

*Adhesion Phenomena at the Interface
between Biomolecules and Silica*

Vom Fachbereich Produktionstechnik

der

UNIVERSITÄT BREMEN

zur Erlangung des Grades

Doktor-Ingenieur

genehmigte

Dissertation

von

Dipl. Phys. Robert Meißner

Gutachter: Professor Dr.-Ing. Lucio Colombi Ciacchi
Professor Robert Latour (Clemson University, USA)

Tag der mündlichen Prüfung: 16.10.2015

Adhesion Phenomena at the Interface between Biomolecules and Silica

ABSTRACT

Estimating the free energy associated with the adsorption of materials-binding peptides or conformational changes in biomolecules is fundamental to quantify their interactions across bio/inorganic interfaces, but is difficult to achieve both experimentally and theoretically. Computer simulations, on the other hand, reveal fundamental insights into processes occurring on the nanoscale with varying level of details and accuracy— needless to say, that this is often possible at a fraction of the expense of an experimental setup with the same resolution. The highest level of detail and accuracy is achieved with *ab initio* quantum chemical calculations for which the Schrödinger Equation is solely solved with natural constants. However, approximations must be made and thus the accuracy strongly depends on the amount of quantum effects considered in the theoretical formulations. In general, the immense computational effort for a description of systems at relatively small length and time scales is the most severe drawback of these methods. On the contrary, classical molecular dynamics simulations facilitate simulations several orders of magnitude higher in time and length scales by employing effective interaction potentials between atoms to describe e.g. van der Waals, Coulomb, or covalent bond interactions. A disadvantage of this approach is that the accuracy of classical force fields crucially depends on the reliability of the underlying parameterization; applications to problems beyond the context of the original parameterization, need to be handled with caution and have to be reviewed thoroughly. Thus, it is necessary to validate results from classical molecular dynamics with experiments by comparing physical or chemical quantities describing distinct properties. The extreme difference in time and length scales between experimental setups in the laboratory and the virtual reality of a computer simulation renders this task even more challenging.

The first part of this work aims on the assessment of a suitable set of force fields to describe biomolecular systems. It is shown that reactive force fields, although, in principle very promising, are not yet sufficiently developed for a realistic description of the interface between dissolved biomolecules and oxide surfaces. Based on these findings, a classical force field, describing a deprotonated silica surface, is developed and examined

with respect to its applicability to describe the interface between water and silica. In a further step, the newly developed force field is applied to simulate interfacial systems, which additionally include dissolved biomolecules.

Challenging in biomolecular simulations is that larger peptides and proteins exhibit a vast number of conformational degrees of freedom, thus impeding a straight-forward prediction of macroscopic properties based on microscopic simulations. If such proteins or peptides are to be considered, regular molecular dynamics simulations are often limited to only one molecule, if at all, due to the restricted system size and accessible time scale. Free energy changes associated with selected reaction coordinates—generalized, one-dimensional system coordinates, representing the progress along a reaction pathway—are widely used in both laboratory experiments and simulations to characterize, e.g., adsorption processes and changes in the structure of biomolecules. A suitably chosen free energy reaction coordinate offers a robust opportunity to verify the applicability of force fields against experiments. However, due to their macroscopic nature a free energy is not easily accessible from computer simulations. A robust, simulation-based prediction of a macroscopic quantity, such as the free energy, requires that all relevant molecular conformations, even those that are separated by large energy barriers, are included in the calculation with the correct statistical weights. A key result of this work is the development of a method to predict circular dichroism (CD) spectra, utilizing a free energy-based reweighting scheme of the molecule’s conformational space in the calculation of the spectra. This process is, amongst others, correlated with the strength of the surface charge and the presence of charged amino acids in the peptide chain. Apart from that, it was found that amino acids which possess a long flexible charged side chain (such as arginine) contribute significantly to the adhesion when they attach to a deprotonated silica surface. A combination of molecular dynamics (MD) simulations and dynamical force-spectroscopy experiments based on atomic force microscopy (AFM) is employed in order to estimate the free energy of adsorption ΔF_{ads} of a tetrapeptide on amorphous SiO_2 in pure water. The results of both equilibrium, advanced sampling MD and non-equilibrium, steered MD are compared with those of two different approaches used to extract ΔF_{ads} from the dependence of experimentally measured adhesion forces on the applied AFM loading rates. In order to obtain unambiguous peak forces and bond loading rates from steered MD trajectories, a novel numerical protocol is developed based on a piecewise-harmonic fit of the adhesion work profile along each trajectory. The interpretation of the experiments has required a thorough quantitative characterization of the elastic properties of polyethylene glycol linker molecules used to tether a repeated GCRL tetrapeptide to AFM cantilevers, and of the polypeptide itself.

Contents

1	WHY SIMULATING THE REAL WORLD?	1
1.1	Biomolecular Adsorption on Technical Glasses	2
1.2	A Short Story about Technical Glasses	8
2	REPRESENTING REALITY WITH MD SIMULATIONS: AN INTRODUCTION	17
2.1	Equations of Motion	19
2.2	Integrating Newton's Equations of Motion	22
2.3	Periodic Boundary Conditions	25
2.4	Constrained Equations of Motion	26
2.5	Empirical Force Fields	30
2.6	Thermostats and Barostats	36
2.7	Comprehensive Force Fields for Silica, Water and Biomolecules	41
2.8	Free Energy Calculations	64
3	SIMULATING THE PHYSICAL WORLD: PROTEINS AT THE SILICA INTERFACE	87
3.1	Conformational Changes of a Peptides Adsorbing to Silica	90
3.2	Adsorption Free Energy Estimates of a Polypeptide on Amorphous SiO_2 using Simulations and Force Spectroscopy Experiments	114
4	CONCLUSIONS AND PERSPECTIVES	143
4.1	A Glance into the Crystal Ball	151
APPENDIX A EXPLORING SECONDARY STRUCTURES OF PEPTIDES WITH FAR-UV CD SPECTROSCOPY		155
A.1	Absorption of Circularly Polarized Light of Chiral Molecules	156
A.2	Electronic Excitation States of an α -helix in Proteins	159
A.3	Solving the Rosenfeld Equation with the Matrix Method	163

APPENDIX B	PHYSICAL AND THEORETICAL CONCEPTS IN MOLECULAR DYNAMICS SIMULATIONS	169
B.1	Force Field Definitions	169
B.2	Derivation of Coulomb Interactions	179
B.3	Derivation of van der Waals Interactions	181
B.4	Treatment of Long-Range Interactions	184
REFERENCES		221

DIESE ARBEIT IST MEINER FAMILIE UND GANZ BESONDERS MEINEM VERSTORBENEN VATER, NORBERT MEIßNER, GEWIDMET - ES ZERREIßT MIR MEIN HERZ, DASS DU DAS NICHT MEHR MITERLEBEN KANNST, OBWOHL DU ES DIR DOCH SO SEHR GEWÜNSCHT HAST.

Acknowledgments

I THANK ALL who supported me in getting this thesis finished. Especially my colleagues at the Fraunhofer IFAM institute, most notably Peter Schiffels, Michael Hoffmann, Mark Amkreutz and Welchy Leite Cavalcanti. Furthermore all the Brazilians I met during my Ph.D. time, especially those with whom I shared an office and shared so much fun - thank you Felipe Favaretto Trombetta, Bruno Guilherme Carvalho, Leandro Goncalves de Andrade Rosa, Alexander Iwami Petzhold, Fernando Machuca Neto, Jonas Aniol, Andreas Volkmann and Hauke Brüning!

I thank Julian Schneider, who was a former Ph.D. student in the Hybrid Materials Interface group of Lucio Colombi Ciacchi, whom I am particularly grateful. Without him many parts of this work wouldn't have been possible.

Of course, I want to thank most especially my family for their steady support, even in rough times. At this point I'd like to thank my many friends, for all their inspiration, their endurance and of course for their welcome distractions. Erik, Janna and Hauke, thanks for putting so much effort into proof-reading this! One very important person in my life deserves a special reference here, but prefers to stay anonymous.

Preface

This dissertation describes the results obtained during my time as a Ph.D. student from the first of August 2010 to Mai 2015 at the University of Bremen. Parts of this thesis have been published as follows:

CHAPTER 2.7.2: Butenuth, A., Moras, G., Schneider, J., Koleini, M., Köppen, S., Meißner, R., Wright, L. B., Walsh, T. R., & Ciacchi, L. C. (2012). Ab initio derived force-field parameters for molecular dynamics simulations of deprotonated amorphous-SiO₂/water interfaces. *Physica Status Solidi B: Basic Solid State Physics*, 249(2), 292–305

CHAPTER 3.1: Hartwig, A., Meissner, R., Merten, C., Schiffels, P., Wand, P., & Grunwald, I. (2013). Mutual influence between adhesion and molecular conformation: Molecular geometry is a key issue in interphase formation. *Journal of Adhesion*, 89(2), 77–95

AND Meißner, R. H., Schneider, J., Schiffels, P., & Colombi Ciacchi, L. (2014). Computational prediction of circular dichroism spectra and quantification of helicity loss upon peptide adsorption on silica. *Langmuir*, 30(12), 3487–3494 ("Reprinted with permission from ref. 146. Copyright 2014, American Chemical Society.", <https://dx.doi.org/10.1021/la500285m>)

CHAPTER 3.2: Meißner, R. H., Wei, G., & Colombi Ciacchi, L. (2015). Estimation of the Free Energy of Adsorption of a Polypeptide on Amorphous SiO₂ from Molecular Dynamics Simulations and Force Spectroscopy Experiments. (Submitted) - Reproduced by permission of The Royal Society of Chemistry.

Two publication discussing the results of Chapter 3.2 and the ongoing circular dichroism experiments as well as simulations of a zinc binding peptide are currently in preparation.

Zusammenfassung

Heutzutage ist es experimentell möglich chemische Prozesse auf atomarer Ebene bildlich darzustellen, wenn auch nur unter Einsatz von immensen Ressourcen und Aufwand. Computersimulationen dagegen erlauben einen fundamentalen Einblick in Prozesse im Nanometerbereich mit einer Genauigkeit abhängig vom nötigen Detaillierungsgrad und den verfügbaren Rechenressourcen – häufig mit einem geringeren Aufwand als bei einem Experiment mit gleicher Auflösung im Labor. Die höchstmögliche Genauigkeit ergibt sich dabei aus Rechnungen die auf quantenmechanischen *ab-initio* Gleichungen, für welche die Schrödingergleichung nur durch Naturkonstanten gelöst wird, beruhen. Nichtsdestotrotz müssen häufig Annahmen und Näherungen gemacht werden um die Komplexität des Problems zu verringern; Die Genauigkeit der Vorhersage von Computersimulationen hängt daher davon ab, wie viele quantenmechanische Effekte in dem zugrunde liegendem Modell berücksichtigt werden. Im Allgemeinen sind die Anforderungen an die Rechenressourcen quantenmechanischer Rechnungen für relativ kleine Längen- und Zeitskalen immens. Im Kontrast dazu erlauben klassische Molekulardynamiksimulationen um mehrere Größenordnungen längere und größere Systeme. Die Beschreibung des Systems beruht dabei auf klassischen Effekten, wie z. B. van der Waals, Coulomb oder harmonische Wechselwirkungen, um nur ein paar zu nennen. Der Nachteil ist hierbei, dass die Genauigkeit der klassischen Kraftfelder von der darunterliegenden Parametrisierung abhängt; das hat zur Konsequenz, dass eine Anwendung auf ein Problem für welches Kraftfeld nicht gedacht ist immer mit Vorsicht zu behandeln ist und eine genaue Überprüfung nötig ist. Es ist daher notwendig die Ergebnisse aus klassischen Molekulardynamiksimulationen mit physikalischen oder chemischen Eigenschaften des realen Systems zu vergleichen. Die unterschiedlichen Zeit- und Längenskalen des Experiments im Labor und am Computer machen den Vergleich allerdings zu einer herausfordernden Aufgabe. Erst wenn die systembestimmenden thermodynamischen Eigenschaften in den Simulationen korrekt wiedergegeben werden, ist eine genauere Betrachtung der Struktur des Systems oder der im System ablaufenden Prozesse auf atomarer Ebene möglich.

Ein Teil dieser Arbeit beschäftigt sich deshalb mit der Bewertung von Kraftfeldern hinsichtlich ihrer Anwendbarkeit auf biomolekulare Systeme an einer Grenzfläche zu einer

Silica-Oberfläche. In dieser Arbeit wird zunächst gezeigt, dass reaktive Kraftfelder, obwohl sie sehr vielversprechend sind, momentan nicht in der Lage sind ausreichend realistisch die Grenzschicht zwischen einem solvatisiertem Biomolekül und einer Silica-Oberfläche zu beschreiben. In einem zweiten Schritt wird deshalb ein klassisches Kraftfeld entwickelt mit der eine deprotonierte Silica-Oberfläche beschrieben werden kann. Anschließend wird überprüft ob das neu entwickelte Kraftfeld die Grenzfläche zwischen Wasser und Silica korrekt beschreiben kann.

Das neu entwickelte Kraftfeld wird in einem nächsten Schritt genutzt um Systeme, die gelöste Biomoleküle und eine Silica-Oberfläche beinhalten, zu beschreiben. Die Berücksichtigung von Wechselwirkungen zwischen Biomolekülen wurde bewusst vernachlässigt, um die Problemstellung nicht noch weiter zu verkomplizieren; In einigen Fällen spielt sie durchaus eine elementare Rolle. Diese Arbeit soll sich aber in erster Linie mit der Betrachtung der Wechselwirkung zwischen Oberfläche und Molekül beschäftigen. Proteine und Peptide besitzen eine große Auswahl an möglichen Konformationen. Die Vorhersage einer makroskopischen Größe aus der Simulation hängt aber von der Berücksichtigung aller Konformationen ab, unabhängig davon ob sie energetisch sinnvoll sind oder nicht – Das macht es unter Umständen extrem schwierig, makroskopische Größen, wie z. B. freie Energien, korrekt aus der Simulation vorherzusagen. Die Bestimmung von freien Energien ist weit verbreitet in Laborexperimenten als auch in Computersimulation, da sie eine robuste Möglichkeit bietet die Simulationen zu verifizieren. Mit Hilfe der freien Energie können z. B. die Fähigkeiten eines Moleküls beschrieben werden, selbständig an einer Oberfläche zu adsorbieren oder eine bestimmte Konformation einzunehmen. Die freie Energie der Adsorption gibt dabei die bei der Adsorption eines Moleküls frei werdende (oder zusätzlich benötigte) Energie an. Häufig werden dafür abstrakte eindimensionale Koordinaten eingeführt, um den Prozess entlang eines Reaktionspfades zu beschreiben. Freie Energien können entweder aus der Simulation gewonnen werden, indem man ausreichend lange simuliert, d. h. über Zeitskalen von Sekunden bis Stunden in Simulationen, oder aber durch Verwendung ausgeklügelter Methoden, um die Einschränkungen einer normalen Molekulardynamiksimulation zu überwinden. Ersteres ist normalerweise durch die natürliche Beschränkung verfügbarer Ressourcen in Computersimulationen nicht möglich. In dieser Arbeit wird deshalb auf fortgeschrittene Methoden, wie paralleles Temperieren und METADYNAMICS, zurückgegriffen.

Ein wesentlicher Punkt dieser Arbeit ist die Vorhersage freier Energien der Adsorption aus der Simulation. Simulationen eines Tetrapeptides werden dafür unter Nicht-Gleichgewichts- sowie Gleichgewichtsbedingungen durchgeführt und mit Ergebnissen der Rasterkraftmikroskopie verglichen. Eine weitere wichtige Errungenschaft dieser Arbeit liegt in der Entwicklung einer Methode zur genauen Bestimmung eines theoretischen Circular dichroismus Spektrums. In dieser Berechnung werden unterschiedliche Konformatio-

nen eines Peptides, die zugehörige freie Energie und ihr entsprechendes Circular dichroismus Spektrum explizit in einem Algorithmus zur Neugewichtung der freien Energie berücksichtigt, um eine möglichst genaue Vorhersage zu treffen.

It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.

Richard P. Feynman

1

Why Simulating the Real World?

A PERSONAL NOTE

When Einstein or other great masterminds thought about their ideas and finally brought them down to paper, they must certainly have felt some kind of relief and happiness – Independent on the resulting impact or even non-impact of my research, I think the same will be true when I have finished my doctorate. Questions about where phenomena like

emotions, the mind or soul and human behavior originates from fascinate me since I know about my own conscious mind.

I'm curious, what are the fundamental reasons of emotions and feelings? Are they some kind of metaphysical or spiritual phenomenons, which are truly intangible, or does everything humanity is made of simply rest upon on chemical and physical processes happening in our brain, body and bloodstream. Understanding such processes on a fundamental level is a way for me to understand the influences, which steer the mechanism behind emotions like faith, anger, joy, love or hate. A small hormone named oxytocin for example consists only of 9 amino acids, or roughly 90 atoms, but has a remarkably attribute: It evidently amplifies the faith between humans and plays an important role in stress regulation¹¹⁸. Knowledge of the occurring processes on a fundamental level can help to allay the pain of severe illnesses such as depressions or similar.

However, there will and always will be some kind of “magic” left over—the purpose, desire and driving force of science—which scientists are not yet able to describe with their rational minds, philosophers will argue about, and believers may fight for.

1.1 BIOMOLECULAR ADSORPTION ON TECHNICAL GLASSES

A thesis about a scientific research topic is always accompanied with ups and downs, frustrations and enthusiasm – I knew that before. In the beginning and throughout the time I increasingly often encountered the pressing question: Why all this? Thus, this chapter is intended to explain my driving forces to examine phenomena at the interface of protein-containing aqueous solutions and the amorphous SiO₂ phase. Moreover, at the end of this chapter a short introduction of the manufacturing process, some applications, and struc-



Figure 1.1: Example of a typical pharmaceutical glass vial for the packaging of protein containing aqueous solutions. The image is under the creative commons licence CC0 and was obtained from the clker.com database (<http://www.clker.com/clipart-160615.html>).

tural properties of technical glasses are given to show the versatility of glasses; making them intriguingly interesting.

Protein-based drugs are often stored in glass vials, particularly when they are dissolved in an aqueous solution. A typical pharmaceutical packaging for protein based drugs is shown in Figure 1.1. Widely used is, for example, the angiogenesis inhibitor bevacizumab that slows down the growth of new blood vessels – a drug with the most comprehensive development program in oncology. Bevacizumab is a humanized monoclonal antibody of the group of immunoglobulines (IgG1) and is prescribed to treat, amongst others, breast, colorectal, or lung cancer. Apart from the packaging of drugs, the issue of protein adsorption onto solid surfaces in a wet environment is of great importance for the design and functionalisation of materials in the context of biomedical implants, biosensors, antifouling surfaces, and microelectromechanical systems^{89,61,193,229}.

Although glasses are known to possess highly inert surfaces, adhesion of the dissolved

active principles of the protein-containing water solution on the container's wall is known to occur. The loss of efficiency and the dangerous side-effects which this may cause, are not at all desirable. Thus, the requirements for primary packagings of new pharmaceutical biological drugs in the pipeline are high and ever increasing. This is a particularly serious problem in the case of low concentrated biomolecule formulations, when a significant amount of dissolved active principles adsorbs at the pharmaceutical packaging surface. In the last twenty years, efforts have been made to reduce the protein adsorption through modification of the drug formulation including additional ingredients, such as sacrificial proteins (typically human serum albumin), salts (NaCl) or surfactants (e.g. tensides like Tween 20)^{104,127,97}. However, this is an expensive and time-consuming route, since any novel reformulation of a drug requires a novel drug certification. Additional ingredients carry the risk of unwanted side-effects. Thus, there is a high market need for bio-compatible pharmaceutical packaging surfaces, which can drastically reduce the adsorption of the active biomolecules on the packaging wall. This becomes particular evident, for example, in the case of the recombinant human interleukin-12 (rhIL-12) protein. One possible usage of this protein is the inhibition of tumor angiogenesis¹. The amount of recovered rhIL-12 after 16 hours incubation time at pH 5.6 and room temperature is illustrated for two different glasses types in Figure 1.2. The amount of recovered rhIL-12 was determined by using Reverse Phase High Performance Liquid Chromatography (RP-HPLC). Most critical is the high loss of active ingredients at low concentrations, since novel biological drugs to treat, e.g., cancer are often stored in glass vials at very low concentrations. Surprisingly, the amount of adsorbed rhIL-12 could be significantly reduced by coating the inner glass walls with a pure SiO₂ phase using a plasma treatment. The recovered rhIL-12 protein concen-

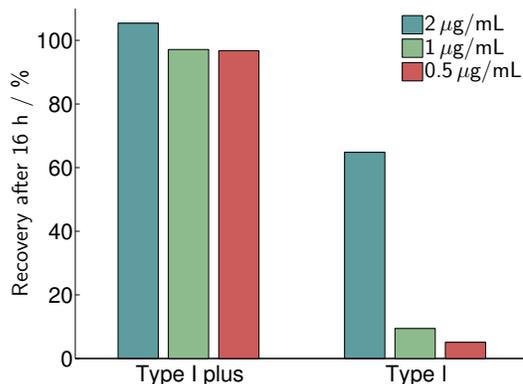


Figure 1.2: Reduction of protein loss through a silicon-dioxide coating developed by Schott AG (“Type I plus”) in comparison with an uncoated standard vial (“Type I”). Used concentrations are given in the legend. Source: Schott AG (Patents: EP821079B1, EP811367B1)

tration was stable even over 36 months at various temperatures. A comparison of the long-term protein concentration stability between “Type I plus” and standard vials is shown in Figure 1.3 for another type of biomolecule, i.e. nicotinic acetylcholine receptors (nAChR). However, this behavior may not be true for other protein-based drugs and furthermore associated phenomena of bimolecular adsorption are still not fully understood. On that account it is necessary to understand underlying physical and chemical processes emerging at the atomistic scale upon adsorption of a biomolecule onto a solid surface from the aqueous phase. Atomic-level investigations of glassy interfaces thus have a high potential impact on society, additional to their significant value measured in purely scientific terms. However, the atomistic details of the interactions mediated by these effects are still unclear and are the object of intense experimental and theoretical investigations^{107,175,176}. As far as atomistic modeling is concerned, most of the present studies either deal with the interactions between single amino acids and idealized metal (or oxide) surface models at the level of Density Functional Theory²⁰⁶, or try to describe the interaction between larger protein

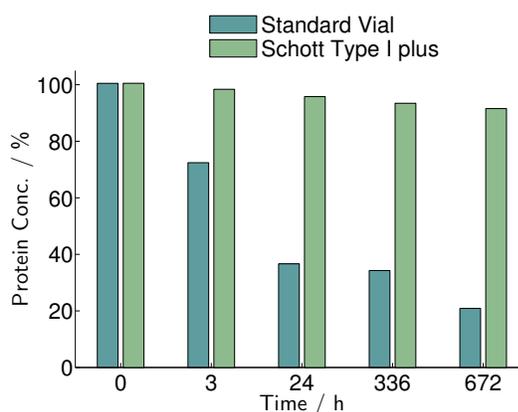


Figure 1.3: nAChR concentration in solution after several hours of incubation in two different glass vials. Source: Dr. André Schrattenholz, Johannes Gutenberg-Universität Mainz, Medical School, Institute of Physiological Chemistry and Pathobiochemistry

fragments and chemically inert surfaces at the level of classical force field potentials^{176,53,180}. Recently, the adhesion of predominantly hydrophobic spider-silk protein domains on a hydrophobic diamond surface has been studied by a combination of atomistic MD modeling with single-molecule AFM spectroscopy experiments¹⁰⁷. This study revealed that MD simulations are capable of predicting the maximum adhesion force between an adsorbed protein and a solid surface with very high precision (within 5 pN of the experimental values), as long as no covalent interactions govern the adhesion behavior. However, the adhesion of peptides to surfaces may involve interactions which go beyond the “weak forces” easily included in force field potentials, and rely on chemical effects stronger than purely ionic or hydrogen bonds¹²⁹.

The interaction of a protein with a solid surface is a complex process. Diffusion or convection driven processes bring molecules in proximity to the surface on which they can adsorb, diffuse, (partially) unfold, or desorb. Driving factors for the adsorption are electrostatic, dispersive¹¹ or hydrophobic¹³⁹ interactions, as well as energetically favorable changes

in the tertiary protein structure^{159,156}. Other factors, such as intrinsic properties of the protein, the physics/chemistry of solid surface, and the composition of the liquid (“formulation”) can have a great impact on the adsorption behavior, as well. Numerous different techniques are used to examine the adsorption process and to evaluate the amount or structural changes of the adsorbed proteins^{117,153}. These include (i) high performance liquid chromatography (HPLC) used for the separation and quantification of components in a mixture; (ii) X-ray photoelectron spectroscopy (XPS) which is used for analyzing the orientation, surface coverage or layer thickness of an adsorbent; (iii) Fourier transform or attenuated total reflection infrared spectroscopy (FTIR or ATR-IR) which find applications in many fields, e.g., in analyzing protein conformations or the quantification of molecular species; (iv) time-of-flight secondary mass spectrometry (ToF-SIMS) which can be used to analyze the adsorption, conformations, and orientations of biomolecules on substrates or (v) atomic force microscopy (AFM) from which adhesion forces, lateral distributions and layer thicknesses with molecular resolution under native conditions, i.e., in buffer or formulation solution, can be obtained. (vi) Ellipsometry and especially (vii) Quartz Crystal Microbalance (QCM) are tools for in-situ experiments to follow the kinetics of the adsorption in real-time. QCM also allows to determine if the process is reversible or not, and the influence of changes in the formulation composition on the adsorption¹³⁵. A major drawback of all experimental techniques listed above is that no single-atom level information can be resolved, and therefore the knowledge about the detailed protein/surface interactions which eventually drive adhesion is severely limited.

In this work, as a considerable progress with respect to the state of the art, experimental results of peptide molecules, synthesized ad hoc, is combined with advanced atomistic mod-

eling of the same peptides. The scope of this work is thus to gain a detailed, atomistic-scale knowledge of chemical and physical processes, taking place at glassy interfacial systems. In particular chemical and physical interactions of peptides at these interfaces are analyzed and possible structural changes, local adhesive interactions and other associated phenomena, occurring upon adsorption, are quantified via molecular dynamics simulations. The physical and chemical informations gained thereof are brought into close relation with results from experimental spectroscopic techniques, i.e. circular dichroism and atomic force microscopy.

1.2 A SHORT STORY ABOUT TECHNICAL GLASSES

Products made out of glass play a significant role not only in our daily live; in its natural occurrence as quartz and its several varieties, SiO_2 is also one of the main constituents of the earth's crust. Manufactured industrially it is used for example in optical glasses as lenses, prisms and mirrors which are essential parts in microscopes, spyglasses or telescopes. Apart from this they can also be found in glass-ceramics of hobs, tumblers, light bulbs, optical fibers, glass fiber-reinforced plastics and as glass panels for cars, buildings and protective layers for solar modules—but, above all, a special focus in this work is on the potential usage of glasses in the pharmaceutical packaging of drugs. In order to illustrate the importance of realistic silica surfaces in simulations, an introduction into topological and chemical heterogeneities occurring on silica surfaces will be given.

Glasses can be classified according to their genesis or chemical composition. The former is divided into in artificial and natural glasses, whereby natural glasses emerge from the melting of sands. Its regular crystal structure can be disturbed by the subsequent effect of a shock wave transforming them into an amorphous crystal. Artificial glasses are mostly

produced by melting raw materials in different smelters or from a so-called sol-gel process, where successively thin layers are deposited on the substrate. A major group in the classification according to their chemical composition is, for example, the siliceous glasses with its most prominent representative the soda lime glass. Some siliceous glasses are doped by other oxides, such as aluminum oxide or sodium carbonate, but all have the same underlying network of amorphous silicon dioxide in common. Without additional constituents in its pure SiO_2 form, siliceous glasses are called vitreous, quartz or silica glass. Silica glass in its pure form is of great importance for this work, since it is often used in pharmaceutical packaging due to its very high chemical inertness. This manifests in its resistance to nearly all elements and compounds including virtually all acids, independent of concentration, but with one exception: hydrofluoric acid which is very reactive even in fairly low concentrations. Broadly speaking, the relatively well-defined structures of pure SiO_2 phases make them a good candidate for computer simulations – in other words, it is a system which allows fundamental insights without the requirement for including effects arising from the introduction of impurities or the like. Impurities could be important for some problems, however, they will increase the complexity of the system which is not absolutely necessary for problems considered in this work.

1.2.1 CRYSTAL STRUCTURE OF PURE SiO_2

Silica occurs naturally in several different forms—depending on temperature and pressure during their formation. Common are for instance quartz, tridymite, cristobalite, coesite, keatite, stishovite, melanophlogite, fibrous, lamellar or amorphous silicas. The temperature and pressure dependence of the different crystallization states of pure SiO_2 can be found

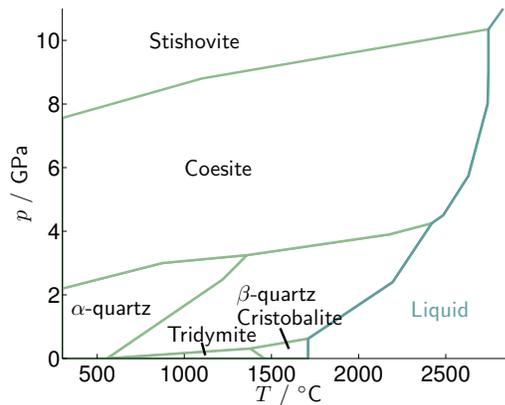


Figure 1.4: Temperature and pressure phase diagram for SiO_2 . The illustration was generated from data taken from Holleman et al.¹⁰³, Rykart¹⁸⁸ and Wenk & Bulakh²³⁶.

in the phase diagram shown in Figure 1.4. With the exception of stishovite, all other above mentioned crystallization forms are composed of tetrahedral building blocks, in which every silicon atom is surrounded by four oxygen atoms. Stishovite, however, differs from the others: it adopts an octahedral coordination geometry in which every silicon atom is surrounded by six oxygen atoms and is even resistant against dissolution in hydrofluoric acid. Despite the common structure of building blocks amorphous silica glass differs from other crystallization forms: it does not possess any regular crystal structure and is metastable at room temperature. A rapid cool-down of the melted silica will lead to an incomplete crystallization because the crystallization cannot follow quickly enough. Owing to its amorphousness in this state, silica has a more diffuse short-range and no long-range order contrary to regular crystals of SiO_2 which possess a very clear short-range and long-range order. This is best observed in the partial radial pair distribution function $g_{ij}(r)$ calculated after Levine et al.¹³⁰ of two example states of SiO_2 , namely α -quartz and amorphous silica, shown in Figure 1.6c. This can also be evaluated experimentally by neutron or x-ray scatter-

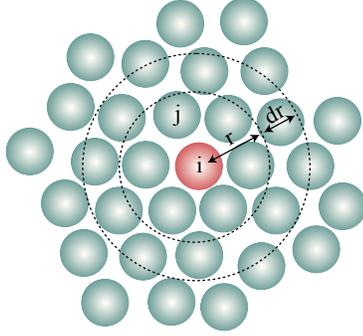


Figure 1.5: Sketch of the radial distribution function $g_{ij}(r)$ calculation. $g_{ij}(r)$ counts the atoms in a thin shell of width dr and distance r to a central atom.

ing allowing a linkage between experiments and simulations. $g_{ij}(r)$ denotes the local density change of species j as a function of the distance r to a reference particle i or simply put, it yields a probability of finding a particle of species j in a thin shell between r and $r + dr$ with particles of species i according to Figure 1.5 and is calculated with Equation (1.1).

$$g_{ij}(r) = \lim_{dr \rightarrow 0} \frac{\langle n(r) \rangle}{4\pi(N_{ij}/V)r^2 dr} \quad \text{with} \quad N_{ij} = \begin{cases} N_i(N_i - 1), & i = j \\ N_i N_j, & i \neq j \end{cases} \quad (1.1)$$

$n(r)$ is the number of pairs found in between r and $r + dr$ of the thin shell in Figure 1.5. The brackets denote an ensemble and time average. This is typically calculated from a thermodynamic ensemble and in the context of molecular dynamics simulations it takes the form

$$\langle n(r) \rangle = \frac{1}{N_{\text{frame}}} \sum_i^{N_{\text{frame}}} \sum_j^{N_j} \sum_{k, k \neq j}^{N_k} \delta(r - r_{ijk}), \quad (1.2)$$

where N_{frame} is the number of frames, r_{ijk} is the distance between atom j and atom k for frame i , and δ is the Dirac delta function. Back to Equation (1.1), N_{ij} is the number of unique

pairs and depends on the definition of i and j : if the radial distribution function is calculated between the same types of atoms than $i = j$ and $N_i = N_j$ yielding $N_{ij} = N_i(N_i - 1)$ otherwise $N_{ij} = N_i N_j$. N_i or N_j represents the total number of atoms with the respective type. Finally, V represents the total volume of the system. Note that the denominator of Equation (1.1) is equal to $\langle n(r) \rangle$ for an ideal gas.

Furthermore, a running coordination number $G_{ij}(r)$, defined as the number of closest neighbors of species j around species i within the radius r , can be obtained by cumulative integration of Equation (1.2) over dr :

$$G_{ij}(r) = \int_0^r 4\pi r'^2 \langle n(r') \rangle dr' , \quad (1.3)$$

For regular crystals this gives a stair-like behavior due to the highly ordered atoms on the lattice, in contrast to amorphous structures where mostly only the first few coordination shells are clearly visible. In Figure 1.6d the coordination number $G_{ij}(r)$ for Si and O is calculated from $g_{ij}(r)$ of Figure 1.6c for α -quartz and amorphous silica to clearly demonstrate the difference between them. It is evident that α -quartz and amorphous silica have nearly the same first coordination shell, organized as a tetrahedral structure with four oxygen atoms around a single silicon atom. The difference becomes apparent from the second coordination shell onward where α -quartz shows the previously mentioned stair-like behavior. A good overview about structural data, such as density, radial distribution functions or bond angle distributions, for silica in its pure form as SiO_2 can be found in the work of Mozzi & Warren¹⁵⁰.

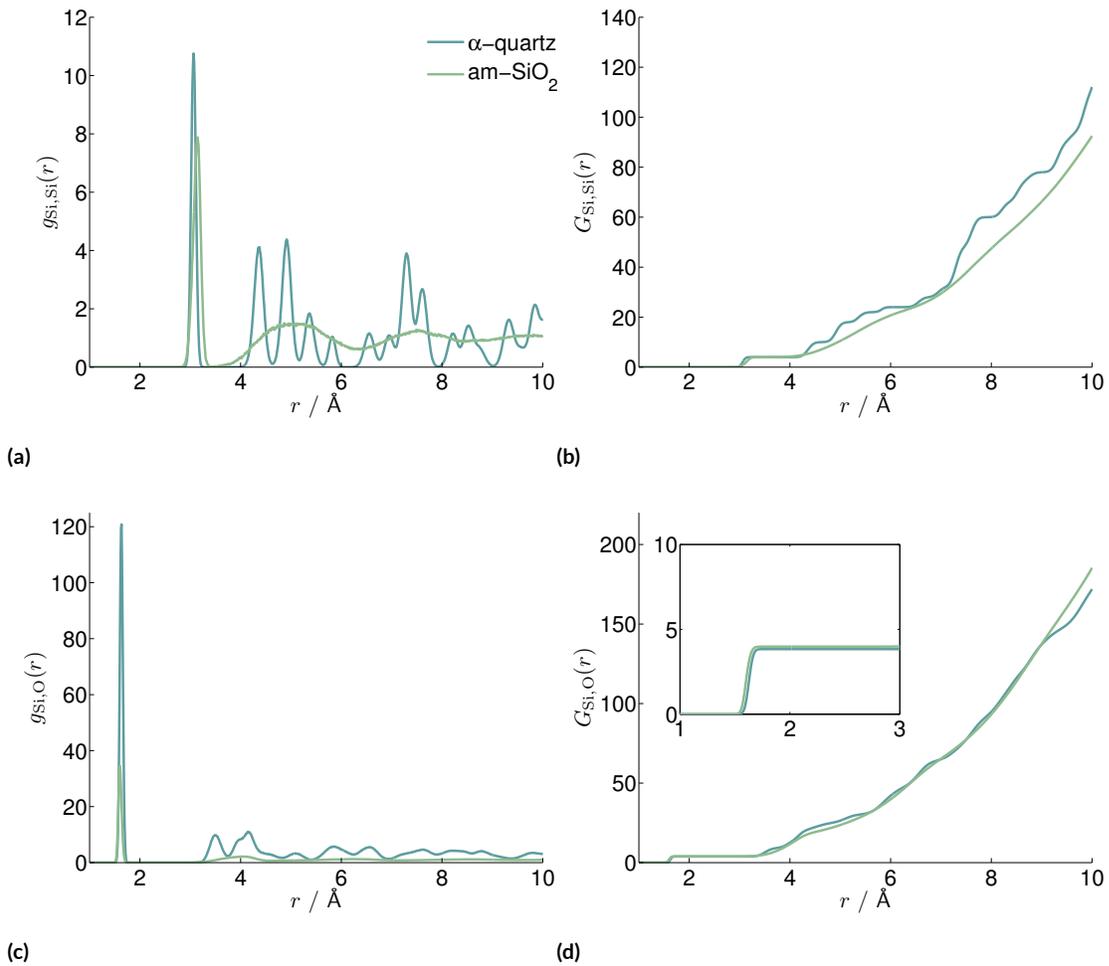


Figure 1.6: Partial radial pair distribution function $g_{Si,Si}(r)$ (a) and $g_{Si,O}(r)$ (c) obtained from molecular dynamics simulations at 300 °K of α -quartz using the BKS potential²²¹ and amorphous silica using the modified Demiralp potential (see Chapter 2.7.2). In (c) and (d) bulk amorphous silica and α -quartz coordination numbers, $G_{ij}(r)$ between Si and O atoms, are shown. An enlarged image for the first coordination shell of SiO₂ is shown in the inset of (d).

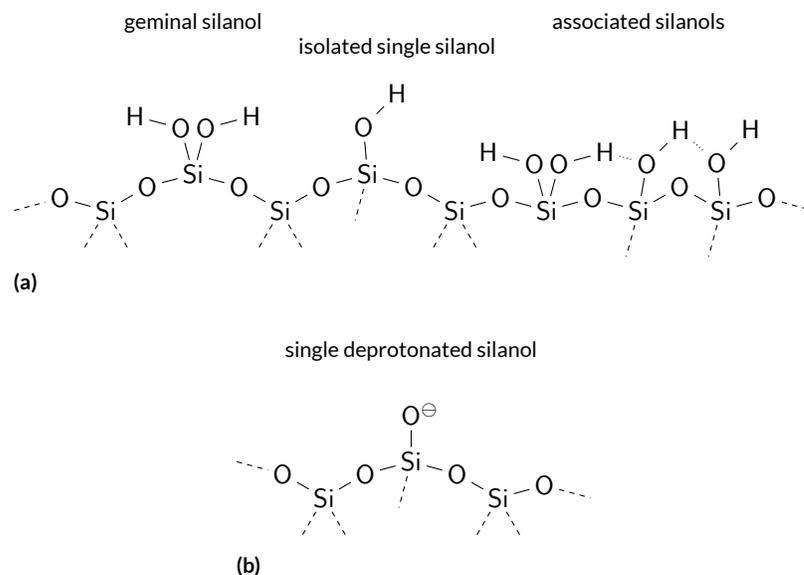


Figure 1.7: (a) Different types of silanol groups occurring on the silica surface. (b) Schematic illustration of a deprotonated silanol group on the surface, which leads to a net surface charge.

1.2.2 SURFACE PROPERTIES OF SILICA

In their utilization as pharmaceutical packaging interfacial effects at the glassy surface become ubiquitous and those differ substantially from its bulk properties. To a great extent the surface characteristics are determined by the number, organization and protonation state of silanol groups on the surface. Thereby Zhuravlev²⁴⁵ calculated an average silanol number of 4.9 OH groups per square nanometer from 100 independent measurements of maximal hydroxylated silica surfaces. As illustrated in Figure 1.7a silanols can be divided into three different classes: geminal, isolated single and associated silanols. In theory four different types of silanols exist, however, only those shown in Figure 1.7a have been found to occur in natural silicas; a silicon atom with three connected hydroxyl groups has not yet been experimentally found in natural silicas. Surface OH groups are subdivided by Zhuravlev²⁴⁶ as follows: (i) isolated free (single silanols), SiOH; (ii) geminal free (geminal silanols or silanediols), Si(OH)₂; (iii) vicinal, or bridged, or OH groups bound through the

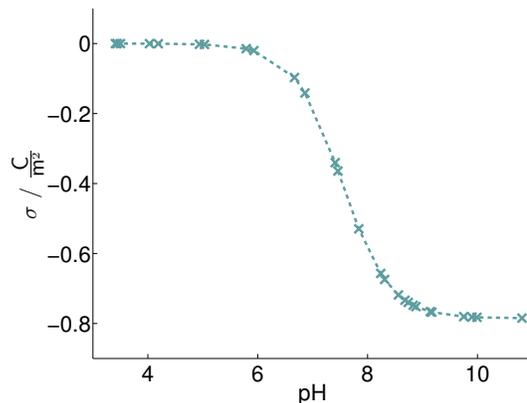


Figure 1.8: Surface charge of amorphous SiO₂ in accordance with pH. The shown data were kindly provided by Ludmilla Derr (unpublished results, ludmilla.derr@uni-bremen.de).

hydrogen bond (H-bonded single silanols, H-bonded geminals, and their H-bonded combinations). As mentioned Si(OH)₃ terminations may occur in theory but have never been found in naturally occurring silicas⁷². At last on the SiO₂ surface there also exist surface siloxane groups with oxygen atoms on the surface.

In contact with water, some of the hydroxyl groups at the surface of the amorphous silica will be deprotonated; this results in a negative surface charge (see Figure 1.7b for a schematic illustration). Only at its isoelectric point at pH 3¹⁶ the surface is neutrally charged. In this case all hydroxyls on the surface are protonated by the free hydrogen atoms from the solution. With further increasing pH more and more silanols will be deprotonated yielding a more negative surface charge until the complete surface is deprotonated. A further pH increase can thus not alter the surface charge as shown in Figure 1.8. The dependence on the surface charge with pH becomes of great importance, if for example the adsorption of a charged nano-particle or a bio-molecule with charged side-chains onto the silica surface is of interest. Moreover, this indirectly influences the adsorption of molecules due to the

formation of a denser water layer in proximity of the surface¹⁹⁰. Molecules with a hydration shell similar to the water layer adsorb easier on the surface compared to others with an incompatible hydration shell.

In addition, different salts can also alter the surface charge as they are able to screen the electrostatic potential of it. The sequence of increasing surface charge reads usually as $\text{Li}^+ < \text{Na}^+ < \text{K}^+ < \text{Rb}^+ < \text{Cs}^+$ and is said to be the normal series in comparison to the reverse one found for other oxides and occasionally for silica. Several explanations have been given for this and the most probable seems to root from coulombic interactions, where cations with a small and strong hydration shell will not approach the charged sites as closely as bigger and poorly dissolved ones. However, this can not explain the occasional occurrence of the reverse series on silica. A possible answer to this could be that this is only true for lower pH conditions or poorly deprotonated surfaces, as with increasing pH normally a stronger water layer will form on the silica surface and will eventually favour cations with tighter hydration shells – justifying the reverse order.

Nature laughs at the difficulties of integration.

Pierre-Simon de Laplace *

2

Representing Reality with MD Simulations: An Introduction

*Quoted in I. Gordon and S. Sorkin, *The Armchair Science Reader*, New York, 1959.

EFFECTS AT THE NANOSCALE become increasingly acknowledged as they are more and more applied to our daily life. However, underlying physical and chemical processes often remain uncomprehended—although nanoscopic effects have been successfully detected and they are measurable in experiments. These experiments are mostly not able to give a fundamental description of processes occurring at the nanoscale, e.g., the initial nucleation of crystal growth²⁴³.

With the famous Born-Oppenheimer approximation the involved particles are described solely by their nuclear coordinates. Tracing their motion with computational methods can augment experimental findings. *Ab-initio* methods yield the most accurate and reliable results and require no previous parametrization for they only depend on natural constants. The increasing amount of particles which have to be considered in the calculations for more complex problems, unfortunately, lead to a more and more computational expensive task. Such calculations are very time consuming since every electron of the involved particles has to be considered. Hence, *ab-initio* calculations are limited to a very small number of atoms.

The advantage of a classical molecular dynamics simulation over *ab-initio* or semi-empirical methods is their transferability to a broader range of problems. This holds, provided a small subset is accurately described by the classical force field in comparison with results from experimental findings or quantum chemical calculations. Parameter and potentials for classical molecular dynamics simulations are usually estimated by fitting the potentials to results from the above mentioned quantum chemical calculations or experimental findings. As a consequence, molecular dynamics simulations can only be as accurate as its underlying assumptions for their potential forms and parameter. The necessary computational effort of molecular dynamics simulations generally scales quadratically with the number

of involved atoms N . With spherical cutoffs or other techniques to treat efficiently the long-range interactions this is even more reduced to the order $\mathcal{O}(N)$. In contrast to this, even *ab-initio* methods which require the least computational effort, scale typically with the cube of atoms in the system, $\mathcal{O}(N^3)$. Depending on the used method they scale with even higher powers of N , e.g., post-Hartree–Fock methods; up to $\mathcal{O}(N^7)$ for coupled cluster calculations. However, recently electronic structure calculation methods have been proposed which employ a linear scaling with the number of atoms in the system²⁸.

A drawback in molecular dynamics simulations is, that it is only hardly possible to incorporate chemical bond breaking and forming processes between atoms—apart from hydrogen bonding. Classical molecular dynamics can only describe non-reactive systems. As a solution, hybrid quantum chemical and molecular mechanics (QM/MM) approaches²³³ and reactive force-fields^{32,222} have been proposed to overcome this problem. However, as long as solely hydrogen bonding has to be regarded it is mostly sufficient to simply use a fully classical molecular dynamics approach with unbreakable bonds which offers a good compromise between accuracy and speed.

2.1 EQUATIONS OF MOTION

Central to molecular dynamics is the stepwise temporal integration of the equations of motion. The evolution of the system consisting of N particles is usually described within the framework of classical Hamiltonian mechanics by a $6N$ dimensional canonical phase space spanned by the particle coordinates $\mathbf{r} = (r_1, \dots, r_{3N})$ and their conjugate momenta $\mathbf{p} = (p_1, \dots, p_{3N})$. Inherent are all possible configurations of the system with a single microscopic configuration in this phase space being depicted by a set of distinct canoni-

cal coordinates $\{\mathbf{r}; \mathbf{p}\}$. A motion from one configuration to another, corresponding to a temporal evolution of the system, can then be described analytically with the concept of Hamilton mechanics and its central quantity, the Hamilton function. It can be calculated via a Legendre transformation of the famous Lagrange function $\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) = T - V$, where T and V embody the kinetic and potential energy, respectively. Which is then

$$\mathcal{H}(\mathbf{r}, \mathbf{p}, t) = \sum_{i=1}^{3N} \dot{r}_i p_i - \mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) \quad (2.1)$$

There is neither an advantage nor disadvantage in solving problems with either method, as both the Lagrange and Hamiltonian formalism are only different interpretations of the same problem. However, historically the Hamilton formalism marks the starting point of the understanding of the relation between classical and quantum mechanics. By partial derivation of Equation (2.1) and under the additional use of the Lagrange equation $\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{r}_i} \right) = \frac{\partial \mathcal{L}}{\partial r_i}$ the canonical or Hamiltonian equations are obtained:

$$\dot{p}_i = \frac{\partial \mathcal{H}}{\partial r_i} \quad , \quad \dot{r}_i = \frac{\partial \mathcal{H}}{\partial p_i} \quad (2.2)$$

As a simple example the particle motion in a potential $V(\mathbf{r}, t)$ is considered. The spatial particle position is given by $\mathbf{r} = (r_x, r_y, r_z)$. In order to obtain the Hamiltonian of this problem the generalized momenta $p_x = \frac{\partial \mathcal{L}}{\partial \dot{r}_x}$, $p_y = \frac{\partial \mathcal{L}}{\partial \dot{r}_y}$ and $p_z = \frac{\partial \mathcal{L}}{\partial \dot{r}_z}$ are derived from the Lagrangian $\mathcal{L} = \frac{m\dot{\mathbf{r}}^2}{2} - V(\mathbf{r}, t)$, so that in turn the Hamiltonian of this problem can be written as

$$\mathcal{H} = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}, t) \quad (2.3)$$

From the canonical equations in (2.2) the equations of motion in vectorial form can be expressed as

$$\dot{\mathbf{p}} = -\vec{\nabla}V(\mathbf{r}), \quad \dot{\mathbf{r}} = \frac{\mathbf{p}}{m} \quad (2.4)$$

which ultimately gives Newton's second law $m\ddot{\mathbf{r}} = -\vec{\nabla}V(\mathbf{r}) = \mathbf{F}$, where \mathbf{F} denotes the force acting on the particle. This simple example holds also for systems comprising of many particles. The force \mathbf{F}_i acting on particle i is then

$$m_i\ddot{\mathbf{r}}_i = -\vec{\nabla}V(\mathbf{r}_i) = \mathbf{F}_i \quad (2.5)$$

Assuming the knowledge of the potential energy $V(\mathbf{r})$ and the initial conditions $\{\mathbf{r}(t = 0); \mathbf{p}(t = 0)\}$ the equations of motion in (2.4) can be integrated and the classical motions for all particles of the system are retrieved for all times t :

$$\mathbf{p}(t) = \int_0^t dt' \vec{\nabla}V(\mathbf{r}(t')) + \mathbf{p}(0) \quad (2.6)$$

$$\mathbf{r}(t) = \int_0^t dt' \frac{\mathbf{p}(t')}{m} + \mathbf{r}(0) \quad (2.7)$$

The form of the inter-atomic potentials yielding the total potential energy $V(\mathbf{r})$, in general, depends strongly on the physical and chemical properties of the system. It is inevitable that assumptions must be made for the functional form of the underlying potential in order to reproduce the correct physical behavior. In most cases interactions are split into bonded, non-bonded and many-body contributions which are discussed more detailed in Chapter 2.5.

2.2 INTEGRATING NEWTON'S EQUATIONS OF MOTION

Predicting the movement of atoms in a quick and accurate manner requires an efficient solving of newton's equations of motion. A short overview is given here by reciting Chapter 6.4 "Solving the equation of motions" from the book of Berendsen¹⁷. Considering a system of N particles with mass m_i , coordinates \mathbf{r}_i and a defined recipe to compute the total potential energy $E_{pot} = V(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = V(\mathbf{r})$. Accordingly, the force on particle i can be evaluated by $\mathbf{F}_i = -\Delta_i V(\mathbf{r}_i)$. The equations of motion are thus simply again Newton's equations:

$$\dot{\mathbf{r}}_i = \mathbf{v}_i \tag{2.8}$$

$$\dot{\mathbf{v}}_i = \mathbf{F}_i/m_i = \mathbf{f}_i \tag{2.9}$$

They will produce a microcanonical ensemble under the assumption that these equations are properly solvable. A microcanonical ensemble is a system with constant amount of particles (N), constant volume (V) and energy (E), abbreviated as NVE. However, in practice one typically encounters numerical errors that are caused by deviations from the ideal analytical behavior. The reason is that Equation (2.8) and (2.9) are not solved analytically and discrete timesteps Δt have to be introduced in order to allow a numerical algorithm. Thus, the total energy will not be exactly conserved. Furthermore, errors in the forces (e.g., due to truncation of interactions which is discussed further in Chapter 2.5.1) will produce pseudorandom disturbances causing energies to drift. Since the temperature is determined by the equipartition theorem it may drift even when equilibrium has been attained. Therefore, to ensure that a trajectory is stable over long timescales, modifications to the Newtonian

equations of motion have to be made. The solution of the equations of motions with a discrete timestep Δt implies several requirements on the integration algorithm; It should be (i) time invariant, (ii) a symplectic integration algorithm and (iii) ideally require only one force evaluation per time step.

In most molecular dynamic simulations the velocity-Verlet algorithm—a simple, robust, reversible and symplectic algorithm—introduced in the following is used. Aside from the velocity-Verlet algorithm, there exist other methods and variants, e.g., the leap-frog algorithm¹⁰², the Reference System Propagator Algorithms (RESPA)²²⁰ and the Gear algorithm⁸⁵. The latter has proven to be quite accurate for small time steps but is not very stable for larger time steps²²⁴ and neither it is time-reversible nor symplectic. Thus, the Gear algorithm has to large extent been superseded by other algorithms.

2.2.1 THE VELOCITY-VERLET ALGORITHM

This algorithm has its origin in the work of Verlet²²⁷ and is a slightly modified version of it²⁰⁷. The velocity-Verlet algorithm employs a discretisation of the second derivative of the Cartesian coordinates of every particle. It is assumed, for notational simplicity, that there is only one degree of freedom $r(t)$. Expanding the position $r(t)$ and $v(t)$ at time $t + \Delta t$ with a Taylor series to second order yields

$$r(t + \Delta t) = r(t) + \Delta t \cdot \dot{r}(t) + \frac{\Delta t^2}{2} \ddot{r}(t) + \mathcal{O}(\Delta t^3) \quad (2.10)$$

$$v(t + \Delta t) = v(t) + \Delta t \cdot \dot{v}(t) + \frac{\Delta t^2}{2} \ddot{v}(t) + \mathcal{O}(\Delta t^3) \quad (2.11)$$

Using the equality $f(t) = F(t)/m = \ddot{r}(t)$ and $v(t) = \dot{r}(t)$ leads to the expressions

$$r(t + \Delta t) = r(t) + \Delta t \cdot v(t) + \frac{\Delta t^2}{2} f(t) + \mathcal{O}(\Delta t^3) \quad (2.12)$$

$$v(t + \Delta t) = v(t) + \Delta t \cdot f(t) + \frac{\Delta t^2}{2} \dot{f}(t) + \mathcal{O}(\Delta t^3) \quad (2.13)$$

Finally, only an expression for $\dot{v}(t)$ is needed. This can be obtained by a Taylor series expansion of

$$\dot{v}(t + \Delta t) = \dot{v}(t) + \Delta t \cdot \ddot{v}(t) + \mathcal{O}(\Delta t^2) \quad (2.14)$$

It is enough to consider terms of the order of Δt^2 here because only an approximation up to order Δt^3 to the quantity $\frac{\Delta t^2}{2} \ddot{v}(t)$ is required¹⁷. Multiplying with $\frac{\Delta t}{2}$ and rearranging gives

$$\frac{\Delta t^2}{2} \ddot{v}(t) = \frac{\Delta t}{2} [f(t + \Delta t) - f(t)] + \mathcal{O}(\Delta t^3) \quad (2.15)$$

By inserting this into Equation (2.13),

$$v(t + \Delta t) = v(t) + \frac{\Delta t}{2} [f(t + \Delta t) + f(t)] + \mathcal{O}(\Delta t^3) \quad (2.16)$$

is obtained. If needed, e.g., to implement a temperature control or to measure kinetic energy, the velocity of the particle in the next iteration step can be computed directly. Furthermore, the velocity-Verlet algorithm has the advantage over the original Verlet algorithm in that it is self-starting and does not need another algorithm to obtain estimates of positions and velocities for the first few steps. Although this algorithm needs the force at the new time step, there is still only one force evaluation per step.

