Chem Soc Rev

High-throughput Computational Screening of Metal-Organic Frameworks

Journal:	*Chemical Society Reviews*
Manuscript ID:	CS-REV-02-2014-000070.R1
Article Type:	Review Article
Date Submitted by the Author:	01-Apr-2014
Complete List of Authors:	Colón, Yamil; Northwestern University, Department of Chemical & Biological Engineering Snurr, Randall; Northwestern University, Department of Chemical Engineering

SCHOLARONE™
Manuscripts

Chem Soc Rev

REVIEW ARTICLE

High-throughput computational screening of metal-organic frameworks

Yamil J. Colón and Randall Q. Snurr[*]

There is an almost unlimited number of metal-organic frameworks (MOFs). This creates exciting opportunities but also poses a problem: How do we quickly find the best MOFs for a given application? Molecular simulations have advanced sufficiently that many MOF properties – especially structural and gas adsorption properties – can be predicted computationally, and molecular modeling techniques are now used increasingly to guide the synthesis of new MOFs. With increasing computational power and improved simulation algorithms, it has become possible to conduct high-throughput computational screening to identify promising MOF structures and uncover structure/property relations. We review these efforts and discuss future directions in this new field.

1. Introduction

Metal-organic frameworks (MOFs)[1-4] are the ultimate designer materials. These novel, crystalline, nanoporous materials are comprised of inorganic subunits (metal ions, clusters, or chains) connected by organic linkers via coordinating groups such as carboxylates, phosphonates, or nitrogen-containing ligands. The modular synthesis approach,[5] along with the great diversity of available building blocks and the ability to introduce chemical functionalities[6-9] into the structures, provides researchers with the opportunity to tune the properties of these materials with great control[10-12] and to design materials for particular applications. As a result, MOFs have been studied for a wide variety of applications, including gas storage,[13, 14] separations,[15-18] sensing,[19, 20] drug delivery,[21-23] light harvesting,[24-26] and catalysis.[27-30]

The number of synthesized MOFs has been increasing exponentially,[31] and given the large number of possible linkers and metal nodes and the various ways of combining them, the number of possible MOF structures is essentially limitless. This presents exciting opportunities, but it also poses a daunting challenge. It is impractical to synthesize and test millions of MOFs for each application of interest; some direction is necessary. Chemical intuition and computational modeling of individual structures will continue to play an important role. However, in the past few years, another powerful tool has emerged: high-throughput computational screening, where the properties of thousands of MOF candidates are evaluated to identify promising candidates and to uncover useful structure/property relations.

The modular building-block approach and the potential for computational screening go to the heart of the newly introduced Materials Genome Initiative (MGI), which aims to "discover,

develop, and deploy new materials twice as fast"[32] as currently possible. MGI seeks to combine theory, computation, synthesis, and characterization to accelerate the discovery of new materials and their release into the market.[33, 34] Undoubtedly, computational high-throughput screening techniques will play a vital role in the development of many classes of materials, not just MOFs.

In this review, we highlight efforts from the recent literature that use high-throughput computational techniques to screen MOFs for gas adsorption and separations. We start with the problem of obtaining the MOF structures, followed by an overview of computational methods for characterizing the structures and predicting their gas adsorption behaviour. Efforts focused on identifying materials for natural gas storage, hydrogen storage, and various separations are reviewed, and data mining techniques are discussed as a way to obtain useful knowledge and insights from the large amount of data generated in high-throughput screening. Finally, we discuss possible future directions and opportunities.

2. Structures

MOF structures for high-throughput computational screening can be taken from experimental crystal structures or (for newly proposed MOFs) generated on the computer. For a single structure, neither option is difficult, but automating these processes for thousands of structures is not a trivial task.

When new MOFs are synthesized, the crystal structures are usually deposited in the Cambridge Structural Database (CSD).[35] A recent paper by Furukawa and co-workers reports that, as of 2011, there were close to 6,000 MOF structures in the CSD (Figure 1).[31] However, these structures are not labelled as "MOFs" and the CSD also contains several hundred

thousand other crystal structures. So, the first problem is to determine which structures in the CSD are MOFs. This can be done by searching for extended structures that contain bonds between metal atoms and elements such as C, B, N, O, Si, P, and S.[36]
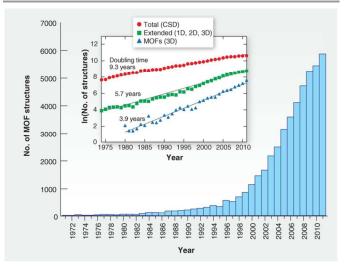

Figure 1. MOF structures reported in the CSD from 1971 to 2011. Reprinted with permission from Ref 31. Copyright 2013 The American Association for the Advancement of Science.

Many of the MOF crystal structures in the CSD contain solvent molecules. In addition, there may be varying degrees of disorder, missing H atoms, overlapping atoms, etc. While removing solvent molecules and correcting a structure are straightforward for a single structure using standard visualization tools, this is not a practical approach for large-scale studies. Therefore, automated methods have been developed by Watanabe and Sholl,[37] and more recently by Goldsmith et al.,[36] to screen through the structures in the CSD to identify MOF structures, remove solvent molecules, fix disorder, etc. Structures that are deemed to be too difficult to fix can be discarded from the screening process. Figure 2 shows a flowchart summarising the process used by Goldsmith et al.[36]
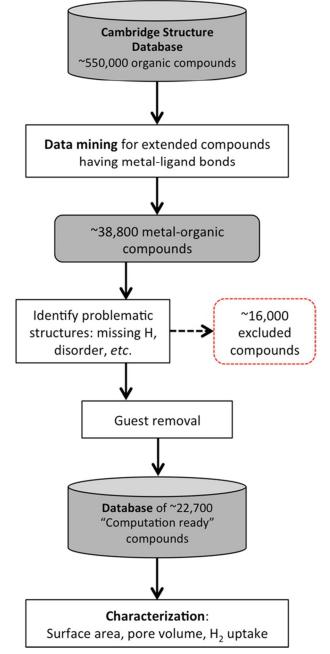

Figure 2. Flowchart of process used by Goldsmith et al. to obtain "computation ready" MOF structures from the CSD. Reprinted with permission from Ref 36. Copyright 2012 American Chemical Society.

An alternative to obtaining MOF crystal structures from the CSD is to take advantage of the building-block nature of MOFs and generate new structures on the computer. Mellot-Draznieks et al.[38] developed an approach known as "automated assembly of secondary building units" or AASBU. Briefly, the building blocks, also known as secondary building units (SBUs), are randomly distributed in a unit cell and given interaction sites at points where they can connect to other building units. These "sticky" sites are parameterized to promote or disfavor certain SBU connections. A simulated annealing Monte Carlo algorithm is used to allow the building

units to rearrange. At each step the cell size and distances between SBUs are allowed to vary to relieve interatomic contacts. One run typically yields ~10^4 trial configurations. Radial distribution functions and simulated diffraction patterns are used to identify duplicates, which are then removed. The configurations are then minimized and any resulting redundancies are removed. This results in a few hundred possible SBU configurations, which are ranked according to a cost function or degree of connectivity, and the symmetry of the arrangement is determined. This provides a set of viable structures that could form from a given set of building units and insight into the topological preferences of certain SBUs. This technique can be used to determine structures of MOFs from powder x-ray diffraction when obtaining large single crystals is difficult.[38]

As an alternative to the energy minimization approach used in the AASBU method, geometric approaches have been developed. These can be classified as "bottom-up" and "top-down." The bottom-up approach consists of sequentially connecting SBUs until a periodic crystal structure is formed. The top down approach starts with a given net or topology, and the appropriate building blocks are then mapped onto the net to generate the structure. Moreover, these techniques allow for the construction of structures that contain more than one linker.[39] Figures 3 and 4 illustrate the top-down and bottom-up approaches, respectively, for generating structures.[40-46] Related methods have been used to generate molecular cages.[47, 48]

For top-down generation, the nets can be obtained from the Reticular Chemistry Structure Resource (RCSR).[2] Several groups have used some of these nets to generate covalent organic frameworks (COFs)[44, 49, 50] and zeolitic imidazolate frameworks (ZIFs)[51-54] using a top-down approach. Lin et al.[52] generated ZIFs using a top-down approach. Using Zeo++[55] zeolites were used as templates for the ZIFs. The unit cell of the corresponding zeolite was scaled by 1.95, which is how many times larger the Zn-imidazole ring distance is than the Si-O distance in zeolites. Oxygen atoms were replaced with imidazole rings and Si with Zn atoms. Resulting geometries were validated using ZIFs with known geometries. More recently, Martin and Haranczyk[56] constructed MOFs based on RCSR topologies, also implemented using Zeo++. Combining this approach with new network generating algorithms may lead to the discovery of new MOFs with nets and topologies not yet synthesized.[57-61]
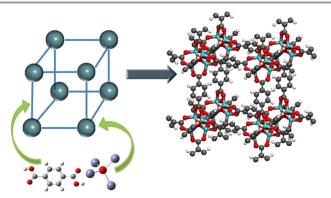

Figure 3. Schematic of the top-down approach. Here, a terephthalic acid linker is mapped onto the edge, and a Zn_4O complex is mapped onto the node of a **pcu** net, forming MOF-5[1]. C = gray, O = red, H = white, Zn = light blue.

