Ephraim Katzir
Miriam Eisenstein

Efrat Ben-Zeev
Noga Kowalsman
Avi Ben-Shimon
Dadi Segal

Development and application of theoretical tools for predicting the structures of molecular complexes

Department of Biological Chemistry

Tel. 972 8 934 **3947** **Fax.** 972 8 934 **8356**
E-mail: ephraim.katzir@weizmann.ac.il, miriam.eisenstein@weizmann.ac.il
Web page: www.weizmann.ac.il/Chemical_Research_Support/molfit

The activity of a living cell can be portrayed as a network of interactions that transfer biological information. Intervention in cellular processes requires thorough understanding of the interactions between the molecules, which can be provided by docking techniques. Our group's goal is to develop reliable docking procedures, which predict the structures of protein-protein complexes given the structures of the component molecules. Such procedures must be able to deal with molecules whose activity is not fully understood and with modeled structures, whose accuracy is limited. Therefore, incorporation of all the available knowledge regarding intermolecular interfaces is important.

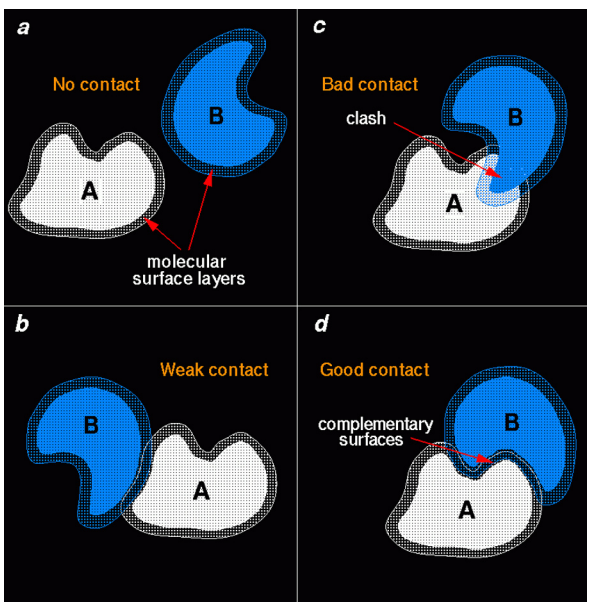

Fig. 1 *A section through the grid representation of two molecules, where the surface layer is distinguished from the interior of the molecule. The figure presents different possibilities in docking: no contact, weak contact, interpenetration and a good contact.*

Our algorithm, implemented in a computer program named MolFit, tests geometric, electrostatic and hydrophobic complementarity. It can also incorporate external data from biological, biochemical and bioinformatics studies. In MolFit, the molecules to be docked (denoted A and B) are rotated with respect to each other and digitized onto 3-dimensional grids. The grids are correlated using Fourier transformations, providing assessment of the quality of the match for different relative positions (see fig. 1). Each grid point in our representation is a complex number. The real part contains the geometric descriptor of the molecular surface, G. Grid points outside the molecule are given the value $G=0$; points on the surface of the molecule are given the value $G=1$ except for the mobile ends of flexible side chains for whom $G<1$; points in the interior of the molecule are given either a negative value (-15), or a positive value (+1), for molecules A and B, respectively. The imaginary part can contain either an electrostatic descriptor, derived from the electrostatic potential, or a hydrophobic descriptor based on the character of the residues nearby, or a weight parameter, which reflects the probability of a given residue to be involved in binding, according to external data.

The structures of oligomers are predicted by combining MolFit with different oligomer forming algorithms, in which symmetry constraints are incorporated. For example: Homo-tetrameric oligomers are formed by combining homo-dimers (produced by MolFit) that comply with the symmetry requirements of the tetramer, either a dimer of dimers (D_2 symmetry) or a planar tetramer (C_4 symmetry).

