

Secondary and Tertiary Structural Fold Elucidation from 3D EM Maps of Macromolecules

Chandrajit Bajaj and Samrat Goswami

Department of Computer Sciences,
Computational Visualization Center,
Institute of Computational Engineering and Sciences,
University of Texas at Austin,
Austin, Texas 78712

Abstract. Recent advances in three dimensional Electron Microscopy (3D EM) have given an opportunity to look at the structural building blocks of proteins (and nucleic acids) at varying resolutions. In this paper, we provide algorithms to detect the secondary structural motifs (α -helices and β -sheets) from proteins for which the volumetric maps are reconstructed at 5 – 10Å resolution. Additionally, we show that when the resolution is coarser than 10Å, some of the tertiary structural motifs can be detected from 3D EM. For both these algorithms, we employ the tools from computational geometry and differential topology, specifically the computation of stable/unstable manifolds of certain critical points of the distance function induced by the molecular surface. With the results in this paper, we thus draw a connection between the mathematically well-defined concepts with the bio-chemical structural folds of proteins.

1 Introduction

Three dimensional Electron Microscopy reconstruction (3D EM) and in particular single particle cryo-EM reconstruction [1], has advanced rapidly over recent years, such that several macromolecules (complexes of proteins and ribo-nucleic acids or RNA) can be resolved routinely at low resolution (10-20 Å) and in some cases at sub-nanometer (intermediate) resolution (7-10 Å) [2]. The ultra-structure of these complexes once elucidated from the 3D EM (henceforth 3D Maps), provide not only insights into individual protein and RNA folds and structural motifs, but even more importantly provide information about how the various structural components interact. In addition, with the increasing capability of determining multiple structural folds and conformers (secondary structures and tertiary arrangements) of a complex [3], there is the promise of studying the dynamics of such interacting systems.

Proteins are polypeptide chains of amino acids, (and nucleic acids are sugar-phosphate chains of acidic bases). The secondary structure of proteins are made up of a set of helical (or crudely, cylindrical) arrangement of sub-chains called α -helices and mostly planar arrangement of sub-chains called β -sheets. The various

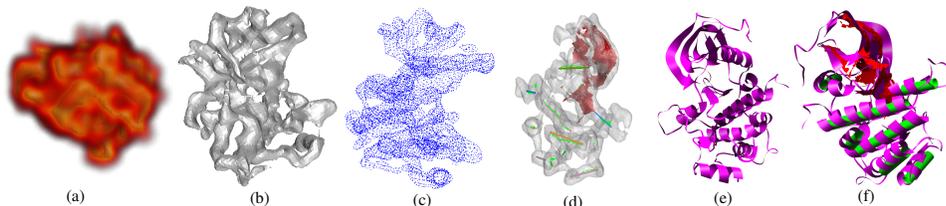


Fig. 1. Secondary structure elucidation algorithm for Insulin Receptor Tyrosine Kinase Domain with pdbid: 1IRK. (a) volume rendering of its blurred 3D Map at 8\AA resolution (b) surface rendering of the protein’s molecular surface (c) pointset sampling of the molecular surface (d) The red patch inside the transparent surface depicts the β -sheet while the straight lines designate the axes of the cylinders which correspond to the α -helices (e) The secondary structural motifs, documented in the Protein Data Bank, where the helices are shown as ribbon coils and the sheets are sets of ribbon strands. (f) combined display of (d) and (e)

structural conformations of sub-groups of helices and sheets, yield the various different tertiary folds [4]. Relatively similar configuration of secondary and tertiary folds arise also in RNA [5].

In this paper, we provide a solution to the problem of automatically elucidating the structural secondary and tertiary folds of proteins (and nucleic acids) from 3D Maps of macromolecular complexes. A 3D Map is akin to a 3D spatial matrix of electron density values. With the improved reconstruction resolution of 3D Maps of macromolecules via 3D EM, the secondary and tertiary structural folds of proteins and RNA can be fully elucidated. Often (as evidenced by structures in the PDB), the atomic resolution structures of individual proteins or RNA, that make up the macromolecule are also discerned via X-ray diffraction and/or Nuclear Magnetic Resonance techniques, allowing us to validate our 3D EM secondary/tertiary structure elucidation algorithms [6]. Depending on the resolution of the 3D Map (in \AA), our goals here are to either detect the secondary structural motifs, or segment the molecule into significant components that can be associated with different tertiary structural folds.

