PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Predicting stabilizing mutations in proteins using Poisson-Boltzmann based models: study of unfolded state ensemble models and development of a successful binary classifier based on residue interaction energies

Jorge Estrada[1,2], Pablo Echenique[2,3,4], Javier Sancho[1,2*]

[1]Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias. Universidad de Zaragoza. Pedro Cerbuna 12, 50009 Zaragoza, Spain

[2]Biocomputation and Complex Systems Physics Institute (BIFI). Joint Unit BIFI-IQFR (CSIC). Mariano Esquillor s/n, Edificio I+D, 50018, Zaragoza, Spain

[3]Instituto de Química Física "Rocasolano", CSIC, Serrano 119, 28006, Madrid, Spain

[4]Departamento de Física Teórica, Universidad de Zaragoza, Pedro Cerbuna 12, 50009, Zaragoza, Spain

[*]Correspondence to: Javier Sancho (jsancho@unizar.es)

Abstract

In many cases the stability of a protein has to be increased to permit its biotecnological use. Rational methods of protein stabilization based on optimizing electrostatic interactions have provided some fine successful predictions. However, the precise calculation of stabilization energies remains challenging, one reason being that the electrostatic effects on the unfolded state are often neglected. We have explored here the feasibility of incorporating to a Poisson-Boltzmann model electrostatic calculations performed on representations of the unfolded state as large ensembles of geometrically optimized conformations calculated with the ProtSA server. Using a data set of 80 electrostatic mutations experimentally tested in two-state proteins, the predictive performance of several such models has been compared to that of a simple one that considers an unfolded structure of non interacting residues. The unfolded ensemble models, while showing correlation between the predicted stabilization values and the experimental ones, are worse than the simple model, suggesting that the ensembles do not capture well the energetics of the unfolded state. A more attainable goal is that of classifying potential mutations as either stabilizing or not, rather than accurately calculating their stabilization energies. To implement a fast classification method that can assist in selecting stabilizing mutations, we have used a still simpler electrostatic model based only in the native structure and have determined its precision using different stabilizing energy thresholds. The binary classifier developed finds 7 true stabilizing mutants in every 10 proposed candidates and can be used as a robust tool to propose stabilizing mutations.

Introduction

In 1982, site-directed mutagenesis was first applied to modify the active site of proteins of known structure and mechanism, marking the effective beginning of Protein Engineering.[1-3] Since then, many biotechnological and biomedical applications have benefited from the advances in the field, using either directed evolution or rational design methods.[4-6] The latter are specially appealing, as they provide a greater control and contribute to the understanding of protein structure-function relationships. In particular, computational methods[7] are key in rational design, allowing to implement and test complex models, and to analyze increasingly large numbers of target proteins. Computational methods greatly vary in complexity. While some use sequence-only models, others consider protein structure; some use machine-learning algorithms and others exploit statistical, semi-empirical or first-principles descriptions of protein energetics. Structure-based, first-principles computational methods are those with a greater potential. Whatever the method used, the goal is to design a protein with the properties needed for the required function. A key property is conformational stability[8,9] because most proteins must be conformationally stable or achieve stability under a given set of conditions. Stability typically depends on the protein solution conditions (pH, ionic strength, solvent, temperature, etc.), and is a function of protein physical-chemical properties. Among the many protein stabilizing methods known (helix optimization, entropic stabilization, disulfide introduction, packing optimization, etc.),[10-16] those focusing on electrostatics[17,18] are central because, to some extent, they constitute the basis of many of the other methods, and are also helpful to understand the interaction of proteins with other molecules.

Electrostatic models of proteins[19] go back to the Linderstrøm-Lang model consisting of a spherical protein with continuous surface charge, embedded on continuous solvent.[20] Its main improvement, the Tanford-Kirkwood model,[21] included point charges and has been used with further refinements as a tool to stabilize proteins in recent years.[22,23] Other continuous-solvent methods are Poisson-Boltzmann[24] (also used in protein stabilization[25]) and Generalized-Born.[26] Some faster to calculate empirical models[17] and some computationally expensive explicit solvent all-atom models or simplified dipolar models[27] have also been used.

Methods based on the Tanford-Kirkwood or the Poisson-Boltzmann models are relatively simple, require a moderate computational cost, and have sometimes been successfully implemented.[28,18] These methods usually focus in designing surface mutations assuming a random coil model of the unfolded state. The high conformational flexibility attributed to folded surface residues allows a less accurate structural modeling of the mutant protein, compared to that in protein core optimization, and tends to minimize complications arising from steric clashes of the mutant residue.[18] However, the previous assumptions do not always hold[18,29] as sometimes other electrostatic (such as hydrogen bonds) and non-electrostatic contributions need

to be to consider. Besides, the unfolded state ensemble may not always be rightly modeled as a random coil of non-interacting residues. Here, we focus on the latter problem and explore a model of the unfolded state consisting in an ensemble of unfolded conformations.

Though Poisson-Boltzmann is considered the reference among the macroscopic continuous methods,[30,31] simpler models such as Tanford-Kirkwood have proved successful[32] and several works have shown that Tanford-Kirkwood residue interaction energies are similar to those calculated with Poisson-Boltzmann models.[33,34] The general strategy used with the Tanford-Kirkwood model consists in first identifying residues with unfavorable electrostatic interactions, and then suggesting charge reversal or charge deletion mutations. In this work, we use Poisson-Boltzmann calculations to detect residues in the wild type protein with unfavorable electrostatic interactions. The method performs, with less computational effort, similarly well as a more detailed version including the calculation of the electrostatic energy of the possible mutants.

Fundamental in the comparison of the different models is to use an adequate metric. Previous works have focused on predicting stabilization energies.[32,35] Although the energy of stabilization is a very informative metric, the precise calculation of stabilization energies still constitutes a difficult quantitative goal for current approaches. A more attainable goal, very important for providing reliable tools for rational protein stabilization, is to accurately classify mutations as either stabilizing or not. Metrics using binary classifications are currently used to evaluate the performance of prediction methods,[36] e.g. general stabilization methods[37] or algorithms used to detect pathological mutations from stability calculations.[38] Using a binary classification, our approach finds 7 true stabilizing mutants in every 10 proposed candidates.

Methods

Energy models

To test different electrostatic models by their capacity to predict changes in Gibbs energy of folding in mutant protein variants relative to wild type ($\Delta\Delta G^{fol}_{wt\rightarrow mut}$), we have made the gross but necessary approximation of considering the non-electrostatic component of the Gibbs energy (ΔG^{fol}_{n-ele}) to be the same in both the wild type and the mutant protein variant.

$$\Delta\Delta G^{fol}_{wt\rightarrow mut} = \left(\Delta G^{fol}_{ele,mut} + \Delta G^{fol}_{n-ele,mut}\right) - \left(\Delta G^{fol}_{ele,wt} + \Delta G^{fol}_{n-ele,wt}\right) \approx \Delta\Delta G^{fol}_{ele,wt\rightarrow mut} \quad (1)$$

To further simplify the model we have dealt with two-state proteins that can be in either the folded or the unfolded state. Accordingly, their electrostatic Gibbs energies of folding can be expressed as:

$$\Delta G^{fol}_{ele} = G^{F}_{ele} - G^{U}_{ele} \quad (2)$$

Our overall goal is to test different protein electrostatics representations and to develop a predictive model for protein stabilization that can be easily implemented and that allows for fast calculations. A single energy model has been used to represent the folded state in which the protein is considered as a continuous dielectric with point charges, surrounded by a continuous solvent (water) with salt ions. On the other hand, four different two-state models have been tested, differing in the description of the unfolded state. In all of them, we have assumed that, in the unfolded state, all the charges are completely exposed and surrounded by the continuous dielectric of the solvent. Our reference unfolded model (called below *Simple* unfolded model) assumes a fully expanded unfolded conformation for the protein where the residues do not interact among them. We foresee that the next simplest model that could treat more realistically the unfolded conformations would simply account for the interactions between residues as in complete exposure to solvent. Further improvements should consider the exact exposure to solvent of each residue and the actual screening due to the ionic strength in the unfolded state, in a similar way as we do for the folded state. In addition to these four two-state models, we have tested and discussed a much simpler predictive model based only in the folded structure of the wild type protein.

