Nanoscale

PAPER

Check for updates

Cite this: DOI: 10.1039/d4nr01390e

Received 29th March 2024, Accepted 30th July 2024 DOI: 10.1039/d4nr01390e

rsc.li/nanoscale

1 Introduction

As the global focus shifts increasingly towards renewable energy, there has been a significant rise in the demand for cost-effective and environmentally sustainable energy storage and transmission methods. Electrochemical water splitting, or water electrolysis, is a sustainable and promising means of evolving H_2 thanks to the wide abundance of water. This process involves two coupled half-reactions: hydrogen evol-

^aWilliam A. Brookshire Department of Chemical and Biomolecular Engineering and Texas Center for Superconductivity (TcSUH), University of Houston, 4226 Martin Luther King Boulevard, Houston, TX 77204, USA. E-mail: rtran17@uh.edu ^bDepartment of Electrical and Computer Engineering, University of Houston, 4226 Martin Luther King Boulevard, Houston, TX 77204, USA. E-mail: jchen84@uh.edu ^cShell Information Technology International Inc., Houston, TX 77082, United Kingdom. E-mail: ligang.lu@shell.com

^eShell Technology Centre Bangalore, Plot#7, Bengaluru Hardware Park KIADB Industrial Park North, Mahadeva Kodigehalli, Bengaluru, Bangalore, Karnataka 562149, India

†Electronic supplementary information (ESI) available. See DOI: https://doi.org/ 10.1039/d4nr01390e



Richard Tran, ^(D)*^{‡a} Liqiang Huang,^{‡b} Yuan Zi,^b Shengguang Wang,^a Benjamin M. Comer,^c Xuqing Wu,^d Stefan J. Raaijman,^f Nishant K. Sinha,^e Sajanikumari Sadasivan,^e Shibin Thundiyil, ^(D)^e Kuldeep B. Mamtani,^e Ganesh Iyer,^e Lars C. Grabow, ^(D)^a Ligang Lu*^c and Jiefu Chen*^b

The efficiency of H₂ production *via* water electrolysis is limited by the sluggish oxygen evolution reaction (OER). As such, significant emphasis has been placed upon improving the rate of OER through the anode catalyst. More recently, the Open Catalyst 2022 (OC22) framework has provided a large dataset of density functional theory (DFT) calculations for OER intermediates on the surfaces of oxides. When coupled with state-of-the-art graph neural network models, total energy predictions can be achieved with a mean absolute error as low as 0.22 eV. In this work, we interpolated a database of the total energy predictions for all slabs and OER surface intermediates for 4119 oxide materials in the original OC22 dataset using pre-trained models from the OC22 framework. This database includes all terminations of all facets up to a maximum Miller index of 1. To demonstrate the full utility of this database, we constructed a flexible screening framework to identify viable candidate anode catalysts for OER under varying reaction conditions for bulk, surface, and nanoscale Pourbaix stability as well as material cost, overpotential, and metastability. From our assessment, we were able to identify 122 and 68 viable candidates for OER under the bulk and nanoscale regime, respectively.

ution reaction (HER) and the significantly slower oxygen evolution reaction (OER) which has primarily been the bottleneck in advancing water splitting technology. The search for a highly active anode catalyst for OER is therefore paramount to the realization of practical water splitting technology.^{1–4}

Transition metal oxides are a promising class of catalysts for OER due to their varying oxidation states allowing for more efficient multi-electron transfer, stability under highly acidic conditions favorable towards OER, and active undercoordinated transition metals sites. In regard to commercialized catalysts, IrO₂ and RuO₂ are the benchmark catalysts for OER, exhibiting low overpotentials (an indicator of activity) of 0.25 to 0.5 V under acidic conditions.⁵ However, the material cost of precious metals (\$18315 and \$155727 per kg for RuO₂ and IrO₂ respectively as of March 2021⁶⁻⁸ with a price variation of ±\$9969 for RuO₂ and +\$80370 for IrO₂ in the last 24 years) limits their widespread adoption. Consequently, there is much desire to identify cheaper materials for catalysts in OER while maintaining similar performance.

Despite the abundance of unary and binary oxides, very few are capable of exhibiting both high catalytic activity and stability under operating conditions. Computational analysis performed by Wang *et al.*⁹ indicated only 68 bimetallic oxides from a pool of 47814 were stable with a Pourbaix decomposition energy (ΔG_{PBX}) of 0.5 eV per atom or less under acidic conditions (pH = 0) and an applied potential cycle between 1.2 and 2.0 V.



View Article Online

^dDepartment of Information Science Technology, University of Houston, 14004 University Boulevard #318, Sugar Land, TX 77479, USA

^fEnergy Transition Campus Amsterdam, Shell Global Solutions International B.V. Grasweg 31, 1031 HW Amsterdam, the Netherlands

[‡] These authors contributed equally to this work.

Gunasooriya and Nørskov¹⁰ calculated the overpotential of these candidate oxides and identified only 11 nonbinary metal oxides with a promising overpotential of less than 0.85 V.

It is possible that many materials filtered out by the highly discriminant electrochemical stability criteria can also exhibit competitive overpotentials if stabilized. Nanoscale stability, elemental doping, and the introduction of oxygen vacancies have been demonstrated to be effective means of improving stability.^{1,11,12} Nanoscaling in particular presents a promising avenue for stabilizing oxides under operating conditions while exposing a greater number of active sites through the increasing surface area-to-volume ratio (SA : V).^{13–15}

However, an accurate evaluation of the nanoparticle formation energy not only requires the thermodynamic contributions of the bulk, but the surface as well. Doing so requires an ensemble of expensive Density Functional Theory (DFT) calculations of the bare surfaces for one material. Likewise, accurate evaluations of the overpotential requires an even larger set of calculations for all the surface intermediates participating in OER. As such, doing so for the massive pool of binary and unary oxide materials available will quickly become computationally expensive and unfeasible.

Machine learning (ML) potentials and screening frameworks have recently contributed significantly in rapidly predicting candidate catalysts without the need for expensive systematic DFT screening.16-20 Among these efforts, the Open Catalyst Project (OCP)²¹ framework has stood out as having the largest dataset of carefully curated DFT calculations for non-oxide slabs and surface intermediates to-date. ML models pre-trained with the OCP has allowed for large scale interpolation efforts to predict binding energies, enabling high-throughput screening efforts to identify viable catalyst candidates²² and explore fundamental surface chemistry.²³ More recently, the Open Catalyst 2022 (OC22) framework has expanded upon this dataset by incorporating random combinations of 4732 oxide materials, Miller indices up to 3, and surface intermediates involved in OER.24 These efforts have yielded predictive ML models with total energy mean absolute error (MAE)s of less than 0.22 and 0.69 eV for in domain (materials observed during training) and out of domain predictions, respectively for the GemNet-OC architecture, respectively.

In this manuscript, we utilized a pre-trained model from the OC22 framework to interpolate the surface energies and OER binding energies for 4119 in domain oxide materials on all facets up to a maximum Miller index (MMI) of 1. To demonstrate the applicability of our interpolated database, we constructed a high-throughput screening framework with a set of progressive criteria that can be modified or expanded upon for ease of customizeability in order to evaluate the commercial and practical viability of each material for OER. Our general framework evaluates materials based on thermodynamic stability, overpotential, and material cost. We also expand upon other screening criteria such as the possibility of nanoscale stabilization or the faceting of surfaces on the equilibrium crystal structure. We propose 190 possible candidates for OER under the bulk and nanoscale regime that warrant further experimental investigation.

2 Methods

All analyses were performed using the python materials genomics $(pymatgen)^{25-28}$ and Atomic Simulation Environment $(ASE)^{29}$ packages.

