Prediction and clarification of structures of (bio)molecules on surfaces-a review

C. Schön„ C. Oligschleger, Juan Cortés

Prediction and clarification of structures of (bio)molecules on surfaces - a review

J. Christian Schön[a], C. Oligschleger[b], J. Cortes[c,d]

[a]Max-Planck-Institute for Solid State Research, Heisenbergstr. 1, D-70569 Stuttgart, Germany; e-mail: c.schoen@fkf.mpg.de
[b]University of Applied Sciences Bonn-Rhein-Sieg, von-Liebigstr. 20, D-53359 Rheinbach, Germany; e-mail: christina.oligschleger@h-brs.de
[c]Laboratoire d'Analyse et d'Application des Systemes CNRS, 7 Avenue du Colonel Roche, F-31400 Toulouse, France; e-mail: juan.cortes@laas.fr
[d]Universite de Toulouse, LAAS, F-31400 Toulouse, France

Dedicated to Professor Wolfgang Jeitschko on the occasion of his 80th birthday

Abstract:

The design of future materials for biotechnological applications via deposition of molecules on surfaces will require not only exquisite control of the deposition procedure. Of equal importance will be our ability to predict the shapes and stability of individual molecules on various surfaces. Furthermore, one will need to be able to predict the structure patterns generated during the self-organization of whole layers of (bio)molecules on the surface. In this review, we present an overview over the current state of the art regarding the prediction and clarification of structures of biomolecules on surfaces using theoretical and computational methods.

1 Introduction

Throughout its history, the field of synthetic chemistry has been divided into several subfields, depending on the way chemistry has been perceived by its practitioners. One of the first divisions, according to classes of compounds, was into organic and inorganic chemistry. Subsequently the latter was subdivided into solid state and molecular chemistry, while the formed spawned fields like polymer chemistry, biochemistry, or

metal-organic chemistry. Other divisions took their lead from the actual physical size of the material at hand, leading to a distinction between molecular, cluster, surface, or (bulk) crystal chemistry.

More recently, the field of nanochemistry with its focus on the synthesis of, and with, nanosized objects [1-4] has produced hybrid-like compounds, where e.g. complex molecules, clusters, or nano-size (mono)-layer flakes, are arranged in layered or three-dimensional patterns. Usually, one still employs classical chemical synthesis techniques to reach such nano-patterned compounds. But the development of advanced deposition techniques [5-7] has allowed us to conceive the generation of compounds, or materials in general, by adding one building block, i.e. molecule, cluster, etc., at a time [8]. A precondition for such a way to produce new compounds is the ability to predict and control the formation of molecular arrangements on well-defined surfaces.

The past couple of decades have seen a plethora of experimental studies of the deposition of molecules of various sizes, ranging from simple gases like N_2 [9], CO [10], Xe [11], etc., to highly complex biomolecules such as proteins [12], spurred at least partly by new measurement probes such as scanning tunneling microscopy (STM) [13] or atomic force microscopy (AFM) [14]. But just as it had originally been the case in other fields of synthetic chemistry, such as molecular or solid state chemistry, the outcome of such a deposition with respect to the structure of the individual molecule or the structural arrangement of many molecules is very hard to predict without prior experimental input. Empirical rules and heuristics can be used to "explain" the structures observed after deposition. But we are still far away from an unbiased structure prediction of thermodynamically and/or kinetically stable structures of molecules on surfaces. Similarly, the a priori design of optimal deposition routes leading to desired specific patterns of molecules on surfaces is still farther in the future.

Or are we? At least with regard to the prediction of (meta)stable structures, we know from our experience with the analogous problem in the fields of molecular [15,16] and solid state [17-19] chemistry, how to solve this problem, in principle: we need to globally study the energy landscape of molecules on surfaces, and identify kinetically and thermodynamically stable regions on this landscape. These so-called locally ergodic regions [18-21] correspond in many cases to local minima on the landscape plus their surrounding basins. Similarly, the powerful global search techniques that have been

applied to, and sometimes specifically developed for, the structure prediction of crystals and single molecules in vacuum or solvent can also be employed for the prediction of the structure(s) of molecules on surfaces.

One curious difference between molecules on surfaces and e.g. in crystals is the amount of firm atom-level structure information available from experiment. While in crystalline compounds one usually knows the positions of all atoms involved from X-ray and/or neutron scattering - unless the crystal is disordered -, only in rare cases one knows more than the rough shape of the molecule on the surface with resolution down to about 1/2 nanometer. Even worse, in many instances the chemical identity of the atoms observed with e.g. a scanning tunneling microscope cannot be determined through this measurement. As a consequence, most of the theoretical work about structures of molecules on surfaces has been aimed at structure clarification, and to a much lesser degree at structure prediction, although the computational tools employed could equally well be used for the latter endeavor, in principle.

In this mini-review, we present an overview over the current state of the field regarding the use of theoretical methods for the clarification and prediction of structures of medium-sized and large organic molecules and biomolecules on surfaces. This is of great interest for basic research and technological applications. Indeed, such hybrid bio-inorganic molecular systems [22] might serve as ingredients of electronic or sensor devices [23], allow for design of nano-drugs [24], tissue engineering and efficient drug delivery [25], improve our understanding of catalysis [26], lead to efficient sequencing procedures [27], and constitute controllable model systems for the dynamics of complex processes such as protein folding [28]. After a quick summary of the range of experimental studies, we shortly describe some of the most popular computational methods employed. This is followed by an overview over the types of systems studied so far using computational approaches, together with an attempt to draw some preliminary conclusions based on these theoretical investigations with respect to their applicability and general understanding of molecules on surfaces.

2. Experiment

On the experimental side, (bio)molecules on surfaces have been studied for a long time. The motivation have been questions like the origin of life [29, 30], neurology [31, 32], understanding self-assembled monolayers (SAM) [25, 33-37], surface engineering [8, 38-40], nanobiotechnological applications [41-43], or catalysis [26]. Proteins and peptides have been deposited on metal surfaces of various orientations, such as the (100), (110) and (111) surfaces of Cu, Ag, Au, Ni, etc. [44-47], on inorganic (oxide) surfaces [29, 30] such as SiO_2 [44, 48], TiO_2 [49], etc., on self-assembled monolayers [50, 51], membranes [52], semiconductors such as Si [53], graphite [44] and graphene [54], just to name a few. These investigations have been performed both in vacuum and in solution [55]. Of great interest has been the formation of ordered structures at low deposition densities (less than one monolayer) ranging from dimers [56] to islands of periodic 2D-structures [57], and the generation of complete self-assembled monolayers [33]. Figure 1 shows three typical experimental results for large flexible molecules (cytochrome C [45], angiotensin II [58], crown-ether [59]) on metal surfaces obtained via scanning tunneling microscopy.

The main tools for the study of the structures the biomolecules exhibit on surfaces have been local probes such as scanning tunneling microscopy (STM) [60, 61] or atomic force microscopy (AFM) [62, 63], STM-vibrational spectroscopy [64], probes that address single molecules, in principle, but without very high spatial resolution such as IR or Raman scattering [65-67], ADXPS [68], NMR [69, 70], LEED [71], X-ray spectroscopy [60], and (small angle) X-ray or neutron scattering that in the context of biomolecules should be suitable for the identification of periodic features of the structures. These methods yield information about structural, electronic and vibrational properties of individual molecules, in principle, and possible structural arrangement patterns of small and large groups of molecules. However, the in-principle atomic resolution these methods could provide is usually only possible to achieve in special instances such as for pentacene on Cu [72]; more representative are spatial resolutions of ca. 1/2 nanometer (c.f. figure 1). Furthermore, the chemical identities of the atoms observed, or even the identities of the molecules themselves, are often not directly available through experimental measurements, although dynamic force microscopy appears to be a potential way to address this issue [73].

Typical deposition processes [74, 75] are deposition from solution [76], physical

and chemical vapor deposition [71, 77], and ion soft landing with matrix assisted laser desorption/ionization (MALDI) [5], or electrospray ionization (ESI) [6, 7, 12, 48, 78] sources. While the deposition from solution has been theoretically analyzed in depth (as we will see below), only few theoretical investigations of the processes involved in ion soft landing have been performed [7].

3. Theoretical methods

3.1 Energy landscape concepts

In the realm of theory, modeling and prediction of the structure of biomolecules has been taking place for about forty years, starting with the so-called secondary structure prediction and the investigation of the protein folding problem (for a review of the early work, c.f. [18, 19] and references therein). The theoretical study of biomolecules on surfaces is a more recent field. Nevertheless, in the last years, theoretical and computational methods have largely paralleled the experiments mentioned above (e.g. [44, 46, 47, 53]), with respect to the large variety of systems investigated, and are often guided by experimental information [79].

As is the case with all chemical systems, the (meta)stable compounds or the kinetically stable conformations of individual molecules or clusters, correspond to so-called locally ergodic regions on the energy landscape of the system [20, 21]. The energy landscape is the hypersurface of the potential energy or enthalpy for all arrangements of the N atoms belonging to the chemical system. Each arrangement corresponds to the vector $\boldsymbol{X} = (\boldsymbol{x}_1,...,\boldsymbol{x}_N)$ in $\boldsymbol{R}^{3N}$ where $\boldsymbol{x}_i$ is the position vector for the i'th atom in the usual three-dimensional space. As discussed in more detail elsewhere [21], subregions R of the state space in which the trajectory of the N-atom system $\boldsymbol{X}(t)$ resides long enough to equilibrate on a given observation times scale t_{obs}, $t_{eq}(R) << t_{obs}$, while at the same time being confined to this region for a sufficiently long time before escaping, $t_{esc}(R) >> t_{obs}$, are called locally ergodic. This implies that we can compute measurements of observables O, $\langle O \rangle_t = (1/t_{obs})\int_0^{tobs} O(\boldsymbol{X}(t))dt$, by using the standard ensemble average formula restricted to the region R, $\langle O \rangle_{ens} = \sum_{\boldsymbol{X}i \text{ in } R} O(\boldsymbol{X}_i)\exp(-E(\boldsymbol{X}_i)/k_B T)/Z(R)$, with $Z(R) = \sum_{\boldsymbol{X}i \text{ in } R} \exp(-E(\boldsymbol{X}_i)/k_B T)$. In particular, we can define a local free energy $F(R) = -k_B T \ln Z(R)$ for each such locally ergodic region (for a given time scale t_{obs}).

Clearly, if we want to predict the stable phases or conformations of a chemical system by determining its locally ergodic regions, we need to first identify such candidates and subsequently verify their stability and equilibration on the observational time scale of interest, and finally compute their local free energy. In practice, in most cases, locally ergodic regions are associated with basins of the energy landscape around local minima, and thus various global optimization techniques (for an overview see e.g [80], and references therein) are used to find candidate regions. This is followed by an analysis of the probability flows between these regions, using barrier exploration methods [81]. In many instances, the stability can be characterized by the generalized barriers surrounding the regions, which include energetic, entropic and kinetic contributions [82, 83], or by rate constants [16, 84] that can be translated into time-scale dependent free energy barriers [85]. However, this comprehensive approach already employed in the structure prediction of molecules, clusters and crystals, has not been realized for molecules on surfaces so far.

3.2 Energy functions

Several types of energy functions can be applied to the study of biomolecules on surfaces, allowing us to deal with a plethora of systems on many time scales and levels of accuracy. We may classify them based on their resolution or degree of accuracy, where the most accurate ones are typically also the computationally most expensive ones.

The most accurate energy functions are those computed via so-called ab initio methods based on quantum mechanics (QM). Among the different types of ab initio methods, density functional theory (DFT) is most commonly used (e.g. [86]). DFT energies are often combined with empirical fitted $1/r^n$ interaction terms that are used to model the van der Waals (vdW) interactions among molecules and between molecules and surfaces, which are difficult to describe with basic QM methods [87-90]. DFT approaches enable us to treat larger systems compared to classical wave-function based ab initio methods such as the Hartree-Fock method. Nevertheless, using current computational resources, QM methods can only be applied in practice to analyze one or

a few small-sized biomolecules on a reduced surface model.

For larger systems, energy functions based on classical mechanics, usually called empirical energy functions or interaction potentials, have to be applied. A plethora of such functions can be found in the literature. The empirical energy functions employed for the modeling of biomolecules on surfaces are predominantly force fields (see e.g. ref. [91, 92]), many of which have been implemented in molecular dynamics codes such as GROMACS [93, 94], AMBER [28, 95, 96], NAMD [97] or DL-POLY [98, 99]. For the atom-level interactions among molecules and between molecule and surface, Coulomb potentials [100, 101] and Lennard-Jones-type potentials - the latter for the description of van der Waals forces - are frequently employed [102, 103]. In the case of metallic surfaces, the induced polarization is incorporated via explicit image charges or polarization dipoles [104-106]. In addition, some potentials are specially designed or adapted to efficiently reproduce the peptide-peptide or peptide-metal (surface) interactions [107].

In many cases, the empirical force fields are parameterized using QM calculations. Very popular has been the recently developed GoIP force field [108] used to describe the interaction between organic molecules and gold surfaces. Here, the potentials are usually fitted to results of DFT calculations since reliable atom-level experimental (structure) data are mostly lacking. Furthermore, empirical potentials are used to describe the interactions between the molecule and the surrounding medium, if a solvent is assumed to be present [109, 110].

A possible intermediary approach is the combination of different types of energy functions, somewhat reminiscent of the QM/MM methods [111]: While the properties of the molecules can be modeled with MM-type force-fields, an embedded atom model [112, 113] can be applied for the underlying surface. For the interaction between the metallic surface and the molecules, a van der Waals and Coulomb-type potential is often assumed to work well. However, the quality of the potential is decisive for the reproducibility of known results [53, 114]. Therefore, it is necessary - due to a frequent lack of atom-level resolution experimental data - to compare the results with ab-initio calculations in order to improve the empirical potential parameters.

