Consistent textural properties for MOFX-DB

1 Introduction

On each MOF webpage, there are several textural properties available, including void fraction (VF), gravimetric surface area (GSA), volumetric surface area (VSA), pore limiting diameter (PLD) and largest cavity diameter (LCD). However, these textural properties were collected from different publications and thus were calculated with inconsistent parameters, procedures, and programs. We refer to these published data as "original set." Here we report a set of recalculated textural properties for all MOFs in the MOFX-DB with a consistent calculation procedure, and we refer to the new data set as "consistent set." We note that there are textural properties missing for some MOFs in the original set, but textural properties are available for all materials in the consistent set.

Both the original set and consistent set can be downloaded from the "Databases" tab in MOFX-DB (https://mof.tech.northwestern.edu/databases), under the name "Download Textural Properties." All data are reported in CSV format. The meaning of each column in the CSV file is explained in Table 1.

Table 1: *Textural properties available in the consistent set and their corresponding columns in CSV file*

Column name	Full name	Definition	Units
id	ID name of the MOF	e.g., for 'hMOF-12.cif' file, '12' is the ID name.	
vf	Helium void fraction (VF)	Volumetric fraction of void space measured by He adsorption	fraction
sa_tot_m2g	Total gravimetric surface area (total GSA)	Gravimetric surface area for both accessible and inaccessible pores	m^2/g
sa_acc_m2g	Accessible gravimetric surface area (accessible GSA)	Gravimetric surface area for accessible pores only	m^2/g
sa_tot_m2cm3	Total volumetric surface area (total VSA)	Volumetric surface area for both accessible and inaccessible pores	m^2/cm^3
sa_acc_m2cm3	Accessible volumetric surface area (accessible VSA)	Volumetric surface area for accessible pores only	m^2/cm^3

pld	Pore limiting diameter (PLD)	Largest probe that can cross the simulation cell in at least one dimension via a diffusive pathway	Å
lcd	Largest cavity diameter (LCD)	Largest cavity size	Å
lcd_free	Largest cavity diameter in the accessible channel (accessible LCD)	Largest cavity size in the accessible channel	Å

2 Calculation details for the consistent set

The surface area, PLD and LCD were calculated by Zeo++ v0.3.[1] The command was

network -ha -r UFF.rad -res -sa 1.86 1.86 1000 <cif_name>

where Universal Force Field (UFF)[2] framework atom radii were used ("-r UFF.rad") with high accuracy ("-ha"). The "UFF.rad" file is available in the "radii" folder. A nitrogen probe with radius of 1.86 Å (corresponding to the size of a nitrogen molecule in the TraPPE force field)[3] was used to assess both the availability of the network and the surface area. A total of 1,000 points were sampled at a distance of 1.86 Å from each framework atom surface, and Monte Carlo integration was used to determine the result. Using 1,000 sampling points roughly gives the accuracy up to the first decimal point. Because this sampling is performed for each framework atom, considering the computational time, 1,000 points is a good choice. In very few cases, we removed "-ha" flag to avoid the failure of the Voronoi volume check.

The helium void fraction, $V_{pore,He}/V$, was measured by Widom insertion method using RASPA2[4–6],

$$\frac{V_{pore,He}}{V} = < exp^{-U(\mathbf{r})/k_B T} > \tag{1}$$

where $V_{pore,He}$ is the pore volume and V is the volume of the entire system. The property in the angular brackets is the average Boltzmann factor, which is estimated by placing a probe helium atom in random locations $\mathbf{r}$ throughout the system and estimating its interactions potential $U(\mathbf{r})$ with the material. Detailed parameters are listed in Table 2. The simulation box size was chosen to be the minimum distance between two opposing cell surface that is larger than twice of the cutoff radius, thus ensuring minimum image convention. We note that there are a few MOF

structures in the CoRE MOF database for which the void fraction calculations failed, possibly due to an unphysical structure.

For the textural property calculation of zeolites, we used a specific force field (LJ size and energy parameters) designed for zeolites,[7,8] and all other calculation parameters remain the same.

