# Modelling components of future molecular devices

**Thomas Trevethan**[1,2,3,4], **Alexander Shluger**[1,2,3,4] **and Lev Kantorovich**[5]

[1] Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[2] London Centre for Nanotechnology, University College London, Gower Street, London WC1E 6BT, UK
[3] The Thomas Young Centre for Theory and Simulation of Materials, University College London, Gower Street, London WC1E 6BT, UK
[4] WPI-AIMR, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai 980-8577, Japan
[5] Department of Physics, Kings College London, Strand, London WC2R 2LS, UK

## Abstract

We discuss challenges involved in modelling different components of molecular devices and give several examples that demonstrate how computer modelling evolved over the last few years to become a comprehensive tool for designing molecules, predicting their adsorption and diffusion at surfaces, simulating atomic force microscopy imaging and manipulation of atoms and molecules at insulating surfaces and studying electron conduction in prototype molecular devices. We describe some of the computational techniques used for modelling adsorption, diffusion, imaging and manipulation of organic molecules at surfaces and challenges pertaining to these studies, give several examples of applications and discuss further prospects for theoretical modelling of complex organic molecules at surfaces.

(Some figures in this article are in colour only in the electronic version)

## 1. Introduction

Understanding the adsorption and diffusion of large organic molecules on surfaces is important to many areas of surface science and nanotechnology, particularly catalysis, coatings, corrosion inhibition, tribology and molecular electronics. In the past few years there has been an explosion of interest in surface structures made from organic molecules for applications in devices such as mechanically flexible computers and displays, and 'molecular' computers, in which individual molecules replace transistors. Since the possibility of using individual molecules to perform computing functions has been proposed more than 30 years ago [1], 'molecular electronics' has evolved into an exciting research area owing to extensive experimental and theoretical efforts (see, for example, [2–5] and references therein). The advent of scanning probes, particularly the scanning tunnelling microscope (STM), opened new opportunities for imaging and manipulating individual atoms and molecules at conducting surfaces and creating prototype molecular devices [6–10]. Over the past two decades, an exquisite level of control has been achieved over the position and motion of individual atoms and molecules adsorbed on metallic and semiconducting surfaces using the STM, often at low temperatures. Determined efforts of organic chemists made possible designing and synthesizing molecules with pre-determined mechanical and electronic properties. It has been demonstrated that using individual custom designed molecules one can construct prototype molecular devices, for example, 'wires' [11, 12], 'switches' [13–16] and molecular diodes, transistors, rectifiers and small circuits [17–20]. Custom designed organic molecules can be also used as building blocks for the self-assembly of functional nanostructured materials [21–23].

Different applications bring special requirements for molecules and their interaction with substrates. In particular, reliable functioning of molecular electronic devices requires that the molecule has to be immobilized and stable at room temperature (RT), but the electronic structure of the molecule must be decoupled from that of the surface. It has been long recognized that the interaction between organic molecules and metallic surfaces can be fairly strong. Therefore, when organic molecules are adsorbed on a metallic substrate, the molecule–

1

surface interaction may induce changes in the molecular structure and properties as well as local modifications in the surface geometric and electronic structure. These properties helped to record excellent STM images of individual molecules on metal surfaces (see, for example, [24, 25]) and semiconductor surfaces [26–28] and to study their electronic properties. However, strong coupling to the substrate may alter the functional properties of a molecule and is detrimental to constructing electronic molecular devices due to electron leakage into the surface. Recently it has been demonstrated that a molecule can be electronically decoupled from a metallic surface by inserting an ultrathin insulating layer [29, 30]. Significant progress has been made on imaging and manipulating the charge state of molecules and atoms at thin insulator films on metal substrates using STM [29]. Molecular orbitals can also be decoupled from the clean silicon surface by its hydrogenation, which results in the passivation of its surface states [31]. However, metallic atoms or wires are likely to couple to a metal or semiconductor substrate [32, 33] and using this setup for real devices is not practical. Due to these reasons, wide gap insulators are more appropriate substrates for designing reliable molecular devices. Very little is known, however, about the adsorption and diffusion of organic molecules on wide gap bulk insulating surfaces or films.

However, to image the surface structure as well as to study properties of adsorbed molecules and manipulate their position at insulating surfaces is experimentally challenging. To study *individual* molecules one should use an atomic force microscope (AFM), which relies on detecting tip–surface forces rather than current. AFMs capable of atomic resolution are much more challenging to design and construct (see, for example, [34–36]) and low temperature UHV microscopes are still very expensive and rare. On the other hand, weak adsorption and high mobility of organic molecules at insulating surfaces often prevents AFM imaging of individual molecules at room temperature. The best studied is perhaps adsorption of perylene and its derivatives (e.g. PTCDA and PTCDI). They represent a prominent class of molecular organic semiconductors and have been imaged on several insulating substrates, such as KBr, KCl, NaCl, mica and TiO$_2$(110) [37–43]. Another molecule which has been studied on insulators is C$_{60}$ [44–48]. However, in this case individual molecules have been observed only rarely and usually associated with surface defects [48].

In order to fully exploit the potential of new nanowires and nanodevices on insulating substrates one needs to learn how to control individual molecules and make perfect molecular films at these surfaces. This requires understanding the diffusion mechanisms and film growth processes at a fundamental level. One of the main challenges is designing molecules with desirable functionality and with predictable and controllable adsorption and diffusion characteristics on a given substrate. As has been pointed out in [49], a major chemistry challenge is in equipping the molecule with lateral chemical groups, not contributing to the function directly but protecting the molecular electronic functionality and assembling and stabilizing the molecule on a given substrate. However, no
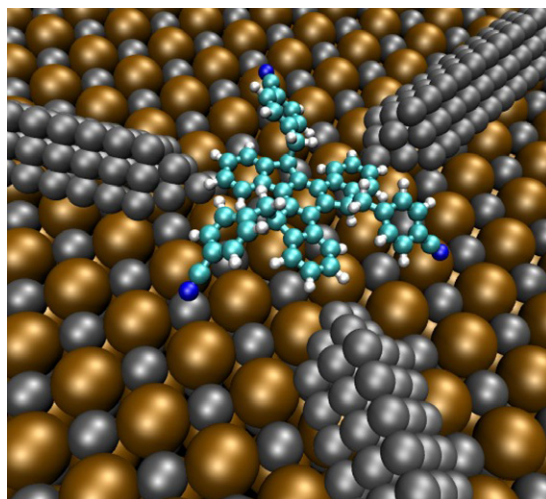


**Figure 1.** Illustration of a prototype molecular electronic device. A molecule on an insulating substrate is connected to several metallic electrodes.

method has been suggested so far for tailoring the molecule–substrate interactions to control the molecular diffusion to produce films or molecular devices with desired morphologies and trials have been mainly empirical.

After such molecules are designed, synthesized and deposited on a particular substrate, they should be connected to conducting wires and their functioning should be tested, see figure 1. The initial deposition may require manipulating molecules to pre-fabricated metal clusters or wires. Building atomic conductive wires to connect to immobile molecules is another challenge, which may require using AFM for manipulating metal atoms into preferred positions [50]. Conductivity in the system is determined by alignment of electronic states and rate of electron transfer between molecules and metal wires [5, 49]. This can be affected by surface defects. The theoretical modelling suggests that defects can be manipulated [51] closer to a molecule to purposely affect its conductivity [52]. It has been demonstrated experimentally that one can detect and change charge state of defects at insulating surfaces using a single electron tunnelling force spectroscopy [53]. However, previous experimental attempts to image, manipulate and control atoms and molecules at insulating surfaces have rarely been successful. Defects have been imaged and occasionally manipulated at several surfaces using non-contact (NC)-AFM (see, for example, [54, 55]) but their chemical identities often remained unclear [54].

It has long been appreciated that modelling is vital for both interpreting the AFM images and designing procedures for manipulating atoms, molecules and charges at surfaces [34, 56]. Computer modelling can also be used for pre-screening and designing molecules with required functionality and which can be deposited and immobilized at insulating surfaces. In this paper we discuss challenges involved in modelling different components of molecular devices and give several examples that demonstrate how computer modelling evolved over the last several years

to become a comprehensive tool for designing molecules, predicting their adsorption and diffusion at surfaces, simulating AFM imaging and manipulation of atoms and molecules at insulating surfaces and studying electron conduction in prototype molecular devices. In section 2 we describe some of the computational techniques used for modelling adsorption, diffusion, imaging and manipulation of organic molecules at surfaces and challenges pertaining to these studies. In section 3 we give several examples of applications and in section 4 discuss further prospects of theoretical modelling of complex organic molecules at surfaces.

## 2. Computational methods and challenges

Theoretical modelling of the adsorption and dynamical behaviour of large organic molecules on insulating surfaces is crucial for the design and implementation of molecular electronic devices and functionalized molecular layers. For a single molecule device, the molecule must be immobile on the surface, often at room temperature, for long periods. In other applications, such as in the growth of thin films and self-assembled layers, the free diffusion of molecules on a surface is a requirement, and controlling the rate of surface diffusion can often significantly affect the morphology of the layer. Using modelling for designing of such systems has huge potential; however there are also significant difficulties in using traditional modelling methods in realizing this aim. The first of these is that the types of systems of interest here are very large on the atomistic scale, and are also often very complex—the molecules involved may be composed of hundreds of atoms, interacting with thousands of atoms in the substrate surface. This means that novel approaches must be employed to understand the behaviour of these systems. The use of a full *ab initio* treatment, such as density functional theory (DFT) [57, 58] is often not feasible for systems of this size (an isolated large organic molecule interacting with a surface will require a simulation cell containing many hundreds or thousands of atoms). Finding all stable configurations of molecules at surfaces and their diffusion paths is also challenging due to the richness and complexity of potential energy landscape. Finally, modelling diffusion and manipulation of molecules at surfaces is only starting to develop. Below we outline several approaches, which have been implemented for modelling realistic systems, and some of the main challenges involved.

