MOFs with the Stability for Practical Gas Adsorption Applications Require New Design Rules

Changhwan Oh[1,2], Aditya Nandy[1,3], Shuwen Yue[1], and Heather J. Kulik[1,3]*

[1]*Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

[2]*Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

[3]*Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

*email: hjkulik@mit.edu

ABSTRACT: Metal-organic frameworks (MOFs) have been widely studied for their ability to capture and store greenhouse gases. However, most chemical discovery efforts use computational study of hypothetical MOFs without consideration of their stability, limiting the practical application of novel materials. We overcome this limitation by screening hypothetical ultrastable MOFs that have predicted high thermal and activation stability, as judged by machine learning (ML) models trained on experimental measures of stability. We enhance this set by computing the bulk modulus as a measure of mechanical stability and filter 1,102 mechanically robust hypothetical MOFs from a database of ultrastable MOFs (USMOF DB). Grand Canonical Monte Carlo simulations are then employed to examine the gas adsorption properties of these hypothetical MOFs, alongside a database of experimental MOFs. We identify privileged building blocks that allow MOFs in USMOF DB to show exceptional working capacities compared to the experimental MOFs. We interpret these differences by training ML models on CO_2 and CH_4 adsorption in these databases, showing how poor model transferability between datasets indicates that novel design rules can be derived from USMOF DB that would not have been gathered through assessment of structurally characterized MOFs. We identify geometric features and node chemistry that will enable the rational design of MOFs with enhanced gas adsorption properties in synthetically realizable MOFs.

1. Introduction.

In recent years, carbon emissions have been on the rise, with 35.8 Gt emitted globally in 2023.[1,2] These escalating levels of greenhouse gases in the atmosphere demand immediate attention as they contribute significantly to climate change.[3] To mitigate these effects, adsorption strategies for capturing and storing greenhouse gases are being studied.[4] Pressure or temperature swing adsorption using porous adsorbents capable of capturing and releasing gases has been proposed as a solution. Metal-organic frameworks[5,6] (MOFs) are an exemplary class of porous materials that have demonstrated promise not just for gas storage[7-14] but also for separation[15-19] and catalysis,[20-24] due to their unique porous structures[25] and tunable properties.[26,27] With their high surface areas and customizable pore sizes, MOFs offer excellent potential for adsorbing and storing greenhouse gases such as methane (CH_4) and carbon dioxide (CO_2). The reticular nature of these MOF materials, consisting of inorganic secondary building units (SBUs) and organic linkers, leads to a large combinatorial space. However, a major limitation hindering the widespread application of MOFs is their poor stability,[28-30] particularly under working conditions for gas adsorption applications, such as high pressure and temperature. Addressing this stability issue is crucial for ensuring the practical viability of MOFs in greenhouse gas capture and other environmental remediation efforts.[31-33]

The large space of candidate MOFs makes experimental identification of highly stable MOFs with target properties challenging and time-consuming. Therefore, computational screening methods have gained prominence as efficient tools for discovering novel MOFs.[34-36] Several databases (DBs) of MOFs exist for computational screening, providing a diverse pool of MOF structures for exploration. Experimental MOFs[37] have been curated by sanitizing single crystal structures of MOFs available from the Cambridge Structural Database (CSD)[38], whereas

hypothetical MOFs[35,36,39-41] are constructed by exhaustively enumerating combinations of building blocks. Some hypothetical DBs include the over 130,000 structures in the pioneering hMOF set,[35] 300,000 structures in BW-DB,[36] ~13,000 diverse structures in ToBaCCo,[39] and around 280,000 structures in ARC-MOF, which aims to unify these disparate sets[40]. Researchers have also curated and screened experimentally characterized MOFs in the roughly 10,000 MOFs identified in the computation-ready experimental MOF 2019 (CoRE MOF 2019) DB[37]. In comparison to experimental MOFs, hypothetical MOFs tend to lack diversity[42] as well as stability.[43]

The ultrastable MOF database (USMOF DB) is a hypothetical database curated by Nandy et al.,[43] which fills this stability gap by virtue of its construction process. The building blocks from the experimental CoRE MOF 2019 DB were extracted and recombined to make a new hypothetical dataset of 54,139 MOFs that comprised of an order of magnitude more topologies and SBUs than any prior hypothetical set. Among these MOFs, 9,524 MOFs were predicted to be "ultrastable" (i.e., activation stable and >1 standard deviation above the average thermal degradation temperature of 359 °C) according to previously trained ML models for thermal and activation stability.[30] The USMOF DB exhibits a 10-fold increase in the number of "ultrastable" MOF structures compared to existing hypothetical databases, while retaining metal diversity compared to the experimental CoRE MOF 2019 DB, making the USMOF DB a valuable resource for computational screening. In addition to thermal and activation stability, mechanical stability is also crucial for MOFs,[44-51] as hydrostatic pressure can lead to amorphization, affecting their porosity and performance. Despite its importance, mechanical stability is one of several crucial measures rarely considered[52] in computational screening efforts.

Previous computational studies have extensively explored the gas adsorption properties of MOFs. Molecular simulations, particularly Grand Canonical Monte Carlo (GCMC) simulations,

have been widely used to predict gas adsorption behavior, such as adsorption isotherms,[53,54] separation,[53,55-58] or working capacity[59,60] with high accuracy for both experimental and hypothetical databases. For experimental database screening, GCMC simulations on the earlier CoRE MOF 2014 database have been used to study the adsorption behavior of mixtures to identify MOFs with high selectivities of CO_2 over H_2O[61] as well as for CO_2/N_2 separation with pressure swing adsorption (PSA).[62] Hypothetical databases are also frequently studied, including the study of MOFs in ToBaCCo for high H_2 deliverable capacities.[63] However, performing molecular simulations on all of the MOFs in these databases can be computationally very demanding. Even when it is straightforward to carry out GCMC simulations exhaustively, hypothetical databases often lack consideration of stability issues, with few exceptions[43], limiting the validity of the structure–property relationship across experimentally realizable materials. Furthermore, design rules for gas adsorption, particularly for greenhouse gases such as CH_4 and CO_2, are not as well-established as those for H_2 adsorption.[64,65]

As a complement to simulation, machine learning (ML) models have emerged as valuable tools for screening high-performance MOFs. ML-based studies using supervised learning have been used for quantitative structure–property relations (QSPRs),[66] prediction of mechanical stability,[52] synthesis condition predictions,[67] and gas adsorption studies.[68,69] Descriptor selection is key in ML model training for MOF adsorption predictions,[69-73] and geometric descriptors have been identified as suitable for H_2 or CH_4[74] but the same descriptors routinely fail for CO_2.[75] As an alternative, complex black-box ML models, such as deep neural networks (DNN)[76] or a crystal graph convolutional neural network (CGCNN)[77] have demonstrated high performance for predicting gas adsorption properties of MOFs, albeit at the cost of interpretability of the models.[78] With recent advancements in natural language processing, transformer[79] models have increasingly

been used in chemistry applications[80]. For example, MOFormer[81] and MOFTransformer[82] have been developed to predict diverse properties including gas adsorption properties based on the string representation of MOF (MOFid) with high precision. Deep learning has also been used for screening CH_4 adsorption of around 247 trillion hypothetical MOFs with 1775 topologies,[83] as well as genetic algorithms to optimize these properties.[84] Choosing appropriate database and feature selection are also critical in ML-based studies.[85] Hypothetical MOF databases are valuable in the sense that they can provide large amount of training data, but the lack of similarity to experimental MOFs[42] often leads to poor transferability. That is, when ML models trained on hypothetical data are applied to a database of experimental MOFs to choose an existing MOF to deploy for a chosen task, they are likely to be error prone.

Here, we directly address the stability challenge in screening materials for CO_2 and CH_4 storage across a set of the hypothetical USMOF structures that capture the diversity of previously synthesized MOFs. Starting from the known subset of 9,524 ultrastable (i.e., in terms of thermal and activation stability) USMOF DB structures, we compute the mechanical strength and select the top 1,102 of these structures in terms of mechanical stability. We obtain CO_2 and CH_4 adsorption properties using GCMC and compare properties of these mechanically stable USMOF DB MOFs to experimental CoRE MOF 2019 DB structures. By training interpretable ML models, we identify and explain unexpected trends in the USMOF DB compared to the experimental CoRE MOF 2019 DB, namely a number of MOFs preferring CH_4 over CO_2 in the USMOF DB. Based on insights gained from these ML models, we propose design rules based on geometric features and node chemistry for MOFs in the USMOF DB to achieve alternately high or low relative CO_2 vs. CH_4 working capacity.

2. Computational Details

2a. Dataset

We screened a hypothetical MOF database referred to as the USMOF DB initially constructed in Ref. [43]. The building blocks of MOFs used in this prior study were extracted from the CoRE MOF 2019 DB[37] and recombined in new combinations on a large number of topologies to create a database of novel MOFs using PORMAKE[83]. In this work, we employ the subset of 9,524 hypothetical "ultrastable" MOFs, which were predicted to possess high thermal stability and stability upon activation. Ultrastability is defined[43] as a thermal decomposition temperature (T_d) greater than one standard deviation (87°C) above the mean T_d (359°C) in the experimental data set (i.e., T_d of 446°C or higher), along with a ML-predicted[30] probability of the structure being stable exceeding 50%. As in prior work, we define edges (i.e., organic building blocks with two connection points) and nodes (i.e., any organic or inorganic component with more than two connection points) instead of linkers and SBUs to provide a clear breakdown of a MOF structure into mutually exclusive building blocks. The 9,524 "ultrastable" subset of USMOF DB consists of MOFs with different configurations of inorganic nodes, organic nodes, and organic edges: (1) one inorganic node and one edge (1inor-1edge, 5,213 MOFs), (2) one inorganic node, one organic node, and one edge (1inor-1org-1edge, 2,918 MOFs), and (3) two inorganic nodes and one edge (2inor-1edge, 1,393 MOFs). Calculations of mechanical properties, which were carried out with the LAMMPS package[86] (ver. 29Sep2021) in prior work[43], were further used to narrow down the set of structures examined in this work. For the 9,524 ultrastable MOFs, the Voigt-Reuss-Hill bulk modulus (K_{VRH}), a measure of mechanical stability, was previously successfully obtained for 7,330 materials. In the present work, we define "ultrastable" mechanical behavior as a MOF with K_{VRH} exceeding 5 GPa. The mechanically stable subset includes 1,102 MOFs, categorized by their connectivity and composition as follows: (1) one inorganic node and one edge (722 MOFs, 65.5%),

(2) one inorganic node, one organic node, and one edge (107 MOFs, 9.7%), and (3) two inorganic nodes and one edge (24.8%) (Figure 1 and Supporting Information Table S1).

