

ARTICLE OPEN



Computational high-throughput screening of alloy nanoclusters for electrocatalytic hydrogen evolution

Xinnan Mao¹, Lu Wang¹✉, Yafeng Xu¹, Pengju Wang², Youyong Li^{1,3}✉ and Jijun Zhao²✉

Here, we report a density functional theory (DFT)-based high-throughput screening method to successfully identify a type of alloy nanoclusters as the electrocatalyst for hydrogen evolution reaction (HER). Totally 7924 candidates of Cu-based alloy clusters of $\text{Cu}_{55-n}\text{M}_n$ ($\text{M} = \text{Co}, \text{Ni}, \text{Ru}, \text{and Rh}$) are optimized and evaluated to screening for the promising catalysts. By comparing different structural patterns, Cu-based alloy clusters prefer the core-shell structures with the dopant metal in the core and Cu as the shell atoms. Generally speaking, the HER performance of the Cu-based nanoclusters can be significantly improved by doping transition metals, and the active sites are the bridge sites and three-fold sites on the outer-shell Cu atoms. Considering the structural stability and the electrochemical activity, core-shell CuNi alloy clusters are suggested to be the superior electrocatalyst for hydrogen evolution. A descriptor composing of surface charge is proposed to efficiently evaluate the HER activity of the alloy clusters supported by the DFT calculations and machine-learning techniques. Our screening strategy could accelerate the pace of discovery for promising HER electrocatalysts using metal alloy nanoclusters.

npj Computational Materials (2021)7:46; <https://doi.org/10.1038/s41524-021-00514-8>

INTRODUCTION

Energy and environmental issues raise imperative demands on the sustainable development of energy. At present, a promising strategy is to produce renewable energy through electrochemically catalytic reactions, which convert some common materials, such as water, carbon dioxide, and nitrogen, into energy carriers. The electrocatalytic hydrogen evolution is one of the most important ways to achieve this goal, which is a critical reaction for a variety of electrochemical processes and has been applied in diverse applications such as hydrogen fuel cells, electrodeposition, corrosion of metals in acids, and energy storage using H_2 medium. For the hydrogen evolution reaction (HER), choosing an appropriate electrocatalyst is most crucial to achieve high efficiency. Precious metals such as platinum are recognized as the best catalysts for electrochemical HER with intrinsically good activity but high cost^{1–4}. Compared to conventional bulk electrocatalysts, nanomaterials could exhibit even more superior performance as benefited from the rapid development of nanoscience and nanotechnology^{5–7}.

Nanoclusters, consisting of a few tens to a few thousands of atoms, exhibit the distinctive optical, electronic, magnetic, and chemical properties definitely different from the bulk materials^{8,9}. Metal clusters have received long-lasting attention owing to their potential applications in catalysis^{10–12}. As known, the catalytic reactions always occur on the surface of catalysts, thus a significant proportion of atoms in the bulk materials cannot be effectively utilized in the catalytic process. The nanoclusters possess compelling advantages like high surface-to-volume ratio and more surface-exposed atoms. In addition, it is possible to increase the activity and selectivity in a catalytic process by carefully controlling the size and structure of metal nanoclusters⁵. As an alternative to the precious metals like Pt and Pd with high expenses and low resources, the alloy nanoclusters prepared by

mixing two or more metal components have been successfully synthesized and display better catalytic activity due to the synergistic effects^{9,13–15}, which could replace or reduce the loading of the precious metals as the efficient electrocatalysts. The formation of core-shell or intermixed nanostructures results in superior catalytic properties with regard to one of the components, such as NiPt, CoPt, CuPd nanoparticles.

For the equilibrium structures of bimetallic alloy nanoclusters, there are four types of geometric patterns^{9,16}. The first one is core-shell structure, where a core of one type of metal atoms is surrounded by a shell of the other kind of metal atoms. The second one is a segregated structure, where two types of metal atoms are separated by distinct boundary. The third one is homogeneously mixed structure with the two metal atoms are distributed in an atomically ordered or a statistically random manner. The fourth one is “onion-like” structure with more than one concentric shell covering the metal core. Among them, the core-shell alloy clusters have received increasing attention since their catalytic activities could be efficiently tuned by controlling the compositions of core and shell atoms.

In order to develop low-cost and environmentally benign HER catalysts to substitute the precious metals, here we performed high-throughput density functional theory (DFT) computations of Cu-based alloy nanoclusters and screened for the promising HER electrocatalysts. Among a total number of 7924 candidate structures, the most stable structures for $\text{Cu}_{55-n}\text{M}_n$ ($\text{M} = \text{Co}, \text{Ni}, \text{Ru}$ and Rh , and $n \leq 22$) nanoclusters adopt the core-shell configurations with Cu as the shell atoms. To avoid the heavy computations, a descriptor has been established to evaluate the HER activity on the nanoclusters, and by training a machine-learning neutral network with a large DFT database, we have achieved a quick and accurate prediction of the hydrogen adsorption free energies on the nanoclusters surface.

¹Institute of Functional Nano & Soft Materials (FUNSOM), Jiangsu Key Laboratory for Carbon-Based Functional Materials & Devices, Soochow University, Suzhou, Jiangsu 215123, China. ²Key Laboratory of Materials Modification by Laser, Ion and Electron Beams (Dalian University of Technology), Ministry of Education, Dalian 116024, China. ³Macao Institute of Materials Science and Engineering, Macau University of Science and Technology, Taipa 999078 Macau SAR, China. ✉email: lwang22@suda.edu.cn; yyli@suda.edu.cn; zhaojj@dlut.edu.cn

RESULTS AND DISCUSSION

As an earth-abundant and inexpensive metal, copper has been widely used in some electrocatalytic reactions. Previously, Nørskov and co-workers¹⁷ have calculated the free energies of hydrogen adsorption on a variety of metal alloy surfaces and evaluated their HER activities. Taking Cu as the host metal, four transition metal dopants (i.e., Co, Ni, Ru, and Rh) in the Cu surface are shown to exhibit superior hydrogen evolution activities. Thus, we choose a medium-sized Cu₅₅ nanocluster (with a diameter of about 1 nm) as the initial structure, and randomly substitute the transition metals for Cu atoms to form binary alloy nanoclusters and screening for the high-efficient HER catalysts. The high-throughput computational workflow is shown in Fig. 1. First, given an initial cluster structure and the doping concentration, the possible doped cluster structures are generated and optimized by DFT calculations. After the stable cluster structures are obtained, the surface-active sites are enumerated and then considered for the H adsorption. Together with the contributions of zero-point energy and entropy, the H adsorption free energies are calculated to identify the optimal active sites and search for the promising HER catalysts.

Geometric structures

For pure Cu₅₅ cluster, the lowest-energy structure is found to be a three-layer icosahedron with I_h symmetry¹⁸. To achieve the low-cost catalysts, it is preferred to contain less precious metals in the catalysts. As the host metal of Cu is remarkably less expensive, the dopant noble metal is the less the better. Here we have considered the dopant concentration is less than half. Firstly, three ratios of the Cu-based alloy nanoclusters of Cu_{55-n}M_n are considered to explore their structures and growth patterns, where M represents Co, Ni, Ru, and Rh, and $n = 7, 13, \text{ and } 16$. By eliminating duplicate individuals using the nearest neighbor matrix method¹⁹, we have obtained totally 6020 configurations for these 12 alloy clusters for further geometric optimization.

