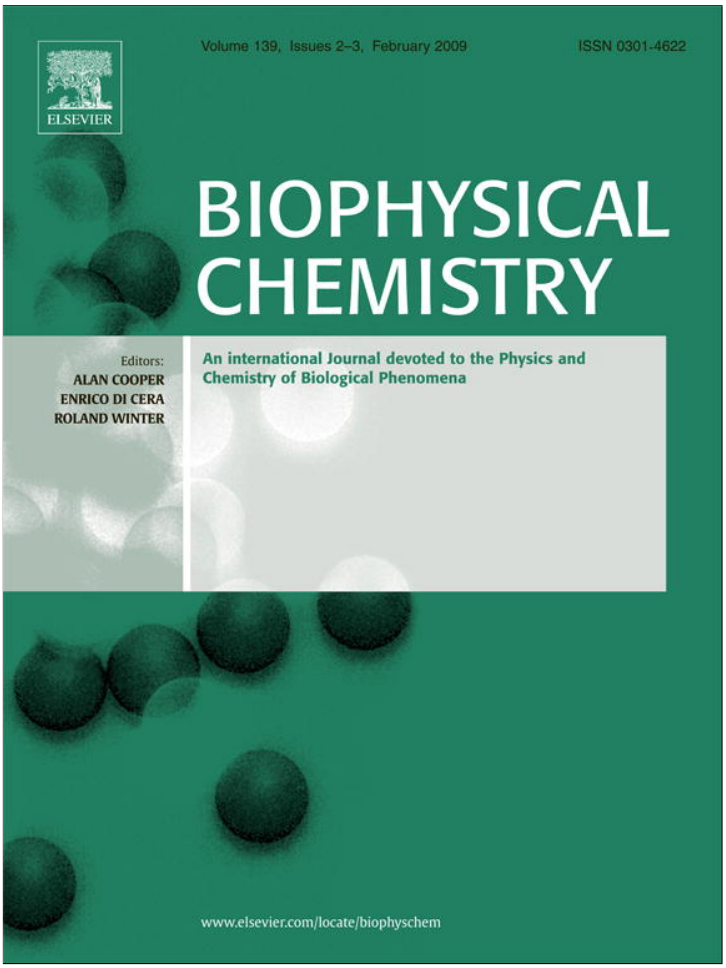
2-2009

Development and Validation of an Empirical Free Energy Function for Calculating Protein–Protein Binding Free Energy Surfaces

Joseph Audie
Sacred Heart University

Recommended Citation

Audie, J. (2009). Development and validation of an empirical free energy function for calculating protein–protein binding free energy surfaces. *Biophysical Chemistry,139*(2/3), 84-91. Doi: 10.1016/j.bpc.2008.10.007



Biophysical Chemistry 139 (2009) 84–91

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Development and validation of an empirical free energy function for calculating protein–protein binding free energy surfaces

Joseph Audie *

Department of Chemistry, Sacred Heart University, Fairfield, CT 06825, United States
CMD Bioscience, Orange, CT, United States

ARTICLE INFO

Article history:
Received 22 September 2008
Received in revised form 18 October 2008
Accepted 22 October 2008
Available online 8 November 2008

Keywords:
Free energy
Docking
Computational
Binding affinity
Scoring function
Protein–protein

ABSTRACT

In a previous paper, we described a novel empirical free energy function that was used to accurately predict experimental binding free energies for a diverse test set of 31 protein–protein complexes to within ≈1.0 kcal. Here, we extend that work and show that an updated version of the function can be used to (1) accurately predict native binding free energies and (2) rank crystallographic, native-like and non-native binding modes in a physically realistic manner. The modified function includes terms designed to capture some of the unfavorable interactions that characterize non-native interfaces. The function was used to calculate one-dimensional binding free energy surfaces for 21 protein complexes. In roughly 90% of the cases tested, the function was used to place native-like and crystallographic binding modes in global free energy minima. Our analysis further suggests that buried hydrogen bonds might provide the key to distinguishing native from non-native interactions. To the best of our knowledge our function is the only one of its kind, a single expression that can be used to accurately calculate native and non-native binding free energies for a large number of proteins. Given the encouraging results presented in this paper, future work will focus on improving the function and applying it to the protein–protein docking problem.

1. Introduction

Protein–protein and protein–peptide interactions are essential to life. Unfortunately, it is often difficult, costly and time consuming to obtain experimentally derived structural and energetic information about protein–protein interactions. Hence, efficient *in silico* or computational methods for predicting and structurally rationalizing protein–protein interactions are required to fill the gaps in our knowledge.

Previously, we developed, tested and reported on an empirical free energy function that was used to blindly and accurately predict experimental binding free energies (ΔG_{bind}), from native state crystallographic coordinates alone, for a large and diverse test set of protein–protein and protein–peptide interactions to within ≈1.0 kcal and with an $R^2 \approx 0.80$. Furthermore, we argued that the function made basic statistical, theoretical and physical sense and that the function was, at least within certain well-defined limits, explanatory. Importantly, the function can be used to estimate binding free energies in a matter of seconds [1].

Our previous analysis, while encouraging, left open the question as to whether or not the same empirical function could be used to accurately predict binding affinities for non-native interactions. Put differently, a generally valid and robust function should (1) make basic physical and theoretical sense, should be (2) predictive for native interactions and should be (3) predictive for non-native interactions too. As mentioned above, we have already argued for (1) and (2); in the present study we extend our previous work and test a slightly improved version of the empirical function according to condition (3).

Unfortunately, unlike the case for condition (2) testing a function according to condition (3) is no simple matter. This is because while experimental binding affinity data is available for evaluating condition (2) no such data is available to evaluate condition (3). Nonetheless, progress can be made by recognizing that, for a given pair of interacting proteins, non-native binding geometries must occupy higher points on the free energy surface than native-like and crystallographic binding geometries and that the larger the calculated native/non-native free energy gap the better. Given this logic, we used the HEX protein–protein docking server to construct 101 member decoy sets for 21 protein–protein complexes. Each of the 21 decoy sets included crystallographic, native-like and non-native interaction geometries. Binding free energies were then calculated for all protein–protein complexes and the results plotted as one-dimensional binding free energy surfaces. For a given protein pair, predicted native and non-native binding free energies were counted as accurate if the crystallographic and native-like binding modes were placed in deep global free energy minima that were also in close agreement with the experimental binding free energy.

At the outset it is important to note that the version of the free energy function reported on in the present paper is slightly different

* Tel.: +1 203 371 7986; fax: +1 203 371 7888.
E-mail address: audiej@sacredheart.edu.

