Journal of Molecular Graphics and Modelling 44 (2013) 208–219

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Characterization and comparison of pore landscapes in crystalline porous materials

Marielle Pinheiro [a], Richard L. Martin [a], Chris H. Rycroft [a,b], Andrew Jones [c], Enrique Iglesia [c,d], Maciej Haranczyk [a,*]

[a] Computational Research Division, Lawrence Berkeley National Laboratory, One Cyclotron Road, Mail Stop 50F-1650, Berkeley, CA 94720-8139, USA
[b] Department of Mathematics, University of California, Berkeley, CA 94720, USA
[c] Department of Chemical and Biomolecular Engineering, University of California, Berkeley, CA 94720, USA
[d] Chemical Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

ARTICLE INFO

Article history:
Received 6 March 2013
Received in revised form 3 May 2013
Accepted 11 May 2013
Available online 17 June 2013

Keywords:
Porous materials
Pore size distribution
Stochastic rays
Pore shape similarity

ABSTRACT

Crystalline porous materials have many applications, including catalysis and separations. Identifying suitable materials for a given application can be achieved by screening material databases. Such a screening requires automated high-throughput analysis tools that characterize and represent pore landscapes with descriptors, which can be compared using similarity measures in order to select, group and classify materials. Here, we discuss algorithms for the calculation of two types of pore landscape descriptors: pore size distributions and stochastic rays. These descriptors provide histogram representations that encode the geometrical properties of pore landscapes. Their calculation involves the Voronoi decomposition as a technique to map and characterize accessible void space inside porous materials. Moreover, we demonstrate pore landscape comparisons for materials from the International Zeolite Association (IZA) database of zeolite frameworks, and illustrate how the choice of pore descriptor and similarity measure affects the perspective of material similarity exhibiting a particular emphasis and sensitivity to certain aspects of structures.

1. Introduction

Crystalline porous materials exhibit complex networks of internal void space, which can permit the adsorption and diffusion of guest species. The shape, size and chemistry of localized regions within these networks determine the interactions that can occur between a guest and the host material. Many applications in the chemical industry take advantage of the pore properties of these materials. For example, zeolites are particularly important as cracking catalysts in oil refining [1], and they also find application as alkylation and isomerization catalysts and materials for separations [2–5]. In addition to zeolites, metal organic frameworks (MOFs) [6,7] and zeolitic imidazolate frameworks (ZIFs) [8] have generated interest for their potential use in gas separation and storage [9–11]. The search for materials with better performance is ongoing. Roughly 194 zeolites with 1400 various chemical compositions have been synthesized [12], while millions of new, computationally generated zeolites await investigation [13–16]. Thousands of MOFs have been synthesized in the last decades [17], and large databases of hypothetical structures are being compiled [18].

Databases of materials can, in principle, be screened to identify materials with useful and superior catalytic, storage or separation properties. The current state-of-the-art computational methodology, based on molecular simulations and electronic structure calculations, can be used to accurately predict adsorption and/or catalytic properties of a particular structure and provide reliable information on its performance in specific applications, such as hydrodewaxing [3] and CO_2 separation [19,20]. However, finding the optimal material for a given application remains a formidable challenge because the number of possible materials is extremely large. Fortunately, recent advances in high-performance parallel computing and void space analysis algorithms have enabled simulation of material sets as large as 100,000 structures [18,20]. At the same time, the development of large material databases and high-throughput molecular simulations have brought new challenges related to material data analysis. New, automated computational and cheminformatics techniques need to be developed to characterize, categorize, and search such large databases as well as enable data-mining for useful knowledge.

A key feature of automated material analysis tools is that they convert structural information describing a material, *e.g.* positions

* Corresponding author. Tel.: +1 510 486 7749; fax: +1 510 486 5812.
E-mail address: mharanczyk@lbl.gov (M. Haranczyk).

1093-3263/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmgm.2013.05.007

of atoms in a periodic unit cell, into one or more sets of descriptors. Descriptors can be symbols, letter codes or numbers that capture important features of a material and thus enable structure comparison. Recently, a number of algorithms and software tools have been developed to allow automatic analysis of a material's structure and its void space. For example, Foster et al. and Haldoupis et al. have presented methods to calculate geometrical parameters describing pores in zeolite materials [21,22], namely the diameter of the largest included (D_i) and the largest free (D_f) spheres [23]. The largest included sphere reflects the size of the largest cavity within a porous material, whereas the largest free sphere corresponds to the largest spherical probe that can diffuse fully through a structure and measures a minimum restricting aperture along a diffusion path. Other descriptors include accessible volume (AV) and accessible surface area (ASA) with efficient Monte Carlo algorithms for calculation thereof explored by Düren et al. [24,25] and Do et al. [26].

