# Automatic Prediction of Surface Phase Diagrams Using Ab Initio Grand Canonical Monte Carlo

Robert B. Wexler,[†] Tian Qiu,[†] and Andrew M. Rappe*

Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, United States

**S** *Supporting Information*

**ABSTRACT:** The properties of a material are often strongly influenced by its surfaces. Depending on the nature of the chemical bonding in a material, its surface can undergo a variety of stabilizing reconstructions that dramatically alter the chemical reactivity, light absorption, and electronic band offsets. For decades, ab initio thermodynamics has been the method of choice for computationally determining the surface phase diagram of a material under different conditions. The surfaces considered for these studies, however, are often human-selected and too few in number, leading both to insufficient exploration of all possible surfaces and to biases toward portions of the composition−structure phase space that often do not encompass the most stable surfaces. To overcome these limitations and automate the discovery of realistic surfaces, we combine density functional theory and grand canonical Monte Carlo (GCMC) into "ab initio GCMC." This paper presents the successful application of ab initio GCMC to the study of oxide overlayers on Ag(111), which, for many years, mystified experts in surface science and catalysis. Specifically, we report that ab initio GCMC is able to reproduce the surface phase diagram of Ag(111) with no preconceived notions about the system. Using nonlinear, random forest regression, we discover that the Ag coordination number with O and the surface O−Ag−O bond angles are good descriptors of the surface energy. Additionally, using the composition−structure evolution histories produced by ab initio GCMC, we deduce a mechanism for the formation of oxide overlayers based on the $Ag_3O_4$ pyramid motif that is common to many reconstructions of Ag(111). In conclusion, ab initio GCMC is a promising tool for the discovery of realistic surfaces that can then be used to study phenomena on complex surfaces such as heterogeneous catalysis and material growth, enabling reliable and insightful interpretations of experiments.
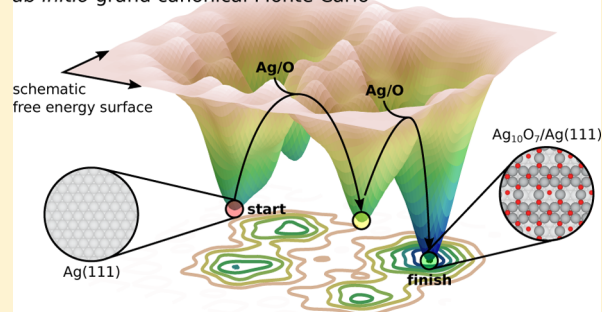
## 1. INTRODUCTION

Theoretical modeling of surface chemical and physical properties often involves making assumptions about the surface structure. However, the physical and chemical properties depend sensitively on these assumptions. The simplest starting point for constructing a surface model is to select a particular facet and then to identify bulk-like terminations from the layering pattern normal to that surface. This approach, however, does not take into account the fact that many bulk-terminated surfaces undergo reconstruction to chemically passivate surface-bound charges and/or saturate surface atom coordination.[1−9] Therefore, the ideal approach involves an exhaustive exploration of all possible surfaces and their reconstructions.

Such an undertaking has two main drawbacks: its computational cost can be prohibitive and the phase space of surface structures is vast and sometimes surprising. Recently, progress has been made toward overcoming these drawbacks by using machine learning to more efficiently traverse surface phase space. For example, genetic algorithms have been developed that programmatically mate different surfaces to explore lower-symmetry phases.[10−12] Additionally, Gaussian process regres-

sion has been employed to learn intermediate surfaces, that is, those that are a mixture of phases from the training set, thereby reducing the number of first-principles calculations necessary.[13] Despite the power of these methods, their main goal is to minimize the surface energy, and they accomplish this using effective but potentially unphysical structural transformations, thus rendering them unable to provide mechanistic information about the natural evolution of the surface.

A simpler and more physically motivated way to explore surface phase space is grand canonical Monte Carlo (GCMC). In GCMC simulations, a system is in contact with both thermal and chemical potential reservoirs, thus allowing fluctuations in the temperature and number of particles. Historically, this technique has been used to study adsorption isotherms: molecules on metals,[14] metal−organic frameworks,[15−17] carbon-based materials,[18,19] zeolites,[20,21] ionic liquids,[22] and activated carbon.[23] GCMC has also been applied to study the bulk phase diagrams of liquids,[24] their mixtures,[25]

alloys,[26−28] fluids,[29] and solvation phenomena.[30,31] In principle, GCMC can be used to generate a collection of surface structures consistent with a predefined temperature and set of chemical potentials of the constituent elements. An application of GCMC to the prediction of surface reconstruction, despite its simplicity and elegance, has never been attempted.

To evaluate the efficacy of GCMC in predicting surface phase diagrams, tests on a well-understood yet complicated material must be performed. One such material that fits these criteria is Ag, which plays an important role in plasmonics,[32,33] catalysis,[34,35] and medicine.[36,37] Since the 1970s, many versions of the Ag(111) surface have been proposed, supported, rejected, and accepted.[4−6,38−43] Early on, low-energy electron diffraction and X-ray photoelectron spectroscopy (XPS) measurements suggested that a $Ag_2O(111)$ overlayer with $p(4 \times 4)$ surface periodicity grows on Ag(111) because of its nearly matching lattice constants.[38,39] With the advent of scanning tunneling microscopy and the re-emergence of ab initio thermodynamics, a host of new structures were proposed, including Ag-deficient and O-enriched variants of the $Ag_2O$ overlayer,[4,40,41] a $Ag_{1.2}O$ cloverleaf-like overlayer,[4] and, most recently, an overlayer consisting of $Ag_9$ islands each connected by two O atoms.[5,42,43] Additionally, surface structures with many other periodicities have been observed experimentally, such as a $c(4 \times 8)$ overlayer, which possesses stripes of base-connected $Ag_3O_4$ triangular pyramids; to date, this $c(4 \times 8)$ pattern offers the lowest surface free energy (for $\Delta\mu_{Ag} = 0$ eV and $-0.64$ eV $\lesssim \Delta\mu_O \lesssim -0.19$ eV) of any Ag(111) reconstruction, as calculated from the density functional theory (DFT).[6]

Here, we report the design of an algorithm and the development of a computer program that implements GCMC in the DFT software package Quantum ESPRESSO.[44] Our implementation of GCMC is open-source, portable, and requires a small number of user inputs.[45] We show that ab initio GCMC, with a small set of simple configurational biases, can independently (re)discover the key features of the oxidized Ag(111) surface phase diagram, which puzzled surface scientists for 5 decades. We also show that by analyzing the ab initio GCMC results with a machine learning model, we can understand and explain the relationships between different structural features and the surface energy. We propose ab initio GCMC as a flexible, general-purpose tool that not only facilitates the discovery of surfaces that are likely to be obtained under different conditions but also generates a rich data set that, upon interrogation, reveals the driving forces behind the formation of different surface structures.

