Supplementary Discussion

Comparison with other methods

We compared our Eris server with other available online prediction servers for all the small-to-large mutations ($\Delta n_\chi < 0$, see below) in our dataset. As shown in Supplementary Table 1, all tested servers show impressive prediction accuracy; the correlations range from 0.45 to 0.66 for these mutants. The Eris flexible-backbone protocol outperforms all other servers mainly due to its ability to model backbone conformational changes induced by mutations.

Comparison of fixed- and flexible-backbone modeling

In order to assess the advantage of backbone flexibility modeling in an unbiased manner, we divided all the single mutations into three classes based on the change of number of side-chain χ angles (Δn_χ) of the mutation and compared the performance of the fixed- and flexible-backbone methods on these three classes separately (Fig. 1b). The mutations with $\Delta n_\chi < 0$ are associated with large-to-small mutations, therefore the protein backbone is expected to be minimally altered; those with $\Delta n_\chi >= 0$ correspond to mutation to residues of the same or larger sizes and backbone adjustment is expected if the mutation site is buried. We found that the fixed and flexible backbone prediction methods worked equally well for $\Delta n_\chi < 0$. However, the flexible-backbone $\Delta\Delta G$ prediction correlated better with experiments for $\Delta n_\chi > 0$ cases, due to its ability to resolve possible side-chain clashes. For small-to-large mutations ($\Delta n_\chi > 0$) we examined the solvent accessibility of the mutant side chains and found that about 68% of the mutated residues were buried or partially buried (with solvent accessibility less than 50%). Therefore, in this case, modeling backbone-flexibility is essential for the accurate $\Delta\Delta G$ calculations (Supplementary Table 2).

Backbone refinement

Our method for evaluating $\Delta\Delta G$ values depends on the protein's three-dimensional structure. When a high-resolution structure is not available, the computed stability changes could significantly deviate from the experimental values. For example, for α spectrin domain R16, when a nuclear-magnetic-resonance (NMR) derived structure (PDB ID: 1aj3) was used in the fixed-backbone calculations, the correlation coefficient between the computed and experimental values was 0.66 for surface mutations but only 0.28 for mutations in the protein core[1]. The overall correlation was only 0.09 (Supplementary Fig. 1a). When we computed the stability changes using a high-resolution x-ray structure (PDB ID: 1u5p), the correlation coefficient increased to 0.75. Since core residues have more contacts than surface residues, more accurate structural information near these sites is required to evaluate the energy contributions. Lack of accurate structural data or imperfections in protein structure models can be compensated by the pre-relaxation step (Supplementary Methods) in Eris. Structural pre-relaxation can be utilized for NMR structure refinement, which often has limited resolution of side-chain packing[2, 3]. To demonstrate this utility of the Eris server, we performed a backbone optimization of the NMR-derived structure. Using the pre-relaxed structure, the correlation between the

recalculated and experimental $\Delta\Delta G$s was dramatically improved (r = 0.70). Such a pre-relaxation step significantly improved the overall correlation coefficient from 0.09 to 0.69, as shown in Supplementary Figure 1b. We found that the root-mean-square deviation between pre-relaxed backbone structure and the original NMR-derived structure to be only 0.7 Å. Such a small adjustment was sufficient to improve side-chain packing of protein core, which consequently lead to more accurate $\Delta\Delta G$ evaluations.

Reference List

1. Scott,K.A., Randles,L.G., & Clarke,J. The folding of spectrin domains II: phi-value analysis of R16. *J. Mol. Biol.* **344**, 207-221 (2004).

2. Kuhlman,B. & Baker,D. Native protein sequences are close to optimal for their structures. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 10383-10388 (2000).

3. Ding,F. & Dokholyan,N.V. Emergence of protein fold families through rational design. *PLoS. Comput. Biol.* **2**, e85 (2006).

Supplementary Table 2.

List of the $\Delta n_\chi > 0$ class mutants, the relative solvent accessibilities (RSA) of the mutated residues, the experimental measured $\Delta\Delta G$, and the evaluated $\Delta\Delta G$ values by various servers. Twenty out of the 31 mutants are buried (solvent accessibility less then 50%) and are highlighted in the table.

