PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Triathlon for energy functions: Who is the winner for design of protein–protein interactions?

Oz Sharabi, Ayelet Dekel, and Julia M. Shifman*

Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

ABSTRACT

Computational prediction of stabilizing mutations into monomeric proteins has become an almost ordinary task. Yet, computational stabilization of protein–protein complexes remains a challenge. Design of protein–protein interactions (PPIs) is impeded by the absence of an energy function that could reliably reproduce all favorable interactions between the binding partners. In this work, we present three energy functions: one function that was trained on monomeric proteins, while the other two were optimized by different techniques to predict side-chain conformations in a dataset of PPIs. The performances of these energy functions are evaluated in three different tasks related to design of PPIs: predicting side-chain conformations in PPIs, recovering native binding-interface sequences, and predicting changes in free energy of binding due to mutations. Our findings show that both functions optimized on side-chain repacking in PPIs are more suitable for PPI design compared to the function trained on monomeric proteins. Yet, no function performs best at all three tasks. Comparison of the three energy functions and their performances revealed that (1) burial of polar atoms should not be penalized significantly in PPI design as in single-protein design and (2) contribution of electrostatic interactions should be increased several-fold when switching from single-protein to PPI design. In addition, the use of a softer van der Waals potential is beneficial in cases when backbone flexibility is important. All things considered, we define an energy function that captures most of the nuances of the binding energetics and hence, should be used in future for design of PPIs.

Proteins 2011; 00:000–000.

Key words: energy functions; computational protein design; optimization; protein–protein interactions.

INTRODUCTION

Virtually all biological processes in the cell are controlled through networks of protein–protein interactions (PPIs). Breaking down of the signaling networks usually results in disease. Hence, finding strategies for rational manipulation of PPIs presents a potential way for disease treatment. Computational protein design could become an invaluable approach for compensating for the disease-associated mutations or for designing inhibitors of aberrant PPIs. In recent years, great progress has been made in this direction.[1] Nevertheless, it remains difficult to accurately predict changes in binding free energy upon introduction of mutations. Up to date, successful predictions for point mutations met variable success, while redesign of the whole binding interfaces frequently produced reduction in affinity.[2–6] These results corroborate that our understanding of the molecular forces governing protein binding remains suboptimal despite the extensive structural analysis of PPIs[7–10] and alanine scanning experiments of binding interfaces.[11–17]

Why is it so difficult to predict the effect of various mutations on binding affinity? One reason is that in most cases, energy functions utilized for such predictions were parametrized on mutational data for monomeric proteins rather than for PPIs. Unlike monomeric proteins that place most hydrophobic residues in the core and most hydrophilic – on the surface, binding interfaces contain both types of amino acids that are buried in the bound state and exposed in the unbound state. Buried salt bridges and hydrogen bonds are very important for molecular recognition,[18] however their impact is hard to calculate with the energy functions optimized for design of monomeric proteins. Additional reasons for possible inaccuracies in calculating the effect of mutations on binding affinity lies in backbone conformational changes that frequently accompany introduced mutations as well as in the presence of water-mediated contacts between the two proteins. Modeling of

Additional Supporting Information may be found in the online version of this article.
Grant sponsor: Israel Science Foundation; Grant number: 1372/10; Grant sponsor: DFG; Grant number: El 423/2-1
*Correspondence to: Julia M. Shifman, Department of Biological Chemistry, The Alexander Silberman Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem 91904, Israel. E-mail: jshifman@cc.huji.ac.il.
Received 5 September 2010; Revised 19 December 2010; Accepted 22 December 2010
Published online 4 January 2011 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.22977

such effects is computationally expensive and could sometimes introduce more error than benefit.[19]

It is widely acceptable that the majority of mutations at protein–protein (PP) interfaces decrease affinity. Yet multiple directed evolution experiments suggest that affinity of practically any PPI could be further improved through mutations, sometimes by several orders of magnitude.[20] As affinity maturation experiments could be time-consuming, it would be of great interest to achieve the same results through computational predictions. Recently, several groups proposed computational approaches for improving affinity of PPIs. One strategy focuses on optimizing solely electrostatic interactions between the two binding partners.[21–27] This strategy proved to be extremely useful in some cases; however it is limited to PPIs whose electrostatics is not already optimized and thus is not applicable to many high-affinity PPIs. Kuhlman and coworkers proposed a different strategy, where enhancement in affinity is achieved by burying more hydrophobic area in the binding interface.[19] High rate of success was demonstrated for this method; however, it cannot be applied to redesign of entire binding interfaces as increase in hydrophobic surface of monomeric proteins would eventually lead to their aggregation. Bockmann and coworkers developed an approach where change in affinity is calculated over a conformational ensemble of a PP complex using an energy function derived from a large experimental dataset of changes in free energy of binding ($\Delta\Delta G_{\text{binding}}$).[28] The advantage of this method over the previous approaches is its generality and incorporation of the backbone flexibility for modeling of PPIs. The disadvantage is that it requires considerable computational time, making design of multiple simultaneous mutations impractical. We set our goal to develop a generalized and fast approach for designing affinity-enhancing mutations at PP interfaces in the framework of the protein design program ORBIT.[29] In particular, we would like to define the best energy function for design of PPIs.

There are three benchmarks that are relevant for PP interface design and could be used to optimize and access the performance of various energy functions. The first benchmark is predicting side-chain conformations of the binding-interface residues as observed in crystal structures of PP complexes. This is an easier task compared to protein design because only one amino acid is considered at each position during such prediction. The advantage of this benchmark is that it represents the true energy minimum, that is, the observed side-chain conformations are guaranteed to exhibit the lowest energy compared to those of all other possible conformations. The second benchmark is recovering native sequences within PP interfaces. Native binding-interface sequences in general do not represent the true energy minimum, as the affinity of various PP complexes could be improved through mutations. Nevertheless, for high-affinity PPIs, such sequences should be close to optimum, and hence should be recovered to a considerable extent by a good energy function. The third benchmark, most closely resembling protein design, is predicting experimentally measured $\Delta\Delta G_{\text{binding}}$ upon introduction of mutations in various PPIs. The disadvantage of this benchmark is that to compile a large enough dataset of $\Delta\Delta G_{\text{binding}}$, one needs to combine data from multiple experiments measured with different methods and sometimes at significantly different conditions. In addition, the large majority of such data includes mutations that decrease binding while only a small minority of data represents mutations that enhance affinity and are more relevant for protein design. In this work, we assess the performance of several energy functions for PP interface design in a triathlon that includes the three described benchmarks: predicting side-chain conformations in PP interfaces, recovering native binding-interface sequences and predicting $\Delta\Delta G_{\text{binding}}$ due to mutations.

Table I
Qualitative Comparison of Per-Term Contributions to Each Tested Energy Function

Energy function	$E_{\text{vdW-rep}}$	$E_{\text{vdW-attr}}$	E_{electr}	E_{HB}	E_{polar}	E_{nonpolar}
E_{monomer}	++	++	+	++	++	++
$E_{\text{PPI-1}}$	+[a]	+++	++[b]	++	−	++
$E_{\text{PPI-2}}$	++	++	++	++	+	++

+++, denotes very large contribution; ++, denotes large contribution; +, denotes small contribution; −, denotes no contribution.
[a]A linear function substitutes the LJ potential for this energy function in contrast to the other two energy functions.
[b]Electrostatic energies were capped at −4 kcal/mol for this energy function.

