PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

A multiscale approach to predicting affinity changes in protein–protein interfaces

Daniel F. A. R. Dourado and Samuel Coulbourn Flores*

Department of Cell and Molecular Biology, Computational and Systems Biology, Uppsala University, 751 24 Uppsala, Sweden

ABSTRACT

Substitution mutations in protein–protein interfaces can have a substantial effect on binding, which has consequences in basic and applied biomedical research. Experimental expression, purification, and affinity determination of protein complexes is an expensive and time-consuming means of evaluating the effect of mutations, making a fast and accurate *in silico* method highly desirable. When the structure of the wild-type complex is known, it is possible to economically evaluate the effect of point mutations with knowledge based potentials, which do not model backbone flexibility, but these have been validated only for single mutants. Substitution mutations tend to induce local conformational rearrangements only. Accordingly, ZEMu (Zone Equilibration of Mutants) flexibilizes only a small region around the site of mutation, then computes its dynamics under a physics-based force field. We validate with 1254 experimental mutants (with 1–15 simultaneous substitutions) in a wide variety of different protein environments (65 protein complexes), and obtain a significant improvement in the accuracy of predicted $\Delta\Delta G$.

Proteins 2014; 00:000–000.

Key words: biologic design; internal coordinate mechanics; multiscale modeling; affinity maturation.

INTRODUCTION

Protein–Protein Interactions (PPIs) are an important field of research, since they underlie molecular recognition and thus are key to catalysis, assembly of complexes, signaling, degradation, and host-pathogen interactions. There is also considerable industrial interest in PPIs, since understanding them is key to designing proteins for therapy and diagnosis (i.e., biologics), as well as for purification and catalysis. Experimentally probing PPIs is an expensive and time consuming process. In particular, experimental affinity maturation is an unreliable process often requiring months of work.

There is thus a considerable demand for an *in silico* method to analyze PPIs, and in particular to determine the changes in protein–protein interaction energy upon mutation ($\Delta\Delta G$). To be successful, such a method must (1) predict the structure of the mutated complex, and (2) evaluate the enthalpic and entropic contributions to change in binding energy ($\Delta\Delta G$). Both steps have been the focus of considerable work in recent years.

Step (1) can be done with full backbone and side chain freedom for all atoms using molecular dynamics (MD) by combining it with a continuum solvent model such as Molecular Mechanics/Poisson–Boltzmann Surface Area,[1–5] or Molecular Mechanics Generalized Born Surface Area.[2] The computational cost for large complexes can be prohibitive given the number of mutations that often need to be evaluated. In some cases such costly methods do more harm than good, since regions distant from the mutation site, which may be largely unaffected by the mutation, are perturbed from their crystallographic configuration.[6] Q addresses this limitation[7] by doing MD calculations in a limited region. However, there can only be one such region at a time and it must be spherical; this is not suitable for complexes which have multiple spatially separated mutations.

Some methods attempt to skip step (1), using coarse-grained representations to reduce the need for structural equilibration. BeAtMuSIC relies on a set of statistical

Additional Supporting Information may be found in the online version of this article.

Grant sponsor: the Cell and Molecular Biology Department, Uppsala University; Grant sponsor: eSSENCE (essenceofscience.se) Swedish Foundation for International Cooperation in Research and Higher Education (STINT) the Wenner-Gren Foundation.

*Correspondence to: Samuel Coulbourn Flores, Department of Cell and Molecular Biology, Computational and Systems Biology, Uppsala university, Biomedical Center Box 596, 751 24 Uppsala, Sweden. E-mail: samuel.flores@icm.uu.se

Received 5 March 2014; Revised 12 June 2014; Accepted 18 June 2014
Published online 26 June 2014 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24634

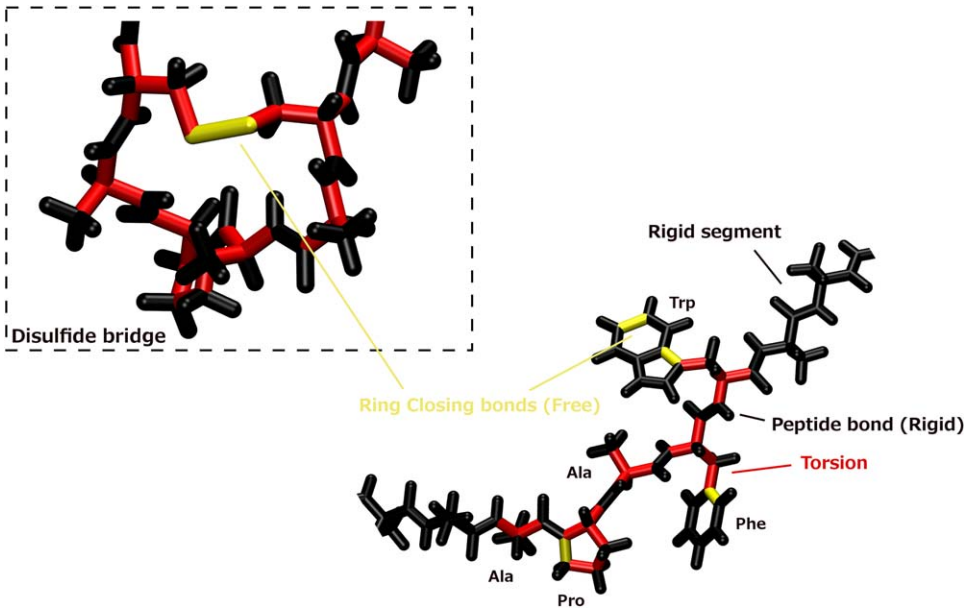

Figure 1

Internal coordinate bond mobilities in proteins. Internal coordinate dynamics uses three types of bond mobilities. Torsion (red bonds) allows dihedral angle changes but keeps bond lengths and angles fixed. Rigid (black bonds) allows no freedom between the connected atoms; two or more atoms can be connected this way to form a single body. Free (yellow bonds) allows bond lengths, angles, and dihedrals to change. Chemically closed rings are represented with *ring closing bonds*. Atoms thus connected are subject to bond stretch, angle bend, and dihedral forces, but are not topologically connected. A set of default bond mobilities applies to all residue types. Most bonds have Torsion bond mobility. Peptide bonds are Rigid, as are guanidinium and amide groups. Hydrogens and double-bonded oxygens are also connected with Rigid bonds. Atoms in the cyclic groups of tryptophan, phenylalanine, tyrosine, and histidine (but not Proline) are connected with rigid bonds and so form a single body.

potentials on a coarse-grained representation of protein structure, to compute binding affinities.[8] mCSM describes the change in protein environment upon mutation using signature vector patterns for each amino acid.[9]

Concoord Poisson–Boltzmann Surface Area (CC/PBSA) generates an ensemble of random protein structures using CONCOORD.[10] It evaluates the electrostatic interaction energy using the Poisson–Boltzmann Surface Area method, implicitly including the entropy of solvation.[11] It uses the Lennard–Jones potential to compute the van der Waals and other contact terms, using a surface-area based approximation to recover the interactions with solvent.[11] However, computer time was at least 4 h/mutant, making it difficult to apply in a high-throughput fashion.

