REVIEW ARTICLE

Check for updates

Advancements in computational approaches for rapid metal site discovery in carbon-based materials for electrocatalysis

Somayeh Faraji, [ID] † Zhiyu Wang,† Paola Lopez-Rivera and Mingjie Liu [ID] *

Heterogeneous electrocatalysts exhibit immense potential for advancing energy technologies. However, the constraints associated with noble metals have sparked a surge of interest in the exploration of single-atom catalysts and metal–carbon hybrids as alternative options. Designing metal sites in carbon-based materials has demonstrated high activity, selectivity, stability, and cost-effectiveness in various electrochemical reactions. In spite of these advantages, the intricate nature of the designed structures and the expansive design space encompassing potential metal site structures pose formidable challenges in terms of experimental characterization and optimization. To address these challenges, computational approaches have emerged as powerful tools to accelerate the discovery of new metal sites in carbon-based materials and understand the structure–catalytic property relationships for electrocatalysis. In this review paper, we provide an overview of the state-of-the-art computational approaches from reported modeled structures, theoretical foundations of computational methods in modeling electrochemical reactions, to the data-driven approaches to accelerate new catalyst design. We summarize the utilization of structure-binding energy relationships, virtual high-throughput screening methods, and machine learning techniques to explore a wide range of metal site structures and identify promising candidates for experimental validation. Furthermore, the review highlights the importance of considering the solvent effect and the impact of spin/oxidation states on extra electron transfer to enhance the accuracy of predicting binding energies. Finally, we summarize the current challenges and offer a brief perspective on future opportunities in the field of computational acceleration for carbon-based catalyst development.

1 Introduction

Electrochemical reactions are critical for sustainable energy and environment, as they form the basis of numerous essential technologies like batteries, fuel cells, and electrocatalysis.[1] Heterogeneous electrocatalysts facilitate electrochemical reactions; thus, developing effective heterogeneous electrocatalysts is crucial for advancing these technologies and promoting sustainable energy and environmental practices. The bulk metals exhibit striking electrocatalytic properties for small molecule activation, for example, Pt for oxygen reduction reaction (ORR) and hydrogen evolution reaction (HER),[2,3] Cu for CO_2 RR,[4,5] and Ru, Pd, and Ag for NRR.[6–9] However, most of them are either noble metals or exhibit poor stability, selectivity, and activity. These limitations emphasize the necessity of developing new electrocatalysts to overcome these challenges.

Department of Chemistry, University of Florida, Gainesville, FL 32611, USA.
E-mail: mingjieliu@ufl.edu
† These authors contributed equally to this work.

Carbon materials, such as graphene and carbon nanotubes (CNT), hold great promise for electrocatalysis due to their high surface area, excellent electric conductivity, and remarkable thermal stability. Moreover, carbon is abundant and environmentally friendly, and has a multitude of allotropes that offer a diverse range of tunable properties through structural variations. The shortcoming of chemical inertness can be addressed by introducing defects and dopants which can act as active sites while the carbon materials behave as the substrate or host materials. For example, single atom catalysts (SACs) as one of the emerging heterogeneous catalysts consist of isolated metal atoms dispersed on a support material (*e.g.*, graphitic carbon, metals or metal oxides).[10] Numerous studies have demonstrated that carbon-based SACs and other metal–carbon hybrid structures (metal clusters on graphene or CNTs) are promising as the next-generation electrocatalysts for various electrochemical reactions, for example, water splitting, CO_2 RR, ORR, and NRR, with the potential to replace precious metal catalysts.[11–13] Therefore, the development of metal-doped carbon-based electrocatalysts has gained significant attention in recent decades

as they have the potential to meet the requirements of high activity, selectivity, stability, sustainability, and cost-effectiveness in electrocatalysis.[14–21]

Computational modeling based on density functional theory (DFT) has become increasingly important in catalyst design as it can offer insights into reaction mechanisms and reveal the structure–property relationship. Owing to advancements in computational power, the virtual screening of hundreds and even thousands of catalyst structures prior to experimental synthesis is now feasible. When combined with machine learning techniques that leverage large datasets, these capabilities can inform the design of new materials with superior catalytic performance.[22] These approaches have revolutionized scientific research and have led to numerous breakthroughs in the development of more efficient catalytic materials.

In this review, our main focus is the application of modeling and computation to expedite the design of carbon-based materials with carefully engineered metal sites. We first overview the well-studied modeled systems and summarize the quantum simulation-based approaches to understand the properties related to the electrochemical reactions (Section 2). Although these approaches can offer detailed insights into the structure–property relationships, enabling the design of efficient catalytic sites, the main challenge lies in the extensive design space and the inherent limitations in accuracy associated with these methods. To tackle these challenges, data-driven approaches such as high-throughput screening and machine learning models are employed. In the next two sections, we review the utilization of data-driven approaches to accelerate the exploration of the extensive design space (Section 3) and the ongoing endeavors to improve the accuracy of the computational models (Section 4). At last, we envision future opportunities from the computational perspective, poised to expedite the development of metal sites in carbon-based materials (Section 5).

2 Modeled structures and computational approaches

2.1 Overview of well-explored structures

Among carbon–metal materials, the SACs are the most well-studied systems. Experimentally, most 3d and 4d metal SACs are successfully synthesized, characterized and applied in several electrochemical reactions.[16–18] A recent experimental library reported 37 SACs embedded in N-doped carbon, characterized by their oxidation state, coordination number, bond length, and coordination element.[23] Computationally, DFT simulations extensively investigated the single metal atom coordinated with four nitrogen atoms (M–N_4) in the graphene system for CO_2RR,[24–28] ORR,[29–31] water splitting,[14,15] and nitrogen reduction reactions,[32–34] as the local structure is analogous to the metal-porphyrin cofactor in enzymes.[35]

In addition to M–N_4, the general M–N_x moieties in the graphene lattice have also been explored. The number of coordinated nitrogen atoms varying from 1 to 6 can be modeled using graphene edges[36,37] or defective graphene.[38–42] Besides graphene as the substrate to host the M–N site, other graphitic carbon structures are also considered, such as CNTs,[43–46] fullerenes,[47–49] graphyne[50,51] and graphdiyne.[52–54] The unique properties (e.g., local curvature, sp/sp² hybridized porous structure) provided by those allotropes can contribute to the enhancement of the catalytic reactivities. As N plays a crucial role in promoting the activity of the metal site, the C_xN_y systems with more N doped carbon systems are also examined (e.g., C_9N_4,[55–57] C_2N,[58–61] CN,[62–64] C_3N_4[65–68]). Through careful examination of both computational simulations and experimental characterizations, it has been determined that the M–N_4 active site may not be the most desirable option. For example, Ni coordinated with one or two N would show better reactivity compared with Ni–N_4 for CO_2RR;[69–72] and Co–N_5 exhibits better selectivity and stability compared to other Co–N sites.[73] In addition, based on DFT calculations, Ni–N_3 has been reported as a more favorable site for *COOH (carboxyl) formation compared to NiN_4 and is active for CO_2RR.[74] Iqbal et al. provided a detailed review on recent advances in the design and synthesis of noble metal (Ru and Au) and non-noble metal (Mn, Fe, Co, Ni, Cu, and Mo) SACs doped with various non-metals (B, N, P, and S) for NRR.[75]

Investigations have also been conducted on alternative coordinated atoms, apart from nitrogen. The most common alternatives are oxygen,[76] sulfur,[77–80] boron,[81] and phosphorus[82,83] with conclusions that some of the dopants coordinated with metal can show better catalytic activity than nitrogen[76,79,83] but the stability may decrease.[77] Dual metals[84] and metal trimers[85] are also investigated on graphene and graphdiyne. However, the metal clusters beyond that are barely explored due to the ambiguous metal cluster structures. Experimentally, the dual metal catalysts have been synthesized with Cu_2, CuNi,[86] FeNi,[87–89] and FeCo[90] and the improved performance for CO_2RR and ORR is explained by synergistic electronic modulation effects.[89–91]

2.2 Computational approaches

Computational approaches based on DFT simulations provide an atomistic understanding of catalytic reactivity and stability. In the next two subsections, we will provide a brief overview of the approaches and their application in understanding these properties.

2.2.1 DFT-based electrochemical modeling. In electrochemical reactions, a series of intermediate steps occur, and each step involves the transfer of electrons and protons. In heterogeneous catalysts, the electrochemical process happens at the electrode/electrolyte interface, where the electrode/electrolyte serves as the conductor for electron/proton respectively. Under equilibrium, an electrochemical system is a grand canonical ensemble as it can exchange electrons with the electrode, while the standard DFT is performed in the canonical ensemble where the charge of the system is fixed. Hence, it is challenging to accurately describe the practical reactivity dependence on the electrode potential with DFT. Here, we focus on reviewing two most commonly used computational methods for modeling electrochemical reactions: the computational hydrogen

electrode (CHE)[92–94] and the grand canonical version of DFT (GCDFT).[95,96]

The CHE approach, pioneered by Norskov et al., is a well-established and widely utilized method for incorporating insights from ab initio calculations into electrochemistry modeling.[94] The core principle of this approach involves defining a reaction pathway and considering that, at each step, the proton–electron pair was transferred together with the chemical potential as half of the gaseous hydrogen at a potential of 0 V. Then the Gibbs free energy for each step along the pathway at zero potential and zero pH is defined as $\Delta G = \Delta E + \Delta ZPE - T\Delta S$, where ΔE and ΔZPE are the reaction and zero-point energy difference of each step, respectively, and ΔS is the change in entropy. All of the values can be obtained from DFT calculations. ΔZPE is defined as $\frac{1}{2}\sum_i h\nu_i$, where ν_i are vibrational frequencies and i goes over all vibrational modes, which can be calculated via DFT with the harmonic oscillator approximation. ΔS could be taken either from standard tables[97] or by calculating translational (S_{trans}), rotational (S_{rot}), and vibrational (S_{vib}) entropies as described in ref. 22. To include the influence of the electrode potential U in an electrochemical step, the CHE method sets the total energy of the electrons in the electrode to $-qU$, where q is the charge of the electron. Therefore, the free energy of each step is

$$\Delta G(U) = \begin{cases} \Delta G(\text{reduction step, 0 V}) + nqU, \\ \Delta G(\text{oxidation step, 0 V}) - nqU, \end{cases} \quad (1)$$

where n is the number of transferred electrons. In addition, the chemical potential of H^+ and e^- is referenced to that of the H_2 molecule using the standard hydrogen electrode $\left(\frac{1}{2}H_2 \leftrightarrow H^+ + e^-\right)$. Finally, at a pH different from 0, the free energy of H^+ ions can be corrected by the concentration dependence of the entropy as $\Delta G(pH) = -kT\ln[H^+]$.[94] The CHE model therefore can be used to plot the free energy diagram, which provides an atomistic picture of the reaction computationally. The free energy diagram can be analyzed to determine the onset potential that can be compared with experimental values directly.

The CHE method could be considered as a post-processing scheme and constant charge approach in which the U-dependency of $\Delta G(U)$ comes only from nqU, assuming the reaction energy of each step is independent of the electrode potential. This simplification makes CHE a valuable tool for a quick assessment of the general thermodynamic trends but would fail to accurately describe the potential dependence in a reaction. Over the past few years, efforts have been made to connect the constant charge scheme to the constant electrode potential scheme by varying the electron in the system,[98] or the size of the supercell[99] so that the constant electrode potential can be extrapolated. Among those efforts, the GCDFT provides a direct simulation scheme to keep the electrode potential constant in the system by controlling the work function. This approach calculates the grand free energy ($\Omega = A - \mu N$, where A is Helmholtz free energy and N is the number of atoms with

chemical potential μ) by solving the Kohn–Sham equations while the number of electrons at the electrode varies to keep a constant Fermi level that corresponds to the applied potential. The details of the algorithm can be found in ref. 96.

Both methods have been applied to study the metal–carbon systems for electrocatalysis. The CHE method with the advantage of computational efficiency has been extensively applied in a wide range of metal–carbon systems for electrocatalysis to quickly screen new catalysts with the sacrificed accuracy requirement. The GCDFT method (or other constant potential approaches) which provides a more accurate description of the electrochemical process has been used specifically in certain systems to fully understand the reaction mechanisms with the potential dependence. However, due to the computational cost of GCDFT, it is challenging to apply it in high-throughput screening of catalyst design. On the one hand, when comparing the results from both models, some trends are consistent. For example, in Fe–, Co–, and Ni–N_4 in graphene, the Ni–N_4 site is the most effective site for CO production, while the Fe–N_4 site can stabilize CO adsorption, which can potentially further reduce CO to other products such as CH_3OH and CH_4.[27,100] On the other hand, the electrode potential-dependent processes, such as CO_2 adsorption and the transition states, are critical in revealing the reaction mechanism. This indicates that the rate-determining step and catalytic selectivity predicted by GCDFT and CHE may be different[101,102] (see Fig. 1 as an example). In this case, the GCDFT is expected to provide more accurate descriptions due to the nonlinear effects of the applied potential on molecular adsorption and transition states.[27,72,102] In particular, in metal–carbon systems, the localized orbital would be dramatically influenced by the electrode potential due to the low density of states in 2D materials.[101] Therefore, the orbital-based design principle needs to be examined carefully by taking the electrode potential into account.

