JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau
Article

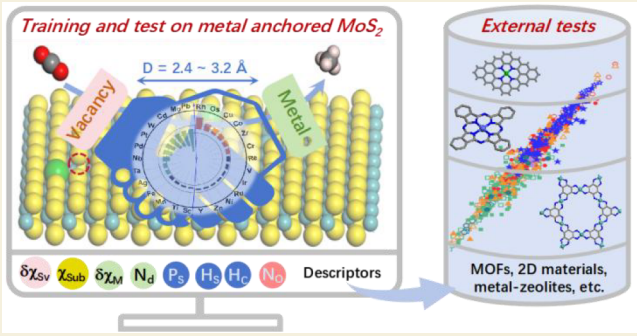
The Synergistic Effect between Metal and Sulfur Vacancy to Boost CO_2 Reduction Efficiency: A Study on Descriptor Transferability and Activity Prediction

Qin Zhu,[§] Yating Gu,[§] Xinzhu Wang, Yuming Gu, and Jing Ma*

Cite This: *JACS Au* 2024, 4, 125−138
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Both metal center active sites and vacancies can influence the catalytic activity of a catalyst. A quantitative model to describe the synergistic effect between the metal centers and vacancies is highly desired. Herein, we proposed a machine learning model to evaluate the synergistic index, P_{Syn}, which is learned from the possible pathways for CH_4 production from CO_2 reduction reaction (CO_2RR) on 26 metal-anchored MoS_2 with and without sulfur vacancy. The data set consists of 1556 intermediate structures on metal-anchored MoS_2, which are used for training. The 2028 structures from the literature, comprising both single active site and dual active sites, are used for external test. The XGBoost model with 3 features, including electronegativity, d-shell valence electrons of metal, and the distance between metal and vacancy, exhibited satisfactory prediction accuracy on limiting potential. Fe@Sv-MoS_2 and Os@MoS_2 are predicted to be promising CO_2RR catalysts with high stability, low limiting potential, and high selectivity against hydrogen evolution reactions (HER). Based on some easily accessible descriptors, transferability can be achieved for both porous materials and 2D materials in predicting the energy change in the CO_2RR and nitrogen reduction reaction (NRR). Such a predictive model can also be applied to predict the synergistic effect of the CO_2RR in other oxygen and tungsten vacancy systems.



KEYWORDS: *CO_2 reduction, metal−vacancy synergistic effect, MoS_2, machine learning, DFT calculations*

INTRODUCTION

The widely used metal−supported catalysts are composed of active metal atoms anchored on supports. Excellent catalytic performance is achieved through the interaction between the metal and support, which affects the electronic structure and morphology of the catalyst.[1−4] Surface vacancies are unavoidably introduced during the preparation of the surface-supported catalysts, tuning the electrical, magnetic, and optical properties.[5−7] For instance, metal−supported catalysts enriched with anion vacancies (oxide, sulfide, and nitride vacancies) effectively promoted charge separation, enhanced oxygen adsorption, and stabilized interfacial structures, thus improving the activity and selectively for surface reactions.[8−11] The vacancies on metal−supported catalysts are also able to promote light absorption, charge separation, and CO_2 conversion.[12−16] The essential role of vacancies, including vacancy type and location, vacancy concentration, and doped metal-based motifs, has been revealed in CO_2 reduction reactions (CO_2RR).[17−19] The oxygen-vacancy-rich MoO_{2-x} nanosea-urchins can promote CO_2 adsorption and activation, displaying extremely strong CO_2 photoreduction ability.[20] In the $NiCo_2O_4$ system, the synergistic effect between oxygen vacancies and Ni facilitates CH_4 generation and vacancy regeneration, showing an attempt at precise control of photocatalytic selectivity and stability.[21] Introducing vacancies and Zn atoms into CoO systems was beneficial in enhancing the stability of the vacancies in photocatalysts. The synergistic effect between metal and vacancies, which was attributed to a significant improvement of CO_2 photocatalytic efficiency.[22] Promotion of CO_2RR activity was achieved by constructing a gradient tungsten vacancy on Bi_2WO_6, promoting rapid electron migration, and reducing the formation barrier of key intermediates.[23] The vacancies not only modulate the local electronic structure but also serve as a docking site to anchor the metal species or adsorbates as a new active site. It is essential to get in-depth insight into the synergistic effect between metal centers and vacancies in metal−supported catalysts for CO_2RR.

Received: September 22, 2023
Revised: December 15, 2023
Accepted: December 18, 2023
Published: January 10, 2024

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

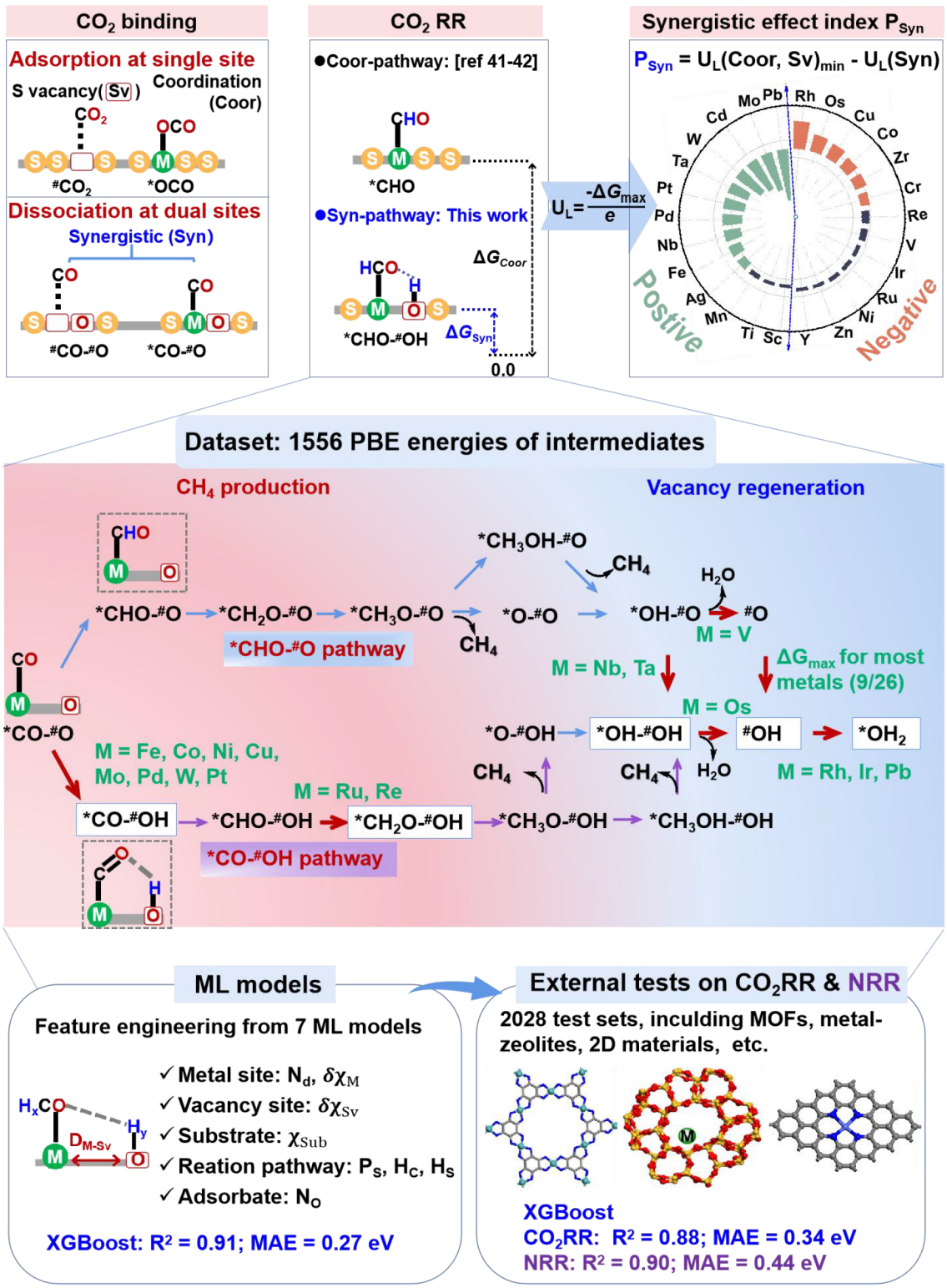


Figure 1. Schematic illustrations of the ML models for predictions of synergistic effect, reaction pathway, free energy change and limiting potential for CO_2 reduction on metal-anchored MoS_2 with and without sulfur vacancy, transferring to the external tests on other systems for both CO_2RR and NRR.

Machine learning (ML) models have been applied to establish the relationship between the reaction activity and catalyst structures.[24−26] The descriptors used for catalytic activity prediction on the CO_2RR are collected in Table S1 of Supporting Information. Among them, some electronic structure descriptors need the aid of density functional theory (DFT) calculations. Instead, there are also some easily available descriptors, such as d-shell valence electron numbers, electron type and number, electronegativity, enthalpy of vaporization, coordination number, and coordination bond

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

length of metal atom to the nearest neighbor atoms.[27−35] Moreover, some DFT-calculated parameters, such as charge transfer, catalyst mode, and work function, could be replaced by easily available descriptors.[36−39] As an illustration, the charge transfer between the catalyst and the adsorbates can be predicted by the formulated atomic ionization energy and electronegativity parameters.[40] To the best of our knowledge, the ML model for describing the interplay between the vacancy and the metal sites has rarely reported.

In this work, the ML models for free energy change (ΔG) and limiting potential (U_L) prediction will be trained from 1556 intermediate energies of 26 kinds of metal-anchored MoS_2 with and without sulfur vacancy (metal@Sv-MoS_2 and metal@MoS_2) on the CO_2RR. As illustrated in Figure 1, after the CO_2 adsorption on the surface, there were two kinds of reaction pathways, called the coordination pathway (Coor-pathway, Figure S1) and synergistic pathway (Syn-pathway, Figure S2), respectively. The Coor-pathway is well recognized for single-atom catalysts (SACs).[41,42] In contrast, the Syn-pathway is usually found for dual-atom catalysts.[43] It will be demonstrated that the vacancy generation steps of $^{\#}O \rightarrow {}^{\#}OH$ are the potential-determining step (PDS) for most kinds of metal@Sv-MoS_2 in this work. Abbreviations used in this article were listed in Table S2 for the catalysts, Table S3 for the calculated limiting potential of CO_2RR intermediates on different pathways, and Table S4 for descriptors used in ML models.

To quantitatively describe the synergistic effect between the metal center and vacancy, the synergistic index, P_{Syn}, is defined to assess the thermodynamic advantages of competing reaction pathways. Some easily accessible features for describing active sites on different substrates and reaction intermediates will be found to have good applicability in prediction of free energy changes of both the CO_2RR and nitrogen reduction reaction (NRR) on 2028 external tests, including metal−zeolites, metal−organic frameworks (MOFs), and 2D materials. The consistency between experimental data and the ML-predicted limiting potential is also exhibited in the synthesized oxygen and tungsten vacancies, which contain with evident synergistic advantage.

