Deep-Learning-Assisted High-Throughput Discovery of Metallophilic MA_2Z_4 Nanomaterials

Peng Zhang,†[1] Xiang Feng,†[1] Dominik Legut,[2,3] Chao Lin,[1] Haohao Dong,[1] Zhihao Li,[1] Zhi Wei Seh,[4] Qianfan Zhang,[1]*

1 School of Materials Science and Engineering, Beihang University, Beijing 100191, P. R. China

2 IT4Innovations, VSB-Technical University of Ostrava, 17.listopadu 2172/15, 708 00 Ostrava, Czsch Republic

3 Department of Condensed Matter Physics, Faculty of Mathematics and Physics, Charles University, Ke Karlovu 3, 121 16 Prague 2, Czech Republic

*4 Institute of Materials Research and Engineering (IMRE), Agency for Science, Technology and Research (A*STAR), 2 Fusionopolis Way, Innovis #08-03, Singapore 138634, Republic of Singapore*

*Corresponding author: *qianfan@buaa.edu.cn.

†These authors contributed equally to this work.

Abstract

With the growing demand for advanced energy materials, finding suitable metallic electrode nanomaterials with strong metallophilicity remains challenging. Machine learning methods offer powerful tools to tackle this issue by enabling efficient exploration of the extensive compositional space and accurate modeling of complex interactions. We investigated the interactions between seven-layered MA_2Z_4 nanomaterials and eight metal atoms using computational simulations, a high-throughput workflow and multitask machine learning (MTL) to explore a vast compositional space. We built a comprehensive dataset of 2,592 MA_2Z_4 nanomaterials and identified 2,018 stable adsorption structures for training models. Using the MTL and crystal graph convolutional neural network (CGCNN), we achieved superior accuracy in predicting adsorption energy of nanomaterials compared to traditional Auto-ML models. MA_2Z_4 nanosheets with low-electronegativity A elements and highly-electronegativity Z elements exhibit high stability and metallophilicity, making them promising electrode nanomaterials. This study highlights the power of integrating MTL and CGCNN methodologies to accelerate the discovery and optimization of novel energy nanomaterials.

Keywords

Metallic electrode nanomaterials; Multitask machine learning framework; Crystal graph convolutional neural network; MA_2Z_4; High-throughput workflow.

1. Introduction

Energy materials, especially electrode materials, hold significant potential for electric vehicles and renewable energy storage.[1] Compared to intercalation electrodes, metal electrodes store and release charge through direct metal deposition and stripping, theoretically providing higher energy densit.[2] However, during charging, metal electrodes tend to form uneven surface structure distribution, leading to dendritic growth, which can cause "isolated metal deposits," resulting in capacity decay and posing safety risks like internal short circuits.[3-10] The interaction between metal electrodes and supporting scaffolds plays a crucial role in this process.[11] To enhance stability and promote uniform metal distribution, scaffolds are used to improve electrode performance by providing a uniform field effect, reducing local current density, and suppressing dendrite formation.[12] Consequently, developing metallophilic surface structures to suppress dendritic growth, lower nucleation overpotentials, and guide uniform deposition has become a key research direction.[13,14]

Layered nanomaterials, as supporting and protecting surface structures, are crucial in metal electrodes, mitigating the effects of volume expansion, and suppressing dendritic growth.[15-18] Among them, MA_2Z_4 materials are composed of transition metals in the middle layer, with the elements A and Z—belonging to the IIIA-VIA groups—distributed on both sides of the transition metal layer.[19,20] These seven-layered 2D nanomaterials have tunable electronic properties and chemical stability, making them suitable as scaffolds for electrode materials in applications requiring high cyclability and volume change resistance.[21,22] However, due to the large compositional space and difficulties in experimental synthesis, utilizing advanced computational simulation methods to screen and optimize these materials has become especially important.[23] These limitations frequently result in a restricted exploration of the extensive compositional space, thereby impeding the discovery of optimal configurations for electrode applications. Machine learning (ML) methods, due to their ability to predict material properties at near-first-principles accuracy but at much faster speeds, are accelerating the design of novel materials.[24] Recent studies have shown that ML algorithms can significantly accelerate the materials discovery process, enabling the identification of promising candidates for specific applications, such as batteries[25] and catalysts[26].

Traditional ML methods often struggle to capture the complex local and non-local interactions

within materials.[27,28] To address this, the Graph Neural Networks (GNNs) model,[29] a type of deep learning framework, can operate on graph-structured data without the need for manual feature engineering, allowing for accurate predictions of material properties such as formation energy, band gaps, and elastic moduli.[29,30] GNNs are particularly advantageous in capturing complex interactions within materials, employing the strengths of deep learning in addressing these challenges. Recent advancements in the GNNs framework have introduced bond-type embedding to better distinguish different chemical bond types and incorporated transformer architectures to more effectively capture long-range interactions in material structures[31-35]. Meng et al.[36] proposed the BNM-CDGNN model achieving state-of-the-art results in formation energy, band gap and adsorption energy. However, GNNs may overfit or underfit specific datasets due to their reliance on neighboring node features[37]—representing local atomic or structural properties. This limitation is particularly significant when predicting properties for extreme data points, as these often challenge the model's ability to capture broader trends[38,39] and multi-task learning (MTL) has been proposed.[40,41] Chen Liang et al.[42] developed a multi-task mixture density network (MTL-MDN) framework to predict adsorption energy (E_b) of CO on Cu-based single-atom alloy catalysts for CO_2 reduction which showed good generalization capabilities on limited datasets and significantly reduced overfitting. However, related methods have limitations when applied to the diverse elemental dataset of 2D materials, especially MA_2Z_4. Predicting their properties is challenging due to the vast compositional space and complex structural features of these materials, especially when dealing with unoptimized structures. This complexity makes it difficult to accurately predict E_b. Therefore, developing models to predict interaction energy in 2D materials, particularly MA_2Z_4, with improved accuracy and generalization, is key for advancing the material design.

Building upon the combined strengths of MTL and CGCNN, we design the MTL-CGCNN model and employ the HTW to construct a comprehensive dataset from 48 selected elements as potential candidates for MA_2Z_4. Specifically, we calculated the adsorption energy (E_b) of 2,310 structures by using DFT calculation, which served as the training set for our model. Utilizing the MTL-CGCNN model, we predicted the E_b for the remaining 62208 unoptimized adsorption structures within the compositional space. Notably, our method innovatively predicts adsorption energy for multi-element adsorption systems without requiring prior structural optimization. This model demonstrates exceptional performance by significantly improving both predictive accuracy

and generalization. It reduces the root mean square error (RMSE) by 0.32 eV and the mean absolute error (MAE) by 35% compared to the baseline automated machine learning (Auto-ML) model, indicating significantly improved prediction accuracy. The training process was stable, as evidenced by the small gap between training and validation MAE values. Notably, the model maintains reliable prediction performance even for structures with higher E_b of 4–8 eV. Finally, we identified promising MA_2Z_4 materials like the α-phase $MoGa_2N_4$. The α-phase is a distinct structural configuration characterized by a seven-layered arrangement, contributing to its high adsorption energy (E_b) and excellent structural stability. Our method effectively addresses the challenges of predicting E_b for complex and unoptimized structures within the vast compositional space of MA_2Z_4 materials.

