Improvement of the Protein–Protein Docking Prediction by Introducing a Simple Hydrophobic Interaction Model: An Application to Interaction Pathway Analysis

Masahito Ohue[1,2], Yuri Matsuzaki[1], Takashi Ishida[1], and Yutaka Akiyama[1]

[1] Graduate School of Information Science and Engineering,
Tokyo Institute of Technology, Tokyo, Japan
[2] Research Fellow of the Japan Society for the Promotion of Science
{ohue,y_matsuzaki,t.ishida}@bi.cs.titech.ac.jp,
akiyama@cs.titech.ac.jp

Abstract. We propose a new hydrophobic interaction model that applies atomic contact energy for our protein–protein docking software, MEGADOCK. Previously, this software used only two score terms, shape complementarity and electrostatic interaction. We develop a modified score function incorporating the hydrophobic interaction effect. Using the proposed score function, MEGADOCK can calculate three physicochemical effects with only one correlation function. We evaluate the proposed system against three other protein–protein docking score models, and we confirm that our method displays better performance than the original MEGADOCK system and is faster than both ZDOCK systems. Thus, we successfully improve accuracy without loosing speed.

Keywords: Protein–Protein Docking, MEGADOCK, Hydrophobic Interaction, Fast Fourier Transform, Protein–Protein Interaction.

1 Introduction

Proteins play a key role in virtually all biological events that take place within and between cells. Many proteins display their biological functions by binding to a specific partner protein at a specific site. Determining the structure of a given complex is one of the most important challenges in molecular biophysical research [1, 2]. In addition, the number of protein 3-D structures stored in the Protein Data Bank (PDB) [3] is currently increasing, allowing protein–protein interactions and complex structures to be connected using computational prediction methods, known as the 3-D interactome concept [4]. Against this background, there has been considerable research on protein–protein docking, which is the computational prediction of protein complex structures.

The goal of protein–protein docking is to determine the protein complex structure in atomic detail, starting from the coordinates of the unbound component molecules. Most current docking methods start with rigid-body docking,

T. Shibuya et al. (Eds.): PRIB 2012, LNBI 7632, pp. 178–187, 2012.

which generates a large number of docked conformations (called "decoys") with good surface complementarity. One of the major methods of simulating protein–protein docking is the Katchalski-Katzir algorithm [5], using a 3-D grid representation and fast Fourier transform (FFT) correlation approach. In the Katchalski-Katzir algorithm, the pseudo interaction energy score (called the docking score) between a receptor protein and a ligand protein is calculated by FFT and inverse FFT (IFFT) using a correlation of two discrete functions, as follows:

$$\begin{aligned} S(\boldsymbol{t}) &= \sum_{\boldsymbol{v}\in\mathbb{N}^3} R(\boldsymbol{v})L(\boldsymbol{v}+\boldsymbol{t}) && (1)\\ &= \mathrm{IFFT}[\mathrm{FFT}[R(\boldsymbol{v})]^*\mathrm{FFT}[L(\boldsymbol{v})]], && (2)\end{aligned}$$

where R and L are the discrete score function of the Receptor and Ligand proteins, $\boldsymbol{v}$ is a coordinate in a 3-D grid space $\mathbb{N}^3$, and $\boldsymbol{t}$ is the parallel translation vector of the ligand protein. In order to find the best docking poses, possible ligand orientations are exhaustively examined at n_θ rotation angles for a given stepsize θ. For each rotation, the ligand protein is translated into $N \times N \times N$ patterns in the $\mathbb{N}^3$ grid space (where $N = |\mathbb{N}|$ is the grid size in each dimension). The decoy that yields the highest value of S for each rotation is recorded. In this manner, a total of $n_\theta \times N^3$ docking poses are evaluated for one protein pair. To directly execute the simple convolution sums in eq. (1), $\mathcal{O}(N^6)$ calculations are required; however, this is reduced to $\mathcal{O}(N^3 \log N)$ using the FFT in eq. (2).

