Supplementary text

Details of the energy model for the folded state

The charging energy, the first term in equation (4), is calculated as follows. The intrinsic pK_a [1] of residue i, $pK_a^{int}(i)$, is the pK_a of that residue in the studied protein, considering no interaction with other ionizable residues. Let P_n be an uncharged state with the protein having all its residues in the neutral state, including residue i, and P_c is a charged state differing from P_n in that residue i is in its charged state. At a given solution pH, the charging energy ΔG_{cha} of residue i is:

$$\Delta G_{cha}(i) = \Delta G_{P_n \to P_c} = \gamma(i) RT(\ln 10)\left(pH - pK_a^{int}(i)\right) \qquad (S1)$$

where $\gamma(i)$ is +1 for positively charged residues, and -1 for negatively charged ones. On the other hand, the ionization of residue i can be related to that of the appropiate model compound[2] using the thermodynamic cycle in figure S1, where M_n and M_c represent, respectively, states where the studied residue i has been separated from the protein to become the model compound in the neutral or charged state. Therefore, the pK_a^{int} can be expressed as a function of the pK_a of the model compounds:

$$pK_a^{int}(i) = pK_a^{mod}(i) - \gamma(i)\frac{\Delta G_{M_c \to P_c} + \Delta G_{P_n \to M_n}}{RT\ln 10} \qquad (S2)$$

To calculate $\Delta G_{M_c \to P_c}$ and $\Delta G_{P_n \to M_n}$ we assume the electrostatic effects are the relevant ones and calculate them from $G_{ele}^{P_n}$, $G_{ele}^{M_n}$, $G_{ele}^{M_c}$, $G_{ele}^{P_c}$. The main non-electrostatic contributions to pK_a^{int}, which cannot be ignored, are represented through the experimental pK_a^{mod} values and the permittivity constants of protein and solvent. After reorganization:

$$\Delta G_{M_c \to P_c} + \Delta G_{P_n \to M_n} \approx \left(G_{ele}^{P_c} - G_{ele}^{P_n}\right) - \left(G_{ele}^{M_c} - G_{ele}^{M_n}\right) \qquad (S3)$$

According to Antosiewic *et al.*[3]:

$$G_{ele}^{P_c} - G_{ele}^{P_n} = \sum_{i \in A}\left(q_i'\phi_i' - q_i\phi_i\right) - \frac{1}{2}\sum_{i \in B}\left(q_i'\phi_i' - q_i\phi_i\right) \qquad (S4)$$

where A represents all the atoms in the protein (including those of the studied residue, which form subset B); ϕ_i represents the electrostatic potential at atom i in state P_n, and ϕ_i' represents the same in state P_c; q_i represent the point charges in state P_n and q_i' those in state P_c. Similarly for the model residue calculations (based on ref. 3, eq. 8):

$$G_{ele}^{M_c} - G_{ele}^{M_n} = \sum_{i \in B}\left(q_i'\phi_i' - q_i\phi_i\right) - \frac{1}{2}\sum_{i \in B}\left(q_i'\phi_i' - q_i\phi_i\right) \qquad (S5)$$

In our approach, differing from ref. 3, the whole model residue is the ionizable group, since the charge of the proton may be distributed among all the atoms of the residue; therefore, B represents all the atoms of the model residue. ϕ_i and q_i correspond to state M_n; and $\phi_i^{'}$ and $q_i^{'}$ correspond to state M_c.

The effective interaction energy, the second term of equation (4), includes all electrostatic interactions among the charged ionizable residues in the protein. Again, following ref. 3, and since the partial charges of the neutral and the charged state for each residue are in the same atomic points:

$$\Delta G_{ele,int}(i,j) = \sum_{x \in res_j} \left(q_x^{'} - q_x\right)\left(\phi_i^{'} - \phi_i\right) \qquad (S6)$$

where x is an index of the atoms in one of the interacting residues (residue j), and $q_x^{'}$ is the charge at atom x in the residue charged state, and q_x the charge in the residue neutral state. $\phi_i^{'}$ represents the electrostatic potential created by residue i charged state at the given atom positions, and ϕ_i represents the potential created by residue i neutral state. As in ref. 3, we assume $\Delta G_{ele,int}(i,j) = \Delta G_{ele,int}(j,i)$. Though this is true in the Poisson-Boltzmann model, it is only an approximation in our model, because $\phi_i^{'}$ and ϕ_i are not the potential energy at the given atom positions, but the potential energy *due to residue* i, which misses the effect of residue j charges at those atom positions both directly and indirectly through the induced charges in the dielectric border. To deal with a single number in the calculations we only consider values of $\Delta G_{ele,int}(i,j)$ with $i < j$ (that is, residue i is first in the protein sequence).

