

Chemical Hieroglyphs: Abstract Depiction of Complex Void Space Topology of Nanoporous Materials

Kevin Theisen,[†] Berend Smit,^{†,‡} and Maciej Haranczyk^{*,§}

Department of Chemistry and Department of Chemical Engineering, University of California, Berkeley, California 94720-1462 and Computational Research Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 50F-1650, Berkeley, California 94720-8139

Received November 19, 2009

In general, most porous materials are so complex that structural information cannot be easily observed with 3D visualization tools. To address this problem, we have developed a special abstract 2D representation to depict all important topological features and geometrical parameters. Our approach involves reducing these structures based on symmetry and perceived building blocks to a compressed, graph representation that allows for quick structure analysis, classification, and comparison.

INTRODUCTION

Porous materials such as zeolites and metal organic frameworks contain complex networks of void channels and cages that are exploited in many different industrial applications, such as catalysis or separations.^{1–4} The properties of these materials are determined by the topology of the void space and their chemical composition. Researchers usually must rely on 3D visualization to gain insight into structural features and the corresponding void space dual. For example, they look for geometrical parameters responsible for suitable guest–host interaction such as the diameters of the channels, cages, and windows that compose the structure. Additionally, knowledge of the void space topology is essential for several reasons, including the determination of spaces inaccessible to guest molecules (e.g., before conducting molecular simulations).

Molecular visualization is employed to present important structural features, along with their connectivity and spatial relation, facilitating conceptual understanding. This understanding provides the insight we perceive between a particular chemical structure and its properties and functions. Visualization of large systems, such as proteins or DNA, usually involves abstract representations as opposed to all atom models, such as ribbons^{5,6} or cartoon diagrams,⁵ as they depict essential structural information that would not be clear otherwise. Beyond proteins and DNA, there are few abstract visualization techniques available for chemical structures. For small molecules, software developers continue to improve molecular models to better visualize molecular features such as cyclic and multibranched components.⁷ Currently, there are no commonly available tools to provide structural information in a facilitated form for other complex systems such as porous materials.

The zeolite class of porous materials is probably the most recognized. Zeolites are nanoporous crystalline materials that

have found wide use in industry since the late 1950s, with common applications as chemical catalysts and membranes for separations and water softeners. They are particularly useful as cracking catalysts in oil refinement.^{2,3} There is an increasing interest in utilizing zeolites as membranes or adsorbents for CO₂ capture applications.⁴ In addition to zeolites, other classes of nanoporous materials, such as metal organic frameworks (MOF)^{8,9} and zeolitic imidazolate frameworks (ZIFS),¹⁰ have enormous potential for gas separations or storage.^{11,12} A key aspect for the success of any nanoporous material is that the chemical composition along with pore topology is optimal at the given conditions for a particular application. It should be noted that the number of possible pore topologies is extremely large. There are approximately 190 unique zeolite frameworks known to exist today in more than 1400 zeolite crystals of various chemical composition and different geometrical parameters.¹³ These experimentally known zeolites constitute only a very small fraction of more than 2.7 million structures that are feasible on theoretical grounds,^{14,15} of which between 314 k and 585 k are predicted to be thermodynamically accessible as aluminosilicates, with the remainder potentially accessible via elemental substitution.¹⁶ Databases of similar or greater magnitude can be developed for other nanoporous materials such as MOFs or ZIFs. New computational and cheminformatics techniques are required to characterize, categorize, and screen such large databases. These new techniques have to rely on facilitated-visual or automatic structure analysis approaches in order to process a large number of structures in reasonable time.

Usually, geometrical analysis of a particular structure involves answering a number of questions: (i) what kind of building blocks are within the structure, (ii) how are they connected, and (iii) what are the topological and geometrical parameters that describe each building block and their interconnectivity. Information about building blocks obtained from (i)–(iii) is then used to analyze structural features on a large scale. One can classify structures as being channel-type or cage-type materials depending on (i)–(iii). Also, the same information is required to determine which cages (if

* Corresponding author e-mail: mharanczyk@lbl.gov; fax: (510) 486 5812.

[†] Department of Chemistry, University of California.

[‡] Department of Chemical Engineering, University of California.

[§] Lawrence Berkeley National Laboratory.

any) are accessible to penetrating molecules. The types of molecules observed in a particular application might be too large to enter some of the cages. This information is not only important for practical application but also essential for performing Grand Canonical Monte Carlo simulations,¹⁷ where penetrating molecules are randomly placed in the void space of a porous material. These molecules should not be placed in regions of the void space that are not accessible to guest species by diffusion due to the small sizes of restricting windows. Such regions have to be excluded from consideration. The same level of caution needs to be taken when studying molecular trajectories with molecular dynamics simulations. In this case, one needs to make sure that the starting position of molecules are in the accessible void space. To facilitate structural analysis of the complex embeddings of porous materials, we have developed an abstract visualization approach to depict building blocks, their connections, and important geometrical parameters. Here, the complex 3D structure is represented as a graphlike 2D structure, where colors, sizes, and patterns of nodes and edges encode the most desired information about the void space in a porous material. Since the most distinctive graphical elements of our abstract representation are rings, the representations are named “topological rings” or just T-rings. The concise T-ring representation allows for quick and easy determination of structure type (channel or cage-type structure) as well as identification of building blocks inaccessible to penetrating molecules due to sterical restrictions coming from the small sizes of connecting windows.

METHODOLOGY

To generate a T-ring, our simplified depiction of void space topology, the following steps need to be performed: (i) analysis of the topological structure of a porous material to retrieve information on existing building blocks and their connectivity, (ii) analysis of each building block to determine geometrical attributes, and (iii) combination of the obtained information into a graphical representation. There is an optional (iv) step for pruning and performing postprocessing of the generated plots to further improve the representation and to maximize the efficiency of structural analysis. All of the steps, (i)–(iv) will be described in the subsections below, followed by a brief summary of technical issues and implementation details.

Selection of a Set of Structural Building Blocks.