■ COMPUTATIONAL METHODS

The first-principles DFT calculations were performed using the Vienna Ab initio Simulation Package (VASP) with the projector-augmented wave (PAW) method.[44,45] Theoretical calculations for the exchange-correlation potentials were based on the generalized gradient approximation (GGA) method in the form of Perdew−Burke−Ernzerhof (PBE) functional.[46−48] The effects of the long-range van der Waals (vdW) corrections were treated using the Grimmes zero-damping DFT-D3 method.[49,50] The plane wave cutoff energy was set to 450 eV. The model of MoS_2 was built in a 5 × 5 × 1 supercell. The vacuum thicknesses were set to 15 Å to reduce interactions between adjacent layers in periodically repeated cells. The spin polarization was taken into consideration. The 2D Brillouin zone was sampled using a 2 × 2 × 1 Monkhorst−Pack k-point grid in reciprocal space. The convergence thresholds for the force and total energy component were set as 4.0×10^{-2} eV/Å and 1.0×10^{-5} eV, respectively.

The binding energies (E_b) of metal anchored MoS_2 with and without sulfur vacancy were evaluated according to eqs 1 and 2:

$$E_b = E_{\text{metal@Sv-MoS2}} - E_{\text{Sv-MoS2}} - E_{\text{metal}} \quad (1)$$

$$E_b = E_{\text{metal@MoS2}} - E_{\text{MoS2}} - E_{\text{metal}} \quad (2)$$

where $E_{\text{metal@Sv-MoS2}}$, $E_{\text{metal@MoS2}}$, $E_{\text{Sv-MoS2}}$, E_{MoS2}, and E_{metal} represent the calculated energies of metal−supported catalysts (metal@Sv-MoS_2 and metal@MoS_2), substrate without metal atom (Sv-MoS_2 and MoS_2), and isolated metal atom in vacuum. When $E_b < 0$ eV, the metal@Sv-MoS_2 and metal@MoS_2 were assumed to be thermodynamically favorable.

The computational hydrogen electrode (CHE) model proposed by Nørskov and co-workers was employed to estimate the thermodynamic free energies of the fundamental reactions.[51,52] Based on this method, the change in Gibbs free energy (ΔG) for the elementary steps in the CO_2RR was obtained by following eq 3. The relative free energy, ΔG, of each reaction intermediate was referred to as energy of the free CO_2 molecule. The free energy difference, $\Delta\Delta G$, was referred to as the free energy change between two successive steps on the CO_2RR process.

$$\Delta G = \Delta E + \Delta E_{ZPE} + T\Delta S + eU + \Delta G_{pH} \quad (3)$$

In the above equation, ΔE represents the electronic energy difference between two reaction species, the ΔE_{ZPE} represents zero-point energy change, which can be obtained by vibrational frequencies. The $T\Delta S$ represents the entropy change, and the simulation temperature T is 298.15 K. The eU term represents the electron transfer with the electrode potential contribution. The last term in eq 3, $\Delta G_{pH} = 2.303k_BT \times pH$, where k_B is the Boltzmann constant. The experimental conditions in acidic medium with pH = 0 and 3 were taken into consideration.

The corresponding limiting potential (U_L) was determined by the potential-determining step, which was the maximum free energy change of the whole process (ΔG_{max}). Therefore, the U_L represents the minimum applied potential to overcome the barrier in the CO_2RR, which was obtained by eq 4.

$$U_L = \Delta G_{max}/e \quad (4)$$

Mo@Sv-MoS_2 was used to test the performance of different DFT functionals, including PBE, Perdew−Wang 91 (PW91), and local density approximation (LDA), as shown in Figure S3 (PBE, PW91, and LDA functionals calculated by using $DMol^3$),[53,54] Figure S4 (PBE, PW91, and LDA functionals calculated using VASP), and Table S5. The reaction step for conversion from *CO−$^{\#}$O to *CO−$^{\#}$OH was predicted to be the potential-determining step for all of the selected DFT functionals. In VASP calculations, the absence of Grimmes zero-damping DFT-D3 corrections on the PW91 and LDA functionals results in large ΔG_{max} values. $DMol^3$ calculations show moderately small ΔG_{max} values using the PBE, PW91, and LDA functionals. It can be assumed that there was little influence from the DFT functional selection on the subsequent machine learning of the DFT-calculated results.

All ML algorithms were conducted by the open-source code Scikit-learn and PyTorch package in the Python3 environment.[55] As shown in Table S6, ten algorithms, including Extreme gradient boosting regression (XGBoost), Gradient boosting regression (GBR), Extra-Trees, Decisiontree (DT), k-nearest neighbor (kNN), Linear ridge, Least absolute shrinkage and selection operator (LASSO), Multiple Linear regression (MLR), Artificial neural network (ANN), and Support vector regression (SVR), were applied for free energy change and limiting potential prediction. To make sure of the generalization and accuracy, the collected data obtained from DFT calculations and external literature were randomly shuffled and divided into the training set and test set in a ratio of 80:20. The normalization preprocessing was applied for the ML model in training and prediction. Three indexes were selected to evaluate prediction errors, the mean absolute error (MAE), the root-mean-square error (RMSE), and the coefficient of determination values (R^2), as described in the eqs 5−7.

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} |y_i - \hat{y}_i| \quad (5)$$

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

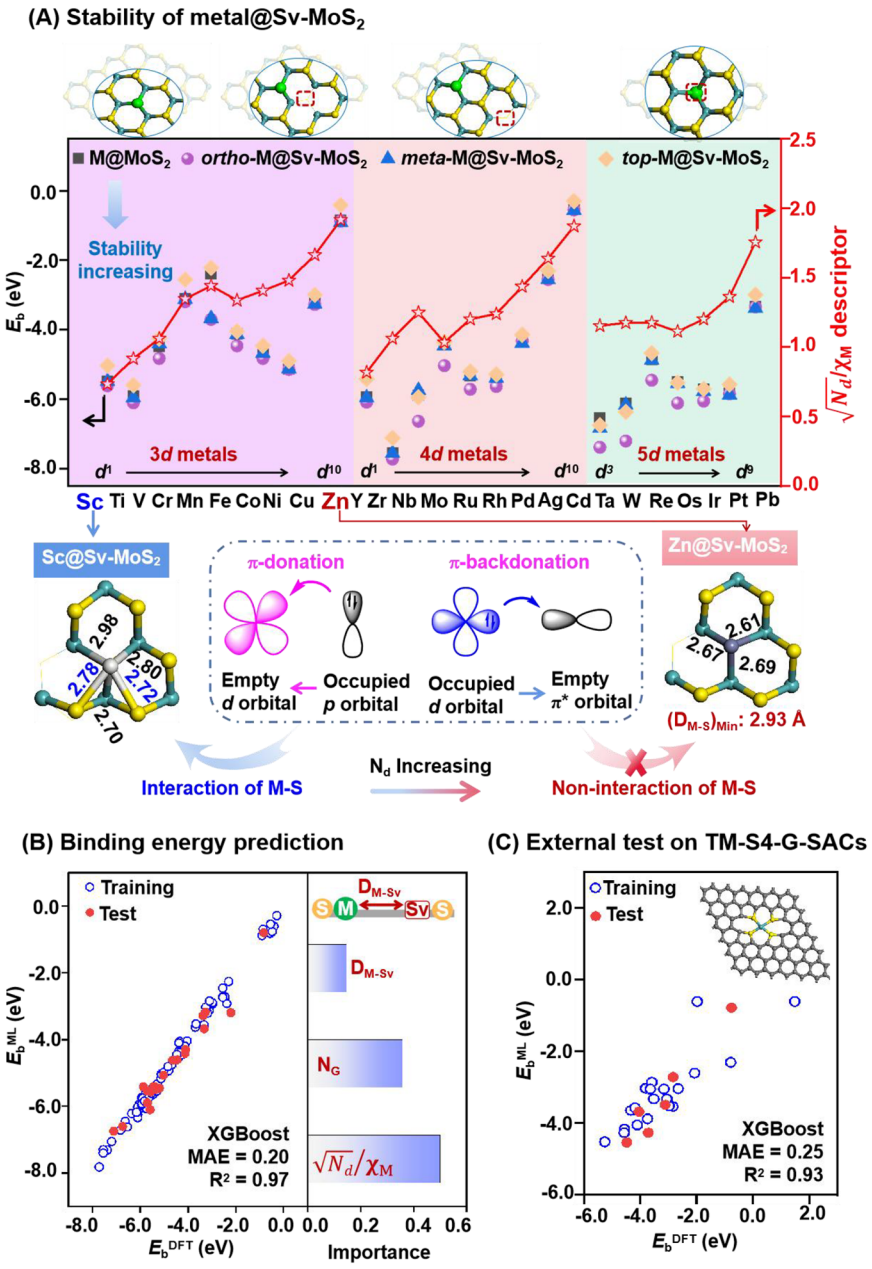


Figure 2. (A) Calculation results for the stability prediction of metal@Sv-MoS_2. The distances between metal and neighboring atom were given in Å. (B) Performance of XGBoost model for the prediction of binding energy using three features. (C) External test on TM-S4-G-SACs for the prediction of binding energy.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{1}^{n}(y_i - \hat{y}_i)^2} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2} \tag{7}$$

In the above equations, y_i, $\hat{y}_i$, and $\bar{y}_i$ mean the ith predicted value by ML, the DFT calculation result, and the average value of DFT calculated results, respectively.

RESULTS AND DISCUSSION

Stability Predictions

The DFT calculations indicated that the studied 26 metal@Sv-MoS_2 systems were thermodynamically stable, with negative binding energies (E_b = −7.72 to −0.55 eV, Figure S5). As shown in Figure S5A, the thermodynamic stability of metal@Sv-MoS_2 decreased with the number of d-shell valence electrons number (N_d) of d^5 and d^{10}. Similarly, the 26 selected metal-anchored MoS_2 without a sulfur vacancy (metal@MoS_2) also exhibited negative binding energies (E_b = −7.55 to −0.48 eV, Figure S5B), demonstrating the thermodynamic stability of the metal@MoS_2.

