Computational simulation for interactions of nano-molecules: The phospho-pivot modeling algorithm for prediction of interactions between a phospho-protein and its receptor

Fumiko Matsuzawa, Shin-ya Ohki, Sei-ich Aikawa, Masumi Eto

Celestar Lexico-Sciences, Inc., MTG D17, Chiba 261-8501, Japan
Japan Advanced Institute of Science and Technology (JAIST), Ishikawa 923-1292, Japan
Center for Cell Signaling and Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Box 800577, Charlottesville, VA 22908-0577, USA

Received 10 November 2004; revised 8 March 2005; accepted 5 April 2005
Available online 28 June 2005

Abstract
Direct docking of nanomolecules, such as proteins, is responsible for biological signal transduction in cells. This physiological interaction is mimicked by various biosensors in nanotechnology. In many cases phosphorylation of protein is involved in protein–protein interaction, and understanding phosphorylation-dependent interaction is necessary to design novel biosensors. Here, we developed and tested a specific method for studying on interaction of phospho-proteins in silico. The algorithm, named phospho-pivot modeling, consists of two parts: first is to generate a library of virtual complexes by pivoting phospho-ligand at the docking site on the receptor, and second is to grade them according to probability in atomic proximity between two molecules. After a 90-min computation by a personal computer, the phospho-pivot modeling yielded an in silico model for the complex of Ser/Thr phosphatase-1 (PP1) and calyculin A, an inhibitory compound of PP1, which was superimposed on the crystal structure in database with r.m.s.d. of 0.23 Å. The phospho-pivot modeling was applied on the prediction for the complex of PP1 and phospho-CPI-17, an inhibitory protein, whose complex structure is unknown. A 1285-min computation selected one converged structure of the PP1-CPI-17 complex out of 186,624 models. The computation time was reduced to 400 min by adding a prescreening process, where virtual complexes with conflicts between main chains were dismissed from the grading process. Thus, phospho-pivot modeling algorithm is sufficient to predict complex structure of proteins, whose monomeric structures have been solved.

Keywords: Biosensor; Computer modeling; Structural biology; Proteomics; Myosin phosphatase; Hypertension; Cerebellar memory

1. Introduction

Biological events are governed by molecular communication network. Physical contact between natural nanomolecules, such as proteins, is highly regulated in TIME and SPACE that maintains entire network of molecular communication. Phosphorylation of Ser/Thr as well as Tyr residues is often found at the molecular interface of signaling proteins, and functions as a joint between molecules. To unravel entire network of molecular communication it is necessary to understand interaction between phospho-protein and its receptor molecule. Besides, the information of the interface between molecules is quite useful to design specific medicine for therapeutics as well as protein-based nanosensors for new technology. Phospho-ester bond is stable enough to mediate signals between molecules in cells, but is too labile to solve 3D-structure by X-ray crystallography or NMR spectroscopy. Recent interactome projects, a part of human genome project, including large-scale yeast-two hybrid screenings, are uncovering entire network of protein-protein interaction in yeast, fruit fly and C. elegans [1–4]. Structural studies on protein-protein interaction are the next step of challenges in structural/functional proteomics.

Computer simulation is one of powerful tools to predict and analyze protein–protein interaction [5,6]. Generally, the first stage of the prediction is to search the docking interface between proteins, where shapes of molecules...
are transformed to voxel grids, and virtual complexes are generated in silico. The second stage is comprised of scoring each interaction of the virtual complex based on predefined parameters, such as van del Waals and/or electrostatic interaction. The initial step of the prediction for the docking site accounts for large portion of calculation, so it requires high-performance computers for modeling. Also, final models largely rely on the initial position and orientation of the ligand at the docking site of the receptor. In case that the docking site between receptor and ligand is known or predicted from experimental data, a specific modeling algorithm can be used for prediction of protein complex structure.

The interaction between Ser/Thr protein phosphatase-1 (PP1) and phospho-CPI-17, a phosphorylation-dependent inhibitory protein, mediates extracellular receptor signals into myosin phosphatase that controls vascular contractility and cerebellar motion memory [7,8]. The phosphate group of CPI-17 directly associates with the catalytic center of PP1 [9]. Therefore, the phosphate of CPI-17 and the catalytic center of PP1 are specified as the docking point. Here, we newly developed modeling algorithm to predict the protein–protein complex using one specified docking site. The simple modeling algorithm, named phospho-pivot modeling, was used to analyze the interaction between PP1 and phospho-CPI-17.

