Staged Training of Machine-Learning Potentials from Small to Large Surface Unit Cells: Efficient Global Structure Determination of the RuO$_2$(100)-c(2 × 2) Reconstruction and (410) Vicinal

Published as part of The Journal of Physical Chemistry C virtual special issue "Machine Learning in Physical Chemistry Volume 2".

Yonghyuk Lee, Jakob Timmermann, Chiara Panosetti, Christoph Scheurer, and Karsten Reuter*

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Machine-learning (ML) potentials trained with density functional theory (DFT) data boost the sampling capabilities in first-principles global surface structure determination. Particular data efficiency is thereby achieved by iterative training protocols that blend the creation of new training data with the actual surface exploration process. Here, we extend this to a staged training from small to large surface unit cells. With many geometric motifs learned from small unit cell data, successively less new DFT structures in computationally demanding large surface unit cells are queried. We demonstrate the fully automated workflow in the context of rutile RuO$_2$ surfaces. For a Gaussian approximation potential (GAP) initially trained on (1 × 1) surface structures, only limited additional data are necessary to efficiently recover only recently identified structures for the RuO$_2$(100)-c(2 × 2) reconstruction. The same holds when retraining this GAP for the (410) vicinal, the optimized structure of which is found to involve c(2 × 2) reconstructed terraces. Due to the high stability of this structure, (410) vicinals appear in the predicted Wulff equilibrium nanoparticle shape.
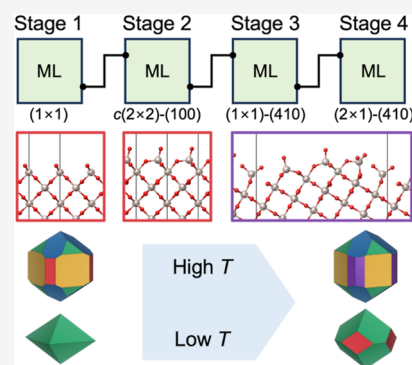
INTRODUCTION

First-principles surface structure determination has traditionally been hampered by the high computational costs of the underlying electronic structure calculations, prevalently performed with density functional theory (DFT) and using periodic boundary condition (PBC) supercells. These costs led to severe limitations in the sampling capabilities of the underlying potential energy surface (PES). Typically, this sampling only involved the local geometry optimization of a batch of trial structures, which are motivated by incomplete experimental characterization or the chemical intuition of the researcher. Whether the true global PES minimum corresponding to the most stable structure is found or not depends, critically, on the selected trial structure batch.

Even worse, detailed geometric insight is most desirable for non-trivial surface structures like reconstructions or (extended) defects including steps or facet edges. Such structures generally need to be described in large surface-area supercells. The exceeding costs of the corresponding DFT calculations then further restrict the sampling (often permitting only a handful of trial structures), while at the increased dimensionality of the PES, ever more extensive sampling is likely necessary to identify the global minimum structure. Till recently, seminal rare attempts to overcome the sampling limitations of first-principles structure determination by state-of-the-art global

geometry optimization techniques have been restricted to rather small unit cell size problems and were possibly at the reliability edge of the employed technique.[1−7] The computational costs dictated, e.g., ultrafast cooling rates in simulated annealing or only limited minima exploration in basin or minima hopping.

Fortunately, this stalemate is now overcome by the advent of machine-learning interatomic potentials (MLIPs).[8−13] The high-cost DFT calculations serve to efficiently train the MLIP as a reliable PES surrogate that ideally maintains the underlying DFT predictive quality.[14] The actual sampling is then performed on the lower-cost surrogate model, typically allowing to increase the number of PES evaluations by several orders of magnitude. Several studies have already heralded the prospects of this MLIP-based first-principles surface structure determination approach for increasingly complex systems.[15−19] Most data-efficient with respect to the required DFT
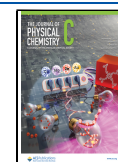
calculations are the emerging active-learning strategies that iteratively extend the DFT training set by promising surface geometries that are identified in the ongoing PES sampling.[20−24] Of particular interest especially for the determination of complex large surface-area structures is to achieve maximum training of the MLIP while sampling a suitable (sub-)space of smaller surface-area structures. When extending the PES search to the larger surface-area cells of true interest, only a minimum number of additional DFT calculations for such computationally demanding structures are then needed to refine the MLIP training and achieve an accurate PES surrogate for the global optimization.

Here, we demonstrate and assess this concept in the context of rutile RuO_2 surfaces. Significant work has been dedicated to this metallic oxide not least because of its activity in thermal and electrochemical oxidation catalysis.[25−28] Despite a plethora of studies, surprisingly many questions nevertheless remain for the detailed surface geometries exposed under the various reaction conditions. Only quite recently, and using MLIP-based global structure determination, did some of us identify a new complexion-type surface structure for the $RuO_2(101)$ facet even in the seemingly trivial space of (1×1) surface unit cells.[16,22,29] As explicitly validated by the experiment, the high stability of this complexion under strongly reducing conditions then leads to a preferred growth along this crystallographic orientation.[16] Similarly, and even more recently, Hess et al.[7] finally resolved a long-time enigmatic $c(2 \times 2)$ reconstruction of the $RuO_2(100)$ facet that is believed to be behind the observed deactivation of the catalyst under O-rich CO oxidation reaction conditions[25,30−32] This study was a tour-de-force using in parts direct DFT-based simulated annealing with a concomitantly rather drastic quenching and in parts manual construction of trial structures. Revisiting this problem within an active-learning MLIP-strategy allows not only to test and validate our small-to-large surface area training protocol, but also to increase the confidence in the direct DFT-based findings.

Indeed, we can fully confirm the findings of Hess et al. within a massively increased and fully automatized sampling, while simultaneously spending a dramatically reduced amount of CPU time. This low computational cost allows us to further extend the iterative learning and readily study the implications of the $c(2 \times 2)$ reconstruction on the structure of the $(110)/(100)$ facet edge—a large surface-area problem that would be completely intractable by direct DFT-based global structure determination. In fact, the high stability of the identified edge structure even changes the predicted equilibrium RuO_2 nanoparticle structure with corresponding (410) vicinals appearing in the Wulff shape, exactly as reported in high-resolution electron microscopy studies.[33]

■ METHODS

Gaussian Approximation Potential. Gaussian approximation potentials (GAPs) are a widely used class of MLIPs based on a sparse variant of Gaussian process regression.[8,11,34] The specific GAP employed in this work is analogous to the one of our previous work,[22] where we described the formalism as well as the iterative-training protocol for (1×1) surface unit cells in detail. We therefore content ourselves here with only a brief recap for self-containment and refer to the previous work for details.

As a PES surrogate, the GAP calculates the total energy E_{GAP} of the system from its atomic coordinates X_n as

$$E_{GAP}(X_n) = \underbrace{\sum_{i,j} \delta_{2B}^2 \sum_{m=1}^{M_{2B}} c_{m,2B} k_{2B}(r_{ij}, r_m)}_{E_{2B}}$$
$$+ \underbrace{\sum_i \delta_{MB}^2 \sum_{m=1}^{M_{MB}} c_{m,MB} k_{MB}(\chi_i, \chi_m)}_{E_{MB}} \tag{1}$$

It thus consists of a two-body (2B) E_{2B} and a many-body (MB) E_{MB} energy contribution, which result from a sum over all pairs of atoms i, j and a sum over all atoms i in the system, respectively. The second sum in both contributions goes over a set of $M_{2B/MB}$ representative data points and includes the regression coefficients $c_{m,2B/MB}$ and the kernel basis functions $k_{2B/MB}$. The $k_{2B/MB}$ measure the similarity between two local geometric descriptors (representations) computed from X_n. In the 2B contribution, these descriptors are simply the interatomic distances r_{ij} up to a specified cutoff radius r_{cut} that enter a squared exponential (Gaussian) kernel. In the MB contribution, the descriptors are translationally, rotationally, and permutationally invariant vectorial representations of the local atomic environment χ_i based on the smooth overlap of atomic positions (SOAP)[35] that enter a second-order polynomial kernel. All technical hyperparameters, including the relative weights of the two contributions, $\delta_{2B/MB}$, the SOAP, and kernel parameters, as well as the force-locality derived radial cutoff r_{cut}, cf. Figure S1, are tabulated in Table S1, with all details of the hyperparameter selection process described in ref 22.

