Computation-Ready, Experimental (CoRE) Metal-Organic Frameworks: A Tool to Enable High-Throughput Screening of Nanoporous Crystals

Yongchul G. Chung,[1] Jeffrey Camp,[2] Maciej Haranczyk,[3] Benjamin J. Sikora,[1] Wojciech Bury,[4, 5] Vaiva Krungleviciute,[6, 7] Taner Yildirim,[6, 7] Omar K. Farha,[4, 8] David S. Sholl,[2*] and Randall Q. Snurr[1*]

[1]Department of Chemical and Biological Engineering, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

[2]School of Chemical & Biomolecular Engineering, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, GA 30332, USA

[3]Computational Research Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 50F-1650, Berkeley, CA 94720-8139, USA

[4]Department of Chemistry, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA

[5]Department of Chemistry, Warsaw University of Technology, Noakowskiego 3, 00-664 Warsaw, Poland

[6]NIST Center for Neutron Research, National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA

[7]Department of Materials Science and Engineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

[8]Department of Chemistry, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia

Abstract

Experimentally refined crystal structures for metal-organic frameworks (MOFs) often include solvent molecules and partially occupied or disordered atoms. This creates a major impediment to applying high-throughput computational screening to MOFs. To address this problem, we have constructed a database of MOF structures that are derived from experimental data but are immediately suitable for molecular simulations. The Computation-Ready, Experimental (CoRE) MOF database contains over 4,700 porous structures with publically available atomic coordinates. Important physical and chemical properties including the surface area and pore dimensions are reported for these structures. To demonstrate the utility of the database, we performed grand canonical Monte Carlo simulations of methane adsorption on all structures in the CoRE MOF database. We investigated the structural properties of the CoRE MOFs that govern methane storage capacity and found that these relationships agree well with those derived recently from a large database of hypothetical MOFs.

Introduction

Metal-organic frameworks (MOFs) are a class of nanoporous crystalline materials synthesized by bonding metal nodes to multitopic organic linkers. In principle, the chemical building blocks can be rationally selected to tailor MOFs for applications ranging from separations[1] to chemical sensing.[2] In practice, it is difficult to predict the complex relationship among the building blocks, the resulting framework structures, and the emergent physical properties prior to synthesis.[3] Before attempting *de novo* design of a new structure for a particular application, it is wise to consider whether any existing MOFs might be suitable.[4] However, synthesis, characterization, and experimental testing of thousands of MOFs to find the best material for a given application would be prohibitively time consuming and expensive.

High-throughput computational screening can guide experimental efforts by identifying top candidate structures for applications of interest.[5] For example, experimental MOF crystal structures from the Cambridge Structural Database (CSD)[6] have been computationally screened to identify top performing MOFs for applications in light gas,[7] noble gas,[8] and CO_2/N_2[9] separations. These large-scale computational screening efforts require a computation-ready database with the crystallographic information for each structure. However, experimentally refined crystal structures reported in the CSD often include solvent molecules and partially occupied or disordered atoms. Such features are crystallographically meaningful but must be removed prior to computer simulations of fully activated, solvent-free structures. To date, the lack of a publicly available database of computation-ready crystal structures has been a major impediment to applying high-throughput computations to MOFs, because each research group interested in such an approach must first construct an appropriate database of materials.

Recently, Goldsmith et al.[10] have developed an in-house database of computation-ready MOF structures derived from the CSD. In their database, solvents have been removed, mislabeled structures have been identified, and some notable MOF structures (e.g., PCN-610 and NU-100) have been manually repaired. These structures were used to predict the theoretical upper bound for hydrogen storage in MOFs. In their database, structures with interpenetrated frameworks and charge-balancing ions were not included. Although the desolvation algorithm they described can be applied automatically, the construction of such a database still requires a significant amount of work. A publicly available database of computation-ready MOF structures would be an invaluable tool for researchers interested in metal-organic frameworks.

The central result of this publication is the availability of a nearly comprehensive set of porous MOF structures that are derived directly from experimental data but are immediately suitable for molecular simulations or visualization. Our Computation-Ready, Experimental (CoRE) MOFs have several characteristics that will make them broadly useful. First, each structure in the database was desolvated by the efficient graph-labeling algorithm described by Goldsmith et al.[10] Second, we removed additional solvent molecules that are bound to unsaturated metal atoms. Third, charge-balancing ions were retained where necessary, so each CoRE MOF structure is charge neutral overall. Fourth, interpenetrated structures were also retained based on a graph-labeling algorithm. Finally, important geometric properties including the helium void fraction, pore limiting diameter, and largest cavity diameter are reported alongside each computation-ready structure.

Not all MOFs that have been reported experimentally are included in the CoRE MOF database. In some instances, incomplete data or a high degree of disorder makes it challenging to generate a plausible computation-ready crystal structure in an automated procedure. Despite not including materials of this kind, the first version of the CoRE MOF database includes over 5,000 porous

structures with a pore-limiting diameter of at least 2.4 Å. Of these, 4,764 structures have been modified in some way from the reported experimental data and are freely available for download (http://dx.doi.org/10.11578/1118280). The remaining structures are computation-ready in the form that is already available from the CSD.[6] Addition of new structures to the CoRE MOF collection is efficient, so we anticipate that the database will readily keep pace with future growth in the number of experimentally reported materials.

As an illustration of the utility of the CoRE MOF database, we performed grand canonical Monte Carlo (GCMC) simulations of methane adsorption in each structure to determine the storage capacity at 65 bar and the deliverable capacity from 65 to 5.8 bar. These pressures correspond to the DOE ARPA-E targets for methane storage in adsorbed natural gas fuel tanks within passenger vehicles.[11] Similar GCMC simulations were previously conducted on a database of over 137,000 hypothetical MOFs (hMOFs) to find structural properties that govern methane storage[12, 13] and deliverable capacity.[13] The best performing hMOF structures were found to have helium void fractions around 0.8 and methane heats of adsorption between 10 and 15 kJ/mol. Our GCMC simulations of methane storage and delivery in the CoRE MOF database demonstrate that the structure-property relationships identified using the hMOF database are also found in real structures.

