Active Learning Accelerated Discovery of Stable Iridium Oxide Polymorphs for the Oxygen Evolution Reaction

Flores, Raul A.; Paolucci, Christopher; Winther, Kirsten T.; Jain, Ankit; Torres, Jose Antonio Garrido; Aykol, Muratahan; Montoya, Joseph; Norskov, Jens K.; Bajdich, Michal; Bligaard, Thomas

Published in:
Chemistry of Materials

Link to article, DOI:
10.1021/acs.chemmater.0c01894

Publication date:
2020

Document Version
Peer reviewed version

Active Learning Accelerated Discovery of Stable Iridium-oxide Polymorphs for the Oxygen Evolution Reaction

Raul A. Flores,[†,‡] Christopher Paolucci,[¶,†,‡] Kirsten T. Winther,[†,‡] Ankit Jain,[§,†,‡] Jose Antonio Garrido Torres,[†,‡] Muratahan Aykol,[‖] Joseph Montoya,[‖] Jens K. Nørskov,[⊥] Michal Bajdich,[*,‡] and Thomas Bligaard[*,#,‡]

† *SUNCAT Center for Interface Science and Catalysis, Department of Chemical Engineering, Stanford University, Stanford 94305, California, USA*

‡ *SUNCAT Center for Interface Science and Catalysis, SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA*

¶ *Department of Chemical Engineering, University of Virginia, Charlottesville, Virginia 22903, United States*

§ *Department of Mechanical Engineering, Indian Institute of Technology Bombay, Powai, India*

‖ *Toyota Research Institute, Los Altos, CA 94022, USA*

⊥ *Department of Physics, Technical University of Denmark, 2800 Kongens Lyngby, Denmark*

\# *Department of Energy Conversion and Storage, Technical University of Denmark, 2800 Kongens Lyngby, Denmark*

E-mail: bajdich@slac.stanford.edu; tbli@dtu.dk

Abstract

The discovery of high-performing and stable materials for sustainable energy applications is a pressing goal in catalysis and materials science. Understanding the relationship between a material's structure and functionality is an important step in the process, such that viable polymorphs for a given chemical composition need to be identified. Machine-learning based surrogate models have the potential to accelerate the search for polymorphs that target specific applications. Herein, we report a readily generalizable active-learning (AL) accelerated algorithm for identification of electrochemically stable iridium-oxide polymorphs of IrO_2 and IrO_3. The search is coupled to a subsequent analysis of the electrochemical stability of the discovered structures for the acidic oxygen evolution reaction (OER). Structural candidates are generated by identifying all 956 structurally unique AB_2 and AB_3 prototypes in existing materials databases (more than 38,000). Next, using an active learning approach we are able to find 196 IrO_2 polymorphs within the thermodynamic amorphous synthesizability limit and reaffirm the global stability of the rutile structure. We find 75 synthesizable IrO_3 polymorphs and report a previously unknown FeF_3-type structure as the most stable, termed α-IrO_3. To test the algorithms performance, we compare to a random search of the candidate space and report at least a twofold increase in the rate of discovery. Additionally, the AL approach can acquire the most stable polymorphs of IrO_2 and IrO_3 with less than 30 density functional theory optimizations. Analysis of the structural properties of the discovered polymorphs reveals that octahedral local coordination environments are preferred for nearly all low energy structures. Subsequent Pourbaix Ir-H_2O analysis shows that α-IrO_3 is the globally stable solid phase under acidic OER conditions and supersedes the stability of rutile IrO_2. Calculation of theoretical OER surface activities reveal ideal weaker binding of the OER intermediates on α-IrO_3 than on any other considered iridium-oxide. We emphasize that the proposed AL algorithm can be easily generalized to search for any binary metal-oxide structure with a defined stoichiometry.

Introduction

To understand or simulate the properties of novel polymorphs of functional materials their crystal structure must first be solved for, which remains a challenging problem in materials science.[1,2] Large high-throughput ab initio datasets[3–6] have enabled approaching many problems in materials research with machine-learning,[7] but these datasets are systematically biased towards known materials or hypothetical materials derived from common crystal prototypes. Thus, there is a need for the systematic exploration of structural diversity at target elemental compositions.

Contemporary approaches to inorganic crystal structure prediction include a variety of methods that explore the expansive potential energy landscape, and include simulated annealing, evolutionary algorithms, and particle swarm optimization.[8–15] These approaches are comprehensive, but become intractable as the number of polymorphic configurations increases exponentially with the number and types of elements considered.[16] Recent materials discovery approaches employing surrogate models in lieu of Density Functional Theory (DFT) calculations include a tight-binding model utilizing genetic algorithms,[17] agent-based rapid generation of phase diagrams in diverse chemistries,[18] and an image-based materials representation procedure from Noh et al.,[19] which was used to find stable vanadium oxide polymorphs. Active learning (AL) frameworks in conjunction with surrogate models have emerged as a computationally efficient approach for problems in materials science, and a potential alternative to established crystal structure prediction (CSP) methods.[20–24]

In this manuscript we report a rapid crystal structure discovery approach that leverages machine learning surrogate models and an AL framework to accelerate the discovery of polymorphs at target chemical compositions. Our method does not rely on the existence of past DFT data, but instead sequentially generates the minimum sized data set to effectively search within a structurally diverse space of candidates generated from crystal structure prototypes. We demonstrate the application of this methodology in the space of iridium oxide polymorphs, an important class of materials with applications in electrochemistry,

but with unresolved structure-activity properties critical for understanding their catalytic activity. In particular, Rutile-IrO_2 (R-IrO_2) (Ir^{4+}), the most stable form of iridium-oxide at standard conditions, is a well-studied electrocatalyst for the oxygen evolution reaction (OER).[25–32] Previous studies on a $SrIrO_3$ OER electrocatalyst demonstrated that Sr leaching might leave behind a highly oxidized Ir^{6+} species which was argued to be responsible for the observed OER activity.[25] Other groups also observed reconstruction of IrO_x catalysts under reaction conditions and subsequent formation of an unknown structure.[33] Highly oxidized IrO_3 phases are also formed as the terminal structure of Li_xIrO_3 anodes.[33] For these reasons, we focused our search on stable polymorphs in the IrO_2 and oxidized IrO_3 stoichiometries.

Here, we first detail the generation of our candidate structures for IrO_2 and IrO_3 and introduce the AL accelerated surrogate model. Next, we demonstrate the application of our AL scheme to the IrO_2 and IrO_3 prototype spaces and evaluate the algorithm's performance towards acquiring of the most stable polymorphs. We analyze the crystallographic motifs of the DFT-relaxed structures and identify structural trends within the most stable polymorphs. Lastly, we incorporate discovered structures into bulk and surface Pourbaix diagrams, and evaluate their catalytic OER performance.

Results and discussion

I. Candidate Space Generation and Active Learning Methodology

Our approach utilizes an active learning framework and surrogate models, whereby a regression model is trained to compute enthalpies of formation (ΔH_f) by iteratively sampling structures from a set of polymorph candidates. Figure 1 shows a schematic overview of the AL loop. We first generate the structure candidate space, followed by an iterative search through the space via a continuously retrained surrogate model using Gaussian Processes Regression (GPR), which is then used to acquire subsequent structures for DFT optimization. No prior DFT training data is required to start the algorithm, eliminating any initial

built-in bias in the model and allowing it to quickly respond to new acquisitions.