The electrostatic interaction energy of a residue i in the folded state is the sum of the interaction energies of that residue with any other ionizable residue j:

$$\Delta G_{ele,int}(i) = \sum_{j \neq i, j \in R} \Delta G_{ele,int}(i,j) \qquad (S7)$$

For a given protonation configuration $\vec{p}$, the interaction energy between two residues i, j is $p(i)p(j)\Delta G_{ele,int}(i,j)$, and then the residue interaction energy is:

$$\Delta G_{ele,int}(i,\vec{p}) = \sum_{j \neq i, j \in R} p(i)p(j)\Delta G_{ele,int}(i,j) \qquad (S8)$$

and for the Boltzmann-weighted average of the protonation configurations:

$$\Delta G_{ele,int}(i) = \sum_{\vec{p}} w(\vec{p}) \sum_{j \neq i, j \in R} p(i)p(j)\Delta G_{ele,int}(i,j) \qquad (S9)$$

where $w(\vec{p})$ is the weight given to configuration $\vec{p}$. Equation (S9) can be rearranged to give:

$$\Delta G_{ele,int}(i) = \sum_{j \neq i, j \in R} \Delta G_{ele,int}(i,j) \sum_{\vec{p}} w(\vec{p})p(i)p(j) \qquad (S10)$$

Since $p(i)p(j)$ is 1 only if both i and j are charged (0 otherwise), the term $\sum_{\vec{p}} w(\vec{p})p(i)p(j)$ is the Boltzmann-weighted fraction of protonation configurations $\vec{p}$ that have both i and j charged. This fraction can be obtained from the Metropolis Monte Carlo sampling used to calculate G^{F}_{ele}. Renaming this fraction as $x(i,j)$, the interaction energy for residue i can be written as:

$$\Delta G_{ele,int}(i) = \sum_{j \neq i, j \in R} x(i,j)\Delta G_{ele,int}(i,j) \qquad (S11)$$

Calculation of electrostatic potentials and interaction energies in the folded state

For each ionizable residue in each protein variant four potential maps corresponding to the P_n, M_n, M_c and P_c states in figure S1 have to be calculated. These potential maps were obtained solving the Poisson-Boltzmann equation using finite differences as implemented in Delphi (see parameters in Table S4). The relative coordinates of the model residue are the same as those when embedded in the protein. To avoid artifacts in the calculations (see discussion on self-energy in ref. 2) we follow the approximation of using the same conformations for the model residue in states M_n and M_c, and for the protein in states P_n, P_c. Using the same solvent-molecule boundary in each state (charged or neutral) also simplifies calculations, as explained above when dealing for radii and charges of atoms in the folded state.

For the P_n and P_c states the processed folded structures described above were used. Then, the difference between the neutral and charged states in the thermodynamic cycle were exclusively modeled through the partial charges assigned to the atoms. Since the protein without the model residue in states M_n and M_c is the same in both cases, we can ignore it for the calculations. The charges used to model each state are: for M_n or M_c the partial atomic charges for the studied residue type in the neutral or charged state, respectively; for P_n or P_c: the partial atomic charges for the studied residue type in the neutral or charged state, respectively. All other atoms in the protein are assigned null partial charges.

