

# Prediction of Protein–compound Binding Energies from Known Activity Data: Docking-score-based Method and its Applications

Yoshifumi Fukunishi,<sup>\*[a]</sup> Yasunobu Yamashita,<sup>[b]</sup> Tadaaki Mashimo,<sup>[b, c]</sup> and Haruki Nakamura<sup>[d]</sup>

**Abstract:** We used protein–compound docking simulations to develop a structure-based quantitative structure–activity relationship (QSAR) model. The prediction model used docking scores as descriptors. The binding free energy was approximated by a weighted average of docking scores for multiple proteins. This approximation was based on a pharmacophore model of receptor pockets and compounds. The weights of the docking scores were restricted to small values to avoid unrealistic weights by a regularization term. Additional outlier elimination improved the results. We applied this method to two groups of targets. The first target was the kinase family. The cross-validation results of 107 kinase proteins showed that the RMSE of

predicted binding free energies was 1.1 kcal/mol. The second target was the matrix metalloproteinase (MMP) family, which has been difficult for docking programs. MMPs require metal-binding groups in their inhibitor structures in many cases. A quantum effect contributes to the metal–ligand interaction. Despite this difficulty, the present method worked well for the MMPs. This method showed that the RMSE of predicted binding free energies was 1.1 kcal/mol. In comparison, with the original docking method the RMSE was 1.7 kcal/mol. The results suggest that the present QSAR model should be applied to general target proteins.

**Keywords:** Binding free energy · ChEMBL · Docking score · Protein–compound docking

## 1 Introduction

The quantitative structure–activity relationship (QSAR) approach is a useful tool for optimizing leads and predicting target/off-target activities and toxicity. QSAR-based affinity predictions are useful for the general drug development process, including the repositioning (repurposing) of already approved drugs, poly-pharmacology, and the prediction of drug–drug interactions.<sup>[1–15]</sup> The recent accumulation of protein–compound affinity data in public repositories, such as the PubChem and ChEMBL projects, has enabled us to carry out proteome-wide target/off-target predictions.<sup>[16,17]</sup> These predictions are based on QSAR models for multiple proteins, just as in conventional computer-aided drug design and virtual screening.

Wide application of QSAR-based models in computer-aided drug development, such as protein–compound binding free energy (affinity) prediction, target/off-target predictions, and counter screening based on QSAR models, has succeeded in many studies, including ours.<sup>[3,4,7,8]</sup> Most QSAR models rely on descriptors with sets of two-dimensional (2D) substructures; the most popular such descriptors are MDL's MACCS key and 0-3D molecular descriptors (e.g., 5,270 descriptors recorded in Dragon (Kode srl, Pisa, Italy)). In our previous studies, we developed QSAR methods for the affinity prediction of a compound by using docking studies against multiple proteins.<sup>[17–19]</sup> We used a protein–compound affinity matrix as the set of descriptors

and applied principal component regression (PCR).<sup>[18]</sup> The  $Q^2$  value of calculated binding free energies was 0.44 and the RMSE was 1.54 kcal/mol for about 97 kinases and 18,491 compounds selected from the ChEMBL database. However, the coefficients of the regression equations for some targets were unrealistically ( $10^3$ – $10^5$ ) higher than those for other

[a] Y. Fukunishi

Molecular Profiling Research Center for Drug Discovery (molprof), National Institute of Advanced Industrial Science and Technology (AIST), 2-3-26, Aomi, Koto-ku, Tokyo 135-0064, Japan  
phone/fax: +81-3-3599-8290/ +81-3-3599-8099  
E-mail: y-fukunishi@aist.go.jp

[b] Y. Yamashita, T. Mashimo

Technology Research Association for Next-Generation Natural Products Chemistry, 2-3-26, Aomi, Koto-ku, Tokyo 135-0064, Japan

[c] T. Mashimo

IMSBIO Co., Ltd., Owl Tower, 4-21-1, Higashi-Ikebukuro, Toshima-ku, Tokyo 170-0013, Japan

[d] H. Nakamura

Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan



Supporting information for this article is available on the WWW under <https://doi.org/10.1002/minf.201700120>



© 2018 The Authors. Published by Wiley-VCH Verlag GmbH & Co. KGaA.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

targets. Either these coefficients should be restricted to a range of realistic values, or the applicability domain should be very restricted around the known experimental data.

In the present study, we applied a combination of the ridge (Tikhonov regularization) regression, robust estimation, and principal component analysis to the protein–compound affinity matrix.<sup>[19–21]</sup> The robust estimation was expected to reduce the problem of error in the experimental data. The present method restricted the coefficients of the generated prediction equation around realistic values. The method was applied to the kinases and the matrix metalloproteinases (MMPs) of the ChEMBL database.<sup>[22–24]</sup>

MMPs, which have zinc ions in the reaction pockets, require metal-binding groups in their inhibitor structures in many cases.<sup>[25]</sup> The metal–ligand interaction shows quantum effects such as electron donation and back donation, that form a weak covalent bond, in addition to the electrostatic and van der Waals interactions.<sup>[26–36]</sup> The quantum effect makes it difficult to estimate binding energy. A method to evaluate metal interaction has long been sought. In the framework of the classical force field, several methods have represented metal interactions. One is to add a metal contact term such as the van der Waals–type potential and the potential considering the coordination number of the central metal ion.<sup>[27–31]</sup> The other method is to modify the parameters of the atomic charge and the van der Waals potential of the original force field.<sup>[32–34]</sup> The metal parameters depend on the environments of the metal atom, so the user should tune the parameters for each protein.<sup>[35,36]</sup> The metal contact terms enable the user to reproduce protein–ligand complex structures and the absolute value of the binding energy. However, protein-dependent parameter tuning has been a time-consuming process, especially when the user analyzes multiple target proteins.

In addition, we evaluated the effect of the elimination of outlier data points from these multiple data points corresponding to each single protein–compound pair, and we improved the present regression model. The method, which we call the “docking-score-based QSAR model”, predicted the protein–compound binding affinities of 107 kinases that have no metals in their pocket, and those of 5 MMPs (MMP2, MMP3, MMP7, MMP9, and MMP13) that have a zinc ion in each pocket. The docking-score QSAR method worked for these various targets. Namely, the RMSE values were ~1 kcal/mol, respectively.

## 2 Materials and Methods

### 2.1 Background of Prediction Models

In the present study, we develop a binding-energy (affinity) prediction method based on the protein–compound dock-

ing scores obtained by a docking program; the present method is a modified version of our QSAR method.<sup>[18]</sup>

In our present and previous QSAR models, the affinity of compounds can be estimated by using a pharmacophore model of the target protein. The IUPAC guidelines define a pharmacophore as “an ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response”.<sup>[37]</sup> Hydrogen donors, hydrogen acceptors, hydrophobic groups of ligands and receptors are called “pharmacophore features” and the functional groups can give these features.<sup>[38]</sup> These features are usually depicted as spheres connected by lines those lengths represent the spatial distances among these features of the ligands and receptors. In the present study, we temporarily define a pharmacophore as a set of spatially distributed pharmacophore features. Each pharmacophore feature represents the probability of existence of a hydrogen-bond donor, a hydrogen-bond acceptor, and both electrostatic and hydrophobic interaction sites. Both receptors and ligands have pharmacophores. We approximate that the receptor–ligand binding energy is given by the sum of interactions between the pairs of the ligand pharmacophore ( $\phi_l$ ) and the receptor pharmacophore ( $\phi_r$ ). Here, we introduce a function for the interaction  $k$  and let  $k(\phi_r, \phi_l)$  be an interaction value between the two pharmacophores.