Some methods such as FoldX[12] and Rosetta[13] do the conformational search using rotamers, which vary the dihedral angles of side chains. Rotamer methods are fast, but as of yet they cannot properly model the backbone conformational changes induced by mutation.[14,15] Antibodies are particularly flexible in the complementarity determining region (CDR), therefore an inadequate treatment of backbone flexibility was identified as one of the main reasons computational design has so far had limited success.[14]

Attempts have been made to recover backbone motion by rotating fragments about axes connecting their endpoint α-carbons (Rosetta Backrub),[6,16,17] but this alone did not improve $\Delta\Delta G$ prediction accuracy.[6] GOBLIN[16] used many conformations generated by Rosetta Backrub,[17] to build a Markov Random Field model, which calculates entropy approximately; however accuracy was not high. Guerois *et al.*[12] used the GROMOS force field to minimize the entire protein, but noted that this did not improve accuracy in the $\Delta\Delta G$ calculation.

These observations suggest that a multiscale approach, which limits conformational sampling to the region of the mutation but leaves distant regions unperturbed, should get superior results. This assumes that the mutations are *perturbative*; that is, they do not induce changes in global tertiary structure.[18] This assumption underlies most modern methods,[19] and has been supported crystallographically.[20,21] Multiresolution modeling (MRM) has recently emerged as a paradigm for merging the accuracy of fine grained methods like MD, with the economy of low-dimensional methods. MRM uses different resolutions for different spatial or temporal regions of a system.[7,22,23] MacroMoleculeBuilder (MMB) is an MRM code, which allows close user control over the flexibility, forces, and constraints of the system (Fig. 1).

It can account for local rearrangement by employing one or more small *flexibility zones*[24,25] in internal coordinates.[22] These rearrangements could in principle be computed by dynamical or Monte Carlo methods.[26] Monte Carlo moves in internal coordinates lead to chain

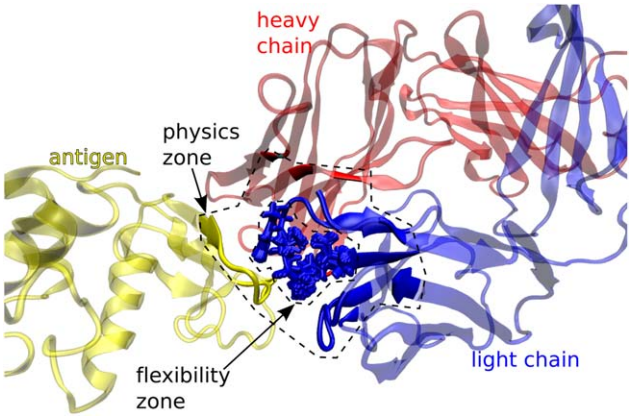

Figure 2

How ZEMu equilibrates the region of a substitution mutation. Residues in the *flexibility zone* (blurred blue licorice) centered about a mutation site have default bond mobilities (Fig. 1). Residues outside the flexibility zone have Rigid bond mobility (Fig. 1) and are fixed to Ground. The *physics zone* consists of residues within 12 Å of the flexibility zone (based on Cα distances), including the flexibility zone itself. Residues in the physics zone have nonbonded (electrostatic and van der Waals) and bonded PARM99 force field terms turned on. The remainder of the protein residues (outside the physics zone) are rigid, fixed to ground, and noninteracting. ZEMu computes the dynamics of this system, for a time span sufficient to settle in an energetically favorable configuration.

closure issues, which must be dealt with by subsequent minimization[27] or other means[28]; for this reason we preferred to equilibrate by dynamics. Our dynamical calculation takes into account physical interactions[29,30] only within the *physics zone*—a small spatial neighborhood about each of the mentioned *flexibility zones* (Fig. 2).[24,25] The approach is economical, avoids perturbing regions distant from the mutation, and still permits most of the relevant motions at the mutation site.

There have also been significant advances for step (2). Some workers compute the single-point (i.e., based on static structure) electrostatic interaction using the Poisson–Boltzmann equation,[14,31] however entropy and van der Waals interactions are ignored. Others have applied free energy perturbation (FEP),[32–35] which "grows" groups of atoms into others and estimates entropy changes using long-time dynamical simulations. However, this method is computationally more costly and convergence becomes more problematic with greater changes in volume or physicochemical property of the side chain. Knowledge based (KB), or empirically trained, potentials such as FoldX[12] compute the effect of substitution mutations on binding energy taking into account the entropy by approximate means, including solvent exposure and physicochemical class.[8,12,19,36–39] KB potentials have become widely accepted for evaluating protein–protein interaction energy.[19] In particular, implicit solvent models, most of which are empirically trained, were used by top scoring groups in Ref. 19. However many KB potentials,[12] and in particular their implicit solvent models,[12,19] are not easily differentiable, thus they cannot be used for dynamics.

ZEMu (zone equilibration of mutants) uses a dynamical equilibration under a physics-based force field for a limited region in internal coordinates[24] for step (1), followed by a FoldX[12] $\Delta\Delta G$ evaluation for step (2). Most leading methods in the literature were validated only against single mutations.[1–6,8,9,12,16,32–38,40,41] However, the lowest (often most interesting) $\Delta\Delta G$'s occur with multiple simultaneous substitutions.[14] We evaluate the method against a dataset that includes a wide range of experimental binding energies ($\Delta\Delta G_{exp}$) from −6.25 to 10.10 kcal/mol, including numerous favorable ($\Delta\Delta G_{exp} < 0$) mutations, as well as multiple mutants (2–15 simultaneous mutations).

Accurately and economically predicting $\Delta\Delta G$ as we do here, for single and multiple mutants, is the key to a future method, which will systematically generate and evaluate many possible mutations with the goal of increasing affinity. Such an *in silico* affinity maturation method will have many applications in basic and applied biomedical research.

