John von Neumann Institute for Computing

Development of a Docking Methodology for Predicting the Structure of Protein-Protein Complexes

I. Meliciani, S. M. Gopal, W. Wenzel

published in

From Computational Biophysics to Systems Biology (CBSB08),
Proceedings of the NIC Workshop 2008,
Ulrich H. E. Hansmann, Jan H. Meinke, Sandipan Mohanty,
Walter Nadler, Olav Zimmermann (Editors),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. **40**, ISBN 978-3-9810843-6-8, pp. 325-328, 2008.

http://www.fz-juelich.de/nic-series/volume40

Development of a Docking Methodology for Predicting the Structure of Protein-Protein Complexes

Irene Meliciani, Srinivasa Murthy Gopal, and Wolfgang Wenzel

Institute of Nanotechnology, Research Centre Karlsruhe,
76344 Eggenstein-Leopoldshafen, Germany
E-mail: {Irene.Meliciani, wenzel}@int.fzk.de

Here we report on the development of a strategy for protein-protein docking using an all-atom refinement protocol based our free-energy on-atom protein forcefield PFF02. We use a set of protein complex conformations generated by a heuristic method (ZDOCK) and score the proposed conformations using short refinement simulations in the all-atom forcefield, which predicts the experimental conformation of 1PPE (complex of bovine beta-trypsin and CMTI-I) and 1KAC (complex of knob domain from adenovirus serotype 12 and its cellular receptor CAR) to within 0.5 Å. Such an approach is much faster than a generation of the decoy set in a full simulation of the protein-complex dynamics.

1 Introduction

Many proteins are able to fulfill their biological function only in complex with other proteins. One of the most important examples for this are G-protein coupled receptors (GPCR) which bind about 40% of all known drugs. For this reason the prediction of protein-protein complexes is an important area of applications of biomolecular simulation. Free-energy based methods, such as POEM@HOME are particularly well suited to address these questions, but all-atom simulations that perform an unbiased search of all possible complex conformations are very time consuming and may not, within a reasonable time frame, even visit the experimentally relevant conformation once. One alternate approach to this problem is the generation of a decoy set, or conformational family, with a relatively inexpensive heuristic method and then to score to the members of this decoy set in the all-atom forcefield. Because models generated by one program are generally not trivially transferable to another, each of these decoys must be subjected to a short relaxation simulation. Here we investigate such a protocol, where decoys for 1PPE and 1KAC were generated with ZDOCK, which were subsequently relaxed using POEM@HOME.

2 Method

The simulations were performed with the POEM[1,2] (Protein Optimization by Energy Minimisation) simulation package using the all-atom protein forcefield PFF02[1,2], which identifies the native structure of protein/protein-complexes as the global minimum of the forcefield. The scoring approach to protein-protein docking requires two principal components (i) a fast and effective method for generating possible orientations of monomers and (ii) an accurate energy function to discriminate native and non-native conformations. We used Zdock[3] for generating the decoys set of possible conformations. A decoy set is a large library of protein conformations generated to approximately span all relevant low energy

regions of the free energy landscape. To measure the predictivity and selectivity of a force-field, the conformations in the library (decoy set) must be ranked according to their energy. If near native conformations emerge lowest in the free-energy function, the force field differentiates between native and near-native conformations. In the limit of completeness of the decoy set, which is rarely reached in practice, this test alone is sufficient to show that the force field stabilizes the native conformation of the protein against all competing metastable conformations and corresponds to the global optimum of the free-energy force field.[4] For decoy sets generated with unbiased methods, the computation of the Z-score (the difference between energies of near-native decoys to the mean energy of the decoy set in units of its standard deviation) gives a quantitative measure of the selectivity of the force field. Zdock is an initial-stage docking programme. It uses an FFT-based grid search to scan for optimal translational conformations. To complement Zdock another server Rdock[5] is used to perform the refinement of the initial-stage predictions.