Because we discuss only the interaction, we do not need to explicitly represent  $\phi$ s, but rather we need only the value of the function  $k(\phi_r, \phi_l)$  for the interaction. Our final  $\Delta G$ -estimation equation does not include  $\phi$ s explicitly but consists of docking scores. In this framework, the interaction between the receptor and ligand pharmacophores gives the binding energy of the pharmacophore  $l$  in a ligand to the pharmacophore  $r$  in a protein,  $\Delta G_l^r$ , as

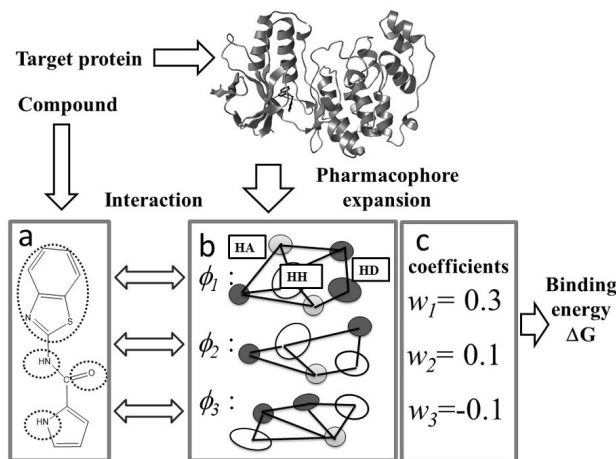
$$\Delta G_l^r = g_l^r k(\phi_r, \phi_l) \quad (1)$$

where  $g_l^r$  is a parameter.

We try to generalize this discussion by introducing a linear combination of a basic pharmacophore. Suppose  $\Phi$  ( $= \{\phi_1, \phi_2, \phi_3, \dots\}$ ) is a set of all pharmacophores. The  $\phi$  functions form the basis set of the pharmacophores of any kind of ligand or receptor. Each pharmacophore ( $\phi_i$ ) does not have to be found in an actual protein structures. In this discussion, pharmacophores work as descriptors of both protein and compound. The pharmacophores are used only to derive a regression model, and we do not calculate them explicitly. And the total number of pharmacophores could be infinite in the present discussion. We suppose that the binding energy of the  $i$ -th compound and the  $r$ -th pharmacophore in a protein is

$$\Delta G_i^r = \sum_{l \in I} g_l^r k(\phi_r, \phi_l) = \sum_{l \in I} \Delta G_l^r, \quad (2)$$

Then the protein–compound binding energy ( $\Delta G_i^a$ ) between the  $a$ -th protein and the  $i$ -th compound is given by the following linear combination of binding energies to pharmacophores  $\{\phi_m; m=1, 2, 3, \dots\}$ , where  $w_r^a$  are the scalar coefficients (see Figure 1).

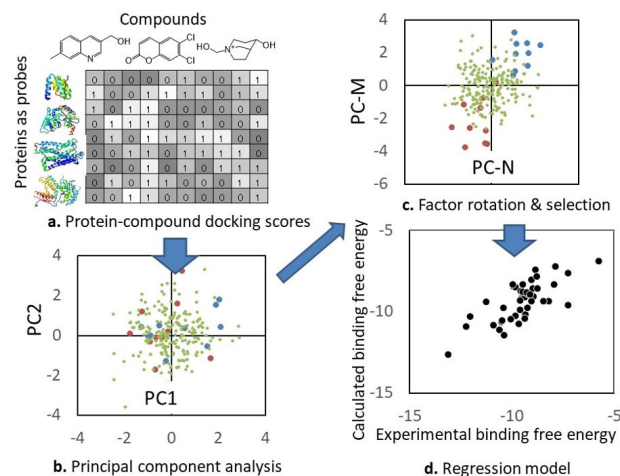


**Figure 1.** Schematic representation of the pharmacophore expansion of a ligand-binding site of a protein. The interaction between the ligand and protein pharmacophores gives the binding energy. **a:** Example of pharmacophore of the ligand pharmacophore. The dotted circles represent the pharmacophore features. **b:** Example of pharmacophores. HH, HA, and HD indicate hydrophobic, hydrogen-bond acceptor, and hydrogen-bond donor sites, shown by open, grey and black circles, respectively. Only three pharmacophores ( $\phi_1$ ,  $\phi_2$ , and  $\phi_3$ ) are depicted. The lines represent the specific distances between pharmacophore features of each pharmacophore. **c:** Example of pharmacophore expansion of the receptor pharmacophore following eqs. 2 and 3.

$$\Delta G_i^a = \sum_{r \in a} w_r^a k \left( \sum_{l \in i} g_l^r k(\phi_r, \phi_l) \right) = \sum_{r \in a} w_r^a \Delta G_i^r \quad (3)$$

Various protein pockets correspond to the various pharmacophores, and they work as probes for a given compound instead of  $\phi$  in eqs. 2 and 3. Thus the binding free energy can be estimated by the regression based on the docking scores for the various protein structures. This is a simplified model, since it does not include the intramolecular interaction or the conformational entropy of the compound.

Figure 2 shows the procedure of the present QSAR method. This method requires a learning set of 3D structures of compounds, the binding energy data between those compounds, and target proteins. We assume that protein–ligand docking programs give  $\Delta G_i^a$ . We proposed an approximation, as follows.<sup>[18]</sup>



**Figure 2.** Schematic representation of the docking-score QSAR method. **a:** The table represents a protein–compound docking score matrix. In the table, the values are depicted in grayscale. **b:** The PC analysis projects the score vector (column vector) of each compound into a point in the PC space. In the PC analysis, each dot represents a compound. The red, green, and blue dots represent the strong, medium, and weak affinity compounds, respectively. In this example, the first and second principal component axes (PC1 and PC2) are not useful to describe the affinity difference. **c:** The factor rotation method selected the representative axes (PC-M and PC-N) from the total  $N_p$  axes. PC-N and PC-M describe the affinity difference clearly. **d:** Finally, the regression model is constructed by using PC-N and PC-M.

$$\Delta G_i^a \approx \sum_b R_b^a \cdot s_i^b + \beta^a \quad (4)$$

Here,  $s_i^b$ ,  $R_b^a$ , and  $\beta^a$  are the docking score of the  $i$ -th compound to the  $b$ -th protein, the weight parameter, and the parameter for fitting, respectively. The set of  $\{b\}$  can include the target protein (the  $a$ -th protein). Equation 4 showed the RMSE of predicted binding energies was 1.5 kcal/mol.<sup>[18]</sup> One of the most serious problems with QSAR models is, in general, the limited range of applicability domains, since these models cannot work for input data that is too different from the training data set.<sup>[39]</sup> Since docking scores have been developed to mimic binding free energy, we assume that a docking score is equal to the binding free energy,  $\Delta G_i^a = s_i^a$ ,  $R_a^a = 1$ ,  $\beta^a = 0$ , and  $R_b^a = 0$  for  $a \neq b$  in eq. 4. In this case, eq. 4 can work without any experimental affinity data, and the problem of identifying an applicability domain is avoided.

