

Part One

Introduction to ALD

1

Theoretical Modeling of ALD Processes

Charles B. Musgrave

1.1

Introduction

This chapter describes simulations of atomic layer deposition (ALD) using quantum chemical electronic structure methods. Section 1.2 provides a brief overview of the quantum chemistry methods useful in the study of chemical processes and materials behavior. Although this section includes a summary of quantum chemical methods, it is not meant to be a comprehensive review of quantum chemistry and the reader is encouraged to examine the many excellent textbooks and reviews of quantum chemistry [1–3]. Section 1.3 overviews the use of quantum chemistry for predicting the properties of molecules and materials, while Section 1.4 specifically overviews the use of these methods to study ALD mechanisms, and Section 1.5 provides several examples of the determination of ALD mechanisms using quantum chemical methods. Again, this section is not meant to provide a comprehensive review of the use of electronic structure theory to study ALD and the motivated reader can examine the literature to explore what has specifically been done in this area. In addition to the various manuscripts our group has published on this topic [4–24], the groups of Esteve and Rouhani [25–29] and Raghavachari [30–35] have published many excellent articles on using quantum chemical methods to study ALD. While the overview of simulations presented here is meant to be helpful in understanding the approaches used to theoretically study the chemistry of ALD, it can be skipped for those who are already familiar with these methods or who do not seek a deeper understanding of electronic structure theory.

1.2

Overview of Atomistic Simulations

While simulations can be aimed at developing a description and understanding of the ALD process or at predicting the properties of the resulting ALD film, we here specifically focus on simulations intended to explain the chemistry of the ALD process. The unique features of ALD specifically rely on the nature of the half-

reactions that describe the chemistry of each half-cycle. Specifically, ALD relies on the self-limiting nature of the surface reactions that result from the use of reagents that each do not self-react and that are introduced into the ALD reactor in separate pulses, temporally separated from each other by a reactor purge [36]. During each precursor pulse, the precursor reacts selectively with the functional groups remaining from the previous precursor pulse, but not with itself. For a successful ALD process, each half-reaction must produce a surface functional group reactive toward the subsequent precursor and deposit at most one atomic layer. Ideally, because reactants do not self-react and are introduced into the reactor separately, chemistry and transport are decoupled in ALD reactors. Thus, the nature of the surface reactions is the central feature of the ALD process that provides it with its ideal attributes – uniformity, conformality, and nanometer-scale control of film thickness and composition. Consequently, a fundamental understanding of an ALD process involves a detailed description of the ALD chemical mechanism, and accordingly, we focus on summarizing methods that describe these surface reactions that provide ALD with its unique advantages.

1.2.1

Quantum Simulations

The most fundamental approach to describing chemical reactions is based on using quantum mechanics to describe the electronic structure of the reacting system. During a chemical reaction, the reactants transform into products by rearranging their atomic coordinates from that of the reactants to that of the products. The actual trajectory followed by the reacting species is variable as atomic motion is a highly dynamical process. Each intermediate structure along the trajectory followed during the reaction involves a unique ground-state electronic wave function. That is, the electron density of the system is dynamically redistributed as the atomic structure evolves from that of the reactants to that of the products. Because the electrons have much lower mass than the nuclei, the reacting system typically stays in its electronic ground state during the course of reaction. In other words, the redistribution of electron density along the trajectory of atomic configurations is adiabatic. Nonetheless, the key point is that this complex process is intrinsically quantum mechanical, despite being electronically adiabatic because the properties along the trajectory of the reaction, and specifically the energy, depend exclusively on the wave function at each atomic configuration.

Because chemical reactions are inherently quantum mechanical, any method used to simulate chemical reactions must either be quantum mechanical or somehow empirically incorporate the quantum mechanical nature of the process of reaction into its description. Classical molecular dynamics (MD) potentials are generally not capable of accurately describing chemical reactions, except in the very few cases where potentials have been designed and trained specifically to describe a particular chemical reactivity. Fortunately, over the past few decades the meteoric rise in computational power together with significant advances in electronic structure methods and algorithmic progress has made the direct application of quantum

mechanics to describing atomistic systems of reasonable size practical, although nowhere as efficient as classical MD. In fact, high-quality quantum calculations that required supercomputers to execute just 20 years ago can now be performed on desktop computers. Another major development that enabled the ongoing revolution in quantum chemistry is the development of high-quality quantum mechanical density functional theory (DFT) methods that have been implemented in reasonably priced simulation packages with relatively simple graphical interfaces. This, combined with the availability of affordable modern computers, has enabled researchers on even the most modest budgets access to the power of quantum chemical calculations. The result is an explosion in the quantity of research conducted using quantum chemical simulations. Thus, the computational approaches we focus on are quantum chemical methods, both because of their now widespread use and because they are the most appropriate and reliable approaches for describing the chemistry of ALD.

A large number of different quantum chemical methods, including DFT, are now available in a variety of software packages. These can be categorized as being either semi-empirical or *ab initio* (from first principles). Semi-empirical methods approximately solve the Schrödinger equation, but use empirically derived parameters to compute the effects of terms that are either neglected or approximated in the specific approximations used within the method. These methods have the advantage of being extremely fast relative to *ab initio* methods. They can also be relatively accurate if the system and property of interest are within or not too different from the training set used to parameterize the method. Unfortunately, semi-empirical methods are generally not reliable for the prediction of activation barriers, which of course are central to predicting the chemistry of a reacting ALD system as the energies of the various possible transition states (TS) determine whether a reaction is active. Consequently, we do not discuss the application of semi-empirical methods to simulate ALD.

In contrast to semi-empirical methods, first principles methods are by definition not empirically fit to any experimental data set. *Ab initio* methods start with a fundamental quantum mechanical description of the system and then employ various approximations to make the solution of the quantum mechanical problem tractable. While approximations are employed, they do not involve fitting to any data. As we will see shortly, many DFT methods involve an empirical component in the development of the exchange correlation (XC) functionals that define each DFT method. Nevertheless, DFT methods are still typically called *ab initio* methods. Despite the fact that most DFT methods do not strictly conform to the definition of *ab initio*, we will also adhere to the general convention of calling these methods first principles methods.