In order to perform simulations of very large systems, or to reduce computing time, coarse-grained energy models have also been proposed that model interactions between rigid (sub)-groups of atoms inside large molecules [62, 115] or even between whole

molecules (described e.g. as a single ellipsoid). For instance, a simple Gō-like model can be used to represent intra-molecular interactions in peptides or proteins at the amino-acid residue level, together with a specific residue-based energy function for the interactions with the surface [116].

In addition to the aforementioned physics-based energy functions, statistics-based approaches can also be applied to model hybrid biomolecule-surface systems [117]. Finally, we mention that terms in physical and statistical energy functions can be combined, as is done in ROSETTA's all-atom score function used for protein structure prediction, which has been extended to consider biomolecule-surface interactions [118].

3.3 Exploration and simulation methods

Concerning the methods used to explore the configuration space of the molecules on the surface, these fall into several general categories: molecular dynamics (e.g. [119-129]) and Monte Carlo simulations [127, 130], global optimization methods [131, 132], and steered MD (e.g. [27]) and metadynamics (e.g. [53]) for barrier investigation methods. Perusing the theoretical studies of molecules on surfaces available up to now, it quickly becomes clear that 90 % or more of the investigations employ molecular dynamics. The reason for this lies both in the intuitively high degree of "realism" of the method regarding the reproduction of the kinetic processes involved in structure formation, and in the wide availability of large and well-developed MD-codes suitable, and often specifically designed for efficiently simulating large molecules in solutions or at the interfaces between solid, liquid, and vacuum.

Since the first simulations performed by Alder and Wainwright in the 1950s [133], the method of (classical) molecular dynamics has become an appropriate instrument to investigate structures, conformations, dynamics, and thermodynamics of atomic systems. In the last decades the growth in the field of applications is coupled with the development of both computer power (hardware and architecture) and fast energy functions (see above). While the first enables the investigation of large systems (up to several millions of atoms [134]) via parallel computing, which can be performed by the linked cell algorithm or the replica data method, the latter enables the modeling of realistic systems.

The systems investigated range from inorganic materials, e.g. crystalline phases or amorphous substances, over organic (polymers) up to complex biomolecules. Using different computational set-ups (external pressures can be applied and the system temperature can be controlled by Nose-thermostats) and environments (for example surface effects / gradients / solutions) will generate appropriate ensembles. These can be applied in order to compare the results of molecular dynamics simulations to those of experiments. In order to determine structural, dynamic and thermodynamic quantities, one typically explores time averages and correlation functions, e.g. determination of the van Hove correlation gives insight into the radial distribution of atomic distances, and the calculation of the velocity autocorrelation [135] or displacement autocorrelation [136] reveals the vibrational spectrum of configurations. This is often combined with periodic local minimizations, in order to gain an overview of possible candidates for kinetically stable structures in the system; this method was first used to identify so-called inherent structures in amorphous solids and liquids.

However, due to internal high frequency vibrations, the time steps used in MD simulations are typically about 1 fs, limiting the simulation time to nano-seconds. Applications of multiple time steps [137] can elongate the simulation time by a factor five. Efforts to reach long simulation times are necessary to address the long-time processes which play a crucial role in biomolecular systems, e.g. protein folding [138]. Furthermore, in the context of many biomolecular simulations a computational set-up which should resemble an experimental one has to include lipids, water molecules or carbohydrates in addition to the biomolecule under investigation. As a consequence, the simulation of such complex systems has led to the development of sophisticated and very specialized interaction potentials, and also to the construction of a number of dedicated MD codes such as AMBER, GROMACS or NAMD.

By now, MD simulations, often incorporating new sampling methods [132, 139-141], have become an established method to reproduce and predict the structures of biomolecules under physiological conditions to elucidate their interactions and to mimic the local dynamics on a time scale of several hundreds ns. Various groups have simulated carbohydrates on metals using DFT [142], the folding of macromolecules on metal surfaces [143], or the adsorption of molecules on various surfaces [144-149].

Concerning the global exploration and optimization methods, we find that most of the other (i.e. not-MD-based) standard procedures have been employed in the context of structure prediction of (bio)molecules or clusters in vacuum [16], but much more rarely or

not at all for molecules on surfaces. Firstly, there are various types of simulated annealing, i.e. Monte Carlo or MD simulations with slowly decreasing temperature [150], multiple local optimizations such as stochastic quenches or gradient minimizations, or basin hopping [151] that is analogous to stochastic simulated annealing but with large changes in configurations followed by local minimizations [152]. Other methods suitable for biomolecules, in vacuum or on surfaces, are e.g. the recently proposed so-called threshold-minimization algorithm for the investigation of flexible molecules [153], and a variety of genetic or evolutionary algorithms [154-156], and numerous hybrid approaches combining elements of several methods.

For the exploration of the connectivity and barrier structure of the energy landscape of molecules, various search methods such as standard saddle-point search procedures [16], the threshold algorithm [157] where random walkers explore the landscape below a given set of energy lids [81, 158], and metadynamics [140] where MD or MC simulations are combined with elements of taboo-searches [159] that prevent a return to previously explored parts of the energy landscape [53], are available and have been employed to study molecular systems. Besides these global search methods, there are procedures for detailed studies of the local barrier structure, such as the nudged elastic band methods [160], the string method [161], transition path sampling [84], discrete path sampling [16], the prescribed path method [162], and the pathopt algorithm [163].

Most of the methods listed above are well known, and they have been employed for many global optimization problems and landscape explorations within and without chemistry (for more details we refer to [18, 19, 21, 80] and references cited therein). However, the limited speed of these very general exploration methods - that can be easily employed for all types of optimization problems, however, - is still a central issue. Thus, of special interest in recent years has been a new class of search methods originating from robotics. Relying on the analogy between problems in robotics and structural biology [164], methods originally developed to compute robot motions have been extended and applied to the simulation of molecular systems. In particular, computationally efficient methods have been developed for sampling and exploring the conformational space of biological macromolecules (see [165] for a survey). Algorithms have been proposed to generate conformational ensembles of flexible segments in proteins (i.e. protein loops) [166-169] and to simulate their motions [170-172].

Combined with methods in computational physics such as normal mode analysis [173], or using appropriate multi-scale molecular models [174], robot path-planning algorithms are able to compute large-amplitude conformational transitions in proteins several orders of magnitude faster than standard simulation methods such as molecular dynamics. In tandem with other optimization methods, robotics-inspired algorithms have also been proposed for the global exploration of the energy landscapes of highly flexible biomolecules [175-177].

3.4 Comparison with experiment

Most often, no precise atom-level structure information is available to easily verify the predicted structures. Therefore, one usually tries to reproduce the actual experimental measurements, at least approximately. For such a direct comparison with experiment, primarily STM images have been calculated [59], using e.g. the Tersoff-Hamann approach [178], where in many instances just the charge distribution of the molecule on the surface (or even without the surface) has been computed. Computations of STM images taking the shape of the tip of the STM-probe explicitly into account have been performed [179, 180], but can be highly problematic since the tip-shape is usually not known with sufficient accuracy.

Another quantity that is, in principle, suitable for a comparison with experiment is the vibrational density of states of molecules on surfaces. This can be calculated either directly from the Hessian for empirical potentials and from approximations to the Hessian via the frozen phonon approximation, or via the Fourier transformation of the velocity-velocity autocorrelation obtained during molecular dynamics simulations [136, 181, 182]. In particular, the combination of both methods might be suitable for studying the vibrational density of states of biomolecules on a surface, since this combination also allows the computation of decay times of the eigenmodes of the molecule on the surface, and a comparison between the modes of the molecule in vacuum or solution and the ones on the surface. Of course, the modes will contain contributions from both the molecule and the surface. But in many instances one can expect that the molecule and the surface are rather weakly coupled, and thus the individual modes will be dominated by either the molecule or the surface. Figure 2 is a snapshot taken from such a

MD-simulation, and shows the deformation of a C_{240} bucky-ball right after hitting a graphene monolayer [183].

4. Procedures and examples

Nearly all theoretical studies of structures of molecules on surfaces have taken place in close association with experiments. Important classes of molecules investigated have been e.g. amino acids, peptides, proteins, and large organic and organometallic molecules [47, 93, 143, 144]. In the appendix, we present a summary of many such examples in form of a table; we note that this list is far from exhaustive, however.

Optimal conformations of individual biomolecules and their multi-molecule patterns [46, 184] on many types of surfaces have been proposed, either taken directly from experiment, from short or moderately long MD simulations, or, more rarely, from global optimizations. Subsequent verification of their stabilities is usually restricted to a local optimization or relatively short molecular dynamics simulations starting from the proposed structure. In most instances, experimental data, within the resolution limits mentioned above, had already been available to guide the theoretical investigations by suggesting e.g. the overall shape of the molecule or the approximate arrangements of groups of molecules. The majority of these studies constitute "structure clarifications", where the goal is to show that the experimental measurements are consistent with a proposed structure and / or that the suggested structure is a likely outcome of the deposition route employed in the experiment. In fact, by now, it has become almost routine to add some kind of structure calculation to interpret experimental data; summaries of such instances going beyond those listed in the appendix can be found in a number of reviews [30, 39, 60, 130, 185-187]. In a number of studies, these kinds of calculations have been expanded to study in-depth the details of the interactions between the molecule and the surface, with the goal to explain the molecule-surface binding mechanism.

Figure 3 shows a generic flow diagram that contains all the elements of a multi-stage set-up that would allow us to derive the structures of molecules on surfaces a priori, based only on the information which type of molecules are to be adsorbed on

which type of surface. Starting from the results of global and local optimizations of the individual molecule in vacuum and/or in a solvent, in a second step we either simulate the actual deposition process on the surface, randomly deposit the molecule(s) on the surface, or use chemical intuition to select specific surface locations for the molecule. The third stage consists of the global search for optimal conformations of the molecule on the surface, and for optimal patterns of many interacting molecules on the surface. The final step is the computation of the physical properties of these optimal configurations, in order to be able to compare them with experimental data and to analyze the physical and chemical reasons underlying the existence and particular features of the observed structures. Essentially all the studies discussed in this review follow at least some partial route of this flow diagram, but only very few are comprehensive enough to merit being classed as unbiased structure predictions.

At the two extremes lie the straightforward minimization of a single proposed structure candidate and the careful long MD simulation of the deposition process (often modeled in a multistage fashion following the flow diagram outlined in figure 3). In some instances, the latter procedure is general and comprehensive enough to be roughly equivalent to a true global optimization that would have predictive power regarding the outcome of the experiment without any prior experimental measurement available. Between these two extremes, we find MD simulations of varying lengths with empirical potentials, resulting in a number of (meta)stable structure candidates that could be re-optimized on ab initio level and compared with experimental data. However, the true prediction - in contrast to the verification of experimental data - of the feasible conformations and patterns of biomolecules on metal surfaces requires exhaustive global searches on highly complex landscapes, as discussed above. In the following, we will discuss some important features of the procedures commonly used for structure clarification and prediction, and mention some exemplary studies.

4.1 Structure clarification

The starting point, either directly or indirectly, of all the structure clarification studies is some experimental information about the structure of the molecule on the surface.

Typically, this data is available with a resolution down to one, perhaps one half, nanometer, if at all. Combined with chemical information and intuition about the molecule - is the molecule rigid or flexible?, would we expect the formation of covalent bonds between (parts of) the molecule and the surface?, can the molecule form metal-organic complexes with individual metal atoms usually expected to be present on the surface?, etc. -, this is often sufficient to propose a candidate structure for the molecule or whole patterns of molecules on the surface. In this situation, the theoretical contribution commonly addresses the following issues: a) If the molecule is flexible, what is its actual shape, and does it differ from the shape in vacuum / solution / crystal? b) How is the molecule aligned / positioned / oriented with respect to the underlying surface? c) Does the surface reorganize due to the presence of the molecule? d) Is the proposed shape / positioning / arrangement pattern of the molecule(s) on the surface a stable configuration?

In the literature, we find a number of procedures, of increasing complexity, that are applied to answer these questions. Very frequently, one or a few positions and/or shapes of the molecule are constructed by hand and locally minimized with respect to energy, often on ab initio level. This is particularly common for rigid molecules where only one or two different shapes are feasible and only minor changes are expected upon placing the molecule onto the surface. An example of this very common procedure, called type A in the table in the appendix, is the formation of supramolecular islands of deprotonated 1,3,5-tris(4-ethynylphenyl) benzene (ext-TEB) on Cu(111) [184]. Here, careful DFT-minimizations of individual molecules and multi-molecule arrangements in agreement with experimental observations were used to analyze and interpret the experimental data and to understand the types of bonding among the molecules and between the molecules and the surface.

An extension of this approach, called type B, uses the minimized candidates as starting points of additional relatively short MD-simulations, usually with an empirical potential, in order to verify the stability of the proposed structure. Examples of this type are more rare. For instance, there is the study of the surface immobilized peptide cecropin P1 (cCP1) on a self-assembled monolayer [124]. Here, two very different conformations in aqueous solution were used as starting points of MD-simulations, demonstrating the different behavior and stability of the deposited or tethered,

respectively, protein, compared to its behavior in solution far away from the surface.

If the molecule exhibits a noticeable degree of flexibility, then systematically using approach A would force us to test an extremely large number of candidates. To avoid this, multistage procedures have been developed, with all the advantages and disadvantages associated with such methods: on the one hand, multistage methods constitute efficient divide-and-conquer approaches to sampling the configuration space of interest, but the price we often pay is a lack of proof that all relevant configurations have been considered. The most basic of these procedures, called type C in this review, consists of a pre-optimization of the shape of the individual molecule in vacuum or solution before placing the most promising pre-optimized polymorph of the molecule onto the surface at the location suggested by experiment, for the final optimization. An example using this type of procedure is the investigation of the role of van der Waals forces in the adsorption of PTCDA (3,4,9,10-perylene-tetracarboxylic-dianhydride) molecules on KBr [88]. Here, the authors used DFT and DFT+vdW energy minimizations to first pre-test a variety of adsorption sites, upon which careful high-accuracy minimizations followed, demonstrating the importance of the van der Waals term for the energy ranking of the binding sites and the diffusion barriers along selected displacement paths.