Our processing pipeline has three macro steps, namely, (i) segmentation of the macromolecular 3D Map into individual protein (or RNA), subvolumes using the techniques of [7] and implemented in the publicly available VolRover tool [8]. (ii) computation of a “distance function” to a suitable molecular surface approximation of the individual protein (or RNA) using the methods detailed in [9, 10] (iii) medial axis computation, classification, and construction of stable and unstable manifolds of the critical points of different indices. We skip a description of the first step, as the most recent developments are summarized in [7]. Step (ii) is sketched in Section 3, for completeness. The main contributions of this paper are: (a) Elucidation of secondary structural motifs (α -helices and β -sheets) from relatively higher resolution ($5 - 10\text{\AA}$) 3D Map of individual protein or RNA molecules, (b) Decomposition of the molecular surface of individual proteins or RNA, into its tertiary structural motifs, for 3D Maps at relatively

coarser resolution (10–15Å). Figure 1 shows the steps of the secondary structure elucidation process on the 3D Map of protein 1IRK at a resolution 8 Å.

Prior Work. There are relatively few published algorithms that detect α -helices and β -sheets of proteins from 3D Maps at coarse resolution ($> 6\text{\AA}$). Wen et al have devised an algorithm called *Helix Hunter* [11] for detecting α -helices in a low resolution map where they modeled the helices as cylinders and the density function of the cylinder is convolved with the original map to detect the peaks of the cross-correlation. The main disadvantage of this technique is that it searches exhaustively over the space of all rigid-body transformations and therefore is very slow. A similar approach for detection of β -sheets was adopted by Kong and Ma [12, 13] who modeled it as a disk like primitive and searched through the input map to find the possible positions of the disk that yielded high cross-correlation. This work, due to its exhaustive search paradigm, is also extremely compute intensive. Recently, Yu and Bajaj [14] have developed a secondary structure elucidation algorithm based on the relative magnitudes of the eigenvalues of the structure tensor computed at various select groups of voxels.

Given a compact surface Σ smoothly embedded in \mathbb{R}^3 , a distance function h_Σ can be assigned over \mathbb{R}^3 that assigns to each point its distance to Σ . $h_\Sigma : \mathbb{R}^3 \rightarrow \mathbb{R}$, $x \mapsto \inf_{p \in \Sigma} \|x - p\|$ In applications, Σ is often known via a finite set of sample points P of Σ . Therefore it is quite natural to approximate the function h_Σ by the function $h_P : \mathbb{R}^3 \rightarrow \mathbb{R}$, $x \mapsto \min_{p \in P} \|x - p\|$ which assigns to each point in \mathbb{R}^3 the distance to the nearest sample point in P . Distance functions have found use in surface reconstruction [9, 15–17], Medial axis approximation [18, 19], shape segmentation and feature analysis [20]. Recently, Goswami et al [21] have presented an algorithm to compute the unstable manifolds of the index 1 and index 2 saddle points of this distance function and demonstrated its use in detecting flat and tubular features of any shape.

2 Preliminaries

Voronoi-Delaunay Diagram: We do not to go over the detail about this well-known datastructure due to space limitation and encourage reader to consult a standard computational geometry textbook [22]. In this work, we primarily use the duality of Voronoi and Delaunay diagram which states that every $k \leq 3$ dimensional Voronoi element is dual to a $3 - k$ dimensional Delaunay simplex.

Critical Points of h_P : The critical points of h_P are the points in \mathbb{R}^3 which lie within the convex hull of its closest points from P . It turns out that the critical points of h_P are the intersection points of the Voronoi objects with their dual Delaunay objects [23].

- *Maxima* are the Voronoi vertices contained in their dual tetrahedra,
- *Index 2 saddles* lie at the intersection of Voronoi edges with their dual Delaunay triangles,

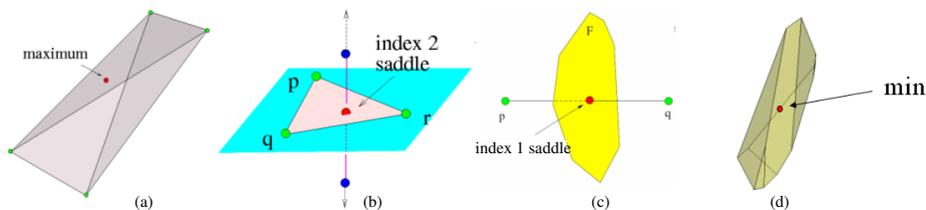


Fig. 2. The relative position of Voronoi and their dual Delaunay objects that results in the generation of critical points.

- *Index 1 saddles* lie at the intersection of Voronoi facets with their dual Delaunay edges, and
- *Minima* are the sample points themselves as they are always contained in their Voronoi cells.