Temperature, pH, and ionic strength are those used in each experiment (Table 1). In cases where we use a value obtained in conditions different to those in which it is applied (such as in the pK_a^{model} values), the fact is indicated.

The energy model for the folded state

This state is considered to have a single structural conformation but multiple protonation configurations. A permittivity constant of 78.54 is used for the solvent (water). As there is no agreement about the best value for the permittivity of the protein[42] (see also the range of values used in the methods reviewed in ref. 43) we have selected a value of 20 because it provided better energy predictions in preliminary tests. The linear Poisson-Boltzmann model has been used to obtain the electrostatic potentials, considering the ionizable residues in table S2. For implementation simplicity, we have not considered the terminal groups as ionizable so that in our model each residue can have, at most, one ionizable group and be either charged (± 1) or neutral. The neutral state of histidine is determined by the hydrogen bond network calculated by the REDUCE software[44] for each protein. Therefore, for a protein with N ionizable residues, there are 2^N protonation configurations, having each configuration $\vec{p}$ an energy $G_{ele}^F(\vec{p})$. Assuming the same non-electrostatic energy in each protonation configuration, the folded state electrostatic energy corresponds to the Boltzmann-weighted average of:

$$G_{ele}^{F} =< G_{ele}^{F}(\vec{\mathrm{p}}) > \quad (3)$$

To avoid time consuming calculations, our models use the Metropolis algorithm as an approximation for sampling the protonation configurations in the folded state. To calculate the electrostatic energy of a given protonation configuration we have followed the description of ref. 45 that breaks the energy into 2 terms: one accounting for the charging of each ionizable residue independently of all other ionizable residues, and the other accounting for the interaction among residues not considered in the first term of eq. (4). Thus, for a protonation configuration $\vec{p}$:

$$G_{ele}^{F}(\vec{p}) = \sum_{i\in\mathcal{R}} p(i)\Delta G_{cha}(i) + \sum_{i\in\mathcal{R}}\sum_{j\in\mathcal{R},i<j} p(i)p(j)\Delta G_{ele,int}(i,j) \quad (4)$$

where $\mathcal{R}$ is the set of all ionizable residues, $p(i)$ is 1 when residue i is charged, and 0 otherwise, and $i < j$ if residue i comes first in the protein sequence.

The calculation of equation (4) follows ref. 45 but differs in the following: 1) for each ionizable residue i, its potential maps ϕ_i and ϕ_i' are calculated using the charges of all its atoms, and not just those of the titratable group; 2) we use AMBER partial charges (see below), extended for dealing with the non-standard protonation state of ionizable residues (see Table S2); we do not yet deal with the ionization of the N- and C-terminal groups; 3) for the model residues, we use the pK_as provided in Table 3 of ref. 46; 4) we do the calculations using Delphi without focusing.

The details of the calculation appear in the supplementary materials. The calculation of $\Delta G_{cha}(i)$ for each ionizable residue i requires obtaining its intrinsic pK_a,[21] which is done through a thermodynamic cycle involving four states ($G_{ele}^{P_n}$, $G_{ele}^{M_n}$, $G_{ele}^{M_c}$, $G_{ele}^{P_c}$) representing, respectively, the neutral protein, the neutral ionizable residue, the charged ionizable residue and the charged protein (see suplementary fig. 1). This requires 4 Poisson-Boltzmann calculations for each ionizable residue (one per state), which are also used for obtaining the second term of equation (4).

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. Being $x(i,j)$ the Boltzmann-weighted fraction of protonation configurations $\vec{p}$ that have both i and j charged, the residue interaction energy is:

$$\Delta G_{ele,int}(i) = \sum_{j\neq i, j\in\mathcal{R}} x(i,j)\Delta G_{ele,int}(i,j) \quad (5)$$

One of our models, the *Native only* one, is entirely based on just the structure of the *wild type* folded state. In this model, we assume $\Delta\Delta G_{wt\to mut}^{fol}$ (termed $\Delta\Delta G_{nat}$ for this model) to be the opposite of the electrostatic interaction energy for the mutated residue i from the wild type: $\Delta\Delta G_{nat}$ =-$\Delta G_{ele,int}(i)$. Therefore, for this model, the analysis of both neutralization and

inversion mutations makes the gross assumption that replacement of an ionizable residue removes its electrostatic effect. The main advantages of this *Native only* model are that it avoids calculating all possible mutants (as it would happen in the *Simple* model) and, above all, that it offers a clear and visual rationale (see figure 4) for selecting stabilizing mutations. Of course, the newly introduced residue could be easily modeled, as it is in the implementation of the *Simple* model, for further checking of the results. On the other hand, the more promising residues according to the *Native only* model predictions could be further studied using the *Simple* model.

Energy models for the unfolded state.

We have used four alternative models for the unfolded state. The first model is termed *Simple* and considers a completely unfolded structure where the residues do not interact among them and only the independent charging energies of the residues are considered. The other three models consider ensembles of 2000 unfolded structures created with the ProtSA server,[47] which constitute a geometrically accurate representations of fully unfolded proteins, as described before.[48] However, because these unfolded ensembles have not been tested for energetic accuracy, we have computed, from the electrostatic energies of the individual conformations, the ensemble electrostatic energy in three different ways that differentiate the models as: *Minimum energy* (the electrostatic energy of the ensemble is taken as the minimum electrostatic energy among all the conformations), A*verage energy* (the electrostatic ensemble energy is taken as the average of the 2000 conformation electrostatic energies), *Boltzmann-weighted energy* (the ensemble electrostatic energy is taken as the Boltzmann-weighted average of the conformations electrostatic energies). In these three ensemble models the ionizable residues are considered to be independently charged and they are allowed to interact as if they were in a completely solvent-exposed environment. This general assumption involves several simplifications that are enumerated in the discussion, and allows to avoid Poisson-Boltzmann calculations for the unfolded ensemble, significantly reducing the computation time needed.

In the *Simple* unfolded model the energy of the different protonation configurations, $\vec{p}$, is represented by equation (4). However, as the ionizable residues do not interact the equation can be simplified to:

$$G_{ele}^{U}(\vec{p}) = \sum_{i \in \mathcal{R}} p(i) \Delta G_{cha}(i) \qquad (6)$$

Furthermore, the structure and solvent exposure of the ionizable residues is considered to be similar to those of their corresponding model compounds, and $pK_a^{int}(i)$ is replaced in eq. (S2) by $pK_a^{mod}(i)$, so that:

$$\Delta G_{cha}(i) = \gamma(i)RT(\ln 10)\left(pH - pK_a^{mod}(i)\right) \quad (7)$$

Since the ionizable residues do not interact, the Boltzmann-weighted average of all the protonation configurations can be calculated directly. The fraction of unfolded proteins having residue i charged, $x(i)$, is given by:

$$x(i) = \frac{e^{\frac{-\Delta G_{cha(i)}}{RT}}}{1 + e^{\frac{-\Delta G_{cha(i)}}{RT}}} \quad (8)$$

and the electrostatic energy of the completely unfolded protein becomes:

$$G_{ele}^{U} = \sum_{i \in \mathcal{R}} x(i)\gamma(i)RT(\ln 10)\left(pH - pK_a^{mod}(i)\right) \quad (9)$$