2.1 Slab generation

We described all facets up to a MMI of 1 containning an atomic and vacuum layer of ca. 12.5 Å thick. The bulk materials used for slab construction in this study were obtained from the Materials Project.³⁰ We also considered all terminations for each facet (sans slabs exceeding 200 atoms) while maintaining equivalent surfaces on both sides of the slab which consequently resulted in non-stoichiometric slabs with respect to the bulk formula. Although the original OC22 dataset covered 4732 distinct bulk oxide materials, the conventional unit cell of some of these materials contains over 100 atoms making the construction of slabs exceed our 200 atom limit for the majority of facets considered. As such, we limit our study to slabs constructed from unit cells of less than 100 atoms. Furthermore, slabs constructed from a select number of materials resulted in the forces in the ML model being unconverged. Thus, we ommitted 609 from the original 4732 materials in the OC22 dataset that exhibited these behaviors, with our final bulk set containing X binary (A-B-O) and X unary (A-O) oxides. For all slabs constructed, we modelled the surface intermediates of O*, OH*, and OOH*. To avoid periodic interactions between the adsorbates, all slabs were expanded along the length and width to at least 8 Å. We assumed all adsorbates bind through the O atom on available undercoordinated metal sites. All bare surface and surface intermediate models were constructed using the python framework adapted from OC22.^{24,31,32}

2.2 DFT and machine learning settings

All DFT calculations were performed using the Vienna *Ab Initio* Simulation Package (VASP)³³⁻³⁶ within the Projected Augmented Wave (PAW)³⁷ approach. We modeled the exchange–correlation effects with the Perdew–Berke–Ernzerhof (PBE) generalized gradient approximation (GGA) functional.³⁸ All calculations were performed with spin-polarization with a plane wave energy cut-off of 500 eV. The energies and atomic forces of all calculations were converged to within 1×10^{-4} eV and 0.05 eV Å⁻¹, respectively. We used Γ -centered *k*-point meshes of $\frac{30}{a} \times \frac{30}{b} \times 1$ for slab calculations, with non-integer values rounded up to the nearest integer. We also apply a Hubbard U correction to chemical systems as suggested by the Materials Project³⁹ to account for missing electron interactions.

We used a pre-trained model for the *Structure to Energy and Forces* (S2EF)-*Total* task from the OC22 framework to perform all machine learning predictions of the relaxed structure and total energy. The entirety of the Open Catalyst 2020 (OC20) dataset (1281040 DFT relaxations) was used to train an S2EF-*Total* model which was subsequently fine-tuned with the OC22 dataset (62331 DFT relaxations) to better predict the total energies of oxide surfaces and surface intermediates. The model was trained using the

GemNet-OC architecture⁴⁰ due to its superior performance in energy predictions when compared to other graph neural network (GNN) architectures as a consequence of its improved capturing of long-range and quadruplet interactions.

Nanoscale

For further details regarding additional parameters used in VASP or the construction of the machine learning model, we refer the reader to Tran *et al.*²⁴

All OC22 S2EF predictions and DFT calculations of slabs were performed with selective dynamics. In regards to surface energy predictions, both the bottom most and topmost layer of atoms within 1.25 Å were allowed to relax in the bare slab in order to ensure both surfaces had equal surface energy contributions. For the adsorption energies, only the topmost layer of atoms within 1.25 Å and any adsorbates were allowed to relax in the bare slab and surface intermediates. To avoid inadvertent desorption and dissociation of adsorbates during our ML relaxation, we applied a spring constant of 7.5 eV $Å^{-2}$ between all adsorbate atoms to preserve the identity of the molecule and between the adsorbate and host surface atoms whenever the adsorbate drifts 2 Å away from its initial position along the axis perpendicular to the surface.^{41,42} Similarly, we also applied the same restorative force to all surface atoms when the relaxed trajectory exceeds 1 Å from the initial position of the ions to avoid drastic surface reconstruction. While this approach yields better interpretability for relaxed adsorption geometries, we recognize the inherent artificiality of these constraints and acknowledge the potential for desorption and dissociation. Consequently, we further implemented ML relaxation without such constraints on relaxed geometries exhibiting low overpotentials to verify the absence of desorption and dissociation phenomena. All values of overpotentials and Gibbs free energies reported in the main manuscript as well as Tables S1-S14[†] are relaxed without these additional constraints unless stated otherwise. Comparisons of the overpotential and Gibbs free energy obtained with and without these constraints can be found in the ESI.†

2.3 Surface thermodynamics

All bare slabs of formula $A_{nx}B_{ny+k}O_{nz+j}$, with A and B being two metal components, are constructed from a bulk ternary oxide of $A_xB_yO_z$. We can calculate the surface energy of any slab of $A_{nx}B_{ny+k}O_{nz+j}$ with the following:

$$\gamma(\mu_{\rm B},\mu_{\rm O_2}) = \frac{E_{\rm slab}^{A_{\rm nx}B_{\rm ny+k}O_{\rm nz+j}} - \left(nE_{\rm bulk}^{A_{\rm x}B_{\rm y}O_{\rm z}} + k\mu_{\rm B} + j\frac{1}{2}\mu_{\rm O_2}\right)}{2A}$$
(1)

where $E_{\text{slab}}^{A_{nx}B_{ny+k}O_{nz+j}}$ is the total energy of the bare slab, $E_{\text{bulk}}^{A_xB_yO_z}$ is the total energy per formula unit of the bulk crystal, and *A* is the surface area. We used correction values from the Materials Project when evaluating γ to account for the mixing of quantities determined with GGA and GGA+U.^{26,39,43} The chemical potentials, μ_i , accounts for any nonstoichiometric species (with respect to bulk stoichiometry) in the slab formula.

The chemical potential of oxygen $(\Delta \mu_{O_2})$ can be referenced to the electrochemical decomposition of water to $O_{2(g)}$:

$$2H_2O_{(g)} \rightarrow 4(H^+ + e^-) + O_{2(g)}:$$
 4.92 eV (2)

which allows us to rewrite μ_{O_2} (and thereby γ) as a function of pH and applied potential (*U*) as such:

$$\Delta \mu_{\rm O_2} = 4.92 + 2\mu_{\rm H_2O}^{\circ} - 4\left(\frac{1}{2}\mu_{\rm H_2}^{\circ} - eU - k_{\rm B}T\,\rm{pH}\,\ln\,10\right) + \Delta G_{\rm corr}^{\rm O^{\circ}}$$
(3)

where $\Delta G_{\rm corr}^{O'}$ corrects for the Gibbs free energy of excess or deficient oxygen at the surface (see Tran *et al.*²⁴ and Gunasooriya and Nørskov¹⁰ for details). We will assume typically employed operating conditions for acidic OER (pH = 1 and U = 1.8 V at $T = 80 \,^{\circ}{\rm C}$ or $60 \,^{\circ}{\rm C}$)^{9,44} ($\Delta \mu_{O_2} = -1.30 \,^{\circ}{\rm eV}$) when assessing the surface energy of all materials as illustrated in Fig. 1(a and b). For slabs containing excess or deficient metal (B) species, the chemical potential of component B is conventionally referenced with respect to the per atom energy of the ground state bulk crystal of pure component B (*e.g.* $\mu_{\rm Fe} = \Delta \mu_{\rm Fe} + E_{\rm BCC,Fe}^{\rm DFT}$). By varying the chemical potential of component B, we can stabilize different surface terminations of the same facet as shown in the surface Pourbaix diagram in Fig. 1(a and b).

