JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

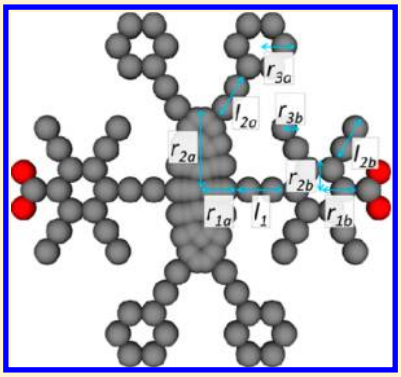
Optimization-Based Design of Metal−Organic Framework Materials

Richard Luis Martin and Maciej Haranczyk*

Computational Research Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 50F-1650, Berkeley, California 94720-8139, United States

Supporting Information

ABSTRACT: Metal−organic frameworks (MOFs) are a class of porous materials constructed from metal or metal oxide building blocks connected by organic linkers. MOFs are highly tunable structures that can in theory be custom designed to meet the specific pore geometry and chemistry required for a given application such as methane storage or carbon capture. However, due to the sheer number of potential materials, identification of optimal MOF structures is a significant challenge. In this contribution we describe an automated technique for MOF design based on mathematical optimization. Optimization is performed on linkers represented by abstract space-filling shapes, in order to generalize the desirable geometric parameters describing linkers, and optimal shapes are projected to real molecules to illustrate realistic MOFs exhibiting the calculated properties. Six examined topologies of MOF and two distinct geometrical pore properties relevant to guest adsorption phenomena are explored. We demonstrate that the optimal shapes of linkers depend on both the topology and the property of interest and moreover that synthetically challenging linkers are not necessary to achieve the most promising candidate materials.



1. INTRODUCTION

Crystalline porous materials have found wide use in industry since the late 1950s. Zeolites, the most well-known class of these materials, are commonly used as chemical catalysts, in particular as cracking catalysts in oil refinement, membranes for separations, and water softeners.[1] Over the past decade, metal−organic frameworks (MOFs) have emerged as a new class of microporous solids exhibiting record surface areas of up to 7000 m^2/g, as well as tunable pore sizes and surface chemistry.[2] MOFs are typically crystalline powders composed of multitopic organic bridging ligands (*linkers*) and metal or metal oxide groups (*secondary building units*, or *SBUs*) in a porous three-dimensional network structure. By adjusting these *building blocks*, it is possible to tune the properties of the material to fulfill its target application requirements. Metal−organic frameworks hold great promise as materials of extraordinary performance in applications related to decreasing the negative environmental impact of energy use such as gas storage,[3] gas separations (as adsorbents or membranes),[4] or catalysis.[5]

However, identification of crystalline porous materials which exhibit the optimal properties for any application remains a serious challenge due to the vast number of possible structures. For example, in the case of all-siliceous zeolites, which contain only one type of building block, the space of possible materials is spanned primarily by variation in topology. Still it has been predicted recently that there are about 330,000 potential all-siliceous zeolite materials with framework energy within 30 kJ of a reference quartz structure.[6] By contrast, MOFs theoretically offer a much greater space of possible structures as organic building blocks may also be varied. Indeed the number of potential building blocks is far larger than the number of feasible topologies, leading to an astronomical number of combinations. Any enumeration-based discovery strategy will clearly be limited to subsets of possible materials. In an enumeration-based design study, Wilmer et al. assembled a database of MOF structures by reusing building blocks of known MOFs.[7] More recently, our group has explored a database of commercially available chemical species as potential MOF building blocks.[8] In general, appropriate selection of sets of building blocks may be critical to the success of screening hypothetical material databases. Successfully identifying promising candidate materials through screening material databases may be considered to depend on the breadth of materials represented therein, which is to say, the *diversity* of the materials represented. A database should be diverse enough to span a large amount of the material space and dense enough to identify optimal materials within families of similar material structures. So far these aspects of materials screening and discovery have not been widely investigated; however, these problems are well-known in the field of pharmaceutical drug discovery,[9] and our group has recently developed similarity/diversity analysis tools for materials discovery.[10] Attempting to satisfy the above criteria for a new material database may lead to a very large data set, posing a computational challenge to a database screening methodology. There is therefore a need for new, alternative design approaches, which can bypass the challenges of large-scale enumeration.

A design strategy based on property optimization with respect to structure, rather than enumeration, is an exciting alternative. In this contribution we introduce and describe our algorithm for optimization-based design of MOF materials. Our

Received: March 28, 2013
Published: May 2, 2013

approach utilizes optimum search algorithms to efficiently navigate the space of possible structures. Here, a number of randomly selected starting structures are modified in an iterative procedure to systematically tune framework properties. In principle, an optimization-based discovery approach avoids computational analysis of large regions of low-performance chemical space, while also removing the specific limitations on the search space imposed by discrete, enumeration-based strategies. This strategy has yet to be tried in the context of porous materials, although attempts have been made in the fields of electronic and optical materials[11−13] as well as molecules.[14] In general, there are two approaches to performing optimization-based design of molecules/materials: 1) perform the search in a *discrete* space of *real* molecular building blocks and employ discrete optimization techniques such as linear programming,[15] simulated annealing,[16] or genetic algorithms;[17] 2) embed the discrete space within a *continuous* space of *abstract* building blocks and employ continuous optimization techniques[12,14] such as steepest-ascent (i.e., gradient-based optimization). The latter approach, based on *alchemical*,[18] steepest-ascent optimization of abstract building blocks, is investigated in our study.

In this contribution we extend our prototype method for design of high gravimetric surface area (GSA) MOFs through optimization[19] and describe a generalized algorithm for optimization-based design of MOFs with any quantifiable geometric property. Rather than enumerate all possible linker molecules and resulting MOFs, of which there are very many, we have devised an approach whereby possible linkers are 1) classified according to their connectivity and branching and 2) represented by a *vector* of geometric properties, which corresponds to a coordinate vector in the chemical space. Utilizing efficient computational methods for MOF assembly and property estimation, optimization can therefore be achieved by modifying linker-defining vectors based on the calculated MOF property gradient, in order to incrementally modify linker geometry and hence iteratively converge to a MOF with optimal properties.

We illustrate the procedure by identifying optimal MOFs with respect to two properties of interest for gas storage applications: a) volumetric surface area and b) volume fraction corresponding to favorable gas-host interaction, defined herein. The optimal molecular geometries obtained by our method for these properties are compared to the optimal GSA linker shapes identified in our initial article.[19] We illustrate that structural design factors correlated with these geometric properties of interest can be readily observed from taking such a high-level view of MOF design and describe how the highest performing materials can be achieved for each of six MOF topologies. We observe that the ranking of topologies with respect to each geometric property is very similar, with the two linker **pcu** net achieving the best performance throughout. Furthermore, we emphasize that this procedure is general in application and can be extended to design MOFs exhibiting other properties of interest, comprising linkers conforming to arbitrary limits of complexity.

2. METHODS

2.1. Overview of Optimization. The purpose of optimization-based design is to automatically discover MOF structures which exhibit an optimal property of interest; for instance, the largest achievable value of a property, or the value closest to a given target, within specified design constraints. Determining a single optimal structure not only enables estimation of the best achievable property within the bounds of the search but, perhaps more importantly, also serves to illustrate which MOF design considerations are critical for high-performing materials. For example, optimization can reveal for which specific MOF topologies branching on organic linker molecules serves to improve performance.