An illustration of the four types of critical points and the relative position of the Voronoi/Delaunay objects resulting these four types is shown in Figure 2.

At any point $x \in \mathbb{R}^3$, one can assign a vector field which is the direction of the steepest ascent of the distance function. The critical points are assigned zero vectors. This vector field induces a *flow*. If a point is allowed to move following the vector field, it traces an *orbit* and ends at a critical point. The set of points whose orbits end at a critical point c is called the *stable manifold* of c . Similarly, a point traces an *inverted orbit* when it follows the steepest descent of the distance function, and ends at a critical point c' . The set of points whose inverted orbits end at c' is called *unstable manifold* of c' .

The stable manifold of a maximum is a three dimensional polytope which is bounded by the stable manifold of critical points of lower indices. Similarly, the unstable manifold of a minimum is a three dimensional polytope which is bounded by the unstable manifold of critical points of higher indices. In this paper, our focus is on stable manifold of maxima and unstable manifold of the index 1 and 2 saddle points.

3 Secondary Structure Identification

The processing pipeline of α -helix and β -sheet identification from 3D Maps, consists of the following steps:

(a) Molecular Surface Extraction and Sampling: Starting with a 3D EM map of a protein (or RNA), a molecular surface is extracted via contouring [24]. For robust contouring, we use the implementation in the publicly available software TexMol [25]. Although several possible isosurfaces can be computed from the 3D Map, we select an isovalue using contour trees [26] (also implemented in TexMol), and topological curvature wherein the isosurface is a single connected component, after removal of smaller completely nested surface components (i.e. voids). We

call this extracted and curated surface component the molecular surface S of the protein (or RNA). Further, a sufficiently dense set of points are sampled from S . We call this pointset P which becomes the input to the latter stages of our algorithm.

- (b) Detection of index 1 and index 2 critical points of h_P (Sec. 2).
- (c) Computing unstable manifold of a subset of those critical points (Sec. 3.1).
- (d) Detection of α -helix and β -sheets from the unstable manifolds (Sec. 3.2).

3.1 Computing U_1 and U_2 from P

Structure and computation of the unstable manifold of an index 1 saddle point (U_1) and an index 2 saddle point (U_2) have been described in detail in [21]. For completeness, we describe it briefly here.

U_1 : Unstable Manifolds of index 1 saddle points are two dimensional. An index 1 saddle point, c lies at the intersection of a Voronoi facet F and a Delaunay edge. For any point $x \in F \setminus c$, h_P increases radially outward from c . Therefore the orbit of one such x hits the Voronoi edges bounding F . Thus F is in $U(c)$. Once the flow hits a Voronoi edge, if the dual Delaunay triangle is acute angled, the flow is along the Voronoi edge, and otherwise, the flow enters the Voronoi facet dual the Delaunay edge opposite to the largest angle of the dual Delaunay triangle. This iterative process computes the unstable manifold of c . The exact computation and its approximation have been described in [21]. Figure 3(a) illustrates an intermediate stage of this computation where the blue facet contains c , yellow facets are currently in $U(c)$ and pink facets are to be included in the subsequent iterations.

U_2 : An index 2 saddle point is generated by the intersection of a Voronoi edge and a Delaunay triangle. The unstable manifold of an index 2 saddle point is one dimensional. It is a polyline with one endpoint at the saddle point and the other endpoint at a local maximum. The polyline consists of segments that are either subsets of Voronoi edges or lie in the Voronoi facets. Due to the later case, the polyline may not be a subcomplex of $\text{Vor } P$. Again, the exact computation and its approximation have been described in [21]. Figure 3(b) illustrates an example. Figure 3(c) shows the unstable manifolds of index 1 and index 2 saddle points on the interior medial axis of the 3D map of molecule 1IRK.

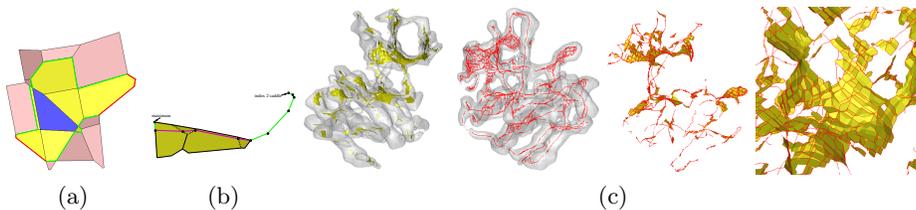


Fig. 3. General structure of U_1 and U_2 is shown in (a,b). (c) A collage of U_1 (yellow) and U_2 (red) of the molecule 1IRK. Note U_2 bounds U_1 .

