IPSJ Transactions on Bioinformatics Vol. 3 10–23 (Mar. 2010)

Original Paper

Improved Prediction Method for Protein Interactions Using Both Structural and Functional Characteristics of Proteins

Tatsuya Yoshikawa,[†1] Shigeto Seno,[†1]
Yoichi Takenaka[†1] and Hideo Matsuda[†1]

To identify protein–protein interaction pairs with high accuracy, we propose a method for predicting these interactions based on characteristics obtained from protein–protein docking evaluations. Previous studies assumed that the required protein affinity strength for an interaction was not dependent on protein functions. However, the protein affinity strength appears to differ with different docking schemes, such as rigid-body or flexible docking, and these schemes may be related to protein functions. Thus, we propose a new scoring system that is based on statistical analysis of affinity score distributions sampled by their protein functions. As a result, of all possible protein pair combinations, a newly developed method improved prediction accuracy of F-measures. In particular, for bound antibody–antigen pairs, we obtained 50.0% recall (= sensitivity) with higher F-measures compared with previous studies. In addition, by combining two proposed scoring systems, Receptor-Focused Z-scoring and Ligand-Focused Z-scoring, further improvement was achieved. This result suggested that the proposed prediction method improved the prediction accuracy (i.e., F-measure), with few false positives, by taking biological functions of protein pairs into consideration.

1. Introduction

Most biological functions involve interactions between several proteins in a cell. Therefore, it is important to elucidate biological phenomena, including cell signaling, enzyme reactions, and gene expression regulation, by analysis of protein–protein interactions (PPIs). In the past, several reviews [1)–3)] have examined the interaction characteristics of known protein complexes and PPI maps, which can play a role in discovering protein partners. Well-known experimental techniques have been developed for the systematic analysis of PPIs, including yeast two-hybrid-based methods [4)], protein-fragment complementation assay [5)–11)], and mass spectrometric identification of isolated protein complexes [12),13)] and protein chips [14)]. Several methods have been developed in computational studies for predicting PPIs using genomic information [15),16)].

Recently, research on PPIs has focused on not only ascertaining their roles in living organisms but also applying this knowledge to medicinal fields such as drug design. In essence, the interactive properties associated with protein conformation need to be thoroughly investigated if we are to analyze the relationships between protein structure and function. Ideally, PPI studies for drug design should be based on protein structural information. Structure-based PPI study includes protein–protein docking. Most previous studies on docking have attempted to solve protein-docking problems; the ultimate objective of these studies was to accurately predict the structures of protein complexes from three-dimensional (3D) structures of individual proteins.

Since accurate prediction involves searching among numerous possible protein–protein docking conformations in 3D space, it is necessary to develop an efficient method for reducing the search space. To identify structures that are most likely to occur in nature, plausible candidates from docking must be ranked using scoring functions. Katchalski-Katzir and Vakser et al. [17)] and subsequently the programs FTDock [18)], 3D-Dock [19)], GRAMM [20)], DOT [21)], ZDOCK [22)], and HEX [23)] pioneered a fast Fourier transform (FFT)-based method for searching rapidly in order to maximize shape and chemical complementarity between a given pair of interacting proteins. Other docking programs such as ROSETTA [24)] and PatchDock [25)] search in a confined space using Monte Carlo methods and Geometric hashing [26)]. Other than shape complementarity, these rigid-body dockings are adequate for various searching methods and scoring functions such as electrostatic and desolvation energies.

In this study, given that protein docking has the potential to decide whether or not a complex actually occurs in nature and for measuring its affinity, we focused on the 3D structures of proteins in order to calculate PPIs via protein–protein docking. This kind of PPI study has been previously discussed by Smith, et al. [27)], although they did not clearly demonstrate the functioning of an actual system

†1 Graduate School of Information Science and Technology, Osaka University

or analysis of results. Recent studies [28)–31)] involved in predicting PPIs on the basis of shape complementarity docking succeeded in up to 23 of 84 predictions. Sacquin-Mora, et al. [32)] successfully predicted 7 out of 10 interaction partners using weighted interaction energies. However, methods for determining protein affinity of only one complex pair have been confined to forecasting a local binding likelihood for each protein–protein pair, therefore have not been considered that the required strength of binding is probably different by biological functions.

In order to solve this problem, we focused on assessing protein interactions by globally considering affinities of one protein with other proteins. In this study, we propose a docking-based prediction method for PPIs by using all possible protein pair combinations. The aim of this study is to improve the prediction accuracy of PPIs.

2. Protein–Protein Interaction Problem

The problem of predicting the interactions between two proteins, designated as a receptor and a ligand (see Section 2.2.3 for the definitions of receptor and ligand), was considered.

2.1 Protein–Protein Interaction Prediction Method

Most of the previous studies in computational approach developed the genome-scale techniques, whereas recently the structure-based studies are considered as an effective mean of applying to drug design. This approach is based on predicting PPIs by protein–protein docking. This method consists of three procedures that are outlined in **Fig. 1**. 3D structures of a receptor and ligand are the input data flow, and the predicted result of the interaction is the output. The squares with "Protein Docking," "Affinity Evaluation," and "Interaction Prediction" are the key procedures, and each input/output port is shown in parentheses.

2.1.1 Protein–Protein Docking

Protein–protein docking is performed for calculating the 3D structure of a protein complex, starting from individual structures of constituent proteins. That is, the general aim of this study is to predict the near-native complex structure of two proteins, which is different from that of a PPI prediction problem that determines whether or not the proteins interact. The reason for considering protein docking as a procedure in a PPI prediction problem is based on the following

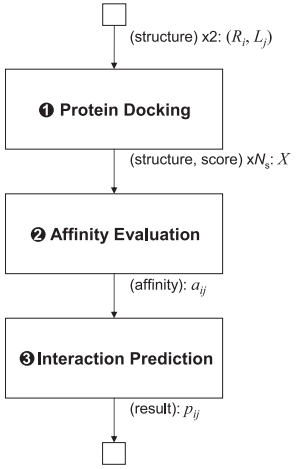

Fig. 1 Flowchart for PPI prediction by protein–protein docking.

assumption. It is assumed that protein affinity plays a role in deciding whether or not proteins interact only when their binding affinity can be calculated accurately. That is, if the affinity of proteins is equal to or more than a certain threshold, then they can be computationally regarded as a protein pair that interacts. Based on this assumption, PPI prediction can be realized by statistically analyzing docking scores from protein docking programs.

A rigid-body model for expressing molecules is exhaustively screened in a six-dimensional rotation and translation space. The procedure starts by rotating a ligand as a probe protein. When the rotational angular step Δ is equal to the widely used 15 degrees, λ (= 3,600) poses are sampled because of rotational symmetry [33)]. A target protein (receptor) and 3,600 probe proteins (ligands) are discretized into 3D grids n^3 by a certain grid pitch γ (e.g., 1.2 Å), where n is the number of grid points in each coordinate. Each grid point is assigned to a structural and chemical property value based on geometric characteristics (i.e., core, surface and cavity areas) and free energies, respectively. When the rotated ligand is translated with respect to the receptor, the docking algorithm calculates the product sum of assigned property values, which is referred to as the docking score (s). Given the grid size of a receptor n^3 and the number of

sampled ligand poses λ, λn^3 docking scores can be obtained by an exhaustive search for only one protein pair. The top-ranked N_s ($N_\mathrm{s} \geq 1$) scores are generally sampled as candidates because of a large number of docking results. The results $X = \{x_k \mid 1 \leq k \leq N_\mathrm{s}\}$ include not only the docking scores s_k but also the structural information of complex candidates:

$$x_k = \big(s_k, t_k\left(t_x, t_y, t_z\right), r_k\left(r_\theta, r_\phi, r_\psi\right)\big), \tag{1}$$

where $t_k(t_x, t_y, t_z)$ and $r_k(r_\theta, r_\phi, r_\psi)$ are the translational distance and the rotational angle of the ligand, respectively.