0301-4622/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2008.10.007

from the function we originally described. Specifically, the function has been re-parameterized, a simple clash term has been added, and a solvent accessible surface area (SASA) requirement has been adopted for weighting the energetic contributions of hydrogen bonds and salt-bridges. Just as importantly, it should be noted that these changes have only improved on the functions physical and theoretical soundness, predictive accuracy for experimental binding affinities and as will be shown below, predictive accuracy for non-native interactions.

In the present paper, then, we report on our continued efforts to develop, test and validate a single mathematical expression for accurately predicting native and non-native protein–protein binding free energies in a fast and physics-based manner. The picture that emerges from the present study is that of an improved and more generally valid empirical function that can be used to consistently place native geometries in deep, well-defined, and physically realistic free energy minima and that can potentially be used to address many applied problems in protein modeling and engineering.

2. Materials and methods

2.1. Theoretical background

Previously, we reported on an empirical expression for calculating native state protein–protein binding free energies (ΔG_{bind}), assuming rigid-body association,

$$\Delta G_{bind} = -0.85\Delta X_{\pm} + 0.067\Delta X_{c/s} - 0.66X_{sb} - 0.90X_{hb} - 0.00087X_{gap} - 0.091\Delta X_{tor} - 0.54. \quad (1)$$

Since that time we re-parameterized the function using the original 24 member training set, minus the 1dhk complex. This resulted in a slightly different but improved version of the original function. The current function, then, has slightly different coefficient magnitudes and is given by,

$$\Delta G_{bind} = -0.79\Delta X_{\pm} + 0.075\Delta X_{c/s} - 0.65X_{sb} - 0.86X_{hb} - 0.00089X_{gap} - 0.089\Delta X_{tor} - 0.33 \quad (2)$$

where $\Delta X_{\pm}$ and $\Delta X_{c/s}$ refer to changes in the number of exposed charged (Asp and Glu carboxyl oxygens; Arg and Lys side chain nitrogens; and all N and O-termini) and hydrophobic atoms (all carbons and sulfurs), with an atom being defined as exposed when its solvent accessible surface area (SASA) > 1.0 Å^2. The third and fourth terms refer to the total number of hydrogen bonds (X_{hb}) and salt bridges (X_{sb}) across the protein–protein interface. Salt bridges were defined for positively and negatively charged receptor and ligand atoms separated by ≤4.0 Å and hydrogen bonds were defined according to the geometric and chemical criteria employed in our previous study. The final two terms refer to the interface gap (X_{gap}) or void volume and the total number of torsions immobilized (ΔX_{tor}) at the interface.

Eq. (2) is readily extended to include the effects of unfavorable interactions that might characterize non-native complex interactions,

$$\Delta G_{bind} = -0.79\Delta X_{\pm} + 0.075\Delta X_{c/s} - 0.65X_{sb} - 0.86X_{hb} - 0.00089X_{gap} - 0.089\Delta X_{tor} - 0.33 + 2.5X_{clash} + \Delta G_{sasa,polar-polar,hb} + \Delta G_{sasa,charge-polar,hb}. \quad (3)$$

Here X_{clash} refers to the total number of non-hydrogen bonded interface atoms that are involved in steric clashes (d < 2.5 Å). An energetic penalty ($\Delta G_{sasa,polar-polar,hb}$) is assessed when the average surface area for all polar–polar atoms involved in all interface hydrogen bonds exceeds a user-defined SASA threshold ($SASA_{polar-polar,hb,avg}$ = 8.0 Å^2); a higher maximum average SASA value is used for assessing the penalty ($\Delta G_{sasa,charge-polar,hb}$) associated with charge–polar hydrogen bonds ($SASA_{charge-polar,hb,avg}$ = 17.0 Å^2). In both cases the user-defined free energy penalties were set at 3.0 kcal ($\Delta G_{sasa,polar-polar,hb} = \Delta G_{sasa,charge-polar,hb}$ = 3.0 kcal). Finally, salt bridges were only counted if the two interacting charged atoms had a $SASA_{sb}$ < 25.9 Å^2. Eq. (3) was used to calculate ΔG_{bind} for all native (crystallographic and native-like) and non-native protein–protein interactions.

2.2. Decoy sets

To evaluate the use of Eq. (3) in predicting native and non-native binding free energies, protein–protein decoy sets had to be constructed. The decoys were generated using the HEX protein–protein docking server (http://www.csd.abdn.ac.uk/hex_server/)[2]. The 100 top scoring HEX complexes were selected as decoys for each of 21 complexes. All 21 complexes are high quality structures, drawn from the training set employed in our previous study, that also appear to satisfy most of the assumptions of our method (rigid body binding, simple charge model, etc.); experimental binding affinity data is also available for all 21 structures. 2tpi and 3cpa were excluded from the present study, as they include small peptide ligands. Moreover, the HEX server failed to return 100 low energy results for them. Information regarding all 21 complexes, including their experimental binding free energies, is listed in Tables 1 and 2.

When submitting jobs to the HEX server, the "shape + electrostatics" option was used for the correlation type and the molecular mechanics (MM) rigid-body minimization option was selected. During docking runs, we only docked bound receptor and ligand conformations and took advantage of the HEX option that allows the user to indicate receptor and ligand interface residues (usually a hydrogen bonding interaction). This is in keeping with the rigid-body assumption of our free energy methodology and had the additional advantage of producing at least some native-like solutions for almost all of the protein partners studied.

In an effort to remove any remaining clashes all decoys were subjected to a second round of rigid body energy minimization using the charmm19 force field and a termination criterion of 0.5 kcal (using only the vdw non-bonded term), as implemented in the molecular modeling package TINKER (http://dasher.wustl.edu/tinker/)[3].

Table 1
ΔG_{bind} values for native ensembles and experiment

PDB[a]	$\Delta G_{bind,native,avg}$[b]	$\Delta G_{bind,exp}$[c]
1acb	−14.8	−13.1
1brs	−16.8	−17.3
1cho	−13.4	−14.4
1cse	−14.6	−13.1
1ppf	−14.0	−13.5
1tec	−13.0	−14.0
1tpa	−16.6	−17.8
1yqv	−13.6	−14.5
2ptc	−17.2	−18.1
2sic	−14.4	−12.7
2tgp	−16.8	−17.8
3tpi	−16.7	−17.3
4tpi	−16.0	−17.7
1ycs	−11.3	−10.3
2sni	−13.6	−15.8
2sec	−14.4	−14.0
3sgb	−13.3	−12.7
4sgb	−12.7	−11.7
1stf	−15.2	−13.5
1vfb	−12.0	−11.5
2pcc	−7.1	−7.0

[a] PDB codes for all 21 proteins tested in the present study; all 21 proteins were drawn from our original training set.
[b] Absolute binding free energies calculated as an average from the best (lowest free energy) and worst (highest native free energy) native state complexes, using Eq. (3).
[c] Experimental binding free energies. The calculated error (rmsd) is ≈ 1.1 kcal and the correlation with experiment (R) is ≈0.92.

