ACS APPLIED MATERIALS & INTERFACES

www.acsami.org

Research Article

Computational Exploration of Adsorption-Based Hydrogen Storage in Mg-Alkoxide Functionalized Covalent-Organic Frameworks (COFs): Force-Field and Machine Learning Models

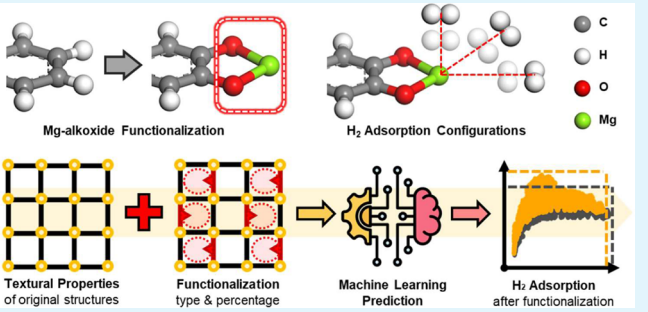
Yu Chen, Guobin Zhao, Sunghyun Yoon, Parsa Habibi, Chang Seop Hong, Song Li, Othonas A. Moultos, Poulumi Dey, Thijs J. H. Vlugt, and Yongchul G. Chung*

Cite This: *ACS Appl. Mater. Interfaces* 2024, 16, 61995−62009

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Hydrogen is a clean-burning fuel that can be converted to other forms. of energy without generating any greenhouse gases. Currently, hydrogen is stored either by compression to high pressure (>700 bar) or cryogenic cooling to liquid form (<23 K). Therefore, it is essential to develop safe, reliable, and energy-efficient storage technology that can store hydrogen at lower pressures and temperatures. In this work, we systematically designed 2902 Mg-alkoxide-functionalized covalent-organic frameworks (COFs) and performed high-throughput (HT) computational screening for hydrogen storage applications at 111, 231, and 296 K. To accurately model the interaction between Mg-alkoxide sites and molecular hydrogen, we performed MP2 calculations to compute the hydrogen binding energy for different types of functionalized models, and the data were subsequently used to fit modified-Morse force field (FF) parameters. Using the developed FF models, we conducted HT grand canonical Monte Carlo (GCMC) simulations to compute hydrogen uptakes for both original and functionalized COFs. The generated data were subsequently used to evaluate the materials' gravimetric and volumetric storage performance at various temperatures (111, 231, and 296 K). Finally, we developed machine learning (ML) models to predict the hydrogen storage performance of functionalized structures based on the features of the original structures. The developed model showed excellent performance with a mean absolute error (MAE) of 0.061 wt % and 0.456 g/L for predicting the gravimetric and volumetric deliverable capacities, enabling a quick evaluation of structures in a hypothetical COF database. The screening results demonstrated that the Mg-alkoxide functionalization yields greater improvements in volumetric H_2 storage capacities for COFs with smaller pores compared to those with larger (mesoporous) pores.



KEYWORDS: *hydrogen storage, metal-alkoxide functionalization, covalent-organic framework, high-throughput screening, machine learning*

1. INTRODUCTION

With increasing awareness of global warming, significant research efforts focus on transitioning from fossil fuels to clean, zero-emission hydrogen energy.[1,2] Hydrogen (H_2) is considered as a promising alternative due to its high energy density and environmental benefits. However, efficiently storing H_2 remains a significant challenge. Due to the high gravimetric and low volumetric energy density of H_2 at ambient conditions, current storage technologies require either cryogenic temperatures or high pressures.[3−5] Cryogenic storage involves cooling H_2 to 20 K, while compressed storage requires pressurizing the storage tank up to 700 bar, both demanding substantial energy and specialized equipment. To address these challenges, H_2 storage via physical adsorption on nanoporous materials, such as metal−organic frameworks (MOFs) and covalent-organic frameworks (COFs), has been proposed.[6−8] COFs, built from organic monomers linked by strong covalent bonds, offer high surface areas and pore volumes, making them attractive for H_2 storage.[9] Experimental and computational studies have proved the outstanding performance at cryogenic conditions (generally 100 bar for adsorption and 5 bar for desorption, at 77 or 160 K).[10−12] Unfortunately, current adsorbents under near-ambient conditions show limited H_2 binding compared to cryogenic

Received: July 22, 2024
Revised: October 12, 2024
Accepted: October 22, 2024
Published: October 30, 2024

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

conditions and fail to meet the U.S. Department of Energy (DOE) technical targets for H_2 storage performance.[13]

One of the strategies to overcome the limitations of current adsorbents is by introducing new functional groups to enhance the H_2 binding and improve the adsorption performance of the adsorbent materials.[14−19] Recent research highlights the effectiveness of introducing open metal sites (OMSs) in MOFs via metal-alkoxide functionalization for integrating metal cations into the framework structures.[16] These OMSs provide free coordination sites that interact directly with H_2 molecules, enhancing the polarization and binding strength. Metal-alkoxide functionalization, demonstrated experimentally, stably incorporates metal cations into MOFs by modifying the phenol linkers postsynthesis.[20] This modification can be achieved by exchanging one or two alcohol H with a metal cation using solvent-based methods or atomic layer deposition (ALD).

Computational screening studies also indicate that Mg-alkoxide exhibits high heat of adsorption at high pressures (adsorption pressure) and moderate heat of adsorption at lower pressures (desorption pressure), making it suitable for pressure-swing storage.[18] Getman et al. built and analyzed various metal-alkoxide sites, parametrizing the force field based on *ab initio* calculated binding energies. Then with the obtained FF models, H_2 loadings of five functionalized MOFs were calculated using GCMC simulations, showing an enhancement of up to 1 wt % at ambient temperature, with Mg-alkoxide catecholate showing the largest heat of adsorption at high pressure. Chen et al.[19] performed high-throughput (HT) screening of over 2700 Mg-alkoxide-functionalized Zr-based MOFs for the theoretical maximum H_2 capacity, assuming four H_2 molecules bind to one Mg site. Using the FF models developed in the aforementioned work by Getman et al., GCMC simulations were conducted to calculate the H_2 capacity of the top 100 MOFs from the initial screening, showing the highest H_2 deliverable capacities (DC) of 7 wt % and 24 g/L at ambient temperature.

When the FF models derived from MP2-calculated H_2 binding energies are adopted for GCMC simulations, selecting the correct FF models is crucial for accurately representing the adsorption behavior of functionalized materials. In the case of Mg-alkoxide functionalization on phenyl rings, functional groups can be added to one side (single-site functionalization) or both sides (two-site functionalization). Using FF models developed for single-site functionalization in two-site simulations can lead to inaccuracies, as the binding and adsorption behavior differs between these two types of functionalization. Therefore, developing precise and specialized FF models for each type of functional site is essential for accurately predicting the adsorption performance of COFs functionalized with Mg-alkoxide groups. With these tailored FF models, we can predict the adsorption performance of functionalized COFs, enabling a computational screening approach to identify high-performance materials for H_2 storage at ambient temperature.

While FF-based simulations could reduce the computational time for estimating adsorption performances in target porous materials, computational screening based on molecular simulations can be integrated with machine learning (ML) to further speed up the evaluation of material properties. The potential of using the ML approach to predict various material properties and performances has been demonstrated to be highly effective for adsorption applications.[21−26] Traditional feature-based ML approaches use lists of material features representing the physical and chemical properties of the structures as inputs, employing linear, tree-based, or simple neural network algorithms to predict material performance, such as hydrogen storage capabilities.[27−30] More recent research reports directly use atom and bond information on the structure itself as features. Advanced approaches, such as crystal graph convolutional neural networks (CGCNNs), can directly predict properties such as band gaps, energy,[31,32] partial atomic charges,[33] and adsorption uptakes[39] from input crystal structures. Especially, Table S1 lists several recent studies[25,34,35] on the ML prediction of H_2 adsorption on COFs.

In this work, we developed specialized FF models for different types of Mg-alkoxide functionalization and estimated the H_2 adsorption performances of the functionalized COFs, which lack abundant OMSs that enable efficient adsorption within framework structures such as MOFs, using GCMC simulations with the specific FF models. HT screening was performed on the COF database to identify high-performance materials at near-ambient temperatures. Pressure-swing DC was estimated at both cryogenic and ambient temperatures. Based on this large data set of H_2 uptakes corresponding to various structures with diverse properties and adsorption performances, ML models were trained to predict H_2 uptake at different temperatures and pressures. Two different types of ML algorithms, including traditional tree-based gradient-boosting regression (GBR) and more recent CGCNN, were adopted. Through the SHAP (SHapley Additive exPlanations) analysis,[36] the contributions of different features to the H_2 gravimetric and volumetric uptakes predictions were investigated, providing detailed insight into the relationship between structural properties and H_2 storage performance. Furthermore, the incorporation of the GBR models enabled larger-scale HT screening of over 90,000 materials from the hypothetical COF database ReDD-COFFEE, expanding our understanding of how Mg-alkoxide functionalization impacts H_2 storage efficiency.

2. METHOD

2.1. Electronic Structure Calculations and Parametrization of FF Models. The H_2 binding energies to Mg-alkoxide groups in functionalized COFs were estimated by using Mg-catecholates. Geometry optimizations and single-point energy calculations of single-site and two-site Mg-catecholate structures were performed at the MP2[37−41]/6-311+G**//MP2/6-311+G** level of theory. Various configurations of H_2 approaching the Mg-alkoxide sites were constructed to simulate different orientations and distances.[18] Single-point calculations were calculated by keeping both H2 and Mg-catecholate molecules rigid along the approach angle. These calculations were carried out using Gaussian 16 software.[42] The H_2 binding energy to Mg-catecholate was calculated as

$$E_{\text{binding}} = E_{\text{total}} - (E_{\text{Mg-catecholate}} + E_{H_2}) \quad (1)$$

where E_{binding} is the binding energy, E_{total} is the total energy of the whole H_2-catecholate system, $E_{\text{Mg-catecholate}}$ is the energy of Mg-catecholate molecule, and E_{H_2} is the energy of the H_2 molecule.