1.2.2

Wave Function-Based Quantum Simulations

Quantum simulation methods are classified as either wave function or density functional methods. *Ab initio* wave function methods involve directly solving the Schrödinger equation using various approximations with the quality of the method depending on the degree of approximation. Thus, *ab initio* wave function methods can be ranked within a hierarchy that depends on the extent of the approximations used,

ranging from the mean field approximation of the Hartree–Fock (HF) method [37–39] to exact methods that include n th-order perturbation theory [40] (in the limit of large n and if the perturbation is small), full configuration interaction (CI), and quantum Monte Carlo methods [41]. The most approximate *ab initio* wave function method is the HF method. The HF wave function is the lowest energy single Slater determinant wave function for an n -electron system, where the wave function is the *determinant* of the n -by- n *matrix* where each column is a different electron orbital and each row a different electron. A determinant form for the wave function guarantees that it is antisymmetric, a requirement for fermions, and that the electrons are treated as identical, indistinguishable particles. Unfortunately, despite being *ab initio*, the mean field approximation that defines the HF method causes it to severely overestimate activation barriers in the vast majority of reactions, among other inaccuracies, which thus makes it unsuitable as a method for accurately describing chemical reactivity. On the other hand, HF serves as the basis for almost all wave function-based methods. Generally, these approaches combine the n -electron HF ground-state wave function with excited HF wave functions constructed by replacing occupied orbitals of the ground state with up to n unoccupied (virtual) orbitals to improve the quality of the wave function. The mixing in of excited HF wave functions with the HF ground state to improve the description of the wave function is usually accomplished using perturbation theory or variationally using configuration interaction.

Because a detailed description of wave function methods is beyond the scope of this chapter and because the application of these methods to simulate ALD has been much more limited than the application of density functional methods, we will forgo a more extensive account of these methods. However, we note that *ab initio* wave function methods have the advantage that they can be systematically improved. That is, we can methodically climb the hierarchical ladder of methods from the base Hartree–Fock method up toward the exact methods by, for example, increasing the order of the many-body perturbation theory (MBPT) or increasing the order of excitations in a configuration interaction method, which include coupled cluster and complete active space methods. Thus, we can robustly determine whether a method accurately describes the system by systematically improving it and analyzing its approach to convergence. Unfortunately, the accurate wave function methods are generally prohibitively expensive for many systems of interest. This is especially true for the case of ALD because the systems of interest are surface reactions, and thus involve a model of the reacting surface containing too many atoms for practical simulation using the higher quality wave function methods. However, they can still play the role of calibrating more approximate methods practical enough for simulating ALD processes to validate the choice of method and provide error bars on the predicted reaction energetics.

1.2.3

Density Functional-Based Quantum Simulations

An alternative to using the electronic wave function as the basis of a quantum mechanical method is to use the electron density. Methods based on this approach are called DFT methods. Hohenberg and Kohn first showed that the electron density

uniquely determines the number of electrons and the location and identity of the atoms, that is, the potential, and thus the system Hamiltonian and its associated Schrödinger equation [42]. Consequently, the electron density uniquely determines the wave function and thus all information contained in the wave function must also be contained in the density! The key is to construct an energy functional equivalent to the Hamiltonian that includes all important contributions to the electronic energy. A functional is essentially a function of a function; in this case, the energy is a function of the electron density, which is itself a function of position. Kohn and Sham proved that the exact energy functional exists [43], but unfortunately proof of its existence did not prescribe an approach to its determination. However, knowing that an exact energy functional exists motivated the development of DFT where the key challenge has been to derive energy functionals that accurately describe the system.

The general approach to developing DFT energy functionals has been to exploit the fact that energy is an extensive property in order to separate the energy functional into its component energy contributions, including the kinetic energy, electron–electron repulsions, electron–nuclear attractions, the exchange energy, and the correlation energy. The explicit inclusion of the exchange and correlation energy through the exchange and correlation density functionals, called the exchange correlation functional when combined, may appear strange here because they are not explicitly included in the Hamiltonian of wave function methods. This is because no anti-symmetry requirement is imposed on the density, so exchange does not naturally arise in DFT as it does when using a Slater determinant form of the wave function and so it must be explicitly added. On the other hand, correlation is an opportunistic addition to the energy functional used to improve the quality of a density functional to help it reproduce the properties of the system. This explicit addition of correlation energy is remarkably computationally efficient compared to the approaches for incorporating additional correlation energy into wave function methods that rely on mixing large numbers of excited states into the ground-state wave function. Furthermore, density functional methods are computationally efficient because they not only use a single Slater determinant description of the electronic structure, but also do not involve calculating any “two-electron integrals,” which is a requirement of all orbital-based wave function methods, making the wave function methods scale with at least the fourth power of the system size.

One drawback of DFT methods is that they are generally not improved systematically; whereas wave function methods calculate exchange and correlation based on expansions, the exchange and correlation density functionals that define the DFT method are based on specific functions. Another drawback is that these functionals often involve parameters that are empirically determined by a fitting process. Although virtually all DFT methods involve XC functionals that are empirically fit and are thus not strictly first principles methods, they are commonly referred to as “*ab initio*.” Because DFT functionals are often empirically determined, great care should be taken in conducting simulations using DFT to ensure that the functional is appropriate for the property and system of interest. Unfortunately, it is common for DFT results to be published in the literature that do not confirm that the choice of DFT method is appropriate for the property of interest and the reader is left with the

question of whether the results are reliable. This issue is often not fully appreciated. This might be partially caused by the confidence in the method instilled in the nonexpert by the use of the term “*ab initio*” or “first principles” to describe DFT methods. A similar caution could be given for wave function methods where the level of approximation within the method may make it incapable of accurately describing the system or phenomena of interest. Again, the user should confirm that the method is appropriate for the property being predicted.

1.2.4

Finite and Extended Quantum Simulations

Computational quantum mechanical methods can also be categorized by whether the method is applied in a finite or extended fashion. A finite calculation involves simulation of a localized system, such as an atom, molecule, or cluster of atoms, whereas an extended calculation models systems such as a surface or bulk material by periodically repeated unit cells. Wave function methods are almost exclusively only applied to finite systems, whereas DFT methods are routinely applied to both finite and extended systems. Extended systems are modeled using periodic boundary conditions (PBC), a good review of which can be found in Ref. [3], even if the underlying system is itself not periodic. The application of PBC requires that a supercell be defined that will be periodically repeated along the lattice vectors. This model naturally applies to periodic systems, such as crystals and their surfaces. However, it can also be applied to systems that do not exhibit periodicity by defining a relatively large supercell containing a locally nonperiodic structure that is repeated on the longer length scale of the chosen supercell. Quantum simulations of surface reactions, such as those of ALD on the growing surface, can be performed using either a finite cluster or periodically extended supercell model of the reacting surface site. Thus, for simulating ALD surface reactions the researcher must decide whether to use a finite cluster or extended model of the reacting surface.