2.1.2 Affinity Evaluation

Based on the docking results, affinity evaluation plays a role in calculating the binding likelihood. The aim is to assess how strong the interactions are. The simplest way is to utilize the maximum value among N_s docking scores as protein affinity such that

$$a = max\big(s_k\big)_{k=1}^{N_\mathrm{s}}. \tag{2}$$

An alternative is to use the statistical characteristics provided by clustering the docking results according to the score or structural similarities between candidates.

2.1.3 PPI Prediction

Interaction prediction makes the final decision as to whether or not proteins interact. The primitive threshold-based approach is used to determine the interactions as follows:

$$p = \begin{cases} 1 & \text{if } a \geq \tau \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

where p is the prediction result that includes the Boolean values (i.e., 1 and 0 indicating positive and negative PPI, respectively), and τ is the affinity threshold for deciding whether or not proteins interact. Here, the threshold τ has to be decided using biological knowledge or statistical characteristics of affinity values, etc.

2.2 Previous Studies

The outlines of previous PPI prediction methods are described based on protein docking evaluations.

2.2.1 Outline of ZDOCK

A previous method for PPI prediction, which we have designated as ZDOCK hereafter, uses docking scores from the protein–protein docking program ZDOCK 3.0.1 [34]. This tentative method was used for comparing the prediction method in Section 2.2.2. The ZDOCK 3.0.1 program can assess structural and chemical complementarity between proteins. It enables us to find binding sites and complex structures using a FFT-based search algorithm with a scoring function that is based on pairwise shape complementarity, electrostatics, and explicit interface atomic contact energies. In affinity evaluation, the maximum score from only one docking simulation among 2,000 docking scores was used to simply assess the affinity of a protein pair. The setting values defined in Section 2.1.1 are as follows: $\Delta = 15$, $\gamma = 1.2$, and $N_\mathrm{s} = 2{,}000$, where Δ is the rotational angular step, γ is the grid pitch, and N_s is the number of samplings for candidates. Here, when applying the ZDOCK 3.0.1 program execution option, all default values (i.e., −N (= 2,000), −S (= no at randomization), and −D (= none)) were used. The threshold τ was determined to maximize the F-measure by ROC analysis described in Section 3.2.3.

2.2.2 Outline of Affinity Evaluation and Prediction (AEP)

A previous study [31] had predicted the interactions by assessing the statistical significance of binding likelihood based on shape complementarity characteristics between protein pairs. **Figure 2** shows the flowchart of this method called *affinity evaluation and prediction* (AEP). In their protein–protein docking procedure, the original docking program with a scoring function of pair-wise shape complementarity was developed. The improvements for the trade-off between time and accuracy, which are important for protein docking, are as follows:

(1) To improve prediction accuracy, protein modeling was controlled by strictly tracing the concavo-convex shape and reducing surface thickness.
(2) To improve computational time, a high-performance FFT library (CONV3D) [35] was used for the actual complementarity search on massively parallel computers.

Here, setting values were $\Delta = 15$, $\gamma = 1.2$, and $N_\mathrm{s} = 512$. In the affinity evaluation procedure, the statistical characteristics of N_s candidates were used

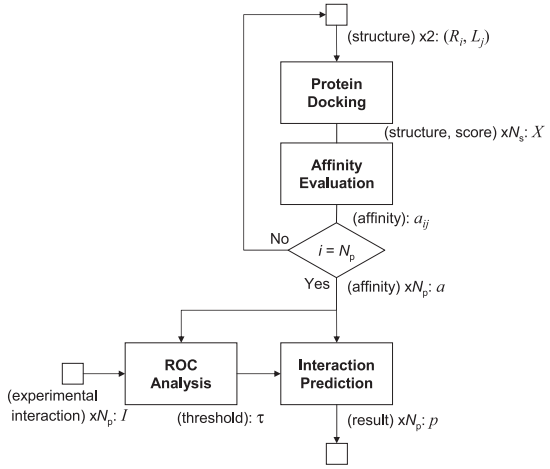

Fig. 2 Flowchart of a previous prediction method (i.e., AEP).

to assess protein affinity. The docking scores were classified into several clusters according to the structural similarities between candidates. Protein affinity in previous studies was calculated by the distribution of docking scores of a representative of each cluster and that of the cluster density. Key parameter settings were also important for determining structural similarities and extracting cluster characteristics in this procedure. In the PPI prediction procedure, the affinity threshold τ was obtained by the receiver operating characteristics (ROC) procedure. The values of key parameters in affinity evaluation and τ are optimized so that the F-measure is maximized.

Using the above procedures, the previous study successfully predicted 23 interaction pairs out of 84. The study assumed that the required protein affinity strength for the interaction was not dependent on the function of proteins. However, the protein affinity strength appears to differ by different docking schemes (i.e., rigid-body or flexible docking), and the scheme may be related to protein functions.

2.2.3 Protein-Pair Data Set

In order to evaluate the performances of ZDOCK and AEP, 168 bound proteins derived from 84 co-crystallized complex structures by Protein–Protein Docking Benchmark 2.0 [36),37)] were used. All proteins were classified as either receptors (R) or ligands (L) according to Weng's dfinition, which resulted in 84 proteins of each type. When the receptor and ligand were derived from the same complex, the receptor molecule was always larger than the ligand. Here, these 84 complexes were classified as follows: 10 pairs of antibody–antigen (functional category A), 12 pairs of bound antibody–antigen (AB), 23 pairs of enzyme–inhibitor or substrate (E), and others (O). In addition, the complexes can be classified according to conformational changes as follows: 63 pairs of rigid-body (conformational category RB), 13 pairs of medium difficulty (MD), and 8 pairs of high difficulty (HD), as shown in **Table 1**. Each complex was assigned with both a protein data bank (PDB) ID and an index from 1 to 84.

To evaluate the prediction method more exhaustively, $N_{\rm p}$ ($N_{\rm R} \times N_{\rm L} = 7{,}056$) possible pairs were constructed by combining $N_{\rm R}$ (= 84) receptors and $N_{\rm L}$ (= 84) ligands, where $N_{\rm p}$ is the total number of receptor–ligand pairs, and $N_{\rm R}$ and $N_{\rm L}$ are the numbers of receptors and ligands, respectively. Therefore, the data set included 84 pairs that have been previously identified experimentally as forming complexes and 6,972 others. Because receptor–receptor or ligand–ligand protein interactions are believed not to occur under normal biological conditions, we only employed receptor–ligand pairs. Here, in order to directly compare our method with previous results, the bound structures of proteins were used.

2.3 Performance Measures

All protein pairs in the data set were classified into either 84 pairs whose interactions have been previously experimentally detected or 6,972 others whose interactions have not been detected. Using the binary values of 1 or 0, the prediction results suggested whether or not proteins interacted, indicating positive or negative for PPIs. **Table 2** shows the logical combinations of experimental interactions and prediction results as a 2×2 contingency table. The four logical combinations (i.e., TP, FN, FP, and TN) are defined in Table 2, and these numbers are represented as tp, fn, fp, and tn, respectively.

In general, performance measures for information retrieval are used for assessing the prediction accuracy of a binary classification problem. Many of the different measures, such as sensitivity ($sens$), recall (rec), precision ($prec$), and F-measure (F), are given by

Table 1 Protein-pair data set consisting of 84 receptors and 84 ligands. "#" indicates the index number from 1 to 84 corresponding to a PDB ID of each complex, "Cat." indicates the conformational category (i.e., RB, MD, and HD).

