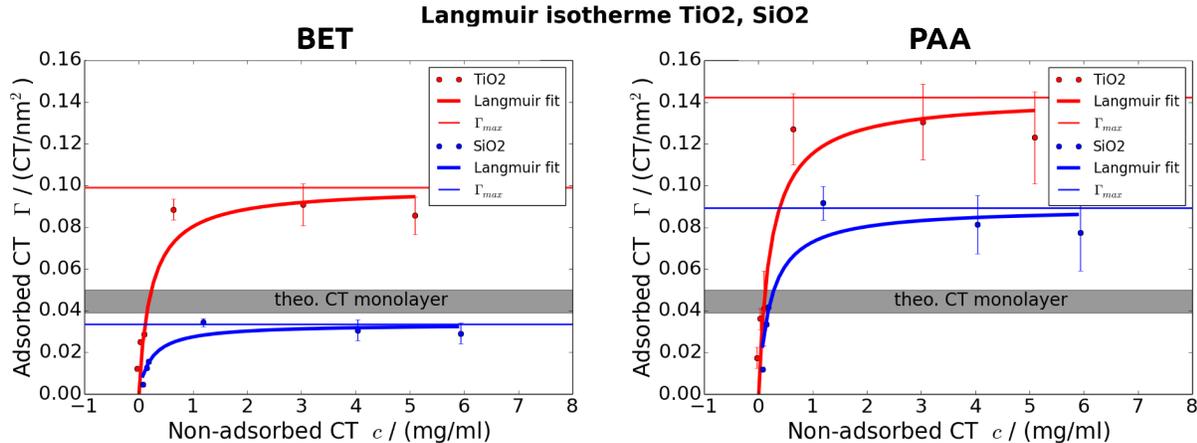


Figure 1.7: Concentration depended BCA protein adsorption experiments of chymotrypsin on TiO_2 and SiO_2 . The adsorbed amount is normalized either on the BET surface, determined by volumetric N_2 adsorption measurements left, or on the protein accessible nanoparticle surface (PAA), determined by the radius found in TEM images (right). The Langmuir isotherms are fitted to the measured data. The maximal adsorbed quantities are shown by the thin colored line (Γ_{max}). The shown theoretical monolayer is calculated by a rectangular packing of the geometries of the proteins. [65]

For both particles types, saturation is reached at about 2 mg ml^{-1} . Γ_{max} indicate that more proteins are adsorbed on TiO_2 . Interestingly, the adsorption free energy ΔG_{ads}^0 (which is independent on the surface normalization) is almost the same for both materials, amounting to $-34.3 \pm 1.0 \text{ kJ mol}^{-1}$ for TiO_2 and $-35.6 \pm 1.0 \text{ kJ mol}^{-1}$ for SiO_2 . To study the origin of the different adsorption quantities by equal adsorption free energies, all-atom molecular dynamics simulations are performed as presented in Chapter 2, 3 and 4.

(III) Surface Anchoring

In order to adsorb on the surface, the protein is in competition with the solvent and its ions. Due to the polarity of the water, the charge of the ions and the entropic properties of both, the solvent exposes dense structures over the charged surface and the protein (Section 1.3.1). The interference of these layers causes free energy barriers or basins. In the case of a negative adsorption free energy, the protein adsorbs on the surface. Dependent on the surface reactive groups, the protein is either chemically adsorbed and forms covalent bonds or is physically adsorbed and forms non-covalent contacts. The physisorption of proteins with a detailed contact analysis, binding motifs and their stability is a central part of the investigations of this thesis (Chapter 2, 3 and 4).

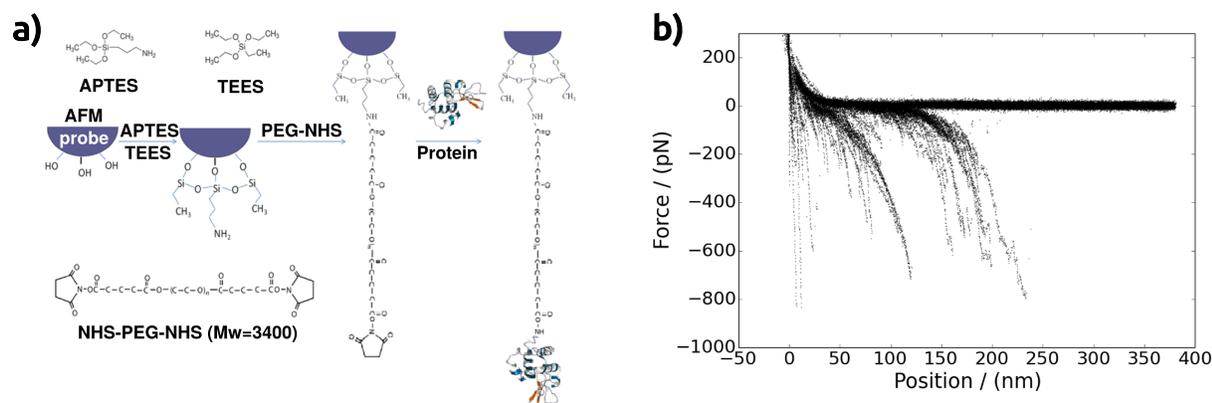


Figure 1.8: Single molecule AFM experiments of lysozyme in contact with SiO_2 . a) The AFM probe is functionalized with lysozyme. The probe is silanized in the desired ratio of APTES to TEES molecules and subsequently conjugated with the PEG linker (3400 Mw). The NHS group of the linker binds to the APTES amino groups [94]. The remaining NHS group of the linker binds to amino groups of lysozyme. b) Superposition of 256 force-distance curves of the retraction of the functionalized probe after the contact with SiO_2 .

The surface contact stability, as well as the binding free energy of the protein at the surface, can be directly measured by single molecule atomic force spectroscopy (AFM). In these experiments, proteins in a diluted concentration are bound to the AFM probe in a sensitive probe preparation (Fig. 1.8 a)) [94].

Due to the atomic resolution of the AFM, a huge variety of different force-distance curves can be measured as shown in the example of lysozyme pulled away from SiO_2 (Fig. 1.8 b)). Depending on the complexity of the system a fraction of the measured curves is assigned to specific protein-surface interactions. The direct interpretation of the force-distance curves can be a challenging task. The superimposed representation of the measured curves may help to identify adsorption patterns or even protein specific fingerprints as observed by the two clustered curves in Fig. 1.8 b)). In Chapter 4 we compare this kind of measured AFM curves with steered MD simulations. In the dynamic non-equilibrium pulling process, the stability of the surface contacts is in competition with the internal stability of the protein including their disulfide bonds. Patterns like a clustering of the measured force-distance curves can either be ascribed to the cleavage of disulfide bonds, or to the presence of more than one lysozyme molecule at the AFM probe. The simulated force-distance curves reveal detachment trajectories on the atomic level and significantly support the interpretation of the measured curves of the AFM experiment.

(IV) Conformational Changes

Once adsorbed on the surface, the ratio between the magnitude of the surface interactions and the protein stability determines surface-induced conformational changes. This process can include a cascade of several stochastically induced angle changes which may result in conformational changes on a relatively long time scale [86]. On the relatively short-time scales examined in atomistic simulations, these conformational changes occur as rare events. In order to sample the rare events often enough to get sufficient statistics, the implementation of enhanced simulation methods is necessary. Further details about the theory of these methods are presented in the Appendix Chapter 7. In this work, a combination of these enhanced methods is adapted to the case of protein adsorption (Chapter 5). These methods are used to investigate surface induced protein denaturations of chymotrypsin on SiO_2 .

Recent experiments have revealed a conformational change of lysozyme and chymotrypsin induced by the adsorption on SiO_2 . In particular, circular dichroism experiments (CD) by Zougrana et al. [24] suggested a loss of the helical structure of chymotrypsin during adsorption on SiO_2 . A similar change of the secondary structure of lysozyme on SiO_2 was measured by CD experiments by various authors [95–97]. Further scanning calorimetry studies by Welzel [98] indicated a reduced denaturation of chymotrypsin adsorbed on SiO_2 in comparison to the adsorption on hydrophobic Teflon.

CD experiments in our own group, performed by Nina Wurzler and Monika Michaelis, show a shift in the spectra during the adsorption of chymotrypsin on SiO_2 , as well as during the heating of the protein in ddH_2O (Fig. 1.9) [99]. The measured spectra are compared to the chymotrypsin spectra of the Protein Circular Dichroism Data Bank (PCDDDB) [100, 101] (dotted lines). The two characteristic minima at about 205 and 230 nm seem to form a fingerprint for chymotrypsin. The adsorption on SiO_2 causes a disappearance of the bulge at 230 nm and a shift of the 205 nm peak to the top-right. A loss of 26% of the helicity and a slight shift of other secondary structures are calculated based on the experimental data (BeStSel [102], tabular insets). A different trend of the CD spectra is observed by heating the solvated protein (Fig. 1.9 b)). An abrupt change of the spectra appears by heating the sample from 55°C to 60°C. The spectra around 205 nm shift to the bottom-left and the prominent bulge at 230 nm disappears. A loss of 50% of the helicity and gain by 5% of the sheet structure is observed.

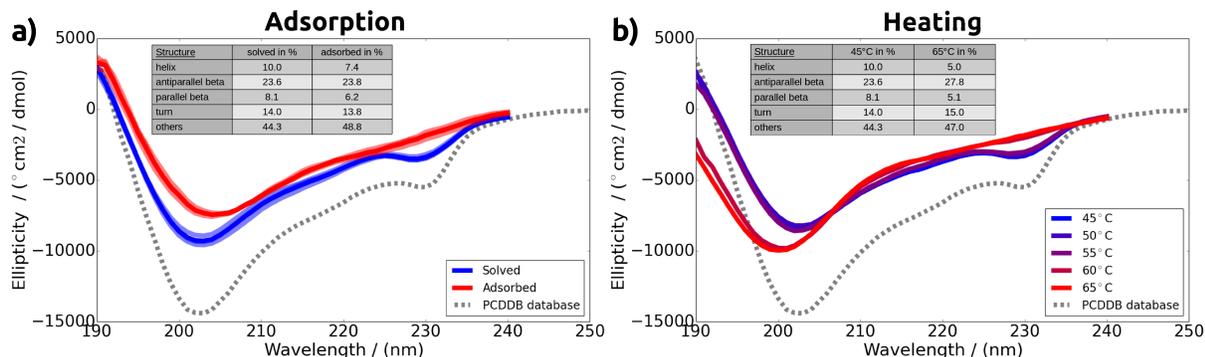


Figure 1.9: a) CD experiments of chymotrypsin in ddH₂O measured after an incubated for 24 h on SiO₂ nanoparticles [99]. The reference dashed line is the chymotrypsin CD spectra taken from the Protein Circular Dichroism Data Bank [100, 101]. The secondary structure fraction is shown in the tabular insets, computed by the BeStSel webserver [102]. b) Same as a) but with the protein heated up to 65°C.

The conformational change of chymotrypsin and lysozyme adsorbed on SiO₂ and TiO₂ is studied by straightforward standard MD simulations in Chapter 2. However, it is assumed that conformational changes of proteins of this size on the hydrophilic surfaces take place over longer time scales than standard MD simulations can provide. Therefore, advanced MD methods are extended and further developed for this particular case of protein adsorption to accelerate the simulations and to analyze the conformational change of proteins, as presented in Chapter 5. The advanced methods are based on a combination of free energy sampling approaches such as Metadynamics, Umbrella Sampling, steered MD and Solute Tempering, which are applied on carefully chosen collective variables of the system (Appendix 7.5).

Lysozyme and Chymotrypsin Binding Motifs on Silica

In order to identify the binding motifs of proteins on silica the adsorption of lysozyme and chymotrypsin is investigated by molecular dynamics (MD) simulations in comparison with adsorption experiments. Besides covalent tethering to the surface, also simple physisorption from concentrated enzyme solutions can lead to sufficiently stable adsorbed protein layers and guarantee enzymatic activity in the adsorbed state [103, 104]. One part of an optimal functionalization process is the prediction of the preferred protein adsorption orientation. An optimal orientation includes an accessibility of the substrate to the enzymatically active site. The functionalization is further optimized by the presence of amino acid motifs in contact with the surface and other proteins to form stable and densely adsorbed enzyme layers.

Several studies of lysozyme adsorption on materials surfaces indicate a major role played by the mutually facing charge density distributions in governing the formation of adsorbed protein layers. Bharti et al. [105] found no adsorption on silica surface at pH of 4.0 or below when the surface is essentially neutral. Instead, protein monolayers form at pH 7 - 8 [97, 105, 106] and multilayers were observed at higher pH, corresponding to the isoelectric point (IEP) of the protein itself of about 10. Measurements of the ζ -potential of silica as a function of the adsorbed protein indicate a progressive shift of the potential from negative (pure silica) to positive as a result of the electrostatic screening and charge reversal induced by the adsorbed proteins [105–107].

After adsorption on silica at low concentration, Su et al. [108] identified a preferred

side-on orientation by means using neutron reflection. For higher protein surface concentrations, a more tilted orientation is observed. Total internal reflection fluorescence (TIRF) measurements of lysozyme on a flat silica surface also indicate side-on adsorption via a positively charged region located on the opposite side with respect to the enzymatic active site [109, 110]. Kubiak-Ossowska et al. [111] performed standard MD simulations of lysozyme on a simplified mica surface model and found preferential side-on binding through the N, C-terminal face, with ARG128, ARG125, ARG5 and LYS1 playing an important role as surface binding sites.

Experimental kinetic studies on the timescale of several minutes to hours indicated that the adsorption of lysozyme on silica is not necessarily stable after the first contact with the surface. Using time-resolved ζ -potential and TIRF measurements Daly et al. [110] identified a two-stage adsorption process with a temporary first contact orientation and a later irreversible reorientation to a final side-on orientation with the active site facing away from the surface. Wertz et al. [112] suggested a similar roll-over behavior of lysozyme on hydrophobic surfaces. Such roll-over may be consistent with molecular dynamics simulations by Wei et al. [113], who suggested a faster adsorption for an end-on orientation, but an eventually more stable adsorption for a side-on orientation. The authors explain this observation in terms of a kinetic barrier associated with the displacement of the surface hydration shell during adsorption, which is smaller when fewer molecules need to be displaced (end-on).

Similar to lysozyme, the adsorption of chymotrypsin on silica has also been measured experimentally. Zoungrana et al. [24] measured adsorption isotherms of the protein on silica using UV-spectroscopy and found a saturation of the protein coverage of about a close-packed monolayer at pH 7.1. The adsorption was found to be irreversible with respect to rinsing of the surface after formation of an adsorbed protein layer.

As for lysozyme, adsorption of chymotrypsin is expected to lead to a measurable change of the surface ζ -potential [106]. However, Vasina et al. [107] pointed out to a different behavior of these two enzymes after adsorption to silica. While lysozyme adsorption leads to an increase of the ζ -potential, adsorption of chymotrypsin, even in monolayer concentration, leave the surface charged negatively. This may point to a very specific orientation of the adsorbed protein with exposure of a negative surface potential region despite the overall positive net protein charge.

These findings motivate the atomic investigations by means of MD simulations of the

protein adsorption process. These methods reveal the calculated DLVO attraction between protein and surface, protein orientations and detailed investigations of the atomic binding motifs [34].

2.1 Results

The technical details of the presented MD simulations are presented in Hildebrand et al. [34] and in the Appendix Chapter 7.

2.1.1 Static Single-Point Calculations of the Protein Orientation

The interaction between single protein molecules and an amorphous silica surface is evaluated by means of static single-point calculations in an implicit solvent, thus effectively taking into account only DLVO interactions. These comprise both van der Waals attraction forces and electrostatic repulsion or attraction forces including entropic properties of the electrolyte between two charged surfaces [66, 67] (Section 1.3). The calculated total energies and forces as a function of the normal distance to the surface are shown in Fig. 2.1 considering 36 different protein orientations. The zero-energy point is set arbitrarily as the average energy value at the maximum distance of 50 Å. The average values of energies and forces over the 36 orientations are represented as bold lines of colors corresponding to the four considered pH values, consistently with the adsorption experiments shown in Fig. 1.6. The semitransparent areas have thicknesses equal to the standard deviations among the different orientations.

As a general trend, predominantly attractive interactions are observed for large height values both in the case of chymotrypsin and of lysozyme. Repulsive interactions, if present, are limited to the region in close surface proximity, between ~ 10 and ~ 5 Å, as expected, due to the entropic repulsion of the ion diffusion layers in the DLVO theory. For both enzymes, the interaction strength and the attraction force are enhanced with increasing pH values, which is consistent with the trend observed in the adsorption experiment. At low pH values, that is for small surface charge densities, similarly small attractive forces (of the order of 5-10 pN at the most) are observed for both chymotrypsin and lysozyme. At the highest pH value, however, the attractive force for lysozyme reaches ~ 64 pN, which is three times higher than the force acting on chymotrypsin (~ 21 pN, on average). Similarly, almost three times larger interaction energy values are observed for

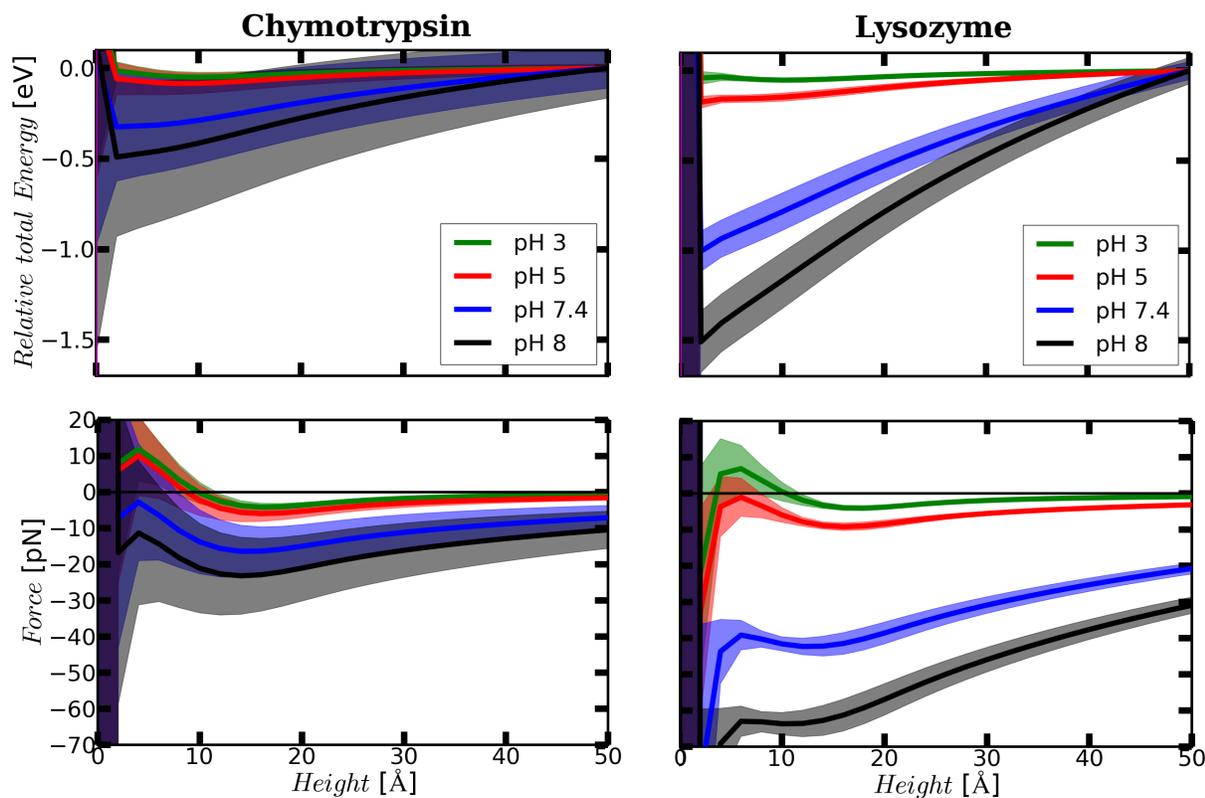


Figure 2.1: Relative total energy and force profiles of chymotrypsin and lysozyme over the silica surface computed within the Generalized Born implicit solvent model at different pH values. The solid lines are the average and the colored semitransparent areas are the standard deviations obtained from 36 different protein orientations. A negative force value is an attractive force onto the protein in the normal direction of the silica surface.

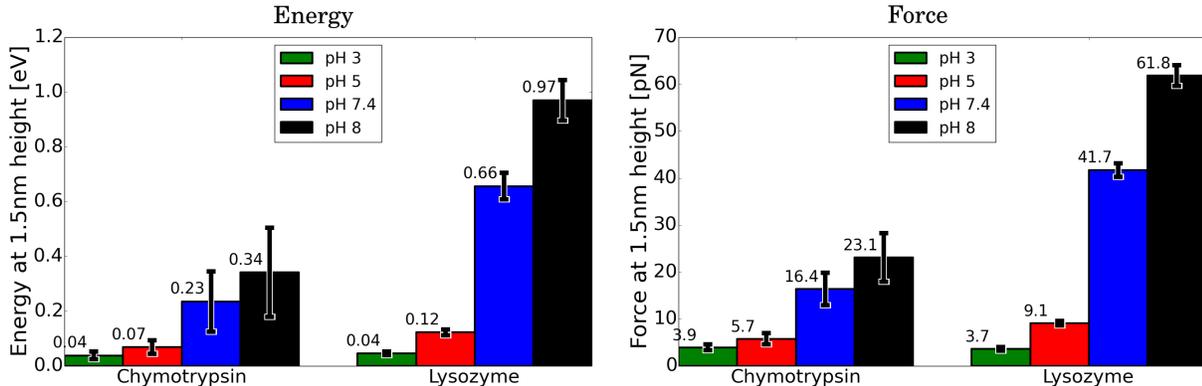


Figure 2.2: The magnitude of energy and force values of chymotrypsin and lysozyme over the silica surface at a height of 15 Å in implicit solvent at different pH values. The error bars show the standard deviation among different protein orientations.

lysozyme than for chymotrypsin. The values computed at the height of 15 Å are reported in Fig. 2.2. Here it becomes evident that the higher total positive net charge of lysozyme results in an overall stronger interaction with the negatively charged surface than the less positively charged chymotrypsin. These results, while consistent with simple electrostatic arguments, are not able to explain the experimental evidence that chymotrypsin adsorbs in a much larger amount than lysozyme at all pH values (Fig. 1.6).

Notably, also the standard deviations increase with pH, as visible both from the increasingly wider force and energy profiles in Fig. 2.1 and from the error bars in Fig. 2.2. This highlights an increased influence of the protein orientation on the attraction as the interaction strength increases, an effect that is much more evident for chymotrypsin than for lysozyme. To investigate more in depth this issue, we depict surfaces of equal electrostatic potential around the proteins at values of ± 1 kBT e^{-1} (or ± 25.9 mV) (Fig. 2.3 a), d)). Lysozyme presents a strong positive potential, due to a high density of positively charged amino acids on the opposite side on the binding pocket, a hydrophobic environment inside the binding pocket and well distributed charged residues surrounding it. Chymotrypsin is instead characterized by a clear separation of a positive from a negative potential region, which is associated with a strong dipole moment of 529 D (Fig. 2.3 b)) (Appendix 7.7). For comparison, the dipole moment of lysozyme is 3.5 times smaller and amounts to 154 D (Fig. 2.3(e)). Due to its strong dipole moment, rotation of chymotrypsin over the negatively charged amorphous SiO₂ surface is expected to give raise to large energy differences, thus explaining the wide interaction profiles in Fig. 2.1.

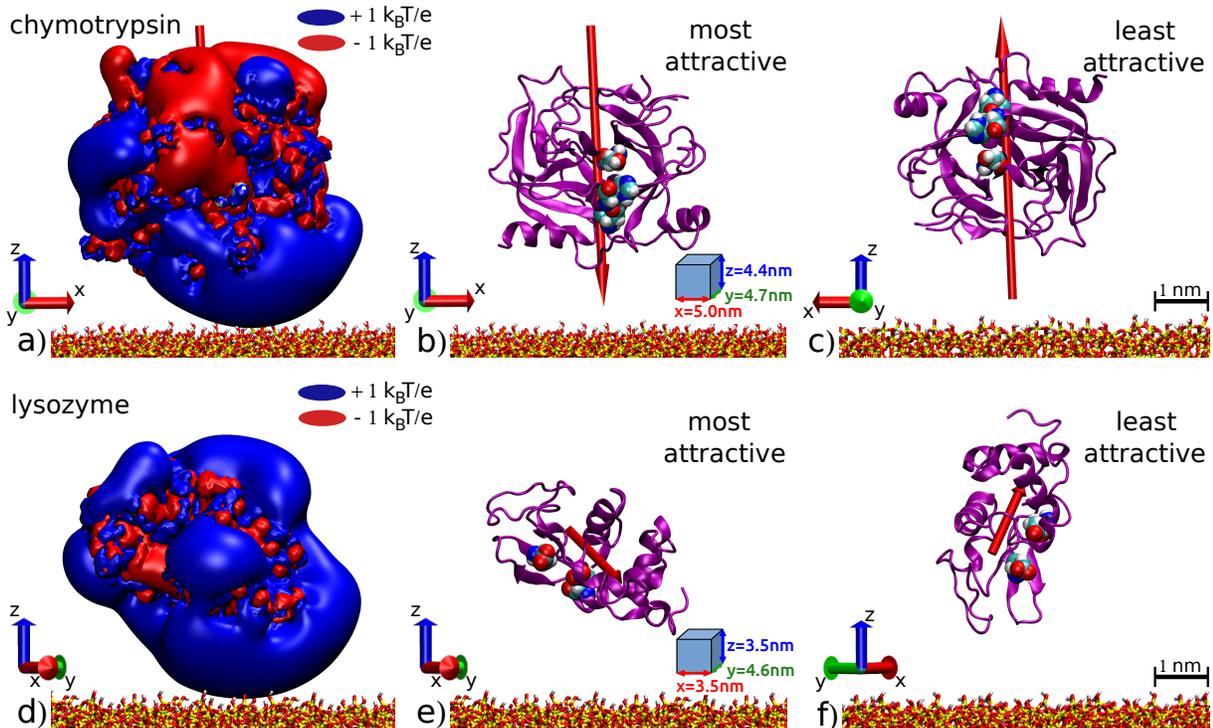


Figure 2.3: Most and least attractive orientations of chymotrypsin (top) and lysozyme (bottom) over the amorphous SiO_2 surface. The amino acids of the enzymatically active sites are depicted as van der Waals spheres (GLU35 and ASP52 for lysozyme, HIS57, ASP102 and SER195 for chymotrypsin). Isovalue surfaces of the electrostatic potential around the proteins at -1 and $+1 k_B T e^{-1}$ are colored red and blue, respectively, in a) and d). The red arrows represent the enzyme dipole moments ($529.1 D$ or $1.76 \cdot 10^{-27} \text{ Cm}$ for chymotrypsin and $153.8 D$ or $0.5 \cdot 10^{-27} \text{ Cm}$ for lysozyme) and are scaled by a factor of 0.1 in both cases. The protein dimensions used to compute the theoretical monolayers in Fig. 1.6 are shown as blue boxes in b) and e), calculated from the maximum atom distances along the shown main axes.

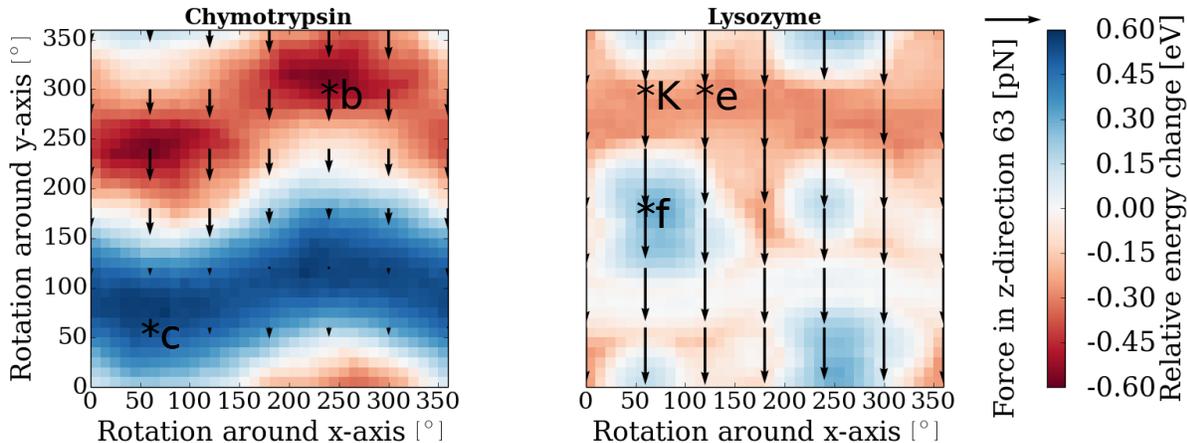


Figure 2.4: Schematic representations of the relative energy changes (colors) and total forces (arrows) experienced by chymotrypsin (left) and lysozyme (right) at pH 8, calculated as in Fig. 2 at the height of 10 Å from the surface. Downward pointing arrows show an attractive force towards the surface. The energy changes are relative to the average energy value of all orientations at this height. The protein orientation is defined by the rotations around the x- and y-axis starting from an arbitrary orientation. The letters b and e mark the most attractive orientations, the letters c and f the least attractive orientation, according to the panel labels in Fig. 2.3. K marks the most attractive orientation found by Kubiak-Ossowska et al. [111].

For each protein, analysis of the interaction profiles enables us to define a “most attractive” and a “least attractive” adsorption orientation. This is done by averaging the computed interaction forces of each orientation over a height range from 2 to 50 Å, and considering the average force value as a representative measure of the interaction strength associated with that orientation. The resulting orientations at pH 8.0 are presented in Fig. 2.3. Not surprisingly, in the most attractive orientations of both proteins a large positive region of the electrostatic potential faces towards the negatively charged surface. Furthermore, the dipole moment of chymotrypsin lies almost perpendicular to the surface and points in opposite directions in the most and least attractive orientations, as one may intuitively expect. In the most attractive orientation, the chymotrypsin active pocket is oriented sideways and the two α -helical strands are facing the surface, whereas the protein is turned upside-down in the least attractive orientation. The smaller dipole moment of lysozyme still have a component favorably oriented towards the surface in the most attractive orientation, but it also has an equally important component parallel to the surface (Fig. 2.3 e)). A side-on orientation with the binding pocket facing the surface is most attractive, while an end-on orientation with the binding pocket pointing sideways is the least attractive.

The landscape of the interaction energy with variable protein orientations at a set height value of 10 Å and pH 8.0 is represented in two-dimensional color maps in Fig. 2.4. The relative energy changes span from -0.6 eV (red) to +0.6 eV (blue) with respect to the average interaction energies at this height, which is set to zero (white) in this representation and correspond to the bold colored lines in Fig. 2.1. Superimposed to the energy landscape are arrows, whose lengths are proportional to the absolute force components perpendicular to the surface.

In the case of chymotrypsin, the energy landscape is composed of two belt-like regions of either high or low relative energy, corresponding to the dipole pointing upwards or downwards with respect to the surface, respectively. An energy difference of about 1.2 eV is needed to rotate the protein from the most attractive to the least attractive orientation (marked with b and c, respectively, as in Fig. 2.3). Instead, rotation around the dipole axis (corresponding to the belt-like red region in Fig. 2.4) is possible without effort and without losing a favorable interaction.

Since the lower energy profiles are also steeper and the higher energy profiles flatter, the force values are less attractive or almost zero (although never repulsive) for orientations of higher energy. In comparison to chymotrypsin, lysozyme has a higher and more homogeneous attraction force distribution around 63 pN. Its energy landscape is flatter and also presents a belt-like region of relatively constant low energy at about 270° around the y-axis. This belt corresponds to a side-on orientation of the protein and contains both our most attractive orientation (marked with e) and the most attractive orientation suggested by Kubiak-Ossowska et al. [111] (marked with K). The least attractive, side-on orientation (marked with f) is within a region of high relative energy (but still of strongly attractive forces), and is separated from the most attractive orientation by an energy difference of 0.6 eV. We can assume that during the approach towards the surface starting from a random orientation, the protein may be able to freely rotate and reach the nearest more favorable orientation. However, to investigate the protein/surface interactions at distances smaller than about 5 Å, it is necessary to perform Molecular Dynamics simulations in explicit water solvent and without constraints on the protein structure, as described in the next section.

2.1.2 Molecular Dynamics Simulations

In the static calculations used to analyze the optimal protein orientations approaching the surface, the protein structure is kept fixed and the solvent described by a continuum theory. To investigate whether the obtained orientations indeed lead to stable surface adsorption, we now perform MD simulations in an explicit solvent under periodic boundary conditions (Appendix 7.1.3). In the starting configurations, the proteins are placed at the height of 10 Å from the surface. During the simulations, we monitor both the RMSD of all protein atoms with respect to each initial (desorbed) configurations and the evolution of selected secondary structure elements such as β -sheet and α -helices. This is shown exemplary for lysozyme in its most attractive orientation in Fig. 2.5 along a 300 ns MD trajectory in which the protein adsorbs to the surface after about 15 ns and remain stably bound therein (see below). The RMSD value increases to 2.0 Å within the first 50-60 ns, and fluctuates by ± 0.5 Å for the rest of the trajectory. These fluctuations are due to thermal noise and are of the same magnitude as in simulations in bulk liquid water, in the absence of the surface. Also, the secondary structure components fluctuate around stable values during the entire simulation, with larger fluctuations visible for α -helices, which are known to be more flexible than β -sheets. The adsorption of chymotrypsin follows the same trend, indicating that in both cases no obvious changes of either tertiary or secondary structure take place on this short time scale. Therefore, in the following we analyze trajectories starting from different orientations lasting each about 50 ns, to keep the total computational time within reasonable limits without a loss of relevant information.

In addition to the least and most attractive orientations displayed in Fig. 2.3, we also consider further four orientations, altogether corresponding to the six sides of an orthorhombic box enclosing the protein (Fig. 2.6 and 2.7, insets of the left columns). First, we concentrate on chymotrypsin, whose configurations at the beginning of each simulation (from I to VI) and after stable adsorption are shown in Fig. 2.6, left column. The evolutions of the height above the surface and of the dipole orientation expressed as the cosine of the angle between the dipole and the surface normal, are plotted in the corresponding graphics in the middle column. When the dynamics are started with chymotrypsin in the “most attractive” orientation (I), it first moves about 10 Å away from the surface and rotates by about 90 degrees. Only after about 20 ns of simulation does the orientation gradually moves back to the initial one (with the dipole pointing towards the surface), while the protein approaches the surface, binds to it after about 30 ns, and eventually

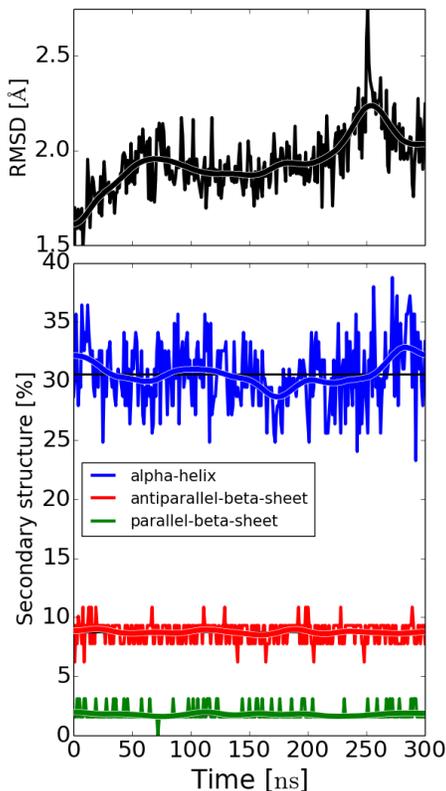


Figure 2.5: Evolution of the RMSD and of the amount of selected secondary structure elements of lysozyme adsorbed on amorphous SiO_2 during an explicit-solvent MD simulation. A Hamming window is used to smooth the data, as shown with superimposed bold lines. The secondary structure components are computed by the DSSP-Method. The averaged values over the whole simulations are $30.5 \pm 2.8\%$ α -helices, $8.7 \pm 0.9\%$ antiparallel β -sheets and $1.8 \pm 0.6\%$ parallel β -sheets.

remains stably adsorbed with the two α -helices in surface contact.

Starting in the “least attractive” orientation (II) results instead in adsorption at the periodically repeated image of the surface slab. Also, in this case, the protein rotates first by about 90° and only later turns back to the original orientation, although, the surface approach is more straight, with the first contacts forming after about 10 ns despite the initial protein-surface distance of about 40 \AA . The final adsorbed geometry and orientation is very similar to the previous case, with the dipole moment facing the surface at a narrow angle to the normal and the α -helices anchoring at the surface. For all other initial orientations with the dipole parallel to the surface, surface adsorption proceeds quickly (within 10 ns or less) either on the pristine slab (simulations IV, V, VI) or its periodical image (simulation III). Adsorption is always accompanied by a rotation of the dipole moment towards the surface until initial adsorption takes place at an angle of at most 45° , after which further rotation is hindered by the presence of stable surface contacts. The resulting absolute value of the cosine of the orientation is always in the range of 0.7 to 1.0, which remarks the stability of the identified most attractive orientation.

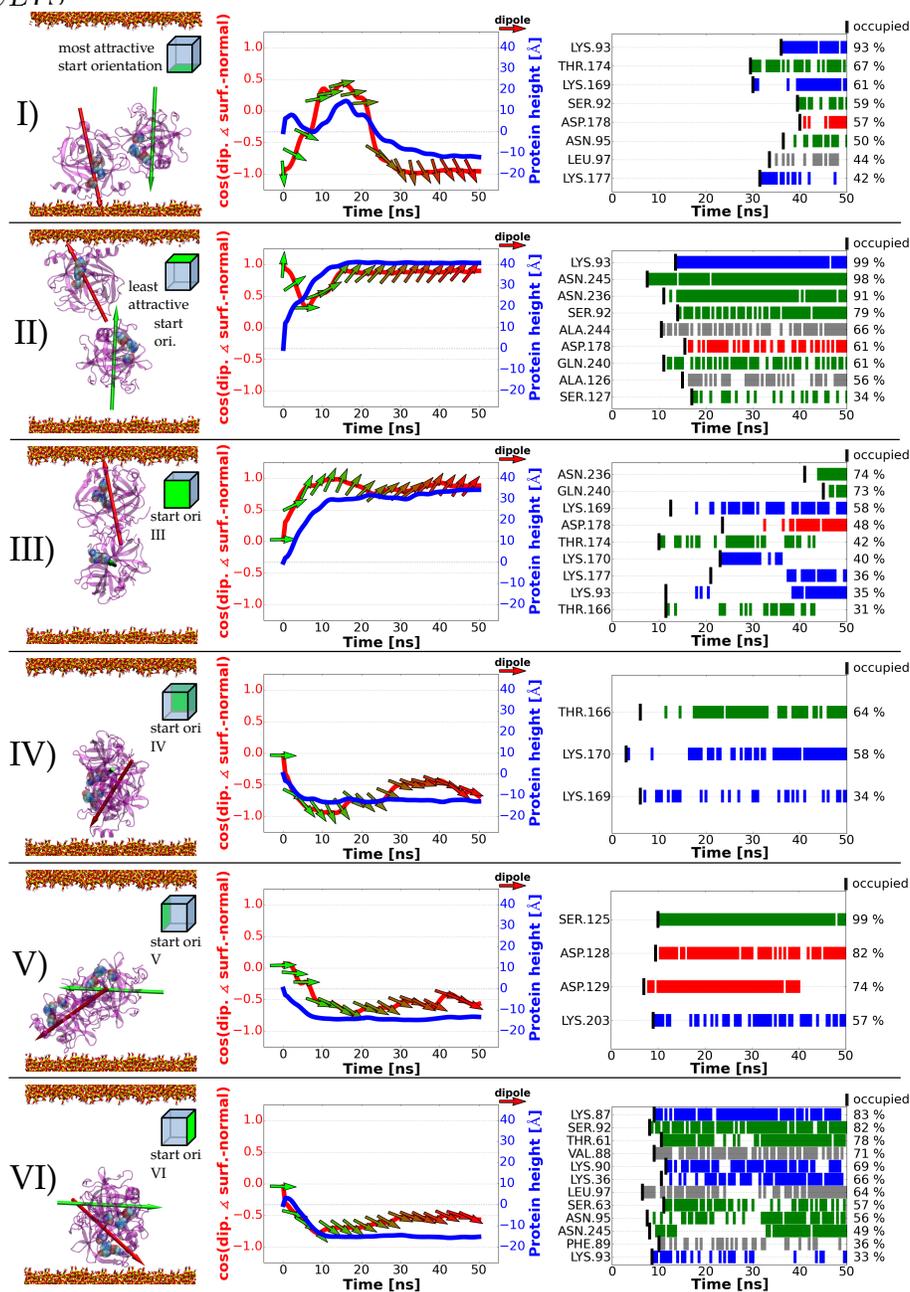


Figure 2.6: (left) Initial and final snapshots of MD simulations of chymotrypsin adsorption starting from six different orientations highlighted by the green faces of an orthorhombic box enclosing the protein (inserts). The dipole moment is green and red in the initial and final configurations, respectively. (middle) Time evolution of the protein height with respect to the initial position (blue curve) and of the cosine of the angle between dipole and surface normal (red curve). The small arrows depict the actual dipole moment orientations. (right) Temporal analysis of the amino acids in contact with the surface, for a threshold distance of 2.4 Å between any atom of the surface and any atom of the amino acid. Contact is marked with a bar colored according to the positive (blue), negative (red), polar (green) and non-polar (gray) character of the side chain. The first contact is marked black. Presented are only the amino acids which keep surface contact for more than 30% of the time after initial contact and for more than 5% of total simulation time, using a frame-rate of 0.5 ns.

In this context, we define as a “surface contact” an interaction between any surface atom and any protein atom which leads to an atom-atom distance of less than an arbitrarily defined cut-off distance of 2.4 Å. This roughly corresponds to the position of the first minimum in the oxygen-oxygen radial distribution function (Appendix 7.6) in bulk water and is thus a representative value under which hydrogen bonds, ion-ion bridges or hydrophobic interactions (involving H atoms) can be considered to be present. An analysis of the contacts between chymotrypsin and the SiO₂ surface during our six MD simulations is shown in Fig. 2.6, right column. As anticipated above, the initial contacts (marked with black bars) take place within 10 ns, with the exception of simulation (I). The major part of contacts is mediated by amino acids with positive (lysine) or polar (asparagine, glutamine, serine, threonine) side chains. Notable, however, is the presence of stable contacts also with a smaller number of negative (aspartic acid) and non-polar (leucine, glycine) residues, despite the overall negative surface charge density and hydrophilicity of amorphous SiO₂ at pH 8.0.

It is useful to classify the residues as “tight binders” if they remain bound to the surface for at least 73 % of the time after the first contact, and as “frequent binders” if they form surface contacts in at least two of the six simulated cases. As shown in Fig. 2.8 a) and c), most of the contacts between chymotrypsin and the surface take place via the α (I) and α (II) helices in the positive region of the protein’s surface potential (see Fig. 2.3). These helices include residues from SER160 to TYR167 and from ARG226 to ALA239 (the C-terminal region), respectively. Interestingly, around the α (II) motif are located residues which are at the same time tight and frequent binders (SER92, LYS93, ASN236, GLN240, ASN245). Around the α (I) motif are located residues which are frequent, but not tight binders (LEU97, THR166, LYS169, LYS170, THR174, LYS177, ASP178). Other residues, on the other side, are tight but not frequent binders (THR61, LYS87, SER125, ASP128, ASP129). This rationalization may give a hint about the more or less important role played by each single amino acid in chymotrypsin adsorption experiments.

Analysis of six MD simulations of the adsorption of lysozyme is presented in Fig. 2.7. Starting from the most attractive orientation (I), the protein moves towards the surface without significant rotation of the dipole moment and adsorbs within 10 ns with the dipole pointing downwards and the binding cleft pointing sideways. In the least attractive orientation (II), despite the fact that attractive forces were predicted from the implicit-solvent simulations, the enzyme is initially repelled by about 20 Å before inverting direction and

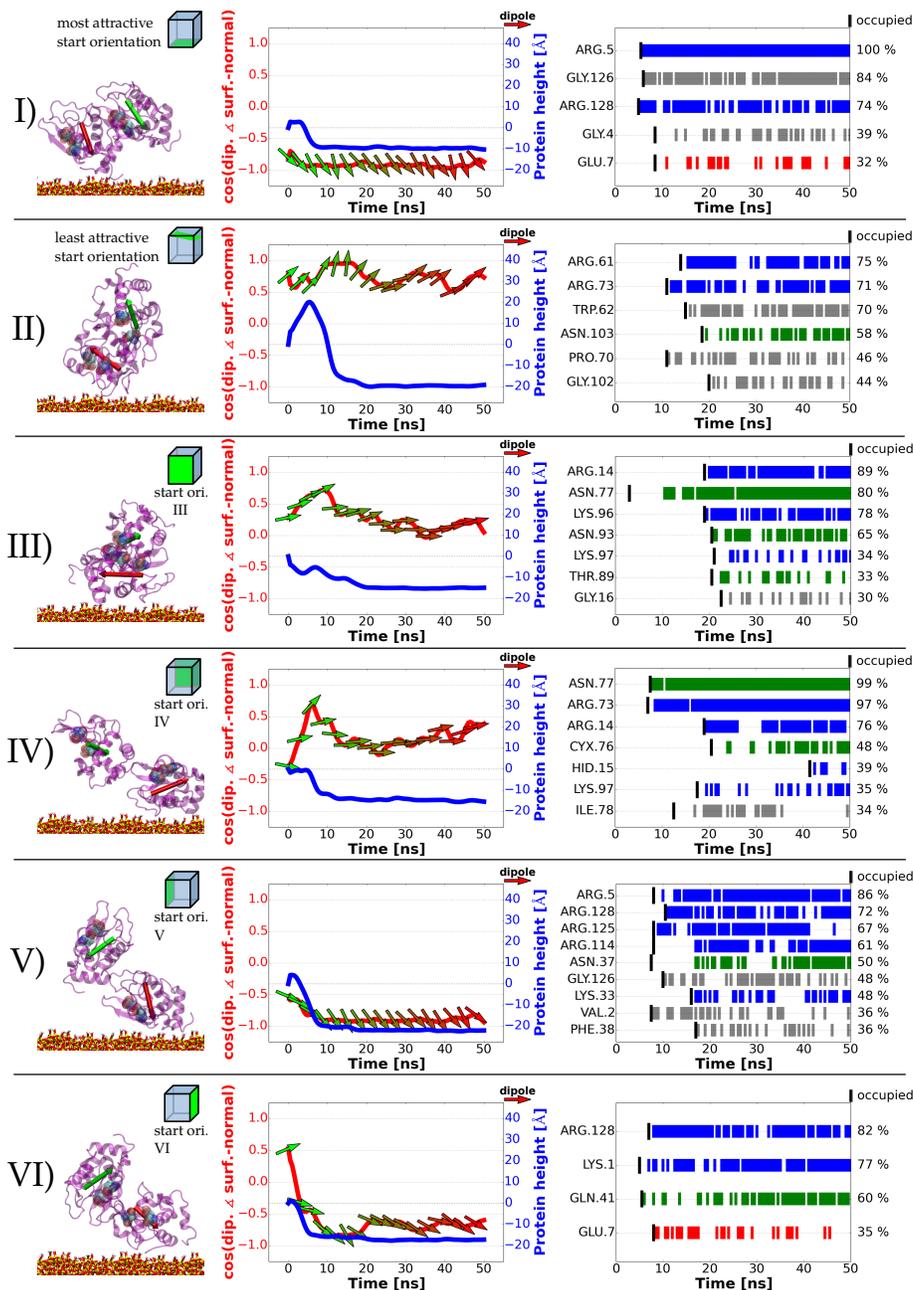


Figure 2.7: Analysis of six MD simulations of lysozyme adsorption. Left: initial and final configurations. Middle: time-dependent evaluations of the dipole orientation and protein height. Right: analysis of protein-surface contacts. See caption of Fig. 2.6 for further details.

