Journal of Materials Chemistry A

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: J. Mater. Chem. A, 2023, 11, 9964

Basal plane activation of two-dimensional transition metal dichalcogenides *via* alloying for the hydrogen evolution reaction: first-principles calculations and machine learning prediction†

Yiqing Chen, [a] Ying Zhao, [a] Pengfei Ou *[ab] and Jun Song *[a]

Two-dimensional transition metal dichalcogenides (2D TMDCs) show promise as potential inexpensive electrocatalysts for the hydrogen evolution reaction (HER). However, their performance is bottlenecked by the inertness of the basal plane. The present study demonstrates alloying as a viable route to address such limitations. A machine learning workflow based on density functional theory (DFT) calculations has been established to predict the HER activity and stability for a series of 2D cation-mixed TMDC alloys of various compositions. The results showed that alloying exhibits a substantial effect in reducing the Gibbs free energy of hydrogen adsorption (ΔG_H) on the basal plane, able to render optimal ΔG_H for the HER for certain TMDC alloys. The stability prediction of these TMDC alloys further showed their potential to be synthesized in experiments. The mechanism underlying this alloying induced basal plane activation originates from the electronic effect, in particular the p-band shifting, resulting from the chemical composition variation. The findings are expected to serve as guidance for the rational design and discovery of TMDC alloys for catalytic applications.

Received 5th March 2023
Accepted 13th April 2023

DOI: 10.1039/d3ta01361h

rsc.li/materials-a

1. Introduction

Hydrogen is the cleanest fuel and is considered an ideal energy carrier for future energy systems.[1] The scalable H_2 production through water splitting requires efficient catalysts for the hydrogen evolution reaction (HER).[2] Conventionally, platinum (Pt) group metals are regarded as the most active catalysts for the HER; nevertheless, the high cost and material scarcity hinder their large-scale application.[3,4] As a result, developing alternative, non-Pt catalysts with low cost and high performance is of great importance.[5]

Of many different candidates, two-dimensional transition metal dichalcogenides (2D TMDCs) have received substantial attention.[6–8] For example, MoS_2, WS_2, and WSe_2 have been intensively studied by both experimental[9] and computational methods[10] as potential HER electrocatalysts. However, the HER performance of these 2D TMDCs is still limited by the low density and poor reactivity of active sites. Consequently, there have been great efforts attempting to overcome such limitations. One route towards further enhancement in the catalytic activity of 2D TMDCs is through structural modification.[11–13] As an example, by exposing a high fraction of edge sites, 2D MoS_2 can catalyze the HER at a moderate overpotential of 0.1–0.2 V with Gibbs free energy of adsorption close to zero (<0.1 eV).[14] Another popular method is the introduction of defects in TMDCs. By introducing sulfur vacancies and strain, the HER performance of MoS_2 has been greatly improved, yielding the optimal ΔG_H close to 0 eV.[15]

In addition to structural modification, it was also reported that substantial enhancement of HER activity can be achieved through composition modification, that is, *via* tailoring the chemical composition of 2D TMDCs by alloying at the metal or chalcogen sites.[16–20] For example, substituting a high concentration of W to Mo in MoS_2 is known to promote the electrocatalytic performance of MoS_2.[21] Many 2D TMDCs have a structure analogous to that of graphite, where transition metal atoms are sandwiched between chalcogen atoms in a configuration of a triangular prism.[22] The similar geometry of various 2D TMDCs enables the fabrication of ternary cation-mixed alloys.[17,19,23–25] These studies suggest alloying in 2D TMDC catalysts as a promising means for tuning reactivity and producing new active sites for the HER. Unlike most previous theoretical and experimental studies on the HER of 2D TMDCs where the performance enhancement was achieved *via* exposed edge sites or localized defect sites, alloying involves activation of the basal plane. The 2D nature of TMDCs makes the basal plane activation particularly attractive as it can provide a high density of active sites on the tendentiously exposed surface while

[a]Department of Mining and Materials Engineering, McGill University, 3610 University St, Montreal, QC H3A 0C5, Canada. E-mail: jun.song2@mcgill.ca

[b]Department of Chemistry, Northwestern University, 2145 Sheridan Rd, Evanston, IL, 60208, USA. E-mail: pengfei.ou@northwestern.edu

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d3ta01361h

retaining the structural integrity.[12] However, to date basal plane activation of 2D TMDC alloys in the context of the HER has not been systematically explored. Thus, it is crucial to develop a fundamental understanding of this aspect to guide the rational design of 2D TMDC alloys for highly active electrocatalysts for the HER.

With the help of theoretical surface science developed over the last few decades, the density functional theory (DFT) method has become an accurate and efficient computational approach towards the investigation of adsorption and reaction processes on material surfaces.[26] Nonetheless, computational exploration of 2D TMDC alloys with multiple transition metal and/or chalcogen constituents by DFT is non-trivial. The complexity of 2D TMDC alloys arising from the multitudinous variations in chemical composition and combinations of various metals necessitates several orders of magnitude more calculations than those needed for pristine single transition metal 2D TMDC structures. This renders it a formidable challenge to model 2D TMDC alloys using the traditional brute-force computational approach. Recently, it has been demonstrated that such challenges may be partially addressed by the application of machine learning (ML) techniques.[27–35] ML techniques can greatly reduce the computation need compared to the traditional approach, particularly for alloy or other complex material systems, promising a solution to bypass the computational bottleneck for *in silico* materials design and exploration. They have found great success in predicting material properties, such as adsorption energies, d-band centers, bandgaps, vibrational free energies, and melting temperatures, among others.[36–42] With adequate input data collected and the ML framework properly devised, the trained ML models have demonstrated high accuracy at a reasonable computational cost.[36,43,44]

In this work, first-principles density functional theory (DFT) calculations combining machine learning models were employed to study the basal plane activation of cation-mixed 2D TMDC alloys for the HER. Based on the DFT results, an ML workflow was developed to successfully predict the Gibbs free energy of hydrogen adsorption (ΔG_H) on basal planes and the stability of 2D TMDC alloys. The significant effect of alloying in reducing ΔG_H has been demonstrated, and the analysis of stability for those alloys has been provided. The alloy formation process and possible factors leading to such reduction were investigated to reveal the underlying mechanism. Our findings were then summarized and their implications for the TMDC alloy-based catalysts were discussed.

2. Computational methods

2.1 Enumeration of hydrogen adsorption configurations

The ternary cation-mixed TMDC alloy system used in this study contains 9 MX_2 units (*i.e.*, 3 × 3 supercell) with the hydrogen atom adsorbed on top of a certain chalcogen site. With the hydrogen atom fixed, atomic simulation environment (ASE)[45,46] was then used to generate all symmetrically inequivalent adsorption configurations by enumerating each transition metal site in the supercell. In total, 8496 alloy configurations were generated from 72 cation-mixed alloys based on 18 MX_2 pristine structures (M = Cr, Mo, W, V, Nb, Ta, Ti, Zr, Hf, X = S, Se). For each type of cation-mixed alloy, to avoid strain on the structures, a pristine structure (*i.e.* MX_2) with the larger unit cell size was chosen as the template for enumeration.

2.2 Density-functional theory (DFT) calculations

After enumeration, a set of more than 1140 adsorption configurations were randomly selected to perform DFT calculations. On top of the random selection, it is also ensured that at least two configurations per concentration per alloy was selected to make sure the data can correspond to various TMDC alloys in the whole concentration range. First-principles DFT calculations were then performed on those adsorption configurations selected. All calculations were carried out employing the Vienna *ab initio* simulation package (VASP)[47,48] using the projector-augmented wave method.[49] The exchange-correlation functional was described by generalized gradient approximation (GGA) parametrized by Perdew, Burke, and Ernzerhof.[50,51] A kinetic cutoff energy of 500 eV was set for the plane-wave basis functions, and the *k*-points for the Brillouin-zone integration are sampled on a mesh grid of 4 × 4 × 1 and 12 × 12 × 1 for geometry optimization and electronic structure calculation. All structures were relaxed until the atomic forces are less than 0.02 eV $Å^{-1}$ and total energies were converged to 10^{-5} eV. To avoid interactions between repeated images along the layer direction within the periodic boundary condition, a vacuum with 20 Å thickness was introduced to the supercell. The effect of van der Waals (vdW) interactions was considered using the DFT-D3 method.[52,53]

The activity trend of the catalyst towards the HER can be described by the Gibbs free energy of hydrogen adsorption on the surface. According to the Sabatier principle, it is desirable for hydrogen adsorption to have a Gibbs free energy of hydrogen adsorption close to the thermoneutral value near zero.[54] The Gibbs free energy of hydrogen adsorption, denoted as ΔG_H in the following, can be calculated as:

$$\Delta G_H = \Delta E_H + \Delta E_{ZPE} - T\Delta S_H \quad (1)$$

where ΔG_H is the hydrogen adsorption energy, defined as:

$$\Delta E_H = E_{alloy-H} - E_{alloy} - \frac{1}{2}E_{H_2} \quad (2)$$

where $E_{alloy-H}$ is the total energy of a hydrogen-adsorbed alloy structure, E_{alloy} is the total energy of an alloy structure without a hydrogen atom, and E_{H_2} denotes the energy of an isolated hydrogen gas molecule. ΔE_{ZPE} is the difference in zero-point energy of hydrogen in the adsorbed state and the gas phase. ΔS_H is the entropy term approximated as half of the entropy of the gas phase H_2 as $\Delta S_H \approx \frac{1}{2}S_{H_2}$.[55] T is the room temperature of 298.15 K.