A bottom-up approach was developed by Wilmer et al.[42] First, building blocks were extracted from the structures of existing MOFs, and a library was created, including the geometries of the building blocks, information on which blocks could combine with each other, and geometric information on how the building blocks connect (Figure 4). To generate a new MOF structure, building blocks were connected in a step-wise fashion. When an atomic overlap occurred, a new building block or connection site was chosen until all possibilities were exhausted. At some point, instead of adding a building block, periodic boundary conditions were imposed. When no more building blocks could be added, the crystal generation process ended.[41] Starting with a library of 102 building blocks, Wilmer et al. generated 137,953 hypothetical MOFs subject to the constraint that each MOF could contain only one type of metal node and one or two types of organic linkers, along with a single type of functional group. Note that no force field or quantum mechanical energy minimizations are involved in this approach.
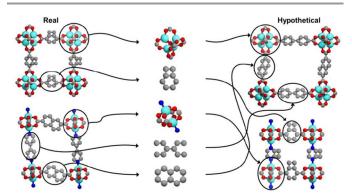

Figure 4. In the bottom-up approach, building blocks are extracted from real MOFs and rearranged into new combinations to generate hypothetical MOFs. C = gray, O = red, H = white, Zn = light blue, N = dark blue. Adapted by permission from Macmillan Publishers Ltd: Nature Chemistry, Ref 42, copyright 2012.

3. Characterization

Given a set of MOF structures, it is useful to calculate their so-called textural properties, such as the surface area and void

fraction. For example, the pore limiting diameter (PLD) and largest cavity diameter (LCD)[37, 46, 62, 63] can be used to narrow down a large set of MOFs to a smaller set with pores large enough to admit a molecule of interest. As shown in Figure 5, the PLD is the size of the largest probe that can traverse through the structure, while the LCD is the largest probe that can fit somewhere within the structure.
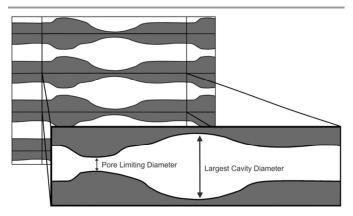

Figure 5. Illustration of pore limiting diameter (PLD) and largest cavity diameter (LCD). Adapted with permission from Ref 62. Copyright 2010 American Chemical Society.

Other useful textural properties include the accessible void volume,[64] He void fraction,[65] accessible surface area,[66-68] and pore size distribution (PSD).[69-72] The accessible void volume can be calculated geometrically using Delaunay tessellation[46, 73, 74] or Voronoi decomposition.[55, 75] In Delaunay tessellation (Figure 6), a collection of points – here, the atoms in a MOF unit cell – are partitioned into the vertices of tetrahedra so as to fill the entire space. Similarly, Voronoi decomposition maps the void space surrounding a set of points by dividing the space into polyhedral cells. (Both of these techniques can also be used to calculate PLD, LCD, accessible surface area, and PSD.) The He void fraction is related to the accessible void volume. The accessible void volume is a purely geometric quantity, while the He void fraction is calculated using Widom insertions of a He probe[76] to mimic how this quantity is measured experimentally using He adsorption.[65] The accessible surface area can be calculated by effectively rolling a probe sphere across the surface of the material.[66-68] The PSD is calculated by randomly selecting points in the structure and recording the radius of the largest sphere containing that point which can fit in the structure.[69]
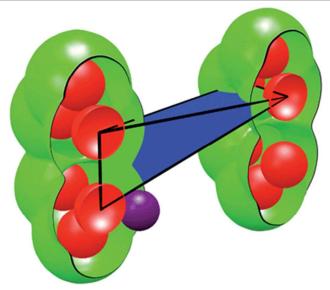

Figure 6. Illustration of Delaunay tessellation. The red spheres represent atoms of a framework which are connected by edges to form tetrahedra. Only one tetrahedron is shown for clarity. The purple sphere represents the probe used to find occupied (red inside tetrahedron), unoccupied (green), and accessible (blue) volume. Reproduced from Ref 46.

Several software packages are available to calculate the textural properties of MOFs and related materials. Zeo++[55] utilizes Voronoi decomposition to calculate the properties. It can calculate PLD, LCD, accessible surface area, accessible void volume, and pore size distributions taking into account inaccessible regions.[77] It can also be used to analyse pore similarity and to generate MOF structures.[52, 56, 72, 78] MOFOMICS[79] is able to identify portals, channels, cages, and connectivity. It identifies portals through k-cycle enumeration, which grows paths iteratively, until they can be closed. Subsequently, channels are identified by the largest void cylinder that can fit between portals. Cages are identified using Delaunay triangulation but only recording the spheres larger than a given threshold. The connectivity is determined by finding "junctions," i.e., places where molecules can change their direction of travel. Then, channel-channel and channel-cage intersections are calculated by intersecting the channels (cylinders) and cages (cylinders). The channels and intersections are examined to find the connectivity between junctions. Poreblazer[70] differs from the previous software packages in that it divides the empty space into cubelets and utilizes them to characterize the pore structure. It can calculate surface area, pore size distribution, connectivity, LCD, and PLD. These software packages can be used to detect guest-inaccessible regions, so that molecules are not inserted in these regions in Monte Carlo simulations.[80]

The TOPOS software[81] can be used to find the underlying topology of a particular structure as well as the cavities in the structures and their sizes. A given MOF structure can be simplified by taking the metal corners as nodes and the organic linkers as edges. Using this criterion, the underlying net may be found. Recently, the developers of this software analysed 6620 3-periodic structures obtained from the CSD and

determined their topologies, finding correlations between specific building blocks and the resulting topology.[81] Figure 7 indicates that **pcu** followed by **dia** are the most frequent nets in the structures analysed. It is also possible to consider parts of the organic linkers as nodes. For instance, a tri-topic linker, which has three connections originating from a central point, could be broken up into three edges (connection sites) and one node (central point).[82]
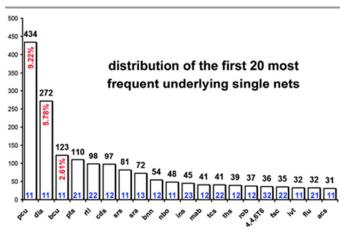

Figure 7. Distribution of first 20 most frequent underlying nets of non-interpenetrated structures analysed by Alexandrov et al[81]. Bottom numbers in blue indicate transitivity (number of unique nodes and unique edges). Reproduced from Ref. 81.

All of these algorithms and software packages are well suited for automated, high-throughput screening of porous materials.[55, 72] Recent efforts involve the use of graphics processing units (GPUs) due to their speed and low price.[46, 78, 83, 84] Calculating the distribution of textural properties for a collection of MOF structures is a useful way to determine the diversity of the structures.[78, 85] This can be important if the goal is to find the best material for some application or to establish widely applicable structure/property relationships. Meeting these goals is facilitated if the full span of physical properties has been covered.

4. Calculation of Adsorption Properties

Molecular modeling can be used to predict properties such as binding energies, adsorption isotherms, and diffusion coefficients for guest molecules in MOFs. The computational methods can broadly be classified as those based on quantum mechanics and those based on classical mechanics and are reviewed elsewhere.[67, 76, 86-90] Briefly, quantum mechanical calculations, such as density functional theory (DFT), typically solve the time-independent Schrödinger equation to find minimum energy structures, binding energies, and details of the electronic structure. Quantum mechanical calculations have been used to screen dozens of MOFs, but not thousands, to date because of the large amount of computer time required. However, quantum mechanical calculations can also be used to parameterize force fields,[91] which can then be used in classical simulations on larger numbers of structures.

Simulations based on a classical picture use the principles of statistical mechanics to calculate macroscopic thermodynamic and transport properties. Grand canonical Monte Carlo (GCMC) simulations can predict enthalpies of adsorption, adsorption isotherms, and (for mixtures) selectivities. Molecular dynamics (MD) simulations can provide diffusion coefficients and other transport properties. A key input to GCMC or MD simulations is a set of equations and parameters describing the energetic interactions among the atoms. The energies among non-bonded atoms are often described by simple Lennard-Jones plus Coulomb potentials.[41, 67] For fairly rigid structures, the framework atoms are usually held fixed at their crystallographic coordinates, but for flexible MOFs, the movements of the framework atoms must be included. Recently, Sarkisov and co-workers[92] developed a simple method to predict whether a structure is flexible or not, which may be quite useful in this regard. Lennard-Jones parameters for framework atoms are often taken from generic force fields such as DREIDING[93] or the Universal Force Field (UFF).[94] Efforts are also underway to develop more sophisticated force fields.[44, 95-98] For the guest molecules, it is often recommended to use force fields fit to match the bulk vapour/liquid equilibria.[99]

5. High-throughput computational screening

High-throughput screening of MOFs has only become possible within the past few years. To date, most efforts have focused on the adsorption of small molecules, motivated by gas storage and separation applications. As we illustrate below, these studies can reveal promising candidates, new structure/property relationships, and possible performance limits for these materials.