PDB ID	Mutation	RSA (%)	Measured $\Delta\Delta G$ (kcal/mol)	$\Delta\Delta G$ predicted by servers (kcal/mol)					
				Eris-flex	Eris-fix	FoldX	Dmutant	I-Mutant	MuPro
1bvc	A 15 L	79.3	-0.20	-0.08 ± 0.6[#]	0.64 ± 0.9	-0.08	-1.16	0.54	0.81
1bvc	A 125 L	100.9	-0.40	1.19 ± 0.7	1.77 ± 0.8	0.18	-0.52	0.48	0.10
1bvc	A 144 L	76	-0.10	-0.98 ± 0.7	0.08 ± 0.7	-0.13	-0.73	0.53	0.31
1bvc	A 130 L	0	2.30	5.43 ± 1.1	*	-2	-0.34	0.56	0.80
1bvc	A 130 K	0	3.70	11.49 ± 0.9	*	0.82	0.5	1.49	1.30
1stn	A 12 V	5.7	1.10	-0.71 ± 0.5	0.57 ± 0.7	-2.31	-1.38	0.22	0.38
1stn	A 17 V	3.9	1.90	-0.31 ± 0.6	-1.58 ± 0.6	-2.6	-1.14	1.36	0.21
1stn	I 18 M	30.9	0.60	-0.14 ± 0.6	0.24 ± 0.7	1.46	0.52	0.22	0.30
1stn	Y 27 E	14	5.00	5.86 ± 0.5	6.24 ± 0.6	4.46	2.98	1.01	0.95
1stn	Y 27 K	14	3.90	8.44 ± 0.6	7.25 ± 0.6	2.58	2.76	0.85	1.45
1stn	Y 27 M	14	2.00	4.23 ± 0.5	3.72 ± 0.6	1.25	1.37	0.37	0.70
1stn	Y 27 Q	14	3.20	7.24 ± 0.5	6.58 ± 0.5	2.45	2.75	1.23	1.16
1stn	Y 27 R	14	2.90	6.07 ± 0.6	6.76 ± 0.6	2.98	2.24	1.36	1.10
1stn	G 29 F	102.9	1.25	3.93 ± 0.5	2.72 ± 0.6	0.28	-0.73	0.97	0.17
1stn	G 50 F	83.6	0.60	2.56 ± 0.5	1.14 ± 0.5	0.4	-0.61	0.45	-0.40
1stn	G 50 V	83.6	1.10	-1.45 ± 0.5	-2.27 ± 0.5	-0.01	0.01	0.81	-0.01
1stn	A 58 V	0	2.80	2.09 ± 0.6	2.97 ± 0.7	-1.82	-0.92	0.76	0.20
1stn	S 59 F	28.9	0.00	-5.27 ± 0.5	-3.48 ± 0.6	-2.13	-1.98	-0.96	0.17
1stn	A 60 F	78.8	0.80	2.12 ± 0.5	1.46 ± 0.5	0.06	-0.99	0	1.37
1stn	A 60 V	78.8	2.80	1.01 ± 0.6	-0.12 ± 0.6	0.32	-0.32	0.83	2.80
1stn	V 66 L	0.3	0.03	-0.58 ± 0.6	-1.31 ± 0.5	-1.01	-1.65	-0.32	-0.70
1stn	V 66 M	0.3	0.20	0.71 ± 0.4	0.49 ± 0.5	-0.31	-0.55	0.54	-0.17
1stn	A 69 T	3	3.08	6.67 ± 0.4	6.69 ± 0.7	-1.5	0.09	1.68	1.06
1stn	A 90 S	0	2.00	1.37 ± 0.5	2.17 ± 0.5	0.55	0.76	2.31	2.40
1stn	A 94 V	3.3	1.30	0.35 ± 0.6	2.67 ± 0.6	-2.76	-1.11	1.08	0.56
1stn	A 102 F	25.5	1.90	4.27 ± 0.5	4.80 ± 0.5	-0.49	-2.33	0.44	0.44
1stn	A 112 F	20.7	1.45	1.76 ± 0.4	1.44 ± 0.5	-0.49	-1.38	0.52	0.65
1stn	A 112 V	20.7	1.50	3.15 ± 0.5	3.38 ± 0.6	-1.13	-0.60	0.25	0.62
1stn	S 128 F	14.1	1.00	-3.03 ± 0.5	-1.58 ± 0.6	-1.15	-0.62	-1.06	-0.18
1stn	A 130 V	75.7	1.20	-0.18 ± 0.4	-0.97 ± 0.6	0.28	0.00	-0.19	0.61
2ci2	T 58 D	96.2	0.04	0.56 ± 0.4	0.31 ± 0.4	0.31	0.08	-0.07	0.06

The uncertainty was estimated from the standard deviation from 20 simulations (see Supplementary Methods).
* Eris fixed-backbone predictions for these mutations are not available because severe atomic clashes are detected.