Our method has three main components (see Figure 1). 1) An initial, random structure is obtained by constructing a unit

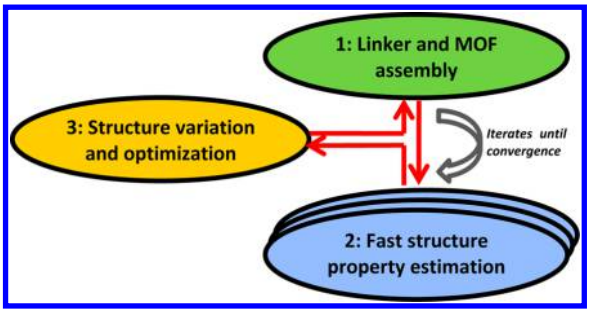


Figure 1. Schematic of the interacting components in MOF optimization.

cell from a specified metal building block and a random linker of the required connectivity. A random linker is constructed by selecting random values for the geometrical parameters which describe each linker molecule, described in **2.2. Linker and MOF Construction and Parameters**. 2) During the iterative optimization algorithm, small modifications are made to each of the linker parameters, and the effect of these modifications upon the *objective function* (e.g., the resulting volumetric surface area of the material) is measured, using algorithms described in **2.3. Property Estimation**. 3) The gradient of the objective function with respect to structure can therefore be determined, and the next linker molecule is obtained by making a larger move in the parameter space along the calculated gradient. This steepest-ascent optimization method terminates once the gradient reaches zero, indicating that a local maximum has been achieved. The optimization method is described in **2.4. Optimization Algorithm**.

2.2. Linker and MOF Construction and Parameters. An efficient abstraction of the space of possible linker molecules is required in order to enable the incremental changes in linker molecule structure necessary for gradient-based optimization of linkers. We note that organic linkers can exhibit many different types of connectivity; for instance, they can simply be linear and two-coordinated (e.g., a dicarboxylic acid as in MOF-5[2a]). or they can exhibit more complex shapes such as a trigonal core, with two coordination sites per arm, leading to a six-coordinated linker (as in the PCN-6X series[20]). Hence, linker molecules are described as being comprised of distinct building units of a particular connectivity, termed *junctions*, and we explore the space of all possible linkers comprised of a central junction and secondary junctions positioned on arms extending from the central junction. Accordingly, a broad range of linker topologies can be achieved, including those of MOF-5 or the PCN-6X series.

Following our initial study,[19] here we have limited our search to a selection of linker connectivity types, each of which, when combined with an appropriate SBU, enables a distinct topology or *net* (i.e., three-dimensional periodic graph, in which atoms or groups thereof are the vertices, and connections between them

are the edges)[21] of MOF to arise. In our approach, MOFs are assembled by iteratively positioning linker and metal building blocks adjacent to one another, fulfilling the requirements of the net; a periodic unit cell is then automatically defined by analyzing the net connections that must be made periodically. We have chosen to design optimal materials within six MOF topologies which have been reported in the literature. Each of the SBUs and nets described below are illustrated in the Supporting Information, and readers are referred to the references cited[21] for additional information on nets and the taxonomy of MOF structures.

A 6-coordinated (6-c) Zn_4O SBU coordinated to linear ditopic carboxylate linkers yields a **pcu** net (e.g., MOF-5[2a]), as does a 6-c Cu_2 *paddle-wheel* with both linear ditopic carboxylate- and nitrogen-terminated linkers (the only topology examined herein which involves more than one linker). A 4-c Cu_2 paddle-wheel (the nitrogen-coordinating sites on the metal assumed to be saturated by water) can produce various nets when combined with distinct carboxylate-terminated linkers: **tbo** when combined with a tritopic linker (e.g., HKUST-1[22]); **ntt** when combined with a hexatopic linker (considered as being comprised of four 3-c vertices, alternatively **rht**[21b] – e.g. the PCN-6X series[20]); **pts** when combined with a tetrahedral tetratopic linker (e.g., MOF-36[23]); **fof** when combined with a rectangular tetratopic linker (considered as being comprised of two 3-c vertices, alternatively **nbo**[21b] – e.g. MOF-505[24]). Furthermore, a single level of interpenetration is explored where permissible by the topology (i.e., **pcu** and **tbo** nets only). Interpenetrated MOFs are achieved by duplicating the underlying material structure and translating it by a fixed fractional coordinate vector, so as to approximate the high symmetry observed experimentally in interpenetrated crystalline materials.[25]

In order to accommodate incremental changes in linker shape during optimization, an abstract space of possible linker molecules of each topology is defined. We abstract real organic linkers using abstract, artificial molecules, which represent possible molecular geometry with an arrangement of space-filling atoms corresponding to junctions, connections between junctions, and branches. For each combination of junction types, nine parameters describe each molecule, and so each molecule is represented in parameter space by a nine-element vector which is utilized in linker optimization. The nine geometric parameters are the distance between the junctions (l_1), and two radii (r_1, r_2), a branch length (l_2), and a branch radius (r_3) for central and outer junctions. This nine-dimensional parameter space is discretized to a 0.1 Å resolution. An example of such a vector is provided in Table 1. The resulting 2-c abstract molecule, comprised of a 2-c central junction and a 2-c outer junction on each arm, is illustrated in Figure 2; abstract molecules of all connectivities based on the same nine parameter values are illustrated in the Supporting Information. We emphasize that optimizing each linker type is considered as an independent optimization problem: linkers are not permitted to transition between topologies during optimization.

Table 1. Example of Nine-Parameter Vector Descriptor for an Abstract Molecule[a]

l_1	r_{1a}	r_{2a}	r_{1b}	r_{2b}	l_{2a}	r_{3a}	l_{2b}	r_{3b}
4.2	1.4	4.2	1.4	1.4	4.2	1.4	2.8	0.0

[a]All values are in Å.

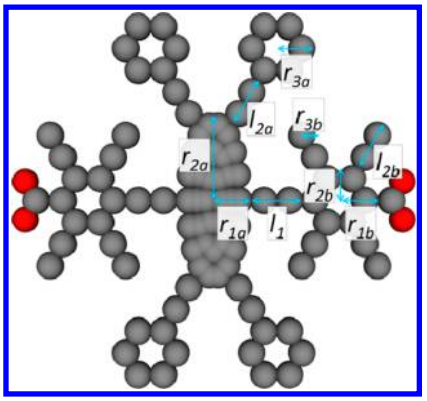


Figure 2. Abstract space-filling molecule constructed from the nine geometric parameters provided in Table 1.

With such an abstraction it is possible to smoothly transition from one molecular shape to another in the space of artificial molecules, and once an optimal structure is achieved for a particular topology, a translation back to real molecules of the discovered optimal geometry can occur. We note that more than one molecule may be mapped to each abstract object. For instance, translation can be performed to achieve molecules exhibiting *backbones* comprised solely of carbon and hydrogen atoms (e.g., Figure 3) or to allow heteroatoms which may permit closer agreement to the geometric shape of the abstract representation.