### 2.1. Interaction of organic molecules with insulating surfaces

The first task concerns simply finding the most stable configuration of an individual molecule on a surface. Brute force approaches rely on performing full quantum mechanical calculations using DFT or less accurate methods, such as tight binding. One of the challenges is the sheer scale of the calculation. This can be overcome for relatively small molecules by using efficient parallelized DFT codes, some of which scale quasi-linearly with the system size. For example, the calculations of the electronic structure and STM images of a single pentacene molecule adjacent to a gold

atom adsorbed on a NaCl film on the Cu(100) substrate have been performed using the VASP code [59]. An extensive study of adsorption, diffusion paths and manipulation of $C_{60}$ molecule on the Si(001) surface have been carried out using the SIESTA code [60, 61]. Calculations employing standard DFT methods and hybrid functionals have been performed for e.g. ruthenium polypyridyl complexes, such as $Ru^{II}(tpy)_2$, attached to $TiO_2$ nanoparticles used as photo-sensitizers for solar energy conversion applications [62].

A hierarchy of the methods combining quantum mechanical and classical atomistic simulation techniques has been developed for studying the structure and properties of complex organic and biological molecules and their interaction with inorganic surfaces. These methods are broadly attributed to the so-called QM/MM (quantum mechanics/molecular mechanics) and embedded cluster techniques and share many common features (see, for example, reviews [66–68]). In these methods, the whole system, e.g. a molecule adsorbed on a crystal surface, is partitioned into two subsystems: one of them is treated quantum mechanically and is linked or embedded into the rest of the system, treated classically. This approach is rarely used for combined systems, such as organic molecules at insulating surfaces, which are of interest here, for several reasons. Firstly, there are few density functionals universally applicable for treating both organic molecules and inorganic, e.g. metallic, surfaces with the same accuracy. Secondly, there are almost no classical force-fields for describing the interactions between organic molecules and inorganic surfaces. Recently, however, the interaction of hydrocarbon and basic polar organic molecules with the MgO(001) surface has been theoretically studied using an embedded cluster model and a hybrid density functional [52]. In this case, the molecules were small enough that the whole molecule and a quantum cluster representing the surface could be treated quantum mechanically. Luckily, the B3LYP density functional is known to give accurate predictions for a wide range of organic molecules as well as for wide band gap insulators, such as MgO. It was found that both methane and benzene molecules do not bind strongly to the perfect surface and will not be deprotonated. Polar organic groups interact more strongly with the surface, with nitromethane and pyridine providing the largest binding energies (0.37 and 0.75 eV, respectively); however, in each case the molecules were physisorbed and not strongly polarized by the surface. Because of the large band gap of MgO, there were no electron transfer between molecules and the surface found [52].

Another approach is to resort to either semi-empirical (e.g. tight binding) or fully classical simulation methods. These methods, however, require external parameterization or derived force-fields, and care must be taken to ensure that these accurately represent the system. There exist many reliable and well-tested force-fields for describing a wide range of isolated organic molecules and complexes of organic molecules, and also of many ionic (or semi-ionic) substrates [63–66]. However, reliable force-fields for the interaction between the two are still rare, due to the novelty of these types of the system.

A relatively weak interaction not involving an electron transfer, as has been found for the interaction of hydrocarbon
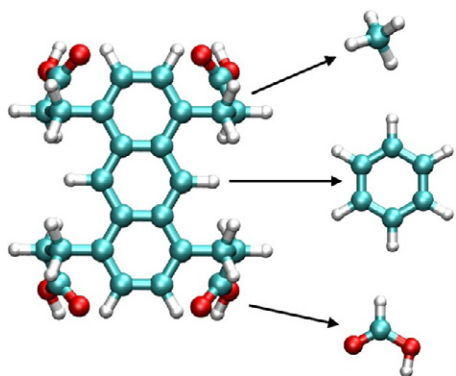
**Figure 2.** Schematic illustrating the splitting of a large organic molecule into its basic components. Here, a functionalized anthracene molecule can be considered as consisting of aromatic rings (benzene), methyl groups (methane) and carboxylic acid groups (formic acid).

and basic polar organic molecules with the MgO(001) surface [52], is characteristic to many other organic molecules at insulating surfaces. This feature combined with the idea of partitioning of large molecules into smaller fragments has been exploited in [69] in order to derive a classical force-field for the interaction of large organic molecules with the (110) surface of rutile $TiO_2$. It has been demonstrated that it is possible to construct a reliable full molecule + surface force-field by breaking a big molecule into simpler molecules (adding additional atoms may be required to terminate their bonds), fitting a force-field for the interaction of each of these smaller molecules with the surface using pair-wise interactions, and then calculating the interaction of the big molecule with the surface as a sum of the pair-wise potentials derived for small molecules. This requires high-quality *ab initio* calculations for the adsorption of an array of small organic molecules on the rutile surface in a wide range of different configurations. The quantum mechanical force-field obtained is then approximated by a sum of pair-wise inter-atomic potentials between each small molecule and surface atoms. The force-field for the entire big molecule is then obtained by combining the pair-wise terms from each smaller molecule and removing contributions due to artificially added atoms (see figure 2). The interactions between atoms inside the big molecule and within the surface, in turn, can be treated by using standard force-fields [70, 71]. The force-field derived in this way is then combined with the force-fields for the intra- and inter-molecular interactions and the intra-surface interactions to yield a force-field for the entire system. This method has been applied to the interaction of large organic carboxylic acids with the $TiO_2(110)$ surface, and has been successfully tested against large *ab initio* calculations to reproduce well both binding energies and diffusion barriers of larger molecules [69, 72].

The approach to deriving the force-fields described above is appropriate for considering physisorption and ionic bonding of organic molecules at the surface. It is unable to describe covalent bond formation/dissociation, i.e., the processes associated with charge transfer between a molecule and the surface. Results in [69, 72] demonstrate that a separate

description of the interactions of individual functional groups of larger molecules with the surface can be used to describe the interactions of the whole complex molecules with the surface. However, this method of deriving the force-field crucially depends on the way in which big molecules are replaced/represented via a set of smaller ones. It is only valid if the electronic structure of the functional groups in the complex molecule does not differ much from the electronic structure of the small molecules representing those groups. For example, the substitution of the hydrogens in benzene with methyls satisfies this requirement. Aliphatic hydrocarbon chains also can be considered as combinations of methyl groups. However, this approach is less applicable for highly conjugated systems such as polycyclic aromatic molecules. These cannot be divided into benzene derivatives since the electronic structure, and therefore the interaction energy with the surface, of the separate parts and the whole molecule can prove to be quite different. The approach implemented so far relies upon chemical intuition and requires further justification and refinement.

## 2.2. Adsorption of large and complex molecules on insulating surfaces

Given that there is a reliable model for a molecule, the surface and the interaction between them, the determination of adsorbed structures is still not straightforward due to complexity often present in this type of system. In the adsorption of relatively small molecules on particular surfaces, it is often sufficient to consider a limited number of initial configurations of a molecule and then to perform structural optimizations of each of these to identify the locations and energies of the accessible minima on the potential energy surface of the system. A good example of this type of study is modelling of $C_{60}$ on the Si(001) surface [73–75]. In recent experiments, $C_{60}$ molecules were deposited on the Si(001) surface and imaged with the STM with submolecular resolution [76]. In addition, it was shown that it is possible to manipulate $C_{60}$ molecules along the dimer rows of the Si(001) surface with the STM tip [76]. In order to understand the nature of the adsorption and the mechanism of manipulation, extensive calculations of this system were performed using the generalized gradient approximation (GGA) and localized density approximation (LDA) of DFT methods. It was found that the molecule forms strong covalent bonds with the surface dimers, which causes the $C_{60}$ cage to significantly distort. Over 20 distinct adsorption sites were found (local minima), corresponding to different orientations of the $C_{60}$ molecule and different positions with respect to the surface structure [61, 77].

Many experiments have also been performed adsorbing $C_{60}$ molecules on pure insulators such as alkali halides, specifically the KBr(001) surface [78], where they form monolayers observed using NC-AFM. Isolated molecules are not observed on these surfaces due to the weak molecule–surface interaction. The atomic and electronic structure of a single $C_{60}$ molecule adsorbed on the KBr(001) surface was investigated in [79] using GGA and LDA DFT. It was found that the molecule preferentially adsorbs above K atoms in the

surface, but with a small adsorption energy of 0.37 eV (LDA) and there is negligible change in the structure of the molecule or surface.

It is important to stress that quantum mechanical calculations do not always give unambiguous answers even in relatively simple cases. In particular, both of the $C_{60}$ studies described above found that the use of LDA functional leads to significantly greater adsorption energies for the $C_{60}$ molecule than obtained using GGA and on the KBr surface GGA leads to no binding at all. As is the case with all standard DFT methods, the contribution of dynamic correlation effects, specifically of the van der Waals attraction, is significantly underestimated.

These calculations treated a relatively rigid $C_{60}$ molecule, which due to its high symmetry, and the symmetry of the surface, allowed all possible adsorption geometries in both cases to be found. However, in larger and more complex systems there are often too many degrees of freedom that need to be included explicitly and many possible configurations may exist, so that derivation of all adsorption sites and configurations for complex molecules within this approach would be not feasible in practice. Moreover, interaction with the surface may significantly alter the conformation of the molecule and thus change its symmetry, which may make it extremely difficult to systematically explore all the adsorption geometries. In this case, global optimization methods and methods to explore the configuration space of the system must be employed to map out all the minima in the potential energy surface (PES) of the system. The latter may be extremely complicated, however, with the number of minima growing exponentially with the number of degrees of freedom [127].

The most simple and common of the global search methods is to use simulated annealing in order to try and escape local states and access low lying potential energy minima, which is implemented either as a Monte Carlo or molecular dynamics (MD) algorithm [80]. In the case of molecular dynamics simulated annealing (MDSA) [81], the system in an initial configuration is run at a high temperature (set so as to initiate structural transitions and diffusion) for a long period, and then slowly cooled down to zero, again over a long period (on MD timescales, typically nanoseconds). This procedure is repeated for many different initial configurations: a good indication that an optimized structure found is a global minimum is that it is arrived at from distinctly different initial configurations. However, there is often no guarantee that the global (thermally accessible) minimum has been found, as this would require an infinitely long simulation time. There are some other methods, such as e.g. genetic algorithms [82] and basin hopping algorithms [83], which have been shown to be very useful in mapping out the PES of the system and finding the global minimum on it, especially for complex molecular systems.

So, in spite of tremendous progress achieved so far for relatively simple systems, complexity of systems that find practical applications requires further developments of methods that allow mapping out the complicated PES. The main problem is in finding an efficient total energy technique, which would be able to locate all essential minima correctly prior to running more precise (and more expensive) calculations to find them with the required precision.