For each protein variant the grid center is that of the parallelepiped containing it, with the parallelepiped faces parallel to the XY, YZ and XZ planes. Then, we create a cube centered on that point, so that the protein volume in it occupies, at most, 80% of the grid (see Table S4). Once the potentials for the four states are obtained, they are introduced in equations (S4) and (S5) using different sets of charges. For protein states —eq (S4)— q_i represents the partial charges corresponding to the neutral state for *all* atoms in the protein, including the studied residue; and q_i' represents the partial charges corresponding to the charged state *only* for the atoms of the studied residue, all the other atoms in the protein being considered in their neutral

state. For model states —eq. (S5)— q_i represents the partial charges of the residue atoms in the neutral state, and $q_i^{'}$ those for the charged state. For calculation of $\Delta G_{ele,int}(i,j)$ using eq. (S6), the calculated potentials are again used, and for $q_x^{'}$ and q_x we used the charged state and neutral state set of atomic charges for residue j.

Sampling of the protonation configurations in the folded state

Each scan has as many Monte Carlo steps as the number of possible transitions (there are as many moves as ionizable residues, plus the number of strongly coupled groups of residues, either couples or triplets), effectively varying from one protein to another, since they will have different numbers of possible transitions. This is similar to the method of Beroza *et al.*[4], though they try all possible transitions in each scan, while we do the same number of transitions per scan, but the transition is selected randomly each time. Each Monte Carlo step consists in selecting a transition in the protonation configuration space of the protein and deciding whether to accept it or to keep the old configuration. Transitions can affect one, two, or three ionizable residues, and the actual number of residues that change their protonation state is based on the probabilities in Table S5 (single_move_probability, double_move_probability, and triple_move_probability). When no double or triple transition is present, their probabilities are set to zero and the other probabilities normalized to add up to 1. For transitions affecting only one residue, any ionizable residue is eligible. For 2 or 3 residues, the transitions allowed depend on whether the ionizable residues interact strongly (as explained in ref. 4). Several residues interact strongly if the only way to change between 2 low energy protonation configurations of those residues is by flipping the state of several of those residues because intermediate steps requiring single flips have high energies. Strong interactions hinder a Metropolis sampling that only uses single-residue transitions. Our method classifies pairs or triplets of residues as strongly interacting if there are at least 2 protonation configurations, $\vec{p}_a$ and $\vec{p}_b$, of those residues that differ in the protonation state of at least one residue (but not all) and whose difference in interaction energy is above a threshold as the following equation shows:

$$\left|E_{int}(\vec{p}_a) - E_{int}(\vec{p}_b)\right| > E_{threshold} \qquad (S12)$$

In the case of triplets, it is besides required that at least a pair of their residues interact strongly. The energy of the interacting residues is approximated by taking, from eq. (4), only those terms referring to the affected residues:

$$E_{int}(\vec{p}) = \sum_{i \in R} p(i)\Delta G_{cha}(i) + \sum_{i \in R} \sum_{j \in R, i<j} p(i)p(j)\Delta G_{ele,int}(i,j) \qquad (S13)$$

with R being the set of strongly interacting residues being considered and $\vec{p}$ is a protonation configuration for those residues. The value of the $E_{threshold}$ parameter and the rest of parameters used in the Metropolis sampling appear in Table S5.

Graphs and data analysis

Graphs and data analysis, including least-squares regression analysis, were done within the IPython environment (http://ipython.org) using pandas 0.12 (http://pandas.pydata.org/), scipy 0.12 (http://scipy.org/scipylib/index.html) and matplotlib 1.3.0 (http://matplotlib.org/).

Effect of the calculated interaction corrections of the ensemble unfolded models on the predicted stabilizing energies

Under the *Simple* model the electrostatic folding energy of a protein is:

$$G_{ele}^{F} - G_{ele,simple}^{U} \qquad (S14)$$

with the folded and unfolded state energies given by equations (3) and (9) respectively.

When considering the unfolded ensemble models (*Minimum energy* model, *Average energy* model and *Boltzmann-weighted energy* model), it can be seen that the unfolded energy is the sum of two terms, the first one corresponding to the first term in equation (11), which is the same as the *Simple* model unfolded energy, and the second term corresponding to the minimum, average or Boltzmann-weighted average (depending on the model) of the second term in equation (12):

$$G_{ele}^{U} = G_{ele,simple}^{U} + G_{ele,int}^{U} \qquad (S15)$$

Based on that, we can express the electrostatic stabilizing energy under those models in relation to the electrostatic stabilizing energy in the *Simple* model as

$$\Delta\Delta G = \left(G_{ele}^{F,mut} - G_{ele}^{U,mut}\right) - \left(G_{ele}^{F,wt} - G_{ele}^{U,wt}\right) = \Delta\Delta G_{simple} + \left(G_{ele,int}^{U,wt} - G_{ele,int}^{U,mut}\right) \qquad (S16)$$