There are a number of software packages using the Katchalski-Katzir algorithm [6–12]. Among them, ZDOCK [11, 12] is a widely used protein–protein docking software [13–15]. ZDOCK uses the original docking scores, which are accurate compared to other software. However, this requires two or more correlation function calculations, with a correspondingly large calculation time. Therefore, it is unrealistic to use ZDOCK in a situation where many docking calculations are needed, e.g., when aimed at predictions of a protein–protein interaction network [16–19] or an ensemble/cross-docking performing an all-to-all docking [20–22].

Our protein–protein docking software, MEGADOCK [23, 24], also uses the Katchalski-Katzir algorithm. By employing an original shape complementarity score function (called rPSC) and a general electrostatic interaction score model, MEGADOCK can calculate the docking score with only one correlation function, and thus exhibits quicker calculation times than ZDOCK. Accordingly, the docking prediction accuracy of MEGADOCK is lower than that of ZDOCK. ZDOCK calculates three physico-chemical effects: shape complementarity, electrostatics, and an empirical potential-based desolvation free energy as a hydrophobic effect, with two or more correlation functions. To improve the docking accuracy of MEGADOCK, we intend to incorporate a hydrophobic interaction effect to our scoring model. However, using the conventional score model employed by ZDOCK would cause an increase in the number of correlation functions to be calculated. Therefore, we need a new score model to make MEGADOCK suitable for varied applications.

In this study, we introduce a hydrophobic interaction effect to MEGADOCK. In particular, looking ahead to the application of an interaction network prediction, which is the final goal of MEGADOCK, we develop a simple hydrophobic interaction model that considers only the receptor protein. This increases the performance of the docking calculation without any detrimental effect on the speed.

2 Materials and Methods

2.1 Previous Score Model

In this subsection, we briefly explain our previously developed docking software, MEGADOCK version 2.5. MEGADOCK 2.5 uses a docking score function that combines two terms: the real Pairwise Shape Complementarity (rPSC) score term and the electrostatics (ELEC) score term, which is defined based on the FTDock force model [6] and the CHARMM19 atomic charge [25]. Each pair of proteins is first allocated a position on the 3-D grid space $\mathbb{N}^3$, which has a grid step size of 1.2 Å. Scores are then assigned to each voxel $\boldsymbol{v} \in \mathbb{N}^3$ according to the location in the protein, such as surface or core.

The rPSC term is defined as follows:

$$\mathrm{rPSC}(\boldsymbol{t}) = \sum_{\boldsymbol{v}\in\mathbb{N}^3} G_R(\boldsymbol{v}) G_L(\boldsymbol{v}+\boldsymbol{t}),$$

$$G_R(\boldsymbol{v}) = \begin{cases} \#\text{ of receptor atoms within } (3.6\text{ Å} + r_{\mathrm{vdW}}) & (\text{open space}) \\ -27 \quad (\text{inside of the receptor}), \end{cases} \tag{3}$$

$$G_L(\boldsymbol{v}) = \begin{cases} 1 & (\text{solvent excluding surface layer of the ligand}) \\ 2 & (\text{core of the ligand}), \end{cases} \tag{4}$$

where G_R and G_L represent the rPSC grid value of the receptor/ligand proteins, r_{vdW} represents the van der Waals atomic radius, and $\boldsymbol{t}$ is the ligand translation vector. We omitted the zero value domain.

The ELEC term from FTDock potential is represented as the electric field $\varphi(\boldsymbol{i})$. $\varphi(\boldsymbol{i})$ is assigned to each voxel $\boldsymbol{i} \in \mathbb{N}^3$ as follows:

$$\varphi(\boldsymbol{i}) = \sum_{\boldsymbol{j}\in\mathbb{N}^3} \frac{q(\boldsymbol{j})}{\varepsilon(r_{\boldsymbol{ij}}) r_{\boldsymbol{ij}}}, \quad \varepsilon(r) = \begin{cases} 4 & (r \le 6\text{ Å}) \\ 38r - 224 & (6\text{ Å} < r < 8\text{ Å}) \\ 80 & (8\text{ Å} \le r), \end{cases}$$