To ensure the predicted 2D TMDC alloys are synthesizable, the stability of these alloys was evaluated by their miscibility, *i.e.*, their ability to avoid phase segregation, at a certain temperature. The miscibility of an alloy at temperature T with concentration x can be evaluated by its free energy of mixing $G_{mix}(x, T)$, which is defined as:

$$G_{\mathrm{mix}}(x, T) = \Delta H(x) - T\Delta S(x) \quad (3)$$

where $\Delta H(x)$ is the formation enthalpy, which can be calculated from DFT. $\Delta S(x)$ is the entropy and was estimated by mean-field approximation:[56,57]

$$\Delta S(x) = -k_{\mathrm{B}}[x \ln x + (1 - x)\ln(1 - x)] \quad (4)$$

Here, k_{B} is the Boltzmann constant. Negative and positive $G_{\mathrm{mix}}(x, T)$ correspond to miscible and immiscible alloys, respectively. Since the entropy $\Delta S(x)$ is always positive, $G_{\mathrm{mix}}(x, T)$ can be lowered by increasing the temperature T until complete miscibility in alloys is achieved. We here used the critical temperature for complete miscibility (T_{CM}) of alloys to determine the stability of 2D TMDC alloys. Since we should consider the complete miscibility of alloys under experimentally achievable temperatures, we set $T_{\mathrm{CM}} < 1000$ K as the criteria for stable alloys, while alloys with $T_{\mathrm{CM}} \geq 1000$ K are regarded as unstable alloys. It is noted that when calculating T_{CM} for some of the 2D TMDC alloys, such as $Mo_{(1-x)}W_xS_2$, due to the limitation of calculations, their T_{CM} can exhibit negative values because those alloys are already stable at 0 K. But these unphysical values do not affect the results because they can also be categorized into stable alloys.

Analyses of the electronic structures of the alloys were performed. We examined the density of states (DOS) associated with the adsorption sites to understand the interaction between hydrogen and the adsorption site on the alloy configurations. According to the widely employed d-band center model of Hammer and Nørskov,[58] the ΔG_{H} of hydrogen adsorption can be analyzed using the d-band center (ε_{d}) of the catalyst. Also, previous studies demonstrate that the p-band center of chalcogen atoms in TMDCs can be used as a descriptor for ΔG_{H}.[13,59] In our case, since H interacts directly with the chalcogen atoms, a p-band center should be considered and is calculated as follows:

$$\varepsilon_{\mathrm{p}} = \frac{\int_{-\infty}^{+\infty} E D_{\mathrm{p}}(E)\mathrm{d}E}{\int_{-\infty}^{+\infty} D_{\mathrm{p}}(E)\mathrm{d}E} \quad (5)$$

where E is the energy level and $D_{\mathrm{p}}(E)$ refers to the p states in catalysts.

To analyze the changes in ε_{p} during the formation process of TMDC alloys, we decomposed the formation process into three steps (volume deformation (VD), charge exchange (CEX), and structural relaxation (SR)) and the composition-weighted average of ε_p for alloy $M1_{(1-x)}M2_xX_2$ at each step is calculated as follows:

$$\varepsilon_{\mathrm{i}} = (1 - x)\varepsilon_{\mathrm{i,M1X_2}} + x\varepsilon_{\mathrm{i,M2X_2}} \quad (6)$$

where x, i are the concentration of alloy and i represents step VD, CEX, or SR. $\varepsilon_{\mathrm{i,M1X_2}}$ and $\varepsilon_{\mathrm{i,M2X_2}}$ are the ε_p of two base lattices at step i.

2.3 Construction of the machine learning (ML) model

The selection of the ML model is a key process to achieve high efficiency of the predictions. In this study, we used an automated machine learning package, TPOT,[60] to evaluate the performance of different machine-learning algorithms and to find the best modeling pipeline for the dataset. The generation size and population size for running TPOT were set as 5 and 20, respectively. 5-Fold cross-validation was used when evaluating pipelines. We used regression to predict the ΔG_{H} of 2D TMDC alloys, while classification was chosen for the prediction of their stabilities. For ΔG_{H} prediction, as recommended by TPOT, we chose the pipeline using gradient boosting regression as implemented in the Scikit-learn Python library.[61] For stability prediction, the pipeline suggested by TPOT contains three estimators, *i.e.*, stacking estimator, recursive feature elimination, and random forest classifier.[60] Detailed information of the ML models can be found in the code provided in this work.

3. Results

3.1 Machine learning (ML) workflow

A typical monolayer TMDC contains metals in groups IV (Ti, Zr, Hf), V (V, Nb, Ta), and VI (Cr, Mo, W) and chalcogens (S, Se).[62] In this study, we first explored HER activity of the basal plane of a series of 2H–MX_2 alloys (M = Cr, Mo, W, V, Nb, Ta, Ti, Zr, Hf; X = S, Se). The template of hydrogen adsorption configuration was constructed by putting one hydrogen atom on top of the chalcogen site, as shown in Fig. 1a. After enumeration at each transition metal site, we obtained 8496 unique adsorption configurations of cation-mixed ternary alloys based on 72 alloy combinations across 18 TMDC pristine structures. Aiming at solving the complexity of these calculations, a ML workflow was developed to predict the hydrogen adsorption energies of TMDC alloys, as shown in Fig. 1b. First, DFT calculations were performed on more than 1140 adsorption configurations. For each alloy, at least two configurations were randomly selected per concentration to ensure that the data set can represent the properties of TMDC alloys in the whole concentration range.

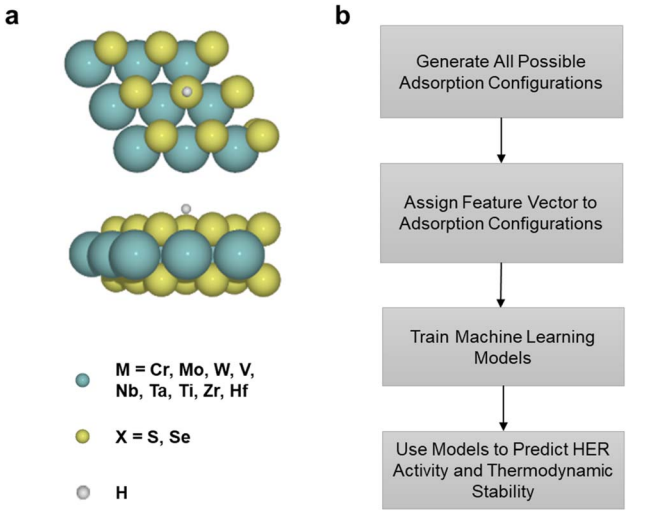

Fig. 1 (a). Template adsorption configuration of hydrogen on a 2D TMDC sheet, where transition metal (M = Cr, Mo, W, V, Nb, Ta, Ti, Zr or Hf) and chalcogen atoms (X = S, Se) are indicated by cyan and yellow spheres, while the hydrogen atom is shown in white. (b) Workflow for predicting ΔG_{H} and thermodynamic stability *via* machine learning.