Methods

Collection of 3-D MOF Structures

The procedure for generating the CoRE MOF database is shown schematically in Figure 1. First, potential MOF structures were collected from the CSD version 5.35, which includes submissions through February 2014.[6] The CCDC Conquest program was used to search for structures with more than one bond between metals and the elements O, N, B, P, S, and C. Additionally, we

required the structures to form any kind of bond from these 6 elements to C, N, P, or S atoms. This search yielded over 60,000 candidate MOF structures.

This set of structures includes 1-D coordination polymers and 2-D hydrogen bonded "planar MOFs" in addition to 3-D MOF structures. We identified the 3-D MOF structures in this set by applying an algorithm previously used to determine the dimensionality of void channels.[14] Here, bonded components in a molecular graph of the structure are analyzed to determine the dimensionality of the basis vectors that describe the connection between the bonded component and its images in neighboring simulation cells. A bonded component of a molecular graph refers to a connected set of atoms based on an internal bond criterion from the Zeo++ software.[14] Two atoms are considered bonded if the distance between them is less than the sum of their covalent radii plus a skin distance. This procedure yielded over 20,000 3-D MOF structures that were further considered for inclusion in the CoRE MOF Database.

Structure Preparation

An automated text editor was then used to remove all atoms marked by special characters ("*" and "?") in the crystallographic information file for each structure. The atoms marked with "*" are symmetry-related copies of atoms already present in the structure. The atoms with "?" are atoms with partial occupancy. In most cases, removing one copy of these partially occupied atoms leaves a single representation of chemical moieties such as aromatic rings. Note that this procedure introduces a degree of order which may not exist in the experimental structure. Each structure was then converted to its primitive unit cell and the symmetry was set to P1 using a perl script that interfaces with the Materials Studio software.

Categorizing Structures

The chemical formulas associated with each structure were searched for "+" and "−" symbols to find MOF frameworks that have associated charge-balancing ions. About half of these charged structures were flagged in the CSD as "disordered" and were discarded after visual inspection showed that most of these structures contain major disorder in the framework atoms or lack resolved ionic coordinates.

Some of the MOF frameworks without associated charge-balancing ions that are flagged as "disordered" only contain disorder in the solvent molecules and were retained in the database. These structures were identified by searching the CSD entries for phrases such as "a N, N-dimethylformamide solvent molecule is disordered." Each of these structures was visually checked for disorder in the framework atoms and manual editing was done where appropriate by referring to the literature. For example, the disordered benzene rings in PCN-68 (CCDC: HABRAF)[15] were manually corrected. The structures without associated charge-balancing ionic species were passed directly to the solvent removal step.

Retention of Charge-balancing Ions

Many MOF structures with associated charge-balancing ions also contain undesirable neutral solvent molecules. To discriminate between ionic species and neutral solvent molecules, the elemental compositions of the bonded components in a molecular graph of each structure were compared to the chemical formulas reported by the CSD. The bonded components are the independent "molecules" within each structure – these include the MOF framework, the ionic species, and any neutral solvent molecules. First, the Atomic Simulation Environment[16] NeighborList module was used to construct the periodic adjacency matrix for each structure. Two atoms were considered bonded if the distance between them is less than the sum of their CSD covalent radii[17] plus a skin distance of 0.3 Å. The skin distance is chosen to be slightly smaller

than the CSD definition (0.4 Å), so that the terminal atom connected to the metal atom does not form another bond with other nearby atoms. The adjacency matrix was then passed to the SciPy connected components module[18] to identify the bonded components in each structure. The bonded components with elemental compositions matching the composition of the ions reported by the CSD were exempted from deletion in the solvent removal step.

Solvent Removal

In the solvent removal step, all bonded components in the molecular graph of each structure other than the MOF framework and charge-balancing ions were removed (Figure 2). The MOF framework was defined as the highest molecular weight bonded component of the graph. Interpenetrated MOF frameworks were retained by identifying the number of atoms, *N*, in the largest bonded component in the structure and retaining all additional components having at least 0.5 *N* atoms.

The bonded component corresponding to the MOF framework often includes undesirable solvent molecules bound to unsaturated metal centers. To remove these *coordinated* solvent molecules, we performed a trial "cut" on all bonds between metal centers and oxygen atoms. If the number of bonded clusters detected by the connected component algorithm remained constant, the bond was restored. If the number of bonded components increased, the entire new component was considered a solvent molecule and removed. For example, the dimethylformamide solvent molecules in the MOF JUC-64 (CCDC: OFODET)[19] that are bonded to manganese metal centers were automatically removed in this step. An exception was built into the algorithm to retain hydroxyl groups bonded to metal centers. The most common coordinated solvents removed by our procedure are listed in Table S5. In some cases, these solvent molecules are necessary to stabilize the MOF frameworks. Our desolvation procedure simply removes these molecules without

considering whether the structural integrity of the framework would be retained upon removal of all solvent molecules.

Manual Structure Editing

Structures flagged as "disordered" were retained if the CSD comment field indicated that only the uncharged solvent atoms have disordered coordinates or if the comment field did not explicitly state that the structure is disordered. Each of these structures was visually examined to confirm that the framework atoms are not disordered. This process identified 106 MOF structures with minor framework disorders that were fixed manually by referring to the original literature for the MOF crystal structures. Missing hydrogen atoms were added automatically to an additional 63 structures using Materials Studio (Accelrys, San Diego, CA). Following manual editing, a geometric optimization was run using the Materials Studio Forcite module. The positions of each atom were allowed to relax to a tolerance of 0.002 kcal/mol while the simulation cell was fixed at the experimental lattice constants. Universal Force Field parameters[20] were assigned to each atom in the simulation cell to model all bonded and non-bonded interactions.

