Title

Diversifying Databases of Metal Organic Frameworks for High-Throughput Computational Screening

Journal

ACS Applied Materials & Interfaces, 13(51)

ISSN

1944-8244

Authors

Majumdar, Sauradeep
Moosavi, Seyed Mohamad
Jablonka, Kevin Maik
et al.

Publication Date

2021-12-29

DOI

10.1021/acsami.1c16220

Peer reviewed

ACS APPLIED MATERIALS & INTERFACES

www.acsami.org

Research Article

Diversifying Databases of Metal Organic Frameworks for High-Throughput Computational Screening

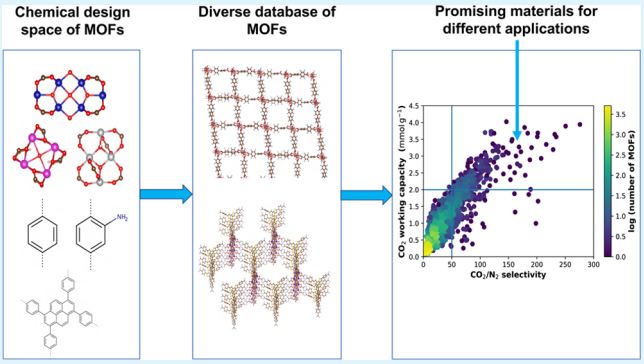
Sauradeep Majumdar, Seyed Mohamad Moosavi, Kevin Maik Jablonka, Daniele Ongari, and Berend Smit*

Cite This: *ACS Appl. Mater. Interfaces* 2021, 13, 61004–61014

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: By combining metal nodes and organic linkers, an infinite number of metal organic frameworks (MOFs) can be designed in silico. Therefore, when making new databases of such hypothetical MOFs, we need to ensure that they not only contribute toward the growth of the count of structures but also add different chemistries to the existing databases. In this study, we designed a database of ∼20,000 hypothetical MOFs, which are diverse in terms of their chemical design space—metal nodes, organic linkers, functional groups, and pore geometries. Using machine learning techniques, we visualized and quantified the diversity of these structures. We find that on adding the structures of our database, the overall diversity metrics of hypothetical databases improve, especially in terms of the chemistry of metal nodes. We then assessed the usefulness of diverse structures by evaluating their performance, using grand-canonical Monte Carlo simulations, in two important environmental applications—post-combustion carbon capture and hydrogen storage. We find that many of these structures perform better than widely used benchmark materials such as Zeolite-13X (for post-combustion carbon capture) and MOF-5 (for hydrogen storage). All the structures developed in this study, and their properties, are provided on the Materials Cloud to encourage further use of these materials for other applications.



KEYWORDS: *MOFs, molecular simulations, machine learning, diversity, carbon capture, hydrogen storage*

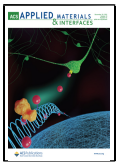
1. INTRODUCTION

Metal organic frameworks (MOFs) have been an exciting class of crystalline nanoporous materials since their discovery about 2 decades ago. By combining metal nodes and organic linkers, one can, in principle, make an infinite number of MOFs.[1] Over 100,000 MOFs have already been currently synthesized.[2−4] Due to characteristics like high surface area, large pore volume, and wide range of pores sized from micro- to mesoscale, MOFs have found applications in several areas like gas storage,[5] catalysis,[6] and nondistillative separations.[7−9]

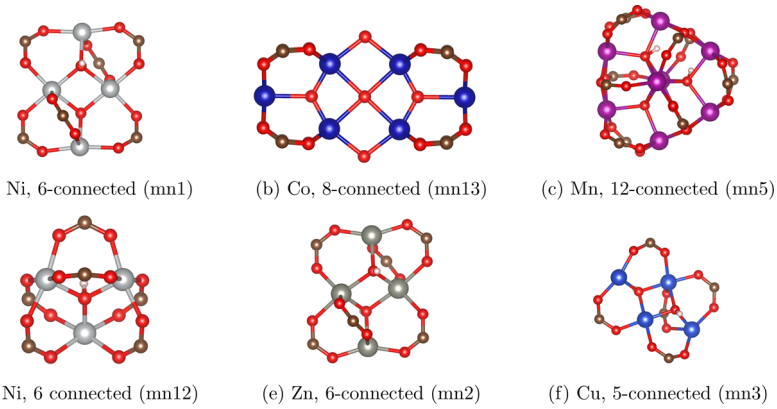
At present, of the millions of possible MOF structures, in practice, we can only synthesize a small fraction of all possible structures. This is because synthesizing a new MOF and then characterizing and testing it could take many months.[10] Therefore, computational researchers have been building databases of hypothetical MOF structures for high-throughput screening purposes. The idea stems from the fact that efficient computational algorithms can help in generating MOF structures and evaluating them for different applications in a less expensive and faster way. Ranking these hypothetical MOFs based on specific material properties then helps in identifying the most promising materials for a specific application.[4,11] Experimentally, we can then focus our efforts on synthesizing only the promising materials. Thus, along with the 100,000 experimental structures, there are also millions of hypothetical MOFs, which have been generated computationally.[10,12−15]

One of the earliest hypothetical MOF databases was generated by Wilmer et al.,[14] which consisted of a database of around 137,000 MOFs constructed using a 'Tinkertoy' algorithm. This database of hypothetical MOFs was used in various screening studies for gas storage and separations, and some of those hypothetical MOFs have been experimentally synthesized.[16] This approach of MOF construction, however, had a limitation. These 137,000 MOFs only sample from six topologies, with most of them having a **pcu** topology. What this algorithm did was it sequentially connected the molecular building blocks (SBUs) until a period crystal was formed, or in

Received: August 24, 2021
Accepted: December 3, 2021
Published: December 15, 2021



https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014



(a) Ni, 6-connected (mn1) (b) Co, 8-connected (mn13) (c) Mn, 12-connected (mn5)

(d) Ni, 6 connected (mn12) (e) Zn, 6-connected (mn2) (f) Cu, 5-connected (mn3)

Figure 1. Some metal nodes used in this study to generate hypothetical MOFs. The metal type, connection type, and the metal node names as we used in this study are provided below each node. (a) 6-connected Ni metal node. (b) 8-connected Co metal node. (c) 12-connected Mn metal node. (d) 6-connected Ni metal node. (e) 6-connected Zn metal node. (f) 5-connected Cu metal node. These metal nodes are shown to highlight different metals, connectivities, and geometries used in this study. The entire list of metal nodes used in this study is provided in the Supporting Information.

other words, it used a bottom-up approach for generating MOFs.

Along with the building blocks, topologies also play an important role in MOF performance. The net of a MOF, also called a topology, represents the underlying connectivity of the metal nodes and organic linkers. Gomez-Gualdron et al.[17] showed in their computational study of Zr-MOFs for volumetric methane storage how Zr-MOFs based on **ftw** topology outperform Zr-MOFs based on **scu** and **csq** topologies, even if the same organic linkers are used for all the three topologies. Subsequent algorithms to generate MOFs explore the different topologies using a top-down MOF construction algorithms. One such topology-based MOF construction algorithm has been developed by Boyd and Woo[12,18] and is called ToBasCCo. This study generated around 325,000 MOFs, which were screened for post-combustion carbon capture, and two of those structures, namely, Al-PMOF and Al-PyrMOF have also been synthesized. Another topology-based MOF construction algorithm was developed by Gomez-Gualdron et al.[13] and is called ToBaCCo. Hypothetical MOFs generated using ToBaCCo have been used for screening applications in hydrogen storage, methane storage, and xenon−krypton separation. Some structures from this database were also synthesized and tested.[13] Recently, a study by Lee et al.[15] presented an algorithm to explore a MOF space of over 100 trillion materials, which was used to find the most optimal structures for methane storage.

At present, there are thus several databases with millions of hypothetical MOF structures in total. The algorithms underlying these databases have been focused on enumerating as many possible structures for a given topology, metal node, organic linker, and functional group. The end result is that we have now reached such a large number of structures that it is practically impossible to screen all structures for a possible application. In addition, as we have an infinite number of possible structures, this is a fundamental problem we cannot solve with faster computers. It is therefore important to take a different approach and carefully select a representative set of diverse structures as a starting point for a screening study and, subsequently, use the strategy of adding only novel structures if in our set of most diverse materials some materials are missing.