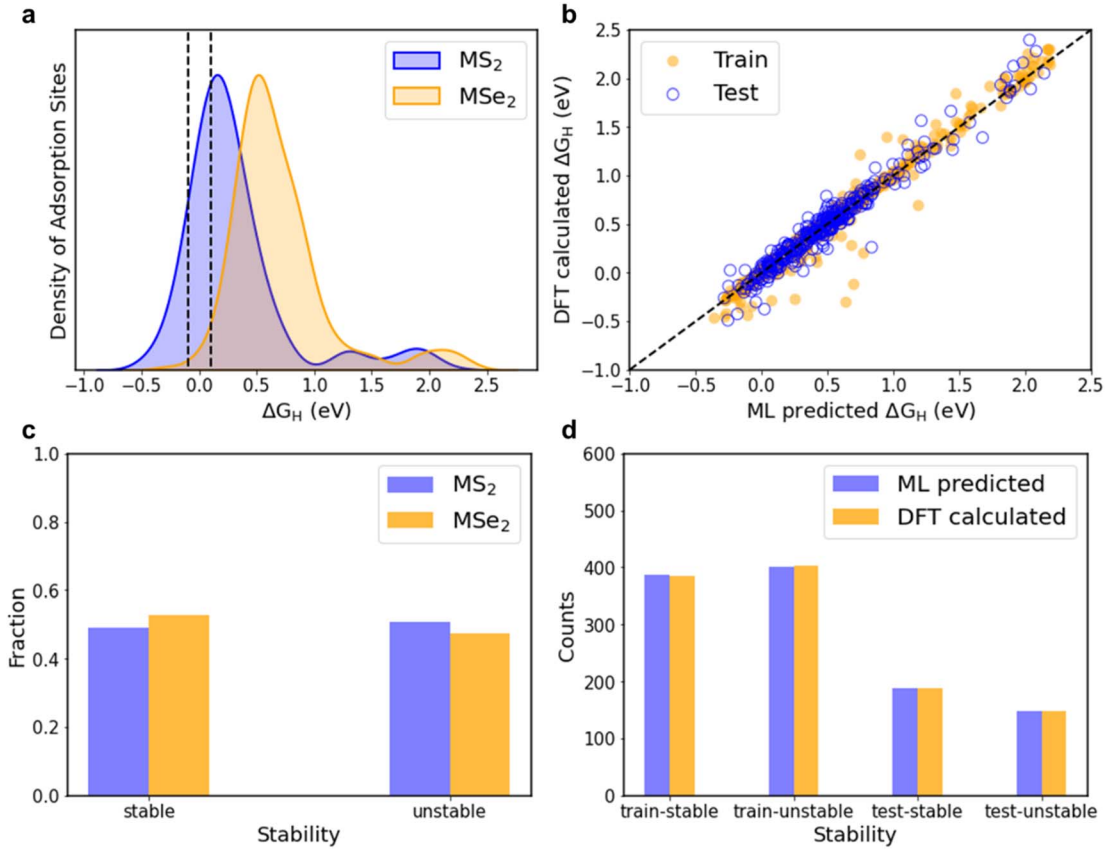

Fig. 2 (a). The normalized distribution of the DFT-calculated ΔG_H for hydrogen adsorption on MS_2 and MSe_2 respectively. Dashed lines indicate the 0.1 eV range around the optimal ΔG_H value of 0 eV. (b) Prediction performance of the training set and the test set for ΔG_H prediction. (c) The fraction of the DFT-calculated stable and unstable alloys for MS_2 and MSe_2 respectively. (d) Prediction performance of the training set and the test set for stability prediction.

The normalized distribution of the DFT-calculated Gibbs free energy (ΔG_H) values of hydrogen adsorption is plotted in Fig. 2a, from which some initial understanding of the HER performance of TMDC alloys can be obtained, as elaborated below. Generally, to achieve good catalytic performance towards the HER, hydrogen adsorption should exhibit ΔG_H close to the thermoneutral value near zero. As seen from Fig. 2a, a good portion of sites offered by TMDC alloys exhibit ΔG_H values close to the optimal $\Delta G_H = 0$ eV. Here, we select the range of $|\Delta G_H| \leq 0.1$ eV to represent the close vicinity of $\Delta G_H = 0$ eV, as the ideal region for effective HER. As demonstrated by the DFT results in Fig. 2a, there is a notable proportion of sites falling into this ΔG_H range. Meanwhile, we can see that these sites are predominantly from the transition metal disulfide alloys (MS_2), rather than the transition metal diselenide alloys (MSe_2). This indicates better potential for the HER from MS_2 than MSe_2. We also note from Fig. 2a that the shape of the ΔG_H distribution of MS_2 is similar to that of MSe_2. This similarity suggests that alloying may induce a similar effect on the HER activity of MS_2 and MSe_2.

We then assigned a feature vector to each TMDC alloy to represent the alloy structure numerically, thus relating the physical and chemical properties to the corresponding structures and using the data we collected to train the ML model. Informed by previous material informatics studies,[38,40,63,64] three of the most relevant properties for predicting ΔG_H of TMDCs were selected, *i.e.*, the atomic number (Z), the Pauling electronegativity (χ), and the Gibbs free energy of hydrogen adsorption on pristine MX_2 (ΔG_{H-MX2}). Here, Z represents the atom type and accounts for the steric effect, and χ accounts for the electronic affinity effect, and its value was obtained from the Mendeleev database.[65] ΔG_{H-MX2} serves as the benchmark reference and was used to estimate the variation in the Gibbs free energy. The ΔG_{H-MX2} of 18 pristine TMDCs were calculated by DFT calculations (see Table S1 in ESI†). Transition metal sites in the supercell were numbered and each site was described by these three properties. In addition, to account for the effect of the chalcogen atoms, Z and χ of S or Se were appended at the end of the feature vector. Since our supercell has 9 transition metal sites and each transition metal has 3 descriptors, a 29-dimensional (27 for transition metals and 2 for chalcogen atoms) vector was created to represent each TMDC alloy (see Section S4 ESI† for more details).

We created our ML model using gradient boosting regression, as selected by the automated ML package, TPOT.[60] All the available DFT data were used to train the ML model. The whole data set was randomly split as training and testing sets according to a 7 : 3 ratio. The distribution of the data in the

training and testing set is presented in Fig. S1 (see the ESI†). To ensure that we have sufficient data for good prediction, the learning curves of our ML model are also presented, in Fig. S3 (see the ESI†). The accuracy of our model prediction is illustrated in the prediction performance plot Fig. 2b, where the error in prediction is indicated by the deviation from the $X = Y$ line. We see that the training set of our model is well fitted, and the test set shows slightly larger deviations than the training set, but still small. As demonstrated, our model successfully predicts ΔG_H values of the HER on TMDC alloys, with a mean absolute error of 0.1 eV for the test set.

Additionally, to discover whether the abovementioned alloy configurations are synthesizable, we further investigated the thermodynamic stability of those 2D TMDC alloys using the ML workflow, similar to that of ΔG_H prediction. However, unlike the ΔG_H prediction workflow which used a regression model to predict continuous valued output, we used a classification pipeline suggested by TPOT to determine whether the alloy configuration is stable or not (see details in the Methods section). The criterion of a stable 2D TMDC alloy is the critical temperature for complete miscibility $T_{CM} < 1000$ K. Fig. 2c plots the distribution of the stability of MS_2 and MSe_2 alloys based on the DFT-calculated data, showing that 49.2% of the MS_2 alloys and 52.7% of the MSe_2 alloys are regarded as stable. We then did the feature vector assignment to each TMDC alloy, similar to the ΔG_H prediction process, except that we switched one material property, *i.e.*, ΔG_{H-MX2}, to the total energy of pristine MX_2 (E_{MX2}) (see Section S4 ESI† for more details) considering the physical properties of our target. During ML training, again, the whole data set was randomly split as training and testing sets according to a 7 : 3 ratio, with data distribution plots presented in Fig. S2.† The learning curves of the ML model show that we have enough data to reach the desired accuracy of the stability prediction (see Section S2 in the ESI† for details). From the prediction performance plot presented in Fig. 2d, we see that our model successfully predicts the stability of TMDC alloys, with an accuracy of 89.3% for the testing set.

3.2 Predictions of electrocatalytic activity of TMDC alloys

The HER activities of various TMDC alloys have been predicted by the ML framework (see Fig. 3 and Fig. S5–S10 in the ESI†). Among these TMDC alloys, only a small fraction has been investigated in experiments,[17,66,67] notably the $Mo_{(1-x)}W_xS_2$ and $Mo_{(1-x)}W_xSe_2$ groups. Experimentally, these two groups of TMDC alloys have been shown to be active for the HER at the edge sites, while it was believed that the basal planes of $Mo_{(1-x)}W_xS_2$ and $Mo_{(1-x)}W_xSe_2$ stay inert.[17,19,23–25] As shown in Fig. 3a and b, all adsorption sites at basal planes of $Mo_{(1-x)}W_xS_2$ and $Mo_{(1-x)}W_xSe_2$ alloys have ΔG_H values much higher than the ideal range for the HER (*i.e.*, $|\Delta G_H| \leq 0.1$ eV) and can be regarded as inert, in good agreement with the experimental findings. Nonetheless, despite the overall basal inertness of $Mo_{(1-x)}W_xS_2$ and $Mo_{(1-x)}W_xSe_2$, we see obviously lower ΔG_H for some alloy configurations compared to those of the corresponding pristine TMDCs. This illustrates an apparent effect of alloying in reducing ΔG_H below those values from a simple interpolation of the ΔG_H of their pristine structures.

Moving to other TMDC alloy groups, a similar effect of alloying-induced reduction in ΔG_H is also observed, which to the best of our knowledge has not yet been studied previously in experiments. Representative plots are shown in Fig. 3c–h. Some of the MS_2 alloys, such as $Mo_{(1-x)}V_xS_2$, $W_{(1-x)}Hf_xS_2$, and $Ti_{(1-x)}Zr_xS_2$, have ΔG_H lies in the ideal region for the HER over a wide range of concentrations, outperforming their pristine structures in electrocatalytic activities (see Fig. S5–S10 in ESI† for other TMDC alloys). For example, as the ΔG_H *vs.* concentration plot is shown in Fig. 3c, $Mo_{(1-x)}V_xS_2$ with the V concentration in the range of 0.4–0.8 yields ΔG_H close to zero. Particularly worth noting (see Fig. 3c) is that the introduction of a small fraction (*e.g.*, 11%) of V atoms into MoS_2 results in a significant drop in ΔG_H, showing a similar effect in enhancing the HER performance at the basal plane as the single-atom catalysis.[68] Meanwhile, the MSe_2 alloys are showing quite similar behaviors to their MS_2 counterparts but exhibiting relatively higher ΔG_H values, resulting in less favorable results for the HER in most cases. In addition, we also note from the results that in general the effect of alloying tends to be more significant for a TMDC alloy when the difference of ΔG_H values of its corresponding pristine TMDCs is more pronounced (see, *e.g.*, Fig. 3e where MoS_2 and VS_2 respectively have ΔG_H values of 2.00 eV and 0.27 eV), possibly because of more significant charge redistribution between atoms in those alloys as elaborated later.