If a cluster is chosen to model the reacting surface, it should be large enough that it includes the changes in the system occurring during the chemical reaction, including all significant interactions with the surrounding material. In other words, any changes induced on the substrate by the ALD surface reactions should be insignificant at the cluster edge. If a cluster model is chosen that is too small to model the ALD surface reaction, the edge of the cluster can introduce nonphysical interactions with the reacting atoms. For example, for a material such as GaAs the broken bonds at the cluster edges should be terminated in such a way as to preserve the electronic structure of the surface. However, a material such as GaAs involves bonds that are partially covalent and bonds that are partially ionic. When a surface is created, the broken bonds preferentially become lone pairs (in the case of As) and empty orbitals (in the case of Ga). Thus, in the case of a cluster, the surfaces that define the edge of the cluster are more difficult to “terminate” as the terminating groups, such as a H atom, will not form a bond to the Ga or the As atom that accurately mimics the Ga–As bonds of the bulk material. Of course, these artifacts can be reduced by using large clusters so that the edges are far from the reactive site of the cluster. Unfortunately,

most quantum chemical methods scale with at least the third power of the system size. It is then important to define cluster models that are just large enough to accurately describe the effects at play during the reaction. Fortunately, the effects involved in chemical reactivity are relatively local phenomena and so the assumption that reactions on a cluster mimic reactions on the surface is usually valid. This is true except in cases of cluster models that are obviously too small because the gas-phase reagent reacts with more than a single surface reactive site, or because of edge effects due to the cluster being too small or designed with poor edge termination.

If a supercell model of the reacting surface is chosen to model the reactive surface site, then a supercell must be defined that is large enough to include the relevant interactions within the cell and where self-interactions resulting from interactions with periodic images in neighboring cells are small. Although the underlying substrate may be periodic, surface reactions break this periodicity as does the growth of the film, which is rarely epitaxial with the underlying substrate. Spurious periodic interactions with images in neighboring unit cells caused by the surface reactions and imposed by PBC can lead to artifacts, and thus like cluster models, care must be taken to avoid edge effects when supercell size is insufficient to isolate the reacting surface site from neighboring reacting sites. Although both approaches have their drawbacks, they have both been successfully applied to simulations of ALD and generally agree with each other when the choices made in model design and method are appropriate so as to not introduce significant artifacts [14].

1.2.5

Basis Set Expansions

In both finite and extended methods, the electronic structure is expanded as linear combinations of basis functions. For finite systems, including clusters of atoms designed to model regions of an extended structure such as a surface, basis functions are localized atomic orbital-like functions with their radial component described using either Gaussians, $\exp(-\zeta r^2)$, or Slater functions, $\exp(-\zeta r)$, where the exponent ζ determines the spatial extent of the basis function. On the other hand, extended systems modeled using supercells within PBC generally expand the electron density over a series of plane waves up to a chosen cutoff energy. This expansion is analogous to a Fourier expansion of a periodic function, in this case the density, and involves the usual difficulties of Fourier expansions, such as describing functions with discontinuities in their derivatives. This challenge can be met by including very high frequency plane waves in the expansion or describing the nucleus and core electrons of the system using pseudopotentials. This also has the effect of reducing the number of electrons that are treated explicitly, greatly reducing the computational cost for systems with large numbers of electrons.

Large basis sets include more basis functions or plane waves to generally allow for a better description of the wave function or the electron density, which is the square of the wave function. When using localized or plane wave basis functions, the number of unknown coefficients to be determined, and thus the computational demand, grows with the size of the basis set. Fortunately, the quality of the calculations typically

converges at reasonably sized basis sets, although it is good practice to confirm this convergence. Methods have also been developed to estimate the energy predicted by infinitely large basis sets using what is called a basis set extrapolation.

1.3

Calculation of Properties Using Quantum Simulations

Quantum mechanical calculations are generally relatively computationally intensive. However, their ability to describe esoteric phenomena and accurately predict system properties, often beyond the capability of experiment, motivates their use. A wide variety of properties useful in understanding an ALD process can be computed using quantum simulations. The most basic and widely calculated property is the system energy. Although the system energy itself has little intrinsic value, differences in energies form the basis for calculating a wide range of properties. These include bond energies, adsorption energies, reaction energies, intermediate energies, activation barriers, electron affinities, ionization potentials, excitation energies, band gaps, and free energies. Furthermore, first and second derivatives of the energy with respect to the atomic coordinates allow one to calculate the forces and vibrational frequencies of the system, and thus optimize geometries. Calculation of forces allows one to predict stationary states (where all forces are zero) including stable structures and transition states, which are first-order saddle points on the system potential energy surface (PES) $E(R_A, \dots, R_{3N})$, where the set $\{R_i\}$ contains the $3N$ coordinates of the N atoms. Furthermore, the calculation of frequencies enables the prediction of experimental vibrational spectra, such as infrared and Raman frequencies. Calculated frequencies also allow for the determination of zero-point energies and vibrational partition functions, from which entropies, enthalpies, heat capacities, and thermal corrections can be calculated, which then enable the calculation of free energies and other temperature-dependent properties.

Additional properties that can be calculated and that are useful for understanding ALD include the electron density, from which valuable insight into the electronic structure of the reacting system can be gleaned, HOMO–LUMO gaps, which are valuable in understanding the interactions involved during reaction, and band gaps and band offsets, which are often useful in predicting the electronic properties of the film and its interface with the substrate or an overlying metal electrode.

1.3.1

Calculation of Transition States and Activation Barriers

Our focus on simulations of the ALD process, specifically the reactions involved in ALD, naturally motivates the desire to predict properties related to the chemical reactivity. These properties include the active chemical mechanism of the process and the associated energetics and kinetics of the mechanism. We have already mentioned that the calculation of forces allows us to optimize structures to determine the energies of reactants, intermediates, products, and TS connecting these stable

structures. If we optimize a series of stable structures from reactants and products along with a series of TS connecting the various intermediate structures, we have a candidate mechanism for a chemical reaction. The differences in energy between the transition states and the structures they connect are the forward and reverse barriers for that step in the mechanism. Finding a low-energy path from reactants to products allows us to state that the pathway is viable. However, we cannot ignore the possibility that a lower energy pathway may exist. This generally motivates the researcher to explore all reasonable pathways, meaning pathways that are chemically rational. Unfortunately, although the optimization of stable structures along the minimum energy path (MEP) of the reaction is straightforward in modern simulation codes, TS searches are still challenging despite the development of various algorithms for locating saddle points on the potential energy surface. One challenge is the sheer number of saddle points in a system with even only a moderate number of atoms and the possibility that incorrect saddle points are located. Another is that saddle points themselves are more subtle topological features on the potential energy surface than the minima of stable structures and involve the more complex constraint that all second derivatives, except the one that corresponds to the reaction coordinate, are zero. Thus, the forces on the atoms near transition states tend to be smaller than those on the atoms near stable structures for equal displacements from the stationary point.

Best practices for determining reaction mechanisms include:

- 1) using chemical intuition to initiate TS searches at starting structures near the TS structure,
- 2) ensuring that the right number of imaginary modes are found,
- 3) visualizing the vibrational mode corresponding to the reaction coordinate to confirm that it moves the geometry toward the two states the TS is intended to connect, and
- 4) performing an intrinsic reaction coordinate calculation where the MEP is followed in both directions from the TS to identify what structures the TS connects.