In this respect, a detailed analysis of the diversity of the computation-ready experimental metal−organic framework (CoRE MOF) database,[3] which represents the synthesized MOFs from the Cambridge Structural Database (CSD),[19] and hypothetical databases was performed by Moosavi et al.[20] Descriptors were built to capture features of a MOF, such as pore geometry, metal chemistry, linker chemistry, and functional groups, which combine to form the chemical design space for a MOF chemist. The chemical diversity of a MOF was then expressed in terms of these features. Moosavi et al.[20] concluded that with respect to pore geometry, linker chemistry, and functional groups, the hypothetical databases seem to sufficiently well covered. However, with respect to metal chemistry, the hypothetical databases turn out to be less explored. The variety of metal chemistry in hypothetical databases was found to be surprisingly low, when compared to those in the experimental databases.[20] Hence, there are many MOF structures corresponding to these missing metal nodes, lying in the less explored regions of the material space. It is therefore important that we are able to generate such structures to study their properties. By harvesting these metal nodes using the algorithms developed in this respect,[15,20,21] one could generate such structures.

In this study, we thus designed a database of ∼20,000 hypothetical MOFs, keeping in mind their chemical diversity in terms of pore geometry, metal chemistry, linker chemistry, and functional groups. We focused on improving the diversity of metal nodes in hypothetical MOF databases by harvesting metal nodes from experimental structures. Diversity of metal nodes can be important for important environmental applications like carbon capture.[20] We are interested in carbon capture and storage because it is considered to be one of the most promising and viable technologies to address the rising CO_2 emissions in the atmosphere.[22,23] In this study, we have specifically looked into post-combustion carbon capture. Another application we have looked into here is the storage of hydrogen, a promising vehicular fuel.[24,25] Promising structures found from this screening study could then be added to the list of already available top-performing structures and the resultant list of structures would thus be chemically more diverse. This would then also help to choose from a wider range of structures and try synthesizing them.

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014

2. METHODS

2.1. Building Block Selection and Structure Generation. Moosavi et al.[20] developed a methodology to mine metal nodes from experimental MOF databases. These are some of the metal nodes, which are not commonly used for structure generation in hypothetical MOF databases. Thus, in this work, we focused on some of these metal nodes, as a proof of concept to validate our argument of improving metal diversity. Figure 1 shows some of these metal nodes. In total, 14 metal nodes have been used and they are all listed in the Supporting Information. We have chosen metal nodes consisting of different metals such as nickel, zinc, cadmium, copper, manganese, cobalt, and lead. Additionally, we have included metal nodes with different connectivities, ranging from 4-connected nodes up to 12-connected nodes, as well as different coordination geometries such as triangular, tetranuclear, square-planar, and others. There are several libraries of organic linkers reported in literature like the ToBaCCo database[9,13] and ToBasCCo database.[12,18] We selected our organic linkers from these reported libraries. We have also used several functional groups to decorate these organic linkers (see Supporting Information). Within ToBaCCo, there is a list of all the topologies from the Reticular Chemistry Structure Resource Database (RCSR)[26] (see Supporting Information for the list of all topologies used in this study). The topologies were selected from this list, based on their compatibility with the building blocks—metal nodes and organic linkers. It is to be noted that the list of topologies used in this study is not exhaustive. One could, in principle, generate even more structures by exploring more topologies. Because our focus in this study was on the diversity of the structures and not on the number of structures in itself, we did not explore all possible topologies for a given set of building blocks. The building blocks along with the topologies were then used to build the hypothetical MOFs using the ToBaCCo algorithm.

2.2. Structure Optimization and Charge Generation. The hypothetical MOFs of our database were optimized using the universal force field (UFF).[27] The optimization of the structures was performed using LAMMPS,[28] the input and data files for which were generated using lammps_interface.[29] The EQeq (extended charge equilibration) method[16] was used to generate the partial charges of the framework atoms of the hypothetical MOFs designed in this work.[30] Additional details on the structure optimization and charge generation process are provided in the Supporting Information.

2.3. Diversity Analysis. To analyze the diversity of MOF databases, we used a set of descriptors to quantify the similarity of MOF structures. Because both pore geometry and material chemistry are important in gas separation applications, we need descriptors for both aspects. Several material descriptors have been developed to characterize different aspects of the similarity of MOF materials.[31−34] We used classic geometric characteristics, such as the largest included sphere, surface area, density, and pore volume to describe the pore geometry. These descriptors were computed using Zeo++.[35,36] We described the chemistry of MOF structures using revised-autocorrelations (RACs). RACs are the product or difference of atomic heuristics, for example, Pauling electronegativity, connectivity, and covalent radii, computed on a molecular or crystal graph.[37] While RACs were initially introduced for machine-learning open shell transition metal complex properties,[37−39] they were recently adapted to MOF chemistry[20] and shown to be successful in capturing structure−property relationships for gas adsorption[20] and photoelectronic properties (e.g., color) of MOFs.[40] In this approach, the MOF structure is described with three groups of features, describing the metal centres, organic linkers, and the functional groups. In total, 156 RAC descriptors were computed using the molSimplify package[38,41] to describe the chemistry of a MOF structure.

We computed variety, balance, and disparity to assess the diversity of the material databases. The diversity metrics were calculated for each aspect of the material chemistry that includes chemistry of metals, linkers, functional groups, and the pore geometry. The chemical and geometric descriptors construct high-dimensional feature spaces. We first split these high-dimensional spaces into 1000 bins using the *k*-means clustering method. In this approach, the structures were assigned to their closest centroid. Then, the three diversity metrics were computed using this binning. Each diversity factor captures different information related to the diversity. These diversity metrics—variety, balance, and disparity—are also used in a wide range of other fields like understanding the stability of ecosystems, social sciences.[42−46] Following Moosavi et al.,[20] variety has been calculated as the percentage of all the bins sampled by a given database, that is, how many district types of structures exist in a database normalized with the 1000 unique bins. The balance of a database gives us an indication of how even is the distribution of structures in a database. For example, let us say that in database 1, we have 100 structures of type A and 2 structures of type B, and in database 2, if we have 70 structures of type A and 50 structures of type B. The variety is the same in both databases, but the balance is very low in database 1. There is thus a bias toward structures of type A in database 1. We used Pielou's evenness,[47] which measures how even the structures are distributed among the sampled bins, as a measure of the balance. Following Moosavi et al.,[20] the evenness of the distribution of structures—balance—could be computed using different methods, which are all transformations of the Shannon entropy.[20] The Shannon entropy is given by

$$H(X) = -\sum P(x_i)\log P(x_i) \tag{1}$$

The maximum entropy would be achieved in case of a uniform distribution. Therefore, normalizing the system entropy with the maximum entropy (when all bins are equally likely) would give us a metric for evenness—relative entropy.

$$H_{\text{rel}}(X) = \frac{\exp(H(X))}{\exp(H_{\text{max}})} \tag{2}$$

One transformation of the entropy was introduced by Pielou

$$\text{PL}_{\text{rel}}(X) = \frac{1 - \exp(H(X))}{1 - \exp(H_{\text{max}})} \tag{3}$$

We used $1 - \text{PL}_{\text{rel}}(X)$ in this study to measure the evenness of distribution, such that 1 is the maximum evenness, that is, uniform distribution. The disparity metric gives us a measure of the spread of the structures in a database. A high value of disparity would mean that the database contains significantly dissimilar structures that are far apart from each other in the material space.To compute disparity, we computed the covered area of the concave hull by a database in the map of the first two principal components. We normalized this number with the area of all databases together. The covered area was computed using the Shapely package with the circumference to area ratio cutoff of 1.[48] A detailed description of the material descriptors and diversity analysis can be found in the previous work.[20]

2.4. Property Calculation. The pore limiting diameter and blocking spheres for each MOF were calculated using Zeo++.[35,36] For blocking spheres, we considered spherical probes with diameters of 3.05 Å for CO_2 (oxygen's sigma in TraPPE), 3.31 Å for N_2 (nitrogen's sigma in TraPPE), and 2.96 Å for H_2 (hydrogen's sigma in the Buch force field[49]). The force-field parameters for the framework atoms were extracted from UFF.[27] CO_2 and N_2 molecules were described by the TraPPE force field,[50] and H_2 was described by the Buch force field[49] with the Feynman Hibbs correction[51] (see Supporting Information for the full list of parameters used). The gas−framework interactions were modeled using Lennard Jones potential, truncated at 12 Å (for CO_2 and N_2) and 12.8 Å (for H_2), with tail corrections.[52] The Lennard Jones interactions between dissimilar atoms were approximated using Lorentz−Berthelot rules.[53] The Coulombic electrostatic interactions were computed using Ewald summation. The gas adsorption calculations were performed in RASPA.[54] Grand-canonical Monte Carlo (GCMC) simulations were used to compute the gas uptake of the MOFs. Each calculation consisted of 10,000 equilibration cycles followed by 10,000 production cycles. In RASPA, a cycle is defined as max(20,*N*) steps where *N* is the number of