Generation of T-ring graphs for porous materials begins with the detection of structural building blocks. A building block is a small, repeating fragment of a periodic structure that carries explicit information about that structure’s void space (pore) topology. Examples of building blocks are cages and segments of channels, which are often used by chemists when discussing the structure of porous materials. Such building blocks are composed from rings of the atoms in the embedded structure. The exact definition of a building block depends on the specific algorithm used to detect these building blocks, in particular, specific criteria of systematic structure segmentation. It must be noted that the building blocks we discuss here do not correspond to IUPAC’s definition of composite building blocks (which may include also chains, not only rings) nor do they exactly correspond to the IUPAC definition of pores, which distinguishes cages,

cavities, and channels.¹⁸ IUPAC’s cage and cavity definitions are polyhedral pores whose windows are respectively, too narrow and wide enough, to be penetrated by a guest species larger than water. Both cages and cavities may be valid building blocks according to our definition, and both will be referred to as cages throughout this article. In turn, IUPAC’s channel is defined as an infinitely extended pore large enough to allow guest species to diffuse, and this entity must be segmented to form valid building blocks.

The task of detecting building blocks has to be performed either by hand with 3D visualization, a task that is often hindered by the structure’s sheer complexity, or by the use of a systematic approach encoded in computer programs. In order to demonstrate our visualization techniques, we employ systematic structure segmentation to produce tilings¹⁹ of periodic nets. Tiles can be considered building blocks within our definition. Recently, rules for the determination of the unique tiling (natural tiling) that carries a given three periodic net have been presented by Blatov and co-workers.²⁰ The vertices and edges of such nets correspond to, respectively, the atoms and bonds of a periodic chemical structure.

It should be stressed that our T-ring representation, in principle, is not restricted to the visualization of a natural tiling obtained with the approach of Blatov and co-workers. The latter is used here because of its convenience and demonstrated success²⁴ in processing the large number of zeolite structures that are in our focus. Our approach can be used, in general, to visualize the topology and properties of structural building blocks coming from any other method, including, for instance, visual detection or tessellation of surfaces representing nanostructures.²¹ The latter are particularly useful when considering porous materials built from carbon nanostructures.

The generation of unique (natural) tilings of the periodic nets representing input porous developed by Blatov et al. is implemented in the TOPOS package.²² The analysis starts with the determination of atom connectivity based on interatomic distance. Then, so-called strong rings are detected by analysis of the resulting molecular graph. Strong rings are defined as rings that are not the sum of any number of smaller cycles. They are then used to divide a structure into separate tiles. Determination of a unique natural tiling of a net relies on a number of rules. The rules are (a) the tiling must have the symmetry of the net (be proper); (b) the faces of tiles must be locally strong rings (no tile has one face larger than the rest); (c) a tile does not have nonface strong rings that do not intersect other such rings; (d) when rules (a)–(c) result in multiple tilings because of intersecting strong rings, only the smaller of a pair is used if the intersecting rings are unequal in size and both are rejected if the intersecting rings are equal in size; and (e) rule (b) may be overridden for tiles with waists of all negative-curvature vertices. For a detailed discussion of these rules, as well as general discussion on tiling of periodic nets, its drawbacks and flaws, we refer to ref 20 and references cited therein. Output from TOPOS provides the tiling, tiling connectivity, individual tile coordinates, and ring coordinates with point group symmetry. By interpreting the information given the symmetry of the system, we can calculate how building blocks are arranged in space and, in particular, the geometrical attributes we need to calculate for the T-ring representation.

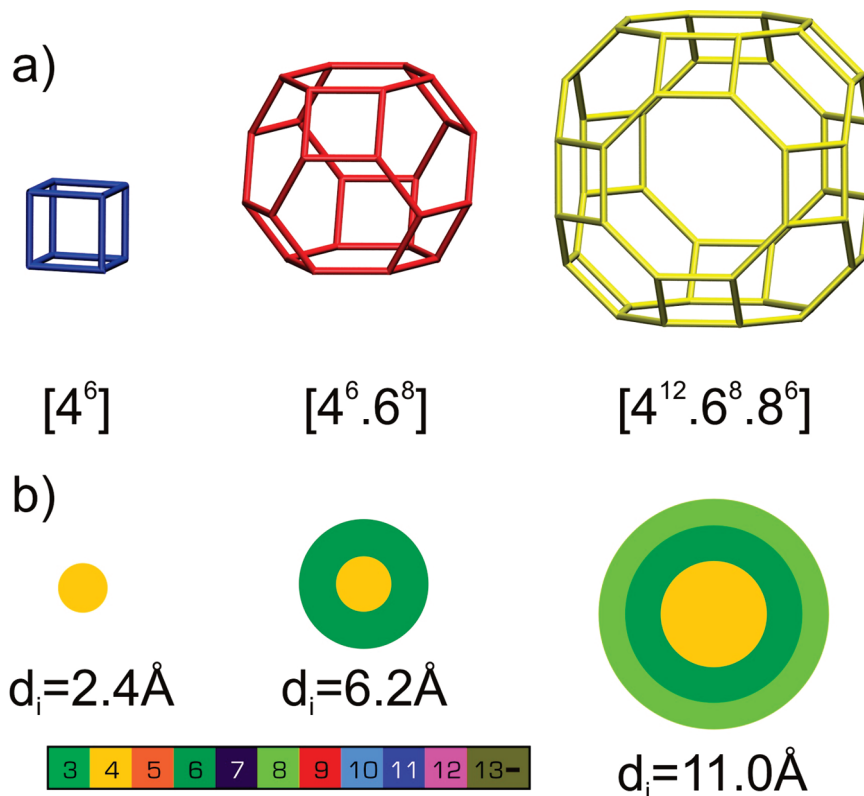


Figure 1. Representations of the building blocks of zeolite LTA in a T-ring graph. (a) Three building blocks of LTA together with the notation used. (b) T-ring nodes: circles representing the building blocks of LTA presented in a color key with diameters proportional to the corresponding maximum included sphere, d_i .