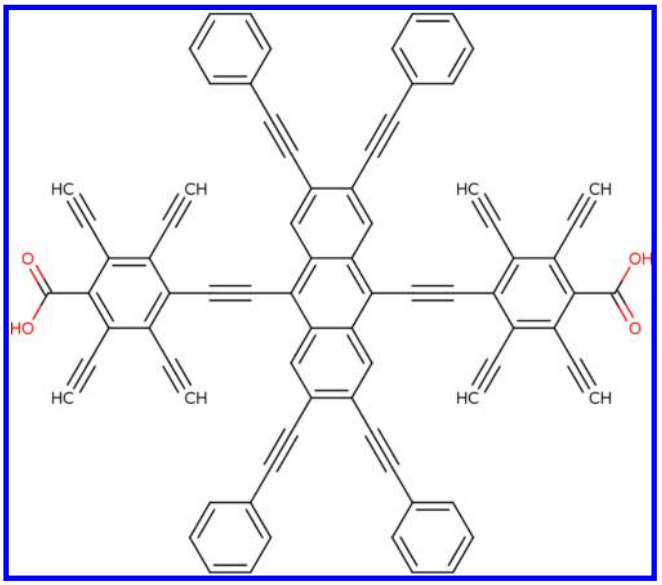


Figure 3. The abstract molecule from Figure 2 projected onto real chemistry, as a dicarboxylic acid molecule exhibiting a backbone comprised of carbon and hydrogen atoms (hydrogen atoms omitted).

It is important, therefore, to ensure that the nine-dimensional geometric parameter space of abstract molecules approximates real chemistry. Accordingly, we examined the eMolecules database[26] of commercially available molecules to determine bounds for junction size, branching length, etc. that are synthetically feasible. Furthermore, based on our previous work,[19] three separate sets of geometric bounds were constructed, representing linkers of increasing synthetic

difficulty. The selected bounds for each set are provided in Table 2.

Table 2. Maximum Length in Å of Each Geometric Parameter, Across Three Sets of Bounds

linker complexity	l_1; l_2	r_{1a}	r_{2b}	r_{1b}; r_{2b}	r_3
set I	5.0	1.4	1.4	1.4	1.4
set II	5.0	2.8	3.5	1.4	1.4
set III	5.0	5.0	5.0	5.0	1.4

2.3. Property Estimation. We have previously described an approach to achieve MOFs optimized with respect to internal, methane-accessible GSA;[19] in this work we generalize this MOF design strategy and demonstrate its application to the discovery of MOF structures optimized with respect to two new objective functions: internal volumetric surface area (VSA) and so-called van der Waals volume fraction (WVF), both of which are calculated with respect to a methane probe molecule, represented using the united atom approximation, i.e. as a single sphere, with radius 1.575 Å.

VSA is a commonly reported material property, defined here as the probe-accessible internal surface area of a material, divided by the total accessible volume, and is provided in the units of m^2/cm^3. On the other hand, WVF is a customized property, designed to illustrate the flexibility of the presented optimization approach to achieve MOFs optimized for any quantifiable property. WVF is defined as the proportion of the probe-accessible volume which is a specific distance from the internal surface of the material and hence is provided as a fraction. Accordingly, the WVF is designed to approximate the proportion of accessible volume corresponding to favorable van der Waals interaction between the guest and the host. The preferential guest−host distance can be approximated by inspecting the energy minimum of an appropriate force field. For the purpose of estimating this distance, we utilize the Lennard-Jones (LJ) interaction model;[27] since linkers are typically carbon-based, the favorable interaction throughout a MOF is approximated by the methane−carbon interaction. The Dreiding force field[28] is utilized for framework carbon atom parameters, while the methane molecule is modeled using the united-atom approximation (i.e., single center) with LJ parameters from the TraPPE force field.[29] In this model, a (favorable) methane−host interaction energy below −15 K is observed in the LJ interaction within the range of 3.75 to 4.6 Å (to the nearest 0.05 Å) from methane center to the framework atom center. This distance range was utilized as representative of the favorable interaction distance range for a guest methane molecule and a framework.

Both VSA and WVF are calculated using our open-source high-throughput porous materials analysis suite *Zeo++*.[30] Algorithms for the calculation of probe-accessible volume and surface area, utilized in the computation of VSA and WVF, are provided in ref 30. VSA and WVF are purely geometric properties and accordingly are computed based on the provided atomic positions and unit cell parameters, with atomic radii as provided by the Cambridge Crystallographic Data Centre.[31]

2.4. Optimization Algorithm. In this work a discretized steepest-ascent optimization method is utilized. The objective function of a material (VSA or WVF), and that of its neighbors in a discretized parameter space, can be used to calculate a numerical gradient of the objective function with respect to linker geometry. If the gradient is zero, the method terminates due to arriving at a local maximum. If the gradient is nonzero however, then the method continues by selecting an appropriate new linker. The new linker is determined by making a move in property space along the calculated gradient and rounding to the nearest discrete position. An arbitrary movement of 2.5 unit steps (i.e., with the specified resolution of 0.1 Å, a movement of 0.25 Å) is performed. In the event that this movement produces a linker outside of the bounds of the property space, iteratively smaller translations are attempted in 0.1 unit step increments until a new linker is successfully achieved, or no movement is possible, in which case the optimization is considered to have converged to a local maximum within the limits of the property space. Finally, the optimization is also terminated if the new linker is identical to one which has been examined previously, to prevent indefinite continuation of the procedure.

It should be emphasized that gradient-based optimization does not guarantee convergence to a global maximum; therefore, in order to mitigate the impact of converging to a local maximum far from the global maximum, ten optimization runs are performed for each topology, initialized with distinct start points. To ensure that the start points are sufficiently distinct from each other, each run begins at a random start point whose parameter values are each distinct from those of any other start point (where permitted by the range of valid geometric parameter values). The structure exhibiting the best calculated property from across the ten runs is selected to represent the best achievable, optimal performance.

For the purposes of optimization, each MOF topology is treated independently; i.e., it is not permitted to transition between MOF topologies during optimization. Similarly, interpenetration of MOF structures, where permissible by the topology in question, is investigated, but transition between interpenetrated and noninterpenetrated structures during optimization is not permitted.

3. RESULTS

For each examined topology of MOF, abstract linkers were optimized in order to identify the maximum VSA and WVF achievable. Interpenetrated and noninterpenetrated cases, as well as junction branching and nonbranching cases, were optimized independently. Three distinct sets of bounds on the space of permissible linkers were investigated, as described above. In total, 156 distinct optimizations were investigated, each of which was performed ten times with a distinct starting structure, with the overall best result selected for each. The single best results for each topology and set of linker bounds are provided in Table 3 and Table 4 and illustrated in Figure 4; full tables of results are provided in the Supporting Information. For comparison, GSA optimization results from our initial article[19] are provided in Table 5 and Figure 5.