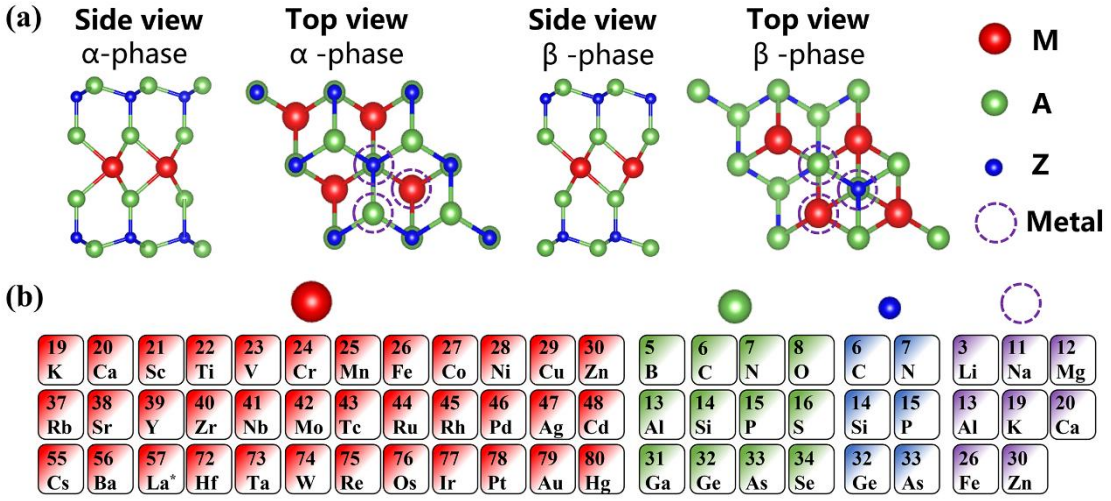

Figure 1. (a) The schematic diagram of MA_2Z_4, where red, green, and blue spheres are M, A, and Z atoms, respectively. The purple circle marks the adsorbed metal atoms. (b) The constituents for structures composed from M, A, Z and metal atoms, respectively.

2. Methods

2.1. HTW

In this study, we primarily investigated the basic structure of seven-layer MA_2Z_4. The side view and top view of two different symmetry types of MA_2Z_4 (the α-phase and the β-phase) are shown (Figure 1a). Based on the arrangement of the interlayer atoms, the stable structures can be classified into α-MA_2Z_4 and β-MA_2Z_4. The α-MA_2Z_4 belongs to P6m2 structure (space group No. 187) with the plane symmetry, while β-MA_2Z_4 belongs to P3m1 structure type, space group (No. 164), being centrosymmetric. According to existing experimental and theoretical work, the M atom should be a

metal atom. We selected 36 M elements from periods 4 to 6 of the periodic table.[12,20] The A and Z elements are chosen from periods 2 to 4, from groups IIIA to VIA with 36 unique A-Z combinations (Figure 1b). Combining the 36 M elements and the two structural phases (α and β) for each composition, the total number of MA_2Z_4 primitive cell structures is 2,592. Furthermore, we examine the interactions of these structures with eight different metal atoms (Li, Na, K, Ca, Mg, Fe, Zn, Al) to form adsorption structures (Figure 2a). In practice, we selected a subset of 770 MA_2Z_4 substrate structures for adsorption energy (E_b) calculation by using the HTW (Figure 2b). Considering three high-symmetry adsorption sites on each substrate, this resulted in 2,310 possible adsorption structures. We performed ab initio screening the resulting structures and identified the stable adsorption configurations for model training (Figure 2c).

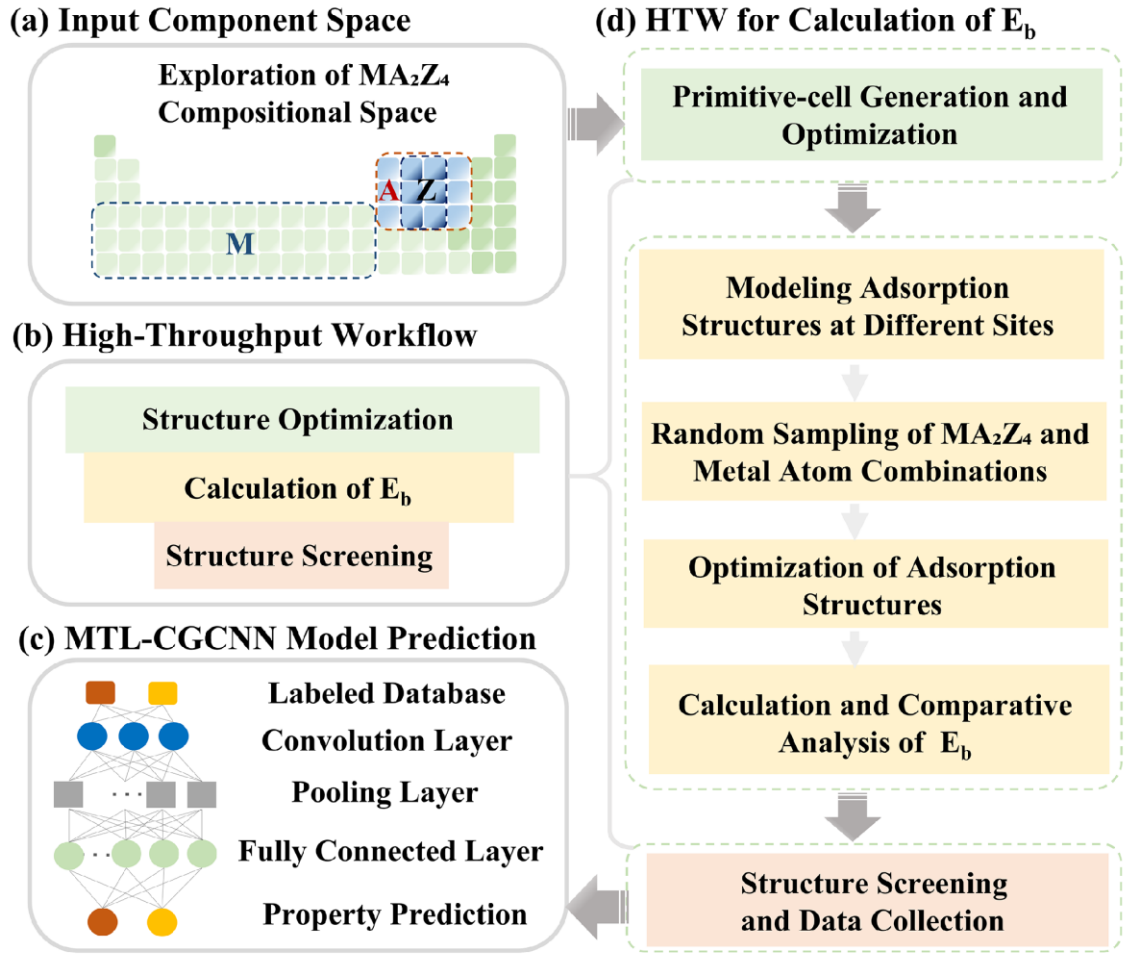

Figure 2. Schematic illustrations about the ideas and methods in this work. (a) Composition space. (b) High-throughput processes. (c) MTL-CGCNN model design. (d) HTW for E_b calculation.