Characterization of Building Blocks. Each building block is composed of rings. The notation $[n^p]$ is used to specify a building block built from p n -member rings. In many cases, building blocks are composed of a number of rings of different sizes and, therefore, each ring type is separated by a dot. For example, a $[n_1^{p_1} \cdot n_2^{p_2}]$ cage consists of p_1 n_1 -member rings and p_2 n_2 -member rings. Rings are built from atoms. Atom coordinates are known and can be used to calculate geometrical parameters. For the application presented in this article, only two parameters are considered: the diameter of the largest sphere included in a block, d_i , and the diameter of a free sphere associated with every ring composing the building block, d_r .²³ These parameters correspond to, respectively, the size of the building block itself and the sizes of the building block's windows. The parameters are approximated and calculated in the following manner. The radius of the largest included sphere is the shortest distance from the centroid of the building block to the van der Waals surface of the closest atom forming the building block. Similarly, the radius of a free sphere associated with a given ring is calculated by determining the shortest distance between the ring's centroid and the van der Waals surface of the closest atom forming the ring. The zeolite structures used in this study are all-siliceous zeolites contained in the IZA database. These zeolite structures have, therefore, only silicon and oxygen atoms. The van der Waals radii for both of these atoms are defined as 1.35 Å, which is the published standard for these atom types in zeolites.²³ It is straightforward to include our types of atoms, with corresponding radii, into the presented framework. We believe that the approximations in the diameters of the largest included and free spheres introduced by our approximation

are reasonable for the purpose of our simplified T-ring representation. In general, as the building block is less convex and spherical, the maximum included sphere may become less accurate. Similarly, as rings deviate from a planar geometry, the free sphere is less likely to be equivalent to the largest sphere at the centroid of the ring; however, the (strong) rings that form the faces of the building blocks are mostly "simple" and fairly planar.

Building T-Rings. A T-ring representation can be generated after calculating the geometrical parameters of a structure's composing building blocks. A T-ring is a 2D graphical representation of a 3D embedding composed of nodes and edges used to describe the building blocks of the structure and their connectivity, respectively. First, representations of cages and segments of channels are depicted (Figure 1). Only one building block of each type is drawn for maximum clarity. Connections between them are then drawn (Figure 2). Please note that Figures 1 and 2 present building blocks of LTA zeolite. We selected this zeolite because of its structural simplicity, which is desired to clearly demonstrate our approach. Building blocks in zeolites do not only correspond to basic Platonics and Archimedean. An interested reader can appreciate their diversity by browsing the IZA database.²⁴

In a T-ring graph, building blocks are represented as circles, with diameter proportional to the corresponding diameter of the largest included sphere (Figure 1). The circle may be composed of a single or multiple colors. The colors correspond to the sizes of the rings that compose the building block and can be interpreted by the numbered color key provided with the T-ring (Figure 1b). The colors in each node are ordered by ascending size from the inside out with each

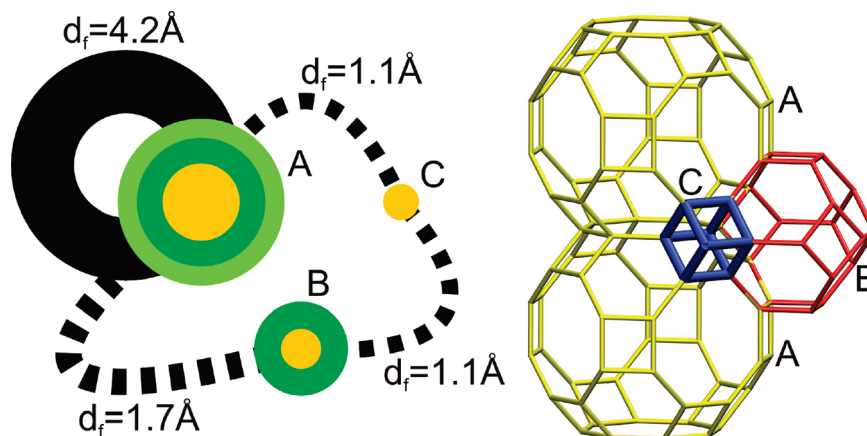


Figure 2. Example T-ring obtained for zeolite LTA. Cages A, B, and C correspond to the $[4^{12} \cdot 6^8 \cdot 8^6]$, $[4^6 \cdot 6^8]$, and $[4^6]$ building blocks, respectively, as presented in Figure 1. Each edge of the T-ring is depicted with a diameter proportional to the diameter of the maximum free sphere corresponding to the connecting window, d_f .

color's width weighted proportionally to the corresponding ring composition of the building block. For instance, a node representing a building block composed of only four-membered rings will be entirely yellow, while a node representing a building block composed of an equal number of four-membered and six-membered rings will have half of its diameter yellow and the other half dark-green. Example cages for the zeolite LTA are presented in Figure 1. Cage $[4^6]$ is composed of only four-member rings and is, therefore, represented by one yellow circle. Cage $[4^6 \cdot 6^8]$ is composed of two types of rings, four-membered and six-membered rings, and therefore is represented by an inner yellow circle and dark green outer circle. The largest of LTA's cages, $[4^{12} \cdot 6^8 \cdot 8^6]$, is composed of three types of rings represented by three circles of light green, dark green, and yellow.

The connectivity between building blocks is also represented in T-rings (Figure 2). If two building blocks share a common ring (building block face) in the chemical system, an edge between the corresponding nodes is depicted on the T-ring graph. The edge is drawn as a black line with a width proportional to the diameter of the largest free sphere that may diffuse between the two building blocks. The length and shape of the edges are not significant. A cutoff diameter may be specified, where all edges less than that diameter will be dashed, to help clarify a T-ring for visual analysis, e.g., these building block connections will not be suitable for diffusion of a probe with diameter larger than the specified threshold. The default threshold diameter is set to 2.8 Å, which is based on a hard sphere often used to represent a water molecule.²³ This threshold parameter can be readjusted to correspond to other molecules, e.g., benzene. In this way, it is straightforward to visually compare the topology of a given porous material accessible to different guest species.