The three ensemble unfolded models are extensions of the *Simple* unfolded model by addition of electrostatic interaction energies between ionizable residues calculated from analysis of 2000 unfolded conformations generated with an offline version of the ProtSA server[47] with the solvent probe radius set to 1.4 Å. In the three models we assume that all unfolded conformations of a given protein have the same charge for the same residue, therefore being the inter-residue distances r_{ij} (see equation 10) the only significant parameter that changes from one conformation to another (the residue charges being the same). We model those charges as point charges placed in a single atom of the ionizable residues and we further assume that they are the same as the mean fractional charges $\gamma(i)x(i)$ previously obtained for the *Simple* unfolded model (this implies that those mean charges per residue do not significantly vary in any of the three unfolded energy models). We have also considered that all the ionizable residues are completely exposed to solvent so that, for the calculation of $\Delta G_{cha}(i)$, they display the pK_a of the model compounds. With these assumptions, the electrostatic interaction energy between residues for a given protein unfolded conformation is:

$$G_{int} = \sum_{i} \sum_{j,i<j} \frac{e^2 N_A}{4\pi\varepsilon_0\varepsilon_r 10^{-10}} \frac{\gamma(i)x(i)\gamma(j)x(j)}{r_{ij}} \quad (10)$$

where N_A is the Avogadro number and ε_r is the solvent relative permittivity. Calculations related to protein mutants use the same unfolded conformations as for the corresponding *wild type* protein. Therefore the atoms used to place the charges of the ionizable residue have been defined (see Table S3). Combining the charging and interaction energy terms, the electrostatic energy of a given unfolded conformation of the ensemble, κ, becomes:

$$G_{ele}^{U}(\kappa) = \sum_{i \in \mathcal{R}} x(i)\gamma(i)RT(\ln 10)\left(pH - pK_a^{mod}(i)\right) + \sum_{i} \sum_{j,i<j} \frac{e^2 N_A}{4\pi\varepsilon_0\varepsilon_r 10^{-10}} \frac{\gamma(i)x(i)\gamma(j)x(j)}{r_{ij}} \quad (11)$$

In the *Minimum energy* model the G_{ele}^{U} representing the unfolded ensemble is the minimum energy among all the electrostatic energies of the 2000 unfolded conformations: $G_{ele}^{U} = \min_{\kappa} G_{ele}(\kappa)$. In the A*verage energy* model: $G_{ele}^{U} = \frac{1}{2000}\sum_{\kappa} G_{ele}^{U}(\kappa)$. In the *Boltzmann-weighted energy* model: $G_{ele}^{U} = \sum_{\kappa} \frac{G_{ele}^{U}(\kappa) e^{-G_{ele}^{U}(\kappa)/RT}}{\sum_{\lambda} e^{-G_{ele}^{U}(\lambda)/RT}}$, where κ and λ are indexes in the number of conformations.

Proteins, mutants and model pK_as

The experiments and proteins studied are shown in Table 1. The experimental conditions indicated are those from the cited reference articles. The table also lists the PDB codes of the folded structures that have been used; these are the same as those used in the reference articles for structural analysis or calculations. In the case of SNase, the experiment seems to have been done with the whole sequence, but the calculations seem to have been done with the smaller sequence in structure 1stn; therefore, for consistency, we have used 1stn for the calculations too. The experimentally determined folding free energy of all mutants with respect to their corresponding wild types are shown in Table S1. Since several electrostatic models will be tested in their capability to suggest stabilizing mutations, we have exclusively selected those mutations affecting ionizable residues (Asp, Glu, Cys, Tyr, Lys, His and Arg) in the wild type structure; this excludes Csp-Bc mutants Gln2Leu, Asn11Ser, Gly23Gln, Ser24Asp, Thr31Ser, Gln53Glu, Asn55Lys, Val64Thr, and Leu66Glu.The following SNase mutants have also been discarded: Glu135Lys (mutant building failed because of steric clashes), Asp143Lys, and Asp143Asn (Asp143 is not modeled in the 1stn PDB file). The proteins selected are two-state, small, monomeric, with no gaps in their sequence, no cofactors, and no ligands. For apoflavodoxin, that behaves as a 2-state protein in urea denaturation but as a 3-state one towards thermal unfolding,[9] the 2-state chemical denaturation data has been always used.

The complete set studied consists of 80 mutations (*complete set*), 22 of them actually representing 11 mutations studied at two different ionic strengths. We have also defined a subset of 56 mutations for which the reference articles include a stability prediction (*comparable set*), which comprises 8 apoflavodoxin, 9 ubiquitin and 39 SNase mutants.

To describe the deprotonation of single residues (model residues) in solution we have used the pK_as provided in ref. 46; see Table S2. Although the experimental conditions used to obtain those pK_as (temperature, ionic strength, and the exact model compounds used to represent the residues) do not sometimes apply to the cases studied here, the choice allows to compare this work with previous reports.

Radii and point charges of atoms in the folded state

To define the solvent-accessible molecular surface separating the protein dielectric region from the solvent dielectric region, we used a single set of van der Waals atomic radii from the Amber parameter set *leaprc.ff03.r1*,[49] with the radii specified in files parm99.dat and frcmod.ff03. Since it specifies a null radius for some hydrogen atoms (such as H^{γ} of serine), we changed those values to 1.0 Å to avoid charges near the protein surface.

The atomic partial charges used were those in the Amber parameter set *leaprc.ff03.r1* (file all_amino03.in). We only used the charges for the standard protonation states from Amber. For obtaining the non-standard protonation state of all the residues, we followed the rules in Table S2, placing point charges in selected atoms to model the addition or removal of a hydrogen atom. Note that no atom is actually removed from the protein structure and the dielectric boundary does not change (this simplifies the calculations for our energy model; see discussion of eq. 10 of ref. 45).

Preparation of protein structures

The PDB structures of the wild type proteins used in the reference articles were considered. They have no gaps, and only 1csp had missing heavy atoms different from the C-terminal oxygen (affecting Glu3, Glu21, Glu36, and Glu66). Those side chains were reconstructed with SCWRL 3.0.[50] All structures consisted of a single chain except 1c9o, where we chose chain A for the calculations. All ligands and ions were removed (sulfate ion in 1ftg, sodium ion and tris buffer molecules in 1c9o). The mutant structures were built from the corresponding wild type structure using SCWRL 3.0. Then hydrogen atoms were added with REDUCE v. 3.10.070818, which follows the method in ref. 44. Finally, for all structures, the N-terminal hydrogens and the C-terminal oxygen were removed if present to deal with them as non-terminal residues so that the terminal groups are neutral.

Ensembles of 2000 wild type unfolded structures were created using the protein sequences as input to an off-line version of the ProtSA server.[47] For the mutants we used the same structures obtained for their wild types, and only the placement or removal of a charge differentiates mutant and wild type.

Sampling of the protonation configurations in the folded state

To calculate the folded state energy from the Boltzmann distribution of the available 2^N protonation configurations we have used the Metropolis Monte Carlo method.[51] The Beroza *et*

al.[52] application of such method has been followed with a modification consisting in adding transitions for triplets of strongly interacting residues, and using a different probability for choosing the transition to try in each step. We start from a random initial protonation state, and run 6000 full scans, discarding the first 1000 ones, considered as equilibration scans. Each scan consists of as many Monte Carlo steps as the number of ionizable residues, plus the number of couples and triplets of strongly interacting residues, therefore effectively running (on average) 6000 tries per ionizable residue (and per strongly interacting groups of residues). Complete details on the method appear in the supplementary materials.

Regression analysis

For the regression analysis only the mutant proteins were considered, and the wild type variants only served as reference points.

