MBenes-supported single-atom catalysts for oxygen reduction and oxygen evolution reactions by first-principles study and machine learning

Erpeng Wang¹,#, Guanjie Wang²,#, Jian Zhou¹,²,* & Zhimei Sun¹,²,*

¹School of Materials Science and Engineering, Beihang University, Beijing 100191, China; ²Center for Integrated Computational Materials Engineering, International Research Institute for Multidisciplinary Science, Beihang University, Beijing 100191, China

#Contributed equally to this work.
*Corresponding authors (emails: jzhou@buaa.edu.cn (Jian Zhou); zmsun@buaa.edu.cn (Zhimei Sun))

Received 20 July 2023; Revised 23 December 2023; Accepted 7 January 2024; Published online 1 February 2024

Abstract: Oxygen reduction reaction (ORR) and oxygen evolution reaction (OER) are key catalytic processes in various renewable energy conversion and energy storage technologies. Herein, we systematically investigated the ORR and OER catalytic activity of the single-atom catalysts (SACs) composed of 4d/5d period transition metal (TM) atoms embedded on MBene substrates (TM-M₂B₂O₂, M = Ti, Mo, and W). We found that TM dominates the catalytic activity compared to the MBene substrates. The SACs embedded with Rh, Pd, Au, and Ir exhibit excellent ORR or OER catalytic activity. Specifically, Rh-Mo₂B₂O₂ and Rh-W₂B₂O₂ are promising bifunctional catalysts with ultra-low ORR/OER overpotentials of 0.39/0.21 V and 0.19/0.32 V, respectively, lower than that of Pt/RuO₂ (0.45/0.42 V). Importantly, through machine learning, the models containing 10 element features of SACs were developed to quickly and accurately identify the superior ORR and OER electrocatalysts. Our findings provide several promising SACs for ORR and OER, and offer effective models for catalyst design.

Keywords: oxygen reduction reaction, oxygen evolution reaction, single-atom catalysts, catalytic activity, machine learning

INTRODUCTION

The oxygen evolution reaction (OER) and oxygen reduction reaction (ORR) are important half-reactions in metal-air batteries and proton-exchange membrane full cells (PEMFC), respectively, which play crucial roles in green energy storage and conversion [1–8]. However, their four-electron reaction process results in overall slow kinetics, requiring efficient catalysts to accelerate the reaction. Conventionally, noble metal Pt and noble metal oxides RuO₂ or IrO₂ are still the most commonly used catalysts to promote ORR or OER, with overpotentials of 0.45 V (Pt for ORR) [9] and 0.42/0.56 eV (RuO₂/IrO₂ for OER) [10–12], but the high cost, scarcity, poor robustness, and low reaction selectivity severely restricted their large-scale practical applications [13–15]. Therefore, it is necessary to design stable, efficient, and non-precious metal ORR and OER electrocatalysts.
Due to the large specific surface area, unusual electronic properties, and high stability, two-dimensional (2D) MBenes have received widespread attention since their proposal by Guo et al. [16] in 2017, especially in the field of electrocatalysts [17–21]. Experimentally, MBenes can be obtained by selectively etching the A atomic layer from the MAB parent phase using different synthesis means (where M represents transition metal, A is the main group element of IIIA or IVA, and B is boron element) [22–26]. The outermost exposed M metal layer of MBenes is easily functionalized by O, OH, and F terminal groups in solution, which greatly reduces the adsorptions of ORR and OER reactants on the surface [19,24,27]. Therefore, increasing the adsorption capacity of ORR and OER reactants by regulating the surface electronic structure of functionalized MBenes is the key to the practical applications of MBenes in electrocatalytic ORR and OER.

Introducing single metal atoms on the surface of 2D materials is one of the commonly used methods to adjust the surface electronic structure [28–32]. In fact, the introduced single metal atoms can directly participate in ORR and OER as the catalytic activity site. On the one hand, due to the specific single active site, the formed single-atom catalysts (SACs) have high reaction selectivity, maximum atom utilization, and high intrinsic catalytic activity. On the other hand, 2D materials as the substrate can not only change the electron distribution of a single atom to activate suitable reactants but also firmly fix the single atom to prevent clusters caused by excessive surface energy of SACs, thus maintaining the overall high stability. However, finding efficient SACs for ORR and OER from numerous combinations of single-atom and 2D substrate materials is a huge challenge, and the mechanism of metal-support interactions also urgently needs to be further revealed.

Recently, driven by artificial intelligence technology, machine learning (ML) has become a powerful assistant for quickly screening superior ORR and OER catalysts. Liu et al. [33] proposed that the number of d electrons in transition metal (TM) atoms \( (N_e) \), the radius of TM atoms \( (r) \), and the charge transfer of TM atoms \( (Q_e) \) are the three most important features affecting the catalytic behavior by implementing the ML algorithm. Performing ML, Anand et al. [34] found that the embedding energy \( (E_e) \) of doped TM atoms, the distance between adsorbed OH and doped TM atoms \( (d-OH) \), and the distance between adsorbed O and doped TM atoms \( (d-O) \), the differences between \( E_e \) and cohesive energy \( (E-C) \) are main features determining the catalytic activity of ORR and OER. These extracted features are based on the density functional theory (DFT) and the architecture of systems, which are time-consuming and complex. Thus, proposing a more convenient descriptor based solely on element features is of great significance.

