Modeling and simulation of protein–surface interactions: achievements and challenges

Musa Ozboyaci1,2*, Daria B. Kokh1, Stefano Corni3 and Rebecca C. Wade1,4,5*

1 Heidelberg Institute for Theoretical Studies (HITS), Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany
2 Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), Heidelberg University, Im Neuenheimer Feld 368, 69120 Heidelberg, Germany
3 Centro S3, CNR Instituto Nanoscienze, via Campi 213/a, 41125 Modena, Italy
4 Zentrum für Molekulare Biologie der Universität Heidelberg, DKFZ-ZMBH Allianz, Im Neuenheimer Feld 282, 69120 Heidelberg, Germany
5 Interdisciplinary Center for Scientific Computing (IWR), Heidelberg University, 69120 Heidelberg, Germany

Quarterly Reviews of Biophysics (2016), 49, e4, page 1 of 45 doi:10.1017/S0033583515000256

Abstract. Understanding protein–inorganic surface interactions is central to the rational design of new tools in biomaterial sciences, nanobiotechnology and nanomedicine. Although a significant amount of experimental research on protein adsorption onto solid substrates has been reported, many aspects of the recognition and interaction mechanisms of biomolecules and inorganic surfaces are still unclear. Theoretical modeling and simulations provide complementary approaches for experimental studies, and they have been applied for exploring protein–surface binding mechanisms, the determinants of binding specificity towards different surfaces, as well as the thermodynamics and kinetics of adsorption. Although the general computational approaches employed to study the dynamics of proteins and materials are similar, the models and force-fields (FFs) used for describing the physical properties and interactions of material surfaces and biological molecules differ. In particular, FF and water models designed for use in biomolecular simulations are often not directly transferable to surface simulations and vice versa. The adsorption events span a wide range of time- and length-scales that vary from nanoseconds to days, and from nanometers to micrometers, respectively, rendering the use of multi-scale approaches unavoidable. Further, changes in the atomic structure of material surfaces that can lead to surface reconstruction, and in the structure of proteins that can result in complete denaturation of the adsorbed molecules, can create many intermediate structural and energetic states that complicate sampling. In this review, we address the challenges posed to theoretical and computational methods in achieving accurate descriptions of the physical, chemical and mechanical properties of protein-surface systems. In this context, we discuss the applicability of different modeling and simulation techniques ranging from quantum mechanics through all-atom molecular mechanics to coarse-grained approaches. We examine uses of different sampling methods, as well as free energy calculations. Furthermore, we review computational studies of protein–surface interactions and discuss the successes and limitations of current approaches.

Key words: Biomolecular adsorption, protein-solid state interactions, bio–inorganic interface, molecular simulation, molecular modeling.

1. Introduction

2. Which types of surfaces can be modeled?
   2.1. Elemental metals and alloys
   2.2. Oxides and minerals
1. Introduction

Protein–inorganic surface interactions have gained increasing attention owing to their widespread occurrence in nature, and their broad range of applications in nanobiotechnology (Choi et al. 2009; Hill et al. 2007; Hu et al. 2005; Jackson et al. 2000; Laera et al. 2011; Manecka et al. 2014; Millo et al. 2009; Park et al. 2008; Qin et al. 2007a, b; Slocik et al. 2011; Xu et al. 2010). Adhesion of proteins on solid substrates is utilized by many organisms, e.g. sea urchins make use of the adsorption of matrix proteins to specific crystal patches on endoskeletal calcite surfaces (Wilt, 1999), and has even been evolutionarily adapted to enable some organisms to live in specific habitats, e.g. for the adhesion of mussels to mineral rocks (Yu et al. 2011). Humans have long used inorganic materials that make direct interactions with proteins. For example, gold crowns as dental prosthetics date back to the ancient Etruscan civilization (Demann et al. 2005), and man-made nanoparticles were used as pigments in ointments by the ancient Romans (Casals et al. 2008). However, it is only quite recently that advances in science and technology have enabled the production of completely new or engineered surfaces and hence allowed new applications. For example, the remarkable structural and mechanical properties of graphene, isolated in 2004 (Novoselov et al. 2004), have drawn increasing attention to carbon allotropes and catalyzed further research into their interactions with proteins, motivated by various biotechnological applications, including efficient biosensors (Alava et al. 2013).
Computational modeling and simulation of biomolecules can help scientists to unravel the mechanisms of molecular-level events and predict the behavior of complex systems at a level of detail that cannot be directly measured in experiments. Since the development of the first modeling and simulation methods for complex molecules, computational research in the field has expanded enormously. Their importance is shown by the Nobel Prize in Chemistry being awarded in 2013 to Martin Karplus, Michael Levitt and Arieh Warshel “for the development of multiscale models for complex chemical systems”. Proteins, nucleic acids, lipids and their interactions in aqueous environments have been widely studied computationally by means of molecular mechanics (MM) force-fields (FFs) (Brooks et al. 1983; Cornell et al. 1995; Oostenbrink et al. 2004) developed and tailored specifically for these types of molecules. However, many of these FFs fall short in reproducing the properties of protein-inorganic surface systems. To alleviate this problem, many useful models (Heinz et al. 2011; Iori & Corni, 2008; Kokh et al. 2010) and FF parameters (De la Torre et al. 2009; Heinz et al. 2013; Iori et al. 2009; Schneider & Colombi Ciacchi, 2010; Wright et al. 2013a) for material surfaces have been introduced that have been designed to be compatible with FFs for biomolecular systems. These FFs are still rather young and their improvement is an area of active research.

A number of reviews addressing different aspects of protein–surface interaction studies have been published previously. Several of these provide a general overview of the adsorption of proteins at solid surfaces (Cohavi et al. 2010; Costa et al. 2013; Horbett & Brash, 1995; Rabe et al. 2011; Qu et al. 2013), whereas others focus on more specific aspects such as the determination of the adsorption kinetics of protein–surface binding by weakly bound mobile precursor states (Garland et al. 2012), and adsorption on various different surface types, such as metallic surfaces (Tomba et al. 2009; Vallee et al. 2010), polymer surfaces (Hahm, 2014; Wei et al. 2014) and protein repellent surfaces (Szott & Horbett, 2011). The physico-chemical properties of nanomaterials, and their applications in medicine, biology and biotechnology, have also been reviewed in several papers (see, e.g. Ansari & Husain, 2012; Dufort et al. 2012; Khlebtsov & Dikman, 2010; Mahmoudi et al. 2011; Mahon et al. 2012; Mandal et al. 2014; Salata, 2004; Saptarshi et al. 2013; Shemetov et al. 2012). Various aspects of the computational methods employed in modeling and simulation of protein–surface interactions are addressed in other reviews, including issues in computational modeling of peptide–surface interactions (Di Felice & Corni, 2011), problems with these simulation techniques (Latour, 2008) and approaches to multiscale modeling of soft matter that are transferable to protein-surface systems (Praprotnik et al. 2008).

This review provides a discussion of the computational models and simulation techniques that have been used in studies of protein–surface interactions. Due to the broad range of models used in these studies, only models of protein–surface interactions based on chemical structures are discussed in this review and, therefore, more abstract models to describe these interactions, such as that developed by Oberle et al. (2015) for the description of competitive adsorption of proteins to nanoparticles, are not discussed further. A brief introduction to various types of material surfaces is provided and some of the properties that need attention from a modeling point of view are discussed. We further give an overview of the different modeling, sampling and free energy calculation techniques employed in recent studies. We discuss the properties that can be computed by these methods and how they can assist and complement experiments. Some of the important findings from applications of these methods are reviewed, and drawbacks and shortcomings of the available techniques are discussed. The paper concludes with a discussion of the general limitations and future directions of the field.

2. Which types of surfaces can be modeled?

The interactions of proteins with inorganic materials are determined not only by the properties of the proteins, but also by the chemical composition, molecular structure, size and shape of the material. Inorganic surfaces possess distinct physico-chemical properties, such as reactivity towards different compounds, material stability and specific adsorption characteristics for different adsorbents. These properties allow different types of surfaces to be employed for different applications, e.g. implantation or chromatography. Understanding the basis of these physico-chemical properties usually requires atomic-level investigation of the surfaces. The types of material surface modeled commonly in computational studies of protein–surface interactions are elemental metals and alloys, metal oxides and minerals, self-assembled monolayers (SAMs), polymers and carbon allotrope surfaces, see Fig. 1.

2.1 Elemental metals and alloys

Protein–metal surface interactions can be studied with experimental techniques such as atomic force microscopy (AFM) (Binnig et al., 1986), surface plasmon resonance (SPR) (Jönsson et al. 1991) and localized SPR (Stuart et al. 2005). Due to their chemical inertness and unique optical properties (Jain et al. 2008), the noble metals, gold and silver, are the most commonly used metals employed as probes or sensors in these techniques. Along with well-known applications of metal surfaces, such as biosensors and implants (Liu et al. 2004), metal surfaces are also used in bioelectronics as electrodes as they allow
controlled exchange of electrons with metalloproteins immobilized on them (Alessandrini et al. 2005; Andolfi et al. 2004). The interactions of proteins with bare metal surfaces have been the subject of many computational studies, which cover a wide range of different elemental and alloy surfaces, such as Cu(100) (Chen & Wang, 2010), Au(111) (Bizzarri, 2006; Hoefling et al. 2011; Siwko & Corni, 2013; Venkat et al. 2007; Zanetti-Polzi et al. 2014), Au(100) (Hagiwara et al. 2009), Au nanoparticle (Todorova et al. 2014), Fe (Zhang et al. 2009b), Ni (Yang & Zhao, 2006), Pd (Coppage et al. 2011), Pt (Kantarci et al. 2005), Ag (Aliaga et al. 2011; Ghosh et al. 2012) and steel (Imamura et al. 2003).

### 2.2 Oxides and minerals

Metal surfaces (excluding noble metals such as gold, platinum and palladium) are oxidized when exposed to water or air, forming metal oxides that are very common in the Earth’s crust. Due to their good mechanical stability, catalytic properties and biocompatibility (Andreeescu et al. 2012), metal oxides and minerals are used in a wide range of applications that include fabrication of biomaterials (Whaley et al. 2000), cellular delivery of drugs and biomolecules (Kievit & Zhang, 2011; Xu et al. 2006), tissue engineering (Shin et al. 2003) and proteomics (Sugiyama et al. 2007). Computational studies to investigate the interactions of oxide and mineral surfaces with proteins or peptides have mostly been carried out for different forms of titanium dioxide, such as rutile and anatase (Carravetta et al. 2009; Kang et al. 2010; Köppen et al. 2008; Monti, 2007; Monti et al. 2008; Sun et al. 2014a; Wu et al. 2013), silicon oxides (Chen et al. 2009a; Nonella & Seeger, 2008; Patwardhan et al. 2012; Rimola et al. 2009; Tosaka et al. 2010), calcite (Wierzbicki et al. 1994) and mica (Kang et al. 2013).

### 2.3 Self-assembled monolayers

SAMs are thin films that coat surfaces by spontaneous adsorption of organic molecules that form ordered molecular assemblies. Typically, in a SAM, molecules are chemisorbed onto a surface substrate through their reactive head groups and are, therefore, very stable. SAMs can be categorized into two groups according to their head-group type: thiol-based and silane-based (Schreiber, 2004). The head-group is attached to a tail group that, by following a fast adsorption phase (seconds),

**Fig. 1.** Simulations of proteins with different types of surface: (a) lysozyme on a polyethylene surface (reprinted with permission from (Wei et al. 2011). Copyright (2011) American Chemical Society), (b) the MRKDV peptide on a bare silver surface (adapted with permission from (Aliaga et al. 2011). Copyright (2011) American Chemical Society), (c) RAD16II on a rutile surface (reprinted with permission from (Monti, 2007). Copyright (2007) American Chemical Society), and (d) NiFe hydrogenase on a SAM surface (reprinted with permission from (Utesch et al. 2013). Copyright (2013) American Chemical Society).
undergoes a slow reorganization (hours), in which interactions with other tail groups increase and the packing is improved (Love et al. 2005). The tail group can be functionalized with small chemical groups or large molecules, such as peptides. These, together with the length of the tail group, allow the physico-chemical interfacial properties, in particular, hydrophobicity and ionization to be adjusted according to the desired application. Alkanethiols are the most common type of SAM. They have the chemical formula S-(CH₂)n-R, where R stands for a functional group, such as -CH₃, -COOH, -NH₂ or -OH.

Bare metals and oxide surfaces are prone to non-specific adsorption of proteins and other organic molecules. These adsorption processes may result in undesirable agglomeration of adsorbates on the surfaces. The self-assembly of organic molecules on a metal surface in a SAM creates a physical barrier between the surface and the adsorbates, acting as an electrical insulator and passivating the surface atoms (Love et al. 2005). The properties of SAMs have been reviewed in (Chaki & Vijaymohan, 2002; Gooding et al. 2003; Love et al. 2005; Senaratne et al. 2005). Modeling of protein–SAM interactions has been reported, mostly for alkanethiol SAMs, in (Alvarez-Paggi et al. 2010; Hsu et al. 2008; O’Mahony et al. 2013; Sun et al. 2005; Utesch et al. 2013; Wang et al. 2010a, b; Xie et al. 2012) and peptide–SAM interactions have been modeled by Nowinski et al. (2012).

2.4 Polymers

Polymer surfaces have attracted much attention in the nanotechnology field due to their mechanical stability, low cost and their wide applicability (Nie & Kumacheva, 2008).

Particularly synthetic polymer-based biomaterials, due to their non-fouling properties, are currently being investigated intensively for applications in controlled drug delivery (Hoffman, 2008), in highly sensitive biosensors (Anker et al. 2008), and in bioelectronics (Senaratne et al. 2005). Nanostructured polymer materials used for bio-related and medicinal research include electrostatic polymer brushes, micelles, layer-by-layer deposition and thin films (Stuart et al. 2010). Polymer brushes are surface modifiers that share many properties in common with SAMs (Senaratne et al. 2005). They are prepared by grafting polymers of the same or varying kinds on surfaces forming homopolymer and mixed brushes, respectively. Polymeric micelles are formed through self-assembly of amphiphilic copolymers and typically have a diameter of size 30–50 nm (Otsuka et al. 2003; Stuart et al. 2010).

These micelles are particularly important due to their lower critical micelle concentration (CMC), higher stability and slower rate of dissociation than surfactant micelles. These properties have allowed polymeric micelles to act as effective cancer treatment tools with high drug deposition at the target site (Otsuka et al. 2003). Polymer surfaces and their applications have been reviewed by (Barbey et al. 2009; Kim et al. 2008; Nie et al. 2010; Otsuka et al. 2003; Senaratne et al. 2005; Stuart et al. 2010). The limited number of computational studies of protein-polymer surfaces to date have been carried out for polymer types including polystyrene, polyethylene and polydimethylsiloxane (Boughton et al. 2010; Jeyachandran et al. 2009; Liu et al. 2012; Lu et al. 1992; O’Brien et al. 2008; Raffaini & Ganazzoli, 2007; Zhang et al. 2009a; Wei et al. 2011).