Our first candidate is the original ORBIT energy function that was trained for design of monomeric proteins and hence is referred to as E_{monomer}. This energy function uses the Lennard-Jones (LJ) potential for description of van der Waals (vdW) interactions, a distance- and angle-dependent hydrogen bond potential, a solvation term that strongly penalizes burial of polar atoms and exposure of nonpolar residues, and a weak Coulombic electrostatic term (see Methods and Table I for details). Our second candidate is a recently reported energy function obtained by an automated optimization procedure that simultaneously re-weighted all the terms in the ORBIT energy function while attempting to reproduce side-chain conformations in PP interfaces.[30] This energy function, termed $E_{\text{PPI-1}}$ (where PPI stands for protein–protein interaction) does not penalize burial of polar atoms and significantly up-weights the electrostatic term in comparison to E_{monomer}. Moreover, it replaces the repulsive LJ potential by a linear function, thus allowing for the presence of small steric clashes in PP interfaces (Table I). A modified LJ potential was used since with the original LJ potential, no meaningful combination of weights could be obtained by the automated optimization procedure. In this study, we introduce a new energy function $E_{\text{PPI-2}}$ that is identical to the original ORBIT energy function in

the physical description of all terms including the term for vdW interactions. The relative weights of each term in $E_{\text{PPI-2}}$ were optimized one at a time, allowing them to vary within the physically meaningful boundaries while reproducing side-chain conformations in PP interfaces. In this manuscript, we discuss strengths and weaknesses of each energy function when competing for the title of a winner in PPI design.

MATERIALS AND METHODS

The energy functions

ORBIT uses an empirical energy function, which linearly combines five terms:[31]

1. The 6–12 Lennard-Jones van der Waals (vdW) potential (denoted E_{vdW}), accounting for both attractive ($E_{\text{vdW}} < 0$, denoted as $E_{\text{vdW-attr}}$) and repulsive ($E_{\text{vdW}} > 0$, denoted as $E_{\text{vdW-rep}}$) interactions.
2. Distance- and angle-dependent hydrogen bond potential (E_{HB});
3. Coulombic electrostatic potential with a distance dependent dielectric constant of 40r (E_{elelctr});
4. Surface area dependent atomic solvation term, which penalizes burial of polar atoms (E_{polar});
5. Surface area dependent atomic solvation term that penalizes exposure and rewards burial of nonpolar atoms (E_{nonpolar}).

The exact equations and the parameters for all terms could be found in.[31]

We split the vdW potential into the attractive and the repulsive components and associate relative weights, w, to each of the terms in the energy function:

$$E_{\text{tot}} = w_{\text{vdW\_rep}} E_{\text{vdW\_rep}} + w_{\text{vdW\_attr}} E_{\text{vdW\_attr}} + w_{\text{HB}} E_{\text{HB}} + w_{\text{electr}} E_{\text{electr}} + w_{\text{polar}} E_{\text{polar}} + w_{\text{nonpolar}} E_{\text{nonpolar}} \quad (1)$$

The original ORBIT energy function, E_{monomer}, was optimized for design of monomeric proteins. For the purpose of this study, we define E_{monomer} by assigning equal weights to each of the energy terms with all the parameters within the terms being optimal for the protein folding problem:[31]

$$E_{\text{monomer}} = 1.0\ E_{\text{vdW\_rep}} + 1.0\ E_{\text{vdW\_attr}} + 1.0\ E_{\text{HB}} + 1.0\ E_{\text{electr}} + 1.0\ E_{\text{polar}} + 1.0\ E_{\text{nonpolar}} \quad (2)$$

The second energy function, $E_{\text{PPI-1}}$, was optimized for interface side-chain repacking in our recent work.[30] $E_{\text{PPI-1}}$ substituted the repulsive vdW term ($E_{\text{vdW}} > 0$) with a softer linear repulsive term. In addition, it capped the contribution of electrostatic interactions at ±4 kcal/mol to prevent very high interaction energies for atoms found in close proximity. The weights for each term of this energy function were simultaneously optimized using the Conditional Random Fields framework,[32,33] while attempting to reproduce side-chain conformations of binding-interface residues observed in X-ray structures of 30 PP complexes. After the optimization, $E_{\text{PPI-1}}$ became:

$$E_{\text{PPI-1}} = 11.2\ E_{\text{lin\_vdW\_rep}} + 4.4\ E_{\text{vdW\_attr}} + 1.2\ E_{\text{HB}} + 6.8\ E_{\text{electr}} + 0.0\ E_{\text{polar}} + 1.0\ E_{\text{nonpolar}} \quad (3)$$

Where $E_{\text{lin\_vdW\_rep}} = 24\ D_o \left[1 - 1.25 \frac{R}{R_0}\right]$, D_0 is the vdW well depth, R is the distance between two atoms and R_0 is the vdW radii.

The new energy function presented in this work, termed $E_{\text{PPI-2}}$, has been also optimized on the ability to predict side-chain conformations in PP interfaces. Unlike $E_{\text{PPI-1}}$, $E_{\text{PPI-2}}$ maintains all terms identical to E_{monomer} with the weights of each term optimized one at a time (see weight optimization).

The dataset for recovering side-chain conformations

99 PDB files containing structures of heterodimeric PP complexes were selected for our dataset from a larger dataset reported by Cohen *et al.*[34] (See Supporting Information in[30] for exact definition). The structures were solved to resolution of 2.6 Å or better. The data set was randomly divided into a training set (30 structures) and a test set (69 structures).

For each protein in a PP complex, residues within 4 Å from the other binding partner were defined as interface. All Prolines and some Cysteins (those that participate in disulphide bonds) were excluded from our analysis. In addition, the nonrotamering Glycines and Alanines were essentially considered part of the fixed structural template, and therefore were also excluded from the analysis. All in all, after these exclusions, our data contained 1096 interfacial residues in the training set (36.53 ± 15.03 per complex, on average) and 2688 residues in the test set (38.96 ± 18.97). The residues within 4 Å from any interfacial residue were defined as shell. These shell residues were allowed to change their side-chain conformation during packing, to permit more flexibility in the PP interface. However, predictions for the shell residues were not used for the evaluation procedure. At each interface and shell position, side chains were modeled using rotamers, each represented by a set of dihedral angles χ_1 to χ_4. The rotamers were defined based on the backbone dependent rotamer library of Dunbrack and Karplus[35] with subrotamers added at ± one standard deviation around χ_1 mean value, therefore tripling the number of rotamers considered at each position.

Weight optimization

To optimize the energy function, we gradually changed one weight at a time, starting from the default value of

1.0. Each obtained energy function was used to compute side-chain/backbone and side chain/side chain interactions for the residues belonging to protein interface and shell. We then used a search algorithm based on the Dead-End Elimination theorem (DEE)[36] to obtain the global minimum energy side-chain configuration (GMEC) for each of the 30 PP interfaces in the learning set of PPIs. We then calculated the RMSD between the locations of heavy atoms (Cβ and further) for interfacial residues in the predicted GMEC structure and those in the actual X-ray structure for all PP complexes in the learning set.

Side-chain repacking

We evaluated the performance of different energy functions on the test set of 69 PP complexes by calculating RMSD as described above and by calculating the percentage of the correctly predicted dihedral angles. For the dihedral angle predictions, we considered a rotamer correctly predicted if its χ angles fall within $\pm$ 20° from that of the rotamer observed in the crystal structure. We compared our success rate for predicting side chain conformations to that obtained by Rosetta 3.1 (fixed backbone design)[37] on the same set of PPIs using the same rotamer library and interfacial and shell residue definitions.