## 2. METHODS

**2.1. Theory.** We work in the grand canonical ensemble, where the chemical potential $\mu$, volume $V$, and temperature $T$ of the system are fixed. The partition function of the grand canonical ensemble is

$$Q(\mu, V, T) = \sum_{N=0}^{\infty} \frac{e^{\mu N/k_B T} V^N}{\Lambda^{3N} N!} \int d\vec{s}^N e^{-U(\vec{s}^N)/k_B T} \tag{1}$$

where $k_B$ is the Boltzmann constant, $N$ is the number of atoms, $\Lambda$ is the thermal de Broglie wavelength, given by $\Lambda = \frac{h}{\sqrt{2\pi m k_B T}}$, $h$ is the Planck constant, $m$ is the mass of the atom, $U$ is the potential energy, and $\vec{s}^N$ are the fractional coordinates of the

atoms. The probability density corresponding to a particular configuration $(\vec{s}^N; N)$ is

$$q_{\mu VT}(\vec{s}^N; \ N) \propto \frac{e^{\mu N/k_B T} V^N}{\Lambda^{3N} N!} e^{-U(\vec{s}^N)/k_B T} \tag{2}$$

There are three different types of actions in unbiased GCMC simulations: move the existing particles in the system, add particles to the system, and remove particles from the system. To ensure that the simulation satisfies detailed balance, the acceptance probability for each action must satisfy

$$q_{\mu VT}(1)\alpha(1, 2)P(1, 2) = q_{\mu VT}(2)\alpha(2, 1)P(2, 1) \tag{3}$$

where 1 and 2 represent configurations $(\vec{s}_1^{N_1}; N_1)$ and $(\vec{s}_2^{N_2}; N_2)$, respectively. $\alpha(1, 2)$ is the probability of attempting a move from configuration 1 to 2 and $P(1, 2)$ is the probability of accepting that move. Because $\alpha(1, 2) = \alpha(2, 1)$, the probability of accepting an attempted "move" step[46] is

$$P_{move} = \min\{1, \ e^{-\Delta U/k_B T}\} \tag{4}$$

where $\Delta U$ is the change in potential energy. For an "exchange" step, if the probabilities of attempting an "add" or "remove" action are equal, that is

$$\alpha(N \text{ particles}, \ N + 1 \text{ particles})$$
$$= \alpha(N + 1 \text{ particles}, \ N \text{ particles}) \tag{5}$$

the acceptance rules are

$$P_{add} = \min\left\{1, \ \frac{V}{(N + 1)\Lambda^3} e^{-(\Delta U - \mu)/k_B T}\right\} \tag{6}$$

and

$$P_{remove} = \min\left\{1, \ \frac{N\Lambda^3}{V} e^{-(\Delta U + \mu)/k_B T}\right\} \tag{7}$$

To focus on the growth of the surface in contact with thermal and chemical potential reservoirs, we replace the "move" action with a structural relaxation after "add" and "remove" actions. The bias introduced by structural relaxation can be countered by replacing the volume in the acceptance probability (see eqs 6 and 7) with an effective volume $V_{eff}$ as discussed in previous works.[47,48] For the "add" action, we choose an element, each with an equal probability, and add it to the system. Instead of randomly selecting the position of the new atom, we include a configurational bias, which prevents the new atom from being too close ($r_{min} = 1.5$ Å) or too far ($r_{max} = 3$ Å) from the closest existing atom. If these criteria are not met, then, we skip this step. This bias has little effect on the detailed balance because all of the configurations we rule out have very high energies and, practically speaking, could never be accepted. For the "remove" action, we randomly choose an atom and remove it from the system. To further restrict the sampling to those phases relevant for surface growth, we added a constraint that atoms can only be inserted at or removed from positions near the top surface (see Figure 1). A flowchart for our ab initio GCMC scheme can be found in the Supporting Information (see Figure S1).

In this work, we study the Ag(111) surface and its reconstructions (see Figure 1). To sample a variety of surface structures and compositions, we set the temperature of the simulations to 500 K and test a range of chemical potentials around the equilibrium between bulk Ag(s) and $Ag_2O(s)$, for
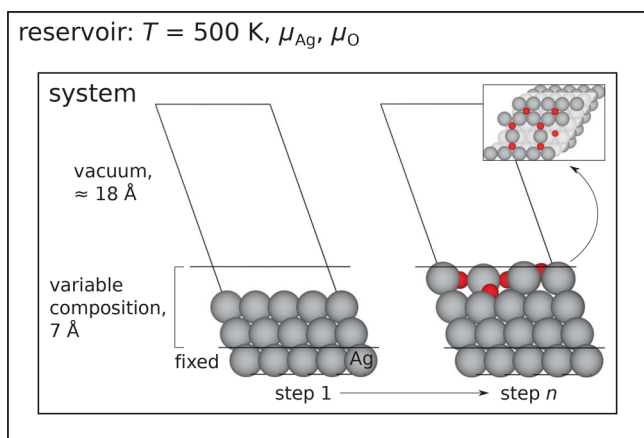
**Figure 1.** $p(4 \times 4)$ Ag(111) slab model for GCMC simulations. We set the temperature and the chemical potentials of Ag and O. The surface is three layers thick, with the bottom layer fixed and ≈18 Å of vacuum. Atoms are only added to or removed from the variable composition region, which extends from 3.5 Å below to 3.5 Å above the top layer of Ag.

which $\mu_{Ag} = \mu_{Ag}^{eq} = G_{Ag}$ and $\mu_O = \mu_O^{eq} = G_{Ag_2O} - 2G_{Ag}$. The free energies of bulk Ag(s) and Ag$_2$O(s) can be approximately written as

$$G_{Ag} \approx U_{Ag}^{DFT} + \Delta G_{Ag}(T)$$

$$G_{Ag_2O} \approx U_{Ag_2O}^{DFT} + \Delta G_{Ag_2O}(T) \tag{8}$$

where the temperature-dependent term is taken from experimental data.[49] We tested five different $\mu_O$ conditions such that $p_{O_2}/p_{O_2}^{eq} = \{10^{-10}, 10^{-6}, 10^{-2}, 1, 10^2\}$. Because the Ag/Ag$_2$O bulk phase boundary corresponds to relatively O-rich conditions, we choose three $p_{O_2}$ lower and only one $p_{O_2}$ higher than $p_{O_2}^{eq}$. The change in the volume from $V$ to $V_{eff}$ in the acceptance probability can be interpreted as a change in the chemical potential, that is, $\delta\mu = k_B T \ln V/V_{eff}$. $V/V_{eff} \approx 10$ because the MC-inserted atoms can access only 10% of $V$, that which is not occupied by the existing atoms. $\delta\mu$ is approximately equal to a 1 order of magnitude change in the partial pressure of O$_2$. Therefore, we can directly use $V$, instead of $V_{eff}$, because our simulation is performed over a range of chemical potentials and it will not influence the result.