2.5 Carbon allotropes

Pure carbon may exist in a number of different allotropes. Owing to their unique thermal, electrical, and mechanical properties, carbon-based nanomaterials have been the subject of numerous applications in analytical chemistry (Scida et al. 2011). These applications mostly focus on carbon nanomaterials with sp²-carbon bonding, such as fullerenes, carbon nanotubes (CNT), graphene and graphite. This is due to the extremely high surface areas of fullerenes and CNTs relative to their size, which makes them suitable for design as highly efficient drug carriers. Furthermore, the excellent electrical properties of graphene and CNTs make them suitable for biosensor applications (Liu & Liang, 2012). In all the four materials, the carbon atoms make three chemical bonds with other carbons in the surface-plane with delocalized π electron clouds in the direction perpendicular to the surface (Scida et al. 2011). This configuration makes the mutual van der Waals interactions between CNTs very strong and hence leads them to be very hydrophobic (Guldi et al. 2006). To alter the hydrophobicity, surface modifications with surface defects and polar groups have been suggested, but these affect the stability of the materials as well as their mechanical and electrical properties (Scida et al. 2011). Computational studies of protein–carbon surface interactions have mostly focused on graphene/graphite (Mereghetti & Wade, 2011; Mücksch & Urbassek, 2011; Raffaini & Ganazzoli, 2010; Kang et al. 2013; Sun et al. 2014b; Yu et al. 2012b), CNT (Balamurugan et al. 2010; Chen et al. 2009b; Tallury & Pasquinelli, 2010; Wang et al. 2003) and fullerenes (Durdagi et al. 2008; Kraszewski et al. 2010; Noon et al. 2002).

3 Which surface properties need consideration in modeling?

The modeling of protein-water-solid surface interfaces poses problems because a variety of distinct physical and chemical properties may be associated with different surface types. Factors such as the size and shape of nanoparticles, the crystal
packing of a surface, the presence of surface defects, the density of SAM molecules, the change of the chemical state of the surface, such as protonation, oxidation, or hydroxylation, due to the presence of water and the environmental conditions, such as the presence of surfactants, all affect the interaction of biomolecules and the solid substrate. Therefore, it is very important to choose the level of microscopic details to be included in the computational model carefully when modeling protein–surface interactions, which must adequately describe the physical and chemical properties of the studied system under experimental conditions.

Several important characteristics of the surfaces that should be considered in modeling protein adsorption are discussed in the next sections: ionization and hydration, polarization, surface reconstruction upon binding, as well as the topography and morphology of surfaces and nanoparticles.

### 3.1 Ionization and hydration

SAMs and metal oxide surfaces may be protonated and/or hydroxylated to varying degrees depending on the environmental conditions and on the material itself, e.g. pH, material shape and size. Many theoretical studies on SAMs do not take into account the ionization states of the functional groups of SAMs. However, a recent study (Utesch et al. 2013) reporting a titration curve of an amino-terminated alkanethiol SAM showed that the level of ionization was very sensitive at pH values around 6 (with 16 ± 5% protonation at pH 7 and 52 ± 9% at pH 6). A positive correlation between the ionization level of a SAM and the strength of adsorption of charged proteins was found (Zhou et al. 2004; Utesch et al. 2013). Therefore, it is crucial to compute and model the surface ionization to obtain reliable results compatible with experiments. The protonation states of SAMs in modeling studies are mostly represented by random assignment of protonated and deprotonated groups (Alvarez-Paggi et al. 2013; Utesch et al. 2012; Zhou et al. 2004). In several studies, they are represented either by a uniform distribution of small partial charges (Sun et al. 2005) or by large partial charges being assigned to functional groups in neutral surfaces (Wang et al. 2010b).

As with SAMs, SiO₂ and TiO₂ surfaces have different levels of ionization of the surface groups depending on the pH of the environment, and this has been shown to determine selectivity of the adsorption of proteins (Patwardhan et al. 2012). Studies have shown that the concentration of silanol groups (Si-OH) and the degree of their ionization define the hydrophobicity of silicon dioxide (silica) surfaces and govern their adsorption properties and thus also, the behavior of silica-based materials in processes such as biomolecular adhesion and biominerilization (Sumper & Brunner, 2008; Voskerician et al. 2003). In computational studies of protein–oxide surface interactions, ionization may be represented explicitly (Friedrichs et al. 2013; Köppen & Langel, 2010; Patwardhan et al. 2012; Tosaka et al. 2010), or by assigning uniform partial charges to selected surface atoms (Kubiak-Ossowska & Mulheran, 2012). Köppen & Langel (2010) showed that the adhesion energies of peptides on titanium dioxide surfaces are sensitive to the values of the partial charges of the surface hydroxyl groups. Therefore, care must be taken to ensure a reliable parameterization of the ionization charges.

Conventional simulations of proteins usually neglect changes in the protonation states of ionizable groups as a function of time. However, an accurate simulation of a protein-surface system may require a more sophisticated approach, such as constant pH simulation, to treat the variation of the protonation states of residues in the interfacial region, as well as of the surface interaction sites. Although constant pH simulation techniques have been applied to various molecular systems, including small chemotherapeutic drug molecules binding to nanodiamonds (Adnan et al. 2011), they have not, as far as we are aware, been applied to simulations of peptide/protein–surface interactions.

The most important issue for modeling the adsorption of biomolecules on oxide surfaces is the treatment of the properties of the surface hydration shell. Due to dissociation, the hydroxyl groups and hydrogen atoms form bonds with unsaturated surface metal (e.g. Ti) and O atoms, respectively. The degree of water dissociation defines the physical properties of the surface and strongly affects the binding properties of biological molecules. Kang et al. (2010) investigated the role of the water in the adsorption process on rutile surfaces and observed that human serum albumin (HSA) adsors on a rutile surface modified with -OH groups more strongly than on an unmodified rutile surface. Hydrogen bond analysis in the same study showed that the bonds formed between the structured hydroxyl groups on the modified surface reduced the possibility of hydrogen bond formation between the surface and the water molecules, hence making it easier for the protein to adsorb onto the modified surface.

It should be noted that water dissociation is reversible on all oxides (Henderson, 2002). Therefore, the concentration and position of hydroxyl groups may change with time. Even though relying on the average distribution of hydroxyl groups on a particular surface may be sufficient, it may not always give accurate results, especially if the protein adsorption kinetics are determined by surface hydration kinetics. This poses a problem for conventional modeling techniques for biomolecules, which ideally have to capture the dynamics of dissociative adsorption and associative desorption of water molecules on oxide surfaces.
3.2 Polarization

The electrostatic potential of solutes and solvent molecules induces an attractive polarization of metal surfaces. Metal polarization is negligible for neutral adsorbates that do not have large dipoles and, since proteins usually have a relatively small total charge, it is often neglected in simulations (Braun et al. 2002). Early studies of metal surfaces also showed that the effect of pure water on polarization of the metal is often negligible (Barabino & Marchesi, 1984; Shelley et al. 1997; Spohr, 1995) and that charge-induced polarization does not cause any change in the structure of the water on the surface (Feng et al. 2011). However, for the adsorption of biomolecules with considerable dipole moments, including some peptides and proteins, the surface polarization contributes to the binding energy and influences the binding mode. Although studies have shown that, for water-metal surface systems, the energy due to polarization is less than 10% of the total binding energy (Feng et al. 2011; Neves et al. 2007; Sieppmann & Sprik, 1995; Vila Verde et al. 2009, 2011), for proteins, in particular, on Au (111) surfaces, the contribution from polarization has been estimated to be about 10–20% of the total binding energy (Heinz et al. 2011). Further, on surfaces such as Au(100), where the van der Waals attraction is weaker, polarization was found to tune the adsorption of proteins (Heinz et al. 2009) and act as a major contributor to the adsorption of highly charged peptides (Heinz et al. 2011). For the simulation of amino acid adsorption on Au(111), polarization effects were also found to be important for reproducing experimental binding propensities (Hoefling et al. 2011).

The polarization of surfaces of other types can also play an important role in the interactions of surfaces with their environments. Schyman & Jorgensen (2013) showed that while a non-polarizable FF is adequate for the description of interactions between water and small hydrocarbons, such as benzene (C₆H₆) and coronene (C₂₄H₁₂), a polarizable FF is required for CNTs and fullerenes in order to reproduce interaction energy values obtained from density functional theory (DFT) calculations. Therefore, the effect of polarization should be carefully considered in computer simulations of biomolecules, in particular, of charged proteins, with metal and carbon surfaces.

3.3 Reconstruction

The atomic structure of the surface of a material generally differs from that of its interior because of differences in the forces acting on the atoms in the vicinity of the surface. The type and the degree of reconstruction of a surface is determined by environmental conditions, such as temperature and pressure (Somorjai & Li, 2011), as well as the structure of the material and may be affected by adsorbing molecules. Ideally, reconstruction of surfaces has to be taken into account in the simulation of adsorption (Ghiringhelli et al. 2009): an enormous task in most cases. On the other hand, it has been reported in several studies that the reconstruction of some surfaces in certain conditions is negligible (Feng et al. 2011; Iori et al. 2008; Raffaini & Ganazzoli, 2012; Wright et al. 2013a). However, other experimental and computational studies show that large scale surface reconstruction may take place after adsorption of small molecules (Eralp et al. 2011; Gibbs et al. 1990; Lal et al. 2004, 2006). Therefore, care must be taken in modeling if a major reconstruction of the surface takes place upon adsorption.

3.4 Topography

The topographic characteristics of a surface or nanoparticle at the micro- to nanometer-scale are important determinants of protein adsorption properties, such as binding affinities and surface saturation values (Fenoglio et al. 2011; Gagner et al. 2011, 2012; Roach et al. 2006). The topography of a surface can be characterized by its exposed crystal planes, its roughness and its defects (due to locally varying chemical composition or the crystalline structure of the surface), as well as kinks, edges and steps that occur during the growth of a crystal. Studies have shown that the same type of protein may have different adsorption energies, varying from highly favorable to highly unfavorable, for surfaces with different lattice structures but the same chemical composition (Heinz et al. 2009; Oren et al. 2005). Although modeling of the intrinsic crystal structure is straightforward, it is important to consider that different crystal planes of a material (which can be kinetically or thermodynamically favored) may be exposed during a surface adsorption process. For example, during the growth of a nanoparticle, different surface planes of the same particle can display different structures at the same time, e.g. (100) and (111) (Korzeniewski et al. 2011).

Similarly, protein binding properties depend on the material structure. Rechendorff et al. (2006) showed that adsorption of fibrinogen on a tantalum surface can be induced by up to 70% by increasing the surface (root mean squared) roughness from 0.2 to 32.9 nm. The increase was much greater than the increase in the surface area due to the surface roughness of around 20%. On the other hand, Rechendorff et al. (2006) found that the adsorption of the more globular protein, bovine serum albumin, to tantalum induced by the surface roughness was similar to the increase in the surface area, thus demonstrating a selective effect of the material structure on the adsorption processes of different proteins. Finally, using molecular dynamics (MD) simulations, Nada (2014) investigated the interactions of an aspartic acid with step edges and kinks, as well as flat regions of a calcite crystal surface. They showed that aspartic acid binds preferentially to an acute step edge and not to an
obtuse one, due to the ordered structure of water near the obtuse step edge that prevented the aspartic acid from binding strongly (Nada, 2014). In summary, these studies demonstrate that the topography of a surface may determine its binding characteristics to proteins and peptides. Although, many topographic features can be neglected in modeling studies when comparing to experiments with well-characterized surfaces, the characteristics of the surface topography should be taken into account to realistically model and simulate the interactions of proteins with the materials used in real-life applications.

### 3.5 Morphology

The size of a nanoparticle, and therefore the curvature of its surface, has a strong impact on both the physico-chemical characteristics of the nanoparticle itself, such as surface polarizability, surface charge (Chiu et al. 2009) and isoelectric point (Suttiponparnit et al. 2010), and the properties of the layer of coating molecules which can determine its acidity (Wang et al. 2011). The curvature of a nanoparticle may change the properties of its hydration shell as well. The solvation free energy of a nanoparticle becomes more favorable as the surface curvature decreases (and the size of the particle increases) owing to the increased water–particle interaction energy, which in turn determines its polarity (Chiu et al. 2009).

The effect of surface curvature on protein adsorption properties, i.e. the kinetics, thermodynamics and structural stability of adsorbed proteins, increases with decreasing particle size (Lacerda et al. 2010), but also depends on the adsorbate. Moreover, a strong difference in binding to spherical nanoparticles and nanorods is often observed (Gagner et al. 2011). A structural adaptation to the curvature of the nanoparticle surface upon adsorption may lead to a loss of enzymatic activity of some proteins (Wu & Narsimhan, 2008), or it may lead to significant changes in secondary or tertiary structures of self-assembling peptides (Shaw et al. 2012) or proteins (Tavanti et al. 2015; Yang et al. 2013) adsorbing onto the surfaces. It has been shown that while surface curvature may help to retain the tertiary structure of some proteins with globular structures adsorbed on small nanoparticles (Vertegel et al. 2004; Lundqvist et al. 2004), it can also cause significant loss in the secondary structure of a protein upon adsorption (Gagner et al. 2011). The effects of the surface curvature of nanoparticles can be neglected in computational studies when the binding of relatively small proteins to large nanoparticles is studied. However, the size and curvature of a nanoparticle may play a major role in the adsorption patterns of proteins not only due to geometric adaptation of the protein to a nanoparticle of similar size, but also due to changes in the physico-chemical properties of the nanoparticle itself. DFT calculations of amino acid adsorption on a single-walled carbon nanotube (SWCNT) showed that glycine adsors more strongly on a nanotube (3,3) than on a flat graphite surface, whereas phenylalanine adsors more strongly on a flat graphite surface, and the amino acids cysteine and histidine, showed no significant change in their adsorption energies (Roman et al. 2006). Binding free energy calculations of amyloidogenic apoC-II(60–70) peptide on fullerene, CNT and graphene using MD simulations showed that the binding affinity was weakest for the fullerene and strongest for the graphene due to reduced efficiency of π-stacking interactions between the aromatic side chains of the peptide and the fullerene and CNT arising from the increased surface curvature (Todorova et al. 2013). Xie et al. (2014) calculated the adsorption free energies of Alzheimer’s β-amyloid peptide fragments (Aβ) to two different fullerene nanoparticle systems, C_{180} and three C_{60} (3C_{60}). They found tighter binding of the peptides on the larger C_{180}.

Raffaini & Ganazzoli (2013) showed that the binding strength of a protein to a carbon nanotube depends on the morphology and that the interaction energy between the protein and the nanoparticle was larger for the concave interior surface of the nanotube than for the convex outer surface. In a similar study, Chen et al. (2009b) showed that the length and the diameter of CNTs affect the energetics of the interactions with a peptide drug. While longer CNTs provide more space to trap the peptide inside the tube, a smaller diameter increases the interaction energy.