2.2.2 Stability assessment. Besides catalytic reactivity, structural stability is another crucial property. In SACs, the metal atom forms strong covalent bonds with the surrounding atoms, known as the covalent metal–support interaction (CMSI),[103] which plays a vital role in stabilizing the metal site within the carbon lattice. However, studies have also shown that certain metals, such as Fe[104] and Cu,[105–107] undergo aggregation in graphene. This mobility and structural transformations make the initially designed structure or active site unstable, consequently affecting the catalytic activity under experimental conditions. Therefore, it is imperative to meticulously evaluate the stability of the catalyst, especially under operando conditions, in order to facilitate a rational computational design and acquire a comprehensive understanding of the durability of these materials. Computationally, there are several common approaches to evaluate the stability, which can be categorized as thermal stability, thermodynamic stability, and electrochemical stability. Thermal stability is commonly examined using molecular dynamics (MD) simulations, which allow for the analysis of the structural changes at varying temperatures.[62,85,108,109] While the assessment of thermal stability is relatively straightforward, we will now provide more
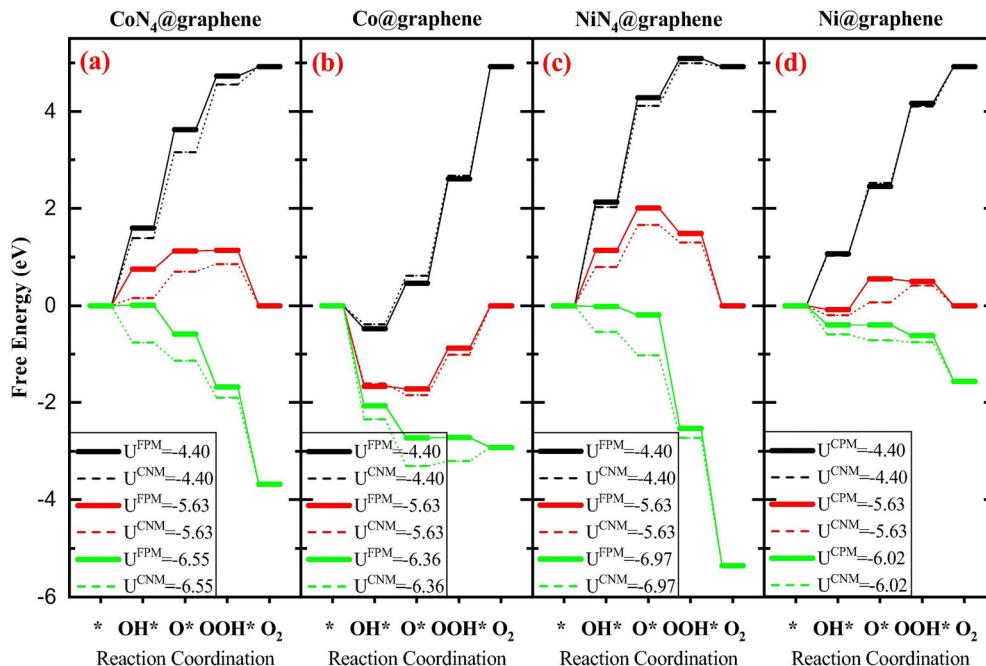
Fig. 1 The free energy diagrams associated with the OER on four different systems: (a) CoN4@graphene, (b) Co@graphene, (c) NiN4@graphene, and (d) Ni@graphene. FPM is GCDFT and CNM is CHE. Reprinted with permission from ref. 101. Copyright 2020 Elsevier.

details regarding thermodynamic stability and electrochemical stability.

Thermodynamic stability. To quantify the thermodynamic stability, the binding energy (E_b), formation energy (E_f), and diffusion barriers are usually calculated as criteria. The first criterion, E_b, is defined as

$$E_b = E_{M@sub} - (E_{sub} + E_M), \qquad (2)$$

where $E_{M@sub}$ and E_{sub} are the total energies of the substrate with and without the metal atom, and E_M is the energy of the isolated metal atom. According to this definition, a negative E_b indicates that the structure is thermodynamically favorable (the more negative the E_b, the more stable the structure).

The second criterion, E_f, employs a different reference energy of atoms to estimate the thermodynamics stability of the structure. For example, in the case of one metal anchored in the divacancy of N-doped graphene, the metal atom is coordinated with four atoms (*i.e.*, $M@N_xC_{4-x}$), where the metal atom is bonded with x nitrogen atoms ($x \leq 4$) and $4 - x$ carbon atoms. The value of E_f can be calculated using the following equation:

$$E_f = E_{M@N_xC_{4-x}} + (x + 2)E_C - (E_{psub} + xE_N + E_{M^0}). \qquad (3)$$

Here, $E_{M@N_xC_{4-x}}$ and E_{psub} are the total energies of the system and the pristine substrate, and E_C, E_N, and E_{M^0} are the energies of one carbon atom, one nitrogen atom, and one metal atom. The energy of atoms can be obtained from the most stable phase of each atom (*e.g.*, E_C as the carbon atom in graphene, E_N as the nitrogen atom in N_2, and E_{M^0} as the metal atom in the bulk phase). The formation energy, depending on the reference

states of the atoms, represents the energy required to form the designed structures from the stable phases of those atoms.

It is also important to consider the stability of the metal atoms against aggregation. This could be examined by comparing the E_b of the metal in the carbon substrate and the metal bulk cohesive energy (E_{coh}) which is defined as $E_{coh} = E_{M^0} - E_M$. Therefore, satisfying both conditions of $E_b < 0$ and $E_b > E_{coh}$ indicates the metal atom's propensity to aggregate. The mobility of the metal atoms in the carbon substrate has been employed as a means to assess their stability from a kinetic perspective. The diffusion barrier of the metal atom migrating in the substrate can be obtained by searching the transition states in the migration path. The lower the diffusion barrier, the easier the metal can migrate, potentially leading to favorable aggregation.

The mentioned criteria were employed in various studies. Some studies have only considered E_f for the stability.[55,110] Others have employed a combination of those criteria.[34,85,108] For example, the stability of Fe–N–C and Mn–N–C can be assessed based on E_f, E_b, diffusion barriers, and E_{coh}.[38,111]

Electrochemical stability. The limitations of thermal and thermodynamic stability lies in their failure to consider the impact of electrode potential, pH, or electrolyte in actual catalytic reactions. In contrast, the evaluation of electrochemical stability specifically takes these factors into account. In the following, we will briefly discuss some attempts to address the electrochemical stability.

One way to evaluate the electrochemical stability is to consider whether the embedded metal atom will dissolve in

solution. This can be examined by the dissolution potential U_{diss} (relative to the SHE),[112] which is calculated as

$$U_{diss} = U_{diss}^0(\text{metal,bulk}) - (E_f/nq). \quad (4)$$

Here $U_{diss}^0(\text{metal,bulk})$ is the standard dissolution potential (pH = 0) of the bulk metal, experimentally determined for pure metals in an aqueous solution; n is the number of electrons transferred during dissolution, which depends on the specific chemical reaction that takes place; and q is the electron charge. According to this definition, a more positive U_{diss} relative to the equilibrium potential indicates that the metal atom binds the substrate strongly, and the dissolution of metal atoms can be avoided under electrochemical conditions. The approach was initially proposed for surface alloys and then applied to SACs in N-doped carbon materials for catalytic reactions such as HER and OER,[113] CO_2RR[114] and ORR.[110,115]

Another way to examine the electrochemical stability under operando conditions is through the Pourbaix diagram. The Pourbaix diagram depicts the thermodynamics stability regions of different chemical species in an aqueous solution as a function of electrode potential and pH.[116] Examples of constructing Pourbaix plots for SACs can be found in the literature.[117–120] For example, Holby et al.[120] proposed an approach based on DFT-calculated properties to explore the electrochemical stability of $Fe–N_4$ in bilayer graphene. The presented methodology can take reaction environment variables into account, such as pH, electrode potential, and reaction intermediates. Experimental observations regarding relative stability in acid vs alkaline environments, dissolution of Fe at low potential, and the possible role of O_2 can be well interpreted within the presented framework. There are also some other proposed approaches to incorporate environmental conditions to make results comparable with experiments. For example, experimentally reported reversible transformation between Cu single atoms and clusters under the realistic reaction conditionss[105,107] can be explained by the constant potential hybrid-solvation dynamic model.[121]

The computational frameworks offer a unique approach to comprehending the structure–property relationship at an atomistic scale. By combining DFT-based simulations with fundamental physical principles, these frameworks can provide valuable design principles to guide experimental synthesis. The current focus in designing metal sites with carbon materials primarily revolves around manipulating the coordination environment surrounding single or dual metal sites within graphitic carbon structures. Through a combination of experimental synthesis and characterization, the incorporation of metal atoms into carbon-based materials exhibits tremendous potential for electrocatalysis. Given the diverse methods available to engineer the molecular interactions between metals and carbon during materials design, as well as the intricacies of electrochemical processes, the development of data-driven and accuracy-driven models has emerged as a solution to address challenges and expedite catalyst design. In the subsequent sections, we will delve into the exploration of these two aspects

separately, specifically in relation to the design of atomic metal sites in carbon-based materials.

We aim to provide an overview of recent successful applications of data-driven models in the field of designing metal atomic sites in carbon catalysts, encompassing high throughput screening, machine learning, and identification of key descriptors. In order to enhance the accuracy of computational modeling, we also review approaches for incorporating solvation effects, as well as electron and proton transfer mechanisms, spin states, and oxidation states into the models.

3 Structure–binding energy relationship

Catalysts are composed of various elements and combinations, which result in a high-dimensional composition space to be explored for specific catalytic reactions. In addition, catalytic reactions often involve multiple intermediates and reaction pathways which make the mechanism complex. Besides, catalyst properties depend on different factors such as surface morphology and electronic properties. Therefore, performing ab initio calculations for all possibilities becomes very challenging because of the high computational cost. The data-driven approaches to address these challenges will be discussed in this section.

3.1 Linear scaling relationship and BEP relationship

The scaling relationship is a fundamental principle in catalyst design. It simplifies complex catalytic reactions which involve multiple reaction steps by finding relationships between those steps. There are two well-accepted scaling relationships for catalyst design: the linear scaling relationship and the Bell–Evans–Polanyi (BEP) relationship. The linear scaling relationship refers to the relationship between the adsorption energies of intermediates that bind to the catalyst surface through the same atom. For example, in CO_2 RR, the possible intermediates are *COOH, *CO, *CHO, *COH, *CH_2O, *OH, and *OCH (* denotes the adsorption site). DFT studies revealed a linear relationship between $E_{ads}[CH_2O]$, $E_{ads}[CHO]$, and $E_{ads}[COOH]$ against $E_{ads}[CO]$ on transition-metal surfaces.[122] Based on the scaling relationship, the complete free energy profile can be depicted by considering the adsorption energy of one (or a few) intermediate(s) on the same adsorption site. From these adsorption energies, the activity and selectivity can be inferred.[123–125] The linear scaling relationship can be explained through the bond order conservation theory.[126–128]

The BEP relationship establishes a connection between the activation energy (energy of the transition state) and reaction enthalpy (the enthalpy change in an elementary reaction). It states that these two quantities are proportional to each other. This idea was first demonstrated by Evans and Polanyi in 1937, who showed that an increase of reaction enthalpy causes a decrease in activation energy in classic molecular reactions.[131] After that, Michaelides et al. found that the linear relationship between activation energy change and enthalpy change was

also established well for heterogeneous catalysts for the elementary reaction.[132] Since surface redox reactions usually involve multiple steps, Cheng and Hu conducted a study demonstrating that multistep surface redox reactions can be effectively simplified as one-step adsorption or desorption processes.[133] As a result, the adsorption energy of the reaction species can be considered as the descriptor for catalyst screening. Based on these scaling relationships, the reactivity of catalysts (*e.g.*, the exchange current density, or overpotential) shows a volcano relationship with the energetic descriptor.[134,135]

Both scaling relationships have been examined in carbon-based SACs for various reactions. For example, in SACs on N-doped graphene and C_2N, the scaling relationship was established between the intermediates of ORR where ΔE_{ads} between *OH and *O exhibited a slope close to 2, and *OH and *OOH showed slopes close to 1.[60,136–138] For CO_2 RR, a linear relationship with a slope close to 1 was observed between the adsorption energy of *CO and *COOH or *CHO on a single metal site with C-, N-, or B-doped graphene and C_2N (Fig. 2(a)).[59,83,129] For NRR, a scaling relationship was established between the adsorption energy of N_2H* and NH_2 with a slope close to 0.8[139–141] on metal (M) in single porphyrin sites and N-doped graphene (M = Cr, Mn, Fe, Co). The BEP relationship was also examined for SACs in carbon-based materials. For example, to examine the activation energy for the transition state in ORR reaction, Wang *et al.* investigated the BEP relationship on zigzag, armchair, and basal plane of graphene and graphene oxide. A clear linear relationship was observed between the energy of the transition state and the chemisorption energy of the O atom with different functional groups.[142] Similar results were also observed in carbon-based SAC systems. Choi *et al.* studied 13 transition metals coordinated with N and C in graphene (denoted as $M@N_xC_y$) for C–H bond activation. The

linear relationship between the activation energy and H* adsorption energy was fitted as $E_a = 0.79E(H*) + 2.07$ for all metal and coordination environments (Fig. 2(b)).[130]

With the scaling relationships demonstrated for metal–carbon materials, the parameters that tune the adsorption energies of the key species can be used to screen new catalysts or catalytic sites that provide the best reactivity. However, it should be noted that the scaling relationships only apply to adsorbates that are adsorbed on the same site through the same type of atom that forms the chemical bonds with the site. Therefore, the application of linear relationships for catalyst design should be complemented by a thorough identification of active sites.[143] Compared to transition metal catalysts, the active sites in metal–carbon materials exhibit greater diversity. The scaling relationship, as a consequence, can be modified if different adsorption sites are introduced.[91,144,145] For example, both metal and non-metal as adsorption sites were observed in $M@N_4C_4$ moieties (M = Fe, Co, Ni) for ORR. *O and *OH were found to preferentially adsorb on the carbon site, while *OOH favored the metal site, circumventing the scaling relationship between their adsorption energies.[14] Similar strategies have also been used for NRR and CO_2 RR on the C_3N monolayer and PC_6.[139,144,146]