These efforts locate and confirm the nature of the TS, but unfortunately do not ensure that this TS is actually that of the MEP. And even if the correct TS is found, it is possible that a competitive pathway exists at the reaction temperature. Consequently, it is important to locate all low-energy pathways that might be active at the reactor conditions. A second issue involves the possibility that even if the reactants follow the MEP through the lowest energy TS, they may not follow the MEP on the product side of the potential energy surface because as the potential energy of the system at the TS is converted to kinetic energy as the system proceeds toward the products, that kinetic energy can enable the system to follow a trajectory different from the MEP along the exit channel. This possibility is relatively rare, but is more probable in cases where the kinetic energy is not quickly thermalized (redistributed to modes other than the reaction coordinate) or where low-energy TS within the exit channel exist. The above issues make locating of TS structures one of the hardest challenges for the user of simulation methods. Fortunately, experience, understanding of the chemical nature of the system, and creative use of the various techniques for finding TS structures will reward the persistent researcher.

Once sufficient confidence in a predicted mechanism is obtained, one can then use the schematic potential energy surface of this mechanism as a basis for understanding the relevant ALD process. While this is valuable, a number of additional useful properties related to the mechanism can also be calculated.

1.3.2

Calculation of Rates of Reaction

In addition to predicting the activation barriers for each step in a mechanism, one can also use the vibrational frequencies of the intermediate (or reactant) and transition states to calculate their partition functions and thus the preexponential factor of the rate constant using transition state theory (TST). Consequently, a quantitative prediction of the rates of reaction for the individual steps and an overall reaction can be calculated. Although the prediction of rate constants using TST with input calculated from quantum calculations is enormously valuable, several words of caution should be provided. The most important is that because the rate law depends exponentially on the ratio of the activation barrier to the thermal energy $k_B T$, where k_B is the Boltzmann constant, through the Boltzmann factor $\exp(-E_A/k_B T)$, a small error in the barrier can lead to considerable errors in the predicted rate. This is exacerbated at low temperatures where $\exp(-\Delta E_A/k_B T)$ is large for relatively small errors in the barrier, ΔE_A . For example, at 298 K an approximately 1.36 kcal/mol error in an activation barrier leads to a factor of 10 error in the rate constant. Unfortunately, most methods do not achieve an error of less than 1.5 kcal/mol in the activation barrier. For example, for most DFT methods errors of 3–4 kcal/mol are relatively typical making errors in the rate constant at 298 K near a factor of 10 000 common. The good news is that most methods tend to be relatively systematic in their under- or overprediction of barriers so that while the predicted rates may be off by a substantial amount, the relative rates between different steps will generally be in error by a significantly lower amount. For example, DFT methods tend to systematically underestimate barriers. Thus, the prediction of branching ratios can be expected to be relatively good compared to the absolute rates. When the competing pathways involve barriers that are different by more than 3–4 kcal/mol, DFT methods can reliably predict the correct active reaction. Of course, more accurate methods can be used to calibrate and verify the predictions of DFT or to provide a correction factor. For example, the CCSD(T) coupled cluster method can predict barriers to within ~ 1 kcal/mol if the reaction possesses what is called “single-reference” character. Because CCSD(T) is computationally demanding, it is usually just used on small systems, for example, small models of the reaction including less than 10 heavy atoms, that are analogous to that of the actual ALD reaction. Another kinetic property that can be calculated is the kinetic isotope effect (KIE) because it is just the ratio of the rate constants of a reaction where in one case the reacting species has been isotopically substituted. Prediction of KIEs for the rate-limiting steps of a reaction thus allows simulation to predict experimental KIEs, providing a powerful approach to confirming the rate-limiting step, and thus the mechanism of a reaction.

A less common quantum simulation approach for investigating ALD involves *ab initio* molecular dynamics. While classical MD is relatively efficient, as mentioned above, it does not usually provide an accurate description of the bond dissociation or formation process. *Ab initio* MD, on the other hand, allows one to follow dynamics trajectories with energies and forces described using various flavors of quantum mechanics, most commonly DFT. Unfortunately, the computational expense of DFT, although low relative to methods of comparable or greater accuracy, still prohibits the simulation of large systems or for processes that take place over relatively long timescales compared to the vibrational timescale of the system. Despite these limitations, DFT-based MD can be quite useful in elucidating the details of an ALD process. The two main varieties of *ab initio* MD are the Born–Oppenheimer (BOMD) and Car–Parrinello (CPMD) methods, reviewed in Ref. [3]. BOMD solves the electronic structure problem explicitly at each time step and uses the resulting energy to calculate forces on the ions. In contrast, CPMD treats the electron coordinates explicitly as dynamical variables in an extended Lagrangian. Both approaches have proven quite useful, although the BOMD approach does not suffer from the electronic drag and nonadiabaticity often encountered with CPMD. These problems are due to how CPMD propagates the electron density forward where it is no longer the correct electronic structure for the “external potential” presented by the ion arrangement.

This brief treatment of quantum simulation methods is only intended to remind the reader of the basic ideas related to these simulation methods. For readers who are inclined to delve into the field of quantum simulations, a thorough review of quantum mechanics and quantum simulations is highly recommended [1–3].

1.4 Prediction of ALD Chemical Mechanisms

As described generally above, electronic structure methods such as DFT are uniquely capable of providing a first principles description of the chemical reactivity of reacting systems. This involves determining the active reaction pathways by identifying and characterizing the intermediates and transition states leading to the products of the ALD reaction. The key challenges are (i) locating the low-lying transition states, (ii) confirming that they indeed connect the states along the reaction path, and (iii) establishing confidence that these transition states are indeed the active pathways. Transition states are first-order saddle points, meaning all forces are zero (they are stationary points) and all second derivatives except one of the energy with respect to the nuclear coordinates are positive. The negative second derivative is the normal mode corresponding to the reaction coordinate for this step. Locating these often subtle topological features in a hyperdimensional surface can be difficult. This is exacerbated by the fact that many low-lying first-order saddle points may reside on this energy landscape, including hindered rotors (e.g., methyl groups) and skeletal modes of larger molecules. Furthermore, at higher temperatures pathways that involve higher lying transition states may become active. Although these obstacles

can make determining ALD mechanisms using quantum chemical methods problematic, it is usually the only practical way to obtain a detailed description of the chemistry, and especially of the nature of the transition states as this type of information is not readily available from experiment. Despite these challenges, researchers have successfully used quantum chemical calculations to explain the mechanisms of a variety of ALD chemistries. We will next discuss the general types of ALD reactions that arise in ALD chemistry and then go through the specific example of the ALD of Al_2O_3 using trimethylaluminum (TMA) and water to illustrate a simple case of determining an ALD mechanism with DFT [20].