Table 3. Highest VSA (in m^2/cm^3) Achieved for Each MOF Topology and Linker Complexity Level

topology	set I	set II	set III
pcu (two linkers)	3911.93	3911.93	3927.89
pcu	3894.79	3894.79	3894.79
fof	2983.87	3027.67	3020.58
tbo	3009.84	3009.84	3009.84
ntt	2536.86	2560.23	2530.38
pts	2671.69	2635.63	2810.23

dx.doi.org/10.1021/ct400255c | J. Chem. Theory Comput. 2013, 9, 2816−2825

Table 4. Highest WVF Achieved for Each MOF Topology and Linker Complexity Level

topology	set I	set II	set III
pcu (two linkers)	0.2421	0.2412	0.2388
pcu	0.2356	0.2409	0.2380
fof	0.1751	0.1766	0.1843
tbo	0.1897	0.1918	0.1890
ntt	0.1411	0.1523	0.1522
pts	0.1739	0.1739	0.1778

Table 5. Highest GSA (in m^2/g) Achieved for Each MOF Topology and Linker Complexity Level

topology	set I	set II	set III
pcu (two linkers)	9673.99	10760.3	11006.5
pcu	9021.72	9978.00	10513.7
fof	5959.05	6318.89	8534.09
tbo	8470.43	9539.93	9842.02
ntt	5355.93	6225.84	8196.53
pts	7229.64	7229.64	9549.47

This approach to navigating the space of MOF structures, utilizing abstract, place-holder molecules and an efficient gradient-based optimization algorithm, enables a high-level view of MOF design considerations. As well as identifying individual optimal MOF structures, MOF topologies can easily be compared, the situational benefit or otherwise of framework interpenetration and linker branching can be elucidated, and the improvement in properties of interest achievable by expanding the space of linkers to include more synthetically challenging molecules can be measured.

It is clear that there is considerable deviation in the highest achievable VSA and WVF across the six distinct MOF topologies explored. In general, the relative performance of the topologies is quite consistent: **pcu** (two linkers) achieves the highest values, closely followed by **pcu**, while **ntt** yields the lowest measured properties. Furthermore, the ranking of MOF topologies is similar to that observed in our previous work identifying optimal GSA frameworks[19] (Figure 5), and indeed a similar relationship is observed between the highest achievable VSA and WVF and the ligand-to-metal ratio, i.e. the ratio of number of ligands (linkers) to number of metal SBUs in a net[19] (Figure 6).

However, the optimal MOFs for high VSA and WVF are significantly different to those exhibiting high GSA. This is implicit from observing that the highest VSA and WVF achievable for each topology remains largely unchanged when linkers are permitted to become more complex (e.g., with larger fused ring systems), in notable contrast to GSA optimization, which yields very large linkers, where permitted.[19] Therefore, complex linkers are shown to be largely unnecessary for achieving high VSA and WVF frameworks. Inspecting the optimal abstract linkers and MOFs achieved for each topology can reveal the reasons for the latter. Optimal linkers for the **pcu** (two linkers) and **ntt** nets are illustrated in Figure 7 and Figure 8; space-filling illustrations of optimal molecules for the remaining topologies are provided in the Supporting Information.

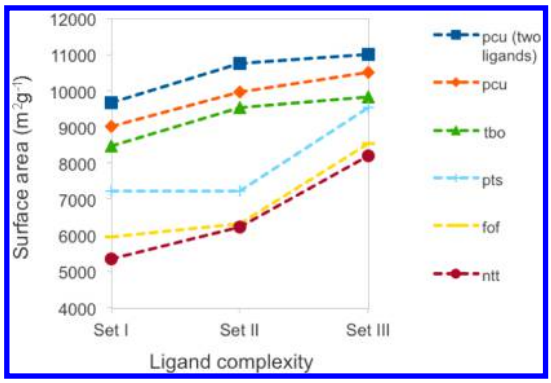


Figure 5. Highest GSA achieved for each topology and linker complexity level.

It is clear that the optimal VSA linkers are geometrically very simple. They are typically small and exhibit little or no branching. This is in stark contrast to optimal GSA linkers, which were typically found to be long and with a high degree of branching.[19] This contrast can be rationalized as follows. GSA (gravimetric surface area) is maximized when a high surface area is achieved with a small mass. Therefore, as few metal SBUs as possible are desired. Increasing the size of the constituent organic linkers serves to improve GSA, and framework interpenetration, which adds SBUs, is a negative agent. By contrast, VSA (volumetric surface area) is maximized when a high surface area is achieved in a small volume. Whether the surface area arises from organic linkers or metal SBUs is irrelevant; the only concern is the optimal occupation of a given volume with framework surface that is exposed to a guest molecule (in this case, methane). Hence, framework interpenetration can be intrinsically beneficial to VSA. Indeed we observe that optimal MOFs for both **pcu** net versions exhibit

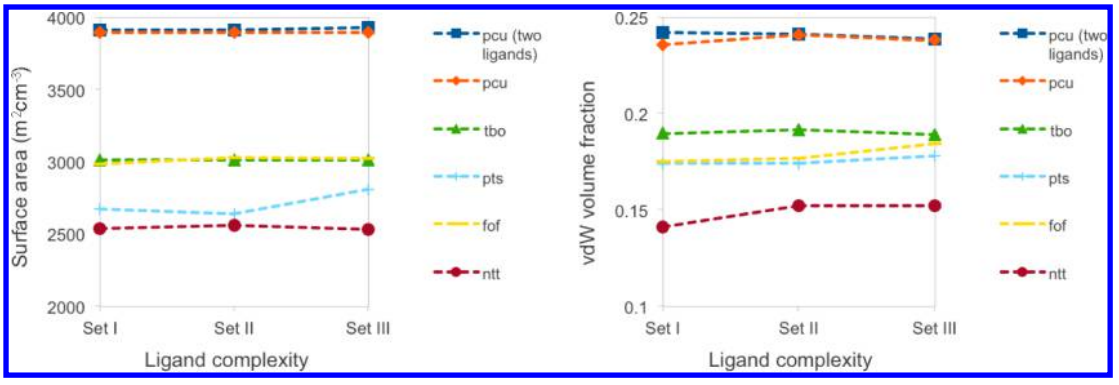


Figure 4. Highest VSA (left, in m^2/cm^3) and WVF (right) achieved for each topology and linker complexity level.

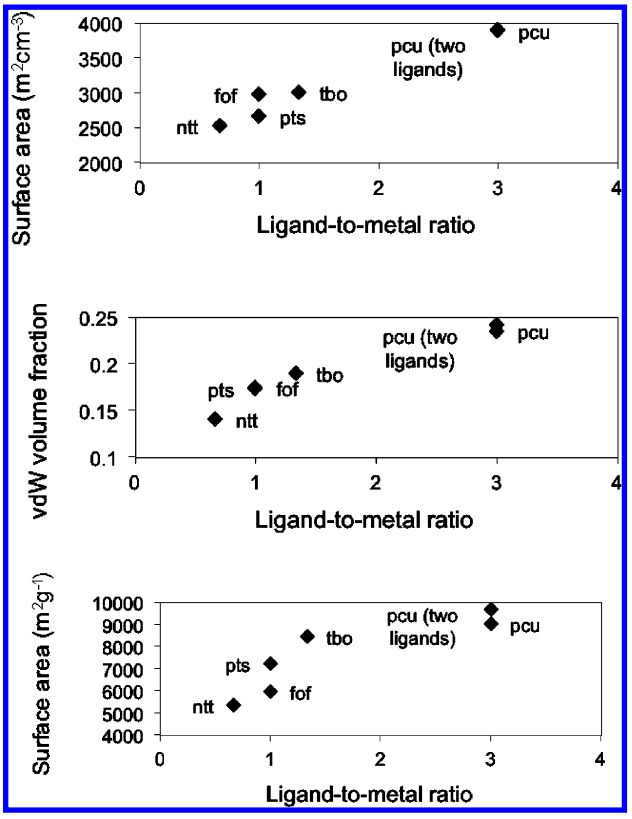


Figure 6. Relationship between the ligand-to-metal ratio and the maximum VSA (above), WVF (center), and GSA (below) achieved for each topology (set I).