3.2 High-throughput screening of metal-centered active sites with binding energies

One of the applications of the scaling relationships is virtual high-throughput screening (VHTS). The VHTS can rapidly screen and evaluate a large number of materials and active sites, and thus the process of catalyst discovery and optimization can be accelerated. In metal–carbon materials, both metal and carbon substrates can be tuned through various strategies. For instance, optimizing the activity can be achieved by using
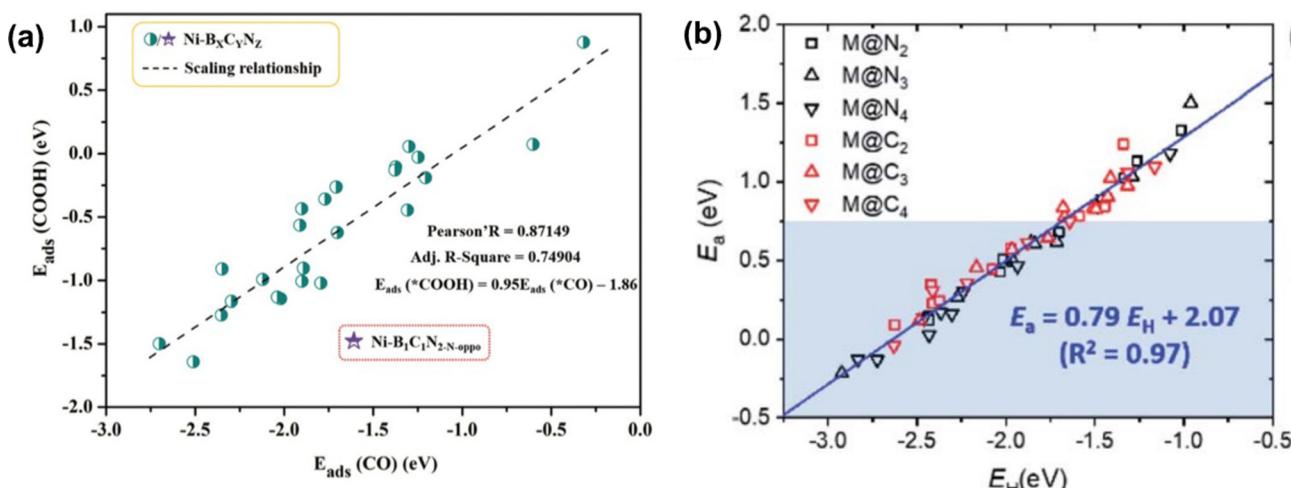
Fig. 2 (a) Scaling relationship between the adsorption energy of *CO and *COOH. Reprinted with permission from ref. 129. Copyright 2021 Royal Society of Chemistry. (b) The scaled relationship between the activation energy and adsorption energy of H atom (BEP relationship) was examined on single-atom catalysts. Reprinted with permission from ref. 130. Copyright 2021 Royal Society of Chemistry.

different metals,[100,136] different host materials,[147] and different coordination environments.[55,56,65,66,68]

Such flexibility in metal–carbon electrocatalysts was probed by Guo *et al.* They screened a total of 210 carbon-based SACs for ORR[110] by exploring 30 different transition metals in combination with various coordination and substrate materials, including defective graphene, N-doped graphene, and phthalocyanine (Pc). Among them, 31 SACs have the potential to break the scaling relationship, and high activity and selectivity can be achieved towards hydroperoxide by screening the adsorption energy of O* (Fig. 3(a) and (b)). In their study, 7 SACs showed higher activity and selectivity than the benchmark $PtHg_4$ in acidic media. $Zn@Pc–N_4$ was recognized as the best catalyst with a small overpotential of 0.15 V, which was supported by experiments.[148] In NRR, multi-step screening strategies have been used to search for the optimized catalysts.[139,149,150] Ling *et al.* screened 540 SACs in nitrogen-doped graphene by varying metal coordinated atoms and single metal species. They found that W_1C_3 site exhibited the most outstanding performance with an extremely low onset potential of 0.25 V.[150] In contrast, fewer VHTS have been applied for CO_2RR.[100,144] Guo *et al.* explored 72 candidates of 12 metals with 6 different coordination environments. By screening the ΔE_{ads} of *CO, *COOH and competing adsorption of *H, they found that $Fe–N_4$, $Ni–N_4$, $Cu–N_4$, $Pd–N_4$, and $Pd–N_4$ are the most selective sites for CO_2RR against HER.

In addition to the energetic screening for the adsorbents, stability has also been considered in the screening process.[68,100,110] In a previous example,[110] before screening the adsorption energy of O*, the structures were first screened regarding the thermodynamic and electrochemical stability. The criterion for stability was $U_{diss} > 0$ and $\Delta E_f < 0$ as we explained in Section 2.2.2. Based on the examination of stability, the substrates such as BN, $Pc–N_4$, and $Py–N_4$ are found to be good supports to host metal atoms. In terms of metal species, transition metals from group 8–12 are more likely to be stabilized under reaction conditions. In total, 146 SACs out of 210 were screened out (Fig. 3c), and $Zn@Pc–N_4$ is verified as the best candidate for both stability and high activity for 2-electron ORR. As a support, these active moieties also have been synthesized successfully including $Ni–N_4–C$, $Fe–N_4–C$, and $Co–N_4–C$.[14]

VHTS is a powerful tool that greatly reduces the effort required for conducting experiments. This approach enables the exploration of various tuning parameters simultaneously. However, performing a large number of DFT calculations is still computationally expensive. Machine learning-assisted high-throughput screening can be further leveraged to accelerate the VHTS.

3.3 Machine learning for adsorption energies

As discussed, the energetic descriptor such as ΔE_{ads} can be used as the property in VHST to screen new catalysts; however, calculating the energetic property from DFT to explore the vast design space is time-consuming. A surrogate model based on machine learning methods provides a faster way to predict
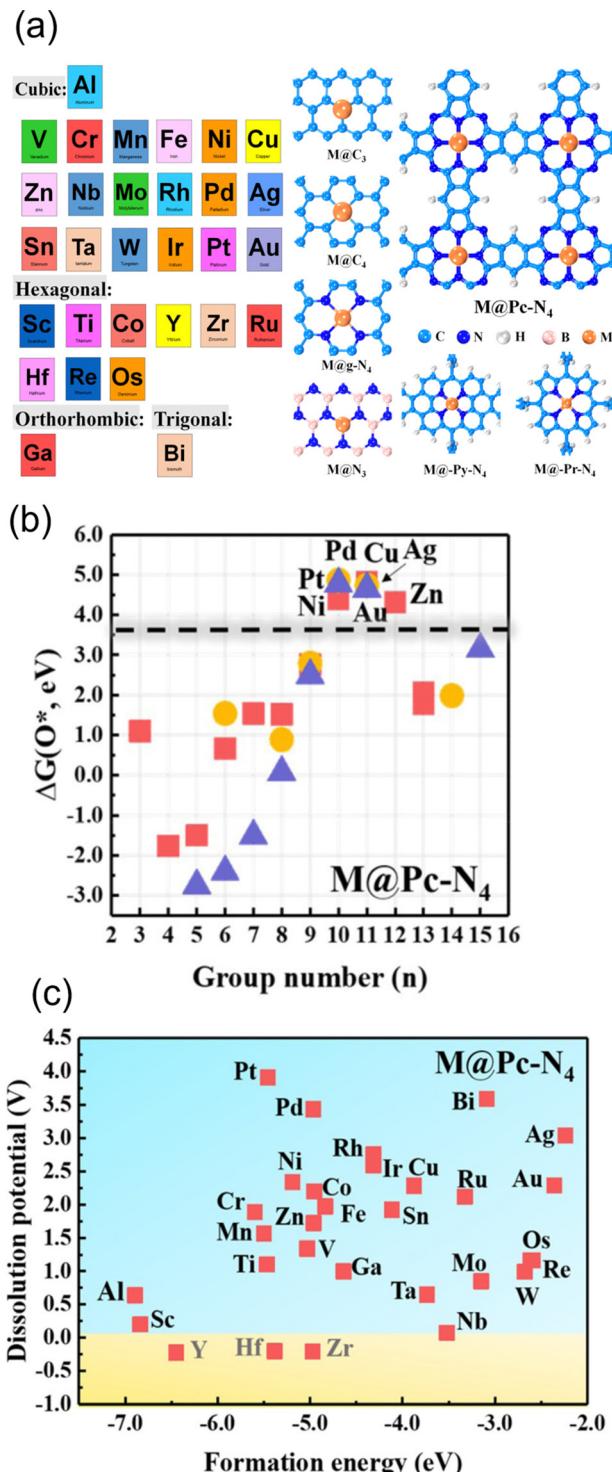
Fig. 3 (a) Schematic illustration of screened metal atoms with their most stable bulk structure and all considered substrates. (b) The O* adsorption energy screening of metal from group 3–15 on $Pc–N_4$ configuration. Squares, circles, and triangles represent Al, 3d transition metal, 4d transition metal, and 5d transition metal, respectively. (c) The stability screen for metal on $Pc–N_4$ configuration. Reprinted with permission from ref. 110. Copyright 2019 American Society of Chemistry.

ΔE_{ads}.[151] The properties that machine learning model get trained to predict adsorption energy are known as features.

By selecting key features, machine learning models can not only predict accurate ΔE_{ads} but also provide insights into the underlying mechanisms and design principles. In SACs, several machine learning models have been explored, including deep neural network (DNN),[152] random forest regression (RFR),[60,153] support vector machine (SVM),[154] gradient boosted regression (GBR)[68,155,156] and sure-independence screening and sparsifying operator (SISSO).[157,158] The training data for these models range from ten to thousands. For example, Kaspe *et al.* combined DFT simulations and machine learning models to predict the ΔE_{ads} of *OH in carbon-based materials for ORR.[154] They explored 2500 sites in 0D (graphene flakes), 1D (graphene nanoribbons), and 2D (graphene) doped with S, P, SO$_2$ and PO$_2$, and applied six machine learning models with 13 electronic and structural features. The SVM achieved the best performance in testing, with an R^2 coefficient of determination of 0.87, followed by multiple linear regressor ($R^2 = 0.84$), K-nearest neighbor (KNN) ($R^2 = 0.84$), RFR ($R^2 = 0.83$), Extreme Gradient Boosting Regressor ($R^2 = 0.79$), and Decision Tree Regressor ($R^2 = 0.78$). Fischer *et al.* studied 1700 transition metals doped in defective N-doped graphene by varying the metal species, the number of dopants, and the number of vacancies for the adsorption energies of O and H adsorbate species. The best performance was given by the RFR model with MAE = 0.31 eV for testing in predicting the ΔE_{ads} of *O compared to SVM with MAE = 0.32 eV. Umer *et al.* screened 364 SACs by anchoring 3d–5d single metal atoms on various substrates. A combined DFT-machine learning framework found that mono- or dual-type

non-metal (B, N, and P) doping in g-C$_3$N$_4$, two-dimensional π-conjugated polymer, graphene and hexagonal boron nitride can substantially improve HER activity and stability. For predicting the adsorption energy for H*, Light Gradient Boosting Machine (LGBM), Histogram-based Gradient Boosting Regressor (HGBR), Extremely Randomized Trees (ERT), Gradient Boosting (GB), Bagging, XGBoost regressor (XGBR), RFR and AdaBoost regressor (ABR) are used. The top three performances are given by LGBR with RMSE = 0.20 eV, HGBR with RMSE = 0.24 eV and ERT with RMSE = 0.26 eV (Fig. 4(a)).

Besides predicting ΔE_{ads}, machine learning models have also been used to reveal the structure–property–performance relationship where there are multiple properties that could influence the catalytic performance. For instance, Zhu *et al.*[155] employed DFT and machine learning to study ORR on dual-metal site catalysts (DMSCs) in the carbon substrate and provide a comprehensive depiction of the relationship between intrinsic properties of these catalysts and their catalytic activity. They found that electron affinity, van der Waals radius, and electronegativity of two metal sites are important properties related to the adsorption energy of OH*. In another example, Deng *et al.* employed a DFT-machine learning study to understand the origin of the activity of bi-atom catalysts on N-doped graphene for ORR. They considered 20 electronic and geometric features in an RFR model, and the ranking of feature importance directly translated from the weight of each feature. Two dominant features, M–M (distance of two metals) and M–N (average distance between the metal and coordinated N atoms),
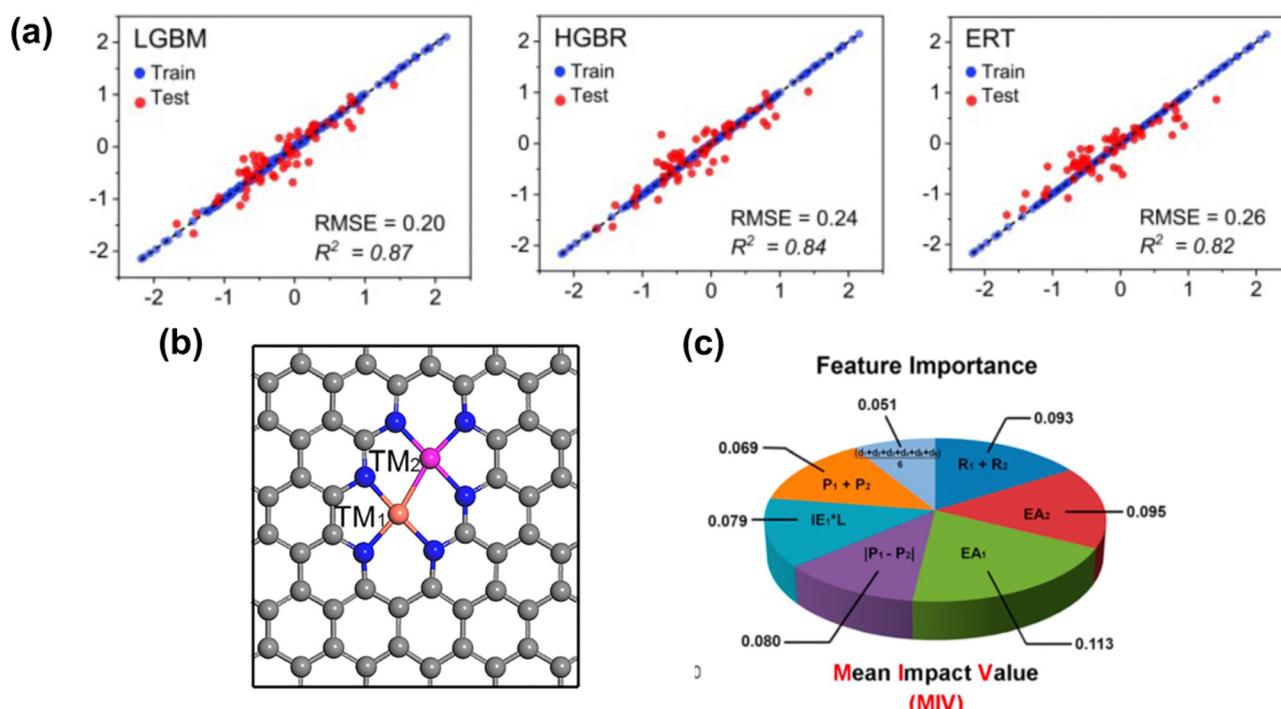
Fig. 4 (a) Machine learning model performance for predicting the adsorption energy of H* on 2D carbon materials. (b) Schematic illustration of the bi-metal site on graphene. (c) Mean impact value (MIV) employed for considering feature importance. Reprinted with permission from ref. 159. Copyright 2022 Royal Society of Chemistry.

are identified. This indicates the ineligible role of local geometry in the ORR activity.[153] Zafari *et al.* performed high-throughput screening among B-doped graphene SACs by using machine learning methods for NRR. They employed principal component analysis (PCA) to reduce the number of features before determining the feature importance by the weight of the mean decrease in accuracy (MDA) obtained from a DNN. The results of the MDA analysis showed that the three most important features in determining the ΔE_{ads} of intermediates were metal coordination number, atomic radius of the metal, and fraction of boron among coordination atoms. It suggested that both geometric and electronic features show a significant impact on ΔE_{ads}.[152]

To summarize, the machine learning models have been employed to predict ΔE_{ads} to screen and discover promising catalysts, and the key properties can be identified to unravel the structure–performance relationship. Such data-driven strategies can screen out optimal catalysts that satisfy certain criteria at a large scale, greatly promoting the rational design of catalysts.