To determine the nanoscale stability of metastable and unstable materials under operating conditions, we assessed the nanoparticle formation energy given by:

$$G_{\rm f}^{\rm NP} = E_{\rm V}({\rm pH}, V, T) \left(\frac{4}{3}\pi r^3\right) + \bar{\gamma}({\rm pH}, V, T, \Delta\mu_{\rm M}) \left(4\pi r^2\right) \qquad (4)$$

whereby $E_{\rm V}$ is the Pourbaix formation energy per volume of the unit cell, $\bar{\gamma}$ is the weighted surface energy of the Wulff shape (an analogue to the nanoparticle morphology), and r is the radius of the nanoparticle. Detailed explanation of these quantities can be found in the ESI.† Fig. 2 demonstrates how a less stable compound (CaTi₂O₅) can become more stable than the ground state compound as nanoparticle size and $\Delta \mu_{\rm Ti}$ decreases. The size effect will change the relative contribution of surface energy and bulk Pourbaix formation energy to $G_{\rm f}^{\rm NP}$



Fig. 1 Surface Pourbaix diagram of the (001) facet of CaCr₂O₄ at T = 80 °C under $\Delta\mu_{Cr} = -4$ eV (a) and $\Delta\mu_{Cr} = 0$ eV (b). The blue dashed line indicates the equilibrium conditions for OER while the black marker indicates reference conditions for OER (pH = 1 and U = 1.8 V). The phases in (a) and (b) are color coded with the corresponding terminations of the (001) facet in (c).



Fig. 2 Nanoscale stability phase diagram for the Ca–Ti–O chemical system plotted using G_f^{NP} as a function of the chemical potential of Ti $(\Delta \mu_{Ti})$ and nanoparticle radius. The red region represents the less stable (higher E_{PBX}) phase (CaTi₂O₅) while the cyan region represents the ground state phase (Ca₂Ti₃O₈). The nanoparticle morphology and surface energy change as a function of $\Delta \mu_{Ti}$ indicated by the inset Wulff shapes.

while $\Delta \mu_{Ti}$ changes the overall particle morphology and thereby surface energy of the particle.

The scope of this study will explore the water nucleophilic attack (WNA) mechanism for OER, a four step mechanism where two water molecules sequentially bind to a metal at the surface and release an electron–proton pair at each step as shown in Fig. 3.⁴⁵ We realize and emphasize that the WNA mechanism is one of many approximations for modeling OER and that it is possible for certain materials to prefer alternative mechanisms *e.g.* the oxo-coupling mechanism or lattice oxygen evolution reaction.^{46–50} However, we focus on the WNA mechanism on account of its ubiquity in computational studies^{51–55} and supporting experimental evidence.^{47,56} We can determine the overpotential for this reaction by identifying

the largest energy difference between each step, reaction energy (ΔG_{rxn}) with the following:

$$\eta = \max(\Delta G^{i}, \Delta G^{ii} - \Delta G^{i}, \Delta G^{iii} - \Delta G^{ii}, 4.92 - \Delta G^{iii})/e$$

- 1.23 V (5)

where ΔG^{i} , ΔG^{ii} , and ΔG^{iii} are the Gibbs free energy of each reaction step, 4.92 eV is the Gibbs free energy to dissociate two water molecules into O₂ and 4(H⁺ + e⁻) shown in eqn (2), and 1.23 V is the equilibrium potential for water decomposition. Here, the step corresponding to the largest value of ΔG_{rxn} is also called the potential determining step (PDS). We can derive the Gibbs free energy of each step listed in eqn (i)–(iv) (see Fig. 3) as such:

$$\Delta G^{i} = E_{\rm ads}^{\rm OH^{*}} + \Delta G_{\rm corr}^{\rm OH^{*}} + \mu_{\rm H^{+}} + \mu_{\rm e^{-}}$$
(6)

$$\Delta G^{\rm ii} = E^{\rm O^*} + \Delta G^{\rm O^*}_{\rm corr} + 2(\mu_{\rm H^+} + \mu_{\rm e^-})$$
(7)

$$\Delta G^{\rm iii} = E^{\rm OOH^*} + \Delta G^{\rm OOH^*}_{\rm corr} + 3(\mu_{\rm H^+} + \mu_{\rm e^-})$$
(8)

where $E_{ads}^{OH^*}$, $E_{ads}^{O^*}$, and $E_{ads}^{OOH^*}$ are the electronic adsorption energies of the intermediates for OER and $G_{corr}^{OH^*}$, $G_{corr}^{O^*}$, and $G_{corr}^{OOH^*}$ are correction terms for the Gibbs free energy derived in the ESI† of OC22.²⁴

To minimize the number of predictions needed, we will begin by using a quick scaling relationship given by $\Delta G^{iii} = \Delta G^i + 3.26^{10,24}$ to estimate ΔG^{iii} . This approach is particularly beneficial for OOH*, where the significantly greater rotational freedom leads to a substantial increase in potential adsorbate placements. We will then perform additional predictions for E^{OOH*} for surfaces exhibiting promising activity ($\eta < 0.75$ V) using eqn (8) to more accurately determine the overpotential. ML relaxations exhibiting dissociation or desorption of intermediates are omitted in any interpretation of η . More details regarding dissociation and desorption events occurring in the dataset as well as a comparison between η obtained with eqn (8) and scaling relationships can be found in the ESI.[†]

Only the most stable site for OH* dictated by $E_{ads}^{OH^*}$ from a set of considered adsorption sites on the same surface is con-



Fig. 3 Reaction diagram for the WNA on an ideal catalyst (dashed lines) and $LaMnO_3$ (100) surface (solid lines) at 0 V (black) and the equilibrium potential of OER (red). An illustration of each reaction step is shown at the bottom (red circles are oxygen while grey circles are hydrogen/protons). Reaction energies for LaMnO₃ were derived from Man *et al.*⁴⁵ The chemical equation between each reaction step (1–4) is listed on the right (i–iv).

sidered when determining η . We maintain the adsorption site corresponding to the most stable site of OH* when considering $E_{ads}^{O^{\circ}}$ and $E_{ads}^{OOH^{\circ}}$.

Further details regarding the derivation of all thermodynamic quantities and scaling relationships can be found in the ESI.[†]

3 Results and discussion

3.1 Database scope and usage

We emphasize that the purpose of both the OC20 and OC22 datasets was to establish a large and diverse set of DFT calculated surfaces and surface intermediates for the purpose of training ML potentials generalized to infer the total DFT of any slab and adsorbate combination. To maximize the diversity of the sample set, the OCP curated DFT calculations of randomly selected combinations of materials, surfaces, and adsorbates. The scope of the dataset does not encompass a comprehensive database for evaluating η directly, as many necessary data points are missing illustrated in Fig. 4. The construction of a comprehensive database is a herculean task that is immeasurably costly and time consuming with DFT alone. However, by consolidating the S2EF-total model to predict thermodynamic overpotentials, we can potentially infer the catalytic activity of large material datasets for OER within a reasonable degree of error with respect to DFT.

Using the previously developed ML models, we systematically extrapolated the total DFT energy of all terminations for



Fig. 4 Scope of surface intermediates for OER calculated using DFT in the OC22 dataset. Overlaps indicate the number of surfaces where different intermediates (OH*, O*, or OOH*) are calculated for the same surface.

all facets up to a MMI = 1 for all 4119 materials considered in this study. We emphasize that although the original training pool does consider selections of bare and adsorbed surfaces up to MMI = 3, the complete dataset is biased towards facets with a MMI = 1 with 39573, 15482, and 7276 data points of MMI = 1, 2 and 3 respectively. We expect this bias to allow for better predictions of the facets (all of which exhibit MMI = 1) considered in this work. We then extrapolated the total energy of all metal adsorption sites on all surfaces for O* and OH* (and OOH* when $\eta < 0.75$ V). Table 1 summarizes the scope of our extrapolated database. In contrast to our interpolation efforts, the size of the OC22 dataset for O* and OH* is only 0.1% of the predicted dataset in this work. Despite requiring orders of magnitude less computational resources than a DFT dataset of the same size, the estimated cost to produce our dataset is still 9473.9 GPU-hrs (an average rate of one prediction every 12.5 seconds), a significant amount of resources. By making this database freely available to the scientific community, users interested in performing similar high-throughput screening exercise or fundamental analysis can do so without the enormous cost in GPUs. The entire database including the initial and relaxed structures and total energies can be accessed through the University of Houston Dataverse Repository.⁵⁷ Details regarding the database metadata are given in the ESI.†