Herein, based on our previous work [19], the most stable O-terminal MBenes (Ti$_2$B$_2$O$_2$, Mo$_2$B$_2$O$_2$, and W$_2$B$_2$O$_2$) are selected as substrates. Using the first principles calculations, a systematic study was conducted on the ORR and OER catalytic activity of 54 SACs composed of 18 kinds of 4d, 5d period TM atoms anchored onto the three substrates mentioned above. By calculating the Gibbs free energy and overpotential of ORR and OER, Pd-Ti$_2$B$_2$O$_2$, Rh-Mo$_2$B$_2$O$_2$, Ir-Mo$_2$B$_2$O$_2$, Rh-W$_2$B$_2$O$_2$, and Pt-W$_2$B$_2$O$_2$ were selected as superior ORR catalysts, with overpotentials of 0.41, 0.39, 0.42, 0.19, and 0.30 V, respectively, comparable to or even lower than that of Pt (0.45 V). Moreover, Rh-Ti$_2$B$_2$O$_2$, Ir-Ti$_2$B$_2$O$_2$, Rh-Mo$_2$B$_2$O$_2$, Au-Mo$_2$B$_2$O$_2$, Rh-W$_2$B$_2$O$_2$, Ir-W$_2$B$_2$O$_2$, and Au-W$_2$B$_2$O$_2$ were selected as superior OER catalysts, with overpotentials of 0.54, 0.51, 0.20, 0.34, 0.31, 0.45, and 0.55 V respectively, comparable to or even lower than that of IrO$_2$/RuO$_2$ (0.56/0.42 V). Specifically, Rh-Mo$_2$B$_2$O$_2$ and Rh-W$_2$B$_2$O$_2$ exhibit high ORR and OER catalytic activity at the same time, and are considered promising bifunctional catalysts. In addition, we found that the adjusting single atom has a more significant effect on catalytic activity compared to adjusting the substrates for SACs.
Finally, combining data from our previous work [1,19], an ML dataset was constructed, which contained 117 SAC systems and 24 material elemental features that can be quickly obtained without complex computations. Subsequently, the feature engineering and decision tree method were used to filter out the top 10 features with lower correlation coefficients. Based on these features, two models were established to rapidly and accurately predict ORR and OER catalytic activity, respectively. Furthermore, based on symbolic regression and genetic algorithms, two symbolic expressions were developed to describe the ORR and OER catalytic activity of SACs, respectively. Our work provides several ORR and OER electrocatalysts with superior catalytic activity and provides simple and effective methods for screening and designing new catalysts.

**COMPUTATIONAL METHODS**

The spin-polarized DFT calculations were employed in Vienna *ab initio* simulation package (VASP) code with the projector augmented wave (PAW) method [35]. The exchange-correlation interaction was described by the Perdew-Burke-Ernzerhof (PBE) functional within the generalized gradient approximation (GGA) [36,37]. The kinetic energy cutoff was set to 500 eV. A $3 \times 3 \times 1$ Monkhorst-Pack $k$-point grid and a denser $7 \times 7 \times 1$ Monkhorst-Pack $k$-point grid was chosen for structural optimization and electronic structure calculations, respectively. The interlayer interactions were eliminated efficiently by setting a vacuum of 25 Å along the $z$-axis. The convergence thresholds for energy and force were set to $10^{-5}$ eV/atom and 0.01 eV/Å, respectively. Given that the studied systems involve surface adsorption and desorption processes, the impact of the van der Waals interactions was treated by using the empirical correction DFT-D3 approach [38]. Microkinetic simulations were carried out using the Catalysis Microkinetic Analysis Package (CATMAP) software package [39]. The thermal stability of SAC structures under aqueous solution was evaluated by performing *ab initio* molecular dynamics (AIMD) simulations at 300 K for 10 ps. It is worth noting that the selection of the above parameters is roughly consistent with our previous work to facilitate subsequent ML [1,19].

The ORR process of MBenes-supported SACs is typically performed by the widely reported four-electron dissociation mechanism, which includes four elementary steps and can be described by the following equations [40]:

\[
\begin{align*}
O_2(g) + H^+ + e^- + * & \rightarrow OOH^+ \\
OOH^+ + H^+ + e^- & \rightarrow O^+ + H_2O(l) \\
O^+ + H^+ + e^- & \rightarrow OH^+ \\
OH^+ + H^+ + e^- & \rightarrow H_2O(l) + *
\end{align*}
\]

where $*$, $H^+$, and $e^-$ represent the active sites of SAC surface, proton, and transferred electron, respectively. $g$ and $l$ denote the gas phase and liquid phase, respectively.

