Geometric and topological methods for visualizing channel and cavity structures in biomolecules

Talha Bin Masood

Dept. of Computer Science and Automation, Indian Institute of Science, Bangalore

Primary Goals

- We aim to devise novel geometric and topological techniques to understand the structure of biomolecules.
- In particular, we focus on the study of cavities and channels in proteins.
- We also aim to design interactive visualization tools which enable biologists to conduct detailed visual analysis of biomolecules.

Motivation

- Biomolecules (e.g. proteins) are the basic building blocks of living systems.
- It has been observed that structure of biomolecule plays an important role in defining its function.
- Analysis of structural features is very important for understanding of structure-function relationships, engineering new proteins with required functional properties, or designing inhibitors for existing proteins

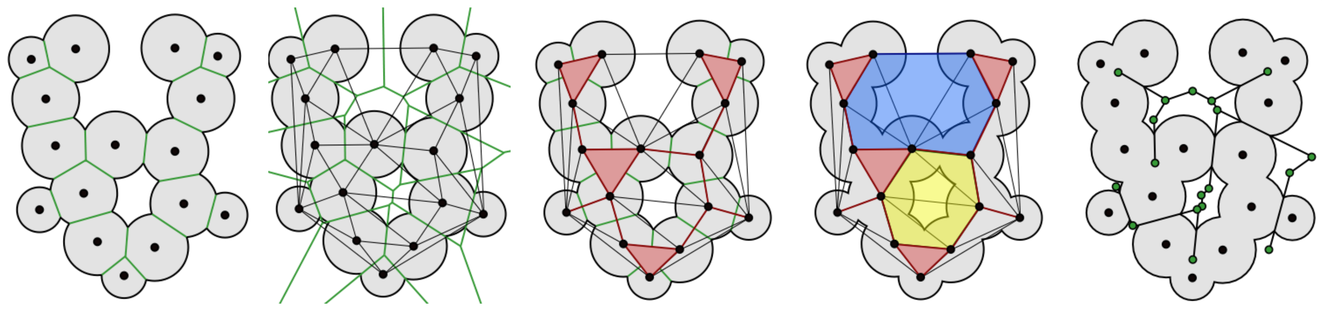
Challenges

- **Size:** Biomolecules consist of thousands of atoms. Identifying interesting features and ranking them based on their significance is non-trivial.
- **Dynamic nature:** Atoms in biomolecules move over time resulting in dynamic structural features.
- **Uncertain data:** Protein structures are obtained experimentally and thus have uncertainty associated with atomic positions and radii.

Alpha Complex based framework

- Molecules can be represented using union of balls model.
- Each atom is represented as a sphere whose radius is *van der Waals* radius.
- Weighted *Voronoi diagram* partitions the space based on proximity to these atoms.
- The dual is called weighted *Delaunay triangulation*.
- *Alpha complex* is a filtration of Delaunay complex.
- Alpha complex at $\alpha = 0$ partitions the molecular space into *Occupied* and *Empty* regions (OR and ER).
- *Cavities* are maximally connected regions in ER.
- *Channels* are pathways in ER.

2D Illustration

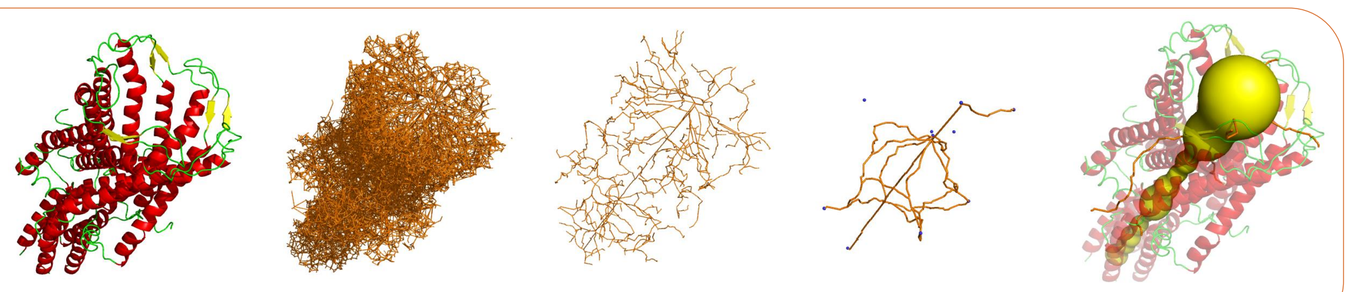


Molecule – as a set of discs | Weighted Voronoi diagram And its dual – Weighted Delaunay triangulation | Alpha complex – A subset of weighted Delaunay triangulation | Cavities are connected subsets of empty region identified by alpha complex | The dual network which we use to identify channels and connect cavities

