Textural Properties of a Large Collection of Computationally Constructed MOFs and Zeolites

Benjamin J. Sikora[a], Randy Winnegar[a], Davide M. Proserpio[b,c], Randall Q. Snurr[a]

[a] *Department of Chemical and Biological Engineering*
Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA
[b] *Università degli Studi di Milano, Dipartimento di Chimica*
Via Golgi, 19 - 20133 Milano, Italy
[c] *Samara Center for Theoretical Materials Science (SCTMS) Samara State University,*
Ac. Pavlov St. 1, Samara 443011, Russia

1. Determination of Conformers

Using the connectivity and composition data of two structures, various methods (Figure S1) can be combined to determine whether they are conformers. The first step involves comparing the compositions of the structures, where structures that do not have equivalent compositions can be ruled out as conformers of one another. The second step uses a number labeling scheme (hash number) based on the structure's connectivity data, where the number generated is likely to be unique to that structure. For example, ethane's atoms are labeled in sequential order in a non-specific way (Figure S2a). The adjacency matrix (Figure S2b), which records which atoms are connected to each other, is obtained from either given or computed connectivity data. The hash matrix (Figure S2c) is then calculated for each atom by first determining the shortest path from a given atom to another atom and recording the value. In the case of ethane, the shortest path from carbon atom 1 to carbon atom 2 is one. The shortest path from hydrogen atom 7 to carbon atom 2 is two, and so on. The overall hash number is calculated by simply summing the values of the hash matrix, where in our example case, ethane has a hash number of 116. The last step, where two structures have equivalent compositions and identical hash numbers, compares "composition connections", where the composition connection is simply a recorded value of the shortest path from all atoms of one atom type to all atoms of another atom type. For example, ethane has 3 composition connections, C-H, C-C, and H-H. The C-C composition connection value is calculated by summing every element in the hash matrix that has a path from a carbon atom to another carbon atom (elements: (1)(1), (1)(2), (2)(1), and (2)(2)). Note that the more diverse in composition the original structure is, the more likely its corresponding hash number and composition connections will be unique only to that structure, resulting in fewer false positives.

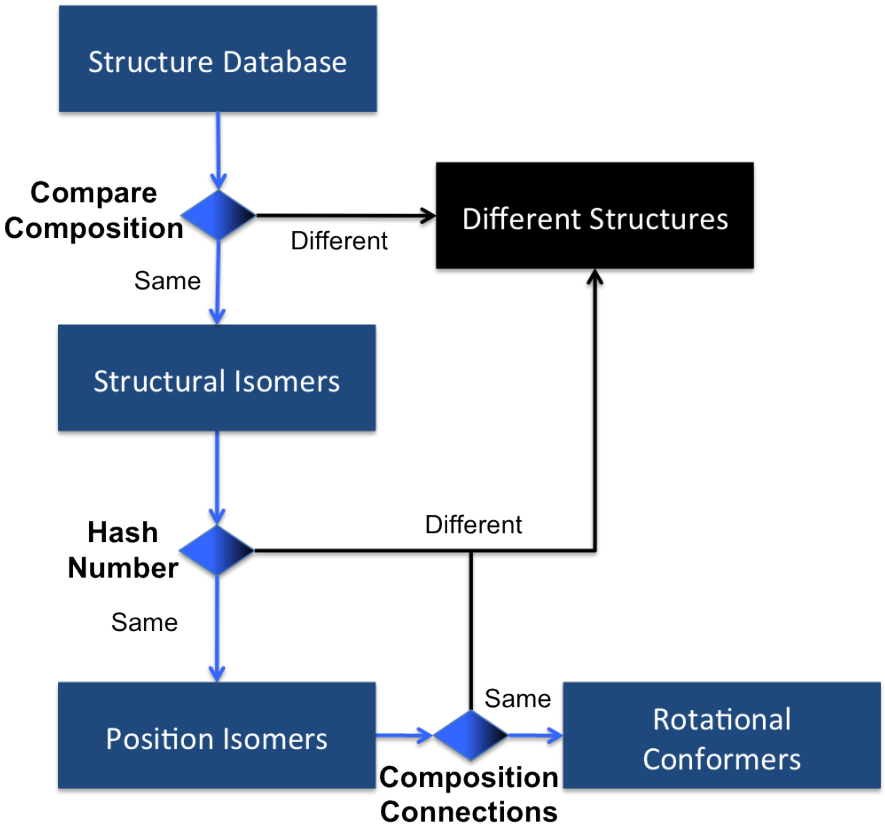
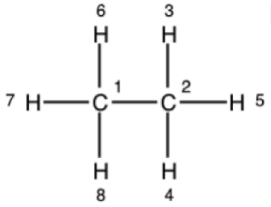


