Fine-Tuning Dual Single-Atom Metal Sites on Graphene Toward Enhanced Oxygen Reduction Reaction Activity

Evan Xie[1] and Xijun Wang*[,2]

[1]Deerfield Academy, 7 Boyden Ln, Deerfield, MA 01342 USA

[2]Northwestern University, 2145 Sheridan Road, Evanston, IL 60208 USA

*Email: *xijun@northwestern.edu*

Keywords: single-atom catalyst, oxygen reduction reaction, density functional theory, machine learning, catalysis

Abstract

The electrocatalytic oxygen reduction reaction (ORR) plays a crucial role in numerous energy and sustainability systems, such as fuel cells, metal-air batteries, and water electrolysers. It holds significant potential for renewable energy generation, transportation, and storage, heralding a cleaner and more sustainable future. Recent trends have shown increased use of single-atom catalysts (SACs), particularly metal-N_4 moieties grown on graphene-based 2D materials, for enhancing ORR efficiency. However, the rational design of SAC for high-performance ORR faces challenges due to unclear structure-property relationships and the limits of conventional experimental trial-and-error approaches. In this study, we harnessed the power of the density functional theory (DFT) calculations, combined with cutting-edge machine learning (ML) techniques, to explore 144 SACs featuring dual interacting M_1-N_4 and M_2-N_4 moieties (M_1, M_2 = Mn, Fe, Co, Ni, Cu, Ru, Rh, Pd, Ag Ir, Pt, Au), denoted as M_1-M_2, grown on graphene. Of all the catalysts we examined, Fe-Pd emerged as the top performer, achieving an impressive overpotential of 0.211 V vs. RHE in alkaline conditions — outperforming most previously reported SACs. Even more striking, 25 of the evaluated SACs surpassed the renowned Fe-N_4 SAC in catalytic efficiency, including more economically viable alternatives like Fe-Ag. Venturing further, we developed three ML models that accurately predict the overpotentials of various M_1-M_2 SACs, showing their strong ability to capture the relationship between single-atom metal site properties and overpotential. These models provide useful navigation toolkits for the rational design of effective electrocatalysts. Our study sheds light on the path toward achieving efficient SAC-catalyzed ORR, contributing to a more sustainable and energy-efficient future.

Oxygen Reduction Reaction (ORR) is a fundamental electrochemical process that plays a crucial role in numerous energy conversion systems and technologies. It involves the reduction of oxygen molecules to form water or other oxygen-containing species.[1] The significance of ORR lies in its direct connection to critical applications, particularly in fuel cells and metal-air batteries, which are considered promising alternatives to conventional combustion-based power sources due to their environmental benefits and higher energy efficiency.[1-4] For example, in the galvanic battery reaction (**Figure 1a**), ORR occurs at the cathode, where oxygen reacts with electrons and protons to produce water (**Figure 1b**). This reaction is crucial for generating electricity in fuel cells, which have the potential to power various transportation modes and provide clean energy for industrial and residential sectors. Similarly, in metal-air batteries,[5] ORR takes place during the discharge phase, where oxygen is reduced to form metal oxides and release electrical energy.

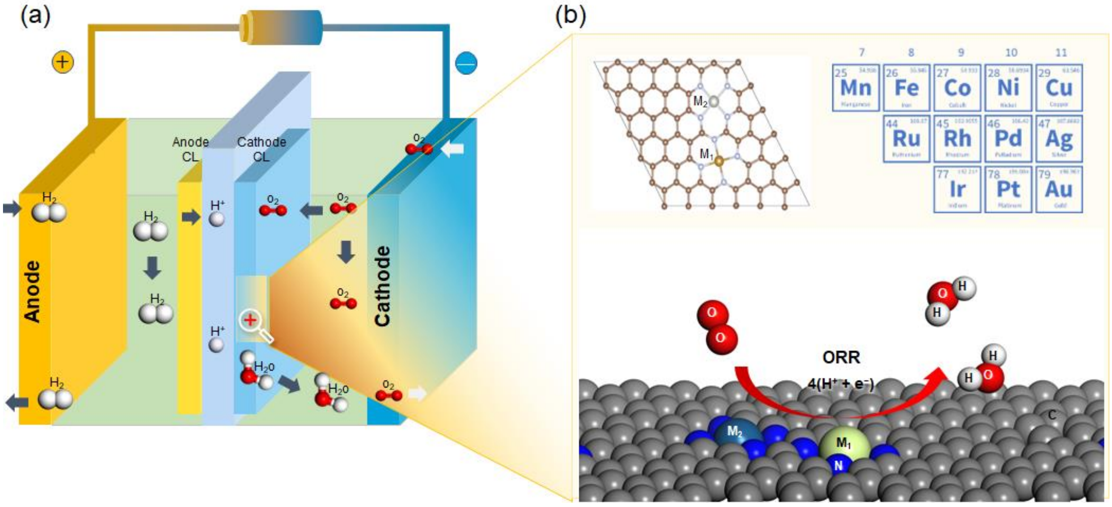


Figure 1. Schematic illustration of (a) the galvanic battery reaction and (b) ORR on M_1-M_2 SACs.

Efficient ORR is paramount for improving the overall performance and longevity of these energy conversion systems. However, the challenge lies in finding cost-effective catalysts that can enhance the kinetics of the ORR process and reduce energy losses. Researchers and engineers continue to explore new materials and design strategies to enhance ORR efficiency. To date, various materials and catalysts have been utilized in this area, including Pt-based electrocatalysts,[6] carbon nanotubes,[7] and transition metals alloys.[8] Among many types of materials, single-atom catalysts (SACs) stand out because of their minimal metal usage, superior efficiency, and enhanced selectivity.[9-11] Recently, Chen, Ji, *et al.* suggested a metal-organic framework (MOF) supported Fe

SAC,[12] which demonstrated excellent stability and catalytic performance with a positive half-wave potential of 0.912 V vs. Reversible Hydrogen Electrode (RHE). Liu *et al.* designed carbon-supported Pt SACs to improve the durability and efficiency of Pt SACs, achieving a positive half-wave potential of 0.87 vs RHE.[13] Rao *et al.* used novel "plasma bombing" strategies to design cheap Co single site SACs for enhanced ORR kinetics, with a Tafel slope of 79 mV dec^{-1}.[14] These pioneering studies highlighted the efficacy and potential of SACs in enhancing ORR.

While researchers have made considerable progress in the development of SACs, the rational design of high-performance SACs for ORR remains a challenging task, primarily due to the limited understanding of the complex structure-property relationships governing their catalytic activity, especially at the atomic level.[15] Conventional experimental trial-and-error approaches often suffer from inefficiency and high costs. The iterative nature of this method necessitates a large number of time-consuming experiments and extensive materials testing, resulting in significant resource expenditure. Moreover, without a thorough comprehension of the underlying mechanisms, the outcomes of these experiments may lack consistency and fail to reach their full potential.

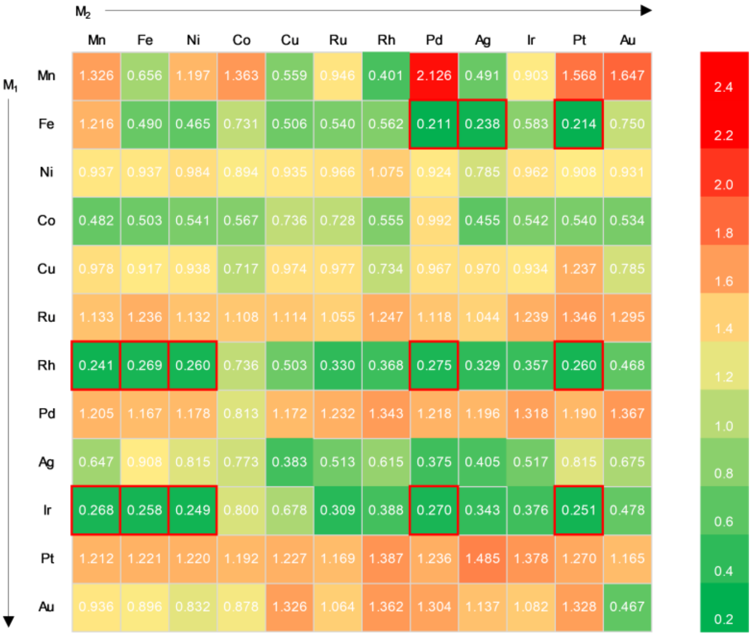
To address these challenges and accelerate SAC design, quantum chemistry methods, especially at the first-principles density functional theory (DFT) level, [16, 17] have been increasingly applied as powerful toolkits to enable a more systematic approach to predict and understand the catalytic behaviors of catalysts.[18,19] These techniques empower researchers to explore a wide range of potential catalyst configurations and identify promising candidates for experimental synthesis, reducing the number of trial-and-error experiments and streamlining the design process. For example, Deng *et al.* adopted DFT to develop Co-based SACs on defective boron nitride (Co/BN) for efficient ORR.[20] Similarly, Yang *et al.* employed DFT to quantify the performance of two-dimensional conjugated aromatic networks (CAN) with a high single-metal-atom-site density that surpassed the performance of a conventional ORR catalyst, Pt/C SAC.[21] Further, Han *et al.* conducted DFT calculations to explore the ORR reaction mechanism and to evaluate the impact of the modulation effect on the ORR performance of Fe-N_4/Pt-N_4 SACs.[22]

In recent years, there has been a notable surge in the utilization of cutting-edge machine learning (ML) techniques for materials design.[23-25] By training predictive models on data from theoretical simulations, researchers have rapidly and accurately screened promising catalytic materials for various applications, including ORR.[26,27] This innovative approach has substantially minimized the requirement for time-consuming DFT calculations and tedious experimental trial-and-error processes. Furthermore, the integration of DFT and ML offers in-depth insights into the intricate structure-property relationships governing catalyst activity, providing essential guidance for the rational design of SACs. This thereby accelerates the development of efficient and sustainable SACs for enhanced ORR, propelling the progress of clean energy technologies and addressing pressing environmental challenges.