As is the convention in the literature,[50] we calculate the surface energy relative to that of Ag(111)

$$\gamma_{slab}^* = \gamma_{slab} - \gamma_{Ag(111)} \tag{9}$$

where $\gamma_{slab}$ is defined as

$$\gamma_{slab} = \frac{1}{A}(U_{slab}^{DFT} - n_{Ag}\mu_{Ag} - n_O\mu_O) \tag{10}$$

Here, $A$ is the surface area and $n$ is the number of atoms. A factor of 2 is missing from the denominator because the bottom layer of each slab is the same, that is, Ag(111), and its contribution to $\gamma_{slab}^*$ cancels out.

**2.2. Computational Details.** DFT[51,52] calculations were performed using Quantum ESPRESSO (version 6.2.1).[44] The generalized gradient approximation of Perdew, Burke, and Ernzerhof was used to treat electron exchange and correlation.[53] Designed nonlocal,[54] optimized,[55] norm-con-

serving pseudopotentials[56] were generated for Ag and O using OPIUM.[57] We used 5s, 5p, and 4d as the valence states for Ag and 2s and 2p for O. We generated a slab model of the $p(4 \times 4)$ Ag(111) surface with three Ag layers and ≈18 Å of vacuum space (see Figure 1). For structural optimizations of the slab model, we fixed the bottom layer and used total energy and force convergence thresholds of 0.01 eV/slab and 0.1 eV/Å, respectively. We sampled the Brillouin zone using a $3 \times 3 \times 1$, Γ-centered $k$-point grid. We also applied a dipole correction along (001) to cancel the artificial electric field across the slab.[58]

Random forests (RFs) were trained using the scikit-learn package (version 0.19.1) for Python (version 3.6.5).[59] Processed data and Python scripts for the machine learning can be found in the Supporting Information. We removed highly correlated and near zero-variance descriptors from our data set. We randomly split the data set into a training and testing set with 2/3 and 1/3 of the data, respectively, so that we could estimate the out-of-sample error in the surface energy prediction.

## 3. RESULTS

**3.1. Surface Phase Diagram of Ag(111).** We perform a series of GCMC simulations, starting from the clean Ag(111) surface, under the conditions described above. Each chemical potential is simulated three times to improve the sampling of surface (composition and structure) phase space. Figure 2
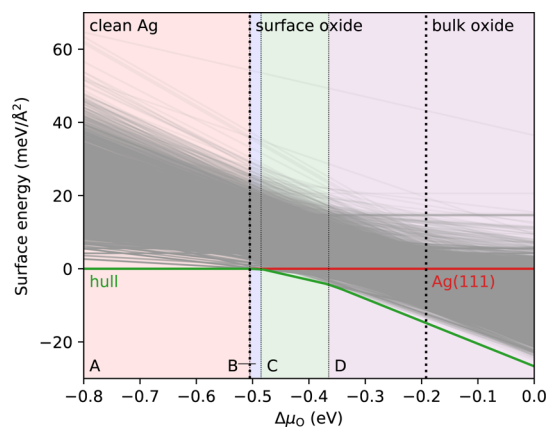


**Figure 2.** Surface phase diagram of Ag(111) exposed to O$_2$, generated by GCMC. There is a gray line for each surface sampled. The red and green lines correspond to Ag(111), that is, the starting point and the surface energy convex hull, respectively. The thick dotted lines separate the three main regions of the phase diagram, and the thin dotted lines separate lightly shaded regions for the four surface phases (A−D, see Figure 3) that constitute the hull.

shows the surface phase diagram generated by GCMC. There are three main regions of this phase diagram with respect to $\Delta\mu_O$ (see thick dotted lines). For $\Delta\mu_O \leq -0.51$ eV, the clean Ag surface is stable (see the red line and Figure 3A). From $-0.51$ eV $\leq \Delta\mu_O \leq -0.19$ eV, surface oxides form. At $\Delta\mu_O = -0.19$ eV, Ag undergoes a bulk phase transition to Ag$_2$O. Over 6000 structures were sampled by the GCMC simulations, and the lines showing their surface free energy versus $\Delta\mu_O$ are shown in gray. Practically speaking, each of the gray lines corresponds to an explicit DFT calculation of $U_{slab}^{DFT}$ in eq 10. We obtained a broad distribution of surface free energies, with values well below and above that of Ag(111) (see red line).
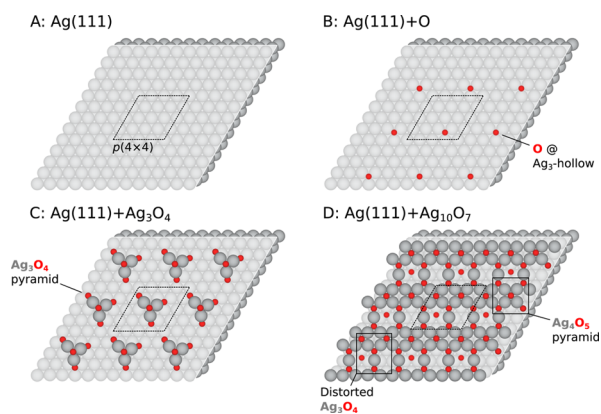
**Figure 3.** Stable Ag(111) surfaces and reconstructions discovered by GCMC: (A) clean Ag, (B) O at a Ag$_3$-hollow site, (C) formation of a Ag$_3$O$_4$ pyramid, and (D) growth of a Ag$_{10}$O$_7$ overlayer. All surfaces have $p(4 \times 4)$ periodicity.

Four different structures make up the surface energy convex hull (see green line). For $\Delta\mu_O \leq -0.51$ eV, Ag(111) is preferred. Between $-0.51$ eV $\leq \Delta\mu_O \leq -0.49$ eV, one O per surface unit cell adsorbs onto a Ag$_3$-hollow site (see Figure 3B). At $\Delta\mu_O$ above $-0.49$ eV and below $-0.37$ eV, surfaces oxides grow in the form of Ag$_3$O$_4$ pyramids (see Figure 3C). O atoms at the corners of these pyramids bind to the surface at Ag$_3$-hollow sites. Under O-rich conditions, that is, $\Delta\mu_O \geq -0.37$ eV, a continuous surface oxide layer forms with the composition Ag$_{10}$O$_7$ (see Figure 3D). This surface consists of edge-sharing, distorted Ag$_3$O$_4$ and symmetric Ag$_4$O$_5$ pyramids. There is also an O atom at one of the two exposed, sublayer Ag$_3$-hollow sites.