OER can be regarded as the reverse reaction process of ORR, which is composed of the following four corresponding elementary steps [41]:

\[
\begin{align*}
H_2O(l) & \rightarrow OH^+ + H^+ + e^- \\
OH^+ & \rightarrow O^+ + H^+ + e^-
\end{align*}
\]
In order to evaluate the ORR catalytic activity, it is necessary to calculate the Gibbs free energy change (\(\Delta G\)) of each elementary step. According to the computational hydrogen electrode (CHE) model proposed by Nørskov et al.\[9\], the general formula of \(\Delta G\) can be described as the following equation for each elementary step.

\[
\Delta G = \Delta E + \Delta ZPE - T \Delta S + \Delta G_{\text{pH}} + \Delta G_U
\]

where \(\Delta E\) is the total energy change before and after the elementary reaction, which can be directly obtained from DFT calculations. \(\Delta ZPE\) and \(T \Delta S\) represent the changes in zero-point energy and entropic energy, respectively. \(\Delta G_{\text{pH}} = -k_B T \ln[H^+] = pH \times k_B T \ln10\) represents the Gibbs free energy correction at the specified pH, where \(k_B\) is the Boltzmann constant. The last term is the correction of electrical potential to Gibbs free energy, denoted as \(\Delta G_U = -eU\), where \(U\) is the applied electrode potential with respect to the standard hydrogen electrode and was set to 0 V, and \(e\) is the elementary charge. Herein, all reactions were assumed to occur under standard conditions \((T = 298.15 \text{ K}, P = 1 \text{ bar and pH} = 0)\). Correspondingly, the free energy changes of the four elementary steps of ORR can be specifically expressed as Equations (10)–(13).

\[
\begin{align*}
\Delta G_a &= \Delta G_{\text{OH}^+} - 4.92 \\
\Delta G_b &= \Delta G_{\cdot} - \Delta G_{\text{OH}^+} \\
\Delta G_c &= \Delta G_{\text{OH}^+} - \Delta G_{\cdot} \\
\Delta G_d &= -\Delta G_{\text{OH}^+}
\end{align*}
\]

For ORR, the elementary step with the maximum energy change would be determined as the potential determining step. To evaluate the catalytic activity of ORR, the well-accepted theoretical overpotential \(\eta^{\text{ORR}}\) was defined as the equation below:

\[
\eta^{\text{ORR}} = \max\{\Delta G_{a}, \Delta G_{b}, \Delta G_{c}, \Delta G_{d}\} / e + 1.23
\]

Similarly, the theoretical overpotential \(\eta^{\text{OER}}\) was defined as

\[
\eta^{\text{OER}} = \max\{-\Delta G_{a}, -\Delta G_{b}, -\Delta G_{c}, -\Delta G_{d}\} / e - 1.23
\]

According to the definition of overpotential, for both ORR and OER, the lower the overpotential, the higher the catalytic activity.

**RESULTS AND DISCUSSION**

**Structural stability**

The long-lasting stability of the catalysts’ structure is indispensable for practical applications. In order to access the feasibility of SAC formation, we first investigated all possible embedding sites of TM in 54 SACs composed of 18 kinds of 4d/5d period TM atoms anchored on three kinds of MBene substrates (donated as TM-M_2B_2O_2, TM = 4d/5d period atoms, M = Ti, Mo and W), and selected the structure with optimal embedding sites for subsequent research, as illustrated in Figure 1. Then, the binding strength of single atoms anchored on different substrates was calculated, which can be quantified in the form of binding energy, defined as
where \( E_{\text{TMMBene}} \), \( E_{\text{MBene}} \), and \( E_{\text{TM}} \) are the total energies of the TM atom anchored SAC systems, the free MBene substrates, and isolated TM atoms, respectively. The more negative binding energy \( E_b \), the more favorable the load of TM atoms. The binding energies of all SAC systems are negative, indicating that all TM atoms can be successfully anchored (Table S1).

As is well known, maintaining uniform and isolated distribution of single atom on substrates is a prerequisite for the advantages of SACs. Thus, to evaluate the aggregation possibility of TM atoms, the metal cluster energy \( E_{\text{cluster}} = E_b - E_{\text{coh}} \) was calculated, where \( E_{\text{coh}} \) is the metal cohesive energy. Herein, \( E_{\text{coh}} \) can be calculated by the energy difference between a metal atom in the most stable bulk phase and in the vacuum, expressed as \( E_{\text{coh}} = E_{\text{TM-bulk}}/n - E_{\text{TM-singles}} \), where \( n \) is the number of atoms in the most stable bulk phase. A value of \( E_{\text{cluster}} \) less than 0 eV is usually considered a powerful criterion for determining the thermodynamic stable of SAC systems. In addition, the dissolution potentials \( U_{\text{diss}} = U_{\text{diss-bulk}} - E_{\text{cluster}} / (eN_e) \) of TM atoms were further calculated to explore the stability of the studied SAC systems in electrochemical environments. \( U_{\text{diss-bulk}} \) represents the standard dissolution potential of TM bulk that can be directly obtained from the handbook of chemistry and physics, and \( N_e \) is the number of transferred electrons [33,42]. Theoretically, a positive value of \( U_{\text{diss}} \) indicates that the SACs are electrochemically stable. The \( E_{\text{cluster}} \) and \( U_{\text{diss}} \) of all SACs are depicted in Figure 2, where the SACs located in the gray background area are thermodynamic and electrochemical stable \( (U_{\text{diss}} > 0 \text{ V and } E_{\text{cluster}} < 0 \text{ eV}) \). In addition, we found that the substrate has a significant impact on the thermodynamic stability and electrochemical stability of SACs, with Mo-based being the optimal substrate materials, which may benefit from the more charge transfer between the Mo-based substrate and metal atoms (Figure S1). Finally, in order to verify the stability of SACs under aqueous solution, the stability of the Rh-Mo\(_2\)B\(_2\)O\(_2\) under aqueous solution composed of 32 water molecules was further performed through AIMD simulations at 300 K. As shown in Figure S2, the AIMD simulations confirm that Rh-Mo\(_2\)B\(_2\)O\(_2\) is thermodynamically stable under aqueous solution.