Eq. 4 gives a linear regression model whose descriptors are docking scores, and the number of parameters is equal to the number of proteins.

In the previous<sup>[18]</sup> and present models, the protein ( $a$ ) – compound ( $i$ ) binding energy  $\Delta G_i^a$  is approximated by the PCR method based on the protein–compound docking scores  $\{s_i^b\}$ . The optimal principal component (PC) axis was selected to maximize the correlation coefficient by the

leave-one-out (LOO) cross-validation test. The PC axes are selected by factor rotation (see Figure 2). The factor rotation method selects the axes that show major contributions in the PC analysis among the total axes. We rewrote eq. 4, as follows.

$$\Delta G_i^a = \sum_{j=1}^{N_{axis}} c_j^a \cdot p_i^j + \beta^a \quad (5)$$

$$p_i^j = \sum_{b=1}^{N_p} d_b^j \cdot (s_i^b - \bar{s}^b). \quad (6)$$

Here,  $c_j^a$ ,  $\beta^a$ ,  $p_i$ , and  $d_b^j$  are the parameter, offset parameter, principal component vector, and loading vector, respectively. The total number of optimal PC axes is  $N_{axis}$ . The upper bar represents an average. The PC axis of the protein–compound docking-score matrix  $s$  gives the loading vector  $d$  and the principal component vector (axis)  $p$ . The parameters  $c$  and  $\beta$  are determined by multilinear regression (MLR).

$N_{axis}$  and  $N_p$  are the number of selected axes by the factor rotation and the total number of proteins used in the docking study as the compound descriptors, respectively. Here,  $N_{axis}$  ( $N_{axis} < N_p$ ) is determined to maximize the correlation coefficient obtained by the LOO cross-validations. The parameters are determined based on the learning set and then are used for prediction.

To calculate the protein–compound docking scores  $s_i^b$ , we used our own program, SievGene,<sup>[40,41]</sup> a protein–ligand flexible docking program for *in silico* drug screening. SievGene is a part of the myPresto system, which is available online (<http://presto.protein.osaka-u.ac.jp/myPresto4/>) and is free for academic use.

## 2.2 New Prediction Model with Restricted Regression Method

In the present study, we introduced a regularization term and robust estimation into the previously developed QSAR model based on docking scores using eqs. 5–6. In eq. 5, the coefficients  $c_j^a$  and  $\beta$  were determined by the multiple linear regression in our previous study.<sup>[18]</sup> In some cases,  $c_j^a$  was unrealistically large ( $10^3$ – $10^5$ ). The value should be 0 or 1 when the target protein structure is used in the docking calculation (we call this  $c_j^a$  value “ideal value”), and the docking score is equal to the binding energy. We would like to restrain  $c_j^a$  to around this values. In the present study, the coefficients were determined by minimizing the objective function by introducing the regularization term. Let  $Objf(a)$  be the objective function for the  $a$ -th target protein for the determined parameters  $c$  and  $\beta$ .

$$Objf(a) = \sum_{i=1}^{N_{cmp}} \left( \Delta G_{exp,i}^a - \left( \sum_{j=1}^{N_{axis}} c_j^a \cdot p_i^j + \beta^a \right) \right)^2 + \lambda \sum_{j=1}^{N_{axis}} (c_j^a - c_{ideal,j}^a)^2 \quad (7)$$

Here,  $c_{ideal,j}^a$  is the ideal value of  $c_j^a$ , and  $\beta^a$  and  $\lambda$  are parameters.  $N_{cmp}$  is the total number of compounds.  $c_{ideal,j}^a$  is unknown.  $\Delta G_{exp}$  represents the experimental  $\Delta G$  value. The last term restrains  $c_j^a$  to around the ideal values. Equation 7 is a generalized version of the Tikhonov regularization.<sup>[19–21]</sup> To estimate the  $\lambda$  value, we considered the following things. Since eq. 4 suggest that  $c_i^a = \Delta G_i^a$  under the ideal conditions ( $\Delta G_i^a = s_i^a$ ,  $R_a^a = 1$ ,  $\beta^a = 0$ , and  $R_b^a = 0$  for  $a \neq b$ , discussed in section 2.1) and that  $\{c_j^a\}$  values should correspond to  $\Delta G$  values. The larger the  $\lambda$  value is, the smaller the  $\{c_j^a\}$  values are. In general, most  $|\Delta G| < 18$  kcal/mol. In the present study,  $\lambda$  was set to 0, 0.0001, 0.0002, ..., 0.01, to satisfy  $\{c_j^a\} < 18$  kcal/mol. Also,  $c_{ideal,j}^a$  was set to 0, since the protein set providing the docking scores does not include the target protein structures in the present study.

In addition, we apply the maximum-likelihood-like estimation (M estimation), a robust estimation method.<sup>[42]</sup> The M estimation method weights the difference between a calculated value and an experimental value considering the predicted experimental error. The M estimation version of our objective function is given as follows:

$$Objf(a) = \sum_{i=1}^{N_{cmp}} w(d_i^a) \left( \Delta G_{exp,i}^a - \left( \sum_{j=1}^{N_{axis}} c_j^a \cdot p_i^j + \beta^a \right) \right)^2 + \lambda \sum_{j=1}^{N_{axis}} c_j^{a^2} \quad (8)$$

where

$$w(d) = \begin{cases} \left(1 - \left(\frac{d}{W}\right)^2\right)^2 & |d| \leq W \\ 0 & |d| > W \end{cases} \quad (9)$$

and

$$d = \sum_{i=1}^{N_{cmp}} \Delta G_{exp,i}^a - \left( \sum_{j=1}^{N_{axis}} c_j^a \cdot p_i^j + \beta^a \right) \quad (10)$$

Here,  $W$  and  $d$  are the upper limit of allowed error and a scalar value, respectively. The  $d$  value is the difference between the experimental value and the fitted value. The parameters  $c$  and  $\beta$  are determined to minimize the objective function. The derivation of eq. 8 is not linear, and we solve eqs. 8–10 by an iterative procedure. The M

estimation method places a higher weight on likely reliable data than unreliable data.  $W=0, 5, 10, 15, \dots, 100$  were examined in the present study. We call this model (eq. 8) the “docking score QSAR model”.

### 2.3 Generation of the Docking-score Index by Protein–compound Docking

The protein–compound docking scores  $s_i^b$  were calculated by the protein–compound docking program SievGene.<sup>[40]</sup> This ligand-flexible program reconstructed about 50% of the receptor–compound complexes in PDB (132 in total) with an accuracy of less than 2 Å root mean square deviation (RMSD) in a self-docking test.<sup>[40]</sup> The computational setup in the present study was exactly the same as that in the previous study. Namely, SievGene generated up to 100 conformers for each compound, and  $200 \times 200 \times 200$  grid potentials were adapted for all proteins. The pocket regions were suggested by the coordinates of the original ligands in the receptor–compound complex structures, and each edge length of the grid was about 35–45 Å. The docking-scoring function is based on the physical chemistry (accessible surface area, van der Waals potential, and electrostatic potential). The estimated error in binding free energy is almost 2.5 kcal/mol.<sup>[42]</sup> It takes 1 second to dock one compound against one protein on a single core of a Xeon 5570 CPU (2.98 GHz).