This phase diagram, generated automatically using GCMC, is in excellent agreement with the experimental and theoretical literature on Ag(111).[4–6,38–43,50,60–62] The Ag$_3$O$_4$ pyramid is common to many of the structures that have been

proposed.[4,6,38–41,50,60] These pyramids can arrange themselves in a variety of geometries, such as Ag$_2$O(111)-like hexagons and shamrocks.[4,38–41,50,60] The Ag$_{10}$O$_7$ surface we find is very similar to a $c(4 \times 8)$ reconstruction that has been synthesized and, to date, has the lowest reported DFT surface energy.[6] The main difference is that this structure contains unconnected chains of edge-sharing Ag$_3$O$_4$ pyramids, whereas in our structure, the chains are connected, which induces 4.52 meV/Å$^2$ increase in the surface energy. In this study, we impose $p(4 \times 4)$ surface periodicity based on historical precedent. However, oxide adlayers with different periodicities have been reported in the literature.[6,43,63] If we had imposed a smaller surface unit cell, Ag$_3$O$_4$ pyramids could still form but may not dimerize, thereby precluding the growth of 2D, continuous surface oxides like the Ag$_{10}$O$_7$ surface. We find that four unit cells perfectly fit pyramid dimers and that multiples of four are necessary to form chains of pyramid dimers. Therefore, while surface periodicity can affect the ground state arrangement of Ag$_3$O$_4$ pyramids, $p(4 \times 4)$ is ideal for computational studies because it is the smallest surface unit cell that can host 2D, continuous overlayers. It is noteworthy that ab initio GCMC, given only a few inputs and without any prior knowledge of the system, is able to reproduce the important features of the Ag(111) surface phase diagram, which took many decades to decipher.

**3.2. Structural Descriptors for the Surface Energy.** The GCMC simulations described in Section 1 generate a large data set composed of structures and energies. This enables the use of machine learning, namely, RF regression, to determine the structural features that govern surface stability. We choose RF regression because we have shown previously that it is a powerful method for the discovery of structural and electronic descriptors for surface chemical properties like catalysis.[64] An RF is an ensemble of decision trees, each trained on a random subset of the data. The decision trees learn by splitting the data based on values of the independent variables
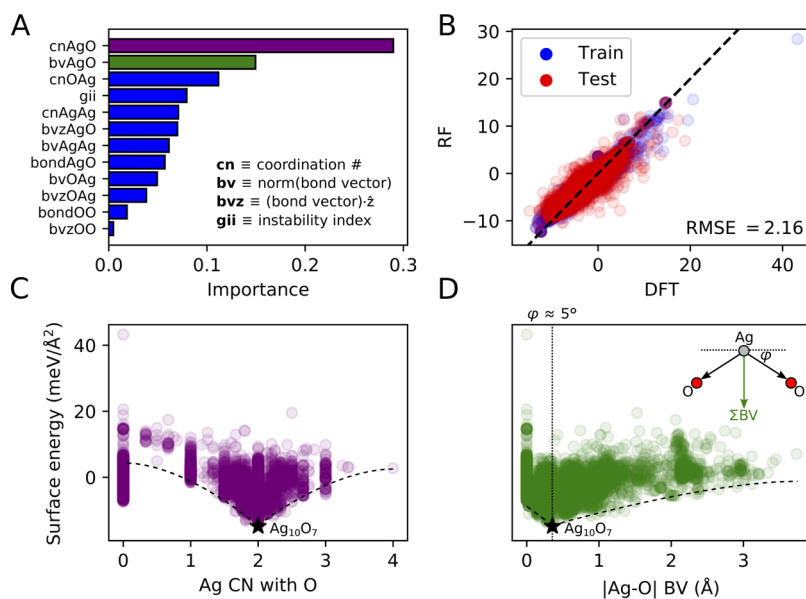


**Figure 4.** Analysis of structural descriptors for the surface energy. (A) Relative importance of descriptors calculated from the RF model. (B) Surface energy predicted by RF vs DFT. The black-dashed line corresponds to perfect agreement. (C) Effect of the Ag coordination number (CN) with O (cnAgO) and (D) effect of the magnitude of the Ag–O bond vector (BV) sum (bvAgO) on the surface energy. cnAgO is the number of O within 3 Å of Ag including bonding to the layer below. The thin, black-dashed lines highlight the trend of increasing free energy for deviations from ideal CN or BV sum. Stars denote the Ag$_{10}$O$_7$ surface.

(e.g., bond length = 2 Å) and then finding which of those splits best separates the data based on the dependent variable (e.g., surface energy). This type of learning is referred to as supervised because we know the value of the output for different sets of inputs. After supervised learning, the RF model can rank the importance of each feature and predict the surface energy (see Figure 4A,B, respectively). Feature importance is a measure of how well splits based on each independent variable separates the data based on the dependent variable. We consider four types of structural features at the surface and calculate their averages: (1) bond length between atoms A and B ("bondAB"), (2) number of atom B within 3 Å of atom A ("cnAB"), (3) magnitude of the sum of the bond vectors (BVs) pointing from atom A to all atom B within 3 Å ("bvAB"), and (4) $z$-component of the BV sum ("bvzAB"). Note that atoms A and B correspond to different elements. In addition to structural features, we calculate the global instability index ("gii"), which measures deviations of each atom from its preferred atomic valence.[65]

Figure 4A,B shows the importance of all of the features and the goodness of fit of the RF model, respectively. The model has a root-mean-squared error (RMSE) of 2.16 meV/Å$^2$, and the data in Figure 4B lie very close to the perfect correlation line. This result shows that we have included features that are excellent descriptors of the surface energy. Scatter plots of surface energy versus the two most important descriptors, that is, cnAgO and bvAgO (see Figure 4C,D, respectively), reveal trends that help rationalize the stability of the Ag$_{10}$O$_7$ surface. Both plots have large spread in the surface energy and concave envelopes tracing the surface energy minima along the descriptor direction (see thin, black-dashed line). The sharpness of these envelopes near the surface energy minimum indicates that surfaces have a clear tendency for cnAgO = 2 and bvAgO ≈ 0.5. The former means that each surface Ag atom tends to form two bonds with O. The preferred value of bvAgO requires a more careful interpretation. bvAgO is zero when Ag either has no O neighbors or the Ag−O BV sums cancel. In the context of twofold coordination of Ag with O, bvAgO is small when the O−Ag−O chain is slightly bent (≈5°, see inset in Figure 4D).