As depicted in Figure 2A, on the metal@Sv-MoS_2 substrate, the sulfur vacancy (Sv) may have three possible positions: (1) Sv located adjacent to the metal (*ortho*-M@Sv-MoS_2), (2) Sv situated in the meta position of the metal (*meta*-M@Sv-MoS_2), and (3) Sv positioned on the top position of the metal (*top*-M@Sv-MoS_2). Three different vacancy configurations were described by the different values of distance between metal and

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

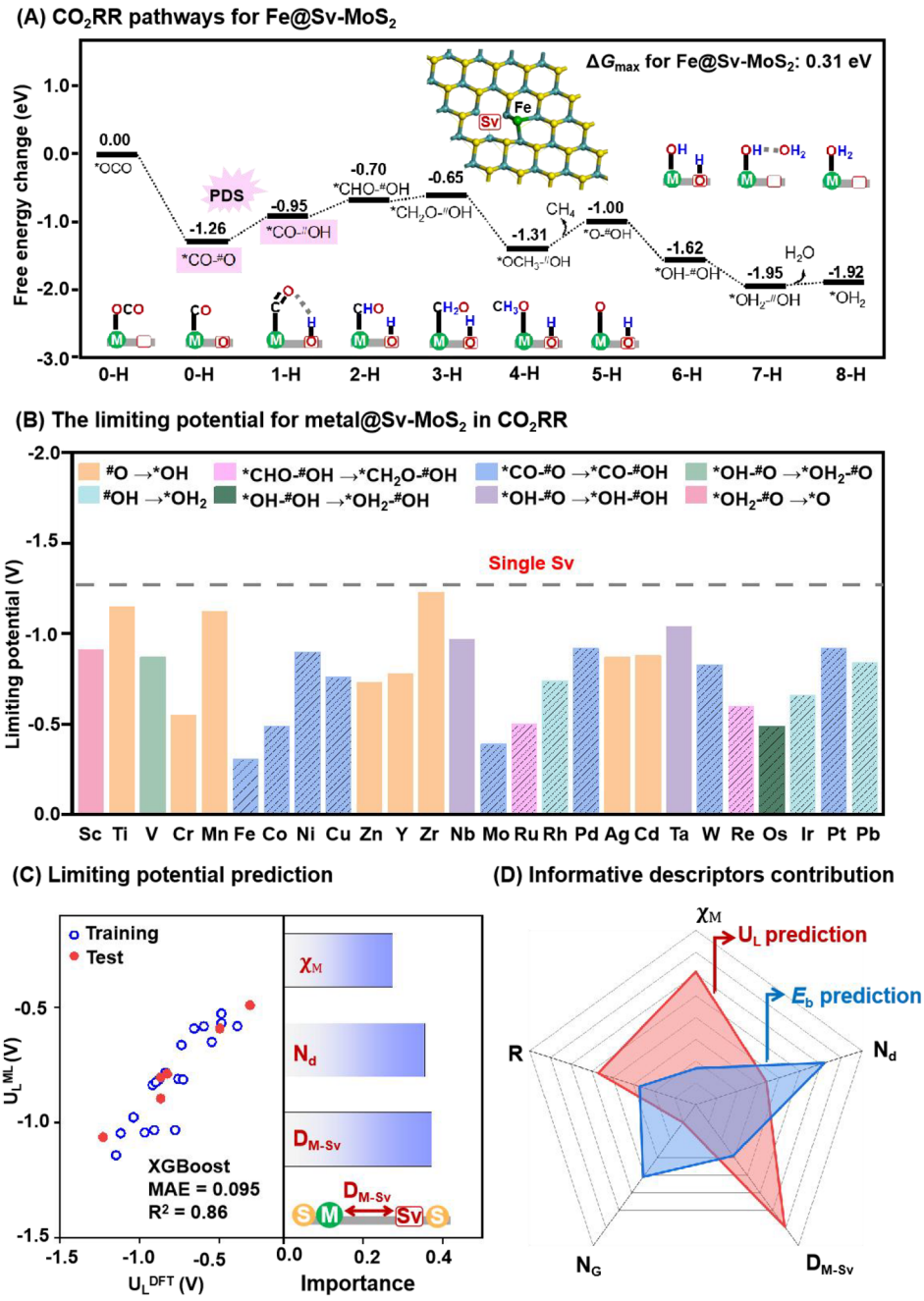


Figure 3. (A) Illustrations of the CO_2RR pathways, the free energy change, ΔG, and the PDS in Fe@Sv-MoS_2. (B) The limiting potential for metal@Sv-MoS_2 in the CO_2RR, the lines represent the *CO−OH pathway, the rest for the *CHO−O pathway. (C) The prediction of the limiting potential (U_L) in the CO_2RR by XGBoost and the feature importance of the XGBoost model. (D) Evaluate the contribution of information descriptors to binding energy (E_b) and limiting potential (U_L) predictions.

vacancy, named as $D_{M\text{-}Sv}$ in inset of Figure 2B. The relative stability and binding energies for the Sv at three possible positions were given in Figures S6 and S7, respectively. The configurations for *ortho*-M@Sv-MoS_2 ($D_{M\text{-}Sv}$ ranging from 2.41−3.25 Å) was more stable than that of *meta*-M@Sv-MoS_2 ($D_{M\text{-}Sv}$ ranging from 5.88−6.37 Å), and *top*-M@Sv-MoS_2 ($D_{M\text{-}Sv}$ ranging from 3.27−4.03 Å, Figure 2A). It can be conceived that the p-type lone pair electrons of the sulfur atom favor coordination with the early transition metal. In turn, the d electrons in transition metals easily form the π-backbonding between metal center and sulfur on metal@Sv-MoS_2. Taking Sc@Sv-MoS_2 as an example, the Sc atom ($3d^1$) exhibits a significant contribution to the projected density of state (PDOS), with π-bonding and π-backbonding bands appearing near the Fermi level (−0.3 to 0.9 eV), as shown in Figure S8A. In contrast, the $3d^{10}$ Zn atoms make a negligible contribution to the electronic states in Zn@Sv-MoS_2 (Figure S8B). The penta-coordinated Sc@Sv-MoS_2 displays a higher oxidation state (Q_{Sc}: 0.88 e) than tricoordinated Zn@Sv-MoS_2 (Q_{Zn}: 0.12 e). We expected that the easily accessible parameter of electronegativity of metal (χ_M) could replace the DFT-calculated values of atomic charge, as demonstrated in previous work.[40] Here, in addition to the value of $D_{M\text{-}Sv}$, the ratio of d-shell valence electrons to the electronegativity of metal center (

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

$\sqrt{N_d}/\chi_M$) was found to be correlated with the stability (the right side in Figure 2A). It is also displayed that with the increase in the number of groups (N_G) of metal atoms in the periodic table, the stability has slightly increased.

Using the above-mentioned three descriptors ($D_{M\text{-}Sv}$, $\sqrt{N_d}/\chi_M$, and N_G), the ML model was then applied to rapidly predict the binding energy for metal-anchored MoS_2 with and without sulfur vacancy (M@MoS_2, *ortho-*, *meta-*, *top*-M@Sv-MoS_2). As illustrated in Figure 2B, the XGBoost model gives a good prediction (MAE = 0.20 eV, R^2 = 0.97) of binding energy values and reaches state-of-the-art results via 10-fold cross-validation (Figure S9). The external test was also conducted on TM-S4-G-SACs[S6] systems using the same 3-feature sets of $\sqrt{N_d}/\chi_M$, N_G, and $D_{M\text{-}Sv}$, yielding satisfactory binding energy prediction results in Figure 2C (MAE = 0.25 eV, R^2 = 0.93).

Adsorption Selectivity Tests

The adsorption energies of CO_2 on metal@Sv-MoS_2 ranged from −0.13 to −0.64 eV, indicating effective capture of CO_2 in *OCO configurations (Figure S10). The CO_2 adsorption energies on the metal active site showed relatively strong adsorption compared to the van der Waals adsorption energies on single sulfur vacancy and double sulfur vacancies (−0.20 eV for both single-Sv and double-Sv), indicating that metal sites enhance CO_2 adsorption and facilitate its activation. The sulfur vacancy on the MoS_2 surface affects the charge distribution around the metal center, promoting diffusion of CO_2 to the vacancy. The synergistic interaction between the Sv and metal center contributes to CO_2 dissociation on metal@Sv-MoS_2 (for the step of *OCO → *CO−#O, $\Delta\Delta G$ was in the range of −2.29 to −0.71 eV), suggesting that the metal center and Sv could lead to spontaneous CO_2 dissociation. The resultant CO molecule was easily captured by metal center, accompanied by the oxidization of sulfur vacancy. Note that the strong CO adsorption on the catalysts leads to further reduction to CH_4 instead of desorption, due to the relatively high desorption energy ($\Delta\Delta G$ = 0.17−1.64 eV).

An efficient CO_2RR catalyst should suppress competing hydrogen evolution reactions (HER) to achieve a high Faraday efficiency. The difference in adsorption energies between *OCO and *H was investigated to understand the competition between the CO_2RR and HER. As depicted in Figure S11, most of the investigated metal@Sv-MoS_2 can avoid *H poisoning, suggesting a high selectivity for the CO_2RR over the HER. The exceptions were found for the Os@Sv-MoS_2 and Ir@Sv-MoS_2 systems. In the following hydrogenation step, early transition metals (M = Sc, Ti, V, Cr, Mn, Y, Zr, Nb, and Ta) and d^{10} metals (M = Zn, Ag, and Cd) predominantly formed the *CHO−#O intermediate, while other kinds of metal@Sv-MoS_2 preferred to produce the *CO−#OH intermediate (Figure 1 and Figure S2). The formation of hydrogen bonding played a crucial role in lowering the reaction energies with hydrogen bonding distance ranging from 1.54−2.82 Å on *CO−#OH, except for d^{10} metal, Cu-, Zn-, and Cd@Sv-MoS_2 (Figure S12). The free energy change for metal@Sv-MoS_2 adsorbed *CHO−#O or *CO−#OH species was much smaller than that of *H, highlighting the potential of metal@Sv-MoS_2 as a highly selective CO_2RR catalyst (Figure S11).

Activity Evaluation

Figure 1 displays two possible pathways for the CO_2RR (*CHO−#O and *CO−#OH). Subsequent hydrogenation reactions occur primarily at the metal active site, followed by the vacancy site. In the *CHO−#O pathway, the d^5 and d^{10} metal@Sv-MoS_2 (M = Mn($3d^5$), Zn($3d^{10}$), Ag($4d^{10}$), and Cd($4d^{10}$)) preferentially produce *CH_3OH−#O species on the fourth hydrogenation, while other metal@Sv-MoS_2 tend to form *O−#O with the release of CH_4. The cases containing late transition metals and main group metal (Pb) prefer the *CO−#OH pathway, where consecutive protonation at the C atom leads to the formation of the *CH_3O−#OH species. In the fifth hydrogenation, transition metals with more d-shell valence electrons number, the metal@Sv-MoS_2 (M = Co($3d^7$), Ni($3d^8$), Cu($3d^{10}$), Rh($4d^8$), Pd($4d^{10}$), and Pt($5d^9$)) were more likely to generate *CH_3OH−#OH species, while other metal@Sv-MoS_2 tends to form *O−#OH, with the release of CH_4 (Figure S13). In most metal@Sv-MoS_2 systems (Figure 3), the hydrogenation step for converting *CO−#O to *CO−#OH (when M = Fe−Cu, Mo, Pd, W, and Pt) and #O to #OH (when M = Ti, Cr, Mn, Zn−Zr, Ag, and Cd) is the potential-determining step during vacancy regeneration.

The Syn-pathway network seems to be more concise than the Coor-pathway network. The synergistic effect was observed between metal center and Sv in three ways: (1) providing C-affinity (metal center) and O-affinity sites (Sv) for CO_2RR, (2) reducing the possible reaction pathways, intermediates, and products, and (3) improving the selectivity for CH_4 production by suppressing HER process.

Taking Fe@Sv-MoS_2 as an example, the CO_2RR on Fe@Sv-MoS_2 follows the *CO−#OH pathway (Figure 3A). The C−O dissociation corresponds to a ΔG of −1.26 eV. During the hydrogenation of CO_2, the conversion of *CO−#O to *CO−#OH was the potential-determining step, with a maximum free energy change (ΔG_{max}) of 0.31 eV. This value was lower than those reported for SACs, such as Fe-supported graphitic carbon nitride (Fe@N_4−C: 0.74 eV, Fe@N_4O−C: 1.02 eV),[29] Fe-zeolites (0.96 eV),[40] and Fe-MOF (0.69 eV)[57] for CH_4 or CH_3OH production, demonstrating the superior activity of metal@Sv-MoS_2 for CO_2RR. The calculated values of limiting potential range from −0.31 to −1.23 V (Figure 3B). Specifically, metal@Sv-MoS_2 has a lower limiting potential value compared to those of single-Sv (−1.29 V) and double-Sv (−1.34 V, Figure S14). It is worth noting that the limiting potential values for metal@MoS_2 were in a range from −0.29 to −1.55 V (Figures S15−S16). Without consideration of vacancy regeneration, the limiting potential values for metal@Sv-MoS_2 ranged from −0.13 to −1.04 V, as plotted in Figure S17. Although vacancy regeneration often occurs during electrocatalytic reactions,[S8] it is important to note that the metal active sites can continue to catalyze CO_2 reduction regardless of whether vacancy was regenerated or not. Therefore, metal@Sv-MoS_2 seems to display superior activity for CH_4 production.