The energies from MP2 calculations were used to parametrize FF models for interactions between H_2 and Mg-alkoxide groups. The nonbonded interaction potentials,

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

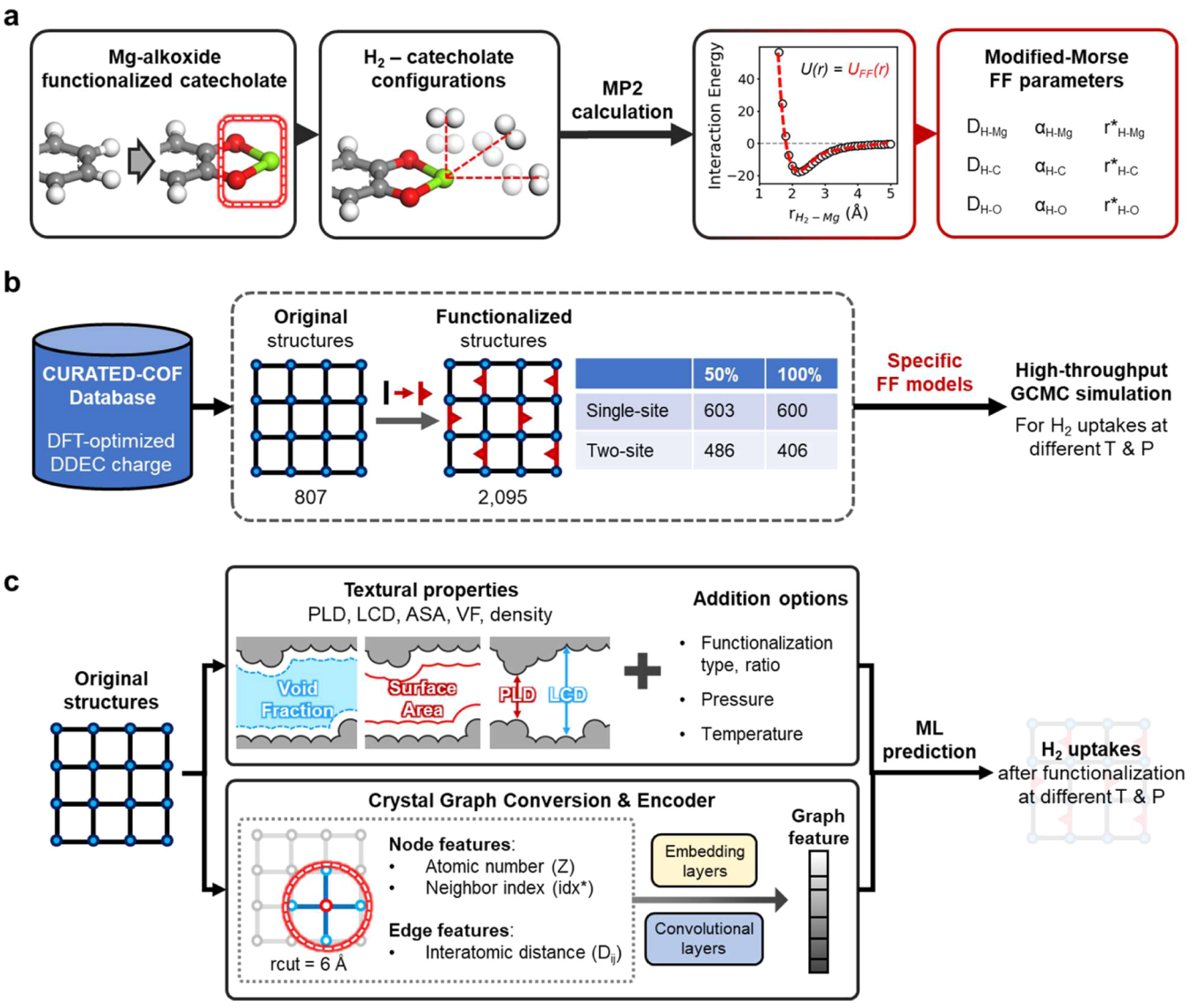


Figure 1. Overall flowchart of (a) construction of different sets of H_2-Mg-alkoxide-functionalized catecholate configurations and FF parametrization based on the MP2-calculated interaction energies. (b) Obtaining COF structures from CURATED-COF database and constructing functionalized structures with different functionalization types and ratios, and high-throughput screening on both original and functionalized COFs for high H_2 deliverable capacity at various temperatures using GCMC simulations with specified FF models developed in (a). (c) Different ML processes of predicting the H_2 uptakes of COFs at certain temperature and pressure points based on the original structures with feature-based ML approach (upper) and CGCNN (lower).

including van der Waals (vdW) and electrostatic potentials, were described by modified-Morse/Lennard-Jones and Coulomb potential models. Interaction energies of six different H_2-catecholate configurations (150 data points) were fitted to these models as follows:

$$U_{\text{forcefield}}(r_{ij}) = U_{\text{LJ}}(r_{ij}) + U_{\text{mod-Morse}}(r_{ij}) + U_{\text{Coulomb}}(r_{ij}) \tag{2.1}$$

$$U_{\text{LJ}}(r_{ij}) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{2.2}$$

$$U_{\text{mod-Morse}}(r_{ij}) = -k_{\text{B}}D_{ij}\left[1 - \left(1 - e^{-\alpha_{ij}(r_{ij}-r_{ij}^{*})}\right)^{2}\right] \tag{2.3}$$

$$U_{\text{Coulomb}}(r_{ij}) = \frac{1}{4\pi\varepsilon_0}\frac{q_i q_j}{r_{ij}} \tag{2.4}$$

where U_{ij} is the interaction energy between atoms i and j and r_{ij} is the distance between atoms i and j. k_{B} is the Boltzmann constant and ε_0 is the vacuum permittivity constant. ε_{ij} and σ_{ij} are the LJ parameters representing the LJ well-depth and the distance at which the particle−particle potential energy is zero (also refers to the diameter of the atom), respectively. D_{ij}, α_{ij}, and r_{ij}* are the modified-Morse parameters representing the well-depth, well-width, and equilibrium bond distance, respectively. q_i and q_j represent the partial atomic charges of different atoms.

For the vdW interactions, the modified-Morse model was adopted to describe the interactions between H atoms in H_2 adsorbates and Mg, O, and the nearest-neighbor C atoms in the Mg-alkoxide group. For the interactions between H atoms

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

in H_2 adsorbates and the other atoms in the framework, the LJ parameters from DREIDING[43] and Universal Force Field (UFF)[44] were applied. The Feynman−Hibbs (FH) correction was used on the LJ potential in H_2 simulations to account for the non-negligible nuclear quantum effects of H_2 molecules at cryogenic temperatures:

$$U_{FH}(r_{ij}) = \frac{\hbar^2}{24\mu k_B T}\nabla^2 U_{LJ}(r_{ij}) \quad (3)$$

where μ is the reduced mass, k_B is the Boltzmann constant, T is the temperature, $\hbar$ is the reduced Planck's constant.

For electrostatic interactions, only interactions between H atoms in H_2 adsorbates and atoms in the Mg-alkoxide group were considered, with partial atomic charges estimated using MP2 calculations, as described in works by Getman et al.[18] Charges on other framework atoms were neglected for the functionalized structures. H_2 molecules were modeled using the rigid 3-site Darkrim−Levesque model,[45] with two H atoms carrying a positive charge of 0.468 and a center-of-mass (COM) pseudoatom carrying a negative charge of −0.936.

A dual annealing approach[46] globally minimized the root-mean-squared error (RMSE) between H_2 binding energies calculated from MP2 and FF models across all data points, determining the modified-Morse parameters. The workflow for FF development based on MP2-calculated binding energies is shown in Figure 1a. Additional details on the metrics, parametrization procedure, and FF model parameters are provided in Sections S2 and S3.

2.2. Molecular Model Construction. We obtained 807 geometry-optimized COFs from the CURATED-COF[47,48] database and constructed 2,095 functionalized COFs by introducing different types (single-site and two-site) and ratios (50 and 100%) of functionalization to the original structures. The functionalization of the structures was implemented by adding the Mg-alkoxide group to the original linkers, as shown in Figure 1a. The geometry of the Mg-alkoxide groups was kept consistent with the optimized Mg-catecholates, and no overall geometry optimization was performed on the structures after functionalization. The Python package MOFUN[49] and the Julia package PoreMatMod[50] were used to perform the modifications on the molecular structures.

The textural properties of the COFs, including pore-limiting diameter (PLD), largest cavity diameter (LCD), void fraction (VF), accessible volumetric surface area (ASA), and density, were computed based on the analysis of the Voronoi network, as implemented in the Zeo++3.0 software.[51−53] The accessible pore volume and surface area were calculated using the probe radius of 1.84 Å, the kinetic radius of nitrogen.[54] Structures with a PLD smaller than the diameter of H_2 adsorbate (2.958 Å) were excluded from the data set. The detailed number of each type of original and functionalized COF structure in the data set is shown in Figure 1b.

2.3. Grand Canonical Monte Carlo Simulation. All GCMC simulations were performed for 6000 equilibration cycles followed by 10,000 production cycles with equal probabilities for translation, reinsertion, rotation, and swap moves in each MC step, as implemented in the open-source RASPA 2.0[55] software. The nonbonded interaction potential between the atoms in COFs and H_2 adsorbates was described by the aforementioned modified-Morse/Lennard−Jones plus Coulomb FF model. The Lorentz−Berthelot mixing rules[56] were applied for the LJ parameters. The vdW interaction cutoff was set to 12.8 Å, and an analytic tail correction was applied beyond the cutoffs. The long-range electrostatic interactions were computed using the Ewald summation method, as implemented in the RASPA 2.0 software. The real-space interactions were truncated at a cutoff distance of 12 Å, while the reciprocal-space contribution was handled using the Ewald summation approach. For original COF structures, the partial atomic charge information was estimated by density-derived electrostatic and chemical (DDEC6) method,[57−60] as included in the crystal definition files provided by the CURATED-COF database. For the functionalized COF structures, only the partial atomic charges on the Mg and O atoms in the Mg-alkoxide functional sites were considered, as listed in Table S3, while the partial atomic charges on all other framework atoms were set to zero. Validations of the impact of different framework charge settings on the final GCMC simulation results are provided in Section S6.

GCMC simulations were conducted at temperatures of 111, 231, 296 K and pressures of 5, 100 bar. The pressure-swing DC was defined as the difference between the uptakes at the adsorption and desorption pressures of 100 and 5 bar, respectively, at each temperature. The definition can be represented as follows:

$$DC = N_{100bar} - N_{5bar} \quad (4)$$

where N_x represents uptake at a certain pressure x.

2.4. Machine Learning. Machine learning (ML) models were trained to predict the H_2 uptakes at different temperatures and pressures for both original and functionalized COFs from the original COF structures and to enable the prediction of the pressure-swing DCs (100 → 5 bar) at various temperatures. We employed both simple tree-based ML models, such as decision tree (DT), gradient-boosting regression (GBR), and random forest (RF), and a more recent deep learning model, the CGCNN model developed by Cao et al.[31] derived from the original model by Xie et al.,[61] for the training process. The schematic shown in Figure 1c describes the training processes of these two different types of ML models. The data set was split in the ratio of 8:2 for the ML model training and testing, respectively. The distributions of the features and targets in the training and test sets are shown in Figure S16.

