Computational Catalysis

RSC Catalysis Series

Titles in the Series:

1: Carbons and Carbon Supported Catalysts in Hydroprocessing
2: Chiral Sulfur Ligands: Asymmetric Catalysis
3: Recent Developments in Asymmetric Organocatalysis
4: Catalysis in the Refining of Fischer–Tropsch Syncrude
5: Organocatalytic Enantioselective Conjugate Addition Reactions: A Powerful Tool for the Stereocontrolled Synthesis of Complex Molecules
6: *N*-Heterocyclic Carbenes: From Laboratory Curiosities to Efficient Synthetic Tools
7: *P*-Stereogenic Ligands in Enantioselective Catalysis
8: Chemistry of the Morita–Baylis–Hillman Reaction
9: Proton-Coupled Electron Transfer: A Carrefour of Chemical Reactivity Traditions
10: Asymmetric Domino Reactions
11: C–H and C–X Bond Functionalization: Transition Metal Mediation
12: Metal Organic Frameworks as Heterogeneous Catalysts
13: Environmental Catalysis Over Gold-Based Materials
14: Computational Catalysis

How to obtain future titles on publication:
A standing order plan is available for this series. A standing order will bring delivery of each new volume immediately on publication.

For further information please contact:
Book Sales Department, Royal Society of Chemistry, Thomas Graham House, Science Park, Milton Road, Cambridge, CB4 0WF, UK
Telephone: +44 (0)1223 420066, Fax: +44 (0)1223 420247
Email: booksales@rsc.org
Visit our website at www.rsc.org/books

Computational Catalysis

Edited by

Aravind Asthagiri
Ohio State University, USA
Email: asthagiri@chbmeng.ohio-state.edu

and

Michael J. Janik
Pennsylvania State University, USA
Email: mjanik@engr.psu.edu

RSC Publishing

RSC Catalysis Series No. 14

ISBN: 978-1-84973-451-6
ISSN: 1757-6725

A catalogue record for this book is available from the British Library

Published by The Royal Society of Chemistry,
Thomas Graham House, Science Park, Milton Road,
Cambridge CB4 0WF, UK

Registered Charity Number 207890

For further information see our web site at www.rsc.org

Preface

The RSC Catalysis Book Series has been publishing books focused on many aspects of catalysis since the 1970's, but to date there has not been a book in the series that has solely focused on computational modeling of heterogeneous catalysis. The importance of computational catalysis has grown over the past two decades and there are an increasing number of young researchers entering this area. The aim of this book is to provide a pedantic presentation of select methods in computational catalysis. Our hope is that this book will prove useful to the graduate student or other researchers already familiar with computer simulations, but interested in applying specific methods to their catalysis research.

In the first chapter, Lars Grabow (University of Houston) discusses the screening of catalysts through the use of first-principles methods. Using density functional theory (DFT), key descriptors and scaling relationships can be identified and incorporated with an appropriate microkinetic model. Such an approach allows for the rapid screening of materials based on DFT calculations.

One of the key challenges in modeling catalysts is the need to predict the appropriate surface structure at reaction conditions. Jason Bray and Bill Schneider (Notre Dame) present a detailed example of a first-principles based thermodynamic model for oxygen adsorption on Pt surfaces. They derive a cluster expansion model, fit to DFT data, which allows for exploring the complex heterogeneous oxygen phase as a function of temperature and oxygen partial pressure using Monte Carlo simulations. These types of simulations also allow for exploring surface reaction behavior under reaction conditions.

In the third chapter, Kuan-Yu Yeh and Mike Janik (Penn State University) present a detailed review of DFT-based modeling of electrocatalysts. The electrochemical interface is one of the more challenging environments to model, and several different models that vary in accuracy and computational expense

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

are presented. With these methods potential dependent reaction energies and barriers can be calculated for elementary steps. Specific examples are presented to illustrate how to apply these various models.

Another important area of computational catalysis is modeling the metal/oxide interface, which is discussed by Tom Senftle, Adri van Duin, and Mike Janik (Penn State). They review several applications, such as the water-gas shift reaction and hydrocarbon activation, and the stability of oxide phases, that applies both DFT-based calculations and charge transfer potentials.

Thomas Manz (New Mexico State University) and David Sholl (Georgia Tech) present the details and application of their charge partitioning method called the density derived electrostatic and chemical (DDEC) method. This method can be used to obtain chemically relevant atomic charges and spin moments for both periodic and non-periodic systems. Such output can assist in understanding the relationship between electronic structure and material properties, and can also be used as input into the fitting of classical potentials.

The last two chapters present details of two classical potentials that incorporate charge transfer. Adri van Duin and co-workers present the details of the ReaxFF potential and discuss several applications. Susan Sinnott and co-workers from the University of Florida present the charge optimized many body (COMB) potentials and its application to molecules and metals on oxide surfaces.

We appreciate the efforts made by the authors to present a wide range of important methods in computational catalysis at a level that can benefit a researcher learning these methods for their research.

Aravind Asthagiri
Michael J Janik

Contents

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

CHAPTER 1

Computational Catalyst Screening

LARS C. GRABOW

Chemical & Biomolecular Engineering, University of Houston, Houston, Texas, 77204-4004, USA
Email: grabow@uh.edu

1.1 Introduction

Brute-force attacks are known in cryptography as (typically illegal) attempts to hack into encrypted data by systematically trying all possible key combinations of letters, digits and special characters until the correct access key or password has been found. Although a brute-force attack is guaranteed to be successful, its application is limited to very small problems because of the time required to generate and test all possible key combinations. For example, a standard 128-bit encryption key has 2^{128} possible permutations. If we simply assume that a typical central processing unit (CPU) can generate 10^9 bit flips per second ($\sim$1 GHz), then the total time that is required to test all possible permutations is $2^{128}/10^9 = 3.4\times10^{29}$ seconds or 10^{22} years! For obvious reasons a brute-force attack is most likely going to fail for this problem and a more targeted strategy is needed.

The above example illustrates the shortcomings of a brute-force attack, but a variation of it is still one of the most widely used strategies for the development of heterogeneous catalysts in practice. By using a combinatorial chemistry approach with completely automated, high-throughput experimentation equipment, one can synthesize and test enormous libraries of catalysts for their catalytic activity for a specific reaction. A good example is the search for advanced water–gas shift catalysts, in which Yaccato *et al.* have synthesized over 50 000 catalysts and

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

tested them in more than 250 000 experiments for low, medium and high temperature water–gas shift conditions.[1] Their effort led to a proprietary noble metal catalyst that can reduce the reactor volume by an order of magnitude without increasing the reactor cost. Although this trial-and-error approach almost always leads to an acceptable catalyst, the search space is restricted by the amount of time and resources available and many, possibly far better, candidates can be missed. The quickly evolving alternative to experimental high-throughput catalyst testing is computational catalyst screening. This approach relies on the fact that the catalyst activity for many catalytic reactions is usually determined by a small number of descriptors, which can be calculated from first-principles density functional theory (DFT) simulations and stored in a large property database. Populating this property database with DFT data is the most time-consuming step in this process, but the resulting database is applicable to any reaction and only has to be generated once. With a comprehensive database in place, it becomes a very easy task to screen thousands of database records in a short amount of time to identify catalyst candidates that possess descriptor values within the optimal range for a given reaction. Although the computational screening process can still be interpreted as a brute-force attack, the complexity of the problem has been greatly reduced. Hence, the number of materials that can be screened computationally increases drastically when compared with the experimental counterpart. The list of catalysts that fall into the desired range of descriptor values may be narrowed down further by using cost, stability, environmental friendliness, or any other applicable criteria.[2,3] The remaining materials can then be synthesized and experimentally tested under realistic reaction conditions. In general, not all computationally screened candidates will be good catalysts, but good catalysts will usually be included in the candidate list.

Somorjai and Li have recently reviewed the major advances in modern surface science that only became possible through the successful symbiosis of theory and surface sensitive experimental techniques.[4] The recent literature also contains several examples where a descriptor-based approach, both theoretically and experimentally, has led to the discovery of new catalytic materials. The following list should not be understood as an exhaustive review, but is meant to serve as inspiration to the reader and to demonstrate the wide applicability of this method. Early on, Besenbacher *et al.* discovered graphite resistant Ni/Au alloy catalysts for steam reforming,[5] Jacobsen *et al.* found an active Co/Mo alloy for ammonia synthesis by interpolation in the periodic table,[6] and Toulhoat and Raybaud showed that the metal–sulfur bond strength can correctly predict trends in hydro-desulfurization activity on metal–sulfide catalysts.[7] These initial successes were followed by other prominent examples that include CO-tolerant fuel cell anodes,[8,9] Cu/Ag alloys as selective ethylene epoxidation catalysts,[10] near-surface alloys for hydrogen activation[11] and evolution[2,12], Ru/Pt core–shell particles for preferential CO oxidation,[13] Ni/Zn alloys for the selective hydrogenation of acetylene,[14] Sc and Y modified Pt and Pd electrodes[15] and mixed-metal Pt monolayer catalysts[16] for electro-chemical oxygen reduction, and the rediscovery of Pt as the most active and selective catalyst for the production of hydrogen cyanide.[17,18]

1.1.1 A Walk through a Computational Catalyst Design Process: Methanation

The most comprehensive example of a success story in computational catalyst design comes from the group of Jens Nørskov, who has pioneered the descriptor-based design approach and has applied it to numerous reactions.[19,20] In several publications his group has studied the methanation reaction ($CO + 2H_2 \rightarrow CH_4 + H_2O$), starting from a detailed electronic structure analysis and leading to the development of a patented technical methanation catalyst based on a Fe/Ni alloy.[21–25] In the beginning of any descriptor-based design study one must first answer the question: "What is the most suitable reactivity descriptor for the reaction?" This question is typically answered by thoroughly studying the underlying reaction mechanism and identifying the rate-limiting step and most abundant surface intermediates. However, intuition can sometimes replace a detailed mechanistic study and a descriptor can be found through an educated guess. In the case of the methanation reaction, CO dissociation is the most critical step in the reaction mechanism. For weakly interacting metal catalysts, the dissociation is rate limiting, whereas for strongly interacting catalysts, the surface is poisoned by adsorbed C and O atoms. This leads to the volcano curve in Figure 1.1(a), which shows the experimentally measured methanation activity as a function of the calculated CO dissociation energy. The top of the volcano corresponds to the maximum methanation activity and indicates the optimal value of the CO dissociation energy, which is the activity descriptor in this case. The next step in the catalyst design process is to screen a database of CO dissociation energies and search for catalysts with CO dissociation energies near the optimum. This screening may be combined with a cost estimation of the resulting material and can further be linked to a stability test. Figure 1.1(b and c) show pareto plots of binary transition metal alloys for which the CO dissociation energy was estimated through a simple interpolation scheme owing to the lack of an existing database. The most active catalysts, characterized by CO dissociation energies close to the optimum, lie to the left of the graph and are connected with a solid line indicating the pareto-optimal set. The pareto-optimal set of Figure 1.1(b) contains the cheapest catalysts for a given value of CO dissociation energies and, similarly, Figure 1.1(c) can be used to screen for alloy stability. Only alloys with a negative alloy formation energy are stable and their stability increases as the alloy formation energy becomes more negative. Upon careful inspection, one notices that $FeNi_3$ is not only contained in both sets, but it is also located at the "knee" of the activity pareto-optimal set, which indicates that neighboring solutions are worse with respect to either activity or cost. Clearly, $FeNi_3$ is a very promising catalyst candidate for the methanation reaction. This catalyst identification step concludes the theoretical design process and experimental verification of the theoretical prediction is necessary. Experimentally obtained methanation rates of Fe/Ni alloys as a function of Ni content are displayed in Figure 1.1(d) and clearly show that the computationally predicted $FeNi_3$ alloy is significantly more active than its components. As an outcome of this *tour de force* in computational catalyst design, a process based on Fe/Ni alloys has been patented for the hydrogenation of carbon oxides.[24]

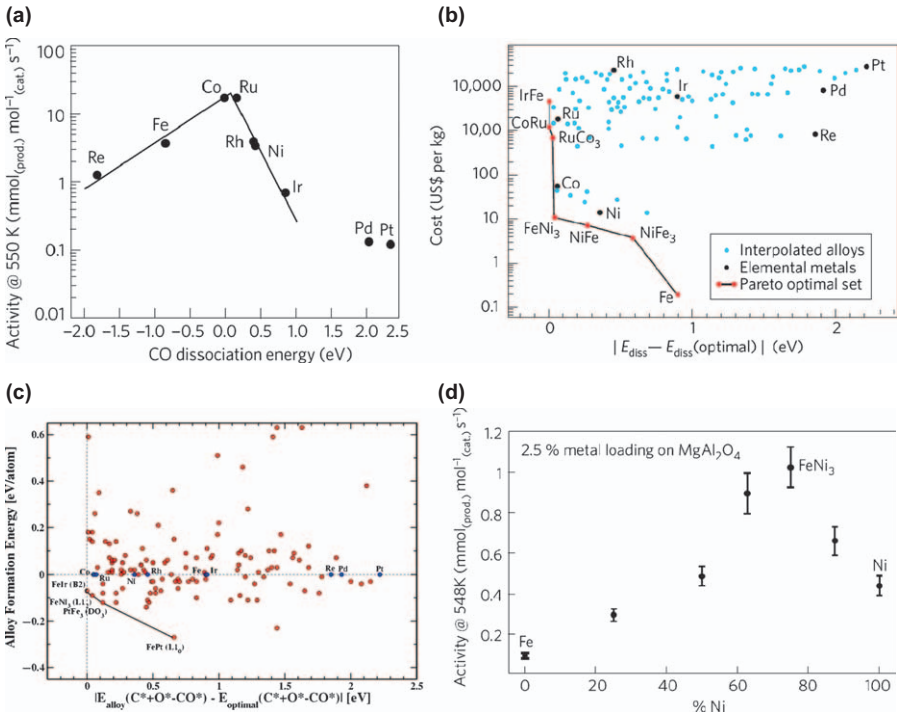


Figure 1.1 Computational design of a technical methanation catalyst. (a) A characteristic volcano curve is obtained when the experimentally determined methanation activity is plotted as a function of the CO dissociation energy. (b) and (c) Pareto-optimal bimetallic catalysts in terms of cost and stability. E_{diss}(optimal) refers to the optimal CO dissociation energy corresponding to the maximum of the activity volcano in panel (a). (d) Measured methanation activity of binary Fe/Ni alloys at T = 548 K, 2% CO in 1 bar H_2 as a function of Ni content. [(a), (b), (d) are reprinted by permission from Macmillan Publishers Ltd: Nature Chemistry, Ref. 15, 2009, (c) is from ref. 25]

In the remainder of this chapter, the background information that leads to the identification of appropriate catalyst descriptors (*e.g. d*-band model, scaling relationships) is reviewed and the basic strategy for successful catalyst screening using various levels of detail (*e.g.* Sabatier rate *vs.* microkinetic model) is outlined. A step-by-step illustration of the method will be given using ammonia synthesis and CO oxidation as examples. The interested reader is encouraged to work through the examples independently at her/his own pace.

1.2 Starting from the Electronic Structure

1.2.1 Density Functional Theory

Computational catalyst screening would not be possible without the existence of a theory that enables us to calculate the chemical properties of the catalyst

and the reaction of interest. Fortunately, in the mid 1960s Hohenberg, Kohn and Sham published two seminal papers formulating two theorems, which led to the development of density functional theory (DFT).[26,27] The contributions of Walter Kohn to the development of this theory were later honored in 1998 with the Nobel Prize in Chemistry. DFT is nowadays widely used in many different areas of science and engineering, including computational chemistry, catalysis, materials science, physics, and geology. The two theorems can be summarized as:

1. The ground state properties of a many-electron system are uniquely determined by the electron density.
2. The total energy of a system has a minimum for the ground state electron density.

DFT provides a solution to the Schrödinger equation:

$$\hat{H}\Psi = E\Psi \tag{1.1}$$

and is in principle an exact theory, but in practice the exact formulation of the kinetic energy term for a system of interacting electrons is unknown. In the Kohn–Sham approach, the kinetic energy is therefore approximated with the kinetic energy of a system of non-interacting electrons and a correction term, E_{XC}, which accounts for exchange and correlation effects in the interacting system. Although approximations for the description of the exchange-correlation energy must be made, DFT has the huge advantage over wave function based methods that the electron density is a function of only three spatial coordinates, while the many-body wave function for N electrons depends on $3N$ coordinates. Thus, DFT significantly reduces the computational intensity of the problem and enables the treatment of systems of several hundred atoms. From the electron density, $n(\mathbf{r})$, all other properties of the system are determined (Theorem 1) and the total energy E is calculated using

$$E[n(\boldsymbol{r})] = T_{\mathrm{KS}}[n(\boldsymbol{r})] + \frac{1}{2}\iint \frac{n(\boldsymbol{r})n(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|}\mathrm{d}\boldsymbol{r}\mathrm{d}\boldsymbol{r}' + \int v_{nucl}(\boldsymbol{r})n(\boldsymbol{r})\mathrm{d}\boldsymbol{r} + E_{XC}[n(\boldsymbol{r})] \tag{1.2}$$

The first term, $T_{\mathrm{KS}}[n(\boldsymbol{r})]$, is the kinetic energy of fictitious, non-interacting electrons and is obtained from the single-electron Kohn–Sham equations

$$\left(\frac{-\hbar^2}{2m}\nabla + \mathrm{v}_{eff}\right)\psi_i = \varepsilon_i\psi_i \tag{1.3}$$

where v_{eff} is the effective field defined by the nuclei and the current electron density. The second and third term in the total energy equation describe the electrostatic electron–electron interactions (Hartree energy) and the electron–nuclei interactions, respectively. The last term, E_{XC}, in the total energy equation depends on the unknown exchange-correlation functional, for which several approximations exist. The simplest approximation is the local density approximation (LDA), which can be derived from the case of a homogeneous

electron gas and only depends on the electron density at a single point. In this case, the exchange contribution in the LDA is exact, but the correlation still has to be approximated. The LDA works remarkably well for bulk materials where the electron density varies slowly, but has insufficient accuracy for most applications in chemistry, including atoms, molecules, clusters, and surfaces.

An obvious extension to the LDA is the generalized gradient approximation (GGA), which depends not only on the local density but on the density gradient. Because the gradient correction can be implemented into a GGA functional in many different ways, there exist a variety of different GGA flavors. The most widely used GGA functionals are the Perdew–Wang 91 (PW91)[28,29] and the Perdew–Burke-Ernzerhof (PBE)[30] functional. Both GGA functionals have good accuracy for a wider range of problems than the LDA because they contain more physical information; however, they are not necessarily always better. The PBE functional was later revised by Hammer, Hansen and Nørskov (RPBE), in order to improve the accuracy of chemisorption energies of atoms and small molecules on transition metal surfaces.[31] These GGA functionals are very good general-purpose functionals and may be used as a starting point for computational catalyst design. However, the GGA still fails for problems such as the accurate prediction of band gaps in semiconductors, systems where van der Waals interactions are dominant, or for electronic structure calculations in materials with strongly correlated electrons, where self-interaction errors can be encountered. Several improvements to the GGA have been suggested (*e.g.* DFT + U, DFT-D, meta-GGA, hybrid-GGA), but many of these functionals are problem specific or contain adjustable parameters that need to be fitted for each system. This empirical nature, along with the increased computational effort, renders these functionals generally unsuitable for computational catalyst screening. Work to improve XC functionals further is ongoing in the community, and in the next few years faster computers and new functionals will have a positive impact on the quality of DFT calculations.

Although DFT calculations are at the heart of computational catalysis, it is not strictly necessary to perform your own calculations for a catalyst design project. DFT calculations for many reactions have already been published and efforts are undertaken to make these data easily available to the whole catalysis community, even on mobile devices. Yes, there is an app for that![32] But even a non-DFT expert should understand the basic principles that underlie the theoretical results, before using them in a research project. For those readers that have a deeper interest in DFT and want to perform their own calculations, the tutorial-style book *Density Functional Theory – A Practical Introduction* by David S. Sholl and Janice A. Steckel is a highly recommended starting point.[33]

1.2.2 The *d*-Band Model

Computational catalyst screening relies on the prediction of correct trends across different catalysts rather than the prediction of quantitative rates and selectivities for each catalyst. Understanding the origin of the observed trends in terms of the underlying electronic structure can therefore be very helpful

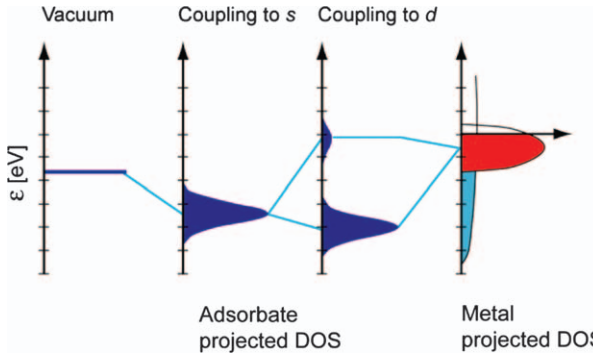


Figure 1.2 Schematic density of states (DOS) illustration of the *d*-band model. The interaction of an adsorbate state with a transition metal can be thought of as a two-step process. The interaction with the broad *s*-band leads to a broadening and downshift of the adsorbate states. The adsorbate states split into bonding and anti-bonding states upon interaction with the narrow transition metal *d*-band. Anti-bonding states that are above the Fermi level remain empty and do not weaken the chemisorption. (Taken from ref. 20)

during the screening process. For transition metal surfaces, trends in reactivity can be very well described and understood in terms of the *d*-band model (Figure 1.2) developed by Hammer and Nørskov.[34,35]

Many of us have likely seen a schematic drawing of the bonding structure in a hydrogen molecule in one of our previous chemistry classes. Upon bond formation, the two atomic orbitals form two new molecular orbitals and they can be distinguished as a bonding and an anti-bonding orbital. In the case of a hydrogen molecule, two electrons can distribute into these orbitals and since each orbital can accommodate up to two electrons, naturally both electrons occupy the lower-lying bonding orbital. The energy that is gained by stabilizing the electrons in this process is the bond energy.

The interactions of adsorbates with transition metals are a bit more complicated but conceptually similar. A transition metal does not possess atomic orbitals, but has a continuous range of available states called a ''band''. In a simplified picture of the band structure in transition metals (Figure 1.2, right) we can imagine that they have a broad *s*-band (turquoise) and a narrow band of localized *d*-electrons (red). Assuming that we can separate the coupling between the adsorbate level and the *s*- and *d*-bands of the transition metal, we can write the chemisorption energy as the sum of both interactions.

$$\Delta E_{\mathrm{Chem}} = \Delta E_s + \Delta E_d \tag{1.4}$$

The interaction of the adsorbate state with the broad *s*-band, ΔE_s, leads to a broadening of the state and a downshift in energy. Since all transition metals have broad and half-filled *s*-bands, the contribution ΔE_s, is approximately the same for all transition metals. In the next step, the broadened molecular state can couple to the narrow *d*-band of the transition metal, which causes a split

into bonding and anti-bonding states, just as in the case of the molecular orbitals of a hydrogen molecule. However, in contrast to the molecular system, where a fixed number of electrons is available to occupy the new orbitals, a transition metal has a large reservoir of electrons. These electrons will fill up all states located below the Fermi level and, consequently, the more anti-bonding states are located below the Fermi level, the weaker the resulting chemisorption bond. For stronger bonds it is desirable that all the anti-bonding states are high in energy and above the Fermi level.

The location of the anti-bonding states relative to the Fermi level and, in turn, the strength of the resulting bond, is largely determined by the position of the *d*-band. Since the *d*-band spans a range of energies, we need to find a convenient way to define the term "position", and therefore introduce the *d*-band center. The *d*-band center, ϵ_d, is the energy-weighted average of the density of *d*-states ρ. It sounds complicated, but it is the same equation that you use to calculate the center of mass (just substitute $r_{\mathrm{COM}} = \epsilon_d$, $r = E$ and $m = \rho$).

$$\epsilon_d = \frac{\int \rho E dE}{\int \rho dE} \tag{1.5}$$

The *d*-band center varies across the transition metals and according to the *d*-band model this will cause a variation in the interaction strength, ΔE_d. If we recall that ΔE_s is constant, the variations in chemisorption energy across transition metals can be attributed to changes in ΔE_d and, thus the *d*-band center. Transition metals with higher-lying *d*-bands have stronger chemisorption properties. Numerous examples supporting the *d*-band model theory exist in the literature (ref. 20 and references 22–37 therein) but, as with most rules, there are some exceptions. In systems including electronegative adsorbates with nearly filled valence shells (*e.g.* OH, F, Cl) on surfaces with almost fully occupied *d*-bands (d^9 and d^{10} metals) there can be a significant repulsive interaction between the transition metal *d*-band and the adsorbate states, which leads to a local reversal of the *d*-band model trend.[36] For computational screening studies, however, we can safely neglect this exception, because it only affects the local fine structure and general trends are preserved.

1.3 Identifying the Right Descriptor Set

If you have ever played the party game "Taboo" you have experienced how hard it can be to describe a word or object without using the five most closely related words listed on the card. On the other hand, if you were allowed to mention only the five forbidden words, your teammates would probably guess the hidden word immediately. Think of these five taboo words as the descriptors that we need to guess the performance of our catalyst. Choosing the right descriptors is the key to a successful descriptor-based catalyst design study and it needs to be done carefully.

The ideal set of descriptors needs to fulfill two conflicting requirements. First and foremost, the descriptor set has to be large enough to enable predictions that

are accurate enough to serve the purpose of the study. For simple reactions, a single descriptor may be sufficient to describe qualitatively the activity trends across different catalyst materials. However, product selectivity, for example, is often more sensitive to the input parameters and may require additional descriptors. If quantitative results are desired, then the number of required descriptors quickly approaches the total number of enthalpy and entropy parameters in a reaction network, which defeats the purpose of reducing the problem complexity by introducing the concept of descriptors. In fact, reducing the complexity is the second most important requirement that the descriptors must meet. The set of descriptors should be as small as needed to capture catalytic trends and enable fast and efficient screening for new catalyst materials.

There is no strict rule for the "correct" choice of descriptors; in fact, for most cases there is more than one set of descriptors and all of them may be equally viable. A descriptor can be any measurable intrinsic quantity of the catalyst (*e.g.* the *d*-band center), but most often descriptors are binding energies of key intermediates inferred from the detailed knowledge of the dominant reaction mechanism and the kinetically relevant steps. This information can be obtained from mechanistic studies using DFT in combination with kinetic measurements and modeling. Alternatively, the existence of scaling relations for surface intermediates and transition states can guide the descriptor selection process. These scaling relations find their physical roots in the *d*-band model and are discussed next.

1.3.1 Scaling Relations for Surface Intermediates

The binding energy of an adsorbed molecule is determined by the number and strength of the chemical bonds that it forms with the surface. To a first approximation we can consider these bonds to form independently of each other and assume that the bond strength only depends on the two types of atom involved in the bond formation. In this simple picture, it would be sufficient to know through which atoms a molecule binds to the catalyst surface in order to predict its adsorption energy. Indeed, it has been shown that the adsorption energies of adsorbates within a family of similar adsorbates can be predicted using this simple idea. The resulting linear scaling relationships can be used to predict unknown adsorption energies from the adsorption energies of related surface species.[37,38] The accuracy of linear scaling relations is generally adequate to predict catalytic trends correctly, but the average errors (0.2–0.3 eV) are too large for quantitative predictions. However in the context of computational screening, in which we are only interested in the relative order between different transition metal catalysts, these errors are acceptable.

1.3.1.1 Hydrogen-containing Molecules

The discussion in this section is fully based on the scaling papers by Abild-Pedersen *et al.*[37] and Fernandez *et al.*,[38] which are highly recommended references on this topic. Each of the four panels in Figure 1.3 shows the

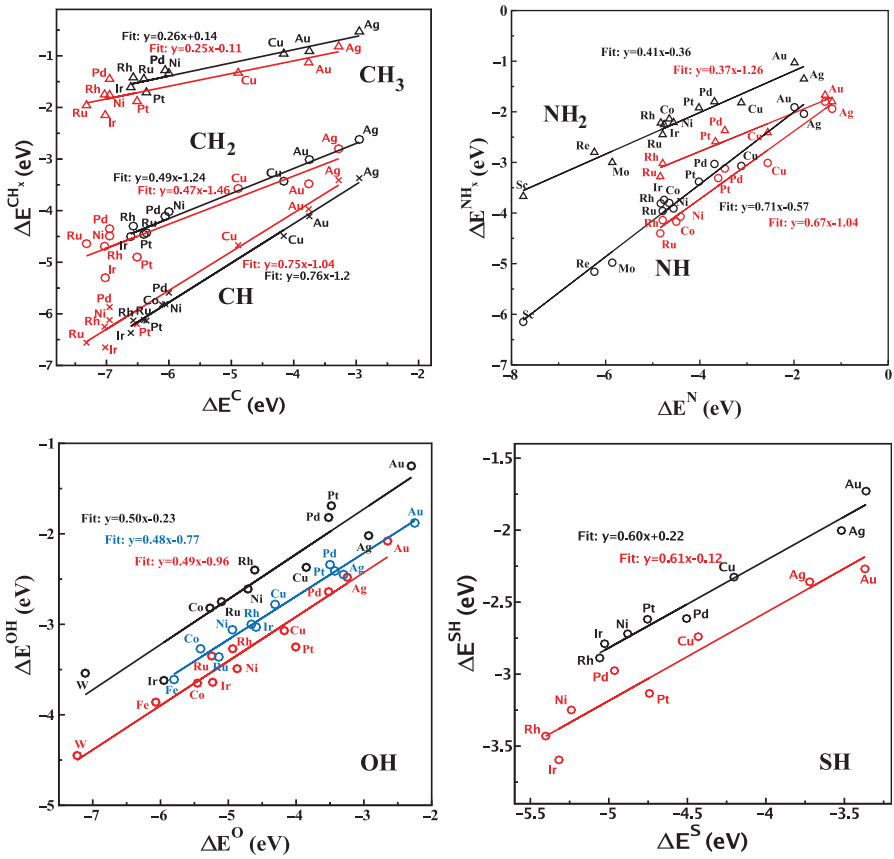


Figure 1.3 Linear scaling relationships on close-packed terraces (black), step sites (red) and the FCC(100) surface (blue) of hydrogen-containing molecules of the type AH_x with A = C, N, O, S.
Reprinted with permission from F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skúlason, T. Bligaard and J. K. Nørskov, *Phys. Rev. Lett.*, 2007, **99**, 016104–016105. Copyright 2007 by the American Physical Society. Available at: http://link.aps.org/doi/10.1103/PhysRevLett.99.016105.[37]

adsorption energies, ΔE^{AH_x}, of the hydrogen-containing intermediates AH_x (A = C, N, O, S) plotted as a function of the adsorption energy, ΔE^{A} of the central atom A for a range of typical transition metal surfaces. The adsorption energies were obtained from periodic DFT calculations on close-packed terraces (black), step sites (red) and, additionally, on the face-centered cubic [FCC(100)] surface (blue) for OH_x. It can be seen that the adsorption energies ΔE^{AH_x} are linearly correlated with ΔE^{A} and given by

$$\Delta E^{AH_x} = \gamma \Delta E^{A} + \xi. \tag{1.6}$$

There is certainly some scatter around the linear scaling lines, but the general trend is correctly captured and the absolute errors are within the 0.2–0.3 eV

range. Upon inspection of Figure 1.3, the observant reader may have noticed that the scaling lines look different for different adsorbates and surface geometries, *i.e.* flat *vs.* stepped surfaces. But is there any physical significance to these differences?

The first observation is that the slope of the scaling relation does not depend on the surface geometry and is constant for any given adsorbate. This is best seen in the case of the OH adsorprtion energy, ΔE^{OH}, which scales with a slope of $\sim 1/2$ with respect to the adsorption energy of oxygen, ΔE^{O}, on all three different surfaces. The same generalization applies to all the other adsorbates and can easily be rationalized using our simple binding picture, where the adsorption energy depends on the number and the type of bonds a molecule forms with the surface. For example, carbon has four valences and likes to form four C–H bonds. If we remove these H atoms stepwise, the unsatisfied carbon valencies start to interact with the catalyst surface and form C–M bonds ('M' denotes a metal atom on the catalyst surface) as depicted in Figure 1.4. Now, as we move from one metal to another, we change the C–M bond strength and we can use the simple bond order method to estimate how this change affects the adsorption energy of the CH_x species.

In general, the scaling slope γ in eqn (1.6) for hydrogen-containing adsorbates AH_x can then be predicted by

$$\gamma(x) = \frac{x_{\max} - x}{x_{\max}} \tag{1.7}$$

where $x_{\max}$ is the maximum number of H atoms that can bind to the central atom A. In particular, $x_{\max} = 4$ for C, $x_{\max} = 3$ for N and $x_{\max} = 2$ for O and S. It should be noted that the linear scaling behavior and characteristic slope of hydrogen-containing adsorbates is not limited to adsorption on metal surfaces. The same linear scaling relations have been discovered for transition metal nitrides, oxides and sulfides, which broadens the range of possible applications of scaling relations far beyond catalysis on transition metal surfaces.[38]

Now that we understand the origin of the slope in the linear scaling relations, we can focus our attention on the intercept ξ. Unfortunately, the intercept cannot be as easily predicted as the slope, but there are two important aspects that need to be mentioned. Using the bond order conservation method to estimate the scaling slope, a single reference calculation can be sufficient to predict adsorption energies on any other catalyst surface. Assume that we calculated ΔE^{AH_x} and ΔE^{A} on a reference catalyst and we know $x_{\max}$ to predict the slope, γ. Then we can use eqn (1.6) to predict ΔE^{AH_x} as a function of ΔE^{A} by simply eliminating the intercept ξ.

$$\Delta E^{AH_x} = \Delta E^{AH_x}_{ref} + \gamma(x) \cdot \left(\Delta E^{A} - \Delta E^{A}_{ref}\right) \tag{1.8}$$

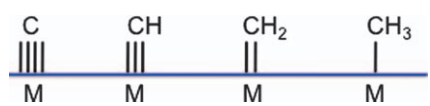


Figure 1.4 Simplified bonding scheme for adsorbed CH_x species.

Structure sensitivity is the other factor that affects the intercept, ξ, which is clearly demonstrated by the offset between the parallel scaling lines for each adsorbate in Figure 1.3. Although structure sensitivity plays an important role in catalysis, its effect on the scaling behavior of adsorbates is not discussed here and the interested reader is referred to the detailed tutorial review by Prof. Nørskov instead.[39]

1.3.1.2 Extension to More Complex Surface Intermediates

The special case of hydrogen-containing AH_x species impressively demonstrates the relationship between the binding energies of similar adsorbates, but in the majority of catalytic reactions more complex surface intermediates are encountered. In the process of establishing linear scaling relations for complex species, we may intuitively start by correlating their adsorption energies to the functional groups contained in the species or simply to the adsorption energies of the atoms through which the molecule binds. This approach is especially useful when the adsorption geometry of the involved surface species is known.

Let us take the bond conservation idea introduced in the previous section for AH_x and apply it to CH_x–NH_y with $x,y = 0$–2 surface intermediates. These intermediates occur, for example, during the Degussa and Andrussow process for HCN synthesis from CH_4 and NH_3.[17,18,40,41] We may assume that the C and N atoms are connected through a single bond and that the CH_x–NH_y intermediates bind to the surface equally through the C atom and the N atom. The adsorption energy of CH_x–NH_y should then scale linearly as in

$$\Delta E^{\mathrm{CH}_x-\mathrm{NH}_y} = \gamma \cdot \left(\Delta E^{\mathrm{CH}_x} + \Delta E^{\mathrm{NH}_y}\right) + \xi \tag{1.9}$$

where $\Delta E^{\mathrm{CH}_x} = \frac{4-1-x}{4}\Delta E^{\mathrm{C}} + \xi_{\mathrm{C}}$ and $\Delta E^{\mathrm{NH}_x} = \frac{3-1-x}{3}\Delta E^{\mathrm{N}} + \xi_{\mathrm{N}}$. Note the '–1' term indicating that one valency of C and N is consumed by the C–N bond formation and is no longer available. If we now substitute the expressions for $\Delta E^{\mathrm{CH}x}$ and $\Delta E^{\mathrm{NH}y}$ into eqn (1.9) and collect the intercepts, we arrive at our final expression for $\Delta E^{(\mathrm{CH}_x-\mathrm{NH}_y)}$.

$$\Delta E^{\mathrm{CH}_x-\mathrm{NH}_y} = \gamma_{\mathrm{CH}_x-\mathrm{NH}_y}\left(\frac{4-1-x}{4}\Delta E^{\mathrm{C}} + \frac{3-1-x}{3}\Delta E^{\mathrm{N}}\right) + \xi_{\mathrm{CH}_x-\mathrm{NH}_y} \tag{1.10}$$

The comparison of CH_x–NH_y adsorption energies estimated from eqn (1.10) with DFT reference calculations given in Figure 1.5(a) shows that this approach adequately reproduces the correct trend, although several outliers with errors of *ca.* 0.5 eV exist. A similar strategy based on bond order conservation was used by Jones *et al.* to derive scaling relations for C_2 hydrocarbons on transition metal surfaces.[42]

Even without deriving scaling relations on paper as done for CH_x–NH_y, one can simply try different scaling relations and evaluate them on the basis of their mean absolute error and/or R^2 value. Figure 1.5(b) shows a very good scaling relation for C_2H_2 and C_2H_4 with methyl (CH_3), and many more such examples exist. Hence, if one has access to an adsorption energy database for many

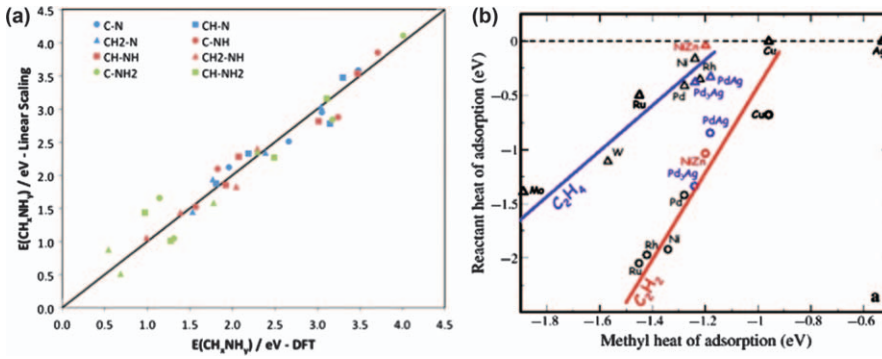


Figure 1.5 Linear scaling relations for complex adsorbates. (a) Parity plot of CH_x–NH_y adsorption energies predicted from scaling and the DFT reference calculations. (b) Linear scaling of C_2H_4 and C_2H_6 with methyl (CH_3) from F. Studt *et al.*, *Science*, **320**, 1320–1322 reprinted with permission from AAAS.[14]

different surface species, it is possible to bypass the bond order conservation idea and just rely on statistical tools. Using this approach, it should be possible to derive a suitable descriptor for any adsorbate, but the physics behind the relationship may not be obvious.

1.3.2 Scaling Relations for Transition States: The Brønsted–Evans–Polanyi Relationship

Transition states are often considered special because they have a very short lifetime. In terms of their adsorption properties, however, they are not different from any other surface species and obey the same physical principles. It is therefore not surprising that scaling laws for transition states exist in the same way as they do for other surface intermediates. The truly remarkable aspect is that transition state scaling for a large number of reactions is universal and can be described by a single linear relationship![43–45] However, before we continue to discuss the beauty of transition state scaling it is important to clarify several technical terms used in the following discussion.

Transition state scaling (TSS) is often used interchangeably with the term Brønsted–Evans–Polanyi or BEP relationship.[46,47] Although both terms refer to the same concept, there is a distinction between them. Transition state scaling refers to a relation between the transition state energy E_{TS} and the energy of either the initial state (E_{IS}) or final state (E_{FS}) of a reaction. The transition state energy, E_{TS}, is not equal to the activation barrier E_a of a reaction, but it can be used to calculate $E_a = E_{TS} - E_{IS}$. A BEP relation, on the other hand, is a relation between the activation energy E_a and the heat of a reaction ΔE, which is defined as $\Delta E = E_{FS} - E_{IS}$. All energy quantities are indicated on a schematic potential energy surface (PES) in Figure 1.6.

A direct comparison between a TSS relation and a BEP relation for an exhaustive data set of (de)hydrogenation reactions over transition metals is

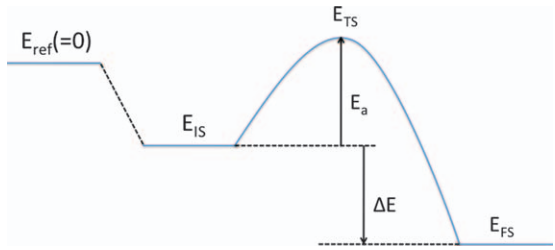


Figure 1.6 A schematic potential energy surface for an arbitrary surface reaction. The nomenclature is E_{ref} = reference energy chosen to be zero, E_{IS} = initial state energy, E_{TS} = transition state energy, E_{FS} = final state energy, E_a = activation energy barrier, ΔE = energy change of the reaction.

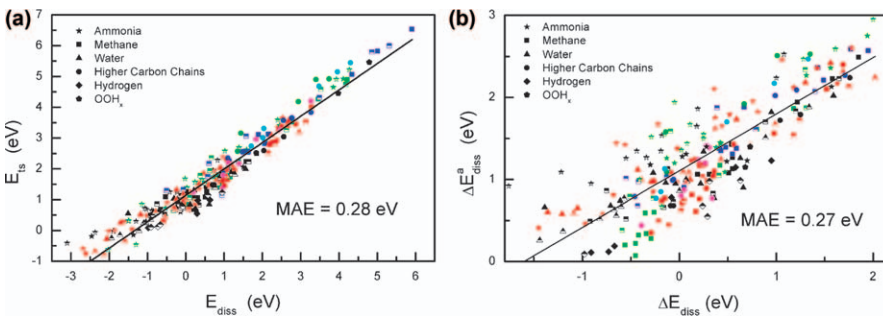


Figure 1.7 Transition state scaling (a) and BEP relation (b) for (de-)hydrogenation reactions.
(Taken from ref. 45)

shown in Figure 1.7.[45] Since all reactions were formulated as dissociation reactions, E_{diss} was used to denote the final state energy E_{FS}; and ΔE_{diss} represents the energy change of the surface reaction, ΔE. Although identical raw data have been used to generate the plots, the transition state scaling in Figure 1.7(a) appears to be much better than the BEP relation in Figure 1.7(b). This initial impression is somewhat deceptive and a comparison of the mean absolute error (MAE) in both plots tells us that the quality, in terms of absolute errors, of both relations is almost identical. The difference lies in the different energy scales used on the *x*- and *y*-axes of the plots, and the BEP graph can be loosely interpreted as a zoomed-in version of the transition state scaling graph.

To convince ourselves that, for most reactions, both graphs contain essentially the same information, we will derive a mathematical conversion of one graph into the other. We start by postulating that the initial and final state of the reaction is described by linear scaling relations that depend on the descriptors E_1 and E_2.

$$E_{IS} = \gamma_1 E_1 + \xi_1 \tag{1.11}$$

$$E_{FS} = \gamma_2 E_2 + \xi_2 \tag{1.12}$$

Using transition state scaling we can find E_{TS} as a function of E_{FS}.

$$E_{TS} = \alpha_1 E_{FS} + \beta_1 \tag{1.13}$$

The quantity that is of practical interest for us is not E_{TS} directly, but the activation energy barrier E_a, which is obtained in TSS from

$$E_a^{TSS} = E_{TS} - E_{IS} = \alpha_1 E_{FS} + \beta_1 - E_{IS} = \alpha_1(\gamma_2 E_2 + \xi_2) + \beta_1 - (\gamma_1 E_1 + \xi_1)$$

$$E_a^{TSS} = \alpha_1\gamma_2 E_2 - \gamma_1 E_1 + \alpha_1\xi_2 + \beta_1 - \xi_1. \tag{1.14}$$

The last three terms are constants and may be summed and replaced with β^{TSS}, which then yields the final expression

$$E_a^{TSS} = \alpha_1\gamma_2 E_2 - \gamma_1 E_1 + \beta^{TSS} \tag{1.15}$$

Alternatively, E_a can be obtained from the BEP relation.

$$E_a^{BEP} = \alpha_2 \Delta E + \beta_2 = \alpha_2(E_{FS} - E_{IS}) + \beta_2 = \alpha_2(\gamma_2 E_2 + \xi_2 - \gamma_1 E_1 + \xi_1) + \beta_2$$

$$E_a^{BEP} = \alpha_2\gamma_2 E_2 - \alpha_2\gamma_1 E_1 + \alpha_2(\xi_2 - \xi_1) + \beta_2$$

$$E_a^{BEP} = \alpha_2\gamma_2 E_2 - \alpha_2\gamma_1 E_1 + \beta^{BEP} \tag{1.16}$$

Although eqns (1.15) and (1.16) look similar, they are not identical if the descriptors E_1 and E_2 are considered to be independent variables. However, for the vast majority of reactions the descriptors are either identical, linearly correlated or at least exhibit a nearly linearly correlated behavior. If we assume a linear correlation and substitute

$$E_2 = \gamma_3 E_1 + \xi_3 \tag{1.17}$$

into eqns (1.15) and (1.16), we can eliminate E_2 and get

$$E_a^{TSS} = (\alpha_1\gamma_2\gamma_3 - \gamma_1)E_1 + \tilde{\beta}^{TSS} \tag{1.18}$$

$$E_a^{BEP} = (\alpha_2\gamma_2\gamma_3 - \alpha_2\gamma_1)E_1 + \tilde{\beta}^{BEP} \tag{1.19}$$

where the constant terms have been included in $\tilde{\beta}^{TSS}$ and $\tilde{\beta}^{BEP}$. By equating E_a^{TSS} and E_a^{BEP} from eqns (1.18) and (1.19) it can easily be seen that they are identical when $\tilde{\beta}^{TSS} = \tilde{\beta}^{BEP}$ and $\alpha_2 = \frac{\alpha_1\gamma_2\gamma_3 - \gamma_1}{\gamma_2\gamma_3 - \gamma_1}$. This shows that, as long as the descriptors E_1 and E_2 used for the initial and final state, respectively, are linearly correlated, it does not matter whether a TSS or BEP relation is used.

Now that we have demonstrated that the choice between TSS and BEP is mostly a personal preference, the author wants to motivate his preference toward BEP over TSS relations. Let us assume that a transition state can be characterized as initial state like, final state like, or anything in between. We define ω such that $\omega = 0$ corresponds to a transition state that is initial state like, while $\omega = 1$ corresponds to a final state-like transition state. The energy of the transition state, E_{TS}, would then scale with E_{IS} and E_{FS} weighted by the

transition state character ω. With β as an arbitrary energy offset, the energy E_{TS} can then be expressed as

$$\begin{aligned} E_{TS} &= (1-\omega)E_{IS} + \omega E_{FS} + \beta = \ \omega(E_{FS} - E_{IS}) + E_{IS} + \beta \\ E_{TS} &= \ \omega\Delta E + E_{IS} + \beta \end{aligned} \tag{1.20}$$

Rewriting eqn (1.20) for the calculation of an activation barrier leads to a BEP relation relating E_a with ΔE and the slope of the BEP relation, α, can be interpreted as the transition state character, ω.

$$E_a = E_{TS} - E_{IS} = \ \omega\Delta E + \beta \tag{1.21}$$

In addition to the physical interpretation of the BEP slope, using a BEP relation has other practical advantages. For the calculation of rate and equilibrium constants E_a and ΔE are needed, which can be directly obtained from a BEP relation. In more advanced kinetic models, the effect of surface coverage on the stability of the surface intermediates may also be explicitly included, but including such an explicit coverage dependency on activation energies is not trivial.[48,49] Assume that we have used DFT to calculate E_a for a surface reaction at low coverage and we are able to predict the initial and final state energies for a high-coverage scenario. Performing additional DFT calculation for E_a at various higher surface coverages is prohibitively time consuming and an approximate method is desired to estimate the effect of coverage on E_a. In this case, we can apply the same idea used to derive eqn (1.8): we eliminate the constant β by using a reference calculation, and we can find E_a.

$$E_a = E_a^{ref} + \omega\left(\Delta E - \Delta E^{ref}\right) \tag{1.22}$$

Now that we can fully appreciate the meaning of TSS and BEP relations, we can discuss their most remarkable property, the fact that these relations are almost universally found for dissociation reactions.[44,45] This is referred to as the Universality Principle in Heterogeneous Catalysis and can be explained by the geometric similarities of the transition states with their respective dissociated states.[43] In Figure 1.7 it has already been shown that the universality principle holds for (de)hydrogenation reactions, and Figure 1.8(a) demonstrates that the principle also applies to many other dissociation reactions. However, the MAE indicated in Figure 1.8(a) of 0.35 eV is larger than the typically accepted MAE (0.3 eV) and the individual TSS/BEP relations for smaller families of reactions listed in refs. 44,45 should be used if possible. For preliminary screening studies, however, it is perfectly acceptable to assume universal TSS/BEP behavior and refine the results later, if needed.

At this point it should be stressed that linear TSS/BEP relations exist for many reactions on a large range of catalyst surfaces, but they are not as "universal" as they have been introduced so far. Even for simple di-atomic dissociation reactions (N_2, NO, O_2) on metal-substituted La-perovskites a significant deviation from the universal TSS/BEP line has been observed.[50] The deviation can clearly be seen in Figure 1.8(b) as one moves to more reactive

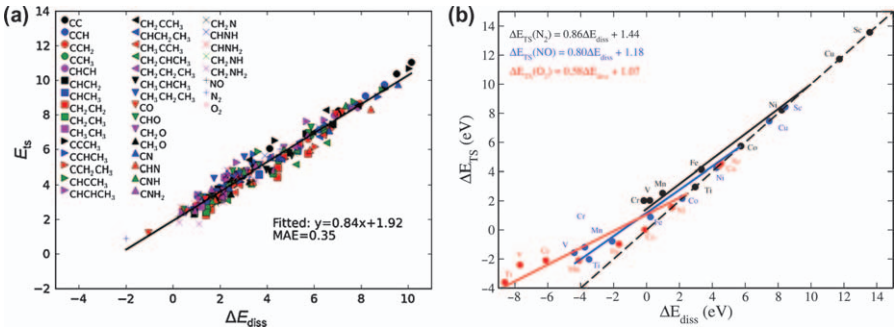


Figure 1.8 (a) Universal TSS relation for C–C, C–O, C–N, N–O, N–N, and O–O dissociation reactions. With kind permission from Springer Science and Business Media,[44] S. Wang *et al.*, *Cat. Lett.*, 2011, **141**, 370–373. (b) Deviations from the "universal" TSS behavior are observed when the nature of the transition state switches between initial and final states alike. From Ref. 50 with permission from A. Vojvodic *et al.*, *J. Chem. Phys.*, 2011, **134**, 244509. Copyright 2011, AIP Publishing LLC.

surfaces (more negative ΔE_{diss}) and can be explained by a continuous change from a late transition state (final state like) to an early transition state on more reactive surfaces. The same phenomena can be seen on highly reactive metallic surfaces, but all transition metals that are typically used as catalysts fall into the linear region of the TSS/BEP line and the relation holds quite well in this range. Nevertheless, care must be taken to choose the appropriate TSS/BEP relation for each reaction and even more attention must be paid when these relations are used to extrapolate beyond the parameter range for which they were obtained.

In summary, scaling and BEP relations are very versatile, broadly applicable, and allow for a significant reduction of the number of energetic parameters to a small set of catalytic descriptors. Given the typical error bars of these relations the results are not meant to be taken quantitatively, but if the scaling and BEP relations are carefully chosen and correctly applied, then catalytic activity and selectivity trends over a broad range of materials can very well be obtained.

1.4 The Sabatier Principle and the Volcano Curve

For most things in life, there is an optimal concentration, amount or frequency in which they should occur; catalysis is no exception. The Nobel Prize-winning French chemist Paul Sabatier qualitatively described the adsorption properties of the optimal catalyst for a reaction as not too weak, such that the reactants don't bind, and not too strong, which would lead to surface poisoning.[51] This statement is known today as the Sabatier principle of catalysis. The principle is easily visualized in volcano-shaped curves that show the catalytic activity as a function of a catalytic descriptor. An example of such a volcano curve was already introduced in Figure 1.1(a) and a schematic representation is shown in Figure 1.9. The top of the volcano corresponds to the highest activity and the descriptor value that is "just right". Finding the optimal value of the

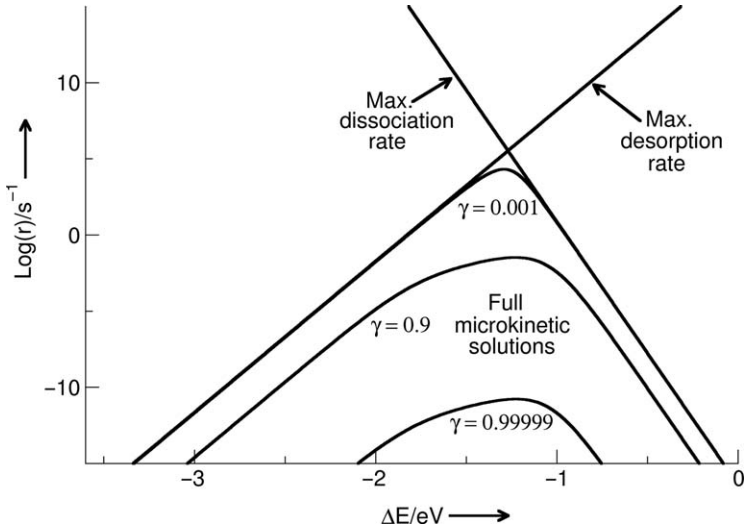


Figure 1.9 Comparison of the volcano curves obtained from a Sabatier analysis and full microkinetic models at different approaches to equilibrium γ. Reprinted from T. Bligaard, *et al.*, *J. Catal.*, 2004, **224**, 206–217 with permission from Elsevier.[21]

descriptor(s) is the central and by far the most critical point in any computational catalyst screening study. Once the optimal value is known, it requires only a simple database search to screen for materials with the desired property. The other features of the volcano, including the activity at the volcano peak and the slopes at which the activity decreases along both sides of the maximum, are of secondary importance. Since we care primarily about the descriptor value at the volcano peak, plus or minus a few tenths of an eV, it is not always necessary to develop a full-blown microkinetic model. Instead, a shortcut method, the Sabatier analysis, can be used to predict upper bounds for the reaction rate. Sabatier analyses have been very successfully used to identify catalytic trends for CO oxidation and NO decomposition.[52,53] As shown in Figure 1.9, the Sabatier rate reproduces the shape of the volcano curve obtained from a microkinetic model well enough to provide an estimate of the optimal descriptor value, and it is significantly easier and faster to implement.

1.4.1 Sabatier Analysis

The Sabatier analysis for surface reactions was first introduced by Bligaard *et al.*[21] and uses a Sabatier rate that is obtained by following this fairly simple recipe:

1. Write the corresponding rate expressions for all reaction steps in the forward direction as a function of the approach to equilibrium $\gamma_i = \Pi_n \nu_n / K_i$, where K_i is the equilibrium constant for step i, a_n is

the activity of species n and ν_n is the corresponding stoichiometric coefficient.

2. For each reaction step in the model, assume optimal surface coverages for the forward reaction.
3. From the overall approach to equilibrium, calculate the approach to equilibrium of step i under the assumption that all other steps are quasi-equilibrated.
4. Use TSS/BEP relations to calculate the forward rate constants k_i and the rate. This rate represents an upper bound on the actual rate of step i and will be denoted as $r_i^{\max}$.
5. The Sabatier rate is the slowest rate of all maximum rates in the mechanism: $r_{\text{Sabatier}} = \min\{r_i^{\max}\}$.

Let us illustrate the procedure using a generic reaction mechanism with two steps as an example. The two reaction steps are a dissociation step, step (1), and a reactive desorption step, step (2):

(1) $A_2 + 2* \rightleftharpoons 2\ A^*$
(2) $A^* + B \rightleftharpoons AB + *$

where the asterisk '*' denotes a free surface site and A* is the adsorbed species A. With θ denoting the surface coverage and γ the approach to equilibrium, the reaction rates in the forward direction are

$$r_1 = 2k_1 P_{\text{A}_2} \theta_*^2 (1 - \gamma_1) \tag{1.23}$$

$$r_2 = k_2 P_{\text{B}} \theta_{\text{A}} (1 - \gamma_2) \tag{1.24}$$

The optimal surface coverage for step (1) is a completely empty surface ($\theta_* = 1$) and for step (2) the surface should be fully covered with A ($\theta_{\text{A}} = 1$). The overall approach to equilibrium γ is closely related to the equilibrium constant and is defined as

$$\gamma = \frac{P_{\text{AB}}^2}{K_{\text{eq}} P_{\text{A}_2} P_{\text{B}}^2} \quad \text{with} \quad \gamma = \gamma_1 \gamma_2^2 \tag{1.25}$$

The allowed range for γ is $0 \leq \gamma \leq 1$, where $\gamma = 1$ indicates a quasi-equilibrated reaction. From eqn (1.25) we can estimate γ_1 and γ_2 under the assumption that the other step is quasi-equilibrated and obtain

$$\gamma_1 = \gamma \text{ and } \gamma_2 = \sqrt{\gamma} \tag{1.26}$$

Assuming that we know k_1 and k_2 we can now write down the upper bounds for r_1 and r_2 and find the Sabatier rate r_{Sabatier}.

$$r_1^{\max} = 2k_1 P_{\text{A}_2} (1 - \gamma) \tag{1.27}$$

$$r_2^{\max} = k_2 P_{\text{B}} (1 - \sqrt{\gamma}) \tag{1.28}$$

$$r_{\text{Sabatier}} = \min\{r_1^{\max}, r_2^{\max}\} \tag{1.29}$$

It is straightforward to apply the same analysis to more complicated reaction mechanisms with a larger number of steps. The resulting volcano would have more than two sides, but the maximum and the corresponding descriptor value can still be found just as easily as in the case presented here. For reaction conditions where the overall approach to equilibrium approaches unity ($\gamma \to 1$), the simple Sabatier analysis can significantly deviate from a full microkinetic model solution, but at low conversions far away from equilibrium where $\gamma \to 0$, the agreement is generally good.

For the sake of completeness, the Sabatier–Gibbs analysis is also briefly mentioned here, although it is not particularly useful for general problems. The Sabatier–Gibbs analysis can only be applied to serial reaction mechanisms that are very simple, and it improves the unrealistic assumption of optimal surface coverages for each step used in the Sabatier analysis. A stricter bound on the coverages can be imposed by considering thermodynamic limits in terms of the approach to equilibrium, but the procedure is somewhat cumbersome and in general one is better off using a microkinetic model if the simple Sabatier analysis fails. For additional information the reader is referred to refs. 52,54.

1.5 Sabatier Analysis in Practice

1.5.1 First Example: Ammonia Synthesis

Now it is finally time to combine all the concepts introduced so far and perform our first real catalyst design. The reaction we will use for the design study is one of the most important heterogeneously catalyzed reactions, namely ammonia (NH_3) synthesis from nitrogen (N_2) and hydrogen (H_2).

$$N_2 + 3\ H_2 \rightleftharpoons 2\ NH_3 \qquad \Delta H^\circ = -45.9\ \text{kJ per mol}(NH_3)$$

This reaction is performed industrially using the Haber–Bosch process over iron (Fe) or ruthenium (Ru) catalysts at high temperatures (>400 °C) and high pressures (>100 bar). Using state-of-the-art surface science[55] and theoretical techniques,[56] the detailed reaction mechanism has been elucidated and a complete potential energy surface (PES) on terrace and step sites of Ru is shown in Figure 1.10.

Upon inspection of the PES for NH_3 synthesis one can simplify the mechanism to two main reaction steps, the dissociative adsorption of N_2 followed by its hydrogenation to NH_3. An extremely simplified reaction mechanism can therefore be written as

(1) $N_2 + 2* \rightleftharpoons 2\ N^*$ (dissociative adsorption)
(2) $N^* + \frac{3}{2} H_2 \rightleftharpoons NH_3$ (recombinative desorption)

This mechanism neglects a number of elementary steps and does not include the possibility of adsorbed H atoms or NH_x species, but it is sufficient to illustrate the Sabatier principle. A catalyst that associates strongly with N will easily

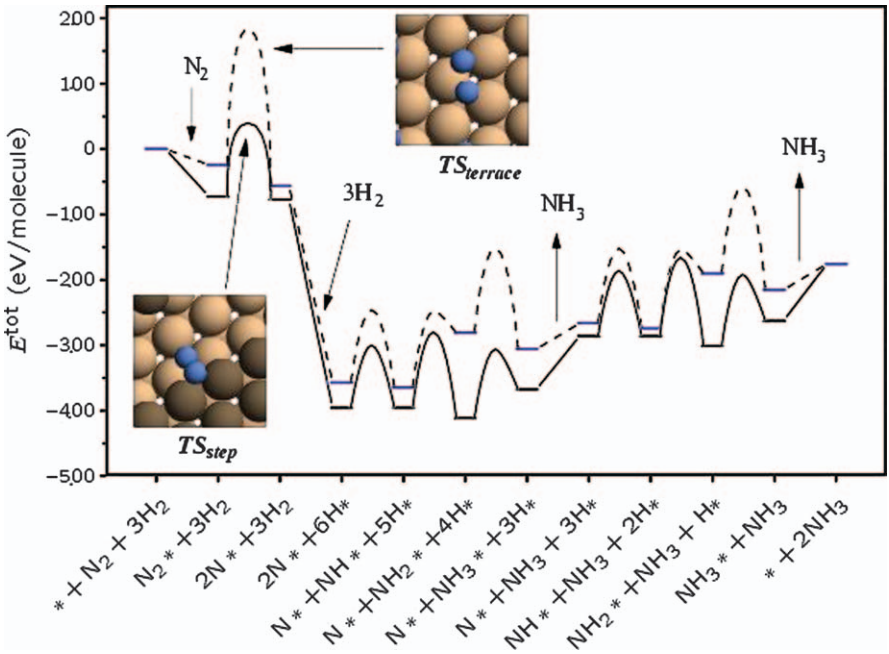


Figure 1.10 Potential energy surface for ammonia synthesis on terrace (dashed) and stepped (solid) sites on Ru.
K. Honkala *et al.*, *Science*, 2005, **307**, 555–558 reprinted with the permission of AAAS.[56]

dissociate N_2, but the desorption step will be very difficult (note that in the desorption step only the energy of N* is affected by the catalyst because H_2 and NH_3 are gas phase species). In contrast, a catalyst interacting weakly with N allows for easy desorption, but the barrier for N_2 dissociation will be large. Hence, there is an optimal binding energy of N*, E_N, at which the dissociative adsorption of N_2 and NH_3 desorption are optimized simultaneously. We will now try to find an estimate of E_N, by performing a Sabatier analysis for typical NH_3 synthesis conditions.

The simplified NH_3 synthesis mechanism is almost identical to the generic mechanism used in the introduction of the Sabatier analysis in Section 1.4. If we substitute $A_2 = N_2$ and $B = 3/2H_2$, we can directly use the previously obtained results and apply them to the current problem. This gives us the following two maximum rates for each step, where the stoichiometric factor of 3/2 for H_2 is included as the exponent to P_{H_2} in the expression for r_2^{max}.

$$r_1^{max} = 2k_1 P_{N_2}(1-\gamma) \tag{1.30}$$

$$r_2^{max} = k_2 P_{H_2}^{\frac{3}{2}}(1-\sqrt{\gamma}) \tag{1.31}$$

For the reaction conditions we may assume that we have a stoichiometric feed composition ($H_2 : N_2 = 3 : 1$), the total pressure is $P = 100$ bar and the

temperature is $T = 673$ K. Using this information we can calculate the partial pressures of hydrogen and nitrogen as $P_{H_2} = 75$ bar and $P_{N2} = 25$ bar. Further, we assume that the reaction is run at the limit of very low conversion where $\gamma \to 0$. This leaves us with only the two unknown rate constants k_1 and k_2. We can obtain the rate constants, k_i, from the standard Arrhenius expression, which depends on the activation energy, $E_{a,i}$, and the pre-exponential factor, ν_i, for each step.

$$k_i = \upsilon_i \cdot \exp\left(\frac{-E_{a,i}}{k_B T}\right) \tag{1.32}$$

The pre-exponential factor ν_i can be calculated using transition state theory from the entropy change between the initial state and the transition state of the reaction, $\Delta S_i^{\ddagger}$.

$$\upsilon_i = \frac{k_B T}{h} \cdot \exp\left(\frac{\Delta S_i^{\ddagger}}{k_B}\right) \tag{1.33}$$

Next, we obtain the activation barrier estimates from published scaling relations, beginning with N_2 dissociation. The dissociation of diatomic molecules was the first group of heterogeneously catalyzed reactions for which a universal BEP relation was identified and the values of $\alpha = 0.87$ and $\beta = 1.34$ eV were reported for stepped surfaces.[43]

$$E_{a,1} = 0.87 \cdot \Delta E_1 + 1.34\ eV \tag{1.34}$$

The second step in the simplified mechanism lumps several hydrogenation steps into one step, therefore it seems appropriate to use the averaged BEP parameters for NH_3, NH_2, and NH dehydrogenation reactions found in Table 2 of ref. 44. Be careful here! The BEP parameters $\alpha = 0.61$ and $\beta = 1.43$ eV are derived for the reverse reaction of our step (2), so we will have to make sure to use them accordingly. The estimated reverse activation energy $E_{a,2}^{rev}$ is then given by

$$E_{a,2}^{rev} = 0.61 \cdot (-\Delta E_2) + 1.43\ \text{eV} \tag{1.35}$$

and we find the forward barrier by using the fact that $\Delta E_2 = E_{a,2}^{fwd} - E_{a,2}^{rev}$. The energy change of step (2), ΔE_2, depends linearly on ΔE_1 owing to the thermodynamic constraint that $\frac{1}{2}\Delta E_1 + \Delta E_2 = \Delta E_r$, where ΔE_r is the overall energy change of the reaction. So far we have exclusively used the electronic ground-state energy, E, at $T = 0$ K in all equations, which is the primary output of DFT calculations, but from thermodynamics we know that free energy changes are the correct quantities to use in the calculation of rate and equilibrium constants. We assume for now that the DFT-calculated electronic energy is approximately equal to the enthalpy and that the entropy change can be calculated separately. This can be justified by the fact

that the electronic energy changes are significantly larger than all other contributions to the enthalpy. Thus, for the present problem we can say that $\Delta E_r \approx \Delta H^\circ = -45.9$ kJ mol^{-1} $= -0.48$ eV. Putting this all together we get the desired BEP relation for $E_{a,2}$.

$$E_{a,2} = \Delta E_2 + E_{a,2}^{rev} = \Delta E_2 + 0.61 \cdot (-\Delta E_2) + 1.43 \text{ eV} = 0.39 \cdot \Delta E_2 + 1.43 \text{ eV}$$

$$= 0.39 \cdot \left(\Delta E_r - \frac{1}{2}\Delta E_1\right) + 1.43 \text{ eV} = -0.195 \cdot \Delta E_1 + 1.24 \text{ eV} \qquad (1.36)$$

By choosing these particular BEP relations to estimate the activation barriers for steps (1) and (2), we have implicitly determined the reactivity descriptor for this reaction. Both barriers depend only on ΔE_1, which is the dissociative chemisorption energy of nitrogen, ΔE_N.

The last piece of missing information comprises the entropy changes between initial and transition states, which are needed to calculate the pre-exponential factors. The entropy values of adsorbed species and transition states depend only weakly on the transition metal surface. Therefore, we may assume the entropy to be independent of the catalyst. For a rough approximation, we can further assume that an adsorbed species or transition state has lost all degrees of freedom (it is completely locked in place on the surface), and the entropy is therefore zero. The gas-phase entropies of N_2 and H_2 at standard state are $S^\circ(N_2) = 192.77$ J mol^{-1} K^{-1} and $S^\circ(H_2)$ 130.68 J mol^{-1} K^{-1}, respectively, and we can now estimate the entropy changes between the initial states and the transition state as $\Delta S_1^\ddagger = -192.77$ J mol^{-1} K^{-1} and $\Delta S_2^\ddagger = -130.68$ J mol^{-1} K^{-1}. For the first step it is assumed that N_2 loses all its gas-phase entropy upon entering the transition state and in the second step, the entropy loss of one H_2 gas-phase molecule is considered when it reacts with N* on the surface. Why did we consider the entropy loss of only one H_2 molecule and not $3/2H_2$ as the stoichiometric coefficient in the *lumped* equation would indicate? Well, the fact that step (2) is a *lumped* step and not an elementary reaction step makes it almost impossible to guess the correct entropy change, but we know that it is a sequence of hydrogen adsorption and hydrogenation steps and for each H_2 molecule adsorbing onto the surface –130.68 J mol^{-1} K^{-1}of entropy are lost. When mechanisms with lumped reaction steps are analyzed, problems with estimating energy barriers and entropy changes will frequently occur, so it is highly recommended to formulate mechanisms that are comprised of only elementary reaction steps! Still, for this case, from eqn (1.33) and the entropy assumptions outlined above, the pre-exponential factors are calculated as $\nu_1 = 1.20 \times 10^3$ s^{-1} and $\nu_2 = 2.10 \times 10^6$ s^{-1}.

At this point the reader may want to open her/his favorite math and graphing software (Excel would suffice), enter the eqns (1.30)–(1.32), (1.34), and (1.36), and calculate the maximum N_2 dissociation rate r_1^{max} and maximum desorption rate r_2^{max} as a function of $\Delta E_1 = \Delta E_N$ and plot the results. The graph should look like the one given in Figure 1.11. The optimal catalyst for ammonia synthesis is found for a descriptor value of $\Delta E_N \approx -0.95$ eV and has a

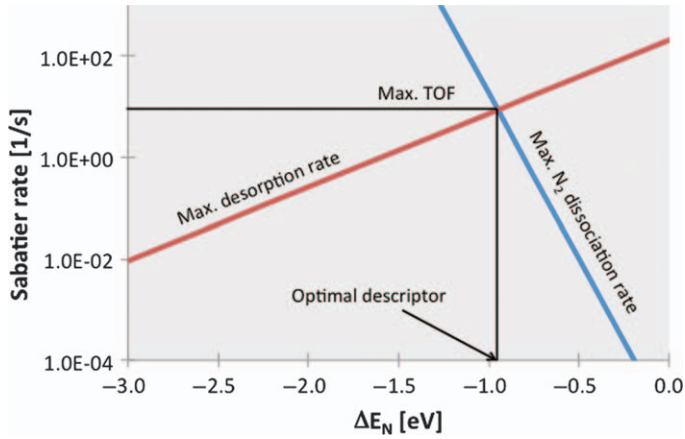


Figure 1.11 Sabatier volcano for NH_3 synthesis.

Table 1.1 Excerpt of the periodic table with dissociative chemisorption energies of N_2 (ΔE_N) given in eV. Reprinted from T. Bligaard, *et al.*, *J. Catal.*, 2004, **224**, 206–217 with permission from Elsevier.[21]

Cr	Mn	Fe	Co	Ni	Cu
–	–	– 1.27	– 0.38	– 0.10	2.88
Mo	Tc	Ru	Rh	Pd	Ag
– 2.76	–	– 0.84	– 0.70	1.78	5.86
W	Re	Os	Ir	Pt	Au
– 4.33	–	–	– 0.59	1.37	5.89

predicted turn over frequency (TOF) of 8.7 s^{-1}. It is usually a lucky coincidence if the estimated TOF is close to an experimentally measured value in such a crude Sabatier analysis, but the location of the maximum with respect to the descriptor value in the volcano curve is typically reproduced within acceptable error bars. With a target value for the catalytic descriptor we can finally proceed to the last step of the example problem and screen for the best catalyst. For this purpose, let us examine the "database" of dissociative chemisorption energies, ΔE_N, of several transition metals listed in Table 1.1, and we find that Fe and Ru are closest to the optimal value of –0.95 eV! Considering the number of assumptions and simplifications that were required to perform the Sabatier analysis, it seems surprising that the two most active catalysts for ammonia synthesis were correctly identified. This certainly emphasizes the strength and robustness of the method. The analysis is helped by the fact that ΔE_N spans a range from *ca.* –4 to +6 eV and neighboring transition metals are often separated by 0.5 eV or more.

In summary, this example of computational screening for an NH_3 synthesis catalyst has demonstrated how BEP relations are applied to reduce the number of parameters to a single descriptor, ΔE_N, and how to perform a Sabatier analysis. As a result we obtained the Sabatier volcano shown in Figure 1.12

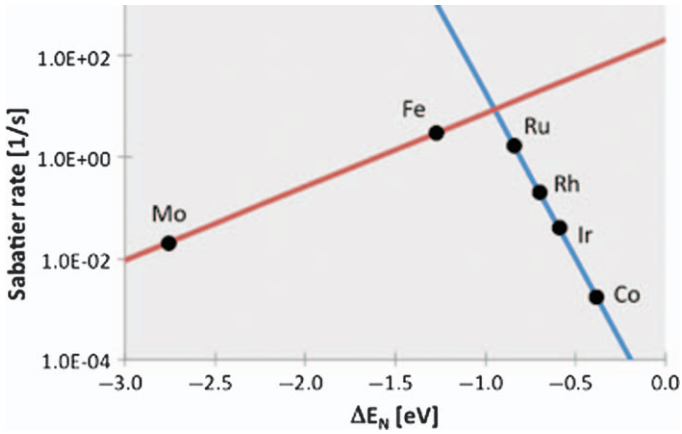


Figure 1.12 Sabatier volcano for NH_3 synthesis with the positions of transition metals labeled on the graph.

from which two suitable catalysts for the reaction can be identified (Fe and Ru). The example also introduced approximations that can be applied to obtain a simple estimate for surface entropies, which will be improved in Section 1.6.2.

1.5.2 Second Example: CO Oxidation

The Sabatier analysis for NH_3 synthesis presented in the previous section captures all the important ideas needed for a computational catalyst design project, but it is very idealized. A somewhat more complicated example is CO oxidation, for which we can derive a volcano surface as a function of two independent descriptors. This analysis was originally completed by Falsig *et al.* and resulted in the prediction of the most active catalysts for high (Pt, Pd) and low (Au) temperature CO oxidation.[53] Here, we discuss only the high temperature CO oxidation with the reaction conditions $T = 600\,K$, $P_{O_2} = 0.33$ bar, $P_{CO} = 0.67$ bar. The low temperature reaction can be studied analogously. For the CO oxidation reaction we start with the following reaction mechanism.

(1) $CO + * \rightleftharpoons CO^*$ (quasi-equilibrated)
(2) $O_2 + * \rightleftarrows O_2^*$ (quasi-equilibrated)
(3) $O_2^* + * \rightleftharpoons 2\,O^*$
(4) $CO^* + O^* \rightleftharpoons CO_2 + 2*$

Following Falsig *et al.*, we assume that the molecular adsorption of CO and O_2 is fast and can be considered as quasi-equilibrated. The two competing steps in the reaction are then the dissociation of O_2* and the reaction between CO* and O* to form CO_2. Assuming very low conversions $\gamma \rightarrow 0$ the maximum rates of steps (3) and (4) are:

$$r_3^{\max} = 2k_3\theta_{O_2}\theta_* \tag{1.37}$$

$$r_4^{\max} = k_4 \theta_{CO} \theta_O \tag{1.38}$$

The assumed optimal coverages that maximize $r_3^{\max}$ are $\theta_{O_2} = \theta_* = 0.5$, and $r_4^{\max}$ is maximized for $\theta_{CO} = \theta_O = 0.5$, leading to our final expression of the Sabatier rate.

$$r_{\text{Sabatier}} = \min\{0.5 \cdot k_3, 0.25 \cdot k_4\} \tag{1.39}$$

Reaction steps (3) and (4) in the mechanism for CO oxidation are both surface reactions, *i.e.* all reactants and the transition state are adsorbed species, and for this type of reactions the pre-exponential factors are typically on the order of $\nu = 10^{13}$ s^{-1}. The underlying assumption is that the entropy of the transition state is similar to the entropy of the initial state ($\Delta S^{\ddagger} = 0$) and eqn (1.33) reduces to $\nu = \frac{k_B T}{h} \approx 10^{13} s^{-1}$ at room temperature. The activation barriers for steps (3) and (4) can be found from the (transition state) scaling relations identified by Falsig *et al.* for fcc(111) surfaces.[53]

$$\Delta E_{\text{TS},3} = 1.39 \cdot \Delta E_{\text{O}} + 1.56 \text{ eV} \tag{1.40}$$

$$\Delta E_{\text{TS},4} = 0.7 \cdot (\Delta E_{\text{CO}} + \Delta E_{\text{O}}) + 0.02 \text{ eV} \tag{1.41}$$

$$\Delta E_{\text{O}_2} = 0.89 \cdot \Delta E_{\text{O}} + 0.17 \text{ eV} \tag{1.42}$$

The TSS equations can be rewritten to yield activation energy barriers.

$$\Delta E_{\text{a},3} = \Delta E_{\text{TS},3} - \Delta E_{\text{O}_2} = 0.5 \cdot \Delta E_{\text{O}} + 1.39 \text{ eV} \tag{1.43}$$

$$\Delta E_{\text{a},4} = \Delta E_{\text{TS},4} - \Delta E_{\text{CO}} - \Delta E_{\text{O}} = -0.3 \cdot (\Delta E_{\text{O}} + \Delta E_{\text{CO}}) + 0.02 \text{ eV} \tag{1.44}$$

The Sabatier rate for CO oxidation in eqn (1.39) can now be obtained from the knowledge of ΔE_{CO} and ΔE_{O}, which have naturally evolved again as the two descriptors that we need to describe the CO oxidation activity. Later in this chapter, a full microkinetic model is used to construct a two-dimensional volcano for CO oxidation that depends on both descriptors, but studying a one-dimensional volcano slice as shown in Figure 1.13 is easier to follow. This volcano slice was prepared under the artificial assumption that all transition metals have a constant binding energy of CO, which was set to $\Delta E_{\text{CO}} = -1.20$ eV. This value is within ± 0.3 eV of the actual CO binding energies on Ru, Rh, Ni, Pt, and Pd. In contrast, Cu, Ag, and Au bind CO much more weakly and their placement on the volcano in Figure 1.13 must be considered extremely approximate. CO oxidation rates on metals to the left of the maximum are limited by step (4), the removal of strongly adsorbed O*, while on the right side of the maximum the limiting step is O_2 dissociation, step (3). The optimal catalyst is the best compromise between all possible rate-determining steps, such that, at the maximum, no single step can be identified as rate-determining. It should also be mentioned that rate-limiting steps can change depending on the catalyst and reaction conditions and they should *not* be considered as a constant in any given reaction mechanism. A more detailed discussion of rate-limiting steps is presented later in this chapter. The important

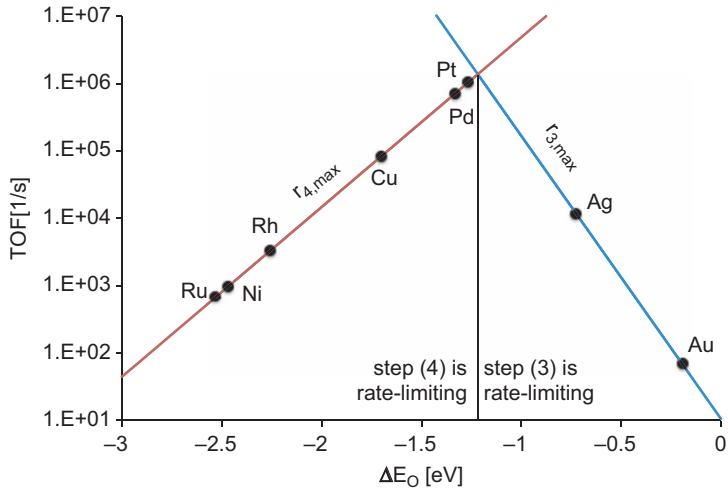


Figure 1.13 A slice through the CO oxidation volcano at $T = 600$ K for a constant value of $\Delta E_{CO} = -1.20$ eV.

observation here is that Pt and Pd are located closest to the top of the volcano, which is exactly what we expected. Given the similarly correct identification of the optimum catalysts for NH_3 synthesis in the earlier example, the success of a Sabatier analysis can no longer be considered a coincidence!

1.6 Notes on Microkinetic Modeling

Although the Sabatier analysis can be applied to reaction mechanisms with any number of elementary steps and even mechanisms with parallel reaction pathways, it is an approximation. More accurate computational screening for complex reaction mechanisms can be performed by implementing a full microkinetic model, which contains information about the dominant reaction mechanism and simultaneously predicts the activity, selectivity and surface coverages. As input to this microkinetic model only the catalytic descriptors are needed and all the missing energy information must be estimated from BEP and scaling relations or assumed. But before we continue with more advanced computational screening examples, it is necessary to introduce some very basic concepts needed for the development of microkinetic models in the context of computational catalyst screening. Gokhale *et al.*,[57] Stoltze and Nørskov,[58] and the standard reference *The Microkinetics of Heterogeneous Catalysis* by Dumesic *et al.*[59] provide useful information for microkinetic modeling.

The microkinetic modeling procedure for catalyst screening will be illustrated for the CO oxidation example, which is *the* prototype reaction in heterogeneous catalysis.[60] The same four-step mechanism used in the Sabatier example will be used here and the net rates of the elementary steps are given by:

$$r_1 = k_1^+ P_{CO} \theta_* - k_1^- \theta_{CO} \tag{1.45}$$

$$r_2 = k_2^+ P_{O_2} \theta_* - k_2^- \theta_{O_2} \tag{1.46}$$

$$r_3 = k_3^+ \theta_{O_2} \theta_* - k_3^- \theta_O^2 \tag{1.47}$$

$$r_4 = k_4^+ \theta_{CO} \theta_O - k_4^- P_{CO_2} \theta_*^2 \tag{1.48}$$

The change in coverage of reaction intermediates is determined by the net rate at which they are produced or consumed in each reaction step.

$$\frac{d\theta_{CO}}{dt} = r_1 - r_4 \tag{1.49}$$

$$\frac{d\theta_{O_2}}{dt} = r_2 - r_3 \tag{1.50}$$

$$\frac{d\theta_O}{dt} = 2r_3 - r_4 \tag{1.51}$$

And the fraction of empty sites on the surface can at any point of time be calculated from the site balance.

$$\theta_* = 1 - \theta_{O_2} - \theta_O - \theta_{CO} \tag{1.52}$$

Eqns (1.49)–(1.51) represent a system of ordinary differential equations (ODE), but we will restrict ourselves to situations where the reaction runs under steady-state conditions, which means that the derivatives of the surface coverages with respect to time in eqns (1.49)–(1.51) are all zero.

$$\frac{d\theta_i}{dt} = 0 \quad \text{for} \quad i = O_2, O, CO \tag{1.53}$$

This converts the system of ODEs into a system of non-linear algebraic equations that can be solved with standard root-finding methods. The steady-state assumption is not strictly necessary when solving a microkinetic model numerically, but the transient start-up and shut-down behavior is typically short in comparison to steady-state operation.

With eqns (1.52) and (1.53) there are four equations for the four unknown coverages θ_{O_2}, θ_O, θ_{CO}, and θ_* and the system of nonlinear algebraic equations may be solved numerically. With currently available CPU speeds numerical solutions to microkinetic models for catalyst screening studies are generally preferred because they avoid the need to make any additional assumptions regarding the mechanism.

Analytical solutions are also valuable, but they almost always require additional assumptions and involve cumbersome pencil and paper math. A very common assumption is that molecular adsorption/desorption steps are fast in comparison to surface reactions and may be assumed to be quasi-equilibrated. We can make this assumption for steps (1) and (2) of our mechanism and then find the surface coverages of CO* and O_2* as a function of the fraction of empty sites.

$$r_1 \approx 0 \Rightarrow k_1^+ P_{CO} \theta_* = k_1^- \theta_{CO} \Rightarrow \theta_{CO} = K_1 P_{CO} \theta_* \tag{1.54}$$

$$r_2 \approx 0 \;\Rightarrow k_2^+ P_{O_2} \theta_* = k_2^- \theta_{O_2} \Rightarrow \theta_{O_2} = K_2 P_{O_2} \theta_* \tag{1.55}$$

There is not a reaction that we can assume to be quasi-equilibrated to derive an equation for the coverage of O*, but from the steady-state assumption and eqn (1.51) we find

$$0 = \frac{\mathrm{d}\theta_O}{\mathrm{d}t} = 2r_3 - r_4 = 2\left(k_3^+ \theta_{O_2} \theta_* - k_3^- \theta_O^2\right) - k_4^+ \theta_{CO} \theta_O + k_4^- P_{CO_2} \theta_*^2 \tag{1.56}$$

which can be solved for θ_O by substituting θ_{O_2} and θ_{CO} from (1.54) and (1.55).

$$\begin{aligned} \theta_O &= \frac{-k_4^+ K_1 P_{CO} + \sqrt{\left(k_4^+ K_1 P_{CO}\right)^2 - 8k_3^- \left(k_4^- P_{CO_2} - 2k_3^+ K_2 P_{O_2}\right)}}{4k_3^-} \theta_* \\ &= W\theta_* \end{aligned} \tag{1.57}$$

In eqn (1.57) we introduced W to denote the long fraction. For an analytical solution one now simply uses the site balance eqn (1.52) to solve for θ_*.

$$\begin{aligned} \theta_* &= 1 - K_2 P_{O_2} \theta_* - W\theta_* - K_1 P_{CO} \theta_* \;\Rightarrow\; \theta_* \\ &= \frac{1}{1 + W + K_1 P_{CO} + K_2 P_{O_2}} \end{aligned} \tag{1.58}$$

Equation (1.58) can be solved for any reaction condition if the equilibrium constants, K_i, for steps (1) and (2), and the forward and reverse rate constants, k_i^+ and k_i^-, for steps (3) and (4) are known. Note that we do not need to know the rate constants of the quasi-equilibrated steps (1) and (2) and the equilibrium constants are sufficient. We will use this analytical solution later and compare it to a full numerical solution of the CO oxidation model.

1.6.1 Numerical Solution Strategies

Most people will tell you that solving a microkinetic model numerically is easy and does not require any special solution strategies: one can simply use a standard ODE solver or root-finding method. This may be the case for microkinetic models that are parameterized for a working catalyst and predict reasonable reaction rates, but when we do a computational screening study we want to explore a large parameter space for our descriptors, which also includes unusual parameter combinations. Specifically, at the edges far away from the volcano maximum, the rate constants in the microkinetic model span many orders of magnitude (40–50 orders is not unusual) and numerical solvers will not always converge. Although there is no general procedure to tackle this problem there are some guidelines.

In a computational screening study, the rate and equilibrium constants calculated using scaling and BEP relations are functions of the chosen descriptors. A volcano plot could be constructed by systematically sweeping through the parameter space for the descriptors. Root-finding methods can quickly solve

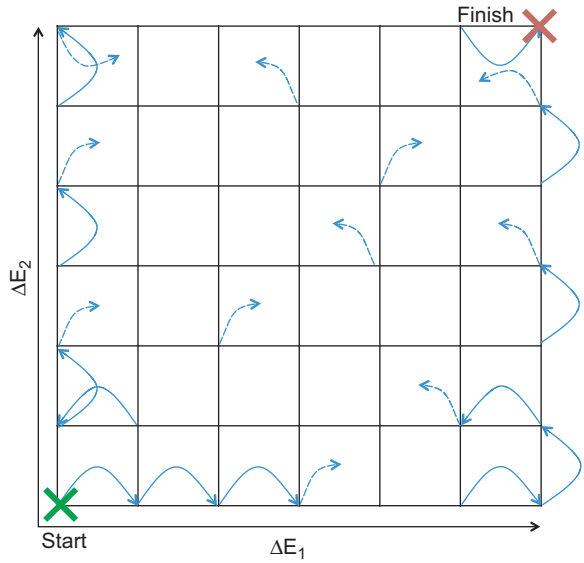


Figure 1.14 Suggested looping technique for the exploration of a two-dimensional descriptor space.

the steady-state equations, but the rate equations in microkinetic models are highly non-linear and convergence is only achieved when good initial guesses are provided. Hence, it is a good idea to use the solution of the microkinetic model for one descriptor set as the initial guess to the neighboring descriptor set. A possible solution path on a grid of two descriptors is suggested in Figure 1.14. Provided that the grid points in the descriptor space are close to each other, the solutions should be sufficiently similar to obtain converged results.

An alternative approach is to integrate the system of ODEs, which is computationally much more intensive, but does not depend as strongly on the initial guesses, as long as they make physical sense, *i.e.* no negative coverages or violation of the site balance. In order to find a steady-state solution the ODEs need to be integrated over a sufficient amount of time to ensure that steady-state has been reached. It is a good practice to verify steady-state convergence by plotting the transient behavior of the surface coverages and ensure that they do not continue to change over time.

Another important aspect to consider is the stiffness of the ODE system, a characteristic caused by the large differences in rate constants. Most ODE solvers, even those specialized for stiff systems, will fail in those cases. If the simpler root-finding strategy also fails, it is possible to reduce the stiffness of the ODE system by artificially slowing down some of the very fast, quasi-equilibrated, reaction steps. But pay attention! If you slow down these steps too much, they may become kinetically relevant and affect the overall rate or the reaction mechanism. This strategy is not recommended for ordinary

problems, but may be attempted with the appropriate care if everything else fails.

1.6.2 Entropy and Enthalpy Corrections

In the previous examples we only considered electronic energy changes and approximated the entropy as all or nothing. In essence, we assumed that gas-phase species have 100% of their standard state entropy and surface species possess no entropy at all. These assumptions can certainly be improved and in order to construct thermodynamically consistent microkinetic models this is not just optional, but absolutely necessary. Entropy and enthalpy corrections for surface species can be calculated using statistical thermodynamics from knowledge of the vibrational frequencies, and the translational and rotational degrees of freedom (DOF). In contrast to gas-phase molecules, adsorbates cannot freely rotate and move across the surface, but the translational and rotational DOF are frustrated within the potential energy well imposed by the surface. In the harmonic limit the frustrated translational and rotational DOF can conveniently be described as vibrational modes, which in turn means that any surface adsorbate will have $3N$ vibrational DOFs that are all treated equally.