where $q(\boldsymbol{j})$ is the charge at grid point $\boldsymbol{j} \in \mathbb{N}^3$, $r_{\boldsymbol{ij}}$ is the Euclid distance between grid points $\boldsymbol{i}$ and $\boldsymbol{j}$, and $\varepsilon(r)$ is a distance-dependent dielectric function. ELEC term is defined as follows:

$$\mathrm{ELEC}(\boldsymbol{t}) = \sum_{\boldsymbol{v}\in\mathbb{N}^3} E_R(\boldsymbol{v}) E_L(\boldsymbol{v}+\boldsymbol{t}),$$

$$E_R(\boldsymbol{v}) = \varphi(\boldsymbol{v}) \quad (\text{open space}),$$

$$E_L(\boldsymbol{v}) = q(\boldsymbol{v}),$$

Table 1. Non-pairwise ACE scores. This table is reproduced from Table 1 of [26] in which Zhang, *et al.* defined the atom types and assigned ACE scores.

atom type	N	C^α	C	O	GC^α	C^β	KN^ζ	KC^δ	DO^δ
ACE score	−0.495	−0.553	−0.464	−0.079	0.008	−0.353	1.334	1.046	0.933
atom type	RN^η	NN^δ	RN^ε	SO^γ	HN^ε	YC^ζ	FC^ζ	LC^δ	CS^γ
ACE score	0.726	0.693	0.606	0.232	0.061	−0.289	−0.432	−0.987	−1.827

where E_R and E_L represent the ELEC grid values of receptor/ligand proteins, determined according to the charge of each voxel $q(\boldsymbol{v})$ in which atoms in the residues are assigned a charge according to CHARMM19.

Considering these two terms, the docking score $S(\boldsymbol{t})$ is represented as:

$$
\begin{aligned}
R(\boldsymbol{v}) &= G_R(\boldsymbol{v}) + iE_R(\boldsymbol{v}),\\
L(\boldsymbol{v}) &= G_L(\boldsymbol{v}) + iw_eE_L(\boldsymbol{v}),\\
S(\boldsymbol{t}) &= \Re\left[\sum_{\boldsymbol{v}\in\mathbb{N}^3} R(\boldsymbol{v})L(\boldsymbol{v}+\boldsymbol{t})\right] = \mathrm{rPSC}(\boldsymbol{t}) - w_e\mathrm{ELEC}(\boldsymbol{t}),
\end{aligned}
$$

where w_e is the weight parameter of ELEC term.

2.2 Proposed Method

In our proposed method, we used a non-pairwise-type atomic contact energy (ACE) score [26] to incorporate a hydrophobic interaction effect. For the current study, we introduce a simple model that considers only the receptor protein because, when both the receptor and ligand are taken into consideration, an increase in the number of correlation functions is unavoidable.

We modify the receptor rPSC value G_R in eq. (3) in order to introduce the ACE score. The new receptor value G'_R is defined as follows:

$$
\begin{aligned}
G'_R(\boldsymbol{v}) &= G_R(\boldsymbol{v}) + w_hH_R(\boldsymbol{v}),\\
H_R(\boldsymbol{v}) &= \begin{cases} \text{sum of ACE scores of receptor atoms} & \\ \qquad\qquad\qquad\qquad \text{within } (3.6\ \text{Å} + r_{\mathrm{vdW}}) & \text{(open space)}\\ 0 \quad \text{(inside of the receptor)}, & \end{cases}
\end{aligned}
$$

where w_h is the weight parameter of H_R. Fig. 1 shows a pattern diagram of the proposed model. We use the ACE values given in Table 1.

This score model attains a value of $G_R(\boldsymbol{v}) + w_hH_R(\boldsymbol{v})$ when the open space near the receptor surface is superposed on the ligand surface. The score of a ligand core of 2 depends on the penalty (−54) at the time of a core collision for enlargement. It is assumed that $2 \times \{G_R(\boldsymbol{v}) + w_hH_R(\boldsymbol{v})\}$ will be obtained by the ligand core, depending on its position, under a situation where the core moves into a pocket that can obtain a high score, because a penalty (−27) is imposed on any collision between the ligand surface and a receptor. Therefore, we do not