For the tree-based ML models, the input features included textural properties from the original structures, functionalization type (single-site or two-site), ratio of functionalization (50 or 100%), temperature, and pressure. For the CGCNN, atomic and bond information served as an input. Node and edge features, including atomic numbers and distances to near-neighbor atoms, were encoded and sent into the convolutional and hidden layers. Any structures that the code failed to convert to feature symbols were excluded. The target for traditional ML models was the H_2 uptake at specific temperatures and pressures; for the CGCNN, the target was a list of 30 uptake values for the original, 50%, and 100% single-site and two-site functionalized structures at two different pressure and three temperature points.

The 5-fold cross-validation was performed in the training process of tree-based ML models to avoid overfitting by holding out parts of the training set as a validation set. The hyperparameter tuning was carried out with Bayesian optimization based on Gaussian Processes, with the goal of minimizing the mean squared error (MSE) of the model in the loops and finding the optimal parameters. For the training

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

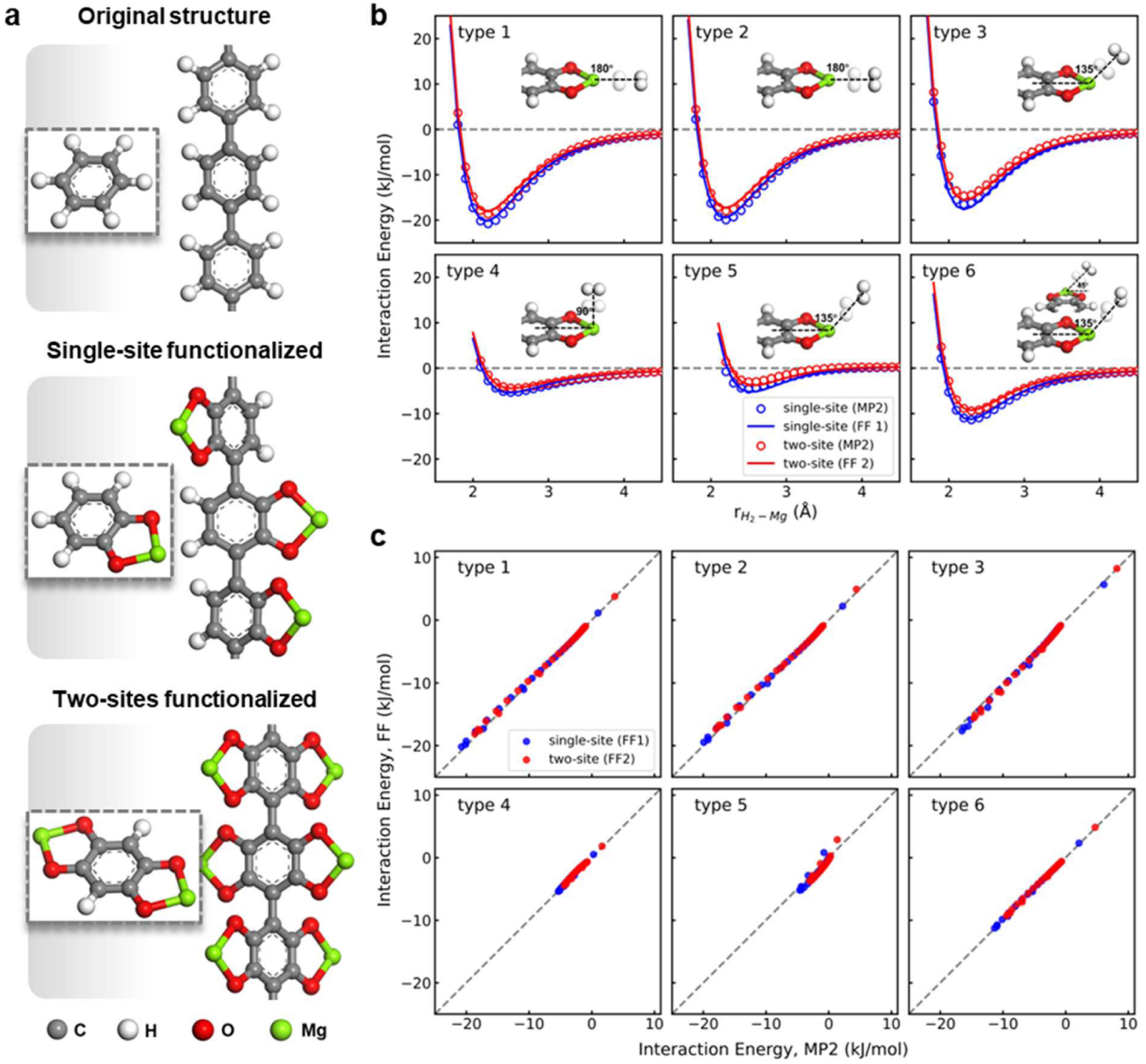


Figure 2. (a) Construction of single- and two-site functionalization with Mg-alkoxide groups on the original catecholate and phenyl rings on the linker in a framework structure. (b) H_2-catecholate interaction energies were calculated with the modified-Morse/Lennard-Jones plus Coulomb potential (lines) compared with energies calculated by MP2/6-311+G** (circles). The energies were calculated from different H_2-catecholate configurations, as shown by the atomic models in each plot. (c) Comparison of the H_2-catecholate interaction energies calculated by the force field (FF) and MP2.

process of CGCNN, one-fourth of the training set was selected to be the validation set (overall train:valid:test = 6:2:2) and MSE is the metric used in forward propagation during model training and Stochastic Gradient Descent (SGD) was used for the weights and bias finetuning. The 5-fold cross-validation was implemented on Python using the scikit-learn[62] library and the hyperparameter tuning was performed using the scikit-optimize library. The data preparation and model training for the CGCNN were performed using the Python packages pymatgen[63] and PyTorch.[64]

3. RESULTS AND DISCUSSION

3.1. H_2-Catecholate Interaction Energies and FF Parametrization. The H_2 binding energies to the Mg-alkoxide groups in the functionalized COFs were estimated by calculating the interaction energies with the Mg-catecholate models. Figure 2a illustrates the introduction of single- and two-site Mg-alkoxide groups on the phenyl rings and ligands in the frameworks. The interaction energies for six different H_2-catecholate configurations are shown in Figure 2b. All configurations followed a similar trend with very high positive interaction energies at small distances, distinct minima at intermediate distances, and energies approaching zero as the distance increases. The variations between the single-site and two-site models reflected differences in binding energies due to the different types of functionalization of the phenyl moiety. The H_2 binding energies for single-site Mg-catecholate were slightly lower than those for two-site models, though the Mg–

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

H_2 distances at the lowest energy were similar in both cases. For instance, in the type 1 configuration, the single-site model's energy was approximately 2 kJ/mol lower than the two-site model. In both cases, the interaction energies reached −20 kJ/mol when the Mg−H_2 distance was approximately 2.2 Å.

Two different sets of modified-Morse model parameters were obtained by fitting to the MP2 energies of the single-site and two-site models, respectively, named FF1 and FF2. The obtained FF parameters and overall fitting RMSE are shown in Table 1. A comparison between the energies estimated by the FF models and MP2 calculations is shown in Figure 2b. The good agreements indicated that the FF models could accurately reproduce the interaction energies of different H_2-catecholate configurations, with overall fitting RMSE of 0.508 and 0.412 kJ/mol for single-site and two-site cases, respectively.

Table 1. Modified-Morse Parameters Fitted from the MP2-Calculated Interaction Energy

FF1		overall fitting RMSE = 0.508 kJ/mol		
atom 1	atom 2	D	α	r^*
H_h2[a]	Mg_cat	910.197	2.027	2.242
H_h2	C_cat	28.629	1.490	3.825
H_h2	O_cat	1.0	0.931	6.335
FF2		overall fitting RMSE = 0.412 kJ/mol		
atom 1	atom 2	D	α	r^*
H_h2	Mg_cat	801.051	2.054	2.257
H_h2	C_cat	30.059	1.515	3.778
H_h2	O_cat	1.0	0.912	6.399

[a] *_h2*: atoms in the H_2 adsorbate; *_cat*: atoms in Mg-alkoxide group in the catecholate.

The comparison between the H_2 binding energies of different functionalized models indicated that, compared to adding Mg sites to both sides of a phenyl ring, the single-site functionalized structure, where the Mg site was only added to one side of each phenyl ring, enabled stronger binding with the H_2 on each Mg site. Since the LJ parameters and partial charges were not specified for different types of functionalization, the differences in the H_2 binding energies to the Mg sites were reflected in the parameters of the specific modified-Morse models FF1 and FF2. The parameter D for the interaction between H_h2 and Mg_cat was shown to be higher for FF1, leading to a deeper potential well with stronger binding at the optimal position. The good agreement with the MP2 energy results demonstrated the flexibility of the modified-Morse potential model in accurately describing the more realistic anharmonic interaction potential. The detailed fitting RMSEs for each configuration are shown in Figure S3.

3.2. High-Throughput Screening of Mg-Alkoxide Functionalized COFs. *3.2.1. Simulations of Two-Site Functionalized COFs with Different FF Models.* To verify the necessity of using specific FF models for different functionalized structures and understand their impact on simulation results, H_2 uptakes of 886 two-site functionalized COF structures at various temperatures and pressures were calculated using GCMC simulations with FF1 and FF2 models. The pressure-swing DCs between 100 and 5 bar at each temperature were also determined. The DC results are shown in Figure 3 and detailed uptakes at adsorption and desorption pressures are shown in Figure S8.