The temperature dependent surface enthalpy, $H(T)$, of an adsorbate is composed of the electronic energy, E_{elec}, the zero-point energy, E_{ZPE}, and a heat capacity correction to account for the difference between the enthalpy at 0 K and at the actual temperature, T.

$$H(T) = E_{\text{elec}} + E_{\text{ZPE}} + \int_{0K}^{T} C_V \, d\tilde{T} \tag{1.59}$$

The electronic energy component, E_{elec}, is directly obtained from either DFT calculations or from scaling relations based on DFT energies. The remaining components can be calculated in the harmonic approximation.[61]

$$E_{\text{ZPE}} = \frac{1}{2} \sum_{i=0}^{\#\text{DOF}} h\nu_i$$

$$\int_{0K}^{T} C_v d\tilde{T} = \sum_{i=0}^{\#\text{DOF}} \frac{h\nu_i}{e^{h\nu_i/k_BT} - 1}$$

$$S(T) = k_B \sum_{i=0}^{\#\text{DOF}} \left[\frac{h\nu_i}{k_BT(e^{h\nu_i/k_BT} - 1)} - \ln\left(1 - e^{-h\nu_i/k_BT}\right) \right] \tag{1.60}$$

The complete set of vibrational modes for an adsorbate is not easily obtained and usually it is necessary to perform an additional vibrational analysis using DFT. In this vibrational analysis it is important to pay attention to the low frequency modes, corresponding to the frustrated translation, for two

reasons: (i) DFT is notoriously inaccurate in the calculation of low frequency modes, and (ii) low frequency modes contribute the most to the surface entropy. It is advised to check the entropy contributions of the translational modes separately, and to limit them to the entropy of a two-dimensional surface gas, which can be derived from the partition function of a particle that can move freely across the surface.

$$S_{\text{trans}}^{2D} = k_B\left[\ln\left(\frac{2\pi m k_B T}{C_S h^2}\right) + 2\right] \tag{1.61}$$

In this equation m denotes the mass of the particle and C_S is the number of sites per unit area, which is roughly 10^{15} sites cm^{-2} for transition metal surfaces. Experience has shown that the entropy and enthalpy corrections do not vary appreciably across different transition metal surfaces and for computational screening it is commonly assumed that these contributions are constant.

1.6.3 Microkinetic Model Analysis

Microkinetic models can provide much more information than "just" surface coverages and reaction rates of individual steps. It is always a good idea to perform a sensitivity analysis on the model input parameters, to obtain additional valuable information. First, more attention should to be paid to parameters to which the model shows a high sensitivity, and these parameters need to be carefully measured or calculated. Non-sensitive parameters, on the other hand, may be roughly estimated or even guessed. Second, the sensitivity of a microkinetic model towards certain parameters can be interpreted as a measure of the degree to which a step is rate controlling (degree of rate control). This information can provide direction for useful catalyst modifications which could optimize performance (degree of catalyst control).

1.6.3.1 Degree of Rate Control

The degree of rate control, X_{RC}, was originally proposed by Campbell as a quantitative measure to identify rate-controlling steps.[62,63] It was later generalized to quantify also the impact of adsorbate binding (poisoning) on the reaction rate.[64] The degree of rate control of step i is formally defined as the normalized partial derivative of the overall rate with respect to the rate constant, k_i, while keeping the equilibrium constant, K_i, and the rate constants, k_j, of all other steps constant.

$$X_{\text{RC},i} = \frac{k_i}{r}\left(\frac{\partial r}{\partial k_i}\right)_{K_i,k_j} = \left(\frac{\partial \ln r}{\partial \ln k_i}\right)_{K_i,k_j} \tag{1.62}$$

An $X_{\text{RC},i}$ value of zero indicates a quasi-equilibrated step, whose rate constant has no effect on the overall rate, while $X_{\text{RC},i}=1$ indicates a single rate-controlling step in a serial reaction mechanism. For serial reaction mechanisms with only one product it has also been found that $\sum_{i=0}^{\#\text{steps}} X_{\text{RC},i} = 1$, but in reaction mechanisms with parallel pathways or more than one possible product

an analogous rule has not been identified. In the context of microkinetic modeling, eqn (1.62) is not easily applicable, but it can be rewritten into a finite-difference equation that can be readily evaluated in a microkinetic model.

$$X_{\mathrm{RC},i} = \frac{k_i}{r}\left(\frac{\partial r}{\partial k_i}\right)_{K_i,k_j} \approx \frac{k_i^0}{r^0}\cdot\frac{r-r^0}{k_i-k_i^0} = \frac{r-r^0}{x\cdot r^0} \quad \text{with} \quad k_i = (1+x)k_i^0 \qquad (1.63)$$

In practice $X_{\mathrm{RC},i}$ is then calculated by changing the forward and reverse rate constants of step i simultaneously by a small amount, typically 10% ($x = 0.1$), and monitoring the change of the overall reaction rate with respect to the reference rate, r^0, for the original values of k_i. *Never* forget to apply this change to *both* rate constants, otherwise K_i and the overall equilibrium constant, K_{eq}, will be altered as well!

Fundamentally, the definition of the degree of rate control in eqn (1.62) implies a change in the Gibbs free energy of the transition state of step i, as indicated for the transition state of CH_4 activation in Figure 1.15, while holding the energy of all other states in the potential energy surface constant. In more mathematical terms, this interpretation can be derived from the Arrhenius expression, which can be written as

$$\ln\left(\frac{k_i}{c}\right) = \frac{-G_{\mathrm{a},i}}{k_B T} \qquad (1.64)$$

where c is a constant ($c = \frac{k_B T}{h}$ in transition state theory) and $G_{\mathrm{a},i}$ is the free energy of activation for step i. Upon substituting eqn (1.64) into the definition of $X_{\mathrm{RC},i}$ in eqn (1.62) we obtain the degree of rate control in terms of the free energy of the transition state of step i, ($G_{\mathrm{TS},i}$).

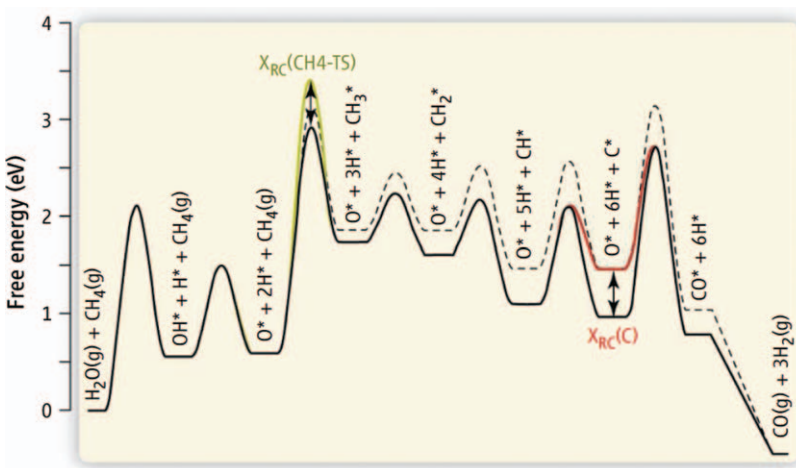


Figure 1.15 Illustration of the generalized degree of rate control X_{RC} and degree of catalyst control X_{CC} (dashed PES) for CH_4 steam reforming. X_{RC}(CH4-TS) = 0.8, X_{RC}(C) = –0.26, X_{CC}(C) = 0.11.
(Reproduced with the kind permission of Science, ref. 65)

$$X_{\mathrm{RC},i} = \left(\frac{\partial \ln r}{\partial \left(\frac{-G_{\mathrm{TS},i}}{k_B T} \right)} \right)_{G_{\mathrm{TS},j \neq i}, G_n} \tag{1.65}$$

Equivalently, one can change the free energy G_n of an intermediate n and calculate its degree of rate control. This is schematically shown for the free energy of C* in Figure 1.15 and one can define the thermodynamic degree of rate control, $X_{\mathrm{TRC},n}$, for an intermediate n as

$$X_{\mathrm{TRC},n} = \left(\frac{\partial \ln r}{\partial \left(\frac{-G_n}{k_B T} \right)} \right)_{G_{\mathrm{TS},i}, G_{m \neq n}} \tag{1.66}$$

Eqns (1.65) and (1.66) are in principle identical and are collectively referred to as the generalized degree of rate control. In typical scenarios, the degree of rate control for transition states is positive, indicating that lowering the transition state energy increases the reaction rate, and the degree of rate control for surface intermediates is negative, which means that stabilizing the intermediate will poison the surface and result in a slower reaction rate. This can be interpreted more generally for heterogeneously catalyzed reactions. On the optimal catalyst the overall reaction proceeds along a ''smooth'' PES from the reactants to the products. Large potential energy barriers or deep wells result in a corrugated PES and slow reaction rates.

1.6.3.2 Degree of Catalyst Control

Campbell's degree of rate control is an extraordinarily useful concept for the analysis of reaction mechanisms, the identification of rate-limiting steps and the poisoning effect of adsorbates, but it implies that the energies of intermediates and transition states can be changed independently, one at a time. While this is the common procedure for a sensitivity analysis, we already know that the stabilities of surface intermediates and transition states are intimately linked through scaling and BEP relations, which find their physical foundation in the *d*-band model. It is therefore unrealistic to assume that we can find catalysts that, for example, reduce the activation energy barrier of a rate-limiting step without simultaneously affecting other parts of the PES. This leads to the idea of the degree of catalyst control X_{CC}, which is a constrained sensitivity analysis of a microkinetic model with respect to a reactivity descriptor E_i.[65] The definition of $X_{\mathrm{CC},i}$ is analogous to the generalized degree of rate control and may be expressed as

$$X_{\mathrm{CC},i} = \left(\frac{\partial \ln r}{\partial \left(\frac{-E_i}{k_B T} \right)} \right)_{E_{j \neq i}, BEP, Scaling} \tag{1.67}$$

In eqn (1.67) it is indicated that all other descriptor values $E_{j \neq i}$ and the implemented scaling and BEP relations are kept constant when taking the partial derivative. The rationale for this definition is that the catalytic

descriptors are the only independent variables and X_{CC} will indicate the direction one must look to identify an improved catalyst. The optimal catalyst is characterized by $X_{CC,i} = 0$ for all descriptors E_i. X_{CC} is proportional to the slope of the volcano curve and following it in the uphill direction will get you to the top!

The difference between X_{RC} and X_{CC} is shown graphically in Figure 1.15 for the sensitivity towards the binding energy of carbon, ΔE_C. Carbon adsorbs quite strongly and is located in a well of the PES, which explains the negative value (–0.26) for X_{RC}(C). Decreasing the depth of the well would smooth out the PES and increase the reaction rate. However, decreasing the stability of C* will also affect the stability of CH*, CH_2*, CH_3* surface intermediates and the associated transition states between them through scaling and BEP relations (dashed line in Figure 1.15). In particular, the transition state energy of CH_4 activation will shift up in energy. The large value of $X_{RC}(CH_4\text{-TS}) = 0.8$ indicates that CH_4 activation is a rate-determining step and increasing the barrier will have a large negative impact on the overall reaction rate. The overall effect of changing the descriptor ΔE_C on the PES is given by $X_{cc}(C) = 0.11$, indicating that a catalyst with stronger ΔE_C will have higher activity. The explanation is that the increased poisoning effect of C* is more than compensated by lowering the barrier for the CH_4 activation step.

The degrees of rate and catalyst control both have their merits and do not conflict with each other. Determining which to use is dependent on the problem at hand. X_{RC} should be used to analyze the reaction mechanism and understand the importance of individual steps and adsorbates, whereas X_{CC} should be used to guide catalyst design.

1.7 CO Oxidation Catalyst Screening

We have already performed a preliminary Sabatier analysis of the CO oxidation reaction in Section 1.5.2, and derived an analytic solution under the assumption that the adsorption of CO and O_2 are quasi-equilibrated in Section 1.6. Now we will formulate a numerical solution to the complete microkinetic model as a function of the descriptors ΔE_{CO} and ΔE_O. We will analyze the reaction mechanism in terms of rate and catalyst control, and at the end of this section, the effect of high surface coverages on the volcano curve will also be briefly addressed.

1.7.1 Numerical Microkinetic Model

Microkinetic models are typically custom-made and can be written in any programming or scripting language that the reader is most familiar with. It is recommended to choose a simulation environment that already provides lots of the functionality needed to solve the model, *i.e.* an ODE solver, a root-finder, and *n*-dimensional curve fitting routines, if parameter optimization is also required. Here we use python, which is an object-oriented scripting language with very good performance, and it is freely available for all popular operating systems (Windows, MacOS, Linux; http://python.org). The python syntax

should be self-explanatory as long as you are aware that comments are declared with '#' and that list and array indices start with '0', not with '1'. Readers who are interested in additional information are referred to the online tutorial available for python newbies (http://docs.python.org/tutorial). Besides the standard python interpreter, we also need the NumPy (http://numpy.scipy.org), SciPy (http://www.scipy.org), mpmath (http://code.google.com/p/mpmath), and matplotlib (http://matplotlib.sourceforge.net) modules as extensions. These modules provide a variety of math routines and plotting functions, very similar to the Matlab environment. All examples here have been tested with Python 2.6, but other versions of python should work as well. A complete python script reproducing all results of this section can be found in the Appendix.

1.7.1.1 Constants and Conditions

It is a good idea to define the reaction conditions and useful physical constants globally at the beginning of the code, so that they are available to all functions. This is also the place to import additional modules, such as numpy.

```
import numpy as np

# Reaction conditions
T    = 600      # K
PCO  = 0.33     # bar
PO2  = 0.67     # bar
PCO2 = 1        # bar

# Physical constants and conversion factors
J2eV = 6.24150974E18            # eV/J
Na   = 6.0221415E23             # mol-1
h    = 6.626068E-34 * J2eV      # in eV*s
kb   = 1.3806503E-23 * J2eV     # in eV/K
kbT  = kb * T                   # in eV
```

1.7.1.2 From Descriptors to Rate Constants

The next important step is to implement the scaling and BEP relations together with any assumptions regarding entropies of transition states and surface states in order to obtain the necessary rate constants. Since we will need this procedure multiple times for each combination of our descriptors, we define it as a function.

```
def get_rate_constants(EO,ECO):
# This function applies scaling and BEP to determine
all rate constants
# depending on the two descriptors, EO and ECO,
provided as input.
```

```
# Gas phase energies referenced to CO and O2
ECOg = EO2g = 0
ECO2g = -3.627              # in eV calculated from DFT

# Scaling for EO2
EO2 = 0.89 * EO + 0.17      # eq. (36)

# Gas phase entropies converted to eV/K
SCOg  = 197.66 * J2eV / Na  # eV/K
SO2g  = 205.0 * J2eV / Na
SCO2g = 213.74 * J2eV / Na

# Surface entropies are assumed to be zero
SCO = SO2 = SO = 0

# Reaction energies
dE = np.zeros(4)            # array initialization
dE [0] = ECO - ECOg
dE [1] = EO2 - EO2g
dE [2] = 2*EO - EO2
dE [3] = ECO2g - ECO - EO

# Entropy changes
dS = np.zeros(4)            # array initialization
dS [0] = SCO - SCOg
dS [1] = SO2 - SO2g
dS [2] = 2*SO - SO2
dS [3] = SCO2g - SCO - SO

# Activation energy barriers from BEP
Ea = np.zeros(4)            # array initialization
Ea [2] = 0.5 * EO + 1.39    # eq. (37)
Ea [3] = -0.3 * (EO + ECO)  # eq. (38)
       + 0.02

# Entropy changes to the transition state
STS = np.zeros(4)           # array initialization
STS [0] = -0.25*SCOg        # loss of CO entropy assumed
STS [1] = -0.25*SO2g        # loss of O2 entropy assumed
STS [2] = 0                 # surface reaction
STS [3] = 0                 # surface reaction

# Calculate equilibrium and rate constants
K  = [0]*4                  # equilibrium constants
kf = [0]*4                  # forward rate constants
```

```
  kr = [0]*4                      # reverse rate constants
  for i in range(4):
      dG = dE[i] - T*dS[i]
      K[i]  = np.exp(-dG/kbT)
      # Enforce Ea > 0, and Ea > dE, independent of what
the descriptors are
      Ea[i] = max([0,dE[i],Ea[i]])
      kf[i] = kbT/h * np.exp(STS[i]/kb) * np.exp(-Ea[i]/
              kbT)
     kr[i] = kf[i]/K[i]           # enforce thermodynamic
                                    consistency
  return (kf,kr)
```

In this function we have used the scaling/BEP relations, eqns (1.42)–(1.44), and we have arbitrarily assumed that CO and O_2 lose 25% of their gas-phase entropy when entering the transition state for adsorption. Using collision theory for the adsorption rate constants[66] of these two steps is a more elegant alternative and the reader is encouraged to see how it affects the result of the model. There are two more noteworthy points. First, the model ensures that the activation barriers, $E_{a,i}$, are all positive, even for descriptor pairs that would otherwise predict negative values for $E_{a,i}$. Second, thermodynamic consistency is enforced throughout the model by calculating the reverse rate constants from the equilibrium constants for each step.

1.7.1.3 Rate Equations

This step is simple. We need rate equations for the individual steps. The rates depend on the coverages, θ_i, and the rate constants. All parameters are passed as an array to the function that returns the rates.

```
def get_rates(theta,kf,kr):
# returns the rates depending on the current coverages
theta

  # Extract elements of theta and assign them
  # to more meaningful variables
  tCO   = theta[0]              # theta of CO
  tO2   = theta[1]              # theta of O2
  tO    = theta[2]              # theta of O
  tstar = 1.0 - tCO - tO2 - tO  # site balance for tstar

  # Caluclate the rates: eqns (39)-(42)
  rate    = [0]*4               # array with 4 zeros
  rate[0] = kf[0] * PCO * tstar - kr[0] * tCO
```

```
  rate [1]  = kf [1]  * PO2 * tstar - kr [1]  * tO2
  rate [2]  = kf [2]  * tO2 * tstar - kr [2]  * tO * tO
  rate [3]  = kf [3]  * tCO * tO - kr [3]  * PCO2 * tstar *
              tstar

  return rate
```

1.7.1.4 System of ODEs vs. Steady-State Equations

As discussed in Section 1.6.1 the microkinetic model may be solved as a system of ODEs or non-linear algebraic equations using the steady-state assumption. It turns out that, regardless of which approach you want to use, the function that must be passed to an ODE solver or numerical root-finding method is the same! Here, the more general case of the ODE system is chosen. Note that we named the previously defined function get_rates().

```
def get_odes(theta,t,kf,kr):
# returns the system of ODEs d(theta)/dt, calculated at
the current value of theta (and time t)

  rate = get_rates(theta,kf,kr)      # calculate the
                                       current rates

  # Time derivatives of theta
  dt  = [0]*3
  dt [0]  = rate [0]  - rate [3]        # d(tCO)/dt
  dt [1]  = rate [1]  - rate [2]        # d(tO2)/dt
  dt [2]  = 2 * rate [2]  - rate [3]    # d(tO)/dt

  return dt
```

1.7.1.5 Solving the Model

All that is left at this point is to set the values of the descriptors and solve the model. For single point calculations using only one descriptor set, it is preferable to use an ODE solver and use a completely empty surface as the initial guess for the coverages. The integration time required to reach steady-state depends on the problem and can vary from 1×10^3 to 1×10^8 seconds. Some might wonder why a reaction could possibly take 1×10^8 seconds, or greater than 3 years, to equilibrate, but keep in mind that our goal is to scan a wide descriptor range, which will result in some very small rate constants. For this

example, we use the Pt values for the descriptors and the odeint ODE solver provided by scipy. Note that the default tolerance and step size had to be tweaked in order for odeint to work for this stiff problem.


```
# Solve the model for Pt with EO = -1.25, ECO = -1.22
ECO = -1.22             # CO oxidation descriptors
EO  = -1.25

(kf,kr) = get_rate_constants # get the rate constants
          (EO,ECO)             for the given descriptors

# Use scipy's odeint to solve the system of ODEs
from scipy.integrate import odeint
# As initial guess we assume an empty surface
theta0 = (0., 0., 0.)

# Integrate the ODEs for 1E6 sec (enough to reach
steady-state)
theta = odeint(get_odes,     # system of ODEs
           theta0,           # initial guess
           [0,1E6],          # time span
           args = (kf,kr),   # additional arguments to
                               get_odes()
           h0 = 1E-36,       # initial time step
           mxstep = 90000,   # maximum number of steps
           rtol = 1E-12,     # relative tolerance
           atol = 1E-15)     # absolute tolerance

print_output(EO,ECO,theta [-1,:])
```

In the last line the helper function `print_output` (given below, but in the actual python script it needs to be defined *before* it is called) is called with the descriptor values and the final coverages. The variable `theta` is a ($t \times c$) matrix with the coverages c for each time step t. Usually, we are only interested in the last row, conveniently accessed by using '-1' as index, but the time evolution is available as well for transient studies or verification that steady-state has been reached.

```
def print_output(EO,ECO,theta):
# Prints the solution of the model
    (kf,kr) = get_rate_constants(EO,ECO)
    rates   = get_rates(theta,kf,kr)
   print "For the descriptors EO=", EO, "and ECO=",
ECO,"the result is:"
```

```
    print
    for r,rate in enumerate(rates):
        print "Step",r,": rate =",rate,", kf =",kf[r],",
kr=",kr[r]
    print
    print "The coverages for CO*, O2*, and O* are:"
    for t in theta:
        print t
```

As a result of this exercise, we find the steady-state coverages on Pt of $\theta_{CO} = 0.215$ ML, $\theta_{O2} = 8.4 \times 10^{-4}$, and $\theta_O = 5.6 \times 10^{-3}$ and a steady-state CO_2 formation rate of 6.16×10^3 s^{-1}.

1.7.2 Degree of Rate and Catalyst Control

Now that we are able to predict a reaction rate for a single catalyst, only a few lines of extra code are needed to calculate the degree of rate control and the degree of catalyst control. However, we need to solve the microkinetic model repeatedly for different parameters and it is more convenient if we first write a function, solve_ode, that takes the rate constants as input and returns the solution of the model. This function may look like this.

```
def solve_ode(kf,kr,theta0=(0.,0.,0.)):
# Solve the system of ODEs using scipy.integrate.odeint
# Assumes an empty surface as initial guess if nothing
else is provided
  from scipy.integrate import odeint

   # Integrate the ODEs for 1E6 sec (enough to reach
steady-state)
  theta = odeint(get_odes,     # system of ODEs
            theta0,            # initial guess
            [0,1E6],           # time span
            args = (kf,kr),    # arguments to get_odes()
            h0 = 1E-36,        # initial time step
            mxstep = 90000,    # maximum number of steps
            rtol = 1E-12,      # relative tolerance
            atol = 1E-15)      # absolute tolerance

  return theta[-1,:]
```

The degree of rate control, X_{RC}, is then calculated by repeatedly calling `solve_ode` with rate constants that are systematically changed for each step.

As input to `calculate_Xrc` we provide the unaltered rate constants and the corresponding rate.

```
def calculate_Xrc(r0,kf0,kr0):
# Calculates Xrc by systematically changing the
# rate constants of each step by 10% around the refer-
ence value

  delta = 0.1                  # change of 10%
  Xrc_rates = np.zeros(4)      # array for storing rates
  for s in range(4):           # loop over all steps
     # initialize rate constants with reference values
     kf = kf0 [:]
     kr = kr0 [:]
     kf [s]  = (1 + delta) * kf0 [s]
     kr [s]  = (1 + delta) * kr0 [s]
     theta = solve_ode(kf,kr)        # Solve with the
modified k' s
     rates = get_rates(theta,kf,kr)  # Get the new rates
     Xrc_rates [s]  = rates [3]     # Save the rate of CO2
production

  # And calculate Xrc for all steps
  Xrc = (Xrc_rates-r0)/(delta*r0)
  return Xrc
```

In our example for CO oxidation on Pt with $\Delta E_O = -1.25$ eV and $\Delta E_{CO} = -1.22$ eV the degrees of rate control for steps (1) and (2) are zero (these steps are quasi-equilibrated), step (3) is almost exclusively rate-controlling with $X_{RC,3} = 0.99$, and step (4) has an $X_{RC,4} = 0.01$. In this example $\sum_i X_{RC,i} = 1$, which is expected for a serial reaction mechanism.

Next we calculate the degree of catalyst control X_{cc} by varying the values of ΔE_O and ΔE_{CO} by 0.05 eV. This step size should generally work well except in close proximity to the volcano peak, where smaller step sizes are preferable. Rather than providing the rate constants directly, we pass the original values of ΔE_O and ΔE_{CO} to the function `calculate_Xcc` and the corresponding rate constants are calculated on the fly.

```
def calculate_Xcc(r0,EO0,ECO0):
# Calculate Xcc by varying EO and ECO around their
# reference values by 0.05 eV

   delta = 0.05               # change of 0.05 eV
   Xcc_rates = np.zeros(2)    # array for storing rates
```

```
  # Start by modifying EO0
  (kf,kr) = get_rate_constants(EO0+delta,ECO0)
  theta = solve_ode(kf,kr)          # Solve with the
modified k' s
  rates = get_rates(theta,kf,kr)    # Get the new rates
  Xcc_rates[0] = rates[3]           # Save the rate of CO2
production

  # Repeat for ECO0
  (kf,kr) = get_rate_constants(EO0,ECO0+delta)
  theta = solve_ode(kf,kr)          # Solve with the
modified k' s
  rates = get_rates(theta,kf,kr)    # Get the new rates
  Xcc_rates[1] = rates[3]           # Save the rate of CO2
production
  Xcc = (np.log(Xcc_rates)-np.log(r0))/(-delta/kbT)
  return Xcc
```

If we perform this analysis for CO oxidation on Pt, we get $X_{cc}(\Delta E_O) = 1.38$ and $X_{cc}(\Delta E_{CO}) = -0.28$, which indicates that by stabilizing O* and destabilizing CO* the reaction rate can be increased. Indeed, on the two-dimensional volcano plot in Figure 1.16 it can clearly be seen that the maximum is located to the left of Pt in the direction of stronger O* binding. The resolution of the volcano plot is however not sufficient to judge whether stronger or weaker CO* binding would increase the activity. It appears as though ΔE_{CO} for Pt is already near to its optimal value.

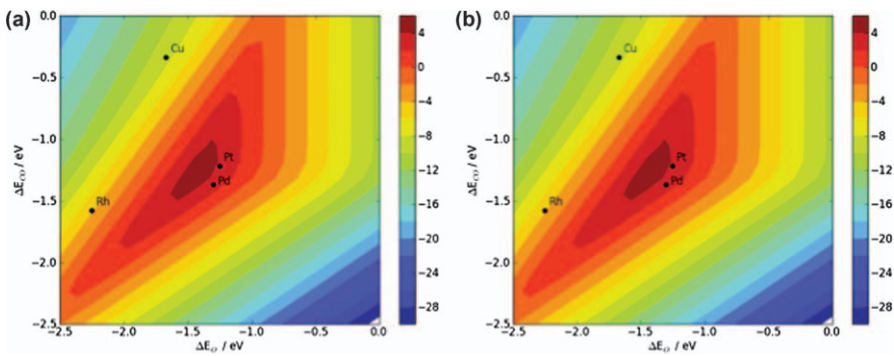


Figure 1.16 CO oxidation activity, $\log_{10}$(TOF), as a function of the two catalytic descriptors ΔE_O and ΔE_{CO}. (a) Numerical solution. (b) Analytical solution derived in Section 1.7, eqns (1.54)–(1.58). The positions of the closed-packed surfaces of Pt, Pd, Cu, and Rh are indicated.

1.7.3 Two-dimensional CO Oxidation Volcano

After this thorough analysis of the microkinetic model results around a single point (Pt), the only remaining task is to calculate the CO oxidation rate over a larger descriptor range and visualize the results as a two-dimensional volcano plot as shown in Figure 1.16. As a comparison to the numerical solution in Figure 1.16(a), the analytical solution from Section 1.6 [eqns (1.54)–(1.58)] is given in Figure 1.16(b). If we recall that the only assumption that was made in the derivation of the analytical solution was that the CO and O_2 adsorption steps are quasi-equilibrated and that the numerically calculated degree of rate control of both steps was zero, it is not surprising that the plots are in perfect agreement.

Figure 1.16 can ideally be generated by simply looping over the entire parameter range, but in practice that will often result in convergence errors due to numerical precision problems with the largely varying rate constants in certain regions of the parameter space. For faster results it is also desirable to use a root-finding method (*e.g.* `mpmath.findroot`), which requires good initial guesses for convergence. Here, the alternating looping technique visualized in Figure 1.14 was used, and the solution of the previous grid point served as an initial guess to the next one. However, even this approach is not fail proof and a sanity check (*e.g.* are all coverages between 0 and 1?) of the solutions in each grid point must be performed before moving on to the next grid point. If the solution is not physical, then one must resort to a slower, but more stable, solver in order to generate a new solution and then possibly continue again with the fast root-finding method. An example of how this may be implemented is provided below.

```
# Generate the data for 2D volcano plots
# and save it in a matrix
gridpoints = 20
num_data = np.zeros( [gridpoints,gridpoints] )

# Generate the initial guess for the first point
(kf,kr) = get_rate_constants(-2.5,-2.5)
theta0 = solve_ode(kf,kr)

# Start the scan through the parameter space
# Alternate directions for alternating columns
EO_range0 = np.linspace(-2.5,0.0,gridpoints)
ECO_range0 = np.linspace(-2.5,0.0,gridpoints)
for i,EO in enumerate(EO_range0):
    if i%2:                                # true if odd
       ECO_range = ECO_range0.copy()[::-1]   #[::-1]
reverses the list
       j_range = range(gridpoints)[::-1]
    else:
       ECO_range = ECO_range0.copy()
       j_range = range(gridpoints)
```

```
    for j,ECO in enumerate(ECO_range):
        (kf,kr) = get_rate_constants(EO,ECO)
        theta = solve_findroot(kf,kr,theta0)
        # Check of solution is physical
        if (theta > 1.0).any() or (theta < 0.0).any():
            # Generate a new initial guess using odeint
and solve again
            theta0 = solve_ode(kf,kr)
            theta = solve_findroot(kf,kr,theta0)
        rates = get_rates(theta,kf,kr)
        # Store the solutions
        num_data [j_range [j],i] = rates [3]
        # Store the numerical solution as initial guess
        # for the next grid point
        theta0 = theta.copy()
```

The numerical rate data collected in `num_data` can then be plotted using the `contourf` function from `matplotlib`. The full example code for solving the microkinetic model including the generation of the two-dimensional volcano graph can be found in the Appendix.

1.7.4 Effect of Lateral Interactions

Throughout this chapter it has been assumed that binding energies and activation energy barriers are constant, but at higher surface coverages these quantities can be strongly influenced by adsorbate–adsorbate interactions. An accurate description of these interactions, specifically when they induce local ordering and lead to island formation or other characteristic patterns, would require the use of kinetic Monte Carlo (kMC) simulations. Since computational catalyst screening is largely based on the use of rather approximate scaling and BEP relations with mean absolute errors of *ca.* 0.2 eV, the use of highly sophisticated kMC methods does not seem appropriate. Adsorbate–adsorbate interactions can also be described within the mean-field approximation, which is less detailed but suitable for inclusion in microkinetic models.[48,67] In the simplest case of interacting atomic adsorbates the lateral interactions have been understood in terms of *d*-band structure changes, which in turn allows for the prediction of these interactions.[68] Just as lateral interactions alter the binding energy of adsorbates, they also affect the stability of transition states and thereby change the activation energy barriers for surface reactions. William Schneider's group has studied this phenomenon in great detail by calculating coverage-dependent activation barriers for NO oxidation, which were later coupled with first principles-based cluster expansions for improved accuracy.[69,70] Current pursuits to systematically improve the description of lateral interactions in microkinetic models are slowly closing the

gap between the mean-field approximation and spatially resolved kMC simulations.[67]

It is well known that adsorbate–adsorbate interactions can significantly change the surface energetics and overall catalytic activity, but the pressing question in computational catalyst screening is: Do lateral interactions alter the predicted reactivity *trends*? This question has been answered for the CO oxidation reaction and the short answer is 'No'. The long answer can be found in ref. 49 and is briefly summarized here. Taking repulsive O–O, CO–CO and CO–O interactions directly into account for adsorbates and indirectly for activation barriers through BEP relations, a coverage-dependent microkinetic model for CO oxidation can be constructed and used for screening purposes. Because the lateral interactions vary across catalysts and no correlation between the interactions and the activity descriptors ΔE_O and ΔE_{CO} could be identified, the volcano in Figure 1.17(a) was calculated under the assumption that the interaction energies for Pt(111) are constant for the whole range of descriptors. The result is that in the low coverage regime there is no difference between coverage corrected and uncorrected volcanoes, and at high coverage the effect of surface poisoning is reduced by the destabilization of surface species. The reduced poisoning effect leads to a higher activity and a broadening of the volcano peak, as clearly seen in Figure 1.17(a), but the position of the peak and the relative order of metals remain constant [not shown here, but a contour plot with finer resolution around the peak is given as Figure 3(c) of ref. 49]. Using catalyst-independent lateral interaction energies is instructive for showing the effect of lateral interactions on the shape of the volcano curve, but the preferred visualization technique is to indicate the shift of the metal positions from their low coverage descriptor values to their high coverage values on the uncorrected

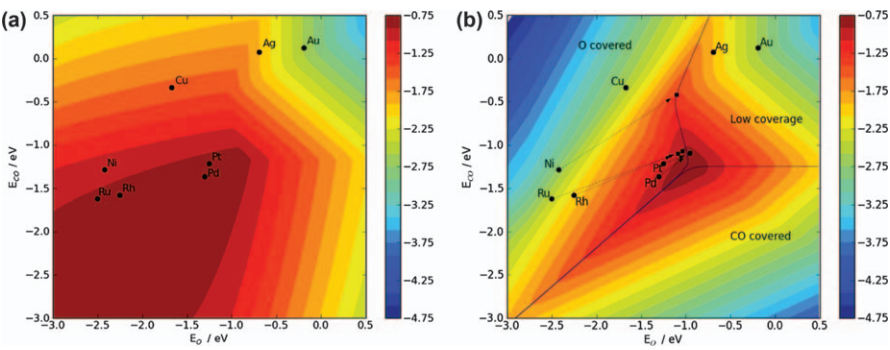


Figure 1.17 The effect of adsorbate–adsorbate interactions on the CO oxidation volcano. (a) Adsorbate interactions characteristic for Pt were included in the generation of the volcano plot. (b) The "standard" activity volcano is shown, but the positions of the metals have been corrected to account for the effect of surface coverage.
L. C. Grabow, *et al.*, 'Understanding Trends in Catalytic Activity: The Effect of Adsorbate Adsorbate Interactions for CO Oxidation Over Transition Metals', *Topic. Catal.*, 2010, **53**, 298–310 is given with kind permission from Springer Science and Business Media.

volcano plot. The result of this technique is shown in Figure 1.17(b). The corrected metal positions all fall onto the line that indicates the transition from the high to the low coverage regime. As long as surface oxide formation can be excluded, you may think of these strongly adsorbing catalysts as operating at their saturation coverage under reaction conditions. The *effective* (or coverage adjusted) descriptor values can then be evaluated at the corresponding surface coverage and used to indicate the coverage-corrected position. While the long answer is fundamentally interesting, the important take-home message is that lateral interactions neither alter the position of the volcano peak nor change the relative order of the activity of metal catalysts.

1.8 Conclusions

This chapter has provided a tutorial style overview of the rapidly emerging field of computational catalyst screening and the reader should feel adequately informed and prepared to attempt a catalyst screening project on his/her own. Although the field is still relatively new, it has already reached such a high level of complexity that several topics could only be mentioned in passing and references for further study had to be provided. The descriptor-based approach relies heavily on assumptions and simplifications, but its success is well founded in the numerous examples of new catalyst discoveries in recent years. It is this success that emphasizes the applicability, strength, and robustness of the method for the search for novel and improved heterogeneous catalysts.

Let us conclude this chapter by stating that computational catalyst screening works in some sense just like a professional dating or matchmaking service. How so? Well, a client (= reaction) who is seeking the help of a professional matchmaker (= us) in order to find her/his perfect partner (= catalyst) will first have to go through an in-depth interview. During the interview, the matchmaker thoroughly analyzes the client's character/personal situation and describes it with certain adjectives (= descriptors), such as attractive (= affinity to adsorbate 'X'), financially stable (= low cost material), emotionally stable (=poisoning resistant), desire to get married (= long-term stability), and so on. Once the matchmaker has developed a full profile of the client, he or she will enter that information into a database and run a search algorithm that systematically screens a database for potential partners. The screening process will yield a list of, perhaps 10, matches that all have a certain compatibility index (= deviation from the optimal descriptor value), and on the following dates (= experimental testing) the client will need to find out whether the love of his/her life (= optimal catalyst) was amongst the identified candidates. The matchmaker's key to success is the creation of an accurate client profile (= identification of the right descriptors), the quality of the matching algorithm (= scaling, BEP, and kinetic model) and the size of the client (= catalyst material) database.

As you can see, there are quite a few parallels between professional matchmaking services and computational catalyst screening. The next time you find yourself in a situation where you need to explain to a non-expert what you do for a living, try starting with: "I am a professional matchmaker and my clients are slow chemical reactions…"!

Appendix

Python Code for the numerical solution of the microkinetic model for CO oxidation and the generation of a two-dimensional volcano plot.


```
#!/usr/bin/env python

import numpy as np

# Reaction conditions
T    = 600      # K
PCO  = 0.33     # bar
PO2  = 0.67     # bar
PCO2 = 1        # bar

# Physical constants and conversion factors
J2eV = 6.24150974E18            # eV/J
Na   = 6.0221415E23             # mol-1
h    = 6.626068E-34 * J2eV      # in eV*s
kb   = 1.3806503E-23 * J2eV     # in eV/K
kbT  = kb * T                   # in eV

def get_rate_constants(EO,ECO):
# This function applies scaling and BEP to determine
all rate constants
# depending the two descriptors EO, ECO provided as
input.

    # Gas phase energies with reference to CO and O2
    ECOg = EO2g = 0
    ECO2g = -3.627           # in eV calculated from DFT

    # Scaling for EO2
    EO2 = 0.89 * EO + 0.17   # eq. (36)

    # Gas phase entropies converted to eV/K
    SCOg  = 197.66 * J2eV / Na          # eV/K
    SO2g  = 205.0 * J2eV / Na
    SCO2g = 213.74 * J2eV / Na

    # Surface entropies are assumed to be zero
    SCO = SO2 = SO = 0

    # Reaction energies
    dE = np.zeros(4)             # array initialization
    dE [0]  = ECO - ECOg
```

```
    dE [1]  = EO2 - EO2g
    dE [2]  = 2*EO - EO2
    dE [3]  = ECO2g - ECO - EO

    # Entropy changes
    dS = np.zeros(4)           # array initialization
    dS [0]  = SCO - SCOg
    dS [1]  = SO2 - SO2g
    dS [2]  = 2*SO - SO2
    dS [3]  = SCO2g - SCO - SO

    # Activation energy barriers from BEP
    Ea = np.zeros(4)           # array initialization
    Ea [2]  = 0.5 * EO + 1.39   # eq. (37)
    Ea [3]  = -0.3 * (EO + ECO) # eq. (38)
            + 0.02

    # Entropy changes to the transition state
    STS = np.zeros(4)          # array initialization
    STS [0]  = -0.25*SCOg      # loss of CO entropy
                                 assumed
    STS [1]  = -0.25*SO2g      # loss of O2 entropy
                                 assumed
    STS [2]  = 0               # surface reaction
    STS [3]  = 0               # surface reaction

    # Calculate equilibrium and rate constants
    K  = [0]*4                 # equilibrium constants
   kf = [0]*4                  # forward rate constants
   kr = [0]*4                  # reverse rate constants
    for i in range(4):
        dG = dE [i]  - T*dS [i]
        K [i]   = np.exp(-dG/kbT)
        # enforce Ea > 0, and Ea > dE
        # independent of what the descriptors are
        Ea [i]  = max([0,dE [i],Ea [i]])
        kf [i]  = kbT/h * np.exp(STS [i]/kb) * np.exp
(-Ea [i]/kbT)
        kr [i]  = kf [i]/K [i]     # enforce thermodynamic
consistency
    return (kf,kr)

def get_rates(theta,kf,kr):
# returns the rates depending on the current coverages
theta
```

```
    # Extract elements of theta and assign them
    # to more meaningful variables
    tCO   = theta [0]                # theta of CO
    tO2   = theta [1]                # theta of O2
    tO    = theta [2]                # theta of O
    tstar = 1.0 - tCO - tO2 - tO     # site balance for
                                       tstar

    # Caluclate the rates: eqns (39)-(42)
    rate    = [0]*4                  # array with 4 zeros
    rate [0]  = kf [0]  * PCO * tstar - kr [0]  * tCO
    rate [1]  = kf [1]  * PO2 * tstar - kr [1]  * tO2
    rate [2]  = kf [2]  * tO2 * tstar - kr [2]  * tO * tO
    rate [3]  = kf [3]  * tCO * tO - kr [3]  * PCO2 * tstar *
tstar

    return rate

def get_odes(theta,t,kf,kr):
# returns the system of ODEs d(theta)/dt
# calculated at the current value of theta (and time t,
not used)

    rate = get_rates(theta,kf,kr)    # calculate the
current rates

    # Time derivatives of theta
    dt  = [0]*3
    dt [0]  = rate [0]  - rate [3]       # d(tCO)/dt
    dt [1]  = rate [1]  - rate [2]       # d(tO2)/dt
    dt [2]  = 2 *  rate [2]  - rate [3]  # d(tO)/dt

    return dt

def print_output(EO,ECO,theta):
# Prints the solution of the model
    (kf,kr) = get_rate_constants(EO,ECO)
    rates   = get_rates(theta,kf,kr)
    print "For the descriptors EO =",EO,"and ECO =",
    ECO,"the result is:"
    print
    for r,rate in enumerate(rates):
        print "Step",r,": rate =",rate,", kf =",kf [r] ,",
kr=",kr [r]
    print
```

```
    print "The coverages for CO*, O2*, and O* are:"
    for t in theta:
        print t

def solve_ode(kf,kr,theta0=(0.,0.,0.)):
# Solve the system of ODEs using scipy.integrate.odeint
# Assumes an empty surface as initial guess if nothing
else is provided
    from scipy.integrate import odeint

     # Integrate the ODEs for 1E6 sec (enough to reach
steady-state)
    theta = odeint(get_odes,         # system of ODEs
                theta0,              # initial guess
                [0,1E6],             # time span
                args = (kf,kr),      # arguments to
                                       get_odes()
                h0 = 1E-36,          # initial time step
                mxstep = 90000,      # maximum number of
                                       steps
                rtol = 1E-12,        # relative tolerance
                atol = 1E-15)        # absolute tolerance

    return theta [-1,:]

def solve_findroot(kf,kr,theta0):
# Use mpmath's findroot to solve the model
    from mpmath import mp, findroot
    mp.dps = 25
    mp.pretty = True

    def get_findroot_eqns(*args):
        return get_odes(args,0,kf,kr)

    theta = findroot(get_findroot_eqns,
                     tuple(theta0),
                     multidimensional=True)
    return np.array(theta)

def solve_analytically(kf,kr):
# Using the analytical solution from section 7, eq.(48)-
(52)
# The equation for W was rewritten to avoid numerical
precision errors
```

```
    x3 = kf[2] * kf[1]/kr[1] * PO2
    y3 = kr[2]
    x4 = kf[3] * kf[0]/kr[0] * PCO
    y4 = kr[3] * PCO2
    numerator = 8*y3*(2*x3+y4)
    denominator = x4**2
    if (numerator < 1.0e-4*denominator):
        W = (x4/(4*y3))*(0.5*(numerator/denominator))
    else:
        W = (x4/(4*y3))*(-1 + np.sqrt(1 + numerator/
denominator))

   tstar = 1. / (1 + W + kf[0]/kr[0] * PCO + kf[1]/kr[1]
* PO2)
    tCO = kf[0]/kr[0] * PCO * tstar
    tO2 = kf[1]/kr[1] * PO2 * tstar
    tO  = W * tstar
    return (tCO,tO2,tO)

def calculate_Xrc(r0,kf0,kr0):
# Calculates Xrc by systematically changing the
# rate constants of each step by 10% around the reference
value

    delta = 0.1                    # change of 10%
    Xrc_rates = np.zeros(4)        # array for storing
                                     rates
    for s in range(4):             # loop over all steps
        # initialize rate constants with reference
values
        kf = kf0[:]
        kr = kr0[:]
        kf[s] = (1 + delta) * kf0[s]
        kr[s] = (1 + delta) * kr0[s]
        # Solve the ODEs with the modified k's
        theta = solve_ode(kf,kr)
        rates = get_rates          # Get the new rates
                (theta,kf,kr)
        Xrc_rates[s] = rates[3]    # Save the rate of CO2
                                     production

    # And calculate Xrc for all steps
    Xrc = (Xrc_rates-r0)/(delta*r0)
    return Xrc
```

```
def calculate_Xcc(r0,EO0,ECO0):
# Calculate Xcc by varying EO and ECO around their
# reference values by 0.05 eV

    delta = 0.05                    # change of 0.05 eV
    Xcc_rates = np.zeros(2)         # array for modified
                                      rates

    # Start by modifying EO0
    (kf,kr) = get_rate_constants(EO0+delta,ECO0)
    theta = solve_ode(kf,kr)        # Solve with the
                                      modified k's
    rates = get_rates(theta,kf,kr) # Get the new rates
    Xcc_rates[0] = rates[3]         # Save the rate of CO2
                                      production

    # Repeat for ECO0
    (kf,kr) = get_rate_constants(EO0,ECO0+delta)
     theta = solve_ode(kf,kr)       # Solve with the
                                      modified k's
    rates = get_rates(theta,kf,kr) # Get the new rates
    Xcc_rates[1] = rates[3]         # Save the rate of CO2
                                      production
    Xcc = (np.log(Xcc_rates)-np.log(r0))/(-delta/kbT)
    return Xcc

# Solve the model for Pt with EO = -1.25, E_CO = -1.22
ECO = -1.22                         # CO oxidation
                                      descriptors
EO  = -1.25

(kf0,kr0) = get_rate_constants      # get the rate
            (EO,ECO)                  constants for the
                                      given descriptor

theta0 = solve_ode(kf0,kr0)         # Solve the model for
the refernce values
print_output(EO,ECO,theta0)         # Print the output

# Sensitivity Analysis
rates = get_rates(theta0,kf0,kr0)
r0 = rates[3]                       # CO2 production rate
                                      as reference

Xrc = calculate_Xrc(r0,kf0,kr0)     # Call the Xrc
                                      function
```

```
print "\nThe Degrees of Rate control are:"
print Xrc

Xcc = calculate_Xcc(r0,EO,ECO)   # Call the Xcc function
print "\nThe Degrees of Catalyst control for EO and ECO
are:"
print Xcc

print "\nStarting to generate 2D data..."
# Generate the data for 2D volcano plots
# and save it in a matrix
# numerical, analytical at the same time
gridpoints = 20
num_data = np.zeros( [gridpoints,gridpoints] )
ana_data = np.zeros( [gridpoints,gridpoints] )

# Generate the initial guess for the first point
(kf,kr) = get_rate_constants(-2.5,-2.5)
theta0 = solve_ode(kf,kr)

# Start the scan through the parameter space
# Alternate directions for alternating columns
EO_range0 = np.linspace(-2.5,0.0,gridpoints)
ECO_range0 = np.linspace(-2.5,0.0,gridpoints)
for i,EO in enumerate(EO_range0):
    if i%2:                                  # true if odd
        ECO_range = ECO_range0.copy()[ ::-1]    # reverses
the list
        j_range = range(gridpoints)[ ::-1]
    else:
        ECO_range = ECO_range0.copy()
        j_range = range(gridpoints)
    for j,ECO in enumerate(ECO_range):
        (kf,kr) = get_rate_constants(EO,ECO)
        theta = solve_findroot(kf,kr,theta0)
        # Check of solution is physical
        if (theta > 1.0).any() or (theta < 0.0).any():
            # Generate a new initial guess using odeint
and solve again
            theta0 = solve_ode(kf,kr)
            theta = solve_findroot(kf,kr,theta0)
        rates = get_rates(theta,kf,kr)
        # Store the solutions
        num_data [j_range [j] ,i] = rates [3]     # the nu-
merical solution
```

```
        theta_ana = solve_analytically(kf,kr)
        ana_data [j_range [j],i]  = get_rates(theta_ana,
kf,kr)[3]
        # Store the numerical solution as
        # initial guess for the next grid point
        theta0 = theta.copy()

# Generate a 2D volcano plot using matplotlib
from matplotlib.pyplot import *
for d,data in enumerate( [num_data, ana_data]):
    figure()
    levels=np.linspace(-30,6,19)

graph=contourf(EO_range0,ECO_range0,np.log10(data),
levels,cmap=cm.jet)
    xlabel(r' $\Delta$E$_O$ / eV')
    ylabel(r' $\Delta$E$_{CO}$ / eV')
    colorbar()
    scatter([-1.67,-1.25,-1.30,-2.25],
            [-0.34,-1.22,-1.37,-1.58],color = 'k')
    dx = 0.04
    dy = 0.04
    text(-1.67+dx,-0.34+dy,'Cu')
    text(-1.25+dx,-1.22+dy,'Pt')
    text(-1.30+dx,-1.37,'Pd')
    text(-2.25+dx,-1.58+dy,'Rh')
    axis([-2.5,0.,-2.5,0.])
    savefig('COox_Volcano_'+str(d)+'.png')
```

References

1. K. Yaccato, R. Carhart, A. Hagemeyer, M. Herrmann, A. Lesik, P. Strasser, A. Volpe, H. Turner, H. Weinberg, R. K. Grasselli, C. J. Brooks and J. M. Pigos, *Combin. Chem. High Throughput Screen.*, 2010, **13**, 318–330.
2. J. Greeley, T. F. Jaramillo, J. Bonde, I. B. Chorkendorff and J. K. Nørskov, *Nature Mat.*, 2006, **5**, 909–913.
3. M. Mavrikakis, *Nature Mat.*, 2006, **5**, 847–848.
4. G. A. Somorjai and Y. Li, *Top. Catal.*, 2010, **53**, 311–325.
5. F. Besenbacher, I. Chorkendorff, B. S. Clausen, B. Hammer, A. M. Molenbroek, J. K. Nørskov and I. Stensgaard, *Science*, 1998, **279**, 1913.
6. C. J. Jacobsen, S. Dahl, B. S. Clausen, S. Bahn, A. Logadottir and J. K. Nørskov, *J. Am. Chem. Soc.*, 2001, **123**, 8404–8405.
7. H. Toulhoat, *J. Catal.*, 2003, **216**, 63–72.

8. P. Strasser, Q. Fan, M. Devenney, W. H. Weinberg, P. Liu and J. K. Nørskov, *J Phys. Chem. B*, 2003, **107**, 11013–11021.
9. A. U. Nilekar, K. Sasaki, C. A. Farberow, R. R. Adzic and M. Mavrikakis, *J. Am. Chem. Soc.*, 2011, **133**, 18574–18576.
10. S. Linic, J. Jankowiak and M. A. Barteau, *J. Catal.*, 2004, **224**, 489–493.
11. J. Greeley and M. Mavrikakis, *Nature Mat.*, 2004, **3**, 810–815.
12. T. F. Jaramillo, K. P. Jørgensen, J. Bonde, J. H. Nielsen, S. Horch and I. Chorkendorff, *Science*, 2007, **317**, 100–102.
13. S. Alayoglu, A. U. Nilekar, M. Mavrikakis and B. Eichhorn, *Nature Mat.*, 2008, **7**, 333–338.
14. F. Studt, F. Abild-Pedersen, T. Bligaard, R. Z. Sorensen, C. H. Christensenand and J. K. Nørskov, *Science*, 2008, **320**, 1320–1322.
15. J. Greeley, I. E. L. Stephens, A. S. Bondarenko, T. P. Johansson, H. A. Hansen, T. F. Jaramillo, J. Rossmeisl, I. Chorkendorff and J. K. Nørskov, *Nature Chem.*, 2009, **1**, 552–556.
16. J. Zhang, M. B. Vukmirovic, K. Sasaki, A. U. Nilekar, M. Mavrikakis and R. R. Adzic, *J. Am. Chem. Soc.*, 2005, **127**, 12480–12481.
17. L. C. Grabow, F. Studt, F. Abild-Pedersen, V. Petzold, J. Kleis, T. Bligaard and J. K. Nørskov, *Ange. Chem. Int. Ed.*, 2011, **50**, 4601–4605.
18. L. C. Grabow, F. Studt, F. Abild-Pedersen, V. Petzold, J. Kleis, T. Bligaard and J. K. Nørskov, *Ange. Chem.*, 2011, **123**, 4697–4701.
19. J. K. Nørskov, T. Bligaard, J. Rossmeisl and C. H. Christensen, *Nature Chem.*, 2009, **1**, 37–46.
20. J. K. Nørskov, F. Abild-Pedersen, F. Studt F and T. Bligaard, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 937–943.
21. T. Bligaard, J. K. Nørskov, S. Dahl, J. Matthiesen, C. H. Christensen and J. Sehested, *J. Catal.*, 2004, **224**, 206–217.
22. M. P. Andersson, T. Bligaard, A. Kustov, K. E. Larsen, J. Greeley, T. Johannessen, C. H. Christensen and J. K. Nørskov, *J. Catal.*, 2006, **239**, 501–506.
23. J. Sehested, K. Larsen, A. Kustov, A. Frey, T. Johannessen, T. Bligaard, M. Andersson, J. Nørskov and C. Christensen, *Topic. Catal.*, 2007, **45**, 9–13.
24. M. Andersson, T. Bligaard, C. H. Christensen, T. Johannessen, A. Kustov, K. E. Larsen, J. K. Nørskov and J. Sehested, "Process and Catalyst For Hydrogenation of Carbon Oxides", US Patent US7790776 B2 (2010).
25. T. R. Munter, D. D. Landis, F. Abild-Pedersen, G. Jones, S. Wang and T. Bligaard, *Comp. Sci. Discov.*, 2009, **2**, 015006.
26. W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.
27. P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.
28. J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1992, **46**, 6671.
29. J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, *Phys. Rev. B*, 1993, **48**, 4978.
30. J. P. Perdew, K. Burke and M. Ernzerhof, M., *Phys. Rev. Lett.*, 1996, **77**, 3865–3868.

31. B. Hammer, L. Hansen and J. Nørskov, *J. Phys. Rev. B*, 1999, **59**, 7413–7421.
32. J. S. Hummelshøj, F. Abild-Pedersen, F. Studt, T. Bligaard and J. K. Nørskov, *Ange. Chem. Int. Ed. Engl.*, 2011, **51**, 272–274.
33. D. S. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction*, Wiley-VCH, Verlag, 2009, vol. 135.
34. B. Hammer and J. K. Nørskov, *Adv. Catal.*, 2000, **45**, 71–129.
35. B. Hammer and J. K. Nørskov, *Surf. Sci.*, 1995, **343**, 211–220.
36. H. Xin and S. Linic, *J. Chem.Phys.*, 2010, **132**, 221101–221104.
37. F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skúlason, T. Bligaard and J. K. Nørskov, *Phys. Rev. Lett.*, 2007, **99**, 016104–016105.
38. E. M. Fernandez, P. G. Moses, A. Toftelund, H. A. Hansen, J. I. Martinez, F. Abild-Pedersen, J. Kleis, B. Hinnemann, J. Rossmeisl, T. Bligaard and J. K. Nørskov, *Ange. Chem.-Int. Ed.*, 2008, **47**, 4683–4686.
39. J. K. Nørskov, T. Bligaard, B. Hvolbaek, F. Abild-Pedersen, I. Chorkendorff and C. H. Christensen, *Chem. Soc. Rev.*, 2008, **37**, 2163–2171.
40. J. Gómez-Díaz and N. López, *J. Phys. Chem. C*, 2011, **115**, 5667–5674.
41. J. Gómez-Díaz, C. Vargas-Fuentes and N. López, *J. Chem. Phys.*, 2011, **135**, 124707.
42. G. Jones, F. Studt, F. Abild-Pedersen, J. K. Nørskov and T. Bligaard, *Chem. Engin. Sci.*, 2011, **66**, 6318–6323.
43. J. K. Nørskov, T. Bligaard, A. Logadottir, S. Bahn, L. B. Hansen, M. Bollinger, H. Bengaard, B. Hammer, Z. Sljivancanin, M. Mavrikakis, Y. Xu, S. Dahl and C. J. H. Jacobsen, *J. Catal.*, 2002, **209**, 275–278.
44. S. Wang, B. Temel, J. Shen, G. Jones, L. C. Grabow, F. Studt, T. Bligaard, F. Abild-Pedersen, C. H. Christensen and J. K. Nørskov, *Catalysis Letters*, 2011, **141**, 370–373.
45. S. Wang, V. Petzold, V. Tripkovic, J. Kleis, J. G. Howalt, E. Skúlason, E. M. Fernández, B. Hvolbæk, G. Jones, A. Toftelund, H. Falsig, M. Björketun, F. Studt, F. Abild-Pedersen, J. Rossmeisl, J. K. Nørskov and T. Bligaard, *Phys. Chem. Chem. Phys.*, 2011, **13**, 20760–20765.
46. J. N. Brønsted, *Chem. Rev.*, 1928, **5**, 231–338.
47. M. G. Evans and Polanyi M., *Trans. Faraday Soc.*, 1938, **34**, 11–24.
48. L. C. Grabow, A. A. Gokhale, S. T. Evans, J.A. Dumesic and M. Mavrikakis, *J. Phys. Chem. C*, 2008, **112**, 4608–4617.
49. L. C. Grabow, B. Hvolbæk and J. K. Nørskov, *Topic. Catal.*, 2010, **53**, 298–310.
50. A. Vojvodic, F. Calle-Vallejo, W. Guo, S. Wang, A. Toftelund, F. Studt, J. I. Martínez, J. Shen, I. C. Man, J. Rossmeisl, T. Bligaard, J. K. Noørskov and F. Abild-Pedersen, *J. Chem. Phys.*, 2011, **134**, 244509.
51. P. Sabatier, *Berichte Deutsch. Chem. Gesells.*, 1911, **44**, 1984–2001.
52. H. Falsig, T. Bligaard, J. Rass-Hansen, A. L. Kustov, C. H. Christensen and J. K. Nørskov, *Topic. Catal.*, 2007, **45**, 117–120.

53. H. Falsig, B. Hvolbæk, I. S. Kristensen, T. Jiang, T. Bligaard, C. H. Christensen and J. K. Nørskov, *Ange. Chem.-Int. Ed.*, 2008, **47**, 4835–4839.
54. H. Falsig, "Understanding Catalytic Activity Trends for NO Decomposition and CO Oxidation using Density Functional Theory and Microkinetic Modeling", PhD Thesis, Technical University of Denmark, Kongens Lyngby, Denmark, 2010.
55. G. Ertl, *Ange. Chem. Int. Ed.*, 2008, **47**, 3524–3535.
56. K. Honkala, A. Hellman, I. N. Remediakis, A. Logadottir, A. Carlsson, S. Dahl, C. H. Christensen and J. K. Nørskov, *Science*, 2005, **307**, 555–558.
57. A. A. Gokhale, S. Kandoi, J. P. Greeley, M. Mavrikakis and J. A. Dumesic, *Chem. Engin. Sci.*, 2004, **59**, 4679–4691.
58. P. Stoltze and J. K. Nørskov, "Theoretical Modelling of Catalytic Reactions" in *Handbook of Heterogeneous Catalysis*, 2nd Edition, Eds.: G. Ertl, H. Knözinger, F. Schüth and J. Weitkamp, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, 2008.
59. J. A. Dumesic, D. F. Rudd, L. L. Aparicio, J. E. Rekoske and A. A. Treviño, *The Microkinetics of Heterogeneous Catalysis*, American Chemical Society, Washington, DC, 1993.
60. H.-J. Freund, G. Meijer, M. Scheffler, R. Schlögl and M. Wolf, *Ange. Chem. Int. Ed.*, 2011, **50**, 10064–10094.
61. C. J. Cramer, *Essentials of Computational Chemistry: Theories and Models*, 2nd edn, John Wiley & Sons Ltd., Chichester, England, 2004.
62. C. T. Campbell, *Topic. Catal.*, 1994, **1**, 353–366.
63. C. T. Campbell, *J. Catal.*, 2001, **204**, 520–524.
64. C. Stegelmann, A. Andreasen and C. T. Campbell, *J. Am. Chem. Soc.*, 2009, **131**, 8077–8082.
65. J. K. Nørskov, T. Bligaard and J. Kleis, *Science*, 2009, **324**, 1655–1656.
66. R. D. Cortright and J. A. Dumesic, *Adv. Catal.*, 2001, **46**, 161.
67. A. Hellman and K. Honkala, *J. Chem. Phys.*, 2007, **127**, 194704.
68. N. İnoğlu and J. Kitchin, *Phys. Rev. B*, 2010, **82**, 1–5.
69. R. B. Getman and W. F. Schneider, *ChemCatChem*, 2010, **2**, 1450–1460.
70. C. Wu, D. J. Schmidt, C. Wolverton and W. F. Schneider, *J. Catal.*, 2012, **286**, 88–94.

CHAPTER 2

First-principles Thermodynamic Models in Heterogeneous Catalysis

J. M. BRAY[a] AND W. F. SCHNEIDER*[a,b]

[a] Department of Chemical and Biomolecular Engineering, 182 Fitzpatrick Hall. University of Notre Dame, Notre Dame, IN 46556, USA;
[b] Department of Chemistry and Biochemistry, 251 Nieuwland Science Hall. University of Notre Dame, Notre Dame, IN 46556, USA
*Email: wschneider@nd.edu

2.1 Introduction

2.1.1 Background

One of the most intriguing aspects of first-principles methods applied to heterogeneous catalysis is the ability to specify the exact constitution of the system – the number and type of atoms and their locations in space. Complete control is held over the system and when results are obtained, whether they be energy, density of states, core electron binding energy, vibrational modes, or any other property, those results are clear and well defined with respect to the specified system. The challenge arises in using these results to understand, interpret, and even predict behavior of experimental systems where the structure is not well defined and is generally quite heterogeneous, measurable properties are averaged over an ensemble of sites, and these characteristics of the system can depend heavily on external conditions and may even change during an experiment or down the length of a reactor. In this chapter we describe tools to

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

connect first-principles results with real, heterogeneous systems, focusing on adsorption at a heterogeneous interface.

Under catalytic reaction conditions a heterogeneous surface is conceived to be in continual flux, with molecules adsorbing, desorbing, diffusing, and reacting.[1] Figure 2.1 illustrates schematically these processes at a catalyst surface. Heterogeneous catalysts are able to bind different molecules to different extents, allowing reactants to adsorb and react, and products to desorb, to complete a catalytic cycle. Adsorption and desorption are central to this surface reactivity, as they are the pathways by which reactants enter and products exit the catalytic cycle. In Langmuir's simplest isotherm model adsorbates bind randomly and independently with fixed adsorption energy at reaction sites.[2] The surface coverage is determined by this adsorption energy, and surface reaction rates are in part a function of this coverage. Reaction rate constants are also related to these adsorption energies. According to the Sabatier principle, an optimal catalyst binds adsorbates with an intermediate binding energy, trading off the tendency to interact with (*e.g.* dissociate) a reactive adsorbate with the ability to desorb a product. This principle plays an important role in understanding reactivity trends in catalytic materials,[1,3] and is often illustrated by so-called volcano curves, which plot reactivity against a descriptor such as the product adsorption energy.[4,5] Consequently, adsorption processes enter into the characterization and screening of catalyst materials in multiple ways. While adsorption properties vary widely across different materials, surface structure and adsorbate coverage also influence adsorption even for the same catalyst material. The catalyst material or composition may in some ways be considered analogous to the coarse focus on a microscope, while adsorbate coverage and catalyst structure can alter adsorption and reactivity over a narrower range similar to the fine focus. However, in reality these various factors are not independent of one another, and ultimately models that accurately capture their effects must be able to incorporate their interrelated nature as well.

Langmuir proposed several adsorption models of varying complexity,[2] but only the simplest of these models are commonly applied today. Although these models are insightful and can be applied in many cases of practical significance,

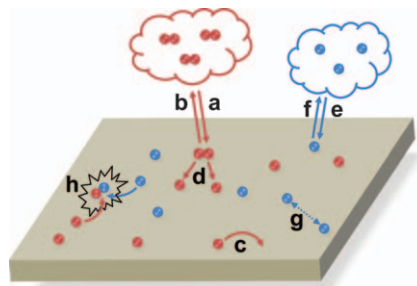


Figure 2.1 Schematic representation of a catalyst surface illustrating: a) molecular adsorption, b) molecular desorption, c) diffusion, d) dissociation, e) atomic adsorption, f) atomic desorption, g) adsorbate–adsorbate interactions, and h) surface reaction.

their primary assumptions – that adsorbates are non- interacting and bind with a constant adsorption energy – are often false for adsorbates on metal surfaces.[6–8] Complicating factors that can cause this Langmuir model to break down include the existence of multiple inequivalent adsorption sites, competition among multiple adsorbed species, incorporation of adsorbates into the surface and/or subsurface lattice, and adsorbate-induced surface reconstructions. Even at a metal surface that retains its integrity as it accumulates adsorbates, adsorbate–adsorbate interactions can cause large deviations from Langmuir behavior.[9] Adsorbates in close proximity tend to interact, leading to coverage-dependent binding energies,[8,10–12] and activation energies sensitive to co-adsorbates.[13] Through-surface electronic effects,[14] surface strain,[12] and through-space electrostatics[15] are but a few of the many mechanisms by which adsorbates can interact.

A complete adsorption model should not only be able to incorporate the ability of adsorbates to bind in a number of different surface (or subsurface) sites with distinct binding energies, but it should also address the energetic consequences of adsorbates interacting with one another and the macroscopic manifestations of these effects. Modern supercell density functional theory (DFT)[16,17] provides access to quantitative descriptions of surface adsorption and can even describe the adsorbate–adsorbate interactions that cause deviations from ideal adsorption behavior. By itself, however, supercell DFT is neither suitable for searching over all adsorbate arrangements possible at a surface nor for constructing ensemble averages over these many configurations present at a heterogeneous surface under reaction conditions.

Were the exact nature and origin of adsorbate–adsorbate interactions well understood for a given system *a priori*, they could be captured by a specific, phenomenological model. In this chapter we describe an approach that requires no prior knowledge of the underlying interaction mechanisms. The DFT results for a number of adsorbate configurations are used to parameterize a cluster expansion (CE) model,[18–27] a type of Ising Hamiltonian, that reduces the adsorption energy to a function of adsorption site occupancies and site–site interactions. Such CE models can give significant insight into the dominant interactions controlling the behavior of the system. Traditional cluster expansions are well suited to capturing short-range interactions but can be elaborated to capture long-range, *e.g.* electrostatic, ones as well.[28] Once in hand, this CE Hamiltonian can be used in equilibrium or non-equilibrium Monte Carlo simulations of extended surface–adsorbate systems to connect system properties with external conditions of temperature and chemical potential.[19,21,27,29,30]

As summarized in Figure 2.2, in this chapter we describe a framework combining these methods to create a comprehensive model that can be used to describe adsorption at a fluid–solid interface. We first explain the major aspects of setting up and characterizing the adsorption system within a widely accepted slab model DFT framework. We then discuss how to initialize and iteratively improve a configuration-dependent CE model of the system until self-consistency is reached. We show how the self-consistent adsorption model

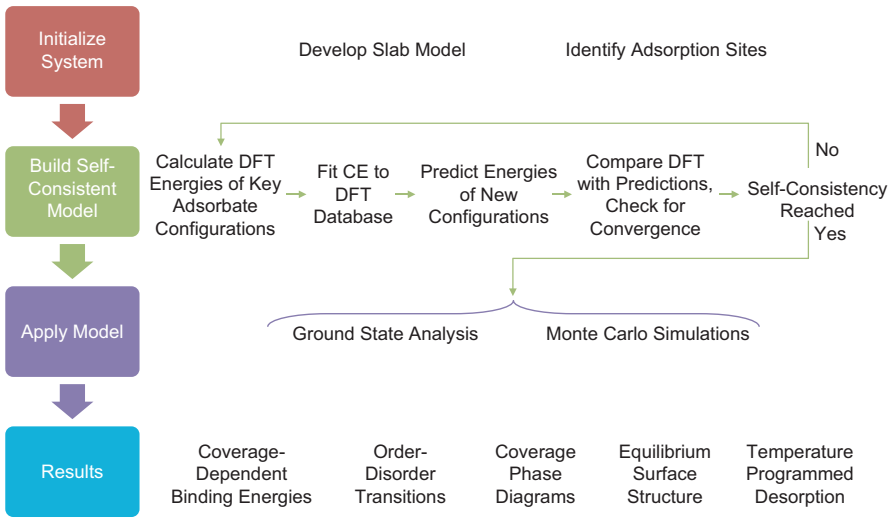


Figure 2.2 Diagram summarizing the first principles approach described in this chapter.

can be applied to understand important thermodynamic features of the system, including ordered ground state structures, equilibrium coverages, surface coverage phase diagrams, order–disorder transitions, and even temperature-programmed desorption (TPD). Lastly, we comment on how these tools can be used to develop coverage-aware kinetic models. We illustrate this procedure and its application using a O–Pt(321) model, a simple, yet non-trivial example of a single atomic adsorbate on a surface possessing atomic-scale heterogeneity.

2.1.2 Background on Oxygen Adsorption on Platinum

Platinum is an effective catalyst for oxidation of H_2, CO, and NO with O_2.[31–35] Pt activates O_2 to form atomic O, but in keeping with the Sabatier principle, does not bind O so strongly that it is unable to react further. The binding strength of O on Pt surfaces has been found to depend strongly on the coverage of O on the surface in both experimental and computational work. Experimentally, coverage-dependent binding energies have been deduced using techniques including temperature programmed desorption (TPD)[36–38] and adsorption calorimetry,[39] and computational studies utilizing periodic DFT calculations of adsorption at a Pt(111) surface agree qualitatively with experimental observations.[7,40] Given that, as mentioned above, the O binding strength can be used as an indicator of reactivity, it is important first to understand the effects of surface structure and coverage on adsorption. The strong coverage effect observed for the O–Pt system makes it an interesting model system for more detailed study.

The questions of structure sensitivity and the "materials gap"[41–44] are often cited as reasons for studying low symmetry or "defect" surfaces, with steps,

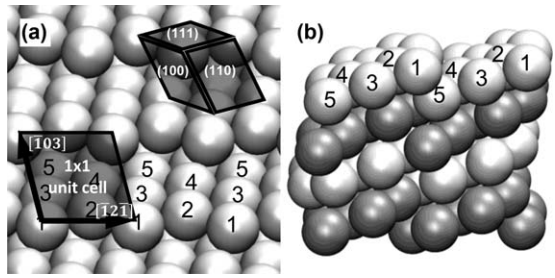


Figure 2.3 Pt(321) surface structure shown from (a) above and (b) the side. Unique Pt atoms are labeled 1–5 beginning with the kink and moving back across the terrace. Metallic coordination values of each Pt are 6, 8, 9, 10, and 11, respectively.
(Adapted with permission from *Langmuir*, 2011, **27**(13), pp. 8177–8186. Copyright 2011 American Chemical Society.)

kinks, or corners.[45] These structural features introduce additional binding sites, often with unique adsorption properties and altered reaction pathways that may change the behavior of the overall surface. Much of the literature relating to the adsorption of oxygen on platinum has been focused on the Pt(111) surface and, as a result, most of the models and understanding of this system have been developed for this specific surface structure. In general, adsorption results for one surface facet do not transfer to all other facets, so it is not clear with what reliability Pt(111) results can be extended to less well-defined systems, particularly supported catalyst nanoparticles. From a pedagogical standpoint as well, low symmetry surfaces present an additional challenge and opportunity to develop first-principles methodologies with enough versatility to handle the additional complexity resulting from the lower symmetry.

To demonstrate the versatility of the coverage-dependent thermodynamic models described in this chapter, as well as to gain insight into the interrelated effects of coverage and structure in O adsorption on Pt, we describe in this chapter results for O adsorption at the Pt(321) surface. The (321) facet exposes steps, kinks, and close-packed terraces with five unique surface atoms, shown in Figure 2.3 and numbered 1–5 beginning at the kink atom and moving back across the terrace.

2.2 Setting up the System

2.2.1 Developing a Slab Model

Periodic DFT codes define a system within a supercell of a particular shape and size that is repeated infinitely in all three dimensions. It is standard practice in simulations of gas–solid interfaces to define a layer, or slab, of atoms and a layer of vacuum within the periodically replicated supercell, such that the physical model is of alternating layers of solid and vacuum. By tuning the thickness of the vacuum, the solid slabs can be separated to the point where

they no longer interact with one another, effectively resulting in an isolated solid surface. The slab thickness can also be tuned to achieve a balance of realistic semi-infinite behavior and computational efficiency.

To determine the positions of atoms in the slab, the bulk lattice constant and atomic positions are optimized using computational parameters consistent with those to be used later for the surface. The bulk structure is then truncated at the desired surface termination. Periodic surfaces are naturally described by two surface vectors and the surface normal, and this coordinate system will usually differ from that used to define the bulk. Use of this new surface coordinate system requires that atomic coordinates in the bulk coordinate system be converted by a series of matrix operations that maintain their relative positions. To compensate for the finite thickness of the slab, the coordinates of some number of layers on one side of the slab are often fixed to their bulk positions, thus mimicking the constraints of a semi-infinite surface.

It costs energy to create a surface, and the surface atoms will generally relax in some way to minimize this energy. Using DFT, it is possible to quantify this relaxation in at least two ways. First, relaxed bond lengths can be compared with the ideal bulk bond lengths. Second, the surface energy, γ, can be calculated for the system with atoms fixed in the bulk positions and again with the surface relaxed, and these energies can be compared. The surface energy can be calculated for a symmetric slab (both sides of the slab either fixed or relaxed, but not one of each) according to eqn (2.1), where E_{slab} is the energy of the slab (either relaxed or fixed) with N_{atom} atoms in it, E_{bulk} is the energy per atom of the bulk material, A_{cs} is the cross-sectional area of one side of the slab, and the extra factor of 2 is required because there are two surfaces of area A_{cs} created.

$$\gamma = \frac{E_{\text{slab}} - N_{\text{atom}} E_{\text{bulk}}}{2 \cdot A_{\text{cs}}} \tag{2.1}$$

The thickness of the slab and the number of layers being fixed or relaxed will also influence the surface energy. Thus, surface energy is one method for determining an appropriate slab setup by calculating the surface energy for a number of different slab thicknesses or number of relaxed layers and finding the slab where the surface energy converges and reaches some limiting value.

We used the Vienna Ab-initio Simulation Package (VASP) code[46] for all DFT calculations described in this chapter. The interactions between valence and core states were treated using the projector-augmented wave (PAW) method with frozen cores,[47,48] electron–electron interactions described within the PW91 generalized gradient approximation (GGA),[49] and planewaves included to a cutoff of 400 eV. An 8×6×1 Monkhorst-Pack k-point mesh[50] was used for the smallest Pt(321) unit cell (denoted 1×1), corresponding to approximately 40 k-points/Å^{-1} in each of the surface directions.

We previously reported surface energy calculations for Pt(321) slabs ranging from two to twelve layers, corresponding to slab thicknesses between 5 and 32 Å, respectively. Surface energies fluctuated with the number of Pt layers, but beyond six layers fluctuations fell below 0.7 meV/A^2. From this analysis, we

found a relaxed Pt(321) surface energy of 110.2 meV/A^2, approximately 20% greater than the 93.327 meV/A^2 for Pt(111).[7] The surface relaxes most at the sites with lowest coordination. In particular, the step atoms are pulled down slightly toward the layer beneath and in towards atoms 3, 4, and 5. The Pt–Pt bond lengths were reduced by as much as 0.170 Å from the bulk value of 2.819 Å (a 6% change). These relaxations have a surface energy benefit[51] of approximately 10 meV/$Å^2$. For all subsequent calculations mentioned here, we used a four-layer slab model of the Pt(321) surface, shown in Figure 2.3b with different layers highlighted, with the bottom two layers fixed in the bulk positions.

2.2.2 Identifying and Characterizing Adsorption Sites

In contrast to physisorption, chemisorption creates localized and directional chemical bonds between the adsorbate and the surface. Chemisorption occurs at specific geometric features of the surface, known as sites. Adsorbates can often bind in more than one type of site, and it then becomes important to identify not only which sites can bind the adsorbate, but also the relative energetics of these binding sites, which influences the likelihood of each site to be populated by the adsorbate.

2.2.2.1 Potential Energy Scan

When investigating a surface with little known about the preferred location of adsorbates, we can obtain an initial glimpse of the potential energy landscape by performing a potential energy scan. Because the scan is only an initial pass in order to get enough information to guide subsequent investigation, quick single point energy calculations without structural relaxations can be performed with relatively low precision computational parameters in order to save time. A grid is set up in the surface plane, often at a fixed distance above the highest surface atom (about 1.9 – 2 Å for O atoms on a Pt surface for example), an adsorbate is placed at each grid point in turn, and the energy is calculated. By observing the relative energies of the adsorbate at each grid point, this process will give a picture of where adsorbates will bind strongly and weakly. If the results of the scan are not definitive, a finer grid may be needed, or the distance above the surface may be inappropriate (either too close or too far) for the particular adsorbate. For oxygen adsorption at the Pt(321) surface, a coarse surface scan revealed about five low energy regions of the surface likely to contain binding sites.

2.2.2.2 Direct Testing of Candidate Sites

While a systematic potential energy scan can be beneficial in cases where little is known about the system in advance and as a way to reduce user bias in exploring possible binding sites, it is not always necessary. High-symmetry,

close-packed metal surfaces, for example, are widely studied, and the types of possible binding sites on these surfaces are generally known.

In exploring the Pt(321) surface, we tested the five potential binding sites discovered from the potential energy scan, and we also tested additional sites, such as hollow or bridge sites, based on intuition and information about other similar systems.[7,52] We tested these sites by manually calculating the coordinates of candidate sites to use as initial guesses and allowing the system to relax to the actual local minima. In this way, we identified six 3-fold hollow sites and three bridge sites capable of binding oxygen stably. An additional atop site was also discovered in the course of later investigations, not discussed here, for a total of 10 binding sites for atomic O on Pt(321),[51] shown in Figure 2.4.

2.2.2.3 *Bidentate Species and Cases of Multiple Conformers*

Larger molecular adsorbates add additional degrees of freedom to the process of identifying important adsorption sites, but the same principles of systematic potential energy scans and direct testing of candidate sites can be combined with additional, potentially automated, searching of different adsorbate conformations, such that a thorough search of the surface can still be accomplished. The methods described in this chapter can be generalized to incorporate large adsorbates and multi-dentate species, but the additional degrees of freedom may of necessity require modifications or extensions of the existing tools and methodology, and this is beyond the scope of the present discussion.

2.2.2.4 *Verifying Stability of Adsorbates*

Occasionally, owing to symmetry constraints or factors relating to the minimization technique, structural optimizations will converge to positions that are not true potential energy minima. Given this, it is important always to verify

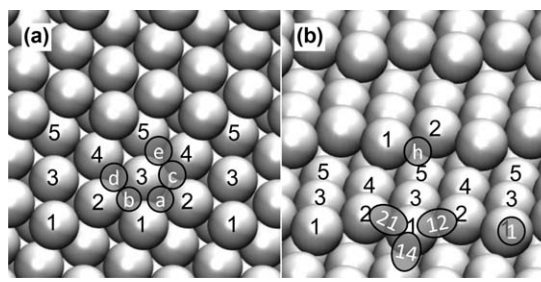


Figure 2.4 (a) Top view or (b) tilted view of Pt(321) surface with 10 stable oxygen adsorption sites identified. Pt atoms are also labeled 1–5 as in Figure 2.3. Hollow sites are labeled with letters, and bridge sites are labeled according to the numbers of their adjacent Pt atoms.
(Adapted with permission from *Langmuir*, 2011, **27**(13), pp. 8177–8186. Copyright 2011 American Chemical Society.)

the stability of a potential binding site by performing a frequency analysis and calculating the normal vibrational modes of the adsorbate and surface.[17] Most DFT packages have the capability to determine these modes within the harmonic approximation. The force constant, or Hessian, matrix is typically constructed by finite differences on the atomic forces evaluated systematically over small Cartesian displacements of all atoms away from their minimum energy positions. The eigenvalues of the mass-weighted force constant matrix are the normal mode force constants, and the eigenvectors are the normal modes themselves. Actual stable binding configurations will have all positive force constants, corresponding to all real vibrational frequencies, while any imaginary modes (negative eigenvalues) indicate that the system can be relaxed further to a more stable state. Additionally, vibrational modes obtained from this analysis can be compared with experimental vibrational spectroscopy results, where available, to look for agreement with the locations of observed vibrational peaks. The stability of each adsorption site in Figure 2.4 was verified by calculating the vibrational modes of the adsorbates and their neighboring metal atoms to ensure that there were no imaginary modes. We give the primary vibrational frequencies for O in each binding site in Table 2.1, and we have shown that these frequencies agree well with the experiments of McClellan *et al.*[53,54] The normal mode frequencies, ν_i, also allow us to calculate the zero-point vibrational energy, E^{ZP}:

$$E^{ZP} = \sum_{i}^{\text{modes}} \frac{1}{2} h\nu_i \tag{2.2}$$

where h is Planck's constant. The zero-point energies for isolated oxygen in each adsorption site are also given in Table 2.1 (below).

Table 2.1 Characterization of each stable adsorption site for O–Pt(321) system; ΔE_{avg} is the average adsorption energy, ΔE_f is the formation energy, and E^{ZP}_{O*} is the zero-point vibrational energy.

Site label (see Figure 2.4 and ref. 51)	*Site type*	ΔE_{avg} *(eV/O)*	ΔE_f *(eV/Pt)*	*Harmonic frequencies* (cm^{-1})	E^{ZP}_{O*} *(eV/O)*
b21	fcc-like bridge	−1.457	−0.0728	545, 480, 154	0.0730
ha	fcc hollow	−1.339	−0.0670	456, 379, 304	0.0707
b12	hcp-like bridge	−1.286	−0.0643	567, 437, 118	0.0695
hb	hcp	−1.140	−0.0570	472, 352, 204	0.0637
b14	(100) bridge	−1.114	−0.0557	576, 463, 145	0.0734
hd	fcc	−1.073	−0.0536	425, 387, 372	0.0734
hh	(110) hollow	−0.887	−0.0443	459, 412, 281	0.0714
hc	hcp	−0.783	−0.0391	434, 355, 329	0.0693
a1	atop	−0.773	−0.0387	795, 67, 30	0.0553
he	fcc	−0.579	−0.0290	471, 419, 347	0.0767

2.2.2.5 Average Adsorption Energies

A molecular adsorption process can be written schematically as:

$$A(g) + {}^{*} \rightarrow A^{*} \tag{2.3}$$

where A(g) is the adsorbate in the gas phase, * represents a vacant surface site, and A* is the adsorbate bound at the surface site. We define the adsorption energy, ΔE_{ads} to be the reaction energy of the above adsorption reaction:

$$\Delta E_{ads} = E_{A^*} - E_{A(g)} - E_{*} \tag{2.4}$$

Consistent computational parameters must be used when calculating each of the above energies in order to cancel systematic errors inherent in supercell implementations of DFT. The resulting adsorption energy can be normalized either to the number of adsorbates or to the total number of sites (*i.e.* surface area). Each of these types of normalization is useful in different contexts, but we begin by discussing the normalization by number of adsorbates, called an average adsorption energy.

Average adsorption energies are the most common measure of the binding strength of adsorbates. Physically, this energy change corresponds to the total internal energy per adsorbate that would be liberated by N_A adsorbates binding in a given configuration, $\boldsymbol{\sigma}$. The average adsorption energy is defined formally in eqn (2.5), where we use ϕ explicitly to represent the specific size and shape of the supercell.

$$\Delta E_{avg}(\boldsymbol{\sigma}) = \frac{1}{N_A(\boldsymbol{\sigma})}\left[E_{\boldsymbol{\sigma}}^{DFT}(\phi) - E_{clean}^{DFT}(\phi)\right] - E_{A(g)}^{DFT} \tag{2.5}$$

In this equation N_A ($\boldsymbol{\sigma}$) is the number of adsorbates in the supercell, $E_{\boldsymbol{\sigma}}^{DFT}(\phi)$ is the energy of the surface with N_A adsorbates in configuration $\boldsymbol{\sigma}$, $E_{clean}^{DFT}(\phi)$ is the energy of the clean surface with no adsorbates, and $E_{A(g)}^{DFT}$ is the energy of the adsorbate in the gas phase. Negative average adsorption energies imply stability relative to the adsorbates in the gas phase, assuming all adsorbates are adsorbed to the surface simultaneously. However, in a real system adsorption occurs in a sequential manner, and the surface may reach saturation – where it becomes endothermic to adsorb a single additional adsorbate – while average adsorption energies remain negative. We use the average adsorption energies of various adsorbate configurations at a given coverage to identify the relative stability of each configuration, with the lowest average adsorption energy being most stable.

Table 2.1 reports average adsorption energies for each of the 10 O adsorption sites on Pt(321) in order of increasing average adsorption energy, calculated in a 2×2 supercell with one adsorbate per supercell ($N_A = 1$) to simulate isolated adsorbates.[51] Atomic energies cannot be accurately evaluated within the

standard DFT formalism and, in the case of O, are not the practically relevant quantity. Here we report adsorption energies relative to $\frac{1}{2}O_2$; $E^{DFT}_{A(g)} = \frac{1}{2}E^{DFT}_{O_2(g)}$.

$$\frac{1}{2}O_2(g) + {}^* \rightarrow O^* \tag{2.6}$$

From Table 2.1, we see that the "b21" site (Figure 2.4) is preferred over all others at low coverage.

2.2.3 Increasing Coverage

The results in Table 2.1 indicate the relative stability of individual adsorption sites but, in general, multiple sites could be occupied at any given coverage. Enumerating and comparing all of the possibilities, even for relatively small supercells, becomes a combinatorial problem that can only be tackled by being able to select intelligently which of the hundreds, or even thousands, of adsorbate configurations should be calculated. Although assumptions regarding which combinations of sites will be most energetically favorable can be made on the basis of isolated adsorption energies such as those in Table 2.1, owing to the potentially significant interactions among adsorbates in nearby sites, the relative stability of each site is subject to change depending on its proximity to other nearby adsorbates.

2.2.3.1 Defining Coverage

The term "coverage" is generally used to refer to the number of occupied sites per number of total sites, but this definition is ambiguous when the number of "sites" is not agreed upon or is altogether unknown. In experiments, coverages are sometimes defined with respect to some maximum, or "saturation," coverage,[53] however this presents two problems. First, the saturation coverage is not necessarily a constant under all conditions, so the coverage scale will vary depending on the conditions under which an experiment is done. Second, it is difficult to relate experimental coverages defined relative to saturation to coverages defined relative to a specific number of sites when the relationship between those two quantities is unknown. Even between computational studies, opinions differ as to the proper way to count sites. Some experiments also report coverages on a per area basis (*e.g.* # adsorbates per cm^2),[53] which is well defined, but does not allow for use of the convenient "monolayer" (ML) notation. Area-based measures of coverage are important when comparing coverages across different surface facets. Here we are concerned with a single surface facet, and to avoid the ambiguity of reporting coverage on a "per site" basis but still preserving the monolayer convention, we choose to define coverage (θ) as the number of adsorbates (N_A) per surface atom (N_M):

$$\theta(\boldsymbol{\sigma}) = \frac{N_A(\boldsymbol{\sigma})}{N_M} \tag{2.7}$$

Defining coverage in this way, $\theta = 1$ ML represents the surface with one adsorbate for each surface atom, irrespective of the number of adsorption sites considered.

2.2.3.2 *Manually Testing Adsorbate Configurations at Higher Coverages*

When little is known in advance about the coverage dependence of the system, or when numerous types of surface or subsurface sites exist, it is possible to test coverage dependence systematically by hand for a limited number of configurations and coverages first and then use those results further to refine subsequent efforts. This is a combinatorial problem, where on a given supercell each combination of adsorbates in two sites is tested, then in three sites, and so on. When numerous types of adsorption site exist, as is the case with 10 adsorption sites for O–Pt(321), or for large enough supercells, chemical intuition can sometimes be applied to reduce the number of calculations to be performed. For example, when two adsorption sites are extremely close together, it may not be necessary to consider cases where both sites are occupied together. However, as we show for the O–Pt(321) case, even making this simplifying assumption is not always sufficient to make the problem tractable.

In the O–Pt(321) system, we performed calculations testing co-adsorption of oxygen in multiple sites on the surface. We began by enumerating by hand all possible combinations of two, three, four, and five sites only for the smallest supercell case (with 10 sites). The number of combinations of n occupied sites ($n = 2,3,4$, or 5) from N possibilities ($N = 10$) is given by the expression:

$$\frac{N!}{n!(N-n)!} \tag{2.8}$$

Even for the smallest Pt(321) supercell, this resulted in over 600 combinations of sites to test, which is far too expensive to be practical. While many of these combinations can be eliminated owing to some sites being too close together to be occupied simultaneously, the problem of too many possible options remains and only gets worse if we then consider larger supercells.

Rather than attempting to calculate every possible combination, we followed a procedure where, still for the smallest supercell, we tested every combination of the one or two lowest energy single sites (“b21” and “ha”) with each of the remaining nine sites. From those tests, we then chose the one or two lowest energy pair configurations and tested them with each of the remaining eight sites, and so on. While clearly limited in its scope, we developed an approximate picture of which sites are more likely to become occupied with changing coverage. This also provided a starting point from which we were able to expand our search to larger supercells.