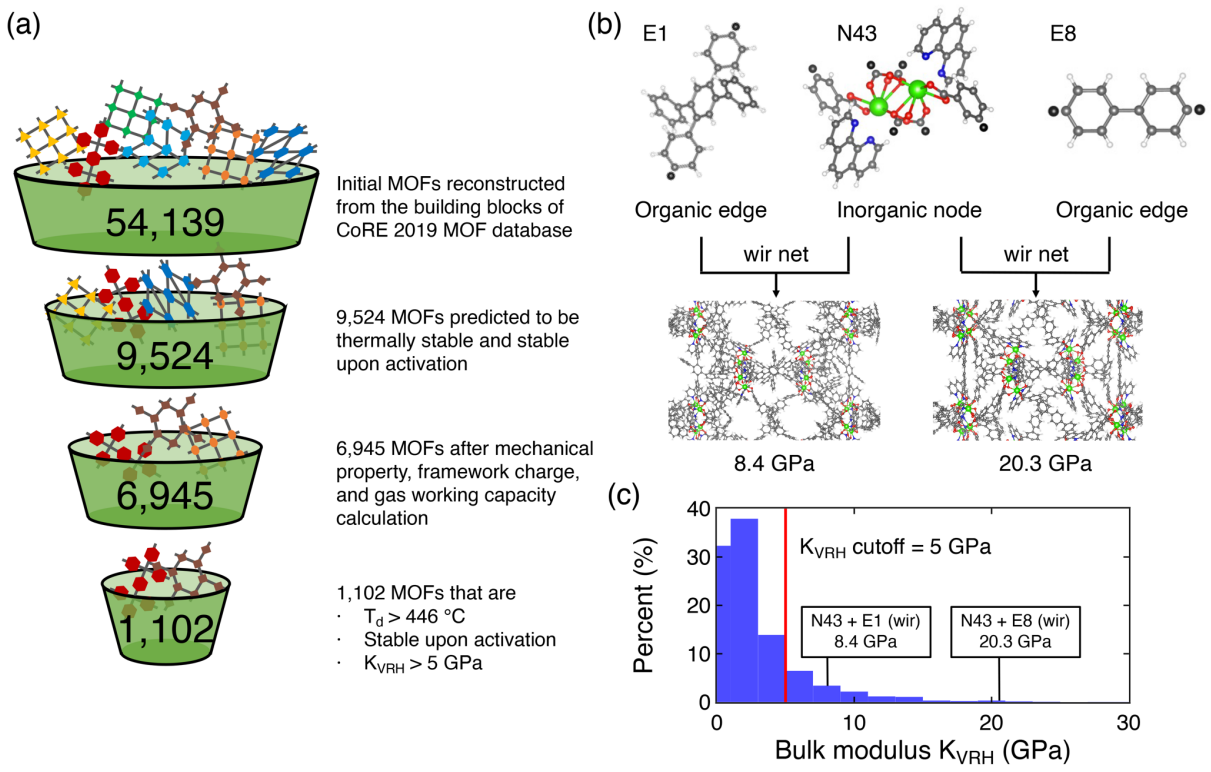

Figure 1. (a) Funnel showing number of hypothetical MOF structures after each screening process. (b) Representative MOFs in the USMOF DB with **wir** net and Gd-based node N43 but with different edges that result in different bulk moduli. C atoms are colored as gray, N atoms as blue, O atoms as red, H atoms as white, Gd atoms as light green, and connection points as black. (c) Distribution of the bulk modulus (K_{VRH}) for MOFs in the USMOF DB. A 5 GPa threshold for mechanically stable structures is depicted as a red line. The bulk moduli of two representative MOFs in (b) are labeled.

2b. Gas Working Capacity Simulations

We performed GCMC simulations using the RASPA2 simulation package[87] for calculating CO_2 and CH_4 uptake in the 7,330 "ultrastable" MOF structures with computed mechanical properties from the USMOF DB. We measured each gas uptake at two different pressures, and the working capacity was obtained by subtracting the uptake at low pressure from the uptake at high

pressure. The pressure range for CH_4 was between 5.8 bar and 65 bar and CO_2 was between 0.15 bar and 16 bar, which were chosen based on the standard CH_4 and CO_2 adsorption and desorption pressures[7]. The temperature was kept at 298 K and a rigid framework assumption was used for all gas uptake simulations. All simulations comprised 5,000 initialization cycles followed by 5,000 equilibration cycles. From all attempted calculations, 385 calculations failed or took longer than 48 hours, leaving 6,945 MOFs after all properties (i.e., mechanical stability, framework charge, and gas working capacity) calculated (Figure 1).

In modeling both gas–gas and framework–gas interactions, Lennard-Jones interactions were truncated and shifted at 12.8 Å. Partial atomic charges of the hypothetical MOF frameworks were calculated using the EQeq method[88] and saved in each MOF .CIF file. Coulombic interactions were computed using these charges with Ewald summation with a default real-space cutoff of 12 Å. We used the Universal Force Field (UFF)[89] and Transferable Potentials for Phase Equilibria (TraPPE)[90] for gas–gas and framework–gas interactions, respectively. The cross-interactions were calculated using the Lorentz–Berthelot mixing rule.[91] The single unit cell was duplicated in select cases to fulfill the requirements of cell lengths to exceed twice the interaction cutoffs, which was required only in a small subset of MOFs typically with one inorganic node and one edge needing up to three duplicates. A complete set of simulation input files is provided in a Zenodo repository.[92]

2c. Machine Learning Model Training

Using gas uptake data, we fit linear regression and random forest models using the scikit-learn 1.3.0 Python package[93]. For all models, 20% of the data were randomly selected as a set-aside test set. From the remaining 80% training set, an exhaustive grid search was used for 5-fold

cross-validation (CV) hyperparameter optimization (Supporting Information Table S2). The error metric for regression tasks was the mean absolute error (MAE).

Geometric features of the MOF databases were computed using the Zeo++ package[94,95]. We used a nitrogen probe molecule with 1.86 Å radius and obtained 14 geometric features (Supporting Information Table S3). For chemical features, we computed revised autocorrelations (RACs)[96] of MOFs using molSimplify[97]. RACs are a class of molecular descriptors that contain relationships of the heuristic atomic properties (i.e., nuclear charge (Z), identity (I), topology (T), electronegativity (χ), and covalent radius (S)). The products and differences of these atomic properties are calculated over a bond difference, with a maximum bond path (d) of three. The scope of the RACs considered include the metal (metal-centered RACs), linker coordinating atoms (linker-connecting RACs), functional group (functional group RACs), full linker (full-scope linker RACs), and full unit cell (full-scope RACs). With a maximum depth of three, there are 160 possible RACs, and removing constant features leaves a total of 134 RACs (Supporting Information Table S4).

The relative importance of each feature was calculated through impurity-based feature importance using scikit-learn. We used Gini importance[98], which determines the significance of a feature by calculating the normalized overall improvement in the criterion achieved by that particular feature.

3. Results and Discussion.

3a. Gas Adsorption Properties in USMOF DB

Along with the thermal and activation stability targeted when constructing USMOF DB, mechanical stability should also be considered for gas storage applications that often involve

applying hydrostatic pressure to the structure.[45,99] From a number of possible measures of mechanical stabilities, such as elastic modulus, shear modulus, or hardness, we used the Voigt-Reuss-Hill average value of bulk elastic modulus (K_{VRH}).[52] We chose K_{VRH} as the representative measure after determining that the bulk modulus and shear modulus (G_{VRH}) show linear correlation with Pearson's $r = 0.88$ in USMOF DB (see Sec. 2 and Supporting Information Figure S1). Due to the high porosity and low density of MOFs, they are relatively soft materials. Although there exist some extreme cases such as the UiO-66 MOF with computed bulk moduli up to 40 GPa[47], typical experimentally realized MOFs fall into the range of 4–30 GPa.[100] Analyzing the structural features of MOFs with K_{VRH} less than 5 GPa, we note that most are unphysically porous with cell volumes and pore volumes up to 10x that of the more stable (i.e., $K_{VRH} > 5$ GPa) structures (Supporting Information Figure S2). MOFs with such large pore volumes are more prone to structural distortion from hydrostatic compression, leading to pressure-induced amorphization.[44] We thus classify a MOF with K_{VRH} greater than 5 GPa as a "mechanically stable" MOF. From a total of 9,524 MOFs in the USMOF DB with "ultrastable" thermal and activation stability, 1,102 MOFs remain after applying this mechanical stability criterion of $K_{VRH} > 5$ GPa. Importantly, upon applying this threshold for mechanical stability, the metal diversity within USMOF DB remains relatively unchanged, with notable enrichment in Co (Figure 2). This enrichment is due to the high survival rate of Co-containing node N41 and N18 (57.8% and 52.5%, respectively) upon applying the mechanical stability criterion (see Sec. 3c).

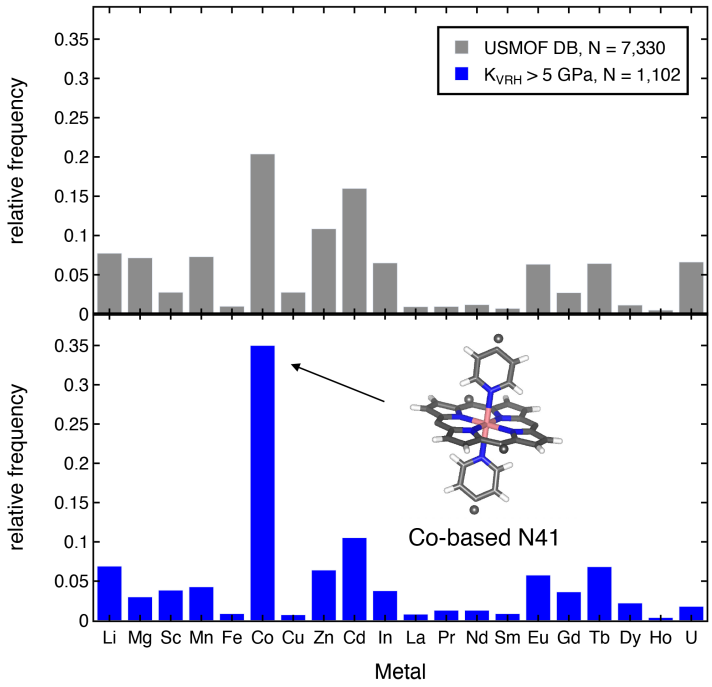

Figure 2. (Top) Distribution of metal atoms in 7,330 thermal- and activation- stable MOFs in USMOF DB for which K_{VRH} could be computed. (Bottom) Distribution of metal atoms in 1,102 thermal-, activation, and mechanically-stable MOFs in USMOF DB. A representative Co-based node (N41) is shown. C atoms are colored as gray, N atoms as blue, H atoms as white, Co atoms as pink, and connection points are shown as black spheres.

We next calculated the gas working capacity (i.e., difference between uptake at low and high pressure, see Sec. 2) of the mechanically stable USMOF DB MOFs using GCMC simulations for both CO_2 and CH_4 (Supporting Information Figure S3). We observe a number of MOFs in USMOF DB that have a CO_2:CH_4 working capacity ratio approaching 0.268 (Figure 3). Notably, this number is the ratio of the standard CO_2 working pressure (i.e., 16 bar – 0.15 bar = 15.85 bar) to the CH_4 standard working pressure (i.e., 65 bar – 5.8 bar = 59.2 bar). These MOFs with CO_2:CH_4 ratio of approximately 0.268 have large pore volumes where gas molecules can reside without interactions with the framework, leading to gas uptake directly proportional to the ratio in pressure differences, with small deviations (i.e., an offset of 100 cm^3(STP)/g framework obtained from fitting) from this behavior due to π-π stacking interactions between the CO_2 molecule and the MOF

pore walls[101]. This trend is even more evident in USMOF DB MOFs prior to applying the high mechanical stability (i.e., K_{VRH} > 5 GPa) criterion because MOFs with extremely large pores often have poor mechanical stabilities and are thus excluded.