In an additional test, we evaluated intermolecular interactions predicted by each energy function and compared them to those observed in the X-ray structures of the 69 PP complexes. For each complex in the test set, we counted the number of strong intermolecular hydrogen bonds and electrostatic interactions (with the energy of −1 kcal/mol or stronger) and compared the total number as well as the number of the correctly recovered interactions for each energy function. An interaction was considered correctly recovered if a pair of given residues forms the same type of interaction in the predicted and the actual X-ray structure. In addition, we calculated the area of polar atoms buried inside the interface for each of the PPI in the test set in the native PP complex structure and in the structure generated by each energy function after side-chain repacking. All the reported values were averaged over the 69 PPIs. Current rotamers were used to evaluate the reported parameters in native PP interfaces and were not used for calculations using the three different energy functions.

Native sequence recovery

We selected ten high-affinity PP complexes with known three-dimensional structures solved to resolution of 2 Å or better (except for 1b41, solved to resolution of 2.7 Å). The list of the pdb files and complex affinities are summarized in Supporting Information Table S3. For each protein in a high-affinity PP complex, the residues within 3 Å from the other binding partner were defined as interface. A shorter cutoff distance was used for this task to reduce the combinatorial complexity of the problem. Glycines, prolines and cysteins that participate in disulphide bonds were excluded from our analysis. Overall, our dataset included 323 interfacial residues.

We then computationally designed the interfacial residues on one partner of each complex, allowing them to mutate to all amino acids except for glycine and proline. At the same time, interfacial residues on the other binding partner could assume different side-chain conformations, but their amino acid identities were kept fixed. We first attempted to design the lowest-energy interfacial sequence using the deterministic algorithm based on the Dead-End Elimination theorem.[36] In cases where our calculations did not converge (50% of cases), FASTER search algorithm[38] was applied to obtain an approximate solution. The designed sequence was then compared to the native interfacial sequence and the percentage of the correctly recovered amino acids was computed. We repeated this procedure for each of the energy functions studied in this work.

The Kullback–Leibler divergence (DKL) from the native amino acid distribution P to the modeled distribution Q, predicted by each of the energy functions, was calculated according to: DKL = sum (P (i) * log (P (i)/Q (i)), where P (i) is the fraction of each amino acid i in the native interface and Q (i) is the fraction of this amino acid in the redesigned interface.

Interfacial mutation dataset and analysis

We generated a dataset consisting of 404 single mutations in 23 PP complexes, for which the free energy of binding has been experimentally measured using quantitative methods (see Supporting Information Tables S1 and S2). Among these, 53 mutations from 8 pdb files came from the protein design studies, where mutations were rationally introduced for the purpose of enhancing protein binding affinity.[19,21,39–41] The rest of mutations came from studies where mostly alanine mutations were introduced in the interface with the purpose of estimating the role of each interfacial residue on binding affinity. For our dataset we selected only single mutations directly in the binding interface from similar datasets of Sammond *et al.*[19] and Benedix *et al.*[28] In addition, we incorporated our own data on PPI design. For each mutation, we first defined the interface residues as those belonging to the other chain that are within 4 Å from the mutated residue and defined the shell residues as those belonging to the same chain that are within 3 Å from the mutated residue. We then introduced the desired mutation at the specified position using ORBIT,[29] while allowing the interface and shell residues to repack. We froze the side-chain conformations, pulled the proteins apart, and calculated the energy of each unbound partner separately. We next computed the

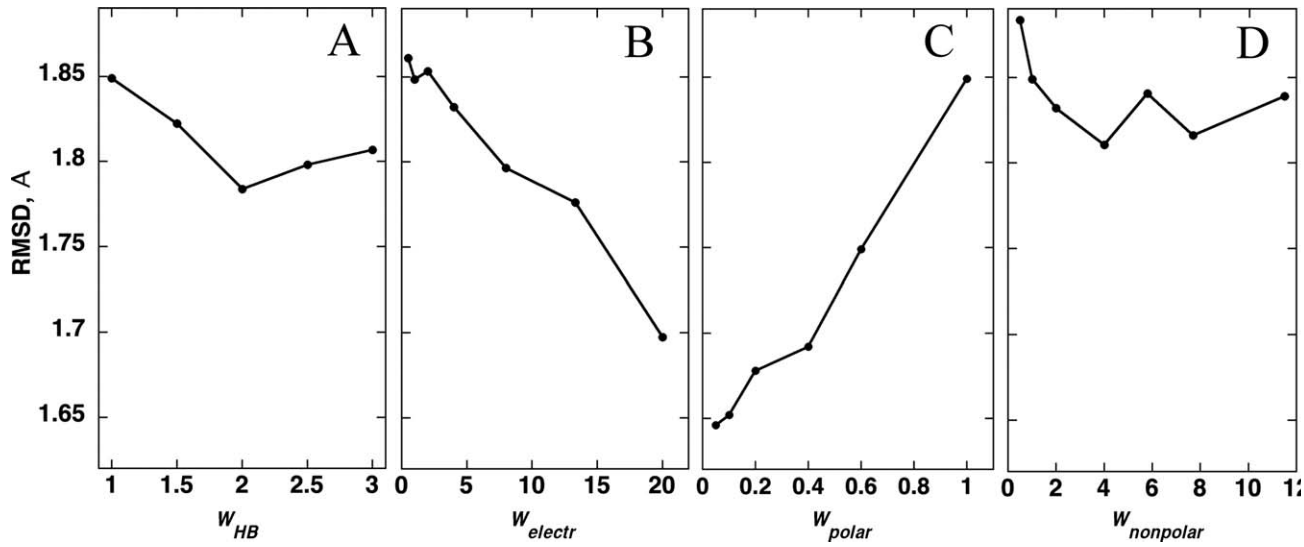

Figure 1

Optimizing the energy function for predicting side-chain conformations in PP interfaces. Each weight of the energy function was varied one at a time, while all other weights were kept at their default value of one. The obtained energy function was used to predict side-chain conformations in a learning dataset of PPIs. The per-complex RMSD in Å between the predicted and the actual positions of heavy atoms in the interface was calculated for each value of a particular weight. Varying the weights for (**A**) hydrogen bond interactions, w_{HB}; (**B**) electrostatic interactions, w_{electr}; (**C**) for polar solvation, w_{polar}; (**D**) for nonpolar solvation, $w_{nonpolar}$.

difference between the total energy of the complex and the energy of the mutated chain. Finally, the change in E_{inter} ($\Delta\Delta E_{inter}$) due to a mutation is simply the difference between the intermolecular energies of the mutated sequence and the WT sequence. All the calculations were repeated using the three energy functions $E_{monomer}$, E_{PPI-1}, and E_{PPI-2}. For this task, we used the rotamer library that adds subrotamers at ± one standard deviation around both χ_1 and χ_2 angles. Water-mediated contacts were defined as those where two polar residues from different chains were linked through a water-mediated H-bond as observed in the PDB file. After exclusion of mutations involving water-mediated contacts our dataset was left with 319 mutations. After exclusion of mutations that significantly destabilize a single chain (by 9 kcal/mol calculated according to E_{PPI-2}), our dataset was left with 312 mutations.

RESULTS AND DISCUSSION

A new energy function for PPI design optimized on side-chain repacking

To optimize an energy function for PPI design, we first compiled a dataset of ninety-nine PP complexes with known high-resolution structures.[30] We divided this database into a learning set of 30 PPIs and a test set of 69 PPIs. For each complex, we defined the binding interface residues as well as the surrounding shell residues. Using ORBIT with the default energy function $E_{monomer}$, we computed the minimal-energy side-chain conformations of all interface and shell residues for every PP complex in the learning dataset. We then calculated the RMSD between the predicted and the actual coordinates for all heavy atoms belonging to the interface residues in the learning dataset to be 1.85 Å. This became a reference number for further improvement.