To train the GAP with DFT data, an iterative refinement workflow is used.[22] This workflow starts with an initial minimal set of training structures composed by the researchers. Different canonical global geometry optimization runs spanning a wide range of surface stoichiometries within the targeted surface unit cell are then conducted based on the resulting preliminary GAP. The structures generated this way are compared to each other and to those already in the training set via a kernel distance

$$\kappa(A, B) = \sqrt{2 - 2 \min_{\substack{a \in A \\ b \in B}}(k_{MB}(\chi_a, \chi_b))} \tag{2}$$

Instead of averaged similarities of the atomic environments in two configurations A and B,[36,37] this kernel checks on the presence of a single hitherto unknown atomic environment by evaluating the minimal similarity $k_{MB}(\chi_a, \chi_b)$ between any two atoms $a \in A$ and $b \in B$. $\kappa_{crit} = 0.075$ is employed as a system-specific parameter to deem a new configuration sufficiently different to be added to the training set after local DFT optimization. After the corresponding augmentation of the training set, the next refinement cycle continues by retraining the GAP and subsequent new global geometry optimization runs. Further cycles are repeated until either no new unknown structures are identified or the out-of-sample error of the last structures added to the training set decreases below a threshold value. As only a small number of structures is added to the training set in each refinement cycle, this out-of-sample error assessed in the form of the root mean square error (RMSE) in energies and forces with respect to the DFT references shows strong fluctuations, cf. Figure S5. We therefore use an exponential moving average (EMA)[19]

$$\text{EMA}(i) = \text{EMA}(i-1) + 0.3(\text{RMSE}(i) - \text{EMA}(i-1)) \tag{3}$$

with $\text{EMA}(0) = \text{RMSE}(0)$ to evaluate the convergence of the training procedure in refinement cycle $i \geq 1$. We deem the GAP to have attained convergence once the EMA drops below 8 meV/atom for energy and 0.15 eV/Å for forces. Having intertwined the GAP training and the actual MLIP-based global geometry optimization, the final set of training structures constitutes also already the outcome of the surface structure search, if the training was stopped because no further unknown structures were found. If the training was stopped because a low EMA indicates that the GAP has learnt all local geometry motifs sufficiently well, we also compute all lowest energy structures found in subsequent global geometry optimization runs explicitly with DFT. In either case, all lowest-energy structures discussed in the text correspond to geometry-optimized DFT structures and their energetics.

Simulated Annealing. The MLIP-based global geometry optimization is performed via molecular dynamics (MD)-based simulated annealing (SA). The MD simulations are carried out with the LAMMPS code,[38] using the velocity Verlet algorithm[39] and 1 fs time steps. Exploiting the small volumetric thermal expansion coefficients for RuO_2,[40] we employ a canonical NVT ensemble, controlling the temperature with an efficient Berendsen thermostat.[41] In the actual SA protocol, the structures are heated from 200 to 1000 K and then quenched back to 200 K with a constant heating and cooling rate of 1.6 K/ps. After each SA, the resulting finite-temperature structure is fully optimized through conjugate gradient minimization with the same convergence threshold as used for DFT calculations described next.

In particular for the rugged (410) vicinal supercells, there is a tendency that stable RuO_4 units desorb from the surface during the highest temperature part of the SA. We suppress this with a harmonic repulsive potential that starts at the topmost layer of the surface slab and rises to a value of 10 eV at 10 Å above the surface. This potential is gradually reduced during the cool-down process of the SA, so that it is no longer present at the end of the cool down and beginning of the geometry optimization. By construction, this confinement potential has therefore no impact on the final structures and their energetics, while fully allowing for any mass transfer of surface species in lateral directions during the high-temperature part of the SA.

Density Functional Theory. All DFT calculations are performed using a plane-wave basis set and SG15 optimized norm-conserving Vanderbilt pseudopotentials[42] as implemented in the Quantum Espresso software package.[43] The semi-local revised Perdew−Burke−Ernzerhof (RPBE)[44] functional is used to describe electronic exchange and correlation. The kinetic cutoff energy for the expansion of the wave function is set to 80 Ry, while that of the charge density is set to 320 Ry. Brillouin-zone integrations are carried out on a grid of k-points with reciprocal distances of 0.02 Å^{-1}, producing an $(11 \times 11 \times 16)$ k-point grid for bulk rutile RuO_2. Optimized lattice parameters for bulk RuO_2 are obtained by minimizing the stress tensor and all internal degrees of freedom until the external pressure falls below 0.5 kbar. Geometry optimization for all slab calculations employs Broyden−Fletcher−Goldfarb−Shanno (BFGS) minimization[45−47] until residual changes in total energy and all force components fell below 1.4×10^{-2} meV and 0.3 meV/Å, respectively. The high numerical

convergence achieved with these computational settings has been demonstrated in ref 22.

All surface calculations are performed using PBC supercell geometries, with a minimum vacuum separation of 20 Å. Symmetric slab models involving at least seven trilayers of RuO_2 units were employed in the previous work that established the GAP for the surface structure determination within the sub-space of (1×1) surface unit cells of all five low-index rutile facets.[22] This GAP is the starting point of the present investigation of the larger $c(2 \times 2)$-$RuO_2(100)$ surface unit cell and the $(100)/(110)$ facet edge. For the DFT calculation of corresponding $c(2 \times 2)$-$RuO_2(100)$ structures identified in the iterative-training protocol, we employed symmetric slabs with at least 13 trilayers and a minimum vacuum region of 20 Å. The facet edge is described within either a (1×1)- or (2×1)-(410) surface unit cell with a minimum thickness of 20 Å and the same 20 Å vacuum region. The size of the (100) facet in the (2×1) surface unit cell is then large enough to establish $c(2 \times 2)$ structures. PES barriers to achieve the $c(2 \times 2)$ reconstruction are calculated via the machine learning accelerated nudged-elastic band (NEB) method AIDNEB.[48−50] These barriers are computed with a maximum uncertainty of the AIDNEB-internal surrogate model of 0.02 eV until a convergence of the forces on the climbing image of less than 0.03 eV/Å.

Ab Initio Thermodynamics. To assess the relative stability of surface structures with differing chemical composition, we adopt the ab initio thermodynamics approach.[51,52] Specifically, we consider the surfaces to be in thermodynamic equilibrium with an oxygen-containing gas phase and calculate the surface free energy $\gamma_{\text{surf}}^{(hkl),\nu_s}$ of a structure with specific Miller index (hkl) and chemical composition ν_s as

$$\gamma_{\text{surf}}^{(hkl),\nu_s} = \frac{1}{2A^{(hkl)}} \left[G_{\text{surf}}^{(hkl),\nu_s} - \sum_s \nu_s^{(hkl),\nu_s} \mu_s \right] \tag{4}$$

Here, $G_{\text{surf}}^{(hkl),\nu_s}$ is the Gibbs free energy of the surface system described by a symmetric slab in a supercell with surface unit cell area $A^{(hkl)}$. The chemical composition is described by ν_s, the number of atoms of various species s ($=$Ru, O), and μ_s is the chemical potential of the species s present in the system.