To calculate thousands of adsorption structures, we developed the HTW to systematically explore the structural and adsorption properties of MA_2Z_4 materials (detailed in section 3 of SI).

The workflow mainly consists three main steps (Figure 2d).

1. Generation of the initial 2,592 MA_2Z_4 primitive cell structures within the entire compositional space of MA_2Z_4. These structures were subject to DFT calculations in order to obtain binding and cohesive energies and determine the stable substrate structures in following steps.
2. In this step, adsorption systems were constructed by building upon the primitive cell structures. Each MA_2Z_4 structure is expanded to identify three highly symmetric adsorption sites, designated as adsorption sites A, B, and C. Combinations of MA_2Z_4 and metal atoms were randomly selected from the entire composition space and the metal atoms are positioned at each of these sites. DFT calculations were performed on the adsorption structures of each combination at the three adsorption sites. And the corresponding E_b was calculated which is given by the equation (1).

$$E_b=-\left(E_{\text{adsorption system}}-E_{MA_2Z_4}-E_{\text{metal atom}}\right) \quad (1)$$

   where $E_{\text{adsorption system}}$, $E_{MA_2Z_4}$, $E_{\text{metal atom}}$ represent the energy of the adsorption system, the substrate energy, and the energy of the single metal atom, respectively. By calculating the E_b at each site, the optimal location with the strongest interaction energy is identified, offering insights into the preferred binding characteristics of the metal atoms. Furthermore, the energy calculations for single metal atoms are conducted within cells with identical lattice constants to ensure consistency.
3. The step of structure screening and data collection involves the E_b at the three sites for each adsorption system. The highest E_b is used as an indicator of the metal's affinity. To maintain uniformity in E_b calculation, structures that exhibit significant structural changes upon adsorption are excluded due to inaccurate E_b. The E_b of all stable adsorption structures is then collected and used to train predictive models, thereby completing the dataset construction (Supplementary material for more details).

This approach allows us to selectively perform DFT calculation on a representative subset of adsorption systems while leveraging machine learning models to predict E_b for the remaining systems. Consequently, this methodology optimizes computational efficiency and resource utilization, significantly reducing the time and cost associated with extensive DFT computations.

2.2 Machine Learning Methods

We explored the metallophilicity of thousands of adsorption structures by utilizing the HTW

method and constructed the training dataset. Three models, namely CGCNN, MTL-CGCNN, and Auto-ML, were individually trained. The confidence intervals, significance tests, performance variance, and uncertainty assessments for the three models were conducted, as detailed in Section 5 of the Supplementary Information (SI). The trained models were used to predict E_b for the unrelaxed adsorption structures, significantly reducing the need for DFT computations and enabling the identification of promising materials with strong metallophilicity.

Both CGCNN and MTL-CGCNN models were trained and compared under the same model parameters. The entire dataset was divided into training, validation, and test sets with a ratio of 7:1:2. To ensure no data leakage and prevent overestimation of model performance, we employed a composition-based splitting strategy rather than random splitting. The model configuration included 64 Gaussian radial basis functions (RBF). It functions as a non-linear transformation that enriches the feature representation by embedding distance information into a higher-dimensional space. The cut-off radius is set to 8 Å (Angstrom) and the maximum number of neighbors is set to 12 to balance computational efficiency and information completeness. Both models include a 216-dimensional features as input and are trained using the Adam optimizer for 400 epochs, where an epoch refers to one complete pass through the entire training dataset. The model architecture is composed of graph information encoding, convolutional layers with batch normalization, activation functions, pooling layers, and fully connected layers which also referred to the multilayer perceptron (Figure S1)[43]. In the convolutional layers, Gaussian RBF was used to represent interatomic distance which is given by the equation (2).

$$\phi_k(d) = \exp\left(-\beta_k (d - \mu_k)^2\right) \tag{2}$$

where, d represents distance, and μ_k is the center of the Gaussian function, while β_k is the width parameter. Batch normalization was employed to stabilize training by adjusting layer outputs to standard normal distributions, which is given by the equation (3).

$$\hat{x}^{(i)} = \frac{x^{(i)} - \mu_B}{\sqrt{\sigma_B^2 + \grave{o}}} \tag{3}$$

where, $\hat{x}^{(i)}$ represents the input features, μ_β is the batch mean, and σ_B^2 is the batch variance, with Ò representing a small positive number to prevent division by zero.

As the comparative model for E_b prediction, we also introduced the Auto-ML model. This

model fully leverages techniques combining multiple models and hyperparameter tuning and is widely used in various regression tasks. After comparing various feature extraction methods, the feature dimension of each adsorption structure was designed to be 336-dimensional Auto-ML features. Atomic properties of each element in the material's chemical formula are used as features for material design. In each layer, atoms can be further grouped based on their distance from the adsorbed metal atom, denoted as d-metal MA_2Z_4. Subsequently, ensemble learning was applied to combine multiple models, and Bayesian optimization was used for hyperparameter tuning.

For all the three models, the performance was evaluated based on root RMSE which is given by the equation (4) and mean absolute error (MAE) which is given by the equation (5). By employing both metrics, we achieved the comprehensive analysis of model accuracy and robustness.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{4}$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{5}$$

where, MAE represents the average absolute difference between predicted and actual value, while RMSE highlights larger deviations by penalizing them more heavily.

2.3 DFT calculation

We constructed the training set with labeled E_b was constructed using the framework of DFT,[44] employing the Vienna Ab-initio Simulation Package (VASP)[45,46] to investigate material properties. The Monkhorst-Pack scheme was used for k-point sampling, with a 12×12×1 grid for the primitive unit cell and a 4×4×1 grid for the 3×3×1 supercell, while the vacuum layer of at least 16 Å was introduced to mitigate interactions between periodic images. The Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional within the GGA framework was used,[47] and the plane-wave cut-off energy was set to 520 eV based on convergence tests.[48,49] Van der Waals interactions were corrected using the DFT-D3 method. We use an automated energy-strain-based high-throughput framework to compute the mechanical properties of materials.[50]

3. Result and discussion

3.1 Adsorption Energy Calculation

This study investigated the E_b of seven-layered MA_2Z_4 structures with eight different metal

atoms. We selected 770 metal-MA_2Z_4 pairs from 20736 metal-MA_2Z_4 pairs. Through HTW, we identified 1697 adsorption structures and analyzed the variations in E_b across different metal atoms and adsorption sites. Among them, the median E_b of different metal atoms displayed noticeable variation from approximately 2.10 eV to 5.87 eV (Figure 3a). For instance, K and Ca exhibited the highest median E_b of 5.87 eV and 5.36 eV, respectively, which suggests the stronger interaction between these atoms and the MA_2Z_4 substrate. In contrast, Zn and Mg had much lower median E_b of 2.10 eV and 3.28 eV, respectively.