The only difference between the leap-frog, Verlet or velocity-Verlet algorithms is that dif-

ferent orders of the Taylor series expansion are considered; hence the underlying theoretical description is the same in every case, which is not immediately obvious²¹.

2.3 PERIODIC BOUNDARY CONDITIONS

The number of atoms that can be treated ranges from a few hundred atoms to about multiples of millions as impressively demonstrated in the work of Zhao et al.; the capsid structure of the mature human immunodeficiency virus-1 (HIV-1) was simulated in an all-atom molecular dynamics simulation with 64 million atoms²⁴⁴. Simulations of even greater macroscopic systems, e.g., the interplay between two capsid structures, are, however, still not feasible with molecular dynamics simulations. Boundary effects can contribute significantly to the system properties in an artificial way as perturbing reflective walls or a vacuum have to be introduced. A practical solution to approximate macroscopic size systems is given by replicating the system under study in all three dimensions. A periodic lattice consisting of cells which are exact copies of the original unit cell is thus created. In the simplest case the central cell is of rectangular shape. Particles in the replicated images of the central cell interact with particles in the original central cell and thus, by using periodic boundary conditions, this mimics a system of infinitely extended size. Walls which prevent atoms from moving through the periodic boundaries are typically neglected. This implies that if an atom crosses, e.g., a lower periodic boundary, it is inserted again on the opposite site. A special case are surface simulations in which one dimension is not periodic or the surface slab is repeated periodically in one direction with enough space between the slabs in order to neglect the interaction between them.

In molecular dynamics usually two kinds of interactions exist: short-range and long-

range interactions. The former obeys the minimum image convention in which each individual particle in the simulation interacts with the closest image of the remaining particles in the periodic lattice. Long-range interactions, such as Coulomb interactions, need a separate treatment. An efficient method for this is introduced in Chapter 2.5.1.

Special precautions must be made to avoid artifacts due to the application of periodic boundary conditions which could lead to, e.g., an unphysical topology. In practice, simulation boxes smaller than half the cut-off of the longest short-range interactions are problematic since an atom can interact with itself. However, this often is a compromise at the expense of accuracy, as artificial errors in the force field are introduced due to a shorter cut-off (cf. Chapter 2.5.1). It is also commonly suggested that, for example in the case of DNA simulations, an at least 1 nm thick solvent is added around the molecule of interest in every direction to prevent artificial interactions of the molecule's "head" and "tail" over the periodic boundary⁶⁴.

2.4 CONSTRAINED EQUATIONS OF MOTION

Constraints are introduced to simulate, for example, rigid molecules. Examples are (the notation $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ is used): (i) distance constraints $\mathbf{r}_{ij}^2 - \mathbf{d}_{ij}^2 = 0$, or (ii) angle constraints $\mathbf{r}_{ij} \cdot \mathbf{r}_{kj} - c = 0$, where $c = d_{ij}d_{kj} \cos \Phi$.

In all-atom molecular dynamics simulation they are often used to allow a larger time step by constraining distinct motions; in most cases covalent bonds involving hydrogens. Bonded interactions between atoms are described typically in biomolecular force fields with harmonic springs or sinusoidal terms, e.g., bonds and bending or dihedral angles. A description of biomolecular force fields can be found in Chapter 2.7.2.

Bonds involving hydrogen atoms normally possess the highest vibrational frequency. This is mainly due to the low atomic mass of hydrogen compared to other atoms and the often lower spring constant in the parameterization compared to other covalent bonds. In general, atomic masses of typical atoms in biomolecules are, at least, tenfold from that of hydrogen.

To roughly estimate vibrational frequencies and their associated minimum timestep of bonds involving hydrogen it is assumed that the bonded partner of the hydrogen atom has an infinite mass (i.e. C_∞ , N_∞ and O_∞). In this view the hydrogen atom is bound via a spring to a fixed wall and its position is given by the coordinate $x(t)$. Solving the differential equation of this system by approximating a harmonic oscillator,

$$F = m_H a = m_H \frac{d^2 x(t)}{dt^2} = -kx(t) \quad (2.17)$$

yields:

$$x(t) = A \cos(\omega t + \Phi) \quad \text{with} \quad \omega = \sqrt{\frac{k}{m_H}} = \frac{2\pi}{T} \quad (2.18)$$

In chemistry vibrational frequencies are usually expressed in terms of the number of vibrations, i.e., wavenumbers, that would occur in the time that light travels one centimeter:

$$\nu = \frac{1}{2\pi c} \sqrt{\frac{k}{m_H}} \quad / \quad (cm^{-1}) \quad (2.19)$$

Using this equation and spring constants k as defined in, e.g., the AMBER force field⁵⁴, approximate wavenumbers for C_∞ -H, N_∞ -H or O_∞ -H vibrations can be estimated to 2072, 2253 and 2543 cm^{-1} , respectively. A precise wavenumber prediction is complicated, since

Group	ν [cm ⁻¹]	Group	ν [cm ⁻¹]
— C - H	3300	≡ C - Br	560
≡ C - H	3020	≡ C - I	500
≡ C - H	2960	— O - H	3600
— C ≡ C —	2050	> N - H	3350
> C = C <	1650	≡ P = O	1295
≡ C - C ≡	900	> S = O	1310
— S - H	2500	≡ C - Cl	650
— N = N —	1600	≡ C - F	1100
> C = O	1700	— C ≡ N	2100

Table 2.1: Infrared group wavenumber table. Data taken from Bernath²².

atoms are mostly not bound to just one another atom. When two or more normal vibrations have the same symmetry a full normal coordinate analysis must be performed²³⁸. Results of this method for a small collection of bonds are shown in Table 2.1.

After the Nyquist theorem, the sampling frequency has to be twice as high as the highest frequency of the signal; a minimum condition for adequate sampling of a signal which also applies for molecular dynamics. The next highest frequency in a biomolecule would be the C=O bond, if all bonds involving hydrogens are constrained and triple bonded carbons are not present. Using 1700 cm⁻¹ from Table 2.1 for C=O type bonds, the maximum timestep size would be 10 fs. Twice as high as if covalent hydrogen bonds were included. Thus, if covalent bonds involving hydrogens are constrained to their equilibrium distance, a two times longer timestep can be chosen. However, imagine that by this approach only two points of an oscillating period are sampled. In reality this is often too less and systems will diverge. Thus, to be on the safe side, a timestep of 2 fs is usually chosen if covalent hydrogen bonds were constrained and 0.5 fs if covalent hydrogen bond vibrations have to be included.

In order to constrain distinct motions, several algorithms have been proposed^{187,2,148,99}. The SHAKE¹⁸⁷ algorithm is a simple and popular method. It is used to solve constraint

equations, $\sigma(\mathbf{r}) = 0$, which could be a bond or angle depending on atomic coordinates.

In conjunction with the original Verlet algorithm²²⁷ and applied in Cartesian coordinates it reads

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} [\mathbf{F}_i^u(t) + \mathbf{F}_i^c(t)] \quad (2.20)$$

where \mathbf{F}^u is the force from the force field and \mathbf{F}^c is the force induced by the constraints. The constraint force on particle i at time t is given by¹⁷

$$\mathbf{F}_i^c(t) = \sum_s \lambda_s(t) \frac{\partial \sigma_s}{\partial \mathbf{r}_i} \quad (2.21)$$

Effectively, the constraint force adds a second contribution to the displacement of particles.

At first the algorithm computes the new position, $\mathbf{r}_i(t + \Delta t)$, without taking any constraints into account:

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} \mathbf{F}_i^u(t), \quad (2.22)$$

and then corrects the positions with $\Delta \mathbf{r}_i$ such that

$$\sigma_s(\mathbf{r}(t + \Delta t) + \Delta \mathbf{r}) = 0, \quad s = 1, \dots, m, \quad (2.23)$$

where

$$\Delta \mathbf{r}_i = \frac{\Delta t^2}{m_i} \sum_s \lambda_s(t) \frac{\partial \sigma_s(\mathbf{r}(t))}{\partial \mathbf{r}_i} \quad (2.24)$$

and m is the amount of constraints. A set of m coupled equations is thus spanned for the λ 's. Usually they are non-linear and must to be solved iteratively. In SHAKE the constraints are solved sequentially and the whole procedure is iterated to convergence. Another approach would be to linearize and then solve them as a set of linear equations. The interested

reader is referred to chapter 15.8.2 of Berendsen¹⁷ for a further description and application to a simple dimer system.

2.5 EMPIRICAL FORCE FIELDS

It is evident that a description of molecular processes is often based on experimental observations of physical and chemical processes. The true nature of such processes is often only understood by interpreting the theoretical framework. With the upcoming of quantum chemical calculations, e.g., density functional theory or more elaborate methods, it was possible to understand comprehensively what occurs on the molecular level and beyond in their electron structure. Several highly complex experiments have recently validated quantum chemical calculations. Very impressive ones are the surveys of Gross et al.⁹¹ where the chemical structure of pentacene is resolved by atomic force microscopy, or Giessibl⁸⁶ and de Oteyza et al.⁶³ where covalent bond structures of single-molecule chemical reactions were imaged. Combining experimental results and quantum chemical calculations is a powerful tool to obtain parameter for empirical force fields. However, the general structure of empirical force fields or rather their potential form and most important their physical meaning is mandatory to represent reality most accurately. Although empirical force fields are only a very scarce approximation to physical and chemical processes at atomistic scales, they often produce reliable results of systems which would not have been feasible with quantum chemical calculations.

Commonly, empirical force fields are split into non-bonded and bonded contributions or more precisely intra- and inter-molecular interactions. Non-bonded interaction act, as the name implies, on atoms which are not linked by covalent bonds (or are separated

by at least 3 bonds in standard biomolecular force fields). It is often enough to approximate their contribution to the potential energy through a pairwise interaction potential $V_2(\mathbf{r}_i, \mathbf{r}_j)$ given in Equation (2.25). Subject to the features one wants to explore also three-body $V_3(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k)$ and higher order terms can be added to this equation.

$$V(\mathbf{r}) = \sum_i V_{\text{ext}}(\mathbf{r}_i) + \sum_i \sum_{j>i} V_2(\mathbf{r}_i, \mathbf{r}_j) + \dots \quad (2.25)$$

The sum runs over all $i = 1 \dots N$ atoms whereas $V_{\text{ext}}(\mathbf{r}_i)$ represents an external field or system walls.

In the simplest form pairwise interaction potentials are composed of Coulombic and van der Waals interactions, in which the latter is often approximated by a Lennard–Jones potential. A detailed explanation is found in appendix B.2 and B.3, respectively. The non-bonded potential is then retrieved by the summation of a Lennard–Jones and Coulomb potential acting between the atoms of species i and j ,

$$V_2(\mathbf{r}_i, \mathbf{r}_j) = 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] + \frac{q_1 q_2}{4\pi\epsilon_p r_{ij}} \quad \text{with } r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|. \quad (2.26)$$

The form of the Lennard–Jones potential has various representations, however, its different representations are all analytically transformable into each other. In general they are composed of two terms: a r^{-12} term representing the repulsion between atoms and a r^{-6} representing Keesom, Debye and London forces (further explained in Chapter B.3). In literature sometimes a Lennard–Jones representation is found which uses a r^{-9} term for the repulsion and a r^{-6} term for the Keesom, Debye and London forces. This expression is typically used for hydrogen bonds, however, in most biomolecular force fields this has been

replaced by the 12-6 representation. In the present case σ is defined as the collision diameter, which is also a separation distance for which the energy is zero. ϵ is defined as an energy given by the well depth. The second term in Equation (2.26) represents the Coulombic interaction, where q_i and q_j are the point charges on the atoms and ϵ_p is the permittivity constant.

One way to represent forces and potentials between two kinds of atoms in polyatomic systems with the above mentioned formulation of non-bonded interactions is to define σ and ϵ for individual atom types. As an example, potential parameter for silicon and oxygen ϵ_{SiO} and σ_{SiO} are obtained by combining the individual parameter ϵ_{Si} , ϵ_{O} , σ_{Si} and σ_{O} using one of the common combination rules. A widely used rule is the geometric (or arithmetic) or Lorentz-Berthelot combination, yielding for

$$\epsilon_{\text{SiO}} = \sqrt{\epsilon_{\text{Si}}\epsilon_{\text{O}}} \quad \text{and} \quad \sigma_{\text{SiO}} = \frac{\sigma_{\text{Si}} + \sigma_{\text{O}}}{2}. \quad (2.27)$$

Lorentz-Berthelot combination is most successful when applied to similar species. Its major failing is, however, that well depths are often overestimated which is owed to the geometric mean rule for ϵ .

As a final remark, in reality it is almost impossible to make a clear distinction between van der Waals and electrostatic interactions. Both contributions occur mostly simultaneously and are hence difficult to separate. Fundamental quantum chemical calculations, advanced theories or complex experiments can offer further information thereof.

Bonded interactions, in contrast, are composed mainly of three contributions: (i) bond, (ii) angle and (iii) torsional terms. Some force fields also have the ability to use improper or cross diagonal terms which link, for example, the influence of a stretched bond to the

change in an angle. On one hand this adds more complexity and on the other hand eventually more accuracy to the force field. A drawback is that bonded interactions in classical molecular dynamics simulations do not allow formation or breaking of bonds which drastically reduces their area of application. Hybrid methods using quantum chemical and molecular mechanics (QM/MM) calculations in combination have been developed to overcome this problematic. However, they are not yet well established and, despite their problems of energy and force conservation, often consume a tenfold of computational resources compared to fully classical simulations.

2.5.1 CUT-OFF SCHEMES AND LONG-RANGE INTERACTIONS

With increasing distance to the atom i for which Equation (2.26) is evaluated the computational effort increases with $\Theta(N^3)$. N represents the total amount of atoms of the simulation. A cutoff in Equation (2.26) reduces the computational effort needed by neglecting contributions from atoms behind the cutoff in the calculations. As a result of which, errors in the potential energy or in the force calculation could arise due to the fact that a fraction of the potential energy is being ignored or, depending on how steep the shift to zero at the cutoff is, forces at the cutoff are highly overestimated. Several cutoff schemes to avoid this circumstance are shortly introduced in the following.

The most common cutoff schemes are, apart from simple truncation, either shifting or switching of the force or the potential energy. Most importantly is that for potentials with long-range characteristics, e.g., Coulomb interactions, a long enough cutoff is chosen. This should be at least 12 Å in the case of Coulomb interactions. Worth to mention is, furthermore, that a potential switching region shorter than 4 Å reduces the protein motion artifi-

cially and results in an unnatural behavior^{133,197}.

In contrast to this, switching or shifting the force at the cutoff yields in many cases more accurate results than introducing a cutoff in the potential energy. Most of the time it is more important to obtain better estimates of forces at the cutoff than potential energies since particle motions are calculated from forces. Therefore, if computational resources are limited and higher order approaches could not be applied, force switching might be the most appropriate cutoff-scheme. It gives a smooth transition to zero forces within the switching region and does not worsen the force upon changing the switching region²⁰². Neither does it change the potential energy landscape as in force or potential shifting methods. Nonetheless, in general it is far more important to employ a long enough cutoff, especially for the case of Coulomb interactions²⁰². A detailed explanation of force shifting and switching cutoff schemes for Coulomb and van der Waals interactions between atoms with fixed charges is given in appendix B.4.

TREATING LONG-RANGE INTERACTIONS WITH AN EWALD SUMMATION

If cutoff schemes are considered unsuitable, other mechanisms to treat effectively the long-range interactions such as Ewald summation or similar approaches have to be used. Periodic boundary conditions are often used to approximate an infinite system and long-range interactions play a significant role in such systems. A general rule of thumb is that interactions decaying no faster than r^{-n} , where n is the dimensionality of the system, have often a signal at distances greater than half the smallest length of the simulation box. If those interactions are not treated correctly, they are often the source of misbehaving atoms and faulty simulations. Especially charge-charge interactions, which decay with r^{-1} , are particularly

problematic. A too short cutoff will neglect a crucial part of the potential energy and a cutoff greater than half of the smallest simulation box length in a simulation where periodic boundary conditions apply will allow atoms to interact with themselves. Both introduce some serious artefacts.

If such interactions have to be treated more accurately, Ewald summation⁷⁶ and similar approaches¹⁰² are the method of choice. Long-range interactions are treated by those methods in a more sophisticated manner. However, if such methods are to be used for an accurate description of long-range Coulomb interactions, they are often restricted to charge neutral systems. The Ewald summation serves in principle as the foundation for nearly all higher order methods and is introduced therefore briefly hereafter. The potential energy of the system is in principle split into a real and fourier space part to avoid the conditional convergence of the true electrostatic potential. The partitioning of the electrostatic potential results in the necessity of introducing correctional terms for the self interactions of atoms and polarizability effects. In the end the total coulombic energy is expressed as

$$\begin{aligned}
V &= V^{(\text{RS})} + V^{(\text{FS})} + V^{(\text{SI})} + V^{(\text{DC})} \\
&= \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} \frac{q_i q_j}{8\pi\epsilon_p} \frac{\text{erfc}(\alpha|\mathbf{r}_{ij} + \mathbf{nL}|)}{|\mathbf{r}_{ij} + \mathbf{nL}|} \\
&+ \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{k} \in \mathcal{M}} \frac{q_i q_j}{2L^3 \epsilon_p k^2} e^{-\frac{k^2}{4\alpha^2}} e^{i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)} \\
&- \sum_{i=1}^N \frac{\alpha q_i^2}{4\pi^{\frac{3}{2}} \epsilon_p} + \frac{1}{L^3 \epsilon_p (4\epsilon_r' + 2)} \left(\sum_{i=1}^N q_i \mathbf{r}_i \right)^2
\end{aligned} \tag{2.28}$$

A detailed derivation of the terms in Equation (2.28) can be found in appendix B.4.2.

2.6 THERMOSTATS AND BAROSTATS

Most experiments are carried out at, for instance, a specific constant temperature. In a simulation, this means that a thermostat is required. However, in a microcanonical simulation the temperature will fluctuate. Only by averaging over the trajectory and using an appropriate thermostat, a true canonical temperature can be obtained. At a later point this is explained in more detail. All there is to say about this for now is, that quantities, such as the temperature, depend on the sampling of the phase space or in other words of all possible configurations of the system. In some cases they rely especially on those which are not energetically favorable, problematic for regular molecular dynamics simulation. By controlling the distribution of energy on the particles, which is exchanged between the heat bath and the system, a virtual temperature can be maintained. In general the temperature is directly related to the kinetic energy, equipartitioned equally with $k_B T/2$ on each degree of freedom,

$$\sum_{i=1}^N \frac{|\mathbf{p}_i|^2}{2m_i} = \frac{k_B T}{2} (3N - N_c), \quad (2.29)$$

In this equation, \mathbf{p}_i is the momentum of particle i and m_i is its mass. N_c is the number of constraints on the system, e.g. when the total linear momentum is constrained to a value of zero, N_c would equal 3. Several different physical interpretations to distribute this energy are proposed and can be classified into four groups¹⁷:

STOCHASTIC METHODS , involving the application of stochastic and friction forces which mimic elastic collisions with light particles forming an ideal gas at a given temperature. This produces an canonical ensemble but disturbs system dynamics.

STRONG-COUPLING METHODS apply a constraint to the quantity coupled to an external

bath. This yields non-Hamiltonian dynamics but can produce, by applying a Gauss thermostat, a canonical distribution in phase space.

WEAK-COUPLING METHODS apply small perturbations to the constraint property, aiming to reduce it smoothly to preset value. Dynamics are non-Hamiltonian, generating an ensemble between microcanonical and canonical for temperature scaling and isochoric and isobaric for coordinate scaling.

EXTENDED SYSTEM DYNAMICS extends system with an extra degree of freedom related to the quantity desired to control. Dynamics of extended system remain fully Hamiltonian, but the dynamics of molecular systems is disturbed.

Moreover, conserving energy is not trivial in molecular dynamics simulation. Due to rounding errors, cut-offs schemes or the choice of the time step a small fraction of the energy stored in the system is lost or could be overestimated. Those numerical errors will accumulate every time energy and forces are computed. Avoiding this unphysical behavior of cooling down until freezing or heating up until, literally speaking, explosion of the system can be achieved, again, by coupling the system to a thermostat. Particular caution is recommended if a barostat is applied to interfacial systems. Different compressibilities of the compartments could cause unrealistic pressure effects. A detailed explanation on this topic can be found in Yancey et al.²⁴².

There exists a few abbreviations typically used to distinct between different applications of thermostats, barostats or both: A system where particles N , volume V and energy E is hold fixed is referred to as a NVE or micro canonical ensemble. In contrast to this, systems which it is allowed to exchange energy with a heat bath are referred to NVT ensembles or

canonical ensembles. Particle number N , volume V and temperature T are held fixed here. Alternatively the temperature of the heat bath can be increased or decreased slightly between to adjacent time steps, which will heat up or cool down the system in a controlled manner. Furthermore, systems which are also allowed to expand or shrink in volume according to the applied pressure p are referred to NpT or isothermal-isobaric ensembles. Somewhat peculiar and more or less seldom used are NPH or isenthalpic ensembles. Only marginally noted here are simulation at constant chemical potential, abbreviated as μVT or grand canonical ensembles. Two widely used thermo- and barostats are introduced, the first belongs to the group of weak-coupling methods and the second to the group of extended system dynamics.

2.6.1 BERENDSEN THERMOSTAT AND BAROSTAT

A Berendsen¹⁹ thermostat will scale the velocities of all particles at each step according to the difference between the temperature of the heat bath T_0 and the actual temperature of the system T at time t , such that it is proportional to the temperature change:

$$\frac{dT}{dt} = \frac{1}{\tau}(T_0 - T). \quad (2.30)$$

τ is the coupling parameter which determines how tightly the bath and system are coupled together. This is accomplished by adding a frictional term to the equations of motion from (2.5):

$$m_i \ddot{\mathbf{r}}_i = \mathbf{F}_i + m_i \frac{1}{2\tau} \left(\frac{T_0}{T} - 1 \right) \dot{\mathbf{r}}_i. \quad (2.31)$$

It represents a proportional scaling of the velocities per time step of

$$\lambda = \left(1 + \frac{\Delta t}{\tau} \left[\frac{T_0}{T} - 1 \right] \right)^{1/2}. \quad (2.32)$$

Similar to this, a coupling to a constant pressure bath can be achieved similar to the approach of applying a thermostat. The resulting pressure scaling becomes then

$$\mu = \left(1 + \frac{\Delta t}{\tau_p} \beta_T (p - p_0) \right)^{1/3}, \quad (2.33)$$

such that coordinates and volume are scaled according to

$$\mathbf{r}' = \mu \mathbf{r} \quad (2.34)$$

$$V' = \mu^3 V. \quad (2.35)$$

τ_p is the coupling strength to the external pressure bath and $\beta_T = -\frac{1}{V} \frac{\partial V}{\partial p}$ is the isothermal compressibility. Weak coupling has the advantage of a smooth response of the system, but on the other hand it poses the drawback of not generating a well-known ensemble and furthermore fluctuations cannot be used.

2.6.2 NOSÉ-HOOVER THERMOSTAT AND BAROSTAT

Widely used is the Nosé-Hoover^{158,157,105} thermostat which is capable of generating trajectories consistent with a canonical ensemble. With this approach, a heat bath is coupled to the simulated system via an additional degree of freedom s in the Hamiltonian of Equation (2.1). In this case the Hamiltonian is defined for a system comprising of many particles

and is thus,

$$\mathcal{H} = \sum_i^N \frac{\mathbf{p}_i^2}{2m_i s^2} + V(\mathbf{r}_1, \dots, \mathbf{r}_N) + \frac{p_s^2}{2Q} + g k_B T \ln(s). \quad (2.36)$$

p_s is the conjugate momentum of s , Q is a parameter of dimension energy \cdot (time)² and behaves as a mass for the motion of s , k_B is Boltzmann's constant and T is the externally set temperature. Furthermore virtual time and coordinates are used in this formulation. They are related to the real coordinates and time as follows:

$$\mathbf{r}' = \mathbf{r}, \mathbf{p}' = \frac{\mathbf{p}}{s} \text{ and } t' = \int^t \frac{d\tau}{s} \quad (2.37)$$

g equals the number of degrees of freedom of the physical system. However, it depends on whether virtual or real time sampling is used; In the first case it should then be chosen to be $3N + 1$ and in the latter $3N$. Assuming now that the Hamiltonian formalism can be applied to Equation (2.36) with its virtual variables, the resulting equations of motion are

$$\dot{\mathbf{r}}_i = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i s^2}, \quad \dot{\mathbf{p}}_i = -\frac{\partial \mathcal{H}}{\partial \mathbf{r}_i} = -\frac{\partial V(\mathbf{r}_1, \dots, \mathbf{r}_N)}{\partial \mathbf{r}_i} \quad (2.38)$$

and for the additional degree of freedom s we obtain

$$\dot{s} = \frac{\partial \mathcal{H}}{\partial p_s} = \frac{p_s}{Q}, \quad \dot{p}_s = -\frac{\partial \mathcal{H}}{\partial s} = -\left(\sum_i^N \frac{\mathbf{p}_i^2}{m_i s^2} - g k_B T \right) \frac{1}{s} \quad (2.39)$$

It can be shown (see Nosé¹⁵⁸ for details of this proof) that an ensemble average with this extended system formalism for $t \rightarrow \infty$ equals that of a canonical ensemble. In contrast to this, with real coordinates and time we obtain non-Hamiltonian dynamics. Nonetheless, it can be shown that the equilibrium phase space distribution in real coordinates and time is

still canonical¹⁰³. However, it has been criticized by Toxvaerd & Olsen²¹⁸ that its behavior is non-ergodic. Hence Martyna et al.¹⁴⁵ formulated an alternative approach by introducing a chain of Nosé-Hoover thermostats, each scaling its predecessor. With real coordinates and time the equations of motion for each particle i are thus

$$\frac{d\mathbf{r}'_i}{dt'} = \frac{\mathbf{p}'_i}{m} \quad , \quad \frac{d\mathbf{p}'_i}{dt'} = -\frac{\partial V}{\partial \mathbf{r}'_i} - s'_1 p'_{s_1} \frac{\mathbf{p}'_i}{Q_1} \quad , \quad (2.40)$$

Similar to the derivation of a Nosé-Hoover thermostat, an analogous barostat is obtained by introducing an additional degree of freedom for the desired pressure in the Hamiltonian of Equation (2.36) by replacing the set temperature T with the set pressure p .

2.7 COMPREHENSIVE FORCE FIELDS FOR SILICA, WATER AND BIOMOLECULES

THERE IS NO STANDARD PROCEDURE to simulate interfacial systems; Usually every part, such as the biomolecule, water or the solid state surface is described with another force field. Careful considerations must be given when combining different force fields, as they are often not compatible to each other. Several classical potentials for crystalline and amorphous SiO_2 have been developed over decades. Examples of simple but widely used interatomic potentials, able to capture some properties of bulk silica, are those developed by van Beest et al.²²¹ and Vashishta et al.²²⁶. They are often used as a basis for interfacial silica/water systems, making it necessary to introduce them briefly. Figure 2.1 shows a graphical comparison of the two-body parts of both potentials. The former has the simple form of a Buckingham-

ham and Coulomb potential acting between atoms i and j ,

$$V(r_{ij}) = A_{ij}e^{-b_{ij}r_{ij}} - \frac{c_{ij}}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon_p r_{ij}}, \quad \text{with } r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \quad (2.41)$$

For details concerning the included parameters in the force field see appendix B.1. At close proximity of atoms the van Beest et al.²²¹ force field fails to describe the repulsive behavior between two atoms, as the Buckingham potential becomes very attractive after a repulsive barrier for interactions were A_{ij} is nonzero. When applied to silica surfaces it is known, that it will generate too many undercoordinated and overcoordinated surface atoms¹⁴⁷, and does not describe well the small-scale (~ 0.5 nm) structural elements of the surface, e.g. small rings, as compared to high-level comparative DFT calculations⁴⁷. The situation is different with the latter potential by Vashishta et al.²²⁶, consisting of two- and three-body terms,

$$V(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{1 \leq i < j \leq N} V_2(r_{ij}) + \sum_{1 \leq i < j < k \leq N} V_3(r_{ij}, r_{jk}, r_{ik}). \quad (2.42)$$

r_{ij} , r_{jk} and r_{ik} are the absolute values of distances between atoms i , j and k ; $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$, etc. The two-body part of the potential, $V_2(r_{ij})$, consists of three terms: Steric repulsion due to ionic sizes, Coulomb interactions to take into account charge transfer, and charge dipole interaction to include the effect of electronic polarizabilities. The following form is used

$$V_2(r_{ij}) = \frac{e^2}{4\pi\epsilon_p} \left(\frac{H_{ij}}{r_{ij}^n} + \frac{q_i q_j}{r} - \frac{\frac{1}{2}(\alpha_i q_j^2 + \alpha_j q_i^2)}{r^4} e^{-r/r_4s} \right). \quad (2.43)$$

e represents the elementary charge and q_i and q_j are given in terms of unitless multiples of it ($q_{Si} = 1.6$ and $q_O = -0.8$). The expression for the three-body interaction is similar to the

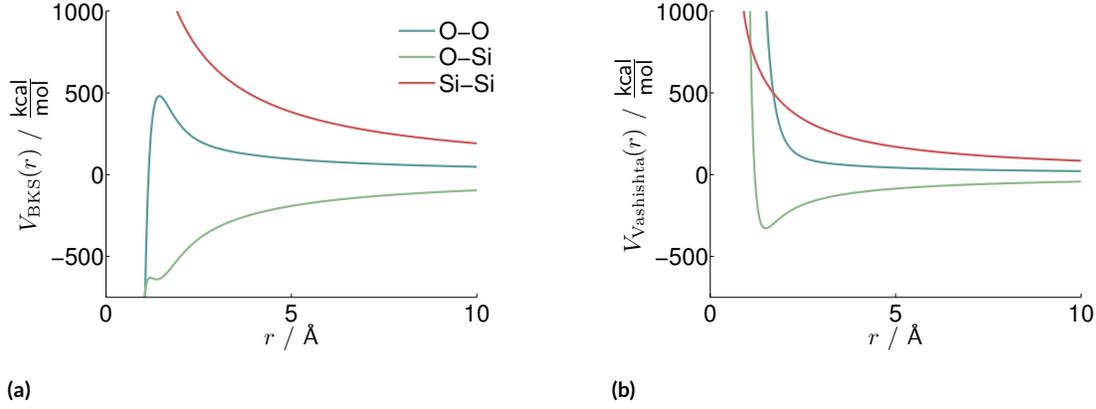


Figure 2.1: Comparison of potential energies for Si-Si, Si-O and O-O interactions between van Beest et al.²²¹ (a) and Vashishta et al.²²⁶ (b) force fields. All used parameter are given in appendix B.1.

formulation of the Stillinger & Weber²⁰³ potential and is given to

$$V_3(r_{ij}, r_{jk}, r_{ik}) = B_{jik} f(r_{ij}, r_{ik}) p(\Theta_{jik}, \bar{\Theta}_{jik}), \quad (2.44)$$

$f(r_{ij}, r_{ik})$ and $p(\Theta_{jik}, \bar{\Theta}_{jik})$ are defined in Equation (2.44) to

$$f(r_{ij}, r_{jk}) = \begin{cases} \exp\left(\frac{1}{r_{ij}-r_0} + \frac{1}{r_{ik}-r_0}\right) & \text{for } r_{ij}, r_{ik} \leq r_0 \\ 0 & \text{for } r_{ij}, r_{ik} > r_0 \end{cases} \quad (2.45)$$

$$p(\Theta_{jik}, \bar{\Theta}_{jik}) = (\cos \Theta_{jik} - \cos \bar{\Theta}_{jik})^2 \quad (2.46)$$

This potential has the advantage over the van Beest et al.²²¹ functional form that it stays repulsive even if atoms come very close. Furthermore it better reproduces Si-O-Si and O-Si-O angle distributions due to the three-body term. Unfortunately this happens at the cost of additional computational time needed for three-body calculations.

It should be noted that Figure 2.1b deviates from Figure 1 of the original publication from Vashishta et al.²²⁶. Even after careful inspection of Equation (2.43) no source of error could be found. It is thus assumed that the original illustration of the two-body interactions in the publication of Vashishta et al.²²⁶ is erroneous.

Another point is that simulations of, generally speaking, siliceous glasses which not only consists of silicon and oxygen by means of classical force fields can be hard task. The complexity of the force fields and its according parameter increase as the number of constituents increase. Pedone¹⁶⁹ proposed additional parameter for the use in his potential to model interactions of, amongst others, for example of boron in borosilicate or sodium in soda-lime glasses. Although parameters are available, this work will focus on silica glasses in its pure SiO₂ polymorphic form, for which compatibility to well established classical biological force fields can be ensured.

On a final note, silica/water force fields should be easily usable in combination with existing biomolecular force fields such as AMBER⁶⁹, CHARMM³³ or OPLS¹⁵ for biology-related simulations.

2.7.1 INTERFACIAL SIMULATIONS USING A REACTIVE FORCE FIELD

It is of crucial importance to represent the real silica surface as good as possible with an atomistic model. As a first attempt, the REAXFF forcefield^{222,223,48} in addition with the recently published parameter for SiO₂ and water⁷⁹ was used; A promising approach to obtain a naturally hydroxylated silica surface without the otherwise usually manual saturation of bridging oxygens or undercoordinated silicon atoms with hydrogen or hydroxyl groups, respectively.

REAXFF is continuously enhanced by adding new functionals or parameter necessary for a better inclusion of additional chemical processes. Hence, several implementations of REAXFF for many other systems have been proposed over the time, e.g. for carbon based systems²²², silicon and silicodioxide polymorphs^{35,34,223}, transition-metal-catalyzed reactions¹⁵⁵, high-energy materials²⁰⁴, or aqueous glycine¹⁷⁷, to name but a few of the options available. However, it also has one great failing: Although they share the same functional principle, different implementations of REAXFF are not necessarily combinable. This is due to the underlying fundamental assumptions and therefore the inclusion of different functional equations. Nonetheless, allowing a chemical description of bond breaking and formation at nearly the same computational cost of a classical molecular dynamics simulation is a major benefit of REAXFF.

Based on the assumption that the bond order BO' between two interacting atoms i and j can be described as a function of their distance r_{ij} , the potential energy in REAXFF is calculated on the basis of bond orders. The bond order is thereby split into contributions from sigma-, pi- and double pi-bonds:

$$\begin{aligned}
 BO'_{ij} &= BO'_{ij}{}^{\sigma} + BO'_{ij}{}^{\pi} + BO'_{ij}{}^{\pi\pi} \\
 &= \exp \left[p_{bo1} \cdot \left(\frac{r_{ij}}{r_0^{\sigma}} \right)^{p_{bo2}} \right] + \exp \left[p_{bo3} \cdot \left(\frac{r_{ij}}{r_0^{\pi}} \right)^{p_{bo4}} \right] + \exp \left[p_{bo5} \cdot \left(\frac{r_{ij}}{r_0^{\pi\pi}} \right)^{p_{bo6}} \right] \quad (2.47)
 \end{aligned}$$

Contributions to the potential energy of, e.g., bonds or angles between atoms are therefore estimated from the bond order. However, the possibility to describe reactive systems is accompanied in this case with a great complexity of the force field. Of particular importance is that bonded energy contributions, e.g. from bond, angle or dihedral terms, vanish

on bond dissociation. Furthermore the parametrization of such force fields is indeed not trivial and lacks often a straightforward approach. Furthermore, all reaction pathways, particularly of complex reactions, have to be included and the outcome is only as good as the information that have been provided in the parameterization process. The total energy is estimated from a vast collection of different energy contributions, which inclusions depends on the respective implementation of REAXFF. A complete energy term description of the REAXFF implementation used in this work for SiO₂ systems can be found in the appendix B.1.3.

Briefly, the following procedure was used to generate a bulk silica model from scratch: Starting from a random distribution of Si and O atoms in the correct stoichiometric ratio, the bulk silica model was then converted into a dry slab by introducing a vacuum layer in between the bulk system. Further generation of a hydroxylated silica surface was then accomplished by introducing water molecules in the vacuum layer. In doing this, the dry surface was allowed to react chemically with the solvent; Allowing the formation of silanol groups on the surface.

GENERATION OF BULK SILICA AND ITS STRUCTURAL FEATURES

A bulk silica model was created from scratch by depositing randomly 794 oxygen and 397 silicon atoms in a periodically repeated box comprising the volume of 30·30·20 Å³. This gives the desired density of 2.2 g/cm³¹⁵⁰. Due to the random deposition of silicon and oxygen atoms, high energy atomic overlaps are present in the system. Removing those artifacts was achieved with a NVE simulation and by employing a soft potential acting between

atoms i and j in the form of

$$V(r_{ij}) = k \left(1 + \cos \left(\frac{\pi r_{ij}}{r_c} \right) \right). \quad (2.48)$$

r_c and the energy constant k were chosen to 10 Å and 57.65 kcal/mol, respectively. The initial application of REAXFF for the equilibration process failed in this case, as the charge equilibration method was unable to assign correct charges. The system was equilibrated for 10000 steps with a timestep of 0.25 fs, yielding a pre-equilibrated system. After using this procedure it was possible to use REAXFF for a further 12.5 ps equilibration. As in Fogarty et al.⁷⁹ described, the same annealing and cooling protocol with a timestep of 0.5 fs was used to obtain a realistic bulk silica model: The equilibrated system was annealed to 4000 K within 75 ns and afterwards cooled down to 300 K within 148 ns using a Nosé-Hoover thermostat. A second annealing and cooling cycle in the same fashion as the previous one followed, this time employing a Berendsen barostat and thermostat. In the end, the same structural data for bulk SiO_2 as with Fogarty et al.⁷⁹ was obtained.

HYDROXILATION OF DRY SILICA

All further simulations are undertaken using a Nosé-Hoover thermostat alone and without the utilization of an additional barostat. This is due to the fact that a barostat used on an interfacial systems would cause unrealistic pressure defects due to the different compressibility of silica and water^{79,242}.

A 40 Å thick slab of vacuum was placed in z-direction of the bulk silica model, resulting in two dry silica surfaces. The system was again heated up to 1000 K and equilibrated at this temperature for further 250 ps to avoid unphysical atomic structures on the surface and re-

lax the system. Once more the system was cooled down to 300 K, this time within 250 ps. Water molecules were added to the vacuum slab, corresponding to a density of approximately 0.998 g/cm^3 . The ability of REAXFF to create or break bonds allowed dissociation of water molecules and silanol formation on the dry silica surface. The underlying chemical reaction can be written as



Hydroxylation of the dry silica surface began thus at the time of starting the dynamics without manual intervention. The resulting hydroxylated surface is shown in Figure 2.3a. Indicated with green and blue colors are surface defects, which will be explained in more detail soon after.

The simulation box size was adjusted in the first one thousand steps to match approximately the desired water density of 0.998 g/cm^3 ¹⁴² in a segment far away from the surface. A further 1 ns production run followed afterwards. Density profiles were obtained by averaging over 0.1 \AA thick density layers in z-direction using the last 400 ps of the simulation. If the desired water density was not reached, the procedure was iteratively repeated until the system comprised the correct volume to reproduce the desired water density. The resulting density profile from the last volume adjustment is shown in Figure 2.2.

Besides regular LAMMPS output REAXFF generates additional information during the simulation. Amongst others, very useful thereof is the bond information of atoms during the simulation. Calculated from this are the temporal evolution of silanol groups on the surface and dissociation of water molecules, the distribution of highly reactive four membered siloxane rings and coordination defects of silicon and oxygen atoms along the z-axis of the simulation box, shown in 2.3. In figures 2.3c to 2.3e the approximate extent of the

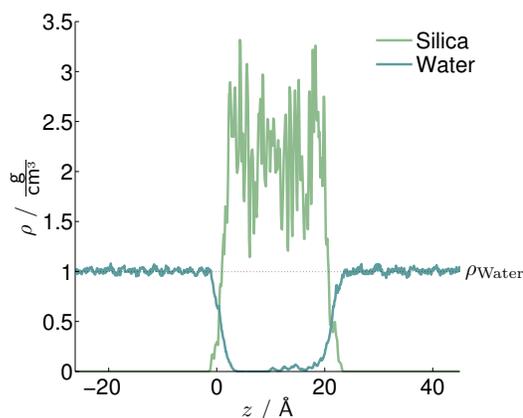


Figure 2.2: Vertical density profile of the hydroxylated SiO_2 slab from the ReaxFF simulation in contact with water. Data was averaged over the last 400 ps of the silica/water ReaxFF simulation.

silica slab is given by the area shaded in grey. Coordination states and molecular fragment distributions are calculated by averaging over the last 400 ps of the simulation to give a correct representation of the long-time stability of the surface defects. A closer look on this data reveals some artificial defects not found in natural silica. Particular problematic are the continuous presence of under- and overcoordinated silicon (see Figure 2.3d) and partly oxygen atoms (see Figure 2.3e), highly reactive four membered siloxane rings (see Figure 2.3c) and the continuous existence of $\text{Si}(\text{OH})_3$ groups on the surface (see Figure 2.2). Over- and undercoordinated silicon atoms and four membered siloxane rings can be found in the bulk, however, they should be abundant on the surface due to the high reactivity with water of these groups^{77,41}. El Shafei⁷² stated that $\text{Si}(\text{OH})_3$ structures can exist in theory, but no positive evidence has been presented to support their actual presence in any known compound. Nonetheless, one example can be found in literature: Severin & Vankan¹⁹⁴ showed the occurrence of silanetriol groups on the surface. However, this group was found to have a low thermal stability; Possibly explaining its absence in natural silicas. Moreover, Brei³¹

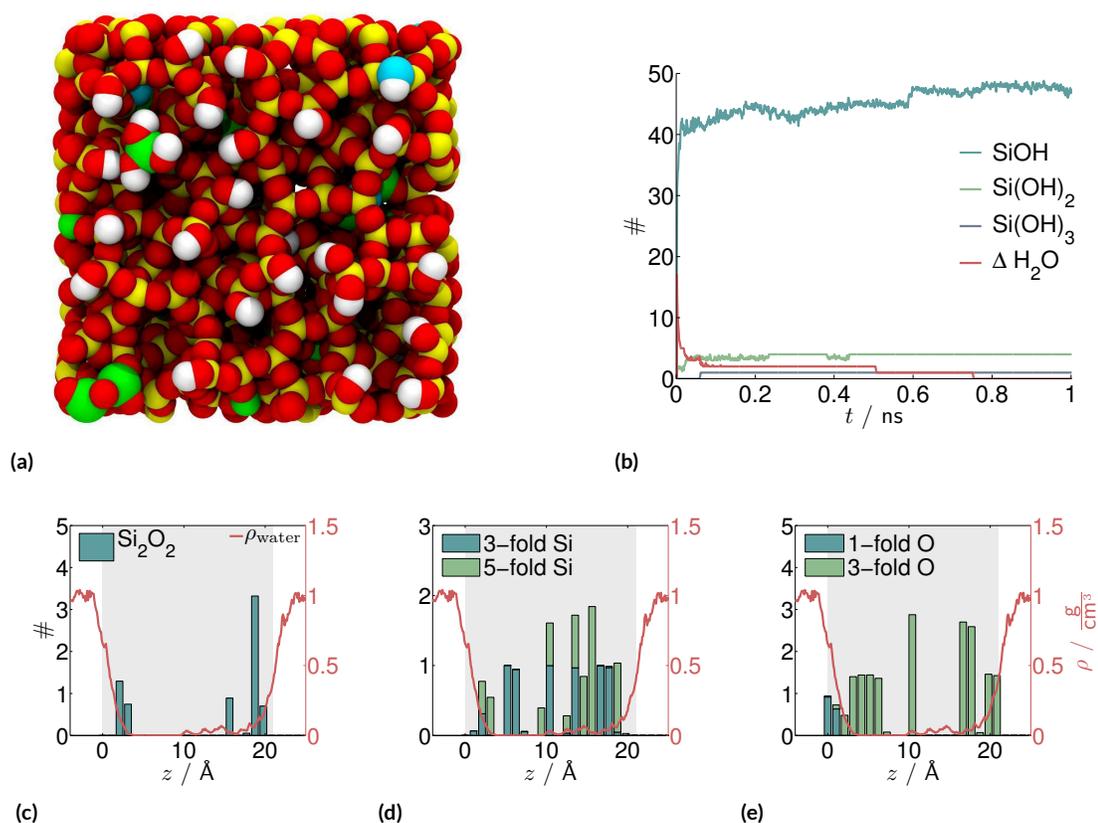


Figure 2.3: (a) Hydroxylated silica surface generated using ReaxFF with a silica/water system⁷⁹. Silicon, oxygen and hydrogen atoms are depicted in red, yellow and white, respectively. Surface defects are highlighted in blue and green, indicating defects on O and Si. (b) Temporal evolution of different silanol groups occurring on the surface and dissociation of water molecules during the course of the ReaxFF surface hydroxylation. Data is smoothed using a moving average filter with a window size of five elements. (c) Highly strained 4-membered siloxane rings, (d) 3- and 5-fold coordinated silicon atoms and (e) 1- and 3-fold coordinated oxygen atoms distributions along z-axis after surface hydroxylation using ReaxFF. Alongside with the histograms is the water density shown (—).

attempted to perform the same reaction on aerosil, but failed to produce any silanetriol. Concerning oxygen defects on the silica surface, it is known that undercoordinated oxygen atoms can occur on the surface, since silica has a negative surface charge at neutral pH associated with a deprotonation of oxygen atoms on the surface⁴³. Overcoordinated oxygen atoms on the surface, however, should not remain in presence when the silica surface is in contact with water.

It was not possible to eradicate the unphysical surface defects with REAXFF alone, e.g. by letting the system react over longer time scales or using a higher temperature. Thus, without further manual surface modifications of the silica slab those defects will not vanish. Reasons for this can be the still insufficient simulation time or the wrong preparation of the bulk silicodioxide structure. For the latter, Fogarty et al.⁷⁹ created their starting structure by tetragonal building blocks randomly positioned in the simulation box. However, as the system is liquefied at 4000 K, this should not have an impact in the end. REAXFF's incapability to reproduce a realistic silica surface when in contact with water and its incompatibility to well-established biomolecular force fields resulted in its discarding of further use. Albeit it should be noted that there has been made progress recently by Rahaman et al.¹⁷⁷ in describing the tautomerization of a single glycine amino acids soluted in water with REAXFF. A combined description of silica and aqueous glycine with REAXFF is, although the same water model is used, still a distant reality due to missing parameterizations of atomic interactions between silica and glycine.

2.7.2 INTERFACIAL SIMULATIONS USING CLASSICAL FORCE FIELDS

In the following, firstly widely used biological force fields are introduced. This is followed by a discussion on how to ensure bio-compatibility of classical force fields for silica. A force field for the interface between water and silica is presented, which is also applicable to biomolecular systems. In the end, a modified silica force field is proposed to be used if a flexible silica surface is desired in the simulations of biomolecular interactions with silica.

INTRODUCTION OF BIOLOGICAL FORCE FIELDS

Several biological force fields have been proposed over decades for the simulation of, e.g., proteins, lipids or nucleic acids. Widely known and often used are the AMBER⁵⁴, CHARMM³³ or OPLS⁵⁵ force fields. The difference between them lies primarily in the parameterization procedure of e.g. charges or van der Waals parameter. Moreover, they have slightly different functional forms and are mainly parameterized for the use with a distinct water model. This is of particular importance as molecules and surfaces interact primarily with each other via their solvation shell. Thus, an accurate description of water molecules is of significant importance. To further increase accuracy especially in terms of free energy estimates, instead of using the associated water model of the biomolecular force field it is often also feasible to use them along with the TIP4P¹⁰⁸ water model¹⁶⁵. This is possible due to the close similarity of van der Waals parameter and charges between TIP3P and TIP4P water models. However, usually a higher accuracy goes along with additional computational time needed since additional charges have to be considered in the calculations.

As the AMBER force field has been used in the present work its functional form is introduced shortly. It normally employs the TIP3P water model in which O-H bonds and

H-O-H angles are constrained to their equilibrium values. It takes the following functional form

$$\begin{aligned}
 V(\mathbf{r}^N) = & \sum_{\text{bonds}} k_b (r_{ij} - r_0)^2 + \sum_{\text{angles}} k_\theta (\theta_{ijk} - \theta_0)^2 \\
 & + \sum_{\text{dihedrals}} \frac{k_\phi}{2} \left[1 + \cos(n\phi_{ijkl} - \gamma) \right] \\
 & + \sum_{i=1}^N \sum_{j=i+1}^N \left(4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_p r_{ij}} \right), \quad (2.50)
 \end{aligned}$$

with $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. $V(\mathbf{r}^N)$ denotes the potential energy which is a function of the positions $\mathbf{r}_1, \dots, \mathbf{r}_N = \mathbf{r}^N$ of all N atoms. Aside from intermolecular terms, i.e. van der Waals and Coulombic terms, this model employs intramolecular terms describing the topology and structure of a molecule. Chemical bonds and angles between two bonds are modeled by a harmonic potential with spring constants k_b and k_θ and their equilibrium constants r_0 and θ_0 , respectively. Dihedral angles describe the torsion of four successive atoms around a central bond by the dihedral force constant k_ϕ , the periodicity n and the phase γ . Closely related to dihedral angles are improper dihedrals. They describe the out-of-plane torsion of a planar group of atoms and have the same functional form as dihedral potentials.

Intermolecular terms govern primarily interactions between different molecules and between topological remote parts of the same molecule. A Lennard–Jones potential is used to approximate interatomic repulsion and van der Waals interactions. Electrostatic interactions are described by a Coulomb potential. Aside from that, it is common practice to use Lennard–Jones potentials to describe all kinds of contributions that are not captured by the aforementioned potentials. In *AMBER* Lennard–Jones coefficients of pair interactions

between two atoms i and j are calculated from atomic coefficients ϵ_i , ϵ_j , σ_i and σ_j by applying Lorentz-Berthelot combination rules. Thus $\epsilon_{ij} = \sqrt{\epsilon_i \epsilon_j}$ and $\sigma_{ij} = \sigma_i + \sigma_j$. Care must be taken, however, if a simulation package other than AMBER is used. Sometimes other combination rules apply for Lennard–Jones interactions of the respective simulation package and coefficients have to be transformed in order to give back the correct pair interactions.

Note that biomolecular force fields often scale non-bonded interactions between bonded atoms: Between neighboring (1-2) and next neighboring (1-3) atoms interactions are often neglected, whereas interactions between atoms separated by three bonds (1-4) are often scaled to 20-50 %, depending on the specific implementation of the force field. AMBER treats remote atom contributions in a molecule by including Lennard–Jones and Coulombic interactions of atoms that are separated by more than three bonds. Those interactions are normally referred to as 1-4 interactions. Contributions from atoms that are separated by three consecutive bonds are scaled by a factor of 0.5 for Lennard–Jones and 0.8333 for Coulombic interactions. Contributions from atoms that are further away than four consecutive bonds will contribute unscaled. Atoms that are only separated by one or two bonds will not contribute via Lennard–Jones or Coulombic interactions. Electrostatic interactions are based on partial charges q_i which are usually determined from ab initio calculations using the RESP¹⁴ method.

A BIOCOMPATIBLE INTERFACIAL SILICA FORCE FIELD

Modifications to biomolecular force fields are not desirable since they have been tested, approved and validated over decades. Of particular importance is, furthermore, that they have been parameterized for the use with a distinct water model. A biocompatible silica/water

force field should thus optimally fulfill compatibility to biomolecular force fields and their associated water models. The need for such force fields is immense and several silica/water force fields have been proposed recently. The Clay force field⁵⁸ describes accurately the quartz/water interface and uses SPC²⁰ water, but is not designed to describe bulk silica as all bonding terms other than those for the surface hydroxyls are absent. The potential of Lopes et al.¹³⁴ is a CHARMM compatible model and captures well the properties of quartz surfaces in combination with the TIP3P^{114,182} water model. Cruz-Chu et al.⁵⁷ proposed a CHARMM silica/water potential by fitting atomic charges and van der Waals parameter of silica to reproduce contact angle measurements of water on a amorphous silica by molecular dynamics simulation. Despite of their similar functional form, the Cygan et al.⁵⁸, Lopes et al.¹³⁴ and Cruz-Chu et al.⁵⁷ empirical force fields have been parameterized differently. A recent comparison of the above mentioned can be found in the work of Skelton et al.^{195,196}. Another attempt to describe accurately the silica/water interface was undertaken by Hassanali & Singer⁹⁵. They augmented the van Beest et al.²²¹ potential with silanol three-body terms and used it in combination with the SPC/E¹⁸ water model. Cole et al.⁵¹ combines the Vashishta et al.²²⁶ potential for silica with a three-body Stillinger & Weber²⁰³ potential for the silicon atoms to describe both silica and natively oxidized silicon in contact with liquid TIP3P water. Last but not least, the force field of Pedone¹⁶⁹ takes explicitly silanol hydrogen-bonding terms on a hydroxylated silica surface into account. Except for REAXFF for silica and water, all of the aforementioned models for SiO₂ describe neutral silica/water interfaces. However, this assumption is valid under strong acidic conditions only, as the isoelectric point of SiO₂ lies between 2 and 3 (see Chapter 1.2.2). The model proposed by Hassanali et al.⁹⁶ includes additional parameter for deprotonated oxygen atoms. How-

ever, basis for their charge assignment has been the van Beest et al.²²¹ force field. Charges therein are defined for silicon and oxygen atoms to 2.4 e and -1.2 e, respectively. These are much larger than those of any of the other force fields introduced hitherto. They appear to poorly perform when used for surface systems and have been used rather more controversially^{147,47,179,4}. Indeed, atomic point charges of the surface atoms should be able to reliably reproduce the electrostatic potential (ESP) in surface proximity calculated by quantum mechanical methods, as done for instance by Lopes et al.¹³⁴. This would ensure optimal mapping of the true electrostatic interaction energy in terms of a simple Coulomb potential, in line with the approach used to develop biomolecular force fields. Moreover, the extent of charge redistribution around the terminal O⁻ atoms following silanol deprotonation has never been studied so far, motivating a thorough analysis based on quantum mechanical calculations, as presented by Butenuth et al.⁴³. Parts of this publication arose in the course of this work. Moreover, it serves as the basis for all further simulations, resulting in the necessity to introduce this force field in more detail herein.