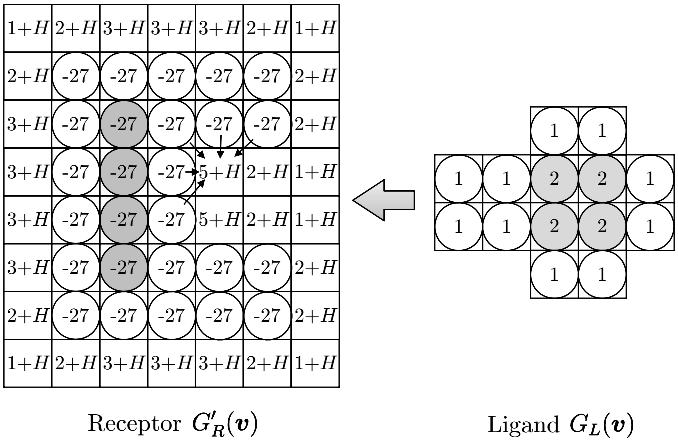
1+H 2+H 3+H 3+H 3+H 2+H 1+H
2+H -27 -27 -27 -27 -27 2+H
3+H -27 -27 -27 -27 -27 2+H
3+H -27 -27 -27 5+H 2+H 1+H
3+H -27 -27 -27 5+H 2+H 1+H
3+H -27 -27 -27 -27 -27 2+H
2+H -27 -27 -27 -27 -27 2+H
1+H 2+H 3+H 3+H 3+H 2+H 1+H
1 1
1 1 2 2 1
1 1 2 2 1
1 1
Receptor $G'_R(\boldsymbol{v})$
Ligand $G_L(\boldsymbol{v})$

Fig. 1. Proposed scoring model $G'_R(\boldsymbol{v})$ and $G_L(\boldsymbol{v})$. The model consists of 3-D grid, but here we show only two dimensions for simplicity. For clarity, grid points with a value of 0 have been omitted. Small arrows indicate the five atoms that are within the cutoff distance of a grid, and thus contribute to its score of $5 + H$, where H means $w_h H_R(\boldsymbol{v})$.

consider this situation to affect the good docking pose of the decoy. ZDOCK 2.3 [11] uses two correlation functions, and ZDOCK 3.0 [12] uses eight correlation functions to consider three effects—shape complementarity, electrostatics, and desolvation free energy—our score model can calculate docking scores under consideration of three effects with only one correlation function, while maintaining an advantage in terms of calculation speed.

2.3 Dataset

The protein complex structures used in this study were retrieved from a standard protein–protein docking benchmark set [27], containing 176 known 3-D structures of complex component proteins in both bound and unbound forms.

2.4 Evaluation of Docking Performance

To evaluate the docking pose prediction performance, we conducted a re-docking and unbound docking experiment using the benchmark dataset. We used the root mean square deviation (RMSD) of the ligand (L-RMSD), which is the RMSD of the predicted ligand position and that of the crystal complex structure calculated for all the atoms when the receptor positions are superimposed, in order to determine the accuracy of the docking predictions. The RMSDs of the unbound structures were only calculated for residues that were aligned by pairwise alignment of the amino acid sequences between the bound and unbound structures. We defined a “near-native decoy” as that for which L-RMSD was less than or equal to 5 Å. We compared the performance of the following docking methods: the proposed method, MEGADOCK 2.5, ZDOCK 2.3, and ZDOCK 3.0. For comparison with ZDOCK, we set parameters of 3,600 decoys per case and $\theta = 15°$ for the ligand rotation step. We compared the following widely used two values [1, 11, 12] to determine the docking performance:

- **Average Hit Count:** The average number of near-native decoys across the set of cases for a given number of top-ranked predictions per test case.
- **Success Rate:** The percentage of cases with near-native decoys for a given number of top-ranked predictions per test case.

3 Results and Discussions

3.1 Optimization of Weight Parameters

For determining parameter values w_e and w_h, we used only the bound dataset to avoid overfitting the unbound structures. We optimized the parameters for maximizing the Success Rate of 100 predictions. We searched the best combination of w_e and w_h, and tested w_e from 0.5 to 1.5 by 0.05 steps and w_h from 0.1 to 2.0 by 0.1 steps. As a result, we found the best values of $w_e = 1.15$ and $w_h = 0.6$.