We apply our algorithm to a large selection of known protein-protein complexes and oligomers, starting from either the bound structures of the component molecules (as they are in the

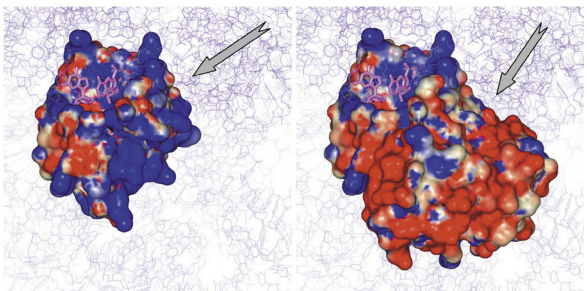

Fig. 2 *Docking model of colE3 to 30S showing the electrostatic potential. The 30S subunit is colored purple except for the cleavage site nucleotides, which are emphasized in pink. The solvent accessible surface of colE3 (left panel) is colored according to the electrostatic potential (red for negative; blue for positive). Note the basic character of the surface that interacts with the 30S ribosomal subunit (where the arrow points). On the right panel the complex colE3:IP is shown after superposition on the colE3 molecule in our model. Note the acidic character of the surface that now faces the 16S rRNA, away from the active site. This suggests that the inhibitory action of IP is due to electrostatic repulsion.*

complex) or their unbound structures (separately determined). Combination of geometric, electrostatic and hydrophobic complementarity measures successfully identifies nearly correct solutions for most of the tested systems and ranks them high. Geometric docking in combination with a dimer-of-dimers forming algorithm successfully predicts the structures of D2 oligomers, starting from one subunit, which can be a homology model. Hydrophobic complementarity further improves the results for such oligomers.

We further test our algorithm by participating in the CAPRI (Critical Assessment of PRediction of Interfaces) blind prediction challenge, together with 25-30 other groups worldwide. Our procedure successfully predicts the structures of most of the targets in CAPRI before the experimental determination. We also apply MolFit to real cases. For example, the complex between the 30S ribosomal subunit and colicin E3 (ColE3) does not crystallize. Our docking procedure provides a model structure that can be used to explain the enzymatic activity of ColE3 and the protective effect of the Immunity Protein (IP) (see fig. 2). Another example is the recognition of crystal surfaces by antibodies, which cannot be studied by the current structure determination techniques. Our predictions provide a rationale for the enantioselectivity of some antibodies and cross reactivity of others.

Selected Publications

Katchalski-Katzir E., Shariv I., Eisenstein M., Friesem A.A., Aflalo C., and Vakser I.A., (1992). Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. PNAS 89, 2195-2199.

Eisenstein M., Shariv I., Koren G., Friesem A.A., and Katchalski-Katzir E., (1997). Modeling supra-molecular helices: extension of the molecular surface recognition algorithm and application to the protein coat of the tobacco mosaic virus. JMB 266, 135-143.

Heifetz A., Katchalsi-Katzir E. and Eisenstein M., (2002). Electrostatics in protein-protein docking. Protein Science 11, 571-587.

Ben-Zeev E. and Eisenstein M., (2003). Weighted geometric docking: Incorporating External Information in the Rotation-Translation Scan. Proteins 52, 24-27.

Heifetz A. and Eisenstein M., (2003). Effect of local shape modifications of molecular surfaces on rigid-body protein-protein docking. Protein Engineering 16, 179-185.

Berchanski A. and Eisenstein M., (2003). Construction of Molecular Assemblies via Docking: Modeling of Tetramers with D2 Symmetry. Proteins, 53, 817-29.

Ben-Zeev E., Berchanski A., Heifetz A., Shapira B. and Eisenstein M., (2003). Prediction of the unknown: Inspiring experience with the CAPRI experiment. Proteins 52, 41-46.

Ben-Zeev E., Zarivach R., Shoham M., Yonath A. and Eisenstein M., (2003). Prediction of the Structure of the Complex between the 30S Ribosomal Subunit and Colicin E3 via Weighted-Geometric Docking. J. BioMolec. Struct. Dyn. 20, 669-675.

Berchanski A., Shapira B. and Eisenstein M., (2004). Hydrophobic complementarity in protein-protein docking. Proteins, in press.

Geva M., Eisenstein M. and Addadi L., (2004). Antibody recognition of chiral surfaces. Structural model of antibody complexes with Leucine-Leucine-Tyrosine crystal surfaces. Proteins, in press.

Acknowledgement:

The Kimmelman Center for Biomolecular Structure and Assembly