When the nodes and edges of a T-ring are explained, the entire graph may be understood (Figure 2). The layout of the nodes is not significant and no topography is implied. To achieve a clear presentation with minimal oversight, we use the following rules for the initial depiction of a T-ring. (1) The nodes are ordered in descending order based on the estimated maximum included sphere size. (2) The edges are ordered in descending order based on the estimated maximum free sphere size. (3) Coordinates are set for each of the nodes in the T-ring in a radial manner. If there is only

one node, it is placed in the center. If there are two nodes, they are placed in order in the following angles from the center: 180°, and 0°. If there are three nodes, they are placed in order in the following angles from the center: 135°, 270°, and 45°. If there are four nodes, they are placed in order in the following angles from the center: 90°, 180°, 270°, and 0°. If there are five or more nodes they are placed in order based on the angle calculated by the equation, $180 - (360/t) \times i$ where t is the number of tiles, and i is the index of the tile in the sorted order starting at index 0. (4) The centroid of all the nodes is calculated. (5) The edges are drawn. For an edge connecting a node to itself, a circle is drawn from the center of the node pointing away from the centroid of all nodes with radius based on the equation $55/(1 + i)$ pixels where i is the number of self-linking edges that have already been painted for that node. For edges between two different nodes, a quadratic curve is drawn from the larger to the smaller, bending away from the centroid of all nodes. The control point for the quadratic curve is placed on the perpendicular bisector of the line connecting the two nodes. The distance of the control point from the center of the line connecting the two nodes is based on the equation $(100 \times t - 2)/i$ where t is the number of tiles, and i is the number of edges that have already been drawn including the larger node. (6) Nodes are drawn. The inner spaces are weighted proportionally to the ring composition of the tile, such that a measurement of the length of each color from one end to the other end of a node will give the exact proportion of the corresponding ring size ratio of the cage.

Postprocessing. Postprocessing is an optional step and can be performed to optimize the clarity of a T-ring for visual analysis. In some cases, several tiles have the same topology but differ in topography. One tile may be a “squashed” version of another, or two tiles may differ by segment angles. These building blocks are essentially different, but to the human eye, this distinction is not significant. Therefore, a compression function was written to discover such cases and merge them. The tiles are checked in a pairwise manner to see if they contain the same rings. If two are found, then the difference in included sphere size is calculated, where a value of 0.3 Å is used as the threshold ceiling for merging. After merging, all associated edges are redirected to the new merged building block, and pairs of edges are then merged

if ring sizes are the same and their corresponding included sphere diameter difference is less than 0.3 Å.

Implementation and Other Technical Issues. The presented procedure for generating T-ring graphs has been implemented in a software package written in Java.²⁵ The T-ring generator requires two output files from the program TOPOS. The first file is the 3dt input file that is generated by the ADS application in TOPOS. It contains the tiling information along with their connectivity and the corresponding rings that form their facets. The second TOPOS output file contains the coordinates of the centroids for the tiles, which are used for the estimation of the maximum included sphere for each tile. For this reason, the corresponding CIF file of the investigated structure must also be provided. Once T-rings are generated, they are written to the standard PDF format. We have developed both a command-line and graphical user interface (GUI) version of the T-ring generator. The first one is convenient for processing a large number of structures. The GUI can be used to analyze, process, and improve presentations by manipulating T-rings individually. A user can easily generate a T-ring and alter the graphical features, such as scale, node location, and restricted probe diameter and whether or not to display building block tags. The use of this tool eases the creation of optimal T-rings for very complex embeddings with many types of building blocks.

RESULTS

Walk-through Interpretation of a T-Ring. We have selected to use the zeolite LTA in a demonstration of a T-ring representation for its geometrical and topological simplicity. A fragment of the zeolite's embedding and the full T-ring representation of LTA are presented in Figure 2. There are a total of three different types of building blocks and four different types of shared rings. Correspondingly, the T-ring contains three nodes and four edges.

The three nodes A, B, and C correspond to the three building block types $[4^{12} \cdot 6^8 \cdot 8^6]$, $[4^6 \cdot 6^8]$, and $[4^6]$, respectively (Figure 1). Building block A is composed of the three different ring sizes 4, 6, and 8. Building block B is composed of the two different ring sizes 4 and 6. Building block C is composed of only 4-membered rings. The colors for the nodes reflect these ring sizes, and the color key is located at the bottom of Figure 1b. The widths of the colored spaces in the nodes are proportional to the ring distribution of the corresponding cage. For building block A, there are 12 four-membered rings, eight six-membered rings, and six eight-membered rings. Therefore, the yellow part of the node has the largest width; the dark green part has the second largest width, and the light green part has the smallest. The widths are in the ratio 12:8:6.

The four edges in the T-ring correspond to the four types of rings that are shared between the cages (Figure 2). Building block A is connected to another building block A via an eight-membered ring of $d_f = 4.2$ Å. The size of the free sphere corresponding to this window is larger than the set threshold; therefore, it is plotted as a solid line. The other connections present in the LTA zeolite are a 6-membered ring that connects building block A to building block B, a four-membered ring connecting building block B to building block C, and another four-membered ring that connects

building block A to building block C. These rings have a corresponding d_f value of 1.7, 1.1, and 1.1 Å, respectively. The previous three sphere diameters are below the set threshold of 2.8 Å, and therefore, the edges corresponding to those rings are plotted using dashed lines.

Although it may not be obvious from Figure 2, T-rings are plotted conserving the size ratio between nodes and edges. Therefore, estimations for free and included sphere diameters may also be measured from the T-ring. In the output of our software, the scale may be printed in addition to the T-ring. So using a ruler, one may easily estimate specific geometric parameters or use his/her sight to judge a particular structure for a particular engineering task.

Having explained how the structure of LTA is converted into the nodes and edges of the corresponding T-ring, the most important topological characteristics may be easily analyzed. From the presence of only a single solid edge, it is understood that there is only one transport pathway in LTA that is derived of A-type building blocks. This transport pathway is connected by windows much narrower in diameter than the maximum probe size that may reside in building block A. Therefore, LTA can be quickly characterized as a cage-type zeolite, as opposed to a channel-type zeolite for instance. (Note that these types have a different meaning from the IUPAC definition.) LTA has been extensively studied, so this characterization is trivial; however, it proves useful for the millions of theoretical structures not yet investigated. In addition to characterizing a porous material, one can also determine that there are spaces unavailable to diffusing guest species corresponding to building blocks B and C since the only edges connected to them are dashed. These spaces may introduce errors in Monte Carlo simulations (e.g., when applied to predict adsorption) if they are not excluded, so a T-ring provides a quick means to understand which spaces should be removed from Monte Carlo studies. In common practice, these spaces are blocked by inserting dummy atoms in them or by rejecting Monte Carlo moves that place a structure in that volume (e.g., see refs 26 and 27).

