

Computational Modeling of Surface Interactions

Dissertation
zur Erlangung des Grades

“Doktor der Naturwissenschaften”

am Fachbereich Chemie, Pharmazie und Geowissenschaften
der Johannes Gutenberg-Universität Mainz

Willem Jan Schravendijk
geb. in Amsterdam, die Niederlande

Mainz, Februar 2007

Zusammenfassung

Die Wechselwirkung zwischen Proteinen und anorganischen Oberflächen fasziniert sowohl aus angewandter als auch theoretischer Sicht. Sie ist ein wichtiger Aspekt in vielen Anwendungen, unter anderem in chirurgischen Implantaten oder Biosensoren. Sie ist außerdem ein Beispiel für theoretische Fragestellungen betreffend die Grenzfläche zwischen harter und weicher Materie. Fest steht, dass Kenntnis der beteiligten Mechanismen erforderlich ist um die Wechselwirkung zwischen Proteinen und Oberflächen zu verstehen, vorherzusagen und zu optimieren.

Aktuelle Fortschritte im experimentellen Forschungsbereich ermöglichen die Untersuchung der direkten Peptid-Metall-Bindung. Dadurch ist die Erforschung der theoretischen Grundlagen weiter ins Blickfeld aktueller Forschung gerückt.

Eine Möglichkeit die Wechselwirkung zwischen Proteinen und anorganischen Oberflächen zu erforschen ist durch Computersimulationen. Obwohl Simulationen von Metalloberflächen oder Proteinen als Einzelsysteme schon länger verbreitet sind, bringt die Simulation einer Kombination beider Systeme neue Schwierigkeiten mit sich. Diese zu überwinden erfordert ein Mehrskalen-Verfahren: Während Proteine als biologische Systeme ausreichend mit klassischer Molekulardynamik beschrieben werden können, bedarf die Beschreibung delokalisierter Elektronen metallischer Systeme eine quantenmechanische Formulierung. Die wichtigste Voraussetzung eines Mehrskalen-Verfahrens ist eine Übereinstimmung der Simulationen auf den verschiedenen Skalen. In dieser Arbeit wird dies durch die Verknüpfung von Simulationen alternierender Skalen erreicht.

Diese Arbeit beginnt mit der Untersuchung der Thermodynamik der Benzol-Hydratation mittels klassischer Molekulardynamik. Dann wird die Wechselwirkung zwischen Wasser und den [111]-Metalloberflächen von Gold und Nickel mittels eines Multiskalen-Verfahrens modelliert. In einem weiteren Schritt wird die Adsorption des Benzols an Metalloberflächen in wässriger Umgebung studiert.

Abschließend wird die Modellierung erweitert und auch die Aminosäuren Alanin und Phenylalanin einbezogen. Dies eröffnet die Möglichkeit realistische Protein-Metall-Systeme in Computersimulationen zu betrachten und auf theoretischer Basis die Wechselwirkung zwischen Peptiden und Oberflächen für jede Art Peptide und Oberfläche vorauszusagen.

Abstract

The interaction of proteins with inorganic surfaces is fascinating from various points of view. As an application, it forms the essential working mechanism in systems like biosensors and surgical implants. As a theoretical problem, it describes a complex interface between hard and soft matter. In all cases, it is clear that theoretical knowledge of the mechanisms involved is needed to understand, predict, and optimize protein-surface interactions.

Recent experimental advancements have enabled the research of direct interactions of peptide groups with metal surfaces, and, with that information as a reference, it becomes possible to investigate the theoretical basis of protein-metal interactions.

One way to study this is via computer simulations. Computer simulations of either solids or biological systems are common, but simulating both systems combined introduces new problems, for which simulations at several levels of detail will be needed. Simulating the behavior of delocalized electrons in the metal will require quantum mechanical treatment, whereas biological systems are best described by classical statistical mechanics. Protein-metal systems form therefore a typical modeling problem for which so-called multiscale simulations are needed. In a multiscale modeling approach, simulations at the multiple scales of interest need to be connected in such a way that a consistent picture can be attained. This will be done in the current work by connecting calculations on the quantum and the atomistic level in a sequential, alternating, manner.

As a start, the thermodynamic properties of hydration of benzene is studied via classical statistical mechanics approaches and computer simulations. Then, the interaction of water with gold and nickel [111]-surfaces is modeled by including quantum calculation data via a newly introduced multiscale procedure. As a next step, these two systems are combined and the multiscale procedure is extended to study benzene surface adsorption in an aqueous environment.

Finally, the modeling is expanded to include the amino acids alanine and phenylalanine, for which differences in the metal adsorption properties can be identified. This opens the way to study realistic protein-metal systems, which will enable the theoretical prediction of surface attraction for a given peptide, at a given surface.

Contents

1	Introduction	5
2	Protein-surface interactions, an overview	11
2.1	Experimental studies of biomolecule-surface interactions	11
2.1.1	Experimental techniques	12
2.1.2	Commonly used inorganic surfaces	12
2.1.3	Experimental research of direct peptide-surface interactions . .	13
2.2	Modeling biomolecule-surface interactions	16
2.2.1	Analytical models	16
2.2.2	Monte Carlo simulations	16
2.2.3	Molecular dynamics simulations	17
2.2.4	Quantum calculations	18
3	Theoretical basis of Molecular Simulation	21
3.1	Mechanics of Molecular Systems	21
3.2	Molecular force fields	22
3.2.1	Electrostatics	22
3.2.2	Dispersion	23
3.2.3	Surface potentials	24
3.3	Molecular Dynamics	25
3.3.1	Time Step Integration	26
3.3.2	Treatment of Bonds and Angles	26
3.4	Periodic Boundary Conditions	27
3.5	Thermal Coupling	28
3.5.1	Langevin Thermostat Molecular Dynamics	28
3.5.2	Berendsen	29
3.5.3	Nosé-Hoover	30
3.6	Pressure Coupling	30

3.6.1	Berendsen	30
3.6.2	Parrinello-Rahman	31
3.7	Long range electrostatics	32
3.7.1	Reaction-field	32
3.7.2	Ewald summation	34
3.8	Quantum calculations	35
3.9	Multiscale Modeling	37
4	Free Energy in Statistical Mechanics	39
4.1	Determining free energy changes	41
4.1.1	Thermodynamic perturbation	41
4.1.2	Thermodynamic integration	42
4.1.3	Obtaining a free energy along an order parameter	43
4.2	Distinction of enthalpy-entropy compensation in the free energy	44
5	A thermodynamic analysis of benzene solvation	49
5.1	Benzene as a hydrophilic molecule	49
5.2	Thermodynamics	51
5.3	Computational details	53
5.3.1	Thermodynamic calculations	53
5.3.2	Simulation Details.	54
5.3.3	Benzene-Water Interaction.	55
5.4	Results and Discussion	56
5.4.1	Hydration Structure: Radial distribution functions	56
5.4.2	Orientalional Distributions	58
5.4.3	Hydration Thermodynamics	61
5.4.4	Solute-Solvent Contributions	62
5.5	Discussion	67
6	A dual-scale modeling of aqueous-metal surface interactions.	69
6.1	Water-surface interactions, an overview	70
6.1.1	Experimental water-metal	70
6.1.2	Quantum calculations of water-metal interactions	71
6.1.3	Atomistic simulation of water near surfaces	71
6.2	Quantum based modeling of liquid water-metal interaction: Basic idea	72
6.3	Simulation details	76

6.4	The modeling idea applied to different water force fields	77
6.5	Discussion	82
7	Modeling the interaction of hydrated Benzene and Phenol with metal surfaces.	85
7.1	Benzene and phenol as side chain analogs	85
7.2	Methodology	86
7.2.1	First principles DFT-based modeling	86
7.2.2	Molecular dynamics implementation	87
7.2.3	Benzene modeling	88
7.2.4	Phenol modeling	89
7.3	Results and Discussion	91
7.3.1	Density profiles	92
7.3.2	PMF and umbrella sampling	93
7.3.3	Influence of the water model	96
7.3.4	Mechanisms	98
7.4	Discussion	101
8	Modeling hydrated amino acids near metal surfaces	103
8.1	Computational Details	103
8.2	Modeling	106
8.2.1	Neutral Alanine	107
8.2.2	Neutral Phenylalanine and Tyrosine	112
8.3	Analysis of Surface Interactions at the Ni(111) / H ₂ O Interface	113
8.3.1	Amino acids	114
8.4	Discussion	121
	Summary and Outlook	123
	List of Publications	129
	Bibliography	131

1 Introduction

This thesis will present a new computer simulation method that uses multiple levels of resolution to study biomolecular interactions with metal surfaces in the presence of water (see figure 1.1). To understand why there is a need to introduce a new method for this purpose, it will be necessary to evaluate the existing methods commonly available in computer simulations, and the requirements needed by the specific application.

At the level of analytical theory, the physics of a molecular system is described by

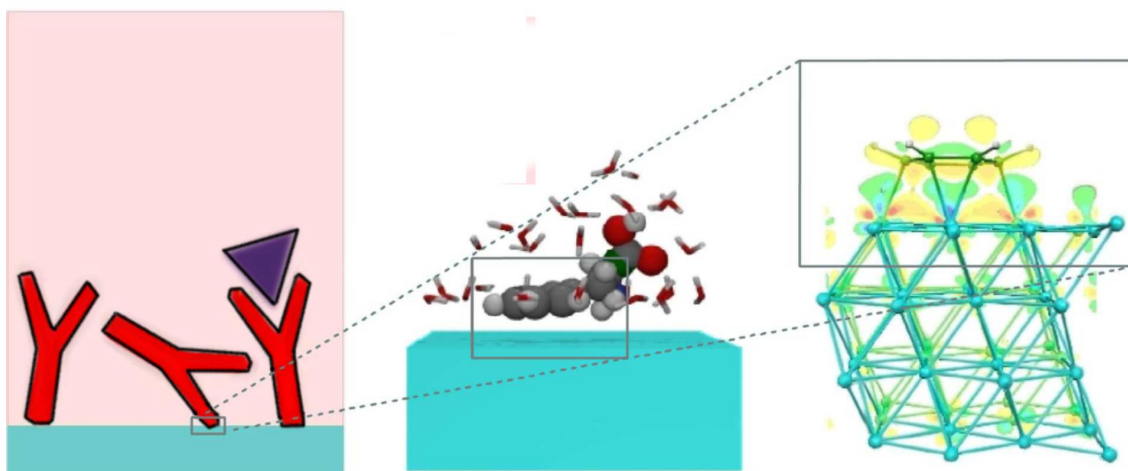


Figure 1.1: A schematic view of the multiscale procedure applied in this thesis. An application could be the study of antibody-surface interactions in a biosensor (left frame). It will at least be necessary to research the interaction mechanism at an atomistic level, for the part of the protein that is closest to the surface (middle frame). For the sake of calculation speed, one could then neglect effects from all other parts of the protein. As will be shown later on, the interaction of a molecule with a metal surface can not simply be seen as a generalized atom-atom interaction, but will depend specifically on the perturbation of the electron-distributions in the metal surface due to the nearby molecule. To obtain chemically and physically correct data, calculations at the quantum level will therefore be necessary (right frame).

a set of equations that can be solved on a computer. These kinds of calculations are useful to obtain general physical properties of molecular systems, such as polymer melts, liquid crystals, and many other systems. Often, the same systems can also be studied by making a simple particle-based model system for the set of molecules and calculating the inter-particle interactions and forces in this system. Atoms or groups of atoms can be represented by 'beads', the connectivity (topology) between these 'beads' is approximated by 'springs'. Both the analytical and bead-and-spring models can be used for fast calculations and deliver a range of useful properties of molecular systems (orientations, viscosity, phase diagrams).

These approaches might not be sufficient for all molecular systems, and one might need to consider the chemical composition of the molecules in the system. In that case, one can represent the different atom types that are present by sets of attractive and repulsive potential energy functions. This concept, called atomistic simulation, has gained considerable popularity in the field of biomolecular simulations. A typical example is the calculation of the interaction of a drug with the receptor protein it should act upon. Even though these systems contain considerable chemical detail, they still contain several approximations, for example molecular topologies and therefore the bonds are fixed, chemical reactions can not take place. The simulation of a chemical reaction will require extensive knowledge about the electronic states of the reacting molecules during the reaction. This is not contained in an atomistic simulation, but can be calculated by quantum calculations. Similar problems arise when studying metal surfaces as substrates for organic/inorganic hybrid materials, a popular combination due to the conducting properties of the surface, enabling the combination of biomolecules with electronic devices. Atomistic and molecular modeling of proteins near these metal surfaces is not straightforward. At the electronic and atomistic level, the interactions between a molecule and a metal surface are driven by the overlap and perturbation of the electrons in the molecular orbitals with the delocalized electrons in the conducting band of the metal. Clearly, quantum simulations are needed on this scale.

The modeling becomes even more difficult for these systems since the chemical and electronic interactions at the quantum scale influence, and in turn are influenced by, the bulk atomistic or larger-scale statistical properties of the many-molecule system, where molecule-molecule interactions and molecular conformations at the surface determine the adsorption properties. Therefore, so-called multiscale

simulations are developed to combine the specificity of more detailed atomistic or quantum simulations with the speed of coarse grained or analytical calculations[1, 2]. Ideally, the separation between the less and more detailed scales would be possible in a fully automated way, and the system would adapt to the desired detail during the simulation[3–5].

As an example, a non-biological system where the usefulness of a multiscale simulation was demonstrated was the simulation study on the interaction of polycarbonate (PC) with a Ni(111) surface.[6–8] There it was shown that the melt morphology close to the surface crucially depends on the choice of chain ends during the polymerization process. These results could be obtained with the use of first-principles Density Functional (DFT) calculations for the surface interaction of subgroups of the PC monomer. These quantum calculations were performed while considering molecular geometries consistent with the topologically allowed PC configurations at the surface, and combined with a study of the conformation properties at the surface as described by a suitably parameterized coarse grained model for the PC chains.

In a similar way also the adsorption out of solution, which is the more appropriate problem for biomolecules, can be tackled. This is however a research area which is still in its infancy. Similar to the situation in polymer melts, the interaction of small organic molecules with a metal surface in the presence of water has to be studied as a first step, to gain a deeper understanding of the underlying processes and mechanisms. This includes researching the thermodynamics of the adsorption-desorption equilibrium as well as retrieving geometrical information near the surface. Once the methodologies have been developed to study this for small molecular fragments, an extension to macromolecules will be possible. A first step, namely the adsorption of benzene, phenol, and several amino acids out of water onto a Ni(111) or Au(111) surface, will be presented in the current thesis. Gold is chosen as it is a commonly used material in hybrid organic/inorganic systems[9]. Nickel was chosen because quantum data and a dual scale modeling had been already developed for this metal, for the research of polymer blend properties near the nickel surface[6, 7].

A more general goal, however, is the rational procedure for the on-demand design of peptides that specifically bind to inorganic surfaces (aptamers). Such a procedure enables the construction of protein-inorganic composite materials, combining optimal solute binding specificity in proteins with the optimal signaling properties of inorganic materials[10]. Applications of these aptamer-inorganic surfaces are abun-

dant, examples include bio-analytical devices such as biochemical sensors (e.g. as applied in Surface Plasmon Resonance,[11, 12]) bio-electronic switches and gates, chemical separation and purification surfaces, and enzymatically controlled electrochemical interfaces[13].

The efficiency and small scales found in biology have been endorsed by material science in its quest to create functional materials and miniaturized devices, exploring ways to mimic biological systems[14–16]. A large amount of work in nanotechnology is therefore focused on combining inorganic materials with biological systems. An additional merit of these systems is that they can be constructed by well-known or even ubiquitous techniques (e.g. a DNA or protein layer can be fixed on a substrate using a standard ink-jet cartridge[17, 18]). As the power of this combination of material science and biotechnology lies in its specific interactions at the molecular level, there is a need for a detailed understanding of its properties at the molecular or even atomic scale. This has driven the interest to study such systems at a theoretical level, and attempts have been made to perform simulations of organic/inorganic interfaces at atomistic and other levels of modeling[19, 20].

The remaining part of this introduction will present a detailed outline of the thesis. Following this introduction, chapter 2 gives an overview of research concerning the interaction of proteins with (metal) surfaces, considering the main questions, applications, experiments, and simulations that have been performed recently.

Methodologies used in molecular simulation are covered in chapter 3. The emphasis will lie on classical atomistic simulation, and introductions to the quantum calculations and multiscale methods that are connected to this thesis. Note that there is no standard recipe on how to unambiguously separate processes occurring at the atomistic and quantum scales. Care should therefore be taken to retain the fundamental physical properties of these two processes and to link them without violating physical or chemical principles.

Next to the multiscale modeling of biomolecule-metal interactions, a recurring theme in the thesis concerns the understanding of macroscopic effects from microscopic phenomena. Atomistic simulation is well suited to describe microscopic phenomena. It will be shown that microscopic quantities calculated in a molecular dynamics simulation can give explanations of macroscopic effects. An important contribution to this is given by the calculation of free energies in classical atomistic simulations, which will therefore be treated separately in chapter 4.

In chapter 5, the quality of an existing force field for benzene solvation is analyzed by calculating a range of thermodynamic properties from simulations and comparing these with experimental values. The free energy of hydration for a realistic model of benzene is calculated in simulations and dissected into entropy and interaction energy terms. In this way it can be researched in how far entropic or energetic arguments determine the lower aqueous solubilities of non-polar molecules compared to polar molecules.

Following that, a series of simulations will introduce the modeling of solute-surface interactions. It starts from the solvated benzene solute of chapter 5, representing an amino acid side chain, adding the surface interaction and leads up to a biomolecule-surface study. A prerequisite for these simulations is the introduction of a multiscale modeling of water at gold and nickel metal surfaces under ambient conditions, as worked out and validated in chapter 6. A main characteristic of this modeling is the use quantum chemical data that specifically take into account the interactions of liquid water with a metal surface. Furthermore care is taken to split-up the water-surface interactions from the water-water interactions, which are treated by existing classical force fields.

A further step is the introduction of benzene and phenol solutes, simple model systems for the amino acid side chains of phenylalanine and tyrosine, in chapter 7. Similarly to the water modeling, the basis will be surface adsorption energy and conformational input from first-principles density functional methods and classical molecular dynamics simulations using atomistic force fields. Surface effects at a Ni(111) and Au(111) surface will be studied, allowing the research of the competition between water adsorption and solute adsorption at different surfaces. This can clarify how geometric effects concerning water layering and molecular configurations are connected to the adsorption mechanism of solutes near metal surfaces in explicit water.

Finally, actual amino acids are introduced in their neutral form in chapter 8. Specifically, systems containing the neutral forms of phenylalanine and alanine are simulated, at the interface with a Ni(111) surface in explicit water at room temperature. This time, an iterative quantum-classical strategy is used, starting with initial quantum mechanical calculations, on which a preliminary chemical modeling scheme is parameterized. This parameterization is used in preliminary atomistic simulations to determine optimal energies and configurations that can be sampled

with this modeling scheme. Additional quantum calculations can be performed to analyze configurations from the atomistic simulations, this will lead to subsequent modeling parameterizations, simulations, and quantum calculations until a consistent modeling is reached. Such simulations give qualitative insight in amino-acid-surface binding. Also here, the importance will be shown of taking the water explicitly into account in the modeling of surface adsorption behavior. Hydration of both solute and surface strongly affect the solute conformation space and cause free energy barriers for surface approach. The multiscale modeling approach used here enable a chemical, quantum-based, simulation of a bulk-to-metal adsorption process of an amino acid in an explicitly hydrated system.

2 Protein-surface interactions, an overview

In order to demonstrate the scientific relevance of the systems studied in this thesis, an overview will be given here of the current state of experimental and computational research of the interaction of proteins with (metal) surfaces. The emphasis will be put on studying direct, steered, protein-surface interactions. That will be followed by a more elaborate description of existing research at various levels of computational modeling for the interaction of proteins with (metal) surfaces. Protein-surface interactions often take place in the hydrated state, and therefore water and its interactions with the protein and the metal should be taken into account in a computational model as well. The analysis and modeling of this water-surface interaction is a problem on itself and will be covered separately in chapter 6.

2.1 Experimental studies of biomolecule-surface interactions

Experimental research on protein-inorganic surface interactions has been concerned mainly with those protein-surface interactions where the size of the systems enables a straightforward acquisition of macroscopic data (e.g. adsorption kinetics[21, 22]), but lacks insight into the driving forces for peptide-surface interactions[23]. Recently, however, it has become possible to determine relative adsorption strengths for series of oligopeptides. Details of these experiments will be shown below, after a short introduction of experimentally available experimental techniques and inorganic surfaces.

2.1.1 Experimental techniques

A range of experimental techniques is available to characterize and obtain quantitative information of protein-surface systems, at various levels of detail[24]. The film thickness of an adsorbed protein layer can be measured with ellipsometry[25], neutron reflectivity can be used to measure the density of an adsorbed layer[26]. In the case of biosensors an enzymatic reaction can be controlled and measured via electrochemistry between the protein and the conductive surface[13]. Atomic force microscopy (AFM) micro-cantilevers can be used by fixing proteins on the cantilever tips. Protein-substrate interactions will change the surface tension on the micro-cantilever, thereby bending the cantilever and creating a measurable signal[27]. Surface Plasmon Resonance (SPR) can be applied for immunoassays where antibody proteins are fixed to a metal surface, and a signal is generated upon antigen binding[11, 28]. Advanced biosensors can be constructed that combine several analytical techniques that are able to give complementary information. For example, SPR can be combined with subsequent Maldi-TOF mass spectrometry[29], or with fiber optic absorbance spectrometry[22], all in a single measurement cell. The technological devices that combine biological systems, mechanics and electronics are known as Biomedical or Biological Micro-Electro-Mechanical Systems (BioMEMS)[30]. Apart from the biochemical analysis tools mentioned above, one can also use the same techniques for e.g. the construction of artificial retinal implants[31].

2.1.2 Commonly used inorganic surfaces

The choice of the inorganic surface substrate in a protein-surface study is based on several practical considerations. Probably the most important consideration is the fact that many biosensor techniques, for example electrochemical or SPR analysis, need a conductive surface [11, 24]. Gold surfaces are traditional metal substrates in liquid/metal systems, due to its ease of use (e.g. patterning) and inertness, which ensures a stable surface during the course of the experiment[9]. This does not imply that gold is the only surface of interest, as also palladium and platinum can be considered interesting substrates for biological systems[9, 23, 32]. Other metals that have been experimentally evaluated as a substrate for peptide binding are Cr[33], Al[23], Ag, Cu[34] and Ti[10, 35].

Apart from metals, semiconductors are systems of high interest for protein-

surface systems, as they can be integrated with electronic circuits[10, 23, 35–38]. Examples of semiconductor surfaces used in experimental peptide binding studies are GaAs, AlGaAs[23, 37], CdS, CdSe, ZnS, ZnSe[38], InP, Ge, GaP, Si[36] and several metal oxide semiconductors as PbO₂, CoO, MnO₂[10], ZnO, Cu₂O, Cr₂O₃ and Fe₂O₃[39].

A special consideration when studying biological systems is the issue of biocompatibility. For this, not only the ability of a surface to bind proteins should be considered, but also its inertness, as the release of metal ions out of a surface can for example disturb protein structure, or have a toxic effect on any living cells that might be present. In the case of Ti, for example, it is known that an oxide layer can form at the contact with air, after which the material becomes effectively inert. Protein binding to this outer TiO₂ layer might then occur via carboxyl groups.[40]. Metals that are not of the noble metal group or able to form an inert layer will generally not be suited for biosensor applications. A well-known example is Cu, which easily oxidizes and is highly toxic[41].

In the current thesis, the noble metal Au will be used for a part of the studies. However, most of the studies will concern a Ni surface, which is too reactive to be considered a noble metal, but was of interest due to experience gained for this system in similar simulation work that was performed for a completely different technological problem concerning the optimization of compact disc materials[6, 7, 42–45]. The nickel surface is modeled here only in its idealized state, as oxidation is not considered, and therefore represents a model inert metal surface. In the water- and benzene-surface interaction studies also gold is taken into account, but the initial studies performed in this thesis indicated that gold was not an interesting substrate to study specific, non-covalent interaction. It will therefore be important to extend the system to surfaces that are likely to be interesting from both an application and a modeling point of view, as for example Pt and Pd.

2.1.3 Experimental research of direct peptide-surface interactions

Ongoing technical developments such as combinatorial peptide engineering[33, 37] have opened the way to study and predict direct protein adhesion by considering adhesive properties of the constituting amino acids. For these studies, either

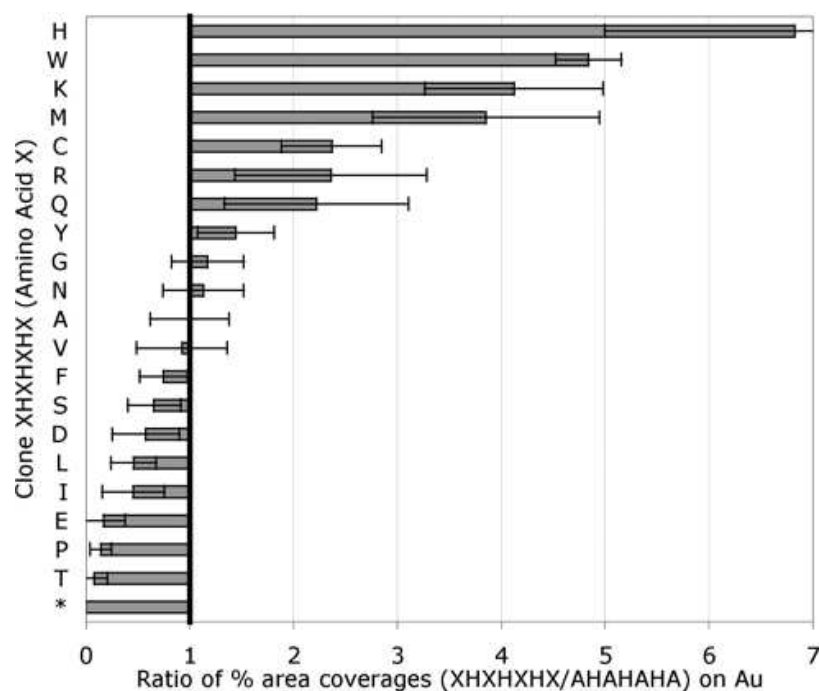


Figure 2.1: Relative binding of different peptide oligomers to a polycrystalline gold surface, as measured by the area coverage of yeast cells with oligomers expressed at the outside of the cells[38]. Each data point is labeled with the one-letter abbreviation used for amino acids (see also figure 2.2). The data point labeled * represents the test performed in the absence of expressed peptide at the yeast cell surface.

existing metal-binding peptides were used as a starting point (e.g. Gold Binding Peptide, GBP[33]), or randomized chains were created and analyzed for their the metal-binding strength[37]. The most recent advancements can be found in the systematic experimental studies that can identify an increase or decrease in peptide surface binding when comparing various peptide groups in synthetic, ordered, poly-peptides[23, 38].


For example, using genetical engineering, it was possible to display oligomer peptides at the outside of yeast cell clones[38]. These yeast cells where then brought into contact with metal plates, and the adhesion of the cells to the metal surface could be quantified by counting individual cells with an optical microscope. Relative area coverages of yeast cells with interdigitated oligopeptides (XHXHXHX, where H is Histidine and X was a different amino acid type for each different clone tested) expressed at the cell surface are shown in figure 2.1. Apparent is that positively charged amino acids (K,R,H) resulted in a relatively stronger binding, whereas

negatively charged amino acids (D,E) had a weakened binding compared with the peptide containing alanine (X=A).


A different study used solubilized peptide decamers and quantified the absorption by labeling each decamer with a single fluorescent group and measuring the fluorescence of the peptides that adhered to the surface. A range of surfaces was tested, a table is given in figure 2.2. The data in table one indicate that charged peptides are the best binders on semiconductor surfaces and some of the metals.

By comparing the data for gold surface binding in the two before-mentioned studies it becomes clear that they are not consistent, which is most likely due to the large difference between the experimental set-up of both studies. Still, the value of studies like these is clear: as the binding strength of a certain peptide subunit will differ from surface to surface, the knowledge about individual peptide binding enables the construction of polypeptides that bind selectively or even specifically to a certain inorganic surface[10, 23, 33, 35–38].


	K Lys	R Arg	H His	D Asp	E Glu	T Thr	S Ser	N Asn	Q Gln	Y Tyr	P Pro	M Met	C Cys	W Trp	G Gly	A Ala	V Val	I Ile	L Leu	F Phe
GaAs	1.4	0.7	<.5	0.9	3.5	2.1	2.1	0.9	<.5	0.7	1.6	1.8	0.7	1.2	0.9	<.5	0.7	<.5	<.5	<.5
Si ₃ N ₄	27	20	22	11	28	4.8	2.1	1.4	<.5	3.2	4.8	1.4	3.9	2.5	3.4	<.5	<.5	<.5	<.5	1.2
SiO ₂	23	17	25	8.7	32	5.2	<.5	0.9	<.5	2.1	4.4	3.5	4.8	0.9	3.3	<.5	<.5	1.8	<.5	1.4
AlGaAs	34	2.1	1.8	16	8.1	2.5	2.1	1.8	2.5	1.2	3.5	3.2	4.8	1.6	2.1	3.4	<.5	3.4	<.5	<.5
Al	3.5	61	1.8	11	5.0	3.5	4.4	2.1	2.3	1.2	1.4	3.5	1.2	3.2	3.0	0.7	3.0	1.4	0.7	0.7
Pt	1.4	1.4	1.4	2.3	1.4	<.5	<.5	<.5	0.7	1.6	<.5	0.7	<.5	1.4	0.7	<.5	<.5	<.5	<.5	<.5
Ti	0.7	1.4	<.5	1.8	1.4	2.3	1.4	0.7	<.5	1.4	0.7	0.7	<.5	<.5	<.5	<.5	<.5	<.5	<.5	<.5
Au	<.5	1.2	<.5	0.9	<.5	1.6	0.9	<.5	<.5	<.5	0.9	<.5	<.5	<.5	<.5	<.5	<.5	0.9	<.5	<.5
Pd	0.9	0.7	1.4	0.9	0.7	2.1	<.5	<.5	<.5	2.7	0.7	0.7	<.5	1.4	0.7	<.5	0.9	<.5	<.5	<.5



Polar -
charged



Polar-
non charged



Non-polar

Figure 2.2: Peptide surface density ($\times 10^3/\mu m^2$) as measured by fluorescence data of peptide decamers containing one end-group fluorescent label[23].

2.2 Modeling biomolecule-surface interactions

Parallel to the advancements in experimental techniques, there is an increased interest in performing modeling studies of water-solvated proteins near (inorganic) interfaces. The molecular modeling of protein-surface interactions can elucidate the underlying principles of the interactions, and thus enable the design of tailored protein-surface interactions[19]. The simulation of such interface systems will however have to cope with methodological problems, for example the at force field level[20], a point that will be addressed in the current thesis. Modeling protein-surface interactions can be done at several levels of detail. In the sections below, ordered by increasing atomistic detail, examples will be given of recent research in these levels.

2.2.1 Analytical models

General interaction properties can be calculated on the analytical level[46, 47]. For example, an analytical approach based on dielectric properties was used to estimate the van der Waals interaction of a slab of protein near a solid[46]. A protein slab was represented by a layer with a refractive index and dielectric permittivity corresponding to protein matter. Other materials present (the metal layer, slabs of water and organic material), were represented by their respective refractive indices and dielectric permittivities. It was found that adding a hydrocarbon layer on top of the metal surface minimized the van der Waals interaction energy between the protein and the metal surface: for a layer corresponding to only 3 carbon atoms in height the interaction energy is already 10 % of the original strength.

2.2.2 Monte Carlo simulations

More detailed structural information at the peptide-surface interface can be obtained from Monte Carlo (MC) simulations. Often, coarse grained peptide models are employed that look for structural transitions that may occur as the peptide model approaches the surface. The most simplified model in these studies is a homopolymer[48], of which a complete folding phase diagram can be constructed[49], or a charged polymer model, used to study pH-dependence of the surface adsorption[50]. The Gō model that performs Monte Carlo steps based on a known global

minimum is often used in protein folding simulations to probe folding paths has also been used in protein-surface simulations, for example to study the influence of the location of the protein-surface binding residue on the folding mechanism[51, 52]. Chemical information can be used in such models, for example by distinguishing between hydrophobic and polar subunits[53]. Several studies address the orientation of proteins at a specific surface, where not only the interaction with the surface[54, 55], but also the spacial arrangement of the protein next to the surface is of importance[56]. Some Monte Carlo studies apply explicit solvent or solvent/co-solvent mixtures to gain orientation and hydration information of small peptides near surfaces[57, 58]. A step higher in complexity is the application of available (empirically parameterized) force-fields and appropriate mixing rules to create a realistic description of both the biomolecule and a nonconducting surface. Although this approach is applicable within Monte Carlo studies, it is more often simulated via molecular dynamics simulations, as that allows to study dynamic behavior.

2.2.3 Molecular dynamics simulations

Interactions in systems containing peptides and solvent molecules near organic and nonconducting inorganic surfaces can in principle be described by classical force-fields. The total solute-surface interaction energy can then be calculated by integrating the solute-surface atom pair interaction potential over all separate surface atoms[59–61].

Another application for this is the simulation of interactions of proteins with self-assembled-monolayers (SAMs), where the SAM is considered to cover its base surface completely[62, 63]. An atomistic description of complete protein-organic surface adsorption processes in explicit water is limited by the large system size and long simulation times required. As an approximation, proteins can at first be energy-minimized onto the surface, after which water is added to the system[59].

The electron-conducting nature of metal surfaces make this the most versatile base for protein-surface composites, but also the most difficult system to study with molecular modeling. The main cause for this is that, contrary to the peptide-organic surface case, one can not construct a generalized force field to parameterize peptide-metal surface interactions. Force fields generally account for the non-bonded interactions between two atoms by applying mixing rules to the Lennard-Jones po-

tential of both atoms, whereas bonded interactions are based on either constrained distances or anharmonic potentials. In the case of metal surfaces however, it will be incorrect to calculate an overall surface interaction of nearby molecules by integrating over all separate pair interactions between solute and metal atoms[20]. A main cause for this is the delocalization of the electrons (the main contributor to interactions) in a metal surface. Perturbations of the electron density caused by a nearby molecule will lead to interaction strengths and optimal interacting configurations that can only be determined via quantum mechanical (QM) calculations and its parameterization into atomistic potential energy functions for classical simulations should be verified on a case-by-case basis[20]. This does not mean that interactions of organic molecules with metal surfaces are not parameterizable in a molecular modeling scheme per se. Even when single atom-metal surface interaction parameters are not readily transferable between different molecule types, we can still model the QM-determined adsorption energies and configurations of chemically active groups that have relatively isolated electronic properties (e.g. polar groups in an alkane chain), and combine any number of those groups in a macromolecule, as long as this combination does not have a heavily altered electronic structure, and the surface-adsorbed groups will be separated far enough to not disturb the electron distribution of the neighboring solute-metal interactions. This methodology has been successfully applied in previous studies concerning the modeling of polycarbonate-nickel interactions[6, 7].

Previous approaches to model metal-adsorbed peptides tried to simplify the problematics in various ways, e.g. by ignoring the peptide-metal interaction altogether[64], or by using a generalized force field for the solvent-metal and protein-metal interaction[65, 66], even though this has the conceptual problems mentioned above. An approach to overcome these difficulties will be introduced in this thesis, as will be worked out in chapters 7 and 8.

2.2.4 Quantum calculations

Several studies describing quantum calculations for the interaction between amino acids in their neutral state and metallic surfaces are available[67–71]. Attempts to study amino acids in their zwitterionic state have also been made[72–75]. Analyzing adsorption energies of the zwitterionic state by quantum calculations is complex,

since the zwitterion is only stable in solvated state, not as an isolated molecule, and no clearly defined reference state is present. It is however possible to obtain a minimal estimate of the adsorption energy, as will be discussed in chapter 8.

A further step is the study of peptide-surface interactions by quantum calculations. The N-methyl-acetamide molecule, which is the smallest representation of a peptide bond, has already been subject of intensive research on the quantum calculation level, most of which are dedicated to the rotational energy profile around the peptide bond[76–84]. The interaction of a peptide bond with a metal surface has been studied by quantum calculations for a system containing a TiO_2 surface and a peptide bond with one of the connected amino acid groups[85].

3 Theoretical basis of Molecular Simulation

Some of the main aspects of molecular simulation will be treated here, with a main focus put on the classical molecular dynamics calculation methods used in the current thesis, and on the special considerations required for simulating aqueous/metal interfaces. Most of the methodologies used are well-established techniques and are extensively explained in classical reference works on molecular simulation[86, 87] and software package references[88]. Therefore only the basic foundations of these subjects will be treated here. Short introductory comments will be made about the quantum calculations to which this work refers to, and on the current status of multiscale simulations methodologies.

3.1 Mechanics of Molecular Systems

Consider a one-component system consisting of N particles. This system can be described by the set of N Cartesian coordinates \mathbf{r}_i and momenta \mathbf{p}_i , which results in a phase space Γ with dimension $6N$.