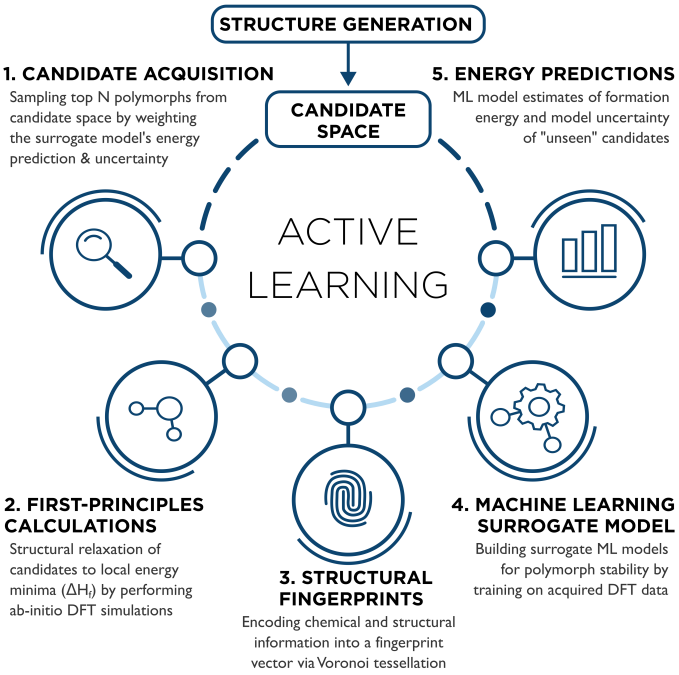

Figure 1: AL accelerated polymorph discovery algorithm diagram. Following the generation of the hypothetical crystal structure data set (candidate space), the AL algorithm proceeds iteratively through: (1) candidate selection in which a subset of structures in the candidate space are selected based on an acquisition function (in lieu of training data, the initial candidates are randomly sampled), (2) structural relaxation into local energy minima (ΔH_f computed), (3) structure featurization to produce numerical vector for input into ML model, (4) Machine learning (ML) model training based on acquired structures and ΔH_f, (5) Prediction of candidate space's ΔH_f distribution via ML model. The algorithm repeats steps (1)-(5) until a suitable stop criteria is reached.

The candidate structure datasets for IrO_2 and IrO_3 were constructed by first obtaining all AB_2 and AB_3 structures in the Materials Project[4] and OQMD[7] databases (in total 7,160 AB_2 and 31,224 AB_3 entries). To reduce the size of the candidate space while maintaining maximum structural diversity, structurally redundant systems were removed via a space group based structural classification scheme developed by Jain et al.[34] In short, a material's structural identity is defined by a unique combination of the element-nonspecific stoichiometry (AB_2, AB_3, etc.), space group symmetry, and Wyckoff positions, collectively referred to as a materials structural prototype. Materials of the same prototype are considered to be structurally equivalent. Eliminating these redundant materials results in orders of mag-

nitude reduction of the search space to 697 and 259 unique prototypes for AB_2 and AB_3, respectively. Finally, only structures containing less than 75 atoms (566 AB_2 and 256 AB_3) were included to reduce the computational expense of subsequent DFT calculations. We next substituted iridium and oxygen for the A and B sites, and these Ir-O adapted polymorphs were isotropically relaxed to accommodate their atomic radii. Bulk DFT optimizations were performed on these systems, yielding 714 relaxed bulk IrO_x polymorphs (466 and 248 structures for IrO_2 and IrO_3, respectively), after discarding 108 non-converged structures. The relatively small size of our candidate space allows us to tractably optimize all structures and allows us to readily benchmark the performance of our algorithm. Full details of the candidate space generation and DFT calculations can be found in the Supporting Information. All structurally unique IrO_x optimized structures (575 in total) can be accessed through the MPcontribs platform.[35]

The active learning algorithm proceeds through a structure featurization scheme based on Voronoi tessellation developed by Ward et al.[36] which produces a 271-length fingerprint vector that is invariant to isotropic lattice changes and insensitive to the precise atomic coordinates. These fingerprints encode both chemical and structural information by constructing attributes from elemental properties which are weighted by the local environment of the structure via the construction of the Wigner-Seitz cell.[37] Since our AL framework focuses on fixed compositions, the dimensionality is reduced to 101 non-zero variance features. We further reduce the dimensionality to 10 features via principal component analysis (PCA),[38] which we found to capture 80% of the variance in the full feature set while also demonstrating an optimal cross-validation mean absolute error (MAE) (see Figure S1).

The active learning algorithm proceeds through iterative generations of ML training, prediction, and acquisition steps that are visualized in Figure 1. To meet our primary goal of identifying the most stable polymorphs within the candidate space, we construct the AL framework to be (1) responsive in improving itself by learning from small batches of newly acquired DFT data, and (2) aware of limitations in its surrogate model by incorporating un-

certainty estimates into the acquisition decision criteria. GPR satisfies both requirements, and we use them here with a Gaussian kernel as implemented in CatLearn.[20,39] In the initial generation (generation 0), the model is trained on a set of randomly sampled candidates (unbiased sampling), and is then used to predict the formation enthalpy (ΔH_f) of all structures in the candidate space. The predicted energy landscape is then used to choose the next polymorphs to acquire (calculate via DFT) by selecting systems that minimize the GP-LCB (Gaussian process lower confidence bound) acquisition function, $U = \mu - \kappa\sigma$.[40] Here, μ and σ are the predicted ΔH_f mean and uncertainty, respectively, and κ is a parameter that weights exploitation vs. exploration of the search-space (set to 1). At every generation of the AL loop, N structures that minimize the acquisition function are acquired for DFT optimization and are subsequently added to the training data set, where N is the AL bin size (here set to 5). The value of N determines the degree of parallelization of the routine. In practice the algorithm can proceed until no more stable polymorphs are found, or after an allocated computational budget is exhausted.

Although initially unique, the structures in the candidate set often relax into one another over the course of the DFT optimization, introducing duplicates in the post-DFT structures. The duplicates are removed during each generation of the AL algorithm by using the structure similarity quantification method of Su et al.[41]

II. Active Learning Algorithm Applied to the Discovery of Stable Iridium-oxide Polymorphs

We next applied the AL algorithm to the discovery of stable and unique polymorphs of IrO_2 and IrO_3, individually. Results for IrO_2 are provided in the supporting information (Figure S2), here we focus on IrO_3, since it is a comparatively unexplored oxide system.

Figure 2a, shows a sequence of snapshots of the AL algorithm applied to IrO_3 at different generations. Each subplot reports the predicted (grey) and DFT-derived (filled red) formation enthalpies (ΔH_f) for each structure, sorted by stability such that structures more

likely to be selected by the acquisition criteria are farther left. As the algorithm acquires DFT data, the GP model's accuracy increases, as evidenced by the decreasing uncertainties when comparing the initial and latter generations (2a.i-v). At the top of each subplot of Figure 2a the x-axis positioning of the ten most stable polymorphs is tracked. Initially, these ten structures are randomly distributed across the entire candidate space due to the lack of training data for the GP model. However, after only three generations (Figure 2a.ii) the GP model is sufficiently accurate to predict the most stable polymorphs as low energy structures. By the sixth generation (40 DFT relaxations) 4/10 of the most stable polymorphs have been acquired, including the globally stable phase of IrO_3, which was found on average in only 4.3 generations (averaged over 100 independent runs). By the 13th generation of the algorithm 10/10 of the most stable structures were acquired.

Seven of the most stable IrO_3 polymorphs discovered are shown in Figure 2b. All of the low energy IrO_3 structures are constructed from octahedrally coordinated units, with a variety of symmetries and packing modes. The globally stable crystal structure consists of a six-atom primitive cell with a space group number of 167 ($R\overline{3}c$) in the rhombohedral crystal system, has exclusively corner-sharing octahedra, and is isomorphic to FeF_3.[42] Herein, this structure will be referred to as α-IrO_3. The second most stable polymorph (Figure 2b.ii) is similar to α-IrO_3, only differing by the stacking of the alternating layers orthogonal to the **c** lattice vector. We label this structure α_2-IrO_3 in Figure 2b and it is only 2 meV/atom less stable than α-IrO_3, well within the margin of error for DFT. The fourth most stable structure (R-IrO_3) is notable for being the first in the series to have mixed edge- and corner-sharing octahedra and is structurally similar to rutile-IrO_2 (R-IrO_2).