Having calculated the DFT energy of a number of adsorbate configurations at different coverages and on different supercells, we calculated the average adsorption energy for each configuration, and these are plotted as a function of

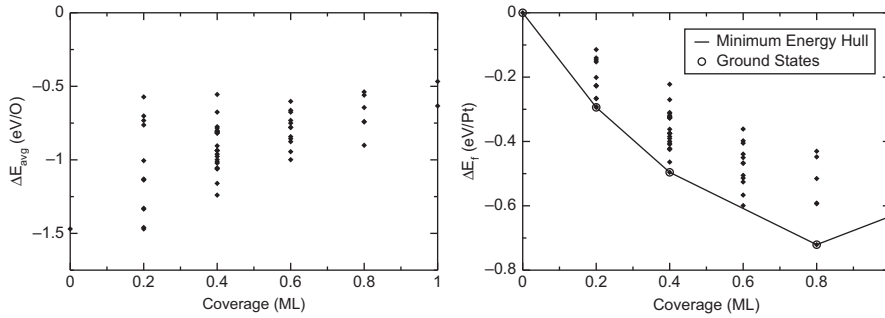


Figure 2.5 Plots of (left) average adsorption energy *vs.* coverage and (right) formation energy *vs.* coverage of several configurations of O on a 1×1 Pt(321) supercell.

coverage in the left panel of Figure 2.5. The average adsorption energy tends to increase with coverage, reflecting an overall repulsive interaction between O atoms at the metal surface. In a non-interacting, single-site system, all configurations of adsorbates at any coverage would result in the same average adsorption energy, while in systems where strong attractive interactions exist, the average adsorption energy would decrease with coverage.

2.2.3.3 Formation Energies

The formation energy of a system is the energy required to create, or form, the system from two or more reference systems. The reference systems should be selected in consideration of the particular application for which the thermodynamic analysis is being performed. In the case of adsorption, we often choose the reference systems to be the clean catalyst surface and the adsorbate in the gas phase, and the formation energy is distinguished from the average adsorption energy in that it is normalized to the surface area rather than the number of adsorbates. With this choice of reference states, the formation energy physically represents the energy change per unit area due to adsorbing a certain number of gas-phase adsorbates to the clean surface into a given configuration, $\boldsymbol{\sigma}$. We define the formation energy as:

$$\Delta E_{\mathrm{f}}(\boldsymbol{\sigma}) = \frac{1}{N_{\mathrm{M}}(\phi)}\left[E_{\boldsymbol{\sigma}}^{\mathrm{DFT}}(\phi) - E_{\mathrm{clean}}^{\mathrm{DFT}}(\phi) - N_{\mathrm{A}}(\boldsymbol{\sigma})E_{\mathrm{A(g)}}^{\mathrm{DFT}}\right] \tag{2.9}$$

Here, N_{M} is the number of surface atoms in the top layer and is proportional to the surface area. The number of surface atoms is commonly used in place of area for normalizing reaction rates and turnover frequencies in catalysis, and we use it here for convenience in the discussion later in the chapter. In this formulation, exothermic binding is signified by a negative formation energy and a positive formation energy indicates endothermic binding.

The formation energies of our preliminary set of O–Pt(321) configurations are plotted in the right panel of Figure 2.5. In this case, the formation energies

first decrease with increasing coverage, but then reach a minimum and begin to increase. The concave up shape of this plot is characteristic of systems with repulsive adsorbate–adsorbate interactions. A formation energy plot for a non-interacting, single-site system would collapse to a straight line, while for a system with attractive interactions, the plot would exhibit a concave down trend.

We calculated the formation energy of adsorbing isolated O atoms into each of the ten adsorption sites. These energies are reported after the average adsorption energies in Table 2.1, and the relative energetic ordering for the 10 sites is the same for both the formation energy and the average adsorption energy, as expected. The average adsorption and formation energies in eqns (2.5) and (2.9) differ only in that the former is normalized to the number of adsorbates (N_A) and the latter to the number of surface atoms (N_M). These can be related using eqn (2.7) to show:

$$\Delta E_{avg}(\boldsymbol{\sigma}) = \frac{\Delta E_f(\boldsymbol{\sigma})}{\theta(\boldsymbol{\sigma})} \qquad \text{or} \qquad \Delta E_f(\boldsymbol{\sigma}) = \theta(\boldsymbol{\sigma}) \cdot \Delta E_{avg}(\boldsymbol{\sigma}) \tag{2.10}$$

2.3 Developing a Self-consistent Cluster Expansion Model

In the previous section, we used DFT calculations to identify potential adsorption sites and evaluate different combinations of sites that may be stable at higher coverages. Examining the preliminary results in Figure 2.5, it is clear that the energy of adsorbate configurations at the same coverage can vary widely. Identifying the adsorption sites that are consistently occupied in the lowest energy structures at each coverage, we found that only 5 out of the original 10 sites were used, namely the "b21", "b14", "ha", "hd", and "hh" sites (see Figure 2.4). These five sites do not correspond to the five lowest energy adsorption sites considered in isolation (compare Table 2.1). Choosing sites in order of their isolated adsorption energy yields sites that are unphysically close. Our preliminary tests of multi-site adsorption configurations over a wide range of coverages were essential to select these five key sites appropriately.

We now describe the construction of a cluster expansion (CE) model fitted to these and other DFT results to predict the relative energy of any adsorbate coverage and configuration.

2.3.1 Cluster Expansion Fundamentals

A cluster expansion is a lattice-based polynomial expansion where a system property of interest, in this case the formation energy, is fitted to a series of site–site interaction terms:

$$E_{\boldsymbol{\sigma}}^{CE} = J_0 + \sum_{i}^{\text{sites}} J_i \sigma_i + \sum_{i>j}^{\text{pairs}} J_{ij} \sigma_i \sigma_j + \sum_{i>j>k}^{\text{3-body}} J_{ijk} \sigma_i \sigma_j \sigma_k + \cdots \tag{2.11}$$

Each term of the expansion consists of an interaction coefficient, J, a particular arrangement of sites (*e.g.* single, pair, 3-body, 4-body, *etc.*) called a cluster, and site occupation variables, σ_i, for each site in the cluster. The sum over pairs is written as $i > j$ and over 3-body clusters as $i > j > k$ to avoid double-counting of identical clusters. Clusters that are equivalent by symmetry are counted separately. The $\boldsymbol{\sigma}$ with no subscript i represents a vector of occupancies, σ_i, of each site in the lattice. Figure 2.6 illustrates several examples of pair and 3-body (or triplet) clusters that can be constructed on a five-site lattice representing the (321) surface of a face-centered cubic metal. In general, clusters can include any number of sites in any arrangement, but in practice most adsorbate CEs can be fitted using primarily pair and 3-body terms.[26,55]

The site occupation variables, σ_i, take on a value of 1 or -1, following the Ising convention, and in this application of the cluster expansion to a two-dimensional (2D) surface–adsorbate system, 1 represents an occupied site and -1 represents a vacant site. Eqn 2.11 is valid for a system with any number of non-equivalent lattice sites, as long as each site is occupied by the same pair of species (in this case the adsorbate and vacancy). Generalizations can be made to multi-component systems with three or more species per site or multi-sublattice systems where different sites may contain different pairs of species,[18,56] but we do not discuss these elaborations in this chapter.

In practical applications, the expansion is truncated to include only those site–site interactions that contribute most significantly to the energy. Thus fitting the CE to a database of configurations (structures) and energies involves two steps, first deciding which interactions (clusters) are significant and should be included, remembering that the number of terms in the expansion must be

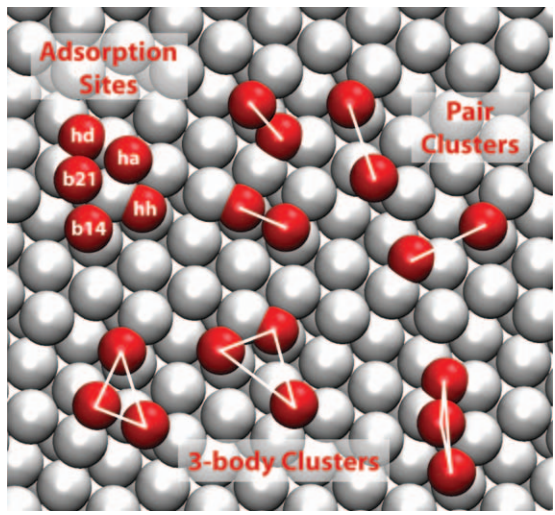


Figure 2.6 Schematic representation of potential pair and three-body clusters that can be formed for a five-site lattice model of the face-centered cubic (321) facet; white = (321) surface, red = O in lattice sites.

significantly less than the number of structures in the fitting database to avoid overfitting, and then using a least-squares regression technique to fit the coefficients to the database of energies. The values of these coefficients, referred to as *effective cluster interactions* or ECIs, are closely related to the energetic contribution of each cluster to the overall energy. As such, the ECIs can provide insight as to which site–site interactions play a major role in determining the overall system energy.

The quality of fit of a CE is measured not only by how well it fits the database energies, but also how well it can predict energies not in the database. To estimate this predictive error, a "leave-one-out" cross-validation score (CV score) is used to measure the quality of fit.[57–60] For a configuration database with n structures, the CV score is calculated by first fitting a CE to $n - 1$ of the database energies, using this CE to predict the energy of the excluded structure, and comparing this prediction to the known energy. This is then repeated for each of the remaining database structures, and the CV score is defined as the root mean square of all n errors generated by the "leave-one-out" fitting:[55]

$$\text{CV score} = \left(\frac{1}{n} \sum_{i}^{n} \left(E_i - \hat{E}_{(i)} \right) \right)^{\frac{1}{2}} \tag{2.12}$$

Here E_i is the DFT energy of structure i and $\hat{E}_{(i)}$ is the predicted energy of structure i using the CE fit without that structure.

Once a satisfactory CE is obtained, it can be applied as a Hamiltonian in lattice-based Monte Carlo (MC) algorithms.[61] This feature of the CE model makes it particularly valuable for predicting thermodynamic properties, as discussed in detail in Section 2.4.3.

2.3.2 Self-consistent Fitting Approach

We used the Alloy Theoretic Automated Toolkit (ATAT)[55,56,59,61] to develop a cluster expansion model of the formation energy of oxygen adsorption at the Pt(321) surface. This suite of codes interfaces with VASP or other DFT packages and automates both the generation of a database of DFT energies for various adsorbate configurations and the fitting of a CE to these energies. For the reduced five-site model, we developed a reliable CE for O adsorption at Pt(321) following an iterative fitting approach outlined in Figure 2.7 and described further in this section.

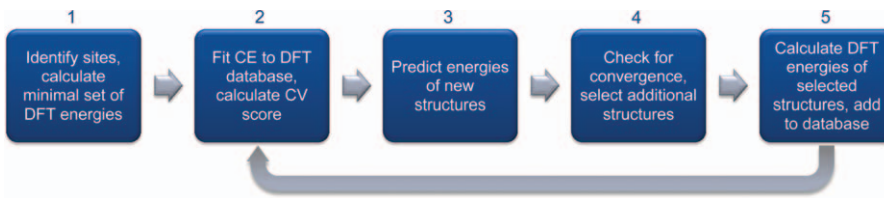


Figure 2.7 Flow chart illustrating the self-consistent cluster expansion fitting approach described in Section 2.3.2.

2.3.2.1 Initialization and Iterative Improvement of Cluster Expansion

As illustrated in Figure 2.7, the system must first be initialized by defining which adsorption sites will be included in the model and calculating energies of a relatively small number of potential adsorbate configurations to form an initial fitting database. The number of structures required for this initial fit depends on the number of sites and the symmetry of the system, and it may range from 4 or 5 structures for a one-site model on a hexagonal lattice to over 20 structures for our five-site O–Pt(321) model. It is most efficient in terms of computational expense to calculate structures for this initial database beginning with the smallest (1×1) unit cell.

Rather than determining possible structures by hand, we use the ATAT software to enumerate systematically all possible unique configurations of adsorbates, beginning with the smallest (primitive) unit cell and increasing the supercell size to a user-specified limit. For the O–Pt(321) case with five sites, we used ATAT to identify eight possible supercells with surface area up to three times that of the primitive cell, and it enumerated 45167 unique adsorbate configurations, or structures, on these eight supercells which we could then potentially calculate with DFT. The symmetry of the surface and the number of possible adsorption sites in the model both influence the size of supercells and number of configurations that can realistically be generated in this way.

Having enumerated all possible adsorbate configurations (structures) within a specified supercell size and calculated the DFT energies of a small, initial subset of these configurations, we then fit a CE to this initial set of energies (Figure 2.7, step 2). We use this CE to predict quickly the energy of the thousands of possible structures (Figure 2.7, step 3) and to identify the ones that are predicted to be low in energy or otherwise improve the database to enable a better fit of the CE (Figure 2.7, step 4). ATAT can recommend new structures based on structural features that are underrepresented in the database and could make the database more representative of the real system by their inclusion. Finally, we calculate the energies of the newly identified structures of interest and update the database (Figure 2.7, step 5). This cycle of calculating DFT energies, fitting a CE, predicting energies of possible new structures, and selecting new structures to calculate and add to the database is repeated until the CE converges to a self-consistent result.

2.3.2.2 Convergence Criteria and Self-consistency

Two primary criteria must be satisfied to ensure that a cluster expansion is reliable. First, the fitting database must contain enough structures to represent fully all of the important features of the real system. Second, significant clusters necessary for describing the system must be included in the CE. One can see that if either criterion is not satisfied the CE will not be reliable: if major low energy structures are missing from the database, then the CE fit to that incomplete database could potentially miss those same structures, and if major

site–site interactions are absent from the CE, then the CE will be unable accurately to predict the energy of arbitrary adsorbate configurations. The challenge is that both criteria are impossible to measure without significant prior knowledge of the system, which is generally not available in the early stages of research.

For this reason, we adopt a criterion of self-consistency as an alternative to the two real criteria for CE reliability. We look for self-consistency in two ways, mimicking the two above-mentioned criteria. First, we compare the DFT database energies to the predicted energies of the entire set of enumerated configurations and evaluate whether the predictions and the known energies are consistent. If new structures are predicted to be more stable than the structures already included in the database, then these structures should be added to the database and iterations continued. Second, as the DFT database increases in size and completeness, we compare subsequent CE fits (with optimal CV score) to ensure that they are using the same clusters with approximately the same interaction coefficients (ECIs). When this is not the case, it suggests that continued iterations are necessary.

Self-consistency is both evaluated and reached by following the iterative fitting process of Figure 2.7. In step 2, we fit a CE to the current DFT database using whichever clusters are required to obtain an optimal CV score. Although the ''leave-one-out'' CV score is not a perfect metric, it does measure predictive capability, which lends confidence that the second criterion is being met and all important clusters are included in the CE, at least for the given database. In steps 3 and 4, this optimal CE is then used to predict the energies of every enumerated structure, and the predicted energies are compared to the fitting database to identify any structures that are predicted to be thermodynamically important (low in energy) but not already in the fitting database. This tests the first criterion regarding the completeness of the database, and by calculating the DFT energies of these highlighted structures and adding them to the fitting database, we continually improve the database. Once self-consistency is reached—that is, once there are no more structures predicted to have low energies that do not already belong to the database and subsequent CE fits have similar functional forms containing the same clusters with similar ECI's—then we assume that the CE is reliable and fits as well as possible.

2.3.2.3 Ordered Ground States

Practical implementation of the above self-consistency criteria requires that the user make some decisions regarding specific rules to determine when convergence is reached. For the second criterion comparing subsequent CEs for consistency in the use of clusters and ECIs, one could ensure that some percentage (*e.g.* 90%) of the clusters are the same and that the ECIs for the clusters that are the same differ by no more than some percentage (*e.g.* 10%). The first criterion of database convergence is more difficult to quantify. The most straightforward measure of self-consistency is when all of the ground state structures determined from the CE energy predictions (called ''predicted

ground states'') agree with the ground states determined from the DFT database energies (called ''true ground states'').

In the context of adsorption, ground states are specific surface orderings of adsorbates that define a convex hull, such as the one in the right panel of Figure 2.5. The convex hull is constructed by drawing lines between all pairs of points on a scatter plot of formation energy *vs.* coverage, with the collection of lowest energy segments defining the hull and the vertices or nodes of the hull being the ordered ground states. There does not have to be a ground state at every coverage represented in the database and, in practice, the ground state structures are usually only a subset of the minimum energy structures at each coverage. Any arrangement of adsorbates can always decrease in energy by decomposing into the nearest ground states of higher and lower coverage.

Because the ground states are the thermodynamically most significant structures, a comparison of the CE ''predicted ground states'' and the DFT ''true ground states'' is usually sufficient to obtain the needed energetic information. In practice, this criterion can be made stricter by placing additional constraints on the comparison, such as requiring the CE predictions to reproduce the correct minimum energy structures at all coverages, not just the ground states, or even reproducing correct ordering among the two or three lowest energy structures, not just the lowest energy structure. Alternatively, the criterion could be made less strict by only requiring a subset of the ground states to agree.

2.3.2.4 Monitoring for Convergence

The iterative process outlined in Figure 2.7 requires some analysis to check for convergence and determine how to proceed. Two key aspects of this analysis are described in the sections below. This analysis is explicitly stated in step 4 of Figure 2.7, but different aspects can be carried out continuously throughout the fitting process in order to determine when the configuration space has been sufficiently explored and the cluster expansion has reached self-consistency.

2.3.2.4.1 Evaluating the Fitting Error. During the process of expanding the fitting database by calculating the energies of additional structures, the CV score of each subsequent CE should decrease slightly as an indication that, by fitting to a more complete database, the new CE will have better predictive accuracy than the previous one. However, the CV score alone can be misleading. In some cases, overfitting by including too many clusters can decrease the CV score to low values when a better CE may actually have a higher CV score. Alternatively, one or two structures that fit very poorly, or may even have some error in the DFT calculations, can significantly increase the CV score, so large CV score values do not necessarily reflect a poor fit, but may actually signify that some of the DFT results should be examined for possible errors. This can be quickly ascertained by plotting the formation energies from DFT together with the formation energies from the CE

predictions, or a plot of the CE prediction *vs.* the DFT energies, and looking for large discrepancies.

The CV score will decrease most significantly early on in the iterative fitting process. Once the database becomes very large, the CE will be able to use more clusters and fit the database with a small CV score, regardless of the incremental changes to the database, as seen in Figure 2.8. Therefore other measurements of error beyond the CV score are also helpful in gauging the quality of fit.

One class of error that might be examined is the error in CE energies of selected structures that are included in the fitting database compared to their DFT energies. This type of comparison will illustrate how well the CE is fitting to the database. Because the ground state structures are most significant thermodynamically, one such metric is the average error between the CE predictions and DFT energies for the ground state structures. For a larger sample to average, the minimum energy structure at each coverage could be used, regardless of whether or not it is a ground state. Averages can be determined from the root mean square (RMS), as is done for the CV score, or a mean absolute error.

The other class of possible error measurements that could be examined comprises the errors in the CE and DFT energies of structures that are not included in the fitting database. This error will give insight into the predictive ability of the CE and can be achieved in two ways. First, given that each cycle throughout the fitting process involves identifying new structures that need to be included in the database and then calculating their DFT energies, the CE prediction of the energy of these selected structures can be compared with the

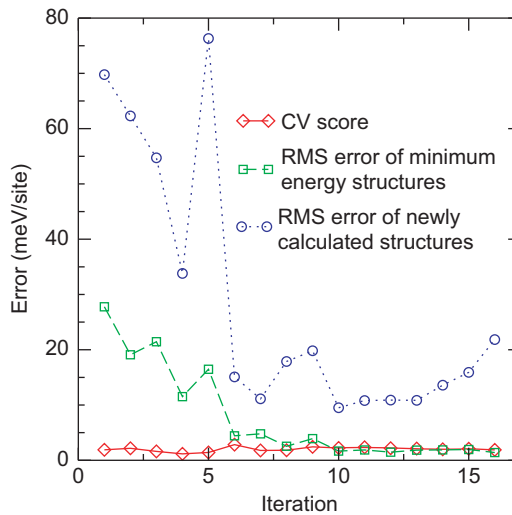


Figure 2.8 Plot showing the RMS error of minimum energy structures and of newly calculated structures. The CV score is also shown for comparison and to illustrate that it may converge much more quickly than the other metrics.

subsequently calculated DFT energies. A second approach would be to set up a testing database of DFT energies that will be excluded from the CE fitting. For each CE, the difference in the CE predictions and DFT energies for the structures in this testing database can be averaged for a prediction error. When using this second method, it is important that the structures in the testing database are selected completely randomly so that no bias is introduced to the fitting. The first approach is often used because it is faster and requires fewer additional calculations, but it introduces bias into the error measurement by only considering those structures that were chosen to be added to the database according to some predetermined criteria, which can make it less reliable than using a completely random testing database. Additionally, the structures used to calculate the prediction error by the first method change with each iteration, making the errors difficult to compare over the course of several iterations.

Regardless of the exact structures compared and the type of averaging done, each of the above-mentioned errors should give a similar view of incremental improvements with each subsequent iteration of the fitting process, though some may converge more quickly than others. By evaluating more than one type of error, a more comprehensive picture can be obtained regarding the quality of fit and ultimate convergence of each progressive cluster expansion. The plot in Figure 2.8 illustrates, for selected iterations of the fitting process of O–Pt(321), how RMS errors of minimum energy structures and of newly calculated structures follow similar overall trends but give slightly different information. It also shows that a low CV score is not sufficient in and of itself to determine whether a CE is converged.

2.3.2.4.2 Visually Comparing Formation Energies of Predicted and Database Structures. In Figure 2.9 we provide two plots of the formation energies for the O–Pt(321) DFT database with the CE predictions overlaid to provide a visual illustration of the concept that when self-consistency is reached, there

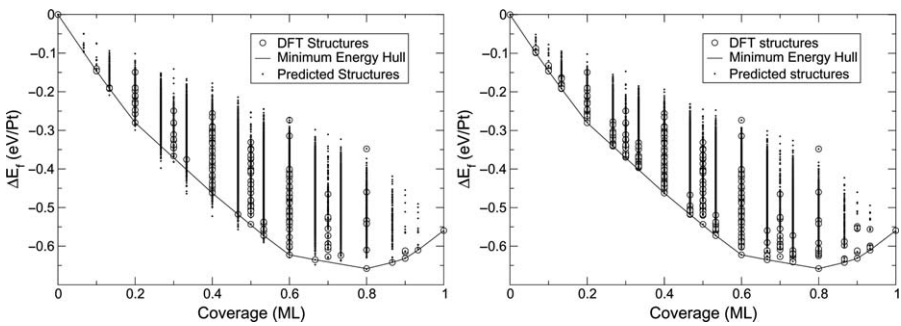


Figure 2.9 Formation energies of the DFT fitting database (open circles) and the predicted energies of uncalculated structures (black dots), illustrating two cases where (left) predictions indicate uncalculated structures with energies below the DFT ground states (corresponding to iteration 1 in Figure 2.8) and (right) predictions show no uncalculated structures with energies below the DFT ground states (corresponding to iteration 16 in Figure 2.8).

will be no new structures lower in energy than the ground states. The left figure shows an unconverged case with 175 structures in the database where several new structures are predicted to be lower in energy than the DFT ground states. The right figure shows the predicted structure energies after self-consistency is reached (354 structures in the database), where no new structures are predicted to have lower energies than those in the DFT database.

2.4 Applying the Model to Obtain Physical Insight

The converged database of DFT energies and the cluster expansion model provide a rich range of information about the surface–adsorbate system. Some useful conclusions can be drawn from the DFT database itself.

2.4.1 Analysis of the DFT Fitting Database

2.4.1.1 O-Pt(321) Ordered Ground States

The formation energies for the O–Pt(321) system are shown in the left plot of Figure 2.10. These energies are referenced to the clean Pt(321) surface and gas-phase molecular O_2, such that $E_{A(g)} = \frac{1}{2}E_{O_2(g)}$. Fourteen ordered ground states were identified for the O–Pt(321) system, shown in Figure 2.11 with letters corresponding to the formation energy plot in Figure 2.10. From examining these structures, we find that at coverages up to 0.2 ML (structures B–E), only the "b21" bridge sites are occupied. This is what we would expect based on the average adsorption energies of the isolated adsorbates listed in Table 2.1, which show that the "b21" site binds oxygen most strongly. Above 0.2 ML, we see the emergence of 4-fold coordinated kink Pt atoms, which gradually occupy more of the surface up to 0.8 ML (structure J), when every kink is saturated with four O atoms. Between 0.8 and 1 ML a fifth oxygen gradually fills in the last

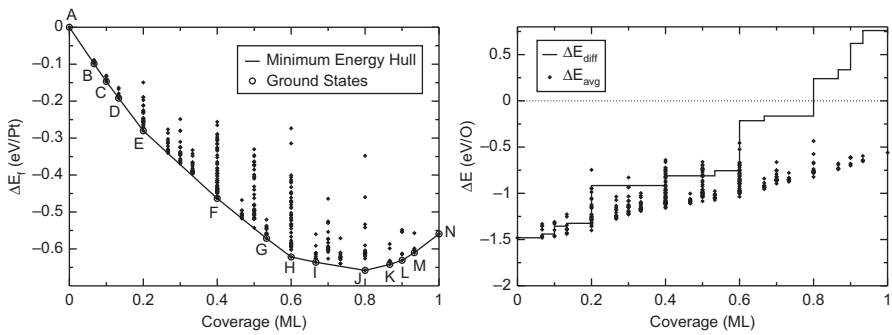


Figure 2.10 Plots showing adsorption energies of 372 configurations of oxygen on Pt(321). Left: formation energy, with letters corresponding to ground state structures shown in Figure 2.11. Right: differential and average adsorption energies.

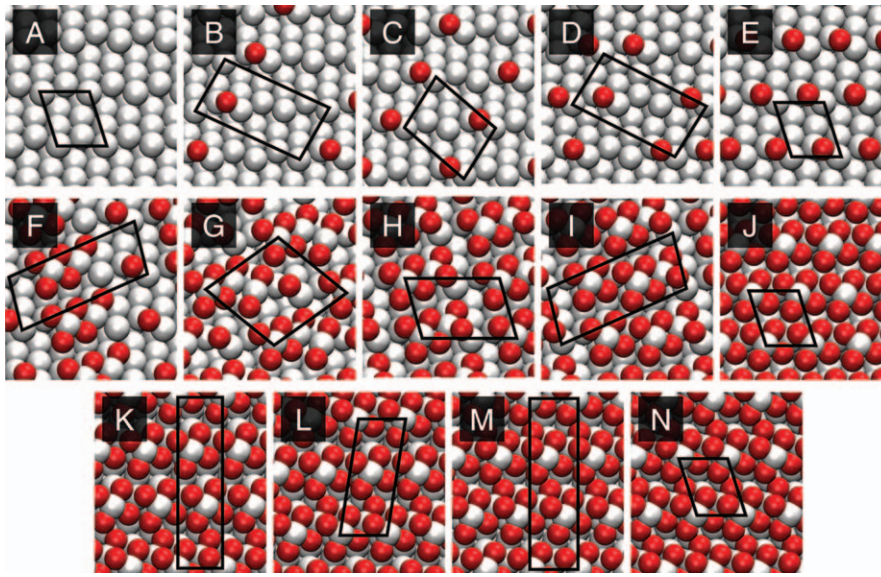


Figure 2.11 Ground states for the O/Pt(321) system (red = O, white = Pt) with integral adsorption energies circled in Figure 2.10 and coverages listed in Table 2.2. The supercell used to calculate each ground state is shown on each image.

adsorption site in each unit cell. Within our five-site model, 1 ML is the maximum coverage possible.

All of the ground states shown in Figures 2.10 and 2.11 are not equal in importance. Some ground states barely break the minimum energy hull and are therefore expected to dominate the surface under a very narrow range of conditions, while others generate large kinks in the hull and will be equilibrium configurations over a wider range of conditions. Hence, the relative importance of each ground state can be directly related to the change in slope of the minimum energy hull at that ground state.[12] As discussed in the next section, the slope of the formation energy hull defines the average differential binding energy for the system as it changes from one ground state to the next. Table 2.2 lists each ground state and its corresponding coverage and slope change.

2.4.1.2 Differential Binding Energy

An adsorbate differential binding energy is the incremental change in energy resulting from the addition of a single adsorbate to a surface with some configuration of adsorbates already present. This energy is calculated by subtracting energies computed between identical supercells that differ by only one adsorbate:

$$\Delta E_{\text{diff}}(\boldsymbol{\sigma}, i) = E_{\boldsymbol{\sigma}}^{\text{DFT}}(\phi) - (E_{\boldsymbol{\sigma}-i}^{\text{DFT}}(\phi) + E_{\text{A(g)}}) \tag{2.13}$$

Table 2.2 List of each ground state with its corresponding coverage and change in slope of the formation energy hull.

Ground state	*Coverage (ML)*	*Δ Slope (eV/O)*
A	0.000	—
B	0.067	0.04
C	0.100	0.09
D	0.133	0.03
E	0.200	0.41
F	0.400	0.11
G	0.533	0.06
H	0.600	0.54
I	0.667	0.05
J	0.800	0.41
K	0.867	0.10
L	0.900	0.29
M	0.933	0.14
N	1.000	—

Here $\boldsymbol{\sigma}$ denotes the final configuration of adsorbates, while $\boldsymbol{\sigma}$–i represents another configuration of adsorbates obtained by removing a single adsorbate from site i of configuration $\boldsymbol{\sigma}$. Thus there can be different differential binding energies for each occupied site i of a configuration $\boldsymbol{\sigma}$.

While differential binding energies can be calculated between any two configurations and are important in constructing statistical averages over surface properties, of direct practical relevance are average binding energy differences between ground state configurations, which as we show below relate to the adsorbate chemical potential. Single crystal adsorption calorimetry[62] and TPD both provide access to average differential binding energies. Section 2.3.2.3 described how the ground states and minimum energy hull are determined from the DFT energies. The average differential binding energy is obtained by differentiating the minimum energy hull of the formation energies with respect to coverage:

$$\Delta E_{\mathrm{diff}}(\theta) = \frac{\mathrm{d}(\Delta E_{\mathrm{f,hull}}(\theta))}{\mathrm{d}\theta} = \frac{\mathrm{d}(E_{\mathrm{hull}}^{\mathrm{DFT}}(\theta))}{\mathrm{d}N_{\mathrm{A}}} - E_{\mathrm{A(g)}}^{\mathrm{DFT}} \tag{2.14}$$

Here $\Delta E_{\mathrm{f,hull}}(\theta)$ is the formation energy and $E_{\mathrm{hull}}^{\mathrm{DFT}}(\theta)$ is the DFT energy of the particular ground state structure at coverage θ. We defined the minimum energy hull in Section 2.3.2.3 as a series of line segments connecting the ordered ground states of the system. As such, the derivative expressed in eqn (2.14) becomes a step function, as seen in the right plot of Figure 2.10. This plot also shows that adsorbates bind less strongly as coverage increases, which indicates repulsive interactions between adsorbates. Because there is only one point on the minimum energy hull at each coverage and by restricting ourselves to adsorbate configurations on this hull, the differential binding energy can now

be expressed directly as a function of coverage, θ, rather than of an arbitrary configuration, $\boldsymbol{\sigma}$.

2.4.1.3 Maximum Equilibrium Coverage at 0 K

In both of the definitions of differential binding energy given above negative values indicate exothermic and positive values endothermic, or unfavorable, adsorption. By this definition, it will continue to be energetically favorable to add adsorbates until the differential binding energy becomes positive. This fact allows us to identify the maximum, or saturation, coverage as the coverage where the differential binding energy crosses zero. Similarly, because the differential binding energy is related to the formation energy, according to eqn (2.14), the same saturation coverage can be identified in a plot of formation energy *vs.* coverage by the minimum point on the minimum energy hull. The saturation coverage defined in this way is dependent on the choice of reference for the adsorbate in the gas phase, $E_{A(g)}$. This saturation coverage is that accessible at equilibrium at 0 K. We will describe in the next section how the effect of temperature on equilibrium coverages can be obtained.

Choosing O_2 as the dosing gas and reference state, we can use Figure 2.10 to identify a saturation coverage for oxygen at the Pt(321) surface of 0.8 ML. As discussed further in Sections 2.4.2.2 and 2.4.2.3, oxygen can also be dosed to the surface by sources other than molecular O_2, such as NO_2 or O_3, and referencing the adsorption energies to these sources will identify the saturation coverage specific to each case.[7]

2.4.2 Analysis of Ordered Ground States

The analysis of formation energies to this point has primarily served to identify and briefly characterize the ground state adsorbate configurations and the maximum equilibrium coverage that can be reached for a given gas-phase reference state. In this section we discuss how to use the 0 K ordered ground states to predict equilibrium surface coverage under various environments and external conditions beyond the 0 K limit already discussed in Section 2.4.1.3. The definition of 0 K ground states is strictly enthalpic in nature, and conclusions relating to limiting equilibrium adsorbate coverages will be an approximation that is valid only at very low temperatures where entropy contributes little to the free energy. At higher temperatures, however, both configurational and vibrational entropy can be significant and must be accounted for in some way. Configurational entropy relates to the tendency for adsorbates to disorder, or rearrange, from the ground state structures at finite temperature. While configurational entropy contributes to both the thermodynamics and kinetics of adsorption,[30] and will be discussed in Section 2.4.3, the absolute value of the contribution is on the order of $RT \ln 2$, small relative to the total free energy, and ignored in first-pass analysis.[7] Even at moderate temperatures the short-range local ordering of the system characteristic of the

ground state structures is not completely lost and the insight gained from a study of only the ground states can still be valuable in understanding the real system.[26] We therefore discuss an analysis of only the ground state structures identified in Figures 2.10 and 2.11, incorporating the temperature dependence of the vibrational degrees of freedom using a standard harmonic approximation treatment from statistical mechanics and DFT-derived vibrational frequencies.

2.4.2.1 *Defining Surface Free Energy*

The surface free energy includes contributions from the 0 K DFT formation energies in Figure 2.10 as well as temperature-dependent corrections to incorporate entropy effects.[29] We apply standard thermodynamic relationships and equilibrium assumptions to define the chemical potential of the adsorbate on the surface, and compare surface free energies of adsorbate configurations to identify the most stable (*i.e.* having minimum surface free energy) under given temperature and chemical potential conditions. For our example of the O–Pt(321) system, we illustrate this technique by identifying the region of stability of each ground state in two oxidizing environments commonly used experimentally to dose oxygen to a surface: an O_2 and an NO/NO_2 environment.[7,13]

The ground state adsorption energies, zero-point energies, and vibrational modes are needed to determine the surface free energy of each ground state configuration. The surface free energy, γ, is defined as:[7]

$$\gamma_{\boldsymbol{\sigma}}(T, \mu_{A^*}) = \gamma_{\text{clean}}(T) + \Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_{A^*}) \tag{2.15}$$

Here $\boldsymbol{\sigma}$ indicates a specific adsorbate configuration, γ_{clean} is the surface energy of the clean surface approximated by eqn (2.1), μ_{A^*} is the chemical potential of the adsorbate A, and T is temperature.

We calculate this energy for each adsorbate configuration and identify the most stable arrangements as those with the lowest surface free energy. In general, whenever adsorbate configurations on different surface facets are compared, then the full surface energy of eqn (2.15) must be considered, particularly if the facets have significantly different surface energies, γ_{clean}. However, differences in surface energy are often small relative to differences in adsorption energy such that the full form of eqn (2.15) is unnecessary. In particular, when all adsorbate configurations being compared are for the same surface facet, then the γ_{clean} term in eqn (2.15) is constant and exactly cancels out, and only $\Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_{A^*})$ is needed to rank the relative surface energies of different configurations:[7]

$$\Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_{A^*}) = \frac{F_{\boldsymbol{\sigma}}(T) - F_{\text{clean}}(T) - N_A(\boldsymbol{\sigma})\mu_{A^*}}{A_{\text{cs}}} \tag{2.16}$$

Here F is the Helmholtz energy and A_{cs} is the cross-sectional area of the DFT supercell. The Helmholtz energy is separable into independent contributions:

$$F(T) = E^{\text{DFT}} + E^{\text{ZP}} + \Delta F(T) \tag{2.17}$$

The DFT energy, E^{DFT}, and zero-point vibrational energy, E^{ZP}, combine to form the 0 K energy, while the temperature dependence is all incorporated into $\Delta F(T)$, which can be expressed in terms of the vibrational partition function, Q^{vib}:

$$\Delta F(T) = F(T) - F(0\,\mathrm{K}) = -k_{\mathrm{B}}T \ln Q^{\mathrm{vib}} = F^{\mathrm{vib}} \tag{2.18}$$

This expression is an approximation valid for surfaces in their electronic ground state where there are no translational or rotational degrees of freedom. Other contributions to the free energy, such as configurational entropy, are assumed to be small in comparison, as previously stated.[7] In Section 2.2.2.4 we discussed that vibrational modes can be calculated within the harmonic approximation. From the normal mode frequencies, ν_i, we can calculate the total vibrational partition function as a product of individual mode partition functions, q_i^{vib}:

$$Q^{\mathrm{vib}} = \prod_i^{\mathrm{modes}} q_i^{\mathrm{vib}} = \prod_i^{\mathrm{modes}} \frac{1}{1 - \exp\left(-\dfrac{h\nu_i}{k_{\mathrm{B}}T}\right)} \tag{2.19}$$

Applying the definitions in eqns (2.17)–(2.19), we can rewrite eqn (2.16):

$$\Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_{\mathrm{A}^*}) = \frac{(E_{\boldsymbol{\sigma}}^{\mathrm{DFT}} - E_{\mathrm{clean}}^{\mathrm{DFT}}) + (E_{\boldsymbol{\sigma}}^{\mathrm{ZP}} - E_{\mathrm{clean}}^{\mathrm{ZP}}) + (F_{\boldsymbol{\sigma}}^{\mathrm{vib}} - F_{\mathrm{clean}}^{\mathrm{vib}}) - \mu_{\mathrm{A}^*} N_{\mathrm{A}}(\boldsymbol{\sigma})}{A_{\mathrm{cs}}} \tag{2.20}$$

To simplify notation, we will use $\Delta_{\boldsymbol{\sigma}}$ to indicate the difference between a particular adsorbate configuration and the clean surface, $X_{\boldsymbol{\sigma}} - X_{\mathrm{clean}}$, where X is any energetic quantity; eqn (2.20) then becomes:

$$\Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_{\mathrm{A}^*}) = \frac{1}{A_{\mathrm{cs}}}\left(\Delta_{\boldsymbol{\sigma}}E^{\mathrm{DFT}} + \Delta_{\boldsymbol{\sigma}}E^{\mathrm{ZP}} + \Delta_{\boldsymbol{\sigma}}F^{\mathrm{vib}} - \mu_{\mathrm{A}^*}N_{\mathrm{A}}(\boldsymbol{\sigma})\right) \tag{2.21}$$

2.4.2.1.1 Zero Point and Vibrational Free Energy. The harmonic oscillator approximation is commonly used to calculate the vibrational contributions to the zero-point energy and temperature dependence of the Helmholtz energy. Examining the $\Delta_{\boldsymbol{\sigma}}F^{\mathrm{vib}}$ term in detail and applying eqns (2.18) and (2.19), we have

$$\Delta_{\boldsymbol{\sigma}}F^{\mathrm{vib}} = F_{\boldsymbol{\sigma}}^{\mathrm{vib}} - F_{\mathrm{clean}}^{\mathrm{vib}} = -k_{\mathrm{B}}T \ln \frac{Q_{\boldsymbol{\sigma}}^{\mathrm{vib}}}{Q_{\mathrm{clean}}^{\mathrm{vib}}} \tag{2.22}$$

Because total vibrational partition functions, Q^{vib}, are products of partition functions of individual modes, q_i^{vib}, and because phonon modes associated with motions of the metal surface are typically well separated energetically from

adsorbate modes, we can factor the product into parts for the surface and the adsorbates:

$$\Delta_{\boldsymbol{\sigma}} F^{\mathrm{vib}} = -k_{\mathrm{B}} T \ln \frac{\prod_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}} N_{\mathrm{A}}} q_i^{\mathrm{vib}} \prod_{i}^{\mathrm{surf}} q_i^{\mathrm{vib}}}{\prod_{i}^{\mathrm{surf}} q_i^{\mathrm{vib}}} \tag{2.23}$$

Here $n_{\mathrm{dof}}^{\mathrm{vib}}$ is the number of vibrational degrees of freedom of a single adsorbate, which has a value of three for atomic adsorbates, but includes additional internal vibrations for adsorbed molecules. If the surface vibrational modes are insensitive to the presence of adsorbates, the surface terms cancel and the expression simplifies to:

$$\Delta_{\boldsymbol{\sigma}} F^{\mathrm{vib}} = -k_{\mathrm{B}} T \ln \prod_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}} N_{\mathrm{A}}} q_i^{\mathrm{vib}} \tag{2.24}$$

Test calculations on the O–Pt(111) system[7] confirm that this approximation influences $\Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_{\mathrm{A}^*})$ by no more than 0.5 meV Å^{-2}.

A similar argument can be made to simplify the zero-point energy terms. If the zero-point energy, defined previously in eqn (2.2), can be separated into parts for the adsorbate and the surface:

$$\Delta_{\boldsymbol{\sigma}} E^{\mathrm{ZP}} = E_{\boldsymbol{\sigma}}^{\mathrm{ZP}} - E_{\mathrm{clean}}^{\mathrm{ZP}} = \left(\sum_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}} N_{\mathrm{A}}} \frac{1}{2} h\nu_i + \sum_{i}^{\mathrm{surf}} \frac{1}{2} h\nu_i \right) - \sum_{i}^{\mathrm{surf}} \frac{1}{2} h\nu_i \tag{2.25}$$

and if the surface phonon modes couple weakly with the adsorbate modes, then the surface terms will cancel, leaving only the sum over modes due to the adsorbates:

$$\Delta_{\boldsymbol{\sigma}} E^{\mathrm{ZP}} = \sum_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}} N_{\mathrm{A}}} \frac{1}{2} h\nu_i \tag{2.26}$$

Each adsorption site has unique, potentially configuration-dependent vibrational modes that in principle should be calculated explicitly. To minimize the computational expense, we can use calculated vibrational modes, zero-point energies, and vibrational partition functions for adsorbates in a handful of key configurations to select an average, or representative, value for the zero-point energy ($E_{\mathrm{A}^*}^{\mathrm{ZP}}$) and vibrational free energy($F_{\mathrm{A}^*}^{\mathrm{vib}}$) of a single adsorbate. These values can then simply be scaled by the number of adsorbates. Doing this, eqns (2.24) and (2.26) become eqns (2.27) and (2.28), respectively.

$$\Delta_{\boldsymbol{\sigma}} F^{\mathrm{vib}} = -k_{\mathrm{B}} T N_{\mathrm{A}} \ln \prod_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}}} q_i^{\mathrm{vib}} = -k_{\mathrm{B}} T N_{\mathrm{A}} \ln q_{\mathrm{A}^*}^{\mathrm{vib}} = N_{\mathrm{A}} F_{\mathrm{A}^*}^{\mathrm{vib}} \tag{2.27}$$

$$\Delta_{\boldsymbol{\sigma}} E^{\mathrm{ZP}} = N_{\mathrm{A}} \sum_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}}} \frac{1}{2} h\nu_i = N_{\mathrm{A}} E_{\mathrm{A}^*}^{\mathrm{ZP}} \tag{2.28}$$

Table 2.1 lists the zero-point energies for an isolated O atom in each Pt(321) adsorption site calculated according to eqn (2.26). For the five sites used in our model, in particular, the zero-point energies are all very similar. Test calculations for the O–Pt(321) system at higher coverages indicate that using a representative single-adsorbate zero-point energy and vibrational partition function influence the surface free energy by less than 1 meV Å^{-2} over the coverage range considered.

2.4.2.1.2 Chemical Potential References. The surface free energy depends on both the temperature and adsorbate chemical potential, yet it is not straightforward to apply eqn (2.21) in its current form using the information available from DFT calculations. It is convenient to redefine the absolute chemical potential, μ_{A*}, as a change in chemical potential relative to some reference state, μ_{ref}. The relative stability of each ground state structure is then a function of $\Delta\mu_A$:

$$\Delta\mu_A = \mu_{A^*} - \mu_{ref} \tag{2.29}$$

In surface adsorption models, an ideal gas of A at 0 K is a common, although not unique, reference.

$$\mu_{ref} = E_{A(g)}(0\,K) = E^{DFT}_{A(g)} + E^{ZP}_{A(g)} \tag{2.30}$$

With this choice of reference, we can further rearrange and simplify the expression for $\Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_{A^*})$ in terms of the formation energy. Because the surface free energy is normalized to the surface area, it is useful to redefine an area-normalized formation energy, $\Delta\tilde{E}_f(\boldsymbol{\sigma})$:

$$\Delta\tilde{E}_f(\boldsymbol{\sigma}) = \frac{1}{A_{cs}}\left[E^{DFT}_{\boldsymbol{\sigma}}(\phi) - E^{DFT}_{clean}(\phi) - N_A(\boldsymbol{\sigma})E^{DFT}_{A(g)}\right] \tag{2.31}$$

Substituting the definitions and simplifications in eqns (2.27), (2.28), (2.30), and (2.31) into eqn (2.21) gives:

$$\Delta\gamma_{\boldsymbol{\sigma}}(T, \Delta\mu_A) = \Delta\tilde{E}_f(\boldsymbol{\sigma}) + \frac{N_A(\boldsymbol{\sigma})}{A_{cs}}\left(E^{ZP}_{A^*} - E^{ZP}_{A(g)} + F^{vib}_{A^*} - \Delta\mu_A\right) \tag{2.32}$$

While this form is significantly simpler than the general case in eqn (2.21), it is important to recognize the assumptions and reference states involved in making this simplification and verify their validity before using eqn (2.32).

Examination of eqn (2.32) reveals that $\Delta\gamma_{\boldsymbol{\sigma}}(T, \Delta\mu_A)$ is a linear function of $\Delta\mu_A$ with slope $-\frac{N_A(\boldsymbol{\sigma})}{A_{cs}}$. As such, the surface free energy for each ground state at a given temperature can be plotted as a straight line *vs.* $\Delta\mu_A$.[29] The plot in Figure 2.12 presents $\Delta\gamma$ at 600 K for all of the O–Pt(321) ground states. This plot quickly allows us to identify which ground state is most stable at a given chemical potential by finding the lowest energy line at that chemical potential. For the choice of reference state in eqn (2.30), the zero on the *x*-axis represents the surface in equilibrium with a 0 K reservoir of the gas-phase adsorbate.

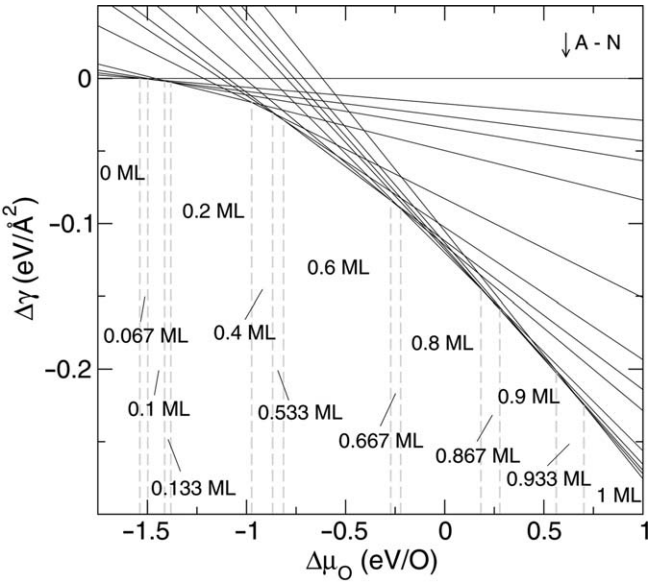


Figure 2.12 Surface free energy *vs.* $\Delta\mu_O$ for the 14 ordered ground states of the O–Pt(321) system at 600 K. Letters correspond to the ground state structures identified in Figures 2.10 and 2.11 and Table 2.2. $\Delta\mu_O$ is defined according to eqns (2.29) and (2.30).

Wherever the minimum energy lines cross indicates a transition from one ordered surface structure to another. By drawing vertical lines down to the *x*-axis from these intersections, we can illustrate the chemical potential ranges where each respective ground state will be observed. For example, the 0.6 ML ground state has the largest range of stability of all of the ground states, followed by the 0.2 ML and 0.8 ML ground states. This is the same trend predicted by our analysis of the change in slope of the formation energy hull for each ground state summarized in Table 2.2.

2.4.2.2 Relating Chemical Potential to a Gas-Phase Reservoir

Because a chemical potential is not easily measured directly, it is desirable to relate it to quantities more directly accessible in an adsorption experiment, such as temperature and pressure. This can be accomplished by assuming that the adsorption system is in equilibrium with a gas-phase reservoir:

$$\mathrm{A(g)}(T,P) + {}^{*} \rightleftharpoons \mathrm{A}^{*} \quad \mu_{\mathrm{A(g)}} = \mu_{\mathrm{A}^{*}} \tag{2.33}$$

This assumption allows the chemical potential of adsorbates to be related to the chemical potential of a gas-phase species. The gas-phase reference chemical

potential is related to T and P through the integrated Gibbs–Duhem equation:[7,63]

$$\mu_{A(g)}(T, P_A) = E^{DFT}_{A(g)} + E^{ZP}_{A(g)} + \Delta G^{\circ}_{A(g)}(T) + k_B T \ln\left(\frac{P_A}{P^{\circ}}\right) \tag{2.34}$$

Here A(g) represents any gas-phase species, E^{DFT} is the DFT electronic energy, E^{ZP} is the zero-point correction due to vibrations, P° is a reference pressure often chosen as 1 bar, and $\Delta G^{\circ}(T) = G(T, P^{\circ}) - G(0\,K, P^{\circ})$ and is calculated either from ideal gas statistical mechanics or from experimental data, often tabulated using the Shomate equation. This form of the Gibbs–Duhem equation is derived assuming ideal gas behavior, which is generally acceptable for the low partial pressures used in catalytic applications, and even for cases where the ideal gas assumption may not be valid it still serves as a first approximation. Applying the 0 K reference state of eqn (2.30), eqn (2.34) can be rewritten as:

$$\Delta\mu_{A(g)}(T, P_A) = \Delta G^{\circ}_{A(g)}(T) + k_B T \ln\left(\frac{P_A}{P^{\circ}}\right) \tag{2.35}$$

The choice of gas-phase reservoir is not unique and may be defined in a number of ways, and this choice affects the relationship between chemical potential and gas pressures. The conceptually simplest case to consider is an adsorbate dosed by its native gas-phase species, A(g), but depending on the particular experimental situation of interest or on the reliability of the available DFT data, it is also appropriate to define a reference to something other than A itself.[7,64] In the O–Pt(321) system, surface O might be dosed by dissociative adsorption of O_2:

$$\frac{1}{2}O_2(g) + {}^{*} \rightleftharpoons O^{*} \tag{2.36}$$

or by NO_2:

$$NO_2(g) + {}^{*} \rightleftharpoons O^{*} + NO(g) \tag{2.37}$$

We discuss these two cases in more detail below, but numerous other scenarios could be imagined, involving dosing by CO_2, H_2O, atomic O, O_3, or even solid oxides.[7] Reference states for adsorbates are not restricted to reservoirs of gas-phase species, but this interpretation is conceptually simple to understand when discussing adsorption at a gas–solid interface, and we refer to the adsorbate reservoir as a ''gas-phase reservoir'' simply for convenience.

2.4.2.2.1 Dissociative O_2 Adsorption. When oxygen is dosed to the surface by dissociation of $O_2(g)$, we define the $O_2(g)$ chemical potential reference state as:

$$\mu_{ref,O_2(g)} = E_{O_2(g)}(0\,K) = E^{DFT}_{O_2(g)} + E^{ZP}_{O_2(g)} \tag{2.38}$$

At equilibrium, the chemical potential of the adsorbed oxygen, O*, is:

$$\mu_{O^*}(T) = \frac{1}{2}\mu_{O_2(g)}(T, P_{O_2}) \tag{2.39}$$

For convenience, we choose the chemical potential reference for adsorbed O, $\mu_{ref,O}$*, to be consistent with our previous choice for O_2(g):

$$\mu_{ref,O^*} = \frac{1}{2}E_{O_2(g)}(0\,K) = \frac{1}{2}\mu_{ref,O_2(g)} \tag{2.40}$$

With this choice of references and the equality given in eqn (2.39) we can use eqn (2.29) to write $\Delta\mu_O$ as:

$$\Delta\mu_O = \mu_{O^*}(T) - \mu_{ref,O^*} = \frac{1}{2}\mu_{O_2(g)}(T, P_{O_2}) - \frac{1}{2}\mu_{ref,O_2(g)} = \frac{1}{2}\Delta\mu_{O_2}(T, P_{O_2}) \tag{2.41}$$

From eqn (2.35), $\Delta\mu_{O_2(g)}$ is:

$$\Delta\mu_{O_2(g)}(T, P_{O_2}) = \Delta G^{\circ}_{O_2(g)}(T) + k_B T \ln\left(\frac{P_{O_2}}{P^{\circ}}\right) \tag{2.42}$$

where the ΔG° term is obtained from tabulated experimental results or calculated. These expressions for $\Delta\mu_O$ and $\Delta\mu_{O_2(g)}$ can be applied to eqn (2.32) to obtain an expression for surface energy as a function of T and P_{O_2}:

$$\begin{aligned}\Delta\gamma_{\boldsymbol{\sigma}}(T, P_{O_2}) = \Delta\tilde{E}_f(\boldsymbol{\sigma}) + \frac{N_O(\boldsymbol{\sigma})}{A_{cs}}\Bigg[E^{ZP}_{O^*} - \frac{1}{2}E^{ZP}_{O_2(g)} + F^{vib}_{O^*} \\ - \frac{1}{2}\left(\Delta G^{\circ}_{O_2(g)}(T) + k_B T \ln\left(\frac{P_{O_2}}{P^{\circ}}\right)\right)\Bigg]\end{aligned} \tag{2.43}$$

2.4.2.2.2 Oxygen Adsorption from NO_2. In this second example, we imagine a surface dosed with oxygen through the equilibrated reaction of NO_2(g) and NO(g), eqn (2.37). The equilibrium assumption allows us to write the oxygen chemical potential, μ_{O^*}, as:

$$\mu_{O^*}(T) = \mu_{NO_2(g)}(T, P_{NO_2}) - \mu_{NO(g)}(T, P_{NO}) \tag{2.44}$$

Each chemical potential can be written as a chemical potential change plus a reference that can, in general, be different for each species:[7]

$$\mu_{NO_2(g)}(T, P_{NO_2}) = \Delta\mu_{NO_2(g)}(T, P_{NO_2}) + \mu_{ref,NO_2(g)} \tag{2.45}$$

$$\mu_{NO(g)}(T, P_{NO}) = \Delta\mu_{NO(g)}(T, P_{NO}) + \mu_{ref,NO(g)} \tag{2.46}$$

$$\mu_{O^*}(T) = \Delta\mu_O(T) + \mu_{ref,O^*} \tag{2.47}$$

Substituting these expressions into the equilibrium relationship above, we obtain:

$$\Delta\mu_{\mathrm{O}}(T) + \mu_{\mathrm{ref,O^*}} = \left(\Delta\mu_{\mathrm{NO_2(g)}}(T, P_{\mathrm{NO_2}}) + \mu_{\mathrm{ref,NO_2(g)}}\right) - \left(\Delta\mu_{\mathrm{NO(g)}}(T, P_{\mathrm{NO}}) + \mu_{\mathrm{ref,NO(g)}}\right) \tag{2.48}$$

Grouping the reference states together and solving for $\Delta\mu_{\mathrm{O}}(T)$ gives:

$$\Delta\mu_{\mathrm{O}}(T) = \Delta\mu_{\mathrm{NO_2(g)}}(T, P_{\mathrm{NO_2}}) - \Delta\mu_{\mathrm{NO(g)}}(T, P_{\mathrm{NO}}) + \left(\mu_{\mathrm{ref,NO_2(g)}} - \mu_{\mathrm{ref,NO(g)}} - \mu_{\mathrm{ref,O^*}}\right) \tag{2.49}$$

If we now choose the reference states for NO_2(g) and NO(g) in keeping with eqn (2.30) and use the same reference for O* defined above in eqn (2.40), then this expression becomes:

$$\Delta\mu_{\mathrm{O}}(T) = \Delta\mu_{\mathrm{NO_2(g)}}(T, P_{\mathrm{NO_2}}) - \Delta\mu_{\mathrm{NO(g)}}(T, P_{\mathrm{NO}}) + \left(E_{\mathrm{NO_2(g)}}(0\,\mathrm{K}) - E_{\mathrm{NO(g)}}(0\,\mathrm{K}) - \frac{1}{2}E_{\mathrm{O_2(g)}}(0\,\mathrm{K})\right) \tag{2.50}$$

The group of reference states above can be written together as the 0 K reaction enthalpy, $\Delta H_{\mathrm{rxn}}(0\ \mathrm{K})$, of the gas-phase reaction:

$$\mathrm{NO(g)} + \frac{1}{2}\mathrm{O_2(g)} \rightleftharpoons \mathrm{NO_2(g)} \tag{2.51}$$

In general, $\Delta H_{\mathrm{rxn}}(0\ \mathrm{K})$ can be calculated directly from DFT or higher level quantum methods, but experimental results can be used where reliable computed values are not available. We can now substitute eqn (2.35) for each of the $\Delta\mu$ terms and replace the chemical potential references with $\Delta H_{\mathrm{rxn}}(0\ \mathrm{K})$ to obtain a complete expression for the oxygen chemical potential:

$$\Delta\mu_{\mathrm{O}}(T, P_{\mathrm{NO_2}}, P_{\mathrm{NO}}) = \left[\Delta G^{\circ}_{\mathrm{NO_2(g)}}(T) + k_{\mathrm{B}}T\ln\left(\frac{P_{\mathrm{NO_2}}}{P^{\circ}_{\mathrm{NO_2(g)}}}\right)\right] - \left[\Delta G^{\circ}_{\mathrm{NO(g)}}(T) + k_{\mathrm{B}}T\ln\left(\frac{P_{\mathrm{NO}}}{P^{\circ}_{\mathrm{NO(g)}}}\right)\right] + \Delta H_{\mathrm{rxn}}(0\,\mathrm{K}) \tag{2.52}$$

As long as the reference pressures, P°, for NO_2(g) and NO(g) are the same, the expression simplifies to:

$$\Delta\mu_{\mathrm{O}}(T, P_{\mathrm{NO_2}}, P_{\mathrm{NO}}) = \Delta G^{\circ}_{\mathrm{NO_2(g)}}(T) - \Delta G^{\circ}_{\mathrm{NO(g)}}(T) + k_{\mathrm{B}}T\ln\left(\frac{P_{\mathrm{NO_2}}}{P_{\mathrm{NO}}}\right) + \Delta H_{\mathrm{rxn}}(0\,\mathrm{K}) \tag{2.53}$$

With the oxygen chemical potential rewritten in terms of the gas-phase reservoir species, we then substitute back into eqn (2.32) for an expression of the surface energy as a function of T, P_{NO_2} and P_{NO}:

$$\Delta\gamma_{\boldsymbol{\sigma}}(T, P_{NO_2}, P_{NO}) = \Delta\tilde{E}_f(\boldsymbol{\sigma}) + \frac{N_O(\boldsymbol{\sigma})}{A_{cs}}\left[E^{ZP}_{O^*} - \frac{1}{2}E^{ZP}_{O_2(g)} + F^{vib}_{O^*} - \left(\Delta G^{\circ}_{NO_2(g)}(T) - \Delta G^{\circ}_{NO(g)}(T) + k_B T \ln\left(\frac{P_{NO_2}}{P_{NO}}\right) + \Delta H_{rxn}(0\,K)\right)\right] \quad (2.54)$$

2.4.2.3 Ground State Phase Diagrams

Having related the adsorbate chemical potential to specific temperature and pressure conditions, we can apply this transformation to the x-axis of the plots in Figure 2.12 to understand better the external conditions that lead to the stability of different coverage regimes. Alternatively, instead of viewing all of the surface energies at one specific temperature condition, we can determine the most stable adsorbate structures at each temperature and pressure of interest and prepare a surface stability phase diagram including both dimensions.

A phase boundary corresponds to equating $\Delta\gamma_{\boldsymbol{\sigma}}(T, \mu_A)$ for a pair of adjacent ground states. The resulting phase diagrams for the two example oxygen reservoirs discussed are shown in Figure 2.13. Each ground state structure is labeled by its coverage, reflecting the assumption that the ground state adsorbate configuration is dominant at the corresponding coverage. This assumption clearly becomes less accurate with increasing temperature, where ultimately the surface will disorder, and this is discussed further in Section 2.4.3.3. In spite of the numerous assumptions involved in preparing these phase diagrams, they do

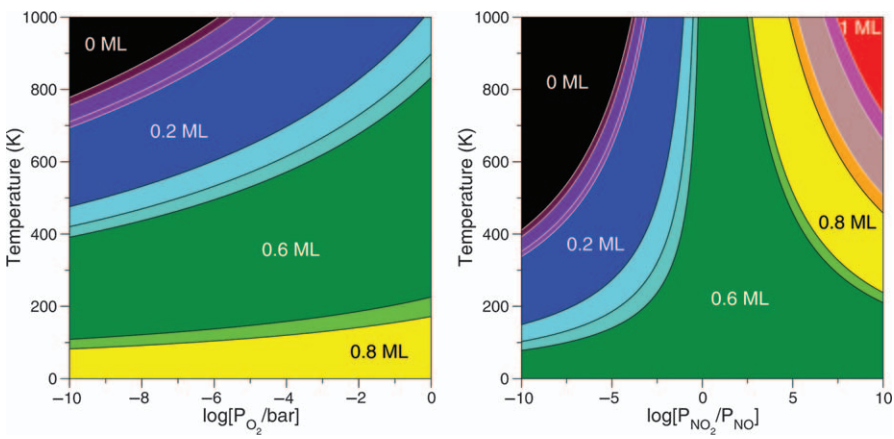


Figure 2.13 Phase diagrams showing the surface oxygen coverage under a variety of temperature and pressure conditions for (left) O_2 and (right) NO/NO_2 environments. The coverage shown refers specifically to the coverage of the ground state configuration from Figure 2.11 that minimizes the surface energy defined by eqn (2.20).

appear to be in agreement with the available experimental results for the O–Pt(321) system where the observed saturation O coverage at 285 K under ultra-high vacuum (UHV) conditions with O_2 dosing is approximately 0.6 ML.[53]

The phase diagram for an O_2 environment on the left of Figure 2.13 exhibits a shape characteristic of systems where the reactants and products have different amounts of gas-phase molecules (see eqn (2.36)). As temperature increases, the additional degrees of freedom of gas-phase O_2 relative to adsorbed O create a strong entropic driving force for oxygen to desorb, resulting in the relatively horizontal orientation of the phases. We contrast this with the phase diagram for an NO/NO_2 environment on the right of Figure 2.13, where the number of gas-phase molecules is a constant (see eqn (2.37)). As temperature increases, there is less entropic driving force for oxygen to desorb, so that the equilibrium coverage is primarily determined by the NO to NO_2 concentration ratio and phase boundaries are more vertical.[7]

2.4.3 Monte Carlo Simulations

The real power of the CE is in modeling adsorption behavior at finite temperature at length scales much greater than accessible directly with DFT. Monte Carlo simulation combined with CE provides access to a variety of equilibrium and even dynamic quantities, as listed at the bottom of Figure 2.2.

2.4.3.1 Details of Grand Canonical Monte Carlo (GCMC) Simulations

Adsorption Monte Carlo simulations are performed in the grand canonical ensemble, with independent variables volume (surface size), temperature, and chemical potential. The natural thermodynamic potential for the grand canonical ensemble is the Legendre transform of the Helmholtz free energy, $F(N, V, T)$, with respect to N. For a single component system, *i.e.* assuming a non-reconstructing surface with a single adsorbate, this transform yields the grand potential ϕ, which is also related to the partition function of the system summing over all configurational, vibrational, and electronic states, i:[61]

$$\begin{aligned}\phi(\mu, V, T) &= F(N, V, T) - \mu N = U - TS - \mu N = -PV \\ &= -k_B T \ln\left[\sum_i^{\text{states}} \exp\left(-\frac{E_i - \mu N_i}{k_B T}\right)\right]\end{aligned} \tag{2.55}$$

2.4.3.1.1 GCMC Implementation. We assume here the reader already has a basic familiarity with statistical mechanics and Monte Carlo simulations, and we therefore do not describe the algorithms in detail. The ATAT package includes a Monte Carlo code designed to interface directly with the CE that automates much of the thermodynamic analysis, including preparing temperature–composition phase diagrams. Van de Walle *et al.* have provided a detailed description of the Monte Carlo algorithms implemented in

ATAT.[61] The GCMC results presented here were performed with code developed in our group, briefly discussed in Section 2.4.3.2.

The acceptance criterion for moves in GCMC simulations is defined on the basis of the grand canonical partition function [eqn (2.55)] and includes the change in energy of the system as well as the change in adsorbate concentration. The acceptance probability P_{accept} is given by:

$$P_{\text{accept}} = \min\left[1, \exp\left(-\frac{\Delta E_{\text{tot}} - \mu_{\text{GCMC}}\Delta N_{\text{A}}}{k_{\text{B}}T}\right)\right] \tag{2.56}$$

where E_{tot} is the total energy of the surface as determined by the chosen Hamiltonian and μ_{GCMC} is the chemical potential input provided by the user. We use this notation to emphasize that we have some flexibility in defining E_{tot}, and this definition will influence the physical interpretation of μ_{GCMC}, as discussed further in the next section. The GCMC algorithm is implemented to sample over different numbers of adsorbates, N_{A}, and different adsorbate configurations, $\boldsymbol{\sigma}$, thus E_{tot} must have a functional form that can account for these variables.

If implemented with the CE as the Hamiltonian, then the change in total energy, ΔE_{tot}, is calculated based on the occupancy of and interactions between lattice sites. The CE always gives energies on a per-area basis, whether that area is defined explicitly by the number of surface metal atoms or cross-sectional surface area or implicitly by the size of the supercell used for calculating the DFT database energies. In order to obtain the total energy of the surface, E_{tot}, the CE energy, E^{CE}, must be scaled by the surface size [in this case, the number of surface metal atoms, N_{M}, as in eqn (2.9)]:

$$E_{\text{tot}} = E^{\text{CE}} \cdot N_{\text{M}} \tag{2.57}$$

The number of adsorbates, N_{A}, is calculated based on the occupancy variables of each lattice site on the surface:

$$N_{\text{A}} = \frac{1}{2}\sum_{i}^{\text{sites}} (\boldsymbol{\sigma}_i + 1) \tag{2.58}$$

Typically, GCMC moves change the occupancy of a single lattice site, so ΔN_{A} equals ± 1. Similarly, the total energy change, ΔE_{tot}, can be obtained more efficiently by recognizing that when changing the occupancy of a single site on a large surface, the energy of most of the surface is unchanged and cancels out and the energy change depends only on those site–site interactions involving the selected site. We can show this explicitly by rewriting eqn (2.11) as a sum of individual site contributions:

$$E_{\sigma}^{\text{CE}} = \sum_{i}^{\text{sites}} E_i^{\text{CE}} \tag{2.59}$$

where

$$E_i^{\text{CE}} = \frac{J_0}{N_{\text{sites}}} + J_i\sigma_i + \frac{\sigma_i}{2}\sum_{j}^{\text{sites}} J_{ij}\sigma_j + \frac{\sigma_i}{3}\sum_{j>k}^{\text{pairs}} J_{ijk}\sigma_j\sigma_k + \cdots \tag{2.60}$$

The total energy change from initial state $\boldsymbol{\sigma}_\mathrm{i}$ to final state $\boldsymbol{\sigma}_\mathrm{f}$ is:

$$\begin{aligned}\Delta E_{\mathrm{tot}}(\boldsymbol{\sigma}) &= N_\mathrm{M}\Delta E^{\mathrm{CE}}_{\boldsymbol{\sigma}} = N_\mathrm{M}\left(E^{\mathrm{CE}}_{\boldsymbol{\sigma}_\mathrm{f}} - E^{\mathrm{CE}}_{\boldsymbol{\sigma}_\mathrm{i}}\right)\\ &= N_\mathrm{M}\left[\sum_i^{\text{sites}}\left(\frac{J_0}{N_{\text{sites}}} + J_i\sigma_i + \frac{\sigma_i}{2}\sum_j^{\text{sites}} J_{ij}\sigma_j + \frac{\sigma_i}{3}\sum_{j>k}^{\text{pairs}} J_{ijk}\sigma_j\sigma_k + \cdots\right)_{\sigma_\mathrm{f}}\right.\\ &\quad\left. - \sum_i^{\text{sites}}\left(\frac{J_0}{N_{\text{sites}}} + J_i\sigma_i + \frac{\sigma_i}{2}\sum_j^{\text{sites}} J_{ij}\sigma_j + \frac{\sigma_i}{3}\sum_{j>k}^{\text{pairs}} J_{ijk}\sigma_j\sigma_k + \cdots\right)_{\sigma_\mathrm{i}}\right]\end{aligned} \tag{2.61}$$

Restricting the change in configurations from $\boldsymbol{\sigma}_\mathrm{i}$ to $\boldsymbol{\sigma}_\mathrm{f}$ to a single site s changing from $\sigma_{s,\mathrm{i}}$ to $\sigma_{s,\mathrm{f}}$, all terms in the initial and final sums cancel out except for those which contain the selected site, s:

$$\begin{aligned}\Delta E_{\mathrm{tot}}(\sigma_s) = N_\mathrm{M}\Bigg[&\Bigg(J_s\,\sigma_{s,\mathrm{f}} + \frac{\sigma_{s,\mathrm{f}}}{2}\sum_j^{\text{sites}} J_{sj}\,\sigma_j + \frac{\sigma_{s,\mathrm{f}}}{2}\sum_i^{\text{sites}} J_{is}\,\sigma_i + \frac{\sigma_{s,\mathrm{f}}}{3}\sum_{j>k}^{\text{pairs}} J_{sjk}\sigma_j\sigma_k\\ &+ \frac{\sigma_{s,\mathrm{f}}}{3}\sum_{i>k}^{\text{pairs}} J_{isk}\sigma_i\sigma_k + \frac{\sigma_{s,\mathrm{f}}}{3}\sum_{i>j}^{\text{pairs}} J_{ijs}\sigma_i\sigma_j + \cdots\Bigg)\\ &- \Bigg(J_s\,\sigma_{s,\mathrm{i}} + \frac{\sigma_{s,\mathrm{i}}}{2}\sum_j^{\text{sites}} J_{sj}\,\sigma_j + \frac{\sigma_{s,\mathrm{i}}}{2}\sum_i^{\text{sites}} J_{is}\,\sigma_i + \frac{\sigma_{s,\mathrm{i}}}{3}\sum_{j>k}^{\text{pairs}} J_{sjk}\sigma_j\sigma_k\\ &+ \frac{\sigma_{s,\mathrm{i}}}{3}\sum_{i>k}^{\text{pairs}} J_{isk}\sigma_i\sigma_k + \frac{\sigma_{s,\mathrm{i}}}{3}\sum_{i>j}^{\text{pairs}} J_{ijs}\sigma_i\sigma_j + \cdots\Bigg)\Bigg]\end{aligned} \tag{2.62}$$

Combining like terms, this energy change simplifies to:

$$\begin{aligned}\Delta E_{\mathrm{tot}}(\sigma_s) &= N_\mathrm{M}\left[\left(\sigma_{s,\mathrm{f}} - \sigma_{s,\mathrm{i}}\right)J_s + \left(\sigma_{s,\mathrm{f}} - \sigma_{s,\mathrm{i}}\right)\sum_j^{\text{sites}} J_{sj}\sigma_j\right.\\ &\quad\left. + \left(\sigma_{s,\mathrm{f}} - \sigma_{s,\mathrm{i}}\right)\sum_{j>k}^{\text{pairs}} J_{sjk}\sigma_j\sigma_k + \cdots\right]\\ &= N_\mathrm{M}\left[\Delta\sigma_s\left(J_s + \sum_j^{\text{sites}} J_{sj}\sigma_j + \sum_{j>k}^{\text{pairs}} J_{sjk}\sigma_j\sigma_k + \cdots\right)\right]\end{aligned} \tag{2.63}$$

where $\Delta\sigma_s = \pm 2$ and the final expression sums over all interaction terms involving site s.

2.4.3.1.2 Interpreting the Simulation Chemical Potential. Some applications of the GCMC approach, such as determining temperature–composition phase diagrams,[59,61] only use the equality of chemical potentials between two phases, but these chemical potentials are not related to any directly observable properties. In such cases an absolute chemical potential is not needed. Adsorption applications, however, typically require the chemical potential to be related to the temperature and pressure of an adsorbing gas. To understand the connection between the chemical potential used in GCMC simulations, μ_{GCMC}, and these gas-phase conditions, we reexamine the grand canonical partition function and show that their relationship depends on how the CE Hamiltonian is developed.

The grand canonical partition function in eqn (2.55) can also be written as a double sum with the first sum over the number of adsorbates and the second sum over all of the microstates of the system for a given number of adsorbates:

$$Z = \sum_{N_{\mathrm{A}}} \sum_{j}^{\text{microstates}} \exp\left[-\frac{E_{N_{\mathrm{A}},j} - \mu_{\mathrm{A}} N_{\mathrm{A}}}{k_{\mathrm{B}} T}\right] \tag{2.64}$$

Here $E_{N_{\mathrm{A}},j}$ is the energy of a given microstate, j, of the system with N_{A} adsorbates. Because $\mu_{\mathrm{A}} N_{\mathrm{A}}$ does not depend on j, we can move it outside of the microstate sum, which can be further divided into a sum of configurations, $\boldsymbol{\sigma}$, of N_{A} adsorbates and vibrational states, v, of the adsorbates on the surface:

$$Z = \sum_{N_{\mathrm{A}}} \exp\left(\frac{\mu_{\mathrm{A}} N_{\mathrm{A}}}{k_{\mathrm{B}} T}\right) \sum_{\boldsymbol{\sigma}}^{\text{configs.}} \sum_{v}^{\text{vibs.}} \exp\left(-\frac{E_{\sigma,v}}{k_{\mathrm{B}} T}\right) \tag{2.65}$$

The microstate energy, $E_{\boldsymbol{\sigma},v}$, can be written as a sum of the configuration electronic energy, $E_{\boldsymbol{\sigma}}^{\mathrm{elec}}$, the vibrational zero point energy, E_v^{ZP}, and the vibrational thermal energy, E_v^{vib}:

$$E_{\boldsymbol{\sigma},v} = E_{\boldsymbol{\sigma}}^{\mathrm{elec}} + E_v^{\mathrm{ZP}}(\boldsymbol{\sigma}) + E_v^{\mathrm{vib}}(\boldsymbol{\sigma}) \tag{2.66}$$

Within the harmonic approximation, the sum over vibrational states, v, of $n_{\mathrm{dof}}^{\mathrm{vib}} N_{\mathrm{A}}$ oscillators simplifies to a product of single-vibration partition functions, $\exp\left[-\frac{h\nu_i(\boldsymbol{\sigma})}{2k_{\mathrm{B}}T}\right] q_i^{\mathrm{vib}}(\boldsymbol{\sigma})$ where q^{vib} was defined previously in eqn (2.19):

$$Z = \sum_{N_{\mathrm{A}}} \exp\left(\frac{\mu_{\mathrm{A}} N_{\mathrm{A}}}{k_{\mathrm{B}} T}\right) \sum_{\boldsymbol{\sigma}}^{\text{configs.}} \exp\left(-\frac{E_{\boldsymbol{\sigma}}^{\mathrm{elec}}}{k_{\mathrm{B}} T}\right) \prod_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}} N_{\mathrm{A}}} \exp\left[-\frac{h\nu_i(\boldsymbol{\sigma})}{2k_{\mathrm{B}} T}\right] q_i^{\mathrm{vib}}(\boldsymbol{\sigma}) \tag{2.67}$$

We can rewrite this product of individual vibrational modes in terms of a configuration-dependent zero-point energy, $E_{\boldsymbol{\sigma}}^{\mathrm{ZP}}$, and vibrational free energy, $F_{\boldsymbol{\sigma}}^{\mathrm{vib}}$, similar to the "coarse-grained" free energy discussed by van de Walle *et al.*[56]

$$\prod_{i}^{n_{\mathrm{dof}}^{\mathrm{vib}} N_{\mathrm{A}}} \exp\left[-\frac{h\nu_i(\sigma)}{2k_{\mathrm{B}} T}\right] q_i^{\mathrm{vib}}(\boldsymbol{\sigma}) = \exp\left(-\frac{E_{\boldsymbol{\sigma}}^{\mathrm{ZP}} + F_{\boldsymbol{\sigma}}^{\mathrm{vib}}}{k_{\mathrm{B}} T}\right) \tag{2.68}$$

In this derivation, we do not treat the metal surface explicitly, so the terms $E_{\boldsymbol{\sigma}}^{\mathrm{ZP}}$ and $F_{\boldsymbol{\sigma}}^{\mathrm{vib}}$ used here would correspond to $\Delta_{\boldsymbol{\sigma}}E^{\mathrm{ZP}}$ and $\Delta_{\boldsymbol{\sigma}}F^{\mathrm{vib}}$, respectively, as used previously in deriving the surface free energy (Section 2.4.2.1). Utilizing these configuration-dependent vibrational energies and recombining all of the exponential terms, we reach the following expression for the grand canonical partition function:

$$Z = \sum_{N_{\mathrm{A}}} \sum_{\boldsymbol{\sigma}}^{\text{configs.}} \exp\left(-\frac{E_{\boldsymbol{\sigma}}^{\mathrm{elec}} + E_{\boldsymbol{\sigma}}^{\mathrm{ZP}} + F_{\boldsymbol{\sigma}}^{\mathrm{vib}} - \mu_{\mathrm{A}} N_{\mathrm{A}}}{k_{\mathrm{B}} T} \right) \tag{2.69}$$

It is important to note that a CE can be fit to any form of the system energy (*e.g.* raw DFT energies, formation energies, average adsorption energies, vibrational free energies, *etc.*). The choice is somewhat arbitrary and should be made taking into account minimization of the DFT error (energy differences are generally more reliable than raw DFT energies) and convenience for the intended application. As noted above, the GCMC acceptance criterion [eqn (2.56)] is based on $\Delta E_{\mathrm{tot}} - \mu_{\mathrm{GCMC}}\Delta N\mathrm{A}$. We use the notation μ_{GCMC} to emphasize the fact that the chemical potential used when running a GCMC simulation is not necessarily equivalent to the absolute chemical potential of the system, μ_{A}. These are equivalent only when the selected Hamiltonian exactly describes the system energy, incorporating all degrees of freedom. However, we also pointed out that the GCMC algorithm only samples over different configurations and numbers of adsorbates, so these additional degrees of freedom (*i.e.* vibrations) must be able to be expressed as a function of N_{A} and $\boldsymbol{\sigma}$. For example, the zero-point and vibrational free energies in eqn (2.69) are configuration dependent, allowing a temperature-dependent cluster expansion to be fit accounting for these additional terms.[56] Use of this temperature-dependent CE as the Hamiltonian in GCMC would allow us to equate μ_{GCMC} and μ_{A} directly.

It is often impractical to use this complete Hamiltonian, and we instead utilize a simplified CE that incorporates only configuration-dependence of the electronic energy, but not vibrations. This simplification is made on the basis of the assumption that the configurational and internal degrees of freedom are separable and the internal parts can be calculated independently. However, when adopting a simplified CE Hamiltonian, the overall physical description of the system must remain the same, so the free energy per adsorbate due to the internal degrees of freedom, $F_{\mathrm{internal}}(T)$, that is ignored in the CE Hamiltonian gets incorporated into the simulation chemical potential term. Additionally, if the CE is fit to some form of the system energy that involves additional reference terms beyond the electronic and vibrational terms given in eqn (2.69), as is the case when fitting to the formation energy [eqn (2.9)], then any energy references, $E_{\mathrm{A}}^{\mathrm{ref}}$, that depend on N_{A} are also incorporated into the simulation chemical potential:

$$\mu_{\mathrm{GCMC}} = \mu_{\mathrm{A}} - F_{\mathrm{internal}}(T) - E_{\mathrm{A}}^{\mathrm{ref}} \tag{2.70}$$

This is illustrated for the case of fitting the CE to the formation energy, where $E_{\mathrm{A}}^{\mathrm{ref}} = E_{\mathrm{A(g)}}^{\mathrm{DFT}}$ and $F_{\mathrm{internal}}(T) = E_{\mathrm{A}^*}^{\mathrm{ZP}} + F_{\mathrm{A}^*}^{\mathrm{vib}}$.

If we cluster expand the formation energy with respect to the clean surface and the adsorbing gas, we can write the argument of the acceptance criterion using the definition of formation energy in eqn (2.9):

$$\begin{aligned}\Delta E_{\mathrm{tot}}(\boldsymbol{\sigma}) - \mu_{\mathrm{GCMC}}\Delta N_{\mathrm{A}} &= N_{\mathrm{M}}\Delta(\Delta E_{\mathrm{f}}(\boldsymbol{\sigma})) - \mu_{\mathrm{GCMC}}\Delta N_{\mathrm{A}} \\ &= N_{\mathrm{M}}\Delta\left[\frac{1}{N_{\mathrm{M}}}\left(E_{\boldsymbol{\sigma}}^{\mathrm{DFT}} - E_{\mathrm{clean}}^{\mathrm{DFT}} - N_{\mathrm{A}}(\boldsymbol{\sigma})E_{\mathrm{A(g)}}^{\mathrm{DFT}}\right)\right] - \mu_{\mathrm{GCMC}}\Delta N_{\mathrm{A}}\end{aligned} \tag{2.71}$$

Comparison with the grand canonical partition function [eqn (2.69)] reveals that:

$$\Delta\left(E_{\boldsymbol{\sigma}}^{\mathrm{DFT}} - E_{\mathrm{clean}}^{\mathrm{DFT}} - N_{\mathrm{A}}E_{\mathrm{A(g)}}^{\mathrm{DFT}}\right) - \mu_{\mathrm{GCMC}}\Delta N_{\mathrm{A}} = \Delta\left(E_{\boldsymbol{\sigma}}^{\mathrm{elec}} + E_{\boldsymbol{\sigma}}^{\mathrm{ZP}} + F_{\boldsymbol{\sigma}}^{\mathrm{vib}} - \mu_{\mathrm{A}}N_{\mathrm{A}}\right) \tag{2.72}$$

The Δs refer to a change in configuration, $\boldsymbol{\sigma}$, and the number of adsorbates, N_{A}, but the clean surface reference is unaffected by these changes and cancels out. The terms $E_{\boldsymbol{\sigma}}^{\mathrm{DFT}}$ and $E_{\boldsymbol{\sigma}}^{\mathrm{elec}}$ are identical in this context and also cancel, giving:

$$-\Delta N_{\mathrm{A}}\left(E_{\mathrm{A(g)}}^{\mathrm{DFT}} + \mu_{\mathrm{GCMC}}\right) = \Delta\left(E_{\boldsymbol{\sigma}}^{\mathrm{ZP}} + F_{\boldsymbol{\sigma}}^{\mathrm{vib}}\right) - \mu_{\mathrm{A}}\Delta N_{\mathrm{A}} \tag{2.73}$$

We showed earlier in Section 2.4.2.1 that by assuming that the vibrational modes of each adsorbate on the surface are approximately equal and independent of configuration we can write:

$$E_{\sigma}^{\mathrm{ZP}} = N_{\mathrm{A}}E_{\mathrm{A}^*}^{\mathrm{ZP}} \tag{2.74}$$

$$F_{\sigma}^{\mathrm{vib}} = N_{\mathrm{A}}F_{\mathrm{A}^*}^{\mathrm{vib}} \tag{2.75}$$

where $F_{\mathrm{A}^*}^{\mathrm{vib}}$ and $E_{\mathrm{A}^*}^{\mathrm{ZP}}$ were defined previously in eqns (2.27) and (2.28), respectively. Substitution of these approximations into eqn (2.73) allows us to factor out ΔN_{A} from each term. Dividing by ΔN_{A} and rearranging to solve for μ_{GCMC}, we obtain:

$$\mu_{\mathrm{GCMC}} = \mu_{\mathrm{A}} - E_{\mathrm{A}^*}^{\mathrm{ZP}} - F_{\mathrm{A}^*}^{\mathrm{vib}} - E_{\mathrm{A(g)}}^{\mathrm{DFT}} \tag{2.76}$$

2.4.3.1.3 Absolute and Relative Chemical Potentials. Eqn (2.76) can be used to recover absolute gas-phase chemical potentials from the GCMC analysis only for the case with $E^{\mathrm{CE}} = \Delta E_{\mathrm{f}}$. Similar expressions can be derived for alternative definitions of E^{CE}. Continuing with this example, the absolute chemical potential is:

$$\mu_{\mathrm{A}} = \mu_{\mathrm{GCMC}} + E_{\mathrm{A}^*}^{\mathrm{ZP}} + F_{\mathrm{A}^*}^{\mathrm{vib}} + E_{\mathrm{A(g)}}^{\mathrm{DFT}} \tag{2.77}$$

Alternatively, a $\Delta\mu_A$ can be defined as introduced in Section 2.4.2.1. Taking $\mu_{ref} = E_{A(g)}(0\ K) = E^{DFT}_{A(g)} + E^{ZP}_{A(g)}$ as was done for the ground state analysis and substituting into eqn (2.77) gives:

$$\Delta\mu_A = \mu_A - E^{DFT}_{A(g)} - E^{ZP}_{A(g)} = \mu_{GCMC} + E^{ZP}_{A^*} + F^{vib}_{A^*} - E^{ZP}_{A(g)} \tag{2.78}$$

As mentioned previously, when using this reference, the condition $\Delta\mu_A = 0$ corresponds to a surface in equilibrium with a gas reservoir at 0 K.

To make the chemical potential change more physically relevant, we can alternatively define the reference potential as the chemical potential of the gas-phase adsorbate at the simulation temperature and a standard state pressure P°:

$$\mu_{ref} = \mu_{A(g)}(T, P^\circ) = \mu^\circ_{A(g)}(T) \tag{2.79}$$

For this choice of reference, we write the chemical potential change as $\Delta\mu^\circ_A$ to distinguish it from the previous case relative to a 0 K reference. In this way, $\Delta\mu^\circ_A = 0$ corresponds to the surface in equilibrium with the gas-phase reservoir at the same temperature and the reference pressure. Negative values of $\Delta\mu^\circ_A$ would indicate equilibrium with a pressure less than P° and positive values with a pressure greater than P°. We have already defined $\mu_{A(g)}(T, P_A)$ in eqn (2.34), which is rewritten here with $P_A = P^\circ$.

$$\mu^\circ_{A(g)}(T) = \mu_{A(g)}(T, P^\circ) = E^{DFT}_{A(g)} + E^{ZP}_{A(g)} + \Delta G^\circ_{A(g)}(T) \tag{2.80}$$

This definition of $\mu^\circ_{A(g)}(T)$ ultimately leads to an expression for $\Delta\mu^\circ_A$ in terms of the simulation chemical potential, μ_{GCMC}:

$$\Delta\mu^\circ_A = \mu_A - \mu^\circ_{A(g)}(T) = \mu_{GCMC} + E^{ZP}_{A^*} + F^{vib}_{A^*} - E^{ZP}_{A(g)} - \Delta G^\circ_{A(g)}(T) \tag{2.81}$$

2.4.3.1.4 Relating Simulation Chemical Potential to Gas-Phase Pressures. Gas pressures are often more interesting than actual chemical potentials. When this is the case, the methods discussed in Section 2.4.2.2 can be applied. The expressions for μ in eqns (2.34) and (2.77) can be equated to show that, for a simple scenario of an ideal gas reservoir of pure A(g):

$$\mu_{GCMC} + E^{ZP}_{A^*} + F^{vib}_{A^*} + E^{DFT}_{A(g)} = E^{DFT}_{A(g)} + E^{ZP}_{A(g)} + \Delta G^\circ_{A(g)}(T) + k_B T \ln\left(\frac{P_A}{P^\circ}\right) \tag{2.82}$$

Rearranging gives an expression for the gas pressure in terms of μ_{GCMC}:

$$\ln\left(\frac{P_A}{P^\circ}\right) = \frac{1}{k_B T}\left[\mu_{GCMC} + E^{ZP}_{A^*} - E^{ZP}_{A(g)} + F^{vib}_{A^*} - \Delta G^\circ_{A(g)}(T)\right] \tag{2.83}$$

This final expression is for the special case where $E^{CE} = \Delta E_f$ with an ideal gas reservoir of pure A(g). Alternative scenarios with different CE definitions or gas

reservoir compositions could also be derived in a similar way as described in this section and Section 2.4.2.2.

2.4.3.2 Chemical Potential vs. Coverage

One result that comes from a GCMC simulation of adsorption is the average adsorbate coverage, a quantity that can also be extracted from temperature programmed desorption (TPD), X-ray photoelectron spectroscopy (XPS), and Auger electron spectroscopy (AES) experiments. At equilibrium the gas-phase reservoir temperature and chemical potential control the surface coverage, and knowing this relationship as well as mapping the chemical potential to gas-phase pressure as discussed in Section 2.4.2.2 enables the prediction of equilibrium surface coverage under various reaction conditions.

As noted in the previous section, the simulation chemical potential, μ_{GCMC}, must be interpreted in the context of how the CE Hamiltonian is defined. Additional insight can be gained from understanding how the simulation chemical potential relates to other measurable quantities as well as to the equilibrium surface coverage.