Several quantum chemical and molecular dynamics calculations of different silica systems were performed to obtain correct charges for silicon, oxygen and hydrogen atoms in the bulk and on the surface. As it is dealt with explicitly deprotonated O⁻ terminal groups, an *ad hoc* charge modification is required. The used models for the calculation of atomic charge differences after the deprotonation of terminal silanol group are (i) a periodically repeated slab model of the (0001) surface of α -quartz, (ii) a small periodically repeated amorphous SiO₂ slab in contact with liquid water, (iii) a deprotonated amorphous SiO₂ cluster carved out from a large amorphous SiO₂ slab and (iv) the amorphous SiO₂ slab itself terminated with OH groups to saturate the artificial dangling bonds. The latter was kindly

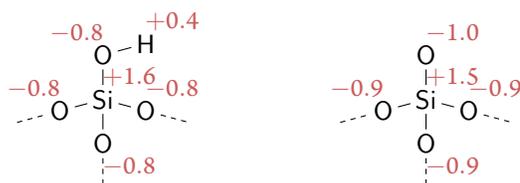


Figure 2.4: Graphical illustration of the novel deprotonated silica surfaces parametrization. Units of charges are in e .

provided by Daniel J. Cole, who also investigated this surface in his work³⁰. The absolute charge values in this force field were determined from best-fitting to the electrostatic potential computed ab-initio (ESP charges). Further analysis of the redistribution of charges after deprotonation of a hydroxyl group via Bader and ESP charge considerations results in the following novel parameterization:

- (i) Charges on silicon and oxygen atoms are set to 1.6 and -0.8 e , respectively.
- (ii) The charge on a hydrogen atom of any terminal OH group is set to 0.4 e .
- (iii) Deprotonated, terminal O^- atoms are assigned a charge of -1.0 e .
- (iv) The first neighbor silicon atom and the three second-neighbor oxygen atoms of the deprotonated, terminal O^- atoms are assigned 1.5 and -0.9 e , respectively.

This is again graphically illustrated in Figure 2.4. It should be noted that the same values can be assigned to neutral or singly-deprotonated geminal $Si(OH)_2$ terminal groups. In the latter case, the oxygen atom of the non-deprotonated hydroxyl group would have a charge of -0.9 e .

Aside from electrostatic, van der Waals type interactions are modeled by a Lennard-Jones potential with parameter summarized in Table 2.2. Lennard-Jones parameter of

	Si	OB	OH	H	OM
This Work					
q	1.600	-0.800	-0.800	0.400	-1.000
σ	1.600	1.762	1.650	1.000	1.650
ϵ	0.300	0.261	0.150	0.021	0.150
Lopes et al. ¹³⁴					
q	1.080	-0.530 ^a	-0.540	0.320	
σ	2.000	1.770	1.770	0.224	
ϵ	0.600	0.152	0.152	0.046	
Cygan et al. ⁵⁸					
q	2.100	-1.050	-0.955	0.425	
σ	1.853	1.777	1.777	0.000	
ϵ	1.86×10^{-6}	0.155	0.155	0.000	
Cruz-Chu et al. ⁵⁷					
q	0.900	-0.450	-0.660	0.430	
σ	2.147	1.750	1.770	0.224	
ϵ	0.300	0.152	0.152	0.046	
Hassanali et al. ⁹⁶					
q	2.400	-1.200	-1.200	0.600	-1.600

Table 2.2: Charges and Lennard–Jones parameters for the interactions between water and SiO₂ surfaces. Parameters used in this work are compared to those of other force fields, taken from Skelton et al.¹⁹⁵ and adapted to match combination rules of the Amber implementation. The units for q , σ and ϵ are electrons, Å and kcal/mol, OB and OM refer to O atoms in bulk SiO₂ and in deprotonated silanol groups, respectively. ^aThe OB charge depends on the selected surface.

other silica/water force fields are given additionally in comparison. The Lennard–Jones parameter definition based on parameter which can be found originally in the work of Cole et al.⁵⁰. However, those parameter could not be simply overtaken and had to be revised for the following reason. The original parameterization included an explicit “hydrogen-bond” interaction between terminal hydroxyl groups and the water solvent, specific to older AMBER force field versions¹⁶⁸. This interaction became recently obsolete and is not included anymore in novel AMBER force field versions, nor in other fixed-charges force fields such as, e.g., CHARMM.

To test the performance of the newly parameterized charges and Lennard–Jones parameter summarized in Figure 2.4 and Table 2.2, energies of a single TIP_{3P} water molecule placed in various orientations over three different SiO₂ structures (neutral amorphous SiO₂ cluster, deprotonated amorphous SiO₂ cluster and a natively oxidized silicon surface) at increasing surface-molecule distances were calculated. A detailed discussion of the excellent results can be found in Butenuth et al.⁴³). Summarized, the novel parameter set reproduces energies of natively oxidized silicon and neutral and deprotonated amorphous SiO₂ surfaces in an unprecedented accuracy as compared to estimates obtained for the same systems with density functional or second order Møller-Plesset perturbation theory.

As a final test of the performance of our parameter set, heat of immersion estimates of the neutral amorphous SiO₂ surface slab from Cole et al.⁵⁰ are computed. The results are shown in Table 2.3. In particular, calculations were performed using three slightly different modifications of the TIP_{3P} water model, namely (i) the so-called modified TIP_{3P} (mTIP_{3P}) model where Lennard–Jones parameter of hydrogen atoms are explicitly considered and (ii) the original TIP_{3P}¹⁴⁴ model and (iii) the modification by Price & Brooks¹⁷³ introduced

Hydroxyl Charges (O, H)	E_{im} [mJ/m ²]		
	mTip3p	Tip3p	Tip3pEw
-0.6, +0.2	451.1	352.5	340.6
-0.7, +0.3	276.7	178.1	166.2
-0.8, +0.4	102.8	4.2	-7.8

Table 2.3: Heat of immersion of neutral am-SiO₂ calculated for different OH charges and Tip3p modifications⁴³.

to take into account small differences in the electrostatic interaction energy calculated with the Ewald method rather than via truncated Coulomb interactions (Tip3pEw). For each water model, the heat of immersion was computed for increasing values of the charges of terminal OH groups and with the Lennard–Jones parameter reported in Table 2.2. The heat of immersion E_{im} is calculated from the difference between the average total energy of the system in contact with water $E_{interface}$ and the average total energies of the two separate components, namely the hydroxylated dry silica surface E_{silica} and the bulk water E_{water} , obtained in MD simulations at constant temperature (300 K) and pressure (1 atm):

$$E_{im} = (-E_{interface} + E_{silica} + E_{water})/A, \quad (2.51)$$

with $A=1.85 \times 10^{-17} \text{ m}^2$ being the total area of the top and bottom surfaces of the silica slab. The computed values of E_{im} (Table 2.3) present a remarkable sensitivity to the surface charges and to the chosen water model, suggesting that absolute heat of immersion values shall be interpreted with great care when performed with even slightly different simulation setups. Interestingly, the strength of the water–surface interaction decreases with increasing charge values of the terminal hydroxyl groups, suggesting that electrostatic repulsion dominates the observed behavior. A precise comparison with experimental values (around 150–250 mJ/m²²¹⁰) is difficult because of the native negative surface charges present in the

experiments and the non-trivial contribution of counterions in MD calculations including charged surface models. However, the values computed with the largest charges, as defined in Figure 2.4, seem to be acceptable, given the expected increase of the E_{im} values arising from the presence of net charged sites.

On a final note, it should be remarked that the silica surface can only interact with water and biomolecules via non-bonded van der Waals and Coulomb forces. As a direct consequence the silica surface must represent a steady state in terms of silanol group formation or other possible chemical reactions. Moreover, this force field was intentionally designed to describe only interfacial interactions between silica, water or biomolecules. It defines no parameter describing intrinsic silica interactions. Thus, without the choice of an appropriate bulk silica force field, all atoms of the silica slab have to be constrained to their origin. In order to represent a realistic silica surface it was decided to use a flexible SiO_2 model instead of constraining positions of all silicon and oxygen atoms not bound to hydrogen to their origin.

A BULK SILICA FORCE FIELD

As a starting point the potential developed by Cole et al.⁵⁰ was then chosen. Its underlying point charge definition for silicon, oxygen and hydrogen atoms is in good agreement with our newly developed potential to describe interfacial systems of water and silica. Moreover, the description of van der Waals interactions between atoms was taken over from this work. Trial simulations revealed, however, some concerns which leads to a rejection of the potential. A very short cutoff of 5.5 Å for the steric repulsion and polarization of atoms in Equation (2.43) has been used, previously introduced by Trioni et al.²¹⁹, in order to save compu-

	MS-Q ^{230,231}	Butenuth et al. ⁴³
Si	1.318	1.4
O	-0.659	-0.8

Table 2.4: Comparison between partial charges obtained for bulk silica with the MS-Q^{66,230,231} and Butenuth et al.⁴³ force fields.

tational time. Unfortunately, several simulations in the course of this work had shown that the resulting density depends crucially on this parameter when used in combination with a barostat, which must not happen.

It turned out that a promising candidate to simulate bulk silica was the MS-Q potential proposed by Demiralp et al.⁶⁶. It consists of a Morse-Stretch potential and employs a charge equilibration method. Point charges of atoms are reassigned every simulation step therein. Its functional form is given by

$$V(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{2} \sum_{i \neq j} \left[D_{ij,0} (e^{-2\alpha_{ij}(r-r_0)} - 2e^{-\alpha_{ij}(r-r_0)}) - \frac{q_i q_j}{r_{ij}} \right], \text{ with } r = |\mathbf{r}_i - \mathbf{r}_j|. \quad (2.52)$$

A charge equilibration has advantages over traditional fixed partial atomic charges, e.g. polarizability effects are included. However, this is a bottleneck in terms of computational efficiency. Owing to the similarity of MS-Q charges obtained for silica^{230,231} to fixed partial charges used in Butenuth et al.⁴³, it was decided to replace the charge equilibration method of the MS-Q force field with fixed partial atomic charges. A comparison between charges obtained for bulk silica with the MS-Q and Butenuth et al.⁴³ force fields is given in table 2.4.

Due to the use of fixed partial charges, slight adaptations to the Morse-Stretch potential had to be made in order to give realistic structural data; In particular the Si-Si interaction

was reparameterized. The remaining parameters are given in table 2.5 to complete the force field definition used in this work.

An amorphous bulk SiO₂ model was then generated starting from a randomly distributed system of silicon and oxygen atoms in the ratio 1:2 and a density corresponding to 2.2 g/cm³. First the system was equilibrated in the NVE ensemble to avoid steric clashes between atoms. Afterwards it was heated up to 6000 K in 150 ps, equilibrated at this temperature for additional 50 ps and then cooled down to 2000 K in 380 ps, using appropriate thermostats. The system was then cooled down to 300 K within 190 ps in the NPT ensemble and eventually equilibrated for additional 50 ps at this temperature. The radial and angular distribution functions in Figure 2.5 are computed from the last 50 ps of the trajectory.

To prove the applicability of the modified Demiralp potential for amorphous silica the resulting bulk SiO₂ structural properties are compared to data available in the literature^{208,79,90,150,228,59,226}. A comparison is outlined in table 2.6.

The obtained RDF features, reported in 2.6, are in remarkably good agreement with available literature values. Notably, the Si-O-Si and O-Si-O angular distribution are also in agreement with previous data (Si-O-Si angle of about 109° with a full width at half maximum in the range of 12° to 21°; O-Si-O angle in the range between 142° and 153° with a full width at half maximum between 21.5° and 35.7°^{150,228,59,226}), although the potential does not include explicitly any three-body terms.

	$D_{ij,0}$ [kcal/mol]	α_{ij} [Å ⁻¹]	$R_{ij,0}$ [Å]
O-O	0.02325	1.37587	3.7835
Si-O	1.99443	2.72548	1.6148
Si-Si	0.01282	1.71743	4.0603

Table 2.5: Parameter of the modified Demiralp potential

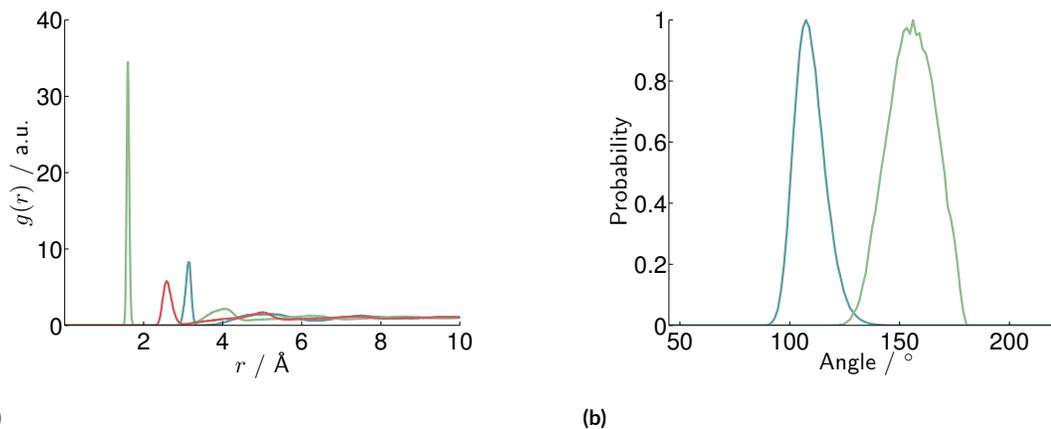


Figure 2.5: (a) Radial distribution functions of Si-Si (blue), Si-O (green) and O-O (red) atoms in bulk SiO₂ obtained with the modified Demiralp potential. (b) O-Si-O (blue) and Si-O-Si (green) angle distributions in bulk SiO₂ obtained with the modified Demiralp potential.

2.8 FREE ENERGY CALCULATIONS

A link between experiments and simulations can be established via calculations of free energy estimates obtained typically only with advanced molecular dynamics simulation techniques. Thermodynamic free energies represents a physical quantity that reflects the amount of work a system can perform or is necessary to change the state of the system from one to another. Thus, chemical or thermal processes can be quantified using this fundamental quantity, finding applications in all fields of engineering and natural sciences. Moreover, it determines phase equilibria, such as melting and boiling points and the pressure of saturated vapors, and chemical equilibria such as solubilities, binding or dissociation constants and conformational changes¹⁷.

As a general rule, the free energy is the internal energy of system minus the amount of energy that cannot be used to perform work. Unusable energy is given by the entropy of a system multiplied by its temperature. Two different formulations exist for systems un-

	Mod.	MD	Exp.
ρ [g/cm ³]	2.215	2.14 ⁷⁹ , 2.23 ²⁰⁸	2.20 ¹⁵⁰
Si-Si RDF 1st peak [Å]	3.15	3.155 ²²⁸	3.077 ¹⁵⁰ , 3.12 ¹⁵⁰
O-O RDF 1st peak [Å]	2.58	2.590 ²²⁸	2.626 ⁹⁰ 2.65 ¹⁵⁰
Si-O RDF 1st peak [Å]	1.604	1.595 ²²⁸	1.608 ⁹⁰ 1.620 ¹⁵⁰
Si-O RDF 2nd peak [Å]	4.06	4.12 ²²⁸	4.15 ¹⁵⁰
Si-O-Si angle [°]	155 (29.5)	150 (21.5) ⁷⁹ , 152 (35.7) ²²⁸	144 (38) ¹⁵⁰ , 153 ⁵⁹
O-Si-O angle [°]	107.45 (15.6)	109.2 (20.9) ⁷⁹ , 108.3 (12.8) ²²⁸	109.4 ¹⁵⁰ , 109.5 ¹⁵⁰

Table 2.6: Comparison between structural results obtained with the modified MS-Q potential, previous results from other MD simulations and experimental findings.

der different conditions. The Helmholtz free energy $F(T, V)$ measures the work obtainable from a closed thermodynamic system at constant temperature and volume, whilst the Gibbs free energy $G(T, p)$ measures the work obtainable from a thermodynamic system at a constant temperature and pressure. They can be expressed as follows

$$F(T, V) = U - TS \quad (2.53)$$

$$G(T, p) = U + pV - TS. \quad (2.54)$$

Unfortunately, in simulations it is practically impossible to calculate U and TS separately, although the potential energy U is directly available from simulation. This is made all the more difficult by the fact that, the entropic term $-TS$ is given by an integral over the *entire* phase space of the system. It makes no difference whether a microstate is energetically favorable, lies behind an energetic barrier or is energetically unfavorable. Energy

contributions from unfavorable microstates are as important to the integral as the contribution from favorable to obtain a realistic estimate of the free energy. However, from regular molecular dynamics simulations those informations are not feasible within reasonable time scales.

In statistical mechanics the phase space $\Gamma = \{\mathbf{r}; \mathbf{p}\}$ is spanned over all possible microstates that a system can theoretically occupy (cf. Chapter 2.1). Since most of the simulations in this work employ a N ose-Hoover thermostat and are thus performed in the canonical N, V, T ensemble the corresponding probability to find the system in a distinct microstate i can be formulated to

$$P(\mathbf{r}_i, \mathbf{p}_i) = \frac{1}{Z} \exp(-\beta \mathcal{H}(\mathbf{r}_i, \mathbf{p}_i)). \quad (2.55)$$

It should be noted that the index i stands here for a microstate and should not be confused with the particle index. Here $\mathcal{H} = T + V$ is the classical Hamiltonian of the system, $\beta = 1/(k_B T)$ is the inverse thermal energy and

$$Z = \iint d\mathbf{r}d\mathbf{p} \exp(-\beta \mathcal{H}(\mathbf{r}, \mathbf{p})) \quad (2.56)$$

is the canonical partition function of the system which normalizes the probabilities. The entropy S of a canonical ensemble is defined as

$$S = \frac{1}{T} U + k_B \ln(Z). \quad (2.57)$$

Rearranging and comparison with the formulation of the Helmholtz free energy (Eq. 2.53)

yields an important relation between statistical mechanics and the thermodynamic free energy:

$$F = -k_B T \ln(Z) . \quad (2.58)$$

Incidentally, in case of a N, p, T ensemble this would yield the Gibbs free energy as it is related to the enthalpy $H = U + pV$.

This strongly illustrates the difficulty of extrapolating free energy estimates from regular molecular dynamics simulations. The potential energy (or energy landscape) of a molecule is a function of all coordinates and has one global minimum. However, depending on the molecule size, it can have a very complex structure with multiple local minima, separated by barriers of various heights. If the system is *ergodic*, it visits in an equilibrium state at temperature T all regions of the configurational space that have an energy within a range of the order of $k_B T$ with respect to the global minimum. Whether all relevant structures in the configurational space of arbitrarily large molecules will be accessed, in order to obtain a robust free energy estimate using Equation (2.55), is thus not necessarily guaranteed in the course of a regular molecular dynamics simulation.

Differences in free energies between two subsystems, which are locally *ergodic*, are often more desirable than absolute values. For this purpose the potential of mean force, which is a free energy with respect to certain defined variables, can be consulted. Its relation to the free energy is explained in the following. Consider two subsystems which have long enough life times and locally ergodic behavior; A reactant state R and a product state P. Those could be, for example, a molecule binding to a surface. The reactant state would be the molecule dissolved in, e.g., a solution and the product state the protein adsorbed on a surface. Each of the states R and P is thus considered as separate thermodynamic states with

Helmholtz free energies of

$$F^{\text{R}} = -k_{\text{B}}T \ln(Z^{\text{R}}) \quad (2.59)$$

$$F^{\text{P}} = -k_{\text{B}}T \ln(Z^{\text{P}}) \quad (2.60)$$

and the partition functions

$$Z^{\text{R}} = \iint_{\text{R}} d\mathbf{r}d\mathbf{p} \exp(-\beta\mathcal{H}(\mathbf{r}, \mathbf{p})) \quad (2.61)$$

$$Z^{\text{P}} = \iint_{\text{P}} d\mathbf{r}d\mathbf{p} \exp(-\beta\mathcal{H}(\mathbf{r}, \mathbf{p})). \quad (2.62)$$

It can immediately be seen that although $Z = Z^{\text{R}} + Z^{\text{P}}$, the total Helmholtz free energy $F \neq F^{\text{R}} + F^{\text{P}}$. Instead it is

$$F = \frac{Z^{\text{R}}}{Z} F^{\text{R}} + \frac{Z^{\text{P}}}{Z} F^{\text{P}} + k_{\text{B}}T \left[\frac{Z^{\text{R}}}{Z} \ln \left(\frac{Z^{\text{R}}}{Z} \right) + \frac{Z^{\text{P}}}{Z} \ln \left(\frac{Z^{\text{P}}}{Z} \right) \right]. \quad (2.63)$$

The latter term stems due to the mixing entropy resulting from the distribution of the system over two states.

By defining a reaction coordinate $\xi(\mathbf{r})$ as a function of particle coordinates which connects the R and P region and separating it from the integration in the partition function Z ,

$$Z = \int d\xi \iint d\mathbf{r}d\mathbf{p} \exp(-\beta\mathcal{H}(\mathbf{r}, \mathbf{p})) \delta(\xi(\mathbf{r}) - \xi) \quad (2.64)$$

is obtained. Constructing now a potential of mean force

$$V_{\text{PMF}} \stackrel{\text{def.}}{=} -k_{\text{B}}T \ln \left(\iint d\mathbf{r}d\mathbf{p} \exp(-\beta\mathcal{H}(\mathbf{r}, \mathbf{p}))\delta(\xi(\mathbf{r}) - \xi) \right) \quad (2.65)$$

such that

$$Z = \int d\xi \exp(-\beta V_{\text{PMF}}(\xi)) \quad (2.66)$$

the Helmholtz free energy can be rewritten to

$$F = -k_{\text{B}}T \ln \left(\int d\xi \exp(-\beta V_{\text{PMF}}(\xi)) \right). \quad (2.67)$$

Once the potential of mean force is known, the Helmholtz free energy of a thermodynamic state can be computed from Equation (2.67) by integrating over the relevant part of the reaction coordinate ξ . The difference in free energy between two states can thus be expressed by

$$\Delta F = -k_{\text{B}}T \ln \left(\frac{Z^{\text{R}}}{Z^{\text{P}}} \right) = -k_{\text{B}}T \ln \left(\frac{\int_{\text{R}} d\xi \exp(-\beta V_{\text{PMF}}(\xi))}{\int_{\text{P}} d\xi \exp(-\beta V_{\text{PMF}}(\xi))} \right) \quad (2.68)$$

using Equation (2.66). The potential of mean force is thus a free energy for the system excluding the reaction coordinates as degrees of freedom¹⁷.

However, it is often not straightforward to evaluate such integrals from simulations. Tractable cases are homogeneous distributions (e.g. ideal gases) and distributions that can be approximated by Gaussian distributions (e.g. harmonic potentials). Free energies of more complex systems can only be obtained by employing special algorithms to increase the sampling along a reaction coordinate. This is mainly done by evaluating derivatives of V_{PMF}

from ensemble averages of multiple simulations. V_{PMF} can then be computed by integrating over those results and will be determined up to an additive constant. Potentials of mean force can be derived from, e.g., thermodynamic integration or perturbation, umbrella sampling, particle insertion or enforcing the system to move from one part of configurational space to another. A few of those algorithms, the most known, are introduced briefly in the following.

2.8.1 UMBRELLA SAMPLING

Pioneered by Torrie & Valleau²¹⁷ this methods adds restrains to the potential energy $V(\mathbf{r})$ by employing usually harmonic umbrella potentials

$$V^{\text{umb}}(\xi) = \frac{1}{2}k(\xi(\mathbf{r}) - \xi_i)^2 \quad (2.69)$$

around coordinates of given configurations. The resulting canonical umbrella distribution $w(\xi)$ will be given by

$$w(\xi) \propto \int d\mathbf{r} \exp(-\beta V(\mathbf{r}) - \beta V^{\text{umb}}(\xi)). \quad (2.70)$$

Consequently, the potential of mean force can be estimated up to an additional constant V_{PMF}^0 to

$$V_{\text{PMF}} = V_{\text{PMF}}^0 - k_{\text{B}}T \ln(w(\xi)) - V^{\text{umb}}(\mathbf{r}). \quad (2.71)$$

By keeping track of the distribution of ξ_i values and correcting the bias caused by the umbrella potentials, the potential of mean force can be reconstructed modulo an immaterial

constant V_{PMF}^0 . This additional constant is in most cases not of interest, because usually one is interested in the form of the free energy profile (or surface) and energy differences such as free energy barriers between two different states and not in the absolute value. Another approach to obtain the free energy estimates from umbrella sampling is by using the weighted histogram analysis method (WHAM)^{123,10}. It involves an additional iterative determination of the additional constants V_{PMF}^0 until the total free energy profile (or potential of mean force) assumes a continuous form.

The advantage over the more simpler thermodynamic integration method is that by this method the spacing of sampling points ξ_i can be chosen larger since the harmonic potentials allow also sampling in the vicinity of the constrained coordinates. Harmonic potentials should, however, possess sufficient overlap to enable a correct sampling along the reaction pathway. This constitutes the main drawback of umbrella sampling—it does not reproduce edges in the free energy profiles very well. Due to better sampling of the molecule in the middle of the umbrella potentials, this leads to many samples on the left or right side of the edge, but not on the edge itself. It is also depending on the path meaning that a molecule can first approach the surface from the solvent or the other way round and both may result in a hysteresis in the free energy profiles.

2.8.2 ADAPTIVE BIAS FORCE METHODS

A popular algorithm to perform adaptive bias force simulations is METADYNAMICS. Many simulations in this work have been performed using this algorithm to augment regular molecular dynamics simulations. This provides a robust method to obtain free energy landscapes by imposing a time dependent bias potential that acts on a selected number of

collective variables (or reactive coordinates). In the following a short summary of the review on METADYNAMICS by Barducci et al.⁷ should be given. An external history-dependent bias potential which is a function of the collective variables is added to the Hamiltonian of the system. METADYNAMICS employs for this reason Gaussians deposited along the system trajectory in the collective variable space in order to discourage the system from revisiting configurations that have already been sampled. Here a version is chosen which acts directly on the microscopic coordinates of the system, though, other versions are also available^{125,110}.

The collective variable $\xi(\mathbf{r}) = (\xi_1(\mathbf{r}), \dots, \xi_d(\mathbf{r}))$ is a set of d function of the microscopic coordinates \mathbf{r} of the system. The time-dependent bias potential is defined by

$$V^{\text{meta}}(\xi, t) = \int_0^t dt' \omega \exp \left(- \sum_{i=1}^d \frac{(\xi_i(\mathbf{r}) - \xi_i(\mathbf{r}(t')))^2}{2\sigma_i^2} \right). \quad (2.72)$$

Basically, Gaussians are deposited with a certain energy rate ω and width σ_i on the collective variable phase space. The energy rate is constant and expressed in terms of a Gaussian height W and a deposition stride τ :

$$\omega = \frac{W}{\tau} \quad (2.73)$$

Using this approach, local free energy minima in which the molecule or particle under consideration could be trapped can easily be overcome by simply adding as much Gaussians as it is needed to compensate local minima of the collective variable. Regions in which the energy landscape is deeper more Gaussians are necessary to compensate the free energy and after a certain time the bias potential compensates eventually completely the free energy landscape (modulo an irrelevant additive constant) of the collective variable which has been

chosen:

$$V^{\text{meta}}(\xi, t \rightarrow \infty) = -V_{\text{PMF}}(\xi) + V_{\text{PMF}}^0. \quad (2.74)$$

However, this can be primarily seen only as a first approximation since the bias potential will oscillate around a certain value as more and more Gaussians are added. This can be avoided by choosing a time-dependent Gaussian height as done in well-tempered METADYNAMICS⁸:

$$W = \omega \tau \exp\left(\frac{V^{\text{meta}}(\xi, t)}{k_B \Delta T}\right). \quad (2.75)$$

In contrast to standard METADYNAMICS the bias deposition rate decreases as $1/t$ with this formulation and the dynamics become thus progressively closer to thermodynamic equilibrium as the simulation proceeds. The bias potential will not fully compensate the underlying potential energy, but it converges to

$$V^{\text{meta}}(\xi, t \rightarrow \infty) = -\frac{\Delta T}{T + \Delta T} V_{\text{PMF}}(\xi) + V_{\text{PMF}}^0 \quad (2.76)$$

And in the long time limit the collective variables probability distributions become

$$w(\xi) \propto \exp\left(\frac{V_{\text{PMF}}(\xi)}{k_B(T + \Delta T)}\right). \quad (2.77)$$

Obviously this represents sampling at a higher temperature $T + \Delta T$. Thus, in the limits of $\Delta T = 0$ or ∞ standard molecular dynamics or standard METADYNAMICS are recovered, respectively. A value in between those limits will therefore regulate the free energy surface exploration and avoids overfilling in a way that low values of ΔT will lead to a quicker convergence of V^{meta} whereas higher values improve the sampling of the collective coordinate.

An example of this method is given in the following for the estimation of the barrier height when a water or $\text{Si}(\text{OH})_4$ molecule penetrate a phosphatidylcholine lipid bilayer. Phosphatidylcholine lipids (POPC) are the main constituent of eukaryotic cell membranes. A graphical illustration of the molecular structure of the system used in the METADYNAMICS simulation is shown in Figure 2.6.

In Figure 2.7a and 2.7b are the temporal evolutions of the free energies of a TIP3P water and a $\text{Si}(\text{OH})_4$ molecule upon penetrating a phospholipid bilayer shown. Furthermore, estimates of the barrier height in order to cross the phospholipid bilayer are given in Figure 2.8 for both molecules. The temporal evolution of the barrier height will fluctuate as the free energy profiles are evolving. An estimate for the barrier height is calculated in Figure 2.8 for every nanosecond of the evolving well-tempered METADYNAMICS simulation.

The energy necessary for a water molecule to cross the barrier is evaluated by averaging the free energy plateaus of either if it is located in bulk water ($-45 \text{ \AA} < z_{\text{com}} < -30 \text{ \AA}$ and $30 \text{ \AA} < z_{\text{com}} < 45 \text{ \AA}$) or in the bulk of the lipid bilayer ($-5 \text{ \AA} < z_{\text{com}} < 5 \text{ \AA}$).

$$\Delta F = \langle V_{\text{PMF}}(z_{\text{com}}) \rangle_{\text{bulk}} - \langle V_{\text{PMF}}(z_{\text{com}}) \rangle_{\text{POPC}} \quad (2.78)$$

Brackets are symbolics for taking the average of the potential of mean force in the respective area. Simultaneously this allows to check the convergence of the free energy profile. By calculating a barrier height estimate for every nanosecond of the trajectory a sufficient convergence is reached for the water molecule after approximately 100 ns. The free energy difference of the water molecule will fluctuate around an estimated barrier height of 6.6 ± 0.5 kcal/mol. A value that has been found also by other studies of a similar system (6.2 kcal/mol for water permeating through a DPPC bilayer¹⁴⁴). However, in the case of

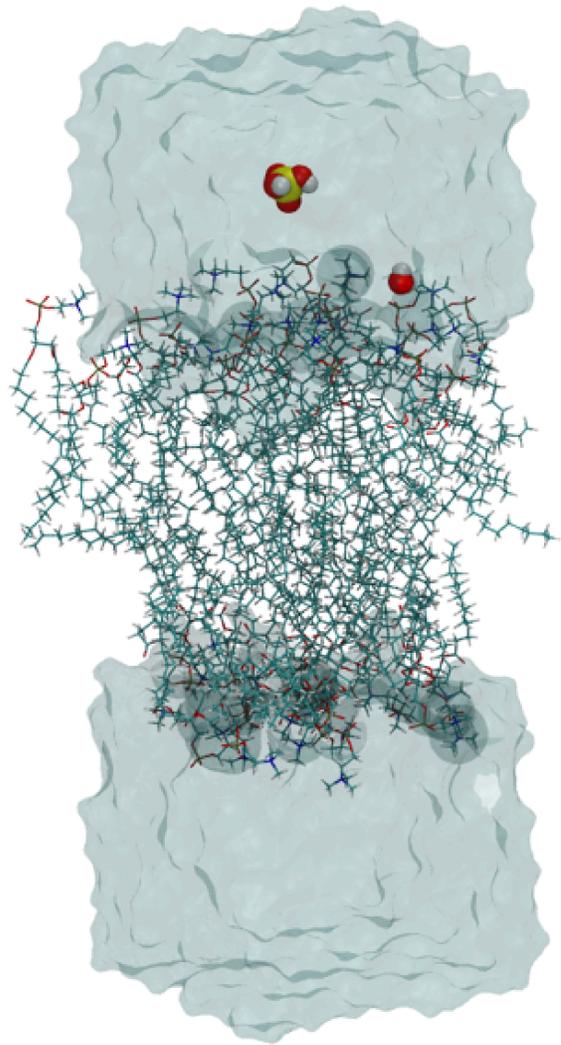


Figure 2.6: Exemplary system used to introduce the functioning of the Metadynamics algorithm. Free energy differences between a water or Si(OH)_4 molecule penetrating a lipid (POPC) bilayer is evaluated using this system.

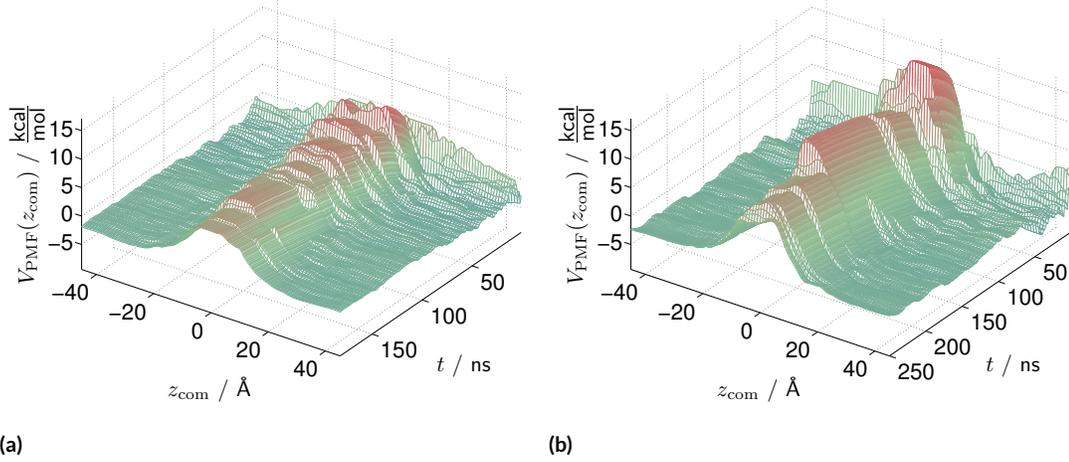


Figure 2.7: Temporal evolution of the potential of mean force of (a) a water (Tip3p) and (b) a Si(OH)₄ molecule penetrating a lipid (POPC) bilayer (cf. Figure 2.6) using the Metadynamics algorithm.

the Si(OH)₄ molecule convergence is hardly reached and the estimated barrier height fluctuates around a value of 10 ± 2 kcal/mol.

The introduction of a history-dependent potential alters the probability distributions of other degrees of freedom ζ and those will be distorted if calculated directly from the simulation trajectory. Since the magnitude of the bias potential continuously increases with time in the well-tempered METADYNAMICS simulation, simple reweighting to recover the unbiased distributions of external collective variables is not possible in a straightforward manner. A suitable reweighting scheme for this situation has been devised in²⁵. The unbiased distribution $w(\chi)$ can be calculated numerically from the temporal evolution of the biased histogram $N_t(\xi, \chi)$ according to

$$w(\chi) = \frac{\sum_s e^{\beta V^{\text{meta}}(\xi, t)} N_t(\xi, \chi)}{\sum_{\xi, \chi} e^{\beta V^{\text{meta}}(\xi, t)} N_t(\xi, \chi)}. \quad (2.79)$$

This can be used in order to calculate probability distributions of, for example, externally

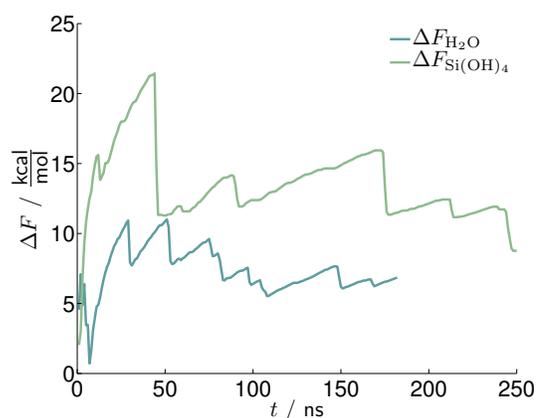


Figure 2.8: Convergence of the free energy barrier height of a water (Tip3p) and $\text{Si}(\text{OH})_4$ molecule penetrating a lipid (POPC) bilayer. The barrier height is calculated from the free energy profiles shown in Figure 2.7a and 2.7b using Eq. (2.78).

calculated collective variables which are not included by default in the METADYNAMICS package.

2.8.3 REPLICA EXCHANGE METHODS

Non-ergodicity is a problem of simulations when for example proteins and nucleic acids in solution should be examined. Such systems needs often macroscopic (say, seconds) to fold into a specific conformation. This time span is not feasible with regular molecular dynamics. Aside from biomolecules, polymers, e.g., copolymers exhibit a similar behavior. Melts of them need long rearrangement times to settle to a structure with minimum free energy. As a third example glasses below glass transition temperature could be named: They will freeze into a subset of possible states, probably not the energetically lowest. All of the aforementioned systems are trapped in a limited set of configurations (or conformations) that is a subset of the complete canonical distribution and will be not ergodic when simulated with a standard molecular dynamics simulation.

A breakthrough in the efficient generation of a representative ensemble of configurations in cases where equilibration is slow was achieved by the introduction of replica exchange methods (REM). With this approach it is sampled over all degrees of freedom, in contrast to the aforementioned approaches in which only the sampling of a distinct set of reaction coordinates ξ is enhanced. Often replica exchange is done in molecular dynamics by exchanging the temperatures of two neighboring replicas (TREM)²⁰⁵ by employing a distinct exchange criteria. An initial replica say at 300 K can thus encompass the whole set of temperatures and is able to easily overcome high energy barriers in a high temperature replica; Each system now rapidly visits a range of temperatures. The result is a much faster relaxation than a single system would have at a low temperature, without distorting the canonical distribution. In the case that the adsorption behavior of proteins or peptides should be evaluated replica exchange methods can facilitate other methods to calculate potentials of mean force. If, for example, the distance of the peptide (or protein) to the surface is considered as a reaction coordinate a potential of mean force can be obtained for the adsorption path with one of the aforementioned methods. However, apart from barriers in the reaction coordinate other barriers not considered explicitly can influence the peptides conformation and could have influence on the adsorption path. Thus, in the worst case, important conformations are excluded in the calculation of the potential of mean force¹⁶¹. Replica exchange methods increase the sampling of the remaining $3N - f$ degrees of freedom not explicitly treated in the method of choice for calculating the potential of mean force. Thus, when used in combination with other methods to calculate free energies, e.g., METADYNAMICS a faster and more reliable free energy estimate can be obtained^{42,190}.

In the temperature replica exchange method it is assumed that the probability of a con-

figuration \mathbf{r} in the m th replica obeys the Boltzmann distribution

$$w_m(\mathbf{r}_m) = \frac{1}{Z_m} \exp(-\beta_m V(\mathbf{r}_m)) \quad , \quad (2.80)$$

with $\beta_m = 1/k_B T_m$, the Hamiltonian of the system $V(\mathbf{r})$ and Z_m as the partition function.

It is assumed that the joint probability distribution of the extended system with M replicas w is represented by the multiplication of the probability of each replica

$$w = \prod_i^M w_i(\mathbf{r}_i) . \quad (2.81)$$

The transition probability that the configuration \mathbf{r}_m of the m th replica is exchanged with configuration \mathbf{r}_n of replica n is written as $W(\mathbf{r}_m, \beta_m; \mathbf{r}_n, \beta_n)$ and for the reverse process it is written as $W(\mathbf{r}_n, \beta_n; \mathbf{r}_m, \beta_m)$. The detailed balance condition for the extended system to reach the extended Boltzmann equilibrium (Eq. (2.81)) should hold

$$\begin{aligned} w[(\dots; \mathbf{r}_m, \beta_m; \mathbf{r}_n, \beta_n; \dots)] W(\mathbf{r}_m, \beta_m; \mathbf{r}_n, \beta_n) = \\ w[(\dots; \mathbf{r}_n, \beta_n; \mathbf{r}_m, \beta_m; \dots)] W(\mathbf{r}_n, \beta_n; \mathbf{r}_m, \beta_m) \end{aligned} \quad (2.82)$$

which directly leads to

$$\frac{W(\mathbf{r}_m, \beta_m; \mathbf{r}_n, \beta_n)}{W(\mathbf{r}_n, \beta_n; \mathbf{r}_m, \beta_m)} = \frac{w[(\dots; \mathbf{r}_m, \beta_m; \mathbf{r}_n, \beta_n; \dots)]}{w[(\dots; \mathbf{r}_n, \beta_n; \mathbf{r}_m, \beta_m; \dots)]} . \quad (2.83)$$

All terms from replicas which were not exchanged cancel out in the extended Boltzmann

equilibrium (Eq. (2.81)), leading to

$$\begin{aligned} \frac{W(\mathbf{r}_m, \beta_m; \mathbf{r}_n, \beta_n)}{W(\mathbf{r}_n, \beta_m; \mathbf{r}_m, \beta_n)} &= \exp[(\beta_m - \beta_n)(V(\mathbf{r}_n) - V(\mathbf{r}_m))] \\ &= \exp(-\Delta). \end{aligned} \quad (2.84)$$

Using the Metropolis acceptance criteria the probability of a temperature exchange between replica m and n is

$$W(\mathbf{r}_m, \beta_m; \mathbf{r}_n, \beta_n) = \begin{cases} 1 & \text{for } \Delta \leq 0 \\ \exp(-\Delta) & \text{for } \Delta > 0 \end{cases}. \quad (2.85)$$

Consequently, the backward transition probability is

$$W(\mathbf{r}_n, \beta_m; \mathbf{r}_m, \beta_n) = \begin{cases} \exp(\Delta) & \text{for } \Delta < 0 \\ 1 & \text{for } \Delta \geq 0 \end{cases} \quad (2.86)$$

with values for W ranging in the interval $[0, 1]$ for both cases. Using this acceptance criterion it can easily be seen that forward and backward transition fulfill Equation (2.84).

Another more general approach is to exchange the Hamiltonians instead of thermostats. Fukunishi et al.⁸⁴ revised in his work replica exchange methods and showed that the temperature replica exchange can be seen as a special case of the more general Hamiltonian replica exchange. Recently an interesting alternative Hamiltonian replica exchange method has been proposed. In this methods instead of exchanging Hamiltonians different bias potentials defined for each replica are exchanged (BEM)¹⁷⁰. It already found application in the

study of Marinelli et al.¹⁴¹ in which a kinetic model of the Trp-Cage folding was derived. Thus, by defining enough bias potentials for varying reaction coordinates in each replica, e.g., via METADYNAMICS, and allowing an exchange of the bias potentials between replicas, the problem of indirect barriers influencing the potential of mean force on the desired reaction coordinate can be circumvented. Another alternative formulation of a Hamiltonian replica exchange method which is very suitable for studying conformational changes of peptides or small proteins is briefly introduced in the following.

REPLICA EXCHANGE WITH SOLUTE TEMPERING (REST)

The Replica Exchange with Solute Tempering (REST) method was originally published by Liu et al.¹³² and optimized by Wang et al.²³². It depends on the Hamiltonian replica exchange method originally published by Fukunishi et al.⁸⁴ which is defined by exchanging the Hamiltonians of two neighboring replicas instead of their temperature. In this case the Boltzmann distribution is defined as

$$w_m(\mathbf{r}_m) = \frac{1}{Z_m} \exp[-\beta V_m(\mathbf{r}_m)] \quad (2.87)$$

The probability of an exchange between the hamiltonian of the m th replica and the n th replica is calculated in the same manner as previously for the temperature replica exchange

method and is therefore

$$W(\mathbf{r}_m, V_m(\mathbf{r}_m); \mathbf{r}_n, V_n(\mathbf{r}_n)) = \begin{cases} 1 & \text{for } \Delta \leq 0 \\ \exp(-\Delta) & \text{for } \Delta > 0 \end{cases}$$

with $\Delta = \beta[(V_m(\mathbf{r}_n) + V_n(\mathbf{r}_m)) - (V_m(\mathbf{r}_m) + V_n(\mathbf{r}_n))]$ (2.88)

The temperature REM can be seen as a special case of the Hamiltonian REM when the inverse temperatures β_m are scaled as $\beta_m = s_m \beta$. In this special case both methods give the same Boltzmann factor $\exp[-\beta s_m E(X)]$ in the Boltzmann distribution of Equation (2.80) and (2.87) meaning that either the temperature could be scaled for a temperature replica exchange or the Hamiltonians for a Hamiltonian replica exchange, both giving identical answers.

In principle, the REST method divides the total energy into three terms and scales them differently according to

$$V_m(\mathbf{r}) = \frac{\beta_m}{\beta_0} V^{PP}(\mathbf{r}) + \sqrt{\frac{\beta_m}{\beta_0}} V^{PW}(\mathbf{r}) + V^{WW}(\mathbf{r}) \quad , \quad (2.89)$$

with β_0 being the inverse temperature in the lowest replica, V^{PP} being the intramolecular energy of the protein, V^{PW} being the interaction energy of protein and water molecules and V^{WW} being the energy of only the interactions between water molecules. This rescaling of Hamiltonians is equivalent to only heating up the temperature of the solute and this is where the name replica exchange with solute tempering (REST) stems from. Effectively the interaction energy for $(\beta_m/\beta_0)E_{PP}$ is achieved by rescaling the bonded interaction terms, the Lennard–Jones ϵ parameter and the charges of the solute atoms by β_m/β_0 , β_m/β_0 and

$\sqrt{\beta_m/\beta_0}$, respectively, and for $(\sqrt{\beta_m/\beta_0})E_{pw}$ by rescaling the nonbonded Lennard–Jones ϵ parameter with $\sqrt{\beta_m/\beta_0}$. The specific heat increase of each replica is thus only a fraction of the case that all interactions would have been scaled. This yields a better overlap of potential energies of each replica allowing a reduced number of replicas necessary and reducing therefore heavily the computational effort needed. Combining Equation (2.88) and 2.89 leads to an expression for the acceptance probability factor

$$\Delta = (\beta_m - \beta_n) \left[(V^{pp}(\mathbf{r}_n)) - V^{pp}(\mathbf{r}_m) + \frac{\sqrt{\beta_0}}{\sqrt{\beta_m} + \sqrt{\beta_n}} (V^{pw}(\mathbf{r}_n) - V^{pw}(\mathbf{r}_m)) \right]. \quad (2.90)$$

One could think of including this expression as a swapping criteria into the simulation code, however, contributions of only fragments of long-range interactions are not trivial to estimate, making an implementation of this equation into molecular dynamics anything but ordinary. Fortunately, it is not necessary to include such energy contributions. Due to Equation (2.88) is still true even for modified Hamiltonians from replica exchange with solute tempering, this exchange criteria will give an identical answer as Equation (2.90) but avoids the necessity of calculating fragments of long-range energy contributions. Equation (2.90) gives us only an insight why the REST method performs more efficiently in terms of computational effort over standard temperature replica exchange methods but is not used in terms of calculating an exchange probability.

In order to demonstrate the superiority of replica exchange with solute tempering over standard temperature replica exchange for the sampling of peptide conformations, respective resulting energy probabilities and temperature positions of replicas for each method are shown in Figure 2.9 for a 10 ns simulation of a small dissolved peptide consisting of four

aminoacids (GCRL, in the one letter amino acid code). Potential energy distributions of different temperatures are much more spread in the case of temperature replica exchange as compared to replica exchange with solute tempering. In Figure 2.9a it can be seen that exchanges between replicas at different temperatures occur thus only very rare, if ever. Whereas in Figure 2.9c an exchange probability of 26 % is achieved for the same temperature range using replica exchange with solute tempering. Hence, in order to achieve the same overlap of potential energies as in replica exchange with solute tempering a significant higher number of replicas has to be used increasing heavily the computational effort necessary (cf. Figure 2.9b and Figure 2.9d).

However, in replica exchange with solute tempering replicas with temperatures higher than the unscaled ground replica cannot be used to extract any thermodynamic information out of them due to the unphysical scaling of potential energies, unlike in regular temperature replica exchange where higher temperature replicas can be used for further evaluation of, e.g., the thermodynamic properties of the desired molecule at higher temperatures (if the force field allows this and is parameterized for higher temperatures). In replica exchange with solute tempering elevated temperature replicas are only used to increase sampling. Furthermore only peptide-peptide and partially peptide-water interactions are scaled in replica exchange with solute tempering which could result in neglecting the sampling of important degrees of freedom of the solvent which is not the case when standard temperature replica exchange is applied. It should be noted that the aforementioned replica exchange with solute tempering simulation was supplemented with a bias potential stemming from METADYNAMICS. This should only have a negligible influence on the exchange behavior in Figure 2.9 and is discussed in more detail in the next chapter.

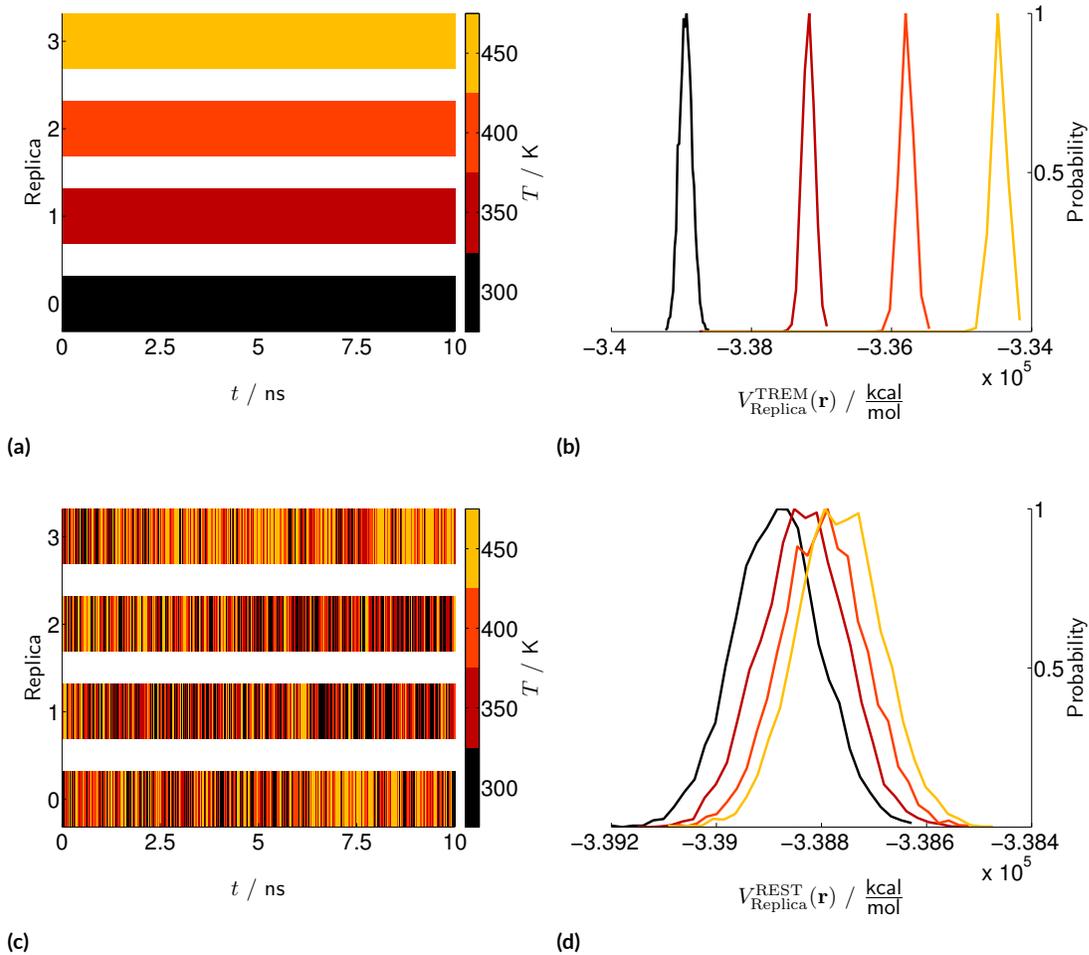


Figure 2.9: Comparison of results from a 10 ns regular temperature replica exchange (a,b) and replica exchange with solute tempering (c,d) simulations. In (a,c) is the respective temperature position of the replica in the course of the simulation is shown, whilst in (c,d) the distribution of the potential energies of the respective method is shown.

I have noticed even people who claim everything is predestined, and that we can do nothing to change it, look before they cross the road.

Stephen Hawking

3

Simulating the Physical World[†]: Proteins at the Silica Interface

[†]Quoted from *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics* by Herman J. C. Berendsen

PROTEINS AND OLIGOPEPTIDES are known to undergo a partial change of their conformations upon interaction with solid materials surfaces, leading to either unfolding or folding of their native structures in solution. This phenomenon governs important biological processes such as blood clotting or amyloid fibre formation³⁰, and determines the behavior of inorganic/organic interfaces during either biomineralization or, inversely, materials recognition by short peptides⁶⁸. Preventing conformational changes of protein-based drugs (such as antibodies) induced by adsorption/desorption at the walls of storage containers, typically SiO₂-coated glass vials, is of paramount importance for pharmaceutical industries⁷⁴, given the costs and risks associated with the lowered or modified activity of the drug molecules in a different folding state. Moreover, the tethering of active enzymes to solid supports, which has emerged as a promising route towards the biotechnological fabrication of environmentally sustainable catalysts, may lead to uncontrolled structural change and reduction of the enzymatic activity¹⁹¹. There is thus a big interest in understanding the interactions between polypeptides and inorganic surfaces at an atomistic level, a difficult task both for current experimental and modeling techniques^{107,167}.

Experimentally, Atomic Force Microscopy (AFM) methods have been used in rare cases for high-resolution imaging⁸² of adsorbed peptides and more often to measure the adhesion forces between polypeptides and materials⁶⁵. The free energy of adsorption can also be estimated via AFM force spectroscopy through the analysis of the adhesion force as a function of the loading rate¹²⁶, or by means of other indirect methods such as quartz crystal microbalance with dissipation (QCM-D)¹⁹² or SPR spectroscopy²³⁴. When the adsorption does not involve the formation or disruption of chemical bonds, classical molecular dynamics modeling have lead to impressive results concerning the prediction of adhesion forces¹⁰⁷

and free energies¹⁹⁰ of adsorbed polypeptides. Large progress towards consistency between simulation results and experiments has recently been achieved by developing and applying advanced simulation techniques, such as replica exchange methods, to explore the vast conformational space of polypeptides interacting with solid surfaces^{161,190,55,146}. Complementing the experimental results by providing atomistic resolution, these simulations have thus significantly advanced our understanding of biomolecular adsorption^{67,212}. More difficult to determine are the conformational changes associated with polypeptide adsorption.