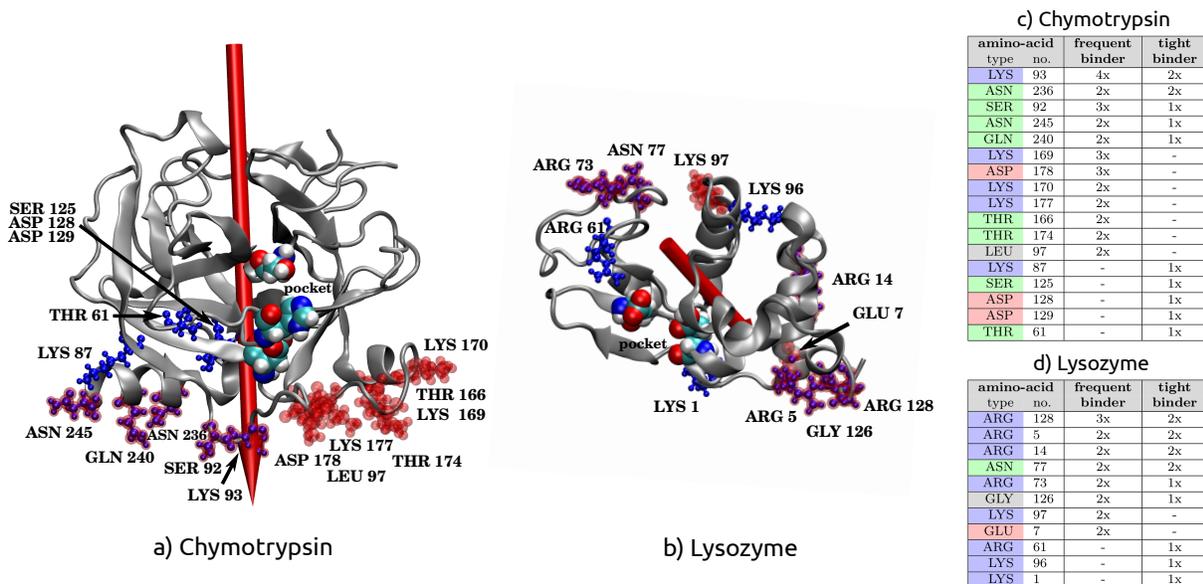


Figure 2.8: Summary of “tight” (blue, cpk style) and “frequent” (red, semitransparent) binder residues observed in MD simulations of chymotrypsin (a,c) and lysozyme (b,d) adsorption on amorphous SiO_2 surfaces. “Frequent binders” form surface contacts in at least two of the six simulated cases (see Figs 2.6 and 2.7); “tight binders” remain bound to the surface for at least 73 % of the time after first contact. The atoms of the active enzymatic sites are highlighted with van der Waals spheres. The tables include the number of simulations in which each identified residue behaves as a frequent or tight binder.

adsorbing in a side-on orientation with the dipole pointing away from the surface. Notable, in this case, is the oscillations of the dipole moment while the protein keeps contact with the surface, revealing a more pronounced mobility of the adsorbed molecule than in the case of chymotrypsin. Similar adsorption patterns are observed for the remaining orientations (III to VI), in all cases leading to adsorption with the binding cleft pointing either sideways or towards the solvent. This orientation is based on the fact that the positive surface charge on lysozyme is well distributed over the protein surface, apart from the region around the binding cleft (see Fig. 2.3), suggesting a less clear preference for a specific binding orientation than in the case of chymotrypsin. This observation is consistent with the less structured potential energy surface in Fig. 2.4 and with the narrower interaction profiles in Fig. 2.1. Indeed, in the final adsorbed configuration the dipole moment does not show a clear preferred direction, in contrast with the results concerning chymotrypsin.

In all of these lysozyme simulations, the first contact with the surface is established within ~ 10 ns. Arginine, which is more often present in the lysozyme sequence compared

to chymotrypsin, is the predominant contact for all orientations (Fig. 2.7, right column). Other contacts are formed by positively charged or neutral polar residues such as lysine, asparagine, glutamine and threonine. As for chymotrypsin, non-polar amino acids also form some relatively stable contacts, with GLY126 behaving as a tight and frequent binder. As shown in Fig. 2.8(b), the binding residues are distributed all around the protein's surface, apart from the region surrounding the binding cleft, as explained above.

In line with the arguments first put forward in Schneider et al., [55] the position of the binders at the SiO₂/water interface is strongly mediated by the local variation of the water density in surface proximity. The interfacial regions of water density 40% higher than the bulk in the absence of adsorbed proteins are depicted in Fig. 2.9 a) with a blue isodensity surface. A qualitative correlation of these regions with the location of the contact points via charged and polar amino acid can be observed after averaging the position of the corresponding residues along all six MD trajectories of lysozyme (Fig. 2.9 b)). Immediately visible is the preferential location of positively charged residues (arginine in particular) around the deprotonated silanol groups of the surface. In contrast, the contacts via negatively charged and polar residues are distributed in between the silanols, whereas contacts via non-polar amino acids are located within regions of water density lower than the bulk.

These effects can be better quantified by analyzing the radial distribution functions (RDF). The RDF of water and different types of amino acid residues around the oxygen atoms of the deprotonated and protonated silanol groups (Fig. 2.10 a) and b)), averaged over all six lysozyme adsorption trajectories shown in Fig. 2.7. The structuring of the hydrophilic residues within the first water shell of the deprotonated silanols (between 1.5 and 3.3 Å) is remarkable, whereas hydrophobic residues remain at larger distances of at least 3.5 Å. This is mostly due to the tight contacts of positively charged arginines to the terminal H atoms of the guanidine heads contributing to the first RDF peak at about 2 Å and the central N atom contributing to the second peak at about 3 Å. Only minor contributions arise from lysine or neutral hydrophilic residues. In contrast, both hydrophilic and hydrophobic residues can penetrate into the first water shell of protonated (neutral) silanol groups. Among the hydrophilic residues, both positively charged and neutral groups contribute to the RDF profiles. Therefore, already from this analysis limited to silanol binding sites it can be concluded that direct ionic interactions do represent a significant contribution to stable adsorption, but are also accompanied by contributions coming from neutral hydrophilic or even hydrophobic amino acids.

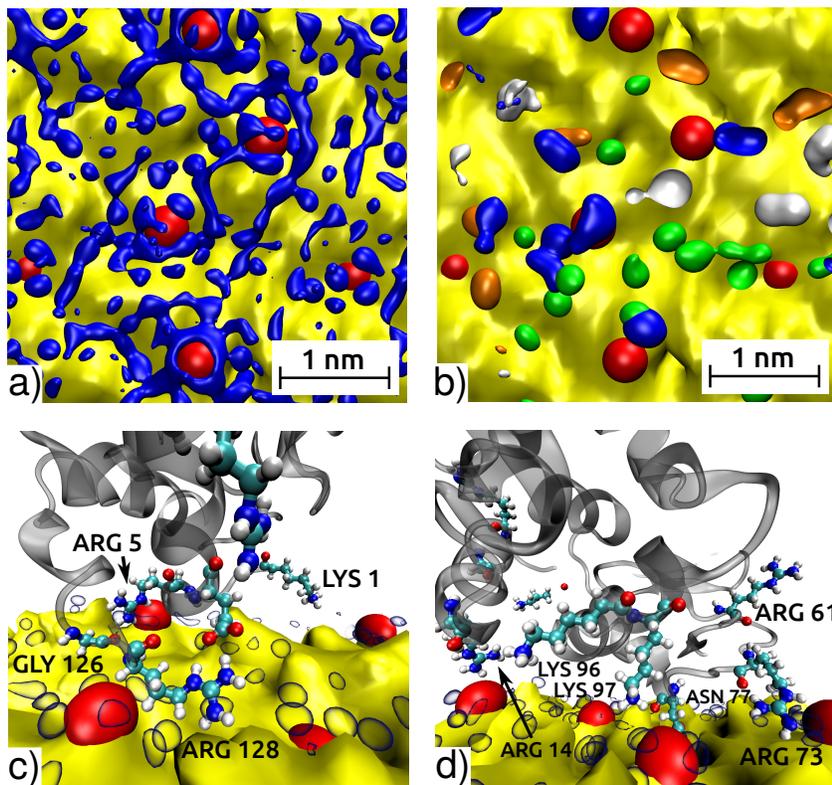


Figure 2.9: Binding motifs of lysozyme adsorbed on amorphous SiO_2 identified in the MD simulations shown in Fig. 2.7. Silica is presented as a yellow Gaussian-density surface of the atomic positions using the QuickSurf tool included in VMD [114]. The negatively charged silanol groups are marked by red VDW-type atoms. (a) Regions of water density 43% higher than the bulk (blue isosurface), averaged over a 50 ns MD simulation in the absence of adsorbed proteins. (b) Density regions of binding residues averaged over all six MD simulations in Fig. 2.7. Positive residues: blue isosurface at $1.2 \cdot 10^{-4} \text{ \AA}^{-3}$. Negative residues: orange isosurface at $1.2 \cdot 10^{-4} \text{ \AA}^{-3}$. Neutral polar residues: green isosurface at $0.5 \cdot 10^{-4} \text{ \AA}^{-3}$. Non-polar residues: white isosurface at $1.5 \cdot 10^{-4} \text{ \AA}^{-3}$. (c) Snapshot from the MD simulation (I) at $t = 40 \text{ ns}$, showing the cooperative binding motif ARG128, ARG5, LYS1 and GLY126. (d) Snapshot from the MD simulation (IV) at $t = 33 \text{ ns}$, showing the cooperative binding motifs (ARG73, ASN77) and (LYS96, LYS97), as well as the two other tight binders ARG14 and ARG61. The blue-grey semitransparent isosurfaces in (b) and (c) include regions of water density twice as high as in the bulk.

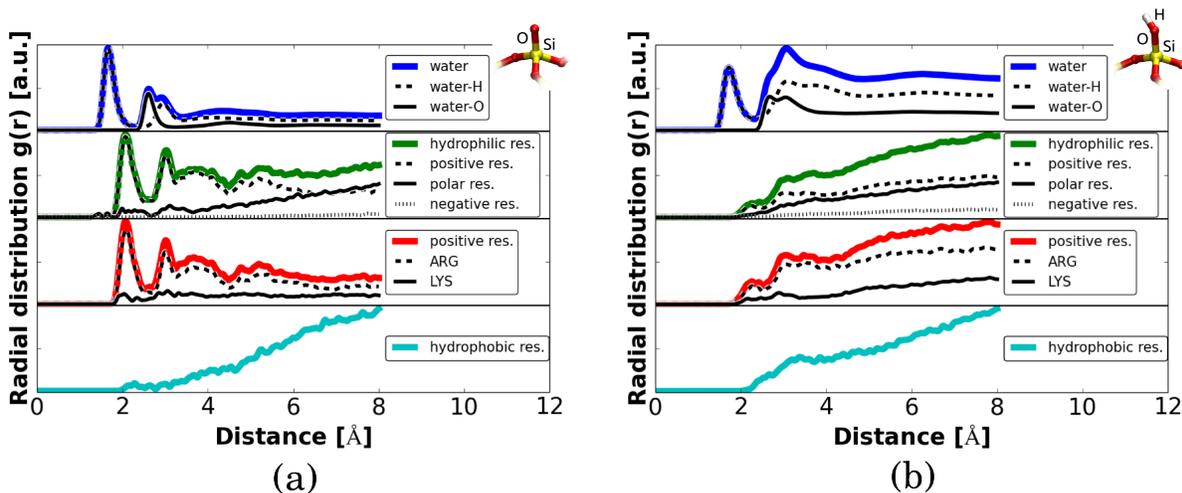


Figure 2.10: Radial distribution functions (RDF) of the water molecules and the amino acid residues (types as labeled) of lysozyme around deprotonated (a) and neutral (b) surface silanol groups, schematically represented in the top right corners of the panels. The RDF are calculated as an average over the six trajectories of lysozyme adsorption shown in Fig. 2.7 using radial shells with thicknesses of 0.01 \AA , and are smoothed by a Hamming algorithm with a window size of 20 successive points. All curves are scaled with respect to the average atom densities of the corresponding residue types.

It is also interesting to observe a cooperative adsorption behavior of the residues, which help to stabilize the anchor points to the surface. For example, ARG128 often adsorbs concomitantly with ARG5, in a few cases supported by LYS1. This is highlighted in the snapshot of simulation (I) shown in Fig. 2.9 c). Another anchoring motif comprises ARG14 together with LYS97 (found in simulations (III) and (IV) of Fig. 2.7), as shown in Fig. 2.9 d). In these cases, the spatial distance between the very flexible positively charged side chains correlates very well with the average distance of the negatively charged, deprotonated surface silanol groups. In another case, ARG73 binds close to a charged silanol whereas ASN77 is anchored to a spot of high-density water between two silanols. By visually analyzing the anchor points, we are also able to explain the unexpected tight-binding behavior of GLY126, which forms a stable surface contact because of its vicinity to ARG128 (Fig. 2.9 c)). This is evident from a comparison between the residues highlighted in Fig. 2.8 a) with the contact point analysis in Fig. 2.6 (right column). Typical binding motifs are for instance ASN245 together with ASN236, GLN2240 and SER92; the lysine triplet LYS177, LYS169, LYS 170; or SER 125 together with ASP128 and ASP129.

2.2 Discussion

The simulations presented in this chapter have been motivated by the puzzling experimental finding that lysozyme, although carrying a higher net positive charge, adsorbs in much lower amount than chymotrypsin on negatively charged amorphous SiO_2 surfaces (see Fig. 1.6). On the other hand, the amount of adsorbed protein does increase with increasing pH (that is, with increasing surface charge density), a fact that can be rationalized regarding classical DLVO forces, and is indeed confirmed by static single-point calculations of protein-surface interaction profiles in implicit solvent (Fig. 2.1). These simulations, however, clearly indicate stronger attractive interactions for lysozyme at all investigated pH values.

The first source of discrepancy between simulations and experiments is the consideration of single proteins and thus the neglect of protein-protein interactions. This assumption is justified in solution, where the employed experimental conditions (see [34]) correspond to proteins with characteristic sizes of about 4 and 5 nm separated by average distances of 23 and 27 nm for lysozyme and chymotrypsin, respectively, and silica particles with a diameter of 180 nm separated by about 700 nm. However, the formation of an adsorbed monolayer, which is experimentally observed for chymotrypsin at high pH (Fig. 1.6 b)), is necessarily associated with a tight packing of proteins and may be thus disadvantageous for proteins carrying a strong net charge [86].

In fact, if we calculate profiles of DLVO protein-protein interactions using the same implicit-solvent model as for the calculations of protein-surface interactions, we can clearly see a difference in the interaction between lysozyme and chymotrypsin. The lysozyme molecules repel each other irrespectively of their mutual orientation, whereas chymotrypsin molecules attract each other and can be brought into tight contact (up to 15 Å) for a large number of orientations (Fig. 2.11). This result is a consequence of the uniformly distributed positive surface potential of lysozyme, as opposed to an apparent charge separation and large dipole moment in the case of chymotrypsin (Fig. 2.3). Therefore, we can safely assume that the formation of a tightly adsorbed layer is favorable for chymotrypsin but not for lysozyme, which is consistent with our experimental observation.

Our implicit solvent simulations are very useful to predict the most attractive orientation of the adsorbing proteins (Fig. 2.3 and 2.4). For chymotrypsin the orientation is

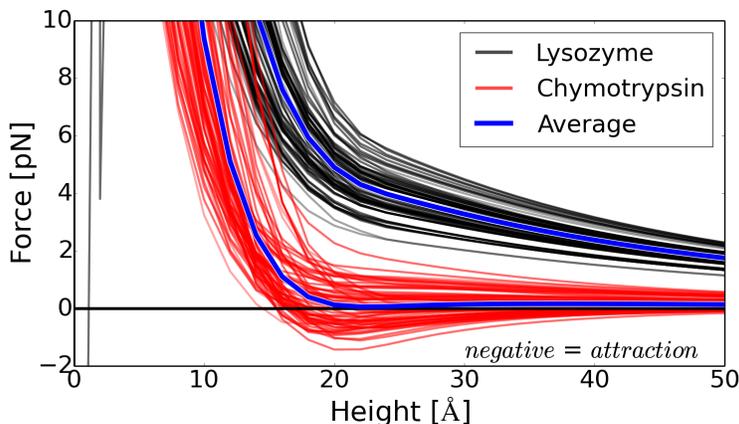


Figure 2.11: *Protein-Protein self-interaction profiles of lysozyme and chymotrypsin. The proteins are rotated around their axes in 90° steps, resulting in a total of 256 mutual orientations. The calculations are performed in a Generalized-Born implicit-solvent model with a salinity of 3 mM monovalent ions and at pH 8.0.*

mainly driven by its large dipole moment, leading to preferential adsorption via its positive surface potential region, exposing to the solvent the negative surface potential region. This orientation seems to be consistent with the experimental observations by Vasina et al., [107] which revealed a negative ζ -potential for a densely packed layer of chymotrypsin adsorbed on a mica surface. In the case of lysozyme, our predicted “most attractive” orientation is in almost perfect agreement with the previous prediction by Kubiak-Ossowska et al. [111] (Fig. 2.4). However, the orientational preference of lysozyme is much less pronounced, due to its more homogeneously distributed surface potential.

A second important point that needs to be considered when discussing adsorption processes is the active role played by the solvent and of its structuring at solid/liquid interfaces [55, 85, 115–119]. When proteins approach surfaces to distances smaller than about 10 Å, implicit-solvent predictions lose their significance, while the molecular nature of the water (and the dissolved ions) needs to be explicitly considered. The presented Molecular Dynamics simulations, in fact, give a more complex and comprehensive picture of the approach and adsorption processes. The predicted adsorption orientations after 50 ns of MD simulations follow the general trend of the static calculations. In particular, it is interesting to observe initial repulsion and rotation of the proteins (especially of chymotrypsin) when starting in the “most attractive” orientation (Fig. 2.6, simulation I) at a distance of 10 Å from the surface. This behavior suggests that attraction/repulsion

patterns between individual amino acids and specific surface sites (leading to an overall rotational moment), as well as solvent restructuring in surface proximity (possibly leading to repulsion in spite of DLVO attraction) largely influence the specific adsorption pathway. It is striking that adsorption takes place faster when the molecule is 30-40 Å away from the surface (Fig. 2.6, simulations II and III) than when it is in closer contact (simulation I). This observation probably suggests the necessity of hovering over the surface until the most favorable pattern of adsorption sites underneath the protein is found and the barrier exerted by the interfacial hydration layer can be overcome [113].

A favorable matching between binding motifs (residue combinations) and adsorption sites, mediated by the vicinal solvent structuring and charge distribution, is a fundamental factor governing the final adsorption configurations of both enzymes (Figs 2.8, 2.9 and 2.10). This matching leads to a preferential side-on adsorption of lysozyme, but without clearly favored orientations of the active site. Our results do not confirm the straight end-on approach and initial adsorption of lysozyme on SiO₂ proposed by Wei et al. [113], but are in agreement with the side-on adsorption orientations suggested by other authors [110–112, 120]. In particular, the MD simulations by Kubiak-Ossowska et al. [111, 120] reveal that ARG128, ARG125, ARG5 and LYS1 are strongly adsorbing amino acids, which we can fully confirm. Once again, the binding motifs appear more evenly distributed for lysozyme than for chymotrypsin, where both frequent and tight binders are located mostly around the two α -helical structure elements on one protein side. This distribution is a consequence of the distinct localization of residues such as lysine and arginine, which both cause a strong dipole moment and thus a clear adsorption orientation preference, and also offer ideal anchoring points to the silica surface.

The evidence of a cooperative surface binding by several spatially close residues, while already observed to be a characteristic feature of protein adsorption at hydrophilic surfaces, [55, 121, 122] is an important outcome of our study. We can postulate that the simultaneous binding via many favorably located residues can stabilize the adsorption of chymotrypsin to a larger extent than lysozyme, given the aforementioned abundant clustering of binding motifs around the α -helices. This could be an additional factor that helps explain why chymotrypsin is able to form a densely packed adsorbed monolayer, while lysozyme is not, and calls for future quantitative investigations of the adhesion forces and free energies in combined experimental and theoretical studies.

As a final remark, we stress that our standard MD simulations are too short to observe

any significant structural changes after adsorption of either protein (Fig. 2.5). Nevertheless, it is interesting to note that adsorption via the α -helices is consistent with a loss of helical content after chymotrypsin binding to silica nanoparticles observed in Circular Dichroism experiments [24]. Moreover, the overall stability of the tertiary structure of SiO₂-adsorbed chymotrypsin has been demonstrated using highly sensitive differential scanning calorimetry [98]. Also, lysozyme has been observed to lose secondary structure ordering partially, once interacting with silica [95–97]. In Chapter 5 we have used and extended advanced MD techniques that enable unbiased predictions of polypeptide structural changes after interactions with heterogeneous interfaces [123] to the case of protein adsorption to address these issues.

Chymotrypsin Binding Motifs on Silica and Titania

The adsorption motifs and details of the adsorption process of two proteins, lysozyme and chymotrypsin, on silica have been investigated in depth in Chapter 2. In this chapter we focus on the influence of the surface on the adsorption process. Therefore, we use and extend the methods presented in Chapter 2 to compare the chymotrypsin adsorption process on both silica (SiO_2) and titania (TiO_2). Chymotrypsin adsorption on SiO_2 can easily reach a monolayer concentration ([24], Fig. 1.6). The irreversible adsorption with respect to rinsing can cause surface-induced conformational changes which may reduce the protein activity [24, 98, 99]. Activity investigations of chymotrypsin on TiO_2 by Derr et al. [39] have shown a reduction of the protein activity in the adsorbed state. Gailite et al. [124] immobilized trypsin on cluster-assembled TiO_2 films and measured the enzyme activity by fluorescent labeling in a microarray-based technique. It was found that the trypsin activity in the adsorbed state drops for an increased adsorption uptake. Gailite et al. suggested that a reduced accessibility of the enzyme active site causes this reduced activity. This possible steric effect depends strongly on the protein concentration and orientation on the surface. The protein orientation and density in a single and double layer configuration on both oxides are investigated in this chapter.

The experimentally measured adsorption amount of chymotrypsin on SiO_2 (Fig. 1.6 [34]) and on TiO_2 [39] has indicated a strong dependency on the pH value. This trend on both surfaces can be explained by interactions based on the DLVO theory as shown in the last chapter and extended in the following. However, the direct experimental comparison

of chymotrypsin adsorption on SiO_2 and TiO_2 nanoparticles has measured an increased amount of adsorbed chymotrypsin on TiO_2 in comparison to SiO_2 (Fig. 1.7, [40,65]). Furthermore, the experiments have shown an adsorption beyond a monolayer concentration on the surface. This effect can not be solely explained by DLVO long-range interactions. Although both surfaces are found to be very hydrophilic [55], they vary by their isoelectric point (IEP) and reactive surface groups. SiO_2 has acidic surface properties while TiO_2 shows an amphoteric character, as already described in Section 1.2. Those surface properties result in a very different surface charge distribution pattern which may result in a different structuring of the water over the surface as also observed by Schneider et al. [55]. An influence of these properties on the final adsorption process can be expected and is investigated in detail in the following.

3.1 Results

Some of the methods used in Chapter 2 are adapted and extended in this chapter. The technical details of the presented MD simulations are listed in [34,65] and in the Appendix Chapter 7.

3.1.1 Long-range Interaction Force of Chymotrypsin over SiO_2 and TiO_2

The long-range interactions of the single protein over both surfaces are investigated to analyse a possible difference in attraction. Static calculations are used to evaluate the force-distance profiles and predict the most favorable adsorption orientation at different pH values (Fig. 3.1), as also presented in Chapter 2.

The force profiles are predominantly defined by the interaction between the strong protein dipole moment and the charge distribution at the surface. For the case of SiO_2 , an average attraction interaction is observed for all pH values. With increasing pH the surface is more and more negatively charged which causes a distinct protein orientation dominated by its dipole moment. On TiO_2 , on the other hand, repulsion is predicted at pH values lower than the surface IEP (6.0) [46], whereas attraction takes place only at pH 7.4 and 8.0. As a consequence of the conversion of the surface charge at the IEP, the protein orientation turns around between pH 5.0 to 7.4 (Fig. 3.1, (*left*)). At pH

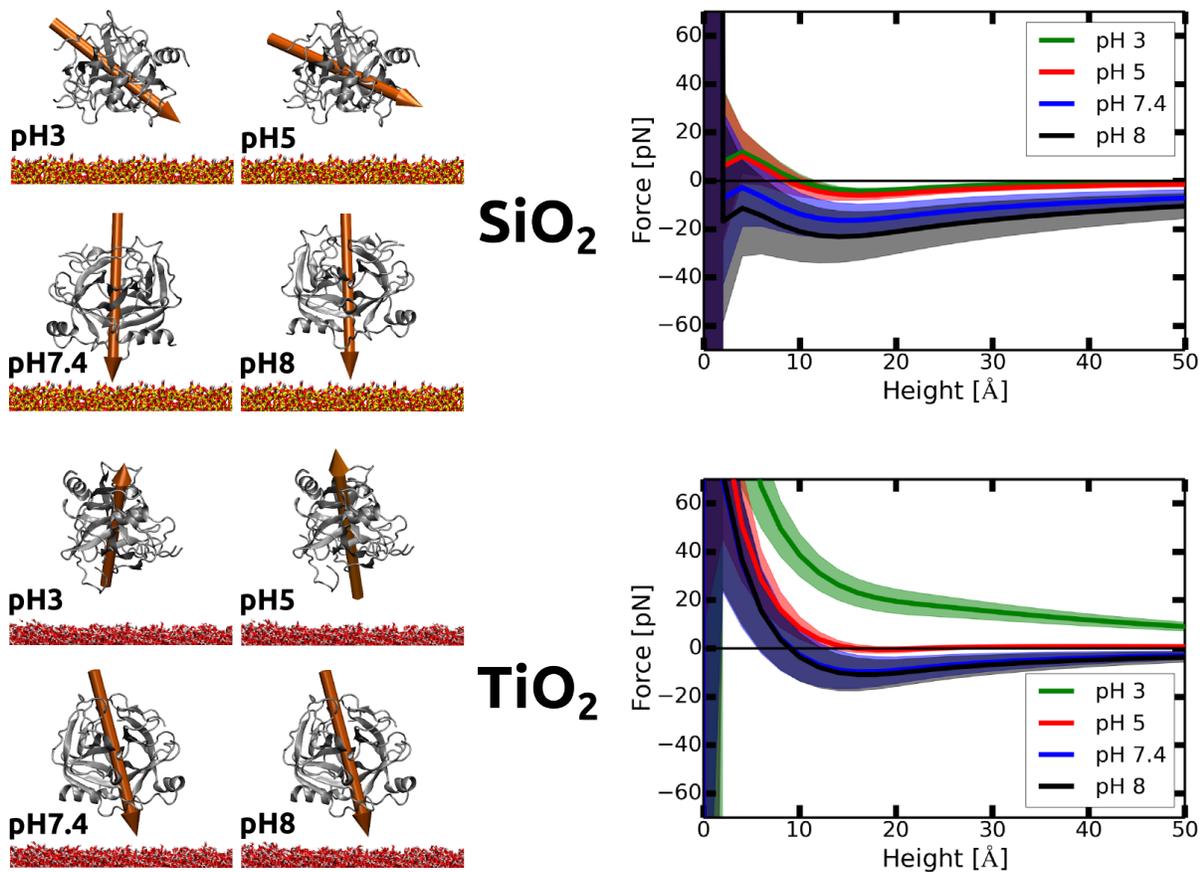


Figure 3.1: (left) Predicted most attractive orientations of chymotrypsin over SiO_2 and TiO_2 at different pH values. The molecular dipole moment is shown as an orange arrow. (right) Force-height profiles calculated in an implicit solvent for 36 different orientations at each pH value. The average of all orientations is drawn with the bold colored lines, and the standard deviation is represented by semitransparent colored regions. Negative values define attractive forces. The height is defined as the distance between the highest surface atom and the lowest protein atom in the surface normal direction.

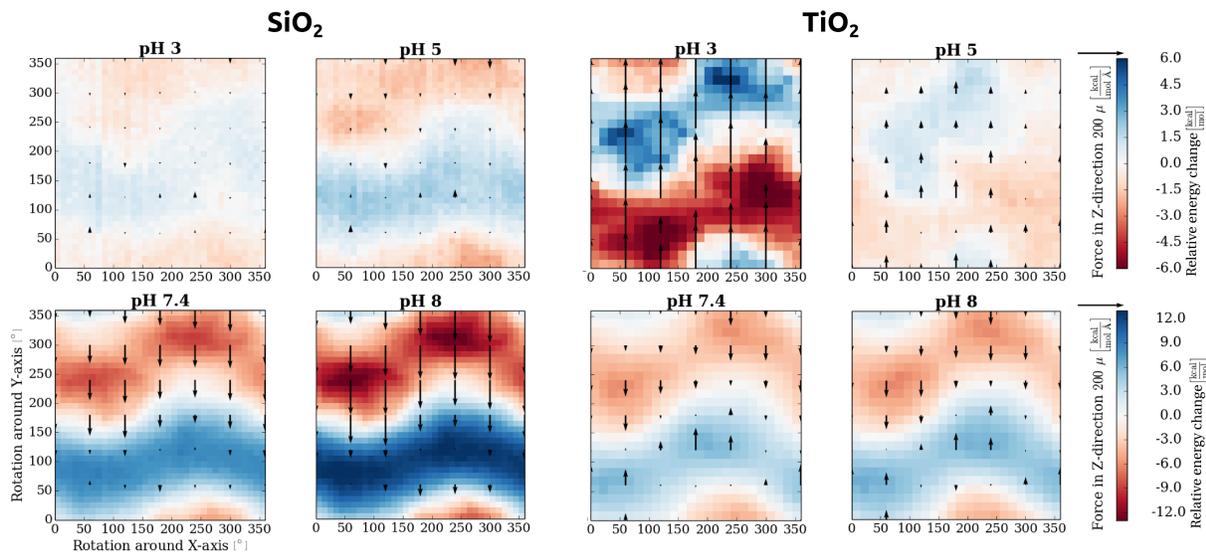


Figure 3.2: Calculated relative energy changes associated with a rotation of chymotrypsin at a height of 1 nm over the surfaces for SiO_2 (left) and TiO_2 (right), each shown for four different pH values. The superimposed arrows represent the force values, downwards for attractive and upwards for repulsive forces.

8, the two α -helix motifs of the protein face towards the surface and the enzymatically active site points sideways with respect to the surface normal. The maximum attractive forces are experienced by the protein at pH 8 at the height of about 1.5 nm over each respective surface. It has to be stressed that these calculations take into account the DLVO interactions in a continuum solvent approximation only and that further protein approaches to the surface below 1.0 nm are strongly mediated by the structuring of the water molecules in close surface proximity ([34, 51, 55] and Section 1.3). The calculated long-range DLVO attractions increases for high pH values, which is also observed in the adsorption experiments shown in Fig. 1.6 [34, 65]. However, the calculated DLVO interactions show a stronger attraction of the protein on SiO_2 compared to TiO_2 . This trend does not correlate with the experimental observation presented in Fig. 1.7. The protein concentration at the surfaces could be influenced by the protein orientation which is investigated in the following.

The calculated interaction profiles over TiO_2 shown in Fig. 3.1 are less deep and more narrow compared to SiO_2 . One can thus expect an easier reorientation of chymotrypsin over TiO_2 due to a less distinct force dependence of the orientation. This is indeed confirmed by analyzing the relative energy changes upon rotation of the proteins around two independent axes at the constant height of 1.0 nm (Fig. 3.2).

For all considered pH values a belt-like feature in the diagrams can be observed (Fig. 3.2). All protein orientations within this region belong to a rotation of the protein around its dipole moment perpendicular to the surface as discussed in detail in the last Chapter 2. This interaction pattern appears less distinct for smaller pH values. The small black arrows shown in the energy diagrams indicate the attractive or repulsive DLVO forces. The differences in the relative energy upon molecular reorientation are about 24 kcal mol⁻¹ over SiO₂ and 12 kcal mol⁻¹ over TiO₂.

3.1.2 Interaction Forces for Chymotrypsin Multilayers

While in Section 3.1.1 the case of a single protein approaching the surface has been investigated, in this section the protein-protein interactions are studied thus taking a preadsorbed protein monolayer into account. All further calculations are performed at a pH of 8 and in case of the implicit solvent calculations at a monovalent salt concentration of 3 mM. A first calculation is performed evaluating the relative energy changes associated with the reorientation of one protein within a preadsorbed protein monolayer with a density of 0.034 molecules nm⁻² (Fig. 3.3 (a)). This concentration is in agreement with a rectangular monolayer density calculated and shown with the measured concentration in Fig. 1.7.

The height of the protein marked in red in Fig. 3.3 (a) is adjusted to 0.3 nm. The interaction energy is calculated upon rotation around the x and y axes. This energy change is shown in Fig. 3.3 (b) for the SiO₂ surface, a visually almost identical map is obtained for TiO₂ as well. Starting from the most attractive adsorption orientation predicted for single proteins (at $x = y = 0^\circ$), the rotation in any direction is associated with an energy increase. An additional energy maximum is obtained for the molecular dipole pointing in the opposite direction ($x = 180^\circ, y = 0^\circ$; or $x = 0^\circ, y = 180^\circ$). The maximum is 30 kcal mol⁻¹ higher than the minimum in the case of SiO₂ and 20 kcal mol⁻¹ higher than the minimum in the case of TiO₂. This suggests that all proteins in the monolayer are oriented in the same direction with the dipole moment facing towards the negatively charged surface. In this orientation the protein α -helices are pointing toward the surface.

In a further simulation the force-distance profiles associated with the approach of a single chymotrypsin molecule towards an already present protein monolayer are calculated,

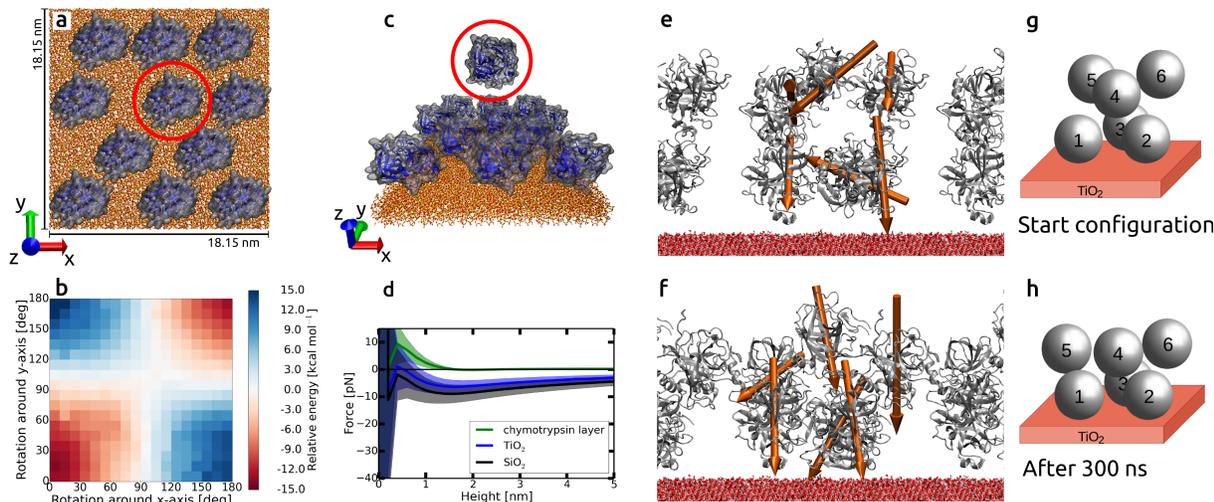


Figure 3.3: Multilayer formations of chymotrypsin over SiO_2 and TiO_2 . (a) Atomistic representation of a protein monolayer at the density of 0.034 molecules nm^{-2} . The red circle marks a protein which is rotated along two axes (x and y) creating the energy map in (b). The force profiles calculated for another protein (c) interacting with the pure protein monolayer and with the underlying two oxide surfaces shown in (d) as green, blue and gray curves respectively. A height of zero is defined when the highest monolayer atom and the lowest atom of the additional protein are at the same height in the direction of the surface normal. The start and final configurations of a protein double layer on TiO_2 (0.077 protein molecules nm^{-2}) in an explicit-solvent MD simulation lasting 300 ns are shown in (e) and (f). The same simulation is presented in the schemes in (g) and (h).

as shown in Fig. 3.3 (c). An additional protein approaching a chymotrypsin monolayer without an underlying oxide surface is mostly repelled by repulsive DLVO forces (Fig. 3.3 (d)). This suggests no favorable driving forces for the spontaneous formation of protein agglomerates in solution. However, if a protein approaches the chymotrypsin monolayer with an underlying SiO_2 or TiO_2 surface, attractive DLVO forces can be observed. The strong protein-surface interactions at the used salt concentration of 3 mM are not completely shielded by the presence of the adsorbed monolayer. The calculated attractive forces are slightly stronger in the case of SiO_2 compared to TiO_2 . An orientation analysis of the protein over the chymotrypsin monolayer reveals an identical preferred protein orientation as observed in the single molecule investigations shown in Fig. 3.1. Therefore it can be suggested that the preferred chymotrypsin protein orientation is always identical with the dipole moment facing towards the surface, no matter whether the protein is isolated over the surface or embedded in a single or a double layer.

To confirm this hypothesis based on the DLVO interactions further simulations of the double layer configuration in an explicit solvent at pH 8 are performed. In particular, a MD simulation starting from a double layer configuration with randomly oriented proteins is carried out (Fig. 3.3 (e) and (g)). After 300 ns of simulation (Fig. 3.3 (f) and (h)) two major trends are observed. First, the proteins collectively move closer to the surface and form a more compact layer on both surfaces. Second, the individual proteins rotate toward the most-attractive orientation presented in Fig. 3.1 (*left*). In some cases, the rotation is hindered by tight protein-protein contacts, which remain stable for the whole simulation time.

In summary, the simulations presented in this section indicate the presence of both long- and short-range interactions. On both surfaces at pH 8 the proteins are attracted to the surfaces in a single or double layer configuration, which is in agreement with the experiment (Fig. 1.7). However, these simulations are still not able to explain differences in the adsorption modes on SiO_2 and TiO_2 resulting in a larger amount of proteins adsorbed on TiO_2 than on SiO_2 . To further explore this issue, explicit solvent MD simulations of adsorbed proteins are presented in the following section.

3.1.3 Explicit-solvent MD Simulations of Chymotrypsin Adsorption on TiO_2 and SiO_2

The atomistic details of the adsorption process are investigated by MD simulations in an explicit solvent at pH 8 following the MD simulation methods presented in Section 2.1.2. On each surface the simulations are started in six different arbitrary orientations at a height of 1 nm over the surface (Fig. 3.4). During the simulation, the protein spontaneously adsorbs on the surface after about 5 ns and establishes a stable adsorption configuration. In this process it performs a rotation towards the most favorable orientation which has also been predicted by the implicit solvent calculations (Fig. 3.4 (a) - (b), Fig. 3.1). The simulations are stopped after 50 ns.

The evolution of the protein-surface contacts during two of the twelve simulations is shown in Fig. 3.4. A protein-surface contact is defined as the distance between one atom of the surface and one atom of the protein of less than 0.24 nm. For both surfaces, the most frequent binding motifs of the protein are the positively charged amino acid lysine as well as the polar amino acid threonine (Fig. 3.4). The most common surface sites involved in the contacts are terminal OH groups on SiO_2 as well as on TiO_2 . While

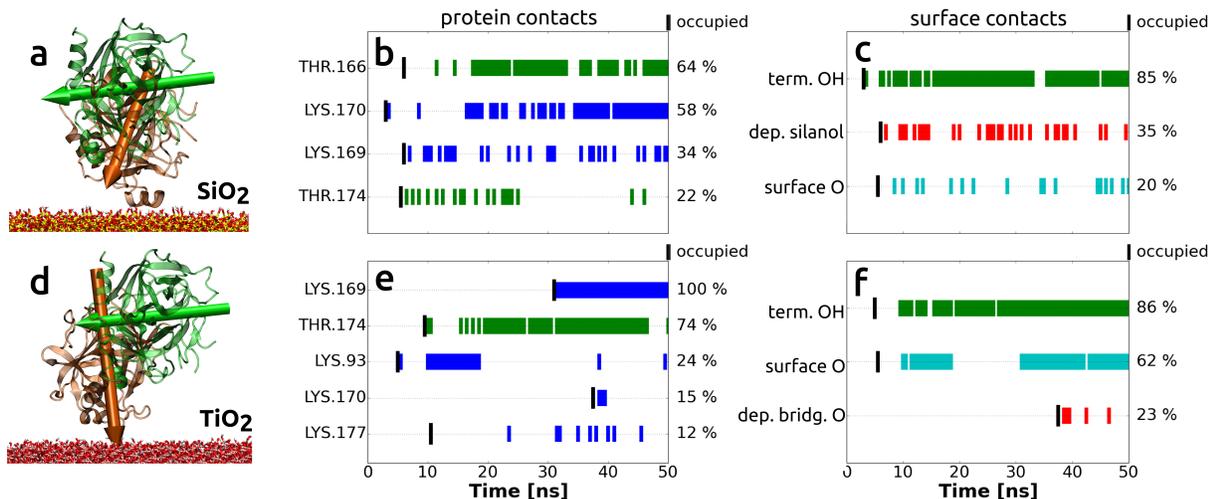


Figure 3.4: Initial (green arrow) and final (orange arrow) chymotrypsin orientations upon adsorption on SiO₂ (a) and TiO₂ (d) obtained in MD simulations in explicit solvent. The arrows represent the dipole moments. The contact analysis of the protein residues on SiO₂ and TiO₂ is shown in (b) and (e). The color code represents positively charged amino acids (blue) and polar amino acids (green). The contact analysis for the surface residues of SiO₂ and TiO₂ is presented in (c) and (f). The percentages reported at the right side of each subfigure indicate the occupancies over time after the first contact (marked with a bold black bar). Groups forming contacts for less than 10% of the time are not shown.

deprotonated OH groups play a crucial role in protein adsorption on SiO₂ they are less frequently involved on TiO₂. Further surface contacts are formed by bridging O sites, which are more frequently formed on TiO₂ than on SiO₂ (Fig. 3.4).

Differences in the experimentally measured adsorption behavior of the two proteins may also be related to their mobility on the surface after adsorption. To investigate this issue, we perform steered MD simulations (SMD) in explicit solvent. In these simulations the adsorbed proteins are pulled in parallel to the surfaces using a harmonic constraint applied to its center of mass (Fig. 3.5). The pulling velocity is 0.5 m s⁻¹ and the constraint force-constant is 1000 kJ mol⁻¹ nm⁻². Two perpendicular pulling directions ($\langle 100 \rangle$ and $\langle 010 \rangle$) are chosen to decrease direction-dependent effects. The proteins are free to roll and/or slide over the surface. A contact analysis of these simulations is shown in Fig. 3.6.

The protein-surface contacts formed during these pulling simulations correspond to a large extent to those contacts reported in Fig. 3.4. Although, more different types of amino acids are involved in surface binding, including hydrophobic ones (Fig. 3.6). More interesting, however, is the analysis of the rolling/sliding behavior of the protein over the

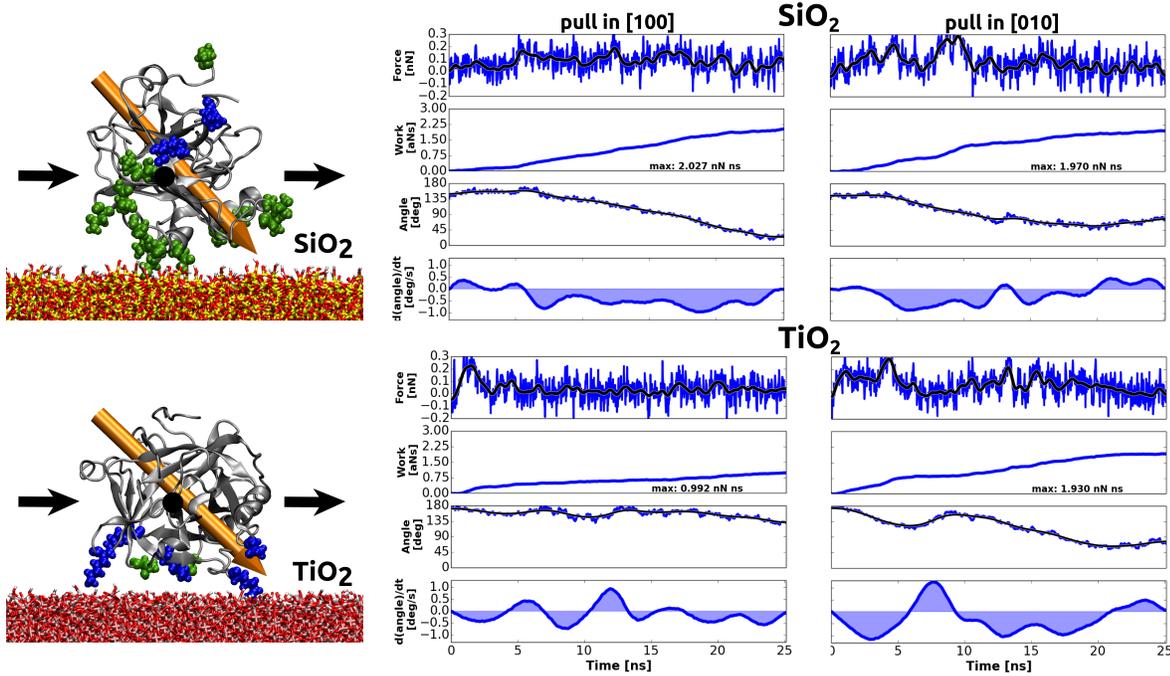


Figure 3.5: (left) Initial configurations of a steered MD simulation of chymotrypsin pulled over the SiO₂ and TiO₂ surfaces. (right) Time evolutions along the corresponding trajectories of the constrained force, the cumulative work, the angle of the dipole moment to the surface normal and its time derivative (from top to bottom in each panel) for SiO₂ (top) and TiO₂ (bottom). The protein is pulled in each case along the $\langle 100 \rangle$ and $\langle 010 \rangle$ directions of the simulation box.

two surfaces, as reported in Fig. 3.5 and visible in the videos of this simulation provided online in the supplementary information of reference [65].

The force required to pull the whole protein over the surface and the associated cumulative work is slightly smaller on the TiO₂ surface. Especially the protein pulled in the $\langle 100 \rangle$ direction on TiO₂ temporarily desorbs between 5 and 15 ns, resulting in negligibly small pulling forces. A striking difference is the kind of protein movement over the surface. They tend to roll over SiO₂ while they slide over TiO₂. This is evident from the evolution of the angle between the dipole moment and the surface normal and its derivative (Fig. 3.5). For SiO₂, the angle varies continuously towards smaller values and the angle derivative is mostly negative, indicating continuous rolling. This is noteworthy if one takes into account that the rotation of the protein is associated with a considerable increase of the potential energy (Fig. 3.3). For TiO₂, the derivative of the angle tends to oscillate around zero which indicates a sliding rather than a rolling movement.