Assuming the surface to be in equilibrium with the underlying bulk rutile RuO_2 constrains the chemical potentials of Ru and O to the Gibbs free energy (per formula unit) of RuO_2 bulk (i.e., $G_{RuO_2,\text{bulk}} = \mu_{\text{Ru}} + 2\mu_{\text{O}}$). In turn, the chemical potential of O is set by the equilibrium with the surrounding O gas-phase reservoir to $\mu_{\text{O}} = \frac{1}{2}E_{O_2} + \Delta\mu_{\text{O}}$. Here, E_{O} is the total energy of an isolated O_2 molecule, including zero point energy (ZPE) contributions.[53,54] The relative O chemical potential $\Delta\mu_{\text{O}}$ depends on the finite temperature T and pressure p by assuming the gas phase molecule as an ideal-gas like reservoir.[51,52] Lastly, in the difference of eq 4, the condensed-phase Gibbs free energies can be approximated by the DFT-computed total energies. In total, this allows to reformulate eq 4 to its working form

$$\gamma_{\text{surf}}^{(hkl),\nu_s}(\Delta\mu_{\text{O}}) = \frac{1}{2A^{(hkl),\nu_s}} [E_{\text{surf}}^{(hkl),\nu_s} - \nu_{\text{Ru}}^{(hkl),\nu_s} E_{RuO_2,\text{bulk}}]$$
$$- \frac{1}{2A^{(hkl),\nu_s}} \left[(\nu_{\text{O}}^{(hkl),\nu_s} - 2\nu_{\text{Ru}}^{(hkl),\nu_s})\left(\frac{1}{2}E_{O_2} + \Delta\mu_{\text{O}}\right) \right] \tag{5}$$
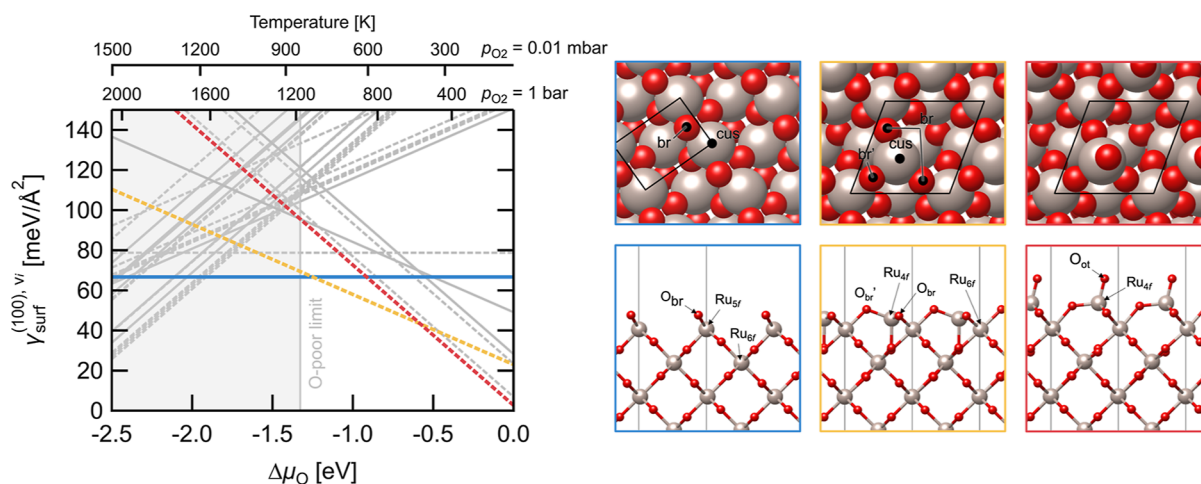
Figure 1. (Left) Surface free energies $\gamma_{surf}^{(100),\nu_i}$ as a function of the oxygen chemical potential $\Delta\mu_O$ for all previously known $RuO_2(100)$-(1×1) terminations (solid lines) and the new structures (dashed lines) identified here. The O-poor limit below which bulk rutile RuO_2 is unstable is indicated by a vertical dotted line.[51] In the upper x-axis, the chemical potential dependence is converted into a temperature dependence at given oxygen pressure. Three structures that are most stable in a range of O chemical potential above the bulk stability limited are highlighted in color: the conventional (1×1) stoichiometric termination (blue line), the $c(2 \times 2)$-* structure (yellow line), and the $c(2 \times 2)$-O structure (red line). (Right) Top and side views of these three structures. Ru and O atoms are depicted as large gray and small red spheres, respectively.

The thermodynamically most stable surface structure for any given facet orientation (hkl) at any given $\Delta\mu_O$ results as the one with the minimum surface free energy. Combining the minimum surface free energies of various facet orientations within a Wulff construction[55−58] allows to predict the equilibrium nanoparticle shape under the corresponding gas-phase conditions.

RESULTS AND DISCUSSION

$RuO_2(100)$-$c(2 \times 2)$ Reconstruction. The starting point of our iterative refinement workflow for the MLIP-based structure search for the $c(2 \times 2)$ reconstruction is the GAP trained in our previous work for the sub-space of low-index RuO_2-(1×1) surfaces.[22] This GAP in turn was itself the outcome of an iterative refinement. Its training was boot-strapped with a DFT training set comprising atomic information, gas phase O_2 dimer data with varying O−O bond lengths, and rutile RuO_2 bulk unit cells at optimized and constrained lattice constants with both optimized and displaced internal coordinates. The set also contained all inequivalent (1×1) surface structures that result from simple bulk truncation at various crystal planes,[59] where each time both the bulk truncated and the locally DFT optimized geometry was added to the pool. The iterative refinement proceeded in nine cycles and was terminated because no further unknown structures were found. In total, the SA runs identified 63 new surface structures, for 43 of which the protocol queried a subsequent DFT geometry optimization. The bulk of the CPU cost for training thus amounted to these DFT relaxations of (1×1) slabs.[22]

Bulk truncation of rutile RuO_2 at different planes along the (100) orientation gives rise to three (1×1) surface terminations with differing relative O/Ru stoichiometries: one metal-rich, one stoichiometric and one oxygen-rich termination. From the perspective of a $c(2 \times 2)$ aka $(\sqrt{2} \times \sqrt{2})R45°$ surface unit cell, these three relative stoichiometries can be described as a metal-rich termination with a 0% (metal-rich, $\nu_O = 2\nu_{Ru} - 2$), 50% (stoichiometric, $\nu_O = 2\nu_{Ru}$), and 100% (O-rich, $\nu_O = 2\nu_{Ru} + 2$) additional O

coverage. In terms of relative O/Ru stoichiometries, the doubled surface area of the $c(2 \times 2)$ cell thus allows for intermediate surface compositions with 25% ($\nu_O = 2\nu_{Ru} - 1$) and 75% ($\nu_O = 2\nu_{Ru} + 1$) additional O coverage. Note that for the sake of the chemical picture, we specified the various ν_i for one surface only. In the case of symmetric slabs with two equivalent surfaces, these numbers need to be adapted to match eqs 4 and 5 above.

Mirroring the above-described bootstrapping, we kick-start the GAP for the $c(2 \times 2)$ structure search by adding the ideal and the subsequently DFT geometry-optimized structure of the two new intermediate surface compositions (25 and 75%) to the training set. Essentially, the ideal structures result from removing three or one O atom(s) from the O-rich bulk-truncated (1×1) termination described in a $c(2 \times 2)$ cell, respectively. Each refinement cycle then spans five parallel SA runs for the five surface compositions (0, 25, 50, 75, and 100%). After 10 cycles, no further new structures are found and the refinement is considered converged. In total, 18 new symmetry-inequivalent structures with $c(2 \times 2)$ symmetry were found across all 5 compositions.