---

Figure 1: The schematic diagram of SAC structures, the considered TM atoms as well as four possible embedding sites of single atoms on the O-terminated MBenes surface, i.e., Hollow, Bridge1, Bridge2, and Top. (A) Top view; (B) side view.
We systematically evaluated the ORR and OER catalytic activities of all the 54 SAC systems in this section. Generally, ORR occurring on the surface of MBene-supported SACs tends to undergo a four-electron association mechanism rather than the dissociation mechanism. This is because that the whole ORR process only involves a single active site, and the fracture of O–O bond needs to overcome a large energy barrier [1,19,43]. Therefore, we only discussed the association mechanism in this work, which is consistent with our previous work [19]. Firstly, the free energies of all adsorbed oxygen intermediates (O*, OH*, and OOH*) in the ORR process were calculated according to Equation (9) and summarized in Tables S2–S4. Furthermore, based on Equations (14) and (15), the overpotentials of all studied SAC systems for ORR and OER were obtained and summarized in Figure 3C and Tables S2–S4. Obviously, Pd-Ti$_2$B$_2$O$_2$, Rh-Mo$_2$B$_2$O$_2$, Ir-Mo$_2$B$_2$O$_2$, Rh-W$_2$B$_2$O$_2$, and Pt-W$_2$B$_2$O$_2$ exhibit superior ORR catalytic activity with overpotentials of 0.41, 0.39, 0.42, 0.19, and 0.30 V, respectively, comparable to or even lower than that of Pt (0.45 V) [9]. Moreover, Rh-Ti$_2$B$_2$O$_2$, Ir-Ti$_2$B$_2$O$_2$, Rh-Mo$_2$B$_2$O$_2$, Au-Mo$_2$B$_2$O$_2$, Rh-W$_2$B$_2$O$_2$, Ir-W$_2$B$_2$O$_2$, and Au-W$_2$B$_2$O$_2$ exhibit superior OER catalytic activity with overpotentials of 0.54, 0.51, 0.20, 0.34, 0.31, 0.45, and 0.55 V, comparable to or even lower than that of IrO$_2$/RuO$_2$ (0.56/0.42 V) [10–12]. After a comprehensive comparison, Rh-Mo$_2$B$_2$O$_2$ and Rh-W$_2$B$_2$O$_2$ exhibit the highest comprehensive performance due to the combination of ultra-low ORR/OER overpotential. It is worth noting that these SAC catalysts with high ORR and/or OER catalytic activity are mainly concentrated in several specific single atoms such as Rh, Ir, Pd, and Au, indicating that the essential regulatory role of single atoms dominates the catalytic activity in SAC systems. In order to further understand the detailed processes of catalysis, the Gibbs free energy diagrams were drawn and shown in Figure 3A and B and Figure S3, from which the energy change of each elementary step and the potential determining steps (PDS) can be clearly observed. For Rh-Mo$_2$B$_2$O$_2$ and Rh-W$_2$B$_2$O$_2$, the PDS of ORR are both the process of O$_2$ protonation to OOH*, and the PDS of OER is both the process of OH* deprotonation to O*. It is worth noting that the absolute value of Gibbs free energy change of the four elementary steps of Rh-Mo$_2$B$_2$O$_2$ and Rh-W$_2$B$_2$O$_2$ are close to the equilibrium potential of 1.23 V for ORR and OER, and thus
they have ultra-low $\eta_{\text{ORR}}$ and $\eta_{\text{OER}}$, exhibiting excellent ORR and OER bifunctional catalytic activity.

**ORR/OER activity trend**

The adsorption strength between TM in SACs and the three intermediates (O*, OH*, and OOH*) generally presents a scaling relationship, which has been reported in many related studies [33,44,45]. Then, we organized the adsorption free energies of three intermediates ($\Delta G_{\text{O*}}$, $\Delta G_{\text{OH*}}$, and $\Delta G_{\text{OOH*}}$) in all SACs and explored their relationships. As shown in Figure 4, both $\Delta G_{\text{OOH*}}$ and $\Delta G_{O*}$ show a positive correlation with the variation of $\Delta G_{\text{OH*}}$. Their correlation equations can be linearly fitted as $\Delta G_{\text{OOH*}} = 0.79\Delta G_{\text{OH*}} + 3.19$ ($R^2 = 0.97$) and $\Delta G_{O*} = 1.4\Delta G_{\text{OH*}} + 1.35$ ($R^2 = 0.76$), respectively.