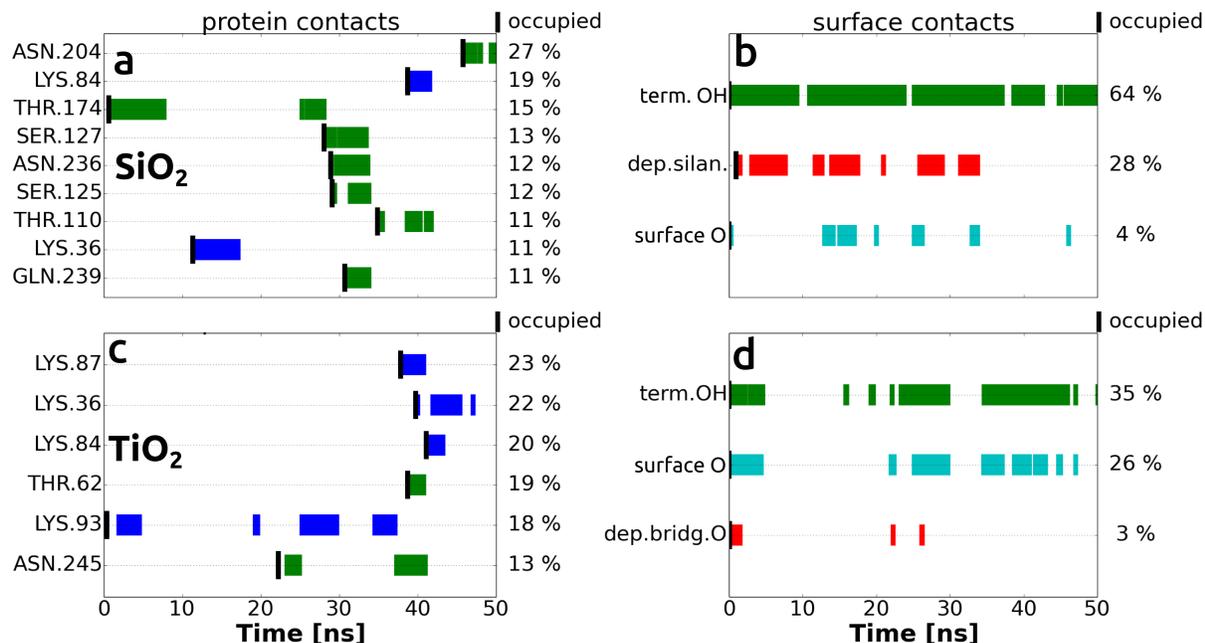


Figure 3.6: Contact analysis of the simulation presented in Fig. 3.5. The analysis of the contacts of the protein residues on SiO₂ (a) and TiO₂ (c). The contact analysis for the surface groups of SiO₂ (b) and TiO₂ (d). The percentages reported at the right edge of each subfigure indicate the occupancy over the time after the first contact is formed. This is marked with a bold black bar. All groups forming a contact for less than 3% of the time after the first contact are not shown.

The differences in the adsorption modes of chymotrypsin on SiO₂ and TiO₂ are further highlighted by the calculation of the radial distribution function (RDF) between atoms of the surface, the water solvent and the protein (Fig. 3.7, Appendix 7.6). Water molecules are known to form shell-like structures around hydrophilic surface terminal groups [34]. The RDF of water around all surface atoms shows a less distinct structuring on SiO₂ (Fig. 3.7 top left) compared to the very evident structuring in two hydration layers on TiO₂ (Fig. 3.7 top right). The RDF shows that predominately hydrophilic residues are in contact with both surfaces whereas only some hydrophobic residues interact with the SiO₂ surface. On TiO₂ the hydrophobic residues are not in close contact with the surface at all. Interestingly a small peak of the protein-surface RDF at 1.5 Å on SiO₂ occurs before the first peak of the water-surface RDF at 1.8 Å. This indicates that some protein residues can deeply penetrate into the first hydration layer and tightly bind to surface sites. Specific surface interactions of positively charged amino acids with deprotonated OH groups are found mainly on the SiO₂ surface and only to a minor extent on TiO₂ (Fig. 3.7

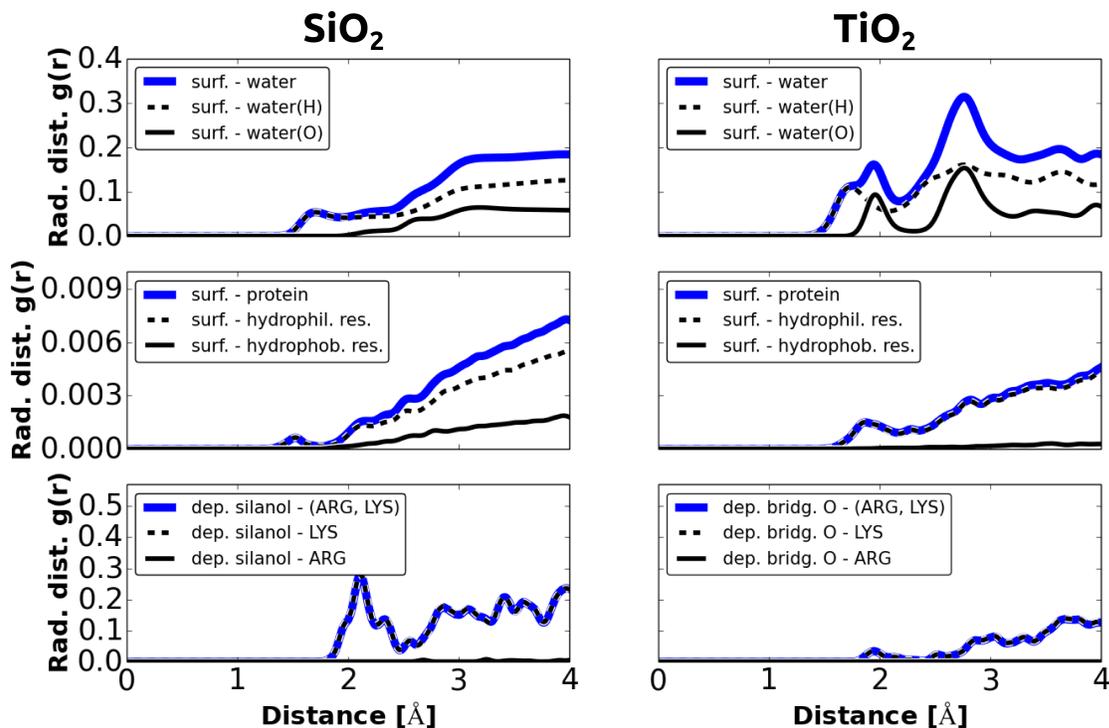


Figure 3.7: Radial distribution function (RDF) of chymotrypsin adsorbed and pulled over SiO_2 (left) and TiO_2 (right): surface-water RDF (top); surface-protein RDF (middle), separating the contributions of hydrophilic and hydrophobic residues. RDF between positively charged amino acids and negatively charged surface sites of deprotonated OH groups (bottom).

bottom). This specific interaction is dominated by the lysine residues whereas arginine plays a minor role which can be simply explained by the reduced number of arginines exposed on the chymotrypsin surface. These binding characteristics of the amino acids are in agreement with the observations in the contacts analysis shown in Fig. 3.4.

3.2 Discussion

Motivated by the need to understand the mechanism of the adsorption of chymotrypsin on ceramics, we have performed a comparative study of the adsorption features of chymotrypsin on SiO_2 and TiO_2 with all-atom MD simulations. The analysis has revealed some common features, such as the same preferred adsorption orientation on the two materials at neutral and basic pH driven by the interaction of the negatively charged surfaces with the strong protein dipole moment (Fig. 3.1). An attraction of the proteins to both surfaces at basic pH is observed for a single and double layer configuration with a rather compact double layer formation revealed by explicit solvent simulations shown in Fig. 3.3. A reduced activity of the adsorbed proteins by steric effects as supposed by Gailite et al. [124] and measured by Derr et al. [39,125], however, can not be confirmed. The active sites of the proteins point parallel to the surface, thus allowing the access of potential ligand molecules. In the activity measurements of Derr et al., p-Nitrophenyl acetate (pNPA) was used as a ligand for chymotrypsin, which contains only 20 atoms and has a size of about 0.1 nm. Therefore, it is probably still small enough to reach the active sites of the densely packed chymotrypsin layers. No significant differences of the described properties could be observed when comparing SiO_2 and TiO_2 . In addition, the measured Langmuir adsorption isotherms reveal a nearly equal free energy of adsorption on both surfaces (Fig. 1.7, [65]). However, some differences have also been identified in the simulations, which we believe to be responsible for the experimentally observed larger amount of adsorbed proteins found on TiO_2 than on SiO_2 .

Anchoring to the SiO_2 surface takes place mostly via positively charged amino acids that interact with both neutral and deprotonated silanol groups. However, also other neutral polar and even hydrophobic residues contribute to frequent and tight surface-protein contacts (Fig. 3.4, 3.6 and 3.7). Anchoring to the TiO_2 surface, which presents a smaller density of deprotonated sites at the pH values relevant to adsorption, is rather dominated by neutral terminal OH and bridging O sites, although involving almost exclusively positively charged amino acids. We believe that the distinct structuring of the surface hydration layers over TiO_2 in which the amino acids are embedded (Fig. 3.7) effectively compensates for the absence of direct ionic interactions between protein and surface. Therefore, the overall free energy of binding to either surface is nearly the same. Different, however, is the ability of the protein to diffuse over the surface once tightly

adsorbed. Our steered MD simulations (Fig. 3.5 and 3.6) reveal that chymotrypsin can easily slide over the TiO_2 surface and even temporarily desorb and adsorb again at another site. On the contrary, no temporary desorption is observed from the SiO_2 surface which is consistent with the larger attraction forces predicted by our static calculations (Fig. 3.1). In this case, rupture of the close surface-protein contacts only takes place if new contacts are concomitantly formed, resulting in a predominant rolling behavior. A change in the dipole orientation due to rolling is associated with a considerable energy loss (Fig. 3.2 and 3.3), so that diffusion over the SiO_2 surface is expected to be more hindered than diffusion over the TiO_2 surface.

Considering these arguments, it is possible that a larger maximum adsorption amount is reached on TiO_2 (Fig. 1.7). While our simulations clearly show that at least a double layer is energetically stable on either surface (Fig. 3.3), its exact formation relies on diffusion of preadsorbed proteins to accommodate new ones from the solution. This process is probably easier on TiO_2 because of the observed more favorable diffusion mobility.

Surface Contact Stability

Once adsorbed, the proteins form contacts with the surface as investigated in depth in the previous chapters. Furthermore, different mobilities at the surface have been identified, including a frequent breaking and forming of new surface contacts (Chapter 3). The stability of these surface contacts and the forces necessary to detach the proteins from the surface are investigated in this chapter by pulling lysozyme away from an amorphous SiO_2 surface. This non-equilibrium pulling process includes not only the detachment of the protein from the surface but also a possible denaturation of the protein caused by the pulling. Thus we are investigating the stability of the surface contacts as well as the internal protein stability, which is among others provided by disulfide bonds (S-S bonds). So far, it is not known whether the pulling of the protein from the surface cause a breaking of the S-S bonds and subsequent denaturation or whether the protein detaches in its native state from the surface. This central question is studied in this chapter.

The mechanism of how proteins are stabilized by disulfide bonds (S-S bonds) is relatively well understood [126, 127]. Covalent bonds between two CYS amino acids reduce the protein entropy and lead to vicinal interactions, thus stabilizing the folded state enthalpically. This can be used in conformational studies of proteins by modification of the S-S bonds [126]. Furthermore, by “S-S bond engineering” the bonds can be artificially introduced into proteins to increase their stability as performed by various authors for T4 lysozyme [128–130]. On the other hand, if the S-S bonds are reduced in a wild-type protein, significant conformational changes can occur. This effect is well investigated in the proteins bovine and human serum albumin (BSA, HSA). They unfold by the action of the reducing agent dithiothreitol (DTT) which selectively cleaves the S-S bonds [131].

An MD study on HSA revealed changes of its secondary structure and hydrogen bonds in the absence of the native S-S bonds [132]. The reduction of some of the 17 S-S bonds caused direct conformational changes whereas the absences of others was compensated by neighboring hydrogen bonds. Another MD simulation by Ning et al. [133] suggests that the missing of the S-S bond CYS 179 - CYS 214 in the prion protein leads to denaturation as well, which could cause prion disease.

The free energy of the reduction of S-S bonds in lysozyme, including the influence of the reducing agent tris(2-carboxyethyl)phosphine (TCEP), has been investigated by density functional theory (DFT) [134]. The calculated energy is strongly dependent on the steric environment of the S-S bonds and on how good it is accessible for the reducing agent. It was found that one of the four S-S bonds of lysozyme is more exposed to the solvent than the others, which are more embedded in the protein. This bond connects the CYS 6 at the N-terminus with the CYS 127 at the C-terminus. Due to the enhanced accessibility and polarizability, this S-S bond has a significant smaller activation energy than the other three bonds. The accessibility can further be influenced by local unfolding and opening events [126, 135]. The same S-S bond at CYS 6-127 in lysozyme was investigated by differential scanning calorimetry [136]. In this study, the wild-type lysozyme was compared to a modified lysozyme in which the CYS 6-127 S-S bond had been reduced and carboxymethylated. Nuclear magnetic resonance spectroscopy (NMR) showed a high conformational similarity of the mutant to the native protein. Still, the thermal unfolding of both proteins revealed one distinct peak of the excess heat capacity at significantly different temperatures. Therefore, the lysozyme with the carboxymethylated S-S bond denaturates at lower temperatures than the wild type. This higher instability of the modified lysozyme suggests a significant contribution of the CYS 6-127 S-S bond to the lysozyme stability.

Single-molecule AFM is particularly suitable to study this protein unfolding as a function of its S-S bonds experimentally. Due to the atomic-scale sensitivity of the AFM, a very specific molecular interaction of ideally only one molecule of interest can be investigated. This technique has successfully been used in the past to study the unfolding of proteins [137, 138] and protein surface interactions [94, 139–141]. Grandbois [142] studied the rupture of covalent bonds on different surface chemistries using an AFM and identified a force of a covalent bond between Au-S of 1.4 nN. Another AFM study on the S-S bonds of proteins by Wiita et al. [143] showed very clearly that a force of 100 pN for a

time of five seconds as well as 165 pN for 400 ms does not break the S-S bonds. Only in the presence of thioredoxin, which catalyzes the detachment of the S-S bonds, a constant force of 100 pN is enough to break the bonds within seconds.

Wei et al. [141] have pulled lysozyme away from a SiO₂ surface using an AFM thus observing four characteristic peaks in the force-distance curves. The authors suspected that those peaks might be related to the breaking of the four S-S bonds of lysozyme. This interesting observation has motivated further AFM experiments which included the attachment of lysozyme to a polyethylene glycol (PEG) linker which is covalently bonded to a silanized AFM probe. The PEG linker can be chosen to be relatively long, e.g. 45 nm [144], to separate nonspecific cantilever-surface-interactions from the specific interactions [145]. However, in the case of long PEG linkers, an unpredictable complex network can be formed. In addition to that, the amino silane surface functionalization can form complex networks of multilayers as well [146]. These networks can result in a wide variety of force distance curves making the interpretation of specific surface interactions very challenging. Moreover, the molecule of interest is not always in contact with the target, reducing the number of specific force distance curves even more [147–149].

A possibility to interpret the experimentally measured force-distance curves is the combination with SMD simulations, which have been used in a variety of these problems [150–155]. Thereby an energy barrier along a distance direction can be sampled, which is hard to overcome in a classical MD simulation. Kubiak-Ossowska et al. [150] have performed SMD simulation of lysozyme being pulled away from a negatively charged surface. In all 14 simulations, the protein preserves its native structure and desorbs without an obvious denaturation. This supports the hypothesis that the S-S bonds are more stable than the surface contacts during the pulling process. Furthermore, it has been observed that the protein detaches stepwise from the surface. The steps are caused by the desorption of ARG residues from the first and second water layers with rupture forces of about 400 - 500 pN. The integration over the force-distance curves results in an activation energy of 1.6 eV. In another simulation of the same study, an activation energy of 2.3 eV has been measured if the protein is not pulled on the ARG 128 but on four different residues.

In this chapter the stability of the surface contacts in comparison to the stability of the S-S bonds during the pulling process is further investigated by comparing SMD simulations with AFM experiments. Observations of patterns in measured force-distance curves

of AFM experiments indicated the denaturation of S-S bonds [141] while SMD simulations described a detachment of the protein without denaturation during the pulling process [150]. The direct comparison between SMD and AFM experiments can be challenging as the observed time scales differ by roughly six orders of magnitude. Furthermore, the pulling velocities vary by three orders of magnitude. Stokes friction as well as thermally activated barrier crossings are a function of the pulling velocity and therefore influence the measured forces [156]. The friction forces vanish for velocities below 10 m/s [155, 157–159] whereas unbinding forces have an additional logarithmic dependency on the pulling velocities [156, 157, 160, 161]. Therefore, activation energies like adsorption or binding energies depend on the pulling velocities used. As a consequence, the direct comparison between experiment and simulation has to be handled with care [157, 162].

4.1 Methods

SMD simulations of hen-egg white lysozyme are pulled away from amorphous SiO_2 in a Generalized Born implicit solvent and in a TIP3P explicit solvent. Further technical details of the setup of the simulations are described in Hildebrand et al. [34]. The described simulations are extended by the following pulling protocol. In implicit solvent the molecules are pulled by a moving harmonic restraint on the side-chain nitrogen atom of LYS with a spring constant of $10 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$. The moving distance restraint between a surface atom and the nitrogen equals a pulling velocity of 5 m s^{-1} . The spring constant in the explicit solvent simulations is set to $1000 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ with a pulling velocity of 0.1 m s^{-1} . The velocity is a compromise between avoiding nonequilibrium effects and keeping computational feasibility [157, 162]. In one simulation a harmonic restraint between the sulfur atoms of two CYSs is set. This restraint replaces the native S-S bonds with variable force constants and an equilibrium distance of 0.204 nm. The restraint is placed with the program PLUMED [163] which is used as a patch on the GROMACS software [164]. Visualization and analysis of the trajectories are performed with VMD [114].

The functionalization of the AFM probes includes its silanization and the attachment of linker and lysozyme according to the references [94, 141, 165]. Single molecule AFM force-distance curves are measured by the protocol presented by Li et al. [94]. The breaking of the S-S bonds of lysozyme is achieved by using 1,4-DL-dithiothreitol (DTT) [141].

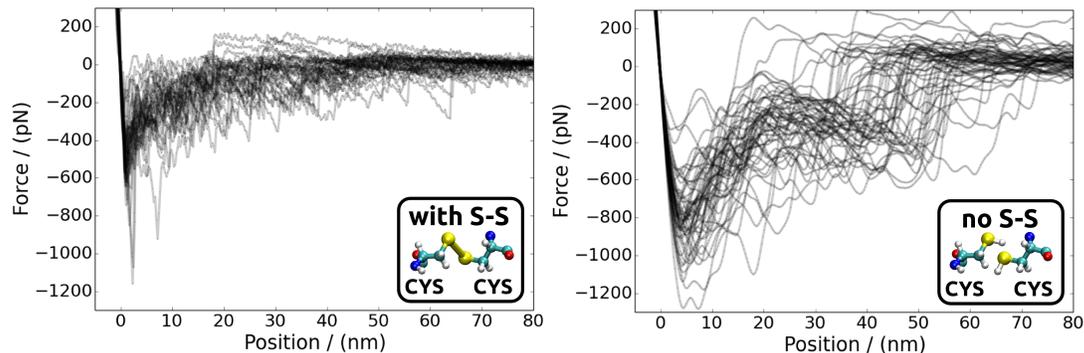


Figure 4.1: Lysozyme is pulled away from the SiO_2 surface without (left) and with (right) the S-S bond reducing agent DTT. The superposition of 33 and 52 AFM retract curves out of 256 in each case are shown. A MAL-PEG-NHS linker with a molecular weight of 1571 g/mol is used to attach lysozyme to the AFM probe.

4.2 Results

Lysozyme is attached by a PEG linker with a molecular mass of 1571 g/mol to an AFM probe. The interactions of lysozyme with SiO_2 in solution are measured by AFM force-distance curves as shown in (Fig. 4.1 left). Here many worm-like chain peaks (WLC) [166] peaks can be observed. If the solution contains the reducing agent dithiothreitol (DTT), the S-S bonds of the protein are selectively cleaved. The usage of DTT changes the shape of the resulting force-distance curves significantly (Fig. 4.1 right). The total force increases and a more plateau-like shape of the first peak is observed. This AFM experiment has been performed by Dr. Gang Wei of our own group.

To interpret the distances in the experimental AFM force-distance curves, the uncoiling of lysozyme and exemplary linker molecules is simulated. The molecules are stretched in SMD simulations. Lysozyme is pulled on the backbone N and C atoms on the terminal amino acids. The simulations are performed in Generalized Born implicit solvent with a pulling velocity of 2 m/s. The four S-S bonds of lysozyme are cut by protonation of the CYS residues to allow the protein to unfold completely in the classical MD simulation. With increasing distance of the terminal groups, the protein first loses its tertiary structure followed by the unfolding of the secondary structures. Finally, the bond angles and bond length are stretched which results in a dramatic increase in the pulling force as can be seen in the kink in the force-distances curve in Fig. 4.2. A positive force in all shown simulations corresponds to a force in the opposite direction to the pulling. In the following we perform a linear regression of the region before and after this change in the force. The intersect

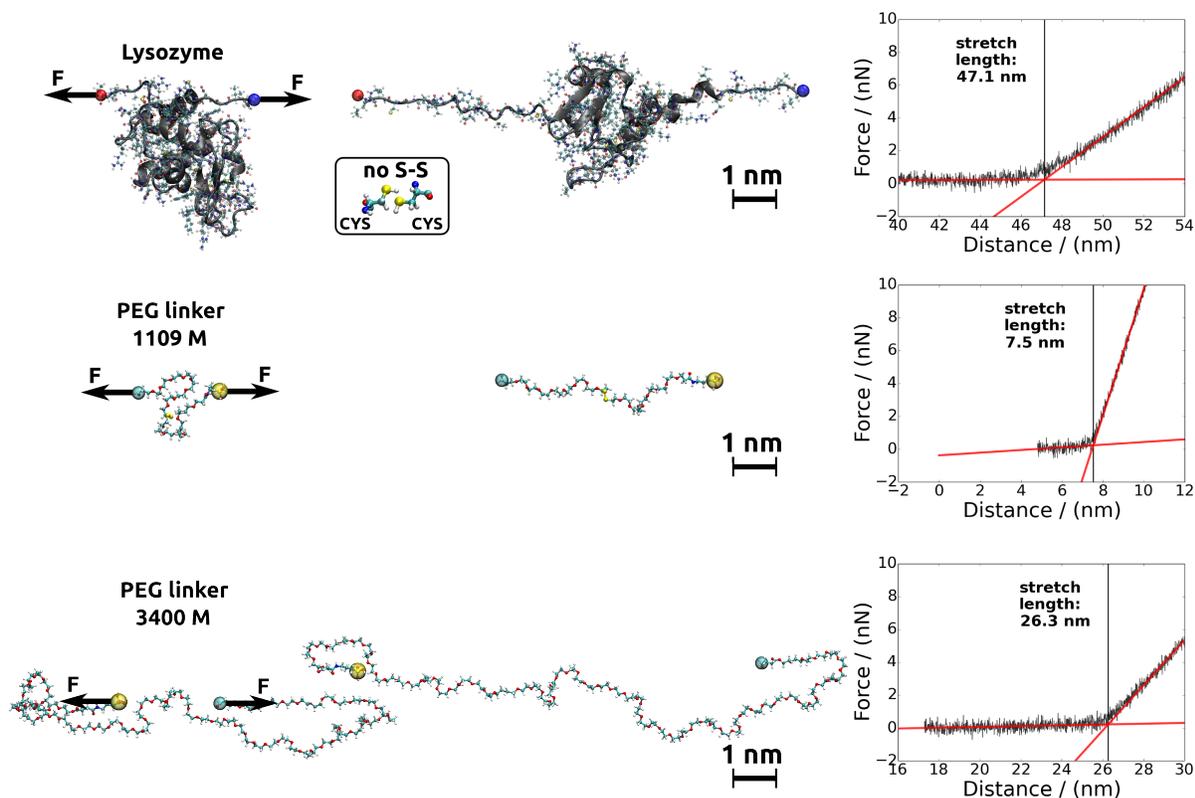


Figure 4.2: Steered MD simulations in implicit solvent to stretch lysozyme and two PEG linkers with molecular weight of 1109 g/mol and 3400 g/mol. (top) Lysozyme with cut S-S bonds is stretched on the N-terminal N and the C-terminal O atom (shown as blue and red atoms). The intersection (crossing of the red fitted lines) of the unfolding force with the force caused by bending of angular and bond potentials of the force field equation is shown in the right column. Therefore the fully stretched lysozyme has a length of about 47 nm. (middle) The 1109 g/mol linker is pulled on one terminal C as well as the Si atom of the attached APTES molecule (shown as cyan and yellow atoms). The succinimidyl ester is cut due to the expected reactive bond to the enzyme during preparation. (bottom) The same is done for the longer linker of 3400 g/mol.

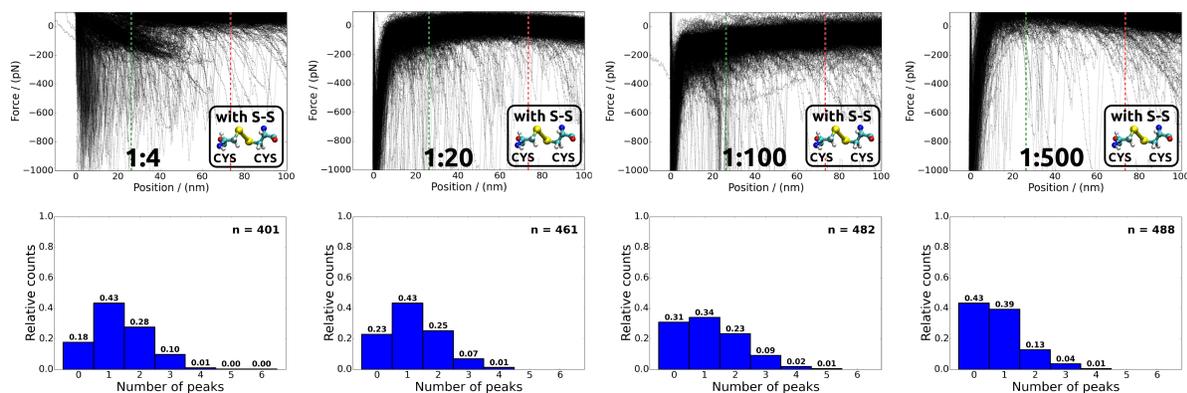


Figure 4.3: Lysozyme attached to an AFM probe is pulled away from a SiO_2 surface in an AFM experiment. Variations of the protein concentration at the AFM probe by a preparation with different APTES to TEES ratios between 1:4 and 1:500 vol.%. Further shown are the superpositions of 401 to 488 force-distance curves from originally 512. The background force of the retract curves is shifted to zero force as well as the initial AFM probe detachment is shifted to zero distance. The histograms show the peak number per curve which are selected by an algorithm described in the text. The NHS-PEG-NHS linker with a molecular weight of 3400 g/mol is used for the conjugation of lysozyme onto the AFM probe.

of both lines (shown in red in Fig. 4.2) is defined as the contour length of the protein or linkers. For lysozyme, this is a length of 47.1 nm. As an example for the stretching of the linkers, we stretch a short and a long linker of a molecular weight of 1109 M and 3400 M, respectively, including one APTES molecule. It is pulled on the terminal C atom as well as the Si atom of APTES. The shorter linker has a contour length of 7.5 nm and the longer one of 26.3 nm, including one APTES molecule with the length of about 0.5 nm from the Si to the N atom. Finally, it can be seen that interactions in the experimental force-distance curves with a distance longer than 73.4 nm are not related to a single-molecule event, as this length represents a combination of the stretch length of lysozyme and the long linker length. The minimum distance with a particular interaction is related to the attachment point of the linker at the AFM probe. Therefore a single protein event can also appear in minimum distances from 0 nm, equal to the initial probe-surface contact.

It is assumed that by the variation of the ratio of APTES to TEES, the concentration of binding sites and therefore of linkers at the AFM probe is varied. This is possibly correlated to the concentration of proteins bound to the probe. In this way, Gang Wei, who has performed this experiment has modified the ratio of APTES to TEES to identify the ratio in which only few proteins can interact with the surface. Ratios of APTES to TEES of 1:4, 1:20, 1:100 and 1:500 vol.% are used (Fig. 4.3). For each concentration 512

force-distance curves are measured. About 20 % of the curves are declared as nonspecific interactions and are therefore discarded, e.g. in case the force-distance curve include interactions in distances longer than the calculated maximum stretch length (Fig. 4.2). The resulting retract curves are shown as superposition in Fig. 4.3. All curves are aligned by shifting the force to zero at a distance of about 500 nm when no surface interactions are expected. A distance of zero is set to the point where the curves intersect + 200 pN. As expected, with an increasing concentration of TEES fewer peaks per curve are measured (Fig. 4.3 from left to right). The peaks do not noticeably cluster at distinctive distances and no obvious specific fingerprint can be observed suggesting complex dynamics during pulling. All individual curves are analyzed by their peak number. The derivative of the force-distance curves are used in an algorithm to automate this peak analysis. Through the derivative we are able to identify the steep slope after a rupture event. In particular, the used algorithm counts one peak if the derivative is higher than an arbitrary threshold, if the peaks are further apart than a distance of 6 nm, and if they are located in a distance range between 15 and 80 nm. All interactions smaller or bigger than this distance window are declared to be nonspecific interactions and are neglected. The peaks found by the algorithm are normalized by the total number of peaks and presented as relative counts in the histograms shown in the second row in Fig. 4.3. The decreasing peak number for increasing TEES concentrations as observed by the visualization of the superimposed force-distance curves can be confirmed by the systematic peak number analysis presented in the histograms. This trend indicates that the preparation with different ratios of APTES to TEES has a certain effect on the molecules attached to the AFM probe.

In order to further understand the experimental findings on an atomic scale, the following MD simulations are performed. First, in implicit solvent simulations lysozyme with intact S-S bonds is pulled away from the SiO₂ surface. Two start configurations are chosen from the result of an adsorption simulation in an explicit solvent shown in Fig. 2.7 in Chapter 2. We pull on the N-atom of one of the six lysines which are present in lysozyme and expected to be the attachment point of the used NHS-PEG linker in the experiment [141]. One of the 12 performed simulations in implicit solvent is shown in Fig. 4.4.

The dynamic simulations include an initial tertiary structure change with a subsequent unfolding of the secondary structure elements. We observe that the protein denatures at the surface within some picoseconds even without pulling, which is not the case for

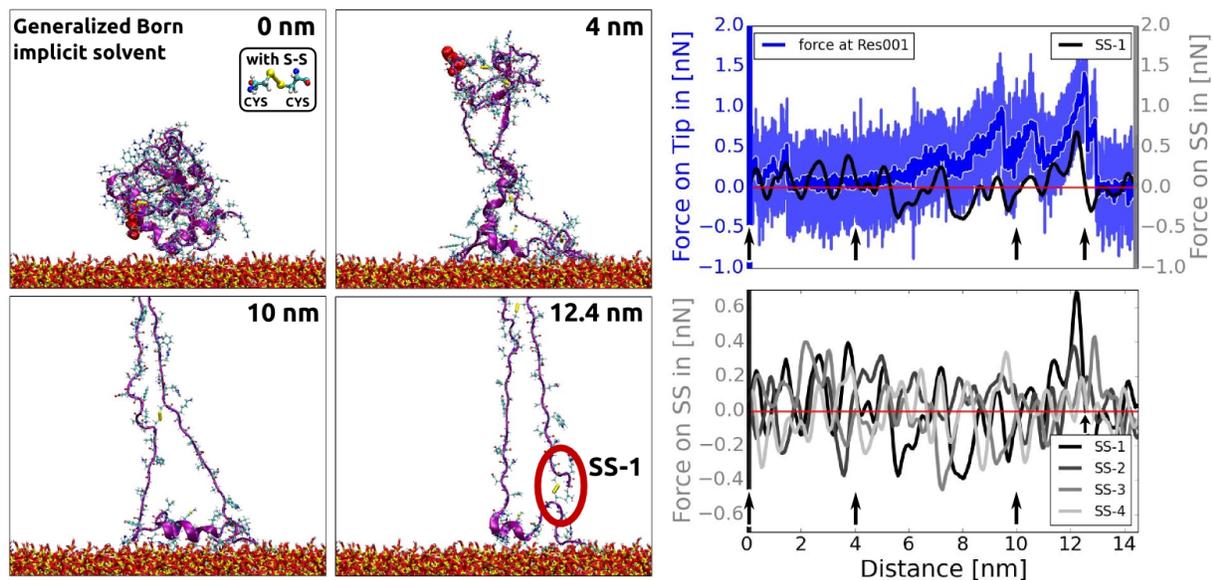


Figure 4.4: Pulling lysozyme of the SiO_2 surface in an implicit solvent. Shown is one exemplary simulation out of 12 performed simulations. The pull restraint is set on one of the 6 LYSs which are highlighted as red balls. All simulations are performed with intact S-S bonds. The four snapshots show distinctive frames of the exemplary trajectory. The start distance between LYS and the surface is set to 0 nm as declared in each shown frame. The force acting on the pulling LYS is shown in the top right as light blue data and smoothed with a Hamming window in dark blue. The curve is superimposed with the force acting on the SS-1 S-S bond to identify a correlation. This bond is additionally shown in the snapshot at 12.3 nm as the small yellow bar. The forces on all S-S bonds are shown at the bottom right subfigure. All forces on these bonds are smoothed using a Hamming window algorithm. The small arrows mark the position of the shown trajectory snapshots.

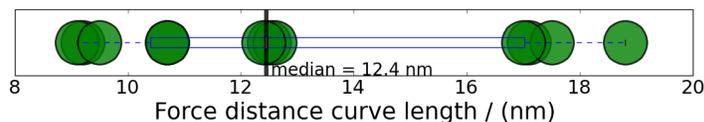


Figure 4.5: Boxplot of the force-distance-lengths directly before the detachment of the protein from the surface based on the 12 SMD simulations in implicit solvent (Fig. 4.4). The median of the pull distance to the detachment is 12.4 nm.

explicit solvent simulations for 300 ns as shown in Fig. 2.5 in Chapter 2. We can analyze the force-distance curves for the possible case that the protein unfolds during the pulling process with intact S-S bonds as shown in Fig. 4.4 in blue for one of the 12 force-distance curves. The related four snapshots include the starting configuration at 0 nm, a protein tertiary structure unfolding event at 4 nm, a detachment of one arginine residues at 10 nm and the snapshot just before the complete protein detachment at 12.4 nm. The bottom right of Fig. 4.4 shows the forces acting on all four S-S bonds of lysozyme, the distances of the snapshots are marked with small black arrows. It can be seen that a correlation between the forces acting on the pulling group and the S-S bonds exists, especially on the SS-1 bond. It is the bond connecting the CYS 6 - CYS 127 at the N and C-terminal of lysozyme which can be seen in the snapshot at 12.4 nm as a small yellow bar. The force acting on this bond is additionally presented as a black curve in the top right figure in Fig. 4.4 for a better comparison to the force acting on the pulling group. The correlation appears to be most significant right before the detachment point at 12.4 nm.

Because of the lack of a sufficient description of the surface-protein interaction in the used implicit solvent model, the direct comparison of the calculated and experimentally measured forces has to be taken with care. However, an estimation of the force-distance length with S-S bonds can be deduced. The median of the 12 distances of the final protein detachment points is 12.4 nm as presented in the boxplot in Fig. 4.5. The blue box marks the interquartile range which shows a pull length between 10.4 and 17.1 nm, in 50% of the cases.

For a more sophisticated description of the protein-surface interactions, in the following we present simulations in TIP3P explicit solvent. Identical to the simulations presented in Fig. 4.4, 12 simulations in two different start configurations are performed. The observed trajectories of these simulations are significantly different from those in the implicit solvent simulations. In all 12 cases, the protein detaches as a whole from the surface, keeping its native tertiary structure (Fig. 4.6). Apparently, the explicit water has a crucial effect on

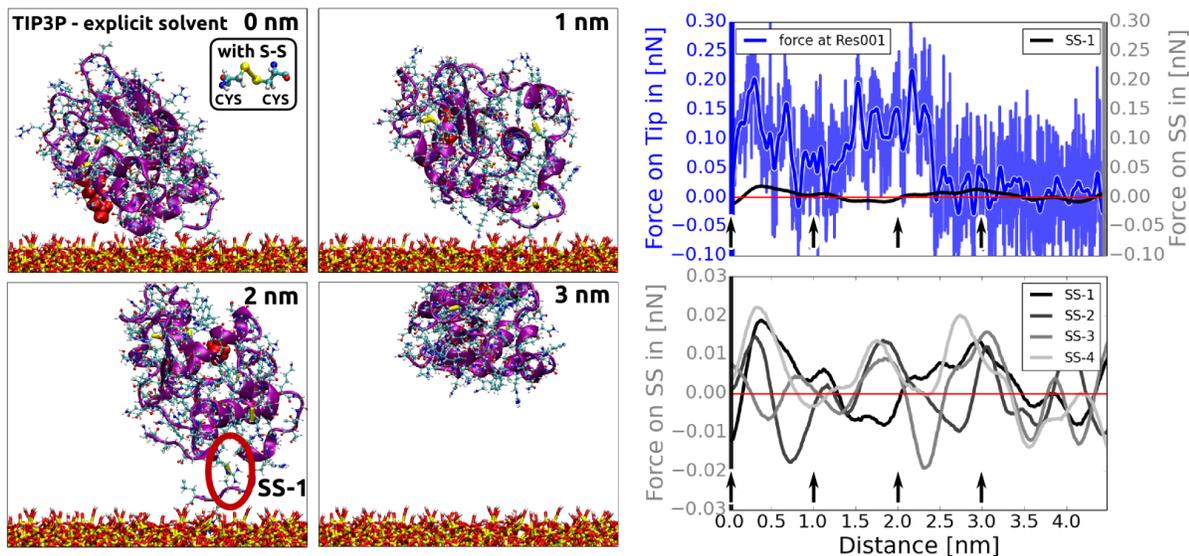


Figure 4.6: SMD simulation of lysozyme pulled away from a SiO_2 surface in an explicit solvent. Four snapshots show distinctive frames of one exemplary trajectory out of 12 simulations, all with native S-S bonds. The pull restraint is set on one of the 6 LYSs highlighted as red balls. The start distance between LYS and the surface is set to 0 nm as declared in each shown frame. The force acting on the LYS residue is shown in the top right as light blue data and smoothed with a Hamming window in dark blue. The curve is superimposed with the force acting on SS-1 to identify a correlation. SS-1 can be seen in the snapshot at 2 nm as the lower yellow bar. The force acting on all S-S bonds is shown at the bottom right subfigure determined by their atom distances. All forces on these bonds are smoothed using a Hamming window algorithm. The small arrows mark the position of the shown trajectory snapshots.

the stability of the protein as well as the protein-surface interactions. Furthermore, the observed force peaks are six times smaller. Snapshots from one representative trajectory are presented in Fig. 4.6. The second snapshot at 1 nm pull-length shows a detachment from one part of the protein from the surface. In the third snapshot at 2 nm the S-S bond SS-1, shown as a small yellow bar, should be under the maximum force load since it forms a bridge between surface and protein body. Even at this point, no significant correlation can be identified between the force acting on SS-1 and the pulling force (Fig. 4.6 top right). The snapshot at 3 nm distance shows the detached protein still in the native structure. No correlation between the force acting on the other S-S bonds and the force acting on the pull molecule at any distance can be identified (see bottom left subfigure in Fig. 4.6).

All distances of the protein detachment are shown in the boxplot in Fig. 4.7. With a median value of 2.0 nm, they are six times smaller than in the implicit solvent simulations

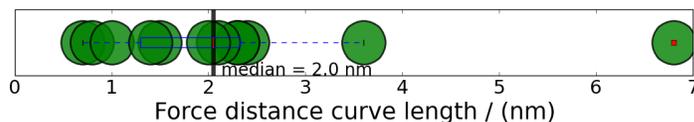


Figure 4.7: Box plot of all force-distance-lengths directly before the detachment of the protein from the surface based on 12 pull simulations in explicit solvent (Fig. 4.6). The median of the pull distance to the detachment is 2.0 nm.

due to the described absence of denaturation. In one of the simulations, the pulling of LYS starts very close to the surface which results in a very long pulling distance of 6.8 nm before the protein detaches. This causes the outlier in Fig. 4.7. The interquartile range of the shown box spans between 1.3 and 2.3 nm.

The explicit solvent simulations are further analyzed to investigate the protein-surface contacts. Fig. 4.8 shows the individual amino acids which are in contact with the surface. A contact with the surface is defined as an atom-atom distance smaller than 2.4 Å (see also Section 2.1.2). The color of the plot as well as the size and color of the dot which marks the last protein detachment are weighted by the peak height of the force curves. The weighting is done with an arbitrary nonlinear rational switching function to have a distinct visual effect. If a high force and the final detachment are in the same region, a big red dot is drawn, highlighting amino acids which contribute to a high force peak. The amino acids of the contact analysis shown in Fig. 4.8 are sorted by the results of the described weighting from top to bottom. As can be seen, ASN19, LYS13, SER81 and LEU84 form contacts which produce the highest forces in the simulations. The most frequent amino acids in this analysis are the positively charged amino acids ARG and LYS. Other types of amino acids are in contact with the surface as well, as similarly observed and described in Chapters 2 and 3. The highest peak in the force-distance curves correspond to the simultaneous detachment of many amino acids at once.

To simulate the DTT experiments, the S-S bonds are cut and the CYS side chains are protonated in an explicit solvent. The simulations are started in two different orientations, identical to the simulations described previously. In the absence of the S-S bonds the protein does not remain in its native structure and starts to unfold as soon as the pulling force is applied (Fig. 4.9). In orientation 1, the main body of the protein detaches while the surface contact remains intact. In the second orientation, the portion of the protein with the applied force starts to detach first, whereas the main body of the protein remains at the surface. The following frames of orientation 1 are shown in Fig. 4.9. At 0.9 nm

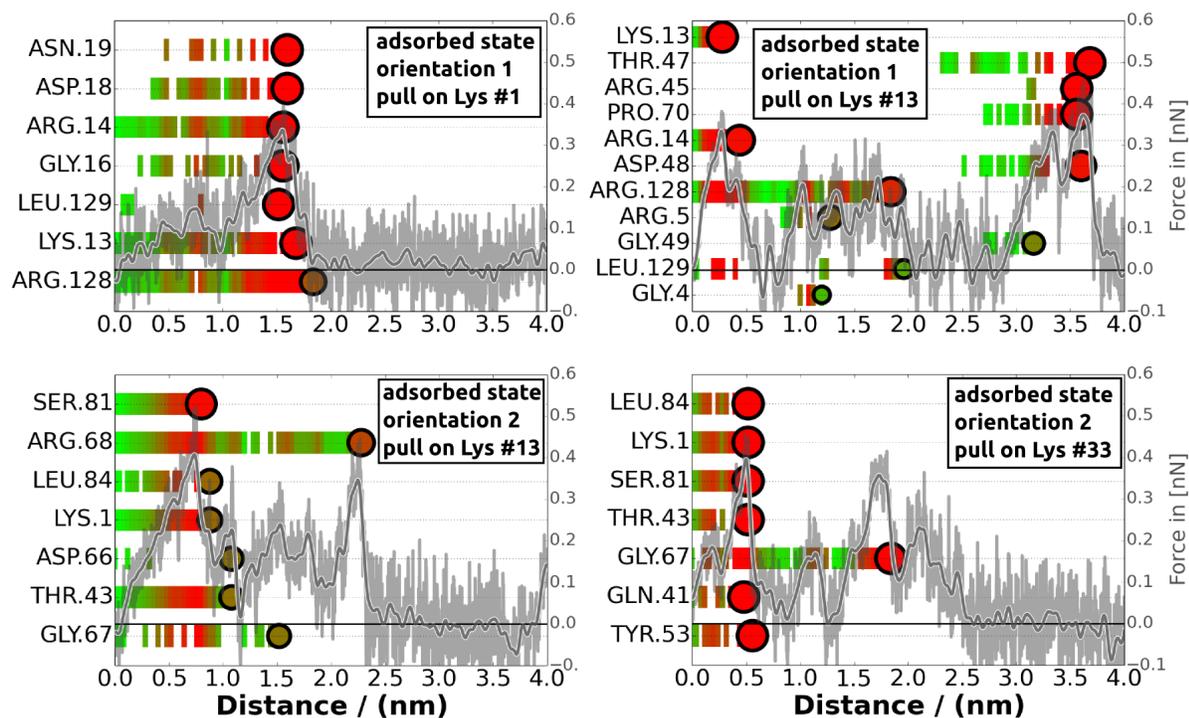


Figure 4.8: Surface-protein contact analysis of four arbitrary chosen simulations out of the 12 performed simulations in explicit solvent. The gray curves are the force-distance curves of lysozyme pulled away from the surface. If one atom of the presented amino acids is in contact with the surface, a small colored bar is drawn. The final detachment point is marked with a circle. The color represents the weights by the force-distance curves with an arbitrary weight function, red being identified with a high force. The final detachment is drawn as a circle, weighted by its size and color according to the intensity of the force. Therefore, very tight surface contacts which result in high forces are presented as big and red circles.

one ARG produces a snap back of the force-distance curve shortly before its detachment from the surface. At 2 nm a strand including one S-S bond detaches from the rest of the protein. At 5.6 nm two helix loops produce another snap back of the force-distance curve right before the separation of the protein. The trajectory of orientation 2 is presented in the frames shown in the middle row in Fig. 4.9. At 1.8 nm one portion including several amino acid - surface contacts is detached from the surface. At 3.3 nm the protein opens a pocket-like cleft which can be seen on the left side. At 4.7 nm the loose strand at the right is losing the contact to the rest of the protein. In both simulations, periodic boundary conditions are used due to the included explicit solvent. Therefore, the box size in the surface normal direction is limited to about 10 nm and the simulations are stopped at a pulling length of about 6 nm.

Comparing Fig. 4.6 and Fig. 4.9 it becomes immediately clear that the presence of the S-S bonds is crucial for the stability of the protein during pulling. In the subsequent simulation, one S-S bond is replaced by a harmonic restraint with four different force constants (Fig. 4.10). The equilibrium distance is set to 0.204 nm, which is the distance of the S-S bond in the used AMBER force field. Similar to the previously described simulations, the protein is pulled away from the surface in an explicit solvent with a pulling velocity of 0.1 m/s. The start configuration is the same as in Fig. 4.9 (orientation 1). This configuration is chosen because the C-terminus is attached to the surface, which stretches the bond SS-1 (CYS6 - CYS127) after a short pulling distance. With increasing pulling length, SS-1 is increasingly stretched until the protein detaches from the surface. A correlation between the relative force acting on SS-1 and the pulling force acting on the LYS is shown in the force-distance curves in the right column in Fig. 4.10. Especially in the case of the highest force constant of 300 kJ/mol/nm², the distinctive correlation right before the detachment can be seen. The relative force is set to a relatively constant distance which is present before the protein detachment. The increase of the relative force between the two CYSs is 10 to 80 pN before the protein detaches from the surface.