3.2 α -helix and β -sheet Selection

The unstable manifold of index 1 and index 2 saddle points restricted to the medial axis of the molecular surface decomposes the medial axis into linear (U_2) and planar (U_1) portions. We call the linear subset M_L and the planar subset M_F . The next task is to select a subset of M_L and M_F that gives the α -helices and β -sheets of the protein. Typically, an α -helix is a cylindrical subvolume of the molecule which is of width approximately 2.5\AA . Also the subvolume does not deviate much from a straight cylinder for proteins [4, 11]. These two conditions, dictate the following computational steps.

The unstable manifold of every index 2 saddle on the medial axis is a polyline with Voronoi vertices at the joints. Every Voronoi vertex has h_P value which can be computed by the circumradius of the dual Delaunay tetrahedron. Locally this gives the radius of the cylinder that best fits the molecular surface in the cylindrical regions. We first populate the set H with the Voronoi vertices whose h_P values fall within 2\AA and 3\AA . A 3\AA neighborhood graph is then computed over H that clusters the points. The choice of 3\AA is dictated by the fact that pitch of the helices is 1.5\AA and usually there is more than 2 turns in every helix. The diametrical point pair in every cluster is then computed. The maximum deviation of any intermediate point from the straight line joining the point pair decided how straight the fitted cylinder to the cluster is. This way we select those clusters from H and the cylinders fitted to these clusters produce the detected α helices. The process is shown in 4(b).

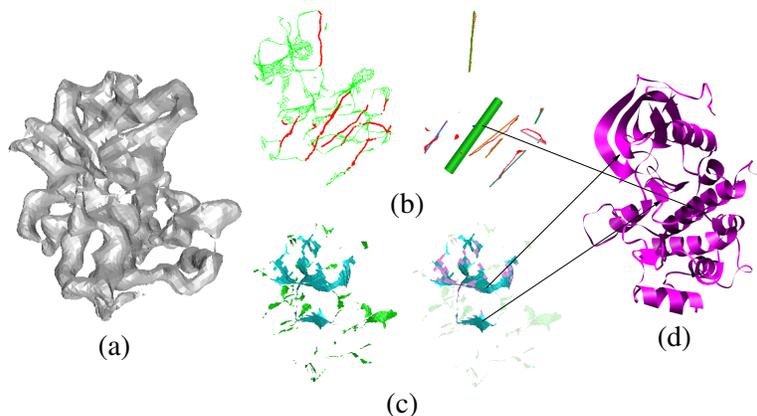


Fig. 4. (a) The molecular surface of 1IRK. (b) The selected Voronoi vertices on U_2 and the fitted cylinder. (c) Filtering out subsets of U_1 which are small (green) or do not satisfy the width test (magenta). (d) shows the secondary structures obtained from the PDB and its correspondence with the computed structure (b,c).

The selection of β -sheets is similar. U_1 gives the possible candidates for β -sheets. First we notice that there are some tiny components that are created due to sampling artifacts, and they do not correspond to real planar substructures of the molecule. We first filter these small clusters out (green patches in Figure 4(c)). After this first stage of filtering, we are left with the planar subsets of the

medial axis (cyan in Figure 4(c)). At this point we apply the knowledge, that β -sheets are of width roughly 1.5\AA [4, 12] and we filter out the planar patches which do not satisfy this “thickness” criterion. The width of a facet in U_1 is easy to check as they are the Voronoi facets and therefore have an 1-1 correspondence with their dual Delaunay edges which cross the medial axis. Therefore, we select only those Voronoi facets from U_1 whose dual Delaunay edges are of length between 1\AA and 2\AA . The portions of U_1 which are filtered out by this test are shown in magenta in Figure 4(c). The rest of U_1 , which qualify for β -sheets, are shown in cyan. The sheets detected in are shown to correspond well with the β -sheets of the molecule 1IRK obtained from Protein Data Bank.

4 Tertiary Fold Elucidation

Tertiary structural folds (or motifs) provide useful information about the conformational and packing arrangement of a protein molecule. Such tertiary folds can be discerned when a coarser resolution 3D Map of the molecule is available. In this section, we show how such information can be gleaned again by looking at the distance function h_P induced by the set of points P sampled on an approximation of the molecular surface S extracted from a coarse resolution 3D Map. Our main focus is on the maxima of h_P . We have already seen that the maxima are the intersection of the Delaunay tetrahedra with their dual Voronoi vertices. In other words, these are the circumcenters of only those Delaunay tetrahedra whose circumcenters lie inside them. As a first step of the tertiary fold detection, we collect the maxima which are the circumcenters of the interior Delaunay tetrahedra.