Based on the scaling relationship, the volcanic curves of $\Delta G_{\text{OH*}}$ versus $-\eta_{\text{ORR}}$, as well as $\Delta G_{O*} - \Delta G_{\text{OH*}}$ versus $-\eta_{\text{OER}}$ were plotted in Figure 5A and B, respectively. As shown in Figure 5A, the PDS of the ORR process in all studied SACs is mostly concentrated in the elementary step of OH* removal to form H$_2$O (OH* + H$^+$ + e$^-$ → * + H$_2$O), which is reflected in the highly linear fitting part of the left branch in the volcanic curve, corresponding to the strong adsorption of OH* intermediate. On the contrary, the right branch part exhibits weak OH* binding strength, corresponding to the formation of OOH* as the PDS in these SACs (* + O$_2$ + H$^+$ + e$^-$ → OOH*). According to Sabatier principle [14], both too strong and too weak adsorption strength between the catalysts and the adsorbent are detrimental to catalytic performance. These catalysts located at the top of the volcano plot demonstrate moderate adsorption strength with the adsorbent and...
exhibit optimal catalytic activity, such as Pd-Ti$_2$B$_2$O$_2$, Rh-Mo$_2$B$_2$O$_2$, and Rh-W$_2$B$_2$O$_2$. Besides, a small amount of catalyst deviates from the volcanic curves, such as Ag-Ti$_2$B$_2$O$_2$, Cd-Mo$_2$B$_2$O$_2$, and Cd-W$_2$B$_2$O$_2$, mainly due to the formation of O* intermediate being their PDS (OOH* + H$^+$ + e$^-$ → O* + H$_2$O). Similarly, for $\Delta G_{O^*} - \Delta G_{OH^*}$ versus $-\eta_{OER}$ volcanic curves, the PDS of the catalysts located on the left and right branches are the formation of OOH* intermediate (O* + H$_2$O → OOH* + H$^+$ + e$^-$) and O* intermediate (OH* → O* + H$^+$ + e$^-$), respectively. The SACs with the best OER catalytic activity were located at the top of the volcanic plots, occupied by Rh-Ti$_2$B$_2$O$_2$, Rh-Mo$_2$B$_2$O$_2$, Au-Mo$_2$B$_2$O$_2$, Rh-W$_2$B$_2$O$_2$, and Ir-W$_2$B$_2$O$_2$. Moreover, the PDS of these SACs located off the volcanic curves is the desorption of O$_2$ (OOH* + H$_2$O → O$_2$ + H$^+$ + e$^-$), such as Hf-Ti$_2$B$_2$O$_2$, Zr-Ti$_2$B$_2$O$_2$, Hf-Mo$_2$B$_2$O$_2$, Zr-Mo$_2$B$_2$O$_2$, Hf-W$_2$B$_2$O$_2$, Zr-W$_2$B$_2$O$_2$, and Ta-W$_2$B$_2$O$_2$. Therefore, the combination of superior OER and ORR catalytic activity renders Rh-Mo$_2$B$_2$O$_2$ and Rh-W$_2$B$_2$O$_2$ SACs promising comprehensive bifunctional catalysts.

Furthermore, in order to more intuitively display the PDS and overpotential distribution regions of all SACs, variables $\Delta G_{OH^*}$ and $\Delta G_{O^*} - \Delta G_{OH^*}$ were selected to depict the catalytic activity of ORR and OER. Specifically, by replacing $\Delta G_{OOH^*}$ with $\Delta G_{OOH^*} = 0.79\Delta G_{OH^*} + 3.19$ ($R^2 = 0.97$) in Equations (10)–(13), the four free-energy changes can be further expressed as the following Equations (17)–(20), containing only $\Delta G_{OH^*}$ and $\Delta G_{O^*} - \Delta G_{OH^*}$ variables.

\[
\begin{align*}
\Delta G_a &= 0.78\Delta G_{OH^*} - 1.74 \\
\Delta G_b &= (\Delta G_{O^*} - \Delta G_{OH^*}) + 0.22\Delta G_{OH^*} - 3.18 \\
\Delta G_c &= -(\Delta G_{O^*} - \Delta G_{OH^*}) \\
\Delta G_d &= -\Delta G_{OH^*}
\end{align*}
\]

Subsequently, the contour map of ORR catalytic activity regarding $\Delta G_{OH^*}$ and $\Delta G_{O^*} - \Delta G_{OH^*}$ is plotted in Figure 5C. Obviously, this contour map is divided into four regions by dashed lines, representing the four elementary reactions of the ORR process. If a catalyst falls in a certain region, it means that the corresponding elementary reaction in that region is the PDS of this catalyst. For early TMs of 4d and 5d periods (such as Y, Zr, Nb, Mo, Tc, Hf, Ta, W, and Re), MBenes-supported SACs exhibit strong OH* adsorption ($\Delta G_{OH^*} < -0.5$ eV), making the PDS solely determined by $\Delta G_{OH^*}$, that is, the desorption process of water in the fourth elementary step. For the later TM (such as Ru, Rh, Pd, Ag, Cd, Os, Ir, Pt, and Au), the further
weakening of the adsorption strength of MBenes-supported SACs for OH* intermediate results in the PDS jointly dominated by $\Delta G_{\text{OH}^*}$ and $\Delta G_{\text{OOH}^*} - \Delta G_{\text{OH}^*}$, and the region where the SACs located correspondingly shifts from Zone 4 (OH* → H2O) to Zone 2 (OOH* → O*) and Zone 1 (O2 → OOH*). Due to the limitations of the scaling relationship, the idea ORR SACs should be located at the boundaries of the four zones. Among all SACs, Pd-Ti2B2O2, Rh-Mo2B2O2, Ir-Mo2B2O2, Rh-W2B2O2, and Pt-W2B2O2 are closest to the boundaries of the four zones and exhibit excellent ORR catalytic activity with an ultra-low overpotential of 0.41, 0.39, 0.42, 0.19 and 0.30 V, respectively.