Figure 1 compares the relative stabilities of all these new and the previously known (1×1) structures from ref 22 within a surface phase diagram. In the assumed equilibrium with a surrounding oxygen gas phase, metal-rich surface compositions ($\nu_O < 2\nu_{Ru}$) have a positive slope with increasing O chemical potential, while O-rich surface compositions ($\nu_O > 2\nu_{Ru}$) have a negative slope. We note that our iterative refinement workflow involves the DFT geometry optimization of all newly identified structures. Figure 1 is completely based on DFT energetics. As can be seen, multiple new structures within the 0% surface composition class (most positive slope) are significantly more stable than the conventional metal-rich (1×1) termination that results after bulk truncation and local optimization. Eventually, the most stable of these new structures becomes even more stable than the conventional stoichiometric (1×1) termination at low O chemical potentials. However, the latter happens only at $\Delta\mu_O$ far

below the O-poor limit, indicating the stability boundary of bulk rutile RuO_2.

Much more relevant are therefore two new O-rich structures found within the 75 and 100% surface composition class (negative slopes), which we will henceforth denote as O-rich $c(2 \times 2)$-* and O-superrich $c(2 \times 2)$-O. For a range of $\Delta\mu_O$ above the O-poor limit, i.e., within the stability range of bulk rutile RuO_2, these two structures exhibit the lowest surface free energies and would correspondingly be identified as thermodynamically most stable in corresponding gas-phase conditions. Top and side views of these two $c(2 \times 2)$ reconstruction models are also shown in Figure 1. A visual comparison suggests that these are indeed exactly the same two structures that were also identified in the direct DFT-based global optimization work of Hess et al.[7] This perception is quantitatively confirmed by the evaluated negligible kernel distance of 0.006, cf. eq 2, between the respective O-rich $c(2 \times 2)$-* structures (unfortunately, no structure file for the O-superrich $c(2 \times 2)$-O was provided in ref 7). To further facilitate the comparison to the previous work, we have added the same labels for the varying prominent surface O and Ru atoms into the structure views in Figure 1 as those used by Hess et al.[7] As discussed there, the high stability of the reconstructions results from the formation of a tetrahedral Ru_{4f} surface complex, a motif that is well known as a highly stable molecule under oxidative conditions.[60]

Hess et al. also employed MD-based simulated annealing for their PES sampling. As this was directly performed with DFT, the excessive computational cost of the corresponding ab initio MD simulations forced them to apply an ultrafast quenching with a cooling rate about 500 times faster than ours (825 vs 1.6 K/ps). In total, they could also only afford three SA runs, at surface compositions 50, 75, and 100%. Even though they thus spent a total of 33,000 DFT single-point calculations of $c(2 \times 2)$ surface slabs for the total number of MD steps in all three runs, this only yielded the O-rich $c(2 \times 2)$-* structure as a result of the 75% SA run. The O-superrich $c(2 \times 2)$-O structure was subsequently manually created by adding one more O at the undercoordinated Ru_{4f} atom of the O-rich $c(2 \times 2)$-* structure. In contrast, the present MLIP-based approach found both structures without human interference and spending only a total of 1307 DFT single-point calculations of $c(2 \times 2)$ surface slabs during the geometry optimizations of new structures in the bootstrapping and refinement cycle starting from the (1×1) GAP-potential from ref 22. Even if one generously accounts for the CPU time spent in the training of this preceding (1×1) potential using undemanding smaller unit cells, this is a reduction in the CPU cost at least by ∼90%. At the same time, the iterative MLIP-based structure search affords a qualitatively increased PES sampling with multiple SA runs for each surface composition and much milder temperature protocols. Indeed, to highlight the benefit of this, we used the converged GAP for two final SA runs at surface compositions 75 and 100%. While during the prior experiment, the O-rich $c(2 \times 2)$-* structure formed readily after 50.7 ps (corresponding to a temperature of ≈280 K), the O-superrich $c(2 \times 2)$-O structure formed in the latter run only after 467.4 ps at around 900 K. Considering that the entire SA runs of Hess et al. only lasted 13.2 ps, this could rationalize why the latter structure could not be found in their search.

We stress though that the high computational efficiency of our iterative MLIP refinement workflow results in part from its clear focus on global geometry optimization. The resulting

GAP is by no means necessarily a comprehensive PES surrogate model. It is deliberately only trained to accurately cover (energetically low-lying) PES minima. Barrier regions should only be described to a level that they can be overcome in MD-based global optimization like simulated annealing. We illustrate this here with the minimum energy profile for the formation of the $c(2 \times 2)$-* reconstruction. Mechanistically, this formation may be initiated by the adsorption of an additional O atom atop an undercoordinated Ru_{5f} atom of the (1×1) stoichiometric termination. A neighboring Ru_{5f} then shifts laterally to form a new bond with this top O atom. Breaking two Ru−O back bonds to lattice oxygen atoms leads to the creation of the tetrahedrally coordinated Ru_{4f} complex.[7] If this process is modeled in a $c(2 \times 2)$ surface unit cell, this corresponds to the entire surface flipping into the reconstruction at once. We therefore also consider a larger $c(4 \times 4)$ surface unit cell, where four such complexes can be formed one at a time. The DFT-calculated barriers for the collective flip in the $c(2 \times 2)$ cell, as well as the first and fourth flip in the $c(4 \times 4)$ cell, are contrasted to the barriers calculated with the converged GAP in Table 1, with the detailed energy profiles

Table 1. Energetic Barriers for the Formation of the $c(2 \times 2)$-* Reconstruction through Flipping into a Ru_{4f} Complex as Calculated by DFT and the Converged GAP[a]

|	energy barrier (eV)		
	collective flip	first flip	fourth flip
DFT	0.43	0.40	0.45
GAP	0.16	0.14	0.10

[a]See text for the three mechanisms considered in different surface unit cells. The at best semi-quantitative agreement highlights that the GAP trained for geometry optimization should not be misinterpreted as a general-purpose RuO_2 force field.

shown in Figure S3. While the barriers for the three different mechanisms do not vary much and are reasonably low to rationalize a fast kinetics towards the reconstruction, the corresponding GAP barriers agree only at best semi-quantitatively. Care has to be taken not to misinterpret the trained GAP as a general-purpose force field for RuO_2. On the other hand, nothing would conceptually prevent the present iterative refinement workflow to be adapted from minima to transition state search. Finally, we note in passing that our DFT barriers for this process are substantially lower than the one reported earlier by Hess et al. (0.92 eV),[7] a discrepancy which we tentatively attribute to an omission of the factor of two in their symmetric slab calculations with two surfaces.

$RuO_2(110)/(100)$ Facet Edge. The demonstrated high computational efficiency of the iterative training workflow should in principle increase further when moving to even larger surface-area problems, as the GAP trained up to this stage has already learned quite a variety of local geometric motifs. As long as the large surface-area problem addressed does not heavily draw on completely new motifs, only a restricted number of additional and costly DFT training structures in the large surface-area supercell should then be required to accurately sample the new problem with the surrogate potential. We demonstrate this idea for the $RuO_2(110)/(100)$ facet edge. The detailed atomic structure of this edge is of particular interest, as high-resolution transmission electron microscopy images indicated the presence of a (410) vicinal between the (110) and (100) facets of high-temperature air-

calcined (900 °C) RuO_2 nanoparticles that showed an improved selectivity for the chlorine evolution reaction.[33]

Figure 2 shows side views of bulk-truncated (1×1) surface unit cells of the (410) vicinal. The metal layers in this
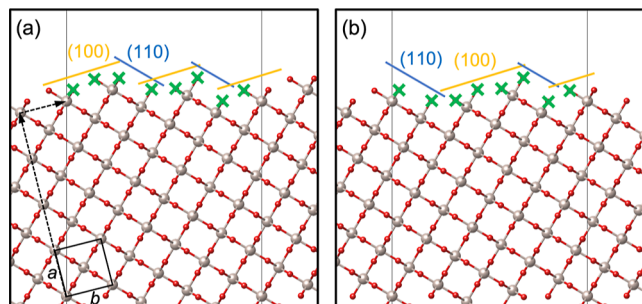

