Nano Energy 115 (2023) 108695

Contents lists available at ScienceDirect

Nano Energy

journal homepage: www.elsevier.com/locate/nanoen

Active-learning accelerated computational screening of A_2B@NG catalysts for CO_2 electrochemical reduction

Xinyu Li [a,1], Haobo Li [b,1], Zhen Zhang [a], Javen Qinfeng Shi [a], Yan Jiao [b], Shi-Zhang Qiao [b,*]

[a] Australian Institute for Machine Learning, The University of Adelaide, Adelaide, SA 5000, Australia
[b] School of Chemical Engineering, The University of Adelaide, Adelaide, SA 5005, Australia

ARTICLE INFO

Keywords:
Few-atom catalysts
Machine learning
Electrochemical CO2 reduction
Computational screening

ABSTRACT

Few-atom catalysts, due to the unique coordination structure compared to metal particles and single-atom catalysts, have the potential to be applied for efficient electrochemical CO_2 reduction (CRR). In this study, we designed a class of triple-atom A_2B catalysts, with two A metal atoms and one B metal atom either horizontally or vertically embedded in the nitrogen-doped graphene plane. Metals A and B were selected from 17 elements across 3*d* to 5*d* transition metals. The structural stability and CRR activity of the 257 constructed A_2B catalysts were evaluated. The active-learning approach was applied to predict the adsorption site of key reaction intermediate *CO, which only used 40% computing resources in comparison to "brute force" calculation and greatly accelerated the large amount of computation brought by the large number of A_2B catalysts. Our results reveal that these triple atom catalysts can selectively produce more valuable hydrocarbon products while preserving high reactivity. Additionally, six triple-atom catalysts were proposed as potential CRR catalysts. These findings provide a theoretical understanding of the experimentally synthesized Fe_3 and Ru_3-N4 catalysts and lay a foundation for future discovery of few-atom catalysts and carbon materials in other applications. A new machine learning method, masked energy model, was also proposed which outperforms existing methods by approximately 5% when predicting low-coverage adsorption sites.

1. Introduction

Electrochemical carbon dioxide reduction represents a promising technique among all carbon-neutral technologies as it converts CO_2 into value-added fuels or chemicals using renewable energy sources under mild conditions. A wide variety of materials, including pure metals [1], alloys (bimetallic [2], trimetallic [3], high-entropy [4]), oxides [5], carbides [6,7], sulfides [8], two-dimension materials [9], nano materials [10] and atomically dispersed catalysts [11] have been experimentally explored as CO_2 electroreduction catalysts, but large scale CO_2 conversion remains challenging. The bottlenecks come from many aspects and innovative material design can improve most of them. For example, it can help engineering of reactions, enhancing product diffusion, and minimizing salt in electrolyte condensation. With the growth of interest in data science and artificial intelligence-enabled material design, a comprehensive database for CO_2 reduction that includes various materials would provide significant benefits for these fields.

Among all explored catalysts, atomically dispersed catalysts have attracted intensive research due to their superior catalytic activity and high atomic utility [11,12]. According to the number of anchored atoms, there are single-atom catalysts (SACs), dual atom catalysts (DACs), triple atom catalysts (TACs) and multiple atom catalysts. Among them, the SACs - disperse one single atom into substrates - have been studied extensively and proved to be effective to catalyse many reactions including hydrogen evolution reaction, oxygen evolution reaction (OER) [13,14], CO_2 reduction reaction (CRR) [15] and nitrogen reduction reaction (NRR) [16]. SACs offer maximum metal utilization as every atom acts as an active site. But, its single-atom active site also brings limitations - the distance between single-atom sites must be sufficient and each site can only bind one intermediate.

To go beyond the limitation of SACs, few atom catalysts (DACs and TACs) have been proposed and experimentally explored. Few-atom catalysts have active-site structures that can enhance or adjust the adsorption strength of reaction intermediates due to their higher coordination number. This capability also allows for the co-adsorption of multiple intermediates during C-C coupling reactions [17], leading to

* Corresponding author.
E-mail address: s.qiao@adelaide.edu.au (S.-Z. Qiao).
[1] These two authors contributed equally

https://doi.org/10.1016/j.nanoen.2023.108695
Received 27 April 2023; Received in revised form 5 July 2023; Accepted 11 July 2023
Available online 12 July 2023

the generation of C_2 or C_3 species in CRR. Qiao group synthesized NiCu DACs which show superior CRR activity and CO selectivity over SACs [18]. Li et al. found that Pt_2-MoS_2 shows great CO_2RR catalytic activity [19]. Guan et al. synthesized Cu and Cu_4 anchored on hydrophobic cyclohexene, and found that Cu SAC shows high-activity and selectivity towards CH_4 production, while Cu_4 few atom catalysts show activity towards C_2 hydrocarbon production [20]. For TACs, Ye et al. has synthesized the Fe, Fe_2 and Fe_3 anchored on N-doped Carbon (NC) and found that Fe_2 exhibits remarkable acid-related ORR performance [13]. Ji et al. has synthesized the Ru_3-N4 through the pyrolysis strategy which shows good 2-aminobenzaldehyde catalytic activity [21]. Theoretical calculations have been used to obtain mechanism insight and optimize few atom catalysts. Pei et al. investigated TACs horizontally embedded in N-doped graphene (NG) substrate (A_3-N6) for CRR [27], and found that A_3-N6 demonstrates high activity and selectivity towards the generation of hydrocarbons, with a limiting potential that is lower than 0.7 eV. Han et al. investigated horizontally embedded Fe_2B-N6 for CRR [22], Zheng et al. and Cui et al. investigated vertically embedded A_3-N4 and A_3@Graphene for NRR, respectively [23,24]. A_2Bs embedded both horizontally and vertically on NG have not been systematically investigated for CRR due to Density Functional Theory (DFT) efficiency limitation.

Computational high-throughput screening has been employed as an alternative to the inefficient trial-and-error method, guiding the development of innovative catalysts such as alloys and SACs [25,26]. Compared with SACs, the computational cost of high-throughput screening of few atom catalysts is significantly increased. This originates from the complexity of the metal element combinations – if the search space is limited to one element (e.g. A_3-N6), it is a single variable optimization with expected search space of N (N is the number of elements for searching, e.g. 17 in the case of this study). However, the complexity dramatically increased to N^2 (e.g., 289 in this study) in the case of two elements search and immediately becomes a bottleneck for finding optimal AB formula, especially considering the high computational cost of DFT. The complexity is even worse if multiple substrates and adsorption sites are considered. For this reason, most of current theoretical studies about few atom catalysts have been limited to one element for both DACs and TACs [23,27–29], while only a few works considered bimetallic compositions [30,31]. Challenge remains in developing computational methodologies to accelerate calculation so as to break the practical limitations in this direction.

In recent years, machine learning (ML) has emerged in the computational chemistry community due to its strong ability to mine the underlying relationships behind the input and output data [32]. For catalysis-related problems, it has been used to predict the adsorption energy [33–35], the activation energy [36], the entropy [37], and the molecular dynamics trajectory [38]. Most ML methods rely on geometrical representations (e.g. coordination number [39]) to encode the adsorption sites and predict their properties by statistical fitting, thereby replace part of time-consuming quantum chemical calculations and largely accelerate the high-throughput screening speed. However, ML method relies on a huge amount of data that forms the same distribution to make precise predictions. For example, if the training set consists of alloys, the test data should be alloys. If the test samples are oxides, it is considered out-of-distribution, and ML predictions based on such data cannot be relied upon. As a result, active learning - a method that selects samples and makes predictions iteratively instead of relying on existing datasets, has become increasingly popular. Ulissi group used coordination number fingerprint and active learning to find alloy catalyst for HER and CRR [2,40]. Zhu et al. used elemental properties (e.g. total valence electrons, atomic number) and gradient boost regression algorithm to predict OH adsorption energy [30]. Wu et al. used elemental properties and topological information as representations, and deep neural network to predict DACs' formation energies and overpotentials for OER and ORR [41].