#	PDB ID	Cat.	Receptor	Ligand
	Antibody–Antigen (10)			
1	1AHW	RB	Fab 5g9	Tissue factor
2	1BGX	MD	Fab	Taq polymerase
3	1BVK	RB	Fv Hulys11	HEW lysozyme
4	1DQJ	RB	Fab Hyhel63	HEW lysozyme
5	1E6J	RB	Fab	HIV-1 capsid protein p24
6	1JPS	RB	Fab D3H44	Tissue factor
7	1MLC	RB	Fab44.1	HEW lysozyme
8	1VFB	RB	Fv D1.3	HEW lysozyme
9	1WEJ	RB	Fab E8	Cytochrome C
10	2VIS	RB	Fab	Flu virus hemagglutinin
	Bound Antibody–Antigen (12)			
11	1BJ1	RB	Fab	vEGF
12	1FSK	RB	Fab	Birch pollen antigen Bet V1
13	1I9R	RB	Fab	Cd40 ligand
14	1IQD	RB	Fab	Factor VIII domain C2
15	1K4C	RB	Fab	Potassium Channel Kcsa
16	1KXQ	RB	camel VHH	Pancreatic alpha-amylase
17	1NCA	RB	Fab	Flu virus neuraminidase N9
18	1NSN	RB	Fab N10	Staphylococcal nuclease
19	1QFW	RB	Fv	Human chorionic gonadotropin
20	2HMI	HD	Fab 28	HIV1 reverse transcriptase
21	2JEL	RB	Fab Jel42	HPr
22	2QFW	RB	Fv	Human chorionic gonadotropin
	Enzyme–Inhibitor or Substrate (23)			
23	1ACB	MD	Chymotrypsin	Eglin C
24	1AVX	RB	Porcine trypsin	Soybean trypsin inhibitor
25	1AY7	RB	Barnase	Barstar
26	1BVN	RB	alpha-amylase	Tendamistat
27	1CGI	RB	Bovine chymotrypsinogen	PSTI
28	1D6R	RB	Bovine trypsin	Bowman-Birk inhibitor
29	1DFJ	RB	Ribonuclease A	Rnase inhibitor
30	1E6E	RB	Adrenoxin reductase	Adrenoxin
31	1EAW	RB	Matriptase	BPTI
32	1EWY	RB	Ferredoxin reductase	Ferredoxin
33	1EZU	RB	D102N Trypsin	Ecotin
34	1F34	RB	Porcine pepsin	Ascaris inhibitor 3
35	1HIA	RB	Kallikrein	Hirustatin
36	1KKL	MD	HPr kinase C-ter domain	HPr
37	1MAH	RB	Acetylcholinesterase	Fasciculin
38	1PPE	RB	Bovine trypsin	CMTI-1 squash inhibitor
39	1TMQ	RB	alpha-amylase	RAGI inhibitor
40	1UDI	RB	Uracyl-DNA glycosylase	Glycosylase inhibitor
41	2MTA	RB	Methylamine dehydrogenase	Amicyanin
42	2PCC	RB	Cyt C peroxidase	Cytochrome C
43	2SIC	RB	Subtilisin	Streptomyces subtilisin inhibitor
44	2SNI	RB	Subtilisin	Chymotrypsin inhibitor 2
45	7CEI	RB	Colicin E7 nuclease	Im7 immunity protein
	Others (39)			
46	1A2K	RB	Ran GTPase	Nuclear transport factor 2
47	1AK4	RB	Cyclophilin	HIV capsid
48	1AKJ	RB	MHC Class 1 HLA-A2	T-cell CD8 coreceptor
49	1ATN	HD	Actin	Dnase I
50	1B6C	RB	FKBP binding protein	TGFbeta receptor
51	1BUH	RB	CDK2 kinase	Ckshs1
52	1DE4	HD	beta2-microglobulin	Transferrin receptor ectodom
53	1E96	RB	Rac GTApase	p67 Phox
54	1EER	HD	Erythropoietin	EPO receptor
55	1F51	RB	Sporulation response factor B	Sporulation response factor F
56	1FAK	HD	Coagulation factor VIIa	Soluble tissue factor
57	1FC2	RB	Staphylococcus Protein A	Human Fc fragment
58	1FQ1	HD	CDK inhibitor 3	CDK2 kinase
59	1FQJ	RB	Gt-alpha	RGS9 MSE
60	1GCQ	RB	GRB2 C-ter SH3 domain	GRB2 N-ter SH3 domain
61	1GHQ	RB	Epstein-Barr virus receptor CR2	Complement C3
62	1GP2	MD	Gi-alpha	Gi-beta,gamma
63	1GRN	MD	CDC42 GTPase	CDC42 GAP
64	1H1V	HD	Actin	Gelsolin
65	1HE1	RB	Rac GTPase	Pseudomonas toxin GAP dom.
66	1HE8	MD	Ras GTPase	PIP3 kinase
67	1I2M	MD	Ran GTPase	RCC1
68	1I4D	RB	Rac GTPase	Arfaptin
69	1IB1	MD	14-3-3 protein	Serotonin N-acteylase
70	1IBR	HD	Ran GTPase	Importin beta
71	1IJK	MD	Botrocetin	Von Willebrand Factor dom. A1
72	1K5D	MD	Ran GTPase	Ran
73	1KAC	RB	Adenovirus fiber knob protein	Adenovirus receptor
74	1KLU	RB	MHC class 2 HLA-DR1	Staphylococcus enterotoxin C3
75	1KTZ	RB	TGF-beta	TGF-beta receptor
76	1KXP	RB	Actin	Vitamin D binding protein
77	1M10	MD	Von Willebrand Factor dom. A1	Glycoprotein IB-alpha
78	1ML0	RB	Viral chemokine binding p. M3	Chemokine Mcp1
79	1N2C	MD	Nitrogenase Mo-Fe protein	Nitrogenase Fe protein
80	1QA9	RB	CD2	CD58
81	1RLB	RB	Transthyretin	Retinol binding protein
82	1SBB	RB	T-cell receptor beta	Staphylococcus enterotoxin B
83	1WQ1	MD	Ras GTPase	Ras GAP
84	2BTF	RB	Actin	Profilin

Table 2 Logical combinations of experimental interactions and prediction results.

		Experimental interaction	
		Detected	Not detected
Prediction result	Positive	True Positive (TP)	False Positive (FP)
	Negative	False Negative (FN)	True Negative (TN)

$$rec(= sens) = \frac{tp}{tp + fn}, \quad prec = \frac{tp}{tp + fp}, \quad F = \frac{2 \cdot prec \cdot rec}{prec + rec}. \tag{4}$$

Sensitivity (= recall) is represented as the proportion of true positives in protein pairs with experimentally validated interactions. That is, a prediction method with high sensitivity can reveal most of the pairs whose interactions have been experimentally detected. Then, the precision, which is the degree of reproducibility, shows the proportion of pairs with experimental interactions from all prediction results. Moreover, the F-measure is defined as the harmonic mean of recall and precision. Its value increases significantly as the values of both these factors increase. Considering the attributes of the used criteria, we employed the F-measure to assess prediction results. Because the F-measure can quantitatively gauge the prediction accuracy relative to the prevalence of a problem, we can evaluate recall and precision as trade-off in the form of a combined value.

3. Method

We propose a method for predicting protein interactions from the docking scores of protein pairs using their functional information.

