Structure
Technical Advance

# MOLE: A Voronoi Diagram-Based Explorer of Molecular Channels, Pores, and Tunnels

Martin Petřek,[1] Pavlína Košinová,[2] Jaroslav Koča,[1] and Michal Otyepka[2,*]
[1]National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno-Bohunice, Czech Republic
[2]Department of Physical Chemistry, Faculty of Science, Palacký University and Center for Biomolecules and Complex Molecular Systems, tr. Svoboby 26, 771 46 Olomouc, Czech Republic
*Correspondence: otyepka@aix.upol.cz
DOI 10.1016/j.str.2007.10.007

## SUMMARY

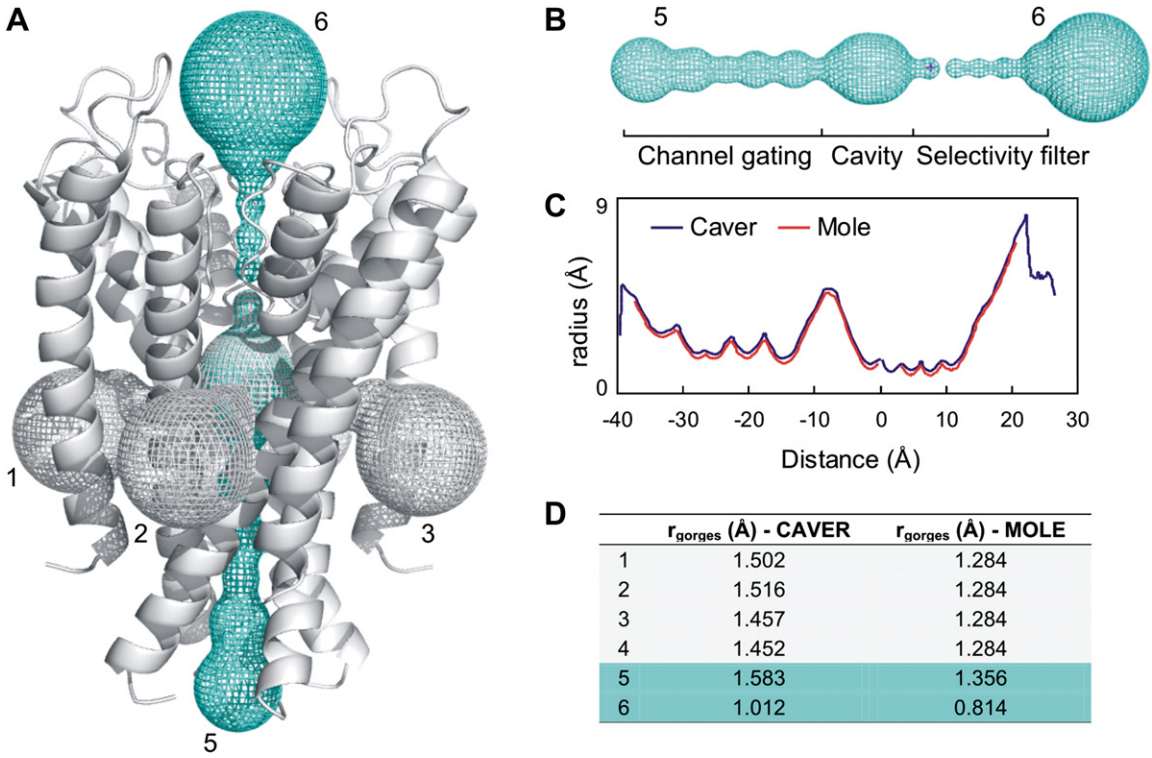
**We have developed an algorithm, "MOLE," for the rapid, fully automated location and characterization of molecular channels, tunnels, and pores. This algorithm has been made freely available on the Internet (http://mole.chemi.muni.cz/) and overcomes many of the shortcomings and limitations of the recently developed CAVER software. The core of our MOLE algorithm is a Dijkstra's path search algorithm, which is applied to a Voronoi mesh. Tests on a wide variety of biomolecular systems including gramicidine, acetylcholinesterase, cytochromes P450, potassium channels, DNA quadruplexes, ribozymes, and the large ribosomal subunit have demonstrated that the MOLE algorithm performs well. MOLE is thus a powerful tool for exploring large molecular channels, complex networks of channels, and molecular dynamics trajectories in which analysis of a large number of snapshots is required.**