Figure 2. Side views of bulk-truncated (1×1) O-rich $RuO_2(410)$ surfaces. (a) Termination with the topmost layer comprising RuO_6 octahedra pointing with their apical axis into the $(1\bar{1}0)$ direction. (b) Termination with the topmost layer comprising RuO_6 octahedra pointing with their apical axis into the (110) direction. Both terminations exhibit alternating micro-facets in (100) and (110) direction, as indicated by yellow and blue lines, respectively. Oxygen atoms that are successively removed to create different relative stoichiometries for the SA runs are marked with green crosses (see text). The bulk unit cell with lattice parameters a and b is shown as a black square, and (400) and (010) vectors are shown as dashed black arrows. Ru and O atoms are depicted as large gray and small red spheres, respectively.

orientation are either comprised of RuO_6 octahedra pointing with their apical axis into the (110) direction or of RuO_6 octahedra pointing with their apical axis into the $(1\bar{1}0)$ direction. Depending on the truncation, this thus gives rise to the two different O-rich terminations shown in Figure 2a,b. By successively removing up to 8 surface O atoms in either of these two terminations, a total of 18 different terminations can be created that span the entire range of relative O−Ru stoichiometries possible within this surface unit cell before the next layering starts. Note that it does not matter which specific or which combination of O atoms are removed, as there is

sufficient mass transport during the high-temperate part of the SA global geometry optimization. Our GAP refinement cycles thus comprise SA runs for each of these 18 relative stoichiometries. After 9 cycles, a low out-of-sample error signals that the GAP has learned all local geometry motifs occurring in this surface unit cell sufficiently well. We then double the size to a (2×1) cell, which is the smallest possible cell size that would allow for the formation of the $c(2 \times 2)$ reconstruction on (100)-oriented mini-facets. Again covering all relative stoichiometries possible in this (2×1) cell, the out-of-sample error stop criterion is reached after another seven refinement cycles. Despite the drastically increased size of these structures, in total, only another 2060 and 1000 DFT single-point calculations for (1×1)- and (2×1)-(410) slabs were queried for geometry optimizations of new training structures during these refinements, respectively. Even though the GAP is then trained sufficiently, new structures are still continuously found in further SA runs. However, these structures only correspond to energetically essentially degenerate rearrangements of the same geometric motifs in the huge possible phase space of (2×1) surface unit cells. While the energetics discussed in the following is thus converged, the specific structures shown should therefore only be seen as representative for a corresponding entire class of structures featuring the various geometric motifs.

Figure 3 compares the surface free energies of the most stable structures identified in the global geometry optimization over the relevant range of oxygen chemical potentials. Specifically, we compare the results obtained in the (2×1) surface unit cell with those obtained in the more restricted (1×1) cell. As apparent, the larger cell allows to stabilize new structures that are significantly more stable. In detail, there are three structures (or better energetically degenerate structure classes) that are each most stable over a certain $\Delta\mu_O$-range. The representative structures for each class shown in Figure 3 reveal that the stabilization is in all cases due to the formation of the tetrahedral Ru_{4f} motifs that are also the stabilizing core of the afore discussed $c(2 \times 2)$-(100) reconstruction. In the first structure stable up to $\Delta\mu_O = -0.87$ eV, protruding Ru surface atoms sink down to form an additional bond to a sub-
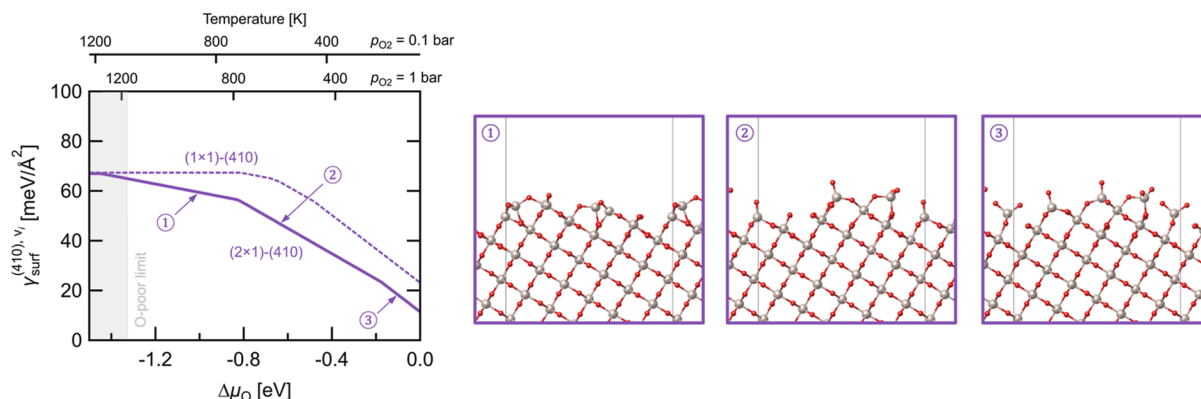
Figure 3. (Left) Surface free energies $\gamma_{surf}^{(410),\nu_i}$ for the (410) vicinal as a function of the oxygen chemical potential $\Delta\mu_O$. Results obtained after global geometry optimization in the (2×1) (solid purple line) and in the restricted (1×1) (dashed purple line) surface unit cells are shown. In each case, only the surface free energy of the most stable structure is drawn at each chemical potential. Kinks in the line thus indicate a change to another structure with a different relative oxygen−Ru stoichiometry. The O-poor limit below which bulk rutile RuO_2 is unstable is indicated by a vertical dotted line.[51] In the upper x-axis, the chemical potential dependence is converted into a temperature dependence at given oxygen pressure. (Right) Side views of three representative stable (2×1)-(410) structures highlighted on the left. Ru and O atoms are depicted as large gray and small red spheres, respectively.

surface O atom, exactly as in the $c(2 \times 2)$-* reconstruction, cf. Figure 1—thereby smoothening out the corrugation of (110) and (100) microfacets. At higher oxygen chemical potentials, the other two structures then exhibit additional tetrahedral Ru_{4f} complexes as in the $c(2 \times 2)$-O reconstruction. These complexes are bound to the surface via either two or three O vertices. In particular, the two-vertex motif has recently been proposed as an intermediate species in the anodic corrosion process.[61] Its increased occurrence in particular in the structure class that is most stable at the higher chemical potentials, $\Delta\mu_O > -0.17$ eV, would then point to a particular role of the (110)/(100) facet edge in the decomposition process.

Implications for the Wulff Particle Shape. The strong reduction of the (100) and (410) surface free energies induced by the $c(2 \times 2)$ reconstructions, cf. Figures 1 and 3, should have bearings on the equilibrium shape of RuO_2 nanoparticles. The corresponding shapes have been computed through a Wulff construction by Wang et al.[66] This study systematically considered all five symmetry-inequivalent low-index rutile facets and all possible (1×1) bulk-truncated terminations at them. Locally geometry optimizing these structures by direct DFT calculations, the surface free energies were obtained within the same ab initio thermodynamics approach as also used here. Figure 4 shows Wulff shapes at varying oxygen
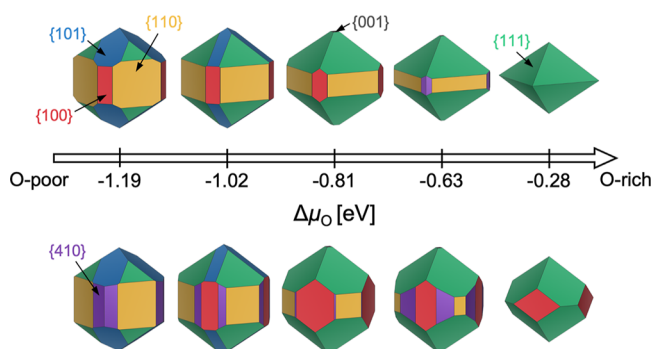
Figure 4. Equilibrium Wulff nanoparticle shapes computed from surface free energies of all five low-index facets and the (410) vicinal, using (top) locally geometry optimized (1×1) structures and (bottom) global geometry optimized structures (see text). Five indicated O chemical potentials correspond to the calcination pretreatment conditions used in experiments by Rosenthal et al.: 1073 K, 1 bar,[62,63] Jirkovský et al.: 873 K, 0.1 bar,[33] Lee et al.: 773 K, 1 bar,[64] and Narkhede et al.: 573 K in 0.1 bar,[65] as well as at standard conditions of 300 K and 1 bar, from left to right.

chemical potentials that we obtain with the present computational setup and drawing on exactly the same set of structures, plus also the 18 locally geometry optimized (1×1) structures that were employed to start the SA runs for the (410) vicinal. The shapes obtained are essentially identical to those reported by Wang et al., with minor quantitative differences mostly arising from the different DFT exchange−correlation functional employed by them (PBE[67] vs the RPBE functional employed here).