Figure S1: Conformational isomer identification routine. This routine takes two structures and attempts to identify whether they are rotational isomers of one another. This is done by comparing the compositions, hash numbers, and composition connections of the structures.

a



b

	1	2	3	4	5	6	7	8
1	0	1	0	0	0	1	1	1
2	1	0	1	1	1	0	0	0
3	0	1	0	0	0	0	0	0
4	0	1	0	0	0	0	0	0
5	0	1	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0
7	1	0	0	0	0	0	0	0
8	1	0	0	0	0	0	0	0

c

	1	2	3	4	5	6	7	8
1	0	1	2	2	2	1	1	1
2	1	0	1	1	1	2	2	2
3	2	1	0	2	2	3	3	3
4	2	1	2	0	2	3	3	3
5	2	1	2	2	0	3	3	3
6	1	2	3	3	3	0	2	2
7	1	2	3	3	3	2	0	2
8	1	2	3	3	3	2	2	0

Figure S2: Labeling scheme **(a)** and the adjacency **(b)** and hash matrices **(c)** for example structure ethane. The hash number is determined by the sum of the values of the hash matrix, which in ethane's case is 116. The composition connection values, determined by recording the shortest path from all atoms of one atom type to all atoms of another atom type, are: C-C = 2, C-H = 36, H-H = 78. The sum of the composition connections is simply the hash number.

2. Additional Textural Properties of Hypothetical Metal-Organic Frameworks

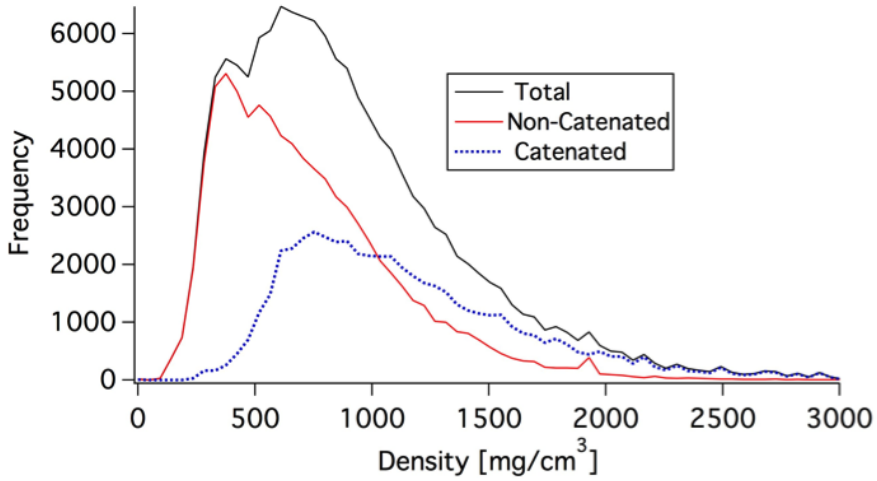


Figure S3: The distribution of framework density for all catenated and non-catenated MOFs in the database of 137,953 hypothetical MOFs. The most dense structures tend to be catenated structures.