Closely related but more complex is procedure D, where the ab initio / high quality empirical potential (local) optimization of a random or e.g. protein-database based starting structure in vacuum (or implicit solvent) is followed by a second local optimization / equilibration in solution via a short MD simulation using an empirical potential. Finally, the solution-equilibrated molecule is placed on top of the surface, usually together with the solution molecules, for the third minimization or local equilibration step. An example of this quite frequently applied procedure is the study of the adsorption orientation and conformation of myoglobin on rutile surfaces [127]. As a first step, the rutile surface was constructed, followed by the construction of the protonated state of the protein in solution. Keeping the protein rigid, parallel tempering Monte Carlo simulations were performed to determine a starting orientation of the protein on the surface. The configuration with the lowest energy was then used as starting point of the MD-simulations of the protein on the surface, where in addition the protein was immersed in a box of water molecules. It was found that depending on the type of rutile surface (001 or 110), the heme group of the

protein was close or far away from the surface, respectively, demonstrating the importance of surface features in the adsorption of proteins.

4.2 Structure prediction

The multistage procedure type D presented in the previous subsection has already the potential to evolve into an unbiased global search procedure if the final minimization is replaced by very long MD simulation that can explore a representative sample of the full configuration space of the molecule-surface system. Of course, up to now such simulations cannot be performed on ab initio level, and thus one is often forced to add a further local optimization on ab initio level afterwards. Furthermore, one would want to also employ global optimization techniques during the first stages of the procedure, and use as many of the outcomes of the early stages as possible for starting points of the long MD simulations. Computational schedules that follow this route at least to some degree are denoted type E procedures in the table, while those that are fully global and unbiased would be a type F procedure when they are based on MD-simulations and type G if they do not involve MD simulations as part of the global optimization. Finally, there are investigations employing essentially one moderately long MD-simulation stage; these studies are summarized in the table as type H procedures.

An example of type E studies which only weakly rely on experimental structure information but are not yet fully global is the modeling of a 2D-molecular self-assembly of 1,3,5-tris(4'-*X*-*Y*-phenyl) benzene molecules (*X* = H, Br; *Y* = mono, bi) on a Si:B surface [53]. In this study, the first step consists of a DFT-investigation of the individual molecules. Next, metadynamics simulations with empirical force fields are performed for single molecules on the surface, yielding elementary building blocks. From these, one constructs by hand two-dimensional lattices (presumably roughly in agreement with experimental images), which serve as starting points of long MD-simulations. Finally, the outcomes of the simulations are compared with the experiment, where one finds that the results of the simulations are in good agreement with what chemical intuition would suggest for the experimentally observed patterns.

We note that procedures F or G (and possibly already C, D or E) usually include a global optimization of the individual molecule in vacuum or solution. Such global

structural optimizations of individual molecules without a surface have been frequently performed in the past for relatively small (< 50 atoms) isolated molecules in vacuum or in solution [152, 155, 184], where the latter one was often modeled as an effective medium [188]. For small systems like e.g. inorganic [81] or intermetallic [158] clusters in vacuum, such searches could be performed on the ab initio energy level [156], but for larger ones, empirical potentials were employed, often combined with ab initio local optimizations [189]. In contrast, no single-step global optimization studies of individual biomolecules or groups of such molecules on surfaces have been performed so far; the closest to this ideal case are studies where multi-stage procedures of types F or G have been employed. One problem is that the computational effort needed to perform a global optimization even of a small molecule on a surface on ab initio level (for both molecule and surface!) is prohibitively high. Thus the situation for molecule + surface systems is comparable to the one in crystal structure prediction twenty years ago when empirical potentials were used for the global search followed by a local ab initio minimization of all candidate structures found.

Nevertheless, there are some studies following the procedure type F, which could be considered global searches. Such a case is the exploration of the statherin-hydroxyapatite system [190]. The procedure used in this study takes its inspiration from the prediction of protein docking and folding in solvent or vacuum, i.e. it is an analogous algorithm for docking and folding of proteins on surfaces. This algorithm folds a peptide from a fully extended conformation to a solution- and an adsorbed-state structure, and repeats this procedure ca. 10^5 times, yielding an ensemble of protein + substrate structure candidates, from which the ones with the lowest energies are selected for a further analysis and comparison with experiment. One finds that large scale features of the molecule in solution and on the surface are predicted reasonably well.

We will close this section with an example of a recent global optimization following procedure G, where no molecular dynamics simulations have been involved, addressing the structure prediction of small sugar molecules on a variety of noble metal surfaces [191]. First, the most representative energy minimum conformations of a single molecule on the surface were determined using a new robotics-inspired stochastic method for global optimization [192]. An empirical potential was used at this level for the energy

computations. Figure 4 shows some low-lying minimum configurations representing major basins on the energy landscape of an individual sucrose molecule on a Cu(111) surface. Then, ab initio local optimizations of the molecule on the surface were performed to improve the quality of the identified conformations. Finally, the self-organization of groups of these molecules on the surface was analyzed using a certain variant of basin hopping simulated annealing (implemented in G42+ [193]). These global optimizations identified stable low-energy structures (such as the one shown in Figure 5) that qualitatively agree with experimentally observed STM images [194].

4.3 Structure explanation

Structure clarification to assist the experiment and structure prediction to guide future experiments are important goals in the theoretical investigation of the structure of molecules on surfaces. Going beyond the issue of correctly assigning a structure model to a given molecule on a specific surface, we would like to also explain the mechanism underlying the structure formation in the molecule+surface system. Obtaining such an explanation would allow us to understand details of the observed structures and of the experimental measurements, and thus guide us in the choice of interesting new systems to explore.

However, a major concern regarding the explanatory power of calculations clarifying and predicting such structures is the choice and quality of the energy function employed. The empirical potentials usually used for the MD-simulations are inherently limited in this respect, even though specially fitted interaction potentials can often generate structures that agree with experiment quite well. But in particular when trying to identify the dominating interactions among molecules, between molecule and solvent, and between molecule and surface, we must employ ab initio energy calculations. In this subsection, we summarize a couple of examples, where such an analysis has been performed.

A system frequently studied is benzene on metal surfaces. For instance, let us consider a density functional theory study [195] using the generalized gradient approximation (GGA) and the Perdew-Wang exchange-correlation functional to

investigate benzene on a Pt(111) surface. The authors concluded that the most favorable arrangement places the aromatic ring on bridge sites parallel to the surface. In this conformation, the molecule is slightly distorted such that it can form six C-Pt bonds with four Pt-atoms at the surface. This contrasts with the alternative, a placement on a hollow site, where six C-Pt bonds are formed but only with three Pt atoms.

Decomposing the adsorption energy into the most important contributions leads to three terms. Two of these must be paid for, i.e. they raise the overall energy of the system: the distortion energy of the molecule relative to the conformation in the gas phase, and similarly the distortion energy of the surface compared to the unperturbed surface. These terms are balanced by the interaction energy of the distorted molecule with the distorted surface, which stabilizes the adsorbed molecule on the surface. The distortion energy is considerably larger on the bridge site (1.51 eV) than on the hollow site (0.87 eV), and the distortion energy of the surface is approximately the same (0.33 eV vs. 0.35 eV). We note that this surface distortion contribution is more important for larger aromatic molecules, such as naphthalene and anthracene, than for benzene. But this energy price is worth paying: the gain in interaction energy equals 2.74 eV for the bridge site, in contrast to only 1.90 eV for the hollow site. The distortion of the molecule also affects its HOMO-LUMO gap, which changes from 5.06 eV in the gas phase to 3.94 eV on the surface, and, of course, leads to a redistribution of the electron density depleting the π and π^* molecular orbitals.

We can contrast this with another study of benzene on a Cu(110) surface [196], where the authors include explicit van der Waals interaction terms in addition to their DFT energy (GGA and PBE exchange-correlation functional). These calculations show again a planar arrangement of the benzene molecule on the metal surface for the optimal configuration. The effect of the van der Waals term barely reduces the distance between the chemisorbed benzene and the surface from 2.43 Å to 2.35 Å, indicating that the partly metallic character of the benzene-metal interaction is not strongly affected by the introduction of the van der Waals interaction. A quantitative analysis shows that the chemical interaction between molecule and surface increases. This effect is not large for benzene, but it becomes quite significant for N-substituted ring molecules such as pyridine and pyrazine. For these molecules, the molecule-surface distance also changes considerably from 2.98 Å to 2.43 Å for pyridine, and from 2.99 Å to 2.59 Å for pyrazine,

respectively. This leads to an increase of the metallic character of the molecule-surface bond and thus changes the character of the adsorption from physisorption to chemisorption in the case of pyridine and pyrazine.

A second popular system is fullerene on metal surfaces. In a study of C_{60} located on top of a surface atom for a Au(111) and Ag(100) surface [197], DFT calculations with the LDA exchange-correlation functional were performed, in order to investigate the C_{60}-surface bonding and the charge transfer involved. It is found that the charge transferred from Au(111) to C_{60} is very small, while it amounts to about 0.2 electrons for C_{60} on Ag(100). While in both cases one would speak of physisorption, it is clear that on Au(111) the bonding mechanism is a charge-neutral polarization, similar to what is usually modelled via e.g. van der Waals interactions, and the mechanism on Ag(100) is a combination of polarization and ionization.

Additional information has been gained from another DFT study [198] using GGA and a Becke-Perdew exchange-correlation functional, this time for C_{60} on a Cu(111) surface. Here, several positions of the fullerene molecule on the surface were compared, showing that several kinetically stable local minimum configurations exist. Regarding the bonding, the authors find a considerable charge transfer of about 0.5 electrons indicating that we are facing chemisorption in the case of C_{60} on copper, in contrast to the gold and silver substrates. However, we note that neither of these two investigations included energy terms that can model the charge localization or electron correlation effects with sufficient accuracy to decide which of the sites is the energetically preferred one. Furthermore, inclusion of such net-attractive terms might change the classification of the C_{60}-Au- and the C_{60}-Ag-surface interactions to chemisorption.

As a final example, we consider a study of the formation of a gadolinium - (4,1',4',1''-terphenyl-1,4''-dicarbolic acid) network on Cu(111) [199]. The ab initio DFT calculations were performed in the generalized gradient approximation for the PBE exchange-correlation functional. It was found that the presence of the surface increased the binding energy of the network by about 1.7 eV per molecule. The molecules are slightly tilted by about 5° with respect to a completely planar arrangement on the surface, and the carboxylic groups are rotated by about 45° to increase the amount of bonding between the Gd- and O-atoms. Furthermore, there is a large charge transfer between gadolinium atoms, terminal oxygen atoms and the surface. Using a Bader analysis, one

finds that there is a positive charge of about +2.1 electrons associated with the Gd-atom, one of -1.1 electrons with each of the oxygen atoms, and a net value of about +0.9 electrons transferred from the surface to the metal-organic assembly unit, respectively. Thus, the authors of the study conclude that there are strong ionic features controlling the optimal structure of the metal-organic network.

4.4 Other interesting non-standard studies

One type of theoretical structure investigations that is very common in the case of solid state chemistry is the so-called structure determination. Here, important structural information is available, such as the unit cell parameters and the unit cell content, and perhaps even a powder diffractogram of the sample, but the positions of the atoms inside the cell are unknown. Combining this experimental information and global optimizations that minimize the energy or a combination of energy and similarity between measured and computed structure factors, one can determine the unknown structure with high probability. Such a type of study is very rare in the molecule-on-surface system, but might occur more often in the future with improvements of e.g. computed and measured STM-images that could be combined to deduce the atom positions in the molecule directly with high confidence.

An example is the study of conformational changes of Cu-tetra-3,5-di-terbutyl-phenyl porphyrin molecules on Cu(211) [200]. Using the elastic scattering quantum chemistry technique for the energy calculation and the sum of the energy and the "difference between computed and experimentally observed STM-images" as the cost function of the global optimization (a so-called Pareto-optimization), the authors could show that two slightly different STM-images could be associated with an ON and OFF state of the molecule, which were characterized by the orientation of one of the legs in the porphyrin molecule.

While many of the studies mentioned in this review have modeled the adsorption process from solution, at least to some degree, only few theoretical studies of some aspects of the deposition/adsorption process during ion-soft landing have been performed [201]. These include the collision of cytochrome C with a Cu(100) substrate [143], the adsorption and self-assembly of pyrene-polyethylene on graphene [144],

deposition of tripetide on rutile surfaces [145], and the adsorption and patterning of DNA-bases on Au(111) [146, 147]. Other interesting studies deal with the dynamics of cytochrome C on Au [148], and the RNA adsorption on SAMs [149].

Finally, we present an example where steered MD is used to show that different DNA-bases can be identified by moving a DNA-fragment through a graphene nanopore, essentially associating each base-nanopore interaction with a kind of energy barrier [27]. To do so, the molecule was placed close to the nanopore and solvated in a box of water. Next, a harmonic spring force was applied to one of the ends of the molecule, forcing it to move through the nanopore. Measuring the resistance force acting on the spring during the steered MD-simulation yielded characteristic force profiles that could be associated with the passage of specific bases (A, T, C, G, and 5-methylcytosine).

5. Conclusion and outlook

While most of the work on (bio)molecules on surfaces in the literature is concerned with experimental investigations, by now also a considerable oeuvre of theoretical studies exists. We can classify these studies according to the degree to which the simulation is guided by available experimental data. The vast majority of these studies deal with the clarification and interpretation of experimental measurements, thus complementing the experimental observations.

In principle, those computational procedures that go beyond local minimizations by employing MD-simulations would be capable of actually predicting the structures of molecules on surfaces without relying on experimental data, in analogy to the by now well-established prediction of crystal structures and molecular / cluster conformations in vacuum or in a solvent. But the computational efficiency of straightforward deterministic MD-simulations is usually not sufficient for this purpose.

Thus, it is encouraging that new fast global optimization techniques and energy landscape exploration methods are being developed that are dedicated to the unbiased structure prediction of molecules on surfaces. Similarly, established stochastic optimization algorithms are being adapted to address this exciting problem. Considering the state of the field described above, it appears that we are fast approaching a break-through point in the

prediction of molecules on surfaces. Thus, we expect that within another five to ten years we will have reached the ability to predict the structures of molecules on surfaces with comparable success to that of crystal structure prediction today.