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014

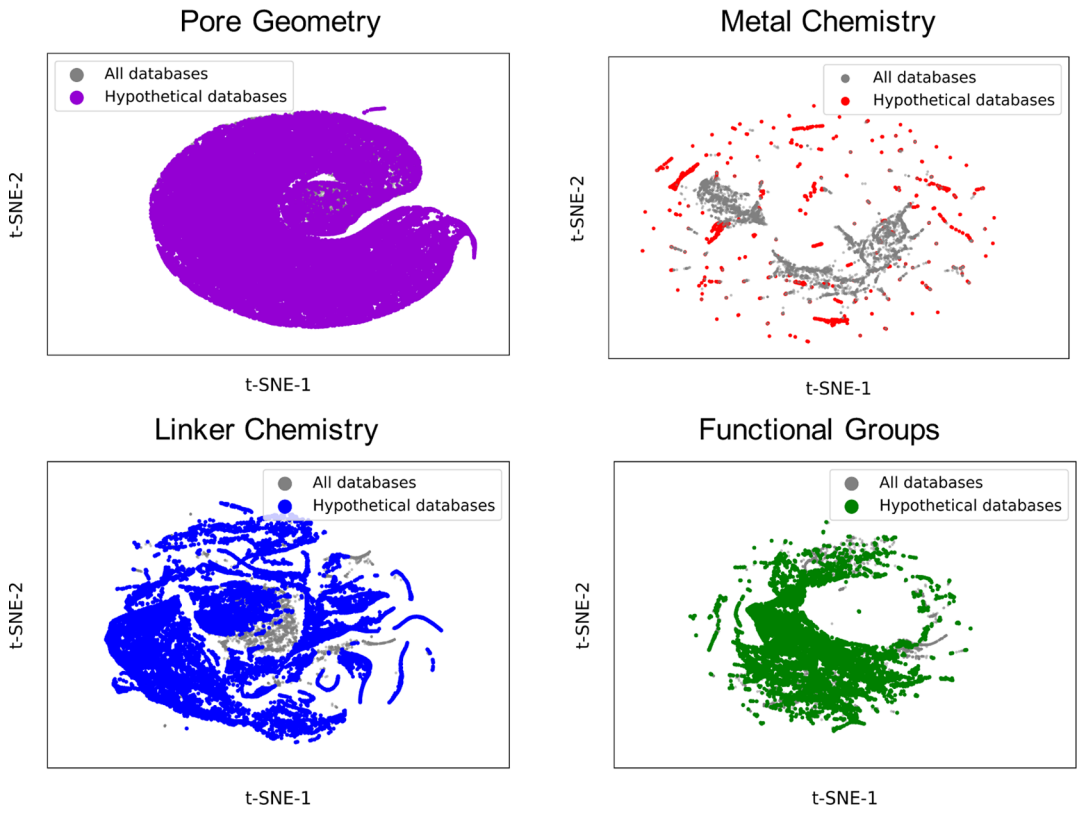


Figure 2. Visualization of the material design space. The t-Distributed Stochastic Neighbour Embedding (t-SNE)[59] method was used to project the pore geometry, metal chemistry, linker chemistry, and functional groups descriptor spaces to 2D maps. The t-SNE method preserves pairwise distances, ensuring that similar structures are mapped close to each other in two dimensions. (See principal component analysis figures in Supporting Information for the global similarities.) Only descriptors up to the second coordination shell were included for metal chemistry to emphasize the local metal chemistry environment. The entire known design space, containing the structures from all databases—experimental and hypothetical—is represented in gray. The structures from all the hypothetical databases were colored and overlaid on this design space. Thus, the colored regions represent those parts of the design space, which are covered by all the hypothetical databases. The gray regions represent those parts of the design space, which are not covered by the hypothetical databases.

molecules.[54] The pure component CO_2 uptakes were calculated at 1 bar and 298 K. We also calculated the uptakes of CO_2 and N_2 for a binary mixture of CO_2 and N_2 in the ratio of 0.15:0.85. For the binary mixture, we considered the flue gas to be adsorbed at 1 bar and 298 K and regenerated at 0.1 bar and 363 K. These conditions have been used in several studies for post-combustion CO_2 capture.[12,55−57] The H_2 uptakes were calculated at 100 bar and 77 K. These conditions have been used in several studies for hydrogen storage.[9,24]

3. RESULTS AND DISCUSSION

3.1. Diversity Analysis. Using the workflow as described in the Methods section, a database of ∼20,000 MOFs was generated. We took the combination of all synthesized MOF structures and hypothetical MOF structures as the current chemical space of MOFs. This chemical space has been described using the high-dimensional pore geometry and chemistry feature vectors, and we have thus made a projection of it on two dimensions to visualize which regions of the material design space our hypothetical MOFs are covering. For experimental structures, we have considered the CoRE-2019 database.[3] For hypothetical structures, we have considered the database developed by Anderson et al.,[58] the ToBaCCo database,[13] a diverse subset of 20,000 structures from the database developed by Boyd and Woo.[12] These databases were used for the analysis in the work by Moosavi et al.[20]

Figure 2 shows a dimensionality reduction visualization of all hypothetical MOF databases, including the database we have developed in this study, when overlaid on the total set of all experimental and hypothetical MOF databases. The distributions of the databases are shown with respect to their pore geometry, metal chemistry, linker chemistry, and functional groups. For pore geometry, linker chemistry, and functional groups, the hypothetical databases are covering and sampling the design space well. For metal chemistry, we find that the sampling of the design space has improved on including the structures from this study, when compared to the previous distribution reported by Moosavi et al.[20] This overall improvement of the diversity in metal chemistry has been quantified below (Table 1).

Figure 3 shows the regions of the material design space we have specifically contributed to through the database of this study. As discussed in the Methods section, for organic linkers, functional groups, and topologies, we have selected them from the respective libraries reported in literature. However, for metal nodes, we have tried to focus on the ones that have not been commonly used in the other hypothetical databases mentioned in this study. Thus, if we look at the metal chemistry map, we find that our metal nodes are complementing different regions of the space as compared to the previously used metal nodes in the other hypothetical databases.[20] Also, when we combine all these metal nodes used in all the hypothetical databases together, we get the map of metal chemistry, as shown in Figure 2.

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014

Table 1. Diversity Metrics for the Different Features of Hypothetical Databases[a]

feature	hypothetical databases	variety	balance	disparity
geometric	excluding this study	0.977	0.849	0.874
	including this study	0.988	0.775	0.933
metal center	excluding this study	0.068	0.334	0.078
	including this study	0.107	0.296	0.104
linker chemistry	excluding this study	0.648	0.617	0.737
	including this study	0.684	0.446	0.798
functional group	excluding this study	0.722	0.213	0.782
	including this study	0.851	0.323	0.834

[a]We first split the high-dimensional spaces into 1000 bins using the k-means clustering method. Variety measures the percentage of all the bins sampled by a given database. Balance measures the evenness of the distribution of the structures among the sampled bins. And, disparity measures the spread of the sampled bins. We normalized these number with the area of all databases together.

We have quantified the diversity of the databases in terms of their variety, balance, and disparity (Table 1). The variety of a database indicates how many distinct types of structures exist in our database. Balance indicates how even the distribution of structures is. Disparity of a database reflects how dissimilar or distinct the structures of our database are. A high disparity would thereby indicate that we have structures from far apart points in the material design space. We thus calculated these metrics for the hypothetical databases in two scenarios: before adding the database of this study and after adding the database of this study.

For the geometric features of the hypothetical databases, we find a slight increase in the variety and the disparity and a slight decrease in the balance. For the metal center features, we see that on adding the structures from the database of this study, the variety of structures have improved. The balance of the structures decreases slightly, and the disparity of the structures also improves. This gives us an indication that the overall diversity of structures with respect to the metal chemistry has improved upon adding the structures from the database of this study. For the linker chemistry, as like the geometric features, we see a slight increase in the variety and the disparity and a decrease in the balance. Also, for functional groups, we see an improvement in all the three diversity metrics.

Now that we have designed a diverse set of hypothetical MOF structures, our next aim was to see if we could use them in some practical applications.