Geometric Characterization and GCMC Simulations

Each CoRE MOF structure was characterized with Zeo++[14] using the high-accuracy setting,[21] which uses Voronoi decomposition to identify probe-accessible regions of void space and calculate the accessible surface area, accessible volume, largest cavity diameter (LCD), and pore limiting diameter (PLD).[7] All Zeo++ calculations used a probe of radius 1.86 Å (corresponding to N_2)[22] and covalent radii from the Cambridge Crystallographic Data Centre for all framework atoms. The reported accessible surface areas only include pore regions accessible through windows large enough to admit N_2. Helium void fractions were calculated via Widom particle insertions as described in the Supporting Information.

Classical GCMC simulations were carried out on all CoRE MOF structures to determine methane adsorption at 0.1, 5.8, and 65 bar and 298 K. The isosteric heat of adsorption was calculated at 0.01 bar. Details of the GCMC simulations can be found in the Supporting Information.

Results & Discussion

Geometric and Physical Characteristics of CoRE MOFs

The CoRE MOF database contains 5,109 3-D MOF structures with pore limiting diameters greater than 2.4 Å, which corresponds to approximately the diameter of a hydrogen molecule. We note that some MOF structures are duplicated in our database because multiple representations of the same MOF are reported in the CSD. Notably, there are at least 13 entries for IRMOF-1 and 50 entries for HKUST-1 in the CSD. The authors of new MOF structures often report more than one crystal structure at different activation conditions, different temperatures, or with different guest molecules inside. Additional crystal structures may also be reported by investigators interested in different applications of the same MOF. The lattice constants of these "duplicate" representations often vary, and this may have an effect on simulated adsorption properties. Since the different structures may be of interest to different researchers, we have not removed duplicates from the CoRE MOF database. For some MOFs in the CSD, the authors have reported the common name. This information is collected and listed in Table S12.

The number of structures in the CoRE MOF database does not represent the number of unique porous MOFs synthesized so far because there are duplicate MOF structures in our database and highly disordered structures are not included. Nevertheless, the CoRE MOF database captures a great deal of the chemical and structural diversity in experimentally synthesized MOFs. To assess the structural diversity of the structures, we determined the underlying topology of more than 2,000

CoRE MOF structures with the TOPOS program[23] (See Supporting Information for further details). Over 350 unique topologies were observed in this subset of the database. Among the structures to which we assigned nets, the most common are **pcu** (16%), **dia** (12%), **ths** (3%), **sql** (3%), **rtl** (3%), **srs** (3%), and **bcu** (3%). This distribution of underlying nets qualitatively agrees with the topological analysis performed by Proserpio and co-workers, who found that the most common topologies in 4,709 non-interpenetrating MOFs are **pcu** (9%) and **dia** (6%).[24] In contrast, only 6 topological nets are represented in a recent database of hypothetical MOFs,[12] where over 90% of structures have the 6-coordinated **pcu** topology.[25]

We determined the 3-D space group of each structure using pymatgen library's Symmetry Finder module after the structure went through the cleaning procedure.[26] The CoRE MOF database contains structures with 190 of the 230 possible 3-D space groups. This represents a remarkable degree of structural diversity. The most common space groups are **$P2_1/c$** (17%), **C2/c** (14%), **P1** (12%), **$R\bar{3}$** (3%), and **$Fm\bar{3}m$** (3%). We also find that 297 out of 5,109 structures have space groups that are different from what was reported in the CSD following our cleaning procedure (Table S11). We have also tabulated the metals associated with each structure in the CoRE MOF database in Table S1. The database includes over 50 types of metal clusters, including lanthanides. The most common metal species in the database are Zn and Cu (Table S6), which is not surprising given the widespread use of ZnO_4 and copper and zinc paddlewheel metal centers for the synthesis of metal-organic frameworks.

Figure 3 shows the distribution of calculated volumetric and gravimetric accessible surface areas in the CoRE MOF database (a, b) and in the database of hypothetical MOFs from Wilmer et al.[12] (c, d). The plots show that the structures from the CoRE MOF database have a flatter distribution for the volumetric surface area but a more peaked distribution for the gravimetric surface area.

Synthesis of large surface area MOFs has been an active area of research for the past decade,[27-30] and the skewed distribution of gravimetric surface areas in the CoRE MOF database might be attributed to the lower stability of MOFs with large surface areas and the difficulty in synthesizing such structures.

Methane Uptake Capacity of CoRE MOFs

GCMC simulations of methane adsorption at 298 K were conducted in each CoRE MOF structure. Details of the simulations are provided in the Supporting Information. Figure 4 shows the simulation results for the absolute methane uptake at 65 bar and the deliverable capacity from 65 to 5.8 bar. We found over 800 structures with predicted methane capacity based on our model greater than 200 vol_{STP} vol^{-1} at 65 bar.

The GCMC simulations predict that the top performing structure in the CoRE MOF database for absolute methane uptake is MIL-53(Al) (CSD: HAFQOW),[31] with an uptake at 65 bar of 267 vol_{STP} vol^{-1}, which exceeds the current methane storage target set by the ARPA-E MOVE program (263 vol_{STP} vol^{-1}), if the packing efficiency loss is ignored.[11] This was a surprising result, since this is a well-known MOF and only moderate methane uptake has previously been reported for this material up to 30 bar.[32] We speculated that if special attention were paid to activating the MOF and maximizing its surface area, we might see very high methane uptake in the lab as predicted by simulation. To test this, we synthesized MIL-53(Al) and activated it with a procedure designed to remove solvent and unreacted organic ligands to produce a high-quality sample (S_{BET} = 1,530 m^2/g). However, as shown in the Supporting Information, subsequent experiments with this high-quality MIL-53(Al) sample failed to confirm the simulation prediction, with the methane uptake measured at 65 bar and 298 K being only 190 vol_{STP} vol^{-1}. To assess the origin of this discrepancy, we identified all of the MIL-53(Al) structures in the CSD, and computed full methane isotherms

from 0 to 65 bar. There are 13 different crystal structures for MIL-53(Al) with somewhat different experimentally resolved lattice parameters and atomic coordinates. We found that there are large variations in the simulated methane isotherms in the 13 MIL-53(Al) structures. At 65 bar and 298 K, the methane storage capacity of these structures varies from 180 to 267 vol_{STP} vol^{-1}. Further discussion about methane adsorption in MIL-53(Al) is provided in the Supporting Information. This case study serves as a cautionary example in using the CoRE MOF structures.