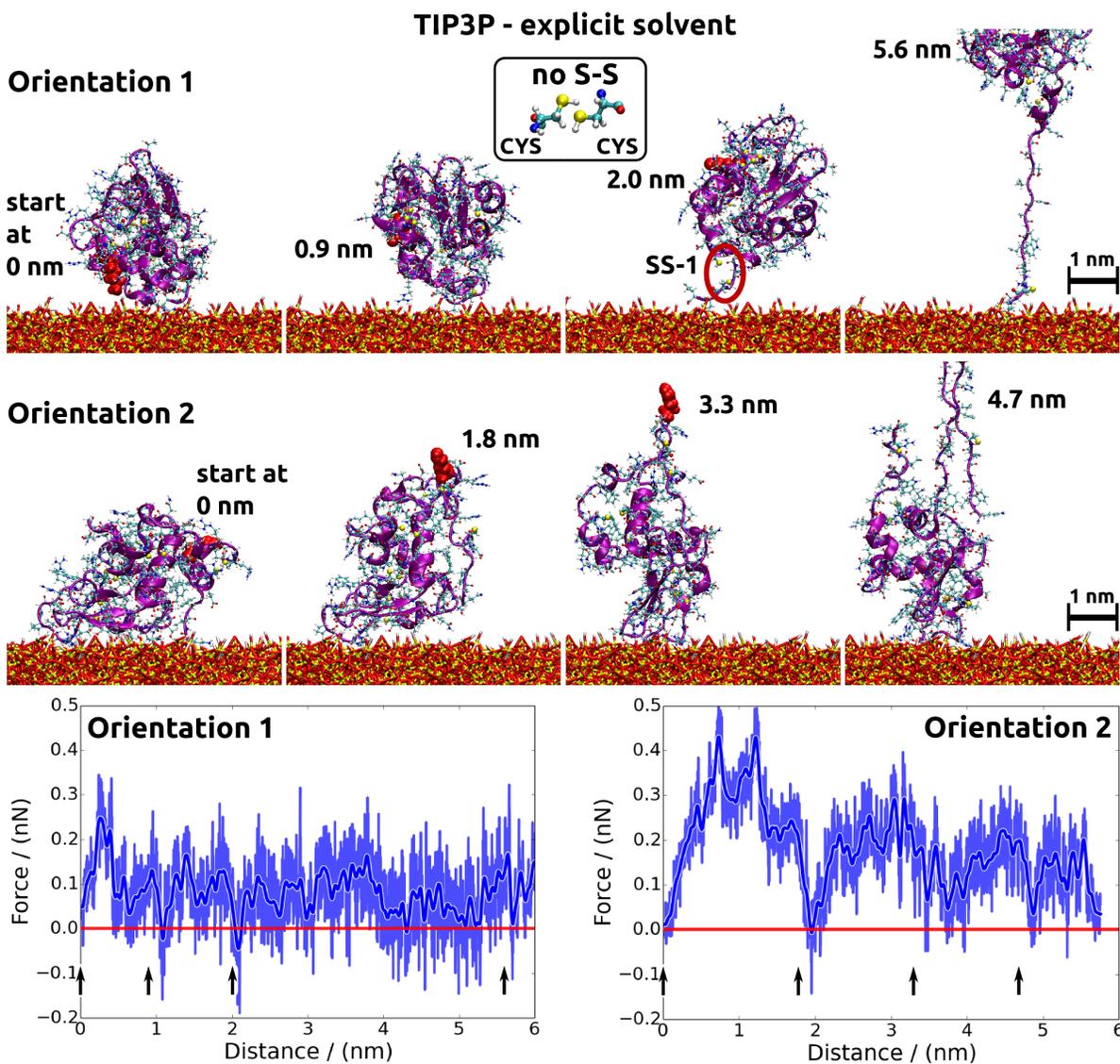


Figure 4.9: Pulling lysozyme off the SiO_2 surface in an explicit solvent with cleaved and protonated S-S bonds. Shown are two different start orientations. The pull restraint is set on one of the 6 LYS highlighted as red balls. The four snapshots show distinctive frames of the trajectories. The start distance between LYS and the surface is set to 0 nm. The force acting on the pulling LYS residues is shown in the bottom two subfigures in blue. The protein denatures in both cases. The small arrows mark the position of the shown trajectory snapshots.

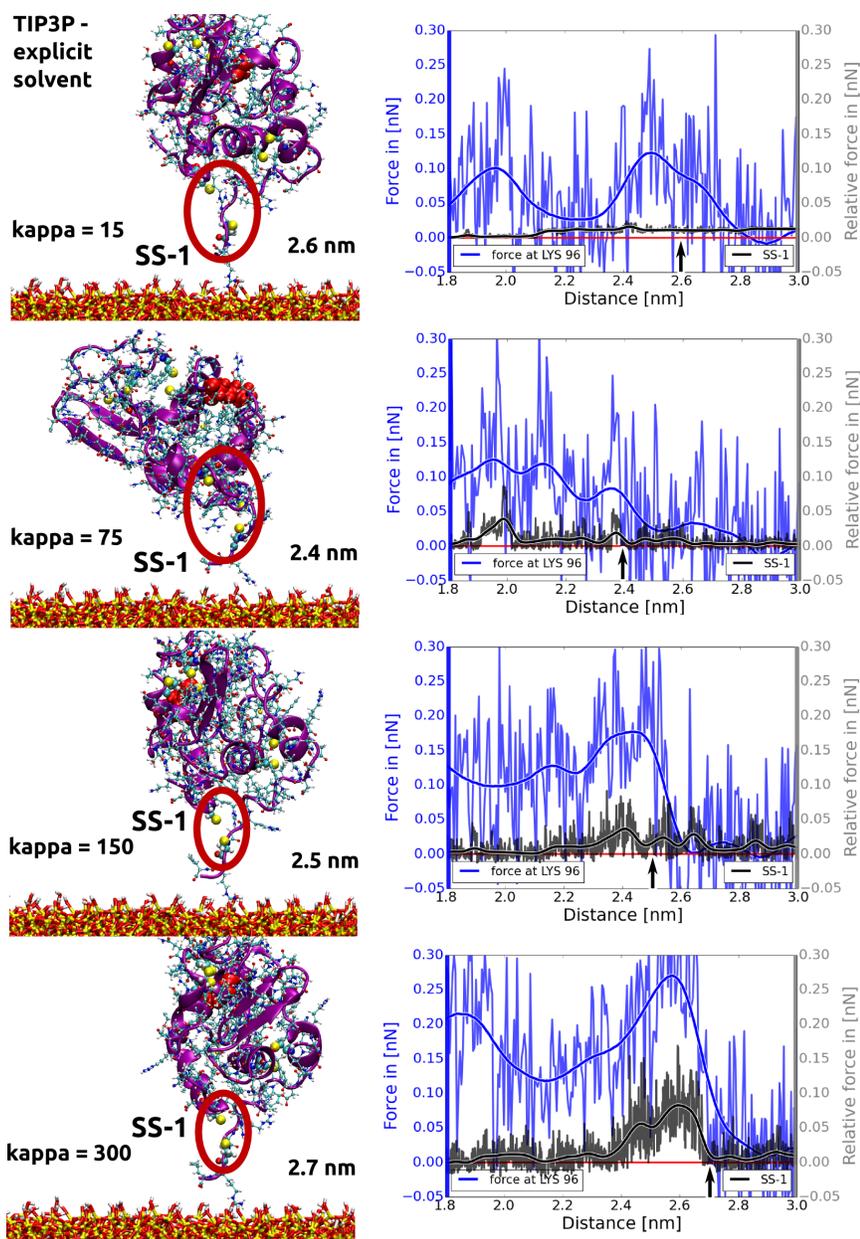


Figure 4.10: The S-S bonds of lysozyme are replaced by harmonic restraints with different force constants (κ , declared in kJ/mol/nm^2) and an equilibrium distance of 0.204 nm . The start configuration in the four simulations is identical to Fig. 4.9 (orientation 1). Lysozyme is pulled away from the SiO₂ surface on one of the 6 LYSs highlighted as red balls. The start distance between LYS and the surface is set to 0 nm . The pulled distance is presented in each snapshot at the right. The right column shows the force acting on the pulling LYS residue in blue. The relative force acting on the S-S bond SS-1 is shown in gray. It is the total force acting on this S-S bond normalized to the force just before the detachment. The snapshots are chosen right before the detachment of the protein in each of the four simulations. This point is marked with an arrow in the force-distance curves.

4.3 Discussion

The SMD simulations of lysozyme being pulled away from the SiO₂ surface in an explicit solvent demonstrate that the protein detaches from the surface without an observable conformational change. The preservation of the conformational structure was also observed by 14 similar simulations performed by Kubiak-Ossowska et al. [150]. In a further simulation shown in Fig. 4.9, the S-S bonds are broken and protonated *a priori*. In this case, the protein denatures during the pulling process. Therefore, the S-S bonds seem to have a crucial effect on the protein stability.

In an additional simulation the distance between the sulfur atoms of two cysteines by a variable force constant has been restrained as shown in Fig. 4.10. A relative force of the magnitude of 10 to 80 pN acting on an S-S bond has been measured shortly before the protein detaches from the surface. The force acting on the protein is 100 - 250 pN at the same time. Wiita et al. [143] have shown that S-S bonds cannot be broken in an AFM experiment with a constant force of 100 pN in seconds and 165 pN in milliseconds. Therefore it should be argued that the proteins shown in the AFM experiments detach before the applied force breaks the S-S bonds.

The presented AFM experiments with and without reducing agent show different shapes of the force-distance curves (Fig. 4.1). Due to the results of the simulations it can be assumed that the protein with cut S-S bonds denatures during the pulling in the AFM experiment. The unfolding results in longer force-distance curves due to the chain-like structure detaching from the surface. The slightly increased forces in the experiment with DTT can be explained by a simultaneous detachment of several lysozyme chains at the same time. The attached chains also cause the plateau-like shape of the force-distance curves.

Wei et al. [141] reported that lysozyme produces a fingerprint of four peaks which were interpreted as the breaking of the four S-S bonds. The presented findings of the simulations suggest that this pattern observed by Wei et al. is rather related to an increased number of proteins at the AFM probe. These proteins detach without denaturation and produce a certain pattern in the force distance curves. The ratio of the active to passive linker binding sites must rather be related to the number of proteins at the probe, which therefore influences the number of peaks in the force-distance curves.

The assumed process of the detachment is shown by the setup of some proteins attached

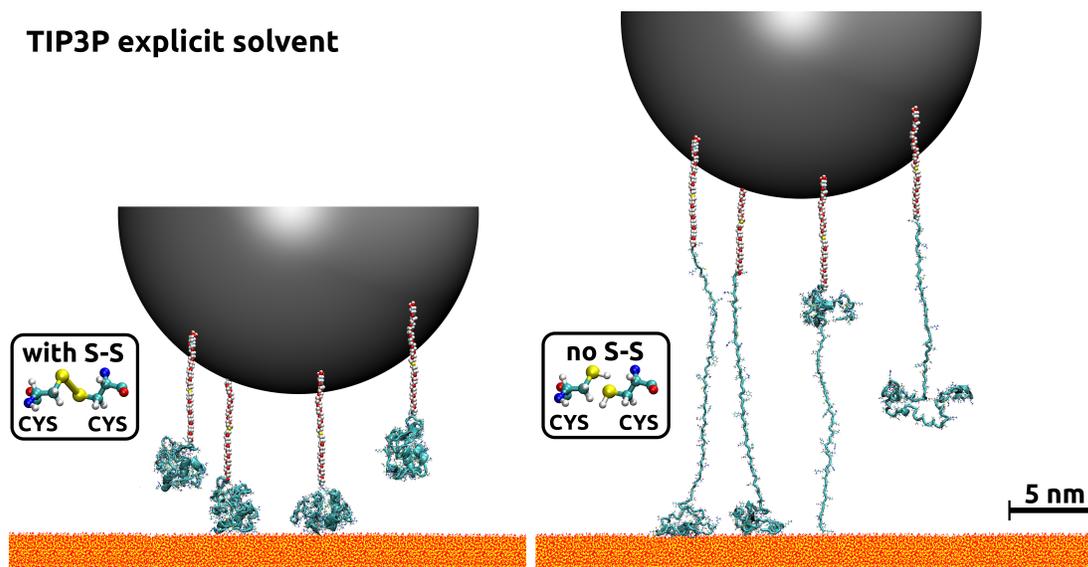


Figure 4.11: Sketch of a lysozyme functionalized AFM probe over the SiO_2 surface. The ball represents a realistic AFM probe radius of 10 nm. The proteins are depicted in turquoise cartoon shape, a linker with the molecular weight of 1109 g/mol in ball-like strings and the surface in yellow-red balls. Shown is the situation without (left) and with DTT (right). The protein denaturates only when the S-S bonds are broken using DTT.

to the AFM probe in Fig. 4.11. With all S-S bonds intact the protein detaches without denaturation (*left*). With broken S-S bonds in the simulation or by adding DTT in the experiment, the protein seems to denaturate at the surface (*right*). Two types of denaturation can be observed: (1) the detachment of the main part of the protein and the unzip of one strand still attached to the surface; (2) the unzip of one strand while the central part of the protein remains at the surface.

In contrast to the simulations in an explicit solvent, the simulations in implicit water show a denaturation of the protein during the pulling process. This observation can be explained by two aspects. On the one hand, the protein-surface interactions seem to be highly increased with the usage of the Generalized Born formalism. This is underlined by the much higher detachment forces compared to the explicit solvent simulations (Fig. 4.4 - 4.6). Probably the water reduces the surface interactions by screening as well as by steric effects. On the other hand, the water forms a network of hydrogen bonds within the first water layer and around the protein, which highly stabilizes the protein as described by Heymann et al. [160]. No correlation of a force acting on the pulling group and the force acting on the S-S bond is observed in the explicit solvent simulations (Fig. 4.6).

Probably the hydrogen bond network and the first water layer around the protein can absorb external forces and mechanical perturbation applied to the protein.

The protein-surface contacts are presented in relation to the force-distance curves in Fig. 4.8. In most cases the detachments of many surface contacts are responsible for the peaks in the force-distance curves. The identified binding motifs are mainly the positively charged amino acids ARG and LYS but also other amino acids: polar uncharged, negatively charged and even hydrophobic ones. This is in agreement with the results presented in Chapter 2 as well as the work of Kubiak-Ossowska et al. [111, 120, 150]. The variety of amino acids involved in the adsorption process can be explained by a simultaneous contribution of nonbinding enthalpic interactions as well as entropic influence of the local water structure at the solid/liquid interface [55, 115, 117, 119]. A differentiated investigation of both types of interactions can be of interest for further studies to identify the leading mechanisms of the protein-surface contacts.

Surface-Induced Conformational Changes

The production of a protein-functionalized surface requires that the protein-surface interactions do not induce a protein denaturation, resulting in a reduced activity of the protein. It is assumed that proteins adsorb tightly on hydrophobic surfaces but have an increased tendency to denature during adsorption in comparison to hydrophilic surfaces [6, 14, 15, 17, 18]. Therefore, on hydrophilic surfaces such as SiO_2 and TiO_2 , as investigated in this dissertation, a decreased probability of conformational changes can be assumed. Nevertheless, even small conformational changes can influence the protein activity or initiate further body reactions. For this reason, their exact knowledge is of interest for the development of applications involving protein adsorption.

As an example for protein adsorption on a hydrophilic surface we are going to investigate the possible conformational changes of adsorbed chymotrypsin on SiO_2 . In Chapter 2 and 3 we have shown that chymotrypsin has a very distinct binding orientation on SiO_2 , namely with the two helix conformations facing the SiO_2 surface. It is demonstrated that chymotrypsin anchors strongly with its positively charged residues in contact with the silanol groups of SiO_2 . This anchoring, however, could change the secondary structure of the helix in contact with the surface. Classical molecular dynamics simulations of 300 ns show no conformational changes of the adsorbed protein (Fig. 2.5). A standard MD of a complex biological system is often trapped in local minima of the energy landscape. In this kind of simulations, the system does not reach any state in a finite number of iterations. Therefore, standard MD cannot capture so-called rare events, as a protein conformational change can be considered. On the atomic scale, a rare event can consist of the statistical crossing of a cascade of several microscopic barriers e.g. angular rotations

at ones [86]. The unlikelihood of these events results in energy barriers bigger than a few $k_B T$, as derived by the Arrhenius equation described in the Appendix in 7.5.1. To reduce this non-ergodicity of the simulation, a large sampling is required in order to overcome these barriers, which does normally not happen in typically short simulation times of ~ 100 ns. In the following more advanced MD simulations are presented which include additional bias potentials on top of the unbiased force field potential to compensate the poor sampling of a standard MD simulation [167].

A very promising advanced MD technique to investigate conformational changes in biomolecules is the usage of Replica Exchange with Solute Tempering (REST) in combination with Metadynamics. The formalism of this method is given in Appendix 7.5. This method has been used to predict the conformational structure of peptides in solution [55, 123, 168]. Bussi [169] presented an extension of this formalism to scale and therefore accelerate a bonded part of a molecule. This idea is adapted and extended to proteins in this chapter. Concerning the REST approach [167], the part of interest of the protein is defined as the solute whereas the remaining system is defined as the solvent (Appendix 7.5.5).

The REST approach enhances the sampling of the canonical ensemble by coupling the system to replicas at higher effective temperatures. The Metadynamics enhances the sampling by accumulating repulsive bias potentials at the current position of the simulated system along a collective variable (CV). The CV is a transformation of a complex set of variables such as the torsional backbone angles in a peptide in one coordinate, for instance the helicity value (Appendix 7.4.1). The combination of the REST method with Metadynamics combines the strengths and compensates for the weaknesses of each method alone [55, 123, 168]. The solute tempering is independent of the chosen CV and samples all slow degrees of freedom. This sampling, however, depends on the height of the effective temperature and is therefore limited for the crossing of very large energy barriers. Metadynamics, on the other hand, samples very efficiently along the CV and is able to cross moderately high barriers. Slow degrees of freedom not reached by the chosen Metadynamics CV can partly be compensated by the solute tempering.

The combination of REST with Well-Tempered Metadynamics [170] increases the sampling of each method alone for two reasons. First, a missing slow degree of freedom, not addressed by the chosen collective variable (CV) of the Metadynamics, might be sampled by REST. Second, the sampling along a thoughtfully chosen CV with Metadynamics works highly effectively even for bigger local free energy barriers, whereas the REST approach

is only able to sample small local barriers in an efficient way. In order to increase the temperature of the CV of the Metadynamics along the replica ladder we have additionally used the Replica Exchange with Collective Variable Tempering (RECT) approach [171] (Appendix 7.5.5). In this case, the small CV temperatures of the lower replicas increase the fine adjustment of the Metadynamics bias whereas the high CV temperatures of the upper replicas compensate very high energy barriers. This combined solute tempering method is compared to two additional methods of free energy simulations which have been further developed in this work.

One of these methods is based on Umbrella Sampling [172] in combination with Well-Tempered Metadynamics [170]. Each umbrella holds the system by a harmonic restraint in one window along the CV. In each of these umbrellas, an individual Metadynamics simulation is performed. This approach increases the sampling within and in particular at the edges of the umbrellas in comparison to standard Umbrella Sampling. The bias at each trajectory frame is the combination of the bias of the Metadynamics and the bias of the umbrella restraint. The sum of both bias potentials is analyzed by the weighted histogram method (WHAM, [173,174]) (Appendix 7.5.2) to produce a free energy surface.

The third extended and used free energy sampling method is also based on the mentioned umbrella method. Two expansions are included in this approach. First, exchanges between the umbrella replicas following the Metropolis-Hastings criterion similarly used in replica exchange simulations. Second, all applied Metadynamics hills contribute to one bias potential, which is called Multiple Walkers Metadynamics [175].

These free energy sampling methods are compared to secondary structure propensities, measured experimentally by circular dichroism experiments (CD). The CD measurements of protein adsorption on SiO₂ nanoparticles are particularly suitable, since these nanoparticles have a non-birefringent character which makes them invisible in the CD signal. Therefore, they do not interfere with the measurement and the adsorbed proteins can be well investigated. CD experiments of chymotrypsin adsorption on SiO₂ performed by Nina Wurzler and Monika Michaelis have shown a change in the CD spectra after adsorption, suggesting a surface-induced conformational change (Fig. 1.9 [99]). A direct link between the experimentally measured spectra and the simulations can be made via the theoretical calculation of the CD spectra based on the atom positions of the sampled trajectory by using the DichroCalc web interface [176].

Other authors have suggested a surface-induced conformational change for chymotrypsin. Zoungrana et al. [24] described a loss of the helical structure during adsorption of chymotrypsin on SiO₂ based on CD experiments. Furthermore, differential scanning calorimetry investigations by Welzel [98] show only a minor conformational change during adsorption. A possible explanation bringing both studies together would be a local loss in helicity without dramatic changes in the tertiary structure while chymotrypsin adsorbs on SiO₂.

In this chapter we focus on possible conformational changes of the helix at the C-terminus of chymotrypsin, which contains 14 amino acids. This C-terminus helix is investigated in three different environments. (1) In an aqueous solvent, (2) the helix natively embedded in the protein and (3) the protein adsorbed on SiO₂ with the helix in direct contact with the surface. A similar environmental effect on the C-helix of lysozyme was studied by Bolin et al. [177] using NMR and CD experiments. In the native protein, the C-helix was found to be helical, whereas in an aqueous solvent it showed a random coil structure. By increasing the concentration of trifluoroethanol, a helix structure could be formed again. Cole et al. [121] performed all-atom standard MD simulations of a collagen fragment on hydrophilic and hydrophobic SiO₂. In solution as well as on the hydrophilic SiO₂ the collagen fragment formed a random coil conformation, whereas on the hydrophobic surface the fragment remained helical and adsorbed via a larger number of surface contacts. Experimentally it was found that the α -helical form of the collagen fragment adsorbed well on a hydrophobic surface [178]. A further study by Long et al. [179] has investigated a small peptide on a hydrophobic surface and confirmed the formation of a helix structure. On the other hand, studies on hydrophilic SiO₂ nanoparticles by Lundqvist et al. [180] as well as Nygren et al. [181] demonstrated that the helix propensity of a peptide can be enhanced during adsorption. The effect of an enhanced propensity towards an helix structure by adsorption seems to depend on a particular combination of the amino acid sequences and the surface properties. Another study by Lundqvist et al. [182] investigated in detail the effect of SiO₂ particle sizes on the tertiary and secondary structure change during protein adsorption. They have found an increased perturbation on the secondary structure, but no trend on the tertiary structure change, with increased nanoparticle diameter. It is suggested that a large interaction area affects the secondary structure strongly than a small one.

5.1 Methods

The MD simulation setup in this chapter is based on the simulation protocols used in Chapter 2 - 4 [34]. Additional specific advanced MD techniques are presented in the following. The theoretical background of these techniques is described in the Appendix 7.5.

5.1.1 Well-Tempered Metadynamics

All presented free energy sampling methods are based on the Well-Tempered Metadynamics introduced by Barducci et al. [170]. The formalism and further details of this method is given in the Appendix 7.5.4. As collective variable (CV) for the Metadynamics, we have used the helicity of the amino acids of one of the two helices of chymotrypsin, as introduced in PLUMED 1 helix loops [183]. We have implemented this CV into PLUMED 2 by using the matheval library [163]. The CV value is calculated by the deviation of the dihedral angles Φ_j and Ψ_j from the ideal angles of a helix $\Phi_0 = -68.75^\circ$ $\Psi_0 = -45.0^\circ$ (Appendix 7.4.1) [123, 183].

In a previous standard MD simulation of chymotrypsin in water for 300 ns, we have analyzed the secondary structure of each amino acid using the Define Secondary Structure of Proteins (DSSP) method by Kabsch and Sander [184], included in AmberTools [76]. All amino acids which remain longer than half of the simulation time in a helical structure are defined as an helix amino acid. By this definition, we have found two persisting helix structures in the native chymotrypsin. One short helix including the amino acids 165 - 172 and one at the C-terminus including the amino acids 231 - 242. In this chapter, we focus on the second helix. To include all amino acids of the C-terminus we have used the 14 amino acids 231 - 244 in the helicity CV. In the present case, an ideal helix of the used 14 amino acids has a CV value of $N - 2 = 12$ (Appendix 7.4.1).

For the Well-Tempered Metadynamics, we have used a hills width of 0.35 rad and a deposition frequency of 1 ps⁻¹. The initial height of the deposited Gaussian is chosen by $h = (k_B \Delta T / \tau) N \Delta t$ [171]. k_B is the Boltzmann constant, ΔT is the boosting temperature and τ is the characteristic time for the bias evolution and chosen equal to 3 ps. The boosting temperature is defined as $\Delta T = T(\gamma - 1)$ [171, 185] with T being the system temperature and γ being the bias factor. γ equals the values along the replica ladder of 4, 4.6, 5.4, 6.2, 7.2, 8.3, 9.6, 11.1, 12.9, 14.9, 17.3, 20. The bias factors result in boosting temperatures ΔT along the CV from 900 to 5700 K according to the latter formula.

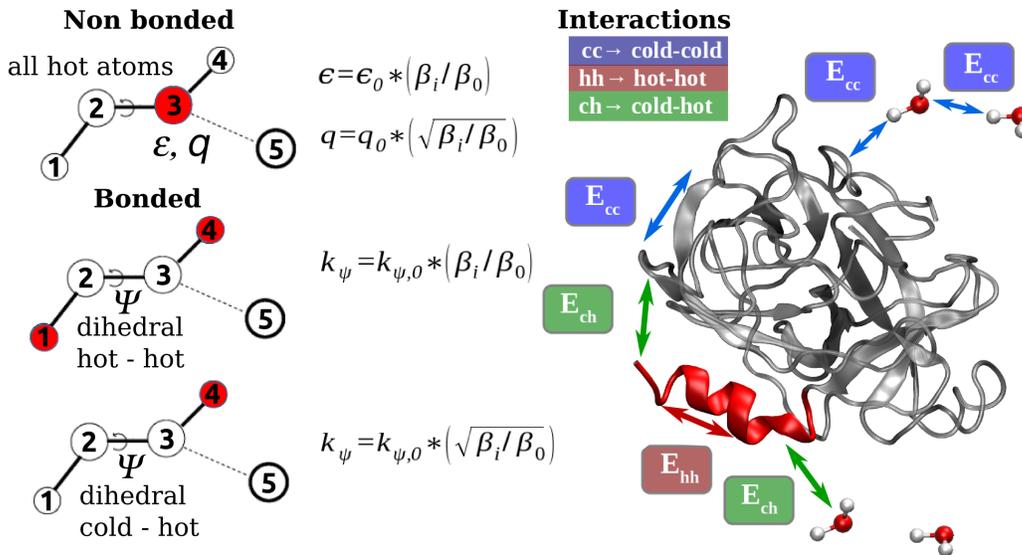


Figure 5.1: Parameter scaling of a solute tempering simulation of one protein part following the REST2 approach. (left) Schematic drawing of four bonded (1-4) and one non-bonded atom (5). Atoms defined as “hot” are red. The Lennard-Jones parameter ϵ as well as the charge q of the “hot” atoms are scaled by the presented formula. The dihedral interactions are scaled differently if both or just one distal atom is “hot”. (right) The C-terminus helix of chymotrypsin is defined as “hot” (red). Interactions are separated into cold-cold, hot-hot and hot-cold interaction types. $\beta_i = 1/(k_B T_i)$.

5.1.2 Solute Tempering - Defining Parts of a Protein as “hot”

A particular chosen part of the protein is scaled by the solute tempering approach to accelerate its dynamics. In this chapter we have used the REST2 method [167, 169] on one helix of chymotrypsin (5.1.3). The helix is scaled and defined as “hot”, while the rest of the system is defined as “cold” [169]. The Lennard-Jones parameters ϵ of the “hot” atoms are scaled by the factor β_i/β_0 with $\beta = 1/(k_B T)$. k_B is the Boltzmann constant and T the effective temperature. β_i is the factor including the effective temperature T_i , while β_0 includes the temperature T_0 of the lowest replica, often chosen as 300 K (Fig. 5.1). All charges q of the “hot” atoms are scaled by the factor $\sqrt{\beta_i/\beta_0}$. If both distal atoms of a dihedral bond (atom 1 and 4 in Fig. 5.1) are “hot” the dihedral force constant is scaled by β_i/β_0 . If one of the distal atoms is “hot” and the other is “cold”, the dihedral constant is scaled by $\sqrt{\beta_i/\beta_0}$ (Fig. 5.1) [169]. “Cold” atoms are not scaled at all.

5.1.3 Replica Exchange with Solute Tempering in Combination with Well-Tempered Metadynamics (RESTmetaD)

Replica exchange simulations include the parallel simulations of several systems with different conditions e.g. different temperatures. Exchanges between the replicas follow a metropolis-hastings acceptance criterion and allow the diffusion of the systems along the replica ladder. In the solute tempering replica exchange approach (REST) each replica is scaled by the protocol introduced in Section 5.1.2 representing effective temperatures (Appendix 7.5.5). In previous exchange probability tests we have identified a suitable usage of 12 replicas. The replicas contain the effective temperatures of 300, 324.0, 350.0, 378.0, 408.3, 440.9, 476.3, 514.4, 555.6, 600.1, 648.1, 700 K following a geometric progression, due to the wider energy distribution of higher temperatures [186]. These temperatures are used to generate the REST2 scaling factors $\beta = 1/(k_B T)$ [167] as described in more detail in Section 5.1.2. Exchanges between the replicas are applied every 400 fs. Based on an overlap of the potential energy distributions of each replica, the average exchange acceptance probability is 46% for the helix system, 47% for the protein system and 48% for the protein/surface system. Furthermore, the average round trip time (rtt) is computed. This is the time one replica needs to move up the complete temperature ladder and back again. The rtt counts 0.48 +- 0.04 ns for the helix system, 0.77 +- 0.84 ns for the protein system and 0.51 +- 0.12 ns for the protein/surface system.

5.1.4 Umbrella Simulation in Combination with Metadynamics (MetaUmbrella)

The motivation of this approach is a forced sampling along the full CV range, independently of its complexity or included local barriers. Parallel independent simulations are performed in smaller regions of the CV which are restrained by the harmonic umbrella. The equilibrium position of the harmonic restraint is shifted for each replica to minimize tensions in the start system. The start value is chosen in equidistant ranges between a CV value of 0 and 12 as shown in Fig. 5.2 b). The system is driven to this start value of the CV by a 1 ns target MD simulation with a harmonic restraint force constant of 100 kJ/mol. A standard Umbrella Sampling is combined with a Well-Tempered Metadynamics simulation applied on each umbrella. Due to the Metadynamics, we expect an

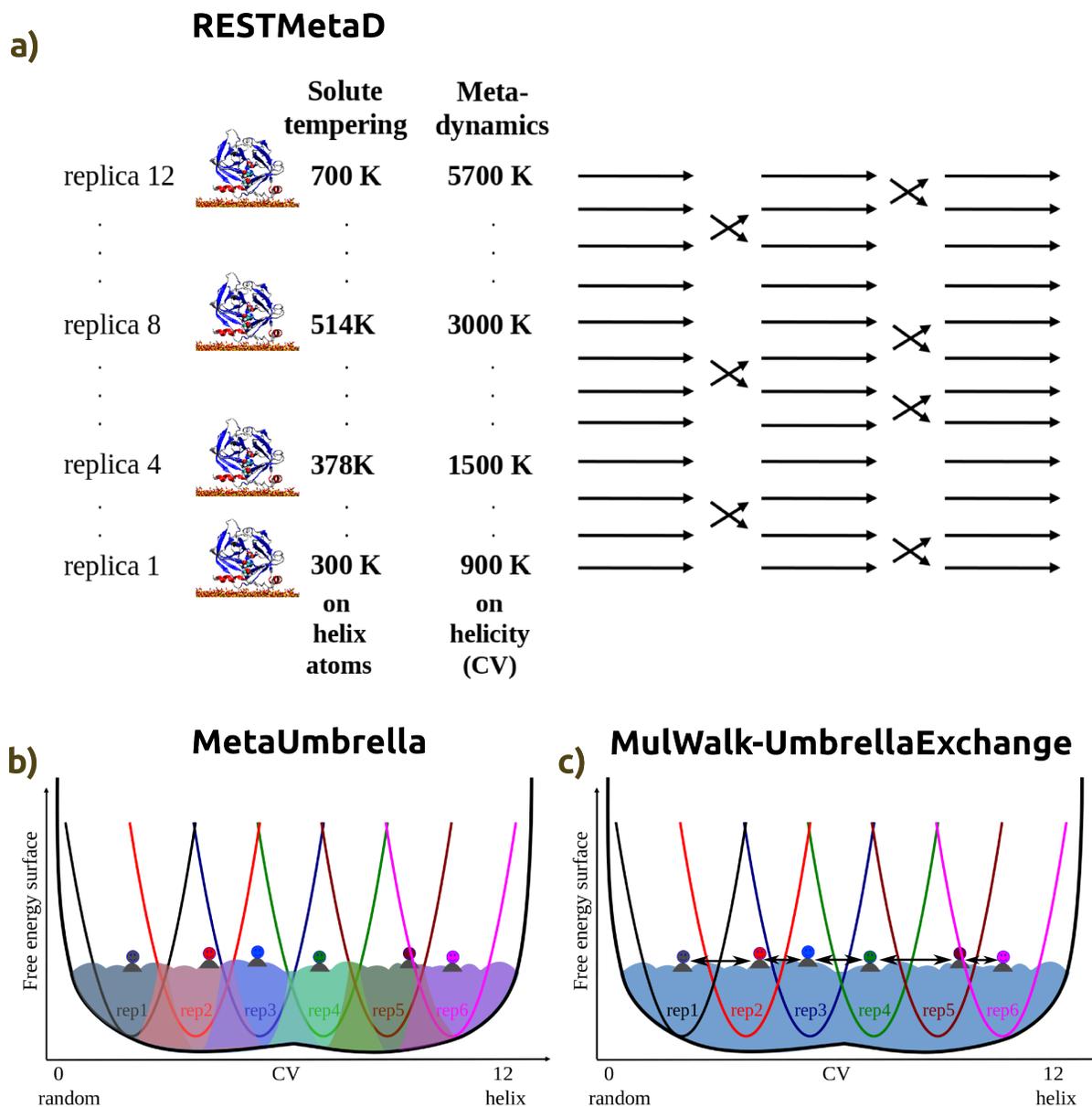


Figure 5.2: **a)** Replica Exchange Solute Tempering in combination with Metadynamics (RESTmetaD). The effective temperature of the 12 replicas span from 300 - 700 K. The Metadynamics boosting temperature on the helicity CV is chosen from 900 - 5700 K. The REST scaling is applied to the helix atoms. **b)** Umbrella sampling in combination with Metadynamics (MetaUmbrella). In our case, 6 out of 25 umbrellas are exemplarily shown. A Metadynamics simulation is applied on each of the 25 individual simulations. The colors mark the Metadynamics bias in each umbrella and the colored dots the current CV value. **c)** The same approach as shown in **b)** is extended by exchanges between the umbrellas and a uniform Metadynamics bias potential by the Multiple Walkers approach (MulWalk-UmbrellaExchange) [163, 175].

increased sampling, particularly at the umbrella edges. The force constant of the harmonic restraint is set to 40 kJ/mol. The Metadynamics simulation parameters consist of a hills width of 0.35 rad, τ is set to 3 ps and the bias factor γ is 4. The hills height is calculated by the formula given in 5.1.3. The deposition rate is 1 ps and exchanges between the replicas are not implemented. The resulting free energy profile is calculated by the weighted histogram method (WHAM) [174]. The bias at each frame is calculated by the sum of the total Metadynamics bias and the bias potential of the umbrella restraint at the CV point. The concatenated trajectories of all replicas are analyzed a posteriori by the driver tool implemented in PLUMED 2 [163].

5.1.5 Umbrella Sampling with Multiple Walker Metadynamics and Exchanges (MulWalk-UmbrellaExchange)

The MetaUmbrella simulation previously described is extended by two additional methods. First, all applied Metadynamics hills at all umbrellas are added up. Hence, further applied hills are added to a bias surface which is identical for all umbrellas. Each Metadynamics simulation at one umbrella is defined as one walker contributing to the same bias surface. The method is called Multiple Walker Metadynamics [175].

The second extension includes exchanges between the umbrella systems following the Metropolis-Hastings criterion similar to the replica exchange simulations (Appendix 7.5.5).

In all three presented simulations we have used the AMBER14SB force field [76] for the protein. The simulations are performed with the PLUMED 2 patch [163] on GROMACS 4.6.7 [164]. In particular, we have used the HREX branch v2.2 of PLUMED 2 [169] to include the Hamiltonian Exchange method (Appendix 7.5.5).

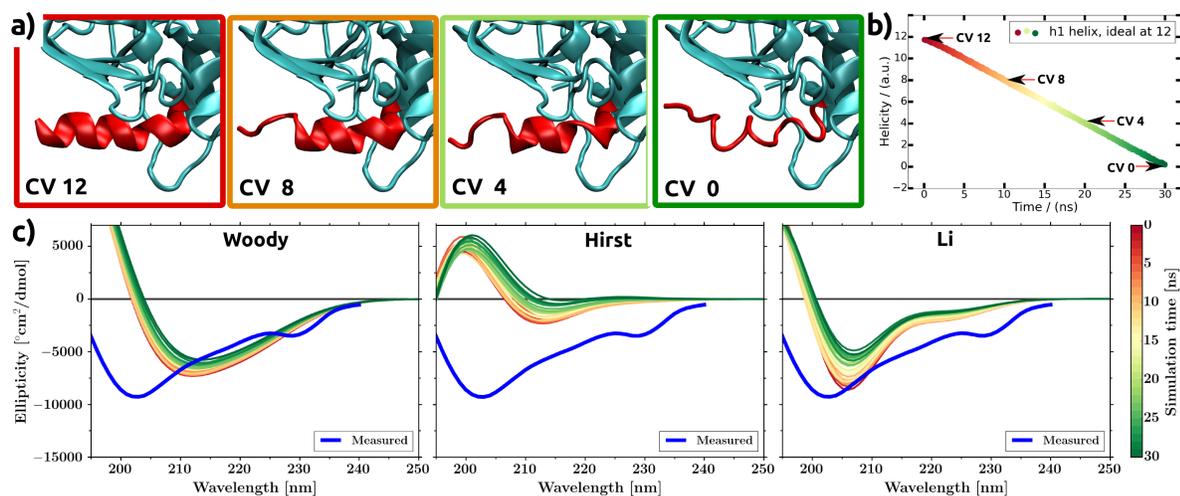


Figure 5.3: Moving restrain MD simulation with the helicity as the CV. The helicity of the helix of chymotrypsin at the C-terminus (shown as red ribbons) is continuously reduced from an ideal helix to a random structure, indicated by the color code from red to green in all subfigures. **a)** Snapshots of the trajectory at CV values of 0 (random structure), 4, 8 and 12 (ideal helix). **b)** The propagation of the CV value over the simulation time of 30 ns. **c)** Three different parameter sets are used to calculate the CD spectra by the atom positions of the trajectory [187–189]. The experimentally measured CD spectra of the solvated protein is shown in blue and taken from Fig. 1.9 [99].

5.2 Results

By using a moving restrain MD simulation, the helicity CV of chymotrypsin is constantly reduced from an ideal helix to a denaturated state as shown in the snapshots in Fig. 5.3 a) and b) (red to green). The atom positions of the trajectory are used for the calculation of the CD spectra using the DichroCalc web interface [176] with three different parameter sets by Woody et al. [187], Hirst et al. [188] and Li et al. [189] Fig. 5.3 c). In comparison to the experiment presented in Fig. 1.9 [99] the simulated results show that the broad spectrum minima between 200 and 230 nm can roughly be represented by the parameter sets of Woody and Li. This minimum is shifted by about 10 nm in the case of the Hirst parameters. The theoretical spectrum at 205 - 240 nm in all parameter sets moves to higher wavelength and smaller CD signals with decreasing helicities.

To investigate a more general conformational change of the protein, MD simulations at different thermostat temperatures are performed. In particular, we have started eight simulations at 300 K, 400 K, 500 K and 600 K in solution and adsorbed on SiO₂. As the used force field is not optimized for these high temperatures, the observed dynamics

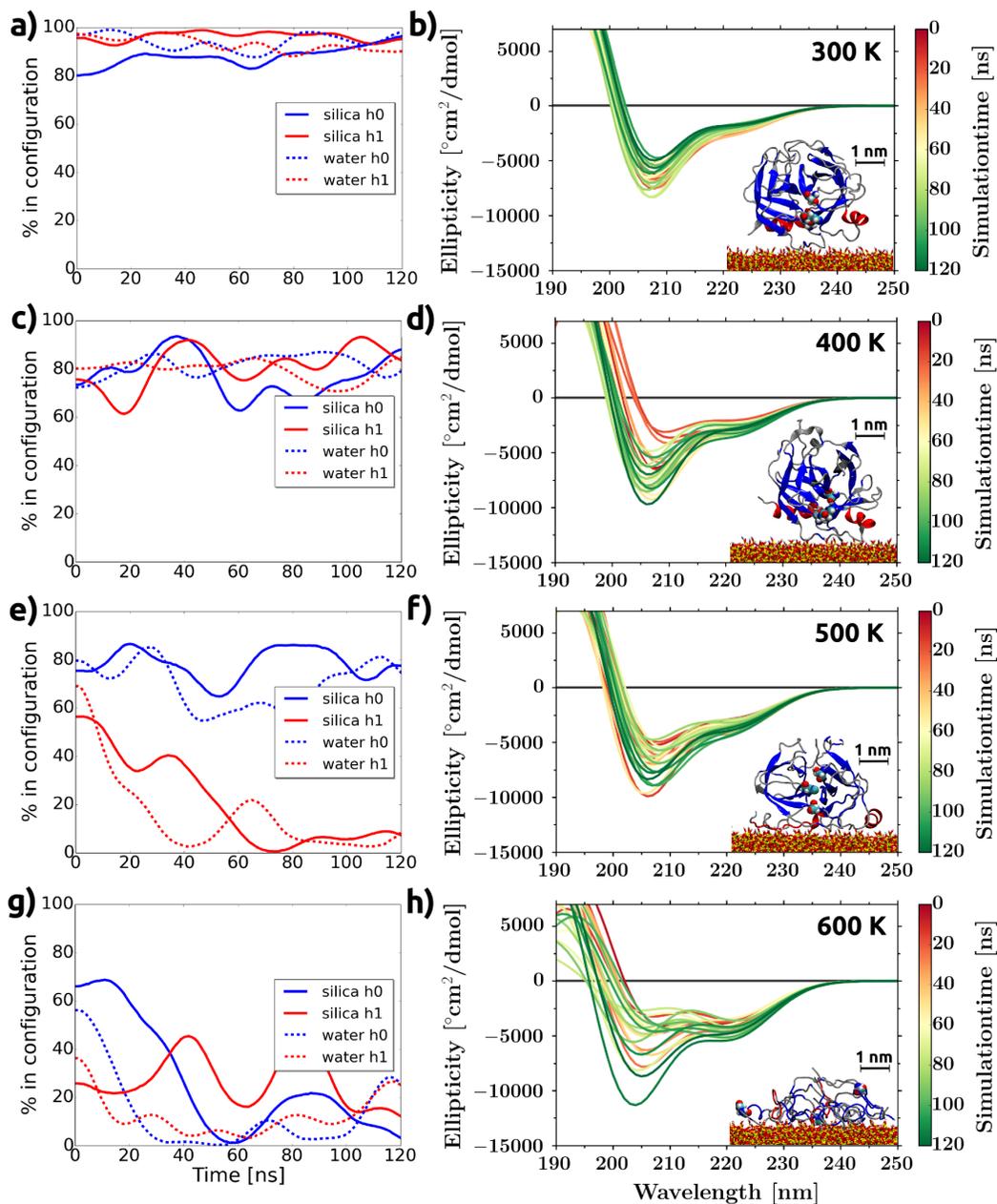


Figure 5.4: Simulated heating of chymotrypsin at 300 - 600 K in an NVT ensemble. (left column) An ideal helix structure of the two helices of chymotrypsin (h0 and h1) refers to 100% in the helix configuration. The curves are smoothed by a Hamming window function. The protein is investigated in solution and on SiO₂. (right column) The insets show the protein on SiO₂ after the simulation time of 120 ns. The shown CD spectra are calculated directly from the simulated frames using the DichroCalc with the Li parameter set [176, 189].

cannot be directly compared to the experiments. However, these simulations provide a simplified method to increase the system energy and overcome local barriers. The subsequent analysis identifies sections of the protein with a high instability or flexibility. The protein configuration at the end of the 120 ns MD simulation is shown in the insets in Fig. 5.4. At 300 K and 400 K no obvious conformational change can be observed in contrast to 500 K and 600 K, which show significant protein denaturations. We have further analyzed the amino acids of the embedded helix configurations. Chymotrypsin has two helix sections, one at the C-terminus (h1) and a second embedded in the protein shell (h0). In Fig. 5.4 the helicity in percentage of the predefined amino acids is analyzed. 100 % refer to an ideal helix structure of h1 and h0. The helicity is calculated by the collective variable (CV) given in Section 5.1.3.

The temperature of 400 K slightly reduces the helix content. At 500 K the helix h1 denaturates, whereas the other helix h0 remains stable. Furthermore, at 600 K both helices lose their secondary structure within 20 - 40 ns. No difference is observed between the content of the solvated protein in water and the protein at the surface (dashed and straight lines). Therefore, the presented simulations do not identify a surface-induced denaturation but reveal the increased instability of the helix h1 compared to h0.

The CD-spectra of the trajectories are calculated by the Li parameter set and DichroCalc [176,189]. No obvious trend of the CD spectra can be observed during the simulation time of 120 ns (Fig. 5.4). For higher temperatures and especially for 600 K an increased variation of the spectra is observed.

In the following, we are going to investigate the secondary structure of the isolated peptide which forms the h1 helix at the C-terminus of chymotrypsin. First insights into the secondary structure are calculated by several online server tools. The sequence of the amino acids 230 to 245 of chymotrypsin is used in the online tool BLAST [190] to find proteins in the protein data bank which contain the same sequence. All proteins with a match of more than 68.8 % can be assigned to trypsin or chymotrypsin and their precursors. The PEP-Fold secondary structure prediction tool is based on an ensemble of elementary prototype conformations [191,192]. It calculates a probability of at least 50 % of the amino acids 233 - 241 to be found in an alpha-helix. The QUARK prediction tool is based on an ab initio protein folding algorithm [193] without homologous templates and predicts a helix conformation except for the last two amino acids at the C-terminus. A similar result is calculated by the I-Tasser tool which uses a combination of structural

templates as well as template fragment assembly simulations [194].

The prediction servers are all based on the pure amino acid sequence and do not take environmental influences such as surface interactions into account. In the following, we present three types of free energy sampling methods which take these influences into consideration. The simulations predict the free energy landscape of the helicity used as the CV. One simulation is based on solute tempering in combination with Metadynamics (RESTmetaD, Section 5.1.3). The other two are based on Umbrella Sampling in combination with Metadynamics with and without exchanges between the umbrellas (MetaUmbrella, Section 5.1.4 and MulWalk-Umbrella Exchange, Section 5.1.5).

Each method is applied to three systems: (1) The helix peptide in solution, (2) the protein in solution and (3) the protein adsorbed on the SiO_2 surface. In all systems, the helicity of the C-terminus helix defined in Section 5.1.3 is used as CV.

(1) The helix peptide in solution

The RESTmetaD simulation is started in two different start configurations, in the native and a denaturated state as shown in Fig. 5.5 . The native state is directly taken from the PDB structure (ID: 4CHA) and has a rather high helicity value. In the denaturated state, the peptide forms a random coil structure and has a low helicity value. After about 200 ns both simulations start to approach the same free energy profile. It can be seen that the final free energy profile contains 13 local minima along the full CV range. The final free energy surface is rather flat with variations of about 12 kJ/mol. In both simulations a shallow global free energy minimum can be identified at a helicity value close to zero. The second, broader minimum between a CV value of 4 and 6 has only a free energy difference of about 1-5 kJ/mol to the global minimum. The sampling of the simulation is shown in Fig. 5.6 which indicates the mobility of the replicas along the CV ladder.

The rough trend of the rather flat energy surface of the RESTmetaD simulation is reproduced by the MetaUmbrella simulation (Fig. 5.5 c)). The first drawn free energy landscape after 12 ns shows a similar shape to the one after the performed simulation time of 127 ns. Especially a local energy minimum close to the CV value of zero followed by a maximum around two can be observed in Fig. 5.5. The start configuration for each of the 25 umbrella simulations is set to its umbrella minimum. The bias overlap of the umbrellas can be seen in the inset in Fig. 5.5 and in the histogram shown in Fig. 5.6. It can be seen that not all umbrellas have a Gauß-like distribution and sampling deficiencies are observed at CV values of about 0.5 and 11.5.