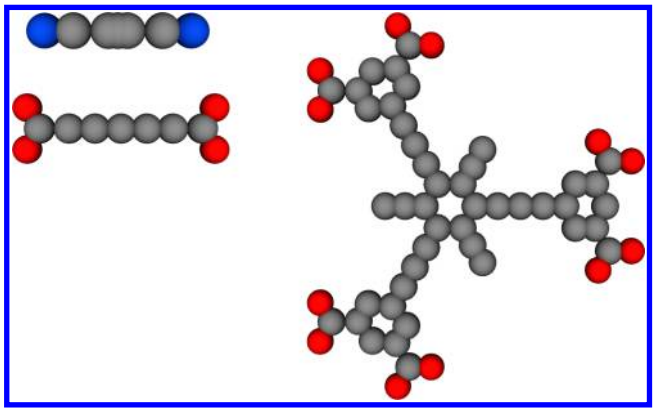


Figure 7. Optimal VSA abstract molecules for pcu (two linkers) net (left) and ntt net (right).

interpenetration. It is interesting that, in general, **tbo** benefits from interpenetration but that the overall best **tbo** MOF is noninterpenetrated; it is clear however from inspecting the optimal shape that it achieves the highest VSA by virtue of being as small in size as possible (hence for this net, the volume can be sufficiently minimized such that a higher VSA is achieved than could be achieved through maximizing surface area). In summary, the optimal linkers will typically be those which are of short length, so as to minimize framework volume. Second, optimal linker shapes will maximize surface area in a manner which is dependent on the net in question, by either a) being comprised of simplistic rodlike shapes if the net interpenetrates (both **pcu** versions and **tbo** nets that involve branching) or b) maximally occupying the available volume

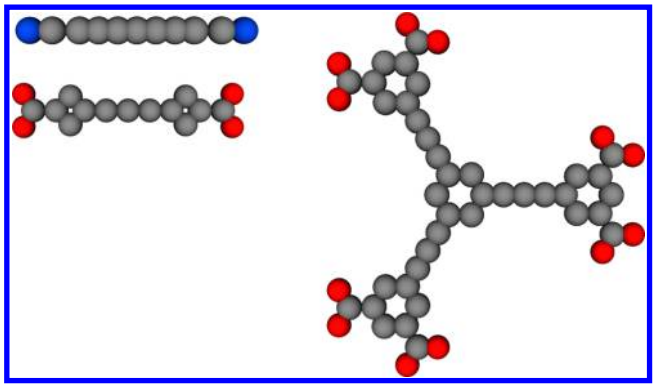


Figure 8. Optimal WVF abstract molecules for pcu (two linkers) net (left) and ntt net (right).

with rings and branching if the net does not interpenetrate (all other nets).

Therefore it is clear that expanding the bounds of linker complexity is in most cases not necessary for improved framework VSA. The notable exception is the **pts** net, which exhibits a nontrivial improvement in VSA when linkers from set III are permitted. For noninterpenetrating nets such as **pts**, case b) above suggests that complex linkers may yield improved VSA if more efficient occupation of volume is thereby achieved. Indeed the optimal pts linker from set III exhibits this property, since it is small in size but exhibits a shape reflecting wide fused ring sections, increasing surface area without a corresponding increase in framework volume (see Figure 9). This effect is not

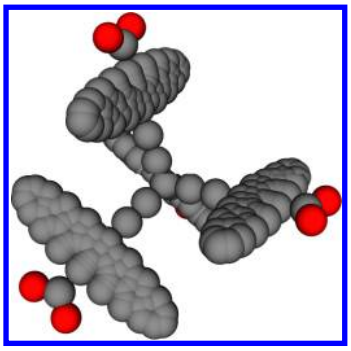


Figure 9. Optimal VSA abstract molecule for pts net (set III).

observed in the other noninterpenetrating nets, **fof** and **ntt**; however, this can also be rationalized by observing that the optimal VSA linker shapes from set I already exhibit branching on the central junction so as to maximize surface area without increasing volume. In these cases, the additional linker breadth permitted in sets II and III is not of benefit, since any further increase in breadth would cause adjacent linkers in the net to be too close to one another, leading to a decrease in surface area.

WVF (van der Waals volume fraction) is, to some extent, a similar framework property to VSA. A high WVF is achieved for MOFs exhibiting a minimized proportion of framework volume which is not adjacent to the surface of the material. Therefore, high WVF can be achieved independently of the amount of framework volume; the only concern is that the volume is occupied efficiently with space that is close to the surface so as to maximize the potential for guest−host interaction. Hence, interpenetration also benefits WVF, and, overall, interpenetrated frameworks are found to be superior in every case (note that the above-mentioned case of the optimal VSA **tbo** net not

dx.doi.org/10.1021/ct400255c | J. Chem. Theory Comput. 2013, 9, 2816−2825

being interpenetrated is not relevant here, since minimizing volume is of no importance).

It is interesting to note that in each case, optimal linkers for achieving high WVF do not include branches. This behavior can be rationalized by considering that although branching on linkers serves to improve surface area, it can reduce the volume fraction which is close to the surface, by a) explicitly removing space by occupying it with atoms of the linker; b) implicitly removing space by causing it to become inaccessible to the guest due to proximity of linker atoms; or c) requiring an overall increase in linker length to accommodate the branches, which causes an increase in total framework volume and hence a reduction in volume fraction near to the surface.

Since WVF represents the fraction of void space within a favorable carbon-methane interaction range, it is an approximation of the true favorable methane interaction volume fraction, since the materials in this work are not entirely comprised of carbon. However, it is presented as a reasonable approximation due to the prevalence of carbon-based linkers in the real molecule projections presented herein and in MOFs, in general, and the small degree of difference in the Dreiding force field[28] between methane interaction with carbon and other common elements such as oxygen and nitrogen. Nevertheless, by the same criteria established above, the most favorable methane−carbon interaction distance range (3.75 to 4.6 Å) is further from the surface than that for methane−oxygen (3.55 to 4.3 Å) or methane−nitrogen (3.7 to 4.35 Å), and so a bias may exist toward structures with fractionally wider pores in order to maximize the void fraction within the most favorable methane−carbon distance range.

It is clear that the value of the framework objective function achievable with abstract molecules does not explicitly reflect the value that can be achieved with real chemistry. Projecting optimal abstract linkers back to real molecules and calculating the VSA or WVF of the resulting frameworks is therefore of critical importance in verifying the observations that have been made in the abstract space. However, since abstract molecules are only geometric shapes, and possess no chemistry, there is therefore a broad scope for assigning real chemistry to the optimal geometry. For certain geometric shapes it may not be possible to exhaustively analyze all feasible chemistry with a similar geometry; furthermore, the abstract geometry may be quite distinct from any realistic chemistry.