The surface morphology not only alters the binding characteristics of a peptide or protein to a surface but also the interactions between biomolecules near a surface or a nanoparticle, thus leading to intermolecular structural changes. Li et al. (2011) performed 3 separate sets of simulations of Aβ(16–22) peptides that abnormally self-assemble into β-rich aggregates: amorphous peptide in solution, amorphous peptide with SWCNT, and prefibrillar peptide with SWCNT. They observed that without the CNTs, the amorphous peptides form β-sheet structures. On the other hand, simulations with SWCNTs showed that the amorphous peptides tend to form disordered coils, whereas the β-sheet structures formed by the prefibrillar peptides were destabilized due to the interactions of the peptides with the CNTs.

Finally, noteworthy to mention is that a change in the morphology of a surface/particle will lead to changes in its physico-chemical properties. In a study by Emami et al. (2014b), it was shown that the size of the silica nanoparticles determines their surface ionization levels. The larger nanoparticles have higher ionization and therefore bind more strongly to peptides with high net positive charge. Further, Baier et al. (2014) recently investigated the binding free energy profile of 12-mer peptides on polar (001) and non-polar (100) ZnO surfaces. Employing an enhanced sampling approach (see Section 10.2), the authors...
showed that there is a positive correlation between the ZnO particle sizes obtained in experiments in the presence of the peptides and the calculated affinities of the peptides for the ZnO surface. Their result showed that the selective adsorption of a peptide can impact the growth of certain nanocrystals. Therefore, one should consider modeling, not only the size and curvature of the material, but also other properties that depend on the morphology, in particular for nanoparticles.

4. Which modeling and simulation techniques are applicable to protein–surface interactions?

The approaches used for modeling protein-inorganic surface systems cover a broad range of scales from sub-atomic quantum mechanical (QM) through classical atomic levels to mesoscopic descriptions and further to continuum descriptions on a macroscopic scale. Figure 2 shows the most common techniques to model and to simulate molecular interactions. Simulations at the QM level are applied to systems of a few hundreds of atoms at most and they do not reach the nanosecond scale, hence they are directly applicable only to small nanoparticle-ligand systems (Mahmoudi et al. 2011).

Many physical processes at the protein-solid surface interface are driven by physisorption, i.e. proceed without formation of chemical bonds between the adsorbate and the solid surface. In contrast to the relatively well-understood chemisorption, the nature and behavior of non-covalent adsorption is often unclear since multiple factors, which depend strongly on the surface type, influence the interactions that govern the adsorption process. Even if chemical binding takes place, physical adsorption drives the first stages of molecular recognition and induces long-time scale structural adaptations of a protein to a solid surface. Non-covalent binding processes can be described within a MM framework, which drastically reduces computational costs compared with QM, and thus enables the simulation of the dynamics of systems that consist of millions of atoms for up to microseconds with solvent molecules modeled explicitly. All-atom MM simulations are important, in particular, for investigation of the dynamic and thermodynamic properties of protein adsorption (Mahmoudi et al. 2011; Utesch et al. 2011). The nano length scale is usually appropriate for studying protein–surface interactions at the molecular level, and hence, all-atom simulations are common methods of choice (Gagner et al. 2012).

In experiments, adsorption events typically take place over time periods from milliseconds to hours which are far from the accessible time scales of all-atom simulations (Mücksch & Urbassek, 2011). The initial stages of protein adsorption to surfaces occur on a sub-second time scale. These may be followed by a slow stage in which large secondary structural changes occur, such as the transition from α-helix towards β-sheet of lysozyme on a SAM surface (Sethuraman & Belfort, 2005), and may entail denaturation over periods lasting up to several hours, or even days (Gray, 2004; Pan et al. 2012). Capturing the time and length scales of a complete adsorption process therefore necessitates employing mesoscopic, coarse-grained (CG) and multiscale approaches (Gray, 2004; Wei et al. 2011). Furthermore, hybrid approaches such as QM/MM to bridge typical time and length scales of conventional approaches, and enhanced sampling simulation techniques to accelerate adsorption and desorption events have been proposed and successfully applied to study protein–inorganic surface interactions (Euston et al. 2008; Utesch et al. 2011; Zhang & Sun, 2010).

5 Quantum mechanics studies of protein–surface interactions

QM-based methods are those where the quantum nature of electrons is explicitly taken into account while the much heavier nuclei are usually considered as classical particles moving in the field generated by the electrons and the other nuclei. Several approaches belong to this class, ranging from those that are relatively fast but rich in adjustable parameters, such as tightbinding and semi-empirical methods, to very accurate but also computationally very expensive, parameter-free calculations, such as coupled-cluster. The QM method that has been applied most extensively so far for studying biomolecules at surfaces is DFT, as it represents the best compromise between accuracy and computational feasibility.

The fundamental idea behind DFT (Martin, 2004), based on the Hohenberg and Kohn theorems (Hohenberg, 1964), is that, for non-degenerate systems, the ground state electron density, \( n(r) \), alone determines the entire behavior of the system, and that such \( n(r) \) minimizes the energy of the system. The most useful approximation of such energy as a function of \( n(r) \) has been proposed by Kohn & Sham (1965). It translates the minimization problem to a non-linear, independent particle approach akin to the Hartree–Fock one. A central quantity within this approach is the so-called exchange correlation functional \( f_{xc} \), i.e. the \( n(r) \)-dependent energy contribution due to quantum-mechanical and many-body deviations from a mean-field description of the electrons. So far, there is neither an exact expression for \( f_{xc} \) nor a one-fits-all approximation. An important and often not-obvious choice to be made for any DFT calculation is therefore which \( f_{xc} \) is the best for the system under study.
Given the geometry of the nuclei in space, DFT allows the forces acting on them to be calculated. This relation can be used to find the nuclei geometries that provide local and global minima of the system energy (i.e. the optimal or relaxed geometry). These are obviously the most important structures, since they are the structures in whose neighborhood the system fluctuates. In the following, we shall refer to this kind of DFT calculation as static. In fact, forces can also be used to simulate the dynamics of the nuclei, and therefore, in principle, the thermodynamics as well, via different propagation algorithms, such as Born-Oppenheimer and Car-Parrinello (Marx & Hutter, 2000). DFT \textit{ab initio} molecular dynamics (AIMD) is often used to collectively refer to this kind of simulations. Several software packages are available and maintained for performing static DFT and/or AIMD calculations. They differ in the numerical approaches implemented to solve the Kohn–Sham equations (e.g. using plane-waves or localized atom-centered functions), in the available fxc, in the level of parallelism (i.e. suitable for highly parallel computers or for a few-core, high memory workstation), in the pre- and post-processing tools offered, and finally in the distribution policy (e.g. open-source versus proprietary, free versus commercial).

Considering proteins on surfaces, static DFT calculations of single amino acids or even di- and tri-peptides, (Lee et al. 2014; Muir et al. 2014) interacting with a surface, with no solvent present, are now affordable and widely used (Arrouvel et al. 2007; DiFelice et al. 2003; Di Felice & Selloni, 2004; Ghiringhelli et al. 2006; Iori et al. 2008; Rimola et al. 2009). They may seem rather distant from biophysically relevant systems, but they can actually provide important information by themselves, or be preliminary to other approaches (e.g. provide the basis for a classical FF parameterization). Among the DFT calculations that directly provide useful information, we mention those aimed at understanding whether a covalent bond is established between an amino acid and an inorganic surface. In fact, when a covalent bond is present (chemisorption), it is unlikely that the protein and solvent environment missing in the calculations dramatically change its nature, although they do modify the details. Notable examples of calculations of this kind are works on Cys–Au(111) interactions (Buimaga-Iarinca & Calborean, 2012; Di Felice et al. 2003; Di Felice & Selloni, 2004; Fajín et al. 2013; Nazmutdinov et al. 2007), often used for protein immobilization (Vigmond et al. 1994). Static DFT calculations on amino acids or simpler molecules representative of chemical groups in natural amino acids on metals (Hong et al. 2009; Iori et al. 2008, 2009) have also highlighted the peculiar nature of the amino acid–metal interaction. Depending on the partners, such an interaction ranges from clear non-bonding (e.g. alkyl side chains on Au(111)) to clear chemisorption (Cys on Au(111)). It also encompasses border-line cases where the interaction has some typical covalent bond characteristics, such as electron sharing and directionality, and yet it is only marginally stronger than a non-bonded interaction (e.g. imidazole on Au(111), see Fig. 3).

DFT calculations have also been used to elucidate amino acid adsorption on silica (Rimola et al. 2013), hydroxyapatite (Jimenez-Izal et al. 2012), alumina (Arrouvel et al. 2007) and titania (Carravetta et al. 2009; Koch et al. 2011) surfaces where electrostatic interactions and hydrogen bonds are important, and where the formation of chemical bonds often implies complex reaction mechanisms. The interaction of quartz and aluminosilicate structures with phospholipids was also studied by DFT to understand why the enzyme phospholipase A2 digests phospholipids faster in the presence of quartz than...
aluminosilicates (Snyder & Madura, 2008). The interaction of amino-acids with graphene has also been the subject of recent DFT studies (Akdin et al. 2013; Wang et al. 2014).

Static DFT simulations have been performed to parameterize classical FFs for protein–surface interactions in water (Belluci et al. 2012; Caravetta & Monti, 2006; Caravetta et al. 2009; Di Felice & Corni, 2011; Ghiringhelli et al. 2006; Iori et al. 2009; Schneider & Colombi Ciacchi, 2010; Wright et al. 2013a, 2013b). The general strategy here is to perform energy calculations on the optimized geometries and/or along specific coordinates, for either entire amino acids or simpler amino acid analogues, on the target surface, and to use the results as a training set. Classical parameters are thus adjusted so as to reproduce as closely as possible the QM interaction energies and geometries in the classical calculations. DFT calculations have also been used to generate partial atomic charges for the surface atoms to be used for Coulombic interactions in the classical models.

AIMD is intrinsically much more expensive than static DFT simulations and, for this reason, its application in the protein–surface field has been less popular. Recently, Motta et al. (2012) investigated the adsorption of glycine on a stepped boehmite (AlO(OH)) surface in water, identifying inner-sphere adsorption, i.e. displacement of surface hydroxyl groups by glycine molecules, as the most favorable. Wright & Walsh (2012) focused on ammonium and acetate ions, which are analogues of common chemical groups in amino acids, at the water/quartz interface. Colombi Ciacchi and colleagues used AIMD to build a model of a native silicon oxide surface (Colombi Ciacchi & Payne, 2005; Cole et al. 2007) that was afterwards used for studying peptide–silica interactions (Schneider & Colombi Ciacchi, 2012).

The current limits of static and AIMD DFT calculations in the field of protein–surface interactions are well illustrated by some recent examples shown in Fig. 4. (Rimola et al. 2012) exploited static DFT to study the adsorption of an entire dodacapeptide on a hydroxyapatite surface, including some key water molecules (the system was composed of around 500 atoms). They discussed the driving force for the adsorption and the role of the surface in determining the peptide folding. A system of similar size (a dodacapeptide on a graphene sheet) has also been investigated by Akdim et al. (2013) to confirm by DFT the adsorbed peptide geometries obtained with a classical force-field. Calzolari et al. (2010) considered a model polysarne β-sheet, periodically replicated, simulated by AIMD on Au(111) in liquid water. This system was composed of approximately 500 atoms and its time evolution could be simulated for 20 ps. Such a simulation time and the composition of the system allowed the investigation of some specific questions, such as the nature of the local β-sheet/Au interactions, the competition between water and the serine side chain for gold, as well as the nature of the β-sheet/water interface. However, it is apparent that several other important questions are inaccessible with this kind of approach. Today, larger systems and longer AIMD simulations are affordable but the most expensive AIMD simulations are still confined to a few thousand atoms and a few hundred ps.

Beside the current limitations related to the computational cost of DFT simulations, one notorious drawback of the fxc functionals used so far is worth discussion here. It is the inability to account for long-range dispersion (London) interactions, that results in an underestimated interaction strength, and even no solute–surface binding when such interactions are the only relevant ones (e.g. for inert metal surfaces and saturated, non-polar molecules). Various corrections have been proposed to solve this problem (Tkatchenko et al. 2010), and some of them have also been tested in the framework of protein–surface interactions with encouraging results. In particular, the DFT-Dn methods (Grimme, 2004; Grimme et al. 2010; Wu et al. 2001) (that add to the DFT energy, empirical atom-atom d⁶ terms, where d is the pairwise interatomic distance, suitably damped for small d) have been used for amino acid and peptide adsorption on mineral surfaces (Folliet et al. 2013; Rimola et al. 2009, 2012) and on graphene (Akdin et al. 2013). The fxc functional, vDW-DF, which does not contain empirical terms (Dion et al. 2004), has been tested against experimental desorption energy data for some small molecules on Au(111) (Wright et al. 2013a). It was used to provide the main data (stable geometries and the related energies for amino acid analogues on Au(111) and Au(100)) needed to parameterize the GoLP-CHARMM FF (Rosa et al. 2014b; Wright et al. 2013a). Tests confirmed the reliability of the vDW-DF adsorption energies, within a few kJ/mol of the experimental values (Fig. 5), and pointed to an already documented tendency of vDW-DF to provide contact distances and Au lattice parameters that are slightly too large by 0.1–0.2 Å (Lee et al. 2010). Other functionals akin to vDW-DF have been proposed to correct this deficiency (Lee et al. 2010; Klimeš et al. 2010, 2011), and are awaiting testing and validation in the field of molecular adsorption, and specifically for protein–surface interactions. The computational overhead in using these functionals is modest, and they have also been applied to other biomolecules adsorbed on gold, such as nucleic acids (Rosa et al. 2012, 2014a, b). AIMD is also possible with these functionals, as exemplified by a recent study of the liquid water/gold interface (Nadler & Sanz, 2012). In this case, dispersion interactions do not change the picture provided by conventional functionals (Cicero et al. 2011). In summary, the DFT limitations connected with the lack of dispersion interactions are currently being overcome by recent methodological developments.
In the future, approaches other than DFT, such as Quantum Monte Carlo (QMC) (Austin et al. 2012), may also become popular for investigating biomolecule–surface interactions. QMC is based on polyelectronic wavefunctions and naturally accounts for long-range dispersion interactions. Such wavefunctions are determined by Monte Carlo (MC) algorithms, which are highly parallelizable. While QMC calculations are currently too expensive to routinely study biomolecule–surface interactions, its intrinsic high parallelism combined with the constantly growing size of modern supercomputers, make its application to this field likely in the years to come.

To conclude, despite their size and time-scale limitations, DFT-based approaches are playing an important role in revealing the physico-chemical basis of protein–surface interactions. They are used either to provide a detailed picture of the local amino acid–surface interactions or as a basis for developing classical atomistic models, i.e. as a source of benchmark data to train classical FFs or of model structures for complex materials such as amorphous silica.