Developing a method to predict the structure and compute the Circular Dichroism (CD) spectra in the near UV wavelength of oligopeptides either dissolved in water or adsorbed on materials surfaces is part of the work in this thesis and is presented in the following. It was recently published and detailed journal information can be found under the descriptor¹⁴⁶ in the bibliography. The method used here is based on molecular dynamics simulations and exploits a RESTMETAD algorithm⁴⁵. It is applied to study in detail the structure and conformational change by means of molecular dynamics simulations of an α -helical oligopeptide adsorbing on an anionic amorphous silica (SiO_2) substrate. Conformational changes of this model have been experimentally examined previously by Burkett & Read⁴⁰, Read & Burkett¹⁸¹.

As it will be seen, the presented simulation technique produces results in quantitative agreement with experiments. Firstly, results of an equilibrium unbiased MD simulations are shown in Chapter 3.1. Using this approach, it was only possible to capture a few microstates of the dissolved and adsorbed states (unless performed over a time scale comparable with the experimental one, which is presently impossible). Therefore, a quantification of the adsorption behavior from the previously unbiased MD trajectory of the model pep-

tide in terms of any experimental observable such as, e.g., the free energy of adsorption is not possible. To overcome this limitation, a thorough exploration of the phase space associated with the processes of adsorption and unfolding using Replica Exchange with Solute Tempering in conjunction with METADYNAMICS (RESTMETAD), as introduced in Camilioni et al.⁴⁵, Schneider & Colombi Ciacchi¹⁹⁰, is carried out in Chapter 3.1.1.

3.1 CONFORMATIONAL CHANGES OF A PEPTIDES ADSORBING TO SILICA

It is not trivial to find an experimentally well studied biological system able to simulate within reasonable time scales. Mostly, interesting systems, e.g. proteins, contain too much atoms to simulate and investigate them further on a high level of complexity.

In contrast, Burkett & Read⁴⁰ and Read & Burkett¹⁸¹ investigated thoroughly conformational changes of small peptides when they undergo adsorption on anionic or cationic silica. The α -helical peptide 4DAR₅, of sequence DDDDAAAARRRR, has been the object of detailed experimental studies concerning its partial unfolding after adsorption on anionic SiO₂ colloidal surface using a combination of CD and ¹H-NMR spectroscopy^{40,181}. In the dissolved state, the folded α -helical structure is stabilized by the central poly-A sequence despite a net dipole arising from the positively charged arginine side and the negatively charged aspartic acid side. The helical loss caused by the adsorption on the negatively charged SiO₂ surface at pH 9 has been quantified to amount to about 40 %⁴⁰. This behavior and the relatively small size in terms of atomic size was reason enough to study this molecule in more detail with molecular dynamics simulations.

First of all the 4DAR₅ peptide model was rebuilt in an attempt to use it in an atomistic simulation; Constructed initially in an ideal α -helical conformation using the LEaP package

of the Amber suite of programs. On a further 2 ns relaxation in bulk water using a molecular dynamic simulation at 300 K, it retained its initially helical structure. According to the experimental protocol of Burkett & Read⁴⁰, the N and C termini were capped by an acetyl group (COCH₃) and an amino cap (NH₂), respectively.

To simulate a colloidal SiO₂ surface a 22 Å thick, hydroxylated amorphous silica slab (devised in the work of Cole et al.⁵¹) was used as a starting point. Although the theoretical force field description is insufficient, as discussed in a previous chapter, the proposed hydroxylated silica model is in good agreement with literature and suits without any problems to the modified MS-Q force field (cf. Chapter 2.7.2). The slab had to be modified in order to represent experimental findings: On each side of the slab 5 OH groups were deprotonated, corresponding to a surface charge density of -0.087 C/m² or to a Si-O⁻ group density of 0.55 nm⁻². This is very close to the value of 0.54 nm⁻² estimated by Burkett & Read⁴⁰ by means of titration experiments with SiO₂ colloids dissolved in water at pH 9.0. Importantly, the form and used atomic charges of the modified MS-Q silica potential in addition with our recently developed potential for interfacial silica systems in contact with water⁴³ ensure that it can be readily used in combination with standard biomolecular force fields.

In summary, peptide-peptide, peptide-water and water-water interactions are described with the all-atom AMBER03 force field^{54,69} combined with the TIP3P water model¹⁴. Peptide and water interactions are described with the SiO₂ parameter set recently published in Butenuth et al.⁴³. It treats with particular care the point charges and the Lennard-Jones parameter of deprotonated silanol groups on the silica surface. Lorentz-Berthelot combination rules were employed to construct the Lennard-Jones interaction potentials between

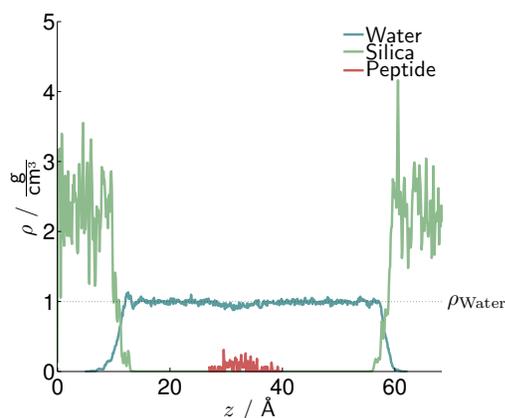


Figure 3.1: Density profiles of the interfacial system components used as starting configuration for further simulations of the helicity loss upon adsorption: a hydroxylated and deprotonated silica surface, Tip3p water and the 4DAR5 peptide.

the atomic species at what flexibility of the complete silica surface slab remained. Si-Si and Si-O interactions within the surface slab are described by the modified MS-Q potential originally introduced by Demiralp et al.⁶⁶. Finally, Na^+ counterions had to be added to the solvent to ensure charge neutrality of the combined silica, peptide and water system in order to allow the use of a long-range PPPM solver^{60,102} for coulombic interactions.

Periodic boundary conditions were imposed in all directions and while the periodicity of the SiO_2 surface slab dictated the size of the simulation cell in the xy plane ($60.5 \times 60.5 \text{ \AA}^2$), its height was manually adjusted to 68.7 \AA in order to reproduce the correct TIP3P water density of 0.998 g/cm^3 far away from the surface slab¹⁴². The resulting averaged density profiles for the silica slab, water and the peptide are given in Figure 3.1.

A graphical representation of the initially non-adsorbed and adsorbed system after 23 ns of molecular dynamic simulation is given in Figure 3.2a and Figure 3.2b, respectively. Averaged density profiles are calculated from a 2 ns simulation in which the peptide and a 1 \AA thick layer in the center of the silica slab were constrained to their initial positions in

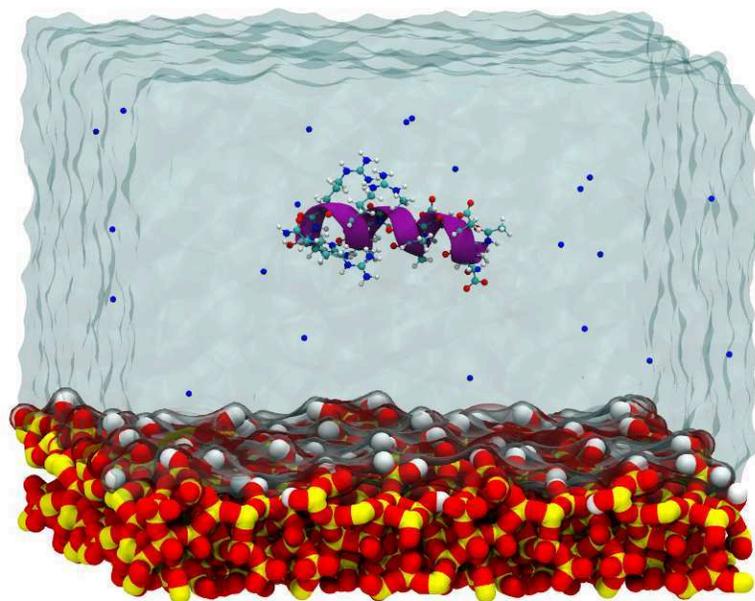
order to not distort the final density profiles. As a starting configuration served for this purpose the system shown in Figure 3.2a.

Molecular dynamic simulations were carried out with the LAMMPS simulation software¹⁷¹ extended with the PLUMED METADYNAMICS package²⁶. Usually METADYNAMICS is used to bias simulations in order to obtain free energy estimates of collective variables. However, it was used here only to monitor the change of peptide helicity during the simulation, without further biasing it. For this purpose, the structural helicity and the distance of the center of mass of the peptide to the surface were used. Both are predefined collective variables in METADYNAMICS, which are otherwise not found in a standalone LAMMPS version. The structural helicity is defined as the number of residues i adopting a (partially) helical conformation, based on their backbone dihedral angles Φ_i and Ψ_i (cf. Figure 3.3 and the correlation with the values of the adjacent residues $i \pm 1$):

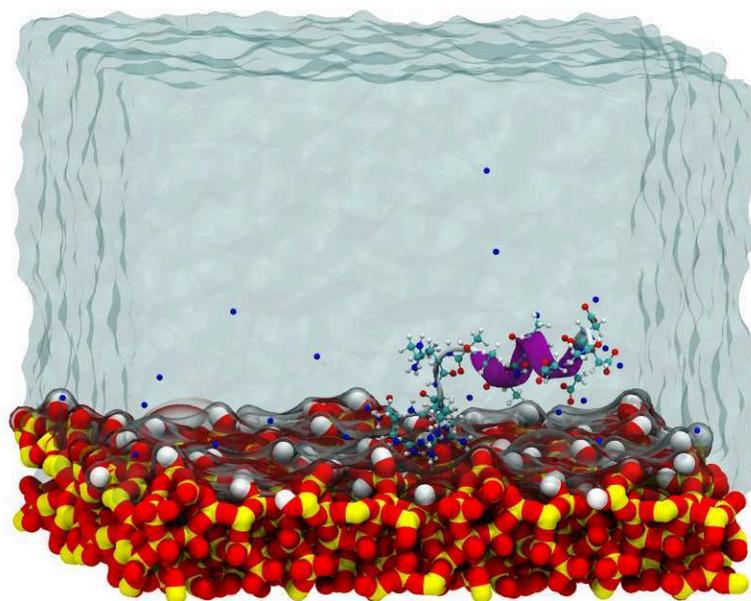
$$H = \sum_{i=2}^{N-1} \prod_{j=i-1}^{i+1} \frac{1}{4} [\cos(\Phi_j - \bar{\Phi}_i) + 1] [\cos(\Psi_j - \bar{\Psi}_i) + 1] . \quad (3.1)$$

In this definition $\bar{\Phi}_i$ and $\bar{\Psi}_i$ represent the target values of residue i associated with an α -helical conformation (-68.75° and -45.0° for all residues, respectively).

A timestep of 2.0 fs was chosen in all simulations and covalent bonds involving hydrogen atoms were constrained to their equilibrium values via the SHAKE algorithm¹⁸⁷. A PPPM Ewald solver^{60,102} was used to calculate long range Coulombic interactions. Pair interactions of the SiO₂ Morse potential as well as the real space part of the Coulombic interactions were truncated at 8.0 Å. As a consequence of the difficulties in applying a barostat to an interfacial system²⁴², all simulations were carried out within the NVT ensemble at 300 K



(a)



(b)

Figure 3.2: Graphical representations of (a) the initial non-adsorbed ($H \approx 9.5$) and (b) the adsorbed system ($H \approx 4.5$) after 23 ns of molecular dynamic simulation.

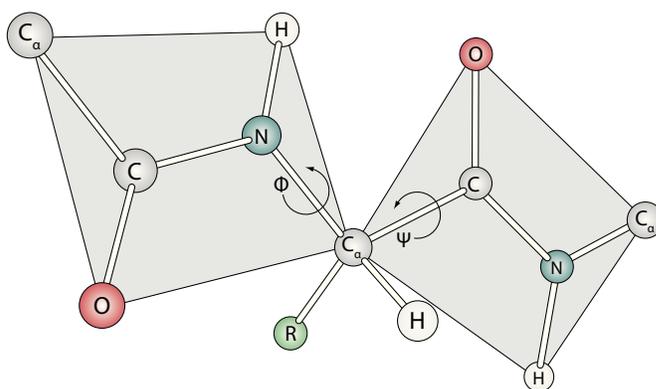


Figure 3.3: Illustration of the definition of backbone dihedral angles in proteins. Φ and Ψ are therein defined as the dihedral angle shared by the atoms C-N-C $_{\alpha}$ -C and N-C $_{\alpha}$ -C-N, respectively. R serves as a placeholder for the side chain of the amino acid.

employing a Nosé-Hoover thermostat.

Simulations were initially started with the fully-folded peptide dissolved in water placed 1.2 nm above a deprotonated amorphous silica surface model with a surface charge density corresponding to the one determined experimentally (Figure 3.2a). Electrostatic interactions drive the spontaneous adsorption of the peptide through the poly-R side within a few ns of MD simulation at constant room temperature. A graphical illustration of this is shown in Figure 3.2b). A more detailed temporal evolution of the helicity change and the peptides distance to the silica surface is given in Figure 3.4a and 3.4b, showing that after the initial adsorption partial unfolding is observed until the simulation is stopped after 23 ns.

z_{ssd} is defined here as the difference between the positions of the peptide's center of mass and the Gibbs' dividing surface along the z direction normal to the surface: $z_{ssd} = z_{com} - z_{GDS}$. The height z_{GDS} of the Gibbs Dividing Surface (GDS), which sets the zero of the collective variable corresponding to the position of the peptide center of mass in the simulations, was chosen so as to ensure that the surface excess of water molecules Γ is zero at

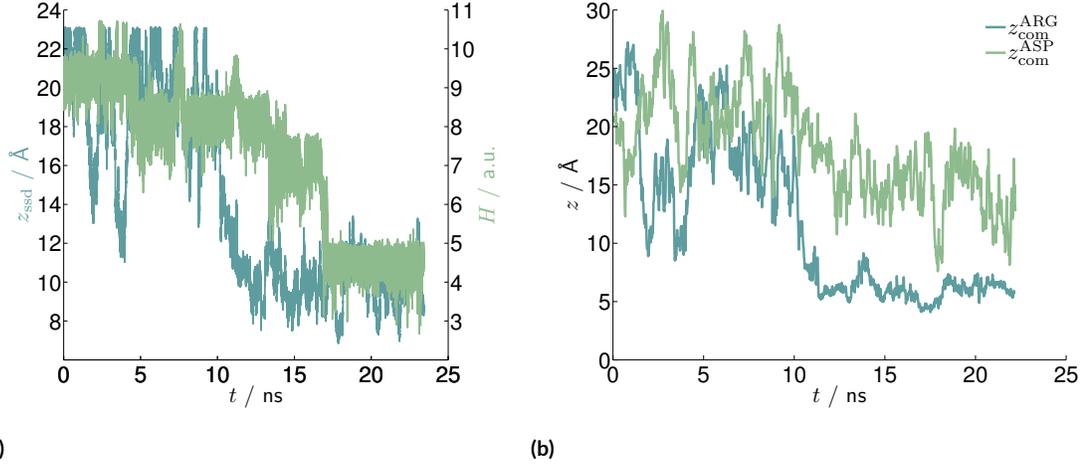


Figure 3.4: (c) z_{ssd} position (green) and helicity (blue) of the 4DAR5 peptide and (d) distance of the center of mass positions of arginine (ARG) and aspartic acid (ASP) residues to the surface during an unbiased MD simulation.

this point¹⁰⁶. This corresponds to fulfilling the condition

$$\Gamma = \int_0^{z_{\text{GDS}}} \rho_{\text{wat}}(z) dz - \int_{z_{\text{GDS}}}^{z_{\text{max}}} \rho_{\text{wat}}^{\text{bulk}} - \rho_{\text{wat}}(z) dz = 0, \quad (3.2)$$

where $\rho_{\text{wat}}(z)$ represents the water density at position z and $\rho_{\text{wat}}^{\text{bulk}} = 0.998 \text{ g/cm}^3$ is the TIP3P water density in liquid bulk. z_{max} was chosen to be 1/3 of the total length of the simulation box in z direction to ensure that z_{GDS} is not affected by the periodically repeated surface slab or by the density contribution of the peptide in the middle of the simulation box (see Figure 3.1). Using this method z_{GDL} is estimated to 11.0 Å with respect to our arbitrary choice of the atomic coordinates in the simulation box.

As mentioned before, It should be noted that equilibrium unbiased MD simulations are only able to capture a few microstates of the dissolved and adsorbed states unless performed over a time scale comparable with experimental ones, which is presently impossible. Therefore, from an unbiased MD trajectory a quantification of the adsorption behavior, in terms

of any experimental observable, is not possible. To overcome this limitation, a method is introduced in the next chapter to overcome this limitation and to enable a thorough exploration of the phase space associated with the processes of adsorption and unfolding.

3.1.1 QUANTIFYING STRUCTURAL CHANGES WITH ADVANCED MD SIMULATIONS

Choosing the distance z between the peptide's center of mass and the SiO_2 surface and the peptide's structural helicity H (as defined by the torsional angles of its backbone, see Eq. 3.1) as two independent internal collective variables, a complete exploration of the energy landscape associated with the helical folding and unfolding during the adsorption process (cf. Figure 3.8) is performed.

A reliable free energy estimate of a given reactive coordinate depends often on more than only itself, probably also on those reactive coordinates which have not been explicitly defined. It relies therefore on a sufficient sampling of the complete phase space necessary to calculate a robust free energy estimate of the desired reaction coordinate. A robust method to obtain free energy estimates of a given reactive coordinate can be achieved by combining METADYNAMICS and Replica Exchange with Solute Tempering (REST) as done previously by Camilloni et al.⁴⁵ or Schneider & Colombi Ciacchi¹⁹⁰. This method is perfectly suited for the purpose of the present topic and is thus used to calculate the free energy surface associated with the conformational change a peptide undergoes upon adsorption. The choice of parameter for this method are explained in the following. To compute free-energy landscapes, adaptive bias potentials were added during the course of the MD runs according to the METADYNAMICS scheme¹²⁴ in the well-tempered ensemble⁸ using a bias factor of 10. Thus, Gaussian hills with a height of 0.7 kcal/mol were added every 0.5 ps. The full width

half maximum of the Gaussians was chosen to be 0.1 and 0.3 Å for the helicity and distance collective variables, respectively.

Additionally, an increased sampling of degrees of freedom not explicitly defined in the METADYNAMICS algorithm was achieved by allowing exchanges between virtually elevated temperature replicas using the Replica Exchange with Solute Tempering method. Simulations were carried out as introduced in the previous chapter. Seven replicas of the 4DAR₅ peptide in contact with water were used. Their potential energy rescaling factors corresponding to solute temperatures of 300, 325, 350, 375, 400, 425, and 450 K. Due to the artificial nature of the rescaled potential energy landscapes in the auxiliary high-temperature replicas, only the results of the system at 300 K were considered for the evaluation, as discussed in Chapter 2.8.3. Exchanges between neighboring replicas were attempted every 0.5 ps, which, as depicted in Figure 3.5a, resulted in uniform occupancies of all temperatures for each replica and an exchange probability of 26 %. As noted in Figure 3.5, due to the limitation in visibility not all 1.5 million exchange events can be shown. Temperatures of the respective replicas are therein given every 1.5 ns, although every 0.5 ps an exchange was allowed. Figure 3.5a and 3.5b show that the condition of a good replica diffusion between all temperatures and a sufficient energy overlap is fulfilled. In Figure 3.6 are the z_{ssd} position and helicity H of the respective replica at 300 K shown, furthermore indicating a sufficient sampling of microstates with respect to the chosen collective variables of the system.

Although in the case of a converged free energy the rate of occurrence for the two collective variables should be ideally flat in 3.6 it can be assumed that a sufficient sampling is nonetheless achieved. Apart from the slightly increased sampling in proximity of the silica surface, which could be improved by increasing the simulation time, a higher sampling of

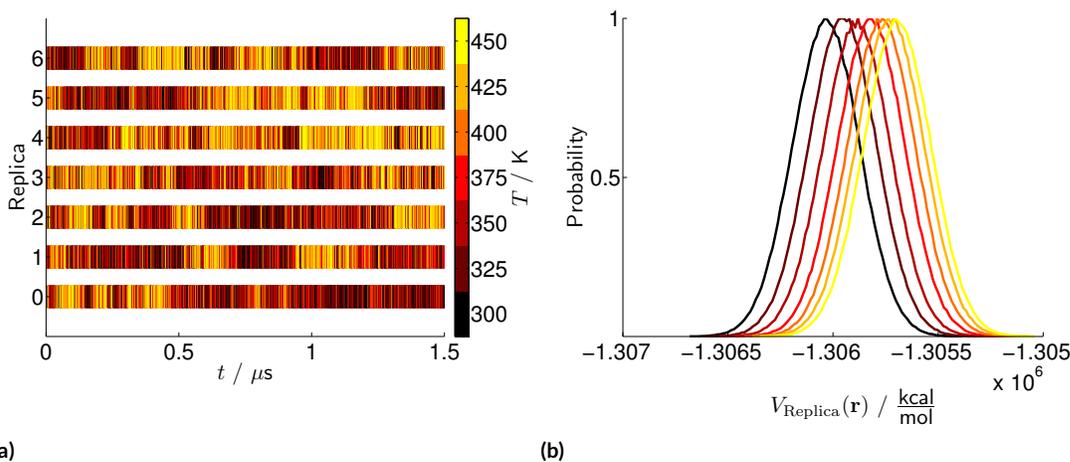


Figure 3.5: (a) Temperatures of the replicas with respect to the simulation time and (b) energy distributions at different temperatures from the RESTmetaD simulation of the 4DAR5 peptide including the anionic silica surface. Replica temperatures are given every 1.5 ns due to the immense amount of data (see discussion in the main text). Colors indicate the temperature from black to yellow, respectively from 300 to 450 K.

microstates with a lower helicity can be observed in Figure 3.6. This is probably due to the higher conformational space of a peptide with lower helicity and, thus, should have no effect on the final free energy surface. Vividly, this can be explained for a peptide with either a helicity of 0 or π .0. In the latter case there exist only one possible conformation, whereas for a helicity of 0 there exist several possible conformations of the peptide.

It can hence be assumed that by using the above mentioned collective variables it is possible to sample over all conformations of a peptide upon adsorption onto a silica surface. It is thus possible to distinguish between structural changes of the peptide induced by the adsorption process by calculating a free energy surface $F(H, z_{ssd})$ associated with the change of two collective variables H and z_{ssd} . The (z_{ssd}, H) free energy surface at 300 K calculated from a well-tempered RESTMETAD simulation using seven independent peptide replicas and lasting 1.5 μ s per replica is shown in Figure 3.7.

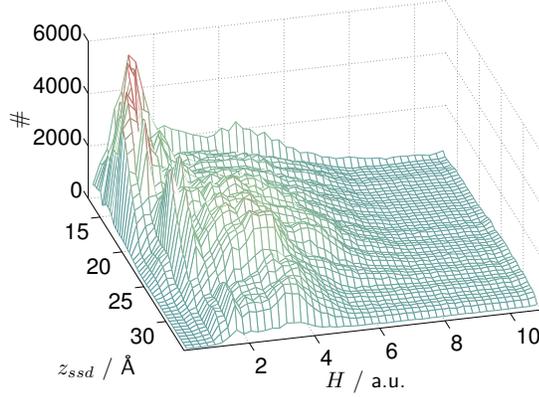


Figure 3.6: Rate of occurrence of the z_{ssd} position and helicity H during the RESTmetaD simulation of the 4DAR5 peptide including the anionic silica surface from the respective replica at 300 K. The mesh grid data is produced by binning all microstates into 30^2 equal sized boxes.

At a first glance, the free energy landscape at large z_{ssd} reveals a shallow minimum in the region of helicity between 6 and 8. A major adsorption channel leads towards the surface retaining the same helicity down to $z_{\text{ssd}} = 7.5 \text{ \AA}$, where a shallow local energy minimum is located. Further surface approach is possible only upon unfolding of the peptide, *i.e.* a decrease in helicity, eventually leading to a more stable minimum at $z_{\text{ssd}} = 2.5 \text{ \AA}$ and H between 1 and 3. Along the adsorption channel the energy barriers encountered are of the order of only 5 kcal/mol, which explains why in the unbiased MD simulation in Figure 3.4a adsorption and partial unfolding took place spontaneously at room temperature.

Individual profiles of the free energy along either z_{ssd} or H are obtained upon integration of the two-dimensional surface $F(H, z_{\text{ssd}})$ over the other variable¹⁹⁹:

$$F(q_1) = -\frac{1}{\beta} \ln \left[\int_{q_2, \min}^{q_2, \max} e^{-\beta F(q_1, q_2)} dq_2 \right], \text{ where } q_{1,2} = H, z_{\text{ssd}}. \quad (3.3)$$

The temporal evolution of $F(z_{\text{ssd}})$ during the RESTMETAD simulation is shown in

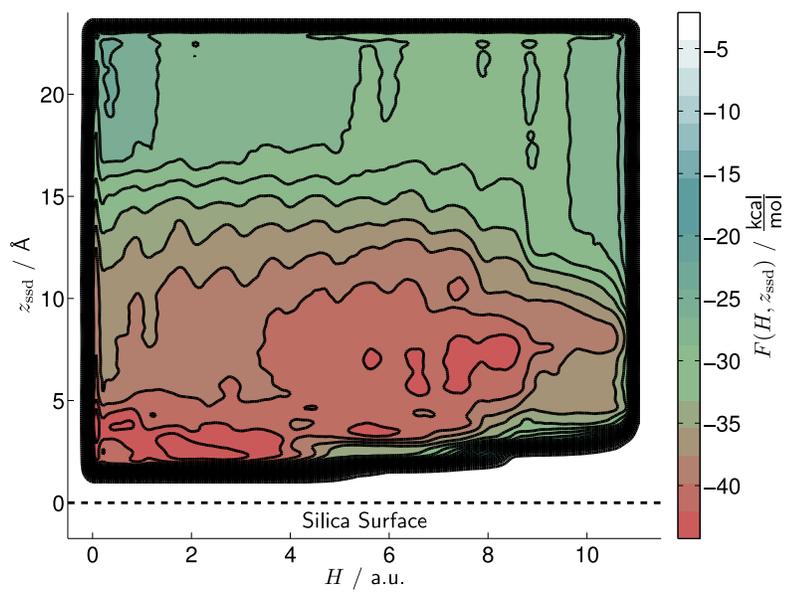


Figure 3.7: Free energy surface $F(H, z_{\text{ssd}})$ of the 4DAR5 peptide in contact with an anionic silica surface.

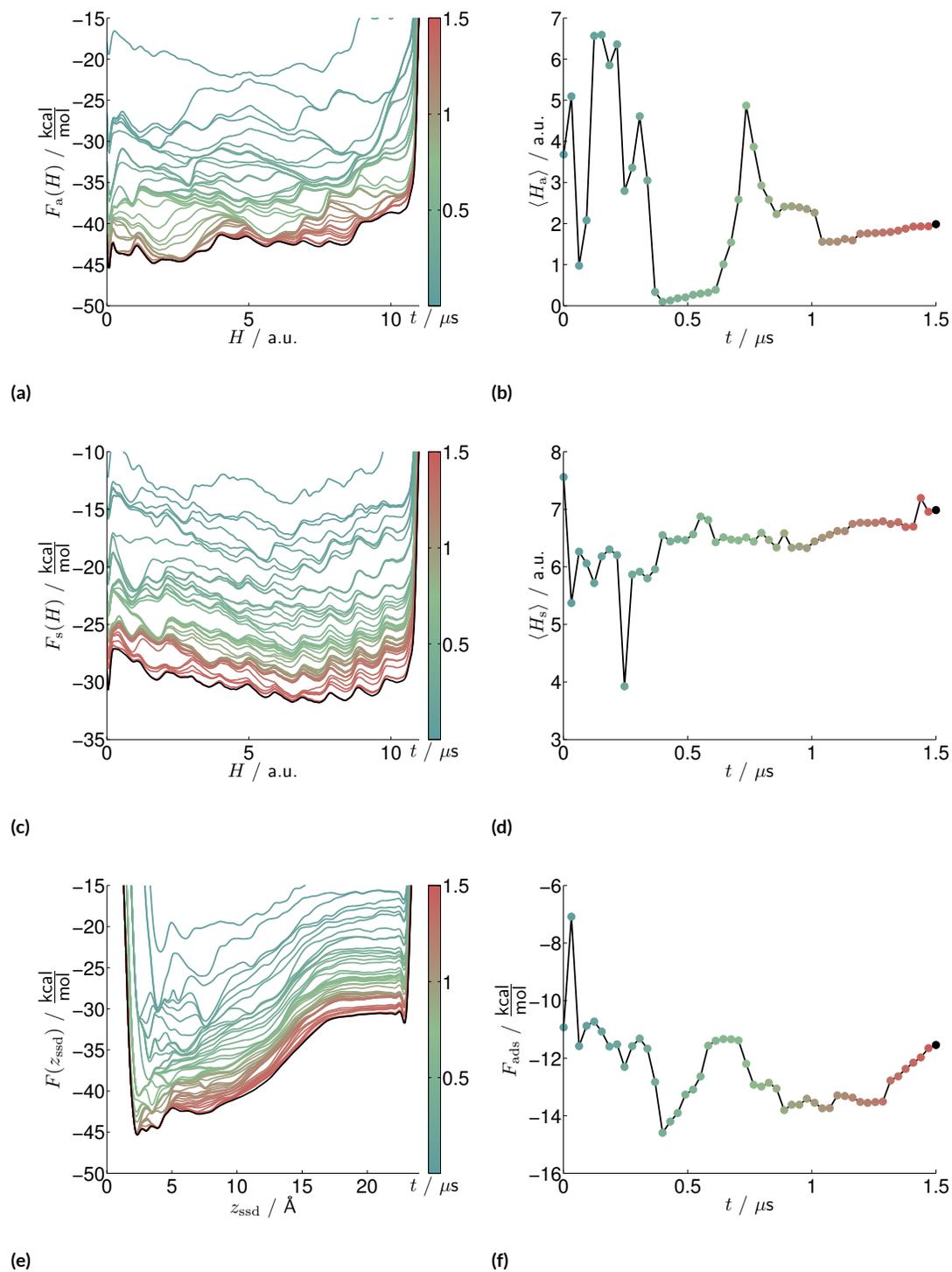


Figure 3.8: Temporal evolution of the free energy profile of the 4DAR5 peptide adsorbing on silica $F(H_a)$ (a), the conformational free energy associated with the helicity in bulk solution $F(H_s)$ (c) and adsorbed on silica $F(z_{\text{ssd}})$ (e); Convergence of the mean adsorbed helicity $\langle H_a \rangle$ (d), the mean dissolved helicity $\langle H_s \rangle$ (e) and the free energy of adsorption ΔF_{ads} (f).

Figure 3.8e, where the color code represents the simulation time (red to blue from 0 to 1500 ns). The shape of the free energy profile at the end of the simulation allows us to define an *adsorbed* state (a) for distances lower than $z_{\text{ssd},0} = 16.5 \text{ \AA}$ and a *dissolved* state (s) for larger distances, where the free energy is flat, indicating that the peptide does not experience any surface interaction and behaves as in bulk solution. The average probability densities of finding the peptide either in the adsorbed or dissolved state are

$$\begin{aligned}\rho_a &= \frac{1}{z_{\text{ssd},0} - z_{\text{ssd},\text{min}}} \int_{z_{\text{ssd},\text{min}}}^{z_{\text{ssd},0}} e^{-\beta F(z_{\text{ssd}})} dz, \\ \rho_s &= \frac{1}{z_{\text{ssd},\text{max}} - z_{\text{ssd},0}} \int_{z_{\text{ssd},0}}^{z_{\text{ssd},\text{max}}} e^{-\beta F(z_{\text{ssd}})} dz.\end{aligned}\quad (3.4)$$

The free energy of adsorption ΔF_{ads} can now be computed as¹⁶¹

$$\Delta F_{\text{ads}} = -k_B T \ln \left(\frac{\rho_a}{\rho_s} \right). \quad (3.5)$$

It should be noted that this definition of free energy difference refers to the molecular concentrations, in bulk solution as well as in proximity of the surface, as the respective standard states. This differs from the commonly used experimental standard states of solution concentration and surface coverage. While the RESTMETAD simulation evolves, ΔF_{ads} varies as shown in Figure 3.8f. Note that even after 1.5 μs convergence is only reached within an error of about 2 kcal/mol at the value of 12 kcal/mol. This indicates that the sampling orthogonal to the collective variables, though greatly enhanced by the REST technique, becomes a limiting factor here.

Profiles of the free energy as a function of H both in the adsorbed and dissolved states $F_a(H)$ and $F_s(H)$ (Fig. 3.8a and 3.8c, respectively), using Equation (3.3) with appropriate integral limits can be furthermore computed. At this point, it is possible to calculate the

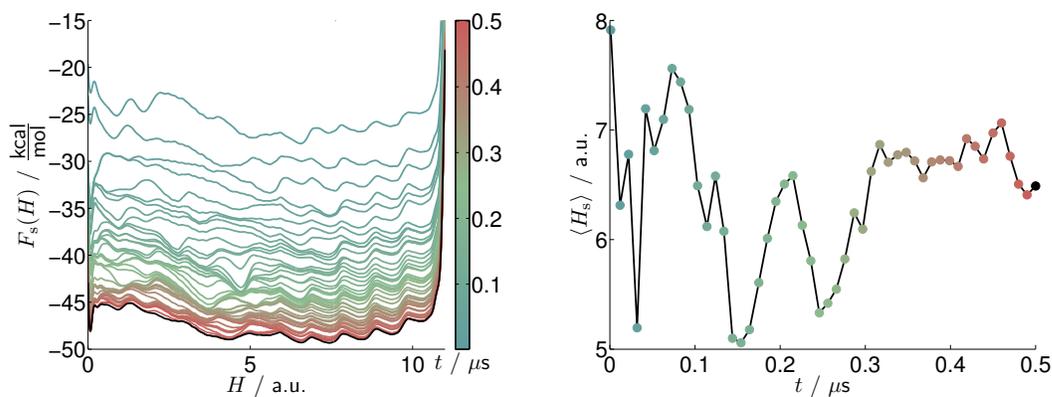
expectation value of the peptide's helicity in either state from

$$\langle H_{a,s} \rangle = Z_{a,s}^{-1} \int_{H_{\min}}^{H_{\max}} e^{-\beta F_{a,s}(H)} \cdot H \, dH, \quad (3.6)$$

where Z_a and Z_s represent the respective partition functions. The evolutions of $\langle H_a \rangle$ and $\langle H_s \rangle$ during the RESTMETAD simulation are shown in Figure 3.8b and 3.8d. The final values of 2.0 and 7.0 appear to be reasonably well converged, with an error of about 0.5 in both cases. To check the convergence in the dissolved state an additional RESTMETAD simulation of the peptide in pure bulk water, obtaining a final value of $\langle H_s \rangle = 6.5$, which is within the previously identified error, has been performed. This was done in the same manner as the previous simulation of the 4DAR₅ peptide in contact with an anionic silica surface, however, with omission of the surface. A graphical representation of the results from this simulation can be found in Figure 3.9. More information on the diffusion of the replicas and the distribution of helicities of the RESTMETAD simulation without considering an anionic silica surface are shown in Figure 3.10 and 3.11, respectively. The exchange probability that two replicas will swap temperatures for each exchange attempt was estimated over the whole simulation to 29 %.

3.1.2 THEORETICAL CD SPECTRA PREDICTION FROM MD SIMULATIONS

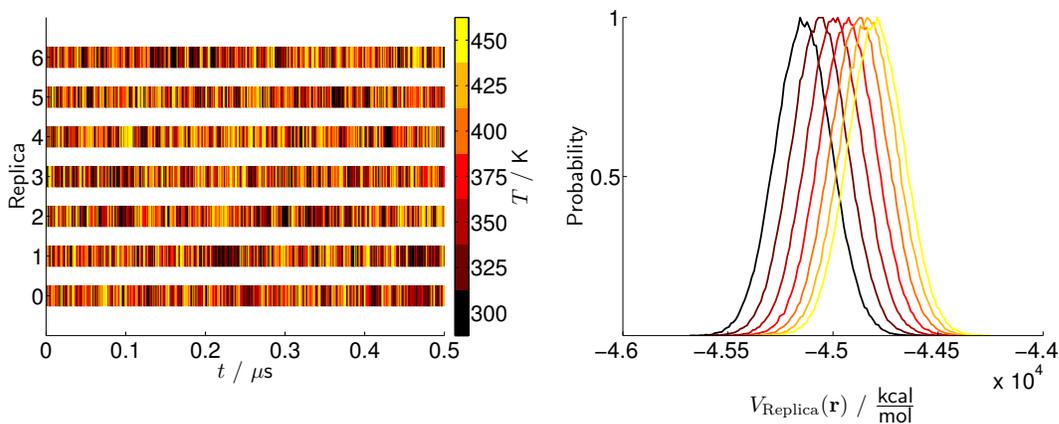
While modeling tools have been developed to compute the CD spectrum associated with a single biomolecular structure³⁷, the knowledge of a measured spectrum is in most cases insufficient to determine uniquely the unknown biomolecule's folding state. This is due to the fact that in a solution of polypeptides, either dissolved or adsorbed to a colloid surface, all possible structures (microscopic conformational states) are present at the same time,



(a)

(b)

Figure 3.9: Evolution of the free energy profile as a function of the helicity during a Replica Exchange with Solute Tempering in combination with Metadynamics simulation of the 4DAR5 peptide in bulk Tip3p water (cubic simulation cell with edge-length of 60 Å) (a). Convergence of $\langle H_s \rangle$ (b). Different colors (from red to blue) represent the temporal evolution (from 0 to 500 ns) of the system.



(a)

(b)

Figure 3.10: (a) Temperatures of the replicas with respect to the simulation time and (b) energy distributions at different temperatures from the RESTmetaD simulation of the 4DAR5 peptide with surface ommision. Replica temperatures are given every 1.5 ns due to the immense amount of data (see discussion in the main text). Colors indicate the temperature from black to yellow, respectively from 300 to 450 K.

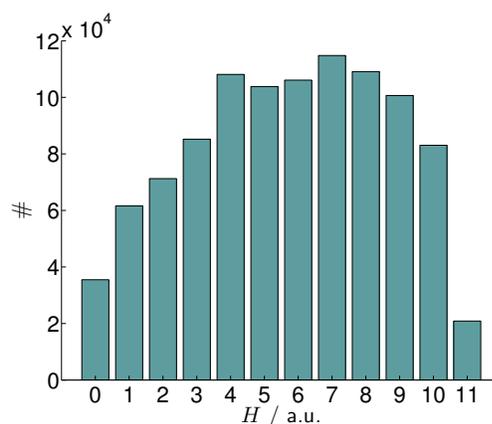


Figure 3.11: Distribution of the helicity H of the 4DAR5 peptide in solution without a silica surface in the lowest temperature replica at 300 K from a 500 ns RESTmetaD simulation with 7 replicas and a temperature difference of 25 K between each replica.

with individual probabilities determined by their relative variations of free energy through the Boltzmann distribution. In other words, particularly for small oligopeptides which exhibit broad and shallow folding/unfolding energy landscapes, one should consider two *ensembles of structures* in the macroscopic dissolved and adsorbed states, each associated with its distinct CD spectrum.

A method is presented in the following to predict the structure and compute the CD spectra of oligopeptides either dissolved in water or adsorbed on materials surfaces. A detailed explanation of the underlying CD spectroscopy theory can be found in appendix A based on the review of Bulheller et al.³⁸. Statistical analysis of the previously obtained converged free energy landscape provides average fractional helicities in the adsorbed and desorbed states matching the experimental estimates previously obtained by ¹H-NMR and CD spectroscopy^{40,181}. Moreover, the CD ellipticity intensity Θ computed with the DichroCalc software^{37,240} for each microscopic state is treated as an external collective variable and is calculated by means of a reweighting procedure (cf. Chapter 2.8.2) applied to the biased

METADYNAMICS trajectory in the (z_{ssd}, H) space⁸. This allows us to predict the full CD spectra associated with the adsorbed and desorbed states, and in particular the ellipticity value at a wavelength of 222 nm, Θ_{222} . Thus it is shown that Θ_{222} is linearly correlated with the peptide helicity, providing a support for this often-employed hypothesis³⁸.

The results demonstrate how CD spectroscopy measurements can be put on an equal footing with atomistic MD modeling, and open up a viable way to link experimental spectra with the amount of secondary structure elements beyond the simple case of a single α -helix. In fact, discrepancies between measured and theoretically computed CD spectra³⁸ are believed not to arise from limitations of the theoretical formalism (or its software implementation) of the underlying CD spectroscopy but from a lack of statistical averaging over the correct ensemble of biomolecular structures.

The knowledge of a vast number of conformational microstates in the (z_{ssd}, H) phase space together with their associated free energy gained through the RESTMETAD simulation allows to predict the experimentally measurable CD spectra of the 4DAR5 peptide in both the adsorbed and dissolved states. In the CD literature, the *fractional helicity* f_{H} of a short peptide containing N amino acids is defined as

$$f_{\text{H}} = \frac{\Theta_{222} - \Theta_{222}^{\text{coil}}}{\Theta_{222}^{\text{hel}}(N) - \Theta_{222}^{\text{coil}}}, \quad (3.7)$$

where Θ_{222} is the CD ellipticity intensity measured at 222 nm, $\Theta_{222}^{\text{hel}}(N)$ is the intensity of a N -mer peptide with ideal α -helical structure, and $\Theta_{222}^{\text{coil}}$ is the intensity of a random-coil polypeptide. Expressing Θ in the usual units of $\text{deg cm}^2 \text{ dmol}^{-1}$, it can be assumed, follow-

ing Gokce et al.⁸⁸, that $\Theta_{222}^{\text{coil}} \approx 250$. It is also agreed that the empirical formula

$$\Theta_{222}^{\text{hel}}(N) = \Theta_{222}^{\infty}(1 - k/N) + cT, \quad (3.8)$$

holds if N is sufficiently large²⁵ (see below). Estimates for Θ_{222}^{∞} range from $-3.7 \cdot 10^4$ to $-4.4 \cdot 10^4 \text{ deg cm}^2 \text{ dmol}^{-1}$ in the literature¹³⁶. Here, the theoretical ellipticity intensity of an infinitely long ideal α -helix is chosen to be $\Theta_{222}^{\infty} = -4 \cdot 10^4 \text{ deg cm}^2 \text{ dmol}^{-1}$. The constants k and c that accounts for non-hydrogen-bonded amide carbonyl groups at the peptide termini³⁸ and the temperature dependence, respectively, are chosen to be $k = 2.5$ and $c = 100 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ K}^{-1}$ ⁸¹.

The validity of the empirical formula above by calculating $\Theta_{222}^{\text{hel}}$ using the DichroCalc software³⁷ has been tested for ideally helical peptides with sequence $(n-1)\text{DAR}n$ (i.e. $\text{D}_{n-1}\text{A}_n\text{R}_{n-1}$) and compared to the results with the values predicted by Equation (3.8). As shown in Figure 3.12, the DichroCalc values start deviating from the empirical prediction for $N < 7$. Therefore, for the case of the $4\text{DAR}5$ peptide considered here, either the actual DichroCalc values or the empirical Equation (3.8) can be used without distinction to obtain the reference $\Theta_{222}^{\text{hel}}(N)$ value corresponding to a fully folded helical conformation.

Using the assumptions above, from a measure of Θ_{222} for a peptide in solution it is possible to compute the fractional helicity f_{H} from Equation (3.7). It is commonly taken for granted that f_{H} is directly related to the *structural helicity* of the peptide, as previously suggested from combined AFM and CD experiments²⁰⁹. However, the contribution of the uncoiled part of peptides to the CD signal is still debated⁸⁸ albeit theoretical CD spectra for fully helical peptides yield reliable results over the whole range of UV-wavelengths³⁸. This hypothesis is put on a firm basis by computing, for all the microscopic states sampled

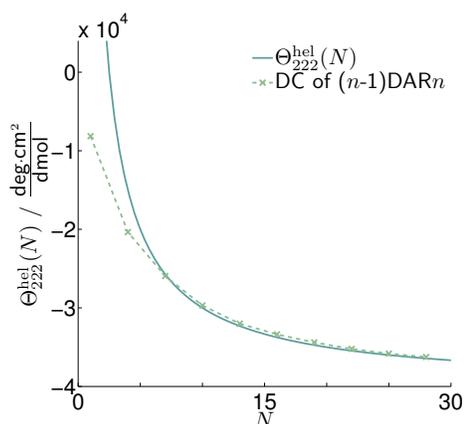


Figure 3.12: CD ellipticities of ideally helical $(n-1)$ DAR n peptides containing $N = 3n - 2$ amino acids computed either with DichroCalc (DC, dashed line) or through the empirical formula of $\Theta_{222}^{\text{hel}}(N)$ from Equation (3.8) (solid line)

during the RESTMETAD simulation of the 4DAR_5 peptide in bulk water, their structural helicity H via Equation (3.1) and their Θ_{222} intensity values via the DichroCalc software. Indeed, as shown in Figure 3.13, the two values are to a very good extent linearly correlated. Note, however, the large spreading of the values of Θ_{222} at a given f_H (of the order of $10^4 \text{ deg cm}^2 \text{ dmol}^{-1}$), or by the values of f_H at a given Θ_{222} (of the order of 0.4). This highlights the fact that evincing a molecular structure from a CD measurement is a non-uniquely defined problem, even for the simplest case of a partially helical and partially random-coiled peptide. The connection between molecular modeling and measured CD spectra should instead be performed by taking into account the whole ensemble of structures defining a macroscopic state of the biomolecule, as it shall be demonstrated in the following.

In the previous section the two-dimensional free-energy surface $F(z_{\text{ssd}}, H)$ was obtained. It consists of two ensembles of conformational microstates arbitrarily separated into adsorbed and dissolved states from the plane at $z_{\text{ssd},0} = 16.5 \text{ \AA}$. The dependence of the free

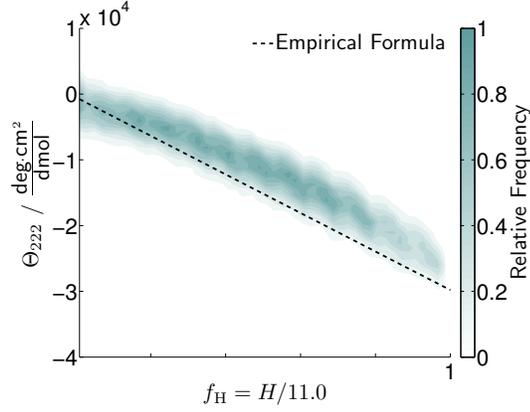


Figure 3.13: Relation between Θ_{222} and the fractional helicity $f_H = H/11.0$. The density distribution is calculated with DichroCalc from 10^5 snapshots of RESTmetaD trajectories for 4DAR5 in bulk water and adsorbed on SiO_2 . The dashed line is drawn according to Equation (3.7) with parameter mentioned in the main text.

energy on any other unbiased collective variable (i.e. different from z_{ssd} or H) can be obtained appropriately by reweighting the biased RESTMETAD trajectory, as put forward in Bonomi et al.²⁵. In particular, the CD ellipticity at a given wavelength, Θ_λ , was considered as an additional collective variable. This quantity can be calculated with the DichroCalc software for each microstate visited during the RESTMETAD simulation (see¹⁴⁶). As a result of the reweighting a two-dimensional free energy profile $F(\Theta_\lambda, z_{\text{ssd}})$ is obtained, out of which we can integrate the one-dimensional profiles $G_a(\Theta_\lambda)$ and $G_s(\Theta_\lambda)$ relative to the adsorbed and dissolved states, respectively (see Eq. (3.3)). For the 222 nm wavelength the free energy surface of $F(\Theta_{222}, z_{\text{ssd}})$ is shown in Figure 3.14. The expectation value of the Θ_λ in each of the two states can be now computed from

$$\langle \Theta_\lambda^{a,s} \rangle = Z^{-1} \int_{\Theta_{\min}}^{\Theta_{\max}} e^{-\beta F_{a,s}(\Theta_\lambda)} \cdot \Theta_\lambda d\Theta_\lambda. \quad (3.9)$$

In summary, for each of the two macroscopic states (a) and (s) we can predict compu-

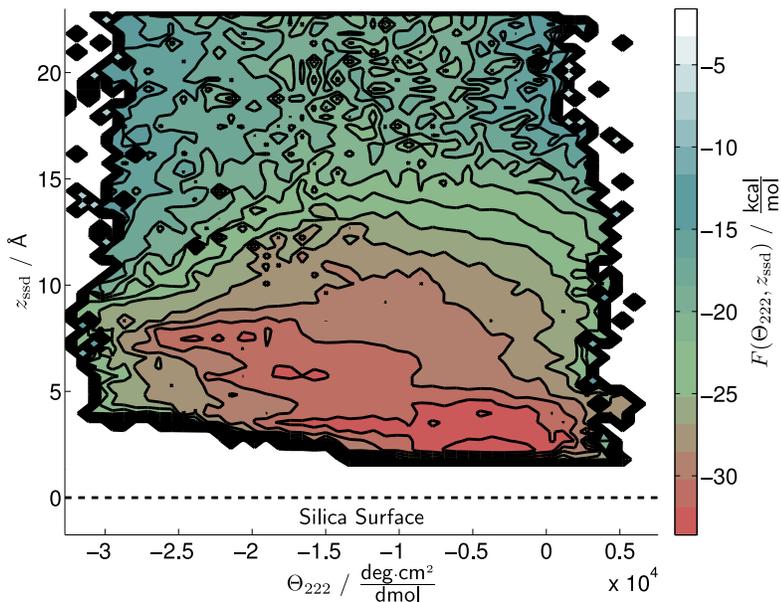


Figure 3.14: Reweighted free energy surface of the unbiased Θ_{222} signal and z_{ssd} from a RESTmetaD simulation. The isolines are separated by 1.68 kcal/mol.

tationally without empirical assumptions (other than the generic force field and DichroCalc parameter sets) both the average helicity $\langle H_{a,s} \rangle$ and the average ellipticity, for instance at 222 nm, $\langle \Theta_{222}^{a,s} \rangle$. This allows us to perform a comparison between the experimental results obtained by CD spectroscopy for the 4DAR₅ peptide both regarding Θ_{222} and the fractional helicity, which is computed as $\langle f_H \rangle = \langle H \rangle / (N - 2)$, where N , the number of amino acids in the peptide, is 13 in this case. In Figure 3.15a open circles represent the experimentally measured values^{40,181} of Θ_{222} versus the corresponding estimated experimental fractional helicities f_H (Eq. (3.7)) for the following cases: (i) an ideal random coil, for which it is assumed experimentally that $\Theta_{222} = 250 \text{ deg cm}^2 \text{ dmol}^{-1}$ and $f_H = 0$ ²¹⁵; (ii) an ideal α -helix ($f_H = 1$), for which Equation (3.8) is assumed to hold; (iii) the dissolved peptide in solu-

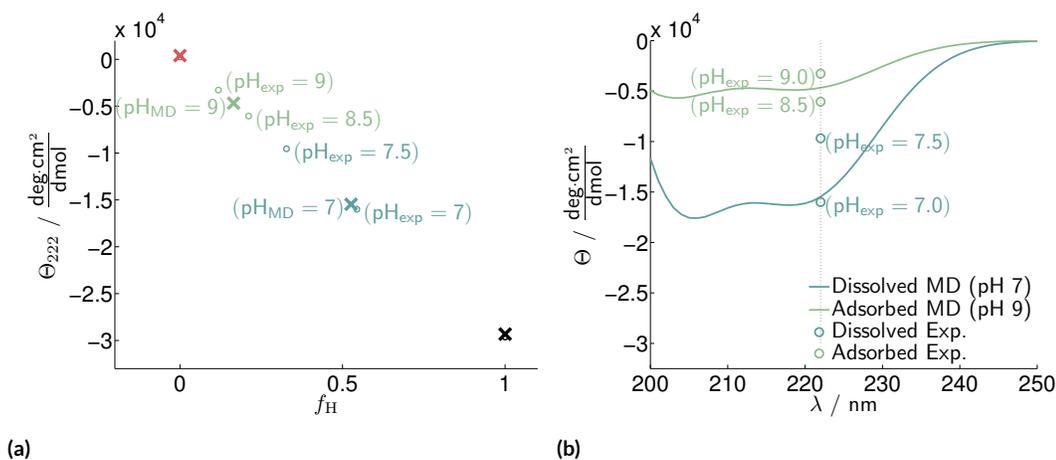


Figure 3.15: (a) $\langle \Theta_{222} \rangle$ values computed from reweighted RESTmetaD trajectories plotted versus the corresponding theoretically predicted $\langle f_H \rangle$ values (crosses) and compared with experimental estimates at various pH^{40,181} (circles). Blue represents adsorbed, green dissolved, red random coil and black fully helical states. (b) Full CD spectra of the 4DAR5 peptide in solution and adsorbed on anionic SiO₂ colloids predicted by the free-energy based method in comparison with the corresponding experimental measurement^{40,181}. In case of experimental results for the dissolved peptide only one point could be displayed.

tions at pH 7.0 and 7.5; (iv) the peptide adsorbed on the SiO₂ colloid surface in solutions at pH 8.5 and 9.0. In the same figure, computational predictions of $\langle \Theta_{222} \rangle$ are indicated with crosses, for the corresponding cases: (i) and (ii) the peptide in bulk water constraining H to either \circ or II ; (iii) the peptide dissolved in pure bulk water (thus formally at pH 7.0); (iv) the peptide in the adsorbed state on a SiO₂ surface with net charge roughly corresponding to the experimental charge density at pH 9.0.

It should be noted, first of all, that all values, both experimental and theoretical, lie on a straight line, again demonstrating the validity of the linear assumption in Equation (3.7). Notable is also the strikingly good agreement between the experimental measurements and the computational predictions of both Θ_{222} and f_H for both the adsorbed and the dissolved state of the 4DAR₅ peptide. The slight deviation for the adsorbed case (which lies between

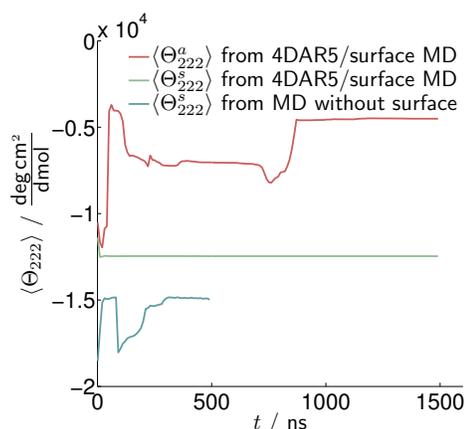


Figure 3.16: Evolution of the reweighted $\langle \Theta_{222} \rangle$ signal during the RESTmetaD simulations of the 4DAR5 peptide in the presence of the silica surface for the adsorbed state (red) and the dissolved state (green) and for the 4DAR5 peptide in pure bulk Tip3p water with surface omission (blue).

the experimental values at pH 9.0 and 8.5) is most probably due either to the imprecise distribution of net charges on the surface model with respect to the experimental reality, or to the inaccuracy of the used force field in the case of interfacial interactions.