For the 12922 surfaces exhibiting $\eta < 0.75$ V, we performed an additional ML relaxation step without the application of constraints on surface relaxation, adsorbate dissociation, and desorption. While only 7% of OH* and 1% of O* intermediates exhibited dissociation and desorption event, an overwhelming number of events (50%) were associated with the OOH* intermediates. Consequently, predictions of η exhibiting these events could not be interpreted and were ignored in our final assessment of overpotential. However, we emphasize that the occurrence of dissociation and desorption does not disqualify the possibility of these surfaces being catalytically active in alternative reaction mechanisms such as the oxo-coupling mechanism or lattice oxygen evolution reaction.^{46–50}

3.2 High-throughput screening

Table 1 Summany of database score

Fig. 5 summarizes the selection criteria that we employed to screen for candidate electrocatalysts for OER. The first criteria

Table I Sum	nary of ualabase scor	Je	
	Prediction	s: 6068572	
	Materia	als: 4119	
	Ave. # slabs p	er material: 47	
OH*	O*	OOH*	*
1,972,166	667,266 Predictions w/o s	3,237,238 spring constraints	191,902
	Predictions	s of η: 12922	
	Dissociation/de	esorption events	
OH*	0,	÷ -	OOH*
908	13	1	6296
Predictions of	n w/o diss./des. event	ts: 6468	

Published on 06 August 2024. Downloaded by University of Houston on 8/28/2024 8:58:33 PM



Fig. 5 Summary of the screening criteria considered in our highthroughput screening framework. The possible number of candidates that have satisfied each tier is listed on the right for bulk Pourbaix stable materials and on the left when nanoscale stability is possible. The second and third criteria can be further modified by changing the environmental parameters (T, U, pH).

in our high-throughput screening framework was to determine if a material exists in the original training dataset of OC22. As previously demonstrated by Tran *et al.*,²⁴ the OC22 framework is capable of predicting the total DFT energies of slabs and surface intermediates of oxides within a mean absolute error of 0.22 eV for the 4119 materials that have been observed in the training dataset. Our database will only interpolate the energies of slabs and surface intermediates amongst these materials (for more information, the reader is redirected to²⁴).

The second criterion describes the Pourbaix stability, i.e., the electrochemical stability of a material in aqueous environment. Here, we can interpret Pourbaix stability under a bulk regime (right side of Fig. 5). We quantify the Pourbaix stability of the bulk using the Pourbaix decomposition energy (E_{PBX}) , which is a function of the temperature (T) applied potential (U) and pH of the environment (at T = 80 °C, U = 1.8 V, and pH = 1).^{27,28} Materials with E_{PBX} = 0 eV per atom are stable under such conditions while materials with $E_{PBX} > 0$ eV per atom are metastable with the likelihood of corrosion increasing with E_{PBX} . It was shown experimentally that metastable materials with $E_{\text{PBX}} \leq 0.2$ eV per atom are less likely to dissolve or corrode.⁵⁸ However, materials with E_{PBX} as high as 0.5 eV per atom have also been shown to be stable, albeit with a degree of surface passivation which can inhibit catalytic activity.²⁷ We allow any material with $E_{PBX} \leq 0.5$ eV per atom under the aforementioned conditions (see Fig. 1a) to satisfy this criterion. Due to the exclusive nature of E_{PBX} , only 1853 of the original 4119 materials will satisfy this criterion.

We adapted our third selection criterion based on the selection criterion from the WhereWulff⁵⁹ high-throughput screening workflow for oxide catalyst discovery. Here, we assess the surface energy of every termination for each facet of each material that is Pourbaix stable with eqn (1). From the surface energy, we were able to construct the Wulff shape which indicates the most prominent facets in an equilibrium crystal. When considering which surfaces are stable and can potentially facilitate OER, we only consider surfaces that appear on the Wulff shape. Depending on the stoichiometry of the slab, the surface energy can vary as a function of $\Delta \mu_{\rm M}$. For simplicity, we will roughly assume a possible chemical potential range of the metal M as $-5 < \Delta \mu_{\rm M} < 0$ eV when interpreting surface energy. Wulff shapes containing negative surface energies are ignored as nonphysical solutions as this indicates the surface is more stable than the bulk (which implies dissolution of the solid). The dissolution of the Wulff shape will consequently lead to 314 additional materials being omitted from our list of candidates, leaving us with 1539 materials with 11918 stable surfaces on the Wulff shape.

In our fourth criterion, we assess the overpotential of each candidate surface with eqn (5). We consider any material with at least two facets on the Wulff shape exhibiting $\eta < 0.75$ V as being potentially competitive with RuO₂ and IrO₂ in regards to catalytic performance. We find that 101 materials (512 surfaces) from our previous 1539 materials will satisfy this criterion.

The fifth criterion assesses the thermodynamic stability of the candidate material *via* the energy above hull (E_{hull}) or the formation energy of a material relative to the ground state. Like with Pourbaix stability, materials with a calculated $E_{hull} >$ 0 eV per atom are metastable with the likelihood of experimental synthesizeability decreasing as E_{hull} increases. Materials with a calculated $E_{hull} \leq 0.1$ eV per atom have been shown to have reasonable rates of demonstrated synthesis in experiment.⁶⁰ From our aforementioned 101 catalytically active candidates, we identify 92 materials that are stable or metastable.

The final criterion assesses the material cost of each compound in \$ per kg. To satisfy this criteria, the cost of a compound must be less than that of RuO₂ or \$18315 per kg (as of March 2021^{6–8}). As a conservative estimate, we will also assume a price variation of \$9969 per kg (based on the lowest and highest prices of Ru in the last 24 years) in the future and as such, all materials must be less than \$8346 per kg. We have identified 81 materials and 66 distinct A–B–O chemical systems from our 101 metastable materials that have satisfied this criterion. A tabulated list of all candidates with the lowest value of η , E_{hull} , E_{PBX} , material cost, space group, and number of facets with low overpotentials is give in Tables S1–S8.†

3.3 Overpotential assessment

Although our set of tiered screening criteria will provide the most economically and industrially viable candidates with respect to thermodynamic stability and material cost, we posit that there is a wider range of materials that can have competitive overpotentials necessary for OER when ignoring the criteria listed above. Fig. 6(a) shows the 53 elements on the periodic table considered in our database. The heatmap indicates the number of materials containing an element that exhibits at least 2 facets with $\eta < 0.75$ V. Oxide combinations containing Cr, Mn, Co, Cu, Bi, Pb, Se, Tl, Ag, and Sb tend to form catalysts that exhibit low overpotentials. Ag-based chemical systems such as Ag–O,⁶¹ Ag–Bi–O,⁶² Ag–Cu–O,⁶³ and Ag–Co–O⁶⁴ have been demonstrated as viable catalysts for OER in experimental



Fig. 6 (a) Periodic table of elements considered in the database. Colormap indicates the number of materials containing the element that exhibit $\eta < 0.75$ V for at least two facets. (b) Grid map for each pair of elements with colors indicating the lowest overpotential amongst all facets on the Wulff shape containing this chemical system (bottom right) and the number of facets on the Wulff shape across material with the same chemical system that exhibit overpotentials less than 0.75 eV (top left). Tick labels on the *x*- and *y*-axis are sorted from the cheapest (Fe) to the most expensive (Rh) element.