6. Challenges in applying biomolecular molecular mechanics force fields to protein–surface interactions

Modeling and simulation of protein–surface interactions brings along the challenges associated not only with modeling the surfaces and proteins separately, but also with modeling the system as a whole (see Fig. 6 for a depiction of protein–surface interactions in aqueous solvent). The FFs routinely used in modeling and simulation studies of proteins are parameterized for interactions between biomolecular fragments or small chemical compounds in aqueous solution. Although the FFs developed for protein simulations may provide a good approximation for modeling the interactions between a protein and a surface in some cases, in general, the force field parameters to be used have to be derived and calibrated for the systems of interest to obtain high quality results.

6.1 Interaction potentials

The classical potential energy functions employed in the all-atom molecular mechanics force fields (MM FFs) for biomolecules, such as AMBER (Cornell et al. 1995), CHARMM (Brooks et al. 1983), GROMOS (Oostenbrink et al. 2004), and OPLS-AA (Jorgensen et al. 1996), are widely used and thoroughly evaluated for simulation of biomolecules in aqueous solution. Most commonly used biomolecular FFs are expressed as a sum of pairwise interaction terms that represent changes of the (i) chemical bond lengths and bond angles of a molecule as harmonic spring functions; and (ii) torsions as periodic functions (dihedral angles, or torsional rotation of atoms around a central bond); and (iii) non-bonded electrostatic and van der Waals intra- and inter-molecular interactions:

\[ E_{\text{Total}} = E_{\text{bond}} + E_{\text{bond-angle}} + E_{\text{tension}} + E_{\text{electrostatic}} + E_{\text{vdW}} \]

Electrostatic interactions between the atoms in the system are approximated by Coulomb’s law with fixed point charges assigned to each atom, whereas van der Waals interactions are typically described by the Lennard–Jones (12–6) potential:

\[ U_{\text{LJ}} = 4\varepsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right] \]

where the two terms represent repulsive and attractive interactions, respectively, and the parameters, and \( \sigma \) are expressed as a combination of parameters of atoms \( i \) and \( j \). If this form of the FF is suitable for the solid surface to be studied, it can also be applied for simulations of protein adsorption.
The FFs derived specifically for simulation of inorganic materials are generally aimed at reproducing the structural and physico-chemical properties of the bulk or surface of a material in vacuum and often have a functional form that is different from that of conventional biomolecular force-fields. Usually, monatomic systems with close-packed structures can be reasonably well described by two-body interaction potentials, whereas a three-body potential must be used for semiconductors (Vashishta et al. 1990). Additional angle-dependent terms can be employed to preserve the directionality of bonds, e.g. in the crystal packing arrangement of a bulk solid (Cruz-Chu et al. 2006). Furthermore, a Buckingham potential of the form $A e^{-B/r^6} - C/r^6$, where $A$, $B$, and $C$ are force field parameters, is often used instead of a Lennard–Jones potential to better describe the repulsion between ions in inorganic materials (see, for example, Van Beest & Kramer, 1990). Moreover, a balanced parameterization of a simple pairwise interaction energy model commonly used in standard biomolecular FFs is not straightforward and is not always possible for the water-material interface. For instance, a combination of two-body and three-body non-bonded potentials was employed in a silica force-field to preserve the tetrahedral structure of the silica glass and to reproduce accurately silica surfaces, pores and surface wettability (Cruz-Chu et al. 2006). Another limitation of a simple pair-wise function is that it is based on a fixed point charge approximation for Coulomb’s interaction, which omits the interactions between higher order multipoles and polarization effects arising from the electrostatic field of the environment (a non-polarizable force-
The partial atomic charges are usually assigned for a molecule in an aqueous environment and, thus, the polarization effect of the aqueous surroundings is implicitly accounted for. Non-polarizable FFs have been extensively validated over the last decades and currently offer a robust description of the equilibrium properties of biomolecules in aqueous solution. However, polarization effects of the surface (such as metal or carbon surfaces, as discussed earlier) may contribute notably to the binding energy of an adsorbate, as well as to the structure of the surface itself.

Polarizable FFs for biological molecules are currently under active development and validation (Halgren & Damm, 2001). However, apart from their notably higher computational cost, polarizable FFs for biomolecules are not designed to reproduce the polarization properties of inorganic surfaces. In Section 7, we will consider several models adapted for simulation of surface polarization effects.

Despite all the limitations mentioned above, the early simulations of protein and peptides on solid surfaces have typically been performed using standard biomolecular FFs. In the last decade, however, much effort has been invested to evaluate and adapt biomolecular FFs to the simulation of interfaces between biomolecules in water and specific inorganic materials, and several new approaches for developing general bio-inorganic FFs have been reported. For example, a dual-scale FF that includes two sets of partial atomic charges, optimized for each phase separately, has been developed (Biswas et al. 2012). In the recently reported INTERFACE FF (Heinz et al. 2013), partial atomic charges and van der Waals parameters have been optimized using a non-polarizable fixed-charge model for a large number of different materials based on the properties of the solid surfaces, e.g. crystal unit parameters, surface tension, water contact angle and interfacial tension, and adsorption energy of selected peptides. Some applications of these new models for specific surfaces will be discussed in Section 7.

6.2 Solvation models

The properties of the hydration shell of a solid substrate, in particular variations of physical properties, such as the density, free energy and dielectric constant of water in a hydration shell, and chemical properties, such as surface reconstruction and...
ionization, often govern adsorption processes. Therefore, a rigorous treatment of water, whether explicit or implicit, is required to account for the effects of water on adsorption. FF parameters for water molecules are usually parameterized for the bulk water state and thus do not take into account the specific features of the surface or protein hydration shell. Moreover, the properties of a solid surface hydration shell differ from those of small solutes and biomolecules in that, generally, more water–water hydrogen bonds are broken and the water is forced to build an extended ordered surface layer. In general, there is still uncertainty as to whether standard explicit water models are able to adequately reproduce adsorption phenomena for proteins and surfaces. Studies of water properties on a wide range of surfaces have been reviewed by Henderson (2002).

The microscopic properties of the water molecules on a surface can be parameterized using DFT ab initio computations, which have been widely used for predicting the adsorption energies of water molecules as well as water dissociation properties on different surfaces. The AIMD technique mentioned earlier is a powerful tool for exploring water properties in a monolayer or even in several layers (e.g. on gold, see Velasco-Velez et al. 2014)). However, it was shown that the spatial and temporal limitations of DFT simulations may prevent accurate modeling of water structural transitions from the hydration shell to bulk, which is important for the modeling of protein adsorption (Große Holthaus et al. 2012). Accordingly, an all-atom MM representation of water has to be used for exploring the adsorption processes of molecules in solution, which usually occur on a much longer time-scale. Likewise, an all-atom MM model of the solvent can be replaced by a continuum solvent model to enable longer simulations of large proteins or protein solutions for investigating their adsorption properties.

Implicit solvation is a common way to reduce the computational cost of simulations that arises in large part from calculating the interactions between the explicit water molecules. In this approach, the contribution of the solvent to the binding energies between the solutes is represented with a mean field model, and is usually decomposed into two major contributions: electrostatic contributions due to charge interactions and the change of dielectric polarization of their surroundings, and non-polar contributions due to the change in entropy and the dispersion energy upon removing water from the solute surface. The non-polar desolvation energy is commonly estimated as a linear function of the solvent accessible surface area. The most common approaches for computing the continuum electrostatic contribution to the binding free energy is to solve the Poisson–Boltzmann (PB) equation or the Generalized Born (GB) model, which provides an approximation to the PB method. The PB method is often used with a rigid molecule approximation because it is relatively computationally expensive to solve the PB equation for every macromolecular configuration during a simulation. For applications of these methods to free energy calculations of protein-surface systems, see Section 10.

The implicit solvent models traditionally used for molecules in solution do not, per se, account for the unique properties of the hydration shell and the interactions of the water in direct contact with the inorganic surface. However, they can be parameterized to account for the hydration shell characteristics of individual materials. An example of such a continuum solvent model is ProMetCS, which was derived for protein-gold (111) interfaces (Kokh et al. 2010). In ProMetCS, the effects of the hydration shell distortion at the protein-surface interface were accounted for by the addition of an analytical function parameterized to reproduce the potential of mean force (PMF) of a probe ion/atom on the gold surface obtained from explicit water MD simulations. The effects arising from the partial replacement of the metal hydration shell by a protein adsorption site are described by a free energy penalty, proportional to the protein-surface contact area, compensating the Lennard–Jones attraction to a large extent. The PMF of an atom on the gold surface computed using the ProMetCS FF reproduces the profile of the PMF function obtained in explicit solvent MD simulations (Fig. 7).

7. All-atom molecular mechanics studies of protein–surface interactions

Some of the factors that are particularly important to take into account when developing or applying MM FFs for simulation of protein-surface adsorption are: (i) compatibility with a standard MM FF for biomolecules; (ii) the ability to provide dynamic and conformational properties of biomolecules in solution as well as on the surface; (ii) the ability to provide physical and chemical properties of the surface such as structure and polarization; (iii) a correct reproduction of the hydration shell properties for a particular surface; and (iv) the ability to model changes in the surface layer upon protein adsorption or surface reconstruction, for example, due to water dissociation or adsorption of ions.

In this section, we will focus on the FFs developed for describing protein–surface, protein–water and surface–water interactions in atomic detail. Further, we will briefly introduce the techniques employed to tackle some of the challenges posed by different surface types. In the following sub-sections, the most commonly studied surface types, namely elemental metal, titanium oxide, silicon oxide, mineral, self-assembled monolayer and sp²-carbon surfaces, are reviewed.
7.1 Metal surfaces

The metal polarization in the presence of the charges of a solute, as discussed earlier, has a significant effect on adsorption and, therefore, efforts have been put into integrating polarization effects into FF models. The simplest way to introduce the induced polarization of the metal surface in a computational model is using a classical image-charge approximation for a charge and a zero-potential metal surface in a continuum dielectric medium with an interaction energy expressed by Coulomb’s law between the charge and its image of the opposite sign (see Fig. 8a). This approximation was shown to give an accurate description for the interaction energy of an Al(111) surface with a charge when the separation distance was more than 2.5 Å (Finnis & Finnis, 1991), while at closer distances to the surface, the image charge model overestimated the interaction potential (Finnis & Finnis, 1991; Smith et al. 1989). The choice of the image plane position was discussed by Heinz et al. (2011) who suggested that the image plane may deviate from the jellium edge (which corresponds to a plane half a lattice spacing above the first layer of surface atoms). The position of this image plane is further away from the surface atoms than the jellium edge and its deviation is represented by a system dependent offset, δ, which is determined by the response of the electron density of the metal (Smith et al. 1989). Heinz et al. (2011) exploited this scheme to estimate polarization effects a posteriori from non-polarizable fully atomistic MD simulations of a water-Au surface, as well as a water-peptide-gold system.

Although implementation of the image-charge approximation in an all-atom MD simulation is straightforward, it is impractical for large systems due to the increased computational load that scales with the number of interacting particles. In an alternative approach, which can be incorporated into any commonly used MD energy function, virtual dipoles or rods that can adjust their position in response to the external electrostatic field are introduced on all the surface atoms. Each virtual dipole is constrained at one end to a real surface atom and, depending on the model, can either change the magnitude of its dipole moment (Drude oscillator) or its orientation (rigid rods, see Fig. 8b) and, hereby screen the external electrostatic field. A model with virtual rigid rods for simulation of the polarization of a metal surface was proposed and implemented by Iori & Corni (2008) and used in the GolP family of FFs optimized for Au surfaces (Iori et al. 2009; Wright et al. 2013a, b), as well as for Ag and graphene.

The first optimized sets of Lennard–Jones parameters to include metal–water and metal–amino acid interactions was developed by Ghiringhelli et al. (2008) on the basis of DFT calculations, and by Heinz et al. (2008) and Vila Verde et al. (2009) to reproduce Au(111) hydrophilicity. In the GolP FF (Iori et al. 2009), which is based on the OPLS FF, additional parameters describing the interactions of biomolecular groups with the gold surface are parameterized from QM calculations and experimental studies of adsorption energy. To reproduce the binding energy and orientation of small molecular fragments, the authors included and optimized a set of additional parameters that describe the van der Waals interactions with gold as well as stronger, chemical-like bonds between aromatic groups and gold atoms in the form of a Lennard–Jones function.
This parameterization can be directly used in standard MM force-fields. The GolP FF has been applied in MD simulations of the adsorption of amino acids, as well as several proteins, on gold (Brancolini et al. 2012; Cohavi et al. 2011; Hoefling et al. 2010a, b, 2011; Kokh et al. 2010). Furthermore, the ProMetCS (Kokh et al. 2010) continuum solvent model describes protein–surface interactions in atomic detail using the GolP Lennard–Jones interaction parameters together with an image charge model for protein-metal surface electrostatic interactions.

Following the same strategy, a new force-field, GolP-CHARMM (Wright et al. 2013a, 2013b), has recently been developed on the basis of the CHARMM FF for studying the interaction of proteins with Au(111) and Au(100) surfaces. In GolP-CHARMM, special attention has been paid to the correct description of the properties of the water molecules adsorbed on gold, with the properties being compared with ab initio MD simulations (Cicero et al. 2011). The new FF improves the quality of the simulation of the hydration shell, reproducing the increased tendency of interfacial water molecules to donate hydrogen bonds to other water molecules compared with those in the bulk water and reproducing the energetics of water molecules on a gold surface.

Heinz et al. (2008) developed Lennard–Jones parameters for face-centered cubic metals based on experimentally determined densities and surface tensions at 298 K under atmospheric pressure for use in modeling of mixed interfaces. They provide two sets of parameters for metal atoms (Ag, Al, Au, Cu, Ni, Pb, Pd, Pt); one set suited for 12–6 Lennard–Jones functions to be used in AMBER, CHARMM, CVFF and OPLS-AA force-fields, and another for 9–6 Lennard–Jones functions to be used in the COMPASS (Sun, 1998) and PCFF force-fields. Penna et al. (2014) used surface models and the Lennard–Jones parameters of Heinz et al. (2008) with a CHARMM FF (MacKerell et al. 1998) to investigate the interactions of peptides with Au and Pt surfaces.

In addition to polarization, surface charges should be taken into account when modeling a surface. Bizzarri (2006) investigated the dynamics and electron transfer (ET) properties of the bacterial ET protein, azurin, anchored to neutral, positively and negatively charged gold surfaces using classical MD simulations. They calculated the ET rate between the azurin copper atom and the sulfur atom of the cysteine surface anchor using classical Marcus theory (Marcus & Sutin, 1985). The computed ET rate was highest on the positively charged surface and lowest on the neutral surface. However, it was not clear to what extent the relatively larger structural changes observed in the molecule anchored on the positively charged surface play a role in this trend.

### 7.2 Titanium oxide surfaces

Because of its high stability and resistance to corrosion, titanium is the material of choice for many medical and technical applications and it is, therefore, one of the most widely used materials in computational modeling. Crystals of TiO₂ (titania) appear in three forms: rutile, anatase and brookite. The rutile form, in particular its interaction with water, has been the most
studied by computer modeling. An oxide layer, predominantly of amorphous structure, is formed at the titanium-water interface (Schneider & Colombi Ciacchi, 2010). The properties of this layer depend strongly on the physico-chemical conditions and the temperature regime of adsorption.