Overall, there is a strong change in the predicted equilibrium shape toward more O-rich conditions. An octagonal cross section at O-poor conditions gradually transforms into a quadratic cross section, which goes hand in hand with a change from a more columnar to an ultimately lenticular crystal habit. The main driver for this is the strong lowering of the surface free energy of O-rich (111) structures with increasing O

chemical potential, as shown in Figure S6 in the Supporting Information. While the capping facets in the O-poor limit are predominantly {101}, only {111} facets are thus eventually exhibited at all at O-rich standard conditions, cf. Figure 4. Notably, the {410} vicinal only appears with a negligible share of the nanoparticle surface in a very small range around $\Delta\mu_O = -0.63$ eV. Under any more O-poor conditions, the apical facets are exclusively formed by directly touching {110} and {100}.

The predicted shapes can be compared to a range of experiments that used different calcination pretreatments and characterized the resulting RuO_2 crystal shapes by electron microscopy.[33,62−65] The specific $\Delta\mu_O$-values chosen in Figure 4 correspond to these pretreatment conditions. By and large, the predicted shapes are consistent with the experimental ones obtained under more O-poor conditions, i.e., at $\Delta\mu_O = -1.19$ eV[62,63] and $\Delta\mu_O = -1.02$ eV.[33] However, neither do the Wulff shapes exhibit the experimentally observed pronounced share of {410} lateral facets and the concomitant irregular octagonal shape in this $\Delta\mu_O$-range, nor was the predicted strong change to fully lenticular habit seen in the experiments with more O-rich pretreatments.[65] While in the latter, lower temperature calcination the experimental shapes could still partially be affected by kinetics, Wang et al.[66] suspected in their original study that the omission of the at the time structurally unknown $c(2 \times 2)$ reconstruction could be behind these differences between computed and measured nanoparticle shapes.

Figure 4 also shows the predicted Wulff shapes, if the most stable (100) and (410) structures identified by global geometry optimization in this study are additionally considered. Indeed, the $c(2 \times 2)$ reconstruction motif thus taken into account has an effect and brings the computed shapes into better agreement with experiment. {410} vicinals are now strongly present as lateral facets for the most O-poor chemical potentials, leading to an irregular octagonal cross section. In addition, the strong (100) surface free energy reduction due to the $c(2 \times 2)$-O structure, cf. Figure 1 and S7 in the Supporting Information, stabilizes the presence of these facets in the Wulff shape up until the highest $\Delta\mu_O$, thereby delaying the complete switch to lenticular habit. Indeed, {410} vicinals exhibiting the two-vertex bound Ru_{4f} motif, cf. Figure 3, also have a more pronounced share at these most O-rich conditions, supporting the suggestion of Hess and Over[61] that these species play a role in the anodic corrosion process. Now that structural models are established for this (110)/(100) facet edge, future detailed mechanistic studies may hence develop an atomic-scale understanding of this degradation, as well as of the connection of this edge with the improved chlorine evolution reaction selectivity reported for the high-temperature, O-poor calcined RuO_2 nanoparticles of Jirkovský et al.[33]

■ CONCLUSIONS

Using ML interatomic potentials as surrogates for first-principles DFT calculations strongly boosts the sampling capabilities in predictive-quality global surface structure determination. However, this boost only becomes effective, if the total human and CPU time invested into the generation of the first-principles training data remains low compared to the costs incurred in a direct DFT-based sampling. The resulting imperative of high data efficiency is difficult to reconcile with static a priori created data bases, given that the objective of the sampling is to identify yet unknown geometric motifs. Iterative or active learning protocols address this by querying new DFT training data and refining the surrogate potential whenever the

global optimization process identifies new geometric motifs not yet appropriately captured in the data set. This way, the algorithm focuses the CPU-expensive DFT calculations on those structures really needed.

Here, we have further considered that the CPU cost rises strongly with the size of the training structure that needs to be computed. In surface structure determination, this applies particularly to large surface unit cell problems that then require huge supercell calculations. We correspondingly extended our previously described iterative training protocol to proceed in stages, where much of the ML surrogate training occurs already during the initial sampling of small unit cell problems. With many geometric motifs thus already learned from such inexpensive small unit cell training data, less data queries and refinement are then necessary during the sampling of the exploding phase space of larger surface unit cell problems.

We have illustrated this automated approach by applying and refining our previously reported GAP potential trained on $RuO_2(1 \times 1)$ low-index surface structures to first explore (100)-$c(2 \times 2)$ cells and then (410)-(1×1) and -(2×1) cells. For both the prior low index and the latter vicinal, a reconstruction involving tetrahedral Ru_{4f} motifs proves critical. This long time enigmatic reconstruction motif had only recently been identified in a mixture of direct DFT-based global optimization and handmade extensions and we here recover it fully automatically at a minute fraction of the CPU cost. The concomitant strong lowering of the surface free energy of both surfaces over a wide range of O chemical potentials extends their presence on the surface of Wulff equilibrium nanoparticles and leads to an improved agreement of the predicted shapes with experimental data. Notably, the (410) vicinal is stabilized as a facet edge between lateral (110) and (100) facets. While the established structural motifs at this edge can serve as the basis for future mechanistic studies on the reported role of this edge in anodic corrosion and in improved selectivities, the degeneracy of structures involving these motifs could also be relevant. The GAP refinement for these larger surface unit cell problems stopped as no further new geometric motifs were identified. However, in particular for the largest (410)-(2×1) cell, the global geometry sampling continued to find more and more structures that all involved different arrangements of the same $c(2 \times 2)$-reconstruction motifs. At finite temperatures, all these energetically essentially degenerate structures will be visited and the concomitant structural fluxionality of the facet edge might be key to its function.

■ ASSOCIATED CONTENT

Data Availability Statement

The data that support the findings of this study can be found in the paper and its supplementary files. The relevant structures and the final training database are available at https://doi.org/10.6084/m9.figshare.23833941. Further data are also accessible upon reasonable request from the corresponding author.

ⓘ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpcc.3c04049.

> Details of force locality test, chosen GAP hyperparameters, the findings of this study, calculated surface free energy of all identified reconstructions for the $RuO_2(100)$ facet, the potential energy profiles obtained from NEB calculations, the evolution of kernel distance

during the simulated annealing runs of the converged GAP model, the learning curves representing the iterative GAP training procedure for RuO_2, and the lowest surface energies of symmetry-inequivalent facets for both local and global optimization scenarios (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Karsten Reuter − *Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany;* ⓞ orcid.org/0000-0001-8473-8659; Email: reuter@fhi-berlin.mpg.de

Authors

Yonghyuk Lee − *Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany;* ⓞ orcid.org/0000-0001-8961-2254

Jakob Timmermann − *Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany*

Chiara Panosetti − *Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany*

Christoph Scheurer − *Fritz-Haber-Institut der Max-Planck-Gesellschaft, D-14195 Berlin, Germany*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jpcc.3c04049

■ REFERENCES

(1) Kresse, G.; Bergermayer, W.; Podloucky, R.; Lundgren, E.; Koller, R.; Schmid, M.; Varga, P. Complex Surface Reconstructions Solved by Ab Initio Molecular Dynamics. *Appl. Phys. A* **2003**, *76*, 701−710.