Results and discussion

We have created a small data set of electrostatics-related mutations experimentally analyzed (Table S1). This is just a first step towards building a benchmark data set for this kind of mutations —see ref. 36 for an analysis of the desired properties of a benchmark data set—. Though ref. 37 established a benchmark data set for stabilizing mutations, it covers any type and not just electrostatic mutations. Similarly, ref. 53 uses a data set including non electrostatic mutations. A related initiative, the pK_a cooperative,[54] does cover electrostatic mutations, but is currently focused on calculating residue pK_a values and not strictly on protein stability. Our *complete* data set contains 80 mutations belonging to 5 different proteins. Based in studies of experimental error[55,37] and in our previous experimental work in the field of protein stability[56] we have chosen a threshold of $-2\ kJmol^{-1}$ in the change in folding free energy (mutant minus wild type) to define a given mutation as stabilizing. There are 20 such stabilizing mutations in the database. The influence of the threshold selected on the accuracy of the classifications provided by our different models for hypothetical mutations as either stabilizing or otherwise will be discussed below.

The 80 mutations of the *complete* data set affect ionizable residues, comprising 47 mutations that correspond to inversions of charge (substitution of an ionizable residue for another one of different sign), and 33 to neutralization (substitution of an ionizable residue for a non ionizable one). We have explored the feasibility of further dividing the neutralization mutations into those that greatly change the size or shape of the original residue and those that do not, as the former ones are likely to significantly affect stability through other means than electrostatic interactions. However, given the small size of the data set, we have not

implemented this additional classification category. Using the *complete* data set we have analyzed the accuracy of the predictions provided by our different models on mutations involving ionizable residues and have compared them to previous work. We are interested both in the quantitative prediction of stability changes and in the rather more practical (and simple) biotechnological problem of classifying potential mutations into stabilizing or not-stabilizing.

Quantitative predictions of stability changes with different models of the unfolded state

The correlation between the stability changes previously determined experimentally for the *complete* set and the corresponding stabilizations calculated by the different models tested is shown is figure 1. The figure also shows what would be perfect correlations (the diagonal lines crossing the graphs), as well as the lines of the least squares fittings, with the numerical values of the fitting (r^2, slope and intercept) in inset boxes. Perfect correlations between measured and calculated stabilization are out of question, not only because of all the simplifications introduced in the model but also because many mutations tend to give rise to stability changes, usually destabilizing, not directly related to electrostatic interactions.[57] It is clear (figure 1) that the *Simple* model outperforms all others, with $r^2 = 0.54$. Given the good correlation for the *complete* set of mutations on ionizable residues found for the *Simple* model, we have separately analyzed neutralization mutations and charge inversion mutations to determine whether either type could be more accurately calculated. No significant difference in correlation is observed ($r^2 = 0.59$ for inversion, 0.57 for neutralization; graphs not shown), which suggests, on one hand, that the common destabilizing effects mentioned above[57] occur to a similar extent in the two groups considered and, on the other hand, that the assignation of charges (Table S2) to the introduced residues in the charge-reversal group are reasonably correct.

Good performance of different implementations similar to our *Simple* model in the quantitative prediction of stability changes had been observed before using the Tanford-Kirkwood model with the solvent-accessibility correction due to Gurd[58] for the apoflavodoxin and ubiquitin mutations,[39,22] or using the Poisson-Boltzmann single site ionization protocol for SNase.[35] To facilitate comparison, we will globally refer to the combined predictions from those models as the *Mixed reference* model predictions. One pertinent question is how the predictions of the present implementation of the *Simple* model compare with those of the *Mixed reference* one. For the 56 mutants in the *comparable set*, the relevant fits that allow the comparison are shown in figure 2. As can be seen, our current implementation performs similarly well (with a slightly higher R^2 of 0.63, compared to 0.59 for the *Mixed reference* model), although with a lower slope of 0.42 compared to 0.69 for the *Mixed reference model*, which makes predictions from the latter closer in magnitude to the experimental values. A more detailed comparison is

not possible because each of those previous works only provided predictions for a subset of the interactions included in the *comparable set*.

Significantly, the three models that introduce electrostatic interactions in the unfolded state, represented by an ensemble of unfolded conformation, lower the correlation found in the *Simple* model (figure 2), which suggests that those ensembles, although geometrically correct[48] do not capture well the energetics of the unfolded state. In addition, the approximations done in the expression of the electrostatic energy of the unfolded state in the ensemble models (equation 11) contribute to the inaccuracy of the calculated energies. Those approximations are: 1) the same 2000 conformations are used to describe the unfolded wild type ensemble and those of the mutants, which only differ from the wild type in the value of the charges; 2) the average charge per residue of the *Simple* model is assumed to apply to each of the 2000 conformations in the ensemble; 3) the ionic strength is not considered; 4) the intrinsic pK_a of each residue is that of its corresponding model pK_a (i.e. the residues are fully exposed); 5) the charge of the residue is located in a point rather than distributed among its atoms.

Of the three ensemble models tested, the one offering the poorest correlation is the *Average-energy* model where the energy of the unfolded ensemble is represented by the arithmetic average of the 2000 conformations analyzed. Using instead the energy of the most stable of the 2000 conformations to represent the energy of the unfolded state of each mutant improves the fit and raises the slope closer to 1. Using the Boltzmann-weighted energy of the ensemble further improves the fit and slope. This improving trend observed in the fit as physically more correct models of the unfolded ensemble are used raises hopes that representing the energetics of the unfolded state of proteins as Boltzmann averages of large, energetically refined ensembles may prove useful. It is possible that energetically refined ensembles, based on those provided by ProtSA or from MD simulations of denatured conformations, may provide a more satisfactory representation of the electrostatic interactions in unfolded proteins and that their Boltzmann-weighted energies, combined with the energy calculated for the folded state, may predict stability changes better that the *Simple* model. Further research on the proper modeling of the unfolded state and its interactions is needed to test this possibility. Towards this end, it should be borne in mind that the calculated interaction corrections introduced on the *Simple* model by any of the three ensemble unfolded models tested —i.e. the second term in equation (11)— are not negligible but rather in the range of the predicted stabilizing energies (see supplementary materials and table S6). The same will very likely apply to more refined ensembles. Therefore, their energies would need to be calculated with great accuracy in order to be of help to improve the correlations.

In addition to the four two-state models discussed above, we have tested an even simpler model (*Native only* model) that can be used to calculate stability changes disregarding the energetics of the unfolded state. This model computes stability changes assuming the

electrostatic effect of the mutated residue is removed by either a neutralization or a charge reversal mutation. This model still shows some correlation, with $r^2 = 0.39$ which, albeit lower than that of the *Simple* model, is higher than that of any of the three ensemble models. Although for the goal of obtaining accurate estimations of the stability changes introduced by mutations involving ionizable residues the *Simple* model is the best among those tested here, the significant correlation shown by the *Native only* model and its great simplicity and speed compared to the others suggest it could be quite useful as a mutation classifier.