Adsorption energy distributions across three adsorption sites (A, B, and C) are illustrated in the violin plot (Figure 3b). It shows that the energy distribution ranges from 2.42 eV to 5.85 eV. Site C emerged as the preferred site, likely due to its stability-enhancing configuration, which minimizes direct adsorption above the outermost atomic layers. Additionally, we analyzed the distribution of adsorption structures across the sites (Figure 3c). Site A hosts 619 structures, site B contains 685 structures, and site C has 714 structures. This uneven distribution suggests potential variations in site accessibility or binding strength, with site C likely being the most favorable for adsorption. These findings highlight the versatility of MA_2Z_4 materials in interacting differently with various metals and adsorption sites, underscoring their potential as promising candidates for metallic electrode applications.

However, the selected adsorption structures represent only 4.2% of the entire compositional space, leaving many adsorption structures with uncalculated E_b. There remain 19865 metal-MA_2Z_4 pairs with uncalculated E_b. To efficiently explore the entire compositional space, machine learning models are crucial in predicting the E_b, thereby significantly reducing computational cost and accelerating the discovery of promising materials.

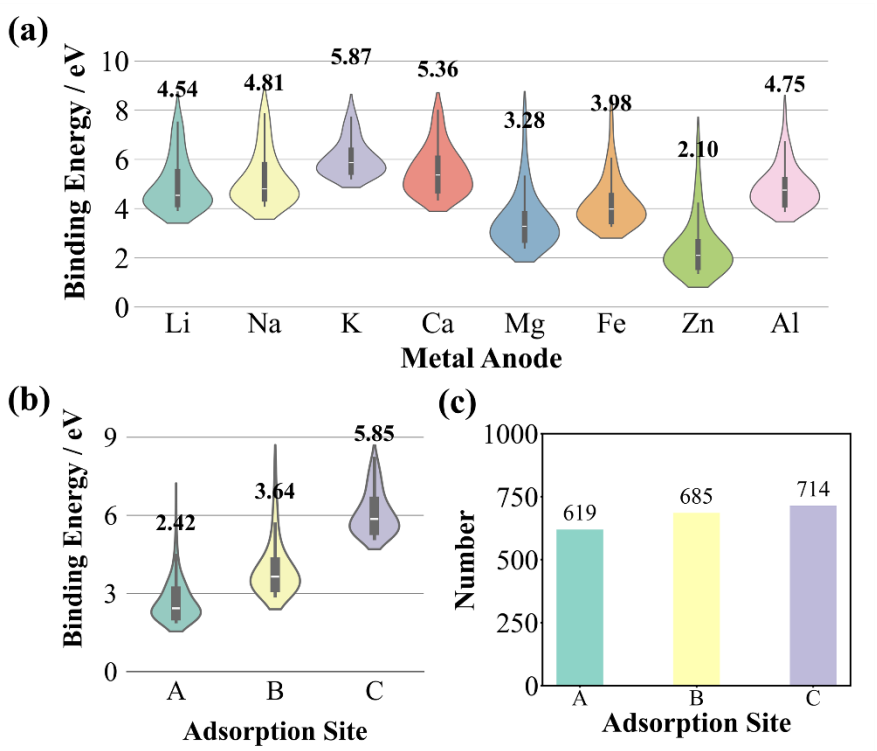

Figure 3. (a) Violin plot for the relationship between DFT calculated E_b and the 8 absorbed metal atoms. (b) Violin plot of E_b distribution by adsorption sites across all metal atoms. (c) Bar chart showing the number of converged structures at different adsorption sites.

3.2 Model Training and Evaluation

To predict E_b, we employed the CGCNN model to capture the adsorption behavior of metal atoms on the surface of MA_2Z_4. In this model, atoms are represented as nodes and interatomic distances as edges. The model integrates elemental properties of atoms and their local environmental information into the 216-dimensional feature space. Consistent with the approach described earlier, we selected 1,357 optimized MA_2Z_4 adsorption structures with adsorption energy labels as the training and validation sets, and 339 structures as the test set. To improve training efficiency, the batch size was set to 800 which means that in each epoch. Each batch undergoes forward and backward propagation, ensuring that the entire dataset is processed once per epoch. With careful parameter design and training, the CGCNN model demonstrated good performance in predicting adsorption energies, achieving a RMSE of 0.78 eV and a MAE of 0.47 eV, indicating generally low prediction error despite some larger deviations (Figure 4a).

The MTL-CGCNN model lies in its encoding process, where tasks are defined by incorporating material system labels (e.g., MA_2Z_4 and MXene) from distinct datasets (in SI Section 3). The primary task is to predict the adsorption energies (E_b) of metal atoms on MA_2Z_4 structures, while

the auxiliary task is to predict the adsorption energy of metal atoms on MXene substrate. To further improve prediction accuracy, we incorporated 7,956 data points of MXene adsorption structures into the training and test sets, adopting a multi-task learning (MTL) strategy with dataset labeling. The inclusion of MXene data enhanced model performance in two key aspects. First, the structural and electronic similarities between MXene and MA_2Z_4 materials allowed the model to leverage shared features, enabling it to capture more complex interatomic interactions and better identify metal adsorption sites. Second, the multi-task learning framework enabled the MTL-CGCNN model to effectively utilize correlations across different adsorption systems, reducing prediction bias and improving its ability to handle extreme values. The performance comparison between the CGCNN and MTL-CGCNN models was conducted using identical hyperparameters. While the CGCNN model achieved an RMSE of 0.78 eV and an MAE of 0.47 eV, it exhibited signs of overfitting, as evidenced by the increasing ΔMAE (difference between training and validation MAE) with training epochs in Figure 4b, indicating divergence in training and validation performance. In contrast, the MTL-CGCNN model maintained a stable ΔMAE, reflecting consistent performance across training and validation sets and better generalization. Additional results are provided in SI (Section 5).