2.4.3.2.1 Relating Simulation Chemical Potential and Binding Energy. In previous sections the adsorbate chemical potentials were defined indirectly by assuming equilibrium with a gas-phase reservoir and using eqn (2.34) to define the gas-phase chemical potentials. In eqn (2.76) we identified how the absolute chemical potential relates to the simulation chemical potential for the specific case where the CE is fitted to the formation energy defined by eqn (2.9). Alternatively, we may also define the chemical potential of the adsorbate on the surface directly by applying the thermodynamic definition of chemical potential as the change in free energy with respect to changes in the concentration of a species, shown in eqn (2.84). Helmholtz free energy is used because the system is at constant volume and temperature.

$$\mu_{\mathrm{A}^*} = \left(\frac{\partial F(T, V, N_{\mathrm{A}})}{\partial N_{\mathrm{A}}}\right)_{T,V} \tag{2.84}$$

Because adsorbates are discrete, we can approximate the differential by finite difference in the Helmholtz energy before and after the addition of one adsorbate to the surface. Applying the definition and simplifications of Helmholtz energy given in eqns (2.17)–(2.19), the chemical potential becomes

$$\begin{aligned}\mu_{\mathrm{A}^*} &= \frac{\langle F(T, V, N_{\mathrm{A}}+1)\rangle - \langle F(T, V, N_{\mathrm{A}})\rangle}{(N_{\mathrm{A}}+1) - N_{\mathrm{A}}} \\ &= \langle E^{\mathrm{DFT}}_{N_{\mathrm{A}}+1} + E^{\mathrm{ZP}}_{N_{\mathrm{A}}+1} + F^{\mathrm{vib}}_{N_{\mathrm{A}}+1}\rangle - \langle E^{\mathrm{DFT}}_{N_{\mathrm{A}}} + E^{\mathrm{ZP}}_{N_{\mathrm{A}}} + F^{\mathrm{vib}}_{N_{\mathrm{A}}}\rangle\end{aligned} \tag{2.85}$$

where $\langle\ \rangle$ indicates an ensemble average. Combining like terms and again assuming all adsorbate vibrational modes are identical and independent of the surface, the expression for μ_{A^*} simplifies to:

$$\mu_{A^*} = \langle E^{DFT}_{N_A+1} \rangle - \langle E^{DFT}_{N_A} \rangle + E^{ZP}_{A^*} + F^{vib}_{A^*} \tag{2.86}$$

At equilibrium $\mu_{A^*} = \mu_A$. Using this equality to combine the above expression for μ_{A^*} with eqn (2.76), we obtain the following result, which applies specifically for cases when the CE is fit to formation energies:

$$\begin{aligned} \mu_{GCMC} &= \mu_{A^*} - E^{ZP}_{A^*} - F^{vib}_{A^*} - E^{DFT}_{A(g)} \\ &= \left(\langle E^{DFT}_{N_A+1} \rangle - \langle E^{DFT}_{N_A} \rangle + E^{ZP}_{A^*} + F^{vib}_{A^*} \right) - E^{ZP}_{A^*} - F^{vib}_{A^*} - E^{DFT}_{A(g)} \\ &= \langle E^{DFT}_{N_A+1} \rangle - \langle E^{DFT}_{N_A} \rangle - E^{DFT}_{A(g)} \end{aligned} \tag{2.87}$$

By returning the $\langle E^{DFT} \rangle$ terms to a differential form, we can rewrite the last part of this expression to show the relationship of μ_{GCMC} with the differential binding energy defined in eqn (2.14):

$$\mu_{GCMC} = \frac{\Delta \langle E^{DFT} \rangle}{\Delta N_A} - E^{DFT}_{A(g)} = \frac{d \langle E^{DFT} \rangle}{dN_A} - E^{DFT}_{A(g)} \tag{2.88}$$

The only difference between eqn (2.14) and eqn (2.88) is that the GCMC chemical potential is based on the ensemble average of the DFT energy calculated over the course of the simulation, which incorporates the configurational variability introduced by finite temperature in the simulation and is not restricted to the 0 K minimum energy hull. However, if we consider the limiting behavior at 0 K, then only the ground states on the minimum energy hull will be accessible in the GCMC simulation; the average energies will converge to the minimum energy hull values, and the 0 K simulation chemical potential will become exactly equal to the 0 K differential binding energy. These results apply specifically when the cluster expansion is fit to formation energies, but a similar derivation could be made for other cases as well, with results differing slightly to match the specific situation.

2.4.3.2.2 O-Pt(321) GCMC Simulation Results. We performed GCMC simulations on the O–Pt(321) system for temperatures ranging from 10 K to 1150 K and chemical potentials ranging from −1.9 to 0.8 eV/O. Each simulation consisted of a period of equilibration followed by collection of statistics. For each T and μ condition tested, two independent simulations were run in parallel, one with an ordered ground state initial configuration and the other with a random disordered configuration. The simulations were considered equilibrated once distributions of ΔE_{tot} due to flipping occupancies of single sites, tabulated independently for each simulation, differed by less than a specified threshold.

2.4.3.2.3 Comparing Chemical Potential References. The average adsorbate coverages obtained from these simulations are plotted for a few temperatures in Figure 2.14 relative to three different definitions of chemical potential discussed above. In the upper plot, we show the raw simulation chemical potential, μ_{GCMC}, the middle plot shows the chemical potential change with respect to a 0 K gas-phase reference, $\Delta\mu_O$, as defined in eqn (2.78), and the

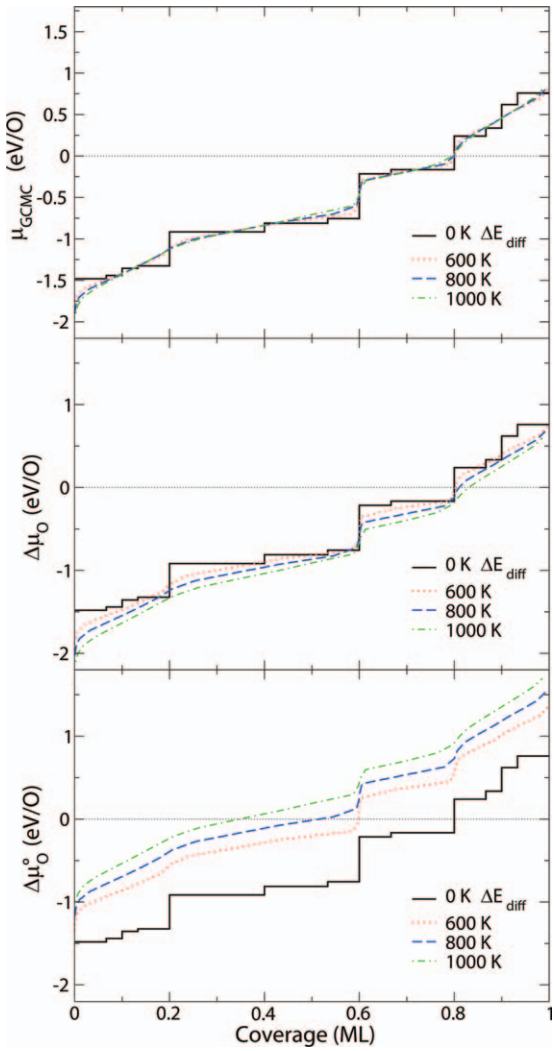


Figure 2.14 Simulation oxygen chemical potential μ_{GCMC} (top), chemical potential change with respect to a 0 K gas-phase reference $\Delta\mu_O$ (middle), and chemical potential change with respect to a standard state gas-phase reference $\Delta\mu_O$ (bottom) *vs.* coverage for various simulation temperatures. The 0 K differential binding energy is also shown for comparison.

lower plot shows the chemical potential change relative to the standard state gas-phase reference, $\Delta\mu_O^\circ$, as given in eqn (2.81).

Plotting the simulation chemical potentials versus the average coverage at different simulation temperatures (Figure 2.14, top) illustrates the relationship between μ_{GCMC} and the 0 K differential binding energy derived in eqn (2.88). The effect of configurational entropy becomes evident with increasing temperature as the μ_{GCMC} *vs.* θ curves become smoother and more linear. If we consider the simulation chemical potential as an average differential binding energy [eqn (2.88)], then the coverage at which μ_{GCMC} is zero represents the maximum equilibrium coverage that can be obtained by dosing with the reference gas [in this case O_2(g)]. In the upper plot of Figure 2.14 we see that, as temperature increases, the smoothing of the curves shifts this maximum coverage from 0.8 ML for the 0 K case down only slightly to 0.797 ML at 1000 K. Since μ_{GCMC}, as defined in eqn (2.88), ignores the contributions of zero-point energy and vibrations, the small change in predicted saturation coverage results solely from configurational effects.

The middle plot in Figure 2.14 shows the relationship of average coverage and the oxygen chemical potential change, $\Delta\mu_O$, relative to 0 K O_2(g). This is defined in eqn (2.78) and expressed here specifically for oxygen adsorption referenced to $\frac{1}{2}O_2$(g):

$$\Delta\mu_O = \frac{1}{2}\left(\mu_{O_2(g)} - E^{DFT}_{O_2(g)} - E^{ZP}_{O_2(g)}\right) = \mu_{GCMC} + E^{ZP}_{O^*} + F^{vib}_{O^*} - \frac{1}{2}E^{ZP}_{O_2(g)} \quad (2.89)$$

The choice of reference state in defining chemical potential changes is arbitrary, so while this definition of $\Delta\mu_O$ is mathematically valid, any physical interpretation of $\Delta\mu_O$ would imply a surface at finite temperature in equilibrium with a 0 K gas. Therefore, this construct cannot be applied meaningfully to a real system. It is essentially a partial correction to account for the vibrational motions of the adsorbates only. Comparing the upper and middle plots of Figure 2.14 in this light, we see that incorporating the zero-point energy and vibrations of adsorbates results in a small increase in the predicted equilibrium coverage at $\Delta\mu_O = 0$, up to 0.82 ML at 1000 K. This agrees neither with intuition that the coverage should decrease with increasing temperature nor with our prior analysis of the 0 K ground states, summarized by the left phase diagram in Figure 2.13. This discrepancy is due to our choice of the reference chemical potential of $\Delta\mu_O$ as the 0 K energy of the gas molecule, which fails to account for the significant loss of entropy that results from taking a molecule from the gas phase and adsorbing it at a surface, and illustrates the problem with attempting to use this "mismatched" $\Delta\mu_O$ to draw physical insight.

For a meaningful interpretation of the GCMC results, we define $\Delta\mu_O^\circ$ from eqn (2.81) referenced to the gas-phase chemical potential at finite temperature and a reference pressure, P°, of 1 bar. This is plotted against average coverage in the lower plot of Figure 2.14. Including the gas-phase entropy contributes significantly to the chemical potential change. As temperature increases, the average coverage where $\Delta\mu_O^\circ = 0$ shifts lower (down to 0.34 ML at 1000 K), in

excellent agreement with the ground state analysis. Physically, this is a consequence of the fact that at higher temperature the additional degrees of freedom in the gas phase compared to the adsorbed phase create an entropic driving force for adsorbates to leave the surface.

2.4.3.2.4 Coverage vs. Pressure. The simulation chemical potential can be related to a gas-phase pressure (or pressures) using eqn (2.83). Treating the pressure of O_2(g) as the independent variable, we plot the average oxygen coverage *vs.* $\ln(P_{O_2})$ for a number of simulation temperatures, shown in Figure 2.15. These results are used further in Section 2.4.3.4 to generate a surface coverage diagram similar to that given previously in the ground state analysis (Figure 2.13) but derived entirely from the GCMC results.

2.4.3.2.5 Modeling Temperature Programmed Desorption. Temperature programmed desorption (TPD) is commonly used to characterize and study adsorption processes. Under certain assumptions, *e.g.* that the barrier to desorption exactly equals the binding energy (*i.e.* non-activated) and that surface diffusion takes place on a much faster time scale than desorption, the coverage-dependent rate of desorption, $r_{des}(\theta, T)$, from a surface can be expressed as:

$$r_{des}(\theta, T) = -\frac{d\theta}{dt} = k(\theta, T)\exp\left(\frac{\mu_{A^*}(\theta)}{k_B T}\right) \quad (2.90)$$

The prefactor, $k(\theta, T)$, is discussed in detail by Kreuzer *et al.* and incorporates such physical processes as the flux of gas molecules to the surface, rotations and vibrations of the adsorbate species, and sticking probabilities for impinging gas molecules.[65–67] Within this formalism the desorption rate depends on the relationship between chemical potential and coverage, which we have shown can

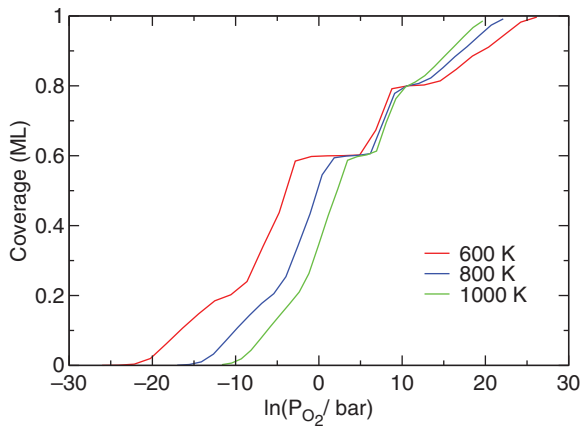


Figure 2.15 Plot of average O coverage *vs.* O_2 pressure at several temperatures for the simple case of a pure O_2 oxygen reservoir. Pressures derived as shown in eqn (2.83).

be obtained from GCMC simulations (Figure 2.14). A full TPD prediction can be obtained from numerous GCMC simulations at closely spaced temperature and chemical potential intervals with interpolation for intermediate conditions. This approach has been successfully used in modeling TPD of systems including CO–Pt(111) and O–Ru(0001),[19,67] however in the particular case of oxygen desorption from the Pt(321) surface, some of the assumptions necessary to use this approach are not valid over the entire range of coverages studied, so we do not present a TPD result for O–Pt(321).

2.4.3.3 Heat Capacity and Order–Disorder Transitions

GCMC simulations produce the average and variance of the energy. The ensemble-averaged variance of the energy, $\langle(\delta E)^2\rangle$, is related to the constant volume heat capacity, C_v, of the system:

$$\frac{\langle(\delta E)^2\rangle}{k_B T^2} = C_v + \delta \tag{2.91}$$

where the δ represents additional terms arising from the fact that C_v is defined for a system with fixed number of particles, but GCMC simulations are performed with fixed chemical potential.[68] Where actual values of C_v are of interest, the δ will be important to include. However, as discussed below, our primary interest in the heat capacity is in observing where it diverges, so the finite correction, δ, to an infinite term will be negligible.

2.4.3.3.1 Divergence of Heat Capacity at Phase Transitions. The heat capacity diverges at phase transitions.[10,69,70] This divergence can be more easily understood by considering the definition of C_v in terms of the system entropy:[71]

$$C_v = T\left(\frac{\mathrm{d}S}{\mathrm{d}T}\right)_V \tag{2.92}$$

Figure 2.16 illustrates schematically how the free energy, entropy, and heat capacity of a closed, constant volume system are related for first-order phase transitions where both phases coexist at the transition temperature, T_p. Each

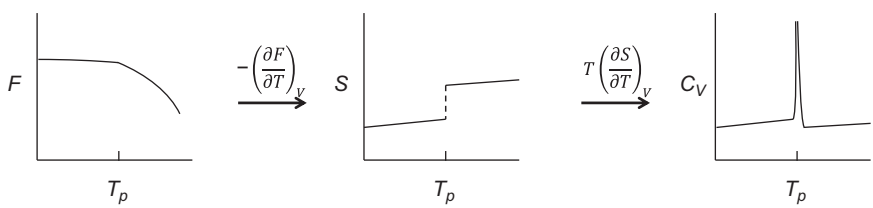


Figure 2.16 Schematic illustration of how free energy, entropy, and heat capacity are related by temperature derivatives. At the transition temperature, T_p, of a first order phase transition, the entropy is discontinuous, leading to the divergent behavior of the heat capacity at that temperature.

phase of a system has an entropy associated with it that relates to the level of ordering in that phase. At T_p both phases are present with different entropies at the same temperature, thus creating a discontinuity in the S *vs.* T plot. The total entropy of the system at that temperature becomes a linear combination of the entropies of the phases in equilibrium and is defined by the relative amounts of each phase in the overall system. The result of this discontinuity is that the heat capacity diverges to infinity at the transition temperature. In continuous phase transitions the heat capacity also diverges, but this results from non-classical behavior near the critical point dominated by fluctuations.[69] Thus, either type of phase transition can be characterized by a divergence of the heat capacity. In practice, it is not possible to distinguish them solely on the basis of the heat capacity behavior owing to the smoothing and broadening that results from finite size effects.[70,72]

Applying eqn (2.91) to our GCMC results for the O–Pt(321) system, we obtain the results in Figure 2.17, where we plot $\frac{\langle(\delta E)^2\rangle}{k_B T^2}$ *vs.* T at a simulation chemical potential of -0.7 eV/O. This chemical potential is shown because it most clearly exhibits a sharp maximum around 800 K, whereas most other results not shown exhibit broad maxima rather than sharp peaks. The presence of a gradual increase and decrease of the heat capacity around the maximum rather than a sharp delta function singularity is a consequence of the finite system size being unable to capture the true behavior at phase transitions.[70,72]

2.4.3.3.2 Identifying and Assigning Phases. To assign the phases identified in this way, we use knowledge of the ordered ground states as well as intuition regarding the relative ordering (and entropy) of each state. An important principle in identifying phases is that a phase, α, stable at low

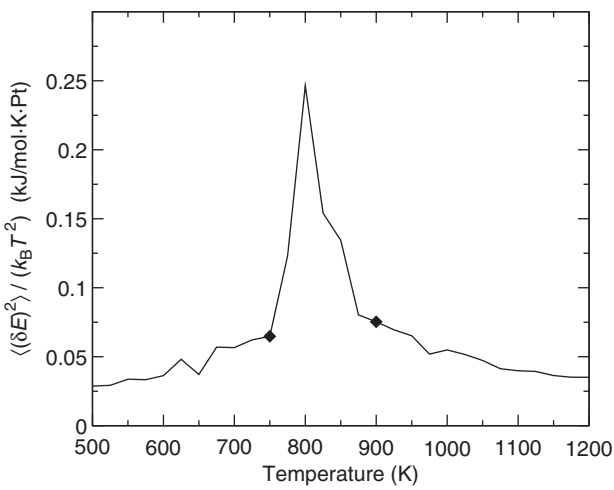


Figure 2.17 Plot of $\frac{\langle(\delta E)^2\rangle}{k_B T^2}$ *vs.* T obtained from GCMC simulations of the O–Pt(321) system run at $\mu_{GCMC} = -0.7$ eV/O. Diamonds correspond to the snapshots shown in Figure 2.18 illustrating the phase change.

temperature must have a lower entropy than a phase, β, stable at higher temperature. This can be shown by considering the Helmholtz free energy of the α and β phases at two temperatures, $T_1 < T_p < T_2$, where T_p is the phase transition temperature. At T_1, the α phase is more stable, so it will have a lower free energy:

$$E^{\alpha} - T_1 S^{\alpha} < E^{\beta} - T_1 S^{\beta} \tag{2.93}$$

At T_2, however, the β phase is more stable:

$$E^{\alpha} - T_2 S^{\alpha} > E^{\beta} - T_2 S^{\beta} \tag{2.94}$$

If we group the energy terms and entropy terms together, and consider a temperature range small enough that E and S are approximately temperature independent, the above inequalities combine to:

$$T_1(S^{\beta} - S^{\alpha}) < E^{\beta} - E^{\alpha} < T_2(S^{\beta} - S^{\alpha}) \tag{2.95}$$

Because $T_1 < T_2$, the entropy difference $S^{\beta} - S^{\alpha}$ must be positive and $S^{\beta} > S^{\alpha}$ in order for the inequality to hold. As a result, we generally assign the highest temperature phase as a completely disordered surface, which normally has the highest entropy. Van de Walle and Ceder point out, however, that while unlikely for metallic systems, some exceptions to this rule of thumb may exist in cases where non-configurational contributions to the entropy (*e.g.* from vibrations) can actually dominate, resulting in an ordered phase becoming favored over the completely disordered phase at higher temperatures.[73]

While it is possible to have multiple phase transitions between different ordered structures before ultimately disordering at high temperature, we generally observe for metal–adsorbate systems that a single transition occurs from the ordered ground states to a disordered phase. Because the lower temperature phases being considered are primarily long-range ordered arrangements of adsorbates on the surface and the higher temperature phases are generally a disordered arrangement of adsorbates, these phase changes are referred to as order–disorder transitions. It is important to confirm the presence and characterization of different phases by examining snapshots from the GCMC simulation on each side of the discontinuity. In Figure 2.18 we present snapshots of the O–Pt(321) system at temperatures below and above the phase change, respectively 750 and 900 K, for the $\mu_{GCMC} = -0.7$ eV/O case in Figure 2.17. These snapshots are near a coverage of 0.6 ML, which has a strong ground state structure with alternating unit cells with four adsorbates and two adsorbates (see Figure 2.11). The ordering can be visually identified by the continuous rows of vacant "ha" and "hh" sites in every other unit cell. In the 750 K (left) snapshot these rows of vacancies, highlighted by green solid lines in the figure, appear at regular intervals and are interrupted by defects only twice. In the 900 K (right) snapshot, however, the rows of vacancies appear at irregular spacings and are interrupted much more frequently. Order parameters may be defined to quantify further what is observed qualitatively in the snapshots.

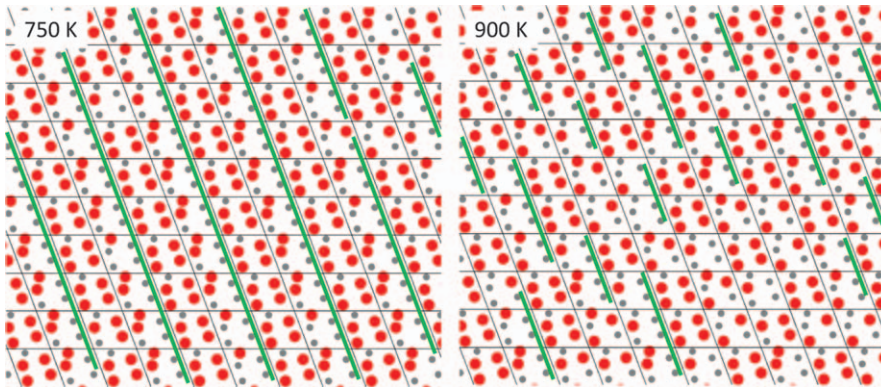


Figure 2.18 Snapshots from GCMC simulations (larger red circles = oxygen, smaller gray circles = vacancy) where $\mu_{GCMC} = -0.7$ eV/O and $T =$ (left) 750 K and (right) 900 K, corresponding to coverages of 0.570 and 0.515 ML, respectively. These conditions are indicated in Figure 2.17 by black diamonds. Green lines highlight ordered segments of more than two "ha" and "hh" vacant sites in a row. At 750 K the adsorbates appear highly ordered in a regular pattern corresponding to the 0.6 ML ground state structure, while at 900 K the 0.6 ML structures are defective and long-range order is lost.

2.4.3.4 Improved Phase Diagrams

Temperature–composition phase diagrams provide information about the region of stability of ordered adsorbate phases. The Monte Carlo methods we have described can also be used to develop such phase diagrams.[21,27,61,74] ATAT has Monte Carlo-based tools to facilitate the construction of temperature–composition phase diagrams. These methods allow the user essentially to trace the boundaries between two phases, keeping the chemical potential of each phase equal while incrementing temperature. By repeating this process for all of the neighboring phases, a complete phase diagram can be obtained, which will identify regions of long-range ordering on the surface as well as the order–disorder transitions mentioned in the previous section. A description of the algorithms, as well as examples of specific applications, is given by van de Walle *et al.*[61]

Temperature–pressure phase diagrams are especially useful for understanding the effect of the reaction environment (temperature and pressure) on the surface phases present. In Section 2.4.2.3 we developed a diagram only for the ordered ground states (Figure 2.13), but this was approximate in the sense that neither configurational variations nor order–disorder transitions were included. The GCMC results incorporate these configurational variations and allow us to identify order–disorder transitions. Such a phase diagram can be constructed from the chemical potential *vs.* coverage results already discussed, and this process is shown schematically in Figure 2.19. The left plot represents the information obtained directly from GCMC simulations, and a corresponding

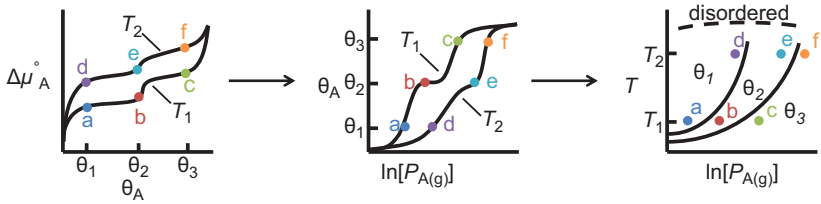


Figure 2.19 Flow chart illustrating the manipulation of GCMC simulation results to obtain a surface coverage diagram. For the O–Pt(321) system the left plot corresponds to the lower plot of Figure 2.14, the middle panel corresponds to Figure 2.15, and the right panel corresponds to Figure 2.20.

result for the O–Pt(321) system was given in the lower plot of Figure 2.14. The chemical potential is converted to pressure according to eqn (2.83) to obtain a plot of coverage *vs.* pressure (Figure 2.19, middle), shown for O–Pt(321) in Figure 2.15. The rightmost plot in Figure 2.19 represents a contour plot of the average coverage at each temperature and pressure. Order–disorder transitions can also be included as discussed in the previous section. Adding these transition temperatures enhances the phase diagram by indicating an upper temperature limit on the regions of stability of each ordered phase. The resulting phase diagram for the O–Pt(321) system is shown in Figure 2.20.

2.4.4 Kinetic Properties from CE/GCMC Methods

It is important to have an understanding of the thermodynamic behavior and limitations of a catalytic system, but ultimately in catalysis we are interested in kinetic properties: rates, rate orders, and activation energies. The cluster expansion approach to modeling thermodynamics can easily be extended to gain access to kinetic properties as well. If the energy differences between two states can be related to the rate of transition between them, then a CE can directly parameterize the rate. There are several examples (see Figure 2.1) where kinetic information on adsorbates might be of interest beyond the thermodynamic properties discussed to this point. For such processes as surface diffusion, desorption, and adsorption, a CE can be used directly to evaluate the energy change between initial and final states. For surface diffusion, this involves moving an adsorbate from one site to another, for desorption changing the occupancy of sites from occupied to unoccupied, and for adsorption changing sites from unoccupied to occupied, but in each case the reaction energy can be obtained directly from a CE. Figure 2.21 illustrates how the CE can be used to obtain kinetic information. We describe the total rate of a surface reaction as a sum of individual reaction events taking place in different "reaction sites" on the surface. This allows us to break down the problem into predicting local rates, r_i, at individual reaction sites as shown in the upper half of Figure 2.21, and finding the probability or likelihood of reaction sites being observed, s_i, under given environmental conditions, illustrated by the lower half of Figure 2.21.

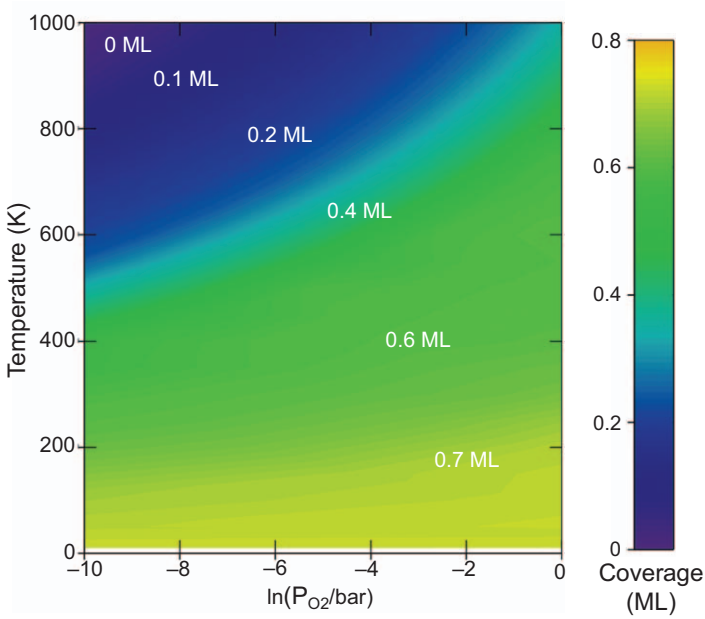


Figure 2.20 Surface coverage diagram identifying the average coverage at a given temperature and pressure for the O–Pt(321) system in equilibrium with an O_2 reservoir. Order–disorder transitions are not shown.

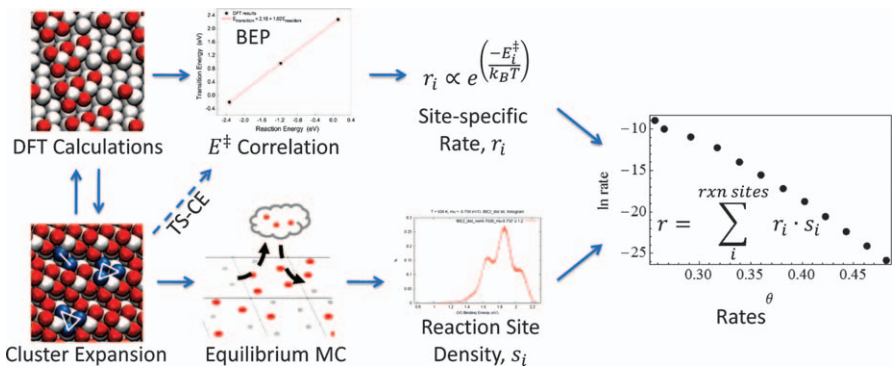


Figure 2.21 Flow chart summarizing how DFT, CE, and MC techniques contribute to deriving kinetic properties of an adsorption system.

There are different ways we might conceive of obtaining transition barriers and local rates from a CE. Owing to the interactions among adsorbates, the activation barrier for surface processes (*e.g.* diffusion, adsorption, *etc.*) in a specific location on the surface with some local coverage and arrangement of nearby adsorbates will depend on the interactions of the diffusing or reacting adsorbates with their spectator neighbors. For this reason, it is possible to cluster expand actual transition state energies (denoted TS-CE in Figure 2.21),

thereby parameterizing the particular adsorbate–adsorbate interactions that influence the transition barrier. Given a database of known transition state energies, we could parameterize a transition state CE to these energies to relate configurations of adsorbates directly to activation energies. This is a non-trivial problem that few people have used in practice but that has been illustrated for some cases.[75] A second approach is to relate activation barriers to either initial or final state energies of the process instead of to specific adsorbate configurations. By this approach, the CE leads to activation barriers *via* a two-step process where a CE fit to stable adsorbate configurations as described in this chapter is used first to calculate reaction energies, and then a second correlation linking reaction and activation energies are used to obtain the barrier. When the correlation between reaction and activation energies is assumed to be linear, these relationships are referred to as Brønsted–Evans–Polanyi (BEP) relationships.[76] BEPs have been used widely in computational catalysis for predicting reaction barriers,[4,44,77] and the use of a BEP relationship in tandem with a CE to predict reaction energies has recently been demonstrated for dissociative O_2 adsorption at a Pt(111) surface.[30] Given a local configuration-specific activation energy, $E_i^{\ddagger}$, *via* either of these two approaches, we write the local configuration-specific rate, r_i, as being proportional to the Arrhenius factor:

$$r_i \propto \exp\left(\frac{-E_i^{\ddagger}}{k_B T}\right) \tag{2.96}$$

Equilibrium Monte Carlo simulations such as those described in Section 2.4.3 provide the ability to average over equilibrium configurations of the system and obtain statistics on the frequency of particular local adsorbate configurations occurring at equilibrium. These local adsorbate configurations, referred to as "reaction sites", are identified by a specified arrangement of one or more lattice sites and an identifying characteristic or property, such as the reaction energy of the process, that distinguishes one reaction site from another. The ability to distinguish reaction sites and count the frequency with which they occur on an equilibrated surface is critical to connecting local, site-specific rates obtained by the methods mentioned in the previous paragraph to a total rate for the whole surface. Local rates, r_i, occurring in different reaction sites, i, can be summed according to the MC-averaged probability of each specific reaction site occurring, s_i, to obtain a prediction of the total rate for the surface:

$$r_{\text{tot}} = \sum_{i}^{\text{rxn sites}} r_i s_i \tag{2.97}$$

Finally, given several total rate predictions obtained in this way for different temperature and chemical potential (*i.e.* pressure) conditions, we can also derive other properties such as apparent activation energies and rate orders.[30]

In systems where the surface cannot be assumed to be at equilibrium, the CE could still be used with alternative MC techniques such as kinetic Monte Carlo (kMC) to access appropriately averaged kinetic properties.

2.5 Conclusions

In this chapter we have discussed a first-principles approach to understanding and appropriately modeling the influence of adsorbate coverage on catalytic systems. We have focused on constructing and using configuration-dependent cluster expansion models of adsorption to identify and characterize thermodynamic ground states and draw connections to the reaction environment. We have demonstrated the versatility of the CE model to perform multiple roles of identifying potentially important, low-energy adsorbate configurations and of ultimately serving as a Hamiltonian for Monte Carlo simulations of large surfaces to obtain average surface properties. These tools allow the user to access thermodynamic properties of adsorption systems such as the ordered ground state structures, coverage-dependent differential binding energies, equilibrium surface coverages, phase diagrams, and order–disorder phase transitions.

We have used the VASP and ATAT codes to apply this first principles approach to oxygen binding at a Pt(321) surface. This low-symmetry surface adds complexity to the adsorption behavior, but is still well within the capability of the methods described to develop a comprehensive adsorption model. The example is somewhat simplified in restricting adsorption to only one species. The problem becomes more complex in the case of multiple adsorbate species coadsorbed on the surface,[23,78] which is always true to at least some extent in catalytic systems. Further complications arise in modeling alloy surfaces, where the metal surface composition itself is heterogeneous and possibly variable in the presence of adsorbates.[22,79,80] More complex multi-component and coupled cluster expansion techniques are necessary to deal with these complications,[56] but the basic concepts illustrated here still apply. The CE technique can even be extended to systems where the surface itself can reconstruct, as has been demonstrated for the missing-row reconstruction.[81] As these models and simulation techniques continue to be developed, improved, and generalized, they will continue to become more widely applied to understanding and modeling the heterogeneous surfaces central to computational catalysis research.

Acknowledgments

We wish to acknowledge funding from the Department of Energy, Basic Energy Sciences under grant DE-FG02-06ER15839 and the Arthur J. Schmitt Foundation, as well as computing resources and technical support from the Notre Dame Center for Research Computing. We also thank Prof. Chris Wolverton, Prof. Chao Wu, Dr. Zhengzheng Chen, Prof. Jean-Sabin McEwen, and Dr. David Schmidt for many helpful discussions and comments.

References

1. I. Chorkendorff and J. W. Niemantsverdriet, *Concepts of Modern Catalysis and Kinetics*, Wiley-VCH, Weinheim, 2007.
2. I. Langmuir, *J. Am. Chem. Soc.*, 1918, **40**, 1361–1403.
3. P. Sabatier, *Ber. Dtsch. Chem. Ges.*, 1911, **44**, 1984–2001.
4. T. Bligaard, J. K. Nørskov, S. Dahl, J. Matthiesen, C. H. Christensen and J. Sehested, *J. Catal.*, 2004, **224**, 206–217.
5. J. K. Nørskov, T. Bligaard, J. Rossmeisl and C. H. Christensen, *Nat. Chem.*, 2009, **1**, 37–46.
6. C. Stampfl, *Catal. Today*, 2005, **105**, 17–35.
7. R. B. Getman, Y. Xu and W. F. Schneider, *J. Phys. Chem. C*, 2008, **112**, 9559–9572.
8. R. A. van Santen, M. Neurock and S. G. Shetty, *Chem. Rev.*, 2010, **110**, 2005–2048.
9. R. I. Masel, *Principles of Adsorption and Reaction on Solid Surfaces*, Wiley, New York, 1996.
10. A. Clark, *The Theory of Adsorption and Catalysis*, Academic Press, New York, 1970.
11. S. D. Miller, N. Inoglu and J. R. Kitchin, *J. Chem. Phys.*, 2011, **134**, 104709.
12. D. J. Schmidt, W. Chen, C. Wolverton and W. F. Schneider, *J. Chem. Theory Comput.*, 2012, **8**, 264–273.
13. R. B. Getman, W. F. Schneider, A. D. Smeltz, W. N. Delgass and F. H. Ribeiro, *Phys. Rev. Lett.*, 2009, **102**, 076101.
14. N. İnoğlu and J. Kitchin, *Phys. Rev. B*, 2010, **82**, 045414.
15. P. Deshlahra, J. Conway, E. E. Wolf and W. F. Schneider, *Langmuir*, 2012, **28**, 8408–8417.
16. R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press, Cambridge, 2004.
17. D. S. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction*, John Wiley & Sons, Hoboken, New Jersey, 2009.
18. J. Sanchez, F. Ducastelle and D. Gratias, *Physica A*, 1984, **128**, 334–350.
19. C. Stampfl, H. J. Kreuzer, S. H. Payne, H. Pfnür and M. Scheffler, *Phys. Rev. Lett.*, 1999, **83**, 2993–2996.
20. M. H. F. Sluiter and Y. Kawazoe, *Phys. Rev. B*, 2003, **68**, 085410.
21. H. Tang, A. van der Ven and B. L. Trout, *Mol. Phys.*, 2004, **102**, 273–279.
22. B. C. Han, A. Van der Ven, G. Ceder and B.-J. Hwang, *Phys. Rev. B*, 2005, **72**, 205409.
23. D. Lerch, O. Wieckhorst, L. Hammer, K. Heinz and S. Müller, *Phys. Rev. B*, 2008, **78**, 121405.
24. D. Lerch, O. Wieckhorst, G. L. W. Hart, R. W. Forcade and S. Mueller, *Model. Simul. Mater. Sci.*, 2009, **17**, 055003.
25. M. Stöhr, R. Podloucky and S. Müller, *J. Phys.-Condens. Mat.*, 2009, **21**, 134017.
26. S. D. Miller and J. R. Kitchin, *Mol. Sim.*, 2009, **35**, 920–927.

27. C. Lazo and F. J. Keil, *Phys. Rev. B*, 2009, **79**, 245418.
28. D. Laks, L. Ferreira, S. Froyen and A. Zunger, *Phys. Rev. B*, 1992, **46**, 12587–12605.
29. K. Reuter, C. Stampfl and M. Scheffler, in *Handbook of Materials Modeling, vol. 1: Fundamental Models and Methods*, ed. S. Yip, Springer, Dordrecht, 2005, pp. 149–194.
30. C. Wu, D. Schmidt, C. Wolverton and W. Schneider, *J. Catal.*, 2012, **286**, 88–94.
31. T. Engel and G. Ertl, *Adv. Catal.*, 1979, **28**, 1.
32. P. R. Norton, in *The Chemical Physics of Solid Surfaces and Heterogeneous Catalysis*, ed. D. A. King and D. P. Woodruff, Elsevier, Amsterdam, **vol. 4**, 1982, 27.
33. J. Wintterlin, S. Volkening, T. Janssens, T. Zambelli and G. Ertl, *Science*, 1997, **278**, 1931–1934.
34. S. Völkening, K. Bedürftig, K. Jacobi, J. Wintterlin and G. Ertl, *Phys. Rev. Lett.*, 1999, **83**, 2672–2675.
35. S. S. Mulla, N. Chen, L. Cumaranatunge, G. E. Blau, D. Y. Zemlyanov, W. N. Delgass, W. S. Epling and F. H. Ribeiro, *J. Catal.*, 2006, **241**, 389–399.
36. D. Parker, M. Bartram and B. Koel, *Surf. Sci.*, 1989, **217**, 489–510.
37. N. Saliba, Y. Tsai, C. Panja and B. Koel, *Surf. Sci.*, 1999, **419**, 79–88.
38. J. F. Weaver, J. J. Chen and A. L. Gerrard, *Surf. Sci.*, 2005, **592**, 83–103.
39. Y. Yeo, L. Vattuone and D. King, *J. Chem. Phys.*, 1997, **106**, 392–401.
40. R. B. Getman and W. F. Schneider, *ChemCatChem*, 2010, **2**, 1450–1460.
41. G. A. Somorjai and J. Carrazza, *Ind. Eng. Chem. Fund.*, 1986, **25**, 63–69.
42. G. Ertl, *Angew. Chem. Int. Ed.*, 1990, **29**, 1219–1227.
43. K. W. Kolasinski, *Surface Science: Foundations of Catalysis and Nanoscience*, John Wiley & Sons, Chichester, 2008.
44. J. K. Nørskov, T. Bligaard, B. Hvolbaek, F. Abild-Pedersen, I. Chorkendorff and C. H. Christensen, *Chem. Soc. Rev.*, 2008, **37**, 2163–2171.
45. T. Jiang, D. J. Mowbray, S. Dobrin, H. Falsig, B. Hvolbaek, T. Bligaard and J. K. Nørskov, *J. Phys. Chem. C*, 2009, **113**, 10548–10553.
46. G. Kresse and J. Furthmüller, *Phys. Rev. B*, 1996, **54**, 11169–11186.
47. P. E. Blöchl, *Phys. Rev. B*, 1994, **50**, 17953–17979.
48. G. Kresse and D. Joubert, *Phys. Rev. B*, 1999, **59**, 1758–1775.
49. J. P. Perdew and Y. Wang, *Phys. Rev. B*, 1992, **45**, 13244–13249.
50. H. J. Monkhorst and J. D. Pack, *Phys. Rev. B*, 1976, **13**, 5188–5192.
51. J. M. Bray and W. F. Schneider, *Langmuir*, 2011, **27**, 8177–8186.
52. P. Feibelman, S. Esch and T. Michely, *Phys. Rev. Lett.*, 1996, **77**, 2257–2260.
53. M. R. McClellan, F. R. McFeely and J. L. Gland, *Surf. Sci.*, 1983, **124**, 188–208.
54. J. M. Bray, J. L. Smith and W. F. Schneider, *Top. Catal.*, 2013, submitted.
55. A. van de Walle, M. Asta and G. Ceder, *Calphad*, 2002, **26**, 539–553.
56. A. van de Walle, *Calphad*, 2009, **33**, 266–278.
57. P. Zhang, *Ann. Stat.*, 1993, **21**, 299–313.

58. J. Shao, *J. Am. Stat. Assoc.*, 1993, **88**, 486–494.
59. A. van de Walle and G. Ceder, *J. Phase Equil.*, 2002, **23**, 348–359.
60. S. Arlot and A. Celisse, *Stat. Surv.*, 2010, **4**, 40–79.
61. A. van de Walle and M. Asta, *Model. Simul. Mater. Sci.*, 2002, **10**, 521–538.
62. W. Brown, R. Kose and D. King, *Chem. Rev.*, 1998, **98**, 797–831.
63. M. V. Bollinger, K. W. Jacobsen and J. K. Nørskov, *Phys. Rev. B*, 2003, **67**, 085410.
64. L. Wang, T. Maxisch and G. Ceder, *Phys. Rev. B*, 2006, **73**, 195107.
65. H. J. Kreuzer and S. H. Payne, in *Computational Methods in Surface and Colloid Science*, ed. M. Borowko, Marcel Dekker, Inc., New York, 2000. pp. 439–480.
66. H. J. Kreuzer, S. H. Payne, A. Drozdowski and D. Menzel, *J. Chem. Phys.*, 1999, **110**, 6982–6999.
67. J.-S. McEwen, S. H. Payne, H. J. Kreuzer, M. Kinne, R. Denecke and H. P. Steinruck, *Surf. Sci.*, 2003, **545**, 47–69.
68. T. L. Hill, *Statistical Mechanics: Principles and Selected Applications*, Dover Publications, New York, 1987.
69. H. B. Callen, *Thermodynamics and an Introduction to Thermostatics*, John Wiley & Sons, New York, 1985.
70. D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics*, Cambridge University Press, Cambridge, 3rd edn, 2009.
71. J. W. Tester and M. Modell, *Thermodynamics and its Applications*, Prentice Hall PTR, New Jersey, 3rd edn, 1997.
72. K. Binder and D. W. Heermann, *Monte Carlo Simulation in Statistical Physics: An Introduction*, Springer, Berlin, 5th Edn, 2010.
73. A. van de Walleand and G. Ceder, *Rev. Mod. Phys.*, 2002, **74**, 11–45.
74. J.-S. McEwen, S. H. Payne and C. Stampfl, *Chem. Phys. Lett.*, 2002, **361**, 317–320.
75. A. Van der Ven, G. Ceder, M. Asta and P. D. Tepesch, *Phys. Rev. B*, 2001, **64**, 184307.
76. M. G. Evans and M. Polanyi, *T. Faraday Soc.*, 1938, **34**, 11.
77. J. K. Nørskov, T. Bligaard, A. Logadottir, S. Bahn, L. B. Hansen, M. Bollinger, H. Bengaard, B. Hammer, Z. Sljivancanin, M. Mavrikakis, Y. Xu, S. Dahl and C. J. H. Jacobsen, *J. Catal.*, 2002, **209**, 275–278.
78. H. Wang and W. F. Schneider, *Catal. Today*, 2011, **165**, 49–55.
79. B. C. Han and G. Ceder, *Phys. Rev. B*, 2006, **74**, 205418.
80. W. Chen, D. Schmidt, W. F. Schneider and C. Wolverton, *J. Phys. Chem. C*, 2011, **115**, 17915–17924.
81. W. Chen, D. Schmidt, W. F. Schneider and C. Wolverton, *Phys. Rev. B*, 2011, **83**, 075415.

CHAPTER 3

Density Functional Theory Methods for Electrocatalysis

KUAN-YU YEH AND MICHAEL J. JANIK*

Pennsylvania State University, Department of Chemical Engineering, University Park, PA 16802
*Email: mjanik@engr.psu.edu

3.1 Introduction

Electrocatalysis involves catalytic reactions occurring in electrochemical systems, where bond breaking and forming on the catalyst surface are coupled with electron and ion transfer. Electrocatalytic reactions occur in fuel cells, with examples such as hydrogen oxidation, methanol oxidation, and oxygen reduction as well as in electrolysis cells, with examples such as hydrogen evolution, water splitting, and carbon dioxide reduction. Density functional theory (DFT) can be used in a similar manner to its application to non-electrochemical catalytic reactions, however, additional complexities arise owing to the electrochemical nature of the catalytic interface. As in typical heterogeneous catalysis, the electrocatalyst is generally a supported nanoparticle, and all of the same challenges in developing appropriate and computationally tractable model systems (use of low-index plane surfaces or small particles as models, for example) apply to electrocatalytic systems. In addition, challenges specific to electrocatalysis arise in representing: (i) the interaction of the catalyst/adsorbate with electrolyte, which is often liquid and contains ionic species, (ii) effects of charging of the surface and separation of charge between the catalytic surface and the electrolyte, (iii) chemical potentials of ionic species that are involved as reactants or products, and (iv) the electrode potential

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

dependent energy of electrons that also serve as reactants or products. We can often take advantage of thermodynamic relationships to overcome challenges (iii) and (iv), but the long length scales associated with charge and ion distribution at the electrochemical interface as well as the long time scales necessary to evaluate dynamics of the liquid/solid interface make all DFT models of electrocatalytic systems an approximation of the true physical system.

In the following section (3.1.1), we introduce the simple example of hydrogen oxidation/evolution to motivate the use of DFT methods to examine electrocatalytic systems. This example illustrates how DFT may be applied in a straightforward manner to examine the relative catalytic properties of various electrocatalysts, while also highlighting the complexities of applying DFT to electrocatalytic systems. Sections 3.1.2–3.1.4 review a few basic electrochemistry principles. This is followed by an overview of DFT modeling approaches used to examine electrocatalytic reaction energetics. After providing this review of the various applications of DFT methods to study electrocatalytic systems, the remaining bulk of the chapter (Section 3.2) is then dedicated to providing detailed examples allowing the reader to work through the application of various DFT models for electrocatalysis investigation.

3.1.1 A Motivating Example: H_2 Oxidation/H_2 Evolution

The oxidation of molecular hydrogen, or the reverse reduction of protons to form H_2 gas, is arguably the simplest electrocatalytic reaction. In the most basic mechanism, the reaction proceeds through two reaction steps:

$$H_{2(gas)} \leftrightarrow 2H^* \tag{3.1}$$

$$H^* \leftrightarrow H^+{}_{(aq)} + e^- \tag{3.2}$$

In step 1 of the oxidation reaction (referred to as the Tafel step), the gaseous hydrogen molecule dissociatively adsorbs to the electrocatalyst surface, where H* represents a surface adsorbed species. In the second step (referred to as the Volmer step), the adsorbed species desorbs to form a proton in the electrolyte [denoted as $H^+{}_{(aq)}$ to signify that the proton is in the electrolyte, often an aqueous solution or hydrated ionic polymer]. An electron is left in the electrode, where it can transfer into an external circuit to provide electrical power to an external device before passing to the cathode where it participates in a reduction reaction. In the reverse hydrogen evolution reaction, these two steps occur in the backward direction with the electron being provided by chemistry occurring at the other electrode and through the external circuit.

Figure 3.1(a) illustrates a free reaction energy diagram of this process occurring over three catalysts. This diagram is drawn at the equilibrium potential for the overall $H_{2(gas)}$ oxidation reaction. As the chemical potentials of the $H_{2(gas)}$ and $H^+{}_{(aq)}$ species are set by their values in the bulk fluid phases, the electrode potential alters only the energy of the electron oxidation product. At the equilibrium potential, the chemical potentials of reactants and products are equivalent, and the overall reaction has no driving force in the forward or

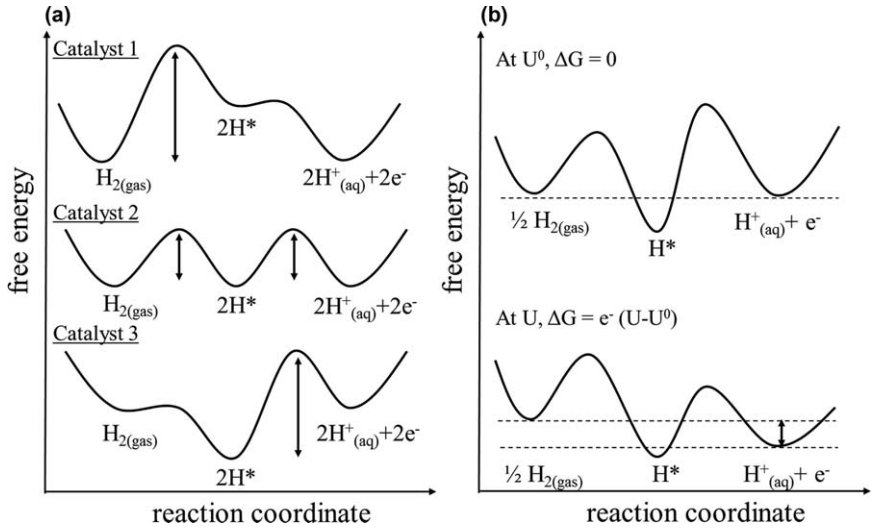


Figure 3.1 Schematic diagram of the potential energy surface for the hydrogen oxidation reaction at (a) three different catalysts and (b) two different potentials.

reverse directions. If the pressure of H_2 gas is 1 bar and the aqueous concentration of protons is 1 molar, the electrode potential (''electron energy'') that makes this reaction free energy equal to zero represents the typical experimental standard hydrogen electrode (SHE) potential. Electrode potentials on a SHE scale then set this potential to 0 and give electrode potentials relative to this value, with negative potentials representing more negatively charged electrodes and thereby less stable electrons. The Nernst equation can be used to adjust the equilibrium potential for the reaction with varying H_2 pressures or proton concentration (pH) values:

$$U = U^0 + \frac{RT}{2F}\ln\frac{\alpha(H^+)^2}{P(H_{2(g)})} \tag{3.3}$$

A ''relative hydrogen electrode'' (RHE) reference provides electrode potentials relative to a 0 set by the equilibrium potential for the overall hydrogen oxidation/evolution reaction at a pH of interest.

The electrocatalyst in this reaction serves to dissociate the hydrogen gas phase molecule and bind atomic hydrogen. Figure 3.1(a) demonstrates how changes in the binding energy of H* on the catalyst surface alter the reaction free energy diagram. Catalyst 1 binds H* to the surface weakest of the three, with the chemical potential of H* (at some coverage that is equivalent on all three surfaces) less stable than the $H_{2(gas)}$. Catalyst 3 binds H* most strongly, with the chemical potential of H* more stable than the $H_{2(gas)}$ species. Catalyst 2 has an intermediate binding energy such that the chemical potential of H* is equal to that of $H_{2(gas)}$. If a Bronsted–Evans–Polanyi (BEP) relationship is assumed, the activation free energy for the dissociation reaction will correlate

linearly with the reaction energy and therefore the H* binding energy. Catalyst 3 would then have the fastest rate of H* formation, and would be expected to establish the highest coverage of H* on the surface. Please see Chapter 1 for a discussion of BEP relationships. To complete the reaction, the H* species will need to desorb from the surface to form the proton–electron pair. Catalyst 3 binds H* the strongest, and therefore may be expected to have the highest barrier and slowest reaction for the desorption step. Catalyst 1 will have the fastest desorption reaction, and Catalyst 2 will have an intermediate rate. The hydrogen oxidation reaction rate will be limited on Catalyst 1 by the rate of H_2 dissociation, whereas Catalyst 3 will be limited by the slow desorption rate. Catalyst 2 is expected to be the optimal catalyst, as it will balance the two attributes to offer the lowest barrier to the rate limiting step. This is a simple illustration of Sabatier's principle, that the optimal catalyst is one with an intermediate binding of a key reaction intermediate, in this case H*. Though this analysis is presented in the context of hydrogen oxidation, the same analysis holds for the reverse hydrogen evolution reaction.

The trade-off in H* binding in determining the optimal catalyst for hydrogen evolution was demonstrated by Trasatti and co-workers,[1,2] as illustrated in Figure 3.2. A "volcano" relationship is observed between the metal–hydrogen binding energy and the hydrogen evolution exchange current. The reduction current is simply the rate of consumption of electrons and therefore a measure of the rate of the reaction. The *y*-axis is labeled as a negative current using the

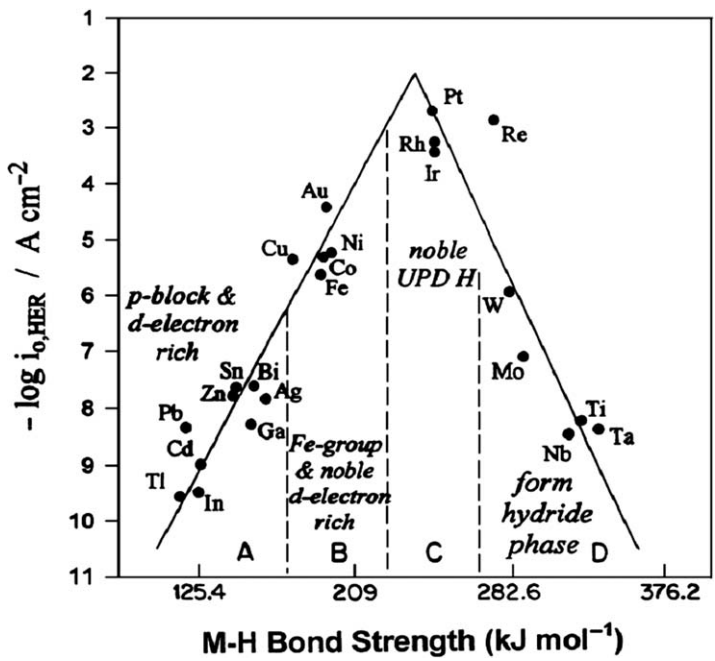


Figure 3.2 Volcano plot for hydrogen evolution as a function of metal–hydrogen bond (M–H) energy.[1,2]

convention that a positive current is an oxidation (production of electrons) current. The exchange current refers to the current at the equilibrium potential, which is a value extracted experimentally by extrapolating the current away from equilibrium back to the equilibrium conditions. Platinum appears near the top of this volcano curve, establishing its high activity for hydrogen oxidation/evolution.

At this point, it is evident how DFT methods may be applied to the evaluation of potential hydrogen oxidation/evolution catalysts. A single energetic descriptor of the optimal catalyst is established, the binding strength of H* to the surface. DFT methods can then be employed to evaluate this binding energy. Nørskov and co-workers have applied DFT methods to investigate the binding energy of H* to numerous bimetallic surfaces.[3–5] With grounding in experimental trends for hydrogen oxidation/evolution rates for single metal surfaces, these calculations can then identify potential bimetallic surfaces that may be closer to the peak of this volcano relationship and provide reaction rates faster than pure platinum. Greeley *et al.*[5] determined that Pt–Bi binary surfaces may offer greater hydrogen evolution rates, a DFT result that was corroborated by subsequent experimental testing. As described in Chapter 1, computational screening procedures such as these are often paired with cost and stability considerations to identify encouraging catalytic materials.

Although the above example shows a simple and successful example of the application of DFT methods to electrocatalyst studies, the DFT calculations employed neglected much of the electrochemical nature of the electrocatalytic interface. This was enabled by the simplicity of the reaction mechanism, empirical knowledge of the key energetic parameter dictating catalyst performance (H* binding), and available experimental data for reaction rates on a number of catalysts. There was no need to identify the reaction mechanism based on DFT analysis, or to employ DFT methods for predicting absolute rates. Coverage effects on dictating catalyst performance could be neglected, as these would only be expected to "flatten" the volcano. Effects of varying electrolyte composition or electrode potential on the reaction rate could be neglected in identifying the optimal catalyst. Despite the simplicity presented above, debate remains on the operable reaction mechanism. In addition to the Tafel and Volmer steps, a Heyrovky step,

$$H^* + H_3O^+_{(aq)} + e^- \leftrightarrow H_{2(gas)} + H_2O_{(aq)} \tag{3.4}$$

may contribute to the observed kinetics. An extensive literature (see for example ref. 6) exists attempting to fit observed electrokinetics of the hydrogen oxidation/evolution reactions to a combination of Tafel–Volmer–Heyrovky reactions. The hydrogen oxidation rate on platinum electrodes (at properly compared electrode potentials) is observed to vary by orders of magnitude between acidic and basic conditions,[7] a result that cannot be explained by the simplified analysis above. More sophisticated models of the electrochemical interface and electrocatalytic reaction energetics are necessary to elucidate more of the subtleties of this or other electrocatalytic reactions.

3.1.2 Electrode Potential Effects on Reaction Energies and Activation Barriers

Prior to further discussion of the development of DFT models for electrocatalytic systems, it is useful to review background material on how electrode potential affects the elementary electrocatalytic reaction free energies and activation barriers. Figure 3.1(b) illustrates a hypothetical reaction free energy diagram for hydrogen oxidation at two different electrode potentials. At the equilibrium potential, U^0, the chemical potentials of the $H_{2(gas)}$ and $H^+ + e^-$ pair are equal. By sign convention, more positive potentials will lower (stabilize, more positively charged electrode) the energy of the electron product. At a potential, U, greater than U^0, the overall reaction will become more favorable, with the reaction free energy (ΔG_{rxn}) reduced by

$$\Delta G_{\mathrm{rxn}} = nFe^-\left(U - U^0\right) \tag{3.5}$$

where e^- represents the elementary charge of an electron, a negative value, and F represents Faraday's constant. The integer n in eqn (3.5) represents the number of electrons generated by reaction, with a negative value used for consumption of electrons. In the example of reaction step 2, n is equal to 1. For convenience, if the electrode potentials are given in the units of volts (V) and the reaction free energy in electronvolts (eV), the elementary charge and Faraday's constant are 1 and energies and potentials can (carelessly) be used interchangeably. Eqn (3.5) provides a simple linear free energy relationship between the reaction free energy and the electrode potential.

The transition state for an electron transfer reaction will occur somewhere between the reactant and product, and thereby can be thought of to involve (potentially) partial electron transfer. Figure 3.3 shows a hypothetical elementary reaction free energy diagram where both reactant and product wells are

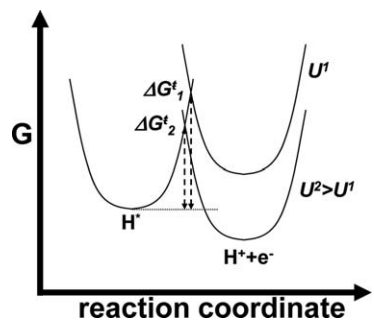


Figure 3.3 A hypothetical free energy relationship for the oxidation of a surface H* species. Both potential energy wells are represented as harmonic, and the product well is drawn at two electrode potentials to illustrate how the reaction barrier change with potential is some fraction of the reaction energy change.

taken to be harmonic. The activation free energy for reaction 2 in the forward direction at potential U^2 can then be written as:

$$\Delta G^t{}_2 = \Delta G^t{}_1 + \beta F\left(U^2 - U^1\right) \tag{3.6}$$

where $F(U^2 - U^1)$ is the elementary reaction free energy change due to the altered electrode potential, and β denotes a symmetry factor. The symmetry factor is similar to a Bronsted–Evans–Polanyi coefficient as it denotes the relationship between the activation barrier change and the reaction energy change, and it is typically between 0.3 and 0.7.[8] The value of U^1 in eqn (3.6) is not necessarily the reaction equilibrium potential. The activation barrier for an electrochemical reaction can therefore be determined at any potential (U^2) once the barrier is known at one potential (U^1) and the symmetry factor β is determined. This is the typical Butler–Volmer formalism of electrochemical reaction kinetics, restated with the reactant being a surface adsorbed species. This is a simplification even for a single elementary step, as β is not necessarily a constant over a substantial potential range. Variable interactions of surface adsorbates with the variable interfacial electric field or potential-dependent changes in the solvation of surface species are two factors that can lead to β varying with electrode potential.

3.1.3 Electrochemical Double-layer Theory

Electrocatalysis most commonly occurs at a metal–solution interface, and the electrochemical potential drop across this interface may affect the stability of species and the rate of elementary reactions. The potential gradient originates from charging of the metal due to the redox chemistry occurring at the electrode or counterelectrode generating charge. Since the metal is a good conductor, any excess charge resides at the surface and attracts ions of opposite sign from the electrolyte. This charge separation creates a potential gradient across the electrode–electrolyte interface. The potential difference across the layer is on the order of 0.1 V for a typical reaction, creating an electric field as large as 10^8 V m^{-1}. This large field provides a driving force for electrode reaction and reorientation of polar solvents, such as water molecules, near the electrode surface.

Several models have been proposed to describe the electrolyte structure, and specifically the charge distribution, within the interfacial region. The simplest model of describing the electrolyte is a layer of oppositely charged ions located at a fixed distance from the electrode surface, as shown in Figure 3.4(a). This model was proposed by Helmholtz.[9] In this model, the electrolyte countercharge to the electrode charge moves as close as possible to the electrode surface without shedding its solvation shell. A sheet of charge equal and opposite to the electrode charge is formed. The two sheets of opposite charge comprise a double layer region that is equivalent to a parallel-plate capacitor. Helmholtz's model predicts a linear potential drop across this double layer region resulting in a constant electric field (the derivative of potential with distance), as well as a

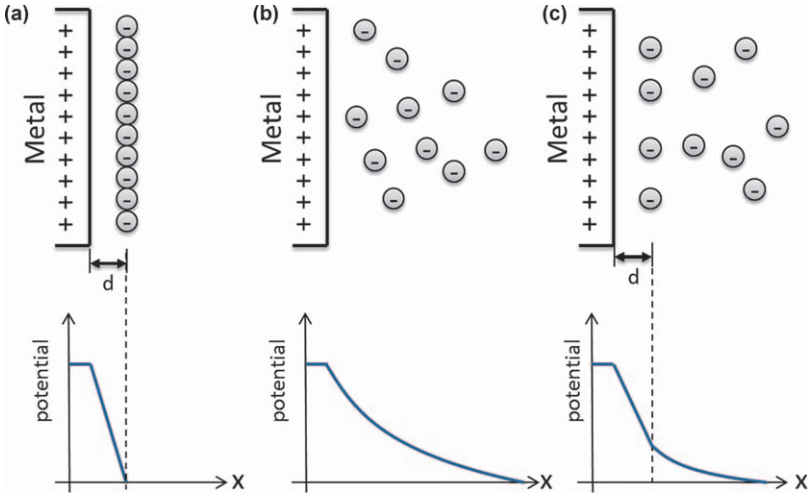


Figure 3.4 Schematic diagrams of the double layer and its corresponding potential profile along the surface normal direction according to the (a) Helmholtz, (b) Gouy–Chapman, and (c) Stern models.

constant capacitance in the double layer region. The Helmholtz model is a convenient simplification of the electrode–electrolyte interfacial structure, however, it does not accurately represent a realistic system.

Even though the charge on the electrode is confined to the surface, the ions in the electrolyte are dynamic as a result of thermal motion, which makes the arrangement of ions less precisely structured than given by the Helmholtz model. This feature was recognized independently by Gouy[10] and Chapman,[11] who incorporated a Boltzmann distribution to describe the thermal fluctuation of ion distribution in the solution. The electrostatic force near the surface is able to overcome the thermal process, leading to the greatest ion concentration. The net ion charge is progressively decreased at greater distance from the surface, with the net charge declining exponentially with distance from the surface. The double layer in this model involves a diffuse layer of charge extending into the solution, as shown in Figure 3.4(b). The Gouy–Chapman theory described the potential dependence of the double layer capacitance on applied potentials and ion concentrations. For a 1 : 1 electrolyte at 25 °C in water, the characteristic thickness of the diffuse layer is $\sim$10 Å. The double layer becomes thicker as the concentration of electrolyte falls. The predicted capacitance typically resembles the behavior of an actual system over a small range of potential and in a very dilute system. The Gouy–Chapman model does not capture the nearly constant capacitance at high potentials or high ion concentration owing to the representation of ionic charges as point charges that can approach infinitely close to the surface. This description is not realistic because ions have a finite size and a layer of solvent on the electrode surface might be considered. Stern[12] combined Helmholtz's and Gouy–Chapman's

ideas and developed a double layer model including a fixed charge plane near the surface and a diffusion layer next to the plane, as shown in Figure 3.4(c). The Helmholtz and Gouy–Chapman theories are two extreme cases of Stern's model. In a system with low electrolyte concentration, the thickness of the diffuse layer is large compared to *d*, and therefore the influence of the Helmholtz layer is not significant. However, at high concentrations or high applied potentials, the ions in solution become more tightly compressed on the electrode surface. The interfacial region resembles the Helmholtz model. Stern's model predicts the general features of ion distribution in real systems.

These macroscopic descriptions for ion distribution capture general qualitative features of the interfacial region, but a precise atomistic representation of electrolyte structure is currently not available for any specific system. The electrolyte structure in the region closest to the electrode will be quite system specific, because it will depend on the solvation structure about the ionic species and interactions of ion or solvent with the electrode surface. We might imagine that the structure about a hydrated polymer electrolyte–electrode interface is quite complex. This interfacial structure is also dynamic on time scales relevant to operation of the electrochemical cell, a requirement as ion transport is essential to operation. Molecular dynamics methods are available to probe this interfacial structure, though no single method allows for accurate electrolyte representation, electrode representation, and consideration of a variable electrode potential in a physically realistic manner. The interested reader might consider starting with a consideration of electrode-charge dynamics.[13] Density functional theory studies of electrocatalysis require either an approximate model of the electrolyte structure (static solvent, a continuum dielectric, prescribed ion/charge distributions, applied electric fields) or assumptions as to the importance (or lack thereof) of adsorbate–electrolyte interaction. The specific approaches are reviewed and illustrated in examples below. It should be highlighted that no single approach is "correct" or necessarily broadly applicable. Further studies are necessary either to provide transferable principles for approximating these interactions or to develop methods that can better take into account the electrode–electrolyte interfacial structure when examining electrocatalytic reaction mechanisms with DFT.

3.1.4 Overview of DFT Models for Electrocatalysis

First-principles quantum chemical methods have allowed elucidation of reaction mechanisms for a variety of heterogeneous catalytic reactions. As discussed above, incorporating the nature of the electrochemical double layer into quantum models is limited by the challenges associated with following the structure and dynamics of the electrolyte over the electrode. Further, to capture electrocatalytic reaction mechanisms accurately using DFT methods, the chemical potential of electrons and ionic species that participate in elementary steps must be evaluated. Several DFT modeling approaches have been developed to include the influence of solvent and/or electrochemical potential on surface reactions and to take into account the chemical potential of ionic species.

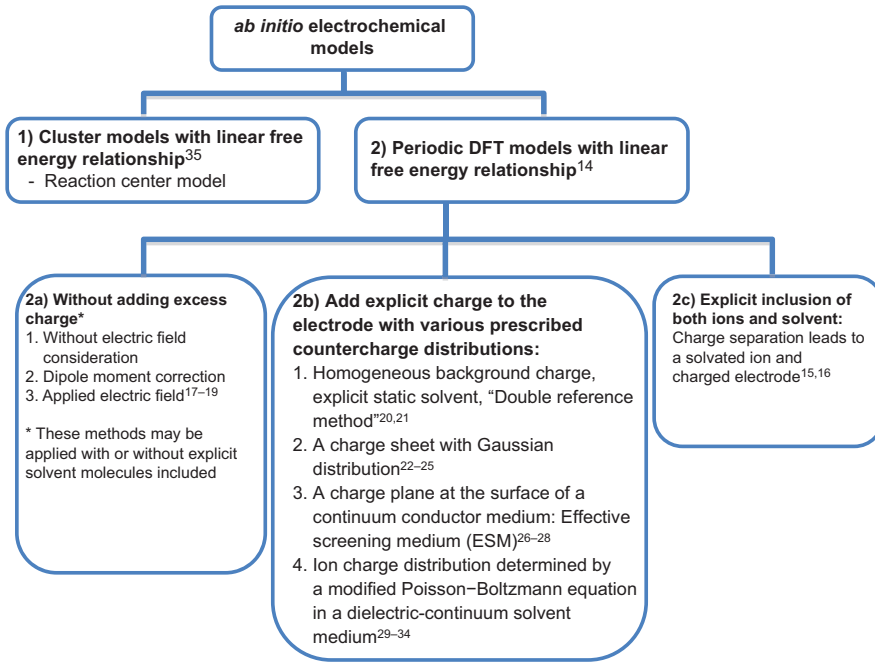


Figure 3.5 A schematic listing of various approaches to applying DFT models to electrocatalytic systems. "Explicit" is used to refer to inclusion of solvent or ions treated with DFT in the model system.

Figure 3.5 schematically lists the various approaches taken in the literature to applying DFT methods to electrocatalytic reactions. These methods are first differentiated by cluster and periodic representation of the electrode surface. The "reaction center" model (Model 1 in Figure 3.5), developed by Anderson and coworkers,[35] is an early attempt to evaluate potential dependent reaction energies and activation barriers. It relies on using a small cluster to represent the "reaction center" of the electrode and evaluates the electron affinity of such cluster. We will not detail this method, because its fidelity is questionable given its arbitrarily small representation of the electrode when considering its electronic structure and lack of scalability to a more accurate electrode representation.

The remaining models to be discussed typically use a periodic representation of the electrode surface, allowing for accurate consideration of the electronic structure of the extended solid. All of the methods take advantage of thermochemical relationships to represent the chemical potential of electrons, ions, or electron–ion pairs. Nørskov *et al.*[14] have employed the linear free energy relationship given in eqn (3.5) and applied a computational reference electrode to perform reaction energy calculations. Because this approach allows the direct use of standard DFT calculations for heterogeneous catalysis, it has been widely used in electrocatalytic systems to predict the reactivity of metal or alloy catalysts. The overview of the theoretical background and examples for this

approach are given in Section 3.2. Here, we provide general discussion of the method differences in their approach to representing the electrified electrode/solvent interface. As discussed above, it is not possible to take into account the dynamic electrolyte structure within a DFT model, and therefore all of the methods surveyed represent an approximation to the actual system.

In addition to the difficulties in representing the time scale of the dynamic electrolyte distribution at the electrode–electrolyte interface, the length scale of relevance is typically unattainable. As mentioned above, uneven ion distribution typically extends to 10 Å or further from the electrode surface. In addition, the lateral spacing between ions typically exceeds that available with the small unit cells used to represent the electrode surface. Therefore, explicit (represented quantum mechanically) inclusion of ions in the electrolyte layer will lead to unphysically high ion concentrations and difficulty in establishing a proper solvation shell about the ion. Model 2c, developed by the Nørskov group, uses explicit H atoms to examine hydrogen oxidation/hydrogen evolution.[15,16] With H atoms added to the DFT unit cell and provided proper solvation with explicit water molecules, the "H electron" will transfer to the metal and an explicit H^+ species will be present in the electrolyte. The electrode potential (energy of the highest energy electron in the electrode) varies with the unit cell size and consequent concentration of hydrogen atoms added. Extrapolation to a large unit cell and low proton concentration was then used to examine hydrogen oxidation/evolution at relevant electrode potentials. Although the explicit inclusion of ions and solvent, albeit in a static solvent structure, is an encouraging advance in DFT model system development, the generalization of this approach to other ions/solvents is unclear.

Generally, there are two different approaches to taking into account the interactions of surface species with electrolyte. The first is to neglect the presence of charge separation within the DFT model or represent it by application of an external electric field (Models 2a). Explicit solvent molecules may still be included to consider solvation of surface species. An electric field may be added along the surface normal direction in a periodic system by adding a dipole sheet in the vacuum region. Alternatively, the electric-field surface adsorbate interaction may be analytically estimated by considering the surface dipole moment change during the reaction and estimating the dipole–field interaction energy difference. The reaction energy change due to the applied electric field and dipole moment variation will be discussed with examples in Section 3.2. The advantage of the applied field or dipole correction approach is that no net charge is introduced, charge neutrality is maintained, and DFT calculations are performed as typically done for heterogeneous catalytic systems. The difficulty of using this type of model is that the connection between electric field and potential relies on the assumption of double layer structure. As the electrode surface charge and counterion distribution are not known *a priori*, assumptions must be made to compare electric field dependent results to electrode potential dependent results. The inclusion of explicit water molecules can be useful in evaluating differences in the solvation stabilization between reactants and products of a surface reaction. The use of a single solvent structure, and

typically with an unconstrained solvent density, makes evaluation of solvation energies approximate. Careful choice of water structures can be used to pick up any significant differences in strongly bound water molecules between adsorbed reactants and products. As a "global minimum" solvation structure is difficult to locate and not necessarily optimal given the unconstrained density, one does have to be careful not to bias the choice of solvation structures towards providing desired results.

The second set of approaches to modeling charge separation at the electrode–electrolyte interface involves explicitly charging the electrode (Models 2b). Explicitly charging the electrode refers to modifying the total charge of the system away from its neutral value. Because the electrode is typically metallic, the added charge (addition or removal of electrons) will segregate to the electrode and occupy surface states (the states about the Fermi level), providing a physically realistic representation of electrode charging. The electrode Fermi level energy will be shifted relative to a constant reference value owing to this charging, both as a consequence of the occupation or depletion of electrons from states near the Fermi level and a shifting of the energy of all states due to their interaction with the added charge. Within a periodic model system, fractional electrons may be added or subtracted, as a fractional electron represents a whole electron distributed over multiple unit cells. Therefore, any desired surface charge density (or electrode potential) may be "dialed" by varying the added charge. Though this provides a physically realistic representation of the charged electrode surface, it leaves two challenges: how to represent the distribution of countercharge and how to ascertain a proper system energy given the varying interactions between the periodic charge and countercharge.

Figure 3.5 lists a number of approaches, with references, for representing the electrolyte within a charged electrode model. These methods differ in the presumed distribution of countercharge as well as in how the solvent is represented. We again emphasize that no method will provide a perfectly accurate representation of any specific system. Charge neutrality in a real system is maintained by localized counterions, as discussed in Section 3.1.3. An approximate model can be developed in the spirit of a Helmholtz double layer, where a fixed charge plane with a Gaussian shape distribution is put at certain distance from the metal electrode (Model 2b.2).[22,25] Instead of putting the countercharge sheet in the vacuum, Otani and Sugino have developed a scheme called effective screening medium (ESM), in which countercharge is added in at the surface of a continuum superconductor above a charged water–metal interface (Model 2b.3).[26–28] These models allow for inclusion of solvent molecules, though the choice of charge plane position must be made with consideration of the water molecule placement. These methods require modification to typical DFT codes.

Jinnouchi and Anderson,[29,30] as well as Goddard and coworkers,[31–33] have instead adopted a Poisson–Boltzmann distribution of countercharge (Model 2b.4). These methods couple this distribution with an implicit continuum solvation model for the solvent (water).[29–33] The continuum model extends the double layer consideration to the diffusion layer region. Jinnouchi and Anderson highlight that strongly bound water molecules must still be included

explicitly to capture reaction energetics accurately.[29,30] The Poisson–Boltzmann countercharge distribution may be viewed as an improvement over the use of a charge-plane however, this remains an imprecise model relative to the actual distribution in any given system. Similarly, although the use of a continuum solvent model may provide for a more accurate representation of a time-averaged solvation of surface intermediates, it requires parameterization. Given that interfacial water dynamics are altered by interaction with the surface, the dielectric constant of water is altered in the interfacial region. Continuum models require parameterization of the water dielectric constant as a function of distance from the electrode. Both the Anderson[29,36] and Goddard[31–33] groups have developed specialized codes to implement this method.

For typical periodic DFT calculations of charged species, a uniform neutralizing background charge is automatically added. Neurock and coworkers[20,21] took advantage of this homogeneous background charge in developing the double-reference method (Model 2b.1). In this method, the electrode potential is varied by adding or subtracting electrons, and the countercharge is represented by this homogeneous background charge distribution. Although the background charge is spread out within the entire unit cell (including within the metal), the added charge influences the orientation of explicit water molecules which provides for screening of the excess electrode charge within a few angstroms from the surface. This effectively mimics the electric field of a Helmholtz double layer region or that created by inclusion of an explicit ion.[21] The method is termed the double-reference method because it uses two reference potentials to establish the effective electrode potential induced for each electrode charge. This approach allows for simulating potentiostatic control of an electrode relative to a reference by varying the electrode charge in a manner similar to an experimental potentiostat. Another advantage of this method is that it requires no modification to typical periodic DFT codes to implement. The model has the disadvantages, as do all of the models discussed in this section, of using a prescribed countercharge distribution that does not necessarily match the actual physical system (homogeneous charge distribution) and an approximate solvation model (a static, typically ice-like water layer). A practical disadvantage is that there is substantial and somewhat complicated ''post-calculation'' analysis required to reference the electrode potential properly and to correct the system energy for the varying number of electrons and interactions with the background charge. The theory behind the double reference method and the practical use of this method are discussed in the case study section (Section 3.2.3). The following section (3.2) provides examples using Models 2a and Model 2b.1.

3.2 Examples Applying DFT Methods to Electrocatalysis

In this section, we discuss examples applying a number of models from Figure 3.5 to electrochemical and electrocatalytic processes. The selection of

reactions focuses on the area of fuel cell electrochemistry, where DFT models are used to understand the molecular details of surface chemical bonding, the reactivity of reaction intermediates, and the reaction mechanisms at the interface between a metal surface and an aqueous solution. Prior to considering surface catalytic reactions, we first examine the electrochemical adsorption of anions. We demonstrate the use of several models, with increasing complexity, to calculate the potential-dependent adsorption energy of anions (Sections 3.2.1 and 3.2.2), and detail how to use these results to simulate linear sweep voltammetry (Section 3.2.3). These models are further extended to consider reaction energies for elementary reaction steps involving proton–electron transfer (Section 3.2.4). The final example illustrates a method for calculating potential-dependent activation barriers.

3.2.1 Simulating the Vacuum–Metal Interface

The simplest DFT model for examination of elementary energetics at an electrode surface is the vacuum slab model, illustrated in Figure 3.6(a) and the most basic form of Model 2a in Figure 3.5. DFT studies of catalysis typically begin by considering the adsorption of reacting species. Often, in electrocatalysis, reactants or products may be ionic species, and calculation of their electrode potential dependent adsorption energy is more complicated than for neutral

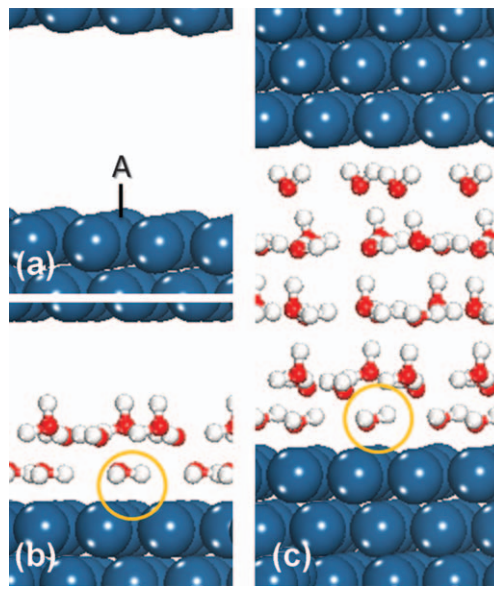


Figure 3.6 Schematic of (a) vacuum slab model, (b) partially solvated model, and (c) fully solvated model. In the vacuum model, *A* represents a surface adsorbate in the solvated models, a water molecule is shown as the adsorbate.

adsorbates. We begin by detailing the calculation of the potential dependent adsorption free energy for a generic A^- anion present in aqueous solution.

To study anion ($A^{n-}_{(aq)}$) adsorption, we initially assume integral electron transfer during the adsorption, described as the reaction:

$$A^{n-}_{(aq)} + * \rightarrow A^* + ne^- \tag{3.7}$$

where * denotes an empty adsorption site and A* is the adsorbed anion. The adsorption free energy of eqn (3.7) (ΔG_{ads}) is determined by subtracting the energy of reactants from that of products:

$$\Delta G_{ads} = G_{A^*} + nG(e^-) - G_{A^{n-}_{(aq)}} - G^* \tag{3.8}$$

Gibbs free energies are used rather than 0 K DFT energies. Within the solvated system, a 0 K state is highly unrealistic. Free energies are more useful as they can directly allow evaluation of equilibrium surface coverage, and the use of free energies allows us to take advantage of experimental thermodynamic equilibrium data in developing computational reference electrode potentials. Except for $G(e^-)$, the Gibbs free energy of each component can be obtained using standard DFT calculations, statistical mechanics equations, and the solvation free energy of the anion. The Gibbs free energy of a gas-phase molecule is calculated as:

$$G_{A^{n-}_g} = E_{DFT} + ZPVE + E(T-0) + PV - (S_{vib} + S_{translation} + S_{rotation})T \tag{3.9}$$

where E_{DFT} is the electronic energy of a system (in this case the isolated anion) determined by DFT calculation, $ZPVE$ is the zero-point vibrational energy correction, PV is the pressure–volume term, and the gas-phase entropy is taken as the sum of the vibrational (S_{vib}), translational ($S_{translation}$), and rotational ($S_{rotation}$) terms at temperature T. The internal energy change at T relative to the 0 K value is calculated through statistical mechanics relationships or taken from tabulated values. The entropic term can also be taken from tabulated data or calculated from standard statistical mechanics relationships, with the only typical DFT input being the calculation of harmonic vibrational frequencies. Likewise, the free energy of an adsorbed species is:

$$G_{A^*} = E_{DFT} + ZPVE - S_{vib}T \tag{3.10}$$

where the vibrational entropy is typically calculated by evaluating the harmonic modes of the adsorbate only and presuming these are sufficiently decoupled from solid or solvent modes. Since ions are dissolved in solution prior to adsorption, the free energy of an aqueous ion ($G_{A^{n-}_{aq}}$) is determined by adding the solvation free energy (ΔG_{sol}) to the gas-phase free energy and correcting for concentration/activity:

$$G_{A^{n-}_{aq}} = G_{A^{n-}_g} + \Delta G^{\circ}_{sol} + RT\ln[A^{n-}] \tag{3.11}$$

ΔG°_{sol} represents the standard solvation energy (at 298.15 K, 1.0 M) and $[A^{n-}]$ is the anion concentration in solution. $RT\ln[A^{n-}]$ is used to correct the concentration deviation from 1 M based on the assumption of ideal solution, though activities are more accurate and might be used if available for the anion of interest. The solvation free energy may be taken from experimental data or estimated using available *ab initio* and continuum solvation approaches.[37]

We can think of eqn (3.8) as providing the adsorption free energy as a function of the electron chemical potential. To a first approximation, only the electron free energy depends on the electrode potential. Corrections can then be added for the potential variance of the G_{A*} value. To make these values relevant to an experimental system, we need to establish a potential reference that is comparable to the experimental potential reference. A simple thermochemical approach is presented in the following section for establishing a reference potential. We use model numbers consistent with Figure 3.5 in labeling each model presented.

3.2.1.1 *Model 2a.1 at Vacuum–Metal Interface*

The energy of a charge in a potential field is given by the charge (ne^-) multiplied by the potential (U) relative to some reference potential (U_{ref}). This gives a simple linear relationship between electrode potential and the electron free energy. Applying the linear free energy relationship, the adsorption free energy (ΔG_{ads}) is potential dependent according to

$$\Delta G_{ads}(U) = \Delta G_{ads}(U_{ref}) + ne^-(U - U_{ref}) = G_{A*} - G_{A^{n-}_{(aq)}} - G^* + ne^-(U - U_{ref}) \tag{3.12}$$

Eqn (3.12) approximates that the free energy of the adsorbed species, G_{A*}, is not potential dependent. This assumption is inherent in Model 2a.1 and is improved upon in subsequent models. The sign convention used for electrode potentials is that negatively charging the electrode leads to a more negative potential. The chemical potential of the "most unstable" electron in the electrode, which will participate in electrochemistry, is the electron chemical potential of interest. This electron becomes less stable as the electrode potential becomes more negative owing to its interaction with the increased negative charge in the electrode. As the anion adsorption process releases an electron, we expect the free energy of adsorption to become less favorable (more positive) as the surface negative charge increases and the electrode potential decreases. Eqn (3.12) correctly represents this relationship when the elementary electron charge is taken as a negative value. For computational convenience, if the free energy values are expressed in electronvolts (eV, with 1 eV equal to 96.48 kJ mol^{-1}), and U is given in volts (V), then the value of e^- is -1.

If the value of U_{ref} in eqn (3.12) is taken as 0, an "absolute reference" is used and the potential U represents the potential relative to that experienced by an electron isolated in vacuum. This reference is also referred to as a "vacuum

reference.'' This is a convenient computational reference, though clearly an isolated electron in vacuum is not convenient experimentally.

Here, we use sulfate anion adsorption to illustrate calculating the potential dependent adsorption energy. Sulfuric acid solution is an important model electrolyte often used in electrochemical experiments. The adsorption of sulfate anions to a Pt electrode is relevant to proton exchange membrane fuel cells (PEMFCs), where the membrane often includes sulfonic acid groups. Figure 3.7 shows the optimized adsorption configurations of the sulfate anion over the Pt(111) electrode. The energetic data for calculating the sulfate ($A^* = SO_4^*$, $n = 2$) adsorption free energy are given in Tables 3.1 and 3.2. These and all data reported herein, unless otherwise mentioned, were obtained using the Vienna ab Initio Simulation Package,[38–40] a 3×3 surface cell of the (111) metal, the PAW approach,[41] and the PW91 exchange-correlation functional.[42,43] Using the metal–vacuum interface model with the data reported in Tables 3.1 and 3.2, the potential dependent sulfate adsorption free energy ($\Delta G_{ads}(U)$) is calculated using eqn (3.13):

$$\begin{aligned}\Delta G_{ads}(U) &= G_{SO_4^*} - G_{SO4^{2-}_{(aq)}} - G^* + 2e^-(U - U_{ref}) \\ &= (-291.28) - (-41.14) - (-261.18) - 2 \times (U - U_{ref})(\text{in eV}) \\ &= 11.04 - 2 \times (U - U_{ref})(\text{in eV}) \end{aligned} \quad (3.13)$$

If the absolute reference is used ($U_{ref} = 0$), any given U is then referenced to the absolute electrode potential (*i.e.* the potential is on a ''vacuum scale'').

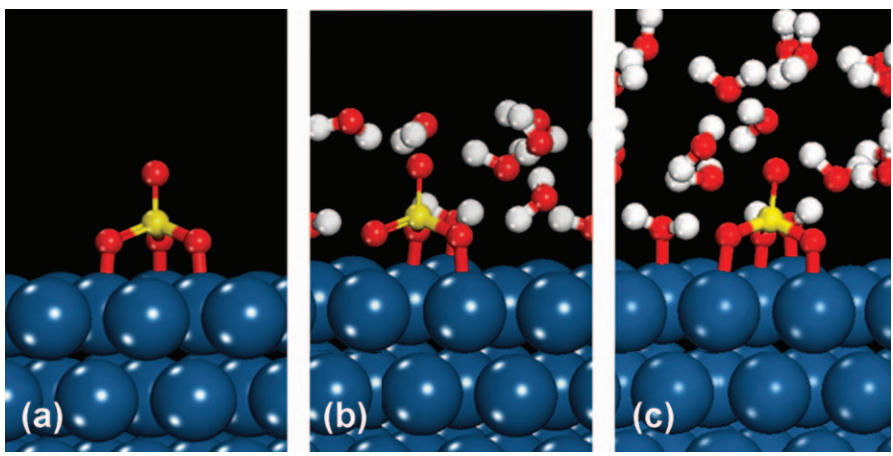


Figure 3.7 Optimized adsorption configurations of sulfate over the Pt(111) surface using (a) a vacuum model, (b) a partially solvated model used in the electric field model, and (c) a fully solvated model used in double reference method.
(Reproduced from ref. 34)

Table 3.1 Calculated DFT energy (E_{DFT}), zero point vibrational energy corrections (*ZPVE*) and entropy (TS_{vib}) contributions to free energy calculations (in eV) for adsorbates over a 3×3 Pt(111) surface.

Adsorbate	E_{DFT}	ZPVE	TS_{vib}	G^a
Pt(111)	−261.18	–	–	−261.18
SO_4*	−291.64	0.46	0.10	−291.28
O_2*	−271.84	0.13	0.09	−271.80
OOH*	−275.76	0.43	0.06	−275.39
OH*	−271.28	0.32	0.01	−270.97
H_2O*	−275.76	0.65	0.05	−275.16

[a]Calculated G for $1/2G_{H_{2(g)}} = -3.46$ eV.

Table 3.2 Calculated DFT energy (E_{DFT}), zero point vibrational energy corrections (*ZPVE*) and entropy (TS_{total}) contributions to free energy calculations (in eV) for aqueous phase Gibbs free energies of a sulfate anion and water molecule.