5.1 Adsorption of small molecules

As noted above, quantum mechanical calculations have not yet been applied in a truly high-throughput manner to screening of MOFs. However, several groups have used quantum mechanical calculations to investigate the effect of substituting different metals into a particular MOF structure. Special attention has focused on MOF-74,[100] also known as CPO-27[101] or M/DOBDC,[102] which has a high density of so-called "open" metal sites where one coordination site is empty. These open metal sites have been demonstrated, both experimentally and via modeling, to interact strongly with various adsorbates. Several groups screened different metals in MOF-74 for their ability to bind CO_2.[103-106] Park et al.[104] predicted that Ti- and V-MOF-74 should have a stronger affinity for CO_2 than Mg-MOF-74. This is a very interesting prediction because the Ti and V versions of MOF-74 have not yet been synthesized, and the Mg version currently shows the strongest binding of CO_2. These calculations also provide insights into the nature of the binding.[103-106] Canepa and co-workers studied the interaction of H_2, CO_2, CH_4, and H_2O with 25 different metals in MOF-74.[103] Starting from Zn-MOF-74, subsequent metals were substituted for Zn and the structures were allowed to fully

relax. The DFT calculations predicted that all of the metals studied except Rh, Pd, Os, Ir, and Pt bind H_2O preferentially over CO_2. Figure 8 illustrates the binding site for H_2O with Rh, Pd, Os, Ir, or Pt.[103] Systems that bind CO_2 preferentially over water may be useful for CO_2 capture under humid conditions.
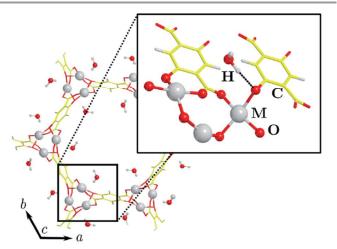

Figure 8. M-MOF-74, where M is one of the noble metals Rh, Pd, Os, Ir, and Pt. Dashed lines indicate a hydrogen bond. Reproduced from Ref. [103].

Methane storage in MOFs has received considerable attention[86, 101, 107-111] driven by energy applications, such as natural gas vehicles. Recently, high-throughput computational screening has been applied to search for better MOFs for natural gas storage. Using a bottom-up structure generation scheme, Wilmer et al.[42] generated 137,953 hypothetical MOF structures and screened them for methane storage. The structures were built using 102 building blocks that were extracted from real MOFs (Figure 4). Methane uptake was calculated for all of the structures using GCMC simulations at 35 bar and 298 K. To speed up the calculations, they were performed in stages. In the first stage, short GCMC simulations were performed for all structures. The structures were ranked from best to worst in terms of methane uptake at 35 bar as shown in Figure 9, and the top 5% were then screened again with longer simulations. Finally, the top 5% from the second stage were subjected to even longer simulations. Using this methodology over 300 hypothetical MOFs were identified which are predicted to adsorb more methane at 35 bar than the world record holder at that time, PCN-14.[112] In addition, several structure/property relationships were identified. For example, Figure 10 shows how methane adsorption at 35 bar correlates with the material's void fraction. It can be seen that, despite a diverse range of textural properties, the best materials all have a void fraction around 0.8. This study illustrates the potential of high-throughput screening techniques to 1) identify promising candidates for synthesis and 2) uncover useful structure/property relationships. The complete database of hypothetical MOFs is accessible online at hmofs.northwestern.edu.
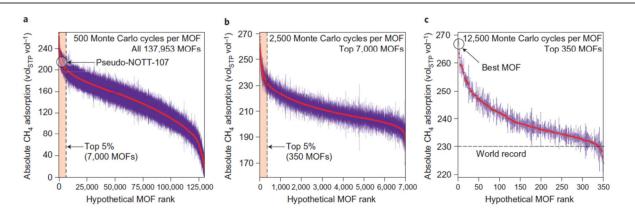

Figure 9. Three-stage screening to identify the best MOFs for methane storage. **a**) In the first stage, 137,953 hypothetical MOFs were screened for methane storage at 35 bar using short GCMC simulations, **b**) In the second stage, the top 5% of structures identified in the first stage were simulated using more Monte Carlo cycles, **c**) In the third stage, the top 5% from the second stage were simulated using even more Monte Carlo cycles. The orange areas in the first two graphs indicates the top 5% of structures in each graph. Purple bars indicate the statistical error. In all graphs, the MOFs are ranked from best to worst according to methane uptake at 35 bar and 298 K in volumetric units. Reprinted by permission from Macmillan Publishers Ltd: Nature Chemistry, Ref 42, copyright 2012.

Chem Soc Rev   RSC**Publishing**
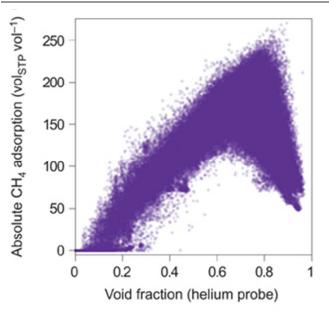
REVIEW ARTICLE

Figure 10. Absolute methane adsorption at 35 bar and 298 K versus void fraction. Optimal values are obtained at a void fraction of 0.8. Adapted by permission from Macmillan Publishers Ltd: Nature Chemistry, Ref 42, copyright 2012.

Hydrogen storage in MOFs has also received considerable attention in the past decade.[13, 87, 113-121] From the literature, it is now known that for room temperature hydrogen storage, the heats of adsorption of MOFs are too low to reach current targets. One strategy to overcome this is to introduce strongly interacting functional groups, such as Mg alkoxides.[122-124] However, it is not readily apparent what combination of MOF topology, pore size, void fraction, etc. is optimal and what density of functional groups should be introduced. To answer these questions, over 18,000 MOFs and porous aromatic frameworks (PAFs)[125, 126] were screened for hydrogen storage.[127] As in the work of Wilmer et al.,[42] the structures were generated in a bottom-up approach. These hypothetical structures contained various numbers of Mg alkoxide sites. Due to the strong interactions between the Mg alkoxide groups and the H_2 molecules, generic force fields are not adequate. Hence, the GCMC simulations, used to calculate hydrogen uptake at 243 K, employed a first principles-derived force field for the hydrogen-Mg alkoxide interactions.[122] Structures were found that are predicted to outperform currently known structures in both gravimetric and volumetric storage. Structure/property relationships were also revealed. For example Figure 11 shows that very high void fractions (0.9) and low Mg densities (0.0 mmol/cm^3 – 0.5 mmol/cm^3) are

optimal for gravimetric uptake, while void fractions around 0.7 and a Mg density of 2.5 mmol/cm^3 are optimal for volumetric uptake.
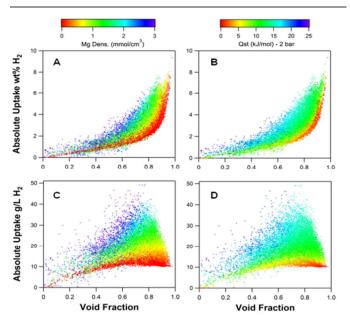

Figure 11. Absolute gravimetric (top) and volumetric (bottom) H_2 uptake versus void fraction obtained from simulated isotherms at 243 K and 100 bar on 18,383 different materials. Colors indicate the Mg alkoxide density (left), and the isosteric heat of adsorption at 2 bar (right). Reprinted with permission from Ref [127]. Copyright 2014 American Chemical Society.

MOFs from the CSD have also been screened for their hydrogen storage potential. Goldsmith et al[36] used data mining techniques to identify MOF structures in the CSD (Figure 2). Instead of performing molecular simulations, they used previously observed correlations with the surface area and pore volume to estimate the hydrogen uptake at 77 K and 35 bar. Promising structures for cryogenic hydrogen storage were identified,[36] and the maximum volumetric hydrogen uptake was found for structures with surface areas around 3100-4800 m^2/g. These authors also explored the trade-off between volumetric capacity and gravimetric capacity as shown in Figure 12. The results show a concave downward relationship between volumetric and gravimetric storage capacities.
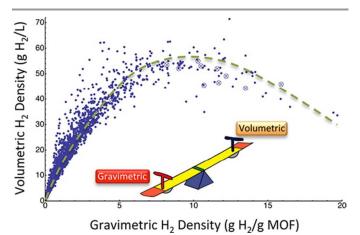
Figure 12. Theoretical absolute H_2 gravimetric and volumetric uptake at 77 K and 35 bar for ~4000 MOFs obtained from the CSD. Crossed circles represent MOFs with incomplete or disordered crystal data in the CSD. These structures were constructed by hand. Reprinted with permission from Ref 36. Copyright 2012 American Chemical Society.

5.2 Separations

Screening for separations applications is more complex than the examples highlighted above, because multiple adsorbates are involved and diffusion effects may be important. In addition, the material will ultimately be incorporated into a process, and the material cannot be optimized without considering this process. This makes it more difficult to determine the selection criteria for the best material for a given separation.

Consider the separation of noble gases. These have a wide range of applications (lasers[128], medicine,[128, 129] etc.), and their separation usually takes place through the use of cryogenic distillation – an energy intensive and costly process. Hence, it is of interest whether MOFs could be used to separate mixtures of noble gases.