The Hamiltonian of this system is given by the sum of the kinetic energy K and the potential energy V , where the former energy term depends on the momentum \mathbf{p} of the particles, the latter on the coordinates \mathbf{r} :

$$H(\mathbf{r}, \mathbf{p}) = K(\mathbf{p}) + V(\mathbf{r}) \tag{3.1}$$

The separation in the Hamiltonian that defines kinetic energy as solely dependent on the momenta, and potential energy solely on the coordinates, enables the

use of a simple set of equations of motion, called Hamilton's equations:

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i} \quad (3.2)$$

$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i} = -\frac{\partial V}{\partial \mathbf{r}_i} = \mathbf{f}_i \quad (3.3)$$

Where \mathbf{f}_i is the force acting on particle i . Given a set of initial conditions, these equations can be numerically integrated in discrete time steps in the molecular dynamics simulation using so-called integrators, as explained in section 3.3. Kinetic energy is the most trivial part of the Hamiltonian, as it is just the sum of kinetic energies for all particles. Potential energy is a more complex issue, however, and will be built up out of a combination of pairwise interactive terms, that are described in the molecular force field used, as described in the following section.

3.2 Molecular force fields

Potential energy within molecular dynamics is defined by a force field, often initially based on initial quantum calculations and then optimized to be able to represent experimental bulk data, leading to so-called 'empirical force fields'. Common force fields describe the interactions between molecules by a combination of dispersion interactions and electrostatic interactions between the atoms or, more generally, sites on the molecules present in the system.

3.2.1 Electrostatics

In the case of ions, electrostatics between an ion i and another ion j is given by the electrostatic interaction energy between the two charges q_i and q_j , which is described by the Coulomb interaction:

$$V_c(r_{ij}) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}} \quad (3.4)$$

Where ϵ_0 represents the dielectric permittivity of vacuum. In atomistic simulations, the total dielectric permittivity of a system (ϵ) can be calculated from fluctuations in the total dipole moment of the system[89]. The ratio between the permittivity of

a system and the permittivity of vacuum (ϵ/ϵ_0) is known as the relative permittivity (ϵ_r) and is often used when calculating coulomb interactions in coarse grained simulations where no explicit medium is present.

The force on charge point q_i due to electrostatic interactions with point q_j can be calculated as:

$$\mathbf{F}_i(\mathbf{r}_{ij}) = \frac{\mathbf{r}_{ij}}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}^3} \quad (3.5)$$

Difficulties arise due to the long distance range of these functions, the special methodologies used to treat this in simulation will be therefore discussed in more detail in section 3.7. For molecules that have no net charge, its atoms can still be assigned partial charges to represent the dipole and higher order electric moments or to reproduce the electrostatic potential on the molecular surface.

3.2.2 Dispersion

Dispersion describes the interactions due to induced dipole interactions in a molecule. These interactions are short range and often described in Lennard-Jones potentials, together with short range repulsive interactions. The Lennard-Jones interaction can either be expressed as a function of radius σ and interaction strength ϵ , or as Lennard-Jones parameters $C^{(6)}$ and $C^{(12)}$.

$$\begin{aligned} V(r_{ij}) &= 4\epsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \\ &= \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6} \end{aligned} \quad (3.6)$$

One advantage of the Lennard-Jones interaction is its computational efficiency. The power terms can be reused for the force calculation, for example:

$$\mathbf{F}_{ij} = \frac{\mathbf{r}_{ij}}{r_{ij}^2} \left[12 \left(\frac{C_{ij}^{(12)}}{r_{ij}^{12}} \right) - 6 \left(\frac{C_{ij}^{(6)}}{r_{ij}^6} \right) \right] \quad (3.7)$$

When calculating the interaction between different particle types, the interaction potential for this set of particles is needed. In most force fields, this interaction potential is determined by applying a mixing rule on the interaction parameters

(interaction strength and optimal distance) of both particle types. The choice of the mixing rule depends on the force field used. In the case of the GROMOS43a1 force field[88], for example, the geometric average method is applied:

$$C_{ij}^6 = (C_i^6 * C_j^6)^{1/2} \quad (3.8)$$

$$C_{ij}^{12} = (C_i^{12} * C_j^{12})^{1/2} \quad (3.9)$$

3.2.3 Surface potentials

In this thesis, the metal surface will be modeled by a set of site-surface potentials. Two popular site-surface potentials are considered. A generic surface interaction can be constructed by 10-4 interactions, that represent an integration of all interactions that would exist between a site and all surface atoms in its vicinity in an all-atom surface description[90]. This potential was used in previous multiscale modeling methods describing polycarbonate-nickel interactions[6]. A graphical representation for all functions is given in figure 3.1.

$$U_{Attr.10-4} = \begin{cases} 2\pi\epsilon \left[\frac{2}{5} \left(\frac{\sigma}{z}\right)^{10} - \left(\frac{\sigma}{z}\right)^4 \right] & , z \leq z_{cutoff} \\ 0 & , z > z_{cutoff} \end{cases} \quad (3.10)$$

$$U_{Rep.10-4} = \begin{cases} 2\pi\epsilon \left[\frac{2}{5} \left(\frac{\sigma}{z}\right)^{10} - \left(\frac{\sigma}{z}\right)^4 + \frac{3}{5} \right] & , z \leq \sigma \\ 0 & , z > \sigma \end{cases} \quad (3.11)$$

An alternative choice for site-surface interaction potentials is the Morse potential. This is originally a bond-like potential[91], and able to describe short-range electron overlap. An advantage to the Morse interaction is that an additional factor is present that can fine tune the steepness with which the interaction strength will decay at increasing distances from the surface.

$$U_{Attr.Morse} = \begin{cases} \epsilon (1 - e^{-a(z-\sigma)})^2 - \epsilon & , z \leq z_{cutoff} \\ 0 & , z > z_{cutoff} \end{cases} \quad (3.12)$$

$$U_{RepMorse} = \begin{cases} \epsilon (1 - e^{-a(z-\sigma)})^2 & , z \leq \sigma \\ 0 & , z > \sigma \end{cases} \quad (3.13)$$

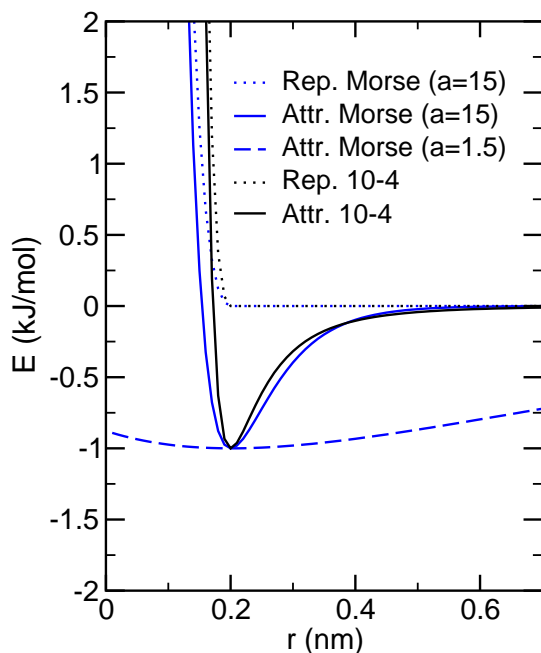


Figure 3.1: A graphical representation of the equations in . A σ value of 0.2 nm was used, ϵ was set to 1 kJ/mol for the Morse potentials, and $\frac{1}{6/5\pi}$ kJ/mol for the 10-4 potentials, both having a well depth of 1 kJ/mol. Two potentials are drawn for the Morse potential, with different Morse parameters. Small Morse parameters will lead to shallow curves, even up to $r = 0nm$.

3.3 Molecular Dynamics

The interactions, forces, velocities, and updated coordinates in a system can be calculated from the interaction potentials via various "integration" schemes. For the current study, only molecular dynamics (MD) was applied. Other integration methods are also available. For example, Monte Carlo calculations employ non-dynamic 'steps' that are based on random movements of the system. This can create more efficient sampling for certain systems, but makes MC less suitable to research dynamic properties of a system.

3.3.1 Time Step Integration

In molecular dynamics simulation, the configuration change in time step Δt are based on velocities and forces in the system as were calculated in the time step before. For infinitely small values of Δt , a continuous and energy-conserving simulation can be performed, but for computational efficiency a time step as high as possible should be chosen that can still retain energy-conservation with acceptable accuracy. Various integrators are available for molecular dynamics, the Gromacs software package used in this thesis[92–94] applies a variation of the Verlet algorithm, called the leap frog algorithm. In this algorithm, the velocity is calculated for the point in time between the current and the next time step ($\mathbf{v}(t + \frac{\Delta t}{2})$). This is calculated from the previously calculated velocity (at $\mathbf{v}(t - \frac{\Delta t}{2})$), plus the acceleration due to the force at the whole time step t . Then, based on that velocity, the coordinates for the next time step are calculated. This scheme is in principle time-reversible.

$$\mathbf{v}(t + \frac{\Delta t}{2}) = \mathbf{v}(t - \frac{\Delta t}{2}) + \frac{\mathbf{F}(t)}{m} \Delta t \quad (3.14)$$

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t + \frac{\Delta t}{2}) \Delta t \quad (3.15)$$

Forces are calculated from the functions describing the distance-derivative of the interaction potentials ($-\frac{\partial V}{\partial \mathbf{r}}$). The simplest form of molecular dynamics describes a system with a constant number of particles, a constant volume, and constant energy (NVE or microcanonical ensemble). This ensemble does not resemble standard experimental conditions, therefore alternative algorithms have to be used to sample in different ensembles. For molecular dynamics several methods are available to control temperature and pressure, as will be shown below.

3.3.2 Treatment of Bonds and Angles

In simulations on the atomistic level it is often sufficient to keep bond lengths fixed during the simulation, thereby avoiding the need of exact calculations of bond length vibrations. Several algorithms are available to retain fixed bond lengths in the simulation, well-known examples are the SHAKE[95] and the LINCS[96] algorithm. The general idea of these bond constraint algorithms is that firstly, all atoms are displaced in an unconstrained way, according to the Verlet algorithm, for example.

After this step most bonds will be offset from their ideal bond lengths. Then, forces are applied to the atoms depending on the deviation of the bonds from their ideal positions. As soon as all atoms are within a certain threshold distance from their ideal positions, the constraints are considered converged and the system is ready for the next time step.

Angles can be treated by simple spring potentials, where the deviation of the angle θ from its optimal value θ_{opt} is used in a harmonic potential, together with a parameterizable force constant k_{angle} . In the case of GROMOS force fields[88, 97, 98], the cosine of the bond angle is used instead of the angle itself:

$$V_{angle}(\theta) = \frac{1}{2}k_{angle}[\cos(\theta) - \cos(\theta_{opt})]^2 \quad (3.16)$$

In the case of rigid water models, no angular vibrations are taken into account, to increase computational speed. Other angular potentials can be used as well. Well-known potentials are dihedral potentials, that can be used to describe the potentials of staggered and eclipsed conformations in alkane chains, and improper dihedrals, that can for example be used to enforce flat conformations in aromatic systems.

3.4 Periodic Boundary Conditions

Basically, a simulation is performed on a single box of molecules, which would mean that a significant amount of the contents is located near a side of the box. To better represent a real-life system, mirror images are considered at each side of the original box, ending up with an infinite system of copies of the central box. Within these periodic boundary conditions, the so-called minimum image approach is normally used, where for each atom in the central periodic box, interactions are calculated with the atoms surrounding it, which can be either in the same box or in the periodic images of box. Normally, the number of calculations that need to be performed is reduced by only considering interactions within a spherical cutoff instead of within the total box, where the atoms within this cutoff are kept in a so-called neighbor lists, which is updated only every n time steps. When studying interactions of solutions with solid surfaces a 3D periodicity is not wanted. Surfaces can mostly be considered infinitely deep when compared to the nanometer scale used in MD, but for the sake of simulation speed, only the first few layers are simulated. To prevent the

calculation of interactions with molecules at the periodic image beyond the surface one can introduce a vacuum space exceeding the neighbor-search cutoff length in one direction. Still, special measures have to be made for long-range electrostatics, as will be discussed in 3.7.

3.5 Thermal Coupling

Molecular dynamics in its simplest form samples a system in the microcanonical ensemble (NVE), with a constant number of particles N , constant volume V , and a constant total energy of the system E . To sample in more common experimental conditions, a constant pressure, temperature *isotherm – isobar* ensemble (NPT) is needed. Defining constant pressure and temperature can be straightforward in Monte Carlo simulations, but is more troublesome under molecular dynamics simulations[99]. In effect one will need to add degrees of freedom used to regulate temperature to the equations of motion, thereby influencing the degrees of freedom.

3.5.1 Langevin Thermostat Molecular Dynamics

In Langevin thermostat molecular dynamics, an effective medium is represented by temperature-dependent noise term and a friction term added to the equations of motion. Langevin thermostat MD is often used in cases where the solvent degrees of freedom are not explicitly modeled. Instead, a solvent 'background' is assumed, that causes random accelerations and friction on the explicitly modeled solutes. The random accelerations are accounted for by a stochastic force, the friction is applied as an additional term to the particle velocities.

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = -m_i \xi_i \frac{d\mathbf{r}_i}{dt} + \mathbf{f}_i(\mathbf{r}_i) + \mathbf{\dot{r}}_i \quad (3.17)$$

Here, ξ is a friction term with unit ps^{-1} , and $\mathbf{\dot{r}}$ is a noise force which depends on the absolute temperature and the friction constant.

The Langevin thermostat MD method generates an exactly defined canonical, NVT, ensemble, but the random noise does disturb the actual dynamics. A special case of the Langevin thermostat is given by Brownian dynamics, where it is assumed that on average no acceleration takes place in the system. This can also be described

as over-damped or non-inertial Langevin dynamics.

The dissipative particle dynamics (DPD) thermostat is taken from the DPD simulation method. DPD is a method for the simulation of hydrodynamic behavior of mesoscopic fluids, and can be seen as a combination of a soft particle interaction potential and a thermostat, [100]. The thermostat is based on a similar idea as the Langevin thermostat as it applies conservative, random, and dissipative forces in the equations of motion. It has however the advantage that it preserves hydrodynamics as the dissipative and random forces are added in a pairwise fashion, which enables the conservation of momentum[101].

3.5.2 Berendsen

The Berendsen coupling scheme was introduced as an alternative to stochastic dynamics, keeping the coupling of temperature to an external heat-bath, while leaving out any noise effects, to not locally disturb dynamics in the simulation[102].

Main part of the Berendsen temperature coupling to a reference temperature T_0 is a change dT in temperature for each time step dt .

$$\left(\frac{dT}{dt}\right)_{bath} = \frac{T_0 - T}{\tau_T} = 2\gamma T_0 - T \quad (3.18)$$

Where γ is a friction term (with unit ps^{-1}), for which in MD programs often the temperature coupling time constant τ_T is used. The temperature change is processed in the equations of motion via a velocity scaling from v to λv , where the scaling factor λ is temperature dependent. In a leap-frog algorithm, velocities for configuration $x(t)$ are known only from half a time step before: $v(t - \frac{1}{2}\Delta t)$. The error due to this difference is small when coupling time τ_T is long compared to time step dt , and therefore λ can be written as:

$$\lambda = \left[1 + \frac{\Delta t}{\tau_T} \left(\frac{T_0}{T(t - \frac{1}{2}\Delta t)} - 1\right)\right]^{1/2} \quad (3.19)$$

After calculation of λ , velocities at half a time step later ($v(t + \frac{1}{2}\Delta t)$) will be scaled.

Berendsen coupling does not maintain an exact constant temperature ensemble, whereas Nose Hoover temperature coupling (described below) does. Still Berendsen coupling has an advantage as one can choose short coupling constants can be well-

suited for quick equilibration.

3.5.3 Nosé-Hoover

In Nosé-Hoover pressure temperature coupling [99], a function added to the equations of motion is chosen in such a way that the static quantities of the system are in accordance with the ensemble. The added degree of freedom is the coupling constant Q , which enters in the definition of heat bath parameter ξ :

$$\frac{d\xi}{dt} = \frac{T - T_0}{Q} \quad (3.20)$$

Note that T is the instantaneous temperature of the system, and T_0 is the reference temperature. The heat bath parameter enters the equation of motion:

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} - \xi \frac{d\mathbf{r}_i}{dt} \quad (3.21)$$

In Gromacs, Q is defined via a coupling time τ_T and reference temperature T_0 , by stating $Q = \frac{\tau_T^2 T_0}{4\pi^2}$. This coupling period τ_T is then a measure of the kinetic energy oscillations between the system and the reservoir.

3.6 Pressure Coupling

Pressure coupling is normally performed on box-vectors and either the particle coordinates (in the case of Berendsen pressure coupling) or particle velocities (in the case of Parrinello-Rahman coupling). As for temperature coupling, Berendsen coupling is more likely to speed up equilibration, but generates an ill-defined ensemble. Parrinello-Rahman introduces longer fluctuations, but generates a well-defined ensemble.

3.6.1 Berendsen

Berendsen pressure coupling[102] applies a weak coupling to coordinates $x(t)$ and box lengths $l(t)$. In the most general case, describing anisotropic systems, pressure needs to be regarded as a tensor, \mathbf{P} . A more straightforward case is formed by

isotropic systems, where pressure can be described as a scalar P , and just as for scalar temperature T in equation 3.18, one can write:

$$\left(\frac{dP}{dt}\right)_{bath} = \frac{P_0 - P}{\tau_P} \quad (3.22)$$

Similar to temperature coupling, P is the instantaneous pressure of the system, and P_0 is the reference pressure. In practice, a scaling factor (μ) is used to perform the pressure coupling. In the case of isotropic coupling, it can be described as:

$$\mu = 1 - \frac{\beta\Delta t}{3\tau_P}(P_0 - P) \quad (3.23)$$

Where β denotes the isothermal compressibility of the system, defined as.

$$\beta = -\frac{1}{V} \left(\frac{dV}{dP}\right)_T \quad (3.24)$$

For anisotropic systems, the scaling factor will take the form of a tensor. In the case of interfaces (for example the modeling of surfaces as performed in this thesis), one can use a tensor consisting of only the diagonal terms μ_{xx} , μ_{yy} , and μ_{zz} . The zz term can then be evaluated independently from the xx and yy terms, for example by setting μ_{zz} to zero, which means that the zz direction will not be scaled during the simulation. Pressure scaling is applied to box-lengths for the next time step ($l(t + \Delta t) = \mu l(t)$); for coordinates scaling is applied only after the coordinates of the next time step have been calculated, resulting in $x(t + \Delta t) = \mu \times x(t + \Delta t)$.

3.6.2 Parrinello-Rahman

Parrinello-Rahman pressure coupling[103] is similar to Nosé-Hoover temperature coupling, as it introduces an additional degree of freedom to the equations of motion. In the case of Parrinello-Rahman, this is an inverse mass parameter matrix W^{-1} . This can be roughly interpreted as using a 'heavy piston' to control the pressure in the system. The mass of the piston influences the fluctuations of the sample and should be set in such a way that the time scale for the volume fluctuation of a sample relates to the length of the sample divided by the speed of sound in the sample[104]. In simulation however, alternative values for the mass parameter can be used[103], for example, Gromacs uses the inverse mass parameter matrix W^{-1} . This matrix

has the unit of inverse mass and is a function of isothermal compressibility β and time constant τ_p . It takes the form of a tensor as it allows for anisotropic coupling, and is defined as $(\mathbf{W}^{-1})_{ij} = \frac{4\pi^2\beta_{ij}}{3\tau_p^2L}$. L denotes the largest box matrix element. For any type of unit cell, the box vectors can be expressed in a matrix \mathbf{b} . In the case of Parrinello-Rahman coupling the box matrix \mathbf{b}' would depend on a function of volume V , transposed box matrix \mathbf{b}' , and \mathbf{W}^{-1} .

$$\frac{d\mathbf{b}^2}{dt^2} = V\mathbf{W}^{-1}\mathbf{b}'^{-1}(\mathbf{P} - \mathbf{P}_{ref}). \quad (3.25)$$

The box matrix can then be used to define a scaling matrix \mathbf{M} , that will enter the equations of motion.

$$\mathbf{M} = \mathbf{b}^{-1} \left[\mathbf{b} \frac{d\mathbf{b}'}{dt} + \frac{d\mathbf{b}}{dt} \mathbf{b}' \right] \mathbf{b}'^{-1} \quad (3.26)$$

$$\frac{d^2\mathbf{r}_i}{dt^2} = \frac{\mathbf{F}_i}{m_i} - \mathbf{M} \frac{d\mathbf{r}_i}{dt} \quad (3.27)$$

3.7 Long range electrostatics

Long range interactions form a complex part of molecular simulations, as an increase in the cut-off radius within which all calculations are performed will scale with the third power for the number of particles that need to be calculated. To reduce computational costs, short-range cut-offs would therefore be desirable. However, at the cut-off distance unphysically high values for the derivative of the potential are present, which will be especially problematic when the cut-off distance is relatively low. This leads not only to an overestimation of the forces in the system, but also impairs the conservation of energy in the system. Common solutions to this problem are based on two popular methods: reaction-field electrostatics or Ewald summation.

3.7.1 Reaction-field

In reaction-field electrostatics[105], the effect of the electrostatic interactions caused by molecules outside the cutoff distance r_c is seen as the effect of a isotropic dielectric continuum with a known dielectric constant ϵ_{rf} . The value for ϵ_{rf} is therefore given

by the dielectric constant of the medium. The interaction potential then becomes:

$$V_{c_{rf}} = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r} \left[\frac{1}{r_{ij}} + k_{rf} r_{ij}^2 - c_{rf} \right] \quad (3.28)$$

In the case for explicit solvent, $\epsilon_r = 1$, in all other cases ϵ_r depicts the relative permittivity as compared to vacuum. The $k_{rf} r_{ij}^2$ term is a gradually decreasing function to making sure the derivative, the force, is zero at the cutoff distance r_c and no unwanted forces are created in the system that would disturb the energy conservation. The k_{rf} factor is determined by the influence of the dielectric continuum outside the cutoff and in its simplest form for a system of zero ionic strength it can be written as [105]:

$$k_{rf} = \frac{1}{r_c^3} \frac{\epsilon_{rf} - \epsilon_r}{(2\epsilon_{rf} + \epsilon_r)} \quad (3.29)$$

The form of the function is such, that the electrostatic interaction is most strongly screened when surrounded by a continuum with a large dielectric constant ($\epsilon_{rf} \gg \epsilon_r$). The last term, c_{rf} , is a constant correction term that makes sure the potential of the system is zero at cutoff distance $r_{ij} = r_c$.

$$c_{rf} = \frac{1}{r_c} + k_{rf} r_c^2 = \frac{1}{r_c} \frac{3\epsilon_{rf}}{(2\epsilon_{rf} + \epsilon_r)} \quad (3.30)$$

The force function determining the force in charge point q_i due to the reaction field coulomb interaction with charge point q_j is:

$$\mathbf{F}_i(\mathbf{r}_{ij}) = \frac{q_i q_j}{4\pi\epsilon_0\epsilon_r} \left[\frac{1}{r_{ij}^3} - 2k_{rf} \right] \mathbf{r}_{ij} \quad (3.31)$$

In the case of anisotropic systems, as liquid/solid interfaces, applying the reaction field method will be incorrect. The dielectric constant will not be constant behind cutoff, but depending on the direction (perpendicular or parallel to the surface). A simple cutoff could be used, but is not very reliable [106], therefore the optimal choice would be a (semi) 2D method, as is possible with techniques based on Ewald summation.

3.7.2 Ewald summation

The treatment of electrostatics with Ewald summation, originally developed for the calculation of charges in crystals[107], is a well-known method for the periodic systems studied in molecular dynamics[108, 109]. It is based on a representation of the electrostatics in a system by coulombic point charges q . For such a system of N point charges the total electrostatic energy can be given as:

$$E = \frac{1}{2} \sum_{i,j=1}^N \sum'_{\mathbf{n} \in Z^3} \frac{q_i q_j}{|\mathbf{r}_{ij} + \mathbf{n}L|} \quad (3.32)$$

The prime $'$ in the sum over periodic images \mathbf{n} indicates that the particle self interaction $i = j$ should be skipped when i and j are in the same periodic image ($\mathbf{n} = 0$). Because of the slow convergence of the electrostatic interaction in equation 3.32, the calculation is split into two parts. The reasoning can be visualized by the following expression:

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1 - f(r)}{r} \quad (3.33)$$

The first part on the right hand side is a Coulomb-like interaction potential within a certain cutoff r_c . Beyond the cutoff this function will be negligible or zero. The second part on the right hand side will be a slowly varying function at all distances r , and is optimized to be represented by a Fourier transfer.

In practice, the charge interactions can be treated by dividing the potential in three terms, a short range real space interaction $E^{(r)}$, a long range interaction that will be calculated in reciprocal space $E^{(k)}$, and a self interaction correction term $E^{(s)}$, which is constant.

$$E = E^{(r)} + E^{(k)} + E^{(s)} \quad (3.34)$$

The reciprocal space part extends infinitely over all periodic images. The actual charge distribution is described by a set of Gaussian charge clouds, for which an interaction in reciprocal space can be written as a sum over the set of \mathbf{k} -vectors:

$$E^{(k)} = \frac{1}{2L^3} \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2} e^{(-k^2/4\alpha^2)} |\rho(\mathbf{k})|^2 \quad (3.35)$$

where α is the inverse length (or: Ewald parameter) that is used to set the relative weight of the real and reciprocal contributions, and $\rho(\mathbf{k})$ is given by:

$$\rho(\mathbf{k}) = \sum_{j=1} q_j e^{-i\mathbf{k}\cdot\mathbf{r}_j} \quad (3.36)$$

The real space part is calculated in a similar way as normal coulomb interactions, except that the Gaussian functions ($\text{erf}(\alpha|\mathbf{r}_{ij} + \mathbf{n}L|)$) are subtracted from the point charges, to cancel the charges added to the system in equation 3.35. Using the function ($\text{erfc}(\dots) \equiv 1 - \text{erf}(\dots)$), this can be written as:

$$E^{(r)} = \frac{1}{2} \sum_{i,j=1} \sum_{\mathbf{n} \in \mathbb{Z}^3} q_i q_j \frac{\text{erfc}(\alpha|\mathbf{r}_{ij} + \mathbf{n}L|)}{|\mathbf{r}_{ij} + \mathbf{n}L|} \quad (3.37)$$

The system now contains both point charges and Gaussian charge clouds around a single charge point, and for a single charge point there can be a charge interaction with itself. To remove this term, a self interaction correction is performed:

$$E^{(s)} = -\frac{\alpha}{\sqrt{(\pi)}} \sum_i q_i^2 \quad (3.38)$$

In computer simulations, an alternative version of Ewald summation is often used, Particle Mesh Ewald (PME)[108,109]. Here, the charge distribution of the particles in the system is discretized onto a grid. Fast Fourier methods can be used to equate the potential for this grid, after which the resulting potential and force is assigned to each particle by interpolating the potential and force on the surrounding grid points.

3.8 Quantum calculations

The principle of standard molecular dynamics simulation is based upon the assumption that quantum effects play no role in the processes that are simulated. For many systems this atom-only description suffices, although the force field parameters that are applied there can be based on quantum calculations. Performing detailed quantum calculations would actually deliver unnecessary information for such problems. In some cases, however, the quantum nature of the system requires detailed in-

formation about the electronic states of the molecules of interest. Examples are chemical reactions, proton transfer, and interactions of solutes with metal surfaces, the subject of this thesis.

Therefore, complementary to the atomistic simulations presented in this thesis, quantum calculations have been performed, as a part of the multiscale modeling procedure. The exact details on the quantum calculations are outside the scope of this thesis and are treated elsewhere[43, 44, 67, 72]. Still, at least an elementary view on the materials is necessary to understand the multiscale modeling presented here, therefore a short introduction on the quantum calculations used is given here.

Ideally, systems would be solved fully by the Schrödinger equation for all electrons and nuclei. A solution would give the ground state wavefunction ψ , and the electron distribution of the system. Exactly solving this many body Schrödinger equation is however practically impossible. As one would like to obtain the electron distribution in the system, approximations have to be made, which will introduce adjustable parameters into the calculations. Methods to calculate the quantum mechanical properties of systems are often called "first principles" or *ab initio* methods as they ideally would not require any adjustable parameters. A popular approximate method is based on the Hartree-Fock methodology, where the complex n -particle wavefunction ψ is approximated by combinations of single orbital elements[110].

The current thesis refers to another commonly used quantum calculation method: first principles Density Functional Theory (DFT). It is based upon the assumption that a functional $E[\rho(\mathbf{r})]$ exists that depends on the electron density $\rho(\mathbf{r})$. Self consistent minimization of $E[\rho(\mathbf{r})]$ with respect to $\rho(\mathbf{r})$ will yield an unique, ground state, electron density[111, 112]. To solve this problem, a so-called pseudopotential plane-wave approach can be applied where the core electrons of the atoms are treated within the pseudopotential[111, 112]. Empirical parameters have to be introduced in DFT to describe electron exchange and correlation effects. These enter the effective potential of the system, $V_{eff}(\mathbf{r})$. Despite that disadvantage of requiring external parameters and thus not being purely *ab initio*, the plane-wave approach will be useful for describing complicated systems, as metallic crystals[112] and surfaces. For these systems, the quality of the calculation, and thus the added empirical parameters, can be readily validated by comparing experimental and simulated lattice constants, cohesive energy, and other properties.

A point of general concern is the treatment of dispersion interactions in quantum

calculations, as these interactions are caused by non-local instantaneous fluctuations, which would require an immense calculation load and is therefore left out of most quantum calculations. Recently, several attempts have been made to quantify the error made by this approximation[113, 114].

3.9 Multiscale Modeling

As mentioned in the previous section, DFT calculations are too slow to perform dynamics with adequate sampling. Similarly, treating chemical reactions or proton transfers in solution will require long time scales for equilibration and detailed treatment of the reaction centers. On a higher scale, when studying large hydrated proteins, a lot of computation time is spent on treating the water. Treatment of water near the protein is essential for realistic simulation, but further away from the protein bulk water could be assumed. All these cases have in common that a relatively small (central) part of the system needs a detailed treatment, while the major (outer) part could be treated by more coarse grained or even analytical descriptions. In these situations, a multiscale approach could form a solution to increase simulation speed[115, 116].

Multiscale methodologies are not without difficulties, often problems concerning the treatment of boundary zones are present, especially in QM/MM multiscale methods. Another problem occurs when a detailed description is needed at dynamic positions in the system, or when the system is dynamically moving around the position where a detailed description is necessary (for example when only molecules close to a surface need detailed treatment). A solution to this last problem can be resolved by a recently developed algorithm that can allow for particles moving from the coarse-grained to the atomistic level within a simulation[3, 4].

The multiscale modeling studies presented in the current thesis treat separate simulations at the various scales in a sequential coarse graining approach. Integration of the two levels is done in two directions. Translation from the quantum to the atomistic level is performed by parameterizing energetic and geometrical data from the molecule-surface interaction into the atomistic force field by a combination of atom-surface and virtual site-surface interactions. Langevin thermostat MD runs can then be used to quickly show the quality of the parameterization. In the other direction, the vacuum Langevin thermostat MD runs can sample stable con-

formations that were not sampled in the initial quantum calculations. Then, these conformations are taken from the Langevin MD runs and used as input files for the consecutive quantum calculations. This way, new minima can be allocated. As the sampling at the atomistic level will be mostly governed by the global surface attraction energy minima conformations, no overly detailed quantum description of local minima is needed. The approach described above will be treated in detail in chapter 8.

In recent years this kind of problem has been successfully addressed for a polycarbonate melt interacting with a metal surface[6]. There, in order to determine the polymer-surface interaction, quantum calculated adsorption energies of polycarbonate sub-molecules were combined with geometry considerations for polymer conformations at the surface, taking properly into account the chain topology and its influence into the adsorption process. This modeling was then incorporated into a coarse-grained model for the polymer interacting with a planar surface and later extended to a step defected surface[117]. Extensions will have to be made when starting from that model to the current modeling requirements. One extension will be the introduction of the modeling of water, which will be treated in chapter 6 of this thesis. Another point of complication will be the fact that more interaction sites and less molecular symmetry can be found in biopolymers, as compared to the synthetic polymers studied earlier. Furthermore, the combination of adsorbing water and biopolymer adsorption will lead to a competition effect in the adsorption process. In chapters 7 and 8, a starting point to deal with these modeling problems will be introduced, based on an approach where all molecular groups of interest are considered separately in a building block-like manner.

4 Free Energy in Statistical Mechanics

Statistical mechanics provides a link between the macroscopic properties and the molecular constitution of a system. It is based on the idea that macroscopic observables can be expressed as statistical averages over the collection of microstates (instantaneous snapshots of particle positions and velocities) sampled by the system. Microstates can be seen as instantaneous snapshots of all positions and momenta in the system (the phase space Γ). An important condition for this is that the system is bound by external constraints that are well defined, for example, density and temperature are given and constant. Such a collection of microstates is referred to as a statistical ensemble.

The ensemble where the number of particles N , the volume V , and the temperature T are fixed, is referred to as the canonical ensemble (or NVT ensemble). The probability $P(\Gamma)$ that the system will visit a particular microstate i , $\Gamma_i = (\mathbf{r}_i, \mathbf{p}_i)$ is given in the NVT ensemble by:

$$P(\Gamma) = \frac{e^{-\beta H(\Gamma)}}{Q_{NVT}} \quad (4.1)$$

where $H(\Gamma)$ is the (classical) Hamiltonian and Q_{NVT} is a normalization factor that is defined as:

$$Q_{NVT} = \int d\Gamma e^{-\beta H(\Gamma)} \quad (4.2)$$

The normalization factor in equation 4.2 is more commonly known as the canonical partition function, and relates to the free energy of an NVT system, $F(N, V, T)$, by:

$$F(N, V, T) = -k_B T \ln Q_{NVT} \quad (4.3)$$

The macroscopic observable $\langle A \rangle$ of a mechanical quantity $A(\mathbf{\Gamma})$ can be expressed in the canonical system as a function of the probability defined in equation 4.1:

$$\begin{aligned}\langle A \rangle_{NVT} &= \int d\mathbf{\Gamma} A(\mathbf{\Gamma}) P(\mathbf{\Gamma}) \\ &= \frac{1}{Q_{NVT}} \int d\mathbf{\Gamma} A(\mathbf{\Gamma}) e^{-\beta H(\mathbf{\Gamma})}\end{aligned}\tag{4.4}$$

By assuming that the ensemble averaged quantity $\langle A \rangle$ can be replaced by a time averaged quantity \bar{A} , it becomes possible to evaluate equation 4.4 in molecular dynamics simulations. When the time average over observation time T equals the ensemble average of a system, the observed system is called ergodic ($\bar{A} = \langle A \rangle$), and \bar{A} can be defined as:

$$\bar{A} = \frac{1}{T} \int_0^T dt A(t)\tag{4.5}$$

Unlike the mechanical properties $\langle A \rangle$ or \bar{A} defined above, the calculation of thermal properties, as for example free energy or entropy, can not be performed by straightforward statistical averaging. The difficulty that arises can be exemplified by expressing free energy 4.3 in the form of an ensemble average (equation 4.4):

$$\begin{aligned}F &= -k_B T \ln \int d\mathbf{\Gamma} e^{-\beta H} \\ &= -k_B T \ln \left[\int d\mathbf{\Gamma} \frac{e^{-\beta H}}{\int d\mathbf{\Gamma} e^{+\beta H} e^{-\beta H}} \right] \\ &= k_B T \ln \left[\frac{\int d\mathbf{\Gamma} e^{+\beta H} e^{-\beta H}}{Q_{NVT}} \right] \\ &= k_B T \ln \langle e^{+\beta H} \rangle_{NVT}\end{aligned}\tag{4.6}$$

It can be seen that the calculation of the ensemble average of $e^{+\beta H}$ is needed to calculate the free energy. This becomes problematic for the states in the phase space where the Hamiltonian H is large. These are unfavorable high-energy states and have a low probability to be sampled (for increasing H , the $e^{-\beta H}$ part of equation 4.1 drops exponentially). These high energy states contribute however significantly to the ensemble average in equation 4.6, leaving the exact calculation of equation 4.6 practically useless. This is a major problem that complicates the calculation of the (absolute) free energy. However, in practice only free energy differences between various states are of interest and not so much the absolute free energy of a single system. Therefore, two frequently used methods to calculate free energy differences

in simulations, will be treated in the following section.

4.1 Determining free energy changes

This thesis will deal for a great part with solvation effects and the corresponding thermodynamic changes. In chapter 5, the free energy change of aqueous solvation is discussed as studied via various methods, whereas in chapters 7 and 8 free energy changes will be retrieved along an order parameter.

4.1.1 Thermodynamic perturbation

As stated before, the parameter of interest is the free energy difference between two states. One can take for example a state B containing a solvated solute, which differs from state A only by a different interaction with its solvent environment. The difference in potential energy U of these states can be written down as:

$$U_B = U_A + \Delta U \quad (4.7)$$

Note that in this example ΔU describes the difference in solute-solvent interaction energy of state A and B. The free energy difference of these two states can be written for an NVT ensemble as (based on equation 4.3):

$$\Delta F_{AB} = -k_B T \ln \frac{Q_B}{Q_A} \quad (4.8)$$

The NVT partition functions can be written fully, as in equation 4.2:

$$\begin{aligned} \Delta F_{AB} &= -k_B T \ln \frac{\int d\mathbf{r}^N e^{-\beta U_B}}{\int d\mathbf{r}^N e^{-\beta U_A}} \\ &= -k_B T \ln \frac{\int d\mathbf{r}^N e^{-\beta U_A} e^{-\beta \Delta U}}{\int d\mathbf{r}^N e^{-\beta U_A}} \\ &= -k_B T \ln \langle e^{-\beta \Delta U} \rangle_A \end{aligned} \quad (4.9)$$

The ensemble average $\langle \dots \rangle_A$ will be taken from a system which is described with the Hamiltonian H_A . Equation 4.9 can successfully be evaluated in case the perturbation potential ΔU is small. In such closely related systems, the probability distributions $P_A(\mathbf{\Gamma})$ and $P_B(\mathbf{\Gamma})$ will have sufficient overlap, and a simulation of only one of the states is necessary. In practice this will only hold for very small changes,

in the case of solute hydration one could perform for example a pure water run (state A) and directly calculate the hydration energy needed to insert small hydrophobic solutes (e.g. He, Ar, CH₄). This leads to a description of state B, the state consisting of the solvent/solute mixture. This method is generally known as Widom's particle insertion method[118]. The limitation of this method is that the bigger the solute will get, the smaller the probability overlap between state A and B will get, which will eventually result in insufficient sampling to reach correct free energy differences between these states.