Structure Analysis and Classification. T-rings can be used to quickly analyze the structure of a porous material and assign it to the channel, cage, and inaccessible cage-type categories. The channel and cage categories correspond to structures where the void space is arranged in such a way that guest species, respectively, are not and are restricted by windows on their passage. An inaccessible cage type is defined by a passage that is blocked by small windows connecting cages. In most cases, zeolites have a mixture of features of different categories and have to be classified using the dominant one. Although topography is not implied in T-rings, a researcher is presented with a very simple method for selecting the zeolites suited for a given engineering task by quickly scanning over T-rings. We present a number of representative T-rings in Figure 3. T-rings obtained for all zeolites currently stored in the International Zeolite Association (IZA) database are provided in the Supporting Information.

All zeolites presented in Figure 3a are cage type, zeolites having channels with restricting windows. In all three cases, channels are formed by the largest cages $[4^{18} \cdot 6^4 \cdot 12^4]$, $[3^{24} \cdot 8^6 \cdot 12^8]$, and $[4^3 \cdot 5^{12} \cdot 6^1 \cdot 8^3]$ for FAU, RWY, and DDR, respectively. Information about restricting windows is ob-

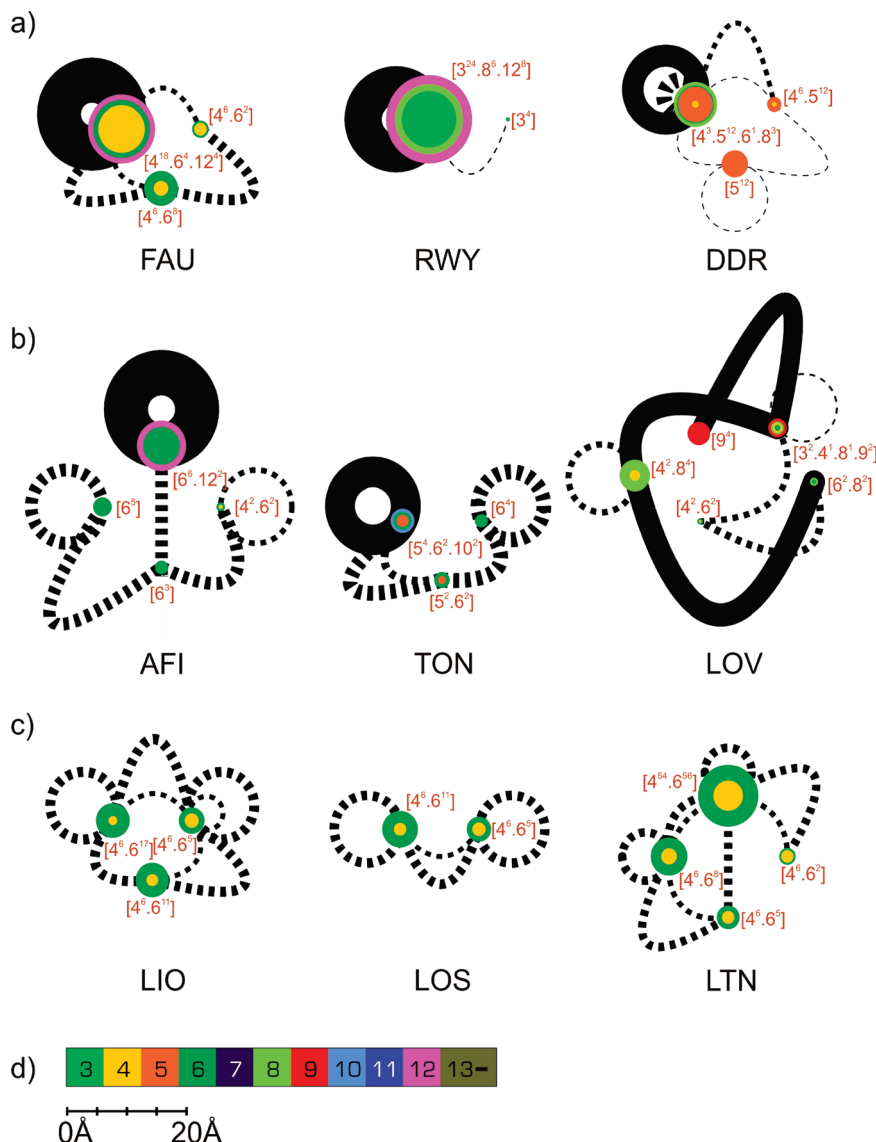


Figure 3. Classification of zeolites based on T-ring representations: (a) cage type, (b) channel type, (c) structures with inaccessible cages, and (d) color key and scale used to generate nodes of T-rings.

tained by comparing the width of these cages with the width of the connected solid black edges, and the latter ones are narrower. Other, smaller cages are not accessible.

All three zeolites presented in Figure 3b are channel-type structures. In this group, the width of the black lines is about the same size as the largest cages presented in the structure. In channel types, just like carbon nanotubes, the channel diameter will fluctuate along the length coordinate, being smaller when the density of atoms is greater in the plane of that coordinate, so the edge width will not be exactly equal to the node width in channel-type T-rings. We define a channel type by having a node with a self-connecting edge or two separate edges that are greater than 85% of the width of the node. Additionally, the TOPOS application generates a tiling where some tiles are small connector fragments that have large openings but small included spheres. These nodes have much smaller diameters than the edges that connect to them.

Another utility is that our T-ring representation offers a means to quickly detect structures that do not allow for practical guest diffusion due to the small size of restricting windows. Examples of such structures are presented in Figure

3c. All connecting windows in the zeolites LIO, LOS, and LTN, have d_i of less than 2.8 Å, and the corresponding edges are plotted as dashed lines. A researcher can quickly go through many T-rings and easily discard these candidates. Additionally, zeolite separations are well implemented by structures that allow for facile diffusion of one species over another. Application of T-rings to this problem will be discussed in the next section.

Case Study: How Molecules “See” Their “Host”. This example illustrates how one can use our software in a practical application. Let us assume that we are interested in the adsorption and diffusion behavior of, say, H_2 , CO_2 , and benzene and suppose we are interested in a material that has a very different behavior for these three molecules. As our software allows for the setting of the restricted diameter cutoff used to determine dashed edges in a T-ring, a researcher may produce several T-rings for a given structure based on the hard sphere estimates of the adsorbents he/she is considering. This may help to determine if one adsorbent will diffuse at all or faster than the others, and if so, it is a good candidate to investigate.