As mentioned above, the distance between the metal center and sulfur vacancy, $D_{M\text{-}Sv}$, is an important descriptor for reflecting stability, electronic structures, and the catalytic activity of the CO_2RR. Again, the feature of $D_{M\text{-}Sv}$ is a prerequisite for evaluating the synergistic effect. For example, when the sulfur vacancy is located in the meta position of the metal ($D_{M\text{-}Sv}$ > 5.88 Å), no synergistic effect was observed between the sulfur vacancy and the metal site (Figure S18). When the sulfur vacancy located in the ortho position of the metal ($D_{M\text{-}Sv}$ < 3.25 Å), the limiting potential value was reduced as $D_{M\text{-}Sv}$ increased, indicating a weakening of the synergistic effect between the sulfur vacancy and the metal site

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

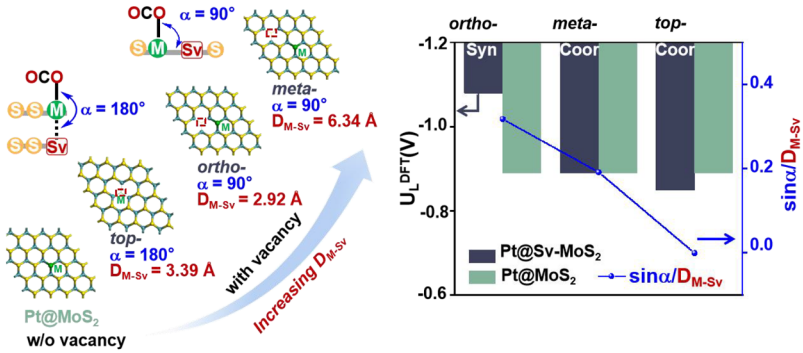


Figure 4. Relationship of limiting potential with orientation angle, α, and distance between metal and vacancy, $D_{\text{M-Sv}}$.

(Figure S19). A ML model was constructed using three features, i.e., N_d, χ_M, and $D_{\text{M-Sv}}$, to predict the limiting potential value for 26 metal@Sv-MoS_2 systems. According to Figure 3C, the XGBoost model demonstrated satisfactory prediction results for limiting the potential value, with a MAE of 0.095 V and a R^2 of 0.86.

Subsequently, we attempted to test the performance of using a combination of physical parameters in two ways. In the first test, a complex descriptor, $F_{\text{M-Sv}}$, was proposed to describe limiting potential, taking into account the contribution of both the metal site $\left(\frac{N_d+\chi_M}{N_d-\chi_M}\right)$ and the vacancy site $\left(\exp(D_{\text{M-Sv}}) - \exp\left(\frac{\chi_M}{R}\right)\right)$, as defined in eq 8.

$$F_{\text{M-Sv}} = \left(\frac{N_d+\chi_M}{N_d-\chi_M}\right) + \left(\exp(D_{\text{M-Sv}}) - \exp\left(\frac{\chi_M}{R}\right)\right) \tag{8}$$

Here, R represents the radius of the metal atom. With the addition of the R into the three feature sets of N_d, χ_M, and $D_{\text{M-Sv}}$, the XGBoost model displayed comparable prediction accuracy (MAE = 0.11 V and R^2 = 0.83) to the 3-feature model (Figure S20). Our results show that reaction activity could be evaluated by four easily obtainable descriptors or their combination in eq 8, with a Pearson correlation coefficient of 0.82 (Figure S21).

The second scheme for building the complex descriptors is stimulated by an investigation of key reaction intermediates. In Coor-pathway, the adsorption free energies of carbon-involved (Figure S22) and oxygen-involved (Figure S23) intermediates were strongly correlated with $\Delta G_{*\text{CHO}}$ and $\Delta G_{*\text{OH}}$, respectively. The values of $\Delta G_{*\text{CHO}}$ and $\Delta G_{*\text{OH}}$ served as indicators to describe the limiting potential associated with the CO_2RR activity (Figures S24 and S25). The sure independence screening and sparsifying operator (SISSO) method[59] was employed to yield two descriptors, F_{CHO} and F_{OH}, which were defined as follows:

$$F_{\text{CHO}} = (N_d - \chi_M)EV^2/e^{(N_d/\chi_M)} \tag{9}$$

$$F_{\text{OH}} = \text{IE}^* N_d/\chi^* \sqrt[3]{e^{N_d}} \tag{10}$$

In the above equations, EV and IE represent the enthalpy of vaporization and the first ionization energy of metal, respectively. The predictions for $\Delta G_{*\text{CHO}}$ and $\Delta G_{*\text{OH}}$ were also satisfactory, with Pearson correlation coefficients of −0.85 and −0.91, respectively (Figure S26).

The usage of the two complex features, F_{CHO} and F_{OH}, and the five descriptor sets of R, N_G, N_d, χ_M, and $D_{\text{M-Sv}}$, gives a good correlation to binding energy and limiting potential. Among them, the distance between the metal center and sulfur vacancy ($D_{\text{M-Sv}}$) has the most significant impact on the CO_2RR activity. The d-shell valence electrons of metal (N_d) has the dominant influence on catalyst stability (Figure 3D). The important role of d-shell valence electrons and electronegativity of metal in stability prediction was also demonstrated in metal−zeolites.[60]

Furthermore, the influence of the relative orientation of the vacancy to the metal was investigated. As shown in Figure 4, different positions of Sv on Pt anchored the Sv-MoS_2 (*ortho*-, *meta*-, and *top*-Pt@Sv-MoS_2), are indicated by the angle, α, between adsorbate on the upmost metal atom and the Sv vacancy site. In the case of *top*-Pt@MoS_2, the angle α is 180° with the linear structure of O···M···Sv. For *ortho*-Pt@Sv-MoS_2 and *meta*-Pt@Sv-MoS_2, the angle α is about 90°. The *top*- and *meta*-Pt@Sv-MoS_2 systems prefer to follow the Coor-pathway (Figure S18). The *ortho*-Pt@Sv-MoS_2 favors the Syn-pathway, which exhibits relatively lower free energy changes of the potential-determining step ($\Delta G_{\max}$ = 0.92 eV) compared to the *top*- ($\Delta G_{\max}$ = 1.15 eV) and *meta*- ($\Delta G_{\max}$ = 1.12 eV) cases. Interestingly, the limiting potential decreases as the value of $\sin\alpha/D_{\text{M-Sv}}$ decreases, highlighting the importance of orientation and distance between the vacancy and metal sites. Here, the use of the sinusoidal function of angle α was capable of transferring the 90° or 180° into the values of 1 and 0, which is much closer to the value of $D_{\text{M-Sv}}$.

Synergistic Effect

To efficiently assess the synergistic effect of metal@Sv-MoS_2 on the CO_2RR, a synergistic effect index, P_{Syn}, was defined to distinguish the activity difference between metal@Sv-MoS_2, metal@MoS_2, and single sulfur vacancy on MoS_2 in eq 11.

$$P_{Syn} = U_L(\text{Coor, Sv})_{\min} - U_L(\text{Syn}) \tag{11}$$

The term $U_L(\text{Coor, Sv})_{\min}$ represented the minimum limiting potential in CO_2RR through either the Coor-pathway for metal@MoS_2 or single sulfur vacancy pathway (Sv-pathway) for vacancy only. The U_L(Syn) represented the limiting potential via the Syn-pathway without considering vacancy regeneration.

When P_{Syn} was greater than 0.10 V, the Coor-pathway was more favorable, indicating a negative synergistic effect on the CO_2RR. When P_{Syn} was within the range of −0.10 to 0.10 V, the Coor-pathway and Syn-pathway were comparable without the synergistic effect in the CO_2RR. When the value of P_{Syn} was smaller than −0.10 V, the Syn-pathway was more favorable, indicating a positive synergistic effect on CO_2RR.

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

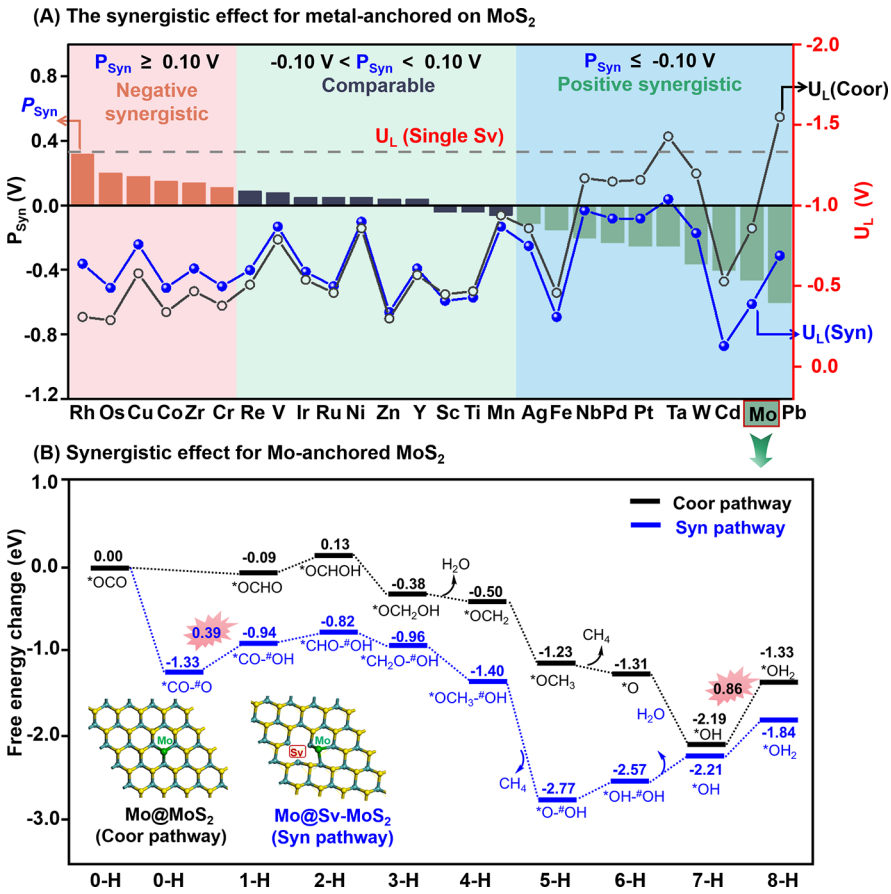


Figure 5. (A) Definition of synergistic effect evaluation index (P_{Syn}) to distinguish the synergistic effect on the Coor-pathway, Sv-pathway, and Syn-pathway in CO_2RR. (B) Illustration of the synergistic effect for Mo-anchored MoS_2.