Channel Extraction and Visualization

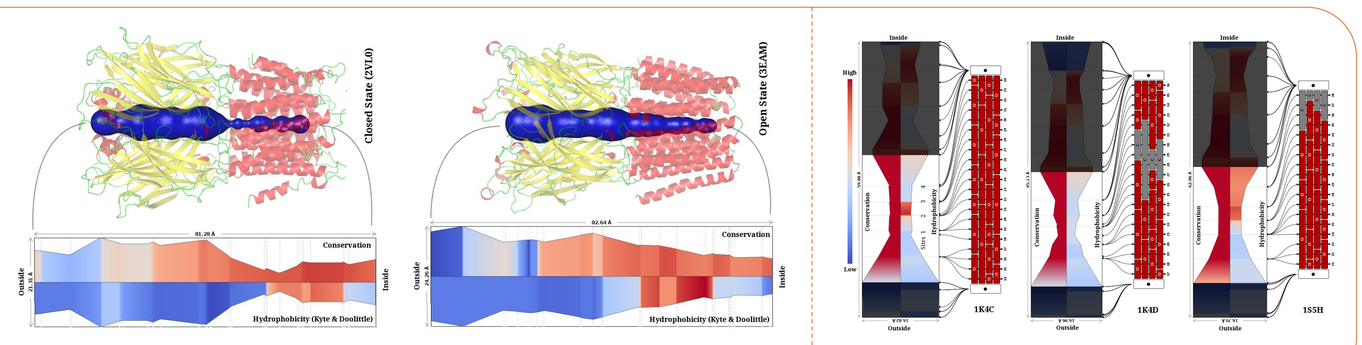
- **Contributions:**
  - Using alpha complex based framework, we design a method to capture all geometrically feasible channels in a concise representation called *channel network* which supports querying for specific channels. The extracted channels are represented as set of connected tetrahedra.
  - We developed novel methods to automatically identify important channels within the network and rank them based on their significance.
  - We also proposed novel visualization methods to facilitate detailed study of the extracted channels.
- **Evaluation:**
  - The integrated channel extraction and visualization framework was successfully used to study multiple transmembrane pores and channels leading to active sites.
  - The channel extraction method was compared with four existing software tools.
- **Web-server:** http://vgl.csa.iisc.ac.in/chexvis/

Method



Input PDB structure | Extracted channel network | Simplified network | Identified set of pores | The top TM pore

Results



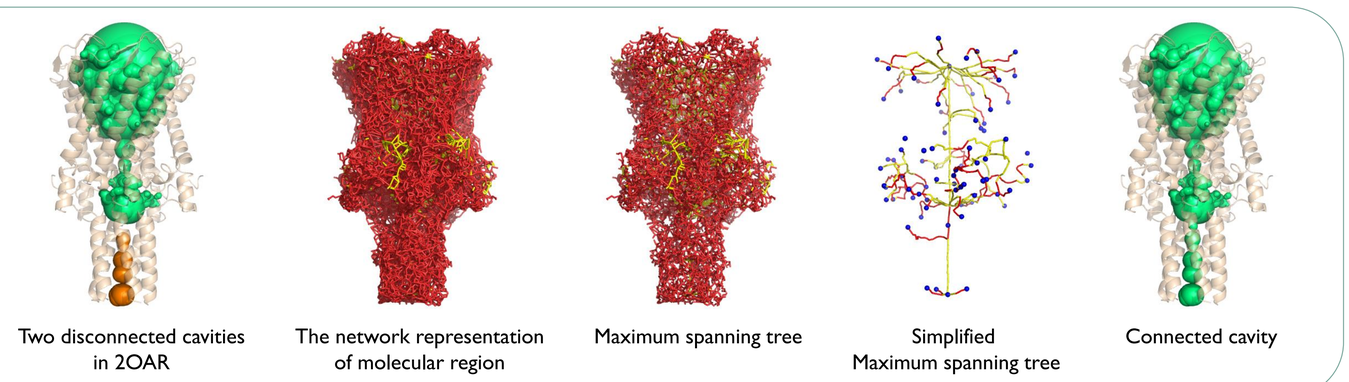
The pentameric ligand-gated ion channels | The comparison of KcsA potassium channels