3.4 Identified descriptors

The selected features, which impact ΔE_{ads}, serve not only to comprehend the catalysts but also to predict their catalytic performance. These selected features, also known as descriptors, play a crucial role in this regard. Significant research efforts have been dedicated to identifying descriptors that exhibit a correlation with catalytic performance. In the following, we summarize the proposed descriptors for carbon-based electrocatalysts.

3.4.1 Physical-based descriptors. Some of the proposed descriptors are based on a combination of fundamental physical principles and chemical intuitions. Sabatier's principle, for instance, provides valuable guidance by stating that the binding energy of reaction intermediates on a catalyst's surface must be neither too strong nor too weak for optimal activity. However, calculating all molecular binding energies from *ab initio* can be expensive. To overcome this problem, descriptors are often considered. One of these descriptors is the Newns–Anderson model,[160] which establishes a connection between the binding energy on transition metal surfaces and the energy of the electronic d-band states relative to the Fermi level. This model offers a physics-based electronic structure characteristic that correlates with the binding energy, known as the d-band model, providing a useful tool to understand catalytic activity. An overview of the d-band model and beyond it for the rational design of electrocatalysts can be found in ref. 161. Following chemical intuition, it has been found that the binding energies of single O and C atoms are proportional to the binding energies of molecular species that coordinate with O or C atoms on the catalyst surface. This finding offers simple scaling relationships[162] that can be utilized to estimate binding energies. In addition to these simple descriptors, more advanced descriptors have been proposed that push the boundaries of catalyst characterization and design. For example, based on the relationship between the charge transfer of the transition metal and the adsorption energy of OH ($\Delta E_{ads}[OH^*]$),

Niu *et al.*[163] explored transition metals in g-C$_3$N$_4$. Based on their studies, they established a new descriptor ϕ for predicting ORR activities as the multiplication of the number of electrons in the d orbital (θ_d) and the electronegativity of the transition metal (E_{TM}) as $\phi = \theta_d \times E_{TM}$. By plotting $\Delta E_{ads}[OH^*]$ *vs.* ϕ, a strong correlation with $R^2 = 0.90$ was shown which indicates that the adsorption strength of OH and other intermediates can be evaluated through this descriptor due to the scaling relationship. To predict ΔE_{ads} of small molecules on different substrates, Gao *et al.*[164] introduced a descriptor that is based on the electronic properties as $E_{ads} = k(S_v^2/\chi^\beta) + b$, where S_v is the number of valence electrons in both d and s orbitals, χ is the electronegativity of the transition metal and β is an index determined by the role of d- and s-orbitals in valence descriptions and electronegativity. In addition to electronic properties, it has been shown that geometric properties, such as bond length, bond angle, and coordination number, could be used as descriptors for ΔE_{ads}. For example, to predict the ΔE_{ads} of *OH and *OOH for ORR activity on N- and B-doped carbon nanotubes, the pyramidalization angle (θ_p) has been used as a geometric descriptor.[165]

3.4.2 Fitting-based descriptors. Moving beyond physical-based descriptors, another approach to obtaining descriptors involves fitting-based methods. In this approach, the descriptor toward target property is fitted by a set of intrinsic properties. For instance, Xu *et al.*[166] proposed a universal descriptor, ϕ, based on the intrinsic properties of the active center to predict the ORR/OER/HER activity through ΔE_{ads} in N-doped defective graphene-based SACs. The descriptor is formulated as $\phi = \theta_d \times$

$$\frac{E_M + \alpha \times (n_N \times E_N + n_C \times E_C)}{E_{O/H}}$$ which combines the valence electrons in the occupied d orbital of the metal atom (θ_d) with the electronegativities of the metal (E_M), nitrogen (E_N), carbon (E_C), oxygen (E_O), and hydrogen (E_H) atoms, as well as the number of nearest-neighbor nitrogen (n_N) and carbon (n_C) atoms. The proposed descriptor could successfully predict the activity order of previously reported experimental SACs on graphene, and it can be used to search for the optimal active centers for different reactions (guiding the synthesis of SACs that outperform precious-metal-based catalysts). It was also found that the descriptor ϕ shows a linear relationship when plotted against the ΔG of OH* and H* and fitted through linear regression: $\Delta G[OH^*] = 0.11\phi - 2.48$, $\Delta G[H^*] = -0.77\phi + 1.27$ when $\phi < 27$ or $\Delta G[H^*] = 0.04\phi - 1.43$ when $\phi > 27$. As another example, Wang *et al.*[114] demonstrated that descriptors combining electronic and geometric properties can effectively screen electrocatalysts for CO$_2$ RR. The proposed descriptor

takes the form $\phi = \dfrac{V_M \left[X_M \left(\prod\limits_{L=1}^{N} \frac{1}{d} X_L \right) \right]^{1/(N+1)}}{X_{C/O}}$, which incorporates the number of d electrons of a metal atom (V_M), the electronegativity of the metal atom (X_M), its nearest atom (X_L), adsorbate C (X_C) or O (X_O) atom connected to the metal atom, and the bond length between the metal atom and its nearest neighbor atom (d). This descriptor accurately describes the $\Delta E_{ads}[*CHO]$ for CO$_2$ RR on M–N$_4$–C SACs, as determined in

Table 1 A summary of reported descriptors (ϕ) for carbon-based catalysts for various reactions to predict adsorption (free) energies and catalytic activity. CNT stands for carbon nanotube

Category	ϕ	Property	System	Reaction	Ref.
Physical-based	d-Band center	$\Delta E_{ads} \propto$ d-band			160
	$\theta_d \times E_{TM}$	$\Delta E_{ads}[OH^*]$	g-C$_3$N$_4$	ORR	163
	$k(S_v^2/\chi^\beta) + b$	ΔE_{ads}	Various substrates	CO$_2$ RR, NRR, ORR	164
	θ_p	$\Delta E_{ads}[OH^*]$, $\Delta E_{ads}[OOH^*]$	N-Doped CNT, B-doped CNT	ORR	165
Fitting-based	$\theta_d \times \dfrac{E_M + \alpha \times (n_N \times E_N + n_C \times E_C)}{E_{O/H}}$	Activity, $\Delta G[OH^*]$, $\Delta G[H^*]$	N-Doped graphene SACs	ORR, OER, HER	166
	$\dfrac{V_M\left[X_M\left(\prod\limits_{L=1}^{N}\dfrac{1}{d}X_L\right)\right]^{1/(N+1)}}{X_{C/O}}$	$\Delta E_{ads}[^*CHO]$	M–N$_4$–C	CO$_2$RR	114
	$(4 \times V_{M_1} \times (E_{M_1} + E_{N/O})/E_{N/O} + V_{M_2} \times (E_{M_2} + E_{N/O})/E_{N/O}) \times L_{M_1-N/O}$	Catalytic activity	Bimetallic 2D MOFs	CO$_2$RR	167
Machine learning-based	$-1.032 \times (e_d/q) + 13.424 \times (1/r_{cov}) + 1.726 \times (e_d \times EN) - 0.045 \times d_{occ}^2 - 9.241$	$\Delta G[H^*]$	N-Doped graphene SACs	HER	157
	$0.69 + 0.00029\sqrt{EA_1} + 0.15EA_1 - 0.067\sqrt{EA_2} + 0.0076EA_2 - 0.012(R_1 + R_2) - 1.19\|P_1 - P_2\|^2$	$\Delta G[OH^*]$	Dual-metal catalysts	ORR	155

a previous study. Additionally, a linear relationship was successfully established between $\Delta E_{ads}[^*CHO]$ and ϕ with $R^2 = 0.70$. Zhang[167] studied bimetallic 2D MOFs for CO$_2$ RR and proposed the descriptor $\phi = (4 \times V_{M_1} \times (E_{M_1} + E_{N/O})/E_{N/O} + V_{M_2} \times (E_{M_2} + E_{N/O})/E_{N/O}) \times L_{M_1-N/O}$, where E, V, and L represent electronegativity, the number of valence electrons, and bond length, respectively. By using this descriptor, they proposed CoPc–Zn–O and CoPc–Co–O as favorable catalysts.

3.4.3 ML-based descriptors. Advancements in machine learning have enabled the development of descriptors that go beyond relying only on physical intuition. Instead, these descriptors are generated through data-driven approaches by applying machine learning algorithms. These data-driven descriptors offer more straightforward alternatives as they do not require a deep understanding of the underlying physical/chemical processes. They automatically learn relationships from data to make predictions or classifications. Also, machine learning-based descriptors can incorporate all possible features in feature space, enabling direct descriptor construction from important features identified by the trained model. For example, Fung et al.[157] applied SISSO for 10 input features and a highly accurate expression for adsorption free energy $\Delta G[H^*]$ was formulated as $\Delta G[H^*] = -1.032 \times (e_d/q) + 13.424 \times (1/r_{cov}) + 1.726 \times (e_d \times EN) - 0.045 \times d_{occ}^2 - 9.241$, where EN is electronegativity, e_d is the d-states center, r_{cov} is the covalent radius, q is the Bader charge and d_{occ} is d electron occupancy. This linear regression model gives a training RMSE of 0.306 eV and a test error of 0.366 eV on N-doped graphene SACs. Similarly, Zhu et al.[155] utilized a gradient-boosting regression model to build a descriptor for ORR on a dual-metal site catalyst. According to their machine learning models, they proposed the free energy of OH adsorption can be predicted by $\Delta G[OH^*] = 0.69 + 0.00029\sqrt{EA_1} + 0.15EA_1 - 0.067\sqrt{EA_2} + 0.0076EA_2 - 0.012(R_1 + R_2) - 1.19|P_1 - P_2|^2$, where EA is

the electron affinity of metal 1 (EA$_1$) and metal 2 (EA$_2$); $(R_1 + R_2)$ is the summation of vDW radius of two metal centers; and $(P_1 - P_2)$ is the difference of Pauling electronegativity of two metal atoms. They also show that as more features are included in the descriptor expression, the RMSE can be improved from 0.395 to 0.283 for $\Delta G[^*OH]$ prediction. We summarized the descriptors for carbon-based catalysts under these three catagories in Table 1.

Based on the reported descriptors, electronegativity and the number of d electrons are recognized as the top important electronic features while bond length between metal and nearest neighbor atoms is reported as the most important geometric feature. So, the mentioned approaches provide consistent results. However, it is important to remember that each approach has its own strengths and weaknesses. For instance, identifying a physics-based descriptor can be challenging for systems that are not well-understood and there is no prior knowledge of the underlying physical principles. On the other hand, machine learning-based descriptors are limited by the quality and quantity of the training data, as well as the difficulty to interpret the underlying physical relationships and mechanisms. It is worth noting that the approaches can complement each other, leading to more accurate and reliable predictions in the physical chemistry field as well as improving interpretability. On the other hand, there are some disadvantages to combining the approaches, such as creating more complex models that are harder to understand and interpret.

4 Catalytic activity beyond binding energy

Electrocatalytic processes occur in aqueous solutions; hence, there are more complex interactions and other physical parameters involved in the actual reaction that have not been fully considered. Besides, in metal–carbon systems, the spin and

oxidation states are usually neglected. This section highlights the importance of accounting for both the solvent effect, which governs the interaction between the electrolyte and the electrocatalyst surface, and the spin state-induced extra electron transfer to improve the accuracy of binding energy predictions.

4.1 Solvation effects

There is direct interaction of solvent molecules with substrates, catalysts, products, intermediates, and transition states. These interactions can profoundly influence electrocatalysis; the solvent not only influences the solubility and transport of reactant molecules,[168] but also interacts with intermediates and catalytic sites;[169,170] thus the reaction mechanisms can differ from the gas phase simulations and include the solvent effects, which further influences the reaction rate and selectivity.[171–173] For example, the adsorption energy of a flat molecule on a clean Pt surface in the gas phase *vs.* solvent is different,[169] and different solvent coverage can alter the kinetics.[174] Those observations are also proved to be true in SACs. Wang *et al.* experimentally studied the solvent effect on N-doped graphene cobalt single-atom by altering the water/ethanol ratio.[175] After adding 30% water to the ethanol solvent, the activity increased from 43.2 to 76.8 h^{-1}, and the aniline selectivity increased from 62.8 to 99.1% for efficient hydrogenation of nitroarenes. Through further DFT study, the introducing of water weakened the binding between nitrobenzene and Co–N$_4$ which benefits the desorption of aniline from the surface. Cao *et al.* studied Mn–N$_4$/C SAC by using *ab initio* MD simulation and DFT calculations to investigate the influence of the solvation effect on the mechanisms of ORR by using explicit water molecules.[176] Fig. 5 illustrates the surface models in the gas phase and liquid phase used in this study, as well as the free-energy pathways. They found that the solvent environment promotes the charge transfer from the substrate to O$_2$, which leads to a dissociative pathway to rupture the O–O bond before protonation. In addition, water molecules could also increase the reactivity of the protonation step for *O and *OH by

elongating the Mn–O bond through a hydrogen bond.[177] As the last example,[178] Dub and Gordon considered metal–ligand bifunctional catalysts. They showed that when eight explicit water molecules are further added to the simulation, the first step (hydride transfer) proceeds similarly, but the relative activation barrier lowers by ~5 kcal mol^{-1}. They also found that in the second step of the reaction (proton transfer), the source of the proton is not C–H functionality, but the solvent. Their conclusions show that the inclusion of solvent effects changes both the reaction mechanism and the actual role of the ligand. Thus, to give an accurate reaction mechanisms, the roles of the solvent molecules must be considered in computational modeling.