Table 2
Summary of key free energy mapping results

PDB[a]	Interaction type[b]	Native rank[c]	$\Delta\Delta G_{native,worst,bind}$[d] (kcal)	$\Delta\Delta G_{native,best,bind}$[e] (kcal)	$(\Delta\Delta G_{native,best,bind} - \Delta\Delta G_{native,worst,bind})$[f] (kcal)
1acb	Protease–inhibitor	1	−4.70	−5.50	−0.80
1brs	Enzyme–inhibitor	1	−4.60	−6.70	−2.10
1cho	Protease–inhibitor	1	−2.30	−2.70	−0.40
1cse	Protease–inhibitor	1	−4.50	−5.50	−1.00
1ppf	Protease–inhibitor	1	−3.10	−3.10	0.00
1tec	Protease–inhibitor	1	−4.00	−6.40	−2.40
1tpa	Protease–inhibitor	1	−4.10	−5.50	−1.40
1yqv	Antibody–antigen	1	−7.30	−7.30	0.00
2ptc	Protease–inhibitor	1	−6.40	−6.80	−0.40
2sic	Protease–inhibitor	1	−2.60	−5.00	−2.40
2tgp	Protease–inhibitor	1	−3.90	−5.40	−1.50
3tpi	Protease–inhibitor	1	−3.40	−5.10	−1.70
4tpi	Protease–inhibitor	1	−3.80	−5.80	−2.00
1ycs	Protein–protein	1	−1.50	−1.50	0.00
2sni	Protease–inhibitor	1	−0.70	−4.30	−3.60
2sec	Protease–inhibitor	1	−0.20	−1.20	−1.00
3sgb	Protease–inhibitor	1	−0.06	−1.50	−1.44
4sgb	Protease–inhibitor	1	−1.50	−1.80	−0.30
1stf	Protease–inhibitor	2	2.80	1.10	−1.70
1vfb	Antibody–antigen	6	0.60	0.60	0.00
2pcc	Cytochrome–cytochrome	>10	4.45	2.95	−1.50

Using Eq. (3), binding free energies (ΔG_{bind}) were calculated and plotted versus rmsd (Å) for 100 decoys and their corresponding X-ray structures, for 21 protein–protein complexes.

[a] Column 1 provides the PDB codes for all 21 proteins tested in the present study; all 21 proteins were drawn from our original training set.

[b] Information regarding the basic biological interaction type for all 21 complexes.

[c] The best rank obtained for a native interaction (out of 100 decoy complexes). A complex is defined as native if its complex geometry has a rmsd < 1.0 from the all-atom crystallographic coordinates.

[d] The free energy gap between the worst native state complex (highest ΔG_{bind} among all native state complexes) and the best non-native complex (lowest ΔG_{bind} among all non-native complexes).

[e] The free energy gap between the best native state complex (lowest ΔG_{bind} among all native state complexes) and the best non-native complex (lowest ΔG_{bind} among all non-native complexes). Negative values for both columns indicate that all native solutions are lower in free energy than non-native ones or that the native state ensemble occupies a global free energy minimum. Larger values imply a deeper global minimum.

[f] Free energy difference between values listed in columns 5 and 6.

Following the second round of minimization, all-atom root-mean-squared deviations (rmsd's) from the relevant crystallographic structures were calculated for each decoy set. A decoy was counted as native-like if its rmsd was < 1.0 from the crystal structure, which is approximately equal to the experimental error expected for an X-ray structure; otherwise it was counted as non-native. Binding free energies (ΔG_{bind}) were then calculated for each decoy set (and the corresponding crystal structure) using Eq. (3). The results were then plotted for all 21 protein complexes as ΔG_{bind} versus rmsd free energy scatter plots. Only solutions with negative ΔG_{bind} values were included in the final plots.

2.3. Calculating ΔG_{bind}

For all structures, ΔG_{bind} was calculated using Eq. (3) as implemented in the Affinity software package, available through CMD Bioscience (http://www.cmdbioscience.com/). The only structural input to the algorithm is a protein–protein complex coordinate file in PDB format. Additionally, users must supply a key-word control file. The control file default settings were used and are described in Materials and methods (see above). Also by default, ΔG_{bind} is reported in kcal and this unit is used in the present paper. Throughout the paper several ΔG_{bind} values are referenced: (1) $\Delta G_{bind,exp}$ refers to experimentally determined binding free energies; (2) $\Delta G_{bind,native}$ refers to the binding free energy calculated for individual native complexes and $\Delta G_{bind,native,avg}$ refers to the average binding free energy for an ensemble of native complexes; (3) $\Delta G_{bind,native,worst}$ and $\Delta G_{bind,native,best}$ refer to the binding affinities calculated for native structures with the highest (worst) and lowest (best) free energies, respectively; (4) $\Delta G_{bind,non\text{-}native}$ refers to binding free energies calculated for non-native protein–protein interactions and $\Delta G_{bind,non\text{-}native,best}$ refers to binding free energies for the best (lowest) non-native protein–protein interactions; (5) finally, $\Delta\Delta G_{bind,native,worst} = \Delta G_{bind,native,worst} - \Delta G_{bind,non\text{-}native,best}$ and $\Delta\Delta G_{bind,native,best} = \Delta G_{bind,native,best} - \Delta G_{bind,non\text{-}native,best}$ refer to the differences between the best non-native and the best and worst native state binding free energies; negative values imply that the native binding interactions exists at free energy minima with respect to non-native binding interactions.

3. Results

The primary aim of the present study was to evaluate the accuracy of Eq. (3) for estimating native and non-native binding affinities. Table 1 summarizes the predicted ensemble average native state binding free energies ($\Delta G_{bind,native,avg}$) and experimental binding free energies ($\Delta G_{bind,exp}$) for all 21 test complexes. By comparing $\Delta G_{bind,native,avg}$ and $\Delta G_{bind,exp}$ we can get a good idea as to how physically accurate, in an absolute sense, the predictions of Eq. (3) are.

Table 2 provides additional information for all 21 test complexes, including complex type and calculated values for $\Delta\Delta G_{bind,native,worst}$ and $\Delta\Delta G_{bind,native,best}$. These latter two values are important for quantitatively determining the presence and depth of native state global free energy minima and help us determine if Eq. (3) can be used to provide physically plausible but relative free energy ($\Delta G_{bind,non\text{-}native}$) rankings for non-native binding modes. Ideally, our calculated $\Delta G_{bind,non\text{-}native}$ values would be compared to experiment or more rigorous calculations but such data is lacking.