(2) Vilhelmsen, L. B.; Hammer, B. A Genetic Algorithm for First Principles Global Structure Optimization of Supported Nano Structures. *J. Chem. Phys.* **2014**, *141*, 044711.

(3) Panosetti, C.; Krautgasser, K.; Palagin, D.; Reuter, K.; Maurer, R. J. Global Materials Structure Search with Chemically Motivated Coordinates. *Nano Lett.* **2015**, *15*, 8044−8048.

(4) Hussein, H. A.; Davis, J. B. A.; Johnston, R. L. DFT Global Optimisation of Gas-Phase and MgO-Supported Sub-Nanometre AuPd Clusters. *Phys. Chem. Chem. Phys.* **2016**, *18*, 26133−26143.

(5) Wexler, R. B.; Qiu, T.; Rappe, A. M. Automatic Prediction of Surface Phase Diagrams Using Ab Initio Grand Canonical Monte Carlo. *J. Phys. Chem. C* **2019**, *123*, 2321−2328.

(6) Zhou, Y.; Scheffler, M.; Ghiringhelli, L. M. Determining Surface Phase Diagrams Including Anharmonic Effects. *Phys. Rev. B* **2019**, *100*, 174106.

(7) Hess, F.; Rohrlack, S.; Knapp, M.; Over, H. Evidence of a Tetrahedrally Coordinated RuO_4 Surface Complex on $RuO_2(100)$: Density Functional Theory and Beyond. *J. Phys. Chem. C* **2022**, *126*, 946−956.

(8) Deringer, V. L.; Caro, M. A.; Csányi, G. Machine Learning Interatomic Potentials as Emerging Tools for Materials Science. *Adv. Mater.* **2019**, *31*, 1902765.

(9) Willatt, M. J.; Musil, F.; Ceriotti, M. Atom-Density Representations for Machine Learning. *J. Chem. Phys.* **2019**, *150*, 154110.

(10) Zuo, Y.; Chen, C.; Li, X.; Deng, Z.; Chen, Y.; Behler, J.; Csányi, G.; Shapeev, A. V.; Thompson, A. P.; Wood, M. A.; et al. Performance and Cost Assessment of Machine Learning Interatomic Potentials. *J. Phys. Chem. A* **2020**, *124*, 731−745.

(11) Deringer, V. L.; Bartók, A. P.; Bernstein, N.; Wilkins, D. M.; Ceriotti, M.; Csányi, G. Gaussian Process Regression for Materials and Molecules. *Chem. Rev.* **2021**, *121*, 10073−10141.

(12) Musil, F.; Grisafi, A.; Bartók, A. P.; Ortner, C.; Csányi, G.; Ceriotti, M. Physics-Inspired Structural Representations for Molecules and Materials. *Chem. Rev.* **2021**, *121*, 9759−9815.

(13) Kocer, E.; Ko, T. W.; Behler, J. Neural Network Potentials: A Concise Overview of Methods. *Annu. Rev. Phys. Chem.* **2022**, *73*, 163−186.

(14) Margraf, J. T.; Jung, H.; Scheurer, C.; Reuter, K. Exploring Catalytic Reaction Networks with Machine Learning. *Nat. Catal.* **2023**, *6*, 112−121.

(15) Bisbo, M. K.; Hammer, B. Efficient Global Structure Optimization with a Machine-Learned Surrogate Model. *Phys. Rev. Lett.* **2020**, *124*, 086102.

(16) Timmermann, J.; Kraushofer, F.; Resch, N.; Li, P.; Wang, Y.; Mao, Z.; Riva, M.; Lee, Y.; Staacke, C.; Schmid, M.; et al. IrO$_2$ Surface Complexions Identified through Machine Learning and Surface Investigations. *Phys. Rev. Lett.* **2020**, *125*, 206101.

(17) Kaappa, S.; Del Río, E. G.; Jacobsen, K. W. Global Optimization of Atomic Structures with Gradient-Enhanced Gaussian Process Regression. *Phys. Rev. B* **2021**, *103*, 174114.

(18) Rønne, N.; Christiansen, M.-P. V.; Slavensky, A. M.; Tang, Z.; Brix, F.; Pedersen, M. E.; Bisbo, M. K.; Hammer, B. Atomistic Structure Search Using Local Surrogate Model. *J. Chem. Phys.* **2022**, *157*, 174115.

(19) Jung, H.; Sauerland, L.; Stocker, S.; Reuter, K.; Margraf, J. T. Machine-Learning Driven Global Optimization of Surface Adsorbate Geometries. *npj Comput. Mater.* **2023**, *9*, 114.

(20) Jinnouchi, R.; Miwa, K.; Karsai, F.; Kresse, G.; Asahi, R. On-The-Fly Active Learning of Interatomic Potentials for Large-Scale Atomistic Simulations. *J. Phys. Chem. Lett.* **2020**, *11*, 6946−6955.

(21) Verdi, C.; Karsai, F.; Liu, P.; Jinnouchi, R.; Kresse, G. Thermal Transport and Phase Transitions of Zirconia by On-The-Fly Machine-Learned Interatomic Potentials. *npj Comput. Mater.* **2021**, *7*, 156.

(22) Timmermann, J.; Lee, Y.; Staacke, C. G.; Margraf, J. T.; Scheurer, C.; Reuter, K. Data-Efficient Iterative Training of Gaussian Approximation Potentials: Application to Surface Structure Determination of Rutile IrO$_2$ and RuO$_2$. *J. Chem. Phys.* **2021**, *155*, 244107.

(23) Vandermause, J.; Xie, Y.; Lim, J. S.; Owen, C. J.; Kozinsky, B. Active Learning of Reactive Bayesian Force Fields Applied to Heterogeneous Catalysis Dynamics of H/Pt. *Nat. Commun.* **2022**, *13*, 5183.

(24) Schaaf, L.; Fako, E.; De, S.; Schäfer, A.; Csányi, G. Accurate Reaction Barriers for Catalytic Pathways: An Automatic Training Protocol for Machine Learning Force Fields. 2023, arXiv:2301.09931. https://doi.org/10.48550/arXiv.2301.09931.

(25) Over, H. Surface Chemistry of Ruthenium Dioxide in Heterogeneous Catalysis and Electrocatalysis: From Fundamental to Applied Research. *Chem. Rev.* **2012**, *112*, 3356−3426.

(26) Over, H. Atomic Scale Insights into Electrochemical Versus Gas Phase Oxidation of HCl over RuO$_2$-Based Catalysts: A Comparative Review. *Electrochim. Acta* **2013**, *93*, 314−333.

(27) Over, H. Fundamental Studies of Planar Single-Crystalline Oxide Model Electrodes (RuO$_2$, IrO$_2$) for Acidic Water Splitting. *ACS Catal.* **2021**, *11*, 8848−8871.

(28) Wang, C.; Jin, L.; Shang, H.; Xu, H.; Shiraishi, Y.; Du, Y. Advances in Engineering RuO$_2$ Electrocatalysts Towards Oxygen Evolution Reaction. *Chin. Chem. Lett.* **2021**, *32*, 2108−2116.

(29) Panosetti, C.; Lee, Y.; Samtsevych, A.; Scheurer, C. Black Box vs Gray Box: Comparing GAP and GPrep-DFTB for Ruthenium and Ruthenium oxide. *J. Chem. Phys.* **2023**, *158*, 224115.

(30) Kim, Y. D.; Schwegmann, S.; Seitsonen, A. P.; Over, H. Epitaxial Growth of RuO$_2$(100) on Ru(1010): Surface Structure and Other Properties. *J. Phys. Chem. B* **2001**, *105*, 2205−2211.