As mentioned earlier, there are four kinds of structures for the bimetallic alloy clusters. In order to describe the geometric distributions of Cu and dopant metal atoms, we define the average radial distance ($D_{\text{radial}, M}$) between the transition metal atoms ($M = \text{Co, Ni, Ru, and Rh}$) and the center atom by the following formula:

$$D_{\text{radial}, M} = \sum_{i=1}^n \frac{|r_i - r_c|}{n} \quad (1)$$

where $|r_i - r_c|$ indicates the radial distance from the i^{th} metal atom to the center atom, and n is the number of transition metal dopant atoms. In the case of the icosahedral Cu_{55-n}M_n clusters, as the $D_{\text{radial}, M}$ value becomes smaller, the dopant atoms distribute closer to the core region, and vice versa. The excess energies (E_{exc}^*) for Cu_{55-n}M_n alloy clusters as a function of the average radial distance of the dopant metal atoms to the core atom are plotted in Fig. 2. These four plots for different transition metal dopants display nearly identical trends. By definition, the lower excess energy indicates a more stable cluster structure. The structures in the left lower quadrant possess relatively lower excess energies, which are the core-shell structures with the core of dopant metal atoms surrounded by the shell of Cu atoms, independent of the doping concentration.

We take Cu₄₂M₁₃ as the representative to illustrate the detailed structural patterns. For the icosahedral structure, the central atom could be either Cu or M. We have systematically compared these two types of core-shell structures together with the other possible structural patterns (homogenous and segregated structures). The excess energies as a function of the average radial distance of the M atoms to the central atom are summarized in Fig. 3. A linear relationship exists between excess energies and $D_{\text{radial}, M}$ for Cu₄₂M₁₃ clusters, and the difference of the central atom (Cu or M) causes a slightly different slope. Among four types of Cu-based alloy clusters, Cu₄₃Ni₁₃ clusters exhibit the smallest difference between the Cu-center and Ni-center structural patterns, as the two fitting lines almost coincide with each other. This indicates the energy differences between Ni-center and Cu-center structures are very small. In contrast, the two fitting lines for Cu-center and Rh-center structures show relatively big difference. According to previous discussions on alloy nanoclusters⁹, as the alloy bond strength is greater, intimate mixing is favored. If the bond strength is lower, segregation of the two metal components is favored; the metal atoms with lower surface energy tend to move to the surface of alloy cluster to form a shell, like Cu.

Before discussing the catalysts for HER, we examine the structural stability of these 55-atom alloy clusters. To confirm the stability of the core-shell structures, we have constructed totally 7924 structures for the Cu_{55-n}M_n clusters with n from 1 to 22, corresponding to the dopant concentrations ranging from 1.8% to 40.0%, and the most stable cluster structures are shown in Supplementary Figs 1–4. The excess energies of structures for each dopant concentration are plotted in Fig. 4. As shown, the CuNi alloy clusters are energetically favorable than the other alloy clusters, that is to say, Ni doping into Cu clusters is most

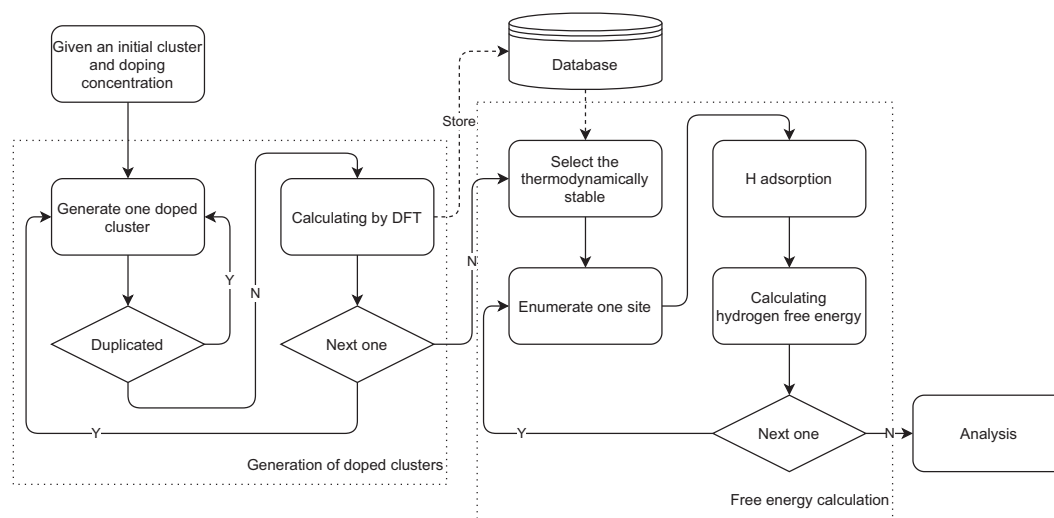


Fig. 1 Workflow of high-throughput screening. The workflow of computational high-throughput screening of HER electrocatalysts on metal alloy nanoclusters.