Components	E_{DFT}	ZPVE	TS_{total}	$\Delta G^{exp}{}_{sol}$ *in 1 M*	G_{aq} *in 1 M*	G_{aq} *in exp. conc.*
$SO_4{}^{2-}{}_{(aq)}$	−30.21	0.41	0.89	−10.36	−40.06	−41.14 (0.007 M)
$H_2O_{(aq)}$	−14.28	0.57	0.59	−0.19	−14.49	−14.35 (55.5 M)

To provide a more direct relationship between $\Delta G_{ads}(U)$ and experimental measurements, the vacuum electrode scale is related to the standard hydrogen electrode (SHE) scale *via* the following equation.[44]

$$U_{SHE} = U - 4.6 \tag{3.14}$$

The value of 4.6 in eqn (3.14) is in the middle of experimental estimates (4.4–4.8),[44] and was calculated using DFT methods and an available H^+ solvation energy.[45] Placing the calculated reaction energy on the standard hydrogen electrode (SHE) scale, eqn (3.13) becomes

$$\Delta G_{ads}(U_{SHE}) = 11.04 - 2(U_{SHE} + 4.6) = 1.84 - 2U_{SHE} \tag{3.15}$$

The equilibrium potential for sulfate adsorption (U_{eq}) is then calculated by

$$\begin{aligned} \Delta G_{ads}(U_{SHE}) &= 1.84 - 2U_{SHE} = 0 \\ U_{eq} &= 0.92\mathrm{V}_{SHE} \end{aligned} \tag{3.16}$$

At a potential more positive than 0.92 V_{SHE}, sulfate anion adsorption on the Pt(111) surface proceeds spontaneously (is exergonic).

The DFT determination of electrode potential dependent ion adsorption energies can be used to evaluate whether electrolyte ions might block catalyst active sites or to evaluate the reactant coverage for ionic reactants. DFT calculations of the acetate potential dependent adsorption free energy were used to

determine that acetate, used often as a fuel for microbial fuel cells (MFCs), likely covers the Pt electrode and thereby limits the sites available for oxygen reduction.[46] The unfavorable adsorption of the borohydride ion to Au(111) at low potentials explains the high overpotential observed experimentally for borohydride electro-oxidation on Au electrodes.[45] Model 2a.1 provides a simple method for evaluating these adsorption free energies, however, it neglects the electrode–electrolyte charge separation that occurs at the interface as well as solvation interactions with the adsorbate. The following sections continue with the sulfate adsorption example to describe how these complexities may be added to the model.

3.2.1.2 *Model 2a.2: Vacuum–Metal Interface with Dipole Moment Correction*

In the first linear free energy approach (Model 2a.1), the energy of the electron(s) transferred on anion adsorption is the only potential dependent term in the adsorption free energy. This approach assumes that the electric field caused by the applied potential (or excess surface charge) has no interaction with the adsorbate or influence on the adsorbate–metal interaction. Two methods to approximate the impact of the electric field effect on elementary surface processes that do not require explicitly charging the metal surface (Models 2a.2 and 2a.3) will be discussed. The dependence of a species' energy within an electric field (F) is

$$E(F) - E(F = 0) = -\mu F + \frac{1}{2}\alpha F^2 + \cdots \tag{3.17}$$

where μ is the species' dipole moment and α is the species' polarizability. In our sign convention, the dipole moment of an adsorbed species that takes on a partial negative charge (relative to a positively charged surface) is positive. For the electrode–electrolyte system, the species of interest is the metal–adsorbate interface. Corrections for the presence of an interfacial electric field may be made if μ (first order) or both μ and α (second order) are known. Adding the linear free energy relationship and energy corrections for the interaction of the interface with an electric field, eqn (3.12) is expanded:

$$\Delta G_{\text{ads}}(U) = \Delta G_{\text{ads}}(U = 0) - eU + (\mu_{\text{A}^*} - \mu^*)F - \frac{1}{2}(\alpha_{\text{A}^*} - \alpha^*)F^2 + \cdots \tag{3.18}$$

Presuming the Helmholtz model is valid for the interfacial double-layer region, the interfacial electric field is given by:

$$F = \frac{U - U_{\text{PZC}}}{d} \tag{3.19}$$

where d is the double-layer thickness and U_{PZC} is the electrode potential at zero charge. The sign convention used here is that a positive field would represent a positively charged metal surface and a net negatively charged electrolyte in the

interfacial region. The potential of zero charge for an electrode will depend on the electrode composition, structure, and the identity of the electrolyte. This value is often not known for the system of interest, and is often approximated as being 0 V_{SHE}. This is a reasonable approximation for Pt(111) in acidic electrolyte. The value of d also must be estimated although, as noted in Section 3.1.3, the Helmholtz model itself is approximate. The value of 3 Å is a reasonable approximation of separation of one solvation shell in water (given a small ion at relatively high concentration), though this value should vary with solvent, ion identity, and ion concentration.

For a periodic surface model, the model is non-periodic in the surface direction. Most DFT codes will allow for determination of the dipole moment along non-periodic directions. For example, the dipole moment is provided in VASP with the keyword IDIPOL = 3, where "3" signifies that the third basis vector is along the surface normal direction. Calculating the surface normal dipole moment for both the bare surface slab (slightly non-zero if the bottom layers are frozen as in a typical surface slab model) and the adsorbed sulfate species then allows for a first order correction to the adsorption energy. The calculated dipole moments are 0.46 and 0.017 e^- Å for adsorbed sulfate and the bare platinum surface, respectively. The positive dipole moment for adsorbed sulfate indicates that the sulfate is partially negatively charged upon adsorption, leaving the surface partially positively charged. With the double-layer thickness approximated at 3 Å and the assumption that U_{PZC} is 0, the sulfate adsorption free energy from eqn (3.15) can be corrected for the dipole–field interaction:

$$\begin{aligned}\Delta G_{ads}(U_{SHE}) &= 1.84 - 2U_{SHE} + \left(\mu_{SO_4^*} - \mu^*\right)\frac{U - U_{PZC}}{d} \\ &= 1.84 - 2U_{SHE} + (0.46 - 0.017)\frac{U_{SHE} - 0}{3} \\ &= 1.84 - 1.85U_{SHE}\end{aligned} \tag{3.20}$$

Adding the dipole moment correction has lessened the impact of a more positive potential in promoting the anion adsorption. This occurs because a more positive potential creates an electric field that destabilizes the adsorbed anion, as adsorption leaves a surface dipole with its negative end pointing towards the electrolyte and positive end towards the surface. The slope of a reaction free energy with electrode potential is equal to the number of electrons transferred upon reaction. With the dipole correction we can conclude that approximately 1.85 electrons are transferred upon sulfate adsorption, with sulfate retaining a partially negative charge. Another interpretation is that anion adsorption generates a Faradaic current of 2 electrons (an oxidation current), but a non-Faradaic current of 0.15 electrons (a reduction current) is necessary to maintain a constant electrode potential owing to the formation of the surface dipole. Though this analysis provided further information on the electrochemical nature of sulfate adsorption, it remains approximate in that we

presumed that the electronic structure (the dipole moment) of the metal–adsorbate interface would not adjust in response to a varying interfacial electric field. The next section removes this approximation.

3.2.1.3 Model 2a.3: Vacuum–Metal Interface with Applied Electric Field

Eqn (3.18) suggests a second order correction by including the polarizability of the interface, however, this value is not typically available. Instead, a series of calculations can be conducted with an explicit and varied external electric field applied (keyword EFIELD in VASP). For each applied field, the structure is re-optimized, allowing the interface to restructure in response to the field. The total energy reported in the applied field calculation then includes the energy of interaction of the electrode–adsorbate system with the applied field. For ion adsorption, both the bare surface and adsorbate–surface system need to be examined at varying applied fields. We can rewrite the adsorption energy from eqn (3.12) as

$$\Delta G_{\mathrm{ads}}(U_{\mathrm{SHE}}) = G_{\mathrm{A}^*}(U) - G_{\mathrm{A}^{n-}_{\mathrm{aq}}} - G^*(U) - n(U_{\mathrm{SHE}} + 4.6) \tag{3.21}$$

where the adsorbate and bare surface energies are expressed as depending on potential presuming eqn (3.19) will be used to relate the applied field to the electrode potential.

Table 3.3 lists the DFT calculated energies at varied applied electric field for the bare surface and the surface with sulfate adsorbed. The sulfate adsorption energy at $F = 0.25\,\mathrm{V\,\AA^{-1}}$ (or $U_{\mathrm{SHE}} = 0.75\,\mathrm{V}$, with $d = 3\,\mathrm{\AA}$ and $U_{\mathrm{PZC}} = 0$ assumed) is

$$\begin{aligned}\Delta G_{\mathrm{ads}}(0.75\mathrm{V_{SHE}}) &= (-291.30) - (-41.14) - (-261.30) \\ &\quad -2(0.75 + 4.6) = 0.44\ \mathrm{eV}\end{aligned} \tag{3.22}$$

Without a local electric field considered, the adsorption energy is 0.34 eV at a potential of 0.75 $\mathrm{V_{SHE}}$ [eqn (3.15)]. The applied electric field causes a 0.10 eV energy difference at 0.75 $\mathrm{V_{SHE}}$, which is very close to the estimation (+0.11 eV) using a dipole–field interaction only. Clearly, the first order dipole correction was sufficient, with higher order corrections slightly lessening the field effect by

Table 3.3 The DFT calculated free energies in (eV) relevant to sulfate adsorption at various applied electric fields. The *ZPVE* and *TS* corrections were presumed unaffected by electric field.

Field [V Å^{-1}]	U*[V]*	G*	G_{SO_4*}	G_{9H_2O*}	$G_{SO_4^*} + 8H_2O*$
–0.5	–1.50	–261.69	–292.07	–393.69	–410.45
–0.25	–0.75	–261.31	–291.54	–392.39	–410.04
0.00	0.00	–261.18	–291.28	–392.44	–409.94
0.25	0.75	–261.30	–291.30	–392.92	–410.18
0.50	1.50	–261.68	–291.61	–392.74	–410.76

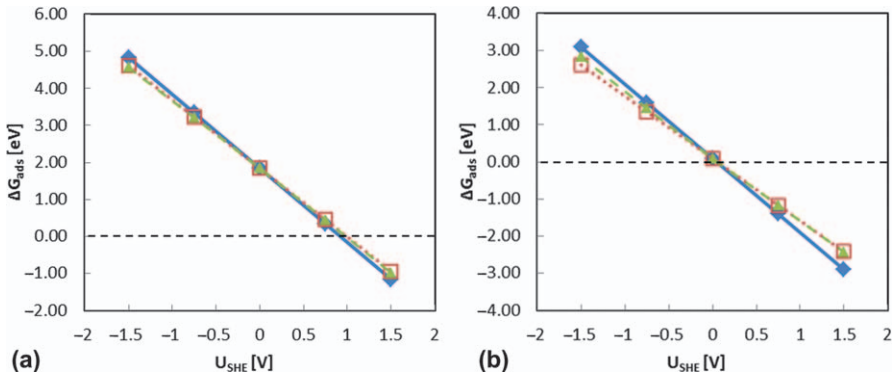


Figure 3.8 The adsorption energy of sulfate anion over Pt(111) at (a) vacuum and (b) partially solvated interfaces. The diamond (◆) with solid line represents results from the linear free energy model (Model 2a.1), the square with dotted line represents (□) the linear free energy model with dipole correction (Model 2a.2), and the triangle (▲) with dashed line represents the electric field model (Model 2a.3).

allowing the interface to polarize in response to the field. The results from Models 2a.1–3 are summarized in Figure 3.8(a). The difference in adsorption free energy among models increases as the electric field increases. The influence of polarizability and other higher order terms is therefore expected to be more important for a larger applied electric field.

3.2.2 Simulating an Aqueous–Metal Interface

The vacuum–metal interface model incorporates the influence of electrode potential on electrochemical reaction energies, but neglects the direct interaction between the electrode or adsorbate and the electrolyte (solvent). The lack of inclusion of solvation interactions limits the accuracy of reaction energetic calculations at an aqueous–metal interface. For the sulfate adsorption calculations presented above, the stabilization of the anion by solvent in solution was included [*via* eqn (3.11)] however, the sulfate was presumed to completely shed any solvent interaction at the surface. When considering reaction energies between two surface-adsorbed species, any substantial differences in interaction with solvent between the two species would lead to errors in approximating reaction energies. As discussed in Section 3.1, DFT methods cannot ''perfectly'' model the time-averaged interaction between adsorbates and solvent. An additional complexity is introduced in that a varying electrode potential (surface charge) can alter the structure and dynamics of the electrolyte. A relatively simple approach to including solvation effects is to add explicit water molecules, treated with DFT, into the model system. We will first introduce a ''partial'' or ''micro-'' solvated model, which builds on the Model 2a approaches, with the inclusion of a solvating water layer. A more sophisticated ''fully solvated'' approach, Model 2b.1, is subsequently introduced.

3.2.2.1 Model 2a with a Partially Solvated Interface

A partial, or micro-, solvated model is generated by adding some number of water molecules to the electrode–adsorbate interface explicitly to solvate the adsorbate. We discuss here the use of a model comprised of one or two water bilayers above the metal surface [6–12 water molecules on a 3×3 (111) surface of an fcc metal]. A water bilayer model with nine water molecules on a 3×3 Pt(111) surface is shown in Figure 3.6(b). Because the "bare" electrode surface is covered with water molecules at an aqueous–metal interface, the adsorption of an anion is taken to occur with an adsorbed water molecule being replaced:

$$A^{n-}_{(aq)} + 9H_2O^* \leftrightarrow (A^* + 8H_2O^*) + H_2O_{(aq)} + ne^- \tag{3.23}$$

The optimized configuration of $(A^* + 8H_2O^*)$ is shown in Figure 3.7(c). Applying the linear free energy relationship [eqn (3.12)] and having the energetic data at zero field (given in Tables 3.2 and 3.3), the potential dependent adsorption free energy can be computed *via*

$$\begin{aligned}\Delta G_{ads}(U_{SHE}) &= G_{8H_2O^*+A^*} + G_{H_2O_{aq}} - G_{A^{n-}_{aq}} - G_{9H_2O^*} - n(U_{SHE} + 4.6) \\ &= (-409.94) + (-14.35) - (-392.44) - (-41.14) - 2(U_{SHE} + 4.6) \\ &= 0.09 - 2U_{SHE}\end{aligned} \tag{3.24}$$

The adsorption energy at 0 V_{SHE} under a partial solvated environment is 0.09 eV. The equilibrium potential shifts from 0.92 (without solvation, Model 2a.1) to 0.05. The value shift of 0.8 eV is due to the energy of strong hydrogen bonds between water and the adsorbed sulfate species, as shown in Figure 3.7(c). Clearly, neglecting solvation of the adsorbed species would severely alter the calculated equilibrium potential for sulfate adsorption.

The dipole moment (Model 2a.2) or applied electric field (Model 2a.3) corrections made to the adsorption energy for the adsorbate–electrode model may also now be made including the water bilayer. The dipole moment calculated for the $A^* + 8H_2O^*$ structure is −0.26 e Å whereas that for the 9 H_2O^* structure is −1.22 e Å. The dipole moment difference of 0.96 e Å leads to a slope of adsorption free energy versus potential of 1.68 (recall 1.85 at the vacuum interface), suggesting that fewer electrons transfer upon adsorption of sulfate in the solvated environment compared to the vacuum interface.

The use of a single local minima water structure, both with the initial water bilayer and the single structure adopted about the adsorbed sulfate, makes it difficult to determine the exact energetic contribution from solvation. The exposure of the water layer to vacuum above may also allow for unrealistic structural relaxations in the water layer. We may conclude that the use of a microsolvated water layer was useful in noting a large (0.8 eV) shift in adsorption potential relative to neglecting the solvation of sulfate, but that it also introduces an imprecision which may grow at potentials far from that of zero charge. The use of micro-solvation models introduces numerous structural degrees of

freedom for the adsorbate–solvent system, and one must be careful to avoid the temptation to tune solvated water structures to provide "desired" values.

The explicit inclusion of an applied electric field at the metal–aqueous interface is considered in the same manner as described in the vacuum condition. System energies as a function of applied field are given for the water bilayer and micro-solvated sulfate in Table 3.3. The potential dependent adsorption energy becomes:

$$\Delta G_{ads}(U_{SHE}) = G_{8H_2O^*+A^*}(U_{SHE}) + G_{H_2O_{aq}} - G_{A^{2-}_{aq}} - G_{9H_2O^*}(U_{SHE}) - 2(U_{SHE} + 4.6) \quad (3.25)$$

It is again presumed that the potential and electric field are interrelated through eqn (3.19). Figure 3.8(b) illustrates the electrode dependent adsorption free energy calculated with the solvated Models 2a.1, 2a.2, and 2a.3. At a potential of 0.75 V_{SHE}, the sulfate adsorption energies are −1.41, −1.17, and −1.17 eV for the linear free energy (Model 2a.1), dipole correction (Model 2a.2), and electric field (Model 2a.3) approaches, respectively. The influence of the electric field on the reaction energy is more significant in the presence of water than for the vacuum condition. Interpretation of the difference between the dipole corrected and applied field models is complicated by significant structural changes in the water layer due to the application of an electric field.

The results represented in Figure 3.8 show that the inclusion of solvation largely facilitates ion adsorption and leads to a substantial shift of adsorption equilibrium potential. Although the use of a bilayer water model represents a more realistic interfacial region than a vacuum model, interpretation of results is complicated by the coupled interactions between the electrode, adsorbate, and solvent. Accuracy of the results relative to the actual dynamic solvated interface is difficult to estimate. The micro-solvated model as presented does not constrain the water density, allowing for possibly non-physical water structures. Density may be constrained by introducing a fixed water layer, but this then requires knowledge of the appropriate water density. The field-induced flip of water molecules from an H-down to the H-up configuration might also complicate the disentanglement of the potential/field effects from solvation effects. Moreover, adsorption or reaction energies are calculated under constant electric field rather than constant potential. Although the electric field and electrode potential in the double layer region can be related using eqn (3.19), the estimation of double layer thickness or potential of zero charge is imprecise. In the next section, we present a "fully solvated" model, which includes water density control, and we also add the ability to control electrode potential directly by electrifying the electrode–electrolyte interface.

3.2.2.2 Model 2b.1: Double Reference Method

The double reference method, developed by the Neurock group,[20,21] allows for explicit charging of the electrode–electrolyte interface and the inclusion of an explicit solvent layer. The vacuum region between two metal slabs is completely

filled with solvent, typically water molecules. The interlayer distance between water layers may be varied to obtain a desired water density, and the energy *vs.* distance can be measured to get the optimal density. This model thereby includes both the local solvating effects offered with a micro-solvated layer and water density control. An ice-like water structure, shown in Figure 3.6(c), is typically used, and optimization can only assure that a local optimum structure is found.

To include an adsorbed species, one or more surface water molecules are removed and the adsorbate added. Figure 3.7(c) illustrates the adsorption structure of sulfate at a fully solvated electrode surface with one water molecule displaced. The choice of cell size, adsorbate site (which also sets position within the water layer), water density, and water structure can all affect the computed results. As these choices can only be loosely based on matching an actual system, and testing of the impact of all of them is often intractable, choices should be made consistently between reactants and products to take advantage of cancellation of variability. For the anion adsorption reaction [eqn (3.7)], adsorption occurs with displacement of some number of water molecules:

$$A^{n-}_{(aq)} + 24H_2O^* \leftrightarrow (A^* + (24 - x)H_2O^*) + xH_2O_{(aq)} + ne^- \qquad (3.26)$$

The potential dependent adsorption energy is given by

$$\Delta G_{ads}(U) = G_{(24-x)H_2O^*+A^*}(U) + xG_{H_2O_{aq}} - G_{A^{n-}_{aq}} - G_{24H_2O^*}(U) - nU \quad (3.27)$$

where we have again presumed an energy unit of eV by leaving out the elementary charge in the last term.

The fully solvated model cannot easily be used with the dipole-correction or applied field methods discussed above. The absence of a vacuum layer in the fully solvated model makes dipole moment evaluation and field application impossible within the periodic model. Hybrid approaches, in which the dipole is measured in an unsolvated model and solvation approximated in the fully solvated model, might be considered. However, the advantage of the fully solvated model is its integration within the direct charging of the electrochemical interface.

To electrify the interface explicitly in a fully solvated model, an excess charge is added into the system. The excess charge allows for direct control of the electrode potential by shifting the Fermi level of the metal electrode. The excess charge segregates to the metal surface and, in the periodic model, countercharge is added to maintain cell neutrality. A homogeneous background charge, distributed evenly through the cell, is added. The structure of the metal/adsorbate/solvent can then be optimized for a variety of system charges. The challenge remains to determine the electrode potential established for each applied charge and to arrive at an intensive energy, independent of the changing number of electrons and removing spurious interactions between the system and the background charge.

The electrode potential established for a given system charge is determined through the use of two reference potentials. Before detailing the first reference

potential, we must establish the relationship between electrode potential and workfunction. For an electrochemical system, the electrode potential (U) is given by the energy of the highest energy electron. This value is analogous to the electrode workfunction (WF). The workfunction is defined as the energy needed to remove an electron from the Fermi level into vacuum. The workfunction therefore provides the energy of the highest electron on an "absolute or "vacuum" scale. To adjust the workfunction to provide the electrode potential on the SHE scale, the electrode potential is calculated *via*

$$U_{\mathrm{SHE}} = \frac{WF}{e} - 4.6 \tag{3.28}$$

where the value of 4.6 was discussed above in context of eqn (3.14).

The first reference in the double reference method is a vacuum reference, in which the electrode workfunction is determined. The determination of a metal workfunction using a surface slab model is well established, however, it is complicated in this case by the absence of a vacuum region. A vacuum layer is added to the unit cell by inserting $\sim$20 Å of vacuum in the middle of the water layer, as illustrated in Figure 3.9(a). This is done using the optimized structure of the solvated system with neutral charge. A single point DFT calculation is then performed on this vacuum-inserted cell. The electrostatic potential at each point in the cell is evaluated, and the xy (surface plane) averaged value is determined as a function of z (surface normal) direction [plotted in Figure 3.9(a)]. The potential at the center of the vacuum region is defined as the vacuum potential [$\emptyset_{\mathrm{vac}} = 0$, see Figure 3.9(a)]. The potential profile at this point is not flat owing to the asymmetry of the water layer however, Taylor *et al.* showed that the use of this value as a vacuum potential is valid if the two solvent layers have equivalent capacitances.[21] This does introduce an error, quoted by Taylor *et al.* as less than $\pm$ 0.25 eV. The entire potential (z) profile is then shifted by a constant such that the value of the potential at the middle of the vacuum layer is set to 0.

The vacuum layer calculation is then used to establish a reference in the uncharged cell such that the workfunction of the "closed" neutral cell can be determined. We refer to the closed neutral cell calculation as "*q0*." To do this, the potential at the middle of the metal layer is determined in the open, vacuum cell, relative to vacuum [$\emptyset_{\mathrm{m,vac}}$ in Figure 3.9(a)]. The potential profile in the closed *q0* cell is then plotted and shifted by a constant [$\Delta\emptyset_{\mathrm{shift}}$ in Figure 3.9(b)] such that the potential at the middle of the metal layer of *q0* [$\emptyset_{m,q0}$, Figure 3.9(b)] is the same as that in the vacuum referenced cell ($\emptyset_{\mathrm{m,vac}}$). The workfunction of the solvated cell can then be calculated by applying the same electrostatic potential "correction" to the DFT determined metal Fermi level ($\emptyset_{\mathrm{DFT_{Fermi}}}$ with reference unspecified):

$$WF = -\left(\emptyset_{\mathrm{DFT_{Fermi}}} - \Delta\emptyset_{\mathrm{shift}}\right)/e \tag{3.29}$$

For the example of $24H_2O$ over a 3×3 Pt(111) surface, the DFT determined Fermi level is 2.55 eV, and the metal potential is shifted from -12.58 to -20.09 ($\Delta\emptyset_{\mathrm{shift}} = 7.51$ V), giving a workfunction of 4.96 eV. The electrode potential of

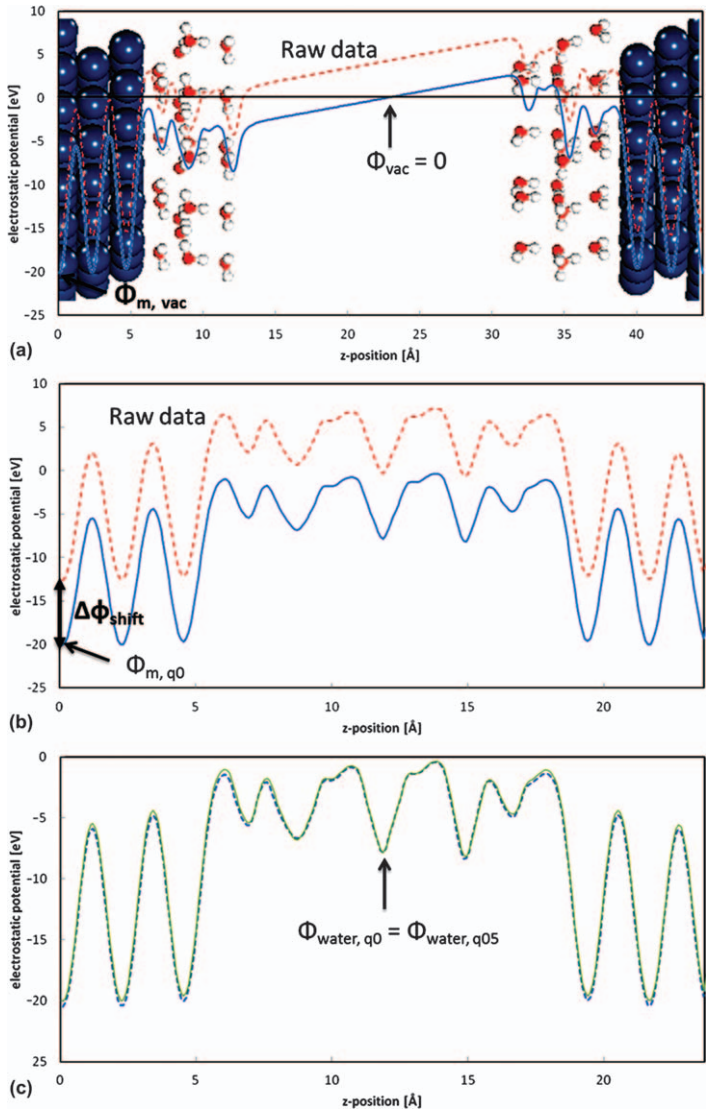


Figure 3.9 An illustration of double reference method: plots of the electrostatic potential averaged across the *xy*-plane as a function of surface normal direction. (a) A vacuum reference cell used to determine the absolute vacuum potential, (b) a neutral (*q0*) cell. In (a) and (b), solid blue lines represent the raw data and red dashed lines represent the data corrected to reference the potential to the vacuum potential. c) The electrostatic potential of neutral (*q0*, solid line) and $+0.5\ |e^-|$ (*q05*, dashed line) cells, both referenced to the vacuum potential.

q0 $24H_2O^*$ on a SHE scale is then $4.95 - 4.6 = 0.35\ V_{SHE}$. This potential represents the static potential of zero charge, though this does not necessarily match an experimental potential of zero charge because the calculated value considered a single static water structure.

For charged systems ($q \neq 0$), a similar vacuum referencing procedure cannot be employed because a neutralizing homogeneous charge would be spread throughout the vacuum region. A second reference is introduced in the double reference method to determine the workfunction (electrode potential) of charged systems. This second reference is set within the water layer with the assumption that the electrostatic potential in the solution far away from the electrode surface remains constant between the charged and uncharged systems. This assumption is reasonable because water molecules and countercharge in the solution region rapidly screen the potential shift generated by a charged electrode. We use the potential at the middle of the water layer as the reference point. The potential is measured relative to vacuum in the neutral *q0* calculation ($\emptyset_{water,q=0}$), and the potential is shifted in the charged calculation ($\emptyset_{water,q}$) such that it equals the uncharged system, as illustrated in Figure 3.9(c). The potential shift applied is again referred to as $\Delta\emptyset_{shift}(q)$ but is now dependent on the system charge q. As the double-referencing procedure effectively shifts the electrostatic potential back to a vacuum reference, the electrode potential of a charged system relative to a standard hydrogen electrode [$U_{SHE}(q)$] is given by

$$U_{SHE}(q) = -\left[\emptyset_{DFT_{Fermi}}(q) - \Delta\emptyset_{shift}\right]/e - 4.6 \tag{3.30}$$

The difference in water potential of *q05* and *q0* systems is -7.78. For example, VASP reports a Fermi energy of -2.43 eV for the *q05* system. A shift of -7.78 V is found necessary to set the potential at the middle of the water layer for *q05* equal to the value at *q0*. The electrode potential of the *q05* system is then found to be $(2.43 - 7.78) - 4.6 = 0.75\ V_{SHE}$. This indicates that, with 24 H_2O filling the vacuum region between slabs, a removal of 0.5 e^- per 3×3 surface cell shifts the potential from the neutral value of 0.35 V_{SHE} to 0.75 V_{SHE}. This small difference in potential between *q0* and *q05* is present in the metal region of Figure 3.9(c), though it is difficult to see owing to the large range of electrostatic potentials in the cell. This procedure may then be replicated for varying system charges, and the electrode potential is established for each charge.

With the electrode potential determined for each charge, the remaining challenge is to relate the DFT calculated energies among various charge systems. The DFT energy must be corrected for the varying number of electrons to provide an intensive system energy, and must be further corrected for system-background charge interactions. Please see Taylor *et al.*[21] for a detailed derivation of these energy corrections. The first correction $\left(E_q(q)\right)$ is the energy necessary to add or remove electrons. This value is simply given by the number of electrons (added or removed) multiplied by the electrode potential ($\emptyset_{Fermi}$ referenced to vacuum)

$$E_q(q) = -q\emptyset_{Fermi} \tag{3.31}$$

The electrostatic interaction of the ions and electrons with the homogeneous background charge must be removed as well. The correction for the background charge ($E_{\text{background}}$) is defined as an integral from 0 to the system charge of the average electrostatic potential in the unit cell. As the VASP reference potential is average electrostatic potential in the cell, this value is equal to the potential shift that was applied within the double-reference method to refer all potentials to vacuum. Therefore, the integrand is replaced with the applied shifts as

$$E_{\text{background}}(Q) = \int_0^Q \Delta\emptyset_{\text{shift}}(q)dq \tag{3.32}$$

When the correction terms are added, total energy of any charged system can be determined *via* the following equation:

$$E(q) = E_{\text{DFT}} + \int_0^Q \Delta\emptyset_{\text{shift}}(q)dq - q\emptyset_{\text{Fermi}} \tag{3.33}$$

The total energy of *q05* of $24H_2O$* *via* eqn (3.33) is:

$$E(q = 0.5) = (-618.30) + (3.82) + (0.5*(-5.35)) = -617.16 \text{ eV} \tag{3.34}$$

where 3.82 represents the numerical integral between the potential shift of 7.51 V for *q0* and 7.78 V for *q05*.

With the double-referencing procedure providing the electrode potential as a function of the system charge and the energy corrections providing a system energy at each charge, the system energy is now established as a function of electrode potential. Calculated energies as a function of electrode potential using double reference method are plotted for the 24 H_2O* system in Figure 3.10(a). Owing to the variation in adsorbate–metal interaction as the system charge is varied, the same net charge does not give the same electrode potential among different adsorbates. To calculate the adsorption or reaction energy at any specific potential, we need a continuous function to describe the free energy–potential relationship. Similar to the energy–electric field relationship, the expansion of the energy about the potential of zero charge gives a quadratic form which arises from the free energy stored in the capacitor set up between the surface charge and the background charge.

$$E(U) = \frac{1}{2}C_{\text{cap}}\left(U - U_{\text{pzc}}\right)^2 + E_{\text{pzc}} = AU^2 + BU + C \tag{3.35}$$

C_{cap} represents the capacitance, E_{pzc} is the zero charge system energy, and A,B, and C are fit constants. The fit constants are typically obtained using five different system charges ($q = -1,\ -0.5,\ 0,\ 0.5,\ 1\ |e|$ per 3×3 unit cell). Table 3.4 lists the fit constants for various systems. Differences in the fit constant indicate differences in metal–absorbate interactions as the potential is varied. We can obtain the potential dependent sulfate adsorption energy by inserting eqn (3.35)

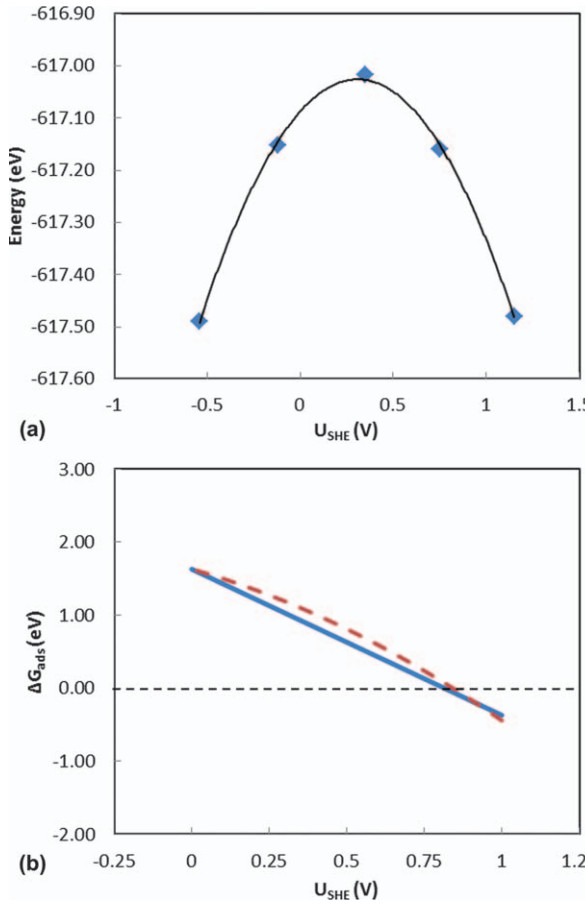


Figure 3.10 (a) The energy calculated with the double reference method using eqn (3.33). Points are for system charges of −1, −0.5, 0, 0.5, and 1 e^- in a Pt(111) 3×3 unit cell of $24H_2O^*$. Solid lines indicate the quadratic fit curves. (b) Calculated reaction energies for sulfate adsorption using the linear free energy method (thick solid line) and the double reference method (dashed line).

(for both $24H_2O^*$ and $SO_4^* + 23H_2O^*$) into eqn (3.27). All potentials in the equations are on the SHE scale.

$$\Delta G_{\mathrm{ads}}(U) = \left(-0.151U^2 + 0.191U - 632.883 + 0.463 - 0.105\right) + G_{\mathrm{H_2O_{aq}}} - G_{\mathrm{A^{n-}_{aq}}} - \left(-0.644U^2 + 0.398U + 617.080 + 0.690 - 0.036\right) - 2U \tag{3.36}$$

At $U = 0.75$ V_{SHE}, the adsorption energy is 0.24 eV (recall 0.44 and −1.17 eV for vacuum and partially solvated models under 0.25 V $Å^{-1}$constant electric field). The adsorption free energy plot shown in Figure 3.9(b) is linear. A small

Table 3.4 Constants of double reference method quadratic fit [eqn (3.35)], and relevant energetic data for free energy calculations

	A	*B*	*C*	*ZPVE*	*TS*
$24H_2O^*$	–0.644	0.398	–617.080	0.690	0.036
$SO_4^* + 23H_2O^*$	–0.151	0.191	–632.883	0.463	0.105
$O_2^* + 23H_2O^*$	–0.604	0.737	–613.233	0.130	0.033
$OOH^* + 23H_2O^*$	–0.584	0.729	–617.594	0.430	0.067
$OH^* + 23H_2O^*$	–0.533	0.614	–612.868	0.343	0.033

slope deviation from two (the linear free energy model) occurs owing to the potential dependence of the interaction between solvent and adsorbate with the electrified interface, as seen in the applied electric field or dipole correction models (shown in Figure 3.8).

In summary, DFT-based electrochemical models with different complexity and computational costs are presented. For the ion adsorption example considered here, the inclusion of solvation stabilizes the adsorbed anion owing to the formation of hydrogen bonds with the adsorbed species. By comparison with solvation, inclusion of dipole moment–electric field interactions, polarization of the interface through an applied field, or a fully electrified interface through the double-reference method result in minor differences. Results of these different models are further compared in the voltammogram simulation in the following section.

3.2.3 Linear Sweep Voltammetry Simulations

Voltammetry is an important tool for evaluating electrochemical and electrocatalytic processes. In a voltammetric experiment, the potential of a working electrode is varied with time relative to a reference electrode. The current of the working electrode is measured and reported as a function of potential. If the potential is swept linearly with time, peaks or waves are observed, which can be attributed to the various electrochemical processes possible in the system. For comparison with experiment, DFT calculated energetics can be used to predict voltammetry results in much the same way microkinetic models are used to predict catalytic kinetics. In the sections above, we have discussed DFT methods to calculate elementary reaction or adsorption free energies as a function of electrode potential. These free energy differences can be used to calculate potential dependent equilibrium constants. Section 3.2.5 will present a method to calculate potential dependent activation barriers. With these values for all possible elementary reaction steps, we could use microkinetic modeling to simulate voltammetry and compare with experiment.

In this section, we detail the approach to modeling the voltammogram associated with specific anion adsorption. Though voltammetry is inherently a kinetic experiment, if we assume that anion adsorption–desorption equilibrium is maintained as the potential is varied, only the adsorption equilibrium

constant is necessary to predict the voltammogram. This assumption would be valid for extremely slow variations of electrode potential with time (low sweep rates). The DFT-calculated ion adsorption free energy is then used to simulate a linear sweep voltammogram (LSV). The equilibrium constant (K) for the anion adsorption eqn (3.7) is given as

$$K(U,\theta) = \exp\left(\frac{-\Delta G_{\text{ads}}(U,\theta)}{RT}\right) = \frac{\theta_{\text{A}*}}{\theta*[\text{A}^{n-}]} \tag{3.37}$$

where $\theta_{\text{A}*}$ is the surface coverage of adsorbed A^{n-}, θ_* is the fraction of empty sites, and $[\text{A}^{n-}]$ is the concentration (or more accurately, the activity) of A^-. In linear sweep voltammetry, the time dependent potential is $U = U_0 + vt$, where v denotes the sweep rate and U_0 represents the starting potential. Inserting the time-dependent potential relationship into eqn (3.37) and adding a site balance, we can solve for the surface coverage of adsorbed A^- ($\theta_{\text{A}*}$) as a function of time or potential. The current density ($j_{\text{A}*}$) due to ion adsorption is determined by

$$j_{\text{A}*} = \left(\frac{\text{d}\theta_{\text{A}*}}{\text{d}t}\right)\left(\frac{m \cdot e \cdot n}{\text{area}}\right) \tag{3.38}$$

In our 3×3 Pt(111) cell, the surface area is 5.98×10^{-15} cm^2 and the maximum number of adsorbed sites per unit cell $m = 3$. Figure 3.11 illustrates the simulated LSVs of sulfuric acid anion adsorption at the Pt(111) surface using various electrochemical models.[34] Because both sulfate and bisulfate anions are present in the sulfuric acid solution, we need simultaneously to solve the equilibrium coverage [eqn (3.37)] for sulfate and bisulfate adsorbtates with the restriction of surface site conservation. As expected, based on the calculated adsorption energies, the partially solvated models predict a sulfate adsorption peak to appear at more negative potentials than the vacuum model. Only a single sharp peak associated with sulfate adsorption is captured using vacuum and partially solvated models. The double-reference method predicts two current peaks, due to bisulfate and sulfate respectively, similar to the observed experimental voltammogram.

3.2.4 Calculation of Surface Reaction Free Energies

The previous sections discussed the calculation of ion adsorption free energies. The methods could be similarly applied to consider surface reaction, however, we have not yet explicitly discussed how to handle surface reactions involving ion transfer. In this section, we present an example involving a surface reaction that occurs with transfer of a proton and electron. The method presented here was initially outlined by Nørskov *et al.*[14]

For a proton–electron transfer reaction, a computational standard hydrogen electrode potential can be used directly. To do this, we take advantage of the equivalence of the chemical potential of hydrogen gas and the proton–electron

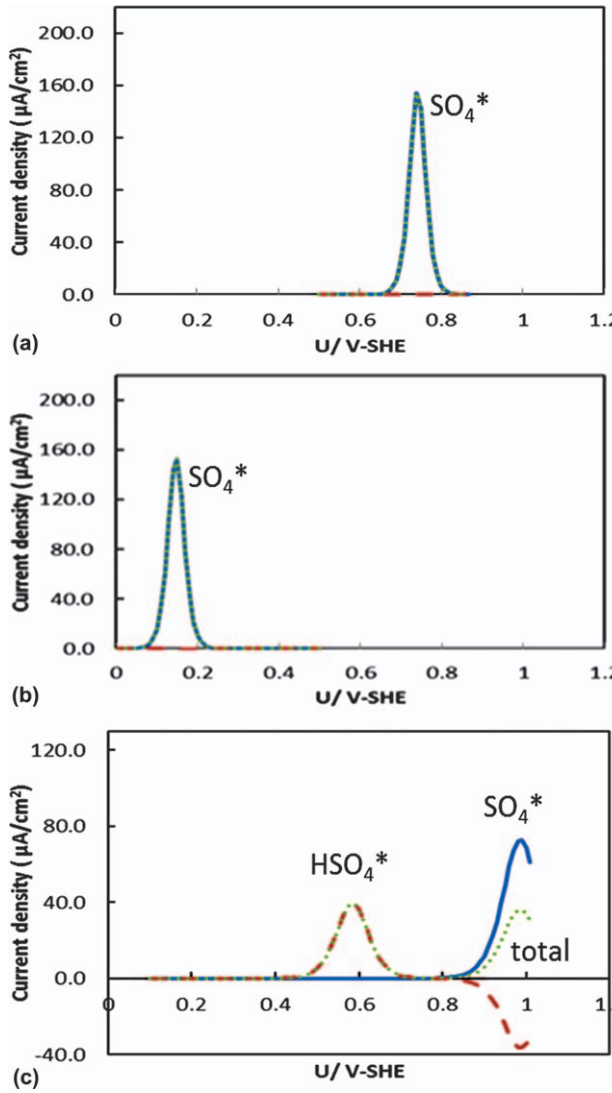


Figure 3.11 Simulated linear sweep voltammograms of sulfate and bisulfate adsorption over the Pt(111) surface using various electrochemical models including the vacuum model (a), partially solvated model (b), and double reference method (c). The currents due to SO_4, HSO_4, and the total current are represented in solid, dashed, and dotted lines. (Reproduced from ref. 34)

pair at the standard hydrogen electrode potential. The overall hydrogen electrode reaction is the sum of reactions 1 and 2:

$$\left(H^{+}_{(aq)} + e^{-}\right) \rightarrow \frac{1}{2} H_{2(g)} \tag{3.39}$$

At the standard condition (298 K, activity(H^+) = 1, H_2 pressure = 1 bar), the reaction free energy is zero at the standard hydrogen electrode potential (0 V_{SHE}). The corresponding equilibrium relationship is written as

$$G_{H^+_{aq}} + G_{e^-} = \frac{1}{2} G_{H_{2(g)}} \tag{3.40}$$

This half reaction equilibrium together with the electron's linear free energy dependence on electrode potential allows for determination of the proton–electron pair free energy without direct DFT consideration of a solvated proton:

$$G_{(H^+_{(aq)}+e^-)}(U_{SHE}) = \frac{1}{2} G_{H_2(g)} - eU_{SHE} \tag{3.41}$$

To apply eqn (3.41), we need only to calculate the free energy of gas-phase hydrogen. At a pH different from 0, the energy of $G_{(H^+_{(aq)}+e^-)}(U_{SHE})$ needs to be corrected by $2.303RT\ln(pH)$.

The relationship in eqn (3.41) can be used to evaluate redox reactions involving a proton–electron pair. For a generic reduction reaction of surface species A*, the half reaction is represented by

$$A^* + (H^+ + e^-) \rightarrow AH^* \tag{3.42}$$

where A* is the adsorbed reactant and AH* is the adsorbed reduction product. The reaction free energy for the reduction reaction is then

$$\Delta G_{red} = G_{AH^*} - G_{A^*} - G_{H^+} - G_{e^-} \tag{3.43}$$

Substituting eqn (3.41) into eqn (3.43), the potential dependent reaction energy for eqn (3.42) becomes

$$\Delta G_{red}(U_{SHE}) = G_{AH^*} - G_{A^*} - \frac{1}{2} G_{H_2} - eU_{SHE} \tag{3.44}$$

A similar expression can be written for an oxidation reaction. The free energy of the adsorbed species may be calculated at a vacuum, micro-solvated, or fully solvated interface. Corrections to the energy difference between the two adsorbed species may be made using the dipole–field correction (Model 2a.2), and applied electric field (Model 2a.3), or the double-reference method (Model 2b.1).

If A* in eqn (3.42) is replaced with an adsorbed oxygen O_2*, the reaction represents the initial reduction step in the oxygen reduction reaction (ORR). The product of this reaction step is OOH*. This reaction occurs at the cathode of many fuel cells. A* may be replaced with OH* with adsorbed H_2O* as the product. This reaction represents an important step in the ORR as well as the reverse of the water oxidation step involved in surface oxidation. The related DFT energetic data for these two reactions are listed in Tables 3.1 and 3.4.

Applying the vacuum model linear free energy approach (Model 2a.1) *via* eqn (3.44), we have

$$\Delta G_{O_2-OOH}(U_{SHE}) = (-275.39) - (271.8) - (-3.46) - eU_{SHE} = -0.13 + U_{SHE} \tag{3.45}$$

Eqn (3.45) gives the reaction energy of the first reduction at any specific potential with respect to the SHE. The reaction is unfavorable ($\Delta G = 1.1$ eV) at the ORR equilibrium potential of 1.23 V_{SHE}.

This linear free energy, vacuum interface approach is the simplest model for a reaction involving proton–electron transfer. The effect of the electric field can also be estimated with the consideration of the dipole moment change between reactants and products (Model 2a.2) or application of an electric field (Model 2a.3). The dipole moments are 0.10 and −0.04 e Å for O_2* and OOH*, respectively, on a 3×3 Pt(111) surface. Assuming the double layer thickness is ~3 Å, the field at a potential of 1 V is 0.33 V $Å^{-1}$. The correction energy due to the electric field is $-0.33*(-0.04 - 0.1) = +0.05$ eV. This correction term is +0.09 for OH* reduction to H_2O* ($\mu = 0.07$ and −0.21, respectively). The effect of an applied electric field could be included in the same manner as the ion adsorption example.

The influence of solvation on the reaction was evaluated using the fully solvated model. The reaction free energy of O_2* + 23H_2O* → OOH* + 23 H_2O* is equal to

$$\Delta G_{O_2\rightarrow OOH}(U_{SHE}) = -0.65 + U \tag{3.46}$$

Compared to the first reduction in vacuum conditions, solvation makes the reaction energy ~0.5 eV more favorable owing to the strong hydrogen bonding between adsorbed OOH and a neighboring water molecule. The equilibrium potential now shifts from 0.13 to 0.65 V_{SHE}.

Using the double reference method (Model 2b.1, data in Table 3.4), the O_2* reduction reaction free energy becomes

$$\begin{aligned}\Delta G_{O_2\rightarrow OOH}(U_{SHE}) = G_{OOH+23H_2O}(U) - G_{O_2+23H_2O}(U) - \frac{1}{2G_{H_2}} + U = (-0.584U^2 \\ + 0.729U - 617.594 + 0.43 - 0.067) - \left(-0.604U^2\right. \\ \left. + 0.737U - 613.233 + 0.13 - 0.033\right) - (-3.46) + U\end{aligned} \tag{3.47}$$

At the ORR equilibrium potential, the double reference method predicts that the initial oxygen reduction reaction becomes favorable at potentials smaller than 0.60 V_{SHE}.

The methods introduced so far can be used to evaluate elementary electrochemical reaction thermodynamics. Evaluation of activation barriers is necessary for consideration of kinetics, which determines the current associated with

electrocatalytic processes. The DFT location of transition states and activation barrier determination are widely used for non-electrochemical reactions. For an electrochemical reaction, the modeling approach must capture the potential dependence of activation barriers. This is challenging, as the transition state represents the highest energy state along the minimum energy path for transfer of an electron from/to the electrode and ion transfer from the bulk electrolyte. In the following section, we will introduce a simple method to approximate potential dependent reaction barriers.

3.2.5 Potential Dependent Activation Barriers

As discussed in Section 3.1.2 and illustrated in Figure 3.3, activation barriers and therefore rate constants for electrochemical elementary reactions vary with electrode potential. A method is introduced here to extrapolate typical DFT, surface slab evaluated activation barriers as a function of potential. This extrapolation takes advantage of the Butler–Volmer formalism [eqn (3.6)]. We first rewrite the reduction of a surface species A* [eqn (3.42)] into the following two equations.

$$A^* + (H^+ + e^-) \leftrightarrow A^* + H^* \tag{3.48}$$

$$A^* + H^* \leftrightarrow AH^* \tag{3.49}$$

The activation barrier (E_a^0) for the non-electrochemical reaction [eqn (3.49)] can be computed using standard DFT methods. Similar to the use of a hydrogen reference, the equilibrium potential (U_0) of eqn (3.48) is determined using the following equation and the computed activation barrier is assigned to this potential.

$$\Delta G(U_{SHE}) = G_{H^*} - \frac{1}{2}G_{H_2} + U_0 = 0 \tag{3.50}$$

The potential dependent reaction barrier is then given by

$$E_a(U) = E_a^0 - \beta(U - U_0) \tag{3.51}$$

where ($U - U_0$) is the elementary reaction free energy change due to the altered electrode potential, and β denotes a symmetry factor. The value of β is expected to be between 0 and 1, and a value of 0.5 may be assumed as a first approximation. Eqn (3.51) describes the reaction barrier as a function of potential.

As an example of application of this method, we apply it to the oxidation of borohydride species bound to a Au(111) electrode in alkaline conditions.[47] In an alkaline condition, an oxidation reaction of surface species AH* may be written as

$$AH^* + OH^- \leftrightarrow A^* + H_2O_{(aq)} + e^- \tag{3.52}$$

Eqn (3.52) can be approximated as a A–H bond-breaking step [eqn (3.53)] followed by an ion transfer step [eqn (3.54)]:

$$AH^* \leftrightarrow A^* + H^* \tag{3.53}$$

$$H^* + OH^- \leftrightarrow H_2O_{(aq)} + e^- + * \tag{3.54}$$

For a 3×3 Au(111) surface, a 1/9 hydrogen coverage and a pH of 14.26, the equilibrium potential for eqn (3.54) is:

$$\Delta G(U_{SHE}) = G_{H_2O(aq)} + G^* - G_{H^*} - G_{OH^-(aq)} - U_0 - 4.6 = 1.19 - U_0 = 0 \tag{3.55}$$

$$U_0 = -1.19\ V_{SHE}$$

When A = BH_4, the reaction represents the activation barrier for the initial surface oxidation reaction within the overall mechanism of borohydride oxidation, which is the reaction occurring at the anode of a direct borohydride fuel cells. Figure 3.12 illustrates the initial, transition, and final states of BH_4* dissociation. The DFT computed activation barrier for the reaction $BH_4^* \rightarrow H^* + BH_3^*$ is 0.37 eV.[45] The potential dependent activation barrier becomes

$$E_a(U) = 0.37 - \beta(U + 1.19) \tag{3.56}$$

With an assumed value of β, a potential dependent activation barrier is determined. This potential dependent barrier was used within a microkinetic model for the overall borohydride oxidation reaction, which allowed for simulation of a borohydride oxidation LSV.[47] A value of β near to 0.5 was found to provide a simulated voltammogram that matched experiment, though the same β value was used for all reaction steps as an approximation.

This method may be improved upon by including similar advances in the model system as those presented above for reaction energy calculations. The inclusion of explicit water molecules can better approximate both the E_a^0 and U_0 values in eqn (3.51). Dipole corrections and applied electric fields will provide corrections to the value of β. Testing of the sensitivity and convergence of activation barriers calculated with this method, along with the application of this method to a variety of reactions, is a current research area in our group.

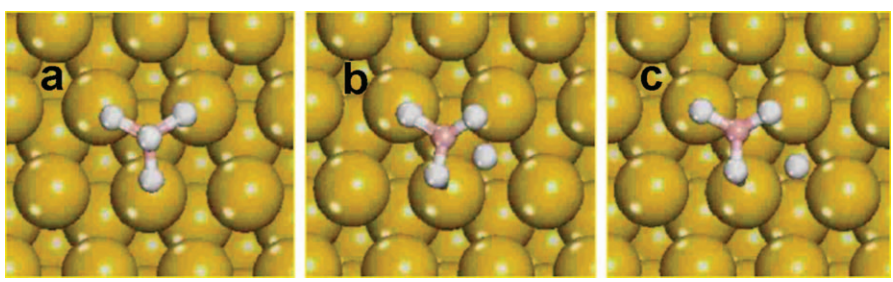


Figure 3.12 Reaction path for BH4* disssociation on the Au(111) surface: (a) initial state BH_4*, (b) transition state, and (c) final state $H^* + BH_3$*.[45]

3.3 Conclusions

This chapter provided a basic introduction to the application of DFT methods to electrocatalysis. Basic electrochemical concepts including the electrochemical double layer and Butler–Volmer kinetics were introduced. An overview of available DFT models for the electrochemical interface was provided, and examples of their application to compute electrochemical reaction energetics were detailed. Through these examples, our hope is that the reader gains an appreciation for the modeling choices made in representing an electrochemical interface with a DFT model. Whereas DFT surface slab models may be taken as a direct representation of an ultra-high vacuum single crystal surface, DFT models of electrocatalytic systems are always an approximation of any real system. The inability to take into account the length and time scales associated with the dynamic electrochemical interface limits model accuracy. Despite this limitation, DFT model systems may be useful both in guiding electrocatalyst design and answering fundamental questions regarding the interaction of reacting species at the electrochemical interface. Advances in the use of coupled DFT–force-field methods may further allow for consideration of the dynamic electrolyte structure at the electrocatalyst surface.

References

1. B. E. Conway and G. Jerkiewicz, Nature of electrosorbed H and its relation to metal dependence of catalysis in cathodic H-2 evolution, *Solid State Ion.*, 2002, **150**(1–2), 93.
2. S. Trasatti, Work function, electronegativity, and electrochemical behaviour of metals: III. Electrolytic hydrogen evolution in acid solutions, *J. Electroanal. Chem.*, 1972, **39**(1), 163.
3. J. Greeley and J. K. Norskov, Large-scale, density functional theory-based screening of alloys for hydrogen evolution, *Surf. Sci.*, 2007, **601**, 1590.
4. J. Greeley, J. K. Norskov, L. A. Kibler, A. M. El-Aziz and D. M. Kolb, Hydrogen evolution over bimetallic systems: understanding the trends, *Chem. Phys. Phys. Chem.*, 2006, **7**, 1032.
5. J. Greeley, T. F. Jaramillo, J. Bonde, I. Chorkendorff and J. K. Norskov, Computational high-throughput screening of electrocatalytic materials for hydrogen evolution, *Nature Mat.*, 2006, **5**, 909.
6. J. X. Wang, T. E. Springer and R. R. Adzic, Dual-pathway kinetic equation for the hydrogen oxidation reaction on Pt electrodes, *J. Electrochem. Soc.*, 2006, **153**(9), A1732.
7. W. C. Sheng, H. A. Gasteiger and Y. Shao-Horn, Hydrogen oxidation and evolution reaction kinetics on platinum: acid vs alkaline electrolytes, *J. Electrochem. Soc.*, 2010, **157**(11), B1529.
8. A. J. Bard and L. R. Faulkner, *Electrochemical Methods: Fundamentals and Applications*, New York, John Wiley & Sons, Inc, 2001.
9. H. F. L. von Helmholtz, *Annal. Phys.*, 1853, **89**, 211.

10. G. Gouy, *Compt. Rend.*, 1910, **149**, 654.
11. D. L. Chapman, *Phil. Mag.*, 1913, **25**, 475.
12. O. Z. Stern, *Elektrochem. Angew. Phys. Chem.*, 1924, **30**, 508.
13. C. G. Guymon, R. L. Rowley, J. N. Harb and D. R. Wheeler, Simulating an electrochemical interface using charge dynamics, *Cond. Matt. Phys.*, 2005, **8**(2), 335.
14. J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard and H. Jonsson, Origin of the overpotential for oxygen reduction at a fuel-cell cathode, *J. Phys. Chem. B*, 2004, **108**(46), 17886.
15. E. Skulason, G. S. Karlberg, J. Rossmeisl, T. Bligaard, J. Greeley, H. Jonsson and J. K. Nørskov, Density functional theory calculations for the hydrogen evolution reaction in an electrochemical double layer on the Pt(111) electrode, *Phys. Chem. Chem. Phys.*, 2007, **9**(25), 3241.
16. J. Rossmeisl, E. Skulason, M. E. Bjorketun, V. Tripkovic and J. K. Norskov, Modeling the electrified solid-liquid interface, *Chem. Phys. Lett.*, 2008, **466**(1–3), 68.
17. M. P. Hyman and J. W. Medlin, Theoretical study of the adsorption and dissociation of oxygen on Pt(111) in the presence of homogeneous electric fields, *J. Phys. Chem. B*, 2005, **109**(13), 6304.
18. A. Panchenko, M. T. M. Koper, T. E. Shubina, S. J. Mitchell and E. Roduner, Ab initio calculations of intermediates of oxygen reduction on low-index platinum surfaces, *J. Electrochem. Soc.*, 2004, **151**(12), A2016.
19. K.-Y. Yeh and M. J. Janik, Density functional theory-based electrochemical models for the oxygen reduction reaction: comparison of modeling approaches for electric field and solvent effects, *J. Comput. Chem.*, 2011, **32**(16), 3399.
20. J. S. Filhol and M. Neurock, Elucidation of the electrochemical activation of water over Pd by first principles, *Ange. Chem. Int. Ed.*, 2006, **45**(3), 402.
21. C. D. Taylor, S. A. Wasileski, J. S. Filhol and M. Neurock, First principles reaction modeling of the electrochemical interface: Consideration and calculation of a tunable surface potential from atomic and electronic structure, *Phys. Rev. B*, 2006, **73**(16), 165402.1.
22. J. Zhao, C. T. Chan and J. G. Che, Effects of an electric field on a water bilayer on Ag(111), *Phys. Rev. B*, 2007, **75**(8), 085435.1.
23. K. P. Bohnen and D. M. Kolb, Charge- versus adsorbate-induced lifting of the Au(100)-(hex) reconstruction in an electrochemical environment, *Surf. Sci.*, 1998, **407**(1–3), L629.
24. C. L. Fu and K. M. Ho, External-charge-induced surface reconstruction on Ag(110), *Phys. Rev. Lett.*, 1989, **63**(15), 1617.
25. S. Schnur and A. Gross, Challenges in the first-principles description of reactions in electrocatalysis, *Catal. Today*, 2011, **165**(1), 129.
26. M. Otani, I. Hamada, O. Sugino, Y. Morikawa, Y. Okamoto and T. Ikeshoji, Structure of the water/platinum interface - a first principles simulation under bias potential, *Phys. Chem. Chem. Phys.*, 2008, **10**(25), 3609.

27. M. Otani and O. Sugino, First-principles calculations of charged surfaces and interfaces: A plane-wave nonrepeated slab approach, *Phys. Rev. B*, 2006, **73**(11), 115407.1.
28. O. Sugino, I. Hamada, M. Otani, Y. Morikawa, T. Ikeshoji and Y. Okamoto, First-principles molecular dynamics simulation of biased electrode/solution interface, *Surf. Sci.*, 2007, **601**(22), 5237.
29. R. Jinnouchi and A. B. Anderson, Electronic structure calculations of liquid-solid interfaces: Combination of density functional theory and modified Poisson-Boltzmann theory, *Phys. Rev. B*, 2008, **77**(24), 245417.1.
30. A. B. Anderson, J. Uddin and R. Jinnouchi, Solvation and zero-point-energy effects on OH(ads) reduction on Pt(111) electrodes, *J. Phys. Chem. C*, 2010, **114**(35), 14946.
31. Y. Sha, T. H. Yu, Y. Liu, B. V. Merinov and W. A. Goddard III, Theoretical Study of solvent effects on the platinum-catalyzed oxygen reduction reaction, *J. Phys. Chem. Lett.*, 2010, **1**(5), 856.
32. Y. Sha, T. H. Yu, B. V. Merinov, P. Shirvanian and W. A. Goddard III, Oxygen hydration mechanism for the oxygen reduction reaction at Pt and Pd fuel cell catalysts, *J. Phys. Chem. Lett.*, 2011, **2**(6), 572.
33. T. H. Yu, Y. Sha, B. V. Merinov and W. A. Goddard III, Improved non-Pt alloys for the oxygen reduction reaction at fuel cell cathodes predicted from quantum mechanics, *J. Phys. Chem. C*, 2010, **114**(26), 11527.
34. K.-Y. Y. Yeh, N. A. Restaino, M. R. Esopi, J. K. Maranas and M. J. Janik, The adsorption of bisulfate and sulfate anions over a Pt(111) electrode: A first principle study of adsorption configurations, vibrational frequencies and linear sweep voltammogram simulations, *Catal. Today*, 2013, **202**, 20.
35. A. B. Anderson, Theory at the electrochemical interface: reversible potentials and potential-dependent activation energies, *Electrochim. Acta*, 2003, **48**(25–26), 3743.
36. R. Jinnouchi and A. B. Anderson, Aqueous and surface redox potentials from self-consistently determined Gibbs energies, *J. Phys. Chem. C*, 2008, **112**(24), 8747.
37. J. Tomasi, B. Mennucci and R. Cammi, Quantum mechanical continuum solvation models, *Chem. Rev.*, 2005, **105**(8), 2999.
38. G. Kresse and J. Furthmuller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, *Comp. Mat. Sci.*, 1996, **6**, 15.
39. G. Kresse and J. Furthmuller, Efficient iterative schemes for *ab initio* total-energy calculations using a plane wave basis set, *Phys. Rev. B*, 1996, **54**(16), 11169.
40. G. Kresse and J. Hafner, *Ab initio* molecular dynamics for liquid metals, *Phys. Rev. B*, 1993, **47**(1), 558.
41. P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B*, 1994, **50**(24), 17953.
42. J. P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, 1996, **77**, 3865.

43. J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation, *Phys. Rev. B*, 1992, **46**(11), 6671.
44. H. Reiss and A. Heller, The absolute potential of the standard hydrogen electrode: a new estimate, *J. Phys. Chem.*, 1985, **89**(20), 4207.
45. G. Rostamikia and M. J. Janik, Borohydride oxidation over Au(111): a first-principles mechanistic study relevant to direct borohydride fuel cells, *J. Electrochem. Soc.*, 2009, **156**, B86.
46. I. S. P. Savizi and M. J. Janik, Acetate and phosphate anion adsorption linear sweep voltammograms simulated using density functional theory, *Electrochim. Act.*, 2011, **56**(11), 3996.
47. G. Rostamikia, A. J. Mendoza, M. A. Hickner and M. J. Janik, First-principles based microkinetic modeling of borohydride oxidation on a Au(111) electrode, *J. Power Sourc.*, 2011, **196**(22), 9228.

CHAPTER 4

Application of Computational Methods to Supported Metal–Oxide Catalysis

THOMAS P. SENFTLE,[a] ADRI C.T. VAN DUIN[b] AND MICHAEL J. JANIK*[c]

[a] Department of Chemical Engineering, Pennsylvania State University, University Park, PA 16802, USA; [b] Department of Mechanical and Nuclear Engineering, Pennsylvania State University, University Park, PA 16802, USA; [c] Department of Chemical Engineering, Pennsylvania State University, University Park, PA 16802, USA
*Email: mjanik@engr.psu.edu

4.1 Introduction

Heterogeneous catalysts featuring metal particles dispersed on an oxide support play an indispensable role in numerous industrial chemical processes. The petrochemical industry relies on supported metal–oxide catalysts for processes that generate industrial chemical feedstocks by reforming the useful byproducts of fossil fuel refinement. Furthermore, many clean energy technologies rely on supported metal–oxide catalysts for the treatment of combustion exhaust and for high-temperature fuel cell applications. Examples of supported metal–oxide catalysis include: catalytic combustion,[1–12] hydrocarbon steam-reforming,[13–24] CO removal from *syngas via* the water–gas-shift (WGS) reaction,[25–34] CO and NO oxidation,[35–41] automotive three-way catalysis,[42–46] solid oxide fuel cell (SOFC) electrodes,[47–53] and selective hydrogenation.[54–59] The activity and

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

selectivity of supported metal–oxide catalysts can be tuned for these applications by altering the metal–support surface morphology.[60]

The computational methods presented in this book can complement experimental efforts in building a molecular level understanding of supported metal–oxide catalysis. Identifying stable surface morphologies, active surface sites, and reaction mechanisms is a difficult task owing to the complexity of competing factors influencing the behavior of the catalyst.[61] To name a few variables, catalyst activity and stability are directly linked to the size and distribution of the metal particles on the oxide surface,[62–64] the oxidation state of the metal,[65–67] charge transfer between the particle and the support,[68–71] the electronic structure of the metal–oxide interface,[70,72–76] the concentration of oxygen-vacancies on the oxide surface,[77–79] the formation of metal–oxide surface phases,[80–83] the incorporation of metal atoms in oxide lattice positions,[4,69,78,84–89] and adsorbate coverage effects.[90–93] Furthermore, these effects are highly dependent on variable reaction conditions: namely, gas phase partial pressures and temperature. A fundamental understanding of how these phenomena are linked to catalytic behavior is essential for making informed design decisions that will yield stable and active catalysts.

This chapter will highlight examples from the literature that apply DFT, *ab initio* thermodynamics, and empirical force-field methods to assess catalytic behavior of supported metal–oxide catalysts. Section 4.2 will provide a brief overview of these computational methods as they pertain to supported metal–oxide systems. Section 4.3.1 considers studies that apply DFT to the design of water–gas-shift catalysts, evaluating the relative energies of metal–oxide surface structures and mapping reaction mechanisms. Section 4.3.2 discusses studies that extend the formalism of DFT to non-zero temperature and pressure *via ab initio* thermodynamics, thus assessing the impact of surface stability on catalytic behavior. Section 4.3.3 focuses on studies that use ReaxFF empirical force-fields in molecular dynamic (MD) and Monte Carlo (MC) simulations to investigate system dynamics at larger time and length scales. Section 4.3.4 highlights initial and ongoing multi-scale studies that integrate the above methods to characterize Pd/ceria catalysts.

4.2 Computational Approaches to Supported Metal–Oxide Catalysis

Advances in computational chemistry over the last few decades have had a great impact on design strategies for obtaining active, selective, and stable catalysts.[94] In particular, density functional theory (DFT)[95–98] has shown tremendous success in modeling catalytic systems, thus enabling *a priori* approaches to designing metal–oxide catalysts. DFT is capable of determining the electronic ground state energy of a system as a function of atomic positions. Thus, one can determine the optimized structure of the catalyst surface during every step of a reaction – allowing determination of stable intermediates and activation barriers on the potential energy surface (PES) of a reaction

coordinate. The computational expense of DFT, however, limits system sizes to around $\sim$100 atoms. For this reason, it is often limited to periodic models of single crystal surfaces, which are used to represent the varying surface facets of large particles. Furthermore, characterizing the interactions between a metal cluster and an oxide support requires highly idealized models, featuring perfectly dispersed metal atoms and clusters, that can only approximate the behavior of actual catalytic systems. Despite this limitation, DFT has proved instrumental for mapping reaction mechanisms and evaluating particle–support interactions at the electronic level.

Reaction conditions, dictated by the temperature and partial pressure of gaseous species, affect the stability and reactivity of the catalyst surface. Understanding this effect is necessary to identify surface phases that are stable and active under *reaction conditions*. As a quantum theory, DFT does not inherently account for the effects of temperature and pressure. DFT can be extended to treat systems at realistic temperatures and pressures through the formalism of *ab initio* thermodynamics.[79,82,98–102] *Ab initio* thermodynamics uses statistical mechanics to incorporate the effect of entropy in systems where solid catalyst surfaces are in equilibrium with a molecular gas phase. This method can calculate the *free energy* of a system, therefore allowing one to determine the relative *stability* of possible surface terminations and structures. With this information, one can construct phase diagrams predicting morphologies that will dominate the catalyst surface as a function of temperature and pressure.

Sufficient models of metal–support structures and dynamics are often beyond the length and time scales that are computationally tractable with DFT. This has provided the impetus to employ classical force-fields, such as the ReaxFF,[103] that are empirically derived from quantum and experimental data. The classical basis for empirical potentials allows for the computational treatment of systems reaching the order of $\sim 10^4$ atoms for nanosecond timescales, which is sufficient to capture system *dynamics* through MD and MC simulations. The application of reactive force-field methods to heterogeneous catalysis is an emerging research area. Herein we highlight the applications of ReaxFF to heterogeneous catalysis over metal and oxide surfaces. The success of these studies has led to ongoing attempts to apply this method to supported metal–oxide systems.

4.3 Selected Applications

The following sections highlight studies employing the computational methods discussed in the previous section to characterize supported metal–oxide systems. It is not intended to be an exhaustive review of the subject, but rather serves to illustrate the strengths and limitations of these methods when applied to supported heterogeneous catalysis. It is broken into three sections corresponding to DFT applied to water–gas-shift (WGS) catalysis (4.3.1), applications of *ab initio* thermodynamics to assess thermodynamic stability of surfaces (4.3.2), and applications of empirical force-fields (4.3.3). The sections

are organized to highlight the individual capabilities of each method, and we have therefore restricted our discussion of each study to the components most directly related to the specific method under consideration. In fact, many of the studies discussed herein apply these methods together, and we ask the reader to keep this in mind as these extensions are not always discussed explicitly. An example of these methods applied in tandem will be discussed in the final section.

The studies discussed in this chapter are chosen to illustrate the capabilities and limitations of computational methods applied to supported metal–oxide catalysis. This chapter is not intended as a comprehensive review of the subject and, as such, numerous notable studies are not discussed in detail. We would nevertheless like to summarize quickly a few especially notable studies before moving to more detailed discussions.

Oxidation of CO is an important industrial process, and has been the subject of many computational studies. Notable contributions have been made by Landman and coworkers, who applied DFT in conjunction with isotopic labeling experiments to determine the oxidation state of Pd clusters supported on MgO during CO oxidation.[104] This allowed them to determine that partially oxidized Pd clusters are active toward CO oxidation, along with the temperature range in which the clusters will remain oxidized. In another study, these authors used quantum (QM) calculations in tandem with infrared spectroscopy to demonstrate that charging effects lead to enhanced CO oxidation activity over Au/MgO catalysts. They show that this charging effect is prominent when Au clusters are adsorbed on oxygen-vacancy defects and is absent in clusters adsorbed on the pristine oxide.[105] Hammer, Molina, and coworkers demonstrate that oxide supports can play an active role in reaction mechanisms in their DFT work analyzing CO oxidation over Au/MgO.[106] They demonstrate the importance of interfacial sites and charge transfer between the oxide support and the metal cluster, which leads to enhanced CO oxidation activity over Au/TiO_2.[107]

The interaction between supported metal clusters and adsorbed hydrogen plays a prominent role in many catalytic processes, such as selective hydrogenation, dehydrogenation, and water–gas-shift. Vayssilov, Rösch, and coworkers used DFT to explore hydrogen saturation on supported Ir, Rh, and Pt clusters.[108,109] They found that the extent of hydrogen uptake is dependent on metal type, and is enhanced in metals that are easily oxidized by the support. Sautet and coworkers conducted multiple studies that investigate structural changes to Pt and Pd clusters adsorbed on γ-alumina caused by hydroxyl groups and H_2.[110,111] Using *ab initio* thermodynamics, they found that in a hydrogen-free atmosphere Pt clusters prefer to adsorb on the oxide in a planar configuration. They subsequently showed that the Pt cluster undergoes reconstruction to a cuboctahedral structure after exposure to hydrogen, induced by the formation of a metal hydride. This restructuring occurs above a critical hydrogen pressure, and greatly affects the catalytic properties of the system. These studies underscore the need to consider interactions between metal clusters and oxide surfaces when assessing catalytic behavior. This type of

analysis is exemplified in the work of Ferrando, Fortunelli, Barcaro, and coworkers who use QM-global optimization methods to model the metal adsorption on oxides.[112–114]

4.3.1 Application of DFT to WGS

DFT is a powerful method for determining reaction mechanisms over metal–oxide systems. We have chosen to review studies that focus on developing catalysts for the water–gas-shift reaction because this is a particularly active research area with numerous examples of DFT application to supported metal–oxide catalysis. The studies first considered herein assess the activity of unsupported gold and copper metal clusters, which can then be compared directly to studies over the analogous oxide-supported systems. The importance of considering particle–support interactions is emphasized, because the oxide support can often play an active role in catalytic mechanisms.

4.3.1.1 Supported Metal–Oxide Catalysts for Water–Gas Shift

There is a growing interest in the production of pure H_2 for use in proton exchange membrane (PEM) fuel cell applications.[34] Typical H_2 production methods implement steam-reforming techniques that convert hydrocarbon feedstocks to a synthesis gas mixture of CO, H_2O, CO_2, and H_2. However, CO is detrimental to the efficiency of PEM fuel cells because it poisons the anode catalyst. For this reason, it is essential to remove CO in a downstream process between the steam-reforming and PEM stages. The water–gas-shift (WGS) reaction removes CO and recovers its energy content in a single process by converting CO and H_2O to CO_2 and usable H_2:

$$CO_{(g)} + H_2O_{(g)} \rightarrow CO_{2(g)} + H_{2(g)} \tag{4.1}$$

Recent DFT studies have characterized the nature of active sites on metal–oxide catalysts for WGS, which emphasize the unique aspects of particle–support interactions in Cu/ceria,[115] Au/ceria systems,[116] and Pt/ceria[117] (among numerous others). The primary intent of this section is to exhibit the utility and limitations of DFT for investigating the many aspects of metal–oxide catalysis, rather than to serve as an exhaustive review of computational work on WGS catalysis in the literature. For a more detailed discussion of the subject, we refer the reader to the recent review of computational work on Au/ceria that can be found in a perspective article by Zhang, Michaelides, and Jenkins.[118]

4.3.1.2 Activity of Au and Cu Nanoparticles

Liu, Rodriguez, and coworkers[115,116,119] applied DFT methods together with experimental studies to investigate how particle–support interactions affect the WGS activity of Au and Cu nanoparticles supported on reducible oxides, such

as ceria and titania. With DFT, the authors calculated reaction energies and activation barriers over unsupported Au and Cu particles.[119] They then compared the activity of Au and Cu particles supported on CeO_2 and ZnO substrates.[115] Their experimental evidence suggests that Au supported on CeO_2 is highly active despite the low activity of both metallic and ZnO supported Au clusters. This contrast suggests that particle–support interactions in CeO_2 supported catalysts play a critical role activating metal clusters for WGS. Computational methods have elucidated the unique behavior of Au/CeO_2 not present on Au/ZnO or unsupported Au-clusters. To accomplish this, the authors utilized DFT methods to evaluate candidate reaction mechanisms by locating optimized intermediate and transition structures on the potential energy surface.

To better understand the effect of the oxide support, the authors first study the activity of unsupported particles in the gas phase. The authors constructed model Au and Cu particles consisting of 29 atoms in a pyramidal geometry, whose shape and size are consistent with those observed by scanning tunneling microscopy[120] and X-ray diffraction.[121] Once the reaction energetics over unsupported particles are understood, they can be directly compared to energetics over the same particles supported on an oxide substrate. The catalytic performance of clean Au(100) and Cu(100) surfaces was also evaluated for comparison against the performance of the nanoparticles. This provides a basis for assessing the importance of edge and corner sites in the reaction mechanism. The authors investigated two reaction mechanisms. The first is a redox mechanism, shown in Figure 4.1, which features the oxidation of the metal surface by the adsorption of CO and H_2O, followed by a subsequent surface reduction as adsorbed O atoms combine with CO before desorbing as CO_2. The second is an associative mechanism that proceeds through either formate or carboxyl surface intermediates. The optimized structures and corresponding reaction energy diagrams for these mechanisms are shown in Figures 4.1 and 4.2.

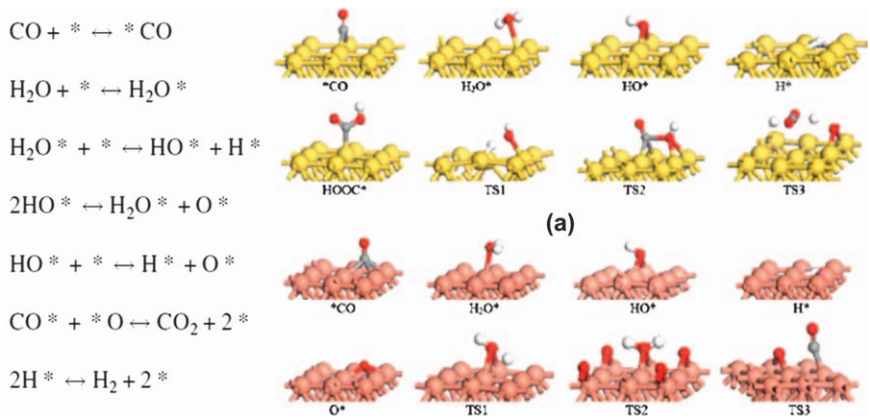


Figure 4.1 WGS redox reaction mechanism (left) and structures (right). (Adapted with permission from ref. 119 Copyright 2007 American Institute of Physics)

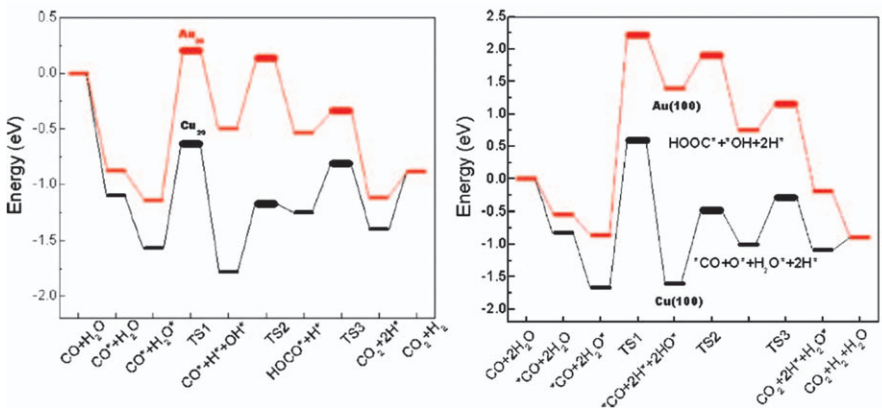


Figure 4.2 DFT energies for WGS reaction path over Au and Cu nanoparticles (left) and surfaces (right), demonstrating that lower energy paths are available over Cu catalysts compared to Au in the absence of support effects). (Adapted with permission from ref. 119 Copyright 2007 American Institute of Physics)

DFT allowed the authors to compare the energetic favorability of the possible mechanisms, thus identifying the likely reaction path over different catalytic surfaces. Cu(100) favors the redox mechanism, while the Au(100) surface (as well as the unsupported Au_{29} and Cu_{29} clusters) favors the associative mechanism. The reaction energy diagrams shown in Figure 4.2 further suggest that the rate-limiting step for the WGS reaction is the dissociation of adsorbed water for mechanisms over both Au and Cu. Determining the magnitude of the reaction barrier for the rate-limiting step allowed the authors to estimate the relative reaction rates over the varying catalytic surfaces. These results show that Cu is consistently more active than Au, and that step and edge sites lead to a further increase in the reaction rate.

If support effects do not influence activity, the above results suggest that supported Cu particles should yield activities higher than supported Au. The authors, however, cite numerous experimental results demonstrating that Au supported on ceria achieves higher conversions and faster reaction rates than the analogous Cu/ceria catalysts. The authors' experimental reaction rates and computed activation barriers demonstrate the low activity of Au compared to Cu, and suggest that high activity of Au/ceria must be the result of support interactions that activate the otherwise inactive Au clusters. Conversely, DFT results predicting low activity over unsupported Au are in qualitative agreement with the experimentally determined activity of Au supported on ZnO, which suggests that ZnO does not activate the particle or participate in the reaction mechanism. Liu, Rodriguez, and coworkers[115] also computed the activity of charged Cu and Au particles to determine whether charge transfer between the cluster and the support could alone alter the activity of the metal clusters. The DFT computed activation barriers over such ionic clusters show that neither cationic nor anionic Au particles are as active as Cu particles, demonstrating that

the activity of Au/ceria is not solely the result of active sites on ceria-stabilized ionic Au species. Since the metal sites alone cannot activate the gas-phase reactants, the reaction may proceed through interfacial sites involving both the metal cluster and oxide support. The support not only activates the metal clusters, but likely plays a role in the reaction mechanism. The authors contrast this observation with the behavior of ZnO supported catalysts, in which the oxide may merely act as an inert substrate supporting the Au cluster for this reaction. These results demonstrate the utility of DFT (applied together with experimental studies) for differentiating between *active* and *spectator* supports.

4.3.1.3 *Characteristics of Au Supported on CeO_2*

DFT methods can model both the structural and the electronic properties of a metal–oxide system. Since the computational expense of DFT scales heavily with system size, it is often necessary to analyze small model systems that, hopefully, capture the fundamental behavior at play in larger systems. This strategy was employed by Liu and coworkers, who conducted systematic DFT calculations investigating the binding trends and electronic properties of a *single* Au atom adsorbed on stoichiometric $CeO_2(111)$, on $CeO_2(111)$ with O-vacancies, and on reduced $Ce_2O_3(0001)$ surfaces.[71] Many properties observed for the single Au atom can be extended to make predictions regarding the properties of larger clusters. Au atoms bind most strongly to O-vacant sites on the partially reduced CeO_2 surface ($\Delta E_{bind} = -1.86$ eV), and they bind least strongly to the fully reduced Ce_2O_3 surface ($\Delta E_{bind} = -0.86$ eV). Binding affinity on the stoichiometric CeO_2 is also considerable in magnitude, with $\Delta E_{bind} = -1.26$ eV. Using a Bader charge analysis,[122] the authors demonstrated that Au adsorbed in an oxygen vacancy on reduced CeO_2 has a negative partial charge (-0.58 e), while Au adsorbed on stoichiometric CeO_2 has a positive partial charge ($+0.35$ e). The negative partial charge on Au adsorbed in an O-vacancy suggests that Au acts as a reduction center, accepting a portion of the two electrons left in the surface after the formation of an O-vacancy. This demonstrates that oxygen vacancies serve as anchor sites that accommodate negative $Au^{\delta-}$ species adsorbed on the ceria surface.

However, the authors note that CO binding energy calculations demonstrate that positive Au species, rather than negative species, are capable of adsorbing CO, and that these sites may play a key role in the WGS reaction mechanism. A density of state (DOS) analysis can determine how ceria stabilizes an active $Au^{\delta+}$ species within $Au_{n>1}$ clusters. The authors posit that the empty band of cerium *f*-orbitals near the Fermi level accepts a Au $6s^1$ electron that would otherwise enter an anti-bonding Au(*6s*)O(*2p*) orbital. A single Au atom can serve as a reduction center, stabilizing an O-vacancy in the ceria lattice. In turn, cerium atoms can subsequently oxidize adsorbed Au atoms by accepting a Au-$6s^1$ electron in the *f*-band. This explains the WGS activity of Au supported on *reduced* CeO_2: the O-vacancy serves as an anchor site for an Au atom, which then serves as a nucleation base for subsequent Au atoms that are oxidized by the surrounding Ce neighbors. This creates $Au^{\delta+}$ sites that favorably adsorb

gas-phase reactants, which can then interact with neighboring sites on both the cluster and the oxide. This result offers guidance for determining the behavior of larger Au clusters on the ceria surface, and how these structures may create active sites for the WGS mechanism. This study demonstrates the value of DFT for analyzing the confluence of electronic and structural particle–support interactions that affect reaction mechanisms.

4.3.1.4 WGS Mechanism over Au Supported on CeO_2

Next we consider studies that use DFT explicitly to evaluate the energetics of candidate reaction mechanisms over various Au/ceria surface configurations. Chen *et al.*[26,123] used DFT + U to probe possible reaction mechanisms that proceed over the Au/ceria interface. In one study they examined redox and formate mechanisms over Au_3 and Au_{10} clusters on $CeO_2(111)$.[123] Both mechanisms feature rate limiting water O–H bond dissociation steps occurring over CeO_2 O–vacancies. The redox mechanism requires H_2O to dissociate after filling an oxygen-vacancy on the ceria surface, leaving H_2 on the metal cluster after oxidizing the support. The CO reactant, which is adsorbed on the metal cluster, then reduces the oxide by removing an O atom from the ceria lattice, thus recreating an O-vacancy. Similarly, the formate mechanism requires an O–H bond breaking step after H_2O adsorbs in an O-vacancy. CO adsorbed on the metal cluster then removes an H atom from the OH group adsorbed in an oxygen-vacancy, creating a CHO group on the cluster. The resultant CHO group then removes an O from the oxide surface and desorbs as CO_2, with an H atom left behind on the metal cluster. This recreates the O-vacancy in the surface, and leaves an H atom that can react with H_2O in the next cycle, thus desorbing H_2 and leaving OH adsorbed in the O-vacancy. The authors consider that breaking OH bonds to refill O-vacancies is rate limiting in both cases. DFT is used to determine the activation barriers for possible elementary steps that accomplish the necessary O-vacancy filling step. The authors conclude that both mechanisms must overcome a reaction barrier greater than 1 eV, which is prohibitive at low temperatures. This suggests that neither mechanism can explain the experimentally observed activity of Au/ceria toward WGS at low temperatures, and for this reason they propose a different mechanism, the carboxyl mechanism, that is unique to the supported system and not subject to this limitation.

In a second study,[26] Chen *et al.* propose a carboxyl mechanism for WGS over Au/ceria in which the rate-limiting H_2O dissociation step occurs at the Au/ceria interface. The mechanism, summarized in the reaction energy diagram shown in Figure 4.3, requires H_2O from the gas phase to fill an oxygen vacancy near an Au cluster. Once in the vacancy, the H_2O molecule dissociates, as shown in Figure 4.4, by allowing H to adsorb on the metal cluster while leaving OH behind on the oxide. The OH group then reacts across the Au/ceria interface with CO adsorbed on the metal cluster, thus forming a carboxyl group and regenerating the oxygen vacancy in the ceria lattice. To support the carboxyl mechanism further, the authors use a microkinetic model to compare the performance of the carboxyl mechanism to the formate and redox mechanisms

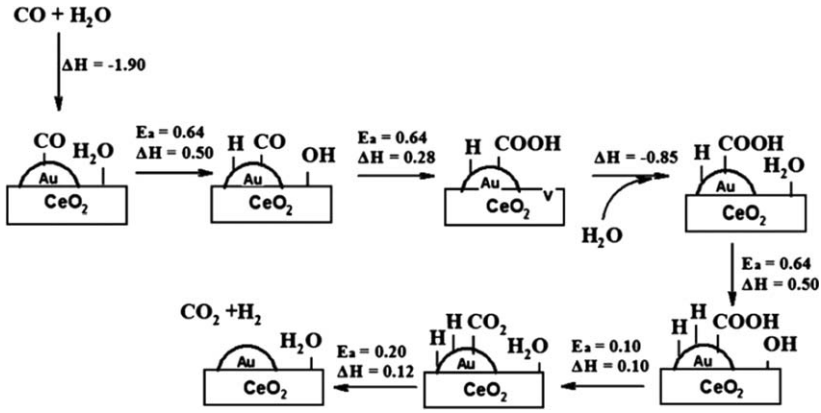


Figure 4.3 Water–gas-shift reaction mechanism proceeding through a carboxyl surface intermediate as proposed by Chen *et al.*
(Reprinted with permission from ref. 26 Copyright 2011 Royal Society of Chemistry)

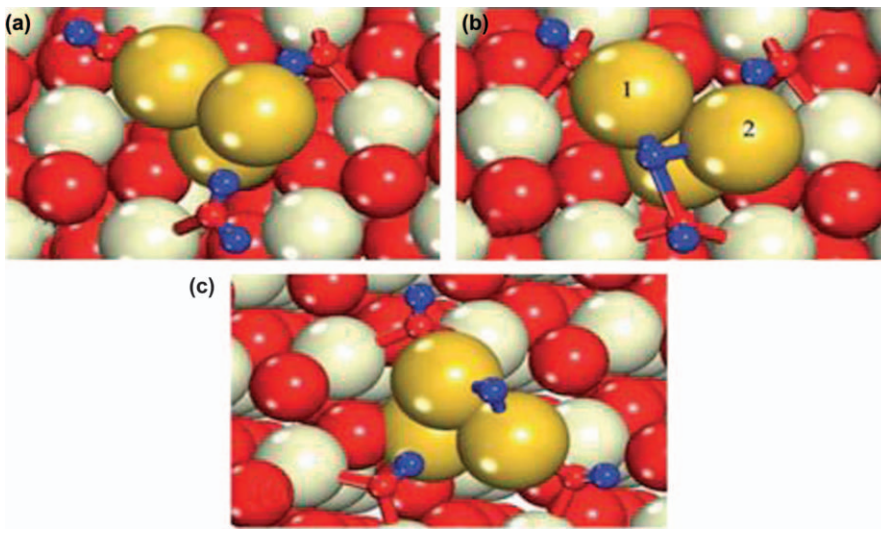


Figure 4.4 (a) Initial, (b) transition, and (c) final structures for the dissociation of H_2O across the Au–CeO_2 interface, where H = blue, Au = yellow, O = red, Ce = white.
(Reprinted with permission from ref. 26 Copyright 2011 Royal Society of Chemistry)

discussed earlier. The result shows that the carboxyl mechanism yields a higher rate than the formate and redox mechanisms, and therefore better explains the high activity of Au/ceria catalysts. The carboxyl mechanism requires sites at the Au–ceria boundary, thus demonstrating the importance of considering

metal–support interactions when postulating reaction mechanisms on supported metal–oxide catalysts.

4.3.2 *Ab Initio* Thermodynamics

The studies discussed in the previous sections did not explicitly consider the effects of a gas-phase environment at operating temperatures and pressures. Although providing useful insight, the methods employed could not model thermodynamic stabilities, adsorbate coverage effects, or *free energy* differences. This section will review studies that apply *ab initio* thermodynamics to accomplish such assessments. We first review the work of Reuter and Scheffler on RuO_2 systems, which was one of the first applications of *ab initio* thermodynamics to catalysis. This study assessed the catalytic behavior of an oxide, which is readily extended to studies of supported metal–oxide catalysts because the oxide support often plays an active role in the catalytic mechanism. We then will consider examples that apply *ab initio* thermodynamics to supported metal–oxide catalysts. In particular, we highlight systems in which particle–support interactions play an important role in overall catalytic activity.