3.2. Post-Combustion Carbon Capture. Figure 4 shows the distribution of the hypothetical MOFs of the current study for the uptake of pure CO_2 at 1 bar and 298 K. A reference line has been drawn to denote the pure CO_2 uptake of Zeolite-13X, which is often used as a benchmark CO_2 adsorbent.[12] From the distribution, we find that there are many structures, which perform as well as Zeolite-13X—pure CO_2 uptake of ~5 mmol g^{-1},[55,60] and there are also many structures—around 800—which surpass the performance of Zeolite-13X.

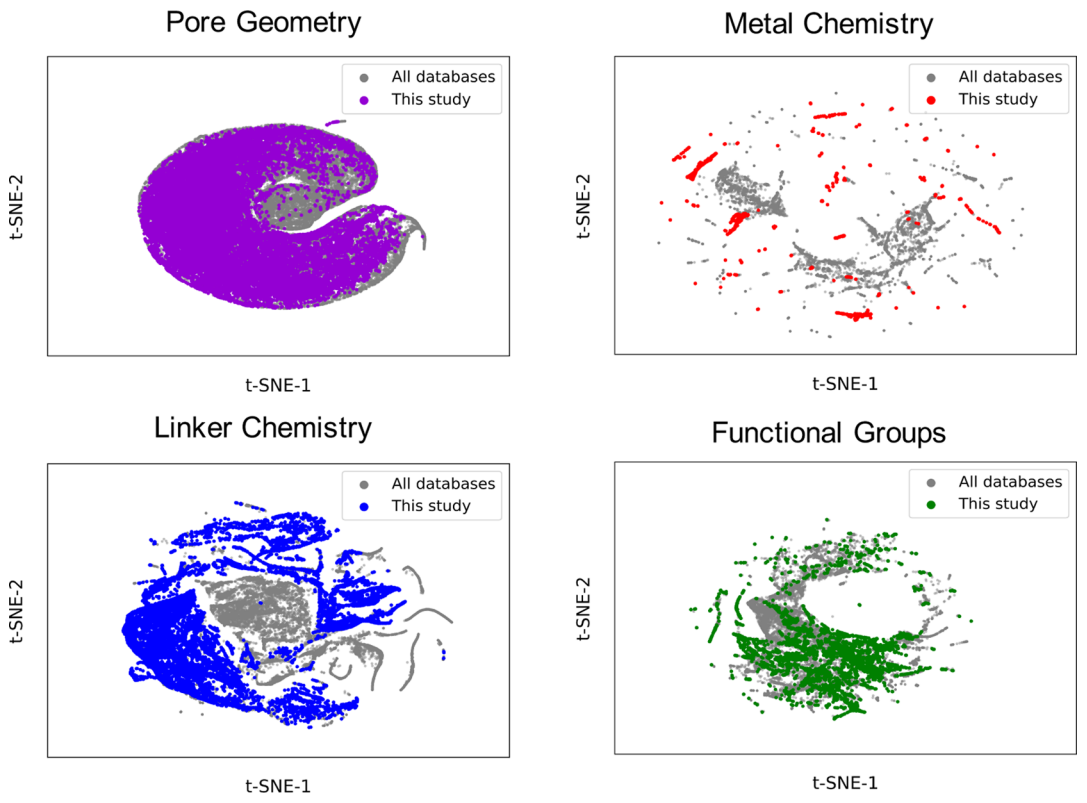


Figure 3. Visualization of the material design space. The t-SNE method was used to project the pore geometry, metal chemistry, linker chemistry, and functional groups descriptor spaces to 2D maps. Only descriptors up to the second coordination shell were included for metal chemistry to emphasize the local metal chemistry environment. The entire known design space, containing the structures from all databases—experimental and hypothetical—is represented in gray. The structures from the hypothetical database developed in this study were colored and overlaid on this design space. Thus, the colored regions represent those parts of the design space, which are covered by the hypothetical database developed in this study. The gray regions represent those parts of the design space that are not covered by the hypothetical database developed in this study.

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014

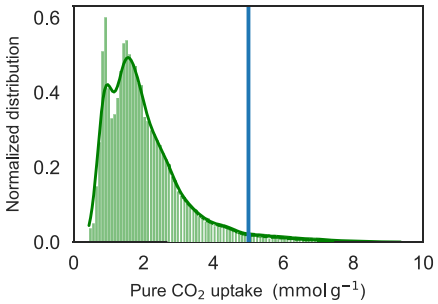


Figure 4. Results from the computational screening of ~20,000 MOFs for post-combustion carbon capture (pure CO_2 adsorption at 1 bar and 298 K). This plot shows the distribution of the pure CO_2 uptake of the MOFs. The blue reference line denotes the pure CO_2 uptake of Zeolite-13X.

We further investigated the structures for their performance in separating CO_2 from flue gas. For this, we considered a binary mixture of CO_2 and N_2. We then calculated the CO_2 working capacity and CO_2/N_2 selectivity of these hypothetical structures. Figure 5 shows how the structures perform. Again,

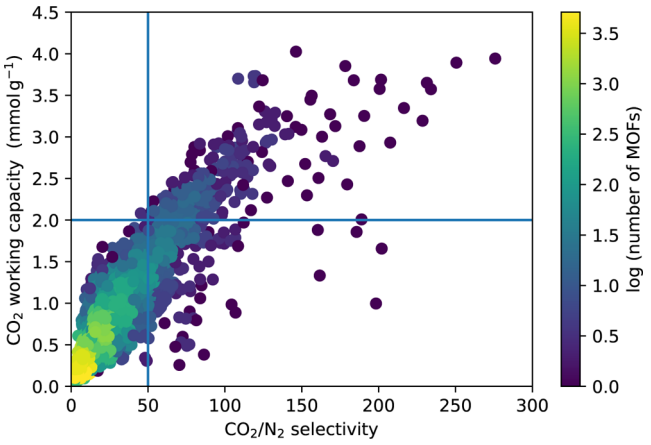


Figure 5. Results from the computational screening of ~20,000 MOFs for post-combustion carbon capture (15:85 CO_2/N_2 mixture with adsorption at 1 bar and 298 K and regeneration at 0.1 bar and 363 K). This plot shows the CO_2 working capacity versus CO_2/N_2 selectivity of the MOFs. The color coding represents the number of MOFs according to the color bar on the right. The blue reference lines denote the CO_2 working capacity and CO_2/N_2 selectivity of Zeolite-13X.

we find that there are many structures—around 250—which surpass the performance of Zeolite-13X under dry conditions—CO_2 working capacity greater than ~2 mmol g^{-1} and CO_2/N_2 selectivity greater than ~50.[12,61]

Based on diversity analysis, Moosavi et al.[20] concluded that for CO_2 adsorption at low pressures, metal chemistry as a factor cannot be ignored. To illustrate the importance of using different metal nodes in our database and to have a look at how some of the metal nodes in our study performed, we plotted the pure CO_2 uptake at 1 bar and 298 K for some of the metal nodes used in this study. For example, Figure 6 shows the distribution of the pure CO_2 uptake for the metal node mn1—a Ni-based metal node, mn3—a Cu-based metal node, mn2—a Zn-based metal node, and mn12—a Ni-based metal node. Nodes mn1, mn2, and mn3 have similar tetranuclear metal clusters. However, in mn3, Cu forms a five-connected cluster—$Cu_4(\mu_3{-}OH)_2(COO)_5$,[62] and in mn1 and mn2, Ni and Zn form a six-connected cluster—$Ni_4(\mu_3{-}OH)_2(COO)_6$ and $Zn_4(\mu_3{-}OH)_2(COO)_6$, respectively.[63,64] Many of the MOF structures of this study containing nodes mn1 and mn2 have pure CO_2 uptakes above 5 mmol g^{-1}, while hardly any of the structures containing node mn12 have pure CO_2 uptakes above 5 mmol g^{-1}. This shows that even if we have metal nodes of similar geometry, the type of metal can impact the formation of a node with different connectivities and different CO_2 adsorption characteristics of the MOF.

If we then compare the distribution of the nodes mn1 and mn12, both metal nodes are made of Ni but have different geometries. In mn1, Ni forms a tetranuclear cluster, while in mn12, Ni forms a six-connected triangular cluster, $Ni_3(\mu_3{-}OH)_2(COO)_6$.[65] Also, the structures made of both these metal nodes perform quite well in their CO_2 uptakes. This shows how the same metal can form two nodes of different geometries and that the structures of both the nodes could be promising for CO_2 capture. Capturing these variations in the metal nodes is important because when we include all these metal nodes, we get the final distribution as shown in Figure 4, which is very different from the individual distributions. Also, the presence of these different metal nodes helps in obtaining many high performing structures. This would also help us to choose from a wider range of metal nodes, while synthesizing new MOFs for carbon capture.