Once the set of interior maxima is populated, we use their stable manifolds to decompose the volume bounded by S into a set of bio-chemically meaningful segments. These stable manifolds are three dimensional solid subsets of the interior of the molecular surface S . Such techniques have been proved useful earlier in segmentation and matching of free-form objects [20]. We follow the algorithm for computing the stable manifold of a maximum approximately, as described in [20].

Given two Delaunay tetrahedra σ and σ' which share a common triangle t , we say $\sigma < \sigma'$, if the circumcenter of σ lies in the half-space defined by t that does not contain the fourth vertex of σ . Figure 5(a) describes this case. Note that σ' is not always unique (Figure 5(b)). However, this can be proved that, a Delaunay tetrahedron σ can have at most two neighbors σ_1 and σ_2 , for which $\sigma < \sigma_1$ and $\sigma < \sigma_2$. Also, it is to be noted, that the Delaunay tetrahedron σ_{max} whose dual Voronoi vertex is a maximum, has none of its neighbors σ' for which $\sigma_{max} < \sigma'$.

Following the above observation the stable manifold of the maxima are approximated. The set of maxima is sorted according to h_P . Starting from the biggest maximum, the algorithm collects all σ which falls under the transitive closure of the relation ‘ $<$ ’. A subtle problem remains. The stable manifolds of the maxima are often numerous, and therefore they need to be clubbed carefully

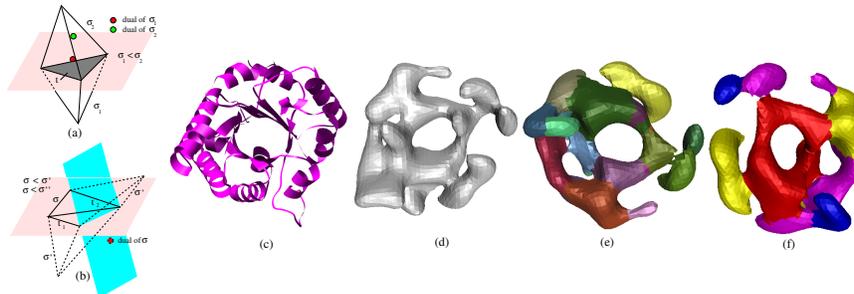


Fig. 5. (a) $\sigma_1 < \sigma_2$. (b) The situation when σ has two neighbors σ' and σ'' for both of which $\sigma < \sigma'$ and $\sigma < \sigma''$. (c) 1TIM: Helices surround the sheets to form the tertiary structure called α/β -barrel. (d) Molecular surface of 1TIM at 15\AA resolution. (e,f) The initial segmentation and further refinement to bring out the β -fold of the barrel from the surrounding helices (yellow, magenta and blue).

to bring out the underlying features of the molecule. To this goal, we apply a merging step, which adjoins two stable manifolds of two maxima sharing a common boundary when the h_P function values at the maxima as well as at a point on the common boundary are comparable to each other.

Figure 5(c-f) shows an example of the stable manifolds and the tertiary folds that they correspond to. It is worth mentioning that, in our experience, tertiary motifs are not always readily decomposable using this algorithm. Nevertheless, we observe that the decomposition algorithm successfully detects the helical region and separates them from the beta regions. The beta regions are sometimes decomposed into more than one component which need to be associated separately to reflect a single fold.

5 Implementation and Results

PDBID	Protein	PDBID	Protein
1IRK	Insulin Receptor Tyrosin Kinase Domain	1TIM	Triose Phosphate Isomerase
1PLQ	Proliferating Cell Nuclear Antigen	1RIE	Rieske Iron-sulphur Protein
1CID	T Cell Surface Glycoprotein CD4	1MBN	Myoglobin
1VDF	Cartilage Oligomeric Matrix Protein	1JPC	Agglutinin
1AOR	Aldehyde Ferredoxin Oxydoreductase	1BBH	Cytochrome C'

Table 1. Name and PDBID of the proteins used in the calibration process.

For calibrating our structure elucidation algorithms, we downloaded atomic level descriptions of proteins from the Protein Data Bank [27]. For each protein, we first “blurred” them into a 3D map (to correspond to a reconstructed 3D EM map) at varying resolutions (5 to 15\AA), using publicly available software EMAN [28]. From these volumetric maps, we extracted the proteins molecular surface

using TEXMOL [25]. Next we collected a pointset sampling of the molecular surface, and used the COCONE software for surface reconstruction and medial axis computation [29]. The next step was to detect the critical points and compute the stable/unstable manifolds of a subset of the critical points. These computations rely on the Voronoi Diagram - Delaunay triangulation of the pointset, and were done using the CGAL library [30].