3.1 Overview

In order to solve the problem of the previous study, we focused on the docking scores of functionally classified pairs. This is because when distributions of scores of pairs with the same functions were checked, the following findings can be revealed: (1) potential protein affinities exist that represent protein bindings, and (2) affinity thresholds were different between particular functionally classified pairs, which determine whether or not proteins interact. **Figure 3** shows a flowchart of the proposed prediction method. The method consists of four procedures, including three key procedures, as shown in Fig. 1, and an additional procedure by ROC analysis. In the ROC procedure, the optimal affinity threshold τ is determined by ROC analysis. Although the flow outline is similar to

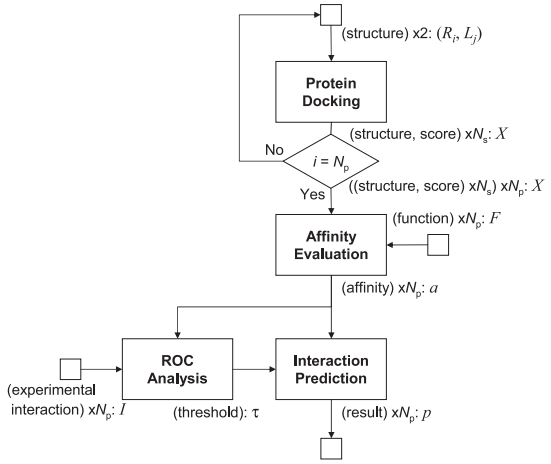

Fig. 3 Flowchart of the proposed prediction method.

that of a previous prediction method (i.e., AEP) in Fig. 2, repeated procedures are quite different. Compared with "Protein Docking" and "Affinity Evaluation" procedures in the previous method, only the "Protein Docking" procedure is repeated N_{p} (i.e., the number of all protein pairs) times. Since the proposed protein affinity is defined not only by docking scores of one pair but also by that of others, the iteration in the "Protein Docking" procedure needs to be completed before the "Affinity Evaluation" procedure. In addition, calculation methods in each procedure and score definitions are different.

3.2 Algorithm

Input and output of the proposed method are as follows:

Input:

3D structures and functions of all receptor–ligand pairs,

Output:

prediction results of computational protein–protein interactions.

The method consists of four steps, as outlined in Fig. 3.

Step 1 (Protein Docking):

Search the structures of complex candidates with docking scores using the 3D structure of one receptor–ligand pair. Step 1 is repeated for all protein pairs.

Step 2 (Affinity Evaluation):
Evaluate protein affinities by statistically inspecting docking scores of pairs using functional information.

Step 3 (ROC Analysis):
By ROC analysis, detect the optimal threshold τ for determining whether or not protein pairs interact.

Step 4 (Interaction Prediction):
Predict computational interactions from protein affinities and functions of pairs.

In order to describe the method, a protein is modeled according to the computational information that represents the 3D structure and biological function, where the protein structure includes the 3D coordinates of its constituent atoms, and biological functions are as follows:

$$F = \{\text{“}A\text{”}, \text{“}AB\text{”}, \text{“}E\text{”}, \text{“}O\text{”}\}. \quad (5)$$

In addition, the experimental interaction of each receptor–ligand pair for determining the threshold τ and measuring the performance of the prediction method was previously detected by

$$I = \begin{cases} 1 & \text{if } \textit{experimental interaction is detected} \\ 0 & \text{otherwise} \end{cases}. \quad (6)$$

3.2.1 Protein–Protein Docking (Protein Docking, Step 1)

In Step 1, we use the protein–protein docking program ZDOCK 3.0.1 to assess structural and chemical complementarity between proteins. As described in Section 2.2.1, ZDOCK 3.0.1 enables us to find binding sites and complex structures using a FFT-based search algorithm with a scoring function based on pair-wise shape complementarity, electrostatics, and explicit interface atomic contact energies. The setting values defined in Section 2.1.1 are as follows: $\Delta = 15$, $\gamma = 1.2$, and $N_{\mathrm{s}} = 2{,}000$. Then, all default values (i.e., −N (= 2,000), −S (= no at randomization), and −D (= none)) of execution options are used. This procedure outputs the top N_{s} scores for each docking score s_{ij} of a receptor–ligand pair (R_i, L_j). By repeating Step 1, docking scores of all N_{p} ($= N_{\mathrm{R}} \times N_{\mathrm{L}}$) pairs are calculated.

3.2.2 Affinity Evaluation (Step 2)

In Step 2, protein affinities are assessed statistically based on functionally classified docking scores. In order to estimate the affinity of a target pair, docking scores of all pairs having the same function are considered. We propose the following Z-scoring systems: *Receptor-Focused Z-scoring* (RFZ) and *Ligand-Focused Z-scoring* (LFZ). The aim of this method was to statistically evaluate each receptor- or ligand-focused group of docking scores and then convert the docking score into an affinity based on the Z-score. That is, RFZ collects those ligands that have scores similar to a specific receptor, while LFZ collects those receptors that have scores similar to a specific ligand. Therefore, either receptor or ligand functions are used for affinity evaluation. When RFZ evaluates protein affinity based on docking score distribution of receptors having the same functions, functional information of ligands is not used. Since the key steps for RFZ and LFZ are the same, the details for only RFZ are shown below.

Input and output of the procedure are as follows:

Input:
docking scores,

Output:
protein affinities.

The method consists of three steps.

Step 2.1 (Preparation of docking scores):
Calculate the maximum score among N_{s} scores for each receptor and ligand pair and obtain a set of docking scores $\{s_{ij} \mid R_i \in R,\ L_j \in L\}$ for assessing protein affinity.

Step 2.2 (Grouping of docking scores):
In this step, the available information for grouping the docking scores depends on receptor–ligand pairs. In this study, the following cases can be considered: (1) both structural and functional information of the confirmed pairs having same functions, and (2) only the structural information of all pairs consisted of proteins with varying functions.

All N_{p} docking scores are classified according to pair functions. For example, a set of receptors in functional category A (i.e., antibody–antigen) is given by:

$$R^{\mathrm{A}} = \{R_i \mid R_i \in R,\ F(R_i) = \text{“}A\text{”}\}, \quad (7)$$

where $F(R_i)$ denotes the biological function of receptor R_i.

Next, $N_{\mathrm{R^A}} \times N_{\mathrm{L}}$ docking scores are split into $N_{\mathrm{R^A}}$ subset of scores of pairs that have the same receptor such that:

$$s_i^{\mathrm{A}} = \{s_{ij} \mid L_j \in L\},\ R_i \in R^{\mathrm{A}}, \tag{8}$$

where $N_{\mathrm{R^A}}$ denotes the number of receptors with the function A among R. Similarly, for other functions AB, E, and O, other subsets s_i^{AB}, s_i^{E}, and s_i^{O} are obtained from s_{ij}.

Step 2.3 (Z-score calculation):

In this step, the Z-score of each score subset is calculated as follows:

$$a_{ij}^{\mathrm{A}} = \frac{s_{ij} - E(s_i^{\mathrm{A}})}{\sigma(s_i^{\mathrm{A}})}, \tag{9}$$

where $E(s_i^{\mathrm{A}})$ and $\sigma(s_i^{\mathrm{A}})$ denotes the mean and standard deviation of scores s_i^{A}, respectively. Thus, the affinities of all pairs, a_{ij}, are determined.

3.2.3 Receiver operating characteristic analysis (ROC Analysis, Step 3)

In Step 3, the affinity threshold τ is decided to maximize the F-measure by ROC analysis. When the optimal τ value is used for determinating PPIs, many prediction results of the 84 protein pairs (R_i, L_j) with experimentally validated interactions (i.e., $I_{ij} = 1$) are correctly evaluated as positives for PPIs (i.e., $p_{ij} = 1$), and the other 6,972 pairs whose interactions have not been experimentally detected (i.e., $I_{ij} = 0$) are correctly predicted as negatives (i.e., $p_{ij} = 0$). In ROC analysis, a cut-off value is generally used as the criterion for separating the two classes (i.e., positives and negatives for PPIs). When the cut-off value is changed by a certain incremental amount from minimum to maximum of all affinities (a), recall and precision values for each cut-off value are obtained. Therefore, the F-measure is obtained from recall and precision values. Although there are various ways of obtaining the optimal threshold, we employed a method based on the balance of recall and precision, as the objective of this study is to maximize the F-measure. The threshold τ was determined so that the value of $\{\mathrm{recall}^2 + (1 - \mathrm{precision})^2\}$ becomes maximum.