METHOD

ZEMu is implemented in MMB and thus works in internal coordinates, explained in Ref. 42. Briefly, this means that molecules consist of *bodies*, which in turn comprise one or more atoms. The *bodies* are connected by *mobilizers*, which can be of type *torsion, rigid*, or *free*. *Torsion* mobilizers grant one degree of freedom; they allow dihedral angle changes, but keep bond lengths and angles fixed. The less common *free* mobilizers grant six degrees of freedom; atoms thus connected can change their dihedrals, bond lengths, and angles. *Rigid* mobilizers grant no degrees of freedom; atoms connected in this way form a single body. All bodies are connected to a parent body in an open tree structure; the ground is the parent of the root body of each molecule. Since the topology must be open, chemically closed cycles are represented with *ring closing bonds,* which comprise bonded *forces* but which do not represent a topological connection. Ring closing bonds are used to represent cysteine bonds and to close proline rings. All amino acids have a default mobilizer configuration, but this can be modified by the user at the level of residues or individual bonds.[42] These concepts are illustrated in Figure 1.

MMB makes it easy to limit flexibility to small regions of interest called *flexibility zones* (Fig. 1). In this work the *flexibility zones* include the position of substitution mutations, plus residues adjacent to that mutation position in sequence (Fig. 2). The atoms in each such *flexibility zone* move under the influence of electrostatic, van der Waals, and other forces due to atoms in the *physics zone*. Each *physics zone* comprises the *flexibility zone* plus all residues within a certain radial distance of the latter.

As mentioned in the *Introduction,* ZEMu consists of (1) creating the substituted structure and equilibrating it dynamically under a physics-based force field, and (2) evaluation of $\Delta\Delta G$ using the FoldX potential. As mentioned, dynamics cannot be done under the FoldX potential; we therefore used the PARM99 force field[29,30] for this equilibration. The potential energy decayed approximately exponentially during the course of the simulations. Simulation time was the same for all runs, chosen to be approximately three exponential decay constants long, to ensure convergence.

MMB's method for creating mutant structures follows. First, the *mutant* structure is created with naïve values for all bond lengths, angles, and dihedrals. The geometry of bonds in said structure, is matched to any corresponding bonds in the *wild-type* structure. Bonds in the substitution sites which have no equivalent in the *wild-type* structure, are left at the default bond geometry. This may result in steric clashes, which are resolved in the subsequent equilibration.

For each mutant to be evaluated, ZEMu equilibrates two complexes: the *wild-type* complex, which is typically has been observed crystallographically, and the *mutant* complex, in which we have introduced one or more substitution mutations using ZEMu. The *mutant* complex usually changes more than the *wild-type* complex during equilibration, because it contains residues that are not in the crystallographically observed structure. The *wild-type* must be processed in precisely the same way as the *mutant* complex, so that any systematic force field bias is introduced into both complexes in the same way.[19]

FoldX has a minimizer which works by a side-chain rotamer search. It is very limited, moving side chains only slightly to eliminate small steric clashes, and cannot be considered a true equilibration.[43] We applied this minimizer to the final PARM99-equilibrated *wild-type* and *mutant* structures described above, as a final structural processing step.

We subsequently dealt with step (2), evaluation of $\Delta\Delta G$. For each proposed mutation locus, after equilibrating under PARM99 and minimizing under FoldX as described, we evaluated ΔG_1 (the binding free energy of the *wild-type* complex[40]) using the FoldX potential.[12] Similarly, for the equilibrated and minimized *mutant* complex, we computed ΔG_2.[40] $\Delta G_2 - \Delta G_1$ is an estimate of $\Delta\Delta G_{exp}$[40]:

$$\Delta\Delta G_{ZEMu} \equiv \Delta G_2 - \Delta G_1$$

Datasets

SKEMPI is a dataset of protein–protein complexes mutants for which $\Delta\Delta G_{exp}$ and PDB structures are available.[44] It comprises 3047 (2792 unique) mutants of 85 complexes. Our dataset consists of 1254 mutants (with 1–15 substitutions/mutant), of 65 different cocrystals, extracted from SKEMPI (Supporting Information Table S2). The dataset has the following characteristics:

1. *Includes a broad range of experimental binding energies.* These range from −6.25 to 10.10 kcal/mol.
2. *Provides a good sampling of stabilizing mutants.* Our dataset has 267 mutants with $\Delta\Delta G_{exp} < 0$.
3. *Includes mutations to all 20 canonical amino acids.*
4. *Includes cocrystals with 1–15 simultaneous mutations.* Most other leading methods were validated exclusively with single mutants.[1–6,8,9,12,16,32–38,40,41]
5. *Avoids complexes where a single position is repeatedly mutated.* ZEMu equilibrates the mutation site in the *wild-type*, in addition to the mutant complex. When a single position is repeatedly mutated any bias in the wild-type complex will affect the final $\Delta\Delta G$ of all the mutants. If in a given position one analyzes all 19 possible amino acid substitutions, accuracy becomes dependent on just 5% of the analysis. All leading flexible backbone methods[11,16,45] obtain their best performance on validation datasets, which exclude such cases. This applies to 977 mutants.
6. *Avoids experimental data which is not peer reviewed.* This applies to 304 mutants.
7. *The dataset should not have redundant entries (more than one $\Delta\Delta G_{exp}$ value for a given mutation).* In the cases this happens one entry is randomly chosen (255 mutants were left out).
8. *Avoids cocrystals/publications with the majority of mutations outside the protein–protein interface. Immune complexes were prioritized.* Since we intend to validate ZEMu for measuring protein–protein binding affinities the dataset should sample cocrystals with a higher incidence of mutations in the protein–protein interface. This applies to 143 mutants.
9. *Avoids mutants in which binding affinity might depend on residues not resolved in the crystal.* This specifically applies to 27 TEM-1 beta-lactamase/BLIP mutants in positions 49 and 50 of chain B (Supporting Information Fig. S2).
10. *Avoids cocrystals having a small number of mutants, and $\Delta\Delta G_{exp}$ near zero (unlikely to affect the results).* This applies to 78 mutants.
11. *Avoids co-crystals for which binding affinities were not measured accurately.* This specifically applies to Bovine Tripsin/BPTI_mutants crystals (total 9 mutants). K_d was experimentally determined at pH 5.0 and not at pH 8.3 like the *wild-type*.[46]

