Volume 12

Washington University
Undergraduate Research Digest

Spring 2017

Geometric Algorithms for Identifying Protein Structures

Dan Zeng
Washington University in St. Louis

Geometric Algorithms for Identifying Protein Structures

Dan Zeng

Mentor: Tao Ju

Understanding the structures of macromolecular assemblies is necessary to describe the mechanics of a wide variety of cellular processes. Such assemblies often consist of hundreds of proteins and nucleic acid, each with a unique shape. Modern imaging techniques, including electron cryo-microscopy, create 3D density maps to portray the shapes of these assemblies. However, computational methods for deriving structural models from these images are not fully developed because structural geometry varies depending on the image resolution. To address this issue, we developed geometric algorithms in Gorgon, a molecular visualization software. First, we created an interface for Pathwalking. This algorithm determines protein backbones by first creating a predetermined number of pseudo-atoms at regions with high densities, then determining a path through these pseudo-atoms with a travelling-salesman heuristic. As long as structural features are resolvable, Pathwalking accurately constructs models even at resolutions as low 7-8 °A. Our interface allows for biomedical researchers to adjust parameters for an algorithm which was previously restricted to the command line. To address the cases in which parameters for Pathwalking are not known and when maps are of higher resolution (<5 °A), we also developed Extremal Curve Skeletonization. Skeletonization uses the geometric profiles of locally maximal density curves and surfaces to identify α-helices and β-sheets. Such secondary structures (SSEs) play in pivotal role in protein interactions. We found that the SSEs which could be accurately obtained through skeletonization can be used as anchors for identifying the rest of the protein structure. Our future work involves creating algorithms that will detect the resolution and automatically adjust the algorithm accordingly. Such resolution-aware algorithms will reduce the labor involved in converting density maps into structural models and reduce the human bias that often results from this process.