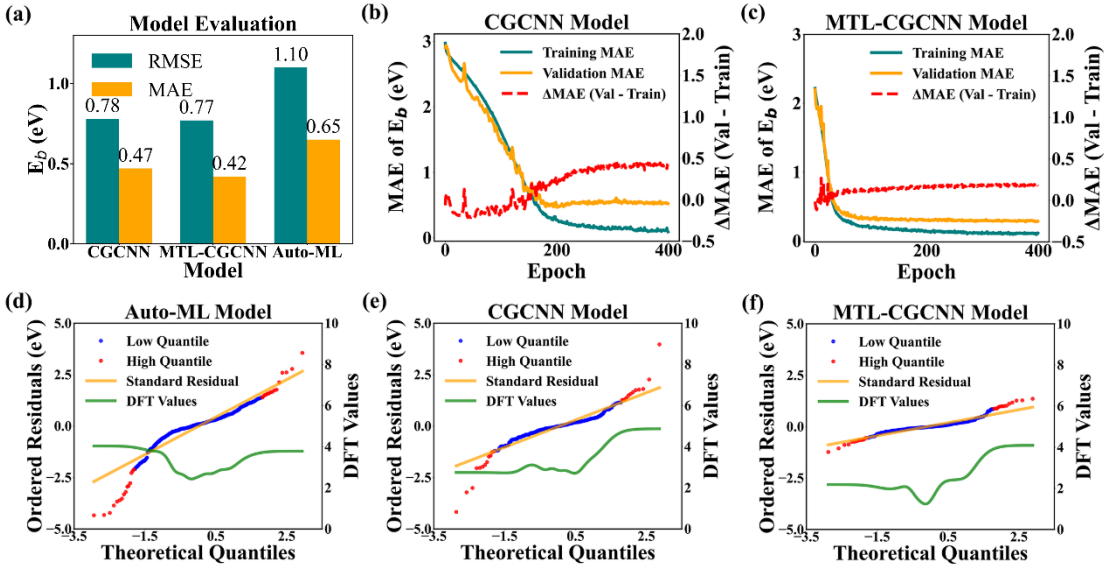

Figure 4. (a) Comparison of MAE and RMSE across different models on training and validation sets. (b) MAE progression over epochs for the CGCNN model during training. (c) MAE progression over epochs for the MTL-CGCNN model during training. (d) Q-Q plot for the Auto-ML model. (e) Q-Q plot for the CGCNN model. (f) Q-Q plot for the MTL-CGCNN model.

We used an Auto-ML model to predict adsorption energy (E_b) for adsorption systems as a benchmark (see Figure 4a and Table S2). We designed 336 descriptors per system, combining atomic properties like electronegativity and ionization energy with structural metrics. Due to dimensional

inconsistencies, the MXene dataset was excluded from the Auto-ML training set. Ablation experiments tested dimensions (64, 128, 192, 216, 336), with 216 dimensions being most efficient, though differences across dimensions were minimal. Using scikit-learn and Bayesian optimization, we tuned an ensemble of 623 models, selecting the top six for the final weighted ensemble predictor of E_b (in SI Section 4). The Auto-ML model achieved a MAE of 0.65 eV for E_b, showing reliable performance in predicting systems with minimal deviations. The RMSE reached 1.10 eV, suggesting that the model struggles with extreme outliers, resulting in a relatively higher overall prediction error. In contrast, the MTL-CGCNN model outperformed the Auto-ML model, reducing the RMSE by ~30% compared to the Auto-ML model, highlighting its superior performance.

Analyzing MAE trends during training and validation highlights key differences in the performance of the models. The MAE trends during training for the CGCNN model (Figure 4b) and the MTL-CGCNN model (Figure 4c) reveal notable difference in their stability and generalization capabilities. The CGCNN model shows a larger gap between training and validation MAE, suggesting potential overfitting. In contrast, the MTL-CGCNN model demonstrates better generalization, with lower and more consistent MAE across training and validation sets, achieving stability in fewer epochs. The model's ability to maintain consistent performance across training and validation sets, as well as its superior handling of outliers, suggests it can make reliable predictions for new instances, including those with high adsorption energy.

In regression analysis, Quantile-Quantile (Q-Q) plots were used to evaluate the normality of residuals for each model—Auto-ML (Figure 4d), CGCNN (Figure 4e), and MTL-CGCNN (Figure 4f). The Q-Q plot shows the theoretical normal distribution as a red line, with blue dots indicating the observed residuals. A strong overlap between the dots and the line implies that the residuals align with a normal distribution. The theoretical quantiles are computed using the Gaussian distribution's quantile formula (equation (6)). Comparative analysis of Q-Q plots indicates that the residuals of the MTL-CGCNN model in extreme E_b regions are closer to the normal distribution compared to the CGCNN model, especially when the theoretical quantiles are less than -1.5. Notably, in regions of high E_b, the MTL-CGCNN model's predicted energy aligned more closely with a normal distribution of residuals. This performance indicates the MTL-CGCNN model's improved ability to handle extreme value and outliers effectively. In contrast, the residuals from the Auto-ML model are more scattered, especially in the high E_b ranges, showing significant deviations. This

suggests that traditional machine learning methods, like the Auto-ML model, struggle to predict E_b in complex two-dimensional materials due to their inability to fully capture the nonlinear relationships in such data. Therefore, compared to the MTL-CGCNN model, the Auto-ML model has limitations in fitting complex data and handling outliers effectively.

$$Q_i = \Phi^{-1}\left(\frac{i-0.5}{n}\right) \tag{6}$$

where Φ^{-1} is the inverse of the cumulative distribution function (CDF) of the standard normal distribution, used to calculate the theoretical quantiles, n is the number of samples.

To validate the effectiveness of MTL-CGCNN, we benchmarked it against three state-of-the-art models such as BE-CGCNN (bond-type embedding), Geo-CGCNN, (Voronoi tessellation and three-body correlations) and Crysco (transformer-based architectures). The results (Table S2) show that MTL-CGCNN outperforms the others in prediction accuracy. The training processes (in SI Figure S4 and Figure S5) demonstrate that MTL-CGCNN exhibits relatively higher training stability. Our benchmarking results show that MTL-CGCNN outperforms BE-CGCNN, Geo-CGCNN, and CrysCo in predictive accuracy and robustness for diverse MA_2Z_4 systems. Additionally, we analyzed the residual distribution of MTL-CGCNN's predictions, which aligns closely with the sample's residual distribution, confirming its reliability and generalization capability across varied material compositions (Figure S6).

Overall, the MTL-CGCNN model exemplifies the benefits of graph-based deep learning and MTL strategy in addressing complex problems in materials science. Its higher prediction accuracy, faster convergence to training stability, reduced overfitting, and improved handling of extreme E_b underscore its potential as a powerful tool for predicting E_b and guiding future research in the field.

3.3 Prediction of Strongest Metallophilicity and the Underlying Mechanism

We performed comprehensive adsorption energy predictions for 62,208 MA_2Z_4 adsorption structures using the selected MTL-CGGNN model. To ensure prediction accuracy, the AutoML model was used to predict lattice constants and interlayer distances with high precision, effectively capturing the relationship between structural parameters and energy through feature extraction. The structural parameter predictions by the Auto-ML model demonstrated high accuracy (Supplementary Information Figure S2 and Table S1).