4.3.2.1 Investigating Oxide Supports

4.3.2.1.1 Background. When conducting DFT to assess catalytic mechanisms over a surface, it is essential to choose an appropriate model. That is, one must choose a model surface that is thermodynamically or kinetically stable under the reaction conditions of interest. Reuter and Scheffler applied the formalism of *ab initio* thermodynamics in a series of studies that assess the structure, stability, and reactivity of the $RuO_2(110)$ surface in equilibrium with a mixed O_2 and CO atmosphere.[100,102] This series of publications was one of the earliest applications of *ab initio* thermodynamics, and effectively demonstrates the methodology for bridging the gap between DFT calculations (at zero T and P) and experimental results under catalytic reaction conditions. They applied *ab initio* thermodynamics to determine the relative stability of $RuO_2(110)$ surface terminations in different oxidation states as a function of T, P_{O_2}, and P_{CO}. Although these studies were explicitly applied to the catalytic properties of RuO_2 toward CO oxidation, the formalism presented therein can be readily extended to studies investigating the stability and activity of oxide surfaces in supported metal–oxide systems. This methodology can be used to determine plausible surface models for detailed analyses of the reaction mechanism energetics.

4.3.2.1.2 Stability of RuO_2 Surface Phases. Reuter and Scheffler first outline the theory and methodology of *ab initio* thermodynamics in a study applied to RuO_2 in equilibrium with an O_2 atmosphere.[100] The authors consider three possible surface terminations of $RuO_2(110)$. The $RuO_2(110)$–O_{bridge} surface is predicted to be the most stable because it has the lowest number of uncoordinated O atoms and has no net dipole. The second

surface considered is $RuO_2(110)$–O_{cus}, which features coordinately unsaturated (cus) oxygen atoms placed directly above the surface row of Ru atoms, creating a 6-fold coordination on all surface Ru atoms. The final structure, $RuO_2(110)$–Ru, has rows of 5-fold coordinated cus-Ru atoms and 4-fold coordinated bridge-Ru atoms on the oxide surface. The authors investigate the stability of the three $RuO_2(110)$ surface terminations by calculating the surface free energy of each termination as a function T and P_{O_2}. These calculations show that at typical operating temperatures for CO oxidation ($T = 600$ K), the O-bridge surface termination is most stable at lower P_{O_2} and that the over oxidized O-cus surface is favored at high P_{O_2}. The Ru terminated surface is never favored in the P_{O_2} ranges of interest. The results reveal that the RuO_2–O_{bridge} surface termination is not always the most stable, and that a different surface phase is likely to form at high P_{O_2}. The authors argue that the emergence of the RuO_2–O_{cus} phase at high O_2 chemical potentials, which was previously unknown, must be considered when interpreting experimental data collected under such conditions. This result highlights the importance of investigating the stability of a system for the entire temperature and pressure ranges of interest before drawing conclusions regarding the nature of active surface sites.

In subsequent studies, the authors incorporated the effect of both P_{O_2} and P_{CO} on surface site occupation.[101] Using the resultant surface phase diagram they predict the regions in T,P space that are likely to display the highest catalytic activity. They conclude that the boundary between the O_{br}/CO_{cus} phase and the O_{br}/O_{cus} phase will balance CO_2 desorption, creating O-vacancies at O_{cus} sites with the subsequent filling of O_{cus} vacancies by gas-phase oxygen. Deep in the O_{br}/O_{cus} phase region, catalytic activity will be hampered by the high O_{cus} vacancy formation energy; whereas, in the CO_{br}/O_{cus} phase, CO_{br} will consume O_{cus} faster than the resultant vacancies can be refilled by the gas phase. These results demonstrate the important role phase coexistence plays in catalytic mechanisms, and they exemplify the utility of *ab initio* thermodynamics for identifying favored reaction paths in T, P regions, bridging the gap between quantum calculations and experimental conditions.

4.3.2.2 Investigating Metal Clusters on Oxide Supports

4.3.2.2.1 Background. The formalism presented in the previous section for predicting the stability of oxide surfaces in equilibrium with a multi-component gas phase is readily extended to systems that contain catalytic metal particles supported on oxide surfaces. Identifying stable particle–support constructions is indispensable for predicting the catalytic activity of the particle–support interface. This section will outline studies on reducible oxides (TiO_2 and CeO_2) that display unique particle–support interactions where the oxide support plays an active role in the catalytic mechanism. These examples demonstrate the ability of *ab initio* thermodynamics to determine the stability of metal clusters on oxide supports under realistic catalytic conditions. Such calculations can be used in concert with DFT reactivity studies

to assess the activity of stable metal–oxide surfaces. Together, these methods offer a powerful means for predicting and characterizing the catalytic activity of a wide range of metal–oxide systems.

4.3.2.2.2 Au/TiO_2. In a study by Laursen and Linic,[68] the authors analyzed the behavior of Au on titanium oxide. They considered two model Au formations on the oxide surface: a nanorod and a 2 mono-layer sheet. They assessed the stability of these formations as a function of oxygen pressure over reduced, stoichiometric, and oxidized surfaces. The results shown in Figure 4.5(c) demonstrate that both the nanorod and sheet constructions are most stable over the oxidized support. Formations over the reduced support can be stabilized at high temperatures in a highly reducing atmosphere at low oxygen chemical potential.

After determining the stability of these model systems, the authors calculated the adsorption and dissociation energies of O_2, along with the adsorption energy of CO, to probe the activity of possible sites on the supported Au clusters. The results of these calculations, shown in Figure 4.5(a), show that the adsorption energy for both O_2 and CO is significantly more exothermic over both the reduced and oxidized catalysts compared to equivalent sites on Au formations supported by a stoichiometric titania surface. The same trend holds for the dissociation barrier of O_2 over all three catalyst models shown in Figure 4.5(b). The authors note that the most favorable adsorption sites involve Au atoms that are adjacent to Au atoms bound to point defects on the reduced and oxidized surfaces (the defects consist of an O-vacancy or an added O-atom, respectively). They explain this trend in terms of bond conservation theory, in which the strengthening of an Au–defect bond will weaken the neighboring Au–Au bond to the next adjacent Au atom, thus creating a chemically active site on the adjacent Au atom for bonding with a gas-phase molecule. The authors support this theory with charge density and local density of states (LDOS) calculations, which show that charge transfer between the surface and the particle (either from Au to the surface for oxidized-TiO_2 or *vice versa* for reduced-TiO_2) results in a strong polar–covalent bond between the Au atom and the point defect. This in turn weakens the bond to the next adjacent Au atom, making adsorption on that atom more favorable. The strong covalent bonds present in both the reduced and oxidized systems are largely absent in the stoichiometric system. Both unsupported Au nanorods and nanorods supported on the stoichiometric surface display low activity because neither system is activated by the covalent-type Au–oxide bonds present in the defected systems. The authors conclude that point defects on the oxide surface not only serve as anchor points for the metal clusters, but that they are also largely responsible for activating the particle through charge transfer. This contrasts with the findings of Corma and coworkers, who found that *neutral* Au particles display higher activity toward the dissociation of H_2, demonstrating the need to assess the relative activity of both charged and neutral clusters.[56,57] These studies exemplify the importance of particle–support interactions for predicting the catalytic activity of a system, and demonstrate how *ab initio*

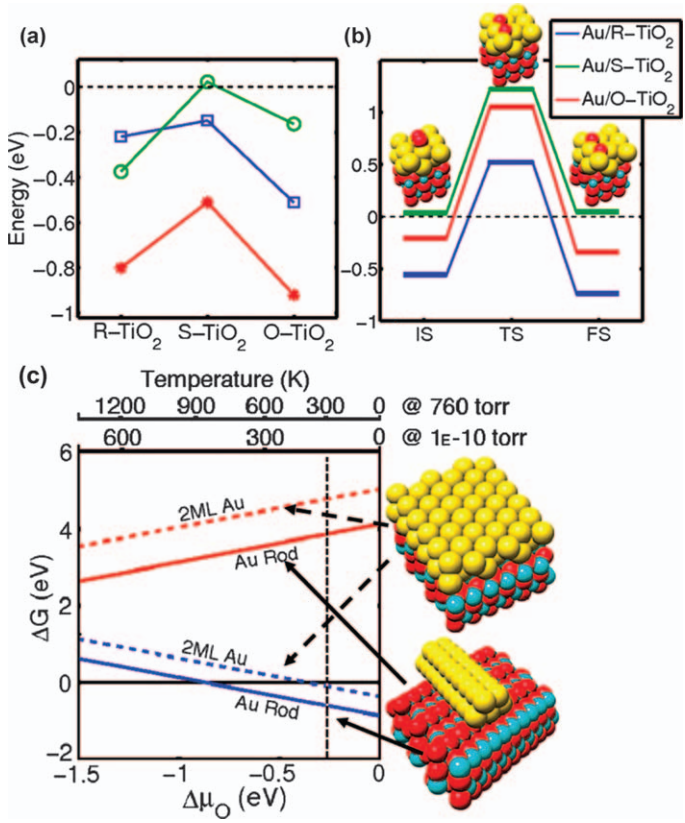


Figure 4.5 (a) Adsorption energy of molecular oxygen and CO with respect to gas-phase species on reduced, stoichiometric, and oxidized Au/TiO_2 surfaces for oxygen on a Au sheet (green, o), oxygen on a Au nanorod (blue, □), and CO on a Au sheet (red, *). (b) Activation barrier for O_2 dissociation over an Au bilayer sheet on reduced, stoichiometric, and oxidized TiO_2. (c) Free energy calculated as a function of oxygen chemical potential for Au rod and bilayer sheet structures on reduced (red, positive slope) and oxidized (blue, negative slope) TiO_2.
(Reprinted with permission from ref. 68 Copyright 2009 American Chemical Society)

thermodynamics – in conjunction with standard DFT energy calculations – can explain the stability and activity of oxide-supported metal clusters.

4.3.2.2.3 Pt/TiO_2. In a similar study, Ammal and Heyden[124,125] investigated the effects of particle–support interactions between Pt and TiO_2 under water–gas-shift T,P conditions. They first determined the effect of Pt clusters on the reducibility of titania by systematically assessing the stability of small Pt_n ($n = 1$–8) clusters adsorbed on a partially reduced $TiO_2(110)$ surface. Using *ab initio* thermodynamics, the authors calculated the free energy

change for forming a single oxygen vacancy in the TiO_2 surface under oxidizing (O_2-rich) and reducing (CO-rich or H_2-rich) atmospheres. Not surprisingly, the authors found that vacancy formation is never favorable under oxidizing conditions. However, they found that Pt clusters greatly increase the favorability of oxygen vacancy formation under CO- and H_2-rich atmospheres. This is demonstrated by the results reproduced in Figure 4.6(a)

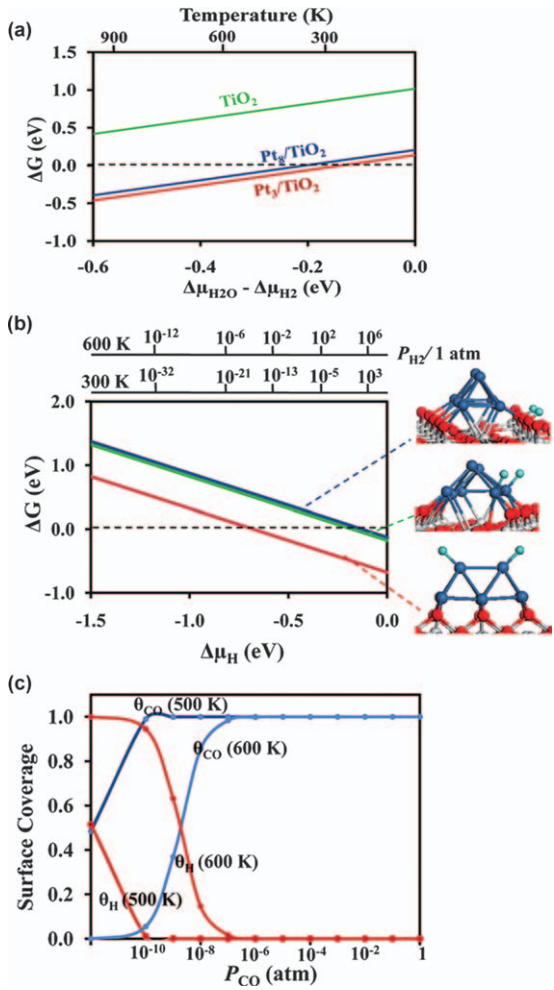


Figure 4.6 (a) Free energy difference for oxygen vacancy formation on the clean TiO_2 surface (green, top), on TiO_2 in the presence of Pt_8 (blue, middle), and in the presence of Pt_3 (red, bottom). (b) Free energy of H_2 adsorption on the oxide (blue, top), the metal–oxide interface (green, middle), and on the metal cluster (red, bottom). (c) Computed equilibrium surface coverage of CO and hydrogen at constant $P_{H2} = 1$ atm.
(Reprinted with permission from refs. 125,126 Copyright 2011 American Chemical Society)

showing vacancy formation free energy as a function of P_{H_2} and T. Vacancy formation energy is significantly lower for surfaces with Pt_n clusters compared with the clean TiO_2 surface, which never favors oxygen vacancy formation in the T,P regions of interest. These results are similar to those of Laursen and Linic discussed above, in which the metal particle is anchored to the reduced TiO_2 surface through a covalent-type interaction that is largely absent over the stoichiometric surface. The subsequent charge transfer from the reduced TiO_2 surface to the Pt particle alters the adsorption behavior of gas phase molecules on the Pt cluster. This further demonstrates the unique effects of particle–surface interactions.

The authors next assessed the stability of H_2 and CO gas phase adsorbates on the Pt_8 cluster, at the Pt–TiO_2 interface, and on the TiO_2 surface adjacent to the Pt cluster. They computed the free energy of adsorption for gas-phase molecules at these sites, as well as calculating the equilibrium surface coverage as a function of partial pressure and temperature. It is possible that the WGS reaction mechanism involves hydrogen spillover from the metal particle to the TiO_2 surface. For this reason, the authors investigated H_2 adsorption in the vicinity of a Pt cluster, which is summarized in Figure 4.6(b). The presence of H atoms at the Pt–TiO_2 interface can have adverse effects on the catalytic performance of the surface if H atoms bind too strongly to the interfacial sites under WGS conditions. The results shown in Figure 4.6(b–c) show that H atoms bind more strongly to Pt cluster sites than to interfacial Pt and O sites, suggesting that interfacial sites will not hinder hydrogen spillover between the Pt cluster and the oxide surface. By calculating the surface coverage of CO and H at various temperatures, the authors also show that CO adsorption on Pt cluster sites will dominate over H adsorption, indicating that Pt cluster sites will be filled by adsorbed CO under WGS conditions. These results demonstrate that interfacial sites will remain open under WGS conditions, and that they may be responsible for the high WGS activity of the catalyst. The authors note that this result corroborates experimental findings,[126,127] in which the WGS reaction rate scales positively with the interfacial boundary length. Using *ab initio* thermodynamics, the authors were able to assess the stability of candidate model systems for the Pt/TiO_2 surface under WGS conditions. Having identified a plausible model system, the authors mention their plans to use DFT to investigate the WGS mechanism over the Pt/TiO_2 interfacial model they identified in this study. We note that the authors also conducted a similar study on an analogous Pt/CeO_2(111) system,[128] which applies similar methods to characterize the effect of Pt–ceria interactions.

4.3.2.2.4 Cu/CeO_2. The previous studies have shown that metal clusters typically enhance the reducibility of metal–oxide surfaces, thus altering the electronic properties of both the metal cluster and the oxide surface. Fabris and coworkers[73] demonstrate that this trend does not hold for all metal–oxide supported systems. They conducted a systematic computational study of Cu/CeO_2 systems that feature Cu atoms adsorbed on a stoichiometric surface, on a surface containing oxygen-vacancies, and on a surface containing

cerium-vacancies. In particular, they use *ab initio* thermodynamics to assess the stability of Cu adsorption on the various CeO_2 surfaces in equilibrium with an oxygen atmosphere. *Ab initio* thermodynamics predicts that Cu typically prefers to adsorb on the stoichiometric CeO_2 surface, rather than on surfaces containing oxygen-vacancies or cerium-vacancies. These results are shown in Figure 4.7, where the free energy of adsorption is plotted against oxygen chemical potential. Under oxidizing conditions, the formation of surface solution phases featuring a Cu atom substituted for a Ce atom is thermodynamically favored over Cu adsorption on the stoichiometric surface.

These observations lead to interesting conclusions regarding the redox behavior of Cu/CeO_2. The authors note that substituted Cu ions form stabilized CuO_4 units that permit the reversible release of oxygen under catalytically relevant thermodynamic conditions. This is reflected by the stability of the $O^{vac}/Cu@Ce^{vac}$ phase that neighbors the $Cu@Ce^{vac}$ phase seen in Figure 4.7. This demonstrates that the surface can act as an oxygen buffer during reactions conducted under oxidizing conditions. Furthermore, it shows that redox processes over Cu/CeO_2 surfaces do not involve the reduction of cerium atoms (from Ce^{4+} to $Ce^{3}+$) neighboring oxygen-vacancies, which is typically characteristic of reaction paths involving a CeO_2 redox process. Other metal–ceria systems (*e.g.* Au/ceria, discussed in previous sections) typically feature a redox mechanism in which metal clusters adsorbed on the surface stabilize the

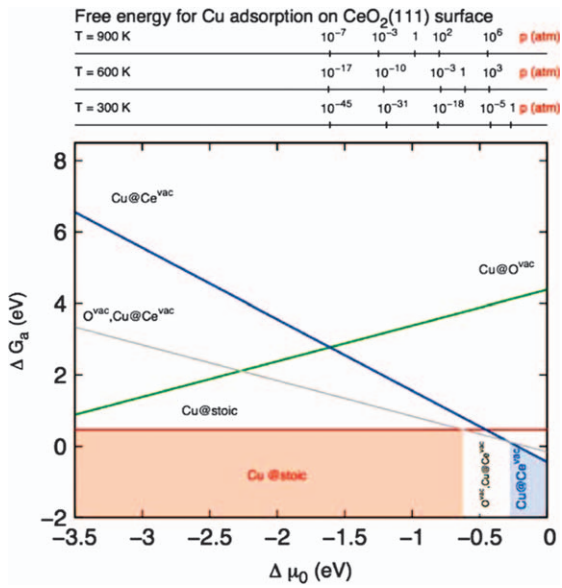


Figure 4.7 Free energy for the adsorption of Cu adatoms on stoichiometric ceria (red), on ceria with an oxygen vacancy (green), on ceria with a cerium vacancy (blue), and on ceria with both an oxygen and a cerium vacancy (gray). (Reprinted with permission from ref. 73 Copyright 2010 American Institute of Physics)

formation of oxygen vacancies. In the case of copper, however, oxygen vacancies are stabilized by substitution defects where Cu atoms fill Ce vacancies. *Ab initio* thermodynamics enabled the authors to determine Cu/CeO_2 structures that are thermodynamically favored under reactive conditions. Together with electronic structure calculations, the authors demonstrate that the redox behavior of Cu/CeO_2 does not follow the typical redox mechanism of other metal/ceria surfaces.

4.3.3 Classical Atomistic Modeling

The computational expense of quantum (QM) methods limits these methods to highly idealized system models, such as the single crystal surfaces and small ($\sim M_{n<8}$) clusters considered in the studies above. However, many catalytically interesting properties of supported metal–oxide systems arise from structural irregularities and possibly from the dynamic effect of surface reconstruction during reaction. Although it is difficult to characterize such effects using QM, it becomes possible with the use of empirical force-fields that have a relatively low computational expense. Reactive force-fields are designed to model bond dissociation and formation, and can therefore be implemented in *reactive* molecular dynamics (RMD) simulations that are capable of describing the dynamic nature of a catalyst at longer length and time scales (up to $\sim 10^4$ atoms and ~ 1 ns). This section will review research examples implementing the ReaxFF potential to model dynamic catalyst behavior under reactive conditions. The work described herein demonstrates that reactive force-field methods can be readily extended to model supported metal–oxide catalysts.

4.3.3.1 The ReaxFF Force Field

4.3.3.1.1 Background. ReaxFF uses bond-length/bond-order relationships to model bond formation and dissociation in covalent systems, making it readily applicable to catalytic systems. This section will summarize two studies that demonstrate the ability of ReaxFF to describe supported metal–oxide catalysts accurately. The first study focuses on the catalytic properties of nickel surfaces and particles interacting with hydrocarbon reactants. The second study demonstrates the use of ReaxFF to describe complicated multi-metal–oxide (Mo/V/O) catalysts under reactive hydrocarbon environments. Together these studies show how ReaxFF extends the reach of computational methods to length and time scales required for characterizing the dynamic behavior of supported metal–oxide catalysts.

4.3.3.1.2 Hydrocarbon Catalysis on Metals. The ReaxFF force-field can model catalytic properties of metal clusters toward C–H and C–C bond activation. This is exemplified by the work of Mueller, van Duin, and Goddard, in which the authors developed a Ni/C/H potential by parameterizing the ReaxFF force-field to reproduce a training set populated with DFT results for hydrocarbon–nickel adsorption energies, activation barriers, and bulk

formation energies.[129] The authors note that surface defects may play an important role in reaction paths that lead to the high activity of the nickel surface, and that surface science experiments of hydrocarbon chemisorption and decomposition often seek to limit the number of defects on the surface in order to characterize the reactivity of ordered low-index nickel surfaces. Similarly, DFT is computationally unable to reach the necessary system sizes for analyzing the effect of defects on irregular surfaces. For these reasons, the authors applied the Ni/C/H force-field in reactive molecular dynamics (RMD) simulations to assess the reactivity of various hydrocarbons over Ni catalyst particles that expose multiple irregular surface terminations, yielding insight into the role defect sites play in Ni-catalyzed hydrocarbon reaction mechanisms.[130]

The study consisted of a series of RMD simulations with a spherical Ni particle consisting of 468 atoms surrounded by a single-species hydrocarbon gas phase in an $80\times80\times80$ Å^3 periodic box. Simulations were conducted with six representative hydrocarbon species (methane, ethyne, ethene, benzene, cyclohexane, and propene) that were chosen to cover a variety of hydrocarbon behaviors arising from varying degrees of saturation. In each simulation, the temperature was ramped from 500 K to 2500 K at a rate of 20 K ps^{-1} over the course of the 100 ps simulation. The authors extracted species population data at each time step, thus identifying reaction intermediates and their corresponding formation and decomposition rates. These data were then used to construct reaction networks from which elementary reaction paths can be identified. The results and analysis of the simulation for propene are reproduced in Figure 4.8. The figure depicts the initial and final structures of the system, as well as the molecular population analysis as the simulation progresses. As seen in the figure, the authors can determine the temperatures at which key reactive events occur. This includes the temperatures at which gas-phase propene begins to chemisorb on the surface (T_A), that at which dehydrogenation commences (T_H), and that at which carbon–carbon bonds are broken (T_C). These temperatures yield insight into the kinetic barriers for hydrocarbon dissociation over nickel particles.

Analysis of the simulation trajectory yields reaction networks that reveal the preferred reaction mechanism over the Ni catalyst. For propene, molecules begin to chemisorb rapidly because C–C π bonds can be readily broken to form C–Ni σ bonds. Correspondingly, chemisorption of saturated hydrocarbons does not occur until higher temperatures, where C–H bonds are broken by dehydrogenation. Analyses of the simulation trajectories show that C–H bond breaking is catalyzed by the insertion of a Ni atom within the C–H bond. The simulations also reveal that dehydrogenation typically precedes C–C bond scission, which is reflected in the figure by a low T_H compared to T_C. In contrast with DFT approaches, reaction paths are identified by MD analysis as opposed to being hypothesized and evaluated explicitly, which allows the method to find reaction paths that are not immediately apparent. These results have many implications for designing Ni catalysts for hydrocarbon treatment. This work exemplifies the ability of empirical force-fields to characterize reactions

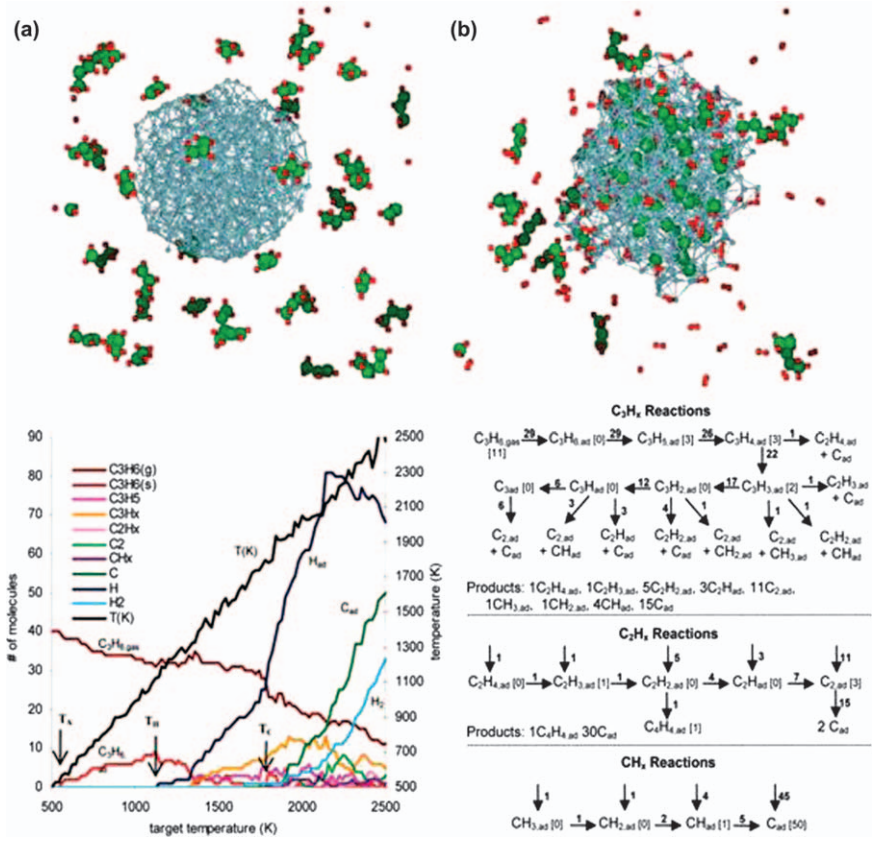


Figure 4.8 (top) Initial and final structures of a RMD simulation demonstrating propene adsorption and decomposition on a nickel nanoparticle. (bottom) Molecular populations plotted against simulation temperature, and the corresponding reaction networks. Numbers in brackets indicate the number of species present at the end of the simulation and numbers over arrows indicate the corresponding number of reactive events. (Adapted with permission from ref. 131 Copyright 2010 American Chemical Society)

involving covalent hydrocarbon species interacting with irregular metal surfaces, which is necessary for modeling supported metal–oxide heterogeneous catalysts at large scales.

4.3.3.1.3 Multi-metal–Oxide Catalysis. The ReaxFF potential has also been utilized to study the catalytic properties of complex metal oxides. Chenoweth *et al*. developed and implemented a V/O/C/H force-field that, when combined with the existing hydrocarbon force-field, can model the interaction between gas-phase hydrocarbons and the vanadium oxide surface.[131,132] For motivation, the authors cite numerous examples in which V_2O_5 is used to catalyze industrial processes that selectively oxidize both

saturated and unsaturated hydrocarbon species, such as the conversion of methanol to formaldehyde. They stress that, although V_2O_5 can be used for the oxidative dehydrogenation of methanol, it lacks selectivity when applied to the partial oxidation of larger hydrocarbons. The selectivity of vanadium oxide can be improved with the addition of appropriate supports and metal dopants (such as V, Te, Ta, and Nb) that alter the surface geometry and electronic structure. The addition of supports and dopants to form multi-metal–oxide catalysts often introduces a confluence of structural and electronic effects that are too complex for experimental efforts to untangle. For this reason, the authors propose that theoretical methodologies, such as RMD simulations *via* ReaxFF, are essential for establishing the necessary atomistic understanding of reaction processes over the complex oxide surface for designing more selective catalysts.

To demonstrate the validity of the ReaxFF potential for modeling catalytic V/O/C/H interactions, the authors simulated the oxidative dehydrogenation of methanol over the $V_2O_5(001)$ surface.[131] They conducted a 250 ps NVT-MD simulation of a three-layer oxide slab surrounded by 30 gas-phase methanol molecules in a $20\times20\times20$ Å^3 periodic box. A dual temperature constraint was applied in which the oxide slab was held at 650 K and the gas-phase molecules were held at 2000 K. The final MD structure, along with the species population analysis, is shown in Figure 4.9. Methanol species adsorb on the oxide surface and, after rearrangement, desorb as formaldehyde. Analysis of molecular populations and the MD trajectory shows that the reaction mechanism proceeds through the abstraction of hydrogen from the methyl group. The authors note that this result seems to contradict experimental evidence (from DRIFTS spectroscopy) that suggests an O–H bond dissociation mechanism. The authors demonstrate, however, that the C–H abstraction path is preferred on the fully

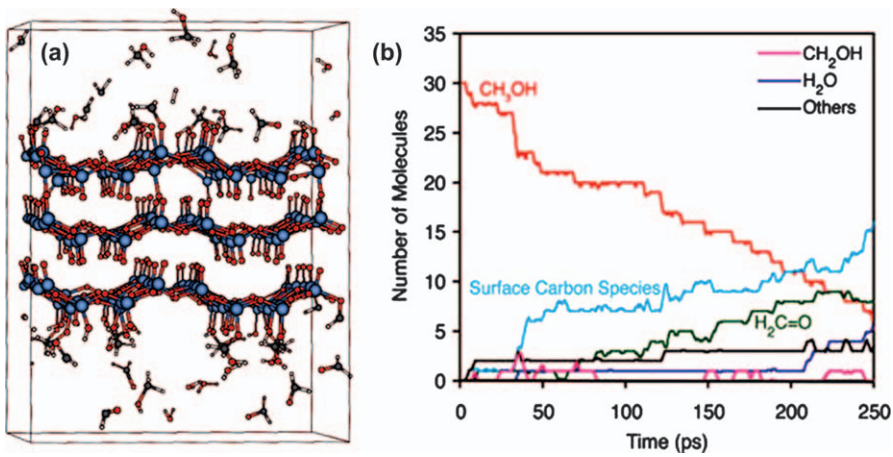


Figure 4.9 (a) Final structure of a NVT–MD simulation of methanol interacting with a V_2O_5 slab and (b) the corresponding species population plot. (Reprinted from ref. 132 Copyright 2008 American Chemical Society)

oxidized $V_2O_5(001)$ surface and that the O–H abstraction path is preferred over a defect site where an oxo- group has been removed from the surface. This result demonstrates the viability of the ReaxFF method for oxide-based catalysis, as well as exhibiting the power of atomistic modeling for complementing experimental efforts to discern structural factors that influence the reaction path.

The ReaxFF method can also be used to determine the structure and nature of active sites on complex oxide surfaces. This is demonstrated by the same authors in a subsequent publication,[132] in which they applied ReaxFF to predict the structure of a highly disordered multi-metal Mo_3VO_x catalyst. As mentioned earlier, the selectivity of vanadium oxide catalysts toward the partial oxidation of hydrocarbons can be altered by the addition of metal dopants to form multi-metal–oxides (MMO). MMO catalysts typically feature partial or mixed occupations of crystallographic sites, making it difficult to characterize the structure and nature of active sites by experimental methods. The authors applied the ReaxFF potential to conduct a combined Monte Carlo/reactive dynamics (MC/RD) procedure to determine the Mo_3VO_x structure. The catalyst surface has metal sites that can be occupied by either Mo or V atoms. The MC/RD scheme identifies the preferred occupation of each site by systematically interchanging two atoms that occupy the same crystallographic site and determining the resultant energy change after re-optimizing the structure through RD. The new structure is either accepted or rejected according to the MC-Metropolis criterion,[133] and the process is repeated until the energy converges at an optimal configuration for the chemical environment of the system.

The authors used the optimized structure to conduct RMD simulations in which hydrocarbons interact with the catalyst. The optimized catalyst displays several channels in the oxide through which hydrocarbons can diffuse to reach active sites. This particular simulation reveals that propane can diffuse through the $C7_2$-labeled channel in the oxide, but not through the $C7_1$-labeled channel. Such observations can reveal structural factors that influence catalyst selectivity, because different hydrocarbon species will tend to diffuse through different channels. This property can be exploited to design highly selective MMO catalysts. This result demonstrates the applicability of the ReaxFF method for characterizing complex metal–oxide catalysts, which can be readily extended to supported metal-cluster/metal-oxide catalysis.

4.3.4 Combined Application: Hydrocarbon Activation over Pd/CeO_2

In the previous sections, we discussed separate applications of DFT, *ab initio* thermodynamics, and empirical force-fields to demonstrate the capabilities and limitations of each computational method applied to supported metal–oxide catalysis. As mentioned earlier, these methods can be applied together to achieve a broader understanding of the factors affecting catalytic activity, scaling from electronic structure to large-scale surface rearrangement and adsorbate coverage effects. This section highlights an ongoing multi-scale study conducted across our laboratory combining these methods in an effort to

identify factors contributing to high activity of Pd/ceria catalysts toward hydrocarbon conversion.

4.3.4.1 Background

Palladium supported on ceria is an effective catalyst for hydrocarbon oxidation and is a promising candidate for application in solid-oxide fuel cell (SOFC) anodes,[48–50] automotive three-way catalysis,[43,46] catalytic combustion,[3,12,87] and water–gas-shift catalysis.[31,134] There is, however, little consensus in the literature regarding the chemical and structural properties of the active sites and reaction mechanisms on the Pd/ceria surface. Under reaction conditions, multiple Pd/ceria surface morphologies are possible, each with unique catalytic activity. Examples of such morphologies include the following: metallic Pd clusters adsorbed on the ceria support, oxidized palladium in PdO surface phases, and palladium incorporated in cerium lattice vacancies as a solid–solution $Pd_xCe_{1-x}O_y$ phase. Experimental evidence suggests that strong electronic interactions between Pd and ceria stabilize oxidized Pd species that in turn are capable of activating C–H bonds in hydrocarbons. This is evident in the work of Colussi *et al.*,[87] which uses high-resolution transmission electron spectroscopy (HRTEM) and DFT to show that Pd/ceria catalysts prepared by solution combustion synthesis (SCS) contain a Pd^{2+}/CeO_x solid solution that is absent in Pd/ceria catalysts prepared by traditional incipient wetness impregnation (IWI). The Pd^{2+}/CeO_x phase forms when Pd atoms are systematically substituted for Ce atoms in the ceria lattice, thus forming a square planar geometry characteristic of Pd^{2+} in bulk PdO. Pd/ceria samples prepared by SCS achieved considerably higher methane combustion rates compared to the analogous catalysts prepared by IWI, suggesting that the presence of Pd^{2+}/CeO_x plays an important role in hydrocarbon activation. Similar results for the combustion of propane and dimethyl ether over Pd/ceria prepared by SCS further emphasize the importance of oxidized $Pd^{\delta+}$ species, while also suggesting the importance of metallic Pd^0 clusters coexisting on the surface with the Pd^{2+}/CeO_x phase.[4] In a similar study, Misch *et al.* used X-ray diffraction (XRD) and X-ray photoelectron spectroscopy (XPS) to characterize $Ce_{1-x}Pd_xO_\delta$ catalysts before and after the catalytic oxidation of methane.[88] Interestingly, their results suggest that metallic Pd^0/CeO_2, rather than an oxidized $Pd^{\delta+}$ species, is essential for high catalytic activity. Furthermore, a recent study by Gorte and coworkers[12] has synthesized Pd-core/ceria-shell catalysts that achieve high methane conversion rates by maximizing the concentration of active Pd/ceria interfacial sites. Together, these results demonstrate the complex nature of Pd/ceria catalysis, in which strong metal–support interactions lead to the coexistence of multiple surface morphologies with unique catalytic properties.

Each of the possible surface morphologies described above may contribute differently to the catalytic activity of the system, and it is essential to identify surface morphologies that are both *stable* and *active* under the reaction conditions of interest. DFT electronic structure calculations and *ab initio* thermodynamics have been applied together to determine the Pd/ceria surface morphologies that are stable under typical operating conditions and that

provide a low energy path for hydrocarbon oxidation. Ongoing efforts will incorporate atomistic simulations *via* the ReaxFF potential to extend our analysis beyond the computational limits of DFT.

4.3.4.2 *Effect of Pd on CeO_2 Surface Reduction*

DFT studies completed in our laboratory examined the nature of oxygen vacancy formation and methane activation over Pd-substituted CeO_2(111), (110), and (100) surfaces.[69] The results demonstrate that methane activation and oxygen-vacancy formation are surface reduction processes, and that the substitution of a Pd atom in the ceria lattice greatly affects the energetics of these processes. The removal of a surface oxygen when forming a vacancy in the ceria lattice results in a *single-spin* gap state above the valence, which is attributed to the single occupation of *4f* states on two cerium atoms (shown in Figure 4.10). This is confirmed by a Bader charge analysis demonstrating that oxygen-vacancy formation results in a negative charge transfer to two cerium atoms in the vicinity of the oxygen-vacancy, further corroborating density of states (DOS) data suggesting that the removal of an oxygen results in the reduction of two cerium atoms from Ce^{4+} to Ce^{3+}. Bader charge analysis confirms that a Ce^{4+} to Ce^{3+} reduction occurs adjacent to CH_3 and H adsorption sites, thus demonstrating that vacancy formation and methane adsorption are surface reduction processes.

A density of states (DOS analysis) can probe the electronic structure changes that occur during surface reductions. This analysis shows the electronic effect of Pd substitution, as well as demonstrates that both O-vacancy formation and dissociative methane adsorption display similar electronic structure rearrangements. The formation of an oxygen-vacancy in Pd-substituted CeO_2(111) results in a spin-paired gap state, in contrast to the single-spin state that appears in clean CeO_2(111). Bader charge analysis shows charge accumulation on Pd, confirming that the spin-paired state is the result of a Pd^{4+} reduction to Pd^{2+}. A comparison of the structure of the Pd-substituted CeO_2 surface before and after reduction shows that, after reduction, Pd assumes a square planar coordination geometry analogous to that of Pd^{2+} in bulk PdO. This demonstrates that the reduction of Pd-substituted surfaces results in one Pd^{4+} to Pd^{2+} reduction, as opposed to two Ce^{4+} to Ce^{3+} reductions over the clean ceria surface. The dissociative adsorption of methane is also a surface reduction process that portrays similar electronic rearrangements, thus leading to the correlation between oxygen vacancy formation energy and methane adsorption energy shown in Figure 4.11. Generally, a substituted Pd atom serves as a reduction center, and the resultant gap state is lower in energy than the analogous gap state over clean CeO_2. This results in lower vacancy formation and methane dissociation energies over Pd-substituted surfaces.

4.3.4.3 *Stability of Pd/Ceria Surface Morphologies*

The DFT method employed, like most QM methods, only applies at zero temperature and pressure, and, correspondingly, cannot evaluate the thermodynamic stability of Pd incorporated surfaces. *Ab initio* thermodynamics must

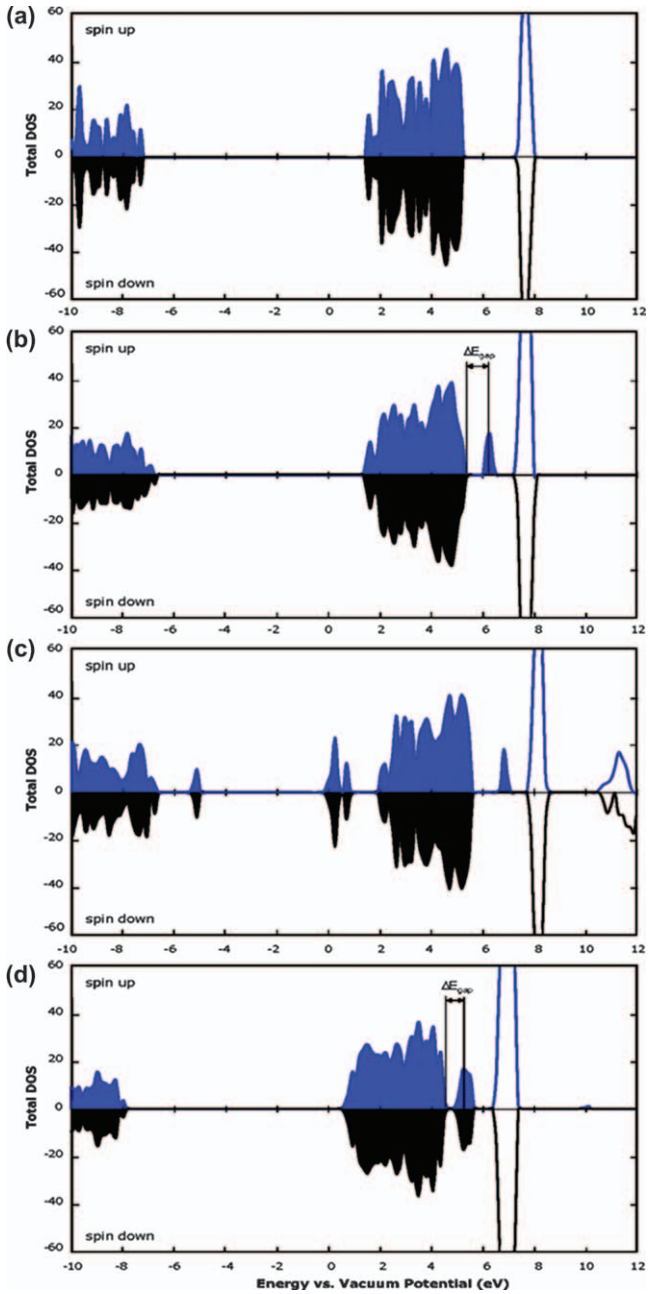


Figure 4.10 Total DOS for (a) clean $CeO_2(111)$, (b) oxygen vacant $CeO_2(111)$, (c) $CeO_2(111)$ with adsorbed $*CH_3$ and *H, and (d) oxygen vacant Pd-substituted $CeO_2(111)$. Occupied states below the Fermi level are highlighted, and the energy axis is referenced to vacuum potential.
(Reprinted with permission from ref. 69 Copyright 2008 American Chemical Society)

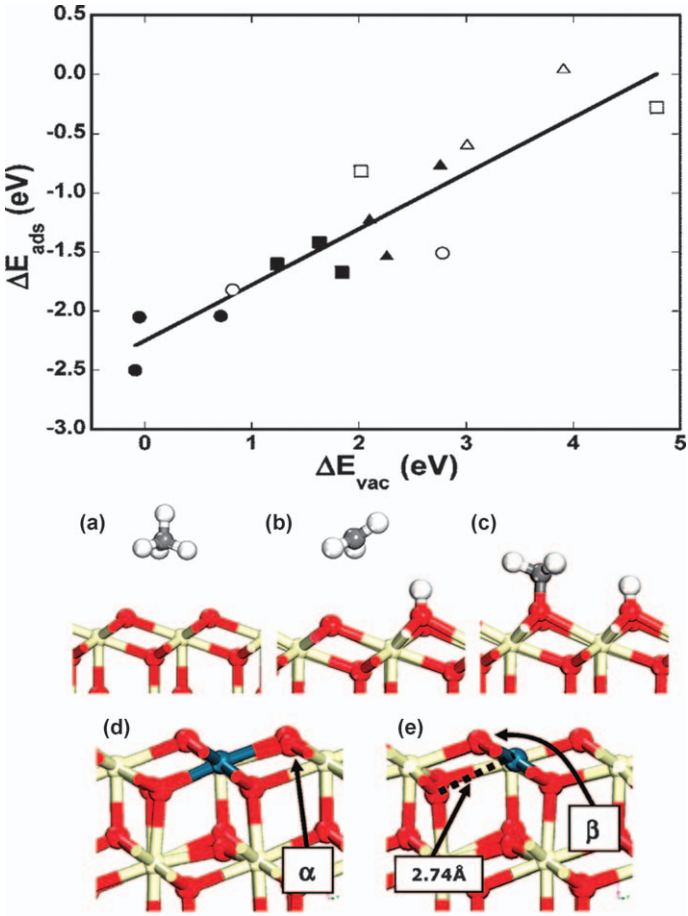


Figure 4.11 Plot demonstrating the correlation between methane adsorption energy, ΔE_{ads}, and oxygen vacancy formation energy, ΔE_{vac}, for Pd-substituted ceria surfaces (●), Zr-substituted surfaces (■), and pure ceria surfaces (▲); open symbols denote pre-reduced surfaces with oxygen vacancies. (a–b) Initial, transition, and final structures for the dissociative adsorption of CH_4 on clean CeO_2(111). (d–e) Structure of Pd-substituted CeO_2(111) before and after oxygen vacancy formation, where α denotes the first oxygen removed and β denotes the second oxygen removed. (Adapted with permission from ref. 69 Copyright 2008 American Chemical Society)

be incorporated to determine the stability of possible Pd/ceria surface morphologies at operating temperatures and pressures. Our laboratory employed such an approach to investigate the stability of single Pd atom states on CeO_2(111), (110), and (100) surfaces.[78] Figure 4.12 contains the resulting phase diagrams showing the temperature and oxygen chemical potential ranges where Pd atoms are stable either: (1) as adsorbed Pd atoms (Pd*), (2) as adsorbed PdO or PdO_2 clusters, or (3) incorporated in Ce lattice positions on either a fully

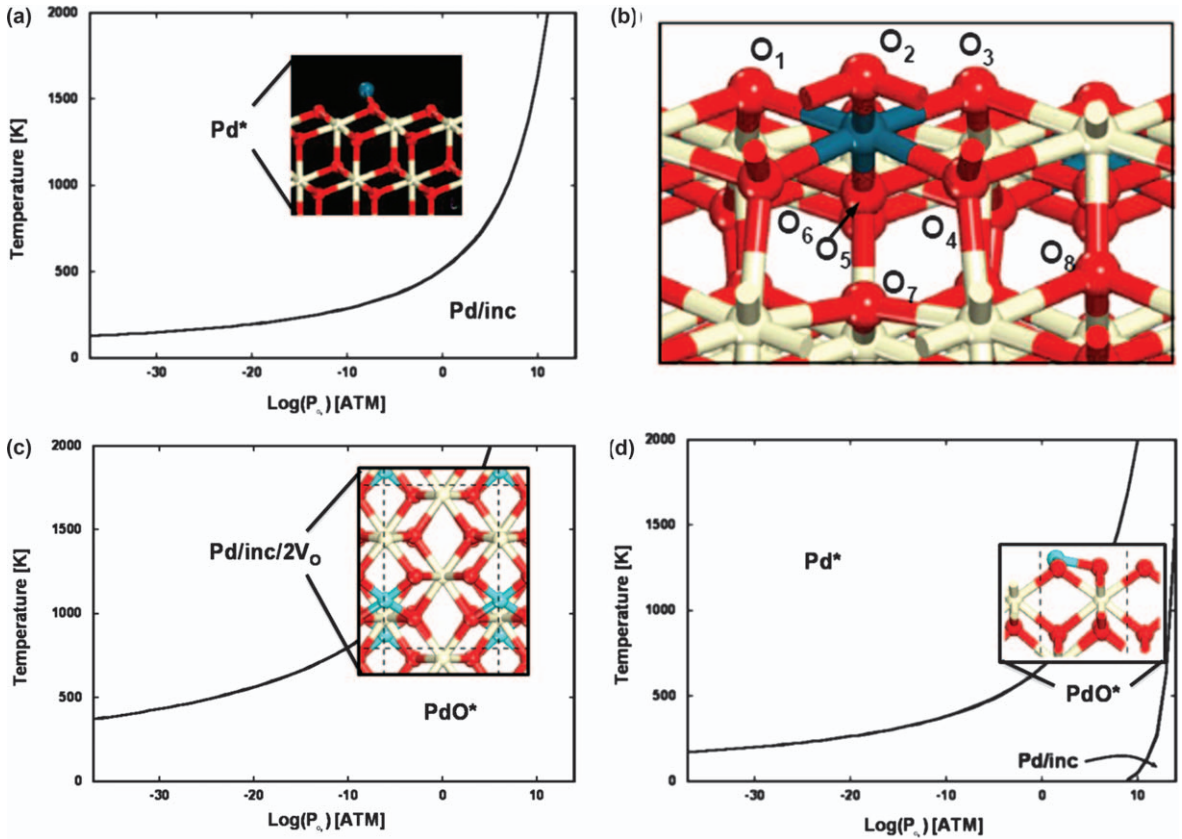


Figure 4.12 *Ab initio* thermodynamic phase diagrams and insets depicting DFT optimized surface structures for Pd/ceria surface morphologies for (a–b) CeO_2(111), (c) CeO_2(110), and (d) CeO_2(100). (b) Octahedral coordination environment formed by oxygen atoms O1–O6, with O7 moving away from the Pd metal center (compared to O8).
(Adapted with permission from ref. 78 Copyright 2009 American Institute of Physics)

oxidized surface (as Pd^{4+}) or an oxygen deficient surface (as Pd^{2+} or Pd^{0}). As seen in the figure, each surface is capable of stabilizing an incorporated Pd atom, and the oxidation state of the incorporated Pd atom varies among ceria terminations. For the $CeO_2(111)$ facet, Pd^{4+} ions are thermodynamically stable when incorporated into the fully oxidized surface under high oxygen pressures or low temperatures. The DFT optimized structure of this surface, shown in Figure 4.12(b), demonstrates that Pd-incorporation is stabilized by a surface reconstruction that provides a favorable octahedral oxygen coordination environment for the d^6 metal center. Both oxygen vacancy formation and methane dissociation over this surface cause a restructuring of the surface, providing a square-planar oxygen coordination that stabilizes the d^8 metal center of Pd^{2+}.

Although these studies demonstrate that Pd-incorporation into the CeO_2 surface may be both stable and active for methane dissociation, optimal performance requires activity for the entire catalytic cycle. The high reducibility of the Pd-incorporated surface could alter the free energy landscape of the reaction mechanism so that re-oxidation of the surface becomes rate limiting. To consider this possibility, *ab initio* thermodynamics approaches were used to derive a free energy diagram that takes into account the impact of varying redox conditions (differing CH_4, O_2 and H_2O chemical potentials) on the reaction energy diagram. Figure 4.13 shows the reaction free energy diagram for methane combustion over $CeO_2(111)$ and $Pd_{incorporated}/CeO_2(111)$ surfaces under catalytic combustion conditions. Pd incorporation reduces the reaction energy for each C–H bond activation step, and therefore leads to a faster overall combustion rate. This allows us to conclude that Pd-incorporation is *stable* under combustion conditions and that it lowers the activation barrier of the rate-determining step, making it *active* toward hydrocarbon oxidation.

4.3.4.4 *Atomistic ReaxFF of Pd/ceria*

Although this previous work has identified a $Pd_xCe_{1-x}O_\delta$ mixed surface oxide as a potentially unique active site, the DFT methods employed are extremely

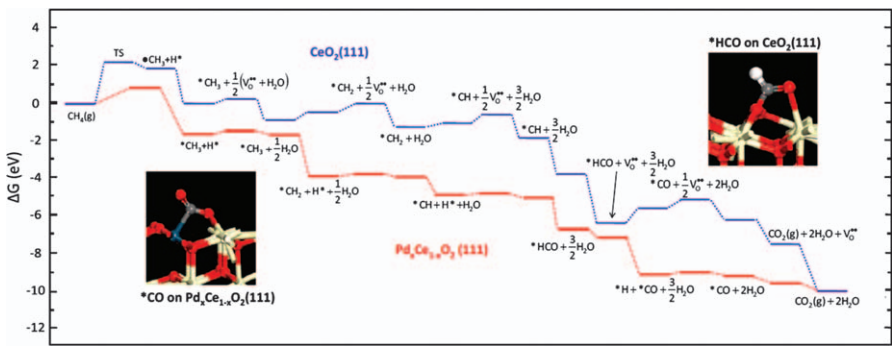


Figure 4.13 Free energy reaction diagrams for complete CH_4 oxidation over (blue) clean $CeO_2(111)$ and (red) Pd-incorporated $Pd_xCe_{1-x}O_\delta$. (Adapted with permission from ref. 89 Copyright 2011 Elsevier)

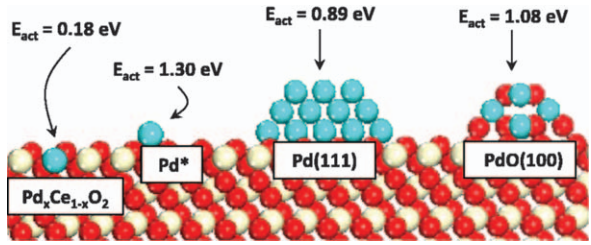


Figure 4.14 DFT-calculated methane activation energies over possible Pd/ceria surface morphologies. Pd-incorporation greatly reduces the C–H bond activation barrier; only highly regular surface structures can be evaluated with DFT.
(Reprinted with permission from ref. 89 Copyright 2011 Elsevier)

limited in the structures that can be considered. Figure 4.14 depicts activation barriers over ordered surface structures that represent limiting cases in which Pd is ideally dispersed throughout the ceria surface, or is entirely contained in pure metal clusters or surface oxides. DFT can only assess these limiting cases separately, and cannot address the combined influence of these separate domains. To model a mixed surface with 2×2 lattice, the surface concentration of Pd atoms is fixed at 25% with Pd atoms incorporated in a specific supercell structure. Lower Pd concentrations, the clustering of Pd dopants, or long range structures involving Pd and O-vacancy clustering are inaccessible with DFT owing to the computational intensity of considering larger supercells. The ReaxFF method will make examination of these possibilities, along with consideration of catalytic reaction chemistry, computationally tractable. As discussed above, the ReaxFF potential is readily suited to handle metal–oxide interfaces under complex hydrocarbon environments. Efforts are currently underway to optimize the ReaxFF potential to treat Ce/Pd/C/O/H systems, and once completed, will be used in tandem with DFT and *ab initio* thermodynamics to determine the combined role of the surface morphologies that were previously only studied as separate limiting cases.

4.4 Conclusions

DFT methods can calculate the ground state electronic structure and energy of a system as a function of nuclear coordinates, allowing one to determine the preferred reaction path by identifying optimal structures on the DFT-calculated potential energy surface (PES). Thus, the utility of DFT lies in its ability accurately to describe the energetics of reactive bond breaking/forming events, which reveals important aspects of the underlying electronic structure, such as oxidation states in atoms, orbital occupancies in molecules, and electronic band structures in solids. *Ab initio* thermodynamics uses statistical mechanics to extend DFT to high temperature and pressure descriptions, which allows one to calculate free energies to assess the stability of the catalyst surface as a function

of chemical environment. Such knowledge is a powerful tool for determining preferred reaction paths over catalytic surfaces, and for building a fundamental understanding of the electronic driving forces behind catalytic mechanisms.

The computational expense of QM limits DFT and *ab initio* thermodynamics to highly idealized models, such as the single crystal surfaces and small clusters considered in the studies above, that only approximate the complex geometry of actual systems. These methods are also limited in their ability to consider dynamic structural transitions at a metal–oxide interface. Although it is difficult to characterize long range and dynamic effects using QM, it becomes possible with the use of empirical force-fields that are based on classical principles and therefore have a relatively low computational expense. The ReaxFF potential is designed to model bond dissociation and formation, and can therefore be implemented in *reactive* molecular dynamics and Monte Carlo simulations that are capable of describing the dynamic nature of a catalyst at longer length and time scales (up to $\sim 10^4$ atoms and ~ 1 ns). Studies utilizing these methods can systematically probe catalytically interesting properties of supported metal–oxide systems that arise from structural irregularities and dynamic effects of surface reconstruction at interfaces.

Together, the computational methods discussed herein allow for a detailed determination of many phenomena influencing the behavior of supported metal–oxide catalysts. The studies considered in this review demonstrate the power of computational methods for assessing the stability and activity of supported metal–oxide catalysts.

References

1. E. Aneggi, C. de Leitenburg, G. Dolcetti and A. Trovarelli, *Catal. Today: Catalytic Control of Diesel Exhaust Emissions*, 2006, **114**, 40–47.
2. R. Burch, P. K. Loader and F. J. Urbano, *Catal. Today*, 1996, **27**, 243–248.
3. L.-H. Xiao, K.-H. Sun, X.-l. Xu and X.-N. Li, *Catal. Commun.*, 2005, **6**, 796–801.
4. S. Colussi, A. Gayen, J. Llorca, C. de Leitenburg, G. Dolcetti and A. Trovarelli, *Ind. Engin. Chem. Res.*, 2012, **51**, 7510–7517.
5. S. Specchia, E. Finocchio, G. Busca, P. Palmisano and V. Specchia, *J. Catal.*, 2009, **263**, 134–145.
6. L. Shi, W. Chu, F. Qu and S. Luo, *Catal. Lett.*, 2007, **113**, 59–64.
7. T. Kuznetsova, V. Sadykov, L. Batuev, E. Moroz, E. Burgina, V. Rogov, V. Kriventsov and D. Kochubey, *J. Nat. Gas Chem.*, 2006, **15**, 149–163.
8. C. Bozo, N. Guilhaume, E. Garbowski and M. Primet, *Catal. Today*, 2000, **59**, 33–45.
9. C. Bozo, N. Guilhaume and J.-M. Herrmann, *J. Catal.*, 2001, **203**, 393–406.
10. M. A. Malecka, L. Kepinski and W. Mista, *Appl. Catal. B: Environ.*, 2007, **74**, 290–298.

11. G. Picasso, M. Gutierrez, M. P. Pina and J. Herguido, *Chem. Engin. J.*, 2007, **126**, 119–130.
12. M. Cargnello, J. J. D. Jaén, J. C. H. Garrido, K. Bakhmutsky, T. Montini, J. J. C. Gámez, R. J. Gorte and P. Fornasiero, *Science*, 2012, **337**, 713–717.
13. G. Jacobs, R. A. Keogh and B. H. Davis, *J. Catal.*, 2007, **245**, 326–337.
14. X. Wang and R. J. Gorte, *Appl. Catal. A: Gen.*, 2002, **224**, 209–218.
15. Q. Zhuang, Y. Qin and L. Chang, *Appl. Catal.*, 1991, **70**, 1–8.
16. E. Nikolla, J. Schwank and S. Linic, *J. Catal.*, 2007, **250**, 85–93.
17. H.-L. Chen, S.-H. Liu and J.-J. Ho, *J. Phys. Chem. B*, 2006, **110**, 14816–14823.
18. Y. Matsumura, W.-J. Shen, Y. Ichihashi and H. Ando, *Catal. Lett.*, 2000, **68**, 181–183.
19. W.-J. S. Y. Matsumura, *Phys. Chem. Chem. Phys.*, 2000, **2**, 1519–1522.
20. A. Yee, S. J. Morrison and H. Idriss, *J. Catal.*, 1999, **186**, 279–295.
21. J. Lahiri, A. Mayernick, S. L. Morrow, B. E. Koel, A. C. T. van Duin, M. J. Janik and M. Batzill, *J. Phys. Chem. C*, 2010, **114**, 5990–5996.
22. H. Idriss, *Platin. Metals Rev.*, 2004, **48**, 105–115.
23. D. Mei, N. A. Deskins, M. Dupuis and Q. Ge, *J. Phys. Chem. C*, 2008, **112**, 4257–4266.
24. D. Mei, N. A. Deskins and M. Dupuis, *Surf. Sci.*, 2007, **601**, 4993–5001.
25. R. Burch, A. Goguet and F. C. Meunier, *Appl. Catal. A: Gen.*, 2011, **409**, 3–12.
26. Y. Chen, H. Wang, R. Burch, C. Hardacre and P. Hu, *Faraday Disc.*, 2011, **152**, 121–133.
27. A. Goguet, F. C. Meunier, D. Tibiletti, J. P. Breen and R. Burch, *J. Phys. Chem. B*, 2004, **108**, 20240–20246.
28. C. M. Y. Yeung, F. Meunier, R. Burch, D. Thompsett and S. C. Tsang, *J. Phys. Chem. B*, 2006, **110**, 8540–8543.
29. G. Jacobs and B. H. Davis, *Appl. Catal. A: Gen.*, 2005, **284**, 31–38.
30. D. Pierre, W. Deng and M. Flytzani-Stephanopoulos, *Topics Catal.*, 2007, **46**, 363–373.
31. X. G. Wang and R.J. Wagner, *J. Catal.*, 2002, **212**, 225–230.
32. X. Wang, J. A. Rodriguez, J. C. Hanson, D. Gamarra, A. Martinez-Arias and M. Fernandez-Garcia, *J. Phys. Chem. B*, 2006, **110**, 428–434.
33. C.-H. Lin, C.-L. Chen and J.-H. Wang, *J. Phys. Chem. C*, 2011, **115**, 18582–18588.
34. R. Burch, *Phys. Chem. Chem. Phys.*, 2006, **8**, 5483–5500.
35. T. Bunlesin, E. S. Putna and R. J. Gorte, *Catal. Lett.*, 1996, **41**, 1–5.
36. M. Boaro, C. de Leitenburg, G. Dolcetti and A. Trovarelli, *J. Catal.*, 2000, **193**, 338–347.
37. D. Teschner, A. Wootsch, O. Pozdnyakova-Tellinger, J. Krohnert, E. M. Vass, M. Havecker, S. Zafeiratos, P. Schnorch, P. C. Jentoft, A. Knop-Gericke and R. Schlogl, *J. Catal.*, 2007, **249**, 318–327.
38. V. Shapovalov and H. Metiu, *J. Catal.*, 2007, **245**, 205–214.
39. S. H. Oh and G. B. Hoflund, *J. Phys. Chem. A*, 2006, **110**, 7609–7613.

40. H. Wang and W. F. Schneider, *Catal. Today*, 2011, **165**, 49–55.
41. R. B. Getman and W. F. Schneider, *ChemCatChem*, 2010, **2**, 1450–1460.
42. W.-J. Zhu, J. Zhang, X.-Q. Gong and G. Lu, *Catal. Today*, 2011, **165**, 19–24.
43. H. W. Jen, G. W. Graham, W. Chun, R. W. McCabe, J. P. Cuif, S. E. Deutsch and O. Touret, *Catal. Today*, 1999, **50**, 309–328.
44. G. W. Graham, H. W. Jen, R. W. McCabe, A. M. Straccia and L. P. Haack, *Catal. Lett*., 2000, **67**, 99–105.
45. E. Aneggi, M. Boaro, C. D. Leitenburg, G. Dolcetti and A. Trovarelli, *J. Alloy. Comp*., 2006, **408–412**, 1096–1102.
46. A. Trovarelli, C. D. Leitenburg, M. Boaro and G. Dolvetti, *Catal. Today*, 1999, **50**, 353–367.
47. M. K. Debe, *Nature*, 2012, **486**, 43–51.
48. S. McIntosh, J. M. Vohs and R. J. Gorte, *Electrochem. Solid-State Lett*., 2003, **6**, A240–A243.
49. S. McIntosh and R. J. Gorte, *Chem. Rev*., 2004, **104**, 4845–4865.
50. R. J. Gorte, J. M. Vohs and S. McIntosh, *Solid State Ion*., 2004, **175**, 1–6.
51. S. An, C. Lu, W. L. Worrell, R. J. Gorte and J. M. Vohs, *Solid State Ion*., 2004, **175**, 135–138.
52. M. Gong, X. Liu, J. Trembly and C. Johnson, *J. Power Sourc*., 2007, **168**, 289–298.
53. J.-H. Wang, M. Liu and M. C. Lin, *Solid State Ion*., 2006, **177**, 939–947.
54. A. Borodziński and G. C. Bond, *Catal. Rev*., 2008, **50**, 379–469.
55. A. Borodziński and G. C. Bond, *Catal. Rev*., 2006, **48**, 91–144.
56. M. Boronat, F. Illas and A. Corma, *J. Phys. Chem. A*, 2009, **113**, 3750–3757.
57. P. Concepción, S. Carrettin and A. Corma, *Appl. Catal. A: Gen*., 2006, **307**, 42–45.
58. N. Khan, S. Shaikhutdinov and H. Freund, *Catal. Lett*., 2006, **108**, 159–164.
59. W. Ludwig, A. Savara, K.-H. Dostert and S. Schauermann, *J. Catal*., 2011, **284**, 148–156.
60. M. W. Tew, M. Janousch, T. Huthwelker and J. A. van Bokhoven, *J. Catal*., 2011, **283**, 45–54.
61. Z.-P. Liu and P. Hu, *J. Am. Chem. Soc*., 2003, **125**, 1958–1967.
62. G. Jacobs, S. Ricote, U. M. Graham, P. M. Patterson and B. H. Davis, *Catal. Today*, 2005, **106**, 259–264.
63. J. H. Holles, R. J. Davis, T. M. Murray and J. M. Howe, *J. Catal*., 2000, **195**, 193–206.
64. J. A. Farmer and C. T. Campbell, *Science*, 2010, **329**, 933–936.
65. G. B. Hoflund, H. A. E. Hagelin, J. F. Weaver and G. N. Salaita, *Appl. Surf. Sci*., 2003, **205**, 102–112.
66. W.-J. Shen and Y. Matsumura, *J. Molec. Catal. A: Chem*., 2000, **153**, 165–168.
67. H. Gabasch, K. Hayek, B. Klötzer, W. Unterberger, E. Kleimenov, D. Teschner, S. Zafeiratos, M. Hävecker, A. Knop-Gericke, R. Schlögl,

B. Aszalos-Kiss and D. Zemlyanov, *J. Phys. Chem. C*, 2007, **111**, 7957–7962.
68. S. Laursen and S. Linic, *J. Phys. Chem. C*, 2009, **113**, 6689–6693.
69. A. D. Mayernick and M. J. Janik, *J. Phys. Chem. C*, 2008, **112**, 14955–14964.
70. H. Borchert, Y. Borchert, V. V. Kaichev, I. P. Prosvirin, G. M. Alikina, A. I. Lukashevich, V. I. Zaikovskii, E. M. Moroz, E. A. Paukshtis, V. I. Bukhtiyarov and V. A. Sadykov, *J. Phys. Chem. B*, 2005, **109**, 20077–20086.
71. Z.-P. Liu, S. J. Jenkins and D. A. King, *Phys. Rev. Lett.*, 2005, **94**, 196102.
72. Z. Yang, G. Luo, Z. Lu, T. K. Woo and K. Hermansson, *J. Phys.: Cond. Matt.*, 2008, **20**, 035210.
73. L. Szabova, M. F. Camellone, M. Huang, V. Matolin and S. Fabris, *J. Chem. Phys.*, 2010, **133**, 234705–234711.
74. M. Nolan, S. Grigoleit, D. C. Sayle, S. C. Parker and G. W. Watson, *Surf. Sci.*, 2005, **576**, 217–229.
75. M. Nolan, S. C. Parker and G. W. Watson, *Surf. Sci.*, 2005, **595**, 223–232.
76. S. Bernal, J. J. Calvino, M. A. Cauqui, J. M. Gatica, C. Larese, J. A. Perez Omil and J. M. Pintado, *Catal. Today*, 1999, **50**, 175–206.
77. B. Wang, D. Weng, X. Wu and J. Fan, *Catal. Today*, 2010, **153**, 111–117.
78. A. D. Mayernick and M. J. Janik, *J. Chem. Phys.*, 2009, **131**, 084701–084712.
79. K. Reuter, D. Frenkel and M. Scheffler, *Phys. Rev. Lett.*, 2004, **93**, 116105.
80. J. F. Weaver, J. A. Hinojosa Jr, C. Hakanoglu, A. Antony, J. M. Hawkins and A. Asthagiri, *Catal. Today*, 2011, **160**, 213–227.
81. H. Zhang, J. Gromek, G. Fernando, H. Marcus and S. Boorse, *J. Phase Equilib. Diffus.*, 2002, **23**, 246–248.
82. J. Rogal, K. Reuter and M. Scheffler, *Phys. Rev. B*, 2004, **69**, 075421.
83. J. Klikovits, E. Napetschnig, M. Schmid, N. Seriani, O. Dubay, G. Kresse and P. Varga, *Phys. Rev. B*, 2007, **76**, 045405.
84. J. A. Rodriguez, X. Wang, J. C. Hanson, G. Liu, A. Iglesias-Juez and M. Fernandez-Garcia, *J. Chem. Phys.*, 2003, **119**, 5659–5669.
85. X. Wang, J. A. Rodriguez, J. C. Hanson, D. Gamarra, A. Martinez-Arias and M. Fernandez-Garcia, *J. Phys. Chem. B*, 2005, **109**, 19595–19603.
86. G. Zhou, P. R. Shah, T. Kim, P. Fornasiero and R. J. Gorte, *Catal. Today*, 2007, **123**, 86–93.
87. S. Colussi, A. Gayen, M. F. Camellone, M. Boaro, J. Llorca, S. Fabris and A. Trovarelli, *Ange. Chem. Int.*, 2009, **48**, 8481–8484.
88. L. M. Misch, J. A. Kurzman, A. R. Derk, Y.-I. Kim, R. Seshadri, H. Metiu, E. W. McFarland and G. D. Stucky, *Chem. Mat.*, 2011, **23**, 5432–5439.
89. A. D. Mayernick and M. J. Janik, *J. Catal.*, 2011, **278**, 16–25.
90. M. L. Drummond, B. G. Sumpter, W. A. Shelton and J. Z. Larese, *J. Phys. Chem. C*, 2007, **111**, 966–976.
91. R. B. Getman, Y. Xu and W. F. Schneider, *J. Phys. Chem. C*, 2008, **112**, 9559–9572.

92. H. Wang and W. F. Schneider, *J. Chem. Phys.*, 2007, **127**, 064706.
93. D. R. Mullins and S. H. Overbury, *Surf. Sci.*, 2002, **511**, L293–L297.
94. J. K. Norskov, T. Bligaard, J. Rossmeisl and C. H. Christensen, *Nat. Chem.*, 2009, **1**, 37–46.
95. P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864–B871.
96. W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
97. W. Kohn, A. D. Becke and R. G. Parr, *J. Phys. Chem.*, 1996, **100**, 12974–12980.
98. D. S. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction*, John Wiley & Sons, Inc., 2009; pp 1–233.
99. C. Stampfl, *Catal. Today*, 2005, **105**, 17–35.
100. K. Reuter and M. Scheffler, *Phys. Rev. B*, 2001, **65**, 035406.
101. K. Reuter and M. Scheffler, *Phys. Rev. B*, 2003, **68**, 045407.
102. K. Reuter and M. Scheffler, *Phys. Rev. Lett.*, 2003, **90**, 046103.
103. A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, *J. Phys. Chem. A*, 2001, **105**, 9396–9409.
104. M. Moseler, M. Walter, B. Yoon, U. Landman, V. Habibpour, C. Harding, S. Kunz and U. Heiz, *J. Am. Chem. Soc.*, 2012, **134**, 7690–7699.
105. B. Yoon, H. Häkkinen, U. Landman, A. S. Wörz, J.-M. Antonietti, S. Abbet, K. Judai and U. Heiz, *Science*, 2005, **307**, 403–407.
106. L. M. Molina and B. Hammer, *Phys. Rev. Lett.*, 2003, **90**, 206102.
107. L. M. Molina, M. D. Rasmussen and B. Hammer, *J. Chem. Phys.*, 2004, **120**, 7673–7680.
108. G. P. Petrova, G. N. Vayssilov and N. Rösch, *J. Phys. Chem. C*, 2007, **111**, 14484–14492.
109. P. S. Petkov, G. P. Petrova, G. N. Vayssilov and N. Rösch, *J. Phys. Chem. C*, 2010, **114**, 8500–8506.
110. C. H. Hu, C. Chizallet, C. Mager-Maury, M. Corral-Valero, P. Sautet, H. Toulhoat and P. Raybaud, *J. Catal.*, 2010, **274**, 99–110.
111. C. Mager-Maury, G. Bonnard, C. Chizallet, P. Sautet and P. Raybaud, *ChemCatChem*, 2011, **3**, 200–207.
112. R. Ferrando, G. Rossi, F. Nita, G. Barcaro and A. Fortunelli, *ACS Nano*, 2008, **2**, 1849–1856.
113. G. Barcaro and A. Fortunelli, *J. Phys. Chem. C*, 2007, **111**, 11384–11389.
114. G. Barcaro, A. Fortunelli, G. Granozzi and F. Sedona, *J. Phys. Chem. C*, 2009, **113**, 1143–1146.
115. J. A. Rodriguez, P. Liu, J. Hrbek, J. Evans and M. Pérez, *Ange. Chem. Int. Ed.*, 2007, **46**, 1329–1332.
116. J. A. Rodriguez, S. Ma, P. Liu, J. Hrbek, J. Evans and M. Pérez, *Science*, 2007, **318**, 1757–1760.
117. A. Bruix, J. A. Rodriguez, P. J. Ramírez, S. D. Senanayake, J. Evans, J. B. Park, D. Stacchiola, P. Liu, J. Hrbek and F. Illas, *J. Am. Chem. Soc.*, 2012, **134**, 8968–8974.
118. C. Zhang, A. Michaelides and S. J. Jenkins, *Phys. Chem. Chem. Phys.*, 2011, **13**, 22–33.

119. P. Liu and J. A. Rodriguez, *J. Chem. Phys.*, 2007, **126**, 164705–164708.
120. J. A. Rodriguez and F. Illas, *Phys. Chem. Chem. Phys.*, 2012, **14**, 427–438.
121. X. Wang, J. A. Rodriguez, J. C. Hanson, M. Perez and J. Evans, *J. Chem. Phys.*, 2005, **123**, 221101–221105.
122. R. F. W. Bader, *Acc. Chem. Res.*, 1985, **18**, 9–15.
123. Y. Chen, J. Cheng, P. Hu and H. Wang, *Surf. Sci.*, 2008, **602**, 2828–2834.
124. S. C. Ammal and A. Heyden, *J. Phys. Chem. C*, 2011, **115**, 19246–19259.
125. S. C. Ammal and A. Heyden, *J. Phys. Chem. C*, 2011, **116**, 1624–1624.
126. T. Ishida, N. Kinoshita, H. Okatsu, T. Akita, T. Takei and M. Haruta, *Ange. Chem. Int. Ed.*, 2008, **47**, 9265–9268.
127. Z. Zhou, S. Kooi, M. Flytzani-Stephanopoulos and H. Saltsburg, *Adv. Funct. Mat.*, 2008, **18**, 2801–2807.
128. S. Aranifard, S. C. Ammal and A. Heyden, *J. Phys. Chem. C*, 2012, **116**, 9029–9042.
129. W. A. Goddard, J. E. Mueller and A. C. T. van Duin, *J. Phys. Chem. C*, 2010, **114**, 4939–4949.
130. J. E. Mueller, A. C. T. van Duin and W. A. Goddard, *J. Phys. Chem. C*, 2010, **114**, 5675–5685.
131. K. Chenoweth, A. C. T. van Duin, P. Persson, M.-J. Cheng, J. Oxgaard and I. I. I. W. A. Goddard, *J. Phys. Chem. C*, 2008, **112**, 14645–14654.
132. K. Chenoweth, A. C. T. van Duin and W. A. Goddard III, *Ange. Chem. Int.*, 2009, **48**, 7630–7634.
133. N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087–1092.
134. P. A. Deshpande, M. S. Hegde and G. Madras, *AIChE J.*, 2010, **56**, 1315–1324.

CHAPTER 5

Computing Accurate Net Atomic Charges, Atomic Spin Moments, and Effective Bond Orders in Complex Materials

THOMAS A. MANZ*† AND DAVID S. SHOLL

School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, 311 Ferst Drive NW, Atlanta, GA 30332-0100
*Email: thomasamanz@gmail.com

5.1 Introduction

There are two types of methods for computing net atomic charges (NACs) and effective bond orders (EBOs). The first type computes these quantities using computationally efficient semi-empirical methods. Some noteworthy methods of this type for computing NACs include electronegativity and charge equalization methods, the split charge equilibration method, and charge-optimized many body (COMB) potentials.[1–6] EBOs can be estimated from semi-empirical bond distance to bond order correlations.[7–9] The second type computes NACs, atomic spin moments (ASMs), and EBOs from detailed quantum chemistry calculations such as density functional theory (DFT) or coupled-cluster theory. Some noteworthy methods of this type include: (a) the natural population

†Present address: Department of Chemical Engineering, New Mexico State University, Las Cruces, NM 88003-8001.

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

analysis and natural bond orbital (NPA/NBO) method, which projects the electron density onto natural atomic and bond orbitals;[10–12] (b) Bader's quantum chemical topology (QCT) which divides space into non-overlapping atomic compartments whose surfaces are perpendicular to $\nabla\rho(\vec{r})$;[13,14] and (c) several atoms-in-molecule (AIM) methods that partition $\rho(\vec{r})$ into overlapping atoms,[15–21] including the density derived electrostatic and chemical (DDEC) method that simultaneously optimizes atomic electron distributions to reproduce the chemical states of atoms in a material and the electrostatic potential, $V(\vec{r})$, surrounding its electron distribution.[18–20] The QCT and DDEC methods can be applied to both periodic and non-periodic materials. The NBO method was recently extended to periodic materials, but has difficulty treating metallic conductors.[67] Bader NACs are chemically meaningful and have been used to study chemical properties of a wide range of materials.[13,14,22] Bader NACs are not well suited for constructing force-fields of porous materials, however, because they do not accurately reproduce $V(\vec{r})$ surrounding a material.[18] The NPA/NBO method is especially useful for studying chemical bonding in molecular systems.[10–12] Excellent reviews of the NPA/NBO and Bader QCT methods have been published.[12,14] Because of its more recent introduction, the DDEC method has been less extensively reviewed and will be the main focus of this chapter.

Our earlier reports have demonstrated the use of the DDEC approach for a wide range of materials.[18–20,23] These include isolated molecules treated with localized basis sets, dense solids treated with plane wave DFT, surfaces of dense solids, crystalline porous materials, and 1-, 2-, and 3-dimensional (D) materials with periodic boundary conditions. These reports also give detailed derivations of the DDEC methods and benchmark these methods against related approaches for assigning NACs and ASMs.

The main purpose of this chapter is to give a practical guide for computing NACs and ASMs using the DDEC method. All calculations in the chapter using DDEC population analysis were performed using the program available at ddec.sourceforge.net. Sections 5.2.1 and 5.2.2 give a concise overview of the equations that define the DDEC approach for NACs and ASMs, respectively. The remainder of this section contains a series of examples to illustrate the use of these methods. Section 5.2.3 shows how NACs and ASMs were determined for the anti-ferromagnetic solid Fe_2SiO_4 using VASP. Section 5.2.4 describes two examples calculated using GAUSSIAN, a small molecule (ozone) and a periodic BN sheet. The final example in Section 5.2.5 illustrates using the DDEC method to assess non-collinear magnetism in a molecular complex. To make it easier for the reader to reproduce these results, input files are given for selected examples. Section 5.3 describes how these NACs can be used to construct force-fields for atomistic simulations, with a focus on the electrostatic potential created by an electron distribution. We consider CO_2 adsorption in metal–organic frameworks as an example.

Section 5.4 turns to a somewhat different topic, namely the calculation of EBOs. We describe new approaches for computing EBOs that are theoretically applicable to materials with periodic boundary conditions and non-collinear magnetism. We summarize remaining work to implement these new strategies for computing EBOs.

The examples used in this chapter were calculated using VASP and GAUSSIAN, but DDEC methods can be applied to the output from a wide range of other quantum chemistry packages. In section 5.5, we briefly comment on this capability.

5.2 Net Atomic Charges and Atomic Spin Moments

5.2.1 The Charge Partitioning Functional

All atoms-in-molecules (AIM) methods partition the total electron density $\rho(\vec{r})$ into a set of components $\{\rho_A(\vec{r}_A)\}$ normally associated with the system's atoms. The system's atoms $\{A\}$ are located at positions $\{\vec{R}_A\}$, in a reference unit cell, **U**. For a non-periodic system (*e.g.* a molecule), **U** is any parallelpiped enclosing the entire electron distribution. The reference unit cell has $k_1 = k_2 = k_3 = 0$ and summation over A means summation over all atoms in this unit cell. For a periodic direction, k_i ranges over all integers with the associated lattice vector $\vec{v}_i$. For a non-periodic direction, $k_i = 0$ and $\vec{v}_i$ is the corresponding edge of **U**. Using this notation, the vector and distance relative to atom A are given by

$$\vec{r}_A = \vec{r} - k_1\vec{v}_1 - k_2\vec{v}_2 - k_3\vec{v}_3 - \vec{R}_A \tag{5.1}$$

and $r_A = |\vec{r}_A|$.[18]

Many AIM methods use a set of spherically symmetric atomic weights $\{w_A(r_A)\}$ to partition $\rho(\vec{r})$ into overlapping atoms:[15–18,20]

$$\rho_A(\vec{r}_A) = w_A(r_A)\rho(\vec{r})/W(\vec{r}) \tag{5.2}$$

where

$$W(\vec{r}) = \sum_{k,A} w_A(r_A) \tag{5.3}$$

$$\rho(\vec{r}) = \sum_{k,A} \rho_A(\vec{r}_A) \tag{5.4}$$

$$\rho_A(\vec{r}_A) \geq 0. \tag{5.5}$$

The number of electrons, n_A, and net atomic charge, q_A, are obtained by integrating $\rho_A(\vec{r}_A)$:

$$n_A = \oint \rho_A(\vec{r}_A) d^3\vec{r}_A = z_A - q_A \tag{5.6}$$

where z_A is the atomic number of atom A. The use of spherically symmetric atomic weights is desirable to give NACs that approximately reproduce $V(\vec{r})$.[18,20] The atomic weight factors are expressed as

$$w_A(r_A) = F_A\left(r_A, \left\{\rho_A^{\text{avg}}(r_A)\right\}\right) \tag{5.7}$$

where different choices for the charge partitioning functional, F_A, produce different AIM methods.[15–18,20]

We now summarize the DDEC/c3 method, which is the most accurate and widely applicable charge partitioning functional of this type developed to date.[20] The DDEC/c3 charge partitioning functional has the form

$$F_A^{c3}(r_A) = \lambda_A G_A(r_A) \tag{5.8}$$

where the constant λ_A is chosen to be the smallest number ≥ 1 such that the number of valence electrons assigned to atom A is non-negative. To ensure the NACs are both chemically meaningful and approximately reproduce $V(\vec{r})$, $G_A(r_A)$ is optimized to resemble the geometric average[20]

$$\sigma_A(r_A) = \left(Y_A^{avg}(r_A)\right)^{3/14}\left(\rho_A^{avg}(r_A)\right)^{11/14} \tag{5.9}$$

between the spherical average of $\rho_A(\vec{r}_A)$ and the conditioned reference density

$$Y_A^{avg}(r_A) = \rho_A^{ref}(r_A, n_A)\left\langle \rho(\vec{r})/\rho^{ref}(\vec{r})\right\rangle_{r_A} \tag{5.10}$$

where

$$\rho^{ref}(\vec{r}) = \sum_{k,A} \rho_A^{ref}(r_A, n_A). \tag{5.11}$$

$\langle\,\rangle_{r_A}$ denotes the spherical average corresponding to radius r_A, and $\rho_A^{ref}(r_A, n_A)$ is the (unconditioned) reference density of an isolated and charge-compensated ion of the same element having the same number of electrons, n_A, as it does in the material of interest.[18] To ensure that buried atoms do not become too diffuse, the buried tail of $F_A^{c3}(r_A)$ is constrained to decay exponentially:[20]

$$\phi_A(r_A) = \frac{dG_A(r_A)}{dr_A} + \eta_A(r_A)G_A(r_A) \leq 0. \tag{5.12}$$

As illustrated in Figure 5.1, this makes $\rho_A(\vec{r}_A)$ decay approximately exponentially with increasing r_A. The function $\eta_A(r_A)$ transitions smoothly from a value of zero where $\rho_A(\vec{r}_A) \approx \rho(\vec{r})$ to a limiting value $b = 1.75$ bohr^{-1} where $\rho_A(\vec{r}_A) \ll \rho(\vec{r})$. This can be accomplished by defining $\eta_A(r_A)$ as:[20]

$$\eta_A(r_A) = b\left(1 - (\tau_A(r_A))^2\right) \tag{5.13}$$

$$\tau_A(r_A) = Y_A^{avg}(r_A)\left\langle Y(\vec{r})^{-1/2}\right\rangle_{r_A} \Big/ \left\langle Y(\vec{r})^{1/2}\right\rangle_{r_A}. \tag{5.14}$$

$G_A(r_A)$ is also constrained to give the same integral over space as $\sigma_A(r_A)$:

$$\varphi_A = \int_0^\infty (G_A(r_A) - \sigma_A(r_A))4\pi(r_A)^2 dr_A = 0. \tag{5.15}$$

The $G_A(r_A)$ that satisfies all of these properties is found by solving

$$\partial H(G_A(r_A))/\partial G_A(r_A) = 0 \tag{5.16}$$

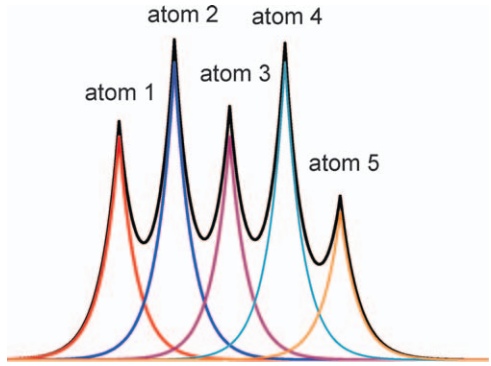


Figure 5.1 Schematic of DDEC/c3 electron density partitioning. Atoms 2, 3, and 4 are "buried" while atoms 1 and 5 are not. Each atom has an exponentially decaying tail accurately describing the atom–atom overlaps.

where

$$H(G_A(r_A)) = \frac{1}{2}\int_0^\infty \frac{(G_A(r_A)-\sigma_A(r_A))^2}{\sqrt{\sigma_A(r_A)}} 4\pi(r_A)^2 dr_A + \int_0^\infty \Gamma_A(r_A)\phi_A(r_A)4\pi(r_A)^2 dr_A - \Phi_A\varphi_A. \quad (5.17)$$

In eqn (5.17), $\Gamma_A(r_A)$ and Φ_A are Lagrange multipliers enforcing constraints (5.12) and (5.15). All of this charge partitioning functional's parameters were theoretically derived,[20] so its results can be regarded as non-empirical.

Solution is by an iterative process.[20] The neutral atom reference densities are used as the initial guess for $w_A(r_A)$. The $w_A(r_A)$ estimate is improved by a series of charge cycles. In each charge cycle, the current $\{w_A(r_A)\}$ estimate is used to compute estimates of $\{\rho_A^{avg}(r_A)\}$ and $\{n_A\}$ which are subsequently used to update the $\{w_A(r_A)\}$ estimate. This process is repeated until a self-consistent solution is reached.

5.2.2 The Spin Partitioning Functional

In this section, we describe how to partition the spin density distribution between atoms in a material. This partitioning is important for understanding the properties of magnetic materials. The spin magnetization density, $\vec{m}(\vec{r})$, is computed by summing the spins of all electrons at position $\vec{r}$ and dividing by the spin magnitude of an individual electron:[19]

$$\vec{m}(\vec{r}) = 2\langle\Psi|\sum_{j=1}^{N}\left(\vec{s}(j)\delta^{dirac}\left(\vec{r}-\vec{e}_j\right)\right)|\Psi\rangle \quad (5.18)$$

$\Psi\left(\{\vec{e}_j\}\right)$ is the multi-electronic wavefunction, $\{\vec{e}_j\}$ are the spatial coordinates of the electrons, and $\vec{s}(j)$ is the operator for measuring the spin of the j^{th}

electron. The factor of 2 occurs because the magnitude of an electron's spin is one-half. The spin density projected onto some measurement direction $\hat{h}$ is

$$\rho\left(\vec{r},\hat{h}\right)=\left(\rho(\vec{r})+\vec{m}(\vec{r})\bullet\hat{h}\right)\Big/2, \tag{5.19}$$

where $\rho(\vec{r})$ and $\vec{m}(\vec{r})$ are expressed in units of electrons per unit volume.[24] A system has collinear magnetism if $\hat{m}(\vec{r})$ is parallel to a global magnetization axis, $\hat{h}_{\text{global}}$. Otherwise, a system has non-collinear magnetism. Collinear magnetism has two independent electron density components, $\rho^{\alpha}(\vec{r})=\rho(\vec{r},\hat{h}_{\text{global}})$ and $\rho^{\beta}(\vec{r})=\rho(\vec{r},-\hat{h}_{\text{global}})$. Non-collinear magnetism has four independent electron density components, $\rho(\vec{r})$ and $\vec{m}(\vec{r})$.