Figure 5A illustrates that the Syn-pathway was more favorable for d^5 (M = Mo) and d^{10} (M = Ag, Cd, and Pd) metal@Sv-MoS_2 in the CO_2RR. For example, in Mo@Sv-MoS_2, the conversion of *CO−#O to *CO−#OH was calculated as the potential-determining step with a ΔG_{max} of 0.39 eV. The ΔG_{max} of Mo@MoS_2 was calculated to be 0.86 eV for the conversion of *OH to *OH_2 (Figure 5B). The ΔG_{max} of the Sv-only pathway on MoS_2 was calculated to be 1.29 eV for the conversion of *O to *OH, implying a positive synergistic effect in Mo@Sv-MoS_2 for CH_4 production.

The external potential and pH value of the solution could change the thermodynamics and electrochemical potential-energy diagrams, since the polar intermediates are sensitive to pH values on CO_2RR.[61] As shown in Figure S27, both Fe@Sv-MoS_2 and Os@MoS_2 required less additional electric field to achieve a high conversion efficiency for CO_2. Increasing the pH value from pH = 0 to pH = 3, the limiting potential for Fe@Sv-MoS_2 (pH = 0, U_L = −0.31 V) and Fe@MoS_2 (pH = 0, U_L = −0.46 V) decreases to −0.49 and −0.64 V, respectively (Figure S28). In fact, the synergistic effect between the metal center and the vacancy (P_{Syn} = −0.15) was not varied by the pH change from 0 to 3.

Free Energy and Limiting Potential Prediction and External Tests

It was demonstrated that the adsorption free energies of key intermediates, such as ΔG_{*OCH2}, ΔG_{*CHO}, and ΔG_{*OH} are highly relevant to the catalytic performance of single-active site catalysts in CO_2RR.[28,40,62] Until now, the synergistic effect of the active metal site and vacancy in adsorption free energies was rarely studied yet. Herein, taking single-active site catalysts (metal@MoS_2) and dual-active site catalysts (metal@Sv-MoS_2) as a prototype, we built ML models for predicting the free energy change of intermediates. To achieve this goal, we defined 3 descriptors for active sites on substrates, including electronegativity difference of metal center and vacancy ($\delta\chi_M$, $\delta\chi_{Sv}$) and global electronegativity on substrates (χ_{sub}). The $\delta\chi_M$ and $\delta\chi_{Sv}$ were applied to reflect the interaction strength for the active site with various intermediates. The descriptor $\overline{\chi_M}$, through the combination of metal center with the nearest coordination atoms, was constructed for the reactive sites. This strategy was applied in our previous work[40] and other literature.[29] The value of $\delta\chi_M$ between catalyst and reduction intermediates in each step of CO_2RR process was easily estimated from eq 12 by the abstraction from $\overline{\chi_M}$ with the average electronegativity in absorbed intermediate ($\overline{\chi_{Int}}$).

$$\delta\chi_M = \overline{\chi_M} - \overline{\chi_{Int}} = \frac{\chi_M + \chi_{Mo}}{\chi_{Mo}} - \frac{\sum_{i=1}^{N_{atom}} N_i\chi_i}{N_{atom}} \quad (12)$$

The descriptor developed for the vacancy site ($\delta\chi_{Sv}$) was calculated by the joint effect of the absorbed intermediate with the nearest neighboring atoms and sulfur vacancy deductions. Furthermore, we added the distance between the metal center and vacancy, $D_{M\text{-}Sv}$, via eq 13 to evaluate the synergistic effects between the metal and vacancy.

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

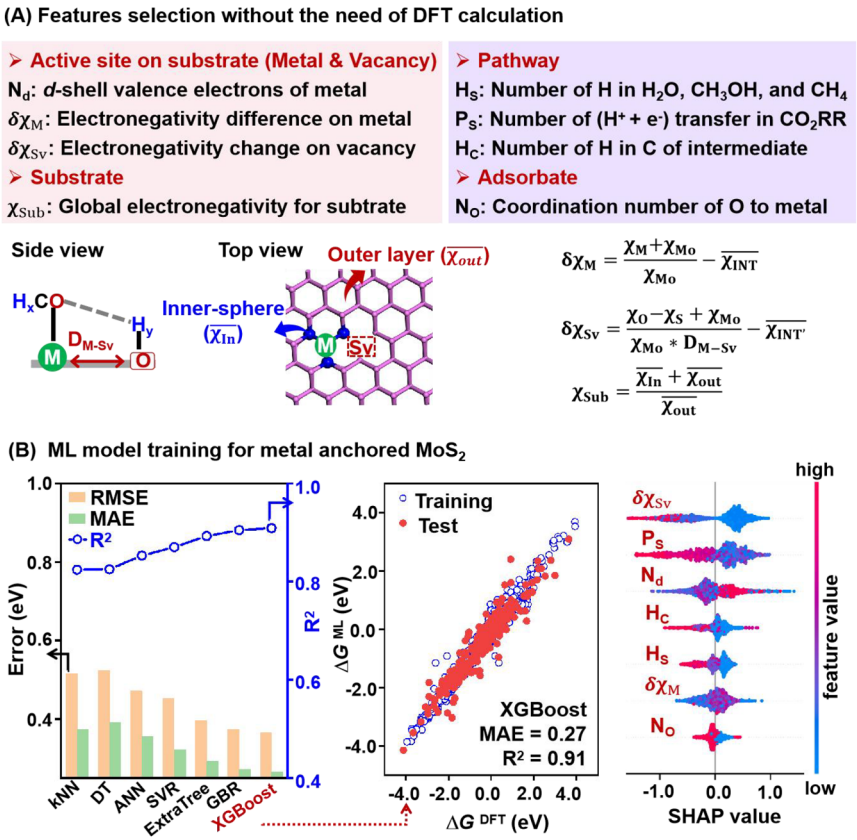


Figure 6. (A) The definition and construction of 8 descriptors in the CO_2RR process. (B) ML model construction for ΔG prediction and comparison of the performance of different algorithms; prediction of ΔG of the CO_2RR process by the XGBoost algorithm; and SHAP values of each sample on CO_2RR.

$$\delta\chi_{Sv} = \frac{\chi_O + \chi_{Mo} - \chi_S}{\chi_{Mo} * D_{M-Sv}} - \frac{\sum_{j=1}^{N_{atom}} N_j\chi_j}{N_{atom}} \quad (13)$$

The global electronegativity of the substrate (χ_{sub}) could reflect the interaction strength between the substrate and the metal center. The χ_{sub} value was estimated from eqs 14 and 15 by dividing the substrate into two groups, the inner sphere consisting of the coordinate atoms (N, O, etc.) in proximity to the metal center ($\overline{\chi_{In}}$) and the out-layer ($\overline{\chi_{Out}}$) in substrate.

$$\overline{\chi_{Out}} = \frac{\sum_{k=1}^{N_{atom}} N_k\chi_k}{N_{atom}} \quad (14)$$

$$\chi_{sub} = \frac{\overline{\chi_{In}} + \overline{\chi_{Out}}}{\overline{\chi_{Out}}} = \frac{\chi_{N/O} + \overline{\chi_{Out}}}{\overline{\chi_{Out}}} \quad (15)$$

For dual-active sites, we defined the average electronegativity value of the metal center site (χ_{sub}) and vacancy (χ_{sub-Sv}) substrate in eqs 16−18.

$$\overline{\chi_{Out-Sv}} = \frac{\sum_{k=1}^{N_{atom}} N_k\chi_k}{N_{atom}} \quad (16)$$

$$\chi_{sub-Sv} = \frac{\overline{\chi_{In'}} + \overline{\chi_{Out-Sv}}}{\overline{\chi_{Out-Sv}}} = \frac{\chi_{Mo} + \overline{\chi_{Out-Sv}}}{\overline{\chi_{Out-Sv}}} \quad (17)$$

$$\chi_{sub} = \frac{\chi_{sub} + \chi_{sub-Sv}}{2} \quad (18)$$

The free energy change in each hydrogenation step was predicted with an 8-feature scheme, including N_d, $\delta\chi_M$, $\delta\chi_{Sv}$, and χ_{sub} for active sites on substrates, P_S, H_C, and H_S for reaction pathways, and N_O for adsorbate (Figure 6A). A grid search was performed in 10 algorithms to determine the appropriate hyper-parameters that would achieve the highest accuracy. The cross-validation metric was selected by RMSE, MAE, and R^2 to obtain a more reliable estimate of the model's performance. The 10-fold cross-validation was applied on the selected 10 ML models, which shows the accuracy of the model on the entire volume of nonlinear data (Tables S7 and S8 and Figure S29). The XGBoost achieved the most satisfactory prediction results with MAE and R^2 values of 0.27 eV and 0.91, respectively (Figure 6B). Based on Pearson correlation analysis, a nonlinear correlation was observed between the features and free energy change (Figure S30). The XGBoost model calculated the average feature importance of each feature (Figure S31) by a 10-fold cross-validation, showing that P_S, H_S, H_C, and $\delta\chi_{Sv}$ were important features. Shapley additive explanations (SHAP)[63] algorithm was combined with the trained model to further analyze the positive or negative correlation between features and predicted results. As shown in Figure 6B and Figure S31, $\delta\chi_{Sv}$, P_S, N_d, H_S, and H_C were the top 5 features with relatively large SHAP values for the free energy change predication. The SHAP analysis reveals a negative correlation among $\delta\chi_{Sv}$, P_S, H_S, and H_C, while N_d exhibits a positive correlation. The XGBoost measures the importance of features based on information gain, while SHAP is a globally accurate additive method that

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

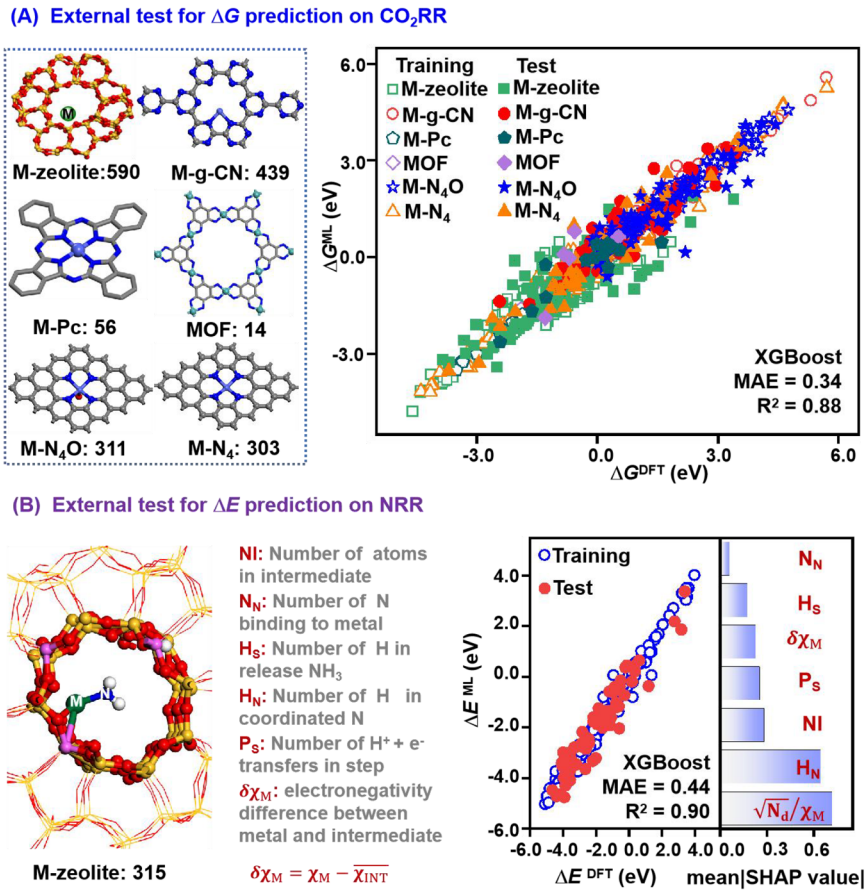


Figure 7. (A) External tests for ΔG prediction on CO_2RR via 8 features by metal−zeolites, M-*g*-CN, M-Pcs, MOFs, M-N_4O, and M-N_4 catalysts. (B) The university descriptors transfer for NRR relative energy change prediction on metal−zeolites with 7 features and feature importance; the absolute SHAP value of each feature was averaged in the data set, called mean |SHAP value|.

identifies the importance of most global features. The different method and parameter in XGBoost model and SHAP resulting in a different ranking of feature importance.