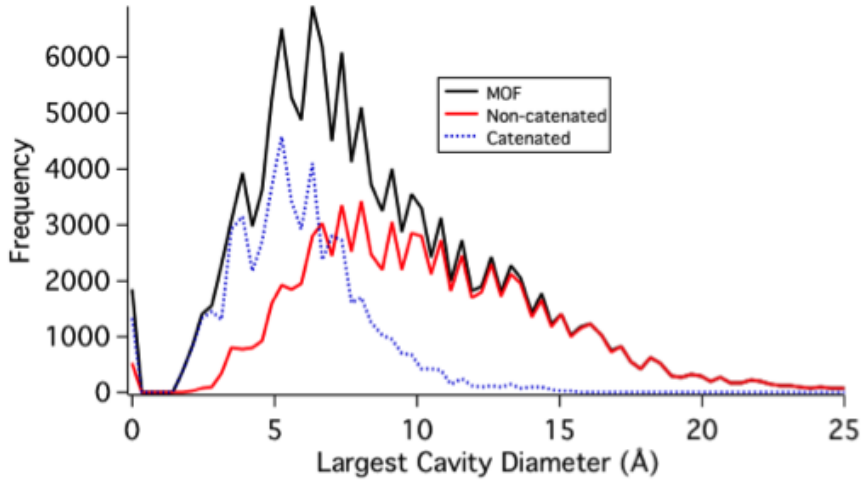


Figure S4: The distribution of the largest cavity diameter (LCD) for all catenated and non-catenated MOFs in the database of 137,953 hypothetical MOFs. Catenated structures tend to have smaller LCDs.

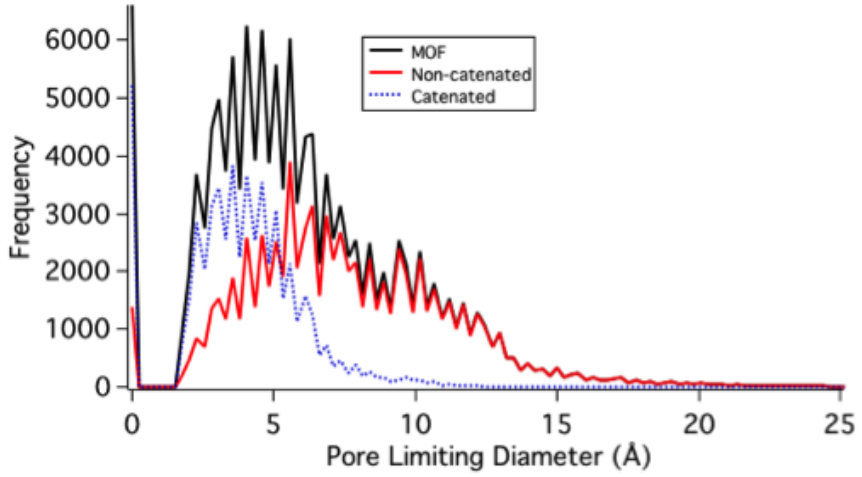


Figure S5: The distribution of the pore limiting diameter (PLD) for all catenated and non-catenated MOFs in the database of 137,953 hypothetical MOFs. Catenated structures tend to have smaller PLDs.

Net	Single # (%)	2-fold	3-fold	4-fold
6-c pcu	73,920 (58.7)	34,601 (27.5)	12,617 (10)	4,843 (3.8)
4-c sra	6,814 (81.7)	1,076 (12.9)	262 (3.1)	186 (2.2)
4-c dia	1,107 (51.8)	957 (44.7)	75 (3.5)	0
3,4-c tbo	1,007 (100)	0	0	0
4-c nbo	404 (100)	0	0	0
12-c fcu	84 (100)	0	0	0
Total	83,336 (60.4)	36,634 (26.6)	12,954 (9.4)	5,029 (3.6)

Table S1: The number of structures for each net type and degree of interpenetration in the database of 137,953 hypothetical MOFs.

3. Energy Minimizaton

In order to better understand what types of atoms (metal corner, linker, or functional group) were moving during the energy minimizations, we looked at the percentage of atoms that moved a given distance separately for the interpenetrated MOFs, the non-interpenetrated functionalized MOFs and the non-interpenetrating unfunctionalized (unmodified) MOFs. In Figure S6, we see for the majority of the cases interpenetrated and non-interpenetrated functionalized MOFs have a higher probability of their atoms moving a significant amount compared to the unmodified MOFs. This indicates that during minimization it is the interpenetrated MOFs and the functionalized portions of non-interpenetrated MOFs that rearrange into lower energy positions.