In the same way, the OER contour map diagram is presented in Figure 5D, which can also be divided into four regions by dashed lines. Accompanied by an increase in $\Delta G_{\text{O}^*} - \Delta G_{\text{OOH}^*}$, the PDS of the SACs shifts from the third step (O* → OOH*) to the fourth step (OOH* → O2) and the second step (OH* → O*). Rh-Mo2B2O2, Au-Mo2B2O2, and Rh-W2B2O2 are closest to the boundary of the four zones and exhibit optimal OER catalytic activity with an ultra-low overpotential of 0.21, 0.34, and 0.32 V, respectively.

Finally, the microkinetic simulations were performed to evaluate the respective turnover frequency (TOF, molecules/sites) of H2O generation for the selected catalysts at 0.9 V based on possible ORR pathways (the specific simulation details are provided in the Supplementary information), as shown in Figure 6A. Due to the strong scaling relationships between $\Delta G_{\text{OH}^*}$ and $\Delta G_{\text{OOH}^*}$, the SACs are dispersed at both sides of the fitting...
line, where Rh-W$_2$B$_2$O$_2$ SACs are closest to the optimal region, exhibiting the highest TOF of H$_2$O production. At the same time, it is not difficult to see that it is worth looking forward to breaking the original scaling relationship through further surface modification to improve the ORR catalytic activity of SACs, that is, the SACs fall into the optimal region. The potential-dependent TOFs of H$_2$O production are plotted and depicted in Figure 6B. Compared with the simulation results of the half-wave potentials on pure Pt for ORR [46], the corresponding half-wave potentials of Rh-Mo$_2$B$_2$O$_2$ SACs and Rh-W$_2$B$_2$O$_2$ SACs exhibit positive shifts of 40 and 140 mV, respectively, further demonstrating the superior ORR catalytic activity of Rh-Mo$_2$B$_2$O$_2$ SACs and Rh-W$_2$B$_2$O$_2$ SACs.

### OER/ORR activity origin

To reveal the essential mechanism of high catalytic activity of the studied systems, the electronic structures of all SACs have been calculated. Since the single transition atoms acting as active sites in all SACs are directly involved in OER and ORR processes, the electronic structure characteristics of the single atoms were given special attention. Previous studies have proved that the d-band center ($\varepsilon_d$) is a reliable descriptor to characterize the catalytic activity of catalysts [5,47–50], typically defined as

$$\varepsilon_d = \frac{\int_{-\infty}^{+\infty} ED(E)dE}{\int_{-\infty}^{+\infty} D(E)dE}$$  \hspace{1cm} (21)$$

where $D(E)$ is the density of states (DOS) of the d orbitals of the TM atoms. Subsequently, the relationship between the adsorption free energy of the intermediates (O*, OH*, and OOH*) and the d-band center of the SACs are depicted in Figure 7A. It can be observed that the $\Delta G_{O^*}$, $\Delta G_{OH^*}$, and $\Delta G_{OOH^*}$ exhibit a similar trend as the value of the d-band center increases. However, the correlation of the $\Delta G_{O^*}$, $\Delta G_{OH^*}$, and $\Delta G_{OOH^*}$ versus $\varepsilon_d$ is relatively poor, making $\varepsilon_d$ unsuitable to describe the OER/ORR catalytic activity for SACs. In fact, the closer the transition metal atom’s electronic states to the Fermi level, the greater its contribution to the interaction with the adsorbents. Therefore, it is necessary to introduce a weight function to quantitatively consider the unequal weight contribution of electronic states at different distances from Fermi energy level.
In this work, the derivative of the Fermi-Dirac function, \( -f'_{\gamma}(E-E_F) \), was chosen as the weight function, which has been validated in many previous studies (Figure 7B) \([51,52]\). Then, the improved d-band center \( \varepsilon_{id} \) is defined as

\[
\varepsilon_{id} = \frac{\int_{-\infty}^{\infty} E D(E) f'_{\gamma}(E-E_F) dE}{\int_{-\infty}^{\infty} D(E) f'_{\gamma}(E-E_F) dE}
\]  

(22)

where \( E_F \) is the energy of the Fermi level and \( f_{\gamma}(E-E_F) \) is the Fermi-Dirac function, expressed as

\[
f_{\gamma}(E-E_F) = \frac{1}{\exp[(E-E_{\gamma})/kT] + 1}
\]  