Accurate classification of mutations for protein engineering purposes

For many biotechnological applications an accurate anticipation of the stabilization afforded by a given mutation may not be essential and it may suffice to know whether a given mutation would stabilize or not a given protein to at least a specified amount. This goal may be much easier to achieve than that of accurately predicting the actual figures of the change in stability. From this perspective, the practical problem becomes one of classification, as found in many other disciplines (ref. 59, chap. 18). From the many evaluation measures that have been proposed[36,60,61], we have focused on *precision*. This is because we are interested in finding stabilizing mutations and, at the same time, in reducing the number of false positives (FP) in the predictions, given the high cost (in time and resources) associated to experimentally testing large numbers of predicted mutants. For a given set of predictions, with true positives (TP) being the number of stabilizing predictions which are experimentally confirmed to be stabilizing, and FP the wrong predictions, precision is defined as $\frac{TP}{TP+FP}$. To do a classification into stabilizing and non-stabilizing mutations we first sort all the calculated mutants by the predictive quantitative stabilization energy assigned by the model. Then, the ordered list is split into two parts by using a threshold value. Predictions with a stabilization energy lower than the threshold (i.e., a mutation predicted to be more stabilizing than those at the threshold), are considered to be stabilizing mutations, and all others to be non-stabilizing. Choosing the right threshold value for a given model is essential for obtaining good precision. A common threshold of 0 $kJmol^{-1}$ for all models might seem a natural choice but small stabilizations are usually of no interest in real applications. Besides, as seen in Figure 1, the different models intercept the y-axis (experimental values) at points different from 0. Past studies have shown actual errors of ± 0.10 to ± 0.60 $kcal\,mol^{-1}$ [55] in protein stability determinations, and the recent work of ref. 37 on protein stability prediction upon mutation found an average absolute error of 0.44 $kcal\,mol^{-1}$ in the experimental data. Based on those reports and on our own previous expertise in the field —see ref. 56 for a discussion of reliability of experimental stability differences in the calculation of equilibrium phi values— we have set the threshold for

experimental data at $-2\ kJ\ mol^{-1}$ (negative values meaning more stable mutants), which provides us with 20 mutations classified as stabilizing in our data set, out of 80 mutations on ionizable residues. On the other hand, we have used as specific thresholds of the different models tested their predicted values corresponding to the single experimental threshold of $-2\ kJmol^{-1}$, rounding at values of half $kJmol^{-1}$. Thus, for the *Simple* model, with *slope* $= 0.45$ and *intercept* $= 0.46\ kJmol^{-1}$, the threshold ($\Delta\Delta G_{thr} = \frac{-2.0-intercept}{slope}$) becomes $-5.82\ kJmol^{-1}$, which we have rounded to $-6\ kJmol^{-1}$ for simplicity. Similarly, for the *Native only* model, the rounded threshold is $-3.5\ kJmol^{-1}$; and for the mix of preditions from the reference articles (*Mixed reference* model), the threshold, in this case obtained using only the 56 mutations of the *comparable set*, is $-4\ kJmol^{-1}$. Achieving a correct prediction of protein stability changes beyond a $2\ kJmol^{-1}$ threshold will prove very useful for the engineering of protein thermostabilization. It is clear, however, that modeling subtle biological processes involving differences in electrostatic energies lower that $2\ kJmol^{-1}$ will require more accurate models.

Table 2 shows the precision values for the *comparable data set* and our best quantitative model *(Simple* two-state model) and the *Native only* model, and compares them to the classification done using the predictions published in the reference articles. The best precision is obtained with the *Simple* model (0.82) but those of the *Mixed reference* model (0.80) and of the *Native only* model (0.75) are close. For our two models, *Simple* and *Native only,* using the larger *complete set* of 80 mutations gives essentially identical results as those obtained for the smaller *comparable data set.* The obtained precisions are much higher than those expected for a random classifying model (0.23 for the *comparable set*) and are relatively insensitive to the actual values of the model thresholds used (see supplementary material and supplementary figure S2).

Given the reasonably high precision of the *Native only* model and its conceptual simplicity and smaller computational cost compared to the *Simple* unfolded state model, we have further characterized its performance towards different types of mutations. The calculated interaction energies of all the mutated residues of the *complete set* in their corresponding wild type structures are shown in Figure 3 (with the sign changed, so that they show predicted stabilization energies upon neutralization) as bars, compared to the experimentally determined changes in folding free energy upon mutation shown as dots (red dots for inversion mutations and green dots for neutralizations). As expected, both the stabilizing or destabilizing effects brought about by mutations at a given position are typically larger in the case of inversions than in neutralizations. However, the dispersion of the data indicates that the effect exerted on protein stability by the different mutations is not only electrostatic, and there are apparently no better predictions for inversions than for neutralizations.

Suggested strategy for predicting stabilizing mutations

From the previous discussion, the *Native only* model is a good candidate to build an automatic tool to predict stabilizing mutations based in generating, for the target protein, residue stabilization profiles similar to the interaction profiles introduced by ref. 33. These profiles (Figure 4) allow to select the residues with greater potential to design stabilizing changes as those with larger predicted stabilization energy. Because proteins differ in electrostatic properties, this method, as any other method conceived to stabilize proteins, will hardly be equally suitable for each protein. In figure 4 we show two extreme cases. Apoflavodoxin, a highly acidic protein with a 2.2 acidic/basic residues ratio, contains a large number of acidic residues exhibiting at neutral pH large, positive interaction energies (Figure 4A). On the other extreme, SNase (83-231), a basic protein with a 1.4 basic/acidic residues ratio, shows a more balanced distribution of charges with only a few residues exhibiting significant positive interaction energies (Figure 4B). Therefore, apoflavodoxin provides many target residues for mutation but SNase provides very few. This explains that, while most mutations in ref. 35 were destabilizing, apoflavodoxin turned out to be an easy target for stabilization.[39]

We suggest to select stabilizing mutations for a target protein in three steps. First the stabilization potential profile for the ionizable residues of the protein will be generated with the *Native only* method. Second, potential target residues will be selected using a predefined threshold ($-3.5\ kJmol^{-1}$ for this model). Third, charge inversion mutations will be designed for the most destabilizing residues, avoiding mutations in residues essential for the biological activity of the protein. Following this procedure, and according to our precision profiles, for every 10 candidate mutations tested, 7 will actually increase the stability of the protein by at least 0.5 kcal/mol. Importantly, this method can be applied in principle to any protein whose three dimensional structure is known, regardless of whether experimental data on its stability is available.

Conclusions

We have shown how a continuous electrostatic model can be used to predict the stabilizing effect of single inversion or neutralization mutations in a set of proteins, using the same parameters for all of them. Though promising, these results are based on a small dataset, and further research will prove whether the models presented here also apply to larger and more diverse sets. The precise modeling of the unfolded state is still a pending challenge. Our proposed unfolded ensemble models, while keeping some correlation between the predicted values and the experimental ones, are worse than the *Simple* unfolded model of non-interacting residues. A likely reason is that although the unfolded ensembles used here are geometrically

consistent with experimental data on unfolded proteins, they still need to be energetically optimized. Additional reasons are the simplifications introduced in the calculation of the unfolded state interaction energies, as detailed in the discussion. On the other hand, the models tested here are two-state models and many proteins exhibit more complex equilibria.[9] In order to stabilize the less stable subdomain of such proteins, the reference state for the calculations will no longer be the unfolded state but the intermediate conformation arising from the corresponding local unfolding event.[39,16] The binary classifiers we have studied here offer a good performance as shown by the *precision* metrics, and they could assist future users in selecting stabilizing mutations, even if the exact stabilization energy is poorly predicted. Specifically, a classifier based on electrostatic residue interaction energies in the folded state computed from a PDB file can be used as a robust tool to propose stabilizing mutations.

Acknowledgements

We acknowledge financial support from grants BFU2013-47064-P from MINECO (Spain) and Protein Targets group from Diputación General de Aragón, (Spain). We thank J. L. Alonso and N. Cremades for many clarifying discussions.

References

1 G. Winter, A. R. Fersht, A. J. Wilkinson, M. Zoller and M. Smith, *Nature*, 1982, **299**, 756–758.

2 G. Dalbadie-MacFarland, L. W. Cohen, A. D. Riggs, C. Morin, K. Itakura and J. H. Richards, *Proc. Natl. Acad. Sci. U. S. A.*, 1982, **79**, 6409–6413.