Supplementary Table 1.

Comparison of Eris $\Delta\Delta G$ calculations with those from other servers. The $\Delta\Delta G$ calculations for all small-to-large mutants (Supplementary Table 2) were submitted to known publicly accessible servers and the resulting correlation coefficients were compared with those from Eris calculations.

Server	Correlation
Eris-flexible backbone	0.75
Eris-fixed backbone	0.71
FoldX[1]	0.45
Dmutant[2]	0.66
I-Mutant[3]	0.57
MutPro[4]	0.57

Reference List

1. Guerois,R., Nielsen,J.E., & Serrano,L. Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *Journal of Molecular Biology* **320**, 369-387 (2002).

2. Zhou,H.Y. & Zhou,Y.Q. Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction (vol 11, pg 2714, 2002). *Protein Science* **12**, 2121 (2003).

3. Capriotti,E., Fariselli,P., & Casadio,R. I-Mutant2.0: predicting stability changes upon mutation from the protein sequence or structure. *Nucleic Acids Research* **33**, W306-W310 (2005).

4. Cheng,J.L., Randall,A., & Baldi,P. Prediction of protein stability changes for single-site mutations using support vector machines. *Proteins-Structure Function and Bioinformatics* **62**, 1125-1132 (2006).

Supplementary Methods

Medusa force field

We used a united atom model and express the free energy of the protein as a weighted sum of van der Waals (VDW) interaction, solvation energy, hydrogen bonding, and backbone-dependent statistical energy for any given amino acid and rotamer state. We assume the unfolded state to be unstructured and fully-solvated and, thus, the reference energy of the unfolded proteins are the linear sum over all the amino acids, which are also parameterized in the Medusa force field. More specifically, the ΔG of a protein was evaluated as,

$$\begin{aligned}\Delta G = & W_{vdw\_attr} E_{vdw\_attr} + W_{vdw\_rep} E_{vdw\_rep} + W_{solv} E_{solv} \\ & + W_{bb\_hbond} E_{bb\_hbond} + W_{sc\_hbond} E_{sc\_hbond} + W_{bb\_sc\_hbond} E_{bb\_sc\_hbond} \\ & + W_{aa|\phi,\psi} E_{aa|\phi,\psi} + W_{rot|\phi,\psi,aa} E_{rot|\phi,\psi,aa} - E_{ref}\end{aligned} \quad (1)$$

Here E_{vdw_attr}, E_{vdw_rep} are the attractive and repulsive part of the VDW interaction, respectively; E_{solv} is the solvation energy; E_{bb_hbond}, E_{sc_hbond} and $E_{bb\_sc\_hbond}$ are the hydrogen bonding energies among backbones, among sidechains, and between backbones and sidechains, respectively. $E_{aa|\phi,\psi}$ and $E_{rot|\phi,\psi,aa}$ correspond to the internal energy for an amino acid (*aa*) in its rotamer state (*rot*) given the backbone dihedrals, phi (ϕ) and psi (ψ). E_{ref} is the reference energy of the unfolded state. The weight coefficients were trained on 34 high-resolution X-ray protein structures so that occurrences of native amino acids sequences were maximized for all proteins.

Fixed-backbone method

In a fixed backbone calculation, we started from the native structure and mutated the native residues. After the substitution, the side chain dihedrals of all residues were minimized using a Monte Carlo (MC) simulated annealing procedure. To search the dihedral space faster and more effectively, all the rotamer angles were divided into discrete rotamer states using the backbone-dependent rotamer library. Due to the stochastic nature of the method, the simulation was repeated 20 times from randomly assigned initial side chain rotamers and the final free energy was taken to be the average over these 20 runs. The $\Delta\Delta G$ value was calculated as the difference between the ΔG of the mutant and the wild type as $\Delta\Delta G = \Delta G^{\text{mutant}} - \Delta G^{\text{wild type}}$.