Table 2: *Simulation parameters for helium void fraction calculation for MOFs.*

Parameters	Values
He atom probe Lennard-Jones σ [Å][9,10]	2.64
He atom probe Lennard-Jones ε/k_B [K][9,10]	10.9
Temperature [K]	298
vdW cutoff radius [Å]	12.8
Framework force field	Universal Force Field
Combining rules	Lorentz-Berthelot
Number of RASPA cycles [a]	10,000

[a] Total number of Widom insertion moves is 200,000 (10,000 * 20). Each Monte Carlo cycle in RASPA includes a minimum 20 steps.

3 Comparison of consistent set with original set

hMOFs. Parity plots comparing the consistent set and original set of textural properties for hMOFs are shown in Figure 1. Definitions of textural properties are available in Table 1. The difference in VF is probably due to using different simulation and force field parameters.[11] For LCD and PLD, differences are probably due to using different algorithms.[12] For LCD (Figure 1c), the data are capped at 25 Å in the original set (It levels off and there is a flat line). We have left it as is to preserve the original published data. We note that the surface area in the original set (for hMOFs and TobaCCo MOFs) is the *total* surface area, thus including the surface area for both accessible and inaccessible pores/channels. For both total VSA and GSA, original values are higher compared to the values in the consistent set overall. There are two possible reasons for this discrepancy: 1) Original data were calculated by RASPA, while consistent set was calculated by Zeo++. 2) A smaller N_2 probe (size of 3.68 Å) was used in the original calculations,[11,13] while current calculations used a slightly larger probe (3.72 Å).
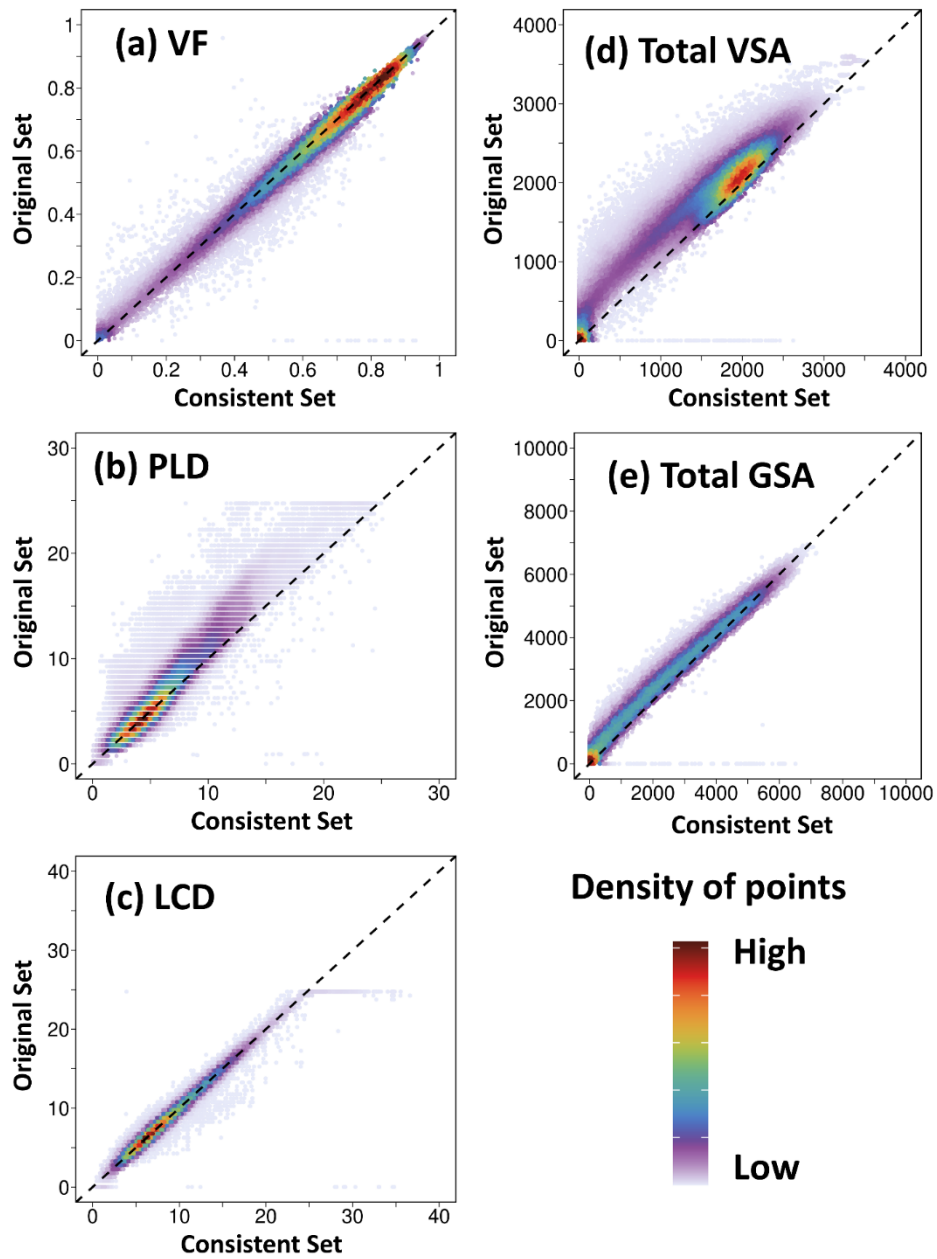
Figure 1: *Parity plots comparing consistent set and the original set of textural properties for hMOFs.*