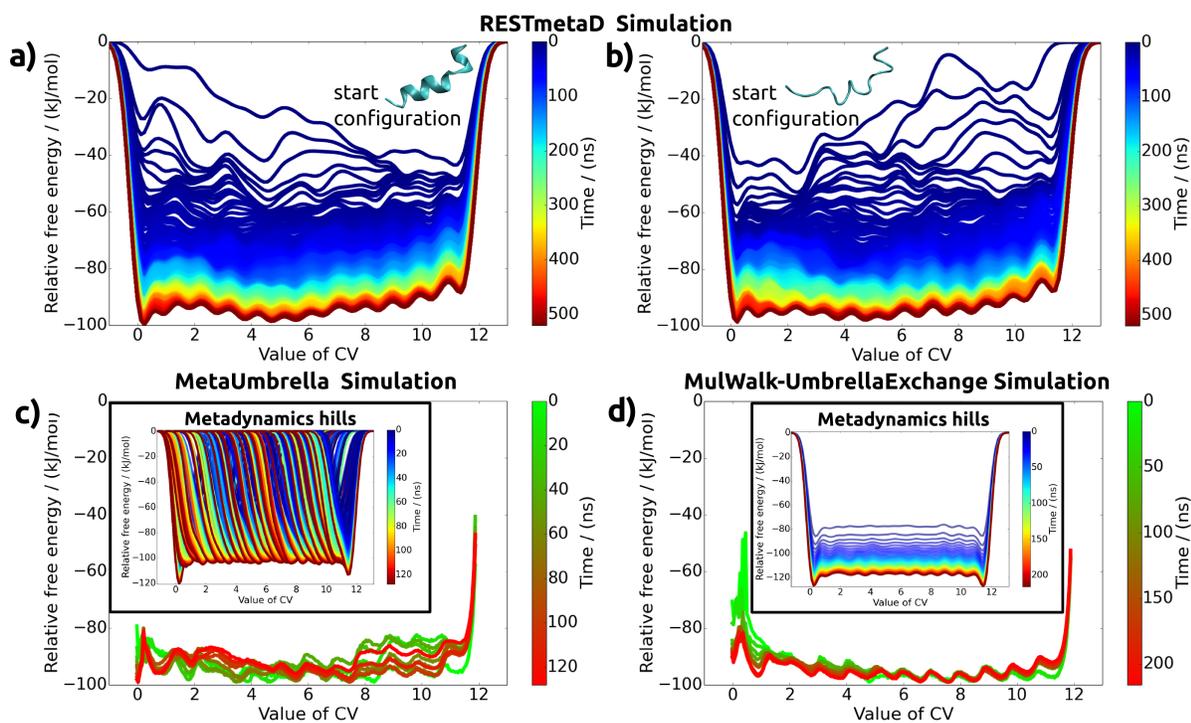


Figure 5.5: a-b) *RESTmetaD* free energy landscape of two different start configurations of the isolated helix of chymotrypsin including the residues 230 - 245 (shown in the inset). a) Ideal helix start configuration and b) a denatured configuration. The colored time evolution shows the free energy landscape. c) *MetaUmbrella* free energy landscape of the same helix. The time evolution is presented in 10 cumulative ranges, whereas the red curves show the complete simulation. The inset displays the sum of the deposited *Metadynamics* hills in free energy units for each of the 25 umbrellas similar to a) and b). d) *MulWalk-UmbrellaExchange* simulations of 25 umbrellas. Ten curves describe the time evolution of the simulation, similar to c). The inset shows the shared deposited *Metadynamics* hills in free energy units of all umbrellas.

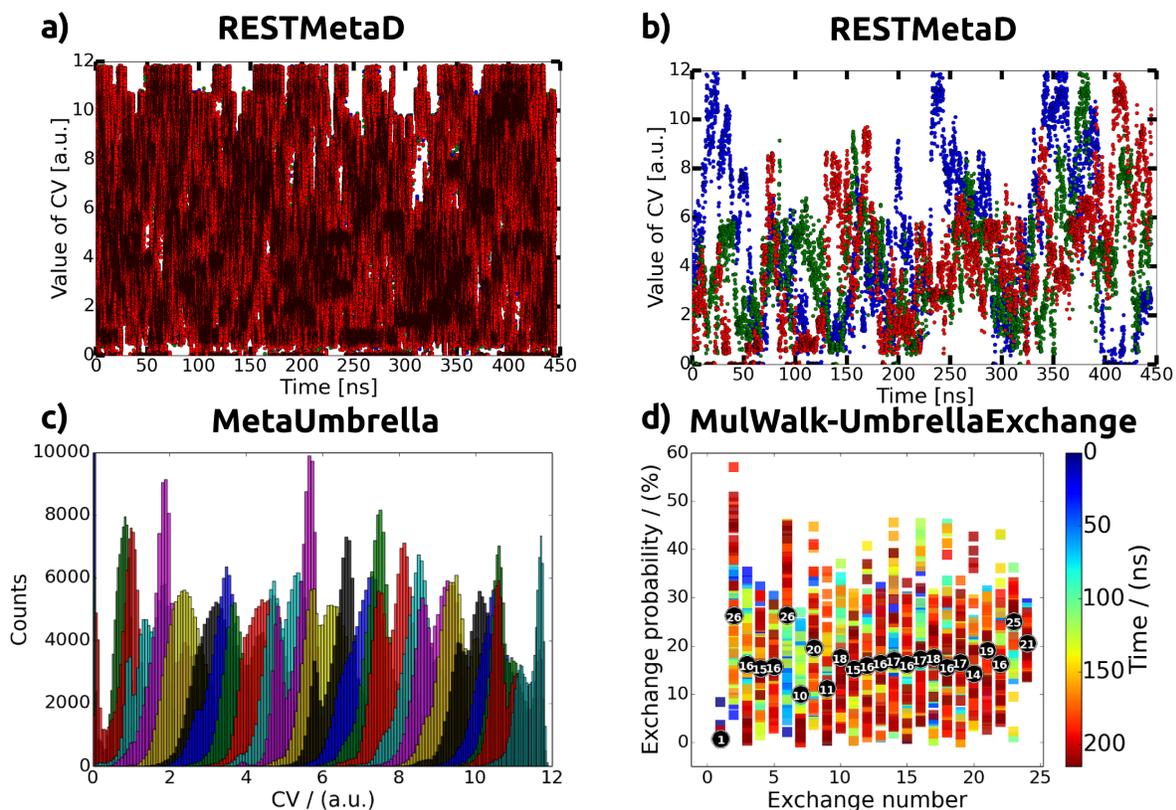


Figure 5.6: Performance analysis of the helix peptide simulation shown in Fig. 5.5 . **a)** Value of the CV during the RESTmetaD simulation. Each dot marks one frame and each color one trajectory (blue, red, green). **b)** Three arbitrary trajectories are sorted by the DEMUX tool included in GROMACS [164] and shown in blue, red and green. **c)** Histogram of the MetaUmbrella simulations. Each color shows one of the 25 umbrellas. **d)** Exchange probabilities of the MulWalk-UmbrellaExchange simulation for each of the 24 exchanges. The color code represents the time evolution. The gray dots indicate the average with its value written inside the dot.

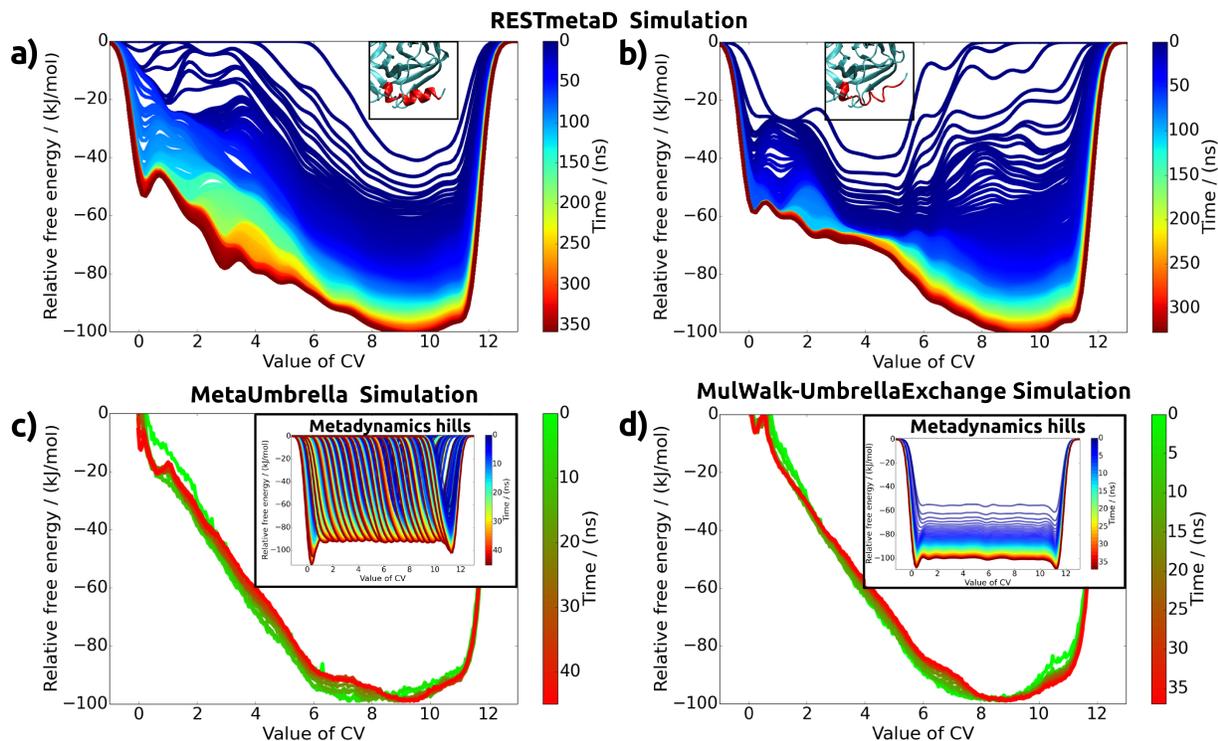


Figure 5.7: Same as Fig. 5.5 but with the complete protein solvated in water. The same helicity collective variable is used for the helix at the C-terminus of chymotrypsin.

The MulWalk-UmbrellaExchange simulation shows a similar free energy surface (Fig. 5.5). The free energy landscapes at the different time periods vary by at least 5 kJ/mol except around a CV value of 0. The reduced convergence around 0 is the result of a reduced sampling in this region (this will be discussed in Section 5.3). The free energy landscape measured after 60 ns is already very close to the final shape observed after 217 ns. The 13 local minima along the full CV range are well resolved, similar to the RESTmetaD simulations. Since a Multiple Walkers approach is used in this simulation, all umbrella simulations share the same Metadynamics bias history. Therefore, the deposited bias hills potentials shown in the inset in Fig. 5.5 form one curve in each shown time period and grow rather parallel over the simulation time. The shape shows the 13 small hillocks similar to the resulting free energy profiles again. At the edge of the bias hills potentials at a CV value of 0 and 12 the free energy surface shows a prominent minimum similar to the inset in Fig. 5.5 c). In the presented exchange probabilities of the simulation shown in Fig. 5.6 a very low exchange rate is observed for the first two replicas.

(2) The protein in solution

The RESTmetaD simulation of the whole protein in water is started in two configurations, similar to the presented pure helix peptide simulations. One simulation is started with a high and one with a low helicity value of the helix at the C-terminus of the protein. The insets in Fig. 5.7 show the part of the protein containing the helix. Both free energy landscapes have a very different shape compared to the helix alone (in comparison to Fig. 5.5). The less prominent parallel growing of the time evolution of the free energy landscape indicates a less ergodic sampling. Both landscapes have a broad minimum at a CV value of about 9 which fits well with the calculated CV value of the used PDB structure (ID: 4CHA). The sampling ability of this RESTmetaD simulation is shown in Fig. 5.8 a) and b). In comparison to the sampling of the solvated helix peptide simulation shown in Fig. 5.6 a) and b), a less homogenous sampling along the collective variable is observed. Especially, the DEMUX analysis in Fig. 5.8 b) reveal longer periods of constant CV values.

The free energy surface of the MetaUmbrella simulation converges faster than in the helix peptide simulation. After 15 ns a very smooth free energy surface is calculated. The comparison of the MetaUmbrella with the RESTmetaD simulation reveals a similar trend with a global free energy minimum at a CV value of about 9. The deposited bias potentials of the Metadynamics simulation overlap well as shown in Fig. 5.7 c) inset) and Fig. 5.8 c).

A comparable time evolution and shape of the free energy landscape is found by the MulWalkUmbrellaExchange in a total simulation time of 37 ns (Fig. 5.7 d)). The bias potential in free energy units grows homogeneously over time as presented in the inset. Similar to the insets in Fig. 5.5 c) and d), the edges in this inset have a prominent free energy minimum. Small hillocks along the full CV range are not present, as shown in Fig. 5.7 a) to d) and in the insets. The exchange probabilities as presented in Fig. 5.8 d) are similar to the probabilities observed for the helix peptide in Fig. 5.6 d).

A separation in hydrophobic and hydrophilic residues of the embedded helix and its surrounding amino acids is shown in Fig. 5.9. The majority of the hydrophobic amino acids of the helix point towards the inside of the protein, whereas a high concentration of hydrophilic amino acids is facing the outside (Fig. 5.9 a)). Therefore, the outer hydrophilic amino acids are accessible to the solvent or potentially a surface. The strong hydrophobic embedding is visualized by a LigPlot [195] interaction scheme in Fig. 5.9 b). The green atoms represent the helix atoms, purple lines show hydrogen bonds and hydrophobic

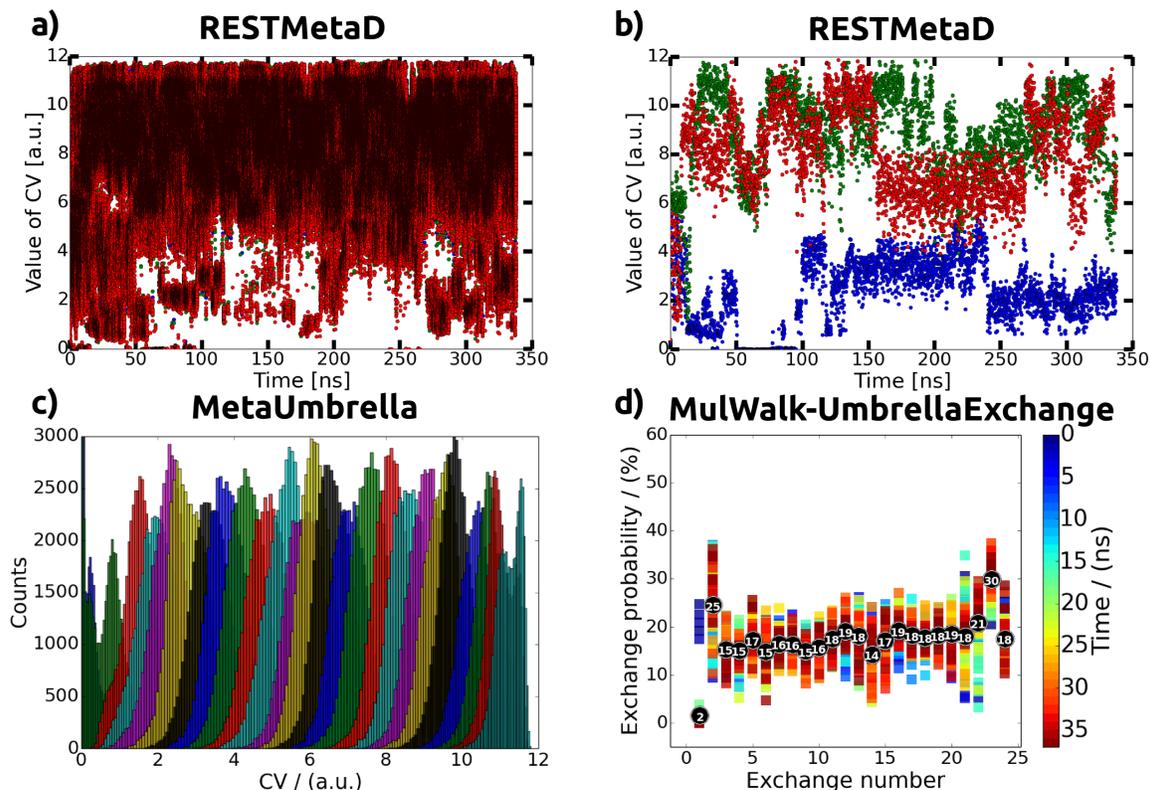


Figure 5.8: Analysis of the free energy simulations of the protein solvated in water shown in Fig. 5.7, identically presented as in Fig. 5.6.

interactions are presented as red circles. The hydrogen bond formation is investigated in more detail in subfigure c) - h) by using the RESTmetaD trajectory shown in Fig. 5.7 a). The number of hydrogen bonds between the helix and the rest of the protein is calculated by the Hbond tool of VMD [114]. An H-bond is defined if an H-bond donor atom (bound to a hydrogen) has a distance smaller than 3 \AA to an acceptor atom and the angle between donor, hydrogen and acceptor atom is between 160° and 180° . Helix internal hydrogen bonds are not included. Each frame contains between 0 and 5 hydrogen bonds. The number of bonds is increased for lower CV values, as displayed by the distribution of the blue dots which mark the number of the bonds in Fig. 5.9 c) - e). Furthermore, the average number of bonds decreases from lower CV values to higher CV values, shown by the red line. The hydrogen bonds and the helix structure of all frames are shown superimposed as blue lines and red ribbons in subfigure f) - h). The protein of the first frame is presented in turquoise as a guide to represent the protein structure. At high CV values, the nearly ideal helix has a large number of internal hydrogen bonds but is located

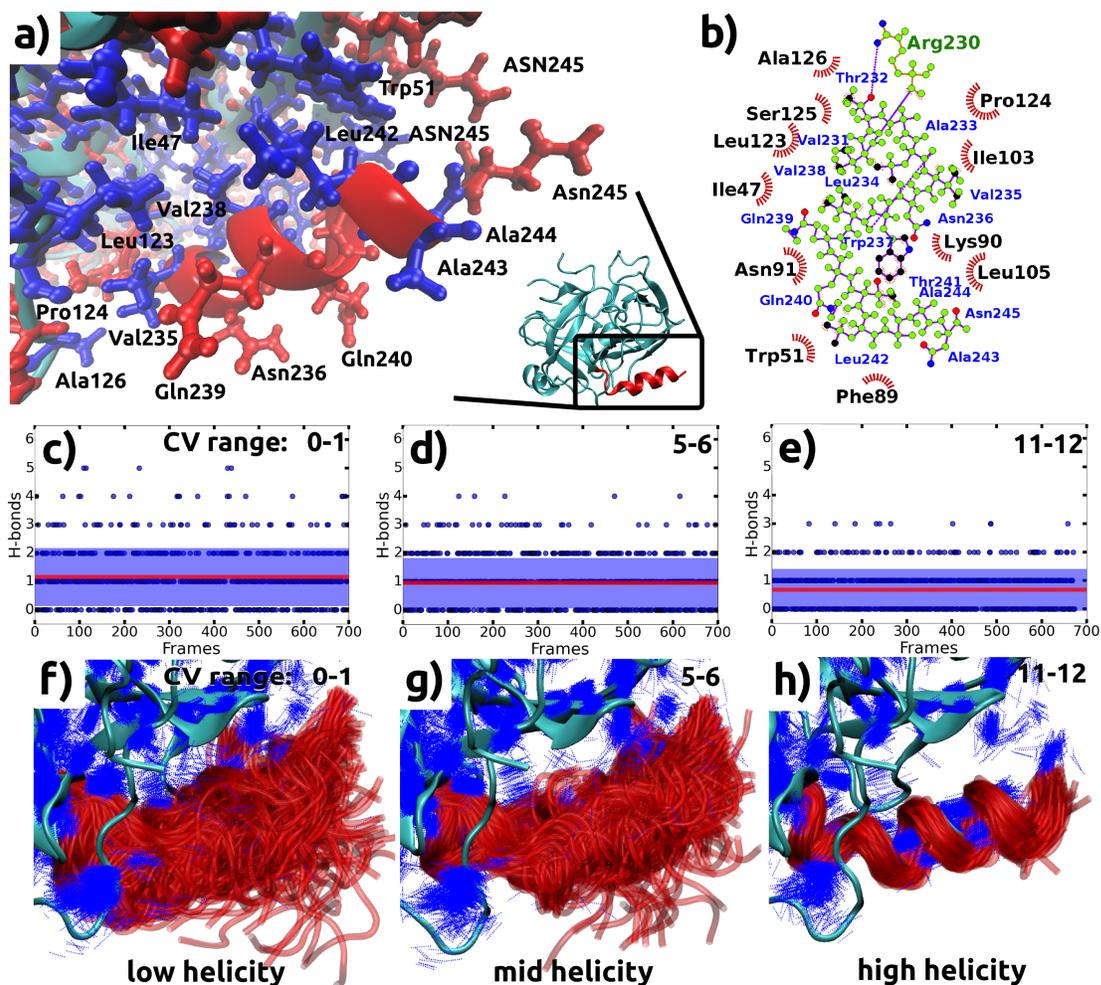


Figure 5.9: Embedding and intra-protein interactions of the investigated helix of chymotrypsin. **a)** The inset shows the location of the helix (red coil) in the protein (turquoise ribbon). Hydrophobic residues are shown in blue and hydrophilic residues in red. **b)** A LigPlot [195] interaction scheme of the helix and surrounding protein. Covalent bonds are shown by straight purple lines and hydrogen bonds by the two dashed lines. Hydrophobic interactions are shown by red circles. Helix residue names and numbers are blue and black for the surrounding protein. **c) - e)** The REStmetaD trajectory shown in Fig. 5.7 a) is cut in three ranges of the CV 0-1, 5-6 and 11-12. In each range the number of H-bonds between helix and protein is computed by VMD [114]. The red line shows the H-bond average values of range 0-1: 1.15 ± 1.02 , range 5-6: 0.95 ± 0.87 and range 11-12: 0.68 ± 0.73 . **f) - h)** The hydrogen bonds (thin blue lines) and the structures (red ribbons) of all frames are shown superimposed.

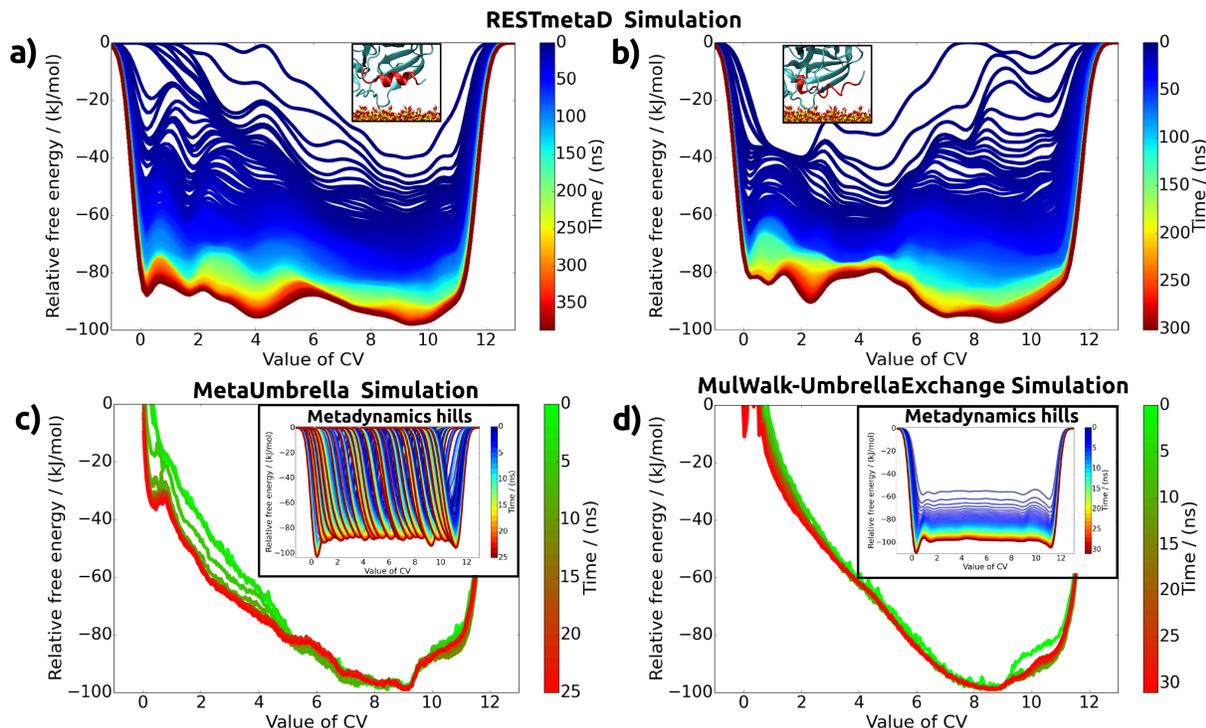


Figure 5.10: Same as Fig. 5.5 and 5.7 for the adsorbed protein on SiO_2 .

further away from the protein and forms fewer bonds to the protein. For lower CV values the helix amino acids can attach along the complete length of the helix sequence to the protein and thus form more hydrogen bonds.

(3) The adsorbed protein on SiO_2

In the following, we are going to include the amorphous SiO_2 surface to investigate whether the distortion induced by the surface compensates the helix embedding and changes the free energy surface of the helix. As a start configuration, the chymotrypsin structure of the adsorption simulation in explicit solvent described in Chapter 2 is used. In a short target MD simulation one configuration of a native and one of a denatured helix are produced. Similar to the previous simulations the two configurations are compared in the RESTmetaD simulations (Fig. 5.10 a) and b) insets). Both RESTmetaD simulations result in an energy minimum at CV values around 9-10 and an overall rather flat free energy landscape of the width of 20 kJ/mol, after a simulation time of 350 ns.

The metaUmbrella simulations reveal a different result with a more steep free energy landscape. At very low CV values close to zero the free energy landscapes are still evolving

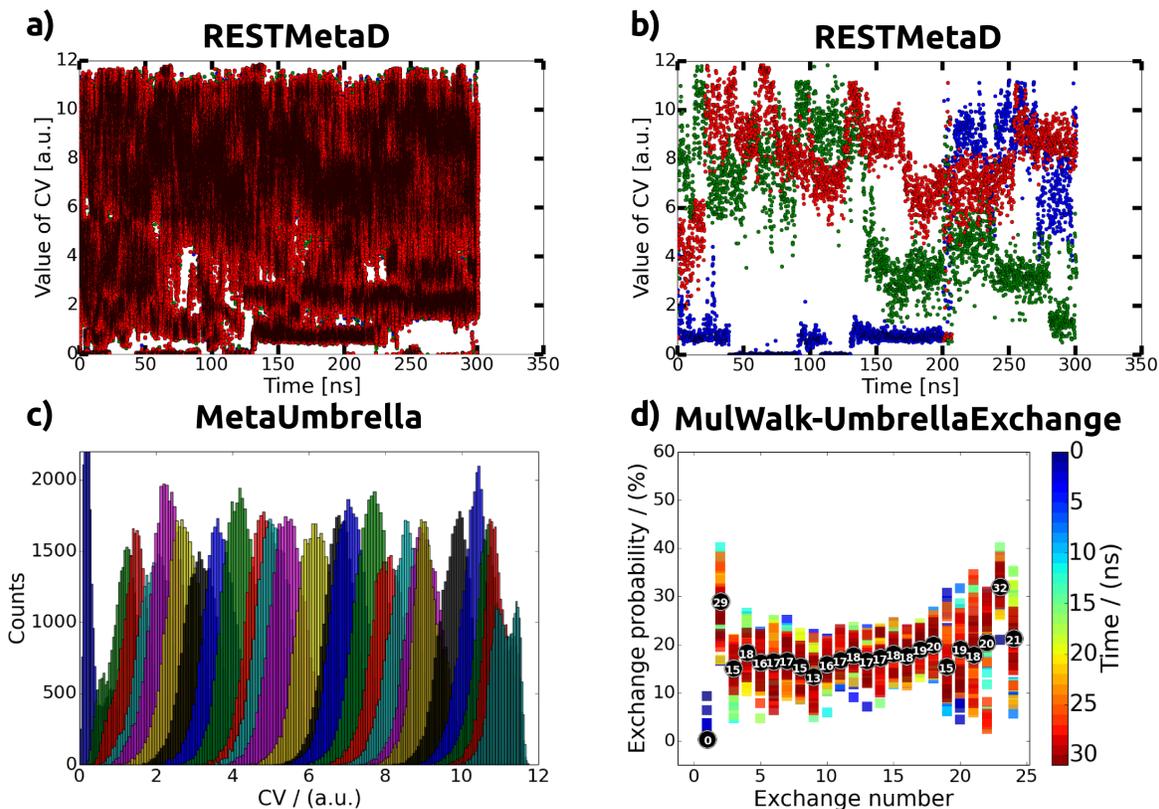


Figure 5.11: Analysis of the free energy simulations of the adsorbed protein shown in Fig. 5.10, identically presented as in Fig. 5.6 and 5.8.

after 44 ns. A broad free energy minimum can be found around a CV value of 8. Compared to the simulation of the solvated protein in Fig. 5.7 c), a kink at the CV value of 10 appears in the free energy landscape.

The MulWalk-UmbrellaExchange simulation shows an even more smooth and extremely quick convergence of the free energy surface compared to the MetaUmbrella calculation (Fig. 5.7). The free energy landscape is already very similar to the final shape after 5 ns compared to the total time of 31 ns. No periodic small hillocks along the CV are observed which appear in the helix peptide simulations shown in Fig. 5.5.

The Umbrella simulations of all three systems show an increased deposition of the Meta-dynamics bias hills at the CV value edges shown in the insets in Fig. 5.5, 5.7 and 5.10. The sampling ability of the adsorbed protein simulations shown in Fig. 5.11 are very similar to the simulation of the solvated protein in Fig. 5.8.

A cluster analysis of all presented free energy simulations is shown in Fig. 5.12 to identify patterns in the configurations visually. The trajectories of all replicas and umbrellas are cut by the CV value of the helix peptide in the ranges 0-1, 3-4, 5-6, 8-9 and 11-12. The helix peptide of those ranges is aligned by a rigid rotation and translation to minimize the RMSD variation. The GROMOS algorithm [196] with an RMSD cutoff of 0.3 nm is used. The clusters are subsequently formed by the RMSD value differences of the individual frames (Appendix 7.8). In the case of the protein simulations, the helix peptide is isolated, analyzed and shown alone.

The clusters of the RESTmetaD helix peptide simulations have a hairpin structure in the CV range of 0-1 and form a more horseshoe-like structure for higher CV values. At values of 8-9, a straight configuration is observed. Due to the inhomogeneous sampling of the entire CV range, each presented range has a different number of frames, up to 4636 in the case of the RESTmetaD simulation. The fraction of the frames which contribute to the cluster is less for smaller values of the CV value, as shown by the declared percentages. The helix formation starts in the middle of the second half of the peptide including the C-terminus and progresses to both termini for higher CV values. The N-terminus can be identified by the Arg230 residue which is shown in blue in Fig. 5.12.

Both Umbrella based simulations show very similar clusters. In all simulations, the helix starts to fold in the second half of the peptide. The hairpin structure observed at a CV value close to zero of the RESTmetaD simulation is transformed to a more horseshoe like structure in both umbrella based simulations. A rough similarity of the percentage of frames which contribute to the clusters is observed along one CV range of all three methods for the helix peptide.

The clusters of the solvated protein are different from the solvated helix peptide as shown in Fig. 5.12 b), e) and h). An overall more straight configuration is observed due to the embedding of the helix in the protein. At low CV values, however, the clusters have a more kinked form. The region around the N-terminal group does not change its configuration significantly. On the other hand, the free C-terminus has a loose tail and varies in its structure within one cluster. Due to the bond to the remaining protein, the first loops of the helix appear rather stable and the helix formation starts from the N-terminus to the C-terminus. Whereas in the helix peptide simulation the helix has started to fold from the middle of the molecule. The percentage of frames that contribute to the clusters is much higher than in the solvated helix simulation which identifies a

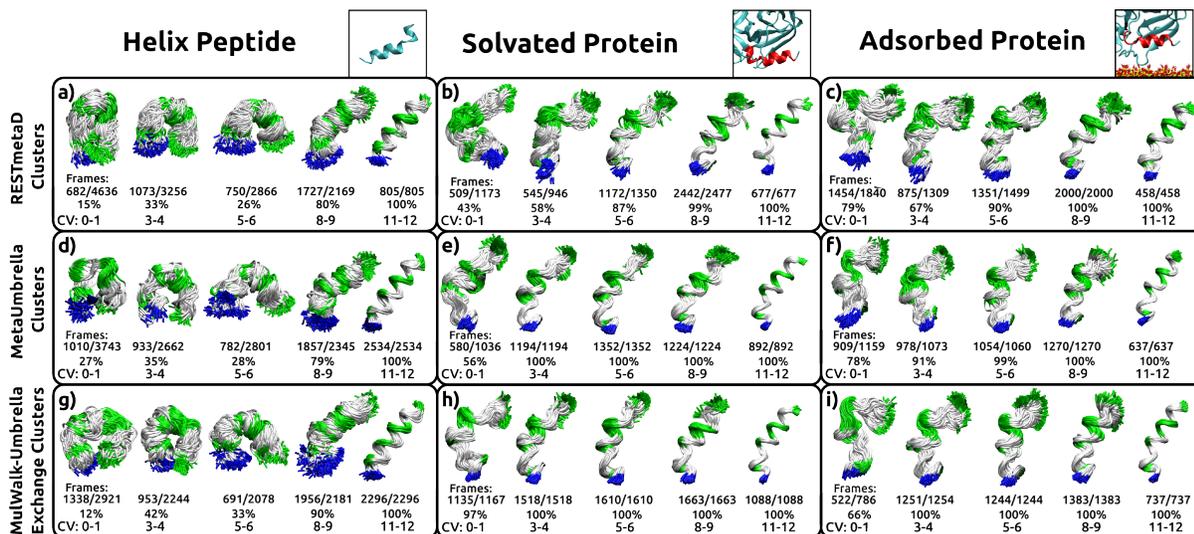


Figure 5.12: Cluster analysis of all three systems and the three used free energy simulation techniques. In each case of these 9 simulations, five ranges of the CVs are shown (see bottom lines). In the case of the solvated and adsorbed protein simulations only the helix is shown and aligned separately. The frame ratio and percentage shown under the visualized clusters represent the number of frames which contribute to the cluster in relation to the available frames of this CV range. The residues of the clusters are shown in the following colors: polar residues (green), hydrophobic residues (gray) and positively charged residues (blue, here only Arg230 which identifies the N-terminus of the helix).

less diverse structure ensemble of the trajectories. This effect is more prominent for the Umbrella based simulations than for the RESTmetaD simulation which seems to have a more diverse sampling.

The cluster analysis of the simulations of the adsorbed protein is shown in Fig. 5.12 c), f), i). Compared to the simulation of the solvated protein no obvious differences can be observed except that the structures within each cluster are more diverse. This effect becomes apparent especially for the CV values of 3 - 6 identified by the percentage of frames that contribute to the clusters.

The presented clusters describe the structural behavior of the solvated helix peptide in water directly. For the other two systems, the helix peptide is embedded in the protein and interacts with the protein or the surface during the simulation. These interactions are not visualized in Fig. 5.12, which is why the entire trajectory of the helix peptide is shown superimposed in Fig. 5.13. The protein in the top row is aligned to its start configuration by a rigid rotation to the RMSD difference minimum. In the case of the adsorbed protein, no alignment is applied. The visualized protein and surface are arbitrarily taken from one

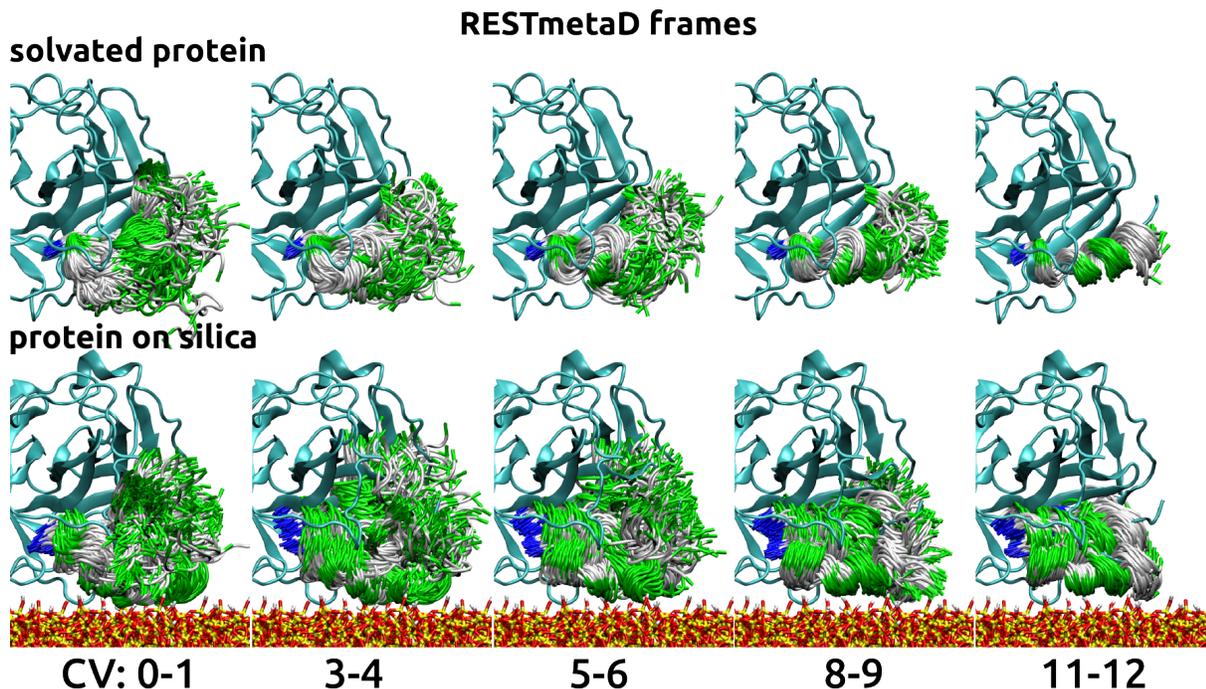


Figure 5.13: Superposition of all helix configurations of all frames of the RESTmetaD simulation. One arbitrary protein structure is presented in turquoise to indicate the position of the superimposed frames of the helix. Shown is the simulation of the protein in water (top) and the protein adsorbed on the surface (bottom) sorted according to their CV value (columns). Each subfigure contains about 1000 structures of the helix peptide colored by the residue type. The N-terminus can be identified by the Arg230 residue shown in blue.

frame to visualize the position of the superimposed helix peptide frames. Identical to the cluster visualization, the orientation of the helix peptide can be identified by the N-terminus Arg230, which is shown in blue.

The loose C-terminus end attaches frequently to the outside of the protein as shown in Fig. 5.13. It seems that the attachment to the protein is preferred to the solvated state of the C-terminus. The helix formation starts at the highly embedded N-terminus and for low CV values, the cluster analysis shows a boomerang-like shape as also presented in Fig. 5.12. It can be assumed that both ends of this boomerang are facing to the top and attach to the protein while the bow like shape is exposed to the solvent.

The trajectory of the adsorbed protein shows a very similar character (Fig. 5.13). For high CV values of 11-12 about two clustered configurations are observed, one is apparently adsorbed at the surface and the other is located above the surface. With decreasing CV values, the helix denaturation starts from the C-terminus as well. Remarkably, half the

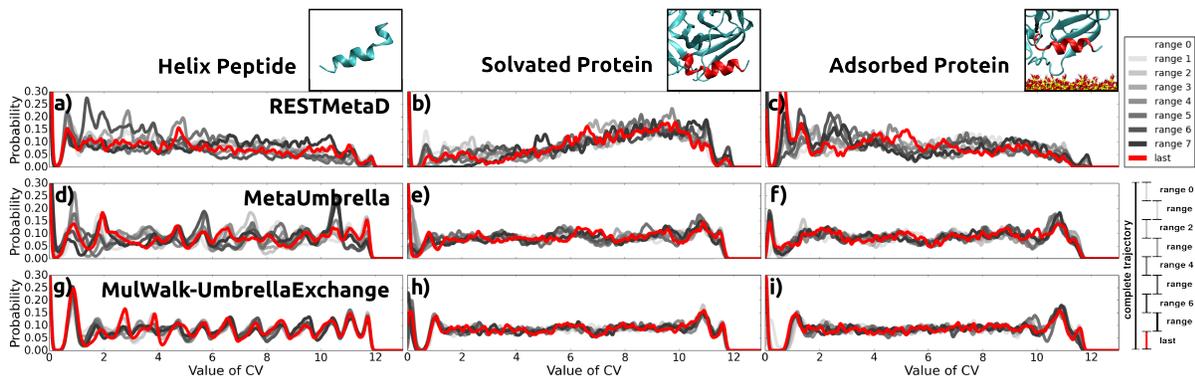


Figure 5.14: Histograms of the three used free energy methods and the three systems. For the calculation of the histograms all replicas in the case of the RESTmetaD simulations and all umbrellas in case of the umbrella based simulations are used. The histograms are calculated in separated and not cumulative periods of the trajectory as shown at the bottom right by the illustrated ranges in different gray scales. The very last period is shown in red. 520 bins are used to calculate the histograms along the CV range of 0 - 13, which are visualized by a Gaussian kernel density estimation with bandwidth of 0.05.

helix peptide attaches to the protein and does not adsorb to the surface for values from 5 to 6. This behavior continues for lower CV values which suggests an energetically preferred adsorption to the protein rather than to the surface. Similar to the solvated protein simulation, the clusters show a boomerang-like shape with the ends in contact to the protein whereas the bow like shape attaches to the surface.

All three free energy methods are compared by their sampling ability along the entire CV range by a histogram analysis shown in Fig. 5.14. The histograms represent the amount of visited CV values by using all replicas or umbrellas of the free energy simulations. Therefore, all available frames of the trajectories are considered. We have chosen to calculate the time evolution of the histograms in separated time periods of the trajectory (see sketch at the right of Fig. 5.14). That way the convergence of the simulation to a homogeneous sampling can be visualized. This homogeneous sampling is expected for Well-Tempered Metadynamics based simulations as included in all three used free energy methods.

The histogram of the RESTmetaD simulation of the helix peptide shows a rather homogeneous sampling with a tilted distribution and slightly increased probabilities for lower CV values. Furthermore, a reduced sampling at a value close to zero is present for all used systems and simulation techniques. A very high histogram count is observed at exactly zero and a reduced count close to the CV value of 12. In the umbrella based simulations

in Fig. 5.14 d) and g) an increased sampling of the 13 hillocks along the CV range can be seen as similarly observed in the free energy profiles in Fig. 5.5. The histogram of the MulWalk-UmbrellaExchange seems more homogeneous compared to the MetaUmbrella simulation.

The reduced sampling for lower CV values in the RESTmetaD simulation of the solvated protein is well visualized in the histogram shown in Fig. 5.14 b). Similar to the final free energy profiles of the system including the protein, the hillocks along the CV range are not visible in the histograms.

The RESTmetaD simulation of the adsorbed protein shows a slightly more homogeneous sampling compared to the same simulation of the protein alone. Still, it seems that the system is stuck at certain CV values of about 1 and 1.5 with increased histogram counts in this regions. The umbrella based simulations reveal a very similar sampling histogram compared to the umbrella simulations of the solvated protein without the surface. The overall comparison of the sampling histograms in Fig. 5.14 show a very homogeneous sampling of the MulWalk-UmbrellaExchange simulation, especially in comparison to the RESTmetaD simulation.

5.3 Discussion

In this chapter, the surface-induced change of the helicity of chymotrypsin on SiO_2 is investigated by using three free energy sampling methods. For this purpose, we have extended and adapted those methods to the case of protein adsorption. In this section, the newly presented free energy sampling methods are discussed first, followed by the discussion of the results of the investigated system.

5.3.1 Free energy MD simulations of proteins

Free energy methods such as Metadynamics, Solute Tempering and Umbrella Sampling have been widely used for small solvated molecules [172, 185, 197]. The application of these methods to proteins or parts of proteins as newly presented in this chapter contains additional challenges. The strong embedding of the region of interest into the protein requires the used advanced free energy sampling methods. By using these methods, the free energy surface of a part of a protein can be predicted. However, the results of this chapter reveal that these methods have some limitations as discussed in detail in the following.

Convergence of RESTmetaD simulations

A necessary criterion of the convergence of the RESTmetaD simulation is a frequent visit of the system at all ranges of the CV. If the system cannot enter or rarely visit certain regions in the CV space, the calculated free energy surface is underestimating these regions. A hint for a converging simulation is a parallel growth of the free energy surfaces along the timeline as shown by the color code in Fig. 5.5. A very similar indicator is the variation of the CV value at the ground replica, as presented in Fig. 5.6 a). In the case of the helix simulations, both indicators show a homogenous sampling suggesting a converged simulation.

Another indicator for the sampling ability is the capability to cross barriers along the CV. To identify the real performed transitions, the replica exchanges are subtracted out of the written trajectory. The newly created trajectory contains one replica which has traveled up and down the replica ladder but contains no exchanges. The tool to sort the trajectory in this way is called DEMUX included in the GROMACS package [164]. If

this sorted trajectory can still cross all barriers in the CV space, it is very likely that the system could converge easily. This trajectory is shown in Fig. 5.6 b) for three exemplary replicas. It can be seen that the sampling is much less broad and that some replicas do not cross along the entire range of the CV space.

Convergence of MetaUmbrella simulations

An important aspect for the convergence of the MetaUmbrella simulation is the overlap of the chosen umbrellas as shown in the inset in Fig. 5.5 and in the histogram in Fig. 5.6. If huge and narrow barriers exist in the CV range, the simulation avoids these regions and a sampling gap appears. These gaps can cause inhomogeneities of the post processing WHAM algorithm while combining the umbrella replicas. The independent Umbrella Samplings allow for an easy integration of further umbrellas a posteriori to fill possible gaps.

Furthermore, the additionally applied Metadynamics compensate small – and if the CV temperature is wisely chosen, it could also compensate bigger – energy minima and drive the system to the very edge of the umbrella restrain. This compensates a possible reduced sampling at the umbrella edges, which is a general weakness of standard Umbrella Samplings.

To visualize the convergence, the simulated trajectory is split in ten parts with increasing length. It can be seen in Fig. 5.5 that even after a relatively short simulation time of about 10 ns the trend of the free energy surface can be estimated. The fast convergence could be a hint that the system is trapped in one specific configuration within the small umbrella window. This particular configuration is not necessarily the only configuration at this CV value. The possible states increase rapidly for small helicity values. At this small values many states have the same CV value. This issue is not distinguished by the used CV. This effect can be seen in the low frame ratios contribution to the clusters for low CV values (Fig. 5.12). For high CV values only one cluster is found which contains all frames, whereas for low CV values many smaller clusters are identified. Additionally, this value is higher for the umbrella based simulation compared to the RESTmetaD simulation (Fig. 5.12). Hence, in the case of the umbrella simulations, the trajectory of each umbrella does not vary much in its configurations. This implies a reduced sampling of the configurations of the umbrella based simulations in comparison to the RESTmetaD simulations and motivates the additional inclusion of the exchanges to increase ergodicity.

Convergence of MulWalk-UmbrellaExchange simulations

The MulWalk-UmbrellaExchange simulations seem to converge rather quickly since no significant changes of the Metadynamics bias potential are observed after 40 ns (Fig. 5.5). The additional exchanges compared to the MetaUmbrella lead to an increased sampling and an even faster convergence as shown in the histograms in Fig. 5.14. Remarkable is an increased hills deposition at the CV edges of the Metadynamics hills as shown in Fig. 5.5 d) in the inset. It is expected that this is caused by an edge effect solely based on the fact that the center of the edge umbrellas at the CV values 0 and 12 have only half the CV space available compared to the other 23 umbrellas in the middle of the CV range. Due to the reduced CV space available and the rather high deposition frequency and hills height, an increased total bias potential is observed. This effect can be seen along all investigated systems but is compensated by the WHAM method in the calculation of the final free energy surface, since it adapts an offset between the umbrella windows.

The sampling in the MulWalk-UmbrellaExchange simulations should be enhanced due to two reasons. First, after each exchange the system is confronted by a bias of another region of the CV range. This bias can disturb the actual configuration and increase the configurational sampling. Due to the Multiple Walkers approach, this effect is fairly small because the bias distribution along the CV range is rather smooth. A more significant effect is based on the new umbrella restraint in which the system enters after each exchange. After each rapidly occurring exchange, the system is confronted with the neighboring umbrella restraint, pulling on the CV value of the system. Both these aspects lead to a sampling of a variety of conformations along the full CV range, as observed in the histograms of the CV values shown in Fig. 5.14.

The Choice of the Collective Variable

All three free energy profiles of the solvated helix peptide contain 13 small energy minima along the profile, which vanish in the simulations of the whole protein (Fig. 5.5, 5.7 and 5.10). These hillocks are caused by hydrogen bonds formed within the helix, as also observed by Meißner et al. [123] . It is expected that not only the internal hydrogen bonds influence the structural formation but also the embedding of the helix in the protein, which involves strong hydrogen bonds and hydrophobic interactions with the surrounding. Therefore, these hillocks are only present for the isolated helix in which the internal

hydrogen bonds dominate, and disappear in the free energy profiles of the protein due to the increased embedding of the helix.

