Marvin Edelman

Mariana Babor
Eran Eyal
Brendan McConkey
Rafael Najmanovich
Vladimir Potapov
Alexander Raskind
Vladimir Sobolev

Structural bioinformatics: Approaches to molecular recognition

Department of Plant Sciences

Tel. 972 8 934 3301 **Fax.** 972 8 946 9124
E-mail: marvin.edelman@weizmann.ac.il

Edelman - Sobolev Bioinformatics Group

Consensus structure for adenine binding sites: We are developing an algorithm to search for similarities in the spatial arrangement of binding pocket atoms from multiple target proteins interacting with a given ligand. A non-redundant dataset of high-resolution ATP-protein structures was analyzed. The adenine rings from each structure were superimposed with concerted movement of the protein atoms in contact with the rings. LPC (http://sgedg.weizmann.ac.il/lpc/) software was used to determine the protein atoms in contact with the ligand and to classify the atom contacts according to their physico-chemical properties. Atoms were defined as belonging to a cluster if they were within a given distance of each other, came from different entries and formed attractive contacts with the ligand. The network of atomic clusters so determined was taken as the consensus binding-site structure for the adenine ring of ATP. We created a program to search for a set of atoms from any PDB file having the spatial characteristics of the consensus binding-site structure. Using the ATP-protein complexes in our dataset, we tested the searching program to determine how many putative ATP binding sites-including the correct one-it finds. The procedure is capable of predicting a small number of putative adenine binding sites for an adenine containing structure that includes the correct site with high probability (Fig. 1).

Ribose recognition: Conserved spatial positions among proteins with different folds: A consensus binding-pocket structure was sought for the semi-flexible furanose ring of the ribose moiety in ATP- and ADP-protein complexes. We analysed atoms forming hydrogen bonds with ribose O2 or O3 atoms, including water bridging interactions. A non-redundant set of high-resolution ligand-protein structures, classified according to furanose ring conformation, was derived. A set of positions in different proteins occupied by atoms forming these hydrogen bonds was found in 2 structural groups. In Group 1, we identified clusters that included more than 80% of the structures, while in Group 2 there were several clusters, each representing more than the 50% of the ATP- and ADP-protein structures. In both groups, more than 90% of the structures were represented by the clusters. The conserved ribose binding positions were independent of amino acidic sequence as well as secondary and tertiary structural elements. Water molecules were found to play a mayor role in mediating polar interactions between the ribose moiety and the protein.

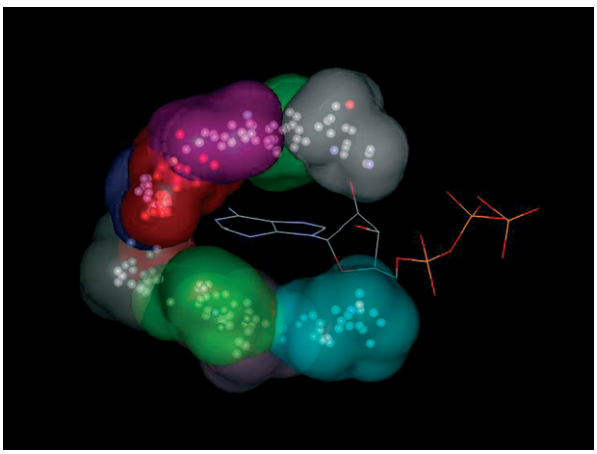

Fig. 1 *Network of atomic clusters defining the concensus structure for adenine binding.*

Protein side chain rearrangement in regions of point mutations: We created a dataset of several hundred high-resolution pairs of protein structures differing in one amino acid (http://bioinfo.weizmann.ac.il/mutaprot/) and analysed it to determine the number and type of residues changing conformation as a result of a point mutation. In 95% of the cases, two or fewer residues change their conformation in the vicinity of the mutation. Surprisingly, the mutation itself is not the major cause for these changes; rather, the inherent flexibility of a specific side chain. This was derived using a control set of proteins, each of whose crystal structure was determined more than once. We conclude that the search space for predicting side chain conformations in the region of a mutation can be effectively restricted. However, the overall ability to predict a particular side chain conformation is still limited. Different amino acids demonstrated a different degree of flexibility near mutation sites: large polar or charged residues and serine are most flexible, while aromatic residues and cysteine are least so. These data are applicable to deriving faster, more accurate algorithms for computational positioning of side chains.