Recent articles give an overview of various theoretical approaches that have been proposed to tackle solvation effects.[179–181] These approaches include the solvation effects within implicit (or continuum), hybrid implicit/explicit (cluster/continuum) or explicit approaches. In the following, we briefly introduce these three main ways to include the solvent effects in quantum mechanical calculations.

Nowadays, most of the DFT in catalysis consider the solvent by continuum models. This methodology has proven to be highly successful in treating the solvent effects. The simplest approximation of the solvent effect is the implicit solvent models. In implicit solvent models, also known as the continuum models, the solvent is treated as a homogeneous medium, described by the dielectric constant ε. In this way, the "*ad hoc*" solvation correction is directly added to the Gibbs free energy of adsorption calculated in vacuum.[94,182] The solvation-free energy can be calculated using various continuum solvation models, such as the polarizable continuum models (PCM)[183,184] and the conductor-like screening models (COSMO).[185,186] These methods are computationally efficient as they avoid sampling over solvent degree of freedom. However, it does not provide an accurate model when solvent molecules directly participate in the reaction. More examples of the success and limitations of continuum solvent models can be found in ref. 187.
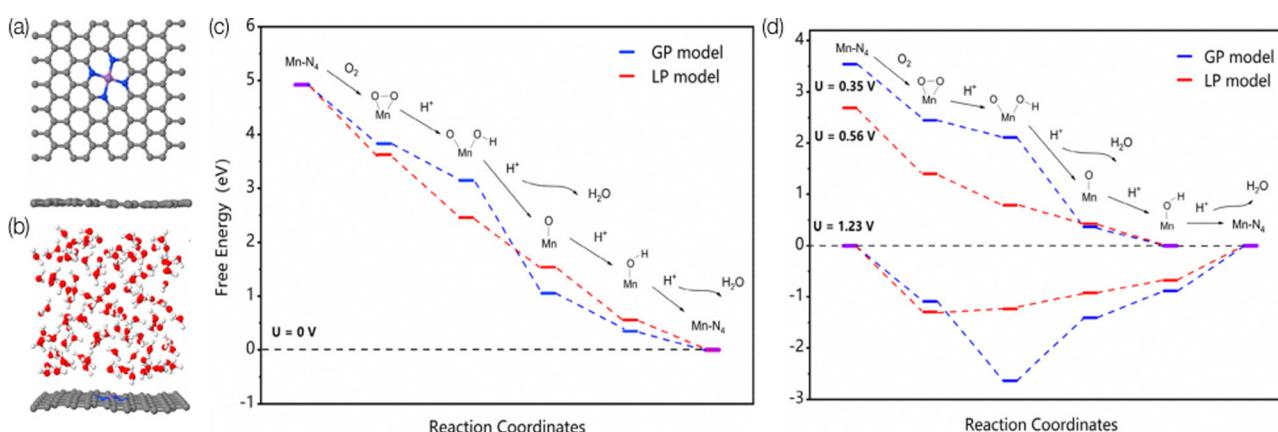
Fig. 5 The surface models of MnN$_4$−graphene in the (a) gas-phase (GP) and (b) liquid-phase (LP) model. The Mn, C, N, O, and H atoms are represented in purple, gray, blue, red, and white, respectively. Predicted free-energy pathways for ORR under different electrode potentials at pH = 0. (c) U = 0 V. (d) U = 1.23 V for both GP and LP models, U = 0.35 V for the GP model, and U = 0.56 V for the LP model. Reprinted with permission from ref. 176. Copyright 2020 American Chemical Society.

The interaction between the electrolyte and the charged surface of the electrode can also be included in the implicit solvation model, for example, based on either the Poisson–Boltzmann (PB) electrostatic or generalized Born (GB) model.[188–191] PB models take into account the solvent and solute charges, as well as the dielectric properties of the solvent and solute. They use the PB equation to calculate the electrostatic potential, which is then used to estimate the solvation free energy. GB models, on the other hand, approximate the model solvent as a continuous structureless medium with a uniform and certain dielectric constant to reduce the total degree of freedom of a system but ignore the potential electrostatic interaction with intermediates.[188–191] They use the Born equation to calculate the solvation free energy which is based on the surface area and volume of the solute, as well as the difference in dielectric constants between the solute and solvent. Both PB and GB models are based on simplifying assumptions and approximations, which can limit their accuracy; e.g., it is highly possible to underestimate the water–adsorbate interaction. Besides, in large-scale simulations, the solution of the complex mathematical equations is computationally expensive. These methods are computationally efficient as they avoid sampling over the solvent molecules. However, since the solvent effect is averaged, it limits the model's accuracy if the solvent molecules play a direct role in the reaction. The limitation can be overcome by including a selected number of explicit solvent molecules in the atomistic calculations which is the basis of the cluster-continuum approach.[192] The theoretical framework used for treating the explicit solvent molecules is the same as the rest of the catalytic system, i.e. the explicit solvent molecules are included in quantum mechanics simulations together with the catalysts. The implicit continuum solvation is then incorporated to account for long-range electrostatic effects. This combination results in hybrid implicit–explicit solvation schemes which have been discussed in detail by Pliego et al.[193] Although it is a cost-effective approach, it is difficult to define the number and the location of the explicit solvent molecules.

A realistic description of the solvent in which catalytic reactions occur is the full inclusion of explicit solvent molecules in the model as individual particles. Then, the computational models can simulate the dynamics of the system which could be done for instance by ab initio MD simulations and kinetic Monte Carlo simulations, which have been developed, improved, and employed by many groups.[194–197] During the simulation, each of the individual particle interacts with the solute molecule through intermolecular forces such as van der Waals and electrostatic interactions. Thus, the calculations are computationally expensive. So far, the widely accepted way to study the solvent effect on interactive interfaces is employing ab initio MD simulation to treat solvent molecules explicitly while maintaining the accuracy of interatomic force.[198] This method allows for thermodynamic and dynamic equilibrium under certain temperatures, with empirically applied force fields. Quantum mechanics/molecular mechanics (QM/MM) has also been proved to be an accurate and affordable strategy for studying electrocatalytic reactions.[199] The impact of the solvent effect on the energy profile can be clearly computed. However, most studies on electrolytes and catalyst surfaces have focused on metal and metal oxide surfaces, and the exploration of SACs still has not been fully understood.

4.2 Spin and charge transfer effects

In electrocatalysis, besides solvent effects, another critical issue that is often neglected in metal–carbon systems is the oxidation and spin states of the active site, which is usually the metal. In the following three paragraphs, we briefly summarize some literature searches that consider only oxidation states, spin states, and both.

Some studies focus solely on the role of oxidation states. For instance, experimental and theoretical investigations on Fe–N–C SACs for CO_2 RR reveal that Fe active sites maintaining +3 oxidation state during electrocatalysis exhibit superior activity for CO_2-to-CO conversion.[200,201] Operando X-ray absorption spectroscopy by Gu et al.[201] revealed that the active sites on Fe–N–C are Fe^{3+} ions, coordinated to pyrrolic nitrogen atoms of the N-doped carbon support, that maintains their +3 oxidation state during electrocatalysis, probably through electronic coupling to the conductive carbon support. Electrochemical data in this study suggest that the Fe^{3+} sites derive their superior activity from faster CO_2 adsorption and weaker CO absorption than that of conventional Fe^{2+} sites.

Detailed analysis of FeN_4 and FeN_5 in Fe–N–C reveals that the Bader charge of Fe in the FeN_5 system (+1.19 e) is higher than that of Fe in FeN_4 (+0.98 e), resulting in improving CO_2 RR on FeN_5 systems.[200] A recent study considered the dynamics of the oxidation state of the Fe SAC in Nafion coated functionalized multi-wall carbon nanotubes (Fe-n-f-CNTs). According to this experimental-theory study, the dynamics of the oxidation state may occur by structural deformation of the Fe–$(O)_3$ configuration, which modifies the coordination configuration due to the absorption of hydrogen and CO intermediates on probable active sites of Fe-n-f-CNTs.[202] This study also shows that, according to the Bader charge analysis obtained from DFT calculations for different configurations of Fe in the Fe SACs in Fe-n-f-CNTs, the Bader charge state of the Fe atom is consistent with experimental results.

Some studies consider only the role of different spin states and demonstrate how these states alter the reaction pathway of catalytic reactions.[104,203–206] For example, spectroscopic studies of Fe–N–C systems consider two distinct moieties, namely high-spin FeN_4C_{12} and low/intermediate-spin FeN_4C_{10}.[104] The results show that both sites initially contribute to the ORR activity but only the low/intermediate-spin moiety substantially contributes after 50 hours of operation. This experimental evidence motivated the examination of oxidation and spin states in computational catalysis. For example, DFT studies on the (salen)Mn(III)-catalyzed epoxidation reaction mechanism reveal that competing channels have different spin states.[203] DFT results on M–N_4–C (M = Mn, Fe, and Co) systems for ORR show that Mn–N_4 and Fe–N_4 centers in graphene exhibit the lowest O_2 dissociation energies across three spin channels,

while for Co–N$_4$ they find two spin channels require the same dissociation energy.[204] Additionally, a study on Fe–N–C systems for ORR indicates that the change of electronic spin moments of Fe and O$_2$ due to molecular-catalyst adsorption scales with the amount of electron transfer from Fe to O$_2$, as shown in Fig. 6, promoting the catalytic activity of C$_2$N–Fe for driving ORR.[206] Based on the relationship between catalytic activity and spin moment variation, they suggested that the electronic spin moment could be a promising catalytic descriptor for Fe SACs.

Some studies consider the effect of both oxidation and spin states in the catalysts.[207,208] For example, an experimental study on Fe–N–C for selective oxidation of C–H bonds provides evidence that among different FeNx (x = 4–6), the medium-spin Fe(III)N$_5$ affords the highest turnover frequency (6455 h^{-1}), and is at least 1 order of magnitude more active than the high-spin and low-spin Fe(III)N$_6$ structures and 3 times more active than the Fe(II)N$_4$ structure, although its relative concentration in the catalysts is much lower than that of the Fe(III)N$_6$ structures.[205] Yang et al.[207] investigated dual-metal atomically dispersed Fe, Mn/N–C catalysts and found that the O$_2$ reduction preferentially takes place on Fe(III) in the FeN$_4$/C system with the intermediate spin state which possesses one e$_g$ electron (t$_{2g}$4e$_g$1) readily penetrating the antibonding π-orbital of oxygen. Both magnetic measurements and the theoretical calculation reveal that the adjacent atomically dispersed Mn–N moieties can effectively activate the FeIII sites by both spin-state transition and electronic modulation, rendering the excellent ORR performances of Fe,Mn/N–C in both alkaline and acidic media. Gong et al.[208] developed a facile strategy to manipulate the cobalt spin state over COF-367-Co by changing the oxidation state of Co, resulting in the regulation of the photocatalytic performance for CO$_2$RR and enhanced selectivity to HCOOH. According to their DFT calculations and experimental results, Co(II) and Co(III) are embedded in COF-367 with $s = \frac{1}{2}$ and 0 spin ground states, respectively.

In electrocatalysis, the transfer of electrons and protons between the electrode, catalyst, and reactants is important. Proton transfer is often coupled with electron transfer in a process known as proton-coupled electron transfer (PCET), which can modulate the pathway that the reactants follow. Therefore, this coupled motion of electrons and protons plays a critical role in a wide range of electrochemical processes such as CO$_2$RR, ORR, and N$_2$RR in carbon and carbon–metal systems.[134,209–214] Fig. 7 provides a representative example demonstrating how electron transfer (ET), proton transfer (PT) and PCET affect the reaction mechanism during the four-electron ORR on nitridated carbon (NC) catalysts. Generally, PCET describes the mechanism of electron and proton transfer from atom to atom, which is also important for catalyst design. As described in Section 2.2, CHE predicts the thermodynamics properties of electrochemical reactions. Thus, PCET reactions can be evaluated using CHE which uses gas phase H$_2$ formed from H$^+$ and e$^-$ as a reference reaction. This intrinsically connects the PCET step to the standard hydrogen electrode (SHE) scale. To understand PCET at interfaces, where there are spatially inhomogeneous internal electric fields and electrostatic potentials within the active site, we need modeling strategies to explore PCET thermodynamics and further develop molecular active sites. Recent works have attempted to provide atomic-scale insight into interfacial PCET reactions, which are key steps in catalyst design. For instance, Hammes-Schiffer et al.[215] provide a first-principles modeling strategy based on DFT calculations. They illustrated that investigating the interfacial PCET at graphite-conjugated catalysts (GCC) bearing organic acid moieties explains the absence of a cyclic voltammetry peak. This combined theoretical-experimental study not only demonstrates the critical role of continuous conjugation and strong electronic coupling between the GCC acid sites and the graphite in enabling PCET at acid cites, but also emphasizes the importance of understanding the connection between the atomic structure of the surfaces and the
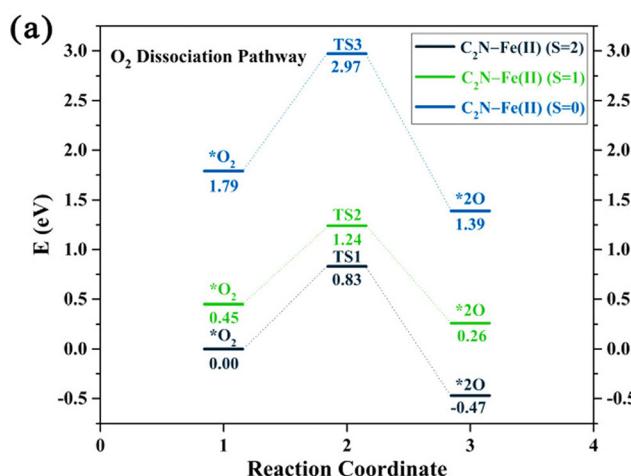