Acknowledgements:

We would like to thank S. Rauschenbach and M. Ternes for helpful comments and suggestions, and S. Rauschenbach for kindly providing figure 1.

Bibliography

[1] G. A. Ozin, A. C. Arsenault, L. Cademartiri, *Nanochemistry: A Chemical Approach to Nanomaterials*, RSC Publishing, London, **2009.**
[2] D. Philp, J. F. Stoddart, *Angew. Chem. Int. Ed*. **1996**, *35*, 1154-1196.
[3] A. Müller, H. Reuter, S. Dillinger, *Angew. Chem. Int. Ed*. **1995**, *34*, 2328-2361.
[4] J. W. Steed, D. R. Turner, K. J Wallace, *Core Concepts in Supramolecular Chemistry and Nanochemistry*, Wiley, Chichester, **2007**.
[5] H. J. Räder, A. Rouhanipour, A. M. Talarico, V. Palermo, P. Samori, K. Müllen, *Nature Mater*. **2006**, *5*, 276-280.
[6] G. E. Johnson, Q. Hu, J. Laskin, *Annu. Rev. Anal. Chem*. **2011**, *4*, 83-104.
[7] J. Cyriac, T. Pradeep, H. Kang, R. Souda, R. G. Cooks, *Chem. Rev*. **2012**, *112*, 5356-5411.
[8] J. V. Barth, *Ann. Rev. Phys. Chem*. **2007**, *58*, 375-407.
[9] D. A. King, M. G. Wells, *Surf. Sci*. **1972**, *29*, 454-482.
[10] R. C. Reuel, C. H. Bartholomew, *J. Catal*. **1984**, *85*, 63-77.
[11] H. J. Jänsch, P. Gerhard, M. Koch, *Proc. Nat. Acad. Sci*. **2004**, *101*, 13715-13719.
[12] Z. Ouyang, Z. Takats, T. A. Blake, B. Gologan, A. J. Guymon, J. M. Wiseman, J. C. Oliver, V. J. Davisson, R. G. Cooks, *Science* **2003**, *301*, 1351-1354.
[13] G. Binnig, H. Rohrer, *Surf. Sci*. **1983**, *126*, 236-244.
[14] G. Binnig, C. F. Quate, C. Gerber, *Phys. Rev. Lett*. **1986**, *56*, 930-933.
[15] D. J. Wales, H. A. Scheraga, *Science* **1999**, *285*, 1368-1372.

[16] D. J. Wales, *Energy Landscapes with Applications to Clusters, Biomolecules and Glasses*, Cambridge University Press, Cambridge, **2003**.
[17] J. C. Schön, M. Jansen, *Angew. Chem. Int. Ed*. **1996**, *35*, 1286-1304.
[18] J. C. Schön, M. Jansen, *Z. Kristallogr*. **2001**, *216*, 307-325.
[19] J. C. Schön, M. Jansen, *Z. Kristallogr*. **2001**, *216*, 361-383.
[20] J. C. Schön in *Proc. RIGI workshop 1998* (Eds.: J. Schreuer), ETH Zürich, Zürich, **1998**, pp. 75-93.
[21] J. C. Schön, M. Jansen, *Int. J. Mater. Res*. **2009**, *100*, 135-152.
[22] E. Ruiz-Hitzky, K. Ariga, Y. Lvov (Eds.), *Bio-inorganic hybrid materials*, Wiley-VCH, Weinheim, **2008**.
[23] C. Sanchez, B. Julian, P. Belleville, M. Popall, *J. Mater. Chem*. **2005**, *15*, 3559-3592.
[24] Z. Yang, S.-G. Kang, R. Zhou, *Nanoscale* **2014**, *6*, 663-677.
[25] D. Mandal, A. N. Shirazi, K. Parang, *Org. Biomol. Chem*. **2014**, *12*, 3544-3561.
[26] J. V. Perales-Rondon, A. Ferre-Vilaplana, J. M. Feliu, E. Herrero, *J. Amer. Chem. Soc*. **2014**, *136*, 13110-13113.
[27] Z. Zhang, J. Shen, H. Wang, Q. Wang, J. Zhang, L. Liang, H. Aagren, Y. Tu, *J. Phys. Chem. Lett*. **2014**, *5*, 1602-1607.
[28] H. Nguyen, J. Maier, H. Huang, V. Perrone, C. Simmerling, *J. Amer. Chem. Soc.* **2014**, *136*, 13959-13962.
[29] R. M. Hazen, D. S. Sholl, *Nature Mater*. **2003**, *2*, 367-373.
[30] J. F. Lambert, *Orig. Life Evol. Biosph.* **2008**, *38*, 211-242.
[31] P. Faller, C. Hureau, G. La Penna, *Acc. Chem. Res.* **2014**, *47*, 2252-2259.
[32] R. Vacha, S. Linse, M. Lund, *J. Amer. Chem. Soc*. 2014, *136*, 11776-11782.
[33] A. Ulman, *Chem. Rev.* **1996**, *96*, 1533-1554.
[34] F. Rosei, M. Schunack, Y. Naitoh, P. Jiang, A. Gourdon, E. Laegsgaard, I. Stensgaard, C. Joachim, F. Besenbacher, *Prog. Surf. Sci*. **2003**, *71*, 95-146.
[35] Z. Nie, E. Kumacheva, *Nature Mater*. **2008**, *7*, 277-289.
[36] A. G. Slater, P. H. Beton, N. R. Champness, *Chem. Sci*. **2011**, *2*, 1440-1448.
[37] J. P. Hill, L. K. Shrestha, S. Ishihara, Q. Ji, K. Ariga, *Molecules* **2014**, *19*, 8589-8609.
[38] J. V. Barth, G. Costantini, K. Kern, *Nature* **2005**, *437*, 671-679.
[39] J. A. A. W. Elemans, S. Lei, S. De Feyter, *Angew. Chem. Int. Ed.* **2009**, *48*, 7298-7332.

[40] L. Bartels, *Nature Chem*. **2010**, *2*, 87-95.
[41] G. M. Whitesides, *Nature Biotech*. **2003**, *21*, 1161-1165.
[42] M. Sarikaya, C. Tamerler, D. T. Schwarz, F. Baneyex, *Ann. Rev. Mater. Res.* **2004**, *34*, 373-408.
[43] A. Lakshmanan, S. Zhang, C. A. E. Hauser, *Trends Biotech*. **2011**, *30*, 155-165.
[44] S. Rauschenbach, G. Rinke, N. Malinowski, R. T. Weitz, R. Dinnebier, N. Thontasen, Z. Deng, T. Lutz, P. M. de Almeida Rollo, G. Costantini, L. Harnau, K. Kern, *Adv. Mater*. **2012**, *24*, 2761-2767.
[45] Z. Deng, N. Thontasen, N. Malinowski, G. Rinke, L. Harnau, S. Rauschenbach, K. Kern, *Nano Lett*. **2012**, *12*, 2452-2458.
[46] G. Rinke, S. Rauschenbach, S. Schrettl, T. N. Hoheisel, J. Blohm, R. Gutzler, R. Rosei, H. Frauenrath, K. Kern, *Int. J. Mass. Spectr*. **2015**, *377*, 228-234.
[47] M. Mahapatra, L. Burkholder, Y. Bai, M. Garvey, J. A. Boscoboinik, C. Hischmugl, W. T. Tysoe, *J. Phys. Chem. C* **2014**, *118*, 6856-6865.
[48] S. Rauschenbach, R. Vogelgesang, N. Malinowski, J. W. Gerlach, M. Benyoucef, G. Costantini, Z. Deng, N. Thontasen, K. Kern, *ACSNANO* **2009**, *3*, 1901-1910.
[49] A. G. Thomas, K. L. Syres, *Chem. Soc. Rev*. **2012**, *41*, 4207-4217.
[50] S. A. Miller, H. Luo, S. J. Pachuta, R. G. Cooks, *Science* **1997**, *275*, 1447-1450.
[51] J. Laskin, P. Wang, O. Hadjar, *Phys. Chem. Chem. Phys*. **2008**, *10*, 1079-1090.
[52] E. Antunes, N. G. Azoia, T. Matama, A. C. Gomes, A. Cavaco-Paulo, *Biointerfaces* **2013**, *106*, 240-247.
[53] G. Copie, Y. Makoudi, C. Krzeminski, F. Cherioux, F. Palmino, S. Lamare, B. Grandidier, F. Cleri, *J. Phys. Chem. C* **2014**, *118*, 12817-12825.
[54] Q. H. Wang, M. C. Hersam, *Nature Chem*. **2009**, *1*, 206-211.
[55] X. Zhao, F. Pan, J. R. Lu, *Prog. Nat. Sci*. **2008**, *18*, 653-660.
[56] S. M. Barlow, R. Raval, *Surf. Sci. Rep.* **2003**, *50*, 201-341.
[57] R. Gutzler, S. Stepanow, D. Grumelli, M. Lingenfelder, K. Kern, *Acc. Chem. Res*. **2015**, *48*, 2132-2139.
[58] S. Abb, L. Harnau, R. Gutzler, S. Rauschenbach, K. Kern, *Nature Comm*. **2016**, DOI: 10.1038/ncomms10335.
[59] N. Thontasen, G. Levita, N. Malinowski, Z. Deng, S. Rauschenbach, K. Kern, *J. Phys. Chem. C* **2010**, *114*, 17768-17772.

[60] F. Klappenberger, *Prog. Surf. Sci*. **2014**, *89*, 1-55.
[61] H. Kong, Q. Sun, L. Wang, Q. Tan, C. Zhang, K. Sheng, W. Xu, *ACSNANO* **2014**, *8*, 1804-1808.
[62] B. Narayanan, G. H. Gilmer, J. Tao, J. J. De Yoreo, C. V. Ciobanu, *Langmuir* **2014**, *30*, 1343-1350.
[63] P. Spijker, T. Hiasa, T. Musso, R. Nishioka, H. Onishi, A. S. Foster, *J. Phys. Chem. C* **2014**, *118*, 2058-2066.
[64] L. J. Lauhon, W. Ho, *J. Phys. Chem. A* **2000**, *104*, 2463-2467.
[65] P. Harder, M. Grunze, R. Dahint, G. M. Whitesides, P. E. Laibinis, *J. Phys. Chem. B* **1998**, *102*, 426-436.
[66] A. Campion, *Ann. Rev. Phys. Chem*. **1985**, *36*, 549-572.
[67] E. C. Le Ru, P. G. Etchegoin, *Annu. Rev. Phys. Chem*. **2012**, *63*, 65-87.
[68] J. T. Koberstein, *J. Polymer Sci*. **2004**, *42*, 2942-2956.
[69] N. F. Breen, T. Weidner, K. Li, D. G. Castner, G. P. Drobny, *J. Amer. Chem. Soc*. **2009**, *131*, 14148-14149.
[70] K. Li, P. S. Emani, J. Ash, M. Groves, G. P. Drobny, *J. Amer. Chem. Soc*. **2014**, *136*, 11402-11411.
[71] J. Ziroff, P. Gold, A. Bendounan, F. Forster, F. Reinert, *Surf. Sci*. **2009**, *603*, 2354-2358.
[72] L. Gross, F. Mohn, N. Moll, P. Lijeroth, G. Meyer, *Science* **2009**, *325*, 1110-1114.
[73] Y. Sugimoto, P. Pou, M. Abe, P. Jelinek, R. Perez, S. Morita, O. Custance, *Nature* **2007**, *446*, 64-67.
[74] S. R. Forrest, *Chem. Rev.* **1997**, *97*, 1793-1896.
[75] V. Grill, J. Shen, C. Evans, R. G. Cooks, *Rev. Sci. Instr.* **2001**, 72, 3149-3179.
[76] K. S. Mali, J. Adisoejoso, E. Ghijsens, I. De Cat, S. De Feyter, *Acc. Chem. Res*. **2012**, *45*, 1309-1320.
[77] M. C. Vasudev, H. Koerner, K. M. Singh, B. P. Partlow, D. L. Kaplan, E. Gazit, T. J. Bunning, R. R. Naik, *Biomacromol.* **2014**, *15*, 533-540.
[78] S. Rauschenbach, F. L. Stadler, E. Lunedei, N. Malinowski, S. Koltsov, G. Costantini, K. Kern, *small* **2006**, *2*, 540-547.
[79] A. D. White, G. A. Voth, *J. Chem. Theor. Comp*. **2014**, *10*, 3023-3030.
[80] A. R. Oganov (Ed.), *Modern Methods of Crystal Structure Prediction*, Wiley-VCH,