### 2.4 Probe Protein Sets

To generate  $\{s_i^b\}$  in eqs. 5–8, we performed a protein–compound docking simulation based on the soluble protein structures registered in the Protein Data Bank (PDB). The probe protein set consisted of 600 arbitrarily selected protein structures, as in our previous study (see APPENDIX A in Supporting Information). All of these structures were protein–ligand complexes. The protein set did not include the present target proteins. For protein sets, the complexes containing a covalent bond between the protein and ligand were removed, and all missing hydrogen atoms were added to form all-atom models of the proteins. All water molecules and cofactors were removed from the protein structures. All Asp and Glu were prepared as negatively charged forms, while Lys and Arg were prepared as positively charged forms. The atomic charges of the proteins were the same as those in AMBER parm99.<sup>[43]</sup> The docking pocket of each protein was indicated by the coordinates of the original ligand.

### 2.5 Training Set: Target Proteins and Compounds

To compare the present and previous results, the compounds and their assay information (compound structures, affinities against kinases) were downloaded from the Kinase

SARfari website (<https://www.ebi.ac.uk/chembl/sarfari/kinase-sarfari/downloads>) in the ChEMBL database, as in our previous work.<sup>[18]</sup> Note that the ChEMBL main page does not link to the KinaseSARfari website directly. The biochemical assay data, namely,  $K_i$ ,  $IC_{50}$ , %residual activity, and/or %inhibition values of human kinase protein-inhibitor systems, were also extracted from the bioactivity table in KinaseSARfari, and these data were converted to binding free energy by the software package used in our previous report.<sup>[18]</sup> The biochemical assay data were translated into the binding free energy by the Cheng–Prusoff equation and others.<sup>[14,44]</sup> We assumed that the experimental conditions were the same in all the assays. The procedure is described in details in APPENDIX B in the Supporting Information.

The first target was the kinase family. As target proteins, 107 kinases were selected. The 3D structures of the compounds were energy-optimized by cosgene<sup>[32]</sup> with the general AMBER force field (GAFF),<sup>[45]</sup> and the atomic charges were calculated by the MOPAC AM1 model using the Hgene program of the myPresto suite. Each functional group in all molecules was set to the dominant ionic form at pH 7. Finally, the filter condition reduced the number of data points used, and 45,663 assay data points of 107 kinases were derived.

The second target was the MMP family. We selected MMP2, MMP3, MMP7, MMP9, and MMP13. The protein structures were extracted from the PDB. The PDB IDs were 1hov, 2y6d, 4g9l, 5b5o, and 5cuh for MMP2, MMP3, MMP7, MMP9, and MMP13, respectively.<sup>[46–50]</sup> The protein structures were prepared in the same manner as the kinases. The zinc atom charges of these MMPs were set to +2, which is the formal charge of the most stable singlet zinc ion. As mentioned in the previous works, the actual charges were smaller than the formal charge and the charge values should differ from each other depending on the pocket structures. The protein–compound interaction data were extracted from the ChEMBL database. The compound structures and the  $\Delta G$  values were prepared in the same manner as the kinases.

### 2.6 Definitions of $Q^2$ and RMSE

The definition of  $Q^2$  and root-mean-square error (RMSE) are determined as follows.

$$Q^2 = 1 - \frac{\sum_{i=1}^{N_{cmp}} (\Delta G_{pred,i} - \Delta G_{exp,i})^2}{\sum_{i=1}^{N_{cmp}} (\Delta G_{exp,i} - \overline{\Delta G_{exp}})^2} \quad (11)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_{cmp}} (\Delta G_{pred,i} - \Delta G_{exp,i})^2}{N_{cmp}}} \quad (12)$$

Here,  $\Delta G_{pred}$ ,  $\Delta G_{expr}$  and the upper bar represent the predicted in validation and experimental  $\Delta G$  values and the average, respectively. In the present study, we do not compare the  $Q^2$  values of kinases to that of MMPs, since the variances of the experimental data were different to each other.<sup>[51]</sup>

### 3 Results and Discussion

#### 3.1 Cross-validation Tests of the Docking-score QSAR Model

Each of the compounds that gave assay data for one or more of the 107 target proteins was docked to all proteins of a protein set to generate the protein–compound docking-score matrix  $s$ . Then we adopted eq. 8 with changes to  $\lambda$  and  $W$  values and the LOO cross-validation test to calculate the  $Q^2$ . In addition, we applied the 4-fold cross validation test in all kinase cases to verify the results.

Table 1 summarizes the RMSE and  $Q^2$  values of  $\Delta G$ s calculated by the docking-score QSAR model with changes to the  $\lambda$  value in eq. 8 without the M estimation. The  $Q^2$  and RMSE depended on  $\lambda$ , and  $\lambda=0.0001$  showed the best  $Q^2$  and RMSE. These values did not change appreciably when  $\lambda$  was  $>0.00001$  and  $<0.005$ . The regularization term worked well, and  $\lambda$  was set to around 0.0001–0.005 in the following calculations. In the present study, our regression model did not use the docking scores for the target proteins and instead used the 600 probe proteins. Indeed, we found the coefficient  $R_b^a$  values reasonably low. Namely, the maximum and minimum values of  $R_b^a$  were 3.9 and  $-3.8$ , respectively, and the average and standard deviations of  $R_b^a$  were 0.01 and 0.1, respectively.

**Table 1.** Average  $Q^2$  and RMSE values obtained by the LOO cross validations of the docking-score QSAR model with various  $\lambda$  and  $W=0$  for all 107 proteins.

$\lambda$	$Q^2$	RMSE (kcal/mol)	$\lambda$	$Q^2$	RMSE (kcal/mol)
0	0.423	1.544	0.0005	0.704	1.074
0.00001	0.702	1.087	0.001	0.702	1.076
0.00002	0.703	1.081	0.002	0.687	1.089
0.00005	0.704	1.077	0.005	0.680	1.101
0.0001	0.705	1.074	0.01	0.669	1.122
0.0002	0.704	1.075	0.1	0.392	1.627

In eq. 8, the estimated error  $d$  was unknown a priori. The  $d$  value was estimated by an iterative solution method. Starting from  $d=0$ , the new  $d$  value was estimated by using the previous  $d$  value. The iteration converged within 4–6 steps in all cases. Tables 2 and 3 summarize the  $Q^2$ , and RMSE of calculated  $\Delta G$ s obtained by the docking-score QSAR model with changes to the  $W$  value in eq. 9. The data

**Table 2.** Average  $Q^2$  and RMSE values obtained by the LOO cross validations of the docking-score QSAR model with various  $W$  for all 107 proteins.