## 2.3. Diffusion of large and complex molecules over long timescales

To model the mobility of a molecule adsorbed on a surface, dynamic methods are employed, which are capable of evaluating the diffusion of the molecule on a surface at a finite temperature. In particular, several $C_{60}$ diffusion pathways were considered along the rows of the Si(001) surface using a constrained minimization method [75]. The molecule was 'pulled' over a potential energy barrier by moving a single coordinate of a carbon atom in the molecule and relaxing all other degrees of freedom at each step. It was found that the $C_{60}$ molecule can 'roll' along the trenches of the surface, by pivoting on the two front bonds formed with the surface dimers. In this system the diffusion of the $C_{60}$ molecule can be approximated as single jumps between neighbouring equivalent adsorption sites, crossing a single potential energy barrier because the molecule is sufficiently rigid and hence does not undergo any conformational change during the transition. The modelling of this transition was facilitated by a very simple choice of the single coordinate that takes the molecule from one energy minimum to the other. The real challenge is to consider diffusion of more complex molecules possessing 'internal' degrees of freedom, which may undergo sufficient changes along the diffusion path (e.g. bending of some groups).

The most direct way of modelling diffusion at a finite temperature is to explicitly follow the trajectory of the atomic coordinates of the system in real time using a molecular dynamics (MD) algorithm [84]. In such a simulation, an isolated molecule, adsorbed on a particular surface in one of its minimum energy configurations, would be evolved at a certain temperature, and the rate of diffusion across the surface can be determined from the mean square displacement of the molecule as a function of time. Given that the MD timestep is small enough, this will follow the dynamical evolution of the system exactly, and account for all vibrational modes in the surface and molecule and all dynamical correlation effects. However, the timescale of MD simulations is limited due to computational expense, and can only be employed directly when diffusion is fairly rapid on a surface (i.e. the barriers to cross between minima on the PES are comparable to $k_B T$). The timestep in an MD simulation is usually 1 fs or less, so as to capture the fast vibrations of the surface atoms and covalent bonds. This limits the length of a run to nanoseconds (millions of steps) or at most micro-seconds (billions of steps), depending on the method used to calculate the forces on atoms (DFT versus force-field based methods). Often in the case of large molecules and the types of the system of interest here, the timescale of surface motion is much longer: residence times of molecules may reach seconds, minutes or hours, even at room temperature, and therefore it is not possible to employ conventional MD to investigate these systems.

Another method of studying the diffusion of adsorbates at surfaces simulates the real dynamics by a sequence of jumps between potential energy minima accessible to the system ('the states'). The transition rates for crossing the energy barrier from one state to the other may be calculated from, e.g., transition state theory (TST) [85]. According to TST, the

rate is determined using an Arrhenius type expression with the activation energy and a prefactor. The latter is calculated from the system vibration frequencies at the minimum and the saddle point. Therefore, both the activation energy and the prefactor can be calculated from atomistic calculations and then used to determine the rate for that transition. This is often useful when considering the behaviour of adatoms and small molecules, when using the symmetry simplifies considerably locating the corresponding transition states (or saddle points). However, in the cases of larger and more complex molecules there may be many different processes and transitions involved in the diffusion mechanism, and finding exactly the saddle point, needed to calculate both the energy barrier and the vibration frequencies, may be problematic. If it is possible to determine all of the thermally accessible states in a system and calculate the rates of all the individual transitions that make up the diffusion mechanism, then the dynamical evolution of the system can be evaluated using a kinetic Monte Carlo (KMC) algorithm [90]. In this method, the dynamics of the system is followed explicitly in real time, and the rate of overall diffusion can be determined from the average behaviour. Due to the simplicity of the KMC method, and the fact that rates are calculated in advance, long timescales are easily accessed.

An alternative method is based on identifying the essential reaction coordinates, $s$, responsible for the transition (sometimes called *collective* coordinates) and then calculating the Landau free energy, $F(s)$, as a function of all collective variables for the given temperature using MD. Then, the transition rate can be calculated from the free energy and its derivatives using the Kramers' formula [86]. Several methods have been developed for calculating the free energy on the subspace of collective coordinates [87]. The recently developed metadynamics method [88] is becoming very popular for mapping out essential energy minima and calculating the free energy, and has a great promise. This method is based on adding dynamically a basin potential to the system in order to 'flood' the visited potential energy minima and hence force the system to explore other minima. The calculation stops when all the minima in the free energy landscape are found. Unfortunately, metadynamics is very expensive computationally as it requires rather long MD runs; it is also limited in practice to systems, which can be adequately described by only 2–3 collective coordinates, although attempts have been made to overcome these limitations [89].

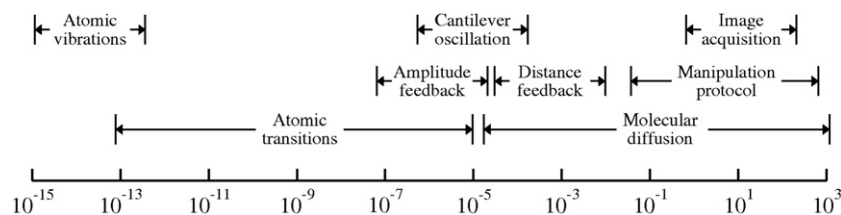### 2.4. Modelling imaging and manipulation with scanning probes

In order to image and manipulate the configuration and position of an adsorbed molecule, a scanning probe method must be employed, and in the case of insulating surfaces this must be the NC-AFM. One of the major problems in NC-AFM imaging of a molecule at a surface is in achieving simultaneous resolution of both the position of the molecule and of the surface atoms, as the interaction of the tip apex with the molecule and the surface can be very different. To understand the origin of contrast in images and protocols for
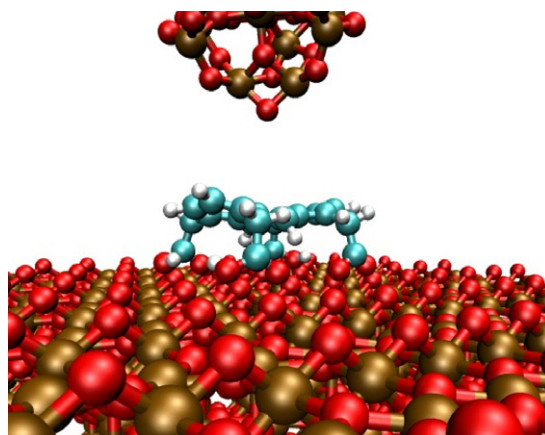


**Figure 3.** Illustration of the configuration of a tip–molecule–surface junction in an atomistic simulation of the AFM imaging/manipulation process.

controlled manipulation, it is essential to model the interaction of the AFM tip with the molecule and the surface and also the dynamical operation of the NC-AFM instrument. To model the interaction of the molecule–surface system with an AFM, an atomistic model of the tip is required, especially of its atomistic structure at its very end: the apex. The exact structure and chemical identity of the tip apex in a NC-AFM experiment is rarely known, due to exposure to the atmosphere and the possibility of transfer of atoms or molecules from/to the surface. Even so, a large number of stable tip apex structures can be envisaged [91–93]. This usually means that the imaging needs to be modelled with several tip models to determine the identity of the tip apex. In addition, modelling of the imaging process with the NC-AFM requires that the entire tip–surface force-field for a particular tip model is calculated—which is computationally much more demanding that simply determining the properties of the molecule–surface system (see figure 3). Here, in all but the most simple and symmetric systems, the use of *ab initio* methods is not feasible, and inter-atomic potentials must be used. These potentials should be capable of reliably reproducing both the interaction of the tip apex with the surface, and with the adsorbed molecule in a wide range of their relative positions.

Modelling the imaging process with the NC-AFM typically involves performing static atomistic calculations of the force-field experienced by a particular tip above the surface–molecule system, from which the frequency shift, as a function of cantilever position, can be determined given experimental parameters [56]. This procedure, however, involves several simplifying assumptions, such as that: (i) the instrumentation controlling the cantilever oscillations and position reacts ideally and instantaneously on the change in the tip–surface interaction (as the surface is scanned); (ii) the tip experiences a conservative force-field due to its interaction with the surface, i.e. that any tip–surface relaxation can be considered adiabatic; (iii) the general atomic configuration of the surface and of the tip do not change during imaging. These assumptions are often valid. However, when structural changes and dynamical processes occur on the surface or tip

**Figure 4.** Illustration of the different timescales (in seconds) relevant to the NC-AFM imaging and manipulation of molecular adsorbates.

during the imaging process, the interpretation and modelling of experiments becomes considerably more complex, and must consider all the timescales spanned in figure 4. Structural changes occurring during the acquisition of an image are often observed experimentally with the NC-AFM. When a structural change at the surface is induced by the tip in a determined way, this constitutes a controlled manipulation process. Random, un-controlled, manipulation of adsorbed atoms or molecules is often unavoidable, but there are many instances when NC-AFM is merely imaging stochastic diffusion of these species. In all these cases, the three conditions outlined above are violated and the tip oscillation frequency, scanning speed and a response time of the electronics controlling the AFM are important for understanding the image.

Consider an example of a molecule rotating between several equivalent minima on the surface. The rate of hoping between different minima will depend on the temperature. In the NC-AFM, the tip is oscillating above the surface and the force-field experienced by the tip due to the interaction with the molecule will be different before and after the molecule changes its position. The inverse of the rate of the molecular transition on the surface (i.e. the jump from one potential energy minimum to another over the barrier) may be comparable to the timescales of the NC-AFM tip oscillations and even the period of taking a scanline, however, the actual transition itself occurs on the timescale of atomic scale relaxation. Therefore, as far as the NC-AFM is concerned, an atomic scale structural change will occur effectively instantaneously with respect to the motion of the tip, and it is sufficient to simply evaluate when transitions will occur on the timescale of this motion (see figure 4). These transitions are stochastic processes, which can be modelled accurately and realistically using a KMC method if the rates are known. The additional problem here is that the rates are not fixed, but change because of the tip oscillations since the tip position with respect to the molecule affects the corresponding energy barriers. These processes can be modelled in real time using a variant of the KMC proposed in [94], which allows for this to be taken into account explicitly. In the methodology described in [95] the change of the potential energy barriers for the surface process as a function of the tip position above the surface were included, which allowed to take account for the role of the tip in affecting the dynamics of the surface process.