That second term, which we call $\Delta\Delta G_{ele,int}^{U}$, represents the effect of the calculated unfolded interaction corrections in the predicted stabilizing energies. As shown in table S6, which presents the mean and standard deviation for $\left|\Delta\Delta G_{ele,int}^{U}\right|$ when considering the 80 mutations in the *complete set*, these values are in the same order as the stabilizing energies in the *Simple* model.

Comments on the precision achieved with the proposed models

Fig. S2 graphically shows how the achieved precision is much higher than the precision one would obtain using a random model for which the precision would correspond to the actual ratio of stabilizing mutants in the *comparable set* (0.23, since 13 out of the 56 mutants are stabilizing; a similar value of 0.25 would be obtained if the whole set of 80 mutations were considered). Notice that such low values correspond to the precisions obtained by placing the

threshold just at the last point of each graph, since that precision value is just the ratio of stabilizing mutants. The graphs also show that the models are relatively insensitive to the actual value for the classifying threshold: slightly moving the threshold line to the left or to the right in the figures would not significantly worsen the precision obtained by the models. Finally, because the graphs show the way each model orders the mutants by predicted stability change (most stabilizing-predicted mutants on the left, less stabilizing-predicted or destabilizing-predicted mutants on the right), the good behavior of the models is clear, since most experimentally stable mutants (black dots) appear towards the left part of the graph.

References

1 C. Tanford and J. G. Kirkwood, *J. Am. Chem. Soc.*, 1957, **79**, 5333–5339.

2 D. Bashford and M. Karplus, *Biochemistry*, 1990, **29**, 10219–10225.

3 J. Antosiewicz, J. M. Briggs, A. H. Elcock, M. K. Gilson and A. McCammon, *J. Comput. Chem.*, 1996, **17**, 1633–1644.

4 P. Beroza, D. R. Fredkin, M. Y. Okamura and G. Feher, *Proc. Natl. Acad. Sci. U. S. A.*, 1991, **88**, 5804–5808.

5 L. A. Campos, M. M. Garcia-Mira, R. Godoy-Ruiz, J. M. Sanchez-Ruiz and J. Sancho, *J. Mol. Biol.*, 2004, **344**, 223–237.

6 D. Perl, U. Mueller, U. Heinemann and F. X. Schmid, *Nat. Struct. Biol.*, 2000, **7**, 380–383.

7 M. Wunderlich and F. X. Schmid, *Protein Eng., Des. Sel.*, 2006, **19**. 355–358.

8 V. V. Loladze, B. Ibarra-Molero, J. M. Sanchez-Ruiz and G. I. Makhatadze, *Biochemistry*, 1999, **38**, 16419–16423.

9 J. M. Schwehm, C. A. Fitch, B. N. Dang, B. Garcia-Moreno and W. E. Stites , *Biochemistry*, 2003, **42**, 1118–1128.

10 M. N. Davies, C. P. Toseland, D. S. Moss and D. R. Flower, *BMC Biochem.*, 2006, **7**, 18.

Supplementary figures and tables

Figure S1. Thermodynamic cycle illustrating the charging of an ionizable residue in a protein. The example shows Glu36 of the cold shock protein from *Bacillus subtilis* (PDB code: 1csp). Only the atoms in that residue have partial charges, and only the partial charges that differ between the neutral and the charged state (in atoms $O^{\varepsilon 1}$, $O^{\varepsilon 2}$, C^{δ}) are shown. All other atoms in the protein have null partial charges. The states in the cycle are: M_n (model residue, neutral: atoms of the ionizable residue are in the neutral form); M_c (model residue, charged: atoms of the ionizable residue are in the charged form); P_n (protein with neutral residue: the whole protein, with all atoms in the residue in the neutral form, and all the other atoms with null partial charges); P_c (protein with charged residue: the whole protein, with all atoms in the residue in their charged form, and all the other atoms with null partial charges). All states in the cycle have continuous solvent and the same ionic strength.