In this study, using DFT and ML, we exhaustively investigated the ORR activity of 144 SACs with two M-N_4 moieties, denoted as M_1-M_2, where M_1 and M_2 = Mn, Fe, Ni, Co, Cu, Ru, Rh, Pd, Ag, Ir, Pt, Au (**Figure 1b**). These late transition metals were selected due to their promising potential in catalytic applications, especially ORR.[28-31] For each of the 144 SACs, we assessed their ORR performances by determining their overpotential values, denoted as η, which refers to the extra voltage necessary beyond the theoretical potential to drive the ORR at a desired rate. Through extensive overpotential calculations under alkaline conditions (pH = 13), we identified 13 highly active SACs that outperform a commonly used benchmark system, Fe-N_4 SAC. This system is typically modeled with an axially coordinating *OH specie on Fe, denoted as Fe(OH)-N_4, to more accurately replicate the experimental onset potential.[32-34] We thus adopted the computed overpotential of Fe(OH)-N_4 (0.304 V) as the criterion to identify promising dual single-metal atom compositions for enhanced ORR activity. Notably, we identified the Fe-Pd system as the most active SAC among all those investigated. Its overpotential (0.211 V) is even slightly lower than that of the previously discovered Fe-Pt SAC.[22] Furthermore, we found economical alternatives such as Fe-Ag, Ag-Cu, and Ag-Ag, which exhibit comparable ORR activity with an overpotential of 0.239 V but at a significantly reduced cost. More intriguingly, we harnessed multiple ML methods to investigate the correlation between the intrinsic properties of the single-atom sites and their corresponding overpotentials. These ML models demonstrated high predictive power in estimating the overpotential, thereby effectively predicting promising SAC materials for ORR. These findings illuminate the ORR capabilities of various SAC materials, greatly broadening our understanding of material design and providing theoretical guidance for the future design and optimization of SAC materials, holding implications that transcend the realm of ORR.

This study primarily focuses on the 4-electron pathway, the most prevalent and desired pathway for ORR. This pathway directly produces water, avoiding the production of intermediate hydrogen peroxide seen in the 2-electron pathway. The overpotential was obtained using the analytical model developed by Nørskov[35] (see details in Section S1.1 in Supporting Information). The DFT calculations were performed using the Vienna ab initio Simulation package (VASP) (see details in **Figure S1**–**S2**, and Section S1.2–1.4),[36] with the Perdew-Burke-Ernzerhof (PBE)[37] exchange-correlation functional and the all-electron projector augmented wave (PAW)[38] pseudopotential. To account for the strong on-site coulomb interactions for d-orbital electrons on the transition metal sites, the GGA+U approach[39] was applied, with The U values and initial spin state of each transition metal adopted from a previous benchmark study.[40]

Overpotential has been widely applied as a quantitative metric for assessing the performance of ORR, which is determined based on the binding free energies of the OOH*, O*, and OH* intermediates. **Figure 2** and **Table S1** summarizes the overpotentials of the 144 catalysts investigated in this work (refer to **Figure S3-S38** for the corresponding free energy profiles). In our simulation study, which aims to explore the structure–properties relationships, we are aware of the competition between M_1 and M_2 sites in binding intermediates. However, we chose to focus on the M_1 site when analyzing ORR activity. This choice is supported by two primary reasons. First, our findings indicate that the ORR activity is roughly equal when comparing atom A in the A-B configuration, where atom A is on the M_1 position, to atom A in the B-A configuration, where atom A is on the M_2 position. For example, the DFT-computed overpotential at the Pt atom in Fe-Pt is 1.086 V,[22] which is closely matched by the overpotential at the Pt atom in Pt-Fe (1.221 V)—a minor difference of just 0.135 V. Similarly, the overpotential at the Fe atom in Fe-Pd (0.211 V) is nearly equivalent to that at the Fe atom in Pd-Fe (0.320 V)—a small discrepancy of 0.109 V. These results suggest that a near-symmetrical relationship between the M_1 site of the A-B configuration and the M_2 site of the B-A configuration, thereby diminishing the need to consider the ORR activity of the M_2 site. Secondly, focusing solely on the M_1 site effectively halves the computational load, saving significant CPU time. It is worth noting that for systems where the metal atom at the M_2 site exhibits higher affinity for binding intermediates, insights into the activity of the M_1 site on structure-property relationships can still be gleaned from these less representative models.



	Mn	Fe	Ni	Co	Cu	Ru	Rh	Pd	Ag	Ir	Pt	Au
Mn	1.326	0.656	1.197	1.363	0.559	0.946	0.401	2.126	0.491	0.903	1.568	1.647
Fe	1.216	0.490	0.465	0.731	0.506	0.540	0.562	0.211	0.238	0.583	0.214	0.750
Ni	0.937	0.937	0.984	0.894	0.935	0.966	1.075	0.924	0.785	0.962	0.908	0.931
Co	0.482	0.503	0.541	0.567	0.736	0.728	0.555	0.992	0.455	0.542	0.540	0.534
Cu	0.978	0.917	0.938	0.717	0.974	0.977	0.734	0.967	0.970	0.934	1.237	0.785
Ru	1.133	1.236	1.132	1.108	1.114	1.055	1.247	1.118	1.044	1.239	1.346	1.295
Rh	0.241	0.269	0.260	0.736	0.503	0.330	0.368	0.275	0.329	0.357	0.260	0.468
Pd	1.205	1.167	1.178	0.813	1.172	1.232	1.343	1.218	1.196	1.318	1.190	1.367
Ag	0.647	0.908	0.815	0.773	0.383	0.513	0.615	0.375	0.405	0.517	0.815	0.675
Ir	0.268	0.258	0.249	0.800	0.678	0.309	0.388	0.270	0.343	0.376	0.251	0.478
Pt	1.212	1.221	1.220	1.192	1.227	1.169	1.387	1.236	1.485	1.378	1.270	1.165
Au	0.936	0.896	0.832	0.878	1.326	1.064	1.362	1.304	1.137	1.082	1.328	0.467

Figure 2. Overpotential of the 144 SACs. Green values indicate low overpotential (high activity), Orange/Red values indicate high overpotential (low activity). Promising SACs with η < 0.304 V are highlighted with red squares.

Given this premise, we observed some distinct patterns that emerged for specific M_1 metals. For example, the SACs with M_1 metals Ir, Rh, and Fe tend to exhibit lower overpotential values for most M_2 species, while catalysts involving Pd, Pt, and Ru generally show higher overpotentials. Nevertheless, we also show that the choice of M_2 is not insignificant – it can impact the catalytic behavior of M_1. This is exemplified by the different performances of Fe-Mn and Fe-Pd. Fe-Mn presents a high overpotential of 1.216 V, denoting poor performance, whereas Fe-Pd demonstrates the lowest overpotential of 0.211 V, indicating superior performance. This observation underlines the importance of tuning the M_1-M_2 composition for optimal ORR activity.

In our exhaustive study of the 144 M_1-M_2 combinations, we successfully pinpointed 13 SACs (highlighted with red squares in **Figure 2**) exhibiting lower overpotential than that of the benchmark catalyst, pristine Fe(OH)-N_4 (0.304V, **Figure S39**).[41] This discovery is especially notable given the well-established efficacy of Fe(OH)-N_4 for ORR. Among the promising catalysts identified, Fe-Pt (with an overpotential of 0.214 V) and Fe-Pd (with an overpotential of 0.211 V) have been validated by previous experimental studies.[22,42] This congruence between computational prediction and experimental validation not only bolsters the reliability of our computational methodology but also substantiates the promising potential of these catalyst candidates. Furthermore, our exploration unveiled four M_1-M_2 combinations (Fe-Ag, Ag-Cu, Ag-Pd, Ag-Ag) that are not only close to the overpotential performance of Fe-Pd but are significantly less expensive. This crucial discovery opens the tantalizing possibility of achieving efficient ORR catalysis at a significantly reduced cost. These advancements could significantly accelerate the development and commercialization of cost-effective, high-performance ORR catalysts.