In this study, we combined ML and DFT to investigate bimetallic A_2B-N6 and A_2B-N4 TACs. The overall workflow of our approach is shown in Fig. 1. Previous research has examined CRR on A_3-N6 or Fe_2B-N6 horizontally embedded structures [22,27], but vertical A_3-N4 was only investigated for NRR [23]. Achieving precise control over the horizontal or vertical A_2B@NG catalysts is challenging. As a result, conducting a mechanism investigation on both can reduce the trial-and-error approach required and provide guidance for experimental chemists to achieve their synthesis targets. With the aid of machine learning, we are able to examine both of them simultaneously. Different from previous works that use ML to predict adsorption energy or limiting potential directly, here we only use ML to predict the most stable adsorption site (adsorption site at a low coverage). This results in halving the number of DFT calculations in comparison to "brute-force" screening, but still guarantees the DFT level accuracy. We investigated the formation energy of these TACs, which refer to the possibility of synthesizing them experimentally; and 257 CO adsorption energy profiles on TACs, the largest number of CO adsorption data set on graphene-based TACs as far as we know. We also investigated the CRR by building a linear scaling relation between limiting potentials and CO adsorption energy. Our results explain the experimentally verified Fe_3@NC and Ru_3-N4 catalysts, and suggest Cu_2Cr-N6, Cu_2Mn-N6, Mn_2Cu-N6, Fe_2Mn-N4, Cu_2Mn-N4 and Fe_2Cr-N4 to be TACs for CRR with low limiting potential.

2. Results and discussion

2.1. Geometries and stability of triple-atom catalysts

We begin our discussion by investigating the geometries and stability of TACs. Fig. 2a-b shows the geometries of the substrate and TACs considered in this study. A $6 \times 6 \times 1$ graphene is used to represent the sparse distribution of triple atom clusters, as shown in Fig. S1a-b. The searching elements include 17 transition metals ranging from Group VI to Group XI elements (Fig. 2c). As shown in Fig. S1c-e, one initial structure was considered for A_2B-N6 and two initial structures for A_2B-N4, considering both previous theoretical and experimental studies [13,21–23,27]. After geometrical optimization, only clusters with geometry close to equilateral triangle were considered, which leads to 256 A_2B clusters. These clusters are put in both N6 and N4 substrates, forming A_2B-N6 and A_2B-N4 TACs, respectively.

Fig. 3 shows the formation energy of A_2B-N6 and A_2B-N4 TACs. For A_2B-N6, 3*d* transition metal TACs (Mn_2B, Fe_2B, Ni_2B, Cu_2B) have low formation energy as indicated by darker color (Fig. 3a), while the formation energy for 4*d* and 5*d* transition metals are much higher shown by lighter color. For A_2B-N4, in addition to the 3*d* transition metals, Pd_2B and Ag_2B in 4*d* metals show low formation energy as well. Generally speaking, the formation energy agrees well with our observation of the structure stability of the physical models, with a few exceptions. For example, some clusters such as Ag_2B and Au_2B (summarized in Fig. S2) moved away from the graphene plane during optimization with an unsymmetrical A_2B-N4 initial structure (Fig. S1e) due to the weak *N adsorption of Ag and Au [42]. The low formation energy of Ag_2B-N4 is inconsistent with this result, and we further calculated the binding energy, as shown in Fig. 3d. It is found that Pd_2B-N6, Pd_2B-N4, Ag_2B-N4 and Au_2B-N4 have weak binding energy, which explains that they are easily to be escaped. In contrast, W_2B-N4, Re_2B-N4, Os_2B-N4 and Ir_2B-N4 have low binding energy but high formation energy. We conclude that the stable and experimentally easily synthesized TACs must have both low formation energy and strong binding energy at the same time. According to this criterion, among all TACs considered, Mn_2B, Fe_2B, Ni_2B, Cu_2B are stable on both N6 and N4 substrates, while Ru_2B and Rh_2B are stable on N4 substrates only. These conclusions from computational perspective agree well with experimental synthesis of Fe_3@NC [13] and Ru_3-N4 [21] structures.

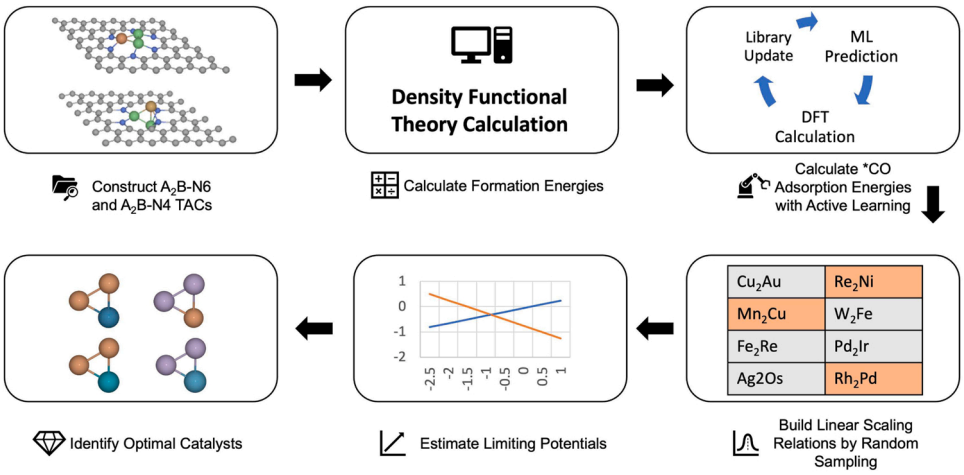

Fig. 1. The overall workflow used in this study.

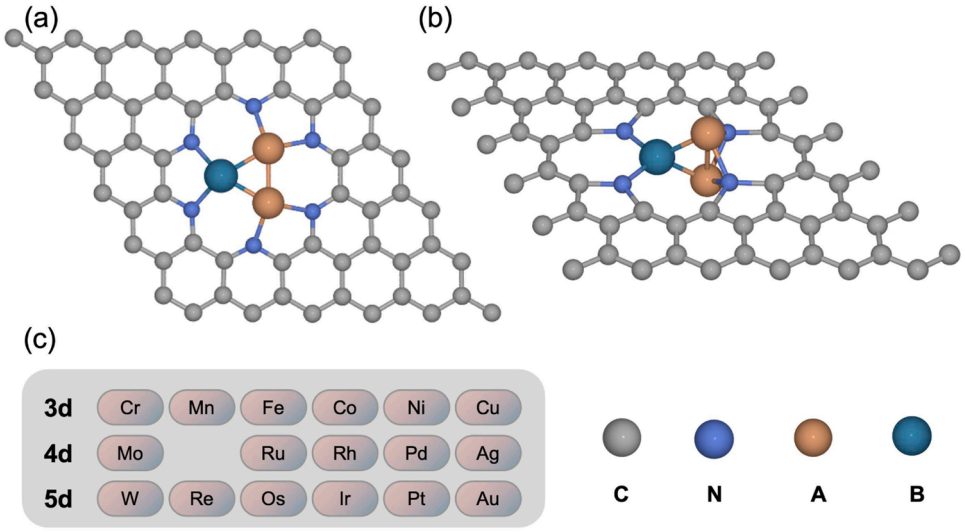

Fig. 2. Illustration of (a) the A_2B-N6, (b) the A_2B-N4 TACs and (c) the transition metals screened as A and B element.

2.2. *Comparing machine learning models in predicting low-coverage site*

Fig. 4 shows the adsorption sites for *CO on unsymmetrical A_2B-N4 (Fig. S3 shows the connectivity graph for adsorption models for A_2B-N4; Fig. S4 shows those for A_2B-N6). The adsorption sites on symmetrical A_2B-N4 should be similar but only contain the first three sites as it is symmetrical in the upper and downside. From Fig. S3 and Fig. S4, we can observe that A_2B-N4 and A_2B-N6 exhibit similar adsorption structures. For instance, Fig. S3b and Fig. S4a showcase identical coordination atom fingerprints. The only distinction arises in Fig. S4e, where the C atom binds to 3 metal atoms, a structure absent in A_2B-N4. Considering that there are 289 A_2B structures which are anchored on N6 and N4 substrates, the total number of DFT adsorption energy calculations is about 2000, which is hardly achievable for "brute-force" calculation. The adsorption site at a low coverage - the site with the lowest adsorption energy when putting only one adsorbate on the slab – is the one that was mostly concerned in theoretical catalysis. In statistical physics, the relative probability of two states (adsorption structure 1 with energy of E_1 and structure 2 with energy of E_2) can be calculated using the Boltzmann factor equation as $\exp(-\frac{E_1-E_2}{K_BT})$, where K_B is the Boltzmann constant and T is the temperature. This implies that a small energy difference can result in a significant difference in probability. Here, we developed a Masked Energy Model (MEM) to predict the most stable site so DFT calculations on other sites are not required, as detailed in Computational Details.