To optimize the energy function, we gradually varied the weight of each energy term (besides w_{vdW}) while the other weights were kept at their default value of 1.0. The side-chain repacking procedure was then repeated using the modified energy function and the RMSD was recorded (Fig. 1). Altering the weights of different energy terms resulted in variable improvements in RMSD. The most significant improvement was achieved when w_{polar} was reduced to almost zero. The second largest improvement was observed when w_{electr} was increased by up to 20-fold. Increasing w_{HB} 2.0-2.5-fold gave small improvements in RMSD, while further increase was not beneficial. Increasing $w_{nonpolar}$ did not produce substantial improvement in RMSD, while decreasing this weight to below 1.0 resulted in worse RMSD.

We next attempted to merge the best weights into one energy function. We first tried to combine the best values of w_{polar} and w_{electr} (Table II). We found that the minimum RMSD of 1.60 Å is obtained when w_{polar} is reduced to 5–10% of its original value and w_{electr} is increased at least 4-fold. Same RMSDs were recorded for several

Table II
Combining the Best Weights Together

w_{polar}[a]	w_{HB}[a]	w_{Electr}[a]	$w_{nonpolar}$[a]	RMSD, Å[b]
0.30	1	20	1	1.62
0.10	1	1	1	1.65
0.10	1	4	1	1.61
0.10	1	8	1	1.60
0.10	1	20	1	1.60
0.05	1	1	1	1.64
0.05	1	4	1	1.60
0.05	1	20	1	1.60
0.05	2	20	1	1.64
0.05	2	10	1	1.62
0.05	2	4	1	1.65

[a]The given values of the weights produce an energy function that was utilized to predict side-chain conformations of interfacial residues in the dataset of 30 PP complexes.
[b]The RMSD between the positions of heavy atoms (Cβ and further) observed in the interfaces of 30 PP complexes and those predicted by the specified energy function.

combinations of w_{polar} and w_{electr}, suggesting that the energy landscape for side-chain repacking is rather shallow. Once burial of polar amino acids is allowed by reducing w_{polar}, these amino acids can form favorable electrostatic and hydrogen bond interactions across the interface and no additional movements of rotamers would result from increasing w_{electr}. An attempt to improve RMSD further by incorporation of higher w_{HB} value was not successful (Table II). The failure to improve predictions by increasing w_{HB} might be due to creation of intramolecular rather than intermolecular hydrogen bonds. Alternatively, it could be due to the limited size of our rotamer library that does not contain correct rotamers for generation of strong hydrogen bonds. Among a few possibilities of weight combinations that gave the best RMSD for side-chain predictions, we chose the one that has been already experimentally tested in designing affinity-enhancing mutations.[41] This energy function, E_{PPI-2} , is defined as:

$$E_{PPI-2} = 1.0\ E_{vdW\_attr} + 1.0\ E_{vdW\_rep} + 1.0\ E_{HB} + 4.0\ E_{electr} + 0.05\ E_{polar} + 1.0\ E_{nonpolar} \quad (4)$$

E_{PPI-2} nearly eliminates the penalty for polar burial and increases the contribution of electrostatic interactions 4-fold, resembling E_{PPI-1} in the description of polar interactions (Table I). Unlike E_{PPI-1} however, E_{PPI-2} retains the LJ-potential for the description of vdW interactions, similar to $E_{monomer}$.

Predicting side-chain conformations in PP interfaces

We next compared the performance of the three energy functions in their ability to predict side-chain conformations of interfacial residues on the unbiased test set of 69 PPIs. The best RMSD of 1.50 Å was obtained by E_{PPI-1}. This is not surprising since this function underwent a comprehensive optimization for this particular task. E_{PPI-2} obtained a somewhat worse prediction with RMSD of 1.59 Å. Both functions however, repacked interfacial residues substantially better in comparison to $E_{monomer}$ (RMSD of 1.80 Å) [Fig. 2(A)]. In comparison, repacking of the same set of PP interfaces using an energy function implemented in Rosetta[37] produced RMSD of 1.70 Å. Next, we compared the percentage of the correctly predicted dihedral angles χ_1 and χ_2 for the binding interface residues in our dataset [Fig. 2(B,C)]. Again, E_{PPI-1} correctly predicts the highest percentage of the dihedral angles (75.2%, 55.4% for χ_1, and χ_1-χ_2, respectively). Slightly worse dihedral angle predictions are obtained by E_{PPI-2}, and considerably worse by $E_{monomer}$. We also compared the ability of each energy function to predict strong intermolecular hydrogen bonds and electrostatic interactions and compared the predictions to the interactions observed in the X-ray structures of the PP complexes [Fig. 2(D–E)]. We found that $E_{monomer}$ predicts similar number of intermolecular hydrogen bonds and about 30% less electrostatic interactions compared to the number observed in native structures. However, only 45% of the predicted hydrogen bonds and 56% of the predicted electrostatic interactions are native. E_{PPI-1} and E_{PPI-2} perform very similarly, predicting slightly more hydrogen bonds and electrostatic interactions across the interface than observed in nature, while recovering ~60% of the native hydrogen bonds and ~75% of the native electrostatic interactions. Lastly, we compared the polar area buried in the predicted versus native interfaces [Fig. 2(F)]. We observe that while $E_{monomer}$ underestimates this area by about 25%, E_{PPI-1} and E_{PPI-2} predict very similar polar buried areas, 9% and 3% larger than those observed in the native structures, respectively. Hence, a dramatic decrease in the polar burial term does not necessarily lead to overestimation of the polar buried area in the interface.

The above results suggest that with a rather crude optimization procedure, we could obtain an energy function E_{PPI-2} that predicts side-chain conformations almost as well as a function, which was obtained with a sophisticated simultaneous weight optimization procedure (E_{PPI-1}). The slightly better predictions for E_{PPI-1} is attributable to its "soft" vdW potential that is more suitable for the discrete rotameric representation of amino acid side chains. The inferior performance of $E_{monomer}$ on the other hand, could be explained by the high penalty it assigns to burial of polar residues. As a result, this function points all polar amino acids towards the surface of the protein complex, thus eliminating all buried H-bonds and salt bridges at interfaces.