4.1.2 Thermodynamic integration

In the case of large perturbation potentials ΔU (equation 4.7), a different method needs to be applied to correctly calculate free energy differences. One of the methods available, thermodynamic integration, can in essence be seen as an extension to thermodynamic integration. It enables the calculation of free energy differences of larger perturbation potentials by dividing the perturbation potential in a collection of smaller perturbation potentials. More exactly put, it introduces a switching parameter λ , which can take values between 0 and 1 (following the previously used definitions, 0 would correspond to state A and 1 would correspond to state B). Then, any state in between A and B has a potential energy described by $U(\lambda)$, which is given as:

$$U(\lambda) = (1 - f(\lambda))U_A + f(\lambda)U_B \quad (4.10)$$

The function $f(\lambda)$ can take various forms but has the boundary conditions that $f(0) = 0$ and $f(1) = 1$, for the sake of simplicity it is considered to be $f(\lambda) = \lambda$ in the rest of the discussion.

Instead of the total perturbation energy ΔU , only the small change of $U(\lambda)$ is considered that occurs at a given λ -point. If the path from state A to state B would be divided into infinitesimally small steps, the energy perturbation at each point λ would be equal to $\frac{\partial U(\lambda)}{\partial \lambda}$. Using that instead of ΔU , and the fact that the partition function (equation 4.3) is now λ -dependent, one can write a λ derivative

of the perturbation free energy:

$$\begin{aligned} \left(\frac{\partial F(\lambda)}{\partial \lambda}\right)_{NVT} &= -k_B T \frac{\partial \ln Q_{NVT}(\lambda)}{\partial \lambda} \\ &= -k_B T \frac{1}{Q_{NVT}(\lambda)} \frac{\partial Q_{NVT}(\lambda)}{\partial \lambda} \end{aligned} \quad (4.11)$$

Rewriting the partition functions and consecutively performing the derivation by λ results in the ensemble average of the factor $\frac{\partial U}{\partial \lambda}$:

$$\begin{aligned} \left(\frac{\partial F(\lambda)}{\partial \lambda}\right)_{NVT} &= \frac{\int d\mathbf{r}^N \frac{\partial U(\lambda)}{\partial \lambda} e^{-\beta U(\lambda)}}{\int d\mathbf{r}^N e^{-\beta U(\lambda)}} \\ &= \left\langle \frac{\partial U}{\partial \lambda} \right\rangle_\lambda \end{aligned} \quad (4.12)$$

Integration over λ results in:

$$\Delta F = \int_0^1 d\lambda \left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda \quad (4.13)$$

In practice, the value $\left\langle \frac{\partial U(\lambda)}{\partial \lambda} \right\rangle_\lambda$ will be calculated for a set of λ values, and then integrated over λ , which results in the total free energy for the perturbation from state A to state B.

4.1.3 Obtaining a free energy along an order parameter

For causes like for example coarse graining, an effective potential might be needed to describe the effective interaction of a medium, including both energetic and entropic effects, for a particle-particle interaction. When applied in the form of a free energy profile along the particle-particle distance coordinate, it can be called a potential of mean force (PMF), and in that interpretation it will be used in the current thesis.

The distance-dependent free energy potential between two particles is a function of the radial distribution function $g(r)$ in the form:

$$V_{PMF}(r) = -k_B T \ln g(r) + \text{constant} \quad (4.14)$$

And therefore a potential of mean force can be obtained by measuring the distribution of particle-particle distance r during a run. However, low sampling may occur for certain values of the particle-particle coordinate due to barriers or other factors, and a bias might be necessary to get a correct PMF profile. A frequently

used method to obtain a PMF in these circumstances is by constraining the particle-particle distance along small intervals over the range of interest. The mean value of constraint force $F_c(r)$, the force needed to replace the particle at its constraint position after being forced away due to forces arising from the other parts of the system, can be calculated over the run. The integral of these mean forces over a distance will result in the potential of mean force. In the current thesis (chapters 7 and 8) this potential will be directed solely along the z -axis, perpendicular to the flat surface that is studied, in which case the free energy potential of mean force can be written as:

$$V_{PMF}(z) = \int_{\infty}^z dz \langle F_c(z) \rangle + constant \quad (4.15)$$

At infinite distance the potential will be zero, and by integrating from the z at the point of zero potential towards $z = 0$, no constant value needs to be added to the integral.

4.2 Distinction of enthalpy-entropy compensation in the free energy

Free energies, as retrieved from experiments or simulations, give an indication of how favorable non-covalent molecule-molecule interactions are. It does not, however, explain which molecular properties, interaction energy or entropy, steer this behavior. Even after splitting the free energy into an entropy and interaction energy, a compensating term can be identified. To retrieve the factual driving forces within the free energy, one needs a redistribution of the contributing terms to the free energy, resulting in a set of terms of which the compensating terms are removed. Even though this would end up as a set of purely theoretical contributions that can not be calculated experimentally, the possibility to calculate every term separately in computer simulations can help understand the driving forces for molecular processes. An example that shows the value of this approach will be given in chapter 5, where the driving forces for hydrophilic vs. hydrophobic solvation will be discussed.

The method discussed here is constructed for the calculation of solute hydration free energy [119, 120]. Instead of defining an overall $\Delta G = \Delta H - T\Delta S$ equation, a separation is made between solvent-solvent (WW) energy and entropy and

solute-solvent energy (SW). The excess potential as defined by Sanchez et al.[121], describing excess interaction of solute particle S with its solution environment as compared to the interaction of the interaction of particle S with its surroundings if that was an ideal gas with the same number density.

$$\mu_{ex,S} = \Delta u_{SW} - T\Delta s_{SW} \quad (4.16)$$

The solute (S) excess chemical potential ($\mu_{ex,S}$) can be defined using Widom's potential distribution theorem[118], which resembles the particle insertion method described previously.

$$\mu_{ex,S} = -\beta^{-1} \ln \langle e^{-\beta\psi} \rangle_A \quad (4.17)$$

In equation 4.17, ψ denotes the interaction energy of the inserted solute molecule with the other (solvent) molecules, $\beta = (k_B T)^{-1}$ with k_B the Boltzmann constant, and the angular brackets a constant volume-temperature ensemble average, in this case in the ensemble including only solvent (state A). By converting equation 4.17 to the ensemble in which both solvent and solute particles are present (state B), the equation can be inverted and decomposed in an energetic and entropic contribution[121, 122]:

$$\begin{aligned} \mu_{ex,S} &= \beta^{-1} \ln \langle e^{\beta\psi} \rangle_B \\ &= \langle \psi \rangle_B + \beta^{-1} \ln \langle e^{\beta(\psi - \langle \psi \rangle_B)} \rangle_B \\ &\equiv \Delta u_{SW} - T\Delta s_{SW} \end{aligned} \quad (4.18)$$

In equation 4.18, $\Delta u_{SW} = \langle \psi \rangle_B$ and therefore $\Delta s_{SW} = -k_B \ln \langle e^{\beta(\psi - \langle \psi \rangle_B)} \rangle_B$. Using the inequality $\langle e^{\beta\psi} \rangle_B \geq e^{\beta\langle \psi \rangle_B}$ one sees that Δs_{SW} is always negative or zero and thus $-T\Delta s_{SW}$ adds positively to $\mu_{ex,S}$. For molecules with attractive interactions Δu_{SW} is always negative. Negative values of $\mu_{ex,S}$ therefore always result from favorable solute-solvent interactions ($\Delta u_{SW} < 0$) overriding the solute-solvent entropy. These two terms can be described as separate processes, where $-T\Delta s_{SW}$ may be interpreted as the work of creating a cavity that has the solvent molecules in the appropriate positions and orientations to accommodate all chemical moieties of the solute. Δu_{SW} can be more intuitively understood as the interaction energy of the solute with the properly formed cavity.

Based on the formalism introduced above, it is possible to find fully compen-

sating terms in the excess chemical potential. Consider a (constant pressure) excess partial molar enthalpy ($h(ex, S)_P$) which will, in addition to the solute-solvent energy Δu_{SW} , contain a contribution arising from changes of solvent-solvent interactions. This contribution is different depending on whether the pressure or volume is kept fixed[123, 124]. The solvent-solvent energy (enthalpy) change in the constant pressure ensemble will be applied as:

$$(\Delta h_{WW})_P = (\Delta u_{WW})_P + p\Delta v \approx (\Delta u_{WW})_P \quad (4.19)$$

This energy change is localized in the solute hydration shell unlike in the constant volume case where it includes a nonlocal bulk response contribution[123, 124]. Below, the subscript $(\dots)_P$ is dropped, but one should keep in mind that in the formulation described here excess partial molar enthalpies and entropies are evaluated at constant pressure. Assuming the energies Δu_{SW} and Δh_{WW} are additive yields:

$$h_{ex,S} = \Delta u_{SW} + \Delta h_{WW} \quad (4.20)$$

Because the excess chemical potential is defined as $\mu_{ex,S} = h_{ex,S} - Ts_{ex,S}$, the excess partial molar entropy is given by:

$$s_{ex,S} = \Delta s_{SW} + \Delta h_{WW}/T \quad (4.21)$$

Hence, $h_{ex,S}$ and $Ts_{ex,S}$ contain a contribution (Δh_{WW}) which never impacts $\mu_{ex,S}$ (exact energy-entropy compensation). Although the definitions of Δu_{SW} and Δs_{SW} (equation 4.18) take slightly different forms in the constant P-T and constant V-T ensembles their values are ensemble-independent.

Equations 4.20 and 4.21 identify Δu_{SW} and Δs_{SW} as the non-compensating terms in the excess potential. However, a physical interpretation of these terms is needed. The interpretation of Δu_{SW} is straightforward, it can be seen as the interaction energy between a solute S and the surrounding solvent W. To better understand the more complicated physics inherent in Δs_{SW} it is useful to think of it as having two contributions of which both are negative. First, Δs_{SW} has a cavity contribution that arises from repulsive solute-solvent interaction energies ($\psi > 0$) with corresponding solvent configurations that violate the excluded volume constraint imposed by the solute. This contribution reflects the tendency of the solvent to close the solute cavity (or "squeeze-out" the solute). The second contribution arises

from fluctuations of the interaction energy in configurations where the interaction is attractive ($\psi \leq 0$). In chapter 5, the alternative formulation of Δs_{SW} as derived by Sanchez et al.[121] will be used:

$$\begin{aligned}\Delta s_{SW}/k_B &= \ln P_{ins} - \ln \langle e^{\beta(\psi - \langle \psi \rangle_B)} \rangle_a - \ln P_a \\ &= \ln P_{ins} - [(\langle \psi^2 \rangle_B - \langle \psi \rangle_B^2)\beta^2/2 + \dots]_a - \ln P_a\end{aligned}\quad (4.22)$$

where the subscript 'a' indicates that the average is taken under the condition that the solute-solvent interaction energy is attractive ($\psi < 0$). The quantity P_{ins} is the probability that a randomly inserted solute molecule into a system of only solvent molecules will experience an attractive or zero interaction ($\psi \leq 0$). P_a is the probability that the solute molecule in the fluid will have an attractive energy. Because P_a will be very close to unity under the given condition, the $\ln P_a$ term can be ignored. The second term on the right hand side of equation 4.22 will always be zero in case ψ is independent of the positions and orientations sampled by solvent molecules vicinal to the solute. In equation 4.22, the fluctuation term $(\langle \psi^2 \rangle_B - \langle \psi \rangle_B^2)\beta^2/2$ expresses the fact that the available configuration space is biased by attractive interactions (i.e. solute-solvent attractive interactions bias positions and orientations of the vicinal solvent molecules). The $(\langle \psi^2 \rangle_B - \langle \psi \rangle_B^2)\beta^2/2$ term is always positive, causing a decrease of the entropy. This can be interpreted as a loss of configuration space to be sampled in presence of attractive interaction, which causes a loss of entropy.

5 A thermodynamic analysis of benzene solvation

Standard force field parameterizations are often based on calculation of the hydration free energy and comparing this with experimental values or estimates. The disadvantage of this method is that a correct solvation free energy may be reproduced by different atomistic force fields with considerable different values for entropy and enthalpy, as these values partly compensate as shown in chapter 4. In the current case, the quality of the force field is validated not only by the hydration free energy, but also by calculating the enthalpy, entropy, and hydration heat capacity. This is not only a well-defined validation of the force field, but can also give insight into the microscopic driving forces for solute hydration. Benzene is chosen as the solute to study, as it combines a slight hydrophobicity with weak hydrogen bond acceptor. The benzene GROMOS43a1 force field is shown to reproduce experimental values of hydration enthalpy, hydration entropy, and hydration heat capacity, and will therefore be used in this thesis in the multiscale study of the adsorption of a solute from bulk solution to a metal surface in an aqueous environment (see chapter 7). Comparison with a hypothetical benzene model that has no hydrophilic interaction will demonstrate differences in the mechanism of hydrophobic and hydrophilic solvation for this type of solute.

5.1 Benzene as a hydrophilic molecule

Hydrophobic hydration is believed to play an important role in protein folding and other self assembly processes in water and therefore a popular object of research.[125–129] Hydrophobic hydration of nonpolar solutes (i.e. aliphatic hydrocarbons), defined as the process of transferring the solute molecule from the gas phase into room temperature water, is characterized by unfavorable (negative) hydration entropies

and favorable, but smaller in magnitude, hydration enthalpies (negative). In room temperature water, first shell water molecules form a hydrogen bonded cage structure surrounding the nonpolar solute. In a cage structure, the water molecules do not "waste" hydrogen bonds by pointing them at the solute; instead they orient their O-H bonds tangential to the solute surface in order to maximize hydrogen bonding with vicinal water molecules.[129] Historically, Frank and Evans[130] proposed that the large entropy that opposes solute transfer into water arises from the cost of ordering the waters in this way.

The mechanism causing hydrophobicity of aromatic hydrocarbons (i.e. benzene, toluene) has a different nature than that causing the hydrophobicity of aliphatic hydrocarbons. In the case of aromatic hydrocarbons the enthalpic contribution predominates over the entropic contribution to the free energy of hydration.[131] This results in a negative value of the hydration free energy of aromatic hydrocarbons, in contrast to the positive values obtained in the case of aliphatic hydrocarbons. In terms of aqueous solubilities, these differences cause the solubility of, e.g., hexane being almost 20 times lower than the solubility of benzene, although these hydrocarbons have similar molecular weight.[132] Moreover, the heat capacity change upon transfer into aqueous solution (the thermodynamic hallmark that indicates the presence of structured hydration water), normalized to the water accessible surface area, is smaller for the aromatic hydrocarbons than the aliphatic ones[133, 134].

Benzene is a slightly polar molecule due to the nature of its π -electron system that acts as a weak hydrogen bond acceptor, which is for this case defined as a bond with an interaction energy high enough to be stable (higher than about 7.5 kJ/mol ($3 k_B T$)) and a small distance between the hydrogen bond donor group (OH) and the benzene center of mass.[135–138] It has been argued that higher aqueous solubilities of aromatic hydrocarbons are probably due to formation of energetically favorable hydrogen bonds between the aromatic hydrocarbons and water.[131] This explanation is in accord with results of computer simulation studies using semi-empirical force fields that showed that the hydration free energy of benzene can only be reproduced when adopting a model that includes a permanent charge distribution of benzene, which can be expressed by an electrostatic multipole expansion where the quadrupole moment is the first nonzero term.[139] A consequence of this charge distribution is that the benzene-water potential energy is strongly dependent on the mutual orientation thereby accounting for weak hydrogen bonding interactions.

Using available experimental and computer simulation data, Graziano and Lee[140] however argued that formation of weak O-H $\cdots\pi$ hydrogen bonds between water and benzene is likely to be largely enthalpy-entropy compensating therefore not explaining the higher solubility of aromatic hydrocarbons. Instead they argue that the large van der Waals interaction energy overwhelms the free energy cost of cavity formation at room temperature in contrast to aliphatic hydrocarbons where the free energy cost of cavity creation dominates.

In the current chapter, the hydration thermodynamics of benzene is obtained by classical molecular dynamics simulations of benzene in simple point charge (SPC [141]) water. To quantify how benzene-water hydrogen bonding changes the hydration thermodynamics, a benzene model that favors formation of weak water-benzene hydrogen bonds ("real" benzene) was compared to a "van-der-Waals-benzene" model obtained by removing all partial charges of the first model while keeping the exchange repulsion and dispersion interaction terms unaffected. In addition to calculating the free energy of hydrating the "real" benzene- and van-der-Waals-benzene models, the enthalpy, entropy, and heat capacity of hydration were calculated. To understand the difference in aqueous solubility between these benzene models better, contributions to the hydration enthalpy and entropy arising from benzene-water interactions (the solute-solvent interaction energy and solute-solvent entropy) and reorganization of the solvent were examined.

5.2 Thermodynamics

For convenience, the main results of the derivation by Yu and Karplus[120] (as discussed in section 4.2) of solute hydration free energy (the Ben-Naim pseudo-chemical potential[119] $\mu_{ex,S}$), and the corresponding hydration enthalpy ($h_{ex,S}$) and entropy ($s_{ex,S}$), are repeated here.

$$\mu_{ex,S} = \Delta u_{SW} - T\Delta s_{SW} \quad (5.1)$$

$$(h_{ex,S})_P = \Delta u_{SW} + (\Delta h_{WW})_P \quad (5.2)$$

$$(s_{ex,S})_P = \Delta s_{SW} + (\Delta h_{WW})_P / T \quad (5.3)$$

In equations (5.1, 5.2, 5.3), Δu_{SW} denotes the solute-water interaction energy

(the solute-solvent energy), Δs_{SW} the entropy change of solute-water interaction (the solute-solvent entropy), and $(\Delta h_{WW})_P$ the (constant pressure) water-water enthalpy change (solvent reorganization enthalpy).

Whereas $\mu_{ex,S}$ is ensemble independent, $h_{ex,S}$ and $s_{ex,S}$ do depend on the insertion condition (constant P,T; constant V,T) with the ensemble dependence occurring in the Δh_{WW} term[123, 124]. Here, only conditions of constant pressure and temperature (at which experiments are usually performed) are considered, and the subscript $(\dots)_P$ is dropped. In hydrophobic hydration, Δh_{WW} is usually small (in comparison to solvent reorganization energies in organic solvents) and positive[142] because water hydrates nonpolar solutes without significantly sacrificing hydrogen bonding.

In computer simulation studies, $\mu_{ex,S}$ may be obtained by thermodynamic integration (TI), perturbation (TP), or related methods [124], some of which have been discussed in chapter 4. Since Δu_{SW} is obtained straightforwardly by keeping an average of the solute-water interaction energy at the end-point simulation ($\lambda=1$) in TI/TP, Δs_{SW} follows directly from eq (5.1). Note that both Δu_{SW} and Δs_{SW} are negative (see section 4.2) while Δh_{WW} is positive. Equations (5.1, 5.2, 5.3) have several important consequences:

- (1) The water-water enthalpy change has no impact on $\mu_{ex,S}$, since Δh_{WW} is excluded from this definition of $\mu_{ex,S}$.
- (2) However, cancellation of Δh_{WW} in $\mu_{ex,S}(= h_{ex,S} - Ts_{ex,S})$ does not mean that the extent to which water-water hydrogen bonds are disrupted is irrelevant in understanding $\mu_{ex,S}$ because Δu_{SW} and Δs_{SW} are defined as ensemble averages containing the water-water interaction implicitly.[120] The solute-solvent entropy change ($-T\Delta s_{SW}$) can conceptually be interpreted as the free energy to create a solute cavity that has all solvent molecules properly oriented to accommodate all chemical moieties of the solute[143], while the solute-solvent energy (Δu_{SW}) may be interpreted as the remaining solute-cavity interaction contribution to $\mu_{ex,S}$. In hydrophobic hydration, first shell water molecules orient to minimize their loss of H-bonds. Any reduction of orientational entropy associated with preferred water orientations (driven by W-W interactions) therefore appears in Δs_{SW} . The cavity formation work ($-T\Delta s_{SW}$) also increases with the excluded volume radius (i.e. loss of translational entropy).

- (3) In hydrophilic hydration, polar solute-solvent interactions (e.g. dipolar interactions, dipole-quadrupole interactions) bias orientations of hydration waters, too. This process, the reduction of orientational freedom due to attractive interaction, may in fact reduce the solute-solvent entropy stronger than the biasing of water orientations close to nonpolar solutes. In that case, higher aqueous solubilities of polar molecules over nonpolar molecules result from the presence of favorable electrostatic interactions with the solvent.

5.3 Computational details

5.3.1 Thermodynamic calculations

Excess chemical potentials $\mu_{ex,S}$ were calculated by TI using at least 50 λ -values. At each new λ , the system was first equilibrated for 50 ps after which the free energy derivative was sampled for 500 ps. A soft-core λ scaling was used[93] to avoid singularities of the free energy derivative at the end-points. The excess partial molar enthalpy was calculated using the expression,

$$h_{ex,S} = \langle U_{SW} + U_{WW} + PV \rangle_{solution} - \langle U_{WW} + PV \rangle_{pure H_2O} \quad (5.4)$$

where U_{SW} denotes the solute-water interaction energy, U_{WW} the sum of interaction energies of the water molecules with all other water molecules, and PV a pressure-volume work term. The brackets denote a constant pressure-temperature ensemble average. Two constant pressure-temperature simulations, one of the aqueous solution (1 solute, N solvent molecules) and one of the neat solvent (N solvent molecules), are performed and the average potential energies are subtracted (the $P\Delta V$ term is usually small). Because both terms on the right hand side of eq 5.4 are of $O(N)$ whereas $h_{ex,S}$ is of $O(1)$, the statistical accuracy obtained in MD runs of several hundreds of picoseconds is usually poor. Therefore, long (90 ns) trajectories were produced to sample the two averages in eq 5.4 resulting in statistical inaccuracies of $h_{ex,S}$ smaller than 0.05 kJ/mol (as calculated by block averaging[144]). Excess partial molar entropies were obtained from the calculated excess chemical potential and enthalpy using $Ts_{ex,S} = h_{ex,S} - \mu_{ex,S}$. Solute-solvent entropies were

obtained from the calculated excess chemical potential and solute-solvent energy using $T\Delta s_{SW} = \Delta u_{SW} - \mu_{ex,S}$ (eq 5.1). The solute-solvent energy was obtained from the simulations by taking the average benzene-water interaction energy. Heat capacities of hydration were calculated from the temperature dependence of $h_{ex,S}$ using the finite difference expression

$$c_{P,ex,S}(T) = \frac{h_{ex,S}(T + \Delta T) - h_{ex,S}(T - \Delta T)}{2\Delta T} \quad (5.5)$$

A relatively large temperature difference $\Delta T = 20K$ was chosen in the current simulations to ensure that a quantifiable difference in molar enthalpies could be found, even though it might be not completely correct if the enthalpy change is nonlinear on this temperature range. The excess partial molar enthalpies $h_{ex,S}(T + \Delta T)$ and $h_{ex,S}(T - \Delta T)$ were calculated using eq 5.4 based on 90 ns trajectories. The statistical inaccuracy of $c_{P,ex,S}$ is determined by that of $h_{ex,S}$ and amounts to $15 - 20 Jmol^{-1}K^{-1}$.

5.3.2 Simulation Details.

All simulations were performed using the Gromacs 3.2.1 simulation package[92, 93] and were based on an equilibrated cubic, periodic simulation box containing 1500 water molecules and 1 benzene molecule. The simple point charge (SPC) model of water was used[141]. A complete force field description of rigid SPC water is given in table 6.1 and figure 6.2 in chapter 6. Benzene was modeled using the GROMOS 43A1 force field parameters.[88] For reasons of comparison, a benzene model based on the recent GROMOS 53A6 parameters was analyzed as well.[98] The non-bonded force field parameters are summarized in Table 5.1. Its geometry is given in figure 7.2(B) in chapter 7. Bonded parameters were identical in the 43A1 and 53A6 GROMOS force field[88, 98], and contained improper dihedrals to enforce the flat conformation of the benzene ring. Intramolecular Lennard-Jones and electrostatic interactions were accounted for between atom pairs separated by more than 3 bonds. The simulation temperature was kept constant at 302 K by weakly coupling to a temperature bath with a relaxation time of 0.1 ps.[102] The pressure was maintained at 1 atmosphere by also applying the weak coupling algorithm with a relaxation time

of 0.5 ps and an isothermal compressibility of $45.75 \times 10^{-5} (kJmol^{-1}nm^{-3})^{-1}$. The equilibrated box length was 3.592 nm. For non-bonded interactions, a twin-range method with cutoff radii of 0.8 and 1.4 nm was used. Outside the longer cutoff radius a reaction field correction was applied with a relative dielectric permittivity of 54.0. The integration time step was 2 fs, the pair list for pairs within the inner cutoff and the energies and forces for pairs between the inner and outer cutoff radii were updated every 10 fs. All bond lengths were kept constant using the SHAKE algorithm[95] using a relative geometrical tolerance of 10^{-4} .

atom	$C_{12}(i, j)$ ($10^{-6}kJmol^{-1}nm^{12}$)	$C_6(i, j)$ ($10^{-3}kJmol^{-1}nm^6$)	q (e)
Benzene (GROMOS 43A1)[88]			
C	3.374569	2.340624	-0.10
H	0.015129	0.084640	0.10
Benzene (Van der Waals)			
C	3.374569	2.340624	0.00
H	0.015129	0.084640	0.00
Benzene (GROMOS 53A6)[98]			
C	4.937284	2.340624	-0.14
H	0.015129	0.084640	0.14
SPC Water [141]			
O	2.634129	2.617346	-0.82
H	0.00	0.00	0.41

Table 5.1: Nonbonded Interaction Parameters (Nonbonded interaction function consisted of the sum of equations 3.6 and 3.28. Combination rules were used as mentioned in equations 3.8 and 3.9).

5.3.3 Benzene-Water Interaction.

The benzene-water interaction potential is strongly orientation dependent. Therefore, the ability of the force field to reproduce this dependency was analyzed. Based on the relative orientations shown in figure 5.1, the interaction energy between a benzene-water (SPC) pair was calculated as a function of the distance r defined in figure 5.1 (using the GROMOS 43A1 parameters in table 5.1). Fig 5.2 shows

the corresponding potential energy curves (denoted here as "real O in" and "real O out"), which also includes the interactions obtained when not accounting for the electrostatic part of the interaction (denoted here as "van-der-Waals O in & O out"). The interaction energies for a water molecule oriented perpendicular to the aromatic plane (figure 5.2 (A)) are in good agreement with the energies reported by Linse[136,145,146] which are based on quantum calculations using a Hartree-Fock self-consistent-field approximation combined with a second order perturbation procedure to account for the dispersion energy. The dashed line in Fig 5.2 a ("real O out"; benzene-water hydrogen bonding) shows a minimum at approximately 3.0 Å with a corresponding energy between 13 and 14 kJ/mol, corresponding to almost 5 k_{BT} at 300K. In quantum calculations,[136,145,146] the minimum is found at the same distance with an energy of 12-13 kJ/mol. The configuration with the oxygen pointing inward (dotted curve in Fig 5.2 (A)) is repulsive and is also in good agreement with the corresponding benzene-water interaction energy reported from quantum calculations in reference [136]. For conformations where the water is aligned in the benzene plane (figure 5.2 (B)), the agreement between the current data and those from the work of Linse et al. is less satisfactorily. In the quantum calculations, the energy minimum for the oxygen-inward conformation is found around 4.5 Å with a corresponding energy of 7-8 kJ/mol, while with the GROMOS 43A1 force field this minimum is located at 5.2 Å with a corresponding energy of 3.6 kJ/mol. For the configuration with the oxygen pointing outward, the quantum calculations show a slightly stronger repulsion than the atomistic force field data presented here.

5.4 Results and Discussion

5.4.1 Hydration Structure: Radial distribution functions

Due to the anisotropic benzene-water interaction potential, the characterization of benzene hydration structure should preferably distinguish between water molecules inside the volume above the benzene plane and water molecules located outside this volume. The approach of Linse et al.[146] to subdivide the space around the benzene was adopted. Water molecules inside the conical volumes above and below the benzene plane and remaining water molecules were considered separately. The

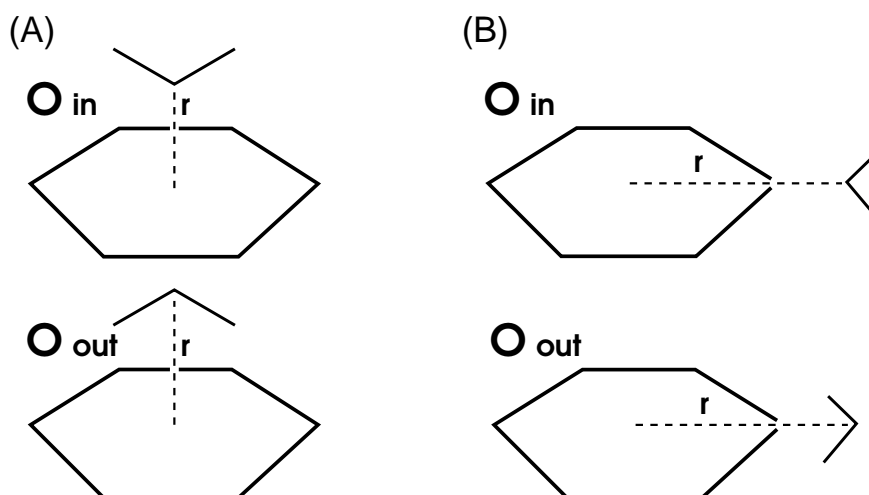


Figure 5.1: Relative orientations used for calculating the water-benzene interaction. The distance r is defined between the water oxygen and the ring center of mass. In orientation "O out", the water dipole moment is pointing toward the solute, in orientation "O in" the water dipole moment is oriented outward. Two symmetry axes are considered: (A) the water molecule is located above the benzene plane. (B) the water molecule is located in the plane of the benzene ring.

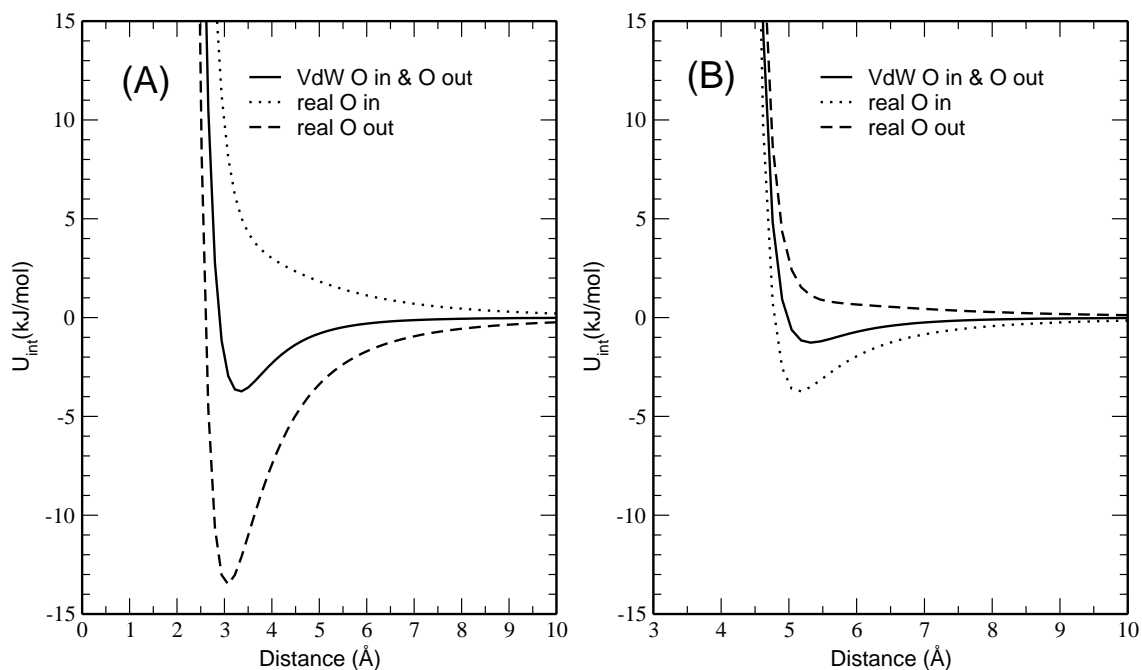


Figure 5.2: Potential energy curve for SPC water - benzene interaction. (A) the water molecule is located above the benzene plane. (B) the water molecule is located in the plane of the benzene ring (see figure 5.1). Solid line: van-der-Waals benzene (the quadrupole-dipole interaction is omitted); Dashed line: water having its dipole pointing inward; Dotted line: water having its dipole pointing outward (see figure 5.1).

cones make a 45° angle with the benzene symmetry axis (see figure 5.3 (A)). Water molecules located inside the conical volumes are more likely hydrogen bonded to benzene than molecules located outside. Fig 5.4 shows the in-cone water-benzene center of mass radial distribution function (RDF) for "real" benzene (figure 5.4 (A)), modeled with the GROMOS 43a1 parameters, and the van-der-Waals-benzene model (figure 5.4 (B)). For the case of the "real" benzene, the first peak of the hydrogen at 2.25 \AA clearly indicates that water OH-groups can approach benzene within short distances, and hydrogen bond formation is plausible. The peak area up to 3 \AA corresponds to 1.0 benzene-water hydrogen bonds. The larger oxygen peak is split in two parts. The first maximum occurs at 3.25 \AA and corresponds to the water molecule donating a hydrogen bond to benzene (i.e. the distance between the first maximum of the hydrogen RDF and that of the oxygen RDF is exactly a OH bond length). The second oxygen peak at 4 \AA corresponds to water not directly hydrogen bonded to benzene. The in-cone RDFs for the van-der-Waals benzene (figure 5.4 (B)) are typical of hydrophobic hydration. The peaks are narrower and the maxima are higher compared to the "real" benzene system, indicating significantly more structured hydration water. Moreover, the first maximum for the oxygen and hydrogen RDFs are located at the same distance, which suggests that water orients one of its OH bonds parallel to the surface of the nonpolar solute. The first peak of the hydrogen RDF has a broad shoulder extending to larger distances, which arises from the second OH bond that orients toward the bulk where it donates a hydrogen bond to water molecules in the second solvation shell. The latter becomes apparent from the second maximum of the oxygen RDF being located closer to the solute than the second maximum of the hydrogen RDF. The out-of-cone RDFs for "real" benzene and van-der-Waals benzene are shown in figs 5.5a and 5.5b, respectively. For both systems the observed structure is typical of hydrophobic hydration with no significant differences between the realistic and van-der-Waals benzene models.

5.4.2 Orientational Distributions

The probability density of orientation of the water OH bonds with respect to the vector connecting the benzene center of mass and the water oxygen is shown in figure 5.6. The orientation angle θ is defined in figure 5.3 (B). The data in figure 5.6 apply to water molecules in the first shell ($r < 6 \text{ \AA}$), which on average contains 26.5 molecules. No distinction is made here between water molecules located inside or

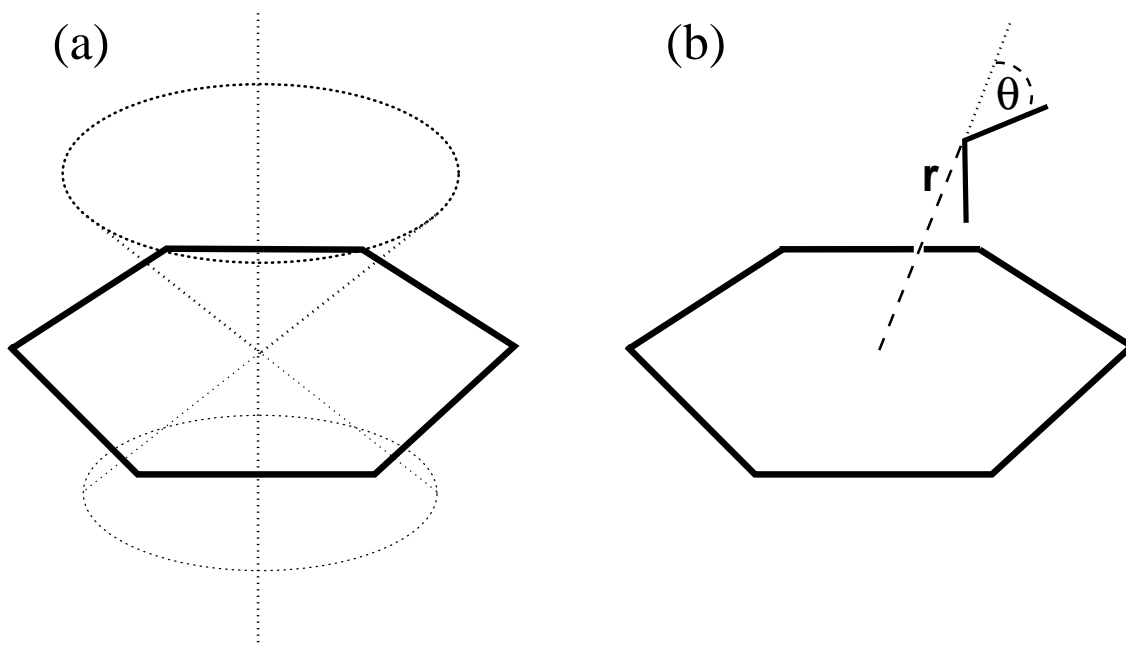


Figure 5.3: (A) Definition of the 45° cone as used in this study and previously in the work of Linse et al. (B) Definition of the angle θ defining the orientation of the OH bond with respect to the vector connecting O to the center of the benzene ring.