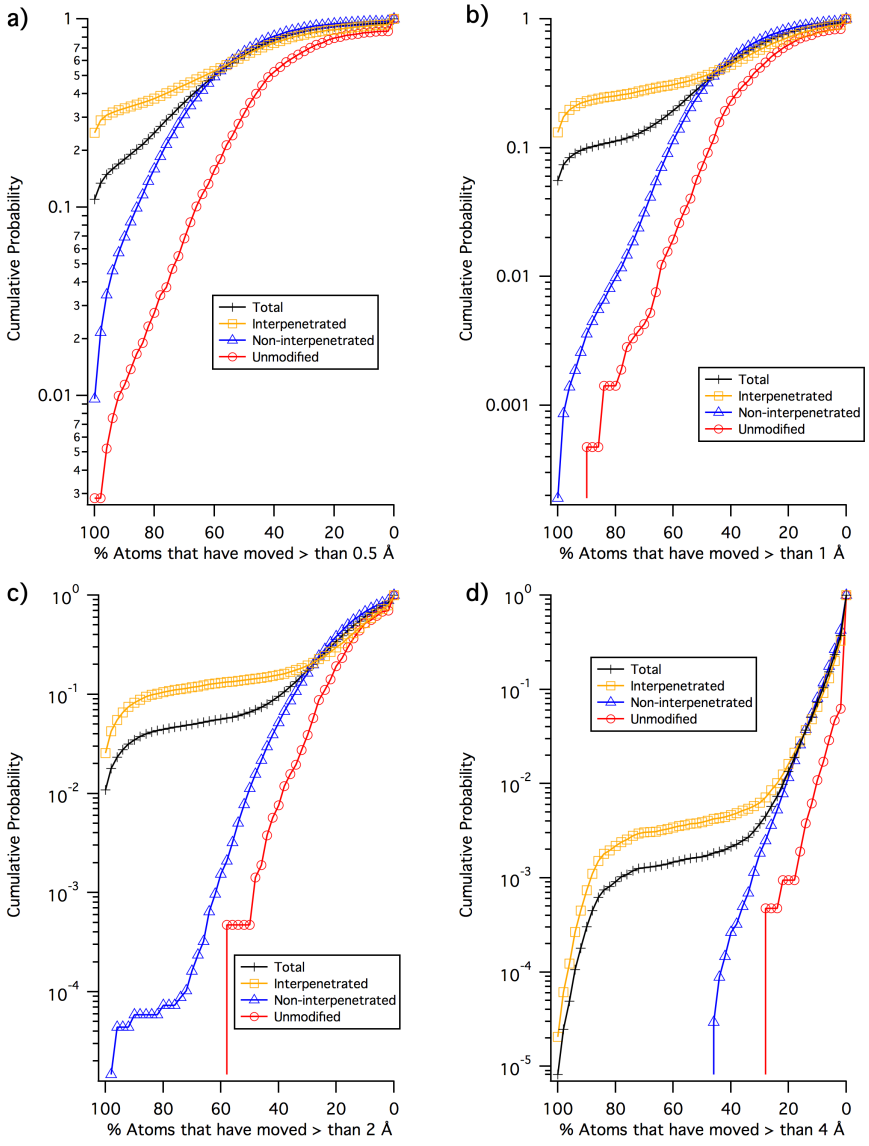


Figure S6: The cumulative normalized frequency of all, interpenetrated, non-interpenetrated, and non-interpenetrated unfunctionalized (unmodified) hypothetical MOFs as a function of the percent of atoms that have moved greater than **(a)** 0.5 Å, **(b)** 1 Å, **(c)** 2 Å, and **(d)** 4 Å.

A comparison of the geometric values for the minimized vs non-minimized structures is shown in Figure S7, where we notice a large variation in the values for the entire database of hypothetical MOFs and a smaller variation in the unmodified hypothetical MOFs. When comparing the database distribution as a whole, we find very little change. An example comparison of the overall distribution can be seen in Figure S8, where we notice an insignificant difference in the distributions for the volumetric surface areas before and after minimization.