As we mentioned before, design of PPIs is a task that is more difficult compared to repacking of amino acid side chains in PP interfaces. We hence went on and tested the reported energy functions in the other two

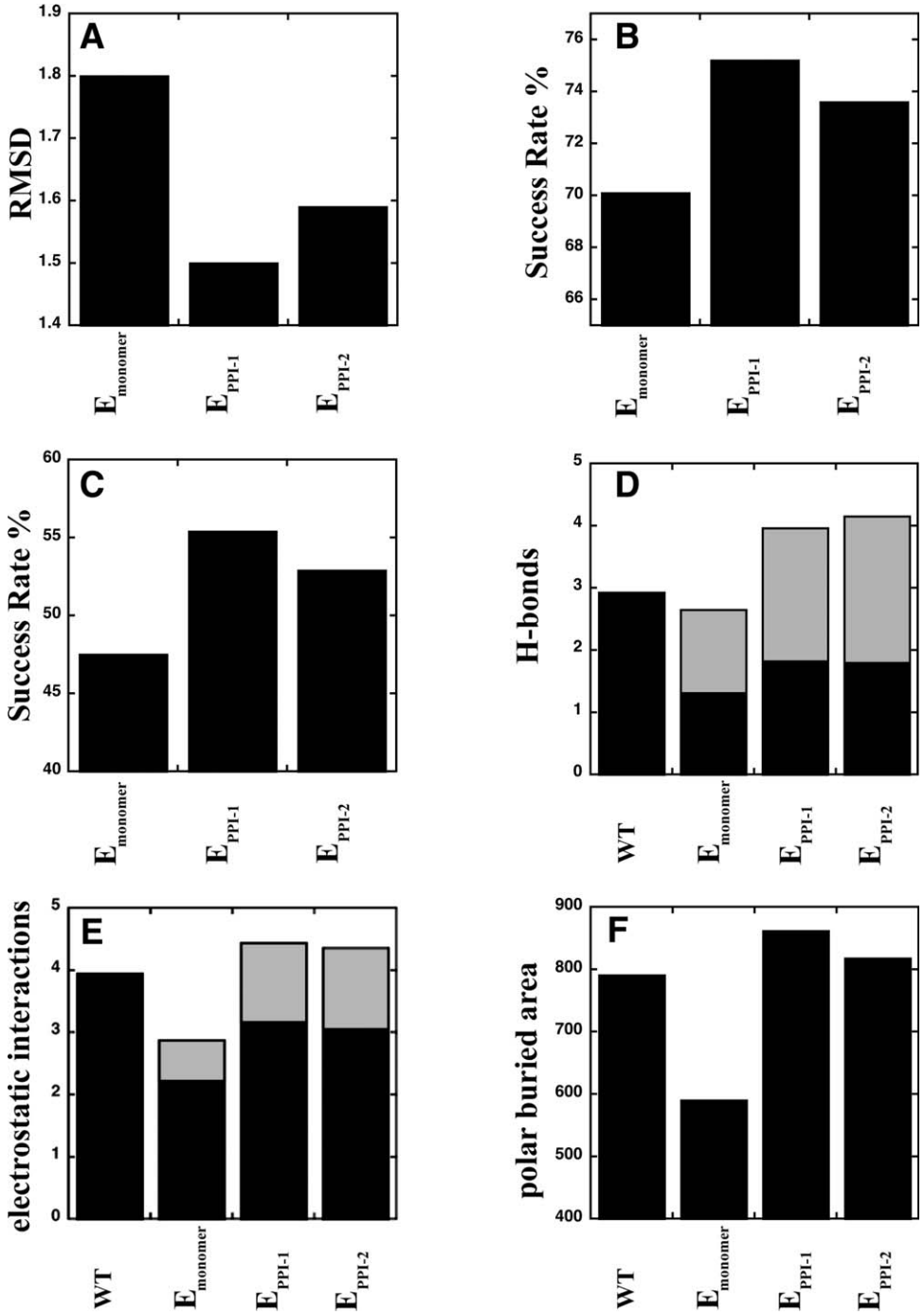

Figure 2

Performance of the three tested energy functions in predicting side-chain conformations in a test set of PPIs. (**A**) RMSD in Å between the actual and the predicted side-chain conformations in PP interfaces. (**B**) and (**C**) show percentage of the correctly predicted dihedral angles, χ_1 (B) and χ_1 and χ_2 (C) in PP interfaces. (**D**) Average number of intermolecular hydrogen bonds per complex observed in native interfaces (WT) or predicted by the specified energy function. Black bars represent the native hydrogen bonds, while gray bars represent non-native ones. Backbone-backbone hydrogen bonds and weak hydrogen bonds (>-1 kcal/mol) were not counted. (**E**) Average number of intermolecular electrostatic interactions per complex observed in native interfaces (WT) or predicted by the specified energy function. Black bars represent the native electrostatic interactions, while gray bars represent non-native ones. Backbone-backbone interactions and weak interactions (>-1 kcal/mol) were not counted. (**F**) Polar buried area (in $Å^2$) in the interface per protein either observed in the native PP interfaces (WT) or predicted by the specified energy function.

Table III
Amino Acid Composition in High-Affinity PP Interfaces: Observed vs. Predicted by Each Energy Function

Amino acids	WT[a]	$E_{monomer}$[b]	E_{PPI-1}[b]	E_{PPI-2}[b]	$\|E_{monomer}-WT\|$[c]	$\|E_{PPI-1}-WT\|$[c]	$\|E_{PPI-2}-WT\|$[c]
Hydrophilic amino acids							
ARG	29	5	62	52	24	33	23
ASN	23	4	7	16	19	16	7
ASP	17	4	16	27	13	1	10
CYS	5	2	1	0	3	4	5
GLN	14	6	11	14	8	3	0
GLU	25	0	26	36	25	1	11
HIS	11	3	5	5	8	6	6
LYS	20	5	27	38	15	7	18
SER	34	7	8	26	27	26	8
THR	20	15	5	19	5	15	1
Total	198	51	168	233	147	30	35
Hydrophobic amino acids							
ALA	14	22	1	2	8	13	12
ILE	15	24	8	10	9	7	5
LEU	18	26	6	10	8	12	8
MET	5	74	13	5	69	8	0
PHE	11	58	6	9	47	5	2
TRP	12	32	69	22	20	57	10
TYR	30	14	44	28	16	14	2
VAL	20	22	8	4	2	12	16
Total	125	272	155	90	147	30	35

[a]Total number observed in interfaces of 10 high-affinity protein-protein complexes selected for this study.
[b]Total number predicted by the specified energy function in the same dataset.
[c]The absolute value of the difference between the total number of the observed and the predicted occurrences of each amino acid by the specified energy function.

categories of our triathlon: native sequence recovery at PP interfaces and prediction of $\Delta\Delta G_{binding}$ due to mutations.

Native sequence recovery test

We performed the native sequence recovery test on a dataset of ten high-affinity PP complexes for which X-ray structures are available (see Supporting Information Table S3). We excluded low- and mid-affinity complexes from our analysis since such complexes are not optimized for binding, and we therefore expected a large majority of their interfacial residues to be mutated during redesign, rendering a comparison between different energy functions meaningless.

For each PP complex in the dataset we redesigned the interfacial residues on one binding partner, while allowing the side chains on the other partner to repack. We then computed the percentage of native amino acids recovered using each energy function. $E_{monomer}$ recovered about 33% of the WT amino acids in the high-affinity PP interfaces, comparable to an average recovery rate of surface residues in a dataset of monomeric proteins.[42] E_{PPI-1} obtained a higher recovery rate of 37.8%. A substantially higher recovery rate of 55% was obtained by E_{PPI-2}. The latter value is slightly higher then an average recovery rate for core residues in a dataset of monomeric proteins,[42] arguing that E_{PPI-2} could reliably reproduce most favorable interactions across PP interfaces.

To better understand how changing different weights affects native sequence recovery, we also calculated the percentage of recovered polarity and amino acid composition for our dataset of high-affinity PP interfaces. For the percentage of recovered polarity, the prediction is defined correct whenever a polar amino acid substitutes a native polar one, or a hydrophobic amino acid substitutes a native hydrophobic one. While $E_{monomer}$ and E_{PPI-1} functions correctly predict 47% and 68% of polarity, respectively, E_{PPI-2} recovers a substantially larger percentage of interface polarity (82%). These findings prove that nearly eliminating the penalty for polar burial and increasing the importance of electrostatic interactions clearly results in designing more native-like sequences in PP interfaces. Nevertheless, a delicate balance between polar versus hydrophobic interactions is important since increasing electrostatics too much produces a decrease in the native sequence recovery (see Supporting Information Table S4).