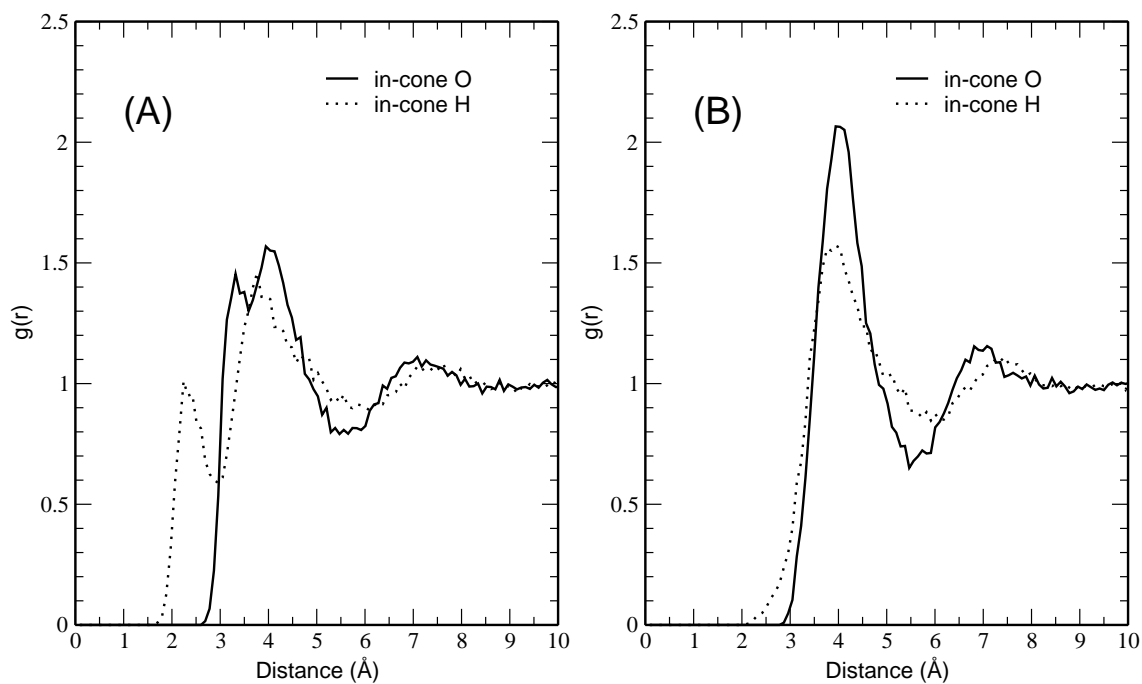


Figure 5.4: In-cone benzene-water radial distribution functions. Solid line: benzene center-of-mass-to-water oxygen RDF; Dashed line: benzene center-of-mass-to-water hydrogen RDF. (A) "real" benzene (GROMOS43a1). (B) van-der-Waals benzene.

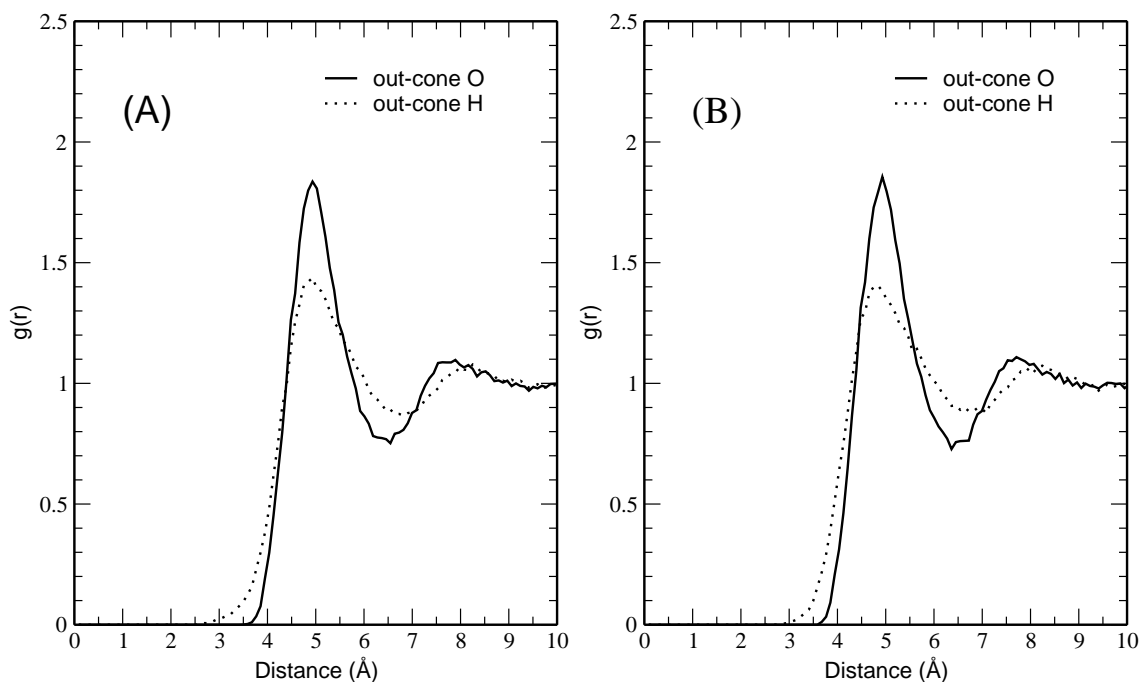


Figure 5.5: Out-of-cone benzene-water radial distribution functions. Solid line: benzene center-of-mass-to-water oxygen RDF; Dashed line: benzene center-of-mass-to-water hydrogen RDF. (A) "real" benzene (GROMOS43a1). (B) van-der-Waals benzene.

outside the conical volumes. The distribution $P'(\theta)$ obtained just from the statistical sampling of the angles was rescaled according to $P(\theta) = P'(\theta)/\sin(\theta)$, accounting for the volume elements associated with the angle θ . For the GROMOS 43A1 van-der-Waals model, a higher occurrence of OH orientation was observed at both 0° (radially outward) and $110-120^\circ$. Postma et al.[147] made similar observations in a study of the hydration of spherical cavities with a thermal radius of 3 \AA (the orientationally averaged cavity radius of van-der-Waals benzene equals $\approx 3.5 \text{ \AA}$). They showed that water molecules may have one OH-bond directed towards the bulk whereas the other OH bond is directed parallel to the cavity surface, but may also have both of their OH-bonds parallel to the cavity surface. The water orientational distribution corresponding to "real" benzene shows somewhat weaker maxima at 0° and 110° and has an additional weak maximum at 180° (OH radially inward), which arises from benzene-water hydrogen bonding. The orientational distribution of second shell hydration waters (dotted line in figure 5.6) neither reveals strong preferential OH directions nor does it show differences for the real- and van-der-Waals solutes.

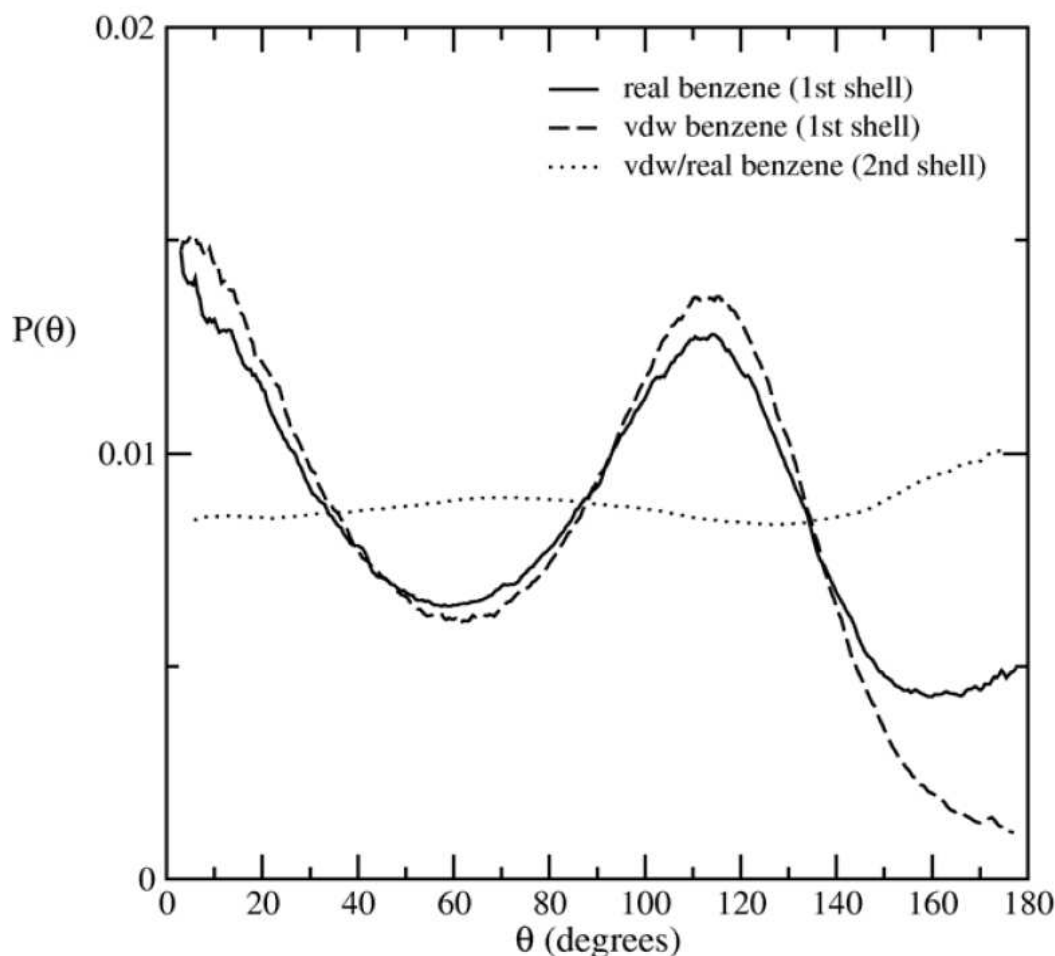


Figure 5.6: Probability density of orientation of the OH direction with respect to the vector connecting the benzene center of mass and water oxygen (see figure 5.3 (B)). Normalization is performed using: $\int P(\theta) \sin \theta d\theta = 1$. Data apply to molecules in the first shell ($r < 6 \text{ \AA}$) or second shell ($6 < r < 9 \text{ \AA}$). Solid line: GROMOS43a1 "real" benzene model (first shell); Dashed line: van-der-Waals benzene (first shell); Dotted line: benzene (second shell).

5.4.3 Hydration Thermodynamics

In table 5.2 all thermodynamic hydration quantities are summarized. For "real" benzene, the calculations were performed using the 43A1[88] and 53A6[98] GROMOS parameters (see table 5.1). The thermodynamic quantities for van-der-Waals benzene were obtained from simulations using 12-6 Lennard-Jones parameters from the 43A1 force field. The best agreement with the experimental data is found for the 43A1 GROMOS parameter set and will be discussed here. Not only the benzene ex-

cess chemical potential agrees satisfactorily with the experimentally reported value, but also the excess partial molar enthalpy and entropy are closely reproduced. Especially the latter quantities provide a good verification of the quality of a particular force field. Although the excess chemical potential can be predicted correctly, it may happen for the wrong reason because errors in the energy may cancel out against errors in the entropy. The isobaric heat capacity of benzene hydration is predicted most accurately as well with the GROMOS 43A1 force field. The hydration thermodynamics of the van-der-Waals benzene model differs from the real model and is typical of hydrophobic hydration: the excess chemical potential is positive and also the isobaric heat capacity of hydration is significantly more positive than for "real" benzene. The excess partial molar entropies of van-der-Waals and "real" benzene do not significantly differ. The excess partial molar enthalpy of real (43A1) benzene is however -4.5 kJ/mol more favorable compared to that of van-der-Waals benzene, hence the larger negative excess chemical potential of "real" benzene is mainly enthalpic. This enthalpy difference compares with a value of -5.4 kJ/mol predicted by Makhatadze and Privalov[131] based on combining experimental data for benzene and aliphatic hydrocarbons, the latter corrected for the benzene accessible surface area. Based on a similar calculation, Makhatadze and Privalov predict a change of the excess chemical potential and excess partial molar entropy of -13.4 kJ/mol and +26.8 J/mol/K, respectively, which the current calculations do not reproduce.

5.4.4 Solute-Solvent Contributions

The solute-solvent interaction energies, Δu_{SW} , solute-solvent entropies, $T\Delta s_{SW}$, and water reorganization enthalpies, Δh_{WW} , are shown in table 5.3 for the 43A1 "real" benzene and the 43A1 van-der-Waals benzene models. The benzene-water energy is -13.8 kJ/mol more favorable than the van-der-Waals-benzene - water energy (T=302 K). Because benzene on average accepts one hydrogen bond from water and the benzene-water hydrogen bond energy equals approximately 13 kJ/mol (figure 5.2 (A)), this result is expected.

The enthalpy change of water reorganization was calculated based on eq 5.4, which can also be written as:

T (K)	$\mu_{ex,S}$ (kJ/mol)	$h_{ex,S}$ (kJ/mol)	$S_{ex,S}$ (J/mol/K)	$c_{p,ex,S}$ (J/mol/K)
Experimental[131]				
278		-35.7	-108.4	318.8
298	-3.6	-29.6	-87.2	291.6
323		-22.6	-64.6	268.1
43A1 benzene				
282		-32.7		
302	-4.8	-29.3	-82.2	277.5
322		-21.6		
53A6 benzene				
282		-34.5		
302	-6.7	-31.1	-81.8	217.5
322		-25.8		
43A1 benzene, van der Waals only				
282		-29.7		
302	1.0	-24.8	-86.5	327.5
322		-16.6		

Table 5.2: Hydration thermodynamics of benzene and "van der Waals benzene" (The excess chemical potentials ($\mu_{ex,S}$) were computed by thermodynamic integration, the excess partial molar enthalpies ($h_{ex,S}$) were computed using eq 5.4 in which total potential energies of the solution (water + benzene) and solvent (water) MD simulations were averaged over 90 ns time periods. The excess partial molar entropies ($S_{ex,S}$) were obtained from $TS_{ex,S} = h_{ex,S} - \mu_{ex,S}$. Heat capacity changes ($c_{p,ex,S}$) were calculated by finite difference (eq 5.5).)

$$\begin{aligned}
h_{ex,S} &= \langle U_{SW} \rangle_{solution} + [\langle U_{WW} + PV \rangle_{solution} - \langle U_{WW} + PV \rangle_{pure H_2O}] \\
&= \Delta u_{SW} + \Delta h_{WW}
\end{aligned}
\tag{5.6}$$

The water reorganization enthalpies (Δh_{ww}) are large and positive contributing significantly to the excess partial molar enthalpies. For dissolving a solute in its own pure liquid at constant P and T , the solvent reorganization enthalpy equals the

T (K)	Δu_{SW} (kJ/mol)	$T\Delta S_{SW}$ (kJ/mol)	Δh_{WW} (kJ/mol)
43A1 benzene			
282	-61.2		28.5
302	-59.4	-54.6	30.1
322	-57.5		35.9
43A1 benzene, van der Waals only			
282	-47.1		17.4
302	-45.6	-46.6	20.8
322	-43.9		27.3

Table 5.3: Solute-solvent energy change (Δu_{SW}), solute-solvent entropy change ($T\Delta S_{SW}$), and solvent reorganization energies (Δh_{WW}) for hydrating "real" benzene and van-der-Waals-benzene.

average potential energy of the liquid.[148] Transferring for example a SPC water molecule from the saturated vapor phase into the liquid SPC water phase involves a solvent reorganization enthalpy of 41.5 kJ/mol at 298 K. For benzene (43A1 model) in its own liquid this enthalpy is 33.2 kJ/mol. If one compares these numbers to the reorganization enthalpies in table 5.3 it shows that, in particular for van-der-Waals benzene, the water reorganization enthalpy is small. For the "real" benzene model, Δh_{WW} is larger because benzene-water hydrogen bonding happens at the expense of water-water hydrogen bonding. The temperature dependencies of Δu_{SW} and Δh_{WW} are interesting because they provide further clues on the different heat capacity changes of real- and van-der-Waals benzene (table 5.2). In the 40 K temperature interval, Δu_{SW} increases with 3.7 kJ/mol and 3.2 kJ/mol for real- and van-der-Waals benzene, respectively. For the "real" benzene model it was found that this energy change was almost exclusively due to a reduction of the benzene-water van der Waals energy with increasing temperature. The benzene-water electrostatic interactions changed only little (+0.7 kJ/mol) in this temperature interval. Interestingly, in the same temperature interval, Δh_{WW} increases with 7.4 kJ/mol ("real" benzene) and 9.9 kJ/mol (van-der-Waals-benzene). Based on these energy changes it can be concluded that the larger positive heat capacity change of hydrating van-der-Waals benzene is entirely due to an increase of the water-water enthalpy, which occurs upon "melting" the cage-like water structure surrounding the nonpolar solute. Note that under conditions of constant pressure, theoretical work[123] and experimental

data[149] have suggested that the mechanism of enthalpy absorption may indeed be localized in the solute hydration shell (whereas under conditions of constant volume, the excess partial molar energy contains a nonlocal, bulk response contribution).[123]

It is interesting to address in some detail the question to what extent reorganization of the solvent occurring in response to introducing solute-solvent electrostatic interactions is enthalpy-entropy compensating in the free energy. Because the solute molecule considered here is rigid, all the entropy change of this process is due to the rearrangement of solvent molecules. Thus, as originally proposed by Lee,[142] the total entropy change can be considered as the solvent reorganization entropy. This quantity will be referred to here as $\Delta s^{reo}(= s_{ex,S}(real) - s_{ex,S}(van\ der\ Waals))$. The solvent reorganization enthalpy for introducing electrostatic interactions is calculated as the difference between solvent reorganization enthalpies for introducing 'real' and van der Waals benzene: $\Delta h^{reo}(= \Delta h_{WW}(real) - \Delta h_{WW}(van\ der\ Waals))$. The free energy change of introducing solute-solvent electrostatic interactions contains the free energy of solvent reorganization ($\Delta h^{reo} - T\Delta s^{reo}$) in addition to the change of solute-solvent interaction energy. In case perfect enthalpy-entropy compensation occurs ($T\Delta s^{reo} = \Delta h^{reo}$), the free energy change of introducing the electrostatic interactions will only be determined by the change of the solute-solvent interaction energy. The solute-solvent energy change (Table 5.3) amounts to $-59.4 + 45.6 = -13.8$ kJ/mol. The free energy change (43A1 benzene, Table 5.2) is however smaller and amounts to $-4.8 - 1.0 = -5.8$ kJ/mol indicating that solvent reorganization contributes unfavorably to the free energy change. From Table 5.2 it can be seen that introducing benzene-water electrostatic interactions causes an entropy change $T\Delta s^{reo} = 1.3$ kJ/mol (302 K). The reorganization enthalpy (Table 5.3) is significantly larger and amounts to $\Delta h^{reo} = 9.3$ kJ/mol. This leads to a free energy of solvent reorganization of 9.3 kJ/mol - 1.3 kJ/mol = 8 kJ/mol. The free energy of solvent reorganization thus consists of an (small) increase of the entropy, which is however over-compensated by a much larger unfavorable increase of the solvent reorganization enthalpy. In this view, the presence of weak solute-solvent hydrogen bonds forces the solvent to assume strained, less stable, conformations relative to the pure solvent leading to a free energy change smaller than the energy gained by solute-solvent hydrogen bonding.

Although the conclusion mentioned above might be satisfactory from a free energy point of view, the microscopic significance of the reorganization entropy remains

difficult to interpret. An alternative way exists to arrive at the same conclusion based on a discussion of the solute-solvent entropy and the solute-solvent energy, whose microscopic significance is discussed in section 4.2. A zero change of the solute-solvent entropy is the condition for perfect enthalpy-entropy compensation of the solvent reorganization process for the introduction of charges to a nonpolar system as discussed above (i.e. $\Delta h^{reo} - T\Delta s^{reo} = T(\Delta s_{sw}(real) - \Delta s_{sw}(van\ der\ Waals))$). In section 4.2 it is shown that:[121]

$$T\Delta s_{SW} = k_B T \ln P_{ins} - [(\langle \psi^2 \rangle - \langle \psi \rangle^2)\beta/2 + \dots]_a \quad (5.7)$$

The solute-solvent entropy is determined by (1) the probability (P_{ins}) that in a system of only solvent molecules a cavity is found where the solute-solvent interaction energy (ψ) is attractive and (2) the fluctuations of the solute-solvent energy in configurations of the solute-solvent system where $\psi < 0$. The second contribution reflects the fact that fluctuations in positions and orientations of solvent molecules vicinal to the solute are biased by attractive solute-solvent interactions resulting in a reduction of configuration space and thus a reduction of the entropy. The process of introducing the electrostatic interactions leads to favorable change of Δu_{SW} of -13.8 kJ/mol while causing a compensating unfavorable change of $T\Delta s_{SW}$ of -54.6+46.6 = -8.0 kJ/mol (see Table 5.3). This observed change of $T\Delta s_{SW}$ results almost completely from changes in the second term on the right hand side of eq 5.7, not from changes in the first term: calculations of $(\langle \psi^2 \rangle - \langle \psi \rangle^2)_a\beta/2$, using the 90 ns trajectories of hydrated "real" benzene and van-der-Waals-benzene, resulted in 12.8 (± 0.2) kJ/mol for "real" benzene and 3.9 (± 0.2) kJ/mol for van-der-Waals benzene, thus contributing -8.9 kJ/mol to the change of $T\Delta s_{SW}$. Benzene-water hydrogen bonding thus introduces a stronger bias on the orientations sampled by hydration waters than the bias introduced in hydrating nonpolar van-der-Waals benzene. The formation of weak benzene-water hydrogen bonds, which energetically favors the hydration of "real" benzene, happens in competition with solvent-solvent interactions that favor water-water hydrogen bonding. In creating these energetically favorable benzene-water hydrogen bonds, the solvent is forced to sacrifice some of its own hydrogen bonds. The corresponding water configurations will have a potential energy significantly larger than configurations representative of pure water and configurations compatible with hydrating van-der-Waals benzene (i.e the solvent reorganiza-

tion enthalpy is positive). Clearly, due to its attempt to maintain hydrogen bonds, pure water will with lower probability sample configurations representative for the hydration structure of "real" benzene than for van-der-Waals benzene. The probability $e^{\Delta s_{sw}/k_B}$ to successfully insert "real" benzene in pure water configurations with an appropriately formed cavity (that satisfies benzene-water hydrogen bonding) will therefore be lower than the probability to successfully insert van-der-Waals benzene. We thus see that the reduction of the solute-solvent entropy arising by introducing electrostatic benzene-water interactions originates from less stable configurations that water has to adopt in order to donate hydrogen bonds to benzene. This automatically implies that the solvent reorganization process is non-compensating and the solvent reorganization enthalpy is positive, and that moderately polar, "real" benzene dissolves better than van-der-Waals-benzene due to energetically favorable electrostatic interactions with the solvent.

5.5 Discussion

Although here it is only shown that the hydration entropies of van-der-Waals- and realistic benzene do not significantly differ, the current results more generally state that lower aqueous solubilities of nonpolar compared to polar molecules are not due to the entropy differences but due to the lack of favorable electrostatic interactions with the solvent. This view, advocated earlier by Gallicchio et al.,[148] is supported by the Monte Carlo simulations of Stone et al.[150] based on which it can be concluded that poor aqueous solubilities correlate with poor solute-solvent electrostatic interaction.

6 A dual-scale modeling of aqueous-metal surface interactions.

Describing the aqueous-metal surface interface is a challenge from both experimental and theoretical point of view. The main possibilities and difficulties of existing research in this area will be reviewed in section 6.1. Clearly, a physically complete description of the liquid water-metal surface for atomistic simulations forms a challenge as it has to combine the fluidity and solvation properties of bulk water with the non-trivial adsorption mechanisms as shown by quantum calculations. For this, equilibrium states of water near the surface are necessary, which require costly calculations on the quantum level. Therefore, a dual-scale method is applied to combine the long sampling times attainable with atomistic MD with the adsorption behavior as acquired from quantum calculations.

The remainder of the chapter is organized as follows: section 6.2 contains a detailed description of the basic ideas used to obtain a dual scale parameterization of the water-surface interaction, section 6.3 presents the technical aspects regarding the classical force fields, and in section 6.4 the independence is shown of both the water-metal surface structure and the adsorption mechanism out of solution from the classical molecular model employed. The discussion will cover the progress of this modeling with respect to pre-existing water-metal classical model and future perspectives.

6.1 Water-surface interactions, an overview

The description of biomolecule-surface interactions under physiological conditions is incomplete without a description of water in these systems. Whereas force field based simulations of biomolecules in bulk water have a long history, water-surface interactions have been researched relatively little.

The interaction of water with surfaces is ubiquitous, of importance in a whole range of processes ranging from biological to industrial. Next to that, it is hard to study experimentally and theoretically. Research in this field is however evolving rapidly, and more and more data is becoming available. A first extensive review of water-surface interaction studies was written almost two decades ago[151]. More recently a range of overviews have been published focusing on either general surface wetting by water[152, 153] or specifically addressing aqueous-solid interfaces, where (contrary to protein-surface studies) experimental and theoretical insight is often closely related[154–156].

For the current thesis, the main interest lies at the effects of room temperature water near metal surfaces, a very common case at first sight, but a very special case from an experimental/theoretical point of view as most studies are feasible only for extreme conditions (e.g. Ultra High Vacuum, UHV). The specific approach that had to be developed to atomistically model aqueous-surface interactions will be discussed in great detail in chapter 6.

6.1.1 Experimental water-metal

From the experimental point of view, at low temperatures, it has been found that deposition of water on metal surfaces forms different structures such as one dimensional chains, bilayers, islands and clusters before bulk ice is formed.[157–159].

More recently, it has become possible to study the aqueous-metal interface. For example, by Fourier transform infrared spectroscopy it was found that water reversibly adsorbs on an MgO surface, and forms a liquid-like layer even for temperatures down to -10°C [160]. A different technique was applied to a metal oxide, Al_2O_3 , where the electron density of the adsorbed water layers could be determined by high-resolution specular X-ray reflectivity[161]. As this metal oxide has oxygen atoms at the outer layer, water specifically coordinated to these sites. This resulted

in the formation of a few coordinated layers of water next to the surface, but this was found to quickly fall off due to the disorder in fluid water, a process which therefore might well be general for all aqueous-surface adsorbed systems. Similarly, it was shown by a high-energy x-ray transmission-reflection method that ice next to a SiO_2 solid will form several intermediate layers, with a thickness of about 1.7 nm, consisting of water molecules in a liquid-like state, a process known as interfacial melting[162]. Recently, the use of scanning tunneling microscopy to study the solid-aqueous interface has been reported and it was possible to study a tunneling behaviour in a several nm thick slit of water between two Au(111) surfaces[163].

6.1.2 Quantum calculations of water-metal interactions

DFT calculations of water on platinum (111) by Meng *et al.* [164] and of water on $3d$, $4d$ and $5d$ (111) transition metal surfaces by Michaelides *et al.*[165] (for the case of nickel these results were reproduced as well, see ref. [44]) show the formation of a bilayer ice-like structure where the adsorption energy compared to a free water molecule does not come from the water-metal interaction but from the water-water interaction. These calculations present an evident limitation. In fact the cell size used is often so small that the effect of the periodic images hinder the system from melting, thus the adsorption properties are suitable to describe an ice/metal interface but not liquid water where, contrary to the case of ice, the fluctuations of the local structure play a crucial role. Obviously to perform calculations with larger clusters, and eventually a more realistic liquid/metal system, would have increased the computational cost massively but, at this stage, would not have offered a clearer and more accurate modeling procedure for the single molecule/surface interaction.

6.1.3 Atomistic simulation of water near surfaces

A recent review of Spohr on computer simulations of aqueous double layers near surfaces pointed out that most simulation studies showed the same layered structure, consisting of three distinctly observable layers with increased water density. After the third and lowest density layer, the water density profile becomes bulk-like.[166]. Another observation was the absence of a clearly visible peak splitting in the first layer in both MD and quantum calculations, even though intuitively, a 'staggered conformation would be expected, to retain hydrogen bonds at the layer closest at

the surface.

In atomistic simulations, several methods to model the strength of the water-surface interaction energy are commonly used. Initially only model surfaces were considered using an estimated guess for the interaction strength of interest[167], but also experimental data have been used[168], as well as ab-initio[44] or cluster calculations[169]. In some cases, a detailed site-dependent potential has been applied[169], but it has been shown that this delivered the same results as less detailed potentials that only contained a distance-dependent part[166].

An image-charge potential could be added to describe the metallic nature of the surface, imaging all charges next to it at a virtual location in the surface[170]. It has however been shown that due to the sheer quantity of partial (opposite) charges present in water, the image charge interactions effectually cancel each other out and it will be more efficient to not take this into account, without losing essential details of the simulation[171].

As noted before, the features of water-surface layers seem very robust against small changes or the addition of more detail in the simulation model. Indeed, a comparison of a polarizable water model near a Mg(111) atomistic crystal surface with a non-polarizable water model near a corrugated potential surface (shaped equal to a Ni(110) surface), showed no significant changes in the interfacial structure due to either corrugation of the surface or polarizability of the water model[172].

6.2 Quantum based modeling of liquid water-metal interaction: Basic idea

The liquid water model developed here should be able to describe both adsorption and solvation properties in combination with solutes, and therefore the focus should deviate from studies aimed at describing catalysis or gas/ice adsorption at a surface. Liquid water has completely different characteristics than ordered ice; as a molecular liquid, its main feature is its fluctuating hydrogen-bonded structure. While the general many body aspect of the liquid can be properly taken into account by several readily available classical force fields, the intrinsic quantum nature of the water-surface interaction must be described by first principle techniques and then inserted into the classical framework. The aim of the quantum-classical modeling

presented here is to design a suitable interaction potential between the molecule and the surface, based on quantum calculations where the local liquid structure and its influence onto the interaction with the surface are properly accounted for.

To this aim, paradoxically, the available quantum studies of isolated water molecules adsorbed on different transition metals are in fact applicable.

For this cause it will be possible to refer to quantum studies of the metal adsorption of isolated water molecules. The interpretation and use of very recent state-of-art Density Functional calculations treating layers with up to 128 ice-like structured molecules on a metal surface[165, 173, 174] however turned out to be rather difficult. The reason is twofold. First, it is necessary to assign the resulting interaction energies to the correct sites of the system. In the case of water, any contribution of the hydrogen bonding network to the total energy of the system should be separated from the surface-induced interaction energy. As pointed out later in this section, this limits quantum calculations to configurations not larger than a water trimer. Thus while large systems can certainly be used a posteriori to analyze the validity of the modeling, they certainly cannot help much in defining a physically well founded, yet simple, molecule-surface interaction. The second reason is that the structures studied resemble in most of the cases ordered layering structures closer to ice, and are therefore far from describing the fluctuating character of liquid water.

The current modeling idea is based on a well accepted statistical property of liquid water, i.e. due to its hydrogen bond network it is locally and instantaneously tetrahedral, also at interfaces[175]. In this case, local water conformations at the surface must consist of full or half tetrahedrons (see figure 6.1) due the confinement of the surface. Among possible arrangements of the like shown in figure 6.1, some are not allowed according to data available from small quantum calculations, i.e. those conformation displaying hydrogen down-like structures. This creates a first screening of possible conformations that must be reproduced by the modeling. The second important point is that, since many relevant configurations need to be explored, many calculations need to be carried out. First principles quantum calculations are known for being computationally very demanding, especially when metal surfaces need to be studied. A solution to this problem comes from the observation that the tetrahedral structure of figure 6.1 can be well described by different combinations of substructures consisting of monomers, dimers and trimers in different conformations. As it is shown in figure 6.1, a so-called "first layer approximation" is made, i.e. only

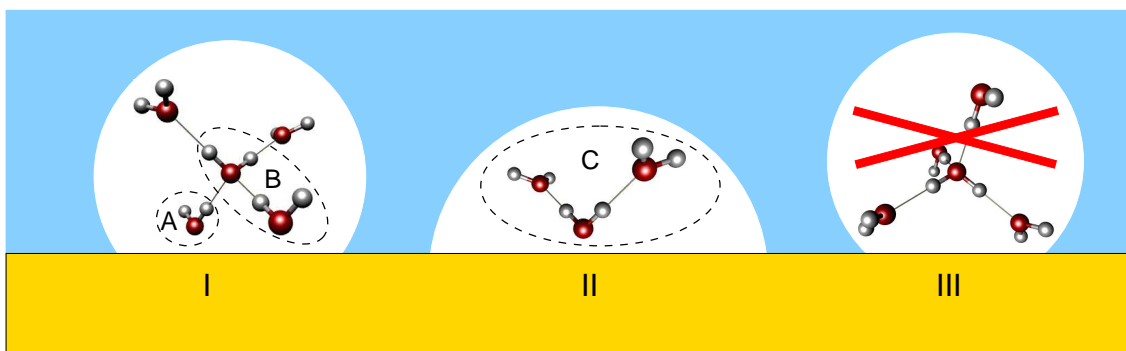


Figure 6.1: A graphical aid showing the rationale behind the set-up of the quantum calculations. The water configurations studied by quantum mechanic DFT calculations were chosen by considering all relevant ways for liquid water to interact with the surface. Some examples are schematically shown here. In white circles: tetrahedral water substructures at the water-metal surface. I: A possible tetrahedral water structure with two waters interacting with the metal surface; II: A trimer, the upper part of a tetrahedral water structure, with one water interacting with the metal surface and two additional hydrogen bonded water molecules; III: One of many configurations that could be immediately discarded as no electronic hydrogen-metal interaction exists. The isolated water tetrahedron is an unstable structure, both in vacuo and metal-bound. Therefore, quantum calculations are performed with smaller subunits. We can represent the tetrahedral structures shown in I and II by using all relevant water-structures consisting of: A: a water monomer, representing one of the symmetry axes of the tetrahedron; B: a water dimer, representing the metal-bound water with its first hydrogen bonded neighbor, the tetrahedral center; C: a water trimer, directly representing structure II. The complete overview on all calculated conformations is given in ref. [44].

the molecule close to the surface and those directly bonded to that participate in the determination of the adsorption strength. This turned to be an important point for three reasons:

(1) First, a straightforward postulation, as recently discussed by Marx[175], clarifies that the structure of water confined by an interface is energetically steered to conformations where as many tetrahedral hydrogen bonds towards the bulk are kept as possible, resulting in partial tetrahedral interfacial or bulk-connected water structures in the interface region.

(2) A second point is that in order to calculate the adsorption energy, i.e. the strength of the interaction of a water structure at the surface, such a structure must

be stable. Adsorption energy E_{ads} is defined here as:

$$E_{ads} = E_{surf+N_{wat}} - E_{surf} - E_{N_{wat}} \quad (6.1)$$

where $E_{surf+N_{wat}}$ is the total energy of the N_{wat} water molecules and the metal surface, E_{surf} is the energy of the surface and $E_{N_{wat}}$ is the the energy of the isolated cluster of N_{wat} water molecules; all such energies are obtained from a geometry optimization process where all the degrees of freedom are allowed to relax. In quantum calculations this requires box sizes that are large enough to prevent hydrogen bonding with neighbors in vicinal boxes, thus preventing crystallization into an ice-like phase. The calculation of E_{ads} up to a trimer does not give problems, but already for a tetramer no reasonable isolated, energy-minimized, cluster-like structure can be found, and one ends up with chain-like molecular structures or pairs of dimers forming and breaking. This makes energetic comparison between the $E_{surf+N_{wat}}$ and the $E_{N_{wat}}$ states impossible for configurations of tetramer and up. The idea of using tetrahedral substructures is in accordance with point (1) for describing bulk liquid water at a metal surface, because if partial tetrahedral structures are present at the surface, then they must locally consist of monomers, dimers or trimers. As clarified in figure 6.1, even for a fully tetrahedral structure this choice of substructures assures a proper conformational description. As this approach manages to describe all possible surface conformations with water structures not bigger than a trimer, this trivially solves also the second problem.

In this line of thought, a Density Functional, Car-Parrinello, study of the adsorption of water monomers dimers and trimers on all high symmetry sites of a nickel (111) surface has been carried out by Sebastiani and Delle Site[44].

(3) The third point affects the integration into the classical model. Water adsorption energies determined by quantum calculations normally include the hydrogen-bonding energies, for example the data of Sebastiani and Delle Site, calculated in that way, showed good agreement with other studies[44]. For the current modeling, however, a recalculation of the water-metal adsorption energies was performed using eq. 6.1 and it was found that all water molecules in the first bilayer (i.e. all water molecules as shown in configuration A, B and C in figure 6.1) have similar optimal interaction energies per water molecule, ca. 24 kJ/mol (0.25 eV). Therefore, a single atom-surface interaction potential can be used in the MD implementation

to describe the metal binding of all molecules in this bilayer. It should be noted that further structuring due to hydrogen bonds is accounted for by the standard force field water-water potentials. This implementation should result in a consistent configuration as used in the initial quantum calculations, which can be verified by analyzing the hydrogen bond network at the surface. Furthermore, it has been found in other quantum calculations and MD modeling studies that the first bilayer, even in the ice-state, results in only one indistinguishable peak in the density profile[166]. This and the fact that in liquid water thermal fluctuations play a larger role than in low-temperature ice, justified the choice for this simple binding potential scheme and redeemed the use of any complex (multibody) potentials to reproduce all details of the bilayer structure near the surface. To sum up, the combined approach of considering electronic binding effects only for the first metal-bounded water and its direct neighbors, and treating both with a single potential in molecular dynamics forms the "first bilayer approximation" presented here.