Van Heest et al. screened over 3000 MOFs extracted from the CSD for the separation of Ar/Kr, Kr/Xe, and Xe/Rn mixtures.[63] PLDs for all structures were calculated, as well as the Henry's constants. Self-diffusivities (D_s) were estimated using transition state theory (TST).[88] With these quantities, adsorption selectivities and permselectivities were calculated. The adsorption selectivity is a strictly thermodynamic quantity and is relevant to adsorption processes such as pressure swing adsorption as well as to membrane separations. The permselectivity applies to membrane applications and takes into account both sorption into the material (here via the Henry's constants) and transport through the membrane (here via the self-diffusivities). The starting list of over 3000 MOFs was reduced to 70 by choosing those structures with selectivities greater than 30 and permselectivities greater than 10. GCMC simulations were then performed on these 70 structures to generate the pure component isotherms, and ideal adsorbed solution theory (IAST)[130] was used to predict the mixture isotherms and selectivities from the pure-component data. Interestingly, some structures showed reverse selectivity (preferential adsorption of the smaller molecule). Figure 13

shows that structures with a fractal dimension above 5 selectively adsorb the smaller of the adsorbates. In other cases, Kr was favoured over Xe (reverse selectivity) but Rn over Xe (normal selectivity). For these cases, the geometric argument is not enough. So, energetic considerations were studied. The researchers found that for distances between 3.92 and 4.03 Å in relation to carbon the interactions are favourable for Kr over Xe and Rn over Xe. If a material has many regions where interactions at these distances take place, the material will be selective for Kr over Xe and Rn over Xe.[63]
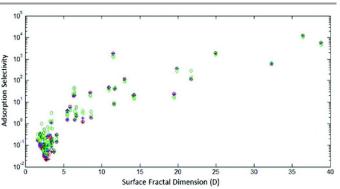

Figure 13. Selectivity for Kr over Xe calculated using IAST for a 80:20 mixture of Kr-Xe versus the surface fractal dimension for probes between the sizes of Kr and Xe. Adapted with permission from Ref. 63. Copyright 2012 American Chemical Society.

Sikora et al.[46] screened the 137,000 hypothetical MOFs generated by Wilmer et al.[42] for Xe/Kr separation. Delaunay tessellation was used to calculate the PLD and LCD of the structures. The calculation was performed using GPUs. Instead of using IAST, multicomponent GCMC simulations were performed to calculate selectivities and adsorption capacities. This large scale study revealed that structures with pore sizes that can fit a single Xe atom along with morphologies resembling tubes (LCD/PLD ratio between 1 and 2) maximize selectivity. See Figure 14.
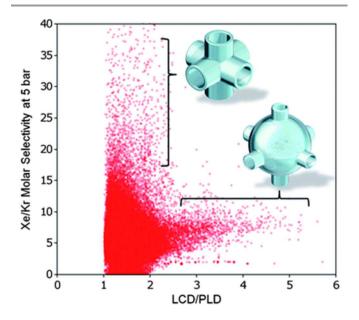
Figure 14. MOFs with tube-like pore morphologies show high selectivity for Xe over Kr. Reproduced from Ref 46.

Kinetic separations based on different rates of molecular diffusion have been studied by Haldoupis et al.[62] and First et al.[131] Haldoupis et al.[62] used a combination of PLDs, LCDs and Henry's constants to study the kinetic separation of CH_4/H_2 mixtures in over 500 MOFs (Figure 15). Figure 16 shows the LCD and PLD for > 500 MOFs. It also shows PLD ranges (arrows) where the adsorbates will have significant diffusion activation energies.[62] From this information, one can identify structures that should be capable of separating adsorbates through molecular sieving (one molecule can go in the structure but the other cannot). Similarly, First et al.[131] characterized the portals, cages, and connectivity (using MOFomics[79]) of over 1800 microporous materials including zeolites, MOFs, ZIFs, and hypothetical MOFs. Subsequent calculation of the energy needed for various adsorbates to pass through portals led to the identification of promising materials for various separations: CO_2/N_2, CO_2/CH_4, CO_2/H_2, O_2/N_2, propane/propylene, ethane/ethylene, styrene/ethylbenzene, and xylene separations.[131]
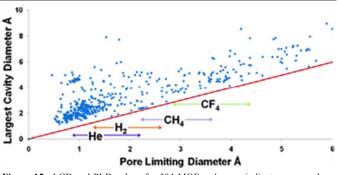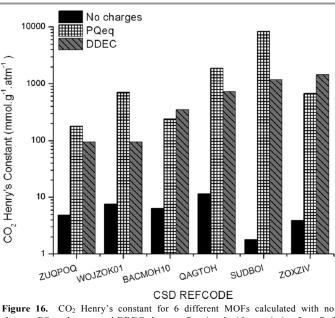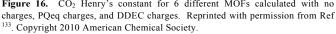

Figure 15. LCD and PLD values for 504 MOFs. Arrows indicate ranges where adsorbates show significant diffusion activation energy. Reprinted with permission from Ref 62. Copyright 2010 American Chemical Society.
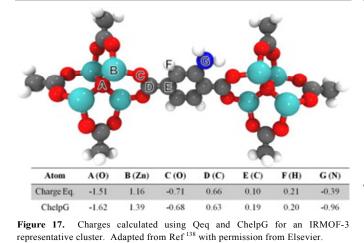
Separations of mixtures containing CO_2 are important for upgrading of natural gas (mainly separating CO_2 from CH_4) and for carbon capture (mainly separating CO_2 from N_2). In contrast to noble gases and methane, which are usually modelled with no charges, Coulombic interactions are important for CO_2. Traditionally, atomic charges for the MOF atoms have been assigned using quantum mechanical calculations.[41, 89, 90, 132] (It should be kept in mind that partial charges are not an experimental observable, and there are a variety of methods for extracting atomic charges from the results of a quantum mechanical calculation.) However, performing quantum mechanical calculations for thousands (or millions) of MOFs may not be feasible. Thus, other techniques for accurately and efficiently assigning charges have been developed in recent years, particularly for high-throughput screening studies. Figure 16 shows that ignoring Coulombic interactions provides very poor estimates of the Henry's constant of CO_2 in some representative MOFs.[133]

Figure 16. CO_2 Henry's constant for 6 different MOFs calculated with no charges, PQeq charges, and DDEC charges. Reprinted with permission from Ref 133. Copyright 2010 American Chemical Society.

Zhong and co-workers developed a very fast method for estimating MOF partial charges known as the connectivity-based atom contribution (CBAC) method.[134, 135] The method is based on the observation that although the number of possible MOF structures is infinite, the elements used are not. The key assumption is that atoms with same bonding connectivity have the same charge in different MOFs. Using a training set of 30 MOFs and a validating set of 13 MOFs, CBAC charges were used to calculate pure component isotherms for CO_2, CO, and N_2 and the isotherms agreed well with those obtained using DFT charges.[134]

Several groups[133, 136-138] have explored the use of charge equilibration methods (Qeq)[139] to calculate the partial charges of MOF atoms. Qeq uses the experimentally determined ionization potential and electronegativity of the atoms and the molecular geometry to predict the charges. Wilmer and co-workers[138] developed their own variant of Qeq and compared the charges on representative fragments of MOFs calculated from Qeq and ChelpG, a quantum mechanical method. As shown in Figure 17, there is reasonable agreement between the charges from the two methods.

Figure 17. Charges calculated using Qeq and ChelpG for an IRMOF-3 representative cluster. Adapted from Ref [138] with permission from Elsevier.

Haldoupis and co-workers[133] introduced a periodic version of Qeq (PQeq)[133, 140] and used it to assign framework charges for 500 MOFs obtained from the CSD. They then calculated Henry's constant to obtain the CO_2/N_2 and CO_2/CH_4 selectivities at low loading (Figure 18). The structures that were deemed promising were subjected to more detailed GCMC and molecular dynamics (MD) simulations.
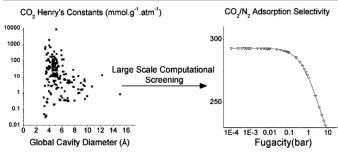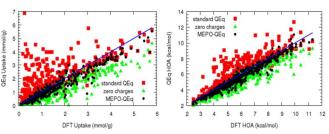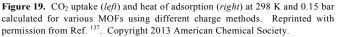

Figure 18. Henry's constants calculated using PQeq charges (left), which were used to narrow down the number of structures and calculate more detailed pure component isotherms using GCMC simulations. IAST was then used to predict mixture isotherms and selectivities (right). Reprinted with permission from Ref. [133] Copyright 2010 American Chemical Society.

Kadantsev et al.[137] developed a Qeq method (MEPO-Qeq) in which the parameters were trained to reproduce DFT-derived electrostatic potentials. A training set of 543 hypothetical MOFs was used, and the parameterization was validated by comparing CO_2 uptake and heats of adsorption calculated using MEPO-Qeq to those calculated using DFT (Figure 19). All of the methods mentioned for calculating partial charges of MOF atoms seek a compromise between time efficiency and the rigor of the method.

Figure 19. CO_2 uptake (*left*) and heat of adsorption (*right*) at 298 K and 0.15 bar calculated for various MOFs using different charge methods. Reprinted with permission from Ref. [137]. Copyright 2013 American Chemical Society.

