The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

A Voxels-based Study of Protein Surfaces: Surface Atoms Extraction and Dock Sites Identification

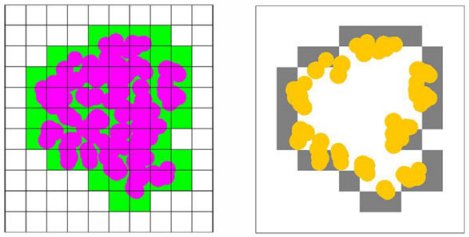
I(C$_2$)S

INTRODUCTION

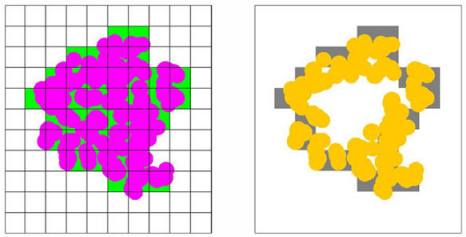
Proteins are formed by chains of amino acids and are essential components of all living organisms. The reactivity of a protein to external binding agents depends on the distribution and arrangement of atoms on its surface. Sequential and structural-based studies of proteins are complemented by protein surface analysis. This research attempts to extract protein surface atoms using a constraint-based voxelisation method, with the extracts improved using statistical analysis. The output obtained is then stored and used as input to the following stage of research which uses computer vision inspired concepts and techniques for the identification of potential dock sites. All research works were carried out on a FK506-bound protein (PDB ID: 1FKF / Human Immunophilin FKBP-12 Complex).

EXTRACTION OF SURFACE ATOMS FROM THE PROTEIN

We have a protein in the voxel space. Voxels containing part of/entire atoms from the protein are highlighted, as shown on the left. In the right image the surface voxels are selected, containing within them the surface atoms as well. However as can be seen not all surface atoms have been selected.

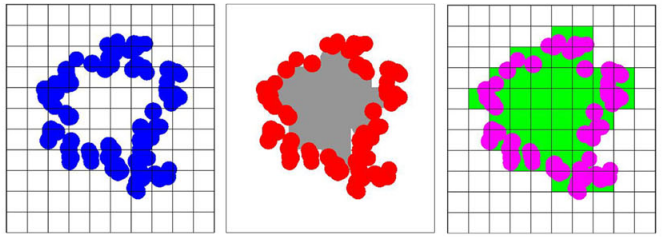


A voxel occupancy of 40%-100% is then applied. The extracts showed improvements.



REFINING THE EXTRACTIONS USING CONSTRAINTS AND STATISTICS

The highlighting of all surface voxels lead to the selection of not only surface atoms, but internal ones as well. Another constraint is introduced to limit the number of internal atoms being included in the form of atom occupancy. Each atom is inspected for its total overlapping area with surface voxels. Any atom with >5% occupancy is shortlisted, as seen in (a).

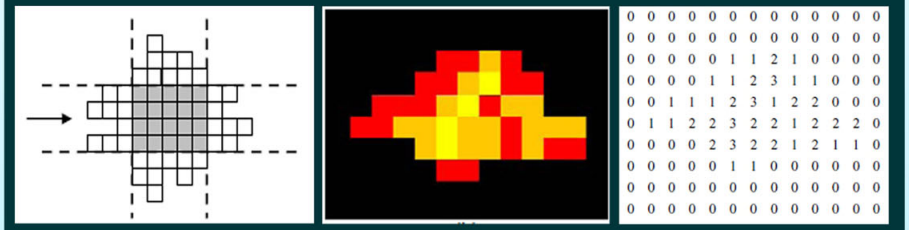


(a) (b) (c)

The next step is the implementation of a peeling algortihm which attempts to remove excess internal atoms. The results are shown in (b). Finally a statistical-based method is introduced for the selection of highly-consistent atoms across a total of 10 rotational cases. This approach is capable of overcoming orientation issues associated with the use of rigid environments even though more cases need to be considered. The results are visualised in (c).

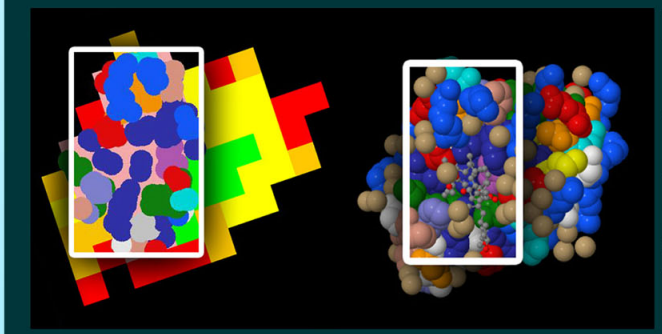
IDENTIFYING THE DOCK SITES AND THEIR CORRESPONDING RESIDUES USING VISION-BASED COLOR-MAPPING

By enclosing the protein within a voxel, we limit the protein to only 6 views - one view for each face of the voxel. When looking at a particular face, one would be able to see protuding regions as well as crevices, with the crevices enclosed by higher regions having darker tones. This motivated the development of a vision-based color mapping of the face. The deepest level (usually the floor of a crevice) begins at Level 1 and the levels continue to build for each higher level of voxels.



(a) (b) (c)

Figure (a) shows a cross-sectional view of a voxeliased protein. The arrow points to a face, and the color-map of the face is given in (b). The corresponding matrix indicating the levels of the voxels is shown in (c). By processing the face using a numbered matrix instead of pixels checking, we reduced the computation time immensely. Dock sites are located at lowest-numbered areas (with the exception of edges) within the voxel cluster. The images on the left show an extract from protein 1FKF with the residues compared against a screenshot from the JMol Viewer in RCSB PDB.



RESEARCH WORK BY

PhD Student : Ling Wei Lee (Malaysia Campus)
Supervisor : Professor Andrzej Bargiela (Jubilee Campus)