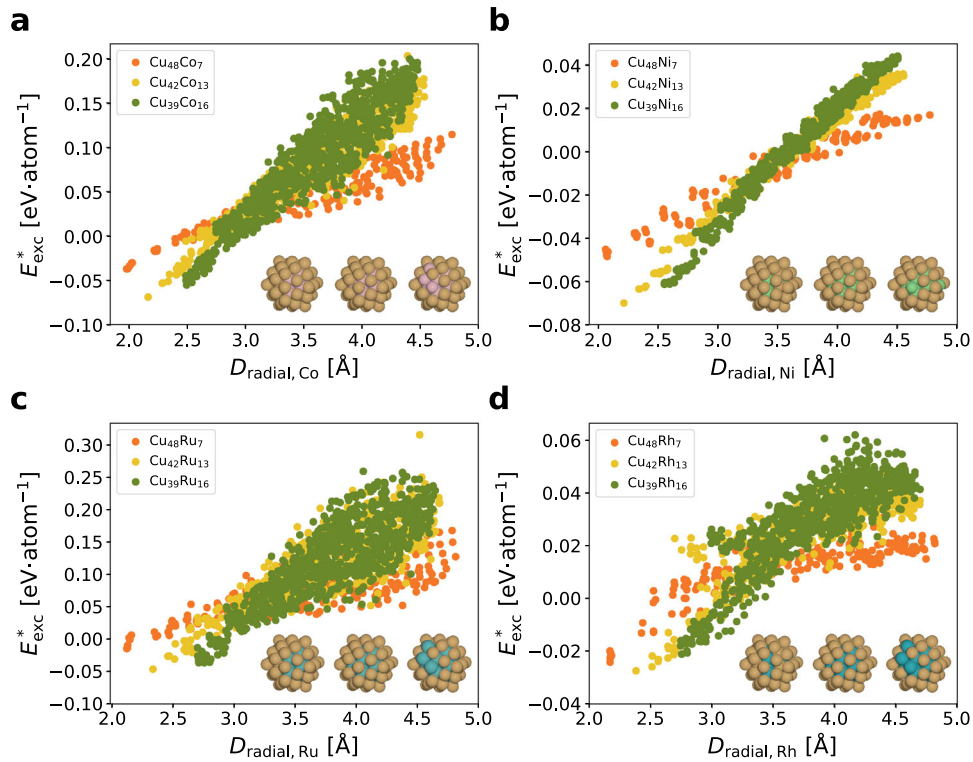


Fig. 2 Energies and structural patterns for alloy clusters. The excess energies (E_{exc}^*) for (a) $\text{Cu}_{55-n}\text{Co}_n$, (b) $\text{Cu}_{55-n}\text{Ni}_n$, (c) $\text{Cu}_{55-n}\text{Ru}_n$, and (d) $\text{Cu}_{55-n}\text{Rh}_n$ alloy clusters with $n = 7, 13$, and 16 as a function of the average radial distance ($D_{\text{radial}, M}$) between the dopant metal atoms and the center atom. The insets show the lowest-energy configurations.

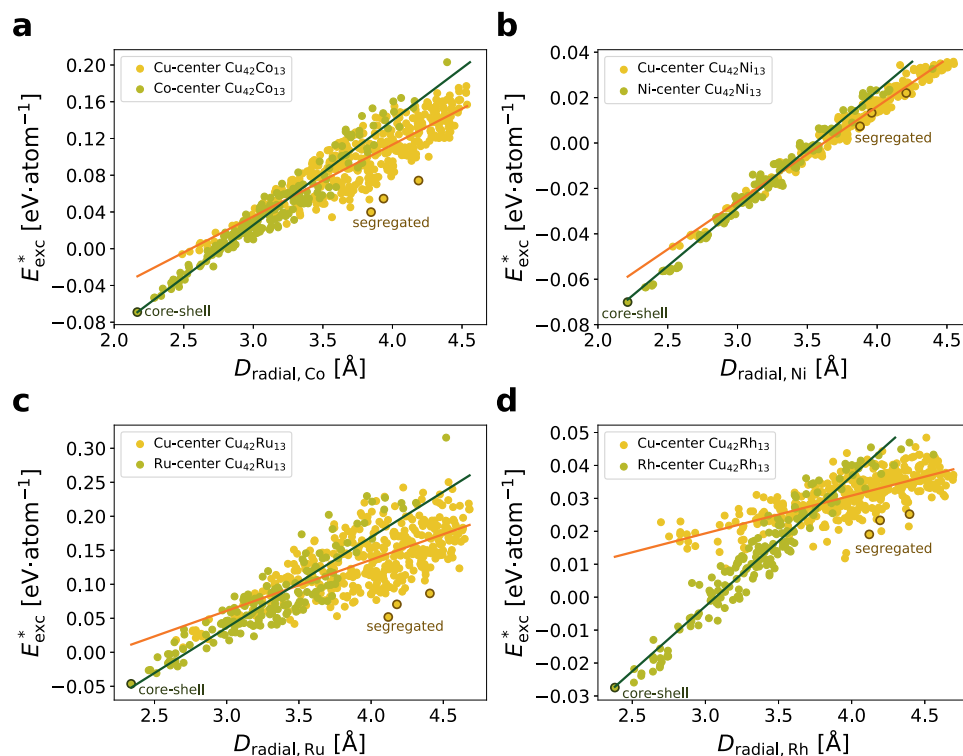


Fig. 3 Energies and structural patterns for $\text{Cu}_{42}\text{M}_{13}$ clusters. The excess energies (E_{exc}^*) for (a) $\text{Cu}_{42}\text{Co}_{13}$, (b) $\text{Cu}_{42}\text{Ni}_{13}$, (c) $\text{Cu}_{42}\text{Ru}_{13}$, and (d) $\text{Cu}_{42}\text{Rh}_{13}$ alloy clusters as a function of the average radial distance ($D_{\text{radial}, M}$) between the dopant metal atoms and the central atom. The core-shell structure and the segregated structure are illustrated.

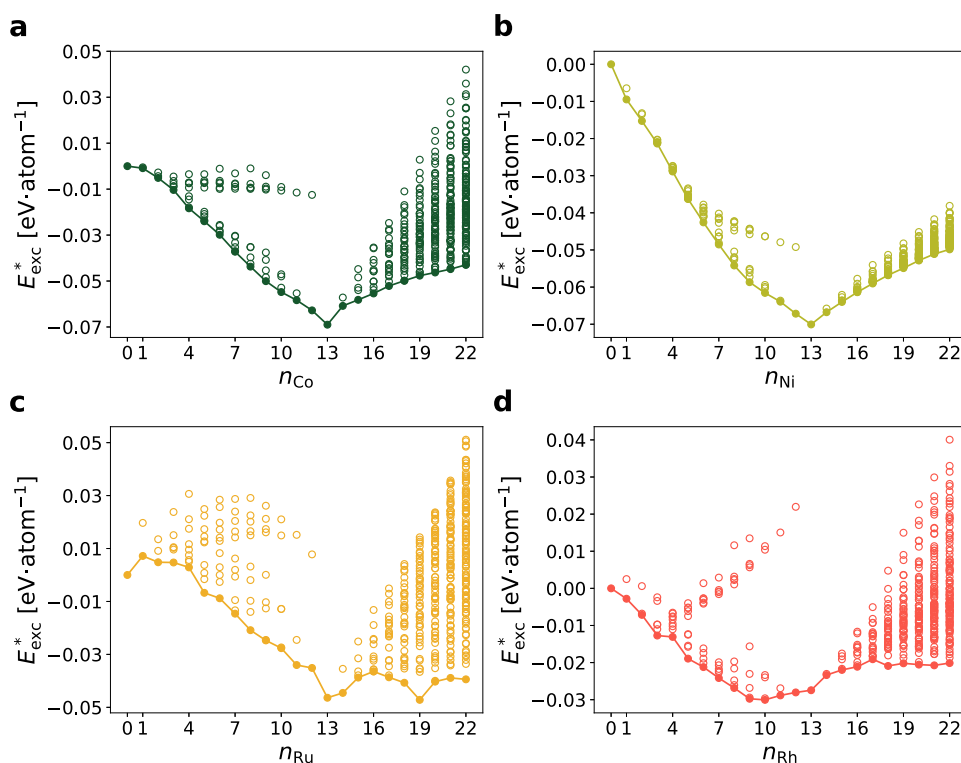


Fig. 4 Energetic stability for core-shell clusters. The excess energies (E_{exc}^*) for core-shell structures of (a) $\text{Cu}_{55-n}\text{Co}_n$, (b) $\text{Cu}_{55-n}\text{Ni}_n$, (c) $\text{Cu}_{55-n}\text{Ru}_n$, and (d) $\text{Cu}_{55-n}\text{Rh}_n$ alloy clusters as a function of dopant concentration n , where n_M is the number of dopant metal atoms in the 55-atom cluster. The lowest excess energies for each dopant concentration are connected by a solid line.