## INTRODUCTION

Molecular channels are key features of a huge range of biomolecules. A few examples include: physiologically important pores through cellular membranes provided by transmembrane proteins (Agre et al., 1995; Doyle et al., 1998; Engel et al., 1994, 2000; Gouaux and MacKinnon, 2005; Jiang et al., 2002; MacKinnon, 2002, 2003; Walz et al., 1994, 1997), the peptide exit channels through which ribosomes release newly synthesized proteins during transcription (Voss et al., 2006), and the tunnels (of various types) that connect the active sites of P450 cytochromes to the cytosol (and markedly affect their substrate specificity) (Cojocaru et al., 2007; Otyepka et al., 2007; Wade et al., 2004, 2005). Thus, voids, channels, tunnels, pores, clefts, and cavities are crucial features of systems studied by researchers seeking to gain a fundamental understanding of the relationships between the structures and activities of diverse biomolecules, e.g., those exploring: the accessibility and anatomy of proteins' active sites and other internal spaces (Kleywegt and Jones, 1994; Lesk, 1986; Liang et al., 1998c; Petrek et al., 2006; Smart et al., 1993), the geometry of ribosomal polypeptide exit channels (Voss et al., 2006), the architecture of biomolecular complexes (Kim et al., 2006; Poupon, 2004), and key features of membrane channels. Furthermore, knowledge of the geometry and other structural aspects of various tunnels, channels, and other voids has a huge range of practical applications, inter alia: locating the centerlines in bronchi and the colon (Bitter et al., 2000, 2001; Kaufman et al., 2005; Wan et al., 2002) during virtual bronchoscopy and colonoscopy, respectively; predicting the catalytic efficiency of heterogeneous catalysts such as zeolites, which is largely governed by pore characteristics (Polarz and Smarsly, 2002; Tao et al., 2006); and the remote navigation of autonomous robots (Garber and Lin, 2002).

Due to the importance of various types of molecular voids, it is not surprising that numerous software packages have been developed in recent years to explore and define them. These include: CAST and ALPHASHAPE; programs for identifying and measuring surface-accessible pockets, interior cavities, and other geometrical features of molecules (Edelsbrunner et al., 1998; Liang et al., 1998a, 1998b; Liang et al., 1998c); and POCKET, which finds protein surface pockets by using the concept of protein-solvent-protein events (Levitt and Banaszak, 1992). Others include: LIGSITE (Hendlich et al., 1997) and LIGSITE$^{csc}$ (Huang and Schroeder, 2006), which extend the POCKET algorithm; PASS, which geometrically characterizes regions of buried volume in proteins (Brady and Stouten, 2000); and energy-based programs, such as Q-SITE and POCKET FINDERs, for predicting protein-ligand binding sites and protein pockets (Laurie and Jackson, 2005). Further option include: VOIDOO, for finding cavities and analyzing their volumes (Kleywegt and Jones, 1994); SURFNET, designed to extract all the channels in a given structure (Laskowski, 1995); and HOLE, a program for analyzing and visualizing the dimensions of holes through ion channels (Smart et al., 1993, 1996).

However, finding access paths from the bulk solvent to a buried protein cavity is not a trivial task. Algorithms based on recursive inspection of a protein surface can be designed in principle for this purpose and the problem can be simplified (provided that the structure of the protein




D

| | $r_{gorges}$ (Å) - CAVER | $r_{gorges}$ (Å) - MOLE |
|---|---|---|
| 1 | 1.502 | 1.284 |
| 2 | 1.516 | 1.284 |
| 3 | 1.457 | 1.284 |
| 4 | 1.452 | 1.284 |
| 5 | 1.583 | 1.356 |
| 6 | 1.012 | 0.814 |

**Figure 1. Analysis of the KscA $K^+$ Channel**

(A) Six paths, represented by gray and cyan meshes, found by the MOLE algorithm in the transmembrane KscA $K^+$ channel.

(B) A mesh representing the channel profile.

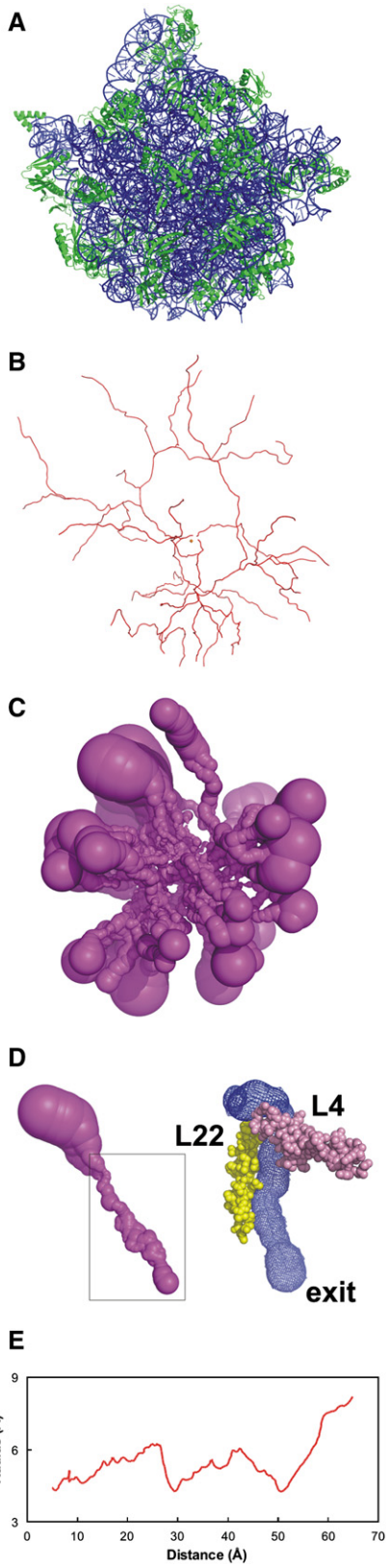
(C) A plot of the channel profiles found by the CAVER (blue line) and MOLE (red line) algorithms.

(D) Table summarizing channel bottlenecks found by the CAVER and MOLE algorithms. The radii of the bottlenecks calculated by HOLE (Smart et al., 1996) for channels 5 and 6 are equal to 1.35 and 0.81 Å, respectively.

cavity is known) by reversing it, i.e., by seeking an exit pathway from a point inside the protein cavity to the exterior. This is the approach adopted here. We begin by defining "the exterior" in order to establish the start and end points of the access route. An ideal algorithm should then identify the shortest, widest, most convenient pathways between these defined points as rapidly as possible.

We recently developed the CAVER program (Petrek et al., 2006), which locates an egress automatically from a given point inside a molecular system represented by the van der Waals radii of its constituents. The CAVER algorithm works by exploring grid nodes utilizing the Dijkstra algorithm. Each grid node is evaluated by a cost function based on the square of the reciprocal distance to the closest atom. Dijkstra's algorithm was used in this manner to select the shortest and cheapest paths (i.e., those with the lowest costs). CAVER, which was developed primarily to find egress paths from protein voids, has also proved capable of locating tunnels in inorganic structures, such as zeolites and channels in DNA quadruplexes (Damborsky et al., 2007). Although it represents a considerable advance in the automatic detection of access and egress paths, the CAVER algorithm has two major limitations: (1) very large demands on processor time and memory when used to explore large channels and (2) small but persistent errors due to grid extrapolation. The exploration of large channels with many internal grid points can become prohibitively expensive with CAVER due to the huge number of possible egress paths that must be analyzed. Thus, it is impractical for the investigation of relatively large access paths, such as ribosomal polypeptide exit channels. The unavoidable errors introduced by CAVER due to grid approximation can be reduced by using a relatively fine grid with small distances, $d$, between nodes (e.g., <0.4 Å). However, a grid that is finer by a factor of $n$ multiplies the number of grid points to be explored by $n^3$, and demands on processor time increase accordingly.

This article presents a strategy for exploring molecular voids based on Voronoi diagrams (Poupon, 2004; Richards, 1974) by using MOLE software, that overcomes many of the limitations of CAVER, since its performance does not depend on the interior volume explored or the use of a finite grid approximation. In addition to a command-line standalone version, the MOLE package can also be used in an easy-to-use graphical environment, either via a plugin for the widely used PyMOL software or an online version, which combines the MOLE algorithm with Jmol graphics. The MOLE software also incorporates a fully automatic clustering algorithm that facilitates analysis of the located tunnels. Applications of the




algorithm to biologically relevant problems are discussed below.

## DISCUSSION

We have thoroughly tested the Voronoi mesh-based MOLE algorithm against the grid-based CAVER algorithm with X-ray structures as well as snapshots taken from MD simulations. Many biomolecular structures, e.g., potassium channels, gramicidines, halorhodopsin, acetylcholine esterases, DNA quadruplexes, and ribozymes, have been analyzed by both approaches (see the Supplemental Data available with this article online). As expected, the results obtained by the MOLE and CAVER algorithms do not differ substantially, for analyses of molecules that they are both capable of handling, although small differences can be seen in the channel profiles, although the MOLE algorithm generally gives smoother channel profiles and more precisely localizes channel bottlenecks, significantly reducing the errors introduced by the finite grid approximation of CAVER. It should be noted that the maximum error introduced by the CAVER algorithm is equal to $\frac{\sqrt{3}}{2}d$, and the mean error to 0.48 $d$, where $d$ is the grid resolution (Petrek et al., 2006). CAVER adds the mean error to the channel profile radii. Differences between the two algorithms can be clearly demonstrated for the KscA $K^+$ channel (PDB: 1K4C—open structure) (Zhou et al., 2001). Here, both algorithms find six channels, the first four of which are symmetrical and aligned perpendicularly to the membrane (Figure 1A, channels 1–4), the next is the channel gating (Figure 1A, channel 5), and the last is the selectivity filter (Figure 1A, channel 6). The channel profiles are almost always slightly overestimated by CAVER (Figure 1C) due to the addition of the mean error, as mentioned above. The bottleneck radii of the symmetrical channels identified by CAVER differ due to the errors introduced by the grid-based algorithm. Clearly, the MOLE algorithm has overcome this shortcoming (as shown by Figure 1D). Nonetheless, both approaches are capable of showing the channel profile, with either a narrow selectivity filter (i.e., allowing the transport of desolvated ions) or a "large" cavity and a channel gating (for transporting solvated ions) with a somewhat larger profile than the selectivity filter (Figure 1C). The MOLE results were also compared with pores in the KscA $K^+$ channel found by HOLE software (Doyle et al., 1998; Smart et al., 1996). The results obtained from MOLE and HOLE were very similar (mean error and

**Figure 2. Analysis of the Large Ribosomal Subunit**

(A) The large ribosomal subunit of *Haloarcula marismortui* (PDB: 1JJ2) represented by secondary structure motifs, where RNA is colored blue, and proteins green.

(B) The web of centerlines of channels found by MOLE.

(C) The same network represented by unions of spheres.

(D) The channel leading through the large ribosomal subunit. The rectangle on the left-hand structure shows the polypeptide exit channel, and the right-hand structure shows the same channel with pink and yellow spheres representing proteins L22 and L4, respectively.

(E) A profile of the ribosomal exit channel showing three bottlenecks, each with a radius of 4.3 Å.

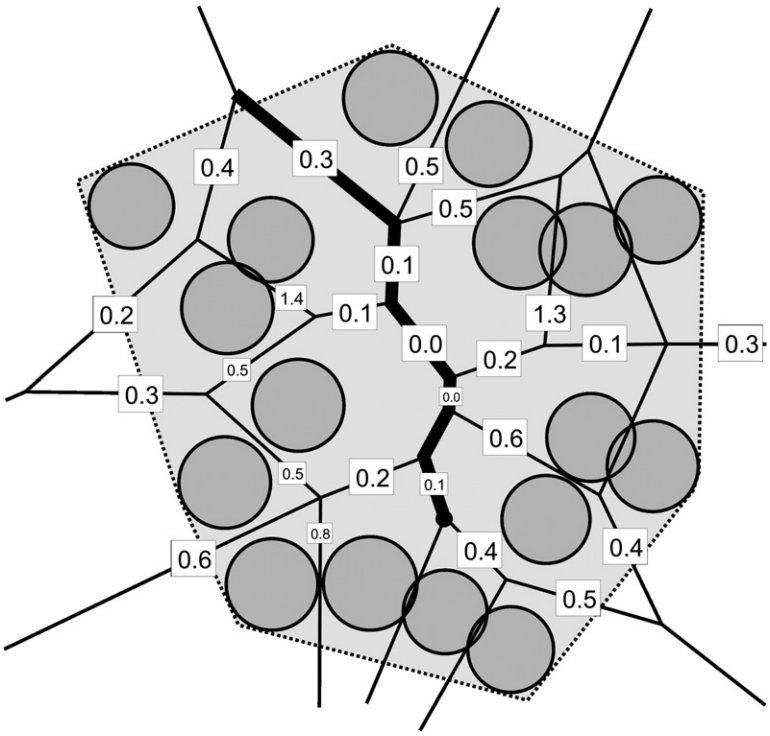



**Figure 3. The Scheme of the 2D Voronoi Diagram**

Scheme shows in 2D the Voronoi diagram (thin lines) and the convex hull (dotted line) of molecule represented by atoms (circles). Each Voronoi edge is evaluated by the cost function (numbers). The thick line represents the optimal path found by the Dijkstra's algorithm from the given starting point (black circle), i.e., centerline of found tunnel.

standard deviation in the channel profile equal to 0.04 Å and 0.16 Å, respectively, using the same set of atomic radii), and the speed of the two software packages is comparable (see the Supplemental Data).

The large ribosomal subunit (PDB: 1JJ2) (Figure 2A) (Klein et al., 2001) from *Haloarcula marismortui* was selected for a more rigorous test of the MOLE algorithm. Due to the enormous demands on processor time, the original CAVER algorithm failed to find an exit path from the ribosomal active site. Like CAVER, HOLE (which performs well on pores) also failed to identify any path from the large ribosomal subunit (Voss et al., 2006). In contrast, the MOLE algorithm presented here found an exit path within ~160 s (CPU Intel PIII 2.6 GHz). All the located paths are represented by a web (Figure 2B) and sets of spheres (Figure 2C), which reveal a complex network of inner ribosomal channels whose biological significance is not yet clearly understood. Among them, one can identify the polypeptide exit channel (Figure 2D), with its opening approximately 65 Å away from the peptidyl transferase center. Three bottlenecks can be distinguished further along the channel, which have very similar radii (4.3 Å) (Figure 2E). MOLE should thus prove to be a valuable tool in what has been termed the "everlasting challenge" (Voss et al., 2006) of exploring the interior of complex molecular networks.

The MOLE algorithm presented here has been shown to be capable of locating the access or egress paths of molecular voids in a rapid and fully automatic manner. Many of the drawbacks of the earlier CAVER algorithm have been eliminated, including errors introduced by a finite grid approximation. We have also implemented automatic clustering of located paths based on their similarity. Robust test data show that the MOLE algorithm provides smoother and more accurate paths than the CAVER approach. The performance of MOLE is demonstrated by its ability to explore, in a relatively short time, (1) large biomolecular channels like the ribosomal peptide exit channel or chaperones (data not shown), (2) complex networks of channels, and (3) structural snapshots taken from molecular dynamic simulations. The MOLE software has been made freely available as a standalone application, plug-in for PyMol software, and online application at http://mole.chemi.muni.cz/.

## EXPERIMENTAL PROCEDURES

### The MOLE Algorithm

A Voronoi diagram divides a metric space according to the distances between discrete sets of specified objects. In the cases considered here, the objects are the centers of protein atoms represented by van der Waals spheres, with radii predefined by AMBER force field (Cornell et al., 1995), and the Voronoi diagram consists of cells representing the set of points closest to the atom in the center of each cell. Voronoi cells of this kind are termed "convex polytopes" or "Dirichlet domains." The boundary of each cell comprises several facets that form the interface between neighboring cells (Figure 3). This simple approach partitions the space solely according to atomic centers and cannot take differing atomic radii into account. However, the facets of this type of Voronoi diagram form planes, and it is usually possible to include differences in atomic radii when constructing a generalized Voronoi diagram. A drawback of this approach is that serious geometric problems can be introduced along the facet boundaries. In order to account for differences in van der Waals atomic radii while retaining the planar geometry of Voronoi facets, an additively weighted Voronoi diagram can be constructed. Since the van der

Waals radii (Cornell et al., 1995) of most heavy atoms that occur in biomolecules (C, N, O, S, P) range from 1.6 to 2.1 Å, the difference between weighted and ordinary Voronoi diagrams is negligible for our purposes (cf. comparison of tunnel paths calculated by MOLE and HOLE presented below) in comparison with other uncertainties, which are introduced via the set radii used, inherent molecular flexibility, and the resolution of most available structures. Other aspects of Voronoi diagram construction will not be considered here since they have been discussed in detail by various other authors (Aurenhammer, 1987; Aurenhammer and Imai, 1987; Kim et al., 2005; Poupon, 2004; Richards, 1974).

The boundaries of Voronoi facets form a three-dimensional mesh (Figure 3) on which the search for molecular channels leading from a given point outside is performed. It is intuitively obvious that an optimal egress path from the interior of a protein will connect points (Voronoi vertices) and segments (Voronoi edges) that are furthest from surrounding atoms. So, in the third step, edges are assigned positive values, representing the relative cost of taking each step along a path and can be thought of as a kind of "highway-toll." An edge is considered convenient if it is both sufficiently far from surrounding atoms and sufficiently short. The cost of an edge is given by a cost function. The selection of a convenient cost function is the key feature of our method; we have modified the original, empirical cost function (Petrek et al., 2006), which did not account for path length. In the modified, MOLE, algorithm the cost of an edge $C(\mathbf{e})$ is given by:

$$C(\mathbf{e}) = l(\mathbf{e}) / \left( d_{\text{closest}}(\mathbf{e})^2 + \varepsilon \right) \quad (1)$$

where $l(\mathbf{e})$ represents the edge length, $d_{\text{closest}}$ is the distance from the edge to the closest atom, and $\varepsilon$ is a small number to avoid division by zero. Dijkstra's graph search algorithm (Dijkstra, 1959) is employed to find and optimize the "cheapest" path on the Voronoi mesh from the starting point outwards. The starting point for a graph exploration is usually defined by the user's specification of the active site; this point should be somewhere in vacant space inside the structure. The starting point can be specified either by Cartesian coordinates or, more flexibly, by a set of groups of residues that form the site of interest. In the latter case, a starting point is calculated from the mass center of each residue. A control subalgorithm with local searching is then carried out to optimize the starting point position; this operation avoids the problem of starting point collisions with surrounding atoms. Optimization is performed by constructing homocentric spheres from the starting point and generating possible starting points on the sphere surfaces with some predefined resolution. The empty space around these surface points is probed to find better candidates for the starting point position, after which the search begins from the Voronoi vertex closest to the optimized starting point position.

The search algorithm explores the Voronoi mesh from an optimized starting point until it reaches one of the so-called boundary points, which lie on a boundary between the structure and its environment. To mark out these points we use a "convex hull" approximation of the structure, with mesh points grouped into interior or exterior categories according to their relationship to the hull. Finally, outer points that have at least one inner-point neighbor are marked as the boundary points.

The MOLE algorithm can be summarized in the following steps: (1) representation of the molecular system under investigation by atom-centered van der Waals spheres; (2) construction of the Voronoi diagram; (3) marking of boundary points by convex hull approximation; (4) optimization of the starting point from a user-defined position; (5) evaluation of edges by a given cost function; and (6) a search for the "cheapest" route from the starting point outward on the Voronoi mesh by Dijsktra's algorithm.

More paths can be located and tracked by continuing to apply the graph-searching algorithm as follows. A large positive "penalty" is added to the Voronoi edges that are parts of a detected path. In subsequent searches, the Dijkstra's algorithm avoids the "penalized" edges due to their high cost, and at the end of the procedure, the new paths are ordered according to cost. Any new paths that do not differ significantly from each other can be grouped to form a set by a new tunnel clustering method; this technique can accelerate the location of qualitatively different tunnels, as described below.

A located path consists of a sequence of joined edges. For each edge, we can locate a small set of the closest atoms that uniquely identify the path. Hence, the similarity of two paths can be measured by a metric suitable for comparing two sets. A metric $D$ is defined to evaluate the similarity of two paths by equation (Equation 2) below:

$$D = \frac{|\mathbf{A} \cap \mathbf{B}|}{\sqrt{|\mathbf{A}| \cdot |\mathbf{B}|}} \quad (2)$$

where **A** and **B** are sets of atoms found along the paths, $|\mathbf{A}|$ is the number of elements of set **A**, and $|\mathbf{A} \cap \mathbf{B}|$ represents the number of elements of **A** that are intersected by set **B**. If $D$ is equal to 1, both sets will be identical, and two paths are considered to be similar if $D \rightarrow 1$. Such an analysis enables all located tunnels to be ordered into path clusters according to their similarity.

## Performance of the MOLE Algorithm

The construction of the Voronoi diagram is inexpensive, and the Qhull software (Barber et al., 1996) used constructs it in $O(n \log n)$ time, where $n$ is a number of points, here atomic centers. Although the theoretical size of the three-dimensional Voronoi vertices generated is $O(n^2)$, in practice it is only $O(n)$, as shown by Dwyer (Dwyer, 1991). The same consideration applies to edges between the vertices. Thus, the evaluation of a cost function is performed quickly over a time that is linearly correlated to the number of edges. The most time-consuming component is the graph search algorithm, and its time complexity is $O(mn)$ for a single source shortest path, where $m$ is the number of edges and $n$ the number of vertices. The MOLE software also benefits from a rapid algorithm for finding additional channels in one structure. This allows fast analyses of molecular dynamic (MD) trajectories in which hundreds or thousands of structural snapshots have to be processed. An analysis of five CYP2C9 channels in MD simulations containing about 250 snapshots takes about 4 hr by CAVER but approximately half an hour by the MOLE algorithm with the same computer.

## Supplemental Data

Supplemental Data include comparison of channel profiles found by MOLE and CAVER, comparison of pore profiles found by MOLE and HOLE, and bottlenecks in the preferred paths, as found by the CAVER and MOLE algorithms in molecular dynamics simulations, and are available online at http://www.structure.org/cgi/content/full/15/11/1357/DC1/.

## ACKNOWLEDGMENTS

We acknowledge grants LC512, LC06030, MSM6198959216, MSM0021622413 (Ministry of Education of the Czech Republic), and GACR 305/06/P139, 204/03/H016 (Grant Agency of the Czech Republic). We thank Sees-Editing, Ltd., (UK) for revising the text of this manuscript. We thank Prof. Jiří Damborský (Masaryk University Brno, Czech Republic) for highlighting the biological significance of tunnels of haloalkane dehalogenases and providing the inspiration and motivation to develop this software. M.P. thanks Petr Kulhánek (Masaryk University Brno, Czech Republic) for his part in the development of a code for adapting AMBER topology.

Received: June 8, 2007
Revised: October 4, 2007
Accepted: October 7, 2007
Published: November 13, 2007

## REFERENCES

Agre, P., Brown, D., and Nielsen, S. (1995). Aquaporin water channels—unanswered questions and unresolved controversies. Curr. Opin. Cell Biol. *7*, 472–483.

Aurenhammer, F. (1987). Power diagrams—properties, algorithms and applications. SIAM J. Comput. *16*, 78–96.

Aurenhammer, F., and Imai, H. (1987). Geometric relations among Voronoi diagrams. In Lecture Notes in Computer Science, *Volume 247,* (Passau, Germany: Springer-Verlag), pp. 53–65.

Barber, B.C., Dobkin, D.P., and Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. ACM Trans. Math. Softw. *22*, 469–483.

Bitter, I., Sato, M., Bender, M.A., Kaufman, A., Wan, M., and Wax, M.R. (2000). Automatic, accurate and robust colon centerline algorithm. Radiology *217*, 370.

Bitter, I., Kaufman, A.E., and Sato, M. (2001). Penalized-distance volumetric skeleton algorithm. IEEE Trans. Vis. Comput. Graph. *7*, 195–206.

Brady, G.P., and Stouten, P.F.W. (2000). Fast prediction and visualization of protein binding pockets with PASS. J. Comput. Aided Mol. Des. *14*, 383–401.

Cojocaru, V., Winn, P.J., and Wade, R.C. (2007). The ins and outs of cytochrome P450s. Biochim. Biophys. Acta *1770*, 390–401.

Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz, J.K.M., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Caldwell, J.W., and Kollman, P.A. (1995). A 2nd generation force-field for simulation of proteins, nucleic-acids and organic-molecules. J. Am. Chem. Soc. *117*, 5179–5197.

Damborsky, J., Petrek, M., Banas, P., and Otyepka, M. (2007). Identification of tunnels in proteins, nucleic acids, inorganic materials and molecular ensembles. Biotechnol. J. *2*, 62–67.

Dijkstra, E.W. (1959). A note on two problems in connection with graphs. Numerische Mathematik *1*, 83–89.

Doyle, D.A., Cabral, J.M., Pfuetzner, R.A., Kuo, A.L., Gulbis, J.M., Cohen, S.L., Chait, B.T., and MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K+ conduction and selectivity. Science *280*, 69–77.

Dwyer, R.A. (1991). Higher-dimensional Voronoi diagrams in linear expected time. Discrete Comput. Geom. *6*, 343–367.

Edelsbrunner, H., Facello, M., and Liang, J. (1998). On the definition and the construction of pockets in macromolecules. Discrete Appl. Math. *88*, 83–102.

Engel, A., Walz, T., and Agre, P. (1994). The aquaporin family of membrane water channels. Curr. Opin. Struct. Biol. *4*, 545–553.

Engel, A., Fijiyoshi, Y., and Agre, P. (2000). The importance of aquaporin water channel protein structures. EMBO J. *19*, 800–806.

Garber, M., and Lin, M.C. (2002). Constraint-based motion planning using Voronoi diagrams. Proc. Fifth International Workshop on Algorithmic Foundations of Robotics (WAFR).

Gouaux, E., and MacKinnon, R. (2005). Principles of selective ion transport in channels and pumps. Science *310*, 1461–1465.

Hendlich, M., Rippmann, F., and Barnickel, G. (1997). LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. J. Mol. Graph. Model. *15*, 359–364.

Huang, B.D., and Schroeder, M. (2006). LIGSITE(csc): predicting ligand binding sites using the Connolly surface and degree of conservation. BMC Struct. Biol. *6*, 19.

Jiang, Y.X., Lee, A., Chen, J.Y., Cadene, M., Chait, B.T., and MacKinnon, R. (2002). Crystal structure and mechanism of a calcium-gated potassium channel. Nature *417*, 515–522.

Kaufman, A.E., Lakare, S., Kreeger, K., and Bitter, I. (2005). Virtual colonoscopy. Commun. ACM *48*, 37–41.

Kim, C.-M., Won, C.-I., Cho, Y., Kim, D., Lee, S., Bhak, J., and Kim, D.-S. (2006). Interaction interfaces in proteins via the Voronoi diagram of atoms. Comput. Aided Des. *38*, 1192–1204.

Kim, D.S., Cho, Y., and Kim, D. (2005). Euclidean Voronoi diagram of 3D balls and its computation via tracing edges. Comput. Aided Des. *37*, 1412–1424.

Klein, D.J., Schmeing, T.M., Moore, P.B., and Steitz, T.A. (2001). The kink-turn: a new RNA secondary structure motif. EMBO J. *20*, 4214–4221.

Kleywegt, G.J., and Jones, T.A. (1994). Detection, delineation, measurement and display of cavities in macromolecular structures. Acta Crystallogr. D Biol. Crystallogr. *50*, 178–185.

Laskowski, R.A. (1995). Surfnet—a program for visualizing molecular-surfaces, cavities, and intermolecular interactions. J. Mol. Graph. *13*, 323–325.

Laurie, A.T.R., and Jackson, R.M. (2005). Q-SiteFinder: an energy-based method for the prediction of protein-ligand binding sites. Bioinformatics *21*, 1908–1916.

Lesk, A.M. (1986). Molecular speleology—the exploration of crevices in proteins for prediction of binding-sites, design of drugs and analysis of surface recognition. Acta Crystallogr. A *42*, 83–85.

Levitt, D.G., and Banaszak, L.J. (1992). Pocket: a computer-graphics method for identifying and displaying protein cavities and their surrounding amino acids. J. Mol. Graph. *10*, 229–234.

Liang, J., Edelsbrunner, H., Fu, P., Sudhakar, P.V., and Subramaniam, S. (1998a). Analytical shape computation of macromolecules: I. Molecular area and volume through alpha shape. Proteins Struct. Funct. Genet. *33*, 1–17.

Liang, J., Edelsbrunner, H., Fu, P., Sudhakar, P.V., and Subramaniam, S. (1998b). Analytical shape computation of macromolecules: II. Inaccessible cavities in proteins. Proteins Struct. Funct. Genet. *33*, 18–29.

Liang, J., Edelsbrunner, H., and Woodward, C. (1998c). Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. Protein Sci. *7*, 1884–1897.

MacKinnon, R. (2002). Nothing automatic about ion-channel structures. Nature *416*, 261–262.

MacKinnon, R. (2003). Potassium channels. FEBS Lett. *555*, 62–65.

Otyepka, M., Skopalik, J., Anzenbacherova, E., and Anzenbacher, P. (2007). What common structural features and variations of mammalian P450s are known to date? Biochim. Biophys. Acta *1770*, 376–389.

Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J., and Damborsky, J. (2006). CAVER: a new tool to explore routes from protein clefts, pockets and cavities. BMC Bioinformatics *7*, 316.

Polarz, S., and Smarsly, B. (2002). Nanoporous materials. J. Nanosci. Nanotechnol. *2*, 581–612.

Poupon, A. (2004). Voronoi and Voronoi-related tessellations in studies of protein structure and interaction. Curr. Opin. Struct. Biol. *14*, 233–241.

Richards, F.M. (1974). The interpretation of protein structures: total volume, group volume distributions and packing density. J. Mol. Biol. *82*, 1–14.

Smart, O.S., Goodfellow, J.M., and Wallace, B.A. (1993). The pore dimensions of gramicidin-A. Biophys. J. *65*, 2455–2460.

Smart, O.S., Neduvelil, J.G., Wang, X., Wallace, B.A., and Sansom, M.S.P. (1996). HOLE: a program for the analysis of the pore dimensions of ion channel structural models. J. Mol. Graph. Model. *14*, 354–360.

Tao, Y.S., Kanoh, H., Abrams, L., and Kaneko, K. (2006). Mesopore-modified zeolites: preparation, characterization, and applications. Chem. Rev. *106*, 896–910.

Voss, N.R., Gerstein, M., Steitz, T.A., and Moore, P.B. (2006). The geometry of the ribosomal polypeptide exit tunnel. J. Mol. Biol. *360*, 893–906.

Wade, R.C., Winn, P.J., Schlichting, E., and Sudarko. (2004). A survey of active site access channels in cytochromes P450. J. Inorg. Biochem. *98*, 1175–1182.

Wade, R.C., Motiejunas, D., Schleinkofer, K., Sudarko, Winn, P.J., Banerjee, A., Kaniakin, A., and Jung, C. (2005). Multiple molecular recognition mechanisms. Cytochrome P450—a case study. Biochim. Biophys. Acta *1754*, 239–244.

Walz, T., Smith, B.L., Agre, P., and Engel, A. (1994). The 3-dimensional structure of human erythrocyte aquaporin chip. EMBO J. *13*, 2985–2993.

Walz, T., Hirai, T., Murata, K., Heymann, J.B., Mitsuoka, K., Fujiyoshi, Y., Smith, B.L., Agre, P., and Engel, A. (1997). The three-dimensional structure of aquaporin-1. Nature *387*, 624–627.

Wan, M., Liang, Z.R., Ke, Q., Hong, L.C., Bitter, I., and Kaufman, A. (2002). Automatic centerline extraction for virtual colonoscopy. IEEE Trans. Med. Imaging *21*, 1450–1460.

Zhou, Y.F., Morais-Cabral, J.H., Kaufman, A., and MacKinnon, R. (2001). Chemistry of ion coordination and hydration revealed by a K+ channel-Fab complex at 2.0 angstrom resolution. Nature *414*, 43–48.