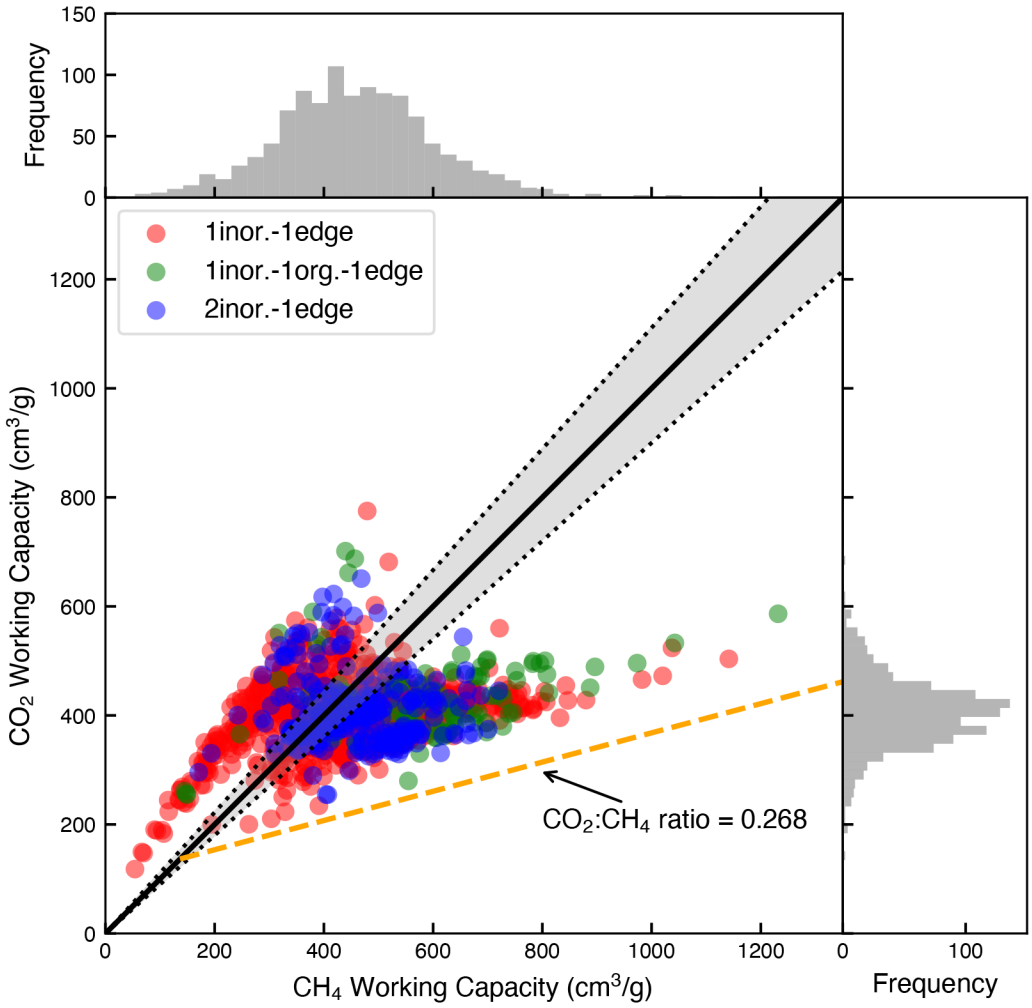

Figure 3. CO_2 working capacity vs. CH_4 working capacity (in cm^3(STP)/g framework) of 1,102 mechanically stable MOFs in USMOF DB. The data is colored by the connectivity of the building blocks: one inorganic node and one edge (1inor.-1edge, in red); one inorganic node, one organic node, and one edge (1inor.-1org.-1edge, in green); and 2 inorganic nodes and one edge (2inor.-1edge, in blue). The dotted lines depict the CO_2:CH_4 ratio of 1.1 and 0.9 for classification of MOF gas adsorption preference. The CO_2:CH_4 ratio = 0.268 and offset of 100 cm^3(STP)/g framework is depicted as a dashed orange line. The distributions of the CO_2 working capacity and CH_4 working capacity are shown as 1D unnormalized histograms as well with a bin width of 30 cm^3(STP)/g framework.

The average and distribution of CO_2 working capacities are similar for all three configurations of node and edges (average for 1inor-1edge: 414.19, 1inor-1org-1edge: 438.13, and 2inor-1edge: 414.81 cm^3(STP)/g framework, Supporting Information Figure S4). However, the 1inor-1org-1edge configuration has a higher average CH_4 working capacity (670.01 cm^3(STP)/g framework) compared to 1inor-1edge and 2inor-1edge configurations (513.47 and 529.33 cm^3(STP)/g framework). This dependence of working capacities on the building block configurations can be understood in the context of geometric factors that differ between CH_4 and CO_2 (see discussion in *Sec. 3c*). In particular, MOFs with 1inor-1org-1edge configuration have 50% greater average gravimetric pore volume (4.70 cm^3/g) compared to the other configurations (3.04 cm^3/g for 1inor-1edge and 3.05 cm^3/g for 2inor-1edge configuration).

We next compared the computed working capacities of CO_2 and CH_4 for the "ultrastable" subset of USMOF DB to those of the experimental CoRE MOF 2019 DB[37]. We used the adsorption data calculated by Moosavi et al.[42] for the latter, where they successfully calculated adsorption properties of 9,525 MOFs. We first classify the MOFs into three groups according to their CO_2:CH_4 working capacity ratios: i) high, where the ratio is greater than 1.1, ii) low, where the ratio is less than 0.9, and iii) intermediate, where ratio is between 0.9 and 1.1. While most of the MOFs from the CoRE MOF 2019 DB (89.4%) have a high ratio (i.e., greater than 1.1), only 27.5% of MOFs in the mechanically stable, "ultrastable" subset of the USMOF DB fall into this category (Supporting Information Table S5). We also note that the working capacity ranges are different between the two databases. Specifically, the majority of MOFs (92.7%) in CoRE MOF 2019 have a CH_4 working capacity below 200 cm^3(STP)/g framework, while very few MOFs (3%) in USMOF DB have a working capacity below that threshold. A similar trend is observed for CO_2 working capacity, as 77.8% of MOFs in CoRE MOF 2019 and only 7 MOFs (0.6%) in the

mechanically "ultrastable" subset of the USMOF DB have a working capacity below 200 cm^3(STP)/g framework for CO_2. Overall, USMOF DB contains some MOFs with exceptionally high CH_4 working capacity, while CoRE MOF 2019 predominantly contains MOFs with preferentially high CO_2 working capacity, albeit with lower working capacity overall than the top-performing MOFs in USMOF DB. Thus, the hypothetical MOFs in USMOF DB can be candidate materials with higher working capacities than MOFs in the CoRE MOF 2019 DB for both CH_4 and CO_2, but many exhibit higher CH_4 working capacities than CO_2.

3b. ML Models for Gas Working Capacity

We next sought to train interpretable machine learning (ML) models to identify essential features that distinguish working capacity for the two gas molecules in the two databases. In order to select the suitable machine learning (ML) model for our study, we performed initial comparison of model accuracy across models with different complexities: linear regression, random forest (RF), and artificial neural network (ANN) models. For the featurization, we use geometry features computed from Zeo++ and RACs, atomic property-centric features computed on the graph of the MOF (see Sec. 2). In the case of CH_4 working capacity for USMOF DB, all models showed high train and test R^2 values (i.e., exceeding 0.97, Supporting Information Table S6). Models predicting the mechanically stable USMOF DB CO_2 working capacity, on the other hand, yielded lower test R^2 values (linear regression: 0.77, RF: 0.83, and ANN: 0.78). Because predicting CO_2 working capacity is the more challenging learning task, we selected an RF model due to its trade-off of accuracy and interpretability.

The RF regressor model trained to predict CH_4 working capacity demonstrates excellent performance for both the mechanically stable subset of USMOF DB and CoRE MOF 2019,

achieving high accuracy (test R^2 = 0.99 and 0.97, respectively) and low mean absolute error (MAE) of test prediction (Supporting Information Figure S5). The models predicting CO_2 working capacity for both databases exhibited somewhat lower test R^2 values, 0.83 and 0.97 for USMOF DB and CoRE MOF 2019 DB, respectively, when compared to the models predicting CH_4 working capacity (Supporting Information Figure S5). The lower accuracy of the CO_2 model trained on USMOF DB than that trained on CoRE MOF 2019 is due to the mechanically stable USMOF DB subset being both smaller (i.e., 1,102 for mechanically stable MOFs in the USMOF DB and 9,525 for CoRE MOF 2019) and having a much wider range of geometric properties (Supporting Information Figure S7).

Next, we carried out feature importance analysis on the RF regressor model trained to predict the CH_4 working capacity of mechanically stable MOFs in USMOF DB or the MOFs in CoRE MOF 2019 (Supporting Information Figure S6). The impurity-based feature importances[98], which we calculated for both databases, underscore the significance of geometric features over chemical features in predicting the CH_4 working capacity. Notably, the gravimetric pore volume emerged as the most important feature, with a feature importance of 89.9% for USMOF DB and 94.7% for CoRE MOF 2019, highlighting how MOFs with large pore volumes can accommodate more CH_4 gas molecules (Supporting Information Figure S6). The smaller, spherical size and absence of dipole and quadrupole moments in the CH_4 molecule restrict its interaction with framework atoms, resulting in a minor contribution from chemical features. This aligns with the Chahine rule[64,65] where hydrogen adsorption is proportional to the BET surface area. Since MOFs in USMOF DB have large pore volumes, CH_4 adsorption is proportional to pore volume rather than surface area (Figure 4). For example, MOF-399 (CSD[38] refcode: BAZGAM) has been identified previously in the literature[102] for its exceptional gas uptake, but its relatively large

gravimetric pore volume of 7.39 cm^3/g is eclipsed by a hypothetical MOF in USMOF DB with Mg-based inorganic node (N54) and long organic edge (E13), which is derived from [p-terphenyl]-4,4"-dicarboxylic acid (H2TPDC, see discussion in Sec. 3c) that possesses a large gravimetric pore volume of 13.67 cm^3/g and exceptionally high CH_4 working capacity of 1141.40 cm^3/g (Figure 4 and Supporting Information Figure S8). Thus, there is a relatively consistent emphasis on pore volume to predict CH_4 uptake when models are trained on either dataset, but the pore sizes are even larger in the mechanically stable subset of USMOF DB than in previously synthesized CoRE MOF 2019 MOFs.

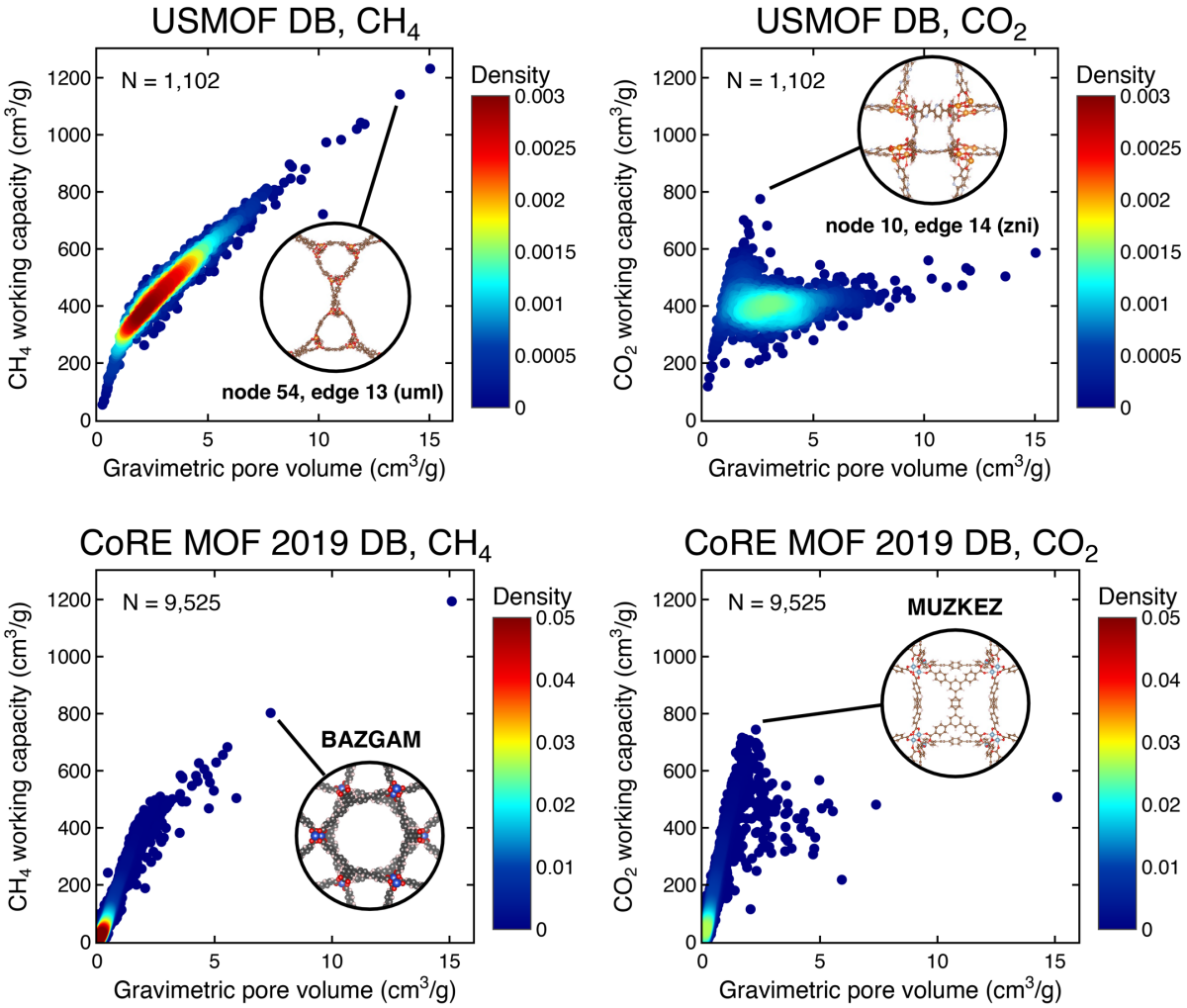

Figure 4. Kernel density estimate (KDE) plots of CH_4 and CO_2 working capacity (in cm^3(STP)/g

framework) vs. gravimetric pore volume (GPOV, in cm^3/g) and top-performing MOF structures for the 1,102 MOFs in the mechanically stable subset of USMOF DB and 9,525 MOFs in CoRE MOF 2019. Higher density is colored in red, while lower density is colored in blue following inset color bars that have distinct ranges for the two datasets.