2. Materials and methods

All molecules were treated as rigid bodies during the modeling. The algorithm of phospho-pivot modeling consists of two stages; First to generate a library of virtual complexes by rotating the ligand structure around a phosphorus atom as a pivot, and second to grade models in the library according to the probability of the interface based on the atomic distances. Pivoting process is illustrated in Fig. 1. Initial position of the phospho-ligand was arbitrarily given on the receptor (Fig. 1, left). A rotational axis G was defined passing through the phosphorus atom and the geometrical center of the ligand. The phosphorus atom of the phospho-ligand was positioned at the origin of the arbitrary XYZ rectangular coordinates, where G-axis was initially set on the Z-axis. X, Y and G were used as rotational axes and the initial position was defined as 0°. The ligand was rotated on the Y-axis by the angle β (0 ≤ β < 360°), and on the X-axis by the angle α (0 ≤ α < 180°), with 5° step, that defines angle of G-axis against the receptor. The ligand was then rotated on G-axis by the angle γ (0 ≤ γ < 360°). Coordinate of the virtual complex was recorded in each single rotation, resulting in 36 × 72 × 72 = 186,624 structures. Structures of virtual complexes were transformed into three-dimensional images following the general graphic method [10]. Models in the library were served to grading process. The grading was based on the probability of the structure, counting conflicting atomic pair between two molecules. Conflicting was defined when atomic distance below the threshold; 2.9 Å (C–C), 2.7 Å (C–O), 2.8 Å (C–N), 2.6 Å (O–O), 2.6 Å (O–N), or 2.6 Å (N–N). The threshold was given according to the experimental data of crystal structures in database [11].

The phospho-pivot modeling program was executed on a single 2 GHz processor of Apple Power Mac G5 computer under Mac OS 10.3.4. Energy minimization was performed with permission of structural change of both molecules using the AMBER force fields energy refinement on SYBYL 6.9 (Tripos, Inc., St Louis, MO). The structure was drawn using PyMoL [12].

Coordinates of the crystal structure of PP1-calycin A (PDB: 1IT6) [13] were obtained from Protein Data Bank. For the prediction of PP1-CPI-17 complex, the NMR structure T38D CPI-17 (22–120) (PDB: 1J2N) [14] mimicking phosphorylated conformation was used, because phospho-CPI-17 structure is not available. Asp38 of the T38D-CPI-17 was virtually replaced with phospho-Thr, using SYBYL 6.9/BIOPOLYMER software (Tripos, Inc.). PP1 was used as a receptor, and calycin A and CPI-17 were used as ligands in this study.

3. Results and discussion

3.1. Phospho-pivot modeling of PP1-calycin A complex

Ser/Thr protein phophatase-1 (PP1) consists of an active site with a bimetal center, which traps a phosphate group of substrates via coordinate bonds. A potent inhibitor compound, calycin A, a phospho-polyketide compound from marine sponge, was cocystalized with PP1 [13]. The phosphate of calycin A was found at the active site close proximity to the bimetal center, mimicking that of phospho-substrates on PP1. To validate the phospho-pivot modeling algorithm we applied it on the PP1-calycin A complex. The structure of calycin A was removed from the crystal structure, and then re-positioned onto the PP1 structure, arbitrarily (Fig. 1, left). Calycin A was rotated at the active site using the phosphate as a pivot (Fig. 1, right). The 5° step pivoting generated 186,624 virtual structures. All atomic
distances between two molecules, total 132,272 atoms of PP1 and 56 atoms of calyculin A, were compared with preset threshold values, and atomic pairs closer than threshold were scored as a conflict. Fig. 2 demonstrates correlation between number of conflicts and the r.m.s.d. value of the calyculin A in each model against that in the crystal structure. Average value was plotted against the number of conflict in models.

Table 1
Effects of pivoting step angle on the modeling

| Angle of step (°) | Number of model | Lowest r.m.s.d. (Å) | Computation (min) |
|------------------|----------------|---------------------|-------------------|
|                  | Total          | Conflict ≤ 2        |                   |
| 10               | 23,328         | 1                   | 0.97              |
| 5                | 186,624        | 11                  | 0.23              |
| 3                | 864,000        | 60                  | 0.13              |

Fig. 2. Correlation between number of conflict and r.m.s.d. value against the crystal structure. The r.m.s.d. value of calyculin A atoms between the crystal structure and models was calculated using SYBYL software. Average value was plotted against the number of conflict in models.

Table 2
Parameter of the phospho-pivot modeling

|                       | Non-prescreening | Prescreening |
|-----------------------|------------------|--------------|
|                       | PPI              | Calyculin A  | P-CPI-17      |
|                       | 2362             | 56           | 766           |
| Number of atom        |                  |              |               |
| Number of atomic pair | 132,272          | 1,809,292    | 421,920       |
| Number of graded model| 186,624          | 186,624      | 499           |
| Computation (min)     | 40               | 1285         | 400           |

Calyculin A in the best model was superimposed with that in the crystal structure with r.m.s.d. of 0.23 Å, suggesting that the phospho-pivot modeling is sufficient to predict the complex between phospho-ligand and the receptor. The entire modeling process was finished for 40 min using a personal computer (Table 1). Lower angle value of the step, such as 3°, reduced the lowest r.m.s.d. value of the model in the library, but it increased number of models in the library, which caused over 10-fold more computation time (Table 1). The modeling with 10° angle of the step was done for 11 min and yielded the lowest r.m.s.d. value of 0.97 Å (Table 1). However, the results depended on the initial position of calyculin A. Therefore, we used 5° step of the pivoting for the modeling of PP1–P–CPI-17 complex.

3.2. Phospho-pivot modeling of PP1–P–CPI-17 complex

Calyculin A in the best model was superimposed with that in the crystal structure with r.m.s.d. of 0.23 Å, suggesting that the phospho-pivot modeling is sufficient to predict the complex between phospho-ligand and the receptor. The entire modeling process was finished for 40 min using a personal computer (Table 1). Lower angle value of the step, such as 3°, reduced the lowest r.m.s.d. value of the model in the library, but it increased number of models in the library, which caused over 10-fold more computation time (Table 1). The modeling with 10° angle of the step was done for 11 min and yielded the lowest r.m.s.d. value of 0.97 Å (Table 1). However, the results depended on the initial position of calyculin A. Therefore, we used 5° step of the pivoting for the modeling of PP1–P–CPI-17 complex.

3.3. Prescreening process

To reduce computation time, we added a prescreening process prior to the grading. The prescreening is based on distances between main chain atoms (1172 (PP1)×360 (CPI-17)=total 421,920 combinations). Computation was stopped and skipped to next model, when the first conflict between main chain atoms was detected in a model.
Distance thresholds for main chain atoms were set at 3.5 Å (Cα–Cα) and 2.0 Å (N–O, N–N, O–O). After this process, 499 models (0.3% in the library) with no main-chain conflicts were selected out of 186,624 structures and were then subjected to the grading process. Fig. 3B shows the distribution of the selected 499 models in angle of α, β and γ. A dominant cluster was found in the angle map (Fig. 3B), which corresponds to the core of the major cluster in Fig. 3A. The 18 models with the fewest conflicts, less than 80, were only found in the dominant cluster, indicating one converged orientation of CPI-17 (Fig. 3B, dark blue spot). Thus, the prescreening process effectively selected models with the fewest number of conflicts. Importantly, execution time for the entire modeling was reduced to 400 min, which is 30% of the computation without the prescreening, even though one extra process for the computation was added in Table 2. Elimination of 99.7% models from the grading process was efficient to reduce the computation for protein–protein interaction without changing the results.

3.4. Evaluation of PP1–P–CPI-17 structure

Fig. 4A shows superimposed structures of top 10 PP1–P–CPI-17 models with the fewest conflicts. Angles of these models were strictly limited in narrow range within 10°; 150 ≤ α < 160°, 180 ≤ β < 190°, and 335 ≤ γ < 345° (Table 3), thereby main chains of CPI-17 in top 10 models were converged with average r.m.s.d. of 2.8 Å for whole molecule, and of 1.1 Å for the segment between Gln31–Leu46 including phospho-Thr38. The top 10 models were served to AMBER force field energy refinement. Two salt bridges between PP1 and P-CPI-17, Asp193–Arg44, and Glu274–Arg36 became evident in these refined structures.

### Table 3

| No.   | Model angle (α, β, γ) | Conflict | Energy (kcal/mol) | Distance (Å) |
|-------|-----------------------|----------|------------------|--------------|
|       |                       |          |                  | D193-R44    | E274-R36   |
| 168621| 160, 185, 340         | 59       | -6187            | 2.5          | 2.6        |
| 163437| 155, 185, 340         | 64       | -6031            | 2.7          | 2.6        |
| 163365| 155, 180, 340         | 67       | -6148            | 3.2          | 2.6        |
| 168549| 160, 180, 340         | 67       | -6114            | 3.8          | 2.6        |
| 168620| 160, 185, 335         | 69       | -6146            | 2.6          | 2.6        |
| 158254| 150, 185, 345         | 71       | -6108            | 2.6          | 2.6        |
| 163366| 155, 180, 345         | 71       | -6101            | 3.5          | 2.6        |
| 158182| 150, 180, 345         | 72       | -6154            | 3.1          | 2.6        |
| 163509| 155, 190, 340         | 72       | -6185            | 2.6          | 2.6        |
| 163438*| 155, 185, 345        | 73       | -6221            | 2.6          | 2.6        |

* Indicates the model illustrated in Fig. 2.
residues in PP1 (Asp193, Arg220, Tyr271, Glu274) are essential for the interaction with P-CPI-17 (Fig. 4B). Mutation of PP1 at these residues reduced the affinity with P-CPI-17, supporting the model predicted by phospho-pivot modeling [15].