Figure 2c reports the discovery rate of the AL algorithm by plotting the number of the ten most stable systems acquired against the number of DFT calculations with the GP-LCB acquisition and a random acquisition scheme to serve as a baseline. The results of Figure 2c are averaged over 100 independent runs of the AL algorithms with the standard deviation shown. Overall, the GP-LCB runs outperform the random acquisition runs, with on average

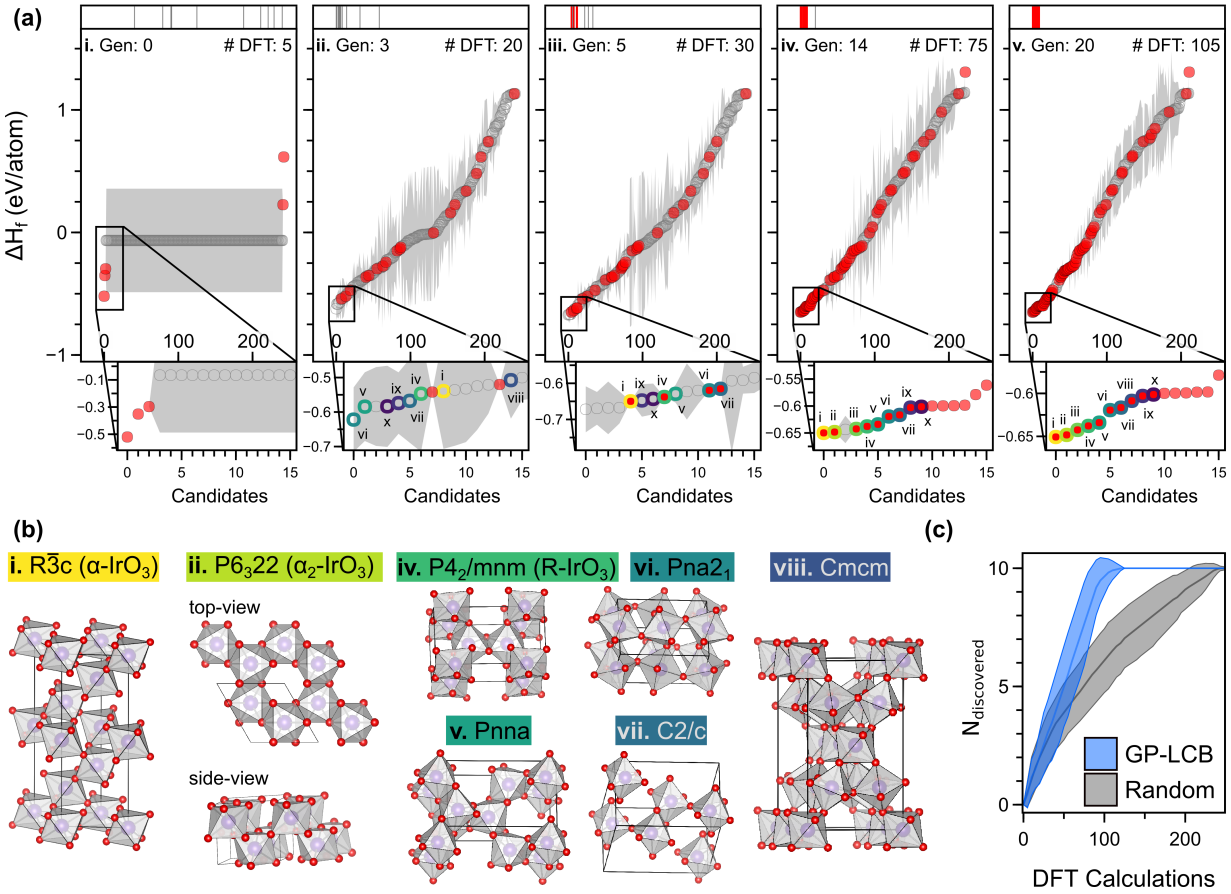

Figure 2: (a) The state of the AL algorithm at five generations. The enthalpy of formation per atom (ΔH_f) is plotted, ordered by stability, against all IrO_3 candidates, with the 1 σ uncertainty estimate shown for each prediction. The number of DFT training points at each generation is displayed. Hollow grey markers indicate a GP model predicted ΔH_f while red indicate a DFT-computed quantity. At the top of subplots a.i-v, the x-axis positions of the ten most stable polymorphs are tracked at each generation by either red (acquired) or grey (not acquired) vertical lines. Insets of the low energy region for each generation is displayed below each subplot. The top ten most stable systems are colored and labeled (i-vii) to indicate their identity. (b) Crystal structures of the 8 most stable IrO_3 polymorphs (structure iii not shown). (c) The number of most stable 10 polymorphs of IrO_3 discovered ($N_{discovered}$) vs. the number of DFT calculations for the GP-LCB (blue) and random (grey) acquisition methods. The results are averaged over 100 independent runs and the 1 σ standard deviation between these runs are displayed. All structurally unique IrO_x DFT optimized structures can be accessed through the MPcontribs platform.[35]

~100 DFT calculations needed to discover the ten most stable structures. This demonstrates over a factor of two improvement in performance compared to random acquisition, which does not acquire the most stable structures until all ~250 candidates are computed. The results for IrO_2 (Figure S2) show a higher discovery rate for GP-LCB compared to the random

acquisition method, although the GP-LCB method "saturated" at 9/10 and was unable to acquire the last structure until the candidate space was exhausted. The performance of GP-LCB relative to random is expected to increase with the size of the candidate space, since the probability of selecting stable structures is inversely proportional to the size of the candidate pool.

We next evaluate the prediction accuracy of the IrO_2 and IrO_3 GP regression models utilizing the full DFT optimized data set of 487 IrO_2 and 249 IrO_3 structures. This dataset corresponds to the final generation of the AL algorithm in which all structures have been acquired. Figure 3 plots the GP model predicted ΔH_f against the DFT-computed values for two cases. Case (1) shows the predictions on the structural fingerprints prior to DFT-optimization (grey), as is done in the regular operation of the algorithm. Case (2) shows the prediction of the same GP model using the post-DFT optimized fingerprints (blue) with 10-fold cross-validation. It is evident that using the pre-optimization fingerprints results in the GP model being highly inaccurate in predicting the post-relaxation ΔH_f's of the candidate space, with a seemingly large MAE of ~1.5 eV/atom. In contrast, the same GP model does comparatively much better at predicting the formation energies of post-DFT optimized structures with an MAE of ~0.2 eV/atom.

The drastic decrease in prediction error is not surprising, since the post-DFT fingerprints directly correspond to the target ΔH_f values, and is primarily due to the large degree of structural drift that occurs during DFT relaxation, the extent of which is not known a priori. In fact, we observe that most of the predictions from pre-DFT features over-predict the formation energy (i.e. less stable than their DFT analogous) and lie above the parity line. This behavior is consistent with what one would expect thermodynamically: structures that are initialized in high energy configurations will naturally reconfigure into a more stable local configuration, resulting in discrepancies between the pre-DFT predicted and final formation energies. In practice, our approach still performs notably well because (1) the energy tends to decrease post-DFT relaxation, meaning favored acquisitions are likely to perform even