**3.3. Mechanistic Analysis of GCMC Composition–Structure Evolution Histories.** While many oxide overlayers have been proposed, the mechanism of their formation remains unclear.[4−6,38−43,50,60−62] It is known that surface oxide formation requires facile O$_2$ dissociation and significant mass transport of Ag and O.[60,61] A benefit of using GCMC is that it produces a composition–structure evolution history, which can be analyzed to reveal the stages of surface reconstruction. Figure 5 shows the path taken by the GCMC simulation to obtain the Ag$_{10}$O$_7$ surface. There are three main stages of the mechanism: chain growth, pyramid formation, and pyramid dimerization. At the beginning of the first stage, pairs of O atoms adsorb onto nearby Ag$_3$-hollow sites (see Figure 5A). At the same time, they extract a Ag atom from the surface, forming surface O−Ag−O chains and subsurface Ag vacancies. Each single chain serves as a nucleation center from which longer chains can grow, such as double and branched chains, through the addition of extra Ag and O from their respective chemical potential reservoirs (see Figure 5B,C, respectively). The latter is a critical intermediate in the formation of Ag$_3$O$_4$ pyramids in the second stage. Here, the branched chain reorients itself via O hopping between Ag$_3$-hollow and Ag$_2$-bridge sites (see Figure 5D). After the pyramid forms (see
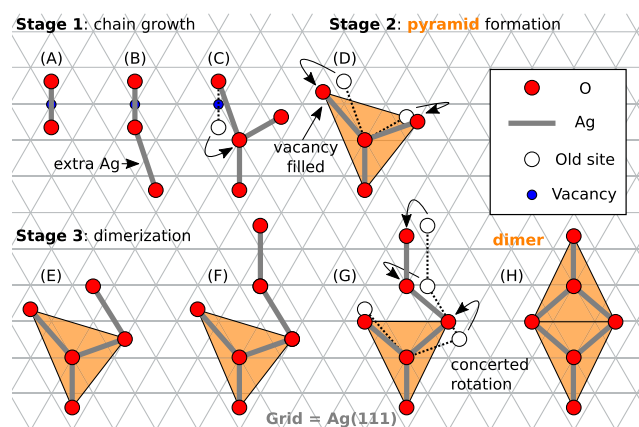


**Figure 5.** Mechanism for the formation of the Ag$_{10}$O$_7$ surface (see Figure 3D). Red, white, and blue circles correspond to O atoms, their previous position, and subsurface Ag vacancies, respectively. Ag atoms are represented by a thick gray line. The mechanism involves three stages: (A−C) chain growth, (D) pyramid formation, and (E−H) dimerization. In the chain growth stage, (A) an O−Ag−O chain and subsurface Ag vacancy form followed by (B) linear and (C) branched chain growth. (D) Next, O atoms jump to new sites and the subsurface Ag vacancy is filled, forming a Ag$_3$O$_4$ pyramid. Finally, in the dimerization stage, (E,F) linear chains grow from the pyramid, which (G) undergoes a concerted rotation. Upon the deposition of a Ag atom, a pyramid dimer is formed.

Figure 5D), the subsurface Ag vacancy is filled. Finally, in the last stage, pyramids dimerize. This starts with chain growth from one of the corners of the pyramid (see Figure 5E). Once a double chain is formed (see Figure 5F), it repositions itself (see Figure 5G) and, upon the deposition of Ag atom, forms a dimer (see Figure 5H), which is the main repeating unit of the Ag$_{10}$O$_7$ surface.

Because many proposed overlayers express the pyramid motif,[4−6,38−43,50,60−62] this is a plausible mechanism for their formation as well, except for the third stage. Recall that the best descriptors of surface energy from the RF model are the Ag−O CN and BV sum. Not only do the most stable surfaces exhibit ideal values for these descriptors but their building blocks, that is, chains and pyramids, do as well. This shows that, within the context of this mechanism, Ag−O CN and BV sum are the key driving forces behind the reconstruction of Ag(111).

# 4. DISCUSSION

Here, we discuss the strengths and weaknesses of ab initio GCMC and provide some recommendations for its future application. Its strengths are that it requires few inputs and has minimal bias toward a particular solution. There are two parameters per element for the configurational bias ($r_{min}$ and $r_{max}$), one parameter defining the dimensions of the variable composition region (see Figure 1) and three parameters for the GCMC simulation ($T$, $\mu_{Ag}$, and $\mu_O$). In ab initio thermodynamics studies of surface reconstructions, it is a common practice to generate a set of reasonable trial structures, sometimes numbering in the hundreds.[4,6,50,60−62] It is difficult, however, to remove bias from this procedure when the structures are human-selected. Such biases are avoided in GCMC because each structure is selected proportionally to its weight in the grand canonical ensemble, which more closely resembles selection in nature. The weaknesses of ab initio GCMC are that it relies on costly ab initio calculations and

only works for surfaces. The first weakness can be overcome by replacing DFT with reactive force fields (e.g., Reaxff,[66] REBO,[67,68] and COMB[69]) or machine-learning atomistic potentials (e.g., aenet[70,71]). There is a trade-off, however, because these methods require careful parameterization and testing and are only available for a small but growing set of systems. A current limitation of our software is that it can only be used for the study of surfaces. We are already in the process of generalizing the code so that it can also be used to study bulk materials and nanoparticles.

Given its success with surfaces, we believe that ab initio GCMC will become an important tool for research in heterogeneous catalysis. Take, for example, the epoxidation of ethylene over Ag, where it is believed that an electrophilic surface O species, seen in XPS measurements, is the key to selective formation of ethylene oxide.[50,60,72−80] For the reconstructions of Ag(111), however, this species is not observed, thus leading to the conclusion that the stable surfaces are not responsible for catalysis.[77] Because GCMC samples both stable and unstable structures, it may find surfaces that do possess electrophilic O species and can therefore catalyze selective epoxidation. In practice, this could involve three steps: (1) reach equilibrium with $\mu_{Ag}$ and $\mu_{O}$, (2) introduce ethylene chemical potential reservoir $\mu_{C_2H_4}$, and (3) reach equilibrium with $\mu_{Ag}$, $\mu_{O}$, and $\mu_{C_2H_4}$. Alternatively, we could apply a bias toward higher free energies that would increase the sampling of surfaces that are less stable but potentially more catalytically active. Other promising applications of ab initio GCMC include the study of binary and ternary materials (e.g., $TiO_2$ and $SrTiO_3$), re-entrant transitions, solvation by including a solvent chemical potential reservoir, and nanoparticle growth for crystal structure prediction.