We have projected the optimal abstract molecules for each topology and objective function onto real molecules. In the absence of an automated technique known to the authors for evaluating this projection while a) ensuring symmetry of the projected molecules and b) constraining the projection to molecules comprising the desired elements, this evaluation is performed manually for the small number of highest-performing abstract molecules. Accordingly, we do not guarantee that the presented real molecules constitute the closest possible matches to the given geometry. First, linkers of similar geometry to the abstract blueprints but exhibiting only carbon and hydrogen atom backbones were constructed. By restricting linker projection in this way, we achieve the simplest possible chemistry, at the potential expense of geometric proximity to the abstract design. Second, alternative linkers were constructed which were permitted to contain nitrogen atoms within the backbone. This small additional degree of flexibility in linker projection allows for more precise tuning of linker shape to reflect localized features of the identified optimal geometry. For simplicity of the resulting projections, no additional chemical complexity was permitted. Optimal hydrocarbon backbone linkers for the **pcu** (two linkers) and **ntt** nets are illustrated in Figure 10 and Figure 11, with optimal linkers for the remaining

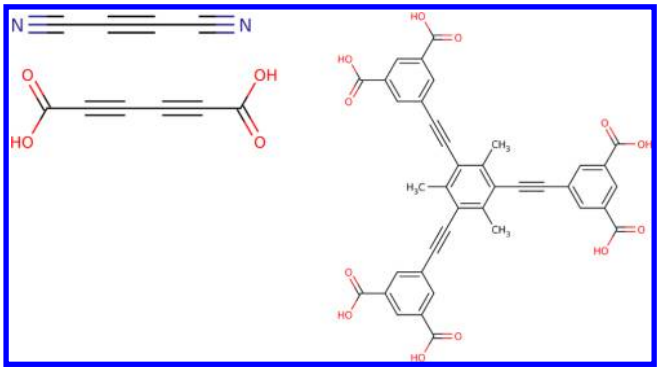


Figure 10. Hydrocarbon-backbone linker projections of optimal VSA abstract molecules for pcu (two linkers) net (left) and ntt net (right); hydrogen atoms omitted.

nets presented in the Supporting Information; in the cases where nitrogen-containing molecules produced MOFs exhibiting a greater VSA or WVF, these molecules are also presented. The overall best VSA or WVF of projected molecules is compared to that of abstract frameworks in Figure 12. We would like to note at this point that the undertaken manual projection approach could be replaced in the future by adapting techniques developed in different contexts, for example, design of structure directing agents.[32]

It can be observed that, in general, projecting from abstract blueprints to real chemistry has a small impact on the calculated VSA or WVF. The maximum deviation in objective function observed between abstract and real frameworks is for the **pcu** net optimized for VSA (projection causing a 15.9% reduction in VSA), followed by the **tbo** net also for VSA (10.9% reduction). In the case of these two frameworks, the poor VSA following projection can be attributed to unrealistic geometry of the abstract molecules arising from their small size and the discrete nature of real chemistry. The abstract **pcu** linker with optimal VSA is simply a rodlike shape, of a particular length that is between one and two acetylenes; the best projection is to a double acetylene, which causes a relative increase in framework volume and hence a reduction in VSA. However, it should be noted that this does not affect the ranking of the topologies and is still a high VSA. The abstract **tbo** linker is of a similar shape but with a trigonal center; no real chemistry of this size exists within the explored limits of projection, and so the closest achievable molecule comprises a ring. In all other cases, projection resulted in less than 10% deviation in objective function.

In certain cases, the incorporation of nitrogen atoms permitted a closer geometric match between real and abstract molecules. For instance, the **pts** net linker optimized for VSA exhibits rings of a challenging geometry. The simplest projection is onto benzene rings; however, the substitution for nitrogen on the two atoms of each ring closest to the center, and the corresponding removal of hydrogen, leads to a closer geometric match to the abstract design. A similar effect is observed for several other nets and enables nitrogen-containing projections to achieve higher objective function values.

Finally, we discuss the behavior of the optimization algorithm, specifically, whether our approach is sufficient to

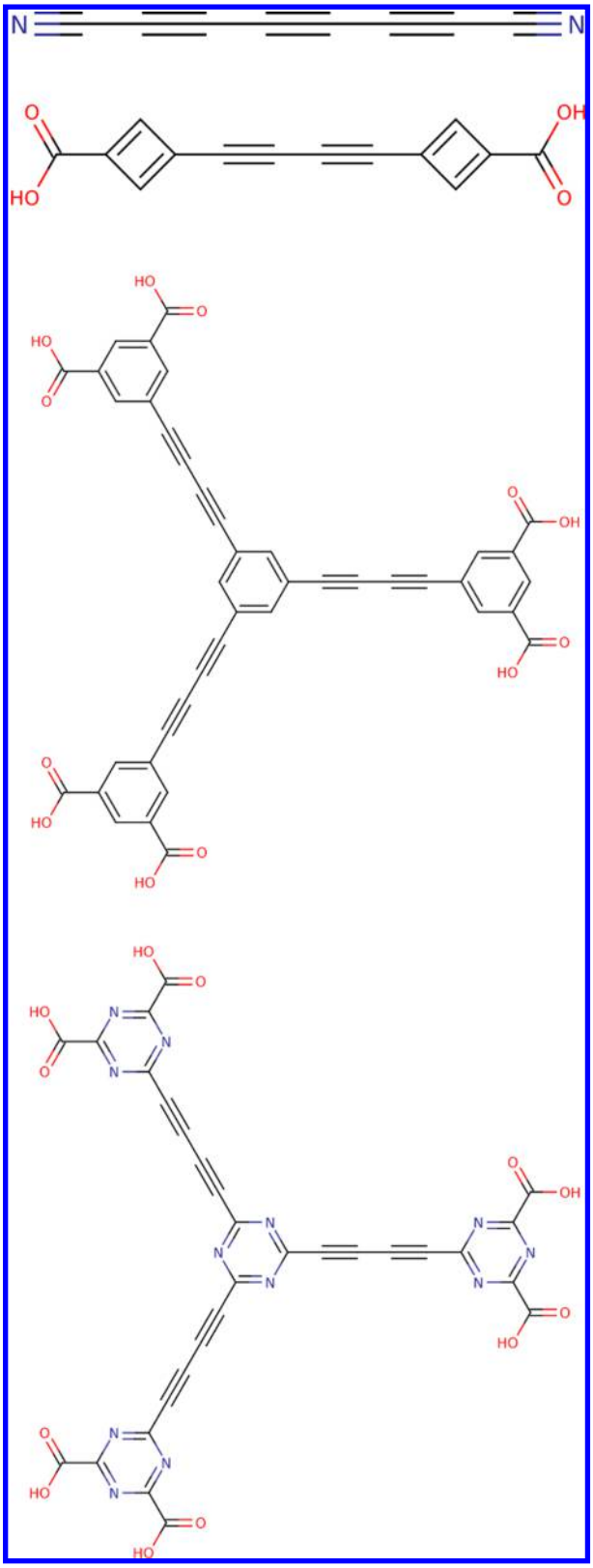


Figure 11. Hydrocarbon-backbone linker projections of optimal WVF abstract molecules for pcu (two linkers) net (above) and ntt net (middle), with superior nitrogen-containing ntt projection (below); hydrogen atoms omitted.

identify a global maximum, and why the maximum achievable objective function is sometimes observed to decrease upon permitting more complex linkers.

As discussed, gradient-based optimization does not guarantee convergence to a global maximum. Based on the position in parameter space at which optimization is initialized, only convergence to a particular local maximum can be guaranteed. By performing optimization for each objective function, topology, etc. ten times, we aim to mitigate this effect and increase the likelihood of identifying a global maximum. In order to estimate the probability of achieving the global maximum after ten runs, we performed a test whereby a much larger number of runs were performed for four distinct optimizations. We selected the **ntt** net, set III (most complex linkers, presenting a larger search space), and performed 100 runs for each objective function, on linkers which a) permitted no branching and b) permitted branching on the central junction only. The 100 objective function values achieved were sorted and are illustrated in Figure 13.