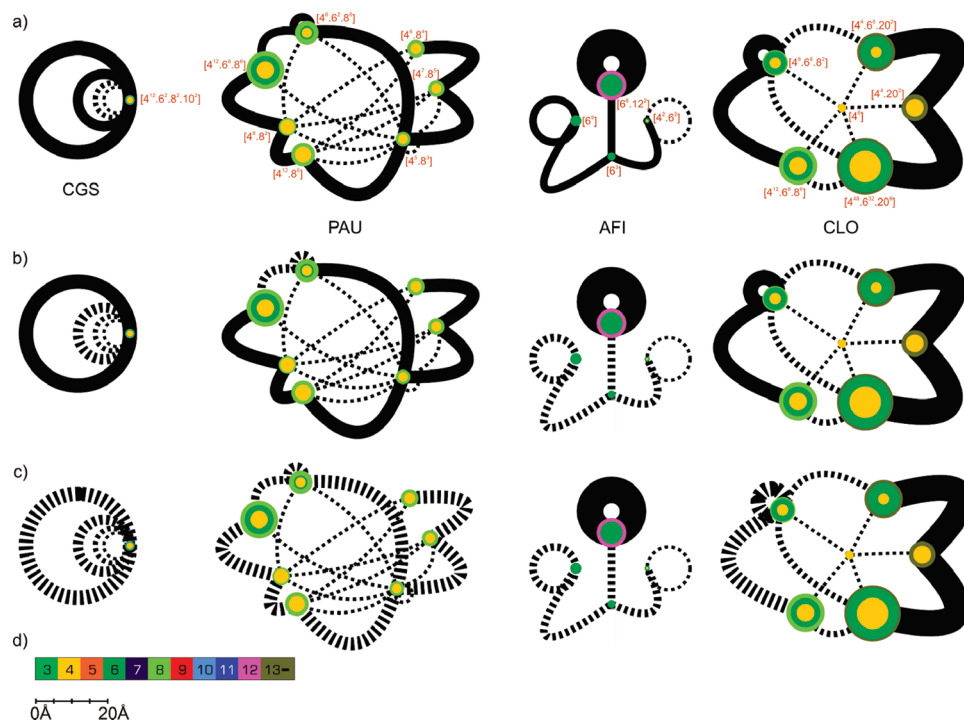


Figure 4. T-ring representations of four zeolites highlighting differences in accessible building blocks. (a,b,c) Correspond to H₂, CO₂, and benzene, respectively; (d) color key and scale used to generate nodes of T-rings.

In Figure 4, four different zeolites are presented using three different restricting diameters for hard sphere diffusion. The diameters chosen are (a) 2 Å to represent a hydrogen molecule, (b) 3 Å to represent a carbon dioxide molecule, and (c) 5.85 Å to represent a benzene molecule.²⁸ The volume accessible to smaller molecules is greater since a smaller size allows for diffusion through small windows. The zeolite CGS is composed of a single building block with several different ring types (Figure 4, left column). Benzene would not diffuse through the small windows, while hydrogen and carbon dioxide can hop through the largest ring. Hydrogen has another option for hopping, which may lead to a selectivity over carbon dioxide. The zeolite PAU (Figure 4, second column from the left side) has a very complex topology, and similarly to CGS, benzene diffusion is not favorable. There are several channel systems available to hydrogen and carbon dioxide. Hydrogen has a few extra options for diffusion trajectories. In zeolite AFI (Figure 4, second column from the right side), the T-ring is identical to both carbon dioxide and benzene, so selectivity between the two cannot be determined from our representations. However, hydrogen has several more pathways to diffuse through, suggesting selectivity over the larger molecules. Zeolite CLO is made up of many large building blocks (Figure 4, right column). The T-rings for hydrogen and carbon dioxide are identical and those guest species have access to a discrete channel system that benzene cannot access. Figure 4 presents how quickly scanning over different probe perspectives for a zeolite allows one to generate simple conclusions about the selectivity to and diffusion of mixtures of guest species.

Other Porous Materials. All examples presented so far are zeolite structures. However, a T-ring can be generated for any porous material such as a metal organic framework (MOF). A T-ring for the MOF, MIL-47,²⁹ is presented in Figure 5. The T-ring depicts a large cavity, [3⁴.8².12²],

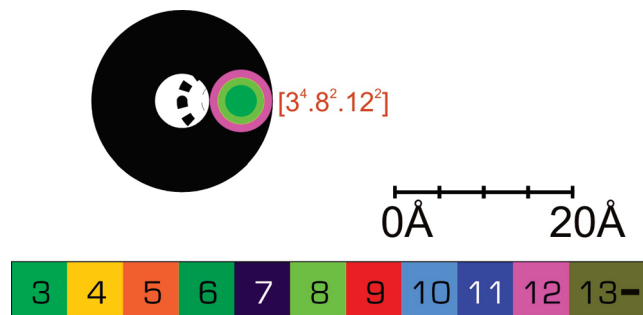


Figure 5. T-ring representation of MIL-47.

which is connected to itself by a similarly large window. Comparisons between MOF structures and zeolites can be made based on this topological analysis, for instance, if one intends to find MOF equivalents for popular zeolite structures.

DISCUSSION

Our T-ring representation provides a simple and fast tool to analyze the void space of a porous material and the related geometrical parameters. T-rings can be used to facilitate screening of structures by quickly removing candidates of no practical use or suggesting potentially suitable ones. As with every tool of this kind, it has some limitations that will be discussed in this section.