Finally, in Figure 3.15b the full CD spectra is reported in the range of wavelengths extending from 200 to 250 nm of the dissolved and adsorbed peptide (see Eq. (3.6)), along with available experimental data at 222 nm^{40,181}. In order to assess an error estimate for the reweighted Θ_{222} signal a similar procedure as proposed by Berteotti et al.²³ was followed. Namely, in Figure 3.16 the temporal evolutions of the reweighted $\langle \Theta_{222} \rangle$ values is shown for the adsorbed and dissolved states, as computed every 10 ns from the REST-METAD simulation including the SiO₂ surface. It is remarkable that the average values $\langle \Theta_{222} \rangle$ converge very quickly. In addition, the temporal evolution of $\langle \Theta_{222} \rangle$ is computed for the dissolved case from a separate simulation of the peptide in pure bulk water, in the absence of the surface. The converged values in panels of $\langle \Theta_{222}^s \rangle$ from the simulation with and without the surface should be ideally the same. However, the difference between them

is $2500 \text{ deg cm}^2 \text{ dmol}^{-1}$, which can be taken as the maximum error associated with the estimates of $\langle \Theta_{222}^s \rangle$. A reason for this discrepancy could arise from the increased sampling in close proximity of the surface, due to the presence of energetically more favorable states, leading to a more sparse sampling of the dissolved states. For what reason it is assumed that the predicted value for $\langle \Theta_{222}^s \rangle$ from the simulation without a surface yields a more accurate estimate.

3.2 ADSORPTION FREE ENERGY ESTIMATES OF A POLYPEPTIDE ON AMORPHOUS SiO_2 USING SIMULATIONS AND FORCE SPECTROSCOPY EXPERIMENTS

The interaction between biomolecules and solid surfaces has become of eminent interest in fields ranging from basic research up to industrial product design^{56,149}. For instance, the development of pharmaceutical packaging for protein-based drugs relies on coating materials that inhibit protein adhesion on the packaging surfaces, and avoid conformational changes of the active components caused by surface adsorption. Anti-ice or anti-fouling coatings are also realized through the immobilization of proteins on solid substrates^{137,151,6}, and novel biomimetic materials can be synthesized by mineralization of short polypeptide sequences that selectively recognize and strongly bind to inorganic solid phases^{211,212,55,201}. Therefore, experimental and simulation effort has been recently spent for a rationalization of the fundamental physical processes that govern the biomolecule-surface interactions at an atomic scale. In this context, several methods that are able to indirectly quantify the free energy of adsorption ΔG_{ads} of short polypeptides on solid materials have been proposed¹⁷⁴. However, an unambiguous, quantitative comparison between different methods, and especially between experiments and simulations, has been achieved only in rare cases^{180,163,107,49}.

Experimental methods that can be used to estimate ΔG_{ads} are for instance quartz-crystal microbalance with dissipation (QCM-D), surface plasmon resonance (SPR) spectroscopy²³⁴, or isothermal titration calorimetry (ITC), concomitantly with the application of adequate adsorption isotherm models, such as the one of Langmuir¹⁹². However, the applicability of SPR or QCM-D rarely goes beyond model systems, since they are limited to nanoscale-thick material coatings on a sensor chip^{119,116,3,214}. Alternatively, force-spectroscopy (FS) methods, for instance based on atomic force microscopy (AFM), can be employed, provided that a relationship between the directly measured adhesion forces and ΔG_{ads} exist^{234,235,214}. An advantage of AFM-based FS (briefly, AFM-FS) is that a variety of substrates and probe molecules can be investigated^{154,225}. Moreover, several models have been proposed to explain the dependencies of the force required to break a chemical or physical bond (within a folded biomolecule, between a receptor and a ligand, or between a molecule and a surface) on the bond loading rate^{75,81,70,71,183}. Indirectly, many of these models are able to provide estimates on ΔG_{ads} , at least under a limited set of conditions, such as under small or large loading rates, or for negligible molecule/surface friction^{122,93}. A comparison between different models applied to the case of the binding forces within an amyloid- β fibre can be found in the work of Hane et al.⁹³

A widely applied model has been introduced by Friddle et al.⁸¹, generalizing the initial approach of Evans & Ritchie⁷⁵ to take into account binding/rebinding equilibria and the presence of multiple bonds. Another approach, based on the original Bell and Evans model, takes into explicit account the contribution of flexible linker molecules to the loading rate^{70,71}. Alternatively, the thermodynamics of the desorption event of long polypeptide molecules from solid/liquid interfaces substrates have been analyzed by various au-

thors^{112,113,166,122}. Particularly interesting is the analysis of Krysiak et al.¹²², since in their model the free energy of adsorption can be estimated without explicit knowledge of the contour length of the linker molecule, which is generally unknown. However, the conclusions of this work hold only for the case of frictionless substrates. Recently, Bullerjahn et al.³⁹ proposed a model which describes reasonably well the widespread spectra of low and high loading rates, which is particularly useful to analyze dynamic force spectra calculated theoretically by means of molecular dynamics simulations.

Simulations methods have also emerged only very recently as a viable way of predicting the adhesion forces and adsorption free energies at bio/inorganic interfaces^{190,146,163,67,161}. Crucial to this regard has been the application of methods that thoroughly sample the conformational space during the adsorption/desorption process, such as the Hamiltonian Replica Exchange⁸⁴ and its variants^{205,132,232}, alone or in combination with METADYNAMICS^{124,45,190,146}. Furthermore, non-equilibrium simulation methods such as steered molecular dynamics (SMD) are able to reveal details of the actual reaction paths leading to the (constrained) detachment of biomolecules from solid surfaces. SMD simulations would thus be in principle directly comparable to AFM-FS experiments, if the bond loading rates applied in the experiments and simulations were the same. Unfortunately, this is not the case since the computational cost of the simulations only allows the molecule to be pulled off the surface at very high speed (of the order of 0.1 m/s or larger), and at reasonably large values of the harmonic spring constant of the pulling constraint. It is also to be noted that extracting equilibrium quantities such as adsorption free energies from non-equilibrium simulations can be an extremely difficult task²⁴, because of the necessity of complete phase-space sampling. Therefore, the famous equality of Jarzynski, that calculates the equilibrium

free-energy difference between two states from the complete set of non-equilibrium work values associated with each individual trajectory that connects the same states, has had only limited practical applicability so far¹²⁰.

The present work concentrates on a model system consisting of a tetrapeptide with sequence GCRL (glycine, cysteine, arginine and leucine)²⁰¹ interacting with an amorphous SiO₂ surface model at neutral pH, for which we have developed a realistic atomistic structure and an accurate force field in previous works^{51,52,50,43}. The adsorption free energy ΔG_{ads} is theoretically predicted using both equilibrium (RESTMETAD⁴⁵) and non-equilibrium (SMD) methods via Jarzynski's equality¹¹³. AFM-FS experiments at variable loading rates are then performed and the results interpreted with the models of Friddle⁸¹ and Krysiak¹²² to provide experimental free energy estimates. As a remark, a similar formalism to that of Krysiak et al.¹²² is found in Paturej et al.¹⁶⁶. Beside comparing the values obtained via the experiments and simulations, a goal of this work is to highlight the advantages and shortcomings of each of the methods employed.

All MD simulations were carried out using the LAMMPS simulation package¹⁷¹ utilizing the AMBER03 force field^{54,69} in combination with the TIP3P water model¹¹⁴. Interactions between the silica surface, biomolecules and water are described using the recently published force field of Butenuth et al.⁴³. The bulk silica is described by an own modified version of the potential of Demiralp et al.⁶⁶, as described in detail in Meißner et al.¹⁴⁶. The input structures for the (GCRL) and (GCRL)₅ peptides are generated using the LEaP suite of the AMBER software package. The N-terminus and C-terminus of the (GCRL) and (GCRL)₅ peptides have been acetylated and capped with a methylamine group, respectively. The deprotonation of a silica surface depends strongly on pH, ionic strength and

particle diameter¹⁶⁹. Taking into account the counterions inserted in the simulations to ensure charge neutrality of the complete system, the surface charge density at pH 7.0 and at a corresponding ionic strength amounts to about 0.55 e/nm^2 ¹⁶. This surface charge density is set by deprotonation of randomly chosen silanol terminal groups (cf. ref. 146).

3.2.1 FREE ENERGY OF ADSORPTION FROM RESTMETAD SIMULATIONS

In Chapter 3.1, the basis for an accurate calculations of the free energy of adsorption of polypeptides on solid surfaces by means of RESTMETAD^{190,146} was demonstrated. In brief, firstly the probabilities ρ_a and ρ_s of finding the peptide in an adsorbed or in a dissolved state are computed, respectively, by Boltzmann integration of the one-dimensional free energy profile $F(z_{\text{ssd}})$, where the collective variable z_{ssd} represents the position of the peptide's center of mass in direction perpendicular to the surface. The simulations based on RESTMETAD are performed as discussed in Schneider & Colombi Ciacchi¹⁹⁰, using a set of 7 independent replicas at temperatures ranging from 300 to 450 K with a ΔT of 25 K. The well-tempered METADYNAMICS algorithm acts on the center of mass position of the peptide by adding Gaussian hills with an initial height of 0.7 kcal/mol and a width of 0.1 Å every 0.5 ps to the corresponding bias potential. The free energy of adsorption is calculated according to the formula previously introduced in Eq. (3.5). However, forces are considered in the following and this should not be confused.

$$\Delta F_{\text{ads}} = -k_B T \ln \left(\frac{\rho_a}{\rho_s} \right), \quad (3.10)$$

The evolution of $F(z_{\text{ssd}})$ along a RESTMETAD trajectory of GCRL adsorbing on silica is shown in Fig. 3.17a, and allows us to define the position $z_0 = 20.0 \text{ Å}$ as the border between

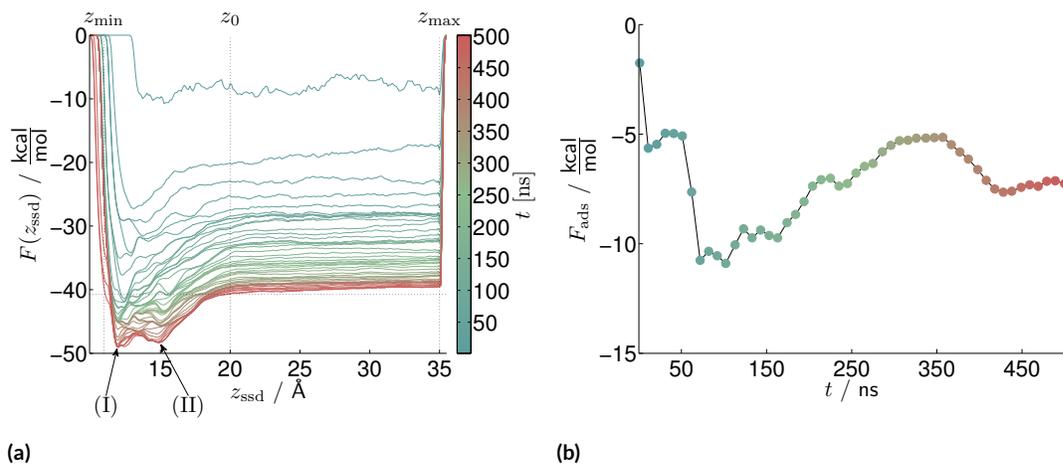


Figure 3.17: (a) Free energy profile of the GCRL peptide adsorbing onto an anionic silica surface as a function of the peptide's center of mass position z_{ssd} in direction perpendicular to the surface, calculated with RESTmetaD. The temporal evolution of the profile is indicated with colors from dark red to blue. (b) Temporal evolution of the free energy of adsorption ΔF_{ads} obtained by Boltzmann integration of the corresponding profile within the limits given in (a), using Eq. (3.5).

the adsorbed state ($z_{\text{min}} < z_{\text{ssd}} < z_0$) and the dissolved state ($z_0 < z_{\text{ssd}} < z_{\text{max}}$), as defined in Fig. 3.17a. While the choice of z_0 is arbitrary, it is important to guarantee that the free energy profile is flat in the dissolved state region, indicating that the peptide does not experience any surface interaction and behaves as in bulk solution. Under this condition, ΔF_{ads} is practically not affected by small changes of the set z_0 value.

The temporal evolution of ΔF_{ads} during the RESTMETAD simulation is shown in Fig. 3.17b. After 500 ns, a final value of -7.3 kcal/mol is reached with an error of about 1.8 kcal/mol, estimated from the fluctuations of ΔF_{ads} in the last 250 ns of simulation. It should be noted that the development of two separate minima (labelled I and II in Fig. 3.17a) in the adsorbed state region takes place only after 350 ns of simulation, pointing towards the importance of long runs to capture essential details of the free energy landscape.

Representative molecular conformations associated with these minima are reported in

Fig. 3.18. In conformation I (Fig. 3.18a,b), all amino acids are in very close contact to the surface, forming hydrogen bonds with the terminal silanol groups. In particular, the positively charged side chain of the arginine residue neighbors a deprotonated silanol. In conformation II (Fig. 3.18c,d), the peptide assumes an upright position, keeping surface contact only via the C-terminus, the leucine and the arginine side chains. Notable is that both polar (and charged) and non-polar amino acid side chains contribute to surface adhesion, as has been previously found in other studies^{50,190,201,146,100}. Also interesting is the fact that the -SH terminal group of cysteine in both cases remains fully hydrated, far from the surface.

3.2.2 ADSORPTION FORCES AND FREE ENERGIES FROM SMD SIMULATIONS

The Steered Molecular Dynamics (SMD) simulations are performed with the help of harmonic potentials of the form

$$V_c = \frac{k_c}{N} \cdot (d - d_0)^2, \quad (3.11)$$

where N is the number of constrained atoms, k_c is the spring constant of the constraint (or virtual AFM cantilever), d is the normal distance of the center of mass of the N atoms to the surface, and d_0 is the variable height of the constraint over the surface, moving at constant speed. 45 random adsorbed configurations of the peptide are obtained by pushing the peptide towards the surface with a speed of 0.01 Å/ps, applying a constraint potential V_c with $k_c = 9.5$ pN/Å on all peptide atoms, until a repulsive force between 500 and 1000 pN is reached. Desorption of the peptide is steered by inverting the constraint velocity direction, using several values of k_c and pulling speeds (vide infra), and applying V_c only to the C atom of the N-terminus of the peptide. Before desorption, the initial position of the harmonic constraint is carefully chosen in order to match the final repulsive force obtained in

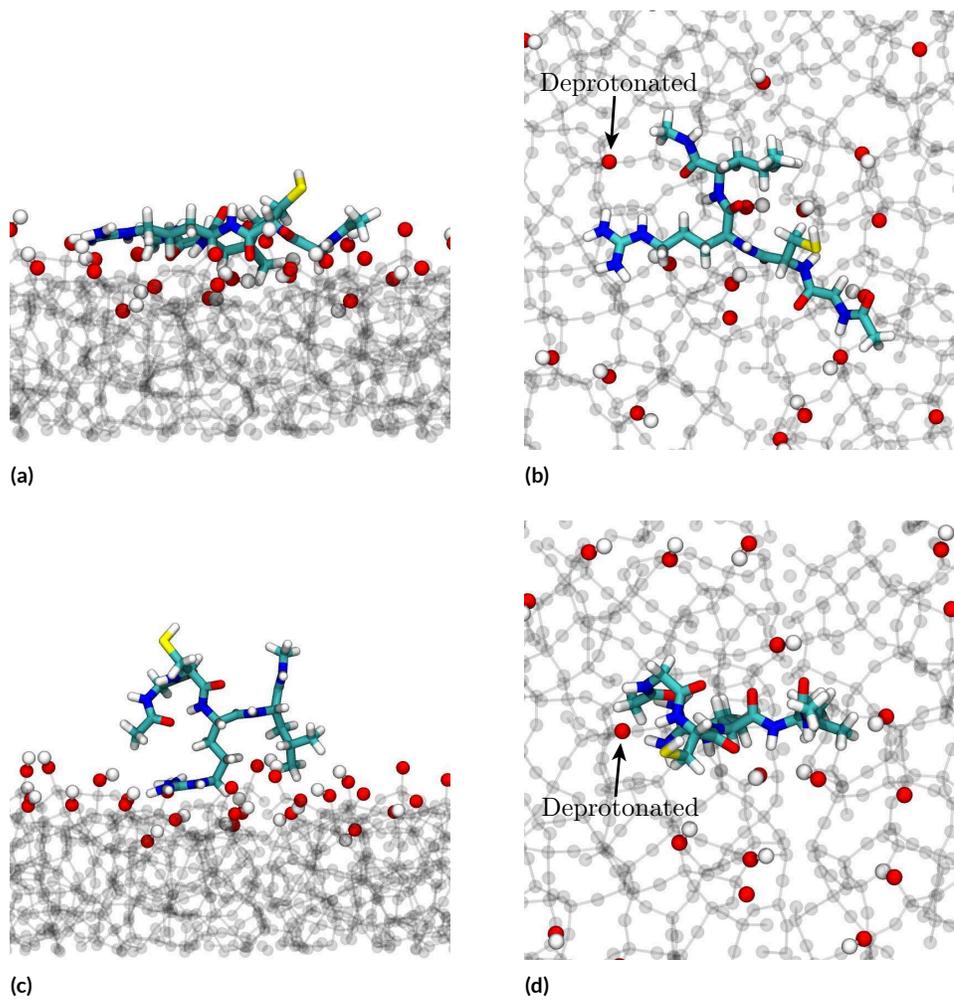


Figure 3.18: Side (a,c) and top (b,d) views of the molecular structures associated with the free energy minima (I) and (II) of Figure 3.17a at $z_{ssd} = 12 \text{ \AA}$ (a,b) and 15 \AA (c,d).

the constrained adsorption. Since the harmonic spring constants of the approach and retraction simulations differ, this requires particular care. This procedure effectively mimics the action of an AFM cantilever functionalized with single peptides.

FREE ENERGY OF ADSORPTION FROM JARZYNSKI'S EQUALITY

In the previous section, the calculation of ΔF_{ads} relies on a complete sampling of the phase space and Boltzmann integration of the free energy profile, which is assumed to describe the thermodynamical equilibrium of the system after reaching convergence. Alternatively, as it is shown in the following, the free energy of adsorption can be estimated from force desorption spectra calculated by out-of-equilibrium steered molecular dynamics (SMD) simulation and application of Jarzynski's equality¹¹³

$$e^{-\Delta F/k_B T} = \overline{e^{-\Delta W/k_B T}}. \quad (3.12)$$

In this equation, ΔF is the free energy difference between two states and ΔW is the work necessary to bring the system from one state to the other. It is important to note that ΔW must be calculated under non-equilibrium conditions, which guarantees correct weighting of individual SMD trajectories¹¹³. To this aim, an extensive set of SMD simulations is carried out pulling the GCRL molecule from an arbitrary adsorbed microstate (generated as described previously) towards a desorbed microstate by means of a harmonic constraint applied to the N-terminal C atom, moving with constant velocity in direction perpendicular to the surface. A total of 810 SMD simulations was performed using three different cantilever spring constants of 69.5, 139.0 and 694.8 pN/Å, and six pulling velocities v_{pull} ranging from 0.001 to 0.5 Å/ps.

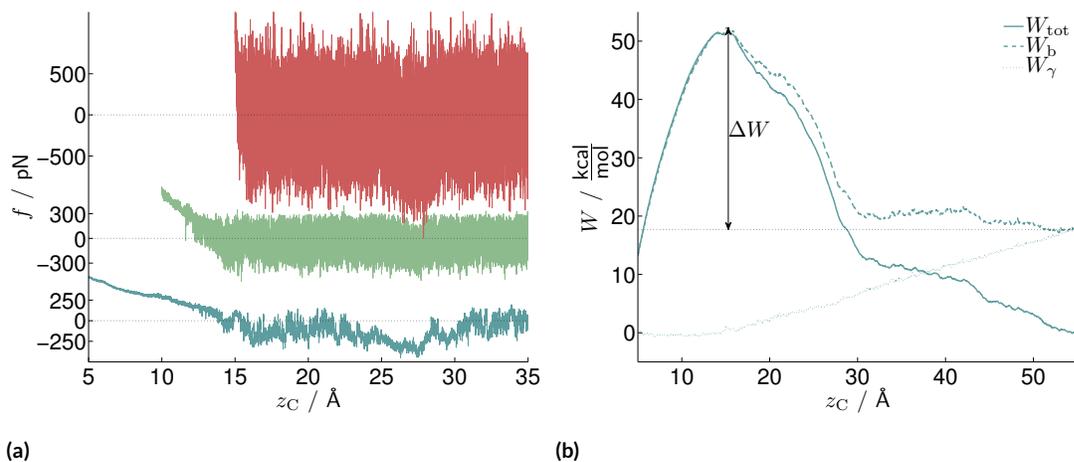


Figure 3.19: (a) SMD retraction curves of the GCRL peptide from the anionic silica surface using pulling speeds of 0.05 (blue), 0.005 (green) and 0.01 Å/ps (red) and spring constants of 69.48, 138.96 and 694.80 pN/Å, respectively. Forces are shifted in y-direction for readability. (b) Cumulative work (solid line) calculated from the blue retraction curve on the left. The frictional and desorption components to the total work are shown as dashed and dashed-dotted lines, respectively. ΔW denotes the desorption work used in Jarzynski's Equality.

Examples of the resulting force-distance curves are reported in Figure 3.19a and show large force fluctuations, scaling with the spring stiffness according to $\sqrt{k_c k_B T}$.⁵ Cumulative numerical integration of these curves leads, nevertheless, to rather smooth work profiles (Fig 3.19b), from which ΔW can be extracted unequivocally. However, especially at high pulling velocity, a frictional contribution to the pulling force due to the motion of the molecule through the viscous water solvent becomes appreciable. This contribution should not enter into Jarzynski's equality if the equilibrium free energy of adsorption is sought for, since the latter is the energy barrier required to detach the molecule from the surface in the limit of zero pulling speeds.

In order to calculate this frictional contribution, SMD simulations of the GCRL peptide are performed in which it is dragged through a periodically repeated box of water with constant velocity v_{drag} using an harmonic spring with stiffness $k_c = 69.5 \text{ pN/Å}$ applied to the

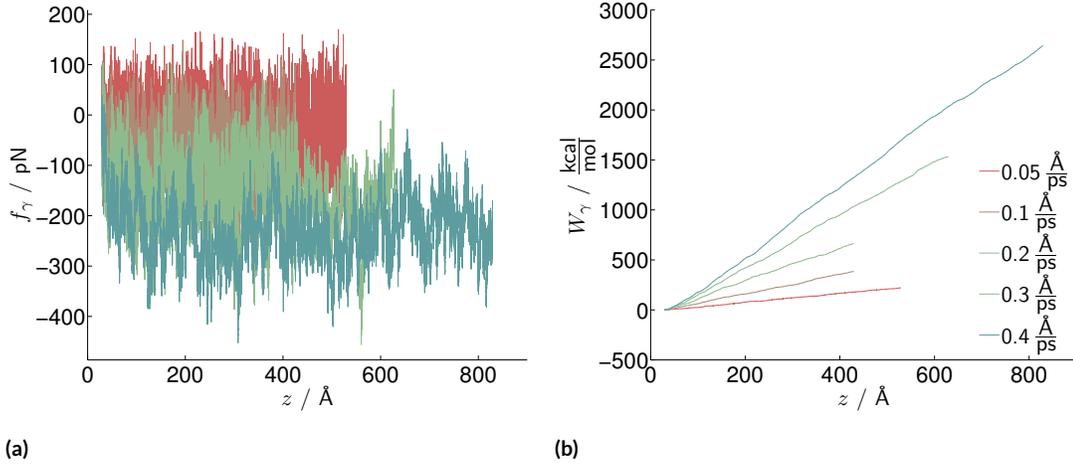


Figure 3.20: Frictional force (a) and work (b) versus path length of the GCRL peptide dragged through Tip3p water with five v_{drag} speeds indicated with different colors.

N terminal C atom. This enables us to perform arbitrarily long simulations without changing the pulling direction, provided that the average velocity of the center of mass of all water molecules is zeroed after each MD step and the system temperature is kept constant via coupling to a Nosé-Hoover thermostat. The frictional coefficient can then be calculated either directly from the obtained average friction force $\langle f_\gamma \rangle$ according to Stoke's law, or via the frictional work $W_\gamma(z) = \frac{1}{z-z_0} \cdot \int_{z_0}^z f_\gamma \cdot dz'$ necessary to drag the peptide from a position z_0 to a position z :

$$\gamma = \frac{\langle f_\gamma \rangle}{v_{\text{drag}}} = \frac{W_\gamma(z)}{\int_{z_0}^z v(z') \cdot dz'}, \quad (3.13)$$

where $v(z)$ is the instantaneous velocity of the C atom to which the harmonic constraint is tethered. The obtained dragging forces and work profiles for v_{drag} values of 0.05, 0.1, 0.2, 0.3 and 0.4 Å/ps are reported in Fig. 3.20, and the calculated values of γ in Table 3.1. If Stoke's law of friction holds, γ should be the same for all different dragging velocities. The variations evident in Table 3.1 are most probably due to the limited simulation times together

$\nu_{\text{drag}} / \frac{\text{\AA}}{\text{ps}}$	$\gamma / \frac{\text{pN}\cdot\text{ps}}{\text{\AA}}$
0.05	612
0.1	670
0.2	579
0.3	593
0.4	576

Table 3.1: Friction coefficient γ for the GCRL peptide dragged through Tip3p water with five ν_{drag} speeds.

with the large force fluctuations, which lead to errors in the average force and average work, especially at low dragging velocities. The molecular friction coefficient γ of the GCRL peptide in water is thus computed by averaging over all dragging velocity and amounts to 606 ± 39 pN ps/ \AA .

With this it is possible to compute the molecular friction contributions W_γ present in the GCRL desorption work profiles by multiplying the velocity of the N-terminus C atom, ν_C , with the molecular friction coefficient γ . This contribution is then subtracted from the work profile W_{tot} to obtain the pure desorption work W_b , from which the desorption barrier ΔW entering into Jarzynski's equality can be obtained (see Fig. 3.19). As expected, the profiles obtained at ν_{pull} below 0.01 $\text{\AA}/\text{ps}$ are hardly affected by the viscous friction, since the corresponding frictional work is smaller than about 0.2 kcal/mol.

To compute the free energy of adsorption, instead of using the pristine Jarzynski's equality (Eq. (3.12)), the approach of Park et al.¹⁶⁴ is followed, which accounts for the effect of finite sampling. Using this approach, the free energy of adsorption is

$$\Delta F_{\text{ads}} = \frac{1}{M} \sum_{i=1}^M \Delta W_i - \frac{1}{2k_B T} \frac{M}{M-1} \left[\frac{1}{M} \sum_{i=1}^M \Delta W_i^2 - \left(\frac{1}{M} \sum_{i=1}^M \Delta W_i \right)^2 \right], \quad (3.14)$$

where M is the number of individual SMD simulations for each k_c and ν_{pull} . The obtained

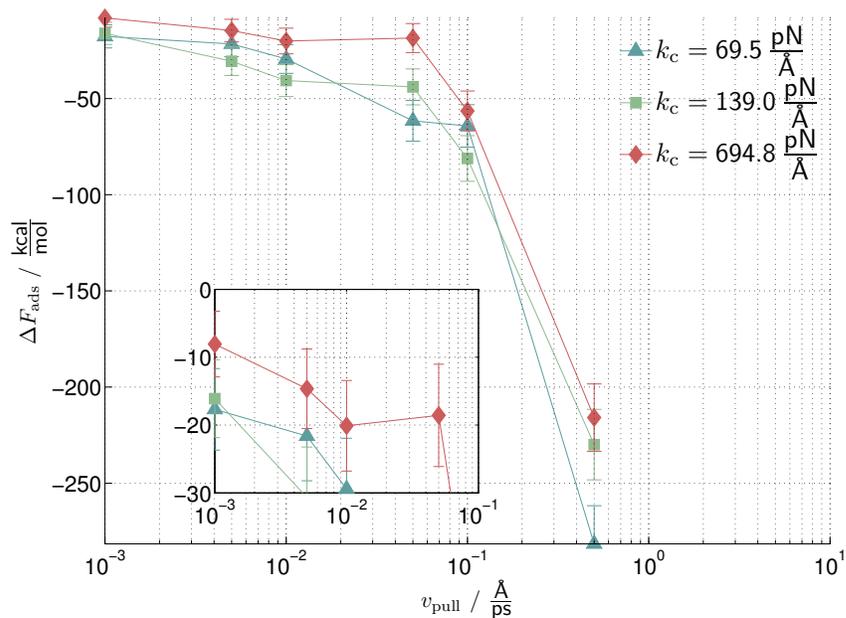


Figure 3.21: Free energy of adsorption ΔF_{ads} obtained with Jarzynski's Equality applied to SMD data using different spring constants and pulling speeds. The inset in (b) magnifies the graphics region for low pulling speeds.

results are shown in Fig. 3.21.

As expected, as v_{pull} decreases, more and more precise estimates of ΔF_{ads} are predicted by Eq. (3.14). Moreover, the convergence is faster for higher k_c , since Jarzynski's equality is exact only in the limit of infinitely stiff springs. In practice, however, too high k_c values are associated with too large fluctuations (see Fig. 3.19), which are a source of error in the estimates of ΔW . For the stiffest harmonic constraint used here, an estimated $\Delta F_{\text{ads}} = -8.0 \pm 4.9$ kcal/mol is obtained, corresponding reasonably well to the value of -7.3 kcal/mol obtained in the RESTMETAD simulation. The error of about 5 kcal/mol is defined by the work fluctuation $\sqrt{\langle \Delta W^2 \rangle - \langle \Delta W \rangle^2}$ in each set of SMD trajectories with the same k_c and v_{pull} , which is often used as a measure of the applicability of Jarzynski's equality^{113,131,109}. This

relatively large error is mostly due to the insufficient sampling, but is comparable with the error bar of the other methods used here (vide infra), and can thus be considered acceptable for the purposes of the present work.

APPLYING THE FRIDDLE⁸¹ MODEL TO SMD FORCE RETRACTION CURVES

Extracting the free energy of adsorption from SMD pulling simulations at different speeds corresponds to performing dynamical force-spectroscopy experiments, typically with an Atomic Force Microscope (AFM). The influence of the loading rate on the desorption forces has been reported previously in several works^{92,138,81}. When an external pulling force is applied to an adsorbed molecule, desorption takes place along a non-equilibrium energy path, which results in a logarithmic dependency of the desorption peak force on the loading rate. The effective loading rate r_{eff} is defined as the product of the pulling velocity v_{pull} and the effective spring constant k_{eff} acting on the surface-molecule bond; $r_{\text{eff}} = v_{\text{pull}} \cdot k_{\text{eff}}$. For soft enough effective spring constants, it follows from the Bell-Evans model that the free energy of adsorption can be obtained from⁸¹

$$\Delta F_{\text{ads}} = \frac{f_{\text{eq}}^2}{2 \cdot \langle k_{\text{eff}} \rangle}, \quad (3.15)$$

where f_{eq} is the limit of the average desorption peak force for zero loading rates. In the following, it is attempted to apply the model of Friddle⁸¹ to the computed SMD data, although that the pulling velocities in the simulations are several order of magnitudes higher than the ones in typical AFM force-spectroscopy experiments.

A problem that is immediately encountered is that the large force fluctuations blur out the force peaks. To overcome this problem an automatic procedure is proposed here to

identify desorption peaks based on a piecewise-linear approximation of the force-distance curves (Fig. 3.22). Firstly, the force-distance curves are smoothed with a moving average filter using a Gaussian window of 1 Å width, and the smoothed force, f_{smooth} , is numerically differentiated with respect to the path length z_C (Fig. 3.22a). The positions of the maxima of the force derivative correspond to the positions at which individual surface-molecule bonds successively break during the pulling process. The original force profile is then cumulatively integrated over z_C , leading to a smooth desorption work profile $W(z_C)$ (Fig. 3.22b). The work profile in each region n between two successive bond breaking events at positions z_n^{low} and z_n^{high} (as identified previously) is nearly harmonic and can thus be least-square fitted by a function

$$W_n(z) = \begin{cases} -k_n \cdot (z - z_n^{\text{off}})^2 + A_n, & \text{for } z_n^{\text{low}} < z < z_n^{\text{high}} \\ 0, & \text{elsewhere} \end{cases} \quad (3.16)$$

The fitting parameter in each harmonic potential region are the effective spring constant k_n , the distance offset z_n^{off} , and the energy offset A_n (see Table 3.2)

A piecewise-harmonic work profile corresponding to the whole desorption process can be now obtained by $W_{\text{fit}}(z_C) = \sum_{i=1}^{\max(n)} W_n(z_C)$ (Fig. 3.22c). Finally, a piecewise-linear force profile (Fig. 3.22d) is obtained upon derivation:

$$f_{\text{fit}}(z_C) = -\frac{d}{dz_C} W_{\text{fit}}(z_C). \quad (3.17)$$

In the example shown in Fig. 3.22, the SMD force-distance curve can be approximated by five linear segments, each corresponding to a bond breaking event, until complete desorp-

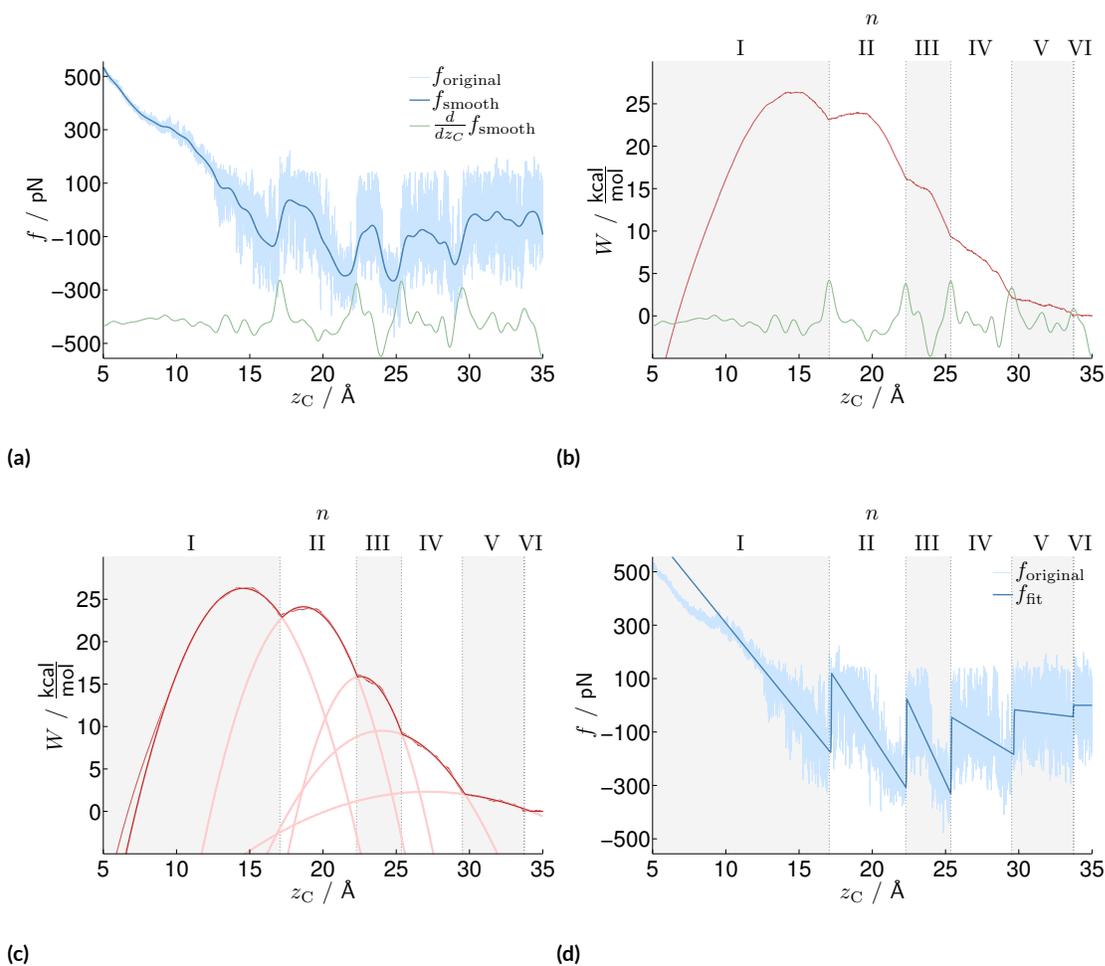


Figure 3.22: Steps of the piecewise-linear fitting procedure applied to a representative SMD force-displacement curve of GCRL desorption (see text). The corresponding fitted parameter are reported in Table 3.2.

tion of the peptide takes place (region VI).

Using this fitting procedure it is now possible to identify the maximum desorption force and its corresponding effective bond stiffness k_{eff} (i.e. the slope of the linear segment preceding the peak) in each of the 810 SMD simulations performed. A mean effective stiffnesses $\langle k_{\text{eff}} \rangle$ for each of the three k_c values (69.5, 139.0 and 694.8 pN/Å) is calculated from

a log-normal distribution fitted to the histogram of all individual k_{eff} values (Figure 3.23a). The maximum forces are plotted as a function of the effective bond loading rates $r_{\text{eff}} = v_{\text{pull}} \cdot k_{\text{eff}}$ in Fig. 3.23b, and fitted separately for each k_c value with the model of Friddle et al.⁸¹. This model provides a functional relationship between peak force and loading rate, having as a fixed input parameter the mean effective stiffness $\langle k_{\text{eff}} \rangle$ and as free parameter the equilibrium force f_{eq} , the transition length x_t and the unbinding constant k_u^0 associated with the surface-molecule bond.

The obtained parameter and confidence intervals for the three fits are summarized in Table 3.3. It should be noted that the values of x_t and k_u^0 may be not physically relevant in this case because of the very high loading rates in the SMD simulations. In fact, the very low value of k_u^0 indicates that rebinding cannot take place, as it should be for a quasi-equilibrium process and assumed in the model of Friddle et al.⁸¹. Nevertheless, x_t is of the typical order of magnitude of values obtained in experimental studies of biomolecular adsorption (see Tables I and SI in ref. 81).

The most important parameter for these purposes is the limit of the peak forces at zero loading rate, f_{eq} , from which the adsorption free energy ΔF_{ads} can be calculated according to Eq. (3.15) (see Table 3.3). While the adsorption free energy should be independent of the in-

n	$z_n^{\text{off}} / \text{\AA}$	$A_n / \frac{\text{kcal}}{\text{mol}}$	$k_n / \frac{\text{pN}}{\text{\AA}}$
I	14.6	26.3	-33.9
II	18.7	24.1	-42.0
III	22.6	15.9	-58.9
IV	24.0	9.5	-16.3
V	27.2	2.3	-3.3

Table 3.2: Fit parameter of harmonic potentials (Eq. (3.16)) in the six regions of the SMD force-displacement curve defined in Figure 3.22.

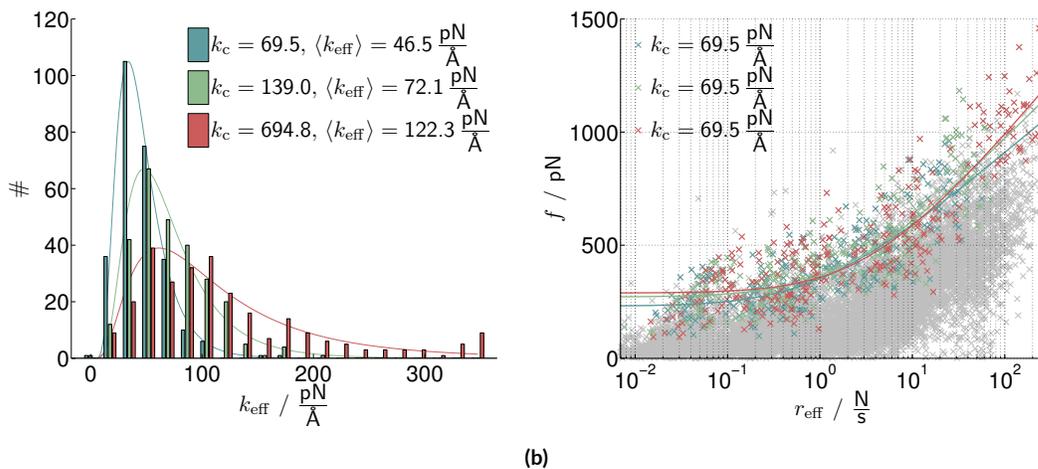


Figure 3.23: (a) Distributions of effective bond spring constants k_{eff} obtained with three different harmonic constraint stiffnesses k_c (colored bars) and their respective mean values obtained from log-normal fits (straight lines). (b) Peak forces from all performed SMD simulations as a function of the corresponding k_{eff} values. The maximum peak forces of each trajectory are indicated in colors, whereas all other identified peaks are indicated with gray crosses. The colored lines represent fits of the maximum forces with the model of Friddle et al.⁸¹. The obtained fitting parameter can be found in Table 3.3.

$k_c / \frac{\text{pN}}{\text{\AA}}$	$\langle k_{\text{eff}} \rangle / \frac{\text{pN}}{\text{\AA}}$	$f_{\text{eq}} / \text{pN}$	$x_t / \text{\AA}$	k_u^0 / ps^{-1}	$\Delta F_{\text{ads}} / \frac{\text{kcal}}{\text{mol}}$
69.5	46.5 ± 32.1	230 ± 14	0.28 ± 0.03	0.0009 ± 0.0004	-8.2 ± 5.7
139.0	72.1 ± 36.2	271 ± 16	0.23 ± 0.03	0.0015 ± 0.0006	-7.3 ± 3.7
694.8	122.3 ± 47.2	288 ± 15	0.19 ± 0.02	0.0029 ± 0.0009	-4.9 ± 2.0

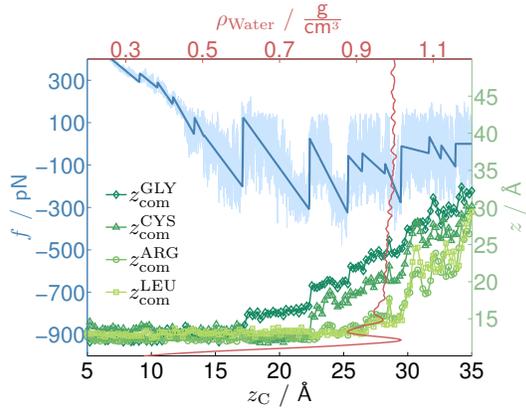
Table 3.3: Parameter and asymptotic standard parameter errors for the model proposed by Friddle et al.⁸¹ fitted to the data in Figure 3.23b, along with the corresponding values of ΔF_{ads} obtained via Eq. (3.15).

dividual cantilever (or effective) spring constant, this is not fully true for the case here. This can be attributed either to the broad scattering of data points in Figure 3.23b or the exceedingly high loading rates, which are not anymore representative of a quasi-equilibrium situation, as noted above. This would also be consistent with the shift of x_t to lower values with increasing cantilever stiffness. Moreover, a further source of uncertainty is the choice of the mean of the k_{eff} distribution, rather than for instance the median or the mode, as the representative value for $\langle k_{\text{eff}} \rangle$. Despite of these facts, the calculated ΔF_{ads} are not too dissim-

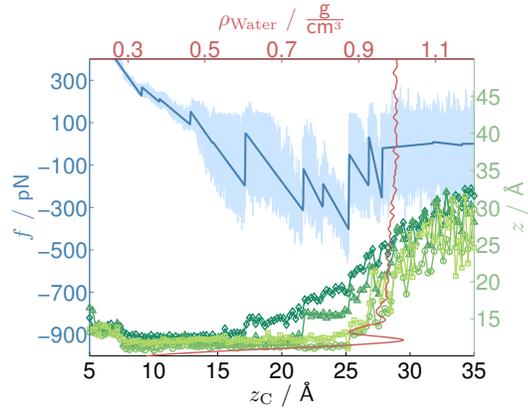
ilar to the values of -7.3 eV and -8.0 eV obtained with the RESTMETAD method and from the Jarzynski's Equality.

ORIGIN OF PEAK FORCES IN THE SMD SIMULATIONS

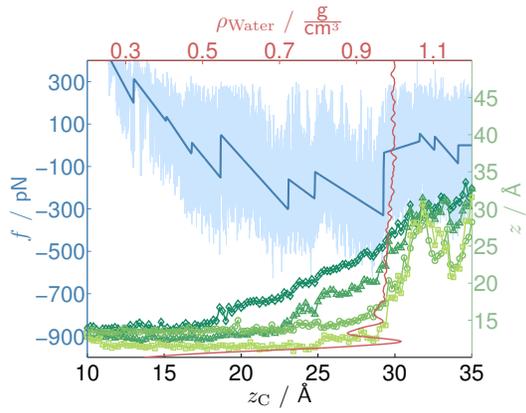
As a direct outcome of the many SMD simulation trajectories, we obtain a clear picture of which surface-molecule interactions contribute to the adsorption and are mainly responsible for the adsorption peak forces. Exemplary SMD force-retraction curves are shown in Figure 3.24 for different k_c and v_{pull} values together with the water density profile near the SiO₂ surface and the evolution of the center of mass positions of single amino acids. Visible is a clear correlation between each peak position and the corresponding detachment of the amino acids from the surface hydration layers. The highest peak forces are mostly observed when the arginine group is released from the first hydration layer close to the surface. However, not only the breakup of the long-range electrostatic interaction between the charged arginine group and deprotonated silanol groups on the surface contribute to adhesion forces. For instance, also the trapping of non-polar residues (here, leucine) in the first water density minimum is an important contribution to the adsorption driving force, as observed in several previous works^{50,190,100,146}. Also to be noted is that not all small peaks resulting from the piecewise linear fitting are attributable to the breaking of surface-molecule interactions, but also to intramolecular rearrangements within the polypeptide during the pulling.



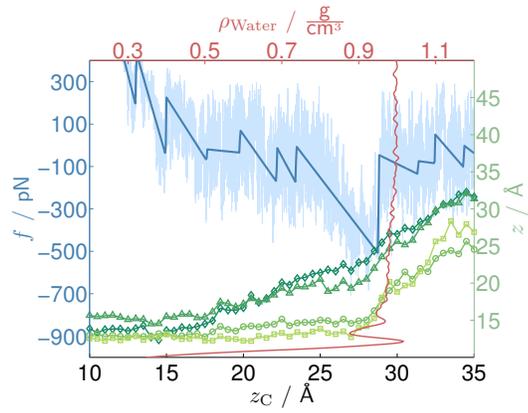
(a) $v_{\text{pull}} = 0.01 \text{ \AA/ps}$



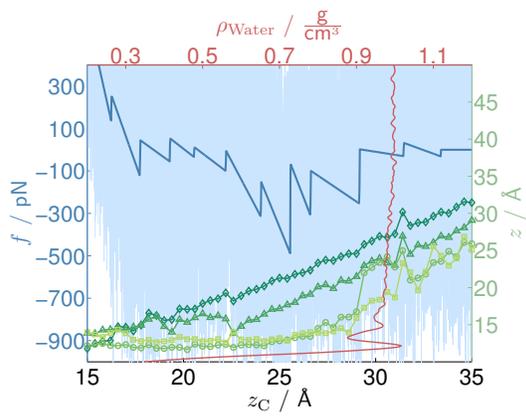
(b) $v_{\text{pull}} = 0.001 \text{ \AA/ps}$



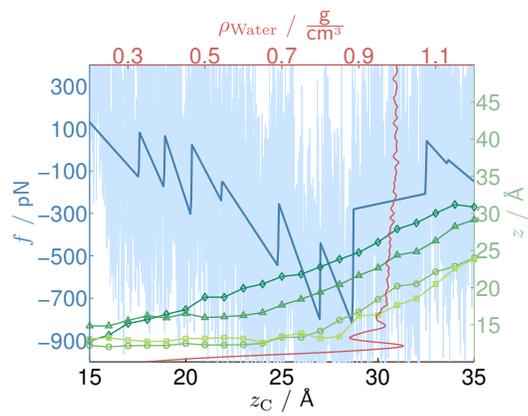
(c) $v_{\text{pull}} = 0.005 \text{ \AA/ps}$



(d) $v_{\text{pull}} = 0.05 \text{ \AA/ps}$



(e) $v_{\text{pull}} = 0.01 \text{ \AA/ps}$



(f) $v_{\text{pull}} = 0.1 \text{ \AA/ps}$

Figure 3.24: Representative SMD force-displacement curves of GCRL desorption (light blue) along with their piecewise-linear fits (blue), superimposed to the evolutions of the center-of-mass positions of selected residues (green) and the water density profile at the anionic silica surface (red). The curves are obtained for $k_c = 694.8 \text{ pN/\AA}$ and different pulling speeds, as indicated.

3.2.3 AFM FORCE SPECTROSCOPY EXPERIMENTS

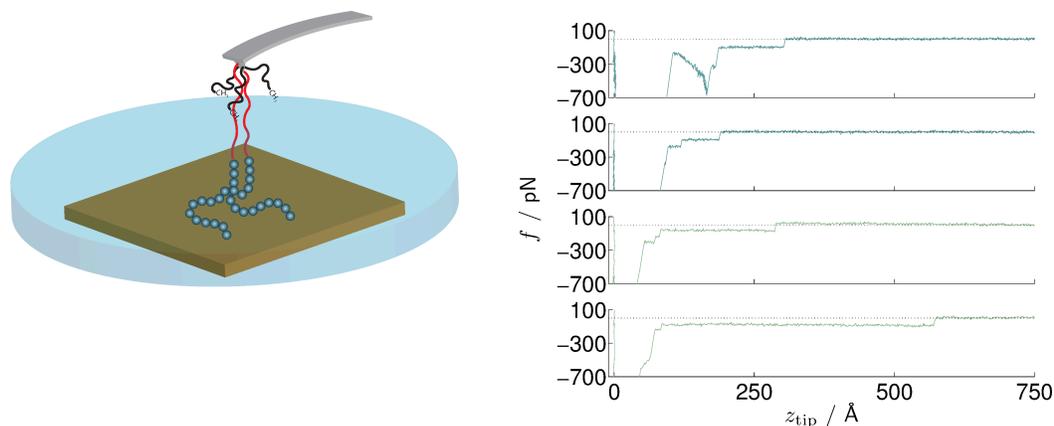
In this section AFM force spectroscopy experiments have been performed to measure the desorption force of the GCRL peptide from amorphous silica in bulk liquid water. To this aim, silicon nitride AFM tips are first covalently functionalized with PEG linker molecules, to which (GCRL)₁₅ polypeptides are attached via standard condensation reactions (see Fig. 3.25a). The force spectroscopy experiments are performed in a liquid cell using a NanoWizard NanoScience atomic force microscope (JPK Instruments AG, Berlin, Germany) with a functionalized Si₃N₄ cantilever (DNP-SiO, Bruker Corporation, France) of spring constant $k_c = 0.42 \pm 0.08$ pN/Å, as determined via its resonance frequency and the equipartition method^{44,78}. Prior to functionalization, the cantilever is cleaned in a freshly prepared Piranha solution for 30 min and washed repeatedly with water and ethanol. It is then immersed into a solution of 3-aminopropyl triethoxysilane (APTES) for 15 minutes and successively incubated in a solution containing a polyethylene glycol n-hydroxysuccinimide ester disulfide (PEG-NHS) and a O-Methyl-O'-[2-(succinylamino)-ethyl]polyethylene glycol N-succinimidyl ester (PEG-Ome) in ratio of 1:20. The PEG-functionalized cantilever is then incubated in a solution containing 0.1 mg/mL (GCRL)₁₅ polypeptides (Selleck Chemicals LLC, Houston, USA), to attach them covalently through their amino terminal to the PEG-NHS linker only. The PEG-Ome linker serves as a spacer to reduce the number of polypeptides tethered to the cantilever tip to a few units, and to reduce the non-specific tip/surface interactions.

Force-displacement curves were collected in ultrapure water after purification with a Mill-Q Integral system against the surface of a fused quartz surface (Hellma Optics GmbH, Jena, Germany) previously cleaned with a Piranha solution and rinsed with abundant

ethanol and water. The roughness of the surface, as determined by AFM imaging, amounted to 0.29 ± 0.01 nm in areas of the order of $2 \times 2 \mu\text{m}^2$. The curves were collected in "force mapping" mode using sets of 16×16 points per each retraction speed value (from 0.05 to 5 $\mu\text{m/s}$), a z-length of 0.4 μm , an extend time of 0.8 s and a delay time on the substrate of 1 s. The reported data are the averages of all curves presenting a clear polypeptide/surface interaction plateau, out of three independent measurement sets. The data were analyzed with the JPK SPM Data processing software (Version 4.3.11).

The functionalized cantilevers are approached to a fused quartz surface until surface contact is established, and are then retracted with constant velocity, leading to force-distance curves which are exemplarily shown in Fig. 3.25b. After an initial non-specific desorption peak arising from the detachment of the PEG-functionalized tip from the surface, a constant force plateau is observed in the majority of the measured curves. This force plateau corresponds to the progressive detachment of individual GCRL units from the surface,^{140,112} and its height is equal to the work of adhesion per unit of length of desorbing polypeptide,¹⁶⁶ under the action of the flexible cantilever (with stiffness k_c) and the elastic PEG linker (with stiffness k_{PEG}).^{162,121,213} In a few cases, plateaus that are much longer than the expected contour length of the combined PEG-(GCRL)₁₅ system are obtained, probably as a result of spurious agglomeration or polycondensation of more than one (GCRL)₁₅ molecule (see Fig. 3.25b, bottommost panel). These cases are discarded from the further analysis.

The goal now is to extract information about the adsorption free energy from the AFM force spectroscopy experiments and, in doing so, to strive a comparison with the simulation results presented in the previous chapter. Estimates of the adsorption free energy will



(a)

(b)

Figure 3.25: (a) Sketch of the AFM experimental setup. Black lines represent the PEG linkers capped by a methanol group and red lines the linkers bound to (GCRL)₁₅ polypeptides, shown as blue-grey beads. The brown surface depicts the flat fused silica surface. (b) Typical force-displacement curves collected at different retraction speeds (green and blue for 0.1 and 0.2 μm/s, respectively). z_{tip} is the tip-sample separation.

be obtained in two ways. First, from the model of Friddle applied to the measured adhesion forces at variable loading rates⁸¹, similarly as done with the SMD simulation results. Second, from the elastic energy contributions stored in the functionalized cantilever, the stretched PEG linker and the desorbed portion of the (GCRL)₁₅ molecule at the point of final detachment of the rest of the adsorbed (GCRL)₁₅ molecule from the surface, according to Krysiak et al.¹²². In both cases, details of the elasticity of the PEG linker and the (GCRL)₁₅ peptide are required, as presented in the next sections.