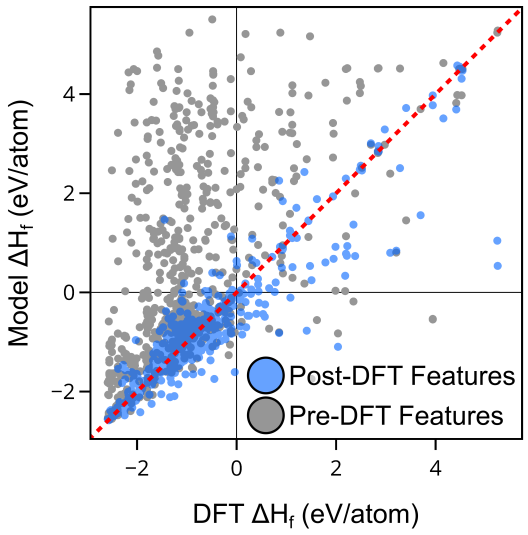

Figure 3: Parity plot of the predicted and actual formation enthalpy (ΔH_f) for the final Gaussian regression ML models for IrO_2 and IrO_3, predicting on either the pre-optimized (grey) or the post-optimized (blue) fingerprints. The final ML regression models are trained on the DFT ΔH_f values for all structurally unique polymorphs and correspond to the final generation of the AL algorithm.

better, and (2) the pre-optimized structures that are similar enough to the most stable final equilibrium structures will not restructure considerably, meaning that their predicted formation energies will be close enough (and low enough) to be quickly picked up by the acquisition criteria. Additionally, the number of duplicates produced during AL is also a factor in increasing the effective performance. For example, there are 8 duplicates of α-IrO_3 produced during the full AL routine due to distinct pre-DFT candidates relaxing into the same energy basin, and this over representation of α-IrO_3 phase effectively increases the chance of it being acquired by a factor of eight.

III. Crystal Coordination Analysis of Discovered Phases

We next assess stability trends and structural variety of the DFT optimized structures, consisting of 384 and 191 unique IrO_2 and IrO_3 polymorphs. Figure 4a,b shows the DFT computed ΔH_f for IrO_2 and IrO_3 plotted against the inverse density, a quantity that is sensitive to crystal porosity and connectivity. To obtain a physically meaningful cutoff

for ΔH_f, we computed the "amorphous limit" of Persson and coworkers for both IrO_2 and IrO_3, which provides a stringent upper bound to polymorph synthesizability.[43] We found the amorphous limit for both IrO_2 and IrO_3 to occur at a ΔH_f of $-0.34\,eV/atom$ and are displayed as horizontal lines in Figure 4a,b. There are 196 and 75 polymorphs for IrO_2 and IrO_3, respectively, that are within the amorphous synthesizability limit.

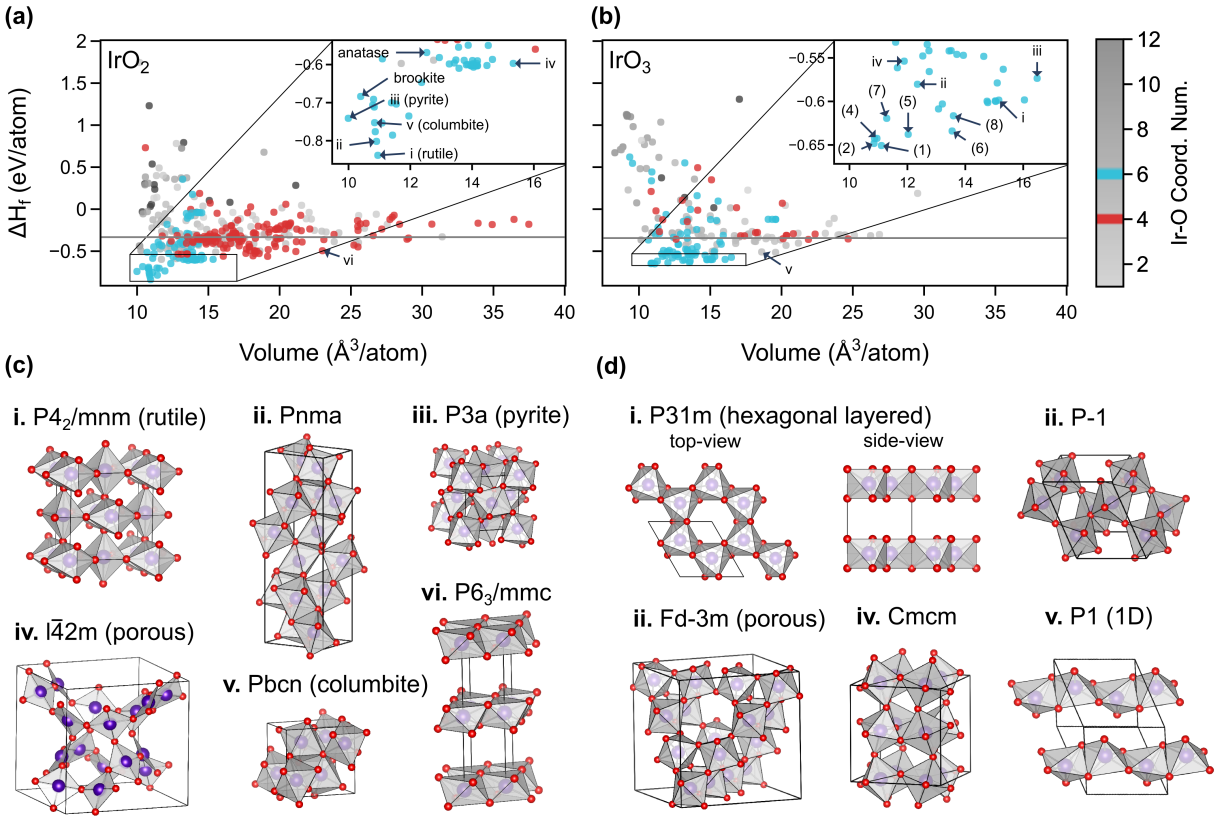

Figure 4: ΔH_f for the 384 IrO_2 (a) and 191 IrO_3 (b) structurally unique DFT optimized structures in the candidate data set plotted against the volume per atom. Insets in the low energy region for (a) and (b) are shown. The color bar represents the average coordination number between Ir and O, with the most common, 6 (octahedra) and 4 (tetrahedral) coordinations, highlighted. For IrO_2 (c) and IrO_3 (d) we highlight the structures of select polymorphs. The amorphous limits for IrO_2 and IrO_3 (Figure S2) defining a strict upper bound for synthesizability are displayed in (a) and (b) as horizontal lines.

Computed materials span a large range of densities and coordination environments. The lowest volume (highest density) structures correspond to an atomic packing factor of roughly 0.50, which is where the most stable structures are found. The highest volume (lowest density) systems sampled have atomic packing factors close to 0.15. However, for IrO_3

there is a comparatively weaker relationship between the energy and volume, such that even highly porous structures are within 0.1 eV/atom of the most stable phase. This is indicative of IrO_3's high degree of polymorphism and ability to readily form layered and/or porous structures.

Ir-O coordination environments were classified (octahedral, square pyramidal, tetrahedral, cubic, etc.) by using the chemEnv package, developed by Waroquiers et. al.[44] as implemented in the Pymatgen software.[45] Our dataset contains structures with coordination numbers ranging from 2 to 10, with coordination numbers of six (octahedral, blue) and four (tetrahedral, red) being the most prevalent (see Figure 4). The vast majority of the most stable (within 0.1 eV/atom) structures adopt an octahedral coordination environment, a common coordination motif found to be favorable in many other transition metal oxides.[44] The arrangement of the octahedral units, which are connected through either corner- or edge-sharing octahedra, can furthermore be used to classify the structures, which typically have a combination of the two. Of the top ten IrO_2 and IrO_3 structures, 9/10 of IrO_2 and 5/10 of IrO_3 have a mixed corner- and edge-sharing octahedral packing. This demonstrates that IrO_2 prefers to form edge-sharing octahedra as a result of having to share more oxygens to maintain its stoichiometry. IrO_3 has comparatively more oxygens per unit cell, and as such can adopt completely corner-sharing arrangements similar to cubic perovskite-type structures.