Next, we moved on to delve deeper into the determinants of overpotential, which hinges on the binding free energies of the OOH*, O*, and OH* intermediates. Through a comprehensive assessment of the ΔG_{OOH*}, ΔG_{OH*}, and ΔG_{O*} binding free energies, we successfully delineated the scaling relationships between these quantities, as depicted in [1]-[3]:

$$\Delta G_{OOH*} = 0.493\ \times \Delta G_{O*} + 1.106 \qquad [1]$$

$$\Delta G_{OH*} = 0.557 \times \Delta G_{O*} - 0.353 \qquad [2]$$

$$\Delta G_{OOH*} = 0.826 \times \Delta G_{OH*} + 1.445 \qquad [3]$$

The strong linear relationships (**Figure 3a**, **Figure S40**) suggested that the binding energies of these intermediates on a given metal site tend to be intertwined, which presents a dilemma: a catalyst cannot simultaneously optimize the binding energy for all intermediates, leading to a trade-off where the optimization of the binding energy for one intermediate may result in a suboptimal binding energy for another intermediate. While overcoming this limitation is an ongoing research pursuit beyond the scope of this study, the discerned correlations offer a powerful analytical tool: they allow for the estimation of ΔG_{OOH*}, ΔG_{O*}, and ΔG_{OH*} binding energies directly from each other, significantly streamlining the analysis of the interplay between binding energies and overpotential.

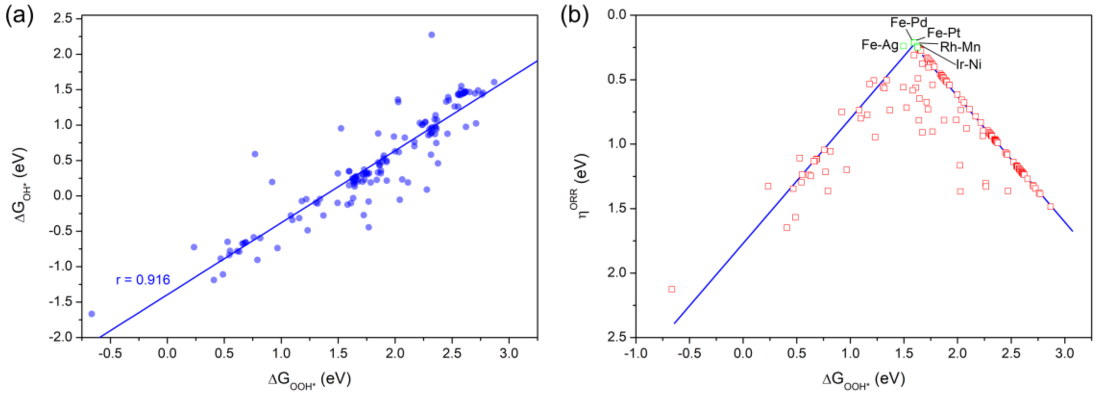
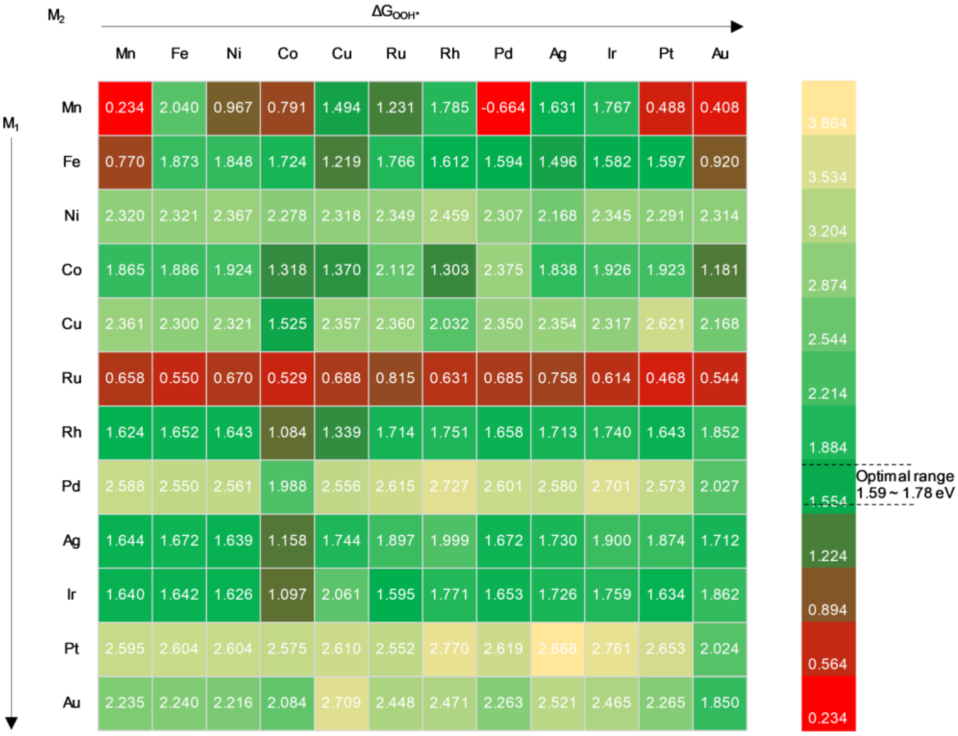


Figure 3. (a) Plot of the relationships between ΔG_{OOH*} and ΔG_{OH*} for the 144 SACs. (b) Overpotential as a function of ΔG_{OOH*} for the 144 SACs. The five best-performing M_1-M_2 SACs are highlighted in green and labeled. The apex of the volcano plot underscores the range of optimal binding energies for ΔG_{OOH*} (~1.59-1.78 eV).

Using one of the three binding free energies – in this case, ΔG_{OOH*} – as the x-axis and plotting the overpotential as the y-axis, we observed a distinctive volcano-shaped trend, also known as the Sabatier principle,[43] shown in **Figure 3b**. The peak of this volcano plot, ranging from approximately 1.59 – 1.78 eV, indicates an optimized region for catalytic activity based on catalysts whose overpotential is affected by ΔG_{OOH*} and has an overpotential ≤ 0.304 V. Catalysts falling within this optimal range of binding energies demonstrate enhanced activity for ORR, prominently displayed by the five best-performing catalysts: Fe-Pd ($\eta = 0.211$ V), Fe-Pt ($\eta = 0.214$ V), Fe-Ag ($\eta = 0.239$ V), Rh-Mn ($\eta = 0.241$ V), and Ir-Ni ($\eta = 0.249$ V). The presence of perfect linear correlations on the right wing of the volcano shape in the ΔG_{OOH*} graph (**Figure 3b**) and on the left wing of the volcano shape in the ΔG_{OH*} graph (**Figure S41a**) signifies that the overpotential-determining step for ORR is primarily contributed either by the first step of *OOH formation or the fourth step of *OH desorption. These findings agree well with previous reports.[44,45] In comparison, the importance of ΔG_{O*} in determining the overpotential is relatively less, as demonstrated in the ΔG_{O*} volcano plot (**Figure S41b**).

Optimizing the ΔG_{OOH*} and/or ΔG_{OH*} values emerge as a compelling strategy to enhance ORR activity. This can be achieved through various strategies, including modifying the coordination environment surrounding the active metal site. For instance, the introduction of ligands, such as an OH group[46] near the active metal site, can alter its electronic structure. This could be achieved by the ligand either donating or withdrawing electrons from the metal center, thereby tuning the metal's electronic properties and enhancing its catalytic activity. Such modifications exert a direct influence on the strength and nature of the interactions between the metal and the adsorbate, thereby shaping the activity of the SACs. This emphasizes the critical significance of precise control over the electronic properties of catalysts, as it empowers the facilitation of efficient and targeted reactions.

Since binding energies play a crucial role in determining the overpotential, we presented them in the form of three distinct heat maps for ΔG_{OOH*} (**Figure 4**), ΔG_{OH*} (**Figure S42a**), and ΔG_{O*} (**Figure S42b**). Most of the Ir-M_2, Fe-M_2, and Rh-M_2 combinations exhibit optimal binding energies that are neither too strong nor too weak. These metal pairings enable sufficiently strong interactions to facilitate the reaction while avoiding overly strong binding that could impede the desorption of reaction intermediates. Conversely, the heat maps indicate that Pd-M_2, Pt-M_2, and Au-M_2 systems generally display relatively weak binding energies with adsorbates. This suggests that the interaction between the active site and their adsorbates may not be strong enough, which could potentially slow down reaction kinetics. On the other hand, the Ru-M_2 combinations show exceedingly strong binding energies with adsorbates. Such intense interactions may hamper the release of adsorbed intermediates, which could result in decreased reactivity and sluggish reaction rates. These findings emphasize the delicate balancing act required in designing efficient SACs for ORR.



	Mn	Fe	Ni	Co	Cu	Ru	Rh	Pd	Ag	Ir	Pt	Au
Mn	0.234	2.040	0.967	0.791	1.494	1.231	1.785	-0.664	1.631	1.767	0.488	0.408
Fe	0.770	1.873	1.848	1.724	1.219	1.766	1.612	1.594	1.496	1.582	1.597	0.920
Ni	2.320	2.321	2.367	2.278	2.318	2.349	2.459	2.307	2.168	2.345	2.291	2.314
Co	1.865	1.886	1.924	1.318	1.370	2.112	1.303	2.375	1.838	1.926	1.923	1.181
Cu	2.361	2.300	2.321	1.525	2.357	2.360	2.032	2.350	2.354	2.317	2.621	2.168
Ru	0.658	0.550	0.670	0.529	0.688	0.815	0.631	0.685	0.758	0.614	0.468	0.544
Rh	1.624	1.652	1.643	1.084	1.339	1.714	1.751	1.658	1.713	1.740	1.643	1.852
Pd	2.588	2.550	2.561	1.988	2.556	2.615	2.727	2.601	2.580	2.701	2.573	2.027
Ag	1.644	1.672	1.639	1.158	1.744	1.897	1.999	1.672	1.730	1.900	1.874	1.712
Ir	1.640	1.642	1.626	1.097	2.061	1.595	1.771	1.653	1.726	1.759	1.634	1.862
Pt	2.595	2.604	2.604	2.575	2.610	2.552	2.770	2.619	2.868	2.761	2.653	2.024
Au	2.235	2.240	2.216	2.084	2.709	2.448	2.471	2.263	2.521	2.465	2.265	1.850

Figure 4. ΔG_{OOH*} for 144 SACs. Optimal binding energies (~1.59 -1.78 eV) are represented in green, overly strong binding energies are marked in red, and excessively weak binding energies are denoted in yellow.