3 I. S. Sigal, B. G. Harwood and R. Arentzen, *Proc. Natl. Acad. Sci. U. S. A.*, 1982, **79**, 7157–7160.

4 M. Leisola and O. Turunen, *Appl. Microbiol. Biotechnol.*, 2007, **75**, 1225–1232.

5 J. A. Brannigan and A. J. Wilkinson, *Nat. Rev. Mol. Cell Biol.*, 2002, **3**, 964–970.

6 U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore and K. Robins, *Nature*, 2012, **485**, 185–194.

7 I. Samish, C. M. MacDermaid, J. M. Perez-Aguilar and J. G. Saven, *Annu. Rev. Phys. Chem.*, 2011, **62**, 129–149.

8 C. N. Pace, S. Treviño, E. Prabhakaran and J. M. Scholtz, *Philos. Trans. R. Soc., B*, 2004, **359**, 1225–1235

9 J. Sancho, *Arch. Biochem. Biophys.*, 2013, **531**, 4-13

10 V. Villegas, A. R. Viguera, F. X. Avilés and L. Serrano, *Folding Des.*, 1996, **1**, 29–34.

11 V. G. H. Eijsink, A. Bjørk, S. Gåseidnes, R. Sirevåg, B. Synstad, B. van den Burg and G. Vriend, *J. Biotechnol.*, 2004, **113**, 105–120.

12 M. Bueno, N. Cremades, J. L. Neira and J. Sancho, *J. Mol. Biol.*, 2006, **358**, 701-712.

13 E. Bae, R. M. Bannen and G. N. Phillips, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 9594–9597.

14 J. K. Myers and S. H. Trevino, *Chim. Oggi*, 2012, **30**, 30-33.

15 R. Graña-Montes, N. S. de Groot, V. Castillo, J. Sancho, A. Velazquez-Campoy and S. Ventura, *Antioxid. Redox Signaling*, 2012, **16**, 1-15.

16 E. Lamazares, I. Clemente, M. Bueno, A. Velázquez-Campoy and J. Sancho, *Sci. Rep.*, 2015, **5**, 9129.

17 C. L. Vizcarra and S. L. Mayo, *Curr. Opin. Chem. Biol.*, 2005, **9**, 622–626.

18 K. L. Schweiker and G. I. Makhatadze, *Methods Enzymol.*, 2009, **454**, 175–211.

19 P. Kukić and J. E. Nielsen, *Future Med. Chem.*, 2010, **2**, 647–666.

20 K. Linderstrøm-Lang, *C. R. Trav. Lab. Carlsberg*, 1924, **15**, 1–29.

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

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

23 A. V. Gribenko, M. M. Patel, J. Liu, S. A. McCallum, C. Wang and G. I. Makhatadze, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 2601–2606.

24 G. Lamm, *Rev. Comput. Chem.,* 2003, ***19***, 147–365.

25 S. Spector, M. Wang, S. A. Carp, J. Robblee, Z. S. Hendsch, R. Fairman, B. Tidor and D. P. Raleigh, *Biochemistry*, 2000, **39**, 872–879.

26 A. Onufriev in *Modeling Solvent Environments: Applications to Simulations of Biomolecules*, ed. M. Feig, Wiley-VCH Verlag GmbH & Co KGaA, Weinheim, 2010, ch. 6,127–165.

27 A. Warshel and A. Dryga, *Proteins: Struc., Funct., Bioinf.*, 2011, **79**, 3469–3484.

28 S. E. Permyakov, G. I. Makhatadze, R. Owenius, V. N. Uversky, C. L. Brooks, E. A. Permyakov and L. J. Berliner, *Protein Eng., Des. Sel.*, 2005, **18**, 425–433.

29 S. Xiao, V. Patsalo, B. Shan, Y. Bi, D. F. Green and D. P. Raleigh, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 11337–11342.

30 F. Fogolari, A. Brigo and H. Molinari, *J. Mol. Recognit.*, 2002, **15**, 377–392.

31 N. A. Baker, D. Bashford and D. A. Case, *Lect. Notes Comput. Sci. Eng.*, 2006, **49**, 263–295.

32 J. M. Sanchez-Ruiz and G. I. Makhatadze, *Trends Biotechnol.*, 2001, **19**, 132–135.

33 G. I. Makhatadze, V. V. Loladze, A. V. Gribenko and M. M. Lopez, *J. Mol. Biol.*, 2004, **336**, 929–942.

34 S. S. Strickler, A. V. Gribenko, A. V. Gribenko, T. R. Keiffer, J. Tomlinson, T. Reihle, V. V. Loladze and G. I. Makhatadze, *Biochemistry*, 2006, **45**, 2761–2766.

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

36 M. Vihinen, *BMC Genomics*, 2012, **13**, S2.

37 V. Potapov, M. Cohen and G. Schreiber, *Protein Eng., Des. Sel.*, 2009, **22**, 553–560.

38 C. Riera, S. Lois and X. de la Cruz, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2014, **4**, 249–268.

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

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

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

42 C. N. Schutz and A. Warshel, *Proteins: Struct., Funct., Genet.*, 2001, **44**, 400–417.

43 E. Alexov, E. L. Mehler, N. Baker, A. M. Baptista, Y. Huang, F. Milletti, J. E. Nielsen, D. Farrell, T. Carstensen, M. H. M. Olsson, J. K. Shen, J. Warwicker, S. Williams and J. M. Word, *Proteins: Struct., Funct., Bioinf.*, 2011, **79**, 3260–3275.

44 J. M. Word, S. C. Lovell, J. S. Richardson and D. C. Richardson, *J. Mol. Biol.*, 1999, **285**, 1735–1747.

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

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

47 J. Estrada, P. Bernadó, M. Blackledge and J. Sancho, *BMC Bioinf.*, 2009, **10**, 104.

48 P. Bernadó, L. Blanchard, P. Timmins, D. Marion, R. W. H. Ruigrok and M. Blackledge, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 17002–17007.

49 *AmberTools Users'Manual. Version 1.0.*, 2008.

50 A. A. Canutescu, A. A. Shelenkov and R. L. Dunbrack, *Protein Sci.*, 2003, **12**, 2001–2014.

51 N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087–1092.

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

53 M. J. Dudek, *Proteins: Struct., Funct., Bioinf.*, 2014, **82**, 2497–2511.

54 J. E. Nielsen, M. R. Gunner and B. García-Moreno, *Proteins: Struct., Funct., Bioinf.*, 2011, **79**, 3249–3259.

55 F. Yi, D. A. Sims, G. J. Pielak and M. H. Edgell, *Journal of the Association for Laboratory Automation*, 2005, **10**, 98–101.

56 L. A. Campos, M. Bueno, J. Lopez-Llano, M. A. Jiménez and J. Sancho, *J. Mol. Biol.*, 2004, **344**, 239-255.

57 S. Ayuso-Tejedor, O. Abián and J. Sancho, *Protein Eng., Des. Sel.*, 2011, **24**, 171–177.

58 B. Ibarra-Molero, V. V. Loladze, G. I. Makhatadze and J. M. Sanchez-Ruiz, *Biochemistry*, 1999, **38**, 8138–8149.

59 S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, Prentice Hall, 2010, third edition (International Edition).

60 T. Loong, *BMJ*, 2003, **327**, 716 – 719.

61 T. Fawcett, *Pattern Recognition Letters*, 2006, **27**, 861–874.

Figure legends

Figure 1. Linear fits of protein stability changes calculated using different models vs. experimental values ($\Delta\Delta G_{exp}$). $\Delta\Delta G_{simple}$ values were calculated with the *Simple* unfolded model (A), $\Delta\Delta G_{avg}$ values with the *Average-energy* model (B), $\Delta\Delta G_{min}$ values with the *Minimum-energy* model (C), $\Delta\Delta G_{Boltzmann}$ values with the *Boltzmann-weighted-energy* model (D), and $\Delta\Delta G_{nat}$ values with the *Native only* model (E). As in the text, negative values of the calculated or the experimental energies indicate stabilization of the mutant.