- T. B. Masood, S. Sandhya, N. Chandra, and V. Natarajan, "ChExVis: a tool for molecular channel extraction and visualization," BMC Bioinformatics, vol. 16, no. 1, pp. 1–19, 2015.

Connecting cavities

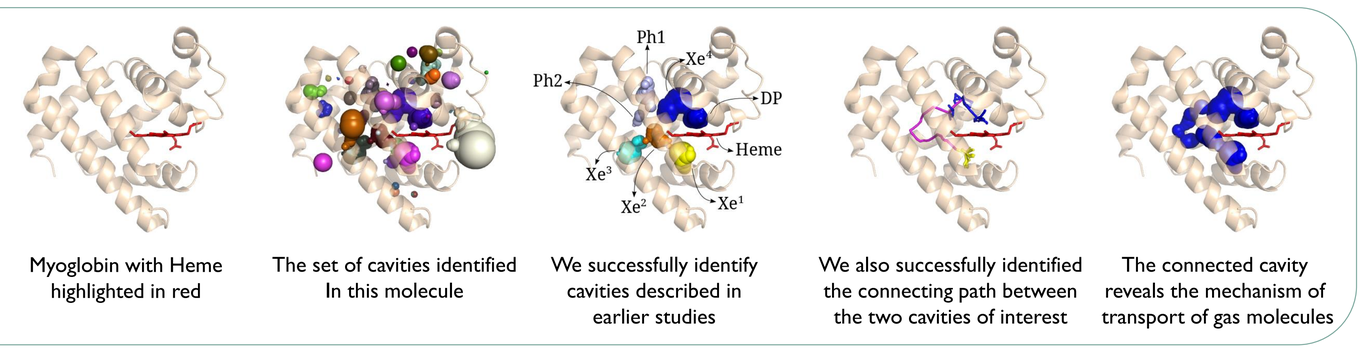
- **Contributions:**
  - We propose a simple and flexible method for extracting cavities in biomolecules from uncertain data with guaranteed bounds on the perturbation required.
  - We also propose efficient algorithms to compute a conduit between user selected cavities that satisfies well defined optimality criteria.
  - We develop an interactive visualization of cavities in a molecule with multiple linked views that facilitates identification of disconnected cavities.
- **Evaluation:**
  - Case studies that demonstrate the benefits of the cavity connection based method.
- **Web-server:** http://vgl.csa.iisc.ac.in/robustCavities/

Method



Two disconnected cavities in 2OAR | The network representation of molecular region | Maximum spanning tree | Simplified Maximum spanning tree | Connected cavity

Results



Myoglobin with Heme highlighted in red | The set of cavities identified In this molecule | We successfully identify cavities described in earlier studies | We also successfully identified the connecting path between the two cavities of interest | The connected cavity reveals the mechanism of transport of gas molecules

- T. B. Masood and V. Natarajan, "An integrated geometric and topological approach to connecting cavities in biomolecules," in 2016 IEEE Pacific Visualization Symposium (PacificVis), 2016.
- R. Sridharamurthy, T. B. Masood, H. Doraiswamy, S. Patel, R. Varadarajan, and V. Natarajan, "Extraction of robust voids and pockets in proteins," Visualization in Medicine and Life Sciences III: Towards Making an Impact , pp. 329–349, 2016.

Acknowledgements

- T. B. Masood was supported by Microsoft Corporation and Microsoft Research India under the Microsoft Research India PhD Fellowship Award.
- This work was partially supported by the Department of Science and Technology, India, under Grant SR/S3/EECE/0086/2012 and the DST Center for Mathematical Biology, IISc, under Grant SR/S4/MS:799/12.