Weinheim, **2011**.
[81] S. Neelamraju, J. C. Schön, K. Doll, M. Jansen, *Phys. Chem. Chem. Phys*. **2012**, *14*, 1223-1234.
[82] J. C. Schön, M. A. C. Wevers, M. Jansen, *J. Phys.: Condens. Matt*er **2003**, *15*, 5479-5486.
[83] K. H. Hoffmann, J. C. Schön, *Found. Phys. Lett*. **2005**, *18*, 171-182.
[84] C. Dellago, P. G. Bolhuis, F. S. Csajka, D. Chandler, *J. Chem. Phys*. **1998**, *108*, 1964-1977.
[85] P. Salamon, D. J. Wales, A. Segall, Y. Lai, J. C. Schön, K. H. Hoffmann, B. Andresen, *J. Non-Eq. Thermodyn*. **2016**, *41*, 13-18.
[86] D. Lee, E. Schwegler, Y. Kanai, *J. Phys. Chem. C* **2014**, *118*, 8508-8513.
[87] M. A. Neumann, M.-A. Perrin, *J. Phys. Chem. B* **2005**, *109*, 15531-15541.
[88] O. H. Pakarinen, J. M. Mativetsky, A. Gulans, M. J. Puska, A. S. Foster, P. Grutter, *Phys. Rev. B* **2009**, *80*, 085401.
[89] J. Klimes, D. R. Bowler, A. Michaelides, *J. Phys.: Condens. Matt*er **2010**, *22*, 022201.
[90] T. Bucko, J. Hafner, S. Lebegue, J. G. Angyan, *J. Phys. Chem. A* **2010**, *114*, 11814-11824.
[91] D. Petrov, B. Zagrovic, *PLOS Comp. Biol*. **2014**, *10*, e1003638.
[92] N. Ding, X. Chen, C.-M. L. Wu, *Env. Sci. Nano* **2014**, *1*, 55-63.
[93] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theor. Comp*. **2008**, *4*, 435-447.
[94] L. B. Wright, N. A. Merrill, M. R. Knecht, T. R. Walsh, *Appl. Mater. Interfaces* **2014**, *6*, 10524-10533.
[95] D. A. Case et al., AMBER 12, **2012**.
[96] L. Xu, X. Wang, S. Shan, X. Wang, *J. Comp. Chem*. **2013**, *34*, 2524-2536.
[97] J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten, *J. Comput. Chem*. **2005**, *26*, 1781-1802.
[98] W. Smith, T. R. Forrester, *J. Mol. Graph*. **1996**, *14*, 136-141.
[99] H. Nada, *J. Phys. Chem. C* **2014**, *118*, 14335-14345.
[100] C. Sagui, T. A. Darden, *Ann. Rev. Biomol. Struct.* **1999**, *28*, 155-179.
[101] M. Dal Peraro, K. Spiegel, G. Lamoureux, M. De Vivo, W. F. DeGrado, M. L. Klein,

J. Struct. Bio. **2007**, *157*, 444-453.
[102] V. P. Raut, M. A. Agashe, S. J. Stuart, R. A. Latour, *Langmuir* **2005**, *21*, 1629-1639.
[103] B. Barz, B. Urbanc, *J. Phys. Chem. B* **2014**, *118*, 3761-3770.
[104] J. Feng, J. M. Slocik, M. Sarikaya, R. R. Naik, B. L. Farmer, H. Heinz, *small* **2012**, *8*, 1049-1059.
[105] M. Hoefling, F. Iori, S. Corni, K.-E. Gottschalk, *ChemPhysChem* **2010**, *11*, 1763-1767.
[106] Z. Xu, S.-L. Yuan, H. Yan, C.-B. Liu, *Coll. Surf. A: Physicochem. Eng. Asp*. **2011**, *380*, 135-142.
[107] D. B. Kokh, S. Corni, P. J. Winn, M. Hoefling, K. E. Gottschalk, R. C. Wade, *J. Chem. Theor. Comp.* **2010**, *6*, 1753-1768.
[108] F. Iori, R. Di Felice, E. Molinari, S. Corni, *J. Comput. Chem.* **2009**, *30*, 1465-1476.
[109] G. Nawrocki, M. Cieplak, *J. Phys. Chem. C* **2014**, *118*, 12929-12943.
[110] J. Sponer, P. Banas, P. Jurecka, M. Zgarbova, P. Kührova, M. Havrila, M. Krepl, P. Stadlbauer, M. Otyepka, *J. Phys. Chem. Lett.* **2014**, *5*, 1771-1782.
[111] H. Lin, D. G. Truhlar, *Theor. Chem. Acc*. **2007**, *117*, 185-199.
[112] M. S. Daw, M. I. Baskes, *Phys. Rev. Lett*. **1983**, *50*, 1285-1288.
[113] M. I. Baskes, *Phys. Rev. B* **1992**, *46*, 2727-2742.
[114] W. Steele, *Chem. Rev*. **1993**, *93*, 2355-2378.
[115] P. W. J. M. Frederix, G. G. Scott, Y. M. Abul-Haija, D. Kalafatovic, C. G. Pappas, N. Javid, N. T. Hunt, R. V. Ulijn, T. Tuttle, *Nature Chem*. **2015**, *7*, 30-37.
[116] S. Wei, T.A. Knotts, *Chem. Phys*. **2013**, *139*, 095102.
[117] Y. Moskovitz, S. Srebnik, *Phys. Chem. Chem. Phys*. **2014**, *16*, 11698-11707.
[118] M. S. Pacella, D. C. E. Koo, R. A. Thottungal, J. J. Gray, *Methods Enzymol*. **2013**, *532*, 343–366.
[119] J. P. Palafox-Hernandez, Z. Tang, Z. E. Hughes, Y. Li, M. T. Swihart, P. N. Prasad, T. R. Walsh, M. R. Knecht, *Chem. Mater*. **2014**, *26*, 4960-4969.
[120] G. Colombo, P. Soto, E. Gazit, *Trends Biotech*. **2007**, *25*, 211-218.
[121] M. J. Penna, M. Mijajlovic, M. J. Biggs, *J. Amer. Chem. Soc*. **2014**, *136*, 5323-5331.
[122] S. Monti, V. Carravetta, C. Battocchio, G. Iucci, G. Polzonetti, *Langmuir* **2008**, *24*, 3205-3214.
[123] J. Guo, X. Yao, L. Ning, Q. Wang, H. Liu, *RSC Adv*. **2014**, *4*, 9953-9962.

[124] Z. Wang, X. Han, N. He, Z. Chen, C. L. Brooks III, *J. Phys. Chem. B* **2014**, *118*, 12176-12185.
[125] S. S. Jang, Y. H. Jang, Y.-H. Kim, W. A. Goddard III, A. H. Flood, B. W. Laursen, H.-R. Tseng, J. F. Stoddart, J. O. Jeppesen, J. W. Choi, D. W. Steuerman, E. DeIonno, J. R. Heath, *J. Amer. Chem. Soc*. **2005**, *127*, 1563-1575.
[126] G. Forte, A. Travaglia, A. Magri, C. Satriano, D. La Mendola, *Phys. Chem. Chem. Phys.* **2014**, *16*, 1536-1544.
[127] C. Yang, C. Peng, D. Zhao, C. Liao, J. Zhou, X. Lu, *Fluid Phase Equil*. **2014**, *362*, 349-354.
[128] S. Kerisit, C. Liu, *Env. Sci. Techn*. **2014**, *48*, 3899-3907.
[129] Z. Yang, Y.-P. Zhao, *Eng. Anal. Bound. Elem*. **2007**, *31*, 402-409.
[130] C.-A. Palma, M. Cecchini, P. Samori, *Chem. Soc. Rev*. **2012**, *41*, 3713-3730.
[131] K. Makrodimitris, D. L. Masica, E. T. Kim, J. J. Gray, J. Amer. Chem. Soc. 2007, 129, 13713-13722.
[132] R. C. Bernardi, M. C. R. Melo, K. Schulten, *Biochim. Biophys. Acta* **2015**, *1850*, 872-877.
[133] B. J. Alder, T. E. Wainwright, *J. Chem. Phys*. **1957**, *27*, 1208.
[134] R. K. Kalia, A. Nakano, P. Vashishta, C. L. Roundtree, L. van Brutzel, S. Ogata, *Int. J. Fract.* **2003**, *121*, 71–79.
[135] J. M. Dickey, A. Paskin, *Phys. Rev*. **1969**, *188*, 1407.
[136] D. Beeman, R. Alben, *Adv. Phys*. **1977**, *26*, 339-361.
[137] H. Grubmüller, B. Heymann, P. Tavan, *Science* **1996**, *271*, 997-999.
[138] Y. Duan, P. Kollman, *Science* **1998**, *282*, 740-744.
[139] M. Karplus, A. McCammon, *Nature Struct. Biol*. **2002**, *9*, 646-652.
[140] A. Laio, M. Parrinello, *Proc. Natl. Acad. Sci. U.S.A*. **2002**, *99*, 12562–12566.
[141] R. A. Latour, *Coll. Surf. B: Biointerfaces* **2014**, *124*, 25-37.
[142] Y. Morikawa, H. Ishii, K. Seki, *Phys. Rev. B* **2004**, *69*, 041403(R).
[143] G. Rinke, S. Rauschenbach, L. Harnau, A. Albarghash, M. Pauly, K. Kern, *Nano Lett*. **2014**, *14*, 5609-5615.
[144] L. Xu, X. Yang, *J. Colloid Interface Sci*. **2014**, *418*, 66-73.
[145] M. Chen, C. Wu, D. Song, K. Li, *Phys. Chem. Chem. Phys*. **2010**, *12*, 406-415.
[146] A. Maleki, S. Alavi, B. Najafi, *J. Phys. Chem. C* **2011**, *115*, 22484-22494.

[147] M. Rosa, S. Corni, R. DiFelice, *J. Chem. Theor. Comp*. **2014**, *10*, 1707-1716.
[148] L. Zanetti-Polzi, I. Daidone, C. A. Bortolotti, S. Corni, *J. Amer. Chem. Soc.* **2014**, *136*, 12929-12937.
[149] J. Liu, G. Yu, J. Zhou, *Chem. Eng. Sci*. **2015**, *121*, 331-339.
[150] S. Kirkpatrick, C. D. Gelatt, Jr., M. P. Vecchi, *Science* **1983**, *220*, 671-680.
[151] D. J. Wales, J. Doye, *J. Phys. Chem. A* **1997**, *101*, 5111–5116.
[152] H. Kusumaatmaja, C. S. Whittleston, D. J. Wales, *J. Chem. Theor. Comp*. **2012**, *8*, 5159-5165.
[153] S. Neelamraju, R. L. Johnston, J. C. Schön, subm. *J. Chem. Theor. Comp*.
[154] J. H. Holland, *Adaptation in Natural and Artificial Systems*, Univ. Mich. Press, Ann Arbor, USA, **1975**.
[155] S. Neelamraju, M. T. Oakley, R. L. Johnston, *J. Chem. Phys*. **2015**, *143*, 165103.
[156] A. Supady, V. Blum, C. Baldauf, *J. Chem. Inform. Model*. **2015**, *55*, 2338-2348.
[157] J. C. Schön, *Ber. Bunsenges. Phys. Chem*. **1996**, *100*, 1388-1391.
[158] C. J. Heard, R. L. Johnston, J. C. Schön, *ChemPhysChem* **2015**, *16*, 1461-1469.
[159] F. Glover, *Interfaces* **1990**, *20*, 74-94.
[160] G. Henkelman, H. Jonsson, *J. Chem. Phys*. **1999**, *111*, 7010.
[161] W. E, W. Ren, E. Vanden-Eijnden, *J. Phys. Chem. B* **2005**, *109*, 6688-6693.
[162] D. Zagorac, J. C. Schön, M. Jansen, *J. Phys. Chem. C* **2012**, *116*, 16726-16739.
[163] C. Grebner, L. P. Pason, B. Engels, *J. Comp. Chem*. **2013**, *34*, 1810-1818.
[164] D. Parsons, J. Canny, in *Proc. Int. Conf. Intel. Sys. Mol. Biol*., ???, **1994**, pp. 322–330.
[165] I. Al-Bluwi, T. Siméon, J. Cortés, *Comp. Sci. Rev*. **2013**, *6*, 125-143.
[166] A. A. Canutescu, R. L. Dunbrack, *Protein Sci*. **2003**, *12*, 963–972.
[167] J. Cortés, T. Simeon, M. Remaud-Simeon, V. Tran, *J. Comput. Chem*. **2004**, *25*, 956-967.
[168] H. Van den Bedem, I. Lotan, J.-C. Latombe, A. M. Deacon, *Acta Crystallogr. D* **2005**, *61*, 2-13.
[169] A. Shehu, C. Clementi, L. E. Kavraki, *Proteins* **2006**, *65*, 164-179.
[170] J. Cortés, T. Simeon, V. R. de Angelo, D. Guieyesse, M. Remaud-Simeon, V. Tran, *Bioinformatics* **2005**, *21(Suppl.1)*, i116-i125.
[171] P. Yao, A. Dhanik, N. Marz, R. Propper, C. Kou, G. Öiu, H. van den Bedem, J.-C.

Latombe, I. Halperin-Landsberg, R. B. Altmann, *IEEE/ACM Trans. Comput. Biol. Bioinfo.* **2008**, *5*, 534-545.
[172] J. Barbe, J. Cortes, T. Simeon, P. Monsan, M. Remaud-Simeon, I. Andre, *Proteins: Struct. Funct. Bioinf.* **2011**, *79*, 2517-2529.
[173] S. Kirillova, J. Cortes, A. Stefaniu, T. Simeon, *Proteins* **2008**, *70*, 131-143.
[174] N. Haspel, M. Moll, M. L. Baker, W- Chiu, L. E. Kavraki, *BMC Struct. Biol*. **2010**, *10* (S1).
[175] D. Devaurs, K. Molloy, A. Shehu, T.Simeon, J. Cortes, *IEEE Trans. Nanobioscience* **2015**, *14*, 545-552.
[176] A. Roth, T. Dreyfus, C. H. Robert, F. Cazals, *J. Comput. Chem*. **2015**, DOI:10.1002/jcc.24256.
[177] L. Jaillet, F. J. Corcho, J.-J. Perez J. Cortes, *J. Comput. Chem*. **2011**, *32*, 3464-3474.
[178] J. Tersoff, D. R. Hamann, *Phys. Rev. Lett.* **1983**, *50*, 1998-2001.
[179] M. Ternes, C. Gonzales, C. P. Lutz, P. Hapala, F. J. Giessibl, P. Jelinek, A. J. Heinrich, *Phys. Rev. Lett*. **2011**, *106*, 016802.
[180] M. Ondracek, C. Gonzales, P. Jelinek, *J. Phys.: Condens. Matt*er **2012**, *24*, 084003.
[181] C. Oligschleger, *Phys. Rev. B* **1999**, *60*, 3182-3193.
[182] C. Oligschleger, J. C. Schön, *Phys. Rev. B* **1999**, *59*, 4125-4133.
[183] D. Liesegang, C. Oligschleger, *J. Mod. Phys.* **2014**, *5*, 149-156.
[184] Y.-Q. Zhang, J. Björk, P. Weber, R. Hellwig, K. Diller, A. C. Papageorgiou, S. C. Oh, S. Fischer, F. Allegretti, S. Klyatskaya, M. Ruben, J. V. Barth, F. Klappenberger, *J. Phys. Chem. C* **2015**, *119*, 9669-9679.
[185] F. S. Tautz, *Prog. Surf. Sci*. **2007**, *82*, 479-520.
[186] S. Khan, C. K. Nandi, *Nanotechnol. Rev*. **2014**, *3*, 347-359.
[187] R. A. Latour, *Biointerphases* **2008**, *3*, FC2-FC12.
[188] R. Notmann, T. R. Walsh, *Langmuir* **2009**, *25*, 1638-1644.
[189] M. Rossi, V. Blum, P. Kupser, G. von Helden, F. Bierau, K. Pagel, G. Meijer, M. Scheffler, *J. Phys. Chem. Lett*. **2010**, *1*, 3465-3470.
[190] D. L. Masica, J. J. Gray, *Biophys. J.* **2009**, *96*, 3082-3091.
[191] J. Cortes, J. C. Schön, B. Andriyevskyy, unpubl.