We now compare the proposed MEM with other models in predicting the low-coverage site. Currently, most studies are focused on predicting adsorption energy since it is the most commonly used descriptor of catalytic activity and selectivity. However, the TACs investigated in this study are different from materials such as pure metals, alloys, oxides, and single-atom catalysts, which have been studied extensively. We cannot benefit from the public data sets for those materials [44,45]. In this case, we used active learning in combination with MEM to address the problem of the limited number of samples and avoid the out-of-distribution prediction. Predicting the low-coverage site is a classification problem, where each site is labeled as either a low-coverage site or not. However, it still can be solved using regression algorithms by identifying the site with the lowest predicted value. Here, we are comparing gradient boosting classification (GBC), gradient boosting regression (GBR), support vector classification (SVC), support vector regression (SVR), GBR-MEM and SVR-MEM by using a toy data set generated in this study (93 slabs on N4 substrate, each slab has at least 2 calculated data points), as shown in Fig. 5. As expected, the accuracy of all methods generally increases with the number of samples. However, the difference in performance is primarily due to the choice of classification or regression, rather than the fundamental algorithm, such

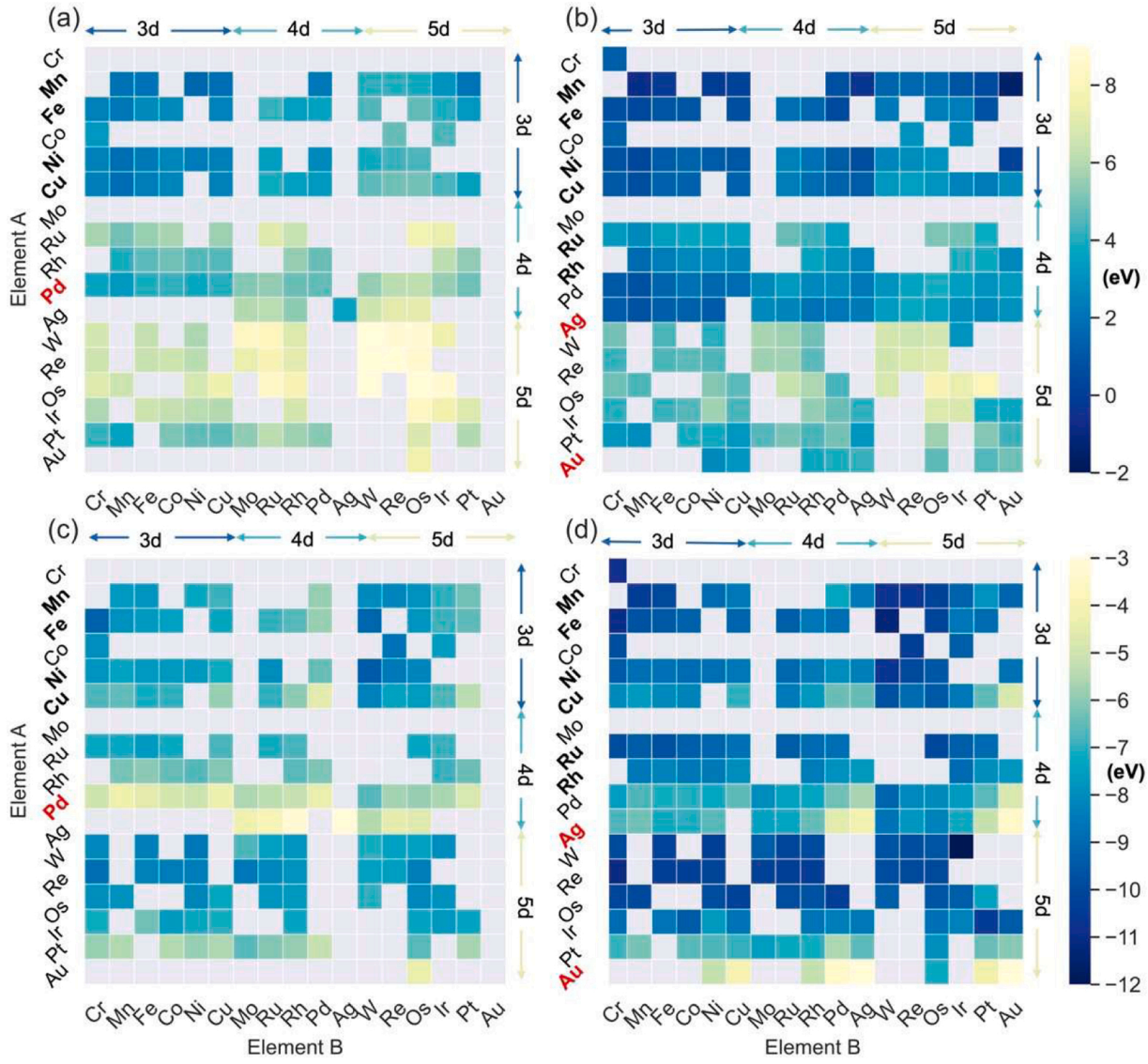

Fig. 3. The formation energy of (a) A_2B-N6, (b) A_2B-N4 and the binding energy of (c) A_2B-N6, (d) A_2B-N4. The *y* axis and *x* axis are the A element and B element in A_2B clusters, respectively. Stable A_2B-N6 and A_2B-N4 are highlighted in black and bold. Low formation energy but high binding energy Pd_2B-N6, Pd_2B-N4, Ag_2B-N4 and Au_2B-N4 are highlighted in red color.

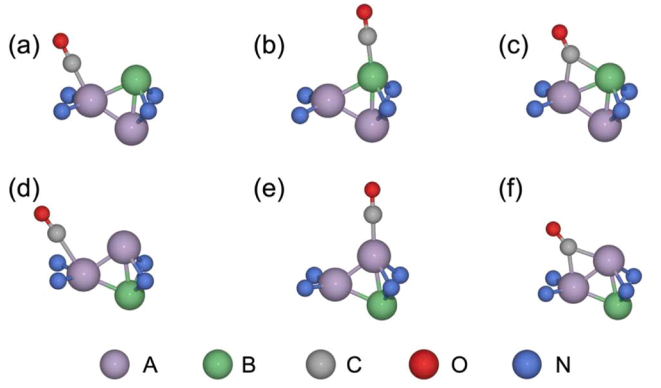

Fig. 4. Illustration of various adsorption sites on the unsymmetrical A_2B-N4. The adsorption sites on symmetrical A_2B-N4 are only (a)-(c) as the upper and downside are symmetrical.

as gradient boost or support vector. While classification algorithms, such as GBC and SVC, only marginally outperformed random choice, regression algorithms, such as GBR and SVR, easily doubled the accuracy in predicting low-coverage sites. Moreover, the accuracy of regression algorithms can be further improved by incorporating the MEM. For instance, GBR-MEM improved the original GBR by 6.7%, and SVR-MEM improved the original SVR by 5.7% with a training ratio of 10%. MEM enhances the prediction of low-coverage sites by providing a better representation of transferred sites. For example, an adsorbate that is placed on the top site but ends up at the hole site after relaxation would be represented by two data points in MEM: a hole site with an energy of −1.0 eV and a masked top site with an energy of −0.5 eV, but only one data point (the hole site) in the original Group and Period based Coordination Atom Fingerprint (GPCAF). MEM's effectiveness is a major reason for the 60% reduction in the number of DFT calculations.