Figure 5 shows a parity plot between experimental (BET or Langmuir) surface areas from the literature and our calculated accessible surface areas for the 53 MOFs ranked highest for methane storage and delivery in our GCMC simulations. As found by Goldsmith et al., the correlation between calculated surface areas and experimentally measured surface areas is poor.[10] Only 16 out of the 53 MOFs examined have experimental surface areas larger than 85% of the calculated value. The discrepancy may be a result of defects in the experimentally synthesized MOFs or incomplete removal of solvents. Our computational analysis assumes that every material can be completely desolvated without resulting in framework collapse. For some materials, it may not be possible to access the porosity at all due to framework collapse (experimental surface areas near zero), while in others the experimental synthesis and activation conditions may not have been optimized. Activation of MOFs is an active area of research and the activation conditions of MOFs often need to be varied to yield high-quality MOF samples.[33] For example, the correlation between the experimental and theoretical surface area of IRMOF-1 has improved as better activation protocols have been introduced.[34] Tables S2 and S3 list the reported experimental activation protocol for our top candidates for methane uptake and delivery.

Comparison between CoRE MOFs and hypothetical MOFs for Methane Uptake

In Figure 6, we compare the simulated methane uptake values at 65 bar with those from the hypothetical MOF (hMOF) database of Wilmer et al.[12] Both the hMOFs and the CoRE MOFs span a wide range of textural properties (surface area, LCD, etc.), but the CoRE MOFs are much more topologically diverse. Nevertheless, as shown in Figure 6, the trends for how methane uptake correlates with different textural properties are remarkably similar for the two sets of MOFs. For example, Figure 6c shows that simulated methane adsorption capacity at 65 bar shows a pronounced maximum at a helium void fraction around 0.8 for both databases.

Figure 6d shows that IRMOF-74-XI (RAVXOD),[35] IRMOF-74-IX (RAVXIX),[35] MOF-399 (BAZGAM),[36] and PCN-21 (YUSWEP)[37] have LCDs greater than all of the structures in the hMOF database because large organic linkers in these MOFs were not included by Wilmer et al. in the library of building blocks used to construct the hMOF database.[12]

Figure 7 shows that the two databases also produce similar trends for methane *deliverable* capacity versus heat of adsorption, with the highest methane deliverable capacities occurring at heats of adsorption between 10 and 15 kJ/mol (vertical lines in Figure 7). Figures 6 and 7 suggest that a wide range of textural properties may be more important for developing structure-property relationships for methane storage in MOFs than a wide range of topologies. This is an unexpected result and suggests future work to see if this tentative conclusion holds for other performance properties of MOFs.

Conclusion

We have constructed a database of over 5,000 computation-ready porous MOF structures that were derived directly from experimental crystal data. Efficient algorithms were used to retain charge-balancing ionic species and remove solvent molecules bound to unsaturated metal centers. Physical

and chemical properties of each structure including surface areas and pore characteristics are tabulated in Table S1. The full atomic coordinates of 4,764 of these structures are available to the public on-line (http://dx.doi.org/10.11578/1118280). These 4,764 structures have been significantly modified with respect to the original structures obtained from the CSD to make them suitable for atomistic simulations. For the remaining 346 structures, we provide the CSD reference codes (see Table S10).

To demonstrate the utility of the CoRE MOF database, we conducted GCMC simulations of methane adsorption to predict the methane storage and deliverable capacities of each material. These calculations predicted that MIL-53(Al) would be the top performing material for methane storage capacity, but subsequent experiments did not confirm this prediction. Simulations of methane adsorption in the 13 different MIL-53(Al) structures in the CSD demonstrated that small variations in experimentally resolved atomic positions and lattice constants can lead to large variations in predicted isotherms.

When using the CoRE MOF database, one should be aware that experimentally synthesized materials may not correspond exactly to the pristine, fully desolvated crystal structures reported in this work. Incomplete activation and material defects can significantly decrease the porosity of real materials and have a dramatic influence on adsorption properties. We found that most of the top-performing structures identified by our methane screening efforts do not have experimentally reported BET or Langmuir surface areas close to the theoretical value.

Simulations of methane adsorption in the CoRE MOF database show that the predicted structure-property relationships agree well with those predicted in a database of hypothetical MOFs, even though the hypothetical MOFs are much less topologically diverse. Notably, both the CoRE MOF

and hypothetical MOF databases predict that methane storage capacity is optimized at a helium void fraction of around 0.8 and that methane deliverable capacity is maximized at heats of adsorption between 10 and 15 kJ/mol. Future work is needed to better understand the role of topology on structure-property relationships in MOFs. For example, are certain combinations of textural properties only accessible with certain topologies? We anticipate that the CoRE MOF database will be useful in answering these questions, in facilitating high-throughput identification of candidate MOFs for future applications, and in revealing structure-property relationships that could suggest design principles for optimum materials.

Acknowledgement

The research was supported by the U.S. Department of Energy, Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences and Biosciences under Award DE-FG02-12ER16362. This research was supported in part through the computational resources and staff contributions provided for the Quest high performance computing facility at Northwestern University which is jointly supported by the office of the Provost, the Office for Research, and Northwestern University Information Technology. We thank Dr. Diego A. Gómez-Gualdrón and Dr. Christopher E. Wilmer for providing methane GCMC data for the hypothetical MOFs.

