Marvin Edelman

Vladimir Sobolev

Moshe Van Dyke
Eran Eyal
Mariana Babor
Vladimir Potapov
Sergei Gerzon

Structural bioinformatics and molecular recognition

Department of Plant Sciences

Tel. 972 8 934 **3301** **Fax.** 972 8 934 **9124**
E-mail: marvin.edelman@weizmann.ac.il vladimir.sobolev@weizmann.ac.il
Web page: www.weizmann.ac.il/~edg

Discrimination of native protein structures using atom-atom contact scoring:

We introduced a method for discriminating correctly folded proteins from well designed decoy structures using atom-atom and atom-solvent contact surfaces. The measure used to quantify contact surfaces integrates the solvent accessible surface and interatomic contacts into one quantity, allowing solvent to be treated as an atom contact (Fig. 1). A scoring function was derived from

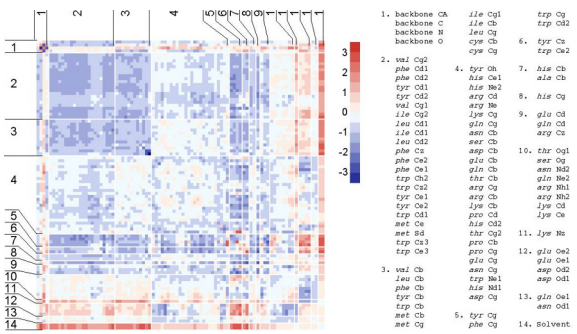

Fig. 1 *Atom-atom contact potentials. Potentials are shown with favorable contacts (blue) as negative values to be consistent with energetic functions. Atom types have been grouped by similar contact potentials and are presented in the given order.*

statistical contact preferences within known protein structures and validated by using established protein decoy sets, including the "Rosetta" decoys and data from the CASP4 structure predictions. The scoring function effectively distinguished native structures from all corresponding decoys in >90% of the cases, using isolated protein subunits as target structures. If contacts between subunits within quaternary structures are included, the accuracy increases to 97%. The contact scoring performed as well or better than existing statistical and physicochemical potentials and may be applied as an independent means of evaluating putative structural models.

Solvent accessibility and contact surfaces in modeling side chain conformations in proteins:

Contact surface area and chemical properties of atoms were used to concurrently predict conformations of multiple amino acid side chains on a fixed protein backbone. Our scoring function is particularly suitable for modeling partially-buried side chains. Both iterative and stochastic searching approaches were used to produce programs SCcomp-I and SCcomp-S, respectively. The two programs, with relatively fast execution times, correctly predict χ_1 angles for 92-93% of buried residues and for 82-84% of all residues, with an RMSD of ~1.7 Å for side chain heavy atoms. We found that the differential between the atomic solvation parameters and the contact surface parameters (including those between non-complementary atoms) is positive; i.e., most protein atoms prefer surface contact with other protein atoms rather than with the solvent (cf. Fig. 1). This might correspond to the driving force for maximizing packing of the protein. The influence of the crystal packing, completeness of rotamer library and precise positioning of C_β atoms on the accuracy of side chain prediction were examined. The SCcomp-I and SCcomp-S programs can be accessed through the web (**http://sgedg.weizmann.ac.il/sccomp.html**).

Protein-protein recognition – The juxtaposition of domain core and interface core positions in sandwich-like proteins:

We are developing a statistical approach to derive the interface core positions of sandwich-like proteins interacting in a sheet–sheet mode. We evaluated a set of 47 immunoglobulin (Ig)

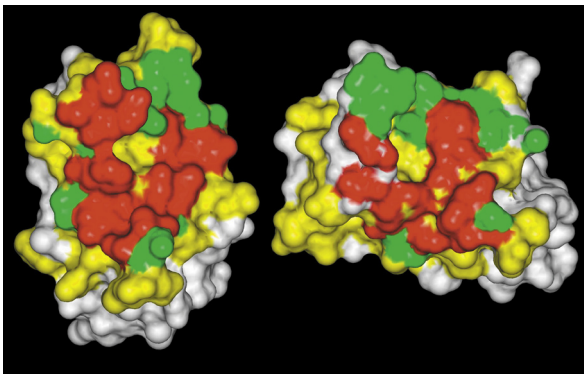

Fig. 2 *The immunoglobin VL-VH interface. Spacefill model of VL (left) and VH (right) domains of pdb entry1a3l. VL-VH interface core (red); entire VL-VH interface (green plus red); virtual VL-VH interface for the complete data set (yellow plus green plus red).*

structures to determine the residue positions that play a primary role in association of the heavy and light chains. Employing a cocktail of criteria (atom–atom contact compatibility, frequency of position in the data set, conservation of residue type and positional conservation in 3D space), we uncovered the existence of highly conserved positional determinants at the two protein surfaces that presumably allow for fast recognition and initial binding of the chains. We further demonstrate that rigidity at the interface surface is geometrically wedded to the domain core, indicating a close relationship between the surface determinants and those involved in protein folding of Ig domains. We find that the same rule of positional connectivity between the rigid domain core and interface core extends to sandwich-like proteins in general that are interacting in a sheet–sheet fashion. Figure 2 illustrates the size diminution of the interface core versus an average Ig interface and versus the cumulative virtual interface of all 47 complexes.

Bioinformatics infrastructure:

Our group also coordinates the activities of the Israeli National Node of EMBnet (INN) **(http://inn.org.il/)** and the Israeli National Center for Bioinformatics Infrastructure (COBI) **(http://cobi.org.il/)**. These national bioinformatics bodies serve over 3000 Israeli researchers.

Selected Publications

Sobolev V., Sorokine A., Prilusky J., Abola E.E. 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.

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

McConkey B.J., Sobolev V. and Edelman M. (2002) The performance of current methods in ligand-protein docking. Curr. Sci., 83, 845-856.

Sobolev V., Niztaev A., Pick U., Avni A., and Edelman M. (2002) A case study in applying docking predictions: Modeling the tentoxin binding sites of chloroplast F1-ATPase. Curr. Sci., 83, 857-867.

McConkey B.J., Sobolev V., and Edelman M. (2002) Quantification of protein surfaces, volumes and atom-atom contacts using a constrained Voronoi procedure. Bioinformatics, 18, 1365-1373.

Babor M., Sobolev V., and Edelman M. (2002) Conserved positions for ribos recognition: importance of water bridging interactions among ATP-, ADP- and FAD-protein complexes. J. Mol. Biol., 323, 523-532.

Eyal E, Najmanovich R, Edelman M., and Sobolev V. (2003) Protein side chain rearrangement in regions of point mutations. Proteins, 50, 272-282.

McConkey B.J., Sobolev V. and Edelman M. (2003) Discrimination of native protein structures using atom-atom contact scoring. Proc. Natl. Acad. Sci. USA, 100, 3215-3220.

Kuttner J., Sobolev V., Raskind A., and Edelman M. (2003) A consensus binding structure for adenine at the atomic level permits searching for the ligand site in a wide spectrum of adenine-containing complexes. Proteins, 52, 400-411.

Eyal E., Najmanovich R., McConkey B.J., Edelman M., Sobolev V. (2004) Importance of solvent accessibility and contact surfaces in modeling side-chain conformations in proteins. J Comput. Chem., 25, 712-724

Acknowledgement:

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

V.S. is a Camea Fellow of the Ministry of Absorption, Division of the Absorption of Scientists