2.3. Adsorption properties on A_2B TACs

The adsorption energy of *CO (ΔE_{CO}) is a commonly used descriptor for CO_2 electrochemical reduction [1,2,28]. Fig. 6 shows the ΔE_{CO} on

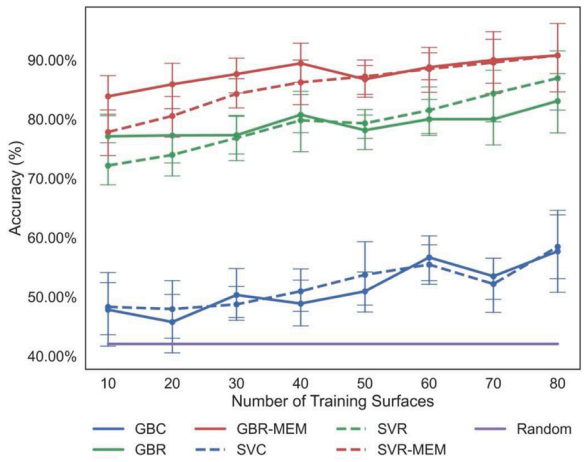

Fig. 5. Learning curves of gradient boosting classification (GBC), gradient boosting regression (GBR), GBR with Masked Energy Model (GBR-MEM), support vector classification (SVC), support vector regression (SVR), SVR with Masked Energy Model (SVR-MEM) and Random choice. Each data point is the average of 10 times independent test, and error bars show ± 1σ. The GPCAF is used as the representation for all ML methods.

the A_2B-N6 and A_2B-N4 TACs. ΔE_{CO} ranges from −4.20 to −1.75 eV, which is stronger than that on Cu surfaces. Previous studies revealed that Group IX and X pure metals (Rh, Ir, Ni, Pd, Pt) have ΔE_{CO} ranging from −2.5 to −1.5 eV, and various Cu surfaces ranging from −0.7–1.0 eV [1]. The CO adsorption on A_2B TACs is more like Group IX and X elements, which would suppress the production of CO. For A_2B-N6, most of the TACs have strong CO adsorption with ΔE_{CO} around −3.1 eV; the only exceptions are Cu_2Cr, Cu_2Mn, Cu_2Fe and Cu_2Cu, which have ΔE_{CO} of −1.92, −2.02, −2.19 and −2.20 eV, respectively. The ΔE_{CO} is co-determined by both the A and B elements for A_2B-N6; for example, A_2Os-N6 has very low ΔE_{CO} values for all A metals.

On the contrary, Fig. 6c shows that ΔE_{CO} is mainly determined by the A element for A_2B-N4 (e.g. Mn_2Os-N4, Fe_2Os-N4, Ni_2Os-N4 and Cu_2Os-N4 have high ΔE_{CO} values). The ΔE_{CO} for A_2B-N4 is peaked at −2.25 eV – about 1.0 eV higher than the peak of A_2B-N6. It is also found that the general trend for these A_2B-N4 is different with the catalytic trend suggested by *d*-band theory. According to the *d*-band theory [43,44], the upper left transition metals in the periodic table (e.g. Cr, Mn) have strong adsorption and lower right transition metals have weak adsorption. For A_2B-N4, upper left elements (e.g. Cr_3-N4 and Mn_3-N4) have weak adsorption and some lower right transition metals (e.g. Ir_3-N4, Pt_2Cu-N4, Ir_2W-N4) have strong adsorption. In contrast, A_2B-N6 still shows that upper left metals (Fe_2B-N6, W_2B-N6) have strong adsorption strength, while right metals (Cu_2B-N6) have weak adsorption strength.

2.4. Limiting potentials on A_2B TACs

Given the large number of A_2B TAC structures in this study, a full DFT-level reaction steps analysis would be extremely complex. Therefore, we focused on the two rate determining steps for CO_2 reduction ($*CO + H^+ + e^- \rightarrow *CHO$ and $CO_2(g) + H^+ + e^- \rightarrow *COOH$). We randomly sampled ~30 A_2B TAC structures on both N6 and N4, and calculated the

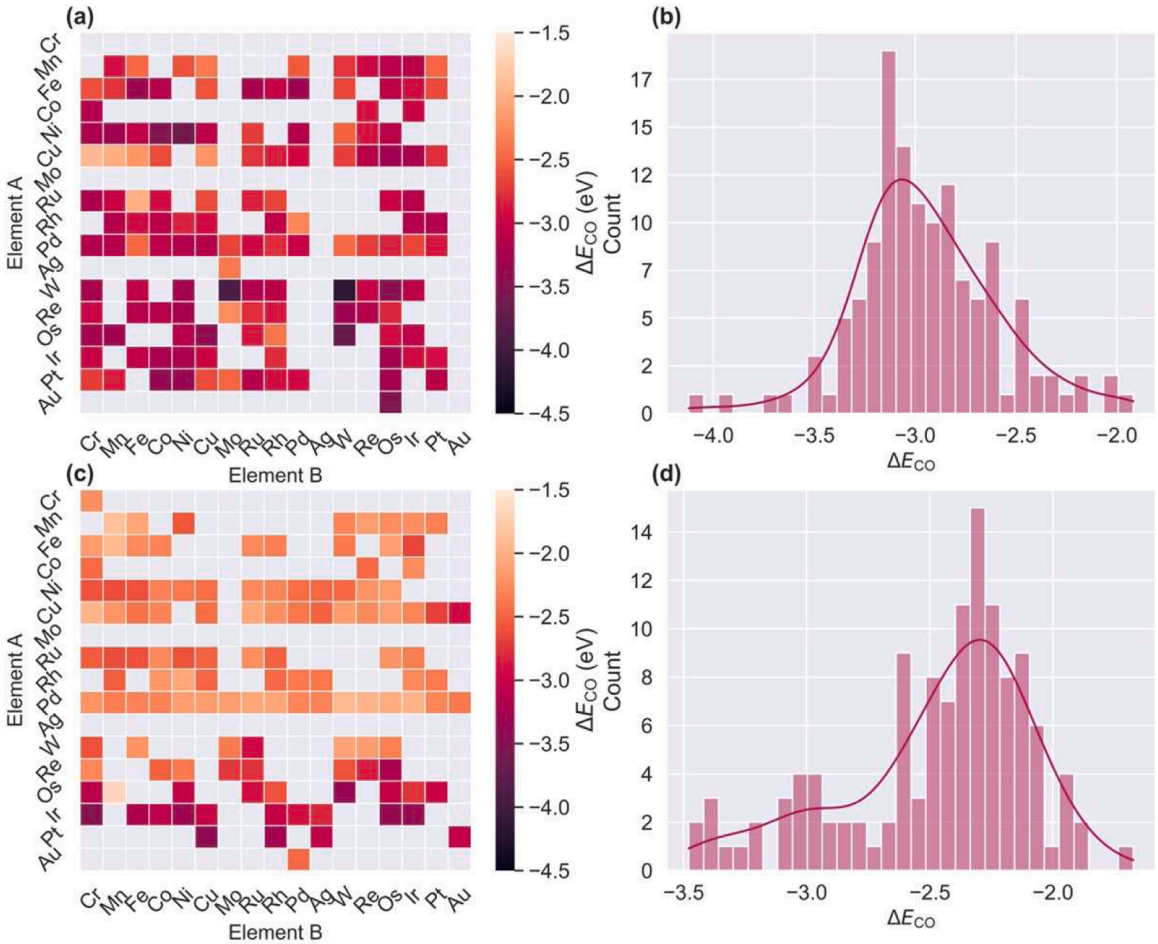

Fig. 6. (a) The heatmap of the adsorption energy of CO (ΔE_{CO}) and (b) the histogram distribution of ΔE_{CO} on A_2B-N6; (c) the heatmap of ΔE_{CO} and (d) the histogram distribution of ΔE_{CO} on A_2B-N4.

*CHO and *COOH adsorption energies on them. Fig. S5 shows the DFT calculated ΔE_{CHO} and ΔE_{COOH} against ΔE_{CO} on both N6 and N4 systems. On pure metal (211) surfaces, the *CO and *CHO binding energies show a linear slope of 0.88, which is seen as the major limitation on pure metals as a surface would stabilize the CHO and CO simultaneously [1]. Here, the slope is only 0.81 and 0.54, for A_2B-N6 and A_2B-N4 structures, respectively, which means the reaction step of $*CO + H^+ + e^- \rightarrow *CHO$ is easier on these TACs, especially on A_2B-N4. The linear scaling relations in Fig. S5 lay the foundation for estimating limiting potentials for all A_2B TACS catalysts considered in this study.