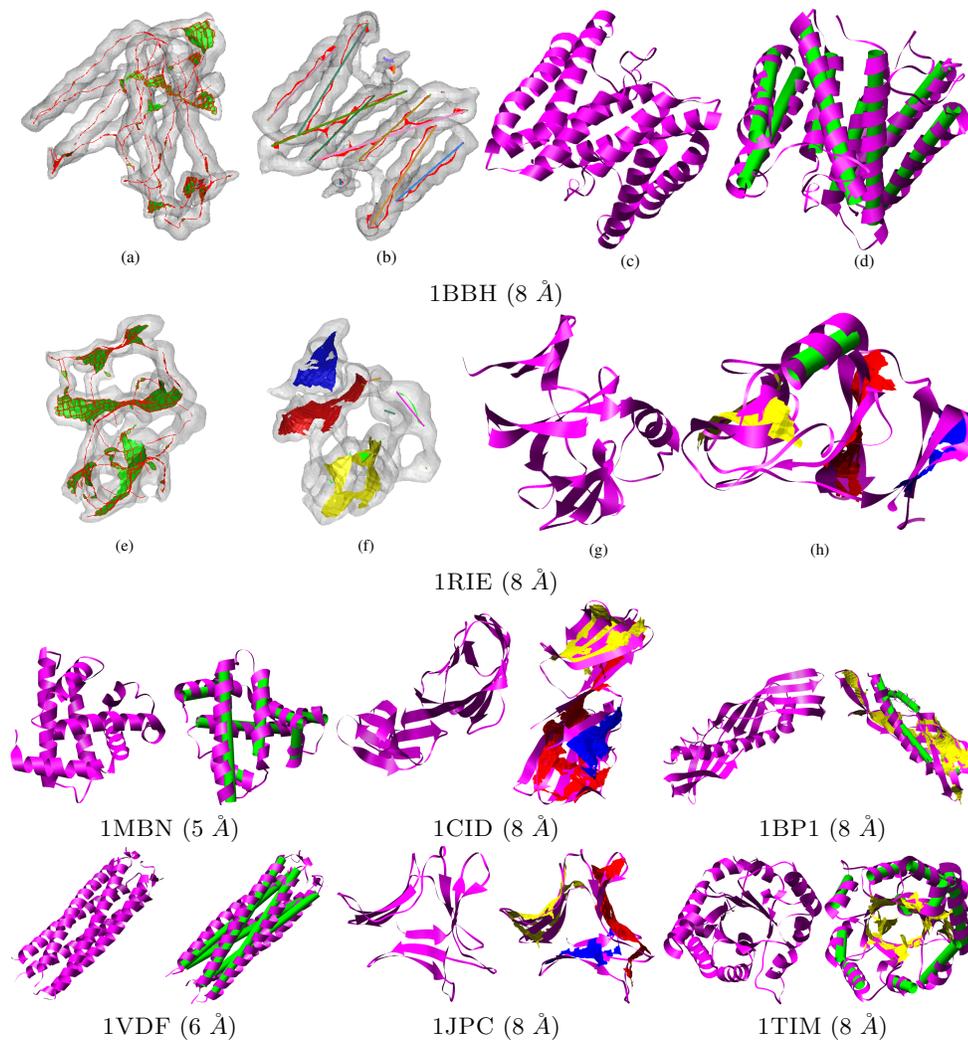


Fig. 6. Performance of our α -helix and β -sheet detection algorithm.

The calibration process is essential before we apply them to 3D EM maps of unknown atomic descriptions. The datasets used in the calibration process are

summarized in Table 1. Details of the molecules are available from the Protein Data Bank (PDB) via the pdbid. Figure 6 shows the snapshots of the key steps of the algorithm. Figure 6(a) shows the U_1 (green) and U_2 (red) of 1BBH. The clusters of points lying on U_2 , selected by the width criterion of α -helix, and the axis of the cylinders to be fitted to the clusters are shown in Figure 6(b). Figure 6(c) shows the secondary structure of 1BBH documented in the PDB. It has total 8 helices and all of them are detected correctly (green cylinders) by our algorithm (Figure 6(d)). The second row in Figure 6 shows similar set of pictures for 1RIE. It has three β -sheets which are identified correctly (red, yellow and blue patches). Third and fourth row show the performance of the algorithm on six more molecules at varying resolutions.