## 5. CONCLUSIONS

In this paper, we introduce our new method of ab initio GCMC for the investigation of surface reconstruction. This method requires a minimal number of selected parameters, enables surfaces to evolve under realistic conditions, and reduces bias associated with the selection of trial structures for surface stability analyses. We show that ab initio GCMC reproduces the salient features of the Ag(111) surface phase diagram, which took decades to unravel, and, in particular, finds a surface ($Ag_{10}O_7$) that is in excellent agreement with the most stable surface reported in the literature. By analyzing the composition−structure evolution histories of GCMC simulations, we propose a mechanism, based on O−Ag−O chain growth and rearrangement, that can explain the formation of $Ag_3O_4$ pyramid building blocks, which are common to a number of nearly-stable reconstructions of Ag(111). We also show the advantages of using GCMC to generate data for the discovery of structural descriptors of the surface energy via machine learning. We find that the most relevant descriptors (CN of Ag with O and norm of the Ag−O BV sum) support our proposed mechanism and therefore are key driving forces for reconstruction. Ab initio GCMC, from structure generation to analysis, is fully transferable to the study of the surfaces of other materials and also holds promise for the exploration of other processes, such as heterogeneous catalysis and nanoparticle growth.

## ■ ASSOCIATED CONTENT

**ⓢ Supporting Information**

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.8b11093.

> XCrysDen structure files (ZIP)
>
> Values of the structural features for each surface structure (TXT)
>
> Additional computational details (DFT and machine learning) and a flowchart for ab initio GCMC (PDF)

## ■ AUTHOR INFORMATION

**Corresponding Author**
*E-mail: rappe@sas.upenn.edu.
**ORCID** ⓘ
Robert B. Wexler: 0000-0002-6861-6421
Tian Qiu: 0000-0001-8510-894X
Andrew M. Rappe: 0000-0003-4620-6496
**Author Contributions**
†R.B.W. and T.Q. contributed equally to this work.
**Notes**
The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

## ■ REFERENCES

(1) Binnig, G.; Rohrer, H.; Gerber, C.; Weibel, E. 7 × 7 Reconstruction on Si(111) Resolved in Real Space. *Phys. Rev. Lett.* **1983**, *50*, 120−123.

(2) Brommer, K. D.; Needels, M.; Larson, B.; Joannopoulos, J. D. Ab initiotheory of the Si(111)-(7 × 7) surface reconstruction: A challenge for massively parallel computation. *Phys. Rev. Lett.* **1992**, *68*, 1355−1358.

(3) Wang, X.-Q. Phases of the Au(100) surface reconstruction. *Phys. Rev. Lett.* **1991**, *67*, 3547−3550.

(4) Michaelides, A.; Reuter, K.; Scheffler, M. When seeing is not believing: Oxygen on Ag(111), a simple adsorption system? *J. Vac. Sci. Technol. A* **2005**, *23*, 1487−1497.

(5) Schmid, M.; Reicho, A.; Stierle, A.; Costina, I.; Klikovits, J.; Kostelnik, P.; Dubay, O.; Kresse, G.; Gustafson, J.; Lundgren, E.; et al. Structure of Ag(111)-p(4 × 4)-O: No silver oxide. *Phys. Rev. Lett.* **2006**, *96*, 146102.

(6) Martin, N. M.; Klacar, S.; Grönbeck, H.; Knudsen, J.; Schnadt, J.; Blomberg, S.; Gustafson, J.; Lundgren, E. High-coverage oxygen-induced surface structures on Ag(111). *J. Phys. Chem. C* **2014**, *118*, 15324−15331.

(7) Kolpak, A. M.; Li, D.; Shao, R.; Rappe, A. M.; Bonnell, D. A. Evolution of the structure and thermodynamic stability of the BaTiO₃(001) surface. *Phys. Rev. Lett.* **2008**, *101*, 036102.

(8) Martirez, J. M. P.; Morales, E. H.; Saidi, W. A.; Bonnell, D. A.; Rappe, A. M. Atomic and electronic structure of the BaTiO₃(001)(5 × 5) R26.6° surface reconstruction. *Phys. Rev. Lett.* **2012**, *109*, 256802.

(9) Morales, E. H.; Martirez, J. M. P.; Saidi, W. A.; Rappe, A. M.; Bonnell, D. A. Coexisting Surface Phases and Coherent One-Dimensional Interfaces on BaTiO3(001). *ACS Nano* **2014**, *8*, 4465−4473.

(10) Zhu, Q.; Li, L.; Oganov, A. R.; Allen, P. B. Evolutionary method for predicting surface reconstructions with variable stoichiometry. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87*, 195317.

(11) Wang, Q.; Oganov, A. R.; Zhu, Q.; Zhou, X.-F. New reconstructions of the (110) surface of rutile $TiO_2$ predicted by an evolutionary method. *Phys. Rev. Lett.* **2014**, *113*, 266101.

(12) Zhou, X.-F.; Oganov, A. R.; Shao, X.; Zhu, Q.; Wang, H.-T. Unexpected reconstruction of the $\alpha$-boron(111) surface. *Phys. Rev. Lett.* **2014**, *113*, 176101.

(13) Ulissi, Z. W.; Singh, A. R.; Tsai, C.; Nørskov, J. K. Automated discovery and construction of surface phase diagrams using machine learning. *J. Phys. Chem. Lett.* **2016**, *7*, 3931−3935.

(14) Senftle, T. P.; van Duin, A. C. T.; Janik, M. J. Determining in situ phases of a nanoparticle catalyst via grand canonical Monte Carlo simulations with the ReaxFF potential. *Catal. Commun.* **2014**, *52*, 72−77.

(15) Walton, K. S.; Snurr, R. Q. Applicability of the BET Method for Determining Surface Areas of Microporous Metal−Organic Frameworks. *J. Am. Chem. Soc.* **2007**, *129*, 8552−8556.

(16) Han, S. S.; Furukawa, H.; Yaghi, O. M.; Goddard, W. A., III Covalent organic frameworks as exceptional hydrogen storage materials. *J. Am. Chem. Soc.* **2008**, *130*, 11580−11581.

(17) Fetisov, E. O.; Shah, M. S.; Long, J. R.; Tsapatsis, M.; Siepmann, J. I. First principles Monte Carlo simulations of unary and binary adsorption: CO2, N2, and H2O in Mg-MOF-74. *Chem. Commun.* **2018**, *54*, 10816−10819.

(18) Dimitrakakis, G. K.; Tylianakis, E.; Froudakis, G. E. Pillared graphene: a new 3-D network nanostructure for enhanced hydrogen storage. *Nano Lett.* **2008**, *8*, 3166−3170.