In comparison to CH_4, the CO_2 molecule possesses a non-zero quadrupole moment and a larger size with a linear shape. Feature importance analysis for the model predicting CO_2 working capacity of MOFs in CoRE MOF 2019 highlights the significance of geometric features, particularly gravimetric pore volume with a feature importance percentage of 89.3% for predicting CO_2 working capacity (Supporting Information Figure S6). Conversely, for the mechanically stable but large-pore MOFs in USMOF DB, both the geometric features including surface area (40.0% importance) and chemical RACs, specifically metal-centered electronegativity RACs (17.3% importance), play crucial roles (Supporting Information Figure S6). Across CoRE MOF 2019, there is a stronger correlation between CO_2 working capacity and gravimetric pore volume ($R^2 = 0.66$) while the mechanically stable MOFs in USMOF DB do not follow this trend as strongly (Figure 4). This difference between the feature importance of two databases can be attributed to the different ranges of gravimetric surface area values (Supporting Information Figure S9). With a smaller range of gravimetric surface area values in CoRE MOF 2019, specific host–guest interactions do not play as important a role as they do in the mechanically stable USMOF DB subset. In the mechanically stable MOFs in USMOF DB, large surface areas may be coupled to specifically favorable adsorption sites, both contributing to adsorption trends.

To assess our expectations of generality based on feature importance commonalities and differences, we evaluated the transferability of the ML models between databases for CH_4 and then for CO_2. As expected, the RF model trained on the USMOF DB, which covers a broader range of working capacity values, exhibits high performance on CoRE MOF 2019, achieving a test R^2 of 0.76 due to the overriding importance of gravimetric pore volume in predicting CH_4 working

capacity for both databases (Supporting Information Figure S10). However, despite the fact that an RF model trained on CoRE MOF 2019 demonstrated excellent performance (test $R^2 = 0.97$) in predicting CH_4 working capacity, the same model encountered challenges in accurately predicting the CH_4 working capacity of mechanically stable USMOF DB MOFs, yielding a test R^2 of -0.18. Specifically, the prediction is accurate for the working capacity range in the domain of the model's training data (0 to 300 cm^3(STP)/g framework) but shows a systematic underestimation over the range greater than the training data (> 300 cm^3(STP)/g framework). The limited transferability from CoRE MOF 2019 to the mechanically stable subset of USMOF DB stems from the absence of data points with high CH_4 working capacity in the former database that are present in the latter (Supporting Information Figure S11). The nature of a RF regressor prediction, where predictions are produced by averaging the predictions of the decision trees in the forest, causes the model to fail in extrapolation, resulting in underestimations in this case. Nevertheless, few, if any, alternative regression models would be expected to extrapolate in a reliable fashion, highlighting the importance of generating data beyond the range of values in synthesized MOFs in order to predict properties of MOFs with exceptional characteristics beyond previously studied materials.

We next repeated the analysis of transferability for CO_2 working capacity. The RF regressor models trained to predict CO_2 working capacity in USMOF DB fail to make accurate predictions on CoRE MOF 2019, and vice versa, with test R^2 values of -0.18 and 0.58, respectively. This outcome is expected since the models emphasized different features to predict the CO_2 working capacity for each database (i.e., pore volume in CoRE MOF 2019 versus surface area and full-scope atomic charge RACs in USMOF DB), as observed through the impurity-based feature importance. The lack of transferability between our *in silico* mechanically stable USMOF DB and CoRE MOF 2019 indicates that they have distinct chemical features despite being derived from

the same building blocks, highlighting opportunities to tune working capacity through continued discovery of novel MOFs such as those represented by the USMOF DB, especially for CO_2.

In USMOF DB, a considerable number of MOFs exhibit a preference for CH_4 over CO_2, while the majority of MOFs in CoRE MOF 2019 (89.4%) demonstrate a higher working capacity for CO_2 over CH_4. To further investigate this trend, we classified MOFs into three groups based on the working capacity ratio (see Sec. 3a). Then, we trained a RF regressor model on the set of MOFs with a high ratio of CO_2:CH_4 working capacity (i.e., > 1.1) and a low CO_2:CH_4 ratio (i.e., < 0.9) to analyze which features contribute the most to the different preferences on adsorbing gas molecules. We omitted the intermediate set with similar working capacities for both species to ensure that the models were learning from distinct sets of MOFs with significant preferences for one or the other molecule. The performance of models predicting CH_4 working capacity trained on each set was marginally reduced (test R^2 = 0.97 and 0.98 for high- and low-ratio models, respectively) with respect to the model trained on the whole mechanically stable subset of USMOF DB, but otherwise was mostly comparable (Supporting Information Table S7). The Gini impurity feature importance analysis reveals that both models emphasize gravimetric pore volume (i.e., the probe accessible volume and pore volume) and gravimetric surface area to predict the working capacity (Supporting Information Table S8). Despite the similarities in the feature importance among models predicting CH_4 working capacity for each set, the transferability between MOFs with high ratios and low ratios is notably poor. The R^2 for the model trained on MOFs with low ratios to predict those with high ratios is 0.18 and it is -0.70 for the reverse case. Even though these models share similar important features, the different range of CH_4 working capacity for each set of MOFs limits model transferability (Supporting Information Figure S12). Along with the RF model, we also investigated a LASSO model in this case since linear model can perform in a

superior fashion in extrapolation when both models share similar important features. As expected, the R^2 for the model trained on low ratios to predict high ratios is 0.94 and 0.45 for the reverse case, showing improved accuracy of LASSO compared to RF (Supporting information Table S10).

We conducted a similar analysis on each set of MOFs for CO_2 working capacities. The performance of CO_2 models trained on high- and low-ratio subsets of the mechanically stable subset of USMOF DB actually improved (test R^2 = 0.89 and 0.84 for high and low ratio, respectively) over the model trained on the entire dataset despite having smaller training set. The RF model trained on MOFs with a high CO_2:CH_4 ratio emphasizes gravimetric surface area (feature importance of 28.2%) and connectivity RACs (f-T-0-all and f-T-1-all, in which T describes the number of atoms to which an atom is bonded, with feature importance of 15.4% and 5.4% respectively, Supporting Information Table S9). On the other hand, the model trained on MOFs with the low-ratio materials predominantly emphasizes the gravimetric surface area (feature importance of 79.2%). These models trained on each subset of the dataset highlight that different features contribute to the CO_2 working capacities between the MOFs with high preference for CH_4 and those that prefer CO_2. Therefore, it is expected that these models should fail to transfer from high to low ratio or vice versa. The RF model applied to the other subset (i.e., high to low) shows test R^2 values of -30.43 (-0.12 for low to high), indicating that distinct features are employed to predict the CO_2 working capacity of the two sets even when the distributions overlap (Supporting Information Table S10). The unexpected disparities in working capacities between the mechanically stable subset of the USMOF DB and CoRE MOF 2019 indicate that even though USMOF DB is derived from CoRE MOF 2019 building blocks, recombining the building blocks with varying connectivity can substantially alter the working capacity. This suggests there are opportunities to tune preferences of MOFs toward CH_4 or CO_2 adsorption that have not yet been

explored experimentally.

3c. Distinct Trends in Working Capacities Between the Two Databases

Even though the building blocks of the mechanically stable USMOF DB MOFs were obtained from CoRE MOF 2019, there are significant differences in properties of the two MOF databases. These differences are evident for both gas molecules but especially apparent for CO_2, both in the distribution of gas working capacity values and in feature importances of the ML models trained on the two sets (see Figure 3). To elucidate the reason for these discrepancies, we next analyzed the MOFs in USMOF DB that exhibit exceptionally high CH_4 working capacity. We observed that certain building blocks appeared frequently in the top 20 MOFs ranked by CH_4 working capacity in the mechanically stable subset of the USMOF DB. Specifically, inorganic node 41 (N41), extracted from CSD refcode: KOZSID in CoRE MOF 2019, appears 11 times in the top 20 MOFs, all exhibiting CH_4 working capacities greater than 790 cm^3/g (Supporting Information Table S11). Although N41 occurs frequently in the USMOF DB, being present in 16.2% (178 out of 1,102) of MOFs, its appearance in the top 20 MOFs is still enriched with respect to its presence in the mechanically stable subset. Chemically, N41 is a Co-based porphyrinic node, the rigid nature of which[103,104] allows the structures to have high mechanical stability, leading to its presence in combination with a number of nets and other linkers in the mechanically stable USMOF DB subset. This finding also aligns with previous research where N41 was frequently observed near the Pareto front of an optimal trade-off between mechanical stability and CH_4 working capacity[43]. We nevertheless note that the original MOF in CoRE MOF 2019 (CSD refcode: KOZSID) from which N41 was extracted has a rather low CH_4 working capacity of only 12.32 cm^3/g (Figure 5). This unexpected discrepancy in CH_4 adsorption arises from the diversity of topologies present in the mechanically stable subset of the USMOF DB, which in turn results

in distinct connectivity (Supporting Information Figure S13). This difference leads to larger variation in pore volumes, which have high importance for predicting the working capacity, when compared with MOFs containing the same node or edge in the mechanically stable subset of the USMOF DB and CoRE MOF 2019 (Supporting Information Figure S14).

Edge 13 (E13, from CSD refcode: UVAHIK) frequently appears in MOFs with high CH_4 working capacity as well (i.e., 10 times in the top 20 MOFs). It contains 3 benzene rings, making it one of the longest edges (12.7 Å) in our organic edge building block set. This allows the structures to have larger pore volumes and higher working capacities. Although there are longer edge building blocks than E13, they do not appear frequently in the mechanically stable set of USMOF DB because they are more susceptible to instability upon application of hydrostatic pressure, resulting in lower mechanical stability.