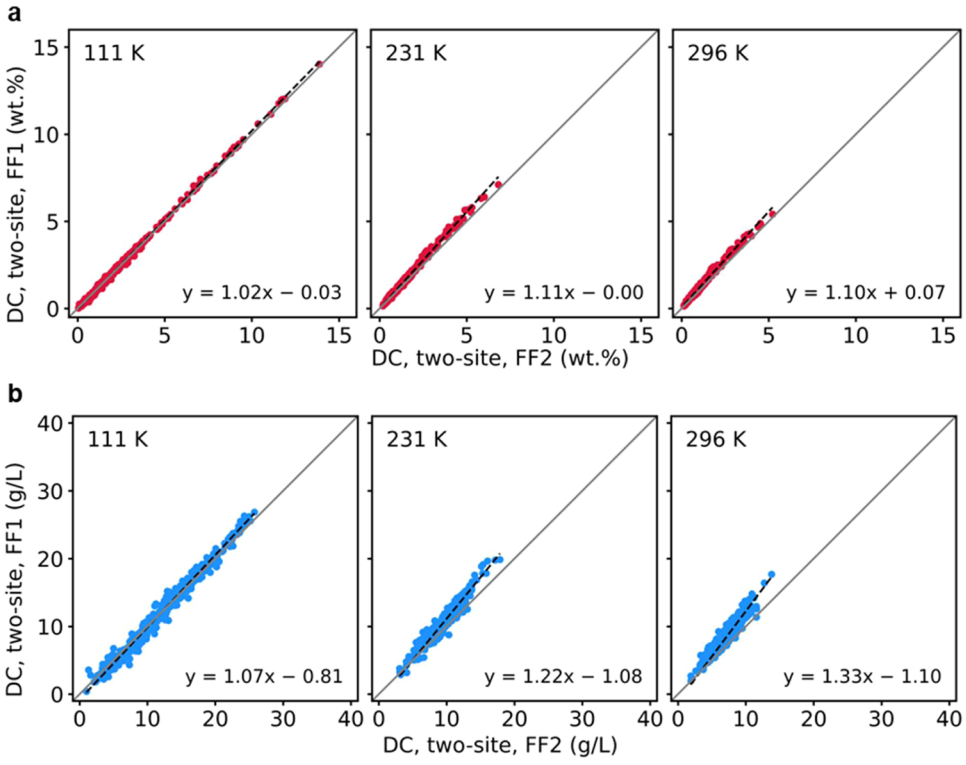


Figure 3. Comparison of (a) gravimetric and (b) volumetric hydrogen deliverable capacity (DC) of two-site functionalized structures calculated using FF1 and FF2 for the three different temperatures. The dashed lines in the subplots with the equations in the lower right corner show the results of the linear regression of the data points.

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

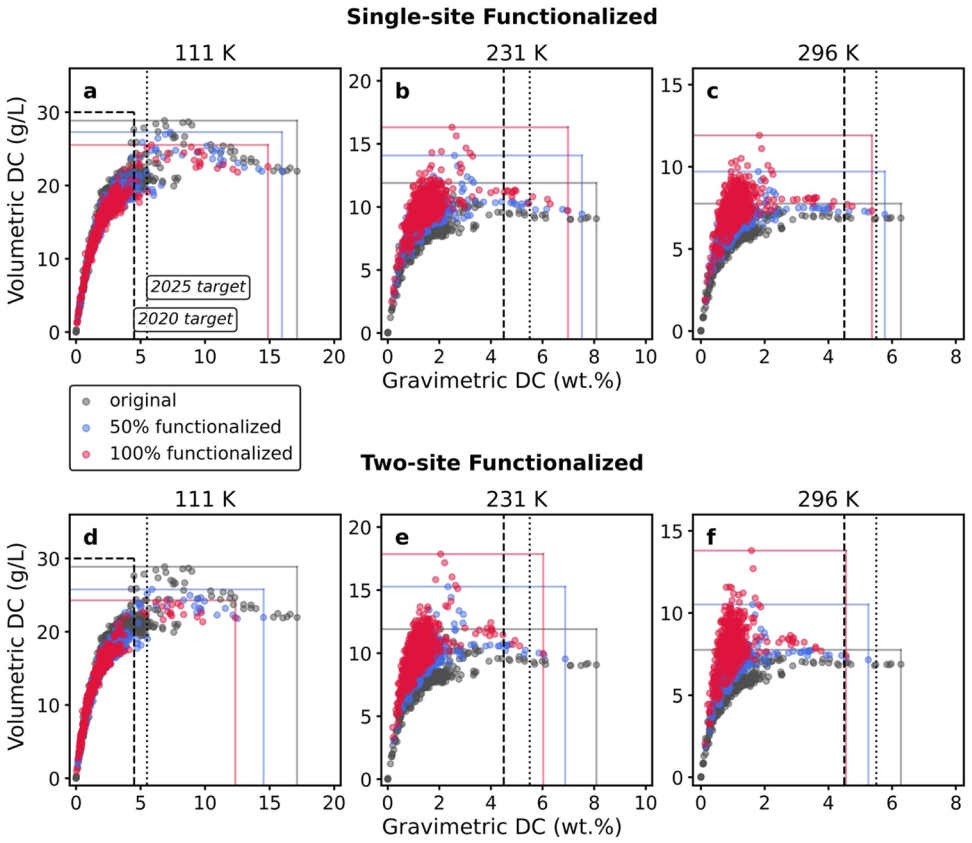


Figure 4. Comparison of gravimetric and volumetric DC for single-site (upper subplots, 111 K (a); 231 K (b); 296 K (c)) and two-sites (lower subplots,111 K (d); 231 K (e); 296 K (f)) functionalized groups at three different temperatures. The scatters were colored according to original structures in gray, 50% functionalized in blue, and 100% functionalized in red. The dashed and dotted lines represent the 2020 and 2025 target standards based on the DOE Technical Targets for Hydrogen Storage[13] shown in Table S6. The solid lines represent the maximum values of the gravimetric and volumetric DC for each data set, colored according to the scatters.

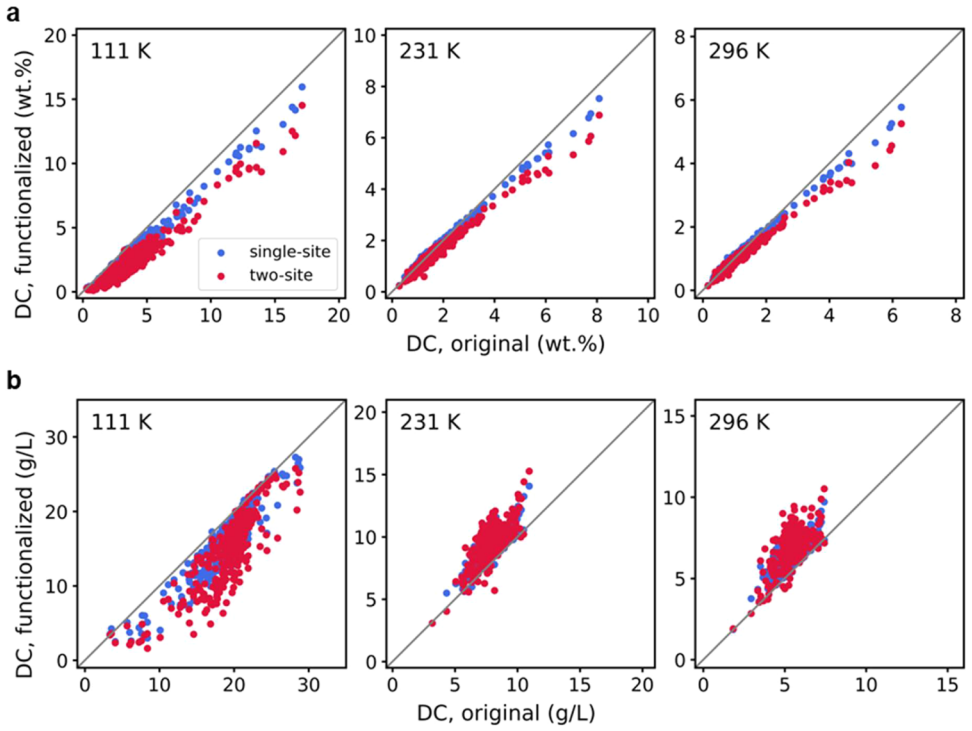


Figure 5. Comparison of the (a) gravimetric and (b) volumetric DC after functionalization (single-site: blue, two-sites: red) compared to the original structures at three different temperatures.

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

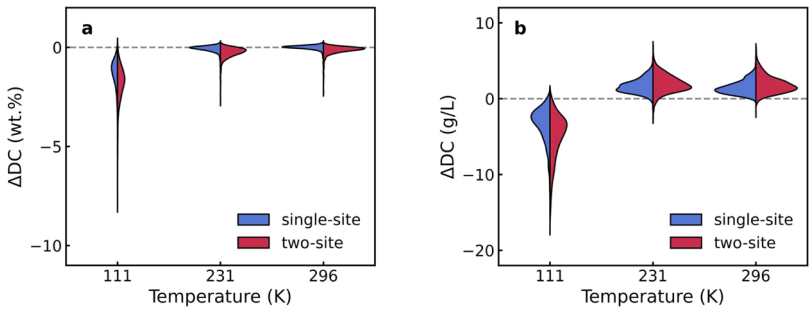


Figure 6. Violin plots showing the distributions of the enhancement of (a) gravimetric and (b) volumetric DC (ΔDC) after functionalization (single-site: blue, two-sites: red) compared to the original structures at three different temperatures.

Due to the stronger H_2 binding with Mg sites in the single-site functionalized structures, FF1 generally overestimated H_2 uptakes compared to FF2. This overestimation was more pronounced at 5 bar than at 100 bar, especially in volumetric uptakes, where the largest deviation reached approximately 5 g/L. At cryogenic temperatures and high pressures, uptakes were near saturation and limited by pore volume, reducing the overestimation by FF1. At ambient temperatures, high-pressure uptakes were not significantly limited by saturation, leading to a greater overestimation by inaccurate FF models.

Consequently, the gravimetric DCs showed no significant difference between FF1 and FF2 across all temperatures due to minor differences in gravimetric H_2 uptakes at adsorption and desorption pressures. Volumetric DCs calculated with FF1 could be overestimated by up to 30% at 296 K, while showing negligible differences at 111 K. This clearly showed that incorrect FF models could cause significant deviations in volumetric DCs at ambient temperatures, emphasizing the importance of using accurate FFs for reliable adsorption performance predictions, particularly at ambient conditions.

3.2.2. Simulations of Original and Functionalized COFs with Specific FFs. High-throughput (HT) GCMC simulations were performed on 2902 structures, including original and functionalized COFs, using the specific FF models as discussed in Section 2. H_2 uptakes at three temperatures (111, 231, 296 K) and two pressures (5, 100 bar) were calculated, and the pressure-swing DCs were then determined. Figure 4 shows the gravimetric and volumetric DCs for the original and functionalized COFs at three different temperatures. The comparison between the original and functionalized structures is illustrated in Figure 5, and the enhancement of DCs (ΔDC) due to functionalization is presented in Figure 6.

At 111 K, original structures showed gravimetric and volumetric DCs higher than those of functionalized ones. The top original structures displayed a volcano-like distribution with a peak gravimetric DC around 17 wt % and a volumetric DC around 28.85 g/L. With a few exceptions, functionalization generally led to a decrease in both gravimetric and volumetric DCs. For instance, while the maximum volumetric DC for original structures was slightly higher than for functionalized ones, the maximum gravimetric DC shows a significant drop postfunctionalization. This indicated that at cryogenic temperatures, original materials performed better in terms of H_2 storage capacity and functionalization did not significantly enhance the DC.