(19) Heyden, A.; Düren, T.; Keil, F. J. Study of molecular shape and non-ideality effects on mixture adsorption isotherms of small molecules in carbon nanotubes: A grand canonical Monte Carlo simulation study. *Chem. Eng. Sci.* **2002**, *57*, 2439−2448.

(20) Snurr, R. Q.; Bell, A. T.; Theodorou, D. N. Prediction of adsorption of aromatic hydrocarbons in silicalite from grand canonical Monte Carlo simulations with biased insertions. *J. Phys. Chem.* **1993**, *97*, 13742−13752.

(21) Fetisov, E. O.; Shah, M. S.; Knight, C.; Tsapatsis, M.; Siepmann, J. I. Understanding the Reactive Adsorption of H2 S and CO2 in Sodium-Exchanged Zeolites. *ChemPhysChem* **2018**, *19*, 512−518.

(22) Mullen, R. G.; Corcelli, S. A.; Maginn, E. J. Reaction Ensemble Monte Carlo Simulations of CO2 Absorption in the Reactive Ionic Liquid Triethyl(octyl)phosphonium 2-Cyanopyrrolide. *J. Phys. Chem. Lett.* **2018**, *9*, 5213−5218.

(23) Müller, E. A.; Rull, L. F.; Vega, L. F.; Gubbins, K. E. Adsorption of water on activated carbons: a molecular simulation study. *J. Phys. Chem.* **1996**, *100*, 1189−1196.

(24) Potoff, J. J.; Siepmann, J. I. Vapor-liquid equilibria of mixtures containing alkanes, carbon dioxide, and nitrogen. *AlChE J.* **2006**, *47*, 1676−1682.

(25) Chen, B.; Potoff, J. J.; Siepmann, J. I. Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols. *J. Phys. Chem. B* **2001**, *105*, 3093−3104.

(26) van de Walle, A.; Asta, M.; Ceder, G. The alloy theoretic automated toolkit: A user guide. *Calphad* **2002**, *26*, 539−553.

(27) Castin, N.; Messina, L.; Domain, C.; Pasianot, R. C.; Olsson, P. Improved atomistic Monte Carlo models based on ab-initio-trained neural networks: Application to FeCu and FeCr alloys. *Phys. Rev. B* **2017**, *95*, 214117.

(28) Poyurovskii, L. V.; Ruban, A. V.; Abrikosov, I. A.; Vekilov, Y. K.; Johansson, B. Application of the Monte Carlo method to the problem of surface segregation simulation. *J. Exp. Theor. Phys.* **2001**, *73*, 415−419.

(29) Wilding, N. B. Critical-point and coexistence-curve properties of the Lennard-Jones fluid: A finite-size scaling study. *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.* **1995**, *52*, 602−611.

(30) Marini, G. W.; Texler, N. R.; Rode, B. M. Monte Carlo simulations of Zn(II) in water including three-body effects. *J. Phys. Chem.* **1996**, *100*, 6808−6813.

(31) Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. Solvent effects on the barrier to isomerization for a tertiary amide from ab initio and Monte Carlo calculations. *J. Am. Chem. Soc.* **1992**, *114*, 7535−7542.

(32) Pendry, J. B. Negative refraction makes a perfect lens. *Phys. Rev. Lett.* **2000**, *85*, 3966−3969.

(33) Kelly, K. L.; Coronado, E.; Zhao, L. L.; Schatz, G. C. The optical properties of metal nanoparticles: the influence of size, shape, and dielectric environment. *J. Phys. Chem. B* **2003**, *107*, 668−677.

(34) Hughes, M. D.; Xu, Y.-J.; Jenkins, P.; McMorn, P.; Landon, P.; Enache, D. I.; Carley, A. F.; Attard, G. A.; Hutchings, G. J.; King, F.; et al. Tunable gold catalysts for selective hydrocarbon oxidation under mild conditions. *Nature* **2005**, *437*, 1132−1135.

(35) Christopher, P.; Xin, H.; Linic, S. Visible-light-enhanced catalytic oxidation reactions on plasmonic silver nanostructures. *Nat. Chem.* **2011**, *3*, 467−472.

(36) Jain, P. K.; Huang, X.; El-Sayed, I. H.; El-Sayed, M. A. Noble metals on the nanoscale: optical and photothermal properties and some applications in imaging, sensing, biology, and medicine. *Acc. Chem. Res.* **2008**, *41*, 1578−1586.

(37) Chen, X.; Schluesener, H. J. Nanosilver: A nanoproduct in medical application. *Toxicol. Lett.* **2008**, *176*, 1−12.

(38) Rovida, G.; Pratesi, F.; Maglietta, M.; Ferroni, E. Effects of oxygen on silver surface structure. *J. Vac. Sci. Technol.* **1972**, *9*, 796−799.

(39) Campbell, C. T. Atomic and molecular oxygen adsorption on Ag(111). *Surf. Sci.* **1985**, *157*, 43−60.

(40) Carlisle, C. I.; King, D. A.; Bocquet, M.-L.; Cerdá, J.; Sautet, P. Imaging the Surface and the Interface Atoms of an Oxide Film onAg{111}by Scanning Tunneling Microscopy: Experiment and Theory. *Phys. Rev. Lett.* **2000**, *84*, 3899−3902.

(41) Carlisle, C. I.; Fujimoto, T.; Sim, W. S.; King, D. A. Atomic imaging of the transition between oxygen chemisorption and oxide film growth on Ag{111}. *Surf. Sci.* **2000**, *470*, 15−31.

(42) Schnadt, J.; Michaelides, A.; Knudsen, J.; Vang, R. T.; Reuter, K.; Lægsgaard, E.; Scheffler, M.; Besenbacher, F. Revisiting the structure of the p (4 × 4) surface oxide on Ag (111). *Phys. Rev. Lett.* **2006**, *96*, 146101.

(43) Schnadt, J.; Knudsen, J.; Hu, X. L.; Michaelides, A.; Vang, R. T.; Reuter, K.; Li, Z.; Lægsgaard, E.; Scheffler, M.; Besenbacher, F. Experimental and theoretical study of oxygen adsorption structures on Ag (111). *Phys. Rev. B: Condens. Matter Mater. Phys..* **2009**, *80*, 075424.

(44) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Buongiorno Nardelli, M.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; et al. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys. Condens. Matter* **2017**, *29*, 465901.

(45) Wexler, R. B.; Qiu, T.; Rappe, A. M. https://github.com/rwexler/gcmc/ (accessed June 28, 2018).

(46) Frenkel, D.; Smit, B. *Understanding Molecular Simulation: From Algorithms to Applications*; Elsevier, 2001; Vol. *1*.

(47) Senftle, T. P.; Meyer, R. J.; Janik, M. J.; Van Duin, A. C. Development of a ReaxFF potential for Pd/O and application to palladium oxide formation. *J. Chem. Phys.* **2013**, *139*, 044109.