settings. The common use of Ag has been attributed to the fact that Ag has the highest electrical conductivity amongst all metals, allowing it to easily facilitate the four electron charge transfer process that takes place during OER.⁶⁴ Although it is not as common and cheap as the 3d transition metals, Ag is still more reasonably priced than the other noble metals. Furthermore, it has been shown that Ag-doping at 1% has been enough to enhance charge transfer in OER,⁶⁵ making Agbased oxides reasonably economical. Similarly, Mn-based chemical systems are known to be promising catalysts for OER. This is owed to the intrinsically high activity and number of polymorphs for compounds of MnO_x⁶⁶ which can be synergistically improved when introducing other components such as Fe,67 Ni,68 and Co.69 Antimonates (Sb-based oxides) have been extensively studied in both computational¹⁰ and

experimental^{62,70–72} settings as promising low-cost anode catalysts for OER owed to their low overpotential and high operational stability as a consequence of Sb–O p–d hybridisation.⁷² Mixed metal Co oxides are abundant in CO^{2+} cations that are useful for OER. The large atomic difference between Co and the larger metal in Co-based Perovskites can also lead to distortions in the structure that can better stabilize either the CO^{2+} or CO^{3+} cation which can allow for tunable active sites.⁷³

Non-noble chemical systems such as Fe–Ni,⁷⁴ Cu–Fe, Fe–Mn,⁷⁵ Fe,⁷⁶ Co–Mn,⁷⁷ and Co–Cr⁷⁸ based oxides have been shown to exhibit competitive overpotentials in the experimental literature despite their absence from our final set of candidates. To account for this discrepancy, we predicted the overpotentials for all Wulff shape facets of all 4119 materials. We summarized the overpotential of each chemical system by plotting the lowest overpotential and the number of facets exhibiting $\eta < 0.75$ eV in Fig. 6(b) across all materials of the same chemical system as a heat map. Upon further inspection, Fe–Mn–O is shown to be relatively unstable in our dataset ($E_{Pbx} = 0.95$ eV) despite having a competitive overpotential corroborating with past experiments. However, this was explored in the context of nanoparticle catalysts.⁷⁵

We have demonstrated 380 chemical systems exhibiting facets on the Wulff shape that are potentially active for OER despite only 76 chemical systems (92 materials) (ignoring material cost) appearing in our set of candidates due to the bulk stability of many materials being inaccessible. Methods for stabilizing bulk oxides such as nanoscale stability (e.g. in the aforementioned case for Fe-Mn-O⁷⁵), elemental doping, and the introduction of oxygen vacancies have been demonstrated to be effective means of improving stability.¹ RuO₂ for example is known to have stability issues in aqueous environments despite being the benchmark for OER catalysts. However doping with other metals such as Cr and Ti has been shown to stabilize this material.¹¹ With these methods of stabilizing the catalyst material and accessing active surfaces, we emphasize that the identification of viable catalysts should not strictly be confined by the Pourbaix stability or even E_{hull} . As an example, we will demonstrate how nanoscale stabilization can potentially expand the material space available for OER to access these materials.

3.4 Alternative screening frameworks

Table 2 shows the number of final candidates identified under different reaction conditions and definitions of Pourbaix stability. In this study we constructed a database of machine learning data and demonstrated how it can be used in a variety of ways to screen for catalyst. First, we identified potential candidates for OER by creating a screening framework based on simple bulk thermodynamic (the energy above hull and Pourbaix energy above hull) and surface stability arguments at pH = 1, U = 1.8 V, and T = 80 °C (see *j* in Table 2). We can easily modify our framework to reflect other reaction conditions and criteria for Pourbaix stability as shown in Table 2. For example, we can modify our framework by introducing an additional layer of complexity to our definition of Pourbaix

Table 2 The different screening frameworks assessed in this study with varying reaction conditions and criteria for Pourbaix stability. All screening frameworks are assumed to occur under pH = 1. Superscript letters are used to label each framework evaluated (see Tables S1–S8† for a list of candidates that were identified in each framework). Numbers in parentheses correspond to conductive candidates with small band gaps ($E_{gap} < 0.1 \text{ eV}$). Framework *j* and *k* correspond to the frameworks investigated in Fig. 5

Temperature (°C)	60	60	80	80
Applied potential (V)	1.8	1.2 to 2.0	1.8	1.2 to 2.0
Bulk	122^{a}	99^{e}	120^{i}	99^{m}
Bulk/Wulff	83^{b}	62^{f}	81^{j}	62^{n}
Bulk/Wulff/nano	111^{c}	83^{g}	121^{k}	84^{o}
Bulk/nano	168^{d}	129^{h}	181^{l}	129^{p}

stability by analyzing the possibility of nanoscale stability under acidic conditions for OER as demonstrated in Fig. 5 (right). Using our nanoscale stability diagrams (see Fig. 2), we were able to identify 2778 Pourbaix stable materials with 886 stabilizing at the nanoscale regime (10 to 100 nm). From our Wulff shape analysis of the 2778 Pourbaix stable materials, we identified 18501 surfaces that appear on the Wulff of 2430 materials. We find that 816 of these surfaces (from 159 materials) also exhibit low overpotentials. From these 159 potentially active materials, 147 are metastable and 121 exhibit a material cost less than RuO_2 . In total we have identified 121 candidates and 40 additional candidates that are potentially commercially viable for OER when synthesized as nanoparticles (see *k* in Table 2).

Another example of how we can modify our screening framework is by changing our criteria for surface stability. As mentioned before, we only considered facets that appear on the Wulff shape within a metal chemical potential between – $5 < \Delta \mu_{\rm M} < 0$ eV. The Wulff shape indicates the most statistically likely facets to appear under equilibrium crystal growth conditions. However, non-equilibrium conditions can potentially force different types of facets to appear.⁷⁹ We can account for this in our screening framework by considering all facets as viable surfaces for OER. With this simple assumption, we identified 181 materials (120 and 61 in the bulk and nanoscale regime respectively) that satisfy all criteria (see *l* in Table 2). Among the A–B–O chemical systems identified as nanostable is Fe–Mn–O which as mentioned previously, has been explored as nanoparticle catalysts for OER.⁷⁵

Other alternative screening frameworks can be explored by simply adding or modifying existing criteria. As a simple example, we can assess Pourbaix stability under different temperature conditions. The operating temperatures for OER are typically between 60 °C and 80 °C with higher temperatures resulting in improved ionic conductivity and kinetics in exchange for lower stability. Although our assessment assumes an operating temperature of 80 °C, we can easily re-assess the second criteria (for bulk stability) in our screening framework at 60 °C instead which will yield 2 additional candidates without considering nanoscale stability (*a* and *b*). When nanoscale stability is considered however, lowering the temperature from 80 °C to 60 °C will result in a loss of 10 (*c*) and 13 (*d*) candidates with and without Wulff shape stability respectively. The bulk Pourbaix formation energy of ground state polymorphs decreases at lower temperatures. This decreases the likelihood for metastable polymorphs to become more stable than the ground state at the nanoscale resulting in fewer nanostable candidates at lower temperatures.

We can also consider more complex and realistic reaction conditions whereby candidates are assessed based on their stability over a range of operating potentials. We adopted the potential range of 1.2-2.0 V for OER for which a material must remain stable (or metastable with $E_{PBX} \leq 0.5$ eV) as suggested by Wang *et al.*⁹ (*e*-*h* at T = 60 °C and *m*-*p* at T = 80 °C). This strict set of operating conditions unsurprisingly results in fewer candidates when compared to a static potential operating condition of 1.8 V. The total number of candidates presented in this work however, still far exceeds the original set of candidates identified by Wang et al.9 and subsequently proposed by Gunasooriya and Nørskov.¹⁰ The OC22 framework is trained on data calculated with the PBE-GGA functional whereas the data presented by Wang et al.9 was obtained using the more accurate SCAN functional⁸⁰ with the addition of correction terms better suited for assessing stability under corrosive conditions. These considerations contributed to a more realistic, albeit pessimistic, set of 11 Pourbaix stable and active candidates subsequently found by Gunasooriya and Nørskov¹⁰ as oppose to the 99 bulk stable candidates identified at 80 °C (m). In a future work, we hope to adapt these functionals and corrections when assessing nanostable catalysts to provide a more accurate expansion to the limited set of bulk stable candidates explored by the references herein.9,10

In total we have identified 190 candidates (122 bulk – and 68 nanostable) with 145 distinct chemical systems when considering all the different screening frameworks listed in Table 2. In the next section, we will validate our findings by comparing to past experimental and computational results as well as our own DFT simulations. We hope this demonstration regarding the ease and variability of how these frameworks can be modified illustrates the utility of these machine learning databases. We highly encourage the scientific community to use our database to explore further possibilities and alternative screening frameworks in the future.