When a jump occurs, the instrument will not react instantaneously to the change in the tip interaction with the surface. To model the complex behaviour of a realistic NC-AFM instrument, a virtual atomic force microscope (VAFM) [96] is employed, which consists of an explicit

numerical simulation of the entire experiment that can be used to model the imaging process in real time. The VAFM is similar in setup to a typical NC-AFM experiment [34, 35] with both automatic gain control (AGC) and automatic distance control (ADC) feedback loops controlling the cantilever excitation and surface position, respectively. The tip moves in a three-dimensional force-field above the surface as the tip is oscillated and the surface scanned, which is determined by interpolating between points on a grid of forces calculated from the atomistic simulations. The imaging scan in a VAFM simulation takes place continuously above the surface and the oscillating tip follows a lateral path collecting line scans in order to produce an image, exactly in the same way as in real NC-AFM experiment. In this way the simulation of the imaging/manipulation process with the VAFM can occur in real time in a self-consistent way, including the effect of structural changes in the system. However, in order to accurately model the effect of the tip on transitions in the surface, the thermally accessible potential energy surface needs to be determined as a function of the tip position above the surface. This can be an extremely expensive calculation, as the barriers for each process need to be calculated for many different tip positions. In many complex systems, including large adsorbed molecules where there are many possible thermally accessible transitions, this may not be computationally feasible.

Therefore, the real challenges here are to extend this technique to adsorbed systems containing many internal degrees of freedom like complex organic molecules, take full account of the dynamic processes taking place at the timescales characteristic for the tip oscillations (and faster) and account for the instrument response in a self-consistent manner.

## 3. Applications of modelling: adsorption, diffusion and manipulation of atoms and molecules

The methods described in the previous section have been applied to several technologically important molecule–surface systems in order to understand the nature of the molecule–surface interaction and mechanisms of surface diffusion. Some examples of these applications are described in the following three subsections. Section 3.1 concerns the adsorption and diffusion of flexible, functionalized molecules adsorbed on the $TiO_2(110)$ surface employing derived inter-atomic potentials. Section 3.2 describes examples of modelling the imaging and manipulation of molecules and adatoms using NC-AFM. Finally, in section 3.3 we describe how one can manipulate the electronic structure of surface defects using the tip of an AFM.
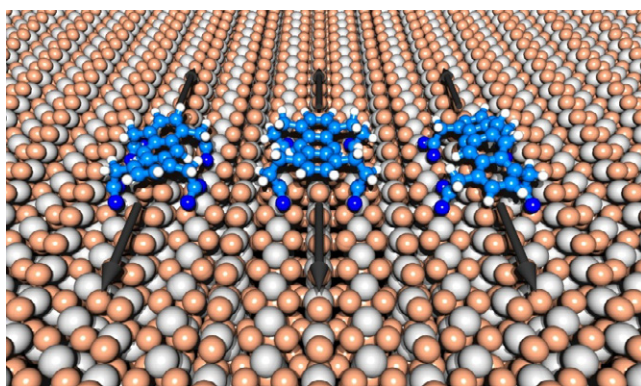
**Figure 5.** Configurations of each of the molecules adsorbed in their ground states on the TiO$_2$(110) surface. Molecule (i) left, (ii) centre and (iii) right.

### 3.1. Adsorption and diffusion of large organic molecules on the TiO$_2$(110) surface

In some of the examples described above, e.g. C$_{60}$ on Si(001), the molecule adsorbed on the surface is relatively rigid and, at the given adsorption site, can occupy only a few molecular conformations, even though there are many different adsorption sites. In the case of more flexible molecules with much softer degrees of freedom, the molecule–surface system can become significantly more complex. In particular, for molecules adsorbed on insulating surfaces, only the polar parts of the molecule will bind strongly to the surface. There may often be several such polar parts in the molecule, so-called 'binding groups'. In this case there will be many more possible configurations of the molecule on the surface, which will be determined by the overall structure of the molecule and the surface. To explore some of the general issues in relation to the mobility of such molecules on crystal surfaces, several structurally similar aromatic molecules adsorbed on the TiO$_2$(110) surface have been studied in [97]. These molecules, which are shown in figure 5, consist of an anthracene or tetracene 'backbone' with four attached carboxylic acid groups (either −CH$_2$COOH or −(CH$_2$)$_2$COOH) in the 1, 4, 5 and 8 positions. This set of molecules was chosen as they have the same functional groups but differ only in the way they are connected together through the molecular 'backbone'. The different backbone structures change both the distance between the groups, the extent to which they are free to move independently of each other, and the strain inside the molecule and inside the surface due to the attachment of the functional binding group to the surface Ti ions (the 'commensurability').

Figure 5 shows the lowest energy configurations of the three molecules (i), (ii) and (iii) adsorbed on the surface, determined from calculations that employ a set of inter-atomic potentials that have been specifically developed for this system and are described in [69]. For molecule (i) the anthracene backbone straddles and is parallel to the oxygen rows and each of the binding groups bonds to two surface titanium atoms in a bi-dentate configuration along the Ti rows, with two groups adjacent to each other in each row. The four pairs of surface Ti atoms that the carboxylic acid groups bind to are

highlighted in the schematic in figure 5. The carboxylic groups are strongly bound to the surface Ti atoms—with a binding energy of approximately 1.5 eV per group—and in this way form a structural motif that defines a potential energy basin on the potential energy surface of the molecule–surface system.

The configuration of the molecule described above is the global minimum energy state of the molecule–surface system. However, there are other thermally accessible energy basins with the binding groups in other positions along the Ti rows. The diffusion of the molecule as a whole consists of a lateral translation of the above configuration into an equivalent position along the rows (either backwards or forwards). It will occur via displacements (or 'steps') of individual binding groups (or 'legs') and in this process the system visits intermediate energy basins, which are separated from each other by potential energy barriers (the molecule 'walks'). The existence of the intermediate states and the heights of the barriers that must be crossed to access them will be determined by the overall structure of the molecule, the freedom with which the binding groups can move in relation to each other and the flexibility of the molecular backbone.

The accessible non-equivalent energy basins and the potential barriers, which separate them for molecule (i) are shown in the schematic in figure 6. In this diagram the configuration of the molecule is depicted in terms of the positions of the binding groups along the two adjacent Ti rows, which define the potential energy basins. Given the complete description of the potential energy surface in terms of the basins characterized by the positions of the binding groups, it is possible to evaluate rates of individual transitions and hence the explicit real time dynamical evolution of this system using the kinetic Monte Carlo method [94], where the rates for individual transitions are determined from the potential energy barriers.

For molecule (i), the lowest rate (highest barrier) for an individual transition (moving into a leading leg state from the ground state) is in the region of a kHz, yet the effective rate for the transition of the molecule as a whole to a neighbouring ground state is only 0.12 Hz—a residence time of 4.3 s. The reason for this complex behaviour is that on average about 44 000 individual intra-molecular transitions occur for every one complete lateral molecular transition. Molecule (ii) has the same anthracene backbone, but the legs that attach to the binding groups are longer, with an additional −CH$_2$− group separating the binding groups from the molecular backbone. This effectively changes the degree of freedom that the binding groups have to move independently and will change the potential energy landscape of the system. The global energy minimum for this molecule has the same binding group structure as for molecule (i) above, and the same number and type of intermediate basins, but the energy barriers are significantly different. Most importantly, the barrier to enter the leading leg state and also the barrier to return to the ground state from it are both significantly higher than for molecule (i), at 0.64 eV and 0.38 eV respectively—a direct consequence of the increased flexibility of the legs. This has a substantial effect on the overall effective rate of diffusion of the molecule, which is now 3.5 Hz, i.e. approximately 30 times faster than for the molecule (i). For this system, on average only 16
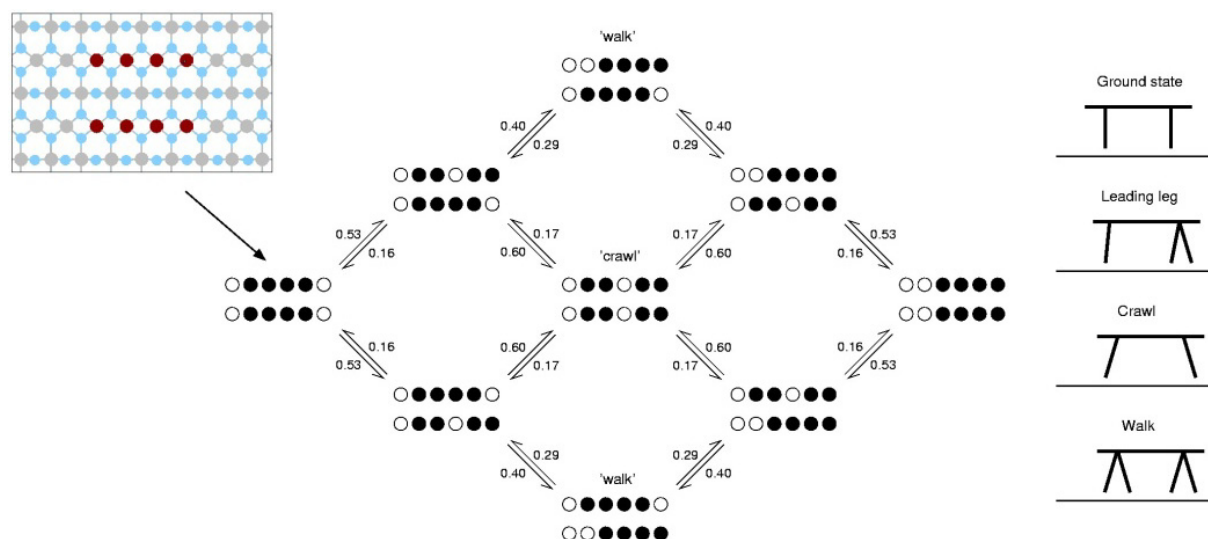
**Figure 6.** A map of the potential energy surface of molecule (i) adsorbed on the TiO$_2$(110) surface, in terms of the position of the binding groups with respect to the Ti atoms along two adjacent rows. The left configuration is the ground state, and the far right is the adjacent equivalent ground state along the same Ti rows.

intra-molecular transitions occur for each complete molecular transition.