Supporting Information Available: The full atomic coordinates of 4,764 of the CoRE MOF structures are available on-line (http://dx.doi.org/10.11578/1118280). Details of molecular simulation, database construction, MIL-53(Al) synthesis and activation protocols, methane uptake measurements for synthesized MIL-53(Al) and simulated results for MIL-53(Al) are available in the Supporting Information (PDF). Physical properties for the CoRE MOF structures are available in a separate file (XLS). This information is available free of charge via the Internet at http://pubs.acs.org/.

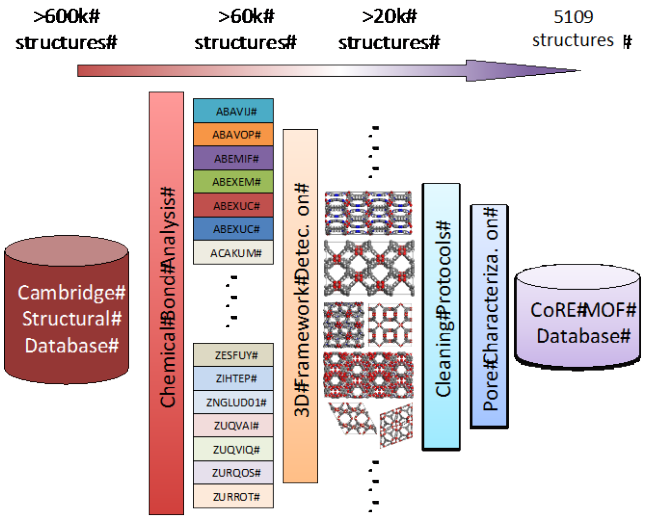

Figure 1. A schematic illustration of the CoRE MOF database construction. Chemical bond analysis was performed using the CCDC Conquest program, 3D framework detection and pore characterization were performed using the Zeo++ Open Source Software, and cleaning protocols were implemented in Python using the ASE and SciPy libraries. All structures in the CoRE MOF database have pore limiting diameters (PLD) > 2.4 Å.

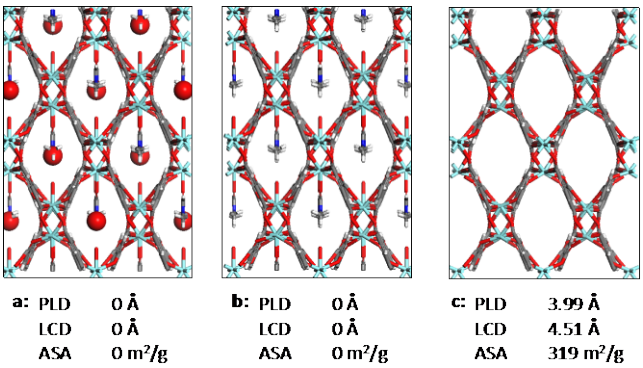

Figure 2. Illustration of solvent removal from a candidate MOF structure (CSD REFCODE: NADZEZ): **a.** Original structure from CSD. Atoms shown with the CPK model are free solvents; **b.** Structure with free solvent removed; **c.** Structure with both free and bound solvent removed. The pore-limiting diameter (PLD), largest cavity diameter (LCD), and gravimetric accessible surface area (ASA) are shown for each structure.

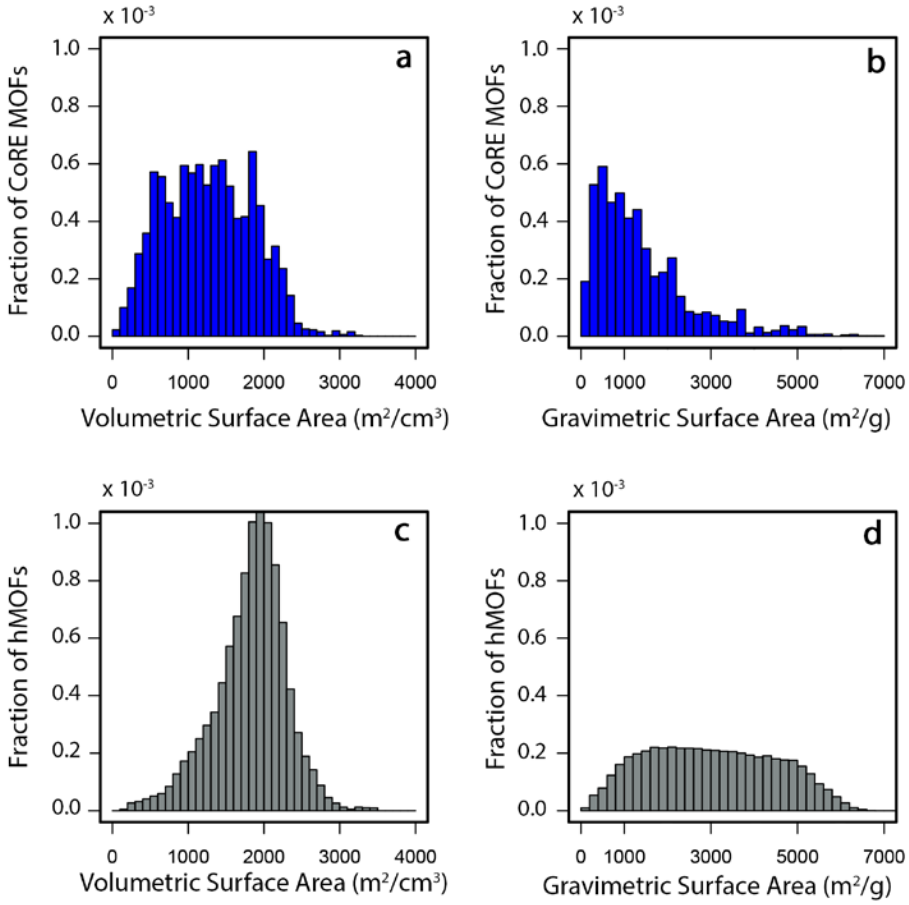

Figure 3. Probability distribution of calculated volumetric and gravimetric accessible surface areas: **a.** Volumetric accessible surface area from the CoRE MOF database; **b.** Gravimetric accessible surface area from the CoRE MOF database; **c.** Volumetric accessible surface area from the hypothetical MOF database of Wilmer et al.; **d.** Gravimetric accessible surface area from the hypothetical MOF database of Wilmer et al. All properties were calculated using Zeo++ with a probe radius of 1.86 Å (corresponding to N_2).