ToBaCCo MOFs. Parity plots comparing the consistent set with the original set for the ToBaCCo MOFs[14] are shown in Figure 2. In general, the original set is in good agreement with the consistent set, reflecting the similar calculation procedures, parameters, and algorithms used in the two cases.
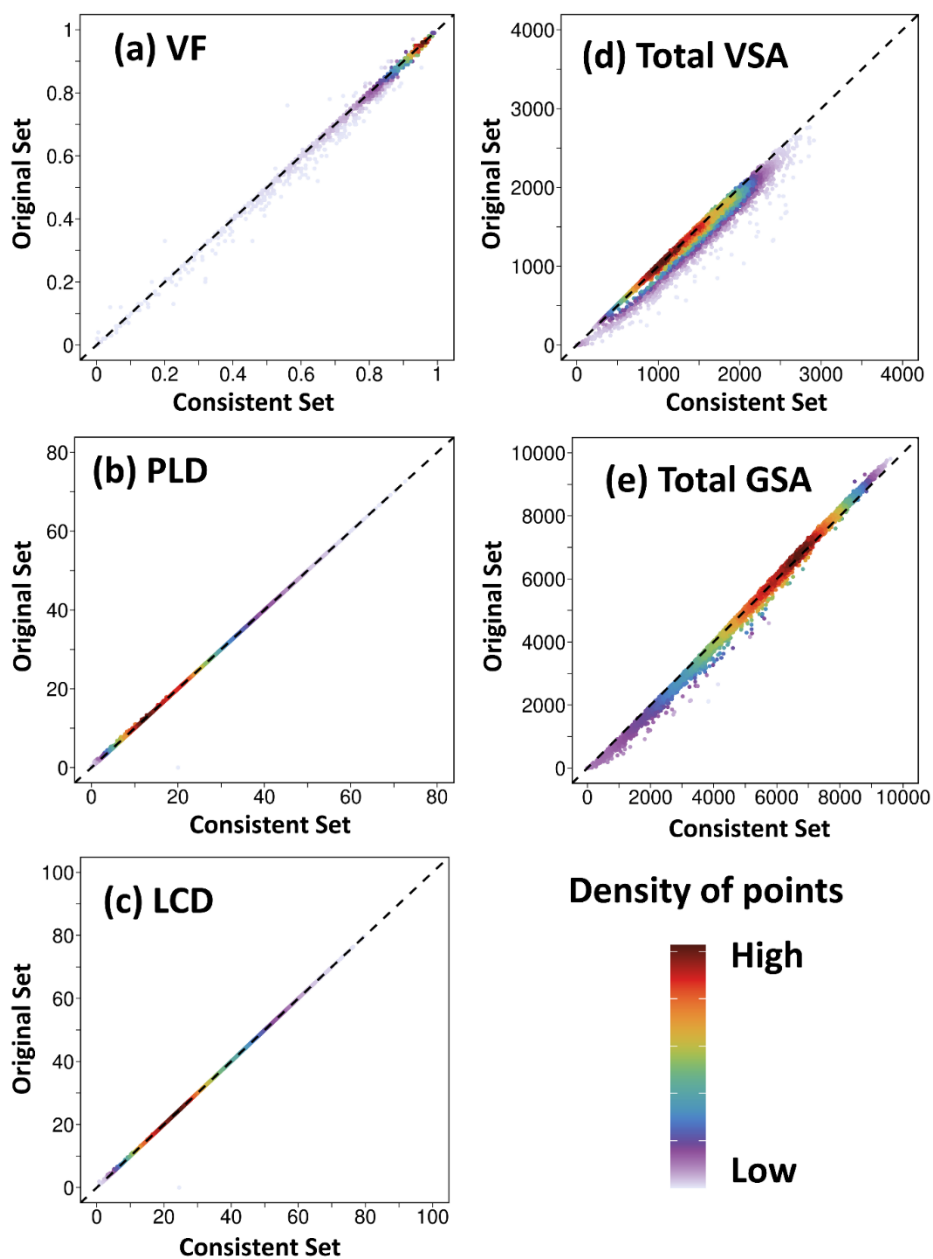