At ambient temperatures (231 and 296 K), the trends shifted from the cryogenic condition. For instance, at 296 K, original structures had gravimetric DCs mostly below 3 wt % and volumetric DCs mostly below 6 g/L, with maximum values of 6.28 and 7.76 g/L, respectively. Functionalized structures consistently showed reduced gravimetric DCs but increased volumetric DCs. The top functionalized structures outperformed the original ones in terms of volumetric DC, with enhancements ranging up to 6.69 g/L. Conversely, gravimetric DC generally exhibited only slight variations, though in some cases, reductions of up to −2 wt % were observed. This indicated that at ambient temperatures functionalized structures showed better volumetric DC performance, though gravimetric DCs were not significantly impacted.

In conclusion, while functionalization did not improve H_2 storage performance at cryogenic temperature and even reduced gravimetric storage at ambient temperatures, it significantly enhanced volumetric storage at ambient conditions. The key challenge remained to identify conditions under which volumetric DC can be increased without substantially compromising gravimetric DC, thereby optimizing the overall storage performance.

3.3. H_2 Adsorption Isotherms of Original and Functionalized COFs. To gain better insight into how functionalization types affect H_2 adsorption behavior, two COF structures were selected from the CURATED-COF database: 21320N2, with a channel-like shape, and 07010N3, with a more complex porous structure. These structures, originally named VCOF-PyrBpy[65] and COF-102,[66] respectively, exhibit differences in their textural properties as shown in Table S7. Specifically, 07010N3 has higher ASA, VF, and more potential functional sites compared with 21320N2, resulting in significant variations in performance before and after functionalization.

H_2 adsorption isotherms of the original and functionalized structures were calculated using specific FF models, and the results are shown in Figures S14 and S15. Introducing heavier O and Mg atoms increased density and decreased gravimetric uptakes of H_2 and DC after functionalization, while volumetric uptakes better reflected the absolute amount of adsorbate since the cell volume remained unchanged.

For the structure 21320N2, adding more Mg sites did not significantly change volumetric uptake but led to continuous decreases in the gravimetric uptake of H_2 at both cryogenic and ambient temperatures. This lack of significant improvement in volumetric DC and the decrease in gravimetric DC was due to the limited functionalization sites, resulting in a minimal enhancement in H_2 binding and adsorption capacity. In contrast, 07010N3 showed increased volumetric uptakes at ambient temperature with more Mg sites. Although at low temperatures and low-pressure adsorption condition, the adsorption saturation limited uptake increases for highly functionalized structures, leading to decreased DC. At ambient

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

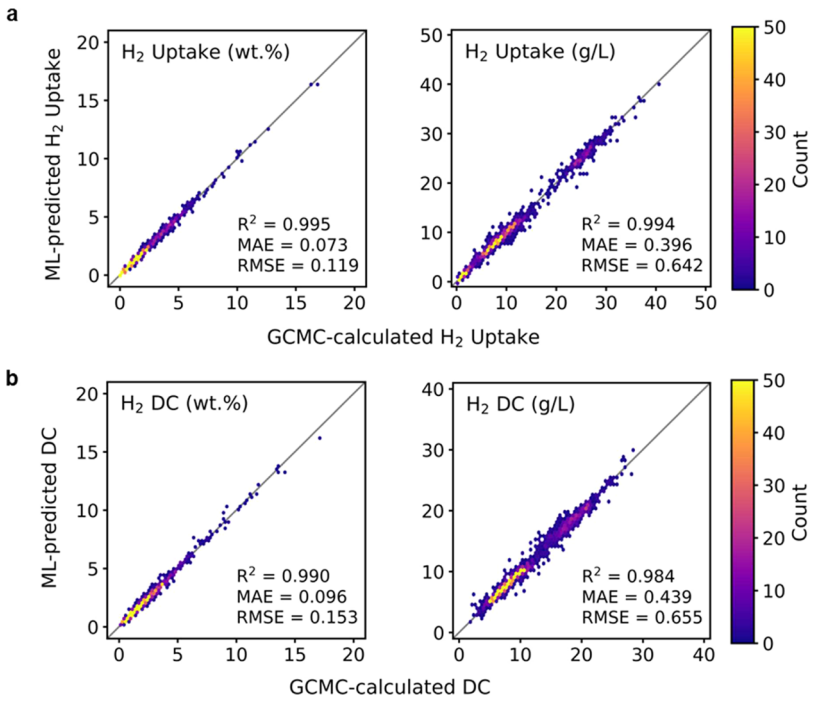


Figure 7. Gravimetric and volumetric H_2 (a) uptakes and (b) deliverable capacity (DC) obtained by GCMC simulations and predicted by GBR models. The figures include data from both original and functionalized structures at 3 temperatures. All data points shown in the plots are from the test set.

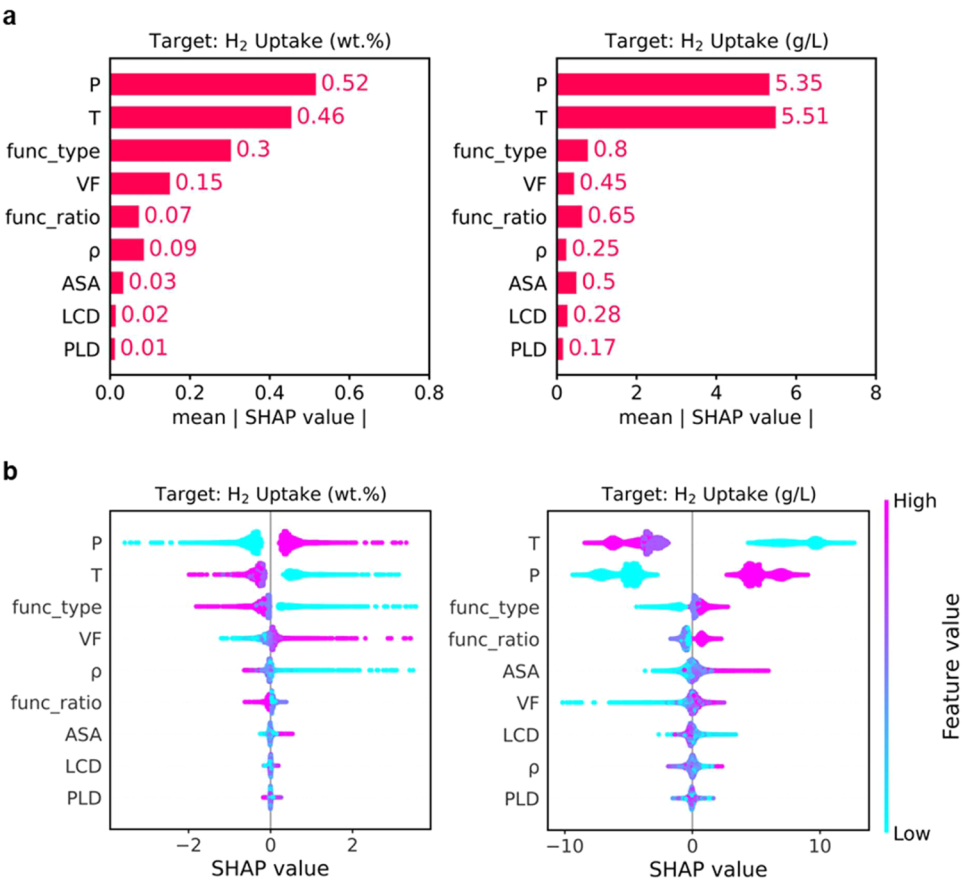


Figure 8. (a) Mean absolute SHAP values and (b) SHAP values of the GBR models with the gravimetric (left, GBR_g) and volumetric (right, GBR_v) H_2 uptake as the target, respectively. In the feature list, func_type stands for the functionalization type, i.e., original (0), single-site (1), or two-site (2); func_ratio stands for the functionalization ratio, i.e., 0% (original), 50%, and 100%. The color bar in (b) represents the value of each feature, from high (magenta) to low (cyan).

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

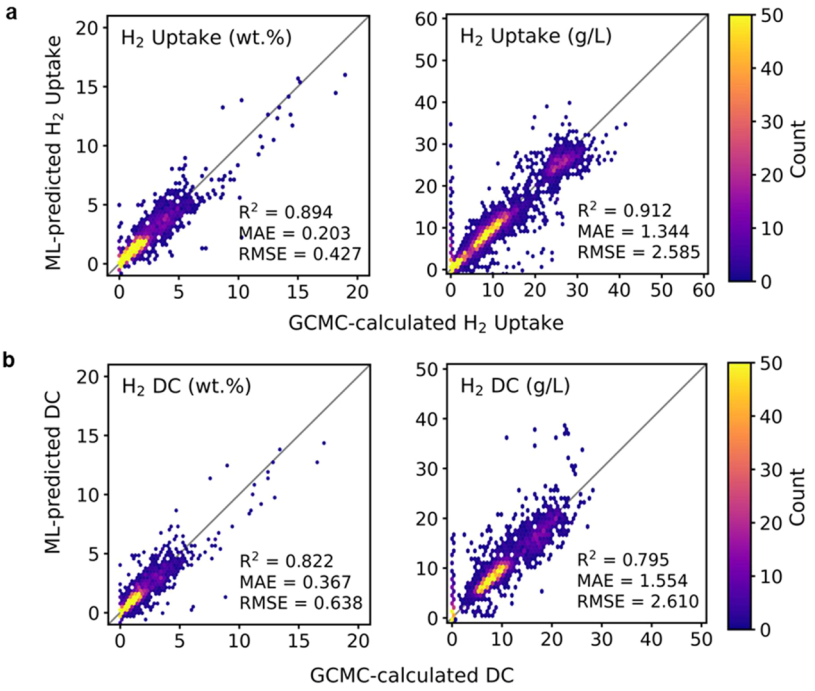


Figure 9. Gravimetric and volumetric H_2 (a) uptakes and (b) deliverable capacity (DC) obtained by GCMC simulations and predicted by CGCNN models. The figures include data from both original and functionalized structures at 3 temperatures. All data points shown in the plots are from the test set.

temperatures, where adsorption did not reach saturation, uptakes at higher pressure were not limited by pore volume, resulting in increased volumetric DC after functionalization.

Furthermore, upon the comparison of 100% single-site and 50% two-site functionalization on the same structure, both configurations introduced the same number of Mg-alkoxide groups. Due to different binding strengths, 100% single-site functionalized structures (i.e., 07010N3_single_all) showed slightly higher uptakes than 50% two-site functionalized structures (i.e., 07010N3_two_half). This difference was more pronounced for 07010N3, where at 296 K, the volumetric DC of 07010N3_single_all exceeded that of 07010N3_two_half by approximately 1.5 g/L.

These observations highlighted how specific structural characteristics and functionalization strategies impact H_2 adsorption performance. While 21320N2 showed limited improvement, 07010N3 demonstrated a notable increase in volumetric DC at ambient temperatures, illustrating the effects of functionalization on different COF structures and providing detailed insights beyond the overall analysis.