[192] J. Cortes, J. C. Schön, M. Vaisset, unpubl.
[193] J. C. Schön, Proc. Appl. Ceram. 2015, 9, 157-168.
[194] S. Abb, L. Harnau, J. C. Schön, J. Cortes, S. Rauschenbach, K. Kern, *Poster at Eur. Conf. Surf. Sci. 31* (Barcelona, Spain), **2015**.
[195] C. Morin, D. Simon, P. Sautet, *J. Phys. Chem. B* **2004**, *108*, 12084-12091.
[196] N. Atodiresei, V. Caciuc, P. Lazic, S. Blügel, *Phys. Rev. Lett*. **2009**, *102*, 136809.
[197] X. Lu, M. Grobis, K. H. Khoo, S. G. Louie, M. F. Crommie, *Phys. Rev. B* **2004**, *70*, 115418.
[198] J. A. Larsson, S. D. Elliott, J. C. Greer, J. Repp, G. Meyer, R. Allenspach, *Phys. Rev. B* **2008**, *77*, 115434.
[199] J. I. Urgel, B. Cirera, Y. Wang, W. Auwärter, R. Otero, J. M. Gallego, M. Alcami, S. Klyatskaya, M. Ruben, F. Martin, R. Miranda, D. Ecija, J. V. Barth, *small* **2015**, DOI: 10.1002/smll.201502761.
[200] F. Moresco, G. Meyer, K.-H. Rieder, H. Tang, A, Gourdon, C. Joachim, *Phys. Rev. Lett*. **2001**, *86*, 672-675.
[201] L. Konermann, R. G. McAllister, H. Metwally, *J. Phys. Chem. B* **2014**, *118*, 12025-12033.
[202] L. Delle Site, K. Kremer, *Int. J. Quant. Chem*. **2005**, *101*, 733-739.
[203] C. Poojari, A. Kukol, B. Strodel, *Biochim. Biophys. Acta* **2013**, *1828*, 327-339.
[204] T. Zheng, C. Wu, M. Chen, *Surf. Sci*. **2013**, *616*, 51-59.
[205] D. Tunega, M. H. Gerzabek, G. Haberhauer, H. Lischka, *Eur. J. Soil Sci*. **2007**, *58*, 680-691.
[206] K. Sheng, Q. Sun, C. Zhang, Q. Tan, *J. Phys. Chem. C* **2014**, *118*, 3088-3092.
[207] C. Liao, Y. Xie, J. Zhou, *RSC Adv*. **2014**, *4*, 15759-15769.
[208] C. L. Freeman, J. H. Harding, *J. Phys. Chem. C* **2014**, *118*, 1506-1514.
[209] R. Mazzarello, A. Cossaro, A. Verdini, R. Rousseau, L.Casalis, M. F. Danisman, L. Floreano, S. Scandolo, A. Morgante, G. Scoles, *Phys, Rev. Lett*. **2007**, *98*, 016102.
[210] M. Böhringer, K. Morgenstern, W.-D. Schneider, R. Berndt, F. Mauri, A. De Vita, R. Car, *Phys. Rev. Lett*. **1999**, *83*, 324-327.
[211] F. S. Emami, V. Puddu, R. J. Berry, V. Varshney, S. V. Patwardhan, C. C. Perry, H. Heinz, *Chem. Mater*. **2014**, *26*, 5725-5734.
[212] J. Yu, M. L. Becker, G. A. Carri, *small* **2010**, *20*, 2242-2245.

[213] I. Brovchenko, G. Singh, R. Winter, *Langmuir* **2009**, *25*, 8111-8116.
[214] T. Onodera, K. Kawasaki, T. Nakakawaji, Y. Higuchi, N. Ozawa, K. Kurihara, M. Kubo, *J. Phys. Chem. C* **2014**, *118*, 5390-5396.
[215] L. Duchesne, G. Wells, D. G. Fernig, S. A. Harris, R. Levy, *ChemBioChem* **2008**, *9*, 2127-2134.
[216] L. Romaner, G. Heimel, J.-L. Bredas, A. Gerlach, F. Schreiber, R. L. Johnson, J. Zegenhagen, S. Duhm, N. Koch, E. Zojer, *Phys. Rev. Lett*. **2007**, *99*, 256801.
[217] A. Michaelides, V. A. Ranea, P. L. de Andres, D. A. King, *Phys. Rev. Lett*. **2003**, *90*, 216102.
[218] Z. Zhao, V. H. Crespi, J. D. Kubicki, D. J. Cosgrove, L. Zhong, *Cellulose* **2014**, *21*, 1025-1039.

Appendix:

Table 1: List of structure simulation studies of (bio)molecules on surfaces. "Molecule", "Solvent" and "Surface" characterize the system that is investigated in paper no. "Ref.". In a number of studies, the effects of different solvents and surfaces are compared, thus several solvents and surfaces are listed. "S/P" refer to whether the structure of single individual molecules (S) or whole structure patterns of many molecules (P), respectively, are investigated in the paper (for example, Y/N is short for Yes/No, with the first entry referring to single molecules and the second to whole patterns). "Type" denotes the simulation strategy according to the classification described in section four in the main text, and "Energy" indicates the kind of energy function(s) employed in the study; here, "emp." refers to any empirical potential. Finally, "Focus" informs about the main focus of the study, as far as the theoretical aspects are concerned. This list aims to give an overview over many typical studies in the field ordered by the name of the molecule involved, but it is far from exhaustive, of course.

Molecule	Surface	Solv.	S/P	Ener.	Type	Focus	Ref.
Adenine	Au(111)	No	Y/Y	DFT, emp.	E	Conformations and aggregation patterns of the DNA-base as function of surface coverage	146
Adenine	Au(111)	No, H_2O	Y/Y	DFT, emp.	E	Energetic and entropic aspects of conformations and monolayer structures of the DNA-base on Au	147
Alanine	Ni(111)	No	Y/Y	DFT, emp.	D	Structure of alanine adsorbed on Ni surface	202
L-Alanine-L-glutamic acid dipeptide (ALA-GLU)	TiO_2 rutile	H_2O	Y/N	emp.	C	Identification of contact points of adsorbed dipeptide on rutile for different conformations	122
L-Alanine-L-lysine dipeptide (ALA-LYS)	TiO_2 rutile	H_2O	Y/N	emp.	C	Identification of contact points of adsorbed dipeptide on rutile for different conformations	122
4-Alkoxybenzoic acids (4-ABA)	HOPG (graphite)	No, 1-octanoic acid	Y/Y	by hand	A	Self-assembly at the liquid-solid interface dependent on substrate / solvent	76
Amino acids (Asp, Glu, Arg, Lys, His, Phe, Trp, Asn, CysH, Gln, Ser, Thr, Ala, Gly, Ile, Leu, Met, Pro, Val)	Au(111)	No	Y/N	emp.	D	Testing force field design by computing binding energies and optimal structures of capped amino acids on Au	107

Amino acids (Asp, Glu, Arg, Lys, His, Phe, Trp, Asn, Cys, Gln, Ser, Thr, Ala, Gly, Ile, Leu, Met, Pro, Val,Tyr)	Au(111)	No, H_2O	Y/N	DFT, emp.	E	Comparison of amino acid conformations and orientations on Au(111) surface	105
Amyloid-beta peptide $A\beta_{42}$	Lipid membrane	H_2O	Y/N	emp.	D	Interaction between peptide and membrane, with surface charge and lipid tail type being crucial for transmembrane stability	203
Models of amyloidogenic peptides	implicit surface	implicit	N/Y	emp.	H	Dependence of growth features of the oligomers on the systems surface/bulk ratio	32
Angiotensin I (DRVYIHPFHL), Angiotensin II (DRVYIHPF)	Au(111)	No	Y/Y	emp.	D	Determination of conformations of individual molecules via simulations of soft-landing + relaxation of single molecules, and stability analysis of multi-molecule patterns	58
Anthracene	Pt(111), Rh(111)	No	Y/N	DFT	A	Comparison of the effect of different locations of the aromatic rings in flat adsorption conformations of anthracene on metal surfaces	195

Arginine-glycine-aspartic acid tripeptide (RGD)	TiO_2 rutile	H_2O	Y/N	emp.	D	Dynamics and adsorption of RGD with electrostatic interactions between charged peptide regions and substrate	145
Aspartic acid	Calcite crystal with surface kinks	H_2O	Y/N	emp.	D	Dependence of the amino acid's conformations and dynamics on the surface features	99
BBA protein: mixed α/β; PDB-ID: 1FME	Graphene	H_2O	Y/N	emp.	D	Effect of graphene surface on structure of adsorbed protein	123
Benzene	Ni(111)	No	Y/Y	DFT, emp.	D	Structure of benzene adsorbed on Ni surface	202
Benzene	Cu(110)	No	Y/N	DFT	A	Role of van der Waals interactions for a flat adsorption geometry of benzene	196
Benzene	Pt(111), Rh(111)	No	Y/N	DFT	A	Comparison of the effect of different locations of the aromatic rings in flat adsorption conformations of benzene on metal surfaces	195
Bisphenol-A-polycarbonate (BPA-PC)	Ni(111)	No	Y/Y	DFT, emp.	D	Structure of adsorbed BPA-PC and interplay of adsorption energies and conformational entropy	202
n-Butane	Mg,Al,Cu, Ru,Ag,Pt, Au	No	Y/N	DFT	A	Dependence of conformations on type of metal surface	142

Cecropin P1 (cCP1) peptide: SWLSKTAKK LENSAKKRIS EGIAIAIQGGP R	Silane-type SAM	H_2O, TFE+ H_2O	Y/N	emp.	B	Stability analysis for two different conformations of cCP1 on the self-assembled monolayer surface in H_2O vs. in TFE/ H_2O-mixture	124
Collagen segment (2KLW)	TiO_2 rutile (110) with step defects	H_2O	Y/N	emp.	D	Adsorption mechanisms for collagen on a hydrated defective rutile surface	204
Collagen molecules (simplified model)	implicit surface	implicit	Y/Y	emp.	H	Elucidation of the interactions that govern the self-assembly of collagen type molecules on a substrate (modeled as implicit surface)	62
Cytochrome C (Cyt C)	Cu(100)	No	Y/N	emp.	D	Simulation of the landing process and relaxation of the unfolded protein Cyt C	143
Cytochrome C (Cyt C)	Cu(001), Au(111), BN on Rh(111)	No	Y/N	emp.	H	Simulation of folding process conformations on surface in vacuum	45
Cytochrome C (Cyt C)	Au(111)	H_2O	Y/N	DFT, emp.	E	Influence of structure and dynamics of Cyt C on the protein's redox-potential	148
Cytosine	Au(111)	No	Y/Y	DFT, emp.	E	Conformations and aggregation patterns of the DNA-base as function of surface coverage	146

Cytosine	Au(111)	No, H_2O	Y/Y	DFT, emp.	E	Energetic and entropic aspects of conformations and monolayer structures of the DNA-base on Au	147
Dibenzo-24-crown-8-alkali complexes (Cs^+, Na^+, H^+)	Cu(100)	No	Y/N	DFT	A	Adsorption geometry as function of alkali cation	59
Dichlorophenoxyacetic acid - Ca^{2+} complex	Montmo-rillonite	H_2O, No	Y/N	DFT	D	Comparison between sorption on dry and hydrated clay surface and intercalation between layers	205
Diethynyl-anthracene (DEAN)	Cu(111)	No	Y/N	DFT	A	Self-assembly based on weak directional inter-molecular interactions - ionic hydrogen bonding - and strong surface-anchoring	184
Dinapthalenyl-benzene (DNYB)	Cu(110), Au(111)	No	Y/Y	DFT	A	Influence of the metal surface on the on-surface self-assembly of DNYB	206
Di(tert-butyl)terphenyl (DTBT)	Cu(110), Au(111)	No	Y/Y	DFT	A	Influence of the tert-butyl group and type of metal surface on the on-surface self-assembly of DTBT	206
Fibronectin (FN-III10)	Hydroxy-apatite (001) with calcium vacancy defects (HAP)	H_2O	Y/N	emp.	D	Mechanism of conformational and orientational changes during the adsorption of fibronectin molecules on HAP	207

Fibronectin (FN-III7-10)	Hydroxy-apatite (001) with calcium vacancy defects (HAP)	H_2O	Y/N	emp.	D	Mechanism of conformational and orientational changes during the adsorption of fibronectin molecules on HAP	207
Fullerene C_{60}	Au(111), Ag(100)	No	Y/N	DFT	A	Comparison of C_{60} substrate bonding and charge transfer on noble metal surfaces	197
Fullerene C_{60}	Cu(111)	No	Y/N	DFT	A	Adsorption geometries of C_{60} molecules on Cu(111) including their orientation	198
Glycyl-glycyl-histidine (Gly-Gly-His) peptide	Au(111)	H_2O	Y/N	emp.	D	Mechanism of adsorption of peptide on Au surface	106
Glycyl-histidine (Gly-His) peptide	Au(111)	H_2O	Y/N	emp.	D	Mechanism of adsorption of peptide on Au surface	106
Glycyl-histidine-glycine (Gly-His-Gly) peptide	Au(111)	H_2O	Y/N	emp.	D	Mechanism of adsorption of peptide on Au surface	106
$Glycine_4$-Glycine-$Glycine_4$ peptide	Alkanethiol SAM on gold, oligoethylene oxide SAM	H_2O with Na/Cl ions	Y/N	emp.	D	Testing force fields for description of peptide adsorption	102
$Glycine_4$-Lysine-$Glycine_4$ peptide	Alkanethiol SAM on gold, oligoethylene oxide SAM	H_2O with Na/Cl ions	Y/N	emp.	D	Testing force fields for description of peptide adsorption	102
Guanine	Au(111)	No	Y/Y	DFT, emp.	E	Conformations and aggregation patterns of the DNA-base as function of surface coverage	146