These four optimizations serve to illustrate that the range of VSA or WVF achieved can vary significantly between runs. For example, in the best case, 23 runs optimizing VSA for the **ntt** net with no branching result in 2488.68 m^2 cm^{-3}, the maximum achieved; in the worst case, only one optimization of the same net for WVF with branching achieves the maximum of 0.1536. The probability of achieving a particular maximum within ten distinct runs can therefore be estimated. In the worst case, the probability of the best result out of ten random runs being within five percent of the maximum achieved in 100 runs is 82.5%.

While linker complexity set I is contained within set II, which itself is contained within set III, Table 4 indicates that the maximum WVF achieved for the **pcu** (two linkers) net decreases as linker complexity is permitted to increase. This counterintuitive behavior can be attributed to poorly converged optimization. By permitting increased linker complexity, the search space of possible linkers expands. Accordingly, with a larger search space, it is less likely that ten runs will be sufficient to identify the global maximum. Nevertheless, in these cases the decrease in objective function observed is very small, and the maximum observed in set I can be considered as a lower bound for that of set II, etc.

4. CONCLUSIONS

In summary we have presented a gradient-based algorithm for the design of MOFs with tailored properties using mathematical optimization. Our approach involves abstracting MOF linkers as space-filling abstract shapes, the geometric properties of which are optimized in order to iteratively improve the properties of the resulting frameworks. We have demonstrated that such a high-level view of MOF design enables 1) optimal linker shapes to be identified by efficiently navigating the space of possible MOF geometries; 2) the maximum VSA and WVF of distinct framework topologies to be easily compared; 3) understanding of the situational benefit of linker design considerations such as branching, fused rings, and overall linker size; and 4) estimation of the maximum achievable framework properties within specified bounds of linker exploration.

Furthermore, we have illustrated specific MOF design considerations that are correlated with framework properties of interest. MOFs exhibiting an optimal volumetric surface area (VSA) are achieved through either a) interpenetrated frameworks (where permissible) comprised of small, rodlike linkers, thereby minimizing overall volume, or b) noninterpenetrated frameworks comprised of relatively short linkers which include

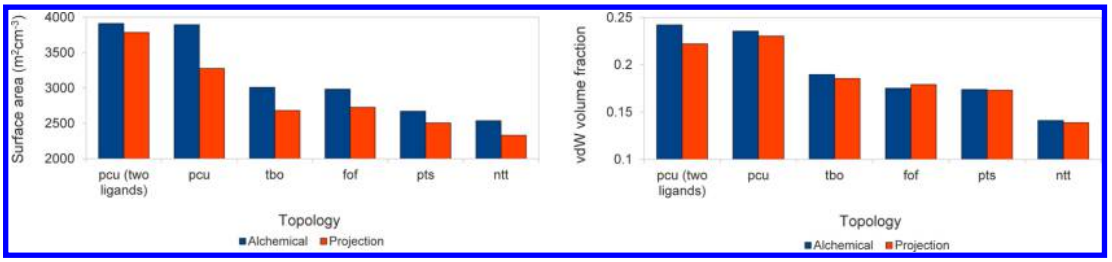


Figure 12. Maximum abstract VSA (left) and WVF (right) compared to projected molecules (set I).

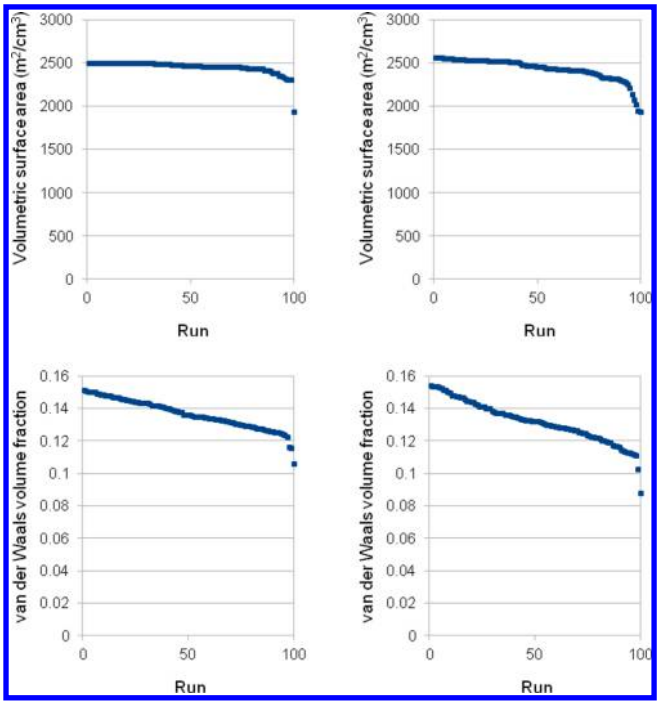


Figure 13. Optimal ntt net objective function values from 100 runs; decreasing order. VSA (above) and WVF (below); no branching (left) and branching on the central junction only (right).

rings and a degree of branching, thereby maximizing surface area. In contrast, an optimal volume fraction close to the internal surface of the material (van der Waals volume fraction, or WVF) is achieved with interpenetrated frameworks (where permissible), and an increase in volume fraction near to the surface through inclusion of rings, but no branches. We note that both versions of the **pcu** net are consistently the highest performing topologies, even in the cases where projecting from abstract to real chemistry causes a decrease in the calculated objective function, and correspondingly the **ntt** net is consistently the lowest performing topology.

Finally, we emphasize that the approach presented here is general in its utility and can readily be deployed to investigate the optimal linkers for any MOF topology, within any arbitrary bounds of linker complexity, and for any material property of interest.

■ ASSOCIATED CONTENT

Supporting Information

Illustrations of metal SBUs, nets, and example abstract linkers are provided in a supporting document. The full results of interpenetrated and noninterpenetrated optimization for each objective function, linker complexity, net and type of linker (junction combination and type of branching) are also provided, along with real molecules projected from the optimal abstract molecules for each objective function and net. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Fax: +1 510 486 5812. E-mail: mharanczyk@lbl.gov.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors wish to thank the reviewer for valuable comments leading to improving the manuscript. This work is supported by the Laboratory Directed Research and Development Program of the Lawrence Berkeley National Laboratory (LBNL). LBNL is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This research used resources of the National Energy Research Scientific Computing Center, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DEAC02-05CH11231.

■ REFERENCES

(1) (a) *Handbook of Zeolite Science and Technology*; Auerbach, S. M., Carrado, K. A., Dutta, P. K., Eds.; Marcel Dekker: New York, USA, 2004. (b) Smit, B.; Maesen, T. L. M. *Nature* **2008**, *457*, 671−677. (c) Smit, B.; Maesen, T. L. M. *Chem. Rev.* **2008**, *108*, 4125−4184. (d) Krishna, R.; van Baten, J. M. *Chem. Eng. J.* **2007**, *133*, 121−131.