In Fig. 1, and for all 21 complexes, the predicted binding free energy (ΔG_{bind}) is plotted against the rmsd (Å) for the various decoy and crystallographic structures. The graphs provide a convenient tool for quickly and visually assessing the performance of Eq. (3). In a best case scenario, all 21 graphs would be characterized by native ensembles (< 1.0 Å rmsd) at deep and global free energy minima, with $\Delta G_{bind,native,avg} \approx \Delta G_{bind,exp}$ and $\Delta\Delta G_{bind,native,worst} \ll \Delta$. $\Delta G_{bind,non\text{-}native,best}$.

4. Discussion

There is a need for an accurate and fast free energy function to solve numerous applied protein design and engineering problems. Previously, we developed and validated a computationally inexpensive empirical free energy equation for predicting experimental binding affinities from the crystallographic coordinates of protein–protein and protein–peptide complexes. Here, we extend that work to include the accurate and fast prediction of native-like and non-native protein–protein interactions. The major result of this paper, then, is Eq. (3). The new function has been tested for its ability to calculate physically plausible free energy surfaces for 21 protein–protein complexes and has been found to perform well. Hence, Eq. (3) represents an improved and more generally valid version of our original free energy function that might also prove suitable as a scoring function for various protein modeling and design applications.

4.1. The free energy expression: evaluating the new coefficients

Eq. (2) expresses our most recent multi-term empirical equation for estimating ΔG_{bind} in terms of regression-weighted physical descriptors. The equation is in fundamental agreement with Eq. (1), our previously described empirical equation, with the only difference being that the regression coefficients were recently re-optimized. The re-parameterization was performed using the original training set, minus a single crystallographic complex (PDB code: 1dhk). Of course, following re-parameterization Eq. (2) was subjected to the

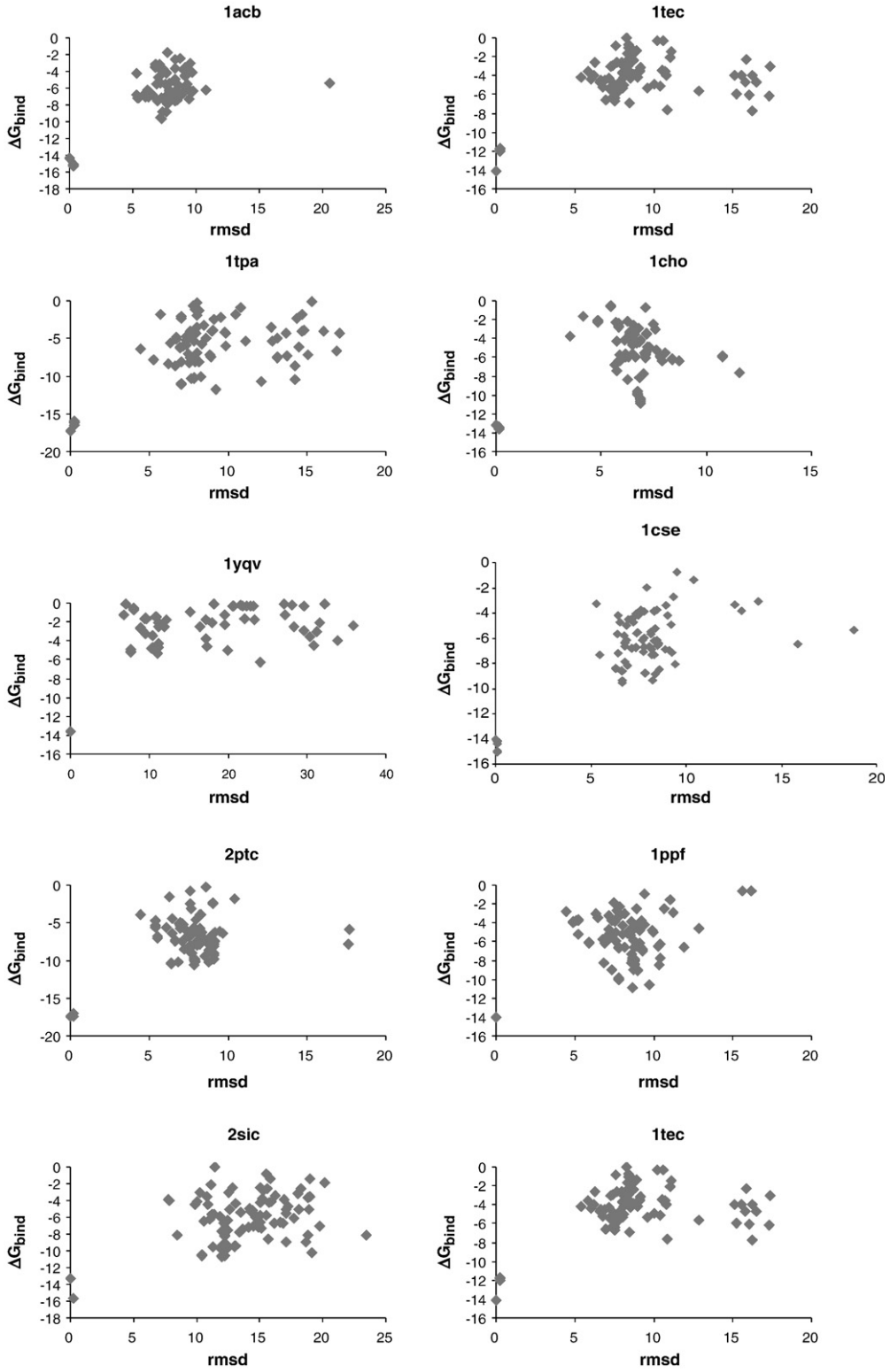


Fig. 1. Plots of protein–protein binding free energy ΔG_{bind} versus root-mean-squared deviation (rmsd) in Angstroms (Å). All ΔG_{bind} values were calculated using Eq. (3). For all 21 complexes, ΔG_{bind} was calculated for the crystallographic complex and 100 decoys. Only negative ΔG_{bind} values are plotted. All decoy structures were generated using the HEX protein–protein docking server. For a given protein complex, all-atom decoy rmsd values were calculated with respect to the relevant crystallographic structure.

same tests as our previous empirical expression (Eq. (1)), giving slightly better results on all tests (see Supplementary data). Thus, any conclusions regarding the form and coefficients of Eq. (1) apply with equal or greater force to Eq. (2). Specifically, our analysis suggests that the form, magnitudes and signs characteristic of Eq. (2) make statistical and physical sense, that Eq. (2) is in basic qualitative and quantitative agreement with theory and experiment, and that Eq. (2) can be used to blindly predict experimental binding free energies