We find that most of the top performing structures contain the metal nodes mn1 (Ni-based)—~40% of the top performing structures, mn12 (Ni-based)—~30%, mn13 (Co-based)—~8%, and mn2 (Zn based)—~5%. For the linkers, we find the top performing structures to have simple two-coordinated linkers like benzene dicarboxylic acids to more complicated three-coordinated linkers like benzene-1,3,5-tricarboxylic acid and triazine to further complicated six-coordinated linkers like bicyclooctanes and a combination of different linkers. Figure 7 shows the structure of one of the high performing MOFs for post-combustion carbon capture.

3.3. Hydrogen Storage.

Figure 8 shows a plot of the gravimetric uptake versus the volumetric uptake of H_2 in our structures at 100 bar and 77 K. This plot shows a volcano-type relationship between the two types of uptakes, as observed in previous studies.[9,24] A reference line has been drawn to denote the gravimetric H_2 uptake—9.20 wt %[24] and volumetric H_2 uptake—52.64 g L^{-1}[24] in MOF-5, a widely used benchmark material for H_2 storage selected by the Hydrogen Storage Engineering Centre of Excellence (HSECoE).[24,66,67] We find that many structures from our database have a gravimetric uptake higher than that of MOF-5. An ideal H_2 adsorbent should however exhibit a balance between high gravimetric uptake and high volumetric uptake.[24] This is because the volumetric uptake of the H_2 storage system has a greater impact on the driving range of fuel cell vehicles (FCVs) than the gravimetric uptake.[24,66−69] Also, in this respect, we find around 50 MOFs from our database, which outperform MOF-5. Figure 9 shows the structure of one such promising hypothetical MOF of this study for hydrogen storage.

Here, we find that almost all of the top performing structures have a **tfz-d** topology, as shown in Figure 10. Metal nodes, which form structures in these topologies, are thereby more prevalent in these top 50 structures. In this study, these are mainly eight connected metal nodes like mn13 (Co-based)—~85% of the top-performing structures and mn4 (Cd-based)—~15% of the top-performing structures. The explora-

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014

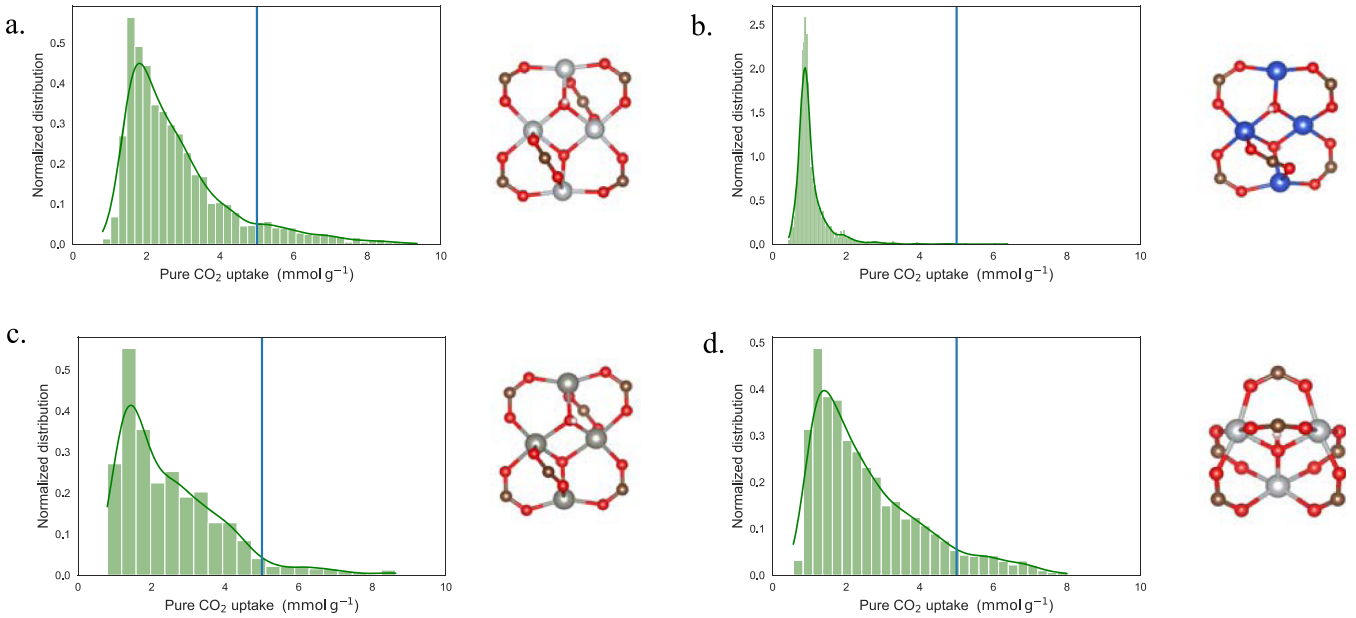


Figure 6. Pure CO_2 uptake distribution for MOFs of different metal nodes at 1 bar and 298 K. (a) Distribution of MOFs of node mn1 (left) and the structure of node mn1 (right). (b) Distribution of MOFs of node mn3 (left) and the structure of node mn3 (right). (c) Distribution of MOFs of node mn2 (left) and the structure of node mn2 (right). (d) Distribution of MOFs of node mn12 (left) and the structure of node mn12 (right). The number of bins for the MOFs of metal node mn2 are less than the other nodes because fewer MOFs were generated with this node mn2 (as it was one of the last metal nodes to be added in our database). The blue reference lines denote the pure CO_2 uptake of Zeolite-13X.

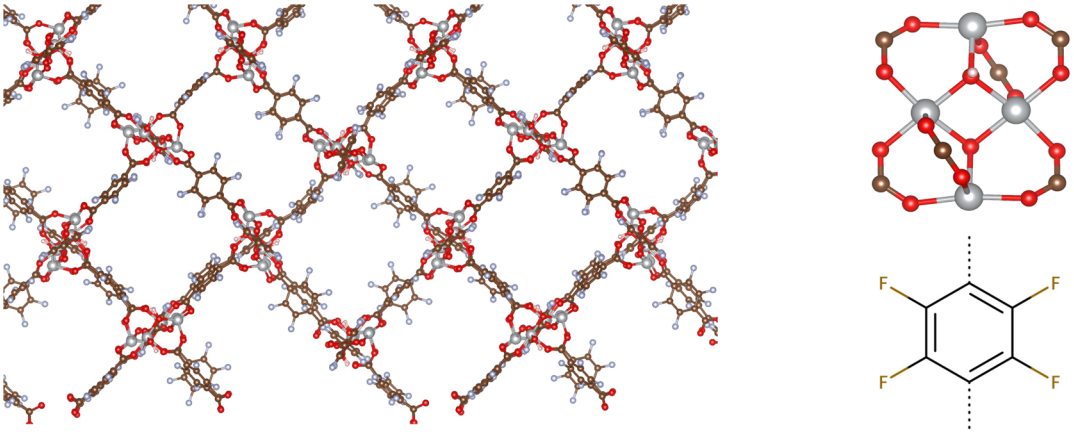


Figure 7. Structure of a top performing MOF for post-combustion carbon capture—ddmof_559—metal node mn1 + organic edge oe31 + topology **snk**.

tion of different topologies with these metal nodes led to the generation of MOFs with different pore geometries, which finally led to some of these structures to be promising for hydrogen storage. This again highlights the importance of having a diverse database. Here, we did not pre-bias the structures for a particular application. We tried to make a diverse set of structures, which covers different aspects of MOF chemistry. Some features of these structures play an important role in one application and some features in other applications. Therefore, some of these structures turn out to be good for one application and some for other. In this case of hydrogen storage, it is the topology of a MOF, which plays a more important role in forming a top-performing storage than metal chemistry.

It is important to note here that our analysis is based on the current state of the art methods used in screening studies, that is, generic force fields and rigid crystals. In our case, we used the UFF, which generally gives good predictions of the adsorption behavior, but for some classes of materials (open metal sites), it is known to underestimate the adsorption.[70] As open metal sites are very sensitive to water,[71] these materials are in practice less interesting for carbon capture applications; so, we did not attempt to correct these results. Also, for hydrogen adsorption at high pressures, UFF is reported to work reasonably well.[72]

4. CONCLUSIONS

In this study, we have designed a database of ~20,000 hypothetical MOFs with the aim to increase the chemical diversity of the existing databases. We show that adding the structures of our database improves the overall diversity metrics of hypothetical databases, especially in terms of metal chemistry.