THE ELASTICITY OF THE PEG LINKER

The stretching stiffness of single PEG molecules is highly non-linear and varies with the externally applied force. The variation of the end-to-end distance $L_{\text{ee}}^{\text{PEG}}$ of a PEG molecule

composed of N_s monomers, upon applying a force f on both ends, can be adequately described using the model proposed by Oosterhelt et al.¹⁶²:

$$L_{ee}^{\text{PEG}}(f) = N_s \cdot \left(\frac{L_{\text{planar}}}{e^{\frac{\Delta F(f)}{k_B T}} + 1} + \frac{L_{\text{helical}}}{e^{-\frac{\Delta F(f)}{k_B T}} + 1} \right) \cdot \left(\coth \left(\frac{f \cdot L_K}{k_B T} \right) - \frac{k_B T}{f \cdot L_K} \right) + N_s \frac{f}{K_s} \quad (3.18)$$

with $\Delta F(f) = (F_{\text{planar}} - F_{\text{helical}}) - f \cdot (L_{\text{planar}} - L_{\text{helical}})$.

This model takes into account the trans-gauche transitions of the PEG backbone through (i) the lengths of the individual trans-trans-gauche and all-trans monomer conformations, L_{helical} and L_{planar} , respectively, and (ii) their associated free energies, F_{helical} and F_{planar} . Further parameter in this model are the Kuhn length of the polymer L_K and the monomer elasticity K_s , whose values have been experimentally determined¹⁶² and are reported in the caption of Fig. 3.26.

This model is applied to a PEG molecule consisting of 18 monomers, which roughly correspond to the length of the used linker (Fig. 3.26). As a result, the spring constant of the PEG linker is estimated at the typical force value of the desorption plateau observed in the force-spectroscopy experiments, $k_{\text{PEG}} = 12.8 \text{ pN}/\text{\AA}$. This lies within the linear force-elongation regime arising from helical unfolding of the molecule¹⁶², where the C-C bonds retain a gauche state typical for PEG dissolved in water¹⁵. It can thus be safely assumed that in this regime (between 70 and 250 pN) the PEG linker behaves like a harmonic spring.

THE ELASTICITY OF THE (GCRL)₁₅ PEPTIDE

The elasticity of the (GCRL)₁₅ polypeptide is determined by means of the WLC model of Bouchiat et al.²⁷ fitted on force-elongation data obtained in a near-equilibrium SMD simu-

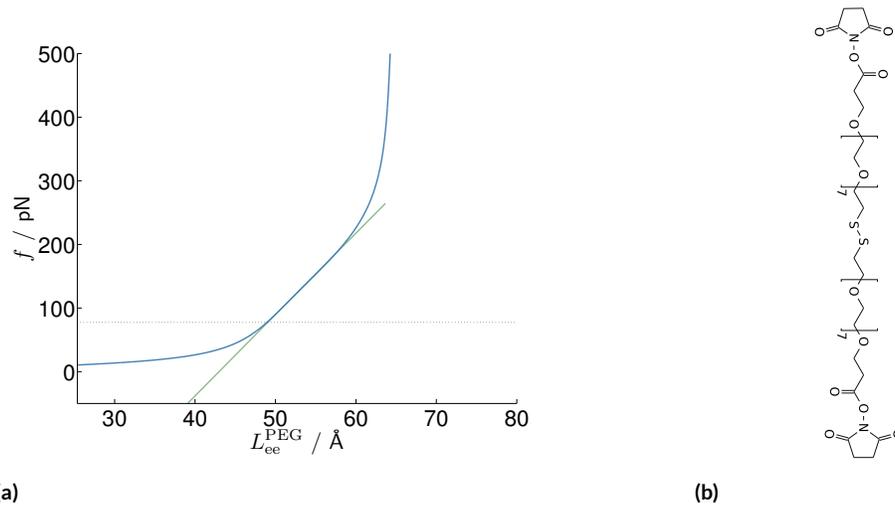


Figure 3.26: (a) Force-extension curve of the PEG linker molecule¹⁶², shown in (b), calculated with Eq. (3.18) (blue). Model parameter are as follows: $N_s = 18$, $K_S = 150 \text{ N/m}$, $G_{\text{planar}} - G_{\text{helical}} = 3 k_B T$, $L_{\text{planar}} = 3.6 \text{ \AA}$, $L_{\text{helical}} = 2.8 \text{ \AA}$, $L_K = 7 \text{ \AA}$. A linear segment (green) is fitted to the supramolecular recognition force regime region to estimate the elasticity of the linker at typical measured desorption forces (dotted line).

lation of a shorter $(\text{GCRL})_5$ polypeptide dissolved in Trp3P water. Namely, the N-terminus and C-terminus are slowly pulled apart at a speed of $0.1 \cdot 10^{-3} \text{ \AA/ps}$ using a harmonic constraint with a spring constant of 20.8 pN/\AA , resulting in an overall simulation time of $2 \mu\text{s}$. The result of this simulation is shown in Figure 3.27 along with the result of the WLC fitting. It should be noted that within the applied WLC model the stretching stiffness of a $(\text{GCRL})_N$ polymer is related to the stiffness of each individual monomer as in a series of Hookean springs: $k_{(\text{GCRL})_N} = k_{\text{GCRL}}/N$. Therefore, from the derivative of the force curve with respect to the end-to-end distance of $(\text{GCRL})_5$ both the stretching stiffness of one (GCRL) monomer, $k_{\text{GCRL}}(f)$, or of $(\text{GCRL})_N$ polymers of arbitrary length, at any given force or extension value are obtained. Moreover, the end-to-end distance of a monomer can be extracted at any given force value, $L_{ee}^{\text{GCRL}}(f) = L_{ee}^{(\text{GCRL})_5}(f)/5$.

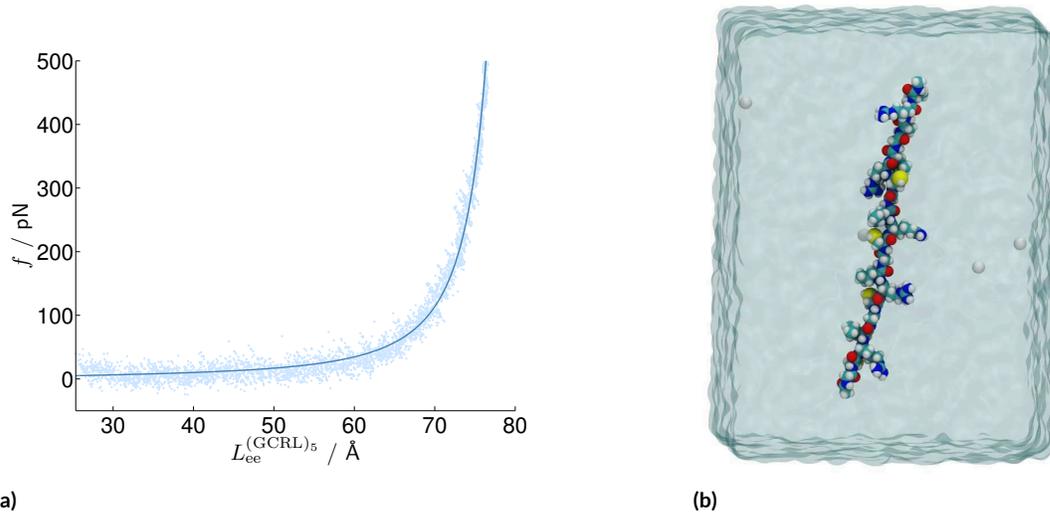


Figure 3.27: (a) WLC model of Marko & Siggia¹⁴³ (green curve) fitted to SMD data (blue dots) of a (GCRL)₅ polypeptide extended in Tip3p water (b). The persistence length and contour length are estimated to 4.15 Å and 82.14 Å, respectively.

FREE ENERGY ESTIMATES FROM SINGLE MOLECULE FORCE SPECTROSCOPY

The average adsorption forces obtained at variable loading rates with the PEG/(GCRL)₁₅ functionalized cantilevers are displayed in Fig. 3.28. Here, the loading rate is computed as the product of the cantilever pulling speed by the effective stretching stiffness of the linker system, $r_{\text{eff}} = v_{\text{pull}} \cdot k_{\text{eff}}$. Since the desorbing force plateau is constant during polypeptide desorption, and the interaction between the PEG molecule and the surface is negligible, the force plateau equals the force required to detach the first GCRL monomer from the surface. For this first detaching event, k_{eff} is determined by the bending stiffness of the cantilever, k_c , the stretching stiffness of the PEG linker, k_{PEG} , and the stiffness of a GCRL monomer, k_{GCRL} :

$$\frac{1}{k_{\text{eff}}} = \frac{1}{k_c} + \frac{1}{k_{\text{PEG}}} + \frac{1}{k_{\text{GCRL}}(f)}. \quad (3.19)$$

In this way, the desorption of the first monomer can be interpreted as a two-state process, for which the model of Friddle can be applied^{93,III}. Indeed, the force spectroscopy data can be well fitted with the Friddle model, using a robust Levenberg-Marquardt algorithm. The extracted parameters are $f_{\text{eq}} = 77.5 \pm 3.0$ pN, $x_t = 0.22 \pm 0.40$ Å, and $k_u^0 = 2589 \pm 3160$ s⁻¹ (Fig. 3.28). Since k_u^0 is the unbinding rate at zero loading rate, consequently $k_{\text{eff}}(f)|_{f_{\text{eq}}}$ is used in the calculations. At this point, from Eq. (3.15) it is possible to estimate an adsorption free energy for GCRL which lies between $\Delta F_{\text{ads}} = -8.8 \pm 2.3$ and -7.4 ± 2.0 kcal/mol, depending on whether the GCRL stiffness is considered or not, respectively. Despite the rough approximation inherent in the application of the Friddle model to this system and in the estimation of the effective linker stiffness, this result is in reasonable agreement with our simulation estimates.

Alternatively, following the considerations of Krysiak et al.¹²², the adsorption free energy is estimated from a balance of the elastic energy stored in the linker/cantilever system and the adsorption energy of the adsorbed portion of the peptides at the moment of the final detachment. This corresponds to the end of the plateau region in the AFM force-displacement curves (see Fig. 3.25). If N_{ads} and N_{des} are the number of adsorbed and desorbed GCRL monomers at the moment of detachment (with $N_{\text{ads}} + N_{\text{des}} = 15$), it can be written:

$$-\Delta F_{\text{ads}} \cdot N_{\text{ads}} = E_c(f_{\text{eq}}) + E_{\text{PEG}}(f_{\text{eq}}) + E_{(\text{GCRL})_{N_{\text{des}}}}(f_{\text{eq}}). \quad (3.20)$$

Here, $E_c(f_{\text{eq}})$ is the elastic energy stored in the bent cantilever, $E_c(f_{\text{eq}}) = f_{\text{eq}}^2 / (2 \cdot k_c) = 1.0$ kcal/mol. The energy stored in the stretched PEG linker is obtained by numerical integration of the inverse function given in Eq. (3.18) up to $L_{\text{ee}}^{\text{PEG}}(f_{\text{eq}})$, yielding $E_{\text{PEG}}(f_{\text{eq}}) = 11.0$ kcal/mol. To calculate the energy stored in the desorbed part of the $(\text{GCRL})_{15}$ polypep-

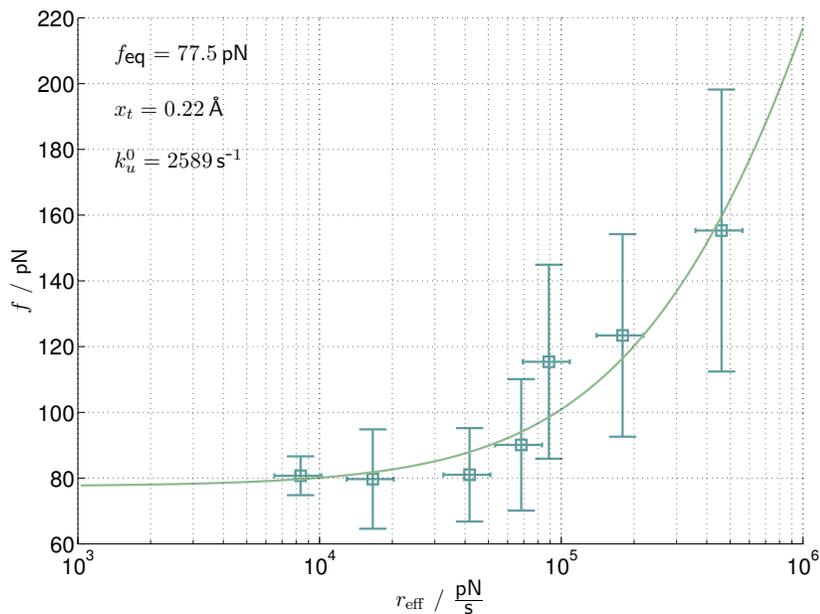


Figure 3.28: Average plateau forces of the (GCRL)₁₅ polypeptide on silica measured by AFM force spectroscopy at several loading rates. The fit with the model of Friddle et al.⁸¹ is shown as a straight line, and the fitting parameter are reported in the inset.

tion, first the amount of desorbed GCRL monomers N_{des} has to be estimated. This is done by subtracting the end-to-end length of the PEG linker at f_{eq} from the average equilibrium plateau length $H_{\text{eq}} = 158 \pm 57 \text{ \AA}$ (Fig. 3.29), and dividing the result by the end-to-end length of a GCRL monomer, $L_{\text{ee}}^{\text{GCRL}} = 13.4 \text{ \AA}$, which leads to $N_{\text{des}}(f_{\text{eq}}) = 8.1 \pm 2.9$. Numerical integration of the WLC model applied to a (GCRL) _{N_{des}} polymer gives $E_{(\text{GCRL})N_{\text{des}}}(f_{\text{eq}}) = 21.7 \pm 10.2 \text{ kcal/mol}$. Insertion of these quantities in Eq. (3.20) with $N_{\text{des}} = 8$ and $N_{\text{ads}} = 7$ finally leads to $\Delta F_{\text{ads}} = -5.1 \pm 4.7 \text{ kcal/mol}$.

It should be considered that the latter estimation of ΔF_{ads} is strongly affected by the broad distribution of plateau end distances (Fig. 3.29), giving an error on the estimation of H_{eq} of the order of a few nm. Moreover, the uncertainty about the precise binding point

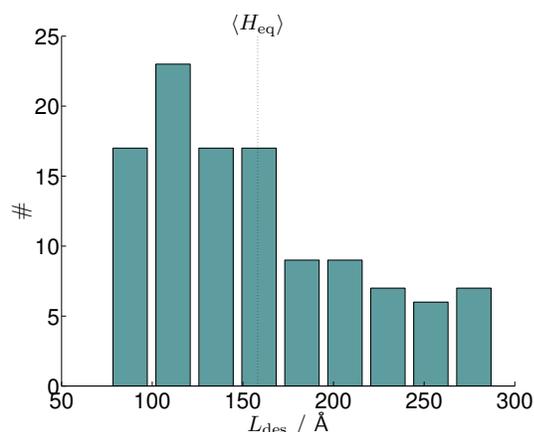


Figure 3.29: Distribution of the plateau lengths for (GCRL)₁₅ desorption from silica, as measured with AFM force spectroscopy at the three smallest loading rate values in Fig. 3.28. The mean plateau length is estimated to $H_{eq} = 158\text{Å}$.

of the PEG linker to the cantilever tip (which could be higher than the tip position) should also be considered, so that the estimate of H_{eq} is actually a lower limit for the real value. This means that so-determined absolute value of ΔF_{ads} is also a lower limit for the true adsorption free energy. Despite of these large uncertainties and rough approximations, however, also in this case the ΔF_{ads} agrees well with the previous estimates of this study, as summarized in Table 3.4.

	RESTmetaD	SMD			AFM	
		$k_c / \frac{\text{pN}}{\text{Å}}$	JE ¹⁶⁴	Friddle ⁸¹	Krysiak ¹²²	Friddle ⁸¹
$\Delta F_{ads} / \frac{\text{kcal}}{\text{mol}}$	-7.3 ± 1.8	69.48	-17.7 ± 6.1	-8.2 ± 7.2	-5.1 ± 4.7	-7.4 ± 2.0
		138.86	-16.1 ± 5.8	-7.3 ± 6.0		-8.8 ± 2.3
		694.80	-8.0 ± 4.9	-4.9 ± 3.6		

Table 3.4: Summarized results of the adsorption free energy obtained with various methods from simulations and experiments.

The truth is rarely pure and never simple

Oscar Wilde: *The Importance of Being Earnest* (1895)

4

Conclusions and Perspectives

IN THE BEGINNING of this thesis I have evaluated several force fields for silica. Among a realistic representation of bulk and surface properties, their ability for the use in combination with well-established biomolecular force fields, e.g., AMBER, CHARMM or OPLS, played a decisive role in their assessment. In this context, particular emphasis has been placed on the

aspect that beyond the perfect crystal surfaces, commonly used in the literature, the present approach includes topological and chemical heterogeneities. In a first attempt to describe such systems a reactive force field (REAXFF)^{222,223,79,177} was chosen, as it was a promising candidate to describe an interfacial system containing dissolved biomolecules with a single force field and furthermore retain the ability of chemical reactions.

It turned out that a complete description of a system containing water, silica and a biomolecule is—for the moment—not feasible with REAXFF due to;

- (i) surface defects emerging upon hydroxylation of a pure SiO₂ phase (cf. Chapter 2.7.1), although at first glance most critical surface features are obtained⁷⁹,
- (ii) the missing compatibility with well-established biomolecular force fields due to their entirely different description of underlying potentials,
- (iii) a lack of parameter for amino acids other than glycine and for the interactions between the solid surface and the nitrogen of the amino acids.

It should be noted that continuous progress in the development of REAXFF towards a complete description of interfacial systems containing dissolved biomolecules is made, yet, in the current state it is not ready to be used with any other REAXFF parameter set available.

A significant step towards a bio-compatible and realistic interfacial force field for silica was made, in cooperation with other researchers, in this work by analyzing the charge distribution upon dissociation of adsorbed water molecules deprotonation of the resulting surface silanol groups. The findings were incorporated into a new potential to describe a natural hydroxylated silica surface in contact with water at neutral pH⁴³. Furthermore, it

has been shown in this work that by adjusting the protonation state of the surface, the solution pH-value could effectively be mimicked without introducing hydrogen ions, e.g., protium, hydronium, oxonium, or Zundel cations.

In order to validate the applicability of this newly developed force field, I calculated the heat of immersion of silica for typical water models used in biological force fields and compared the results to corresponding experimental values found in literature, yielding reasonable agreement.

If surface flexibility is a desired property in simulations, a reliable bulk SiO₂ force field is required. The majority of available silica force fields, however, describe interfacial effects of water without ensuring compatibility to existing biomolecular force fields. Most problematic is the description of electrostatic interactions, e.g., by using point or polarizable charges or by applying a charge equilibration scheme—in almost all cases the resulting descriptions do not fulfill the requirements of charge calculations used for biomolecular force field parameterization. Based on these considerations, I decided to modify the MS-Q force field developed by Demiralp et al.⁶⁶, for the use with our newly developed point charges for silica and to assess its reliability towards biomolecular simulations. Thus, a complete set of force fields for the different compartments of a realistic interfacial biomolecular system is made available for simulations as part of this work.

After its validation, the newly developed force fields have been applied to study experimentally well characterized model systems. In the first case, the secondary structure of a small peptide in bulk solution and adsorbed to colloid silica is well-defined from circular dichroism (CD) spectroscopy^{40,181}. CD spectroscopy is one of the few experimental techniques sensitive to the structural changes that peptides undergo when they adsorb on in-

organic material surfaces, a problem of deep significance in medicine, biotechnology, and materials science. Although the theoretical calculation of the CD spectrum of a molecule in a given conformation can be routinely performed, the inverse problem of extracting atomistic structural details from a measured spectrum is not uniquely determined. Especially complicated is the case of oligopeptides, whose folding/unfolding energy landscapes are often very broad and shallow. This means that the CD spectra measured for either dissolved or adsorbed peptides arise from a multitude of different structures, each present with a probability dictated by their relative free-energy variations, according to Boltzmann statistics. A modeling method based on replica exchange with solute tempering in combination with METADYNAMICS has been presented, which allows us to predict both the helicity loss of a small peptide upon interaction with silica colloids in water and to compute the full CD spectra of the adsorbed and dissolved states. The CD ellipticity Θ for any given wavelength λ is calculated as an external collective variable by means of reweighting the biased trajectory obtained using the peptide–SiO₂ surface distance and the structural helicity as two independent, internal collective variables. Results from this method are in quantitative agreement with experimental measurements (as shown in Figure 3.15a). As an interesting additional result, it was shown that the often-employed hypothesis that the CD intensity Θ at $\lambda = 222$ nm is linearly correlated with the peptides' fractional helicity is correct, although only precisely defined in the case of *macroscopic* conformational states. In the case of single structures, i.e. of *microscopic* states, a linear correlation could also be found (see Figure 3.13), but it is not uniquely defined because of the large variances associated to the different contributions of various secondary structure elements. This means that a large number of individual structures with largely different structural helicities are associated

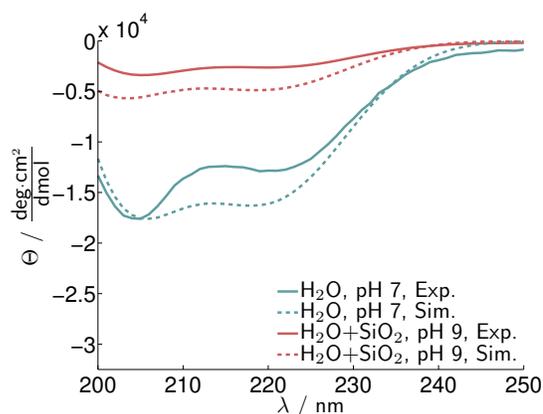


Figure 4.1: Shown are CD spectra of the 4DAR5 peptide in bulk water (blue) and adsorbed on silica (red) at pH 7 and 9 obtained by means of advanced molecular dynamics simulations (dashed lines) and experimental CD spectroscopy (solid lines). The shown experimental data were kindly provided by Nina Würzler (unpublished results, wurzler.nina@gmail.com)

with the same CD ellipticity value. Vice versa, different structures with the same structural helicity can lead to different CD ellipticities. In other words, the *fractional helicity* of a peptide in solution, inferred from CD measurements via Equation (3.7), must be interpreted as an average value of a distribution of structures, each with its own, and different, structural helicity. It should be noted, that a complete CD spectra of the 4DAR5 peptide in aqueous solution at pH 7 was, unfortunately, not published. A further CD analysis of the peptide at neutral pH done in our laboratory revealed, however, a good agreement to theoretically predicted spectra. Results from the experimental CD analysis and theoretical simulations are compared once more in Figure 4.1. To reach acceptable convergence of the free energy landscape, it was necessary to employ seven system replicas and carry out a biased molecular dynamics simulation lasting 1.5 μ s, which represents a very heavy computational effort even for fixed-charges force fields. The application of the method is thus presently confined to short oligopeptides or, possibly, to small proteins with relatively rigid, globular structures.

Combining measured data from atomic force microscopy (AFM) and steered molecular dynamics (SMD) provides a rare opportunity to seamlessly extend the resolution of microscopic experiments towards even smaller, atomic length scales. In this work I have attempted a comparison between different methods for the estimation of the adsorption free energy of short polypeptides at solid/liquid interfaces by means of both Molecular Dynamics (MD) simulations and AFM force spectroscopy experiments. As a model system, the adsorption of the (GCRL) peptide sequence on amorphous silica in contact with water at pH 7 has been considered. This sequence had been identified in a previous experimental work²⁰¹ as a possible (weak) binder for silica, but its surface binding affinity remained uncertain. Indeed, the results of the present investigation suggest a relatively low free energy of adsorption ΔG_{ads} between -5 and -9 kcal/mol, depending on the method used.

Among the employed simulation methods, it is believed that the RESTMETAD^{190,146,45,87} approach is the most suitable method for an accurate estimation of ΔF_{ads} , since it is limited only by the availability of computational resources to perform simulations which are long enough to reach adequate convergence. Obviously, whether the achieved result can be trusted or not is strictly dependent on the accuracy of the employed force field, and this is an issue that will still require further comparative studies with adequately performed experiments, e.g., by using another substrate or peptide. However, extracting accurate adsorption energy values from experimental studies of peptide adsorption is far from trivial, as it relies on the interpretation of rough data through models which may introduce large errors and uncertainties. Dynamical force spectroscopy experiments have been widely used in the past for the investigation of surface-molecule adhesion forces, and several competing models have been developed to extract estimates of the binding free energy from measurements

of adhesion forces. A particularly suitable model is the one of Friddle and De Yoreo⁸¹, in which the analytic expression linking the loading rate on the bond with the force required to break it can be readily fitted to many different non-covalently bonded systems. In order to apply the Friddle model, I have considered the elasticity of the PEG linker molecule and of the GCRL polypeptide itself, and obtained a rough estimate of the free energy of binding that is in reasonable agreement with the RESTMETAD reference (Table 3.4). I have also applied the same model to three sets of simulation data obtained via Steered MD. This has required the development of a procedure to reliably extract force peaks from noisy force-distance curves obtained in a large number of MD trajectories, after subtraction of the friction work arising from pulling the peptide through water at very high speed (see Fig. 3.22). The newly developed procedure gives both information on the peak forces and on the effective spring constant acting on each surface molecule bond, and thus on the effective bond loading rates. Especially applying intermediate harmonic springs to pull the molecule off the surface results in estimates of both the adsorption free energy and of the bond breaking transition length in very good agreement with the corresponding experimental estimates (-7.3 vs -7.4 eV -and 0.23 vs 0.22 Å, respectively). The agreement is impressive given the several orders of magnitude difference in the typical loading rates of the experiments (10^3 to 10^6 pN/s) and simulations (10^{10} to 10^{14} pN/s). However, this is well reflected in the fitted value of the unbinding constant k_u^0 (2.6 ms^{-1} in the experiments, 1.5 ns^{-1} in the simulations), which is dictated by the pulling speed in the latter case. This means that the simulations are actually out of equilibrium, so that the application of the Friddle model shall be performed with care and without giving a strong physical meaning to the fitted kinetic constant values³⁹. More appropriate in this case is the direct use of the Jarzynski's Equality¹⁶⁴

to extract the free energy of adsorption from SMD simulations. This method relies on the stochastic presence of small work values to detach the molecule from the surface, which are essential for a robust free energy estimate using this approach. Therefore, it is not sufficient to use starting adsorption configurations where the peptide is in a few deep local minima, but to use a large number of independent (although realistic) configurations. This was achieved in an ideal way by repeatedly pushing and pulling the probe molecule towards and off the surface with the help of appropriate harmonic constraints, which effectively mimics AFM force-spectroscopy experiments. However, as expected, the convergence of the free energy values with the pulling speed is only very rough (see Fig. 3.21), and a ΔF_{ads} value in reasonable agreement with the RESTMETAD reference is only obtained for an intermediate spring and the lowest pulling speed employed.

Finally, the force-spectroscopy experiments have been also interpreted in terms of an energy balance between the adsorption free energy of the biomolecule and the elastic energy stored in the cantilever system. In this case, the main limitation of the method is the uncertainty about the precise contour length of the PEG linker due to the random attachment position to the AFM cantilever on the one side and the extremely wide distribution of measured detaching lengths on the other side. While an accurate model has been developed to deal with this situation for the case of negligible surface-molecule friction¹²², the model is not directly applicable to the case of a polypeptide in contact with an oxide surface considered here. The fact that a very reasonable estimate of the free energy is obtained also in this case should be considered as almost fortuitous, as indicated by the very large error bar associated with it.

4.1 A GLANCE INTO THE CRYSTAL BALL

CD is a viable technique to investigate and characterize structural changes of peptides or proteins and is, with limitations, applicable to study adsorption phenomena of biomolecules. On this account CD spectroscopy offers the possibility to, e.g., inspect a correct immobilization of proteins, preferably enzymes, and small peptides in terms of conformational changes upon adsorption and the function of active sites. This technique becomes even more powerful when linked to molecular dynamics simulations, providing atomistic insight into the origin of CD spectra of biomolecules. Widely available deconvolution algorithms for CD spectra yielding furthermore information about the amount of secondary structure elements²³⁷. However, they do not allow precise insights into the conformational states in which the system lies. It is believed that shifting the focus from individual secondary structure elements to ensembles (clusters) of structures giving rise to the same CD ellipticity may help in the interpretation of biomolecular CD spectra. Conceivable is the development of cluster-analysis methods based on similar CD spectral shapes over selected wavelength ranges rather than on similar (secondary) structural elements. In particular, this may contribute to the development of empirical relationships between CD ellipticity and other (averaged) secondary structure elements, which is especially useful to understand on an atomic scale the complex conformational changes of biomolecules associated with processes of surface adsorption or biomineralization. In addition, the method presented here is also highly eligible to investigate the somewhat still unclear random coil CD spectra⁸⁸ due to the detailed knowledge of individual CD spectra for each conformational microstate and their associated free energy. Further research on the structure of a zinc binding peptide, typ-

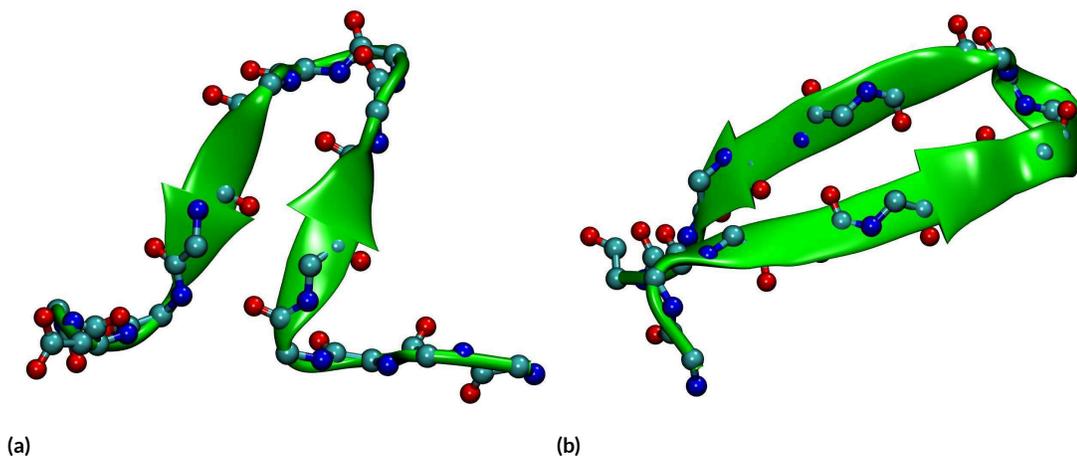


Figure 4.2: Representative minimum free energy conformations of the GB-1 peptide without (a) and with (b) caps. The shown molecular snapshots were kindly provided by Christian Perl (unpublished results, christian.perl@uni-bremen.de)

ically in a random coil conformation, is already initiated, alongside with the validation of this approach to other secondary structural elements. Yet another aspect which is currently addressed from both sides, i.e., from simulation and CD spectroscopy, is the influence of the termination (or capping) of the N- resp. C-terminus in peptides. The GB-1 peptide, a prototypical β -hairpin molecule, was chosen as an ideal study object for this, being well-characterized regarding its stability in the absence of end-caps. The preliminary results of this system appear to be very promising and show a remarkable and surprising behavior. It seems that the capping of the termini with neutral groups leads to a stabilization of the β -hairpin character of the peptide (see Figure 4.2).

Due to the highly versatile implementation of the replica exchange with solute tempering method in LAMMPS it is possible to use alternative replica exchange methods, such as the promising bias-exchange method¹⁷⁰, within this framework. The bias-exchange method is currently tested for its ability to predict minimal free energy conformational structures of

the GB-1 molecule (see Fig.4.2), for a later comparison with CD measurements.

While advanced molecular dynamics simulations can be used to achieve very precise estimates of the surface-molecule adsorption free energy within a given force field, experimental determination of the same quantity by means of dynamic force spectroscopy is still associated with large uncertainties. An alternative approach to assess the adsorption free energy of a single short peptide is by employing a long linker, for which the unspecific peak will not interfere with the adhesion signal of the molecule. Furthermore, available models shall be extended both to take into account strong surface-molecule friction and to remove the explicit dependence of quantities that are not readily experimentally accessible, such as the contour length and elasticity of the linker. The latter could be achieved by including a theoretical model for the linker molecule, e.g., the model of Oesterhelt et al.¹⁶², in the equations of the force-spectroscopy model proposed by Friddle et al.⁸¹ (similar to the formulation of the model proposed by Dudko et al.⁷¹). Furthermore, the model of Friddle et al.⁸¹ for polymeric peptides could be optimized by including the transition energy contributions for a two-dimensional, confined peptide on the surface upon desorption to an increased dimensionality of the peptide in solution, as proposed by Iliafar et al.¹¹¹. In order to make the predictions among experiments and between experiments and simulations more robust, comparisons to other experimental approaches, such as quartz crystal microbalance or isothermal titration calorimetry, which can be performed with short peptide sequences, could also be attempted. However, particular attention should be paid to the functionalization of the tip, in such that only very few functionalized long linker molecules, which are connected to a peptide, are being attached to the tip, whilst the remainder of the tip area is passivated by, e.g., a shorter linker terminated by a methyl group. Aside from this, the

force smoothing algorithm developed within the framework of this thesis already finds application in other fields of material sciences, for instance in the evaluation of forces between TiO_2 nanoparticles^{189,128}.



Exploring Secondary Structures of Peptides with far-UV CD Spectroscopy

Circular Dichroism can be seen as a key technique in the structural characterisation of proteins. Especially for second order structure determination of, e.g., α -helices, β -sheets and to a limited extent random coil structures and double helix structures of nucleic acids.

It is also one of the few experimental techniques sensitive to the structural changes that peptides undergo when they adsorb on inorganic materials surfaces, a problem of deep significance in medicine, biotechnology and materials science. It can be applied to peptide/nanoparticle suspensions at least in the case of colloids in the size-range of a few tens of nanometers, for which light adsorption and scattering do not interfere with the CD signal of the biomolecules^{49,181}.

A.1 ABSORPTION OF CIRCULARLY POLARIZED LIGHT OF CHIRAL MOLECULES

This technique refers to the differential adsorption of left and right circularly polarized light exhibited in the adsorption bands of optically active chiral molecules as schematically shown in Figure A.1a. Of the standard α -amino acids, all but glycine can exist in either of two enantiomers, called L or D amino acids, which are mirror images of each other and have thus differential adsorption of left and right circularly polarized light. This is especially useful in the case of proteins. Simply put, since circularly polarized light itself is “chiral”, it interacts differently with chiral molecules. Circularly polarized light is generated by the superposition of two linearly polarized light beams that are oscillating perpendicular to each other and propagating with a phase difference of $\pi/2$ radians and a wavelength of λ ; The magnitude of the light vector stays the same while its orientation changes by rotating around its propagation direction. If the electric field vector of the light beam forms a right-handed helix upon propagation, it is right circularly polarized light and vice versa. Upon impinging of circularly polarized light on a protein its electronic structure gives rise to characteristic bands in specific regions in the CD spectrum associated with electronic excitation energies. Left and right polarized lights are absorbed differently by chiral molecules due to

different refractive indices for both sources of light. As a consequence left or right polarized light travels at different speeds through the solution of chiral molecules and are adsorbed to different extents at each energy. The molar extinction coefficients are thus different for left and right polarized light, $\epsilon_L \neq \epsilon_R$. This effect is called Circular Dichroism (CD) and by CD spectroscopy the molar differential absorbance,

$$\Delta\epsilon = \epsilon_L - \epsilon_R \quad (\text{A.1})$$

of left and right circularly polarized light is measured for wavelengths λ . Although $\Delta\epsilon$ is typically measured in $M^{-1} \cdot \text{cm}^{-1}$, with M being the molar concentration, for historical reasons most measurements are reported in degrees of molar ellipticity. This is easily interconvertible with the Equation

$$\Theta = 3298.2 \cdot \Delta\epsilon \text{ in } \frac{\text{deg} \cdot \text{cm}^2}{\text{dmol}} . \quad (\text{A.2})$$

Introducing the magnitudes of the electric field vectors of both kinds of circularly polarized light, E_R and E_L , the dependence of the ellipticity θ is given to

$$\tan \theta = \frac{E_R - E_L}{E_R + E_L} . \quad (\text{A.3})$$

It should be noted that $\Theta = \theta / (c \cdot l)$ (in $\text{deg} \cdot \text{cm}^2 / \text{dmol}$) is the molar ellipticity, which is corrected for the molar concentration c (in g/L) and pathlength l (in cm), whereas θ is the uncorrected ellipticity. The situation of two interfering circular polarized lights is shown in Figure A.1b. It should be noted that a path difference stemming from the different refrac-

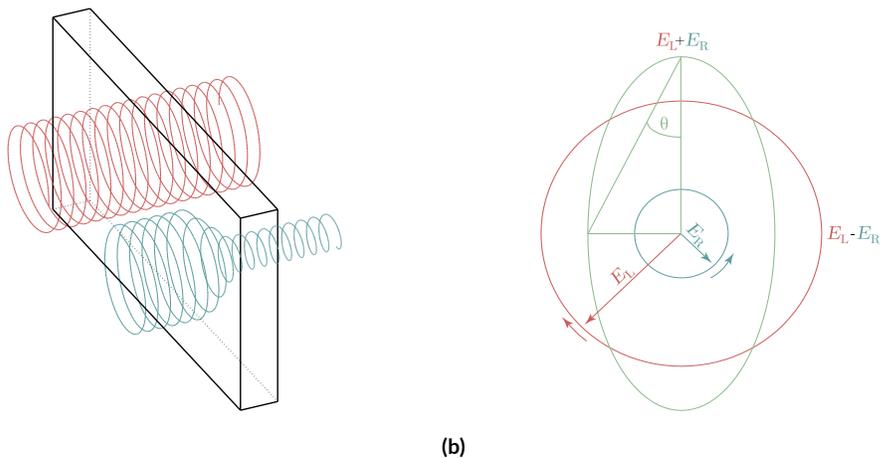


Figure A.1: (a) Schematic Illustration of the functional principle of circular dichroism: Left (red) and right (blue) circularly polarized lights have a differential absorption by chiral molecules. (b) Superposition of left (red) and right (blue) circularly polarized lights and the resulting elliptically deformed circular polarized light (green)

tive indices will tilt the resulting ellipse in Figure A.1b. The two extrema of Equation ((A.3)) are when the magnitude of E_R equals E_L , meaning no difference in the absorbance of right and left circular polarized light. θ will then be 0° and the resulting light beam is linearly polarized. In the other case, when either E_R or E_L is equal to zero, meaning complete absorbance of one of the circularly polarized lights, θ will become 45° and the resulting light beam will be again circularly polarized. In between those two extrema the resulting light will be deformed elliptically characterized by the angle θ .

Circular Dichroism is, among other related methods, such as ORD¹⁹⁸ and LD¹⁸⁵, the most popular type of chiroptical spectroscopy. It can be applied in different energy regimes to excite for example vibrational¹⁵² or electronic transitions³⁸. In this work it is only referred to the latter. On impinging of circular polarized radiation on the protein, the protein's electronic structure gives rise to characteristic bands in specific regions of the CD spectrum^{12,160}, reflecting the electronic excitation energies; α -helices, β -sheets, β -turns and random coil

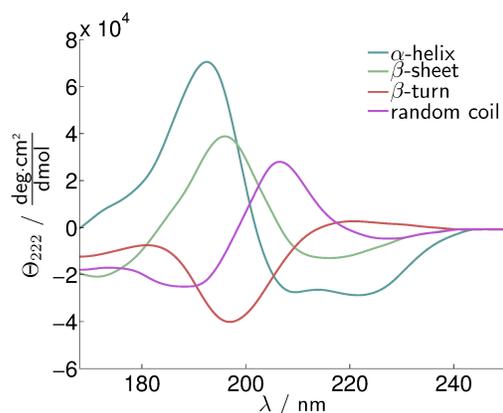


Figure A.2: Expected circular dichroism ellipticities of several secondary protein structures. Corresponding secondary structure types are given in the legend, redrawn after Brahms & Brahms²⁹.

structures possess all a characteristic signal in the far-ultraviolet CD spectrum which shape and magnitude is governed by the associated secondary structure. In Figure A.2 typical spectra for several secondary structure elements are shown.

However, random coil structures are still problematic in interpreting and remain the least well-defined circular dichroism spectral form, as they have a wide range of possible conformational states. Tiffany & Krimm²⁵ stated that it is probable that those structures represents a highly dynamic feature similar to the poly(Pro)II (PPII) helix. Although this statement was proposed in the seventies, there is still a lot ambiguity about this topic⁸⁸.

A.2 ELECTRONIC EXCITATION STATES OF AN α -HELIX IN PROTEINS

As in this work the α -helix is of most interest, the reasons for its characteristic band structure in the far-ultraviolet Circular Dichroism spectrum is explained by reference to Figure A.3 in the following. The spectrum possess an intense peak at 190 nm and a negative band at around 208 nm arising from the exciton splitting of electronic transitions from the

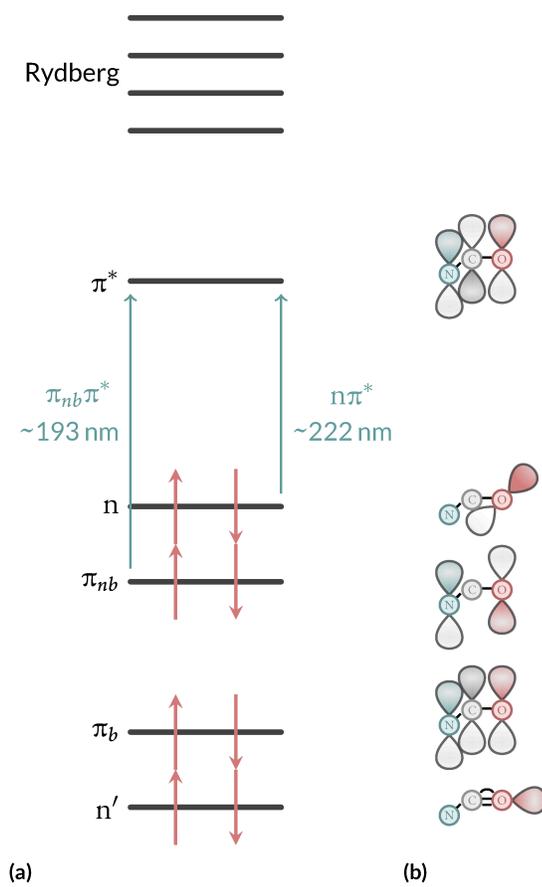


Figure A.3: Electronic transitions of the amide group in the far-ultraviolet region (a). Molecular orbitals shown are the bonding, nonbonding and antibonding π orbitals (π_b, π_{nb} and π^*) and the two lone pairs on the oxygen atom (n and n') (b), redrawn from Bulheller et al.³⁸

amide non-bonding π orbital, π_{nb} , to the anti-bonding π orbital, π^* ³⁸. Another negative band is located at about 222 nm, arising from the electronic transition from an oxygen lone pair orbital, n , to the π^* orbital. Each transition in the protein CD spectrum is characterized by an energy and a rotational strength, given by the Rosenfeld Equation¹⁸⁶.

The rotational strength is experimentally determined by integrating over the intensity $\Delta\epsilon$ of the CD band²⁰⁰

$$R = \frac{hc\pi^3 N_A}{32} \int \frac{\Delta\epsilon}{\lambda} d\lambda. \quad (\text{A.4})$$

A transition in the protein CD spectrum is quantum mechanically treated as an electronically excited oscillating dipole (a linear displaced charge) induced by the electric transition dipole moment, $\boldsymbol{\mu}$, and a circulation of charge due to the magnetic transition dipole moment \mathbf{m} . The probability of a transition from the ground state $0 \rightarrow k$ to occur for oscillating a charge is represented by the integral $\langle \Psi_k | \boldsymbol{\mu} | \Psi_0 \rangle$. Consequently the probability for a transition of a circulation of charge is accordingly $\langle \Psi_k | \mathbf{m} | \Psi_0 \rangle$. The brackets in the Equations represents the standard Dirac notation widespread used for describing quantum states. The combination of both probabilities describes thus a helical displacement of a charge, however, $\boldsymbol{\mu}$ is a real operator whereas \mathbf{m} is imaginary as it describes the rotation of charge in a complex coordinate system⁴⁶:

$$\mathbf{m} = \frac{e}{2mc} (\mathbf{r} \times \mathbf{p}), \quad \mathbf{p} = \frac{\hbar}{i} \nabla, \quad (\text{A.5})$$

where e is the charge, m the mass of the electron, c the speed of light, \mathbf{r} the position operator of the electron, \hbar is Planck's constant divided by 2π , $i = \sqrt{-1}$, ∇ is the gradient operator and \mathbf{p} is the linear momentum operator. The rotational strength, R_k , of a transition from

the ground state 0 to an electronically excited state k is the product of the electric transition dipole moment, $\boldsymbol{\mu}$, and the magnetic transition dipole moment, \mathbf{m} :

$$R_k = \Im (\langle \Psi_0 | \boldsymbol{\mu} | \Psi_k \rangle \langle \Psi_k | \mathbf{m} | \Psi_0 \rangle) , \quad (\text{A.6})$$

where Ψ_0 and Ψ_k denote the wave functions of the ground state and the excited state, respectively, and \Im the imaginary part of the product. Equation (A.6) is the famous Rosenfeld Equation¹⁸⁶. This theoretical approach will yield a line spectrum, that is, values for the rotational strength and a wavelength for each transition. A connection to the experimental spectrum can be made by considering that every line in the calculated spectrum represents an integrated CD band area of the molar ellipticity $\Delta\epsilon_k$ over λ for distinct transitions K (cf. (A.4)). In an experimental CD spectrum, molar ellipticity is broadened due to the uncertainty principle, unresolved vibronic components, and the interactions of protein chromophores with their environment, including other chromophores and the solvent³⁶. An overlay of approximately Gaussian-shaped bands is observed. Thus, for a Gaussian band with $\Delta\epsilon_{k,\max}$ as the peak intensity, λ_k as the band center and σ_k as half-width at 1/e of the peak height, the intensity distribution is given as:¹⁷²

$$\Delta\epsilon_k(\lambda) = \Delta\epsilon_{k,\max} \cdot \exp\left(-\frac{\lambda - \lambda_k}{\sigma_k}\right)^2 . \quad (\text{A.7})$$

Gaussian bandshapes give better results than Lorentzian curves¹². Substitution of Equation (A.7) in Equation (A.4) gives an expression for the peak band intensity $\Delta\epsilon_k$ of a transition k ²⁰⁰

$$\Delta\epsilon_k = \frac{2.278 \lambda_k R_k}{\sigma_k} \quad (\text{A.8})$$

A typical value for σ_k found in literature consistently used for all transitions is 12.5 nm^{38} . Using the peak intensity $\Delta\epsilon_{k,\text{max}}$ and bandwidth σ_k the CD spectral intensity, at any wavelength, can be simulated with a Gaussian intensity profile. In practice, the intensity distribution from all transitions will have to be summed to obtain the intensity at a given wavelength, due to the broadening of the rotational strength of electronic transitions.

A.3 SOLVING THE ROSENFELD EQUATION WITH THE MATRIX METHOD

The following chapter follows closely the explanation given in the dissertation of Buller³⁶. In the matrix method^{241,239,13}, the protein is split into M independent chromophores, with a monomer wave function Φ_{is} for each chromophoric group i and electronic state s . The proteins k^{th} excited state wavefunction, Ψ_k , is assumed to be, in a first order approximation, the linear combination of electronic configurations

$$\Phi_{i,a} = \varphi_{1,0} \cdots \varphi_{i,a} \cdots \varphi_{j,0} \cdots \varphi_{M,0}, \quad (\text{A.9})$$

in which only one chromophoric group, i , is in an excited state, a , and the others are in the ground state 0^{200} . The proteins wavefunction is thus composed of

$$\Psi_k = \sum_i^M \sum_a^{n_i} c_{i,a}^k \Phi_{i,a}. \quad (\text{A.10})$$

Consequently, the protein ground state is defined as

$$\Psi_0 = \varphi_{1,0} \cdots \varphi_{i,0} \cdots \varphi_{j,0} \cdots \varphi_{M,0} \quad (\text{A.11})$$

In order to solve the Rosenfeld Equation¹⁸⁶ the expansion coefficients $c_{i,a}^k$ for each state, k , have to be determined. For the simple case of a dipeptide with two excited states on each amino acid, the lowest energy backbone $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions this is illustrated in the following. Ψ_k is then written as

$$\Psi_k = c_{1,n\pi^*}^k \Phi_{1,\pi\pi^*} + c_{1,\pi\pi^*}^k \Phi_{2,n\pi^*} + c_{2,n\pi^*}^k \Phi_{2,n\pi^*} + c_{2,\pi\pi^*}^k \Phi_{1,\pi\pi^*} \quad (\text{A.12})$$

$$\begin{aligned} \Phi_{1,n\pi^*} &= \varphi_{1,n\pi^*} \cdot \varphi_{2,0} \\ \text{where } \Phi_{1,\pi\pi^*} &= \varphi_{1,\pi\pi^*} \cdot \varphi_{2,0} \\ \Phi_{2,n\pi^*} &= \varphi_{1,0} \cdot \varphi_{2,n\pi^*} \\ \Phi_{2,\pi\pi^*} &= \varphi_{1,0} \cdot \varphi_{2,\pi\pi^*} \end{aligned} \quad (\text{A.13})$$

The corresponding energies of the four excited states ($k = 1 \dots 4$) can be calculated by solving the Schrödinger Equation

$$\mathcal{H}\Psi_k = E_k\Psi_k. \quad (\text{A.14})$$

The Hamiltonian is constructed by separating it into a contribution from M local chromophores, independent from each other, \mathcal{H}_i , and a sum of all intergroup potentials V_{ij} of the entire molecule:

$$\mathcal{H} = \underbrace{\sum_{i=1}^M \mathcal{H}_i}_{\mathcal{H}_0} + \underbrace{\sum_{i=1}^{M-1} \sum_{j=i+1}^M V_{ij}}_V. \quad (\text{A.15})$$

Combining Equation (A.10),(A.14) and (A.15) gives an Equation in terms of chromophore wavefunctions and the coefficients $c_{i,a}^k$. It is convenient to convert this into a matrix formalism. The Hamiltonian matrix is diagonalized by a unitary transformation using the matrix U :

$$U^{-1} \cdot \mathcal{H} \cdot U = \mathcal{H}_{\text{diag}}. \quad (\text{A.16})$$

Diagonal elements of the resulting matrix $\mathcal{H}_{\text{diag}}$ are the excited state energies (the ground state is set to be zero) and the eigenvectors, $c_{i,a}^k$, form the unitary matrix U . Electric and magnetic transitions of non-interacting chromophores, $\boldsymbol{\mu}_a^0$ and \mathbf{m}_a^0 , respectively, can be transformed into the interacting system using the unitary matrix from Equation (A.16) Sreerama & Woody²⁰⁰:

$$\begin{aligned} \boldsymbol{\mu}_i &= \sum_a U_{a,i} \boldsymbol{\mu}_a^0, \\ \mathbf{m}_i &= \sum_a U_{a,i} \mathbf{m}_a^0. \end{aligned} \quad (\text{A.17})$$

The rotational strength is readily calculated using this approach (cf. Equation (A.6)). For example, considering again the dipeptide previously introduced, the Hamiltonian matrix constructed using the Hamiltonian of Equation (A.15) has the form

$$\mathcal{H} = \begin{pmatrix} E_{1,n\pi^*} & V_{1,n\pi^*;1,\pi\pi^*} & V_{1,n\pi^*;2,n\pi^*} & V_{1,n\pi^*;2,\pi\pi^*} \\ V_{1,n\pi^*;1,\pi\pi^*} & E_{1,\pi\pi^*} & V_{1,n\pi^*;2,\pi\pi^*} & V_{1,\pi\pi^*;2,\pi\pi^*} \\ V_{1,n\pi^*;2,n\pi^*} & V_{1,\pi\pi^*;2,n\pi^*} & E_{2,n\pi^*} & V_{2,n\pi^*;2,\pi\pi^*} \\ V_{1,n\pi^*;2,\pi\pi^*} & V_{1,\pi\pi^*;2,\pi\pi^*} & V_{2,n\pi^*;1,\pi\pi^*} & E_{2,\pi\pi^*} \end{pmatrix}, \quad (\text{A.18})$$

where $V_{i,a;i,b}$ denotes interactions between states on the same group and $V_{i,a;j,b}$ are interactions between different groups. The off-diagonal elements, $V_{i,a;i,b}$ and $V_{i,a;j,b}$, are the reason

for the dependence of CD spectra on the proteins' secondary (and tertiary) structure. A transition from $0 \rightarrow a$ on group i and $0 \rightarrow b$ on group j has the form

$$V_{i,0a;j,0b} = \int_i \int_j \Phi_{i,0} \Phi_{i,a} V_{i,j} \Phi_{j,0} \Phi_{j,b} d\tau_i d\tau_j . \quad (\text{A.19})$$

If the interaction between chromophores is regarded as an electrostatic interaction between charge densities the matrix elements could be evaluated exactly from the monomer wave functions using integral evaluation routines available in many quantum chemical simulation packages. It should be noted that in the matrix method orbitals on different chromophores are assumed not to overlap so that no inter-chromophore charge-transfer occurs. However, to make calculations computational tractable the *monopole-monopole approximation* is used. Equation (A.19) is thus re-casted as a sum of Coulomb interactions²¹⁶:

$$V_{i,0a;j,0b} = \sum_{s=1}^{N_s} \sum_{t=1}^{N_t} \frac{q_s q_t}{r_{st}} , \quad (\text{A.20})$$

where q_s and q_t are the point charges on chromophores i and j , and N_s and N_t are the total number of charges on the chromophore. At its heart, the matrix method calculations depends on the magnitudes and locations of the monopoles^{241,239}. The dependence on monopoles and their distance as well as their orientation makes the CD calculation independent on ad-hoc definitions of the secondary structure. It has been shown that this method is fairly successful in calculating the CD spectra of proteins. Furthermore, proteins with a high amount of α -helicity can be calculated almost quantitatively¹⁰¹.

Implementing this algorithm as a computer program is straightforward and has been done in the Dichrocalc algorithm. According to the above formulation the only informa-

tions necessary are the atomic coordinates and the position of the chromophores in the protein. The electrostatic potential of each chromophore is then parameterized by a set of monopoles and consequently for every group in the protein the respective set of monopoles is superposed on the chromophore's atoms. The Hamiltonian matrix is hence constructed by calculating the interaction between all the different electronic excitations and the rotational strengths R_k are determined.

B

Physical and Theoretical Concepts in Molecular Dynamics Simulations

B.1 FORCE FIELD DEFINITIONS

In the following the functional form and according parameter of force fields used within the framework of this thesis are given. If fixed point charges are used they are given in terms

of multiples of the elementary charge e .