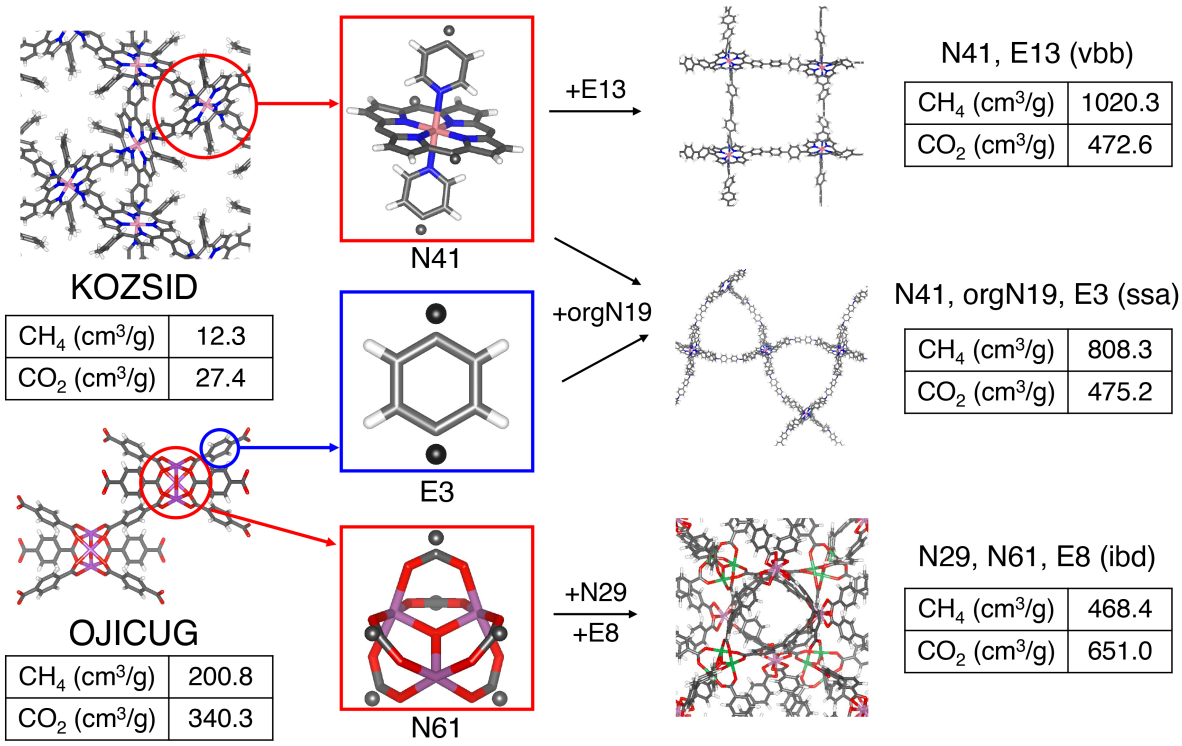

Figure 5. (Left) Example MOFs (CSD refcodes: KOZSID and OJICUG) from CoRE MOF 2019 and their calculated CH_4 and CO_2 working capacities. (Middle) Extracted nodes (N41 and N61) and an edge (E3, derived from benzene-1,4-dicarboxylic acid; BDC) from the CoRE MOF 2019 structures shown at left. (Right) Hypothetical MOFs constructed from recombining extracted nodes and edges, along with additional nodes and edges indicated on the arrows, and their calculated CH_4 and CO_2 working capacities.

Next, we repeated our analysis on the MOF components with high occurrence in the top 20 MOFs in the mechanically stable subset of the USMOF DB for CO_2 working capacity (Supporting Information Table S12). In comparison to CH_4, no single inorganic node dominates as consistently, with node 61 (N61, CSD refcode: OJICUG, 6 times) and node 29 (N29, CSD refcode: TAGTUT, 5 times) each occurring about 25% of the time. Notably, N61 is consistently selected in structures that include other inorganic or organic nodes to form either a 2 inorganic node–1 edge or a 1 inorganic node–1 organic node–1 edge configuration (Figure 5). All 11 MOFs containing these nodes have CO_2 working capacities greater than 580 cm^3(STP)/g framework. The original CO_2 working capacities of the CoRE MOF 2019 MOFs, from which N61 and N29 were taken, were 340.3 cm^3/g and 295.9 cm^3/g, respectively. Although these values are lower than the working capacities in USMOF DB, they still exceed one standard deviation (127.7 cm^3/g) above the average value of CO_2 working capacity (average: 133.6 cm^3/g + 1 std. dev. is 261.3 cm^3/g) in CoRE MOF 2019. The less dramatic difference for CO_2 compared to the case of CH_4 is expected because ML model feature importance indicates CO_2 working capacity depends on both geometric features and the chemical identity of nodes (see Sec. 3b). Thus, recombining the nodes with alternate linkers and on new nets should result in a wide range of geometric features that could alter working capacities, but it has a smaller impact on CO_2 working capacity compared to CH_4. The enrichment of privileged edges to produce high CO_2 working capacity is also less significant than for the CH_4 case. Edge 8 (E8, CSD refcode: KIFKEQ), derived from biphenyl-4,4-dicarboxylic acid, appears 8 times among the top 20 MOFs, and this edge is also the most

frequently appearing edge in USMOF DB (145 out of 1,102, Supporting Information Figure S8). Unlike the case of the preferred edge for CH_4 working capacity, E8 consists of two benzene rings and is only 8.63 Å long, suggesting its length is not a major contributing factor to higher CO_2 uptake.

Now we analyze the building blocks (i.e., edges and nodes) and their connectivity (i.e., nets) to identify their relationship with the adsorption preferences of CO_2 and CH_4 for MOFs in the USMOF DB. We first investigate the edges, examining whether their lengths are correlated with the adsorption preferences of gas species. Our preceding analysis revealed distinct trends within the USMOF DB, with MOFs exhibiting high CH_4 working capacity demonstrating longer edge lengths, while no strong correlation was observed between CO_2 working capacity and edge length. Now, we observe a notable difference in average edge length between MOFs with CO_2:CH_4 ratios below 0.9 (10.03 ± 2.82 Å) and those above 1.1 (6.93 ± 2.02 Å,), indicating that edge length may serve as a useful design knob for enhancing CH_4 adsorption preference over CO_2 (Figure 6). Despite the smaller standard deviation for the average length in the high-ratio case, longer edges serve as a better tool for suggesting what the MOF working capacity ratio will be. That is, most of the MOFs (87.5%) with longer edges (> 11 Å) have a low CO_2:CH_4 ratio, whereas only a slight majority (58.7%) of MOFs with shorter edges (< 7 Å) exhibit a high ratio. Distinct building block configurations (i.e., 1inor-1org-1edge) contribute to this trend: large-pore MOFs that feature longer linkers combined with organic nodes give rise to the lowest CO_2:CH_4 ratios, while a short linker combined with other components will not necessarily lead to smaller-pore MOFs.

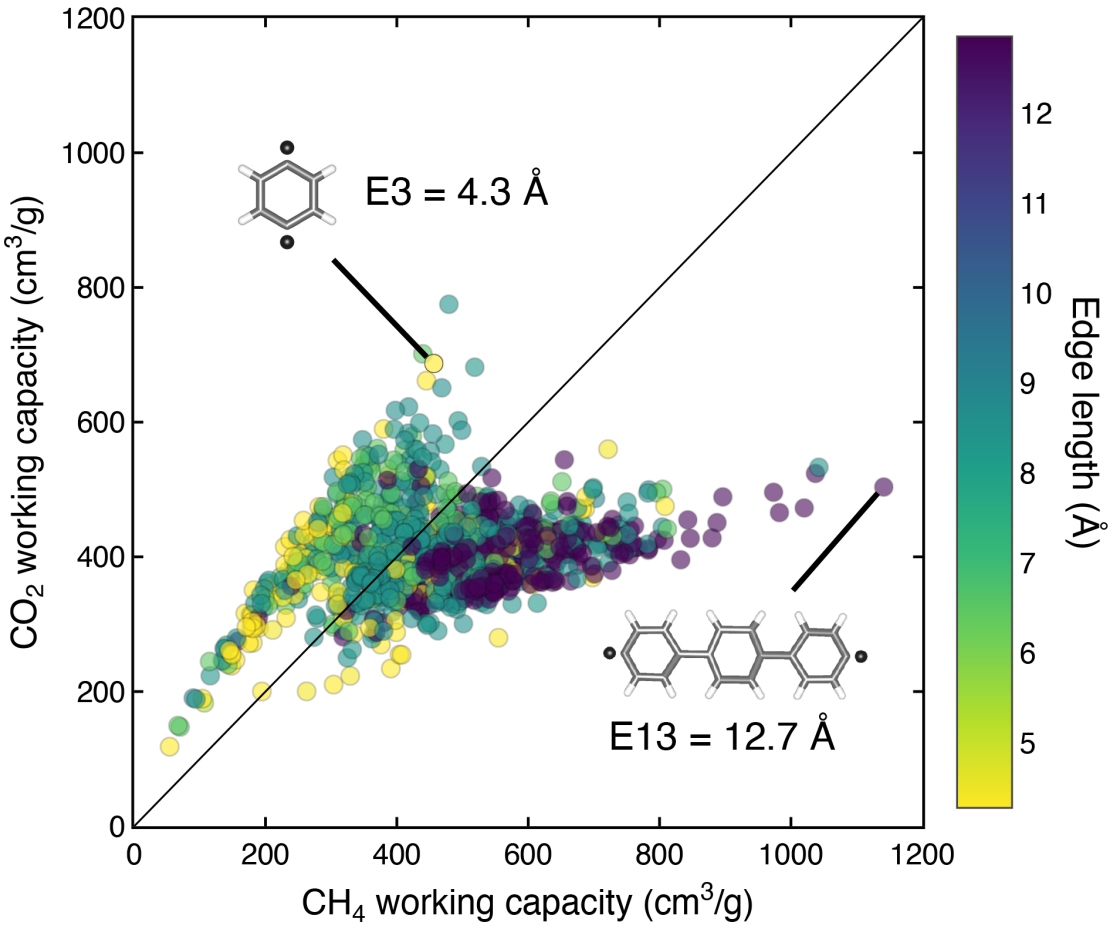

Figure 6. CO_2 vs. CH_4 working capacity (in cm^3(STP)/g framework) colored by edge length (Å). The points are colored by edge length according to inset color bar, with shorter edges colored in yellow and longer edges colored in purple. Representative edges E3 and E13 (BDC and H2TPDC, respectively) are shown with their edge lengths annotated.

We also considered the connectivity net as another factor influencing adsorption preferences, which is particularly necessary because the distribution of connectivity nets observed between the mechanically stable USMOF and CoRE MOF 2019 databases differs (Supporting Information Figure S13). Notably, MOFs characterized by topologies that give rise to small gravimetric pore volume (GPOV) and large volumetric surface areas (VSA), such as **wiv** and **ibd**, exhibit a preferentially high working capacity for CO_2 over CH_4 in the USMOF DB and the opposite is true for topologies such as **myd** and **ssa** (Supporting Information Figure S15). For

example, edge E8 (8.6 Å) on an **ibd** connectivity net appears in 30% of the top 20 MOFs with the highest CO_2 working capacity in the mechanically stable subset of USMOF DB (Supporting Information Table S11). Therefore, shorter edges with connectivity nets that result in small pore volume and large surface area can be expected to lead to higher CO_2 working capacity over CH_4, but many of the USMOF DB MOFs have the converse features that instead give rise to higher CH_4 working capacity.

Next, we investigated the influence of node identity on the high- and low-ratio gas adsorption preferences. There are distinct trends among different nodes, with some nodes exhibiting a preference for CO_2 (e.g., nodes N38 and N61) while others favor CH_4 (e.g., nodes N41 and N79). Most other nodes are present in both CO_2- and CH_4-favoring MOFs. The N38 and N61 nodes exhibit identical structures characterized by trinuclear metal clusters centered on a μ_3-oxygen [$M_3O(O_2C\text{-})_6$], albeit with distinct metals: iron for N38 and scandium for N61. These clusters feature strongly bound Fe^{3+} and Sc^{3+} centers coordinated by a μ_3-O^{2-} donor, resulting in exceptional mechanical stability. Each metal atom is coordinated by five oxygen atoms in an octahedral coordination environment, leaving one open metal site available for adsorption. The presence of these highly accessible open metal sites renders MOFs containing N38 and N61 particularly suitable for adsorbing CO_2 over CH_4. In addition to the aforementioned Co-based porphyrinic N41, N79 is also a CH_4-favoring node. N79 is a $Li_4(OR)_4$ (R = pyridine) cubane cluster. This large node contains eight pyridines, resulting in a high void volume conducive to CH_4 adsorption. These nodes do not exhibit open metal sites, explaining their relatively low CO_2 adsorption. Of the four highlighted nodes, N41 and N61 have the greatest apparent influence on gas uptake preference. For 96% of N61-containing MOFs (65 out of 68), the CO_2:CH_4 ratio is greater than 1.1, while 88% of N41-containing MOFs (135 out of 153) have a ratio less than 0.9

(Supporting Information Figure S16). This observation is consistent with earlier analysis where these nodes appeared frequently in the top 20 MOFs ranked by working capacity of CH_4 and CO_2, respectively.