Figure 4c,d shows a selection of metastable structures for IrO_2 and IrO_3, respectively. For IrO_2, we reaffirm the rutile ground state. Additionally, the experimentally synthesized high-pressure pyrite phase of IrO_2 was found in our dataset and has a ΔH_f ~0.1 eV/atom greater than rutile, in agreement with theoretical and experimental calorimetric data.[46,47] Several common AB_2 crystal structures were found within the dataset, including brookite,[48] anatase[49] and the recently synthesized columbite-IrO_2 phase.[50] For IrO_3 the eight most stable systems are reported in Figure 2, and labeled as (1)-(8) in Figure 4b. In addition to the most thermodynamically stable systems, we have identified several interesting metastable

structures, including two dimensional (i), highly porous (iii) and one dimensional (v) polymorphs with varying degrees of porosity and connectivity, which are important structural properties for applications as battery cathodes and ionic conductors.[33,51]

IV. Electrochemical OER Application

We next performed ab-initio thermodynamic analyses to test the OER electrochemical operational stability and activity of the most stable IrO_x in aqueous solution. In particular, we compare the stability and activity of R-IrO_2 to our newly discovered α-IrO_3 and R-IrO_3 polymorphs. In addition, we have also computed the stability and activity of a delithiated form of a recently reported β-Li_xIrO_3 structure, referred here to as β-IrO_3.[33,51] The OER activity was computed assuming the most common single-site associative OER mechanism utilizing the thermodynamic limiting potential analysis with the computational hydrogen electrode as described extensively in numerous previous works[52–55] (see also Supporting Information for details).

The calculated bulk Pourbaix diagram of the Ir-H_2O system is shown in Figure 5. The diagram was constructed by considering the thermodynamic equilibrium between the following species: Ir, R-IrO_2, α-IrO_3, R-IrO_3, β-IrO_3, and an aqueous dissolved IrO_4^- species. To obtain free energies, we utilized a free energy correction to our calculated values to reproduce the known experimental ΔH_f and ΔG_f of R-IrO_2.[56] While Ir and R-IrO_2 are most stable at low bias, α-IrO_3 becomes the thermodynamically dominant phase under the relevant conditions for the OER (potentials $> 1.23\ V_{\mathrm{RHE}}$ and in an acidic environment). The stability regions for the less stable β-IrO_3 and R-IrO_3 polymorphs (in the absence of other IrO_3 phases) are also included (unfilled solid lines). It can be seen that these phases have a reduced, but sufficiently large, stability window relative to IrO_2 and the IrO_4^-. Removal of the IrO_3 phases from the bulk Pourbaix diagram results in a completely different thermodynamic picture of IrO_2 stability (Figure S4). In total, we have discovered 21 unique IrO_3 polymorphs with a non-zero bulk Pourbaix stability region ($0 \leq \mathrm{pH} \leq 14$). Interest-

ingly, these 21 structures are more stable than the most stable IrO_3 structure in Materials Project[57] (which is also present in our dataset, Figure 4d.iv). We note that thermodynamic driving force toward these IrO_3 phases under OER conditions suggests that these structures may spontaneously form under reaction conditions, especially for systems that undergo a large degree of surface reconstruction.[25]

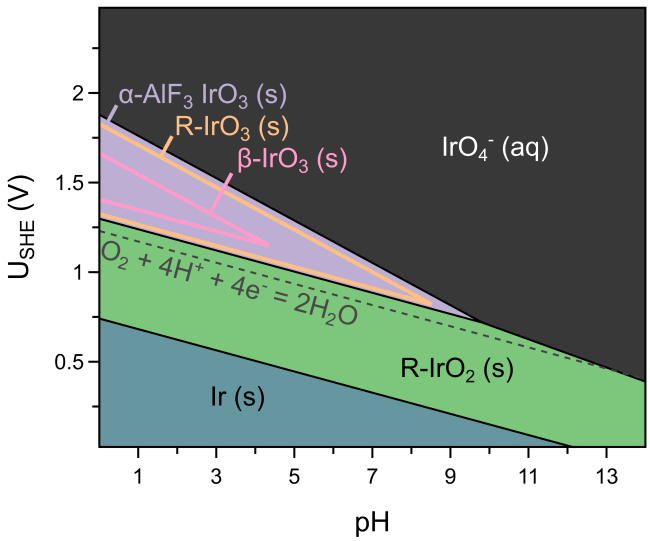

Figure 5: Revised bulk Pourbaix diagram of the Ir-H_2O system as a function of applied potential (U_{SHE}) and pH. The diagram was constructed with Ir(s) (blue), R-IrO_2 (green), various IrO_3 polymorphs and a dissolved IrO_4^- ion species (dark grey). The stability regions corresponding to the metastable R-IrO_3 and β-IrO_3 polymorphs (see text), in the absence of any competing IrO_3 phase, are displayed as yellow and pink lines, respectively. The thermodynamic onset of OER (water equilibrium at 1.23 V_{RHE}) is also shown. To be compared to Figure S4 without IrO_3 phases. See Table S3 for the bulk formation energies (ΔG_f) used to construct the diagram.

We next computed the surface energy Pourbaix plots and OER activity for various surface facets at select coverages (for simplicity we only choose bare and one monolayer of OH* and O*) of all four systems from Figure 5; results are summarized in Figure 6. For each polymorph, surfaces were constructed by cleaving along the Miller indices with the highest calculated diffraction peaks, corresponding to planes with higher density of atoms. The surface free energy Pourbaix plots identify which surface facets and surface coverage species are thermodynamically preferred under OER conditions. Our results show that most of the

facets prefer to have a high surface coverage of O*, therefore we consider mainly oxygen terminated surfaces for the OER analysis. Our results are comparable to previous studies on the electrochemical stability of IrO_2 surfaces,[58,59] but without considering highly reconstructed facets such as (101). The surface stability analysis is therefore crucial for accurate determination of the activity.

The calculated OER activities of relevant OER stable surfaces are plotted against the ΔG_O-ΔG_{OH} descriptor and are shown in Figure 6b. There, we display two thermodynamic limiting potential volcanos based on (1) the standard universal[52] (black) and (2) fitted (grey) scaling relations between the OER intermediates (Figure S5). Additionally, we have also added a kinetic OER volcano (dashed line) from Dickens et al.[63] based on the detailed microkinetic model developed for rutile systems. The kinetic volcano is constructed at the potential required to reach $10\,\mathrm{mA\,cm^{-2}}$. The thermodynamic and kinetic volcanos agree remarkably well in the strong binding portion (left hand side) of the plot and exhibit similar optimum value, ΔG_O-$\Delta G_{OH} \approx$ 1.55-1.65 eV. The corresponding surface structures for selected systems featuring high oxygen coverage are visualized in Figure 6c.

In general, the R-IrO_2 surfaces bind the OER intermediates relatively strongly, with theoretical limiting potentials of ~1.8 V_{RHE} (overpotential of 0.57 V_{RHE}) having the *O to *OOH potential limiting step, in agreement with previous theoretical studies.[59,64,65] The predicted overpotentials of our R-IrO_2 systems are also within the range of experimentally observed values (horizontal lines).[25,60] The surfaces of the three IrO_3 polymorphs have ΔG_O-ΔG_{OH} values shifted to higher energies, indicative of overall weaker binding energetics (see also Figure S5). On average, the adsorption of OH* and O* is weakened by 0.7 and 1.2 eV relative to IrO_2 (Table S4), respectively. The highest performing systems include the α-IrO_3 (100), (110), and (211), followed by β-IrO_3 (101), and then R-IrO_3 (110). These surfaces have overpotentials of ~0.4 V_{RHE}, which represents a ~0.2 V_{RHE} improvement over R-IrO_2, mirroring the observed shift in experimental onset potentials (horizontal lines).[25,60] The primary driver for the improved OER activity is the higher oxidation state of IrO_3 compared