B.1.1 VAN BEEST ET AL.²²¹ FORCE FIELD DEFINITION

$$V(r_{ij}) = A_{ij}e^{-b_{ij}r_{ij}} - \frac{c_{ij}}{r_{ij}^6} + \frac{q_i q_j}{4\pi\epsilon_p r_{ij}}, \quad \text{with } r_{ij} = |\mathbf{r}_i - \mathbf{r}_j| \quad (\text{B.1})$$

$i - j$	A_{ij} ($\frac{\text{kcal}}{\text{mol}}$)	b_{ij} (\AA^{-1})	c_{ij} ($\frac{\text{kcal}}{\text{mol}} \text{\AA}^6$)	q_i
O-O	32025.8580	2.76000	4035.5948	-1.2
Si-O	415176.3981	4.87318	3079.4610	2.4
Al-O	369165.4813	4.79667	3010.9204	1.4
P-O	208333.7325	5.19098	458.4274	3.4

Table B.1: van Beest et al.²²¹ force field parameter

B.1.2 VASHISHTA ET AL.²²⁶ FORCE FIELD PARAMETER

$$V(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_{1 \leq i < j \leq N} V_2(r_{ij}) + \sum_{1 \leq i < j < k \leq N} V_3(r_{ij}, r_{jk}, r_{ik}) \quad (\text{B.2})$$

$$V_2(r_{ij}) = \frac{e^2}{4\pi\epsilon_p} \left(\frac{H_{ij}}{r_{ij}^{\eta_{ij}}} + \frac{q_i q_j}{r} - \frac{\frac{1}{2}(\alpha_i q_j^2 + \alpha_j q_i^2)}{r^4} e^{-r/r_{4s}} \right) \quad (\text{B.3})$$

$$V_3(r_{ij}, r_{jk}, r_{ik}) = B_{jik} f(r_{ij}, r_{ik}) p(\Theta_{jik}, \bar{\Theta}_{jik}) \quad (\text{B.4})$$

$$f(r_{ij}, r_{jk}) = \begin{cases} \exp\left(\frac{l}{r_{ij}-r_0} + \frac{l}{r_{jk}-r_0}\right) & \text{for } r_{ij}, r_{jk} \leq r_0 \\ 0 & \text{for } r_{ij}, r_{jk} > r_0 \end{cases} \quad (\text{B.5})$$

$$p(\Theta_{jik}, \bar{\Theta}_{jik}) = (\cos \Theta_{jik} - \cos \bar{\Theta}_{jik})^2 \quad (\text{B.6})$$

$i - j$	η_{ij}	$H_{ij} (\frac{\text{kcal}}{\text{mol}} \cdot (\text{\AA})^{\eta_{ij}})$	$r_{4s} (\text{\AA})$	q_i	α_i
Si-Si	11	18.9150	4.43	1.6	0.0
Si-O	9	3778.6757	4.43	1.6	0.0
O-O	7	17153.5352	4.43	-0.8	2.4

Table B.2: Vashishta et al.²²⁶ two-body force field parameter

$j - i - k$	$B_{jik} (\frac{\text{kcal}}{\text{mol}})$	l	$\bar{\Theta}_{jik} (\text{deg})$	$r_0 (\text{\AA})$
Si-O-Si	464.5777	1.0	141.00	2.6
O-Si-O	116.1444	1.0	109.47	2.6

Table B.3: Vashishta et al.²²⁶ three-body force field parameter

B.1.3 A CLOSER LOOK ON REAXFF

An uncorrected bond order is derived from Equation (2.47). Necessary for this is the uncorrected overcoordination, defined as the difference between the bonding electrons around the atom and its total bond order

$$\Delta'_i = -\text{Val}_i + \sum_{j=1}^{\text{neighs}(i)} \text{BO}'_{ij} \quad . \quad (\text{B.7})$$

This is used to calculate the corrected bond order BO_{ij} with the scheme pointed out in Equation (B.9). In order to softening corrections of atoms bearing lone electron pairs, a second overcoordination definition Δ_i^{boc} is introduced

$$\Delta_i^{\text{boc}} = -\text{Val}_i^{\text{boc}} + \sum_{j=1}^{\text{neighs}(i)} \text{BO}'_{ij} \quad , \quad (\text{B.8})$$

and used in Equation (B.9). This allows nitrogen and oxygen, which bear lone pair electrons after they filled up their valence, to break up these lone pair electrons and let them involve into bonding without obtaining a full bond order correction.

$$\begin{aligned}
\text{BO}_{ij}^{\sigma} &= \text{BO}_{ij}'^{\sigma} \cdot f_1(\Delta'_i, \Delta'_j) \cdot f_4(\Delta'_i, \text{BO}'_{ij}) \cdot f_5(\Delta'_j, \text{BO}'_{ij}) \\
\text{BO}_{ij}^{\pi} &= \text{BO}_{ij}'^{\pi} \cdot f_1(\Delta'_i, \Delta'_j) \cdot f_1(\Delta'_i, \Delta'_j) \cdot f_4(\Delta'_i, \text{BO}'_{ij}) \cdot f_5(\Delta'_j, \text{BO}'_{ij}) \\
\text{BO}_{ij}^{\pi\pi} &= \text{BO}_{ij}'^{\pi\pi} \cdot f_1(\Delta'_i, \Delta'_j) \cdot f_1(\Delta'_i, \Delta'_j) \cdot f_4(\Delta'_i, \text{BO}'_{ij}) \cdot f_5(\Delta'_j, \text{BO}'_{ij}) \\
\text{BO}_{ij} &= \text{BO}_{ij}^{\sigma} + \text{BO}_{ij}^{\pi} + \text{BO}_{ij}^{\pi\pi}.
\end{aligned} \tag{B.9}$$

The bond order corrections f_i are given by

$$\begin{aligned}
f_1(\Delta_i, \Delta_j) &= \frac{1}{2} \left(\frac{\text{Val}_i + f_2(\Delta'_i, \Delta'_j)}{\text{Val}_i + f_2(\Delta'_i, \Delta'_j) + f_3(\Delta'_i, \Delta'_j)} \right. \\
&\quad \left. + \frac{\text{Val}_j + f_2(\Delta'_i, \Delta'_j)}{\text{Val}_j + f_2(\Delta'_i, \Delta'_j) + f_3(\Delta'_i, \Delta'_j)} \right) \\
f_2(\Delta'_i, \Delta'_j) &= e^{-p_{\text{boc1}} \cdot \Delta'_i} + e^{-p_{\text{boc1}} \cdot \Delta'_j} \\
f_3(\Delta'_i, \Delta'_j) &= -\frac{1}{p_{\text{boc2}}} \ln \left(\frac{1}{2} \left[e^{-p_{\text{boc2}} \cdot \Delta'_i} + e^{-p_{\text{boc2}} \cdot \Delta'_j} \right] \right) \\
f_4(\Delta'_i, \text{BO}'_{ij}) &= \frac{1}{1 + e^{-p_{\text{boc3}} \cdot (p_{\text{boc4}} \cdot \text{BO}'_{ij}{}^2 - \Delta'_i{}^{\text{boc}}) + p_{\text{boc5}}}} \\
f_5(\Delta'_j, \text{BO}'_{ij}) &= \frac{1}{1 + e^{-p_{\text{boc3}} \cdot (p_{\text{boc4}} \cdot \text{BO}'_{ij}{}^2 - \Delta'_j{}^{\text{boc}}) + p_{\text{boc5}}}}
\end{aligned} \tag{B.10}$$

Furthermore a corrected overcoordination Δ_i can be derived from the corrected bond orders in Equation (B.9) by

$$\Delta_i = -\text{Val}_i + \sum_{j=1}^{\text{neighs}(i)} \text{BO}_{ij} \tag{B.11}$$

As in classical models usual the energy is splitted into partial energy terms which describes for example torsional, Coulomb or van der Waals interactions. For the REAXFF model there are different implementations of amount and style of energy terms which summarized describing the total energy. According to Fogarty et al.⁷⁹ the total energy of the system is composed of the following energy terms:

$$E_{\text{system}} = E_{\text{bond}} + E_{\text{lp}} + E_{\text{over}} + E_{\text{under}} + E_{\text{val}} \quad (\text{B.12})$$

$$E_{\text{pen}} + E_{\text{tors}} + E_{\text{conj}} + E_{\text{vdWaals}} + E_{\text{Coulomb}}$$

Not all energy terms are necessary for the simulation of SiO_2 and water. Thus only those needed are discussed very briefly below. A more detailed discussion on the different energy contributions can be found in Chenoweth et al.⁴⁸.

BOND ENERGY

The energy of the bond between two atoms is calculated in REAXFF via the corrected bond order BO_{ij} defined in Equation (B.9):

$$E_{\text{bond}} = -D_e^\sigma \text{BO}_{ij}^\sigma \cdot e^{p_{\text{be1}}(1 - (\text{BO}_{ij}^\sigma)^{p_{\text{be2}}})} - D_e^\pi \text{BO}_{ij}^\pi - D_e^{\pi\pi} \text{BO}_{ij}^{\pi\pi} \quad (\text{B.13})$$

With this formulation of a bond energy it is ensured that on bond dissociation, i.e. when the bond orders tends to zero, its contribution to the system energy vanishes.

LONE PAIR ENERGY PENALTY

The difference between the total number of outer shell electrons, e.g. 6 for oxygen, 4 for silicon and 1 for hydrogen, and the sum of bond orders around an atomic center is described by

$$\Delta_i^e = -\text{Val}_i^e + \sum_{j=1}^{\text{neighs}(i)} \text{BO}_{ij}. \quad (\text{B.14})$$

Thus the number of lone pairs around an atom is given by

$$n_{\text{lp},i} = \text{int} \left(\frac{\Delta_i^e}{2} \right) + \exp \left(-p_{\text{lp}1} \left(2 + \Delta_i^e - 2 \cdot \text{int} \left[\frac{\Delta_i^e}{2} \right] \right)^2 \right). \quad (\text{B.15})$$

The deviation from an optimal value is thus

$$\Delta_i^{\text{lp}} = n_{\text{lp,opt}} - n_{\text{lp},i} \quad (\text{B.16})$$

and finally the energy penalty for an atom bearing lone pair electrons is given to

$$E_{\text{lp}} = \frac{p_{\text{lp}2} \cdot \Delta_i^{\text{lp}}}{1 + e^{-75 \cdot \Delta_i^{\text{lp}}}} \quad (\text{B.17})$$

OVERCOORDINATION ENERGY PENALTY

An atom is overcoordinated if Δ_i is greater zero. Thus an energy penalty on the system is introduced. However, the degree of overcoordination has to be decreased if an atom contains

a broken-up lone electron pair. Summarized this leads to the equation

$$E_{\text{over}} = \frac{\sum_{j=1}^{\text{nbond}} p_{\text{ovun1}} D_e^\sigma \text{BO}_{ij}}{\Delta_i^{\text{lpcorr}} + \text{Val}_i} \Delta_i^{\text{lpcorr}} \left(\frac{1}{1 + e^{p_{\text{ovun2}} \cdot \Delta_i^{\text{lpcorr}}}} \right) \quad (\text{B.18})$$

and the coordination number corrected for lone pairs

$$\Delta_i^{\text{lpcorr}} = \Delta_i - \frac{\Delta_i^{\text{lp}}}{1 + p_{\text{ovun3}} \cdot e^{p_{\text{ovun4}} \sum_{j=1}^{\text{neighs}(i)} (\Delta_j - \Delta_j^{\text{lp}}) \cdot (\text{BO}_{ij}^\pi + \text{BO}_{ij}^{\pi\pi})}} \quad (\text{B.19})$$

UNDERCOORDINATION ENERGY PENALTY Undercoordinated atoms satisfy the condition Δ_i smaller zero and a penalty should only apply for the resonance of π -electrons between attached undercoordinated atomic centers. Consequently this is formulated to

$$E_{\text{under}} = - \frac{p_{\text{ovun5}} \left(1 - e^{p_{\text{ovun6}} \cdot \Delta_i^{\text{lpcorr}}} \right)}{\left(1 + e^{p_{\text{ovun2}} \cdot \Delta_i^{\text{lpcorr}}} \right) \cdot \left(1 + p_{\text{ovun7}} \cdot e^{p_{\text{ovun8}} \sum_{j=1}^{\text{neighs}(i)} (\Delta_j - \Delta_j^{\text{lp}}) \cdot (\text{BO}_{ij}^\pi + \text{BO}_{ij}^{\pi\pi})} \right)} \quad (\text{B.20})$$

VALENCE ANGLE TERMS

ANGLE ENERGY E_{val} Similar to bond energies, valence angle contributions have to vanish on dissociation. However, this expression is more complex, as it depends additionally on the sum of π -bond orders (SBO), which defines the equilibrium angle Θ_0 between three successive atoms. As already mentioned before, only a brief description of the energy terms should be given here without further discussion. A more detailed discussion on this topic can be found in Chenoweth et al. ⁴⁸. The valence angle energy contribution is given to

$$E_{\text{val}} = f_7(\text{BO}_{ij}) \cdot f_7(\text{BO}_{ik}) \cdot f_8(\Delta_j) \left(p_{\text{val1}} - p_{\text{val1}} e^{-p_{\text{val2}} (\Theta_0(\text{BO}) - \Theta_{ijk})^2} \right) \quad (\text{B.21})$$

The corrections f_7 and f_8 are calculated with

$$f_7(\text{BO}_{ij}) = 1 - e^{-p_{\text{val}3} \cdot \text{BO}_{ij}^{p_{\text{val}4}}}$$

$$f_8(\Delta_j) = p_{\text{val}5} - \frac{(p_{\text{val}5} - 1) \cdot (2 + e^{p_{\text{val}6} \cdot \Delta_j^{\text{angle}}})}{1 + e^{p_{\text{val}6} \cdot \Delta_j^{\text{angle}}} + e^{p_{\text{val}7} \cdot \Delta_j^{\text{angle}}}}. \quad (\text{B.22})$$

The equilibrium angle between atoms i, j and k is given by

$$\Theta_0(\text{BO}) = \pi - \Theta_{0,0} \left(1 - e^{-p_{\text{val}10} \cdot (2 - \text{SBO}_2)} \right). \quad (\text{B.23})$$

SBO_2 in this equation is obtained by evaluation of

$$\text{SBO}_2 = \begin{cases} 0, & \text{SBO} \leq 0 \\ \text{SBO}^{p_{\text{val}9}}, & 0 < \text{SBO} < 1 \\ 2 - (2 - \text{SBO})^{p_{\text{val}9}}, & 1 < \text{SBO} < 2 \\ 2, & \text{SBO} \geq 2 \end{cases}, \quad (\text{B.24})$$

in which SBO is calculated by

$$\text{SBO} = \sum_{n=1}^{\text{neighs}(i)} (\text{BO}_{jn}^{\pi} + \text{BO}_{jn}^{\pi\pi}) + \left(1 - \prod_{n=1}^{\text{neighs}(j)} e^{-\text{BO}_{jn}^8} \right) \cdot \left(-\Delta_j^{\text{angle}} - p_{\text{val}8} \cdot n_{\text{ip},j} \right). \quad (\text{B.25})$$

And ultimately Δ_j^{angle} in Equation (B.25) is defined by

$$\Delta_j^{\text{angle}} = -\text{Val}_j^{\text{angle}} + \sum_{n=1}^{\text{neighs}(j)} \text{BO}_{jn} \quad (\text{B.26})$$

PENALTY ENERGY E_{pen} Although this term should be of negligible influence for SiO_2 simulations, it is included in $\text{REAXFF}_{\text{SiO}_2}$. Intended to reproduce the stability of systems with two double bonds sharing an atom in a valency angle, e.g. for allenes, it is given by

$$E_{\text{pen}} = p_{\text{pen1}} f_9(\Delta_j) e^{-p_{\text{pen2}} \cdot (\text{BO}_{ij}-2)^2} e^{-p_{\text{pen2}} \cdot (\text{BO}_{jk}-2)^2} \quad (\text{B.27})$$

$$\text{with } f_9(\Delta_j) = \frac{2 + e^{-p_{\text{pen3}} \cdot \Delta_j}}{1 + e^{-p_{\text{pen3}} \cdot \Delta_j} + e^{-p_{\text{pen4}} \cdot \Delta_j}}. \quad (\text{B.28})$$

TORSIONAL ANGLE TERMS

TORSIONAL ANGLE ENERGY E_{tors} Again, the energy contribution of torsional angles should vanish on bond dissociation. This is ensured by employing the equation

$$\begin{aligned} E_{\text{tors}} = & f_{10}(\text{BO}_{ij}, \text{BO}_{jk}, \text{BO}_{kl}) \sin \Theta_{ijk} \sin \Theta_{jkl} \\ & \cdot \left[\frac{1}{2} V_1 (1 + \cos \omega_{ijkl}) \right. \\ & + \frac{1}{2} V_2 e^{p_{\text{tor1}} \cdot (\text{BO}_{jk}^\pi - 1 + f_{11}(\Delta_j, \Delta_k))^2} (1 - \cos(2\omega_{ijkl})) \\ & \left. + \frac{1}{2} V_3 (1 + \cos(3\omega_{ijkl})) \right] \end{aligned} \quad (\text{B.29})$$

in conjunction with

$$\begin{aligned} f_{10}(\text{BO}_{ij}, \text{BO}_{jk}, \text{BO}_{kl}) = & (1 - e^{-p_{\text{tor2}} \cdot \text{BO}_{ij}}) \cdot (1 - e^{-p_{\text{tor2}} \cdot \text{BO}_{jk}}) \cdot (1 - e^{-p_{\text{tor2}} \cdot \text{BO}_{kl}}) \\ \text{and } f_{11}(\Delta_j, \Delta_k) = & \frac{2 + e^{-p_{\text{tor3}} \cdot (\Delta_j^{\text{angle}} + \Delta_k^{\text{angle}})}}{1 + e^{-p_{\text{tor3}} \cdot (\Delta_j^{\text{angle}} + \Delta_k^{\text{angle}})} + e^{-p_{\text{tor4}} \cdot (\Delta_j^{\text{angle}} + \Delta_k^{\text{angle}})}}. \end{aligned} \quad (\text{B.30})$$

FOUR BODY CONJUGATION TERM E_{conj} Conjugation effects are accounted for with

$$E_{\text{conj}} = p_{\text{cot1}} f_{12}(\text{BO}_{ij}, \text{BO}_{jk}, \text{BO}_{kl}) \left(1 + (\cos^2(\omega_{ijkl}) - 1) \sin \Theta_{ijk} \sin \Theta_{jkl} \right) \quad (\text{B.31})$$

$$f_{12}(\text{BO}_{ij}, \text{BO}_{jk}, \text{BO}_{kl}) = e^{-p_{\text{cot2}}(\text{BO}_{ij}-1.5)^2} e^{-p_{\text{cot2}}(\text{BO}_{jk}-1.5)^2} e^{-p_{\text{cot2}}(\text{BO}_{kl}-1.5)^2} \quad (\text{B.32})$$

as, e.g., in benzene and other aromatics.

NONBONDED INTERACTIONS

In REAXFF van der Waals and Coulomb interactions are included for all atom pairs, without taking any scaling of 1-2, 1-3 or 1-4 interactions between bonded atoms into account, to avoid alterations in the energy description upon bond dissociation.

TAPER CORRECTIONS Long-range interactions between to atoms i and j are switched smoothly to zero in REAXFF with a taper function, taking the form of a 7th order polynomial.

$$T = \frac{20}{r_c^7} r_{ij}^7 - \frac{70}{r_c^6} r_{ij}^6 + \frac{84}{r_c^5} r_{ij}^5 - \frac{35}{r_c^4} r_{ij}^4 + 1. \quad (\text{B.33})$$

Multiplied with van der Waals and Coulomb energies and forces, it is ensured that the first, second and third derivatives are continuous and go to zero at the cutoff boundary r_c .

VAN DER WAALS INTERACTIONS Van der Waals interactions take the form of a Morse potential

$$E_{\text{vdW}} = T \cdot D_{ij} \left(e^{\alpha_{ij} \left(1 - \frac{f_{13}(r_{ij})}{r_{\text{vdW}}} \right)} - 2e^{\alpha_{ij} \left(1 - \frac{f_{13}(r_{ij})}{r_{\text{vdW}}} \right)} \right)$$

$$\text{with } f_{13}(r_{ij}) = \left(r_{ij}^{p_{\text{vdW}1}} + \gamma^{-p_{\text{vdW}1}} \right)^{\frac{1}{p_{\text{vdW}1}}}. \quad (\text{B.34})$$

COULOMB INTERACTIONS Atomic point charges are calculated with a charge equilibration method (QEq)¹⁷⁸. This allows the description of polarizable effects, an advantage towards classical force fields with fixed point charges. In addition a shielded Coulomb-potential is used to adjust for orbital overlap between atoms at close distances.

$$E_{\text{Coul}} = T \cdot C \cdot \frac{q_i q_j}{\left[r_{ij}^3 + \left(\frac{1}{\gamma_{ij}} \right)^3 \right]^{\frac{1}{3}}} \quad (\text{B.35})$$

B.2 DERIVATION OF COULOMB INTERACTIONS

Electrostatic interactions in classical molecular dynamics simulations are mostly described by fixed point charges q_i at the center of each atom i , as for example electronegative elements attract electrons more than less electronegative elements. The charge density $\rho(\mathbf{r})$ of a system consisting of N particles can be written to

$$\rho(\mathbf{r}) = \sum_{i=1}^N q_i \delta(\mathbf{r} - \mathbf{r}_i), \quad (\text{B.36})$$

for which the electrostatic potential at point \mathbf{r} is obtained by solving the Poisson equation of a homogenous medium

$$\Delta\Phi(\mathbf{r}) = -\frac{\rho(\mathbf{r})}{\epsilon_p}. \quad (\text{B.37})$$

ϵ_p represents here the linear permittivity $\epsilon_p = \epsilon_r\epsilon_0$ of a homogeneous material, which is the product of the relative permittivity ϵ_r of the material, often also referred to as the dielectric constant, and the permittivity of free space ϵ_0 . The analytical solution of Equation (B.37) is then

$$\Phi(\mathbf{r}) = \sum_{i=1}^N \frac{q_i}{4\pi\epsilon_p} \frac{1}{|\mathbf{r} - \mathbf{r}_i|}. \quad (\text{B.38})$$

The electrostatic energy of this potential is related to

$$\begin{aligned} V(\mathbf{r}_1, \mathbf{r}_i, \dots, \mathbf{r}_N) &= \frac{1}{2} \sum_{i=1}^N q_i \Phi(\mathbf{r}_i) \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^{N'} \frac{q_i q_j}{4\pi\epsilon_p} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \end{aligned} \quad (\text{B.39})$$

Where the prime after the second sum indicates that the summation is not over elements where $i = j$. Looking at the the interaction between a pair of particles i and j separately, the electrostatic energy between those is thus

$$V(\mathbf{r}_i, \mathbf{r}_j) = \frac{q_i q_j}{4\pi\epsilon_p} \frac{1}{r_{ij}} \quad \text{with} \quad r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|. \quad (\text{B.40})$$

However, a drawback is that with fixed point charges polarization effects are not included. Although polarizable force-fields exists, these require often far more computational resources and fixed point charges mostly suffices to describe accurately the electrostatics.

B.3 DERIVATION OF VAN DER WAALS INTERACTIONS

An electrostatic description of charge neutral elements, such as rare gases, cannot be made with fixed point charges as the name suggests. It is a matter of fact that they will hardly react with any other elements including themselves to form compounds. However, those elements feature a weak interaction among themselves and with other elements through attractive dispersive London forces in the long-range part, stemming from quantum fluctuations in the electronic orbitals, and strong Pauli repulsion at closer distances due to overlapping electronic orbitals. Generally speaking, those interactions are, amongst others, typically summarized into the group of van der Waals forces which is characterized by contributions that are neither covalent nor hydrogen bonds and do not stem from electrostatic interactions of ions with one another or with other neutral or charged molecules. In the following a short line-up of all van der Waals group components is given

PAULI FORCES Results from the Pauli exclusion principle to prevent the collapse of molecules on close contact and is usually approximated by a r^{-12} dependence.

KEESOM FORCES Attractive or repulsive electrostatic interactions between permanent charges of multipoles. Those are also known as orientational forces.

DEBYE FORCES Attractive interaction between a permanent multipole and an induced multipole on another atom or molecule, which are sometimes called induction forces.

LONDON FORCES Attractive interaction between any pair of molecules. Arising from the fluctuation of instantaneous multipoles in the electronic clouds, which can in turn induce a multipole in the neighboring atoms.

This work focus only on dipole interactions and will neglect higher order terms of more complex multipoles, as a result this leads in general to a r^{-6} dependence of Keesom, Debye and London forces. Dispersive London forces are mostly the main contributor to the total energy of van der Waals interactions and are therefore described in more detail hereinafter. In principle, dispersive London interactions can be derived from quantum mechanics via a second-order perturbation theory. The perturbation is the Coulomb interaction between the electrons and nuclei of a dimer yielding a r^{-6} dependence. A more simpler model, which is explained in more detail here, was proposed by Drude. It consists of two molecules, labeled as 1 and 2. Each consists of two charges, $+q$ and $-q$, where the centers of mass of the two molecules are separated by the distance r . In both molecules the negative charge is rotating around fixed positive charge at a distance z from the z -axis with angular frequency ω . The Schrödinger equation for this system is thus¹⁸⁴:

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial z_1^2} - \frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial z_2^2} + \left[\frac{1}{2} k z_1^2 - \frac{1}{2} k z_2^2 - v(\mu_1, \mu_2) \right] \Psi = E \Psi, \quad (\text{B.41})$$

with the dipole-dipole interaction energy

$$v(\mu_1, \mu_2) = -\frac{2\mu_1\mu_2}{4\pi\epsilon_0 r^3} = -\frac{2z_1 z_2 q^2}{4\pi\epsilon_0 r^3}, \quad (\text{B.42})$$

where z_1 and z_2 are the separation distance of the charges in each molecule. By making the following substitutions

$$\begin{aligned} a_1 &= \frac{z_1 + z_2}{\sqrt{2}}, & a_2 &= \frac{z_1 - z_2}{\sqrt{2}}, \\ k_1 &= k - \frac{2q^2}{4\pi\epsilon_0 r^3}, & k_2 &= k + \frac{2q^2}{4\pi\epsilon_0 r^3}, \end{aligned} \quad (\text{B.43})$$

this equation can be transformed into the Schrödinger equation of two independent oscillators

$$-\frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial a_1^2} - \frac{\hbar^2}{2m} \frac{\partial^2 \Psi}{\partial a_2^2} + \left[\frac{1}{2} k_1 a_1^2 + \frac{1}{2} k_2 a_2^2 \right] \Psi = E \Psi, \quad (\text{B.44})$$

with the resulting frequencies

$$\omega_1 = \omega \sqrt{1 - \frac{2q^2}{4\pi\epsilon_0 r^3 k}} \quad \text{and} \quad \omega_2 = \omega \sqrt{1 + \frac{2q^2}{4\pi\epsilon_0 r^3 k}} \quad (\text{B.45})$$

where ω is the angular frequency of an isolated Drude molecule. The ground state energy is hence given to $E_0 = \frac{1}{2} \hbar(\omega_1 + \omega_2)$. Substituting ω_1 and ω_2 into the ground state energy E_0 gives

$$E_0(r) = \hbar\omega - \frac{q^4 \hbar \omega}{2(4\pi\epsilon_0)^2 r^6 k^2} + \dots \quad (\text{B.46})$$

Subtracting now from Equation (B.46) the energy of an infinitely separated system with $r \rightarrow \infty$ results in the interaction energy of the two oscillators

$$V(r) = -\frac{q^4 \hbar \omega}{2(4\pi\epsilon_0)^2 r^6 k^2}, \quad (\text{B.47})$$

showing the desired r^{-6} dependency. It can also be shown that the force constant k is related to the polarizability α of the Drude molecule. Under the assumption that a Drude molecule is in an electric field \mathbf{E} , a force $q\mathbf{E}$ will act on the molecule and separate the charges until an equilibrium with the force arising due to the stretching of the bond between them is reached. Bringing this to math, in equilibrium it can be stated that $qE = kz$ and a static dipole of magnitude $\mu_{ind} = qz = q^2E/k$ is induced. Vice versa is this dipole related to the polarisability times the electric field $\mu_{ind} = \alpha\mathbf{E}$ yielding $\alpha = q^2/k$. Thus Equation (B.47) can be rewritten to

$$V(r) = -\frac{\alpha^4 \hbar \omega}{2(4\pi\epsilon_0)^2 r^6} \quad (\text{B.48})$$

In contrast to this the typically chosen r^{-12} dependence to describe Pauli repulsion has yet no physical justification, but it approximates the repulsion behavior well enough and is often used because of its computational efficiency of calculating r^{-12} as the square of r^{-6} .

B.4 TREATMENT OF LONG-RANGE INTERACTIONS

B.4.1 FORCE SHIFTING OF COULOMB AND VAN DER WAALS INTERACTIONS

In order to achieve a smooth transition of the force to zero at the cutoff, Coulomb and van der Waals energy contributions of Equation (2.26) are separated in order to enable the ability to apply two different cutoff lengths for each, as they have normally varying long-range characteristics. The unaltered potential energy function of Coulomb and van der Waals interactions are denoted by $V_{\text{Coul}}(r)$ and $V_{\text{vdW}}(r)$, respectively. By multiplying both of them with a typical shifting function e.g. from the CHARMM forcefield³³ $S(r) = \left(1 - \frac{r}{r_c}\right)^2$, it is ensured that the force is shifted by a constant factor of either $\frac{q_i q_j}{4\pi\epsilon_p r_c}$ or $4\epsilon \left((\sigma/r_c)^{12} - (\sigma/r_c)^6 \right)$

in case of Coulomb or van der Waals interactions, so that both are shifted continuous to zero at the cutoff r_c . For the shifted Coulomb potential and force this leads to

$$\begin{aligned} V_{\text{Coul}}^{\text{sh}}(r) &= V_{\text{Coul}}(r) \cdot S(r) \\ &= \frac{q_i q_j}{4\pi\epsilon_p} \left(\frac{1}{r} - \frac{2}{r_c} + \frac{r}{r_c^2} \right) \end{aligned} \quad (\text{B.49})$$

$$F_{\text{Coul}}^{\text{sh}}(r) = -\frac{d}{dr} V_{\text{Coul}}^{\text{sh}}(r) = \frac{q_i q_j}{4\pi\epsilon_p} \left(\frac{1}{r^2} - \frac{1}{r_c^2} \right) \quad (\text{B.50})$$

and correspondingly for the van der Waals energy and force this yields

$$\begin{aligned} V_{\text{vdW}}^{\text{sh}}(r) &= V_{\text{vdW}}(r) \cdot S(r) \\ &= 4\epsilon \left(\sigma^{12} \left(\frac{1}{r^{12}} - \frac{2}{r^{11}r_c} + \frac{1}{r^{10}r_c^2} \right) - \sigma^6 \left(\frac{1}{r^6} - \frac{2}{r^5r_c} + \frac{1}{r^4r_c^2} \right) \right) \end{aligned} \quad (\text{B.51})$$

$$F_{\text{vdW}}^{\text{sh}}(r) = -4\epsilon \left(\sigma^{12} \left(\frac{12}{r^{13}} - \frac{22}{r^{12}r_c} + \frac{10}{r^{11}r_c^2} \right) - \sigma^6 \left(\frac{6}{r^7} - \frac{10}{r^6r_c} + \frac{4}{r^5r_c^2} \right) \right). \quad (\text{B.52})$$

A Drawback is, however, that this method will vary the potential energy and forces independent of the distance, because the constant shifting factor is added on the whole range of interactions until it reaches the cutoff point, which could be indeed a source of errors due to the underestimation of forces or energies.

FORCE SWITCHING OF COULOMB AND VAN DER WAALS INTERACTIONS

On the other hand, force or potential switching avoids this problematic by only applying a force or potential alternating switching function

$$S(r) = \begin{cases} 1, & r \leq r_{\text{on}} \\ \frac{(r_{\text{off}}^2 - r^2)^2 (r_{\text{off}}^2 + 2r^2 - 3r_{\text{on}}^2)}{(r_{\text{off}}^2 - r_{\text{on}}^2)^3}, & r_{\text{on}} < r \leq r_{\text{off}} \end{cases} \quad (\text{B.53})$$

to a small part of the original function leaving the rest unaltered. Nonetheless it is of great importance that the switching zone is in the case of potential switching sufficiently broad, depending on how steep the function is switched to zero, otherwise huge force errors in this region are the result due to the fact that the force is the first derivative of the potential energy and the derivative of a very steep function has thus also a very high signal. In Figure B.1 the effects of potential and force switching on the potentials of oxygen and hydrogen and on a water dimer are shown. Obviously, potential switching has a dramatic influence on the forces in the switching zone. Although potentials are switched in a very smooth way, forces are showing a remarkable bump in this region which can cause an improper behavior of the simulated molecules. This becomes even more evident for the simple example of a water dimer system, where two water molecules are consecutively pulled apart from each other and each time the total energy is recorded, shown on the upper and lower panel on the right side of Figure B.1. Two artificial energy minima are created around the switching zone and it can easily happen that in this case water molecules are energetically trapped at distances around the switching zone leading to a nonphysical pseudo-freezing of the water molecules in the simulation.

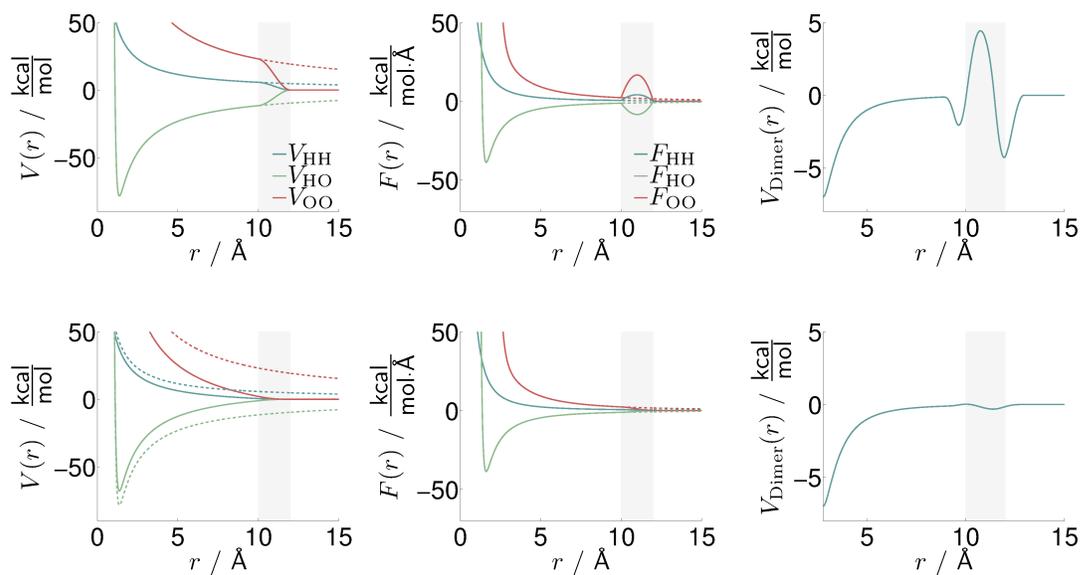


Figure B.1: Effects of potential (upper) and force (lower) switching on the non-bonding potentials and forces of Charmm-Tip3p hydrogen and oxygen and on the total energy of a water dimer system. Dashed lines represent unaltered forces and potentials. The switching zone is displayed by a grey shaded area.

As already mentioned, force switching shows altogether less of the previously mentioned artefacts and is explained in the following in more detail. It leads to the most robust estimate of dynamical and structural properties of the system whenever a spherical cutoff is needed and energies are of minor importance. A drawback is, however, that potential energies are shifted similar as in the force shifting approach due to the integration of the switched force, thus the total potential energy calculated from the shifted single potentials of every species interaction is not consistent with the true unscaled total potential energy as one can guess from Figure B.1.

Following the standard definition of force switching from the CHARMM force field²⁰² Coulomb and van der Waals forces are separately multiplied with the previously defined switching function $S(r)$ from Equation (B.53). As a consequence of the force switching the potential energy is now the integral over the switched force. Summarised this results to

$$\begin{aligned} F_{\text{Coul}}^{\text{sw}}(r) &= F_{\text{Coul}}(r) \cdot S(r) \\ &= \frac{q_i q_j}{4\pi\epsilon_p r^2} \cdot S(r) \end{aligned} \quad (\text{B.54})$$

$$V_{\text{Coul}}^{\text{sw}}(r) = \begin{cases} V_{\text{Coul}}(r) + \Delta V_1(r), & r \leq r_{\text{on}} \\ - \int_{r_{\text{off}}}^r dr' S(r') F_{\text{Coul}}^{\text{sw}}(r'), & r_{\text{on}} < r \leq r_{\text{off}} \end{cases} \quad (\text{B.55})$$

with the approximation from²⁰² this yields

$$V_{\text{Coul}}^{\text{sw}}(r) = \begin{cases} \frac{q_i q_j}{4\pi\epsilon_p} \left(\frac{1}{r^2} + 8 \frac{\left((r_{\text{on}} r_{\text{off}})^2 (r_{\text{off}} - r_{\text{on}}) - \frac{(r_{\text{off}}^5 - r_{\text{on}}^5)}{5} \right)}{\gamma} \right), & r \leq r_{\text{on}} \\ \frac{q_i q_j}{4\pi\epsilon_p} (A(r^{-1} - r_{\text{off}}^{-1}) + B(r_{\text{off}} - r) \\ \quad + C(r_{\text{off}}^3 - r^3) + D(r_{\text{off}}^5 - r^5)), & r_{\text{on}} < r \leq r_{\text{off}} \end{cases} \quad (\text{B.56})$$

where $A = r_{\text{off}}^4 (r_{\text{off}}^2 - 3r_{\text{on}}^2) / \gamma$, $B = 6(r_{\text{on}} r_{\text{off}})^2 / \gamma$, $C = -(r_{\text{on}}^2 + r_{\text{off}}^2) / \gamma$, $D = 2 / (5\gamma)$ and $\gamma = (r_{\text{off}}^2 - r_{\text{on}}^2)^3$.

Switched Lennard–Jones forces and their according potentials are similarly calculated as the switched Coulomb forces and its potentials. By multiplying the Lennard–Jones forces with the switching function $S(r)$ from Equation (B.53) in between the interval $r_{\text{on}} < r \leq r_{\text{off}}$ this yields for the switched forces

$$\begin{aligned} F_{\text{vdW}}^{\text{sw}}(r) &= F_{\text{vdW}}(r) \cdot S(r) \\ &= 24\epsilon \left(2 \frac{\sigma^{12}}{r^{13}} - \frac{\sigma^6}{r^7} \right) \cdot S(r) \end{aligned} \quad (\text{B.57})$$

and the corresponding potential

$$V_{\text{vdW}}^{\text{sw}}(r) = \begin{cases} V_{\text{vdW}}(r) + \Delta V_2(r), & r \leq r_{\text{on}} \\ - \int_{r_{\text{off}}}^r dr' S(r') F_{\text{vdW}}^{\text{sw}}(r'), & r_{\text{on}} < r \leq r_{\text{off}} \end{cases} \quad (\text{B.58})$$

For the sake of simplicity, the Lennard–Jones potential is separated into two contribution, one term to the power of twelve and the other to the power of six ($n = 12, 6$). Which yields

with the approximation from²⁰²

$$V_{\text{vdW}}^{\text{sw},n}(r) = \begin{cases} V_{\text{vdW}}(r) + \Delta V_2, & r \leq r_{\text{on}} \\ k(r^{-\beta} - r_{\text{off}}^{-\beta})^2, & r_{\text{on}} < r \leq r_{\text{off}} \end{cases} \quad (\text{B.59})$$

with $\Delta V_2 = \frac{-c}{(r_{\text{on}}r_{\text{off}})^\beta}$ and $k = c \frac{r_{\text{off}}^\beta}{r_{\text{off}}^\beta - r_{\text{on}}^\beta}$, where β is defined as $n/2$ depending if n is defined as 6 or 12 identifying which Lennard–Jones term should be switched and $c = 4\epsilon\sigma^n$ as the corresponding prefactor of the respective Lennard–Jones term. Hence the switched potential energy is

$$V_{\text{vdW}}^{\text{sw}}(r) = \begin{cases} 4\epsilon \left(\sigma^{12} \left(\frac{1}{r^{12}} - \frac{1}{(r_{\text{on}}r_{\text{off}})^6} \right) - \sigma^6 \left(\frac{1}{r^6} - \frac{1}{(r_{\text{on}}r_{\text{off}})^3} \right) \right), & r \leq r_{\text{on}} \\ 4\epsilon \left(\sigma^{12} \left(\frac{1}{r^6} - \frac{1}{r_{\text{off}}^6} \right)^2 \cdot C_1 - \sigma^6 \left(\frac{1}{r^3} - \frac{1}{r_{\text{off}}^3} \right)^2 \cdot C_2 \right), & r_{\text{on}} < r \leq r_{\text{off}} \end{cases} \quad (\text{B.60})$$

with the two prefactors $C_1 = \left(\frac{r_{\text{off}}^6}{r_{\text{off}}^6 - r_{\text{on}}^6} \right)$ and $C_2 = \left(\frac{r_{\text{off}}^3}{r_{\text{off}}^3 - r_{\text{on}}^3} \right)$.

B.4.2 EWALD SUMMATION OF COULOMBIC INTERACTIONS

Basically, the potential energy is now composed of particle-particle interaction stemming from the central simulation cell and from interactions between the particles of the central simulation cell and an infinite array of periodically repeated image cells. For a simulation box comprising of N atoms at the cubic lattice position $\mathbf{nL} = (n_xL, n_yL, n_zL)$, where L is the box length, the charge density is thus given to

$$\rho(\mathbf{r}) = \sum_{i=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} q_i \delta(\mathbf{r} - \mathbf{r}_i + \mathbf{nL}). \quad (\text{B.61})$$

The Poisson Equation (B.37) is analytically solvable for this charge density. Hence, the electrostatic potential is given by

$$\Phi(\mathbf{r}) = \sum_{i=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} \frac{q_i}{4\pi\epsilon_p} \frac{1}{|\mathbf{r} - \mathbf{r}_i + \mathbf{nL}|} \quad (\text{B.62})$$

and the total energy of charge-charge interactions is calculated according to

$$\begin{aligned} V(\mathbf{r}_1, \mathbf{r}_i, \dots, \mathbf{r}_N) &= \frac{1}{2} \sum_{i=1}^N q_i \Phi(\mathbf{r}_i) \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum'_{\mathbf{n} \in \mathbb{Z}^3} \frac{q_i q_j}{4\pi\epsilon_p} \frac{1}{|\mathbf{r}_{ij} + \mathbf{nL}|}. \end{aligned} \quad (\text{B.63})$$

The prime indicates here that the summation does not include the non desired $i = j$ interaction of the same atom in the central box ($|\mathbf{n}| = 0$). Problematic is that this summation converges extremely slow and is de facto conditionally convergent. Fortunately, this circumstance can be avoided by splitting the electrostatic energy into a real and a Fourier space contribution. Fourier, discrete Fourier and their inverse transformations are here denoted by

$$\mathcal{F}[f(\mathbf{r})]_{\mathbf{k}} = \tilde{f}(\mathbf{k}) = \int_{\mathbb{R}^3} f(\mathbf{r}) e^{-i\mathbf{r} \cdot \mathbf{k}} d^3 r \quad (\text{B.64})$$

$$\widehat{\mathcal{F}}[f(\mathbf{s})]_{\mathbf{k}} = \hat{f}(\mathbf{k}) = \sum_{\mathbf{s} \in \mathbb{G}} f(\mathbf{s}) e^{-i\mathbf{k} \cdot \mathbf{s}} d^3 k \quad (\text{B.65})$$

$$f(\mathbf{r}) = \mathcal{F}^{-1}[\tilde{f}(\mathbf{k})]_{\mathbf{r}} = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^3} \tilde{f}(\mathbf{k}) e^{i\mathbf{r} \cdot \mathbf{k}} \quad (\text{B.66})$$

$$f(\mathbf{s}) = \widehat{\mathcal{F}}^{-1}[\hat{f}(\mathbf{k})]_{\mathbf{r}} = \frac{1}{V} \sum_{\mathbf{s} \in \mathbb{G}} \hat{f}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{s}}, \quad (\text{B.67})$$

where \mathbb{G} is an arbitrary discrete set. Folding the point charge density $\rho(\mathbf{r})$ of Equation (B.61) with a Gaussian distribution, typically $\gamma(\mathbf{r}) = \frac{\alpha^3}{\pi^{3/2}} e^{-\alpha^2 r^2}$, yields for the charge densities in real (RS) and Fourier space (FS)

$$\rho^{(\text{RS})}(\mathbf{r}) = \rho(\mathbf{r}) - (\gamma * \rho)(\mathbf{r}) \quad (\text{B.68})$$

$$\rho^{(\text{FS})}(\mathbf{r}) = (\gamma * \rho)(\mathbf{r}) \quad (\text{B.69})$$

whereby $\rho^{(\text{RS})} + \rho^{(\text{FS})} = \rho$ and the $*$ -operation denotes the three dimensional folding $(f * g)(\mathbf{r}) = \int_{\mathbb{R}^3} f(\mathbf{r})g(\mathbf{r} - \mathbf{r}')d^3r'$. Recalling Equation (B.37), the Poisson equation can now separately be solved for the charge distributions $\rho^{(\text{RS})}$ and $\rho^{(\text{FS})}$ of the real and Fourier space, respectively:

$$-\Delta\Phi^{(\text{RS})}(\mathbf{r}) = \frac{\rho^{(\text{RS})}(\mathbf{r})}{\epsilon_p}; \quad V^{(\text{RS})} = \frac{1}{2} \sum_{i=1}^N q_i \Phi^{(\text{RS})}(\mathbf{r}_i) \quad (\text{B.70})$$

$$-\Delta\Phi^{(\text{FS})}(\mathbf{r}) = \frac{\rho^{(\text{FS})}(\mathbf{r})}{\epsilon_p}; \quad V^{(\text{FS})} = \frac{1}{2} \sum_{i=1}^N q_i \Phi^{(\text{FS})}(\mathbf{r}_i) \quad (\text{B.71})$$

Beginning with the derivation of the real space energy $V^{(\text{RS})}$, its contribution to the energy and its according potential can be solved analytically with the use of the complementary error function $\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$. Thus it is calculated to

$$\Phi^{(\text{RS})}(\mathbf{r}) = \sum_{i=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} \frac{q_i}{4\pi\epsilon_p} \frac{\text{erfc}(\alpha|\mathbf{r} - \mathbf{r}_i + \mathbf{nL}|)}{|\mathbf{r} - \mathbf{r}_i + \mathbf{nL}|} \quad (\text{B.72})$$

$$V^{(\text{RS})} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{n} \in \mathbb{Z}^3} \frac{q_i q_j}{4\pi\epsilon_p} \frac{\text{erfc}(\alpha|\mathbf{r}_{ij} + \mathbf{nL}|)}{|\mathbf{r}_{ij} + \mathbf{nL}|} \quad (\text{B.73})$$

The Fourier space contribution to the energy $V^{(\text{FS})}$ is in turn calculated by a Fourier transformation of the Poisson equation and application of the convolution and shifting theorem, which gives

$$\begin{aligned}\tilde{\Phi}^{(\text{FS})}(\mathbf{k}) &= \frac{1}{\epsilon_p k^2} \mathcal{F}[(\gamma * \rho)(\mathbf{r})]_{\mathbf{k}} = \tilde{g}(\mathbf{k}) \tilde{\gamma}(\mathbf{k}) \tilde{\rho}(\mathbf{k}) \frac{1}{\epsilon_p} \\ &= \sum_{i=1}^N \frac{q_i}{\epsilon_p k^2} \exp\left(-\frac{k^2}{4\alpha^2}\right) e^{i\mathbf{k}\cdot(\mathbf{r}-\mathbf{r}_i)},\end{aligned}\quad (\text{B.74})$$

with the Fourier transformed Green function $\tilde{g}(\mathbf{k}) = 1/k^2$. Furthermore can the summation over \mathbf{n} be neglected in the Fourier space by employing the symmetry of the simulation box. In other words, this is due to the fact that $\mathbf{k} = 2\pi\mathbf{n}/L$ is a reciprocal vector and $\exp(-i\mathbf{k} \cdot \mathbf{n}L)$ equals thus one. Discrete inverse Fourier transformation of Equation (B.74) and utilizing in Equation (B.71) gives the Fourier space contribution to the electrostatic energy

$$V^{(\text{FS})} = \frac{1}{2L^3} \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{k} \in \mathcal{M}} \frac{q_i q_j}{\epsilon_p k^2} \exp\left(-\frac{k^2}{4\alpha^2}\right) e^{i\mathbf{k}\cdot(\mathbf{r}_i-\mathbf{r}_j)}, \quad (\text{B.75})$$

with $\mathcal{M} = \{2\pi\mathbf{n}/L \mid \mathbf{n} \in \mathbb{Z}^3 \setminus \{0\}\}$. For the time being, it shall be assumed that Equation (B.75) for $\mathbf{k} = 0$ of a charge neutral system, i.e. $\sum_{i=1}^N q_i = 0$, is zero; A direct consequence of the conditional convergence of the Ewald sum. This assumption is only correct if the periodic system is embedded in a medium with infinite dielectric constant, as it shall be seen later upon the derivation of the dipole correction term.

Closer examination of this equations reveals the necessity of additional correction terms. One is due to the inclusion of self interaction terms stemming from contributions when $i = j$ and $|\mathbf{n}| = 0$ in the Fourier space summation. In the real space summation of Equa-

tion (B.73) this is explicitly not accounted for, whereas the Fourier space summation is not restricted that rigorous. Thus a correction term for the self interaction $V^{(SI)}$ has to be introduced:

$$V^{(SI)} = \frac{1}{2} \sum_{i=1}^N q_i \lim_{\mathbf{r}_i \rightarrow 0} \Phi^{(SI)}(\mathbf{r}_i) \quad (\text{B.76})$$

$\Phi^{(SI)}(\mathbf{r})$ is obtained by solving the Poisson equation for a Gaussian charge distribution centered at the position \mathbf{r} using the spherical symmetry of the Gaussian charge cloud.

$$\rho^{(SI)}(\mathbf{r}) = -q_i(\gamma * \delta(\mathbf{r})) \quad (\text{B.77})$$

$$\Phi^{(SI)}(\mathbf{r}) = -\frac{q_i}{4\pi\epsilon_p} \frac{1}{r} \text{erf}(\alpha r) \quad \text{with } r = |\mathbf{r}|, \quad (\text{B.78})$$

where the definition of the error function $\text{erf} = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ was employed. To compensate the additional Gaussian for $i = j$ in Fourier space when $\mathbf{n} = 0$ a Gaussian with opposite sign is positioned at the center of charge i in real space. Hence, in the case of self interaction the distance \mathbf{r}_i to the compensating charge distribution is zero. Thus $\lim_{\mathbf{r} \rightarrow 0} \Phi^{(SI)}(\mathbf{r}) = -2q_i \frac{\alpha}{4\pi^{\frac{3}{2}} \epsilon_p}$ and the spurious self interaction energy is then

$$V^{(SI)} = -\frac{1}{2} \sum_{i=1}^N \frac{\alpha q_i^2}{2\pi^{\frac{3}{2}} \epsilon_p} \quad (\text{B.79})$$

The other term considers a correction for the dielectric boundary conditions in infinity $V^{(DS)}$. It can be traced back to the negligence of $\mathbf{k} = 0$ in the Fourier space summation of Equation (B.75) in case of a relative permittivity ϵ'_r of the surrounding medium not infinite, i.e. if non-metallic boundary conditions are assumed^{62,98}. A detailed explanation of this term would break the mold, wherefore only a formal derivation should be given as done

by Frenkel & Smit⁸⁰. It may appear *prima facie* a bit curious that the potential energy depends on the nature of the boundary conditions in infinity, but it has a simple physical interpretation: Varying charge distribution in the unit cell results in a fluctuating total dipole moment of the same unit cell,

$$\mathbf{M} = \sum_{i=1}^N \boldsymbol{\mu}_i = \sum_{i=1}^N \mathbf{r}_i q_i. \quad (\text{B.80})$$

On the surface of the infinitely large Ewald sphere, constructed by periodic repetitions of the unit cell in all three dimensions, a surface charge will emerge due to the normalized polarization per unit volume $\mathbf{P} = \mathbf{M}/V$ and will in turn introduce a homogeneous electric field in the Ewald sphere

$$\mathbf{E} = -\frac{\mathbf{P}}{3\epsilon_p}. \quad (\text{B.81})$$

The surrounding medium responds with a counter charge distribution, inducing, in turn, a field trying to counteract the electric field in the unit cell. In case of metallic boundary conditions, that is ϵ'_r of the surrounding medium is infinite, the electric field emerging from the surface charge of the surrounding medium will compensate completely the electric field stemming from the dipole fluctuations and no correction is necessary. However, if the surrounding medium has a relative permittivity ϵ'_r other than infinity the electric field from Equation (B.81) cannot be compensated fully. Following the so-called reaction field theory⁸³ the arising depolarization field is given by

$$\mathbf{E}_{\text{RF}} = \frac{2(\epsilon'_r - 1)}{2\epsilon'_r + 1} \mathbf{E}. \quad (\text{B.82})$$

In case of vacuum boundary conditions, i.e. $\epsilon'_r = 1$, this term will vanish. The total electrical field is thus obtained by combining Equation (B.81) and (B.82)

$$\mathbf{E} \rightarrow \mathbf{E} + \mathbf{E}_{\text{RF}} = \frac{\mathbf{P}}{\epsilon_p(2\epsilon'_r + 1)} \quad (\text{B.83})$$

The resulting electric field \mathbf{E} is independent of the radius of the sphere, as well it is constant within the sphere. The electrostatic potential is thus given by

$$\Phi^{(\text{DC})}(\mathbf{r}) = -\mathbf{r} \cdot \mathbf{E} = \frac{1}{\epsilon_p(2\epsilon'_r + 1)} \frac{\mathbf{r} \cdot \mathbf{M}}{V}. \quad (\text{B.84})$$

Hence the dipolar correction term $V^{(\text{DC})}$ is calculated according to

$$\begin{aligned} V^{(\text{DC})}(\mathbf{r}) &= \frac{1}{2} \sum_{i=1}^N q_i \Phi^{(\text{DC})}(\mathbf{r}_i) \\ &= \frac{1}{2L^3 \epsilon_p (2\epsilon'_r + 1)} \left(\sum_{i=1}^N q_i \mathbf{r}_i \right)^2. \end{aligned} \quad (\text{B.85})$$

Summing up all contributions, the total coulombic energy is finally given to

$$\begin{aligned} V &= V^{(\text{RS})} + V^{(\text{FS})} + V^{(\text{SI})} + V^{(\text{DC})} \\ &= \sum_{i=1}^N \sum_{j=1}^N \sum'_{\mathbf{n} \in \mathbb{Z}^3} \frac{q_i q_j \operatorname{erfc}(\alpha |\mathbf{r}_{ij} + \mathbf{nL}|)}{8\pi \epsilon_p |\mathbf{r}_{ij} + \mathbf{nL}|} \\ &\quad + \sum_{i=1}^N \sum_{j=1}^N \sum_{\mathbf{k} \in \mathcal{M}} \frac{q_i q_j}{2L^3 \epsilon_p k^2} e^{-\frac{k^2}{4\alpha^2}} e^{i\mathbf{k} \cdot (\mathbf{r}_i - \mathbf{r}_j)} \\ &\quad - \sum_{i=1}^N \frac{\alpha q_i^2}{4\pi^{\frac{3}{2}} \epsilon_p} + \frac{1}{L^3 \epsilon_p (4\epsilon'_r + 2)} \left(\sum_{i=1}^N q_i \mathbf{r}_i \right)^2 \end{aligned} \quad (\text{B.86})$$

In the course of time more sophisticated and faster algorithms began to evolve, based upon the initial approach derived originally by Ewald⁷⁶. Traditional Ewald summation for example scales as $N^{3/2}$, whereas recent algorithms, such as Particle-mesh Ewald (PME)⁶⁰, smooth particle-mesh Ewald (SPME)⁷³ or particle-particle particle-mesh (PPPM)¹⁰² methods, all scale with $N \log N$. Those are used commonly and in this work as well; The latter mentioned method is used here in all simulation where long-range Coulomb interactions have to be treated correctly .