Due to this strong embedding, the sampling ability of the protein simulations are significantly reduced compared to the solvated helix peptide simulations which contain a relatively homogeneous sampling. As the chosen CV does not influence those interactions directly, high bias energies are necessary to break the associated bonds. Thus we have introduced rather high CV temperatures up to 5700 K of the Metadynamics simulations and effective temperatures of the solute tempering of 700 K at the highest replica. Especially the solute tempering approach decreases the interactions between the helix and its surrounding, which is comparable to a CV which uses contacts or hydrogen bonds. Additionally, the MetaUmbrella simulation reduces the CV space and simplifies the sampling supported by the additional Well-Tempered Metadynamics acting only on the reduced CV space. Despite these high temperatures and the reduced CV space a reduced sampling of the protein based simulations is observed, thus emphasizing the strong and complex embedding.

The sampling ability of the three free energy methods and used systems is shown by the histograms of all visited CV values (Fig. 5.14). An incomplete sampling appears at a value close to zero across all histograms, which is possibly caused by the chosen helicity CV. This effect comes along with a very small exchange probability of the MulWalk--UmbrellaExchange simulation between the first two umbrellas (Fig. 5.6 d), 5.8 d), 5.11 d)). It seems that at a value close to zero the dihedral angles ψ and ϕ of the helix CV are intrinsically very unlikely for this helix peptide. If the angles had been presented in a Ramachandran plot, these values would form a kind of an exclusion zone.

Transferability

The RESTmetaD technique has been used several times before [55, 123, 168] but to our knowledge, its combination with RECT [171] and its application to parts of a protein are entirely new. Furthermore, the introduced umbrella-based simulations are a novel combination of free energy sampling methods. These techniques are suitable for the sampling of very large collective variables (CVs). They can also be applied if a particular CV can be measured experimentally, such as the helicity in the CD experiments in our case. In both cases, the umbrella based methods provide the possibility to calculate the enthalpic part along the CV easily. However, the method might neglect the sampling of all

available conformational states of the system in some cases. In this context, the presented MulWalk-UmbrellaExchange has an advantage over the MetaUmbrella simulations due to the Metropolis-Hastings exchanges, which increase the sampling and therefore the entropic contribution.

The computational effort of these free energy sampling methods is rather high due to the usage of several replicas. Including tests during the development phase of these methods, the presented simulations consumed about $4 \cdot 10^6$ core hours. The RESTmetaD simulations required rather long serial computation to sample the complete phase space, whereas the Umbrella based simulations can be easily parallelized by including a high number of umbrellas. Due to the availability of supercomputer clusters, the parallelization is much easier to realize than long serial computations.

Experimental techniques such as X-ray and NMR structural analysis have the advantage of a reliable result of the molecule structure. Still, it can be an extremely challenging task to crystallize the desired molecules. Techniques such as CD and FTIR reveal globally averaged structural informations under certain environmental effects such as adsorption. For complex molecules undergoing several conformational changes, the interpretation of the spectra of these techniques is often left to speculations. The presented free energy sampling methods can support those experimental methods significantly, giving further details about interactions at the atomic level. Furthermore, within their limits, the presented simulations provide an enormous potential for the flexibility of an externally applied perturbation, e.g a surface, different solvents or the presence of other molecules. Even nonphysical conditions are applicable and still allow the discussion of structural information.

5.3.2 Surface-induced changes of the helicity of chymotrypsin

The described three simulation methods are used to evaluate the free energy surface of the helicity of the C-terminus of chymotrypsin in three different surroundings. Across all three simulation methods, the final free energy surface of the helix peptide in solution shows a very shallow profile with a free energy difference between random structure and ideal helix configuration of about 10 kJ/mol (Fig. 5.5). Furthermore, in the case of the RESTmetaD simulation, the global free energy minima describe a random structure. The free energy difference to a helicity of about 50 % ($CV = 5$) is about 2 kJ/mol ($k_B T$ at 300

$K = 2.5$ kJ/mol). Therefore it can be assumed that a significant amount of molecules of the peptide in solution exists in a random structure or short-lived helices.

The “prediction server” calculations of the helix configuration can be separated into PDB-based predictions (homology sampling), which include the complete protein structure, and “ab-initio” calculations, which calculate the conformation based on the isolated helix. Both the homology sampling and the ab-initio calculation predict a helix structure of the sequence of the C-terminus of chymotrypsin but do not directly provide the free energy gain of the most likely configuration. In contrast, the free energy methods include the solvent and its entropic properties explicitly. Furthermore, they use a well established force field of the atom interactions which could have advantages against the simplified interaction potentials used by the prediction servers. To find further explanations for the contradiction of the server predictions and the calculated free energy surface, the extended free energy sampling methods are applied to the helix peptide embedded in the native protein.

The resulting free energy surfaces reveal an evident difference between the protein-embedded helix and the peptide in solution (Fig. 5.7). The free energy minimum is shifted towards an ideal helix structure based on a strong stabilizing effect of the surrounding protein. A clear global free energy minimum can be observed at a CV value between 8 and 10 in all three simulation methods. This broad free energy minimum is consistent with the CV value of 8.4 of the used PDB structure (4CHA.pdb).

The shape of the clusters of the protein system is more straight compared to the peptide alone (Fig. 5.12). This is caused by the embedding in the protein structure, which hinders more bent configuration. The helix peptide is embedded in the protein by an extensive network of hydrophobic interactions and hydrogen bonds in the interface between helix and protein (Fig. 5.9). It seems that the embedding network thus stabilizes the helix formation and prevents the formation of random coil structures considerably. Hydrophilic amino acids of the helix are facing to the outside of the protein and may form potential contacts to SiO_2 or another hydrophilic media. Indeed, positively charged and polar hydrophilic amino acids form strong contacts to the SiO_2 surfaces as investigated in detail in Chapter 2 and 3.

The presence of the surface results in a flatter free energy surface compared to the solvated protein in RESTmetaD simulation (Fig. 5.10). Here, a clear difference between the RESTmetaD and the umbrella based simulations can be observed. We assume that

the difference is caused by two aspects. First, the RESTmetaD is not fully converged and frequently stuck in local energy minima. Second, the umbrella based methods sample the enthalpic contribution at all CV values, but probably cannot reach all states, especially at low helicity values, and therefore underestimate the entropic contributions.

Compared to the solvated protein free energy landscape, a small kink can be observed around a CV value of 10 in the case of the adsorbed protein. The cluster analysis indicates that at this CV the middle part of the helix is formed while the tail is in a random structure (Fig. 5.12 c) f) i)). The direct representation of the frames of the trajectory shows that this C-terminus tail is not in contact with the surface (Fig. 5.13). Instead, the central part of the helix peptide is in contact with the surface whereas the C-terminus is in contact with the protein, which results in a bow-like shape consistent with the cluster analysis. It can be assumed that this surface contact is the reason for the described kink in the free energy surface at the CV value of 10.

CD experiments by Zoungrana et al. [24] showed a vanishing of the peak in the spectra at 222 nm during SiO₂ adsorption and assumed a loss of the helicity during adsorption. Similar results have been found in CD experiments by Wurzler [99] (Fig. 1.9), which reveal a computed reduction of the protein helicity from 10 % to 7.4 %. The free energy simulations including the surface also suggest only a slight surface-induced influence on the protein helicity. Calculations of the CD spectra, while one of the helices is progressively denaturated, shows the same trend of peak shift at the wavelength of about 210 nm as measured in the adsorption experiments. Furthermore, our investigations presented in Chapter 2 have demonstrated an orientation of chymotrypsin with its helice structures, in contact with the surface, which makes likely a denaturation due to direct contact. The simulations at high temperatures identify a rather stiff core of the protein which withstands high temperatures and therefore might not denaturate during the adsorption process either. This stiffness is caused by two naturally stable beta barrel structures shifted by 90° with respect to each other (Fig. 5.1). Furthermore, studies by Welzel [98] suggest a reduced denaturation of chymotrypsin adsorbed on SiO₂ compared to Teflon so that at least parts of the protein seem to maintain their native structure during adsorption on SiO₂.

In summary, it can be concluded that the chymotrypsin adsorption on SiO₂ induces only a partial denaturation of the helix in the protein, while the majority of the chymotrypsin structure keeps its native conformation.

Conclusion and Perspective

6.1 Conclusion

The adsorption of lysozyme and chymotrypsin on silica and titania has been investigated by Molecular Dynamics (MD) simulations in comparison to experiments. The MD simulations performed in this work reveal an atomic resolution of the complete adsorption process. This includes the long-range interactions of the proteins with the surface, a possible adsorption and desorption, the surface anchoring as well as a possible surface-induced conformational changes. Methodological recipes have been developed and applied, especially for the purpose of the direct comparison to experiments.

The long-range interactions between the protein and the surface have been calculated by force and energy profiles based on the Derjaguin-Landau-Verwey-Overbeek (DLVO) theory. The profiles have revealed general attractive and repulsive forces between protein and surface and thus have identified the proteins preferred adsorption orientation. In particular, the orientation and attraction of lysozyme and chymotrypsin on silica and titania have been computed. The found orientations have been confirmed by several explicit solvent simulations. This consistency discloses that from DLVO calculations one can obtain first insights on the interactions and orientations, comparable to those obtained with more advanced and computationally costly simulation methods or even from an experiment.

It was found that chymotrypsin has a distinct adsorption orientation on silica. Driven by its dipole moment, the helix region of the protein faces towards the negatively charged

silica surface. Since silica exposes an acidic surface, this orientation dominates in the whole investigated pH range between 3 and 8. Titania, on the other hand, has an amphoteric surface. Hence, it has been found that the mentioned preferred orientation of chymotrypsin dominates above the isoelectric point and an opposite orientation can be observed below the isoelectric point. On the other hand, Lysozyme on silica exhibits an orientation less influenced by the dipole moment.

The calculated long-range interactions have suggested a clear trend of attraction as a function of the pH value of the system. This trend has been confirmed by protein quantification adsorption experiments. However, the experiments have revealed a different adsorption quantity of lysozyme and chymotrypsin on silica, which cannot be explained by DLVO interactions exclusively. Further performed simulations have suggested that protein-protein interactions and the detailed surface-protein contact formation at the interface have to be taken into account as well.

A similar observation has been made by comparing the adsorption isotherms of chymotrypsin on silica and titania. The Langmuir isotherms have revealed comparable adsorption free energies on both materials but a highly increased total adsorption amount of the enzyme on titania. At the investigated pH, the computed DLVO interactions have similar attractive forces. Preliminary simulations of single and double layer formation of chymotrypsin over silica have shown that the protein monolayer does not screen the surface and allows for the formation of a double layer. Furthermore, DLVO and explicit solvent simulations have demonstrated that the favored orientation of the first and second protein layer over silica and titania is identical. Only extensive explicit solvent simulations and the analysis of several radial distribution functions have identified a difference of the adsorption process on both surfaces. On silica, the proteins anchor at particular surface reactive groups whereas on titania very ordered dense water layers at the surface mediate the adsorption process. This diverse adsorption behavior may explain the different adsorption quantities measured in the adsorption experiments.

The details of the protein anchoring at the surface have been investigated by extensive contact analysis. The atomic binding motif of lysozyme and chymotrypsin on silica and titania includes mainly positively charged protein residues on silica but also negatively charged, polar and even non-polar residues on both surfaces. The most frequent and

tightest surface binding sites are the silanol and hydroxyl groups of silica as well as terminal oxygen atoms and hydroxyl groups of titania. The calculation of the radial distribution of water and protein atoms around these surface groups has clearly identified increased water densities at the adsorption sites. The overlap of the hydration shell of the protein atoms and these sites form additional entropically driven adsorption interactions.

The stability of these contacts has been investigated by a comparison between steered MD simulations and atomic force spectroscopy (AFM) experiments, in which lysozyme has been pulled away from a silica surface. The non-equilibrium simulations in implicit solvent resulted in a quick denaturation of the protein within picoseconds. The comparison to explicit solvent simulations, on the other hand, has shown that the protein is highly stabilized by a surrounding water shell as well as by hydrogen bonds. In particular, the protein denatured during the desorption in implicit solvent whereas the native structure was preserved in explicit solvent. A similar stabilizing effect has been observed for the disulfide (S-S) bonds of lysozyme. If the S-S bonds were cleaved, the protein denatured during desorption, which is not the case with intact native S-S bonds. Experimentally, the S-S bonds were selectively cleaved by a reducing agent, which resulted in different force-distance curves, suggesting a more loose protein structure. An adsorption pattern or protein fingerprint has not been observed in the experimental force-distance curves, supporting the observation of the simulation that the protein desorbs in its native structure. It can be suggested that the adsorption free energy of the protein and therefore the stability of the surface contacts are not strong enough to break the S-S bonds before the protein desorbs from the surface.

The high resolution of AFM and the additional possibility to functionalize the cantilever tip with molecules of interest is a very powerful tool for investigations at the atomic level. However, due to the single event measurement at this resolution level, a huge variety of possible traces produces very diverse force-distance curves. The traces are based on many events, whose interpretation can be challenging due to the complexity of the system. The combination with MD simulations provides access to the kinetics at the molecular level. Both techniques rely on a limited number of force-distance curves, which is why a meaningful statistical interpretation needs to be handled with care. Still, the described combination of both methods forms a powerful tool not only to investigate protein desorption but also to help analyze any molecular AFM experiment. It can be challenging

for a steered MD simulation to reach ergodicity in order to calculate the free energy of the system. Therefore, we have further developed and used advanced MD simulations to analyze conformational changes of proteins as described in the following.

The surface-induced conformational changes have been investigated exemplarily on the chymotrypsin adsorption on silica. A standard unbiased MD simulation has revealed no conformational change of the adsorbed protein after a simulation time of over 300 ns. For a protein of this size and amount of degrees of freedom, the surface-induced conformational changes could still appear on longer time scales. To accelerate the standard MD simulations, we have therefore increased the simulation temperature, used target MD simulations and have further developed and used three kinds of advanced simulation methods. The first method is a Replica Exchange Solute Tempering method in combination with Metadynamics (RESTmetaD). The second method is an Umbrella Sampling in combination with Metadynamics (MetaUmbrella). The third method is an Umbrella Sampling with a Multiple Walkers Metadynamics combined with exchanges between the umbrellas (MulWalk-UmbrellaExchange).

To our knowledge, the combination of these methods and their application to proteins as presented in this work is new. The methods can be used in any scenario to investigate the free energy minima of a particular region of a molecule. These studies could include the configuration analysis of a protein binding cleft, or the structure of a binding motif undergoing any type of interaction. The methods based on Umbrella Sampling can force the simulation to sample along a chosen collective variable. This can be of interest if the collective variable is measured experimentally, such as the secondary structure by circular dichroism (CD). Due to multiple parallel simulations in all umbrella windows, this kind of simulations is easy to be performed in parallel on short total computational times scales. Therefore, these simulations are especially suitable for supercomputing clusters including a high number of processors.

These methods can be transferred to various interaction studies. The transferability includes not only protein adsorption but also other types of molecular interactions in comparison to a reference. The identification of attractive interactions, binding motifs and molecule orientations can be applied to systems such as ligand-acceptor, protein-protein or oligomer-complex interactions.

In this work, the presented MD simulations have been compared to CD experiments.

Clear CD spectra shifts were observed by experiments at different temperatures and by adsorption on silica. These trends have partly been reproduced by a target MD simulation with the helicity as the collective variable and a subsequent theoretical calculation of the CD spectra. Furthermore, our investigations have shown a distinct adsorption orientation of the helical content in direct contact with the silica surface, which supports arguments for a surface-induced denaturation of the helix. Additionally, the analysis of 400 - 600 K MD simulations has shown an increased instability of the C-terminus helix of chymotrypsin. Therefore, this C-terminus helix has been chosen for extensive free energy calculations on three systems. These include the isolated C-terminus helix in solution, the protein in solution and the protein in the adsorbed state on silica.

In the case of the isolated helix, all three methods have shown a very flat free energy surface as a function of its helicity. A shallow global minimum in the denaturated state has been observed for the RESTmetaD simulation. The natively embedded helix, on the other hand, is highly stabilized by the protein environment. The observed free energy minimum is comparable to the native state found in the PDB structure. Only a small change of the helicity has been observed when the protein adsorbs on silica. In conclusion, the results of this work suggest that the majority of the chymotrypsin structure is conserved during the adsorption on silica and that the protein helicity is only slightly reduced by the contact with the surface.

Neutral hydrophilic surfaces are promising candidates for biomedical applications such as implants [23]. In the literature, it has been assumed that surfaces with these properties have a high biocompatibility and that blood plasma proteins such as fibrinogen are poorly adsorbed on these surfaces, without strong conformational changes [6, 19, 22]. The ability of a surface to change the conformation of blood plasma proteins is assumed to play a major role in its biocompatibility [8, 13]. The conformational change of fibrinogen, for example, is associated with its activation and a subsequent blood coagulation. This work has shown a reduced adsorption capacity for the two investigated proteins on hydrophilic surfaces with decreasing surface charge. Furthermore, the detailed investigation of surface-induced conformational changes of chymotrypsin adsorbing on the negatively charged hydrophilic silica has shown only small conformational changes as well. Therefore, the presented results are in line with the requirements mentioned above. Further research would need to study the influence of the exact type and surface of the material as well

as the size of the nanoparticles, as those factors might have a significant effect on the biocompatibility [198].

6.2 Perspective

As stressed above, the developed methods can be applied to multiple other research projects. Therefore, in the course of this dissertation, several cooperations have been established, which includes the calculation of theoretical properties of alkaline phosphatase, the prediction of suitable mutation sites of beta-glucosidase with a surface specific binding tag and the interactions of an amino silane functionalized surface with the MS2 virus. To give a perspective for possible further research work, these projects are briefly introduced in the following.

6.2.1 Characterization of Alkaline Phosphatase

Aminian et al. [199] investigated initial cell adhesions of human osteoblasts and stromal cells on covalently immobilized alkaline phosphatase on oxides. Alkaline phosphatase is known to play a major role in hard tissue formation by osseous tissue mineralization [199–201]. The exact atomic adsorption process is unknown. In order to get an initial understanding of this process, the protein is characterized by its theoretical properties using the methods presented in this work.

The dipole moment of the protein dimer shown in red in Fig. 6.1 is bigger than the one of chymotrypsin, which is a hint for a corresponding adsorption orientation to a charged surface (Chapter 2 and 3). The outer region of the protein shows primarily unordered and flexible residue structures, as presented in Fig. 6.1 a). The flexibility of those residues is analyzed by the RMSD deviation of the individual amino acids in a short simulation in water which is shown in the color code from red to blue in Fig. 6.1 c). Amino acids marked in red are the most flexible residues. The flexible outside might lead to an induced fit of the protein to potential binding partners while the rigid core with predominantly beta sheets and helical parts gives the protein its mechanical stability.

The overall negative charge of the protein (-14 e) is also visible in the electrostatic isosurface shown in red in Fig. 6.1 b) and suggests a favored adsorption on positively charged surfaces such as titania at low pH (Chapter 3). The active sites identified by the Zn^{2+} , Mg^{2+} and PO_4^{3-} ions [202] are surrounded by areas of positively charged isosurfaces.

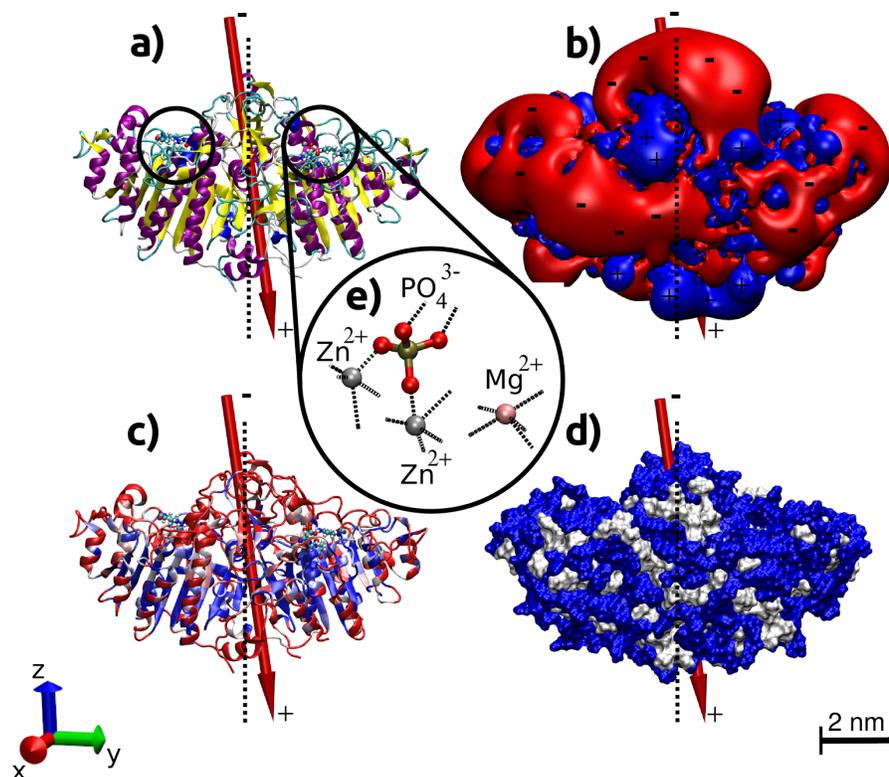


Figure 6.1: (a) The secondary structure distribution of alkaline phosphatase. The mirror symmetry of the dimer is indicated by the dotted line in all subfigures. Both monomers contain one active site marked with black circles. The protein dipole moment is shown as a red arrow (655 D). (b) A more detailed charge distribution can be derived from the electrostatic isosurface of $\pm 125 k_B T e^{-1}$ shown in blue and red. (c) More flexible residues of the protein are colored red and more rigid ones blue. The individual residues are colored by the RMSD change during a short MD simulation (0.8 - 1.4 Å in blue to red). (d) The Solvent-Accessible-Surface-Area (SASA) of the protein is colored by the underlying hydrophobic (white) and non-hydrophobic (blue) residues. (e) (center) The ions of one of the two symmetric active sites are shown in CPK-style.

This remarkable distribution might be related to the enzyme activity which is removing phosphate groups from molecules.

In order to predict hydrophilic or hydrophobic interactions of the protein with the surface the amount of accessible hydrophobic residues is estimated as well. The calculation of the solvent-accessible surface area (SASA) reveals that 21% of the accessible surface is composed of hydrophobic residues visualized in Fig. 6.1 d). In comparison to lysozyme with 26% and chymotrypsin with 42% (Chapter 2 and 3), it therefore seems that alkaline phosphatase is a rather hydrophilic protein.

These findings thus give a first idea of the atomic distribution of the protein size, the electrostatic distribution and its hydrophilicity. In order to understand the adsorption process of this enzyme in more depth, further simulations are required, taking into account the surface and the adsorption environment.

6.2.2 Mutation of Beta-Glucosidase

The group of Nils Kröger and Nidhi Dubey from the Technical University of Dresden is able to replace one lysine in the protein beta-glucosidase with a silica specific binding tag. The binding tag is composed of six arginine amino acids (Fig. 6.2 a)). Arginine as a positively charged amino acid has a high affinity to silica as presented in Chapter 2 and 3. Since the synthesis of the mutated protein including the binding tag is very complex, not all 35 lysine positions can be evaluated experimentally to find the most suitable mutation site. Therefore, we present in the following a simulation strategy to predict a suitable binding site for the tag.

Each lysine position of the protein is rated by:

- (i) The solvent accessibility, to find the mutation site with the smallest influence on the native tertiary structure.
- (ii) The pocket accessibility for ligands, to ensure protein activity after the mutation.
- (iii) The binding affinity based on DLVO interactions, so that the tag is supported by the binding affinity of the protein.

All three contributions are translated into scaled values of 0 - 1. A value of one is associated with a high solvent accessibility, a high pocket accessibility and a high binding affinity and vice versa.

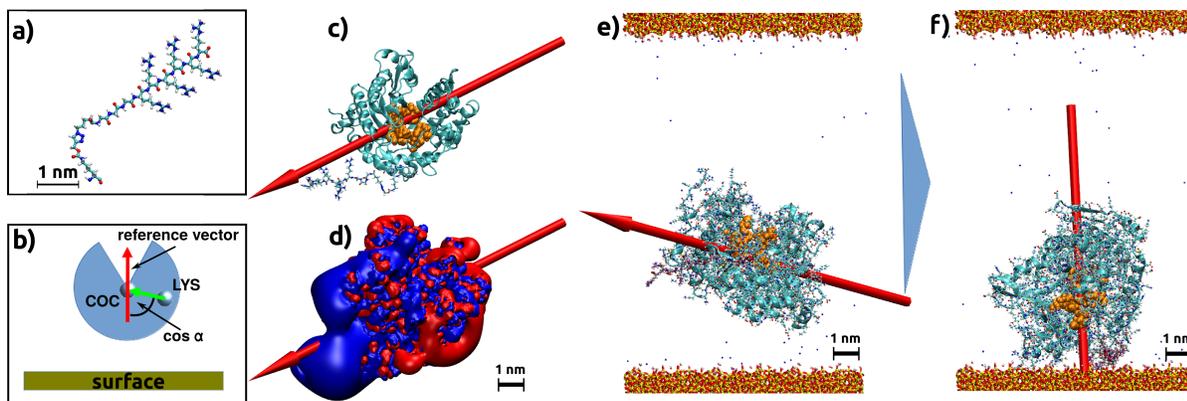


Figure 6.2: a) Silica surface specific binding tag with six arginine residues. As a mutation, it replaces one lysine residue in the protein. b) Scheme for the calculation of suitable mutation sites. Blue is the protein with the active site marked by the blank triangle. c) and d) show the protein with a tag mutation of the lysine 440. The isoelectric surface is presented by ± 43.9 mV in blue and red. e) and f) show the start and end position of a 50 ns MD simulation in explicit water. The simulation is adapted from the procedure introduced in Chapter 2 Fig. 2.6.

(i) The solvent accessibility is identical to the embedding of the residues in the protein structure. It is calculated by the EDTSurf [203] tool, which calculates the solvent-accessible surface and a quantity per amino acid which is representative for its embedding. This quantity value is normalized by its maximum to values from 0 to 1 for each lysine.

(ii) The pocket accessibility for ligands after the mutation is based on orientation arguments. If a lysine is located on the opposite side of the active site a high pocket accessibility is assumed. A representative value for the accessibility is calculated by the geometric ratios as shown in Fig. 6.2 b) and by:

$$v_{ac,i} = \frac{\cos(\alpha_i) + 1}{2} .$$

α_i is the angle between the lysine i , the center of coordinates and the surface normal (0 0 1) of the protein. In summary, we suggest that a $v_{ac,i}$ value of 1 is associated with a high ligand accessibility and probably high activity of the protein and vice versa for a value of 0.

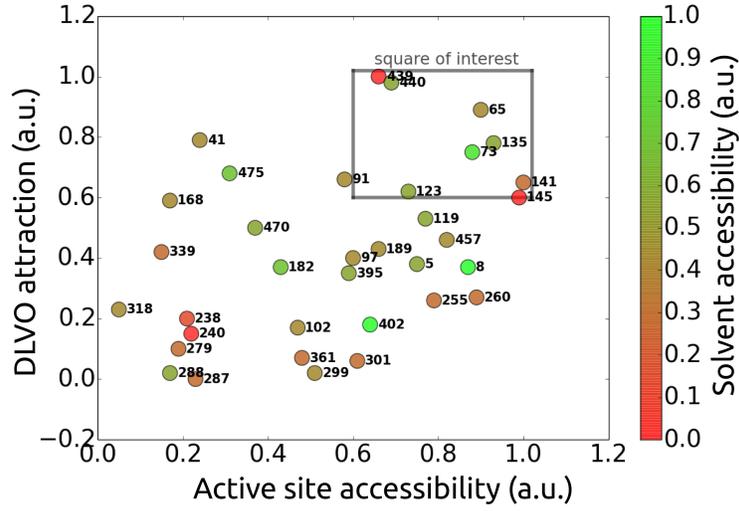


Figure 6.3: The suitability of a mutation site is quantified by values from 0 - 1 of the DLVO attraction (ordinate), active site accessibility (abscissa) and solvent accessibility (color code). The numbers indicate the lysine residue numbers. The “square of interest” marks the most suitable mutation sites.

(iii) The silica surface affinity of the protein is calculated by DLVO interaction profiles as introduced in Chapter 2 Fig. 2.1. If the calculated DLVO forces support the adsorption of a lysine position facing towards the surface, an attraction of this lysine position is assumed. The DLVO affinity $v_{af,i}$ for each lysine i is calculated by:

$$v_{af,i} = \sum_j (-F_{1nm,j}) \cdot \frac{\cos(\alpha_{i,j}) + 1}{2} .$$

$F_{1nm,j}$ is the DLVO force on the protein for the orientation j in a height of 1 nm above the surface. $\alpha_{i,j}$ is the angle between the lysine, the center of coordinates and the surface normal (0 0 1) of the protein in the orientation j . Therefore, the value $v_{af,i}$ is weighted by the force of the orientation and also by the lysine position of this orientation. In summary, we assume that a $v_{af,i}$ value of 1 is associated with supporting DLVO forces within a possible adsorption and vice versa for a value of 0.

All three values from (i) - (iii) are visualized in Fig. 6.3. The visualization in one figure allows for a comparison of all three contributions. A green dot in the top right is the most favorable mutation site. Therefore, a square of interest is defined in this region containing the lysine sites with the most promising adsorption properties after adding the binding tag

(Fig. 6.3). Based on these definitions, lysine 440 is selected as a suitable candidate and lysine 288 as an unsuitable one.

In order to compare these predicted mutation sites, two standard MD simulations of 50 ns are performed with lysines 440 and 288 being replaced by the binding tag, respectively. The structure and the electrostatic isosurface of the protein modified at lysine 440 as well as the start and end configuration of the Lys440 MD simulation are shown in Fig. 6.2. The simulations confirm that the protein with the mutation at Lys440, which has been predicted to be the better choice for the mutation, adsorbs straight away with the surface specific tag. The mutant with the tag at Lys288, which has been predicted to be an unsuitable choice for the mutation, adsorbs with the opposite side rather than the binding tag. The predicted properties of the Lys440 mutant such as the increased surface binding affinity and maximal protein activity are currently verified in ongoing experiments.

6.2.3 Amino Silane Functionalized Surfaces

Julia Bartels [204] recently demonstrated the successful filtering of a MS2 virus-contaminated aqueous solution by using an amino silane functionalized ceramic. Due to the large pore size of the ceramic, a very high flow rate and a higher virus retention were accomplished when compared to commercially available polymer filters [204]. However, neither the precise structure of the amino silane coating nor the interactions responsible for the virus surface adsorption are known.

The MS2 viruses are composed of a protein shell (the capsid) with well-defined pores, a rather short single-stranded RNA core and a maturation protein (M-Protein) embedded in the capsid [205]. The schematic structure is shown in Fig. 6.4 a) and an atomic representation of the side view of one of the pores is shown in Fig. 6.4 d).

The experimental results revealed a very promising functionalization with mono- and tri-amino silane molecules [204], as shown Fig. 6.4 c). The silanization can form a complex network of several covalent and ionic binding types beyond a monolayer coverage of the ceramic [146]. Based on these conditions three types of interactions might be possible, which are investigated in further simulations.

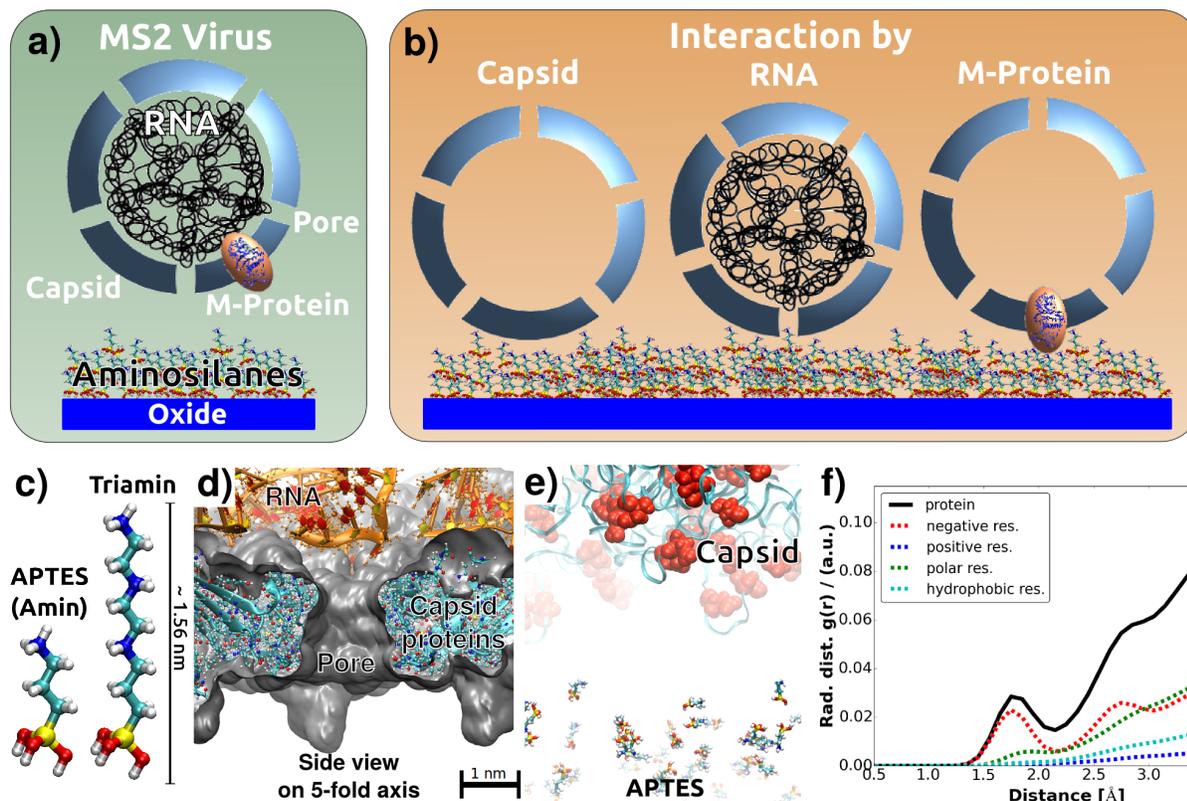


Figure 6.4: Interaction of an MS2 virus with an amino silane functionalized ceramic filter membrane. a) Schematic overview of the RNA virus with its prominent holes in the protein capsid and its embedded maturation Protein (M-Protein). b) Possible types of interactions via capsid, RNA or M-Protein with the surface. c) Two types of amino silanes with one and three amino groups. d) Atomic presentation of the 5-fold axis of the capsid and parts of the RNA genome based on PDB data. e) Multiple Walkers simulation of 225 amino silane molecules interacting with the virus capsid. f) Radial distribution function of the capsid amino acids with the amino silane molecules of the simulation shown in e).

- (i) An adsorption based on capsid interactions (Fig. 6.4 b) *left*).
- (ii) A binding of the positively charged aminosilane to the negatively charged genome through the virus pores (Fig. 6.4 b) *middle*).
- (iii) A surface-virus interaction by the M-Protein, which is essential for the adsorption to the native host cell [206] (Fig. 6.4 b) *right*).

One of the simulations investigating the presented interactions is based on a Multiple Walker Steered MD simulation as shown in Fig. 6.4 e) and f) which is an advancement of the multiple walker simulations presented in Chapter 5. For this simulation, 225 amino silane molecules are pulled periodically towards the capsid and back again in individual

steered MD simulations (Fig. 6.4 e)). All 225 amino silanes are in one simulation box without additional replicas. Still, each amino silane molecule samples one steered MD interaction. This kind of simulation uses the fact that the simulation box has enough space to contain many small molecules of interest at the same time. Hence, the sampling compared to an ordinary single molecule steered MD simulation is increased by the factor of 225. To our knowledge, this kind of method is new and contains a high potential to be used in many ligand-receptor or peptide-adsorption interaction studies.

The simulation is preliminarily analyzed by the radial distribution function (RDF) as introduced in Chapter 2 and 3. The RDF reveals the averaged radial distribution of the amino silanes with certain types of amino acids as shown in Fig. 6.4 f). The negatively charged amino acids of the capsid proteins dominate the near-field interactions at distances of about 1.7 Å, suggesting strong electrostatically driven interactions between virus capsid and the functionalization.

Even stronger electrostatic interactions can be expected and could be investigated by the positively charged amino silanes with the negatively charged genome through the pores of the MS2 virus. Additional investigations would include the construction of a realistic surface and a virus model with M-protein and genome using all-atom, alchemical and free energy sampling methods. A final goal of this project is the development of a computer-aided design of an optimal and selective surface functionalization using a combination of bioinformatics, docking simulations and MD tools.

The three presented examples in Section 6.2.1, 6.2.2 and 6.2.3 demonstrate close cooperations with experimental projects highly suitable for the kind of MD simulations introduced in this thesis. The combination of experiments with MD simulations is a very powerful tool to explain an experiment at the atomic level. Otherwise, the events at the atomic level are sometimes left to a speculative debate in many stand-alone experimental methods. On the other hand, independent MD simulations are often left to an argumentative debate about the reliability of the used model e.g. the force field [207, 208]. In combination, however, both techniques support each other. The observed effect in the well-performed experiment can be accurately explained by the simulated model. Furthermore, the all-atom MD simulations provide the possibility of further analysis at the atomic level, like the identification of binding motifs or structural clusters.

In most of the joint experimental and simulation projects, the simulations represent a supporting method to explain the properties found in the experiments, as frequently applied in this thesis. For time and cost reasons, however, in many cases it is not possible to perform large variations of experimental measurements. Therefore, another approach is the prediction of a possible outcome of the experiment by the simulation, as described in the research projects of the beta-glucosidase and the MS2 virus interactions. This strategy significantly enriches the level of usability of the simulations. In this case, the rough screening of massive datasets is left to the simulation, while the still necessary experiment can focus on the validation of the most promising candidates. As long as the predictive approach is realized *in silico*, it can be powerfully scaled with regard to sample screening and system complexity.

APPENDIX: Molecular Dynamics

Simulations

The biological interface as investigated in this dissertation is particularly suitable to study via standard molecular dynamics (MD) due to the following reason. We expect no significant chemical reactions during the adsorption process. The system chemistry of MD simulations are mostly pre-defined and no reactive force fields are used. The dynamics are driven by Coulomb and Lennard-Jones interactions of non-bonded and Hook's law interactions of bonded atoms. Each atom is explicitly represented. Dependent on computational resources a typical system size in MD simulations measures $10 \times 10 \times 10 \text{ nm}^3$ and lasts ns to μs in system time. Many biological molecules and macromolecules have these dimensions and are rather flexible so that dynamical conformational changes are possible to meet the time scale of the simulation. Equally, the diffusion of the solute in an aqueous solvent, as well as the water and ion diffusion itself, is nicely observable by this kind of simulations.

7.1 Standard Molecular Dynamics Simulations

7.1.1 Interactions Described by Force Field Equations

In standard MD simulations the interactions of the atoms are represented by equations of standard mechanics. The parameters of these equations are summarized in the force field. In the following the AMBER force field equation is described as used in this disser-

tation [81]. In the AMBER approach the atom interactions are divided in bonded and non-bonded interactions. Three types of bonds (length, angle, dihedral) and two types of non-bonded (Coulomb, Lennard Jones(LJ)) interactions are typically used. The potential of the AMBER force field equation is given as [76]:

$$\begin{aligned}
 V &= \overbrace{V_{bond} + V_{angle} + V_{dihedral}}^{\text{bonded}} + \overbrace{V_{Coulomb} + V_{LJ}}^{\text{non-bonded}} \\
 &= \overbrace{\sum_{1-2 \text{ pairs}} k_r (r - r_{eq})^2}^{\text{bonds}} + \overbrace{\sum_{1-3 \text{ pairs}} k_\theta (\theta - \theta_{eq})^2}^{\text{angles}} + \overbrace{\sum_{1-4 \text{ pairs}} k_\phi (1 - \cos(n\phi + \delta))}^{\text{dihedrals}} \\
 &+ \overbrace{\sum_{\substack{\text{non-bonded} \\ 1-4 \text{ scaled}}} \left(\frac{q_i q_j}{\epsilon r_{ij}} \right)}^{\text{Coulomb}} + \overbrace{\sum_{\substack{\text{non-bonded} \\ 1-4 \text{ scaled}}} \epsilon_{ij} \left(\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right)}^{\text{LJ}}
 \end{aligned} \tag{7.1}$$

The bond length potential is defined by a harmonic potential identical to Hook's law (Equation 7.1). The equilibrium length r_{eq} and spring constant k_r are the two parameters individual for the atom type in the force field. Similar to all bonded atoms this parameters can be determined by the vibrational modes of experimentally measured infrared spectroscopy or quantum mechanics based simulations. In case the bond deformation exceeds distances beyond the equilibrium state the retracting forces on the bond become very high. In this case the Morse potential is often used since it decreases exponentially to a constant potential value for higher bond distances. The harmonic angle potential parameters consists of the equilibrium angle θ_{eq} and the force constant k_θ , it describes three direct neighboring atoms. The non harmonic dihedral part describes the rotation of four bonded atoms. It is described by a periodic cos function and formed by two parameters: the force constant k_ϕ and the phase shift δ , which is identical to the equilibrium angle. The number n is the periodicity of ϕ . These parameters are often calculated by a direct comparison to ab-initio based simulation such as density functional theory (DFT). Typically this parameter set is adjusted at the very last so that it acts like a collecting vessel by adjusting to the quantum mechanical based potential and compensates e.g. for slightly wrong angle parameters.

The electrostatic interactions are described by the Coulomb law (Equation 7.1). ϵ is the dielectric constant. Each atom in the simulation has a partial charge ($q_{i,j}$). Based on this charge the Coulomb potential between two atoms decreases with the inverse distance.

Therefore, in theory, these interactions have an infinite range. To save computational effort a cut-off of typically 0.8 to 1.2 nm is used, so that no interactions beyond this cut-off are considered.

Furthermore, all non-bonded atoms and bonded atoms which are further apart than the four next neighbors within the cut-off are described by the Lennard-Jones (LJ) potential. In the AMBER force field equation, a 12-6 Lennard-Jones potential is used (Equation 7.1). The attractive part is caused by Van-der-Waals forces and dipole-dipole interactions. These interactions are summarized by an attractive potential proportional to r^{-6} . The empirically chosen repulsive part of the LJ equation is based on the Pauli repulsion and is set proportional to r^{-12} . Both non-bonded terms, Coulomb and LJ are additionally employed in a scaled form in the AMBER force field for the 1-4 interactions [81]. This additional term acts on all atoms which are bonded with two partners in between, similar to the dihedral atoms. The Coulomb potential on these atoms is scaled down by a factor of 1.2 and the LJ potential by a factor of 2.0.

7.1.2 From the Force Field Equation to the Dynamics

By the calculation of the potential of the system as described in Section 7.1.1 the physics of the atoms interactions are already defined. With the potential of the system $V(\vec{r})$ the force of each atom (F_i) can be determined by the Newton's equation of motion:

$$\vec{F}_i = -\nabla V(\vec{r}_i) = m_i \ddot{\vec{r}}_i$$

The integration over time reveal the velocities $\vec{v}_i(t) = \dot{\vec{r}}_i$. The numeric time integration, essential for the MD, is described in the following.

The dynamics of a molecular system involves interactions of several neighboring atoms. Therefore, the atoms do not follow straight-line trajectories like in a hard sphere simulation of few collisions. Thus a trajectory can only be calculated iteratively by a finite-difference scheme in small time-steps after which the potential is updated. From a known conformation $\vec{r}(t_0)$ the positions at $\vec{r}(t_0 + \Delta t)$ are calculated. The finite-difference method is based in most of the cases on a truncated Taylor expansion [209]. The simplest, truncated after the first-order term is Euler's method.

$$\vec{r}(t + \Delta t) = \vec{r}(t) + \dot{\vec{r}}(t)\Delta t \quad (7.2)$$

With a known or predicted position and velocity at time t the position at time $t + \Delta t$ can be computed. This Euler method is the basis of the Runge-Kutta method (RK) [210] and forms a first-order RK method. Other RK schemes vary by the calculation of the slope $\dot{\vec{r}}(t)$. A fourth order RK e.g. calculates the slope at four points at fractions of Δt and computes the average which is used for Equation 7.2. The evaluations of the intermolecular forces and consequently the slope is by far the most time-consuming calculation [209]. Hence, the fourfold calculation of the force interactions for one timestep is very inefficient.

A more efficient way is the usage of positions and velocities from previously calculated steps. This approach is used in the more popular Verlet algorithm [211]. The Verlet algorithm is a combination of two Taylor expansions from position of time t forward to $t + \Delta t$ and another backward to $t - \Delta t$ [209].

$$\begin{aligned}\vec{r}(t + \Delta t) &= \vec{r}(t) + \frac{d\vec{r}(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\vec{r}(t)}{dt^2}\Delta t^2 + \frac{1}{3!}\frac{d^3\vec{r}(t)}{dt^3}\Delta t^3 + \mathcal{O}(\Delta t^4) \\ \vec{r}(t - \Delta t) &= \vec{r}(t) - \frac{d\vec{r}(t)}{dt}\Delta t + \frac{1}{2}\frac{d^2\vec{r}(t)}{dt^2}\Delta t^2 - \frac{1}{3!}\frac{d^3\vec{r}(t)}{dt^3}\Delta t^3 + \mathcal{O}(\Delta t^4)\end{aligned}$$

Adding both expansions eliminates the odd-order terms and yield the Verlet's algorithm for positions:

$$\vec{r}(t + \Delta t) = 2\vec{r}(t) - \vec{r}(t - \Delta t) + \frac{d^2\vec{r}(t)}{dt^2}\Delta t^2 + \mathcal{O}(\Delta t^4)$$

The velocities do not explicitly appear in the algorithm, they can be calculated in various ways, e.g. by an estimate at the half-step $t + \frac{1}{2}\Delta t$ [212]:

$$\vec{v}(t + \frac{1}{2}\Delta t) = \frac{\vec{r}(t + \Delta t) - \vec{r}(t)}{\Delta t} \quad (7.3)$$

This algorithm estimates the positions at $t + \Delta t$ and $t - \Delta t$. Therefore, it is not self-starting, it needs next to $\vec{r}(0)$ and $\vec{v}(0)$ also $\vec{r}(-\Delta t)$ which needs to be backward calculated e.g. by an Euler method [209]. Further variants of this methods exists such as the Leapfrog scheme in which the velocity is defined at the half-step similar to Equation 7.3 [213, 214] and the velocity-Verlet method [215] which is a self starting-algorithm [216].

7.1.3 The Simulated Solvent and Periodic Boundary Conditions

Atomistic MD simulations can be performed in three ways, in vacuum, with an implicit or an explicit solvent. Biomolecules in a natural environment are nearly exclusively in

contact with a solvent. In comparison to vacuum, the effect of the solvent is crucial in most of the cases. The solvent influences and often stabilizes the biomolecule by a network of electrostatic, hydrophobic and entropic properties. Therefore, in biological systems, a solvent different from vacuum is frequently used. The implicit solvent imitates the solvent by a continuous analytical representation which influences the electrostatics as shown in Section 7.3 and 1.3.1. An explicit solvent, however, surrounds the solute and fills the simulation box. Therefore, in explicit solvent simulations, periodic boundary conditions are used. The periodicity prevents nonphysical edge effects and also allows the solute to diffuse and interact across the box edges. This makes it necessary to include enough pure solvent regions in the system to prevent interactions of the solute with itself.

A method to handle the non-bonded interactions across the periodic boundary conditions is the Particle-mesh-Ewald (PME) method based on the Ewald summation [217,218]. The crucial trick of this method is to divide the non-bonded interactions in short and long range interactions treated in real and Fourier space respectively. The sum of the short-range interactions is chosen to converge quickly in real space whereas the long-range interactions converge quickly in Fourier space. Crucial for the convergence of the electrostatic long-range interactions is a system charge neutrality. A detailed derivation of the Ewald sum can be found in the book of A.R. Leach [212].