Our structural dataset comprises the following PDB accession codes[44]: 1DVF, 1JTG, 1MLC, 1VFB, 1JRH, 1EAW, 3BN9, 3NPS, 2QJA, 2QJB, 2QJ9, 1KTZ, 1CBW, 1REW, 3BK3, 1GC1, 1DAN, 1EMV, 2WPT, 1FR2, 2GYK, 2VLO, 2VLP, 2VLN, 2VLQ, 1Z7X, 1A4Y, 2JEL, 1IAR, 1A22, 2A9K, 1FC2, 1FCC, 1NMB, 1AHW, 1NCA, 1KIQ, 1KIP, 1KIR, 2VLJ, 1JCK, 1DQJ, 1LFD, 1HE8, 2PCC, 2PCB, 1XD3, 1S1Q, 2OOB, 1TM1, 1ACB, 1CSE, 1F47, 1UUZ, 2B42, 1E96, 1AK4, 2VLR, 2J12, 2J1K, 1KAC, 1P69, 1P6A, 2I26, 2I9B. To better understand ZEMu's

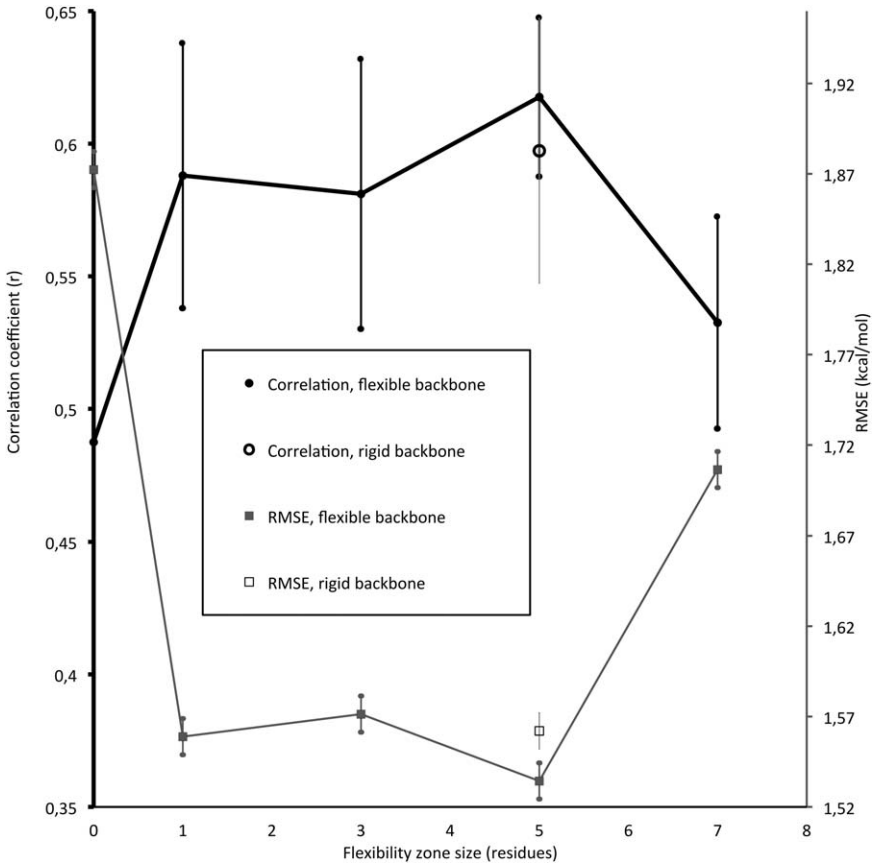

Figure 3

Optimization of the *flexibility zone*. Correlation (values on left axis) and RMSE (values on right axis) versus various sizes of flexibility zones, ranging from zero (corresponding to the FoldX-only result) to 7 residues centered about each mutation site. Note that in the case of multiple mutants near each other in sequence, two or more *flexibility zones* could merge into a single one. We also computed performance with five flexible side chains and a rigid backbone (empty square and circle).

performance we divided the dataset into *single* and *multiple* mutants. The single mutants were further divided in 5 different subsets as follows. The set of *charged* residues includes all the positions mutated where the wild-type and/or substituted residue is a His, Arg, Lys, Asp, or a Glu. Similar logic applies to *polar* (Asn, Gln, Ser, Thr, Tyr, Cys), *nonpolar* (Val, Gly, Pro, Ala, Ile, Leu, Phe, Trp), *bulky* (Trp, Phe, Tyr), and *special* (Pro, Gly) residues. The *multiple* mutants were divided in two subsets, with 2–3 and 4–15 simultaneous substitutions.

We also test ZEMu on the CC/PBSA dataset,[11] composed of 244 single and 123 multiple mutants of 9 cocrystals. The performance of leading methods has been benchmarked against this dataset.[45]

RESULTS

To optimize the *flexibility zone* we varied its size from zero to seven residues, while maintaining a fixed *physics zone* of 12 Å. For each size of the *flexibility zone* we computed $\Delta\Delta G_{ZEMu}$ and compared to $\Delta\Delta G_{exp}$ for the entire dataset described above. The best results are with five flexible residues, since root mean square error (RMSE) (see equation in Supporting Information) and correlation deteriorate for *flexibility zones* larger than this (Fig. 3). We also evaluated performance for *physics zones* of different sizes, and found that accuracy plateaued at 12 Å (Fig. 4). This is consistent with the common practice in the MD community, of cutting off interactions at interatomic distances greater than 8–12 Å.[47] For five flexible residues and a physics zone of 12 Å we obtained an RMSE of 1.54 kcal/mol and a correlation of 0.62. RMSE and correlation are taken to be converged within ±0.01 kcal/mol and ±0.03, under those conditions for this dataset (Table I).

We also repeated the above, except with a *rigid* backbone, leaving only the side chains of the flexibility zone residues flexible. With a *flexibility zone* of 5 residues and a rigid backbone, the performance was marginally worse than with a flexible backbone (Table I). The analyses of the single and multiple mutants separately revealed

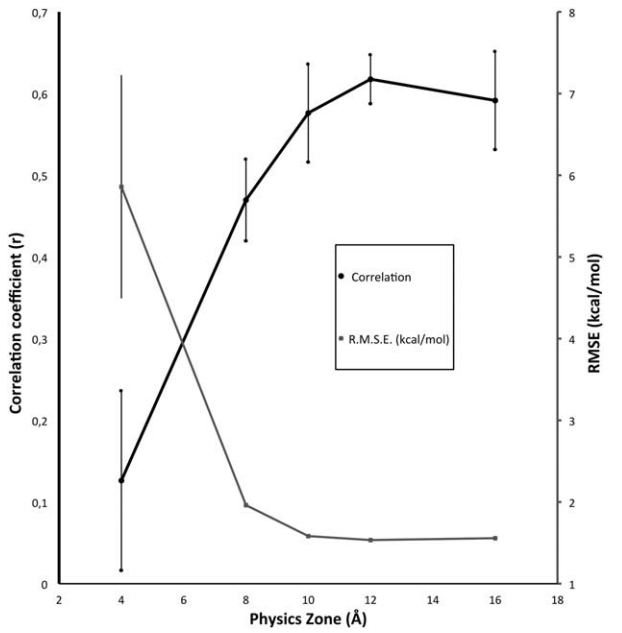

Figure 4

Optimizing the size of the *physics zone*. We varied the size of the *physics zone* from 4 to 16 Å about the *flexibility zone*. The size of the *flexibility zone* is fixed at 5 residues. In this work we used a physics zone of 12 Å since no further improvement in correlation or RMSE was obtained for zones larger than this.

interesting facts. For the single mutants the rigid backbone is actually slightly better than the flexible backbone (Table II). However, in the multiple mutants the performance of the flexible backbone is considerably higher (Table I). Thus adding backbone flexibility improves accuracy with respect to the *rigid* backbone, 5-residue flexibility zone.