Flexible-backbone method

In flexible-backbone stability prediction, we performed similar calculations as a fixed-backbone one except that we allowed the backbone dihedral angles to be relaxed if the acceptance rate during the MC loop was below a pre-defined threshold. The threshold we used in Eris protocol was 0.05; such a low acceptance rate usually indicates strains on the backbone by the mutated residue. The relaxation was done by a conjugate-gradient minimization over all backbone dihedral angles at the end of the MC loop.

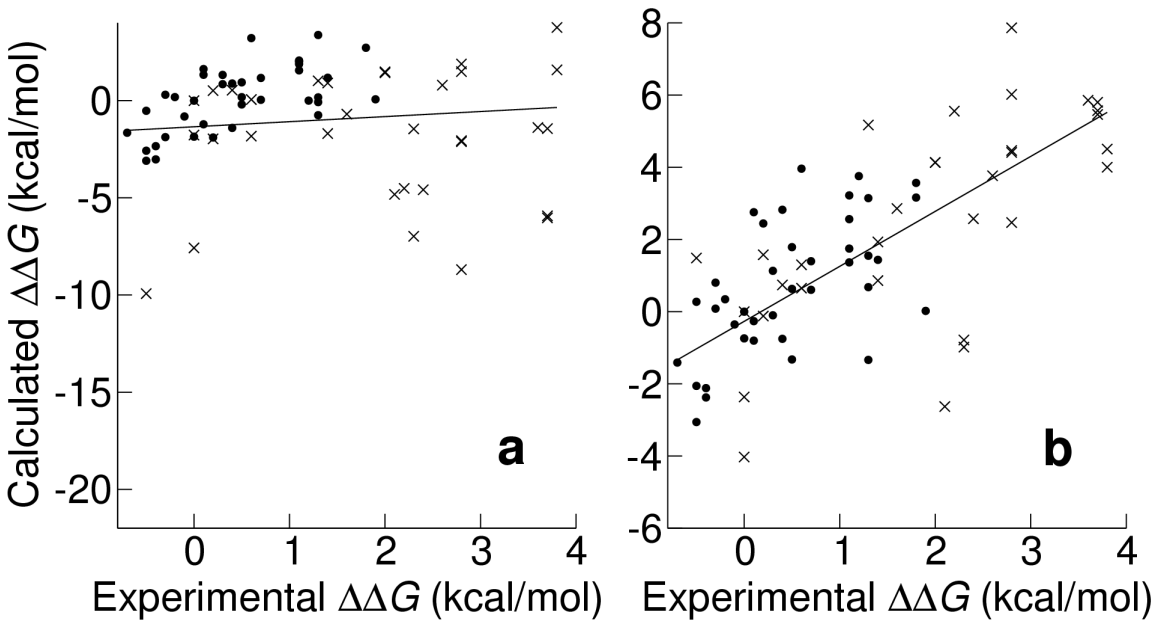
Pre-relaxation

The backbone pre-relaxation was implemented by performing the flexible-backbone calculation on the wide-type protein structure for 20 times. The redesigned structure with the lowest free energy was selected as the pre-relaxed structure.

Dataset

We performed $\Delta\Delta G$ calculations on 595 mutants and compared the computational predictions with experimental measurements. The mutants are of six proteins, which include FK506 binding protein (PDB ID: 1fkj), apomyoglobin (PDB ID: 1bvc), staphylococcal nuclease (PDB ID: 1stn), chymotrypsin inhibitor 2 (PDB ID: 2ci2), and α spectrin domain repeat 16 (PDB ID: 1u5p).

Supplementary Figure 1.



Eris-aided structure pre-relaxation relieves backbone strains of the α-spectrin initial structure. (**a**) The scatter plot of the calculated and experimental $\Delta\Delta G$ values obtained using NMR-determined structure (PDB ID 1aj3). The Pearson correlation coefficient between the computed and experimental $\Delta\Delta G$ values is 0.66 for surface mutants (•) and only 0.28 for core mutants (×). **(b)** The scatter plot of the calculated and experimental $\Delta\Delta G$ values obtained using the Eris pre-relaxed structure. Significant correlations are obtains for both surface (•; r = 0.60) and core mutations (×; r = 0.70). The lines are the linear regression fits.