6.3 Simulation details

The MD simulations applied the Gromacs 3.3 code[93,94]. The Gromacs code was adapted to take into account two, structureless, flat metal surfaces at opposite sides of the simulation box, having the z-axis as the surface normal. These surfaces were modeled by adding benzene-wall and water-wall interactions as external potentials (see below), resulting in a slit of water with the same average density as a bulk system. Within the accuracy we can achieve here, it is sufficient to employ a structureless surface. Even in the case of strong interaction (benzene Ni, and water Ni) the energy barriers to move from one minimum to the next are small compared to the overall interaction energy.

The box contained 3000 water molecules yielding a starting geometry with a slab width of 7.6 nm, twice as big as the x and y box dimensions to ensure a bulk liquid state in the middle of the box. The Gromacs pseudo-2d PME algorithm for electrostatics was used and a semi-isotropic NPT ensemble was applied by coupling x and y pressure components with Parrinello-Rahman pressure coupling[103,176], keeping the z-distance constant, and coupling temperature with the Nose-Hoover thermostat[99,177]. A time step 0.002 ps was used and every fifth time step a pair list update was performed.

As classical water models the SPC[141], SPC/E[178], TIP3P[179], TIP4P[179] and TIP5P[180] models were studied. All water models used were rigid, meaning that besides the constraining of bond distances, no angular vibrations were taken into account. The force field data and spatial configurations are given in table 6.1 and figure 6.2.

Water-solute interactions were taken from the GROMOS-specified pair interactions[88, 98] for SPC and SPC/E, or by using geometric mixing rules, see equations 3.8 and 3.9, describing interaction parameters with the TIP models.

6.4 The modeling idea applied to different water force fields

Figure 6.3 shows the normalized water densities near a Au(111) and a Ni(111) surface for the SPC, SPC/E, TIP3P, TIP4P, and TIP5P water models. Note that the hydrogen density (dashed line) shows most structure because hydrogen atoms belonging to water molecules in the first adsorbed layer either correspond to OH bonds aligning the surface or OH bonds hydrogen bonded to water in the second adsorbed layer. Hydrogen atoms belonging to water molecules in the second adsorption layer, in turn, donate hydrogen bonds to waters in the first layer and waters in the bulk. The large first oxygen peak is of equal magnitude for all water models. At this distance, the water density is determined by the water-metal interaction (see table 6.2 for the quantum data and table 6.3 for the atomistic modeling parameters). The first peak of the water hydrogen is located at 2.4 \AA as well and has equal peak height for all models except TIP5P. The 3D tetrahedral configuration of TIP5P causes a weaker alignment of water OH bonds parallel to the surface. The water oxygen density in the second peak is identical for the SPC, SPC/E, and TIP5P models while TIP3P and TIP4P show a slightly lower and higher density, respectively.

The hydrogen density in the second peak (first layer water donating H-bonds to second layer waters) is equal for SPC, SPC/E, and TIP3P while being larger for TIP4P and TIP5P. The hydrogen densities in the third (second layer water donating H-bonds to first layer water) and fourth (second layer water donating H-bonds to the bulk) peaks are of equal magnitude for all water models. Generally, the surface-water density correlations are identical for all classical water models, except for small

	SPC[141]	SPC/E[178]	TIP3P[179]	TIP4P[179]	TIP5P[180]
$q_O(e)$	-0.82	-0.8476	-0.834		
$q_H(e)$	+0.41	+0.4238	+0.417	+0.520	+0.241
$q_V(e)$				-1.04	-0.241
$\sigma_O(nm)$	0.31656	0.31656	0.31506	0.31536	0.312
$\epsilon_O(kJ/mol)$	0.65017	0.65017	0.63639	0.64895	0.66989
$r_{OH}(nm)$	0.100	0.100	0.09572	0.09572	0.09572
$r_{OV}(nm)$				0.0150	0.070
θ_{HOH} (deg)	109.47	109.47	104.52	104.52	104.52
θ_{VOV} (deg)					109.47

Table 6.1: Definition of the rigid water force fields used. TIP4P contains one virtual site, V , located inbetween the two hydrogens. TIP5P contains two virtual sites V , positioned like lone pairs around the oxygen atom, creating a tetrahedral configuration in combination with the hydrogen atoms. The spatial configurations are shown in figure 6.2

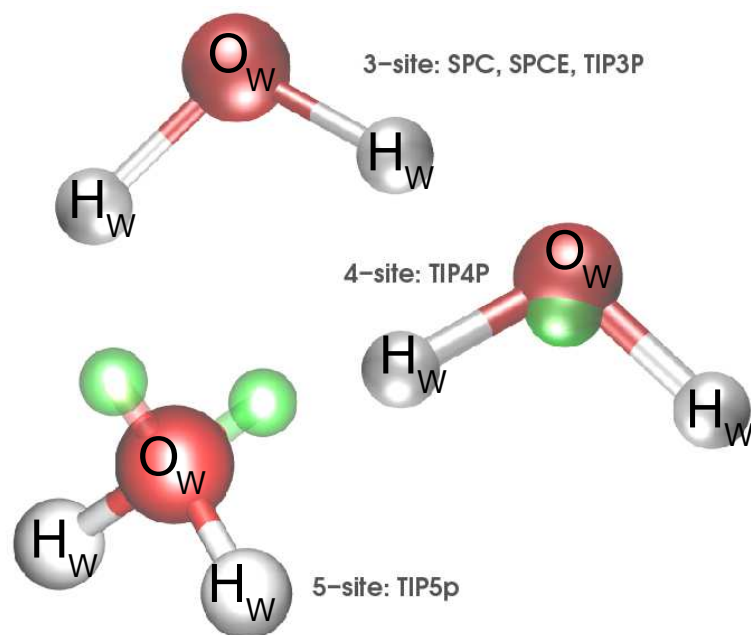


Figure 6.2: Spatial configuration of the rigid water models used. The exact force fields parameters are listed in table 6.1

Interaction	E_{ads} (kJ/mol)	d_{opt} (nm)
H ₂ O-Au	9.65	0.31
H ₂ O-Ni	24.1	0.24

Table 6.2: Maximal first principles DFT interaction energies (E_{ads}) and optimum water-surface distances (d_{opt}).[42]

Interaction	Potential Type	ϵ (kJ/mol)	σ (nm)
Water-Au(111)			
O _w -Au	attractive 10-4	2.56	0.31
H _w -Au	no interaction		
Water-Ni(111)			
O _w -Ni	attractive 10-4	6.40	0.24
H _w -Ni	no interaction		

Table 6.3: Modeling of water-surface potential parameters. See figure 6.2 for the corresponding structures and equation 3.12 for the potential function used.

differences in peak heights.

The difference between the water adsorption on the Ni(111) and the Au(111) surfaces is mainly a question of adsorption energy, which can be seen by the lowered peaks for Au(111) in figure 6.3, whereas the other characteristics of the adsorption profiles near Au(111) and Ni(111) are the same. The density profile at the surface can not be used for a quantitative comparison, however, and therefore the surface area per water molecule was compared for both surfaces (table 6.4). Even though there is a factor difference in peak heights for the first peak adjacent to the surface of Au(111) and Ni(111), the factual difference in area per molecule is only about 10%.

To verify the consistency between the bulk-like hydrogen-bonding network at the surface that was assumed in the initial quantum modeling, a hydrogen bonding profile for water near the surface was calculated. The hydrogen bonding profile for Ni(111) is shown in figure 6.4. Near the surface the amount of hydrogen bonds per molecule oscillate slightly, between 3.2 and 3.6 hydrogen bonds per molecule, but over the whole range the amount of hydrogen bonding is close to the bulk value of

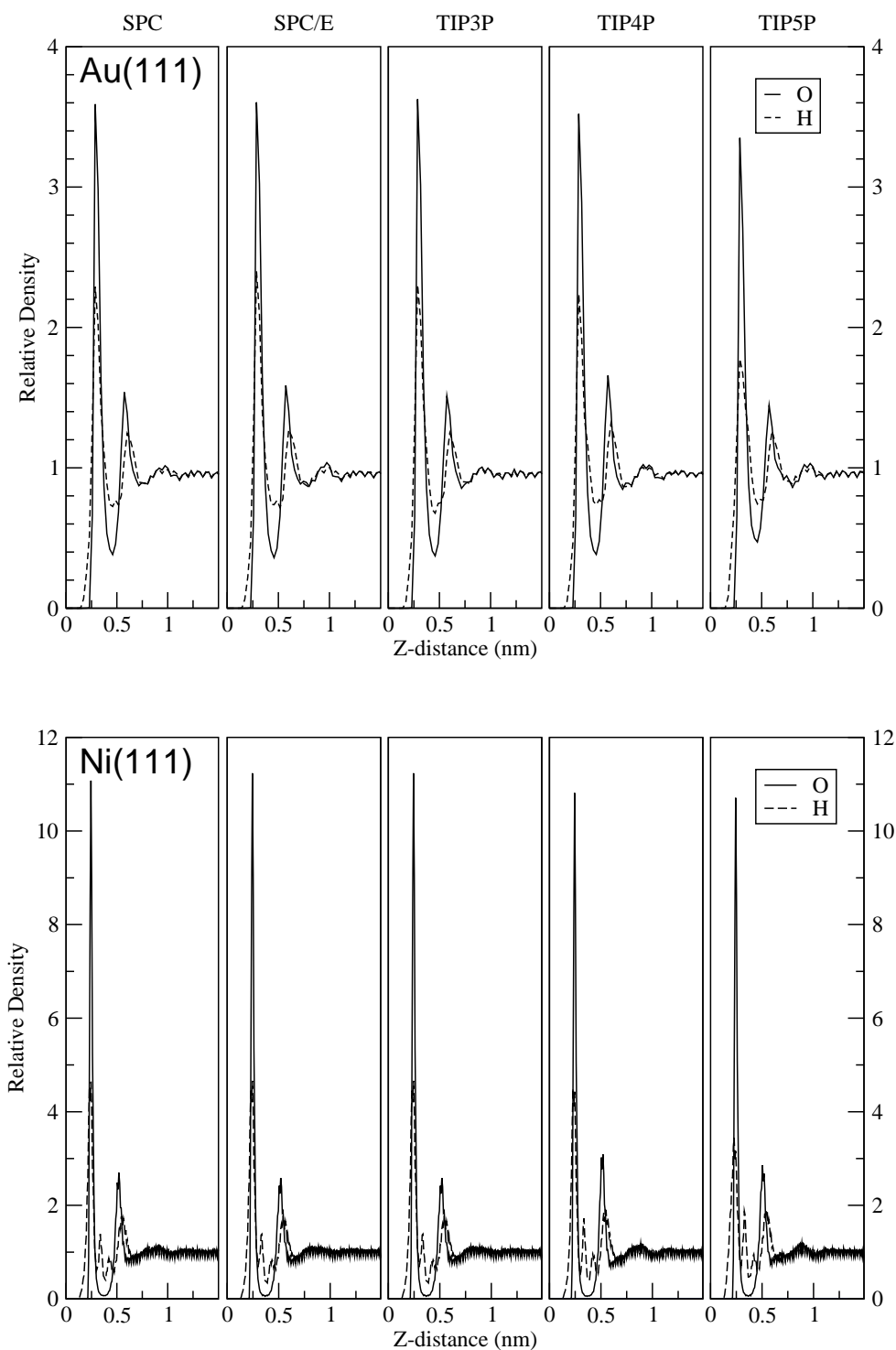


Figure 6.3: Normalized water densities near an Au(111) (top) and a Ni(111) surface (bottom). Water models used were: SPC, SPC/E, TIP3P, TIP4P, and TIP5P. Solid line: water oxygen. Dashed line: water hydrogen.

Water model	Au(111)	Ni(111)
	Area per molecule (\AA^2)	
SPC	9.52	8.64
SPC/E	9.27	8.37
TIP3P	9.46	8.61
TIP4P	9.61	8.79
TIP5P	9.86	8.92

Table 6.4: Surface area per water molecule in the first adsorbed layer, as determined by the integral of the oxygen density of the first peak adjacent to the surface in figure 6.3.

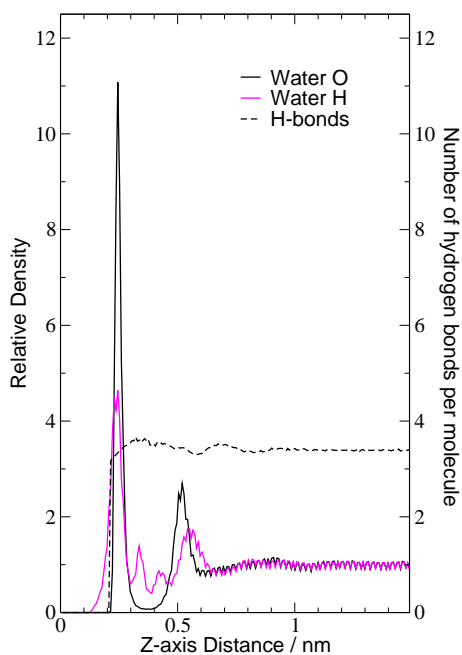


Figure 6.4: Hydrogen bonding profile for SPC near the Ni(111) surface. Hydrogen bonds were calculated for a hydrogen bond donor by counting the amount of surrounding water molecules whose oxygen atom is closer than 0.35 nm and for which one of its hydrogen atoms make an $O_{\text{donor}} - O_{\text{acceptor}} - H_{\text{acceptor}}$ angle that is between -60 and $+60$ degrees. For each water molecule in each snapshot it was calculated how many times it was involved in a hydrogen bond as either a donor or an acceptor. Solid black line: water oxygen. Solid grey line: water hydrogen. Dashed black line: hydrogen bonds per water molecule.

3.4 hydrogen bonds per water molecule. This supports the assumption used in the quantum calculations that a bulk-like hydrogen bonding structure exists near the surface.

In chapter 7, it will be shown that the potential of mean force (PMF) for displacing a solute (benzene) along the z-coordinate from bulk liquid water to the surface does not only give an idea about adsorption/desorption free energies and intermediate free-energy barriers, but also help in understanding the critical steps in the adsorption/desorption mechanism[42]. Any crucial effect that the choice of water model would have on the adsorption mechanism would therefore be visible in the corresponding PMF. However, as will be shown there, the PMFs of all models are relatively similar, the shape of all graphs have the same characteristics and differences appear only in the relative height of the barrier and therefore as well the free energy difference between bulk and the distances close to the wall.

6.5 Discussion

Even though most quantum calculations of water-metal adsorption are considering only the ice state[165, 173, 174], and most classical simulation studies of these systems are based on simplified model systems[181, 182] or semi-empirical models[168], a combination of quantum calculations and molecular dynamics simulations to model the water-metal interactions seems logical[183, 184]. However, as in all multiscale approaches, it is a non-trivial task to end up with a model where the connection of the both scales is as simple and as physically sound as possible.

The work presented here demonstrates the first elaborated approach that can deliver a general quantum-statistical model of the water-metal surface to be used in molecular dynamics simulations. Water conformations near the surface, water-metal interactions and water-water hydrogen bonding interactions are distinctively taken into account on both the quantum and the classical force field level, to ensure that every interaction is physically justified, and no double counting of interactions occurs. The only limiting factor to the current modeling is that of possible contributions that are not covered in the quantum calculations. One of those is the dispersion energy contribution to water adsorption, but a recent study by Feibelman[113] comparing theoretical and experimental work already indicated that this will be only a secondary factor compared to the strong interactions that determine

the mechanisms of interest for these systems (Feibelman mentions a ± 4 kJ/mol contribution from dispersion interaction on a total interaction of ± 19 kJ/mol). More direct quantum calculations of dispersion near a surface, as the calculation performed for phenol near a graphite surface by Chakarova-Käc et al.[185], will be needed to obtain quantitative data on the dispersion interaction near a surface. Another factor not taken into account yet could be water dissociation at the surface, however quantum calculations have shown that the dissociation in a water bilayer is not the most stable conformation for almost all of the transition metals tested (Ag,Pt,Cu,Pd,Ni,Rh), excluding one (Ru)[165].

This water-metal surface atomistic model and the general modeling procedure used form the basis for a broad range of modeling problems concerning liquid-metal surfaces in aqueous environments, as will be treated in the following chapters.

7 Modeling the interaction of hydrated Benzene and Phenol with metal surfaces.

Building further on the water-surface modeling introduced in chapter 6, now a solute is introduced in the system, as a step towards the multi-scale modeling approach of amino acids. Benzene is chosen as a model solute, as the hydration of this molecule was already a matter of extensive simulations (see chapter 5) and the metal adsorption of this molecule has been an object of study in previous research papers in the theory group[6–8]. As a model solute it resembles the amino acid side chain of phenylalanine, and the similar phenol solute relates to the side chain of tyrosine.

7.1 Benzene and phenol as side chain analogs

In this chapter, the adsorption of fully hydrated benzene on Au(111) and benzene and phenol molecules onto Ni(111) surfaces will be studied. A bottom-up multiscale strategy is applied in which the interaction energies of isolated molecular fragments (i.e. amino acid residues) with transition metal surfaces are determined by first principles Density Functional Theory (DFT) quantum calculations. These energies are then used to parameterize potential energy functions that are implemented in a fully classical, atomistically detailed molecular dynamics (MD) simulation of the system. In these MD simulations the remaining solute-water and water-water interactions are described by empirical force fields, which are parameterized based on hydration free energies of similar molecular fragments (e.g. the GROMOS force field).[98] Despite the assumptions made in these modeling steps, this approach allows to describe surface phenomena at a level of detail that cannot be reached by

any alternative computational method available today.

It will be demonstrated that benzene adsorption on Ni(111) and (to a lesser extent) Au(111) is governed by solvation effects related to near-surface ordering of water molecules. Despite the fact that the benzene-Ni(111) interaction energy is approximately 40 kJ/mol more favorable compared to the benzene-water interactions energy of a fully hydrated benzene, it is balanced by an unfavorable entropy change of transferring a benzene molecule from the aqueous bulk solution into contact with the Ni(111) surface. Moreover, it will be shown that kinetic barriers exist causing benzene adsorption on and desorption from the Ni(111) surface being a rare event. Sampling of this event will therefore require accelerated sampling techniques (introduced in chapter 4), of which in this case constrained dynamics with force averaging and umbrella sampling will be used.

7.2 Methodology

7.2.1 First principles DFT-based modeling

First principles DFT calculations for benzene/nickel(gold)[7, 67, 186] and water/nickel(gold)[44] were carried out using the Free Energy Molecular Dynamics (FEMD) approach of Alavi[187, 188], implemented into the Car-Parrinello Molecular Dynamics package (CPMD)[189]. For benzene on Ni the results are also in good agreement to recent experimental studies[190, 191].

The adsorption energy for benzene/metal and water/metal was calculated as:

$$E_{ads} = E_{surf+mol} - E_{surf} - E_{mol} \quad (7.1)$$

, where all energies correspond to relaxed geometries. To account for the possible molecular conformations at the surface, the distance dependence and the angular dependence of the adsorption energy were determined, where the angle corresponds to the inclination of the carbon ring with respect to the surface. For gold, due to the relatively weak interaction, this study turned out not to be necessary.

The distance and angular dependence of the benzene-nickel interaction is shown in figure 7.1, combined with the purely atom-surface based MD potential discussed below.

Interaction	E_{ads} (kJ/mol)	d_{opt} (nm)
Benzene-Au	4.82	0.33
Benzene-Ni	105.2	0.20

Table 7.1: Maximal first principles DFT interaction energies (E_{ads}) and optimum solute-surface distances (d_{opt}).

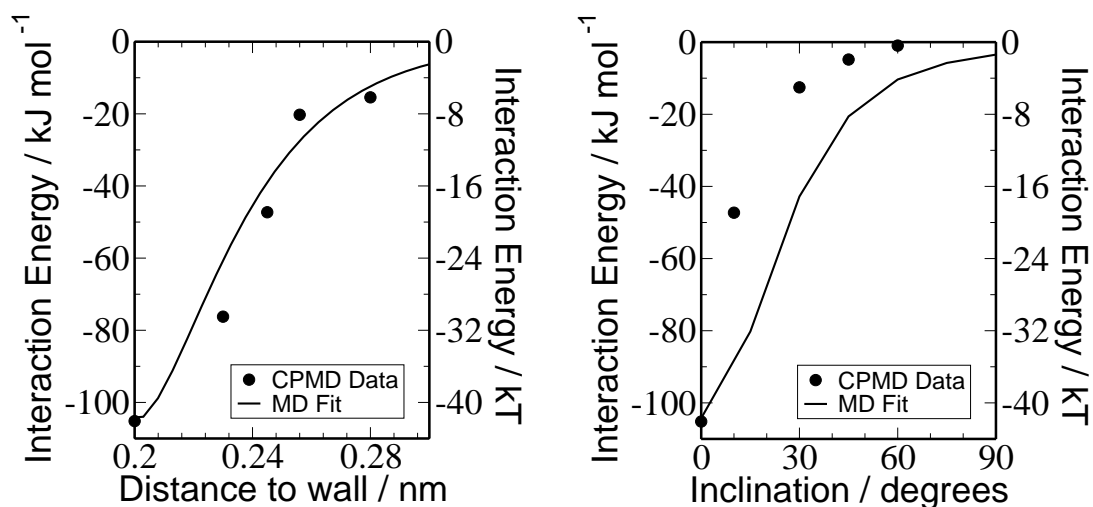


Figure 7.1: First principles DFT values (dots) and corresponding MD potential (solid line) for the distance and inclination dependence of the benzene-nickel interaction energy. Distances are based on parallel orientation of the benzene and the metal surface. The tilt angle is defined as the angle between the normal of the metal surface and the normal on the plane of the benzene ring, upon rotation around one of the benzene hydrogens, starting from the optimal distance (0.2 nm) at zero inclination. Note that the MD potential does not contain an inclination-dependent part in the potential, but instead the inclination dependence follows from the combination of atom-site potentials. In that way, an inclined configuration at the surface will contain several atoms that are further away from the surface than the optimal distance, thereby lowering the total interaction energy of the molecule.

7.2.2 Molecular dynamics implementation

The rigid, non-polarizable, simple point charge (SPC) water model has been used in describing water-water and solute-water interactions.[141] The effect of electronic polarizability has been shown to have only little influence on the structure of water

near a metal surface and can therefore be neglected.[172]

7.2.3 Benzene modeling

Benzene-water non-bonded interaction parameters as well as the benzene bonded parameters have been extensively investigated in chapter 5. Molecular dynamics simulations were performed similarly to the simulations of chapter 6, however now using the Gromacs 3.2.1 MD simulation package.[92,93], and a system containing 1500 water molecules at constant volume and temperature ($T=298\text{K}$), using a Berendsen weak-coupling thermostat,[102] with relaxation time $\tau = 0.1$ ps. Other differences include a twin-range cutoff scheme with 0.8 nm and 1.4 nm cutoff radii that was applied here for solute-water and water-water electrostatic interactions[88]. The non-bonded interactions in the range between these radii were updated every fifth time step. In the current simulations, bond lengths were constrained by the SHAKE algorithm.[95]

Due to the different nature and strength of the benzene-metal interactions for gold and nickel, different benzene-surface pair potentials (equations 3.10-3.12) were used to model the first principles DFT data (table 7.1) in atomistic site-surface potentials (table 7.2 and figure 7.2). An attractive 10-4 potential was chosen for the interaction between benzene C-atoms and gold, between water O-atoms and gold, and between water O-atoms and nickel (as described in table ?? in chapter 6). This potential is appropriate to describe these relatively weak non-covalent interactions. They take the form of equation 3.10 with z_{cutoff} is 1.4 nm in the current simulations. The stronger, more covalent-like, benzene-nickel interaction was modeled by a Morse potential on the benzene C-atoms (U_{BC-Ni} , equation 3.12), combined with a repulsive 10-4 potential on the H-atoms (U_{BH-Ni} , equation 3.11), to fine-tune the angular and distance dependence to the one found in the quantum calculation (figure 7.1). The repulsive H-surface potential corresponds to the fact that the H atoms are slightly repelled by the surface, as indicated by the quantum calculations. The Morse parameter was optimized to get the best compromise between corresponding distances and angles ($a = 35 \text{ nm}^{-1}$). The resulting curve (figure 7.1) however somewhat underestimates the tilt-dependency and therefore will overestimate the binding energy for tilted conformations at the surface. A larger Morse potential leading to a steeper decrease of the Morse potential would have improved

the inclination-dependent correspondence to quantum data of the overall potential. However, it would have impaired the distance-dependent correspondence to quantum data, resulting in much weaker binding at all intermediate distances between 0.2 and 0.3 nm. The final form, shown in figure 7.1, was adopted as the best trade-off for the distance and the inclination correspondence of MD data with quantum data. In calculating the free energy of binding the error made by such an approximation will be of the order of $k_B T$, which results from a larger entropy of the tilt angle vibration. Taking the entropy difference between a classical freely rotating benzene and a benzene fixed to the flat configuration with respect to the wall (at constant energy) leads to an entropy difference in the order of $k_B T$ (about 2.5 kJ/mol) for the two degrees of freedom, disregarding any interaction effect. Thus the over estimate of the benzene entropy close to the surface is within the order of the intrinsic error bars of the quantum calculations. As we will see later, this is negligible compared to the other energies, which dominate the properties.

7.2.4 Phenol modeling

Phenol on Ni (111) has a maximum surface interaction energy that is almost 15% lower than the maximum surface interaction energy of benzene (see table 7.1 and ref. [72]). Its conformation at the surface has the hydroxyl oxygen lifted by 0.5 Å, which has some effect on the adjacent carbon atom as well. The hydroxyl hydrogen is pointing down, ending up at the same height as the phenylic carbon atoms (see fig 7.3). As a basis of the atomistic modeling of phenol, the benzene model as discussed in chapter 5 is chosen. To describe the interaction with the surface, Morse potentials are used on the aromatic ring carbon atoms (C_r) (equation 3.12), and repulsive 10-4 potentials (equation 3.11) are used on the ring hydrogens (H_r). The model parameters are given in table 7.2, atom types are defined in fig. 7.2. For all attractive atom types a z_{cutoff} of 1.4 nm was used.

Besides the substitution of one of the hydrogens by a hydroxyl group, only a reduction of interaction energy was needed which was done by scaling down all carbon-wall interactions. The carbon next to the hydroxyl group (C_p) was scaled down to half the interaction energy of the other carbon atoms (C_r). A weak 10-4 repulsion was put on the hydroxyl oxygen (O_p), an attractive 10-4 potential (see equation 3.10) with similar strength was put on the hydroxyl hydrogen (H_p).

Interaction	Potential Type	$\epsilon(kJ/mol)$	$\sigma(nm)$
Benzene-Au(111)			
C _r -Au	attractive 10-4	0.21	0.33
Benzene-Ni(111)			
C _r -Ni	Morse (a=35nm ⁻¹)	17.5	0.20
H _r -Ni	repulsive 10-4	4.27	0.20
Phenol-Ni(111)			
C _r -Ni	Morse (a=35nm ⁻¹)	15.8	0.20
H _r -Ni	repulsive 10-4	4.27	0.20
C _p -Ni	Morse (a=35nm ⁻¹)	7.96	0.20
O _p -Ni	repulsive 10-4	1.00	0.25
H _p -Ni	attractive 10-4	0.70	0.22

Table 7.2: Modeling of molecule-surface potential parameters. See figure 7.2 for the corresponding structures and equations 3.12-3.13 for the potential functions used.

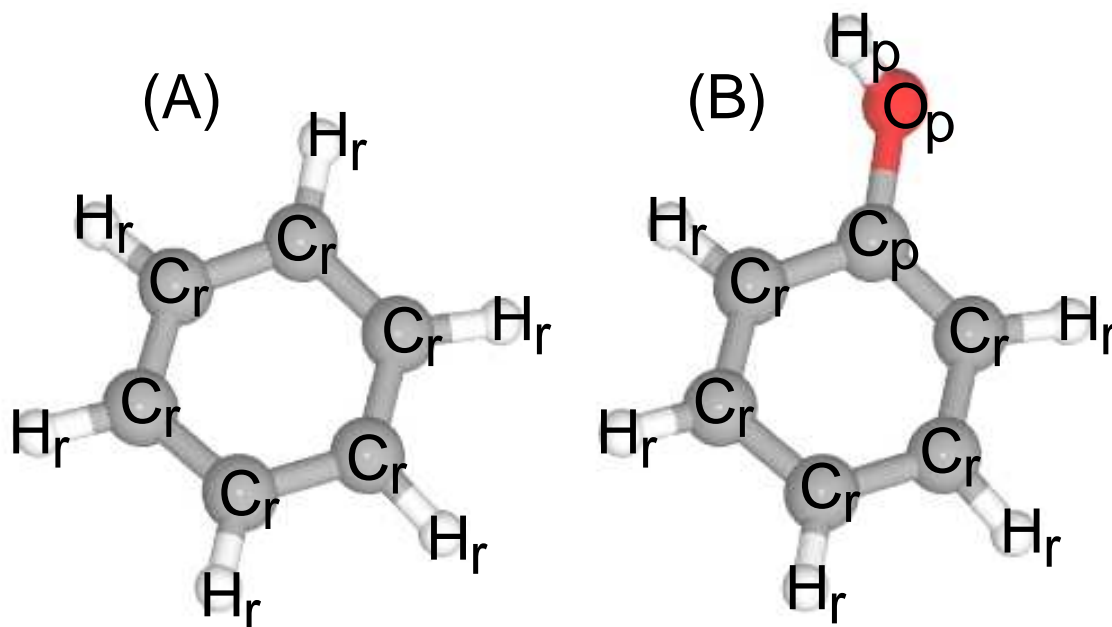


Figure 7.2: Shown are: (A) Benzene (B) Phenol. Atom names refer to atom names in table 7.2. Note that the water hydrogens H_w do not have any interaction with the surface.

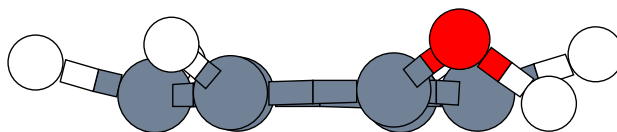


Figure 7.3: Minimal energy phenol configuration when adsorbed on Ni(111), as found by quantum calculations[67]. The green area represents the location of the nickel surface.

Configurations of phenol with the OH group pointing away from the surface did not need to be taken into account in QM calculations, as the perturbing effect of the OH group will be lower the further it is away from the surface, and any interaction would be similar to the inclination dependence of benzene, which was modeled previously[42]. The configuration where phenol is pointing with the OH group towards the Ni(111) surface is currently being investigated by quantum calculations. Preliminary calculations indicated a weak interaction, not significantly higher than the error margin of the quantum calculation. Currently, no special modeling considerations were made to account for this orientation in the classical atomistic simulation.

7.3 Results and Discussion

MD simulations were performed for 2500 water molecules (oxygen and hydrogen atoms) with one single benzene molecule (carbons and hydrogen atoms). To show that the simulations are well equilibrated, typically density profiles are shown for the full system. The overall size of the simulation box is then 5nm, with walls at $z = 0nm$ and $z = 5nm$ and periodic boundary condition in the x and y direction.

7.3.1 Density profiles

Figure 7.4 shows the normalized densities for the 2500 water (oxygen and hydrogen atoms) molecules and the single benzene (carbon and hydrogen atoms) molecule along the z -coordinate normal to the surface. The density profile for water is highly symmetric, while for the benzene molecule the averaging is somewhat poorer, due to the much smaller statistics. Both, the Ni(111) and Au(111), surfaces are significantly wetted and two pronounced water layers are observed. The wetting profile of water at the Au(111) and Ni(111) surface has been discussed extensively in chapter 6. The peak heights depend clearly on the metal-water interaction strength: water adsorption on Au(111) (see inset) is significantly weaker than on Ni(111). Analyses of the in-plane and out-of-plane diffusion showed that in both cases the adsorbed water remained liquid-like. In the middle of the cell, the water concentration of the Ni(111) forms a slight 'dip', due to the dense layering at the surface after starting from the correct bulk density. This is due to the fact that a constant volume was used for the simulation shown here, and can be prevented by performing constant pressure simulations as the ones performed in chapters 6 and 8. However, for the surface structure the results are indistinguishable.

Due to the presence of only a single benzene molecule in the system, the benzene density profile is not fully symmetric, however, the sampling is sufficient to extract its most important features. The closer to the surface, the less probable it is to have a benzene present. Most likely, steric (entropy) effects inhibit the benzene molecule to penetrate the adsorbed water layers, preventing it from approaching the surface close enough ($z \leq 0.3nm$) to experience the short-ranged attractive benzene-wall interaction (cf. figure 7.1). Benzene adsorption on nickel, which due to a strong benzene-nickel interaction of $40 k_B T$ at room temperature ($z = 0.2 nm$) may thermodynamically be favorable, is kinetically prevented within the currently used simulation times. To test if a pre-coating of the benzene molecule onto the surface would be (meta)stable, a second 50 ns MD simulation starting with a benzene molecule coated onto the nickel surface was performed, indeed showing benzene to stay at the surface. From a pre-coated gold surface, however, benzene moves away into bulk water within the first few picoseconds of the simulation. Below, we analyze the thermodynamic aspects in more detail.

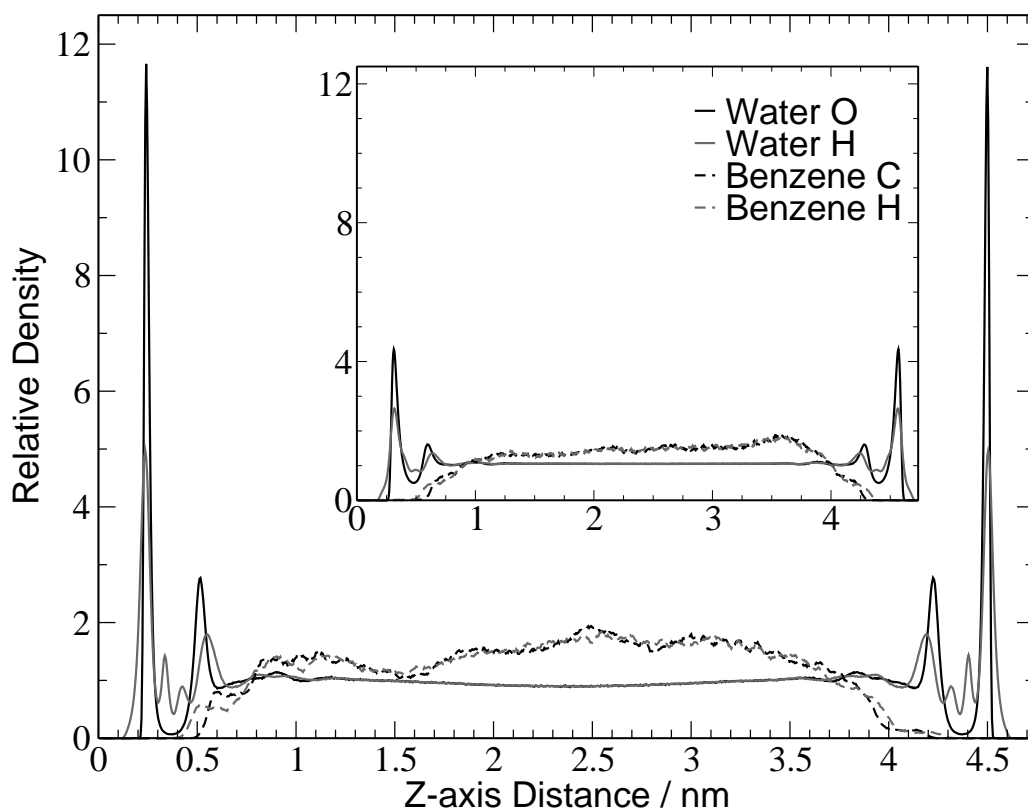


Figure 7.4: Z-axis density profile between two modeled Ni(111) walls. Solid black line: water O; solid grey line: water H; dashed black line: benzene C; dashed grey line: benzene C. System consists of 2500 water molecules and 1 benzene molecule, runtime is 50 ns. The relatively symmetrical density profile for benzene between the walls indicates that sampling is almost sufficient. Main picture: Nickel walls at both z-axis ends. Inset: Similar system, but with Au(111) walls.