The values of τ vary with prediction methods and functional categories. That is, $\tau_{\mathrm{RFZ}}^{\mathrm{A}}$ for predicting the interactions of pairs in functional category A by RFZ is different from $\tau_{\mathrm{RFZ}}^{\mathrm{AB}}$ for predicting AB pairs by the same method. In addition, $\tau_{\mathrm{RFZ}}^{\mathrm{A}}$ is not equal to $\tau_{\mathrm{LFZ}}^{\mathrm{A}}$ in spite of same functional categories. When the prediction method uses a combination of RFZ and LFZ, called RFZ×LFZ and RFZ+LFZ (see Section 3.2.4 for the definitions of RFZ×LFZ and RFZ+LFZ), a threshold $\tau_{\mathrm{RFZ\times LFZ}}^{\mathrm{A}}$ for protein pairs in A from RFZ×LFZ shows a pair of thresholds for both RFZ and LFZ such as $(\tau_{\mathrm{RFZ}}^{\mathrm{A}}, \tau_{\mathrm{LFZ}}^{\mathrm{A}})$.

In the previous method, the entire protein-pair data set, including 84 target protein pairs, was used for determining the threshold τ by ROC analysis. In order to directly compare the previous results, the entire data set was also used in this study.

3.2.4 Interaction Prediction (Step 4)

In Step 4, the interactions of all receptor–ligand pairs are finally predicted. Protein affinities are assessed by comparisons with the optimal threshold τ. If the affinity value is more than or equal to τ, the prediction result p_{ij} is set to 1, indicating "positive;" Otherwise, the value is set to 0, indicating "negative," such as

$$p_{ij} = \begin{cases} 1 & \text{if } a_{ij} \geq \tau \\ 0 & \text{otherwise} \end{cases}. \tag{10}$$

Here, we propose additional scoring systems, RFZ×LFZ and RFZ+LFZ, which predict the interactions by combining the results of RFZ and LFZ. That is, the prediction result of each protein pair, $p_{ij}^{\mathrm{RFZ\times LFZ}}$ is defined by the logical AND operation of p_{ij}^{RFZ} and p_{ij}^{LFZ}, while $p_{ij}^{\mathrm{RFZ+LFZ}}$ is defined by the logical OR operation.

4. Results and Discussion

The important findings of this study are as follows: (1) achieving improved prediction accuracy by assessing protein affinities by considering biological functions, and (2) achieving additional refinement of prediction accuracy by combining proposed scoring systems, RFZ and LFZ. We discuss these points in this order below.

4.1 Prediction Performance of RFZ and LFZ

For comparison with previous methods, we employed the protein-pair data set in Section 2.2.3. At the first evaluation for improving prediction accuracy

Table 3 Comparison of proposed methods, RFZ and LFZ, and previous methods. Prediction performances (%) are determined by the F-measure, Recall (= Sensitivity), and Precision. τ indicates the threshold for determining protein–protein interactions.

Method	Performance measures						
	F-measure	Recall	Precision	*tp*	*fp*	*fn*	*tn*
RFZ ($\tau = 2.984$)	28.3	20.2	47.2	17	19	67	6,953
LFZ ($\tau = 3.103$)	21.9	16.7	31.8	14	30	70	6,942
AEP ($\tau = 4.520$)	6.3	27.4	3.5	23	629	61	6,343
ZDOCK ($\tau = 1{,}364.018$)	4.0	73.8	2.1	62	2,941	22	4,031

using the proposed method, the entire data set ($84 \times 84 = 7{,}056$ protein pairs) was used. **Table 3** shows comparison of prediction accuracy for the proposed scoring systems (RFZ and LFZ) and previous methods (AEP and ZDOCK). AEP indicates the previous method detailed in Section 2.2.2, and the best performance of AEP is given by optimizing key parameters, as shown in Table 3. ZDOCK is a simple PPI prediction method that directly uses ZDOCK 3.0.1 for protein-protein affinity calculations, as described in Section 2.2.1. The proposed RFZ and LFZ greatly improved prediction accuracy, with the obtained F-measures higher than those obtained by previous methods. This improvement arose as the proposed method could reduce many false positives; the number of false positives obtained by RFZ was only 19 among all 6,972 negative examples, compared with 629 by AEP. As a result, using only structural information of proteins, RFZ and LFZ correctly predicted the interactions of 17 and 14 pairs among 84, respectively.

For predictions using both structural and functional information of proteins, **Fig. 4** summarizes the comparison results of prediction methods in each functional category. Horizontal and vertical axes denote the F-measure value or recall/precision and the prediction method, respectively. The detailed value of each criterion and the numbers of TP, FP, FN, and TN are shown in **Table 4**. The previous methods, ZDOCK and AEP, indicate prediction results only using protein pairs having same functions. Each threshold τ was determined so as to maximize the F-measure by ROC analysis. Here, it was not always necessary that the proposed method, which used affinities based on functionally classified docking scores, could improve the prediction accuracy. One reason was that several pairs having different functions and high affinities were present in all 7,056 pairs. These high scores were greater than those of pairs with the same functions.

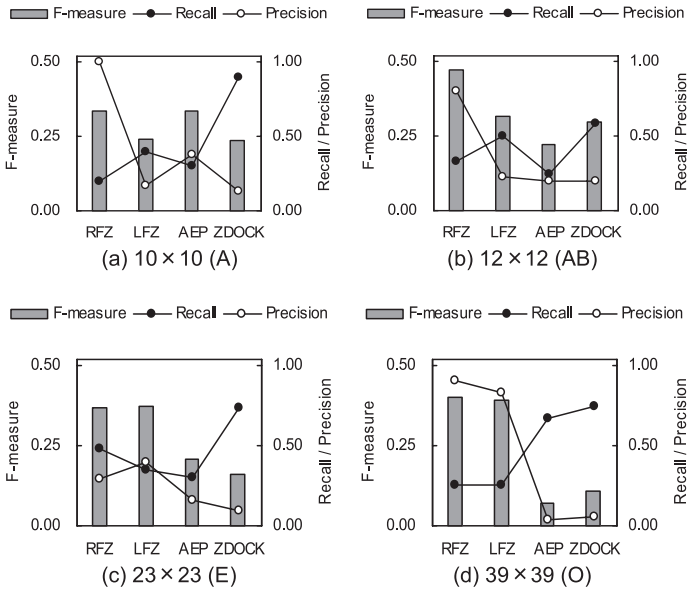

Fig. 4 Comparison of prediction accuracy in respective functional categories. The bar chart and line graph denote the values of the F-measure and recall/precision, respectively.

An example of a high affinity pair is R_3 and L_{26}, where R_3 denotes a receptor protein of 1BVK (index: 3, functional category: A, conformational category: RB) [38], which is the Fv fragment of a humanized anti-hen egg white lysozyme Ab (HuLys), and L_{26} denotes a ligand protein of 1BVN (26, E, RB) [39], which is the alpha-amylase inhibitor Hoe-467A (Tendamistat). In a preliminary experiment, the affinity of $a(R_3, L_{26})$ was 1.43 times greater than that of the 10 affinities of $\{a(R_3, L_1), a(R_3, L_2), \cdots, a(R_3, L_{10})\}$ (minimum: -1.28, maximum: 0.91) for pairs having the same receptor R_3 in functional category A. For all receptors R_1–R_{10} in category A, there were 122 high affinity pairs among 740 pairs that had different functions, that is, 16.5% of all pairs. As shown in Table 4, both proposed scoring systems improved the performance in F-measures in almost all cases compared to that by previous methods. Only in the case of LFZ, the 10×10 subset of functional category A was less than that of AEP (Table 4 (a)). As shown in Table 4, LFZ gave a prediction accuracy of 24.2% with

Table 4 Comparison of proposed methods, RFZ and LFZ, and previous methods in respective functional categories. Prediction performances (%) are determined by the F-measure, Recall (= Sensitivity), and Precision. τ indicates the threshold for determining protein–protein interactions.