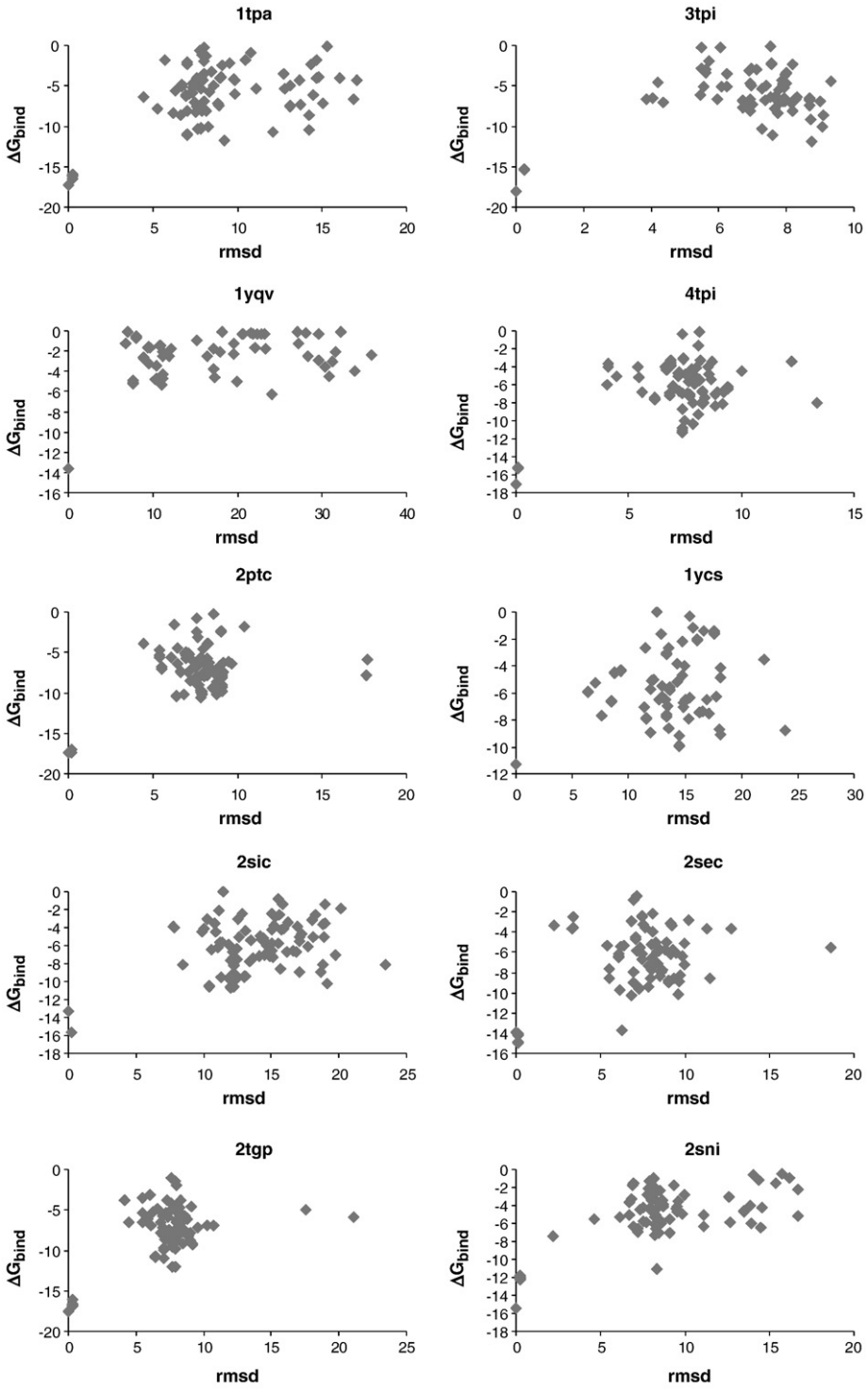


Fig. 1 (*continued*).

from crystallographic complexes to within ≈1.0 kcal [1]. Henceforth, our discussion will focus on Eq. (2).

4.2. The free energy expression: evaluating the physical reasonableness of the new SASA hydrogen-bonding and salt-bridge criteria

Eq. (1) and its re-parameterized cousin (Eq. (2)), while shown to be valid for a large and diverse number of native interactions, are not valid for non-native interactions. This is because non-native interfaces often include various energetically unfavorable interface interactions and such interactions were not considered in our original analysis. This is an important consideration, for applied protein design and modeling problems often involve the modeling and scoring of non-native interactions. Indeed, preliminary testing indicated that while Eq. (1) and (2) could be used to consistently rank native-like decoys in the top 10, the free energy surface was plagued by closely spaced local minima (data not shown).

A mathematically and algorithmically simple method for extending Eq. (2) to include non-native interactions is given by Eq. (3). Here our emphasis was on the basic physics of the problem. Thus,

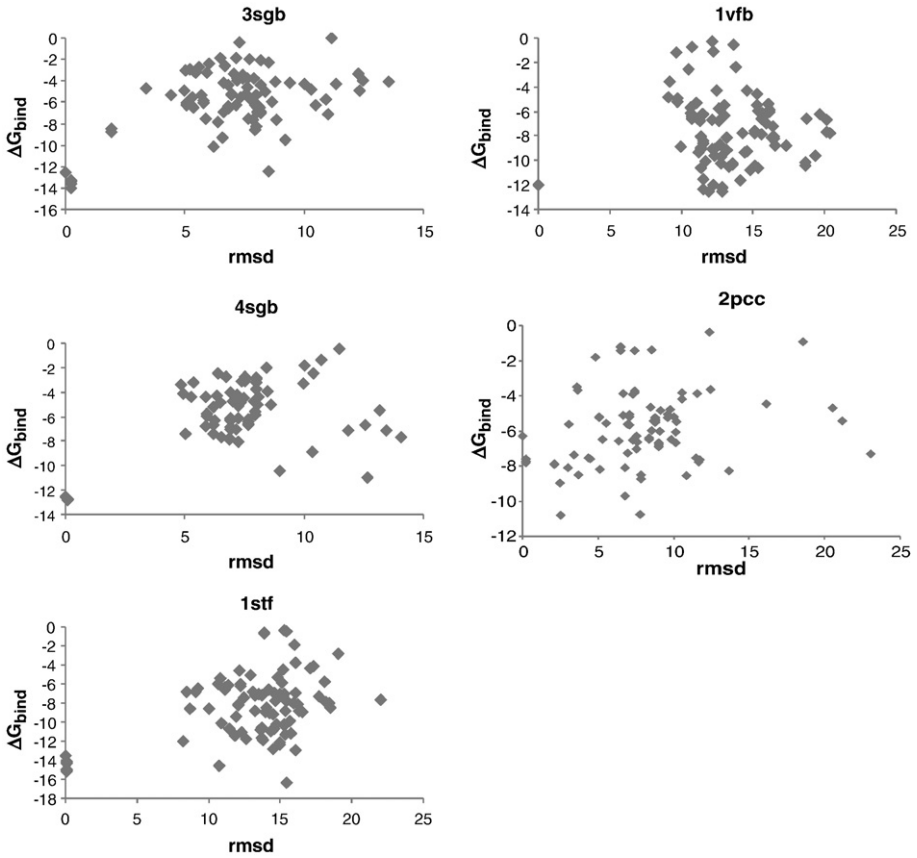


Fig. 1 (*continued*).