7.3.2 PMF and umbrella sampling

The results of the previous section indicates clearly that there is no tendency for the benzene to strongly adsorb onto the Au(111) surface and replace the water. For the case of the Ni(111) surface the situation is more complicated. Taking the plain energy information, as it comes out of the CPMD calculations, the energies are not much different if one considers one benzene replaced by about five water molecules. Thus a plain energetic consideration is certainly not sufficient. In addition the results of the previous section clearly show, that there is a huge barrier for the benzene to penetrate the water layer and adsorb onto the Ni. Thus, to make quantitative

statements on benzene adsorption on nickel as well as water-mediated free energy barriers near the surface we calculated the potential of mean force (PMF) based on the distance z between the nickel surface and the benzene center of mass. The PMF was constructed using a quasi-static approach, measuring mean constraint forces required to constrain the distance z between the surface and the center of mass of the benzene molecule. Since only the distance z between the surface and the center of mass of the benzene was constrained and fluctuations in the x and y direction were permitted, this methodology should deliver the correct measure of the water- and wall-induced forces working on the benzene molecule. The constraint was applied for distances between 0.2 nm and 1.0 nm using 0.01 nm intervals along z , and 0.02 nm intervals between 1.0 nm and 2.0 nm. At every z -point, 3 ns runs were performed to sample the mean force. Forces were stored each time step, the average force and integrated error in the calculation was estimated by block averaging over 1000 time steps. The location and relative magnitudes of the extremes in the PMF provides insights in the mechanism of benzene adsorption as well as possible metastability of a benzene-coated nickel surface.

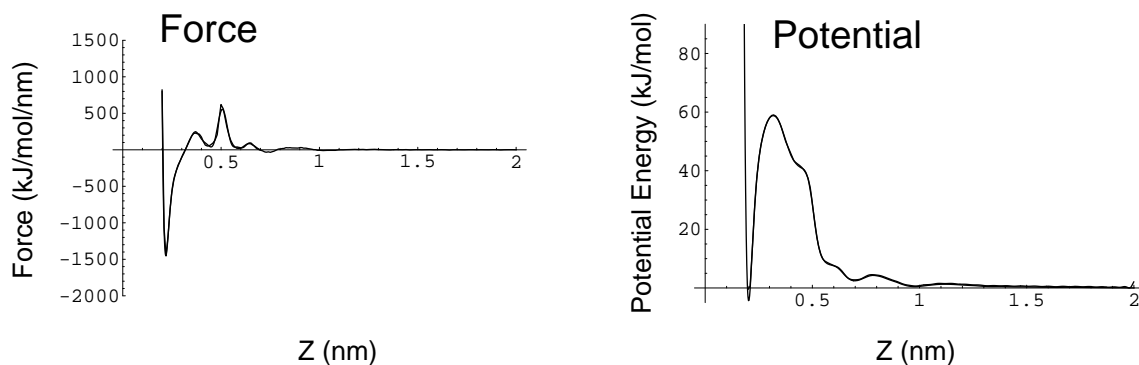


Figure 7.5: Construction of the umbrella forces and potential: The mean forces obtained from the constraint mean force calculations are plotted on the left. A completely overlapping force function (also plotted) was constructed by fitting a function built from Fourier series with 30 independent parameters. The force function was used to subtract forces during the simulation. Similarly, by integrating the fitted force function, a potential could be obtained that completely overlapped with the potential of mean force obtained by the constrained mean force calculations (both fit and original data shown in right graph).

To assess the quality of the resulting potential of mean force, we used this as an input potential for a 100 ns umbrella sampling run, in order to get a more precise estimate of free energy barrier. For this a Fourier series fit of the mean forces ob-

tained in the distance constraint runs with force averaging was made (figure 7.5). In the umbrella sampling simulation, the corresponding z-distance dependent umbrella force and potential was subtracted from the actual force and potential values. In the 100 ns umbrella sampling run the umbrella force on the benzene center of mass was redistributed over the benzene interaction sites as mass-weighted forces. After the simulation, a z-dependent free energy profile was constructed from the z-dependent distribution of the benzene center of mass, corrected by adding the umbrella potential. Figure 7.6 shows the PMF obtained by applying the constraints as well as the one obtained by umbrella sampling. The agreement between the two PMFs is excellent. The integration of the error on the mean constraint force ($z = 2 \text{ nm}$ to $z = 0.2 \text{ nm}$) results in an uncertainty of the free energy of adsorption of 3 kJ/mol , about $1.3 k_B T$ at room temperature. Within the error of our calculation, the adsorption free energy of benzene on nickel, which corresponds to $\text{PMF}(z = 0.2 \text{ nm}) - \text{PMF}(z = 2.0 \text{ nm})$, is negligible. The slightly favorable energy difference driving benzene adsorption apparently is compensated by unfavorable entropy changes. These entropy changes are associated with loss of benzene orientational degrees of freedom near the surface, as well as a more significant loss of solvent translational degrees of freedom in creating a solute-sized cavity close to the surface relative to creating it in the bulk. Due to the "rigid" water structure close to the hydrophilic surface (see figure 7.4), density fluctuations resulting in the formation of a suitable cavity are suppressed thereby imposing strong excluded volume restrictions on the benzene solute to contact the surface (by performing test-particle insertions of smaller, methane-sized molecules we observed that insertions were hardly ever successful close to the surface while in the bulk they were readily inserted). Although it is outside the scope of this thesis to quantify these entropy contributions, we note that these are reminiscent to those observed in co-solvent/water mixtures where the interactions of solutes with co-solvent molecules are more strongly opposed by the entropy change if the co-solvent molecules are stronger hydrated.[192, 193] To adsorb or remove benzene from Ni(111) a barrier of $24 k_B T$ must be surmounted, leading to a metastable benzene layer on a nickel surface in the presence of water. For gold no such metastable regime can be found. Even though both water and benzene have a weak interaction with the Au(111) surface, the competition between water and benzene adsorption onto the surface is dominated by the small size of water and its hydrogen bonded network at the surface. This will prevent any metastable state of benzene at the gold surface.

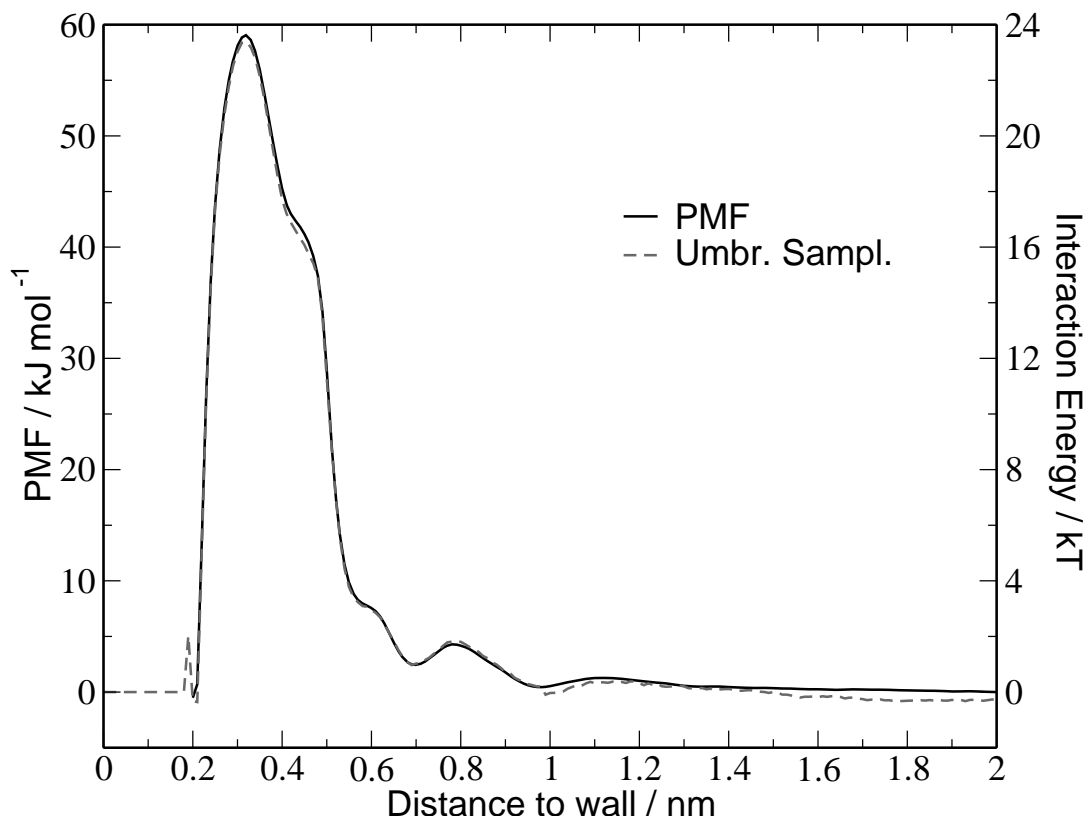


Figure 7.6: PMF as obtained by constraining benzene-wall distances (solid black line) and the potential retrieved from the density-profile of an umbrella sampling run, with a correction for the umbrella potential (dashed grey line). The umbrella sampling potential was applied between 0.203 and 0.981 nm, the discontinuity of the curve around 0.20 is a result of one constrained PMF data point lying below the umbrella region.

7.3.3 Influence of the water model

In chapter 6, the water density profile at the surface was shown for 5 classical water models (see fig 6.3).

Both the benzene and phenol z -dependent PMFs (Fig 7.7, see the appendix for a detailed description) are qualitatively very similar. The structural details of the PMFs follow the water oxygen density fluctuations (Fig 6.3). Upon approaching $z = 0.5$ nm (corresponding to the position of the second oxygen peak in the water density profile) from larger distances, the benzene (phenol) molecule starts to expel water from the second water adsorption layer and the PMF increases rapidly. The PMF profiles correlate with the water density data from figure 6.3, where we can

find a first water layer around $z = 0.2\text{-}0.3$ nm, a second water layer around $z = 0.5\text{-}0.6$ nm, and a very weak third water layering around $z = 0.8\text{-}0.9$ nm.

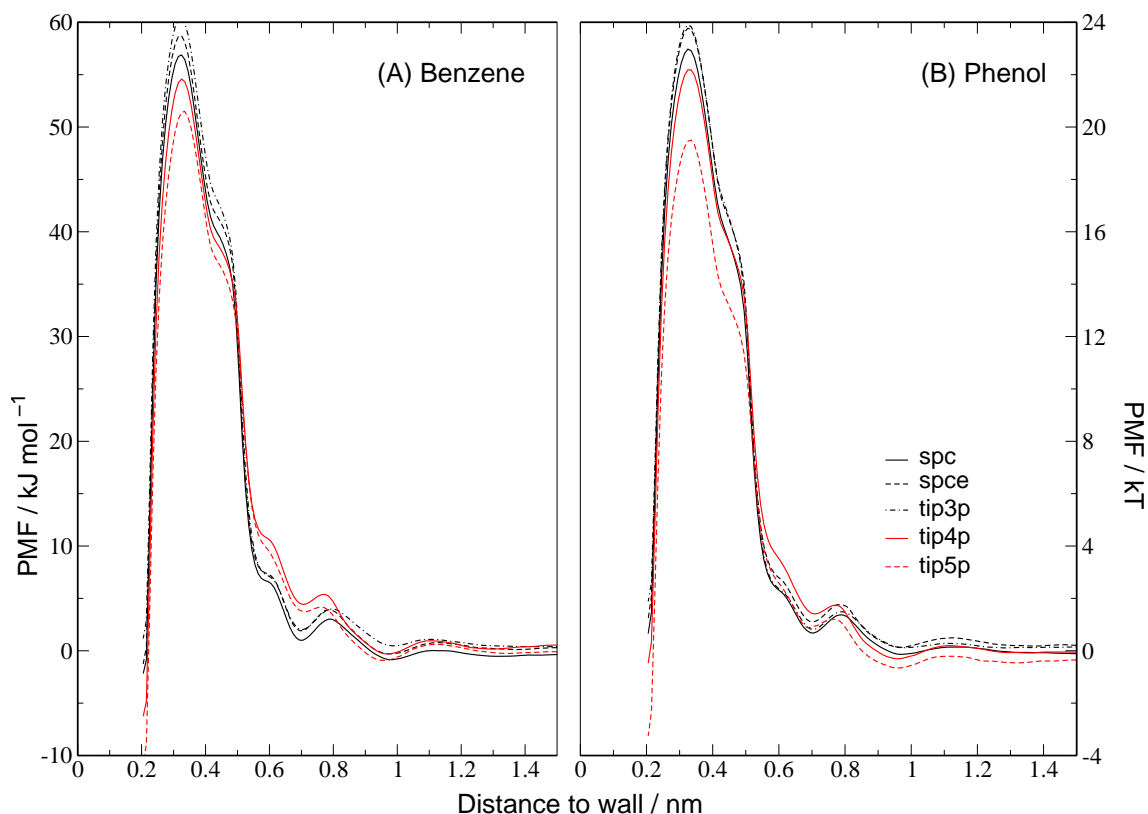


Figure 7.7: Solute-surface PMFs for displacing the geometrical center of the phenyl ring of (A) benzene and (B) phenol perpendicular to a surface, in liquid water (300 K) described with five classical water models. PMFs were obtained by integrating the average constraint force on the solute center of mass along 140 discrete points between 0.20 and 3.00 nm from the metal surface. At each point, 3 ns MD runs were performed to sample the mean constraint force. Solid black line: SPC water[141]; dashed black line: SPC/E water[178]; dashed/dotted black line: TIP3P water[179]; solid grey line: TIP4P water[179]; dashed grey line: TIP5P water[180]. The errors of these PMF calculations ranged between 4 and 8 kJ/mol at $z=0.2$ nm, and was estimated by calculating the block average error estimate[144] for every constraint distance and taking the square root of the integral of the square of all error estimates from bulk ($z = 3.0$ nm) towards the closest distance to the surface that was sampled ($z = 0.20$ nm).

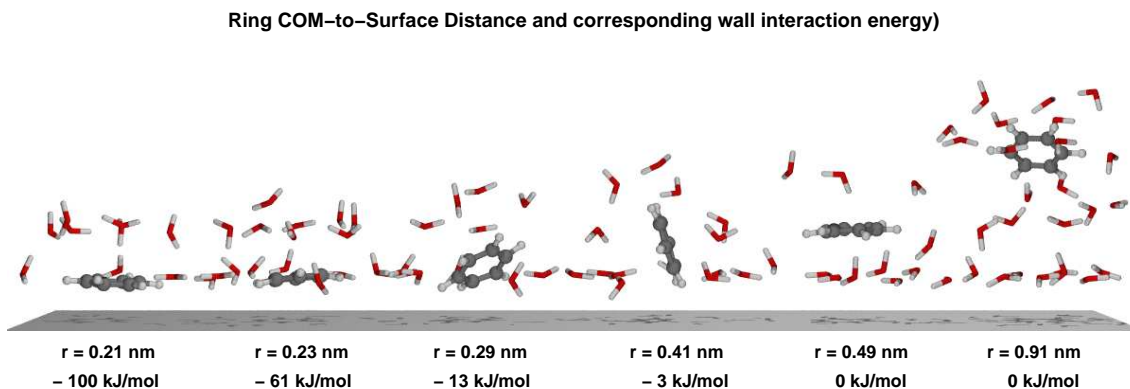


Figure 7.8: Desorption mechanism, from left to right, as shown by snapshots from the umbrella sampling run. The first solvation shell and all water molecules between the molecule and the surface are shown for six conformations at various distances from the surface.

7.3.4 Mechanisms

To get an impression of the desorption mechanism, several snapshots of benzene conformations near the surface were made (figure 7.8). Note that due to the liquid nature of the surrounding of the benzene molecule and the structureless adsorbing surface, the plain distance dependence trivially defines the natural "reaction path". From these snapshots it appears that the first step in desorption is a fluctuation in the benzene-surface distance, moving it some fractions of an Ångström away from the surface. Due to the rapid decay of the interaction energy with the benzene-wall distance (figure 7.1), the interaction energy already drops down significantly at this point, however does not allow for any water to fill into the gap. As a second stage, the benzene starts inclining and the interaction energy drops down further, until the benzene is oriented perpendicular to the surface and almost no interaction energy is left. At distances further away from the surface, water (5.3 molecules on average) will completely fill up the gap left by benzene, re-adsorption from that point is clearly hindered by the closed water-layer.

In figure 7.9 (A), 2-dimensional PMFs for benzene and phenol are shown as a function of z and the inclination angle between the aromatic ring- and surface normals. A selection of snapshots of benzene at various surface distances is provided in fig. 7.9 (B). In this 2-dimensional free energy landscape, a more detailed picture

is shown, and we can subdivide several zones when approaching the surface ($z = 0$ nm) from the bulk solvent ($z > 1$ nm). In the bulk zone going from $z > 1$ nm towards $z = 0.8$ nm the sampling is evenly distributed over all angles. For benzene (less clear for phenol), it can be seen that around 0.7 nm (in between the third and second water layer; see Fig 6.3) already some structural effects occur and the parallel conformation ($\cos(\theta) = 1$) is slightly less sampled. At distances corresponding to the second adsorbed water layer (between $z = 0.6$ nm and $z = 0.5$ nm), benzene (phenol) preferentially orients parallel to the metal surface (the distribution of $\cos(\theta)$ is narrowed down to 0.9-1.0 in fig 7.9 (A)). A snapshot of the parallel conformation can be seen in fig. 7.9 (B), at 0.49 nm). This orientation is favored energetically due to O-H $\cdots\pi$ hydrogen bonding involving water molecules in the first and second surface hydration layers. This type of weak water-aromatic hydrogen bonding is properly described by the force field [194]. In figure 7.7, this causes a shoulder to appear in the PMF at $z=0.58$ nm. A rather flat landscape with respect to the orientational degree of freedom is found in the region between the second and the first adsorbed water layer (0.4 nm $<z<$ 0.5 nm). Though, towards $z= 0.4$ nm, perpendicular orientations are favored in comparison to parallel ones, which become energetically unfavorable. In figure 7.7, a shoulder is observed at these distances, in figure 7.9 (B) one can see how a perpendicular orientation minimizes the displacement of water molecules in both the first and second adsorbed water layer. As the ring center of mass approaches the first adsorbed water layer, the benzene (phenol) ring normal gets significantly tilted with respect to the surface normal and finally lines up with the surface normal, driven by an energetic stabilization of 1 eV due to the overlap of benzene(phenol)- π -orbitals with free electrons in the surface. In this process, first layer water molecules are expelled from the hydrophilic nickel surface, even before significant benzene(phenol)-surface binding interaction is present. It is clear that if the benzene would adsorb in a solvent-free environment, the z -dependent angular distribution plot would show a random orientation for distances of ca. 0.5 nm and higher, and will show a similar profile to the solvated state only for short distances below 0.3 nm, as here the orientation of the ring is governed by the presence of the surface.

The mechanism above originates from the very short-ranged and strong benzene-surface interaction, combined with a strong wetting of the transition metal surface by water. Without a need for additional simulations one can predict the same mechanism to hold for other transition metals in case strong, short-ranged surface

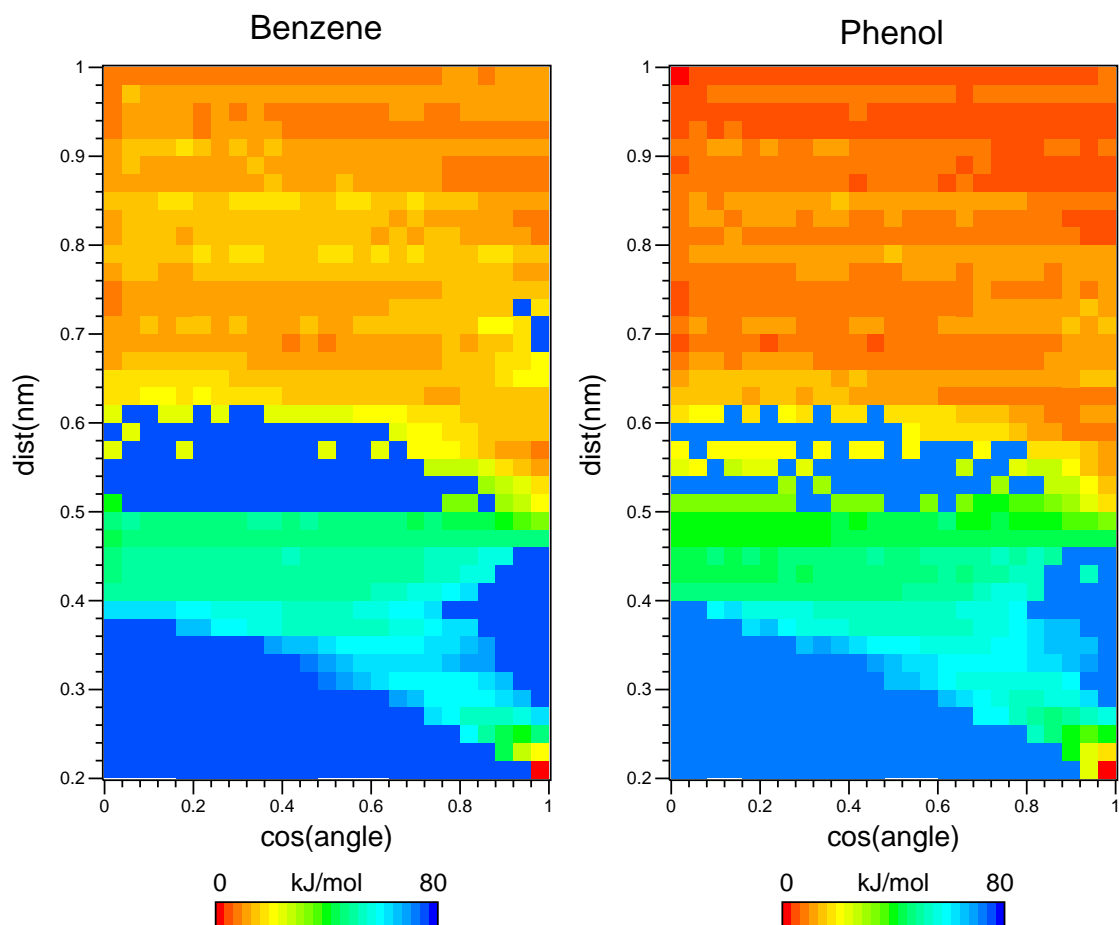


Figure 7.9: Figure 7.9 Two-dimensional PMFs for benzene (left) and phenol (right) at Ni(111) in SPC/E water (298 K). The solute (center-of-mass)-to-surface distance is plotted vertically, the cosine of the angle between the solute- and surface normal vectors is plotted horizontally. The 2D-PMF, $G(z, \theta) = G(z) + G(\theta | z)$, was calculated from constrained MD. By applying a constraint to keep the z -coordinate fixed, $G(z)$ (see fig 7.7) is obtained by integrating the mean constraint force along the z -coordinate. $G(\theta | z) = -k_B T \ln P(\theta | z)$ is obtained from the conditional distribution function, $P(\theta | z)$, sampled in the constrained MD runs. Compare with the snapshots in figure 7.8.

interactions combine with approximately equal ratios of benzene and water adsorption energies (determined by either Density Functional calculations or experiment). If one considers the energies listed in table 7.3, one likely expects the mechanism in figure 7.8 to occur also at the (111)-surfaces of Ni, Pt, Pd, and Rh. Although in our quantum modeling we do not have the contribution of dispersion forces, these will not change the observed mechanism, because they are weaker and much longer-ranged. This will most probably remain true also for the noble metal surface

E_{ads} (kJ/mol)	H ₂ O	C ₆ H ₆
Ni(111)	24.1	105.2
Pt(111)	33.8 [195]	86.8[196]
Pd(111)	31.8 [195]	114.8[196]
Rh(111)	40.5 [195]	147.6[196]
Au(111)	9.65	4.82

Table 7.3: Energy data from quantum calculations for a series of metal surfaces

Au(111), since including dispersion effects in the simulation will increase both the water and benzene adsorption simultaneously. Only in the case if dispersion for benzene would be several factors higher than for water, the inclusion of dispersion would change the overall picture shown in the current chapter. No dispersion data of these molecules near a gold surface currently exist, and further research will be necessary for an exact clarification.

7.4 Discussion

The simulations performed with benzene at Au(111) and benzene and phenol at the Ni(111) showed that water-mediated effects are crucial in modeling the interaction of benzene with metal surfaces. This illustrates that the current dual scale approach that combines electronic structure quantum calculations and classical descriptions of solvation is an appropriate tool to study the adsorption phenomena out of solution for small molecules.

In the following chapter, the same technique will therefore be applied to a small set of amino acids, as a next step towards the full multiscale modeling of biological macromolecules.

8 Modeling hydrated amino acids near metal surfaces

The modeling methodology for solute-surface interactions in the hydrated state that was constructed in the previous chapters is extended here to model the interaction of the three amino acids alanine, phenylalanine, and tyrosine, with the Ni(111) surface. The neutral forms of phenylalanine and alanine will be atomistically simulated in explicit water at room temperature, at the interface with a surface representing Ni(111). The increased conformational degrees of freedom for these molecules will be shown to complicate the simulation and the interpretation of its results, however still several general properties of the adsorption can be obtained and compared to related experimental work.

8.1 Computational Details

The model presented here extends the 'molecular building block' approach that has been already successfully applied to organic-inorganic interfaces[6, 42] in earlier work, as for example the previous chapter of this thesis. To reiterate, note that the main idea in this approach is that macromolecule-surface interactions are described at two levels. Firstly, building block molecules are chosen that describe the recurring parts in the macromolecule. The small sizes of these building blocks allow for quantum mechanical calculations of the building block-surface interactions, and the resulting data are used for parameterizing atomistic solute-surface interaction potentials. On the other, classical atomistic, level, this quantum-based parameterization describes building block-surface interactions, not only for the single building blocks, but also for the structures such as (oligo)peptides or proteins that can be constructed from combinations of these building blocks. The current chapter will treat the neutral forms of alanine, phenylalanine and tyrosine, and study their in-

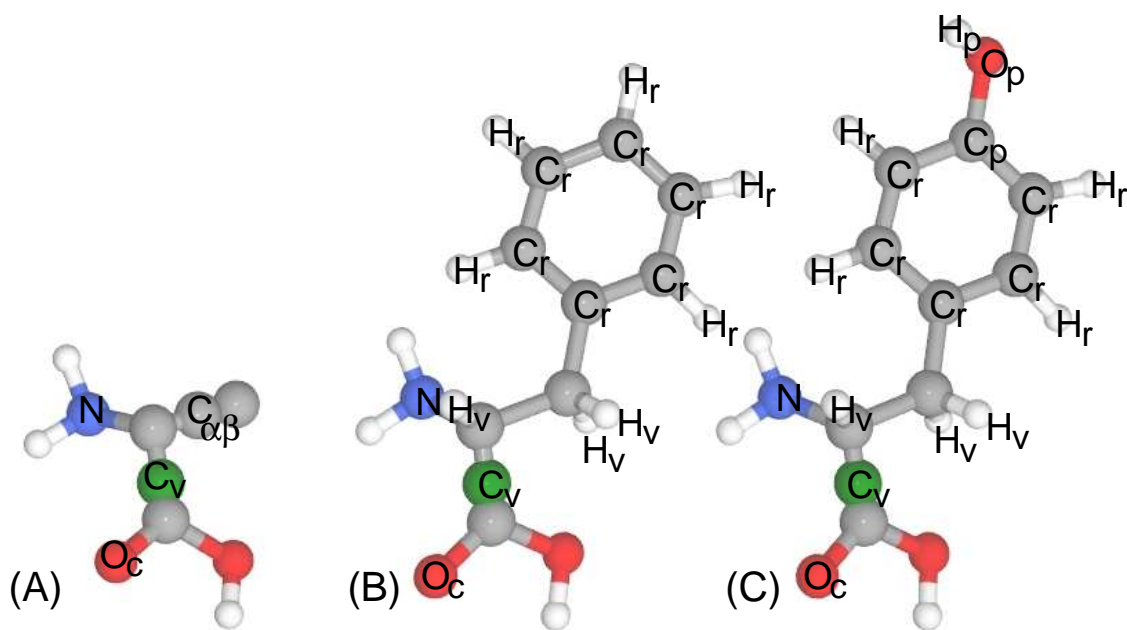


Figure 8.1: Molecules used in the modeling study: (A) neutral Alanine (B) neutral Phenylalanine (C) neutral Tyrosine. Atom names refer to atom names in table 8.2. The sites $C_{\alpha\beta}$ and C_V are specially introduced for the current modeling: $C_{\alpha\beta}$ is a virtual site located exactly at the center-of-mass of C_α and C_β . C_V is a repulsive virtual site needed to prevent simultaneous adsorption of the N and carbonyl O (O_C) atoms within the same molecule, as explained in the text. Note that the water hydrogens H_W do not have any interaction with the surface. Unlabeled atoms have the general repulsion 'R'.

teraction with a Ni(111) surface from the bulk-hydrated state. As Ni is taken only as an ideal model surface, any oxidation effects at the surface will be ignored.

Quantum calculations of the amino acids alanine and phenylalanine have been performed using the DFT based finite-electronic temperature method of Alavi *et al.* [187] (FEMD), as implemented in the CPMD code [189]. Note that amino-acids under physiological conditions exist mainly (> 99%) in the zwitterionic state, whereas in the case of peptides the end-groups are condensated to form peptide bonds and are therefore of different chemical nature. The neutral amino acids used here are not the exact representations of any of those two states, but still are of interest, as the neutral state (as well as the zwitterionic state) has been shown to bind to a Ni(111) surface[72]. The zwitterionic state of amino acids as well as the peptide form are currently under investigation[72] and will be a matter of further research.

Simulations were carried out in a system similar to the one described in chapter 6. As solutes, the molecules described above were used, for which the solute-solute, solute-solvent, and solvent-solvent intermolecular potentials were described, as in the previous chapters, by the GROMOS 43a1 force field[88]. As a solvent the SPC/E water model was used[178]. Note that in chapter 6 it was already shown that the choice of the water model has no influence on the water density profile at the surface. Repeating the same study with a different water model would therefore be possible without significant differences in the water-mediated adsorption behavior.

In addition to MD simulations, several simulations applying a Langevin thermostat (at 300K) were performed in which the solute interacts with the metal surface in a vacuum environment. The Langevin thermostat, available in Gromacs[94], was used with a friction coefficient of 1 ps^{-1} .

Solute-surface potentials of mean force (PMFs) in water were obtained via constraint-biased simulation with force averaging[197] as provided by the Gromacs package. The constraint direction was chosen perpendicular to the surface (i.e. in the z -direction). For each PMF, 142 defined constraint distances from the surface were chosen, in steps of 0.01 nm from 0.19 nm to 1.00 nm, in 0.02 nm steps from 1.00 to 2.00 nm and in 0.10 nm steps for distances up to 3.00 nm. In all cases, this proved to be far enough to reach a bulk hydrated state of the molecule as indicated by a constant value for the PMF at these distances. Several constraint sites on the solute molecule were analyzed, each separately, to see if any site-dependence for the PMF is present. The following interaction sites were chosen as constraint sites: the center of the ring for benzene and phenol; for alanine the carbonylic oxygen and the amine nitrogen; for phenylalanine the ring center, carbonylic oxygen and amine nitrogen. It is important to note that only the z -component of the constraint site has been fixed. The constraint site can move in the xy directions, and the remaining parts of the molecule can move in any directions as long as this movement will not displace the z -distance of the constraint site.

The starting conformations at the 142 distances mentioned above were generated by first solvating the molecule in the middle of the slit, following by pulling it to an adsorbed state (at $z=0.19 \text{ nm}$). From there the z -distance was increased in a sequence of short pull runs. Once all starting points were generated, 3 ns production runs were performed for each z -distance, enabling the calculation of distance-dependent mean forces, and mean surface interaction energies.

8.2 Modeling

The quantum-atomistic multiscale modeling of amino-acid-metal surface interactions is more complex than the previous multiscale modeling of benzene/polycarbonate adsorption (see chapter 5 and references [6, 42]). This is largely due to the low symmetry and relatively large number of interaction sites in amino acids, and the fact that these interaction sites are located too close to each other to be treated separately. The low symmetry greatly affects the load of quantum calculations necessary: Even though alanine has fewer atoms than benzene, there are many more possible geometries in which it can be oriented relative to the surface. A complete quantum analysis of all these configurations at all possible metal lattice sites would require a very expensive computational effort. An expansion to the multiscale modeling methodology is introduced here that can overcome the before-mentioned problems in a way that actually enhances the methodology's general validity and applicability.

It is important to note that the attention lies on finding global minimum free energy conformations, as these will be the dominant contribution to the sampling statistics during long runs at the atomistic level. Calculations on intermediate, higher-energy states would provide only superfluous information at the current modeling precision. Therefore, the starting point of the modeling will be a selection of initial configurations that aims to give us the strongest binding conformations. The electronic properties of the metal binding of these configurations is calculated by a series of quantum density functional calculations (geometry optimization)[72], considering per configuration the various possible positions on the metal lattice. The data retrieved from these calculations give us the parameters (interaction sites, interaction strength, optimal distances) based on which an atomistic model is constructed. However, since the molecules studied here have a large number of orientational degrees of freedom, the molecule is likely to get trapped in local energy minima during quantum calculations. This is resolved by performing Langevin thermostat MD simulations of the molecule in vacuum, applying the modeled molecule-surface interaction parameters, and sample its configuration space at a range of distances perpendicular to the surface, using the constraint method and sites described in the Methodology section. This set of Langevin thermostat MD runs can sample a larger amount of the phase space than could be done by quantum optimization runs in a fraction of the time. If the Langevin thermostat MD simulation generates (only) the surface adsorbing conformations that were already found in the initial

quantum calculations, the classical model is considered finished. If, however, the Langevin thermostat MD simulation generates adsorbing (low-energy) conformations that were not previously found in quantum calculations, these conformations are used as starting configurations in a new series of quantum-based structure optimizations. If these quantum calculations show that the conformations found differ in their interaction properties from the Langevin thermostat MD simulations, the atomistic modeling is adjusted. Then, the procedure is repeated (Langevin thermostat run with new modeling, quantum calculation of Langevin run output), until consistency between the quantum and the atomistic level is reached and all Langevin run lowest energy conformations correspond to stable conformations from quantum calculations with respect to the energy and the configuration.

8.2.1 Neutral Alanine

For neutral alanine four distinct conformations were used in the quantum calculations. The structures are shown in figure 8.2, energies and optimal surface interaction site distances of all configurations are shown in table 8.1, the exact description of the quantum calculation part of the simulations is described in ref. [72]. One of the structures (number 4 in table 8.1 and figures 8.2 and 8.3) turned out to be nonbonding.

No.	Conf.	E_{ads} (eV)	(kJ/mol)	N (nm)	O_c (nm)	C_β (nm)
1	$N_{\text{down}}CH_{\text{up}}$	-0.57	-55	0.22	0.32	0.46
2	$N_{\text{down}}CH_{\text{down}}$	-0.32	-31	0.24	0.33	0.36
3	$O_{\text{down}}CH_{\text{up}}$	-0.37	-36	0.46	0.25	0.57
4	$O_{\text{down}}CH_{\text{down}}$	non-bonded		0.39	0.22	0.27

Table 8.1: Properties of four neutral alanine conformations that were evaluated in quantum mechanical DFT calculations[72]. Shown are the adsorption energy on the Ni(111) surface, and the atom - to - top Ni layer distance of three of the atoms whose surface distance most strongly affects the surface interaction.

The amine nitrogen (N) and the carbonyl oxygen atom (O_c) interact attractively with the surface. Surface attraction at these sites is of the order of one to two hydrogen bond strengths and is modeled with attractive 10-4 potentials. An interesting feature is the influence of the methyl side-group. Even when not directly

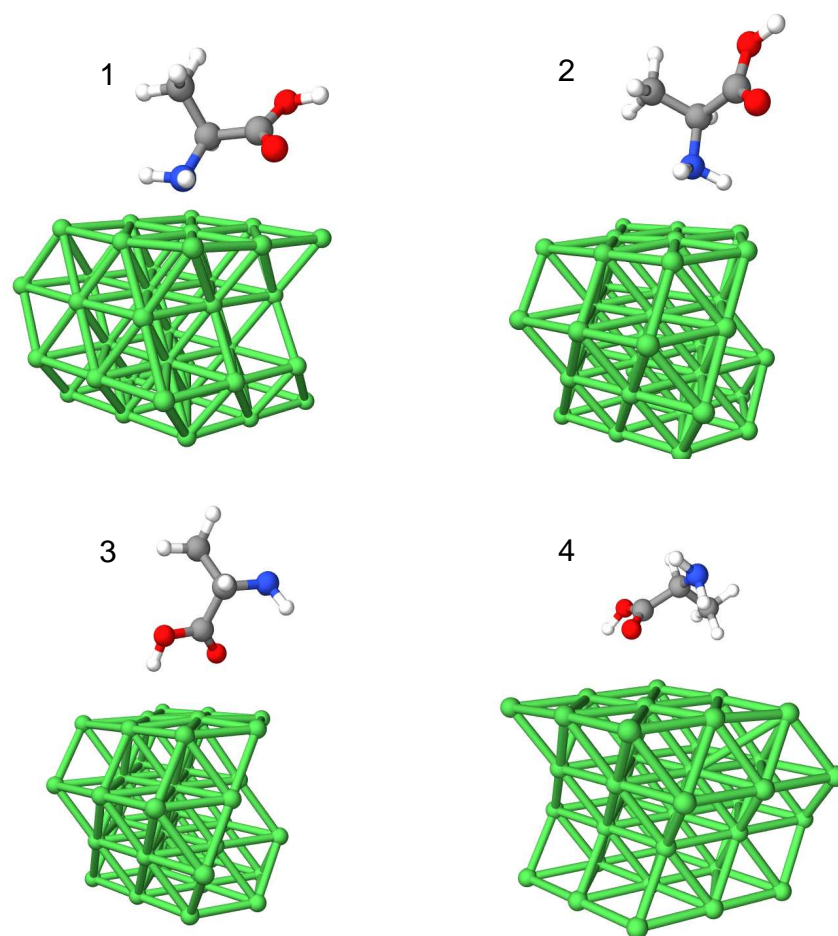


Figure 8.2: Alanine-Ni(111) conformations as mentioned in table 8.1 and figure 8.3. Main characteristic of conformation 1 is the binding of the amine nitrogen to a top site of the Ni(111) surface and the methyl group pointing away from the surface. The carbonylic oxygen is pointing slightly towards the surface. Conformation 2 is similar, but with the methyl group closer to the surface. Conformation 3 has the carbonyl oxygen bound to the surface, with the methyl group pointing away and the bond connecting the C-alpha and the amine nitrogen almost parallel to the surface. The difference between conformations 3 and 4 is almost solely the fact that in conformation 4 the methyl group is closer to the surface. An analysis of these structures at the quantum level is given in ref. [72].

interacting with the surface, the methyl-surface distance is found to influence the total interaction energy, although relatively weakly.