Figure S2. Precision depending on the threshold chosen for the subset of 56 mutants in the *comparable data set.* Experimentally stabilizing mutants are shown in black, non-stabilizing ones in white. Each dot represents the *precision* (stabilizing success rate) obtained after placing the classifying threshold exactly at the predicted stabilizing energy for the given mutant (predicted energies are represented on the x-axis). $\Delta\Delta G_{\text{simple}}$ is the *Simple* model (A), $\Delta\Delta G_{\text{pred}}$ is the *Mixed reference* model (B), and $\Delta\Delta G_{\text{nat}}$ is the *Native only* model (C). Vertical lines indicate the actual threshold used to calculate the precision of the different models, as shown in Table 2.

Figure S1

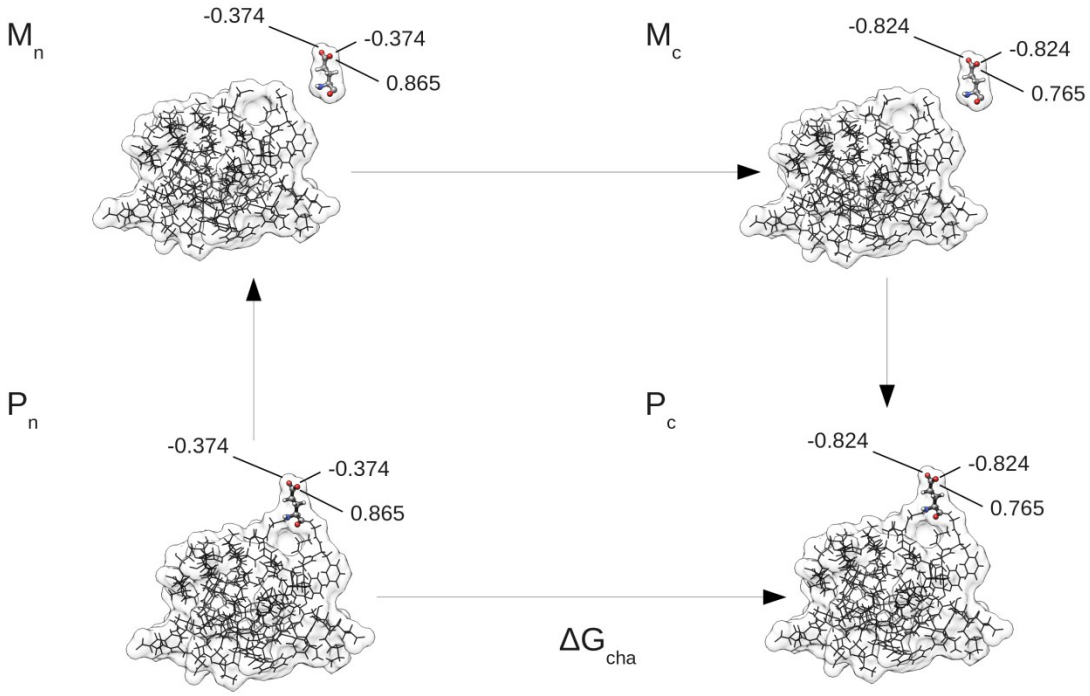
M n
-0.374
-0.374
0.865
M c
-0.824
-0.824
0.765
P n
-0.374
-0.374
0.865
P c
-0.824
-0.824
0.765
ΔG cha