we included a physically plausible clash (+2.5X_{clash}) term and, given the well-known screening effect solvent exposure can have on electrostatic interactions, a SASA condition ($SASA_{sb}$=25.9 Å) for neutralizing salt bridges. It is important to note that interface clashes were absent from the complexes that characterized our original training and test sets and the $SASA_{sb}$ value is a maximum value that was derived from an analysis of our original training and test sets. Thus, according to our new model for calculating ΔG_{bind}, non-hydrogen bonded atoms involved in steric clashes entail a free energy penalty of 2.5 kcal and highly exposed salt-bridges fail to contribute to binding [4].

The available experimental data and theoretical calculations indicate that deeply buried hydrogen bonds are characteristic of native protein–protein interfaces and that the energetic contributions of hydrogen bonds depend upon the degree of solvent accessibility [5,6]. In a manner analogous to that of salt bridges, an interface characterized by buried hydrogen bonds is expected to exhibit stronger binding than one characterized by solvent exposed hydrogen bonds (everything else being equal). Similar reasoning suggests that charge–polar hydrogen bonds can tolerate more solvent exposure than polar–polar hydrogen bonds before an energetic penalty is incurred. Indeed, an analysis of our training and test sets produced maximum average SASA values of 8.0 $Å^2$ and 17.0 $Å^2$ for all polar–polar and charge–polar hydrogen bonds, respectively. This is the basic logic that supports the inclusion of $\Delta G_{sasa,polar-polar,hb}$=3.0 kcal and $\Delta G_{sasa,charge-polar,hb}$=3.0 kcal as the last two terms in Eq. (3). As will be argued below, the inclusion of SASA hydrogen bonding criteria proved to be the single most important step in accurately estimating non-native binding affinities.

A few final points are worth noting. The criteria chosen for the clash, salt bridge and hydrogen bonding penalty contributions were not adjusted or tweaked so as to improve our ability to calculate and map native and non-native interactions. Rather, each penalty magnitude, while somewhat arbitrary, was selected *a priori* to simply ensure that a modest and physically plausible unfavorable or neutral free energy contribution was added to ΔG_{bind} for interfaces that involved clashes, highly exposed salt bridges, and exposed hydrogen bonds. Future work will focus on optimizing the various penalty weights. Finally, because the SASA threshold values represent maximum values obtained from our original training and test sets and because the only differences between Eq. (2) and (3) has to do with the new penalty terms, the predictive performance of Eq. (3) for clash free native interactions (experimental binding free energy prediction) is exactly the same as for Eq. (2).

4.3. Testing the new function: binding free energy surfaces calculated using Eq. (3)

An equation for estimating ΔG_{bind} can be tested for its ability to (1) accurately predict crystallographic or experimental binding free energies for a large and diverse set of protein interactions, (2) for its basic statistical, physical and theoretical validity, and (3) for its ability to accurately rank binding affinities for a large and diverse number of native-like and non-native protein–protein interaction ensembles. If an equation performs well in all three tests the equation can be said to be more generally accurate than an equation that does not perform well on all three tests, at least within the limits established by any simplifying assumptions and approximations.

We have argued that Eq. (2) performs well in the first two tests but not in the third (data not shown). We have also argued that the various penalty contributions introduced to transform Eq. (2) into Eq. (3) make basic physical and theoretical sense and were introduced in such a way so as to ensure that Eq. (3) is as good as Eq. (2) when it

comes to predicting experimental binding free energies. What remains to be seen, then, is the performance of Eq. (3) in estimating native-like and non-native binding affinities.

In what follows we evaluate Eq. (3) for its ability to accurately map, in absolute and relative terms, ΔG_{bind} as a function of distance (rmsd) from crystallographic binding geometries. Essentially, we are testing Eq. (3) to see if it can be used to consistently place native-like and crystallographic binding modes in global and physically plausible free energy minima relative to non-native binding modes. Given the limits of our experimental knowledge and put quantitatively, we are testing Eq. (3) for its ability to estimate that (1) $\Delta G_{bind,native,avg} \approx \Delta G_{bind,exp}$ and that (2) $\Delta G_{bind,native,worst} < \Delta G_{bind,non\text{-}native,best}$, where it's understood that $\Delta G_{bind,native,worst} \ll \Delta G_{bind,non\text{-}native,best}$ is probably more realistic. Success in this round of preliminary testing will help to further establish the basic validity of our methodology and justify continued research.

It is reasonable to assume that Eq. (3) should predict similar binding affinities for crystallographic and native-like (≈1.0 Å rmsd) interactions and the predicted binding affinities should be close to experiment. For all 21 protein complexes tested, there is good agreement between $\Delta G_{bind,native,avg}$ and $\Delta G_{bind,exp}$. This is made evident by the tight and well-defined native ensembles that tend to collect around $\Delta G_{bind,exp}$ (see Fig. 1 and Tables 1 and 2). In all cases, the average binding free energy calculated for each native ensemble ($\Delta G_{bind,native,avg}$) is in excellent agreement with the corresponding experimentally determined values ($\Delta G_{bind,exp}$), with an overall rmsd of only ≈1.0 kcal and R=0.92. Thus, for all 21 complexes $\Delta G_{bind,native,avg} \approx \Delta G_{bind,exp}$, with an error of ≈1.0 kcal.

Ideally, we would like to compare the predictions made using Eq. (3) for non-native (>1.0 Å rmsd) interactions with experimental data or binding affinity calculations derived from fundamental theory. Unfortunately, such data is lacking. Thus, we must content ourselves with evaluating Eq. (3) according to its ability to provide relative free energy rankings that are physically reasonable or that place all non-native interactions at higher points on the free energy surface than native ones.