Method	Performance measures						
	F-measure	Recall	Precision	*tp*	*fp*	*fn*	*tn*
(a) Antibody–Antigen (10 × 10)							
RFZ ($\tau = 2.001$)	33.3	20.0	100.0	2	0	8	90
LFZ ($\tau = 0.445$)	24.2	40.0	17.4	4	19	6	71
AEP ($\tau = 4.630$)	33.3	30.0	37.5	3	5	7	85
ZDOCK ($\tau = 1{,}489.529$)	23.7	90.0	13.6	9	57	1	33
(b) Bound Antibody–Antigen (12 × 12)							
RFZ ($\tau = 1.751$)	47.1	33.3	80.0	4	1	8	131
LFZ ($\tau = 0.949$)	31.6	50.0	23.1	6	20	6	112
AEP ($\tau = 4.590$)	22.2	25.0	20.0	3	12	9	120
ZDOCK ($\tau = 1{,}831.057$)	29.8	58.3	20.0	7	28	5	104
(c) Enzyme–Inhibitor or Substrate (23 × 23)							
RFZ ($\tau = 1.536$)	36.7	47.8	29.7	11	26	12	480
LFZ ($\tau = 1.751$)	37.2	34.8	40.0	8	12	15	494
AEP ($\tau = 4.530$)	20.9	30.4	15.9	7	37	16	469
ZDOCK ($\tau = 1{,}320.669$)	16.2	73.9	9.1	17	170	6	336
(d) Others (39 × 39)							
RFZ ($\tau = 3.131$)	40.0	25.6	90.9	10	1	29	1,481
LFZ ($\tau = 3.483$)	39.2	25.6	83.3	10	2	29	1,480
AEP ($\tau = 3.970$)	7.1	66.7	3.8	26	665	13	817
ZDOCK ($\tau = 1{,}359.312$)	10.9	74.4	5.9	29	463	10	1,019

an F-measure < 33.3% of AEP because of the increase from 5 false positives by AEP to 19 by LFZ, whereas the number of true positives was slightly increased by one. In contrast, RFZ provided maximum performance in each category. The considerable improvement by RFZ was owing to the decrease of many false positives. This is supported by the fact that precision was greater than recall in three of four categories. In addition, although there was naturally a trade-off between recall and precision, LFZ in the AB category achieved the best recall (= sensitivity) that was 50.0% more than AEP (Table 4 (b)). These results indicate that the proposed scoring system could provide a significant improvement in prediction accuracy using both structural and functional information of proteins.

As described in Section 3.2.3, in order to make direct comparisons with previous results, all protein-pair data sets, including target protein pairs, were used for determining the threshold τ by ROC analysis. Here, we also investigated the prediction results using the threshold τ by leave-one-out cross-validation (LOOCV). **Table 5** shows a comparison of prediction accuracy of the proposed scoring systems (i.e., RFZ and LFZ) and other methods based on LOOCV (i.e., RFZ_LOOCV and LFZ_LOOCV). RFZ and LFZ indicate prediction accuracy using the optimum threshold τ determined by all data sets, including target protein pairs. Because the processes for determining the threshold τ were based on the training data and evaluating the validation data using this threshold τ were repeated for a required number of times (e.g., $10 \times 10 = 100$ times in functional category A), τ_{ave} denotes the mean of these thresholds. In functional category A, although prediction results of RFZ_LOOCV were the same as that of RFZ, LFZ_LOOCV gave a prediction accuracy of 6.7% F-measure that was less than 24.2% by LFZ because of the decrease from four true positives by LFZ to one by LFZ_LOOCV (Table 5 (a)). In functional category E, 33.9% F-measure of RFZ_LOOCV is less than 36.7% of RFZ because of the decrease from 11 true positives by RFZ to 10 by RFZ_LOOCV. The number of true positives by LFZ_LOOCV ($tp = 6$) is less than that by LFZ ($tp = 8$) (Table 5 (b)). The number of false positives in each functional category was not changed by the LOOCV method. These results indicate that the performance of the proposed scoring

Table 5 Comparison of prediction accuracy by LOOCV and other methods using all data sets, including target pairs in functional categories A and E. RFZ and LFZ denote the proposed method with the optimum threshold τ by ROC analysis, simply using all data sets. RFZ_LOOCV and LFZ_LOOCV denote RFZ and LFZ with the average threshold τ_{ave} from the LOOCV method.

Method	Performance measures						
	F-measure	Recall	Precision	*tp*	*fp*	*fn*	*tn*
(a) Antibody–Antigen (10 × 10)							
RFZ ($\tau = 2.001$)	33.3	20.0	100.0	2	0	8	90
LFZ ($\tau = 0.445$)	24.2	40.0	17.4	4	19	6	71
RFZ_LOOCV ($\tau_{ave} = 1.969$)	33.3	20.0	100.0	2	0	8	90
LFZ_LOOCV ($\tau_{ave} = 0.550$)	6.7	10.0	5.0	1	19	9	71
(b) Enzyme–Inhibitor or Substrate (23 × 23)							
RFZ ($\tau = 1.536$)	36.7	47.8	29.7	11	26	12	480
LFZ ($\tau = 1.751$)	37.2	34.8	40.0	8	12	15	494
RFZ_LOOCV ($\tau_{ave} = 1.536$)	33.9	43.5	27.8	10	26	13	480
LFZ_LOOCV ($\tau_{ave} = 1.752$)	29.3	26.1	33.3	6	12	17	494

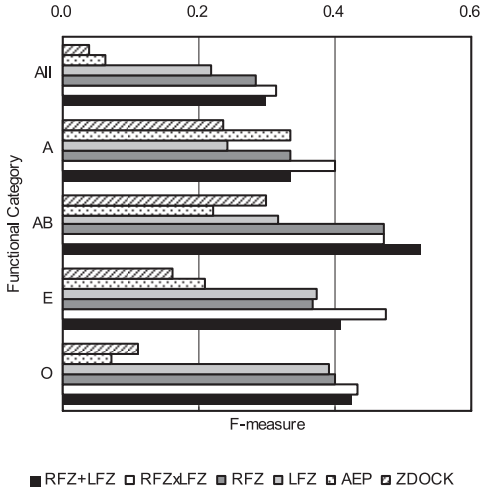

Fig. 5 Comparison with prediction accuracy of F-measure. A, AB, E, and O indicate respective functional categories. "All" indicates prediction accuracy using all data sets.

system RFZ and LFZ decreases based on thresholds by the LOOCV method.

4.2 Effect of combining RFZ and LFZ

For further analysis, we investigated the influence of further improvement by combining RFZ and LFZ. **Figure 5** shows a comparison with prediction accuracy based on the F-measure. For predictions using all data sets and considering only the structural information of proteins, the use of both RFZ×LFZ and RFZ+LFZ exceeded RFZ and LFZ over previous methods. In addition, regarding the results of function-related predictions, F-measures of proposed methods were more than or equal to those of RFZ and LFZ in each functional category. Comparing RFZ×LFZ and RFZ+LFZ performances, RFZ×LFZ was better than RFZ+LFZ in three (i.e., A, E, and O) of four functional categories. These results indicate that combining RFZ and LFZ is effective for improving prediction accuracy.