3.5 DFT and literature validation

We used DFT to validate 85 values for Gibbs adsorption energy corresponding to eqn (6)–(8) for 33 Pourbaix stable compounds in our database. The MAE of the test set shown in Fig. 7(a) is 0.42 which is larger than the MAE of 0.239 eV obtained from the validation set in the original OC22 assessment.²⁴ We find no difference in the amount of error when validating ML data points with (square) and without (circle) desorption/dissociation event. However, ML values corresponding to O* adsorption on Ag₂SeO₃ and Na₂Se₂O₇ and OOH* adsorption on Ag₂SeO₃ will underestimate ΔG_{ads} relative to DFT. These data points exhibit better agreement with DFT

Nanoscale



Fig. 7 (a) DFT calculated data points for reaction energy plotted against the corresponding ML predicted quantities. (b) DFT calculated data points for overpotential plotted against the corresponding ML predicted quantities. Overpotentials of the benchmark materials are highlighted in blue (IrO_2 and RuO_2). Overpotentials of less than 0.75 eV are considered catalytically active. Square data points indicate desorption or dissociation of the intermediate. Transparent data points correspond to ML predicted data points with spring constraints applied to prevent desorption and dissociation events.

when applying a spring constant constraint to prevent dissociation/desorption events (transparent data points). This implies that the model from OC22 may overestimate the tendency for desorption/dissociation in some combinations of intermediates and surface and could be a potential point of improvement in future iterations of the model.

Next we assess the predictability of the ML inferred overpotential (eqn (5)) using the predicted and calculated values of ΔG_{ads} from Fig. 7(a). Fig. 7(b) once again plots the DFT calculated data points against the corresponding ML quantities. The majority of data points sampled will lie within the MAE of 0.22 V, however overpotentials corresponding to KSnO₂, Na₅ReO₆, Ag₂SeO₃, and Na₂Se₂O₇ will lie outside the MAE. Ag₂SeO₃ and Na₂Se₂O₇ will exhibit overpotentials closer to parity with DFT when applying a spring constant constraint to prevent the dissociation/desorption of O^{*} and OOH^{*} in Ag₂SeO₃ and O^{*} in Na₂Se₂O₇ as shown in Fig. 7(a). Despite the large deviation from MAE however, we find that most data points with predicted low overpotentials will still exhibit low overpotentials close to or less than 0.75 V using DFT with the exception of KSnO₂ and Na₅ReO₆.

From our DFT calculations, we were able to identify 6 data points with overpotentials below the soft theoretical limit (η_{Th} > 0.3 V) imposed by scaling relationships: Mn₂₃FeO₃₂, HgSeO₄, Na₂Co₂O₃, Cd₂PbO₄, MnTlO₃, and KBiO₂. The Mn–Fe–O chemical system is well explored in the literature with experimental overpotentials as low as 0.47 V.^{75,81,82} Although no studies have investigated ordered structures of Na–Co–O, the doping of layered CoO₂ with Na has resulted in overpotentials as low as 0.24 V.⁸³ Although we have predicted HgSeO₄ and Cd₂PbO₄ as having overpotentials less than η_{Th} , Pb, Cd, and Hg are known to possess potential health risks⁸⁴ and as such, caution is advised for any further investigation of candidates containing these elements. As far as we are aware, the Mn–Tl– O and Bi–K–O chemical systems have yet to be explored in the literature. Lastly, we compared our predicted values of overpotential to those obtained by Gunasooriya and Nørskov¹⁰ for various facets of 11 materials in Fig. 7(b). The majority of our predicted datapoints lie within a MAE of 0.38 V. ML data points tend to underestimate the DFT values. However, those that lie well beyond 0.38 V, such as IrO_2 , $Ni(SbO_3)_2$, $TiSnO_4$, $Sn(WO_4)_2$, $Sn(WO_4)_2$, FeSbO₄ will again exhibit overpotentials much closer to parity with DFT when applying a spring constant constraint to prevent the dissociation/desorption. A key difference in the methods used to obtained both datasets is that the surfaces observed by Gunasooriya and Nørskov¹⁰ can be covered completely by O* or OH* whereas all surfaces considered in this study only considered the adsorption of a single O* or OH* intermediate at a time. Despite this, we still find many data points terminated by O* or OH* within the MAE.

Tables S9-S14[†] lists all 190 candidate materials we have identified in this work along with references to the experimental literature where available and the PDS of the surface with the lowest overpotential. In total, we have identified 102 out of 145 unique chemical systems (127 out of 190 materials) that have yet to be explored with 77 non-toxic (does not contain Pb, Cd, Hg, or Cr) chemical systems (98 materials). Although not considered in our study, conductivity is also required for the operation of electrocatalysts.9 Of these 98 non-toxic and unexplored candidates, 27 have a band gap of less than 0.1 eV according to the Materials Project³⁰ which satisfies this additional criteria. These conductive candidates are: Cu₃ (SbO₃)₄, Ba₈(Bi₂O₇)₃, BaBiO₃, Cu₃Mo₂O₉, Mg(BiO₃)₂, CuMoO₄, Li (Bi₃O₅)₄, Ce₂Mo₄O₁5, Ce₉YO₂₀, AgSnO₃, Ce₂ (GeO₃)₃, Ag₄GeO₄, MnTlO₃, LuCoO₃, BaMn₂O₃, Li(CuO)₂, MnBiO₃, Mn₂BeO₄, VZn₂O₄, ScMn₂O₄, LiMn₃O₄, Mn₃NiO₄, Mn₂NiO₃, VSbO₄, Ag₃RuO₄, TiCu₃O₄, and TlCuO₂. The other 71 candidates, although non-conducting, can potentially be considered as anodes in photocatalytic mechanisms for OER. The discovery of a potential candidate for OER demonstrates the potential of MLassisted screening techniques in identifying novel catalysts.

4 Conclusions

In this contribution, we employed pre-trained machine learning potentials from the Open Catalyst Project to develop a publicly available database of O*, OH*, and OOH* surface intermediates for 4119 oxide materials. We demonstrated the utility and variability of this database by presenting several easily implemented high-throughput screening frameworks for identifying thermodynamically stable and catalytically active (low overpotential) materials for the oxygen evolution reaction from our initial pool of 4119 candidates. Our first screening framework identified 81 candidates that are Pourbaix stable in the bulk regime with catalytic surfaces that appear on the Wulff shape. By slightly modifying this framework to account for the possibility of nanoscale stability, we identified 121 additional candidates. Additional modifications to the reaction conditions and our definition of Pourbaix stability yields a total of 190 candidates with 27 unexplored candidates fulfilling additional criteria of being non-toxic and conductive. Furthermore, we were able to validate our predictions with DFT calculations from the literature as well as our own. When ignoring material cost and bulk stability, we find that oxides containing Cr, Mn, Co, Cu, Se, Sb, Bi, Pb, and Tl tend to exhibit overpotentials that are possibly competitive with current benchmark materials (IrO₂ and RuO₂). We plan to further expand our database to include other potential reaction mechanisms for OER such as the oxo-coupling mechanism and lattice oxygen evolution. We hope this database will encourage future investigators to develop their own highthroughput screening frameworks.