Figure S2

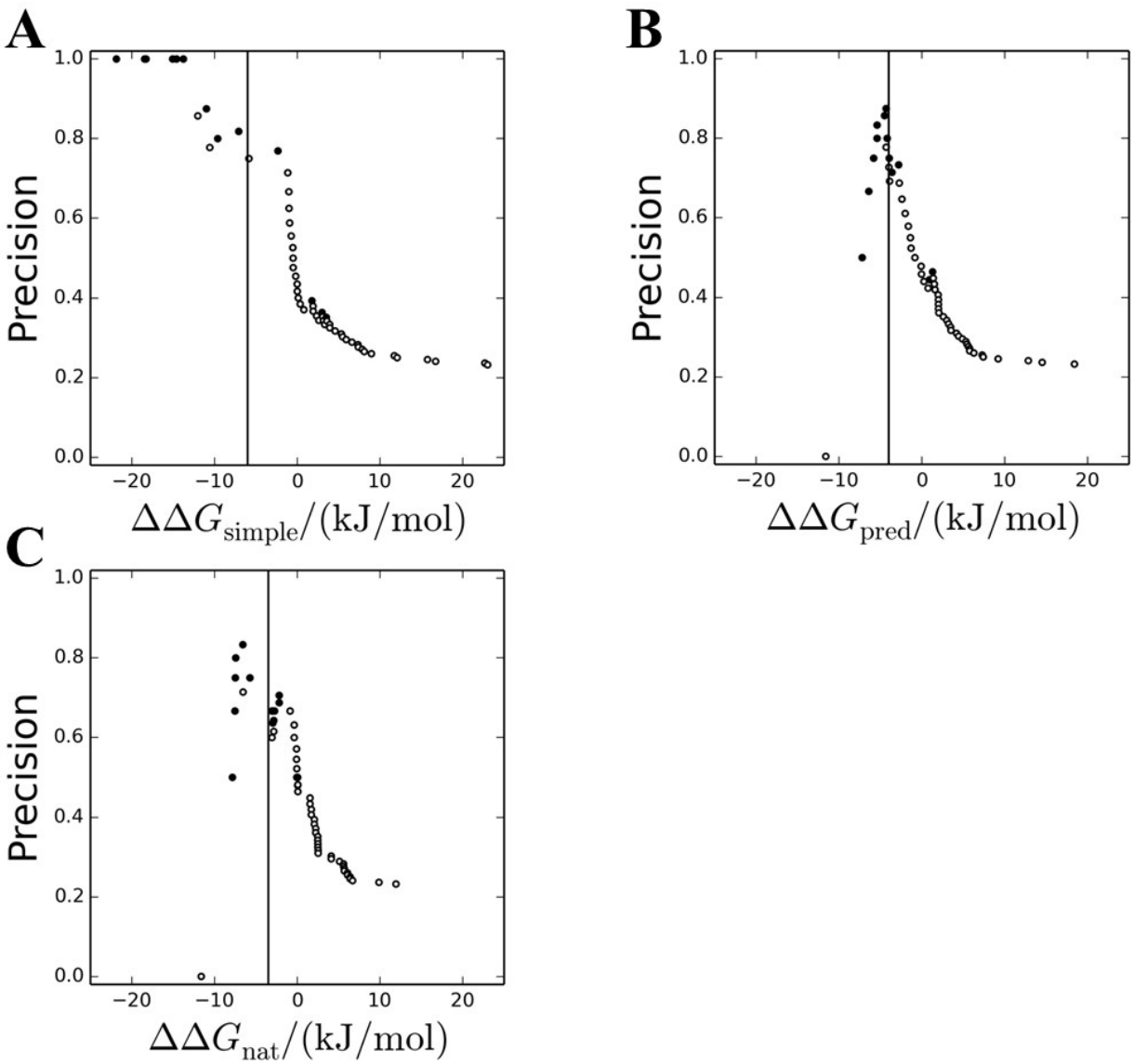
A
B
C
Precision
1.0
0.8
0.6
0.4
0.2
0.0
−20
−10
0
10
20
$\Delta\Delta G_{\mathrm{simple}}/(\mathrm{kJ/mol})$
$\Delta\Delta G_{\mathrm{pred}}/(\mathrm{kJ/mol})$
$\Delta\Delta G_{\mathrm{nat}}/(\mathrm{kJ/mol})$

Table S1. Experimental differential folding energies.