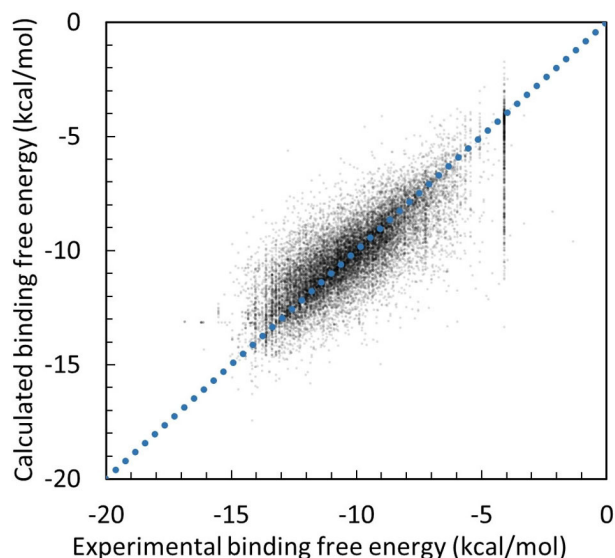
$W$ (kcal/mol)	$\lambda=0.0001$		$\lambda=0.0002$	
	$Q^2$	RMSE (kcal/mol)	$Q^2$	RMSE (kcal/mol)
–	0.702	1.087	0.704	1.075
5	0.622	1.27	0.623	1.268
10	0.698	1.099	0.698	1.097
15	0.704	1.077	0.704	1.078
20	0.704	1.076	0.704	1.075
50	0.705	1.074	0.705	1.074
100	0.704	1.076	0.705	1.074

**Table 3.** Average  $Q^2$  and RMSE values obtained by the 4-fold cross validations of the docking-score QSAR model with various  $W$  for all 107 proteins.

$W$ (kcal/mol)	$\lambda=0.002$		$\lambda=0.005$	
	$Q^2$	RMSE (kcal/mol)	$Q^2$	RMSE (kcal/mol)
–	0.629	1.236	0.629	1.225
5	0.577	1.345	0.579	1.334
10	0.627	1.240	0.627	1.230
15	0.629	1.235	0.629	1.225
20	0.628	1.236	0.629	1.224
50	0.628	1.236	0.629	1.224
100	0.629	1.235	0.629	1.224

in Tables 2 and 3 were obtained by the LOO and 4-fold cross validation tests, respectively. The  $Q^2$  and RMSE values were improved when  $20 \text{ kcal/mol} < W < 100 \text{ kcal/mol}$  in many cases compared to the result by eq. 7 ( $W$  value is “–” in Tables 2 and 3) and the prediction with  $W=5 \text{ kcal/mol}$  gave the worst  $Q^2$  and RMSE values (Figure S1). The  $Q^2$  and RMSE values depended on  $W$  weakly, and the results with  $W > 20 \text{ kcal/mol}$  were almost equal to each other. The M estimation improved the results slightly.

Equation 8 with  $\lambda=0.0002$  and  $W=20 \text{ kcal/mol}$  in the LOO cross validation and that with  $\lambda=0.005$  and  $W=20 \text{ kcal/mol}$  in the 4-fold cross validation gave the best  $Q^2$  and RMSE values. Figure 3 shows the results of the LOO cross-validation test with  $\lambda=0.0002$  and  $W=20 \text{ kcal/mol}$  for all 107 kinases. The average error reached 1.08 kcal/mol and  $Q^2=0.70$  in the binding free energy. Some of the datasets showed very high accuracy, considering that the thermal fluctuation was about 0.6 kcal/mol at room temperature (Table S1). The results of the 4-fold cross-validation test with  $\lambda=0.005$  and  $W=20 \text{ kcal/mol}$  for all 107 kinases showed qualitatively the same trends as those in Figure 3 (Table S2). The  $W$  value was originally in the acceptable error range, and the regression/analysis ignored the data points with the estimation error ( $d$ )  $> W$ . The optimal  $W$  value was much bigger than the standard deviation of  $\Delta G$  values that are multiply observed for each target protein–compound pair



**Figure 3.** Correlation between the experimental and prediction data for all 107 kinase proteins obtained by the docking-score QSAR model with  $\lambda=0.0002$  and  $W=20$  kcal/mol. The dots represent the predicted data points by the LOO cross-validation test.

(see section 3.2), meaning that the regression used all the data points with almost equal weight.

### 3.2 Effect of Elimination of Outliers from the Experimental Data

There are multiple experimental affinity data for some single protein–compound pairs in the database. This is because the experimental data depend on the experimental conditions such as pH, temperature, density of buffer salt, and cell line, and then the unique protein–compound pairs in the database correspond to these different experimental affinity data points under different experimental conditions. Some kinases are anti-cancer drug targets, and the amino-acid mutation of kinases causes drug resistance. The database includes such kinase data, and these mutants with different affinities share unique protein IDs in the database.

In this case, a drug-resistant protein should show weaker binding affinity to the same compound than to a native protein. The enzyme activity could depend on the effector protein in the protein–protein interaction network. In this case, the observed enzyme activity depends on the experimental environment, such as whether it is *in vivo* or *in vitro*. Also, the protein activities depend on the temperature, pH, and density of the salt of the buffer solvent. These conditions are not shown in the ChEMBL database clearly.

The average standard deviation of the  $\text{Log}_{10}K_i$  affinity data was about 0.24 kcal/mol, that of  $\text{Log}_{10}IC_{50}$  was about 0.44 kcal/mol, and that of  $\text{Log}_{10}$  (%-inhibition or %-residual

activity) was 0.093 kcal/mol. When all the experimental data were translated to  $\Delta G$ , the average deviation of the binding affinity was about 0.73 kcal/mol (Table S3).

We examine the effect of eliminating outliers among multiple data on the prediction result. When multiple affinity data correspond to a single pair of a protein ID and a compound ID, we eliminate outliers among the multiple data. Let the number of multiple data points for the  $a$ -th protein and the  $i$ -th compound be  $M_a^i$  and the  $k$ -th affinity value be  $E_a^i(k)$ . We define  $E_a^i(k)$  as an outlier of the data set when  $E_a^i(k)$  satisfies the following relationship.

$$|E_a^i(k) - \Delta G_a^i| > N \cdot \sigma_a^i \quad (13)$$

$$\sigma_a^i = \sqrt{\frac{\sum_{m=1}^{M_a^i} (E_a^i(m) - \Delta G_a^i)^2}{M_a^i - 1}} \quad (14)$$

where  $\Delta G_a^i$  is the average value of a set of  $E_a^i(m)$ ,  $m=\{1, \dots, M_a^i\}$ , and  $\sigma_a^i$  is the deviation of the data. Equation 14 defines the  $k$ -th compound as an outlier when the  $k$ -th compound satisfies this condition. In the present study, we removed outliers from all experimental data trying  $N=0.2, 0.4, 0.5, 0.6, 0.8, 1, 2,$  and  $3$  before the following cross-validation tests.