thermodynamically favorable compared to the other metal doping. Doping Co into Cu clusters leads to the second low-energy alloy clusters. The Ru and Rh doped Cu clusters are not as stable as Co and Ni, which may be attributed to the relatively larger volume expansion of the alloy clusters (the covalent radii of Ru and Rh atoms are 1.46 Å and 1.42 Å, respectively, larger than that of Cu atom with 1.32 Å). The diameter of the pure Cu_{55} clusters is 9.606 Å. Taking $\text{Cu}_{33}\text{M}_{22}$ cluster as an example, the equilibrium diameters for $\text{Cu}_{33}\text{Co}_{22}$, $\text{Cu}_{33}\text{Ni}_{22}$, $\text{Cu}_{33}\text{Ru}_{22}$, and $\text{Cu}_{33}\text{Rh}_{22}$ clusters are 9.489 Å, 9.548 Å, 9.939 Å, and 10.058 Å, respectively. The Ru and Rh doping into the Cu clusters result in the largest volume expansion to increase energy. Therefore, doping Ni into Cu clusters is most feasible to form the CuNi alloy clusters in experiments, compared to the other alloy clusters considered here.

For the icosahedral core-shell structures of the $\text{Cu}_{55-n}\text{M}_n$ clusters, the outer shell consists of 42 atoms distributed in two different sites, i.e., 30 atoms at the edge sites and 12 atoms at the vertex sites. Meanwhile, there are 13 atoms in the core region. When the number of dopant metal atoms is less than 13, the dopants prefer to occupy the core positions, and the Cu atoms tend to stay on the outer shell. As the dopant concentration increases ($n > 13$), Co, Ni, and Rh atoms start to occupy the edge sites on the shell, while Ru atoms seem to prefer both sites. For the $\text{Cu}_{55-n}\text{Ni}_n$ clusters with dopant concentration of $13 < n_{\text{Ni}} \leq 19$, Ni atoms are distributed dispersedly and symmetrically on edge sites; but when $n_{\text{Ni}} > 19$, Ni atoms prefer to form small island on the cluster surface. Our results have demonstrated that, the most stable structural pattern of the $\text{Cu}_{55-n}\text{M}_n$ clusters ($M = \text{Co}, \text{Ni}, \text{Ru}, \text{and Rh}$) is the core-shell structure, in which the dopant metal atoms occupy the core and the Cu atoms are distributed as the shell. It is really important to consider the structural evolution of the 55-atom clusters with high concentration doping. To this point, we have compared our icosahedron structures to another typical cuboctahedron structures. Because the structural evolution

is more possible to occur in the high doping concentration, here we have investigated the $\text{Cu}_{33}\text{M}_{22}$ ($M = \text{Co}, \text{Ni}, \text{Ru}, \text{and Rh}$) clusters, which is the largest doping concentration considered in our work. Upon structural optimization (40 structures for each dopant), most of the cuboctahedron structures are reconstructed and transformed into icosahedron structures. The total energies of the low-energy cuboctahedron structures with different mixing patterns are summarized in Supplementary Table 1. The core-shell structural pattern in cuboctahedron configuration also possesses the lowest energy, but still higher by more than 2 eV in total energy than the icosahedron configuration. Therefore, the Cu_{55} clusters with high doping concentration (less than half) will remain the same structure as Cu_{55} cluster.

Hydrogen adsorption free energy

Based on the most stable structures of the $\text{Cu}_{55-n}\text{M}_n$ clusters, we further investigate the HER activities on the surface of these alloy clusters. To evaluate the HER activity, the H adsorption free energy is commonly used as a reasonable descriptor for a wide range of metals and alloys^{20–22}. The HER process includes two steps. The first step involves bonding of H to the catalyst $\text{H}^+ + \text{e}^- + * \rightarrow \text{H}^*$, where * denotes a site on the surface able to bind to H. The second step is the release of molecular H_2 through one of the two processes: $2\text{H}^* \rightarrow \text{H}_2 + *$ or $\text{H}^+ + \text{e}^- + \text{H}^* \rightarrow \text{H}_2 + *$. If H binds too weakly on the catalyst surface, it is difficult to activate in the first step; but when the H binds too strongly on the catalyst, the active sites will be occupied and suppress the second step of releasing H_2 . Thus, the optimal H adsorption free energy should be close to zero. The H adsorption free energy is calculated at a potential $U = 0$ relative to the standard hydrogen electrode at $\text{pH} = 0$. The free energy of $\text{H}^+ + \text{e}^-$ is by definition the same as that of $1/2 \text{H}_2$ at standard conditions. Therefore, herein we have evaluated the adsorption free energy for hydrogen (ΔG_{H}) on all the possible adsorption sites for the stable structures of $\text{Cu}_{55-n}\text{M}_n$ nanoclusters.

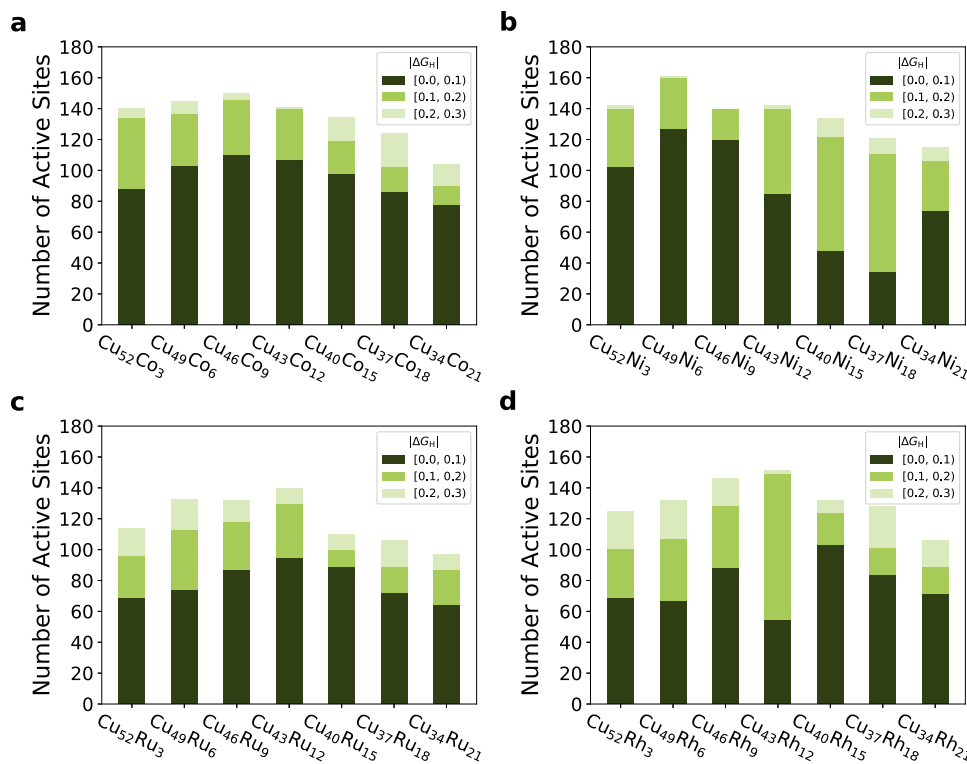


Fig. 5 HER activities of alloy catalysts. Total number of active sites counted in different ranges of $|\Delta G_H|$ values as a function of various compositions for (a) Cu_{55-n}Co_n, (b) Cu_{55-n}Ni_n, (c) Cu_{55-n}Ru_n, and (d) Cu_{55-n}Rh_n alloy clusters. The dark green, green, and light green colors indicate the total number of active sites with the $|\Delta G_H|$ values in the range of [0.0, 0.1), [0.1, 0.2), [0.2, 0.3), respectively.