Spin partitioning is the process of assigning a spin magnetization density, $\vec{m}_{\text{A}}(\vec{r}_{\text{A}})$, to each atom in the system such that

$$\vec{m}(\vec{r})=\vec{\Delta}(\vec{r})+\sum_{k,\text{A}}\vec{m}_{\text{A}}(\vec{r}_{\text{A}}). \tag{5.20}$$

$$\vec{\Delta}(\vec{r})=\vec{0} \tag{5.21}$$

$$\kappa_{\text{A}}(\vec{r}_{\text{A}})=\rho_{\text{A}}(\vec{r}_{\text{A}})-m_{\text{A}}(\vec{r}_{\text{A}})\geq 0 \tag{5.22}$$

Eqns (5.20) and (5.21) ensure that the $\{\vec{m}_{\text{A}}(\vec{r}_{\text{A}})\}$ sum to $\vec{m}(\vec{r})$ at each point in space. Constraint (5.22) is required to ensure that the results are chemically reasonable.[19] For example, one cannot assign more spin up electrons to a particular atom than the total number of electrons assigned to that atom. Since $\rho_{\text{A}}(\vec{r}_{\text{A}})$ decays approximately exponentially with increasing r_{A}, constraint (5.22) ensures that $m_{\text{A}}(\vec{r}_{\text{A}})$ is localized in the region of space occupied by atom A.[19] In a previous publication,[19] we showed that excellent results are obtained when $\{\vec{m}_{\text{A}}(\vec{r}_{\text{A}})\}$ minimizes the functional

$$\begin{aligned}H^{\text{spin}}=\sum_{\text{A}}\Bigg(&\oint_{\omega}\oint\rho_{\text{A}}\left(\vec{r}_{\text{A}},\hat{h}\right)\left(\ln\left(\frac{\rho_{\text{A}}\left(\vec{r}_{\text{A}},\hat{h}\right)}{w_{\text{A}}^{\text{spin}}\left(\vec{r}_{\text{A}},\hat{h}\right)}\right)\right)d^3\vec{r}_{\text{A}}d^2\omega\\&-\oint\nu_{\text{A}}(\vec{r}_{\text{A}})\kappa_{\text{A}}(\vec{r}_{\text{A}})d^3\vec{r}_{\text{A}}\Bigg)+\int_{U}\vec{\Lambda}(\vec{r})\bullet\vec{\Delta}(\vec{r})d^3\vec{r}\end{aligned} \tag{5.23}$$

where

$$w_{\text{A}}^{\text{spin}}\left(\vec{r}_{\text{A}},\hat{h}\right)=\sqrt{\rho_{\text{A}}^{0}\left(\vec{r}_{\text{A}},\hat{h}\right)\rho_{\text{A}}^{\text{avg}}\left(r_{\text{A}},\hat{h}\right)} \tag{5.24}$$

$$\rho_{\text{A}}^{\text{avg}}\left(r_{\text{A}},\hat{h}\right)=\frac{1}{4\pi(r_{\text{A}})^2}\oint\rho_{\text{A}}\left(\vec{r}'_{\text{A}},\hat{h}\right)\delta^{\text{dirac}}(r_{\text{A}}-r'_{\text{A}})d^3\vec{r}'_{\text{A}} \tag{5.25}$$

$$\rho_A\left(\vec{r}_A,\hat{h}\right)=\left(\rho_A(\vec{r}_A)+\vec{m}_A(\vec{r}_A)\bullet\hat{h}\right)\Big/2 \tag{5.26}$$

$$\rho_A^0\left(\vec{r}_A,\hat{h}\right)=\left(\rho_A(\vec{r}_A)+\vec{m}_A^0(\vec{r}_A)\bullet\hat{h}\right)\Big/2 \tag{5.27}$$

$$\vec{m}_A^0(\vec{r}_A)=\vec{m}(\vec{r})\rho_A(\vec{r}_A)/\rho(\vec{r}). \tag{5.28}$$

$\vec{\Lambda}(\vec{r})$ and $\nu_A(\vec{r}_A)\geq 0$ and are Lagrange multipliers enforcing constraints (5.21) and (5.22); ω is the unit sphere surface comprised of all possible endpoints for $\hat{h}$, and integration over ω means integration over all possible choices for $\hat{h}$.[19] This spin partitioning functional has a unique minimum that is found by an iterative solution algorithm.[19]

5.2.3 Example using VASP Software

In this section, we give an example of using the method for an extended solid with VASP, a widely used code for plane wave DFT calculations.[25,26] The geometry of the anti-ferromagnetic spinel phase of crystalline Fe_2SiO_4 was optimized using the PBE functional. The following VASP POSCAR specified the initial geometry estimate:

======VASP POSCAR======

```
Fe2SiO4 Example
  8.391
    0.000000    0.500000    0.500000
    0.500000    0.000000    0.500000
    0.500000    0.500000    0.000000
    8   2   4
Selective Dynamics
Direct
0.236 0.236 0.236 T T T
-0.208 0.236 0.236 T T T
0.236 -0.208 0.236 T T T
0.236 0.236 -0.208 T T T
-0.236 -0.236 -0.236 T T T
0.208 -0.236 -0.236 T T T
-0.236 0.208 -0.236 T T T
-0.236 -0.236 0.208 T T T
-0.125 -0.125 -0.125 T T T
0.125 0.125 0.125 T T T
0.5 0.5 0.5 T T T
0.0 0.5 0.5 T T T
0.5 0.0 0.5 T T T
0.5 0.5 0.0 T T T
```

The flags T T T indicate that all atomic positions are optimized. The *k*-point mesh was specified using the following file:

```
= = = = = = VASP KPOINTS = = = = = =

Automatic mesh
0
Monkhorst-Pack
4  4  4
0  0  0
```

The *k*-point mesh is normally chosen so the product of *k*-points and unit cell volume is $\geq$4000 Å^3. In this case, 4×4×4 *k*-points multiplied by a unit cell volume of $\sim$137 Å^3 gives the product 4×4×4×137 Å^3 = $\sim$8700 Å^3. The VASP INCAR file was set up in the usual way, including the keywords: (a) ISYM = 0 (turn off symmetry), (b) ISIF = 3 (optimize ionic positions and unit cell volume and shape), (c) ENCUT = 400.00 eV (plane-wave cutoff value), (d) NSW = 200 (maximum number of ionic steps), (e) IBRION = 1 or 2 (geometry optimization algorithm), (f) ISPIN = 2 (perform spin-polarized calculation), (g) MAGMOM = 0 0 0 0 0 0 0 0 0 0 0 3.33 −3.33 3.33 −3.33 (initial guesses for the atomic magnetic moments), (h) smearing parameters were specified using the ISMEAR and SIGMA keywords, and (i) if desired the wavefunction and charge files are saved using LCHARG = .TRUE. and LWAVE = .TRUE. The INCAR, KPOINTS, POSCAR, and POTCAR files are the complete set of input files for this geometry optimization. Optimizing the geometric parameters until the forces on every atom were $\leq$0.03 eV Å^{-1} gave the lattice vectors a = 5.793 Å, b = 5.806 Å, c = 5.793 Å, $\alpha = 59.9^\circ$, $\beta = 59.9^\circ$, $\gamma = 59.9^\circ$.

After the geometry was optimized, a second calculation was performed to generate the files necessary for atomic population analysis. These files must be generated in a calculation that holds the geometry constant. To do this, the CONTCAR file containing the optimized geometry was renamed POSCAR. If the wavefunction was stored from geometry optimization, this can be read by placing ISTART = 2 in the INCAR file. The following keywords were placed in the INCAR to generate the required charge density files: (a) PREC = Accurate, (b) LCHARG = .TRUE., and (c) LAECHG = .TRUE. The INCAR contained the keyword NSW = 0 to indicate that no ionic relaxation is performed in this calculation. Other parameters in the INCAR, such as the smearing parameters, initial guesses for magnetic moments, energy cutoffs, *etc*., were specified in the usual manner. If desired, the electrostatic potential can be output to the LOCPOT file by including LVTOT = .TRUE. in the INCAR. After this VASP calculation was finished, the AECCAR0 (all-electron core density), AECCAR2 (all-electron valence density), CHG (pseudo-valence density and magnetization density), and POTCAR files were used for subsequent charge partitioning. As described in the following section, the LOCPOT file can be used to assess the accuracy of point charge models for reproducing $V(\vec{r})$ in the material's pores.

Charge partitioning was performed by running the DDEC program (ddec.source.forge.net) on these VASP output files. The type and version of input file

(*e.g.* VASP 5) and the net unit cell charge were specified as inputs when running the DDEC program. The DDEC program generated an output file containing a running summary of the calculation's progress. Before charge partitioning, the DDEC program always checks to make sure that all valence and core electrons are accounted for and the input file grids are adequate. Typically, no problems are encountered at this stage if the above procedures are followed. Next, the program used an iterative process to solve for the atomic core electron distributions.[19] Then an iterative process solved for the DDEC/c3 charge distributions and NACs.[20] The file DDEC_net_atomic_charges.xyz was generated when charge partitioning was completed. This file contained the geometry, NACs, atomic dipoles and quadrupoles, a summary of calculation details, and a least squares regression of each atom's tail density to an exponential decay function. Next, the program performed spin partitioning and generated the file DDEC_atomic_spin_moments.xyz. This file contained the geometry, ASMs, total spin moment of the unit cell, and a summary of calculation details. The DDEC program was finished when spin partitioning was completed. Table 5.1 lists key results for this system: (a) XYZ coordinates in Å, (b) NACs, (c) atomic dipole magnitude, μ, in atomic units, and (c) ASMs. The ASM values show that the number of spin up and spin down electrons differs by ± 3.6 electrons for each Fe atom. The atomic dipole moments are negligible on the Fe and Si atoms, and small on the O atoms. The small differences in NAC, μ, and ASM values for crystallographically equivalent atoms (*e.g.* ASM values of -3.644 *vs.* -3.636 for two equivalent Fe atoms) are not significant and arose from the finite grid spacing used to integrate the electron and spin densities.

The .xyz files containing the NACs and ASMs were visualized using the Jmol program (jmol.sourceforge.net). Jmol has commands that allow (a) the atoms to be colored according to the NAC or ASM value and (b) the numeric NAC or ASM value to be printed next to each atom as a label. These were used in the

Table 5.1 DDEC results for the anti-ferromagnetic spinel phase of crystalline Fe_2SiO_4. Cartesian lattice vectors: (0.009, 4.093, 4.099), (4.106, 0.003, 4.105), (4.100, 4.093, 0.009).

Atom	*X*	*Y*	*Z*	*NAC*	μ	*ASM*
O	1.976	1.988	1.977	−1.164	0.048	−0.031
O	1.982	4.148	4.176	−1.158	0.047	0.031
O	4.185	1.998	4.186	−1.158	0.046	−0.027
O	4.177	4.146	1.981	−1.158	0.047	0.029
O	6.239	6.202	6.237	−1.164	0.048	−0.031
O	6.233	4.042	4.038	−1.158	0.047	0.031
O	4.029	6.191	4.028	−1.158	0.046	−0.027
O	4.038	4.043	6.233	−1.158	0.047	0.029
Si	7.188	7.165	7.186	2.208	0.001	0.000
Si	1.027	1.024	1.027	2.208	0.001	0.000
Fe	4.107	4.095	4.107	1.225	0.000	3.659
Fe	4.112	6.141	6.156	1.213	0.000	−3.644
Fe	2.054	4.093	2.054	1.211	0.000	3.617
Fe	6.157	6.141	4.111	1.209	0.000	−3.636

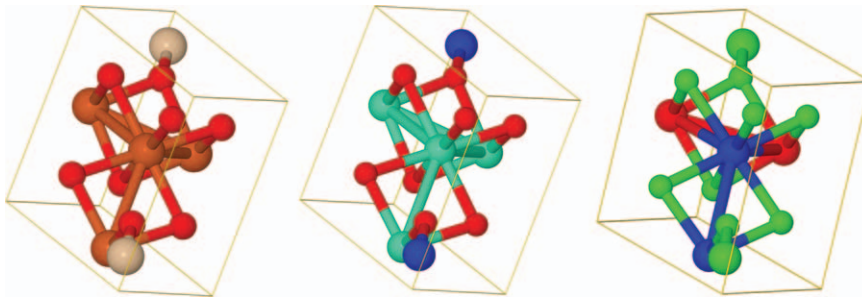


Figure 5.2 DDEC results for the anti-ferromagnetic spinel phase of crystalline Fe_2SiO_4. *Left*: Atoms are colored by element: orange (Fe), creme (Si), and red (O). *Middle*: Atoms are colored by NAC on a continuous color scale: red (−1.2) to light blue (1.2) to dark blue (2.2). *Right*: Atoms are colored by ASM on a continuous color scale: red (−3.6) to green (0.0) to blue (3.6). Lines indicate the unit cell boundaries.

following manner. After opening the DDEC_net_atomic_charges.xyz file with Jmol, the Jmol Script Editor window was opened using the File » Script Editor menu option. To color the atoms by NAC, the command "color property partialcharge" (without quotes) was typed in the Jmol Script Editor, and then the run button was pushed to execute the command. To label the atoms with the numeric NAC value, the command "label %P" (without quotes) was executed in the Jmol Script Editor. Graphical results for Fe_2SiO_4 are shown in Figure 5.2. For clarity, the molecular images shown in Figure 5.2 were created using the "color property partialcharge" command, not the "label %P" command. To color the atoms by ASM or to label the atoms with the numeric ASM value, the file DDEC_atomic_spin_moments.xyz was opened and then the command "color property partialcharge" or "label %P" was used. These Jmol images were saved using the File » Export menu options.

5.2.4 Examples using GAUSSIAN Software

In this section we give an example that uses the method defined above with GAUSSIAN applied to a molecule and to a material using periodic boundary conditions. Since the development of GAUSSIAN 09 Revision B, the natural orbitals, their occupancies, and the molecular geometry can be written to a .wfx file.[27] The .wfx format was specifically designed to contain all the information required for AIM charge partitioning.[28] The keywords Output = WfX, Density = Current, and Pop = NOAB (for spin polarized systems) or Pop = NO (for spin unpolarized systems) were added to the route line in the GAUSSIAN input file to generate the .wfx file. The keyword Density = Current specifies that the correlated density (rather than the uncorrelated Hartree–Fock density) should be written to the .wfx file. For DFT calculations, the correlated (Kohn–Sham) density is written to the .wfx file irrespective of whether

Density = Current is specified. Pop = NOAB specifies that separate orbitals for spin up and spin down electrons should be written to the .wfx file. Since GAUSSIAN 09 Revision C, Density = Current is included by default. The desired name for the .wfx output file must be specified at the bottom of the GAUSSIAN input file.

As an example, consider a coupled-cluster calculation of the ozone triplet state. The molecular geometry was optimized using the CCSD method with aug-cc-pVTZ basis sets, no frozen core electrons (full), and no symmetry constraints (NoSymm). (The keywords (full,SaveAmplitudes), scf = (fermi, maxcycle = 400), and geom = connectivity are convenient options but not required to run the calculation.) The file ozone_triplet_CCSD.wfx was written at the completion of the calculation. Below is an example GAUSSIAN input file for performing these tasks:

======GAUSSIAN input file======

```
%chk=ozone_triplet_CCSD.chk  <-- name for checkpoint file
%mem=8000MB  <-- amount of memory requested
%nproc=4   <-- number of processors requested
# opt CCSD(full,SaveAmplitudes)/aug-cc-pVTZ scf=(fermi,-
maxcycle=400)  <-- methods
# geom=connectivity  <-- initial geometry info
# density=current pop=NOAB Output=WfX  <-- keywords for
generating the .wfx file

ozone triplet example  <-- title

0 3  <-- net charge and spin multiplicity
 O   <-- the following lines specify the initial geometry
 O                    1              B1
 O                    1              B2    2            A1

   B1             1.34000000
   B2             1.34000000
   A1            98.00000000

 1 2 1.0 3 1.0
 2
 3

ozone_triplet_CCSD.wfx   <-- name of .wfx file to write
   <-- a blank line is required at the end of file
```

The optimal O–O bond length was 1.32 Å and the optimal O–O–O angle was 97.5°.[19] The NACs and ASMs were computed by running the DDEC program on the ozone_triplet_CCSD.wfx file. After reading the .wfx file, the DDEC

program sets up the core, valence, and spin density grids and checks to make sure that all electrons are properly accounted for. Once the grids are set up, core, valence, and spin partitioning are performed by the same iterative processes described for the crystalline Fe_2SiO_4 example above. The NAC (ASM) for the center atom was 0.230 (0.425) and for each outer atom was −0.115 (0.7875). The atomic dipole magnitudes, μ, were 0.162 (center atom) and 0.055 (each outer atom) in atomic units.

We now consider a hexagonal boron–nitride sheet as an example system with two-dimensional periodic boundary conditions. This calculation used the PW91 exchange-correlation functional with LANL2DZ basis sets. An automatically generated density-fitting basis set was requested by the keyword /auto. Because the geometry was optimized in a previous calculation, the calculation to generate the .wfx file was performed at fixed geometry. The GAUSSIAN input file is shown below. The Tv symbols denote the lattice vectors.

======GAUSSIAN input file======

```
%chk=BN_plane_lanl2dz_auto.chk
%mem=16000MB
%nproc=8
# pw91pw91/lanl2dz/auto output=wfx scf=tight pop = no

BN_plane

0 1  <-- net charge and spin multiplicity
 B        -0.625290   -0.368260    0.000000  <-- atom XYZ
coordinates(Å)
 N         0.625290    0.368260    0.000000
 Tv        2.513682    0.021689    0.000000    <-- lattice
vectors (Å)
 Tv        1.238058    2.187906    0.000000

BN_plane_lanl2dz_auto.wfx  <-- name of .wfx file to generate
  <-- file must terminate with a blank line
```

Because the original .wfx format did not include lattice vectors, these were manually added to the .wfx file using a text editor. Specifically, the lines

```
<Number of Translation Vectors>
2
</Number of Translation Vectors>
<Translation Vectors>
  4.75034738E+00 -6.93889390E-18  0.00000000E+00
  2.37517655E+00  4.11420313E+00  0.00000000E+00
</Translation Vectors>
```

were inserted into the .wfx file. These lines can be inserted anywhere in the .wfx file that is outside the <fieldname> </fieldname> block of another field. The translation vectors are in units of bohr and must be in the same XYZ coordinate system as the atomic positions already listed in the following lines of the .wfx file.

```
<Nuclear Cartesian Coordinates>
 -1.187587226830e+00 -6.856894978828e-01 0.000000000000e+00
  1.187587226830e+00 6.856894978828e-01  0.000000000000e+00
</Nuclear Cartesian Coordinates>
```

Translation vectors satisfying these conditions can be found by opening the GAUSSIAN .fchk file, searching for the string "Translation vectors", and then reading the set of numbers on the following lines. We hope the .wfx format specification will be upgraded to include translation vectors automatically. After adding the translation vectors to the .wfx file, the DDEC program was run as described above for the ozone triplet. The computed NACs were 0.968 (B) and −0.968 (N). The DDEC program did not perform spin partitioning for this system, because all electrons were paired and only one set of orbitals was present in the .wfx file. By definition, all ASMs are zero in systems containing only paired electrons. The DDEC program automatically performs spin partitioning if separate orbital sets for spin up and down electrons are present in the .wfx file.

Calculations can be performed using either all-electron basis sets or basis sets that replace some core electrons with relativistic effective core potentials (RECPs). According to the .wfx format specification, information about core electrons replaced by RECPs is supposed to be contained in the .wfx file. Tests should be performed on single atom systems to ensure that this has been done correctly, before using a basis set that includes RECPs for studying large systems. Specifically, if one wants to use the SDD basis set (which includes RECPs) for atoms of a particular element (*e.g.* Br) in a large system, we recommend that a calculation first be performed for a single atom of this element. The DDEC program should be run on the .wfx file for this single atom to check whether the RECP information is included in the .wfx file. If this information is not included, the DDEC program will abort with a warning that all electrons have not been properly accounted for. In this case, one should choose a different basis set for the atom and repeat the test. If the DDEC program runs without errors, the basis set can be used with confidence for larger systems. In our experience, the RECP information is being properly written to the .wfx file when using GAUSSIAN 09 Revision C or later.

Various *ab initio* methods in GAUSSIAN can generate .wfx files for DDEC analysis. These *ab initio* calculations can be performed in a vacuum or in an implicit solvent. DDEC analysis should work with any of these *ab initio* methods. We strongly recommend using methods such as DFT, coupled-cluster theory, configuration interaction, *etc.* that accurately account for electron–electron correlations. Low-level theories that do not accurately account for

electron–electron correlations may be useful for generating $\rho(\vec{r})$ for computing DDEC charges, but only for very large systems where a more accurate *ab initio* method is computationally infeasible. Second, particular care must be exercised when computing spin densities and ASMs for multi-reference systems such as open shell singlets.[29–34] For such systems, the spin density distribution that results from the *ab initio* calculation should be visualized in a program such as GaussView to make sure that it has the correct symmetry and the *ab initio* calculation has converged to the correct spin state. If this is not the case, the calculation should be redone using more appropriate methods. Once the *ab initio* calculation has converged to the correct spin state and spin density symmetry, DDEC population analysis can be performed with confidence to obtain accurate ASMs.

5.2.5 VASP Non-collinear Magnetism Example

Non-collinear magnetism occurs when the spin magnetization density, $\vec{m}(\vec{r})$, is not parallel to the same axis at all points in space. Oshio *et al*. synthesized and characterized the ferrous cube complex $Fe_4C_{40}H_{52}N_4O_{12}$ {commonly called $[Fe_4(sae)_4(MeOH)_4]$} as the first example of a single-molecule magnet involving an Fe(II) cluster.[35,36] DFT computations showed that this material has non-collinear magnetism with $\sim$2 unpaired electrons on each Fe atom and negligible spin polarization on the other atoms.[19,20] ASMs on the four Fe atoms are grouped into two pairs of anti-parallel vectors, where the two pairs are approximately orthogonal.[19] Figure 5.3 summarizes a procedure for finding the ground state geometry and magnetic configuration of materials with non-collinear magnetism.[19] Using an initial geometry guess, energies are computed for several sets of random spin moments on the magnetic atoms. Following this, geometry optimization is performed for several of these structures having comparatively low energies. The lowest energy converged geometry is further optimized to converge its magnetic degrees of freedom more tightly. After the final geometry and magnetic degrees of freedom have been fully optimized, the final geometry and four charge density components are saved for subsequent DDEC and spin–orbit coupling analysis. The spin–orbit coupling potential energy surface for this material (Figure 5.4) gives a computed magnetic anisotropy barrier of 2.9 meV,[19] which is in good agreement with the experimental value of 2.4 meV.[35,36]

These computations were performed in VASP using the PW91 functional. Example input files and further computational details are given in the Supporting Information of a previous report.[19] All exchange-correlation functionals that depend only on $\{\rho(\vec{r}), m(\vec{r})\}$ have the limitation that

$$\vec{B}_{XC}(\vec{r}) = \frac{\partial E_{XC}}{\partial \vec{m}(\vec{r})} \tag{5.29}$$

is always parallel to $\vec{m}(\vec{r})$.[19] This limitation has traditionally applied to all local spin density approximation (LSDA) and generalized gradient approximation

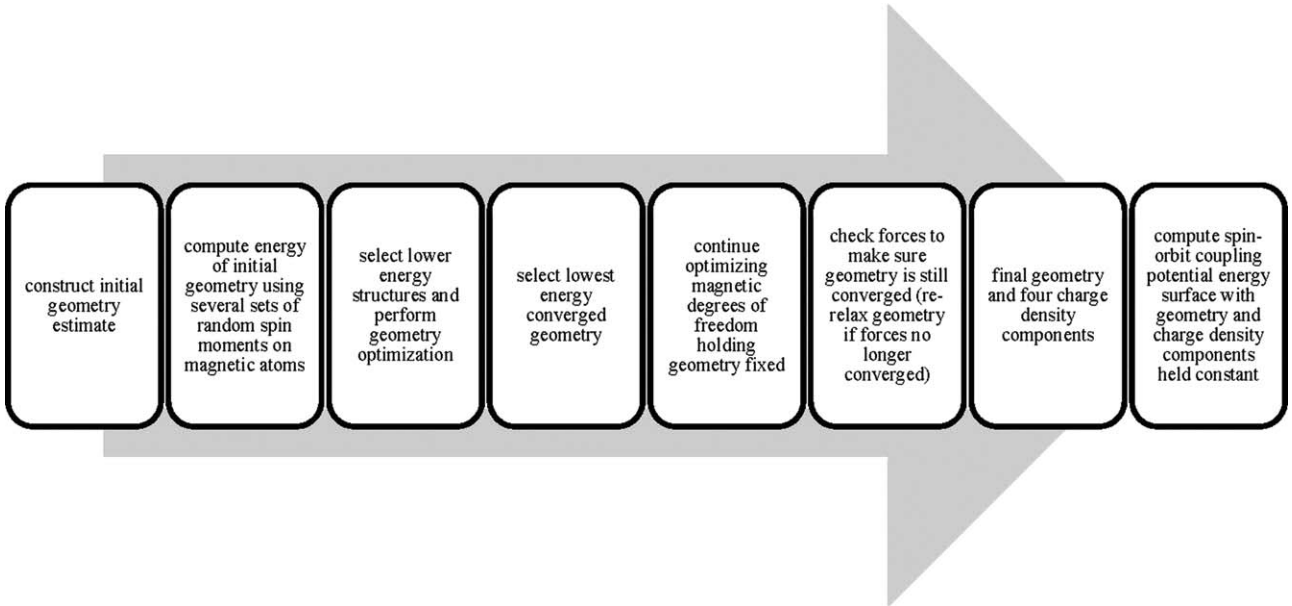


Figure 5.3 Method for computing the ground state of systems with highly non-collinear magnetism. (Reprinted with permission from T. A. Manz and D. S. Sholl, *J. Chem. Theory Comput.*, 2011, **7**, 4146–4164. Copyright American Chemical Society)

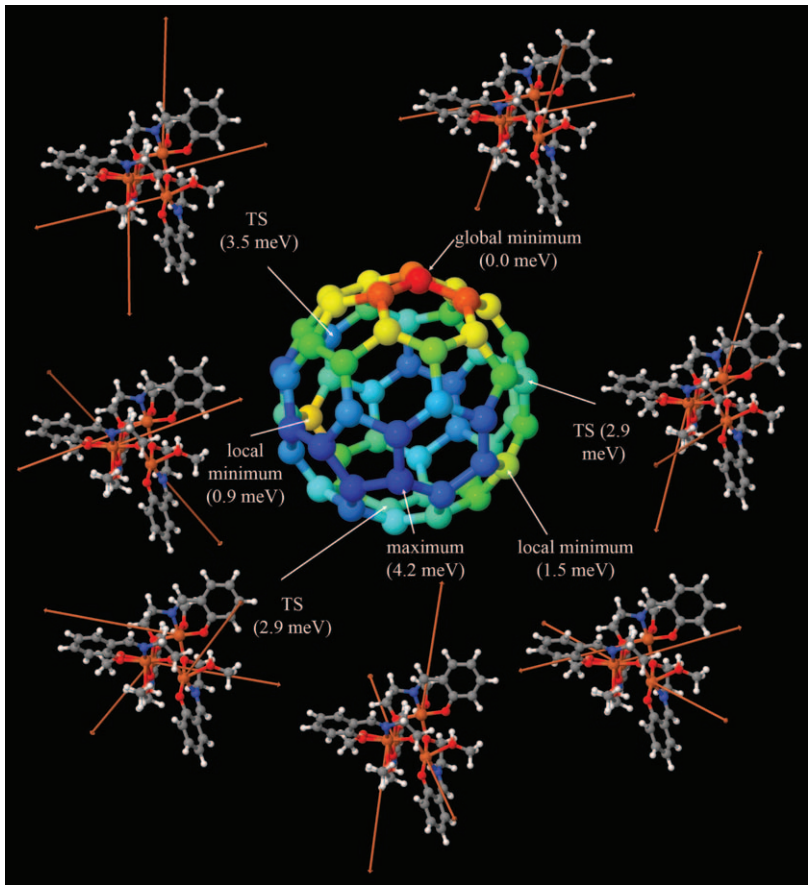


Figure 5.4 Spin–orbit coupling potential energy surface of the ferrous cube complex $Fe_4C_{40}H_{52}N_4O_{12}$. The center shows 60 points forming the truncated icosahedron used to sample the unit sphere of possible rotations of the spin axes, where color indicates the relative energy of each rotation. Adjacent to each labeled minimum, transition state, and maximum, the ASMs are displayed as vectors on the molecular structure, where color indicates the element (orange: Fe; red: O; blue: N; gray, C; white: H). ASMs are insignificant on all atoms except Fe. The large orange vectors show the direction and magnitude of Fe ASMs. The magnitudes and relative angles between ASMs are constant, but they rotate with respect to the molecular structure.
(Reprinted with permission from T. A. Manz and D. S. Sholl, *J. Chem. Theory Comput*., 2011, **7**, 4146–4164. Copyright American Chemical Society)

(GGA) functionals,[37,38] but recently an extension of GGA functionals to non-collinear magnetism has been proposed that does not have this limitation.[66] Hybrid functionals, which mix a portion of Hartree-Fock exchange with a GGA functional, also do not have this limitation, because Hartree-Fock exchange is not a functional of $\{\rho(\vec{r}), m(\vec{r})\}$ only. However, hybrid functionals

have rarely been used to study non-collinear magnetism, because of their higher computational cost.[24] We believe improvements in computer processing speed and efficiency will make the application of hybrid functionals to the study of non-collinear magnetism attractive in the near future.

From a user's perspective, performing DDEC analysis on systems with non-collinear magnetism differs in only a few respects from performing DDEC analysis on systems with collinear magnetism. Necessary files for performing non-collinear spin partitioning are generated by a single-point calculation on the final geometry using the INCAR keywords NSW = 0 (no ionic relaxations), LAECHG = .TRUE., LCHARG = .TRUE., LWAVE = .TRUE., Prec = Accurate, LNONCOLLINEAR = .TRUE., and LMAXMIX = 4 or 6. The saved CHGCAR or WAVECAR must be read from the previous geometry optimization using either the keywords ISTART = 2 or ICHARG = 1 to ensure that the initial guess of the single-point calculation has the correct magnetic structure. If desired, spin–orbit coupling calculations can be performed in the following manner. After the single-point calculation described above writes the final CHG, AECCAR0, AECCAR2, and WAVECAR files, subsequent spin–orbit coupling energies can be computed by varying the SAXIS value and using the INCAR keywords NSW = 0, LAECHG = .FALSE., LCHARG = .FALSE., LWAVE = .FALSE., Prec = Accurate, LNONCOLLINEAR = .TRUE., LSORBIT = .TRUE., ISTART = 2 (read saved WAVECAR), ICHARG = 11 (use density in saved CHGCAR), and LMAXMIX = 4 or 6. In these spin–orbit coupling calculations, the LCHARG and LWAVE keywords must be set to .FALSE. to avoid over-writing the CHGCAR and WAVECAR files. Examples of these input files can be found in the Supporting Information of a previous publication.[19] After VASP generates the CHG file, the user must manually rename it to CHG_noncollinear to alert the DDEC program to load four charge density components instead of only the two required for collinear magnetism. The SAXIS value in the VASP and DDEC programs only has meaning for spin–orbit coupling calculations, and the SAXIS value in the DDEC program should be changed only if spin–orbit coupling calculations were performed. Changing the SAXIS value in the DDEC program to the same SAXIS value for a VASP spin–orbit coupling calculation will produce the ASM vectors corresponding to that particular orientation of the spin magnetization density. For this to work properly, the above procedures must be closely followed. After DDEC analysis completes, the DDEC_atomic_spin_ moments.xyz file contains the following information for each atom: (a) atomic symbol, (b) XYZ coordinates in Å, (c) ASM magnitude, and (d) XYZ components of the ASM vector. The ASM magnitudes can be visualized in Jmol using the commands "label %P" or "color property partialcharge". The ASM vectors can also be visualized in Jmol. By default, Jmol animates the vectors as vibrations. This animation was turned off using the Tools ≫ Vibrate ≫ Stop vibration menu command. Once this was done, the spin moment vectors displayed automatically for each atom. The size of the vectors was changed using the option under the menu command Display ≫ Vector. Each molecular image shown in Figure 5.4 was created in this way.

5.3 Modeling the Electrostatic Potential Surrounding a Material

5.3.1 Atom-centered Distributed Multipole Expansion

There are a variety of applications where it is necessary to quantify the electrostatic potential outside the volume filled by atoms in a material. The NACs defined by the DDEC approach are well suited to this task because they were specifically derived to reproduce the electrostatic potential defined by a material's full electron distribution. In this section, we describe how the quality of the electrostatic potential defined by DDEC NACs can be assessed for calculations with periodic materials using VASP and localized molecules using GAUSSIAN.

The electrostatic potential, $V(\vec{r})$, is the sum of atomic electrostatic potentials

$$V(\vec{r}) = \sum_{k,\mathrm{A}} V_{\mathrm{A}}(\vec{r}_{\mathrm{A}}), \tag{5.30}$$

where $V_{\mathrm{A}}(\vec{r}_{\mathrm{A}})$ is defined as the electrostatic potential due to atom A's charge distribution,

$$V_{\mathrm{A}}(\vec{r}_{\mathrm{A}}) = \frac{z_{\mathrm{A}}}{r_{\mathrm{A}}} - \oint \frac{\rho_{\mathrm{A}}\left(\vec{r}_{\mathrm{A}}{}'\right) d^3\vec{r}_{\mathrm{A}}{}'}{\left|\vec{r}_{\mathrm{A}} - \vec{r}_{\mathrm{A}}{}'\right|} = \frac{q_{\mathrm{A}}}{r_{\mathrm{A}}} + B + C. \tag{5.31}$$

The terms B and C are due to atomic multipoles (AMs) and penetration of the atom's electron cloud, respectively.[18] The penetration term, C, decays approximately exponentially with increasing r_{A} and is approximately zero where $\rho(\vec{r})$ is negligible.[39]

The accuracy of a truncated multipole expansion for reproducing $V(\vec{r})$ can be quantified by the root mean squared error (RMSE),

$$\mathrm{RMSE} = \sqrt{\sum_{\mathrm{grid}} \left[\left(V(\vec{r}) - V_{\mathrm{model}}(\vec{r}) - \langle V(\vec{r}) - V_{\mathrm{model}}(\vec{r}) \rangle_{\mathrm{grid}} \right)^2 \Big/ N_{\mathrm{grid}} \right]} \tag{5.32}$$

where $V(\vec{r})$ is the *ab initio* electrostatic potential, $V_{\mathrm{model}}(\vec{r})$ is the electrostatic potential of the net atomic charges $\{q_{\mathrm{A}}\}$ and (optionally) atomic multipoles up to some chosen order, and N_{grid} is the number of grid points. Because a distributed multipole expansion approximately reproduces $V(\vec{r})$ only in regions where C is small, the grid points are chosen in regions where $\rho(\vec{r})$ is small. This can be accomplished by using a set of uniformly spaced points between surfaces defined by γ_{inner} and γ_{outer} times the van der Waals (vdW) radii, where $(\gamma_{\mathrm{inner}}, \gamma_{\mathrm{outer}}) = (1.4, 2.0)$ for non-periodic materials and $(\gamma_{\mathrm{inner}}, \gamma_{\mathrm{outer}}) = (1.3, 20.0)$ for periodic materials.[23,40,41] Alternatively, a set of grid points can be chosen for locations where $\rho(\vec{r})$ is below some threshold value (*e.g.* $\leq 10^{-3}$ e bohr^{-3}).[1] The relative root mean squared error (RRMSE) is defined as the RMSE value for an atomic charge/multipole model divided by the RMSE when all atomic charges/multipoles are set to zero.[23,40–42]

The DDEC program can compute the RMSE and RRMSE values for systems involving non-periodic materials and porous materials with three periodic dimensions. For porous systems with three periodic dimensions, the keyword LVTOT = .TRUE. is added to the VASP INCAR to generate a LOCPOT file containing the electrostatic potential, as described above. To compute the RMSE, the LOCPOT and POTCAR files are placed in a folder with a file called "partial_charges_data" that lists the NAC and atomic dipole moment for each atom. The NACs and atomic dipoles can be obtained from any desired model. This is useful, for example, to compute the RMSE values for a series of different point charge models describing the same material. In this case, one creates a separate "partial_charges_data" file for each point charge model and runs the calculation on each of these files. When the RMSE program is run, it separately computes and prints the RMSE corresponding to (a) just point charges, (b) point charges plus atomic dipoles, (c) no point charges or atomic dipoles. If a model does not include atomic dipole information, these variables should be set to zero in the "partial_charges_data" file. The program automatically selects optimal Ewald summation parameters for computing the model electrostatic potential, so no user input is required to do this. If desired, the user may alter the default values of γ_{inner} and γ_{outer}. The only other parameter of interest is the variable "skip" which allows the grid to be thinned. Setting skip = 1 uses every grid point in the LOCPOT file. Setting skip = 2 uses every other grid point along each lattice direction, which reduces the number of grid points by a factor of eight. This can be useful to decrease computational time. The total number of valid grid points is printed in the output file at end of the calculation. This number should be at least several thousand to ensure that a sufficient number of grid points have been used to compute meaningful RMSE and RRMSE values.

For non-periodic materials, a similar script is available for computing the RMSE from a GAUSSIAN cube file called "potential.cube" containing the electrostatic potential. This file is generated by running the GAUSSIAN cubegen program on the.fchk file. Below is an example command line to generate the potential.cube file:

```
C:\G03W> cubegen 0 Potential=type C:\B4N4.fchk
potential.cube -1
-1     -10.0     -10.0     -10.0
-102     0.2     0.0     0.0
102     0.0     0.2     0.0
102     0.0     0.0     0.2
```

In the above "Potential = type" phrase, type should be replaced with SCF, MP2, CI, or CC to specify the type of density to use for generating the electrostatic potential. The value of −1 at the end of the command line prompts GAUSSIAN to read the cube parameters, which are manually input in the following format:[27]

-1, X_0, Y_0, Z_0
Output unit number and initial point.

-N_1, X_1, Y_1, Z_1
Number of points and step-size along first directional vector: $\vec{v}_1=(N_1X_1, N_1Y_1, N_1Z_1)$.

N_2, X_2, Y_2, Z_2
Number of points and step-size along second directional vector: $\vec{v}_2=(N_2X_2, N_2Y_2, N_2Z_2)$

N_3, X_3, Y_3, Z_3
Number of points and step-size along the third directional vector: $\vec{v}_3=(N_3X_3, N_3Y_3, N_3Z_3)$

All cube parameters are entered in atomic units (bohr). A large enough volume should be specified by the N_1, N_2, and N_3 parameters to include enough space around the molecule to completely enclose the surface defined by γ_{outer} times the vdW radii. The RMSE script for non-periodic materials is run in a similar manner to the RMSE script for periodic materials described above.

5.3.2 Applications to Force-fields used in Atomistic Simulations

Force-fields used in molecular dynamics, Monte Carlo, and other types of atomistic simulations typically contain several types of interaction: (a) electrostatic interactions due to the atomic charges and (optionally) atomic multipoles, (b) van der Waals interactions, and (optionally) (c) bond, angle, and torsion force parameters to describe the material's flexibility.[6,43–45] If part of a structure is kept rigid, a force-field for the atoms in this portion of the material can be constructed using the NACs and a vdW model such as the Lennard–Jones or Morse potentials. For rigid materials, the electrostatic potential grid from an *ab initio* calculation can also be used directly as input for performing subsequent atomistic simulations.[23] Watanabe *et al.* used this approach to compute CO_2 adsorption isotherms in metal–organic frameworks.[23,68] Adsorption isotherms computed with several different point charge models were compared to the isotherm computed using the DFT electrostatic potential energy surface.[23] Computed isotherms for the metal–organic framework ZIF-90 are shown in Figure 5.5. During these calculations, the ZIF-90 framework was held rigid at the DFT-optimized geometry. The isotherm computed without electrostatic interactions (labeled "No Charge") dramatically underestimated CO_2 adsorption. Isotherms computed with the REPEAT and DDEC methods slightly underestimated the CO_2 adsorption. The REPEAT method computes NACs in porous periodic materials by minimizing the RMSE over a chosen set of grid points.[42] This leads to accurate adsorption isotherms, but the REPEAT charges are highly sensitive to the choice of grid points.[23,40] Table 5.2 compares RMSE and RRMSE values for several metal–organic frameworks using different point charge models. Although the DDEC charges give a higher RMSE than the REPEAT charges,[23] the DDEC charges are more chemically meaningful and less sensitive to the choice of grid points.[18–20,23] For these reasons, we expect the DDEC charges to be more suitable than REPEAT charges for constructing force-fields of porous materials.

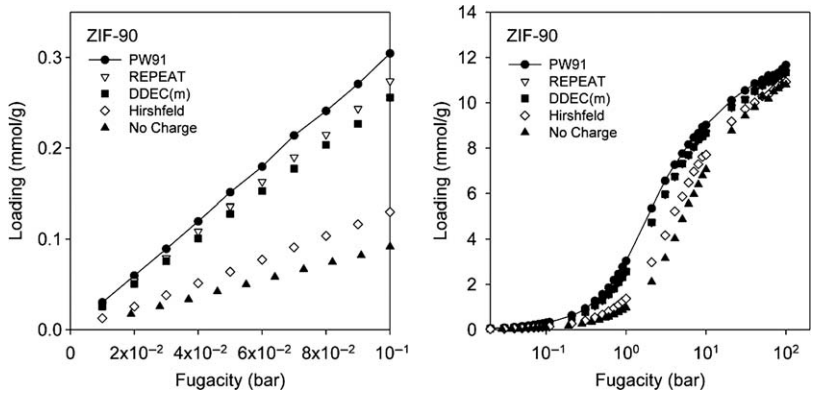


Figure 5.5 CO_2 adsorption isotherm for the metal organic framework ZIF-90 computed with several different point charge models. The symbol PW91 denotes the isotherm computed directly from the DFT electrostatic potential energy surface.
(Reprinted with permission from T. Watanabe, T. A. Manz and D. S. Sholl, *J. Chem. Phys. C*, 2011, **115**, 4824–4836. Copyright American Chemical Society).

Table 5.2 RRMSE (dimensionless) and RMSE (kJ mol^{-1} in brackets []) for several metal organic frameworks. DDEC/c3(d) includes the NACs and atomic dipoles, while DDEC/c3(m) includes only the NACs. (From refs. 23 and 20)

	REPEAT	*DDEC/c3(d)*	*DDEC/c3(m)*	*Hirshfeld*	*No Charge*
IRMOF-1	0.1 [1.3]	0.2 [2.4]	0.3 [2.7]	0.6 [6.3]	1.0 [11]
ZIF-8	0.2 [1.7]	0.5 [3.0]	0.6 [3.7]	0.9 [7.2]	1.0 [7.7]
ZIF-90	0.1 [3.2]	0.1 [3.4]	0.1 [3.5]	0.5 [15]	1.0 [32]
$Zn(nicotinate)_2$	0.1 [1.1]	0.3 [1.9]	0.5 [4.0]	0.7 [6.2]	1.0 [9.3]

5.4 Effective Bond Orders

Chemical bonds are conventionally thought to arise from the hybridization of two atomic orbitals of like spin centered on two adjacent atoms. This concept can be quantified by defining effective bond order (EBO) descriptors that measure the number of like-spin electrons shared between two atoms. The second-order density matrix is important for conceptually understanding the origins of the EBO. Let $\rho_{\alpha\alpha}(\vec{e}_1, \vec{e}_2)$ be the probability density of simultaneously finding a spin up electron at position $\vec{e}_1$ and another spin up electron at $\vec{e}_2$. Similarly, let $\rho_{\beta\beta}(\vec{e}_1, \vec{e}_2)$ be the probability density of simultaneously finding a spin down electron at $\vec{e}_1$ and another at $\vec{e}_2$. We define the EBO in general form for collinear magnetism as

$$EBO_{AB} \equiv \left\langle \left(P^{\alpha}_{A}(\vec{e}_1)P^{\alpha}_{B}(\vec{e}_2) + P^{\alpha}_{B}(\vec{e}_1)P^{\alpha}_{A}(\vec{e}_2)\right)\left(\rho_{\alpha}(\vec{e}_1)\rho_{\alpha}(\vec{e}_2) - \rho_{\alpha\alpha}(\vec{e}_1, \vec{e}_2)\right)\right\rangle + \left\langle \left(P^{\beta}_{A}(\vec{e}_1)P^{\beta}_{B}(\vec{e}_2) + P^{\beta}_{B}(\vec{e}_1)P^{\beta}_{A}(\vec{e}_2)\right)\left(\rho_{\beta}(\vec{e}_1)\rho_{\beta}(\vec{e}_2) - \rho_{\beta\beta}(\vec{e}_1, \vec{e}_2)\right)\right\rangle. \quad (5.33)$$

The Fermi hole density of the spin up electrons,

$$\rho^{\alpha}(\vec{e}_1)\rho^{\alpha}_{\text{hole}}(\vec{e}_1,\vec{e}_2) = \rho^{\alpha}(\vec{e}_1)\rho^{\alpha}(\vec{e}_2) - \rho^{\alpha\alpha}(\vec{e}_1,\vec{e}_2) \geq 0, \tag{5.34}$$

is the probability density of excluding a spin up electron at $\vec{e}_2$ when a spin up electron is at $\vec{e}_1$ and satisfies the well-known normalization condition

$$\oint (\rho^{\alpha}(\vec{e}_1)\rho^{\alpha}(\vec{e}_2) - \rho^{\alpha\alpha}(\vec{e}_1,\vec{e}_2))d^3\vec{e}_2 = \rho^{\alpha}(\vec{e}_1). \tag{5.35}$$

An analogous normalization condition holds for the Fermi hole of the spin down electrons. The projector $P^{\alpha}_{\text{A}}(\vec{e}_1) \geq 0$ defines the probability of assigning a spin up electron at position $\vec{e}_1$ to atom A, and as a probability must satisfy

$$\sum_{\text{A}} P^{\alpha}_{\text{A}}(\vec{e}_1) = 1. \tag{5.36}$$

Examining eqns (5.33) and (5.34), EBO_{AB} is always ≥ 0.

Different definitions of the projectors, $\{P^{\alpha}_{\text{A}}(\vec{e}_1)\}$, lead to different EBO_{AB} descriptors. To be chemically meaningful, the projectors should be formally invariant of the basis sets used for calculating $\rho(\vec{r})$. At least three types of basis-set independent projectors have been described in the prior literature:

1. Projection onto the natural atomic orbitals of atom A:

$$P^{\alpha}_{\text{A}} = \sum_i \left|\mu^{\alpha}_i\right\rangle_{\text{A}}\left\langle\mu^{\alpha}_i\right| \tag{5.37}$$

   where $\left\{\left|\mu^{\alpha}_i\right\rangle_{\text{A}}\right\}$ are the NAOs centered on atom A.[10–12] To the best of our knowledge, all the descriptions and applications of this projector in the prior literature apply it to one-electron density matrices,[10–12] rather than the full two-electron expansion shown in eqn (5.33). Recently, Bultinck *et al*. developed AIM projectors that appear to yield similar information to NPA, using a different computational framework.[46]
2. Projection onto a volumetric compartment assigned to atom A:

$$P_{\text{A}}(\vec{r}) = \begin{cases} 1 \text{ if } \rho_{\text{A}}(\vec{r}_{\text{A}}) = \max(\{\rho_{\text{B}}(\vec{r}_{\text{B}})\}) \\ 0 \text{ otherwise} \end{cases} \tag{5.38}$$

   which has commonly been used with Bader compartments.[47–55] It can, however, be used with atomic basins computed using any AIM method. DDEC basin projectors result when the DDEC $\rho_{\text{A}}(\vec{r}_{\text{A}})$ is inserted into eqn (5.38).
3. Projection using the same atomic weights used to partition $\rho(\vec{r})$ into overlapping atoms,[46,53,56–58]

$$P_{\text{A}}(\vec{r}) = w_{\text{A}}(r_{\text{A}})/W(\vec{r}). \tag{5.39}$$

The above projectors and similar ones can be expressed in matrix form and diagonalized to obtain "bonding orbitals".[10–12,46,50,53–55,58,59,71,72] This allows one to identify the occupancies and shapes of independent bonding modes between atom pairs.[10–12,46,50,53–55,58,59,71,72]

Because the H_2 molecule contains just one electron of each spin, its EBO can be computed analytically and is identically equal to 1 for any of the above choices of projector. The closely related shared electron distribution index (SEDI) measures the number of electrons shared between two atoms irrespective of the spins at $\vec{e}_1$ and $\vec{e}_2$.[53,58] Because this correlation decreases electron delocalization, it necessarily leads to a SEDI less than 1 for the H_2 molecule.[69] This provides a compelling reason to define the EBO as sharing of like spin electrons.[69]

Owing to the large number of components contained in the second-order density matrix, it is convenient to develop approximations to it that allow EBOs to be computed using just the first-order density matrix.[11,12,46–49,52,54,56,58,60] The first-order density matrix is routinely available from quantum chemistry calculations in the form of its eigenvectors (the natural spin-orbitals, $|\psi_i(\vec{r},\sigma)\rangle = |\phi_i(\vec{r})\rangle|\sigma_i(\vec{r})\rangle$) and eigenvalues (the natural spin-orbital occupations, $\{f_i\}$). Most EBO measures, including NPA/NBO analysis, use the first-order density matrices as input. However, there have been a few reports of EBOs computed for small molecules using explicitly correlated second-order density matrices.[53,54,70] When the multielectronic wavefunction is a single Slater determinant, the second-order density matrix can be explicitly written in terms of the first-order density matrix. Because theories such as Hartree–Fock and Kohn–Sham DFT typically use a single Slater determinant for the wavefunction, the EBOs can be computed directly from the first-order density matrices when using these methods.

Recently, Wheatley and Gopal reported a linear correlation between iterative stockholder atom (ISA) overlap populations and covalent bond orders:

$$EBO_{AB} \approx K \oint \frac{w_A(\vec{r}_A) w_B(\vec{r}_B)}{(W(\vec{r}))^2} \rho(\vec{r}) d^3\vec{r} \tag{5.40}$$

with an empirically determined value $K \approx 3.32$.[61] Computed bond orders were within 0.4 of the expected values for all tested systems, and within 0.1 of the expected value half the time.[61] An underlying theoretical reason for $K \approx 3.32$ was not given. This descriptor has several advantages and would be easy to apply to several types of systems for which existing EBO methods have limitations. Specifically, this EBO descriptor could be readily computed for materials with periodic boundary conditions and multi-determinant wavefunctions. Using the EBO definition of eqn (5.33), we now derive the physical meaning of the proportionality factor K. Because the systems examined by Wheatley and Gopal included only paired electrons, the spin-free

projector of eqn (5.39) can be used. Under these conditions, eqn (5.33) can be rewritten as:

$$\mathrm{EBO_{AB}} = 2 \oint \oint \frac{w_\mathrm{A}(\vec{r}_\mathrm{A}) w_\mathrm{B}(\vec{r}'_\mathrm{B})}{W(\vec{r}) W(\vec{r}')} \rho(\vec{r}) \rho_\mathrm{hole}(\vec{r}, \vec{r}') d^3\vec{r} d^3\vec{r}'. \tag{5.41}$$

where $\rho_\mathrm{hole}(\vec{r}, \vec{r}')$ is defined by eqn (5.34) and it follows

$$\oint \rho_\mathrm{hole}(\vec{r}, \vec{r}') d^3\vec{r}' = 1. \tag{5.42}$$

Comparing eqns (5.41) and (5.42) gives

$$\mathrm{EBO_{AB}} = 2 \oint \frac{w_\mathrm{A}(\vec{r}_\mathrm{A})}{W(\vec{r})} \left\langle \frac{w_\mathrm{B}(\vec{r}'_\mathrm{B})}{W(\vec{r}')} \right\rangle_\mathrm{hole} \rho(\vec{r}) d^3\vec{r}. \tag{5.43}$$

Comparing eqns (5.40) and (5.43) gives

$$K = 2 \oint \frac{w_\mathrm{A}(\vec{r}_\mathrm{A})}{W(\vec{r})} \left\langle \frac{w_\mathrm{B}(\vec{r}'_\mathrm{B})}{W(\vec{r}')} \right\rangle_\mathrm{hole} \rho(\vec{r}) d^3\vec{r} \Bigg/ \oint \frac{w_\mathrm{A}(\vec{r}_\mathrm{A}) w_\mathrm{B}(\vec{r}_\mathrm{B})}{(W(\vec{r}))^2} \rho(\vec{r}) d^3\vec{r}. \tag{5.44}$$

The value of K can now be estimated by assuming that only two atoms overlap at each point in space, so that

$$0 \le \frac{w_\mathrm{A}(\vec{r}_\mathrm{A})}{W(\vec{r})} = \left(1 - \frac{w_\mathrm{B}(\vec{r}_\mathrm{B})}{W(\vec{r})}\right) \le 1. \tag{5.45}$$

If we assume that $w_\mathrm{A}(\vec{r}_\mathrm{A})/W(\vec{r})$ is uniformly distributed over the range 0 to 1 and independent of $\rho(\vec{r})$, then eqn (5.44) becomes

$$K \approx 2 \left(\int_0^1 \mathrm{d}x \right) \left(\int_0^1 (1-x) \mathrm{d}x \right) \Bigg/ \int_0^1 x(1-x) \mathrm{d}x = 3 \tag{5.46}$$

Thus, $K \approx 3$ arises as an approximation of eqn (5.33) and agrees closely with the empirical value $K \approx 3.32$. This shows that the correlation suggested by Wheatley and Gopal has a firm physical basis.

For selected diatomic molecules, Table 5.3 shows EBOs computed with three different projectors. All three of these methods follow the general scheme of eqn (5.33). The DDEC basin EBOs use the projector of eqn (5.38). The natural atomic orbital projector (NOAP) EBOs use the projector of eqn (5.37). Finally, the overlap EBOs use eqn (5.40), which results from simplifying the projector of eqn (5.39) in the manner described above. The overlap EBO values were taken directly from the literature.[61] All three methods give similar EBOs, but the NAOP results correlate most closely with the EBOs from simple molecular orbital theory. For these systems, the geometry and electron density were computed using the PW91 exchange-correlation functional with 6-311+G* basis sets.

In every case, the EBO computed using DDEC basins was greater than or equal to the NAOP EBO. Because core electron orbitals are confined to a single

Table 5.3 Computed EBOs of diatomic molecules.

	Bond length (Å)	*EBO from simple MO theory*	*EBO*		
			DDEC basins	*NAOP*	*Overlap*
H_2	0.75	1	1.00	1.00	0.94[b]
O_2	1.22	2	2.24	2.03	[c]
CO	1.14	>2 and <3[a]	2.56	2.30	2.41[b]
N_2	1.11	3	3.04	3.03	2.70[b]
NaCl	2.39	0	0.32	0.17	0.27[b]
Kr_2	4.36	0	0.03	0.00	[c]
Cl_2	2.06	1	1.32	1.02	1.37[b]

[a]For CO, two electron pairs are shared between C and O and a third electron pair formally originating from the O atom is partially shared with the C atom leading to a bond order intermediate between 2 and 3.
[b]From ref. 61.
[c]Not computed.

basin and comprised of a single atom's NAOs, they do not contribute to either the basin projector or NAOP bond orders. Because a valence lone pair is more diffuse than core electrons, it may cross the boundary of an atom's basin leading to a small but non-negligible contribution to the basin projector bond order. Because this valence lone pair is approximately described by a single NAO, its contribution to the NAOP bond order is negligible. From this, we infer that the main difference between the basin projector and the NAOP bond orders is due to the exchange of lone pairs across basin boundaries. Examining Table 5.3, the basin projector EBO tends to be slightly larger than the accepted covalent bond order. The basin EBO can be expressed as two times the trace of a matrix, and equals the sum of the eigenvalues (populations) of this matrix, where the eigenvectors represent the bond orbitals. Because this matrix is real and symmetric, all of its eigenvalues, λ_k^{AB}, are positive. Accordingly, a potential strategy for removing lone pair interactions is to separate the basin EBO into a "covalent bond order" (CBO) and a "weak bond order" (WBO)

$$\mathrm{EBO_{AB}} = \mathrm{CBO_{AB}} + \mathrm{WBO_{AB}} \tag{5.47}$$

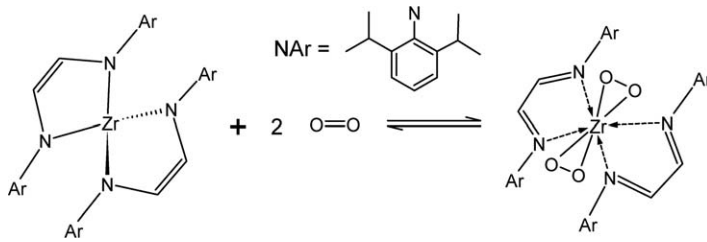
using a smooth function $f(s)$ that monotonically decreases from $f(1)=1$ to $f(0)=0$:

$$\mathrm{CBO_{AB}} = 2\sum_k \left(\lambda_k^{AB} f\left(\lambda_k^{AB}\right)\right) \tag{5.48}$$

$$\mathrm{WBO_{AB}} = 2\sum_k \left(\lambda_k^{AB}\left(1 - f\left(\lambda_k^{AB}\right)\right)\right). \tag{5.49}$$

Additional work is needed to determine whether an explicit form for $f(s)$ can be found such that the *CBO* values closely correspond to the NAOP bond orders. Examining Table 5.3, it would be useful if $f(s)$ is chosen such that the CBO is $\sim$1.0 ($\sim$1.0) and the WBO is $\sim$0.3 ($\sim$0.0) for the Cl_2 (H_2) molecule.

As a final example, we consider the reaction of Stanciu *et al.* in which a Zr organometallic complex adsorbs two O_2 molecules to form a bisperoxy complex, as shown in Scheme 5.1.[62] This reaction is accompanied by a change of bond orders in the ligands.[62] Previously, we reported DDEC NACs on the B3LYP/LANL2DZ optimized geometry of this complex before and after O_2 adsorption, which showed that there is electron transfer from the ligands to the O_2 molecule.[18,20] The geometric parameters and NAOP bond orders are summarized in Table 5.4. These results clearly show that the O–O bond is single



Scheme 5.1 Electron transfer reaction in which $ZrN_4C_{52}H_{72}$ reacts with two oxygen molecules to produce a bisperoxy complex. During this reaction, the bond orders in the Ar–N–C=C–N–Ar ligand backbone are converted to Ar–N=C–C=N–Ar.

Table 5.4 Comparison of experimental and computed structural parameters. The bond lengths and EBOs for each row correspond to the two atoms shown in underlined type in the first column. Ar denotes the adjacent aromatic C atom.

	Zr bisperoxy complex			*Zr bare complex*		
Two atoms	*X-ray[a] length(Å)*	*DFT length(Å)*	*NAOP EBO*	*X-ray[a] length(Å)*	*X-ray[a] length(Å)*	*NAOP EBO*
Zr–O	2.03, 2.03, 2.04, 2.04	2.09, 2.09, 2.09, 2.09	0.6	–	–	–
O–O	1.50, 1.51	1.54, 1.54	1.0	–	(isolated O_2: 1.27)	(isolated O_2: 2.03)
Zr–N	2.43, 2.44, 2.44, 2.45	2.47, 2.47, 2.47, 2.47	0.2	2.05, 2.06, 2.07, 2.09	2.08, 2.08, 2.13, 2.13	0.6
N–C–C–N	1.47, 1.48	1.47, 1.47	1.1	1.36, 1.37	1.38, 1.38	1.6
N–C–C–N	1.26, 1.26, 1.27, 1.27	1.30, 1.30, 1.30, 1.30	1.8	1.40, 1.40, 1.41, 1.41	1.43, 1.43, 1.43, 1.43	1.1
N–Ar	1.46, 1.46, 1.46, 1.46	1.46, 1.46, 1.46, 1.46	1.0	1.43, 1.43, 1.43, 1.44,	1.44, 1.44, 1.45, 1.45	1.0
Zr–N–C–C–N	3.18, 3.19, 3.19, 3.20	3.22, 3.22, 3.22, 3.22	0.0	2.50, 2.51, 2.51, 2.53,	2.64, 2.64, 2.65, 2.65,	0.1
Zr–N–Ar	3.54, 3.54, 3.54, 3.57	3.59, 3.59, 3.59, 3.59	0.0	3.33, 3.33, 3.39, 3.40,	3.34, 3.34, 3.43, 3.43	0.0
Aromatic C–C	~1.40	~1.42	1.4	~1.40	~1.40	1.4
Ar–H	~0.95	~1.09	0.9	~0.95	~1.09	0.9
Aliphatic C–H	~0.98	~1.10	0.9	~0.98	~1.10	0.9
Aliphatic C–C	~1.53	~1.55	1.0	~1.53	~1.55	1.0

[a]From ref. 62.

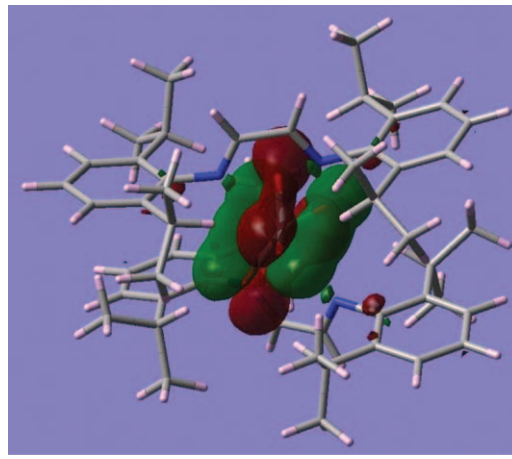


Figure 5.6 Molecular orbital # 227 in the Zr bisperoxy complex with overlaps between Zr *d*-orbital and O *p*-orbitals.

order in the bisperoxy complex and that the ligand backbone has bond orders of Ar–N–C=C–N–Ar in the bare complex and Ar–N=C–C=N–Ar in the bisperoxy complex. The experimental and computed bond lengths are in accord with these bond order changes. The molecular orbital in Figure 5.6 shows one of the major components of bonding between the Zr metal and the two adsorbed O_2 units.

This progress is encouraging, but some issues still need to be resolved. For materials with periodic boundary conditions or non-collinear magnetism, EBOs will probably be easiest to compute using a modified form of the projector in eqn (5.43) that explicitly accounts for spin degrees of freedom. This will require only the electron and spin density distributions as inputs. Tests also need to be performed to see whether explicit averaging over the hole, as shown in eqn (5.43), increases the EBO accuracy compared with the simplified correlation shown in eqn (5.40). Since eqn (5.40) was only tested for spin unpolarized systems,[61] additional tests need to be performed for spin polarized systems. For non-periodic systems, EBOs computed with proposed projectors should be carefully compared to NAOP bond orders. Because the Kohn–Sham DFT spin-orbitals form a single Slater determinant, the basin EBO has the form

$$\mathrm{EBO}_{\mathrm{AB}} = 2\sum_{i\in occ}\sum_{j\in occ}\oint_{\mathrm{A}}\psi_i^{\dagger}(\vec{r})\psi_j(\vec{r})d^3\vec{r}\oint_{\mathrm{B}}\psi_i^{\dagger}(\vec{r}')\psi_j(\vec{r}')d^3\vec{r}' \qquad (5.50)$$

where $\psi_i(\vec{r})$ is a Kohn–Sham spin-orbital, $\psi_i^{\dagger}(\vec{r})$ is its conjugate transpose, the spatial integrals are performed over individual atomic volumes, and the summations over i and j are performed for the occupied orbitals. Eqn (5.50) is directly applicable to systems with either collinear or non-collinear magnetism. We believe that, with sufficient development, these approaches might lead to

the widespread use of computed bond orders for materials with periodic boundary conditions and non-collinear magnetism.

5.5 Conclusions

Net atomic charges (NACs), atomic spin moments (ASMs), and effective bond orders (EBOs) are widely used concepts in the chemical sciences. These properties can be computed using a variety of atomic population analysis methods. The DDEC method is applicable to both periodic and non-periodic materials and simultaneously optimizes the atomic charges to reproduce chemical states in a material as well as the electrostatic potential outside the material's electron distribution. These properties make DDEC NACs especially suitable for constructing transferable force-fields used in atomistic simulations. DDEC ASMs are ideally suited for assigning spin to different atoms in collinear and non-collinear magnetic systems. Natural population analysis (NPA) and the related natural bond orbital (NBO) and Adaptive Natural Density Partitioning methods are useful for computing EBOs of molecular systems, and have recently been extended to periodic materials.[12,67,73,74] Although basis set invariant EBOs are commonly used to study molecular systems, they have only rarely been computed for periodic materials.[47,63–65,67,74] Several different indices that measure the number of electrons shared between atomic compartments are potential ways to compute EBOs for periodic materials. A recently reported correlation between EBOs and overlap populations for iterated stockholder atoms is an appealing strategy, because it can compute the EBOs directly from $\rho(\vec{r})$ without using the natural orbitals as input.[61] However, this strategy has yet to be tested on periodic materials. A more accurate and widely applicable method for computing EBOs could transform our understanding of periodic materials. For these reasons, atomic population analysis methods will remain an important field of research in the near future.

Finally, we comment on performing DDEC analysis on output from various quantum chemistry packages. In principle, the .wfx file format is applicable to all periodic and non-periodic calculations performed with localized gaussian basis sets. Any combination of atom-centered and non-atom centered (''ghost atom'') gaussian basis sets is supported. For quantum chemistry calculations performed with localized gaussian basis sets, .wfx is the preferred file type for performing DDEC analysis. The DDEC program can also directly analyze stored valence and spin density grids, which makes it applicable to output from any quantum chemistry program regardless of basis set type.

Acknowledgments

This material is based on work supported as part of the Center for Atomic Level Catalyst Design, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001058 (Center for Atomic Level Catalyst Design). Supercomputing resources were provided by the Georgia Institute of

Technology and the Extreme Science and Engineering Discovery Environment (XSEDE project grant TG-CTS100027). XSEDE is funded by NSF grant OCI-1053575.

References

1. T. Verstraelen, S. V. Sukhomlinov, V. Van Speybroeck, M. Waroquier and K. S. Smirnov, *J. Phys. Chem. C*, 2012, **116**, 490–504.
2. T. Verstraelen, E. Pauwels, F. De Proft, V. Van Speybroeck, P. Geerlings and M. Waroquier, *J. Chem. Theory Comput.*, 2012, **8**, 661–676.
3. W. L. Jolly and W. B. Perry, *J. Am. Chem. Soc.*, 1973, **95**, 5442–5450.
4. A. K. Rappe and W. A. Goddard, *J. Phys. Chem.*, 1991, **95**, 3358–3363.
5. R. A. Nistor, J. G. Polihronov, M. H. Muser and N. J. Mosey, *J. Chem. Phys.*, 2006, **125**, 094108.
6. T. R. Shan, B. D. Devine, J. M. Hawkins, A. Asthagiri, S. R. Phillpot and S. B. Sinnott, *Phys. Rev. B*, 2010, **82**, 235302.
7. L. Pauling, *J. Am. Chem. Soc.*, 1947, **69**, 542–553.
8. S. S. Glad and F. Jensen, *J. Org. Chem.*, 1997, **62**, 253–260.
9. K. N. Houk, S. M. Gustafson and K. A. Black, *J. Am. Chem. Soc.*, 1992, **114**, 8565–8572.
10. A. E. Reed, R. B. Weinstock and F. Weinhold, *J. Chem. Phys.*, 1985, **83**, 735–746.
11. J. P. Foster and F. Weinhold, *J. Am. Chem. Soc.*, 1980, **102**, 7211–7218.
12. A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899–926.
13. R. F. W. Bader, P. J. Macdougall and C. D. H. Lau, *J. Am. Chem. Soc.*, 1984, **106**, 1594–1605.
14. R. F. W. Bader, *Acc. Chem. Res.*, 1985, **18**, 9–15.
15. P. Bultinck, C. Van Alsenoy, P. W. Ayers and R. Carbo-Dorca, *J. Chem. Phys.*, 2007, **126**, 144111.
16. F. L. Hirshfeld, *Theor. Chim. Acta*, 1977, **44**, 129–138.
17. T. C. Lillestolen and R. J. Wheatley, *Chem. Comm.*, 2008, **45**, 5909–5911.
18. T. A. Manz and D. S. Sholl, *J. Chem. Theory Comput.*, 2010, **6**, 2455–2468.
19. T. A. Manz and D. S. Sholl, *J. Chem. Theory Comput.*, 2011, **7**, 4146–4164.
20. T. A. Manz and D. S. Sholl, *J. Chem. Theory Comput.*, 2012, **8**, 2844–2867.
21. D. Geldof, A. Krishtal, F. Blockhuys and C. Van Alsenoy, *J. Chem. Theory Comput.*, 2011, **7**, 1328–1335.
22. R. F. W. Bader, *J. Phys. Chem. A*, 2007, **111**, 7966–7972.
23. T. Watanabe, T. A. Manz and D. S. Sholl, *J. Phys. Chem. C*, 2011, **115**, 4824–4836.
24. J. E. Peralta, G. E. Scuseria and M. J. Frisch, *Phys. Rev. B*, 2007, **75**, 125119.
25. J. Hafner, *J. Comput. Chem.*, 2008, **29**, 2044–2078.
26. G. Kresse and J. Furthmuller, *Phys. Rev. B*, 1996, **54**, 11169–11186.
27. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakasuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov,

J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakamjima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. J. Montgomery, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Ciolowski and D. J. Fox, GAUSSIAN 09, Gaussian, Inc., Wallingford, CT, 2009.

28. The .wfx format is described at http://aim.tkgristmill.com/wfxformat.html.
29. J. Grafenstein, E. Kraka and D. Cremer, *Chem. Phys. Lett.*, 1998, **288**, 593–602.
30. J. Grafenstein, E. Kraka, M. Filatov and D. Cremer, *Int. J. Mol. Sci.*, 2002, **3**, 360–394.
31. A. I. Krylov, *Chem. Phys. Lett.*, 2001, **338**, 375–384.
32. F. A. Evangelista, *J. Chem. Phys.*, 2011, **134**, 224102.
33. S. Grimme and M. Waletzke, *J. Chem. Phys.*, 1999, **111**, 5645–5655.
34. L. V. Slipchenko and A. I. Krylov, *J. Chem. Phys.*, 2002, **117**, 4694–4708.
35. H. Oshio, N. Hoshino and T. Ito, *J. Am. Chem. Soc.*, 2000, **122**, 12602–12603.
36. H. Oshio, N. Hoshino, T. Ito and M. Nakano, *J. Am. Chem. Soc.*, 2004, **126**, 8805–8812.
37. K. Capelle, G. Vignale and B. L. Gyorffy, *Phys. Rev. Lett.*, 2001, **87**, 206403.
38. S. Sharma, J. K. Dewhurst, C. Ambrosch-Draxl, S. Kurth, N. Helbig, S. Pittalis, S. Shallcross, L. Nordstrom and E. K. U. Gross, *Phys. Rev. Lett.*, 2007, **98**, 196405.
39. M. A. Freitag, M. S. Gordon, J. H. Jensen and W. J. Stevens, *J. Chem. Phys.*, 2000, **112**, 7300–7306.
40. D. L. Chen, A. C. Stern, B. Space and J. K. Johnson, *J. Phys. Chem. A*, 2010, **114**, 10225–10233.
41. U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, 1984, **5**, 129–145.
42. C. Campana, B. Mussard and T. K. Woo, *J. Chem. Theory Comput.*, 2009, **5**, 2866–2878.
43. J. R. Hill, C. M. Freeman and L. Subramanian, in *Reviews in Computational Chemistry*, ed. K. B. Lipkowitz and D. B. Boyd, Wiley-VCH, New York, 2000, vol. 16, p. 141.
44. G. Kaminski and W. L. Jorgensen, *J. Phys. Chem.*, 1996, **100**, 18010–18013.
45. J. W. Ponder and D. A. Case, in *Protein Simulations*, ed. V. Daggett, Elsevier, Amsterdam, 2003, vol. 66, pp. 27–85.
46. D. Vanfleteren, D. Van Neck, P. Bultinck, P. W. Ayers and M. Waroquier, *J. Chem. Phys.*, 2012, **136**, 014107.

47. G. La Penna, S. Furlan and M. Sola, *Theor. Chem. Acc.*, 2011, **130**, 27–36.
48. R. L. Fulton, *J. Phys. Chem.*, 1993, **97**, 7516–7529.
49. R. L. Fulton and S. T. Mixon, *J. Phys. Chem.*, 1993, **97**, 7530–7534.
50. R. Ponec, *J. Math. Chem.*, 1998, **23**, 85–103.
51. R. F. W. Bader and M. E. Stephens, *J. Am. Chem. Soc.*, 1975, **97**, 7391–7399.
52. J. G. Angyan, M. Loos and I. Mayer, *J. Phys. Chem.*, 1994, **98**, 5244–5248.
53. P. Bultinck, D. L. Cooper and R. Ponec, *J. Phys. Chem. A*, 2010, **114**, 8754–8763.
54. D. L. Cooper and R. Ponec, *Phys. Chem. Chem. Phys.*, 2008, **10**, 1319–1329.
55. R. Ponec, G. Yuzhakov and M. R. Sundberg, *J. Comput. Chem.*, 2005, **26**, 447–454.
56. I. Mayer and P. Salvador, *Chem. Phys. Lett.*, 2004, **383**, 368–375.
57. D. Vanfleteren, D. Ghillemijn, D. Van Neck, P. Bultinck, M. Waroquier and P. W. Ayers, *J. Comput. Chem.*, 2011, **32**, 3485–3496.
58. D. Vanfleteren, D. Van Neck, P. Bultinck, P. W. Ayers and M. Waroquier, *J. Chem. Phys.*, 2010, **132**, 164111.
59. D. Vanfleteren, D. Van Neck, P. Bultinck, P. W. Ayers and M. Waroquier, *J. Chem. Phys.*, 2010, **133**, 231103.
60. A. M. K. Muller, *Phys. Lett.*, 1984, **105A**, 446–452.
61. R. J. Wheatley and A. A. Gopal, *Phys. Chem. Chem. Phys.*, 2012, **14**, 2087–2091.
62. C. Stanciu, M. E. Jones, P. E. Fanwick and M. M. Abu-Omar, *J. Am. Chem. Soc.*, 2007, **129**, 12400–12401.
63. N. Suzuki, T. Yamasaki and K. Motizuki, *J. Phys. Soc. Jpn.*, 1989, **58**, 3280–3289.
64. A. I. Baranov and M. Kohout, *J. Comput. Chem.*, 2011, **32**, 2064–2076.
65. R. W. Godby, G. A. Benesh, R. Haydock and V. Heine, *Phys. Rev. B*, 1985, **32**, 655–661.
66. G. Scalmani and M. J. Frisch, *J. Chem. Theory Comput.*, 2012, **8**, 2193–2196.
67. B. D. Dunnington and J. R. Schmidt, *J. Chem. Theory Comput.*, 2012, **8**, 1902–1911.
68. T. Watanabe and D. S. Sholl, *Langmuir*, 2012, **28**, 14114–14128.
69. I. Mayer, *J. Comput. Chem.*, 2007, **28**, 204–221.
70. E. Matito, M. Sola, P. Salvador and M. Duran, *Faraday Discuss.*, 2007, **135**, 325–345.
71. J. Cioslowski and S. T. Mixon, *J. Am. Chem. Soc.*, 1991, **113**, 4142–4145.
72. A. Michalak, R. L. DeKock and T. Ziegler, *J. Phys. Chem. A*, 2008, **112**, 7256–7263.
73. D. Y. Zubarev and A. I. Boldyrev, *Phys. Chem. Chem. Phys.*, 2008, **10**, 5207–5217.
74. T. R. Galeev, B. D. Dunnington, J. R. Schmidt and A. I. Boldyrev, 2013, **15**, 5022–5029.

CHAPTER 6

A Reaxff Reactive Force-field for Proton Transfer Reactions in Bulk Water and its Applications to Heterogeneous Catalysis

ADRI C.T. VAN DUIN,*[a] CHENYU ZOU,[a] KAUSHIK JOSHI,[a] VYASCHESLAV BRYANTSEV[b] AND WILLIAM A. GODDARD[b]

[a] Department of Mechanical and Nuclear Engineering, Penn State University, University Park, PA 16802, USA; [b] Material and Process Simulation Center, Beckman Institute, California Institute of Technology, Pasadena, CA 91125, USA
*Email: acv13@psu.edu

6.1 Introduction

Water is essential to the very existence of life, playing its important role in a myriad of physical, chemical, and biological processes. Despite having a simple molecular structure, it forms one of the most complex substances.[1–4] Specific interactions amongn water molecules in the condensed phase are responsible for its anomalous behavior. Extensive atomistic simulations have been performed to connect the microscopic structure of water to its macroscopic properties. The accuracy of simulation results strongly depends on the quality of the applied intermolecular potentials. To this end, more than 50 empirical water potentials have emerged in the literature, broadly differing in the

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

representation of the intra- and intermolecular energetics and the range of properties that can be reliably reproduced. Comprehensive reviews of ever-growing research effort in this field and comparison of the performance of several popular water potentials can be found in Refs. 5–11.

The existing interaction potentials for water vary from the simplest rigid pairwise additive three-site models to much more complicated flexible polarizable potentials with additional interaction sites, smeared charges and dipoles, fluctuating charges, Drude oscillators, *etc*. The classification of the potentials into pairwise additive (two-body) and polarizable (many-body) depends on whether the interaction energy for an assembly of water molecules contains only a sum of pair potentials or includes many-body correction terms resulting from the polarization of the monomer by its neighbors. The pair and polarizable potentials have their own merits and shortcomings. The former category of potentials implicitly incorporates many-body effects through the use of enhanced charges or monomer dipole moments. They are aimed at reproducing experimental results for the range of temperatures, pressures, and phases considered in their parametrization, and thus regarded as effective models having limited transferability to dissimilar environments (clusters, liquid water, ice). The latter category of potentials can be thought of as transferable potentials expected to perform well across different environments with regard to description of cooperative effects. However, the main practical drawback of this class of models, when fitted at the bottom of the potential well for water clusters, is the necessity to incorporate computationally expensive quantum dynamical simulation protocols for including intermolecular zero-point effects (ZPE) in order rigorously to derive macroscopic thermodynamic observables.[12]

There is growing interest in and need for generalizing water potentials to describe various phenomena in inhomogeneous environments, for example in the context of solvation of neutral and ionic solutes, solute transport, molecular structure of an organic–water interface, protolytic dissociation, *etc*. Owing to the highly complex nature of these processes, there still remains a question as to whether empirical potentials are able to reproduce the aforementioned processes with sufficient accuracy. It has been shown that using polarizable potential models for both water and monatomic inorganic ions is paramount for modeling solvation and transport of ions across organic–water and air–water interfaces.[13–15] Although the importance of the polarization and charge transfer in the polar and inhomogeneous environments has been recognized, it has been difficult to include these effects explicitly in the broadly applicable polarizable force field. Ongoing research in this area has been recently documented in a special issue.[16] The vast majority of simulations for biomolecular systems are still being performed using non-polarizable force field, which can be especially problematic for describing interactions involving ions, charged residues, and π-electron systems.

Most of the flexible water potentials employ stiff (harmonic) bond-stretching and angle-bending potentials and do not allow for bond forming and breaking. The solvation of single structures, either H_3O^+ or OH^-, in the aqueous solution and at interfaces has been modeled with reasonable success using

non-reactive potentials.[13–15] However, the important processes of H^+ and OH^- migration *via* a structural diffusion (hopping) through water cannot be described using such methods. In response to this deficiency, several empirical models for water that can dissociate have been published.[17–22] However, the dissociative properties of water (if any) have been illustrated only on small clusters and the applicability of these potentials to model charge migration and various protolytic reactions in the aqueous bulk has yet to be tested. Meanwhile, other interesting methodologies to model proton transfer reactions in aqueous solutions have emerged, with a water molecule represented as still undissociable but protonizable moiety.[23–30]

Voth and coworkers developed a multistate empirical valence bond (MS-EVB) method[23–25] to simulate proton solvation and transport in water and biomolecular systems. The EVB Hamiltonian matrix constructed on the basis of the chemically relevant protonation configurations is diagonalized in order to determine the excess proton delocalization and shuttling through water molecules for any given nuclear coordinates and time step in the molecular dynamics (MD) run. This approach was successfully applied for investigation of proton transport in bulk water, proton wires, proton channels of biomolecular systems, and polymer electrolyte membranes.[26] Lill and Helms introduced the stochastic proton hopping scheme (Q-HOP) that is capable of simulating transport of a single proton in aqueous solutions by combining classical MD simulations with instantaneous proton hopping events.[27] Proton transfer probability is calculated on the basis of the distance between donor and acceptor and the environmental influence of the surrounding groups. Surprisingly, these and other[28–30] developed schemes are not parameterized to describe the dissociation of water and the structural migration of the OH^-, and therefore do not provide dissociation pathways for a variety of base-catalyzed reactions in aqueous solution.

Many of the questions raised here can, in principle, be addressed in the framework of Car–Parrinello[31] and *ab initio* molecular dynamics where the parametrization of the potential is avoided by solving Hartree–Fock or Kohn–Sham equations at each time step.[32–36] However, such simulations become computationally intractable for systems with a few hundred water molecules and simulation times of more than 100 ps. Recent controversy on the structural and dynamic properties of liquid water, apart from the limited accuracy of the modern exchange-correlation functionals, raised the issue of using the proper simulation protocols with the sufficient length of equilibration and production runs and the combination of pseudopotentials and basis sets yielding converged results.[35,36] As such, they provide valuable benchmark data for validation of more approximate methods having much broader applicability for realistic and complex systems.

Our current research centers on the development of a computational approach to model chemical reactions in aqueous solution, which has been a long-standing and almost formidable challenge within the realm of empirical models. The important step toward this goal is to build effective potential for water that can allow for autoionization and intermolecular proton and hydroxyl transfer. In addition, the ability to model water dissociation, both homolytically and heterolytically, can provide all the necessary reaction pathways under different conditions.

The reactive force-field (ReaxFF) platform contains all essential ingredients accurately to model various dissociation pathways and formation of chemical bonds. Since the first publication in 2001 on hydrocarbons,[37] ReaxFF has been successfully applied for modeling complex chemical transformations involving high-energy materials,[38–40] silicon/silicon oxides,[41] polymer decompositions,[42] metals, and transition metals[43,44] in the gas phase, in the solid phase, and at the interface. In the present chapter we have extended the ReaxFF method to describe liquid water and reactions involving proton transfer in small clusters and the water phase. Since the ReaxFF water potential is fully compatible with our previous parametrization, considering the interaction between water and other molecules of interest will allow for reactive molecular dynamic simulations of a wide range of complex chemical processes in the aqueous phase.

6.2 Methods

6.2.1 Quantumchemical Methods

The structures and energies for a series of the clusters used in the training set were obtained with the Jaguar 7.0 program package using the X3LYP/6-311G** level of theory. All structures were fully optimized.

6.2.2 Force-field Optimization

The force-field parameterization was performed with the ReaxFF standalone program, using a single-parameter, parabolic approximation fitting procedure. Parameter correlation was accounted for by looping multiple times over the force-field parameters. Each parameter was allowed to modify within a set range, thus ensuring physically realistic and mathematically relevant final values. Forty-five different parameter combinations were obtained by fitting against the static DFT-data – these different combinations were obtained by modifying the weights of the individual training set components and re-parameterization. The results in this manuscript are all from parameter combination 41, which gave the best results for the bulk water dynamical (Section 6.3.2) and static properties (Section 6.3.1).

6.3 Results and Discussion

6.3.1 Force-field Development

The following sections describe the various QM-cases that were either used to parameterize the ReaxFF water description or to validate the potential.

6.3.1.1 Bond Dissociation, Angle Distortion and Charges

To ensure transferability to other ReaxFF descriptions we not only considered O–H bond dissociation but we also trained the ReaxFF parameters against

H–H dissociation in H_2 and O–O single- and double-bond dissociation in the H_2O_2 and O_2 molecules. For these three cases we derived adiabatic QM bond expansion/compression energies, where the expansion was taken all the way to the dissociative limit. Most weight was put on the QM-data close to the equilibrium and to the dissociation energies. Figure 6.1(a–c) shows the QM and ReaxFF results for these cases.

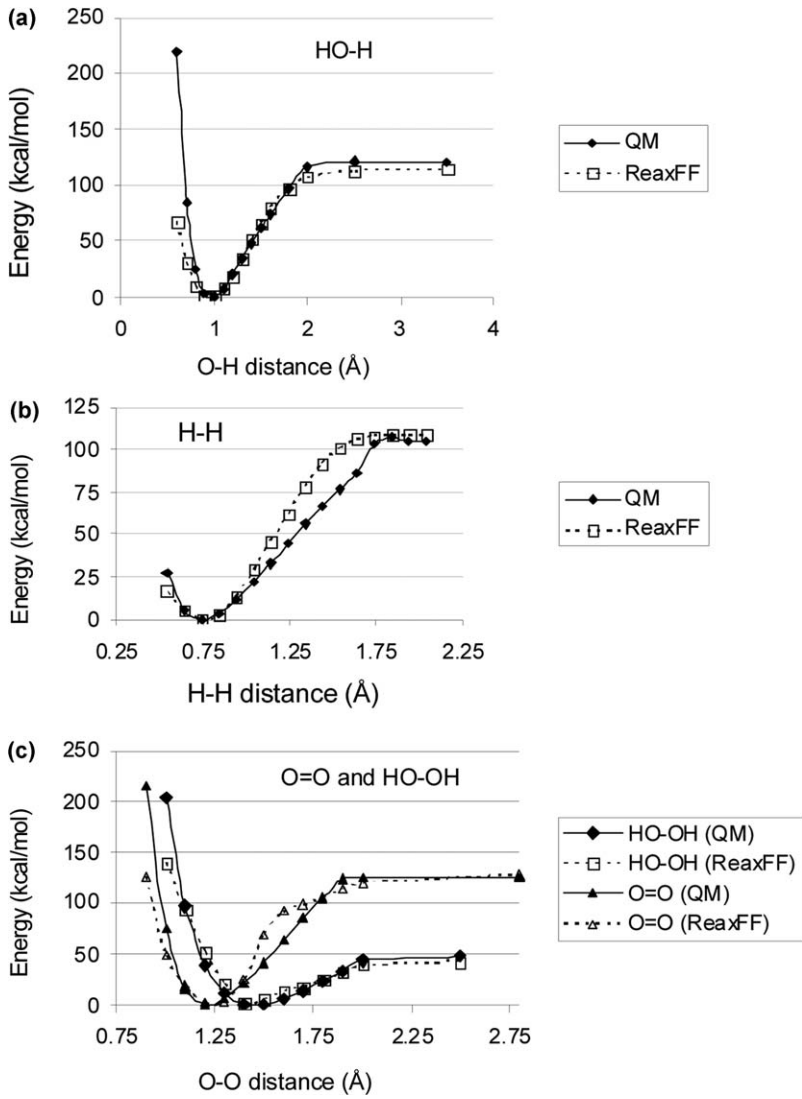


Figure 6.1 QM and ReaxFF bond dissociation energies for the O–H dissociation in H_2O (a), the H–H dissociation in H_2 (b), and the O–O single and double bond dissociation in H_2O_2 and O_2(triplet) (c).

To defined the ReaxFF angle parameters we performed a QM-simulation on the H_2O molecule where the H–O–H angle was adiabatically distorted from 72° to 134°. Figure 6.2 shows the QM- and ReaxFF results for this angle distortion energy. ReaxFF obtains an equilibrium angle of 103.3° for the isolated water molecule; this angle opens up to a value of 105.6° in the Ice(cmc)-phase.

Figure 6.3 compares the QM/Mulliken and ReaxFF average charges on the oxygen atoms for the H_2O-monomer up to the H_2O-decamer cases and for Ice(pna21). We put a relatively low weight on these Mulliken charges during the force-field optimization, which allowed the charges and polarization to also be determined by the QM-energy data. This resulted in ReaxFF charge distributions that are systematically more polarized than the QM/Mulliken charges. Figure 6.4 compares the ReaxFF and QM/Mulliken charges for the $H_3O^+/[H_2O]_6/OH^-$ cluster; for this cluster we find that both QM and ReaxFF

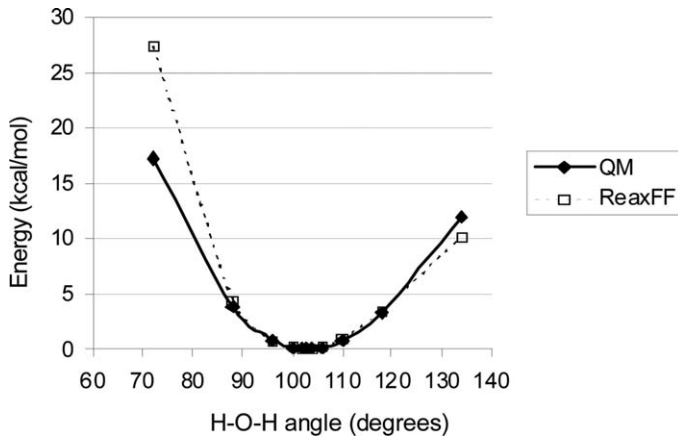


Figure 6.2 QM and ReaxFF H–O–H angle strain energies for H_2O.

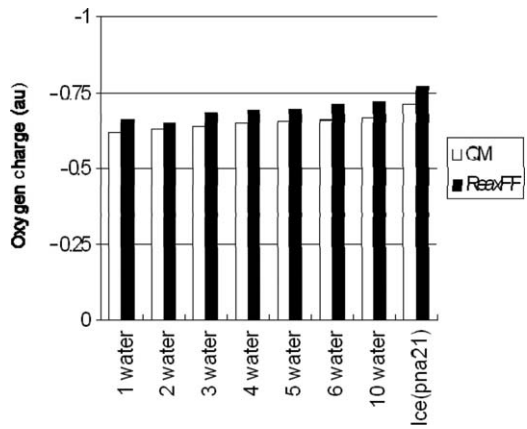


Figure 6.3 QM/Mullken charges and ReaxFF charges on oxygen for water clusters and for the Ice(pna21)-condensed phase.

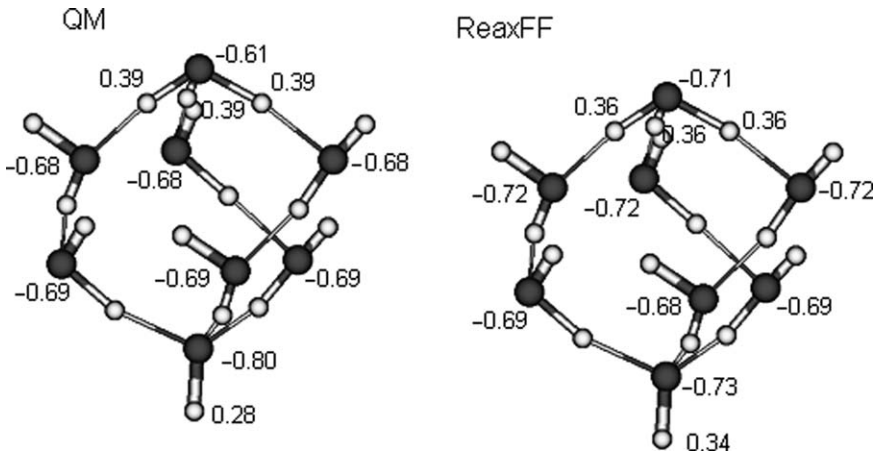


Figure 6.4 QM/Mullken and ReaxFF oxygen and cation/anion hydrogen charges for a $[H_3O]^+[OH]^-[H_2O]_6$-cluster.