For 18 of the 21 proteins tested ΔG_{bind} calculations made using Eq. (3) imply a global minimum for the native state ensembles, as indicated by the first 18 graphs from Fig. 1 and their corresponding negative values for $\Delta\Delta G_{native,best,bind}$ and $\Delta\Delta G_{native,worst,bind}$ (see Table 2). For 18 of the cases, then, Eq. (3) placed all native state structures at a lower free energy than all non-native structures. Put differently, in almost 90% of the cases tested Eq. (3) was used to make predictions that imply $\Delta G_{bind,native,worst} < \Delta G_{bind,non\text{-}native,best}$.

Importantly, 13 out of 18 native ensembles clearly occupy deep and physically plausible global free energy minima (≈3–7 kcal, average ≈ 4.2 kcal) while the remaining 5 protein native ensembles occupy less pronounced global minima (≈0.5–1.5 kcal, average=0.8 kcal). Fully 15 of the 18 native ensembles would qualify as deep if we were to ignore two isolated and anomalous non-native data points for 2sec and 2sni. The results are even better if we focus on the $\Delta G_{bind,native,best} < \Delta G_{bind,non\text{-}native,best}$ data. In that case, the 13 and 5 member native ensembles occupy even deeper global free energy minima (average ≈5.4 kcal and average=2.1 kcal, respectively). These results are all the more encouraging given that the error associated with Eq. (3) for estimating experimental and native-like binding free energies is roughly 1.0 kcal. Hence, a reasonable interpretation of the data is that Eq. (3) worked extremely well for 13 proteins and acceptably for 18 proteins and that Eq. (3) is more generally valid than our earlier expressions for ΔG_{bind}.

For two proteins (1stf and 1vfb) Eq. (3) placed the native state structures in local minima. For these two proteins $\Delta G_{bind,native,worst} > \Delta G_{bind,non\text{-}native,best}$ and Eq. (3), strictly speaking, failed in its predictions. Despite this, the methodology can still be interpreted as having achieved qualified successes in that the crystallographic and native-like structures are ranked reasonably well and are very close to their respective global minima. The only complete failure is for 2pcc, which exhibits a very bumpy and physically unrealistic binding free energy surface. Presumably, the failure to account for 2pcc binding and the partial failures to account for 1stf and 1vfb binding have something to do with interaction terms that are incorrectly parameterized, ignored or assumed away in Eq. (3). Future work will focus on the explicit consideration of additional interactions and parameter re-optimization to try and improve the non-native predictive performance of Eq. (3).

4.4. Comparing Eq. (2) and Eq. (3): hydrogen bond exposure as the key to accurately ranking native and non-native interactions

In preliminary testing, Eq. (2) failed to produce global minima populated exclusively by native state complexes (data not shown). This means success at predicting experimental binding free energies from the coordinates of crystallographic complexes does not imply or even suggest success at predicting non-native binding affinities. When Eq. (2) is augmented with a clash term ($+2.5X_{clash}$) and SASA weighted salt bridge term the observed free energy surfaces, while somewhat improved, remain unacceptably rugged (data not shown). Of course, this makes sense given that our protocol for constructing decoys all but guaranteed protein interfaces with good surface complimentarity. It is only after Eq. (2) is fully transformed into Eq. (3) through the addition of SASA weighted hydrogen bonding penalty terms ($\Delta G_{sasa,polar-polar,hb}$ and $\Delta G_{sasa,charge-polar,hb}$) that the encouraging results summarized in Fig. 1 and Tables 1 and 2 are obtained. Thus, solvent inaccessible intermolecular hydrogen bonds appear to be one of the key interactions that distinguishes native from non-native protein–protein complexes and this inference is in good agreement with at least one previous docking study [7].

4.5. Comparison with previous research

Many methods have been developed to score and rank native and non-native protein–protein interactions, mostly within the context of the protein–protein docking problem. In fact, these scoring schemes often do quite a good job at placing native-like binding modes in global or near-global calculated minima. Indeed, there might even be a reduced version of Eq. (3) that can be used to consistently place native interactions in some kind of global minima. These calculated minima and the other points on the calculated protein–protein interaction surface, however, do not represent binding free energies, although they certainly represent components of the binding free energy [8–15]. Thus, direct comparisons with the results presented in the present paper are difficult to make. Suffice it to say, the results reported here – when viewed purely in terms of the relative rankings of crystallographic, native and non-native binding modes – appear comparable to what has been reported in previous studies, with the added advantage that the values calculated by Eq. (3) can be reasonably interpreted as binding free energy estimates.

About a decade ago, Vajda et al. did, however, publish a series of important papers exploring the use of free energy functions for estimating experimental binding affinities and for ranking native and non-native binding geometries [16–18]. These studies do provide a basis for a more direct comparison with the results generated using Eq. (3) and summarized in Fig. 1. and Tables 1 and 2. Very briefly, it would seem that Eq. (3) and the results reported here can be interpreted as reproducing, building upon and extending that foundational work. More recently, only Ma et al. seem to have continued work on developing a free energy function for use on the docking problem. Using their empirical equation Ma and co-workers scored and ranked the docking output for 10 protein complexes and while they obtained encouraging results, their methodology failed to rank native solutions first [19].

4.6. Practical implications for the protein–protein docking problem

It is important to point out that while the present study was, in part, motivated by the protein–protein docking problem its primary purpose was to further test our free energy methodology and not to solve or even directly address the docking problem. Of course, given that the testing we performed involved the estimation of native and non-native binding free energies the present study has obvious implications for the docking problem, particularly the predictive docking problem or the problem of predicting *if* and *how* two proteins will react to form a complex. In light of our methodology and results, at least five things can be said about Eq. (3) and its possible use as a scoring function in protein–protein docking algorithms.

First and most importantly, unlike most other protein–protein scoring strategies Eq. (3) provides an estimate of the binding free energy, the quantity that actually controls binding reactions. Second, Eq. (3) is relatively simple and fast and this makes it an attractive scoring function. Third, our data suggests that if a docking algorithm can produce binding modes within ≈1.0 Å rmsd of the actual complex, Eq. (3) would stand a good chance of predicting them. However, the high ΔG_{bind} values for the near-native binding modes from 2sec, 2sni and 3sgb (see Fig. 1) suggest that the exclusive use of Eq. (2) could be problematic, even for binding modes within ≈2.0–3.0 Å of the native complex. Of course, this does not imply that Eq. (2) is inaccurate; it could be the case that nature is quite sensitive to deviations from structural ideality, as is suggested by our results. Fourth, it is not entirely clear that our results can be generalized beyond protease–inhibitor complex interactions. Fifth, Eq. (3) is, strictly speaking, limited to the problem of rigid-body docking and should be modified or otherwise supplemented for use in flexible docking applications.