We also compared the electrocatalytic performance of various TMDC alloys by calculating their respective fractions of enumerated adsorption sites that have optimal ΔG_H values ($|\Delta G_H| \leq 0.1$ eV) over the whole concentration range, with the mapping presented in Fig. 4a and Fig. S11a (see the ESI†) for MS_2 and MSe_2 alloys, respectively. By comparing two active site mappings, we see that the sites of optimal ΔG_H values are predominantly from MS_2 alloys, while HER active sites on MSe_2 alloys are negligible. This observation agrees well with the DFT calculated results in Fig. 2a. Therefore, we narrow down our exploration for HER catalysts to only MS_2 alloys. Also noted from Fig. 4a is that MS_2 alloys with metal atoms exclusively from group-VI elements provide no effective adsorption sites, thus predicted to be undesirable toward the HER. For cases involving elements from group-IV and/or group-V, some MS_2 alloys show promising performance towards the HER. As seen in Fig. 4a, a number of alloys exhibit a high fraction of effective adsorption sites, with a fraction value close to or beyond 50%. Examining these alloys in detail, we note that some high fraction values, *e.g.*, those of $Hf_{(1-x)}Nb_xS_2$ (71%) and $Ti_{(1-x)}V_xS_2$ (69%), may be attributed to their base metal disulfides, *i.e.*, NbS_2 and TiS_2, being active towards the HER. Nonetheless, there are also cases among those where their base metal disulfides are inactive, *e.g.*, $Zr_{(1-x)}V_xS_2$ (62%) and $Zr_{(1-x)}Ta_xS_2$ (47%). It is also interesting to note that the benefit of alloying in enhancing HER activity in general is most significant when there are mixtures of metal elements from different groups.

We further screened the MS_2 alloys by considering their stability, with the fractions of stable active sites for each MS_2

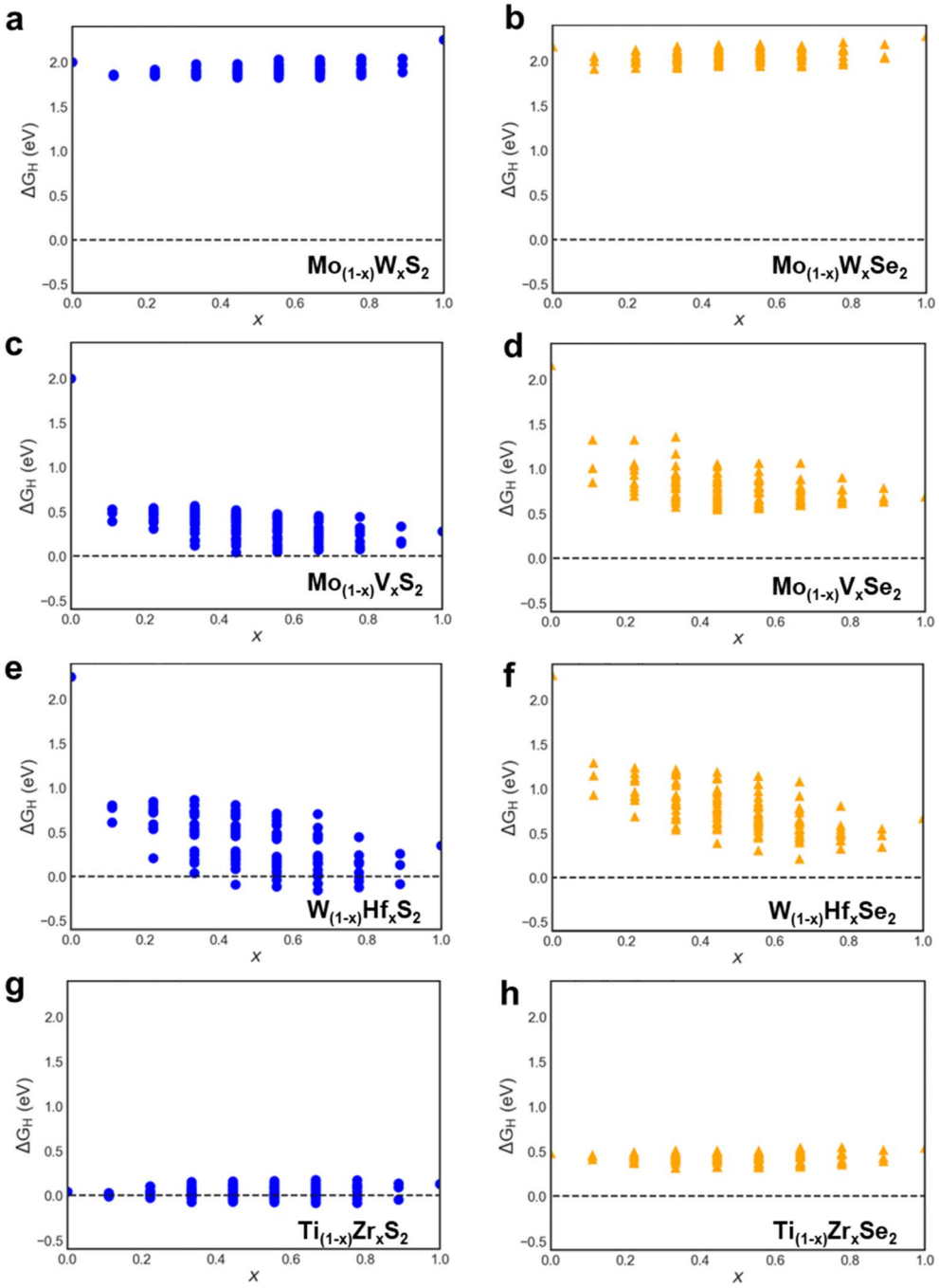

Fig. 3 (a–h) Machine learning predictions for ΔG_H of 8 representative TMDC ternary alloys at different concentrations. Black dash lines indicate the optimal ΔG_H to achieve the best HER performance.

alloy presented in Fig. 4b. We see that around half of the active adsorption sites have been filtered, with the number of active adsorption sites on MS_2 dropping significantly from originally 1133 to only 448. This observation agrees well with the DFT-calculated results presented in Fig. 2c. Although from Fig. 4a we have learned that MS_2 alloys with metal atoms from different groups result in enhanced HER activities, those alloys may suffer from low stability probably due to the large lattice mismatch between their base metal disulfides. For the extreme cases, *i.e.*, alloying between elements from group-IV and group-VI, almost all MS_2 alloys are found to be unstable, leading to nearly no effective adsorption sites in this region as shown by the mapping from Fig. 4b. Combining the effects of both basal plane activation and stability, in general, alloying between metal elements from group-IV and group-V is found to be most effective towards the HER. In addition, in order to guide the synthesis of MS_2 alloys, the fractions of stable active sites for each MS_2 alloy at different concentrations are provided (see

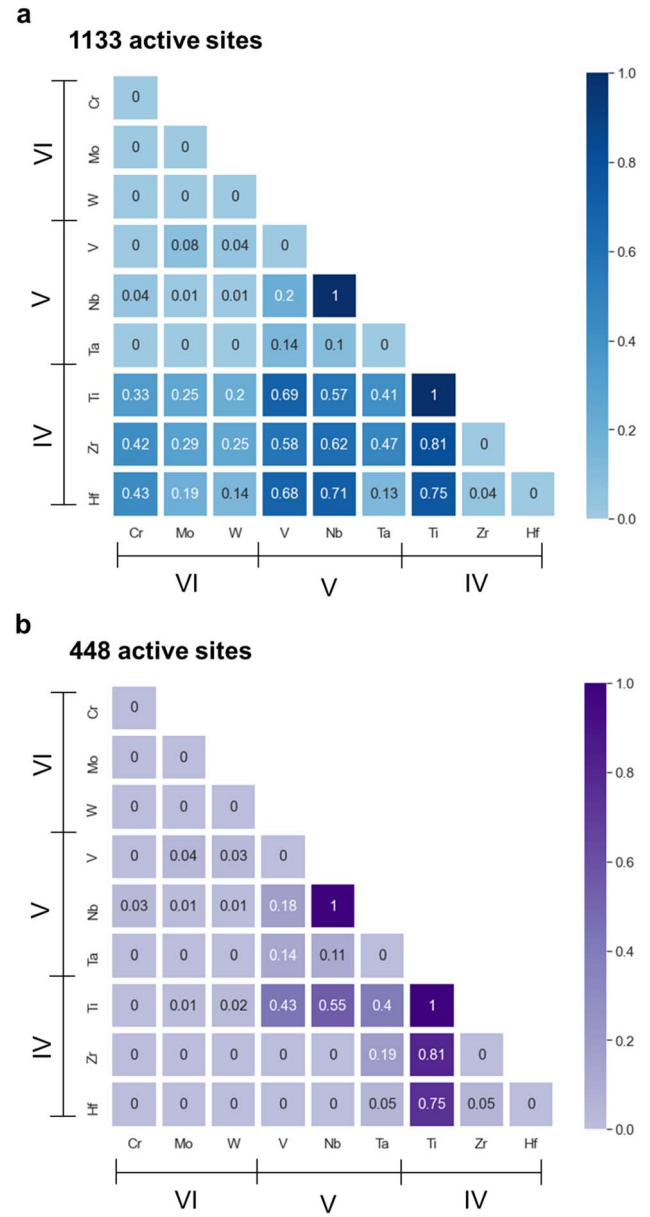

Fig. 4 HER activity heatmaps for MS_2 alloys showing (a). The fraction of adsorption sites of ΔG_H within the optimal range of (−0.1, 0.1) eV for each alloy, and (b). The fraction of adsorption sites that are both stable and have optimal ΔG_H for each MS_2 alloy. The number indicates the fraction value an alloy exhibits, which is also reflected by the color.