Another approach to the analysis of porous materials is a fragment-based approach where the basic building units for a material are identified and characterized. For example, Blatov and coworkers have pursued the concept of natural tiling of periodic networks to find and classify primitive building blocks in zeolites [27]. First et al. have developed ZEOMICS and MOFOMICS tools [28] that fragment and represent the void space inside porous materials as a set of geometrical blocks such as cylinders and spheres. Another approach, which combines simple structure representations and local information about the void space, is characterization by pore size distribution (PSD) histograms. PSD histograms indicate the fraction of the void space volume that corresponds to certain pore sizes. They lack, however, information about pore connectivity. Algorithms for PSD and their implementations were recently discussed by Do et al. [26] and Sarkisov et al. [29], and involved computational geometry and grid-based approaches, respectively.

Our group has also contributed algorithms and software for high-throughput geometry-based analysis of materials and their void space [30–32]. The core of our approaches is a computational geometry technique, the Voronoi decomposition [33]. In the Voronoi decomposition, the space surrounding n atoms is divided into n polyhedral cells such that each cell face is a plane equidistant from the two atoms sharing that face. Edges of such cells overlap with lines equidistant to three neighboring atoms, whereas vertices of cells, the Voronoi nodes, are equidistant from four neighboring atoms (in a general asymmetric case). The Voronoi network, a three-dimensional graph comprising such nodes and edges, maps the void space surrounding the atoms. Analysis of such a network can provide parameters such as D_i and D_f and detailed information about void space geometry and topology [32]. For example, void space regions inaccessible to a given probe can be identified, and this information can be utilized in the calculation of ASA and AV [32]. Moreover, the Voronoi decomposition served as a starting point for the development of a novel structure descriptor, which encodes the entire shape of the void space for a given material: the Voronoi hologram [34]. The Voronoi hologram is a histogram representation of the guest-accessible portion of the Voronoi network. Its development allowed us to establish a framework for efficient navigation through the chemical space of all-siliceous zeolites [34]. Our algorithms have been implemented in the Zeo++ package [35] and use the Voro++ software library [36,37]. The Voronoi decomposition has also been used by others in the analysis of crystalline materials [38] and their voids [39], as well as membranes [40] and has been suggested as a tool to investigate ion transport pathways in crystals [41].

In the current contribution, we present new algorithms that allow high-throughput automatic analysis of pore landscapes of porous materials and generation of the corresponding descriptors. In particular, we demonstrate an efficient algorithm for PSD calculation as well as highlight recently developed stochastic ray-tracing histograms [42]. Finally, we illustrate and discuss applications of these and previously developed descriptors in calculating pore-shape similarities, which then can be used in database searching and diverse structure sampling.

2. Methods

2.1. Calculation of the Voronoi network and prediction of void space accessibility

The calculation of the Voronoi network and its analysis to predict void space accessibility have been described in detail in Ref. [32] and will therefore only be briefly introduced here. The Voronoi network is computed using a modified version of Voro++, an open source library for three-dimensional Voronoi calculations [36,37]. The library is based upon individually computing the Voronoi cell associated with each atom, which is stored internally as a collection of edges and vertices. During computation of the Voronoi network, several important parameters are tabulated for use in later analysis. For example, for each edge and node (vertex), the minimum distance to an atom (or its surface) is stored and we denote these as node radii. The library can also handle chemical systems with atoms of various radii by making use of the radical Voronoi tessellation, which is a common approach to handle polydisperse particle arrangements and has been previously shown to be a good method of constructing a network for porosity calculations of unequally sized spheres [43].

The obtained Voronoi network represents the void space in a porous material. Analysis of such a Voronoi network using Dijkstra-like graph algorithms [44] allows the detection of probe-accessible subnetworks. This information can then be used to determine if any point inside a periodic unit cell is accessible to a probe. Here, an accessible point is defined as a point that can be reached by the center of a spherical probe of a given radius. In practice, the Voronoi decomposition is performed on a copy of the atom network in which all atom radii are increased by an amount equivalent to the probe radius. Therefore, any point outside of such expanded particles can accommodate the spherical probe. The procedure for determining accessibility of points is described in detail in Ref. [32], and was employed in Monte Carlo sampling of accessible volume. The same procedure provides the basis for new algorithms presented herein.

2.2. Pore landscape characterization

2.2.1. Pore size distribution

The pore size distribution (PSD) provides information about the fraction of void space that is occupied by pores of certain size. The PSD can be calculated using a Monte Carlo approach similar to accessible volume calculations [32], but with each sampled point described by both its accessibility and pore size. The key features of our algorithm focus on the determination of the largest sphere that encapsulates the sampled point without overlapping any atoms of the structure. In the discussion below, the atoms are treated as having zero radii and the regular Voronoi tessellation is used. The generalization to take into account atom radii is presented later.

Let the atoms in a given structure have positions $\mathbf{x}_i$ for $i = 1, \ldots, n$ and let $d(\mathbf{x},\mathbf{y})$ be the Euclidean distance between two positions $\mathbf{x}$ and $\mathbf{y}$. For a given atom i, the Voronoi cell is defined to be

$$d(\mathbf{x}, \mathbf{x}_i) < d(\mathbf{x}, \mathbf{x}_j)$$

for all j not equal to i. The Voronoi cells form irregular polyhedra that tessellate the space. Here, we consider a general asymmetric case, wherein each Voronoi cell face is equidistant to two atoms, each