Fig. 6 The potential energy profiles for the O$_2$ dissociation pathway calculated for two systems: (a) C$_2$N–Fe(II) with spin states S = 2, 1, and 0, and (b) C$_2$N–Fe(III) with spin states S = 5/2, 3/2, and 1/2. The relative energies represented in the profiles correspond to Gibbs free energies, with the energies of the spin ground states set as a reference point at 0 eV. Reprinted with permission from ref. 206. Copyright 2021 American Chemical Society.
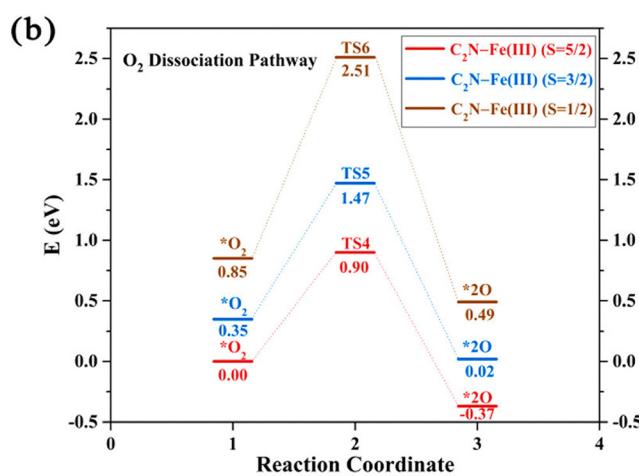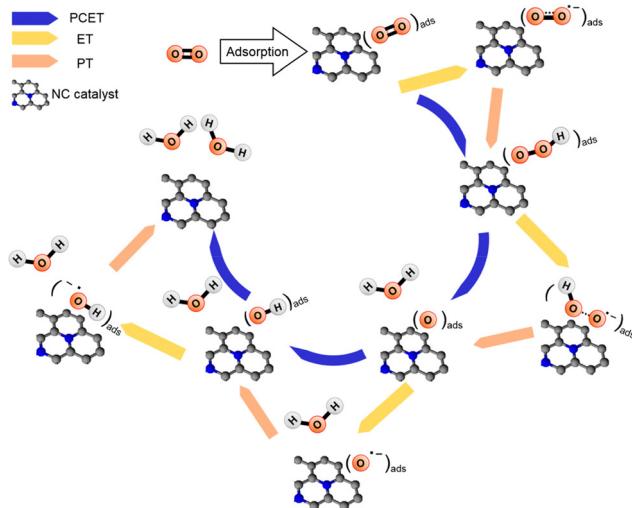
Fig. 7 Representative possible elementary steps during four-electron ORR ($O_2 + 4H^+ + 4e^- \rightarrow 2H_2O$) on carbon-based electrocatalysts. Black and blue balls represent carbon and nitrogen atoms, respectively. Reprinted with permission from ref. 210. Copyright 2018 American Chemical Society.

interfacial electrostatic potentials and fields that govern PCET thermochemistry. Furthermore, the study of these GCC catalysts in the presence of defects, such as heteroatom dopants in the graphitic surface, reveals that the lowest unoccupied electronic states can serve as a descriptor for changes in PCET thermodynamics.[216] Such studies highlight the need to consider PCET mechanisms in the computational screening of SACs.

The above-mentioned case studies provide evidence that oxidation states, spin states, and interfacial PCET are critical issues that cannot be ignored in catalyst design. However, providing computationally accurate results for such properties is challenging. For instance, it has been shown that the choice of DFT exchange–correlation functional can highly affect the binding energy of CO on TM atoms embedded on nitrogen-doped graphene.[217] Similarly, the absolute value of the spin magnetic moment of Fe, the spin magnetic moment of *O_2 ($M*O_2$), and the d-band center gap of spin states can also be affected.[218] For the (salen)Mn(III)-catalyzed epoxidation reaction mechanism, Abachkin et al. clearly showed how the choice of DFT functional and model can dramatically change the results.[203] Hence, these challenges can generally be addressed for specific systems with careful treatment of each effect. Nevertheless, the solution is still computationally expensive and not feasible for widely applying in large structural space search.

5 Perspectives and future opportunities

In Section 2.1, we have summarized the computational work on single atom and multiple metal sites in carbon-based materials, covering the exploration of coordination configurations, carbon

substrate structures, alternative dopants, and their influences on catalytic activity and stability. However, to find an optimal M–X_x (X = N, P, S, B, etc.) moiety that possesses high activity, selectivity and stability remains demanding and not straightforward, particularly in the exploration of moieties in different graphitic carbon structures to host M–X sites. Additionally, the curvature of the carbon substrates can also affect the properties of the electrocatalysts. Unfortunately, most modeled structures are primarily based on flat graphene substrates, with only a few studies demonstrating the importance of the curvature effects. For example, fullerenes which exhibit reactivity changes due to their curvatures[219] are rarely explored. In contrast, studies on N-doped fullerenes indicate the potential of such structures as cathode catalysts for hydrogen fuel cells,[220] CO oxidation,[221,222] and ORR.[223] In light of the challenges encountered in the exploration of optimal configurations for metal embedded in carbon-based catalysts, there are compelling opportunities to advance the field through the development of innovative computational tools and collaborative initiatives. One such opportunity lies in the creation of automated virtual high-throughput screening workflows that would be capable of swiftly exploring the huge design space. The workflows that can automatically generate diverse structures, efficiently perform DFT simulations, and subsequently conduct data-driven assessments would revolutionize the search for stable moieties that exhibit enhanced catalytic stability and selectivity.

Furthermore, accurately describing the reactivity through DFT simulations presents another significant challenge. A benchmark from the comparison with experimental data to assess the accuracy of different computational models is needed. Also, modeling the catalysts under operando conditions by considering the solvent effects, electrode potential, and structural evolution during reactions increases both the complexity of the model and the computational costs dramatically. As a consequence, attaining high accuracy in virtual high-throughput screening continues to be a persistent challenge. One potential solution is to train an accurate machine learning model to replace the computationally expensive DFT simulations. The machine learning models are expected to be trained based on the database that would be obtained from simulations with high accuracy. The establishment of such database plays a pivotal role in accelerating the research in the field. The database could serve as a repository with valuable information on structural, electronic and catalytic properties regarding the metal embedded in carbon-based materials.

Overall, computational approaches such as DFT simulations and data-driven models have been extensively leveraged in discovering sustainable, cost-effective, and highly reactive electrocatalysts derived from metal–carbon materials. However, there is still a need for efficient and predictive machine learning models. Training those models requires automated virtual high-throughput screening tools as well as accessible databases, which serve as crucial components in advancing the field. In conclusion, the continued advancement and integration of computational techniques, coupled with experimental efforts, hold tremendous promise for accelerating the

development and optimization of metal sites in carbon-based materials for electrocatalysis, propelling us closer to realizing efficient and scalable energy technologies.

Author contributions

Somayeh Faraji: literature curation, formal analysis, investigation, methodology, writing – original draft, writing – review and editing; Zhiyu Wang: literature curation, formal analysis, investigation, methodology, writing – original draft, writing – review and editing; Paola Lopez-Rivera: literature curation; Mingjie Liu: conceptualization, supervision, writing – review and editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 V. R. Stamenkovic, D. Strmcnik, P. P. Lopes and N. M. Markovic, *Nat. Mater.*, 2017, **16**, 57–69.
2 Y. Bing, H. Liu, L. Zhang, D. Ghosh and J. Zhang, *Chem. Soc. Rev.*, 2010, **39**, 2184–2202.
3 C. Li and J.-B. Baek, *ACS Omega*, 2019, **5**, 31–40.
4 Y. Hori, R. Takahashi, Y. Yoshinami and A. Murata, *J. Phys. Chem. B*, 1997, **101**, 7075–7081.
5 J. Resasco and A. T. Bell, *Trends Chem.*, 2020, **2**, 825–836.
6 A. J. Medford, A. Vojvodic, J. S. Hummelshøj, J. Voss, F. Abild-Pedersen, F. Studt, T. Bligaard, A. Nilsson and J. K. Nørskov, *J. Catal.*, 2015, **328**, 36–42.
7 H. Tao, C. Choi, L.-X. Ding, Z. Jiang, Z. Han, M. Jia, Q. Fan, Y. Gao, H. Wang and A. W. Robertson, *et al.*, *Chem*, 2019, **5**, 204–214.
8 G. Deng, T. Wang, A. A. Alshehri, K. A. Alzahrani, Y. Wang, H. Ye, Y. Luo and X. Sun, *J. Mater. Chem. A*, 2019, **7**, 21674–21677.
9 H. Huang, L. Xia, X. Shi, A. M. Asiri and X. Sun, *Chem. Commun.*, 2018, **54**, 11427–11430.
10 S. K. Kaiser, Z. Chen, D. Faust Akl, S. Mitchell and J. Perez-Ramirez, *Chem. Rev.*, 2020, **120**, 11703–11809.
11 Y. Wang, H. Su, Y. He, L. Li, S. Zhu, H. Shen, P. Xie, X. Fu, G. Zhou and C. Feng, *et al.*, *Chem. Rev.*, 2020, **120**, 12217–12314.
12 Y. Chen, L. Wang, Z. Yao, L. Hao, X. Tan, J. Masa, A. W. Robertson and Z. Sun, *Acta Phys.-Chim. Sin.*, 2022, **38**, 2206020.
13 D. Yu and J. Wang, *Chin. J. Struct. Chem.*, 2022, **41**, 2207002.
14 H. Fei, J. Dong, Y. Feng, C. S. Allen, C. Wan, B. Volosskiy, M. Li, Z. Zhao, Y. Wang and H. Sun, *et al.*, *Nat. Catal.*, 2018, **1**, 63–72.
15 X. Zeng, J. Shui, X. Liu, Q. Liu, Y. Li, J. Shang, L. Zheng and R. Yu, *Adv. Energy Mater.*, 2018, **8**, 1701345.
16 Y.-N. Chen, X. Zhang and Z. Zhou, *Small Methods*, 2019, **3**, 1900050.
17 Y. Yang, Y. Yang, Z. Pei, K.-H. Wu, C. Tan, H. Wang, L. Wei, A. Mahmood, C. Yan and J. Dong, *et al.*, *Matter*, 2020, **3**, 1442–1476.
18 Q. Zhang and J. Guan, *Adv. Funct. Mater.*, 2020, **30**, 2000768.
19 S. Back, J. Lim, N.-Y. Kim, Y.-H. Kim and Y. Jung, *Chem. Sci.*, 2017, **8**, 1090–1096.
20 T. Liu, G. Wang and X. Bao, *J. Phys. Chem. C*, 2021, **125**, 26013–26020.
21 K. Rossi and R. Buonsanti, *Acc. Chem. Res.*, 2022, **55**, 629–637.
22 D. R. Alfonso, D. N. Tafen and D. R. Kauffmann, *Catalysts*, 2018, **8**, 424.
23 L. Han, H. Cheng, W. Liu, H. Li, P. Ou, R. Lin, H.-T. Wang, C.-W. Pao, A. R. Head and C.-H. Wang, *et al.*, *Nat. Mater.*, 2022, **21**, 681–688.
24 K. Zhao, X. Nie, H. Wang, S. Chen, X. Quan, H. Yu, W. Choi, G. Zhang, B. Kim and J. G. Chen, *Nat. Commun.*, 2020, **11**, 1–10.
25 X. Li, W. Bi, M. Chen, Y. Sun, H. Ju, W. Yan, J. Zhu, X. Wu, W. Chu and C. Wu, *et al.*, *J. Am. Chem. Soc.*, 2017, **139**, 14889–14892.
26 Y. Yang, J. Li, C. Zhang, Z. Yang, P. Sun, S. Liu and Q. Cao, *J. Phys. Chem. C*, 2022, **126**, 4338–4346.
27 P. Brimley, H. Almajed, Y. Alsunni, A. W. Alherz, Z. J. Bare, W. A. Smith and C. B. Musgrave, *ACS Catal.*, 2022, **12**, 10161–10171.
28 W. Ju, A. Bagger, G.-P. Hao, A. S. Varela, I. Sinev, V. Bon, B. Roldan Cuenya, S. Kaskel, J. Rossmeisl and P. Strasser, *Nat. Commun.*, 2017, **8**, 1–9.
29 A. Kumar, S. Ibraheem, T. A. Nguyen, R. K. Gupta, T. Maiyalagan and G. Yasin, *Coord. Chem. Rev.*, 2021, **446**, 214122.
30 K. L. Svane, M. Reda, T. Vegge and H. A. Hansen, *ChemSusChem*, 2019, **12**, 5133–5141.
31 C. Yin, H. Tang, K. Li, Y. Yuan and Z. Wu, *New J. Chem.*, 2018, **42**, 9620–9625.
32 Z. Jiang, X. Zhang, H.-S. Chen, X. Hu and P. Yang, *ChemCatChem*, 2019, **11**, 4558–4567.
33 W. Shan and G. Wang, *J. Phys. Chem. C*, 2021, **125**, 16004–16012.
34 C. Choi, S. Back, N.-Y. Kim, J. Lim, Y.-H. Kim and Y. Jung, *ACS Catal.*, 2018, **8**, 7517–7525.
35 H. Yang, R. Shi, L. Shang and T. Zhang, *Small Struct.*, 2021, **2**, 2100007.
36 G. Gao, S. Bottle and A. Du, *Catal. Sci. Technol.*, 2018, **8**, 996–1001.
37 Y. Zhou, G. Gao, Y. Li, W. Chu and L.-W. Wang, *Phys. Chem. Chem. Phys.*, 2019, **21**, 3024–3032.

38 W. Yang, M. Zhao, X. Ding, K. Ma, C. Wu, I. D. Gates and Z. Gao, *Phys. Chem. Chem. Phys.*, 2020, **22**, 3983–3989.

39 X. Rong, H.-J. Wang, X.-L. Lu, R. Si and T.-B. Lu, *Angew. Chem.*, 2020, **132**, 1977–1981.

40 X. Wang, Z. Chen, X. Zhao, T. Yao, W. Chen, R. You, C. Zhao, G. Wu, J. Wang and W. Huang, *et al.*, *Angew. Chem.*, 2018, **130**, 1962–1966.