To understand how varying M_1-M_2 combinations result in different binding free energies, we directed our attention toward 12 systems with Fe as the M_1 site (Fe-M_2). Previous work by Rosen *et al.*[47] suggested a strong correlation between the binding energies of a molecule on different metal sites and the group number of the metal elements, which is understandable given that the number of outermost electrons tends to increase as one moves across the periodic table. It is therefore reasonable to infer that an increase in M_2's group number may subtly modify M_1's behavior, leading to a gradual change in the binding affinity of intermediates on M_1. In light of this hypothesis, we plotted ΔG_{OOH*} against M_2's group number (**Figure S43a**). Contrary to expectations, a careful examination of the data from group 7 to group 11 yielded a pattern resembling a volcano plot rather than a clear linear trend. The absence of a linear correlation suggests that different M_2 elements interact uniquely with Fe, beyond simple electron donation or withdrawal mechanisms. Further complicating our understanding, no discernible pattern emerged when comparing the M_2's group number with overpotential (**Figure S43b**). These observations prompt us to look beyond intrinsic properties and delve into the more complex interactions between the M_1 and M_2 sites. More in-depth analysis, such as molecular orbital hybridization analysis, might offer a more nuanced perspective on the complex interactions between M_1 and M_2, and, in turn, shed more light on how these interactions influence the binding of intermediates on M_1 sites.

To further illuminate how binding free energies of the reaction intermediates can influence overpotential, we turned our focus to two extreme cases, Fe-Pd and Mn-Pd, which represent the best and worst-performing SACs, respectively. Their free energy profiles along the 4-electron pathway are shown in **Figure 5a** and **5b**. The Mn-Pd catalyst displays a significantly higher overpotential of 2.126 V compared to the Fe-Pd catalyst (0.211 V), demonstrating its inferior performance. Under $U = 1.23$ V, it is apparent that energy increments at each step are more smoothly transitioned in the Fe-Pd catalyst compared to the Mn-Pd catalyst. For the latter, the energy steps on the three binding configurations are significantly more stable than the former, especially for the OH* desorption step. In fact, Mn-Pd showcases the smallest ΔG_{OH*} value (-1.67 eV) amongst all the 144 systems studied, leading to a prohibitive barrier in the OH* desorption step, leading to a notable decrease in its activity. Conversely, the Fe-Pd catalyst exhibits an optimal ΔG_{OH*} value (0.35 eV), corresponding to an enhanced ORR activity.

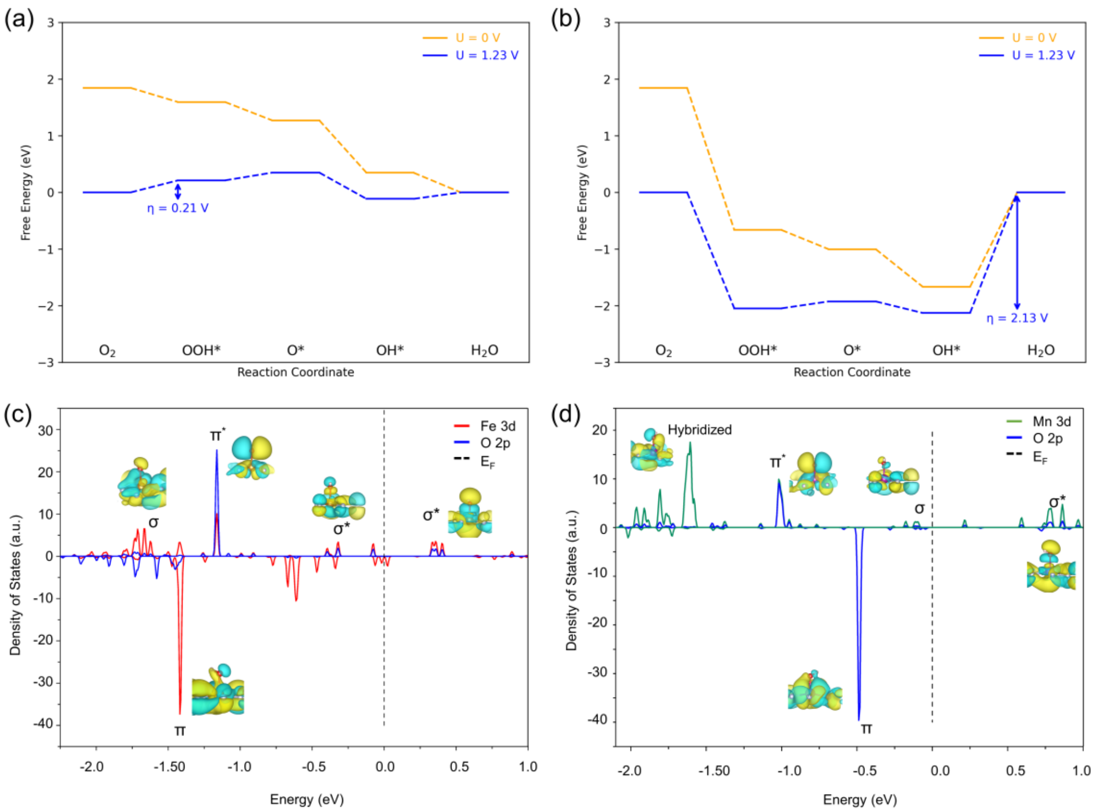


Figure 5. Comparison of the free energy profiles of (a) the best-performing SAC (Fe-Pd) and (b) the worst-performing SAC (Mn-Pd). Computed PDOS for (c) Fe-Pd and (d) Mn-Pd, after the binding of OH* binding. The bonding molecular orbitals, denoted as σ and π, correspond to bonding interactions between the d-orbitals of Fe or Mn and the p-orbitals of O. σ* and π* denote the antibonding molecular orbitals arising from interactions between the d orbitals of Fe or Mn and p orbitals of O.

To rationalize the huge difference in ΔG_{OH*} between Fe-Pd and Mn-Pd, we computed the projected density of states (PDOS) for the d-orbitals of M_1 sites (Fe and Mn) and p-orbitals of O after OH* binding in both systems. Specifically, we visualized the spatial distribution of molecular orbitals for the electronic states with high density or those close to the Fermi level (E_F). These electronic states are depicted in **Figure 5c** and **5d**, as they are believed to play a critical role in determining the bonding strength of M_1–OH*. In the case of OH* binding on Fe-Pd (**Figure 5c**), the frontier orbitals for Fe-O bonding are primarily the π* and σ* antibonding orbitals, leading to relatively weak Fe-OH bonding. In contrast, the orbital energy levels and spatial distributions are remarkably different in the Mn-Pd system (**Figure 5d**). Near the Fermi level, we observed σ and π bonding orbitals, which considerably strengthen the Mn-O bonding. This leads to a significantly more negative ΔG_{OH*} value of -1.67 eV, which in turn results in a much higher overpotential of 2.126 V (**Figure 5c**). This detailed analysis of molecular orbital hybridization provides further insight into how different M_1-M_2 combinations can dramatically affect the binding free energies and the overall ORR activity.

Similarly, we also emphasized the advantages of tailoring the interactions between dual metal-N_4 moieties for optimizing the binding strength toward intermediates, in contrast to a single metal-N_4 moiety. As evidence, we studied the free energy profile for ORR on single Fe-N_4 moiety and compared it to Fe-Pd, as shown in **Figures S44a** and **b**. We observed a minor reduction in the free energy rise for the OOH* binding step in Fe-Pd. This suggests that OOH* binds more strongly to the Fe site in Fe-Pd than in Fe-N_4. This behavior can be explained by the diminishing of frontier antibonding orbitals and an increase in bonding orbitals between Fe and O in the Fe-Pd system, as depicted in **Figures S44c** and **d**. Similar effects on molecular orbitals due to neighboring metal-N_4 sites have been noted in recent reports.[22,48-50]

While molecular orbital analysis provides qualitative insights into the bonding strength between metal and intermediate reactants, quantitatively predicting overpotential from the readily available intrinsic atomic properties or easy-to-compute electronic properties of single-atom catalysts (SACs) remains a formidable task. This necessitates an in-depth, quantitative grasp of the structure-property relationship, a challenge intensified by the intricate interplay among, in our context, M_1-M_2, and M_1-intermediates. Although DFT is invaluable in analyzing these relationships, its usage is constrained by its time-consuming nature. Consequently, we turn to employing data-driven methodologies like machine learning (ML), as a potentially efficient and accurate alternative for overpotential prediction. As illustrated in **Figure 6a**, we considered 16 intrinsic or electronic features, including inherent properties such as electronegativity (χ), atomic radius (r_a), atomic mass (m_a), ionization energy (E_I) of M_1 and M_2 metal atoms (denoted by subscripts $_1$ and $_2$), alongside DFT-derived electronic properties like spin-up and spin-down d-band centers ($\varepsilon_{d\uparrow}$ and $\varepsilon_{d\downarrow}$), partial charge (q), spin density (ρ) on M_1 and M_2, and the HOMO-LUMO gap (E_g) of the M_1-M_2 systems. We employed two types of ML algorithms - Sure Independence Screening and Sparsity Optimization (SISSO)[51] for feature space dimensional reduction (see details in Section S1.5), and the Extra Trees Regressor (ETR)[52], a decision-tree-based regression approach (see details in Section S1.6). These algorithms were trained using 80% of the dataset comprising of 144 SACs and tested on the remaining 20% of the data.