The SISSO[59] was used to generate the formulas describing the relationship between the features and free energy change. The feature reconstruction by SISSO for free energy change prediction based on the 1D, 2D, and 3D descriptors exhibits the results of R^2 = 0.41, 0.62, and 0.71, respectively (Figure S32). The synergistic effect between the metal center and vacancy can be described by the term $\delta\chi_M - \delta\chi_{Sv}$, which appeared in 2D and 3D descriptors from the SISSO method.

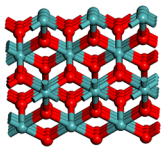
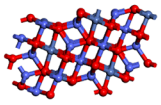
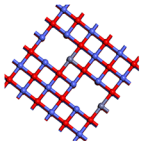
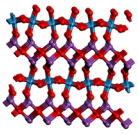
Other algorithms, including GBR (Figure S33), ExtraTree (Figure S34), ANN, SVR, and DecisionTree (DT), also showed satisfactory prediction performance with R^2 values greater than 0.80 and MAE values less than 0.40 eV (Figure S35). The SHAP analysis results based on different models, such as GBR, and ExtraTrees, were consistent with the XGBoost model in prediction of the free energy change (Figure 6B and Figures S36 and S37). Partial dependence plot[64] was used to compare with SHAP analysis. As shown in Figure S38, the range of $\delta\chi_{Sv}$ was large (the blue line was steeper), which has a significant impact on the free energy prediction. In addition, the distribution of SHAP values of P_S, N_d, H_S, and H_C has the same tendency for partial dependence plots.

Linear models, such as MLR, linear ridge, and LASSO, were subjected to a grid search to find the appropriate hyper-parameters. More detailed parameters of models can be found in Tables S9−S12. However, even after this optimization process, these models still yielded poor prediction results, with low R^2 values and high the mean absolute errors (R^2 < 0.5 and MAE > 0.7 eV, as shown in Tables S12 and Figure S39). As a result, the top three ensemble algorithms (XGBoost, GBR, and ExtraTree) were selected for the prediction of energy properties for both the CO_2RR and NRR in the following subsection. More importantly, the number of atoms in the intermediates (NI) can be used as an effective adsorbate descriptor to predict the free energy change, with MAE and R^2 values of 0.25 eV and 0.92, respectively (Figure S40).

Again, the three formulated descriptors, $F_{M\text{-}Sv}$, F_{CHO}, and F_{OH}, were applied for the reaction energy prediction. As shown in Figure 6 and Figure S41, The XGBoost performance with 7 descriptors (MAE = 0.27 eV, R^2 = 0.91) was comparable to the models with 8 descriptors (in which $F_{M\text{-}Sv}$ was added, MAE = 0.29 eV, R^2 = 0.90) and 9 descriptors (in which F_{CHO} and F_{OH} were added, MAE = 0.27 eV, R^2 = 0.89). The electronegativity of the active metal site plays a crucial role in predicting free energy changes in CO_2RR. Previous studies have demonstrated that the use of other electronic descriptors, such as band gap (E_g),[40] the lowest unoccupied and highest occupied molecular orbitals (E_{LUMO} and E_{HOMO}), the vertical ionization potential (IP),[65] etc., is efficient for predicting CO_2RR and radical reactions. By replacing the electronegativity in $\delta\chi_{Sv}$ and $\delta\chi_M$ with the first ionization energy (δIE_{Sv} and δIE_M), a comparable

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

Table 1. Application of the ML Model to Experimentally Reported Systems with Different Vacancy Characteristics

System	Experiments		DFT Results		ML prediction (This work)
	product	Selectivity	P_{Syn} ($\Delta G1_{max}^{DFT}/\Delta G2_{max}^{DFT}$)	PDS	P_{Syn} ($\Delta G1_{max}^{ML}/\Delta G2_{max}^{ML}$)
 Ov-MoO_{2-X}	CH_4	~80% ref [20]	-0.66 (1.10 / 0.44)	Pristine & Ov *CO → *CHO *COOH → *CO	-0.98 (1.34 / 0.36)
 Ov-$NiCo_2O_4$	CO	99% ref [21]	-0.66 (1.37 / 0.71)	Pristine & Ov *CO_2 → *COOH *CO_2 → *COOH	-0.72 (1.18 / 0.46)
 Ov-Zn-CoO	CH_4	96.3% ref [22]	-0.46 (0.85 / 0.39)	Pristine & Ov *CO_2 → *COOH *CO_2 → *COOH	-1.06 (1.38 / 0.33)
 Wv-BWO	CH_4	63.8% ref [23]	-0.75 (1.21 / 0.46)	Pristine & Wv *CO_2 → *COOH *CO_2 → *COOH	-0.72 (1.18 / 0.46)

prediction ability was achieved, with MAE = 0.29 eV and R^2 = 0.91 (Figure S42).

Different ratios for dividing the training set and test set (70:30, 80:20, and 90:10) were also investigated (Figure S43). The training set and test set randomly scrambled into 70:30 (MAE = 0.29 eV, R^2 = 0.88) and 90:10 (MAE = 0.26 eV, R^2 = 0.92, Figure S44) displayed performance comparable to that of 80:20 (MAE = 0.27 eV, R^2 = 0.91).

Such 8 descriptors can be further transferred to other metal−supported catalysts, both in two- and three-dimensional topologies. The combination of 1713 external tests, including metal−supported graphitic carbon nitride (M-*g*-CN),[28] porpyridinic-like M−N*x* complexes (M−N_4O and M−N_4),[29] metal−zeolites,[40] metal−phthalocyanines (M-Pcs),[66] and MOFs[57,67] on CO_2RR, were incorporated into the ML models to predict free energy change. Again, the XGBoost model demonstrated satisfactory performance with an MAE of 0.34 eV and R^2 of 0.88 (Figure 7A). The substrate descriptor displayed a relatively high feature importance, highlighting its significant role in the CO_2RR process (Figures S45 and S46). Additionally, the GBR (MAE = 0.34 eV, R^2 = 0.88) and ExtraTree (MAE = 0.41 eV, R^2 = 0.86) models also exhibited prediction performance comparable to that of XGBoost (Figure S47).

Furthermore, the proposed descriptors in this work can be utilized to evaluate the performance of metal−zeolites in nitrogen reduction reactions (NRR) with 315 data samples.[68] As depicted in Figure 7B, the XGBoost model exhibits a satisfactory prediction capability (MAE = 0.44 eV, R^2 = 0.90) for the relative energy value of the reaction intermediates (ΔE) with 7 descriptors. These descriptors include the metal center descriptor ($\sqrt{N_d}/\chi_M$), electronegativity difference between metal center and intermediate ($\delta\chi_M$), numbers of the nitrogen binding to metal center (N_N), number of hydrogen atoms on the coordinated nitrogen (H_N), number of proton−electron pair transfer in the reduction step (P_S), the number of atoms in the intermediates (NI), and number of hydrogen atoms in released NH_3 (H_S). The SHAP analysis based on the XGBoost model on NRR reveals that the complex metal center descriptor ($\sqrt{N_d}/\chi_M$) was the most significant feature in regulating the catalyst performance (Figure 7B and Figure S48). The performance was comparable to the reported XGBoost model (MAE = 0.51 eV, R^2 = 0.84)[68] trained from metal−zeolites. In addition, both GBR and ExtraTree models exhibit similar prediction capabilities (GBR: MAE = 0.44 eV and R^2 = 0.89; ExtraTree: MAE = 0.51 eV and R^2 = 0.87, Figures S49−S51) to XGBoost. Meanwhile, using the aforementioned 7 descriptors, the XGBoost model also provides a good prediction of energy difference between two successive steps ($\Delta\Delta E$) on the NRR process in metal−zeolites, with a MAE of 0.46 eV and an R^2 value of 0.83 (Figure S52).

The present ML model for predicting synergistic effect evaluation was applied to metal−supported catalysts with different types of vacancy sites. Table 1 shows the selected systems, such as Ov-MoO_{2-x},[20] Ov-$NiCo_2O_4$,[21] Ov-Zn-CoO,[22] and Wv-BWO,[23] which represent the influence of various vacancy characteristics, including type, position, concentration, and doped metal, on the synergistic effect. All these selected systems have been demonstrated to have high production rates and selectivity for CO_2 conversion in experiments, consistent with the predicted positive synergistic effect (P_{Syn} < −0.10 V) from our ML model. The performance of DFT calculation and

https://doi.org/10.1021/jacsau.3c00558
JACS Au 2024, 4, 125−138

ML prediction of the maximum free energy change on 26 metal-anchored MoS_2 with and without sulfur vacancy were presented in Table S13 for P_{Syn} prediction. The predicted ΔG_{max} for Sc@MoS_2 and Ag@Sv-MoS_2 was slightly higher than the DFT calculated ΔG_{max}, whereas the predicted ΔG_{max} for the Mo@MoS_2 was smaller than the DFT calculated ΔG_{max}. It should be mentioned that the data set we used to construct the ML model was derived from the free energy changes of reducing intermediates (ΔG, referring to free CO_2 molecules and catalyst), rather than the free energy difference between two consecutive steps ($\Delta\Delta G$). The achievement of quantitative power of ML models for free energy change prediction is still a big challenge. Further effort to improve the prediction accuracy of P_{Syn} based on $\Delta\Delta G$ is of great significance for better guiding experimentalists in rational designing and screening of catalysts.

CONCLUSIONS

The synergistic effect between the metal and vacancy has been revealed in a general model for CO_2 activation in metal-anchored MoS_2 and other systems. The stability of the sulfur vacancy at different positions and catalytic activity of metal@Sv-MoS_2 are related to the number of d-shell valence electrons, group number, electronegativity of the metal center, and distance between the sulfur vacancy and the metal site. The Os@MoS_2 (−0.29 V) and Fe@Sv-MoS_2 (−0.31 V) are the most promising CO_2RR catalysts, with very low limiting potentials and high selectivity against the HER. A quantitative model for synergistic effect evaluation was proposed to assess the thermodynamic advantages of CH_4 conversion in three possible pathways (Syn, Coor, and Sv). Descriptors related to active sites on substrates, reaction pathways, and adsorbates were used to predict the change in free energy for CO_2RR (R^2 = 0.91), and successfully transferred to an external test for NRR (R^2 = 0.90). An interpretable ML model based on SHAP analysis reveals that $\delta\chi_{Sv}$, $\delta\chi_{Sub}$, P_S, and N_d are the most significant descriptors affecting the catalyst activity. The predicted limiting potential value and synergistic evaluation index are in agreement with the observed synergistic advantage for oxygen and tungsten vacancy systems in experiment, showing applicability of the present ML models to the prediction of the CO_2RR reactivity on other metal−supported catalysts. This work offers a practical tool for the design of efficient CO_2RR and NRR catalysts.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.3c00558.