4.3 High-Affinity Complexes

Figure 6 shows the predicted and non-predicted protein pairs by the proposed methods. The PDB index from 1 to 84 on the horizontal axis indicates protein pairs that have been identified experimentally to form complexes. The relationships between PDB IDs of these complexes and their corresponding indices are

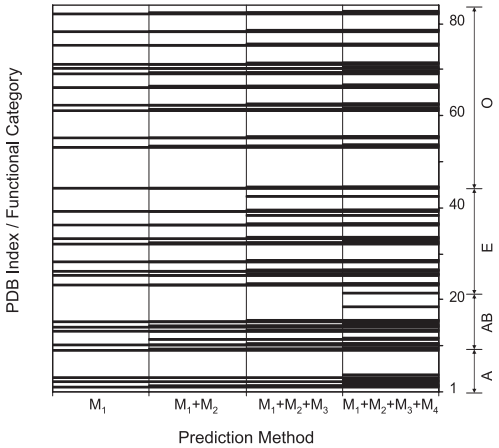

Fig. 6 Predicted and non-predicted protein pairs by all proposed methods. M_1, M_2, M_3, and M_4 are abbreviations for RFZ×LFZ, RFZ+LFZ, RFZ, and LFZ, respectively.

shown in Table 1. The vertical axis represents the proposed prediction methods of M_1, M_2, M_3, and M_4 indicating RFZ×LFZ, RFZ+LFZ, RFZ, and LFZ, respectively. The high-affinity pairs are shown as straight lines parallel to the vertical axis. Thus, the low-affinity pairs that were not predicted by any method are shown as white spaces in $M_1 + M_2 + M_3 + M_4$. Each method predicted similar protein pair interactions, such as 1BGX (index: 2, functional category: A, conformational category: MD) [40], 1BJ1 (11, AB, RB) [41], 1AVX (24, E, RB) [42], and 1EER (54, O, HD) [43]. **Table 6** shows the results of a number of high- and low-affinity complexes. The complexes of MD or HD in conformational categories are shown in parentheses. The number of pairs successfully predicted by at least one method is the highest at 58.3% of total pairs in functional category AB. In addition, predictions by all methods were very similar in E and O categories; that is, between one and three methods can predict several protein pairs such as 1F34 (34, E, RB) [44] and 1IJK (71, O, MD) [45]. On the other hand, all 84 complexes include 21 complexes in conformational categories MD or HD. About half of these 21 complexes were predicted by all methods, and the remaining 10 pairs were not predicted by any method. Especially in category O, nine pairs predicted by all methods were in MD or HD conformational categories. These

Table 6 Results of the number of high-affinity complexes predicted by proposed methods. A, AB, E, and O indicate respective functional categories. The protein pairs of medium difficulty (MD) or high difficulty (HD) in conformational categories are shown in parentheses.

Complex		Num. of high-affinity complexes		
Cat.	Num.	0 methods	1–3 methods	4 methods
A	10 (1)	6 (0)	2 (0)	2 (1)
AB	12 (1)	5 (1)	4 (0)	3 (0)
E	23 (2)	12 (2)	3 (0)	8 (0)
O	39 (17)	28 (7)	2 (1)	9 (9)
Total	84 (21)	51 (10)	11 (1)	22 (10)

results suggest that although all four methods predict similar protein pairs, the effect of combining RFZ and LFZ can be obtained by several different pairs by each method (i.e., RFZ and LFZ). In addition, our proposed methods can predict protein pair interactions with conformational changes at high accuracy.

5. Conclusions

We have proposed a method for predicting protein interactions from the docking scores of protein pairs that have same functions. We developed PPI prediction scoring systems, RFZ and LFZ, to statistically evaluate separate receptor- or ligand-focused groups of docking scores and convert the docking score into protein affinity based on the Z-score. The proposed method assessed the improvement in prediction accuracy using a protein-pair data set. The following conclusions were derived from the results and discussion.

(1) By an analysis of biological functions of protein pairs, the prediction accuracy is significantly improved without changing the prediction algorithm itself.

(2) Further refinement of prediction accuracy was achieved by combining RFZ and LFZ.

In the near future, we plan to study the identification of protein functions using the existing docking programs with various scoring functions. We also aim at extending the new findings for application to related research such as drug design.

References

1) Jones, S. and Thornton, J.M.: Principles of protein-protein interactions, *Proc. Natl. Acad. Sci. U.S.A.*, Vol.93, No.1, pp.13–20 (1996).
2) Lo Conte, L., Chothia, C. and Janin, J.: The atomic structure of protein-protein recognition sites, *J. Mol. Biol.*, Vol.285, No.5, pp.2177–2198 (1999).
3) Norel, R., Petrey, D., Wolfson, H.J. and Nussinov, R.: Examination of shape complementarity in docking of unbound proteins, *Proteins*, Vol.36, No.3, pp.307–317 (1999).
4) Fields, S. and Song, O.: A novel genetic system to detect protein-protein interactions, *Nature*, Vol.340, No.6230, pp.245–246 (1989).
5) Giot, L., Bader, J.S., Brouwer, C., et al.: A protein interaction map of Drosophila melanogaster, *Science*, Vol.302, No.5651, pp.1727–1736 (2003).
6) Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S. and Sakaki, Y.: Toward a protein-protein interaction map of the budding yeast: A comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins, *Proc. Natl. Acad. Sci. U.S.A.*, Vol.97, No.3, pp.1143–1147 (2000).
7) Rain, J.C., Selig, L., De Reuse, H., et al.: The protein-protein interaction map of *Helicobacter pylori*, *Nature*, Vol.409, No.6817, pp.211–215 (2001).
8) Tarassov, K., Messier, V., Landry, C.R., et al.: An *in vivo* map of the yeast protein interactome, *Science*, Vol.320, No.5882, pp.1465–1470 (2008).
9) Uetz, P., Giot, L., Cagney, G., et al.: A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*, *Nature*, Vol.403, No.6770, pp.623–627 (2000).
10) Walhout, A.J., Sordella, R., Lu, X., Hartley, J.L., Temple, G.F., Brasch, M.A., Thierry-Mieg, N. and Vidal, M.: Protein interaction mapping in C. elegans using proteins involved in vulval development, *Science*, Vol.287, No.5450, pp.116–122 (2000).
11) Wojcik, J. and Schachter, V.: Protein-protein interaction map inference using interacting domain profile pairs, *Bioinformatics*, Vol.17, No.Suppl.1, pp.S296–S305 (2001).
12) Gavin, A.C., Bosche, M., Krause, R., et al.: Functional organization of the yeast proteome by systematic analysis of protein complexes, *Nature*, Vol.415, No.6868, pp.141–147 (2002).
13) Ho, Y., Gruhler, A., Heilbut, A., et al.: Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry, *Nature*, Vol.415, No.6868, pp.180–183 (2002).
14) Zhu, H., Bilgin, M., Bangham, R., et al.: Global analysis of protein activities using proteome chips, *Science*, Vol.293, No.5537, pp.2101–2105 (2001).
15) Galperin, M.Y. and Koonin, E.V.: Who's your neighbor? New computational ap-