Guanine	Au(111)	No, H_2O	Y/Y	DFT, emp.	E	Energetic and entropic aspects of conformations and monolayer structures of the DNA-base on Au	147
Histidine (Hys)	Au(111)	H_2O	Y/N	emp.	D	Mechanism of adsorption of peptide on Au surface	106
λ-repressor: α-helices; PDB-ID: 1LMB	Graphene	H_2O	Y/N	emp.	D	Effect of graphene surface on structure of adsorbed protein	123
Mannose	Calcite (10.4) surface	H_2O, No	Y/N	emp.	E	Computation of the entropy of adsorption via thermodynamic integration from solvent-free system to system with solvent	208
Methanoic acid	Calcite (10.4) surface	H_2O, No	Y/N	emp.	E	Computation of the entropy of adsorption via thermodynamic integration from solvent-free system to system with solvent	208
Methylthiolate CH_3S	Au(111)	No	Y/N	DFT	B	Formation of superstructures in self-assembled monolayers of CH_3S on Au	209
Myoglobin	TiO_2 rutile (110), (001)	H_2O	Y/N	emp.	D	Dependence of orientation and conformation of adsorbed myoglobin on type of rutile surface	127

Napthalene	Pt(111), Rh(111)	No	Y/N	DFT	A	Comparison of the effect of different locations of the aromatic rings in flat adsorption conformations of naphthalene on metal surfaces	195
p-Nitro-analine	p-Nitro-analine	Octanol, H_2O	Y/Y	DFT, emp.	D	Ordering phenomena at the solid-liquid interface (e.g. hydration structures) in different solvents	63
Nitronaphthalene (NN)	Au(111)	NO	Y/Y	DFT, emp.	B	Influence of hydrogen bonds and electrostatic (repulsive) forces on self-assembly	210
Oligoyne derivatives with six acetylene units	Au(111)	No	Y/N	DFT	A	Structure and electron density map of molecule in gas phase (representing soft-landed molecule)	46
Peptide AFILPTG (overall neutral)	Na-silica	H_2O + ions	Y/N	emp.	H	Prediction of adsorption of peptides on various silica surfaces as function of pH and particle size	211
Peptide A3 (AYSSGAPPMPPF)	Pt(111)	H_2O	Y/N	emp.	D	Identification of a molecular level mechanism for peptide adsorption for uncharged surfaces	121
Peptide A3 (AYSSGAPPMPPF)	Au-nano-particles: Au(111) surfaces	H_2O	Y/N	emp.	H	Influence of the shape of the nano-particle surface on the adsorption	104

Peptide AYSSGAPPMPPF (A3)	Au nano-particle	H_2O	Y/N	emp.	D	Influence of the peptide's structure on hydrated Au nano-particles of different sizes	212
Peptide FLG-Na_3 (DYKDDDDK - 3 Na^+)	Au-nano-particles: Au(111) surfaces	H_2O	Y/N	emp.	H	Influence of the shape of the nano-particle surface on the adsorption	104
Peptide EQLGVRKELRGV (AgBP2)	Ag(111), Au(111)	H_2O	Y/N	emp.	H	Comparison of peptide adsorption on Ag and Au surfaces including solvent and surface contributions to free energies and binding energies	119
Peptide Gly_2-Lys-Gly_2-Lys-Gly_2-His_6	Ni,Cu,Au (100)	H_2O	Y/N	emp.	C	Strength of the interactions between peptide and metal surface, and their effect on adsorption geometries	129
Peptide GNNQQNY (part of yeast prion Sup35)	Paraffin-like(hydrophobic), silica-like (hydrophilic) model surfaces	H_2O	N/Y	emp.	D	Influence of water surface interaction on the structure of the peptides and their degree of self-aggregation	213
Peptide KLPGWSG (positively charged)	Na-silica	H_2O + ions	Y/N	emp.	H	Prediction of adsorption of peptides on various silica surfaces as function of pH and particle size	211
Peptide LDHSLHS (negatively charged)	Na-silica	H_2O + ions	Y/N	emp.	H	Prediction of adsorption of peptides on various silica surfaces as function of pH and particle size	211

Peptide LE10	Lipidic bilayers	H_2O	N/Y	emp.	D	Interaction of the peptide with the membrane, in particular via electrostatic interactions	52
Peptide NFGAIL (part of islet amyloid peptide)	Paraffin-like(hydrophobic), silica-like (hydrophilic) model surfaces	H_2O	N/Y	emp.	D	Influence of water surface interaction on the structure of the peptides and their degree of self-aggregation	213
Peptide SD152	Pt(111)	H_2O	Y/N	emp.	D	Identification of a molecular level mechanism for peptide adsorption for uncharged surfaces	121
Peptide Ser_{12} (SSSSSSSSSSSS)	Au-nano-particles: Au(111) surfaces	H_2O	Y/N	emp.	H	Influence of the shape of the nano-particle surface on the adsorption	104
Peptide TGIFKSARAMRN (AgBP1)	Ag(111), Au(111)	H_2O	Y/N	emp.	H	Comparison of peptide adsorption on Ag and Au surfaces including solvent and surface contributions to free energies and binding energies	119
Peptide WAGAKRLVLRRE (AuBP1)	Ag(111), Au(111)	H_2O	Y/N	emp.	H	Comparison of peptide adsorption on Ag and Au surfaces including solvent and surface contributions to free energies and binding energies	119

Peptide WALRRSIRR QSY (AuBP2)	Ag(111), Au(111)	H_2O	Y/N	emp.	H	Comparison of peptide adsorption on Ag and Au surfaces including solvent and surface contributions to free energies and binding energies	119
Peptide Tyr_{12} (YYYYYYYYY YYY)	Au-nano-particles: Au(111) surfaces	H_2O	Y/N	emp.	H	Influence of the shape of the nano-particle surface on the adsorption	104
Peptide fragments of the nerve growth factor, NGF(1-14), and of the brain derived neurotrophic factor, BDNF(1-12)	Au_2O_3	H_2O	Y/Y	emp.	D	Conformations, orientations, and aggregates of individual peptide fragments and of the homo- and hetero-dimers of the peptide fragments, and mechanisms of adsorption of the polypeptides	126
Perylene-tetra carboxylic-dia nhydride (PTCDA)	KBr(001)	No	Y/N	DFT	C	Importance of van der Waals forces in controlling diffusion and adsorption	88
Phenanthrene	Pt(111), Rh(111)	No	Y/N	DFT	A	Comparison of the effect of different locations of the aromatic rings in flat adsorption conformations of phenanthrene on metal surfaces	195
Phenylalanine	Ni(111)	No	Y/Y	DFT, emp.	D	Structure of phenylalanine adsorbed on Ni surface	202

Polybrominated diphenyl ether (PBDE) congeners: BDE 15, BDE 28, BDE 47, BDE 99, BDE 153, BDE 154, BDE 183, BDE 194, BDE 209	Graphene	H_2O	Y/N	DFT, emp.	B	Interaction between PBDE and graphene during adsorption, and the effect of bromination on the adsorption process	92
Polyethylene	Mg,Al,Cu, Ru,Ag,Pt, Au	No	Y/N	DFT	A	Dependence of conformations on type of metal surface	142
Polytetrafluoroethylene (PTFE)	Al_2O_3 for different surface atoms: F, OH, Al	No	Y/Y	emp.	H	Structure and formation of a PTFE transfer film between an Al_2O_3-surface and a PTFE crystal	214
Protein CALNNGKGALVPRGSKGTAK, with/without Bis(sulfosuccinimidyl)suberate (BS3) linker	Peptide SAM nano-particle	No	Y/N	emp.	B	Influence of the linker molecule on the size of the space of molecule conformations accessible to the peptide on the surface	215
Pyrazine $C_4N_2H_4$	Cu(110)	No	Y/N	DFT	A	Role of nitrogen heteroatom and van der Waals interactions for a flat adsorption geometry of pyrazine	196
Pyrene-polyethylene glycol (Py-PEG_9, Py-PEG_{45}, Py-PEG_{113})	Graphene	H_2O / No	Y/N	emp.	D	Dynamics of adsorption on graphene, and analysis of the influence of the solvent on the process	144

Pyridine C_5NH_5	Cu(110)	No	Y/N	DFT	A	Role of nitrogen heteroatom and van der Waals interactions for a flat adsorption geometry of pyridine	196
Ribonuclease A (RNase A)	SAMs: $HS(CH_2)_{10}COOH$, $HS(CH_2)_{10}NH_2$	H_2O	Y/N	emp.	E	Dependence of orientation and conformation of adsorbed RNase A on the charge of the surface	149
[2]Rotaxane	Au(111)	No	Y/Y	DFT, emp.	D	Structure and properties of disulfide tethered rotaxane self-assembled monolayers on Au surfaces	125
Statherin	Hydroxy-apatite crystal	implicit	Y/N	emp.	F	Approach to structure prediction of foldable proteins on surfaces	190, 131
Gd:Terphenyl-dicarboxylic acids (Gd-TDA)	Cu(111)	No	Y/Y	DFT	A	Rationalization of supramolecular networks	199
Cu-Tetra-3,5 di-ter-butyl-phenyl porphyrin (Cu-TBPP)	Cu(211)	No	Y/N	DFT + experimental data	A	Orientation and intramolecular conformation of Cu-TBPP on Cu-surface via combining DFT + experimental data	200

Tetrafluoro-tetracyanoquinodimethane (F4-TCNQ)	Cu(111)	No	Y/N	DFT	A	Explanation of the interfacial electronic structure by a combination of bidirectional charge transfer between molecule and metal and geometric distortions of the molecule	216
Thymine	Au(111)	No	Y/Y	DFT, emp.	E	Conformations and aggregation patterns of the DNA-base as function of surface coverage	146
Thymine	Au(111)	No, H_2O	Y/Y	DFT, emp.	E	Energetic and entropic aspects of conformations and monolayer structures of the DNA-base on Au	147
Tri-peptides (8000 different combinations of 20 amino acids)	No	H_2O	N/Y	emp.	F	Development of a set of design rules for self-assembling sequences via joint optimization of aggregation propensity and hydrophilicity	115
Tris(biphenyl) benzene (THBB)	Si(111):B	No	Y/Y	DFT, emp.	E	Energetics of different patterns of molecular lattices on passivated semiconductors	53
Tris(bromobiphenyl) benzene (TBBB)	Si(111):B	No	Y/Y	DFT, emp.	E	Energetics of different patterns of molecular lattices on passivated semiconductors	53

Tris(bromophenyl) benzene (TBB)	Si(111):B	No	Y/Y	DFT, emp.	E	Energetics of different patterns of molecular lattices on passivated semiconductors	53
Tris(ethynylphenyl) benzene (Ext-TEB)	Cu(111)	No	Y/N	DFT	A	Self-assembly based on weak directional inter-molecular interactions - ionic hydrogen bonding - and strong surface-anchoring	184
Tris(phenyl) benzene (THB)	Si(111):B	No	Y/Y	DFT, emp.	E	Energetics of different patterns of molecular lattices on passivated semiconductors	53
Uranyl	Kaolinite, orthoclase	H_2O	Y/N	emp.	H	Adsorption of uranyl on mineral surfaces	128
Uranyl carbonate	Kaolinite, orthoclase	H_2O	Y/N	emp.	H	Adsorption of uranyl carbonate on mineral surfaces	128
Water (H_2O)	Ru(0001) (111) of Rh,Pd,Pt Cu,Ag,Au	No	Y/N	DFT	A	Exploration of the binding mode of H_2O during adsorption on a variety of metal surfaces	217
Water (H_2O)	Functionalized Si	H_2O	N/Y	DFT	H	Influence of surface adsorbates on interfacial H_2O on several nonpolar hydrophobic surfaces	86
WW domain: β-strands; PDB-ID: 1E0L	Graphene	H_2O	Y/N	emp.	D	Effect of graphene surface on structure of adsorbed protein	123

Xyloglucan (several oligosaccharides consisting of three repeating units, with each unit consisting of four glucosyl type residues)	Cellulose microfibrils	H_2O	Y/N	emp.	D	Comparison of adsorption of various xyloglucan molecules on cellulose studying the effects of surface hydrophobicity and side-chain variation	218

Figure captions:

Figure 1: Large molecules with conformational freedom on surfaces observed by scanning tunneling microscopy. (a) Unfolded proteins (cytochrome c) on a Cu(100) surface. Figure based on ref. [45]. (b) Self-assembled network of eight-aminoacid peptide angiotensin II molecules on a Au(111) surface. Figure based on ref. [58]. (c) Host-guest interaction of crown ether di-benzo-24-crown-8 with different central ions Cs^+ and Na^+ on a Cu(100) surface. Figure based on ref. [59].

Figure 2: C_{240}-buckyball hitting a graphene surface, Note the deformation of the C_{240} molecule and the dip on the graphene surface. Figure based on ref. [183].