3.2 Docking Prediction Accuracy

The Average Hit Count is shown in Fig. 2 since bound dataset was used for optimization of weight parameters, the results of unbound dataset are more important than bound dataset. We can see that our proposed method performed better than MEGADOCK 2.5 with both the bound and unbound sets. In addition, the proposed method displays an equivalent performance to ZDOCK 2.3 for the unbound set and is broadly similar for the bound set. However, our method is still less accurate than ZDOCK 3.0 for both sets. The performance of ZDOCK 3.0 is mainly due to its pairwise potential function, although this performance is obtained at the expense of calculation speed.

A similar trend is observed in the Success Rate of each method, as shown in Fig. 3. We see that the Success Rate of our proposed method is again better than that of MEGADOCK 2.5 for both sets. However, our proposed method is noticeably worse than ZDOCK 2.3 for the bound set. We think that G_R and H_R require further tuning using more complex structures in the PDB.

3.3 Calculation Time

Table 2 shows the average computation time for the benchmark dataset. All the calculations were conducted on the TSUBAME 2.0 supercomputing system, Tokyo Institute of Technology, Japan, which consists of two Intel Xeon 2.93 GHz (6 cores × 2) processors and 32 GB RAM, with operational nodes connected via an InfiniBand and Gigabit Ethernet. An average of 14.2 min was required for

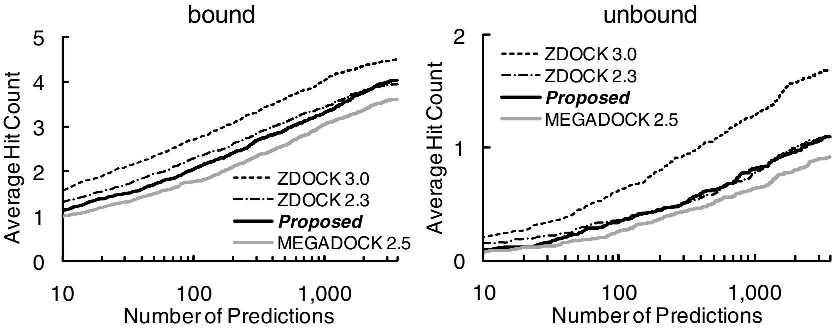

Fig. 2. Average Hit Count for all test cases of benchmark dataset. The Average Hit Count was defined as the average number of near-native decoys across the set of cases for a given number of top-ranked predictions per test case.

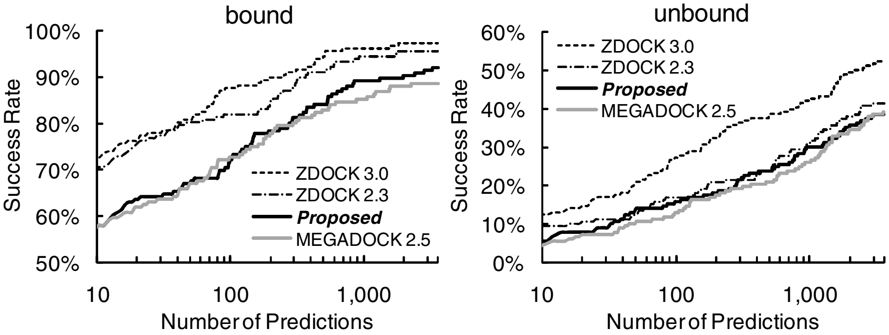

Fig. 3. Success Rate for all test cases of benchmark dataset. The Success Rate was defined as the percentage of cases with near-native decoys for a given number of top-ranked docking predictions per test case.

Table 2. Total time for 176 docking calculations using the benchmark dataset

	Proposed	MEGADOCK 2.5	ZDOCK 2.3	ZDOCK 3.0
time (hr)	41.7	41.6	157.3	365.6
speedup from ZDOCK 2.3	3.77	3.78	(1.0)	0.43
speedup from ZDOCK 3.0	8.77	8.79	2.32	(1.0)

each docking calculation using one CPU core. The proposed method obtained the almost same calculation speed as MEGADOCK 2.5 (only 0.7% of calculation time increase), some 3.8 times faster than ZDOCK 2.3 and 8.8 times faster than ZDOCK 3.0. Since FFT takes most of the execution time of MEGADOCK and the proposed method, if we increase the correlation function to 2 or 3 to get better performance of docking, calculation time will also increase 2- or 3-fold.

3.4 Application to Pathway Analysis

We also performed a case study using a biological interaction network by applying our proposed docking method to the protein–protein interaction prediction problem of bacterial chemotaxis pathways, which represents a typical target of signal transduction in the field of systems biology [28]. Docking and protein–protein interaction prediction were undertaken for 101 × 101 = 10,201 pairs corresponding to the constituent protein data of the 13 protein species present in the chemotaxis pathway [17].

We used the method of Matsuzaki *et al.* [17], with the improved MEGADOCK in place of ZDOCK 3.0. The docking score of 101 × 101 combinations was calculated for 101 protein structures and their affinity scores based on the literature [17]. We obtained an F-measure of 0.45 for this system, which is similar to that found in the previous study using ZDOCK 3.0 (F-measure of 0.49).

4 Conclusion

In this study, we added a hydrophobic interaction model to the protein docking software MEGADOCK. This additional component, which considers only the receptor protein, was combined with the considerations of shape complementarity and electrostatic interaction without increasing the calculation time. The proposed method succeeded in achieving the better level of accuracy as previous MEGADOCK. Although we need more better level of accuracy in bound cases, the proposed method achieved the same level of accuracy as ZDOCK 2.3 in unbound cases. It was also 3.8 times faster than ZDOCK 2.3 and 8.8 times faster than ZDOCK 3.0. However, to enhance the accuracy of the proposed model, further tuning of some system parameters is necessary in future. ACE was introduced only into the receptor side in the study because receptor term of rPSC was easy of introducing some atomic effects. We are attempting to develop a new score model with both receptor and ligand ACE term using only one correlation function. Additionally, we will apply our method to other large analyses, such as the interaction network prediction problem of other biological systems or the cross-docking of ensemble structures.

Acknowledgements. This work was supported in part by a Grant-in-Aid for JSPS Fellows (23·8750), a Grant-in-Aid for Research and Development of The Next-Generation Integrated Life Simulation Software, and Education Academy of Computational Life Sciences, all from the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT).

References

1. Pons, C., Grosdidier, S., Solernou, A., Pérez-Cano, L., Fernández-Recio, J.: Present and future challenges and limitations in protein–protein docking. Proteins 78(1), 95–108 (2010)
2. Wass, M.N., David, A., Sternberg, M.J.E.: Challenges for the prediction of macromolecular interactions. Curr. Opin. Struct. Biol. 21(3), 382–390 (2011)
3. Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., et al.: The Protein Data Bank. Nucleic Acids Res. 28(1), 235–242 (2000)
4. Stein, A., Mosca, R., Aloy, P.: Three-dimensional modeling of protein interactions and complexes is going ‘omics. Curr. Opin. Struct. Biol. 21(2), 200–208 (2011)
5. Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A.A., Aflalo, C., et al.: Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. Proc. Natl. Acad. Sci. USA 89, 2195–2199 (1992)
6. Gabb, H.A., Jackson, R.M., Sternberg, M.J.E.: Modelling protein docking using shape complementarity, electrostatics and biochemical information. J. Mol. Biol. 272(1), 106–120 (1997)
7. Vakser, I.A.: Evaluation of GRAMM low-resolution docking methodology on the hemagglutinin-antibody complex. Proteins (suppl. 1), 226–230 (1997)
8. Mandell, J.G., Roberts, V.A., Pique, M.E., Kotlovyi, V., Mitchell, J.C., et al.: Protein docking using continuum electrostatics and geometric fit. Protein Eng. 14(2), 105–113 (2001)
9. Cheng, T.M.-K., Blundell, T.L., Fernández-Recio, J.: pyDock: electrostatics and desolvation for effective scoring of rigid-body protein–protein docking. Proteins 68(2), 503–515 (2007)
10. Kozakov, D., Brenke, R., Comeau, S.R., Vajda, S.: PIPER: an FFT-based protein docking program with pairwise potentials. Proteins 65(2), 392–406 (2006)
11. Chen, R., Li, L., Weng, Z.: ZDOCK: an initial-stage protein-docking algorithm. Proteins 52(1), 80–87 (2003)
12. Mintseris, J., Pierce, B., Wiehe, K., Anderson, R., Chen, R., et al.: Integrating statistical pair potentials into protein complex prediction. Proteins 69(3), 511–520 (2007)
13. Hwang, H., Vreven, T., Pierce, B.G., Hung, J.-H., Weng, Z.: Performance of ZDOCK and ZRANK in CAPRI rounds 13–19. Proteins 78(15), 3104–3110 (2010)
14. Uchikoga, N., Hirokawa, T.: Analysis of protein–protein docking decoys using interaction fingerprints: application to the reconstruction of CaM-ligand complexes. BMC Bioinformatics 11(236) (2010)