proaches for functional genomics, *Nat. Biotechnol.*, Vol.18, No.6, pp.609–613 (2000).
16) Huynen, M., Snel, B., Lathe, W. 3rd and Bork, P.: Predicting protein function by genomic context: Quantitative evaluation and qualitative inferences, *Genome Res.*, Vol.10, No.8, pp.1204–1210 (2000).
17) Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, 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. U.S.A.*, Vol.89, No.6, pp.2195–2199 (1992).
18) Gabb, H.A., Jackson, R.M. and Sternberg, M.J.: Modelling protein docking using shape complementarity, electrostatics and biochemical information, *J. Mol. Biol.*, Vol.272, No.1, pp.106–120 (1997).
19) Sternberg, M.J., Aloy, P., Gabb, H.A., Jackson, R.M., Moont, G., Querol, E. and Aviles, F.X.: A computational system for modelling flexible protein-protein and protein-DNA docking, *Proc. Int. Conf. Intell. Syst. Mol. Biol.*, Vol.6, pp.183–192 (1998).
20) Vakser, I.A.: Protein docking for low-resolution structures, *Protein Eng.*, Vol.8, No.4, pp.371–377 (1995).
21) Mandell, J.G., Roberts, V.A., Pique, M.E., Kotlovyi, V., Mitchell, J.C., Nelson, E., Tsigelny, I. and Eyck, L.F.T.: Protein docking using continuum electrostatics and geometric fit, *Protein Eng.*, Vol.14, No.2, pp.105–113 (2001).
22) Chen, R., Li, L. and Weng, Z.: ZDOCK: An initial-stage protein-docking algorithm, *Proteins*, Vol.52, No.1, pp.80–87 (2003).
23) Ritchie, D.W. and Kemp, G.J.: Protein docking using spherical polar Fourier correlations, *Proteins*, Vol.39, No.2, pp.178–194 (2000).
24) Gray, J.J., Moughon, S., Wang, C., Schueler-Furman, O., Kuhlman, B., Rohl, C.A. and Baker, D.: Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations, *J. Mol. Biol.*, Vol.331, No.1, pp.281–299 (2003).
25) Duhovny, D., Nussinov, R. and Wolfson, H.J.: Efficient unbound docking of rigid molecules, *WABI '02: Proc. 2nd International Workshop on Algorithms in Bioinformatics*, Springer-Verlag, pp.185–200 (2002).
26) Wolfson, H.J. and Rigoutsos, I.: Geometric hashing: An overview, *IEEE Comput. Sci. Eng.*, Vol.4, No.4, pp.10–21 (1997).
27) Smith, G.R. and Sternberg, M.J.: Prediction of protein-protein interactions by docking methods, *Curr. Opin. Struct. Biol.*, Vol.12, No.1, pp.28–35 (2002).
28) Tsukamoto, K., Yoshikawa, T., Hourai, Y., Fukui, K. and Akiyama, Y.: Development of an affinity evaluation and prediction system by using the shape complementarity characteristic between proteins, *J. Bioinf. Comput. Biol.*, Vol.6, No.6, pp.1133–1156 (2008).
29) Yoshikawa, T., Tsukamoto, K., Hourai, Y. and Fukui, K.: Parameter tuning and evaluation of an affinity prediction using protein-protein docking, *MMACTEE '08: Proc. 10th WSEAS International Conference on Mathematical Methods and Computational Techniques in Electrical Engineering*, pp.312–317, World Scientific and Engineering Academy and Society (WSEAS) (2008).
30) Tsukamoto, K., Yoshikawa, T., Yokota, K., Hourai, Y. and Fukui, K.: The development of an affinity evaluation and prediction system by using protein-protein docking simulations and parameter tuning, *Adv. Appl. Bioinf. Chem.*, Vol.2, pp.1–15 (2009).
31) Yoshikawa, T., Tsukamoto, K., Hourai, Y. and Fukui, K.: Improving the accuracy of an affinity prediction method by using statistics on shape complementarity between proteins, *J. Chem. Inf. Model.*, Vol.49, No.3, pp.693–703 (2009).
32) Sacquin-Mora, S., Carbone, A. and Lavery, R.: Identification of protein interaction partners and protein-protein interaction sites, *J. Mol. Biol.*, Vol.382, No.5, pp.1276–1289 (2008).
33) Lattman, E.: Optimal sampling of the rotation function, *Acta Crystallogr., Sect. B: Struct. Sci.*, Vol.28, No.4, pp.1065–1068 (1972).
34) Mintseris, J., Pierce, B., Wiehe, K., Anderson, R., Chen, R. and Weng, Z.: Integrating statistical pair potentials into protein complex prediction, *Proteins*, Vol.69, No.3, pp.511–520 (2007).
35) Nukada, A., Hourai, Y., Nishida, A. and Akiyama, Y.: High performance 3D convolution for protein docking on IBM Blue Gene, *ISPA '07: Proc. 5th International Symposium on Parallel and Distributed Processing and Applications*, pp.958–969, Springer (2007).
36) Chen, R., Mintseris, J., Janin, J. and Weng, Z.: A protein-protein docking benchmark, *Proteins*, Vol.52, No.1, pp.88–91 (2003).
37) Mintseris, J., Wiehe, K., Pierce, B., Anderson, R., Chen, R., Janin, J. and Weng, Z.: Protein-Protein Docking Benchmark 2.0: An update, *Proteins*, Vol.60, No.2, pp.214–216 (2005).
38) Holmes, M., Buss, T. and Foote, J.: Conformational correction mechanisms aiding antigen recognition by a humanized antibody, *J. Exp. Med.*, Vol.187, No.4, pp.479–485 (1998).
39) Wiegand, G., Epp, O. and Huber, R.: The crystal structure of porcine pancreatic α-amylase in complex with the microbial inhibitor Tendamistat, *J. Mol. Biol.*, Vol.247, No.1, pp.99–110 (1995).
40) Murali, R., Sharkey, D.J., Daiss, J.L. and Murthy, H.M.: Crystal structure of Taq DNA polymerase in complex with an inhibitory Fab: The Fab is directed against an intermediate in the helix-coil dynamics of the enzyme, *Proc. Natl. Acad. Sci. U.S.A.*, Vol.95, No.21, pp.12562–12567 (1998).
41) Muller, Y.A., Chen, Y., Christinger, H.W., Li, B., Cunningham, B.C., Lowman, H.B. and de Vos, A.M.: VEGF and the Fab fragment of a humanized neutralizing antibody: Crystal structure of the complex at 2.4 Å resolution and mutational analysis of the interface, *Structure*, Vol.6, No.9, pp.1153–1167 (1998).

42) Song, H.K. and Suh, S.W.: Kunitz-type soybean trypsin inhibitor revisited: Refined structure of its complex with porcine trypsin reveals an insight into the interaction between a homologous inhibitor from Erythrina caffra and tissue-type plasminogen activator, *J. Mol. Biol.*, Vol.275, No.2, pp.347–363 (1998).
43) Syed, R.S., Reid, S.W., Li, C., et al.: Efficiency of signalling through cytokine receptors depends critically on receptor orientation, *Nature*, Vol.395, No.6701, pp.511–516 (1998).
44) Ng, K.K., Petersen, J.F., Cherney, M.M., et al.: Structural basis for the inhibition of porcine pepsin by Ascaris pepsin inhibitor-3, *Nat. Struct. Biol.*, Vol.7, No.8, pp.653–657 (2000).
45) Fukuda, K., Doggett, T.A., Bankston, L.A., Cruz, M.A., Diacovo, T.G. and Liddington, R.C.: Structural basis of von Willebrand factor activation by the snake toxin botrocetin, *Structure*, Vol.10, No.7, pp.943–950 (2002).

(Received October 20, 2009)
(Accepted December 7, 2009)
(Released March 8, 2010)

(Communicated by *Tatsuya Akutsu*)

Tatsuya Yoshikawa received his B.E. and M.E. degrees from Kobe University in 1998 and 2000, respectively. He is currently a Ph.D. student at the Graduate School of Information Science and Technology, Osaka University. His research interests include computational analysis of protein–protein interactions, massively parallel computing, and machine learning. He is a member of IPSJ and JSBi.

Shigeto Seno is an Assistant Professor of the Graduate School of Information Science and Technology, Osaka University. He received his B.E., M.E. and Ph.D. degrees from Osaka University in 2001, 2003 and 2006 respectively. He is a member of IPSJ, ISCB, JSBi and MBSJ.

Yoichi Takenaka received his M.E. and Ph.D. in 1997, and 2000 from Osaka university, respectively. He worked for Osaka Universitiy from 2000 to 2002 as assistant professor, and now he is an associate professor at Graduate School of Information Science and Technology, Osaka University. His research intersts include Bioinformatics, DNA computing, and Neural Networks.

Hideo Matsuda is a Professor of the Department of Bioinformatic Engineering, the Graduate School of Information Science and Technology, Osaka University. He received his B.S., M.Eng., and Ph.D. degrees from Kobe University in 1982, 1984 and 1987, respectively. His research interests include computational analysis of genomic sequences and integrated biological databases. He is a member of JSBi, ISCB, IEEE CS and ACM.