Feature selection for catalytic activity prediction, coordination and synergistic pathway on CO_2RR, the abbreviation explanations in this work, free energy change calculation with different functionals, binding energies and CO_2 adsorption energy for metal-anchored MoS_2 with and without sulfur vacancy, machine learning for limiting potential prediction, descriptors for ΔG_{*CHO} and ΔG_{*OH} prediction, the external potential and pH effects, machine learning for free energy change prediction and external test (PDF)

AUTHOR INFORMATION

Corresponding Author

Jing Ma − *Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, P. R. China;* orcid.org/0000-0001-5848-9775; Email: majing@nju.edu.cn

Authors

Qin Zhu − *Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, P. R. China; State Key Laboratory of Organic Electronics and Information Displays (SKLOEID), Institute of Advanced Materials (IAM), Nanjing University of Posts & Telecommunications, Nanjing 210023, P. R. China*

Yating Gu − *Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, P. R. China*

Xinzhu Wang − *Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, P. R. China*

Yuming Gu − *Key Laboratory of Mesoscopic Chemistry of Ministry of Education, School of Chemistry and Chemical Engineering, Nanjing University, Nanjing 210023, P. R. China*

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.3c00558

Author Contributions

[§]Q.Z. and Y.G. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (22033004 and 22373049), National Key Laboratory (2009DS690095) and the Jiangsu Provincial Excellent Postdoctoral Program (2023ZB655). We are grateful to the High Performance Computing Centre of Nanjing University for providing the IBM Blade cluster system, Mr. Lifeng Zheng, Jiawei Chen, and Xu Gu for the help of building dataset webpage.

REFERENCES

(1) Babucci, M.; Guntida, A.; Gates, B. C. Atomically Dispersed Metals on Well-Defined Supports including Zeolites and Metal-Organic Frameworks: Structure, Bonding, Reactivity, and Catalysis. *Chem. Rev.* **2020**, *120*, 11956−11985.

(2) Shi, Y.; Lyu, Z.; Zhao, M.; Chen, R.; Nguyen, Q. N.; Xia, Y. Noble-Metal Nanocrystals with Controlled Shapes for Catalytic and Electrocatalytic Applications. *Chem. Rev.* **2021**, *121*, 649−735.

(3) Zhang, J.; Wang, M.; Gao, Z.; Qin, X.; Xu, Y.; Wang, Z.; Zhou, W.; Ma, D. Importance of Species Heterogeneity in Supported Metal Catalysts. *J. Am. Chem. Soc.* **2022**, *144*, 5108−5115.

(4) Jing, W. T.; Shen, H.; Qin, R. X.; Wu, Q. Y.; Liu, K. L.; Zheng, N. F. Surface and Interface Coordination Chemistry Learned from Model Heterogeneous Metal Nanocatalysts: From Atomically Dispersed Catalysts to Atomically Precise Clusters. *Chem. Rev.* **2023**, *123*, 5948−6002.

(5) Zhu, K.; Shi, F.; Zhu, X.; Yang, W. The Roles of Oxygen Vacancies in Electrocatalytic Oxygen Evolution Reaction. *Nano Energy* **2020**, *73*, No. 104761.
(6) Guo, X.; Song, E.; Zhao, W.; Xu, S.; Zhao, W.; Lei, Y.; Fang, Y.; Liu, J.; Huang, F. Charge Self-Regulation in 1T‴-MoS_2 Structure with Rich S Vacancies for Enhanced Hydrogen Evolution Activity. *Nat. Commun.* **2022**, *13*, 5954.
(7) Wu, X.; Zhang, W.; Li, J.; Xiang, Q.; Liu, Z.; Liu, B. Identification of the Active Sites on Metallic MoO_{2-x} Nano-Sea-Urchin for Atmospheric CO_2 Photoreduction Under UV, Visible, and Near-Infrared Light Illumination. *Angew. Chem, Int. Ed.* **2023**, *62*, No. e202213124.
(8) Gong, Y. N.; Jiao, L.; Qian, Y.; Pan, C. Y.; Zheng, L.; Cai, X.; Liu, B.; Yu, S. H.; Jiang, H. L. Regulating the Coordination Environment of MOF-Templated Single-Atom Nickel Electrocatalysts for Boosting CO_2 Reduction. *Angew. Chem., Int. Ed.* **2020**, *59*, 2705−2709.
(9) Liang, J. X.; Lin, J.; Liu, J.; Wang, X.; Zhang, T.; Li, J. Dual Metal Active Sites in an Ir_1/FeO_x Single-Atom Catalyst: A Redox Mechanism for the Water-Gas Shift Reaction. *Angew. Chem., Int. Ed.* **2020**, *59*, 12868−12875.
(10) Mitchell, S.; Pérez-Ramírez, J. Atomically Precise Control in the Design of Low-Nuclearity Supported Metal Catalysts. *Nat. Rev. Mater.* **2021**, *6*, 969−985.
(11) Yi, J.-d.; Gao, X.; Zhou, H.; Chen, W.; Wu, Y. Design of Co-Cu Diatomic Site Catalysts for High-efficiency Synergistic CO_2 Electroreduction at Industrial-level Current Density. *Angew. Chem., Int. Ed.* **2022**, *61*, No. e202212329.
(12) Guo, H.; Li, L.; Wang, X.; Yao, G.; Yu, H.; Tian, Z.; Li, B.; Chen, L. Theoretical Investigation on the Single Transition-Metal Atom-Decorated Defective MoS_2 for Electrocatalytic Ammonia Synthesis. *ACS Appl. Mater. Interfaces* **2019**, *11*, 36506−36514.
(13) Yu, H. J.; Chen, F.; Li, X. W.; Huang, H. W.; Zhang, Q. Y.; Su, S. Q.; Wang, K. Y.; Mao, E. Y.; Mei, B.; Mul, G.; et al. Synergy of Ferroelectric Polarization and Oxygen Vacancy to Promote CO_2 Photoreduction. *Nat. Commun.* **2021**, *12*, 4594.
(14) Hu, J.; Yu, L.; Deng, J.; Wang, Y.; Cheng, K.; Ma, C.; Zhang, Q. H.; Wen, W.; Yu, S.; Pan, Y.; Yang, J.; Ma, H.; Qi, F.; Wang, Y.; Zheng, Y.; Chen, M.; Huang, R.; Zhang, S.; Zhao, Z.; Mao, J.; Meng, X.; Ji, Q.; Hou, G.; Han, X.; Bao, X.; Wang, Y.; Deng, D. Sulfur Vacancy-Rich MoS_2 as a Catalyst for the Hydrogenation of CO_2 to Methanol. *Nat. Catal.* **2021**, *4*, 242−250.
(15) Luo, L.; Fu, L.; Liu, H.; Xu, Y.; Xing, J.; Chang, C.; Yang, D.; Tang, J. Synergy of Pd Atoms and Oxygen Vacancies on In_2O_3 for Methane Conversion under Visible Light. *Nat. Commun.* **2022**, *13*, 2930.
(16) Bui, T. S.; Lovell, E. C.; Daiyan, R.; Amal, R. Defective Metal Oxides: Lessons from CO_2RR and Applications in NO_xRR. *Adv. Mater.* **2023**, *35*, No. 2205814.
(17) Jiang, K.; Siahrostami, S.; Akey, A. J.; Li, Y.; Lu, Z.; Lattimer, J.; Hu, Y.; Stokes, C.; Gangishetty, M.; Chen, G.; Zhou, Y.; Hill, W.; Cai, W.-B.; Bell, D.; Chan, K.; Nørskov, J. K.; Cui, Y.; Wang, H. Transition-Metal Single Atoms in a Graphene Shell as Active Centers for Highly Efficient Artificial Photosynthesis. *Chem.* **2017**, *3*, 950−960.
(18) Yu, H.; Li, J.; Zhang, Y.; Yang, S.; Han, K.; Dong, F.; Ma, T.; Huang, H. Three-In-One Oxygen Vacancies: Whole Visible-Spectrum Absorption, Efficient Charge Separation, and Surface Site Activation for Robust CO_2 Photoreduction. *Angew. Chem., Int. Ed.* **2019**, *58*, 3880−3884.
(19) Zu, X.; Zhao, Y.; Li, X.; Chen, R.; Shao, W.; Wang, Z.; Hu, J.; Zhu, J.; Pan, Y.; Sun, Y.; Xie, Y. Ultrastable and Efficient Visible-Light-Driven CO_2 Reduction Triggered by Regenerative Oxygen-Vacancies in $Bi_2O_2CO_3$ Nanosheets. *Angew. Chem., Int. Ed.* **2021**, *60*, 13840−13846.
(20) Wu, X.; Zhang, W.; Li, J.; Xiang, Q.; Liu, Z.; Liu, B. Identification of the Active Sites on Metallic MoO_{2-x} Nano-Sea-Urchin for Atmospheric CO_2 Photoreduction Under UV, Visible, and Near-Infrared Light Illumination. *Angew. Chem., Int. Ed.* **2023**, *62*, No. e202213124.
(21) Ni, M.; Zhu, Y.; Guo, C.; Chen, D.-L.; Ning, J.; Zhong, Y.; Hu, Y. Efficient Visible-Light-Driven CO_2 Methanation with Self-Regenerated Oxygen Vacancies in $Co_3O_4/NiCo_2O_4$ Hetero-Nanocages: Vacancy-Mediated Selective Photocatalysis. *ACS Catal.* **2023**, *13*, 2502−2512.
(22) Chen, K.; Jiang, T.; Liu, T.; Yu, J.; Zhou, S.; Ali, A.; Wang, S.; Liu, Y.; Zhu, L.; Xu, X. Zn Dopants Synergistic Oxygen Vacancy Boosts Ultrathin CoO Layer for CO_2 Photoreduction. *Adv. Funct. Mater.* **2022**, *32*, No. 2109336.
(23) Wang, Y.; Hu, J.; Ge, T.; Chen, F.; Lu, Y.; Chen, R.; Zhang, H.; Ye, B.; Wang, S.; Zhang, Y.; et al. Gradient Cationic Vacancies Enabling Inner-to-Outer Tandem Homojunction: Strong Local Internal Electric Field and Reformed Basic Sites Boosting CO_2 Photoreduction. *Adv. Mater.* **2023**, *35*, No. 2302538.
(24) Tran, K.; Ulissi, Z. W. Active Learning Across Intermetallics to Guide Discovery of Electrocatalysts for CO_2 Reduction and H_2 Evolution. *Nat. Catal.* **2018**, *1*, 696−703.
(25) Zhong, M.; Tran, K.; Min, Y.; Wang, C.; Wang, Z.; Dinh, C. T.; De Luna, P.; Yu, Z.; Rasouli, A. S.; Brodersen, P.; Sun, S.; Voznyy, O.; Tan, C. S.; Askerka, M.; Che, F.; Liu, M.; Seifitokaldani, A.; Pang, Y.; Lo, S. C.; Ip, A.; Ulissi, Z.; Sargent, E. H. Accelerated Discovery of CO_2 Electrocatalysts using Active Machine Learning. *Nature* **2020**, *581*, 178−183.
(26) Lin, X.; Wang, Y.; Chang, X.; Zhen, S.; Zhao, Z.; Gong, J. High-Throughput Screening of Electrocatalysts for Nitrogen Reduction Reactions Accelerated by Interpretable Intrinsic Descriptor. *Angew. Chem., Int. Ed.* **2023**, *62*, No. e202300122.
(27) Yuan, H.; Li, Z.; Zeng, X. C.; Yang, J. Descriptor−Based Design Principle for Two−Dimensional Single-Atom Catalysts: Carbon Dioxide Electroreduction. *J. Phys. Chem. Lett.* **2020**, *11*, 3481−3487.
(28) Wang, J.; Zheng, M.; Zhao, X.; Fan, W. Structure-Performance Descriptors and the Role of the Axial Oxygen Atom on M−N_4−C Single-Atom Catalysts for Electrochemical CO_2 Reduction. *ACS Catal.* **2022**, *12*, 5441−5454.
(29) Zhang, Q.; Hu, P.; Xu, Z.; Tang, B.; Zhang, H.; Xiao, Y.; Wu, Y. Unravelling Intrinsic Descriptors Based on a Two Stage Activity Regulation of Bimetallic 2D c-MOFs for CO_2RR. *Nanoscale* **2023**, *15*, 4991−5000.
(30) Gong, L.; Zhang, D.; Lin, C. Y.; Zhu, Y.; Shen, Y.; Zhang, J.; Han, X.; Zhang, L.; Xia, Z. Catalytic Mechanisms and Design Principles for Single-Atom Catalysts in Highly Efficient CO_2 Conversion. *Adv. Energy Mater.* **2019**, *9*, No. 1902625.
(31) Yang, Z.; Gao, W.; Jiang, Q. A Machine Learning Scheme for the Catalytic Activity of Alloys with Intrinsic Descriptors. *J. Mater. Chem. A* **2020**, *8*, 17507−17515.
(32) Ren, C.; Lu, S.; Wu, Y.; Ouyang, Y.; Zhang, Y.; Li, Q.; Ling, C.; Wang, J. A Universal Descriptor for Complicated Interfacial Effects on Electrochemical Reduction Reactions. *J. Am. Chem. Soc.* **2022**, *144*, 12874−12883.
(33) Chen, A.; Zhang, X.; Chen, L.; Yao, S.; Zhou, Z. A Machine Learning Model on Simple Features for CO_2 Reduction Electrocatalysts. *J. Phys. Chem. C* **2020**, *124*, 22471−22478.
(34) Chen, Z. W.; Gariepy, Z.; Chen, L.; Yao, X.; Anand, A.; Liu, S.-J.; Tetsassi Feugmo, C. G.; Tamblyn, I.; Singh, C. V. Machine-Learning-Driven High-Entropy Alloy Catalyst Discovery to Circumvent the Scaling Relation for CO_2 Reduction Reaction. *ACS Catal.* **2022**, *12*, 14864−14871.
(35) Gariepy, Z.; Chen, G.; Xu, A.; Lu, Z.; Chen, Z.; Singh, C. V. Machine Learning Assisted Binary Alloy Catalyst Design for the Electroreduction of CO_2 to C_2 Products. *Energy Adv.* **2023**, *2*, 410−419.
(36) Wan, X.; Zhang, Z.; Niu, H.; Yin, Y.; Kuai, C.; Wang, J.; Shao, C.; Guo, Y. Machine-Learning-Accelerated Catalytic Activity Predictions of Transition Metal Phthalocyanine Dual-Metal-Site Catalysts for CO_2 Reduction. *J. Phys. Chem. Lett.* **2021**, *12*, 6111−6118.