1.4.1

Gas-Phase Reactions in ALD

The first possible reaction in the ALD system might involve gas-phase self-reactions that could occur when the precursor is introduced into the ALD chamber. Although ALD precludes self-reactions, one could first calculate the barriers to possible gas-phase self-reactions to estimate the extent of the CVD component in this ALD chemistry as a function of temperature. Ideally, the CVD component is very low, but prediction of these barriers could provide evidence to show this. Often, experiment has shown that the precursors of the ALD chemistry being investigated do not self-react at reasonable temperatures and this step is skipped.

Other gas-phase reactions that could be studied as a prelude to the ALD surface chemistry are the reactions between the complementary ALD precursors. Although these species are introduced into the reactor in pulses temporally separated by intervening reactor purges to prevent them from reacting anywhere but on the substrate, the predicted reaction energies and TS energies can provide a reasonable estimate for the thermodynamics and kinetics of the ALD surface reactions. This is because these reactions are often isodesmic to the surface reactions. That is, they involve breaking and forming the same number and types of bonds and often proceed through a similar TS. Consequently, they are small analogues of the surface reactions and thus their chemistry can be calculated relatively quickly to screen and develop understanding of the ALD surface chemistry. In addition to providing insight into the driving forces and rates of the ALD reactions, they can also be used to facilitate the search for the TS of the analogous surface reactions.

1.4.2

Surface Reactions in ALD

1.4.2.1 Adsorption Reactions in ALD

The ALD process can be generally treated as a series of surface chemical reactions. Although the surface reactions that define ALD are self-limiting, they can be treated in the same fashion as the surface reactions of chemical vapor deposition, organic functionalization of surfaces, and even heterogeneous catalysis. This entails exploring the individual steps that may be active in the reacting system. A typical first step in an ALD mechanism is adsorption, which can be either dissociative or molecular.

Usually adsorption of the ALD precursor on the reactive substrate is molecular because of the self-limiting requirement of the surface chemistry. To calculate the adsorption energy, one calculates the energy of adsorbed state and subtracts the energies of the initial state, which is the isolated precursor and the substrate, with the structures of each of these states optimized to their minimum energy geometries. Adsorption may or may not involve going over a barrier. In the majority of cases, molecular adsorption either is barrierless or has an insignificant barrier [5, 8, 9, 14, 18–21]. In either case, the PES that describes the energy as a function of the adsorption reaction coordinate can be calculated.

Although adsorption is usually one of the simplest steps in an ALD mechanism to calculate, a few points of caution should be exercised. One is that the molecularly adsorbed state might involve van der Waals interactions, which are poorly described by most DFT methods not specifically designed to describe dispersive interactions. A second issue relates to the possibility that a precursor may adsorb to more than one surface site and that these sites might be different. If this is the case, the model of the reacting substrate must include these sites. This point becomes even more important when the precursor is relatively large or possesses ligands that can bond to various possible surface sites. Similarly, this can also be important in cases where different substrate conditions provide various possible surface reactive sites. For example, in the ALD of metal oxides using water as the oxygen source, a purge subsequent to the water pulse may leave the substrate with different relative concentrations of OH groups, $-O-$ bridge sites, and even molecularly adsorbed H_2O [14, 44]. Finally, caution should be exercised when the interaction between the adsorbed precursor and the substrate can lead to nonlocal effects. For example, in the ALD of metal oxides and nitrides, metal precursors can act as Lewis acids and dative bond to surface Lewis base groups to form Lewis acid–base complexes. Similarly, oxygen and nitrogen sources can act as Lewis bases and dative bond to metal atoms of the surface, which act as Lewis acids. In the latter case, the donation of an electron lone pair from the adsorbing species to the surface can lead to charge transfer to neighboring sites, which requires that the model include those neighboring sites to describe this effect.

Although adsorption may appear less interesting than the reactions directly involved in removing precursor ligands to deposit the atoms of interest, it can play an important role in both establishing the initial structures from which ligand exchange reactions follow and often determining the energetics of the ALD process. For example, if the initial ALD adsorbed precursor–substrate complex is too stable, it can act as a trapped intermediate [18, 19]. Furthermore, the stability of this complex, together with the stability of the subsequent intermediate, affects the barrier connecting these states. If the TS to proceed to the following intermediate lies above the energy of the entrance channel (i.e., if the forward barrier is higher than the barrier to desorption), the ALD reaction may be relatively slow and have inefficient precursor utilization [18, 19].

1.4.2.2 Ligand Exchange Reactions

A key step in any ALD process is replacing the functional groups that cover the reacting surface with the functional groups of the reacting species. The surface reactions that are often responsible for this transformation are ligand exchange

reactions. For example, the adsorbed precursor complex often involves a dative bond between the initial surface functional group and the precursor. If this initial functional group is not a component of the deposited film, it must be transformed to remove any atoms that are not constituents of the final material. For example, if the reactive surface site is an OH group and the ALD process deposits a metal oxide, the H atom must be removed by the reactions of the ALD mechanism. In this example, this can be achieved by a hydrogen transfer from the OH to one of the ligands of the adsorbed precursor to form a volatile by-product to carry away the H together with one of the precursor ligands. This ligand exchange reaction thus converts the OH functionalized surface to one functionalized by the ligands of the precursor. Similarly, ligand exchange can replace these functional groups by OH groups via a H transfer from adsorbed or absorbing H₂O and the surface ligand.

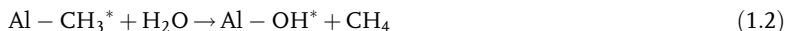
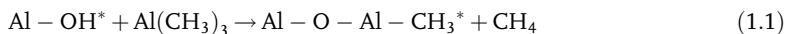
Although a ligand exchange reaction replaces the ligands deposited by the previous half-cycle with those of the pulsed precursor, the by-product of that reaction must be volatile. Often, this species is physisorbed to the surface, for example, through a dative bond arising from a dipolar interaction. This binding must not be so strong as to severely inhibit the desorption of the product, otherwise these species may require long purges at relatively high temperatures to remove them. A purge that does not completely remove these by-products can lead to film contamination and nonuniform growth through site blocking. Although the by-product should have a reasonably low desorption barrier, ALD does not require that each half-cycle be exothermic. In fact, each half-reaction can be relatively endothermic because ALD is run under nonequilibrium flow conditions. Thus, the chemical potential of the reactants remains relatively high because the pulse continues to supply the precursors while the chemical potential of the by-products is kept relatively low by the fact that as the by-product is formed and desorbed from the surface, it is removed from the reactor chamber by the flow, providing the thermodynamics to drive the reaction forward toward film growth despite the endothermicity of the reaction.

1.5

Example of a Calculated ALD Mechanism: ALD of Al₂O₃ Using TMA and Water

The ALD of Al₂O₃ using trimethylaluminum (TMA, Al(CH₃)₃) and water (H₂O) is one of the most commonly employed ALD processes, both because of the wide number of applications of thin films of Al₂O₃ and because this ALD process is one of the most likely to succeed for any given substrate. In fact, the ALD of Al₂O₃ using TMA and water is sometimes considered the prototypical ALD process [45–47].