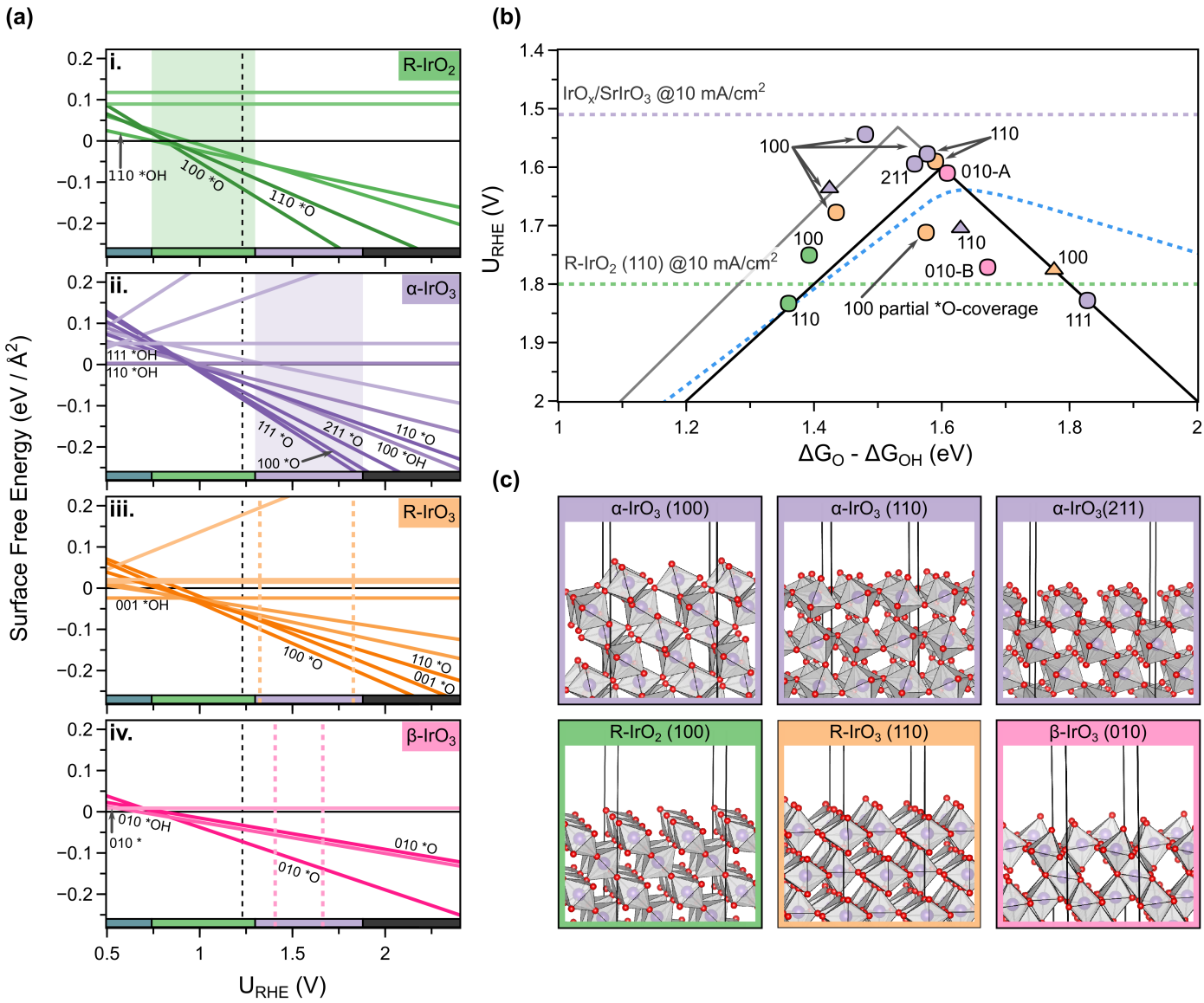

Figure 6: Summary of OER results for the following four bulk structures of IrO_x: R-IrO_2 (green), α-IrO_3 (purple), R-IrO_3 (orange), and β-IrO_3 (pink). (a) Surface energy Pourbaix diagrams for each structure, with the surface energy of various facets and coverages shown as a function of applied potential (V_{RHE}). The bulk Pourbaix diagram's bounds of stability at pH 0 are superimposed as horizontal bars at the bottom of each subplot. The pseudo-stability regimes for the metastable β-IrO_3 and R-IrO_3 are indicated by dashed vertical lines. (b) OER activity volcano for IrO_x systems considered utilizing the ΔG_O-ΔG_{OH} descriptor. The horizontal lines correspond to recent experimental OER limiting potentials for R-IrO_2 (110)[60] and $SrIrO_3$,[25] at 10 mA cm^{-2} (extrapolated values). (c) Corresponding structural models for selected OER surfaces at one mono-layer O* coverage used for calculation of the overpotentials. Color legend: oxygen (red), purple (iridium), coordination motif (white). Computational cell is displayed by black lines. All OER slab models and corresponding DFT energies are freely available under the "FloresActive2020"[61] dataset at Catalysis-hub.org.[62]

to IrO_2, having only three $5d$-electrons for Ir^{6+} as opposed to five $5d$-electrons in Ir^{4+}, respectively. Oxygen saturated IrO_3 systems thus bind OER intermediates more weakly, which leads to positive shift in ΔG_O-ΔG_{OH}. IrO_2 and RhO_2 are generally overbinding for OER[63] so there is consequent improvement in OER activity when compared to these oxides. These results are consistent with Back et al., who recently computed elevated activity in highly oxidized IrO_3 catalysts.[66] An added feature of α-IrO_3 is comparably higher density of active sites due to completely corner-sharing geometry. The exact improvement in the theoretical overpotential is slightly dependent on the DFT level of theory and the inclusion of spin polarization, and has been discussed recently.[64,65]

Conclusion

We have described a cogent procedure for generating and searching a structurally diverse candidate space of bulk structural prototypes with a desired composition. Once this space is enumerated, we show how it can be efficiently searched using an algorithm with an active learning loop without a prior knowledge of accurate atomic positions. In most cases, the DFT optimization of only a fraction of the candidates leads to identification of the most stable polymorphs. In particular, this approach is well-suited for discovery in structurally diverse structures, such as metal oxides and other metal-ligand bulk systems, where there exits a large degree of structural diversity. The current dataset includes octahedral, tetrahedral, square-pyramidal, cubic, and square-planar Ir-O conformers. We also note, that our AL algorithm is capable of discovering experimentally known phases such as pyrite, columbite and layered IrO_2 and several recently discovered layered IrO_3 phases formed by Li^+ deintercalation. In particular, we have identified a number of previously unknown IrO_3 polymorphs below the amorphous synthesizability limit, including a new globally stable α-IrO_3 phase. This high valency Ir^{6+} phase is stable under OER relevant conditions and has an ideal 100% corner-sharing octahedral structure, a short Ir-O bond length of 1.93 Å, and also has a very

high surface coverage of active oxygens. Calculations of surface thermodynamics reveal this structure and other OER stable IrO_3 phases have much higher theoretical OER activity than a benchmark rutile IrO_2. The thermodynamic stability and high OER activity of the α-IrO_3 phase may provide clues as to the nature of the yet uncharacterized structures reported after reconstruction of $SrIrO_3$ and IrO_x precursors under OER reaction conditions. Methods combining diverse structural generation, AL-enabled accelerated searches, and ab-initio simulation of material performance could open up new avenues for in silico material design with application tailored structural properties.

Acknowledgement

This work was supported by the Toyota Research Institute through the Accelerated Materials Design and Discovery program. JAGT and MB would also like to acknowledge the support by the U.S. Department of Energy, Office of Science, Office of Basic Energy Science, via Grant DE-SC0008685 to the SUNCAT Center of Interface Science and Catalysis. The authors would like to acknowledge the use of the computer time allocation for the Transition metal-oxide and metal surfaces: applications and reactivity trends in catalysis at the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Additionally, the authors would like to thank Stanford University and the Stanford Research Computing Center (Sherlock cluster) for providing computational resources and support that contributed to these research results. The authors would also like to acknowledge Patrick Huck for enabling the inclusion of our IrO_x crystal data set into the Materials Project Contributions and Materials Project databases. Lastly, we thank Trevor Mowry for assistance with the Table of Contents (TOC) figure.