(48) Lachet, V. r.; Boutin, A.; Tavitian, B.; Fuchs, A. H. Grand canonical Monte Carlo simulations of adsorption of mixtures of xylene molecules in faujasite zeolites. *Faraday Discuss.* **1997**, *106*, 307−323.

(49) Wagman, D. D.; Evans, W.; Parker, V.; Halow, I.; Bailey, S. M. *Selected Values of Chemical Thermodynamic Properties. Tables for Elements 35 through 53 in the Standard Order of Arrangement*, 1969.

(50) Michaelides, A.; Bocquet, M.-L.; Sautet, P.; Alavi, A.; King, D. A. Structures and thermodynamic phase transitions for oxygen and silver oxide phases on Ag{1 1 1}. *Chem. Phys. Lett.* **2003**, *367*, 344−350.

(51) Hohenberg, P.; Kohn, W. Inhomogeneous electron gas. *Phys. Rev.* **1964**, *136*, B864−B871.

(52) Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **1965**, *140*, A1133−A1138.

(53) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77*, 3865.

(54) Ramer, N. J.; Rappe, A. M. Designed nonlocal pseudopotentials for enhanced transferability. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 12471−12478.

(55) Rappe, A. M.; Rabe, K. M.; Kaxiras, E.; Joannopoulos, J. D. Optimized pseudopotentials. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1990**, *41*, 1227−1230.

(56) Hamann, D. R.; Schlüter, M.; Chiang, C. Norm-conserving pseudopotentials. *Phys. Rev. Lett.* **1979**, *43*, 1494−1497.

(57) http://opium.sourceforge.net/ (accessed June 20, 2016).

(58) Bengtsson, L. Dipole correction for surface supercell calculations. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1999**, *59*, 12301−12304.

(59) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.

(60) Li, W.-X.; Stampfl, C.; Scheffler, M. Subsurface oxygen and surface oxide formation at Ag (111): A density-functional theory investigation. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2003**, *67*, 045408.

(61) Reuter, K. *Nanocatalysis*; Springer, 2007; pp 343−376.

(62) Jones, T. E.; Rocha, T. C. R.; Knop-Gericke, A.; Stampfl, C.; Schlögl, R.; Piccinin, S. Thermodynamic and spectroscopic properties of oxygen on silver under an oxygen atmosphere. *Phys. Chem. Chem. Phys.* **2015**, *17*, 9288−9312.

(63) Reicho, A.; Stierle, A.; Costina, I.; Dosch, H. Stranski-Krastanov like oxide growth on Ag(111) at atmospheric oxygen pressures. *Surf. Sci.* **2007**, *601*, L19−L23.

(64) Wexler, R. B.; Martirez, J. M. P.; Rappe, A. M. Chemical pressure-driven enhancement of the hydrogen evolving activity of $Ni_2P$ from nonmetal surface doping interpreted via machine learning. *J. Am. Chem. Soc.* **2018**, *140*, 4678−4683.

(65) Brown, I. D. Chemical and steric constraints in inorganic solids. *Acta Crystallogr. B.* **1992**, *48*, 553−572.

(66) van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: a reactive force field for hydrocarbons. *J. Phys. Chem. A* **2001**, *105*, 9396−9409.

(67) Tersoff, J. New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 6991−7000.

(68) Tersoff, J. Empirical interatomic potential for carbon, with applications to amorphous carbon. *Phys. Rev. Lett.* **1988**, *61*, 2879−2882.

(69) Yu, J.; Sinnott, S. B.; Phillpot, S. R. Charge optimized many-body potential for the $Si/SiO_2$ system. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2007**, *75*, 085311.

(70) Artrith, N.; Urban, A. An implementation of artificial neural-network potentials for atomistic materials simulations: Performance for $TiO_2$. *Comput. Mater. Sci.* **2016**, *114*, 135−150.

(71) Artrith, N.; Urban, A.; Ceder, G. Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *96*, 014112.

(72) Bukhtiyarov, V. I.; Hävecker, M.; Kaichev, V. V.; Knop-Gericke, A.; Mayer, R. W.; Schlögl, R. Atomic oxygen species on silver: Photoelectron spectroscopy and x-ray absorption studies. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2003**, *67*, 235422.

(73) Kaichev, V. V.; Bukhtiyarov, V. I.; Hävecker, M.; Knop-Gercke, A.; Mayer, R. W.; Schlögl, R. *Kinet. Catal.* **2003**, *44*, 432−440.

(74) Rocha, T. C. R.; Oestereich, A.; Demidov, D. V.; Hävecker, M.; Zafeiratos, S.; Weinberg, G.; Bukhtiyarov, V. I.; Knop-Gericke, A.; Schlögl, R. The silver-oxygen system in catalysis: new insights by near ambient pressure X-ray photoelectron spectroscopy. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4554−4564.

(75) Jones, T. E.; Wyrwich, R.; Böcklein, S.; Rocha, T. C. R.; Carbonio, E. A.; Knop-Gericke, A.; Schlögl, R.; Günther, S.; Wintterlin, J.; Piccinin, S. Oxidation of ethylene on oxygen reconstructed silver surfaces. *J. Phys. Chem. C* **2016**, *120*, 28630−28638.

(76) Jones, T. E.; Rocha, T. C. R.; Knop-Gericke, A.; Stampfl, C.; Schlögl, R.; Piccinin, S. Insights into the electronic structure of the oxygen species active in alkene epoxidation on silver. *ACS Catal.* **2015**, *5*, 5846−5850.

(77) Carbonio, E. A.; Rocha, T. C. R.; Klyushin, A. Y.; Píš, I.; Magnano, E.; Nappini, S.; Piccinin, S.; Knop-Gericke, A.; Schlögl, R.; Jones, T. E. Are multiple oxygen species selective in ethylene epoxidation on silver? *Chem. Sci.* **2018**, *9*, 990−998.

(78) Bocquet, M.-L.; Sautet, P.; Cerda, J.; Carlisle, C. I.; Webb, M. J.; King, D. A. Specific Ethene Surface Activation on Silver Oxide Covered Ag{111} from the Interplay of STM Experiment and Theory. *J. Am. Chem. Soc.* **2003**, *125*, 3119−3125.

(79) Linic, S.; Barteau, M. A. Formation of a stable surface oxametallacycle that produces ethylene oxide. *J. Am. Chem. Soc.* **2002**, *124*, 310−317.

(80) Linic, S.; Barteau, M. A. Control of ethylene epoxidation selectivity by surface oxametallacycles. *J. Am. Chem. Soc.* **2003**, *125*, 4034−4035.