Many studies of separations in MOFs focus on the selectivity as a metric for ranking materials. However, the selectivity is not the only property that determines the effectiveness of a material in a separation process. As noted above, it is ultimately the performance of the combined material plus process that matters. To avoid the need for a full process design to evaluate each candidate material, researchers have developed various short-cut metrics for materials screening. Bae and Snurr[18] discussed five adsorbent evaluation criteria from the engineering literature and used them to assess over 40 MOFs for their potential in four related CO_2 separations. To calculate the adsorbent evaluation criteria, they used experimental, pure-component isotherm data for CO_2, CH_4, and N_2 from the literature. The evaluation criteria are described in Table 1. None of them is perfect, and they are best considered together. Recently, Wilmer et al.[141] used these metrics to screen their database of 137,000 hypothetical MOFs. Framework charges for the MOFs were calculated very quickly using an extended charge equilibration method (EQeq) that they developed.[136] Using GCMC simulations, pure component adsorption data were obtained for CO_2, CH_4, and N_2. The results were then used to calculate the five adsorbent evaluation criteria for four different separation cases.[141]

Table 1. Adsorbent evaluation criteria. The subscripts 1 and 2 indicate CO_2 and the other, more weakly adsorbing component, respectively. The superscripts ads and des indicate adsorption and desorption conditions, respectively, and *y* is the mole fraction in the gas phase.

Criterion	Definition
CO_2 uptake (mol kg^{-1})	N_1^{ads}
Working capacity (mol kg^{-1})	$\Delta N_1 = N_1^{ads} - N_1^{des}$
Regenerability (%)	$R = \Delta N_1 / N_1^{ads} \times 100\%$
Selectivity	$\alpha_{12}^{ads} = (N_1^{ads}/N_2^{ads})/(y_1/y_2)$
Sorbent selection parameter	$S = (\alpha_{12}^{ads})^2/(\alpha_{12}^{ads})(\Delta N_1/\Delta N_2)$

Both Bae and Snurr[18] and Wilmer et al.[141] used their data to look for relationships between the adsorbent evaluation criteria and the physical properties of the MOFs. Figure 20 shows an example relating the amount of CO_2 adsorbed at 2.5 bar and the isosteric heat of adsorption. As shown in the figure, it can be difficult to establish whether any relationships exist if there are only a small number of data points. However, with over

137,000 data points, clear trends emerge. This highlights one of the biggest advantages and potential impacts of high-throughput computational screening: the ability to discover structure/property relationships that were previously impossible to discern due to the small sample size available.[141]
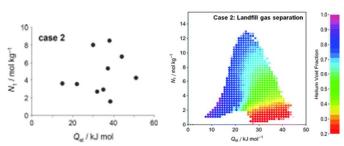

Figure 20. CO_2 uptake at 2.5 bar versus the isosteric heat of adsorption. The graph on the left plots experimental data collected from the literature by Bae and Snurr and shows no clear trend between uptake and heat of adsorption. The graph on the right shows simulation results from over 130,000 hypothetical MOFs and shows a clear trend. Left adapted from Ref 18. Copyright 2011 WILEY-VCH Verlag GmBH & Co. KGaA, Weinheim. Right adapted from Ref 141.

In evaluating materials for CO_2 capture from power plant exhaust, Lin et al. adopted another approach for material evaluation.[52] They calculated the parasitic energy, i.e., the additional electrical energy needed from the power plan to operate the process for separating CO_2 from the flue gas. They screened both real and hypothetical zeolites and ZIFs to find materials with minimum parasitic energy. Charges for the structures were determined using the CBAC method, and Widom insertions were used to calculate Henry's coefficients and isosteric heats of adsorption at low loading. Using the Henry's coefficients and saturation loadings obtained from a correlation with the pore volume, single or dual-site Langmuir models were fit for pure-component N_2 and CO_2 isotherms. In contrast to the studies highlighted above, the mixture isotherms were predicted using competitive Langmuir isotherms instead of IAST or multicomponent GCMC simulations. Lin et al. found that materials should have CO_2 binding energies that are strong enough to be selective but not so strong that the CO_2 cannot be desorbed, to avoid an energy penalty in the regeneration of the material (Figure 21).[52] Furthermore, this screening established a theoretical limit for the lowest parasitic energy of this particular class of materials. This highlights another attractive feature of large-scale, high-throughput screening: performance limits of a material class may be found.
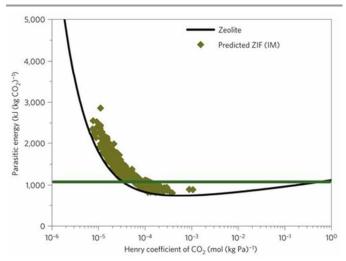

Figure 21. Parasitic energy for CO_2 capture versus Henry coefficient of CO_2. The green line gives the parasitic energy of current MEA technology, while the black line is the minimal parasitic energy calculated in the all-silica zeolite structures. Diamonds are predicted ZIF structures; only a diverse, representative set are shown. Reprinted by permission from Macmillan Publishers Ltd: Nature Materials Ref. 52, copyright 2012.

6. Data mining

An interesting aspect of large-scale, high-throughput screening is the large amount of data that is generated. Often, even plotting the data can prove difficult because of the high dimensionality of the data sets. Simple plots such as those in Figures 11, 12, and 14 can be used to test hypotheses about how different variables are correlated. However, it may be unclear which variables to plot. More sophisticated data mining tools can be very useful in obtain new insights and understanding from the large amount of data generated in high-throughput screening. For instance, Fernandez et al.[142] employed quantitative structure-property relationship (QSPR) tools to analyse methane uptake data in 137,000 hypothetical MOF structures.[42] Nonlinear supported vector machines (SVMs) were found to predict methane storage adequately. The two descriptors that were found to be most strongly correlated with methane uptake were the void fraction and pore diameter. Figure 22 shows the results obtained from the SVM model.[142] Interestingly, the predictions showed a maximum in an unexplored regime (red arrow in Figure 22).
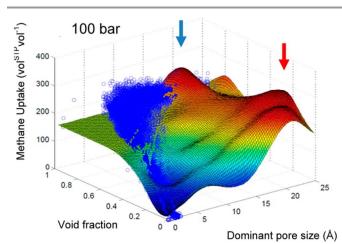
Figure 22. Response surface of SVM model for methane storage at 100 bar using void fraction and dominant pore size. Blue dots are GCMC results. Color of surface represents methane storage value: blue is low and red is high. Arrows indicate maxima. Reprinted with permission from Ref. 142. Copyright 2013 American Chemical Society.

Similarly, Wu et al.[143] developed QSPR models to predict CO_2/N_2 selectivity. The important descriptors in the model were the difference in heat of adsorption between CO_2 and N_2 (ΔQ^o_{st}) and the porosity (φ) of the structure. Simultaneously increasing the difference in heat of adsorption and decreasing the porosity was found to be a promising strategy as shown in Figure 23.
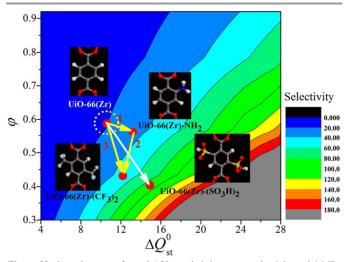

Figure 23. Interplay map of φ and ΔQ^o_{st} on their impact on selectivity at 0.1 MPa for CO_2/N_2 mixture in MOFs. Reprinted with permission from Ref. 143 Copyright 2012 American Chemical Society.

Other descriptors have also been introduced and used to predict the isosteric heat of adsorption, such as the number of functional groups, dipole moment of the adsorbed gas, boiling temperature of the adsorbed gas and the mean curvature of the pore.[144, 145] These descriptors are nice because they can be calculated more quickly than Q_{st}. Another descriptor that has been introduced is the atomic property radial distribution function (AP-RDF), which is tailored for large scale QSPR.[146]

Approximately 58,000 hypothetical MOFs were used to calibrate correlation models for CH_4, N_2, and CO_2 uptake capacities obtained from GCMC simulations. These predictive tools can be found on-line via MOF informatics analysis (MOFIA).[146]

7. Summary and future directions

With the increasing number of MOF structures being generated both computationally and experimentally, high-throughput computational screening techniques are poised to play an important role in the development of new MOFs for particular applications. The structures of existing MOFs can be obtained from the CSD, and new hypothetical structures can be generated computationally using bottom-up or top-down approaches. Structural characterization in an automated and high-throughput fashion has been the area of most development in this burgeoning field, with various software packages readily available. These characterization tools have been used to pre-screen and narrow down the list of materials for more detailed simulations. For some simple classes of molecules, reliable force fields allow for high-throughput simulations with results that have good predictive power. A growing number of studies dealing with the adsorption of methane, hydrogen, and CO_2 have employed high-throughput screening and suggested promising new candidates for gas storage and separations. In addition, these studies have revealed useful structure/property relationships. For molecules such as CO_2, where Coulombic interactions are important, there have been significant efforts to develop methods to calculate MOF framework charges in an efficient, but accurate, manner. Data mining techniques are proving useful for obtaining new insights and understanding from the enormous amount of data generated in high-throughput screening. Table 2 summarizes some of the resources that are readily available on-line.