Molecule (iii), which has a tetracene backbone, but the same length 'legs' as molecule (i), is an example of keeping the same flexibility, but changing the separation of the binding groups and hence the commensurability of the molecule with the surface. The structural motif of the ground state of this molecule adsorbed on the surface is the same as in the previous two cases, as is the number and type of accessible intermediate states, with only the energies of the states and the separating barriers differing. In this system the barriers for many transitions are significantly less than for the other two molecules, and this has the effect of making this molecule highly mobile on the surface resulting in an effective diffusion rate of 3.7 kHz. The main reason for this is that, due to the different spacing of the legs, this molecule is significantly less commensurate with the surface than molecule (i).

These dramatic differences in the overall rate of diffusion of the three molecules on a surface that reduces their motion to one dimension show how the mechanical properties, shape and flexibility of a large molecule can significantly affect its behaviour. The differences in the mobility of the molecules described here are not due to changing the way the molecule interacts with the surface (which occurs directly through the binding groups), but by modifying the potential energy landscape of the system through the overall structure of the molecule. In this system, the hydrocarbon part of the molecule, which is the only part that is varied, interacts very weakly with the surface and changing it slightly affects the interaction of the molecule with the surface. In particular, both molecules (i) and (ii) have very similar adsorption energies, however the actual effective diffusion rates that result from transitions via the intermediate states differ by two orders of magnitude.

The method employed here makes it fairly straightforward to see how in these relatively simple systems the different structures result in the corresponding potential energy surfaces

and energy barriers, and how these in turn result in the different dynamical behaviours of the systems. For other molecules and surfaces this of course may be less intuitive and more challenging to model. However, these calculations demonstrate how molecular structure can be designed to tailor the mobility of a molecule for a specific function. At the same time, these simulations pave a way to investigating diffusion of more complex molecules strongly bound to surfaces.

### 3.2. Imaging and manipulation of atoms and molecules

There have been several attempts to model the imaging and manipulation of molecules adsorbed at surfaces with NC-AFM, however these calculations are usually very limited due to the system size and unknown nature of the tip apex. Since an accurate description of the tip–molecule interaction is essential to understanding the contrast in these systems, fully *ab initio* methods are usually employed. The imaging of relatively small molecules on the oxide surfaces TiO$_2$(110) and MgO(001) was investigated in [98]. The adsorption of the formate ion and 3-{4-[tris-(3,5-di-tert-butyl-phenyl)-methyl]-phenoxy}-propionic acid (C$_{52}$H$_{72}$O$_3$) was considered on both surfaces employing several types of oxide tip models. The results of this modelling suggested that flat molecules can be identified by their shape, but that simultaneous atomic resolution inside the molecule and on the substrate under the same imaging conditions is not feasible in this type of systems using a purely topographic mode.

The imaging and manipulation by the tip at close approach of the larger C$_{60}$ molecule adsorbed on the Si(001) surface was modelled using DFT and a silicon tip model [75, 99, 100]. The imaging simulations [100] demonstrated a possibility of obtaining submolecular resolution of the molecule with the NC-AFM so that two particular orientations of the C$_{60}$ became distinguishable. The manipulation mechanism consists of the lowering of the energy barrier for the C$_{60}$ diffusion due to

the interaction of the molecule with the AFM tip and the subsequent thermally induced movement of the molecule over this barrier. This was modelled by performing numerical simulations of these energy barriers for a series of tip positions relative to the molecule to show how the barriers change with the tip position, and then using the VAFM to simulate the reaction of the instrument on the successful manipulation of the molecule jumping (in fact, 'rotating') from one minimum to the other. The results demonstrate that manipulation of the molecule, which is covalently bound to the surface, is possible with the NC-AFM, even though there is no continuous tip–molecule contact, opposite to the case of the STM manipulation.

The manipulation of a much smaller, benzene molecule on the Cu(110) surface was considered with several tip models and a full DFT method in [101]. It was found that a clean silicon tip is able to push the adsorbed benzene molecule from one adsorption site to another at close approach. However, a copper terminated tip binds too strongly to the benzene molecule at close approach and removes the molecule from the surface as the tip retracts. This calculation illustrates the importance of the tip model when modelling manipulation processes, and also the importance of the correct description of the tip–molecule interaction.

The smallest and ultimately the most difficult species to manipulate experimentally is an individual adatom or an atomic defect in the surface layer. The mechanism of lateral manipulation of single atom defects with the NC-AFM has been considered in several studies, e.g. [92, 102], however the most complete treatment that also includes the operation of the instrument was applied to Pd atoms adsorbed on the MgO(001) surface and is described in [95]. A single Pd atom is adsorbed above oxygen atoms on the MgO(001) surface and can move into one of the four equivalent nearest neighbour oxygen sites in a single jump (see figure 6). The main challenge here is to be able to control the stochastic diffusion of the Pd atom by the tip, i.e. manipulate it into a desirable direction. The atom can be 'pushed' (i.e. with a repulsive short range interaction) to a specific neighbouring oxygen site by the close approach of the tip (see figure 7). To model a single Pd atom on the MgO(001) surface and its interaction with an AFM tip, a set of classical pair-wise inter-atomic potentials was employed. In these simulations the tip is represented by an MgO cube terminated with a single Mg atom.

Each O atom in the MgO(001) surface (which has the rock-salt structure) is the next nearest neighbour to four other O atoms, and therefore the Pd adatom can move in one of four directions in a single transition by following a minimum energy path to an adjacent O site. The two-dimensional potential energy surface (PES) for the Pd adatom as a function of its lateral position above a single oxygen unit cell on the MgO(001) surface is shown in figure 6. The manipulation of the Pd atom to a neighbouring minimum occurs when the potential energy surface, and hence energy barriers, are modified by the interaction with the AFM tip.

The nature of the interaction of the Mg terminated tip and the Pd adatom is critical to the mechanism of manipulation. In this case the terminating Mg atom has a very weak attractive
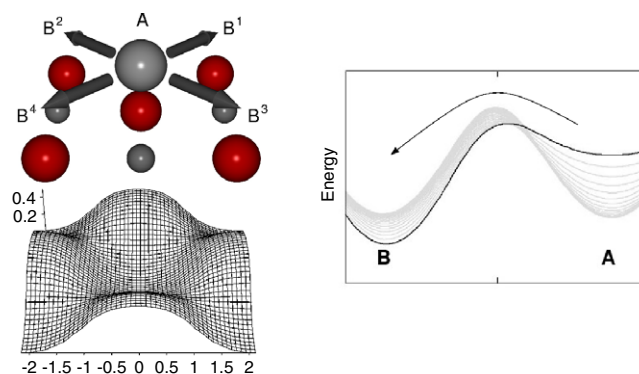


**Figure 7.** The potential energy surface of a Pd atom adsorbed on the MgO(001) surface. Also shown (right) is an illustration of how the barriers are modified by the tip at close approach (for the tip directly above the Pd atom).

interaction with the Pd atom, and is therefore able to 'push' the Pd atom at close approach through Pauli repulsion (see figure 7). The minimum energy path for the Pd atom to move from its initial state into each of the neighbouring states is determined for a given fixed position $(x, y, z)$ of the tip above the surface. This calculation is then performed for a series of fixed positions of the tip above the surface in order to derive the four energy barriers as a function of the tip position. Each of the energy barriers can then be determined for any arbitrary tip position $(x, y, z)$ using interpolation, resulting in four energy barrier 'fields'. With these barrier fields it is then possible to simulate the evolution of the system in real time as the tip oscillates above the surface using the VAFM [96] with the specifically developed KMC algorithm [94] mentioned in section 2.4, which takes account of the change of the energy barriers with time due to tip oscillations. These calculations demonstrate that imaging and controlling stochastic processes using NC-AFM tip can be very complicated.

A detailed study of the NC-AFM imaging that explicitly takes into account thermally induced dynamical processes occurring at the surface in real time was carried out in [103]. Two model defect systems were studied atomistically: a Pd atom adsorbed on MgO(001) which can diffuse across the surface, and a water molecule adsorbed on the CeO$_2$(111) surface, which can rotate about an oxygen atom. Each of these processes is evolved in real time using the KMC method, while simultaneously being imaged using the VAFM, which performs a simulation of the entire instrumentation. The results of these simulations show how dynamical processes can significantly change the contrast seen in NC-AFM images. In particular, mobile defects present on the surface may not be visible in the image and mobile adsorbed molecules can have different images depending on the temperature.

### 3.3. Manipulating electronic structure using the tip of an AFM

Interpretation of image features in NC-AFM images and chemical identification of defects and adsorbed species is still the biggest challenge for the force microscopy. One way of establishing the chemical identity of defect species with confidence could be to correlate the AFM image with
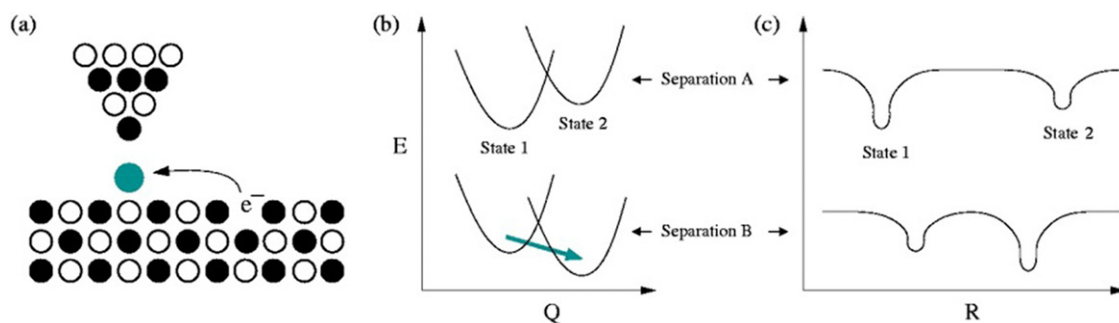
**Figure 8.** (a) Schematic illustrating the manipulation of charge from a neutral oxygen vacancy in the MgO(001) surface to a nearby adsorbed metal adatom. (b) Schematic adiabatic curves representing the initial and the final state of the electron transfer for two different separations between the vacancy and the adatom; the upper set corresponds to long separation and the lower set to the short separation causing the energy shift and the electron transfer. $E$ is a total energy and $Q$ is a configuration coordinate. (c) Schematic of the potential energy of an electron in the vacancy and on the adatom for the long (upper set) and short (lower set) separations between the centres, $R$.

other properties of the imaged species. In particular, the AFM tip could be used to probe and/or modify some defect properties. For example, in studies of bulk defects, one can use different spectroscopies, such as optical spectroscopy and paramagnetic resonance, and one can correlate them via kinetic measurements or magneto-circular dichroism. However, atomic resolution has not yet been achieved combining the analogous SPM methods, such as scanning near-field optical microscopy (SNOM) [104] and tip induced magnetic resonance [105, 106]. A STM tip has been used as a source of both electrons and holes to study injection luminescence from different materials and nanostructures (e.g. [107–112]). Band gap radiation and emission from radiative deep levels were resolved in optical spectra. For example, spatial maps of the resulting integrated light intensity acquired simultaneously with the conventional topographic images in CdS [108] revealed subnanometer scale image contrasts, which has been attributed to defects, such as dislocations or radiative deep levels. However, in insulators only single electron injection into defect states is possible [53, 113] and the electric or magnetic field of a tip is used as a primary source of perturbation to affect the defect properties. As demonstrated above, the electric field of an ionic tip can cause mechanical manipulation of atoms and molecules at surfaces. Interestingly, one can also use the localized perturbation produced by an AFM tip to modify electronic structure locally on the surface or in an adsorbed species or molecule, which could result in the control of their electronic processes and transitions, and help in their chemical identification.