There is controversy between theory and experiment regarding the extent to which water dissociates on TiO$_2$ simulations often show a larger degree of dissociation than experiment (Diebold, 2003; Huang et al. 2014; Sun et al. 2010). It is believed that for rutile (110), dissociation takes place predominantly on defects, while water can also dissociate to some extent on a defect-free (100) surface, and effectively dissociates on a (001) surface. Accordingly, the density distribution of water strongly varies for different TiO$_2$ surfaces from two sharp maxima on a rutile surface to a smooth distribution on hydrophobic titania (Huang et al. 2014). The mixture of hydroxyl groups and molecularly adsorbed water may form a partially ordered solvation layer that influences interactions of peptides with the surface (Li et al. 2012).

The behavior of water on different types of TiO$_2$ surfaces, including partial dissociation of water, was investigated using the reactive FF ReaxFF (Huang et al. 2014; Kim et al. 2013) and AIMD (Carravetta et al. 2009; Schneider & Colombi Ciacchi, 2011). More details on theoretical studies of titania–water interactions can be found in the review by Sun et al. (2010).

As mentioned before, water dissociation on oxide surfaces is reversible and cannot be simply described by classical MM force-fields. Bandura & Kubicki (2003) pointed out that the two-body interaction potential is not suitable for capturing the correct directionality of hydroxyl groups on a TiO$_2$ surface and that this problem can be solved by using a polarizable FF or by introducing bond stretching and angle bending terms for Ti–O–H. They used the Matsui & Akaogi (1991) model which incorporated additional bond-stretching and angle-bending terms for O–Ti–H into the Buckingham potential for the atoms of the (110) surface of rutile. The FF was later refined by Predota et al. (2004) and then revised for anatase (101), (112) and rutile (110) surfaces. In the FF developed by Borodin et al. (2003), an exponential term for modeling short-range repulsion was added to the standard Lennard–Jones function together with an additional polarization term, and the FF was optimized against the results of QM calculations for polyethyleneoxide on small TiO$_2$H$_4$ clusters and then applied to TiO$_2$ surfaces.

Since the use of AIMD and reactive FFs is too time-consuming for simulations of large systems and/or long time-scales, TiO$_x$ parameters have been incorporated into standard MM FFs employed in biomolecular simulations. A classical FF with a Finnis–Sinclair type many-body potential for the surface coupled to a Buckingham potential for the short-range repulsive interactions was developed by Schneider and Colombi Ciacchi and applied to simulating peptide interactions on a Ti/TiO$_2$ interface (Schneider & Colombi Ciacchi, 2010). The parameters of the AMBER FF were optimized for TiO$_2$ interactions with collagen peptides in a study by Köppen & Langel (2010). The ratio of protonated, hydroxylated and stoichiometric units on the surface was adapted to the physiological pH value and retained during MD simulations. The charges in the bulk of TiO$_2$ were fixed, but different charges on the surface oxygen and hydrogen atoms were tested. The binding of Glu and Lys to TiO$_2$ was found to be mediated by hydroxyl groups on the surface, with the adsorption energy strongly dependent on the charge model used.

Friedrichs & Langel (2014) re-parameterized the Matsui–Akaogi potential to the standard Lennard–Jones form for a description of the interactions between the rutile surface atoms and used their model to simulate adsorption of peptides on the rutile surface. Although this enabled standard MD software to be used, the simplicity of the form of the potential gave rise to some deviation of physical properties, such as crystal structure parameters and permittivity, from those observed in experiments and simulations with the original potential.

The adsorption of different peptides onto a rutile surface, TiO$_2$(110), was simulated in a series of studies by Monti and colleagues of the EAK16 and RAD16 oligopeptides (Carravetta & Monti, 2006) and (Monti et al. 2008), and β-sheet conformations of the EAK16 and RAD16 oligopeptides (Carravetta et al. 2009). The FF for biomolecule-surface and water–surface interactions used in these studies was optimized against data from QM calculations for a small titania cluster and experimental data (Carravetta & Monti, 2006). The TIP3P water model and a combination of a Buckingham potential for the interaction between the surface and surface-water atoms and a Lennard–Jones potential for most of the rest of the atoms were employed. Peptide binding to the surface was found to preferentially occur through the backbone atoms; particularly carboxyl oxygens and the C-terminal carboxy group on the peptides interacted with exposed Ti atoms. Water molecules were also observed to mediate peptide-surface binding through hydrogen bonds. Since the dissociation of water was not taken into account, the contribution of hydroxyl groups to peptide binding was not analyzed. Finally, the adsorption of glycine (Li et al. 2012) and cysteine (Li et al. 2014) onto TiO$_2$ was explored using the reactive force-field, ReaxFF (Van Duin et al. 2001), extended to treat the interaction of solid inorganic substrates with biomolecules. It is interesting that these adsorbing molecules, especially glycine, tend to form self-assembled clusters and that only weak adsorption was observed. Additionally, the proton transfer reactivity of the amino acid is observed to be enhanced by the presence of the surface.
It should be noted that MM simulations of proteins on reactive surfaces like titanium oxide should be performed with caution, because the presence of defects may change the surface properties dramatically, and thus affect the adsorption behavior. In a study by Utesch et al. (2011), the BMP-2 protein interacted unexpectedly weakly with the TiO2 surface, because it was unable to penetrate the highly-ordered water layer formed on the defectless surface.

7.3 Silicon oxide surfaces

A number of FFs have been developed for MD simulation of bulk silicon oxides and their interfaces with water due to the wide range of technological applications of silica-based materials (see reviews by Butenuth et al. (2012) and by Rimola et al. (2013)). There are, however, some challenges in building a FF for silicon oxide surfaces. In particular, for modeling the tetragonal bulk structure of SiO2 (silica), many-body potentials are often used, which are generally not compatible with common biomolecular force-fields. At the SiO2-water interface, water molecules dissociate and saturate dangling Si and O groups, which leads to formation of different types of exposed functional groups: hydrophilic silanols (Si-OH), and hydrophobic siloxanes (with oxygen buried, Si-O-Si) and silanes (with silica atoms saturated by hydrogen). Silanols groups can make hydrogen bonds to each other or two silanol hydroxyl groups can bind to one Si atom. Since they are hydrophilic in character, these groups form hydrogen bonds with surface water molecules. Crystalline silica has a dense ordered network of geminal hydrogen-bonded silanol groups leading to ordered water layers (Notman & Walsh, 2009) (see Fig. 9), while amorphous silica has isolated geminal and hydrogen-bonded silanol groups resulting in only locally ordered regions of water molecules and hydroxyl groups (Aarts et al. 2005).

Silica has an isoelectric point of about 2–3 (Parks, 1965), and, at neutral pH, between 5% and 20% of silanol groups are deprotonated, leading to a negative surface charge. Experimental studies of peptide adsorption on silica nanoparticles in aqueous solution (Patwardhan et al. 2012; Puddu & Perry, 2012) showed that binding was dominated by electrostatic interactions between negatively charged siloxide groups and ammonium groups at N-termini, and Lys and Arg containing peptides and, to a lesser extent, hydrogen bonds between negatively charged silanol and siloxide groups and polar groups in Ser, His and Asp-containing peptides. Binding of hydrophobic peptides was observed at low pH (Puddu & Perry, 2012). A FF for MD simulations of a deprotonated negatively charged SiO2 surface in solution at neutral pH was reported by Butenuth et al. (2012).

Recently, several FFs for silica with biomolecular FFs have been developed. A FF based on the CHARMM FF was parameterized by Lopes et al. (2006) to obtain a structural and dynamic representation of water in the vicinity of neutral quartz crystalline surfaces using MD simulations. This FF was used for a set of peptides known to be strong and weak binders to the quartz (100) surface and showed that Pro, Trp and Leu were the main residues forming close contacts with the surface (Oren et al. 2010).

The GLASSFF_2.01 FF is compatible with CHARMM and with TIP3P water parameters and was developed by Cruz-Chu et al. (2006) for simulation of amorphous silica surfaces and nanopores. This FF employs standard two-body Lennard–Jones and Coulomb potentials as well as a three-body directional term for modeling the tetrahedral arrangement of bulk silica. To construct amorphous silica, the authors applied annealing cycles in MD simulations, which initiated surface reconstruction from crystal to amorphous silica accompanied by formation of dangling atoms (oxygens with less than two bonds and silicons with less than four bonds) that were required for modeling the wetting properties of silica correctly.

A comparison of FFs for the prediction of the physical and chemical properties of bulk silica and its aqueous interface such as atomic charge, bond length and angles, density of silanol groups on the surface and degree of ionization of silanol groups can be found in the papers by Skelton et al. (2011) and by Emami et al. (2014a). In the latter study, a data set with more than 20 different silica surface models that cover most important types of silica surface chemistry at different surface ionization states was assembled. Then, atomic parameters for the different chemical groups in silica were derived to reproduce a large set of chemical and physical properties of bulk and surface silica as well as the water contact angle, heat of water immersion and water adsorption isotherms observed in experiments. The parameters were integrated into several FFs (AMBER, CHARMM, COMPASS, INTERFACE (Heinz et al. 2013)) compatible with the TIP3P and SPC water models. In particular, the CHARMM-INTERFACE FF (Heinz et al. 2013) was applied to simulation of peptide adsorption on silica nanoparticles of different size at different pH values (Emami et al. 2014b), showing good agreement of nanoparticle coverage with experimental data (Puddu & Perry, 2012).

The Dual-FF (Biswa et al. 2012), (discussed in more detail in Sec.7-5) was optimized to reproduce the experimental binding energies of TGTG-X-GTGT peptides, with X = N, D, G, K, F, T, W and V, on hydroxylated quartz (100) and glass surfaces in MD simulations (Snyder et al. 2012). In the Dual-FF, CHARMM parameters were used for biomolecules, while interfacial
parameters, such as partial charges and Lennard–Jones parameters for silica–peptide and silica–water interactions, were tuned to obtain the best agreement with peptide adsorption energies.

7.4 Mineral surfaces

The main difficulty in simulating mineral-biomolecule interfaces in aqueous solution is that, while both minerals and biomolecules may have atom types in common, e.g. oxygen, the FF for bonded terms are different in minerals and biomolecules. Specifically, minerals are ionic systems and are best represented by ionic bonds based on Coulomb electrostatics and Pauli repulsion terms (Hauptmann et al. 2003). In contrast, biomolecules are covalently bonded. As for oxide surfaces, water may dissociate on a mineral surface leaving adsorbed hydroxyl groups at charged group sites on the surface with a partial occupancy of about 50% of the charged sites (De Leeuw, 2010). This process cannot be modeled in a standard MM FF, but the presence of hydroxyl groups/hydrogens on the surface may strongly affect the adsorption properties of biomolecules.

A number of FFs for modeling the bulk and surface properties of minerals have been developed (see (Mishra et al. 2013) and references therein), and several non-reactive FFs for the simulation of mineral-water interfaces have been reported. A general FF, CLAYFF, for modeling the interface of a hydrated multicomponent mineral with aqueous solution was proposed by Cygan et al. (2004), where charges on oxygen atoms and hydroxyl groups vary according to their structural environment. This FF includes a flexible SPC-based model for water and hydroxyl groups and treats most interatomic interactions as non-bonded with Lennard–Jones and Coulomb terms.

A methodology for the generation of a general bio-compatible FF for minerals was proposed by Freeman et al. (2007). In this work, existing FFs and models were used for the various components of the system (including the TIP3P model for water), while new parameters were derived for cross-species interaction terms, for example, for the interaction of a hydroxyl group with a Ca2+ ion. This methodology was used for simulation of small organic molecules on calcite and magnesite surfaces (Freeman et al. 2009) by combining the AMBER FF for organic molecules with inter-species terms proposed by Freeman et al. (2007); and for simulation of the binding of ovocleidin-17 protein to calcium carbonate nanoparticles (Freeman et al. 2010) and surfaces (Freeman et al. 2011). Furthermore, Katti et al. reported parameters for modeling clay minerals (NaSi16(Al6FeMg)O40(OH)8) in the CHARMM FF (Katti et al. 2005b) and amino-acid adsorption to these clay minerals (Katti et al. 2005a).

Hydroxyapatite (HAP, Ca10 (PO4)6 (OH)2) is the main mineral component of bones and teeth and thus a promising material for application in bone replacement. A FF for HAP has been developed with a Born–Mayer–Huggins potential with an exponential repulsion term similar to that of Buckingham potential instead of the standard Lennard–Jones plus Coulomb expression for the non-bonded interactions. The FF parameters were derived by fitting the QM electrostatic field in a 3D crystal environment and experimental crystal parameters (Hauptmann et al. 2003). This FF was subsequently re-parameterized for monoclin hydroxyapatite using an energy function consistent with common biomolecular FFs (Bhowmik et al. 2007). Dong et al. (2007) adapted these HAP parameters for the standard biomolecule FF for modeling the adsorption dynamics of the BMP-2 protein on a HAP (001) surface. In other studies, the Lennard–Jones and Coulomb parameters proposed by Hauptmann et al. (2003) were combined with CHARMM parameters for the protein using the standard mixing rule in order to study the adsorption of fibronectin on HAP (001) (Shen et al. 2008) and with OPLS-AA parameters for simulating the HAP-water interface (Zahn & Hochrein, 2003).
7.5 Self-assembled monolayer surfaces

The applicability of several commonly used biomolecular FFs for the simulation of peptide adsorption on hydrophobic and hydrophilic SAMs has been systematically evaluated by Latour and colleagues for CHARMM19 (Sun & Latour, 2006; Sun et al. 2007; Vellore et al. 2010), and CHARMM22, OPLS-AA, and AMBER94 (Collier et al. 2012) FFs. In the latter study by Collier et al. (2012), the FF parameters and partial charges of the SAMs (-OH and -COOH) were assigned by analogy to amino acids with similar functional groups. The authors compared experimental findings with their computed free energies of adsorption and qualitative behavior of peptides on different SAM surfaces such as the change in conformation upon adsorption and the orientation on the surface. This study (Collier et al. 2012) demonstrated that although some FFs perform reasonably well (the best match to experiment was obtained in simulations using CHARMM22 and AMBER94 (Cornell et al. 1995)), none of the FFs capture the specific interaction properties of the SAM-water interface. In particular, systematic overestimation of the binding strength of hydrophobic peptides and underestimation for negatively charged peptides was observed. On the other hand, Collier et al. (2012) also noted that altering FF parameters to reproduce the properties of adsorbed peptides unavoidably led to alteration of peptide behavior in solution. These results suggested that for accurate simulations of peptide adsorption a new FF strategy was required.