41 X.-F. Li, Q.-K. Li, J. Cheng, L. Liu, Q. Yan, Y. Wu, X.-H. Zhang, Z.-Y. Wang, Q. Qiu and Y. Luo, *J. Am. Chem. Soc.*, 2016, **138**, 8706–8709.

42 J. Liu, X. Kong, L. Zheng, X. Guo, X. Liu and J. Shui, *ACS Nano*, 2020, **14**, 1093–1101.

43 Y. Qin, Y. Li, W. Zhao, S. Chen, T. Wu and Y. Su, *Nano Res.*, 2023, **16**, 325–333.

44 F. Pan, B. Li, E. Sarnello, Y. Fei, Y. Gang, X. Xiang, Z. Du, P. Zhang, G. Wang and H. T. Nguyen, *et al.*, *ACS Nano*, 2020, **14**, 5506–5516.

45 A. V. Kuzmin and B. A. Shainyan, *Int. J. Quantum Chem.*, 2021, **121**, e26809.

46 G. Han, X. Zhang, W. Liu, Q. Zhang, Z. Wang, J. Cheng, T. Yao, L. Gu, C. Du and Y. Gao, *et al.*, *Nat. Commun.*, 2021, **12**, 6335.

47 S. Yang, C. Zhao, R. Qu, Y. Cheng, H. Liu and X. Huang, *RSC Adv.*, 2021, **11**, 3174–3182.

48 S. Yang, Y. Cheng, H. Liu and X. Huang, *Appl. Organomet. Chem.*, 2022, **36**, e6577.

49 A. Omidvar, *Synth. Met.*, 2017, **234**, 38–46.

50 X. Gao, Y. Zhou, Y. Tan, S. Liu, Z. Cheng and Z. Shen, *Appl. Surf. Sci.*, 2019, **492**, 8–15.

51 W. Song, K. Xie, J. Wang, Y. Guo, C. He and L. Fu, *Phys. Chem. Chem. Phys.*, 2021, **23**, 10418–10428.

52 Y. Xue, B. Huang, Y. Yi, Y. Guo, Z. Zuo, Y. Li, Z. Jia, H. Liu and Y. Li, *Nat. Commun.*, 2018, **9**, 1460.

53 Z. Feng, R. Li, Y. Ma, Y. Li, D. Wei, Y. Tang and X. Dai, *Phys. Chem. Chem. Phys.*, 2019, **21**, 19651–19659.

54 T. He, S. K. Matta, G. Will and A. Du, *Small Methods*, 2019, **3**, 1800419.

55 Z. Xue, X. Zhang, J. Qin and R. Liu, *J. Energy Chem.*, 2021, **55**, 437–443.

56 J.-L. Chen, H.-J. Hu and S.-H. Wei, *Chin. Phys. B*, 2022, **31**, 107306.

57 J. Huang, C. Zhou, Z. Chu, X. Liu and X. Duan, *Phys. Chem. Chem. Phys.*, 2021, **23**, 1868–1873.

58 B. He, J. Shen and Z. Tian, *Phys. Chem. Chem. Phys.*, 2016, **18**, 24261–24269.

59 X. Cui, W. An, X. Liu, H. Wang, Y. Men and J. Wang, *Nanoscale*, 2018, **10**, 15262–15272.

60 Y. Ying, K. Fan, X. Luo, J. Qiao and H. Huang, *J. Mater. Chem. A*, 2021, **9**, 16860–16867.

61 S. Liu and J.-y Liu, *J. Power Sources*, 2022, **535**, 231449.

62 H. Niu, Z. Zhang, X. Wang, X. Wan, C. Shao and Y. Guo, *Adv. Funct. Mater.*, 2021, **31**, 2008533.

63 B. Huang, Y. Wu, B. Chen, Y. Qian, N. Zhou and N. Li, *Chin. J. Catal.*, 2021, **42**, 1160–1167.

64 J.-J. Zhong, S.-P. Huang, J.-F. Gu, Y. Li, K.-N. Ding, Y.-F. Zhang, W. Lin and W.-K. Chen, *Appl. Surf. Sci.*, 2023, **609**, 155280.

65 C. Guo, T. Zhang, X. Deng, X. Liang, W. Guo, X. Lu and C.-M. L. Wu, *ChemSusChem*, 2019, **12**, 5126–5132.

66 C. Yang, Z. Zhao and Q. Liu, *Appl. Surf. Sci.*, 2021, **566**, 150697.

67 S.-L. Li, H. Yin, X. Kan, L.-Y. Gan, U. Schwingenschlögl and Y. Zhao, *Phys. Chem. Chem. Phys.*, 2017, **19**, 30069–30077.

68 H. Niu, X. Wan, X. Wang, C. Shao, J. Robertson, Z. Zhang and Y. Guo, *ACS Sustainable Chem. Eng.*, 2021, **9**, 3590–3599.

69 H. Zhao, H. Cao, Z. Zhang and Y.-G. Wang, *ACS Catal.*, 2022, **12**, 11380–11390.

70 X. Zhao and Y. Liu, *J. Am. Chem. Soc.*, 2020, **142**, 5773–5777.

71 Y. Wang, L. You and K. Zhou, *Chem. Sci.*, 2021, **12**, 14065–14073.

72 M. D. Hossain, Y. Huang, T. H. Yu, W. A. Goddard and Z. Luo, *Nat. Commun.*, 2020, **11**, 1–14.

73 Y. Pan, R. Lin, Y. Chen, S. Liu, W. Zhu, X. Cao, W. Chen, K. Wu, W.-C. Cheong and Y. Wang, *et al.*, *J. Am. Chem. Soc.*, 2018, **140**, 4218–4221.

74 Q. Fan, P. Hou, C. Choi, T.-S. Wu, S. Hong, F. Li, Y.-L. Soo, P. Kang, Y. Jung and Z. Sun, *Adv. Energy Mater.*, 2020, **10**, 1903068.

75 M. S. Iqbal, Z.-B. Yao, Y.-K. Ruan, R. Iftikhar, L.-D. Hao, A. W. Robertson, S. M. Imran and Z.-Y. Sun, *Rare Met.*, 2023, **42**, 1075–1097.

76 Y. Li, Z.-S. Wu, P. Lu, X. Wang, W. Liu, Z. Liu, J. Ma, W. Ren, Z. Jiang and X. Bao, *Adv. Sci.*, 2020, **7**, 1903089.

77 M. S. Ritopečki, A. S. Dobrota, N. V. Skorodumova and I. A. Pašti, *Nanomaterials*, 2022, **12**, 4309.

78 Y. Qin, S. Zhang, G. Gao, S. Ding and Y. Su, *Mol. Catal.*, 2022, **517**, 112048.

79 Z. Wen, H. Lv, D. Wu, W. Zhang, X. Wu and J. Yang, *J. Phys. Chem. Lett.*, 2022, **13**, 8177–8184.

80 X. Bai, L. Wang, B. Nan, T. Tang, X. Niu and J. Guan, *Nano Res.*, 2022, **15**, 6019–6025.

81 M. Guo, M. Ji and W. Cui, *Appl. Surf. Sci.*, 2022, **592**, 153237.

82 X.-Y. Xu, S. Lin, H. Xu, H. Guo and C. Zhao, *Appl. Surf. Sci.*, 2021, **556**, 149776.

83 Y. Yang, Z. Yang, C. Zhang, J. Zhou, S. Liu and Q. Cao, *Inorg. Chem.*, 2022, **61**, 12012–12022.

84 Y. Li, W. Shan, M. J. Zachman, M. Wang, S. Hwang, H. Tabassum, J. Yang, X. Yang, S. Karakalos and Z. Feng, *et al.*, *Angew. Chem., Int. Ed.*, 2022, **61**, e202205632.

85 X. Zheng, Y. Liu and Y. Yao, *Chem. Eng. J.*, 2021, **426**, 130745.

86 S. Li, A. Guan, C. Yang, C. Peng, X. Lv, Y. Ji, Y. Quan, Q. Wang, L. Zhang and G. Zheng, *ACS Mater. Lett.*, 2021, **3**, 1729–1737.

87 W. Wan, Y. Zhao, S. Wei, C. A. Triana, J. Li, A. Arcifa, C. S. Allen, R. Cao and G. R. Patzke, *Nat. Commun.*, 2021, **12**, 5589.

88 W. Ren, X. Tan, W. Yang, C. Jia, S. Xu, K. Wang, S. C. Smith and C. Zhao, *Angew. Chem., Int. Ed.*, 2019, **58**, 6972–6976.

89 J. Zhu, M. Xiao, D. Ren, R. Gao, X. Liu, Z. Zhang, D. Luo, W. Xing, D. Su and A. Yu, *et al.*, *J. Am. Chem. Soc.*, 2022, **144**, 9661–9671.

90 J. Wang, Z. Huang, W. Liu, C. Chang, H. Tang, Z. Li, W. Chen, C. Jia, T. Yao and S. Wei, et al., J. Am. Chem. Soc., 2017, 139, 17281–17284.

91 L. Li, K. Yuan and Y. Chen, Acc. Mater. Res., 2022, 3, 584–596.

92 Z. Ao and F. Peeters, J. Phys. Chem. C, 2010, 114, 14503–14509.

93 J. K. Nørskov, T. Bligaard, J. Rossmeisl and C. H. Christensen, Nat. Chem., 2009, 1, 37–46.

94 J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard and H. Jonsson, J. Phys. Chem. B, 2004, 108, 17886–17892.

95 J. D. Goodpaster, A. T. Bell and M. Head-Gordon, J. Phys. Chem. Lett., 2016, 7, 1471–1477.

96 R. Sundararaman, W. A. Goddard and T. A. Arias, J. Chem. Phys., 2017, 146, 114104.

97 P. Atkins, P. W. Atkins and J. de Paula, Atkins' physical chemistry, Oxford University Press, 2014.

98 E. Skúlason, V. Tripkovic, M. E. Björketun, S. Gudmundsdóttir, G. Karlberg, J. Rossmeisl, T. Bligaard, H. Jónsson and J. K. Nørskov, J. Phys. Chem. C, 2010, 114, 18182–18197.

99 J. Rossmeisl, E. Skúlason, M. E. Björketun, V. Tripkovic and J. K. Nørskov, Chem. Phys. Lett., 2008, 466, 68–71.

100 C. Guo, T. Zhang, X. Liang, X. Deng, W. Guo, Z. Wang, X. Lu and C.-M. L. Wu, Appl. Surf. Sci., 2020, 533, 147466.

101 G. Gao and L.-W. Wang, J. Catal., 2020, 391, 530–538.

102 H. Cao, Z. Zhang, J.-W. Chen and Y.-G. Wang, ACS Catal., 2022, 12, 6606–6617.

103 B. Qiao, J.-X. Liang, A. Wang, C.-Q. Xu, J. Li, T. Zhang and J. J. Liu, Nano Res., 2015, 8, 2913–2924.

104 J. Li, M. T. Sougrati, A. Zitolo, J. M. Ablett, I. C. Oğuz, T. Mineva, I. Matanovic, P. Atanassov, Y. Huang and I. Zenyuk, et al., Nat. Catal., 2021, 4, 10–19.

105 D. Karapinar, N. T. Huan, N. Ranjbar Sahraie, J. Li, D. Wakerley, N. Touati, S. Zanna, D. Taverna, L. H. Galvão Tizei and A. Zitolo, et al., Angew. Chem., Int. Ed., 2019, 58, 15098–15103.

106 C. E. Creissen and M. Fontecave, Nat. Commun., 2022, 13, 2280.

107 H. Xu, D. Rebollar, H. He, L. Chong, Y. Liu, C. Liu, C.-J. Sun, T. Li, J. V. Muntean and R. E. Winans, et al., Nat. Energy, 2020, 5, 623–632.

108 M. Chen, M. Luo, C. Liu, X. Qi, S. G. Peera and T. Liang, Comput. Theor. Chem., 2020, 1187, 112945.

109 Z. Lu, G. Xu, C. He, T. Wang, L. Yang, Z. Yang and D. Ma, Carbon, 2015, 84, 500–508.

110 X. Guo, S. Lin, J. Gu, S. Zhang, Z. Chen and S. Huang, ACS Catal., 2019, 9, 11042–11054.

111 L. Xu, L.-M. Yang and E. Ganz, Theor. Chem. Acc., 2018, 137, 1–13.

112 J. Greeley and J. K. Nørskov, Electrochim. Acta, 2007, 52, 5829–5836.

113 M. Ha, D. Y. Kim, M. Umer, V. Gladkikh, C. W. Myung and K. S. Kim, Energy Environ. Sci., 2021, 14, 3455–3468.

114 J. Wang, M. Zheng, X. Zhao and W. Fan, ACS Catal., 2022, 12, 5441–5454.

115 Z. Huang and Q. Tang, J. Phys. Chem. C, 2022, 126, 21606–21615.

116 M. Pourbaix, Atlas of electrochemical equilibria in aqueous solutions, 1966.

117 L. Yang, H. Xu, H. Liu, X. Zeng, D. Cheng, Y. Huang, L. Zheng, R. Cao and D. Cao, Research, 2020, 2020, 7593023.

118 J. H. Zagal and M. T. Koper, Angew. Chem., Int. Ed., 2016, 55, 14510–14521.

119 A. S. Dobrota, N. V. Skorodumova, S. V. Mentus and I. A. Pašti, Electrochim. Acta, 2022, 412, 140155.

120 E. F. Holby, G. Wang and P. Zelenay, ACS Catal., 2020, 10, 14527–14539.

121 X. Zhao and Y. Liu, J. Am. Chem. Soc., 2021, 143, 9423–9428.

122 F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skulason, T. Bligaard and J. K. Nørskov, Phys. Rev. Lett., 2007, 99, 016105.