Author contributions

Richard Tran implemented the screening framework and performed DFT calculations. Liqiang Huang designed the parameters and software to manage and automate predictions and performed ML predictions. Yuan Zi, Shengguang Wang, and Benjamin M. Comer defined the overall goal of the project and provided scientific insight and solutions. Lars Grabow defined the DFT aspect of the work. Jiefu Chen and Xuqing Wu defined machine learning aspect of the work. Jiefu Chen and Lars Grabow acquired funding for the project. Ligang Lu initiated and funded the project and recruited Shell Subject Matter Experts to the project. Ligang Lu, Stefan J. Raaijman, Nishant K. Sinha, Sajanikumari Sadasivan, Shibin Thundiyil, Kuldeep B. Mamtani, and Ganesh Iyer defined the industrial benchmarks and goals of the project. All authors contributed to and commented on the manuscript.

Data availability

Data for this article, including the ML predicted initial and relaxed structures and total energies are available at the University of Houston Dataverse Repository at https://doi.org/ **10.18738/T8/APJFTM.**⁵⁷ Details regarding the database metadata are given in the ESI.† The code for all analysis and data generation can be found on Github with **https://doi.org/ 10.5281/zenodo.12616291.**

Conflicts of interest

This project is supported by Shell US.

Acknowledgements

The authors gratefully acknowledge financial support from Shell US. The authors also gratefully acknowledge financial support from the University of Houston Energy Transition Institute. This research was performed using computational resources from the Carya, Opuntia, and Sabine clusters provided by the Research Computing Data Core at the University of Houston.

References

- 1 M. I. Jamesh and X. Sun, J. Power Sources, 2018, 400, 31-68.
- 2 N. Yuan, Q. Jiang, J. Li and J. Tang, *Arabian J. Chem.*, 2020, 13, 4294–4309.
- 3 J. S. Kim, B. Kim, H. Kim and K. Kang, *Adv. Energy Mater.*, 2018, 8, 1–26.
- 4 N. T. Suen, S. F. Hung, Q. Quan, N. Zhang, Y. J. Xu and H. M. Chen, *Chem. Soc. Rev.*, 2017, **46**, 337–365.
- 5 Y. Lee, J. Suntivich, K. J. May, E. E. Perry and Y. Shao-Horn, *J. Phys. Chem. Lett.*, 2012, **3**, 399–404.
- 6 Free Metal Price Tables and Charts, https://www.dailymetalprice.com/.
- 7 Precious Metals Dealer, https://www.apmex.com/.
- 8 Latest and Historical Metal Prices, https://www.metalary.com.
- 9 Z. Wang, Y. R. Zheng, I. Chorkendorff and J. K. Nørskov, *ACS Energy Lett.*, 2020, 5, 2905–2908.
- 10 G. T. K. Gunasooriya and J. K. Nørskov, ACS Energy Lett., 2020, 5, 3778–3787.
- J. Abed, J. Heras-Domingo, M. Luo, R. Sanspeur, W. Alnoush, D. Meira, H.-T. Wang, J. Wang, J. Zhou, D. Zhou, *et al.*, Research Square, 2023, preprint, arXiv:0904.1138, DOI: 10.21203/rs.3.rs-2410178/v1.
- 12 J. Fu, X. Jiang, W. Han and Z. Cao, *Energy Storage Mater.*, 2021, **42**, 252–258.
- 13 S. Kang, Y. Mo, S. P. Ong and G. Ceder, *Nano Lett.*, 2014, 14, 1016–1020.
- 14 W. Sun, Ph.D. thesis, Massachusetts Institute of Technology, 2016.
- 15 A. Navrotsky, ChemPhysChem, 2011, 12, 2207-2215.
- 16 Z. Li, L. E. Achenie and H. Xin, ACS Catal., 2020, 10, 4377– 4384.
- 17 K. K. Rao, Q. K. Do, K. Pham, D. Maiti and L. C. Grabow, *Top. Catal.*, 2020, **63**, 728–741.

- 18 K. Abdelfatah, W. Yang, R. Vijay Solomon, B. Rajbanshi, A. Chowdhury, M. Zare, S. K. Kundu, A. Yonge, A. Heyden and G. Terejanu, *J. Phys. Chem. C*, 2019, **123**, 29804–29810.
- 19 A. Malek, Q. Wang, S. Baumann, O. Guillon, M. Eikerling and K. Malek, *Front. Energy Res.*, 2021, **9**, 1–15.
- 20 S. Back, J. Yoon, N. Tian, W. Zhong, K. Tran and Z. W. Ulissi, J. Phys. Chem. Lett., 2019, 10, 4401–4408.
- 21 L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho, W. Hu, A. Palizhati, A. Sriram, B. Wood, J. Yoon, D. Parikh, C. L. Zitnick and Z. Ulissi, *ACS Catal.*, 2021, **11**, 6059–6072.
- 22 R. Tran, D. Wang, R. Kingsbury, A. Palizhati, K. A. Persson, A. Jain and Z. W. Ulissi, *J. Chem. Phys.*, 2022, **157**, 074102.
- 23 C. C. Price, A. Singh, N. C. Frey and V. B. Shenoy, *Sci. Adv.*, 2022, **8**, 1–12.
- 24 R. Tran, J. Lan, M. Shuaibi, B. M. Wood, S. Goyal, A. Das, J. Heras-Domingo, A. Kolluru, A. Rizvi, N. Shoghi, A. Sriram, F. Therrien, J. Abed, O. Voznyy, E. H. Sargent, Z. Ulissi and C. L. Zitnick, *ACS Catal.*, 2022, **13**, 3066–3084.
- 25 S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson and G. Ceder, *Comput. Mater. Sci.*, 2013, 68, 314–319.
- 26 K. a. Persson, B. Waldwick, P. Lazic and G. Ceder, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2012, **85**, 1–12.
- A. K. Singh, L. Zhou, A. Shinde, S. K. Suram, J. H. Montoya, D. Winston, J. M. Gregoire and K. A. Persson, *Chem. Mater.*, 2017, 29, 10159–10167.
- 28 A. M. Patel, J. K. Nørskov, K. A. Persson and J. H. Montoya, *Phys. Chem. Chem. Phys.*, 2019, 21, 25323–25327.
- 29 A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. SchiØtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, *J. Phys.: Condens. Matter*, 2017, 29, 1– 30.
- 30 A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder and K. A. Persson, *APL Mater.*, 2013, 1, 0110021.
- 31 W. Sun and G. Ceder, Surf. Sci., 2013, 617, 53-59.
- 32 J. H. Montoya and K. A. Persson, *npj Comput. Mater.*, 2017, 3, 14.
- 33 G. Kresse and J. Hafner, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1993, 47, 558-561.
- 34 G. Kresse and J. Hafner, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **49**, 14251–14269.
- 35 G. Kresse and J. Furthmüller, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**, 11169–11186.
- 36 J. Kresse and G. Furthmuller, Phys. Rev. B: Condens. Matter Mater. Phys., 1996, 54, 169.
- 37 P. E. Blöchl, Projector augmented-wave method, 1994.
- 38 J. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, 77, 3865–3868.