A new approach, the dual-FF, was proposed by Biswas et al. (2012). In this FF, different sets of non-bonded parameters, i.e. atomic partial charges and parameters of the Lennard–Jones potential, were used to represent intra-phase (i.e. between peptides or between water molecules) and inter-phase (peptide–SAM) interactions simultaneously in simulations. Furthermore, the non-polarizable TIP3P (Jorgensen et al. 1983) water model was reparameterized for water-SAM interfaces based on the electrostatic properties of water molecules in the interface computed using the polarizable TIP4P-FQ (Rick et al. 1994) FF. In addition, extended experimental benchmarks of binding free energies of peptides with different sequences on various functionalized SAM surface interfaces by Wei & Latour (2009, 2010) were used for optimization of the parameters for amino-acids.

Utesch et al. (2012) studied the adsorption of sulfite oxidase (SO) on a functionalized SAM surface under different ionic strength conditions using the CHARMM32 FF (MacKerell et al. 1998) for parameters of both the protein and the SAM surface. The catalytic activity of SO was earlier experimentally observed to be dependent on the flexibility of the tether connecting the molybdenum cofactor (Moco) domain and the cytochrome b5 domain of SO. In line with experimental findings, their results indicated that the adsorption of the enzyme’s cytochrome b5 domain is inhibited at high ionic strength (750 mM), whereas under much lower ionic strength conditions (100 mM), stable interactions with the surface take place, leading to a loss of flexibility of the tether.

Effects of water on adsorption on SAMs have been reported in several studies. Residence times ($\tau_r$) and self-diffusion coefficients ($D_S$) of interface water molecules on several different functionalized SAM surfaces of type S(CH$_2$)$_n$-F were inspected by Wang et al. (2010a, b) and found to be in the order: $\tau_{\text{COOH}} > \tau_{\text{NH2}} > \tau_{\text{OH}} > \tau_{\text{CH3}} > \tau_{\text{bulk}}$ and $D_{\text{COOH}} < D_{\text{OH}} < D_{\text{NH2}} < D_{\text{CH3}} < D_{\text{bulk}}$. This ranking was found to be directly correlated with the number of hydrogen bonds formed between the interfacial water and the modified SAM surfaces. Further, it was shown that the structure and the dynamics of the water determine the adsorption behavior of amyloid-$\beta$ peptide on the modified SAM surfaces and that the energy barrier for the adsorption of the peptide is lower on hydrophobic CH3-SAM than on the hydrophilic OH-SAM.

The mobility of interfacial water on the zwitterionic sulfobetaine (SBT)-SAM surface was investigated in a separate study by Xie et al. (2012). They found that the lower density packing and the higher flexibility of the SBT-SAM allowed water molecules to penetrate into the surface, thereby increasing the number of interactions of SBT-SAM with the water molecules and decreasing the mobility of the interface water molecules significantly compared with that of pure bulk water ($D_{\text{SBT}} < D_{\text{bulk}}/20$). The mean force-distance profile showed that the affinity of the SBT-SAM surface towards neuromedin-B peptide was smaller than for the other two SAM surfaces (CH$_3$-OH-SAM). The observations support the experimental results showing the anti-fouling properties of SBT-SAM surfaces (Xie et al. 2012).

7.6 sp$^2$-Carbon surfaces

The ability of a biomolecular FF to predict the interaction properties of a single water molecule described by a standard TIP3P or TIP4P model, and several ions on CNT and fullerene surfaces was explored by Schyman & Jorgensen (2013). Adsorption energies, as well as the binding site on the surface and the orientation of the water molecule obtained from the simulations using MM FFs, were compared with those of DFT calculations. The non-polarizable OPLS-AA FF was found to be adequate for the description of water properties on benzene (C$_6$H$_6$) and coronene (C$_{24}$H$_{12}$). However, with increasing surface size above (C$_{54}$H$_{18}$) and especially for C$_{60}$, the simulations became inaccurate, indicating that surface polarization played an
important role. Indeed, a polarizable biomolecular FF with induced dipoles on all non-hydrogen atoms (OPLS-AAP) was shown to yield reasonably good agreement with QM calculations for the interaction properties of water molecules and ions. Another polarizable FF, AMOEBA (Ren & Ponder, 2002, 2003), has been extended for the description of the non-bonded interactions between peptides and CNT (De Miranda Tomasio & Walsh, 2007) and graphite (Walsh, 2008), and exploited for simulations of peptide binding on CNT (Tomásio et al. 2009). These studies demonstrated that surface polarization caused a strong interaction between the carbon surface and aromatic rings of the side chains of the residues, particularly of tryptophan, placed in the surface plane, which caused strong binding of peptides rich in aromatic residues. The binding of peptides on CNT was also shown to be affected by the interplay between the ability of the aromatic rings to align on a non-flat surface and the total peptide-surface contact area that is increased due to surface curvature. The importance of π-stacking between aromatic side chains and carbon surfaces was also shown in a study of peptide adsorption on a single-layer graphene substrate by Akdim et al. (2013) using the polarizable AMOEABIO-09 FF implemented in TINKER 5.1. (Saint Louis, Washington. Software Tools for Molecular Design.). Although these studies demonstrate that polarizable MM FFs are able to provide reasonably accurate descriptions of the interaction properties of peptides on carbon substrates, such simulations are still computationally very demanding for the use of protein adsorption simulations. A viable approach is the inclusion of polarizability in graphene simulations via the rigid rod model, which was employed by Hughes et al. (2014) to calculate the adsorption free energies of single amino acids on a aqueous graphene surface.

The role of the water in peptide binding to a graphene surface was explored by Camden et al. (2013), who simulated GXG tripeptides for 20 amino acid residues (X) with an explicit water model. In the TEAM FF used in these simulations, the carbon atoms of the graphene surface were not polarizable, and thus important polarization-driven interactions between aromatic groups and the surface were missing. However, this study demonstrated that desolvation effects may play a very important role in peptide binding to a graphene surface. The dynamics of water on carbon-based hydrophilic graphite-COOH and hydrophobic graphite-CH₃ surfaces was studied by Li et al. (2005) using the standard OPLS-AA FF. They observed that the diffusion of water was slowed dramatically in the vicinity of the surfaces, and this effect extended up to 15 Å from the surface. The water diffusion coefficient within a 3 Å distance from a hydrophilic surface (-COOH) was 4 times smaller than that of a hydrophobic surface (-CH₃).

Implicit solvent models offer an alternative way to explore biomolecule-carbon interfaces in aqueous solution. It can be expected that the hydration shell on carbon surfaces has a much less defined structure than on hydrophilic surfaces and thus, is better suited for a continuum model. Indeed, no strong influence of structural water on the conformation of a test peptide was observed in evaluation of implicit and explicit water models (Walther et al. 2001). In an implicit solvent model, however, one has to account for hydrophobic desolvation effects, which may promote binding of biomolecules to carbon surfaces, explicitly. This was demonstrated by Mereghetti & Wade (2011) in a study of hydrophobin adsorption on graphite.

Raffaini and Ganazzoli reported simulations of protein adsorption on graphite surfaces in a series of studies using the consistent valence FF (CVFF) (Dauber-Osguthorpe et al. 1988) with a Morse potential applied to bonded interactions (Raffaini & Ganazzoli, 2003, 2004a, b, 2006, 2007, 2010). In their simulations, many proteins were observed to denature partially or completely upon adsorption to the hydrophobic surface.

Raffaini & Ganazzoli (2010) also carried out kinetic calculations for lysozyme spreading on a graphite surface by fitting the time evolution of the interaction energies, protein-surface center of mass distance and components of the radius of gyration of the protein to a function consisting of an exponential term and a stretched exponential term (Kohlrausch function):

\[ \gamma = A + B \exp\left(-\frac{t}{t_0}\right) + C \exp\left[-\left(\frac{t}{t_1}\right)^\delta\right] \]

where \( \gamma \) stands for an arbitrary quantity that decreases with time \( t \), \( t_0 \) and \( t_1 \) are relaxation times, and \( A \), \( B \), \( C \) and \( \delta \) are fitting parameters. In one simulation, they found that, upon adsorption, lysozyme underwent a liquid-like spreading on the surface suggested by a very fast initial decrease of distance between the center of mass of the protein and the surface with \( t_0 = 28 \) ps and by loss of most of the secondary structure. The initial kinetic step was followed by a longer stage with a \( t_1 \sim 1 \) ns in which all the secondary structure was lost. The authors point out that the fast kinetics of spreading, observed in the study may be attributable to the implicit solvent model used in the simulations.
8. Coarse-Grained molecular mechanics modeling of protein–surface interactions

Protein adhesion on a surface is often a combination of hierarchical processes occurring on a wide range of different time-scales. Moreover, an understanding of adhesion processes in real life applications often requires considering the interplay between protein–protein and surface-protein interactions. Thus, increasing the time- and length-scales of simulations beyond the all-atom MD limits, while keeping the atomic details of the protein-water-surface interface in the simulation model, is required. Coarse-graining methodologies offer a trade-off between computational speed and accuracy (Tozzini, 2005). However, apart from the established limitations of the CG models developed for biological molecules, there are other fundamental problems in developing and applying a CG FF for adsorption processes. One is the reliable modeling of the hydration shell on a surface that depends strongly on the chemistry of the surface and the surrounding solution. Taking into account that even standard all-atom FFs have notable deficiencies for modeling the surface-water interface, developing a CG water model to satisfactorily describe all accessible water states, and yet reduce the number of degrees of freedom, is challenging.

CG simulations are often employed in studies of adsorption of model homopolymer chains on surfaces. Due to the greater structural and dynamic complexities of peptides and proteins, the physical characteristics obtained from these polymer simulations cannot be directly extrapolated to peptide adsorption and spreading on a surface. To build a simplified description for proteins, CG models of protein-like AB polymers, in which each of the monomers is represented with types A (with hydrophilic properties) or type B (with hydrophobic properties) beads, have been introduced and applied to study adsorption (Liu & Chakrabarti, 1999).

Studies of CG models with improved system descriptions have been reported for protein adsorption on flat surfaces and on nanoparticles. The binding of the negatively charged proline rich salivary protein, PRP-1, to a negatively charged surface was studied by MC simulation using a model in which each residue was described by one bead that was either positively or negatively charged, or uncharged (Skepö et al. 2006). The protein changes conformation depending on salt concentration, and therefore, an implicit solvent model was used in which ions were treated explicitly. The effects of salt on protein conformation and consequently on adsorption properties were investigated in the simulations. Similar united residue models for peptide chains were used in several other papers (Evers et al. 2012; Knotts et al. 2005, 2008; Skepö, 2008; Xie et al. 2010). In two of these studies, short-range attraction potentials were added to model the binding of statherin protein on pure hydrophobic, charged, and mixed surfaces (Skepö, 2008), and the binding of β-casein on hydrophobic and charged surfaces (Evers et al. 2012). Further, MC simulations of bovine serum albumin (BSA) and HSA on silver nanoparticles using a CG model elucidated conformational changes of the protein upon adsorption, which were found to be in agreement with measured tryptophan adsorption spectra (Voicescu et al. 2012).

Tavanti et al. applied a Gō-type model, which takes only the interactions that are present in a molecule in its natively-folded structure into account, to study corona formation by ubiquitin molecules on gold nanoparticles of various sizes (10, 16, 20 and 24 nm diameter) in the presence of both implicit and explicit citrate (Tavanti et al. 2015). They used a FF and parameters developed originally for protein and RNA folding described by (Clementi et al. 2000; Pincus et al. 2008) for ubiquitin, and the parameters described by Ding et al. (2013) for citrate. With this setup, they showed that the aggregation characteristics of the proteins depend on the size of the nanoparticle and that a loss in secondary structure of the proteins is more prominent in nanoparticles of smaller size, i.e. 10 and 16 nm diameter. Further, they observed that the protein corona starts forming together with a slow phase of protein reorientation on the nanoparticle surface to optimize interactions on the nanoparticle surface.

In several studies, the MARTINI FF (Marrink et al. 2007; Seo et al. 2012), in which each residue is described by several beads, was employed to investigate the adsorption of proteins to surfaces (Griepenku et al. 2008; Liang et al. 2012) and to nanoparticles (Hung et al. 2011). MARTINI, which was originally developed for lipid bilayers, has been extended for simulation of proteins in solvent by applying an elastic network model to keep the protein near the native protein structure.

Most of the studies mentioned above employed CG FFs evaluated for biomolecules or surfaces alone, and hence their accuracy for modeling the protein-solvent-surface interface was not evaluated. Moreover, none of these CG models represents all the effects of the surface hydration shell that can be explored in atomic detail MD simulations. A step towards the inclusion of such solvent effects was performed by Carrillo-Parramon et al. (2013), who exploited the surface desolvation terms as defined in the ProMetCS implicit solvent FF in an essential dynamics CG model of ubiquitin interacting with a gold surface. Their model reproduced fluctuation characteristics of the protein obtained with classical atomistic MD.

Models developed specifically for protein structures near interfaces have been proposed in several studies. Bhirde et al. (2014) modeled the aggregation of gold nanoparticles in an albumin solution using a rigid-body CG model of inter-nanoparticle, inter- and intra-protein and nanoparticle–protein interactions consisting of a 12–6 Lennard–Jones term for all three types
of interactions and an electrostatic term for protein–protein interactions based on the screened Coulomb potential implicit solvent model (SCPISM) (Cardone et al. 2013; Hassan & Steinbach, 2011). They investigated the role of nanoparticle size and coating of its surface in adsorption of albumin by performing a number of MC simulations with different resolutions of coarse-graining and different values of the Lennard–Jones parameters (σ and ε) and the nanoparticle concentration. This study showed that the albumin molecules formed aggregates with the nanoparticles, and the stability of these aggregates depended on the surface coating characteristics and the nanoparticle concentration. Wu and Narsimhan proposed an implicit solvent CG model in which 1–3 beads were used per protein residue, and applied it to the unfolding of lysozyme on a neutral silica surface, investigating the dependence on temperature and ionic strength (Wu & Narsimhan, 2009). The authors evaluated their approach by comparing with all-atom MD simulations of protein dynamics in solution.

Ravichandran & Talbot (2000) investigated the kinetics of adsorption and the structure of the adsorbed layer of hen egg white lysozyme (HEWL) on a solid surface. The HEWL molecules were represented as spheres with a uniform charge distribution on each of the molecules. The authors pointed out that the simplified description of the molecules would not work with proteins that are not globular, or that have an anisotropic charge distribution. In a separate study, Ravichandran et al. (2001) investigated the early stages of the adsorption process and the binding orientations of HEWL on a positively charged surface with a full atomic detail representation of the protein. Although HEWL has an overall positive charge, it could bind to the charged surface due to the non-uniform distribution of the charges on the protein. Their result shows the necessity of less CG representations in certain systems.

9. Applications of sampling methods to protein–surface interactions

Simulations of biological molecules give many insights into the molecular mechanisms of interactions and association, as well as their thermodynamic properties and kinetics. To perform a simulation, descriptions of the structures present in the system (e.g. a set of coordinates of the atoms for all-atom models), a reliable FF and a sampling method are required. The sampling technique used in a simulation should be chosen according to the properties of the system to be investigated. In this section, we will discuss advantages and disadvantages of conventional and enhanced sampling techniques in the context of their application to protein–surface interactions.