Figure 2: *Parity plots comparing consistent set and the original set of textural properties for ToBaCCo MOFs.*

CoRE MOF 2014. Parity plots comparing the consistent set with the original set for CoRE MOF 2014[15] are shown in Figure 3. We note that data from the original set were also calculated by Zeo++ software but with CCDC radii used for framework atoms (new calculations use the atomic radii from UFF). In addition, according to the original CoRE MOF 2014 paper, surface area reported in the original set is the accessible surface area which only includes that of the accessible pores through windows/channels. This explains the horizontal line of points at the bottom of the parity plot (see Figure 3, Total VSA, Total GSA); these points correspond to zero in the original set (*i.e.*, structures determined as inaccessible to the probe) but non-zero in the consistent set (*i.e.*, total surface area, values also include the surface area from inaccessible pores). When we compare accessible surface area from both sets (see Figure 3, Accessible VSA), there are both vertical and horizontal line of points shown at the left bottom corner. This can be attributed to the use of different framework atom radii in both sets, and the accessibility of pores was therefore determined inconsistently, even though same probe size (3.72 Å) was implemented in both sets.
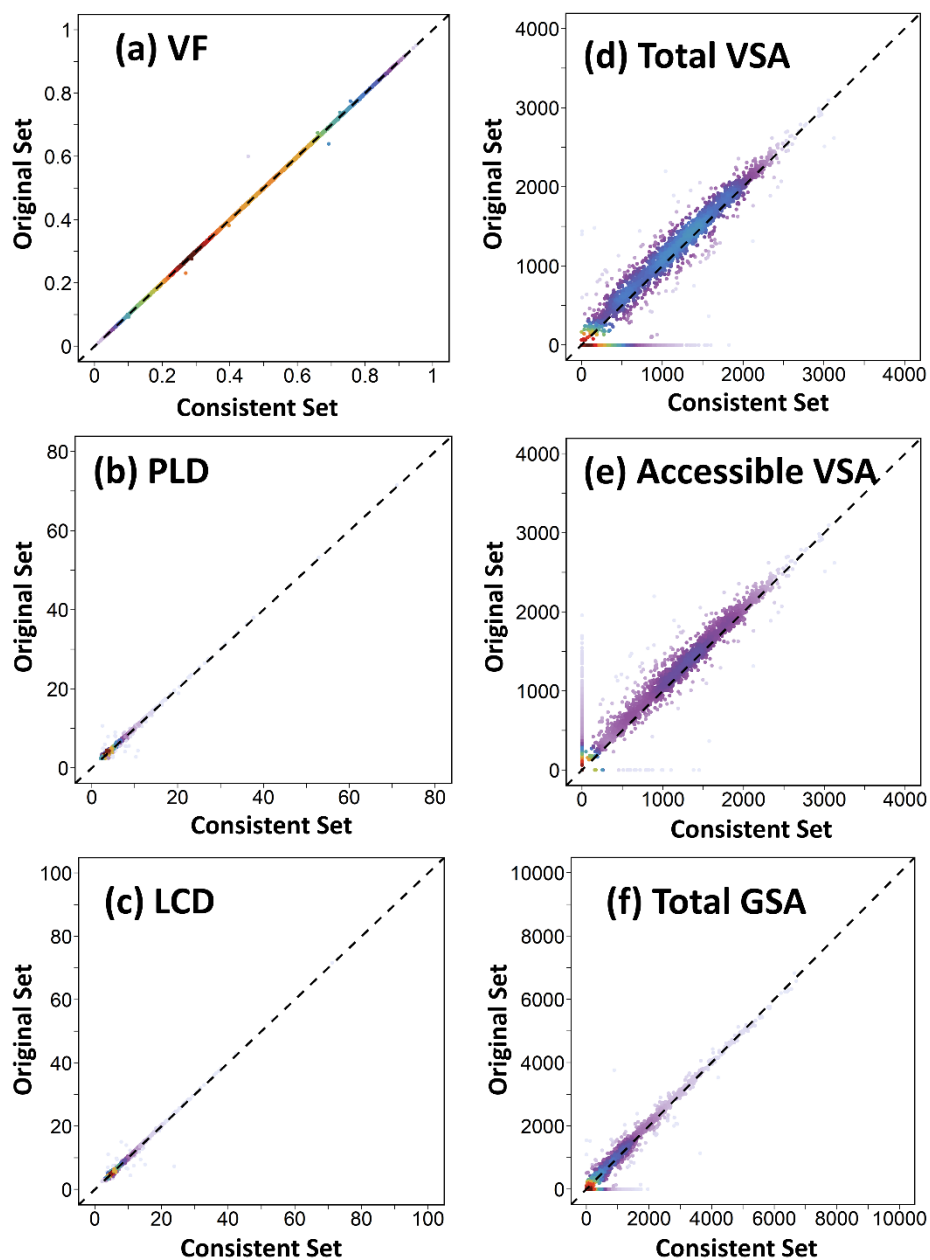
Figure 3: *Parity plots comparing consistent set and the original set of textural properties for CoRE MOFs 2014. Color gradient indicates the density of points (purple to red is low to high).*