In [114, 115] it has been proposed that one can study individual surface defects by correlating changes in defect optical properties to changes in the electric field produced by the AFM tip. In particular, one can significantly alter the oscillator strength and energies of optical transitions of some surface impurity centres by varying the position of or external voltage applied to an SPM tip. A favourable case could be an excitation of a defect, which, in the absence of the tip, has a very low or zero luminescence yield and is hence invisible. In this case, the perturbation of the tip leads to an increase in the luminescence yield, which can be monitored, the defect 'blinks'. The correlation of topographical and optical SFM signals may provide an additional evidence for a successful

chemical identification of the defect. Further, if the defect nature can be firmly established from other data, one could use spectroscopic properties of such impurities as probes for local electric fields at surfaces.

This extra information offers one route to the identification of certain defect or impurity species using scanning probe microscopy. As a case study, the $Cr^{3+}$ ion in the $Mg^{2+}$ lattice site at the MgO(001) surface was considered. The calculations described in [114, 115] show how the SFM tip can affect the impurity's optical properties. The NC-AFM topographic image was predicted using classical atomistic simulation methods; the effect of the tip on the defect spectroscopic properties was studied using an *ab initio* quantum mechanical embedded cluster method. The electrostatic force due to the applied bias and the polarization of the conducting electrodes because of the image interaction were included self-consistently in the calculation of the system geometry and tip force; the polarization was also accounted for in calculating both the ground and exited potential energy surfaces of the defect, and hence its optical energies and corresponding oscillator strengths. The electronic structure calculations demonstrated that an oxidized tip can significantly affect the oscillator strength and energy of the well-localized Cr ion d–d transitions. These effects can be used to identify a topographic defect image with a specific luminescence signal. The defect spectroscopic properties can depend strongly on the local electric field, significantly altering the branching ratios between radiative and non-radiative transitions. This effect could also be used to study local electric fields at surfaces due to proximity of surface steps or dislocations.

The described modification of surface electronic structure by the AFM tip also suggests that it should be possible to manipulate the location of individual electrons on an insulating surface. In a certain sense this would be achieved in a similar way to a single atom manipulation with the AFM, in that the tip would induce a change in the potential energy surface for a single electron trapped at the crystal surface as opposed to a change in a structural degree of freedom or atomic position [116]. The field produced by the tip at close approach may modify both the relative energies of two distinct electronic states as well as the potential energy barrier separating them, inducing the transfer of a single electron (see figure 8). The

realization of these ideas in an experiment would mean that a single electron transfer process could be directly observed and controlled in an experiment, and would herald a new regime of control at the atomic scale.

Calculations presented in [116] show how it is possible for a pair of defects consisting of a metal adatom and a vacancy on the MgO(001) surface, separated by some distance, to exist in two distinct states: one where each defect is neutral (and hence there are two electrons localized in the vacancy) and one where they are oppositely charged (one electron from the vacant site jumped on the metal adatom). The total energy difference between these two states is dependent on the type of the metal adatom (Pd or Pt) and on the distance between the adatom and the vacancy. The width of the potential barrier for the tunnelling of the electron between the two states will decrease as the metal adatom is moved closer to the vacancy. Therefore the rate of electron tunnelling will be a function of the distance between the two defects. As has been shown in [95], it should be possible to manipulate this type of adatom on this surface with the NC-AFM, which would then allow for control over both the relative energy of the two states and also the rate of electron transfer (through the defect separation, and hence barrier width-see figure 8). Since the transfer of the electron will be clearly visible in the change of the image contrast, it will in principle be possible to determine the transfer rate— through many repeated experiments.

In addition to controlling the relative energies of the final and initial states through the separation of the two defects, the results presented in [116] indicate that the ionization potential of the oxygen vacancy and the electron affinity of the metal adatom can also be modified by the electric field produced by an AFM tip apex at close approach. Therefore there are two possibilities to induce a single electron transfer on the surface: to manipulate the position of a defect (in this case a metal adatom) to vary the total energy difference between the initial and final states and to use the electric field produced by the tip to temporarily vary the total energies and potential energy barrier to induce an electronic transition due to a close approach of the tip to the surface.

These are model calculations on an ideal system and this type of manipulation on a real system has yet to be realized experimentally. However, they do suggest ways in which local electronic structure and the localization of individual electrons can be controlled in an atomic system. Achieving this level of control for realistic systems will be required to realize single molecule electronic devices in practice.

## 4. Perspectives

We have summarized above some of the current state-of-the-art in the theory and modelling of organic molecules on insulating surfaces, focusing on the methods that are required to model components of single molecule devices and help in their design. In particular, we surveyed the modelling of adsorption and diffusion of organic molecules on insulating surfaces, their interaction with surface defects and how surface adsorbates can be imaged and controlled with the NC-AFM tip. However, there is still a lot of scope for both further

methods developments and applications of simulations to realistic systems. Many of the results presented here concern highly idealized and model systems, and while these can be very useful to test new ideas and concepts, more quantitative and realistic simulations will be needed in order to optimize the construction and functioning of real molecular devices. Below we outline some of the prospects for such calculations.

The development of *ab initio* methods, and specifically DFT, over the last decade combined with a huge increase in computational power made it possible to study the adsorption of large molecules on surfaces from the first principles. However, this type of calculations still remains very computationally expensive and for realistic systems only a limited number of configurations can be explored and dynamical simulations over long periods of time are not possible. However, as computational power increases and algorithms improve, there will certainly be more scope for the use of *ab initio* methods to treat such large systems even using sophisticated quantum-chemistry methods [117]. The recent developments in linear scaling DFT methods [118] hold promise for the treatment of very large systems as computational power continues to increase. Other approaches that are particularly applicable to the type of systems considered here are QM/MM and embedding techniques [119, 120].

An alternative way of dealing with composite systems, such as a large molecule on a surface, is to replace a full *ab initio* calculation by a set of smaller calculations of overlapping subsystems treated at the same level of accuracy or by a set of simpler calculations. These methods have been recently reviewed and developed further in [121]. The gain in efficiency with respect to performing a full calculation on the entire system using one of the existing quantum chemical or DFT methods is achieved due to a nonlinear scaling of these methods. This makes a calculation of the whole system more efficient. However, the accuracy of the calculated total energy with respect to the full calculation depends on the system, on the availability and accuracy of the required full-atom pseudopotentials and on the scheme used for partitioning the full system into smaller fragments [121].

The method developed in [121] has also a promise of reducing, or even completely eliminating, the problem inherent to any QM/MM technique related to link atoms needed to terminate the quantum region. The sophistication of classical force-fields is also improving, such as in the use of multipoles on atomic sites and charge transfer potentials to describe more accurately molecules and their interactions [122]. The fragmentation method [121] also has a potential of becoming a rigorous platform for constructing molecule–surface force-fields as it allows, in a controllable and rigorous way, to split a large molecule into smaller fragments and produce the force-field describing the interaction of each fragment with the surface individually. The application of these methods to molecular adsorption on surfaces, calibrated from more extensive and sophisticated quantum chemical calculations would enable more convincing use of classical models.

One important problem common to all the *ab initio* DFT methods based on standard density functionals is the

lack of non-local correlation effects. This is manifested itself in the significant underestimation of the van der Waals attraction between the molecule and surface. The van der Waals interaction is expected to play a significant role in the binding of a molecule to a surface for the types of systems discussed here, where the interaction with the surface is usually ionic in nature. This may not necessarily be significant for the mechanisms and rates of diffusion on a surface; however it needs to be accounted for to provide quantitative predictions for the functioning of real devices. Typically in quantum chemical calculations, the correlation can be accurately included using post-Hartree–Fock methods such as Møller–Plesset perturbation theory [123] or the coupled cluster approach. Although these methods are extremely computationally expensive, and are nowadays directly only feasible for small molecules and complexes, recent advances in the method of increments [124] and in parallel computing suggest that high-quality methods may soon be routinely used for larger systems as well, including molecules on crystal surfaces. Another interesting recent progress is related to the development of the first principles density functional (vdW-DF), which incorporates the non-local part of the correlation energy responsible for the van der Waals interaction [125, 126]. This functional enables fully self-consistent calculation of systems energies and atomic forces and is already showing a promise of producing correct binding energies and geometries for systems where standard DFT functionals notoriously failed [127]. Also, several semi-empirical methods have been developed, that can calculate the van der Waals interaction between molecules and clusters atomistically and relatively cheaply; an appropriate contribution can then be simply added to the forces calculated from a DFT calculation. Such methods include that of dynamic polarizabilities [128] and dispersion corrected DFT [129], which are beginning to be widely applied to the study of inter-molecular interactions. These methods, especially for large systems, will likely to become a promising alternative to the first principles methods for the incorporation of the van der Waals forces into calculations of molecule–surface interactions, due to their low computational cost and a possibility to achieve a reasonable accuracy.