As demonstrated by Fig. 7, the limiting potential for two rate-limiting steps is depicted as a function of CO adsorption energies, revealing the high activity and selectivity that can be achieved through these TACs. It shows that these TACs have superior activity for CRR with limiting potential lower than 0.60 eV due to steeper overpotential slope for the rate-limiting step of $*CO + H^+ + e^- \rightarrow *CHO$ (0.19 for N6 and 0.46 for N4). A few A_2B with ΔE_{CO} of −2.0 eV can have the estimated limiting potentials of < 0.60 eV even if they can not sit atop of the volcano plot. These values are lower than both Cu(211) and Cu-C_3N_4, which sit atop of the TM(211) and A-C_3N_4 volcano plot [28]. Besides, these TACs also shows high selectivity for hydrocarbon production since A_2B-NG has a ΔE_{CO} of ~−2.0 eV, much stronger than the Cu(211) and Cu-C_3N_4; thus, the production of CO would be suppressed and the CRR would result in more valuable hydrocarbon products.

With the stability and activity analysis, we now make recommendations for A_2B@NG for CRR. First, Mn_2B-N6, Fe_2B-N6, Ni_2B-N6 and Cu_2B-N6 are the most stable ones among both N6 and N4 supported TACs, and Ru_2B-N4 and Rh_2B-N4 are also stable as N4 TACs. In addition, the limiting potential analysis suggests that a weak CO adsorption energy would have a low limiting potential, therefore increasing CRR activity. Among all stable A_2Bs, Cu_2Cr-N6, Cu_2Mn-N6, Mn_2Cu-N6, Fe_2Mn-N4, Cu_2Mn-N4 and Fe_2Cr-N4 have weak CO adsorption energy and low limiting potential, and are recommended as TACs for CRR. Especially, Cu_2Mn shows high activity with both N6 and N4 systems, which means multiple configurations would be effective for CRR therefore increase the overall efficiency.

Of course, there are still some challenges in the synthesis of these TACs. Firstly, controlling the number of atoms to be 3 and the A:B ratio of 2:1 is important for the CRR activity, as it has been demonstrated in this study that the CRR activity on A_2B and B_2A TACs are significantly different. A rational precursor selection, as introduced in Ref [13] which precisely synthesized Fe_1 to Fe_3 few-atom catalysts may help in achieving this control. Secondly, controlling the direction of these TACs - whether they are vertically or horizontally embedded, is vital for precise synthesis of A_2B-N4 and A_2B-N6. The application of electrostatic catalysis, [45] which utilizes an external electric field, may prove useful in controlling the direction. This is due to the polar nature of the A_2B clusters, which align themselves along the electric field. As a result, it facilitates the synthesis of consistent vertical A_2B-N4 or horizontal A_2B-N6 TACs. The existence of these challenges implies that A_2B TACs offer unique opportunities and difficulties for experimental verification.

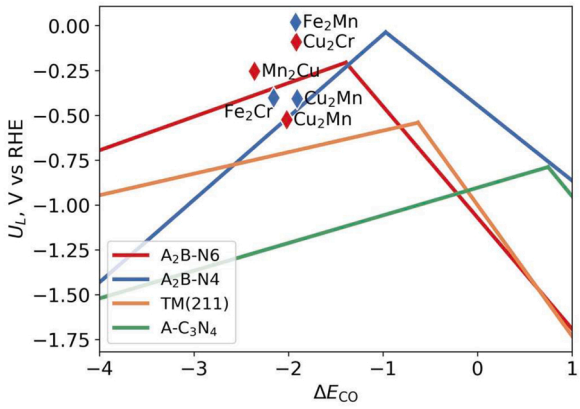

Fig. 7. Limiting potentials (U_L) for the two rate-limiting steps as a function of adsorption energy of *CO (ΔE_{CO}) on A_2B-N6(red), A_2B-N4 (blue), transition metal (211) surfaces [1] (orange) and single atom supported by g-C_3N_4 catalysts (A-C_3N_4) [28] (green). The line to the left of the peak represents the $*CO + H^+ + e^- \rightarrow *CHO$ reaction step and that to the right represents $CO_2(g) + H^+ + e^- \rightarrow *COOH$ step.

3. Conclusion

In summary, we systematically investigated the stability and CRR activity on bimetallic A_2B-N6 and A_2B-N4 TACs using DFT calculation and machine learning. We found that Mn_2B, Fe_2B, Ni_2B, Cu_2B are stable N6 TACs, and Mn_2B, Fe_2B, Ni_2B, Cu_2B, Ru_2B and Rh_2B are stable N4 TACs, which is consistent with the experimental evidence that Fe_3@NC and Ru_3-N4 have been fabricated experimentally. We propose a machine learning method, Masked Energy Model (MEM), for predicting low-coverage sites. MEM achieves an accuracy of ~5% higher than previous methods. By using active learning, we also calculated the CO adsorption energies on these TACs with 60% less computing resources used. Furthermore, a linear scaling relationship was built by randomly sampling N6 and N4 TACs and calculating the limiting potential for two rate-limiting steps on them. A volcano plot of limiting potential was also constructed for computational screening of optimal TACs. The volcano plot suggests that a weaker CO adsorption would result in a lower limiting potential on both N4 and N6 systems. Although none of the TACs can be located atop of the volcano, TACs with the highest CO adsorption energy (~−2.0 eV) can have a limiting potential lower than 0.6 V, outperforming state-of-the-art catalyst materials including Cu (211) and Cu-C_3N_4, while preserving high selectivity towards hydrocarbon production. Among all TACs explored, six TACs with weak CO adsorption and low limiting potentials are recommended for experimental verification. The finding of these stable TACs and the calculation of their CO adsorption energies sets a solid foundation for further investigation of carbon-related reactions such as carbon-nitrogen coupling and biomass reforming. Also, the active learning framework developed for the TACs discovery can be applied to uncover more intricate few-atom catalysts and carbon materials for various applications.

4. Computational methods

4.1. DFT calculations

All DFT calculations were performed using the Vienna ab initio simulation package (VASP) [46,47] with the Perdew-Burke-Ernzerhof (PBE) [48] exchange correlation functional, a 420 eV planewave cut-off energy and a Monkhorst-Pack k-point mesh of 4 × 4 × 1. Geometry optimization was carried out using the BFGS algorithm implemented in the Atomistic Simulation Environment (ASE) code [49] until the maximum force less than 0.05 eV/Å. Following Ref [50], spin polarization was considered for calculations involved with Mn, Fe, Co and Ni, with initial magnetic moments of 3, 3, 2, 1 μ_B, respectively. Dispersion correction was considered using DFT-D3 [51]. Different DFT settings, including DFT+U and an implicit solvation model,[52] have been tested for CO adsorption on A_2B TACs. As shown in Table S1, these settings have slight influence on the CO adsorption energy calculated in this study.

4.2. Machine learning and active learning

Group and Period based Coordination Atom Fingerprint (GPCAF) [35] is used as the representation of adsorption site in this study (Fig. S6 shows the construction of GPCAF for the structures in Fig. 4). Gradient Boost Classification, Gradient Boost Regression [53], Support Vector

Classification [54] and Support Vector Regression [55], as implemented in the *Scikit-learn* package [56], are used as the machine learning algorithms. Different with previous studies which intend to predict the adsorption energy from initial configuration, the models used here are designed to predict the adsorption site with the lowest adsorption energy. This is done by giving those unstable sites a 0.5 eV higher adsorption energy (e.g. if a CO was initially put in the top site of Ni(111) surface and finished at the bridge site with an adsorption energy of −1.5 eV, then a 'fake' data point of top site with an adsorption energy of −1.0 eV was added to the training set. This model is denoted as Masked Energy Model, MEM). The MEM approach is expected to be effective for catalysts where site transferring occurs, such as pure metals, alloys and oxides. However, there is an exception when it comes to catalysts with surface defects, where adsorbates strongly prefer the defect site. In these cases, MEM would not make a substantial contribution to the prediction accuracy.