We further analyzed the amino acid composition of the redesigned high-affinity interfaces for all of the tested energy functions (Table III). Interestingly, the native sequences of high-affinity interfaces are enriched in Ser, Tyr, and Arg; these amino acids are also found extremely frequently in affinity maturation experiments.[43] The interfaces redesigned by $E_{monomer}$ consist mostly of hydrophobic amino acids such as Met and Phe, while completely or nearly devoid of some hydrophilic amino acids (e.g., Glu and Lys). E_{PPI-1}, on the other hand, does a good job when predicting hydrophobic and many polar amino acids but highly overestimates the presence of large amino acids – Trp and Arg. This is not surprising since with its "soft" vdW potential it has a tendency for over-packing. E_{PPI-2} does pretty well in reproducing most hydrophobic and polar uncharged amino acids, but overestimates the occurrence of all charged amino acids. An overall comparison between the

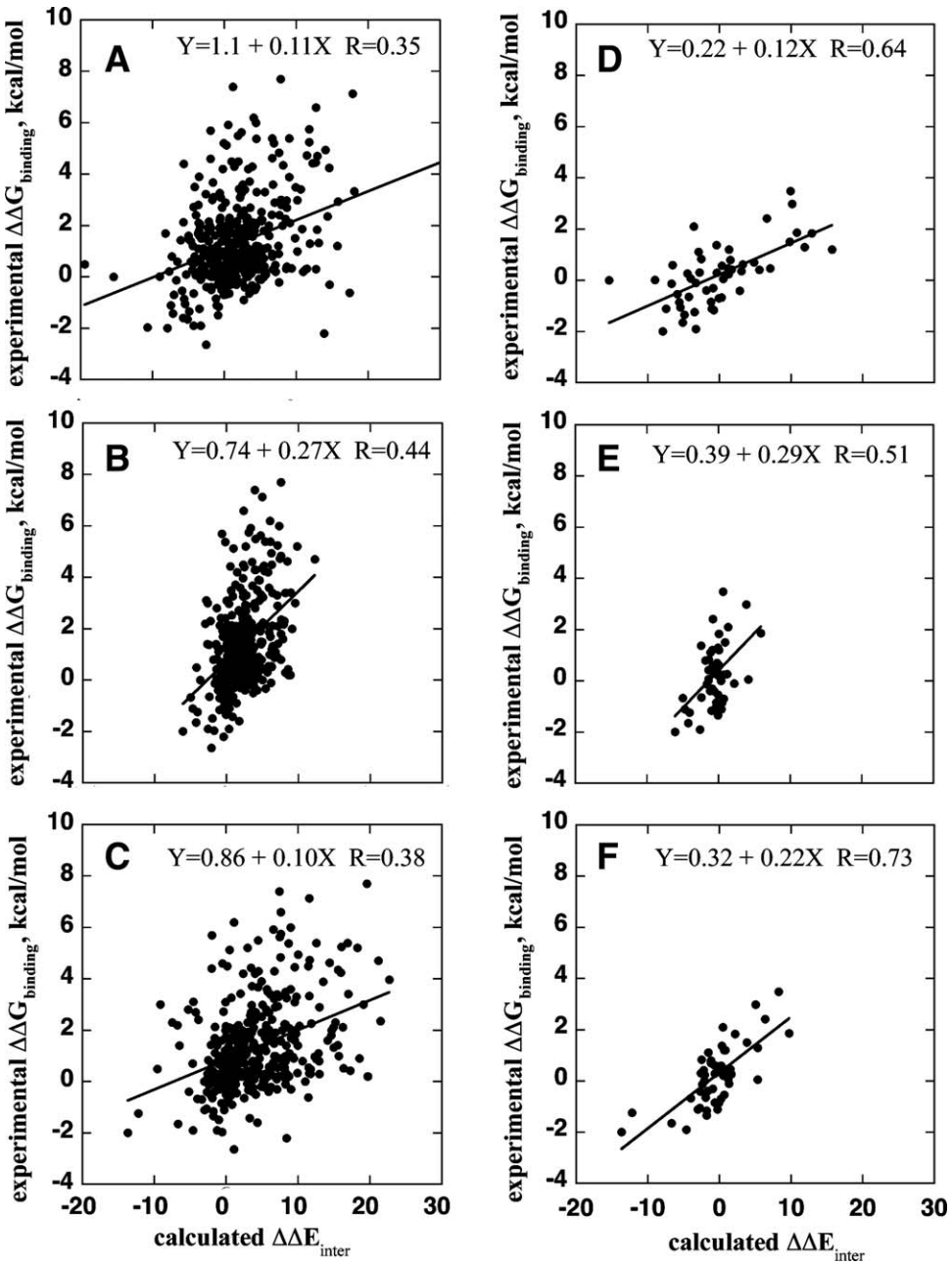

Figure 3

Correlation between the calculated change in the intermolecular energy ($\Delta\Delta E_{inter}$) and the experimentally measured change in free energy of binding ($\Delta\Delta G_{binding}$) upon introduction of single mutations. (**A**)–(**C**) show the data for the whole set of 404 mutations while (**D**)–(**F**) show the data for the subset of 53 designed mutations. (A) and (D) show the results for $E_{monomer}$; (B) and (E) for E_{PPI-1}; (C) and (F) for E_{PPI-2}. The data were fit to a linear function and the correlation coefficient R was calculated for each fit. No points were excluded from the fit.

native amino acid composition and the composition predicted by each energy functions was evaluated using the Kullback–Leibler divergence (DKL), which is zero for equal distributions of amino acids and is a large number for distant distributions. The DKL score was calculated to be 1.73, 0.67, and 0.38 for $E_{monomer}$, E_{PPI-1}, and E_{PPI-2}, respectively, verifying the overall superiority of E_{PPI-2} in reproducing amino acid composition of high-affinity interfaces.

Calculating $\Delta\Delta G_{binding}$ for single interfacial mutations

Next, we tested the performances of the three energy functions in predicting changes in free energy of binding ($\Delta\Delta G_{binding}$) resulting from single mutations located directly at PP interfaces. For this purpose, we compiled a dataset of 404 single mutations for which $\Delta\Delta G_{binding}$ was

Table IV
Correlation Coefficient R Between Experimental $\Delta\Delta G_{binding}$ and $\Delta\Delta E_{inter}$ Calculated Using Different Energy Functions

Energy function	Whole set of mutations (404)[a]	Exclude water-mediated contacts[b] (319)[a]	Exclude mutations destabilizing a single chain[c] (312)[a]	Designed mutations (53)[a]
$E_{monomer}$	0.35	0.38	0.41	0.64
E_{PPI-1}	0.44	0.46	0.46	0.51
E_{PPI-2}	0.38	0.39	0.45	0.73

[a]Total number of mutations used for a particular analysis is given in parenthesis.
[b]Calculated after excluding mutations involving residues that participate in water-mediated intermolecular interactions as observed in the pdb file.
[c]Calculated after excluding all the mutations that destabilize a single protein by more than 9 kcal/mol according to E_{PPI-2}. Results for different cutoff values are summarized in Supporting Information Table S5.

experimentally measured. These mutations came from 23 PP complexes with available high-resolution structures. We calculated the changes in the intermolecular energy of the PP complex due to mutations, using the three tested energy functions. We next determined the correlation coefficient R between the experimental and the predicted value of $\Delta\Delta G_{binding}$ assuming a linear relationship between these two variables [Fig. 3(A–C)]. Without excluding any points from our analysis, the highest R-value of 0.44 is obtained by E_{PPI-1}, the medium value of 0.38 is obtained by E_{PPI-2} and the lowest R-value of 0.35 is exhibited by $E_{monomer}$ (Table IV). The relatively low R-values obtained by the three energy functions are comparable to those reported in a similar test for other protein design energy functions that have not been trained on predicting $\Delta\Delta G_{binding}$.[28] Suboptimal predictions could be in part explained by inconsistency of the experimental dataset that compiles $\Delta\Delta G_{binding}$ values from multiple experiments.