3.4. ML Model for Predicting Deliverable Capacity of Functionalized COFs. *3.4.1. Gradient-Boosting Regression.* Among the three tree-based ML algorithms selected, namely, GBR, RF, and DT—the GBR models demonstrated the highest prediction accuracy for both gravimetric and volumetric uptakes. Consequently, this section will focus on the GBR models, with results from the other algorithms provided in Section S8. The optimal sets of hyperparameters for the GBR models are shown in Table S9.

The GBR models accurately predicted H_2 uptakes with R^2 values of 0.995 and 0.994 for GBR_g and GBR_v models, respectively. This is illustrated in Figure 7a, where GBR-predicted uptakes closely matched GCMC-calculated values, exhibiting acceptable errors. The pressure-swing DC between 100 and 5 bar, calculated from the ML-predicted uptakes, also aligned well with GCMC simulation results, as shown in Figure 7b. The good agreements and low MAE and RMSE values demonstrated that the GBR models can reliably predict pressure-swing DCs based on textural properties, functionalization type and ratio, temperature, and pressure.

The analysis of the SHAP values in Figure 8 highlights the importance of various features in the GBR models. For gravimetric uptake, pressure and temperature were identified as the most influential features, followed by functionalization type, void fraction (VF), and density. The SHAP value plot further confirmed that higher pressure and lower temperature positively influenced H_2 uptake with functionalization parameters also playing significant roles. Similarly, for volumetric uptake, pressure and temperature again dominated, with significant contributions from functionalization type and ratio and VF. These insights emphasized the necessity of optimizing pressure and temperature conditions, along with appropriate functionalization strategies, to enhance H_2 storage performance.

The inclusion of temperature and pressure as input features allowed the ML models to predict DC under various conditions, such as pressure-swing, temperature-swing, and combined pressure−temperature-swing scenarios. This capability enabled predictions of DC across a wide range of operating conditions, from 111 K, 100 bar to 296 K, 5 bar, facilitating comprehensive high-throughput screening of materials. Although minor numerical inaccuracies may exist, the GBR models offer valuable guidance for identifying high-performance materials and selecting the top candidates for H_2 storage applications.

3.4.2. Crystal Graph Convolutional Neural Network. In addition to the feature-based ML models, CGCNN models were trained to predict H_2 uptakes using only the crystal structure definition file as an input. Unlike the GBR models, which can predict uptake under specific conditions by

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

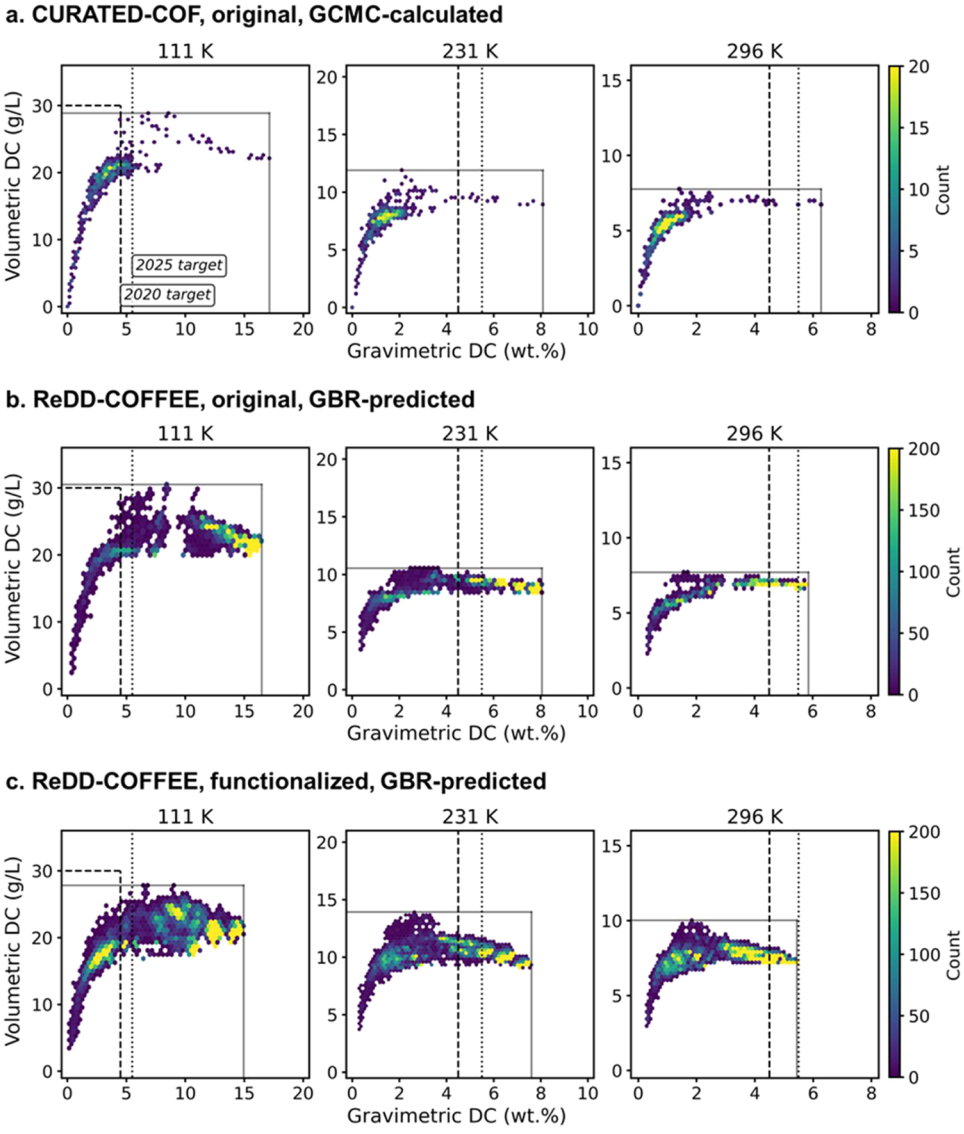


Figure 10. Gravimetric and volumetric DCs for (a) original structures from CURATED-COF database, (b) original and (c) functionalized structures from ReDD-COFFEE database at three different temperatures. Colors of the hexagonal bins represent the number of data points, as shown in the color bars. The dashed and dotted lines represent the 2020 and 2025 DOE targets. The solid lines represent the maximum values of the gravimetric and volumetric DC for each data set.

providing additional inputs, the CGCNN models generally do not allow for the inclusion of other inputs. Instead, in this work, the CGCNN models predicted a list of 30 values, representing the uptakes of original and functionalized structures with various functionalization types and ratios at three different temperatures and two pressure points. Two separate CGCNN models were developed for predicting gravimetric and volumetric uptakes, named CGCNN_g and CGCNN_v, respectively.

The performance of the CGCNN models, as shown in Figure 9, was lower than that of the GBR models, with lower R^2 values and higher MAE and RMSE metrics. For gravimetric H_2 uptake, the CGCNN model achieved an R^2 of 0.894, an MAE of 0.203 wt %, and an RMSE of 0.427 wt %. For volumetric uptake, it showed an R^2 of 0.912, an MAE of 1.344 g/L, and an RMSE of 2.585 g/L. The errors of the CGCNN predictions were nearly 3 times those of the GBR predictions. This discrepancy was even more pronounced when calculating the DC values from the CGCNN-predicted uptakes, resulting in lower R^2 values and significant deviations in MAE and RMSE.

The CGCNN models tended to overestimate volumetric uptakes for materials expected to have negligible adsorption. This overestimation was evident in the vertical clustering of data points near the origin for volumetric uptake, and DC predictions are shown in Figure 9. Efforts to exclude materials with low volumetric uptakes from the data sets proved challenging, as each target array included at least one low uptake value (<1 g/L), particularly at low pressures. This resulted in no data being available for training unless we excluded data at high temperatures like 296 K or included temperature and pressure as part of the inputs, as done for GBR model training.

The less accurate performance of the CGCNN models could be attributed to their reliance on mostly local features, such as atomic number and interatomic distance, without incorporating global structural features such as pore size, pore volume, density, and surface area. These global features are more

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

related to the overall adsorption performance and crucial for accurately predicting adsorption uptakes. Previous studies, such as MOFormer,[31] MOF-CGCNN,[67] etc., have shown that combining local and global features can enhance prediction accuracy for various properties, suggesting a potential path for improving CGCNN models. However, in our study, feature-based ML models, such as GBR, have already provided high accuracy using global textural properties and operational conditions.

In conclusion, while CGCNN models offer a novel approach by utilizing crystal structure definition files alone, their prediction accuracy for H_2 storage performance is currently limited compared to traditional feature-based ML models. The analysis underscored the importance of including global structural features to enhance prediction accuracy. Therefore, feature-based models remained more reliable for predicting the H_2 uptake and DC of functionalized COFs, providing valuable insights for the high-throughput screening and selection of top-performing materials.

3.5. Screening of Hypothetical COF Structures Using an ML Model. To further investigate the H_2 adsorption performance of various COFs before and after Mg-alkoxide functionalization, we used the GBR model developed in this work to quickly evaluate the performance of the hypothetical COFs from ReDD-COFFEE[68] database (DB). First, we identified 95,860 structures containing phenyl rings, excluding materials without the potential functionalization site from the original ReDD-COFFEE DB. Several key textural properties, such as PLD, LCD, surface area, void fraction, and density, were extracted from the ReDD-COFFEE DB and organized as input features for the GBR predictions. Figure 10 shows the ML model predicted H_2 pressure-swing DCs at various temperatures, both before and after functionalization.

The screening results show trends consistent with those observed in the CURATED-COF DB. At near cryogenic temperatures (111 K), both gravimetric and volumetric DCs decreased after functionalization. However, at higher temperatures (231 and 296 K), functionalization led to improvements in volumetric DCs for some materials. Some of the original and functionalized COF structures from the ReDD-COFFEE DB could meet gravimetric targets at ambient temperatures but not the volumetric targets. Additionally, functionalization did not significantly enhance volumetric DCs in this data set. When comparing with the CURATED-COF database, unfunctionalized materials from the ReDD-COFFEE DB exhibited relatively lower volumetric DCs at 231 and 296 K.

Notably, more COFs from the ReDD-COFFEE DB demonstrated higher gravimetric DCs, as represented by the light-yellow regions in Figure 10b,c. This contrasted with CURATED-COF, where most data points were concentrated in regions with lower gravimetric and volumetric DCs, and only a few materials showed strong performance with high DCs. This difference could be attributed to the distinct structural characteristics of COFs in the ReDD-COFFEE database, which generally feature larger pores and larger void fractions. As highlighted in previous studies and demonstrated by our comparison of structures 21320N2 and 07010N3, H_2 adsorption improvements tended to be more pronounced in materials with smaller pores, like 07010N3. In contrast, the larger pores and higher void fractions in many COFs from the ReDD-COFFEE DB resulted in lower volumetric H_2 adsorption performance with limited improvement following Mg-alkoxide functionalization.