Figure 2. Comparison of fits to experimental values ($\Delta\Delta G_{exp}$) for the *Simple* model of the unfolded state ($\Delta\Delta G_{simple}$) (A), and the *Mixed reference model* ($\Delta\Delta G_{pred}$) (B), considering the 56 mutants of the *comparable data set*.

Figure 3. Calculated stability changes ($\Delta\Delta G_{nat}$) for all ionizable *mutated* residues in the *complete data set* (at the different experimental conditions) using the *Native only* model and assuming a neutralization or inversion mutation, ordered from greatest to lowest expected stabilization (shown as bars; axis is on the left). Superimposed, the experimental stability changes ($\Delta\Delta G_{exp}$) of inversion (red points) or neutralization (green points), placed at the same horizontal position as the interaction energy of the mutated residue (axis is on the right).

Figure 4. Residue-specific expected stabilization profiles calculated with the *Native only* model. A. Apoflavodoxin (1ftg) at pH 7, $\mathbf{298}\,\boldsymbol{K}$, $\mathbf{0.02}\,\boldsymbol{M}$ ionic strength. B. SNase (1stn) at pH 7, $\mathbf{293}\,\boldsymbol{K}$, $\mathbf{0.14}\,\boldsymbol{M}$ ionic strength.

Figure 1

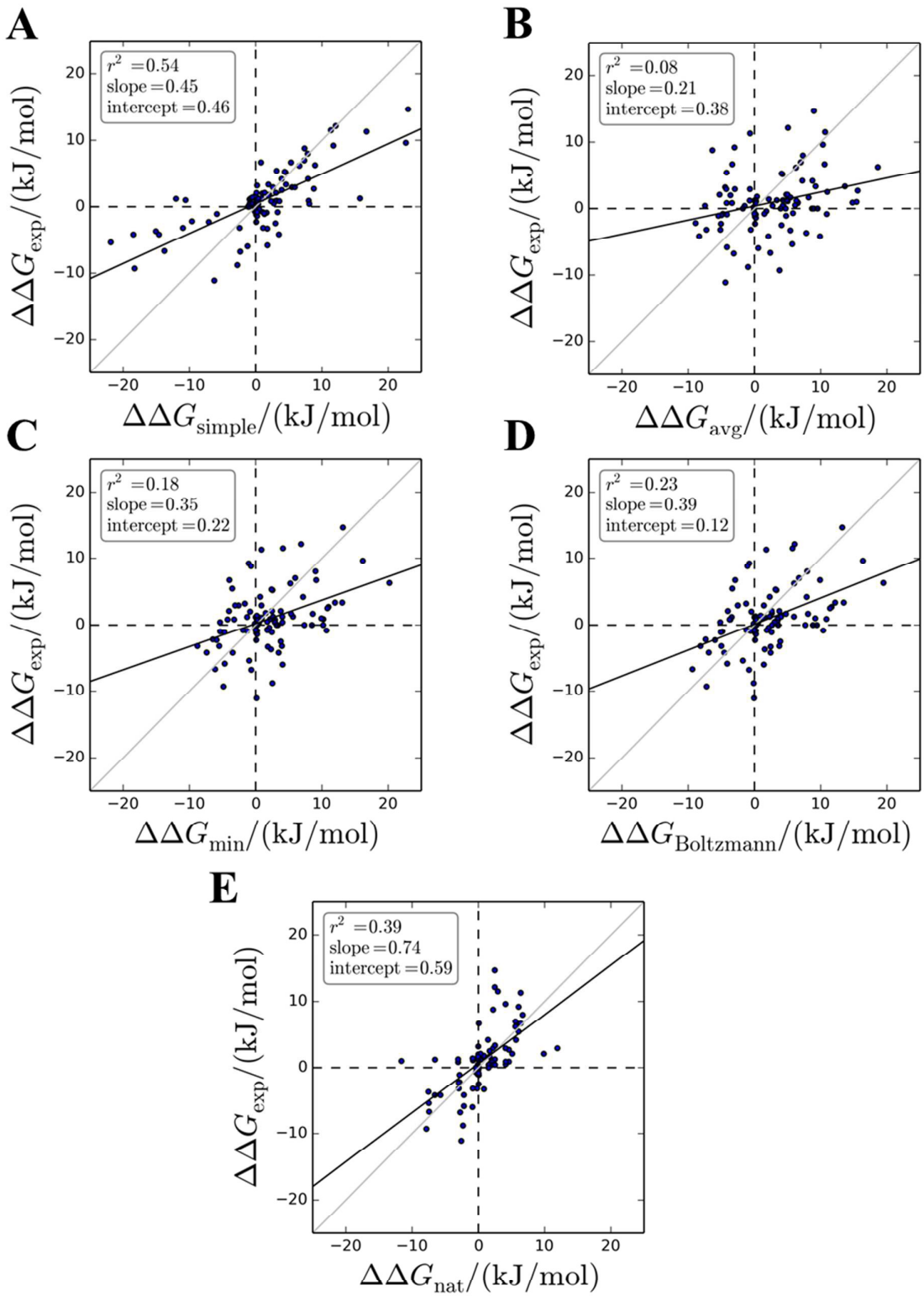

Figure 2

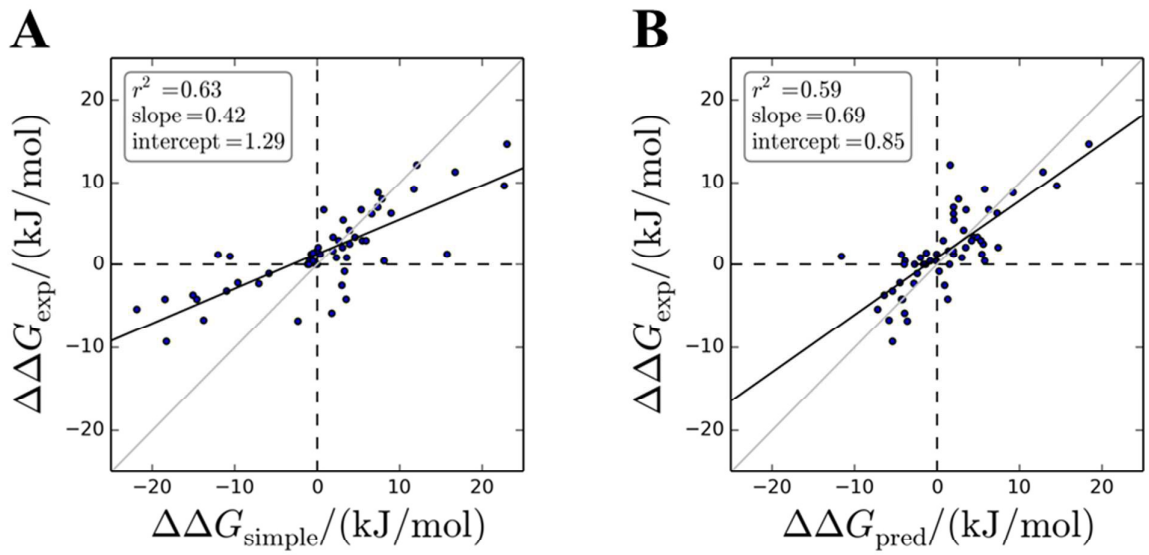

Figure 3

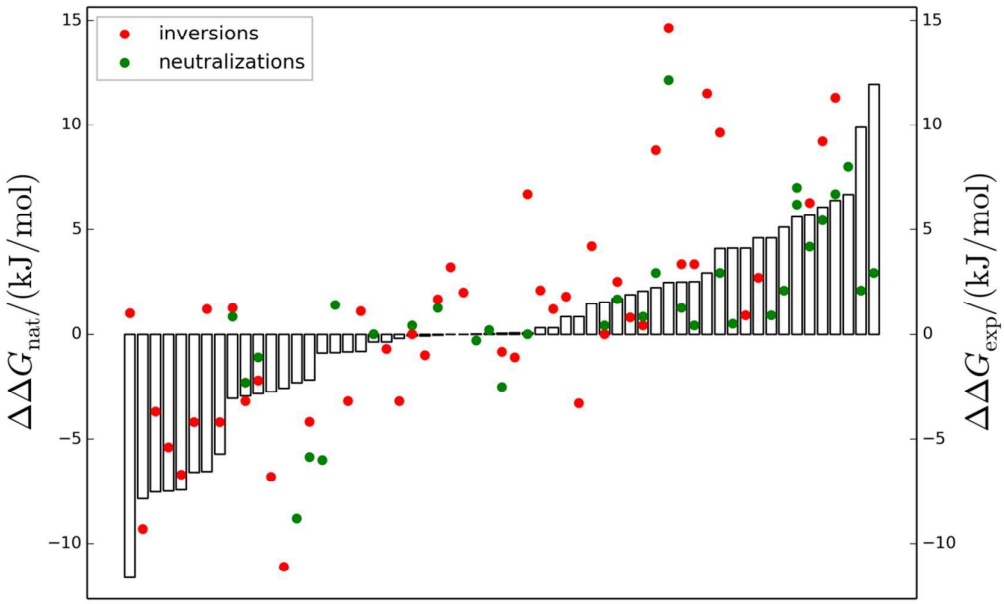

Figure 4

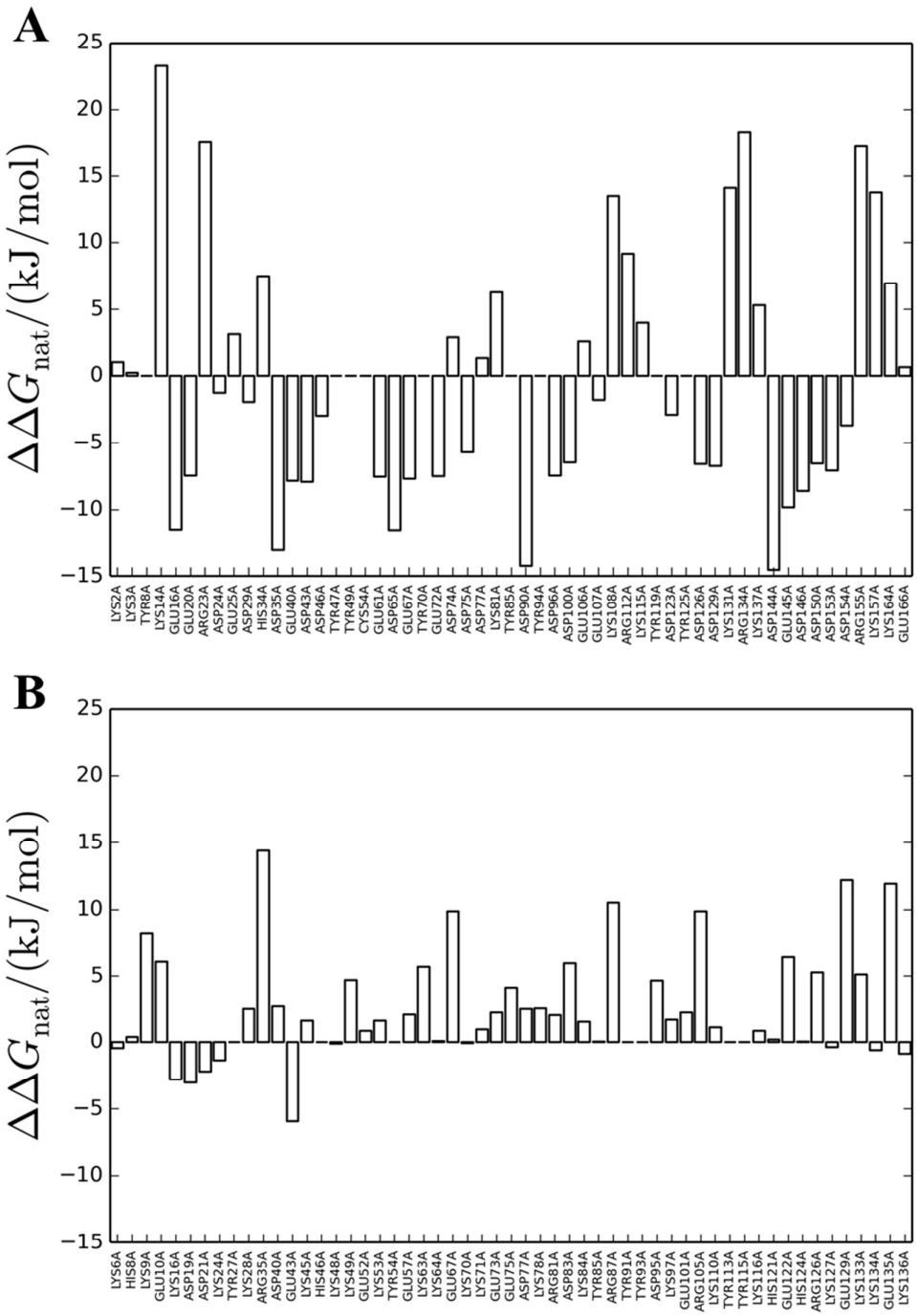

Table 1. Stability experiments considered