The origin of this influence might be an effect of the position of the side chain on the surface orientation of the interacting amine and carboxyl groups. A simple

scheme to model this effect is chosen that includes a weak repulsive Morse potential (equation 3.13) on the center-of-mass of the C_α and the C_β atoms: this additional interaction site is referred to as ($C_{\alpha\beta}$) side-chain (see figure 8.1).

The optimal distances (σ -values) for these potentials are retrieved from the binding conformations in quantum calculations. Binding energies for the potentials (ϵ -values) are chosen such that the sum of all site-surface potentials (eqs. 3.12-3.13) acting on the molecule will reproduce the total interaction energy as found in the quantum calculation, for every analyzed conformation. These are only initial values, as the interplay between all binding energies and intramolecular interaction energies might eventually lead to binding energies and conformations in the atomistic modeling that deviate from stable states that can be found in quantum calculations. Atoms with no specific surface interaction were given a simple repulsive 10-4 potential (with parameters based on the repulsive hydrogens in the benzene ring, see group R in table 8.2) to prevent the occurrence of unphysical conformations corresponding to a penetration of the surface. The first iteration of this modeling resulted in minimal energy conformations from the Langevin thermostat MD runs that had, both, the amino nitrogen and the carbonyl oxygen interaction sites at optimum distance. The total interaction energy corresponded to the sum of both interactions. Quantum calculations of the Langevin thermostat run output structures showed that this was actually a non-bonding conformation. Consequently, a modification of the initial modeling was needed to assure that only one interaction site at a time will be able to bind with optimal interaction energy. This was reached by introducing a repulsive site (C_v) at a position between the amino nitrogen and the carboxylic oxygen (at 40 % of the bond length between carboxylic carbon and C_α , see figure 8.1), leading to a seesaw-like mechanism which allows for the binding of either the amino or the carbonyl group.

Langevin thermostat runs with this modeling lead to optimal configurations close to the previous quantum calculations, see figure 8.3, and this modeling was therefore chosen for the MD production runs. The final potential parameters are given in table 8.2. Note that even though only one point (minimized energy) per conformation is used, the fact that four different conformations were used for the potential fit, with various groups interacting simultaneously for every conformation, makes this more elaborate than a one-point parameterization. The shape of the average potential energy curves of the Langevin thermostat MD runs of the atomistic modeling in

Interaction	Potential Type	$\epsilon(kJ/mol)$	$\sigma(nm)$
Non-charged Ala			
O _c -Ni	attractive 10-4	8.90	0.23
N-Ni	attractive 10-4	15.0	0.22
C _{$\alpha\beta$} -Ni	Rep. Morse (a=6.0nm ⁻¹)	4.0	0.58
C _v -Ni	repulsive 10-4	10.0	0.38
R-Ni	repulsive 10-4	4.27	0.20
Non-charged Phe			
O _c -Ni	attractive 10-4	8.90	0.23
N-Ni	attractive 10-4	15.0	0.22
H _v -Ni	repulsive 10-4	4.27	0.17
C _v -Ni	repulsive 10-4	10.0	0.38
C _r -Ni	Morse (a=35nm ⁻¹)	17.5	0.20
H _r -Ni	repulsive 10-4	4.27	0.20
R-Ni	repulsive 10-4	4.27	0.20
Non-charged Tyr			
O _c -Ni	attractive 10-4	8.90	0.23
N-Ni	attractive 10-4	15.0	0.22
H _v -Ni	repulsive 10-4	4.27	0.17
C _v -Ni	repulsive 10-4	10.0	0.38
C _r -Ni	Morse (a=35nm ⁻¹)	15.8	0.20
H _r -Ni	repulsive 10-4	4.27	0.20
C _p -Ni	Morse (a=35nm ⁻¹)	7.96	0.20
O _p -Ni	repulsive 10-4	1.00	0.25
H _p -Ni	attractive 10-4	0.70	0.22
R-Ni	repulsive 10-4	4.27	0.20

Table 8.2: Modeling of molecule-surface potential parameters. See figure 8.1 for the corresponding structures and equations 3.12-3.13 for the potential functions used. The 'R' covers all interactions with no specific interaction with the surface (unlabeled atoms in figure 8.1). The difference between phenylalanine and tyrosine corresponds to the difference in parameterization of benzene and phenol as shown in table 7.2 and figure 7.2.

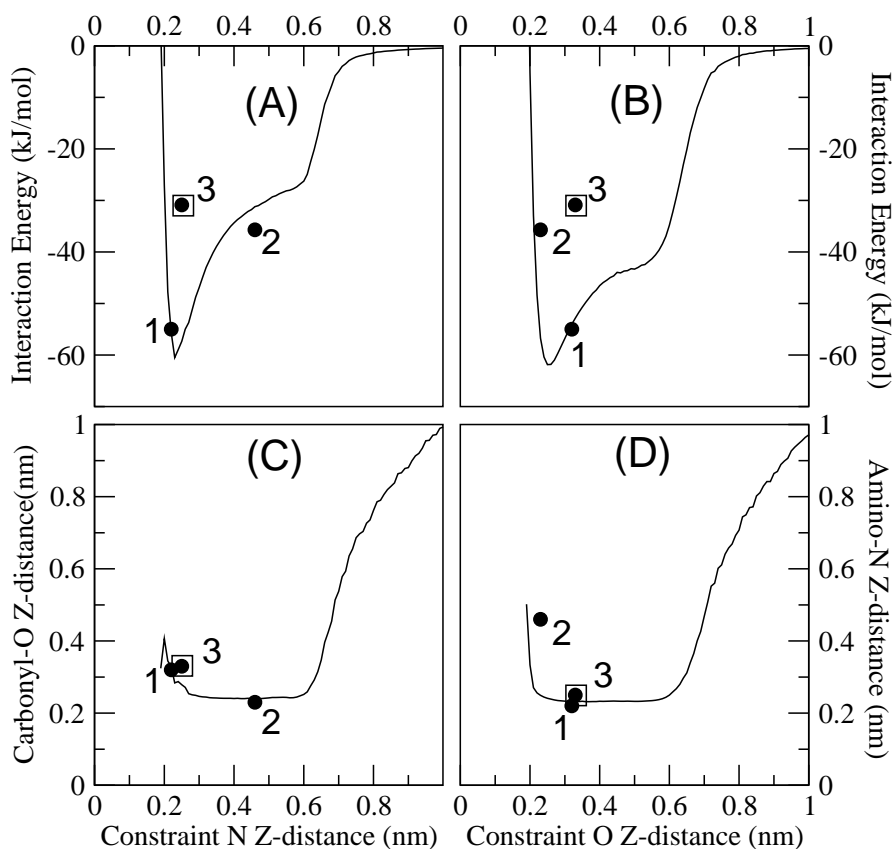


Figure 8.3: Justification of the atomistic modeling. Dots: optimal energy ((A) and (B)) and optimal distance ((C) and (D)) of zero-temperature quantum calculation optimized structures. The numbers of the dots correspond to the numbered configurations in table 8.1. Solid line: average alanine-surface interaction energy ((A) and (B)) and site-surface distance ((C) and (D)) obtained from a series of 3 ns constrained atomistic Langevin thermostat simulations using the modeling parameters from figure 7.2 and table 8.2, in vacuum at 300K. In the left column, for a given constrained distance of the amino N (horizontal axis), the total interaction energy (A) or the average distance of the carbonylic O (C) is shown on the vertical axis. In the right column, for a given constrained distance of the carbonylic O (horizontal axis), the total interaction energy (B) or the average distance of the amino N (D) is shown on the vertical axis. Configurations 1 and 2 are sampled by the atomistic runs either correctly, or within a 0.05 nm distance. Configuration 3 (shown by dots in open squares) has similar positions for O and N as configuration 1, but has a lower interaction energy due to the position of the methyl group close to the surface. Therefore, configuration 1 was the preferred configuration during the Langevin thermostat MD sampling and the interaction energy of configuration 3 is not retrieved in the atomistic sampling. As configuration 4 was non-bonding, it is not shown in this diagram.

figure 8.3 requires additional clarification. There is a clear difference in the shape of graphs (A) and (B), where in the case of (A) the potential seems to reach a shoulder in its interaction energy of about -30 kJ/mol, at 0.6 nm. For (B), at a similar distance, this shoulder reaches about -40 kJ/mol. This can be explained as follows. In the case of (A), the strongest interacting site (N) is constrained at the given z-distance, whereas the rest of the molecule is allowed to move freely. When the N-site is located below 0.6 nm, the other interacting site (O) is able to reach its site of maximum interaction at about 0.23 nm. This can be seen in graph (C), where the carbonyl O distance remains constant at 0.23 nm for the range of constrained amino N distances between about 0.24 and 0.6 nm. At this range, the carbonyl oxygen will have a stable interaction with the surface, resulting in the shoulder in interaction energy observed in (A). Below constrained amino N distances of 0.24, the seesaw mechanism sets in, pushing carbonyl O away from its optimal distance. The interpretation of the shoulder in the case if the constrained carbonyl oxygen simulations (B) is similar, with the only difference that a stronger adsorption energy can be reached when the amino N group will bind the surface for the range of carbonyl O constrained between 0.24 and 0.6 nm.

8.2.2 Neutral Phenylalanine and Tyrosine

Phenylalanine and tyrosine can, in a building block manner, be constructed by combining characteristics of the alanine and the benzene/phenol modeling. However, the side-chain residue in phenylalanine and tyrosine were considered here to be of a different nature than in alanine and discarded the $C_{\alpha\beta}$ interaction site present in the alanine model. Instead, as repulsive aliphatic hydrogen atoms proved to be a necessary modeling element to prevent unphysical conformations with C_{β} hydrogens penetrating the surface, these were introduced (one on the C_{α} , two on the C_{β}) as virtual sites in the phenylalanine and tyrosine molecules. The terminology 'virtual' sites is used because, apart from experiencing surface repulsion, they do not interact with the solvent; in the GROMOS 43a1 force field[88] hydrogens connected to aliphatic C_{α} and C_{β} carbons are absorbed into united-atom potentials centered on the carbon positions. Interactions of other atoms (e.g. solvent) with these aliphatic CH_n groups are described using these united atom potentials. Virtual sites in Gromacs are built from the coordinates of 2, 3, or 4 vicinal atoms; any forces on the virtual sites are spread out over these atoms at the end of each time-step. The bond

lengths and angular potentials of the virtual sites applied here were taken from the all-atom OPLS force field[198]. In the resulting lowest energy conformation obtained from Langevin thermostat simulations in vacuum, we found that, both, the amine nitrogen and the phenyl ring could reach their positions of maximum interaction energy, leading to a total interaction energy of -153 kJ/mol (-1.59 eV). This is different from the initial quantum calculation corresponding to a -106 kJ/mol (-1.1 eV) binding conformation, where the amino and carboxyl group were chosen to point away from the surface, to minimize surface contact of the aliphatic C_β hydrogens[43]. Following the iterative multiscale modeling procedure introduced here, the lowest energy Langevin thermostat conformation was analyzed in a consecutive quantum calculation, and it was found to be a stable conformation -145 kJ/mol (-1.5 eV). One aliphatic C_β hydrogen is close to the surface, but can apparently find a stable position within a hollow site of the surface. As the optimal interaction energy and conformation found in Langevin thermostat runs was close enough to the optimal conformation found by quantum calculations (within the ± 5 -10 kJ/mol ($\approx \pm 0.05$ -0.1 eV) error of the quantum calculation), no further optimization of the modeling was needed. The final combination of atom-surface potentials is given in table 8.2 and figure 8.1.

8.3 Analysis of Surface Interactions at the Ni(111) / H₂O Interface

For benzene and phenol, an in-depth analysis of the surface interaction mechanism can be made. In chapter 7 it was shown how the distance between the surface and the geometrical center of the phenyl ring was a suitable order parameter to obtain free energy and geometric information for the absorption process of benzene and phenol.

Following the surface interaction mechanisms of hydrated amino-acids will be more problematic, not only because of their low symmetry as compared to a benzene ring, but also because of their many degrees of conformational freedom and various interaction sites. Therefore, several order parameters to follow the process can be chosen, all of which will however be interdependent: at any surface distance of a given interaction site, all other interaction sites will contribute to the PMF at

that particular distance. The PMF curves will be given below, but because their interpretation is not straightforward, it will be more clarifying to analyze the free energy difference between the bulk-hydrated state and a selection of states with strong surface interaction, and a description of the interaction energies involved.

To better understand hydration effects the explicit solvent MD simulations will be compared with the Langevin thermostat simulations of the amino acids in vacuum that were performed during the modeling procedure. The total combination of relative surface interaction energies for the various amino acids and comparisons of surface interactions under solvated and non-solvated conditions, can give essential information to understand the main chemical factors that determine metal surface interaction of amino acids, and in this way facilitate the design of surface-binding peptides.

8.3.1 Amino acids

To study surface-interacting amino acid conformations, we performed several MD runs, in which a single interaction site-surface distance (order parameter) was constrained and a PMF as a function of that order parameter was obtained by integrating the mean force from $z = 3$ nm backwards towards the surface. Using this method, the distances with the minimal free energy were determined, as can be seen in figure 8.4. Snapshots of corresponding conformations as well as the average energies and free energies (relative to $z=3$ nm) at these distances are shown in figures 8.5 and 8.6.

Comparison of the free energies in fig. 8.5(A) and 8.5(B) as well as fig. 8.6(A) and 8.6(B) shows that, independent of whether the distance constraint is applied on the oxygen or nitrogen, similar values are obtained, which is an indication that sufficient sampling has been reached. The snapshots show that the amino acid conformations in fig. 8.5(A) and fig. 8.5(B) as well as those in fig. 8.6(A) and fig. 8.6(B) are similar. The free energies for phenylalanine viewed from the amino N (fig. 8.6(A)) and carbonylic O (fig. 8.6(B)) pathways are lower than for the alanine cases (about 9 kJ/mol). This is rather counterintuitive, as the bulky phenyl ring did not contribute to the surface interaction energy for these paths, but its excluded volume requires the displacement of more water molecules from the surface as compared to alanine. An explanation could be that in the case of alanine, the small methyl

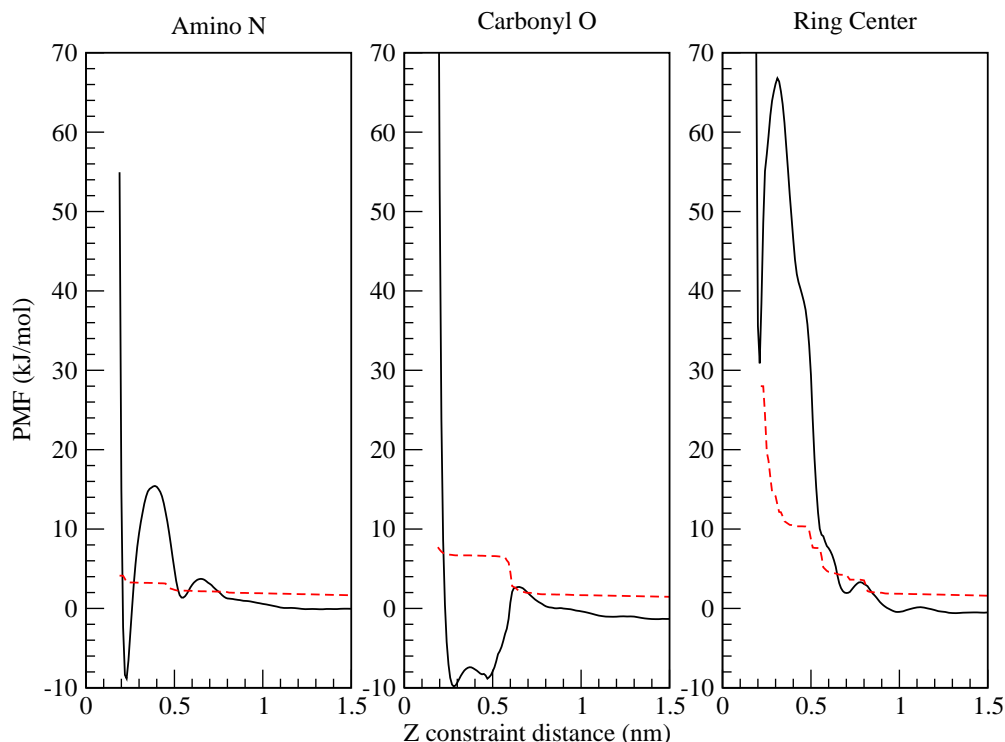


Figure 8.4: PMF curves (solid black lines) and the cumulative error calculated from bulk ($z \leq 3$) to a given z (see also the caption of figure 7.7), as obtained for the three constraint sites: Amino N, Carbonyl O, and Ring Center. For the Amino N and Carbonyl O different energy landscapes are found, with optimal distances at different sites and a ≈ 15 kJ/mol barrier in the case of Amino N. However, it is clear by comparing figure 8.6(A) and 8.6(B) that in both cases a similar state is reached, with a same value for the free energy minimum. The Ring Center constraint site has a profile related to the benzene and phenol PMFs from figure 7.7, though has a unfavorable free energy at the local minimum close to the surface.

group is able to come close to the surface thereby weakening the interaction of amine and carbonyl groups. In contrast, the bulky phenyl group of phenylalanine will not be able to pass the hydrations layers present at the surface, and binding of the amine and carbonyl groups is free. Note, however, that these ΔG values can change dramatically if another set of potentials is needed, for example when studying alanine within a peptide chain. The values presented here should therefore

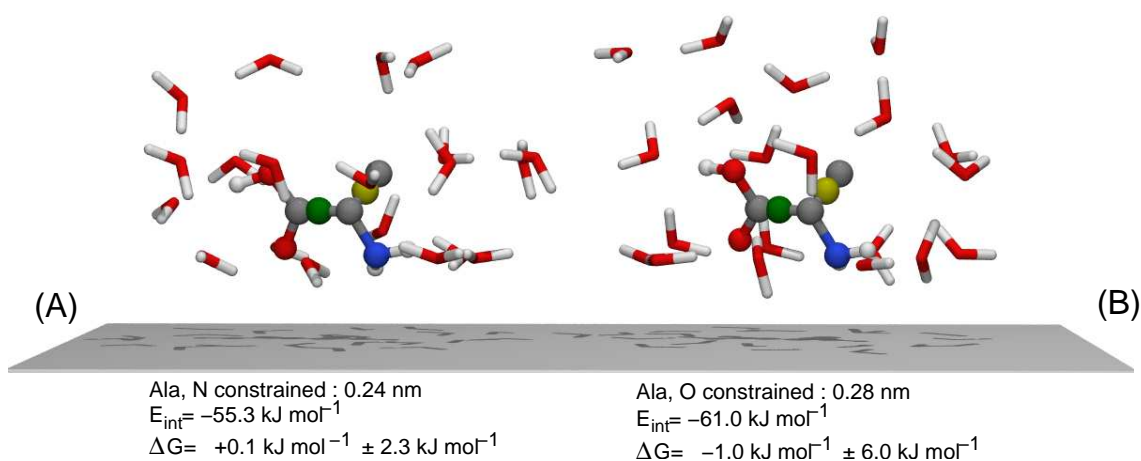


Figure 8.5: Snapshots of amino acids and parts of the surrounding water at the free energy minimum distance of the constraint sites chosen, determined by the PMF method as explained in the text. (A) Alanine, amino N constrained at $z = 0.24 \text{ nm}$ (B) Alanine, carbonyl O constrained at $z = 0.28 \text{ nm}$. E_{int} denotes the sum of interaction energies for all solute interaction sites with the surface. ΔG denotes the free energy of surface interaction, taken by the difference in the PMF between the distance given here, and the bulk state at $z = 3.0 \text{ nm}$. Error bars are calculated by calculating the block average error estimate[144] for every constraint distance and taking the square root of the integral of the square of all error estimates from bulk ($z = 3.0 \text{ nm}$) towards the constraint distance mentioned here. ΔG for a given molecule is found to be independent of the interaction site chosen for the distance constraint order parameter (N for (A) or O for (B)).

be only used as a relative indication to help understanding adsorption processes.

The large positive free energy obtained by displacing the phenyl ring in phenylalanine towards the surface (fig 8.6(C)) requires additional comments. This conformation has the lowest total solute-surface interaction energy of all structures that interact with the surface studied here, but it is the least favorable from a free energy point of view ($\Delta G_{ads} = +32.6 \text{ kJ/mol} \pm 19.8 \text{ kJ/mol}$). A reasoning to explain this observation would be to consider the fact that bringing the ring to the surface requires a displacement of more water molecules than in all other cases. In figures 8.7, 8.8, and 8.9, the surface interaction energies for water, phenylalanine, and the interaction sites of phenylalanine are plotted against the distance of a constraint site used in the calculation. For the simulation with a constrained geometric center of the phenyl ring, the contribution of water-surface interactions decreases over the bulk-to-surface distance with about 140 kJ/mol (see figure 8.9), about 60 kJ/mol more than for the other cases, where, instead of the phenyl ring, the O (figure 8.8), or

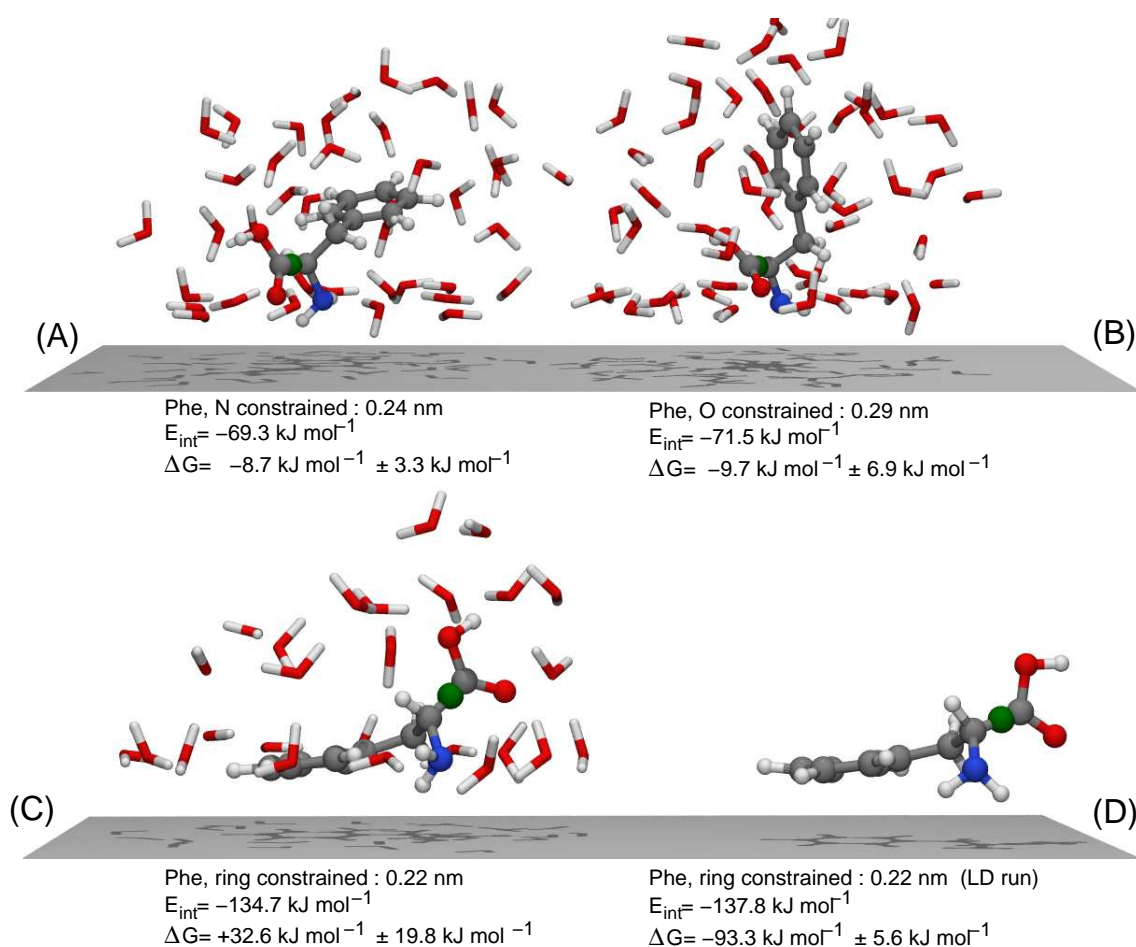


Figure 8.6: Snapshots of phenylalanine and parts of the surrounding water at the free energy minimum distance of the constraint sites chosen. All conditions as mentioned in the caption of figure 8.5. (A) Phenylalanine, amino N constrained at $z = 0.24 \text{ nm}$. (B) Phenylalanine, carbonyl O constrained at $z = 0.29 \text{ nm}$. (C) Phenylalanine, geometrical center-of-ring constrained at 0.22 nm . (D) As in (C), but taken from a Langevin thermostat MD simulation in vacuum. Also here, ΔG is virtually independent of the choice of the order parameter, but by comparing (C) and (D) it becomes clear that ΔG is dramatically dependent on the presence of water.

N (figure 8.7) interacts with the surface. One can get an impression of this solvent effect when comparing the phenylalanine surface interaction in explicit solvent MD (fig 8.6(C)) with Langevin thermostat runs in vacuum (fig 8.6(D)), where free energy minima were found at the same distance (geometrical center of the ring at 0.22 nm), and similar interaction energies (-134.7 kJ/mol in explicit solvent MD and -137.8 kJ/mol in vacuum Langevin thermostat MD), but largely different free energy values ($+32.6 \text{ kJ/mol}$ in explicit solvent MD vs. -93.3 kJ/mol in vacuum

Langevin thermostat MD). Interestingly, in the 'water-free' Langevin thermostat

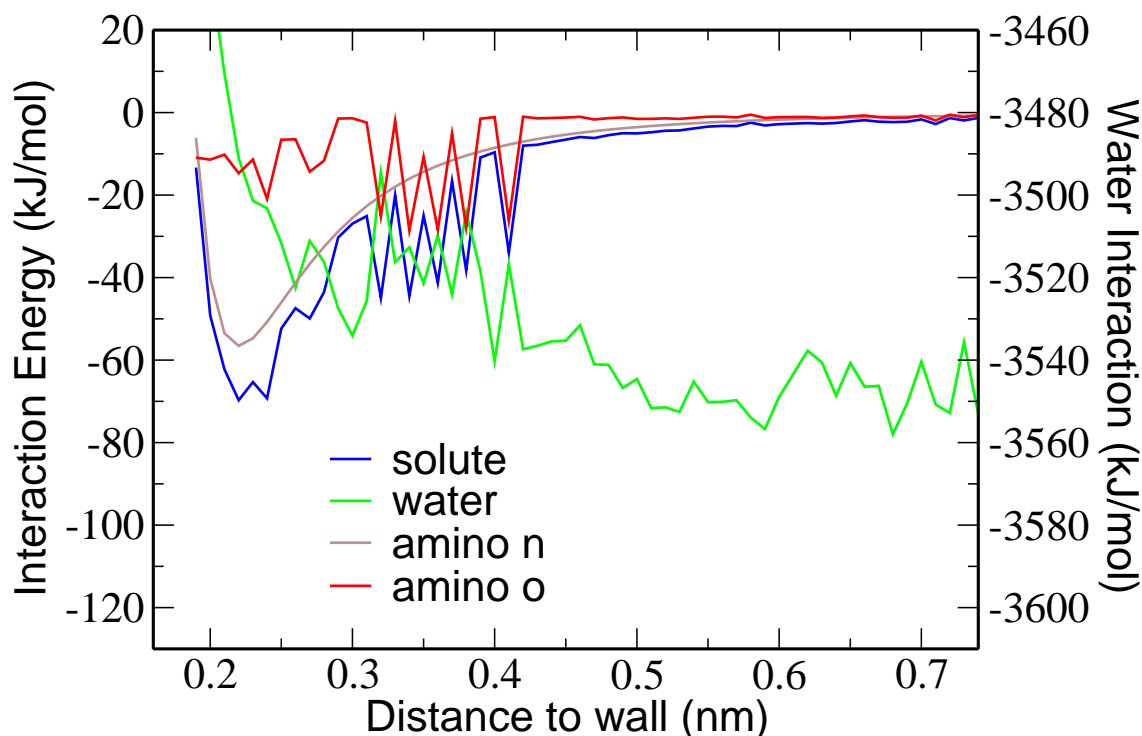


Figure 8.7: Dependence of surface interaction energies (averaged over 3 ns simulations) of phenylalanine and water on the amino nitrogen-surface z -distance at 300K. Shown are the overall solute-surface interaction energy (blue line), the water-surface interaction energy (green line, right vertical axis), carbonyl oxygen-surface interaction energy (red line) and amino nitrogen-surface interaction energy (brown line). The interaction energy between the phenyl ring and the surface was zero for all distances as the phenyl ring did not approach the surface close enough to interact with it.^a When going from bulk ($z \leq 3.0$ nm) to $z = 0.24$ nm (see snapshot (A) in figure 8.6), the solute gains a 69.3 kJ/mol stabilization of surface interaction energy, whereas the water-surface interaction energy is destabilized over the same range by only 49.7 kJ/mol. In the range from about 0.3 Å to 0.4 Å, a strong competition between carbonyl O-surface and water-surface adsorption can be observed.

By looking at the water- and solute-surface interaction energies in figure 8.9, it is apparent that a competition between water and nitrogen is present: every increase in solute-surface interaction energy is accompanied by a decrease of water-surface interaction energy, and vice versa. The destabilization of water-surface interaction energy never exceeds the gain in solvent-surface interaction energy, and the difference

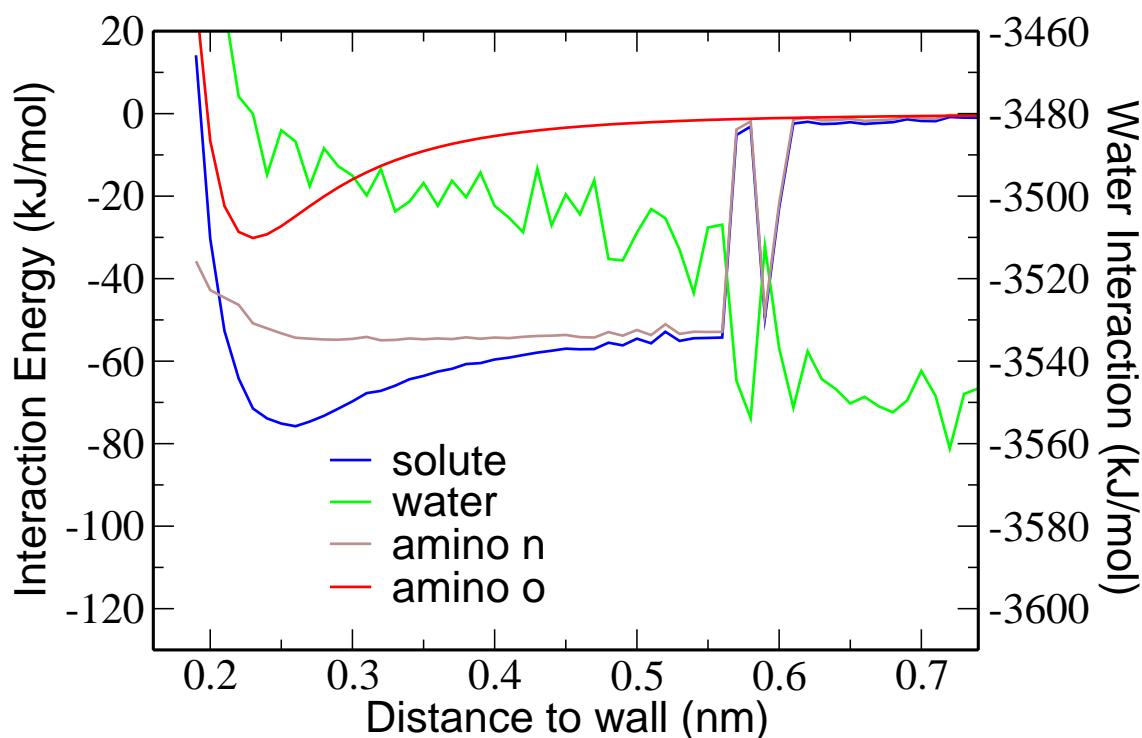


Figure 8.8: Dependence of surface interaction energies (averaged over 3 ns simulations) of phenylalanine and water on the carbonyl oxygen-surface z -distance at 300K. Shown are the overall solute-surface interaction energy (blue line), the water-surface interaction energy (green line, right vertical axis), carbonyl oxygen-surface interaction energy (red line) and amino nitrogen-surface interaction energy (brown line). The interaction energy between the phenyl ring and the surface was zero for all distances as the phenyl ring did not approach the surface close enough to interact with it. When going from bulk ($z \leq 3.0$ nm) to $z = 0.29$ nm (see snapshot (B) in figure 8.6), the solute gains a 71.5 kJ/mol stabilization of surface interaction energy, whereas the water-surface interaction energy is destabilized over the same range by 60.3 kJ/mol. Interestingly, hardly any competing effects between surface adsorbing groups in the solute and water-surface adsorption is seen, which might explain the absence of a barrier for this constraint site in figure 8.4.

between those two is in the same order for all three sets of distance constrained runs. Still, the absence of a -150 kJ/mol surface interaction state in the explicit solvent MD run shows that in this competition the water interaction eventually overrides the chance of an amino N having a maximum interaction with the surface. The similarities in the surface interaction mechanism of phenol and benzene (fig 7.9) indicates that the tyrosine-surface interaction is similar to that of phenylalanine.

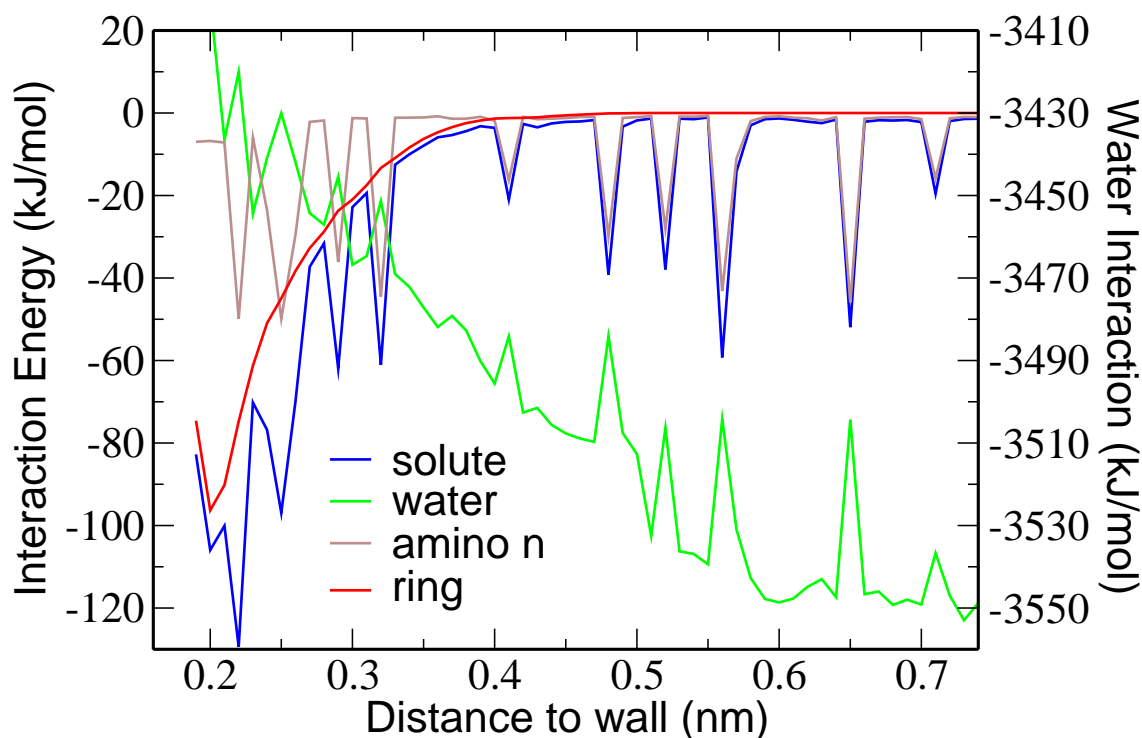


Figure 8.9: Dependence of surface interaction energies (averaged over 3 ns simulations) of phenylalanine and water on the phenyl-ring-surface z -distance at 300K. Shown are the overall solute-surface interaction energy (blue line), the water-surface interaction energy (green line, right vertical axis), phenyl-ring-surface interaction energy (red line) and amino nitrogen-surface interaction energy (brown line). Water is expelled as the ring approaches the surface, causing a loss in water-surface interaction energy of approximately 140 kJ/mol over the distance from 0.75 nm to 0.20 nm. Over the same distance the solute gains a maximum of 134.7 kJ/mol at 0.22 nm (see snapshot (C) in figure 8.6), whereas the water-surface interaction energy is destabilized over the same range by 125.5 kJ/mol. The statistical noise in the water-surface and solute-surface interaction energy profiles is correlated due to a competition of water and amino nitrogen binding. The large statistical uncertainty (19.8 kJ/mol) in the free-energy calculation (snapshot (C) in figure 8.6) is due to the noisy energy profile. The free energy barrier in figure 7.7 can be explained from the water-surface and solute-surface interaction energies: for $0.35 \text{ nm} < z < 0.6 \text{ nm}$ the water-surface interaction energy decreases and is not compensated by an increase of solute-surface interaction energy.