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014

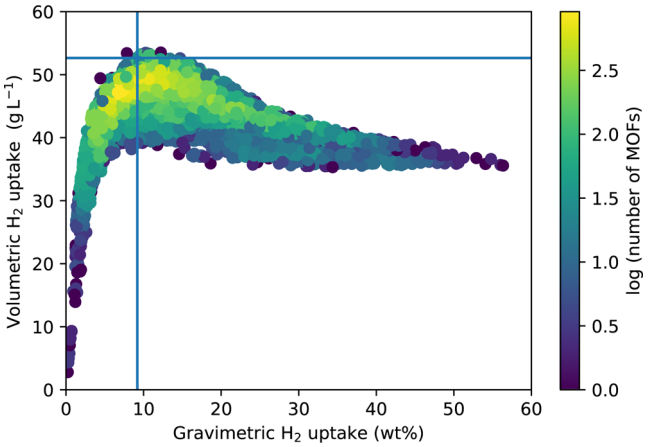


Figure 8. Results from the computational screening of ~20,000 MOFs for hydrogen storage (pure H_2 adsorption at 100 bar and 77 K). The colour coding represents the number of MOFs according to the colour bar on the right. The blue reference lines denote the volumetric H_2 uptake and gravimetric H_2 uptake of MOF-5.

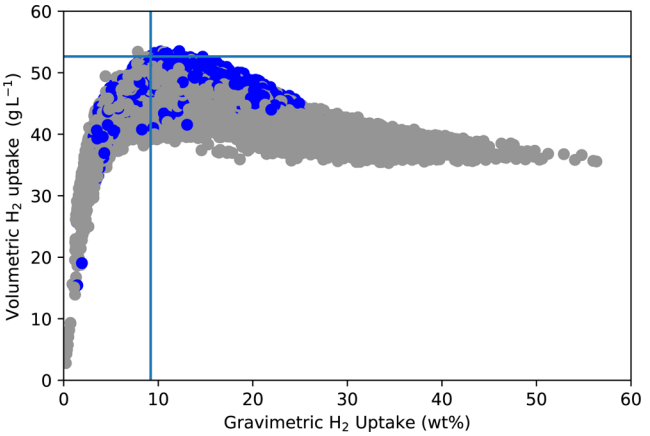


Figure 10. Results from the computational screening of ~20,000 MOFs for hydrogen storage (pure H_2 adsorption at 100 bar and 77 K). The structures with the **tfz-d** topology are highlighted in blue and all the remaining structures are in gray.

To highlight the usefulness of these diverse structures, we evaluated their performance for two important environmental applications—post-combustion carbon capture and hydrogen storage. In the case of post-combustion carbon capture, we find that many of these structures outperform Zeolite-13X, a widely used benchmark material for carbon capture, in terms of their pure CO_2 uptake—around 800 structures, CO_2 working capacity and CO_2/N_2 selectivity—around 250 structures. For hydrogen storage, we find around 50 structures, which outperform MOF-5, a widely used benchmark material for hydrogen storage, in terms of their balance between gravimetric uptake of H_2 and volumetric uptake of H_2. For post-combustion carbon capture, we find that including different metal nodes help in obtaining high performing structures. In the case of hydrogen storage, we find that it is the topology of the MOF, which plays the more dominant role for a structure to be high-performing. The promising structures found in this study could be added to the existing list of promising structures in literature and this would provide us with a more diverse range of materials to choose from, for synthesizing.

Through this study, we thus show that on starting with a relatively small but diverse set of materials, one could still obtain interesting materials for different applications. This would help us to locate the interesting regions of the material space. To avoid brute-force screening of an infinite number of possible MOFs, as a next step, one could then explore around these interesting regions using active learning, Bayesian optimization,[73,74] or generative models.[75,76]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsami.1c16220.

Additional details for MOF structure generation—list of metal nodes, organic linkers, functional groups, and topologies used to generate the hypothetical MOFs in this study; additional details for structure optimization and charge generation calculations; force-field parameters used for GCMC simulations; additional figures for diversity analysis; and additional figures for hydrogen storage (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Berend Smit − *Laboratory of Molecular Simulation (LSMO), Institut des Sciences et Ingénierie Chimiques, École*

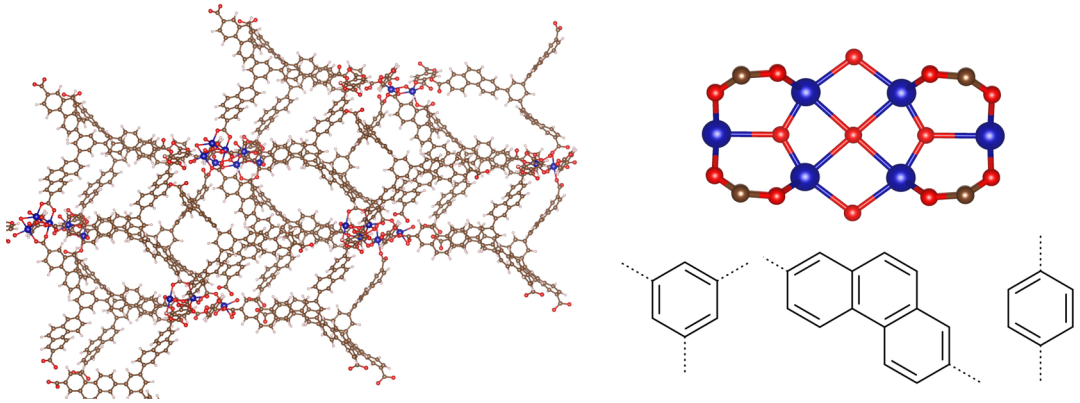


Figure 9. Structure of a top performing MOF for hydrogen storage—ddmof_6749—metal node mn13 + organic node on1 + organic edges oe33, oe68 + **tfz-d** topology.

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014

Polytechnique Fédérale de Lausanne (EPFL), CH-1951 Sion, Valais, Switzerland; orcid.org/0000-0003-4653-8562; Email: berend.smit@epfl.ch

Authors

Sauradeep Majumdar − *Laboratory of Molecular Simulation (LSMO), Institut des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), CH-1951 Sion, Valais, Switzerland;* orcid.org/0000-0002-2095-3082

Seyed Mohamad Moosavi − *Laboratory of Molecular Simulation (LSMO), Institut des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), CH-1951 Sion, Valais, Switzerland;* orcid.org/0000-0002-0357-5729

Kevin Maik Jablonka − *Laboratory of Molecular Simulation (LSMO), Institut des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), CH-1951 Sion, Valais, Switzerland;* orcid.org/0000-0003-4894-4660

Daniele Ongari − *Laboratory of Molecular Simulation (LSMO), Institut des Sciences et Ingénierie Chimiques, École Polytechnique Fédérale de Lausanne (EPFL), CH-1951 Sion, Valais, Switzerland;* orcid.org/0000-0001-6197-2901

Complete contact information is available at: https://pubs.acs.org/10.1021/acsami.1c16220

Notes

The authors declare no competing financial interest.

All of the MOF structures developed in this study and their properties are available on the Materials Cloud platform[77] at https://doi.org/10.24435/materialscloud:yn-de. The structures and properties of the experimental and hypothetical MOFs used for diversity analysis were collected from the Materials Cloud platform at https://doi.org/10.24435/materialscloud:3y-gr.

■ ACKNOWLEDGMENTS

S.M. would like to acknowledge Peter G. Boyd for his help and guidance in the MOF structure generation process. This research was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement 666983, MaGic), by the NCCR-MARVEL, funded by the Swiss National Science Foundation, and by the PrISMa Project (299659), funded through the ACT Programme (Accelerating CCS Technologies, Horizon 2020 Project 294766). Financial contributions from the Department for Business, Energy & Industrial Strategy (BEIS) together with extra funding from the NERC and EPSRC Research Councils, United Kingdom, the Research Council of Norway (RCN), the Swiss Federal Office of Energy (SFOE), and the U.S. Department of Energy are gratefully acknowledged. Additional financial support from TOTAL and Equinor is also gratefully acknowledged.

■ REFERENCES

(1) Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M. Systematic Design of Pore Size and Functionality in Isoreticular MOFs and Their Application in Methane Storage. *Science* **2002**, *295*, 469−472.