Figure 3: Generic flow diagram of the computational elements of a multi-stage procedure to predict the structures of molecules on surfaces, i.e. conformations of individual molecules and arrangement patterns of assemblies of molecules. The arrows indicate that the calculations in the next stage are assumed to be able to start from configurations derived in the previous stage. The theoretical studies discussed in this review usually include only some stages of the flow diagram, e.g. the investigation might start with a global optimization of the molecule in the solvent, followed by a deposition by hand on the surface, and finishes with a local optimization of the molecule's conformation on the surface. Alternatively, one might perform a local optimization of the single molecule in vacuum, which is then placed at random on the surface in vacuum, and globally optimized on the surface, followed by a local optimization of a chemically intuitive pattern of many molecules created by hand. Of course, the original starting point of all studies (not shown explicitly in the diagram) is the choice of the organic molecule or biomolecule that can usually be assumed to preserve its topology (the pattern of bonds) throughout the process, where its structure is taken from some database. Note that while most empirical potentials force the molecule's topology to remain unchanged during the simulations, global optimizations on ab initio level can lead to the break-up (or even merger) of molecules, in principle, although such a massive reconstruction of the molecule will only rarely occur.

Figure 4: Four low-energy minima conformations belonging to different major basins on the energy landscape of an individual sucrose molecule on a Cu(111) surface. Colors: White = H-atom, red = O-atom, blue = C-atom, golden = Cu-atom. Starting point of the search was a sucrose molecule that had been optimized in vacuum (following the procedure described in ref. [153]) and then placed on the copper surface. The global optimization of the molecule on the surface allowed the variation of all dihedral angles and of the position and orientation of the molecule with respect to the Cu-surface. Figure based on ref. [191].

Figure 5: Low-energy minimum structure pattern of many sucrose molecules on a Cu(111) surface, observed during global optimization using a basin-hopping simulated annealing type algorithm. Starting point of the global search were several molecules with the conformation shown in Figure 4a), which had been placed at random on the substrate inside a variable simulation box with periodic boundary conditions in the two dimensions parallel to the substrate surface. During the search, simulation cell parameters, and position and orientation of the sucrose molecules were allowed to vary. Note that the aim of this figure is to show the overall pattern observed as a result of the global optimization, not the atom-level structural details of how the sucrose molecules individually are positioned on the surface. Figure based on ref. [191].

Figure 1:

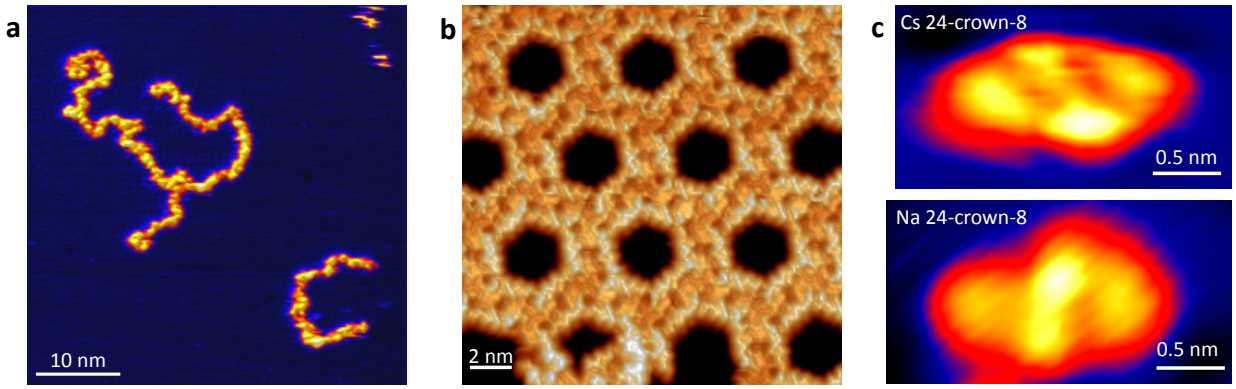


Figure 2:

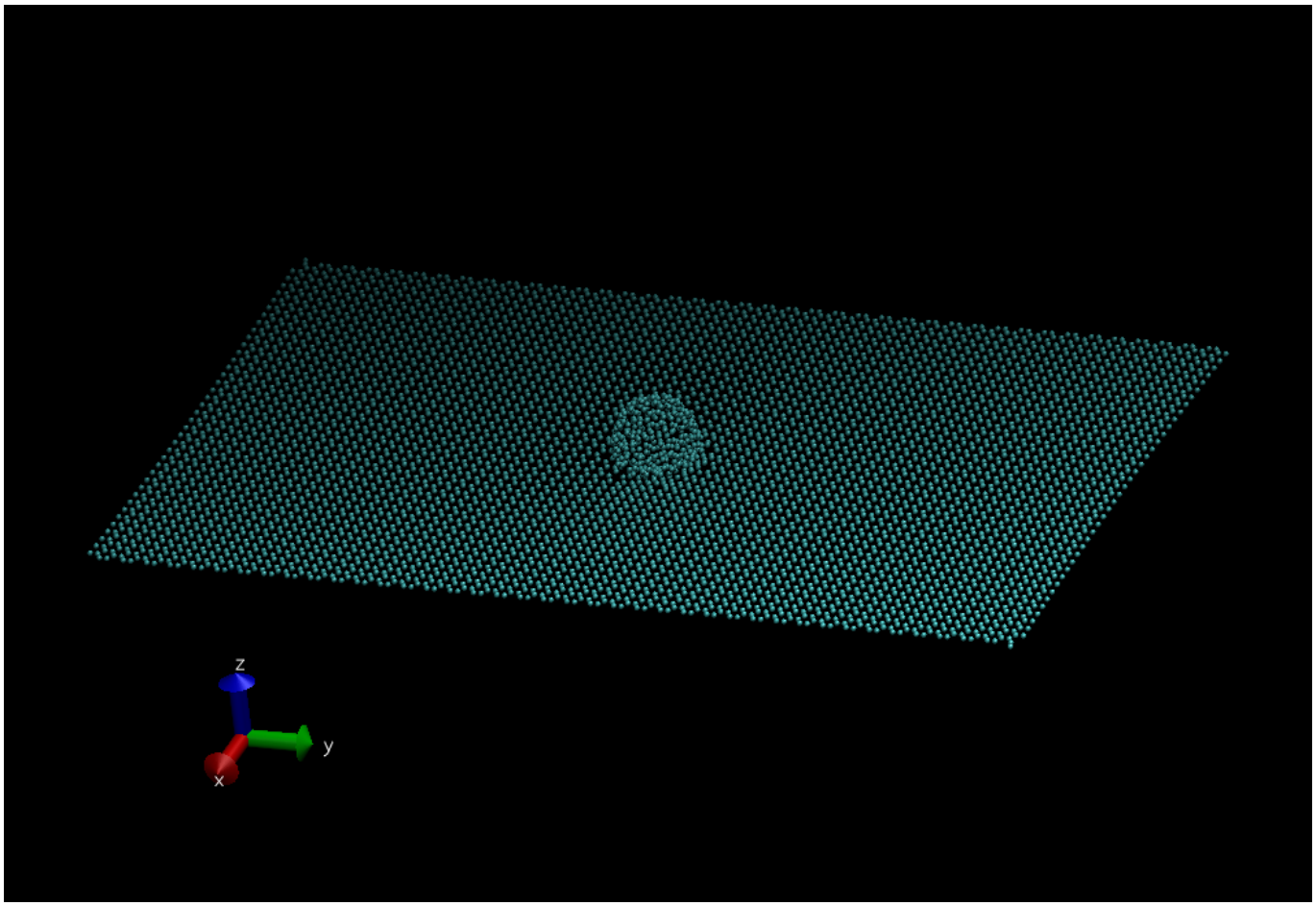


Figure 3:

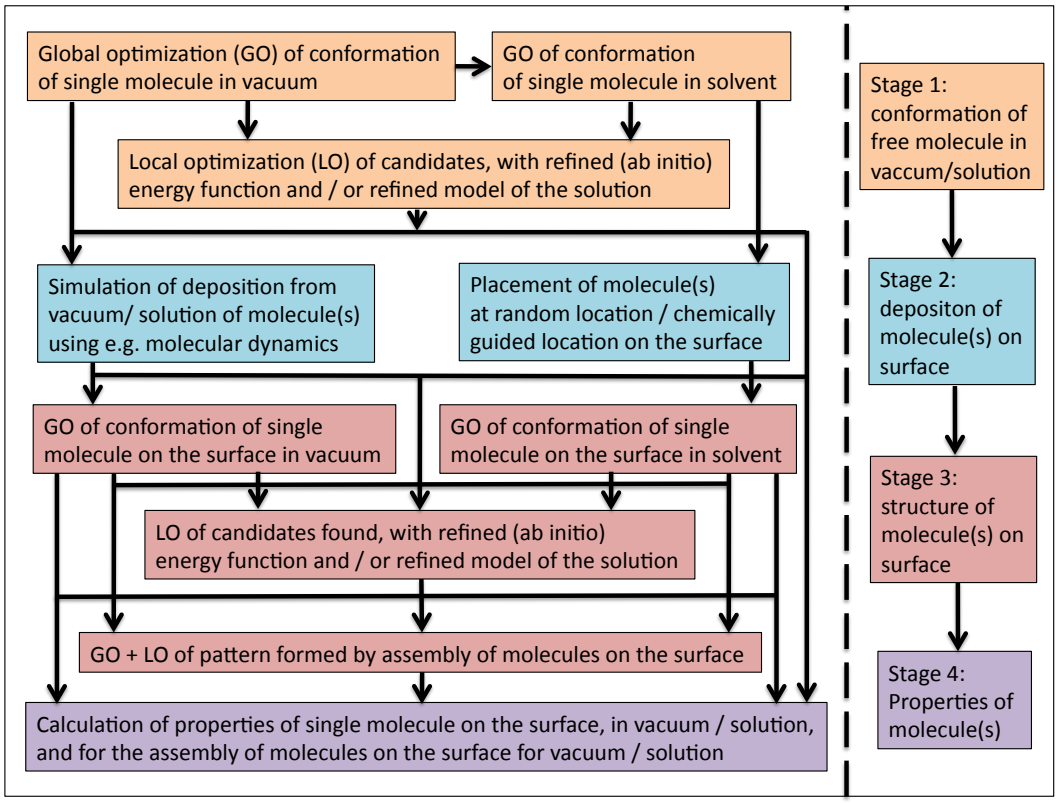
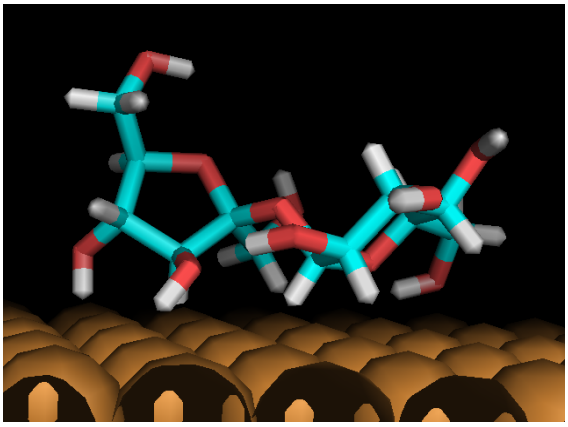
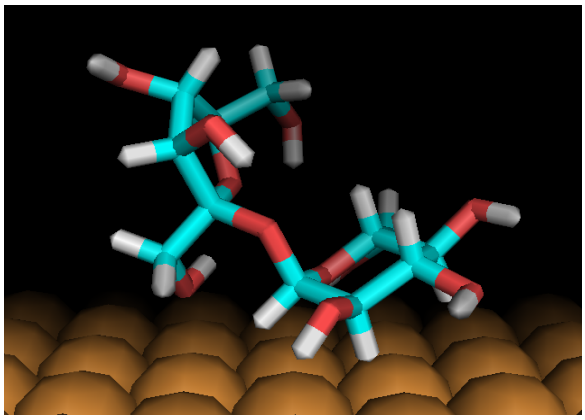
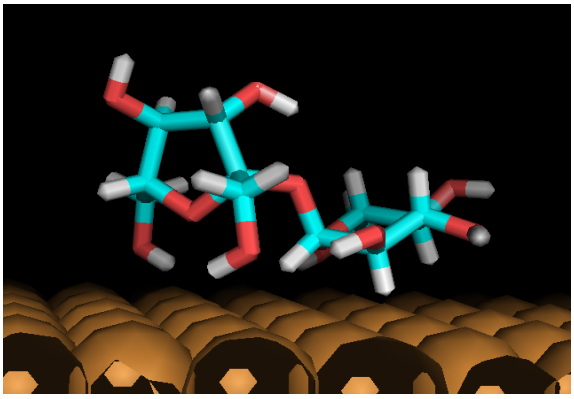

Global optimization (GO) of conformation of single molecule in vacuum
GO of conformation of single molecule in solvent
Local optimization (LO) of candidates, with refined (ab initio) energy function and / or refined model of the solution
Simulation of deposition from vacuum/ solution of molecule(s) using e.g. molecular dynamics
Placement of molecule(s) at random location / chemically guided location on the surface
GO of conformation of single molecule on the surface in vacuum
GO of conformation of single molecule on the surface in solvent
LO of candidates found, with refined (ab initio) energy function and / or refined model of the solution
GO + LO of pattern formed by assembly of molecules on the surface
Calculation of properties of single molecule on the surface, in vacuum / solution, and for the assembly of molecules on the surface for vacuum / solution
Stage 1: conformation of free molecule in vaccum/solution
Stage 2: depositon of molecule(s) on surface
Stage 3: structure of molecule(s) on surface
Stage 4: Properties of molecule(s)

Figure 4:

a) 

b) 

c) 

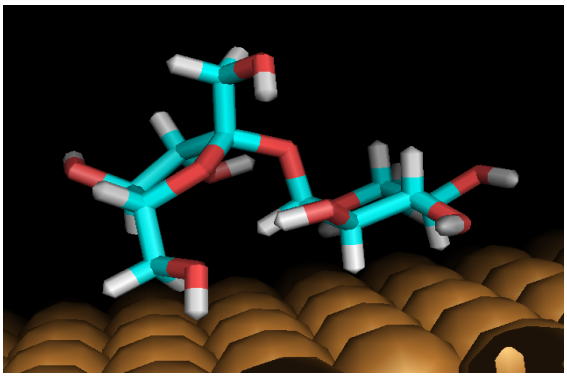
d) 

Figure 5: