Voronoia: analyzing packing in protein structures

Kristian Rother[1],*, Peter Werner Hildebrand[2], Andrean Goede[3], Bjoern Gruening[4] and Robert Preissner[4]

[1]International Institute for Molecular and Cell Biology, ul. ks. Trojdena 4, 02-109 Warszawa, Poland, [2]Institut for Medical Physics and Biophysics, Charité - University Medicine Berlin, Ziegelstr. 5-9, 10098 Berlin, [3]Institute for Biochemistry, Charité - University Medicine Berlin, Monbijoustrasse 2, 10117 Berlin and [4]Institute of Molecular Biology and Bioinformatics, Charité - University Medicine Berlin, Arnimallee 22, 14195 Berlin, Germany

ABSTRACT

The packing of protein atoms is an indicator for their stability and functionality, and applied in determining thermostability, in protein design, ligand binding and to identify flexible regions in proteins. Here, we present Voronoia, a database of atomic-scale packing data for protein 3D structures. It is based on an improved Voronoi Cell algorithm using hyperboloid interfaces to construct atomic volumes, and to resolve solvent-accessible and -inaccessible regions of atoms. The database contains atomic volumes, local packing densities and interior cavities calculated for 61 318 biological units from the PDB. A report for each structure summarizes the packing by residue and atom types, and lists the environment of interior cavities. The packing data are compared to a nonredundant set of structures from SCOP superfamilies. Both packing densities and cavities can be visualized in the 3D structures by the Jmol plugin. Additionally, PDB files can be submitted to the Voronoia server for calculation. This service performs calculations for most full-atomic protein structures within a few minutes. For batch jobs, a standalone version of the program with an optional PyMOL plugin is available for download. The database can be freely accessed at: http://bioinformatics.charite.de/voronoia.

INTRODUCTION

The packing of protein atoms is an important indicator for their stability and functionality. Measuring the packing of protein structures has been successfully applied to calculate the intrinsic compressibility of proteins (1,2), ligand binding and protein design (3), to find structural determinants for thermostability (4) and to identify flexible regions in proteins (5). Protein packing has recently received importance in assessing the quality of models in tertiary structure prediction (6), and to find water molecules not resolved in the crystal structure. However, tools for the analysis of packing are sparse. Here, we present Voronoia, a database for atomic-scale packing density analysis, also available as a downloadable application.

Proteins are not homogeneously packed; in their interior, they contain many atom-sized cavities that preferably locate close to the surface (7). Often, water molecules that reside in these cavities are not detected in crystal structures. We have observed an accumulation of cavities at functionally important sites of membrane channels and transporters (2,8). This conforms to an earlier finding that cavities are crucial for the conformational flexibility of photosynthetic proteins (9). Knowing the position of cavities allows to predict the effects of mutations (10), and thus to engineer proteins with increased stability (11). In neurotransmitter receptors, interior cavities are believed to form the site of action for some anesthetics (12).

Several procedures to assess the spatial proximity of atoms in 3D space have been developed, most allocating space to Voronoi polyhedra (13). Atom radii were introduced to apply the Voronoi procedure to proteins (14). This method was further improved by the Voronoi Cell method, using curved instead of planar interfaces and treating atoms at cavities and at the protein surface separately (15). An alternative approach is the Alpha Shape (AS) method (16) which models atoms as self-inflating spheres. Mathematically, this is a generalization of the Voronoi method applicable to a multitude of domains other than biomolecules, but it lacks the ability to calculate packing densities for atoms exposed to the solvent. Finally, the Occluded Surface (OS) method (17) calculates the length of rays protruding from an atom, but overestimates densities for atoms adjacent to cavities (2), making it more difficult to detect subtle packing differences in these regions.

---

*To whom correspondence should be addressed. Tel: +48 60 70 95 012; Fax: +48 22 597 07 15; Email: krother@genesilico.pl
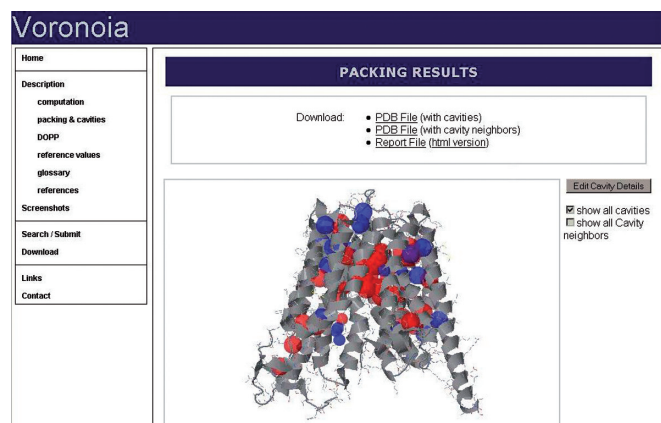
Figure 1. Screenshot of the 3D visualization of an entry in Voronoia. Here, the glycerol-3-phosphate transporter (PDB code 1pw4) is shown. It has a transmembrane domain, in which a number of nonpolar (blue) and polar (red) cavities indicate regions of enhanced flexibility.