Protein	Mutant	$\Delta\Delta G$ $(kJmol^{-1})$	Protein	Mutant	$\Delta\Delta G$ $(kJmol^{-1})$
apoflavodoxin	E20K[d]	-6.7	SNase	E10K[h]	9.2
	E40K[d]	-9.3		E10Q[h]	5.4
	E61K[d]	-3.7		D19K[h]	1.3
	E72K[d]	-5.4		D19N[h]	0.84
	D65K[d]	1.0		E21K[h]	-4.2
	D75K[d]	-4.2		E21N[h]	-5.9
	D126K[d]	-4.2		K28E[h]	3.3
	D150K[d]	1.2		K28Q[h]	1.3
CspB-Bc	R3E[a, e]	11.5		K48E[h]	0.0
	Y15F[a, e]	0.2		K48Q[h]	0.4
	E46A[a, e]	0.9		E57K[h]	0.4
	R3E[b, e]	4.2		E57Q[h]	0.8
	Y15F[b, e]	-0.3		K63E[h]	6.3
	E46A[b, e]	0.5		K63Q[h]	4.2
	E12K[a, f]	1.8		K64E[h]	6.7
	E21K[a, f]	1.1		K64Q[h]	0.0
	H29E[a, f]	3.2		E67Q[h]	2.1
	E36K[a, f]	0.8		K70E[h]	1.7
	E46K[a, f]	2.7		K70Q[h]	1.3
	E50K[a, f]	2.1		E73K[h]	8.8
	R56E[a, f]	-3.2		E73Q[h]	2.9
	E12K[b, f]	1.2		E75K[h]	9.6
	E21K[b, f]	-0.7		E75Q[h]	2.9
	H29E[b, f]	2.0		D77K[h]	14.6
	E36K[b, f]	-3.3		D77N[h]	12.1
	E46K[b, f]	0.9		K78E[h]	3.3
	E50K[b, f]	-1.1		K78Q[h]	0.4
	R56E[b, f]	-1.0		K84E[h]	0.0
CspB-Bs	E3R[a, e]	-11.1		K84Q[h]	0.4
	E66L[a, e]	-8.8		K97E[h]	2.5
	E3R[b, e]	-3.2		K97Q[h]	1.7
	E66L[b, e]	-6.0		E122K[h]	11.3
ubiquitin	K6E[g]	-2.2		E122Q[h]	6.7
	K6Q[g]	-1.1		H124E[h]	-0.8
	K27Q[g]	8.0		H124Q[h]	-2.5
	K29Q[c, g]	6.2		K127E[h]	0.0
	K29N[c, g]	7.0		K127Q[h]	0.0
	R42E[g]	-6.8		K133Q[h]	2.1
	H68Q[g]	-2.3		E135Q[h]	2.9
	H68E[g]	-3.2			
	R72Q[g]	1.4			

[a] I=0.1 M. [b] I=2.1 M. [c] Energy of mutants K29Q and K29N are swapped from those in the original article because fig. 2 of that article and other values in the table in that article (C1/2) show those values were mistakenly swapped. [d] From ref. 5. [e] From ref. 6. [f] From ref. 7. [g] From ref. 8. [h] From ref. 9.

Table S2 Ionizable residues, their pK_a^{mod} (from ref. 10), and their protonation states with the corresponding *Amber* residue type and the point charge changes needed to obtain the non-standard protonation variants.

Ionizable residue[a]	pK_a^{mod}	**Protonation variants and their point charges**		
Asp	4.0	With $H^{\delta 2}$	0	ASP, adding 0.1 to C^{γ}, and 0.45 to $O^{\delta 1}$ and $O^{\delta 2}$.
		Without $H^{\delta 2}$	-1	ASP
Glu	4.4	With $H^{\epsilon 2}$	0	GLU, adding 0.1 to C^{δ}, and 0.45 to $O^{\epsilon 1}$ and $O^{\epsilon 2}$.
		Without $H^{\epsilon 2}$	-1	GLU
Cys	9.5	In a disulfide bond	0	CYX, point charges specified by Amber force field.
		With H^{γ}	0	CYS
		Without H^{γ}	-1	CYS, null charge in H^{γ}, and the rest of charge until a total of -1, is taken from S^{γ}.
His	6.3	With $H^{\delta 1}$	0	HID
		With $H^{\epsilon 2}$	0	HIE
		With $H^{\delta 1}$ and $H^{\epsilon 2}$	$+1$	Either HID adding $+1.0$ to $N^{\epsilon 2}$, or HIE adding $+1.0$ to $N^{\delta 1}$.
Lys	10.4	With $H^{\zeta 1}$, $H^{\zeta 2}$ and $H^{\zeta 3}$	$+1$	LYS
		With $H^{\zeta 1}$ and $H^{\zeta 2}$	0	LYS, adding -1 to N^{ζ}.
Tyr	10.0	With H^{ζ}	0	TYR
		Without H^{ζ}	-1	TYR, null charge in H^{η}, and the rest (up to -1), from O^{η}.
Arg	12.0	With 4 hydrogens H^{η}	$+1$	ARG
		With 3 hydrogens H^{η}	0	ARG, reduce 0.25 the charge in all H^{η} atoms.