The publicly available large data sets like OC20 [57] are mostly focused on alloy catalysts; therefore, machine learning model in this study cannot be beneficial from the large number of samples in the public data sets and here we use active learning to automatically select data points. Fig. S7 illustrates the active learning process used in this study. At each active learning iteration, the machine learning model is trained and used to predict uncalculated samples. The query strategy then selects some uncalculated data points for calculation using DFT. In this study, a query strategy that combines greedy query and random query is used. The greedy query requires the data points that were predicted to be low coverage site, and the random query requires data points randomly to avoid overfitting. At every active learning iteration, 60% of samples were selected by greedy query and 40% by random query. After the DFT calculation, the dataset is updated and the next iteration begins. Our active learning ends when all machine learning predicted sites have been calculated using DFT.

4.3. Screening criteria for catalyst stability

The stability of the A_2B@NG catalyst is evaluated by the formation energy E_{f,A_2B-NG}, which is calculated by:

$$E_{f,A_2B@NG} = E_{A_2B@NG} - E_{NG} - 2E_{A(bulk)} - E_{B(bulk)} \quad (1)$$

where $E_{A_2B@NG}$ is the total energy of the whole TAC structure, E_{NG} is the energy of the defected N6 and N4 substrate. $E_{A(bulk)}$ and $E_{B(bulk)}$ are the reference energies of element A and B, respectively. Here the energy per atom of the bulk metal structures at room temperature is used as reference energies.

Following Ref [27], we also calculated the binding energy which measures the energy difference before and after a A_2B cluster being anchored on the NG, as:

$$E_{b,A_2B@NG} = E_{A_2B@NG} - E_{NG} - E_{A_2B} \quad (2)$$

where E_{A_2B} is the energy of A_2B cluster without the substrate.

4.4. Screening criteria for CRR activity

The catalytic activity is evaluated by comparing their limiting potentials (U_L). To obtain limiting potential, the adsorption energy of relevant reaction intermediate is first calculated, as:

$$\Delta E_{ads} = E_{ads-slab} - E_{slab} - \sum_{S\in\{C,H,O\}} N_S E_S \quad (3)$$

where $E_{ads-slab}$ is the DFT energy of adsorbate/slab complex, E_{slab} is the DFT energy of pure slab, N_S and E_S are the number and reference energy of element S in the adsorbate. In this study, the reference energies of carbon, oxygen and hydrogen are defined as , $E_O = E_{CO} - E_C$, $E_H = 1/2E_{H_2}$. Hydrogen evolution reaction (HER) is a well-known competing reaction of CRR. Pei et al. have shown that most A_3-N6 catalysts have weaker H adsorption than CO_2. Here, we also compared the CO_2 and H adsorption on some representative A_2B catalysts (Fig. S8) and showed that the trend is consistent with A_3-N6, [27] as the majority of A_2B catalysts have weaker H binding strength, except for the Cu_3-N6 tested here. This indicates that this type of catalyst structures is generally more selective to CRR than HER.

To reduce the calculation amount for a full reaction network, we focused on the two rate-limiting steps [1,27]:

$$*CO + H^+ + e^- \rightarrow *CHO \quad (4)$$

$$CO_2(g) + H^+ + e^- \rightarrow *COOH \quad (5)$$

where * denotes an adsorption site.

For each reaction step, the reaction free energy ΔG is calculated by:

$$\Delta G = \Delta E + \Delta ZPE - T\Delta S + \Delta \int C_p dT \quad (6)$$

where ΔE is the total DFT energy difference between reactants and products of each reaction, ΔZPE is the zero-point energy correction, ΔS is the vibrational entropy change, $\Delta \int C_p dT$ is the heat capacity change. T is set as 298.15 K in our study. To accelerate the screening, a fixed numerical correction was used with ΔZPE, ΔS and $\Delta \int C_p dT$ on Cu(211) surfaces [58]. The energy of $H^+ + e^-$ is referenced to ½ H_2(g) according to computational hydrogen electrode [59].

The limiting potential (U_L) is defined as:

$$U_L = \max\{\Delta G_1, \Delta G_2\}/e \quad (7)$$

where ΔG_1 and ΔG_2 are the reaction free energy of the two rate-limiting steps (Eq. 4 and Eq. 5), and e is the elementary negative charge. We have verified all the possible reaction steps in CRR towards methane production on Cu_2Mn-N6 and Cu_2Mn-N4. The free energy diagram in Fig. S9 shows that the $*CO + H^+ + e^- \rightarrow *CHO$ step is the rate-limiting step. The calculated limiting potentials are well correlated with the fixed numerical correction values (−0.48 V vs −0.52 V for Cu_2Mn-N6 and −0.46 V vs −0.40 V for Cu_2Mn-N4).

CRediT authorship contribution statement

S.-Z.Q., Y.J. and J.Q.S. conceived and supervised this research; X.L. and H.L. designed and carried out the DFT calculations and electrochemical analysis; X.L. and Z.Z. designed and programmed the machine learning models; S.-Z.Q. and Y.J. acquired research and computational resources; The manuscript was written through contributions of all authors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

X.L. and H.L. contributed equally to this work. This work was funded by the Center for Augmented Reasoning at the Australian Institute for Machine Learning. X.L. would like to thank Miss Sijia Fu for sharing her data on A-C_3N_4 catalysts. S.Z.Q. acknowledges financial support of Australian Research Council (ARC) through Discovery Project program (DP230102027, DP220102596, FL170100154). Y.J. acknowledges funding support through ARC (FT190100636). Computations for this

research were under-taken with the assistance of supercomputing resources provided by the Phoenix HPC service at the University of Adelaide and the resources and services from the National Computational Infrastructure (NCI) supported by the Australian Government.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.nanoen.2023.108695.

References

[1] A.A. Peterson, J.K. Nørskov, Activity descriptors for CO_2 electroreduction to methane on transition-metal catalysts, J. Phys. Chem. Lett. 3 (2) (2012) 251–258.

[2] M. Zhong, K. Tran, Y. Min, C. Wang, Z. Wang, C.-T. Dinh, P. De Luna, Z. Yu, A. S. Rasouli, P. Brodersen, S. Sun, O. Voznyy, C.-S. Tan, M. Askerka, F. Che, M. Liu, A. Seifitokaldani, Y. Pang, S.-C. Lo, A. Ip, Z. Ulissi, E.H. Sargent, Accelerated discovery of CO_2 electrocatalysts using active machine learning, Nature 581 (7807) (2020) 178–183.

[3] X. Ma, F. Sun, L. Qin, Y. Liu, X. Kang, L. Wang, D.-E. Jiang, Q. Tang, Z. Tang, Electrochemical CO_2 reduction catalyzed by atomically precise alkynyl-protected Au_7Ag_8, Ag_9Cu_6, and $Au_2Ag_8Cu_5$ nanoclusters: probing the effect of multi-metal core on selectivity, Chem. Sci. 13 (34) (2022) 10149–10158.

[4] S. Nellaiappan, N.K. Katiyar, R. Kumar, A. Parui, K.D. Malviya, K.G. Pradeep, A. K. Singh, S. Sharma, C.S. Tiwary, K. Biswas, High-entropy alloys as catalysts for the CO_2 and CO reduction reactions: experimental realization, ACS Catal. 10 (6) (2020) 3658–3663.

[5] D. Wu, C.K. Dong, D.Y. Wu, J.Y. Fu, H. Liu, S.W. Hu, Z. Jiang, S.Z. Qiao, X.W. Du, Cuprous ions embedded in ceria lattice for selective and stable electrochemical reduction of carbon dioxide to ethylene, J. Mater. Chem. A 6 (20) (2018) 9373–9377.

[6] C. Griesser, H. Li, E.-M. Wernig, D. Winkler, N. Shakibi Nia, T. Mairegger, T. Götsch, T. Schachinger, A. Steiger-Thirsfeld, S. Penner, D. Wielend, D. Egger, C. Scheurer, K. Reuter, J. Kunze-Liebhäuser, True nature of the transition-metal carbide/liquid interface determines its reactivity, ACS Catal. 11 (8) (2021) 4920–4928.