CoRE MOF 2019. Parity plots comparing the consistent set with the original set for CoRE MOFs 2019[16] are shown in Figure 4. A discrepancy is observed for the void fraction (VF) between the consistent set and the original set. The reason is that the original set was calculated in a purely geometric way using Zeo++ software (*i.e.*, geometric VF), while the consistent set was calculated using Widom insertion method (*i.e.*, helium VF). We note that there are some CoRE MOF structures with calculated helium VF larger than 1 (not shown). In principle, VF should range from 0 to 1 as indicated by Eq. 1. In these cases, the high VF is due to some unphysical structures in CoRE MOF 2019 database where solid atoms overlap with each other forming a high-density structure. This might be due to disordered rings in the structures which are resolved from X-ray diffraction or other issues with the structure. This leads to unrealistically strong attractive interactions between the probe helium atom and the framework and results in a higher estimated VF than is realistic. We found there are 63 structures in the CoRE MOF 2019 database with VF larger than 1.

The original CoRE MOF 2019 paper provides surface area values for both accessible and inaccessible pores, we took the sum of them when uploading data to the MOFDB. Therefore, the surface area in the original set is the *total* surface area. We observe that, in general, surface area in the original set is higher than that in the consistent set. This is because a probe with a smaller diameter (3.31 Å) was used in the original set to evaluate the surface area, compared to the probe with a slightly larger diameter (3.72 Å) in the consistent set.
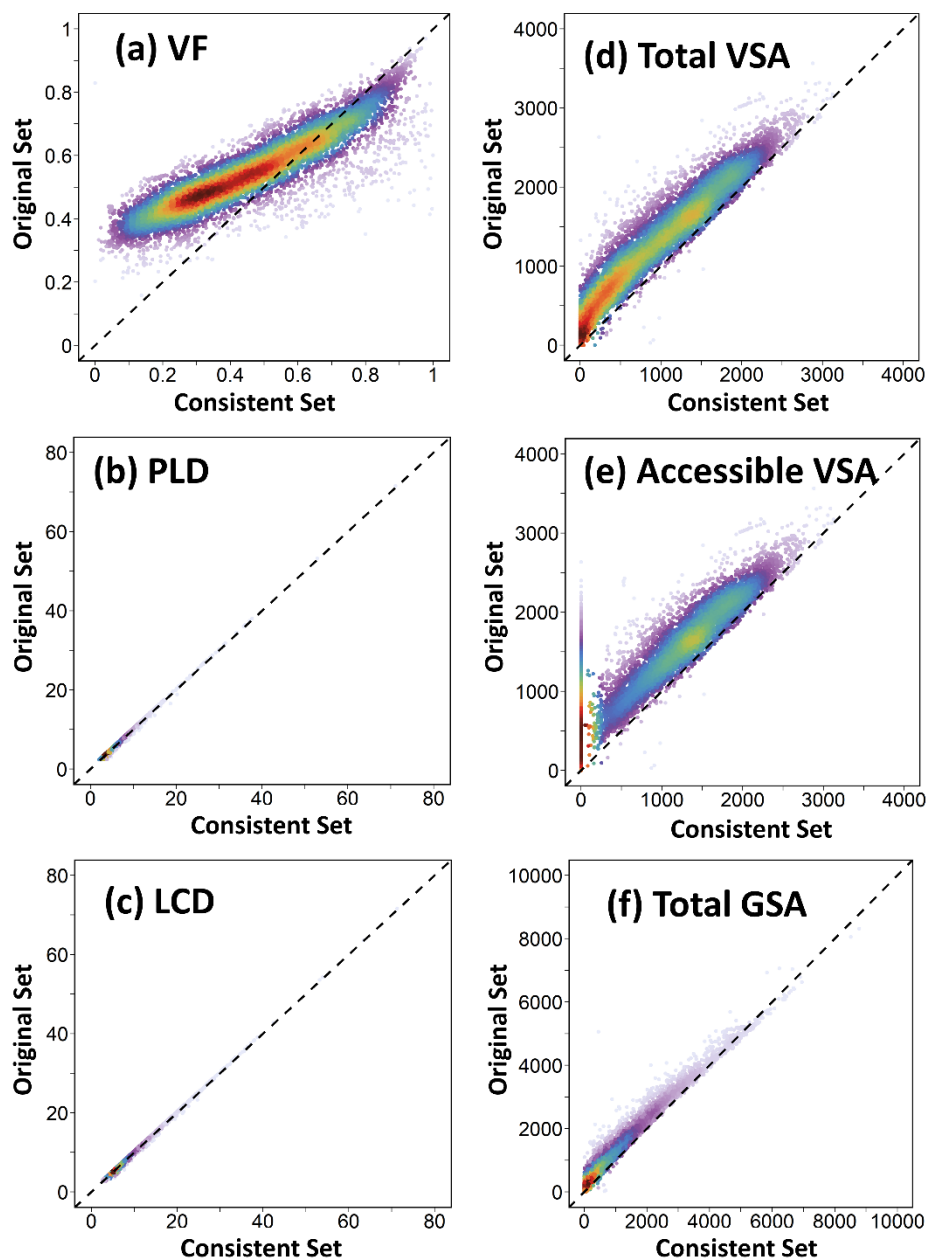
Figure 4: *Parity plots comparing consistent set and the original set of textural properties for CoRE MOFs 2019. Color gradient indicates the density of points (purple to red is low to high).*

References

(1)   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* (1), 134–141. https://doi.org/10.1016/j.micromeso.2011.08.020.

(2)   Rappe, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard, W. A.; Skiff, W. M. UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **1992**, *114* (25), 10024–10035. https://doi.org/10.1021/ja00051a040.

(3)   Bae, Y.-S.; Yazaydın, A. Ö.; Snurr, R. Q. Evaluation of the BET Method for Determining Surface Areas of MOFs and Zeolites That Contain Ultra-Micropores. *Langmuir* **2010**, *26* (8), 5475–5483. https://doi.org/10.1021/la100449z.