4.7. Limitations of the present study

An interesting question that goes unanswered in the present study has to do with the absolute accuracies of the binding free energies calculated for non-native interactions. Can it be argued that the $\Delta G_{bind,non\text{-}native}$ predictions made using Eq. (3) are accurate to within ≈1.0 kcal in analogy with the $\Delta G_{bind,native}$ values? Vajda et al. also encountered this question, ultimately concluding that the balance of the evidence indicated the accuracy of their predicted binding affinity values to within ≈5% [18]. We too are tempted in this direction, especially for the 13 complexes that exhibit deep global minima, but leave it to future work to try and settle the matter more decisively.

Another possible limitation of the present study is the apparently small and homogenous nature of our 21 member test set. The exclusive emphasis on protein–protein interactions, the use of rigorous protein structural quality selection criteria, the need for high quality experimental binding data, and a desire to minimize confounding influences and maintain continuity with our past work all combined to reduce the data and structural space available for analysis. With this in mind, however, it is important to note that some 2100 diverse protein–protein interactions were ultimately considered and that (to the best of our knowledge) the present study represents the largest of its kind completed to date. Moreover, there was no fitting involved in the present study so all 2100 predictions qualify as truly blind. As such, we think it is clear that "chance" can be ruled out as an explanation for our ≈60–90% success rate, although care should be exercised in generalizing our results. In the future, we would like to analyze a larger and more heterogeneous data set.

4.8. Conclusions and future research

Our results in this somewhat preliminary round of testing suggest that Eq. (3) can be used to predict binding free energies (ΔG_{bind}) for native protein–protein binding interactions to within approximately 1.0 kcal. In approximately 90% of the cases studied, Eq. (3) was used to place native binding modes in global free energy minima and in ≈60% of the cases Eq. (3) placed native-like and crystallographic binding modes in physically realistic, deep global minima. Thus, Eq. (3) represents a more generally valid and robust method for binding free energy prediction than its predecessor functions (Eqs. (1), (2)). To the best of our knowledge Eq. (3) is the only function currently available that can be used to quickly and accurately predict experimental binding free energies and accurately rank native and non-native binding modes according to the free energy of binding. Finally, a comparison between Eq. (2), and Eq. (3), suggests buried intermolecular hydrogen bonds are vital to complex formation and to differentiating native from non-native complexes.

When the relative simplicity and speed of Eq. (3) is taken into account, Eq. (3) emerges as a viable scoring function for (rigid-body) protein–protein docking algorithms. In particular, our results suggest that Eq. (3) can be used to address the predictive docking problem, the problem of predicting if and how two proteins will interact. Future work will focus on the further development, optimization and evaluation of Eq. (3) and on applying Eq. (3) to protein modeling and design problems, such as the protein–protein docking problem.

Acknowledgements

I would like to thank Lisa Yan for proof reading the manuscript and for making helpful suggestions. I would also like to thank Kevin Hannigan and Kruti Naik for generating the docking decoys and for producing many of the figures used in the paper.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bpc.2008.10.007.

References

[1] J. Audie, S. Scarlata, A novel empirical free energy function that explains and predicts protein–protein binding affinities, Biophys. Chem. 129 (2007) 198.
[2] D.W. Ritchie, Evaluation of protein docking predictions using Hex 3.1 in CAPRI rounds 1 and 2, Proteins 52 (2003) 98.
[3] P. Ren, J.W. Ponder, Consistent treatment of inter- and intramolecular polarization in molecular mechanics calculations, J. Comput. Chem. 23 (2002) 1497.
[4] F. Dong, H.X. Zhou, Electrostatic contributions to T4 lysozyme stability: solvent-exposed charges versus semi-buried salt bridges, Biophys. J. 83 (2002) 1341.
[5] A.V. Efimov, E.V. Brazhnikov, Relationship between intramolecular hydrogen bonding and solvent accessibility of side-chain donors and acceptors in proteins, FEBS Lett. 554 (2003) 389.
[6] D. Xu, C.J. Tsai, R. Nussinov, Hydrogen bonds and salt bridges across protein–protein interfaces, Protein Eng. 10 (1997) 999.
[7] R. Norel, F. Sheinerman, D. Petrey, B. Honig, Electrostatic contributions to protein–protein interactions: fast energetic filters for docking and their physical basis, Protein Sci. 10 (2001) 2147.
[8] C.J. Camacho, H. Ma, P.C. Champ, Scoring a diverse set of high-quality docked conformations: a metascore based on electrostatic and desolvation interactions, Proteins 63 (2006) 868.
[9] C.J. Camacho, C. Zhang, FastContact: rapid estimate of contact and binding free energies, Bioinformatics 21 (2005) 2534.
[10] J.J. Gray, S. Moughon, C. Wang, O. Schueler-Furman, B. Kuhlman, C.A. Rohl, D. Baker, Protein–protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations, J. Mol. Biol. 331 (2003) 281.
[11] I. Halperin, B. Ma, H. Wolfson, R. Nussinov, Principles of docking: an overview of search algorithms and a guide to scoring functions, Proteins 47 (2002) 409.
[12] B. Pierce, Z. Weng, ZRANK: reranking protein docking predictions with an optimized energy function, Proteins 67 (2007) 1078.
[13] D.W. Ritchie, Recent progress and future directions in protein–protein docking, Curr. Protein Pept. Sci. 9 (2008) 1.
[14] J.G. Mandell, V.A. Roberts, M.E. Pique, V. Kotlovyi, J.C. Mitchell, E. Nelson, I. Tsigelny, L.F. Ten Eyck, Protein docking using continuum electrostatics and geometric fit, Protein Eng. 14 (2001) 105.
[15] C.J. Camacho, D.W. Gatchell, S.R. Kimura, S. Vajda, Scoring docked conformations generated by rigid-body protein–protein docking, Proteins 40 (2000) 525.
[16] S. Vajda, Z. Weng, R. Rosenfeld, C. DeLisi, Effect of conformational flexibility and solvation on receptor-ligand binding free energies, Biochemistry 33 (1994) 13977.
[17] Z. Weng, C. Delisi, S. Vajda, Empirical free energy calculation: comparison to calorimetric data, Protein Sci 6 (1997) 1976.
[18] Z. Weng, S. Vajda, C. Delisi, Prediction of protein complexes using empirical free energy functions, Protein Sci. 5 (1996) 614.
[19] X.H. Ma, C.X. Wang, C.H. Li, W.Z. Chen, A fast empirical approach to binding free energy calculations based on protein interface information, Protein Eng. 15 (2002) 677.