[7] H. Li, K. Reuter, Ab initio thermodynamic stability of carbide catalysts under electrochemical conditions, ACS Catal. 12 (16) (2022) 10506–10513.

[8] A. Yamaguchi, M. Yamamoto, K. Takai, T. Ishii, K. Hashimoto, R. Nakamura, Electrochemical CO_2 reduction by Ni-containing Iron Sulfides: how Is CO_2 electrochemically reduced at bisulfide-bearing deep-sea hydrothermal precipitates? Electrochim. Acta 141 (2014) 311–318.

[9] Z.W. Chen, L.X. Chen, M. Jiang, D. Chen, Z.L. Wang, X. Yao, C.V. Singh, Q. Jiang, A triple atom catalyst with ultrahigh loading potential for nitrogen electrochemical reduction, J. Mater. Chem. A 8 (30) (2020) 15086–15093.

[10] C. Xu, A. Vasileff, B. Jin, D. Wang, H. Xu, Y. Zheng, S.-Z. Qiao, Graphene-encapsulated nickel–copper bimetallic nanoparticle catalysts for electrochemical reduction of CO_2 to CO, Chem. Commun. 56 (76) (2020) 11275–11278.

[11] C. Xu, X. Zhi, A. Vasileff, D. Wang, B. Jin, Y. Jiao, Y. Zheng, S.-Z. Qiao, Highly selective two-electron electrocatalytic CO_2 reduction on single-atom Cu catalysts, Small Struct. 2 (1) (2021), 2000058.

[12] Z.W. Chen, L.X. Chen, C.C. Yang, Q. Jiang, Atomic (single, double, and triple atoms) catalysis: frontiers, opportunities, and challenges, J. Mater. Chem. A 7 (8) (2019) 3492–3515.

[13] W. Ye, S. Chen, Y. Lin, L. Yang, S. Chen, X. Zheng, Z. Qi, C. Wang, R. Long, M. Chen, J. Zhu, P. Gao, L. Song, J. Jiang, Y. Xiong, Precisely tuning the number of Fe atoms in clusters on N-doped carbon toward acidic oxygen reduction reaction, Chem 5 (11) (2019) 2865–2878.

[14] J.Q. Shan, C. Ye, S.M. Chen, T.L. Sun, Y. Jiao, L.M. Liu, C.Z. Zhu, L. Song, Y. Han, M. Jaroniec, Y.H. Zhu, Y. Zheng, S.Z. Qiao, Short-range ordered iridium single atoms integrated into cobalt oxide spinel structure for highly efficient electrocatalytic water oxidation, J. Am. Chem. Soc. 143 (13) (2021) 5201–5211.

[15] L.M. Wang, W.L. Chen, D.D. Zhang, Y.P. Du, R. Amal, S.Z. Qiao, J.W. Bf, Z.Y. Yin, Surface strategies for catalytic CO_2 reduction: from two-dimensional materials to nanoclusters to single atoms, Chem. Soc. Rev. 48 (21) (2019) 5310–5349.

[16] X.F. Li, Q.K. Li, J. Cheng, L.L. Liu, Q. Yan, Y.C. Wu, X.H. Zhang, Z.Y. Wang, Q. Qiu, Y. Luo, Conversion of dinitrogen to ammonia by FeN_3-embedded graphene, J. Am. Chem. Soc. 138 (28) (2016) 8706–8709.

[17] X. Zhi, Y. Jiao, Y. Zheng, S.Z. Qiao, Key to C_2 production: selective C-C coupling for electrochemical CO_2 reduction on copper alloy surfaces, Chem. Commun. 57 (75) (2021) 9526–9529.

[18] D. Yao, C. Tang, X. Zhi, B. Johannessen, A. Slattery, S. Chern, S.-Z. Qiao, Inter-metal interaction with a threshold effect in NiCu dual-atom catalysts for CO_2 electroreduction, Adv. Mater. 35 (2023), 2209386.

[19] H.L. Li, L.B. Wang, Y.Z. Dai, Z.T. Pu, Z.H. Lao, Y.W. Chen, M.L. Wang, X.S. Zheng, J.F. Zhu, W.H. Zhang, R. Si, C. Ma, J. Zeng, Synergetic interaction between neighbouring platinum monomers in CO_2 hydrogenation, Nat. Nanotechnol. 13 (5) (2018) 411–417.

[20] A.X. Guan, C. Yang, Q.H. Wang, L.P. Qian, J.Y. Cao, L.J. Zhang, L.M. Wu, G. F. Zheng, Atomic-level copper sites for selective CO_2 electroreduction to hydrocarbon, ACS Sustain. Chem. Eng. 9 (40) (2021) 13536–13544.

[21] S. Ji, Y. Chen, Q. Fu, Y. Chen, J. Dong, W. Chen, Z. Li, Y. Wang, L. Gu, W. He, C. Chen, Q. Peng, Y. Huang, X. Duan, D. Wang, C. Draxl, Y. Li, Confined pyrolysis within metal–organic frameworks to form uniform Ru_3 clusters for efficient oxidation of alcohols, J. Am. Chem. Soc. 139 (29) (2017) 9795–9798.

[22] B. Han, H. Meng, F. Li, Supported bimetallic trimers Fe_2M@NG: triple atom catalysts for CO_2 electroreduction, ACS Omega 7 (18) (2022) 16080–16086.

[23] G. Zheng, L. Li, Z. Tian, X. Zhang, L. Chen, Heterogeneous single-cluster catalysts (Mn_3, Fe_3, Co_3, and Mo_3) supported on nitrogen-doped graphene for robust electrochemical nitrogen reduction, J. Energy Chem. 54 (2021) 612–619.

[24] C.N. Cui, H.C. Zhang, R. Cheng, B.B. Huang, Z.X. Luo, On the nature of three-atom metal cluster catalysis for N_2 reduction to ammonia, ACS Catal. 12 (24) (2022) 14964–14975.

[25] J. Greeley, T.F. Jaramillo, J. Bonde, I.B. Chorkendorff, J.K. Norskov, Computational high-throughput screening of electrocatalytic materials for hydrogen evolution, Nat. Mater. 5 (11) (2006) 909–913.

[26] F.Y. Li, Y.F. Li, X.C. Zeng, Z.F. Chen, Exploration of high-performance single-atom catalysts on support M_1/FeO_x for CO oxidation via computational study, ACS Catal. 5 (2) (2015) 544–552.

[27] W. Pei, S. Zhou, J. Zhao, X. Xu, Y. Du, S.X. Dou, Immobilized trimeric metal clusters: a family of the smallest catalysts for selective CO_2 reduction toward multi-carbon products, Nano Energy 76 (2020), 105049.

[28] S.J. Fu, X. Liu, J.R. Ran, Y. Jiao, Theoretical considerations on activity of the electrochemical CO_2 reduction on metal single-atom catalysts with asymmetrical active sites, Catal. Today 397 (2022) 574–580.

[29] N. Karmodak, S. Vijay, G. Kastlunger, K. Chan, Computational screening of single and Di-atom catalysts for electrochemical CO_2 reduction, ACS Catal. 12 (9) (2022) 4818–4824.

[30] X. Zhu, J. Yan, M. Gu, T. Liu, Y. Dai, Y. Gu, Y. Li, Activity origin and design principles for oxygen reduction on dual-metal-site catalysts: a combined density functional theory and machine learning study, J. Phys. Chem. Lett. 10 (24) (2019) 7760–7766.

[31] D.C. Chen, Z.W. Chen, Z.L. Lu, J. Tang, X.X. Zhang, C.V. Singh, Computational screening of homo and hetero transition metal dimer catalysts for reduction of CO_2 to C_2 products with high activity and low limiting potential, J. Mater. Chem. A 8 (40) (2020) 21241–21254.

[32] H. Li, Y. Jiao, K. Davey, S.-Z. Qiao, Data-driven machine learning for understanding surface structures of heterogeneous catalysts, Angew. Chem. Int. Ed. 62 (9) (2023), e202216383.