Table 2. Resources available on-line for high-throughput screening of MOFs

Resource	Web address
Cambridge Structural Database (CSD)	http://www.ccdc.cam.ac.uk/
Reticular Chemistry Structure Resource (RCSR)	http://rcsr.anu.edu.au/
Hypothetical MOF Database	http://hmofs.northwestern.edu/
MOF-5 Analogues and other porous materials	http://www.nanoporousmaterials.org/databases/
Zeo++	http://www.maciejharanczyk.info/Zeopp/
Poreblazer	http://www.see.ed.ac.uk/~lsarkiso/
MOFOMICS	http://helios.princeton.edu/mofomics/
TOPOS	http://www.topos.ssu.samara.ru/

MOFIA	http://titan.chem.uottawa.ca/woolab/MOFIA/

There is a need for continued development of a robust software infrastructure for high-throughput computational screening. For example, obtaining MOF structures from the CSD sounds simple enough, but if structures have high degrees of disorder, or other problems, they are sometimes discarded in high-throughput screening. Finding automated ways to "fix" these structures would allow for more comprehensive screening. For structures generated computationally, there are still open questions about how well the generated structures agree with experimental structures. This has been tested for only a limited number of MOFs, and wider testing is needed. In addition, proposed structures are sometimes energetically minimized after the generation process using a variety of methods, including classical mechanics,[85] semi-empirical methods,[56] and quantum mechanics.[51, 54] Will this always be required, or could improved generation schemes make this unnecessary? Are the structures really better after minimization, especially if a generic force field is used? These are still open questions. Methods for quantifying the diversity[78, 85] of a given set of structures are also needed and could be used to improve the diversity of future databases of hypothetical MOFs.

Continued development of predictive and efficient simulation methods is also needed. This includes both efficient simulation algorithms and accurate and transferable force fields. Already GCMC simulations are being performed on GPUs, as a way of speeding up the simulations.[147] Coarse grained models have been developed for adsorption in some systems and shown to agree well with fully atomistic simulations while offering 2 to 3 orders-of-magnitude acceleration for non-polar and polar adsorbates.[95] This sort of approach seems promising for high-throughput applications. Other methods for accelerating GCMC simulations, such as grand canonical-transition matrix Monte Carlo (GC-TMMC), could also be promising. With this technique, an entire adsorption isotherm can be generated from a single GC-TMMC simulation.[148] This may hold particular promise for adsorbates where adsorption isotherms are time consuming and difficult to equilibrate. Improved force fields will improve the accuracy of simulations and allow screening of materials with open-metal sites and other interesting functional groups. Generating force fields from quantum mechanical results in an automated manner is an active area of research that will greatly benefit high-throughput screening.

Other areas of need include new metrics to rank material performance and better descriptors for QSPR and ways to generate the descriptors automatically rather than relying on intuition about the important properties.[149] Computational methods or descriptors that predict the stability (water, thermal, chemical) of MOFs would be very useful – not only for high-throughput screening.

It is likely that in the near future, we will see new MOFs identified by high-throughput computational screening synthesized and tested experimentally for gas storage and separation applications. Additionally, high-throughput screening techniques are being applied to amorphous structures such as porous polymer networks (PPNs).[150] Already these methods are providing new insights and structure/property relationships that small scale studies simply cannot. Furthermore, high-throughput computational screening can tell us the ultimate performance limits of MOF materials for particular applications.

Acknowledgements

Notes and references

^a Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA email: snurr@northwestern.edu

1. H. Li, M. Eddaoudi, M. O'Keeffe and O. M. Yaghi, *Nature*, 1999, **402**, 276-279.
2. M. O'Keeffe, M. A. Peskov, S. J. Ramsden and O. M. Yaghi, *Accounts of Chemical Research*, 2008, **41**, 1782-1789.
3. G. Ferey, *Chemical Society Reviews*, 2008, **37**, 191-214.
4. S. Horike, S. Shimomura and S. Kitagawa, *Nature Chemistry*, 2009, **1**, 695-704.
5. M. Eddaoudi, D. B. Moler, H. Li, B. Chen, T. M. Reineke, M. O'Keeffe and O. M. Yaghi, *Acc. Chem. Res.*, 2001, **34**, 319-330.
6. B. Arstad, H. Fjellvåg, K. Kongshaug, O. Swang and R. Blom, *Adsorption*, 2008, **14**, 755-762.
7. K. L. Mulfort, O. K. Farha, C. L. Stern, A. A. Sarjeant and J. T. Hupp, *J. Am. Chem. Soc.*, 2009, **131**, 3866-3868.
8. D. Himsl, D. Wallacher and M. Hartmann, *Angew. Chem. Int. Ed.*, 2009, **48**, 4639-4642.
9. C. Volkringer, T. Loiseau, N. Guillou, G. r. Férey, M. Haouas, F. Taulelle, E. Elkaim and N. Stock, *Inorganic Chemistry*, 2010, **49**, 9852-9862.
10. O. K. Farha, I. Eryazici, N. C. Jeong, B. G. Hauser, C. E. Wilmer, A. A. Sarjeant, R. Q. Snurr, S. T. Nguyen, A. Ö. Yazaydın and J. T. Hupp, *Journal of the American Chemical Society*, 2012, **134**, 15016-15021.
11. L. Sarkisov, *Advanced Materials*, 2012, **24**, 3130-3133.
12. R. L. Martin and M. Haranczyk, *Chemical Science*, 2013, **4**, 1781-1785.
13. L. J. Murray, M. Dinca and J. R. Long, *Chemical Society Reviews*, 2009, **38**, 1294-1314.
14. J. Sculley, D. Yuan and H.-C. Zhou, *Energy & Environmental Science*, 2011, **4**, 2721-2735.
15. J.-R. Li, R. J. Kuppler and H.-C. Zhou, *Chemical Society Reviews*, 2009, **38**, 1477-1504.
16. J. An, S. J. Geib and N. L. Rosi, *J. Am. Chem. Soc.*, 2010, **132**, 38.
17. J.-R. Li, J. Sculley and H.-C. Zhou, *Chemical Reviews*, 2011, **112**, 869-932.
18. Y.-S. Bae and R. Q. Snurr, *Angewandte Chemie International Edition*, 2011, **50**, 11586-11596.
19. M. D. Allendorf, C. A. Bauer, R. K. Bhakta and R. J. T. Houk, *Chemical Society Reviews*, 2009, **38**, 1330-1352.
20. L. E. Kreno, K. Leong, O. K. Farha, M. Allendorf, R. P. Van Duyne and J. T. Hupp, *Chemical Reviews*, 2011, **112**, 1105-1125.
21. P. Horcajada, C. Serre, M. Vallet-Regí, M. Sebban, F. Taulelle and G. Férey, *Angewandte Chemie*, 2006, **118**, 6120-6124.