(2) Moghadam, P. Z.; Li, A.; Wiggin, S. B.; Tao, A.; Maloney, A. G. P.; Wood, P. A.; Ward, S. C.; Fairen-Jimenez, D. Development of a Cambridge Structural Database Subset: A Collection of Metal-Organic Frameworks for Past, Present, and Future. *Chem. Mater.* **2017**, *29*, 2618−2625.

(3) 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.* **2019**, *64*, 5985−5998.

(4) 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*, 1890−1900.

(5) Sculley, J.; Yuan, D.; Zhou, H.-C. The Current Status of Hydrogen Storage in Metal−Organic Frameworks—Updated. *Energy Environ. Sci.* **2011**, *4*, 2721−2735.

(6) Farha, O. K.; Shultz, A. M.; Sarjeant, A. A.; Nguyen, S. T.; Hupp, J. T. Active-Site-Accessible, Porphyrinic Metal Organic Framework. *J. Am. Chem. Soc.* **2011**, *133*, 5652.

(7) Simon, C. M.; Mercado, R.; Schnell, S. K.; Smit, B.; Haranczyk, M. What Are the Best Materials To Separate a Xenon/Krypton Mixture? *Chem. Mater.* **2015**, *27*, 4459−4475.

(8) Sikora, B. J.; Wilmer, C. E.; Greenfield, M. L.; Snurr, R. Q. Thermodynamic Analysis of Xe/Kr Selectivity in over 137000Hypothetical Metal −OrganicFrameworks. *Chem .Sci.* **2012**, *3*, 2217−2223.

(9) 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.

(10) Boyd, P. G.; Lee, Y.; Smit, B. Computational Development of the Nanoporous Materials Genome. *Nat. Rev. Mater.* **2017**, *2*, 17037.

(11) Moosavi, S. M.; Jablonka, K. M.; Smit, B. The Role of Machine Learning in the Understanding and Design of Materials. *J. Am. Chem. Soc.* **2020**, *142*, 20273−20287.

(12) Boyd, P. G.; Chidambaram, A.; García-Díez, E.; Ireland, C. P.; Daff, T. D.; Bounds, R.; Gładysiak, 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 CO2 Capture. *Nature* **2019**, *576*, 253−256.

(13) Gómez-Gualdrón, D. A.; Colón, Y. J.; Zhang, X.; Wang, T. C.; Chen, Y.-S.; Hupp, J. T.; Yildirim, T.; Farha, O. K.; Zhang, J.; Snurr, R. Q. Evaluating Topologically Diverse Metal−Organic Frameworks for Cryo-Adsorbed Hydrogen Storage. *Energy Environ. Sci.* **2016**, *9*, 3279−3289.

(14) 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.* **2012**, *4*, 83−89.

(15) 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.

(16) Wilmer, C. E.; Kim, K. C.; Snurr, R. Q. An Extended Charge Equilibration Method. *J. Phys. Chem. Lett.* **2012**, *3*, 2506−2511.

(17) Gomez-Gualdron, D. A.; Gutov, O. V.; Krungleviciute, V.; Borah, B.; Mondloch, J. E.; Hupp, J. T.; Yildirim, T.; Farha, O. K.; Snurr, R. Q. Computational Design of Metal-Organic Frameworks Based on Stable Zirconium Building Units for Storage and Delivery of Methane. *Chem. Mater.* **2014**, *26*, 5632−5639.

(18) Boyd, P. G.; Woo, T. K. A Generalized Method for Constructing Hypothetical Hanoporous Materials of Any Net Topology from Graph Theory. *CrystEngComm* **2016**, *18*, 3777−3792.

(19) 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*, 171−179.

(20) 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.

(21) Bucior, B. J.; Rosen, A. S.; Haranczyk, M.; Yao, Z.; Ziebel, M. E.; Farha, O. K.; Hupp, J. T.; Siepmann, J. I.; Aspuru-Guzik, A.; Snurr, R. Q. Identification Schemes for Metal−Organic Frameworks to Enable Rapid Search and Cheminformatics Analysis. *Cryst. Growth Des.* **2019**, *19*, 6682−6697.
(22) Bui, M.; Adjiman, C. S.; Bardow, A.; Anthony, E. J.; Boston, A.; Brown, S.; Fennell, P. S.; Fuss, S.; Galindo, A.; Hackett, L. A.; Hallett, J. P.; Herzog, H. J.; Jackson, G.; Kemper, J.; Krevor, S.; Maitland, G. C.; Matuszewski, M.; Metcalfe, I. S.; Petit, C.; Puxty, G.; Reimer, J.; Reiner, D. M.; Rubin, E. S.; Scott, S. A.; Shah, N.; Smit, B.; Trusler, J. P. M.; Webley, P.; Wilcox, J.; Mac Dowell, N. Carbon Capture and Storage (CCS): The Way Forward. *Energy Environ. Sci.* **2018**, *11*, 1062−1176.
(23) Smit, B.; Reimer, J. A.; Oldenburg, C. M.; Bourg, I. C. *Introduction to Carbon Capture and Sequestration*; World Scientific, 2014; Vol. *1*.
(24) Ahmed, A.; Liu, Y.; Purewal, J.; Tran, L. D.; Wong-Foy, A. G.; Veenstra, M.; Matzger, A. J.; Siegel, D. J. Balancing Gravimetric and Volumetric Hydrogen Density in MOFs. *Energy Environ. Sci.* **2017**, *10*, 2459−2471.
(25) 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*, 2844−2854.
(26) 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*, 1782−1789.
(27) Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A., III; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024−10035.
(28) Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular Dynamics. *J. Comput. Phys.* **1995**, *117*, 1−19.
(29) Boyd, P. G.; Moosavi, S. M.; Witman, M.; Smit, B. Force-Field Prediction of Materials Properties in Metal-Organic Frameworks. *J. Phys. Chem. Lett.* **2017**, *8*, 357−363.
(30) Ongari, D.; Boyd, P. G.; Kadioglu, O.; Mace, A. K.; Keskin, S.; Smit, B. Evaluating Charge Equilibration Methods to Generate Electrostatic Fields in Nanoporous Materials. *J. Chem. Theory Comput.* **2018**, *15*, 382−401.
(31) 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*, 8066−8129.
(32) Fernandez, M.; Boyd, P. G.; Daff, T. D.; Aghaji, M. Z.; Woo, T. K. Rapid and Accurate Machine Learning Recognition of High Performing Metal Organic Frameworks for CO2 Capture. *J. Phys. Chem. Lett.* **2014**, *5*, 3056−3060.
(33) Lee, Y.; Barthel, S. D.; Dłotko, P.; Moosavi, S. M.; Hess, K.; Smit, B. Quantifying Similarity of Pore-Geometry in Nanoporous Materials. *Nat. Commun.* **2017**, *8*, 15396.
(34) Moosavi, S. M.; Xu, H.; Chen, L.; Cooper, A. I.; Smit, B. Geometric Landscapes for Material Discovery within Energy−Structure−Function Maps. *Chem. Sci.* **2020**, *11*, 5423−5433.
(35) 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*, 14529−14538.
(36) 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.
(37) 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.
(38) Nandy, A.; Duan, C.; Janet, J. P.; Gugler, S.; Kulik, H. J. Strategies and Software for Machine Learning Accelerated Discovery in Transition Metal Chemistry. *Ind. Eng. Chem. Res.* **2018**, *57*, 13973−13986.
(39) Nandy, A.; Duan, C.; Taylor, M. G.; Liu, F.; Steeves, A. H.; Kulik, H. J. Computational Discovery of Transition-metal Complexes: From High-throughput Screening to Machine Learning. *Chem. Rev.* **2021**, *121*, 9927−10000.
(40) Jablonka, K. M.; Moosavi, S. M.; Asgari, M.; Ireland, C.; Patiny, L.; Smit, B. A Data-Driven Perspective on the Colours of Metal−Organic Frameworks. *Chem. Sci.* **2021**, *12*, 3587−3598.
(41) Ioannidis, E. I.; Gani, T. Z. H.; Kulik, H. J. MolSimplify: A Toolkit for Automating Discovery in Inorganic Chemistry. *J. Comput. Chem.* **2016**, *37*, 2106−2117.
(42) van Dam, A. Diversity and its Decomposition into Variety, Balance and Disparity. *R. Soc. Open Sci.* **2019**, *6*, 190452.
(43) Daly, A.; Baetens, J.; De Baets, B. Ecological Diversity: Measuring the Unmeasurable. *Mathematics* **2018**, *6*, 119.
(44) Purvis, A.; Hector, A. Getting the Measure of Biodiversity. *Nature* **2000**, *405*, 212−219.
(45) Page, S. *On Diversity and Complexity*; Princeton University Press, 2010; pp 16−53.
(46) Stirling, A. A General Framework for Analysing Diversity in Science, Technology and Society. *J. R. Soc.* **2007**, *4*, 707−719.
(47) Pielou, E. C. The Measurement of Diversity in Different Types of Biological Collections. *J. Theor. Biol.* **1966**, *13*, 131−144.
(48) Shapely: Manipulation and Analysis of Geometric Objects. https://github.com/Toblerity/Shapely, using March 4, 2021 release throughout this study.
(49) Buch, V. Path Integral Simulations of Mixed Para-D2 and Ortho-D2 Clusters: The Orientational Effects. *J. Chem. Phys.* **1994**, *100*, 7610−7629.
(50) Potoff, J. J.; Siepmann, J. I. Vapor−Liquid Equilibria of Mixtures containing Alkanes, Carbon Dioxide, and Nitrogen. *AIChE J.* **2001**, *47*, 1676−1682.
(51) Feynman, R. P.; Hibbs, A. R.; Styer, D. F. *Quantum Mechanics and Path Integrals*; Courier Corporation, 2010; pp 279−287.
(52) Jablonka, K. M.; Ongari, D.; Smit, B. Applicability of Tail Corrections in the Molecular Simulations of Porous Materials. *J. Chem. Theory Comput.* **2019**, *15*, 5635−5641.
(53) Lorentz, H. A. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Ann. Phys.* **1881**, *248*, 127−136.
(54) 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*, 81−101.
(55) Hu, Z.; Wang, Y.; Shah, B. B.; Zhao, D. CO2 Capture in Metal−Organic Framework Adsorbents: An Engineering Perspective. *Adv. Sustainable Syst.* **2019**, *3*, 1800080.
(56) Wang, L.; Liu, Z.; Li, P.; Wang, J.; Yu, J. CO2 Capture from Flue Gas by Two Successive VPSA Units using 13XAPG. *Adsorption* **2012**, *18*, 445−459.
(57) Tlili, N.; Grévillot, G.; Vallières, C. Carbon Dioxide Capture and Recovery by Means of TSA and/or VSA. *Int. J. Greenhouse Gas Control* **2009**, *3*, 519−527.
(58) Anderson, R.; Rodgers, J.; Argueta, E.; Biong, A.; Gómez-Gualdrón, D. A. Role of Pore Chemistry and Topology in the CO2 Capture Capabilities of MOFs: from Molecular Simulation to Machine Learning. *Chem. Mater.* **2018**, *30*, 6325−6337.
(59) Van der Maaten, L.; Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2509−2605.
(60) Cavenati, S.; Grande, C. A.; Rodrigues, A. E. Layered Pressure Swing Adsorption for Methane Recovery from CH4/CO2/N2 Streams. *Adsorption* **2005**, *11*, 549−554.
(61) Ho, M. T.; Allinson, G. W.; Wiley, D. E. Reducing the Cost of CO2 Capture from Flue Gases using Pressure Swing Adsorption. *Ind. Eng. Chem. Res.* **2008**, *47*, 4883−4890.
(62) Wang, Y.-X.; Li, Z.-Y.; Zhu, J.; Wang, X.-Q.; Xu, L.-L.; Lachgar, A.; Zhang, J.-J. Four (5, 5)-Connected Three-Dimensional Metal Organic Materials Based on Pentacarboxylate Ligand: Synthesis, structures and characterization. *CrystEngComm* **2013**, *15*, 6395−6402.
(63) Liu, F.; Zhang, L.; Wang, R.; Sun, J.; Yang, J.; Chen, Z.; Wang, X.; Sun, D. Five MOFs with Different Topologies Based on