For 1/5 of the dataset we also tested a *flexibility zone* that incorporated *all* residues in the protein–protein interface (defined as all residues in either partner that are within 6 Å of the *other* partner), in addition to the 5-residue window, still with a rigid backbone throughout. The RMSE was 1.53 kcal/mol while the correlation slightly increased to 0.66. However, it does not improve accuracy with respect to the *flexible* backbone, 5-residue flexibility zone. It is thus better *not* to flexibilize the interface, and just stay with the *flexible* backbone, 5-residue flexibility zone, since the former needs some 460-fold more computer time than the latter.

We compared the $\Delta\Delta G$ obtained with ZEMu ($\Delta\Delta G_{\mathrm{ZEMu}}$) against that obtained using only FoldX's limited side-chain rotamer search minimization plus KB energy evaluation (FoldX-only, or $\Delta\Delta G_{\mathrm{FoldX}}$). The results of running FoldX-only cannot be directly compared with the ones published[12] since FoldX was trained and validated against a dataset of changes in protein stability[12] rather than protein–protein interaction. Further, the dataset contained only single mutants, and very few (<1%) stabilizing mutants ($\Delta\Delta G_{\mathrm{exp}} < 0$ kcal/mol).

Full dataset

For the full dataset the correlation between $\Delta\Delta G_{\mathrm{FoldX}}$ and $\Delta\Delta G_{\mathrm{exp}}$ is 0.49, with a slope of 0.62 and a RMSE of 1.88 kcal/mol (Table I), in line with its recently reported performance.[6,16] For ZEMu the correlation increases to 0.62, the slope increases to 0.66 (Fig. 5) and the RMSE decreases to 1.54 kcal/mol (Table I). Antibodies have particularly flexible CDRs.[14] ZEMu performance on the 30 immune complexes (Table I) is quite similar to that on the entire dataset. For the biomedically relevant case of the 267 stabilizing mutants, $\Delta\Delta G_{\mathrm{FoldX}}$ shows an RMSE of 1.92 kcal/mol, while for $\Delta\Delta G_{\mathrm{ZEMu}}$ the RMSE is just 1.22 kcal/mol (Table I).

Single mutants

For the single mutants, FoldX-only shows an RMSE of 1.55 kcal/mol and correlation of 0.39, while with ZEMu we obtain a lower RMSE of 1.34 and a higher correlation of 0.51 (Table II). For the nonpolar and polar residue subsets the performance of ZEMu does not differ

Table I

Comparison of FoldX-Only and ZEMu $\Delta\Delta G$ Evaluation

			FoldX-only		ZEMu		ZEMu with rigid backbone	
Dataset	Number of mutants	Number of complexes	Correlation	RMSE (kcal/mol)	Correlation	RMSE (kcal/mol)	Correlation	RMSE (kcal/mol)
All mutants (1–15 simultaneous mutations)	1254	65	0.49	1.88	0.62	1.54	0.60	1.57
Immune complexes	480	29	0.52	1.70	0.59	1.53	0.56	1.59
Stabilizing mutants	267	38	0.17	1.92	0.31	1.22	0.39	1.17
All multiple mutants (2–15)	273	30	0.49	2.76	0.64	2.11	0.61	2.29
2–3 simultaneous mutations	237	21	0.63	2.06	0.64	2.09	0.60	2.14
4–15 simultaneous mutations	36	9	0.49	5.43	0.69	2.24	0.78	3.10

ZEMu considerably increases correlation and decreases RMSE as compared to FoldX-only. Results are for optimum parameters, namely a flexibility zone of 5 residues centered about each mutation position, and a physics zone including residues within 12 Å of the flexibility zone. Data for an alternative ZEMu protocol with rigid backbone and flexible side chain is also shown. By randomly dividing the dataset in 4 equal portions and recomputing RMSE and correlation we estimate RMSE is converged within ±0.01 kcal/mol while the correlation is converged within ±0.03.

Table II
Comparison of FoldX-Only and ZEMu $\Delta\Delta G$ evaluation for single mutant subsets

Dataset	Number of mutants	Number of complexes	FoldX-only Correlation	FoldX-only RMSE (kcal/mol)	ZEMu Correlation	ZEMu RMSE (kcal/mol)	ZEMu with rigid backbone Correlation	ZEMu with rigid backbone RMSE (kcal/mol)
All single mutants	981	62	0.39	1.55	0.51	1.34	0.49	1.30
Charged residues	434	53	0.31	1.85	0.46	1.44	0.45	1.44
Polar residues	326	47	0.49	1.22	0.49	1.23	0.49	1.21
Nonpolar residues	221	48	0.50	1.31	0.59	1.30	0.55	1.27
Bulky residues	150	36	0.48	1.46	0.48	1.51	0.50	1.48
Special residues	48	24	0.68	1.62	0.73	1.59	0.78	1.41

Flexibility zone of 5 residues; Physics zone of 12 Å about the flexibility zone. Data for an alternative ZEMu protocol with rigid backbone and flexible side chain is also shown. RMSE and correlation are converged within an estimated ±0.01 kcal/mol and ±0.03, respectively (Table I).

significantly from FoldX-only (Table II). However, for the charged residues the difference is remarkable. ZEMu shows a RMSE of 1.44 kcal/mol and a correlation of 0.46, while FoldX-only has an RMSE of 1.85 kcal/mol and a correlation of 0.31. For the bulky residues ZEMu RMSE is slightly worse than FoldX-only, while the correlation is identical (Table II).