The AS and OS methods are available as downloadable programs. While these methods produce reasonable data, there are only few possibilities to report and visualize the results, especially for nonprogrammers. Other software comes with comfortable user interfaces, but has limited usage: SURFNET (18), CastP (19) and CAVER (20) calculate cavities and surface pockets in detail but no packing. Voro3D offers a wide range of functions but is restricted to residue-wise calculation of packing (21).

The Voronoi Cell method

These limitations are overcome by Voronoia, an implementation of the state-of-the-art Voronoi Cell algorithm (15). It includes construction of hyperboloid instead of planar interfaces, assigns individual radii for each atom, and separates the atomic volume between a Van der Waals sphere and a solvent-excluded portion. While the allocation of space is similar to that made by the AS method, the definition of a solvent excluded volume provides an extra level of detail, especially in protein regions with large packing defects. The volumes are calculated numerically by applying a cubic lattice to the protein and calculating each cubiculum separately. This approach reduces computation time to a few minutes for most structures. The resulting per-atom measure is the packing density:

$$PD = V_{vdw}/(V_{vdw} + V_{se})$$

where V_{vdw} is the assigned atomic volume inside the atoms' Van der Waals radius and V_{se} is the remaining solvent excluded volume. The Voronoi Cell procedure uses the empirical ProtOr (22) set of atom radii, and distinguishes between buried atoms, surface atoms and atoms neighboring internal cavities.

Voronoia web interface

The calculation of packing files at higher accuracy is a time-consuming process. To allow the user to quickly visualize the packing of a particular PDB structure, we created an online database of packing data for all biological units from the PDB using the Voronoia program.

In total, 61 318 pre-calculated packing files calculated with a grid distance of 0.02 Å, the highest accuracy possible. The results may be displayed in textual form, summarizing packing densities for buried atoms in the structure, and for each residue separately. Also, a complete list of interior cavities and their environment is given. For visual inspection, a Jmol-based 3D plugin allows to distinguish tightly and loosely packed regions in a structure, and to locate interior cavities (Figure 1). To query the database, a live search of PDB headers is enabled by an AJAX-based mechanism.

Reference packing data calculated from a current non-redundant set of 744 structures based on SCOP superfamilies is provided. For each structure in the database, its deviation from the reference values is given as a *z*-score RMS. If a particular file cannot be calculated, e.g. because the source data violates the PDB format or the file size reaches hundreds of MB, the user will be informed. This database will be updated every 6 months by means of an automatic script. A web interface has been set up at that allows performing the packing analysis for arbitrary input structures. The Voronoia web service performs calculations for most full-atomic protein structures within one to a few minutes. A lattice resolution for the Voronoi Cell algorithm of up to 0.1 Å is available on the web server, which is sufficient for most applications.

A standalone version of the Voronoia program is available with and without graphical interface. It uses regular PDB structure files as input, and creates modified PDB files containing packing densities for each atom. Both cavities and their surroundings may be written as separate structure files. Textual reports that summarize all this information and the deviation of packing from reference data are created for each structure. Voronoia is capable of creating these reports for a batch of structures in one go. This also generates average packing values, which may be used as a new reference set. Parameters such as the resolution of the cubic lattice, and the subset of atoms considered may be specified. Options going beyond those of the web database include choosing atom radii between the empirical ProtOr set (22) and a set optimized for molecular dynamics (23). Using Voronoia as a plugin for the PyMOL molecular viewer, the packing density, its deviation from reference values and cavities can be visualized immediately upon calculation.

CONCLUSION

Concluding, Voronoia is the first comprehensive easy-to-use toolkit for packing analysis. It allows browsing of pre-calculated packing densities and cavities, and *de novo* calculation for single protein structures and data sets, making it useful for a wide range of applications. Voronoia is available at http://bioinformatics.charite.de/voronoia.

ACKNOWLEDGEMENTS

FUNDING

Conflict of interest statement. None declared.