[33] X. Li, R. Chiong, Z. Hu, D. Cornforth, A.J. Page, Improved representations of heterogeneous carbon reforming catalysis using machine learning, J. Chem. Theory Comput. 15 (12) (2019) 6882–6894.

[34] X. Li, R. Chiong, Z. Hu, A.J. Page, Low-cost Pt alloys for heterogeneous catalysis predicted by density functional theory and active learning, J. Phys. Chem. Lett. 12 (30) (2021) 7305–7311.

[35] X. Li, R. Chiong, A.J. Page, Group and period-based representations for improved machine learning prediction of heterogeneous alloy catalysts, J. Phys. Chem. Lett. 12 (21) (2021) 5156–5162.

[36] A.R. Singh, B.A. Rohr, J.A. Gauthier, J.K. Nørskov, Predicting chemical reaction barriers with a machine learning model, Catal. Lett. 149 (9) (2019) 2347–2354.

[37] L. Kollias, G. Collinge, D. Zhang, S.I. Allec, P.K. Gurunathan, G. Piccini, S.F. Yuk, M.-T. Nguyen, M.-S. Lee, V.-A. Glezakou, R. Rousseau, Chapter One - Assessing entropy for catalytic processes at complex reactive interfaces, in: D.A. Dixon (Ed.), Annual Reports in Computational Chemistry, Elsevier, Amsterdam, Netherlands, 2022, pp. 3–51.

[38] O.T. Unke, S. Chmiela, H.E. Sauceda, M. Gastegger, I. Poltaysky, K.T. Schutt, A. Tkatchenko, K.R. Muller, Machine learning force fields, Chem. Rev. 121 (16) (2021) 10142–10186.

[39] F. Calle-Vallejo, J. Tymoczko, V. Colic, Q.H. Vu, M.D. Pohl, K. Morgenstern, D. Loffreda, P. Sautet, W. Schuhmann, A.S. Bandarenka, Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors, Science 350 (6257) (2015) 185–189.

[40] K. Tran, Z.W. Ulissi, Active learning across intermetallics to guide discovery of electrocatalysts for CO_2 reduction and H_2 evolution, Nat. Catal. 1 (9) (2018) 696–703.

[41] L.P. Wu, T. Guo, T. Li, Data-driven high-throughput rational design of double-atom catalysts for oxygen evolution and reduction, Adv. Funct. Mater. 32 (31) (2022), 2203439.

[42] F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T.R. Munter, P.G. Moses, E. Skulason, T. Bligaard, J.K. Norskov, Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces, Phys. Rev. Lett. 99 (1) (2007), 016105.

[43] B. Hammer, J.K. Norskov, Why gold is the noblest of all the metals, Nature 376 (6537) (1995) 238–240.

[44] F. Abild-Pedersen, M.P. Andersson, CO adsorption energies on metals with correction for high coordination adsorption sites - a density functional study, Surf. Sci. 601 (7) (2007) 1747–1753.

[45] A.C. Aragones, N.L. Haworth, N. Darwish, S. Ciampi, N.J. Bloomfield, G. G. Wallace, I. Diez-Perez, M.L. Coote, Electrostatic catalysis of a Diels-Alder reaction, Nature 531 (7592) (2016) 88–91.

[46] G. Kresse, J. Furthmüller, Efficient iterative schemes forab initiototal-energy calculations using a plane-wave basis set, Phys. Rev. B 54 (16) (1996) 11169–11186.

[47] G. Kresse, J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. 6 (1) (1996) 15–50.

[48] J.P. Perdew, K. Burke, M. Ernzerhof, Generalized gradient approximation made simple, Phys. Rev. Lett. 77 (18) (1996) 3865.
[49] A.H. Larsen, J.J. Mortensen, J. Blomqvist, I.E. Castelli, R. Christensen, M. Dulak, J. Friis, M.N. Groves, B. Hammer, C. Hargus, E.D. Hermes, P.C. Jennings, P. B. Jensen, J. Kermode, J.R. Kitchin, E.L. Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J.B. Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiotz, O. Schutt, M. Strange, K.S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z.H. Zeng, K.W. Jacobsen, The atomic simulation environment-a Python library for working with atoms, J. Phys.: Condens. Matter 29 (27) (2017), 273002.
[50] O. Mamun, K.T. Winther, J.R. Boes, T. Bligaard, High-throughput calculations of catalytic properties of bimetallic alloy surfaces, Sci. Data 6 (1) (2019) 76.
[51] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, J. Chem. Phys. 132 (15) (2010), 154104.
[52] K. Mathew, V.S.C. Kolluru, S. Mula, S.N. Steinmann, R.G. Hennig, Implicit self-consistent electrolyte model in plane-wave density-functional theory, J. Chem. Phys. 151 (23) (2019).
[53] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (5) (2001) 1189–1232.
[54] V.N. Vapnik, The nature of statistical learning theory, Springer,, New York, 1995.
[55] A.J. Smola, B. Schölkopf, A tutorial on support vector regression, Stat. Comput. 14 (3) (2004) 199–222.
[56] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: machine learning in python, J. Mach. Learn. Res. 12 (2011) 2825–2830.
[57] L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho, W.H. Hu, A. Palizhati, A. Sriram, B. Wood, J. Yoon, D. Parikh, C. L. Zitnick, Z. Ulissi, Open catalyst 2020 (OC20) dataset and community challenges, ACS Catal. 11 (10) (2021) 6059–6072.
[58] A.A. Peterson, F. Abild-Pedersen, F. Studt, J. Rossmeisl, J.K. Norskov, How copper catalyzes the electroreduction of carbon dioxide into hydrocarbon fuels, Energy Environ. Sci. 3 (9) (2010) 1311–1315.
[59] J.K. Norskov, T. Bligaard, A. Logadottir, J.R. Kitchin, J.G. Chen, S. Pandelov, J. K. Norskov, Trends in the exchange current for hydrogen evolution, J. Electrochem. Soc. 152 (3) (2005) J23–J26.

Xinyu Li received his Ph.D. degree from the School of Information and Physical Sciences, the University of Newcastle, Australia in 2022. He is now a Postdoctoral Researcher at the Australian Institute for Machine learning, the University of Adelaide. His research interests include machine learning, computational chemistry and their applications in energy material discovery.

Haobo Li received her Ph.D. in 2017 from Dalian Institute of Chemical Physics (DICP). In 2018 she moved to Technical University of Munich (TUM) as an Alexander von Humboldt postdoctoral fellow. She is currently a Research Fellow at the School of Chemical Engineering, The University of Adelaide, Australia. Her research interest focus on artificial-intelligence assisted design of catalysts.

Zhen Zhang received the bachelor's and Ph.D. degrees in computer science from Northwestern Polytechnical University, Xi'an, China. He is a Postdoctoral Research Fellow with the Australian Institute for Machine learning, the University of Adelaide. His research interests include graph neural networks, graph models, metalearning, and their applications in computer vision and scientific problems.

Javen Qinfeng Shi is a Professor in the School of Computer and Mathematical Sciences, University of Adelaide, and is the Founding Director of Probabilistic Graphical Model Group, and Director in Advanced Reasoning and Learning of Australian Institute for Machine Learning (AIML). His core expertise includes Probabilistic Graphical Models, Causation, and Deep Learning. Google Scholar ranks him 7th globally in Probabilistic Graphical Models. He has transferred his research to diverse industries including agriculture, mining, sport, manufacturing, bushfire, water utility, health and education.

Yan Jiao received her Ph.D. in 2012 from the University of Queensland. She is currently an Associate Professor and ARC Future Fellow at the School of Chemical Engineering of the University of Adelaide. Her current research area is computational electrochemistry, which involves investigating the nature and origin of electrocatalytic activity on nanostructured materials and developing novel catalysts for clean energy conversion reactions using Density Functional Theory computations.

Shi-Zhang Qiao received his PhD degree in Chemical Engineering from the Hong Kong University of Science and Technology in 2000, and is currently a Chair Professor and ARC Laureate Fellow at the School of Chemical Engineering of the University of Adelaide, Australia. He is also an elected Fellow of Australian Academy of Science. His research interests include synthesis and characterisation of nanomaterials as well as their applications in catalysis, energy storage and conversion (fuel cell, batteries, electrocatalysis and photocatalysis).