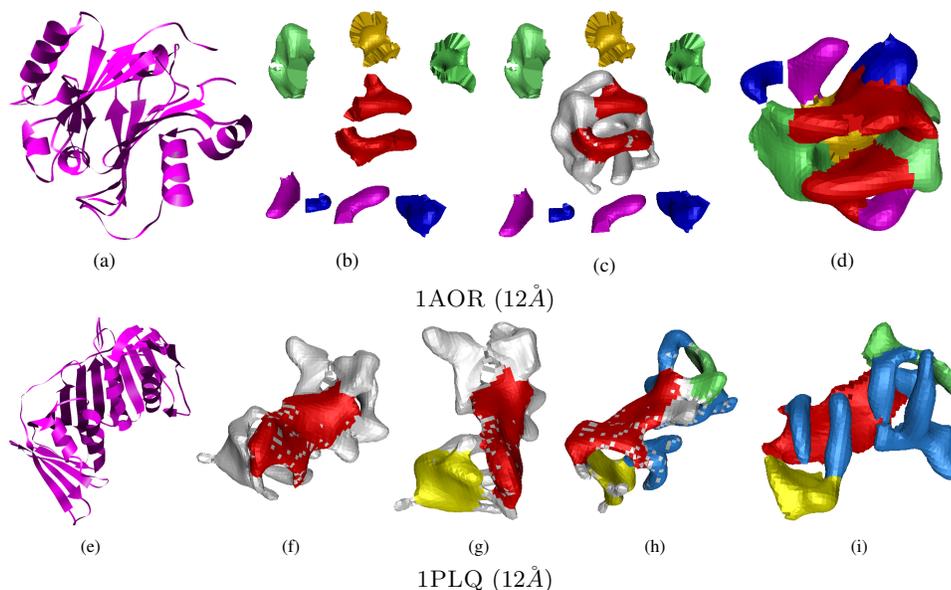


Fig. 7. Performance of the tertiary fold elucidation algorithm.

Figure 7 shows the performance of our tertiary fold detection algorithm. Top row shows the secondary structure of 1AOR as provided by PDB along with the ensemble of the segments of the protein at 12 Å. The tertiary structure of 1AOR is named as β sandwich. The two red segments (in c) correspond to the two β -sheets. The bottom row shows similar results for 1PLQ (at 12 Å). The tertiary fold elucidation can successfully separate the three sheets (red, yellow, green) from the helices (blue).

6 Conclusions

We have presented an algorithm for secondary and tertiary fold elucidation of a protein from 3D EM maps at varying resolutions. Similar constructions are applicable for ultra-structure elucidation of RNA's. The algorithms work by

analyzing the stable and unstable manifolds of a subset of the critical points of the distance function, computed from the molecular surface pointset sampling of the protein.

The algorithm presented in this paper relies on a suitable approximation of the molecular surface. We plan to further investigate the choice of such surfaces based on the bond lengths of the atoms present to narrow down the possible range of selection.

Also we believe the critical points of h_P which lie outside the molecular surface, carry useful information about the tertiary structure. For example, presence of α/β barrels accommodate a through hole in protein surface which can be characterized by exterior index-2 saddle points.

Acknowledgments

The authors are supported in part by NSF grants ITR-EIA-0325550, CNS-0540033 and NIH grants P20 RR020647, R01 GM074258-021 and R01-GM073087. We thank the Jyamiti group at The Ohio State University for providing the Cocone software for surface reconstruction and medial axis approximation. We also thank Dr. Zeyun Yu for valuable discussion on secondary structure elucidation of virus proteins.

References

1. Frank, J.: Three-Dimensional Electron Microscope of Macromolecular Assemblies. San Diego: Academic Press (1996)
2. Baker, T.S., Olson, N.H., Fuller, S.D.: Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbio. and Mol. Bio. Reviews* **63**(4) (1999) 862–922
3. Zhou, Z.H., Baker, M.L., Jiang, W., Dougherty, M., Jakana, J., Dong, G., Lu, G., Chiu, W.: Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus. *Nature Structural Biology* **8**(10) (2001) 868–873
4. Branden, C., Tooze, J.: *An Introduction to Protein Structure*. Garland Publishing Inc., New York (1999)
5. Schneider, B., Morveki, Z., Berman, H.: RNA conformational classes. *Nucleic Acids Research* **32**, **5** (2004) 1666–1677
6. Bajaj, C., Yu, Z.: Geometric processing of reconstructed 3D maps of macromolecular complexes. In Aluru, S., ed.: *Handbook of Computational Molecular Biology*. Chapman and Hall, CRC Press (2005) Chap 45
7. Yu, Z., Bajaj, C.: Automatic ultra-structure segmentation of reconstructed cryo-em maps of icosahedral viruses. *IEEE Transactions on Image Processing* **14**, **9** (2005) 1324–1337
8. CVC, UT Austin: Volrover. (<http://cvcweb.ices.utexas.edu/software/guides.php>)
9. Bajaj, C., Bernardini, F., Xu, G.: Automatic reconstruction of surfaces and scalar fields from 3D scans. In: *ACM SIGGRAPH*. (1995) 109–118
10. Bajaj, C., Lee, H., Merkert, R., Pascucci, V.: Nurbs based b-rep models from macromolecules and their properties. In: *Proc. 4th Sympos. on Solid Modeling and Applications*. (1997) 217–228