The first limitation is related to the fact that free and included spheres are used to approximate guest molecules passing through and contained within a building block. In practice, most molecules of interest, even the simplest solvents or gases, rarely have a spherical shape, and treating them as such may lead to errors. For example, using the length of the molecule as a spherical probe diameter leads to a non-negligible overestimation of the required size of a channels diameter (defined by d_f) in a material considered for diffusion. One has to consider these shortcomings when

analyzing a T-ring before drawing any conclusions. At this moment, it is worth mentioning that we have recently proposed an advanced approach,³⁰ in which a spherical probe is replaced with one resembling the shape and flexibility of a “real” molecule, a complex object built from solid blocks connected by flexible links. Such advanced probes are able to change shape during the traversal of a porous material, reaching areas not accessible to either a single large spherical probe or rigid molecular-shaped probes. We presented the problem as an Eikonal equation in configuration space, in which the cost of entering each point in configuration space corresponds to the local geometry. Although this approach can give accurate predictions of accessible building blocks and volumes inside a porous material, its computational cost is much greater. We believe that, ultimately, one would like to use both approaches: T-rings to visually prescreen large sets of materials and, then, a more advanced approach to refine the prescreening results and verify structures before conducting even more computationally intensive molecular simulations.

Another limitation of T-ring representations comes from the fact that T-rings contain only topological information of building blocks and their local geometrical parameters. These limitations are a natural consequence of dimension reduction, which was a goal of developing our simplified representation. In particular, this representation does not provide information about the spatial location of the blocks, their quantity, and location within the periodic unit cell of a material. This limitation has three consequences: (i) it is not possible to classify channels by their dimensionality, nor distinguish zigzag channels; (ii) it is not possible to reconstruct 3D structure given a T-ring graph only; (iii) it is not possible to tell, judging only from a T-ring graph, if the represented void space is accessible to a probe with a diameter larger than the smallest d_f within the T-ring. Although we cannot point out any specific examples within the IZA database that would fail (iii), we can imagine a system where a large void space is surrounded by impenetrable building blocks. In such a case, a corresponding T-ring may look very similar to other T-rings presented within this article. Fortunately, problem (iii) could be easily detected by 3D visualization and further structure inspection, which is probably conducted before performing any further characterization.

Finally, we would like to discuss the influence of input data on the availability and quality of T-ring representations. It is required that building blocks of a given structure are known before a T-ring is generated. The porous materials used as examples throughout this article are structures for which natural tiling can be obtained. In particular, we have used the approach and software of Blatov and co-workers to obtain such tiling. In a general case, however, there may be structures for which tiling cannot be obtained, e.g., highly disordered structures or structures with very large rings (in the aforementioned procedure (strong) rings need to be detected in order to determine a natural tiling). These rare cases will require either manual detection of building blocks or writing dedicated software, and they go beyond the scope of this article. We should also mention here the influence of crystal structure on the obtained T-ring representation. As mentioned in the introduction, there may be a number of zeolite structures that correspond to one framework topology. These structures usually differ by their chemical composition,

and as a result, their geometry varies. These differences in geometries will be reflected in T-ring representations. For example, we recently reported¹³ that for ITE zeolite the diameter of the largest free sphere can vary up to 0.36 Å depending on a selection of a (crystal) structure. Clearly, such a large difference can be noticed in T-rings representations.

CONCLUSIONS

Most porous materials are so complex that structural information cannot be easily observed with 3D visualization tools. To address this problem, we have developed a special abstract representation, where all important topological features and geometrical parameters of these complex structures are depicted in a 2D plot, facilitating their analysis. This approach may be performed on large numbers of structures, allowing for quick structure analysis, classification, and comparison.

In the depiction, the complex 3D structure is represented as a 2D graph, where colors, sizes, and the patterns of nodes and edges encode the most desired information about a void space in a porous material. Since the most distinctive graphical elements of our abstract representation are rings, they are simply named “topological rings” or just T-rings. The concise T-ring representation can allow for quick and easy determination of structure type (channel or cage type structure) as well as the identification of building blocks inaccessible to penetrating molecules due to sterical restrictions coming from the small size of connecting windows.

Our future pursuits include adding more visual cues to the T-rings to depict topographical features. One example is to discover rings that have a geometry optimized for interacting with certain guest species, such as how CO₂ interacts with restricting cage windows to block the diffusion of other guest species.³¹ These special features can be flagged by changing the colors of the corresponding T-ring edges. Therefore, it would be convenient to browse over T-rings and limit ones scope to comparing only T-rings with this tag present. Additionally, several new material informatics descriptors will be generated for porous materials based on these methods, such as %cage/%channel/%inaccessible volume ratios and numerical descriptors that describe the extent of different types of connectivity. Another interesting direction of research is to investigate the possibility of the use of information about structural building blocks to interpret infrared (IR) spectra of zeolites. This idea is inspired by a recent article,³² where it was proposed to analyze an IR spectrum by comparing the computed spectrum of the framework with that of the asymmetric unit that was used to generate this framework. We believe that similar analysis could be performed for secondary building blocks (e.g., computed spectra of secondary building blocks are compared with the spectrum of the framework).

ACKNOWLEDGMENT

We would like to thank Prof. Vladislav Blatov and Prof. Davide Proserpio for provision of the zeolite structures contained in IZA database in a format suitable for the TOPOS package as well as their invaluable help regarding its use. M.H. is a 2008 Seaborg Fellow at Lawrence Berkeley National Laboratory. This research was supported in part (to

M.H.) by the U.S. Department of Energy under contract DE-AC02-05CH11231. This work was also supported in part (to M.H.) jointly by DOE Office of Basic Energy Sciences and the Office of Advanced Scientific Computing Research through SciDAC project #CSNEW918 entitled "Knowledge guided screening tools for identification of porous materials for CO₂ separations". K.T. and B.S. were supported as part of the Center for Gas Separations Relevant to Clean Energy Technologies, an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences under Award Number DE-SC0001015.