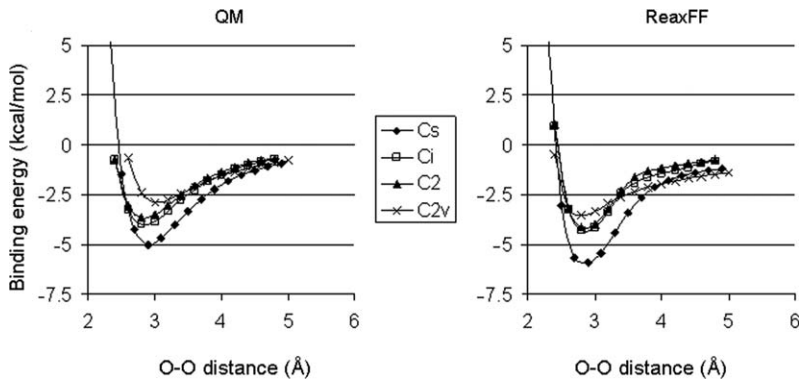


Figure 6.5 QM and ReaxFF water dimer binding energies as a function of oxygen–oxygen distance.

do not completely localize the positive and negative charges on the H_3O^+ cations or OH^- anions. This delocalization of the formal cation/anion charges is more substantial in ReaxFF compared with QM/Mulliken.

6.3.1.2 Water Clusters and Ice Condensed Phase Binding Energies

Figure 6.5 compares the ReaxFF and QM-results for water dimer dissociation. We included the C_s, C_i, C_2 and C_{2v} water dimer configurations in the ReaxFF training. Figure 6.4 shows that ReaxFF properly predicts the dimer stability order ($C_s > C_i, C_2 > C_{2v}$). ReaxFF overbinds the C_s dimer: it gives a binding energy of 5.9 kcal mol^{-1} where QM gives 5.0 kcal mol^{-1}. The ReaxFF C_s-dimer

has an equilibrium O–O distance of 2.9 Å, which is in exact agreement with the QM-geometry.

Figure 6.6 shows the ReaxFF and QM binding energies for $[H_2O]_n$ clusters, with $n=2$ to 34. While ReaxFF over-binds the H_2O-dimer it actually under-binds the $[H_2O]_3$ to $[H_2O]_{11}$ clusters. For larger clusters we find good agreement between ReaxFF and QM. The larger clusters were given a relatively high weight in the force-field fitting, as these are more relevant for a bulk water description. Figure 6.7 shows the QM and ReaxFF binding energies to $H_3O^+[H_2O]_n$ and $OH^-[H_2O]_n$ clusters. For both the positively as well as the negatively charged clusters we find that for $n=1$ there is a significant over-binding, related to overpolarization of the one solvating water molecule. With increasing water solvation the ReaxFF results agree well with the QM, indicating that ReaxFF should properly describe both H_3O^+ as well as OH^- solvation in bulk water.

We also tested ReaxFF for the Ice(cmc) and Ice(pna21) binding energies; Figure 6.8 shows that ReaxFF consistently underestimates the QM binding energies by about 10%. ReaxFF predicts an equilibrium density of 0.93 kg dm^{-3} for Ice(cmc) while QM gives an equilibrium density of 0.899 kg dm^{-3}.

6.3.1.3 *Reactions*

To test and train the ReaxFF parameters for reactive events we considered a number of hydrogen and proton transfer reactions for neutral, cation, and anion systems. Figure 6.9 and Table 6.1 compare the QM and ReaxFF barriers

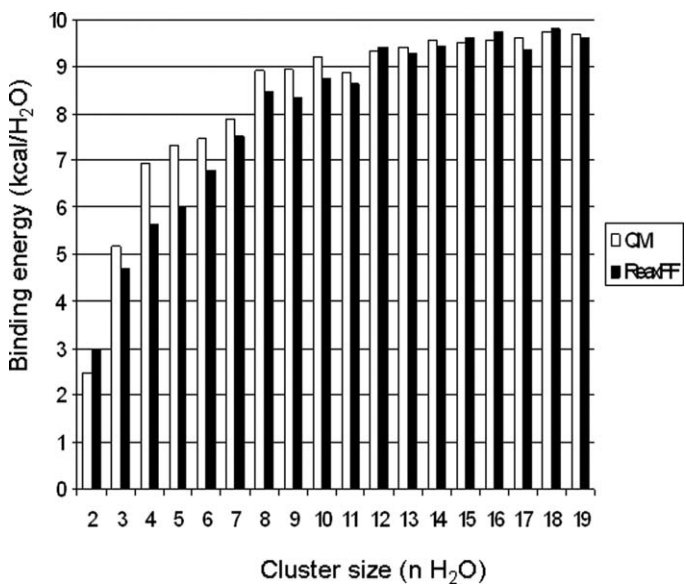


Figure 6.6 QM and ReaxFF water cluster binding energies as a function of cluster size.

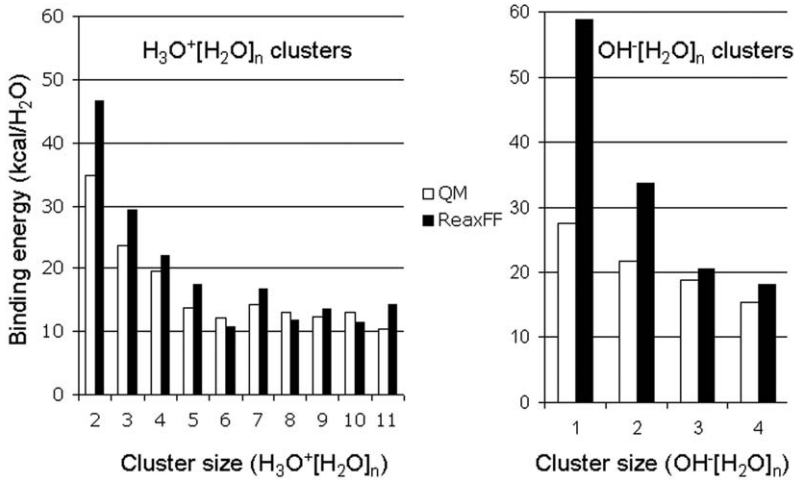


Figure 6.7 QM and ReaxFF water binding energies to $H_3O^+[H_2O]_n$ and $OH^-[H_2O]_n$ clusters.

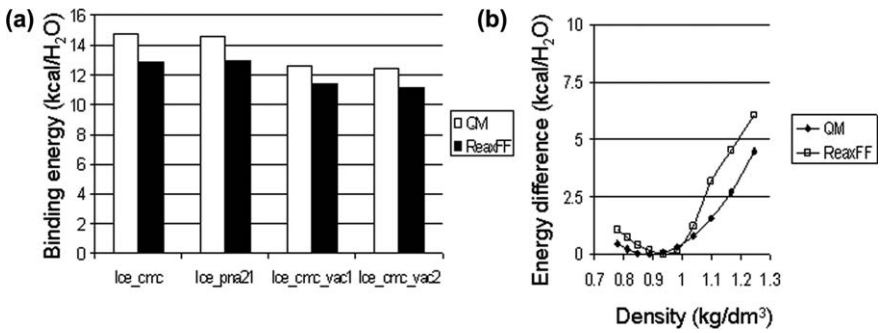


Figure 6.8 (a) QM and ReaxFF water binding energies in Ice(cmc), Ice(pna21) and Ice(cmc) with one water vacancy in two different locations. (b) QM and ReaxFF water energies of Ice(cmc) as a function of density.

for a concerted hydrogen shift reaction in $[H_2O]_n$ systems, where $n = 2$ to 6. These reactions are very relevant for water-catalyzed reactions in, for example, enzymes and on surface catalysis. We find good qualitative and quantitative agreement between ReaxFF and QM for these reactions; both methods agree that the four-water case has the lowest concerted hydrogen-shift barrier, while this barrier goes up substantially, owing to increase angle strain, when fewer water molecules are involved in the concerted event. Increasing the number of water molecules in the water-cycle from four to six leads to a modest increase in barrier.

To test the ReaxFF potential for proton transfer we considered a $H_3O^+ - H_2O$ dimer (Figure 6.10). In this system, we monitored the barrier for

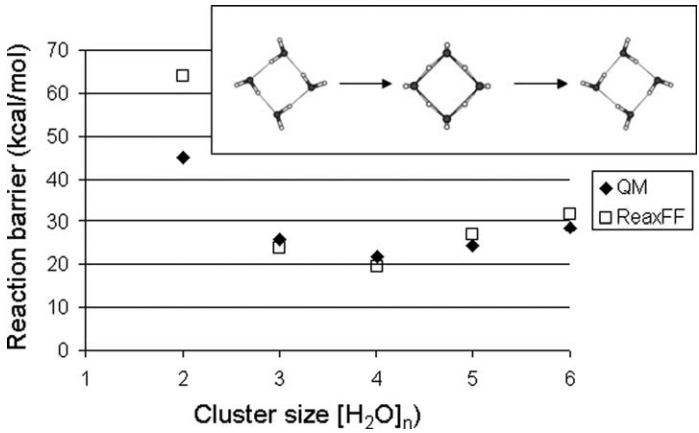


Figure 6.9 QM and ReaxFF barriers for concerted hydrogen transfer in $[H_2O]_n$ clusters. The inset shows the reaction for $n = 4$.

Table 6.1 QM and ReaxFF barriers (in kcal mol^{-1}) for concerted hydrogen transfer in $[H_2O]_n$ as a function of the number of water molecules.

n(H_2O)	*Barrier(QM)*	*Barrier(ReaxFF)*
2	45.0	64.0
3	25.7	23.8
4	21.9	19.4
5	24.5	27.0
6	28.6	31.9

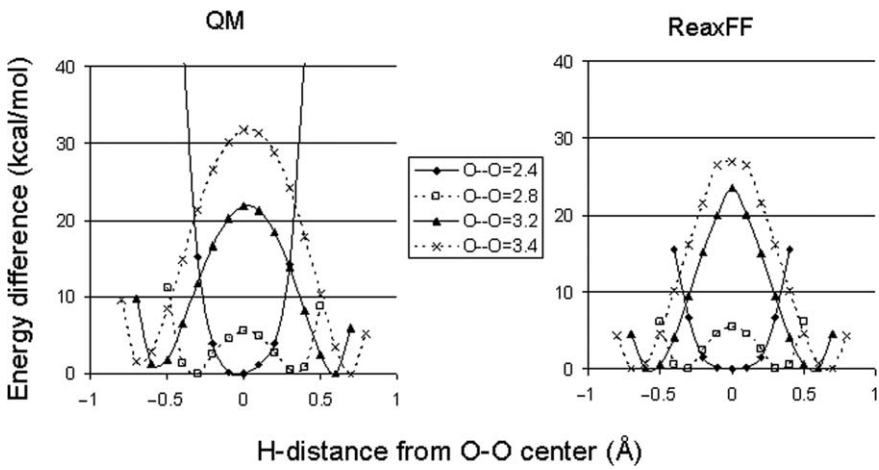


Figure 6.10 QM and ReaxFF barriers for proton migration as a function of O—O distance in an $[H_2O–H–OH_2]^+$ cluster.

proton transfer at different O–O distances. QM and ReaxFF agree that the global minimum for this dimer is at a O–O distance of 2.8 Å. For this configuration both ReaxFF and QM find proton transfer barriers of around 5 kcal mol^{-1}. Shortening the O–O distance to 2.4 Å increases the overall energy of the system but removes the barrier for proton transfer; at this O–O distance both QM and ReaxFF find that the most stable configuration is with the proton equidistant from both oxygen atoms. An increase in O–O distance to 3.2 and 3.6 Å results in a considerable increase in the proton transfer barrier.

Figure 6.11 shows a similar test case for an anionic system; here we transfer a proton in a OH^-–H_2O dimer. In this case the most stable configuration of the dimer has a O–O distance of 2.48 Å. In contrast to the H_3O^+–H_2O case discussed previously, for the anionic dimer the proton is shared between the oxygen atoms in the global energy minimum. At increased O–O distances a barrier to proton transfer appears; this barrier increase to more than 60 kcal mol^{-1} for an O–O distance of 3.6 Å.

The cases presented in Figures 6.9–6.11 were all considered in the ReaxFF parameterization. After the parameterization we performed two additional validation cases on proton transfer reactions. Figure 6.12 shows the ReaxFF proton transfer reaction energies for two different migration paths in an $H_3O^+[H_2O]_2$ system. In path 1, the H_3O^+ remains solvated by two water molecules, while path 2 results in a 1-fold solvated H_3O^+ that is about 3 kcal mol^{-1} higher in energy than the reactant. Figure 6.13 shows that ReaxFF predicts an automatic proton transfer in response to a solvation change, which is a key mechanism for proton transfers in bulk water.

6.3.2 Molecular Dynamics Simulations

6.3.2.1 Density and Cohesive Energy

In order to validate the ReaxFF bulk water description we performed a series of NVT/MD simulations at five different water densities (0.8, 0.9, 1.0, 1.1, and

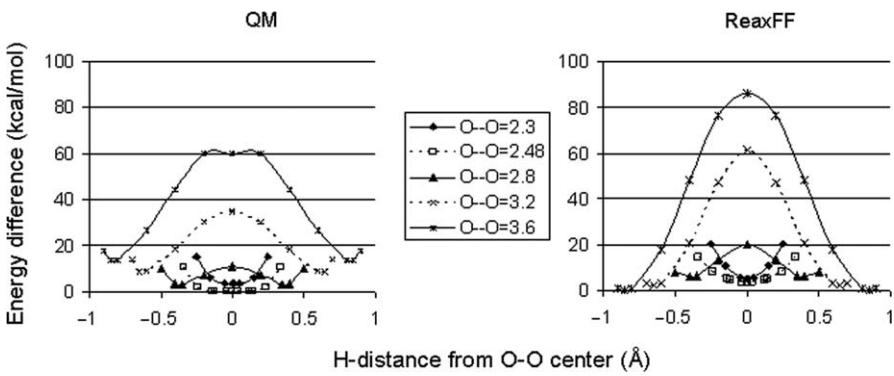


Figure 6.11 QM and ReaxFF barriers for proton migration as a function of O–O distance in an $[HO–H–OH]^-$ cluster.

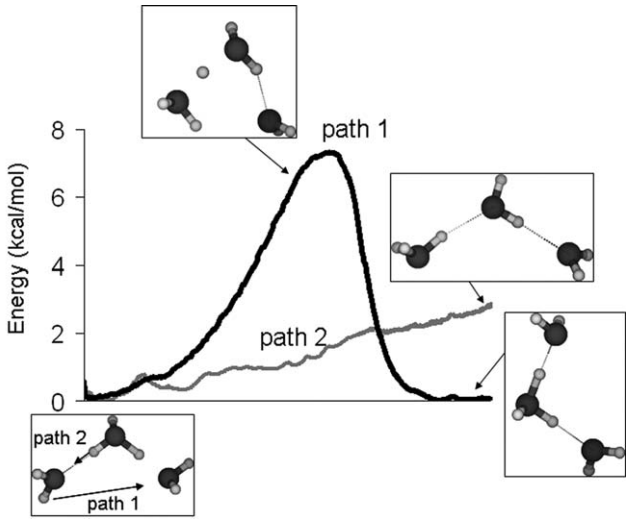


Figure 6.12 ReaxFF path for indirect (path 1) and direct (path 2) proton transfer in a $[H_2O–H_3O–H_2O]^+$ system.

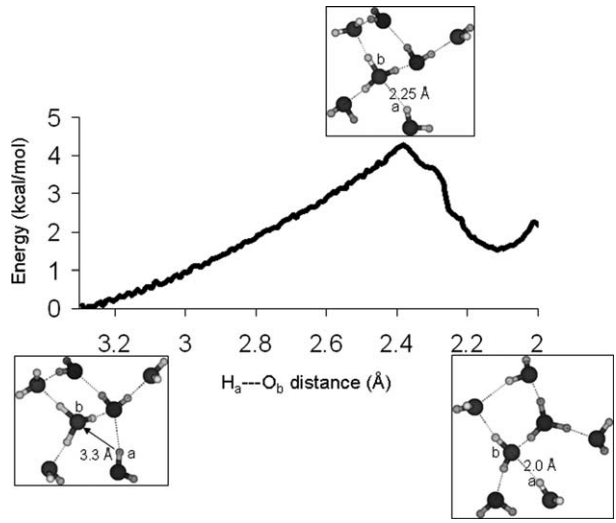


Figure 6.13 ReaxFF barrier for proton transfer resulting from a forced solvation change in a $[(H_2O)_6H_3O]^+$ system. External restraints were only imposed on the solvating waters.

1.2 kg dm^{-3}). These MD simulations were performed on periodic systems containing 100 water molecules using a Berendsen thermostat with a temperature damping constant of 100 fs. Figure 6.14 (right panel) shows the average potential energy per water molecule obtained from these simulations. We observe that the 1.0 kg dm^{-3} density produces the lowest average potential

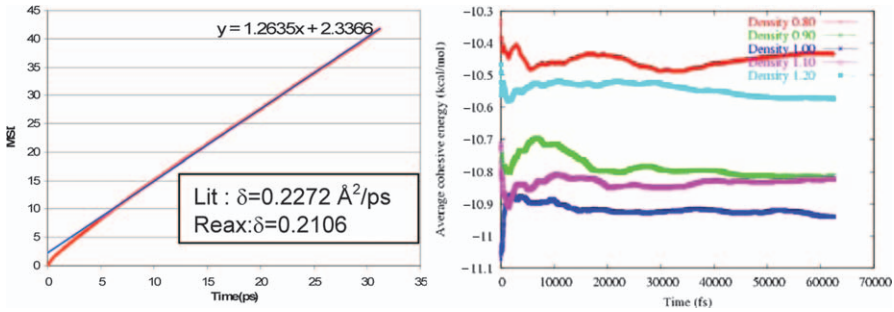


Figure 6.14 (left) Mean square displacement *vs.* time for an 800-water system at density = 1.00 kg dm^{-3} and $T = 300$ K. (right) Average cohesive energy versus density for an 800-water system.

energy (– 10.9 kcal per H_2O); this energy is in good agreement with the experimental heat of vaporization for water (– 10.5 kcal mol^{-1}). These results indicate that these ReaxFF water parameters can reproduce both the cohesive energy and the density of liquid water in an NVT-environment. In MD-simulations with an NPT-ensemble we observe a density decrease for liquid water at room temperature to a value of 0.94 kg dm^{-3}.

6.3.2.2 *Diffusion Constants*

To determine the water self-diffusion constant and the proton and anion diffusion constants we performed MD/NVE simulations at $T = 300$ K on a 100 water system, a 99 water/H_3O^+ and a 99 water/OH^- system at densities of 1.00 kg dm^{-3}. The 99 water/H_3O^+ and 99 water/OH^- were performed using a overall neutral simulation box. These simulations were performed for 1 ns with a time-step of 0.25 fs; unfolded system coordinates were saved every 250 iterations (62.5 ps). This yielded 4000 system configurations from which an averaged mean square displacement (MSD) was obtained. For the 100 H_2O system the MSD was obtained by analyzing and averaging the displacement of all the oxygen atoms; for the anion and proton systems both non-reactive and reactive diffusion was taken into account by locating the position of the OH^- or H_3O^+ molecule in the system box, using the ReaxFF bond order information, and by tracking the displacement of the oxygen atom in the anion or proton. From the MSD analysis for the proton and anion diffusion we can clearly observe jumps in the proton or anion coordinates, associated with a proton transfer reaction. Figure 6.14 (left panel) shows the MSD for the water, proton and anion diffusion, resulting in diffusion constants of, respectively, 0.2106 Å^2 ps^{-1}, 0.400 Å^2 ps^{-1} and 0.500 Å^2 ps^{-1}. While the water and proton diffusion constants are in good agreement with experimental values (0.2272 Å^2 ps^{-1} and 0.4200 Å^2 ps^{-1}),[45] we find that ReaxFF overestimates the anion diffusion constant (experiment: 0.3200 Å^2 ps^{-1}) by almost a factor of 2. We initially speculated that the ReaxFF over-binding between the OH^- and its solvation shell (see Figure 6.7) might have been responsible for this fast anion diffusion rate.

To test this hypothesis we increased the weights of the $OH^-/[H_2O]_n$ QM-data in the ReaxFF training set and re-parameterized the force-fields. While this resulted in a substantial improvement of the ReaxFF reproduction of the $OH^-/[H_2O]_n$ solvation energies, at the expense of the ReaxFF performance for other segments of the training set, MD-simulations with these re-optimized parameters did not result in a marked improvement of the anion diffusion rates.

6.3.2.3 *Radial Distribution Analysis*

Figure 6.15(a–c) compares the ReaxFF and experimental H–H, O–H and O–O radial distribution functions for a bulk liquid water system. These figures were obtained from a 100 ps 800-water NVE MD simulation at $T = 300$ K. System configurations were saved every 1 ps, resulting in 100 system configurations, from which the average Radial Distribution Function (RDF) were obtained. We observe for all three rdf-pairs that ReaxFF gives a good reproduction of the experimental data.[46] Of particular relevance for the quality of the bulk water structure obtained by ReaxFF is a accurate reproduction of the O–O radial distribution data, indicating that ReaxFF recognizes the optimal O–O distance for the first (2.8 Å), second (4.4 Å) and third (6.7 Å) solvation shells.

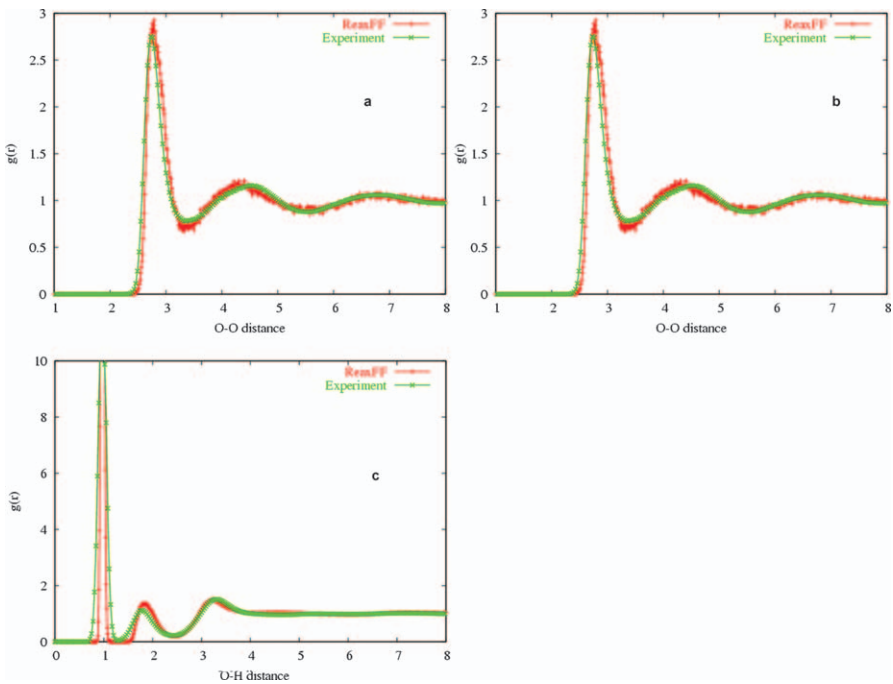


Figure 6.15 ReaxFF and experimental H/H (a), O/O (b), and O/H (c) radial distribution functions for bulk water. The ReaxFF radial distribution functions were obtained for an 800-water system using an NVE-ensemble at a density of 1.00 kg dm^{-3} and a temperature of 300 K.

6.3.2.4 Dynamical Properties as Direct Training Set Entries

The ReaxFF description of water, as given in this manuscript, was revised many times to incorporate different types of chemical reaction involving water. In all those revisions, the force-field optimization was restricted to static properties of water involving energy minimization calculations. However, there were no means to optimize for dynamic properties of water such as the diffusion coefficient, viscosity, *etc*. As such, we simply generated over 40 different parameter combinations that gave similar results for the static cases, tested these parameter combinations for dynamics properties and picked the O/H parameters that gave the best mixture of static and dynamic properties. In order to circumvent this time-consuming force-field development scheme in the future, we recently implemented python script, which allows the user to include both static and dynamic properties of water in force-field optimization calculations. For every force-field parameter which we want to optimize, the python script first performs molecular dynamics simulation to evaluate desired dynamical properties of water and calculates the force-field error. Following this, it performs regular energy minimization computations on all the geometries to evaluate the force-field error from static part. After finishing both MD and EM simulations, it combines both sets of force-field errors. Based on the total error, the script then searches for the optimum value of the parameter using a simple parabolic extrapolation search method, which tries to minimize the total force-field error. The detailed algorithm of the python wrapper is shown in Figure 6.16. This approach of including MD simulations in force-field

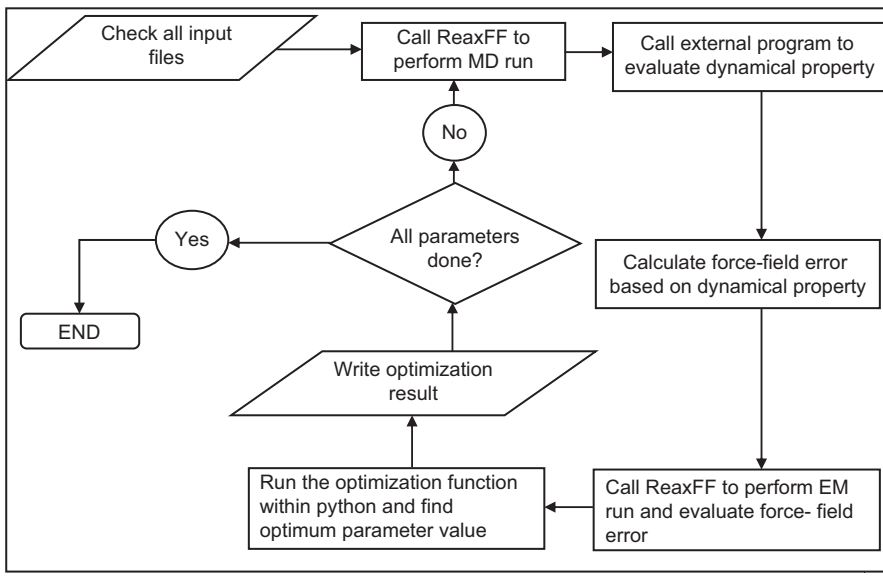


Figure 6.16 Schematic overview of a Python-based ReaxFF force-field development scheme, including dynamical properties.

optimization increases the time required for force-field parameterization, but it makes sure that dynamical properties are not compromised during force-field optimization.

6.3.3 Heterogeneous Catalysis

6.3.3.1 Overview of Applications of this ReaxFF Water Description

Water-related reactions are highly relevant to a broad range of heterogeneous catalysis systems, either as a direct contributor to the active site or as an exchange-reaction with the catalysis support, providing a reservoir of protons or hydroxyl-groups for reactions at the active site. Over the last three years, the ReaxFF water model presented here has been integrated in a number of metal/metal oxide descriptions relevant to heterogeneous catalysis. Our first successful metal oxide/water interface calculation involved a Zn/O/H ReaxFF description, which managed to reproduce experimentally observed water dissociation trends involving transitions between surface–zinc oxides and zinc hydroxides,[47,48] and was later used to simulate the collapse of Zn-based metal organic frameworks (MOFs) in the presence of water.[49] We also managed to integrate the O/H water parameters presented here with a previously published Si/O description,[39] enabling simulations of water dissociation on silica materials.[50] This silica/water ReaxFF description was recently extended with Ca/O/H parameters[51] and Al/O/H[52,53] parameters, enabling simulations on clays,[54] aluminum metal surfaces[52] and calcium silicates.[55] Besides the relatively straightforward Ca/H_2O interactions, we also managed to derive ReaxFF parameters for the more complex Cu cations in water; for these copper cations we managed to reproduce the non-classical Jahn–Teller distortion of Cu/water complexes.[56]

Parallel to these metal/metal oxide force-field development efforts, we also extended the first- and second-row elements that can incorporate the water description provided here. These extensions include carbon and nitrogen, in a ReaxFF description for glycine reactions in water,[57] sulfur, related to the catalytic cleavage of S–S bridges,[58] and phosphates.[59] Recently, we have merged the ReaxFF glycine description with a recently developed parameter set for titania/water interactions, enabling reactive simulations of peptides on titania surfaces.[60]

6.3.3.2 Heterogeneous Catalysis Application Example

The long-term structural evolution of catalytic surfaces is very important in the evaluation of new catalyst candidates. Studying this evolution is very difficult using solely *ab initio* methods, owing to their limited size and time scales. ReaxFF, however, provides a sufficient size and time range to study relatively slow structural and chemical changes in the surface and bulk composition of catalysts. An important example of a catalyst that undergoes structural changes

is the iron metal catalyst, which has been used in a wide range of hydrocarbon formation and conversion reactions. Iron metal has a relatively high chemical reactivity – it oxidizes readily and can easily integrate carbon and hydrogen atoms in its structure. As such, iron surfaces very rarely remain metallic during long-term catalytic use, and water-dissociation and -formation reactions provide a key reaction step in these iron surface reconstruction events. In order to simulate these events we have developed a Fe/O/H ReaxFF description,[61,62] using the O/H water parameters described in this chapter. This Fe/O/H description was trained against metal iron, iron oxide, and iron hydride DFT-data, thus covering a wide range of iron surface chemistry relevant to Fisher–Tropsch and water/gas shift catalytic conversions.

Figures 6.17 and 6.18 show an example of the capability of ReaxFF to describe the dynamic property of water formation and decomposition involving the oxidation of an Fe/FeH cluster. The cluster contains an iron hydride core with a diameter of 12 Å and a bare iron metal shell with a thickness of 6 Å. The total atom number of the cluster is 658 with a Fe/H ratio of 8.68: 1. The cluster was placed in the center of a 40 Å*40 Å*40 Å periodic box surrounded by 300 oxygen molecules (Figure 6.17). After energy minimization, the system was subjected to molecular dynamic simulation in the NVT ensemble for 250 ps at 1000 K. The temperature was controlled using a Berendsen thermostat with a damping constant of 100 fs. The oxygen and hydrogen atoms formed hydroxyl groups soon after the simulation started, as indicated in Figure 6.18. These hydroxyl groups can form water molecules while diffusing over the particle surface. Figure 6.18 also shows the continuous formation and dissociation of water molecules, indicating that the iron oxide surface is actively converted to a surface hydroxide during the simulation. These simulations indicate that ReaxFF can describe the changes in catalytic structure as a function of gas-phase composition.

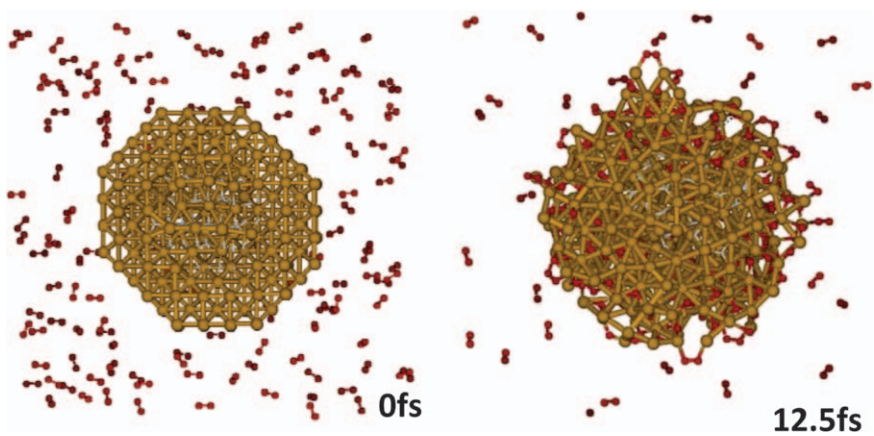


Figure 6.17 Cross-sections of the initial (left) and final (right) system composition in the iron hydride/O_2 molecular dynamics simulations.

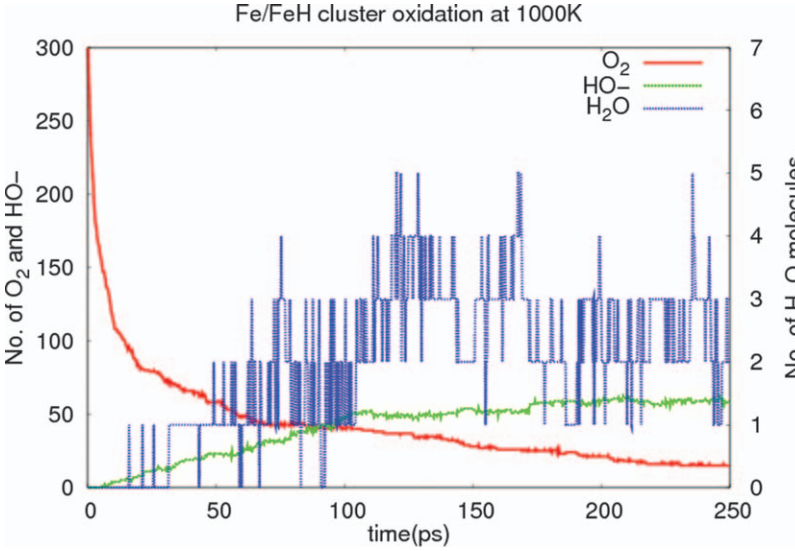


Figure 6.18 Analysis of the O_2(g), surface hydroxyl groups and gas-phase water molecules as observed during the ReaxFF MD-simulation on iron–hydride conversion in O_2.

6.4 Conclusions

By parameterization against a large QM-based training set, consisting of bond dissociation energies, angle scans, water-dimer and oligomer binding energies, H_3O^+/water and OH^-/water oligomer binding energies, proton migration barriers and ice binding energies, and equations of state, we have developed a ReaxFF reactive force-field capable of describing proton migration in water. This force-field was tested in molecular dynamics simulations and was found to give an excellent reproduction of bulk water properties, including density, cohesive energy, self-diffusion and structural features. We also found good agreement with experiment for the proton diffusion coefficient, which includes non-reactive as well as reactive aspects; however, we found that the force-field over-predicts the OH^- diffusion. Given that this ReaxFF description can also describe H–H and O–O single and double bonds it can be combined with existing ReaxFF descriptions, enabling large-scale ($>1\ 000\ 000$ atoms) nanosecond-scale simulations on reactive processes in the water phase. Based on the O/H ReaxFF parameters presented here we have already developed ReaxFF parameters for aqueous-phase simulations on proteins, phosphates, sulfonates, silicates and homogeneous and heterogeneous zinc/copper/iron systems.

References

1. D. Eisenberg and W. Kauzmann, *The Structure and Properties of Water*, Oxford University Press, Oxford, 1969.

2. G. W. Robinson, S.-B. Zhu, S. Singh and M. W. Evans, *Water in Biology, Chemistry and Physics, Experiment Overviews and Computational Methodologies*, World Scientific, Singapore, 1996.
3. F. Franks, *Water and Matrix of Life*, Royal Society of Chemistry, Cambridge, 2000.
4. M. Chaplin, *Water Structure and Science*, http://www.lsbu.ac.uk/water/, 2007.
5. A. Wallqvist and R. D. Mountain, Molecular models of water: derivation and description, in *Reviews in Computational Chemistry*, ed. K. B. Lipkowitz and D. B. Boyd, Wiley, New York, 1999, p. 183.
6. B. Guillot, *J. Mol. Liq.*, 2002, **101**, 219.
7. S. S. Xantheas, *Struct. Bond*, 2005, **116**, 119.
8. C. Vega, C. McBride, E. Sanz and J. L. F. Abascal, *Phys. Chem. Chem. Phys.*, 2005, **7**, 1450.
9. C. Vega and J. L. F. Abascal, *J. Chem. Phys.*, 2005, **123**, 144504.
10. M. W. Mahoney and W. L. Jorgensen, *J. Chem. Phys.*, 2000, **112**, 8910.
11. Y. Wu, H. L. Tepper and G. A. Voth, *J. Chem. Phys.*, 2006, **124**, 024503.
12. G. S. Fanourgakis, G. K. Schenter and S. S. Xantheas, *J. Chem. Phys.*, 2006, **125**, 141102.
13. M. Mucha, T. Frigato, L. M. Levering, H. C. Allen, D. J. Tobias, L. X. Dang and P. Jungwirth, *J. Phys. Chem. B*, 2005, **109**, 7617.
14. R. Vacha, V. Buch, A. Milet, J. P. Devlin and P. Jungwirth, *Phys. Chem. Chem. Phys.*, 2007, **9**, 4736.
15. C. D. Wick, I.-F. W. Kuo, C. J. Mundy and L. X. Dang, *J. Chem. Theory Comput.*, 2007, **3**, 2002.
16. W. L. Jorgensen, *J. Chem. Theory Comput.*, 2007, **3**, 1877.
17. F. H. Stillinger and C. W. David, *J. Chem. Phys.*, 1978, **69**, 1473.
18. C. W. David, *J. Chem. Phys.*, 1996, **104**, 7225.
19. J. W. Halley, J. R. Rustad and A. Rahman, *J. Chem. Phys.*, 1993, **98**, 4110.
20. L. R. Corrales, *J. Chem. Phys.*, 1999, **110**, 9071.
21. E. Lussetti, G. Pastore and E. Smatgiassi, *Chem. Phys. Lett.*, 2003, **381**, 287.
22. S. H. Garofalini and T. S. Mahadevan, *J. Phys. Chem. B*, 2007, **111**, 8919.
23. U. W. Schmitt and G. A. Voth, *J. Chem. Phys.*, 1999, **111**, 9361.
24. T. J. F. Day, A. V. Soudackov, M. Cuma, U. W. Schmitt and G. A. Voth, *J. Chem. Phys.*, 2003, **117**, 5839.
25. F. Wan and G. A. Voth, *J. Chem. Phys.*, 2005, **122**, 144105.
26. G. A. Voth, *Acc. Chem. Res.*, 2006, **39**, 143.
27. M. A. Lill and V. Helms, *J. Chem. Phys.*, 2005, **115**, 7993.
28. S. R. Billeter and W. F. van Gunsteren, *J. Phys. Chem. A*, 1998, **102**, 4669.
29. S. R. Billeter and W. F. van Gunsteren, *J. Phys. Chem. A*, 2000, **104**, 3276.
30. S. R. Billeter, S. P. Webb, T. Iordanov, P. K. Agarwal and S. Hammes-Schiffer, *J. Chem. Phys.*, 2001, **114**, 6925.
31. P. Car and M. Parrinello, *Phys. Rev. Lett.*, 1985, **55**, 2471.
32. D. Marx, *ChemPhysChem*, 2006, **7**, 1848.

33. P. L. Geissler, C. Dellago, D. Chandler, J. Hutter and M. Parribello, *Science*, 2001, **291**, 2121.
34. T. Todorova, A. P. Seitsonen, J. Hutter, I.-F. W. Kuo and C. J. Mundy, *J. Phys. Chem. B*, 2006, **110**, 3685.
35. P. H.-L. Sit and N. Marzari, *J. Chem. Phys.*, 2005, **122**, 204510.
36. H.-S. Lee and M. E. Tuckerman, *J. Chem. Phys.*, 2007, **126**, 164501.
37. A. C. T. van Duin, S. Dasgupta, F. Lorant and W. A. Goddard, *J. Phys. Chem. A*, 2001, **105**, 9396.
38. A. C. T. van Duin and J. S. S. Damste, *Organ. Geochem.*, 2003, **34**, 515.
39. A. C. T. van Duin, A. Strachan, S. Stewman, Q. S. Zhang, X. Xu and W. A. Goddard, *J. Phys. Chem. A*, 2003, **107**, 3803.
40. Q. Zhang, T. Cagin, A. van Duin, W. A. Goddard, Y. Qi and L. G. Hector, *Phys. Rev. B*, 2004, **69**, 045423.
41. A. Strachan, A. C. T. van Duin, D. Chakraborty, S. Dasgupta and W. A. Goddard, *Phys. Rev. Lett.*, 2003, **91**, 098301.
42. A. Strachan, E. M. Kober, A. C. T. van Duin, J. Oxgaard and W. A. Goddard, *J. Chem. Phys.*, 2005, **122**, 054502.
43. K. D. Nielson, A. C. T. van Duin, J. Oxgaard, W. Q. Deng and W. A. Goddard, *J. Phys. Chem. A*, 2005, **109**, 493.
44. S. Cheung, W. Q. Deng, A. C. T. van Duin and W. A. Goddard, *J. Phys. Chem. A*, 2005, **109**, 851.
45. D. Eisenberg and W. Kauzman, *The Structure and Properties of Water*, Oxford University Press, Oxford, 1969.
46. A. K. Soper and P. J. Rossky, *Chemical Physics, Special Issue*, 2000, **258**, 121–137.
47. D. Raymand, A. C. T. van Duin, D. Spangberg, W. A. Goddard and K. Hermansson, *Surf. Sci.*, 2010, **604**, 741.
48. D. Raymand, A. C. T. van Duin, W. A. Goddard, K. Hermansson and D. Spangberg, *J. Phys. Chem. A*, 2011, **115**, 8573.
49. S. S. Han, S.-H. Choi and A. C. T. van Duin, *Chem. Commun.*, 2010, **46**, 5713.
50. J. C. Fogarty, H. M. Aktulga, A. Y. Grama, A. C. T. van Duin and S. A. Pandit, *J. Chem. Phys.*, 2010, **132**, 174704/1.
51. H. Manzano, R. Pellenq, F.-J. Ulm, M. J. Buehler and A. C. T. van Duin, *Langmuir*, 2012, **28**, 4187.
52. M. Russo, R. Li, M. Mench and A. C. T. van Duin, *Int. J. Hydr. Energ.*, 2011, **36**, 5828.
53. B. Narayanan, A. C. T. van Duin, B. D. Kappes, I. E. Reimanis and C. Ciobanu, *Model. Simul. Mat. Sci. Engin.*, 2012, **20**, 015002/1.
54. M. C. Pitman and A. C. T. van Duin, *J. Am. Chem. Soc.*, 2012, **134**, 3042.
55. H. Manzano, F.-J. Ulm, A. van Duin, R. Pellenq, F. Marinelli and S. Moeni, *J. Am. Chem. Soc.*, 2012, **134**, 2208.
56. A. C. T. van Duin, V. S. Bryantsev, M. S. Diallo, W. A. Goddard, O. Rahaman, D. Doren, J. Raymand and D. Hermansson, *J. Phys. Chem. A*, 2010, **114**, 9507.

57. O. Rahaman, A. C. T. van Duin, W. A. Goddard, III and D. J. Doren, *J. Phys. Chem. B*, 2011, **115**, 249.
58. S. Keten, C.-C. Chou, A. C. T. van Duin and M. J. Buehler, *J. Mech. Behav. Biomed. Mater.*, 2012, **5**, 32.
59. J. Quenneville, R. S. Taylor and A. C. T. van Duin, *J. Phys. Chem. C*, 2010, **114**, 18894.
60. S. Monti, A. C. T. van Duin, S.-Y. Kim and V. Barone, *J. Phys. Chem.*, 2012, **116**, 5141.
61. M. Aryanpour, A. C. T. van Duin and J. D. Kubicki, *J. Phys. Chem. A*, 2010, **114**, 6298.
62. C. Zou, A. C. T. van Duin and D. Sorescu, *Topics Catal.*, in press.

CHAPTER 7

Charge Transfer Potentials

YU-TING CHENG, TAO LIANG, SIMON R. PHILLPOT AND SUSAN B. SINNOTT*

Department of Materials Science and Engineering, University of Florida, Gainesville, FL, 32611-6400
*Email: ssinn@mse.ufl.edu

7.1 Introduction

The last two decades have witnessed a dramatic rise in computational resources that has facilitated tremendous progress in computational science. In particular, this progress has enabled the application of quantum-based methods such as Hartree–Fock (HF) theory and density functional theory (DFT) to compute the potential energy surfaces of numerous complex reactions that are critical to understanding catalytic reactions. These approaches provide high fidelity because of their explicit treatment of electronic structure; however, their computational cost increases rapidly with system size. Therefore, they are limited to a relatively small number of atoms (<500). To overcome this limitation, classical empirical methods (also known as interatomic potentials) that model molecules and materials at the atomic scale without explicitly treating electrons have been developed and have been employed in molecular dynamics (MD) and Monte Carlo (MC) simulations. Such simulations have been employed to examine catalysis at length and time scales beyond the reach of quantum-based approaches.

The main strength of classical empirical potentials is their low computational cost relative to electronic-structure calculations. Recently, systems with billions of atoms have been modeled in MD simulations.[1] In MD simulations, the motions of the particles in a system are predicted by solving Newton's equation of motion. Therefore, it is important that the interatomic potential, which is

RSC Catalysis Series No. 14
Computational Catalysis
Edited by Aravind Asthagiri and Michael J. Janik

Published by the Royal Society of Chemistry, www.rsc.org

used to calculate energies and forces, is accurate and transferable in order to describe correctly the chemical and physical properties of materials under a variety of conditions. Over the last few decades, empirical potentials have been successfully developed that are specific to the type of chemical bonding within a given system. For example, Lennard–Jones (LJ) potentials[2,3] are typically used to characterize van der Waals (vdW) interactions; embedded atom method (EAM) potentials[4,5] are typically used to describe metallic bonding; and Buckingham potentials[6,7] are typically used to model ionically bonded systems. Although these potentials and many others[8,9] describe the specific interactions for which they are parameterized very well, an obvious limitation is that most empirical interatomic potentials are applicable to only one type of bonding, which prevents their application to multicomponent systems with different bonding types. In most cases, as illustrated in Figure 7.1 below, catalytically driven chemical reactions involve several different interactions such as recant–metal and metal–metal oxide that may contribute to overall catalytic activity and selectivity. Therefore, a great deal of effort has been invested in recent years in developing a robust potential scheme that is applicable to multiple types of bonding within a given system. Such a scheme has been built on two key developments.

The first key development was the bond-order type of interatomic potential, such as the Abell potential.[10] Because the bond-order function not only takes the distances among atoms into account but also their local atomic

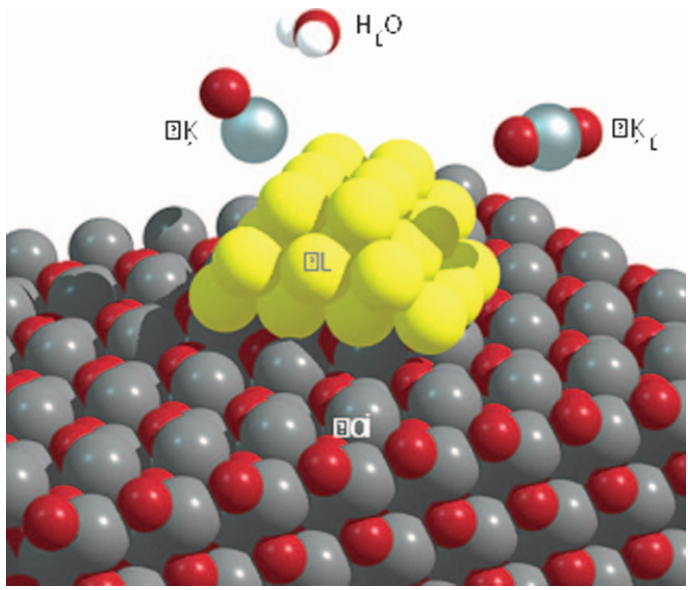


Figure 7.1 Snapshot of a molecular dynamics simulation of water, carbon monoxide, and carbon dioxide molecules interacting with a Cu catalyst supported by a zinc oxide substrate. The elements of Zn (gray), O (red), Cu (yellow), C (light blue) and H (white) are distinguished by color.

environment such as angles, bonding type and number, and chemical identities of other surrounding atoms, it is among the most flexible formalisms in common use today and has been adapted with slight modification to covalently bound semiconductors, hydrocarbons, and metals. Today, the Tersoff bond-order potential[11–13] is one of the most widely used for silicon. Building on the work of Tersoff, Brenner[14] extended the formalism to radical and conjugated hydrocarbon bonds by introducing additional terms to the bond-order function for hydrocarbon systems. This potential is known as the Brenner potential or the reactive empirical bond order (REBO) potential. The reactive term in the name indicates that this potential has the capacity to model bond breaking and bond formation. The second-generation REBO potential (REBO2)[15] provided even more accurate descriptions of short-range bonding in solid state carbon materials and hydrocarbon systems, and was subsequently further extended to C–H–O,[16–18] C–H–F,[19] and C–H–S[20] systems. Most recently, Liang *et al.*[21] parameterized REBO2 to model the metallic, ionic and covalent bonding present in Mo–S systems. While bond-order type potentials have been successfully applied to numerous systems, the traditional bond-order formalism lacks electrostatic interactions, which leads to difficulties in examining the effect of charge on reactions and limits their ability to describe chemical changes in a multicomponent system with dominant ionic features.

The second key development for a robust potential scheme applicable to multiple types of bonding was a method for equilibration of charge. Although fixed-charged potentials work well for bulk material systems, they cannot describe systems that require charge equilibration, such as a heterogeneous interface between dissimilar materials. Variable-charge electrostatics, in which the partial charges of individual atoms are not fixed but instead are determined in a self-consistent manner based on the principle of electronegativity equilibration (EE),[22–25] compensates for this limitation. Because variable electrostatics only describes electrostatic energies between interacting ion pairs, it is normally coupled with a charge-free potential to describe the non-electrostatic energy to the total energy; this was implemented in the Electrostatics + approach (ES +) of Streitz and Mintmire.[26] In particular, they successfully coupled the charge equilibration (QEq) approach to the EAM potential to model Al/Al_2O_3 interfaces, at which an Al atom is charge neutral in the metal but has an ionic charge in the oxide phase.

Yasukawa[27] coupled the bond-ordered potential (extended-Tersoff potential) for short-range interactions with the QEq variable charge approach for Coulomb electrostatics to describe successfully Si/SiO_2 interfacial systems. Subsequently, Iwasaki and Miura[28] explicitly employed the Yasukawa potential to simulate metal–ceramic interfaces and found that the predicted adhesion energies are in excellent agreement with experiment. Building on these developments, Yu *et al.*[29] developed the charge-optimized many-body (COMB) formalisms. After several extensions and modifications of the formalism, the COMB potential has been successfully parameterized for Si/SiO_2,[30] Hf/HfO_2,[31] Cu/Cu_2O,[32] Cu/ZnO,[33] and hydrocarbon systems[34] and has been used to address problems that require accurate modeling of chemistry, such as oxidation

and surface modification. In addition, an important feature as well as a significant challenge for the COMB potential is that all the element-specific parameters are kept the same for all bond types in metallic, organic, and ionic materials. It thus provides a flexible and transferable interatomic potential to model multicomponent systems without time-consuming parameterization.

7.2 Variable Charge Reactive Potentials: COMB Potentials

7.2.1 A General Form of the COMB Potentials

This section presents the general formula of the COMB potentials and describes how they may be used to obtain a proper description of energy and charge for a multicomponent system. The complete details of the evolution and latest generation of COMB potentials are given in Refs. 34 and 35. As shown in eqn (7.1), the total potential energy ($U^{\text{tot}}[\{q\},\{r\}]$) of a system as described by the COMB potential is composed of electrostatic energies ($U^{\text{es}}[\{q\},\{r\}]$), short-range interactions ($U^{\text{short}}[\{q\},\{r\}]$), van der Waals interactions ($U^{\text{vdW}}[\{r\}]$), and correction terms ($U^{\text{corr}}[\{r\}]$), where $\{q\}$ and $\{r\}$ represent the charge array and coordinate array of the system, respectively.

$$U^{\text{tot}}[\{q\},\{r\}] = U^{\text{es}}[\{q\},\{r\}] + U^{\text{short}}[\{q\},\{r\}] + U^{\text{vdW}}[\{r\}] + U^{\text{corr}}[\{r\}] \quad (7.1)$$

7.2.2 Electrostatic Energies

The electrostatic energies ($U^{\text{es}}[\{q\},\{r\}]$) include the self-energy ($U^{\text{self}}[\{q\},\{r\}]$), the charge–charge interactions ($U^{qq}[\{q\},\{r\}]$), the charge–nucleus interactions ($U^{qZ}[\{q\},\{r\}]$), and the energies associated with atomic polarizability ($U^{\text{polar}}[\{q\},\{r\}]$), which are defined in eqn (7.2). The repulsion between nuclei is not explicitly formulated because the overall effects are assumed to be included in the charge-dependent short-range interactions ($U^{\text{short}}[\{q\},\{r\}]$).

$$U^{\text{es}}[\{q\},\{r\}] = U^{\text{self}}[\{q\},\{r\}] + U^{qq}[\{q\},\{r\}] + U^{qZ}[\{q\},\{r\}] + U^{\text{polar}}[\{q\},\{r\}] \quad (7.2)$$

The self-energy ($U^{\text{self}}[\{q\},\{r\}]$) describes the energy required to form a charge on an atom. This can be expressed by the summation of the ionization or electron affinity energy of an isolated atom ($V^{\text{ionize}}(q_i)$) and a correction function ($\text{V}^{\text{field}}(r_{ij},q_i)$), termed the field effect,[36] that reflects the change of electronegativity and atomic hardness of the atom with its environment, as specified in eqns (7.3) to (7.5):

$$U^{\text{self}}[\{q\},\{r\}] = E_i^0(0) + \chi_i q_i + \left(J_i + \sum_{j\neq i}^{N} V_{ij}^{\text{field}}(r_{ij},q_j)\right) q_i^2 + K_i q_i^3 + L_i q_i^4 \quad (7.3)$$

$$V_i^{\text{ionize}}(q_i) = E_i^0(0) + \chi_i q_i + J_i q_i^2 + K_i q_i^3 + L_i q_i^4 \quad (7.4)$$

$$V_{ij}^{\text{field}}(r_{ij}, q_j) = \frac{1}{2\pi\varepsilon_0} \sum_{j \neq i}^{NN} \left(\frac{P_{ij}^{\chi} q_j}{r_{ij}^3} + \frac{P_{ij}^{J} q_j^2}{r_{ij}^5} \right) \tag{7.5}$$

In eqn (7.4), as outlined by Toufar *et al.*,[36] the parameters χ is identified with the electronegativity, while J is associated with chemical hardness or self-Coulombic interactions,[25] which are inherent to each element. Therefore, χ_i, J_i, K_i, and L_i are treated as atomic parameters and fitted to the ionization energies and electron affinities for each of the elements. In contrast, the parameters P_{ij}^{χ} and P_{ij}^{J} in eqn (7.5) are associated with specific bonds to describe the environmental dependence of field effects.

The long-range electrostatic interaction between two ions is normally described by Coulomb's Law, which is defined in eqn (7.6):

$$U_{ij}^{\text{Coul}}(q_i, q_j) = \frac{1}{4\pi\varepsilon_0} \frac{q_i q_j}{r_{ij}} \tag{7.6}$$

However, as the distance between two ions approaches zero, the Coulomb expression in eqn (7.6) results in infinite values, which has been called the "Coulomb catastrophe". To avoid the Coulomb catastrophe which is introduced by a point charge model, COMB potentials adopt the Streitz–Mintmire[26] charge density function. Specifically, the charge density of an atom is taken to be a function of its charge, spatial location (r) and atomic position (r_i):

$$\rho_i(r, q_i) = Z_i \delta(|r - r_i| + (q_i - Z_i) f_i(|r - r_i|)) \tag{7.7}$$

$$f_i(|r - r_i|) = \xi_i^3 \pi^{-1} \exp(-2\xi_i |r - r_i|) \tag{7.8}$$

where Z_i is an effective point core charge treated as a fitted parameter, δ is the Dirac delta function, $f(r)$ is a function that captures the radial decay of the electron density of the S-type orbital,[37] and ξ is an orbital exponent that controls the length scale associated with this decay. Further, the charge–charge interactions ($U^{qq}[\{q\},\{r\}]$) can be calculated as indicated in eqn (7.9) by integration over electron densities between pairs of atoms through the Coulomb integral operator, J_{ij}^{qq}, defined in eqn (7.10).

$$U^{qq}[\{q\}, \{r\}] = \sum_i \sum_{j \neq i} q_i J_{ij}^{qq} q_j \tag{7.9}$$

$$J_{ij}^{qq} = [\rho_i | \rho_j] = \int d^3 r_1 \int d^3 r_2 \frac{\rho_i(r_1) \rho_j(r_2)}{r_{12}} \tag{7.10}$$

where r_1 and r_2 indicate the centers of $\rho_i(r)$ and $\rho_j(r)$, and r_{12} is the distance between density distributions. Similarly, the charge–nucleus interactions ($U^{qZ}[\{q\},\{r\}]$) are calculated in eqn (7.11) through the charge–nuclear coupling operator, J_{ij}^{qZ}, which is a shielded nuclear attraction integral and defined in eqns (7.12) and (7.13):

$$U^{qZ}[\{q\}, \{r\}] = \sum_i \sum_{j \neq i} \left(q_i J_{ij}^{qZ} q_j + q_j J_{ji}^{qZ} q_i \right) \tag{7.11}$$

$$J_{ij}^{qZ} = [j|\rho_i] - [\rho_j|\rho_i] \tag{7.12}$$

$$[j|\rho_i] = \int d^3r \frac{\rho_i(r)}{|r - r_i|} \tag{7.13}$$

The Coulombic interactions in eqns (7.9) and (7.11) are solved *via* a direct Wolf summation method,[38] which is more efficient than the Ewald summation[39] for computing the electrostatic interactions.

The polarizability P represents the relative tendency of an equilibrium charge distribution to be distorted in response to an external field, which is defined as the ratio of the induced dipole moment, $\vec{\mu}$, of an atom to the electric field, $\vec{E}$. The inclusion of the polarization effects has been proven to be useful in classical simulations to stabilize complex oxide structures such as α-alumina[40] and to model the interactions with small polarizable molecules such as water and oxygen molecules. Using the fluctuation dipole model by Sterne *et al.*,[41] as specific in eqn (7.14), the polarization vector can be directly calculated from the electrostatic field generated by the atomic charges, $\vec{E}_i^q$, and the neighboring induced dipoles:

$$\vec{\mu}_i = P_i \vec{E}(\vec{r}) = P_i \left(\vec{E}_i^q + \sum_{j=1, j\neq i}^{N} T_{ij'} \vec{\mu}_j \right) \tag{7.14}$$

$$\vec{E}_i^q = \frac{1}{4\pi\varepsilon_0} \sum_{j\neq i}^{N} q_j \frac{\partial J_{ij}^{qq}}{\partial r} \frac{\vec{r}_{ij}}{\left|\vec{r}_{ij}\right|} \tag{7.15}$$

where P_i is the polarizability tensor and T_{ij} is the induced dipole field tensor. Because the atomic polarizability is assumed to be isotropic, P_i thus reduces to a scalar value that is only associated with the elemental nature of the atom. The induced dipole field tensor (T_{ij}) is employed as a damped function that diminishes as atoms overlap.

$$T_{ij} = \frac{1}{4\pi\varepsilon_0 \left|\vec{r}_{ij}\right|^3} \left(1 - 3\frac{\vec{r}_{ij} \cdot \vec{r}_{ij}}{\left|\vec{r}_{ij}\right|^2} \right) \left[1 - e^{-2\xi_i r} (1 + 2\xi_i r_{ij} + 2\xi_j^2 r_{ij}^2) \right] \tag{7.16}$$

Lastly, as shown in eqn (7.17), the energies contributed from atomic polarizability are the dipole self-energy, the dipole–charge interaction, and the dipole–dipole interaction, respectively.

$$U^{\text{polar}}[\{q\}, \{r\}] = \sum_i \frac{\vec{\mu}_i^{\,2}}{2P_i} + \sum_i \vec{\mu}_i \cdot \vec{E}_i^q + \sum_i \sum_{j\neq i} \vec{\mu}_i T_{ij} \vec{\mu}_j \tag{7.17}$$

7.2.3 Short-range Interactions

The bond-order type short-range interactions, ($U^{short}[\{q\},\{r\}]$), where the short-range repulsion energy, $V^R(r_{ij},q_i,q_j)$, and attraction energy, $V^A(r_{ij},q_i,q_j)$, are

based on the Tersoff[11] and Yasukawa[27] potentials, as briefly described in eqn (7.18). In the original Tersoff potential, the short-range interaction is dependent on the interatomic distance (r). However, the short-range interactions in COMB potentials not only decay exponentially with interatomic distance (r) but vary with charge (q). In eqns (7.19) and (7.20), the charge-dependent correction functions, $D_i(q_i)$, are added to the exponential term of the repulsion and attraction energies to reflect the change in atomic radius with charge. The change in the bond order with charge is reflected in the charge dependent function, $B_{ij}^{*}(q_i, q_j)$. In addition, the cutoff function, as shown in eqn (7.21), adopts the cosine-decay formalism to terminate the short-range interactions smoothly, where $R_{ij}^{\max}$ and $R_{ij}^{\min}$ are the upper and lower cutoff distances.

$$U^{\text{short}}[\{q\},\{r\}] = \frac{1}{2}\sum_i \sum_{j\neq i} \left\{ f_c(r_{ij}) \left[V^R(r_{ij}, q_i, q_j) - b_{ij} V^A(r_{ij}, q_i, q_j) \right] \right\} \quad (7.18)$$

$$V^R(r_{ij}, q_i, q_j) = A_{ij} \times \exp\left\{ -\lambda_{ij} r_{ij} + \frac{1}{2}\left[\lambda_{ii} D_i(q_i) + \lambda_{jj} D_j(q_j) \right] \right\} \quad (7.19)$$

$$V^A(r_{ij}, q_i, q_j) = B_{ij} B_{ij}^{*}(q_i, q_j) \times \exp\left\{ -\alpha_{ij} r_{ij} + \frac{1}{2}\left[\alpha_{ii} D_i(q_i) + \alpha_{jj} D_j(q_j) \right] \right\} \quad (7.20)$$

$$f_c(r_{ij}) = \begin{cases} 1 & r_{ij} \le R_{ij}^{\min} \\ \frac{1}{2}\left(1 + \cos\left(\frac{r_{ij} - R_{ij}^{min}}{R_{ij}^{max} - R_{ij}^{min}} \pi \right) \right) & R_{ij}^{\min} < r_{ij} \le R_{ij}^{\max} \\ 0 & r_{ij} > R_{ij}^{\max} \end{cases} \quad (7.21)$$

The bond-order term (b_{ij}) uses a similar formalism to that used in the REBO potential,[14,15] and includes contributions from the bond angle (b^{angle}), coordination (b^{coord}), and additional torsion (b^{torsion}), and conjugation ($b^{\text{conjugation}}$) effects from hydrocarbon systems, which are used to capture many-body effects. Here, owing to the complex bond environment of hydrocarbon systems, the bond-order term only shows the contribution from bond angle and coordination without torsion and conjugation effects, as shown in eqn (7.22). The detailed parameters and fitting procedure of the COMB potential for hydrocarbon systems are described in Ref. 34.

$$b_{ij} = \left[1 + \left(\sum_{k\neq i,j} \zeta(r_{ij}, r_{ik}) g_{ij}(\cos(\theta_{ijk})) + P_{ij} \right)^{\eta_i} \right]^{-1/2\eta_i} \quad (7.22)$$

The asymmetric term $\zeta(r_{ij}, r_{ik})$ to weaken the longer bond, the angular function $g_{ij}(\theta_{ijk})$ and the coordination function P_{ij} are described from eqns (7.23) to (7.25).

$$\zeta(r_{ij}, r_{ik}) = f_c(r_{ik}) \exp\left[\beta_{ij}^{m_i} (r_{ij} - r_{ik})^{m_i} \right] \quad (7.23)$$

$$g_{ij}(\theta_{ijk}) = \sum_{n=0}^{6} b_{ij}^{n} \cos^n(\theta_{ijk}) \quad (7.24)$$

$$P_{ij} = c_0 \Omega_i + c_1 e^{c_2 \Omega_i} + c_3 \quad (7.25)$$

7.2.4 van der Waals Interactions

The long-range vdW interactions ($U^{vdW}[\{r\}]$) in COMB potentials are captured by the classic LJ formula,[2,3] and take the form:

$$U^{\mathrm{vdW}}[\{r\}] = \sum_{i}^{NN}\sum_{j\neq i}^{NN} V^{\mathrm{vdW}}(r_{ij}) = \sum_{i}^{NN}\sum_{j\neq i}^{NN} 4\varepsilon_{ij}^{\mathrm{vdW}}\left[\left(\frac{\sigma_{ij}^{\mathrm{vdW}}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}^{\mathrm{vdW}}}{r_{ij}}\right)^{6}\right] \tag{7.26}$$

In this expression, ${\varepsilon_{ij}}^{\mathrm{vdW}}$ and ${\sigma_{ij}}^{\mathrm{vdW}}$ reflect the strength and equilibrium distance of the vdW interactions, respectively. The vdW interactions are truncated and shifted to zero at the Coulombic cutoff radii. To avoid extremely high repulsion between short-range bonded atoms, a cubic spline function is added to smoothly terminate the vdW interactions at the upper end of the short-range cut-off radii. For any binary vdW bonds, the ${\varepsilon_{ij}}^{\mathrm{vdW}}$ and ${\sigma_{ij}}^{\mathrm{vdW}}$ are decided by considering the geometric and arithmetic means of the values of element-type vdW bonds, respectively.

7.2.5 Correction Terms

The energy terms described in this section are primarily used to modify the energy contribution from specified bond angles, and are therefore designated as correction terms ($U^{\mathrm{corr}}[\{r\}]$). The set of correction terms consist of Legendre polynomials (*LPs*) up to sixth order and a bond bending (*BB*) term:

$$U^{\mathrm{corr}}[\{r\}] = \sum_{i}\sum_{j\neq i}^{N}\sum_{k\neq i}^{N}\left\{\sum_{n=1}^{6} K_{ijk}^{LP_n}\left[\cos(\theta_{ijk})\right] + K_{ijk}^{BB}\left[\cos(\theta_{ijk}) - \cos\left(K_{ijk}^{\theta}\right)\right]^{2}\right\} \tag{7.27}$$

The LPs, whose detailed forms can be found in Ref. 42, are a set of symmetric energy penalties on the bond angle. In contrast, the bond-bending term provides an asymmetric energy penalty on a specific bond angle K_{ijk}^{θ}. More detailed information on the applications of correction terms in COMB potentials can be found in Refs. 29 and 43. All the parameters in eqn (7.27) are three-dimensional, where the first subscript character (i) represents the element type of the central atom.

7.2.6 Parameterization of the COMB Potential

As indicated in Sections 7.2.1 to 7.2.5, although the COMB potential has a complex formalism for the potential energy, each component is indispensable when modeling multiple types of bonding within a given system. Based on the

origin of the potential energies, the general procedure for parameterizing the COMB potential is performed as follows: (i) creating the database from published experimental data and/or electronic structure calculations; (ii) parameterizing the pure systems (starting with the parameters of short-range interactions and then those of electrostatic energies); (iii) parameterizing the binary systems (starting with the parameters of short-range interactions).

7.2.6.1 *Creating the Database*

The parameterization of the COMB potential is performed by optimizing its parameters against the fitting database. In order to allow the COMB potential to transfer among various systems, the fitting database, including wide ranges of pure and binary phases, needs to be considered in the parameterization process. Normally, the fitting database for a pure system consists of bond lengths (or lattice constants), elastic constants, and cohesive energy, and for binary systems consists of atomic charge and heat of formation. The fitting database should also comprise data from systems with a variety of coordination numbers to provide a good description of the bond-order dependence, which consists of different defected and surface structures. In addition, for studies of different focus, a database, such as one containing the transition state energy for atomic diffusion, can be added into the parameterization process.

The properties of these systems are obtained from published experimental data and/or the electronic structure calculations. For crystalline structures, the electronic structure calculations are performed with plane-wave DFT calculations using the software VASP (Vienna Ab initio Simulation Package) with appropriate pseudopotential and exchange-correlation functionals (US-LLDA or PAW-PBE). For molecular systems, the electronic structure calculations are performed using the Gaussian09 computational chemistry software package. Table 7.1 shows the fitting database and predicted values from the COMB potential for Cu.

7.2.6.2 *Parameterization of Pure Systems*

Because the electrostatic energy terms for an uncharged pure system (such as a bulk structure) is zero, the energy contribution of COMB for such a system is from the short-range interactions [eqn (7.18)] and the formalism is reduced to the Tersoff type of potential. This case is straightforward for the parameterization of the COMB potential, which only fits parameters in the pairwise term and bond-order terms, as shown in Table 7.2. The van der Waals interaction will be considered for hydrocarbon systems.

After fitting the parameters in the short-range interactions, the second step is to determine the charge-associated parameters of an atom. The parameters involved in this step are listed in Table 7.3. The fitting process for this step is achieved through the following three substeps: (i) fitting X_i, J_i, K_i, L_i to the electron affinity, first, second, and third order of ionization energies of an isolate atom; (ii) fitting the atomic polarizability, P_i, from the atomic

Table 7.1 Comparison of the properties of Cu predicted by the COMB potential with values obtained from experiments and DFT calculations.

	Experimental	*DFT*	*COMB*
Lattice properties			
A_0(Å)	3.615	3.640	3.610
E_0(eV atom^{-1})	−3.54	−3.50	−3.58
B (GPa)	138	140	142
C_{11}(GPa)	170	173	179
C_{12}(GPa)	123	123	123
C_{44}(GPa)	76	80	48
Other structures			
E(hcp) (eV atom^{-1})		0.006	0.008
E(bcc) (eV atom^{-1})		0.038	0.015
E(sc) (eV atom^{-1})		0.470	0.572
E(dia) (eV atom^{-1})		1.039	0.999
Surface			
γ_s(111) (mJ m^{-2})	1780	1294	1473
γ_s(100) (mJ m^{-2})	1780	1478	1515
γ_s(110) (mJ m^{-2})	1780	1609	1620
Point defect: E_f (eV)			
Vacancy	1.27–1.28		1.18
Interstitial	2.8–4.2		2.79

Table 7.2 List of fitting parameters for the pure systems.

Parameters	*Descriptions*	*Locations*
A_{ii}, B_{ii}, λ_{ii}, α_{ii}	pairwise term	Eqns (7.19–7.20)
η_i, m_i	on bond order	Eqns (7.22–7.23)
β_{ii}	asymmetric term	Eqn (7.23)
b^{ang_0}– b^{ang_6} for ii bond	bond angle term	Eqn (7.24)
c_0–c_3 for ii bond	coordination term	Eqn (7.25)
ε_{ii}^{vdW}, σ_{ii}^{vdW}	vdW interaction	Eqn (7.26)
$K_{iii}^{LP\_1}$ to $K_{iii}^{LP\_6}$, K_{iii}^{BB} and K_{iii}^{θ}	correction functions	Eqn (7.27)

Table 7.3 List of fitting parameters for the charge-dependent energy terms.

Parameters	*Descriptions*	*Locations*
X_i, J_i, K_i, L_i	ionization energy	Eqn (7.4)
P_{ii}^{χ}, P_{ii}^{J}	field effects	Eqn (7.5)
ξ_i, Z_i	Coulomb interaction	Eqns (7.7–7.8)
P_i	atomic polarizability	Eqn (7.14)
D_{Ui}, D_{Li}, Q_{Ui}, Q_{Li}	charge-dependence on short range	Eqsn (7.19–7.20)

polarization in a bonded dimer system; (ii) fitting the rest of the parameters to the reaction energies associated with inserting or detaching an electron from the dimer system.

7.2.6.3 Parameterization of the Binary Systems

Lastly, the COMB poetical is fit to the binary system with multiple phases. Similar to the parameterization process for a pure system, the potential starts with fitting the pairwise terms to the phase orders and charge of a variety of binary systems, which is followed by fitting the parameters in many-body terms to the properties of the binary system. The parameters involved in the binary system are listed in Table 7.4.

Generally, the COMB potential has 30 one-dimensional (also called element-type) parameters for each element, 32 two-dimensional (bond-type) parameters for each bond, and additional three-dimensional (bond-angle-type) parameters for correcting the energy contribution from specified bond angles. This design of parameters is important to balance parameter transferability and flexibility in a multicomponent system. There is no re-parameterization needed for the previously existing force-field when the potential is expanded to include more components or to add new force-fields. Relatively speaking, the parameterization of the COMB potential is more challenging than that of traditional reactive potentials because of the considerable fitting database and parameters. Currently, the COMB potential is available in the Large-scale Atomic/Molecular Massively Parallel Simulator[44] (LAMMPS) programs. A detailed comparison between the COMB potential and other many-body potentials in LAMMPS is given in Table 7.5.

Table 7.4 List of fitting parameters for the binary system.

Parameters	*Descriptions*	*Locations*
P_{ij}^{χ}, P_{ij}^{J}, P_{ji}^{χ}, P_{ji}^{J}	field effects	Eqn (7.5)
A_{ij}, B_{ij}, λ_{ij}, α_{ij}	pairwise short range	Eqns (7.19–7.20)
β_{ij}, β_{ji}	on symmetric term	Eqn (7.23)
b^{ang_0} to b^{ang_6} for *ij* and *ji* bonds	bond angle term	Eqn (7.24)
c_0-c_3 for *ij* and *ji* bonds	coordination term	Eqn (7.25)
$K_{ijk}^{LP\_1}$ to $K_{ijk}^{LP\_6}$, K_{ijk}^{BB} and K_{ijk}^{θ}	correction functions	Eqn (7.27)

Table 7.5 Comparison of computational costs of many-body potentials in LAMMPS.[45]

Potential	*Year*	*Materials*	*Benchmark*	*Ratio(to LJ)*
EAM	1983	fcc metals	Cu	2.3×
MEAM	1987	metals	Ni	20×
Tersoff	1988	covalent solids	Si	4.6×
REBO	1990	carbon nanotubes (CNTs)	polyethylene	8.7×
BOP	1999	covalent solids	CdTe	33×
AIREBO	2000	Multi-wall CNTs	polyethylene	54×
ReaxFF[a]	2001	universal	PETN crystal	256×
COMB[a]	2007	oxides, interfaces	SiO_2	585×

System size (32 000 atoms).
[a]Variable-charge potential.

7.3 Applications

7.3.1 Ethyl Radical Deposition on the Cu(111) Surface

The deposition of clusters and molecules from the vapor onto a surface is a widely used approach to grow thin films, create heterogeneous interfaces, or chemically modify the substrate surface. Such processes also take place in the catalytic reaction. Here, we used COMB potentials to examine the mechanisms by which ethyl radicals (CH_3CH_2) react following deposition on the Cu(111) surface, which is extremely challenging to probe experimentally on its reaction mechanism because of the instability of the ethyl group on transition metal surfaces. The binding energy, which is generally used as an indicator to describe the interaction between molecules and substrate, is one of the important fitting targets in the development of COMB potential. Table 7.6 lists the binding energies of an ethyl radical on Cu surfaces determined by COMB potentials and compares with the results of DFT calculations.[46] The binding energies predicted by COMB are about −35.5, −39.3 and −32.5 kcal mol^{-1} for Cu(100), Cu(110), and Cu(111), which are in good agreement with those obtained by DFT calculations. This comparison indicates that COMB is able to describe organic–inorganic surface chemistry reasonably well.

Further, to analyze the dynamic interactions more realistically, CH_3CH_2 molecules are given a kinetic energy in the direction normal to the Cu substrate with dimensions of 3.1 × 3.1 × 5.0 nm. Such a simulation is possible but much more computationally expensive by DFT-MD. Figure 7.2 shows snapshots

Table 7.6 Binding energies (kcal mol − 1) of CH_3CH_2 on Cu surfaces.

Surfaces	*DFT*	*COMB*
Cu(100)	−36.2	−35.5
Cu(110)	−39.9	−39.3
Cu(111)	−30.9	−32.5

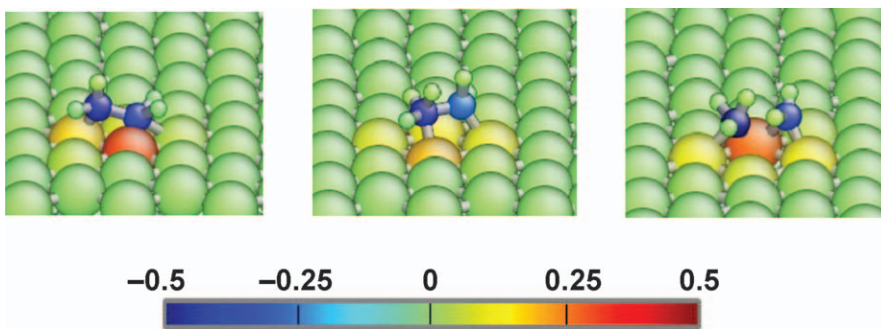


Figure 7.2 Example of a MD simulation with COMB potential for the interaction between CH_3CH_2 and a Cu(111) surface. All the images are colored by charge.

from MD simulations of initially neutral CH_3CH_2 being deposited on Cu(111) and undergoing charge transfer followed by dissociation. As indicated in Figure 7.2(a), the C on the CH_2 initially impacts a surface Cu atom and the instantaneous charges predicted by COMB are $+0.5\ e$ on Cu, $-0.3\ e$ on the C of the CH_3, and $-0.4\ e$ on the C of the CH_2. Because of charge transfer, this leads to a destabilization in the electrostatic energy of C–C interaction, which enables the breaking of the C–C bond in CH_3CH_2, as indicated in Figure 7.2(b–c). From the view of the energy barrier, in the absence of a Cu substrate, the dissociation energy of CH_3CH_2 ($CH_3CH_2 \Rightarrow CH_3 + CH_2$) is predicted to be about 110 kcal mol − 1 by DFT and about 100 kcal mol − 1 by COMB. In contrast, in the presence of a Cu substrate, this energy is decreased to about 90 kcal mol − 1 by DFT and about 85 kcal mol − 1 by COMB. Through the MD simulation with the COMB potential, a dynamic reaction that is computationally expensive for first-principles approaches can be easily simulated. More importantly, the dynamic charge transfer can be presented, which is valuable information for characterizing chemical reactions associated with bond breaking and new bond formation.

7.3.2 Cu/ZnO Heterogeneous System

Heterogeneous interfacial systems are not only ubiquitous in such applications as electronic devices, but are also widely present in catalytic systems. The Cu/ZnO system has been used to catalyze CO_2 hydrogenation to methanol, but questions remain about the nature of the active site and the role of Cu–ZnO interactions in the catalyst performance. Cu metal clusters, Cu^+, the Cu–ZnO interface, and Cu–Zn alloys have all been proposed as the active site but are difficult to model directly with first-principles calculation because of the limitation of size. Possibly, there may be multiple active sites contributing to the overall catalyst performance. Here, the inquiry about different states of copper and zinc oxide may be answered by employing the COMB potential. Such a system is exemplified by Cu clusters on the ZnO surface. Figure 7.3 shows two Cu clusters with different shapes supported by ZnO substrate. The atomic charge of Cu was predicted to vary with its environment within the clusters. For instance, a higher positive charge on the Cu is predicted when the Cu is closer to a surface oxygen atom.

Further, the resulting different states of the Cu/ZnO system are used to determine the mechanisms of chemical reactions between CO_2 and H_2O, as illustrated in Figure 7.4. Although these simulations do not predict the formation of new chemical products, they do predict that the active interactions between CO_2 and H_2O occur near the interface of Cu and ZnO, where there is the highest degree of charge transfer between the surface and the supported metal cluster. Similarly, interfaces have also been proposed as likely reaction sites in the Au/TiO_2 system for CO oxidation.[47] This case study demonstrates that the capability of COMB potentials enables the analysis of catalysis in a multicomponent system at the atomic scale in a more realistic manner than was previously possible without charge transfer potentials.

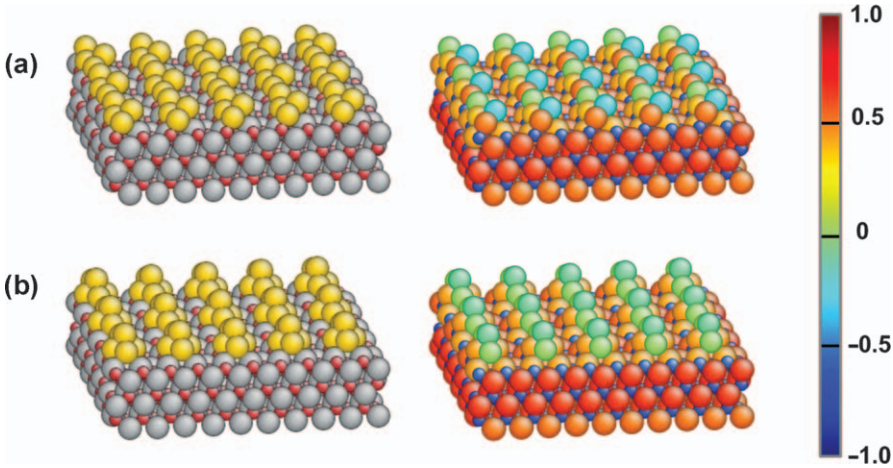


Figure 7.3 Relaxed structures of (a) planar and (b) stack-shaped Cu_5 clusters supported by ZnO surface (yellow atoms are Cu, red atoms are O, and gray atoms are Zn). Below, atoms are color-coded by their charges with the color indicating the charge values.

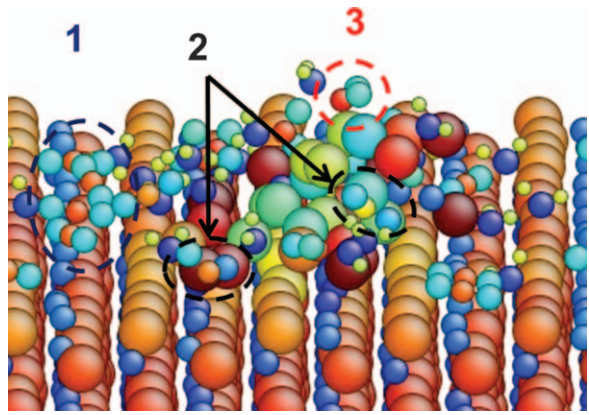


Figure 7.4 Example of a MD simulation with COMB potential for the chemical reaction of CO_2 and H_2O interacting on the Cu clusters supported by ZnO substrate. 1) CO_2 and H_2O adsorb on the ZnO surface, 2) dissociated CO_2 molecules, 3) oxidized CO_2 molecule. All the images are colored by charge.

7.4 Conclusions

The empirical, variable charge COMB potential can be used to model processes such as the dynamic reactions associated with organic–inorganic surface chemistry and interfaces between two dissimilar materials. This information is complementary to the results of DFT calculations and experimental data. Because traditional MD simulations can only simulate events efficiently for

only up to a few nanoseconds, processes such as diffusion and chemical reactions that occur at longer time scales require approaches such as kinetic Monte Carlo (kMC) schemes.[48,49] Considering that the energetics of all likely transitions must be known ahead of time in a typical kMC simulation, the adaptive kMC (akMC) simulation method,[50,51] employing the dimer method[52] for finding saddle points without the knowledge of the final state of the transition, has been adopted in conjugation with the COMB potential.[33] Because the dimer method uses only the first derivatives of the potential and the initial state, it is computationally efficient for the description of complex reaction systems. This approach has qualitatively captured the migration barriers of Cu atoms on the anisotropic ZnO$(10\bar{1}0)$ surface,[33] which are validated against the results of density functional theory calculations and experiments. We expect that this approach can be applied not only to the study of dynamic structural changes but also to reaction behaviors over longer time scales.

Acknowledgments

The work of YTC and SRP was supported by the National Science Foundation (DMR-1005779) while the work of TL and SBS was supported by the National Science Foundation (CHE-0809376) and as part of the Center for Atomic Level Catalyst Design, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001058.

References

1. K. Kadau, T. C. Germann and P. S. Lomdahl, *Int. J. Mod. Phys.*, 2004, **15**, 193.
2. J. E. Jones, *Proc. R. Soc. London, Ser. A*, 1924, **106**, 441.
3. J. E. Jones, *Proc. R. Soc. London, Ser. A*, 1924, **106**, 463.
4. M. S. Daw and M. I. Baskes, *Phys. Rev. Lett.*, 1983, **50**, 1285.
5. M. S. Daw and M. I. Baskes, *Phys. Rev. B*, 1984, **29**, 6443.
6. R. A. Buckingham, *Proc. R. Soc. London, Ser. A*, 1938, **168**, 264.
7. G. V. Lewis and C. R. A. Catlow, *J. Phys. C: Solid State Phys.*, 1985, **18**, 1149.
8. M. W. Finnis and J. E. Sinclair, *Phil. Mag. A*, 1984, **50**, 45.
9. F. Cleri and V. Rosato, *Phys. Rev. B*, 1993, **48**, 22.
10. G. C. Abell, *Phys. Rev. B*, 1985, **31**, 6184.
11. J. Tersoff, *Phys. Rev. Lett.*, 1986, **56**, 632.
12. J. Tersoff, *Phys. Rev. B*, 1988, **37**, 6991.
13. J. Tersoff, *Phys. Rev. B*, 1989, **39**, 5566.
14. D. W. Brenner, *Phys. Rev. B*, 1990, **42**, 9458.
15. D. W. Brenner, O. A. Shenderova, J. A. Harrison, S. J. Stuart, B. Ni and S. B. Sinnott, *J. Phys.: Condens. Matter*, 2002, **14**, 783.

16. B. Ni, K.-H. Lee and S. B. Sinnott, *J. Phys.: Condens. Matter*, 2004, **16**, 7261.
17. A. F. Fonseca, G. Lee, T. L. Borders, H. Zhang, T. W. Kemper, T.-R. Shan, S. B. Sinnott and K. Cho, *Phys. Rev. B*, 2011, **84**, 075460.
18. T. Kemper and S. B. Sinnott, *Plasma Proc. Poly.*, 2012, **9**, 690.
19. I. Jang and S. B. Sinnott, *J. Phys. Chem. B*, 2004, **108**, 18993.
20. T. W. Kemper and S. B. Sinnott, *J. Phys. Chem. C*, 2011, **115**, 23936.
21. T. Liang, S. R. Phillpot and S. B. Sinnott, *Phys. Rev. B*, 2009, **79**, 245110.
22. R. G. Parr, R. A. Donnelly, M. Levy and W. E. Palke, *J. Chem. Phys.*, 1978, **68**, 3801.
23. R. T. Sanderson, *J. Am. Chem. Soc.*, 1983, **105**, 2259.
24. W. J. Mortier, K. van Genechten and J. Gasteiger, *J. Am. Chem. Soc.*, 1985, **107**, 820.
25. A. K. Rappe and W. A. Goddard III, *J. Phys. Chem.*, 1991, **95**, 3358.
26. F. H. Streitz and J. W. Mintmire, *Phys. Rev. B*, 1994, **50**, 11996.
27. A. Yasukawa, *Jsme Int. J. Ser. A*, 1996, **39**, 313.
28. T. Iwasaki and H. Miura, *J. Mater. Res.*, 2001, **16**, 1789.
29. J. G. Yu, S. B. Sinnott and S. R. Phillpot, *Phys. Rev. B*, 2007, **75**, 085311.
30. T.-R. Shan, B. D. Devine, J. W. Hawkins, A. Asthagiri, S. R. Phillpot and S. B. Sinnott, *Phys. Rev. B*, 2010, **82**, 235302.
31. T.-R. Shan, B. D. Devine, T. W. Kemper, S. B. Sinnott and S. R. Phillpot, *Phys. Rev. B*, 2010, **81**, 125328.
32. B. Devine, T. R. Shan, Y. T. Cheng, A. J. H. McGaughey, M. Lee, S. R. Phillpot and S. B. Sinnott, *Phys. Rev. B*, 2011, **84**, 17.
33. Y. T. Cheng, T. R. Shan, T. Liang, B. D. Devine, B. B. Hinojosa, S. R. Phillpot, A. Asthagiri and S. B. Sinnott, *Surf. Sci.*, 2012, **606**, 1280.
34. T. Liang, B. Devine, S. R. Phillpot and S. B. Sinnott, *J. Phys. Chem. A*, 2012, **116**, 7976.
35. T. R. Shan, B. D. Devine, T. Liang, Y. T. Cheng, S. R. Phillpot, S. B. Sinnott, M. Noordhoek, Y. Li and Z. Lu, *Mater. Sci. Eng. Rep.*, 2012, in press.
36. H. Toufar, K. Nulens, G. O. A. Janssens, W. J. Mortier, R. A. Schoonheydt, F. DeProft and P. Geerlings, *J. Phys. Chem.*, 1996, **100**, 15383.
37. J. C. Slater, *Phys. Rev.*, 1930, **36**, 57.
38. D. Wolf, P. Keblinski, S. R. Phillpot and J. Eggebrecht, *J. Chem. Phys.*, 1999, **110**, 8254.
39. P. P. Eward, *Ann. Phys*, 1921, **369**, 253.
40. M. Wilson, M. Exner, Y. M. Huang and M. W. Finnis, *Phys. Rev. B*, 1996, **54**, 15683.
41. H. A. Stern, G. A. Kaminski, J. L. Banks, R. H. Zhou, B. J. Berne and R. A. Friesner, *J. Phys. Chem. B*, 1999, **103**, 4730.
42. B. J. Thijsse, *Nucl. Instrum. Meth. Phys. Res. B*, 2005, **228**, 198.
43. J. Yu, S. B. Sinnott and S. R. Phillpot, *Phil. Mag. Lett.*, 2009, **89**, 136.
44. S. J. Plimpton, *J. Comput. Phys.*, 1995, **117**, 1.
45. S. J. Plimpton and A. P. Thompson, *MRS Bull.*, 2012, **37**, 513.

46. X. Li, A. J. Gellman and D. S. Sholl, *J. Chem. Phys.*, 2007, **127**, 144710.
47. I. X. Green, W. Tang, M. Neurock and J. J. T. Yates, *Science*, 2011, **333**, 736.
48. A. B. Bortz, M. H. Kalos and J. L. Lebowitz, *J. Comput. Phys.*, 1975, **17**, 10.
49. A. F. Voter (ed.), *Introduction to the Kinetic Monte Carlo Method*, Springer, Berlin, 2005.
50. G. Henkelman and H. Jónsson, *J. Chem. Phys.*, 2001, **115**, 9657.
51. L. Xu and G. Henkelman, *J. Chem. Phys.*, 2008, **129**, 114104.
52. G. Henkelman and H. Jónsson, *J. Chem. Phys.*, 1999, **111**, 7010.

Subject Index