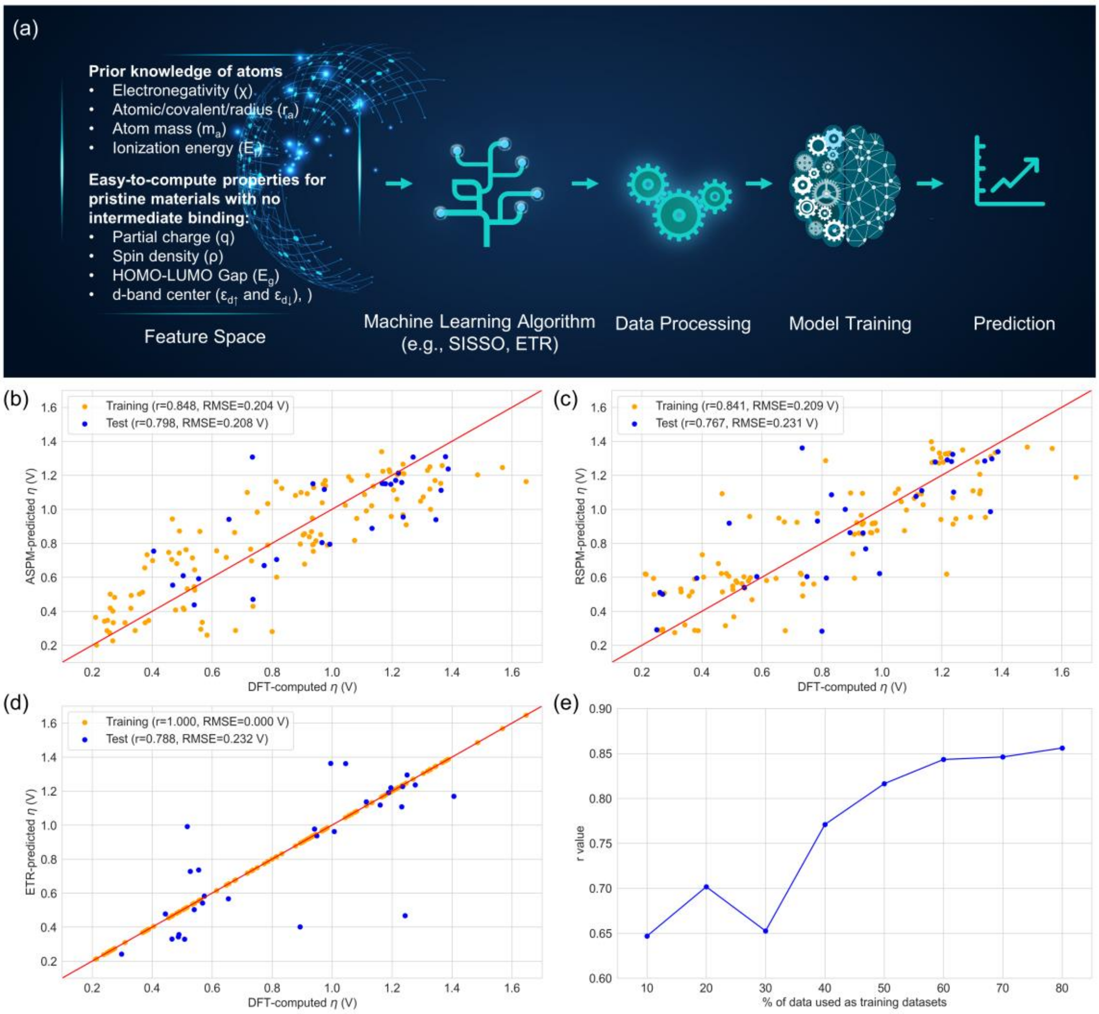


Figure 6. (a) Workflow of the ML process for predicting overpotential from intrinsic atomic properties and DFT-computed electronic properties. Prediction performance of the (b) ASPM SISSO, (c) RSPM SISSO, and (d) ETR models for the training (80%) and testing (20%) datasets. (e) Illustration of the incremental improvement in the predictive ability of the ETR model with increasing size of training datasets.

Firstly, we employed the SISSO method[51] to bridge the input features – namely, the intrinsic atomic properties and DFT-computed electronic properties – and the target property (ORR overpotential). Using all the above-mentioned features, we name the resultant model as the all-features-trained SISSO prediction model (ASPM), with the best-performing mathematical expression (eq. [4]) shown below:

$$\eta_{ASPM-Predicted} = -0.381 \times \left[\cos\left(\frac{m_{a1}}{\chi_1}\right)\right] - 0.238 \times cos(E_{I2} \times q_1) + 0.199 \times \left[\frac{\chi_1 - \chi_2}{\varepsilon_{d\downarrow 1}}\right] + 0.886 \qquad [4]$$

This equation demonstrated statistically significant accuracy in predicting overpotential, with a significant correlation coefficient (*r*) of 0.848 for the training datasets (**Figure 6b**). The root mean square error (RMSE) value of 0.204 V falls within the computational error of DFT, signifying a strong agreement between SISSO-predicted and DFT-computed values. When predicting for the testing datasets, the performance slightly deteriorated (r = 0.798, RMSE = 0.208 V, **Figure 6b**), hinting at possible overfitting. This may be attributed to including too many unrelated or interrelated features or a lack of sufficient datasets (in this instance, 144). Despite comprehensive attempts to improve the prediction accuracy of the SISSO model — such as incorporating more informative features (refer to Section S2.1 for more details) and exploring alternative approaches to overpotential prediction (refer to Section S2.2 for more details) — we failed to achieve significant improvement. This outcome underscores the complexity of the problem and invites further investigation into novel predictive methodologies, particularly in the approach to minimize the RMSE value. Nevertheless, the results emphasize the relative precision of the descriptors generated by SISSO in capturing the quantitative relationship between features and overpotential.

More interestingly, upon closer inspection, eq. [4] is mainly comprised of M_1 properties, which is expected since M_1 is the site directly bonding with the reaction intermediates and should, therefore, govern the overpotential. The equation includes three key terms. The first term, $\cos(\frac{m_{a1}}{\chi_1})$, relates to M_1's intrinsic properties, i.e., mass and electronegativity, which naturally influence the catalytic properties of the M_1 site. The second term, $cos(E_{I2} \times q_1)$, captures the ionization energy of M_2 atom and the M_1 site's partial charge. This offers descriptions of M_2's electron donation capability and M_1's atomic charge, hinting at charge exchange dynamics between M_1 and M_2. The third term's numerator, $\chi_1 - \chi_2$, is the electronegativity difference between M_1 and M_2, also conveying the information of charge transfer between the metal sites, while the denominator, $\varepsilon_{d\downarrow 1}$, carries molecular orbital information of M_1. These terms validate our hypothesis that the catalytic properties of M_1-M_2 are impacted by the interactions between the M_1-N_4 and M_2-N_4 sites. This happens even though the M_1-N_4 and M_2-N_4 sites are not directly connected, and extends beyond simple electron donation or withdrawal mechanisms. Importantly, these complex interactions can be quantified using the SISSO-generated descriptors.

To mitigate the overfitting resulting from the small data problem, it is potentially beneficial to evaluate the relevance and significance of each feature. Less important features may introduce noise or bias, potentially leading to inaccurate predictions. Consequently, we refined the initial sixteen features used for training the model. The importance analysis for each feature was performed using the ETR algorithm, a commonly utilized decision tree-based ML method. This allowed us to assess the relative importance of each feature in determining the overpotential (**Table 1**). The refined features were selected based on a mean importance cutoff of 0.05, ensuring only the most informative and impactful features were retained. The refined feature set included the ionization energy of M_1 (E_{I1}), d-band centers ($\varepsilon_{d\uparrow}$ and $\varepsilon_{d\downarrow}$), partial charge of M_1 (q_1), electronegativity of M_1 (χ_1), and atomic mass of M_1 (m_{a1}). Using these features, another SISSO model, namely the refined-feature-trained SISSO prediction model (RSPM), was produced, with the best-performing expression as follows:

$$\eta_{RSPM-Predicted} = -0.322 \times \left[\cos\left(\frac{m_{a1}}{\chi_1}\right)\right] + 0.165 \times cos(\varepsilon_{d\uparrow 1} + \varepsilon_{d\downarrow 1}) - 0.199 \times cos(E_{I1} \times q_1) + 0.878 \quad [5]$$

The prediction performance of RSPM is comparable to that of ASPM, achieving r = 0.841 and RMSE = 0.209 V on training datasets, and r = 0.768 and RMSE = 0.231 V on test datasets, as illustrated in **Figure 6c**. This suggests that the dimensionality reduction feature of SISSO could have eliminated less important features in ASPM, leading RSPM to exhibit almost equivalent predictive accuracy as ASPM.