8.4 Discussion

To get a complete theoretical picture of peptide adsorption on metal surfaces, four factors are necessary at the level of computer simulations. First, any molecule-metal surface interaction has to be parameterized by quantum calculations and can not be represented by generalized force-fields. Therefore, one of the goals of this chapter was to develop an efficient methodology to obtain a correct parameterization. A second factor is the competitive adsorption energies of the solvent and solute, where one should take into account the number of solvent molecules that have to be displaced by the solute[42, 199]. This implies that, as shown by the example of the phenyl and the amino nitrogen binding, for two molecular groups with similar surface attraction but different excluded volume, the smaller molecule will generally have a stronger surface attraction. It has become apparent, however, that taking these competitive effects alone does not provide the full description of the system, because it does not properly account for solute hydration and the surface hydrophilicity[42]. This is the third factor that should be taken into account: depending on surface hydrophilicity, adsorbed water layers may exist close to the surface. Intrusion of these mutually hydrogen bonded layers causes energy barriers for surface approach. Therefore, explicit solvent (atomistic) simulations are necessary, using timescales long enough to allow for solvent rearrangements. The fourth factor concerns the geometry and orientation. Especially for longer molecules like polypeptides, many conformations exist next to the surface and a correct sampling has to be performed to find all possible surface-interacting conformations. Due to solvent effects, the outcome may well be non-trivial.

The multiscale simulation approach presented here combines configurational and chemical information needed for engineering surface-interacting peptides. The simulations provide insights in mechanisms of surface interactions in hydrated systems, and can therefore directly be applied to support and explain observations in experimental studies. Obviously, the approach described here requires extension. QM calculations of the peptide group (CONH) and water interacting with Pt surfaces are currently being performed by us. Future work on additional modeling of amino acid residue interactions with this surface will result in a molecular construction set opening the way to the modeling of a variety of peptide-surface systems.

Amino acid interactions with the nickel surface modeled in this work are trans-

ferable to surfaces of different chemical composition. The surface interaction energies of benzene and water can be ordered in a series with comparable energies for Ni, Pt, Pd, and Rh[42]. Therefore, the mechanisms described here for nickel are likely to be similar on Pt, Pd and Rh. Several preliminary generalizations comparing the current modeling with experimental results can therefore be made. First of all, these simulations explain why phenylalanine, albeit QM density functional calculations indicate that it binds strongly to a metal surface (i.e. Ni, Pt, Pd), is not found within the strong metal binders experimentally (see figures 2.1 and 2.2) [23, 38]. The strongest binding configuration of a phenylalanine molecule in vacuo (the interacting aromatic ring oriented parallel to the surface) is shown by us to correspond with a highly unfavorable free energy in the presence of solvent. Although, in these simulations, phenylalanine is a better binder than alanine (contrary to experimental findings on oligo-peptides[23, 38]), one has to take into account that this was observed for the single amino acid; we will at a later point extend the simulations to peptide chains. Tyrosine is experimentally found to be a relatively strong binder among the uncharged amino acids[23, 38]. Because the simulations excluded the planar-ring conformation to bind to a hydrated surface, the most reasonable explanation will be an interaction of the phenol hydroxyl group in tyrosine with the metal surface. Due to electronic polarization effects, surface defects will contribute to this interaction. Here it should also be kept in mind that in most experiments polycrystalline noble metals are used,[23, 38] hence interactions with alternative crystal planes and surface defects require attention in future calculations. QM density functional calculations of the adsorption of water[44] and benzene[186] on metal surface defects have already been performed recently.

Summary and Outlook

A multiscale methodology for biomolecule-metal surface interactions under aqueous ambient conditions, based on an interplay of classical molecular dynamics simulation with quantum calculations, has been introduced. Aside from the advantage of large system size and long sampling times attainable in molecular dynamics simulation, it is shown that free energy methodologies can relate microscopic effects as measurable in simulations to macroscopic effects that can be measured experimentally. As a starting point, a validation of the benzene force field used was performed by calculating benzene excess chemical potential, excess partial molar enthalpy and entropy, as well as the hydration heat capacity change. A good overall agreement with the corresponding experimental values was found. A more in-depth study was then performed, comparing the hydration structure and thermodynamics between a realistic benzene model capable of forming weak hydrogen bonds with water and a non-realistic (van der Waals) benzene model incapable of accepting H-bonds. Calculations were performed using an all-atom benzene model taken from the GROMOS43a1 force field and the SPC water model. The hydration heat capacity of the van-der-Waals model was found to be larger than that of the realistic model. This is as expected from classical hydrophobic hydration models, where the water surrounding a hydrophobic solute will form a network that will minimize the loss of hydrogen bonds, and therefore closely relates to pure water when it comes to the amount of enthalpy that can be adsorbed by breaking hydrogen bonds.

The GROMOS43a1 model for hydrated, realistic, benzene was then used as an initial solute for the study of biomolecular interaction with metal surface in aqueous environments. This first required the formulations of a correct description of the water-metal interaction. As the simulations done should represent ambient conditions, special care was taken to make sure water described a liquid-like state both in the quantum calculations and in the atomistic parameterization. Important considerations for this were the ability to move freely at the surface and form instantaneous

tetrahedrons as in the bulk environment.

In quantum calculations there is no unambiguous way to separate binding energies of parts of the system of interest. For the case of metal surface adsorption of clusters of water molecules, this will lead to the problem that the adsorption energy calculated includes both water-metal interactions and water-water hydrogen bonds. This was addressed by considering only water clusters up to a trimer, where a stable reference state in absence of the surface could be found. This way, the direct water-surface adsorption energy could be retrieved and the water-water interactions could therefore be treated fully independently from the water-surface interactions. The hydrogen bonding energies in the modeling were taken into account by the atomistic water model, initially only the 3-site SPC water model. However, to evaluate if the atomistic parameterization on the atomistic water model used is model dependent, a comparison was made with the same quantum-atomistic parameterization performed on five rigid water models that are often used in simulation and parameterization of biological systems. No significant differences for the adsorption behavior were found, even between the 3-site models and the explicitly tetrahedral 5-site model.

The combination of first principles density functional calculations and classical molecular dynamics simulations used in the water-surface modeling was then applied to study adsorption of hydrated benzene and phenol on Ni(111) and benzene on Au(111) surfaces. As the quantum calculations indicated an overlap between benzene π orbitals and delocalized electrons in the metal, and the distance-dependent quantum data showed a steep decrease of the interaction energy within the first few Å above the surface, a Morse potential was chosen to fit the weak bond-like behavior. Repulsive 10-4 potentials on the benzene hydrogens were added to improve the dependency of the benzene-metal interaction on the benzene-surface inclination. interaction of benzene with metal surfaces. Snapshots of benzene, phenol, and hydration water near the surface show that the water layering combined with the geometry of the molecule (a disc-like shape in the case of benzene), form an important contribution to the total surface adsorption process. This mechanism could be elucidated in even more detail by the calculation of surface inclination angle-distance distributions of benzene and phenol. These calculations help to understand the factors influencing thermodynamics and kinetics of adsorption processes, and it is found that water-mediated effects form a significant contribution here. Indeed, by comparing the benzene-metal and water-metal interaction for different metals it is

found that the ratio of the surface binding strengths for solute and solvent will determine if surface adsorption might be found or not. The inclusion of explicit water is therefore a crucial part of the surface interaction modeling.

In the final part of this thesis, the multiscale methodology was extended into an iterative modeling strategy where quantum calculations are used to determine precise interaction data and to verify the atomistic modeling results. On the atomistic level simulations overcome energetic barriers and help finding global minima. After consistency between the atomistic modeling and the quantum calculations has been reached, the fast sampling obtainable with atomistic simulation is used to study adsorption properties of biological molecules to metal surfaces in bulk water. Following this approach, it is possible to model the optimal adsorption energy and orientations as attainable in quantum calculations for interactions of amino acids with metal surfaces. By treating the water-solute and water-surface interactions present in hydrated systems as pairwise additive, fully hydrated systems can be treated using the current methodology. These advancements are essential in approaching a realistic modeling of experimental peptide-surface systems. As an application, the multiscale modeling of hydrated phenol, alanine, phenylalanine and tyrosine has been performed, and several adsorption properties (adsorption energy, free energy, structural information) have been obtained. Despite that no unambiguous order parameter related to the adsorption process for amino acids exists, the various order parameters used here (surface-site distance for various molecular sites) provided internally consistent results.

Additionally, several general conclusions concerning the chemistry of surface adsorption can be drawn from the current study. Most importantly, it is found that quantum-based binding energies alone do not suffice to understand thermodynamic aspects of protein-surface interactions. Instead, one should account for the competing effects of solvent and other adsorbing groups, which can be a source for free energy barriers for surface approach. It is clear how molecular simulation can aid the research of biomolecular adsorption on metals by coupling macroscopic effects as adsorption free energy or adsorption kinetics, to microscopic phenomena, such as the hydration effects shown here for benzene, or the structural and orientational dependent interactions found for biomolecules at hydrated metal surfaces.

A next step will be the extension to describe larger, fully hydrated systems by combining surface potentials of molecular fragments ('building blocks'), and com-

bine these building blocks to construct a realistic (bio)molecule-surface system. This will require additional parameterizations, but also the introduction of new methodologies. Most straightforward will be the parameterizations similar to the ones performed here, to model solute-surface interactions for all amino acid side chains and the backbone. Looking further than the alanine, phenylalanine, and tyrosine presented here, it will be relatively straightforward to model the amino acids with aliphatic side chains as glycine, valine, leucine, isoleucine, as the side chain will not involve specific polar interactions with the surface that require additional modeling. Initial quantum calculations for an amino acid with charged amino and carboxyl end groups have already been performed[72], atomistic modeling schemes for those groups will closely resemble the modeling of the charged-sidechain amino acids aspartic acid, glutamic acid, and lysine. Already with this subset of amino acids it will be possible to study basic surface adsorption effects, such as the experimentally measurable relative up- and down-modulation for a given combination of amino acids in a peptide[23, 38], or one could imagine creating small model-peptides with features of secondary structure (e.g. alpha-helix or beta-sheet) and studying the effect of the surface on the stability of the structure. A model for the peptide bond is still needed though, and this will be the next point of attention in the multiscale modeling project this thesis belongs to. Finally, apart from the readily attainable modeling of the amino acids mentioned above, still various amino acids exist that might be problematic to analyze in quantum calculations and atomistic modeling, as for example the large tryptophan amino acid and histidine, which contains a heterocyclic, aromatic, base.

The amino acid and peptide adsorption parameterizations will likely be done for metals other than nickel. Gold is an option, but apart from the covalent thiol binding, solute-metal interactions were found to be low. More interesting information from a modeling point of view is therefore expected for platinum or palladium surfaces, both of which are known surfaces in experimental work considering proteins immobilized on metals. As experimental conditions often involve multi-crystalline surfaces, parameterization of solute interaction with surfaces different from the (111) metal plane is also wished for. This will require a more elaborate approach to the modeling of the metal surface. Another extension to improve the metal surface modeling will be the introduction of dispersion interactions, about which still only little quantitative data is present.

Finally, more scales will have to be added to the multiscale approach, where also the biopolymer itself will be treated on different levels of detail, since it is impossible to treat large scale conformational adjustments of the molecules close to the surface on an all-atom basis. This will allow computer simulations for the size ranges at which all the main interactions in biopolymer/surface systems take place, and direct comparisons between experiments and simulations will be possible.

List of Publications

- P. Schravendijk, L. M. Ghiringhelli, L. Delle Site, N. F. A. van der Vegt.
J. Chem. Phys C, Accepted.
'Interaction of Hydrated Amino Acids with Metal Surfaces: A Multiscale Modeling Description'.
- L. M. Ghiringhelli, P. Schravendijk, L. Delle Site.
Phys. Rev. B, 74:035437, 2006.
'Adsorption of Alanine onto Ni(111) Surface: a Multiscale Modelling Oriented Density Functional Study'.
- P. Schravendijk, N. van der Vegt, L. Delle Site and K. Kremer.
ChemPhysChem, 6:1866, 2005.
'Dual Scale Modeling of Benzene Adsorption onto Ni(111) and Au(111) Surfaces in Explicit Water'.
- P. Schravendijk and N. F. A. van der Vegt.
J. Chem. Theory & Comp., 1:643, 2005.
'From Hydrophobic to Hydrophilic Solvation: An Application to Hydration of Benzene'.

Bibliography

- [1] P. Nielaba, M. Mareschal, and G. Ciccotti, editors. *Bridging time scales: Molecular Simulations for the next decade, Lecture Notes in Physics, 605*. Springer-Verlag, 2002.
- [2] S. Attinger and P. Koumoutsakos, editors. *Multiscale Modelling and Simulation, Lecture Notes in Computational Science and Engineering, 39*. Springer-Verlag, 2004.
- [3] M. Praprotnik, L. Delle Site, and K. Kremer. *J. Chem. Phys.*, 123:224106, 2005.
- [4] M. Praprotnik, L. Delle Site, and K. Kremer. *Phys. Rev. E*, 73:066701, 2006.
- [5] M. Praprotnik, K. Kremer, and L. Delle Site. *Phys. Rev. E*, 75:017701, 2007.
- [6] L. Delle Site, C. F. Abrams, A. Alavi, and K. Kremer. *Phys. Rev. Lett.*, 89:156103, 2002.
- [7] C. F. Abrams, L. Delle Site, and K. Kremer. *Phys. Rev. E*, 67:021807, 2003.
- [8] L. Delle Site, S. Leon, and K. Kremer. *J. Am. Chem. Soc.*, 126:2944, 2004.
- [9] J. C. Love, L. A. Estroff, J. K. Kriebel, R. G. Nuzzo, and G. M. Whitesides. *Chem. Rev.*, 105:1103–1169, 2005.
- [10] K.-I. Sano, H. Sasaki, and K. Shiba. *Langmuir*, 21:3090, 2005.
- [11] W. Knoll. *Annu. Rev. Phys. Chem.*, 49:569, 1998.
- [12] J. Homola, S. S. Yee, and G. Gauglitz. *Sensors and Actuators B*, 54:3, 1999.
- [13] I. Willner and E. Katz. *Angew. Chem. Int. Ed.*, 39:1180, 2000.

- [14] C. Sanchez, H. Arribart, and M. M. Giraud Guille. *Nature Materials*, 4:277, 2005.
- [15] Y. Astier, H. Bayley, and S. Howorka. *Curr. Opin. Chem. Biol.*, 9:576, 2005.
- [16] M. Sarikaya, C. Tamerler, A. K.Y. Jen, K. Schulten, and F. Baneyx. *Nature Materials*, 2:577, 2003.
- [17] A. Bietsch, M. Hegner, H. P. Lang, and C. Gerber. *Langmuir*, 20:5119–5122, 2004.
- [18] T. Okamoto, T. Suzuki, and N. Yamamoto. *Nature Biotechnology*, 18:438, 2000.
- [19] J. J. Gray. *Curr. Opin. Struc. Biol.*, 14:110, 2004.
- [20] J. H. Harding and D. M. Duffy. *J. Mater. Chem.*, 16:1105, 2006.
- [21] R. J. Green, R. A. Frazier, K. M. Shakesheff, M. C. Davies, C. J. Roberts, and S. J.B. Tendler. *Biomaterials*, 21:1823, 2000.
- [22] F. Xu, G. Zhen, E. Kuennemann, M. Textor, and W. Knoll. *J. Am. Chem. Soc.*, 127:13084, 2005.
- [23] R. L. Willet, K. W. Baldwin, K. W. West, and L. N. Pfeiffer. *Proc. Natl. Acad. Sci. USA*, 102:7817–7822, 2005.
- [24] C. M. Halliwell. *Analyst*, 129:1166, 2004.
- [25] H. Elwing. *Biomaterials*, 19:397, 1998.
- [26] R. Levicky, T. M. Herne, M. J. Tarlov, and S. K. Satija. *J. Am. Chem. Soc.*, 120:9787, 1998.
- [27] N. Backmann, C. Zahnd, F. Huber, A. Bietsch, A. Plückthun, H.-P. Lang, H.-J. Güntherodt, M. Hegner, and C. Gerber. *Proc. Natl. Acad. Sci USA*, 102:14587, 2005.
- [28] W. M. Mullett, E. P. C. Lai, and J. M. Yeung. *Methods*, 22:77, 2006.
- [29] R. W. Nelson, D. Nedelkov, and K. A. Tubbs. *Electrophoresis*, 21:1155, 2000.

- [30] R. Bashir. *Adv. Drug Deliv. Rev.*, 56:1565, 2004.
- [31] J.-U. Meyer. *Sensors and Actuators A*, 97-98:1, 2002.
- [32] N. Turner, M. Armitage, R. Butler, and G. Ireland. *Cell. Biol. Int*, 28:541–547, 2004.
- [33] S. Brown. *Nature Biotechnol.*, 15:269, 1997.
- [34] T. E. Jones and C. J. Baddeley. *Langmuir*, 21:9468, 2005.
- [35] T. Hayashi, K.-I. Sano, K. Shiba, Y. Kumashiro, K. Iwahori, I. Yamashita, and M. Hara. *Nano Lett.*, 6:515, 2006.
- [36] K. Goede, P. Busch, and M. Grundmann. *Nano Letters*, 4:2115, 2004.
- [37] S. R. Whaley, D. S. English, E. L. Hu, P. F. Barbara, and A. M. Belcher. *Nature*, 405:665, 2000.
- [38] B. R. Pelle, E. M. Krauland, K. D. Wittrup, and A. M. Belcher. *Langmuir*, 21:6929, 2005.
- [39] M. Sarikaya, C. Tamerler, D. T. Schwartz, and F. Baneyx. *Annu. Rev. Mater. Res.*, 34:373, 2004.
- [40] M. A. Strehle, P. Rösch, R. Petry, A. Hauck, R. Thull, W. Kiefer, and J. Popp. *Phys. Chem. Chem. Phys.*, 6:5232, 2004.
- [41] G. Schmalz, U. Schuster, and H. Schweikl. *Biomaterials*, 19:1689–1694, 1998.
- [42] P. Schravendijk, N. van der Vegt, L. Delle Site, and K. Kremer. *ChemPhysChem*, 6:1866, 2005.
- [43] L. Delle Site and K. Kremer. *Int. J. Quant. Chem.*, 101:733, 2005.
- [44] D. Sebastiani and L. Delle Site. *J. Chem. Theory Comput.*, 1:78, 2005.
- [45] T. Murakhtina, L. Delle Site, and D. Sebastiani. *ChemPhysChem*, 7:1215, 2006.
- [46] V. P. Zhdanov and B. Kasemo. *Langmuir*, 17:5407, 2001.

- [47] P. M. Biesheuvel, M. van der Veen, and W. Norde. *J. Phys. Chem. B*, 109:4172, 2005.
- [48] S. R. Euston and Md. A. Nasr. *Langmuir*, 21:4227, 2005.
- [49] M. Bachmann and W. Janke. *Phys. Rev. Lett.*, 95:058102, 2005.
- [50] M. Lund, T. Åkesson, and B. Jönsson. *Langmuir*, 21:8385, 2005.
- [51] M. Friedel, A. Baumketner, and J-E Shea. *Proc. Natl. Acad. Sci. USA*, 103:8396, 2006.
- [52] T. A. Knotts IV, N. Rathore, and J. J. de Pablo. *Proteins*, 61:385, 2005.
- [53] N. Gupta and A. Irbäck. *J. Chem. Phys.*, 120:3983, 2004.
- [54] J. Zhou, H.-K. Tsao, S. Chen, and S. Jiang. *J. Chem. Phys.*, 121:1050, 2004.
- [55] J. Zhou, S. Chen, and S. Jiang. *Langmuir*, 19:3472, 2003.
- [56] Y. Dai and J. S. Evans. *J. Chem. Phys.*, 112:5144, 2000.
- [57] D. Song and D. Forciniti. *J. Chem. Phys.*, 115:8089, 2001.
- [58] A. Mungikar and D. Forciniti. *Biomacromolecules*, 7:239, 2006.
- [59] G. Raffaini and F. Ganazzoli. *Phys. Chem. Chem. Phys.*, 8:2765, 2006.
- [60] F. Ganazzoli and G. Raffaini. *Phys. Chem. Chem. Phys.*, 7:3651, 2005.
- [61] G. Raffaini and F. Ganazzoli. *Langmuir*, 19:3403, 2003.
- [62] J. Zheng, L. Li, H-K Tsao, Y-J Sheng, S. Chen, and S Jiang. *Biophys. J.*, 89:158, 2005.
- [63] V. P. Raut, M. A. Agashe, S. J. Stuart, and R. A. Latour. *Langmuir*, 21:1629, 2005.
- [64] Y. Dou, N. Winograd, B. J. Garrison, and L. V. Zhigilei. *J. Phys. Chem. B*, 107:2362, 2003.
- [65] N. Kantarci, C. Tamerler, M. Sarikaya, T. Haliloglu, and P. Doruker. *Polymer*, 46:4307, 2005.

- [66] E. E. Oren, C. Tamerler, and M. Sarikaya. *Nano Lett.*, 5:415, 2005.
- [67] L. Delle Site, A. Alavi, and C. F. Abrams. *Phys. Rev. B*, 67:193406, 2003.
- [68] R. Di Felice, A. Seloni, and E. Molinari. *J. Phys. Chem. B*, 107:1151, 2003.
- [69] R. Di Felice and A. Selloni. *J. Chem. Phys.*, 120:4906, 2004.
- [70] D. I. Sayago, M. Polcik, C. L. A. Lamont, and D. P. Woodruff. *Surf. Sci.*, 590:76, 2005.
- [71] M. Nyberg, J. Hasselström, O. Karis, N. Wassdahl, M. Weinelt, A. Nilsson, and L. G. M. Pettersson. *J. Chem. Phys.*, 112:5420, 2000.
- [72] L. Ghiringhelli, P. Schravendijk, and L. Delle Site. *Phys. Rev. B*, 74:035437, 2006.
- [73] W. Langel and L. Menken. *Surf. Sci.*, 538:1, 2003.
- [74] C. Boehme and D. Marx. *J. Am. Chem. Soc.*, 125:13362, 2003.
- [75] S. Pulkkinen, M. Noguera, L. Rodríguez-Santiago, M. Sodupe, and J. Bertran. *Chem. Eur. J.*, 6:4393, 2000.
- [76] J. Gao and M. Freindorf. *J. Phys. Chem. A*, 101:3182, 1997.
- [77] S. Yang and M. Cho. *J. Chem. Phys.*, 123:134503, 2005.
- [78] P. L. Polavarapu, Z. Deng, and C. S. Ewig. *J. Phys. Chem.*, 98:9919, 1994.
- [79] G. Cuevas, V. Renugopalakrishnan, G. Madrid, and A. T. Hagler. *Phys. Chem. Chem. Phys.*, 4:1490, 2002.
- [80] H.-A. Yu, M. Pettitt, and M. Karplus. *J. Am. Chem. Soc.*, 113:2425, 1991.
- [81] Y. A. Mantz, H. Gerard, R. Iftimie, and G. J. Martyna. *J. Am. Chem. Soc.*, 126:4080, 2004.
- [82] A. G. Martínez, E. T. Vilar, A. G. Fraile, and P. Martínez-Ruiz. *J. Phys. Chem. A*, 106:4942, 2002.
- [83] N. G. Mirkin and S. Krimm. *J. Am. Chem. Soc.*, 113:9742, 1991.

- [84] S. K. Gregurick, G. M. Chaban, and R. B. Gerber. *J. Phys. Chem. A.*, 106:8696, 2002.
- [85] V. Carravetta and S. Monti. *J. Phys. Chem. B*, 110:6160, 2006.
- [86] M. P. Allen and D. J. Tildesley. *Computer Simulation of Liquids*. Oxford University Press, New York, 1990.
- [87] D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Academic Press, 2002.
- [88] W. F. van Gunsteren, S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott, and I. G. Tironi. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. vdf Hochschulverlag AG an der ETH Zürich, 1996.
- [89] T. N. Heinz, W. F. van Gunsteren, and P. H. Hünenberger. *J. Chem. Phys.*, 115:1125, 2001.
- [90] I. Bitsanis and G. Hadziioannou. *J. Chem. Phys.*, 92(6):3827, 1990.
- [91] P. M. Morse. *Phys. Rev.*, 34:57, 1929.
- [92] H. J. C. Berendsen, D. van der Spoel, and R. van Drunen. *Comput. Phys. Commun.*, 91:43, 1995.
- [93] E. Lindahl, B. Hess, and D. van der Spoel. *J. Mol. Model.*, 7:306–317, 2001.
- [94] D. van der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, and H. J. C. Berendsen. *J. Comput. Chem.*, 26:1701, 2005.
- [95] J. P. Ryckaert, G. Ciccotti, and H. J. C. Berendsen. *J. Comput. Phys.*, 23:327, 1977.
- [96] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. *J. Comput. Chem.*, 18:1463, 1997.
- [97] L. D. Schuler, X. Daura, and W. F. van Gunsteren. *J. Comput. Chem.*, 22:1205, 2001.
- [98] C. Oostenbrink, A. Villa, A. E. Mark, and W. F. van Gunsteren. *J. Comput. Chem.*, 25:1656, 2004.

- [99] S. Nosé. *Mol. Phys.*, 52:255, 1984.
- [100] T. Soddemann, B. Dünweg, and K. Kremer. *Phys. Rev. E*, 68:046702, 2003.
- [101] R. D. Groot and P. B. Warren. *J. Chem. Phys.*, 107:4423, 1997.
- [102] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. *J. Chem. Phys.*, 81:3684, 1984.
- [103] M. Parrinello and A. Rahman. *J. Appl. Phys*, 52:7182, 1981.
- [104] H. C. Andersen. *J. Chem. Phys.*, 72:2384, 1980.
- [105] I. G. Tironi, R. Sperb, P. E. Smith, and W. F. van Gunsteren. *J. Chem. Phys.*, 102:5451, 1995.
- [106] D. van der Spoel and P. J. van Maaren. *J. Chem. Theory Comput.*, 2:1, 2006.
- [107] P. P. Ewald. *Ann. Rev.*, 64:253, 1921.
- [108] T. Darden, D. York, and L Pedersen. *J. Chem. Phys*, 98:10089, 1993.
- [109] M. Deserno and C. Holm. *J. Chem. Phys*, 109:7678, 1998.
- [110] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. McGraw-Hill, New York, 1989.
- [111] P. Hohenberg and W. Kohn. *Phys. Rev.*, 136:864, 1964.
- [112] W. Kohn and L. J. Sham. *Phys. Rev.*, 140:1133, 1965.
- [113] P. J. Feibelman. *Phys. Rev. B.*, 72:113405, 2005.
- [114] J. M. Gottfried, E. K. Vestergaard, P. Bera, and C. T. Campbell. *J. Phys. Chem. B*, 110:17539, 2006.
- [115] F. Müller-Plathe. *ChemPhysChem.*, 3:754, 2002.
- [116] K. Kremer. *Macromol. Chem. Phys.*, 204:257, 2003.
- [117] L. Delle Site, S. Leon, and K. Kremer. *J. Phys. Condens. Matter*, 17:L53, 2005.
- [118] B. Widom. *J. Chem. Phys.*, 39:2808, 1963.

- [119] A. Ben-Naim and Y. Marcus. *J. Chem. Phys.*, 81:2016, 1984.
- [120] H.-A. Yu and M. Karplus. *J. Chem. Phys.*, 89:2366, 1988.
- [121] I. C. Sanchez, T. M. Truskett, and P. J. In 't Veld. *J. Phys. Chem. B.*, 103:5106, 1999.
- [122] B. Widom. *J. Phys. Chem.*, 86:869, 1982.
- [123] N. Matubayasi, L. H. Reed, and R. M. Levy. *J. Phys. Chem.*, 98:10640, 1994.
- [124] R. M. Levy and E. Gallicchio. *Annu. Rev. Phys. Chem.*, 49:531, 1998.
- [125] K. P. Murphy, P. L. Privalov, and S. J. Gill. *Science*, 247:559, 1990.
- [126] G. Hummer, S. Garde, A. García, M. E. Paulaitis, and L. R. Pratt. *J. Phys. Chem. B*, 102:10469, 1998.
- [127] K. A. Dill. *Biochemistry*, 29:7133, 1990.
- [128] G. Hummer, S. Garde, A. García, and M. E. Paulaitis. *Proc. Natl. Acad. Sci. U.S.A.*, 95:1552, 1998.
- [129] N. T. Southall, K. A. Dill, and A. D. Haymet. *J. Phys. Chem. B*, 106:521, 2002.
- [130] H. S. Frank and M. E. Evans. *Chem. Phys.*, 13:507, 1945.
- [131] G. I. Makhatadze and P. L. Privalov. *Biophys. Chem.*, 50:285, 1994.
- [132] C. McAuliffe. *J. Phys. Chem*, 70:1267, 1966.
- [133] G. I. Makhatadze and P. L. Privalov. *J. Solution Chem.*, 18:1267, 1989.
- [134] G. I. Makhatadze and P. L. Privalov. *J. Mol. Biol.*, 213:375, 1990.
- [135] M. Levitt and M. F. Perutz. *J. Mol. Biol.*, 201:751, 1988.
- [136] P. Linse. *J. Am. Chem. Soc.*, 112:1744, 1990.
- [137] J. L. Atwood, F. Hamada, K. D. Robinson, G. W. Orr, and R. L. Vincent. *Nature*, 349:683, 1991.

-
- [138] S. Suzuki, P. G. Green, R. E. Bumgarner, S. DAsgupta, W. A. Goddard III, and G. A. Blake. *Science*, 257:942, 1992.
- [139] W. L. Jorgensen and D. L. Severance. *J. Am. Chem. Soc.*, 112:4768, 1990.
- [140] G. Graziano and B. Lee. *J. Phys. Chem. B*, 105:10367, 2001.
- [141] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. In B. Pullman, editor, *Intermolecular Forces*, pages 331–338. Reidel, Dordrecht, 1981.
- [142] B. Lee. *Biopolymers*, 24:813, 1985.
- [143] C. L. Brooks III, M. Karplus, and B. M. Pettitt. In I. Prigogine and S.A. Rice, editors, *Advances in Chemical Physics*, volume 71. 1988.
- [144] B. Hess. *J. Chem. Phys.*, 116:209, 2002.
- [145] G. Karlström, P. Linse, A. Wallqvist, and B. Jönsson. *J. Am. Chem. Soc.*, 106:3777, 1983.
- [146] P. Linse, G. Karlström, and B. Jönsson. *J. Am. Chem. Soc.*, 106:4096, 1984.
- [147] J. P. M. Postma, H. C. Berendsen, and J. R. Haak. *Faraday Symp. Chem. Soc.*, 17:55, 1982.
- [148] E. Gallicchio, M. M. Kubo, and R. M. Levy. *J. Phys. Chem. B*, 104:6271, 2000.
- [149] S. J. Gill, S. F. Dec, G. Olofsson, and I. Wadsö. *J. Phys. Chem.*, 89:3758, 1985.
- [150] M. T. Stone, P. J. In 't Veld, Y. Lu, and I. C. Sanchez. *Mol. Phys.*, 100:2773, 2002.
- [151] P. A. Thiel and T. E. Madey. *Surf. Sci. Rep.*, 7:211, 1987.
- [152] M. A. Henderson. *Surf. Sci. Rep.*, 46:1–308, 2002.
- [153] A. Verdaguer, G. M. Sacha, H. Bluhm, and M. Salmeron. *Chem. Rev.*, 106:1478, 2006.

- [154] R. Guidelli and W. Schmickler. *Electrochim. Acta*, 45:2317, 2000.
- [155] L. J. Michot, F. Viliéras, M. François, I. Bihannic, M. Pelletier, and J-M Cases. *Geoscience*, 334:611, 2002.
- [156] C. D. Taylor and M. Neurock. *Curr. Opin. Solid. St. M.*, 9:49, 2005.
- [157] M. Morgenstern, T. Michely, and G. Comsa. *Phys. Rev. Lett.*, 77:703, 1996.
- [158] A. Glebov, A. P. Graham, A. Menzel, and J. P. Toennies. *J. Chem. Phys.*, 106:9382, 1997.
- [159] G. Held and D. Menzel. *Phys. Rev. Lett.*, 74:4221, 1995.
- [160] M. Foster, M. D'Agostino, and D. Passno. *Surf. Sci.*, 590:31, 2005.
- [161] J. G. Catalano, C. Park, Z. Zhang, and P. Fenter. *Langmuir*, 22:4668, 2006.
- [162] S. Engemann, H. Reichert, H. Dosch, J. Bilgram, V. Honkimäki, and A. Snigirev. *Phys. Rev. Lett.*, 92:205701, 2004.
- [163] M. Hugelmann and W. Schindler. *Surf. Sci. Lett.*, 541:643, 2003.
- [164] S. Meng, L. F. Xu, E. G. Wang, and S. Gao. *Phys. Rev. Lett.*, 89:176104, 2002.
- [165] A. Michaelides, A. Alavi, and D. A. King. *Phys. Rev. B*, 69:113404, 2004.
- [166] E. Spohr. *Electrochim. Acta*, 49:23, 2003.
- [167] S. H. Lee and P. J. Rossky. *J. Chem. Phys.*, 100:3334, 1993.
- [168] Y. Dou, L. V. Zhigilei, N. Winograd, and B. J. Garrison. *J. Phys. Chem. A*, 105:2748–2755, 2001.
- [169] E. Spohr, G. Tóth, and K. Heinziger. *Electrochim. Acta*, 41(14):2131, 1996.
- [170] S.-B. Zhu and M. R. Philpott. *J. Chem. Phys.*, 100:6961, 1993.
- [171] J. C. Shelley, G. N. Patey, D. R. Bérard, and G. M. Torrie. *J. Chem. Phys.*, 107:2122, 1997.
- [172] A. Kohlmeyer, W. Witschel, and E. Spohr. *Chem. Phys.*, 213:211, 1996.

-
- [173] P. Vassilev, R. A. van Santen, and M. T. M. Koper. *J. Chem. Phys.*, 122:054701, 2005.
- [174] S. Meng, E. G. Wang, and S. Gao. *Phys. Rev. B*, 69:195404, 2004.
- [175] D. Marx. *Science*, 303:634, 2004.
- [176] S. Nosé and M. L. Klein. *Mol. Phys.*, 50:1055, 1983.
- [177] W.G. Hoover. *Phys. Rev. A.*, 31:1695, 1981.
- [178] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. *J. Phys. Chem.*, 91:6269, 1987.
- [179] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. *J. Chem. Phys.*, 79:926, 1983.
- [180] M. W. Mahoney and W. L. Jorgensen. *J. Chem. Phys.*, 112:8910, 2000.
- [181] E. Spohr. *J. Mol. Liquids*, 64:91, 1995.
- [182] M. R. Philpott and J. N. Glosli. *Chem. Phys.*, 198:53–61, 1995.
- [183] J. I. Siepmann and M. Sprik. *J. Chem. Phys.*, 102:511–524, 1995.
- [184] H. Sakuma, T. Tsuchiya, K. Kawamura, and K. Otsuki. *Surf. Sci.*, 536:L396, 2003.
- [185] S. D. Chakarova-Käck, Ø. Borck, E. Schröder, and B. I. Lundqvist. *Phys. Rev. B*, 74:155402, 2006.
- [186] L. Delle Site and D. Sebastiani. *Phys. Rev. B*, 70:115401, 2004.
- [187] A. Alavi, J. Kohanoff, M. Parrinello, and D. Frenkel. *Phys. Rev. Lett.*, 73:2599, 1994.
- [188] A. Alavi. In *Monte Carlo and Molecular Dynamics of Condensed Matter Systems*, chapter 25, page 649. Italian Physical Society, Bologna, 1996.
- [189] J. Hutter, A. Alavi, T. Deutsch, M. Bernasconi, S. Goedecker, M. Tuckerman, and M. Parrinello. *CPMD v. 3.4.1*, 1995-1999.
- [190] F. Mittendofer and J. Hafner. *Surf. Sci.*, 472:133, 2001.

- [191] S. Yamagishi, S. J. Jenkins, and D. A. King. *J. Chem. Phys.*, 110:5765, 2001.
- [192] N. F. A. van der Vegt, D. Trzesniak, B. Kasumaj, and W. F. van Gunsteren. *ChemPhysChem*, 5:144, 2004.
- [193] N. F. A. van der Vegt and W. F. van Gunsteren. *J. Phys. Chem. B*, 108:1056, 2004.
- [194] P. Schravendijk and N. van der Vegt. *J. Chem. Theory Comp.*, 1:643, 2005.
- [195] A. Michaelides, V. A. Ranea, P. L. de Andres, and D. A. King. *Phys. Rev. Lett.*, 90(21):216102, 2003.
- [196] C. Morin, D. Simon, and P. Sautet. *J. Phys. Chem. B*, 108:5653, 2004.
- [197] I. R. McDonald and J. C. Rasaiah. *Chem. Phys. Lett.*, 34:382, 1975.
- [198] W. L. Jorgensen and J. Tirado-Rives. *J. Am. Chem. Soc.*, 110:1657, 1988.
- [199] T. G. Cooper and N. H. de Leeuw. *Langmuir*, 20:3984, 2004.