As with any ALD process, the reactions between the precursors and the surface can be separated into two half-reactions:



where the asterisks denote the surface species. Although this reaction appears to be not balanced, it is written to illustrate the new bonds that form and the surface termination during each half-cycle. To determine the atomistic detail of the ALD mechanism represented by these two half-cycles, and to predict the associated thermochemistry and kinetics of the TMA and H_2O half-reactions, the B3LYP [48] gradient-corrected exchange DFT method combined with the 6-31 + G(d,p) double-zeta plus polarization and diffuse functions basis set was used by Widjaja and Musgrave [20]. The geometry of each species was optimized by finding stationary points followed by frequency calculations needed to identify the nature of the stationary points on the potential energy surface and to calculate the zero-point energy corrections.

The $\text{Al}(\text{OH})_2\text{-OH}$, $\text{Al}(\text{OH})(\text{CH}_3)_2$, and $\text{Al}(\text{OH}_2)\text{-CH}_3$ clusters were used to represent the Al-OH^* and Al-CH_3^* reactive surface sites. In addition, the larger $\text{Al}[\text{O-Al}(\text{OH})_2]_2\text{-OH}$, $\text{Al}[\text{O-Al}(\text{OH})_2]_2\text{-CH}_3$, and $\text{Al}[\text{O-Al}(\text{OH})_2](\text{CH}_3)_2$ clusters were used to determine the effect of cluster size by mimicking the effects of the surrounding material of a model Al_2O_3 surface site. Because the clusters were fully relaxed, the calculations may underestimate the strain energy of the system. The clusters used are shown in Figure 1.1. The PES of the $\text{Al}(\text{OH})_2\text{-OH} + \text{Al}(\text{CH}_3)_3$ reaction, representing reactions during the TMA pulse at Al-OH^* surface sites, is shown in Figure 1.2. TMA first adsorbs molecularly at an Al-OH^* with an adsorption energy of 14.1 kcal/mol. The adsorbed complex involves a Lewis acid–base interaction, with TMA acting as the Lewis acid and OH^* acting as the Lewis base. Next, one H atom from the surface -OH^* group transfers to the methyl group of the adsorbed TMA– OH^* complex to form CH_4 , which then desorbs. This transition state lies below the entrance channel: the TS energy is lower than that of the reactants, although it lies 12.0 kcal/mol above the chemisorbed TMA. Overall, the TMA half-reaction is exothermic by 25.1 kcal/mol and results in adding an Al layer to the surface

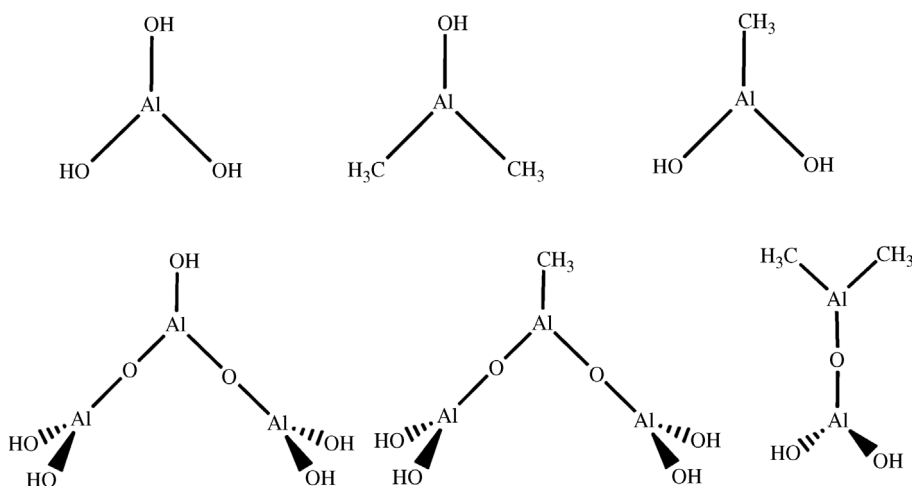


Figure 1.1 Schematic illustration of the clusters used to model reactive sites on the Al_2O_3 surface.

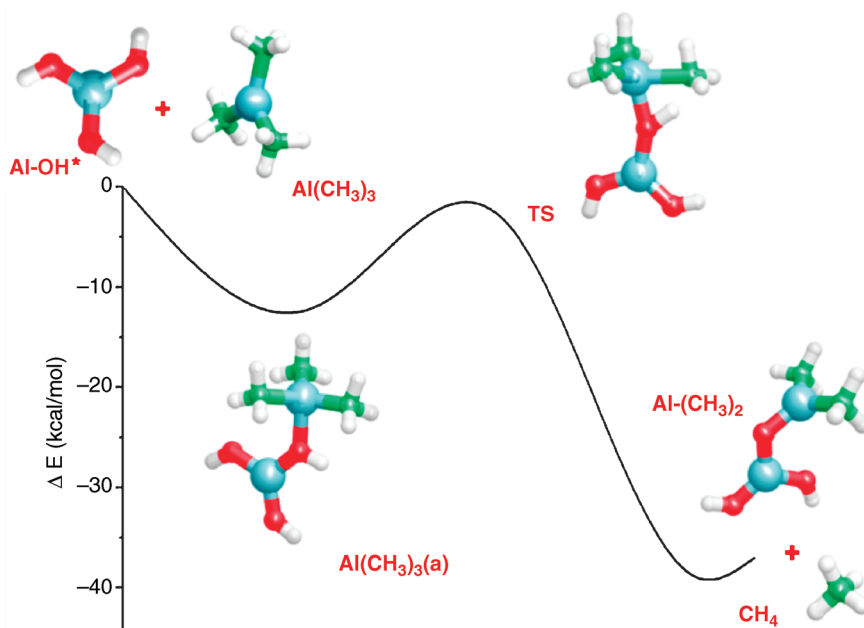


Figure 1.2 Schematic potential energy surface for the first half-reaction of TMA + an OH^* site on the growing Al_2O_3 surface. Note that the barrier lies ~ 2 kcal/mol below the reactants. See the text and Ref. [20] for more details.

while replacing surface hydroxyl groups with surface methyl groups. Following a reactor purge, H_2O is then introduced into the ALD chamber.