The methods discussed above are simply employed to determine the interactions between atoms in the molecule and surface, however determining the adsorbed structure of the molecule (and hence the adsorption energy) requires exploration of the high-dimensional potential energy surface (PES) of the system. As discussed above, this typically involves 'guessing' likely adsorption configurations and performing a gradient based optimization to locate a local minimum on the PES. For the case of small adsorbed molecules and atoms, the number of possible adsorption sites and configurations is limited and it is often possible to perform an exhaustive search, however in the cases of larger molecules and/or surfaces with complicated structure this is not possible due to the size and complexity of the system. In addition to different adsorption sites, in the cases of flexible molecules possessing collective internal degrees of freedom (such as a bending angle, for instance), there are also often many possible conformations of the molecular structure on the surface.

The most common method for exploring these configurations and determining the accessible global minimum energy of a system is the simulated annealing, however, finding the global minimum using this method in general may take a very long computational time, and hence may not be efficient in many types of systems. Especially when the evaluation of the PES is expensive (as is the case for *ab initio* methods), there is a need for faster converging algorithms. Some methods that have shown to be particularly useful when applied to complex molecular systems include genetic algorithms [82], various basin hopping algorithms [83, 130] and metadynamics [87], and are often able to find solutions more quickly than simulated annealing [131].

Once the possible adsorbed configurations have been determined, it is then essential for the design process of a nanodevice to be able to determine the residence time of the molecule in that configuration and how the molecule will diffuse, i.e. the mobility of the molecule on the surface need to be considered. As previously mentioned, this can be achieved directly using the MD method to explicitly follow the trajectory of the atomic coordinates, however this method is only computationally feasible over timescales of nanoseconds (picoseconds if an *ab initio* method is being employed directly), whereas many molecular transitions may occur over much longer timescales. However in many types of complex systems, the configurations that will be accessed during the evolution may not be intuitively obvious and may be too numerous to determine. In this case other novel methods can be used to explore the configuration space more efficiently. In temperature accelerated molecular dynamics a higher temperature is used during the MD run, which enables to increase the rates of all processes [132], run longer time simulations and explore the PES faster. To recover the dynamics of the system at the lower temperature, the residence time in each state is extrapolated from the energy barriers calculated for each detected transition. This method has an advantage of finding possible transitions automatically, and also that attempt pre-factors can be determined from the residence time in each state at the high temperature. This method had been applied successfully to the diffusion of atoms and clusters on surfaces, however it has yet to be applied to a molecular system with 'soft' degrees of freedom and a hierarchy of different timescales. One possible problem is that low barriers associated with such 'soft' degrees of freedom would be crossed too rapidly making the algorithm very inefficient due to the computational expense involved in evaluating these transitions, however possible solutions to this problem have been suggested [133], which could make this method suitable for looking at the long timescale simulation of molecular diffusion.

Once the behaviour of individual molecules on a surface has been determined, the interaction of the system with an AFM tip must be investigated in order to simulate the process of imaging and manipulation. This adds additional complexity to the system and increases the expense of the calculations significantly. In this paper we have summarized some of the more recent advances in the modelling of AFM imaging, which includes the diffusion and manipulation of a

single adatom as well as imaging diffusion and manipulation of a simple molecule in a single configuration. These in themselves are significant efforts, as in the first case the entire PES as a function of the tip position needs to be determined and combined with a real time simulation of the experimental apparatus, while in the second the force-field for the entire adsorbed molecule interacting with the tip is required with a realistic tip model. To fully understand the mechanisms and dynamics of the manipulation of a large complex molecule (using the VAFM), the accessible potential energy surface of the system (like that considered in a KMC simulation) would need to be determined for every tip position, and then the real time dynamics during imaging investigated. This would be a truly colossal task for most realistic systems. However, investigations using model systems with limited degrees of freedom may be feasible and could shed light on generic mechanisms.

Finally, simulations of the contrast formation in images of static molecules are still important for understanding better the interaction of the tip with the molecular structure and for improving our understanding of the information obtained from experiments. One major problem in imaging of adsorbed molecules on insulators with the AFM is that it is rare to achieve simultaneous contrast of both the adsorbed molecules and the atomic structure of the surface, which prevents the determination of adsorption sites. This is due to the different interaction of the tip with surface atom and the molecule, and the different imaging parameters the two require. One solution to this is to use 'hybrid' imaging modes (i.e. a mixture of constant height and topography) to capture both structures simultaneously. To model this, an explicit simulation of the instrument is required, as implemented in the VAFM. This modelling could be extremely useful in informing experimental efforts to image adsorbed molecules. The other problem is related to the unknown structure of the tip apex which forces to perform imaging and manipulation simulations with several tip models. More effort is needed for understanding the relationships between the tip preparation techniques used in real experiments and possible tip apex structures. Understanding of this link will eventually improve the predictability of theory.

In summary, the use of modelling in the design and implementation of single molecule devices on insulating surfaces is essential. The work undertaken so far has provided many insights into the processes that occur during adsorption of a large molecule on a surface—both electronic and structural—and has indicated ways in which these can be controlled. Modelling of simplified, but realistic, systems has suggested design rules for the molecular structure to optimize the mobility of a molecule for a specific function. Modelling of AFM imaging and manipulation has provided insights into imaging mechanisms and mechanisms of manipulation and has been able to suggest optimum protocols for controlled atomic scale manipulation. In order for this success to be built upon, this modelling must be extended to more realistic and demanding systems in order to provide quantitative guidance for experimental efforts to engineer these systems on the atomic scale.

## References

[1] Aviram A and Ratner M 1974 *Chem. Phys. Lett.* **29** 277
[2] Joachim C, Gimzewski J, Schlittler R R and Chavy C 1995 *Phys. Rev. Lett.* **74** 2102
[3] Joachim C, Gimzewski J and Aviram A 2000 *Nature* **408** 541
[4] Joachim C 2002 *Nanotechnology* **13** R1
[5] Heath J R and Ratner M A 2003 *Phys. Today* **56** 43
[6] Eigler D M and Schweizer E K 1990 *Nature* **344** 524
[7] Crommie M F, Lutz C P and Eigler D M 1993 *Science* **262** 218
[8] Heinrich A J, Lutz C P, Gupta J A and Eigler D M 2002 *Science* **298** 1381
[9] Sugimoto Y, Abe M, Hirayama S, Oyabu N, Custance O and Morita S 2005 *Nat. Mater.* **4** 156
[10] Ternes M, Lutz C P, Hirjibehedin C F, Giessibl F J and Heinrich A J 2008 *Science* **319** 1066
[11] Gourdon A 1998 *Eur. J. Org. Chem.* **12** 2797
[12] Nilius N, Wallis T M and Ho W 2002 *Science* **297** 1853
[13] Miwa J A, Weigelt S, Gersen H, Besenbacher F, Rosei F and Linderoth T R 2006 *J. Am. Chem. Soc.* **128** 3164
[14] Weigelt S, Busse C, Petersen L, Rauls E, Hammer B, Gothelf K V, Besenbacher F and Linderoth T R 2006 *Nat. Mater.* **5** 112
[15] Alemani M, Peters M V, Hecht S, Rieder K H, Moresco F and Grill L 2006 *J. Am. Chem. Soc.* **128** 14446
[16] Iancu V and Hla S W 2006 *Proc. Natl Acad. Sci. USA* **103** 13718
[17] Perepichka D F, Bryce M R, Pearson C, Petty M C, Mcinnes E J L and Zhao J P 2003 *Angew. Chem. Int. Edn* **42** 4635
[18] Perepichka D F and Bryce M R 2005 *Angew. Chem. Int. Edn* **44** 5370
[19] Bendikov M, Wudl F and Perepichka D F 2004 *Chem. Rev.* **104** 4891
[20] Oleynik I I, Kozhushner M A, Posvyanskii V S and Yu L 2006 *Phys. Rev. Lett.* **96** 096803
[21] Barth J V, Weckesser J, Cai C, Günter P, Bürgi L, Jeandupeux O and Kern K 2000 *Angew. Chem. Int. Edn Engl.* **39** 1230
[22] Hecht S 2003 *Angew. Chem. Int. Edn Engl.* **42** 24
[23] Theobald J A, Oxtoby N S, Phillips M A, Champness N R and Beton P H 2003 *Nature* **424** 1029
[24] Grill L, Rieder K H, Moresco F, Rapenne G, Stojkovic S, Bouju X and Joachim C 2007 *Nat. Nanotechnol.* **2** 95
[25] Katano S, Kim Y, Hori M, Trenary M and Kawai M 2007 *Science* **316** 1883
[26] Lastapis M, Martin M, Riedel D, Hellner L, Comtet G and Dujardin G 2005 *Science* **308** 1000
[27] Sloan P A, Hedouin M F G, Palmer R E and Persson M 2003 *Phys. Rev. Lett.* **91** 118301
[28] Lopinski G P, Wayner D D M and Wolkow R A 2000 *Nature* **406** 48
[29] Repp J, Meyer G, Stojkovic S, Gourdon A and Joachim C 2005 *Phys. Rev. Lett.* **94** 026803
[30] Villagomez C J, Zambelli T, Gauthier S, Gourdon A, Barthes S S C and Joachim C 2007 *Chem. Phys. Lett.* **450** 107

[31] Bellec A, Ample F, Riedel D, Dujardin G and Joachim C 2009 *Nano Lett.* **9** 144

[32] Pacchioni G, Giordano L and Baistrocchi M 2005 *Phys. Rev. Lett.* **94** 226104

[33] Yoon B and Landman U 2008 *Phys. Rev. Lett.* **100** 056102

[34] Morita S, Wiessendanger R and Meyer E (ed) 2002 *Non-Contact Atomic Force Microscopy* (*Nanoscience and Technology Series*) (Berlin: Springer) p 450

[35] Giessibl F 2003 *Rev. Mod. Phys.* **75** 949

[36] García R and Pérez R 2002 *Surf. Sci. Rep.* **47** 197

[37] Schlettwein D, Back A, Schilling B, Fritz T and Armstrong N R 1998 *Chem. Mater.* **10** 601

[38] Nony L, Bennewitz R, Pfeiffer O, Gnecco E, Baratoff A, Meyer E, Eguchi T, Gourdon A and Joachim C 2004 *Nanotechnology* **15** S91

[39] Proehl H, Nitsche R, Dienel T, Leo K and Fritz T 2005 *Phys. Rev. B* **71** 165207

[40] Proehl H, Dienel T, Nitsche R and Fritz T 2004 *Phys. Rev. Lett.* **93** 097403

[41] Kunstmann T, Schlarb A, Fendrich M, Wagner Th, Möller R and Hoffmann R 2005 *Phys. Rev. B* **71** 121403(R)