Totally, we have 242 H adsorption sites on each cluster structure (more details can be found in Supplementary Fig. 5, and those with H adsorption free energy ($|\Delta G_H|$) less than 0.1 eV are defined as the optimal active sites.

Here we utilize the number of the optimal active sites with $|\Delta G_H| < 0.1$ eV as the descriptor to evaluate the HER performance of nanoalloy catalysts. The total number of active sites counted in different ranges of $|\Delta G_H|$ for pure Cu₅₅, M₅₅, and Cu_{55-n}M_n alloy clusters with various compositions are summarized and plotted in Fig. 5 and Supplementary Fig. 6. Compared to pure metal clusters, the amount of active sites with $|\Delta G_H| < 0.1$ eV on the Cu_{55-n}M_n alloy clusters is greatly increased, indicating the superior electrocatalytic HER performance. Among the four kinds of alloy clusters, the CuCo and CuNi alloy clusters are predicted to have more active sites with $|\Delta G_H| < 0.1$ eV, in contrast to CuRu and CuRh clusters. It is worthy to note that pure Ni clusters exhibit very poor HER activity, but the CuNi alloy clusters display the best catalytic efficiency. As we have discussed above, Ni-doped Cu₅₅ clusters are the most stable alloy clusters; thus the CuNi alloy clusters not only possess the stable structures but also exhibit the best HER activity. By increasing the dopant concentrations, the HER performance of the Cu_{55-n}M_n alloy clusters varies. But it is interesting to note that, the HER catalytic efficiency has a maximum value for each type of Cu_{55-n}M_n alloy clusters, such as Cu₄₆Co₉ for CuCo alloy clusters and Cu₄₉Ni₆ for CuNi alloy clusters, respectively. With further increasing the dopant concentration, the amount of the active sites with $|\Delta G_H| < 0.1$ eV decreases, indicating a remarkable drop in HER performance.

We now check the optimal active site on the surface of these alloy clusters. For the icosahedral clusters, we have totally 242 adsorption sites; among them, we have 42 on-top sites on V (vertex) or E (edge), 120 bridge sites on VE or EE, 80 three-fold sites on VEE or EEE. The H adsorption free energies between

-0.1 eV and 0.1 eV are defined as the optimal values for HER. After examining carefully, it is found that the majority of the optimum sites with the H free energies between -0.1 eV and 0.1 eV are bridge sites of VE and three-fold sites of VEE on Cu atoms. The adsorbed H atom prefers to form a ring structure around the shell Cu atoms (see Fig. 6), which is remarkably different from the pure Cu clusters. The binding of H atoms on top of V or E sites are too weak with more positive ΔG_H values. As an H atom is adsorbed on the bridge sites of EE, it will spontaneously move to the EEE or VEE sites. In addition, most of the optimal active sites are Cu sites instead of the dopant metal sites, which may account for why the HER activity starts to decrease as the dopants concentration further increasing. As the dopants concentration increases, the dopant metal atoms are gradually exposed on the cluster surface, and the H adsorption free energy on these dopant metal sites are mostly strong with the ΔG_H values more negative. Among the alloy clusters that we have considered here, the core-shell Cu₄₉Ni₆ alloy cluster is recognized as the promising electrocatalyst for HER with high stability and the most abundant active sites. Some related experiments on CuNi nanoparticle as the high-efficient HER catalyst have been reported recently and proved our theoretical prediction^{14,15}.

Descriptor for the structure-activity relationship

To deeply understand the modulation mechanism of alloy metal clusters, Bader charge distributions of the Cu_{55-n}M_n clusters have been further calculated and discussed. Since the HER occurs on the cluster surface, we focus on the charge distributions of the shell metal atoms. The metal atoms in the pure Cu₅₅ cluster with icosahedral structure possess two inequivalent sites on the surface, i.e., vertex site and edge site. For the pure Cu₅₅ cluster, the charge of Cu atoms on the vertex site (Q_{Cu}) is -0.05 e, and that of Cu atoms on the edge site (Q_{Cu}) is +0.002 e; thus the outer shell

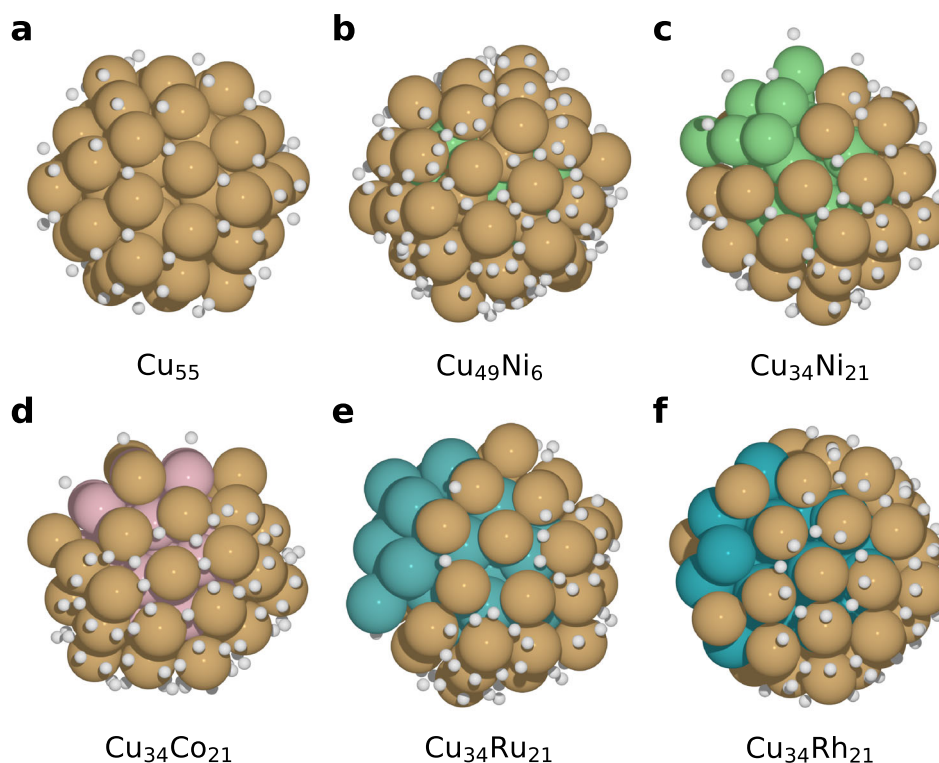


Fig. 6 Active sites for HER. The optimal adsorption sites for H on the surface of (a) pure Cu_{55} , (b) $\text{Cu}_{49}\text{Ni}_6$, (c) $\text{Cu}_{34}\text{Ni}_{21}$, (d) $\text{Cu}_{34}\text{Co}_{21}$, (e) $\text{Cu}_{34}\text{Ru}_{21}$, and (f) $\text{Cu}_{34}\text{Rh}_{21}$ clusters with $|\Delta G_{\text{H}}| < 0.1$ eV.