It is important to acknowledge that the GBR model was trained using adsorption data obtained from the structures reported in the CURATED-COF DB, which might not fully capture the range of properties found in the ReDD-COFFEE DB, as shown in Figure S24. This limitation could lead to inaccuracies in the predictions for certain structures in ReDD-COFFEE. Moving forward, incorporating models that better learn from a broader range of features or expanding the training data set with more diverse materials could improve the predictive accuracy of H_2 storage performance.

4. CONCLUSIONS

In this work, we investigated Mg-alkoxide-functionalized COF structures for H_2 pressure-swing storage applications. H_2 binding energies were estimated using MP2-calculated interaction energies with Mg-catecholates and specialized FF models were developed for GCMC simulations. HT screening on the CURATED-COF database using these FF models shows that Mg-alkoxide functionalization does not enhance H_2 storage at 111 K but improves the performance at 231 and 296 K, particularly for volumetric capacity. The enhancement at higher temperatures is more pronounced in materials with smaller pores, where Mg-alkoxide functionalization significantly boosts the volumetric H_2 storage performance. Using the developed ML model, we screened hypothetical COF structures from ReDD-COFFEE database. The additional screening further confirmed that the Mg-alkoxide functionalization yields greater improvements in volumetric H_2 storage capacities for COFs with smaller pores compared with those with larger (mesoporous) pores. This conclusion provides valuable material design guidelines for improving the volumetric H_2 storage capacity under ambient conditions.

■ ASSOCIATED CONTENT

Data Availability Statement

The simulation files for Gaussian 16 and RASPA 2 and the Python scripts used in the ML prediction workflow are available in the following Github repository: https://github.com/mtap-research/data-repository/tree/master/2024-ACSAMI-H2-COF.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsami.4c11953.

Functionalized COFs. (ZIP)

Metrics used in this work; electronic structure calculations and FF parametrization details; molecular models construction details; high-throughput GCMC simulations of original and functionalized structures with specific FF models; GCMC simulations of original and functionalized structures with specific FF models of two selected materials; recent studies on the ML prediction of H_2 adsorption on COFs; traditional tree-based ML and crystal graph convolutional neural network (CGCNN) model training. (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Yongchul G. Chung − *School of Chemical Engineering, Pusan National University, Busan 46241, Republic of Korea;* orcid.org/0000-0002-7756-0589; Email: drygchung@gmail.com

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009

Authors

Yu Chen − *School of Chemical Engineering, Pusan National University, Busan 46241, Republic of Korea;* orcid.org/0000-0002-6530-0974

Guobin Zhao − *School of Chemical Engineering, Pusan National University, Busan 46241, Republic of Korea;* orcid.org/0000-0002-7728-4211

Sunghyun Yoon − *School of Chemical Engineering, Pusan National University, Busan 46241, Republic of Korea*

Parsa Habibi − *Engineering Thermodynamics, Process & Energy Department, Faculty of Mechanical Engineering, Delft University of Technology, 2628 CB Delft, The Netherlands*

Chang Seop Hong − *Department of Chemistry, Korea University, Seoul 02841, Republic of Korea;* orcid.org/0000-0002-4329-4745

Song Li − *Department of New Energy and Science Engineering, School of Energy and Power Engineering, Huazhong University of Science and Technology, Wuhan 430074, China*

Othonas A. Moultos − *Engineering Thermodynamics, Process & Energy Department, Faculty of Mechanical Engineering, Delft University of Technology, 2628 CB Delft, The Netherlands;* orcid.org/0000-0001-7477-9684

Poulumi Dey − *Materials Science and Engineering Department, Faculty of Mechanical Engineering, Delft University of Technology, 2628 CD Delft, The Netherlands;* orcid.org/0000-0003-4679-1752

Thijs J. H. Vlugt − *Engineering Thermodynamics, Process & Energy Department, Faculty of Mechanical Engineering, Delft University of Technology, 2628 CB Delft, The Netherlands;* orcid.org/0000-0003-3059-8712

Complete contact information is available at: https://pubs.acs.org/10.1021/acsami.4c11953

Author Contributions

Y.C.: software, methodology, formal analysis, data curation, visualization, and writing—original draft; G.Z.: software, methodology, and writing—original draft; S.Y.: software and methodology; P.H.: software, methodology, and writing—reviewing and editing; C.S.H.: funding acquisition; S.L.: writing—reviewing and editing; O.A.M.: software, methodology, and writing—reviewing and editing; P.D.: writing—reviewing and editing; T.J.H.V.: writing—reviewing and editing; Y.G.C.: conceptualization, writing—reviewing and editing, resources, supervision, project administration, and funding acquisition.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea Government (MSIT) (No. 2021M3I3A1084664 and RS-2024-00449431). The computational resources provided by the KISTI Supercomputing Center (Project No. KSC-2023-CRE-0166) are gratefully acknowledged. The authors thank Dr. Haoyuan Chen for the help with force field implementation in RASPA.

■ REFERENCES

(1) Ma, Y.; Wang, X. R.; Li, T.; Zhang, J. W.; Gao, J.; Sun, Z. Y. Hydrogen and ethanol: Production, storage, and transportation. *Int. J. Hydrogen Energy* **2021**, *46* (54), 27330−27348.

(2) Zhang, W. Y.; Fang, X. M.; Sun, C. W. The alternative path for fossil oil: Electric vehicles or hydrogen fuel cell vehicles? *J. Environ. Manage.* **2023**, *341*, No. 118019.

(3) Durbin, D. J.; Malardier-Jugroot, C. Review of hydrogen storage techniques for on board vehicle applications. *Int. J. Hydrogen Energy* **2013**, *38* (34), 14595−14617.

(4) Chalk, S. G.; Miller, J. F. Key challenges and recent progress in batteries, fuel cells, and hydrogen storage for clean energy systems. *J. Power Sources* **2006**, *159* (1), 73−80.

(5) Elberry, A. M.; Thakur, J.; Santasalo-Aarnio, A.; Larmi, M. Large-scale compressed hydrogen storage as part of renewable electricity storage systems. *Int. J. Hydrogen Energy* **2021**, *46* (29), 15671−15690.

(6) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The chemistry and applications of metal-organic frameworks. *Science* **2013**, *341*, No. 1230444.

(7) Rosi, N. L.; Eckert, J.; Eddaoudi, M.; Vodak, D. T.; Kim, J.; O'Keeffe, M.; Yaghi, O. M. Hydrogen storage in microporous metal-organic frameworks. *Science* **2003**, *300*, 1127−1129.

(8) Wu, H.; Simmons, J. M.; Srinivas, G.; Zhou, W.; Yildirim, T. Adsorption sites and binding nature of H2, CO2, and CH4 in metal-organic frameworks: A high-resolution inelastic neutron scattering and first-principles calculation study. *J. Phys. Chem. Lett.* **2010**, *1*, 1946−1951.

(9) Kim, D. W.; Chen, Y.; Kim, H.; Kim, N.; Lee, Y. H.; Oh, H.; Chung, Y. G.; Hong, C. S. High Hydrogen Storage in Trigonal Prismatic Monomer-Based Highly Porous Aromatic Frameworks. *Adv. Mater.* **2024**, *36* (26), No. 2401739.

(10) Dincă, M.; Haney, C. M.; Liu, Y.; Dailly, A.; Long, J. R. Hydrogen storage in a microporous metal-organic framework with exposed Mn2+ coordination sites. *J. Am. Chem. Soc.* **2006**, *128*, 16876−16883.

(11) Li, H.; Wang, K.; Sun, Y.; Lollar, C. T.; Li, J.; Zhou, H.-C. Recent advances in gas storage and separation using metal-organic frameworks. *Mater. Today* **2018**, *21*, 108−121.

(12) Lin, X.; Telepeni, I.; Blake, A. J.; Dailly, A.; Brown, C. M.; Simmons, J. M.; Zoppi, M.; Walker, G. S.; Thomas, K. M.; Mays, T. J.; et al. High capacity hydrogen adsorption in Cu(II) tetracarboxylate frameworks. *J. Am. Chem. Soc.* **2009**, *131*, 2159−2171.

(13) DOE Technical Targets for Onboard Hydrogen Storage for Light-Duty Vehicles. https://energy.gov/eere/fuelcells/doe-technical-targets-onboard-hydrogen-storage-light-duty-vehicles (accessed July 23, 2024).

(14) Canossa, S.; Wuttke, S. Functionalization Chemistry of Porous Materials. *Adv. Funct. Mater.* **2020**, *30* (41), No. 2003875.

(15) Colón, Y. J.; Snurr, R. Q. The effect of co-adsorbed solvent molecules on H(2) binding to metal alkoxides. *Phys. Chem. Chem. Phys.* **2019**, *21* (18), 9218−9224.

(16) Dincă, M.; Long, J. R. Hydrogen storage in microporous metal-organic frameworks with exposed metal sites. *Angew. Chem., Int. Ed.* **2008**, *47* (36), 6766−6779.

(17) Zhou, D. S.; Sun, H. F.; Guo, S. H.; Zhao, D. L.; Li, J.; Zhang, Y. H. Hydrogen storage properties of Mg-based alloys modified with metal-organic frameworks and carbon-based porous materials: A review and summary. *Int. J. Hydrogen Energy* **2024**, *57*, 1373−1388.

(18) Getman, R. B.; Miller, J. H.; Wang, K.; Snurr, R. Q. Metal Alkoxide Functionalization in Metal-Organic Frameworks for Enhanced Ambient-Temperature Hydrogen Storage. *J. Phys. Chem. C* **2011**, *115* (5), 2066−2075.

(19) Chen, H. Y.; Snurr, R. Q. Computational Screening of Metal-Catecholate-Functionalized Metal-Organic Frameworks for Room-Temperature Hydrogen Storage. *J. Phys. Chem. C* **2021**, *125* (39), 21701−21708.

(20) Himsl, D.; Wallacher, D.; Hartmann, M. Improving the Hydrogen-Adsorption Properties of a Hydroxy-Modified MIL-53(Al) Structural Analogue by Lithium Doping. *Angew. Chem., Int. Ed.* **2009**, *48* (25), 4639−4642.