22. J. Della Rocca, D. Liu and W. Lin, *Accounts of Chemical Research*, 2011, **44**, 957-968.
23. M. C. Bernini, D. Fairen-Jimenez, M. Pasinetti, A. J. Ramirez-Pastor and R. Q. Snurr, *Journal of Materials Chemistry B*, 2014, **2**, 766-774.
24. C. A. Kent, B. P. Mehl, L. Ma, J. M. Papanikolas, T. J. Meyer and W. Lin, *Journal of the American Chemical Society*, 2010, **132**, 12767-12769.
25. C. A. Kent, D. Liu, L. Ma, J. M. Papanikolas, T. J. Meyer and W. Lin, *Journal of the American Chemical Society*, 2011, **133**, 12940-12943.
26. C. Y. Lee, O. K. Farha, B. J. Hong, A. A. Sarjeant, S. T. Nguyen and J. T. Hupp, *Journal of the American Chemical Society*, 2011, **133**, 15858-15861.
27. D. Farrusseng, S. Aguado and C. Pinel, *Angewandte Chemie-International Edition*, 2009, **48**, 7502-7513.
28. L. Ma, C. Abney and W. Lin, *Chemical Society Reviews*, 2009, **38**, 1248-1256.
29. J. Lee, O. K. Farha, J. Roberts, K. A. Scheidt, S. T. Nguyen and J. T. Hupp, *Chemical Society Reviews*, 2009, **38**, 1450-1459.
30. O. K. Farha, A. M. Shultz, A. A. Sarjeant, S. T. Nguyen and J. T. Hupp, *Journal of the American Chemical Society*, 2011, **133**, 5652-5655.
31. H. Furukawa, K. E. Cordova, M. O'Keeffe and O. M. Yaghi, *Science*, 2013, **341**.
32. *Materials Genome Initiative*, http://www.whitehouse.gov/mgi, Accessed March, 2014.
33. J. J. de Pablo, B. Jones, C. Lind-Kovacs, V. Ozolins and A. Ramirez, *THE MATERIALS GENOME INITIATIVE THE INTERPLAY OF EXPERIMENT, THEORY AND COMPUTATION June 23, 2013*, 2013.
34. A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, S. Cholia, D. Gunter, D. Skinner, G. Ceder and K. A. Persson, *APL Materials*, 2013, **1**, 011002.
35. F. Allen, *Acta Crystallographica Section B*, 2002, **58**, 380-388.
36. J. Goldsmith, A. G. Wong-Foy, M. J. Cafarella and D. J. Siegel, *Chemistry of Materials*, 2013, **25**, 3373-3382.
37. T. Watanabe and D. S. Sholl, *Langmuir*, 2012, **28**, 14114-14128.
38. C. Mellot Draznieks, J. M. Newsam, A. M. Gorman, C. M. Freeman and G. Férey, *Angewandte Chemie International Edition*, 2000, **39**, 2270-2275.
39. X. Kong, H. Deng, F. Yan, J. Kim, J. A. Swisher, B. Smit, O. M. Yaghi and J. A. Reimer, *Science*, 2013, **341**, 882-885.
40. M. Li, D. Li, M. O'Keeffe and O. M. Yaghi, *Chemical Reviews*, 2013, **114**, 1343-1370.
41. C. Wilmer and R. Snurr, in *Topics in Current Chemistry*, Springer Berlin Heidelberg, 2013, pp. 1-33.
42. C. E. Wilmer, M. Leaf, C. Y. Lee, O. K. Farha, B. G. Hauser, J. T. Hupp and R. Q. Snurr, *Nature Chem.*, 2012, 83-89.
43. S. Amirjalayer, M. Tafipolsky and R. Schmid, *The Journal of Physical Chemistry C*, 2011, **115**, 15133-15139.
44. S. Bureekaew and R. Schmid, *CrystEngComm*, 2013, **15**, 1551-1562.
45. R. Chakrabarty, P. S. Mukherjee and P. J. Stang, *Chemical Reviews*, 2011, **111**, 6810-6918.
46. B. J. Sikora, C. E. Wilmer, M. L. Greenfield and R. Q. Snurr, *Chemical Science*, 2012, **3**, 2217-2223.
47. K. E. Jelfs, E. G. B. Eden, J. L. Culshaw, S. Shakespeare, E. O. Pyzer-Knapp, H. P. G. Thompson, J. Bacsa, G. M. Day, D. J. Adams and A. I. Cooper, *Journal of the American Chemical Society*, 2013, **135**, 9307-9310.
48. M. E. Briggs, K. E. Jelfs, S. Y. Chong, C. Lester, M. Schmidtmann, D. J. Adams and A. I. Cooper, *Crystal Growth & Design*, 2013, **13**, 4993-5000.
49. B. Lukose, A. Kuc, J. Frenzel and T. Heine, *Beilstein journal of nanotechnology*, 2010, **1**, 60-70.
50. B. Lukose, A. Kuc and T. Heine, *J Mol Model*, 2013, **19**, 2143-2148.
51. I. A. Baburin and S. Leoni, *CrystEngComm*, 2010, **12**, 2809-2816.
52. L.-C. Lin, Berger, A. H., Martin, R. L., Kim, J., Swisher, J. A., Jariwala, K., Rycroft, C. H., Bhown, A. S., Deem, M. W., Haranczyk, M., and Smit B., *Nature Materials*, 2012, **11**, 9.
53. H. Hayashi, A. P. Cote, H. Furukawa, M. O'Keeffe and O. M. Yaghi, *Nature Materials*, 2007, **6**, 501.
54. D. W. Lewis, A. R. Ruiz-Salvador, A. Gomez, L. M. Rodriguez-Albelo, F.-X. Coudert, B. Slater, A. K. Cheetham and C. Mellot-Draznieks, *CrystEngComm*, 2009, **11**, 2272-2276.
55. T. F. Willems, C. H. Rycroft, M. Kazi, J. C. Meza and M. Haranczyk, *Microporous and Mesoporous Materials*, 2012, **149**, 134-141.
56. R. L. Martin and M. Haranczyk, *Crystal Growth & Design*, 2014.
57. O. D. Friedrichs, A. W. M. Dress, D. H. Huson, J. Klinowski and A. L. Mackay, *Nature*, 1999, **400**, 644-647.
58. S. T. Hyde, O. Delgado Friedrichs, S. J. Ramsden and V. Robins, *Solid State Sciences*, 2006, **8**, 740-752.
59. G. L. McColm, W. E. Clark, M. Eddaoudi, L. Wojtas and M. Zaworotko, *Crystal Growth & Design*, 2011, **11**, 3686-3693.
60. D. J. Tranchemontagne, J. L. Mendoza-Cortes, M. O'Keeffe and O. M. Yaghi, *Chemical Society Reviews*, 2009, **38**, 1257-1283.
61. M. O'Keeffe and O. M. Yaghi, *Chemical Reviews*, 2011, **112**, 675-702.
62. E. Haldoupis, S. Nair and D. S. Sholl, *Journal of the American Chemical Society*, 2010, **132**, 7528-7539.
63. T. Van Heest, S. L. Teich-McGoldrick, J. A. Greathouse, M. D. Allendorf and D. S. Sholl, *The Journal of Physical Chemistry C*, 2012, **116**, 13183-13195.
64. E. M. Sevick, P. A. Monson and J. M. Ottino, *The Journal of Chemical Physics*, 1988, **88**, 1198-1206.
65. A. L. Myers and P. A. Monson, *Langmuir*, 2002, **18**, 10261-10273.
66. T. Düren, F. Millange, G. Ferey, K. S. Walton and R. Q. Snurr, *J. Phys. Chem. C*, 2007, **111**, 15350.
67. T. Düren, Y.-S. Bae and R. Q. Snurr, *Chemical Society Reviews*, 2009, **38**, 1237-1247.
68. Y.-S. Bae, A. O. z. r. Yazaydın and R. Q. Snurr, *Langmuir*, 2010, **26**, 5475-5483.
69. L. D. Gelb and K. E. Gubbins, *Langmuir*, 1998, **15**, 305-308.
70. L. Sarkisov and A. Harrison, *Molecular Simulation*, 2011, **37**, 1248-1257.
71. D. D. Do, L. F. Herrera and H. D. Do, *Journal of Colloid and Interface Science*, 2008, **328**, 110-119.
72. M. Pinheiro, R. L. Martin, C. H. Rycroft, A. Jones, E. Iglesia and M. Haranczyk, *Journal of Molecular Graphics and Modelling*, 2013, **44**, 208-219.
73. L. R. Dodd and D. N. Theodorou, *Molecular Physics*, 1991, **72**, 1313-1345.
74. M. L. Greenfield and D. N. Theodorou, *Macromolecules*, 1993, **26**, 5461-5472.
75. M. Tanemura, T. Ogawa and N. Ogita, *Journal of Computational Physics*, 1983, **51**, 191-207.
76. A. R. Leach, *Molecular Modelling: Principles and Applications, 2nd ed.*, Prentice Hall, Harlow, England, 2001.
77. M. Pinheiro, R. L. Martin, C. H. Rycroft and M. Haranczyk, *Crystal Engineering Communications*, 2013, **15**, 7531-7538.
78. R. L. Martin, B. Smit and M. Haranczyk, *Journal of Chemical Information and Modeling*, 2011, **52**, 308-318.
79. E. L. First and C. A. Floudas, *Microporous and Mesoporous Materials*, 2013, **165**, 32-39.
80. M. Haranczyk and J. A. Sethian, *Journal of Chemical Theory and Computation*, 2010, **6**, 3472-3480.
81. E. V. Alexandrov, V. A. Blatov, A. V. Kochetkov and D. M. Proserpio, *CrystEngComm*, 2011, **13**, 3947-3958.
82. T. R. Cook, Y.-R. Zheng and P. J. Stang, *Chemical Reviews*, 2012, **113**, 734-777.
83. R. L. Martin, Prabhat, D. D. Donofrio, J. A. Sethian and M. Haranczyk, *International Journal of High Performance Computing Applications*, 2012, **26**, 347-357.
84. J. Kim, R. L. Martin, O. Rübel, M. Haranczyk and B. Smit, *Journal of Chemical Theory and Computation*, 2012, **8**, 1684-1693.
85. B. J. Sikora, R. Winnegar, D. M. Proserpio and R. Q. Snurr, *Microporous and Mesoporous Materials*, 2014, **186**, 207-213.
86. R. B. Getman, Y.-S. Bae, C. E. Wilmer and R. Q. Snurr, *Chemical Reviews*, 2011, **112**, 703-723.

87.   S. S. Han, J. L. Mendoza-Cortes and W. A. Goddard, *Chemical Society Reviews*, 2009, **38**, 1460-1476.

88.   D. Dubbeldam and R. Q. Snurr, *Molecular Simulation*, 2007, **33**, 305-325.

89.   S. Keskin, J. Liu, R. B. Rankin, J. K. Johnson and D. S. Sholl, *Industrial & Engineering Chemistry Research*, 2009, **48**, 2355-2371.

90.   R. Q. Snurr, A. O. Yazaydin, D. Dubbeldam and H. Frost, in *Metal-Organic Frameworks: Design and Application*, ed. L. R. MacGillivray, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2010.

91.   H. Fang, H. Demir, P. Kamakoti and D. S. Sholl, *Journal of Materials Chemistry A*, 2014, **2**, 274-291.

92.   L. Sarkisov, R. L. Martin, M. Haranczyk and B. Smit, *Journal of the American Chemical Society*, 2014, **136**, 2228-2231.