(4)   Talu, O.; Myers, A. L. Reference Potentials for Adsorption of Helium, Argon, Methane, and Krypton in High-Silica Zeolites. *Colloids Surfaces A Physicochem. Eng. Asp.* **2001**, *187–188*, 83–93. https://doi.org/10.1016/S0927-7757(01)00628-8.

(5)   Dubbeldam, D.; Calero, S.; Ellis, D. E.; Snurr, R. Q. RASPA: Molecular Simulation Software for Adsorption and Diffusion in Flexible Nanoporous Materials. *Mol. Simul.* **2016**, *42* (2), 81–101. https://doi.org/10.1080/08927022.2015.1010082.

(6)   Sarkisov, L.; Bueno-Perez, R.; Sutharson, M.; Fairen-Jimenez, D. Materials Informatics with PoreBlazer v4.0 and the CSD MOF Database. *Chem. Mater.* **2020**, *32* (23), 9849–9867. https://doi.org/10.1021/acs.chemmater.0c03575.

(7)   Bai, P.; Tsapatsis, M.; Siepmann, J. I. TraPPE-Zeo: Transferable Potentials for Phase Equilibria Force Field for All-Silica Zeolites. *J. Phys. Chem. C* **2013**, *117* (46), 24375–24387. https://doi.org/10.1021/jp4074224.

(8)   Sun, Y.; DeJaco, R. F.; Li, Z.; Tang, D.; Glante, S.; Sholl, D. S.; Colina, C. M.; Snurr, R. Q.; Thommes, M.; Hartmann, M.; Siepmann, J. I. Fingerprinting Diverse Nanoporous Materials for Optimal Hydrogen Storage Conditions Using Meta-Learning. *Sci. Adv.* **2021**, *7* (30), eabg3983. https://doi.org/10.1126/sciadv.abg3983.

(9)   Hirschfelder, J. O.; Curtiss, C. F.; Bird, R. B. *Molecular Theory of Gases and Liquids*; Wiley: New York, 1964. https://doi.org/10.1063/1.3061949.

(10)   Ongari, D.; Boyd, P. G.; Barthel, S.; Witman, M.; Haranczyk, M.; Smit, B. Accurate

Characterization of the Pore Volume in Microporous Crystalline Materials. *Langmuir* **2017**, *33* (51), 14529–14538. https://doi.org/10.1021/acs.langmuir.7b01682.

(11)   Bobbitt, N. S.; Chen, J.; Snurr, R. Q. High-Throughput Screening of Metal–Organic Frameworks for Hydrogen Storage at Cryogenic Temperature. *J. Phys. Chem. C* **2016**, *120* (48), 27328–27341. https://doi.org/10.1021/acs.jpcc.6b08729.

(12)   Sikora, B. J.; Wilmer, C. E.; Greenfield, M. L.; Snurr, R. Q. Thermodynamic Analysis of Xe/Kr Selectivity in over 137 000 Hypothetical Metal–Organic Frameworks. *Chem. Sci.* **2012**, *3* (7), 2217. https://doi.org/10.1039/c2sc01097f.

(13)   Walton, K. S.; Snurr, R. Q. Applicability of the BET Method for Determining Surface Areas of Microporous Metal−Organic Frameworks. *J. Am. Chem. Soc.* **2007**, *129* (27), 8552–8556. https://doi.org/10.1021/ja071174k.

(14)   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* (11), 5801–5810. https://doi.org/10.1021/acs.cgd.7b00848.

(15)   Chung, Y. G.; Camp, J.; Haranczyk, M.; Sikora, B. J.; Bury, W.; Krungleviciute, V.; Yildirim, T.; Farha, O. K.; Sholl, D. S.; Snurr, R. Q. Computation-Ready, Experimental Metal–Organic Frameworks: A Tool To Enable High-Throughput Screening of Nanoporous Crystals. *Chem. Mater.* **2014**, *26* (21), 6185–6192. https://doi.org/10.1021/cm502594j.

(16)   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* (12), 5985–5998. https://doi.org/10.1021/acs.jced.9b00835.