Fig. S12 in the ESI†). Previous studies on TMDCs mostly focus on alloying within group-VI elements while other cation-mixed alloys have been largely neglected. The results we presented here clearly suggest the necessity of expanding the exploration into other element groups in the design and development of TMDC alloys towards the HER.

3.3 The origin of alloying in enhancing HER activity

The results in Fig. 3 and 4 have clearly demonstrated the promising role of alloying in enhancing HER activity. To better utilize the information for rational TMDC alloy design towards the HER, it is important to understand the underlying mechanism. As suggested by previous studies,[69,70] the influence of alloying on surface adsorption properties may be described by either electronic (or ligand) and/or geometrical (or ensemble) effects. Since hydrogen always adsorbs on the chalcogen atom (S or Se) in a nearly identical geometrical fashion, effectively there is no geometrical effect in our study (see Fig. S16 in the ESI†). Therefore, it is expected that the influence of alloying on TMDCs would be dominated by the electronic effect resulting from the variation in chemical composition.

The electronic effect can be analyzed by examining the intrinsic electronic structures. In particular, we found that variations in chemical composition in TMDC alloys can significantly modify the p states of chalcogen atoms. One good metric to quantify such modification in the p state is the p-band center (denoted as ε_p in the following) of the chalcogen atom (see details in the Methods section). The band centers, *e.g.*, d-band and p-band centers, have also been used by previous studies[13,59] in describing characteristics of electronic structures. In Fig. 5a, we plotted the ΔG_H values of 140 randomly selected TMDC alloys *versus* their corresponding ε_p values. Overall we see a good linear correlation between ΔG_H and ε_p, for both the MS_2 and MSe_2 alloy groups, despite some outliers. Those outlier data points might be partially attributed to the fact of the p bands of chalcogen atoms are broad and in irregular shapes,[71] which consequently causes variation in ε_p. Nevertheless, the overall trends observed in Fig. 5a confirm that the electronic effect is related to the electronic structure of the adsorption (chalcogen) site, and a chalcogen site with a higher (lower) p-band center exhibits stronger (weaker) affinity to hydrogen. Meanwhile, we see that the linear fitting of the $\Delta G_H - \varepsilon_p$ data yields similar slopes for both MS_2 and MSe_2 alloy groups. Since the slope is indicative of the strength of the electronic effect,[71] this suggests that alloying results in a similar degree of electronic effect for both MS_2 and MSe_2, consistent with our previous observations in Fig. 2a. In addition, to consider the influence of configuration changes on the electronic structures, using $W_{(1-x)}V_xSe_2$ alloy as an example, we plotted the ΔG_H values of all the possible alloy configurations of $W_{(1-x)}V_xSe_2$ *versus* their corresponding ε_p values, as shown in Fig. S13 (in the ESI†). It is found that the variations in ε_p by the configuration changes still follow the trend observed in Fig. 5a. In accordance with the results shown in Fig. 5a, we then proceeded to examine how alloying modifies the p-band center. Fig. 5b plots the ε_p of the adsorption sites with the best (lowest) ΔG_H at each concentration for two representatives, *i.e.*, $Mo_{(1-x)}V_xS_2$ and $W_{(1-x)}V_xSe_2$ (see ε_p *vs.* x plots for all adsorption sites in $Mo_{(1-x)}V_xS_2$ and $W_{(1-x)}V_xSe_2$ in Fig. S14†). Alloying is observed to have the apparent effect of enhancing ε_p from that of the pristine base metal disulfide/diselenide. The increased ε_p indicates the shifting of p bands to the Fermi energy, which stabilizes hydrogen adsorption and therefore leads to the enhanced HER activity of alloys.

As suggested by previous studies,[57,72,73] a more in-depth understanding of the effect of alloying on ε_p can be obtained by analyzing the changes in the properties of TMDC alloys during their formation processes. In particular, the formation process can be decomposed into three steps, namely (i) volume

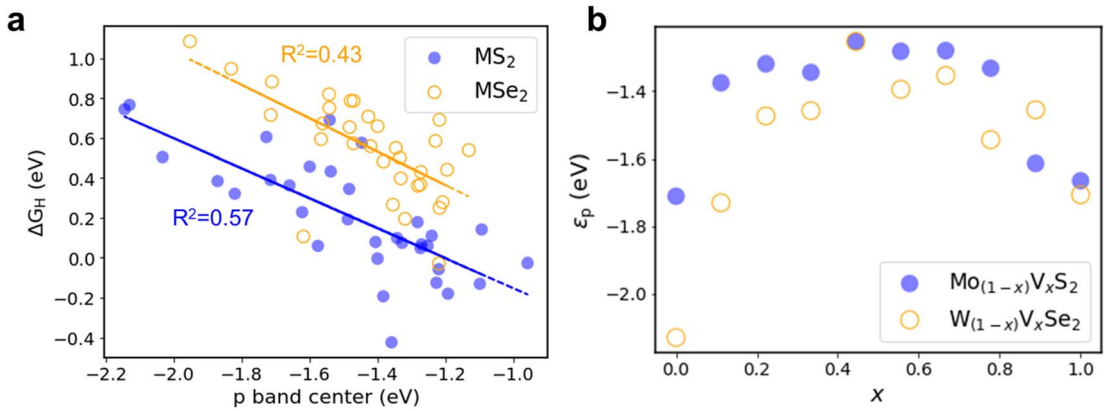

Fig. 5 (a). Relationship between ε_p values and ΔG_H for MS_2 and MSe_2. (b) ε_p of the adsorption sites in $Mo_{(1-x)}V_xS_2$ and $W_{(1-x)}V_xSe_2$ alloys with the best (lowest) ΔG_H as a function of the composition (*i.e.*, atomic fraction of V).

deformation (VD), where the corresponding pristine TMDC was compressed or expanded from its equilibrium lattice constant to the alloy lattice constant, (ii) charge exchange (CEX), where mixing of atoms takes place and introduced into the unrelaxed alloy lattice constructed in step (i), and subsequently (iii) structural relaxation (SR) where the alloy lattice obtained in step (ii) was fully relaxed. The composition-weighted average of ε_p was calculated (see the Computational methods section for details) at each step for different TMDC alloys. Fig. 6a shows the evolution of ε_p for six representative alloys, considering different concentrations, transition metals, and chalcogen types. We note that the changes in ε_p in the VD and SR steps are negligible in some cases, *e.g.*, $W_{0.22}V_{0.78}S_2$, but significant in other cases, *e.g.*, $W_{0.22}V_{0.78}Se_2$. However, an interesting

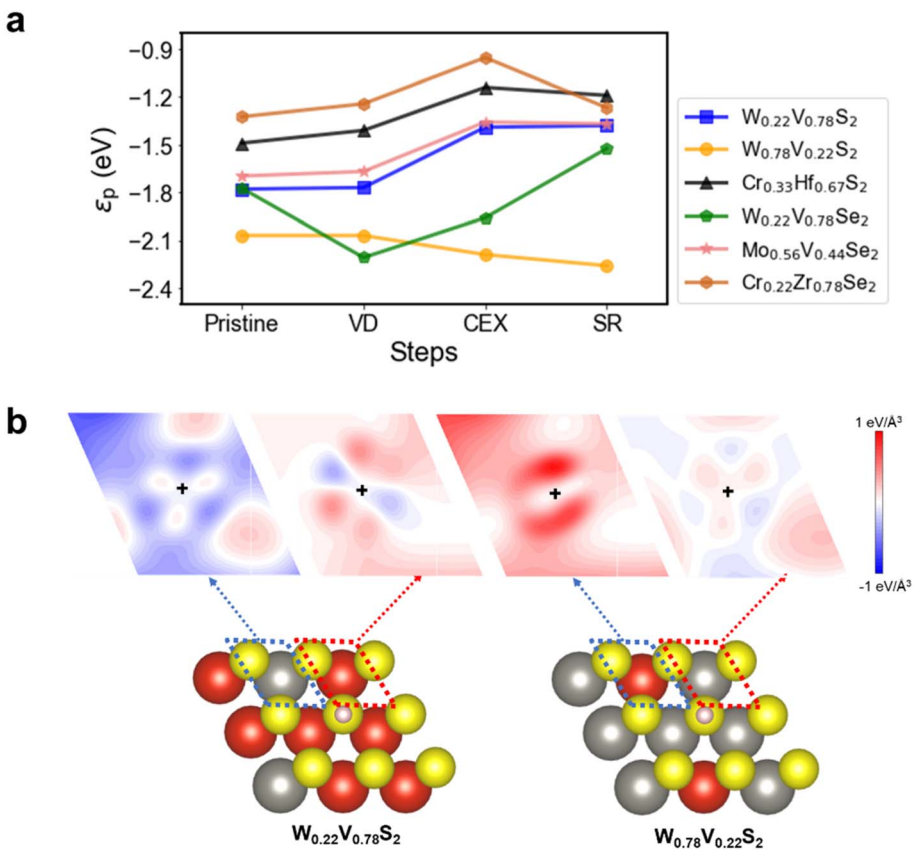

Fig. 6 (a). The evolution of composition-weighted average of ε_p at the VD, CEX, and SR steps. (b) Representative hydrogen adsorption configurations on $W_{0.22}V_{0.78}S_2$ and $W_{0.78}V_{0.22}S_2$ alloys, with related charge density difference maps of metal atoms at the vicinity of the adsorption site. Atoms are colored with V (red), W (grey), S (yellow), and hydrogen (pink). The black cross symbols in charge density difference maps indicate the center of the metal atom (W or V).