We looked for additional reasons that could explain the relatively low correlation coefficients between the experimental and the predicted $\Delta\Delta G_{binding}$. A source of error in our predictions could come from water molecules that are not modeled in our study yet are quite frequent in PP interfaces where they mediate hydrogen bonds across the binding interface. We hence excluded 85 mutations that are involved in water-mediated intermolecular contacts from our dataset and recalculated the correlation coefficient (Table IV). The correlation was improved but only slightly for all the tested energy functions. This is in agreement with recent findings that only a few very conserved water molecules are important for PPI affinity.[44] We further noticed that in some cases improvement in the predicted $\Delta\Delta G_{binding}$ comes solely or largely from significant destabilization of a mutated protein chain rather than from improvement in the intermolecular interactions. In many such cases, the experimentally measured $\Delta\Delta G_{binding}$ shows reduction in affinity, in contrast to our predictions. We hence thought that large destabilization of single chains could lead to partial protein unfolding (not measured in the binding experiments) or to backbone conformational changes that are not modeled in our predictions. Thus, we excluded from our analysis mutations that significantly destabilize single chains remaining with 312 mutations in our dataset. Using this subset of mutations, we got a significant improvement in the R-values for two out of three energy functions ($E_{monomer}$ and E_{PPI-2}) reaching 0.45 for E_{PPI-2} (Table IV). In contrast, the correlation coefficient for E_{PPI-1} was improved only slightly to 0.46. This lack of improvement for E_{PPI-1} is not unexpected since the function does not predict high destabilization of single chains for any of the mutations due to its "soft" vdW potential. When no destabilization of single chains is observed, E_{PPI-1} and E_{PPI-2} success rate in $\Delta\Delta G_{binding}$ prediction is nearly the same.

To finish our analysis, we tested the three energy functions on the subset of 53 mutations (from 8 pdb files) that were computationally designed in attempt to improve binding affinity of a PPI [Fig. 3(D–F)]. While only some of them proved successful experimentally, all such mutations should be optimal for the particular backbone conformation of the PP complex and hence are very unlikely to result in unfolding or conformational changes of monomeric proteins. Indeed, for the designed set of mutations, R-values were much higher than those calculated for the whole set of mutations. The highest R-value of 0.73 is obtained by E_{PPI-2}. Lower correlation of 0.64 is observed for $E_{monomer}$. Interestingly, E_{PPI-1} that performed the best on the whole dataset, performed the worst on the designed mutations with $R = 0.51$. We conclude that E_{PPI-1} is the best in reproducing $\Delta\Delta G_{binding}$ for mutations that are likely to lead to conformational changes in proteins. Such changes however, could be modeled directly by generating different backbone conformations or expanding the number of considered side-chain conformations.[45,46] When conformational changes are unlikely, as in the case of rationally designed mutations, E_{PPI-1} does not recapitulate some of the favorable intermolecular interactions necessary for affinity enhancement. E_{PPI-2}, in contrast, could predict the effects of such mutations reliably, proving that this function could be successfully used for PPI design. The obtained correlation equation between the predicted and the experimental $\Delta\Delta G_{binding}$ provides a guideline for selecting affinity-enhancing mutations for experimental testing. While small negative values of $\Delta\Delta E_{inter}$ are insignificant, $\Delta\Delta E_{inter}$ of less than -4.0 (according to E_{PPI-2}) identify beneficial mutations with very high probability.

CONCLUSIONS

In this work we compared the performance of three energy functions on the benchmarks related to design of PPIs: side chain repacking at PP interfaces, native

sequence recovery at interfaces and predicting $\Delta\Delta G_{binding}$. We conclude that optimizing for one task (in our case repacking of amino acid side chains in PP interfaces) improves the function's performance on other tasks since both E_{PPI-1} and E_{PPI-2} outperform $E_{monomer}$. However, there is no function that performs best at all three tasks simultaneously. E_{PPI-1} is the best at tasks where backbone and side-chain flexibility are important including side-chain repacking and prediction of $\Delta\Delta G_{binding}$ for the whole dataset of mutations. It does worse, however in the native sequence recovery test and in predicting $\Delta\Delta G_{binding}$ for the designed set of mutations. E_{PPI-2}, which comes second in the side-chain repacking competition, is superior at problems that most closely resemble protein design, i. e. native sequence recovery and predicting $\Delta\Delta G_{binding}$ for the designed set of mutations using a fixed backbone approximation. Since protein flexibility could be incorporated directly through a number of strategies, we recommend E_{PPI-2} for future experimental studies in altering affinity or specificity of various PPIs.

Our best energy function for PPI design nearly eliminates the penalty for burial of polar atoms and enhances electrostatic interactions with respect to the energy function trained on monomeric proteins. Our results are in agreement with previous findings by Nussinov and coworkers that reported a strong correlation between the number of hydrophilic intermolecular bridges across the interface and the free energy of binding. The authors explained the distinct contribution of polar residues to binding and to folding by differences in solvation of the reference states for the two problems.[18] While in protein folding, polar groups are fully solvated in the unfolded state, in protein binding, the same groups are only partially solvated in the unbound state, hence reducing the desolvation cost for protein binding relative to protein folding.[18] In addition, electrostatic interactions are more important in binding since they drive the formation of the initial encounter complex and determine the rate of association between the two proteins.[47] In folding, on the contrary, formation of the initial folded state is governed mostly by hydrophobic interactions. Thus, while the fundamental molecular forces governing protein folding and protein binding remain the same, the simplified energy functions for design of monomeric proteins and protein complexes should be tuned to catch the nuances of each particular problem.

ACKNOWLEDGMENTS

The authors thank B. Raveh for helping to run Rosetta.