Supporting Information Available: T-ring representations for all zeolite structures that are currently stored in the International Zeolite Association (IZA) database. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- (1) Auerbach, S. M.; Carrado, K. A.; Dutta, P. K., Eds. *Handbook of Zeolite Science and Technology*; Marcel Dekker: New York, 2004.
- (2) Smit, B.; Maesen, T. L. M. Towards a molecular understanding of shape selectivity. *Nature* **2008**, *457*, 671–677.
- (3) Smit, B.; Maesen, T. L. M. Molecular Simulations of Zeolites: Adsorption, Diffusion, and Shape Selectivity. *Chem. Rev.* **2008**, *108*, 4125–4184.
- (4) Krishna, R.; van Baten, J. M. Using molecular simulations for screening of zeolites for separation Of CO₂/CH₄ mixtures. *Chem. Eng. J.* **2007**, *133*, 121–131.
- (5) Richardson, J. S. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* **1981**, *34*, 167–218.
- (6) (a) Carson, M.; Bugg, C. Algorithm for ribbon models of proteins. *J. Mol. Graph.* **1986**, *4*, 121–122. (b) Carson, M. Ribbon models of macromolecules. *J. Mol. Graph.* **1987**, *5*, 103–106.
- (7) Cross, S.; Kuttel, M. M.; Stone, J. E.; Gain, J. E. Visualisation of cyclic and multi-branched molecules with VMD. *J. Mol. Graphics Modell.* **2009**, *28*, 131–139.
- (8) Millward, A. R.; Yaghi, O. M. Metal-organic frameworks with exceptionally high capacity for storage of carbon dioxide at room temperature. *J. Am. Chem. Soc.* **2005**, *127*, 17998–17999.
- (9) Walton, K. S.; Millward, A. R.; Dubbeldam, D.; Frost, H.; Low, J. J.; Yaghi, O. M.; Snurr, R. Q. Understanding inflections and steps in carbon dioxide adsorption isotherms in metal-organic frameworks. *J. Am. Chem. Soc.* **2008**, *130*, 406–407.
- (10) Banerjee, R.; Phan, A.; Wang, B.; Knobler, C.; Furukawa, H.; O'Keeffe, M.; Yaghi, O. M. High-throughput synthesis of zeolitic imidazolate frameworks and application to CO₂ capture. *Science* **2008**, *319*, 939–94.
- (11) Sumida, K.; Hill, M. R.; Horike, S.; Dailly, A.; Long, J. R. Synthesis and Hydrogen Storage Properties of Be-12(OH)(12)(1,3,5-benzenetribenzoate)(4). *J. Am. Chem. Soc.* **2009**, *131*, 15120.
- (12) Choi, H. J.; Dinca, M.; Long, J. R. Broadly hysteretic H-2 adsorption in the microporous metal-organic framework Co(1,4-benzenedipyrazolate). *J. Am. Chem. Soc.* **2008**, *130*, 7848.
- (13) Zimmermann, N. E. R.; Haranczyk, M.; Sharma, M.; Liu, B.; Smit, B.; Keil, F. J. Adsorption and Diffusion in Zeolites: The Pitfall of Isotypic Crystal Structures. *Microporous Mesoporous Mater.*, submitted.
- (14) Foster, M. D.; Treacy, M. M. J. <http://www.hypotheticalzeolites.net> (accessed Nov 13, 2009).
- (15) Earl, D. J.; Deem, M. W. Toward a Database of Hypothetical Zeolite Structures. *Ind. Eng. Chem.* **2006**, *45*, 5449–5454.
- (16) Deem, M. W.; Pophale, R.; Cheeseman, P. A.; Earl, D. J. Computational Discovery of New Zeolite-Like Materials. *J. Phys. Chem. C* **2009**, *113*, 21353–21360.
- (17) Frenkel, D.; Smit, B. *Understanding molecular simulations*, 2nd ed.; Academic Press, San Diego, 2002, pp 23–62.
- (18) McCusker, L. B.; Liebau, F.; Engelhardt, G. Nomenclature of structural and compositional characteristics of ordered microporous and mesoporous materials with inorganic hosts (IUPAC Recommendations 2001). *Pure Appl. Chem.* **2001**, *73*, 381–394.
- (19) Delgado Friedrichs, O.; Dress, A. W. M.; Huson, D. H.; Klinowski, J.; Mackay, A. L. Systematic enumeration of crystalline networks. *Nature* **1999**, *400*, 644–647.
- (20) Blatov, V. A.; Delgado-Friedrichs, O.; O'Keeffe, M.; Proserpio, D. M. Three-periodic nets and tilings: natural tilings for nets. *Acta Crystallogr.* **2007**, *A63*, 418–425.
- (21) Diudea, M. V. Nanoporous carbon allotropes by septupling map operations. *J. Chem. Inf. Model.* **2005**, *45*, 1002–1009.
- (22) Blatov, V. A. *IUCr CompComm Newsletter* **2006**, *7*, 4–38, <http://www.topos.ssu.samara.ru/> (accessed Nov 13, 2009).
- (23) Foster, M. D.; Rivin, I.; Treacy, M. M. J.; Delgado, O. A geometric solution to the largest-free-sphere problem in zeolite frameworks. *Microporous Mesoporous Mater.* **2006**, *90*, 32–38.
- (24) <http://www.iza-online.org/> (accessed Feb 13, 2010).
- (25) The code is available from the authors upon request.
- (26) Dubbeldam, D.; Smit, B. Computer simulation of incommensurate diffusion in zeolites: Understanding window effects. *J. Phys. Chem. B.* **2003**, *107*, 12138.
- (27) Bates, S. P.; van Well, W. J. M.; van Santen, R. A.; Smit, B. Energetics of n-alkanes in zeolites: A configurational-bias Monte Carlo investigation into pore size dependence. *J. Am. Chem. Soc.* **1996**, *118*, 6753.
- (28) Baertsch, C. D.; Funke, H. H.; Falconer, J. L.; Noble, R. D. Permeation of aromatic hydrocarbon vapors through Silicalite-Zeolite membranes. *J. Phys. Chem.* **1996**, *100*, 7676–7679.
- (29) Barthelet, K.; Marrot, J.; Riou, D.; Ferey, G. A breathing hybrid organic-inorganic solid with very large pores and high magnetic characteristics. *Angew. Chem., Int. Ed.* **2002**, *41*, 281–284.
- (30) Haranczyk, M.; Sethian, J. A. Navigating molecular worms inside chemical labyrinths. *Proc. Natl. Acad. Sci. U.S.A. (PNAS)* **2009**, *106*, 21472–21477.
- (31) Krishna, R.; van Baten, J. M. Segregation effects in adsorption of CO₂-containing mixtures and their consequences for separation selectivities in cage-type zeolites. *Sep. Purif. Technol.* **2008**, *61*, 414–423.
- (32) Praprotnik, M.; Hočevár, S.; Hodošček, M.; Penca, M.; Janežič, D. New all-atom force field for molecular dynamics simulation of an AlPO₄-34 molecular sieve. *J. Comput. Chem.* **2008**, *29*, 122–129.

CI900451V