Studies have shown that some MOFs with open metal sites, such as MOF-74-Mg (CPO-27-Mg, Mg2(dobdc))[105], show enhanced CO_2 adsorption properties.[106-108] Therefore, we also examine the presence of open metal sites more broadly within the nodes of MOFs in USMOF DB beyond the privileged nodes we identified with high CO_2:CH_4 ratios (i.e., N38 and N61). While there are notable examples with open metal sites that prefer CO_2, the presence of an open metal site alone is not the sole deciding factor in determining that MOFs will prefer CO_2 over CH_4. When considering only the most mechanically stable MOFs by requiring a bulk modulus (K_{VRH}) greater than 10 GPa (i.e., rather than just 5 GPa), the set reduces to 340 MOFs. However, among these MOFs, MOFs with open metal sites do indeed show a strong preference for CO_2 (Supporting Information Figure S17). Thus, observation of the nodes with open metal site preference for CO_2 cannot be fully decoupled from mechanical stability considerations, which removes less mechanically stable MOFs with large pore volumes that often exhibit high CH_4 working capacities.

Finally, we summarize these design principles to identify strategies for enhancing the selectivity of MOFs for either CO_2 or CH_4 gas adsorption. For high-throughput screening, our RF models revealed that computed void volume and surface area of MOFs should serve as predictive indicators of gas adsorption preferences. Additionally, since trinuclear metal clusters [$M_3O(O_2C-)_6$] (i.e., N38 and N61) show a preference for CO_2 over CH_4, substituting these metal clusters with different metal atoms also in a +3 oxidation state (e.g., V, Cr, Mn, and Co) could be a strategy to tailor MOFs for high selectivity for CO_2 while introducing metal diversity. To construct MOFs exhibiting a preference for high CH_4 working capacity over CO_2, employing longer edges and

connectivity nets characterized by large void volumes and small surface areas should serve as the most effective design strategies. As a secondary strategy, node identities are also informative: open metal sites are likely to enhance CO_2 adsorption. Conversely, robust nodes contribute to mechanically stable structures even when combined with long edges in order to create large void volumes that contribute to CH_4 adsorption enhancement. These principles cannot be observed in CoRE MOF 2019 due to the absence of MOFs with higher CH_4 working capacities in previously synthesized MOFs. However, we anticipate these design principles should aid in the future identification of candidate MOFs as synthetic targets or for assessment of synthesized MOFs that have not been deposited in structural databases.

4. Conclusions

To advance the design of novel MOFs tailored for specific gas adsorption properties, our study focused on the CO_2 and CH_4 adsorption properties of two MOF databases: mechanically stable MOFs in the *in silico* ultrastable MOF database (USMOF DB) and CoRE MOF 2019. The USMOF DB differs from other *in silico* databases by incorporating considerations for thermal stability and stability upon solvent removal during its construction. In addition, to ensure mechanical stability, we computed the bulk moduli of MOFs in the USMOF DB and selected those demonstrating sufficient mechanical strength. We showed that the gas working capacities of these two databases showed significantly different distributions: a significantly higher fraction of MOFs in the mechanically stable subset of the USMOF DB exhibited preferentially high CH_4 working capacity, while most MOFs in CoRE MOF 2019 exhibited higher CO_2 working capacity.

Furthermore, leveraging machine learning techniques, we trained random forest models on the gas working capacities of these databases and investigated the features that determine the

adsorption properties. Interestingly, while geometric features such as pore volume and surface area predominantly determined both CO_2 and CH_4 gas working capacities in CoRE MOF 2019, the CO_2 working capacity in the USMOF DB was influenced by both geometric features and the electronegativity of atoms near the metal centers in the secondary building units. We observed that these models had limited transferability across databases, except for the case of CH_4 from USMOF to CoRE MOF 2019. The USMOF DB CH_4 capacities spanned those of CoRE MOF 2019, often leading to good interpolative transferability, whereas the same was not true of CO_2 due to different feature importance in the models to predict CO_2 working capacity. This underscores the need for continued expansion of datasets to include novel experimental and hypothetical MOFs in order to obtain more universal design rules for tailoring MOFs for high adsorption of one gas species over another.

Moreover, our investigation into the structural characteristics of MOFs provided valuable insights. While MOFs with longer edge lengths expectedly tended to prefer CH_4 over CO_2, topologies containing a small gravimetric pore volume and large volumetric surface area (e.g., **wiv** and **ibd**) exhibited a preference for CO_2 over CH_4, and vice versa (e.g., **myd** and **ssa**). By analyzing top-performing CO_2 adsorbing MOFs, we observed the MOFs containing SBUs with trinuclear metal clusters centered on a μ_3-oxygen demonstrated a high preference for CO_2 adsorption due to the accessible open metal site of this cluster. We were able to confirm the generality of this observation that MOFs with open metal sites and high mechanical stability facilitate higher CO_2 adsorption. These findings contribute to the understanding of key design principles for engineering MOFs with tailored gas adsorption properties, thereby advancing their applications in gas separation and storage.

ASSOCIATED CONTENT

Supporting Information.

Dataset statistics for mechanical properties calculation; hyperparameter set for random forest models; description of geometric features and RAC features; relationship between bulk and shear elastic modulus; relationship between unit cell volume and bulk elastic modulus; uptake and working capacities of CH_4 and CO_2; distribution of working capacities for each configurations of nodes and edges; number of MOFs in USMOF and CoRE MOF 2019 grouped by working capacities; model performances of linear regression, random forest, and artificial neural network model trained on CH_4 and CO_2 working capacities; feature importances of random forest regressor predicting working capacities; principal component analysis of geometric features of USMOF DB and CoRE MOF 2019; distribution of geometric features, working capacities, and net in each DB; model transferability between DBs; top MOFs in the USMOF DB with high CH_4 or CO_2 working capacities; model performance and feature importance for predicting the ratio of CO_2:CH_4 working capacities; distribution of geometric features for connectivity nets; distribution of working capacities of representative nodes (N61 and N41); distribution of working capacities with mechanical stability cutoff.

This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*email: hjkulik@mit.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was supported by the Center for Enhanced Nanofluidic Transport, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award DE-SC0019112 (to C.O. and S.Y.). C.O. was partially supported by an MIT Portugal Seed Fund during the initial stages of this work and subsequently by a seed grant from the Abdul Latif Jameel Water and Food Systems Lab at the Massachusetts Institute of Technology. A.N. was partially supported by the National Science Foundation Graduate Research Fellowship Program (Grant Number #1122374). The authors acknowledge the MIT SuperCloud and Lincoln Laboratory Supercomputing Center for providing HPC resources that have contributed to the research results reported within this article. H.J.K. was supported by a Simon Family Faculty Research Innovation Fund and a Sloan Fellowship in Chemistry. The authors acknowledge Adam H. Steeves for providing a critical reading of the manuscript.

REFERENCES

(1) Specht, E.; Redemann, T.; Lorenz, N. Simplified Mathematical Model for Calculating Global Warming through Anthropogenic Co2. *Int. J. Therm. Sci.* **2016,** *102*, 1-8.

(2) Barbera, A. C.; Vymazal, J.; Maucieri, C. In *Encyclopedia of Ecology (Second Edition)*; Fath, Brian, Ed.; Elsevier: Oxford, 2019,329-333

(3) Grubb, M.; Okereke, C.; Arima, J.; Bosetti, V.; Chen, Y.; Edmonds, J.; Gupta, S.; Köberle, A.; Kverndokk, S.; Malik, A.; Sulistiawati, L. " Introduction and Framing. In Ipcc, 2022: Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group Iii to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change," 2022.

(4) Ben-Mansour, R.; Habib, M. A.; Bamidele, O. E.; Basha, M.; Qasem, N. A. A.; Peedikakkal, A.; Laoui, T.; Ali, M. Carbon Capture by Physical Adsorption: Materials, Experimental Investigations and Numerical Modeling and Simulations – a Review. *Appl. Energy* **2016,** *161*, 225-255.

(5) O'Keeffe, M.; Peskov, M. A.; Ramsden, S. J.; Yaghi, O. M. The Reticular Chemistry Structure Resource (Rcsr) Database of, and Symbols for, Crystal Nets. *Acc. Chem. Res.* **2008,** *41*, 1782-9.

(6) Zhou, H. C.; Long, J. R.; Yaghi, O. M. Introduction to Metal-Organic Frameworks. *Chem. Rev.* **2012,** *112*, 673-4.

(7) Simon, C. M.; Kim, J.; Gomez-Gualdron, D. A.; Camp, J. S.; Chung, Y. G.; Martin, R. L.; Mercado, R.; Deem, M. W.; Gunter, D.; Haranczyk, M.; Sholl, D. S.; Snurr, R. Q.; Smit, B. The Materials Genome in Action: Identifying the Performance Limits for Methane Storage. *Energy Environ. Sci.* **2015,** *8*, 1190-1199.

(8) Sumida, K.; Rogow, D. L.; Mason, J. A.; McDonald, T. M.; Bloch, E. D.; Herm, Z. R.; Bae, T. H.; Long, J. R. Carbon Dioxide Capture in Metal-Organic Frameworks. *Chem. Rev.* **2012,** *112*, 724-81.
(9) Han, S.; Huang, Y.; Watanabe, T.; Nair, S.; Walton, K. S.; Sholl, D. S.; Carson Meredith, J. Mof Stability and Gas Adsorption as a Function of Exposure to Water, Humid Air, So2, and No2. *Microporous Mesoporous Mater.* **2013,** *173*, 86-91.
(10) 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.; Hubberstey, P.; Champness, N. R.; Schroder, M. High Capacity Hydrogen Adsorption in Cu(Ii) Tetracarboxylate Framework Materials: The Role of Pore Size, Ligand Functionalization, and Exposed Metal Sites. *J. Am. Chem. Soc.* **2009,** *131*, 2159-71.
(11) Murray, L. J.; Dinca, M.; Long, J. R. Hydrogen Storage in Metal-Organic Frameworks. *Chem. Soc. Rev.* **2009,** *38*, 1294-314.
(12) He, Y.; Zhou, W.; Qian, G.; Chen, B. Methane Storage in Metal-Organic Frameworks. *Chem. Soc. Rev.* **2014,** *43*, 5657-78.
(13) Peng, Y.; Krungleviciute, V.; Eryazici, I.; Hupp, J. T.; Farha, O. K.; Yildirim, T. Methane Storage in Metal-Organic Frameworks: Current Records, Surprise Findings, and Challenges. *J. Am. Chem. Soc.* **2013,** *135*, 11887-94.
(14) Morris, R. E.; Wheatley, P. S. Gas Storage in Nanoporous Materials. *Angew. Chem., Int. Ed. Engl.* **2008,** *47*, 4966-81.
(15) Carsch, K. M.; Huang, A. J.; Dods, M. N.; Parker, S. T.; Rohde, R. C.; Jiang, H. Z. H.; Yabuuchi, Y.; Karstens, S. L.; Kwon, H.; Chakraborty, R.; Bustillo, K. C.; Meihaus, K. R.; Furukawa, H.; Minor, A. M.; Head-Gordon, M.; Long, J. R. Selective Adsorption of Oxygen from Humid Air in a Metal-Organic Framework with Trigonal Pyramidal Copper(I) Sites. *J. Am. Chem. Soc.* **2024,** *146*, 3160-3170.
(16) Zhao, X.; Wang, Y.; Li, D. S.; Bu, X.; Feng, P. Metal-Organic Frameworks for Separation. *Adv. Mater.* **2018,** *30*, 1705189.
(17) Qian, Q.; Asinger, P. A.; Lee, M. J.; Han, G.; Mizrahi Rodriguez, K.; Lin, S.; Benedetti, F. M.; Wu, A. X.; Chi, W. S.; Smith, Z. P. Mof-Based Membranes for Gas Separations. *Chem. Rev.* **2020,** *120*, 8161-8266.
(18) Watanabe, T.; Keskin, S.; Nair, S.; Sholl, D. S. Computational Identification of a Metal Organic Framework for High Selectivity Membrane-Based Co2/Ch4 Separations: Cu(Hfipbb)(H2hfipbb)0.5. *Phys. Chem. Chem. Phys.* **2009,** *11*, 11389-94.
(19) Gonzalez, M. I.; Kapelewski, M. T.; Bloch, E. D.; Milner, P. J.; Reed, D. A.; Hudson, M. R.; Mason, J. A.; Barin, G.; Brown, C. M.; Long, J. R. Separation of Xylene Isomers through Multiple Metal Site Interactions in Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2018,** *140*, 3412-3422.
(20) Yang, D.; Gates, B. C. Catalysis by Metal Organic Frameworks: Perspective and Suggestions for Future Research. *ACS Catal.* **2019,** *9*, 1779-1798.
(21) Goetjen, T. A.; Liu, J.; Wu, Y.; Sui, J.; Zhang, X.; Hupp, J. T.; Farha, O. K. Metal-Organic Framework (Mof) Materials as Polymerization Catalysts: A Review and Recent Advances. *Chem. Commun.* **2020,** *56*, 10409-10418.
(22) Gascon, J.; Corma, A.; Kapteijn, F.; Llabrés i Xamena, F. X. Metal Organic Framework Catalysis: Quo Vadis? *ACS Catal.* **2013,** *4*, 361-378.