4. Conclusion

Our phospho-pivot modeling visualized the probable 3D structure of PP1·P–CPI-17 complex in silico, which has not been solved in vitro. Compared with powerful algorithms, such as genetic algorithm [16,17], the phospho-pivot modeling algorithm uses only atomic distance for the evaluation and does not include extensive computational feedback. Advantage of the phospho-pivot modeling is the simplicity on the simulation of the interaction between a phospho-protein and its receptor. This unique simplicity enables us to use a personal computer in our office for the modeling. A half of total proteins, about 25,000 proteins, are detected as phospho-proteins. We are expecting that the phosphate-pivot modeling will be applied on proteins consisting of a phosphate binding site, such as 14–3–3 protein, Tyr-phosphatases, and P-Tyr binding domains in expanding database for protein structure. Furthermore, the algorithm would be applied on the interaction of other nanomolecules, by replacing the pivot from phosphate to other chemical groups, such as an acetyl group. This simple molecular simulation method for nanomolecules will be utilized in development of novel biosensors.

Acknowledgements

This work was supported by a Scientist Development Grant from AHA National Center (to M.E.) and a grant from Center for Nanomaterials and Technology (CNMT), JAIST (to S.O.).

References

[1] P. Uetz, et al., A comprehensive analysis of protein–protein interactions in Saccharomyces cerevisiae, Nature 403 (2000) 623–627.
[2] T. Ito, T. Chiba, R. Ozawa, M. Yoshida, M. Hattori, Y. Sakaki, A comprehensive two-hybrid analysis to explore the yeast protein interactome, Proc. Natl Acad. Sci. USA 98 (2001) 4569–4574.
[3] L. Giot, et al., A protein interaction map of Drosophila melanogaster, Science 302 (2003) 1727–1736.
[4] S. Li, et al., A map of the interactome network of the metazoan C. elegans, Science 303 (2004) 540–543.
[5] G.R. Smith, M.J. Sternberg, Evaluation of the 3D-Dock protein docking suite in rounds 1 and 2 of the CAPRI blind trial, Proteins 52 (2003) 74–79.
[6] A.H. Elcock, D. Sept, J.A. McCammon, Computer simulation of protein–protein interactions, J. Phys. Chem. B 105 (2001) 1504–1518.
[7] T. Kitazawa, A. Polzin, M. Eto, CPI-17-deficient smooth muscle of chicken, J. Physiol. 557 (2004) 515–528.
[8] M. Eto, R. Bock, D. Brautigan, D.J. Linden. Cerebellar long-term synaptic depression requires PKC-mediated activation of CPI-17, a myosin/moesin phosphatase inhibitor. Neuron 36 (2002) 1145–1158.
[9] M. Eto, T. Kitazawa, D.L. Brautigan, Phosphoprotein inhibitor CPI-17 specificity depends on allosteric regulation of protein phosphatase-1 by regulatory subunits, Proc. Natl Acad. Sci. USA 101 (2004) 8888–8893.
[10] J.D. Foley, A. van Dam, S.K. Feiner, J.F. Hughes, Computer Graphics: Principles and Practice in C, second ed., Addison-Wesley, Reading, MA, 1996.
[11] G.N. Ramachandran, V. Sasisekharan, Conformation of polypeptides and proteins, Adv. Protein Chem. 23 (1968) 283–438.
[12] W.L. De Lano, The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, 2002.
[13] A. Kita, S. Matsunaga, A. Takai, H. Kataiwa, T. Wakimoto, N. Fusetai, M. Isebe, K. Miki, Crystal structure of the complex between calyculin A and the catalytic subunit of protein phosphatase 1, Structure (Camb) 10 (2002) 715–724.
[14] S. Ohki, M. Eto, M. Shimizu, R. Takada, D.L. Brautigan, M. Kainosho, Distinctive solution conformation of phosphatase inhibitor CPI-17 substituted with aspartate at the phosphorylation-site threonine residues, J. Mol. Biol. 326 (2003) 1539–1547.
[15] F. Matsuzawa, S. Aikawa, S. Ohki, M. Eto, Phospho-pivot modeling predicts specific interactions of protein phosphatase-1 with a phospho-inhibitor protein CPI-17, J. Biochem. 137 (5) (2005) 633–641.
[16] Y. Maki, et al., An integrated comprehensive workbench for inferring genetic networks: voyagene, J. Bioinform. Comput. Biol. 2 (3) (2004) 533–550.
[17] S. Kimura, et al., Inference of S-system models of genetic networks using a cooperative coevolutionary algorithm, Bioinformatics 21 (7) (2005) 1154–1163.