9.1 Molecular dynamics

Among the different sampling methods available, classical MD is by far the most widely used technique to investigate the characteristics of protein–surface interactions. Examples include simulations for structural refinement of docked complexes (Aliaga et al. 2011; Alvarez-Paggi et al. 2013; Brancolini et al. 2012, 2015; Imamura et al. 2007), and for the investigation of the binding orientations of proteins on surfaces (Alvarez-Paggi et al. 2010; Boughton et al. 2010; Coppage et al. 2013); the kinetic mechanisms of adsorption (Raffaini & Ganazzoli, 2010); the ET pathways and properties of adsorbed proteins (Bizzarri, 2006; Siwko & Corni, 2013; Utesch et al. 2012; Zanetti-Polzi et al. 2014; Zhou et al. 2004); the effects of pH (Emami et al. 2014b; Imamura et al. 2003; Tosaka et al. 2010; Utesch et al. 2013) and ionic strength (Bizzarri, 2006) on adsorption; the role of ions in mediating adsorption (Wu et al. 2013); and structural and energetic aspects of adsorption of proteins on surfaces (Apicella et al. 2013; Jose & Sengupta, 2013; Hoeflimg et al. 2011; Hung et al. 2011; Kubiak-Ossowska & Mulheran, 2010a, b; O’Mahony et al. 2013; Vila Verde et al. 2009, 2011; Wang et al. 2010a, b; Yu et al. 2012a; Steckbeck et al. 2014; Sun et al. 2014a, Sun et al. 2014b). Further, conventional MD can be used to simulate physical perturbations, such as mechanical or electrical forces exerted on molecules in experiments. Examples include voltage-biased MD simulations of peptides on a gold surface to validate the signals resulting from a surface enhanced Raman spectroscopy (SERS) experiment (Chen et al. 2012). Mimicking the experimental conditions helps to understand how experiments work and how their results should be interpreted.

Classical MD simulations are very helpful in understanding many molecular phenomena at atomic detail. However, the capabilities and limitations of the MD method, as of other sampling methods, must be known in order to set up a simulation and interpret the results obtained properly. When comparing MD results with the thermodynamic and kinetic properties of the same system obtained from experiments, it should be remembered that a trend observed in an MD simulation may not always agree with the expected average properties. A system kinetically trapped in an energy well, may stay in its metastable state for the rest of the simulation and appear as if in equilibrium (Wei et al. 2011). In fact, a single simulation corresponds to a single trajectory of a system, and it may not always ensure an adequate sampling. Drawing conclusions from a statistically inadequate sampling will thus often lead to incorrect interpretations. To improve the statistical significance of their sampling, Penna et al. (2014) chose a direct approach and performed more than 240 explicit solvent MD simulations for the investigation of adsorption mechanisms of a platinum-binding peptide and a gold-binding peptide on neutral Pt(111) and Au (111) surfaces, respectively. From these simulations, they obtained statistics on many binding characteristics, such as
anchoring events between each of the peptide residues and the surfaces and the distribution of transition times from the anchoring phase to the initial lockdown phase.

Another important sampling problem is the simulation time required to reproduce an adsorption event on a molecular scale. Even though a simulation that captures the dynamics of a system for several nanoseconds is helpful for structural refinement of molecules or for understanding the initial stages of a process, many molecular events take place over much longer time scales. In their paper, Wei et al. (2011) stressed the necessity of very long simulations to be able to investigate a complete adsorption process on a surface in detail. To bypass the energy barriers that normally require long simulation times to overcome, Emami et al. (2014b) employed 1 ns-long annealing simulations at 500 K prior to their production simulations at 298 K. From these classical MD simulations, they investigated the adsorption of three peptides differing in net charge to silica surfaces under four different pH values (with changing ionization levels of the silica surface depending on the pH value). The authors showed that the times spent by each of the peptides on the silica surfaces in the simulations could be used as a relative measure of adsorption strength as a function of pH (see Fig. 10). Even though the trends in adsorption level with pH for each of the peptides from the simulations and the experiments are in good agreement, the time spent on the surface by each of the peptides fails to fully reproduce the difference in relative adsorption levels between the negatively charged and overall neutral peptides measured experimentally.

To improve the sampling characteristics of conventional MD simulations, replica exchange MD (REMD) (Sugita et al. 1999) was developed based on a MC simulation method called parallel tempering (Hansmann, 1997) (see the following section). In conventional MD simulations, structures may become trapped in local energy minima on the potential energy landscape, hence, sampling of configurational phase space is limited. To overcome high energy barriers, the REMD method makes use of sampling of the configurational space with independent MD simulations (i.e. replicas) of a system at different replica conditions, e.g. at different temperatures or with a different Hamiltonian. By exchanging the replicas, configurations that are trapped in local energy minima can explore other parts of the energy landscape and an improved Boltzmann-weighted sampling can be achieved. REMD methods are useful to sample a large set of configurations in different potential wells separated by high energy barriers, which otherwise is not possible in a classical MD simulation. Indeed, temperature-based REMD (T-REMD) simulations were used in several studies to accelerate the sampling of peptide–surface interactions (Corni et al. 2013; Li et al. 2011; Oren et al. 2010) and of interactions of basic fibroblast growth factor (bFGF), a small protein, with a hydroxyapatite (001) surface (Liao & Zhou, 2014). Liao and Zhou observed that while the protein displaces the surface hydration shell and binds tightly to the surface in the T-REMD simulations (five replicas in the range from 310 K to 2500 K), bFGF did not contact the surface directly in classical MD simulations at 310 K (Liao & Zhou, 2014). However, the T-REMD method is usually not applicable to large systems, e.g. a large protein and a surface with explicit solvent, due to the number of replicas required which scales as $O(f^{1/2})$, with $f$ being the system’s total number of degrees of freedom (Fukunishi et al. 2002; Wright & Walsh, 2013).

Typically, simulation of a protein-surface system in atomic detail requires solvation in tens of thousands of water molecules. The large number of water molecules becomes problematic because the water–water interaction energy term ($E_{\text{ww}}$) dominates the energy terms due to interactions between the solute molecules and between the solute and water molecules, thereby demanding more replicas in a temperature-based REMD simulation (Huang et al. 2007). To overcome the poor scaling of the T-REMD method with the size of the system, replica exchange with solute tempering (REST) was proposed (Liu et al. 2005). In this approach, the potential energy function is tailored such that the $E_{\text{ww}}$ term is eliminated from the acceptance probability.

This increases the chances of acceptance of exchanges between replicas compared with regular TREMD simulations. Wright & Walsh (2013) investigated the transferability of the REST method to peptide–surface interactions by using a quartz binding protein and a fully hydroxylated α-quartz system as a benchmark for the REST and the T-REMD simulations. The results showed that the REST approach was 50% less expensive in terms of CPU time than the T-REMD method using the same simulation parameters. Furthermore, up to about 80% of the CPU time could be saved with the REST method, owing to the smaller number of replicas used in the simulations. Using the REST parameters optimized for peptide–surface interaction simulations (Wright & Walsh, 2013), Tang et al. (2013) performed REST simulations for numerous gold binding peptides to obtain an accurate ensemble of peptide conformations adsorbed on a gold surface.

The REMD and REST methods sample the phase space with a canonical distribution (fixed number of atoms, volume and temperature) and therefore fall short in exploring the very low probability states. Although investigating the dynamics of a system does not necessarily require an extensive exploration of these states, it may lead to erroneous results, particularly when calculating free energies. This problem can be greater for systems with peptides strongly interacting with surfaces (Wright & Walsh, 2013). To obtain more exhaustive sampling of the phase space, replica exchange methods can be used in conjunction with methods employing biasing potentials such as metadynamics (Bussi et al. 2006a).
Monte Carlo methods

MC methods are stochastic techniques used to solve problems that obey certain probability functions. MC methods have long been used in molecular simulations as an alternative to MD, owing to their efficiency in sampling of conformational ensembles with a Boltzmann distribution. From these ensembles, the geometric averages and thermodynamic properties of a system can be approximated (Schlick, 2002). The MC methods applied to biomolecular systems have been reviewed (see Hansmann & Okamoto, 1999; Taylor et al., 2002), and information on MC methods can be found elsewhere (Landau & Binder, 2009).

Commonly employed for atomistic and CG simulations of protein-surface systems, Metropolis sampling (Metropolis et al., 1953) is a simple and yet very efficient Markov chain Monte Carlo (MCMC) scheme. The Metropolis scheme is used to create a random walk on a set of points with an importance-weight based on a probability distribution (Frenkel & Smit, 2001). MC simulations with the Metropolis algorithm have been used to explore the adsorbent conformations and protein denaturation on the surfaces (Anand et al., 2009; Castells et al., 2002; Euston & Naser, 2005; Skepö, 2008; Skepö et al., 2006; Zhdanov & Kasemo, 2000a), to investigate ordering of the proteins adsorbed on the surface (Zhdanov & Kasemo, 2000b), and to calculate the binding free energy of adsorption (Lund et al., 2005).

Al-Mekhnaqi et al. (2009) applied MC simulations with an energy minimization scheme to predict the conformations of a HSA subdomain A on a graphite surface with transition probabilities determined using the Metropolis method. The protein structure was modeled with a united atom approximation, and the structures obtained during MC simulations were minimized using a direct search scheme, which is a technique used for solving optimization problems without constraints. The simulations started with the peptide chain fully extended. In each step, one dihedral angle was chosen randomly and given a random value. After an energy minimization step, a new step was performed and this conformation was accepted with the Metropolis criterion. Simulations were ended after a predefined number of perturbations was achieved. Using this method, they were able to reproduce very similar conformations of the albumin subdomain adsorbed to a graphite surface to those obtained by all-atom molecular simulations (Raffaini & Ganazzoli, 2003).

The Metropolis MC method can fail in sampling the configuration space of a system if the system becomes trapped in a low-energy state at low temperatures. Therefore, typically, many independent Metropolis MC simulations of a system are carried out. Alternatively, an energy-based MC method, parallel tempering Monte Carlo (PTMC), (Swendsen & Wang, 1986) can be used. It is the basis for the REMD method. Every time a certain number of MC moves have been performed, the configurations are exchanged between a replica and its neighboring replicas that have the closest temperature values. The exchange operation is ruled by the Boltzmann distribution and done with the Metropolis acceptance probability. The PTMC method has been applied to protein–surface interactions in several studies, including prediction of the binding orientations of proteins (Xie et al., 2010), and investigation of thermal and mechanical properties of proteins tethered on surfaces using a hybrid PTMC method with MD (Knotts et al., 2005).

Xie et al. (2010) compared the performance of the PTMC method with that of conventional serial Metropolis MC simulations of lysozyme on charged surfaces using the same parameters. Figure 11a shows the distribution of the orientations of the lysozyme on a negatively charged surface obtained from the different MC methods. A comparison of these distributions with the
The adsorption energy landscape (Fig. 11b) shows that with the PTMC method, the two lowest energy minima were found. In the four serial MC simulations, on the other hand, each simulation revealed a different energy minimum: two local energy minima in addition to the two lower minima found with PTMC.

In two other studies, the same group applied the PTMC method to obtain the most probable orientations of fibronectin on a HAP surface (Liao et al. 2014) and of protein G on -COOH and -NH₂ functionalized SAMs (Liu et al. 2013). These orientations were used as starting orientations for MD simulations to investigate the adsorption mechanisms of the proteins.

Finally, basin-hopping Monte Carlo (BHMC) (Li & Scheraga, 1987; Wales & Doye, 1997) is worth mentioning, since it is one of the most reliable MC methods for searching the lowest energy configurations of molecules (Iwamatsu & Okabe, 2004). Although the BHMC method has been successfully applied to conformational sampling of peptides in implicit solvent (Strodel & Wales, 2008), to the best of our knowledge, this method has yet to be employed for peptides or proteins near inorganic surfaces.

9.3 Brownian dynamics

The diffusive dynamics of a Brownian particle suspended in a solution can be modeled by the BD technique. BD is commonly applied for simulation of diffusion-driven binding processes leading to the formation of ‘diffusional encounter complexes’. Further information on details of BD methods, and their applications can be found in other reviews (Allison et al. 1986; Gabdoulline & Wade, 2002; Madura et al. 1994).

BD has been used in many studies of adsorption of proteins on surfaces with CG (Ravichandran & Talbot, 2000) and all-atom (Ravichandran et al. 2001) representations of the proteins. Simplified BD simulation methods designed for CG models have been proposed (De la Torre et al. 2009; Gorba & Helms, 2003) and one (Gorba & Helms, 2003) was tested for the diffusional dynamics of cytochrome c on a charged surface. As an alternative to CG models, several BD studies of protein adsorption to different surfaces have been performed with rigid, atomic detail models of the solutes (Brancolini et al. 2012; Mereghetti & Wade, 2011; Kokh et al. 2010). Mereghetti and Wade (Mereghetti & Wade, 2011) applied BD to the simulation of hydrophobin proteins on a graphite surface. They explained the high affinity of the protein for the graphite surface from the average proximity and orientation of hydrophobin on the surface with its hydrophobic face towards the surface obtained from the simulations.

Brancolini and coworkers employed BD simulations for docking of ubiquitin (Brancolini et al. 2012) and of human β2-microglobulin (Brancolini et al. 2014, 2015; Brancolini et al. 2015) on bare, citrate covered and thiol-protected gold nanoparticles using the ProMetCS model (Kokh et al. 2010). The most abundant orientations of ubiquitin on the gold obtained from BD simulations were used as starting orientations for fully flexible MD simulations. By using a combination of BD and MD methods, they were able to reduce the number of possible initial orientations for MD simulations, evaluate the stability of each orientation with MD simulations and observe the dynamics of adsorption in detail.

Overall, BD is an effective method for simulation of biomolecules as well as interactions of molecules with surfaces. An implicit representation of the solvent and a rigid-body treatment of the molecules help to reduce the time required for calculations, allowing thousands of simulations to be performed within a few hours with current computer technology. However,
while the rigid-body model may be appropriate for globular proteins, it is less suitable for peptides and unstructured proteins. For these, flexible CG models of the solute can be employed in BD simulations.

10. Applications of free energy calculation methods to protein–surface interactions

The computation of binding free energies is used for determining binding affinities between proteins and surfaces, as well as the kinetics and transition free energies of adsorption processes. Free energy calculations are, therefore, important for bridging between theoretical and experimental studies. To date, many methods have been developed, which differ in their accuracy and complexity, for use in molecular simulation (Swanson et al. 2004). Most of these methods employ MD or MC techniques for sampling of the phase space. The free energy calculation methods can be classified as end-point methods and pathway methods. In end-point methods, only the reference unbound and the final bound states are sampled to obtain the free energy difference between the states. A commonly used end-point method is the MM Poisson–Boltzmann Surface Area (MM-PBSA) method (Kollman et al. 2000; Srinivasan et al. 1998). The MM-PBSA method has been applied to a horse heart cytochrome c-COOH-SAM surface system to compute the adsorption free energy (Alvarez-Paggi et al. 2010). Although easy to implement and perform, this method was shown to have a relatively large error range (Singh & Warshel, 2010) and variable performance (Hou et al. 2011) for protein–ligand interactions, and to require long MD simulations and extensive PB calculations.