Finally, for the special residues subset ZEMu overcomes FoldX-only (Table II). However, ZEMu rigid-backbone is better than flexible-backbone performance (Table II). If we analyze this subset in detail we can attribute rigid backbone over-performance to mutations in sequences GP (Gly-Pro), GGP (mutants AD488G, GD489A, and GD489V of 1AK4) and PPP (mutant PA417S of 1KAC). For these mutants rigid backbone RMSE is 2.4 kcal/mol while flexible backbone RMSE is 3.4 kcal/mol. In fact, these particular sequences, which have very unique backbone arrangement, are better described by the crystallographic backbone arrangement, since FoldX-only RMSE is just 1.7 kcal/mol.

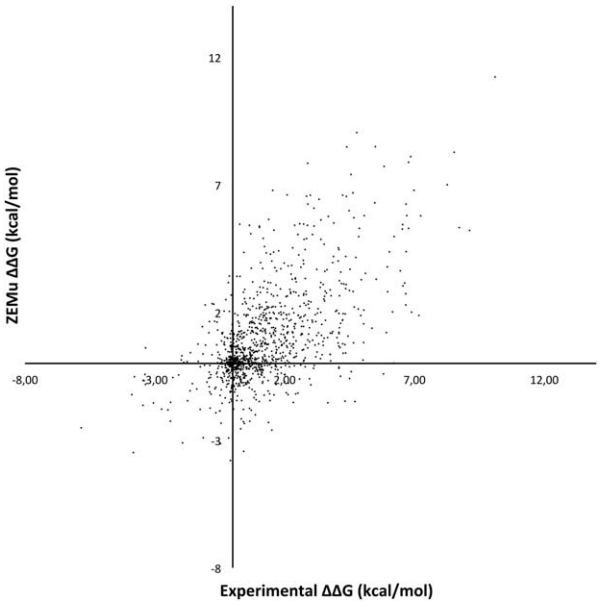

Figure 5
ZEMu versus experimental $\Delta\Delta G$. Note substantial improvements in correlation, slope and RMSE compared to FoldX-only (see Tables I and II and Supporting Information Fig. S1).

Multiple mutants

The analysis of the multiple mutant subsets indicates that for double and triple mutants ZEMu and FoldX-only performance is essentially the same (Table I). For the 4–15 simultaneous substitution subset ZEMu is considerably better than FoldX-only (Table I). FoldX-only RMSE for three of these mutants is a staggeringly high 17.6 kcal/mol. For the same mutants ZEMu with the flexible backbone protocol shows an RMSE of just 1.84 kcal/mol. We created a new subset without the worst three mutants of ZEMu and FoldX-only (total of 6 mutants) and the former still outperforms the latter—with RMSE of 1.84 and 1.98 kcal/mol, respectively. For the entire dataset, minus the six outliers, the RMSE becomes 1.51 kcal/mol and 1.67 kcal/mol, respectively.

Computational cost

For single mutants (5 flexible residues), ZEMu equilibration takes between 5 and 10 min, while FoldX evaluation required about 2–4 min on a single core of a 3.00 GHz AMD Opteron 6220 processor. For multiple mutants the equilibration time was longer (up to 25 min for a typical case of a quadruple mutant), but the FoldX evaluation time was roughly the same.

Comparison with flexible backbone methods

Rosetta variants

GOBLIN[16] was validated with a dataset composed of 704 single mutants of 8 protein–protein complexes. They obtained an RMSE of 1.63 kcal/mol and a correlation of 0.56.[16] Variants of Rosetta[6] use side chain repacking, all-atoms minimization, and reweighting of energy terms

Table III
Performance of several leading methods on the CC/PBSA[11] dataset

Method	Number of mutants (Single/Multiple mutants)	Single mutants		Multiple mutants	
		RMSE(kcal/mol)	Correlation	RMSE(kcal/mol)	Correlation
BeAtMuSIC[45]	242/NA	1.36	0.52	NA	NA
FoldX	244/123	1.64	0.44	3.21	0.53
ZEMu	244/123	1.54	0.53	2.43	0.65
CC/PBSA[11]	244/123	1.13	0.71	1.37	0.74
Li et al.(pred2)[45]	242/NA	1.20	0.69	NA	NA
Li et al.(pred1)[45]	242/123	1.27	0.62	1.90	0.63

Methods are in order of ascending computational cost. BeATMuSIC results were obtained by Li et al.[45]; CC/PBSA results were taken from Ref. 11.

against a dataset of 1210 mutants, obtaining a correlation of 0.73. A protein stability (rather than protein–protein binding) dataset was used. No RMSE data was published.[6] Both these variants and GOBLIN were only validated against single mutants.

MD-based methods

CC/PBSA[11] obtained good results for protein–protein binding affinity, though a later independent test[6] obtained a correlation of just 0.56.

Li *et al.*[45] published recently an interesting modified MM-PBSA energy function with 4 or 11–13 weight factors. When they trained one variant of method on the SKEMPI-derived dataset (4 weight factors) and then tested it on the CC-PBSA dataset the correlation was 0.62 and the RMSE was 1.27 kcal/mol for the single mutants, while for the multiple mutants they obtained a correlation of 0.63 and RMSE of 1.90 kcal/mol (Table III).

ZEMu adds no new weighting parameters, and is considerably faster than both MD-based methods. CC/PBSA takes 249 min to compute a protein of just 149 residues on a 3.2 GHz Intel Xeon processor, while the method of Li *et al.* needs 15 min on 16 (2.8 GHz Intel EMT64) processors for equilibrating a small 160-residue protein, plus 10 min for calculating binding energy on a single 2.6 GHz AMD Opteron processor.

ZEMu is substantially faster than MD-based methods. Importantly, MD-based methods have computer time requirements that increase rapidly with the size of the complex, while ZEMu's multiscale approach is largely independent of size.

Comparison with rigid backbone methods

BeAtMuSIC's[8] 10-fold cross-validation on a dataset of 2007 single mutants yields an RMSE of 1.80 kcal/mol and a correlation of 0.40, while on the CC/PBSA dataset the RMSE is 1.36 kcal/mol and the correlation 0.52.[45] mCSM,[9] using the BeAtMuSIC dataset and 10-fold cross-validation as BeAtMuSIC, yields an RMSE of 1.55 kcal/mol and a correlation of 0.58. These methods have been validated only for single mutants and introduce fitting parameters. The computational cost of both methods is quite low. Using BeAtMuSIC web server it is possible to compute a single mutant in ~1 s, while the mCSM server requires ~3 s.

DISCUSSION

In this work we describe and validate ZEMu. ZEMu finds energetically favorable *local* rearrangements following mutation at protein–protein interfaces. Key to the success of ZEMu is our use of a limited *flexibility zone*, which leads to efficient sampling in the neighborhood of the mutation, and importantly also avoids perturbing regions distant to that mutation. This leads to considerable improvement in the performance of the popular FoldX[12] potential.