Protein and origin	Short name	Reference	Experimental conditions.	PDB	Number of single mutants	Num. res. in PDB
Apoflavodoxin *Anabaena PCC 7119*	apoflavodoxin	ref. 39	Urea unfolding; pH 7, 50 mM MOPS. 298.15 K. $I = 19.78\ mM$	1ftg	8	168
Cold shock protein *Bacillus caldolyticus*	CspB-Bc	ref. 40	Thermal unfolding. pH 7, 0.1 M Na cacodylate/HCl. With and without 2 M NaCl. 343.15 K. $I = 0.1\ M$ or $I = 2.1\ M$	1c9o	6	66
Cold shock protein *Bacillus subtilis*	CspB-Bs	ref. 40	Thermal unfolding. pH 7, 0.1 M Na cacodylate/HCl. With and without 2 M NaCl. 343.15 K. $I = 0.1\ M$ or $I = 2.1\ M$	1csp	4	67
Cow blood ubiquitin	ubiquitin	ref. 22	Urea unfolding. pH 5, 50 mM sodium acetate. 298.15 K. $I = 31.84\ mM$	1ubq	9	76
Cold shock protein *Bacillus caldolyticus*	CspB-Bc	ref. 41	Thermal unfolding. pH 7.0, 0.1 M Na cacodylate/HCl. With and without 2 M NaCl. 348.15 K. $I = 0.1\ M$ or 2.1 M.	1c9o	14	66
Staphylococcal nuclease (83-231)	SNase	ref. 35	GuHCl unfolding. pH 7, 100 mM NaCl and 25 mM sodium phosphate. 293.15 K. $I = 0.144\ M$	1stn	39	136

Table 2. Hit rates of different models on the *comparable set* of **56** mutants after applying the established thresholds[a]

Model	Total (TP + FP)	TP	FP	Precision ($\frac{TP}{TP+FP}$)
Native only	8	6	2	0.75
Simple	11	9	2	0.82
Mixed reference	10	8	2	0.80

[a]Thresholds are: $-3.5\ kJmol^{-1}$ for the *Native only* model, $-6.0\ kJmol^{-1}$ for the *Simple* model, and $-4.0\ kJmol^{-1}$ for the *Mixed reference* model.