observation from the ε_p evolution results is that in the cases where the changes in ε_p in the VD and SR steps are significant, the changes in the two steps are always in the opposite directions and thus cancel each other, *e.g.*, the cases of $W_{0.22}V_{0.78}Se_2$ and $Cr_{0.22}Zr_{0.78}Se_2$. As a result, the CEX step effectively becomes the critical step responsible for the overall change in ε_p. Additionally, we investigated the changes in ε_p in the CEX step ($\Delta\varepsilon_{CEX}$) with the concentration x using the adsorption sites with the best (lowest) ΔG_H at each concentration for $Mo_{(1-x)}V_xS_2$ and $W_{(1-x)}V_xSe_2$. As shown in Fig. S15,† $\Delta\varepsilon_{CEX}$ varies with the x, and this trend is similar to that shown in Fig. 5b, further proving that the CEX step is indeed the critical step responsible for the overall change in ε_p.

With a more focused analysis of the CEX step, we found that the change in ε_p during this step is strongly correlated with the charge transfer around the adsorption site. In particular, depending on the charge transfer direction, negative or positive change in ε_p is expected. Such correlation is elaborated below using $W_{0.22}V_{0.78}S_2$ and $W_{0.78}V_{0.22}S_2$ systems as examples, where the Bader[74] charge transfer analysis and charge density difference mapping (see Fig. 6b) were performed. In the case of $W_{0.22}V_{0.78}S_2$, on average each V atom has a charge accumulation of about 0.02e while each W atom shows a charge depletion of about 0.06e. The electron depletion and accumulation around W and V atoms, respectively, are also clearly revealed in the charge density different map shown in Fig. 6b, indicative of the electron transfer from W to V. The electron gain at V shifts its d-band center away from the Fermi level to lead to weaker interaction between V and S atoms,[69,75] which in return shifts the p-band center of the S atom towards the Fermi level and therefore strengthens the H–S bond.[76] That is, charge gain at a transition metal would lead to an increased ε_p for the nearby chalcogen atoms. For instance, if we consider the case of hydrogen adsorption on the chalcogen atom surrounded by three V atoms, as shown in Fig. 6b, the chalcogen atom undergoes a substantial increase in ε_p (*i.e.*, $\Delta\varepsilon_p = 0.38$ eV). Similar charge transfer and redistribution behaviors are also observed for the $W_{0.78}V_{0.22}S_2$ system, where each V atom and W atom on average show charge gain and depletion of 0.04e and 0.02e, respectively. If we examine the case of hydrogen adsorption on the chalcogen atom surrounded by three W atoms, the chalcogen atom exhibits a decrease in ε_p (*i.e.*, $\Delta\varepsilon_p = -0.19$ eV) due to charge depletion from the surrounding W atoms. Since in $W_{0.22}V_{0.78}S_2$ and $W_{0.78}V_{0.22}S_2$, the hydrogen adsorbed chalcogen is more likely to be surrounded by V (with charge gain) and W (with charge depletion) respectively, the composition-weighted average of ε_p is expected to increase and decrease during the CEX step, respectively, in agreement with the trends shown in Fig. 6a.

From the above, we see that charge transfer is a critical factor affecting HER activity. Consequently, it also plays a large role in prescribing the local chemical ordering of adsorption sites yielding optimal HER activities. Taking $W_{(1-x)}V_xS_2$ as an example, since local charge transfer occurs from W to V, it can thus be hypothesized that hydrogen adsorption would prefer a chalcogen site with more V presence in its surrounding to achieve a charge accumulation center for maximized ε_p enhancement, and thus lower ΔG_H. Fig. 7 shows several representatives of the adsorption configurations offering the lowest ΔG_H at different compositions, alongside the corresponding ΔG_H values as the composition x varies, obtained from DFT calculations, which corresponds quite well with our hypothesis.

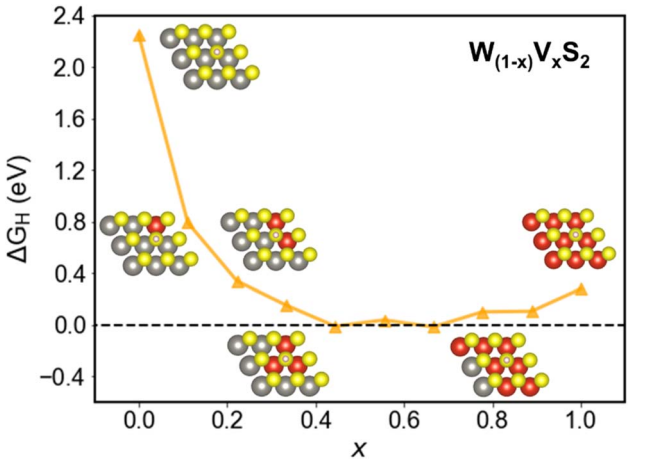

Fig. 7 Lowest ΔG_H values for hydrogen adsorption on $W_{(1-x)}V_xS_2$ as the composition (x) varies. The insets illustrate sample adsorption configurations yielding the lowest ΔG_H. V, W, S, and hydrogen atoms are shown in red, grey, yellow, and pink respectively.

4. Conclusions

In summary, first-principles DFT calculations in conjunction with ML were performed to investigate the possibility of basal plane activation in 2D TMDCs *via* alloying for the hydrogen evolution reaction (HER). An ML workflow was created based on the results from 1000+ DFT calculations, and feature vectors were then assigned accounting for the physical and chemical properties of both transition metals and chalcogen atoms. The Gibbs free energy of hydrogen adsorption ΔG_H and the stability of a series of 2D cation-mixed TMDC alloys of various compositions were shown to be accurately predicted by the ML model. Our results demonstrated that alloying leads to a substantial reduction in ΔG_H on the basal plane of TMDCs. For a sizable fraction of the TMDC alloys examined, alloying was found to render ΔG_H in the ideal range for the HER. Moreover, for those TMDC alloys exhibiting good HER activity, half of them were found to be stable, confirming the thermodynamic feasibility of the alloying route for basal plane activation.

The alloying effect was found to be mostly dependent on the transition metals, while not much affected by the chalcogen atom type. It was then further revealed to mainly originate from the electronic effect resulting from the chemical composition variation. In particular, the electronic effect comes from p-band center shifting of the adsorption sites, attributed to the alloying produced local charge exchange. The present study provides essential mechanistic insights for a new route towards basal plane activation of 2D TMDCs for enhanced electrocatalytic performance. The insights coupled with the ML model will enable machine-derived suggestions, crucial for developing an

efficient roadmap of rational design and exploration of 2D TMDC alloy catalysts for the HER.

Code availability

The code used to perform this work is available at **https://github.com/YiqingChen524/2D_TMDC_alloys**.

Author contributions

Y. C. and J. S. initiated the project. Y. C. conducted the first-principles calculations and created the machine learning workflow. Y. C., P. O., and Z. Y. analyzed the results. Y. C., P. O., and J. S. wrote the manuscript with inputs from all the co-authors.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (grant #: NSERC RGPIN-2017-05187), and the McGill Engineering Doctoral Award (MEDA). The authors would also like to acknowledge Compute Canada and Calcul Québec for providing computing resources.