We report the appropriate sizes of the flexibility (Fig. 3) and physics zones (Fig. 4). We varied the size of the *flexibility zone* from zero (corresponding to FoldX-only) to seven flexible residues, with a *physics zone* comprising all residues within 12 Å of the flexible residues. In this run both backbone and side chains were flexible. The accuracy resulting from use of the five-residue *flexibility zone* may be related to the minimum degrees of freedom required to guarantee chain closure.[48] A five-residue *flexibility zone* contains 10 backbone torsional degrees of freedom (each residue has free φ and ψ angles), thus five on each side. This may be sufficient when (as here) the rearrangements are small.

We found that even a single-residue *flexibility zone* improved accuracy of $\Delta\Delta G$. However, the *flexibility zone* should not be too large, as this would perturb regions distant from the mutation; the ideal size is five residues. This finding further supports the idea that most substitution mutations have a perturbative rather than global effect (Supporting Information Fig. S4).[6,18,20,21]

We then repeated the five-residue *flexibility zone* calculation with the backbone rigid, allowing side chain motions only, and obtained moderately lower $\Delta\Delta G$ accuracy (Fig. 3). Expanding the *flexibility zone* to include the entire interface closed the accuracy gap, but at too high a cost in computer time to be practical. Nonetheless, the finding that the *flexibility zone* can be larger if the backbone is rigid, underscores the idea that not perturbing

global structure[6] is about as important as getting good sampling at the mutation site.

Consistent with common MD practice, we found that the *physics zone* need not be larger than 12 Å (Fig. 4) to sufficiently represent the tails of electrostatic interactions with neighbors.[47] Our equilibration does not add weighting factors, indicating that our method should also improve the performance of rigid-backbone potentials other than FoldX. Our validation on a large and diverse dataset further underscores the transferability of the results to other complexes and applications.

The analysis of the single- and multiple-mutant subsets gave insight into the reasons for ZEMu's accuracy improvements. ZEMu's equilibration is able to significantly improve the description of electrostatic interactions. This appears to also be demonstrated by the analysis of a small number of mutants where the crystallographic structure of the wild-type and mutant and complexes is available (Supporting Information Table S1). For a relatively large, crystallographically observed conformational change induced by mutation to a charged residue, ZEMu recapitulated the local rearrangement at the mutation site, more accurately than FoldX-only (Supporting Information Fig. S4 and Table S1). However clear examples of this were few, since the rearrangements were typically small and tended to be drowned out by small but numerous differences in conformation between the independently solved crystallographic structures.[6] ZEMu is better than FoldX in describing simultaneous substitutions and that improvement is particularly marked when the substitutions are consecutive, due to the flexible backbone. Around 72% of the mutants in the 4–15 simultaneous substitutions subset have consecutive substitutions. In this case, ZEMu's flexible-backbone clearly outperforms its rigid-backbone protocol. So, it seems that a flexible backbone is essential to efficiently describe the rearrangements involved in accommodating simultaneous consecutive substitutions.

ZEMu's economy makes it useful for use on a laptop, or in a future *in silico* affinity maturation method. For example, to evaluate all possible single mutants of some 14 residues of the CDR of an antibody to all 19 possible alternatives would cost only about 32 core-hours. The best of these single mutations (those of highest affinity) could then be combined to generate multiple mutants of even higher affinity at modest cost.

AVAILABILITY

MMB 2.14 software and documentation are free for academics, and can be downloaded from simtk.org/home/rnatoolbox. A ZEMu tutorial is provided as a supplement.

ACKNOWLEDGMENTS

Daniel Larsson, John Löfblom, Johan Åqvist, Alwyn Jones, and David Bejker provided helpful critiques and discussion. Michael Sherman made crucial Simbody upgrades. J. Schymkowitz provided FoldX advice. Calculations were done on resources provided by the Swedish National Infrastructure for Computing (SNIC) at UPPMAX.

REFERENCES

1. Massova I, Kollman PA. Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies. J Am Chem Soc 1999;121:8133–8143.
2. Massova I, Kollman PA. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect Drug Discov 2000;18:113–135.
3. Huo S, Massova I, Kollman PA. Computational alanine scanning of the 1:1 human growth hormone-receptor complex. J Comput Chem 2002;23:15–27.
4. Moreira IS, Fernandes PA, Ramos MJ. Unravelling hot spots: a comprehensive computational mutagenesis study. Theor Chem Acc 2007;117:99–113.
5. Moreira IS, Fernandes PA, Ramos MJ. Unraveling the importance of protein-protein interaction: application of a computational alanine-scanning mutagenesis to the study of the IgG1 streptococcal protein G (C2 fragment) complex. J Phys Chem B 2006;110:10962–10969.
6. Kellogg EH, Leaver-Fay A, Baker D. Role of conformational sampling in computing mutation-induced changes in protein structure and stability. Proteins 2011;79:830–838.
7. Marelius J, Kolmodin K, Feierberg I, Aqvist J. Q: a molecular dynamics program for free energy calculations and empirical valence bond simulations in biomolecular systems. J Mol Graph Model 1998;16:213–225, 261.
8. Dehouck Y, Kwasigroch JM, Rooman M, Gilis D. BeAtMuSiC: prediction of changes in protein-protein binding affinity on mutations. Nucleic Acids Res 2013;41(Web Server issue):W333–W339.
9. Douglas EV, Pires DBA, Blundell TL. mCSM: predicting the effects of mutations in proteins using graph-based signatures. Bioinformatics 2014;30:335–342.
10. de Groot BL, van Aalten DM, Scheek RM, Amadei A, Vriend G, Berendsen HJ. Prediction of protein conformational freedom from distance constraints. Proteins 1997;29:240–251.
11. Benedix A, Becker CM, de Groot BL, Caflisch A, Bockmann RA. Predicting free energy changes using structural ensembles. Nat Methods 2009;6:3–4.
12. Guerois R, Nielsen JE, Serrano L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. J Mol Biol 2002;320:369–387.
13. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein-protein complexes. Proc Natl Acad Sci U S A 2002;99:14116–14121.
14. Lippow SM, Wittrup KD, Tidor B. Computational design of antibody-affinity improvement beyond in vivo maturation. Nat Biotechnol 2007;25:1171–1176.
15. Sivasubramanian A, Maynard JA, Gray JJ. Modeling the structure of mAb 14B7 bound to the anthrax protective antigen. Proteins 2008;70:218–230.
16. Kamisetty H, Ramanathan A, Bailey-Kellogg C, Langmead CJ. Accounting for conformational entropy in predicting binding free energies of protein-protein interactions. Proteins: Struct Funct Bioinformatics 2011;79:444–462.
17. Smith CA, Kortemme T. Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. J Mol Biol 2008;380:742–756.
18. Ackers GK, Smith FR. Effects of site-specific amino acid modification on protein interactions and biological function. Annu Rev Biochem 1985;54:597–629.
19. Moretti R, Fleishman SJ, Agius R, Torchala M, Bates PA, Kastritis PL, Rodrigues JP, Trellet M, Bonvin AM, Cui M, Rooman M, Gillis