93.   S. L. Mayo, B. D. Olafson and W. A. Goddard, *Journal of Physical Chemistry*, 1990, **94**, 8897.

94.   A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. G. III and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024-10035.

95.   K. Yu, J. G. McDaniel and J. R. Schmidt, *The Journal of Chemical Physics*, 2012, **137**, 244102.

96.   T. Pham, K. A. Forrest, P. Nugent, Y. Belmabkhout, R. Luebke, M. Eddaoudi, M. J. Zaworotko and B. Space, *The Journal of Physical Chemistry C*, 2013, **117**, 9340-9354.

97.   J. G. McDaniel, K. Yu and J. R. Schmidt, *The Journal of Physical Chemistry C*, 2011, **116**, 1892-1903.

98.   J. S. Grosch and F. Paesani, *Journal of the American Chemical Society*, 2012, **134**, 4207-4215.

99.   M. G. Martin and J. I. Siepmann, *J. Phys. Chem. B*, 1999, **103**, 4508-4517.

100.   N. L. Rosi, J. Kim, M. Eddaoudi, B. L. Chen, M. O'Keeffe and O. M. Yaghi, *Journal of the American Chemical Society*, 2005, **127**, 1504.

101.   P. D. C. Dietzel, V. Besikiotis and R. Blom, *Journal of Materials Chemistry*, 2009, **19**, 7362-7370.

102.   S. R. Caskey, A. G. Wong-Foy and A. J. Matzger, *Journal of the American Chemical Society*, 2008, **130**, 10870.

103.   P. Canepa, C. A. Arter, E. M. Conwill, D. H. Johnson, B. A. Shoemaker, K. Z. Soliman and T. Thonhauser, *Journal of Materials Chemistry A*, 2013, **1**, 13597-13604.

104.   J. Park, H. Kim, S. S. Han and Y. Jung, *The Journal of Physical Chemistry Letters*, 2012, **3**, 826-829.

105.   H. S. Koh, M. K. Rana, J. Hwang and D. J. Siegel, *Physical Chemistry Chemical Physics*, 2013, **15**, 4573-4581.

106.   D. Yu, A. O. Yazaydin, J. R. Lane, P. D. C. Dietzel and R. Q. Snurr, *Chemical Science*, 2013, **4**, 3544-3556.

107.   Y. Peng, G. Srinivas, C. E. Wilmer, I. Eryazici, R. Q. Snurr, J. T. Hupp, T. Yildirim and O. K. Farha, *Chemical Communications*, 2013, **49**, 2992-2994.

108.   T. Düren, L. Sarkisov, O. M. Yaghi and R. Q. Snurr, *Langmuir*, 2004, **20**, 2683-2689.

109.   M. Eddaoudi, J. Kim, N. Rosi, D. Vodak, J. Wachter, M. O'Keeffe and O. M. Yaghi, *Science*, 2002, **295**, 469.

110.   S. Ma, D. Sun, J. M. Simmons, C. D. Collier, D. Yuan and H.-C. Zhou, *J. Am. Chem. Soc.*, 2008, **130**, 1012-1016.

111.   S. Ma and H.-C. Zhou, *Chemical Communications*, 2010, **46**, 44-53.

112.   S. Q. Ma, D. F. Sun, J. M. Simmons, C. D. Collier, D. Q. Yuan and H. C. Zhou, *Journal of the American Chemical Society*, 2008, **130**, 1012.

113.   Y.-S. Bae and R. Q. Snurr, *Microporous Mesoporous Mater.*, 2010, **132**, 300-303.

114.   A. Blomqvist, C. M. Araujo, P. Srepusharawoot and R. Ahuja, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 20173.

115.   B. Chen, N. W. Ockwig, A. R. Millward, D. S. Contreras and O. M. Yaghi, *Angew. Chem. Int. Ed.*, 2005, **44**, 4745-4749.

116.   D. J. Collins and H. C. Zhou, *Journal of Materials Chemistry*, 2007, **17**, 3154.

117.   M. Dinca, A. Dailly, Y. Liu, C. M. Brown, D. A. Neumann and J. R. Long, *J. Am. Chem. Soc.*, 2006, **128**, 16876-16883.

118.   M. Dinca and J. R. Long, *Angew. Chem. Int. Ed.*, 2008, **47**, 6766-6779.

119.   H. Frost and R. Q. Snurr, *J. Phys. Chem. C*, 2007, **111**, 18794.

120.   S. S. Han, W. Q. Deng and W. A. Goddard, *Angewandte Chemie-International Edition*, 2007, **46**, 6289-6292.

121.   M. P. Suh, H. J. Park, T. K. Prasad and D.-W. Lim, *Chemical Reviews*, 2011, **112**, 782-835.

122.   R. B. Getman, J. H. Miller, K. Wang and R. Q. Snurr, *J. Phys. Chem. C*, 2011, **115**, 2066-2075.

123.   S. K. Brand, Y. J. Colón, R. B. Getman and R. Q. Snurr, *Microporous and Mesoporous Materials*, 2013, **171**, 103-109.

124.   T. Stergiannakos, E. Tylianakis, E. Klontzas and G. E. Froudakis, *The Journal of Physical Chemistry C*, 2010, **114**, 16855-16858.

125.   J. Lan, D. Cao, W. Wang, T. Ben and G. Zhu, *The Journal of Physical Chemistry Letters*, 2010, **1**, 978-981.

126.   T. Ben, C. Pei, D. Zhang, J. Xu, F. Deng, X. Jing and S. Qiu, *Energy & Environmental Science*, 2011, **4**, 3991-3999.

127.   Y. J. Colón, D. Fairen-Jimenez, C. E. Wilmer and R. Q. Snurr, *The Journal of Physical Chemistry C*, 2014, **118**, 5383-5389.

128.   J. Marshall and A. C. Bird, *British Journal of Ophthalmology*, 1979, **63**, 657-668.

129.   S. C. Cullen and E. G. Gross, *Science*, 1951, **113**, 580-582.

130.   A. L. Myers and J. M. Prausnitz, *AIChE Journal*, 1965, **11**, 121.

131.   E. L. First, C. E. Gounaris and C. A. Floudas, *Langmuir*, 2013, **29**, 5599-5608.

132.   Q. Yang, D. Liu, C. Zhong and J.-R. Li, *Chemical Reviews*, 2013, **113**, 8261-8323.

133.   E. Haldoupis, S. Nair and D. S. Sholl, *Journal of the American Chemical Society*, 2012, **134**, 4313-4323.

134.   Q. Xu and C. Zhong, *The Journal of Physical Chemistry C*, 2010, **114**, 5035-5042.

135.   C. Zheng and C. Zhong, *The Journal of Physical Chemistry C*, 2010, **114**, 9945-9951.

136.   C. E. Wilmer, K. C. Kim and R. Q. Snurr, *The Journal of Physical Chemistry Letters*, 2012, **3**, 2506-2511.

137.   E. S. Kadantsev, P. G. Boyd, T. D. Daff and T. K. Woo, *The Journal of Physical Chemistry Letters*, 2013, **4**, 3056-3061.

138.   C. E. Wilmer and R. Q. Snurr, *Chemical Engineering Journal*, 2011, **171**, 775-781.

139.   A. K. Rappé and W. A. Goddard, *J. Phys. Chem.*, 1991, **95**, 3358-3363.

140.   S. Ramachandran, T. G. Lenz, W. M. Skiff and A. K. Rappé, *The Journal of Physical Chemistry*, 1996, **100**, 5898-5907.

141.   C. E. Wilmer, O. K. Farha, Y.-S. Bae, J. T. Hupp and R. Q. Snurr, *Energy & Environmental Science*, 2012, **5**, 9849-9856.

142.   M. Fernandez, T. K. Woo, C. E. Wilmer and R. Q. Snurr, *The Journal of Physical Chemistry C*, 2013, **117**, 7681-7689.

143.   D. Wu, Q. Yang, C. Zhong, D. Liu, H. Huang, W. Zhang and G. Maurin, *Langmuir*, 2012, **28**, 12094-12099.

144.   H. Amrouche, B. Creton, F. Siperstein and C. Nieto-Draghi, *RSC Advances*, 2012, **2**, 6028-6035.

145.   E. J. Garcia, J. Perez-Pellitero, C. Jallut and G. D. Pirngruber, *Physical Chemistry Chemical Physics*, 2013, **15**, 5648-5657.

146.   M. Fernandez, N. R. Trefiak and T. K. Woo, *The Journal of Physical Chemistry C*, 2013, **117**, 14095-14105.

147.   J. Kim and B. Smit, *Journal of Chemical Theory and Computation*, 2012, **8**, 2336-2343.

148.   D. W. Siderius and V. K. Shen, *The Journal of Physical Chemistry C*, 2013, **117**, 5861-5872.

149.   T. Le, V. C. Epa, F. R. Burden and D. A. Winkler, *Chemical Reviews*, 2012, **112**, 2889-2919.

150.   R. L. Martin, C. M. Simon, B. Smit and M. Haranczyk, *Journal of the American Chemical Society*, 2014.