References

1 M. S. Dresselhaus and I. L. Thomas, Alternative energy technologies, *Nature*, 2001, **414**(6861), 332–337.
2 Z. W. Seh, J. Kibsgaard, C. F. Dickens, I. Chorkendorff, J. K. Nørskov and T. F. Jaramillo, Combining theory and experiment in electrocatalysis: Insights into materials design, *Science*, 2017, **355**(6321), eaad4998.
3 S. Trasatti, Work function, electronegativity, and electrochemical behaviour of metals: III. Electrolytic hydrogen evolution in acid solutions, *J. Electroanal. Chem. Interfacial Electrochem.*, 1972, **39**(1), 163–184.
4 N. Cheng, S. Stambula, D. Wang, M. N. Banis, J. Liu, A. Riese, B. Xiao, R. Li, T. K. Sham, L. M. Liu, G. A. Botton and X. Sun, Platinum single-atom and cluster catalysis of the hydrogen evolution reaction, *Nat. Commun.*, 2016, **7**, 13638.
5 Y. Jiao, Y. Zheng, M. Jaroniec and S. Z. Qiao, Design of electrocatalysts for oxygen- and hydrogen-involving energy conversion reactions, *Chem. Soc. Rev.*, 2015, **44**(8), 2060–2086.
6 Y. Chen, K. Yang, B. Jiang, J. Li, M. Zeng and L. Fu, Emerging two-dimensional nanomaterials for electrochemical hydrogen evolution, *J. Mater. Chem. A*, 2017, **5**(18), 8187–8208.
7 S. Jayabal, G. Saranya, J. Wu, Y. Liu, D. Geng and X. Meng, Understanding the high-electrocatalytic performance of two-dimensional MoS2 nanosheets and their composite materials, *J. Mater. Chem. A*, 2017, **5**(47), 24540–24563.
8 X. Chia and M. Pumera, Layered transition metal dichalcogenide electrochemistry: journey across the periodic table, *Chem. Soc. Rev.*, 2018, **47**(15), 5602–5613.
9 A. Ambrosi, Z. Sofer and M. Pumera, 2H?> 1T phase transition and hydrogen evolution activity of MoS2, MoSe2, WS2 and WSe2 strongly depends on the MX2 composition, *Chem Commun.*, 2015, **51**(40), 8450–8453.
10 D. B. Putungan, S. H. Lin and J. L. Kuo, A first-principles examination of conducting monolayer 1T'-MX(2) (M = Mo, W; X = S, Se, Te): promising catalysts for hydrogen evolution reaction and its enhancement by strain, *Phys. Chem. Chem. Phys.*, 2015, **17**(33), 21702–21708.
11 Y. R. An, X. L. Fan, Z. F. Luo and W. M. Lau, Nanopolygons of Monolayer MS2: Best Morphology and Size for HER Catalysis, *Nano Lett.*, 2017, **17**(1), 368–376.
12 J. Zhu, Z. C. Wang, H. Dai, Q. Wang, R. Yang, H. Yu, M. Liao, J. Zhang, W. Chen, Z. Wei, N. Li, L. Du, D. Shi, W. Wang, L. Zhang, Y. Jiang and G. Zhang, Boundary activated hydrogen evolution reaction on monolayer MoS2, *Nat. Commun.*, 2019, **10**(1), 1348.
13 Y. Ouyang, C. Ling, Q. Chen, Z. Wang, L. Shi and J. Wang, Activating Inert Basal Planes of MoS2 for Hydrogen Evolution Reaction through the Formation of Different Intrinsic Defects, *Chem. Mater.*, 2016, **28**(12), 4390–4396.
14 B. Hinnemann, P. G. Moses, J. Bonde, K. P. Jørgensen, J. H. Nielsen, S. Horch, I. Chorkendorff and J. K. Nørskov, Biomimetic Hydrogen Evolution: MoS2 Nanoparticles as Catalyst for Hydrogen Evolution, *J. Am. Chem. Soc.*, 2005, **127**(15), 5308–5309.
15 H. Li, C. Tsai, A. L. Koh, L. Cai, A. W. Contryman, A. H. Fragapane, J. Zhao, H. S. Han, H. C. Manoharan, F. Abild-Pedersen, J. K. Norskov and X. Zheng, Corrigendum: Activating and optimizing MoS2 basal planes for hydrogen evolution through the formation of strained sulphur vacancies, *Nat. Mater.*, 2016, **15**(3), 364.
16 Y. Lei, S. Pakhira, K. Fujisawa, X. Wang, O. O. Iyiola, N. Perea Lopez, A. Laura Elias, L. Pulickal Rajukumar, C. Zhou, B. Kabius, N. Alem, M. Endo, R. Lv, J. L. Mendoza-Cortes and M. Terrones, Low-temperature Synthesis of Heterostructures of Transition Metal Dichalcogenide Alloys (WxMo1-xS2) and Graphene with Superior Catalytic Performance for Hydrogen Evolution, *ACS Nano*, 2017, **11**(5), 5103–5112.
17 H. Wang, L. Ouyang, G. Zou, C. Sun, J. Hu, X. Xiao and L. Gao, Optimizing MoS2 Edges by Alloying Isovalent W for Robust Hydrogen Evolution Activity, *ACS Catal.*, 2018, **8**(10), 9529–9536.
18 S. M. Tan and M. Pumera, Composition-Graded MoWSx Hybrids with Tailored Catalytic Activity by Bipolar Electrochemistry, *ACS Appl. Mater. Interfaces*, 2017, **9**(48), 41955–41964.
19 M. Zhuang, L.-Y. Gan, M. Zou, Y. Dou, X. Ou, Z. Liu, Y. Ding, I. H. Abidi, A. Tyagi and M. Jalali, Engineering sub-100 nm Mo (1−x) W x Se 2 crystals for efficient hydrogen evolution catalysis, *J. Mater. Chem. A*, 2018, **6**(7), 2900–2907.
20 O. E. Meiron, V. Kuraganti, I. Hod, R. Bar-Ziv and M. Bar-Sadan, Improved catalytic activity of Mo 1− x W x Se 2

alloy nanoflowers promotes efficient hydrogen evolution reaction in both acidic and alkaline aqueous solutions, *Nanoscale*, 2017, **9**(37), 13998–14005.

21 J. Greeley and M. Mavrikakis, Alloy catalysts designed from first principles, *Nat. Mater.*, 2004, **3**(11), 810–815.

22 L. M. Xie, Two-dimensional transition metal dichalcogenide alloys: preparation, characterization and applications, *Nanoscale*, 2015, **7**(44), 18392–18401.

23 Q. Gong, L. Cheng, C. Liu, M. Zhang, Q. Feng, H. Ye, M. Zeng, L. Xie, Z. Liu and Y. Li, Ultrathin MoS2(1–x)Se2x Alloy Nanoflakes For Electrocatalytic Hydrogen Evolution Reaction, *ACS Catal.*, 2015, **5**(4), 2213–2219.

24 Q. Fu, L. Yang, W. Wang, A. Han, J. Huang, P. Du, Z. Fan, J. Zhang and B. Xiang, Synthesis and enhanced electrochemical catalytic performance of monolayer WS2 (1–x) Se2x with a tunable band gap, *Adv. Mater.*, 2015, **27**(32), 4732–4738.

25 V. Kiran, D. Mukherjee, R. N. Jenjeti and S. Sampath, Active guests in the MoS2/MoSe2 host lattice: efficient hydrogen evolution using few-layer alloys of MoS(2(1-x))Se(2x), *Nanoscale*, 2014, **6**(21), 12856–12863.

26 B. Hammer and J. K. Nørskov, Theoretical surface science and catalysis—calculations and concepts, *Adv. Catal.*, 2000, **45**, 71–129.

27 J. Schmidt, M. R. G. Marques, S. Botti and M. A. L. Marques, Recent advances and applications of machine learning in solid-state materials science, *npj Comput. Mater.*, 2019, **5**(1), 83.

28 K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev and A. Walsh, Machine learning for molecular and materials science, *Nature*, 2018, **559**(7715), 547–555.

29 K. Omata, Screening of New Additives of Active-Carbon-Supported Heteropoly Acid Catalyst for Friedel–Crafts Reaction by Gaussian Process Regression, *Ind. Eng. Chem. Res.*, 2011, **50**(19), 10948–10954.

30 A. Mannodi-Kanakkithodi, M. Y. Toriyama, F. G. Sen, M. J. Davis, R. F. Klie and M. K. Chan, Machine-learned impurity level prediction for semiconductors: the example of Cd-based chalcogenides, *npj Comput. Mater.*, 2020, **6**(1), 1–14.

31 C. T. Chen and G. X. Gu, Generative Deep Neural Networks for Inverse Materials Design Using Backpropagation and Active Learning, *Adv. Sci.*, 2020, 1902607.

32 C. Wang, K. Aoyagi, P. Wisesa and T. Mueller, Lithium Ion Conduction in Cathode Coating Materials from On-the-Fly Machine Learning, *Chem. Mater.*, 2020, **32**(9), 3741–3752.

33 X. Zhang, Y. Tian, L. Chen, X. Hu and Z. Zhou, Machine Learning: A New Paradigm in Computational Electrocatalysis, *J. Phys. Chem. Lett.*, 2022, **13**(34), 7920–7930.

34 L. Chen, X. Zhang, A. Chen, S. Yao, X. Hu and Z. Zhou, Targeted design of advanced electrocatalysts by machine learning, *Chin. J. Catal.*, 2022, **43**(1), 11–32.

35 L. Chen, Y. Tian, X. Hu, S. Yao, Z. Lu, S. Chen, X. Zhang and Z. Zhou, A Universal Machine Learning Framework for Electrocatalyst Innovation: A Case Study of Discovering Alloys for Hydrogen Evolution Reaction, *Adv. Funct. Mater.*, 2022, **32**(47), 2208418.

36 T. Toyao, Z. Maeno, S. Takakusagi, T. Kamachi, I. Takigawa and K.-I. Shimizu, Machine Learning for Catalysis Informatics: Recent Applications and Prospects, *ACS Catal.*, 2020, **10**(3), 2260–2297.

37 T. Toyao, K. Suzuki, S. Kikuchi, S. Takakusagi, K.-I. Shimizu and I. Takigawa, Toward effective utilization of methane: machine learning prediction of adsorption energies on metal alloys, *J. Phys. Chem. C*, 2018, **122**(15), 8315–8326.