7.1.4 Thermostats

If a standard MD simulation is treated without any thermostat, it is mostly in an NVE ensemble. The energy, number of atoms and system volume is constant but not the overall velocities nor the pressure in the system. In the case of a non-equilibrated system hot spots of potential energies can heat up the system or even produce very high local velocities. Temperature and pressure thermostats are used to equilibrate the system. The thermostats influence the atom velocities or the system pressure by adjusting the system size. The direct thermostat control by increasing the temperature is an easy way to increase the system energy, to produce e.g. a glassy phase from a crystal or to denaturate a protein from the native structure (as used in Chapter 5).

The simplest way to control the temperature is to scale the velocities of all atoms by a factor λ [219]. This approach is called velocity rescaling. The system temperature is

proportional to the average kinetic energy by the relation [212]:

$$\langle \mathcal{K} \rangle = \frac{3}{2} N k_B T$$

with the degrees of freedom N . The kinetic energy is given by $\mathcal{K} = \frac{1}{2} \sum_{i=1}^n m_i v_i^2$, with the number of atoms n . Thus, the temperature difference between the target and the current temperature:

$$\Delta T = T_{tar} - T_{cur} = \frac{1}{2} \sum_{i=1}^n \frac{2 m_i (\lambda v_i)^2}{3 N k_B} - \frac{1}{2} \sum_{i=1}^n \frac{2 m_i (v_i)^2}{3 N k_B}$$

$$\Delta T = (\lambda^2 - 1) T_{cur}$$

$$\lambda = \sqrt{T_{tar}/T_{cur}} \quad (7.4)$$

To control the temperature the velocities are multiplied by this λ value with the current temperature calculated from the kinetic energy.

An alternative way to control the temperature is to couple the system to an external heat bath. This method was introduced by Berendsen et al. [220]. The rate of change of the temperature of this Berendsen thermostat is proportional to the difference of the current and the bath temperature [212]:

$$\frac{dT_{cur}(t)}{dt} = \frac{1}{\tau} (T_{bath} - T_{cur}(t)) \quad (7.5)$$

τ is the coupling constant between the heat bath and the system. Solving this differential equation by an exponential function shows directly that the temperature decays exponentially to the target temperature. The rate (Eq. 7.5) expressed in the difference of temperature at the timestep Δt is:

$$\Delta T = \frac{\Delta t}{\tau} (T_{bath} - T_{cur}(t))$$

With Equation 7.4 follows the scaling factor λ for the velocities:

$$\lambda^2 = \frac{\Delta t}{\tau} \left(\frac{T_{bath}}{T_{cur}(t)} - 1 \right) + 1$$

τ , the coupling constant has the unit of a time. If this number is large a weak coupling is applied. If $\tau = \Delta t$, the timestep, the algorithm is identical to a direct scaling as presented in Equation 7.4. This method does not generate temperature fluctuations within canonical averages [212].

If not stated differently in all the results presented in this dissertation a method based on both, the described velocity rescaling and Berendsen approach is used. This V-rescale thermostat is described by Bussi et al. [221] and included in the GROMACS simulation package [164]. This velocity rescaling works similar to Equation 7.4 but with a modified target temperature T_{tar} . T_{tar} is modified from the canonical equilibrium distribution for the kinetic energy just as the resulting temperature fluctuates within the canonical averages. This procedure ensures a proper canonical ensemble. Similar to the Berendsen approach the same coupling constant τ is used to influence the coupling of the thermostat to the system.

Further popular thermostats are the Anderson, Nosé-Hoover and Langevin thermostat. The Anderson thermostat is a stochastic collision method [212, 222]. The velocity of a randomly chosen atom is reassigned by a random selection from the Maxwell-Boltzmann distribution. The Nosé-Hoover thermostat is an extended system method [212, 223, 224]. Here, an additional reservoir is introduced with an own potential and kinetic energy, the coupling to the real system influences the temperature fluctuations which conserves the canonical ensemble of the original system. The Langevin thermostat modifies Newton's equation of motion by the addition of a friction term and a random force [225].

Similar to the temperature control, it can be of interest to control the system pressure. Experiments are often performed at constant temperature and pressure conditions which are mimicked in many biological simulations. Furthermore, it can be of interest to control the pressure to study phenomena of pressure-induced transitions [212]. The procedures of the barostats are similar to the temperature control methods in many cases. In particular the coupling of the system to a 'pressure bath' of the Berendsen barostat [220]. Analog to the temperature control the rate of change of pressure is given by [212]:

$$\frac{dp_{cur}(t)}{dt} = \frac{1}{\tau_p}(p_{bath} - p_{cur}(t))$$

with the coupling constant τ_p between the pressure of the bath p_{bath} and the current

pressure $p_{cur}(t)$. The pressure adapts exponentially to the bath pressure. Therefore a thermodynamic ensemble is not represented. Still, the system equilibrates quickly to the desired system pressure.

An example of an extended system pressure-coupling is the Anderson barostat [222]. This barostat can be associated with a piston acting on the system [212]. A similar approach is chosen in a self-developed procedure to adjust the pressure in the systems presented in this dissertation. This approach is used in case the system includes an oxide surface. In this case, the bulk atoms are frozen in the simulation. Since periodic boundary conditions are used and the surface reaches the simulation box edges the compressibility in the X-Y-plane is zero. To adjust the water density, the simulation box is shrunk in the Z-direction. In particular, the approach includes the reduction of the box size in Z-direction, an equilibration of the water and a measurement of the water density which is applied iteratively. In this thesis, we have used this method to adjust the water density to the density at 300 K and 1 bar pressure in this case of the TIP3P density of 983 ± 1 kg/m³.

7.2 Explicit Solvent

Several different water models are available for standard MD simulations. Most of them are rigid, some are flexible, polarizable or dissociable [226]. The majority of models vary slightly by their bond distances and angles of the molecule but mainly by the charge associated with the oxygen atom. All these models have a partial charge at the position of the hydrogens in common but vary by the position of the second charge. It is either located at the position of the oxygen (SPC [227], TIP3P [228]), at the outside of the triangle molecule (PPC [229]), at the inside of the triangle (TIP4P [228], OPC [230]) or two charges at the outside (TIP5P [228], ST2 [231]). An overview of popular water models can be found in [232] and on the website [233].

In this work, we exclusively use a rigid TIP3P model [228]. An NPT MD simulation for 1 ns of a box filled with this water model and the dimensions of $3 \times 3 \times 3$ nm³ are performed with a reference pressure of 1 bar using a Berendsen barostat. This simulation revealed a TIP3P density of 983 ± 1 kg/m³.

7.3 Implicit Solvent Simulations

An explicit solvent of an MD simulation can be imitated by an implicit solvent. In this case, the water is represented by a continuous model based on the solution of analytical equations. The representation of the electrical double layer (see Section 1.3.1) of these models are an essential requirement since most of the interactions in an MD simulation appears at the interface or in close distance of the solute atoms. To represent the electrostatic interactions, two models are widely used in MD simulations. The Poisson-Boltzmann (PB) (Section 1.3.3) and Generalized Born (GB) approach (Section 1.3.4). From a simulation point of view, the following crucial properties need to be mentioned. The PB approach calculates the electrostatic potential in the most accurate way available in the most common simulation packages. The properties of the Stern layer resulting from effects of the hydration shell of the ions are not included in this model (Section 1.3.1). It can be computationally expensive to solve the differential PB Equation 1.1 for each step in the MD simulation. Furthermore, the numerical calculation is performed on a predefined grid. For the calculation e.g. of an interaction profile above a charged surface, the usage of this grid can lead to periodic artifacts visible by a wavy appearing energy profile. Depending on the used algorithm and grid the calculation of the PB approach can easily be as expensive as an explicit solvent calculation.

The GB approach, on the other hand, is an approximation to the PB method and can be computed analytically (Section 1.3.4). This approach simplifies and accelerates the calculation notably. Crucial parameters in the GB Equation 1.2 are the effective Born radii, which are calculated in each MD step.

If an explicit solvent is not included in the simulated system, the free energy of solvation of the solute needs to be computed in another way. This free energy of solvation is given by [81]:

$$\Delta G_{solv} = \Delta G_{el} \Delta G_{non-el}$$

The described PB and GB approaches represent the electrical contribution to the solvation energy, in other words, the energy costs to bring a charge on a solute. The embedding of the solute in the solvent requires a further energy. First, the unfavorable cost of breaking the structure of the solvent by “digging a hole in the solvent” and second, the formation of a favorable VdW interaction between solute and solvent. This interaction is summarized by ΔG_{non-el} . In the used AMBER simulation package ΔG_{non-el} is taken to

be proportional to the solvent accessible surface area (SASA) which needs to be computed for each simulation step [81]. The computation is performed by the LCPO algorithm [234].

These implicit solvent methods are not exclusively used for the replacement of an explicit solvent. Instead, these methods can compute the statistical distribution of the ions in the system in one step. This is a tremendous advantage against explicit simulations because typically the included ions are few in its number and statistic distributions are unlikely. Still, due to its high partial charge of the ions, their influence can be strong. This is the reason why the implicit solvent models are still in use in many alchemical free energy sampling methods like thermodynamic integration or Molecular Modelling Poisson-Boltzmann Surface Area (MM-PBSA) methods.

7.4 Collective Variables

A collective variable (CV) is a function of coordinates. Consequently, the multidimensional coordinates are simplified in one dimension. Multidimensional CVs are also possible but not discussed and used here. A very simple CV variable is e.g. the distance between two atoms. A more complex CV is the e.g. the Radius of Gyration including the calculation of the center of mass (COM) and the distance of all atoms to the COM weighted by their masses. A CV is sometimes mixed up with a reaction coordinate which is a CV used to follow a reaction from A to B.

In this dissertation, the helicity of a peptide or protein is used as a CV which is described in the following.

7.4.1 HELIX Loops

The helix loops is a CV to describe the helicity of a molecule including amino acids. It is presented in the MD patch PLUMED 1 [183]. With the idea of a Ramachandran-Plot the helicity is declared as the deviation from the ideal dihedral angles of the backbone $\Phi_0 = -68.75^\circ$ $\Psi_0 = -45.0^\circ$. The helicity H is given by:

$$H = \sum_{i=2}^{N-1} \prod_{j=i-1}^{i+1} \frac{1}{4} [\cos(\Phi_j - \Phi_0) + 1] [\cos(\Psi_j - \Psi_0) + 1]$$

i is the index of the amino acids with their total number N . An ideal helix of all amino acids of the molecule has the CV value of $N - 2$.

An alternative to the helix loops is the alphaRMSD CV presented by Pietrucci et al. [235]. This CV uses the root-mean-square-deviation (RMSD) comparison of the N, C $_{\alpha}$, C $_{\beta}$, O and C atom positions to ideal helix positions obtained by the average of several X-ray/NMR PDB structures. The RMSD difference of six continuous residues is calculated in segments and in a switching function, typically a rational function to calculate the CV value of the helicity. If N is the number of all amino acids of the molecule the ideal alphaRMSD helix is formed at a CV value of $N - 5$.

7.5 Free Energy Sampling Methods

Since standard MD simulation of a complex biological system is often trapped in local minima of the free energy landscape more advanced MD simulation methods are presented in the following. Additional bias potentials are added on top of the unbiased force field potential along a chosen CV to compensate the quasi-ergodic sampling of the standard MD simulation.

7.5.1 Arrhenius Equation

Two states A and B along a particular reaction coordinate can be separated by an energy barrier E_a (Section 7.4). In this case, the rate k of the transition between A and B is defined by the Arrhenius equation:

$$k = Ae^{-\frac{E_a}{k_B T}} \quad (7.6)$$

with the rate k in s $^{-1}$, the energy barrier E_a in kcal/mol, a frequency prefactor A , for molecular processes about 10 12 s $^{-1}$ [150], the Boltzmann constant k_B and the temperature T in K. It can be seen that this equation is based on the Boltzmann distribution $\propto e^{-\frac{E_a}{k_B T}}$. Thus, the number of states for higher temperatures decreases with an exponential function. Therefore, it is very unlikely that a high energy barrier is crossed for low temperatures in a short period of time. An energy barrier in the range of $k_B T$ at 300 K equals $E_a \approx k_B T \approx 0.6$ kcal/mol. With Equation 7.6 follows that $k \approx 3.7 \cdot 10^{11}$ 1/s. Hence, the energy barrier is crossed in average in 2.7 ps, which is on the timescale of MD

simulations. In comparison, an energy barrier of $15 k_B T$ at 300 K is crossed in $\approx 3.3 \mu\text{s}$, which is not necessarily reached in an MD simulations.

This reduced ability to cross energy barriers in standard MD simulations at room temperature motivates the free energy sampling methods described in the following. These methods use one of two tricks to cross higher energy barriers by either decreasing the energy barriers in a defined way or by increasing the system energy.

7.5.2 Umbrella Sampling

In case a high energy barrier along a certain CV, as described in Section 7.5.1, exists, umbrella sampling can be used to analyze this barrier. The method requires the *a priori* knowledge of the energy barrier. A bias potential is set to steer the system close to the barrier. The bias potential acts on a chosen CV by an additional harmonic restraint potential $V(s(q)) = k/2(s(q) - s_0)^2$. In case only one restraint is used, the simulation is equal to a target MD simulation. Assuming the system is in a canonical ensemble the probability distribution without the additional bias is given by [163]:

$$P(s) \propto \int dq e^{-\frac{U(q)}{k_B T}} \delta(s - s(q))$$

with the coordinates q the CV named s , temperature T and Boltzmann constant k_B . This probability can be used to calculate the free energy F which is given by:

$$F(s) = -k_B T \ln P(s) + F_0 \quad (7.7)$$

The biased probability distribution is given by:

$$P'(s) \propto \int dq e^{-\frac{U(q)+V(s(q))}{k_B T}} \delta(s - s(q)) \propto e^{-\frac{V(s(q))}{k_B T}} P(s)$$

Therefore, the unbiased probability can be easily computed from the biased distribution as found in the simulation by:

$$P(s) \propto e^{\frac{V(s(q))}{k_B T}} P'(s)$$

In other words the biased histogram (based on $P'(s)$) can be weighted by the given exponential factor to get a unbiased histogram, which can be further used by Equation 7.7 to compute the free energy. This free energy is also called the potential of mean force.

In most of the applied cases, umbrella sampling is used to include more than one umbrella along the chosen CV. With a good overlap of the umbrellas in the CV space, the free energy along the complete CV can be computed using the Weighted Histogram Analysis Method (WHAM) [173, 174].

WHAM - Weighted Histogram Analysis Method

This method is based on the principle mentioned in the last section that a weighted histogram with discrete bins provides the possibility to compute the relative probability of observing a state of interest. It is shown in Equation 7.7 that as long as the probability distribution is known the relative free energy can easily be computed. The WHAM algorithm is applied for those cases of more than one umbrella applied on the same CV. The relative shifts F_0 (see Eq. 7.7) and the probabilities $P(s)$ are unknown. The algorithm does not compute directly the free energy but adjusts the probabilities across the umbrellas to generate one smooth continuous probability along the complete CV including all umbrellas. The algorithm uses an iterative determination to adjust the probabilities. The essential WHAM equation is the best estimate of the unbiased probability distribution $P(s)$ given as [173, 174]:

$$P(s) = \frac{\sum_{i=1}^{N_{umb}} n_i(s)}{\sum_{i=1}^{N_{umb}} N_i e^{\frac{F_{0,i} - V_i(s)}{k_B T}}}$$

with:

$$F_{0,i} = -k_B T \ln \left(\sum_{s_{bins}} P(s) e^{-\frac{V_i(s)}{k_B T}} \right)$$

with the number of the umbrellas N_{umb} , the number of counts in histogram bin associated with CV value s , biasing potential $V_i(s)$ and free energy shift $F_{0,i}$ from umbrella i . $F_{0,i}$ and $P(s)$ are the two unknown parameters which are determined by iterations of the WHAM equation.

7.5.3 Steered Molecular Dynamics Simulations

A similar method as the umbrella approach to overcome a priori known high energy barriers is Steered Molecular Dynamics (SMD) simulations. In this approach, a harmonic restraint is continuously moved during the simulation. In other words - SMD simulations

are moving umbrella simulations. The moving restraint pulls the system over the barrier of interest and is given by:

$$V(s(q, t)) = k/2(s(q) - s_0(t))^2$$

with a time-dependent center of the restraint s_0 of the CV s . By integration over the distance, the work of the simulated path can be computed. This work can be used to calculate the free energy difference using the Jarzynski relationship [236] which interconnects both values.

$$e^{-\frac{\Delta G}{k_B T}} = \left\langle e^{-\frac{W}{k_B T}} \right\rangle$$

This means that the average of the work W within the Boltzmann-factor is equal to the free energy difference ΔG within the Boltzmann-factor. The free energy difference is addressed between two states A and B, the start and stop points. The SMD simulation is a non-equilibrium simulation except in the case of an infinite slow pulling. The Jarzynski equation is particularly valid for this kind of non-equilibrium simulations.

The convergence of SMD simulations is crucially dependent on the path the system takes in each pulling simulation. Since molecular systems can easily span over several bonded atoms with many dihedral angles, the simulated traces can span over a wide variety. That is the reason why SMD simulations do often not converge easily even for hundreds of performed pull traces.

7.5.4 Metadynamics

In Metadynamics simulations, the applied bias which accelerates the sampling along the CV is adapted in its height during the simulation. In the presented target, umbrella or SMD simulation, the bias needs to be chosen *a priori*. In Metadynamics simulations, on the other hand, the system starts in an unbiased state. After a selected period of time, a repulsive bias potential, typically of the shape of a Gaussian, is applied at the current position of the CV. With the deposition stride of the Gaussians τ and a natural number k the bias potential, which is applied on the trajectory is given by [163, 237]:

$$V(s(q), t) = \sum_{k\tau < t} W(k\tau) \exp \left(- \sum_{i=1}^d \frac{[s_i - s_i(q(k\tau))]^2}{2\sigma_i^2} \right) \quad (7.8)$$

with the height and width of the Gaussian $W(k\tau)$ and σ_i respectively. d is the total number of the collective variables s_i . The longer the system is trapped at local free energy minima the more bias is added at this position in the collective variable space. Therefore, the native free energy landscape is compensated by the increasing bias potential over time. After the free energy landscape is compensated the system is diffusive along the complete phase space. Still more Gaussians are added so that the system fluctuates around a constant bias value dependent on the Gaussian height. Furthermore the total bias value is increasing, thus the added bias potential equals minus the free energy plus an offset constant:

$$V(s(q), t \rightarrow \infty)_{bias} = -F(s(q)) + C$$

To decrease the fluctuations of this method the Well-tempered approach is introduced by Barducci et al. [170]. In this method the Gaussian height is decreased with respect to the placed bias potential. This means that the Gaussian height decreases over time dependent on the total bias potential which reduces the fluctuations and increases the convergence of the method. The Gaussian height of the Well-tempered method is given by [163, 170]:

$$W(k\tau) = W_0 \exp\left(\frac{-V(s(q), k\tau)_{bias}}{k_B \Delta T}\right)$$

with the initial Gaussian height W_0 and ΔT which has the dimension of a temperature and is further described in the following. In the long time limit the potential smoothly converges to:

$$V(s(q), t \rightarrow \infty)_{bias} = -\frac{\Delta T}{T + \Delta T} F(s(q)) + C$$

with the temperature of the system T . The temperature along the CV is given by $T + \Delta T$. To adjust the decrease rate of the Gaussian height to the height of the expected free energy barrier a user defined ‘‘Biasfactor’’ γ is used which is given by:

$$\gamma = \frac{T + \Delta T}{T} \tag{7.9}$$

For $\Delta T \rightarrow \infty$ the simulation is equal to a standard Metadynamics simulation and for $\Delta T \rightarrow 0$ equal to a standard MD simulation.

In some cases, it can be useful to run multiple Metadynamics simulations of the same system which share the developing bias potential. This trick is called multiple walker Metadynamics and can increase the sampling significantly [175] (see Section 5.1.5).

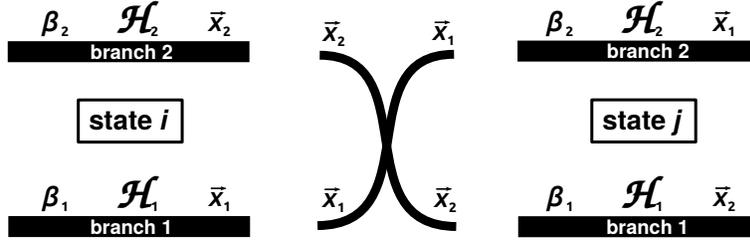


Figure 7.1: i, j are the states before and after the exchange of the replica coordinates $\bar{x}_{1,2}$ in Replica Exchange simulations. Presented are two replicas: branch 1 and 2 with individual $\beta_{1,2} = \frac{1}{k_B T_{1,2}}$ with the Boltzmann constant k_B and temperature $T_{1,2}$ as well as Hamiltonian function $\mathcal{H}_{1,2} = E_{1,2}^{pot} + T_{1,2}^{kin}$ with potential and kinetic energies.

7.5.5 Replica Exchange Methods

Another way to overcome the limitations of a standard MD simulation is the usage of several replicas of the system. The replicas usually include a “ground” replica at 300 K. The other replicas are made in the manner that energy barriers in the system are reduced. One concept is the usage of increasing temperatures (Temperature Replica Exchange Method (T-REM)) another is the reduction of the potential energy landscape (Hamiltonian Replica Exchange Method (H-REM)). Both approaches exchange the replicas after a particular simulation time.

Exchanged are the atom coordinates \vec{x} whereas the conditions of each replica branch stay constant. If the exchange satisfies the Detailed-Balance condition, the ground replica is still in a canonic ensemble, even though after exchanges with other branches [238, 239]. The Detailed-Balance condition implies that in the case of infinite sampling the number of exchanges in one direction is equal to the number in the reverse direction.

$$P_i P_{i \rightarrow j} = P_j P_{j \rightarrow i} \quad (7.10)$$

where $P_{i,j}$ is the probability that the replica is in a certain state i, j and $P_{i \rightarrow j, i \rightarrow j}$ the probability of the exchange between i, j . i is the state before the exchange and j after. By using the Boltzmann distribution with Equation 7.10 follows:

$$\begin{aligned} e^{-\beta_1 \mathcal{H}_1(\bar{x}_1)} e^{-\beta_2 \mathcal{H}_2(\bar{x}_2)} P_{i \rightarrow j} &= e^{-\beta_1 \mathcal{H}_2(\bar{x}_1)} e^{-\beta_2 \mathcal{H}_1(\bar{x}_2)} P_{j \rightarrow i} \\ \frac{P_{i \rightarrow j}}{P_{j \rightarrow i}} &= e^{\beta_1 [\mathcal{H}_1(\bar{x}_1) - \mathcal{H}_1(\bar{x}_2)] + \beta_2 [\mathcal{H}_2(\bar{x}_2) - \mathcal{H}_2(\bar{x}_1)]} \\ &=: e^{(-\Delta)} \end{aligned} \quad (7.11)$$

$\exp(-\Delta)$ is the acceptance ratio of the exchange. By assuming that the average kinetic energy in each branch is independent of the coordinate distribution $\vec{x}_{1,2}$ the kinetic energy is always canceling out in Equation 7.11 and the Hamiltonian function becomes $\mathcal{H}_{1,2} = E_{1,2}^{pot} = E_{1,2}$, leading to:

$$\frac{P_{i \rightarrow j}}{P_{j \rightarrow i}} = e^{\beta_1[E_1(\vec{x}_1) - E_1(\vec{x}_2)] + \beta_2[E_2(\vec{x}_2) - E_2(\vec{x}_1)]} =: e^{(-\Delta)} \quad (7.12)$$

With a general Metropolis criterion follows [240]:

$$P_{acceptance} = \min \{1, e^{-\Delta}\} \quad (7.13)$$

This criterion is initially introduced in Monte Carlo algorithms e.g. [239]. Equation 7.13 is used in the following way. If $e^{-\Delta}$ is bigger than 1 the exchange is always accepted. Otherwise, a random number between 0-1 is generated and only if the generated number is smaller than the probability $e^{-\Delta}$, the exchange is accepted.

Temperature Replica Exchange Method (T-REM)

In the most classic replica exchange method the replicas of the system run at different temperatures [241] whereas the potential energy functions of each replica are the same. The probability of the simulated system to cross energy barriers is proportional to the Boltzmann distribution $P(E) \propto e^{-\beta E}$ with $\beta = 1/k_B T$. It is distinctly increased for a given barrier energy E and higher temperature T . The potential energy at each branch is equal (Fig. 7.1), therefore is $E_{1,2}(\vec{x}_{1,2}) = E(\vec{x}_{1,2})$. Equation 7.12 becomes:

$$\frac{P_{i \rightarrow j}}{P_{j \rightarrow i}} = e^{\beta_1[E(\vec{x}_1) - E(\vec{x}_2)] + \beta_2[E(\vec{x}_2) - E(\vec{x}_1)]} = e^{[\beta_1 - \beta_2][E(\vec{x}_1) - E(\vec{x}_2)]} =: e^{(-\Delta)} \quad (7.14)$$

By using this acceptance ratio of the exchange in the criterion 7.13 an exchange acceptance can be explained as follows. T_1 is the temperature of β_1 and T_2 of β_2 . If $T_1 < T_2$ and $E_1 < E_2$ which is mostly the case than $\exp(-\Delta) < 1$. A random number between 0-1 is generated. If the generated number is smaller than the probability $\exp(-\Delta)$, the change is accepted. In the more unlikely case if $T_1 < T_2$ and $E_1 > E_2$ follows $\exp(-\Delta) > 1$, than the exchange is always accepted because the generated random number is always smaller than 1. Therefore, a high acceptance ratio is given when the potential energy distributions

of the replicas overlap. This is given by close temperature differences between the replicas.

Fukunishi et al. [242] used the criterion so that the energy fluctuation of a replica $\Delta E_{1,2}$ should be equal or bigger than the difference of the mean value of the replica energies $\bar{E}_1 - \bar{E}_2$ to get a sufficient exchange probability. Furthermore he estimated the replica temperatures by assuming that the energy fluctuations scale by $\Delta E_{1,2} \propto \sqrt{f}k_B T_{1,2}$ and the mean energy of a replica $\bar{E}_{1,2} \propto f k_B T_{1,2}$, with the degrees of freedom f . This results directly in a differential equation with the solution of the replica temperatures [242]:

$$T_m \propto e^{m/\sqrt{f}}$$

where m is the number of the replica. With the temperature range between T_{min} and T_{max} follows the number of replicas M :

$$M \propto \sqrt{f} \ln \left(\frac{T_{max}}{T_{min}} \right) \quad (7.15)$$

Therefore, the number of replicas scale with the square root of the degrees of freedom of the system.

A more detailed estimation of the replica temperatures can be found by Patriksson et al. [243] following the same scaling. By using the implemented online server with the protein chymotrypsin with about 3500 atoms and 21000 water molecules and a temperature range of 300 to 500 K the Patriksson estimation results in 144 replicas for an exchange probability of 20 %. A simulation with the bare protein, without the solvent, would contain only 23 replicas. Hence, the solvent has a very high influence on the degrees of freedom of the system and therefore also on the exchange probability. Therefore, a standard T-REM scales poorly with the system size and is not suitable for bigger biomolecules like proteins, especially in an explicit solvent.

Hamiltonian Replica Exchange Method (H-REM)

In the H-REM method, the potential energy functions of the replicas are different. The potential energy of the ground replica is not scaled and is identical with the T-REM ground replica. All other replicas can be treated by choice of the user. Typically, parts of the force field equation are scaled to lower potential energy barriers. In T-REM simulations, the kinetic energy of the system is increased by higher temperatures while in H-REM, the potential energy barriers are reduced. As long as the Detailed-Balance condition Eq. 7.10

of the exchanges are satisfied, the ground replica is still in a canonical ensemble. In most of the cases, the potential energy of the solute but not the solvent is scaled, which decreases the exchange probability of the replicas significantly as shown in the following section.

Using Equation 7.12 with $\beta_1 = \beta_2 = \beta$, because the temperatures of all replicas are identical, the exchange probability becomes:

$$\frac{P_{i \rightarrow j}}{P_{j \rightarrow i}} = e^{\beta\{[E_1(\vec{x}_1)+E_2(\vec{x}_2)]-[E_1(\vec{x}_2)+E_2(\vec{x}_1)]\}} =: e^{(-\Delta)} . \quad (7.16)$$

As first described by Fukunishi et al. [242]. The Metropolis like criteria of Equation 7.13 is used here as well.

Replica Exchange with Solute Tempering (REST)

Based on Fukunishi et al. [242], Liu et al. [240] suggested to divide the potential energy of the system in a central group (e.g. the protein p) and the bath (e.g. the water w). The total potential energy of replica m becomes:

$$E_m = E_{pp} + \left[\frac{\beta_0}{\beta_m} \right] E_{ww} + \left[\frac{\beta_0 + \beta_m}{2\beta_m} \right] E_{pw}$$

with the internal interactions of the central group of the system E_{pp} , the internal interactions of the bath E_{ww} and the interaction between both parts E_{pw} . $\beta_0 = 1/k_B T_0$ with the temperature T_0 of the ground replica. $\beta_m = 1/k_B T_m$ with the temperature T_m of the higher replicas. A thermostat is acting on the whole system and is set to the increasing target temperatures for the replicas as in the case of T-REM. Hence, the thermostat temperature is different for each replica. The scaling by the temperature dependent factors of the individual parts are keeping the bath (solvent) cold whereas the solute is exposed to the higher temperatures. Thereby the method is called replica exchange with solute tempering (REST).

The shown potential energy E_m plugged in the exchange probability Equation 7.16 indicate that the E_{ww} part is canceling out. This results in the tremendous advantage of this method that the degrees of freedom of the solvent do not enter in the exchange probability which highly increases the Metropolis exchanges. In other words, the strong fluctuations of the bath molecules are neglected. This reduction of degrees of freedom

leads to an increased overlap of the energy distribution of neighboring replicas. Therefore, much fewer replicas compared to T-REM are needed.

In an alanine dipeptide system, Liu et al. [240] estimated an eight times more efficient sampling of REST compared to standard T-REM.

Wang et al. [167] presented a slightly different method which is also used in this work. It is called “REST2” approach, in comparison to the above mentioned “REST1” method. The scaling of the potential energy of replica m becomes:

$$E_m = \left[\frac{\beta_m}{\beta_0} \right] E_{pp} + E_{ww} + \sqrt{\left[\frac{\beta_m}{\beta_0} \right]} E_{pw} \quad (7.17)$$

Next to the different scaling as shown in Equation 7.17 the defining difference is that the thermostat of all replicas is set to the ground temperature T_0 . In contrast to REST1, in REST2, the increased solute temperatures of the replicas are implemented by lowering the potential energies of the solute by a temperature dependent factor. The temperature in the factor β is no longer a real temperature in terms of a thermostat parameter and is called effective temperature instead. By using the exchange probability Equation 7.16 the E_{ww} part is canceling out as well, but the exchange probability varies to REST1 which results in a more efficient sampling of REST2 compared to REST1 [167].

The atoms which are scaled by the presented solute tempering approach are defined as “hot” [169]. In an alanine dipeptide system e.g. all atoms of the peptide would be defined as “hot” whereas the water is defined as “cold”. The presented REST2 scaling in Equation 7.17 is satisfied if the force field parameters are scaled as follows [167, 169]:

- Partial charges of atoms in the ”hot” region by $\sqrt{\lambda}$
- Lennard-Jones parameter ϵ of atoms in the ”hot” region by λ
- Dihedral force constant: if atom 1 and atom 4 are in the ”hot” region by λ
- Dihedral force constant: if atom 1 or atom 4, not both, are in the ”hot” region by $\sqrt{\lambda}$

with the quotient $\lambda = \frac{T_m}{T_0}$. T_0 is the temperature of the ground replica whereas T_m is the temperature of a higher replica m . Hence, the quotient λ is always smaller one. T is defined as an effective temperature [169] since it is imitated by the scaling not by a thermostat as mentioned.

Table 7.1: Exchange probabilities of 14 replicas. The complete protein chymotrypsin is scaled by the REST2 protocol from 300 to 600K. The system contains 80 000 atoms with 3500 in the "hot" region.

Repl.	0/1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11	11/12	12/13
%	3	3	3	4	4	5	5	5	5	4	6	5	6

We are going to use this method to accelerate the sampling of a part of a protein as presented in Chapter 5. This part of the protein is accelerated by a REST2 scaling and therefore defined as "hot". In the following, it is shown that the free energy sampling of a complete protein of the size of chymotrypsin can be extremely challenging and that the reduction of only a part of the protein is required. Exemplarily we analyze and discuss a simulation in which all atoms of chymotrypsin with about 3500 atoms are scaled. The degrees of freedom of this system is crucially influencing the numbers of required replicas as shown in Equation 7.15.

We have used 14 replicas with the following effective temperatures 300, 316, 334, 352, 371, 392, 413, 436, 460, 485, 511, 539, 569, 600 K. The temperatures are calculated by following a geometric progression due to the wider energy distribution of higher temperatures [186]. The resulting exchange probabilities are shown in Tab. 7.1.

The average exchange probability over all replicas is 4.4 %. The ability of the replicas moving along the replica ladder can be described by the round trip time (rtt). This is the time one replica goes up and down the complete ladder. Due to the very small exchange probability, not even one replica performs a full rtt cycle in the performed simulation time of 7.6 ns. Hence the rtt value is > 7.6 ns.

The small exchange probabilities are based on huge energy differences between the replicas (Equation 7.12). Due to the high number of atoms in the scaled region, small effective temperature distances between the replicas are necessary to get high exchange probabilities. Therefore, a large number of replicas are required to increase the exchange probabilities which on the other hand decreases the round trip time due to the long replica ladder. One solution is simply the reduction of the atoms in the scaled region as shown in the following.

Now, only one helix of chymotrypsin is scaled by the REST2 approach, as also presented in detail in Chapter 5. This method can be associated with an increased sampling ability of the helix structure and a standard MD of the rest of the system. This helix contains

Table 7.2: Exchange probabilities of 12 replicas. One helix of chymotrypsin with 257 atoms is scaled by the REST2 protocol from 300 to 700K.

Repl.	0/1	1/2	2/3	3/4	4/5	5/6	6/7	7/8	8/9	9/10	10/11
%	44	45	46	47	47	48	49	50	51	52	53

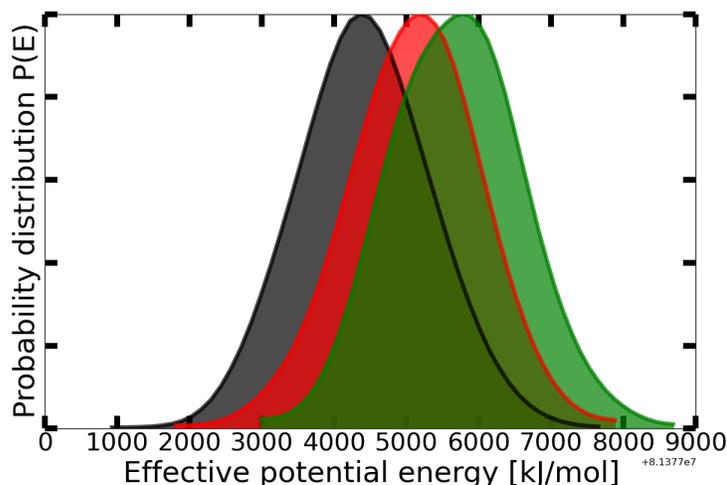


Figure 7.2: Normalized effective potential energy of the first (gray), a middle (red) and the last (green) replica of a 12 replica simulation in which one helix of chymotrypsin is scaled by the REST2 approach.

14 amino acids with 257 atoms. Here 12 replicas from 500 to 700 K are used with the effective temperatures: 300, 324, 350, 378, 408, 441, 476, 514, 556, 600, 648, 700 K. The exchange probabilities are shown in Table 7.2.

The average exchange probability over all replicas is 48.4 %. The average round trip time of the replicas is 0.5 ± 0.1 ns. The normalized effective potential energy distribution of this simulation is shown in Fig. 7.2. Shown are the first, a middle and the last replica. A huge overlap of the potential energy causes the high exchange probabilities of this system.

In summary, the scaled number of atoms in the system crucially influence the exchange probabilities of the REST simulations. With a replica number around ten, the simulated systems with a size of oligopeptides (in solution or as part of a bigger molecule) allow a good diffusion along the replica ladder.

Replica Exchange with Flexible Tempering (REFT)

The REFT method is technically identical to the REST2 approach but is applied in a special way. Identical to the nomenclature used in Section 7.5.5, atoms with a scaled potential are called “hot” whereas the non-scaled atoms are called “cold”. In the REFT approach, the molecule, e.g. a protein, is divided in a flexible “hot” and a rigid “cold” part. Moors et al. [244] presented this method on T4 lysozyme and set the random coil secondary structure “hot” and left the helices and sheets as well as the water “cold”. The hypotheses of Moors is that the conformational change along a particular reaction coordinate is achieved by a hinge-like-movement of the rigid helix and sheet structures around the flexible random like secondary structures. In the REFT approach, the hinge-like-movement is even accelerated by setting the hinge-like parts, the coils, “hot”. In the case of T4 lysozyme Moors found an increased sampling of the opening and closing of the binding cleft by REFT compared to a REST2 scaling of the whole protein. The increased sampling is explained by the fact that in the REFT approach only the degrees of freedom necessary for the opening and closing of the binding cleft is involved. In REST2 more degrees of freedoms are involved not directly related to the opening and closing [167,244].

Replica Exchange Solute Tempering with Metadynamics (RESTmetaD)

The combination of both methods complements each other. The Metadynamics samples moderately high barriers along a preselected CV. However, slow degrees of freedom not reached by the CV can partly be compensated by the solute tempering which is independent of the chosen CV.

The effective temperatures of the REST approach reaches typically from about 300 - 450 K (Schneider et al. [55] and Meißner et al. [123]), 270 to 695 K (Bussi et al. [168]). Whereas the CV temperature ($\Delta T + T$) (Section 5.1.1) calculated by Equation 7.9 is typically 3000 K (Schneider et al. [55] and Meißner et al. [123]). These high temperatures along the CV provided by the Metadynamics demonstrate the sampling potential compared to REST alone.

Replica Exchange with Collective Variable Tempering (RECT)

Another approach of free energy sampling methods is the Replica Exchange with Collective Variable Tempering (RECT) introduced by Gil-Ley et al. [171]. The method includes

a different sampling acceleration of the replicas compared to solute tempering. An individual Well-tempered Metadynamics is applied on each replica. The central idea of this approach is the application of different bias factors on the replicas. In particular, the Well-tempered Gaussian heights decrease slowly for high bias factors (top replica) and drop fast for low bias factors (ground replica)(Section 5.1.1). Therefore, high energy barriers are crossed by the top replicas, whereas the ground replica samples fine free energy fluctuations. The method allows the usage of several CV within the replicas. Each replica contains the same CVs with individually applied Metadynamics. Therefore, e.g. all dihedral angles of a small peptide and additional system-specific CVs can be applied on each replica. The number of CVs are only limited by the calculation of the CV itself, which could slow down the simulation in the case of a complex computed CV.

7.6 Radial Distribution Function (RDF)

The radial distribution function (RDF) $g(r)$ calculates the probability to find a chosen pair of particles, atoms or molecules at a distance r . In this way, $g(r)$ displays the statistical distribution of the local neighbors of the species of interest. $g(r)$ is only a function of the distance, the analysis, however, is performed in all space dimensions. Therefore, $g(r)$ is a transformation of a typically three-dimensional distribution to one dimension, the distance between the pairs. This approach is a useful tool to analyze local order phenomena, like dense water structures around polar groups as performed in our case. Another application is the analysis of the grade of an amorphous material since the RDF displays the crystalline near-order and long-range disorder. To calculate $g(r)$, the algorithm searches in a shell around one species for the chosen pair and counts each match in a histogram. The shell spans from r to $r + dr$ and its volume can be approximated with $4\pi r^2 dr$ for small dr . Consequently, the RDF is defined as [245]:

$$g(r) = \lim_{dr \rightarrow 0} \frac{\rho(r)}{4\pi r^2 dr} \frac{V_{total}}{N_{all-pairs}}$$

with $\rho(r)$ as the number of pairs found in the shell which spans from r to $r + dr$. V_{total} is the total volume of the system and $N_{all-pairs}$ all possible pair between the two species of interest. By this means $g(r)$ is a density normalized by the total density of all pairs per volume $\frac{N_{all-pairs}}{V_{total}}$ and therefore unitless.

7.7 Dipole Moment

The dipole moment visualizes and quantifies the separation of positive and negative charges in the system. It is used for systems of at least two partial charges. It summarizes the general trend of the charge distribution in two dimensions. In the case of a discrete charge distribution, the dipole moment \vec{d} is defined as:

$$\vec{d} = \sum_{i=1}^N q_i \vec{r}_i$$

with the sum of all partial charges N with the individual charge q_i and the vector to this charge \vec{r}_i . Per definition of this equation, it points from the negatively to the positively charged part. To describe this interestingly short equation in words, first, assume only charges of equal amount q but with different signs $\pm q$. The sum forms one vector pointing from all negative to all positive charges, by that it also “weights” with their distance to each other. In the case of different charges q , the sum “weights” each partial charge point also by its charge value. In this way, high charges far away from the center of the charge distribution influence the dipole moment the most.

7.8 Cluster Analysis

The cluster analysis is a helpful data mining tool to visualize similarities of structures in trajectories. The RMSD clustering sorts structures by their RMSD differences. The rigid structures are aligned by translation and rotation to their minimal RMSD difference.

An RMSD cluster analysis can be performed by using many different algorithms and distance metrics. A review of many different approaches is reported by Shao et al. [246]. In this work, the GROMOS nearest neighbors algorithm is used [196], which is included in the GROMACS package [164]. This algorithm calculates the RMSD differences between all frames, after alignment. Subsequently, the number of frames within a chosen cut-off is counted, and the cluster with the largest number of frames is defined as the first found cluster. The remaining frames define the new pool of structures. From this pool, the most prominent cluster is extracted again and form the second found cluster. The procedure is repeated until no frames are left.

Publications

In the following, all peer-reviewed publications are named with a contribution of Nils Hildebrand the author of this dissertation which are originated during his time as a PhD student at the University of Bremen:

A. Aminian, B. Shirzadi, Z. Azizi, K. Maedler, E. Volkmann, N. Hildebrand, M. Maas, L. Treccani, K. Rezwan, *Enhanced cell adhesion on bioinert ceramics mediated by the osteogenic cell membrane enzyme alkaline phosphatase*, Materials Science and Engineering C 69, 184-194 (2016)

L. Derr, N. Hildebrand, S. Köppen, S. Kunze, L. Treccani, R. Dringen, K. Rezwan, L. Colombi Ciacchi, *Physisorption of α -chymotrypsin on SiO_2 and TiO_2 : A comparative study via experiments and molecular dynamics simulations* *Biointerphases*, 11, 011007 (2016)

L. Derr, R. Dringen, L. Treccani, N. Hildebrand, L. Colombi Ciacchi, K. Rezwan, *Physisorption of enzymatically active chymotrypsin on titania colloidal particles*, Journal of Colloid and Interface Science 455, 236-244 (2015)

N. Hildebrand, S. Köppen, L. Derr, K. Li, M. Koleini, K. Rezwan, L. Colombi Ciacchi, *Adsorption orientation and binding motifs of lysozyme and chymotrypsin on amorphous silica*, Journal of Physical Chemistry C 119, 7295-7307 (2015)

In the dissertation the results from the supervision of the following students works' are included:

Simon Kunze, Bachelor thesis: *Molekulardynamische Simulation der Adsorption von Enzymen an Oxidoberflächen im Kontext der Entwicklung von biofunktionellen Hybridmaterialien*, Faculty of production engineering, University of Bremen, 2014

Contribution: Production of the data for Fig. 3.1 and development of the charge distribution of the TiO₂ surface.

Nina Wurzler, Master thesis: *Untersuchungen der Struktur und des Adsorptionsverhaltens von Biomolekülen auf Nanopartikeln mittels Zirkulardichroismus-Spektroskopie*, Faculty of production engineering, University of Bremen, 2015

Contribution: Production of the data for Fig. 1.9.

Johannes Stoedter-Rosien, Bachelor thesis: *Computergestützte Vorhersage einer optimalen Oberflächenfunktionalisierung eines keramischen Virenfilters*, Faculty of production engineering, University of Bremen, 2017

Contribution: Classical Molecular Dynamics simulations of the CC-dimer capsid protein on an amino silane functionalized silica surface and docking simulations of silane compounds on the CC-dimer as described in Section 6.2.3.

Acknowledgements

Thanks, Lucio that you have recruited me straight away and given me the trust to work in your group. Your support has always been constructive, direct and open minded. I have profited not only scientifically but also from you personally and the atmosphere you have created in the group. The past years have been great!

Thanks, Susi for all your support scientifically but also regarding any other circumstances. You have offered me help no matter whether you have been on parental leave or right before the birth of your child. Thanks a lot.

Jens, it has been an honor to share the office with you for the last years. I have loved our small challenges to be more productive or to finish tasks in time. We had a very constructive working atmosphere with the whiteboard always full of current scientific problems from one of us. I have worked together with a friend not a colleague.

Thanks, Meike and Ilona, for your warm, professional and incredibly quick help in any issues we have bothered you with.

Thanks, Stefan and Steffen for solving most of the computer problems in no time and providing us always with well maintained hard- and software.

I thank all the people of the HMI group I have worked with. We had an excellent time together also besides work. Thanks, Robert, Monika, Gang, Felipe, Barbara, Lena, Magdalena, Qing, Jingfeng, Johannes, Laura, Julen, Tatjana and Sascha.

Thanks to Massimo, Jens, Sebastian, Steffen and Gesine for proofreading chapters of this work.

Thanks to Prof. Dirk Zahn for agreeing to act as the co-examiner.

Thanks to all cooperation partners I have had the pleasure to work with, in particular: Ludmilla Derr, Kurosch Rezwan, Laura Treccani, Julia Bartels, Stephen Kroll, Ralph Dringen, Nils Kröger, Nidhi Dubey, Giovanni Bussi and Alejandro Gil-Ley.

This work was supported by the Deutsche Forschungsgemeinschaft under grants KO 3811/3-1 and CI 144/2-1. Computational resources were provided by the North-German Supercomputing Alliance system (HLRN).

Finally, I thank my family, friends and most of all Gesine for all the support they have given me. Thanks, that I can always rely on you.