References

(1) Woodley, S. M.; Catlow, R. Crystal structure prediction from first principles. 2008.

(2) Graser, J.; Kauwe, S. K.; Sparks, T. D. Machine Learning and Energy Minimization Approaches for Crystal Structure Predictions: A Review and New Horizons. 2018; https://pubs.acs.org/doi/10.1021/acs.chemmater.7b05304.

(3) Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD). *JOM* **2013**, *65*, 1501–1509.

(4) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Materials* **2013**, *1*.

(5) Curtarolo, S.; Setyawan, W.; Hart, G. L.; Jahnatek, M.; Chepulskii, R. V.; Taylor, R. H.; Wang, S.; Xue, J.; Yang, K.; Levy, O.; Mehl, M. J.; Stokes, H. T.; Demchenko, D. O.; Morgan, D. AFLOW: An automatic framework for high-throughput materials discovery. *Computational Materials Science* **2012**, *58*, 218–226.

(6) Mamun, O.; Winther, K. T.; Boes, J. R.; Bligaard, T. High-throughput calculations of catalytic properties of bimetallic alloy surfaces. *Scientific data* **2019**, *6*, 1–9.

(7) Kirklin, S.; Saal, J. E.; Meredig, B.; Thompson, A.; Doak, J. W.; Aykol, M.; Rühl, S.; Wolverton, C. The Open Quantum Materials Database (OQMD): Assessing the accuracy of DFT formation energies. *npj Computational Materials* **2015**, *1*.

(8) Oganov, A. R.; Glass, C. W. Crystal structure prediction using ab initio evolutionary techniques: Principles and applications. *Journal of Chemical Physics* **2006**, *124*, 244704.

(9) Lyakhov, A. O.; Oganov, A. R.; Valle, M. How to predict very large and complex crystal structures. *Computer Physics Communications* **2010**, *181*, 1623–1632.

(10) Glass, C. W.; Oganov, A. R.; Hansen, N. USPEX-Evolutionary crystal structure prediction. *Computer Physics Communications* **2006**, *175*, 713–720.

(11) Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. CALYPSO: A method for crystal structure prediction. *Computer Physics Communications* **2012**, *183*, 2063–2070.

(12) Pickard, C. J.; Needs, R. J. Ab initio random structure searching. 2011; `https://iopscience.iop.org/article/10.1088/0953-8984/23/5/053201/meta`.

(13) Pickard, C. J.; Needs, R. J. High-pressure phases of silane. *Physical Review Letters* **2006**, *97*, 045504.

(14) Avery, P.; Toher, C.; Curtarolo, S.; Zurek, E. XTALOPT Version r12: An open-source evolutionary algorithm for crystal structure prediction. *Computer Physics Communications* **2019**, *237*, 274–275.

(15) Lonie, D. C.; Zurek, E. XtalOpt: An open-source evolutionary algorithm for crystal structure prediction. *Computer Physics Communications* **2011**, *182*, 372–387.

(16) Stillinger, F. H. Exponential multiplicity of inherent structures. *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* **1999**, *59*, 48–51.

(17) Van Den Bossche, M.; Grönbeck, H.; Hammer, B. Tight-Binding Approximation-Enhanced Global Optimization. *Journal of Chemical Theory and Computation* **2018**, *14*, 2797–2807.

(18) Montoya, J. H.; Winther, K.; Flores, R. A.; Bligaard, T.; Hummelshj, J. S.; Aykol, M. Autonomous Intelligent Agents for Accelerated Materials Discovery. **2020**,

(19) Noh, J.; Kim, J.; Stein, H. S.; Sanchez-Lengeling, B.; Gregoire, J. M.; Aspuru-Guzik, A.; Jung, Y. Inverse Design of Solid-State Materials via a Continuous Representation. *Matter* **2019**, *1*, 1370–1384.

(20) Hansen, M. H.; Torres, J. A. G.; Jennings, P. C.; Wang, Z.; Boes, J. R.; Mamun, O. G.; Bligaard, T. An Atomistic Machine Learning Package for Surface Science and Catalysis. *arXiv preprint arXiv:1904.00904* **2019**,

(21) Torres, J. A. G.; 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. *Physical review letters* **2019**, *122*, 156001.

(22) Jennings, P. C.; Lysgaard, S.; Hummelshøj, J. S.; Vegge, T.; Bligaard, T. Genetic algorithms for computational materials discovery accelerated by machine learning. *npj Computational Materials* **2019**, *5*, 46.

(23) Podryabinkin, E. V.; Shapeev, A. V. Active learning of linearly parametrized interatomic potentials. *Computational Materials Science* **2017**, *140*, 171–180.

(24) Bassman, L.; Rajak, P.; Kalia, R. K.; Nakano, A.; Sha, F.; Sun, J.; Singh, D. J.; Aykol, M.; Huck, P.; Persson, K.; Vashishta, P. Active learning for accelerated design of layered materials. *npj Computational Materials* **2018**, *4*, 1–9.

(25) Seitz, L. C.; Dickens, C. F.; Nishio, K.; Hikita, Y.; Montoya, J.; Doyle, A.; Kirk, C.; Vojvodic, A.; Hwang, H. Y.; Norskov, J. K.; Jaramillo, T. F. A highly active and stable IrOx/SrIrO3 catalyst for the oxygen evolution reaction. *Science* **2016**, *353*, 1011–1014.

(26) 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. *Journal of Physical Chemistry Letters* **2012**, *3*, 399–404.

(27) McCrory, C. C. L.; Jung, S.; Ferrer, I. M.; Chatman, S. M.; Peters, J. C.; Jaramillo, T. F. Benchmarking Hydrogen Evolving Reaction and Oxygen Evolving Reaction Electrocatalysts for Solar Water Splitting Devices. *Journal of the American Chemical Society* **2015**, *137*, 4347–4357.

(28) Trotochaud, L.; Ranney, J. K.; Williams, K. N.; Boettcher, S. W. Solution-cast metal oxide thin film electrocatalysts for oxygen evolution. *Journal of the American Chemical Society* **2012**, *134*, 17253–17261.

(29) Danilovic, N.; Subbaraman, R.; Chang, K. C.; Chang, S. H.; Kang, Y. J.; Snyder, J.; Paulikas, A. P.; Strmcnik, D.; Kim, Y. T.; Myers, D.; Stamenkovic, V. R.; Markovic, N. M. Activity-stability trends for the oxygen evolution reaction on monometallic oxides in acidic environments. *Journal of Physical Chemistry Letters* **2014**, *5*, 2474–2478.

(30) Carmo, M.; Fritz, D. L.; Mergel, J.; Stolten, D. A comprehensive review on PEM water electrolysis. *International Journal of Hydrogen Energy* **2013**, *38*, 4901–4934.

(31) Miles, M.; Klaus, E.; Gunn, B.; Locker, J.; Serafin, W.; Srinivasan, S. The oxygen evolution reaction on platinum, iridium, ruthenium and their alloys at 80C in acid solutions. *Electrochimica Acta* **1978**, *23*, 521–526.

(32) Beni, G.; Schiavone, L. M.; Shay, J. L.; Dautremont-Smith, W. C.; Schneider, B. S. Electrocatalytic oxygen evolution on reactively sputtered electrochromic iridium oxide films [7]. **1979**, *282*, 281–283.