(23)

where \( k \) is the Boltzmann’s constant, \( T \) represents the temperature, and \( kT \) represents the nominal electron temperature. Obviously, the spreading of the weight function can be adjusted through the \( kT \) parameter. Through systematic testing of \( kT \) values ranging from 0.1 to 1 eV, it was confirmed that the best fitting effect was achieved when the \( kT \) was set to 0.5 eV. As shown in Figure 7C, the relationships between the adsorption free energies of intermediate adsorbents (O*, OH*, and OOH*) and \( \varepsilon_{id} \) were established. Compared to the adsorption free energies versus \( \varepsilon_d \) (Figure 7A), the adsorption free energies versus \( \varepsilon_{id} \) exhibit a higher linear correlation. According to the d-band center theory \([53]\), the interaction between the sp orbitals of adsorbed molecules and the d orbitals of metals would cause hybridization of orbitals and form binding and anti-bonding states. With the decrease of the d-band center, the anti-bonding energy band generated by coupling
decreases accordingly and is then filled with electrons, reducing the bond strength and further leading to the decrease of the adsorption free energy. This also indicates that the results of this work follow the d-band center criterion, that is, the d-band center is negatively correlated with the adsorption free energy. Nevertheless, the fact is that the improved d-band center still cannot effectively describe the catalytic activity of the studied SACs.

To provide a clearer explanation of the adsorption behavior mechanism of adsorbents, the bonding strength between OH* intermediate and metal atoms for all SACs was first investigated by the bond length and charge transfer, as shown in Figure S4. It can be seen that the bond length and charge transfer from the metal atoms to the substrates cannot accurately describe the Δ\(G_{OH^*}\) due to the different radii and valence electron structures for different metal atoms. Then, the crystal orbital Hamiltonian population (COHP) of OH* intermediate for all SACs was further calculated. Furthermore, by integrating COHP, the parameter ICOHP can be obtained to quantify the bond strength between OH* and metal atoms. A more intuitive relationship between ICOHP and Δ\(G_{OH^*}\) is depicted in Figure 8A. It can be observed that there is a linear correlation between ICOHP and Δ\(G_{OH^*}\). The linear relationship between ICOHP and adsorption free energy has also been reflected in our previous work (Figure S5) [19]. Besides, ICOHP presents a similar pattern as the increase of the atomic number of TM atoms for a given period, that is, it decreases first and then increases, except for Pt-M\(_2\)B\(_2\)O\(_2\) (M = Ti, Mo, and W), as shown in Figure 8B. The more negative ICOHP, the smaller Δ\(G_{OH^*}\), and the stronger the bond strength between OH* and metal atoms. The SACs embedded with Rh, Ir, Pd, and Au atoms exhibit suitable ICOHP values, thus displaying superior ORR or OER catalytic activity, which also explains the reason why single metal atoms are the main factor affecting catalytic activity.

**Machine learning methods**

Based on the above findings, it is known that the catalytic activity of SACs is closely related to their adsorption surface structures and adsorbed TM atoms. In order to obtain a more explicit and intuitive relationship expression, the models containing 10 intrinsic elements features of SACs were established, and
the quantitative expressions of ORR and OER catalytic activity were further obtained through symbolic regression. In this work, the construction of the models and the symbolic regression expressions for the OER and ORR performance mainly encompasses the following three core aspects:

**Dataset for machine learning**

The dataset used for ML consists of a total of 117 initial data, sourced from 54 SAC systems in this work, 27 SAC systems of MBenes-supported TM atoms (TM-M$_2$B$_2$O$_2$, TM = Ti, V, Mn, Fe, Co, Ni, Cu, and Zn, M = Ti, Mo, and W) [19], 24 SAC systems of Ti$_2$CO$_2$-supported TM atoms (TM = Ti, V, Cr, Mn, Fe, Co, Ni, Cu, Zr, Nb, Mo, Tc, Ru, Rh, Pd, Ag, Hf, Ta, W, Re, Os, Ir, Pt, and Au) and 12 SAC systems of Ti$_3$C$_2$O$_2$-supported TM atoms (TM = Mn, Fe, Co, Ni, Cu, Ru, Rh, Pd, Ag, Ir, Pt, and Au) [1]. Furthermore, based on the cross-validation method, the aforementioned dataset was divided into a training set containing 75% data and a validation set containing 25% data. With the training set, 70% data were allocated for training, and the remaining 30% was used for the testing dataset.

**Descriptors**

Material descriptors, often termed as “features” in the ML community, are typically used to transform material composition or structural information into ML input datasets. In this study, 37 features (descriptors) were assigned for each given data-point, which can be directly extracted from element characteristics. These features include both numerical and categorical variables (the configuration of extra-nuclear electrons). Through feature engineering, 13 features with similar maximum and minimum values are cleared, and the remaining 24 features are listed in Table S5. Their correlation coefficients are showcased in Figure 9A. A notable number of these features demonstrate a high correlation coefficient, implying that these features are not independent variables (Figure S6). Therefore, the features with an absolute value of correlation coefficients exceeding 0.85 have been removed, and the subsequent 10 independent features are represented in
Figure 10 Comparison between ML-predicted and DFT-calculated (A) $\eta^{\text{ORR}}$, (B) $\eta^{\text{OER}}$ using models and (C) $\eta^{\text{ORR}}$, (D) $\eta^{\text{OER}}$ using symbolic regression models.