[a]For His, the pK_a^{mod} corresponds to the protonation of either HID or HIE to become charged. N-terminal and C-terminal groups were not considered as ionizable in this work and are not shown. The pK_a^{mod} values are from ref. 10.

Table S3. Atoms from the mutated wild type variant residue used to place the charge of an ionizable residue in the mutant.

Wild type residue	Atom holding charge in mutant	Wild type residue	Atom holding charge in mutant
Asp	C^{γ}	Glu	C^{δ}
Arg	C^{ζ}	His	C^{γ}
Tyr	O^{η}	Lys	N^{ζ}
Cys	S^{γ}	Gly	C^{α}
Ala	C^{β}	Ser	O^{γ}
Thr	C^{β}	Leu	C^{γ}
Ile	C^{β}	Val	C^{β}
Asn	C^{γ}	Gln	C^{δ}
Trp	$C^{\zeta 3}$	Phe	C^{ζ}
Pro	C^{β}	Met	C^{ϵ}

Table S4. Parameters used for Delphi calculations [a]

Program version	Delphi V. 4 Release 1.1
Main equation	Linear Poisson-Boltzmann
ε_{in}	20
ε_{out}	78.54
$r_{solvent}$	1.4 Å
r_i	Amber's *leaprc.ff03.r1* with hydrogen atoms of null radius set to 1.0 Å radius
Surface type	Solvent accessible surface
z_i	$\pm 1e\ C$
$r_{ion,i}$	2.0 Å
$\rho^f(r)$	Amber's *leaprc.ff03.r1* with special rules for protonated variants
Numerical method	Finite differences
Convergence criteria	Changes less than $10^{-4} k_B T\ kJ$ in all grid points, or a maximum of 800 iterations.
Grid points	As many as needed so that the major dimension of the studied protein occupies, at most, 80% of the grid.
Grid spacing	0.5 Å
Grid center	Center of the parallelepiped containing the studied protein (grid edges are parallel to the coordinate axis). That parallelepiped is later enlarged to become the cubic grid.
Boundary condition	*coulombic* (Debye-Hückel approximation).

[a] Atomic charges and radii are discussed in the Methods section.

Table S5. Parameters used in the Metropolis Monte Carlo calculations.

Parameter	**Value**
T	Depends on the case studied.
pH	Depends on the case studied.
equilibration_scans	1000
full_scans	5000
single_move_probability	0.95
double_move_probability	0.03
triple_move_probability	0.02
$E_{threshold}$	$3RT\ln 10\ kJmol^{-1}$
Pseudo-random number generator	Linear congruential

Table S6. Mean and standard deviation of the absolute values of $\Delta\Delta G_{ele,int}^{U}$ (equation S16), the contribution to the stabilizing energy due to the electrostatic interaction energy between residues in the three ensemble unfolded models. The data refers to the 80 mutants of the complete set. The mean and standard deviation of the absolute predicted stabilizing energy of the *Simple* model ($\Delta\Delta G_{\text{simple}}$) is shown for comparison.

Model	**Mean absolute value** $(kJmol^{-1})$	**Std. dev. of the absolute value** $(kJmol^{-1})$
Minimum energy ($\Delta\Delta G_{ele,int}^{U}$)	6.09	5.00
Boltzmann energy ($\Delta\Delta G_{ele,int}^{U}$)	5.74	4.80
Average energy ($\Delta\Delta G_{ele,int}^{U}$)	7.35	6.48
Simple unfolded ($\Delta\Delta G$)	5.57	5.89