(37) Xing, M.; Zhang, Y.; Li, S.; He, H.; Sun, S. Prediction of Carbon Dioxide Reduction Catalyst using Machine Learning with a Few-Feature Model: WLEDZ. *J. Phys. Chem. C* **2022**, *126*, 17025−17035.
(38) Feng, H.; Ding, H.; He, P.; Wang, S.; Li, Z.; Zheng, Z.; Yang, Y.; Wei, M.; Zhang, X. Data-Driven Design of Dual-Metal-Site Catalysts for the Electrochemical Carbon Dioxide Reduction Reaction. *J. Mater. Chem. A* **2022**, *10*, 18803−18811.
(39) Wang, D.; Cao, R.; Hao, S.; Liang, C.; Chen, G.; Chen, P.; Li, Y.; Zou, X. Accelerated Prediction of Cu-Based Single-Atom Alloy Catalysts for CO_2 Reduction by Machine Learning. *Green Energy Environ.* **2023**, *8*, 820−830.
(40) Zhu, Q.; Gu, Y.; Liang, X.; Wang, X.; Ma, J. A Machine Learning Model to Predict CO_2 Reduction Reactivity and Products Transferred from Metal-Zeolites. *ACS Catal.* **2022**, *12*, 12336−12348.
(41) Wang, Y.; Su, H.; He, Y.; Li, L.; Zhu, S.; Shen, H.; Xie, P.; Fu, X.; Zhou, G.; Feng, C.; Zhao, D.; Xiao, F.; Zhu, X.; Zeng, Y.; Shao, M.; Chen, S.; Wu, G.; Zeng, J.; Wang, C. Advanced Electrocatalysts with Single-Metal-Atom Active Sites. *Chem. Rev.* **2020**, *120*, 12217−12314.
(42) Liang, X.; Fu, N.; Yao, S.; Li, Z.; Li, Y. The Progress and Outlook of Metal Single-Atom-Site Catalysis. *J. Am. Chem. Soc.* **2022**, *144*, 18155−18174.
(43) Tang, T.; Wang, Z.; Guan, J. Structural Optimization of CarbonBased Diatomic Catalysts towards Advanced Electrocatalysis. *Coord. Chem. Rev.* **2023**, *492*, No. 215288.
(44) Kresse, G.; Furthmüller, J. Efficient Iterative Schemes for Ab Initio Total−Energy Calculations using a Plane-Wave Basis Set. *Phys. Rev. B* **1996**, *54*, 11169−11186.
(45) Kresse, G.; Hafner, J. Ab Initio Molecular Dynamics for Liquid Metals. *Phys. Rev. B* **1993**, *47*, 558−561.
(46) Blöchl, P. E. Projector Augmented-Wave Method. *Phys. Rev. B* **1994**, *50*, 17953−17979.
(47) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(48) Perdew, J. P.; Ernzerhof, M.; Burke, K. Rationale for Mixing Exact Exchange with Density Functional Approximations. *J. Chem. Phys.* **1996**, *105*, 9982−9985.
(49) Grimme, S. Semiempirical Gga-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(50) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate Ab Initio Parametrization of Density Functional Dispersion Correction (DFT−D) for the 94 Elements H−Pu. *J. Chem. Phys.* **2010**, *132*, No. 154104.
(51) Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jonsson, H. Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *J. Phys. Chem. B* **2004**, *108*, 17886−17892.
(52) Peterson, A. A.; Abild-Pedersen, F.; Studt, F.; Rossmeisl, J.; Nørskov, J. K. How Copper Catalyzes the Electroreduction of Carbon Dioxide into Hydrocarbon Fuels. *Energy Environ. Sci.* **2010**, *3*, 1311−1315.
(53) Delley, B. From Molecules to Solids with the $DMol^3$ Approach. *J. Chem. Phys.* **2000**, *113*, 7756−7764.
(54) *Materials Studio*, version 7.0; Accelrys Inc.: San Diego, CA, 2013.
(55) Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(56) Wen, Z.; Lv, H.; Wu, D.; Zhang, W.; Wu, X.; Yang, J. Sulfur-Coordinated Transition Metal Atom in Graphene for Electrocatalytic Nitrogen Reduction with an Electronic Descriptor. *J. Phys. Chem. Lett.* **2022**, *13*, 8177−8184.
(57) Tang, M.; Shen, H.; Sun, Q. Two-Dimensional Fe-Hexaaminobenzene Metal-Organic Frameworks as Promising CO_2 Catalysts with High Activity and Selectivity. *J. Phys. Chem. C* **2019**, *123*, 26460−26466.
(58) Wang, X.; Wu, J.; Zhang, Y.; Sun, Y.; Ma, K.; Xie, Y.; Zheng, W.; Tian, Z.; Kang, Z.; Zhang, Y. Vacancy Defects in 2D Transition Metal Dichalcogenide Electrocatalysts: From Aggregated to Atomic Configuration. *Adv. Mater.* **2023**, *35*, No. 2206576.
(59) 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*, No. 083802.
(60) Gu, Y.; Gu, Y.; Tao, Q.; Wang, X.; Zhu, Q.; Ma, J. Machine Learning for Prediction of $CO_2/N_2/H_2O$ Selective Adsorption and Separation in Metal-Zeolites. *J. Mater. Inf* **2023**, *3*, 19.
(61) Birdja, Y. Y.; Pérez-Gallent, E.; Figueiredo, M. C.; Göttle, A. J.; Calle-Vallejo, F.; Koper, M. T. M. Advances and Challenges in Understanding the Electrocatalytic Conversion of Carbon Dioxide to Fuels. *Nat. Energy* **2019**, *4*, 732−745.
(62) Peterson, A. A.; Nørskov, J. K. Activity Descriptors for CO_2 Electroreduction to Methane on Transition-Metal Catalysts. *J. Phys. Chem. Lett.* **2012**, *3*, 251−258.
(63) Chen, H.; Covert, I. C.; Lundberg, S. M.; Lee, S.-I. Algorithms to Estimate Shapley Value Feature Attributions. *Nat. Mach. Intell.* **2023**, *5*, 590−601.
(64) Apley, D. W.; Zhu, J. Visualizing the Effects of Predictor Variables in Black Box Supervised Learning Models. *J. R. Stat. Soc.* **2020**, *82*, 1059−1086.
(65) Wang, G.; Zhu, Q.; Zheng, H.; Zhang, S.; Ma, J. Binding Selectivity Between Diazobenzene and Different Nucleophilic Reagents: Covalent and Noncovalent Interactions. *Chem. J. Chinese Universites* **2021**, *42*, 2238−2244.
(66) Kour, G.; Mao, X.; Du, A. Computational Screening of Transition Metal-Phthalocyanines for the Electrochemical Reduction of Carbon Dioxide. *J. Phys. Chem. C* **2020**, *124*, 7708−7715.
(67) Cui, Q.; Qin, G.; Wang, W.; Geethalakshmi, K. R.; Du, A.; Sun, Q. Novel Two Dimensional MOF as a Promising Single-Atom Electrocatalyst for CO_2 Reduction: A Theoretical Study. *Appl. Surf. Sci.* **2020**, *500*, No. 143993.
(68) Gu, Y.; Zhu, Q.; Liu, Z.; Fu, C.; Wu, J.; Zhu, Q.; Jia, Q.; Ma, J. Nitrogen Reduction Reaction Energy and Pathway in Metal-zeolites: Deep Learning and Explainable Machine Learning with Local Acidity and Hydrogen Bonding Features. *J. Mater. Chem. A* **2022**, *10*, 14976−14988.