



ConSurf: A Server for the Identification of Functional Regions in Proteins by Surface-Mapping of Phylogenetic Information - *Award Winning Poster*

Glaser, F.¹, Pupko, T.², Paz, I.¹, Bechor, D.¹, Martz, E.³ and Ben-Tal, N.¹

¹ Department of Biochemistry, George S. Wise Faculty of Life Sciences, Tel Aviv University

² The Institute of Statistical Mathematics, Minami-Azabu, Minato-ku, Tokyo 106-8569, Japan

³ Department of Microbiology, University of Massachusetts, Amherst, MA, USA

Mutual interactions between proteins and between proteins and peptides, nucleic acids or ligands play a vital role in every biological process. A detailed understanding of the mechanism of these processes requires the identification of functionally important amino acids that are responsible for these interactions.

It is often difficult to determine the three-dimensional (3D) structure of protein complexes, and sometimes only the structures of the unbound proteins are available. Moreover, crystal contacts between proteins are not always indicative of biologically relevant interactions. Thus, it is common to carry out tedious mutagenesis studies to determine functionally important residues. Because of the amount of work required to determine the functionality of the protein, many entries in the Protein Data Bank have only partial information about their function. The relative fraction of such entries is expected to increase rapidly due to recent high throughput studies to determine protein structures.

Recently, we developed algorithmic tools for the identification of functionally important regions in a protein with known 3D-structure, by estimating the degree of conservation of the amino acid sites within its close sequence homologues. The degree of conservation at each amino acid site is similar to the inverse of the site's rate of evolution; slowly evolving sites are evolutionarily conserved while fast evolving sites are variable.

Projecting the conservation grades onto the molecular surface of the protein usually reveals patches of highly conserved (or occasionally highly variable) residues that are often of important biological function.

Here we report the development of a web server, ConSurf, which automates these algorithmic tools and is available to the scientific community at <http://bioinfo.tau.ac.il/ConSurf/>. Providing a protein structure in PDB format, the server extracts the sequence of the selected polypeptide chain (subunit). It then automatically carries out a PSI-BLAST search for close sequence homologues and multiply aligns them using CLUSTAL W. Alternatively, the user can provide a previously made multiple sequence alignment (MSA). In any event, the server builds a phylogenetic tree consistent with the MSA and calculates the conservation grades taking into account the evolutionary relations between the homologues. The protein, with the conservation grades color-coded onto its surface, can finally be visualized on-line using the Protein Explorer engine.

The ConSurf server enables easy, high throughput studies of proteins with known 3D-structure, and we hope that it will become a standard tool in structural biology studies, in biochemical and molecular biology laboratories.

website: <http://bioinfo.tau.ac.il/ConSurf/>