38 L. Bassman, P. Rajak, R. K. Kalia, A. Nakano, F. Sha, J. Sun, D. J. Singh, M. Aykol, P. Huck, K. Persson and P. Vashishta, Active learning for accelerated design of layered materials, *npj Comput. Mater.*, 2018, **4**(1), 74.

39 M. O. J. Jäger, E. V. Morooka, F. Federici Canova, L. Himanen and A. S. Foster, Machine learning hydrogen adsorption on nanoclusters through structural descriptors, *npj Comput. Mater.*, 2018, **4**(1), 37.

40 K. Tran and Z. W. Ulissi, Active learning across intermetallics to guide discovery of electrocatalysts for CO2 reduction and H2 evolution, *Nat. Catal.*, 2018, **1**(9), 696–703.

41 K. T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K.-R. Müller and E. K. Gross, How to represent crystal structures for machine learning: Towards fast prediction of electronic properties, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2014, **89**(20), 205118.

42 F. Legrain, J. Carrete, A. van Roekeghem, S. Curtarolo and N. Mingo, How chemical composition alone can predict vibrational free energies and entropies of solids, *Chem. Mater.*, 2017, **29**(15), 6220–6227.

43 Z. Yang, W. Gao and Q. Jiang, A machine learning scheme for the catalytic activity of alloys with intrinsic descriptors, *J. Mater. Chem. A*, 2020, **8**(34), 17507–17515.

44 P. Schlexer Lamoureux, K. T. Winther, J. A. Garrido Torres, V. Streibel, M. Zhao, M. Bajdich, F. Abild-Pedersen and T. Bligaard, Machine Learning for Computational Heterogeneous Catalysis, *ChemCatChem*, 2019, **11**(16), 3581–3601.

45 A. Hjorth Larsen, J. Jørgen Mortensen, J. Blomqvist, I. E. Castelli, R. Christensen, M. Dułak, J. Friis, M. N. Groves, B. Hammer, C. Hargus, E. D. Hermes, P. C. Jennings, P. Bjerre Jensen, J. Kermode, J. R. Kitchin, E. Leonhard Kolsbjerg, J. Kubal, K. Kaasbjerg, S. Lysgaard, J. Bergmann Maronsson, T. Maxson, T. Olsen, L. Pastewka, A. Peterson, C. Rostgaard, J. Schiøtz, O. Schütt, M. Strange, K. S. Thygesen, T. Vegge, L. Vilhelmsen, M. Walter, Z. Zeng and K. W. Jacobsen, The atomic simulation environment—a Python library for working with atoms, *J. Phys.: Condens. Matter*, 2017, **29**(27), 273002.

46 M. Ångqvist, W. A. Muñoz, J. M. Rahm, E. Fransson, C. Durniak, P. Rozyczko, T. H. Rod and P. Erhart, ICET – A Python Library for Constructing and Sampling Alloy Cluster Expansions, *Adv. Theory Simul.*, 2019, **2**(7), 1900015.

47 G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1999, **59**(3), 1758–1775.

48 G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis

set, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1996, **54**(16), 11169–11186.

49 P. E. Blöchl, Projector augmented-wave method, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1994, **50**(24), 17953–17979.

50 J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.*, 1996, **77**(18), 3865–3868.

51 J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh and C. Fiolhais, Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1992, **46**(11), 6671–6687.

52 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, 2010, **132**(15), 154104.

53 S. Grimme, S. Ehrlich and L. Goerigk, Effect of the damping function in dispersion corrected density functional theory, *J. Comput. Chem.*, 2011, **32**(7), 1456–1465.

54 G. Rothenberg, *Catalysis: Concepts and Green Applications*, John Wiley & Sons, 2017.

55 C. Tsai, F. Abild-Pedersen and J. K. Nørskov, Tuning the MoS2 Edge-Site Activity for Hydrogen Evolution via Support Interactions, *Nano Lett.*, 2014, **14**(3), 1381–1387.

56 W. R. L. Lambrecht and B. Segall, Anomalous band-gap behavior and phase stability of c-BN–diamond alloys, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1993, **47**(15), 9289–9296.

57 J. Kang, S. Tongay, J. Li and J. Wu, Monolayer semiconducting transition metal dichalcogenide alloys: Stability and band bowing, *J. Appl. Phys.*, 2013, **113**(14), 143703.

58 B. Hammer and J. K. Nørskov, Electronic factors determining the reactivity of metal surfaces, *Surf. Sci.*, 1995, **343**(3), 211–220.

59 Y. Chen, P. Ou, X. Bie and J. Song, Basal plane activation in monolayer MoTe2 for the hydrogen evolution reaction via phase boundaries, *J. Mater. Chem. A*, 2020, **8**(37), 19522–19532.

60 R. S. Olson, N. Bartley, R. J. Urbanowicz and J. H. Moore, *Evaluation of a Tree-Based Pipeline Optimization Tool for Automating Data Science, Proceedings of the Genetic and Evolutionary Computation Conference 2016*, Association for Computing Machinery, Denver, Colorado, USA, 2016, pp. 485–492.

61 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss and V. Dubourg, Scikit-learn: Machine learning in Python, *J. Mach. Learn Res.*, 2011, **12**, 2825–2830.

62 M. Chhowalla, H. S. Shin, G. Eda, L.-J. Li, K. P. Loh and H. Zhang, The chemistry of two-dimensional layered transition metal dichalcogenide nanosheets, *Nat. Chem.*, 2013, **5**(4), 263–275.

63 Z. Li, S. Wang, W. S. Chin, L. E. Achenie and H. Xin, High-throughput screening of bimetallic catalysts enabled by machine learning, *J. Mater. Chem. A*, 2017, **5**(46), 24131–24138.

64 S. J. Davie, N. Di Pasquale and P. L. Popelier, Kriging atomic properties with a variable number of inputs, *J. Chem. Phys.*, 2016, **145**(10), 104104.

65 M. Lukasz, *Mendeleev–A Python Resource for Properties of Chemical Elements, Ions and Isotopes*, 2014.

66 Z. Hemmat, J. Cavin, A. Ahmadiparidari, A. Ruckel, S. Rastegar, S. N. Misal, L. Majidi, K. Kumar, S. Wang, J. Guo, R. Dawood, F. Lagunas, P. Parajuli, A. T. Ngo, L. A. Curtiss, S. B. Cho, J. Cabana, R. F. Klie, R. Mishra and A. Salehi-Khojin, Quasi-Binary Transition Metal Dichalcogenide Alloys: Thermodynamic Stability Prediction, Scalable Synthesis, and Application, *Adv. Mater.*, 2020, **32**(26), 1907041.

67 X. Gan, R. Lv, X. Wang, Z. Zhang, K. Fujisawa, Y. Lei, Z.-H. Huang, M. Terrones and F. Kang, Pyrolytic carbon supported alloying metal dichalcogenides as free-standing electrodes for efficient hydrogen evolution, *Carbon*, 2018, **132**, 512–519.

68 H. Duan, C. Wang, G. Li, H. Tan, W. Hu, L. Cai, W. Liu, N. Li, Q. Ji, Y. Wang, Y. Lu, W. Yan, F. Hu, W. Zhang, Z. Sun, Z. Qi, L. Song and S. Wei, Single-Atom-Layer Catalysis in a MoS2 Monolayer Activated by Long-Range Ferromagnetism for the Hydrogen Evolution Reaction: Beyond Single-Atom Catalysis, *Angew. Chem., Int. Ed.*, 2021, **60**(13), 7251–7258.

69 P. Liu and J. K. Nørskov, Ligand and ensemble effects in adsorption on alloy surfaces, *Phys. Chem. Chem. Phys.*, 2001, **3**(17), 3814–3818.

70 A. Ramstad, F. Strisland, T. Ramsvik and A. Borg, CO adsorption on the Pt/Rh(100) surface studied by high-resolution photoemission, *Surf. Sci.*, 2000, **458**(1), 135–146.

71 M. Mavrikakis, B. Hammer and J. K. Nørskov, Effect of Strain on the Reactivity of Metal Surfaces, *Phys. Rev. Lett.*, 1998, **81**(13), 2819–2822.

72 J.-H. Yang and B. I. Yakobson, Unusual Negative Formation Enthalpies and Atomic Ordering in Isovalent Alloys of Transition Metal Dichalcogenide Monolayers, *Chem. Mater.*, 2018, **30**(5), 1547–1555.

73 L. G. Ferreira, S.-H. Wei and A. Zunger, First-principles calculation of alloy phase diagrams: The renormalized-interaction approach, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1989, **40**(5), 3197–3231.

74 G. Henkelman, A. Arnaldsson and H. Jónsson, A fast and robust algorithm for Bader decomposition of charge density, *Comput. Mater. Sci.*, 2006, **36**(3), 354–360.

75 F. Gao and D. W. Goodman, Pd–Au bimetallic catalysts: understanding alloy effects from planar models and (supported) nanoparticles, *Chem. Soc. Rev.*, 2012, **41**(24), 8009–8020.

76 C. Tsai, K. Chan, J. K. Nørskov and F. Abild-Pedersen, Theoretical insights into the hydrogen evolution activity of layered transition metal dichalcogenides, *Surf. Sci.*, 2015, **640**, 133–140.