(31) Knapp, M.; Seitsonen, A. P.; Kim, Y. D.; Over, H. Catalytic Activity of the RuO$_2$(100) Surface in the Oxidation of CO. *J. Phys. Chem. B* **2004**, *108*, 14392−14397.

(32) Aßmann, J.; Crihan, D.; Knapp, M.; Lundgren, E.; Löffler, E.; Muhler, M.; Narkhede, V.; Over, H.; Schmid, M.; Seitsonen, A. P.; et al. Understanding the Structural Deactivation of Ruthenium Catalysts on an Atomic Scale under Both Oxidizing and Reducing Conditions. *Angew. Chem., Int. Ed.* **2005**, *44*, 917−920.

(33) Jirkovský, J.; Hoffmannová, H.; Klementová, M.; Krtil, P. Particle Size Dependence of the Electrocatalytic Activity of Nanocrystalline RuO$_2$ Electrodes. *J. Electrochem. Soc.* **2006**, *153*, No. E111.

(34) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Phys. Rev. Lett.* **2010**, *104*, 136403.

(35) Bartók, A. P.; Kondor, R.; Csányi, G. On Representing Chemical Environments. *Phys. Rev. B* **2013**, *87*, 184115.

(36) Stocker, S.; Csányi, G.; Reuter, K.; Margraf, J. T. Machine Learning in Chemical Reaction Space. *Nat. Commun.* **2020**, *11*, 5505−5511.

(37) Cheng, B.; Griffiths, R.-R.; Wengert, S.; Kunkel, C.; Stenczel, T.; Zhu, B.; Deringer, V. L.; Bernstein, N.; Margraf, J. T.; Reuter, K.; et al. Mapping Materials and Molecules. *Acc. Chem. Res.* **2020**, *53*, 1981−1991.

(38) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.

(39) Swope, W. C.; Andersen, H. C.; Berens, P. H.; Wilson, K. R. A Computer Simulation Method for the Calculation of Equilibrium Constants for the Formation of Physical Clusters of Molecules: Application to Small Water Clusters. *J. Chem. Phys.* **1982**, *76*, 637−649.

(40) Music, D.; Stelzer, B. Intrinsic Thermal Shock Behavior of Common Rutile Oxides. *Physics* **2019**, *1*, 290−300.

(41) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−3690.

(42) Hamann, D. R. Optimized Norm-Conserving Vanderbilt Pseudopotentials. *Phys. Rev. B* **2013**, *88*, 085117.

(43) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; et al. QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.: Condens.Matter* **2009**, *21*, 395502.

(44) Hammer, B.; Hansen, L. B.; Nørskov, J. K. Improved Adsorption Energetics within Density-Functional Theory Using Revised Perdew-Burke-Ernzerhof Functionals. *Phys. Rev. B* **1999**, *59*, 7413−7421.

(45) Broyden, C. G. The Convergence of a Class of Double-Rank Minimization Algorithms: 2. The New Algorithm. *IMA J. Appl. Math.* **1970**, *6*, 222−231.

(46) Goldfarb, D. A Family of Variable-Metric Methods Derived by Variational Means. *Math. Comput.* **1970**, *24*, 23−26.

(47) Shanno, D. F. Conditioning of Quasi-Newton Methods for Function Minimization. *Math. Comput.* **1970**, *24*, 647−656.

(48) Garrido Torres, J. A.; Jennings, P. C.; Hansen, M. H.; Boes, J. R.; Bligaard, T. Low-Scaling Algorithm for Nudged Elastic Band Calculations Using a Surrogate Machine Learning Model. *Phys. Rev. Lett.* **2019**, *122*, 156001.

(49) Koistinen, O.-P.; Dagbjartsdóttir, F. B.; Ásgeirsson, V.; Vehtari, A.; Jónsson, H. Nudged Elastic Band Calculations Accelerated with Gaussian Process Regression. *J. Chem. Phys.* **2017**, *147*, 152720.

(50) Garijo del Río, E.; Mortensen, J. J.; Jacobsen, K. W. Local Bayesian Optimizer for Atomic Structures. *Phys. Rev. B* **2019**, *100*, 104103.

(51) Reuter, K.; Scheffler, M. Composition, Structure, and Stability of RuO$_2$(110) as a Function of Oxygen Pressure. *Phys. Rev. B* **2001**, *65*, 035406.

(52) Reuter, K. Ab Initio Thermodynamics and First-Principles Microkinetics for Surface Catalysis. *Catal. Lett.* **2016**, *146*, 541−563.

(53) Irikura, K. K. Experimental Vibrational Zero-Point Energies: Diatomic Molecules. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389−397.

(54) Herzberg, G. *Molecular Spectra and Molecular Structure II. Infrared and Raman Spectra of Polyatomic Molecules*; D. Van Nostrand company, 1945; p 632.

(55) Wulff, G. Xxv. zur Frage der Geschwindigkeit des Wachsthums und der Auflösung der Krystallflächen. *Z. Kristallogr.* **1901**, *34*, 449−530.

(56) Herring, C. Some Theorems on the Free Energies of Crystal Surfaces. *Phys. Rev.* **1951**, *82*, 87−93.

(57) Rogal, J.; Reuter, K.; Scheffler, M. Thermodynamic Stability of PdO Surfaces. *Phys. Rev. B* **2004**, *69*, 075421.

(58) Li, H.; Reuter, K. Ab Initio Thermodynamic Stability of Carbide Catalysts under Electrochemical Conditions. *ACS Catal.* **2022**, *12*, 10506−10513.

(59) Lee, Y.; Scheurer, C.; Reuter, K. Epitaxial Core-Shell Oxide Nanoparticles: First-Principles Evidence for Increased Activity and Stability of Rutile Catalysts for Acidic Oxygen Evolution. *ChemSusChem* **2022**, *15*, No. e202200015.

(60) Kötz, R.; Stucki, S.; Scherson, D.; Kolb, D. M. In-Situ Identification of RuO_4 as the Corrosion Product During Oxygen Evolution on Ruthenium in Acid Media. *J. Electroanal. Chem. Interfacial Electrochem.* **1984**, *172*, 211−219.

(61) Hess, F.; Over, H. Coordination Inversion of the Tetrahedrally Coordinated Ru_{4f} Surface Complex on $RuO_2(100)$ and Its Decisive Role in the Anodic Corrosion Process. *ACS Catal.* **2023**, *13*, 3433−3443.

(62) Rosenthal, D.; Girgsdies, F.; Timpe, O.; Blume, R.; Weinberg, G.; Teschner, D.; Schlögl, R. On the CO-Oxidation over Oxygenated Ruthenium. *Z. Phys. Chem.* **2009**, *223*, 183−208.

(63) Rosenthal, D.; Girgsdies, F.; Timpe, O.; Weinberg, G.; Schlögl, R. Oscillatory Behavior in the CO-Oxidation over Bulk Ruthenium Dioxide - The Effect of the CO/O_2 Ratio. *Z. Phys. Chem.* **2011**, *225*, 57−68.

(64) Lee, Y.; Suntivich, J.; May, K. J.; Perry, E. E.; Shao-Horn, Y. Synthesis and Activities of Rutile IrO_2 and RuO_2 Nanoparticles for Oxygen Evolution in Acid and Alkaline Solutions. *J. Phys. Chem. Lett.* **2012**, *3*, 399−404.

(65) Narkhede, V.; Aßmann, J.; Muhler, M. Structure-Activity Correlations for the Oxidation of CO over Polycrystalline RuO_2 Powder Derived from Steady-State and Transient Kinetic Experiments. *Z. Phys. Chem.* **2005**, *219*, 979−995.

(66) Wang, T.; Jelic, J.; Rosenthal, D.; Reuter, K. Exploring Pretreatment−Morphology Relationships: Ab Initio Wulff Construction for RuO_2 Nanoparticles Under Oxidising Conditions. *ChemCatChem* **2013**, *5*, 3398−3403.

(67) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.