15. Fleishman, S.J., Whitehead, T.A., Strauch, E.-M., Corn, J.E., Qin, S., et al.: Community-wide assessment of protein-interface modeling suggests improvements to design methodology. J. Mol. Biol. 414(2), 289–302 (2011)
16. Wass, M.N., Fuentes, G., Pons, C., Pazos, F., Valencia, A.: Towards the prediction of protein interaction partners using physical docking. Mol. Syst. Biol. 7(469) (2011)
17. Matsuzaki, Y., Matsuzaki, Y., Sato, T., Akiyama, Y.: In silico screening of protein–protein interactions with all-to-all rigid docking and clustering: an application to pathway analysis. J. Bioinform. Comput. Biol. 7(6), 991–1012 (2009)
18. Tsukamoto, K., Yoshikawa, T., Hourai, Y., Fukui, K., Akiyama, Y.: Development of an affinity evaluation and prediction system by using the shape complementarity characteristic between proteins. J. Bioinform. Comput. Biol. 6(6), 1133–1156 (2008)
19. Yoshikawa, T., Tsukamoto, K., Hourai, Y., Fukui, K.: Improving the accuracy of an affinity prediction method by using statistics on shape complementarity between proteins. J. Chem. Inf. Model. 49(3), 693–703 (2009)
20. Chaleil, R.A.G., Tournier, A.L., Bates, P.A., Kro, M.: Implicit flexibility in protein docking: Cross-docking and local refinement. Proteins 69(4), 750–757 (2007)
21. Dobbins, S.E., Lesk, V.I., Sternberg, M.J.E.: Insights into protein flexibility: The relationship between normal modes and conformational change upon protein–protein docking. Proc. Natl. Acad. Sci. USA 105(30), 10390–10395 (2008)
22. Venkatraman, V., Ritchie, D.W.: Flexible protein docking refinement using posedependent normal mode analysis. Proteins 80(9), 2262–2274 (2012)
23. Ohue, M., Matsuzaki, Y., Matsuzaki, Y., Sato, T., Akiyama, Y.: MEGADOCK: an all-to-all protein–protein interaction prediction system using tertiary structure data and its application to systems biology study. IPSJ TOM 3(3), 91–106 (2010) (in Japanese)
24. Ohue, M., Matsuzaki, Y., Akiyama, Y.: Docking-calculation-based method for predicting protein-RNA interactions. Genome Inform. 25(1), 25–39 (2011)
25. Reiher III, W.H.: Theoretical studies of hydrogen bonding. Ph.D. Thesis at Harvard University (1985)
26. Zhang, C., Vasmatzis, G., Cornette, J.L., DeLisi, C.: Determination of atomic desolvation energies from the structures of crystallized proteins. J. Mol. Biol. 267(3), 707–726 (1997)
27. Hwang, H., Vreven, T., Janin, J., Weng, Z.: Protein–protein docking benchmark version 4.0. Proteins 78(15), 3111–3114 (2010)
28. Baker, M.D., Wolanin, P.M., Stock, J.B.: Systems biology of bacterial chemotaxis. Curr. Opin. Microbiol. 9(2), 187–192 (2006)