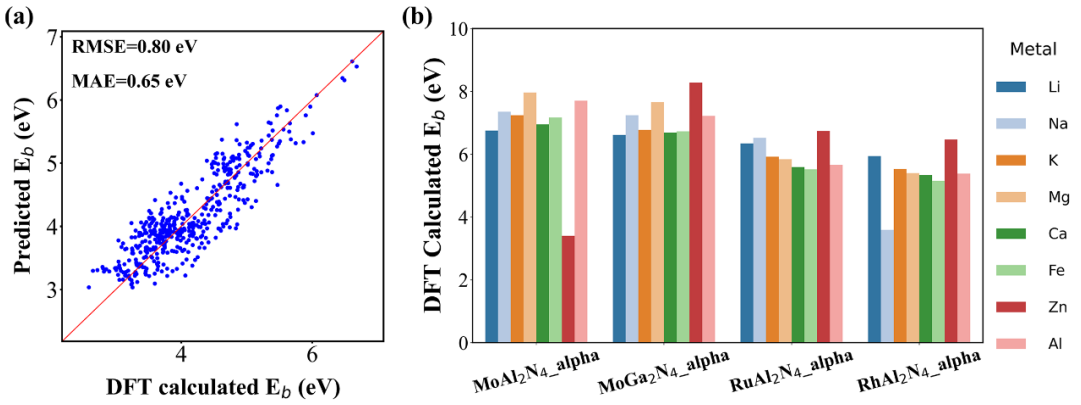

Figure 5. (a) Scatter plot comparing DFT-calculated E_b with MTL-CGCNN-predicted E_b for the top ~1% of structures. (b) DFT-calculated E_b for selected MA_2Z_4 compounds with the strongest metallophilicity in each metal category.

Based on the Auto-ML, we constructed the initial structures for the adsorption system. Then the predictions from the MTL-CGCNN model were then used to screen the E_b for various MA_2Z_4 materials. For each adsorbed metal atom, we categorized the E_b across the entire compositional space and ranked them from high to low. Higher E_b corresponds to stronger metallophilicity. For each metal, we selected the top ~1% of structures, yielding 623 adsorption structures in total. DFT validation then confirmed that 568 of these were stable. Adsorption structures with high E_b, as predicted by the MTL-CGCNN model, were highly consistent with DFT-calculated value (Figure 5a). This further confirms the model's robustness and reliability in predicting complex adsorption behaviors. Among these, four substrate structures ($MoAl_2N_4$_alpha, $MoGa_2N_4$_alpha, $RuAl_2N_4$_alpha, and $RhAl_2N_4$_alpha) exhibit the strongest metallophilic property, showing the highest affinity for the adsorbed metal atoms. The phonon spectra for these four substrate structures were simulated, with no imaginary frequencies found (Supplementary Figure S5). These results demonstrate their dynamic stability. The adsorption behaviors of different metals vary across the above four different MA_2Z_4 systems (Figure 5b). For instance, Mg and Zn show high value of E_b in most systems, with $MoGa_2N_4$_alpha exceeding 8 eV. $MoGa_2N_4$_alpha and $MoAl_2N_4$_alpha substrates also exhibit relatively high E_b for Li, Na, and K. And Ca and Fe have similar E_b across these materials. Furthermore, we calculated the mechanical properties of materials with reliable stability and included these results in Section 7 of the Supplementary Information (SI).

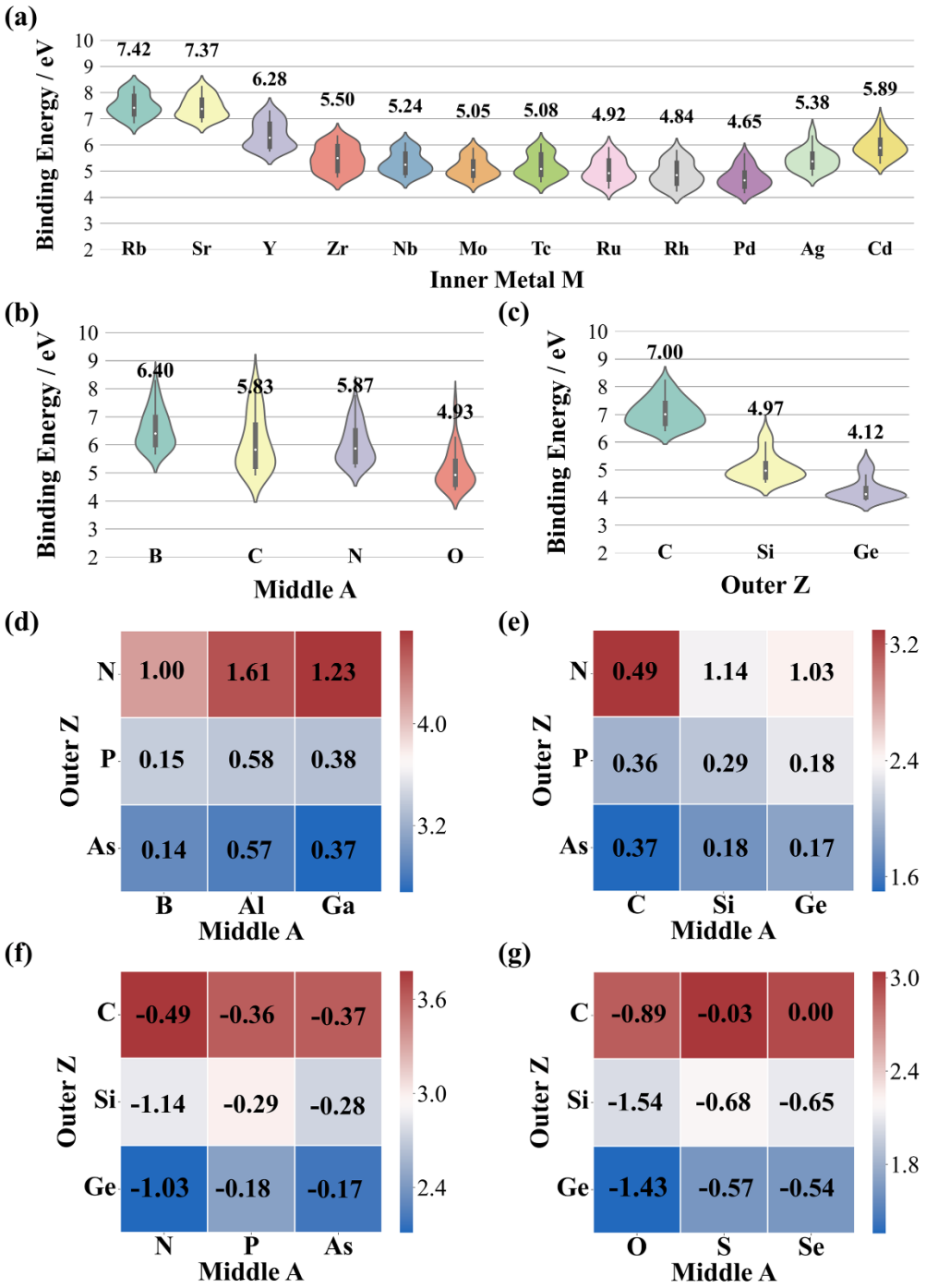

Figure 6. Violin plots of adsorption energy distributions for MA_2Z_4 configurations. (a) Different inner metal M. (b) Different middle A elements. (c) Different outer Z elements. Heatmaps (d-g) present the average E_b values for different combinations of A and Z elements, with M varies across all possible metals. The numbers in the heatmaps represent the $\Delta\chi$. (d) Z = N, P, As. A = B, Al, Ga. (e) Z = N, P, As. A = C, Si, Ge. (f) Z = C, Si, Ge. A = N, P, As. (g) Z = C, Si, Ge. A = O, S, Se.