The water pulse reacts with methyl-terminated Al_2O_3 to replace methyl groups with hydroxyl groups while adding a layer of oxygen to the surface. Determining the mechanism of this half-reaction is analogous to the calculations described above to study the TMA half-cycle. In fact, not only is the approach similar, but the reaction is also isodesmic to that of the first half-cycle. In the water half-cycle, the OH bond of water and the Al–C bond of the surface $\text{Al}-(\text{CH}_3)_2^*$ site are broken while an Al–O bond and a C–H bond are formed to make $\text{Al}-(\text{CH}_3)(\text{OH})^*$ and CH_4 . This involves forming and breaking the same bonds as the TMA half-reaction, and consequently, the energetics are different to the degree that effects beyond the active atoms in the reaction are important, which is a general expectation of isodesmic reactions. To predict the reaction mechanism of H_2O with $\text{Al}-(\text{CH}_3)_2^*$ sites, we calculated the states along a reaction path similar to that of the first half-reaction. The resulting PES is shown in Figure 1.3. As Figure 1.3 shows, H_2O first adsorbs molecularly by forming a Lewis acid–base complex with the $\text{Al}-(\text{CH}_3)_2^*$ surface site. The complex energy is 13.1 kcal/mol below the energy of the reactants. The adsorption of water is followed by CH_4 formation and desorption, which involves a barrier of 16.1 kcal/mol and an overall reaction energy of -21.0 kcal/mol, both relative to the adsorbed H_2O complex. As expected for the ALD process, reaction of $\text{Al}-(\text{CH}_3)_2^*$ with H_2O results in $\text{Al}-(\text{CH}_3)(\text{OH})^*$ sites and a second reaction with water is required to remove the

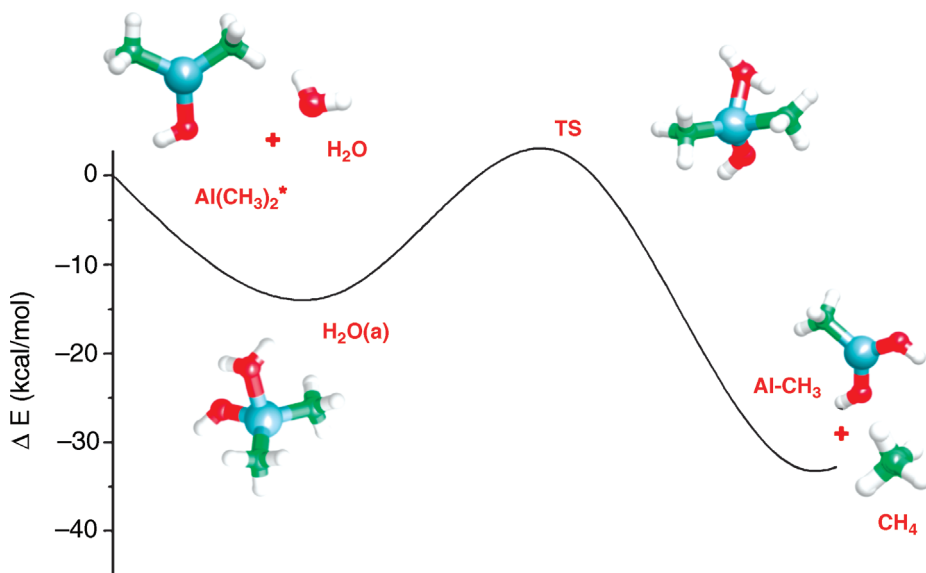


Figure 1.3 Schematic PES of the reaction of water with the Al-CH₃* sites on the methyl-terminated surface. Unlike the reaction of TMA with Al-OH* surface sites, the TS lies above the entrance channel. Note that the water reaction

is slightly less exothermic than the TMA reaction and because the barrier lies above the entrance channel, it is kinetically slower than the TMA reaction. See the text and Ref. [20] for more details.

remaining methyl group. Reaction of the Al-CH₃ group with H₂O is subsequently investigated. Similar to the reaction of Al-(CH₃)₂ with H₂O, the reaction follows a trapping-mediated pathway. Although the reaction is analogous to reaction at the site with two methyl groups, replacing one methyl by a hydroxyl group affects the energetics of the reaction. We found that the adsorbed water complex energy is -17.1 kcal/mol, 4.0 kcal/mol more stable than for the Al-(CH₃)₂* surface site. We then calculated that the ligand exchange reaction barrier was 21.0 kcal/mol relative to the H₂O adsorbed state, ~5 kcal/mol larger than for reaction at the Al-(CH₃)₂* site. Finally, we saw that the reaction was exothermic by 12.9 kcal/mol, which was about 8 kcal/mol less exothermic than for water reacting at the Al-(CH₃)₂* surface site. Despite these differences, the qualitative nature of the chemistry is quite similar. Although the differences in the reactivity may suggest that the effects are nonlocal, it should be remembered that the two cases shown here differed by exchanging a methyl group directly bound to the active Al surface atom with an OH group. In the case of using different clusters to model surface reactions, the effects are less direct as the clusters are modified further from the reactive site of the cluster, as shown next.

We also calculated the reaction energies using the larger Al(O-Al(OH)₂)₂-OH, Al[O-Al(OH)₂](CH₃)₂, and Al[O-Al(OH)₂]₂-CH₃ clusters. The reaction mechanisms are identical to those found using the analogous smaller clusters. Furthermore, the calculated energies are very similar to those calculated using the smaller clusters, with energy differences of less than 2.5 kcal/mol. This showed that the ALD surface

reactions for the case of TMA and H₂O on the Al₂O₃ surface do not exhibit significant nonlocal effects and that relatively small clusters are good models of the surface active sites because they are sufficient to describe the surface reactions.

Similar to the ALD of Al₂O₃ using TMA and water, formation of complexes has also been predicted for the ALD of ZrO₂ and HfO₂ with H₂O and ZrCl₄ and HfCl₄ precursors [5, 12, 14, 18, 19]. In those cases, an oxygen lone pair is donated to an empty d orbital of the metal center to form a Lewis acid–base complex with an energy below that of the final products of the half-cycle. In contrast to the complexes formed in the cases of ZrO₂ and HfO₂ ALD, the energy of the TMA–water complex is above the energy of the final product. As a result, Al₂O₃ ALD is exothermic relative to both the initial state and the complex, so the H₂O complex should not be observed at room temperature, unlike the cases of ZrO₂ and HfO₂. This is confirmed experimentally for water reacting with methyl-terminated Al₂O₃ with infrared spectroscopy where the H₂O bending mode in the region between 1380 and 1650 cm⁻¹ is not observed [49]. In the case of Al₂O₃ deposition, long purge sequences to drive the reaction toward the products are probably not necessary, because equilibrium favors the reaction products, in contrast to ALD of ZrO₂ and HfO₂. However, if excess water is introduced beyond that necessary to remove methyl groups, then a long purge may be required.

This example illustrates the basic use of quantum chemical methods to investigate an atomistic mechanism, namely, the simple chemistry of Al₂O₃ ALD using TMA and H₂O. It is quite common to find that the ALD half-reactions for various metal oxides and nitrides proceed through a stable intermediate complex before the reaction removing the H atom of the OH or NH₂ surface reactive site. These complexes result from the interaction between the oxygen or nitrogen lone pair electrons and the empty p orbital of the metal atom.