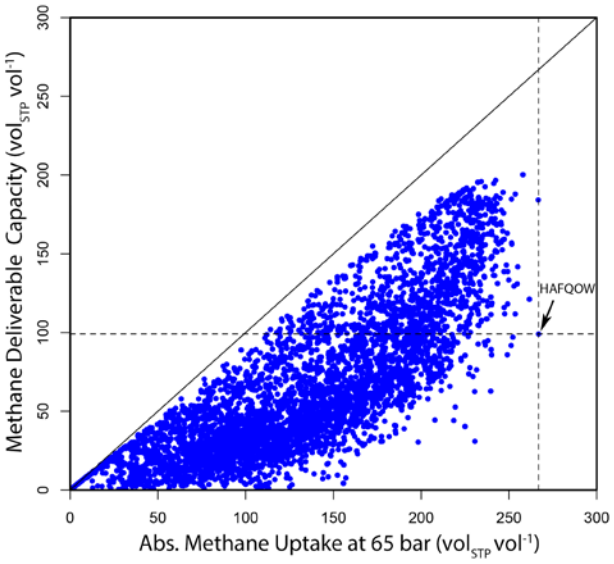

Figure 4. Methane deliverable capacity from 65 to 5.8 bar as a function of absolute methane uptake at 65 bar calculated from GCMC simulations for the structures in the CoRE MOF database at 298 K. The data point at the intersection between the vertical and horizontal dotted lines is HAFQOW (MIL-53(Al)). HAFQOW is the best MOF in terms of methane storage (267 vol_{STP} vol^{-1}) but not in terms of methane deliverable capacity (~100 vol_{STP} vol^{-1}).

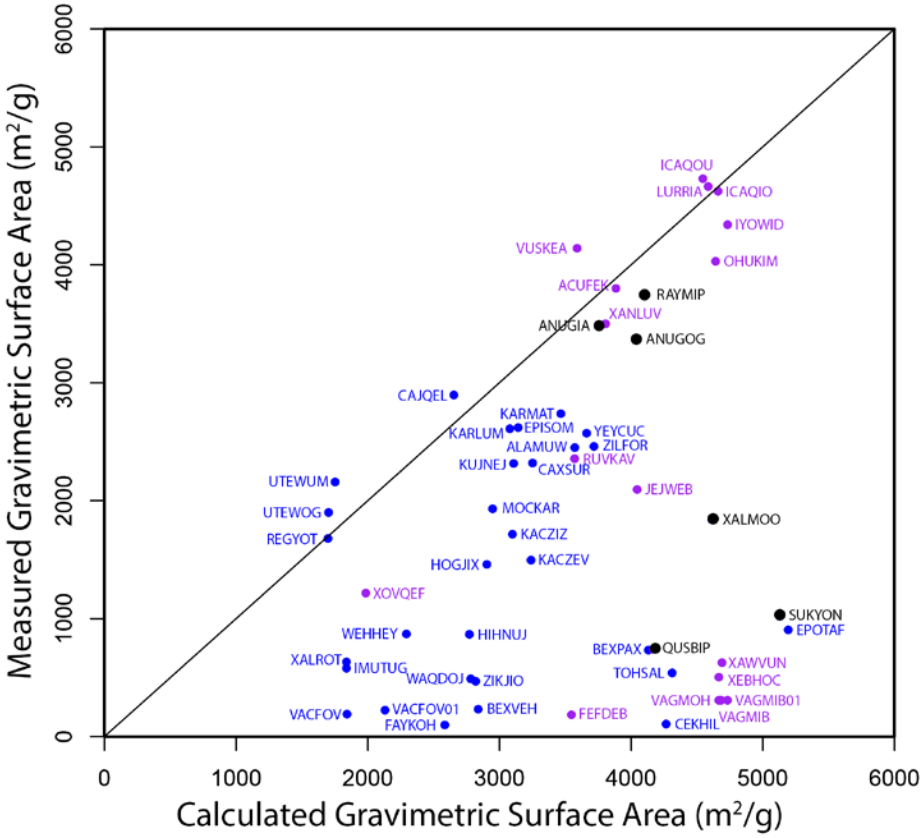

Figure 5. Comparison between gravimetric surface areas from the experimental literature and those calculated geometrically from the crystal structures for structures with methane deliverable capacity greater than 180 vol_{STP} vol^{-1} or methane uptake greater than 240 vol_{STP} vol^{-1}. CSD Reference Codes are shown next to the data points: purple for top methane deliverable capacities, blue for top methane uptake values, and black for both. If the experimental BET surface area was not reported, the Langmuir surface area was used.