11. Wen, J., Baker, M., Ludtke, S., Chiu, W.: Bridging the information gap: computational tools for intermediate resolution structure interpretation. *J. of Mol. Bio.* **308** (2001) 1033–1044
12. Kong, Y., Ma, J.: A structural-informatics approach for mining β - sheets: locating sheets in intermediate-resolution density maps. *J. of Mol. Bio* **332** (2003) 399–413
13. Kong, Y., X.Zhang, T.S.Baker, Ma, J.: A structural-informatics approach for tracing β -sheets: building pseudo-Ca traces for β -strands in intermediate-resolution density maps. *J. of Mol. Bio.* **339** (2004) 117–130
14. Yu, Z., Bajaj, C.: Computational approaches for automatic structural analysis of large bio-molecular complexes. Technical Report TR-06-05, Department of Computer Sciences, UT Austin (2006)
15. Edelsbrunner, H.: Surface reconstruction by wrapping finite point sets in space. In Aronov, B., Basu, S., Pach, J., Sharir, M., eds.: *Ricky Pollack and Eli Goodman Festschrift*. Springer-Verlag (2002) 379–404
16. Chaine, R.: A geometric convection approach of 3-d reconstruction. In: *Proc. Eurographics Sympos. on Geometry Processing*. (2003) 218–229
17. Giesen, J., John, M.: The flow complex: a data structure for geometric modeling. In: *Proc. 14th ACM-SIAM Sympos. Discrete Algorithms*. (2003) 285–294
18. Chazal, F., Lieutier, A.: Stability and homotopy of a subset of the medial axis. In: *Proc. 9th ACM Sympos. Solid Modeling and Applications*. (2004) 243–248
19. Siddiqi, K., Shokoufandeh, A., Dickinson, J., Zucker, S.: Shock graphs and shape matching. *Computer Vision* (1998) 222–229
20. Dey, T.K., Giesen, J., Goswami, S.: Shape segmentation and matching with flow discretization. In Dehne, F., Sack, J.R., Smid, M., eds.: *Proc. Workshop Algorithms Data Structures (WADS 03)*. LNCS 2748, Berlin, Germany (2003) 25–36
21. Goswami, S., Dey, T.K., Bajaj, C.L.: Identifying flat and tubular regions of a shape by unstable manifolds. In: *Proc. 11th Sympos. Solid and Physical Modeling*. (2006) 27–37
22. de Berg, M., van Kreveld, M., Overmars, M., Schwarzkopf, O.: *Computational Geometry: Algorithms and Applications*. Springer-Verlag, Berlin (1997)
23. Siersma, D.: Voronoi diagrams and morse theory of the distance function (1999)
24. Lorensen, W., Cline, H.: Marching Cubes: A High Resolution 3D Surface Construction Algorithm. In: *SIGGRAPH*. (1987) 163–169
25. Bajaj, C., Djeu, P., Siddavanahalli, V., Thane, A.: Interactive visual exploration of large flexible multi-component molecular complexes. In: *Proc. of the Annual IEEE Visualization Conference*. (2004) 243–250
26. van Kreveld, M., van Oostrum, R., Bajaj, C., Pascucci, V., Schikore, D.: Contour trees and small seed sets for isosurface generation. In Rana, S., ed.: *Topological Data Structures for Surfaces*. John Wiley & Sons (2004) Chap 5, pg 71 –86
27. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., Shindyalov, I., Bourne, P.: The Protein Data Bank. *Nucleic Acids Research* (2000) 235–242
28. Ludtke, S., Baldwin, P., Chiu, W.: EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. of Struct. Bio.* (**128**) 82–97
29. Cocone: Tight Cocone Software for surface reconstruction and medial axis approximation. (<http://www.cse.ohio-state.edu/~tamaldey/cocone.html>)
30. CGAL Consortium: CGAL: Computational Geometry Algorithms Library. (<http://www.cgal.org>)