Anthracene Functionalized Tetracarboxylic Acid: Syntheses, Structures, and Properties. *CrystEngComm* **2014**, *16*, 2917−2928.
(64) Li, X.-J.; Jiang, F.-L.; Wu, M.-Y.; Chen, L.; Qian, J.-J.; Zhou, K.; Yuan, D.-Q.; Hong, M.-C. Construction of Two Microporous Metal−Organic Frameworks with flu and pyr Topologies Based on Zn4 (μ3-OH) 2 (CO2) 6 and Zn6 (μ6-O)(CO2) 6 Secondary Building Units. *Inorg. Chem.* **2014**, *53*, 1032−1038.
(65) Liu, S.-J.; Han, S.-D.; Chang, Z.; Bu, X.-H. Cluster-and Chain-Based Magnetic MOFs Derived from 3d Metal Ions and 1, 3, 5-Benzenetricarboxylate. *New J. Chem.* **2016**, *40*, 2680−2686.
(66) Veenstra, M.; Purewal, J.; Xu, C.; Yang, J.; Blaser, R.; Sudik, A.; Siegel, D.; Ming, Y.; Liu, D.; Chi, H.; Gaab, M.; Arnold, L.; Muller, U. Ford/BASF/UM Activities in Support of the Hydrogen Storage Engineering Center of Excellence. 2015 https://www.osti.gov/biblio/1296578 (accessed August 7, 2021).
(67) Purewal, J.; Liu, D.; Sudik, A.; Veenstra, M.; Yang, J.; Maurer, S.; Müller, U.; Siegel, D. J. Improved Hydrogen Storage and Thermal Conductivity in High-Density MOF-5 Composites. *J. Phys. Chem. C* **2012**, *116*, 20199−20212.
(68) Mason, J. A.; Veenstra, M.; Long, J. R. Evaluating Metal−Organic Frameworks for Natural Gas Storage. *Chem. Sci.* **2014**, *5*, 32−51.
(69) DOE Technical Targets for Onboard Hydrogen Storage for Light-Duty Vehicles. https://www.energy.gov/eere/fuelcells/doe-technical-targets-onboard-hydrogen-storage-light-duty-vehicles, 2017 (accessed August 20, 2021).
(70) Dzubak, A. L.; Lin, L.-C.; Kim, J.; Swisher, J. A.; Poloni, R.; Maximoff, S. N.; Smit, B.; Gagliardi, L. Ab Initio Carbon Capture in Open-Site Metal−Organic Frameworks. *Nat. Chem.* **2012**, *4*, 810−816.
(71) Peng, X.; Lin, L.-C.; Sun, W.; Smit, B. Water Adsorption in Metal−Organic Frameworks with Open-Metal Sites. *AIChE J.* **2015**, *61*, 677−687.
(72) Basdogan, Y.; Keskin, S. Simulation and Modelling of MOFs for Hydrogen Storage. *CrystEngComm* **2015**, *17*, 261−275.
(73) Jablonka, K. M.; Jothiappan, G. M.; Wang, S.; Smit, B.; Yoo, B. Bias Free Multiobjective Active Learning for Materials Design and Discovery. *Nat. Commun.* **2021**, *12*, 2312.
(74) Deshwal, A.; Simon, C. M.; Doppa, J. R. Bayesian Optimization of Nanoporous Materials. *Mol. Syst. Des. Eng.* **2021**, *6*, 1066−1086.
(75) Yao, Z.; Sánchez-Lengeling, B.; Bobbitt, N. S.; Bucior, B. J.; Kumar, S. G. H.; Collins, S. P.; Burns, T.; Woo, T. K.; Farha, O. K.; Snurr, R. Q.; Aspuru-Guzik, A. Inverse design of nanoporous crystalline reticular materials with deep generative models. *Nat. Mach. Intell.* **2021**, *3*, 76−86.
(76) Kim, B.; Lee, S.; Kim, J. Inverse Design of Porous Materials using Artificial Neural Networks. *Sci. Adv.* **2020**, *6*, No. eaax9324.
(77) Talirz, L.; Kumbhar, S.; Passaro, E.; Yakutovich, A. V.; Granata, V.; Gargiulo, F.; Borelli, M.; Uhrin, M.; Huber, S. P.; Zoupanos, S.; Adorf, C. S.; Andersen, C. W.; Schütt, O.; Pignedoli, C. A.; Passerone, D.; VandeVondele, J.; Schulthess, T. C.; Smit, B.; Pizzi, G.; Marzari, N. Materials Cloud, a Platform for Open Computational Science. *Sci. Data* **2020**, *7*, 299.

https://doi.org/10.1021/acsami.1c16220
ACS Appl. Mater. Interfaces 2021, 13, 61004−61014