**Machine learning model and symbolic regression**

With the advancements in computational capability and GPU acceleration, we have separately established two ML models for ORR and OER performances using the Random Forest algorithm. These models consist of 100 decision tree sub-models, and use the mean-square error (MSE) as the accuracy measurement standard of the model. The minimum number of samples required for node splitting in the decision tree is 2, and the minimum number of samples for leaf nodes is 1. A maximum of one feature is considered during each decision tree branching. Overfitting is prevented by setting the maximum search depth of the decision tree to 10 and incorporating cross-validation. As a result, we have constructed two models for predicting ORR and OER performances, with accuracies of 92.17% and 91.72%, respectively, as depicted in Figure 10A and B. These models do not rely on any computational parameters and can rapidly and accurately predict the ORR and OER performances of materials based on ten element-related features.

Furthermore, to quantitatively describe the mathematical relationship between material features and target performances, we have integrated the Random Forest method with genetic algorithms to construct a symbolic regression approach, as detailed in Github link (https://github.com/ALKEMIE-Lab/sraes). In symbolic regression, the cross-validation method is used to ensure that the data does not overfit. Initially, we obtained the
importance of each feature through the decision tree method. The importance of each feature for ORR and OER is depicted in Figure S7. It can be found that $N_\text{e}$ in TM atoms is the dominant factor associated with ORR and OER performance, followed by the covalent bond of the TM element ($r_{\text{cov}}\text{TM}$), the atomic number of TM ($Z_{\text{TM}}$), and the number of valence electrons in A and B atoms ($V_A$ and $V_B$), etc. Furthermore, targeting the top 10 significant feature descriptors, we constructed quantitative expressions to describe the catalytic activity of ORR and OER, as illustrated in Equations (24) and (25). The mean square errors for the aforementioned equations were 0.095 and 0.094, respectively, corresponding to the prediction accuracy of 79.91% and 75.36% for ORR and OER, respectively, as depicted in Figure 10C and D. The accuracy of symbolic regression models is comparable to that of the relevant work [54,55]. The aforementioned equations have proposed new ideas for exploring the physicochemical laws between material features by using compact mathematical expression.

$$\eta_{\text{predicted}}^{\text{OER}} = \sqrt{0.825 + \sqrt{Z_{\text{TM}} - I_{\text{TM}}}} - N_\text{e} + \sqrt{Z_{\text{TM}} - N_B} - N_\text{e}$$

$$\eta_{\text{predicted}}^{\text{ORR}} = \sqrt{\frac{r_{\text{cov}}\text{TM}}{N_\text{e}}} - N_\text{e} + \frac{Z_{\text{TM}}}{N_B}$$

**CONCLUSION**

In summary, a series of SACs composed of 4d and 5d period transition atoms anchored onto three MBenes ($\text{Ti}_2\text{B}_2\text{O}_2$, Mo$_2$B$_2$O$_2$, and W$_2$B$_2$O$_2$) were designed to comprehensively study their ORR and OER catalytic activity based on first principles calculations and ML. The results show that Rh-Mo$_2$B$_2$O$_2$, Ir-Mo$_2$B$_2$O$_2$, Rh-W$_2$B$_2$O$_2$, and Pt-W$_2$B$_2$O$_2$ have superior ORR catalytic activity, with overpotentials of 0.41, 0.39, 0.42, 0.19, and 0.30 V, respectively, comparable to or even lower than that of Pt (0.45 V). In addition, Rh-Ti$_2$B$_2$O$_2$, Ir-Ti$_2$B$_2$O$_2$, Rh-Mo$_2$B$_2$O$_2$, Au-Mo$_2$B$_2$O$_2$, Rh-W$_2$B$_2$O$_2$, Ir-W$_2$B$_2$O$_2$, and Au-W$_2$B$_2$O$_2$ have superior OER catalytic activity, with overpotentials of 0.54, 0.51, 0.20, 0.34, 0.31, 0.45, and 0.55 V, respectively, comparable to or even lower than that of IrO$_2$/RuO$_2$ (0.56/0.42 V). Specifically, Rh-Mo$_2$B$_2$O$_2$ and Rh-W$_2$B$_2$O$_2$ exhibit both ultra-low overpotentials of ORR and OER, making them promising bifunctional catalysts. Moreover, we found that the influence of transition metal atoms on the catalytic activity of SACs is dominant compared to the influence of the MBene substrates. Combining the scaling relationships of intermediate species, volcano maps and contour maps were established to explore the adsorption behavior of the SACs. Furthermore, the analysis of d-band center and ICOHP explained that the moderate interaction strength between the SACs and intermediate species is a source of excellent catalytic activity. Finally, through ML, the models containing 10 intrinsic element features of SACs were established, and the symbolic expressions describing the catalytic activity were provided to quickly and accurately identify catalysts with excellent ORR and OER catalytic activity. Our work provides several novel and promising ORR and OER electrocatalysts, and develops efficient models to reasonably guide and design new catalysts.

**Acknowledgements**

We are very grateful for the support of the high-performance computing (HPC) resource platform of Beihang University.
Funding
This work was supported by the National Key Research and Development Program of China (2022YFB3807200).

Author contributions
J.Z. and Z.S. proposed and supervised the project. E.W. and G.W. conceived and wrote the manuscript. All authors read and approved the submission of the manuscript.

Conflict of interest
The authors declare no conflict of interest.

Supplementary information
The supporting information is available online at https://doi.org/10.1360/nso/20230043. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References
28206–28216.


46 Hammer B, Nørskov JK. Why gold is the noblest of all the metals. *Nature* 1995; **376**: 238–240.