Bibliography

- [1] S. Rana, A. Bajaj, R. Mout, and V. M. Rotello. Monolayer coated gold nanoparticles for delivery applications. *Advanced Drug Delivery Reviews*, 64(2):200–216, 2012.
- [2] R. Mout, D. F. Moyano, S. Rana, and V. M. Rotello. Surface functionalization of nanoparticles for nanomedicine. *Chemical Society Reviews*, 41(7):2539, 2012.
- [3] I. H. El-Sayed, X. Huang, and M. A. El-Sayed. Surface plasmon resonance scattering and absorption of anti-EGFR antibody conjugated gold nanoparticles in cancer diagnostics: Applications in oral cancer. *Nano Letters*, 5(5):829–834, 2005.
- [4] R. Elghanian. Selective colorimetric detection of polynucleotides based on the distance-dependent optical properties of gold nanoparticles. *Science*, 277(5329):1078–1081, 1997.
- [5] E. Wintermantel and S.-W. Ha. *Medizintechnik: life Science Engineering (German Edition)*. Springer, 2009.
- [6] P. Thevenot, W. Hu, and L. Tang. Surface chemistry influence implant biocompatibility. *Current Topics in Medicinal Chemistry*, 8(4):270–280, 2008.
- [7] C. J. Wilson, R. E. Clegg, D. I. Leavesley, and M. J. Pearcy. Mediation of biomaterial–cell interactions by adsorbed proteins: a review. *Tissue Engineering*, 11(1-2):1–18, 2005.
- [8] M. Shen, I. Garcia, R. V. Maier, and T. A. Horbett. Effects of adsorbed proteins and surface chemistry on foreign body giant cell formation, tumor necrosis factor alpha

- release and procoagulant activity of monocytes. *Journal of Biomedical Materials Research*, 70A(4):533–541, 2004.
- [9] D. A. Pankowsky, N. P. Ziats, N. S. Topham, O. D. Ratnoff, and J. M. Anderson. Morphologic characteristics of adsorbed human plasma proteins on vascular grafts and biomaterials. *Journal of Vascular Surgery*, 11(4):599–606, 1990.
- [10] J. D. Andrade and V. Hlady. Plasma protein adsorption: the big twelve. *Annals of the New York Academy of Sciences*, 516(1):158–172, 1987.
- [11] S.-H. Teoh. *Engineering Materials for Biomedical Applications (Biomaterials Engineering and Processing Series)*. World Scientific Publishing Company, 2004.
- [12] N. Völcker, D. Klee, H. Höcker, and S. Langefeld. Functionalization of silicone rubber for the covalent immobilization of fibronectin. *Journal of Materials Science: Materials in Medicine*, 12(2):111–119, 2001.
- [13] N. Nath, J. Hyun, H. Ma, and A. Chilkoti. Surface engineering strategies for control of protein and cell interactions. *Surface Science*, 570(1-2):98–110, 2004.
- [14] A. Higuchi, K. Shirano, M. Harashima, B. O. Yoon, M. Hara, M. Hattori, and K. Imamura. Chemically modified polysulfone hollow fibers with vinylpyrrolidone having improved blood compatibility. *Biomaterials*, 23(13):2659–2666, 2002.
- [15] D. E. MacDonald, N. Deo, B. Markovic, M. Stranick, and P. Somasundaran. Adsorption and dissolution behavior of human plasma fibronectin on thermally and chemically modified titanium dioxide particles. *Biomaterials*, 23(4):1269–1279, 2002.
- [16] B. G. Keselowsky, D. M. Collard, and A. J. García. Surface chemistry modulates focal adhesion composition and signaling through changes in integrin binding. *Biomaterials*, 25(28):5947–5954, 2004.
- [17] D. R. Lu and K. Park. Effect of surface hydrophobicity on the conformational changes of adsorbed fibrinogen. *Journal of Colloid and Interface Science*, 144(1):271–281, 1991.
- [18] T. O. Collier and J. M. Anderson. Protein and surface effects on monocyte and macrophage adhesion, maturation, and survival. *Journal of Biomedical Materials Research*, 60(3):487–496, 2002.

- [19] M. Lestelius, B. Liedberg, and P. Tengvall. In vitro plasma protein adsorption on ω -functionalized alkanethiolate self-assembled monolayers. *Langmuir*, 13(22):5900–5908, 1997.
- [20] S. Kamath, D. Bhattacharyya, C. Padukudru, R. B. Timmons, and L. Tang. Surface chemistry influences implant-mediated host tissue responses. *Journal of Biomedical Materials Research Part A*, 86A(3):617–626, 2008.
- [21] C. D. Tidwell, S. I. Ertel, B. D. Ratner, B. J. Tarasevich, S. Atre, and D. L. Allara. Endothelial cell growth and protein adsorption on terminally functionalized, self-assembled monolayers of alkanethiolates on gold. *Langmuir*, 13(13):3404–3413, 1997.
- [22] M. Lindblad. Cell and soft tissue interactions with methyl- and hydroxyl-terminated alkane thiols on gold surfaces. *Biomaterials*, 18(15):1059–1068, 1997.
- [23] E. Ostuni, R. G. Chapman, R. E. Holmlin, S. Takayama, and G. M. Whitesides. A survey of structure-property relationships of surfaces that resist the adsorption of protein. *Langmuir*, 17(18):5605–5620, 2001.
- [24] T. Zoungrana, G. H. Findenegg, and W. Norde. Structure, stability, and activity of adsorbed enzymes. *Journal of Colloid and Interface Science*, 190(2):437–448, 1997.
- [25] R. Cegielska-Radziejewska, G. Leńnierowski, and J. Kijowski. Properties and application of egg white lysozyme and its modified preparations - a review. *Polish Journal Of Food And Nutrition Sciences*, 58(1):5–10, 2008.
- [26] R. Ulbrich, R. Golbik, and A. Schellenberger. Protein adsorption and leakage in carrier-enzyme systems. *Biotechnology and Bioengineering*, 37(3):280–287, 1991.
- [27] O. A. Mogil'naya and V. S. Bondar. Comparative study of antibacterial properties of lysozyme upon its adsorption and covalent binding to nanodiamonds. *Nanotechnologies in Russia*, 7(11-12):658–665, 2012.
- [28] S. Landgraeber, M. Jäger, J. J. Jacobs, and N. J. Hallab. The pathology of orthopedic implant failure is mediated by innate immune system cytokines. *Mediators of Inflammation*, 2014:1–9, 2014.

- [29] R. Smeets, A. Henningsen, O. Jung, M. Heiland, C. Hammächer, and J. M. Stein. Definition, etiology, prevention and treatment of peri-implantitis – a review. *Head & Face Medicine*, 10(1), 2014.
- [30] M. Vallet-Regí and F. Balas. Silica materials for medical applications. *The Open Biomedical Engineering Journal*, 2(1):1–9, 2008.
- [31] S. B. Goodman, Z. Yao, M. Keeney, and F. Yang. The future of biologic coatings for orthopaedic implants. *Biomaterials*, 34(13):3174–3183, 2013.
- [32] W. S. Price, F. Tsuchiya, and Y. Arata. Lysozyme aggregation and solution properties studied using PGSE NMR diffusion measurements. *Journal of the American Chemical Society*, 121(49):11503–11512, 1999.
- [33] R. E. Canfield. The amino acid sequence of egg white lysozyme. *Journal of Biological Chemistry*, 238(8):2698–2707, 1963.
- [34] N. Hildebrand, S. Köppen, L. Derr, K. Li, M. Koleini, K. Rezwan, and L. Colombi Ciacchi. Adsorption orientation and binding motifs of lysozyme and chymotrypsin on amorphous silica. *The Journal of Physical Chemistry C*, 119(13):7295–7307, 2015.
- [35] N. Ui. Isoelectric points and conformation of proteins. *Biochimica et Biophysica Acta (BBA) - Protein Structure*, 229(3):582–589, 1971.
- [36] S. N. Timasheff and K. C. Aune. Dimerization of α -chymotrypsin. I. pH dependence in the acid region. *Biochemistry*, 10(9):1609–1617, 1971.
- [37] H. Tsukada and D. M. Blow. Structure of α -chymotrypsin refined at 1.68 Å resolution. *Journal of Molecular Biology*, 184(4):703–711, 1985.
- [38] H. Kitano, Y. Maeda, and T. Okubo. Kinetic study of the effects of solvation on the dimerization process of α -chymotrypsin. *Biophysical Chemistry*, 33(1):47–54, 1989.
- [39] L. Derr, R. Dringen, L. Treccani, N. Hildebrand, L. C. Ciacchi, and K. Rezwan. Physisorption of enzymatically active chymotrypsin on titania colloidal particles. *Journal of Colloid and Interface Science*, 455:236–244, 2015.

- [40] L. Derr. *Interaction between enzymes and oxide colloidal particles and their influence on enzymatic activity*. PhD thesis, Universität Bremen: Produktionstechnik, University of Bremen, 2016.
- [41] U. Hilleringmann. *Silizium-Halbleitertechnologie: Grundlagen mikroelektronischer Integrationstechnik (German Edition)*. Vieweg+Teubner Verlag, 2008.
- [42] J. Sonnefeld, A. Göbel, and W. Vogelsberger. Surface charge density on spherical silica particles in aqueous alkali chloride solutions. *Colloid and Polymer Science*, 273(10):926–931, 1995.
- [43] P. M. Dove and C. M. Craven. Surface charge density on silica in alkali and alkaline earth chloride electrolyte solutions. *Geochimica et Cosmochimica Acta*, 69(21):4963–4970, 2005.
- [44] M. Kobayashi, F. Juillerat, P. Galletto, P. Bowen, and M. Borkovec. Aggregation and charging of colloidal silica particles: Effect of particle size. *Langmuir*, 21(13):5761–5769, 2005.
- [45] A. P. Legrand, Y. Grillet, A. Burneau, E. Papirer, and B. Fubini. *The Surface Properties of Silicas*. John Wiley and Sons, Inc., New York, 1998.
- [46] S. Köppen and W. Langel. Simulation of the interface of (100) rutile with aqueous ionic solution. *Surface Science*, 600(10):2040–2050, 2006.
- [47] D. E. Yates and T. W. Healy. Titanium dioxide-electrolyte interface. Part 2. - Surface charge (titration) studies. *Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases*, 76(0):9, 1980.
- [48] J. P. Holmberg, E. Ahlberg, J. Bergenholtz, M. Hassellöv, and Z. Abbas. Surface charge and interfacial potential of titanium dioxide nanoparticles: experimental and theoretical investigations. *Journal of Colloid and Interface Science*, 407:168–176, 2013.
- [49] G. D. Panagiotou, T. Petsi, K. Bourikas, C. S. Garoufalis, A. Tsevis, N. Spanos, C. Kordulis, and A. Lycourghiotis. Mapping the surface (hydr)oxo-groups of titanium oxide and its interface with an aqueous solution: the state of the art and a new approach. *Advances in Colloid and Interface Science*, 142(1-2):20–42, 2008.

- [50] C. E. Giacomelli, M. J. Avena, and C. P. De Pauli. Aspartic acid adsorption onto TiO₂ particles surface. experimental data and model calculations. *Langmuir*, 11(9):3483–3490, 1995.
- [51] D. J. Cole, M. C. Payne, G. Csányi, S. M. Spearing, and L. Colombi Ciacchi. Development of a classical force field for the oxidized Si surface: application to hydrophilic wafer bonding. *The Journal of Chemical Physics*, 127(20):204704, 2007.
- [52] A. Butenuth, G. Moras, J. Schneider, M. Koleini, S. Köppen, R. Meißner, L. B. Wright, T. R. Walsh, and L. Colombi Ciacchi. Ab initio derived force-field parameters for molecular dynamics simulations of deprotonated amorphous-SiO₂/water interfaces. *Physica Status Solidi B*, 249(2):292–305, 2011.
- [53] M. Kosmulski. *Chemical Properties of Material Surfaces*. CRC Press, 2001.
- [54] J. Laskowski and J. Kitchener. The hydrophilic-hydrophobic transition on silica. *Journal of Colloid and Interface Science*, 29(4):670–679, 1969.
- [55] J. Schneider and L. Colombi Ciacchi. Specific material recognition by small peptides mediated by the interfacial solvent structure. *Journal of the American Chemical Society*, 134(4):2407–2413, 2012.
- [56] N. N. Greenwood. *Chemistry of the Elements*. Elsevier Science Ltd, 1997.
- [57] F. Ullmann. *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH: Weinheim, 2001.
- [58] L. Mädler, H. Kammler, R. Mueller, and S. Pratsinis. Controlled synthesis of nanostructured particles by flame spray pyrolysis. *Journal of Aerosol Science*, 33(2):369–389, 2002.
- [59] S. Salameh, J. Schneider, J. Laube, A. Alessandrini, P. Facci, J. W. Seo, L. C. Ciacchi, and L. Mädler. Adhesion mechanisms of the contact interface of TiO₂ nanoparticles in films and aggregates. *Langmuir*, 28(31):11457–11464, 2012.
- [60] C. Oviedo. Oxidation kinetics of pure titanium at low pressures. *Journal of Physics: Condensed Matter*, 5(33A):A153, 1993.

- [61] Y. Takakuwa, S. Ishidzuka, A. Yoshigoe, Y. Teraoka, Y. Yamauchi, Y. Mizuno, H. Tonda, and T. Homma. Real-time monitoring of oxidation on the Ti (0001) surface by synchrotron radiation photoelectron spectroscopy and RHEED-AES. *Applied Surface Science*, 216(1-4):395–401, 2003.
- [62] J. Schneider. *Molecular Dynamics Simulations of Biological Molecules on the Natively Oxidized Titanium Surface*. PhD thesis, Universität Bremen: Produktionstechnik, University of Bremen, 2011.
- [63] T. Hiemstra, J. De Wit, and W. Van Riemsdijk. Multisite proton adsorption modeling at the solid/solution interface of (hydr)oxides: a new approach. *Journal of Colloid and Interface Science*, 133(1):105–117, 1989.
- [64] S. Kunze, *Molekulardynamische Simulation der Adsorption von Enzymen an Oxidoberflächen im Kontext der Entwicklung von biofunktionellen Hybridmaterialien*, Bachelor thesis, University of Bremen, 2014.
- [65] L. Derr, N. Hildebrand, S. Köppen, S. Kunze, L. Treccani, R. Dringen, K. Rezwan, and L. Colombi Ciacchi. Physisorption of α -chymotrypsin on SiO₂ and TiO₂: a comparative study via experiments and molecular dynamics simulations. *Biointerphases*, 11(1):011007, 2016.
- [66] B. Derjaguin and L. Landau. Theory of the stability of strongly charged lyophobic sols and of the adhesion of strongly charged particles in solutions of electrolytes. *Progress in Surface Science*, 43(1-4):30–59, 1993.
- [67] E. J. W. Verwey. Theory of the stability of lyophobic colloids. *The Journal of Physical Chemistry*, 51(3):631–636, 1947.
- [68] I. C. Bourg and G. Sposito. Molecular dynamics simulations of the electrical double layer on smectite surfaces contacting concentrated mixed electrolyte (NaCl - CaCl₂) solutions. *Journal of Colloid and Interface Science*, 360(2):701–715, 2011.
- [69] K. Hayes and J. Leckie. Modeling ionic strength effects on cation adsorption at hydrous oxide/solution interfaces. *Journal of Colloid and Interface Science*, 115(2):564–572, 1987.

- [70] J. N. Israelachvili. *Intermolecular and Surface Forces, Third Edition: revised Third Edition*. Academic Press, London, 2011.
- [71] M. Gouy. Sur la constitution de la charge électrique à la surface d'un électrolyte. *Journal de Physique Théorique et Appliquée*, 9(1):457–468, 1910.
- [72] D. L. Chapman. A contribution to the theory of electrocapillarity. *Philosophical Magazine Series 6*, 25(148):475–481, 1913.
- [73] D. C. Grahame. The electrical double layer and the theory of electrocapillarity. *Chemical Reviews*, 41(3):441–501, 1947.
- [74] S. Alvarez. A cartography of the van der waals territories. *Dalton Transactions*, 42(24):8617, 2013.
- [75] J. H. Dusenbury. Book review: the collected papers of peter j. w. debye. new york, interscience publishers, inc. *Textile Research Journal*, 24(10):940–940, 1954.
- [76] D. Case, D. Cerutti, T. Cheatham, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, Y. D.M., and P. Kollman. *AMBER 2015*. University of California, San Francisco, 2015.
- [77] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. *The Journal of Physical Chemistry*, 100(51):19824–19839, 1996.
- [78] G. D. Hawkins, C. J. Cramer, and D. G. Truhlar. Pairwise solute descreening of solute charges from a dielectric medium. *Chemical Physics Letters*, 246(1-2):122–129, 1995.
- [79] A. Onufriev, D. Bashford, and D. A. Case. Modification of the generalized born model suitable for macromolecules. *The Journal of Physical Chemistry B*, 104(15):3712–3720, 2000.

- [80] A. Onufriev, D. Bashford, and D. A. Case. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics*, 55(2):383–394, 2004.
- [81] D. Case, D. Cerutti, T. Cheatham, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. Roe, A. Roitberg, C. Sagui, C. Simmerling, W. Botello-Smith, J. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, L. Xiao, Y. D.M., and P. Kollman. *AMBER 12 Reference Manual*. University of California, San Francisco, 2012.
- [82] W. C. Still, A. Tempczyk, R. C. Hawley, and T. Hendrickson. Semianalytical treatment of solvation for molecular mechanics and dynamics. *Journal of the American Chemical Society*, 112(16):6127–6129, 1990.
- [83] J. Srinivasan, M. W. Trevathan, P. Beroza, and D. A. Case. Application of a pairwise generalized born model to proteins and nucleic acids: inclusion of salt effects. *Theoretical Chemistry Accounts: Theory, Computation, and Modeling (Theoretica Chimica Acta)*, 101(6):426–434, 1999.
- [84] B. Mueller, M. Zacharias, and K. Rezwan. Bovine serum albumin and lysozyme adsorption on calcium phosphate particles. *Advanced Energy Materials*, 12(1-2):B53–B61, 2010.
- [85] D. M. Togashi, A. G. Ryder, and G. Heiss. Quantifying adsorbed protein on surfaces using confocal fluorescence microscopy. *Colloids and Surfaces B: Biointerfaces*, 72(2):219–229, 2009.
- [86] M. Rabe, D. Verdes, and S. Seeger. Understanding protein adsorption phenomena at solid surfaces. *Advances in Colloid and Interface Science*, 162(1-2):87–106, 2011.
- [87] I. Langmuir. The adsorption of gases on plane surfaces of glass, mica and platinum. *Journal of the American Chemical Society*, 40(9):1361–1403, 1918.
- [88] P. Atkins. *Atkins Physical Chemistry, 9th edition*. WH Freeman, 2010, 2010.

- [89] G. K. Gomma and M. H. Wahdan. Effect of temperature on the acidic dissolution of copper in the presence of amino acids. *Materials Chemistry and Physics*, 39(2):142–148, 1994.
- [90] R. Dringen, Y. Koehler, L. Derr, G. Tomba, M. M. Schmidt, L. Treccani, L. Colombi Ciacchi, and K. Rezwani. Adsorption and reduction of glutathione disulfide on α -Al₂O₃ Nanoparticles: Experiments and modeling. *Langmuir*, 27(15):9449–9457, 2011.
- [91] A. Serro, K. Degiampietro, R. Colao, and B. Saramago. Adsorption of albumin and sodium hyaluronate on UHMWPE: a QCM-D and AFM study. *Colloids and Surfaces B: Biointerfaces*, 78(1):1–7, 2010.
- [92] S. Brunauer, P. H. Emmett, and E. Teller. Adsorption of gases in multimolecular layers. *Journal of the American Chemical Society*, 60(2):309–319, 1938.
- [93] Wikipedia. Sorptionsisotherme. <https://de.wikipedia.org/wiki/Sorptionsisotherme>. Accessed 05.09.2016.
- [94] Q. Li, M. Michaelis, G. Wei, and L. Colombi Ciacchi. A novel aptasensor based on single-molecule force spectroscopy for highly sensitive detection of mercury ions. *Analyst*, 140(15):5243–5250, 2015.
- [95] F. Felsovalyi, P. Mangiagalli, C. Bureau, S. K. Kumar, and S. Banta. Reversibility of the adsorption of lysozyme on silica. *Langmuir*, 27(19):11873–11882, 2011.
- [96] Y. Wei, A. A. Thyparambil, and R. A. Latour. Quantification of the influence of protein-protein interactions on adsorbed protein structure and bioactivity. *Colloids and Surfaces B: Biointerfaces*, 110:363–371, 2013.
- [97] A. A. Vertegel, R. W. Siegel, and J. S. Dordick. Silica nanoparticle size influences the structure and enzymatic activity of adsorbed lysozyme. *Langmuir*, 20(16):6800–6807, 2004.
- [98] P. B. Welzel. Investigation of adsorption-induced structural changes of proteins at solid/liquid interfaces by differential scanning calorimetry. *Thermochimica Acta*, 382(1-2):175–188, 2002.

- [99] N. Wurzler, *Untersuchungen der Struktur und des Adsorptionsverhaltens von Biomolekülen auf Nanopartikeln mittels Zirkulardichroismus-Spektroskopie*, Master thesis, University of Bremen, 2015.
- [100] L. Whitmore, B. Woollett, A. J. Miles, D. P. Klose, R. W. Janes, and B. A. Wallace. PCDDDB: The protein circular dichroism data bank, a repository for circular dichroism spectral and metadata. *Nucleic Acids Research*, 39:D480–D486, 2010.
- [101] J. G. Lees, A. J. Miles, F. Wien, and B. A. Wallace. A reference database for circular dichroism spectroscopy covering fold and secondary structure space. *Bioinformatics*, 22(16):1955–1962, 2006.
- [102] A. Micsonai, F. Wien, L. Kernya, Y.-H. Lee, Y. Goto, M. Réfrégiers, and J. Kardos. Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences*, 112(24):E3095–E3103, 2015.
- [103] L. Treccani, M. Maiwald, V. Zöllmer, M. Busse, G. Grathwohl, and K. Rezwani. Antibacterial and abrasion-resistant alumina micropatterns. *Advanced Engineering Materials*, 11(7):B61–B66, 2009.
- [104] S. Kroll, C. Brandes, J. Wehling, L. Treccani, G. Grathwohl, and K. Rezwani. Highly efficient enzyme-functionalized porous zirconia microtubes for bacteria filtration. *Environmental Science and Technology*, 46(16):8739–8747, 2012.
- [105] B. Bharti, J. Meissner, and G. H. Findenegg. Aggregation of silica nanoparticles directed by adsorption of lysozyme. *Langmuir*, 27(16):9823–9833, 2011.
- [106] K. Rezwani, L. P. Meier, and L. J. Gauckler. A prediction method for the isoelectric point of binary protein mixtures of bovine serum albumin and lysozyme adsorbed on colloidal titania and alumina particles. *Langmuir*, 21(8):3493–3497, 2005.
- [107] E. N. Vasina and P. Déjardin. Adsorption of α -chymotrypsin onto mica in laminar flow conditions. adsorption kinetic constant as a function of tris buffer concentration at pH 8.6. *Langmuir*, 20(20):8699–8706, 2004.

- [108] T. J. Su, J. R. Lu, R. K. Thomas, Z. F. Cui, and J. Penfold. The adsorption of lysozyme at the silica-water interface: a neutron reflection study. *Journal of Colloid and Interface Science*, 203(2):419–429, 1998.
- [109] J. L. Robeson and R. D. Tilton. Spontaneous reconfiguration of adsorbed lysozyme layers observed by total internal reflection fluorescence with a pH-sensitive fluorophore. *Langmuir*, 12(25):6104–6113, 1996.
- [110] S. M. Daly, T. M. Przybycien, and R. D. Tilton. Coverage-dependent orientation of lysozyme adsorbed on silica. *Langmuir*, 19(9):3848–3857, 2003.
- [111] K. Kubiak-Ossowska and P. A. Mulheran. Mechanism of hen egg white lysozyme adsorption on a charged solid surface. *Langmuir*, 26(20):15954–15965, 2010.
- [112] C. F. Wertz and M. M. Santore. Adsorption and reorientation kinetics of lysozyme on hydrophobic surfaces. *Langmuir*, 18(4):1190–1199, 2002.
- [113] T. Wei, M. A. Carignano, and I. Szleifer. Molecular dynamics simulation of lysozyme adsorption/desorption on hydrophobic surfaces. *The Journal of Physical Chemistry B*, 116(34):10189–10194, 2012.
- [114] W. Humphrey, A. Dalke, and K. Schulten. VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33–38, 1996.
- [115] W. Friedrichs, S. Köppen, and W. Langel. Titanium binding dodecapeptides and the impact of water structure. *Surface Science*, 617:42–52, 2013.
- [116] S. Monti and T. R. Walsh. Molecular dynamics simulations of the adsorption and dynamical behavior of single DNA components on TiO₂. *The Journal of Physical Chemistry C*, 115(49):24238–24246, 2011.
- [117] R. A. Latour. Thermodynamic perspectives on the molecular mechanisms providing protein adsorption resistance that include protein-surface interactions. *Journal of Biomedical Materials Research Part A*, 78A(4):843–854, 2006.
- [118] O. Cohavi, S. Corni, F. De Rienzo, R. Di Felice, K. E. Gottschalk, M. Hoeffling, D. Kokh, E. Molinari, G. Schreiber, A. Vaskevich, and R. C. Wade. Protein-surface interactions: challenging experiments and computations. *Journal of Molecular Recognition*, 23:259–262, 2010.

- [119] M. J. Penna, M. Mijajlovic, and M. J. Biggs. Molecular-level understanding of protein adsorption at the interface between water and a strongly interacting uncharged solid surface. *Journal of the American Chemical Society*, 136(14):5323–5331, 2014.
- [120] K. Kubiak-Ossowska and P. A. Mulheran. Multiprotein interactions during surface adsorption: a molecular dynamics study of lysozyme aggregation at a charged solid surface. *The Journal of Physical Chemistry B*, 115(28):8891–8900, 2011.
- [121] D. J. Cole, M. C. Payne, and L. Colombi Ciacchi. Water structuring and collagen adsorption at hydrophilic and hydrophobic silicon surfaces. *Physical Chemistry Chemical Physics*, 11(48):11395–11399, 2009.
- [122] S. Köppen, B. Ohler, and W. Langel. Adsorption of collagen fragments on titanium oxide surfaces: a molecular dynamics study. *Zeitschrift für Physikalische Chemie*, 221(1):3–20, 2007.
- [123] R. H. Meißner, J. Schneider, P. Schiffels, and L. Colombi Ciacchi. Computational prediction of circular dichroism spectra and quantification of helicity loss upon peptide adsorption on silica. *Langmuir*, 30(12):3487–3494, 2014.
- [124] L. Gailite, P. E. Scopelliti, V. K. Sharma, M. Indrieri, A. Podestà, G. Tedeschi, and P. Milani. Nanoscale roughness affects the activity of enzymes adsorbed on cluster-assembled titania films. *Langmuir*, 30(20):5973–5981, 2014.
- [125] L. Derr, S. Steckbeck, R. Dringen, L. Colombi Ciacchi, L. Treccani, and K. Rezwan. Assessment of the proteolytic activity of α -chymotrypsin immobilized on colloidal particles by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Analytical Letters*, 48(3):424–441, 2014.
- [126] W. J. Wedemeyer, E. Welker, M. Narayan, and H. A. Scheraga. Disulfide bonds and protein folding. *Biochemistry*, 39(15):4207–4216, 2000.
- [127] C. N. Pace, G. R. Grimsley, J. Thomson, and B. Barnett. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *The Journal of Biological Chemistry*, 263(24):11820–5, 1988.
- [128] A. A. Dombkowski, K. Z. Sultana, and D. B. Craig. Protein disulfide engineering. *FEBS Letters*, 588(2):206–212, 2013.

- [129] R. Wetzel, L. J. Perry, W. A. Baase, and W. J. Becktel. Disulfide bonds and thermal stability in T4 lysozyme. *Proceedings of the National Academy of Sciences*, (85(2)):401–405, 1988.
- [130] L. Perry and R. Wetzel. Disulfide bond engineered into T4 lysozyme: stabilization of the protein toward thermal inactivation. *Science*, 226(4674):555–557, 1984.
- [131] C. David, S. Foley, C. Mavon, and M. Enescu. Reductive unfolding of serum albumins uncovered by raman spectroscopy. *Biopolymers*, 89(7):623–634, 2008.
- [132] M. M. Castellanos and C. M. Colina. Molecular dynamics simulations of human serum albumin and role of disulfide bonds. *The Journal of Physical Chemistry B*, 117(40):11895–11905, 2013.
- [133] L. Ning, J. Guo, N. Jin, H. Liu, and X. Yao. The role of Cys179 - Cys214 disulfide bond in the stability and folding of prion protein: Insights from molecular dynamics simulations. *Journal of Molecular Modeling*, 20(2):2106, 2014.
- [134] C. David and M. Enescu. Free energy calculations on disulfide bridges reduction in proteins by combining ab initio and molecular mechanics methods. *The Journal of Physical Chemistry B*, 114(8):3020–3027, 2010.
- [135] Z. Feng, M. C. Butler, S. L. Alam, and S. N. Loh. On the nature of conformational openings: native and unfolded-state hydrogen and thiol-disulfide exchange studies of ferric aquomyoglobin. *Journal of Molecular Biology*, 314(1):153–166, 2001.
- [136] A. Cooper, S. J. Eyles, S. E. Radford, and C. M. Dobson. Thermodynamic consequences of the removal of a disulphide bridge from hen lysozyme. *Journal of Molecular Biology*, 225(4):939–943, 1992.
- [137] M. Rief. Reversible unfolding of individual titin immunoglobulin domains by AFM. *Science*, 276(5315):1109–1112, 1997.
- [138] L. Tskhovrebova, J. Trinick, J. A. Sleep, and R. M. Simmons. Elasticity and unfolding of single molecules of the giant muscle protein titin. *Nature*, 387(6630):308–312, 1997.

- [139] P. Hinterdorfer and Y. F. Dufrêne. Detection and localization of single molecular recognition events using atomic force microscopy. *Nature Methods*, 3(5):347–355, 2006.
- [140] F. Oesterhelt, M. Rief, and H. E. Gaub. Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. *New Journal of Physics*, 1(1):6, 1999.
- [141] G. Wei, S. Steckbeck, S. Köppen, and L. Colombi Ciacchi. Label-free biosensing with single-molecule force spectroscopy. *Chemical Communications*, 49(31):3239, 2013.
- [142] M. Grandbois. How strong is a covalent bond? *Science*, 283(5408):1727–1730, 1999.
- [143] A. P. Wiita, R. Perez-Jimenez, K. A. Walther, F. Gräter, B. J. Berne, A. Holmgren, J. M. Sanchez-Ruiz, and J. M. Fernandez. Probing the chemistry of thioredoxin catalysis with force. *Nature*, 450(7166):124–127, 2007.
- [144] Y. Razvag, V. Gutkin, and M. Reches. Probing the interaction of individual amino acids with inorganic surfaces using atomic force spectroscopy. *Langmuir*, 29(32):10102–10109, 2013.
- [145] A. Fuhrmann and R. Ros. Single-molecule force spectroscopy: a method for quantitative analysis of ligand-receptor interactions. *Nanomedicine*, 5(4):657–666, 2010.
- [146] M. Zhu, M. Z. Lerum, and W. Chen. How to prepare reproducible, homogeneous, and hydrolytically stable aminosilane-derived layers on silica. *Langmuir*, 28(1):416–423, 2012.
- [147] E. M. Puchner, A. Alexandrovich, A. L. Kho, U. Hensen, L. V. Schäfer, B. Brandmeier, F. Gräter, H. Grubmüller, H. E. Gaub, and M. Gautel. Mechanoenzymatics of titin kinase. *Proceedings of the National Academy of Sciences*, 105(36):13385–13390, 2008.
- [148] S. Maity, M. Mazzolini, M. Arcangeletti, A. Valbuena, P. Fabris, M. Lazzarino, and V. Torre. Conformational rearrangements in the transmembrane domain of CNGA1 channels revealed by single-molecule force spectroscopy. *Nature Communications*, 6:7093, 2015.

- [149] R.-P. Baumann, M. Schranz, and N. Hampp. Bending of purple membranes in dependence on the pH analyzed by AFM and single molecule force spectroscopy. *Physical Chemistry Chemical Physics*, 12(17):4329, 2010.
- [150] K. Kubiak-Ossowska and P. A. Mulheran. Protein diffusion and long-term adsorption states at charged solid surfaces. *Langmuir*, 28(44):15577–15585, 2012.
- [151] H. Lu and K. Schulten. Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins*, 35(4):453–463, 1999.
- [152] S. Köppen and W. Langel. Simulation of adhesion forces and energies of peptides on titanium dioxide surfaces. *Langmuir*, 26(19):15248–15256, 2010.
- [153] L. Peplowski, K. Kubiak, and W. Nowak. Mechanical aspects of nitrile hydratase enzymatic activity. Steered molecular dynamics simulations of pseudonocardia thermophila. *Chemical Physics Letters*, 467(1-3):144–149, 2008.
- [154] F. Colizzi, R. Perozzo, L. Scapozza, M. Recanatini, and A. Cavalli. Single-molecule pulling simulations can discern active from inactive enzyme inhibitors. *Journal of the American Chemical Society*, 132(21):7361–7371, 2010.
- [155] H. Grubmüller, B. Heymann, and P. Tavan. Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*, 271(5251):997–999, 1996.
- [156] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophysical Journal*, 72(4):1568–1581, 1997.
- [157] B. Heymann and H. Grubmüller. AN02/DNP-hapten unbinding forces studied by molecular dynamics atomic force microscopy simulations. *Chemical Physics Letters*, 303(1-2):1–9, 1999.
- [158] M. Rief. Single molecule force spectroscopy on polysaccharides by atomic force microscopy. *Science*, 275(5304):1295–1297, 1997.
- [159] B. Heymann and H. Grubmüller. Chair-boat transitions and side groups affect the stiffness of polysaccharides. *Chemical Physics Letters*, 305(3-4):202–208, 1999.

- [160] B. Heymann and H. Grubmüller. Elastic properties of poly(ethylene-glycol) studied by molecular dynamics stretching simulations. *Chemical Physics Letters*, 307(5-6):425–432, 1999.
- [161] E. Evans and K. Ritchie. Dynamic strength of molecular adhesion bonds. *Biophysical Journal*, 72(4):1541–1555, 1997.
- [162] R. H. Meißner, G. Wei, and L. C. Ciacchi. Estimation of the free energy of adsorption of a polypeptide on amorphous SiO₂ from molecular dynamics simulations and force spectroscopy experiments. *Soft Matter*, 11(31):6254–6265, 2015.
- [163] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi. PLUMED 2: New feathers for an old bird. *Computer Physics Communications*, 185(2):604–613, 2014.
- [164] S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7):845–854, 2013.
- [165] Z. Jin, Y. Lv, H. Cao, J. Yao, J. Zhou, W. He, and L. Yin. Core-shell nanocarriers with high paclitaxel loading for passive and active targeting. *Scientific Reports*, 6(1):27559, 2016.
- [166] J. F. Marko and E. D. Siggia. Stretching DNA. *Macromolecules*, 28(26):8759–8770, 1995.
- [167] L. Wang, R. A. Friesner, and B. J. Berne. Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). *The Journal of Physical Chemistry B*, 115(30):9431–9438, 2011.
- [168] G. Bussi, F. L. Gervasio, A. Laio, and M. Parrinello. Free-energy landscape for β hairpin folding from combined parallel tempering and metadynamics. *Journal of the American Chemical Society*, 128(41):13435–13441, 2006.
- [169] G. Bussi. Hamiltonian replica exchange in GROMACS: a flexible implementation. *Molecular Physics*, 112(3-4):379–384, 2013.

- [170] A. Barducci, G. Bussi, and M. Parrinello. Well-tempered metadynamics: a smoothly converging and tunable free-energy method. *Physical Review Letters*, 100(2):020603, 2008.
- [171] A. Gil-Ley and G. Bussi. Enhanced conformational sampling using replica exchange with collective-variable tempering. *Journal of Chemical Theory and Computation*, 11(3):1077–1085, 2015.
- [172] J. Kästner. Umbrella sampling. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(6):932–942, 2011.
- [173] S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen, and P. A. Kollman. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *Journal of Computational Chemistry*, 13(8):1011–1021, 1992.
- [174] A. Grossfield. Wham: the weighted histogram analysis method, version 2.0.9, <http://membrane.urmc.rochester.edu/content/wham>, WHAM: the weighted histogram analysis method, Version 2.0.9, <http://membrane.urmc.rochester.edu/content/wham>, Accessed 03.07.2016.
- [175] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello. Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *The Journal of Physical Chemistry B*, 110(8):3533–3539, 2006.
- [176] B. M. Bulheller and J. D. Hirst. Dichrocalc - circular and linear dichroism online. *Bioinformatics*, 25(4):539–540, 2009.
- [177] K. A. Bolin, M. Pitkeathly, A. Miranker, L. J. Smith, and C. M. Dobson. Insight into a random coil conformation and an isolated helix: Structural and dynamical characterisation of the C-helix peptide from hen lysozyme. *Journal of Molecular Biology*, 261(3):443–453, 1996.
- [178] R. Montserret, E. Aubert-Foucher, M. J. McLeish, J. M. Hill, D. Ficheux, M. Jaquinod, M. van der Rest, G. Deléage, and F. Penin. Structural analysis of the heparin-binding site of the NC1 domain of collagen XIV by CD and NMR. *Biochemistry*, 38(20):6479–6488, 1999.

- [179] J. R. Long, N. Oyler, G. P. Drobny, and P. S. Stayton. Assembly of α -helical peptide coatings on hydrophobic surfaces. *Journal of the American Chemical Society*, 124(22):6297–6303, 2002.
- [180] M. Lundqvist, P. Nygren, B.-H. Jonsson, and K. Broo. Induction of structure and function in a designed peptide upon adsorption on a silica nanoparticle. *Angewandte Chemie International Edition*, 45(48):8169–8173, 2006.
- [181] P. Nygren, M. Lundqvist, K. Broo, and B.-H. Jonsson. Fundamental design principles that guide induction of helix upon formation of stable peptid nanoparticle complexes. *Nano Letters*, 8(7):1844–1852, 2008.
- [182] M. Lundqvist, I. Sethson, and B.-H. Jonsson. Protein adsorption onto silica nanoparticles: Conformational changes depend on the particles curvature and the protein stability. *Langmuir*, 20(24):10639–10647, 2004.
- [183] M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia, and M. Parrinello. PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications*, 180(10):1961–1972, 2009.
- [184] W. Kabsch and C. Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22(12):2577–2637, 1983.
- [185] A. Barducci, M. Bonomi, and M. Parrinello. Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5):826–843, 2011.
- [186] N. Rathore, M. Chopra, and J. J. de Pablo. Optimal allocation of replicas in parallel tempering simulations. *The Journal of Chemical Physics*, 122(2):024111, 2005.
- [187] R. W. Woody and N. Sreerama. Comment on “improving protein circular dichroism calculations in the far-ultraviolet through reparametrizing the amide chromophore”. *The Journal of Chemical Physics*, 111(6):2844–2845, 1999.
- [188] N. A. Besley and J. D. Hirst. Theoretical studies toward quantitative protein circular dichroism calculations. *Journal of the American Chemical Society*, 121(41):9636–9644, 1999.

- [189] Z. Li, D. Robinson, and J. D. Hirst. Vibronic structure in the far-UV electronic circular dichroism spectra of proteins. *Faraday Discussions*, 177:329–344, 2015.
- [190] C. Camacho, G. Coulouris, V. Avagyan, N. Ma, J. Papadopoulos, K. Bealer, and T. L. Madden. BLAST+: architecture and applications. *BMC Bioinformatics*, 10(1):421, 2009.
- [191] P. Thevenet, Y. Shen, J. Maupetit, F. Guyon, P. Derreumaux, and P. Tuffery. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. *Nucleic Acids Research*, 40(W1):W288–W293, 2012.
- [192] A. Camproux, R. Gautier, and P. Tuffry. A hidden markov model derived structural alphabet for proteins. *Journal of Molecular Biology*, 339(3):591–605, 2004.
- [193] D. Xu and Y. Zhang. Ab initio protein structure assembly using continuous structure fragments and optimized knowledge-based force field. *Proteins: Structure, Function, and Bioinformatics*, 80(7):1715–1735, 2012.
- [194] J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, and Y. Zhang. The I-TASSER suite: protein structure and function prediction. *Nature Methods*, 12(1):7–8, 2014.
- [195] A. C. Wallace, R. A. Laskowski, and J. M. Thornton. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Protein Engineering, Design and Selection*, 8(2):127–134, 1995.
- [196] X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren, and A. E. Mark. Peptide folding: when simulation meets experiment. *Angewandte Chemie International Edition*, 38(1-2):236–240, 1999.
- [197] X. Huang, M. Hagen, B. Kim, R. A. Friesner, R. Zhou, and B. J. Berne. Replica exchange with solute tempering: Efficiency in large scale systems. *The Journal of Physical Chemistry B*, 111(19):5405–5410, 2007.
- [198] M. Adabi, M. Naghibzadeh, M. Adabi, M. A. Zarrinfard, S. S. Esnaashari, A. M. Seifalian, R. Faridi-Majidi, H. T. Aiyelabegan, and H. Ghanbari. Biocompatibility and nanostructured materials: applications in nanomedicine. *Artificial Cells, Nanomedicine, and Biotechnology*, 45(4):833–842, 2016.

- [199] A. Aminian, B. Shirzadi, Z. Azizi, K. Maedler, E. Volkmann, N. Hildebrand, M. Maas, L. Treccani, and K. Rezwan. Enhanced cell adhesion on bioinert ceramics mediated by the osteogenic cell membrane enzyme alkaline phosphatase. *Materials Science and Engineering: C*, 69:184–194, 2016.
- [200] L. T. de Jonge, J. J. van den Beucken, S. C. Leeuwenburgh, A. A. Hamers, J. G. Wolke, and J. A. Jansen. In vitro responses to electrosprayed alkaline phosphatase/calcium phosphate composite coatings. *Acta Biomaterialia*, 5(7):2773–2782, 2009.
- [201] A. Aminian, K. Pardun, E. Volkmann, G. Li Destri, G. Marletta, L. Treccani, and K. Rezwan. Enzyme-assisted calcium phosphate biomineralization on an inert alumina surface. *Acta Biomaterialia*, 13:335–343, 2015.
- [202] J. E. Coleman. Structure and mechanism of alkaline phosphatase. *Annual Review of Biophysics and Biomolecular Structure*, 21(1):441–483, 1992.
- [203] D. Xu and Y. Zhang. Generating triangulated macromolecular surfaces by euclidean distance transform. *Public Library of Science (PLOS) ONE*, 4(12):e8140, 2009.
- [204] J. Bartels, M. Souza, A. Schaper, P. Árki, S. Kroll, and K. Rezwan. Amino-functionalized ceramic capillary membranes for controlled virus retention. *Environmental Science and Technology*, 50(4):1973–1981, 2016.
- [205] K. Toropova, G. Basnak, R. Twarock, P. G. Stockley, and N. A. Ranson. The three-dimensional structure of genomic RNA in bacteriophage MS2: implications for assembly. *Journal of Molecular Biology*, 375(3):824–836, 2008.
- [206] K. Dent, R. Thompson, A. Barker, J. Hiscox, J. Barr, P. Stockley, and N. Ranson. The asymmetric structure of an icosahedral virus bound to its receptor suggests a mechanism for genome release. *Structure*, 21(7):1225–1234, 2013.
- [207] K. Lindorff-Larsen, P. Maragakis, S. Piana, M. P. Eastwood, R. O. Dror, and D. E. Shaw. Systematic validation of protein force fields against experimental data. *Public Library of Science (PLOS) ONE*, 7(2):e32131, 2012.

- [208] E. A. Cino, W.-Y. Choy, and M. Karttunen. Comparison of secondary structure formation using 10 different force fields in microsecond molecular dynamics simulations. *Journal of Chemical Theory and Computation*, 8(8):2725–2740, 2012.
- [209] J. M. Haile. *Molecular Dynamics Simulation: elementary Methods*. Wiley-Interscience, 1997.
- [210] C. A. Angell. *Dynamic Aspects of Structural Change in Liquids and Glasses (Annals of the New York Academy of Sciences)*. New York Academy of Sciences, 1986.
- [211] L. Verlet. Computer experiments on classical fluids. I. Thermodynamical properties of lennard-jones molecules. *Physical Review*, 159(1):98–103, 1967.
- [212] A. Leach. *Molecular Modelling: principles and Applications (2nd Edition)*. Pearson, 2001.
- [213] R. W. Hockney. The potential calculation and some applications. *Methods in Computational Physics*, 9:136–210, 1970.
- [214] T. Schlick. *Molecular Modeling and Simulation: an Interdisciplinary Guide: 21 (Interdisciplinary Applied Mathematics)*. Springer New York, 2010.
- [215] W. C. Swope. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. *The Journal of Chemical Physics*, 76(1):637, 1982.
- [216] R. Meißner. *Adhesion Phenomena at the Interface between Biomolecules and Silica*. PhD thesis, University of Bremen, University of Bremen, 2015.
- [217] P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 369(3):253–287, 1921.
- [218] T. Darden, L. Perera, L. Li, and L. Pedersen. New tricks for modelers from the crystallography toolkit: the particle mesh ewald algorithm and its use in nucleic acid simulations. *Structure*, 7(3):R55–R60, 1999.
- [219] L. Woodcock. Isothermal molecular dynamics calculations for liquid salts. *Chemical Physics Letters*, 10(3):257–261, 1971.

- [220] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684, 1984.
- [221] G. Bussi, D. Donadio, and M. Parrinello. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101, 2007.
- [222] H. C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384, 1980.
- [223] S. Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2):255–268, 1984.
- [224] W. G. Hoover. Canonical dynamics: equilibrium phase-space distributions. *Physical Review A*, 31(3):1695–1697, 1985.
- [225] R. J. Loncharich, B. R. Brooks, and R. W. Pastor. Langevin dynamics of peptides: the frictional dependence of isomerization rates of N-acetylalanyl-N-methylamide. *Biopolymers*, 32(5):523–535, 1992.
- [226] B. Guillot. A reappraisal of what we have learnt during three decades of computer simulations on water. *Journal of Molecular Liquids*, 101(1-3):219–260, 2002.
- [227] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. Interaction models for water in relation to protein hydration. *Intermolecular Forces*, pages 331–342, 1981.
- [228] M. W. Mahoney and W. L. Jorgensen. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *The Journal of Chemical Physics*, 112(20):8910, 2000.
- [229] P. G. Kusalik and I. M. Svishchev. The spatial structure in liquid water. *Science*, 265(5176):1219–1221, 1994.
- [230] S. Izadi, R. Anandakrishnan, and A. V. Onufriev. Building water models: a different approach. *The Journal of Physical Chemistry Letters*, 5(21):3863–3871, 2014.
- [231] F. H. Stillinger. Improved simulation of liquid water by molecular dynamics. *The Journal of Chemical Physics*, 60(4):1545, 1974.

- [232] W. L. Jorgensen and J. Tirado-Rives. Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. *Proceedings of the National Academy of Sciences*, 102(19):6665–6670, 2005.
- [233] M. Chaplin. Water structure and science. http://www1.lsbu.ac.uk/water/water_models.html. Accessed 17.11.2016.
- [234] J. Weiser, P. S. Shenkin, and W. C. Still. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *Journal of Computational Chemistry*, 20(2):217–230, 1999.
- [235] F. Pietrucci and A. Laio. A collective variable for the efficient exploration of protein beta-sheet structures: Application to SH3 and GB1. *Journal of Chemical Theory and Computation*, 5(9):2197–2201, 2009.
- [236] C. Jarzynski. Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14):2690–2693, 1997.
- [237] A. Laio and M. Parrinello. Escaping free-energy minima. *Proceedings of the National Academy of Sciences*, 99(20):12562–12566, 2002.
- [238] Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1-2):141–151, 1999.
- [239] K. Hukushima and K. Nemoto. Exchange monte carlo method and application to spin glass simulations. *Journal of the Physical Society of Japan*, 65(6):1604–1608, 1996.
- [240] P. Liu, B. Kim, R. A. Friesner, and B. J. Berne. Replica exchange with solute tempering: a method for sampling biological systems in explicit water. *Proceedings of the National Academy of Sciences*, 102(39):13749–13754, 2005.
- [241] H. Nymeyer, S. Gnanakaran, and A. E. Garcia. Atomic simulations of protein folding, using the replica exchange algorithm. *Methods in Enzymology*, 383:119–149, 2004.
- [242] H. Fukunishi, O. Watanabe, and S. Takada. On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: application to protein structure prediction. *The Journal of Chemical Physics*, 116(20):9058, 2002.

- [243] A. Patriksson and D. van der Spoel. A temperature predictor for parallel tempering simulations. *Physical Chemistry Chemical Physics*, 10(15):2073, 2008.
- [244] S. L. C. Moors, S. Michielssens, and A. Ceulemans. Improved replica exchange method for native-state protein sampling. *Journal of Chemical Theory and Computation*, 7(1):231–237, 2011.
- [245] B. G. Levine, J. E. Stone, and A. Kohlmeyer. Fast analysis of molecular dynamics trajectories with graphics processing units - Radial distribution function histogramming. *Journal of Computational Physics*, 230(9):3556–3569, 2011.
- [246] J. Shao, S. W. Tanner, N. Thompson, and T. E. Cheatham. Clustering molecular dynamics trajectories: 1. characterizing the performance of different clustering algorithms. *Journal of Chemical Theory and Computation*, 3(6):2312–2334, 2007.

Erklärung

Hiermit erkläre ich, dass ich

- die Arbeit ohne unerlaubte fremde Hilfe angefertigt habe;
- keine anderen als die von mir angefertigten Quellen und Hilfsmittel benutzt habe;
- die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Bremen, den 23.06.2017

Nils Alexander Hildebrand