123 J. Greeley, Annu. Rev. Chem. Biomol. Eng., 2016, 7, 605–635.

124 Z.-J. Zhao, S. Liu, S. Zha, D. Cheng, F. Studt, G. Henkelman and J. Gong, Nat. Rev. Mater., 2019, 4, 792–804.

125 J. Zhang, H. B. Yang, D. Zhou and B. Liu, Chem. Rev., 2022, 122, 17028–17072.

126 G. Lendvay, J. Phys. Chem., 1989, 93, 4422–4429.

127 E. Shustorovich, Surf. Sci. Rep., 1986, 6, 1–63.

128 E. Shustorovich, J. Am. Chem. Soc., 1984, 106, 6479–6481.

129 M. Ma, F. Li and Q. Tang, Nanoscale, 2021, 13, 19133–19143.

130 C. Choi, S. Yoon and Y. Jung, Chem. Sci., 2021, 12, 3551–3557.

131 M. Evans and M. Polanyi, Trans. Faraday Soc., 1938, 34, 11–24.

132 A. Michaelides, Z.-P. Liu, C. Zhang, A. Alavi, D. A. King and P. Hu, J. Am. Chem. Soc., 2003, 125, 3704–3705.

133 J. Cheng, P. Hu, P. Ellis, S. French, G. Kelly and C. M. Lok, J. Phys. Chem. C, 2008, 112, 1308–1311.

134 L. Zhong, L. Zhang and S. Li, ACS Mater. Lett., 2020, 3, 110–120.

135 J. Tymoczko, F. Calle-Vallejo, W. Schuhmann and A. S. Bandarenka, Nat. Commun., 2016, 7, 1–6.

136 Y. Xiao and W. Zhang, Electrocatalysis, 2020, 11, 393–404.

137 L. Zhong and S. Li, ACS Catal., 2020, 10, 4313–4318.

138 S. Wang, L. Li, J. Li, C. Yuan, Y. Kang, K. S. Hui, J. Zhang, F. Bin, X. Fan and F. Chen, et al., J. Phys. Chem. C, 2021, 125, 7155–7165.

139 J. Long, X. Fu and J. Xiao, J. Mater. Chem. A, 2020, 8, 17078–17088.

140 Y. Guo, G. Wang, S. Shen, G. Wei, G. Xia and J. Zhang, Appl. Surf. Sci., 2021, 550, 149283.

141 X. Liu, Y. Jiao, Y. Zheng, M. Jaroniec and S.-Z. Qiao, J. Am. Chem. Soc., 2019, 141, 9664–9672.

142 Z. Wang, B. Yang, Y. Wang, Y. Zhao, X.-M. Cao and P. Hu, Phys. Chem. Chem. Phys., 2013, 15, 9498–9502.

143 S. Vijay, G. Kastlunger, K. Chan and J. K. Nørskov, J. Chem. Phys., 2022, 156, 231102.

144 C. Li, X. Liu, F. Xu, D. Wu, H. Xu and G. Fan, Electrochim. Acta, 2022, 426, 140764.

145 P. Ma, C. Feng, Y. Kong, D. Wang, M. Zuo, S. Wang, R. Wang, L. Kuang, X. Ding and S. Zhou, et al., Chem Catal., 2022, 2, 2764–2777.

146 R. Hu, Y. Yu, Y. Li, Y. Wang, J. Shang, Y. Nie and X. Jiang, *Nano Res.*, 2022, **15**, 8656–8664.

147 Y. Wang, R. Hu, Y. Li, F. Wang, J. Shang and J. Shui, *Nano Res.*, 2022, **15**, 1054–1060.

148 P. Sun, Z. Qiao, S. Wang, D. Li, X. Liu, Q. Zhang, L. Zheng, Z. Zhuang and D. Cao, *Angew. Chem.*, 2023, **135**, e202216041.

149 Z. Xue, X. Zhang, J. Qin and R. Liu, *Nano Energy*, 2021, **80**, 105527.

150 C. Ling, Y. Ouyang, Q. Li, X. Bai, X. Mao, A. Du and J. Wang, *Small Methods*, 2019, **3**, 1800376.

151 M. Wang, T. Wang, P. Cai and X. Chen, *Small Methods*, 2019, **3**, 1900025.

152 M. Zafari, D. Kumar, M. Umer and K. S. Kim, *J. Mater. Chem. A*, 2020, **8**, 5209–5216.

153 C. Deng, Y. Su, F. Li, W. Shen, Z. Chen and Q. Tang, *J. Mater. Chem. A*, 2020, **8**, 24563–24571.

154 S. Kapse, N. Barman and R. Thapa, *Carbon*, 2023, **201**, 703–711.

155 X. Zhu, J. Yan, M. Gu, T. Liu, Y. Dai, Y. Gu and Y. Li, *J. Phys. Chem. Lett.*, 2019, **10**, 7760–7766.

156 X. Wan, Z. Zhang, H. Niu, Y. Yin, C. Kuai, J. Wang, C. Shao and Y. Guo, *J. Phys. Chem. Lett.*, 2021, **12**, 6111–6118.

157 V. Fung, G. Hu, Z. Wu and D.-E. Jiang, *J. Phys. Chem. C*, 2020, **124**, 19571–19578.

158 S. Zhang, S. Lu, P. Zhang, J. Tian, L. Shi, C. Ling, Q. Zhou and J. Wang, *Energy Environ. Mater.*, 2023, **6**, e12304.

159 M. Umer, S. Umer, M. Zafari, M. Ha, R. Anand, A. Hajibabaei, A. Abbas, G. Lee and K. S. Kim, *J. Mater. Chem. A*, 2022, **10**, 6679–6689.

160 D. Newns, *Phys. Rev.*, 1969, **178**, 1123.

161 S. Jiao, X. Fu and H. Huang, *Adv. Funct. Mater.*, 2022, **32**, 2107651.

162 G. Jones, T. Bligaard, F. Abild-Pedersen and J. K. Nørskov, *J. Phys.: Condens. Matter*, 2008, **20**, 064239.

163 H. Niu, X. Wang, C. Shao, Y. Liu, Z. Zhang and Y. Guo, *J. Mater. Chem. A*, 2020, **8**, 6555–6563.

164 W. Gao, Y. Chen, B. Li, S.-P. Liu, X. Liu and Q. Jiang, *Nat. Commun.*, 2020, **11**, 1196.

165 K. Yang, J. Zaffran and B. Yang, *Phys. Chem. Chem. Phys.*, 2020, **22**, 890–895.

166 H. Xu, D. Cheng, D. Cao and X. C. Zeng, *Nat. Catal.*, 2018, **1**, 339–348.

167 Q. Zhang, P.-A. Hu, Z.-Y. Xu, B.-B. Tang, H.-R. Zhang, Y.-H. Xiao and Y.-C. Wu, *Nanoscale*, 2023, **15**, 4991–5000.

168 U. K. Singh and M. A. Vannice, *Appl. Catal., A*, 2001, **213**, 1–24.

169 N. Singh and C. T. Campbell, *ACS Catal.*, 2019, **9**, 8116–8127.

170 S. Liu, M. G. White and P. Liu, *J. Phys. Chem. C*, 2016, **120**, 15288–15298.

171 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.

172 P. Suppan, *J. Chem. Soc. A*, 1968, 3125–3133.

173 F. Calle-Vallejo, A. Krabbe and J. M. Garca-Lastra, *Chem. Sci.*, 2017, **8**, 124–130.

174 S. Mukherjee and M. A. Vannice, *J. Catal.*, 2006, **243**, 131–148.

175 H. Wang, Y. Wang, Y. Li, X. Lan, B. Ali and T. Wang, *ACS Appl. Mater. Interfaces*, 2020, **12**, 34021–34031.

176 H. Cao, G.-J. Xia, J.-W. Chen, H.-M. Yan, Z. Huang and Y.-G. Wang, *J. Phys. Chem. C*, 2020, **124**, 7287–7294.

177 W. Zhang, F.-L. Sun, Q.-J. Fang, Y.-F. Yu, J.-K. Pan, J.-G. Wang and G.-L. Zhuang, *J. Phys. Chem. C*, 2022, **126**, 2349–2364.

178 P. A. Dub and J. C. Gordon, *ACS Catal.*, 2017, **7**, 6635–6655.

179 Q. Li, Y. Ouyang, S. Lu, X. Bai, Y. Zhang, L. Shi, C. Ling and J. Wang, *Chem. Commun.*, 2020, **56**, 9937–9949.

180 Y. Basdogan, A. M. Maldonado and J. A. Keith, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2020, **10**, e1446.

181 J. J. Varghese and S. H. Mushrif, *React. Chem. Eng.*, 2019, **4**, 165–206.

182 Z.-D. He, S. Hanselman, Y.-X. Chen, M. T. Koper and F. Calle-Vallejo, *Importance of solvation for the accurate prediction of oxygen reduction activities of Pt-based electrocatalysts*, 2017.

183 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117–129.

184 D. M. Chipman, *Theor. Chem. Acc.*, 2002, **107**, 80–89.

185 A. Klamt, *J. Phys. Chem.*, 1995, **99**, 2224–2235.

186 A. Klamt, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2018, **8**, e1338.

187 G. Norjmaa, G. Ujaque and A. Lledós, *Top. Catal.*, 2022, **65**, 118–140.

188 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Phys. Chem. B*, 2009, **113**, 6378–6396.

189 K. Mathew, R. Sundararaman, K. Letchworth-Weaver, T. Arias and R. G. Hennig, *J. Chem. Phys.*, 2014, **140**, 084106.

190 M. Fishman, H. L. Zhuang, K. Mathew, W. Dirschka and R. G. Hennig, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2013, **87**, 245402.

191 G. Scalmani and M. J. Frisch, *J. Chem. Phys.*, 2010, **132**, 114110.

192 R. B. Sunoj and M. Anand, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12715–12736.

193 J. R. Pliego Jr and J. M. Riveros, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2020, **10**, e1440.

194 J. A. Gauthier, C. F. Dickens, H. H. Heenen, S. Vijay, S. Ringe and K. Chan, *J. Chem. Theory Comput.*, 2019, **15**, 6895–6906.

195 J. Zhang, H. Zhang, T. Wu, Q. Wang and D. van der Spoel, *J. Chem. Theory Comput.*, 2017, **13**, 1034–1043.

196 M. M. Montemore, O. Andreussi and J. W. Medlin, *J. Chem. Phys.*, 2016, **145**, 074702.

197 S. Schnur and A. Groß, *New J. Phys.*, 2009, **11**, 125003.

198 R. Iftimie, P. Minary and M. E. Tuckerman, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6654–6659.

199 Z. Wang and E. H. Sargent, *Matter*, 2021, **4**, 12–14.

200 H. Zhang, J. Li, S. Xi, Y. Du, X. Hai, J. Wang, H. Xu, G. Wu, J. Zhang and J. Lu, *et al.*, *Angew. Chem.*, 2019, **131**, 15013–15018.

201 J. Gu, C.-S. Hsu, L. Bai, H. M. Chen and X. Hu, *Science*, 2019, **364**, 1091–1094.

202 K. Lakshmanan, W.-H. Huang, S. A. Chala, B. W. Taklu, E. A. Moges, J.-F. Lee, P.-Y. Huang, Y.-C. Lee, M.-C. Tsai and W.-N. Su, *et al.*, *Adv. Funct. Mater.*, 2022, **32**, 2109310.

203 Y. G. Abashkin, J. R. Collins and S. K. Burt, *Inorg. Chem.*, 2001, **40**, 4040–4048.

204 W. Orellana, *J. Phys. Chem. C*, 2013, **117**, 9812–9818.

205 W. Liu, L. Zhang, X. Liu, X. Liu, X. Yang, S. Miao, W. Wang, A. Wang and T. Zhang, *J. Am. Chem. Soc.*, 2017, **139**, 10790–10798.

206 W. Zhong, Y. Qiu, H. Shen, X. Wang, J. Yuan, C. Jia, S. Bi and J. Jiang, *J. Am. Chem. Soc.*, 2021, **143**, 4405–4413.

207 G. Yang, J. Zhu, P. Yuan, Y. Hu, G. Qu, B.-A. Lu, X. Xue, H. Yin, W. Cheng and J. Cheng, *et al.*, *Nat. Commun.*, 2021, **12**, 1734.

208 Y.-N. Gong, W. Zhong, Y. Li, Y. Qiu, L. Zheng, J. Jiang and H.-L. Jiang, *J. Am. Chem. Soc.*, 2020, **142**, 16723–16731.

209 A. M. Verma, K. Honkala and M. M. Melander, *Front. Energy Res.*, 2021, **388**, 606742.

210 K. Sakaushi, M. Eckardt, A. Lyalin, T. Taketsugu, R. J. Behm and K. Uosaki, *ACS Catal.*, 2018, **8**, 8162–8176.

211 Y. Li, Z. Qiao, Y. Cao, H. Wang, H. Liang, H. Yu and F. Peng, *ChemSusChem*, 2019, **12**, 1133–1138.

212 K. Sun, W. Xu, X. Lin, S. Tian, W.-F. Lin, D. Zhou and X. Sun, *Adv. Mater. Interfaces*, 2021, **8**, 2001360.

213 Y. Li, S. Zhang, J. Yu, Q. Wang, Q. Sun and P. Jena, *Nano Res.*, 2015, **8**, 2901–2912.

214 J. C. Peters, *Faraday Discuss.*, 2023, **243**, 450–472.

215 R. E. Warburton, P. Hutchison, M. N. Jackson, M. L. Pegis, Y. Surendranath and S. Hammes-Schiffer, *J. Am. Chem. Soc.*, 2020, **142**, 20855–20864.

216 P. Hutchison, R. E. Warburton, Y. Surendranath and S. Hammes-Schiffer, *J. Phys. Chem. Lett.*, 2022, **13**, 11216–11222.

217 Q. Wu, G. Wang and M. Liu, *ChemPhysChem*, 2022, **23**, e202100787.

218 K. Liu, J. Fu, Y. Lin, T. Luo, G. Ni, H. Li, Z. Lin and M. Liu, *Nat. Commun.*, 2022, **13**, 2075.

219 T. Lin, W.-D. Zhang, J. Huang and C. He, *J. Phys. Chem. B*, 2005, **109**, 13755–13760.

220 F. Gao, G.-L. Zhao, S. Yang and J. J. Spivey, *J. Am. Chem. Soc.*, 2013, **135**, 3315–3318.

221 I.-H. Lin, Y.-H. Lu and H.-T. Chen, *J. Comput. Chem.*, 2017, **38**, 2041–2046.

222 E. Shakerzadeh, H. Hamadi and M. D. Esrafili, *Inorg. Chem. Commun.*, 2019, **106**, 190–196.

223 B. Modak, K. Srinivasu and S. K. Ghosh, *Int. J. Hydrogen Energy*, 2017, **42**, 2278–2287.