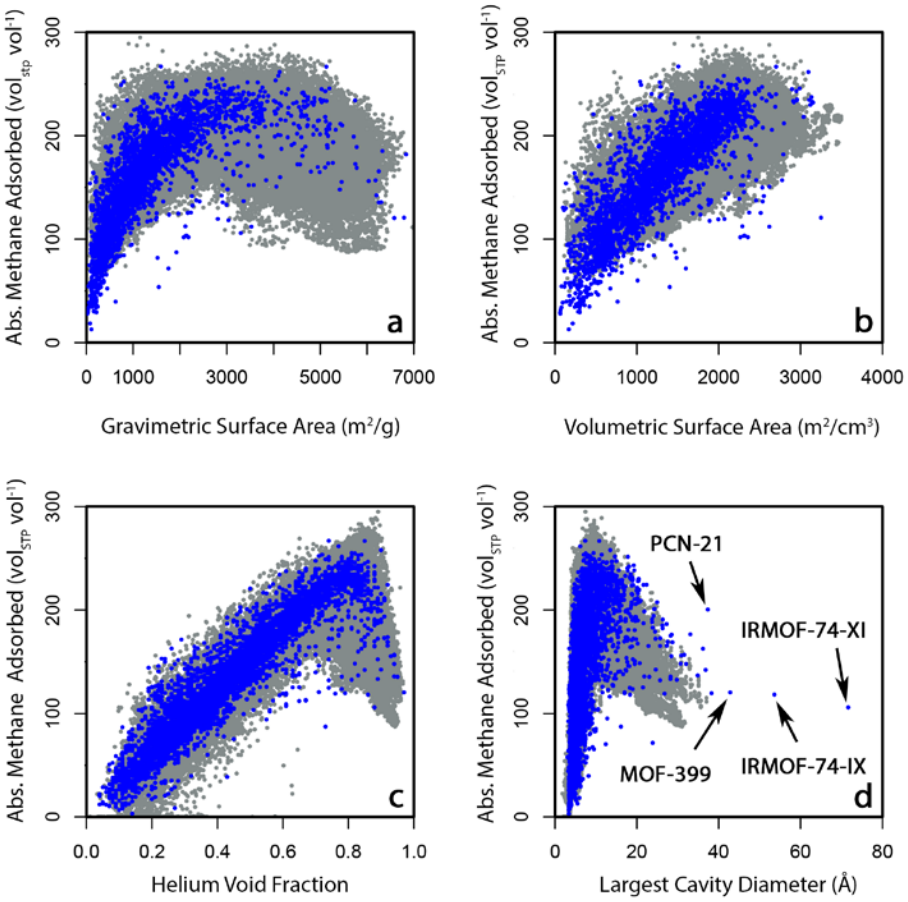

Figure 6. Absolute methane storage capacity at 65 bar for the CoRE MOF structures (blue) and the hMOF structures (grey) plotted as a function of: **a**. gravimetric accessible surface area, **b**. volumetric accessible surface area, **c**. helium void fraction, and **d**. largest cavity diameter.

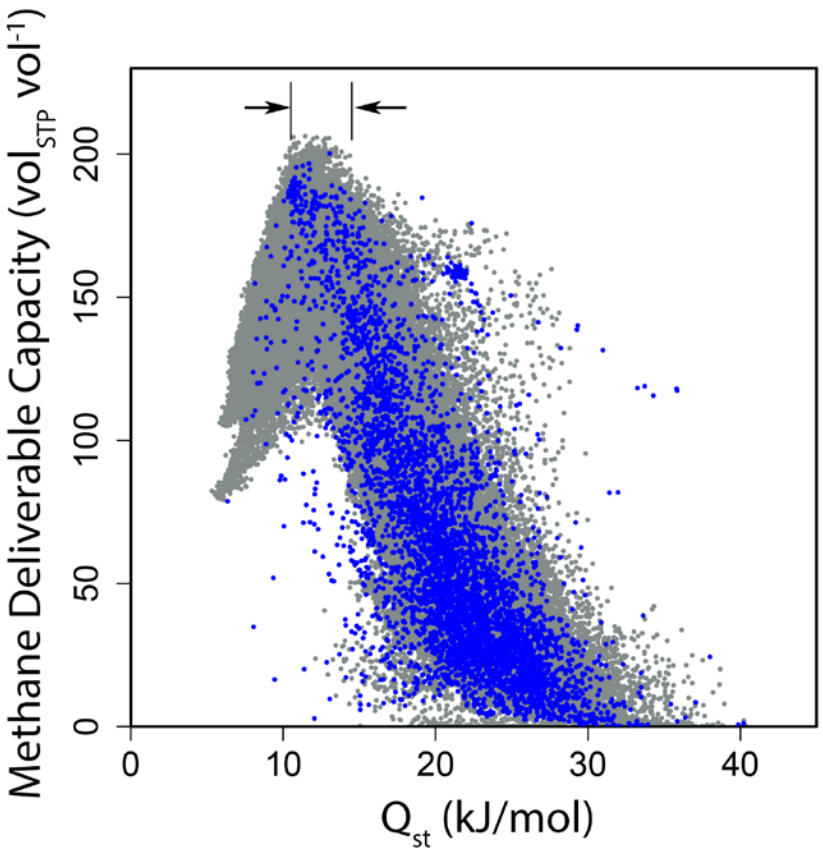

Figure 7. Methane deliverable capacity from 65 to 5.8 bar for the CoRE MOF database (blue) and the hMOF database of Wilmer et al.[12] (grey) as a function of the heat of adsorption (Q_{st}) at 0.01 bar.

REFERENCES

1. Li, J.-R.; Sculley, J.; Zhou, H.-C. *Chemical Reviews* **2011,** 112, (2), 869-932.

2. Kreno, L. E.; Leong, K.; Farha, O. K.; Allendorf, M.; Van Duyne, R. P.; Hupp, J. T. *Chemical Reviews* **2011,** 112, 1105-1125.

3. O'Keeffe, M.; Eddaoudi, M.; Li, H.; Reineke, T.; Yaghi, O. M. *Journal of Solid State Chemistry* **2000,** 152, (1), 3-20.

4. Furukawa, H.; Cordova, K. E.; O'Keeffe, M.; Yaghi, O. M. *Science* **2013,** 341, 1230444.

5. Colon, Y. J.; Snurr, R. Q. *Chemical Society Reviews* **2014**, 43, 5735 - 5749.

6. Allen, F. *Acta Crystallographica Section B* **2002,** 58, 380-388.

7. Haldoupis, E.; Nair, S.; Sholl, D. S. *Journal of the American Chemical Society* **2010,** 132, 7528-7539.

8. Van Heest, T.; Teich-McGoldrick, S. L.; Greathouse, J. A.; Allendorf, M. D.; Sholl, D. S. *The Journal of Physical Chemistry C* **2012,** 116, (24), 13183-13195.