Table 1. Evaluation of the mean importance for sixteen input features using the ETR Algorithm. Features with mean importance greater than 0.05 (bolded) were selected for the RSPM.

Features	Feature Mean Importance
M_1-Partial-charge (q_1)	0.404
M_1- Ionization (E_{I1})	0.214
M_1- Electronegativity (χ_1)	0.152
Band-Up ($\varepsilon_{d\uparrow}$)	0.110
M_1- Radius (r_{a1})	0.061
Band-Down ($\varepsilon_{d\downarrow}$)	0.051
M_1-Spin-density (ρ_1)	0.025
Gap Up ($E_{g\uparrow}$)	0.024
M_1-Mass (m_{a1})	0.023
M_2-Mass (m_{a2})	0.019
M_2-Electronegativity (χ_2)	0.016

Gap Down ($E_{g\downarrow}$)	0.014
M_2-Radius (r_{a2})	0.012
M_2- Partial-charge (q_2)	0.012
M_2-Spin-density (ρ_2)	0.006
M_2-Ionization (E_{I2})	0.004

Subsequently, we utilized the ETR algorithm to predict overpotential based on the 16 features. A significant disparity was noted when comparing the performance of the ETR-based model on the training data (80%) and the testing data (20%). The model performed perfectly on the training data, showing an ideal correlation coefficient of $r = 1.000$ and an RMSE value of 0.000 V (**Figure 6d**). However, its performance declined when applied to the testing data, producing a correlation coefficient of $r = 0.788$ and an RMSE value of 0.233 V (**Figure 6d**). This divergence is reflective of overfitting, a classic challenge encountered with smaller datasets, where the model could excessively adapt to the training data at the expense of generalization. We see a clear path for the future advancement of our ETR model. We can expect an improvement in the model prediction power with an increase in the size of the training datasets. **Figure 6e** depicts that as the training dataset size grows, there is a marked increase in prediction accuracy. This suggests that expanding the training data could be a fruitful direction for the continued refinement and evolution of our model.

Table 2 outlines the performance of the above-mentioned three ML models – the ASPM and RSPM SISSO models, and the ETR model. These models exhibit sufficient performance, demonstrating the utility and effectiveness of ML in material design. Particularly, we applied these three models to predict promising M_1-M_2 combinations, using pristine Fe(OH)-N_4 ($\eta = 0.304$ V) as a benchmark. As detailed in **Table 3** and **Table 4**, the ASPM, RSPM, and ETR models predicted 10, 10, and 13 promising materials, respectively. DFT validations of these predictions highlighted 5, 5, and 13 successful matches, translating to success rates of 50%, 50%, and a stellar 100%, respectively, demonstrating the robust predictive capabilities of these models.

Notably, considering the potential error margin in DFT calculations, which usually falls between ±0.1[53] and ±0.5 V.[54] Using ±0.1 V as an illustrative point, the benchmark overpotential for

Fe(OH)-N_4, originally at 0.304 V, is adjusted to 0.404 V. When factoring in this adjustment, the accuracy of the ML models in predicting promising M_1-M_2 compositions improves substantially. Specifically, the success rates rise to 74% for ASPM and 69% for RSPM, as detailed in **Table 4** (see Section S2.3 and **Table S4** for more details), underscoring the effectiveness of the ML models in reducing time-consuming DFT calculations and costly experimental trial-and-error.

These findings affirm that while ML models may not be flawless, they can significantly alleviate the burden of time-consuming DFT computations and expensive experimental procedures. The enhanced performance of the RSPM model in predicting top catalysts underscores the potential of refining such models to improve prediction accuracy, thus driving the field of materials science toward more efficient and precise discovery and development.

Table 2. Performance comparison of ASPM, RSPM, and ETR in predicting overpotentials for the 144 M_1-M_2 SACs.

Models	Coefficient (r)	r^2	RMSE	MAE
ASPM (Training)	0.848	0.720	0.204	0.583
ASPM (Testing)	0.798	0.637	0.208	0.573
RSPM (Training)	0.841	0.707	0.209	0.783
RSPM (Testing)	0.768	0.589	0.231	0.627
ETR (Training)	1.000	1.000	0.000	0.000
ETR (Testing)	0.788	0.621	0.232	0.776

Table 3. Comparison between DFT-computed overpotentials and those predicted by ASPM, RSPM, and ETR for the top 13 SACs against the criterion overpotential of 0.304 V. * denotes the data points sourced from the test dataset for ETR.

M_1-M_2	DFT-computed η	ASPM Predicted η	RSPM Predicted η	ETR Predicted η
Fe-Pd	0.211			0.211
Fe-Pt	0.214	0.202		0.214
Fe-Ag	0.238	-0.017		0.238
Rh-Mn	0.241			0.298*
Ir-Ni	0.249	0.286	0.292	0.249
Ir-Pt	0.251		0.289	0.251

Ir-Fe	0.258	0.274	0.290	0.258
Rh-Pt	0.260			0.260
Rh-Ni	0.260			0.260
Ir-Mn	0.268	0.227	0.288	0.268
Rh-Fe	0.269			0.269
Ir-Pd	0.270		0.295	0.270
Rh-Pd	0.275			0.275

Table 4. Performance comparison of ASPM, RSPM, and ETR in predicting promising SACs. SACs are deemed promising when their overpotentials fall below the specified criterion. The first three rows display results without an error bar, while rows 4–6 present results with an error bar of 0.1 V in the criterion of overpotential.

	Criterion of overpotential (V)	**Total # of good catalysts predicted** (A)	**# of good catalysts confirmed by DFT** (B)	**% of good catalysts confirmed by DFT** (C) = (B)/(A)
ASPM	0.304	10	5	50%
RSPM	0.304	10	5	50%
ETR	0.304	13	13	100%
ASPM	0.404	19	14	74%
RSPM	0.404	13	9	69%
ETR	0.404	19	19	100%

Furthermore, we delved into the performance of the ML models in predicting overpotential for SACs containing metal sites that have never been learned, termed as ‘extrapolation’. While extrapolation allows predictions beyond known data, it is inherently challenging in ML due to the assumptions made about unencountered behavior. It is contrasted with ‘interpolation’, where predictions occur within known data boundaries, as highlighted previously. In our tests on the M_1-Zn and Re-M_2 systems, the extrapolation predictions were subpar, further detailed in Section 2.4 in the Supporting Information. This difficulty predominantly stems from our limited training data. To illustrate, using the ASPM model as an example, its predictability improves similarly to the ETR model (**Figure 6e**) as the dataset grows (**Figure S45**). The correlation coefficient r stabilizes when the data reaches a certain volume. Such behavior suggests that the SISSO algorithm excels in interpolation when given sufficient training data. It is thus reasonable to expect that enriching the training data will consistently boost the predictive power of ML models. Hence, we aspire for this paper to act as a spark, encouraging further ventures into data augmentation, thus progressively improving the predictive efficacy of ML models and, in the long run, encompassing the majority of metals in the periodic table.

In summary, we applied a synergistic approach of DFT calculations and ML techniques to explore the ORR performance of 144 SACs. We discovered 13 materials that outperformed the pristine Fe(OH)-N_4 benchmark (η = 0.304 V), offering superior catalytic activity. Of particular interest, the Fe-Pd, Fe-Pt, and Fe-Ag catalysts emerged as top performers, delivering impressive overpotentials of 0.211 V, 0.214 V, and 0.238 V respectively, under alkaline conditions (pH = 13). This surpasses the performance of most, if not all, previously reported materials, and sets a promising precedent for the field. Especially noteworthy is the economic feasibility of Fe-Ag, given the abundant availability and cost-effectiveness of both iron (Fe) and silver (Ag). Additional catalysts, such as Ag-Cu, Ag-Pd, and Ag-Ag, showed promising performance alongside economic efficiency.

Our comprehensive analysis of adsorbate binding energies helped us pinpoint an optimal range for ΔG_{OOH*} (~1.59-1.78 eV) and ΔG_{OH*} (~0.15-0.37 eV). This range correlates with high catalytic performance, thereby providing valuable insight for future experimental optimization. Therefore, the potential to enhance overall catalyst performance through structural modifications that optimize active site binding energies is notable. To further delve into these insights, we conducted an orbital analysis on our best performing catalyst (Fe-Pd) and the least effective one (Mn-Pd). The analysis highlighted that the ideal binding strength of reaction intermediates arises from a balanced interaction between the bonding and anti-bonding orbitals in the frontier orbitals.

We trained ML models using a suite of atomic and electronic properties pertaining to the single-atom metal sites as input features to predict the overpotential of the SACs. Impressively, these models demonstrated satisfying predictive accuracy: 50% of the SACs identified as top performers by the SISSO models were validated through DFT calculations. When accounting for possible DFT errors, the success rate could climb to 74%. These results highlighted the potential of ML as a tool to reduce the need for costly DFT computations and to streamline experimental trial-and-error processes.