[42] Dienel T, Loppacher C, Mannsfeld S C B, Forker R and Fritz T 2008 *Adv. Mater.* **20** 959

[43] Schütte J, Bechstein R, Rahe P, Rohlfing M, Kühnle A and Langhals H 2009 *Phys. Rev. B* **79** 045428

[44] Murray P W, Gimzewski J K, Schlittler R R and Thornton G 1996 *Surf. Sci.* **367** L79

[45] Burke S A, Mativetsky J M, Hoffmann R and Grütter P 2005 *Phys. Rev. Lett.* **94** 096102

[46] Burke S A, Mativetsky J M, Fostner S and Grütter P 2007 *Phys. Rev. B* **76** 035419

[47] Fukui K and Sakai M 2006 *J. Phys. Chem. B* **110** 21118

[48] Loske F, Bechstein R, Schütte J, Ostendorf F, Reichling M and Kühnle A 2009 *Nanotechnology* **20** 065606

[49] Joachim C and Ratner M A 2005 *Proc. Natl Acad. Sci. USA* **102** 8801

[50] Repp J, Meyer G, Paavilainen S, Olsson F E and Persson M 2006 *Science* **312** 1196

[51] Trevethan T *et al* 2007 *Phys. Rev. Lett.* **98** 028101

[52] Trevethan T and Shluger A L 2007 *J. Phys. Chem. C* **111** 15375

[53] Bussmann E and Williams C C 2006 *Appl. Phys. Lett.* **88** 263108

[54] Hirth S, Ostendorf F and Reichling M 2006 *Nanotechnology* **17** S148

[55] Enevoldsen G H, Foster A S, Christensen M C, Lauritsen J V and Besenbacher F 2007 *Phys. Rev. B* **76** 205415

[56] Hofer W, Foster A S and Shluger A L 2003 *Rev. Mod. Phys.* **75** 1287

[57] Lanzani G *et al* 2007 *Theor. Chem. Acc.* **117** 805

[58] Groß A 2008 *J. Comput. Theor. Nanosci.* **5** 894

[59] Repp J, Meyer G, Paavilainen S, Olsson F E and Persson M 2006 *Science* **312** 1196

[60] Keeling D L, Humphry M J, Fawcett R H J, Beton P H, Hobbs C and Kantorovich L 2005 *Phys. Rev. Lett.* **94** 146104

[61] Martsinovich N, Hobbs C, Kantorovich L, Fawcett R H J, Humphry M J, Keeling D L and Beton P H 2006 *Phys. Rev. B* **74** 085304

[62] Lundqvist M J, Nilsing M, Lunell S, Åkermark B and Persson P 2006 *J. Phys. Chem. B* **110** 20513

[63] Mackerell A D *et al* 1998 *J. Phys. Chem. B* **102** 3586

[64] Halgren T A 1996 *J. Comput. Chem.* **17** 553

[65] Catlow C R A, Bell R G and Gale J D 1994 *J. Mater. Chem.* **4** 781

[66] Bush T S *et al* 1994 *J. Mater. Chem.* **4** 831

[67] Lin H and Truhlar D 2007 *Theor. Chem. Acc.* **117** 185

[68] Bo C and Maseras F 2008 *Dalton Trans.* **22** 2911

[69] Sushko M L, Gal A Y and Shluger A L 2006 *J. Phys. Chem. B* **110** 4853

[70] Bandura A V and Kubicki J D 2003 *J. Chem. Phys. B* **107** 11072

[71] Cornell W D, Cieplak P, Balyly C I, Gould I R, Merz K, Ferguson D M, Spellmeyer D C, Fox T, Caldwell J W and Kollman P A 1995 *J. Am. Chem. Soc.* **117** 5179

[72] Watkins M, Trevethan T, Sushko M L and Shluger A L 2008 *J. Phys. Chem. C* **112** 4226

[73] Hobbs C and Kantorovich L 2006 *Surf. Sci.* **600** 551

[74] Hobbs C, Kantorovich L and Gale J D 2005 *Surf. Sci.* **591** 45–55

[75] Martsinovich N and Kantorovich L 2008 *Phys. Rev. B* **77** 115429

[76] Moriarty P *et al* 1998 *Phys. Rev. B* **57** 362

[77] Godwin P D, Kenny S D and Smith R 2003 *Surf. Sci.* **529** 237

[78] Kim Y *et al* 1997 *Surf. Sci.* **385** L945–51

[79] Jung S C and Kang M H 2008 *Surf. Sci.* **602** 1916

[80] Pannetier J 1990 *Inst. Phys. Conf. Ser.* **107** 23

[81] Hartke B and Carter E A 1993 *Chem. Phys. Lett.* **216** 324

[82] Clark D E and Westhead D R 1996 *J. Comput.-Aided Mol. Des.* **10** 337

[83] Goedecker S 2004 *J. Chem. Phys.* **120** 9911

[84] Allen M P 2002 *Computer Simulation of Liquids* (Oxford: Oxford University Press)

[85] Laidler K and King C 1983 *J. Phys. Chem.* **87** 2657

[86] Weidenmuller H A and Zhang J S 1984 *J. Stat. Phys.* **34** 191

[87] Laio A and Parrinello M 2002 *Proc. Natl Acad. Sci. USA* **99** 12562

[88] Iannuzzi M, Laio A and Parrinello M 2003 *Phys. Rev. Lett.* **90** 238302

[89] Ensing B *et al* 2006 *Acc. Chem. Res.* **39** 73

[90] Voter A F 2007 *Radiation Effects in Solids* (Berlin: Springer)

[91] Ghasemi S A, Goedecker S, Baratoff A, Lenosky T, Meyer E and Hug H J 2008 *Phys. Rev. Lett.* **100** 236106

[92] Sugimoto Y, Pou P, Abe M, Jelinek P, Pérez R, Morita S and Custance O 2007 *Nature* **446** 64

[93] Foster A S, Barth C, Shluger A L, Nieminen R M and Reichling M 2002 *Phys. Rev. B* **66** 235417

[94] Kantorovich L 2007 *Phys. Rev. B* **75** 064305

[95] Trevethan T *et al* 2007 *Phys. Rev. B* **76** 085414

[96] Polesel-Maris J and Gauthier S 2005 *J. Appl. Phys.* **97** 044902

[97] Trevethan T and Shluger A L 2008 *J. Phys. Chem. C* **112** 19577

[98] Sushko M L, Gal A Y, Watkins M and Shluger A L 2006 *Nanotechnology* **17** 2062

[99] Martsinovich N and Kantorovich L 2008 *Nanotechnology* **19** 235702

[100] Hobbs C and Kantorovich L 2006 *Surf. Sci.* **600** 551

[101] Atodiresei N, Caciuc V and Holscher H 2008 *Phys. Rev. B* **77** 153408

[102] Dieška P, Štich I and Pérez R 2005 *Phys. Rev. Lett.* **95** 126103

[103] Watkins M, Trevethan T, Shluger A L and Kantorovich L N 2007 *Phys. Rev. B* **76** 245421

[104] Ohtsu M (ed) 2005 *Progress in Nano-Electro-Optics IV: Characterization of Nano-Optical Materials and Optical Near-Field interactions* (Berlin: Springer) p 206

[105] Rugar D, Budakian R, Mamin H J and Chui B W 2004 *Nature* **430** 329

[106] Degen C L, Poggio M, Mamin H J, Rettner C T and Rugar D 2009 *Proc. Natl Acad. Sci.* **106** 1313 doi:10.1073/pnas.0812068106 Early edition

[107] Gimzewski J K *et al* 1988 *Z. Phys. B* **72** 497

[108] Berndt R and Gimzewski J K 1992 *Phys. Rev. B* **45** 14095

[109] Berndt R, Gaisch R, Schneider W D, Gimzewski J K, Reihl B, Schlittler R R and Tschudy M 1994 *Surf. Sci.* **307–309** 1033

[110] Uchioda S 1997 *Appl. Surf. Sci.* **113/114** 335
[111] Comtet G, Dujardin G, Mayne A J and Riedel D 2006
*J. Phys.: Condens. Matter* **18** S1927
[112] Mayne A J, Dujardin G, Comtet G and Riedel D 2006 *Chem. Rev.* **106** 4355
[113] Bussmann E B, Zheng N and Williams C C 2006 *Nano Lett.* **6** 2577
[114] Kantorovich L N, Shluger A L and Stoneham A M 2000 *Phys. Rev. Lett.* **85** 3846
[115] Kantorovich L N, Shluger A L and Stoneham A M 2001 *Phys. Rev. B* **63** 184111
[116] Trevethan T and Shluger A L 2009 *Nanotechnology* **20** 264019
[117] Shao Y *et al* 2006 *Phys. Chem. Chem. Phys.* **8** 3172
[118] Li Z Y, He W and Yang J L 2005 *Prog. Chem.* **17** 192
[119] Altoè P P, Stenta M, Bottoni A and Garavelli M 2007 *Theor. Chem. Acc.* **118** 219
[120] Harding J H, Duffy D M, Sushko M L, Rodger P M, Quigley D and Elliott J A 2008 *Chem. Rev.* **108** 4823

[121] He J, di Paolo C and Kantorovich L N 2009 *J. Chem. Phys.* **130** 144104
[122] Stone A J and Misquitta A J 2007 *Int. Rev. Phys. Chem.* **26** 193
[123] Møller C and Plesset M S 1934 *Phys. Rev.* **46** 618
[124] Paulus B 2006 *Phys. Rep.* **428** 1
[125] Gulans A, Puska M J and Nieminen R M 2009 under preparation
[126] Dion M *et al* 2004 *Phys. Rev. Lett.* **92** 246401
[127] Langreth D C *et al* 2009 *J. Phys.: Condens. Matter* **21** 084203
[128] Misquitta A J and Stone A J 2008 *Mol. Phys.* **106** 1631
[129] Lin I C and Rothlisberger U 2008 *Phys. Chem. Chem. Phys.* **10** 2730
[130] Wales D 2003 *Energy Landscapes* (Cambridge: Cambridge University Press)
[131] Ali M M, Smith R and Hobday S 2006 *Comp. Phys. Commun.* **175** 451
[132] Sorensen M R and Voter A F 2000 *J. Chem. Phys.* **112** 9599
[133] Montalenti F and Voter A F 2002 *J. Chem. Phys.* **116** 4819