9. Watanabe, T.; Sholl, D. S., *Langmuir* **2012,** 28, (40), 14114-14128.

10. Goldsmith, J.; Wong-Foy, A. G.; Cafarella, M. J.; Siegel, D. J. *Chemistry of Materials* **2013,** 25, 3373-3382.

11. ARPA-E Methane Opportunities for Vehicular Energy (MOVE), **2012**, http://arpa-e-foa.energy.gov (DE-FOA-000672).

12. Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. *Nature Chemistry* **2011,** 4, 83-89.

13. Gómez-Gualdrón, D. A.; Wilmer, C. E.; Farha, O. K.; Hupp, J. T.; Snurr, R. Q. *The Journal of Physical Chemistry C* **2014,** 118, (13), 6941-6951.

14. Willems, T. F.; Rycroft, C. H.; Kazi, M.; Meza, J. C.; Haranczyk, M. *Microporous and Mesoporous Materials* **2012,** 149, 134-141.

15. Yuan, D.; Zhao, D.; Sun, D.; Zhou, H.-C. *Angewandte Chemie International Edition* **2010,** 49, (31), 5357-5361.

16. Bahn, S. R.; Jacobsen, K. W. *Computing in Science & Engineering* **2002,** 4, 56-66.

17. Cordero, B.; Gomez, V.; Platero-Prats, A. E.; Reves, M.; Echeverria, J.; Cremades, E.; Barragan, F.; Alvarez, S. *Dalton Transactions* **2008**, 2832-2838.

18. Pearce, D. J. Technical report, Victoria University, Wellington, NZ: 2005.

19. Xue, M.; Zhu, G.; Li, Y.; Zhao, X.; Jin, Z.; Kang, E.; Qiu, S. *Crystal Growth and Design* **2008,** 8, (7), 2478-2483.

20. Rappé, A. K.; Casewit, C. J.; Colwell, K. S.; Goddard Iii, W. A.; Skiff, W. M. *Journal of the American Chemical Society* **1992,** 114, 10024-10035.

21. Pinheiro, M.; Martin, R. L.; Rycroft, C. H.; Haranczyk, M. *CrystEngComm* **2013,** 15, (37), 7531-7538.

22. Bae, Y.-S.; Yazaydın, A. Ö.; Snurr, R. Q. *Langmuir* **2010,** 26, 5475-5483.

23. Blatov, V. A. *Structural Chemistry* **2012,** 23, (4), 955-963.

24. Alexandrov, E. V.; Blatov, V. A.; Kochetkov, A. V.; Proserpio, D. M. *CrystEngComm* **2011,** 13, 3947-3958.

25. Sikora, B. J.; Winnegar, R.; Proserpio, D. M.; Snurr, R. Q. *Microporous and Mesoporous Materials* **2014,** 186, 207-213.

26. Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. *Computational Materials Science* **2013,** 68, (0), 314-319.

27. Koh, K.; Wong-Foy, A. G.; Matzger, A. J. *Journal of the American Chemical Society* **2010,** 132, (42), 15005-15010.

28. Furukawa, H.; Ko, N.; Go, Y. B.; Aratani, N.; Choi, S. B.; Choi, E.; Yazaydin, A. Ö.; Snurr, R. Q.; O'Keeffe, M.; Kim, J.; Yaghi, O. M. *Science* **2010,** 329, (5990), 424-428.

29. Farha, O. K.; Yazaydın, A. Ö.; Eryazici, I.; Malliakas, C. D.; Hauser, B. G.; Kanatzidis, M. G.; Nguyen, S. T.; Snurr, R. Q.; Hupp, J. T. *Nature Chemistry* **2010,** 2, 944-948.

30. Farha, O. K.; Eryazici, I.; Jeong, N. C.; Hauser, B. G.; Wilmer, C. E.; Sarjeant, A. A.; Snurr, R. Q.; Nguyen, S. T.; Yazaydın, A. Ö.; Hupp, J. T. *Journal of the American Chemical Society* **2012,** 134, 15016-15021.

31. Potts, S. V.; Barbour, L. J.; Haynes, D. A.; Rawson, J. M.; Lloyd, G. O. *Journal of the American Chemical Society* **2011,** 133, (33), 12948-12951.

32. Férey, G.; Mellot-Draznieks, C.; Serre, C.; Millange, F.; Dutour, J.; Surblé, S.; Margiolaki, I. *Science* **2005,** 309, (5743), 2040-2042.

33. Mondloch, J. E.; Karagiaridi, O.; Farha, O. K.; Hupp, J. T. *CrystEngComm* **2013,** 15, (45), 9258-9264.

34. Kaye, S. S.; Dailly, A.; Yaghi, O. M.; Long, J. R. *Journal of the American Chemical Society* **2007,** 129, (46), 14176-14177.

35. Deng, H.; Grunder, S.; Cordova, K. E.; Valente, C.; Furukawa, H.; Hmadeh, M.; Gándara, F.; Whalley, A. C.; Liu, Z.; Asahina, S.; Kazumori, H.; O'Keeffe, M.; Terasaki, O.; Stoddart, J.R.; Yaghi, O. M. *Science* **2012,** 336, (6084), 1018-1023.

36. Furukawa, H.; Go, Y. B.; Ko, N.; Park, Y. K.; Uribe-Romo, F. J.; Kim, J.; O'Keeffe, M.; Yaghi, O. M. *Inorganic Chemistry* **2011,** 50, (18), 9147-9152.

37. Zhuang, W.; Ma, S.; Wang, X.-S.; Yuan, D.; Li, J.-R.; Zhao, D.; Zhou, H.-C. *Chemical Communications* **2010,** 46, (29), 5223-5225.

Table of Contents Image