of Cu_{55} are negatively charged. By doping transition metal atoms into Cu clusters, the charge distributions on the surface of alloy clusters have been changed accordingly. In the alloy clusters, the edge-Cu atoms become positively charged, and finally enlarging the charge difference between the adjacent E_{Cu} and V_{Cu} . Taking $\text{Cu}_{49}\text{Ni}_6$ cluster as an example, it is found that the charge values on E_{Cu} reach $+0.03$ e, while the on-site charges of V_{Cu} are around -0.07 e, inducing a substantial charge difference of 0.1 e between the edge-Cu and the adjacent vertical-Cu atoms. As discussed above, most of the optimal sites with high HER activity are the bridge sites of VE on the shell Cu atoms, so the H adsorption free energies should be associated with the charge values of two adjacent Cu atoms locating on the vertex site and edge site. Hence, we suggest the average charge difference between two adjacent E_{Cu} and V_{Cu} as a descriptor to probe the structure-activity relationship and screening for the HER electrocatalyst, which is defined by:

$$\Delta Q_{\text{Cu-Cu}} = \frac{\sum_i^n \left(\sum_j^m |q_{\text{CuV}_j} - q_{\text{CuE}_j}| \right)}{N_{\text{VE}}} \quad (2)$$

where q_{CuV_j} is the charge of j^{th} Cu atom on vertex site, and q_{CuE_j} is the charge of j^{th} Cu atom located on shell next to the i^{th} Cu atom. n is the number of Cu atoms on vertex site, and m is the number of the i^{th} Cu atom's neighboring Cu atoms on cluster shell. N_{VE} is chosen as 60, which is calculated from the total number of Cu-Cu bonds between Cu atoms on vertex and edge sites ($m \cdot n$) in Cu_{55} .

The proposed descriptor is controlled by two factors: the first one is the amount of active metal atoms on the cluster surface (i.e., the surface Cu atoms in CuM clusters), and the second one is the charge difference between two adjacent Cu atoms on vertex and edge sites. We have calculated the $\Delta Q_{\text{Cu-Cu}}$ values for the $\text{Cu}_{55-n}\text{M}_n$ alloy clusters. The $\Delta Q_{\text{Cu-Cu}}$ values on the $\text{Cu}_{55-n}\text{M}_n$ alloy clusters vary with the increasing dopant concentration, and there is a maximum value for each type of $\text{Cu}_{55-n}\text{M}_n$ alloy clusters, as shown in Fig. 7a. After reaching the maximum value, the $\Delta Q_{\text{Cu-Cu}}$ values start decreasing, because more dopant metal atoms are

distributed on the shell and less Cu-Cu bonds are preserved on the cluster surface. The variations of the charge difference $\Delta Q_{\text{Cu-Cu}}$ values exhibit almost identical trends to the total number of active sites on the alloy clusters with $|\Delta G_{\text{H}}| < 0.1$ eV. For example, the $\text{Cu}_{46}\text{Co}_9$ and $\text{Cu}_{49}\text{Ni}_6$ clusters exhibit superior HER activities with relatively more active sites of $|\Delta G_{\text{H}}| < 0.1$ eV, and they also possess the larger $\Delta Q_{\text{Cu-Cu}}$ values. To make it clear, we have calculated the relationship between the $\Delta Q_{\text{Cu-Cu}}$ values and the number of active sites on the $\text{Cu}_{55-n}\text{M}_n$ alloy clusters and found a nearly linear dependency between them, as displayed in Fig. 7b. As the charge differences between two adjacent surface Cu atoms become larger, the HER activity described by the number of optimal active sites almost linearly increases. Therefore, $\Delta Q_{\text{Cu-Cu}}$ value could be identified as a descriptor to evaluate the HER activity. Compared to the traditional descriptor (e.g., H adsorption free energy) for evaluating the HER performance and confirming the optimal active sites, our proposed descriptor of charge difference is remarkably easier to compute, which greatly enhances the computational screening efficiency and facilitates the HER evaluations on the more complicated alloy nanoclusters.

Machine-learning prediction

Machine-learning technique combined with the DFT calculations holds an immense potential for the acceleration of catalyst discovery²³. As we discussed above, the charge distribution around the active site is closely related to the strength of the hydrogen adsorption, so here we introduce the machine-learning techniques based on our large data set of high-throughput DFT calculations. The neural network utilizes the fingerprints of the H adsorption site, and the features consist of the coordination number and the mean Bader charge of the first-neighbor and the second-neighbor atoms of the active site on the outer atomic shell. The H adsorption free energies for different sites on the Cu-based alloy nanoclusters are predicted with a subsequently increasing training set. A learning curve in Fig. 8a shows that

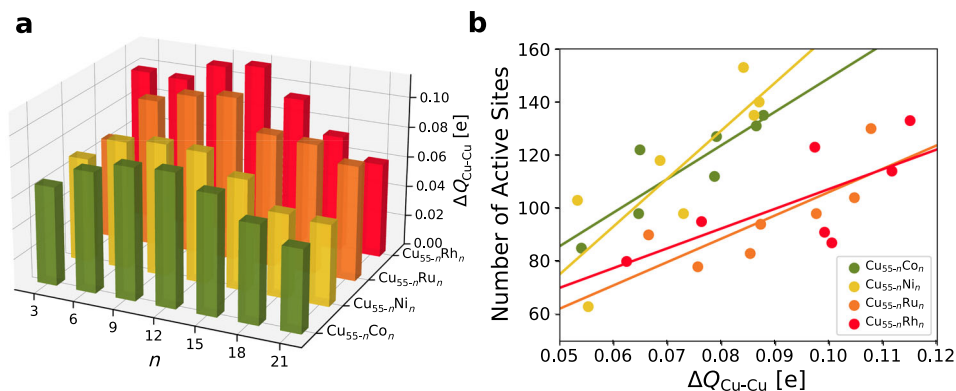


Fig. 7 Descriptor to evaluate HER performance. **a** The average charge difference $\Delta Q_{\text{Cu-Cu}}$ between two adjacent Cu atoms on the edge site and vertical site for $\text{Cu}_{55-n}\text{M}_n$ ($M = \text{Co}, \text{Ni}, \text{Ru}, \text{and Rh}$) alloy clusters, **b** linear relationship between the number of active sites ($|\Delta G_{\text{H}}| < 0.15 \text{ eV}$) and the average charge difference $\Delta Q_{\text{Cu-Cu}}$.