D, Dehouck Y, Moal I, Romero-Durana M, Perez-Cano L, Pallara C, Jimenez B, Fernandez-Recio J, Flores S, Pacella M, Kilambi KP, Gray JJ, Popov P, Grudinin S, Esquivel-Rodriguez J, Kihara D, Zhao N, Korkin D, Zhu X, Demerdash ON, Mitchell JC, Kanamori E, Tsuchiya Y, Nakamura H, Lee H, Park H, Seok C, Sarmiento J, Liang S, Teraguchi S, Standley DM, Shimoyama H, Terashi G, Takeda-Shitaka M, Iwadate M, Umeyama H, Beglov D, Hall DR, Kozakov D, Vajda S, Pierce BG, Hwang H, Vreven T, Weng Z, Huang Y, Li H, Yang X, Ji X, Liu S, Xiao Y, Zacharias M, Qin S, Zhou HX, Huang SY, Zou X, Velankar S, Janin J, Wodak SJ, Baker D. Community-wide evaluation of methods for predicting the effect of mutations on protein-protein interactions. Proteins 2013;81: 1980–1987.

20. Ishizuka J, Stewart-Jones GB, van der Merwe A, Bell JI, McMichael AJ, Jones EY. The structural dynamics and energetics of an immunodominant T cell receptor are programmed by its Vbeta domain. Immunity 2008;28:171–182.
21. Keeble AH, Joachimiak LA, Mate MJ, Meenan N, Kirkpatrick N, Baker D, Kleanthous C. Experimental and computational analyses of the energetic basis for dual recognition of immunity proteins by colicin endonucleases. J Mol Biol 2008;379:745–759.
22. Charles D. Schweiters GMC. Internal coordinates for molecular dynamics and minimization in structure determination and refinement. J Magn Reson 2001;152:288–302.
23. Flores SC, Bernauer J, Shin S, Zhou R, Huang X. Multiscale modeling of macromolecular biosystems. Brief Bioinform 2012;13:395–405.
24. Flores S, Zemora G, Waldsich C. Insights into diseases of human telomerase from dynamical modeling. Pac Symp Biocomput 2013; 18:200–211.
25. Flores SC. Fast fitting to low resolution density maps: elucidating large-scale motions of the ribosome. Nucleic Acids Res 2014;42:e9.
26. Jorgensen WL, TiradoRives J. Monte Carlo vs molecular dynamics for conformational sampling. J Phys Chem-Us 1996;100:14508–14513.
27. Maiorov V, Abagyan R. A new method for modeling large-scale rearrangements of protein domains. Proteins 1997;27:410–424.
28. Sim AY, Levitt M, Minary P. Modeling and design by hierarchical natural moves. Proc Natl Acad Sci U S A 2012;109:2890–2895.
29. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules (vol 117, pg 5179, 1995). J Am Chem Soc 1996;118:2309–2309.
30. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25:1157–1174.
31. Rocchia W, Alexov E, Honig B. Extending the applicability of the nonlinear Poisson-Boltzmann equation: multiple dielectric constants and multivalent ions. J Phys Chem B 2001;105:6507–6514.
32. Zwanzig RW. High-temperature equation of state by a perturbation method. I. Nonpolar gases. J Chem Phys 1954;22:1420–1426.
33. Jorgensen WL. Free-energy calculations—a breakthrough for modeling organic-chemistry in solution. Acc Chem Res 1989;22:184–189.
34. Kollman P. Free-energy calculations—Applications to chemical and biochemical phenomena. Chem Rev 1993;93:2395–2417.
35. Brandsdal BO, Smalas AO. Evaluation of protein-protein association energies by free energy perturbation calculations. Protein Eng 2000; 13:239–245.
36. Kortemme T, Baker D. A simple physical model for binding energy hot spots in protein-protein complexes. Proc Natl Acad Sci U S A 2002;99:14116–14121.
37. Dehouck Y, Gilis D, Rooman M. A new generation of statistical potentials for proteins. Biophys J 2006;90:4010–4017.
38. Abagyan R, Totrov M, Kuznetzov D. ICM—a new method for protein modeling and design: applications to docking and structure prediction from the distorted native conformation. J Comput Chem 1994;15:488–506.
39. Moal IH, Fernandez-Recio J. Intermolecular contact potentials for protein-protein interactions extracted from binding free energy changes upon mutation. J Chem Theory Comput 2013;9:3715–3727.
40. Brandsdal BO, Aqvist J, Smalas AO. Computational analysis of binding of P1 variants to trypsin. Protein Sci 2001;10:1584–1595.
41. Almlof M, Aqvist J, Smalas AO, Brandsdal BO. Probing the effect of point mutations at protein-protein interfaces with free energy calculations. Biophys J 2006;90:433–442.
42. Flores SC, Sherman MA, Bruns CM, Eastman P, Altman RB. Fast Flexible modeling of RNA structure using internal coordinates. IEEE ACM Trans Comput Biol Bioinformatics 2011;8:1247–1257.
43. The FoldX manual, http://foldx.crg.es/manual3.jsp.
44. Moal IH, Fernandez-Recio J. SKEMPI: a structural kinetic and energetic database of mutant protein interactions and its use in empirical models. Bioinformatics 2012;28:2600–2607.
45. Li M, Petukh M, Alexov E, Panchenko AR. Predicting the impact of missense mutations on protein–protein binding affinity. J Chem Theory Comput 2014;10:1770–1780.
46. Krowarsch D, Dadlez M, Buczek O, Krokoszynska I, Smalas AO, Otlewski J. Interscaffolding additivity: binding of P1 variants of bovine pancreatic trypsin inhibitor to four serine proteases. J Mol Biol 1999;289:175–186.
47. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 2008;4:435–447.
48. Coutsias EA, Seok C, Jacobson MP, Dill KA. A kinematic view of loop closure. J Comput Chem 2004;25:510–528.