REFERENCES

1. Karanicolas J, Kuhlman B. Computational design of affinity and specificity at protein–protein interfaces. Curr Opin Struct Biol 2009;19:458–463.
2. Song G, Lazar GA, Kortemme T, Shimaoka M, Desjarlais JR, Baker D, Springer TA. Rational design of intercellular adhesion molecule-1 (ICAM-1) variants for antagonizing integrin lymphocyte function-associated antigen-1-dependent adhesion. J Biol Chem 2006;281:5042–5049.
3. Shifman JM, Mayo SL. Exploring the origins of binding specificity through the computational redesign of calmodulin. Proc Natl Acad Sci USA 2003;100:13274–13279.
4. Clark LA, Boriack-Sjodin PA, Eldredge J, Fitch C, Friedman B, Hanf KJM, Jarpe M, Liparoto SF, Li Y, Lugovskoy A, Miller S, Rushe M, Sherman W, Simon K, Van Vlijmen H. Affinity enhancement of an in vivo matured therapeutic antibody using structure-based computational design. Prot Sci 2006;15:949–960.
5. Kortemme T, Baker D. Computational design of protein–protein interactions. Curr Opin Chem Biol 2004;8:91–97.
6. Palmer AE, Giacomello M, Kortemme T, Hires SA, Lev-Ram V, Baker D, Tsien RY. Ca^{2+} indicators based on computationally redesigned calmodulin-peptide pairs. Chem Biol 2006;13:521–530.
7. Lo Conte L, Chothia C, Janin J. The atomic structure of protein–protein recognition sites. J Mol Biol 1999;285:2177–2198.
8. Nooren IMA, Thornton JM. Diversity of protein–protein interactions. EMBO J 2003;22:3486–3492.
9. Keskin O, Ma B, Rogale K, Gunasekaran K, Nussinov R. Protein–protein interactions: organization, cooperativity and mapping in a bottom-up systems biology approach. Phys Biol 2005;2:S24–S35.
10. Nooren IM, Thornton JM. Diversity of protein–protein interactions. EMBO J 2003;22:3486–3492.
11. Cunningham BC, Wells JA. High-resolution epitope mapping of hGH-receptor interactions by alanine-scanning mutagenesis. Science 1989;244:1081–1085.
12. Cunningham BC, Wells JA. Comparison of a structural and a functional epitope. J Mol Biol 1993;234:554–563.
13. DeLano WL. Unraveling hot spots in binding interfaces: progress and challenges. Curr Opin Struct Biol 2002;12:14–20.
14. Pal G, Ultsch MH, Clark KP, Currell B, Kossiakoff AA, Sidhu SS. Intramolecular cooperativity in a protein binding site assessed by combinatorial shotgun scanning mutagenesis. J Mol Biol 2005;347:489–494.
15. Skelton NJ, Koehler MF, Zobel K, Wong WL, Yeh S, Pisabarro MT, Yin JP, Lasky LA, Sidhu SS. Origins of PDZ domain ligand specificity. Structure determination and mutagenesis of the Erbin PDZ domain. J Biol Chem 2003;278:7645–7654.
16. Yang J, Swaminathan CP, Huang Y, Guan R, Cho S, Kieke MC, Kranz DM, Mariuzza RA, Sundberg EJ. Dissecting cooperative and additive binding energetics in the affinity maturation pathway of a protein–protein interface. J Biol Chem 2003;278:50412–50421.
17. Zhang Z, Palzkill T. Determinants of binding affinity and specificity for the interaction of TEM-1 and SME-1 beta-lactamase with beta-lactamase inhibitory protein. J Biol Chem 2003;278:45706–45712.
18. Xu D, Lin SL, Nussinov R. Protein binding versus protein folding: the role of hydrophilic bridges in protein associations. J Mol Biol 1997;265:68–84.
19. Sammond DW, Eletr ZM, Purbeck C, Kimple RJ, Siderovski DP, Kuhlman B. Structure-based protocol for identifying mutations that enhance protein–protein binding affinities. J Mol Biol 2007;371:1392–1404.
20. Dufner P, Jermutus L, Minter RR. Harnessing phage and ribosome display for antibody optimisation. Trend Biotech 2006;24:523–529.
21. Selzer T, Albeck S, Schreiber G. Rational design of faster associating and tighter binding protein complexes. Nat Struct Biol 2000;7:537–541.
22. Selzer T, Schreiber G. Predicting the rate enhancement of protein complex formation from the electrostatic energy of interaction. J Mol Biol 1999;287:409–419.
23. Selzer T, Schreiber G. New insights into the mechanism of protein–protein association. Prot Struct Funct Gen 2001;45:190–198.
24. Lippow SM, Wittrup KD, Tidor B. Computational design of antibody-affinity improvement beyond in vivo maturation. Nat Biotechnol 2007;25:1171–1176.

25. Kangas E, Tidor B. Optimizing electrostatic affinity in ligand-receptor binding: theory, computation, and ligand properties. J Chem Phys 1998;109:7522–7545.
26. Lee LP, Tidor B. Optimization of electrostatic binding free energy. J Chem Phys 1997;106:8681–8690.
27. Gilson MK, Sharp KA, Honig BH. Calculating the electrostatic potential of molecules in solution—method and error assessment. J Comp Chem 1988;9:327–335.
28. Benedix A, Becker CM, de Groot BL, Caflisch A, Bockmann RA. Predicting free energy changes using structural ensembles. Nat Meth 2009;6:3–4.
29. Dahiyat BI, Mayo SL. De novo protein design: fully automated sequence selection. Science 1997;278:82–87.
30. Sharabi O, Yanover C, Dekel A, Shifman JM. Optimizing energy function for protein–protein interface design. J Comp Chem 2011;32:23–32.
31. Gordon DB, Marshall SA, Mayo SL. Energy functions for protein design. Curr Opin Struct Biol 1999;9:509–513.
32. Sutton C, McCallum A. An introduction to conditional random fields for relational learning. In: Getoor La TB, editor. Introduction to statistical relational learning. Cambridge, MA: MIT Press; 2007.
33. Yanover C, Schueler-Furman O, Weiss Y. Minimizing and learning energy functions for side-chain prediction. J Comput Biol 2008;15:899–911.
34. Cohen M, Reichmann D, Neuvirth H, Schreiber G. Similar chemistry, but different bond preferences in inter versus intra-protein interactions. Prot Struct Funct Bioinf 2008;72:741–753.
35. Dunbrack RL, Karplus M. Backbone-dependent rotamer library for proteins. Applications to side-chain predictions. J Mol Biol 1993;230:543–574.
36. Desmet J, De Maeyer M, Hazes B, Lasters I. The dead-end elimination theorem and its use in side chain packing problem. Nature 1992;356:539–542.
37. Rohl CA, Strauss CEM, Misura KMS, Baker D. Protein structure prediction using Rosetta. Numerical computer methods, Pt D. Methods Enzymol 2004;383:66–93.
38. Allen BD, Mayo SL. Dramatic performance enhancements for the FASTER optimization algorithm. J Comp Chem 2006;27:1071–1075.
39. Filchtinski D, Sharabi O, Rüppel A, Vetter IR, Herrmann C, Shifman JM. What makes Ras an efficient molecular switch: a computational, biophysical, and structural study of Ras-GDP interactions with mutants of Raf. J Mol Biol 2010;399:422–435.
40. Yosef E, Politi R, Choi MH, Shifman JM. Computational design of calmodulin mutants with up to 900-fold increase in binding specificity. J Mol Biol 2009;385:1470–1480.
41. Sharabi O, Peleg Y, Mashiach E, Vardy E, Ashani Y, Silman I, Sussman JL, Shifman JM. Design, expression, and characterization of mutants of fasciculin optimized for interaction with its target, acetylcholinesterase. Prot Eng Des Sel 2009;22:641–648.
42. Kuhlman B, Baker D. Native protein sequences are close to optimal for their structures. Proc Natl Acad Sci USA 2000;97:10383–10388.
43. Fellouse FA, Wiesmann C, Sidhu SS. Synthetic antibodies from a four-amino-acid code: a dominant role for tyrosine in antigen recognition. Proc Natl Acad Sci USA 2004;101:12467–12472.
44. Reichmann D, Phillip Y, Carmi A, Schreiber G. On the contribution of water-mediated interactions to protein-complex stability. Biochem 2008;47:1051–1060.
45. Georgiev I, Keedy D, Richardson JS, Richardson DC, Donald BR. Algorithm for backrub motions in protein design. Bioinf 2008;24:I196–I204.
46. Humphris EL, Kortemme T. Prediction of protein–protein interface sequence diversity using flexible backbone computational protein design. Structure 2008;16:1777–1788.
47. Schreiber G, Fersht AR. Rapid, electrostatically assisted association of proteins. Nat Struct Biol 1996;3:427–431.