Collectively, our research lays a firm foundation for the intelligent design and discovery of SACs. We have not only illuminated the possibilities for a more sustainable and eco-friendly future but also set a foundation for further exploration of ORR performance and the application of ML in material design. This opens up the possibility for the development of highly efficient, cost-effective catalysts that could revolutionize energy conversion and storage applications, extending far beyond catalysis.

Associated Content

Supporting Information

Computational details; schematic illustration of structural model construction and DFT workflow; free energy diagrams of 144 investigated SACs and the benchmark Fe(OH)-N_4; DFT-computed data on catalytic properties of 144 SACs; plot of the relationships between ΔG_{OOH^*} vs ΔG_{OH^*} and ΔG_{OH^*} vs ΔG_{O^*} for 144 SACs; ΔG_{OH^*} and ΔG_{O^*} volcano plots for 144 SACs; ΔG_{OOH^*} and ΔG_{O^*} heat maps for 144 SACs; Correlation between group number of M_2 and ΔG_{OOH^*} for Fe-M system; Correlation between group number of M_2 and overpotential for Fe-M system; Evaluation of prediction performance of electronic descriptor ψ; Performance of SISSO model using binding energies (ΔG_{OOH^*}, ΔG_{O^*}, and ΔG_{OH^*}) to predict overpotential; impact of DFT error bars in SACs prediction; Extrapolation Capability of ASPM.

Acknowledgment

The numerical calculations have been carried out on the supercomputing system provided by Beijing Paratera Co., Ltd. We thank Prof. Xiyu Li at Songshan Lake Materials Laboratory and Prof. Li Yang at Anhui University for their valuable methodological and software support.

References

(1) Ge, X.; Sumboja, A.; Wuu, D.; An, T.; Li, B.; Goh, F. W. T.; Hor, T. S. A.; Zong, Y.; Liu, Z. Oxygen Reduction in Alkaline Media: From Mechanisms to Recent Advances of Catalysts. *ACS Catal.* **2015**, *5* (8), 4643–4667.

(2) Katsounaros, I.; Cherevko, S.; Zeradjanin, A. R.; Mayrhofer, K. J. J. Oxygen Electrochemistry as a Cornerstone for Sustainable Energy Conversion. *Angew. Chem., Int. Ed.* **2013**, *53* (1), 102–121.

(3) Zaman, S.; Huang, L.; Abdoulkader Ibro Douka; Yang, H.; You, B.; Bao Yu Xia. Oxygen Reduction Electrocatalysts toward Practical Fuel Cells: Progress and Perspectives. *Angew. Chem., Int. Ed.* **2021**, *60* (33), 17832–17852.

(4) Liu, M.; Xiao, X.; Li, Q.; Luo, L.; Ding, M.; Zhang, B.; Li, Y.; Zou, J.; Jiang, B. Recent Progress of Electrocatalysts for Oxygen Reduction in Fuel Cells. *J. Colloid Interface Sci.* **2022**, *607*, 791–815.

(5) Ma, R.; Lin, G.; Zhou, Y.; Liu, Q.; Zhang, T.; Shan, G.; Yang, M.; Wang, J. A Review of Oxygen Reduction Mechanisms for Metal-Free Carbon-Based Electrocatalysts. *Npj Comput. Mater.* **2019**, *5* (1).

(6) Ma, Z.; Cano, Z. P.; Yu, A.; Chen, Z.; Jiang, G.; Fu, X.; Yang, L.; Wu, T.; Bai, Z.; Lu, J. Enhancing Oxygen Reduction Activity of Pt-Based Electrocatalysts: From Theoretical Mechanisms to Practical Methods. *Angew. Chem.* **2020**, *132* (42), 18490–18504.

(7) Lu, Y.; Li, X.; Compton, R. G. Oxygen Reduction Reaction at Single Entity Multiwalled Carbon Nanotubes. *J. Phys. Chem. Lett.* **2022**, *13* (16), 3748–3753.

(8) Wu, Z.-P.; Caracciolo, D. T.; Maswadeh, Y.; Wen, J.; Kong, Z.; Shan, S.; Vargas, J. A.; Yan, S.; Hopkins, E.; Park, K.; Sharma, A.; Ren, Y.; Petkov, V.; Wang, L.; Zhong, C.-J. Alloying–Realloying Enabled High Durability for Pt–Pd-3d-Transition Metal Nanoparticle Fuel Cell Catalysts. *Nat. Commun.* **2021**, *12* (1), 859.

(9) Li, L.; Chang, X.; Lin, X.; Zhao, Z.-J.; Gong, J. Theoretical Insights into Single-Atom Catalysts. *Chem. Soc. Rev.* **2020**, *49* (22), 8156–8178.

(10) Guo, J.; Liu, H.; Li, D.; Wang, J.; Djitcheu, X.; He, D.; Zhang, Q. A Minireview on the Synthesis of Single Atom Catalysts. *RSC Adv.* **2022**, *12* (15), 9373–9394.

(11) Wang, A.; Li, J.; Zhang, T. Heterogeneous Single-Atom Catalysis. *Nat. Rev. Chem.* **2018**, *2* (6), 65–81.

(12) Chen, Y.; Ji, S.; Zhao, S.; Chen, W.; Dong, J.; Cheong, W.-C.; Shen, R.; Wen, X.; Zheng, L.; Rykov, A. I.; Cai, S.; Tang, H.; Zhuang, Z.; Chen, C.; Peng, Q.; Wang, D.; Li, Y. Enhanced Oxygen Reduction with Single-Atomic-Site Iron Catalysts for a Zinc-Air Battery and Hydrogen-Air Fuel Cell. *Nat. Commun.* **2018**, *9* (1).

(13) Liu, J.; Jiao, M.; Lu, L.; Barkholtz, H. M.; Li, Y.; Wang, Y.; Jiang, L.; Wang, Y.; Liu, D.-J.; Zhuang, L.; Ma, C.; Zeng, J.; Zhang, B.; Dang Sheng Su; Zou, M.-H.; Xing, W.; Xu, W.; Jiang, Z.; Sun, G. High Performance Platinum Single Atom Electrocatalyst for Oxygen Reduction Reaction. *Nat. Commun.* **2017**, *8* (1).

(14) Rao, P.; Wu, D.; Wang, T.-J.; Li, J.; Deng, P.; Chen, Q.; Shen, Y.; Chen, Y.; Tian, X. Single Atomic Cobalt Electrocatalyst for Efficient Oxygen Reduction Reaction. *eScience* **2022**, *2* (4), 399–404.

(15) Li, J.; Chen, C.; Xu, L.; Zhang, Y.; Wei, W.; Zhao, E.; Wu, Y.; Chen, C. Challenges and Perspectives of Single-Atom-Based Catalysts for Electrochemical Reactions. *J. Am. Chem. Soc.* **2023**, *3* (3), 736–755.

(16) Jain, A.; Shin, Y.; Persson, K. A. Computational Predictions of Energy Materials Using Density Functional Theory. *Nat. Rev. Mater.* **2016**, *1* (1), 1–13.

(17) Fu, C.; Liu, C.; Li, T.; Zhang, X.; Wang, F.; Yang, J.; Jiang, Y.; Cui, P.; Li, H. DFT Calculations: A Powerful Tool for Better Understanding of Electrocatalytic Oxygen Reduction Reactions on Pt-Based Metallic Catalysts. *Comput. Mater. Sci.* **2019**, *170*, 109202.

(18) Makkar, P.; Ghosh, N. N. A Review on the Use of DFT for the Prediction of the Properties of Nanomaterials. *RSC Adv.* **2021**, *11* (45), 27897–27924.

(19) Florian Kraushofer; Parkinson, G. S. Single-Atom Catalysis: Insights from Model Systems. *Chem. Rev.* **2022**, *122* (18), 14911–14939.

(20) Deng, C.; He, R.; Shen, W.; Li, M.; Zhang, T. A Single-Atom Catalyst of Cobalt Supported on a Defective Two-Dimensional Boron Nitride Material as a Promising Electrocatalyst for

the Oxygen Reduction Reaction: A DFT Study. *Phys. Chem. Chem. Phys.* **2019**, *21* (13), 6900–6907.

(21) Yang, S.; Yu, Y.; Dou, M.; Zhang, Z.; Dai, L.; Wang, F. Two-Dimensional Conjugated Aromatic Networks as High-Site-Density and Single-Atom Electrocatalysts for the Oxygen Reduction Reaction. *Angew. Chem., Int. Ed.* **2019**, *58* (41), 14724–14730.

(22) Han, A.; Wang, X.; Tang, K.; Zhang, Z.; Ye, C.; Kong, K.; Hu, H.; Zheng, L.; Jiang, P.; Zhao, C.; Zhang, Q.; Wang, D.; Li, Y. An Adjacent Atomic Platinum Site Enables Single-Atom Iron with High Oxygen Reduction Reaction Performance. *Angew. Chem., Int. Ed.* **2021**, *60* (35), 19262–19271.

(23) Choudhary, K.; DeCost, B.; Chen, C.; Jain, A.; Tavazza, F.; Cohn, R.; Park, C. W.; Choudhary, A.; Agrawal, A.; Billinge, S. J. L.; Holm, E.; Ong, S. P.; Wolverton, C. Recent Advances and Applications of Deep Learning Methods in Materials Science. *Npj Comput. Mater.* **2022**, *8* (1).