3 Results

The goal of this study was to identify a suitable protocol to generate the best conformations of the complex. As a start we focused on rigid protein docking, where individual monomers do not change conformation on binding. In order to arrive at a meaningful com-

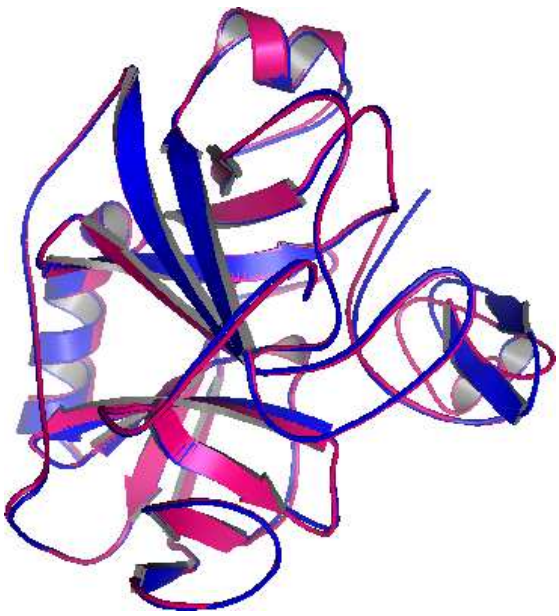

Figure 1. Docking of 1PPE to a RMSD of 0.5 Å.

parison of the energies we relaxed approximately 1000 decoys for 2 proteins in the decoy library in PFF02. This procedure maps each decoy to a local minimum of the force field of similar structure, the average change in RMSD between the starting and relaxed conformation was less than 0.02 Å, i.e. the decoys are not changed in the relaxation process.

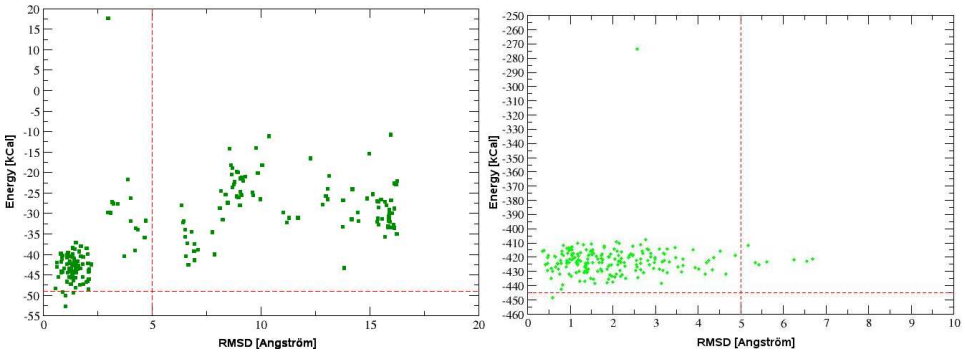

Figure 2. Results for 1PPE (left) and 1KAC (right). Free energy versus bRMSD of all accepted conformations in the simulation (25000 steps, maximum translation shift per move 0.5 Å, maximum rotational shift 0.05 rad).

Using this protocol, both proteins 1PPE (Fig. 1) and 1KAC show very good results between the native structure and the best decoys. The bRMSD-energy plot of all accepted conformations during the simulation (Fig. 2) demonstrates that the simulation explores a wide variety of conformations, with regard to their free-energy and their deviation from the native conformation. The best protocol results in 25000 steps, where we have also included rigid body translational and rotational moves to treat the aggregates. The translational moves are sampled from an equidistributed interval with maximum change of 0.5 Å, whereas the rotational moves are sampled from an equidistributed interval with maximum change of 0.05 Rad.

Acknowledgments

We thank the Leonardo da Vinci project Unipharma Graduates coordinated by Sapienza University of Rome (`www.unipharmagraduates.it`), the DFG (WE 1863/10-2) and the volunteers of POEM@HOME (`http://boinc.fzk.de`) for supporting this work.

References

1. A. Verma, W. Wenzel, *All-atom protein folding using an improved free-energy forcefield*, Biophys. J: Suppl. **S**, 214, (2007).
2. T.Herges, W. Wenzel, *An all-atom forcefield for teriary structure prediction of helical proteins*, Biophys. J **87**, 3100, (2004).
3. R. Chen, W. Tong, J. Mintseris, L. Li *ZDOCK: An Initial-stage Protein-Docking Algorithm*, Proteins **52**, 80–87, (2003).
4. C. B. Anfinsen, *Principles that govern the folding of protein chains.*, Science **181**, 223–230, 1973.
5. L. Li, R. Chen,Z. Weng, *RDOCK: Refinement of Rigid-body Protein Docking Predictions*, Proteins **53**, 693–707, (2003).