References

- [1] Akhtar, N., Padilla, M. L., Dickerson, E. B., Steinberg, H., Brent, M., Auerbach, R., & Helfand, S. C. (2004). Interleukin-12 inhibits tumor growth in a novel angiogenesis canine hemangiosarcoma xenograft model. *Neoplasia*, 6(2), 106 – 116.
- [2] Andersen, H. C. (1983). RATTLE: A “velocity” version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*, 52(1), 24–34.
- [3] Ansorena, P., Zuzuarregui, A., Pérez-Lorenzo, E., Mujika, M., & Arana, S. (2011). Comparative analysis of {QCM} and {SPR} techniques for the optimization of immobilization sequences. *Sensors and Actuators, B: Chemical*, 155(2), 667–672.
- [4] Bakaev, V. A. & Steele, W. A. (1999). On the computer simulation of a hydrophobic vitreous silica surface. *Journal of Chemical Physics*, 111(21), 9803–9812.
- [5] Balsera, M., Stepaniants, S., Izrailev, S., Oono, Y., & Schulten, K. (1997). Reconstructing potential energy functions from simulated force-induced unbinding processes. *Biophysical Journal*, 73(3), 1281–1287.
- [6] Banerjee, I., Pangule, R. C., & Kane, R. S. (2011). Antifouling coatings: Recent developments in the design of surfaces that prevent fouling by proteins, bacteria, and marine organisms. *Advanced Materials*, 23(6), 690–718.
- [7] Barducci, A., Bonomi, M., & Parrinello, M. (2011). Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(5), 826–843.
- [8] Barducci, A., Bussi, G., & Parrinello, M. (2008). Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Physical Review Letters*, 100, 020603.
- [9] Barisik, M., Atalay, S., Beskok, A., & Qian, S. (2014). Size dependent surface charge properties of silica nanoparticles. *Journal of Physical Chemistry C*, 118(4), 1836–1842.

- [10] Bartels, C. (2000). Analyzing biased Monte Carlo and molecular dynamics simulations. *Chemical Physics Letters*, 331(5–6), 446–454.
- [11] Basiuk, V. A. (2007). *Encyclopedia of Surface and Colloid Science*, volume 1, chapter Adsorption of Biomolecules at Silica, (pp. 359–375). Taylor & Francis.
- [12] Bayley, P. (1973). The analysis of circular dichroism of biomolecules. *Progress in Biophysics and Molecular Biology*, 27, 1–76.
- [13] Bayley, P. M., Nielsen, E. B., & Schellman, J. A. (1969). Rotatory properties of molecules containing two peptide groups: Theory. *Journal of Physical Chemistry*, 73(1), 228–243.
- [14] Bayly, C. I., Cieplak, P., Cornell, W., & Kollman, P. A. (1993). A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: The RESP model. *Journal of Physical Chemistry*, 97(40), 10269–10280.
- [15] Begum, R. & Matsuura, H. (1997). Conformational properties of short poly(oxyethylene) chains in water studied by IR spectroscopy. *Journal of the Chemical Society, Faraday Transactions*, 93, 3839–3848.
- [16] Behrens, S. H. & Grier, D. G. (2001). The charge of glass and silica surfaces. *Journal of Chemical Physics*, 115(14), 6716–6721.
- [17] Berendsen, H. (2007). *Simulating the Physical World: Hierarchical Modeling from Quantum Mechanics to Fluid Dynamics*. Cambridge University Press.
- [18] Berendsen, H. J. C., Grigera, J. R., & Straatsma, T. P. (1987). The missing term in effective pair potentials. *Journal of Physical Chemistry*, 91(24), 6269–6271.
- [19] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *Journal of Chemical Physics*, 81(8), 3684–3690.
- [20] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., & Hermans, J. (1981). *Intermolecular Forces*, volume 14 of *The Jerusalem Symposia on Quantum Chemistry and Biochemistry*, chapter Interaction Models for Water in Relation to Protein Hydration, (pp. 331–342). Springer Netherlands.
- [21] Berendsen, H. J. C. & van Gunsteren, W. F. (1986). Practical algorithms for dynamic simulations. In G. Ciccotti & W. Hoover (Eds.), *Molecular Dynamics Simulation of Statistical Mechanical Systems* Amsterdam, North Holland.

- [22] Bernath, P. (1995). *Spectra of Atoms and Molecules*. Topics in physical chemistry series. Oxford University Press.
- [23] Berteotti, A., Barducci, A., & Parrinello, M. (2011). Effect of urea on the β -hairpin conformational ensemble and protein denaturation mechanism. *Journal of the American Chemical Society*, 133(43), 17200–17206.
- [24] Beveridge, D. L. & DiCapua, F. M. (1989). Free energy via molecular simulation: Applications to chemical and biomolecular systems. *Annual Review of Biophysics and Biophysical Chemistry*, 18(1), 431–492.
- [25] Bonomi, M., Barducci, A., & Parrinello, M. (2009a). Reconstructing the equilibrium boltzmann distribution from well-tempered metadynamics. *Journal of Computational Chemistry*, 30(11), 1615–1621.
- [26] Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. A., & Parrinello, M. (2009b). PLUMED: A portable plugin for free-energy calculations with molecular dynamics. *Computer Physics Communications*, 180(10), 1961–1972.
- [27] Bouchiat, C., Wang, M., Allemand, J.-F., Strick, T., Block, S., & Croquette, V. (1999). Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophysical Journal*, 76(1), 409–413.
- [28] Bowler, D. R. & Miyazaki, T. (2012). $\mathcal{O}(n)$ methods in electronic structure calculations. *Reports on Progress in Physics*, 75(3), 036503.
- [29] Brahms, S. & Brahms, J. (1980). Determination of protein secondary structure in solution by vacuum ultraviolet circular dichroism. *Journal of Molecular Biology*, 138(2), 149–178.
- [30] Brambilla, D., Verpillot, R., Le Droumaguet, B., Nicolas, J., Taverna, M., Kóňa, J., Lettiero, B., Hashemi, S. H., De Kimpe, L., Canovi, M., Gobbi, M., Nicolas, V., Scheper, W., Moghimi, S. M., Tvaroška, I., Couvreur, P., & Andrieux, K. (2012). PE-Gylated nanoparticles bind to and alter amyloid- β peptide conformation: Toward engineering of functional nanomedicines for alzheimer’s disease. *ACS Nano*, 6(7), 5897–5908.
- [31] Brei, V. V. (1994). ^{29}Si solid-state NMR study of the surface structure of aerosil silica. *Journal of the Chemical Society, Faraday Transactions*, 90, 2961–2964.

- [32] Brenner, D. W. (1990). Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. *Physical Review B: Solid State*, 42(15), 9458–9471.
- [33] Brooks, B. R., Brooks, I. C. L., Mackerell, J. A. D., Nilsson, L., Petrella, R. J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caffisch, A., Caves, L., Cui, Q., Dinner, A. R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R. W., Post, C. B., Pu, J. Z., Schaefer, M., Tidor, B., Venable, R. M., Woodcock, H. L., Wu, X., Yang, W., York, D. M., & Karplus, M. (2009). CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, 30(10, Sp. Iss. SI), 1545–1614.
- [34] Buehler, M. J., Tang, H., van Duin, A. C. T., & Goddard, W. A. (2007). Threshold crack speed controls dynamical fracture of silicon single crystals. *Physical Review Letters*, 99, 165502.
- [35] Buehler, M. J., van Duin, A. C. T., & Goddard, W. A. (2006). Multiparadigm modeling of dynamical crack propagation in silicon using a reactive force field. *Physical Review Letters*, 96, 095505.
- [36] Bulheller, B. M. (2009). *Circular and Linear Dichroism Spectroscopy of Proteins*. PhD thesis, The University of Nottingham.
- [37] Bulheller, B. M. & Hirst, J. D. (2009). DichroCalc - circular and linear dichroism online. *Bioinformatics*, 25(4), 539–540.
- [38] Bulheller, B. M., Rodger, A., & Hirst, J. D. (2007). Circular and linear dichroism of proteins. *Physical Chemistry Chemical Physics*, 9, 2020–2035.
- [39] Bullerjahn, J. T., Sturm, S., & Kroy, K. (2014). Theory of rapid force spectroscopy. *Nature Communications*, 5.
- [40] Burkett, S. L. & Read, M. J. (2001). Adsorption-induced conformational changes of α -helical peptides. *Langmuir*, 17(16), 5059–5065.
- [41] Burneau, A., Barres, O., Gallas, J. P., & Lavalley, J. C. (1990). Comparative study of the surface hydroxyl groups of fumed and precipitated silicas. 2. characterization by infrared spectroscopy of the interactions with water. *Langmuir*, 6(8), 1364–1372.
- [42] Bussi, G., Gervasio, F. L., Laio, A., & Parrinello, M. (2006). Free-energy landscape for β -hairpin folding from combined parallel tempering and metadynamics. *Journal of the American Chemical Society*, 128(41), 13435–13441.

- [43] Butenuth, A., Moras, G., Schneider, J., Koleini, M., Köppen, S., Meißner, R., Wright, L. B., Walsh, T. R., & Ciacchi, L. C. (2012). Ab initio derived force-field parameters for molecular dynamics simulations of deprotonated amorphous-SiO₂/water interfaces. *Physica Status Solidi B: Basic Solid State Physics*, 249(2), 292–305.
- [44] Butt, H. J. & Jaschke, M. (1995). Calculation of thermal noise in atomic force microscopy. *Nanotechnology*, 6(1), 1–7.
- [45] Camilloni, C., Provasi, D., Tiana, G., & Broglia, R. A. (2008). Exploring the protein G helix free-energy surface by solute tempering metadynamics. *Proteins: Structure, Function, and Bioinformatics*, 71(4), 1647–1654.
- [46] Cantor, C. (1980). *The conformation of biological macromolecules*. San Francisco: W. H. Freeman.
- [47] Ceresoli, D., Bernasconi, M., Iarlori, S., Parrinello, M., & Tosatti, E. (2000). Two-membered silicon rings on the dehydroxylated surface of silica. *Physical Review Letters*, 84, 3887–3890.
- [48] Chenoweth, K., van Duin, A. C. T., & Goddard, W. A. (2008). ReaxFF reactive force field for molecular dynamics simulations of hydrocarbon oxidation. *Journal of Physical Chemistry A*, 112(5), 1040–53.
- [49] Cohavi, O., Corni, S., De Rienzo, F., Di Felice, R., Gottschalk, K. E., Hoefling, M., Kokh, D., Molinari, E., Schreiber, G., Vaskevich, A., & Wade, R. C. (2010). Protein–surface interactions: challenging experiments and computations. *Journal of Molecular Recognition*, 23(3), 259–262.
- [50] Cole, D. J., Payne, M. C., & Ciacchi, L. C. (2009). Water structuring and collagen adsorption at hydrophilic and hydrophobic silicon surfaces. *Physical Chemistry Chemical Physics*, 11, 11395–11399.
- [51] Cole, D. J., Payne, M. C., Csányi, G., Spearing, S. M., & Colombi Ciacchi, L. (2007). Development of a classical force field for the oxidized Si surface: Application to hydrophilic wafer bonding. *Journal of Chemical Physics*, 127(20), 204704.
- [52] Colombi Ciacchi, L., Cole, D. J., Payne, M. C., & Gumbsch, P. (2008). Stress-driven oxidation chemistry of wet silicon surfaces. *The Journal of Physical Chemistry C*, 112(32), 12077–12080.

- [53] Cormack, A. N., Lewis, R. J., & Goldstein, A. H. (2004). Computer simulation of protein adsorption to a material surface in aqueous solution: Biomaterials modeling of a ternary system. *Journal of Physical Chemistry B*, 108(52), 20408–20418.
- [54] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., & Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19), 5179–5197.
- [55] Corni, S., Hnilova, M., Tamerler, C., & Sarikaya, M. (2013). Conformational behavior of genetically-engineered dodecapeptides as a determinant of binding affinity for gold. *Journal of Physical Chemistry C*, 117(33), 16990–17003.
- [56] Couto, S. R. & Herrera, J. L. T. (2006). Industrial and biotechnological applications of laccases: A review. *Biotechnology Advances*, 24(5), 500–513.
- [57] Cruz-Chu, E. R., Aksimentiev, A., & Schulten, K. (2006). Water-silica force field for simulating nanodevices. *Journal of Physical Chemistry B*, 110(43), 21497–21508.
- [58] Cygan, R. T., Liang, J.-J., & Kalinichev, A. G. (2004). Molecular models of hydroxide, oxyhydroxide, and clay phases and the development of a general force field. *Journal of Physical Chemistry B*, 108(4), 1255–1266.
- [59] Da Silva, J. R. G., Pinatti, D. G., Anderson, C. E., & Rudee, M. L. (1975). A refinement of the structure of vitreous silica. *Philosophical Magazine*, 31(3), 713–717.
- [60] Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *Journal of Chemical Physics*, 98(12), 10089–10092.
- [61] Davis, D. H., Giannoulis, C. S., Johnson, R. W., & Desai, T. A. (2002). Immobilization of RGD to <111> silicon surfaces for enhanced cell adhesion and proliferation. *Biomaterials*, 23(19), 4019–4027.
- [62] de Leeuw, S. W., Perram, J. W., & Smith, E. R. (1980). Simulation of electrostatic systems in periodic boundary conditions. I. Lattice sums and dielectric constants. *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, 373(1752), 27–56.
- [63] de Oteyza, D. G., Gorman, P., Chen, Y.-C., Wickenburg, S., Riss, A., Mowbray, D. J., Etkin, G., Pedramrazi, Z., Tsai, H.-Z., Rubio, A., Crommie, M. F., & Fischer,

- F. R. (2013). Direct imaging of covalent bond structure in single-molecule chemical reactions. *Science*, 340(6139), 1434–1437.
- [64] de Souza, O. & Ornstein, R. (1997). Effect of periodic box size on aqueous molecular dynamics simulation of a DNA dodecamer with particle-mesh Ewald method. *Biophysical Journal*, 72(6), 2395–2397.
- [65] De Yoreo, J. J., Chung, S., & Friddle, R. W. (2013). In situ atomic force microscopy as a tool for investigating interactions and assembly dynamics in biomolecular and biomineral systems. *Advanced Functional Materials*, 23(20), 2525–2538.
- [66] Demiralp, E., Çağın, T., & Goddard, W. A. (1999). Morse stretch potential charge equilibrium force field for ceramics: Application to the quartz-stishovite phase transition and to silica glass. *Physical Review Letters*, 82(8), 1708–1711.
- [67] Di Felice, R. & Corni, S. (2011). Simulation of peptide–surface recognition. *Journal of Physical Chemistry Letters*, 2(13), 1510–1519.
- [68] Dickerson, M. B., Sandhage, K. H., & Naik, R. R. (2008). Protein- and peptide-directed syntheses of inorganic materials. *Chemical Reviews (Washington, DC, United States)*, 108(11), 4935–4978.
- [69] Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., & Kollman, P. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry*, 24(16), 1999–2012.
- [70] Dudko, O. K., Hummer, G., & Szabo, A. (2006). Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys. Rev. Lett.*, 96, 108101.
- [71] Dudko, O. K., Hummer, G., & Szabo, A. (2008). Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 105(41), 15755–15760.
- [72] El Shafei, M. S. (2000). Silica surface chemical properties. In *Adsorption on Silica Surfaces*, Surfactant Science chapter 3, (pp. 35–62). Taylor & Francis.
- [73] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh Ewald method. *Journal of Chemical Physics*, 103(19), 8577–8593.

- [74] Eu, B., Cairns, A., Ding, G., Cao, X., & Wen, Z.-Q. (2011). Direct visualization of protein adsorption to primary containers by gold nanoparticles. *Journal of Pharmaceutical Sciences*, 100(5), 1663–1670.
- [75] Evans, E. & Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. *Biophysical Journal*, 72(4), 1541–1555.
- [76] Ewald, P. P. (1921). Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik (Berlin, Germany)*, 369(3), 253–287.
- [77] Feuston, B. P. & Garofalini, S. H. (1988). Empirical three-body potential for vitreous silica. *Journal of Chemical Physics*, 89(9), 5818–5824.
- [78] Florin, E.-L., Rief, M., Lehmann, H., Ludwig, M., Dornmair, C., Moy, V., & Gaub, H. (1995). Sensing specific molecular interactions with the atomic force microscope. *Biosensors and Bioelectronics*, 10(9–10), 895–901.
- [79] Fogarty, J. C., Aktulga, H. M., Grama, A. Y., van Duin, A. C. T., & Pandit, S. A. (2010). A reactive molecular dynamics simulation of the silica-water interface. *Journal of Chemical Physics*, 132(17), 174704.
- [80] Frenkel, D. & Smit, B. (2001). *Understanding molecular simulation: From algorithms to applications*. Elsevier.
- [81] Friddle, R. W., Noy, A., & De Yoreo, J. J. (2012). Interpreting the widespread nonlinear force spectra of intermolecular bonds. *Proceedings of the National Academy of Sciences of the United States of America*, 109(34), 13573–13578.
- [82] Friddle, R. W., Weaver, M. L., Qiu, S. R., Wierzbicki, A., Casey, W. H., & De Yoreo, J. J. (2010). Subnanometer atomic force microscopy of peptide–mineral interactions links clustering and competition to acceleration and catastrophe. *Proceedings of the National Academy of Sciences of the United States of America*, 107(1), 11–15.
- [83] Fröhlich, H. (1949). *Theory of Dielectrics: Dielectric Constant and Dielectric Loss*. Monographs on the physics and chemistry of materials. Clarendon Press.
- [84] Fukunishi, H., Watanabe, O., & Takada, S. (2002). On the Hamiltonian replica exchange method for efficient sampling of biomolecular systems: Application to protein structure prediction. *Journal of Chemical Physics*, 116(20), 9058–9067.
- [85] Gear, C. W. (1971). *Numerical Initial Value Problems in Ordinary Differential Equations*. Englewood Cliffs, N.J., USA.

- [86] Giessibl, F. J. (2013). Seeing the reaction. *Science*, 340(6139), 1417–1418.
- [87] Gil-Ley, A. & Bussi, G. (2015). Enhanced conformational sampling using replica exchange with collective-variable tempering. *Journal of Chemical Theory and Computation*, 11(3), 1077–1085. PMID: 25838811.
- [88] Gokce, I., Woody, R. W., Anderluh, G., & Lakey, J. H. (2005). Single peptide bonds exhibit Poly(Pro)II (“Random Coil”) circular dichroism spectra. *Journal of the American Chemical Society*, 127(27), 9700–9701.
- [89] Grayson, A., Shawgo, R., Johnson, A., Flynn, N., Li, Y., Cima, M., & Langer, R. (2004). A biomems review: MEMS technology for physiologically integrated devices. *Proceedings of the IEEE*, 92(1), 6–21.
- [90] Grimley, D. I., Wright, A. C., & Sinclair, R. N. (1990). Neutron scattering from vitreous silica IV. Time-of-flight diffraction. *Journal of Non-Crystalline Solids*, 119(1), 49–64.
- [91] Gross, L., Mohn, F., Moll, N., Liljeroth, P., & Meyer, G. (2009). The chemical structure of a molecule resolved by atomic force microscopy. *Science*, 325(5944), 1110–1114.
- [92] Grubmüller, H., Heymann, B., & Tavan, P. (1996). Ligand binding: Molecular mechanics calculation of the streptavidin-biotin rupture force. *Science*, 271(5251), 997–999.
- [93] Hane, F. T., Attwood, S. J., & Leonenko, Z. (2014). Comparison of three competing dynamic force spectroscopy models to study binding forces of amyloid- β (1-42). *Soft Matter*, 10, 1924–1930.
- [94] Hartwig, A., Meissner, R., Merten, C., Schiffels, P., Wand, P., & Grunwald, I. (2013). Mutual influence between adhesion and molecular conformation: Molecular geometry is a key issue in interphase formation. *Journal of Adhesion*, 89(2), 77–95.
- [95] Hassanali, A. A. & Singer, S. J. (2007). Model for the water–amorphous silica interface: The undissociated surface. *Journal of Physical Chemistry B*, 111(38), 11181–11193.
- [96] Hassanali, A. A., Zhang, H., Knight, C., Shin, Y. K., & Singer, S. J. (2010). The dissociated amorphous silica surface: Model development and evaluation. *Journal of Chemical Theory and Computation*, 6(11), 3456–3471.

- [97] Hawe, A., Friess, W., Sutter, M., & Jiskoot, W. (2008). Online fluorescent dye detection method for the characterization of immunoglobulin G aggregation by size exclusion chromatography and asymmetrical flow field flow fractionation. *Analytical Biochemistry*, 378(2), 115–122.
- [98] Herce, H. D., Garcia, A. E., & Darden, T. (2007). The electrostatic surface term: (I) periodic systems. *Journal of Chemical Physics*, 126(12), 124106.
- [99] Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. E. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), 1463–1472.
- [100] Hildebrand, N., Köppen, S., Derr, L., Li, K., Koleini, M., Rezwan, K., & Colombi Ciacchi, L. (2015). Adsorption orientation and binding motifs of lysozyme and chymotrypsin on amorphous silica. *Journal of Physical Chemistry C*, 119(13), 7295–7307.
- [101] Hirst, J. D. & Besley, N. A. (1999). Response to “Comment on ‘improving protein circular dichroism calculations in the far-ultraviolet through reparameterizing the amide chromophore’ ” [J. Chem. Phys. 111(6):2844-2845, 1999]. *Journal of Chemical Physics*, 111(6), 2846.
- [102] Hockney, R. & Eastwood, J. (2010). *Computer Simulation Using Particles*. Taylor & Francis.
- [103] Holleman, A., Wiberg, E., & Wiberg, N. (1995). *Lehrbuch der anorganischen Chemie*. de Gruyter.
- [104] Holmberg, K. (2007). Proteins at interfaces: Control of adsorption. In *Encyclopedia of Surface and Colloid Science, Second Edition* chapter 419, (pp. 5288–5298). Taylor & Francis.
- [105] Hoover, W. G. (1985). Canonical dynamics: Equilibrium phase-space distributions. *Physical Review A: Atomic, Molecular, and Optical Physics*, 31, 1695–1697.
- [106] Horinek, D., Herz, A., Vrbka, L., Sedlmeier, F., Mamatkulov, S. I., & Netz, R. R. (2009). Specific ion adsorption at the air/water interface: The role of hydrophobic solvation. *Chemical Physics Letters*, 479(4–6), 173–183.
- [107] Horinek, D., Serr, A., Geisler, M., Pirzer, T., Slotta, U., Lud, S. Q., Garrido, J. A., Scheibel, T., Hugel, T., & Netz, R. R. (2008). Peptide adsorption on a hydrophobic

- surface results from an interplay of solvation, surface, and intrapeptide forces. *Proceedings of the National Academy of Sciences of the United States of America*, 105(8), 2842–2847.
- [108] Horn, H. W., Swope, W. C., Pitner, J. W., Madura, J. D., Dick, T. J., Hura, G. L., & Head-Gordon, T. (2004). Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *Journal of Chemical Physics*, 120(20), 9665–9678.
- [109] Hummer, G. & Szabo, A. (2001). Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 98(7), 3658–3661.
- [110] Iannuzzi, M., Laio, A., & Parrinello, M. (2003). Efficient exploration of reactive potential energy surfaces using car-parrinello molecular dynamics. *Physical Review Letters*, 90, 238302.
- [111] Iliafar, S., Vezenov, D., & Jagota, A. (2013). Brownian dynamics simulation of peeling a strongly-adsorbed polymer molecule from a frictionless substrate. *Langmuir*, 29(5), 1435–1445.
- [112] Iliafar, S., Wagner, K., Manohar, S., Jagota, A., & Vezenov, D. (2012). Quantifying interactions between dna oligomers and graphite surface using single molecule force spectroscopy. *Journal of Physical Chemistry C*, 116(26), 13896–13903.
- [113] Jarzynski, C. (1997). Nonequilibrium equality for free energy differences. *Physical Review Letters*, 78(14), 2690–2693.
- [114] Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79(2), 926–935.
- [115] Jorgensen, W. L., Maxwell, D. S., & Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45), 11225–11236.
- [116] Kim, J., Kim, S., Ohashi, T., Muramatsu, H., Chang, S.-M., & Kim, W.-S. (2010). Construction of simultaneous spr and qcm sensing platform. *Bioprocess and Biosystems Engineering*, 33(1), 39–45.

- [117] Kingshott, P. & Höcker, H. (2007). Adsorption of proteins: Measurement methods. In *Encyclopedia of Surface and Colloid Science, Second Edition* chapter 49, (pp. 669–694). Taylor & Francis.
- [118] Kosfeld, M., Heinrichs, M., Zak, P. J., Fischbacher, U., & Fehr, E. (2005). Oxytocin increases trust in humans. *Nature*, 435(7042), 673–676.
- [119] Kößlinger, C., Uttenthaler, E., Drost, S., Aberl, F., Wolf, H., Brink, G., Stanglmaier, A., & Sackmann, E. (1995). Comparison of the {QCM} and the {SPR} method for surface studies and immunological applications. *Sensors and Actuators B: Chemical*, 24(1-3), 107–112.
- [120] Kosztin, I., Barz, B., & Janosi, L. (2006). Calculating potentials of mean force and diffusion coefficients from nonequilibrium processes without Jarzynski's equality. *The Journal of Chemical Physics*, 124(6), 064106.
- [121] Kreuzer, H., Payne, S., & Livadaru, L. (2001). Stretching a macromolecule in an atomic force microscope: Statistical mechanical analysis. *Biophysical Journal*, 80(6), 2505–2514.
- [122] Krysiak, S., Liese, S., Netz, R. R., & Hugel, T. (2014). Peptide desorption kinetics from single molecule force spectroscopy studies. *Journal of the American Chemical Society*, 136(2), 688–697. PMID: 24328353.
- [123] Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., & Kollman, P. A. (1992). The weighted histogram analysis method for free-energy calculations on biomolecules. I. the method. *Journal of Computational Chemistry*, 13(8), 1011–1021.
- [124] Laio, A. & Gervasio, F. L. (2008). Metadynamics: A method to simulate rare events and reconstruct the free energy in biophysics, chemistry and material science. *Reports on Progress in Physics*, 71(12), 126601.
- [125] Laio, A. & Parrinello, M. (2002). Escaping free-energy minima. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20), 12562–12566.
- [126] Landoulsi, J. & Dupres, V. (2011). Probing peptide–inorganic surface interaction at the single molecule level using force spectroscopy. *ChemPhysChem*, 12(7), 1310–1316.
- [127] Lassen, B. & Malmsten, M. (1997). Competitive protein adsorption at plasma polymer surfaces. *Journal of Colloid and Interface Science*, 186(1), 9–16.

- [128] Laube, J., Salameh, S., Kappl, M., Mädler, L., & Colombi Ciacchi, L. (2015). Contact forces between TiO₂ nanoparticles governed by an interplay of adsorbed water layers and roughness. (submitted).
- [129] Lee, H., Scherer, N. F., & Messersmith, P. B. (2006). Single-molecule mechanics of mussel adhesion. *Proceedings of the National Academy of Sciences of the United States of America*, 103(35), 12999–13003.
- [130] Levine, B. G., Stone, J. E., & Kohlmeyer, A. (2011). Fast analysis of molecular dynamics trajectories with graphics processing units—Radial distribution function histogramming. *Journal of Computational Physics*, 230(9), 3556–3569.
- [131] Liphardt, J. (2002). Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski’s equality. *Science*, 296(5574), 1832–1835.
- [132] Liu, P., Kim, B., Friesner, R. A., & Berne, B. J. (2005). Replica exchange with solute tempering: A method for sampling biological systems in explicit water. *Proceedings of the National Academy of Sciences of the United States of America*, 102(39), 13749–13754.
- [133] Loncharich, R. J. & Brooks, B. R. (1989). The effects of truncating long-range forces on protein dynamics. *Proteins: Structure, Function, and Bioinformatics*, 6(1), 32–45.
- [134] Lopes, P. E. M., Murashov, V., Tazi, M., Demchuk, E., & MacKerell, A. D. (2006). Development of an empirical force field for silica. Application to the quartz–water interface. *Journal of Physical Chemistry B*, 110(6), 2782–2792.
- [135] Lubarsky, G., Davidson, M., & Bradley, R. (2007). Hydration–dehydration of adsorbed protein films studied by AFM and QCM-D. *Biosensors and Bioelectronics*, 22(7), 1275–1281.
- [136] Luo, P. & Baldwin, R. L. (1997). Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from trifluoroethanol/water mixtures back to water †. *Biochemistry*, 36(27), 8413–8421.
- [137] Lv, J., Song, Y., Jiang, L., & Wang, J. (2014). Bio-inspired strategies for anti-icing. *ACS Nano*, 8(4), 3152–3169.
- [138] Maitra, A. & Arya, G. (2010). Model accounting for the effects of pulling-device stiffness in the analyses of single-molecule force measurements. *Physical Review Letters*, 104, 108301.

- [139] Malmsten, M. (1995). Ellipsometry studies of the effects of surface hydrophobicity on protein adsorption. *Colloids and Surfaces, B: Biointerfaces*, 3(5), 297–308.
- [140] Manohar, S., Mantz, A. R., Bancroft, K. E., Hui, C.-Y., Jagota, A., & Vezenov, D. V. (2008). Peeling single-stranded dna from graphite surface to determine oligonucleotide binding energy by force spectroscopy. *Nano Letters*, 8(12), 4365–4372.
- [141] Marinelli, F., Pietrucci, F., Laio, A., & Piana, S. (2009). A kinetic model of trp-cage folding from multiple biased molecular dynamics simulations. *PLoS Computational Biology*, 5(8), e1000452.
- [142] Mark, P. & Nilsson, L. (2001). Structure and dynamics of the TIP3P, SPC, and SPC/E water models at 298 K. *Journal of Physical Chemistry A*, 105(43), 9954–9960.
- [143] Marko, J. F. & Siggia, E. D. (1995). Stretching DNA. *Macromolecules*, 28(26), 8759–8770.
- [144] Marrink, S.-J. & Berendsen, H. J. C. (1994). Simulation of water transport through a lipid membrane. *Journal of Physical Chemistry*, 98(15), 4155–4168.
- [145] Martyna, G. J., Klein, M. L., & Tuckerman, M. (1992). Nosé–Hoover chains: The canonical ensemble via continuous dynamics. *Journal of Chemical Physics*, 97(4), 2635–2643.
- [146] Meißner, R. H., Schneider, J., Schiffels, P., & Colombi Ciacchi, L. (2014). Computational prediction of circular dichroism spectra and quantification of helicity loss upon peptide adsorption on silica. *Langmuir*, 30(12), 3487–3494.
- [147] Mischler, C., Kob, W., & Binder, K. (2002). Classical and ab-initio molecular dynamic simulation of an amorphous silica surface. *Computer Physics Communications*, 147(1–2), 222–225.
- [148] Miyamoto, S. & Kollman, P. A. (1992). Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8), 952–962.
- [149] Moses, J. E. & Moorhouse, A. D. (2007). The growing applications of click chemistry. *Chem. Soc. Rev.*, 36, 1249–1262.
- [150] Mozzi, R. L. & Warren, B. E. (1969). The structure of vitreous silica. *Journal of Applied Crystallography*, 2(4), 164–172.

- [151] Müller, Werner, E., Wang, X., Proksch, P., Perry, Carole, C., Osinga, R., Gardères, J., & Schröder, Heinz, C. (2013). Principles of biofouling protection in marine sponges: A model for the design of novel biomimetic and bio-inspired coatings in the marine environment? *Marine Biotechnology*, 15(4), 375–398.
- [152] Nafie, L. A., Cheng, J. C., & Stephens, P. J. (1975). Vibrational circular dichroism of 2,2,2-trifluoro-1-phenylethanol. *Journal of the American Chemical Society*, 97(13), 3842–3843.
- [153] Nakanishi, K., Sakiyama, T., & Imamura, K. (2001). On the adsorption of proteins on solid surfaces, a common but very complicated phenomenon. *Journal of Bio-science and Bioengineering*, 91(3), 233–244.
- [154] Neuman, K. C. & Nagy, A. (2008). Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nature Methods*, 5(6), 491–505.
- [155] Nielson, K. D., van Duin, A. C. T., Oxgaard, J., Deng, W.-Q., & Goddard, W. A. (2005). Development of the ReaxFF reactive force field for describing transition metal catalyzed reactions, with application to the initial stages of the catalytic formation of carbon nanotubes. *Journal of Physical Chemistry A*, 109(3), 493–499.
- [156] Norde, W. (2000). Proteins at solid surfaces. In *Physical Chemistry of Biological Interfaces* (pp. 115). Taylor & Francis.
- [157] Nosé, S. (1984a). A molecular dynamics method for simulations in the canonical ensemble. *Molecular Physics*, 52(2), 255–268.
- [158] Nosé, S. (1984b). A unified formulation of the constant temperature molecular dynamics methods. *Journal of Chemical Physics*, 81(1), 511–519.
- [159] Nylander, T. (2003). *Biopolymers at Interfaces, Second Edition*, chapter Protein Adsorption in Relation to Solution Association and Aggregation. Surfactant Science. CRC Press.
- [160] Oakley, M. T., Bulheller, B. M., & Hirst, J. D. (2006). First-principles calculations of protein circular dichroism in the far-ultraviolet and beyond. *Chirality*, 18(5), 340–347.
- [161] O'Brien, C. P., Stuart, S. J., Bruce, D. A., & Latour, R. A. (2008). Modeling of peptide adsorption interactions with a poly(lactic acid) surface. *Langmuir*, 24(24), 14115–14124.

- [162] Oesterhelt, F., Rief, M., & Gaub, H. E. (1999). Single molecule force spectroscopy by afm indicates helical structure of poly(ethylene-glycol) in water. *New Journal of Physics*, 1, 6–6.
- [163] Oren, E. E., Notman, R., Kim, I. W., Evans, J. S., Walsh, T. R., Samudrala, R., Tamerler, C., & Sarikaya, M. (2010). Probing the molecular mechanisms of quartz-binding peptides. *Langmuir*, 26(13), 11003–11009.
- [164] Park, S., Khalili-Araghi, F., Tajkhorshid, E., & Schulten, K. (2003). Free energy calculation from steered molecular dynamics simulations using jarzynski's equality. *Journal of Chemical Physics*, 119(6), 3559.
- [165] Paschek, D., Day, R., & Garcia, A. E. (2011). Influence of water-protein hydrogen bonding on the stability of trp-cage miniprotein. a comparison between the TIP3P and TIP4P-Ew water models. *Physical Chemistry Chemical Physics*, 13, 19840–19847.
- [166] Paturej, J., Dubbeldam, J. L. A., Rostiashvili, V. G., Milchev, A., & Vilgis, T. A. (2014). Force spectroscopy of polymer desorption: theory and molecular dynamics simulations. *Soft Matter*, 10, 2785–2799.
- [167] Patwardhan, S. V., Emami, F. S., Berry, R. J., Jones, S. E., Naik, R. R., Deschaume, O., Heinz, H., & Perry, C. C. (2012). Chemistry of aqueous silica nanoparticle surfaces and the mechanism of selective peptide adsorption. *Journal of the American Chemical Society*, 134(14), 6244–6256.
- [168] Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., III, T. E. C., DeBolt, S., Ferguson, D., Seibel, G., & Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications*, 91(1–3), 1–41.
- [169] Pedone, A. (2009). Properties calculations of silica-based glasses by atomistic simulations techniques: A review. *Journal of Physical Chemistry C*, 113(49), 20773–20784.
- [170] Piana, S. & Laio, A. (2007). A bias-exchange approach to protein folding. *Journal of Physical Chemistry B*, 111(17), 4553–4559.
- [171] Plimpton, S. (1995). Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1), 1–19.
- [172] Polavarapu, P. L. (2005). Kramers–Kronig transformation for optical rotatory dispersion studies. *Journal of Physical Chemistry A*, 109(32), 7013–7023.

- [173] Price, D. J. & Brooks, C. L. (2004). A modified TIP3P water potential for simulation with Ewald summation. *Journal of Chemical Physics*, 121(20), 10096.
- [174] Rabe, M., Verdes, D., & Seeger, S. (2011). Understanding protein adsorption phenomena at solid surfaces. *Advances in Colloid and Interface Science*, 162(1-2), 87–106.
- [175] Raffaini, G. & Ganazzoli, F. (2004). Molecular dynamics simulation of the adsorption of a fibronectin module on a graphite surface †. *Langmuir*, 20(8), 3371–3378.
- [176] Raffaini, G. & Ganazzoli, F. (2007). Understanding the performance of biomaterials through molecular modeling: Crossing the bridge between their intrinsic properties and the surface adsorption of proteins. *Macromolecular Bioscience*, 7(5), 552–566.
- [177] Rahaman, O., van Duin, A. C. T., Goddard, W. A., & Doren, D. J. (2011). Development of a ReaxFF reactive force field for glycine and application to solvent effect and tautomerization. *Journal of Physical Chemistry B*, 115(2), 249–261.
- [178] Rappe, A. K. & Goddard, W. A. (1991). Charge equilibration for molecular dynamics simulations. *Journal of Physical Chemistry*, 95(8), 3358–3363.
- [179] Rarivomanantsoa, M., Jund, P., & Jullien, R. (2001). Classical molecular dynamics simulations of amorphous silica surfaces. *Journal of Physics: Condensed Matter*, 13(31), 6707.
- [180] Raut, V. P., Agashe, M. A., Stuart, S. J., & Latour, R. A. (2005). Molecular dynamics simulations of peptide–surface interactions. *Langmuir*, 21(4), 1629–1639.
- [181] Read, M. J. & Burkett, S. L. (2003). Asymmetric α -helicity loss within a peptide adsorbed onto charged colloidal substrates. *Journal of Colloid and Interface Science*, 261(2), 255–263.
- [182] Reiher, W. E. I. (1985). *Theoretical Studies of Hydrogen Bonding*. PhD thesis, Harvard University.
- [183] Rico, F., Gonzalez, L., Casuso, I., Puig-Vidal, M., & Scheuring, S. (2013). High-speed force spectroscopy unfolds titin at the velocity of molecular dynamics simulations. *Science*, 342(6159), 741–743.
- [184] Rigby, M., Smith, E. B., Wakeham, W. A., & Maitland, G. C. (1986). *The forces between molecules*. Clarendon Press Oxford.

- [185] Rodger, A., Rajendra, J., Marrington, R., Ardhammar, M., Nordén, B., Hirst, J. D., Gilbert, A. T. B., Dafforn, T. R., Halsall, D. J., Woolhead, C. A., & et al. (2002). Flow oriented linear dichroism to probe protein orientation in membrane environments. *Physical Chemistry Chemical Physics*, 4(16), 4051–4057.
- [186] Rosenfeld, L. (1929). Quantenmechanische Theorie der natürlichen optischen Aktivität von Flüssigkeiten und Gasen. *Zeitschrift für Physik*, 52(3–4), 161–174.
- [187] Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3), 327–341.
- [188] Rykart, R. (1995). *Quarz-Monographie: die Eigenheiten von Bergkristall, Rauchquarz, Amethyst, Chalcedon, Achat, Opal und anderen Varietäten*. Ott.
- [189] Salameh, S., Schneider, J., Laube, J., Alessandrini, A., Facci, P., Seo, J. W., Ciacchi, L. C., & Mädler, L. (2012). Adhesion mechanisms of the contact interface of tio₂ nanoparticles in films and aggregates. *Langmuir*, 28(31), 11457–11464.
- [190] Schneider, J. & Colombi Ciacchi, L. (2012). Specific material recognition by small peptides mediated by the interfacial solvent structure. *Journal of the American Chemical Society*, 134(4), 2407–2413.
- [191] Secundo, F. (2013). Conformational changes of enzymes upon immobilisation. *Chemical Society Reviews*, 42, 6250–6261.
- [192] Serro, A., Degiampietro, K., Colaço, R., & Saramago, B. (2010). Adsorption of albumin and sodium hyaluronate on UHMWPE: A QCM-D and AFM study. *Colloids and Surfaces B: Biointerfaces*, 78(1), 1–7.
- [193] Service, R. F. (2002). Can sensors make a home in the body? *Science*, 297(5583), 962–963.
- [194] Severin, J. W. & Vankan, J. M. J. (1990). Evidence for observation of silanetriol groups on silica surfaces by solid-state NMR. *Philips Journal of Research*, 45(1), 35–39.
- [195] Skelton, A. A., Fenter, P., Kubicki, J. D., Wesolowski, D. J., & Cummings, P. T. (2011a). Simulations of the quartz(10 $\bar{1}$ 1)/water interface: A comparison of classical force fields, ab initio molecular dynamics, and x-ray reflectivity experiments. *Journal of Physical Chemistry C*, 115(5), 2076–2088.

- [196] Skelton, A. A., Wesolowski, D. J., & Cummings, P. T. (2011b). Investigating the quartz (10 $\bar{1}$ 1)/water interface using classical and ab initio molecular dynamics. *Langmuir*, 27(14), 8700–8709.
- [197] Smith, P. E. & Pettitt, B. M. (1991). Peptides in ionic solutions: A comparison of the Ewald and switching function techniques. *Journal of Chemical Physics*, 95(11), 8430–8441.
- [198] Snatzke, G. (1968). Circular dichroism and optical rotatory dispersion— principles and application to the investigation of the stereochemistry of natural products. *Angewandte Chemie, International Edition in English*, 7(1), 14–25.
- [199] Spiwok, V. & Tvaroška, I. (2009). Metadynamics modelling of the solvent effect on primary hydroxyl rotamer equilibria in hexopyranosides. *Carbohydrate Research*, 344(12), 1575–1581.
- [200] Sreerama, N. & Woody, R. W. (2004). Computation and analysis of protein circular dichroism spectra. *Methods in Enzymology*, (pp. 318–351).
- [201] Steckbeck, S., Schneider, J., Wittig, L., Rischka, K., Grunwald, I., & Ciacchi, L. C. (2014). Identification of materials' binding peptide sequences guided by a maldi-tof ms depletion assay. *Analytical Methods*, 6(5), 1501.
- [202] Steinbach, P. J. & Brooks, B. R. (1994). New spherical-cutoff methods for long-range forces in macromolecular simulation. *Journal of Computational Chemistry*, 15(7), 667–683.
- [203] Stillinger, F. H. & Weber, T. A. (1985). Computer simulation of local order in condensed phases of silicon. *Physical Review B: Solid State*, 31, 5262–5271.
- [204] Strachan, A., van Duin, A. C. T., Chakraborty, D., Dasgupta, S., & Goddard, W. A. (2003). Shock waves in high-energy materials: The initial chemical events in nitramine RDX. *Physical Review Letters*, 91, 098301.
- [205] Sugita, Y. & Okamoto, Y. (1999). Replica-exchange molecular dynamics method for protein folding. *Chemical Physics Letters*, 314(1–2), 141–151.
- [206] Sushko, M. L., Gal, A. Y., & Shluger, A. L. (2006). Interaction of organic molecules with the TiO₂ (110) surface: Ab initio calculations and classical force fields. *Journal of Physical Chemistry B*, 110(10), 4853–4862.

- [207] Swope, W. C., Andersen, H. C., Berens, P. H., & Wilson, K. R. (1982). A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. *Journal of Chemical Physics*, 76(1), 637–649.
- [208] Takada, A., Richet, P., Catlow, C., & Price, G. (2004). Molecular dynamics simulations of vitreous silica structures. *Journal of Non-Crystalline Solids*, 345–346(0), 224–229.
- [209] Takeda, S., Ptak, A., Nakamura, C., Miyake, J., Kageshima, M., Jarvis, S. P., & Tokumoto, H. (2001). Measurement of the length of the α -helical section of a peptide directly using atomic force microscopy. *Chemical and Pharmaceutical Bulletin*, 49(12), 1512–1516.
- [210] Takei, T. & Chikazawa, M. (1998). Origin of differences in heats of immersion of silicas in water. *Journal of Colloid and Interface Science*, 208(2), 570–574.
- [211] Tamerler, C. & Sarikaya, M. (2009). Genetically designed peptide-based molecular materials. *ACS Nano*, 3(7), 1606–1615.
- [212] Tang, Z., Palafox-Hernandez, J. P., Law, W.-C., E. Hughes, Z., Swihart, M. T., Prasad, P. N., Knecht, M. R., & Walsh, T. R. (2013). Biomolecular recognition principles for bionanocombinatorics: An integrated approach to elucidate enthalpic and entropic factors. *ACS Nano*, 7(11), 9632–9646.
- [213] Thormann, E., Hansen, P. L., Simonsen, A. C., & Mouritsen, O. G. (2006). Dynamic force spectroscopy on soft molecular systems: Improved analysis of unbinding spectra with varying linker compliance. *Colloids and Surfaces, B: Biointerfaces*, 53(2), 149–156.
- [214] Thyparambil, A. A., Wei, Y., & Latour, R. A. (2012). Determination of peptide/surface adsorption free energy for material surfaces not conducive to spr or qcm using afm. *Langmuir*, 28(13), 5687–5694.
- [215] Tiffany, M. L. & Krimm, S. (1972). Effect of temperature on the circular dichroism spectra of polypeptides in the extended state. *Biopolymers*, 11(11), 2309–2316.
- [216] Tinoco, I. (1962). Theoretical aspects of optical activity part two: Polymers. *Advances in Chemical Physics*, (pp. 113–160).

- [217] Torrie, G. & Valleau, J. (1977). Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling. *Journal of Computational Physics*, 23(2), 187–199.
- [218] Toxvaerd, S. & Olsen, O. H. (1990). Canonical molecular dynamics of molecules with internal degrees of freedom. *Berichte der Bunsengesellschaft für physikalische Chemie*, 94(3), 274–278.
- [219] Trioni, M., Bongiorno, A., & Colombo, L. (1997). Structural properties of silica surface: A classical molecular dynamics study. *Journal of Non-Crystalline Solids*, 220(2–3), 164–168.
- [220] Tuckerman, M., Berne, B. J., & Martyna, G. J. (1992). Reversible multiple time scale molecular dynamics. *Journal of Chemical Physics*, 97(3), 1990–2001.
- [221] van Beest, B. W. H., Kramer, G. J., & van Santen, R. A. (1990). Force fields for silicas and aluminophosphates based on ab initio calculations. *Physical Review Letters*, 64(16), 1955–1958.
- [222] van Duin, A. C. T., Dasgupta, S., Lorant, F., & Goddard, W. A. (2001). ReaxFF: A reactive force field for hydrocarbons. *Journal of Physical Chemistry A*, 105(41), 9396–9409.
- [223] van Duin, A. C. T., Strachan, A., Stewman, S., Zhang, Q., Xu, X., & Goddard, W. A. (2003). ReaxFFSiO reactive force field for silicon and silicon oxide systems. *Journal of Physical Chemistry A*, 107(19), 3803–3811.
- [224] van Gunsteren, W. & Berendsen, H. (1977). Algorithms for macromolecular dynamics and constraint dynamics. *Molecular Physics*, 34(5), 1311–1327.
- [225] Varghese, N., Mogera, U., Govindaraj, A., Das, A., Maiti, P. K., Sood, A. K., & Rao, C. N. R. (2009). Binding of dna nucleobases and nucleosides with graphene. *ChemPhysChem*, 10(1), 206–210.
- [226] Vashishta, P., Kalia, R. K., & Rino, J. P. (1990). Interaction potential for SiO₂: A molecular-dynamics study of structural correlations. *Physical Review B: Condensed Matter and Materials Physics*, 41(17), 12197–12209.
- [227] Verlet, L. (1967). Computer “experiments” on classical fluids. I. Thermodynamical properties of lennard-jones molecules. *Physical Review*, 159, 98–103.

- [228] Vollmayr, K., Kob, W., & Binder, K. (1996). Cooling-rate effects in amorphous silica: A computer-simulation study. *Physical Review B: Condensed Matter and Materials Physics*, 54, 15808–15827.
- [229] Voskerician, G., Shive, M. S., Shawgo, R. S., Recum, H. v., Anderson, J. M., Cima, M. J., & Langer, R. (2003). Biocompatibility and biofouling of MEMS drug delivery devices. *Biomaterials*, 24(11), 1959–1967.
- [230] Wang, C., Kuzuu, N., & Tamai, Y. (2002). Effects of charge transfer on α -SiO₂ surface structure: A molecular dynamics study. *Journal of Applied Physics*, 92(8), 4408–4413.
- [231] Wang, C., Tamai, Y., & Kuzuu, N. (2003). A molecular dynamics study on vibration spectra of α -SiO₂ surface. *Journal of Non-Crystalline Solids*, 321(3), 204–209.
- [232] Wang, L., Friesner, R. A., & Berne, B. J. (2011). Replica exchange with solute scaling: A more efficient version of replica exchange with solute tempering (REST2). *Journal of Physical Chemistry B*, 115(30), 9431–9438.
- [233] Warshel, A. & Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *Journal of Molecular Biology*, 103(2), 227–249.
- [234] Wei, Y. & Latour, R. A. (2008). Determination of the adsorption free energy for peptide–surface interactions by SPR spectroscopy. *Langmuir*, 24(13), 6721–6729.
- [235] Wei, Y. & Latour, R. A. (2010). Correlation between desorption force measured by atomic force microscopy and adsorption free energy measured by surface plasmon resonance spectroscopy for peptide/surface interactions. *Langmuir*, 26(24), 18852–18861.
- [236] Wenk, H. & Bulakh, A. (2004). *Minerals: Their Constitution and Origin*. Cambridge University Press.
- [237] Whitmore, L. & Wallace, B. A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Research*, 32(suppl 2), W668–W673.
- [238] Wilson, E. B. (1941). Some mathematical methods for the study of molecular vibrations. *Journal of Chemical Physics*, 9(1), 76–84.

- [239] Woody, R. W. (1968). Improved calculation of the $n\pi^*$ rotational strength in polypeptides. *Journal of Chemical Physics*, 49(11), 4797.
- [240] Woody, R. W. & Sreerama, N. (1999). Comment on “Improving protein circular dichroism calculations in the far-ultraviolet through reparametrizing the amide chromophore” [J. Chem. Phys. 109(2):782–788, 1998]. *Journal of Chemical Physics*, 111(6), 2844–2845.
- [241] Woody, R. W. & Tinoco, I. J. (1967). Optical rotation of oriented helices. III. Calculation of the rotatory dispersion and circular dichroism of the α - and 3_{10} -Helix. *Journal of Chemical Physics*, 46(12), 4927–4945.
- [242] Yancey, J., Vellore, N., Collier, G., Stuart, S., & Latour, R. (2010). Development of molecular simulation methods to accurately represent protein-surface interactions: The effect of pressure and its determination for a system with constrained atoms. *Biointerphases*, 5, 85–95.
- [243] Zahn, D. (2004). Atomistic mechanism of nacl nucleation from an aqueous solution. *Phys. Rev. Lett.*, 92, 040801.
- [244] Zhao, G., Perilla, J. R., Yufenyuy, E. L., Meng, X., Chen, B., Ning, J., Ahn, J., Gronenborn, A. M., Schulten, K., Aiken, C., & et al. (2013). Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. *Nature*, 497(7451), 643–646.
- [245] Zhuravlev, L. (1993). Surface characterization of amorphous silica—a review of work from the former USSR. *Colloids and Surfaces, A: Physicochemical and Engineering Aspects*, 74(1), 71–90.
- [246] Zhuravlev, L. (2000). The surface chemistry of amorphous silica. Zhuravlev model. *Colloids and Surfaces, A: Physicochemical and Engineering Aspects*, 173(1–3), 1–38.

Erklärung

Hiermit erkläre ich, dass ich

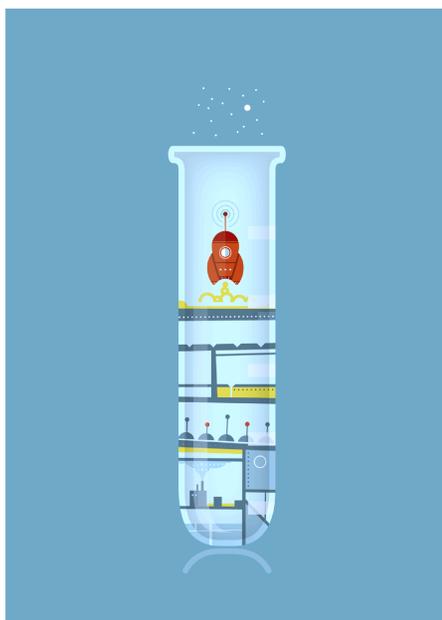
- die Arbeit ohne unerlaubte fremde Hilfe angefertigt habe;
- keine anderen als die von mir angegebenen Quellen und Hilfsmittel benutzt habe;
- die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

In der vorliegenden Arbeit sind Ergebnisse enthalten, die im Rahmen der Betreuung folgender studentischer Arbeiten entstanden sind:

- Nina Wurzler, Untersuchungen der Struktur und des Adsorptionsverhaltens von Biomolekülen auf Nanopartikeln mittels Zirkulardichroismus-Spektroskopie, 2015
- Christian Perl, Strukturprognose des Peptids GB-1 - Metadynamics kombiniert mit Replica Exchange und Solute Tempering, 2014

Bremen, den 29. Juni 2015

Robert Meißner



THIS THESIS WAS TYPESET using \LaTeX , originally developed by Leslie Lamport and based on Donald Knuth's \TeX . The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. The above illustration, "Science Experiment 02", was created by Ben Schlitter and released under CC BY-NC-ND 3.0. A template that can be used to format a PhD thesis with this look and feel has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate or from its author, Jordan Suchow, at suchow@post.harvard.edu.