(33) Pearce, P. E.; Perez, A. J.; Rousse, G.; Saubanère, M.; Batuk, D.; Foix, D.; McCalla, E.; Abakumov, A. M.; Van Tendeloo, G.; Doublet, M. L.; Tarascon, J. M. Evidence for anionic redox activity in a tridimensional-ordered Li-rich positive electrode β-Li2IrO3. *Nature Materials* **2017**, *16*, 580–586.

(34) Jain, A. PHYSICAL REVIEW B 98 , 214112 (2018) Atomic-position independent descriptor for machine learning of material properties. **2018**, *214112*, 1–7.

(35) MPContribs Active Learned IrOx Polymorphs. `https://portal.mpcontribs.org/active_learned_irox_polymorphs/`, 2020.

(36) Ward, L.; Liu, R.; Krishna, A.; Hegde, V. I.; Agrawal, A.; Choudhary, A.; Wolverton, C. Including crystal structure attributes in machine learning models of formation energies via Voronoi tessellations. *Physical Review B* **2017**,

(37) Wigner, E.; Seitz, F. On the constitution of metallic sodium. *Physical Review* **1933**, *43*, 804–810.

(38) Tipping, M. E.; Bishop, C. M. Probabilistic principal component analysis. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **1999**,

(39) CatLearn: An environment for atomistic machine learning in Python for applications in catalysis. 2019; `https://github.com/SUNCAT-Center/CatLearn`.

(40) Cox, D. D.; John, S. A statistical method for global optimization. Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics. 1992; pp 1241–1246.

(41) Su, C.; Lv, J.; Li, Q.; Wang, H.; Zhang, L.; Wang, Y.; Ma, Y. Construction of crystal structure prototype database: Methods and applications. *Journal of Physics Condensed Matter* **2017**, *29*.

(42) Hepworth, M. A.; Jack, K. H.; Peacock, R. D.; Westland, G. J. The crystal structures of the trifluorides of iron, cobalt, ruthenium, rhodium, palladium and iridium. *Acta Crystallographica* **1957**, *10*, 63–69.

(43) Aykol, M.; Dwaraknath, S. S.; Sun, W.; Persson, K. A. Thermodynamic limit for synthesis of metastable inorganic materials. *Science Advances* **2018**, *4*.

(44) Waroquiers, D.; Gonze, X.; Rignanese, G. M.; Welker-Nieuwoudt, C.; Rosowski, F.; Göbel, M.; Schenk, S.; Degelmann, P.; André, R.; Glaum, R.; Hautier, G. Statistical analysis of coordination environments in Oxides. *Chemistry of Materials* **2017**, *29*, 8346–8360.

(45) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science* **2013**, *68*, 314–319.

(46) Bolzan, A. A.; Fong, C.; Kennedy, B. J.; Howard, C. J. Structural studies of rutile-type metal dioxides. *Acta Crystallographica Section B: Structural Science* **1997**, *53*, 373–380.

(47) others,, et al. Synthesis, crystal structure, and electronic properties of high-pressure PdF2-type oxides MO2 (M= Ru, Rh, Os, Ir, Pt). *Inorganic chemistry* **2014**, *53*, 11616–11625.

(48) Pauling, L.; Sturdivant, J. H. XV. The crystal structure of brookite. *Zeitschrift für Kristallographie-Crystalline Materials* **1928**, *68*, 239–256.

(49) Parker, B. L. I. Zur Kristallographie von Anatas und Rutil. *Zeitschrift für Kristallographie-Crystalline Materials* **1923**, *59*, 1–54.

(50) Lee, K.; Flores, R. A.; Lui, Y.; Wang, B. Y.; Bajdich, M.; Hikita, Y.; Sinclair, R.; Hwang, H. Y. Epitaxial Stabilization and Oxygen Evolution Reaction Activity of Metastable Columbite Iridium Oxide. *To be submitted* **2020**,

(51) Pearce, P. E.; Yang, C.; Iadecola, A.; Rodriguez-Carvajal, J.; Rousse, G.; Dedryvère, R.; Abakumov, A. M.; Giaume, D.; Deschamps, M.; Tarascon, J. M.; Grimaud, A. Revealing the Reactivity of the Iridium Trioxide Intermediate for the Oxygen Evolution Reaction in Acidic Media. *Chemistry of Materials* **2019**, *31*, 5845–5855.

(52) Man, I. C.; Su, H. Y.; Calle-Vallejo, F.; Hansen, H. A.; Martínez, J. I.; Inoglu, N. G.; Kitchin, J.; Jaramillo, T. F.; Nørskov, J. K.; Rossmeisl, J. Universality in Oxygen Evolution Electrocatalysis on Oxide Surfaces. *ChemCatChem* **2011**, *3*, 1159–1165.

(53) Rossmeisl, J.; Qu, Z. W.; Zhu, H.; Kroes, G. J.; Nørskov, J. K. Electrolysis of water on oxide surfaces. *Journal of Electroanalytical Chemistry* **2007**, *607*, 83–89.

(54) Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jónsson, H. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *Journal of Physical Chemistry B* **2004**, *108*, 17886–17892.

(55) Bajdich, M.; García-Mota, M.; Vojvodic, A.; Nørskov, J. K.; Bell, A. T. Theoretical investigation of the activity of cobalt oxides for the electrochemical oxidation of water. *Journal of the American Chemical Society* **2013**, *135*, 13521–13530.

(56) Barin, I. In *Thermochemical Data of Pure Substances*; Barin, I., Ed.; Wiley-VCH Verlag GmbH: Weinheim, Germany, 1995.

(57) IrO3 mp-1097041. `https://materialsproject.org/materials/mp-1097041/`, Accessed: 2020-01-20.

(58) Nattino, F.; Marzari, N. Operando XANES from first-principles and its application to iridium oxide. *Submitted to Phys. Chem. Chem. Phys.* **2019**,

(59) Raman, A. S.; Patel, R.; Vojvodic, A. Surface stability of perovskite oxides under OER operating conditions: A first principles approach. Publisher: The Royal Society of Chemistry.

(60) Kuo, D. Y.; Kawasaki, J. K.; Nelson, J. N.; Kloppenburg, J.; Hautier, G.; Shen, K. M.; Schlom, D. G.; Suntivich, J. Influence of Surface Adsorption on the Oxygen Evolution Reaction on IrO2(110). *Journal of the American Chemical Society* **2017**, *139*, 3473–3479.

(61) Catalysis Hub Surface Slab Dataset. https://www.catalysis-hub.org/publications/FloresActive2020.

(62) Winther, K. T.; Hoffmann, M. J.; Boes, J. R.; Mamun, O.; Bajdich, M.; Bligaard, T. Catalysis-Hub.org, an open electronic structure database for surface reactions. *Scientific data* **2019**, *6*, 75.

(63) Dickens, C. F.; Kirk, C.; Nørskov, J. K. Insights into the electrochemical oxygen evolution reaction with ab initio calculations and microkinetic modeling: Beyond the limiting potential volcano. *Journal of Physical Chemistry C* **2019**, *123*, 18960–18977.

(64) Briquet, L. G. V.; Sarwar, M.; Mugo, J.; Jones, G.; Calle-Vallejo, F. A New Type of Scaling Relations to Assess the Accuracy of Computational Predictions of Catalytic Activities Applied to the Oxygen Evolution Reaction. *ChemCatChem* **2017**, *9*, 1261–1268.

(65) Strickler, A. L.; Flores, R. A.; King, L. A.; Nørskov, J. K.; Bajdich, M.; Jaramillo, T. F. Systematic Investigation of Iridium-Based Bimetallic Thin Film Catalysts for the Oxygen Evolution Reaction in Acidic Media. *ACS Applied Materials and Interfaces* **2019**, *11*, 34059–34066.

(66) Back, S.; Tran, K.; Ulissi, Z. W. Toward a Design of Active Oxygen Evolution Catalysts: Insights from Automated Density Functional Theory Calculations and Machine Learning. *ACS Catalysis* **2019**, *9*, 7651–7659.

Graphical TOC Entry