(21) Clayson, I. G.; Hewitt, D.; Hutereau, M.; Pope, T.; Slater, B. High Throughput Methods in the Synthesis, Characterization, and

Optimization of Porous Materials. *Adv. Mater.* **2020**, *32* (44), No. 2002780.
(22) Schmidt, J.; Marques, M. R. G.; Botti, S.; Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **2019**, *5*, No. 83, DOI: 10.1038/s41524-019-0221-0.
(23) Zhao, G.; Chen, Y.; Chung, Y. G. High-Throughput, Multiscale Computational Screening of Metal-Organic Frameworks for Xe/Kr Separation with Machine-Learned Parameters. *Ind. Eng. Chem. Res.* **2023**, *62* (37), 15176−15189.
(24) Jablonka, K. M.; Ongari, D.; Moosavi, S. M.; Smit, B. Big-Data Science in Porous Materials: Materials Genomics and Machine Learning. *Chem. Rev.* **2020**, *120* (16), 8066−8129.
(25) Anderson, G.; Schweitzer, B.; Anderson, R.; Gómez-Gualdrón, D. A. Attainable Volumetric Targets for Adsorption-Based Hydrogen Storage in Porous Crystals: Molecular Simulation and Machine Learning. *J. Phys. Chem. C* **2019**, *123* (1), 120−130.
(26) Cho, E. H.; Lin, L.-C. Nanoporous Material Recognition via 3D Convolutional Neural Networks: Prediction of Adsorption Properties. *J. Phys. Chem. Lett.* **2021**, *12* (9), 2279−2285.
(27) Ahmed, A.; Siegel, D. J. Predicting hydrogen storage in MOFs via machine learning. *Patterns* **2021**, *2* (7), No. 100291.
(28) Borboudakis, G.; Stergiannakos, T.; Frysali, M.; Klontzas, E.; Tsamardinos, I.; Froudakis, G. E. Chemically intuited, large-scale screening of MOFs by machine learning techniques. *npj Comput. Mater.* **2017**, *3* (1), 40.
(29) Bucior, B. J.; Bobbitt, N. S.; Islamoglu, T.; Goswami, S.; Gopalan, A.; Yildirim, T.; Farha, O. K.; Bagheri, N.; Snurr, R. Q. Energy-based descriptors to rapidly predict hydrogen storage in metal−organic frameworks. *Mol. Syst. Des. Eng.* **2019**, *4* (1), 162−174.
(30) Thornton, A. W.; Simon, C. M.; Kim, J.; Kwon, O.; Deeg, K. S.; Konstas, K.; Pas, S. J.; Hill, M. R.; Winkler, D. A.; Haranczyk, M.; Smit, B. Materials genome in action: identifying the performance limits of physical hydrogen storage. *Chem. Mater.* **2017**, *29* (7), 2844−2854.
(31) Cao, Z. L.; Magar, R.; Wang, Y. Y.; Farimani, A. B. MOFormer: Self-Supervised Transformer Model for Metal-Organic Framework Property Prediction. *J. Am. Chem. Soc.* **2023**, *145*, 2958−2967.
(32) Rosen, A. S.; Iyer, S. M.; Ray, D.; Yao, Z.; Aspuru-Guzik, A.; Gagliardi, L.; Notestein, J. M.; Snurr, R. Q. Machine learning the quantum-chemical properties of metal−organic frameworks for accelerated materials discovery. *Matter* **2021**, *4* (5), 1578−1597.
(33) Zhao, G.; Chung, Y. G. PACMAN: A Robust Partial Atomic Charge Predicter for Nanoporous Materials Based on Crystal Graph Convolution Networks. *J. Chem. Theory Comput.* **2024**, *20* (12), 5368−5380.
(34) Desgranges, C.; Delhommelle, J. Ensemble Learning of Partition Functions for the Prediction of Thermodynamic Properties of Adsorption in Metal−Organic and Covalent Organic Frameworks. *J. Phys. Chem. C* **2020**, *124* (3), 1907−1917.
(35) Aksu, G. O.; Keskin, S. Advancing CH 4/H 2 separation with covalent organic frameworks by combining molecular simulations and machine learning. *J. Mater. Chem. A* **2023**, *11* (27), 14788−14799.
(36) Lundberg, S. M.; Erion, G.; Chen, H.; DeGrave, A.; Prutkin, J. M.; Nair, B.; Katz, R.; Himmelfarb, J.; Bansal, N.; Lee, S. I. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.* **2020**, *2* (1), 56−67.
(37) Head-Gordon, M.; Pople, J. A.; Frisch, M. J. MP2 energy evaluation by direct methods. *Chem. Phys. Lett.* **1988**, *153* (6), 503−506.
(38) Sæbø, S.; Almlöf, J. Avoiding the integral storage bottleneck in LCAO calculations of electron correlation. *Chem. Phys. Lett.* **1989**, *154* (1), 83−89.
(39) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A direct MP2 gradient method. *Chem. Phys. Lett.* **1990**, *166* (3), 275−280.
(40) Frisch, M. J.; Head-Gordon, M.; Pople, J. A. Semi-direct algorithms for the MP2 energy and gradient. *Chem. Phys. Lett.* **1990**, *166* (3), 281−289.
(41) Head-Gordon, M.; Head-Gordon, T. Analytic MP2 frequencies without fifth-order storage. Theory and application to bifurcated hydrogen bonds in the water hexamer. *Chem. Phys. Lett.* **1994**, *220* (1), 122−128.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Gaussian 16, Revision C.01. 2016, Gaussian, Inc.: Wallingford CT.
(43) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. Dreiding - a Generic Force-Field for Molecular Simulations. *J. Phys. Chem. A* **1990**, *94* (26), 8897−8909.
(44) Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. Uff, a Full Periodic-Table Force-Field for Molecular Mechanics and Molecular-Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024−10035.
(45) Darkrim, F.; Levesque, D. Monte Carlo simulations of hydrogen adsorption in single-walled carbon nanotubes. *J. Chem. Phys.* **1998**, *109* (12), 4981−4984.
(46) Xiang, Y.; Sun, D. Y.; Fan, W.; Gong, X. G. Generalized simulated annealing algorithm and its application to the Thomson model. *Phys. Lett. A* **1997**, *233* (3), 216−220.
(47) Ongari, D.; Yakutovich, A. V.; Talirz, L.; Smit, B. Building a Consistent and Reproducible Database for Adsorption Evaluation in Covalent-Organic Frameworks. *ACS Cent. Sci.* **2019**, *5* (10), 1663−1675.
(48) Ongari, D.; Talirz, L.; Smit, B. Too Many Materials and Too Many Applications: An Experimental Problem Waiting for a Computational Solution. *ACS Cent. Sci.* **2020**, *6* (11), 1890−1900.
(49) Boone, P.; Wilmer, C. E. MOFUN: a Python package for molecular find and replace. *Digital Discovery* **2022**, *1* (5), 679−688.
(50) Henle, E. A.; Gantzler, N.; Thallapally, P. K.; Fern, X. Z.; Simon, C. M. PoreMatMod.jl: Julia Package for Postsynthetic Modification of Crystal Structure Models. *J. Chem. Inf. Model.* **2022**, *62* (3), 423−432.
(51) Ongari, D.; Boyd, P. G.; Barthel, S.; Witman, M.; Haranczyk, M.; Smit, B. Accurate Characterization of the Pore Volume in Microporous Crystalline Materials. *Langmuir* **2017**, *33* (51), 14529−14538.
(52) Pinheiro, M.; Martin, R. L.; Rycroft, C. H.; Haranczyk, M. High accuracy geometric analysis of crystalline porous materials. *CrystEngComm* **2013**, *15* (37), 7531−7538.
(53) Willems, T. F.; Rycroft, C.; Kazi, M.; Meza, J. C.; Haranczyk, M. Algorithms and tools for high-throughput geometry-based analysis of crystalline porous materials. *Microporous Mesoporous Mater.* **2012**, *149* (1), 134−141.
(54) Ismail, A. F.; Khulbe, K. C.; Matsuura, T.Fundamentals of Gas Permeation Through Membranes. In *Gas Separation Membranes: Polymeric and Inorganic*; Wiley, 2015; pp 11−35.
(55) Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. RASPA: molecular simulation software for adsorption and diffusion in flexible nanoporous materials. *Mol. Simul.* **2016**, *42* (2), 81−101.
(56) Allen, M. P.; Tildesley, D. J. *Computer Simulation of Liquids*; Oxford University Press, 2017.
(57) Limas, N. G.; Manz, T. A. Introducing DDEC6 atomic population analysis: part 4. Efficient parallel computation of net atomic charges, atomic spin moments, bond orders, and more. *RSC Adv.* **2018**, *8* (5), 2678−2707.

(58) Manz, T. A. Introducing DDEC6 atomic population analysis: part 3. Comprehensive method to compute bond orders. *RSC Adv.* **2017**, *7* (72), 45552−45581.
(59) Limas, N. G.; Manz, T. A. Introducing DDEC6 atomic population analysis: part 2. Computed results for a wide range of periodic and nonperiodic materials. *RSC Adv.* **2016**, *6* (51), 45727−45747.
(60) Manz, T. A.; Limas, N. G. Introducing DDEC6 atomic population analysis: part 1. Charge partitioning theory and methodology. *RSC Adv.* **2016**, *6* (53), 47771−47801.
(61) Xie, T.; Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. *Phys. Rev. Lett.* **2018**, *120* (14), No. 145301.
(62) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825−2830.
(63) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Comput. Mater. Sci.* **2013**, *68*, 314−319.
(64) Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.et al. In *PyTorch: An Imperative Style, High-Performance Deep Learning Library*, Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019.
(65) Bu, R.; Zhang, L.; Liu, X.-Y.; Yang, S.-L.; Li, G.; Gao, E.-Q. Synthesis and Acid-Responsive Properties of a Highly Porous Vinylene-Linked Covalent Organic Framework. *ACS Appl. Mater. Interfaces* **2021**, *13* (22), 26431−26440.
(66) El-Kaderi, H. M.; Hunt, J. R.; Mendoza-Cortés, J. L.; Côté, A. P.; Taylor, R. E.; O'Keeffe, M.; Yaghi, O. M. Designed Synthesis of 3D Covalent Organic Frameworks. *Science* **2007**, *316* (5822), 268−272.
(67) Wang, R.; Zou, Y.; Zhang, C.; Wang, X.; Yang, M.; Xu, D. Combining crystal graphs and domain knowledge in machine learning to predict metal-organic frameworks performance in methane adsorption. *Microporous Mesoporous Mater.* **2022**, *331*, No. 111666.
(68) De Vos, J. S.; Borgmans, S.; Van Der Voort, P.; Rogge, S. M.; Van Speybroeck, V. ReDD-COFFEE: a ready-to-use database of covalent organic framework structures and accurate force fields to enable high-throughput screenings. *J. Mater. Chem. A* **2023**, *11* (14), 7468−7487.

https://doi.org/10.1021/acsami.4c11953
ACS Appl. Mater. Interfaces 2024, 16, 61995−62009