(24) Xu, P.; Ji, X.; Li, M.; Lu, W. Small Data Machine Learning in Materials Science. *Npj Comput. Mater.* **2023**, *9* (1).

(25) Wang, S.; Jiang, J. Interpretable Catalysis Models Using Machine Learning with Spectroscopic Descriptors. *ACS Catal.* **2023**, *13* (11), 7428–7436.

(26) Zhang, X.; Liu, J.; Li, R.; Jian, X.; Gao, X.; Lu, Z.; Yue, X. Machine Learning Screening of High-Performance Single-Atom Electrocatalysts for Two-Electron Oxygen Reduction Reaction. *J. Colloid Interface Sci.* **2023**, *645*, 956–963.

(27) Ghanekar, P. G.; Deshpande, S.; Greeley, J. Adsorbate Chemical Environment-Based Machine Learning Framework for Heterogeneous Catalysis. *Nat. Commun.* **2022**, *13* (1).

(28) Xie, X.; Peng, L.; Yang, H.; Waterhouse, G. I. N.; Shang, L.; Zhang, T. MIL-101-Derived Mesoporous Carbon Supporting Highly Exposed Fe Single-Atom Sites as Efficient Oxygen Reduction Reaction Catalysts. *Adv. Mater.* **2021**, *33* (23), 2101038.

(29) Shang, H.; Zhou, X.; Dong, J.; Li, A.; Zhao, X.; Liu, Q.; Lin, Y.; Pei, J.; Li, Z.; Jiang, Z.; Zhou, D.; Zheng, L.; Wang, Y.; Zhou, J.; Yang, Z.; Cao, R.; Sarangi, R.; Sun, T.; Yang, X.; Zheng, X. Engineering Unsymmetrically Coordinated Cu-S1N3 Single Atom Sites with Enhanced Oxygen Reduction Activity. *Nat. Commun.* **2020**, *11* (1).

(30) Xiao, M.; Zhu, J.; Li, G.; Li, N.; Li, S.; Cano, Z. P.; Ma, L.; Cui, P.; Xu, P.; Jiang, G.; Jin, H.; Wang, S.; Wu, T.; Lu, J.; Yu, A.; Su, D.; Chen, Z. A Single-Atom Iridium Heterogeneous Catalyst in Oxygen Reduction Reaction. *Angew. Chem., Int. Ed.* **2019**, *58* (28), 9640–9645.

(31) Jeong, H.; Shin, D.; Kim, B.; Bae, J.; Shin, S.; Choe, C.; Han, J. W.; Lee, H. Controlling the Oxidation State of Pt Single Atoms for Maximizing Catalytic Activity. *Angew. Chem.* **2020**, *132* (46), 20872–20877.

(32) Xu, H.; Chen, S.; Chen, L.; Tian, Y.; Yao, S.; Zhengliang Lü; Zhang, X.; Zhou, Z. What Is the Real Origin of the Activity of Fe–N–c Electrocatalysts in the O_2 Reduction Reaction? Critical Roles of Coordinating Pyrrolic N and Axially Adsorbing Species. *Journal of the American Chemical Society* **2022**, *144* (39), 18144–18152.

(33) Wang, Y.; Tang, Y.; Zhou, K. Self-Adjusting Activity Induced by Intrinsic Reaction Intermediate in Fe–N–c Single-Atom Catalysts. *Journal of the American Chemical Society* **2019**, *141* (36), 14115–14119.

(34) Jia, Q.; Ramaswamy, N.; Hafiz, H.; Tylus, U.; Strickland, K.; Wu, G.; Barbiellini, B.; Bansil, A.; Holby, E. F.; Piotr Zelenay; Mukerjee, S. Experimental Observation of Redox-

Induced Fe–N Switching Behavior as a Determinant Role for Oxygen Reduction Activity. *ACS Nano* **2015**, *9* (12), 12496–12505.

(35) 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. *J. Phys. Chem. B* **2004**, *108* (46), 17886–17892.

(36) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes Forab Initiototal-Energy Calculations Using a Plane-Wave Basis Set. *Phys. Rev. B.* **1996**, *54* (16), 11169–11186.

(37) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868.

(38) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B.* **1994**, *50* (24), 17953–17979.

(39) Himmetoglu, B.; Floris, A.; de Gironcoli, S.; Cococcioni, M. Hubbard-Corrected DFT Energy Functionals: The LDA+U Description of Correlated Systems. *Int. J. Quantum Chem.* **2013**, *114* (1), 14–49.

(40) Xu, H.; Cheng, D.; Cao, D.; Zeng, X. C. A Universal Principle for a Rational Design of Single-Atom Electrocatalysts. *Nat. Catal* **2018**, *1* (5), 339–348.

(41) Chen, Y.; Ji, S.; Wang, Y.; Dong, J.; Chen, W.; Li, Z.; Shen, R.; Zheng, L.; Zhuang, Z.; Wang, D.; Li, Y. Isolated Single Iron Atoms Anchored on N-Doped Porous Carbon as an Efficient Electrocatalyst for the Oxygen Reduction Reaction. *Angew. Chem., Int. Ed.* **2017**, *56* (24), 6937–6941.

(42) Li, Q.; Luo, L.; Xu, C.; Song, S.; Wang, Y.; Zhang, L.; Tang, Y.; Texter, J. Palladium Enhanced Iron Active Site – an Efficient Dual-Atom Catalyst for Oxygen Electroreduction. *Small* **2023**.

(43) Medford, A. J.; Vojvodic, A.; Hummelshøj, J. S.; Voss, J.; Abild-Pedersen, F.; Studt, F.; Bligaard, T.; Nilsson, A.; Nørskov, J. K. From the Sabatier Principle to a Predictive Theory of Transition-Metal Heterogeneous Catalysis. *J. Catal.* **2015**, *328*, 36–42.

(44) Yang, L.; Zhang, Y.; Huang, Y.; Deng, L.; Luo, Q.; Li, X.; Jiang, J. Promoting Oxygen Reduction Reaction on Carbon-Based Materials by Selective Hydrogen Bonding. *Chem. Eur.* **2023**.

(45) Mitchell, S.; Pérez-Ramírez, J. Single Atom Catalysis: A Decade of Stunning Progress and the Promise for a Bright Future. *Nat. Commun.* **2020**, *11* (1), 4302.

(46) Li, B.; Xie, H.; Yang, C.; Shi, C.; He, C.; Zhao, N.; Liu, E. Unraveling the Mechanism of Ligands Regulating Electronic Structure of MN4 Sites with Optimized ORR Catalytic Performance. *Appl. Surf. Sci.* **2022**, *595*, 153526.

(47) Rosen, A. S.; Notestein, J. M.; Snurr, R. Q. Structure–Activity Relationships That Identify Metal–Organic Framework Catalysts for Methane Activation. *ACS Catal.* **2019**, *9* (4), 3576–3587.

(48) Li, L.; Yuan, K.; Chen, Y. Breaking the scaling relationship limit: From single-atom to dual-atom catalysts. *Acc. Mater. Res.* **2022**, 3 (6), 584-596.

(49) Li, R.; Wang, D. Superiority of dual-atom catalysts in electrocatalysis: one step further than single-atom catalysts. *Adv. Energy Mater.* **2022**, 12 (9), 2103564.

(50) Fang, C.; Zhou, J.; Zhang, L.; Wan, W.; Ding, Y.; Sun, X. Synergy of dual-atom catalysts deviated from the scaling relationship for oxygen evolution reaction. *Nat. Commun.* **2023**, 14 (1), 4449.

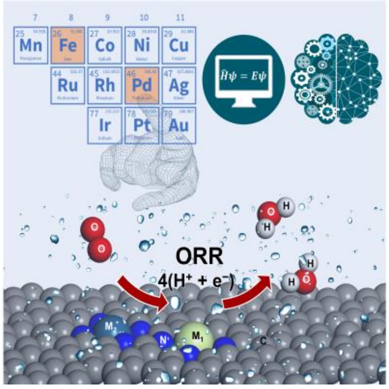
(51) Ouyang, R.; Curtarolo, S.; Ahmetcik, E.; Scheffler, M.; Ghiringhelli, L. M. SISSO: A Compressed-Sensing Method for Identifying the Best Low-Dimensional Descriptor in an Immensity of Offered Candidates. *Phys. Rev. Mater.* **2018**, *2* (8).

(52) Geurts, P.; Ernst, D.; Wehenkel, L. Extremely Randomized Trees. *Mach. Learn.* **2006**, *63* (1), 3–42.

(53) Christensen, R.; Hansen, H. A.; Dickens, C. F.; Nørskov, J. K.; Vegge, T. Functional independent scaling relation for ORR/OER catalysts. *J. Phys. Chem. C* **2016**, 120 (43), 24910-24916.

(54) Urrego-Ortiz, R.; Builes, S.; Calle-Vallejo, F. Impact of intrinsic density functional theory errors on the predictive power of nitrogen cycle electrocatalysis models. *ACS catal.* **2022**, 12 (8), 4784-4791.

TOC


7
8
9
10
11
25 Mn
26 Fe
27 Co
28 Ni
29 Cu
44 Ru
45 Rh
46 Pd
47 Ag
77 Ir
78 Pt
79 Au
$\hat{H}\psi = E\psi$
O
H
ORR
$4(H^+ + e^-)$
M
N
M_1
C