(23) Adamji, H.; Nandy, A.; Kevlishvili, I.; Roman-Leshkov, Y.; Kulik, H. J. Computational Discovery of Stable Metal-Organic Frameworks for Methane-to-Methanol Catalysis. *J. Am. Chem. Soc.* **2023,** *145*, 14365-14378.
(24) Lee, J.; Farha, O. K.; Roberts, J.; Scheidt, K. A.; Nguyen, S. T.; Hupp, J. T. Metal-Organic Framework Materials as Catalysts. *Chem. Soc. Rev.* **2009,** *38*, 1450-9.
(25) Furukawa, H.; Ko, N.; Go, Y. B.; Aratani, N.; Choi, S. B.; Choi, E.; Yazaydin, A. O.; Snurr, R. Q.; O'Keeffe, M.; Kim, J.; Yaghi, O. M. Ultrahigh Porosity in Metal-Organic Frameworks. *Science* **2010,** *329*, 424-8.
(26) Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. The Chemistry and Applications of Metal-Organic Frameworks. *Science* **2013,** *341*, 1230444.
(27) Wei, R.; Gaggioli, C. A.; Li, G.; Islamoglu, T.; Zhang, Z.; Yu, P.; Farha, O. K.; Cramer, C. J.; Gagliardi, L.; Yang, D.; Gates, B. C. Tuning the Properties of Zr6o8 Nodes in the Metal Organic Framework Uio-66 by Selection of Node-Bound Ligands and Linkers. *Chem. Mater.* **2019,** *31*, 1655-1663.
(28) Ding, M.; Cai, X.; Jiang, H. L. Improving Mof Stability: Approaches and Applications. *Chem. Sci.* **2019,** *10*, 10209-10230.
(29) Yuan, S.; Feng, L.; Wang, K.; Pang, J.; Bosch, M.; Lollar, C.; Sun, Y.; Qin, J.; Yang, X.; Zhang, P.; Wang, Q.; Zou, L.; Zhang, Y.; Zhang, L.; Fang, Y.; Li, J.; Zhou, H. C. Stable Metal-Organic Frameworks: Design, Synthesis, and Applications. *Adv. Mater.* **2018,** *30*, e1704303.
(30) Nandy, A.; Duan, C.; Kulik, H. J. Using Machine Learning and Data Mining to Leverage Community Knowledge for the Engineering of Stable Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2021,** *143*, 17535-17547.
(31) Farha, O. K.; Hupp, J. T. Rational Design, Synthesis, Purification, and Activation of Metal-Organic Framework Materials. *Acc. Chem. Res.* **2010,** *43*, 1166-75.
(32) Zhang, X.; Chen, Z.; Liu, X.; Hanna, S. L.; Wang, X.; Taheri-Ledari, R.; Maleki, A.; Li, P.; Farha, O. K. A Historical Overview of the Activation and Porosity of Metal-Organic Frameworks. *Chem. Soc. Rev.* **2020,** *49*, 7406-7427.
(33) Dodson, R. A.; Wong-Foy, A. G.; Matzger, A. J. The Metal–Organic Framework Collapse Continuum: Insights from Two-Dimensional Powder X-Ray Diffraction. *Chem. Mater.* **2018,** *30*, 6559-6565.
(34) Ockwig, N. W.; Delgado-Friedrichs, O.; O'Keeffe, M.; Yaghi, O. M. Reticular Chemistry: Occurrence and Taxonomy of Nets and Grammar for the Design of Frameworks. *Acc. Chem. Res.* **2005,** *38*, 176-82.
(35) Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. Large-Scale Screening of Hypothetical Metal-Organic Frameworks. *Nat. Chem.* **2011,** *4*, 83-9.
(36) Boyd, P. G.; Chidambaram, A.; Garcia-Diez, E.; Ireland, C. P.; Daff, T. D.; Bounds, R.; Gladysiak, A.; Schouwink, P.; Moosavi, S. M.; Maroto-Valer, M. M.; Reimer, J. A.; Navarro, J. A. R.; Woo, T. K.; Garcia, S.; Stylianou, K. C.; Smit, B. Data-Driven Design of Metal-Organic Frameworks for Wet Flue Gas Co(2) Capture. *Nature* **2019,** *576*, 253-256.
(37) Chung, Y. G.; Haldoupis, E.; Bucior, B. J.; Haranczyk, M.; Lee, S.; Zhang, H.; Vogiatzis, K. D.; Milisavljevic, M.; Ling, S.; Camp, J. S.; Slater, B.; Siepmann, J. I.; Sholl, D. S.; Snurr, R. Q. Advances, Updates, and Analytics for the Computation-Ready, Experimental

Metal–Organic Framework Database: Core Mof 2019. *J. Chem. Eng. Data* **2019,** *64*, 5985-5998.
(38) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. The Cambridge Structural Database. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2016,** *72*, 171-9.
(39) Colón, Y. J.; Gómez-Gualdrón, D. A.; Snurr, R. Q. Topologically Guided, Automated Construction of Metal–Organic Frameworks and Their Evaluation for Energy-Related Applications. *Cryst. Growth Des.* **2017,** *17*, 5801-5810.
(40) Burner, J.; Luo, J.; White, A.; Mirmiran, A.; Kwon, O.; Boyd, P. G.; Maley, S.; Gibaldi, M.; Simrod, S.; Ogden, V.; Woo, T. K. Arc–Mof: A Diverse Database of Metal-Organic Frameworks with Dft-Derived Partial Atomic Charges and Descriptors for Machine Learning. *Chem. Mater.* **2023,** *35*, 900-916.
(41) Majumdar, S.; Moosavi, S. M.; Jablonka, K. M.; Ongari, D.; Smit, B. Diversifying Databases of Metal Organic Frameworks for High-Throughput Computational Screening. *ACS Appl. Mater. Interfaces* **2021,** *13*, 61004-61014.
(42) Moosavi, S. M.; Nandy, A.; Jablonka, K. M.; Ongari, D.; Janet, J. P.; Boyd, P. G.; Lee, Y.; Smit, B.; Kulik, H. J. Understanding the Diversity of the Metal-Organic Framework Ecosystem. *Nat. Commun.* **2020,** *11*, 4068.
(43) Nandy, A.; Yue, S.; Oh, C.; Duan, C.; Terrones, G. G.; Chung, Y. G.; Kulik, H. J. A Database of Ultrastable Mofs Reassembled from Stable Fragments with Machine Learning Models. *Matter* **2023,** *6*, 1585-1603.
(44) Chapman, K. W.; Halder, G. J.; Chupas, P. J. Pressure-Induced Amorphization and Porosity Modification in a Metal-Organic Framework. *J. Am. Chem. Soc.* **2009,** *131*, 17546-7.
(45) Tan, J. C.; Cheetham, A. K. Mechanical Properties of Hybrid Inorganic-Organic Framework Materials: Establishing Fundamental Structure-Property Relationships. *Chem. Soc. Rev.* **2011,** *40*, 1059-80.
(46) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. Reticular Synthesis and the Design of New Materials. *Nature* **2003,** *423*, 705-14.
(47) Wu, H.; Yildirim, T.; Zhou, W. Exceptional Mechanical Stability of Highly Porous Zirconium Metal-Organic Framework Uio-66 and Its Important Implications. *J. Phys. Chem. Lett.* **2013,** *4*, 925-30.
(48) Hu, Y. H.; Zhang, L. Amorphization of Metal-Organic Framework Mof-5 at Unusually Low Applied Pressure. *Phys. Rev. B* **2010,** *81*, 174103.
(49) Erkartal, M.; Durandurdu, M. Pressure-Induced Amorphization of Mof-5: A First Principles Study. *ChemistrySelect* **2018,** *3*, 8056-8063.
(50) Su, Z.; Miao, Y. R.; Mao, S. M.; Zhang, G. H.; Dillon, S.; Miller, J. T.; Suslick, K. S. Compression-Induced Deformation of Individual Metal-Organic Framework Microcrystals. *J. Am. Chem. Soc.* **2015,** *137*, 1750-3.
(51) Bennett, T. D.; Cheetham, A. K. Amorphous Metal-Organic Frameworks. *Acc. Chem. Res.* **2014,** *47*, 1555-62.
(52) Moghadam, P. Z.; Rogge, S. M. J.; Li, A.; Chow, C.-M.; Wieme, J.; Moharrami, N.; Aragones-Anglada, M.; Conduit, G.; Gomez-Gualdron, D. A.; Van Speybroeck, V.; Fairen-Jimenez, D. Structure-Mechanical Stability Relations of Metal-Organic Frameworks Via Machine Learning. *Matter* **2019,** *1*, 219-234.