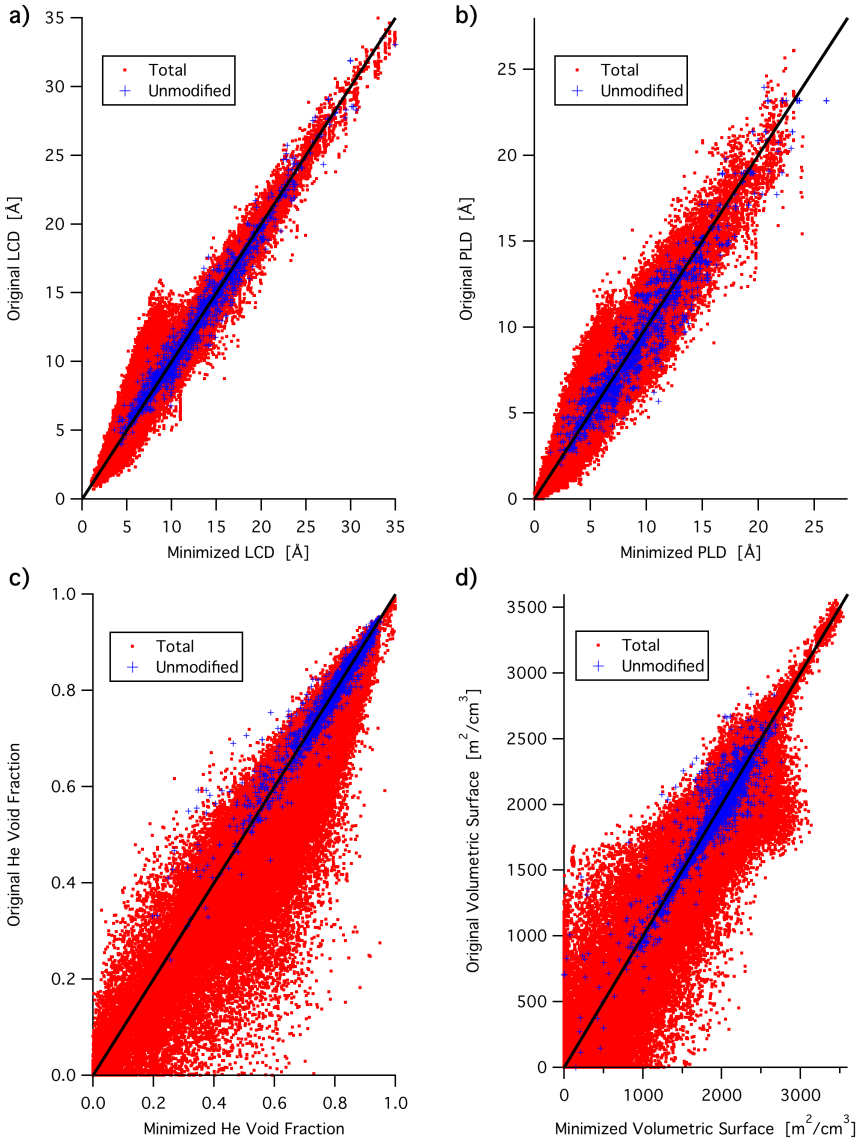


Figure S7: Parity plots of the LCD, PLD, He void, and volumetric surface area for the original structures and minimized structures for the entire database of hypothetical MOFs and for the unmodified hypothetical MOFs.

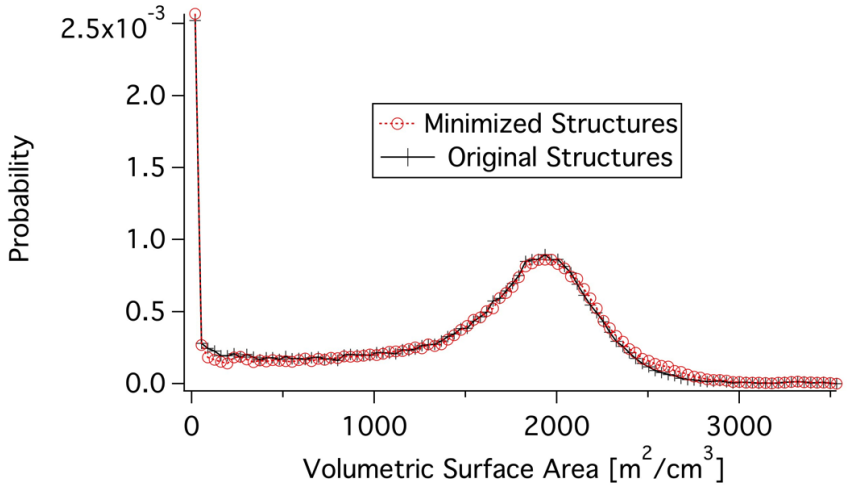


Figure S8: The normalized volumetric surface area distribution for both the original hypothetical MOFs and the geometrically optimized hypothetical MOFs.