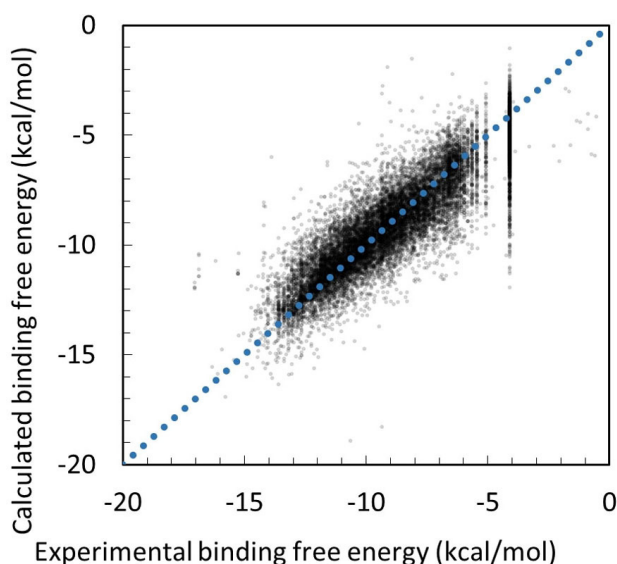
Table 4 summarizes the cross-validation results for  $N=0.2, 0.4, 0.5, 0.6, 0.8, 1, 2,$  and  $3$ . In Table 4, the regression models used were eq. 8 with  $\lambda=0.0002$  (LOO cross validation) and  $\lambda=0.005$  (4-fold cross validation), respectively. In both tables,  $W=20$  kcal/mol. In all cases, the elimination of outliers in the test data set improved the prediction results with decreasing  $N$ . Elimination of outliers worked well when  $N \sigma$  was set to  $N < 0.8$ . Figure 4 shows the result of the 4-fold cross-validation test for all 107 kinases. The result by the LOO cross validation was similar to Figure 4. Some of the data points are located vertically at  $-4$  kcal/mol. These data points correspond to 0% inhibitions or 100% residual activities. Except for these data points, the predicted data points clearly correlate to the experimental data. Also, we examined the 4-fold cross validation case with  $\lambda=0.002$ , and the result was close to that summarized in Table 5 (Table S4).

When  $N=0.2, 0.5, 0.8,$  and  $1.0$ , eqs. 13–14 eliminated 23%, 21%, 17%, and 4% of the data, respectively, out of the total of 45,663 data points. Since the standard deviation of the experimental  $\Delta G$  values was 0.7 kcal/mol, the results of  $\text{RMSE} < 0.7$  kcal/mol should be the accurately predicted cases. The outlier elimination increased the accurately predicted cases ( $\text{RMSE} < 0.7$  kcal/mol) from 20 to 42 cases out of the 107 targets in the LOO cross validation.

We selected 24 targets arbitrarily from the total of 107 kinases examined above (Table S1) and analyzed the individual results carefully. For these targets, 12 protein structures were available in the PDB (Table S5). These protein coordinates were prepared in the same manner as

**Table 4.** Average  $Q^2$  and RMSE values obtained by the LOO cross validation ( $W=20$  kcal/mol and  $\lambda=0.0002$ ) and 4-fold cross validation ( $W=20$  kcal/mol and  $\lambda=0.005$ ) tests of the docking-score QSAR model with various  $N$  for all 107 proteins.

$N\sigma$	Total no. of compounds	LOO cross validation		4-fold cross validation	
		$Q^2$	RMSE (kcal/mol)	$Q^2$	RMSE (kcal/mol)
0.2 $\sigma$	35050	0.762	0.919	0.678	1.132
0.4 $\sigma$	35736	0.763	0.913	0.680	1.134
0.5 $\sigma$	36249	0.763	0.915	0.679	1.133
0.6 $\sigma$	36804	0.758	0.918	0.678	1.136
0.8 $\sigma$	38124	0.760	0.913	0.685	1.117
$\sigma$	44063	0.704	1.052	0.645	1.192
2 $\sigma$	45549	0.694	1.079	0.631	1.220
3 $\sigma$	45650	0.693	1.084	0.631	1.223



**Figure 4.** Correlation between experimental and prediction data for all 107 kinase proteins obtained by the docking-score QSAR model eliminating outliers with  $\lambda=0.005$ ,  $W=20$  kcal/mol, and  $N=0.8$ . The dots represent the predicted data points by the 4-fold cross-validation test.

**Table 5.**  $Q^2$  and RMSE data obtained by the SievGene docking program and the LOO cross validation tests of the docking-score QSAR model over all MMPs.

Protein	No of ligands	Naïve docking		Docking-score QSAR	
		$Q^2$	RMSE (kcal/mol)	$Q^2$	RMSE (kcal/mol)
MMP-2	489	0.004	1.81	0.774	1.23
MMP-3	369	0.029	2.50	0.672	1.48
MMP-7	98	0.058	0.99	0.922	0.46
MMP-9	445	0.048	1.82	0.64	1.61
MMP-13	148	0.044	1.34	0.903	0.62
Average	309.8	0.037	1.692	0.782	1.08

described in sections 2.3 and 2.4. We applied the docking-score QSAR method ( $\lambda=0.005$ ,  $W=20$  kcal/mol,  $N=0.8$ , 4-fold cross validation) and the naïve docking. In each individual target, the docking-score QSAR result was better than that by the naïve docking study (Figure S2). Namely, the average RMSE values by the naïve docking and the docking-score QSAR were 1.6 and 1.2 kcal/mol, respectively. The average  $Q^2$  values were 0.11 and 0.65 for the naïve docking and the docking-score QSAR method, respectively (Table S5).

In some cases, the outlier elimination did not work and the prediction accuracy remained low. The data sets with more than 1000 data points showed particularly low accuracy, such as the sets for cyclin-dependent kinase 2, epidermal growth factor receptor erbB1, tyrosine-protein kinase SRC, vascular endothelial growth factor receptor 2, and MAP kinase p38 alpha. The reason for this was unclear. When a target protein has multiple ligand binding sites, such as orthosteric and allosteric sites, the linear regression model is not suitable. In this case, nonlinear regression, such as logistic regression and neural networks, may solve the problem. But the present model is based on eqs. 2–4. When the target protein structure gave the docking score without computational error, then  $\Delta G_i^a = s_i^a$ . The nonlinear regression must satisfy this simple condition. This problem is somewhat troublesome. We will examine it in the future.

The ChEMBL database did not provide details on the experimental conditions (density of native ligand, temperature, etc.). The  $\Delta G$  values should be precise by using the actual experimental information. But this method requires natural language analysis, which is an expensive approach. One possible improvement of the present method might be batch effect correction.<sup>[52]</sup> That is, we could determine some optimal artificial values as the experimental parameters, such as the densities of ligands and substrates, in order to minimize the prediction error. In this approach, first the present method generates a prediction model based on the experimental data with the standard experimental values, then the batch-effect correction method would optimize the experimental parameters of each batch to minimize the computational error of the set of data points of the batch.

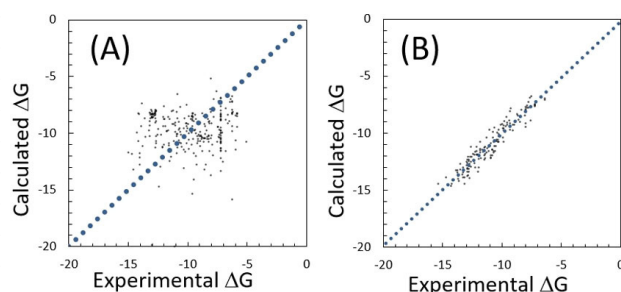
The outlier elimination improved accuracy more than the M-estimation did. This gap might be attributable to the fact that the standard deviations of  $\Delta G$  values differ from each other among the different targets and different compound sets, while the  $W$  value is consistent throughout all the data. This difference could be the reason why the outlier elimination improved the results better than the M-estimation.