A computationally faster, though less accurate, alternative to MM-PBSA, is MM Generalized-Born Surface Area (MM-GBSA) (Kollman et al. 2000; Srinivasan et al. 1998). This method is based on solving the GB equation, which provides the solvation free energy of each individual atom and, therefore, is suitable for modeling the interactions of flexible solutes. Guo et al. (2014) computed binding free energies of three different proteins on a graphene surface from classical MD simulations using MM-GBSA. For each of the protein-graphene systems, they performed two separate simulations, each with a different protein orientation with respect to the surface. The results showed that the calculated free energy values for the two simulations may vary up to 65 kcal mol$^{-1}$ with standard error values around 11 kcal mol$^{-1}$. Although these calculations give insights into the adsorption strength of a binding mode from a trajectory, a better sampling of configuration phase space is required for MD simulations. To reduce the effect of the insufficient sampling on the energy calculations, Xie et al. (2014) used the REMD method for sampling the adsorption of the Ab peptide to fullerenes. After clustering of the conformers, the MM-GBSA method was employed to obtain binding free energies for the largest clusters.

The end-point methods provide good approximations to free energy differences. Pathway methods are formally exact. Based on a predefined reaction coordinate, pathway methods employ either non-physical (alchemical) or physical intermediates. A commonly used method is free energy perturbation (FEP) (Mori et al. 2013). Although FEP has been applied to study binding of a formate anion (HCOO$^-$) to a rutile surface (Mori et al. 2013), to the best of our knowledge, this method has not been applied to study protein–surface interactions. Other pathway methods based on physical reaction coordinates and applied to protein–surface interactions are discussed below. These methods can be categorized into two groups depending on how the system is sampled: equilibrium and non-equilibrium simulation methods.

10.1 Equilibrium methods

A commonly used equilibrium sampling technique to obtain a free energy profile of a system is umbrella sampling (Torrie & Valleau, 1974, 1977). In this method, a reaction coordinate is defined that connects two thermodynamic states for which the free energy difference is computed. The reaction coordinate used is usually selected based on a geometric entity, e.g. distance, angle, etc. The reaction coordinate is then divided into distinct windows in which a bias (umbrella) potential is applied to keep the system near that coordinate point by means of restraints. For each window, a separate simulation of the system is performed for sampling around the corresponding coordinate point. Then, the simulations are combined with a reweighting procedure since they are performed in biased ensembles. Two reweighting procedures are weighted histogram analysis method (WHAM) (Kumar et al. 1992; Souaille & Roux, 2001) and umbrella integration (Kästner & Thiel, 2005). Using the final probability distribution of the configurations, the PMF and hence, the free energy profile of the system, is obtained.

The umbrella sampling technique has been successfully employed for the calculation of the free energy of adsorption of a model surfactant on hydrophobic and hydrophilic silica surfaces (Xu et al. 2008), nanoparticles on phospholipid membranes (Li & Gu, 2010), ions on hydrophobic surfaces (Horinek et al. 2008) and amino acids on a ZnO surface (Navrocki & Cieplak, 2013). Umbrella sampling combined with WHAM has also been applied to peptide–surface interactions. Examples include adsorption of various tripeptides to a CH$_2$-SAM (Sun et al. 2007), and an RGD peptide to a titanium oxide surface (Schneider & Colombe Ciacchi, 2010). Although, as in these two examples, the distance of the peptide from the surface is typically used as the reaction coordinate for sampling, other definitions can also be employed. For instance, Boughton
et al. (2010) used the angular orientation of the \(\alpha\)-helix of magainin 2, an \(\alpha\)-helical amphiphilic peptide, on a polystyrene surface as a reaction coordinate to compute its adsorption free energy by umbrella sampling combined with WHAM.

Umbrella sampling, even though it restricts the spatial sampling necessary to calculate free energy differences, can still be subject to conformational sampling problems. Wang et al. (2008) therefore combined umbrella sampling with a biased REMD technique. In this method, an initial umbrella sampling of the system is employed to obtain a rough estimate of the PMF profile of the system. The negative of the PMF is then used as a biasing potential for REMD simulations. A final PMF profile is constructed from a probability-ratio method analysis for the REMD simulations. Comparison of the estimated PMF profile of a \(\mathrm{Na}^+\) ion on a carboxylic acid functionalized SAM surface with the theoretical profile, showed that the biased REMD gives much better agreement with the theoretical profile compared with those obtained from conventional REMD and umbrella sampling simulations. The method was applied in separate studies to adsorption of the G_{44}-K-G_{44} peptide on a polylactic acid surface for the investigation of binding free energy profiles (O’Brien et al. 2008), for the assessment of the protein FF parameters to be used in protein-surface simulations (Yancey et al. 2010), and for investigation of the effect of pressure on the calculation of protein-surface binding free energies from MD simulations (Yancey et al. 2010).

Another common approach used in molecular simulations to calculate the free energy landscape as well as to investigate new reaction pathways is metadynamics (Bussi et al. 2006b; Laio & Gervasio, 2008; Laio & Parrinello, 2002). In this approach, the dynamics of the system is steered by thermodinamic forces and compensated by a history-dependent biased potential (Laio & Parrinello, 2002). Owing to the bias potential, the frequency at which a state is visited during a simulation typically decreases linearly with its free energy (Bussi et al. 2006a). Accurate free energy profiles can be obtained from metadynamics with an appropriate reweighting scheme (Bononi et al. 2009) to recover an unbiased probability distribution (Meissner et al. 2014). To date, metadynamics has been successfully applied for calculation of the free energy of binding of a formate anion to a TiO\(_2\) surface (Mori et al. 2013) and of amino acids to silver and gold surfaces (Palafox-Hernandez et al. 2014), and for calculation of the free energy landscapes of alanine dipeptide in its free and gold surface bound states from 2μs-long simulations (Bellucci & Corni, 2014).

The combination of metadynamics with other enhanced sampling techniques allows a better exploration of low probability states, hence increasing the accuracy of the free energy calculations (Bussi et al. 2006a). Qin & Buehler (2014) investigated the adsorption free energy of mussel adhesion proteins onto a silica surface using metadynamics starting from configurations obtained from REMD simulations.

Schneider & Colombi Ciacchi (2012) have successfully applied a hybrid REST and metadynamics to study the adsorption of small peptides on Si and Ti surfaces. With only four replicas at temperatures of 300, 350, 400 and 450 K, they were able to obtain agreement with the experimental adsorption free energy of the RKLPA peptide on a Ti surface (experimental: 38-0 ± 8 kJ mol\(^{-1}\), computed: 38-6 ± 3-9 kJ mol\(^{-1}\)). The same approach was employed recently by Meissner et al. (2014) to estimate the circular dichroism (CD) spectra of a peptide in its free state and bound to a silica surface. CD spectroscopy is a useful technique to monitor secondary structure content in biomolecules, in particular, \(\alpha\)-helix content in peptide structures. Meissner et al. calculated the CD ellipticity values of conformations from simulation snapshots and used these values as external collective values in their reweighting procedure. The estimated fractional helicity values of the free and adsorbed peptides obtained from the simulations show good agreement with experiments. Even though these results are very promising, application of this method to protein adsorption requires careful selection of parameters, e.g. adequate number of replicas, appropriate selection of temperatures.

Thermodynamic integration is another approach used in free energy calculations of protein–surface interactions. This method, like umbrella sampling, requires a predefined reaction path between the initial and the final states. Juffer et al. (1996) applied thermodynamic integration to compute the adsorption free energies of the enzyme cutinase and its variants to a charged surface in the presence of explicit ions. Hoelling et al. (2010a) applied the method to obtain the PMF profiles of amino acids adsorbing on an Au(111) surface. Further, Schneider & Colombi Ciacchi (2010) employed thermodynamic integration, along with the umbrella sampling/WHAM method, to compare the free energy profiles of adsorption of RGD peptides on oxidized Ti obtained by the two methods. Their results showed a perfect agreement between the results of the two methods. Finally, Friddle et al. (2011) used the adaptive biasing force simulation method, which is a technique based on thermodynamic integration to obtain free energy profiles using a biasing force, for the adsorption of a 12-mer C-terminal fragment of the amelogenin protein on various different crystal terminations of two different hydroxyapatite surfaces (\{100\} and \{001\}). By complementing AFM measurements with the free energy calculations, Friddle et al. (2011) were able to predict the crystal termination of the hydroxyapatite surface used in experiments and identify the interactions governing the adsorption of the amelogenin peptide to it.
10.2 Non-equilibrium methods

Observing adsorption/desorption events of large molecules using conventional MD methods requires very long simulation times. These limitations can be overcome by steered molecular dynamics (SMD) simulations (Utesch et al. 2011), which is the most commonly used non-equilibrium simulation technique for biomolecular systems. In a typical SMD simulation, a protein or a peptide is pulled with a non-physical external force along a predefined reaction coordinate with a constant velocity or with a constant force, thereby accelerating adsorption or desorption events. Further, SMD allows simulation of systems under mechanical stress such as stretching, shearing and bending, and can hence be used to predict the effects of external disturbances on protein-surface systems (Hamdi et al. 2008). These disturbances may lead to uncertainties in experimental measurements. Particularly, SMD simulations conducted to investigate the effect of the tip of an AFM showed that pushing a molecule towards a surface with an AFM tip will bias the results by increasing adhesion, as the force exerted on the molecule leads it to spread more on the surface (Horinek et al. 2008; Mücksch & Urbassek, 2011). With SMD simulations, the adsorption mechanisms of peptides and proteins on different types of surfaces have been investigated (Alvarez-Paggi et al. 2010; Dong et al. 2007; Emami et al. 2014b; Friedrichs et al. 2013; Hamdi et al. 2008; Shen et al. 2008; Utesch et al. 2011; Yang & Zhao, 2007) and compared with experimental measurements (Schneider & Colombi Ciacchi, 2010, 2012). In addition to accelerating the sampling of certain events, the SMD method is used to calculate free energy differences. It was shown by Jarzynski (1997) that a free energy difference can be obtained from a non-equilibrium process following the equation: $e^{-\beta\Delta F} = \langle e^{-\beta W} \rangle$ where $\beta = 1/(k_BT)$ ($k_B$ is the Boltzmann constant and $T$ is the absolute temperature) and $\Delta F$ and $W$ stand for the free energy difference and the non-reversible work done along a reaction coordinate, respectively. The average is taken on all (in principle) the realizations of the non-equilibrium process that connect the initial and final states of interest. Based on this equation, a free energy difference can be obtained using a population of SMD simulations that start from an equilibrium ensemble. Binding free energies of peptides and proteins to various different surfaces computed using the Jarzynski equation have been reported in a number of papers (Chen et al. 2009b; Kang et al. 2009; Mijailovic et al. 2013). Baier et al. (2014) investigated the adsorption free energies of ZnO-binding peptides on ZnO surfaces. To this end, they applied a hybrid steered MD with umbrella sampling method. In this approach, the peptides were steered towards the surface with a constant force applied to the peptides. After an equilibration period without any external forces applied upon adsorption, the peptides were pulled back from the ZnO surface. Selected conformers were then simulated for another 5 ns with a harmonic potential applied to their centers of mass. The PMF profiles were obtained using the WHAM method. Combined SMD simulations with umbrella sampling were also used in a study to calculate the adsorption free energy of a cationic peptide on a silica surface (Emami et al. 2014b).

11. Outlook and future directions

The modeling and simulation studies reviewed in this paper suggest that the adsorption of aqueous peptides and/or proteins to surfaces is governed by a number of properties that determine the strength and specificity of the interactions. These properties can be summarized as: pH, solvated ion types and surfactants, and ionic strength of the solution; ionization levels, physical character (i.e. polar, non-polar, charged, etc.), size, shape, thickness, structural and compositional homogeneity/heterogeneity, chemical modifications and molecular structure of the surface; and flexibility, physical character and intra–peptide interactions determined by the sequence and affinity of the peptide/protein. To derive design rules for proteins and surfaces with desired binding characteristics, the studies often oversimplify adsorption mechanisms, focusing on only one or a few of these properties as determinants of the specificity or the affinity of the proteins for the surfaces. The transferability of the simple design rules derived in these studies to systems with different surface/protein types under different solution conditions is, hence, to be considered with caution. To be able to draw a complete picture of these interactions and thus to draw universal design rules, one has to take all of these properties into account and investigate the significance of each of them in the system of interest thoroughly.

The current FFs used for biomolecular interactions were developed and optimized specifically for their interactions in aqueous environment. The major FFs used for simulations of biomolecule–inorganic surface interactions, on the other hand, are based on mixing the parameters of separately parameterized biomolecular and inorganic FFs. Although their applicability to inorganic interfaces has been tested and validated to a certain extent, many more studies are needed for an extensive calibration of the parameter sets. The number of experimental studies providing structural information on protein–surface interactions is currently limited. However, advances in experimental capabilities for application to these interactions are very promising, for instance probing the 3D structures of peptides adsorbed on metal oxides (Mirau et al. 2011) and identification of sites on ubiquitin engaged in binding gold nanoparticles (Calzolai et al. 2010) by means of nuclear magnetic resonance (NMR) techniques.
A bottleneck in the simulation of protein-surface systems is sampling. Undoubtedly, the most commonly used simulation method at present is classical MD. However, with classical MD simulations, complete sampling of the phase space and, therefore, the adsorption dynamics is not possible. Enhanced sampling methods (e.g., replica exchange methods) are therefore invaluable tools for capturing the less probable states of the system that might otherwise require dozens of classical MD simulations. Even though simulations of peptide–surface interactions by means of enhanced sampling methods have been reported in a number of studies (discussed in previous sections), simulations of protein–surface interactions with these methods are still often not feasible due to the large size and complexity of proteins. Further studies are necessary to develop simulation protocols, evaluate them and optimize suitable parameters.

Advances in the simulation of protein–surface interactions are tied to the general advances in simulation methods. For instance, we discussed the significance of the representation of a change in ionization states of SAM and oxide surfaces and hence the need for constant pH simulations earlier in this review. However, many constant pH simulations suffer from convergence issues and pose even more issues with explicit solvation (Mongan & Case, 2005). Therefore, improved simulation methods are required for accurate modeling and simulation of protein-inorganic surface systems.

Finally, none of the simulation techniques covered in this review is able to provide an accurate picture of protein adsorption events that take place on a large range of time and length scales by itself. Therefore, appropriate multi-scale modeling and simulation approaches should be developed and employed in a concerted manner.

Acknowledgements

We thank Dr. Neil J. Bruce for his careful reading of the manuscript.

Financial support

M. O. acknowledges the support of Heidelberg Graduate School of Mathematical and Computational Methods for the Sciences (HGS MathComp), Heidelberg University. S. C. acknowledges funding from MIUR through PRIN 2012A7LMS3003. Our work is supported by the Klaus Tschira Foundation.

References