(53) Altintas, C.; Keskin, S. Molecular Simulations of Mof Membranes and Performance Predictions of Mof/Polymer Mixed Matrix Membranes for Co(2)/Ch(4) Separations. *ACS Sustainable Chem. Eng.* **2019,** *7*, 2739-2750.
(54) Granato, M. A.; Martins, V. D.; Ferreira, A. F. P.; Rodrigues, A. E. Adsorption of Xylene Isomers in Mof Uio-66 by Molecular Simulation. *Microporous Mesoporous Mater.* **2014,** *190*, 165-170.
(55) Yang, Q.; Zhong, C. Molecular Simulation of Carbon Dioxide/Methane/Hydrogen Mixture Adsorption in Metal-Organic Frameworks. *J. Phys. Chem. B* **2006,** *110*, 17776-83.
(56) Liu, B.; Smit, B. Comparative Molecular Simulation Study of Co2/N2 and Ch4/N2 Separation in Zeolites and Metal-Organic Frameworks. *Langmuir* **2009,** *25*, 5918-26.
(57) Águeda Maté, V. I.; Delgado Dobladez, J. A.; Álvarez-Torrellas, S.; Larriba, M.; Martínez Rodríguez, Á. Modeling and Simulation of the Efficient Separation of Methane/Nitrogen Mixtures with [Ni3(Hcoo)6] Mof by Psa. *Chem. Eng. J.* **2019,** *361*, 1007-1018.
(58) Gutierrez-Sevillano, J. J.; Caro-Perez, A.; Dubbeldam, D.; Calero, S. Molecular Simulation Investigation into the Performance of Cu-Btc Metal-Organic Frameworks for Carbon Dioxide-Methane Separations. *Phys. Chem. Chem. Phys.* **2011,** *13*, 20453-60.
(59) Gómez-Gualdrón, D. A.; Wilmer, C. E.; Farha, O. K.; Hupp, J. T.; Snurr, R. Q. Exploring the Limits of Methane Storage and Delivery in Nanoporous Materials. *J. Phys. Chem. C* **2014,** *118*, 6941-6951.
(60) Hu, J.; Gu, C.; Liu, J. Post-Synthetic Metalation of Porous Framework Materials for Achieving High Natural Gas Storage and Working Capacity: A Gcmc Simulation Study. *Microporous Mesoporous Mater.* **2021,** *315*, 110931.
(61) Li, S.; Chung, Y. G.; Snurr, R. Q. High-Throughput Screening of Metal-Organic Frameworks for Co(2) Capture in the Presence of Water. *Langmuir* **2016,** *32*, 10368-10376.
(62) Leperi, K. T.; Chung, Y. G.; You, F.; Snurr, R. Q. Development of a General Evaluation Metric for Rapid Screening of Adsorbent Materials for Postcombustion Co2 Capture. *ACS Sustainable Chem. Eng.* **2019,** *7*, 11529-11539.
(63) Chen, H.; Snurr, R. Q. Computational Screening of Metal-Catecholate-Functionalized Metal–Organic Frameworks for Room-Temperature Hydrogen Storage. *J. Phys. Chem. C* **2021,** *125*, 21701-21708.
(64) Bénard, P.; Chahine, R. Storage of Hydrogen by Physisorption on Carbon and Nanostructured Materials. *Scr. Mater.* **2007,** *56*, 803-808.
(65) Poirier, E. C., R.; Bose, T.K. Hydrogen Adsorption in Carbon Nanostructures. *Int. J. Hydrogen Energy* **2001,** *26*, 831-835.
(66) Fernandez, M.; Woo, T. K.; Wilmer, C. E.; Snurr, R. Q. Large-Scale Quantitative Structure–Property Relationship (Qspr) Analysis of Methane Storage in Metal–Organic Frameworks. *J. Phys. Chem. C* **2013,** *117*, 7681-7689.
(67) Luo, Y.; Bag, S.; Zaremba, O.; Cierpka, A.; Andreo, J.; Wuttke, S.; Friederich, P.; Tsotsalas, M. Mof Synthesis Prediction Enabled by Automatic Data Mining and Machine Learning. *Angew. Chem., Int. Ed. Engl.* **2022,** *61*, e202200242.
(68) Fanourgakis, G. S.; Gkagkas, K.; Tylianakis, E.; Froudakis, G. E. A Universal Machine Learning Algorithm for Large-Scale Screening of Materials. *J. Am. Chem. Soc.* **2020,** *142*, 3814-3822.

(69) Anderson, R.; Biong, A.; Gomez-Gualdron, D. A. Adsorption Isotherm Predictions for Multiple Molecules in Mofs Using the Same Deep Learning Model. *J. Chem. Theory. Comput.* **2020,** *16*, 1271-1283.
(70) 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*, 40.
(71) Pardakhti, M.; Moharreri, E.; Wanik, D.; Suib, S. L.; Srivastava, R. Machine Learning Using Combined Structural and Chemical Descriptors for Prediction of Methane Adsorption Performance of Metal Organic Frameworks (Mofs). *ACS Comb. Sci.* **2017,** *19*, 640-645.
(72) Fanourgakis, G. S.; Gkagkas, K.; Tylianakis, E.; Froudakis, G. A Generic Machine Learning Algorithm for the Prediction of Gas Adsorption in Nanoporous Materials. *J. Phys. Chem. C* **2020,** *124*, 7117-7126.
(73) Altintas, C.; Altundal, O. F.; Keskin, S.; Yildirim, R. Machine Learning Meets with Metal Organic Frameworks for Gas Storage and Separation. *J. Chem. Inf. Model.* **2021,** *61*, 2131-2146.
(74) Haldoupis, E.; Nair, S.; Sholl, D. S. Efficient Calculation of Diffusion Limitations in Metal Organic Framework Materials: A Tool for Identifying Materials for Kinetic Separations. *J. Am. Chem. Soc.* **2010,** *132*, 7528-39.
(75) Fernandez, M.; Trefiak, N. R.; Woo, T. K. Atomic Property Weighted Radial Distribution Functions Descriptors of Metal–Organic Frameworks for the Prediction of Gas Uptake Capacity. *J. Phys. Chem. C* **2013,** *117*, 14095-14105.
(76) Yang, C.-T.; Pandey, I.; Trinh, D.; Chen, C.-C.; Howe, J. D.; Lin, L.-C. Deep Learning Neural Network Potential for Simulating Gaseous Adsorption in Metal–Organic Frameworks. *Mater. Adv* **2022,** *3*, 5299-5303.
(77) 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*, 111666.
(78) Du, M.; Liu, N.; Hu, X. Techniques for Interpretable Machine Learning. *Commun. ACM* **2019,** *63*, 68-77.
(79) Vaswani, A. S., N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; Polosukhin, I. In *Advances in neural information processing systems*, 2017.
(80) Wang, J.; Liu, J.; Wang, H.; Zhou, M.; Ke, G.; Zhang, L.; Wu, J.; Gao, Z.; Lu, D. A Comprehensive Transformer-Based Approach for High-Accuracy Gas Adsorption Predictions in Metal-Organic Frameworks. *Nat. Commun.* **2024,** *15*, 1904.
(81) Cao, Z.; Magar, R.; Wang, Y.; Barati Farimani, A. Moformer: Self-Supervised Transformer Model for Metal–Organic Framework Property Prediction. *J. Am. Chem. Soc.* **2023,** *145*, 2958-2967.
(82) Kang, Y.; Park, H.; Smit, B.; Kim, J. A Multi-Modal Pre-Training Transformer for Universal Transfer Learning in Metal–Organic Frameworks. *Nature Machine Intelligence* **2023,** *5*, 309-318.
(83) Lee, S.; Kim, B.; Cho, H.; Lee, H.; Lee, S. Y.; Cho, E. S.; Kim, J. Computational Screening of Trillions of Metal-Organic Frameworks for High-Performance Methane Storage. *ACS Appl. Mater. Interfaces* **2021,** *13*, 23647-23654.

(84) Bao, Y.; Martin, R. L.; Simon, C. M.; Haranczyk, M.; Smit, B.; Deem, M. W. In Silico Discovery of High Deliverable Capacity Metal–Organic Frameworks. *J. Phys. Chem. C* **2014,** *119*, 186-195.
(85) Demir, H.; Daglar, H.; Gulbalkan, H. C.; Aksu, G. O.; Keskin, S. Recent Advances in Computational Modeling of Mofs: From Molecular Simulations to Machine Learning. *Coord. Chem. Rev.* **2023,** *484*, 215112.
(86) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. Lammps - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales. *Comput. Phys. Commun.* **2022,** *271*, 108171.
(87) Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. Raspa: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials. *Mol. Simul.* **2015,** *42*, 81-101.
(88) Wilmer, C. E.; Kim, K. C.; Snurr, R. Q. An Extended Charge Equilibration Method. *J. Phys. Chem. Lett.* **2012,** *3*, 2506-11.
(89) Addicoat, M. A.; Vankova, N.; Akter, I. F.; Heine, T. Extension of the Universal Force Field to Metal-Organic Frameworks. *J. Chem. Theory. Comput.* **2014,** *10*, 880-91.
(90) Martin, M. G.; Siepmann, J. I. Transferable Potentials for Phase Equilibria. 1. United-Atom Description of N-Alkanes. *J. Phys. Chem. B* **1998,** *102*, 2569-2577.
(91) Lorentz, H. A. Ueber Die Anwendung Des Satzes Vom Virial in Der Kinetischen Theorie Der Gase. *Ann. Phys. (Berlin, Ger.)* **1881,** *248*, 127-136.
(92) Oh, C.; Nandy, A.; Yue, S.; Kulik, H. J. Mofs with the Stability for Practical Gas Adsorption Applications Require New Design Rules. *Zenodo* **2024**, DOI:https://doi.org/10.5281/zenodo.13146670 https://doi.org/10.5281/zenodo.13146670.
(93) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011,** *12*, 2825-2830.
(94) Martin, R. L.; Smit, B.; Haranczyk, M. Addressing Challenges of Identifying Geometrically Diverse Sets of Crystalline Porous Materials. *J. Chem. Inf. Model.* **2012,** *52*, 308-18.
(95) Willems, T. F.; Rycroft, C. H.; 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*, 134-141.
(96) Janet, J. P.; Kulik, H. J. Resolving Transition Metal Chemical Space: Feature Selection for Machine Learning and Structure-Property Relationships. *J. Phys. Chem. A* **2017,** *121*, 8939-8954.
(97) Ioannidis, E. I.; Gani, T. Z.; Kulik, H. J. Molsimplify: A Toolkit for Automating Discovery in Inorganic Chemistry. *J. Comput. Chem.* **2016,** *37*, 2106-17.
(98) Breiman, L. Random Forests. *Machine Learning* **2001,** *45*, 5-32.
(99) Tan, J. C.; Bennett, T. D.; Cheetham, A. K. Chemical Structure, Network Topology, and Porosity Effects on the Mechanical Properties of Zeolitic Imidazolate Frameworks. *Proc. Natl. Acad. Sci. U. S. A.* **2010,** *107*, 9938-43.

(100) McKellar, S. C.; Moggach, S. A. Structural Studies of Metal-Organic Frameworks under High Pressure. *Acta Crystallogr., Sect. B: Struct. Sci., Cryst. Eng. Mater.* **2015,** *71*, 587-607.
(101) Yang, L.; Chang, G.; Wang, D. High and Selective Carbon Dioxide Capture in Nitrogen-Containing Aerogels Via Synergistic Effects of Electrostatic in-Plane and Dispersive Pi-Pi-Stacking Interactions. *ACS Appl. Mater. Interfaces* **2017,** *9*, 15213-15218.
(102) Ahmed, A.; Seth, S.; Purewal, J.; Wong-Foy, A. G.; Veenstra, M.; Matzger, A. J.; Siegel, D. J. Exceptional Hydrogen Storage Achieved by Screening Nearly Half a Million Metal-Organic Frameworks. *Nat. Commun.* **2019,** *10*, 1568.
(103) Zou, C.; Wu, C. D. Functional Porphyrinic Metal-Organic Frameworks: Crystal Engineering and Applications. *Dalton Trans.* **2012,** *41*, 3879-88.
(104) Burnett, B. J.; Barron, P. M.; Choe, W. Recent Advances in Porphyrinic Metal–Organic Frameworks: Materials Design, Synthetic Strategies, and Emerging Applications. *CrystEngComm* **2012,** *14*, 3839.
(105) Dietzel, P. D. C.; Blom, R.; Fjellvåg, H. Base-Induced Formation of Two Magnesium Metal-Organic Framework Compounds with a Bifunctional Tetratopic Ligand. *Eur. J. Inorg. Chem.* **2008,** *2008*, 3624-3632.
(106) Wu, H.; Simmons, J. M.; Srinivas, G.; Zhou, W.; Yildirim, T. Adsorption Sites and Binding Nature of Co2 in Prototypical Metal−Organic Frameworks: A Combined Neutron Diffraction and First-Principles Study. *J. Phys. Chem. Lett.* **2010,** *1*, 1946-1951.
(107) Petit, C. Present and Future of Mof Research in the Field of Adsorption and Molecular Separation. *Curr. Opin. Chem. Eng.* **2018,** *20*, 132-142.
(108) Maia, R. A.; Louis, B.; Gao, W.; Wang, Q. Co2 Adsorption Mechanisms on Mofs: A Case Study of Open Metal Sites, Ultra-Microporosity and Flexible Framework. *React. Chem. Eng.* **2021,** *6*, 1118-1133.

For Table of Contents Use Only