References

- 1 Szabo, A. and Ostlund, N.S. (1996) *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Dover, Mineola, NY.
- 2 Cramer, C. (2004) *Essentials of Computational Chemistry*, John Wiley & Sons, Ltd, Chichester.
- 3 Payne, M.C., Teter, M.P., Allan, D.C., Arias, T.A., and Joannopoulos, J.D. (1992) *Rev. Mod. Phys.*, **64**, 1045.
- 4 Ardalan, P., Musgrave, C.B., and Bent, S.F. (2009) *Langmuir*, **25**, 2013.
- 5 Han, J.H., Gao, G.L., Widjaja, Y., Garfunkel, E., and Musgrave, C.B. (2004) *Surf. Sci.*, **550**, 199.
- 6 Heyman, A. and Musgrave, C.B. (2004) *J. Phys. Chem. B*, **108**, 5718.
- 7 Kang, J.K. and Musgrave, C.B. (2002) *J. Appl. Phys.*, **91**, 3408.
- 8 Kelly, M.J., Han, J.H., Musgrave, C.B., and Parsons, G.N. (2005) *Chem. Mater.*, **17**, 5305.
- 9 Mui, C. and Musgrave, C.B. (2004) *J. Phys. Chem. B*, **108**, 15150.
- 10 Mui, C., Senosiain, J.P., and Musgrave, C.B. (2004) *Langmuir*, **20**, 7604.
- 11 Mui, C., Widjaja, Y., Kang, J.K., and Musgrave, C.B. (2004) *Surf. Sci.*, **557**, 159.
- 12 Mukhopadhyay, A.B. and Musgrave, C.B. (2006) *Chem. Phys. Lett.*, **421**, 215.
- 13 Mukhopadhyay, A.B. and Musgrave, C.B. (2007) *Appl. Phys. Lett.*, **90**, 173120.
- 14 Mukhopadhyay, A.B., Musgrave, C.B., and Sanz, J.F. (2008) *J. Am. Chem. Soc.*, **130**, 11996.

- 15 Mukhopadhyay, A.B., Sanz, J.F., and Musgrave, C.B. (2007) *J. Phys. Chem. C*, **111**, 9203.
- 16 Musgrave, C.B., Han, J.H., and Gordon, R.G. (2003) *Abstr. Pap. Am. Chem. Soc.*, **226**, U386.
- 17 Musgrave, C.B., Mukhopadhyay, A., and Sanz, J.F. (2007) *Abstr. Pap. Am. Chem. Soc.*, **234**.
- 18 Widjaja, Y. and Musgrave, C.B. (2002) *J. Chem. Phys.*, **117**, 1931.
- 19 Widjaja, Y. and Musgrave, C.B. (2002) *Appl. Phys. Lett.*, **81**, 304.
- 20 Widjaja, Y. and Musgrave, C.B. (2002) *Appl. Phys. Lett.*, **80**, 3304.
- 21 Xu, Y. and Musgrave, C.B. (2004) *Chem. Mater.*, **16**, 646.
- 22 Xu, Y. and Musgrave, C.B. (2005) *Surf. Sci.*, **591**, L280.
- 23 Xu, Y. and Musgrave, C.B. (2005) *Appl. Phys. Lett.*, **86**, 192110.
- 24 Xu, Y. and Musgrave, C.B. (2005) *Chem. Phys. Lett.*, **407**, 272.
- 25 Dkhissi, A., Esteve, A., Mastail, C., Olivier, S., Mazaleyrat, G., Jeloica, L., and Rouhani, M.D. (2008) *J. Chem. Theory Comput.*, **4**, 1915.
- 26 Dkhissi, A., Mazaleyrat, G., Esteve, A., and Rouhani, M.D. (2009) *Phys. Chem. Chem. Phys.*, **11**, 3701.
- 27 Esteve, A., Rouhani, M.D., Jeloica, L., and Esteve, D. (2003) *Comput. Mater. Sci.*, **27**, 75.
- 28 Fetah, S., Chikouche, A., Dkhissi, A., Esteve, A., Rouhani, M.D., Landa, G., and Pochet, P. (2010) *Thin Solid Films*, **518**, 2418.
- 29 Jeloica, L., Esteve, A., Rouhani, M.D., and Esteve, D. (2003) *Appl. Phys. Lett.*, **83**, 542.
- 30 Fenno, R.D., Halls, M.D., and Raghavachari, K. (2005) *J. Phys. Chem. B*, **109**, 4969.
- 31 Halls, M.D. and Raghavachari, K. (2003) *J. Chem. Phys.*, **118**, 10221.
- 32 Halls, M.D. and Raghavachari, K. (2004) *J. Phys. Chem. A*, **108**, 2982.
- 33 Halls, M.D. and Raghavachari, K. (2004) *J. Phys. Chem. B*, **108**, 4058.
- 34 Halls, M.D., Raghavachari, K., Frank, M.M., and Chabal, Y.J. (2003) *Phys. Rev. B*, **68**, 161302.
- 35 Raghavachari, K. and Halls, M.D. (2004) *Mol. Phys.*, **102**, 381.
- 36 George, S.M., Ott, A.W., and Klaus, J.W. (1996) *J. Phys. Chem.*, **100**, 13121.
- 37 Hartree, D.R. (1928) *Proc. Camb. Philos. Soc.*, **24**, 89.
- 38 Slater, J.C. (1951) *Phys. Rev.*, **81**, 385.
- 39 Slater, J.C. (1930) *Phys. Rev.*, **35**, 0210.
- 40 Bartlett, R.J. (1981) *Annu. Rev. Phys. Chem.*, **32**, 359.
- 41 Grimm, R.C. and Storer, R.G. (1971) *J. Comput. Phys.*, **7**, 134.
- 42 Hohenberg, P. and Kohn, W. (1964) *Phys. Rev. B*, **136**, B864.
- 43 Kohn, W. and Sham, L.J. (1965) *Phys. Rev.*, **140**, 1133.
- 44 Mukhopadhyay, A.B., Sanz, J.F., and Musgrave, C.B. (2006) *Chem. Mater.*, **18**, 3397.
- 45 George, S.M., Sneh, O., Dillon, A.C., Wise, M.L., Ott, A.W., Okada, L.A., and Way, J.D. (1994) *Appl. Surf. Sci.*, **82–83**, 460.
- 46 Dillon, A.C., Ott, A.W., Way, J.D., and George, S.M. (1995) *Surf. Sci.*, **322**, 230.
- 47 Ott, A.W., McCarley, K.C., Klaus, J.W., Way, J.D., and George, S.M. (1996) *Appl. Surf. Sci.*, **107**, 128.
- 48 Stephens, P.J., Devlin, F.J., Chabalowski, C.F., and Frisch, M.J. (1994) *J. Phys. Chem.*, **98**, 11623.
- 49 Hakim, L.F., Blackson, J., George, S.M., and Weimer, A.W. (2005) *Chem. Vapor Depos.*, **11**, 420.