Modeling tentoxin binding sites of chloroplast F1-ATPase: Tentoxin is a specific inhibitor of plastid CF1-ATPase. It interacts with alpha and beta subunits of the enzyme inhibiting activity at a high affinity site and stimulating it at lower affinity site(s). The quasi-symmetrical structure of bovine mitochondria F1-ATPase was used to model tentoxin binding at multiple sites in plastid CF1-ATPase. We located and analyzed 3 theoretical binding sites for tentoxin. Complementarity to tentoxin was sensitive to the nucleotide occupancy state of the β subunit. The main interactions stabilizing the putative complexes were determined. The residues most likely involved in the direct binding of tentoxin in CF1-ATPase were identified. The predicted binding pocket residues for the high affinity site (Fig. 2) are at the alphaTP/betaTP interface and include residue Glu-83β (previously identified by us as a molecular-genetic determinant for response to tentoxin). Biologically relevant pockets were predicted for the low affinity site(s): one located entirely within the alphaTP subunit and sharing one residue with the nucleotide binding site; the other situated at the alphaTP/betaE/gamma interface and sharing one residue with the conserved DELSEED sequence. Non-catalytic residues in these putative pockets represent potential targets for mutational analysis.

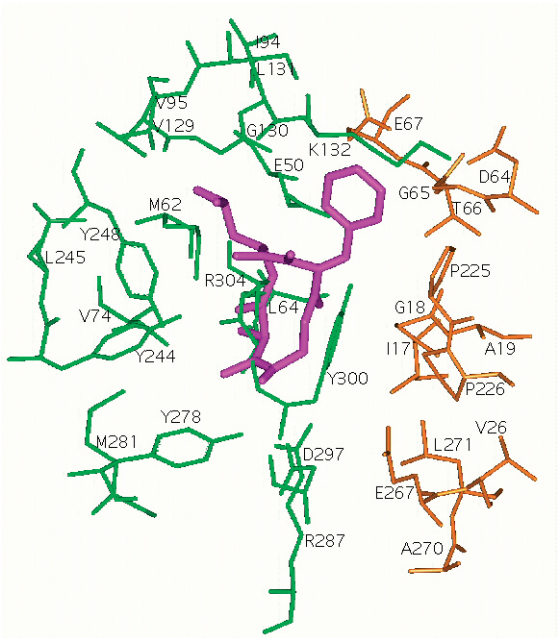

Fig. 2 *Proposed high-affinity binding site for tentoxin.*

Bioinformatics infrastructure

Our group coordinates the National Center for Bioinformatic-Genetic Infrastructure, which, in collaboration with Biological Services, oversees the activities of INN, the Israeli National Node of EMBnet (http://inn.org.il/). We also coordinate the UNESCO-sponsored International Center for Cooperation in Bioinformatics network (ICCBnet), providing hands-on training in developing areas such as India, Turkey and Poland (http://www.iccbnet.org/).

Selected Publications

Sobolev V, Moallem TM, Wade RC, Vriend G and. Edelman M (1997) CASP2 molecular docking predictions with the LIGIN software. Proteins Suppl. 1, 210-214

Sobolev V, Sorokine A, Prilusky J,. Abola EE. and Edelman M (1999) Automated analysis of interatomic contacts in proteins. Bioinformatics 15, 327-332

Najmanovich R., Kuttner J, Sobolev V and Edelman M (2000) Side chain flexibility in proteins upon ligand binding. Proteins 39, 261-268

Eran E, Najmanovich R, Sobolev V and Edelman M (2001) MutaProt: a web interface for structural analysis of point mutations. Bioinformatics 17, 381-382

Kuttner YY, Babor M, Edelman M and Sobolev V (2001) Structural commonality in protein binding sites for ATP. In: Currents in Computational Molecular Biology Ed: El-Manrouk N, Lengauer T, Sankoff D.Les Publ.CRM, Montreal; p. 91-92

McConkey BJ, Sobolev V and Edelman M (2001) The performance of current methods in ligand-protein docking. Curr. Sci. (in press).

Acknowledgements

M.E. is incumbent of the Siegmund Warburg Chair in Agricultural Molecular Biology