To gain insights into the underlying factors influencing the metallophilicity of MA_2Z_4 materials, we statistically analyzed the median E_b distributions with respect to the M, A, and Z elements. The median E_b values are displayed above each violin plot for clarity (Figure S3). We began by analyzing the effect of M elements in the fourth period of the Periodic Table on E_b (Figure 6a). Elements like

Ru/Sr in the main group and Ag/Cd in the subgroup exhibit higher median E_b. This trend suggests that M elements with partially filled s-orbitals or nearly full d-orbitals are more likely to accept or share electrons with neighboring A atoms, leading to higher E_b. For the A elements, E_b decreases as the electronegativity of A increases (Figure 6b). Low-electronegativity A elements tend to donate electrons to Z, which strengthens the interaction between Z and the adsorbed metals. The Z elements occupy the outermost layer in the MA_2Z_4 structure and directly influence the interaction between the substrate and the adsorbed metal atoms (Figure 6c). We found that relatively higher electronegative Z elements like C lead to strong electron redistribution at the interface, resulting in high E_b.

We further focus on the electronegativity difference between outer Z and middle A ($\Delta\chi = \chi_A - \chi_Z$), as it plays a crucial role in determining the strength of interactions with adsorbed metal atoms. We used heatmaps to present the distributions of the average E_b for different A and Z combinations (Figure 6d-g). Deeper red tones correspond to higher E_b values, while lighter blue tones indicate lower E_b. This visual representation clearly illustrating the positive correlation between $\Delta\chi$ and E_b. When Z = N/P/As and A = B/Al/Ga, the large difference in $\Delta\chi$ leads to generally high adsorption strength, with N-Ga exhibiting the highest $\Delta\chi$ among these combinations and the highest E_b of 4.7 eV (Figure 6d). On the contrary, when A = C/Si/Ge, its higher electronegativity reduces $\Delta\chi$, thereby lowering E_b (Figure 6e). Similarly, when Z = C/Si/Ge, the $\Delta\chi$ values become negative, but the overall trend between $\Delta\chi$ and adsorption strength remains consistent (Figure 6f-g). When Z = C and A = N/P/As, the large negative $\Delta\chi$ corresponds to the E_b of about 3.8 eV, indicating strong adsorption. For A = O/S/Se, the magnitude of the negative $\Delta\chi$ decreases, resulting in lower E_b. To further validate this trend, we fixed the M = Mn as the metal and investigated the E_b for different combinations of A and Z elements with differing $\Delta\chi$ Figure S4). The charge difference density maps show that systems with larger $\Delta\chi$ exhibit more pronounced electron accumulation and depletion at the interface, indicating stronger metal–substrate interactions. The analysis of Bader charge transfer further confirms the positive correlation between $\Delta\chi$ and and E_b (Figure S10). Thermodynamic parameters (E_b) and electronic structure analyses ($\Delta\chi$) both show that the electronegativity difference between the outer Z and middle A elements is a key factor affecting adsorption strength. Synergistically with Bader charge and charge density, consistently shaping adsorption energy across element combinations.

Our analysis revealed that the E_b of MA_2Z_4 materials is primarily determined by the $\Delta\chi$ between the Z and A elements. A larger $\Delta\chi$ results in stronger adsorption strength due to the increased electron transfer between layers. This relationship holds true across various element combinations, providing a valuable guideline for designing MA_2Z_4 materials with optimized adsorption properties for applications in energy storage and catalysis.

4. Conclusion

In this study, we systematically investigated the metallophilicity of seven-layered MA_2Z_4 materials with eight different metal atoms using the HTW approach and the MTL framework. We efficiently generated 2592 MA_2Z_4 primitive cell structures and screened 62208 possible adsorption structures. From these, 2310 adsorption configurations were selected and identified 1697 stable adsorption structures for model training. We then used the trained MTL-CGCNN model to predict the E_b of 20736 metal–MA_2Z_4 pairs, which greatly reduced the computational resources for DFT calculations. The MTL-CGCNN model demonstrated superior performance, particularly in regions with high E_b, outperforming traditional Auto-ML models. From the thermodynamic and kinetic stable materials, we identified 50 candidates exhibiting exceptional mechanical properties (e.g., $E_{xy} > 300$ GPa, G_z up to 280 GPa) and high E_b (e.g., $MoGa_2N_4$_alpha > 8 eV). Our results indicate that the adsorption energy of MA_2Z_4 materials is primarily determined by the $\Delta\chi$ between the Z and A elements, where A larger $\Delta\chi$ leads to stronger adsorption. These results reveal the potential of MA_2Z_4 materials as versatile metallic electrodes with strong interactions, suitable for various applications. Furthermore, this study underscores the potential for extending the model to explore the properties of a broader range of two-dimensional materials, such as graphene and BN, across diverse application scenarios.

Supplementary materials

Supplementary material associated with this article can be found, in the online version.

Data availability

Data for this article are available on request from the authors.

Author contributions

Q.Z., P.Z, D.L. and X.F. conceived and designed the project. P.Z. and X.F. designed the HTAW and performed DFT simulations. P.Z. and H.D. designed the ML features and performed the ML study and deep learning study. P.Z. and Q.Z. prepared the manuscript and all the authors discussed the results and commented on the manuscript.

Conflict of interest

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

Acknowledgements

Q.Z. acknowledges the support by the Science and Technology Major Project of Hebei Province (22284402Z), the Beijing Natural Science Foundation (2192029), the National Key Research and Development Program of China (2017YFB0702100). Z.W.S. acknowledges the support of the Singapore National Research Foundation (NRF Investigatorship NRF-NRFI09-0002). D.L. acknowledges the projects e-INFRA CZ (ID:90254) and LUASK22099 by the Ministry of Education, Youth and Sports of the Czech Republic.

Notes and references

1 D. Lin, Y. Liu and Y. Cui, *Nat. Nanotechnol.*, 2017, **12**, 194-206.

2 E. Fan, L. Li, Z. Wang, J. Lin, Y. Huang, Y. Yao, R. Chen and F. Wu, *Chem. Rev.*, 2020, **120**, 7020-7063.

3 J. Xiao, *Science*, 2019, **366**, 426-427.

4 B. Lee, E. Paek, D. Mitlin and S. W. Lee, *Chem. Rev.*, 2019, **119**, 5416-5460.

5 C. Wei, Y. Tao, H. Fei, Y. An, Y. Tian, J. Feng and Y. Qian, *Energy Storage Mater.*, 2020, **30**, 206-227.

6 J. H. Kwak, Y. Jeoun, S. H. Oh, S. Yu, J.-H. Lim, Y.-E. Sung, S.-H. Yu and H.-D. Lim, *ACS Energy Lett.*, 2021, **7**, 162-170.

