A Novel System for Assessment of Macromolecular Interaction in Condensed Phases. 2) Interaction Site Inference by Molecular Shape and Electrostatic Complementarity

Atsushi Yoshimori yosimori@translell.eco.tut.ac.jp

Carlos Adriel Del Carpio Munoz carlos@translell.eco.tut.ac.jp

Department of Ecological Engineering, Toyohashi University of Technology, Tempaku, Toyohashi 441-8580, Japan

1 Introduction

The authors are involved in the development of the system for automatic assessment of macromolecular interaction MIAX described elsewhere [1].

In the present study we undertake the problem of reduction of the huge configuration search space for the complex in order to improve processing times of MIAX. Here we describe an algorithm to infer interaction sites for the monomers composing the complex which dictate plausible starting points for the optimization process performed by MIAX.

The algorithm is based on the computation of macromolecular complementarity, geometrical as well as electrophysical. Shape complementarity is calculated by a modification of the algorithm proposed by Katchalski-Katzir *et al.* [2]. Electrophysical complementarity is evaluated by means of a self organizing map (SOM) [3] algorithm which renders a nonlinear projection of the high dimensional configuration of the electrostatic potentials on the surface of the monomers into a lower dimensional map of neurons.

We evaluate the algorithm predicting plausible interaction sites for several complexes whose coordinates are found in the PDB, and comparing them with the inferred candidate interaction regions.

2 Methodology

2.1 Electrostatic Complementarity

To examine electrostatic complementarity among two interacting macromolecules the electrostatic potential on a probe charge of +1 is placed on every point of a grid on the surface of each molecule obtained by dividing the two angular components of the polar coordinates in intervals of 5 degrees. The electrostatic energy was calculated by the Coulomb law:

$$Eelc_i = 332 \sum_j \frac{q_j}{\epsilon r_{ij}} \tag{1}$$

where $Eelc_i$ represents the electrostatic energy on the molecular surface at a point where the angular components are i(ϕ,θ). q_j is the atomic charge and r_{ij} is the distance between the atomic center and the probe on the molecular surface. A dielectric constant of 4 was used in the calculation.

The calculation is performed for both monomers composing the complex. To evaluate electrostatically complementary regions on both molecules, the three dimensional grid of electrostatic potentials of the first molecule was projected onto a two dimensional map of neurons by means of a self organizing map algorithm. The two dimensional map derived by this learning process for the first molecule is then used as a template to evaluate the complementarity (in fact similarity by changing the electrostatic potentials signs for the second molecule) with the grid of the second molecule by computing the

distance between each grid point and the vectors of each of the neurons of the two dimensional array. Neurons with points from both molecules represent regions of complementarity.

2.2 Geometrical Complementarity

The algorithm of Katchalski-Katzir *et al.* [2], for geometric fit of ligands into proteins is based on a representation of molecular shape by discrete three dimensional functions (derived from its atomic coordinates) which discriminate among the interior and surface of the molecules. A correlation function based on Fourier transforms of those functions to assess the degree of overlap of two molecules (as one of them is rotated and translated respect the other in three dimensions) leads to correlation values indicating the extent of geometric match between the surfaces of the molecules.

Here we have extended the technique to infer sites geometrically complementary in interacting protein molecules. Combination of the algorithm with an a priori analysis of the electrostatic complementarity reduces processing time and the number of inferred complementary interaction sites.

3 Results and Discussion

Validation of the algorithm proposed was carried out by comparing inferred complementary regions for monomers composing complexes whose structures are recorded in the PDB. Here, we show results for Ubiquitin (PDB-CODE:1AAR). Inferred interaction sites are in high agreement with the crystal structure, as shown in Fig.1 for one candidate site. Results for other systems are also in good agreement with the observed ones, yielding in this way appropriate starting positions for the optimization process in MIAX, reducing thoroughly the configurational search-space for the complexes.

Figure 1: Comparison of an inferred interaction site (right) and crystal structure (left).

References

[1] Del Carpio,C.A. and Yoshimori, A., MIAX: a novel system for assessment of macromolecular interaction in condensed phases. 1) description of the interaction model and simulation algorithm, *Genome Informatics 1999*, Universal Academy Press, 1999.

[2] Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesen,A.A., Aflalo, C., and Vakser, I.A., Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques, *Proc. Natl. Acad. Sci. USA*, 89:2195–2199, 1992.

[3] Kohonen, T., The self-organizing map, *Proc. IEEE*, 78:1464–1480, 1990.