REFERENCES

1. Paci,E. and Marchi,M. (1996) Intrinsic compressibility and volume compression in solvated proteins by molecular dynamics simulation at high pressure. *Proc. Natl Acad. Sci. USA*, **93**, 11609–11614.
2. Hildebrand,P.W., Rother,K., Goede,A., Preissner,R. and Frömmel,C. (2005) Molecular packing and packing defects in helical membrane proteins. *Biophys. J.*, **88**, 1970–1977.
3. Liang,J., Edelsbrunner,H. and Woodward,C. (1998) Anatomy of protein pockets and cavities: measurement of binding site geometry and implications for ligand design. *Protein Sci.*, **7**, 1884–1897.
4. Glyakina,A.V., Garbuzynskiy,S.O., Lobanov,M.Y. and Galzitskaya,O.V. (2007) Different packing of external residues can explain differences in the thermostability of proteins from thermophililic and mesophilic organisms. *Bioinformatics*, **23**, 2231–2238.
5. Jernigan,R.L. and Kloczkowski,A. (2007) Packing regularities in biological structures relate to their dynamics. *Methods Mol. Biol.*, **350**, 251–276.
6. Pawlowski,M., Gajda,M.J., Matlak,R. and Bujnicki,J.M. (2008) MetaMQAP: a meta-server for the quality assessment of protein models. *BMC Bioinformatics*, **9**, 403.
7. Rother,K., Preissner,R., Goede,A. and Frömmel,C. (2003) Inhomogeneous molecular density: reference packing densities and distribution of cavities within proteins. *Bioinformatics*, **19**, 2112–2121.
8. Hildebrand,P.W., Gunther,S., Goede,A., Forrest,L., Frommel,C. and Preissner,R. (2008) Hydrogen-bonding and packing features of membrane proteins: functional implications. *Biophys. J.*, **94**, 1945–1953.
9. Shlyk-Kerner,O., Samish,I., Kaftan,D., Holland,N., Sai,P.S., Kless,H. and Scherz,A. (2006) Protein flexibility acclimatizes photosynthetic energy conversion to the ambient temperature. *Nature*, **442**, 827–830.
10. Cuff,A.L. and Martin,A.C. (2004) Analysis of void volumes in proteins and application to stability of the p53 tumour suppressor protein. *J. Mol. Biol.*, **344**, 1199–1209.
11. Teerinen,T., Valjakka,J., Rouvinen,J. and Takkinen,K. (2006) Structure-based stability engineering of the mouse IgG1 Fab fragment by modifying constant domains. *J. Mol. Biol.*, **361**, 687–697.
12. Liu,R., Loll,P.J. and Eckenhoff,R.G. (2005) Structural basis for high-affinity volatile anesthetic binding in a natural 4-helix bundle protein. *FASEB J.*, **19**, 567–576.
13. Voronoi,G.F. (1908) Nouvelles applications des paramètres continus à la théorie de formes quadratiques. *J. angew. reine Math.*, **134**, 198–287.
14. Richards,F.M. (1974) The interpretation of protein structures: total volume, group volume distributions and packing density. *J. Mol. Biol.*, **82**, 1–14.
15. Goede,A., Preissner,R. and Frömmel,C. (1997) Voronoi Cell: new method for allocation of space among atoms: elimination of avoidable errors in calculation of atomic volume and density. *J. Comp. Chem.*, **18**, 1113–1123.
16. Liang,J., Edelsbrunner,H., Fu,P., Sudhakar,P.V. and Subramaniam,S. (1998) Analytical shape computation of macromolecules: I. Molecular area and volume through alpha shape. *Proteins*, **33**, 1–17.
17. Pattabiraman,N., Ward,K.B. and Fleming,P.J. (1995) Occluded molecular surface: analysis of protein packing. *J. Mol. Recognit.*, **8**, 334–344.
18. Laskowski,R.A. (1995) SURFNET: a program for visualizing molecular surfaces, cavities and intermolecular interactions. *J. Mol. Graph.*, **13**, 323–330.
19. Binkowski,T.A., Naghibzadeh,S. and Liang,J. (2003) CASTp: Computed Atlas of Surface Topography of proteins. *Nucleic Acids Res.*, **31**, 3352–3355.
20. Petrek,M., Otyepka,M., Banás,P., Kosinová,P., Koca,J. and Damborský,J. (2006) CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics*, **7**, 316.
21. Dupuis,F., Sadoc,J.F., Jullien,R., Angelov,B. and Mornon,J.P. (2005) Voro3D: 3D Voronoi tessellations applied to protein structures. *Bioinformatics*, **21**, 1715–1716.
22. Tsai,J., Taylor,R., Chothia,C. and Gerstein,M. (1999) The packing density in proteins: standard radii and volumes. *J. Mol. Biol.*, **290**, 253–266.
23. Stouten,P.F.W., Frömmel,C., Nakamura,H. and Sander,C. (1993) An effective solvation term based an a atomic occupancies for use in protein simulations. *Mol. Simulat.*, **10**, 97–120.