### 3.3 Application to MMPs

We studied MMP2, MMP3, MMP7, MMP9, and MMP13 by the docking-score QSAR method and the naïve protein–compound docking simulations. We applied the



naïve protein–ligand docking calculation by Sievgene.<sup>[40]</sup> Figure 5A shows the correlation between the experimental and the calculated protein–compound binding free energies of MMP2 by Sievgene. The averaged values of the  $Q^2$  and RMSE were 0.037 and 1.692 kcal/mol, respectively. Table 5 summarizes the total number of ligands as well as the  $Q^2$  and RMSE values of MMPs by naïve docking. The accuracy was poor and the results did not show any experimental trends, since the Sievgene docking program does not have a metal-contact term to support the metal ions.



**Figure 5.** Correlation between experimental and prediction data for MMP2 (unit in kcal/mol). (A) obtained by the naïve protein–ligand docking calculation by Sievgene, and (B) obtained by the docking-score QSAR model with  $\lambda=0.0002$ ,  $W=20$  kcal/mol. The dots represent the predicted data points by the LOO cross-validation test.

Next, we applied the docking-score QSAR method to the same data sets. We applied the LOO cross-validation test to the MMPs. Figure 5B shows the correlation between the experimental and calculated protein–compound binding free energies of MMP2 by eq. 8. The  $Q^2$  and RMSE values by the LOO cross-validation test were 0.88 and 1.08 kcal/mol, respectively. The RMSE values by the docking-score QSAR method were much better than those by the naïve docking, while both methods used the same docking program (Table 5).

We checked the individual correlation results. The results obtained by the docking-score QSAR method were much better than those by the naïve docking in many cases. The overall trends of the results were the same as those in the kinase cases (Figure S3). Namely, the average RMSE value by the naïve docking studies was 1.7 kcal/mol. These values were close to the deviation of the experimental  $\Delta G$  values. On the other hand, the docking-score QSAR showed an average RMSE of 1.1 kcal/mol and a better  $Q^2$ .

The docking-score QSAR method used only the Sievgene docking scores and was exactly the same program used in the above naïve docking study without any parameter tuning for the quantum mechanical interaction between the metal and the ligand. The docking-score QSAR method should take into account the quantum mechanical

interaction implicitly to improve the prediction results. The RMSE values were similar to those given in Section 3.2.

The docking-score QSAR method was based on the docking program without consideration of the quantum effect between the metal atom and the ligands, but it improved the RMSE by 0.6 kcal/mol, the same as in the kinase cases. Thus, the present method predicted  $\Delta G$  values with an RMSE of 1 kcal/mol even for the difficult targets like the MMPs, and it could be applied to general target proteins.

The force-field parameters for the metal–ligand interaction are generally poor, because such interactions should include quantum effects, which depend on the environment. Consequently, the naïve docking simulations could not provide good scores for the MMP systems. The current docking-score QSAR model is based on a docking score that does not include a quantum effect, but this QSAR model is a completely different approach from that of the naïve docking study. The machine learning procedures provided smaller RMSE value than the naïve docking study.

## 4 Conclusions

We developed a docking-score QSAR model based on combinations of multiple docking scores from protein–drug docking simulations, and applied this model to 107 kinase proteins from the ChEMBL public database. The prediction model employed a descriptor-based weighted PCR with a regularization term and robust estimation (M estimation) methods in order to realize more realistic prediction than that by the ordinal multilinear regression model. The compound descriptor was a set of docking scores against many nontarget proteins. The LOO and 4-fold cross-validation tests showed that the addition of the regularization term improved the RMSE from 1.5 kcal/mol to 1.1 kcal/mol. In addition, the data preparation with outlier elimination worked to improve the results. We also applied the present method to MMPs that were difficult targets because of the quantum mechanical interaction. The present method improved the RMSE values for the MMPs without any manual parameter tuning, the same as in the kinase cases. The LOO cross-validation tests showed that the docking-score QSAR method improved the RMSE from 1.7 kcal/mol by the naïve docking calculation to 1.1 kcal/mol. These results suggested that this method is applicable to general target proteins. The analysis performed was based on the cross-validation tests only, and there was no prospective experimental validation in this study. Further analysis of the validation tests designed for extrapolation should be performed.

## Supporting Information

The appendices, Figures S1–S3, and Tables S1–S5 were supplied as described in the Supporting Information.

## Abbreviations

QSAR: quantitative structure–activity relationship  
LOO: leave-one-out  
PCA: principal component analysis  
PCR: principal component regression  
RMSE: root-mean-square deviation  
 $Q^2$ : coefficient of determination of prediction

## Conflict of Interest

None declared.

## Acknowledgments

This work was supported by grants from the National Institute of Advanced Industrial Science and Technology (AIST), the Japan Agency for Medical Research and Development (AMED), and the Ministry of Economy, Trade, and Industry (METI) of Japan.