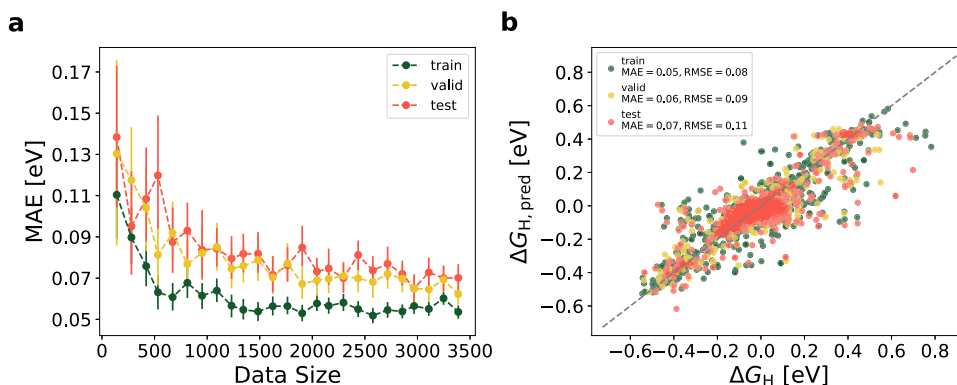


Fig. 8 Machine-learning prediction. **a** Learning curve of the neural network. The errors are averaged over 25 randomized runs and the error bars indicate the 95% confidence interval; **b** parity plots between predicted and DFT-calculated ΔG_{H} values.

about 700 adsorption free energies from DFT calculations (including 420 training data) is enough to achieve an ideal mean absolute error (MAE) of 0.10 eV from our features. We obtain a similar machine-learning precision to a recent machine-learning work with an accuracy of 0.11 eV MAE by using the other features²⁴. Based on this, we use 3388 adsorption free energies to predict the H adsorption free energy on nanoclusters with a MAE error of 0.07 eV and a root-mean-square error (RMSE) of 0.11 eV on the test set, respectively, as shown in Fig. 8b. Our input properties for machine-learning models are ideally available without heavy computation from the H adsorption relaxation. Therefore, the charge of atoms around the adsorption site is closely correlated with the H adsorption free energies, which could be a useful descriptor to estimate the HER activity and screening for the high-efficient HER electrocatalysts.

In summary, we have performed high-throughput DFT calculations to screening for the efficient HER electrocatalysts on the alloy nanoclusters. The following important conclusions have been addressed. First, considering different structural patterns, the core-shell structures with M-core and Cu-shell are energetically preferred for Cu-based alloy clusters of $\text{Cu}_{55-n}\text{M}_n$ ($M = \text{Co}, \text{Ni}, \text{Ru}, \text{and Rh}$), and CuNi is the most stable alloy clusters among these four kinds of alloy clusters. Second, the total number of the active sites with $|\Delta G_{\text{H}}| < 0.1 \text{ eV}$ is utilized as a descriptor to evaluate the HER performance of the alloy catalysts, which could be efficiently modulated by the dopant concentration. Among all the explored alloy clusters, $\text{Cu}_{49}\text{Ni}_6$ exhibits the best HER activity, which is significantly improved in comparison with the pure Cu_{55} . Third, we propose a descriptor to screen for the HER electrocatalysts by combining DFT calculations and machine-learning techniques.

Finally, our high-throughput screening strategy could be widely applicable to the other alloy nanoclusters, and also extend to the other electrocatalytic reactions.

METHODS

DFT calculations

The high-throughput computations were performed using DFT as implemented in the Vienna Ab-initio Simulation Package (VASP)^{25,26}. The projector augmented wave (PAW) method was used to describe the ion-electron interactions^{27,28}. The valence states of the transition metals are considered as $3d^{10}4p^1$ for Cu atom, $3d^84s^1$ for Co atom, $3d^84s^1$ for Ni, $4d^75s^1$ for Ru, and $4d^85s^1$ for Rh, respectively. The Perdew-Burke-Ernzerh (PBE) functional within the generalized gradient approximation (GGA) was used to describe the exchange-correlation interactions^{29,30}, and the D3 Grimme's scheme was adopted to account for London dispersion interactions³¹. Each cluster was placed in a cubic box of $20 \text{ \AA} \times 20 \text{ \AA} \times 20 \text{ \AA}$ to minimize the interaction from periodical images. The Γ -point was adopted for sampling the Brillouin zone of simulation supercell throughout the calculations. The kinetic energy cutoff for plane wave was set to 350 eV. All the geometries were fully optimized until the maximal components of force on each atom converged to less than $5 \times 10^{-2} \text{ eV \AA}^{-1}$.

To compare the relative energies among different compositions of clusters, the excess energy is defined by⁹:

$$E_{\text{exc}}^* = \frac{E_{\text{Cu}_{55-n}\text{M}_n} - (55-n)E_{\text{Cu}} - nE_{\text{M}}}{55} \quad (3)$$

where $E_{\text{Cu}_{55-n}\text{M}_n}$ is the total energy of $\text{Cu}_{55-n}\text{M}_n$ cluster, E_{Cu} and E_{M} are the total energies of the Cu_{55} cluster and M_{55} cluster ($M = \text{Co}, \text{Ni}, \text{Ru}, \text{and Rh}$), respectively.

The adsorption free energy for hydrogen is defined as:

$$\Delta G_{\text{H}} = \Delta E_{\text{H}} + \Delta E_{\text{ZPE}} - T\Delta S \quad (4)$$

where the ΔE_{H} is the adsorption energy of H species on the cluster surface from DFT calculations. The ΔE_{ZPE} and ΔS are the difference for the zero-point energy and entropy between the adsorbed H and half of H_2 molecule in the gas phase, respectively. The zero-point energy and entropy are calculated at the standard conditions corresponding to the pressure of 101325 Pa (~1 bar) at the temperature of 298.15 K. The coverage of H adsorbate is difficult to consider for the cluster structures due to more expensive calculations; but for the similar systems (the Cu-based alloy clusters in our work possess the same structure), the H adsorption free energy could provide a reasonable trend under the assumption of same coverage.

The machine-learning protocol

The neural network consists of one input layer, three hidden layers and one output layer. The number of hidden layer neurons are 32, 32, and 64. The Rectified Linear Unit activation function³² was chosen for each hidden layer. The overfitting issue was prevented using the L_2 regularization³³. The neural network is subjected to a supervised training scheme by using a backpropagation algorithm³⁴ implemented in the TensorFlow framework³⁵. All the input features are normalized to avoid the use of raw variables with different range of values. The data is transformed with $x' = \frac{x_i - \mu}{\sigma}$, where x' , x_i , μ , and σ represent the normalized data, input data, mean and sample standard deviation values of the input data, respectively. The Adam optimizer³⁶ is selected with the learning rate of 0.001. Prediction errors are evaluated by mean absolute errors, and the cross-validation technique³⁷ is applied to estimate the accuracy and robustness of the neural network training. The DFT results of hydrogen adsorption are divided into 60%, 20%, and 20% for training, validation, and test sets, respectively.

DATA AVAILABILITY

The data obtained in this study are available from the authors upon reasonable request.