7 S. D. Pu, C. Gong, X. Gao, Z. Ning, S. Yang, J.-J. Marie, B. Liu, R. A. House, G. O. Hartley and J. Luo, *ACS Energy Lett.*, 2020, **5**, 2283-2290.

8 Q. Yang, Q. Li, Z. Liu, D. Wang, Y. Guo, X. Li, Y. Tang, H. Li, B. Dong and C. Zhi, *Adv. Mater.*, 2020, **32**, 2001854.

9 Y. Zhang, Y. Tian, Z. Wang, C. Wei, C. Liu, Y. An, B. Xi, S. Xiong and J. Feng, *Chem. Eng. J.*, 2023, **458**, 141388.

10 Y. Long, H. Li, M. Ye, Z. Chen, Z. Wang, Y. Tao, Z. Weng, S.-Z. Qiao and Q.-H. Yang, *Energy Storage Mater.*, 2021, **34**, 194-202.

11 Z. Liang, D. Lin, J. Zhao, Z. Lu, Y. Liu, C. Liu, Y. Lu, H. Wang, K. Yan and X. Tao, *Proc. Natl. Acad. Sci. U.S.A.*, 2016, **113**, 2862-2867.

12 X. Feng, R. Dong, Y. Li, X. Liu, C. Lin, T. Wang, Z. W. Seh and Q. Zhang, *Energy Storage Mater.*, 2023, **63**, 103035.

13 G. Zheng, S. W. Lee, Z. Liang, H.-W. Lee, K. Yan, H. Yao, H. Wang, W. Li, S. Chu and Y. Cui, *Nat. Nanotechnol.*, 2014, **9**, 618-623.

14 K. Yan, H.-W. Lee, T. Gao, G. Zheng, H. Yao, H. Wang, Z. Lu, Y. Zhou, Z. Liang and Z. Liu, *Nano Lett.*, 2014, **14**, 6016-6022.

15 A. S. Arico, P. Bruce, B. Scrosati, J.-M. Tarascon and W. Van Schalkwijk, *Nat. Mater.*, 2005, **4**, 366-377.

16 Y. Li, W. Y. Lieu, T. Ghosh, L. Fu, X. Feng, A. J. Y. Wong, A. Thakur, B. C.

Wyatt, B. Anasori and Q. Zhang, *Small Methods*, 2023, **7**, 2201598.

17 H. Song, J. Su and C. Wang, *Adv. Energy Mater.*, 2021, **11**, 2003685.

18 B. Sun, P. Li, J. Zhang, D. Wang, P. Munroe, C. Wang, P. H. Notten and G. Wang, *Adv. Mater.*, 2018, **30**, 1801334.

19 L. Wang, Y. Shi, M. Liu, A. Zhang, Y.-L. Hong, R. Li, Q. Gao, M. Chen, W. Ren and H.-M. Cheng, *Nat. Commun.*, 2021, **12**, 2361.

20 C. Lin, X. Feng, D. Legut, X. Liu, Z. W. Seh, R. Zhang and Q. Zhang, *Adv. Funct. Mater.*, 2022, **32**, 2207415.

21 K. Wang, P. Wang, Y. Qian, X. Wang, J. Luo, X. Tao and W. Li, *Mater. Today Energy*, 2024, **42**, 101565.

22 Y. Liu, Y. Ji, Y.-M. Ding, Y. Li and S.-T. Lee, *2D Mater.*, 2023, **10**, 025020.

23 J. Du, X. Zhou, X. Cheng and G. Jiang, *J. Colloid Interface Sci.*, 2025, **678**, 150-158.

24 Y. Liu, T. Zhao, W. Ju and S. Shi, *J. Materiomics*, 2017, **3**, 159-177.

25 M.-F. Ng, J. Zhao, Q. Yan, G. J. Conduit and Z. W. Seh, *Nat. Mach. Intell.*, 2020, **2**, 161-170.

26 J. R. Kitchin, *Nat. Catal.*, 2018, **1**, 230-232.

27 M. Karamad, R. Magar, Y. Shi, S. Siahrostami, I. D. Gates and A. Barati Farimani, *Phys. Rev. Mater.*, 2020, **4**, 093801.

28 X. Feng, Q. Zhang and Z. W. Seh, *Adv. Mater. Technol.*, 2023, **8**, 2200616.

29 T. Xie and J. C. Grossman, *Phys. Rev. Lett.*, 2018, **120**, 145301.

30 G. Cong and V. Fung, *Mach. Learn. Sci. Technol.*, 2023, **4**, 035030.

31 K. Bang, D. Hong, Y. Park, D. Kim, S. S. Han and H. M. Lee, *Nat. Commun.*, 2023, **14**, 3004.

32 S. Pablo-García, S. Morandi, R. A. Vargas-Hernández, K. Jorner, Ž. Ivković, N. López and A. Aspuru-Guzik, *Nature Computational Science*, 2023, **3**, 433-442.

33 J. Cheng, C. Zhang and L. Dong, *Communications Materials*, 2021, **2**, 92.

34 M. Madani, V. Lacivita, Y. Shin and A. Tarakanova, *npj Computational Materials*, 2025, **11**, 15.

35 C. W. Park and C. Wolverton, *Phys. Rev. Mater.*, 2020, **4**, 063801.

36 K. Meng, C. Huang, Y. Wang, Y. Zhang, S. Li, Z. Fang, H. Wang, S. Wei and S. Sun, *J. Chem. Inf. Model.*, 2023, **63**, 6043-6052.

37 K. M. Jablonka, D. Ongari, S. M. Moosavi and B. Smit, *Chem. Rev.*, 2020, **120**, 8066-8129.

38 A. Merchant, S. Batzner, S. S. Schoenholz, M. Aykol, G. Cheon and E. D. Cubuk, *Nature*, 2023, **624**, 80-85.

39 L. Chanussot, A. Das, S. Goyal, T. Lavril, M. Shuaibi, M. Riviere, K. Tran, J. Heras-Domingo, C. Ho and W. Hu, *ACS Catal.*, 2021, **11**, 6059-6072.

40 S. Ding and Q. Qian, *Comput. Mater. Sci.*, 2024, **237**, 112897.

41 Y. Zeng, H. Cao, Q. Ouyang and Q. Qian, *Mater. Today Commun.*, 2021, **27**, 102314.

42 X. Zou, C. Liang, B. Wang, S. Hao, G. Chen and P.-A. Heng, 2022.

43 S. D. Xue and Q. J. Hong, *Materials*, 2024, **17**, 4176.

44 G. Kresse and J. Furthmüller, *Comput. Mater. Sci.*, 1996, **6**, 15-50.

45 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

46 P. E. Blöchl, *Phys. Rev. B*, 1994, **50**, 17953.

47 G. Kresse and D. Joubert, *Phys. Rev. B*, 1999, **59**, 1758.

48 S. Grimme, J. Antony, S. Ehrlich and H. Krieg, *J. Chem. Phys.*, 2010, **132**, 154104.

49 S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.*, 2011, **32**, 1456-1465.

50 S. H. Zhang and R. F. Zhang, *Computer Physics Communications*, 2017, **220**, 403-416.