## References

- [1] Y. Yang, S. J. Adelstein, A. I. Kassis, *Drug Discovery Dev.* **2009**, *14*, 147–154.
- [2] C. G. Wermuth, *Drug Discovery Dev.* **2006**, *11*, 160–164.
- [3] J. H. Nettles, J. L. Jenkins, A. Bender, Z. Deng, J. W. Davies, M. Glick, *J. Med. Chem.* **2006**, *49*, 6802–6810.
- [4] Y. Fukunishi, S. Kubota, H. Nakamura, *J. Chem. Inf. Model.* **2006**, *46*, 2071–2084.
- [5] R. P. Sheridan, K. Nam, V. N. Maiorov, D. R. McMasters, W. D. Cornell, *J. Chem. Inf. Model.* **2009**, *49*, 1974–1985.
- [6] S. Nijjima, A. Shiraiishi, Y. Okuno, *J. Chem. Inf. Model.* **2012**, *52*, 901–912.
- [7] E. Lounkine, M. J. Keiser, S. Whitebread, D. Mikhailov, J. Hamon, J. L. Jenkins, P. Lavan, E. Weber, A. K. Doak, S. Côté, B. K. Shoichet, L. Urban, *Nature* **2012**, *486*, 361–367.
- [8] A. Peragovics, Z. Simon, I. Brandhuber, B. Jelinek, P. Hari, C. Hetenyi, P. Czobor, A. Malnasi-Csizmadia, *J. Chem. Inf. Model.* **2012**, *52*, 1733–1744.
- [9] Z. Simon, A. Peragovics, M. Vigh-Smeller, G. Csukly, L. Tombor, Z. Yang, G. Zahoranszky-Kohalmi, L. Vegner, B. Jelinek, P. Hari, C. Hetenyi, I. Bitter, P. Czobor, A. Malnasi-Csizmadia, *J. Chem. Inf. Model.* **2012**, *52*, 134–145.
- [10] M. C. Cobanoglu, C. Liu, F. Hu, Z. N. Oltvai, I. Bahar, *J. Chem. Inf. Model.* **2013**, *53*, 3399–3409.
- [11] A. Peragovics, Z. Simon, L. Tombor, B. Jelinek, P. Hari, P. Czobor, A. Malnasi-Csizmadia, *J. Chem. Inf. Model.* **2013**, *53*, 103–113.
- [12] H. B. Engin, O. Keskin, R. Nussinov, A. Gursoy, *J. Chem. Inf. Model.* **2012**, *52*, 2273–2286.
- [13] Y. Pan, T. Cheng, Y. Wang, S. H. Bryant, *J. Chem. Inf. Model.* **2014**, *54*, 407–418.
- [14] J. Tang, A. Szwajda, S. Shakyawar, T. Xu, P. Hintsanen, K. Wennerberg, T. Aittokallio, *J. Chem. Inf. Model.* **2014**, *54*, 735–743.
- [15] M. Lindh, F. Svensson, W. Schaal, J. Zhang, C. Skold, P. Brandt, A. Karlen, *J. Chem. Inf. Model.* **2015**, *50*, 343–353.
- [16] Y. Wang, J. Xiao, T. O. Suzek, J. Zhang, S. H. Bryant, *Nucleic Acids Res.* **2009**, *37*, W623–W633.
- [17] A. Gaulton, L. J. Bellis, A. P. Bento, J. Chambers, M. Davies, A. Hersey, Y. Light, S. McGlinchey, D. Michalovich, B. Al-Lazikani, P. Overington, *Nucleic Acids Res.* **2011**, *40*, D1100–D1107.
- [18] Y. Fukunishi, S. Yamasaki, I. Yasumatsu, K. Takeuchi, T. Kurosawa, H. Nakamura, *Mol. Inf.* **2017**, *36*, 1600013.
- [19] R. Tibshirani, *J. R. Stat. Soc. B.* **1996**, *58*, 267–288.
- [20] A. E. Hoerl, R. W. Kennard, *Technomet* **1970**, *12*, 69–82.
- [21] G. H. Golub, M. Heath, G. Wahba, *Technomet* **1979**, *21*, 215–223.
- [22] O. P. J. van Linden, A. J. Kooistra, R. Leurs, I. P. de Esch, C. de Graal, *J. Med. Chem.* **2014**, *57*, 249–277.
- [23] T. Anastassiadis, S. W. Deacon, K. Devarajan, H. Ma, J. Peterson, *Nat. Biotechnol.* **2011**, *29*, 1039–1046.
- [24] M. I. Davis, J. P. Hunt, S. Herrgard, P. Ciceri, L. M. Wodicka, G. Pallares, M. Hocker, D. K. Treiber, P. P. Zarrinkar, *Nat. Biotechnol.* **2011**, *29*, 1046–1052.
- [25] C. T. Supran, A. Scozzafava, in *Proteinase and peptidase inhibition: recent potential targets for drug development*, Chapter 3 (Eds: H. J. Smith and C. Simons), Taylor and Francis, London **2002**, pp. 35–61.
- [26] O. G. Dadrass, A. M. Sobhani, A. Shafiee, M. Mahadoudian, *DARU* **2004**, *12*, 1–10.
- [27] R. Wu, Z. Lu, Z. Cao, Y. Zhang, *J. Chem. Theory Comput.* **2011**, *7*, 433–443.
- [28] P. Li, L. F. Song, K. M. Merz Jr., *J. Phys. Chem. B.* **2015**, *119*, 883–895.
- [29] D. S. Martins, S. Forli, M. J. Ramos, A. J. Olson, *J. Chem. Inf. Model.* **2014**, *54*, 2371–2379.
- [30] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, R. D. Taylor, *Proteins Struct. Funct. Bioinf.* **2003**, *52*, 609–623.
- [31] B. Seebeck, I. Reulecke, A. Kamper, M. Rarey, *Proteins Struct. Funct. Bioinf.* **2008**, *71*, 1237–1254.
- [32] Y. Fukunishi, Y. Mikami, H. Nakamura, *J. Phys. Chem. B.* **2003**, *107*, 13201–13210.
- [33] T. Singh, O. A. Adekoya, B. Jayaram, *Mol. BioSyst.* **2015**, *11*, 1041–1051.
- [34] R. H. Stote, M. Karplus, *Proteins Struct. Funct. Bioinf.* **1995**, *23*, 12–31.
- [35] X. Hu, W. H. Shelver, *J. Molec. Graph. Modell.* **2003**, *22*, 115–126.
- [36] X. Hu, S. Balaz, W. H. Shelver, *J. Molec. Graph. Modell.* **2004**, *22*, 293–307.
- [37] C. G. Wermuth, C. R. Ganellin, P. Lindberg, L. A. Mitscher, *Pure Appl. Chem.* **1988**, *70*, 1129–1143.
- [38] G. Wolber, T. Langer, *J. Chem. Inf. Model.* **2005**, *45*, 160–169.
- [39] M. Masuda, H. Kaneko, K. Funatsu, *Ind. Eng. Chem. Res.* **2014**, *53*, 8553–8564.
- [40] Y. Fukunishi, Y. Mikami, H. Nakamura, *J. Mol. Graph. Model.* **2005**, *24*, 34–45.
- [41] Y. Fukunishi, Y. Mikami, S. Kubota, H. Nakamura, *J. Mol. Graph. Model.* **2006**, *25*, 61–70.
- [42] M. Cohen, S. R. Dalal, J. W. Tukey, *Appl. Statist.* **1993**, *42*, 339–353.

- [43] D. A. Case, T. A. Darden, T. E. III. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, B. Wang, D. A. Pearlman, M. Crowley, S. Brozell, V. Tsui, H. Gohlke, J. Mongan, V. Hornak, G. Cui, P. Beroza, C. Schafmeister, J. W. Caldwell, W. S. Ross, P. A. Kollman, AMBER 8, University of California, San Francisco **2004**.
- [44] Y. Cheng, W. H. Prusoff, *Biochem. Pharmacol.* **1973**, *22*, 3099–3108.
- [45] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- [46] A. Topai, P. Breccia, F. Minissi, A. Padova, S. Marini, I. Cerbara, *Bioorg. Med. Chem.* **2012**, *20*, 2323–2337.
- [47] Y. Feng, J. J. Likos, L. Zhu, H. Woodward, G. Munie, J. J. McDonald, A. M. Stevens, C. P. Howard, G. A. De Crescenzo, D. Welsch, H. S. Shieh, W. C. Stallings, *Biochim. Biophys. Acta.* **2002**, *1598*, 10–23.
- [48] B. D. Belviso, R. Caliandro, D. Siliqi, V. Calderone, F. Arnesano, G. Natile, *Chem. Commun.* **2013**, *49*, 5492–5494.
- [49] C. Camodeca, E. Nuti, L. Tepshi, S. Boero, T. Tuccinardi, E. A. Stura, A. Poggi, M. R. Zocchi, A. Rossello, *Eur. J. Med. Chem.* **2016**, *111*, 193–201.
- [50] H. Nara, K. Sato, T. Naito, H. Mototani, H. Oki, Y. Yamamoto, H. Kuno, T. Santou, N. Kanzaki, J. Terauchi, O. Uchikawa, M. Kori, *J. Med. Chem.* **2014**, *57*, 8886–8902.
- [51] D. L. J. Alexander, A. Tropsha, D. A. Winkler, *J. Chem. Inf. Model.* **2015**, *55*, 1316–1322.
- [52] A. Vandenbon, V. H. Dinh, N. Mikami, Y. Kitagawa, S. Teraguchi, N. Ohkura, S. Sakaguchi, *Proc. Nat. Acad. Soc.* **2016**, *113*, E2393–E2402.

Received: September 27, 2017

Accepted: January 22, 2018

Published online on February 14, 2018