- 39 A. Jain, G. Hautier, S. P. Ong, C. J. Moore, C. C. Fischer, K. A. Persson and G. Ceder, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2011, 84, 045115.
- 40 J. Gasteiger, A. Sriram, S. Günnemann and Z. Ulissi, *Transactions on Machine Learning Research*, 2022, pp. 1–22.
- 41 A. A. Peterson, Top. Catal., 2014, 57, 40-53.
- 42 J. Musielewicz, X. Wang, T. Tian and Z. Ulissi, *Mach. Learn.: Sci. Technol.*, 2022, **3**, 03LT01.
- 43 A. Wang, R. Kingsbury, M. McDermott, M. Horton, A. Jain, S. P. Ong, S. Dwaraknath and K. A. Persson, *Sci. Rep.*, 2021, 11, 1–10.
- 44 G. Shi, T. Tano, D. A. Tryk, M. Yamaguchi, A. Iiyama, M. Uchida, K. Iida, C. Arata, S. Watanabe and K. Kakinuma, *ACS Catal.*, 2022, 12, 14209–14219.
- 45 I. C. Man, H.-Y. Su, F. Calle-Vallejo, H. A. Hansen, J. I. Martínez, N. G. Inoglu, J. Kitchin, T. F. Jaramillo, J. K. Nørskov and J. Rossmeisl, *ChemCatChem*, 2011, 3, 1159–1165.
- 46 E. Fabbri, A. Habereder, K. Waltar, R. Kötz and T. J. Schmidt, *Catal.: Sci. Technol.*, 2014, 4, 3800–3821.
- 47 T. Reier, H. N. Nong, D. Teschner, R. Schlögl and P. Strasser, *Adv. Energy Mater.*, 2017, 7, 1–18.
- 48 H. Dau, C. Limberg, T. Reier, M. Risch, S. Roggan and P. Strasser, *ChemCatChem*, 2010, 2, 724–761.
- 49 E. Fabbri and T. J. Schmidt, ACS Catal., 2018, 8, 9765–9774.
- 50 D. González, J. Heras-Domingo, M. Sodupe, L. Rodríguez-Santiago and X. Solans-Monfort, J. Catal., 2021, 396, 192– 201.
- 51 J. Zhang, H. B. Tao, M. Kuang, H. B. Yang, W. Cai, Q. Yan, Q. Mao and B. Liu, ACS Catal., 2020, 10, 8597–8610.
- 52 J. C. Fornaciari, L. C. Weng, S. M. Alia, C. Zhan, T. A. Pham, A. T. Bell, T. Ogitsu, N. Danilovic and A. Z. Weber, *Electrochim. Acta*, 2022, 405, 1–8.
- 53 J. Huang, M. Li, M. J. Eslamibidgoli, M. Eikerling and A. Groß, *JACS Au*, 2021, 1, 1752–1765.
- 54 T. Nishimoto, T. Shinagawa, T. Naito and K. Takanabe, J. Catal., 2020, **391**, 435–445.
- 55 J. T. Mefford, Z. Zhao, M. Bajdich and W. C. Chueh, *Energy Environ. Sci.*, 2020, **13**, 622–634.
- 56 T. Naito, T. Shinagawa, T. Nishimoto and K. Takanabe, Inorg. Chem. Front., 2021, 8, 2900–2917.
- 57 R. Tran, Replication Data for: Rational design of oxide catalysts for OER with OC22, 2023, DOI: 10.18738/T8/APJFTM..
- 58 A. Jain, Z. Wang and J. K. Nørskov, ACS Energy Lett., 2019, 4, 1410–1411.
- 59 R. Y. Sanspeur, J. Heras-Domingo, J. R. Kitchin and Z. Ulissi, *J. Chem. Inf. Model.*, 2023, **63**, 2427–2437.
- 60 M. Aykol, S. S. Dwaraknath, W. Sun and K. A. Persson, *Sci. Adv.*, 2018, 4, 1–8.
- 61 K. S. Joya, Z. Ahmad, Y. F. Joya, A. T. Garcia-Esparza and H. J. De Groot, *Nanoscale*, 2016, 8, 15033–15040.
- 62 D. Simondson, M. Chatti, J. L. Gardiner, B. V. Kerr, D. A. Hoogeveen, P. V. Cherepanov, I. C. Kuschnerus, T. D. Nguyen, B. Johannessen, S. L. Chang, D. R. Macfarlane, R. K. Hocking and A. N. Simonov, ACS Catal., 2022, 12, 12912–12926.

- 63 X. Huang, M. Xie, Y. Chen, Q. Zong, Z. Liu and Y. Jin, RSC Adv., 2015, 5, 26150–26156.
- 64 M. Yu, G. hee Moon, R. G. Castillo, S. DeBeer, C. Weidenthaler and H. Tüysüz, *Angew. Chem., Int. Ed.*, 2020, **59**, 16544–16552.
- 65 X. Zhao, H. Zhang, Y. Yan, J. Cao, X. Li, S. Zhou, Z. Peng and J. Zeng, *Angew. Chem., Int. Ed.*, 2017, **56**, 328–332.
- 66 L. Tian, X. Zhai, X. Wang, J. Li and Z. Li, J. Mater. Chem. A, 2020, 8, 14400–14414.
- 67 Y. Teng, X. D. Wang, J. F. Liao, W. G. Li, H. Y. Chen, Y. J. Dong and D. B. Kuang, *Adv. Funct. Mater.*, 2018, 28, 1–8.
- 68 J. S. Hong, H. Seo, Y. H. Lee, K. H. Cho, C. Ko, S. Park and K. T. Nam, *Small Methods*, 2020, 4, 1–7.
- 69 C. Hu, L. Zhang, Z. J. Zhao, A. Li, X. Chang and J. Gong, *Adv. Mater.*, 2018, **30**, 1–8.
- 70 L. Zhou, Y. Wang, K. Kan, D. M. Lucana, D. Guevarra,
 Y. Lai and J. M. Gregoire, *ACS Sustainable Chem. Eng.*, 2022,
 10, 15898–15908.
- 71 L. Zhou, H. Li, Y. Lai, M. Richter, K. Kan, J. A. Haber, S. Kelly, Z. Wang, Y. Lu, R. S. Kim, X. Li, J. Yano, J. K. Nørskov and J. M. Gregoire, ACS Energy Lett., 2022, 7, 993–1000.
- 72 S. Luke, M. Chatti, A. Yadav, B. V. Kerr, J. Kangsabanik, T. Williams, P. V. Cherepanov, B. Johannessen, A. Tanksale, D. R. MacFarlane, R. K. Hocking, A. Alam, A. Yella and A. N. Simonov, *J. Mater. Chem. A*, 2021, 9, 27468–27484.

- 73 S. Gupta, R. Fernandes, R. Patel, M. Spreitzer and N. Patel, *Appl. Catal.*, A, 2023, 661, 23.
- 74 C. Si, Y. Zhang, C. Zhang, H. Gao, W. Ma, L. Lv and Z. Zhang, *Electrochim. Acta*, 2017, 245, 829–838.
- 75 M. Li, Y. Xiong, X. Liu, X. Bo, Y. Zhang, C. Han and L. Guo, *Nanoscale*, 2015, 7, 8920–8930.
- 76 M. Müllner, M. Riva, F. Kraushofer, M. Schmid,
 G. S. Parkinson, S. F. Mertens and U. Diebold, *J. Phys. Chem. C*, 2019, **123**, 8304–8311.
- 77 S. Hirai, S. Yagi, A. Seno, M. Fujioka, T. Ohno and T. Matsuda, *RSC Adv.*, 2016, 6, 2019–2023.
- 78 M. Al-Mamun, X. Su, H. Zhang, H. Yin, P. Liu, H. Yang,
 D. Wang, Z. Tang, Y. Wang and H. Zhao, *Small*, 2016, 12, 2866–2871.
- 79 Z. Quan, Y. Wang and J. Fang, Acc. Chem. Res., 2013, 46, 191–202.
- 80 Z. Wang, X. Guo, J. Montoya and J. K. Nørskov, *npj Comput. Mater.*, 2020, 6, 1–7.
- 81 Y. Konno, T. Yamamoto and T. Nagayama, *Nanoscale*, 2021, 13, 12738–12749.
- 82 J. Han, M. Zhang, X. Bai, Z. Duan, T. Tang and J. Guan, *Inorg. Chem. Front.*, 2022, 3559–3565.
- 83 L. Sun, Z. Dai, L. Zhong, Y. Zhao, Y. Cheng, S. Chong,
 G. Chen, C. Yan, X. Zhang, H. Tan, L. Zhang, K. N. Dinh,
 S. Li, F. Ma and Q. Yan, *Appl. Catal.*, *B*, 2021, 297, 120477.
- 84 Toxic Metals Overview, https://www.osha.gov/toxic-metals.