(2) (a) Eddaoudi, M.; Kim, J.; Rosi, N.; Vodak, D.; Wachter, J.; O'Keeffe, M.; Yaghi, O. M. *Science* **2002**, *295*, 469. (b) Matsuda, R.; Kitaura, R.; Kitagawa, S.; Kubota, Y.; Belosludov, R. V.; Kobayashi, T. C.; Sakamoto, H.; Chiba, T.; Takata, M.; Kawazoe, Y.; Mita, Y. *Nature* **2005**, *436*, 238. (c) Millward, A. R.; Yaghi, O. M. *J. Am. Chem. Soc.* **2005**, *127*, 17998. (d) Furukawa, H.; Miller, M. A.; Yaghi, O. M. *J. Mater. Chem.* **2007**, *17*, 3197. (e) Morris, R. E.; Wheatley, P. S. *Angew. Chem., Int. Ed.* **2008**, *47*, 4966. (g) Murray, L. J.; Dinca, M.; Long, J. R. *Chem. Soc. Rev.* **2009**, *38*, 1294 and references therein.

(3) Sumida, K.; Hill, M. R.; Horike, S.; Dailly, A.; Long, J. R. *J. Am. Chem. Soc.* **2009**, *131*, 15120.

(4) (a) Demessence, A.; D'Alessandro, D. M.; Foo, M. L.; Long, J. R. *J. Am. Chem. Soc.* **2009**, *131*, 8784. (b) Yazaydin, A. Ö.; Snurr, R. Q.; Park, T.-H.; Koh, K.; Liu, J.; LeVan, M. D.; Benin, A. I.; Jakubczak, P.; Lanuza, M.; Galloway, D. B.; Low, J. J.; Willis, R. R. *J. Am. Chem. Soc.* **2009**, *131*, 18198. (c) Britt, D.; Furukawa, H.; Wang, B.; Glover, T. G.; Yaghi, O. M. *Proc. Natl. Acad. Sci.* **2009**, *106*, 20637.

(5) Lee, J.; Farha, O. K.; Roberts, J.; Scheidt, K. A.; Nguyen, S. T.; Hupp, J. T. *Chem. Soc. Rev.* **2009**, *38*, 1450−1459.

(6) (a) Deem, M. W.; Pophale, R.; Cheeseman, P. A.; Earl, D. J. *J. Phys. Chem. C* **2009**, *113*, 21353−21360. (b) Pophale, R.; Cheeseman, P. A.; Deem, M. W. *Phys. Chem. Chem. Phys.* **2011**, *13*, 12407−12412.

(7) Wilmer, C. E.; Leaf, M.; Lee, C. Y.; Farha, O. K.; Hauser, B. G.; Hupp, J. T.; Snurr, R. Q. *Nat. Chem.* **2012**, *4*, 83−89.

(8) Martin, R. L.; Lin, L.-C.; Jariwala, K.; Smit, B.; Haranczyk, M. *J. Phys. Chem. C* **2013**, DOI: 10.1021/jp401920y.
(9) (a) Snarey, M.; Terrett, N. K.; Willett, P.; Wilton, D. J. *J. Mol. Graphics Modell.* **1997**, *15*, 372−385. (b) Leach, A. R.; Gillet, V. J. *An Introduction to Chemoinformatics*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2003; pp 103−104, 134−136.
(10) (a) Martin, R. L.; Smit, B.; Haranczyk, M. *J. Chem. Inf. Model.* **2012**, *52* (2), 308−318. (b) Martin, R. L.; Willems, T. F.; Lin, L.-C.; Kim, J.; Swisher, J. A.; Smit, B.; Haranczyk, M. *ChemPhysChem* **2012**, *13* (16), 3595−3597. (c) Pinheiro, M.; Martin, R. L.; Rycroft, C. H.; Jones, A.; Iglesia, E.; Haranczyk, M. *J. Mol. Graph. Model* (in press).
(11) (a) Franceschetti, Z.; Zunger, A. *Nature* **1999**, *402*, 60−63. (b) Kim, K.; Graf, P. A.; Jones, W. B. *J. Comput. Phys.* **2005**, *208*, 735−760.
(12) Rinderspacher, B. C.; Andzelm, J.; Rawlett, A.; Dougherty, J.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2009**, *5*, 3321−3329.
(13) O'Boyle, N. M.; Campbell, C. M.; Hutchison, G. R. *J. Phys. Chem. C* **2011**, *115* (32), 16200−16210.
(14) Wang, M.; Hu, X.; Beratan, D. N.; Yang, W. *J. Am. Chem. Soc.* **2006**, *128* (10), 3228−3232.
(15) Dantzig, G. B. *Linear Programming and Extensions*; Princeton Univ. Press: Princeton, NJ, 1963.
(16) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. *Science* **1983**, *220*, 671−680.
(17) Muhlenbein, H.; Gorgesschleuter, M.; Kramer, O. *Parallel Comput.* **1988**, *7*, 65−85.
(18) (a) von Lilienfeld, O. A.; Tuckerman, M. E. *J. Chem. Phys.* **2006**, *125*, 154104. (b) Pérez, A.; von Lilienfeld, O. A. *J. Chem. Theory Comput.* **2011**, *7* (8), 2358−2369.
(19) Martin, R. L.; Haranczyk, M. *Chem. Sci.* **2013**, *4*, 1781−1785.
(20) Yuan, D.; Zhao, D.; Sun, D.; Zhou, H.-C. *Angew. Chem., Int. Ed.* **2010**, *49*, 5357−5361.
(21) (a) O'Keeffe, M.; Peskov, M. A.; Ramsden, S. J.; Yaghi, O. M. *Acc. Chem. Res.* **2008**, *41*, 1782−1789. (b) O'Keeffe, M.; Yaghi, O. M. *Chem. Rev.* **2012**, *112*, 675−702.
(22) Chui, S. S.-Y.; Lo, S. M.-F.; Charmant, J. P. H.; Orpen, A. G.; Williams, I. D. *Science* **1999**, *283*, 1148−1150.
(23) Kim, J.; Chen, B.; Reineke, T. M.; Li, H.; Eddaoudi, M.; Moler, D. B.; O'Keeffe, M.; Yaghi, O. M. *J. Am. Chem. Soc.* **2001**, *123*, 8239−8247.
(24) Chen, B.; Ockwig, N. W.; Millward, A. R.; Contreras, D. S.; Yaghi, O. M. *Angew. Chem., Int. Ed.* **2005**, *44*, 4745−4749.
(25) Uribe-Romo, F. J.; Hunt, J. R.; Furukawa, H.; Klöck, C.; O'Keeffe, M.; Yaghi, O. M. *J. Am. Chem. Soc.* **2009**, *131* (13), 4570−4571.
(26) http://www.emolecules.com (accessed 10/02/12).
(27) LJ interaction model: $U(r_{ij}) = 4\epsilon_{ij}[(\delta_{ij}/r_{ij})^{12}-(\delta_{ij}/r_{ij})^{6}]$.
(28) Mayo, S. L.; Olafson, B. D.; Goddard, W. A. *J. Phys. Chem.* **1990**, *94* (26), 8897−8909.
(29) Martin, M. G.; Siepmann, J. I. *J. Phys. Chem. B* **1998**, *102* (14), 2569−2577.
(30) Willems, T. F.; Rycroft, C. H.; Kazi, M.; Meza, J. C.; Haranczyk, M. *Microporous Mesoporous Mater.* **2012**, *149*, 134−141.
(31) (a) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441−452. (b) Rowland, R. S.; Taylor, R. *J. Phys. Chem.* **1996**, *100*, 7384−7391. (c) Allen, F. H. *Acta Crystallogr., Sect. B* **2002**, *58*, 380−388.
(32) Pophale, R.; Daeyaert, F.; Deem, M. W. *J. Mater. Chem. A* **2013**, DOI: 10.1039/C3TA10626H.