CODE AVAILABILITY

The codes developed in this study are available from the authors upon reasonable request.

Received: 12 August 2020; Accepted: 8 March 2021;
Published online: 06 April 2021

REFERENCES

- Sultan, S. et al. Single atoms and clusters based nanomaterials for hydrogen evolution, oxygen evolution reactions, and full water splitting. *Adv. Energy Mater.* **9**, 1900624 (2019).
- Seh, Z. W. et al. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science* **355**, eaad4998 (2017).
- Nørskov, J. K. & Christensen, C. H. Toward efficient hydrogen production at surfaces. *Science* **312**, 1322–1323 (2006).
- Turner, J. A. Sustainable hydrogen production. *Science* **305**, 972–974 (2004).
- Wang, P. et al. Precise tuning in platinum-nickel/nickel sulfide interface nanowires for synergistic hydrogen evolution catalysis. *Nat. Commun.* **8**, 14580 (2017).
- Gong, Q. et al. Ultrasmall and phase-pure W_2C nanoparticles for efficient electrocatalytic and photoelectrochemical hydrogen evolution. *Nat. Commun.* **7**, 13216 (2016).
- Cabán-Acevedo, M. et al. Efficient hydrogen evolution catalysis using ternary pyrite-type cobalt phosphosulphide. *Nat. Mater.* **14**, 1245–1251 (2015).
- Tyo, E. C. & Vajda, S. Catalysis by clusters with precise numbers of atoms. *Nat. Nanotechnol.* **10**, 577–588 (2015).
- Ferrando, R., Jellinek, J. & Johnston, R. L. Nanoalloys: from theory to applications of alloy clusters and nanoparticles. *Chem. Rev.* **108**, 845–910 (2008).
- Huang, X. et al. High-performance transition metal-doped Pt_3Ni octahedra for oxygen reduction reaction. *Science* **348**, 1230–1234 (2015).
- Liu, L. & Corma, A. Metal catalysts for heterogeneous catalysis: from single atoms to nanoclusters and nanoparticles. *Chem. Rev.* **118**, 4981–5079 (2018).
- Gawande, M. B. et al. Core-shell nanoparticles: synthesis and applications in catalysis and electrocatalysis. *Chem. Soc. Rev.* **44**, 7540–7590 (2015).
- Stamenkovic, V. R. et al. Trends in electrocatalysis on extended and nanoscale Pt-bimetallic alloy surfaces. *Nat. Mater.* **6**, 241–247 (2007).
- Lin, Z. et al. Manipulating the hydrogen evolution pathway on composition-tunable CuNi nanoalloys. *J. Mater. Chem. A* **5**, 773–781 (2017).
- Li, Z. et al. Mesoporous hollow Cu–Ni alloy nanocage from Core–Shell Cu@Ni nanocube for efficient hydrogen evolution reaction. *ACS Catal.* **9**, 5084–5095 (2019).
- Sankar, M. et al. Designing bimetallic catalysts for a green and sustainable future. *Chem. Soc. Rev.* **41**, 8099–8139 (2012).
- Greeley, J., Jaramillo, T. F., Bonde, J., Chorkendorff, I. & Nørskov, J. K. Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. *Nat. Mater.* **5**, 909–913 (2006).
- Darby, S., Mortimer-Jones, T. V., Johnston, R. L. & Roberts, C. Theoretical study of Cu–Au nanoalloy clusters using a genetic algorithm. *J. Chem. Phys.* **116**, 1536–1550 (2002).
- Lysgaard, S., Landis, D. D., Bligaard, T. & Vegge, T. Genetic algorithm procreation operators for alloy nanoparticle catalysts. *Top. Catal.* **57**, 33–39 (2014).
- Nørskov, J. K. et al. Trends in the exchange current for hydrogen evolution. *J. Electrochem. Soc.* **152**, J23–J26 (2005).
- Liu, P. & Rodriguez, J. A. Catalysts for hydrogen evolution from the [NiFe] hydrogenase to the Ni_2P (001) surface: the importance of ensemble effect. *J. Am. Chem. Soc.* **127**, 14871–14878 (2005).
- Hinnemann, B. et al. Biomimetic hydrogen evolution: MoS_2 nanoparticles as catalyst for hydrogen evolution. *J. Am. Chem. Soc.* **127**, 5308–5309 (2005).
- Schlexer Lamoureux, P. et al. Machine learning for computational heterogeneous catalysis. *ChemCatChem* **11**, 3581–3601 (2019).
- Jäger, M. O. J., Ranawat, Y. S., Canova, F. F., Morooka, E. V. & Foster, A. S. Efficient machine-learning-aided screening of hydrogen adsorption on bimetallic nanoclusters. *ACS Comb. Sci.* **22**, 768–781 (2020).
- Kresse, G. & Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **54**, 11169–11186 (1996).
- Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).
- Kresse, G. & Hafner, J. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B* **47**, 558–561 (1993).
- Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
- Perdew, J. P. & Wang, Y. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B* **45**, 13244–13249 (1992).
- Perdew, J. P., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**, 154104 (2010).
- Nair, V. & Hinton, G. E. Rectified linear units improve restricted Boltzmann machines. in *Proceedings of the 27th International Conference on International Conference on Machine Learning* 807–814 (Omnipress, 2010). <https://doi.org/10.1145/1015330.1015435>.
- Ng, A. Y. Feature selection, L_1 vs. L_2 regularization, and rotational invariance. in *Proceedings of the Twenty-First International Conference on Machine Learning* 78 (Association for Computing Machinery, 2004). <https://doi.org/10.1145/1015330.1015435>.
- Hirose, Y., Yamashita, K. & Hijiya, S. Back-propagation algorithm which varies the number of hidden units. *Neural Netw.* **4**, 61–66 (1991).
- M. Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems. <https://www.tensorflow.org/> (2015).
- Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. in *International Conference for Learning Representations* (2015).
- Zhang, G., Hu, M. Y., Patuwo, B. E. & Indro, D. C. Artificial neural networks in bankruptcy prediction: general framework and cross-validation analysis. *Eur. J. Oper. Res.* **116**, 16–32 (1999).

ACKNOWLEDGEMENTS

This work was supported by the National Key Research and Development Program of China (Grant nos. 2018YFB0703900, 2017YFA0204800, and 2017YFB0701600) and the National Natural Science Foundation of China (91961204 and 21973067). The calculations were carried out at National Supercomputer Center in Tianjin, and the calculations were performed on TianHe-1 (A). This project was also supported by the Fund for Collaborative Innovation Center of Suzhou Nano Science & Technology, the Priority Academic Program Development of Jiangsu Higher Education Institutions.

AUTHOR CONTRIBUTIONS

L.W. conceived and designed the project. X.M. performed all the DFT and machine-learning calculations, plotted the figures and prepared the initial draft of the paper. Y.X. and P.W. contributed to the results analysis. L.W., J.Z., and Y.L. supervised the project and revised the paper.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41524-021-00514-8>.

Correspondence and requests for materials should be addressed to L.W., Y.L. or J.Z.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021