A Stability Pattern of Protein Hydrophobic Mutations that Reflects Evolutionary Structural Optimization

Raquel Godoy-Ruiz, Raul Perez-Jimenez, Beatriz Ibarra-Molero, and Jose M. Sanchez-Ruiz
Departamento de Quimica Fisica, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

ABSTRACT We have determined the effect of mutations involving isoleucine and valine (i.e., mutations I→V and V→I) on the stability of *Escherichia coli* thioredoxin. Despite the similarity in chemical structure (V and I differ only in a methyl group), we find that many environments are optimized to a significant extent for either V or I. We find, furthermore, that a plot of effect of hydrophobic mutations on stability versus packing density shows a strikingly simple pattern that clearly reflects evolutionary structural optimization. The existence of such patterns suggests the possibility of rationalizing (and perhaps even predicting) mutation effects on protein stability on the basis of evolutionary models. By "evolutionary model" we specifically refer in this context to a model for mutation effects on stability in which certain physical features of the mutated residue environments are evaluated from an assumption regarding how such environments have been selected during protein evolution (as opposed to a purely "physical model" in which those features would be derived from some kind of energetics analysis of the protein structural characteristics). To illustrate this novel approach and provide general guidelines for its application, we develop here a simple evolutionary model that successfully explains the effect of the I↔V mutations on thioredoxin stability.

INTRODUCTION

Since the advent of genetic engineering, numerous physicochemical studies into the effects of mutations on proteins have been carried out. The interpretation of such studies in terms of the physical forces responsible for the native protein structure has provided a basis for many current efforts in protein design.

However, many features of natural proteins are expected to be the result of natural selection, which could also operate on stability. In fact, the marginal thermodynamic stability of natural proteins (experimental values of the order of a few tens of kilojoules per mol for the denaturation free energy at physiological temperature) has been sometimes suggested to have been selected during evolution due to several potential advantages (1). For instance, low stability may be associated with the degree of flexibility perhaps required for function, to facilitate degradation, and to help avoid trapping the protein in incorrectly folded states during folding (2). Also, a certain number of specific destabilizing interactions may be required for functional reasons or to guarantee structural uniqueness (3,4).

More surprisingly, recent work highlights that the relation between protein evolution and protein stability is apparent even at the level of individual-residue environments. Thus, we reported recently (1) that all glutamate (E) to aspartate (D) mutations in wild-type *Escherichia coli* thioredoxin were destabilizing, as well as most of the aspartate to glutamate mutations. We also found a robust correlation between the effects of these mutations on thioredoxin stability and the frequencies of occurrence of the involved residues in several hundred sequence alignments derived from a BLAST search. These results indicate that environments of charged residues in protein surfaces may be optimized for stabilizing interactions to a remarkable degree of specificity (to the point of discriminating between glutamate and aspartate).

Here, we study the effect of mutations involving isoleucine (I) and valine (V) on the stability of *E. coli* thioredoxin and compare the measured mutation effects with the frequency of occurrence of I and V residues in sequence alignments. These I→V and V→I mutations are expected to be conservative and have very small effects on protein structure (in a well-packed hydrophobic region, an I→V mutation is expected to create a very small cavity and a V→I mutation is expected to introduce a small amount of strain). In fact, I and V residues can easily exchange in the course of protein evolution as shown by the I/V coefficients in PAM and BLOSUM matrices (5).

The study of the effect of I→V and V→I mutations on stability allows us to explore whether the type of evolutionary optimization we previously found for exposed charged residues may also occur in protein hydrophobic cores. In addition, and most important, the analysis of hydrophobic mutations leads naturally to a structural interpretation of the evolutionary optimization. The reason is that the environment of buried hydrophobic residues can be described, as a first approximation at least, in terms of a single parameter: the hydrophobic packing density. (This is in sharp contrast with the description of the environment of charged residues, which must necessarily involve several parameters, related to long-distance electrostatic interactions, extent of charged-atoms solvation, ion pairing, hydrogen bonding, etc.)

Indeed, we find in this work that the plot of effect of hydrophobic mutations on stability versus packing density shows a strikingly simple pattern that clearly reflects evolutionary structural optimization. The existence of such a

Submitted May 23, 2005, and accepted for publication July 28, 2005.

Address reprint requests to Jose M. Sanchez-Ruiz, Facultad de Ciencias, Departamento de Quimica Fisica, Universidad de Granada, Campus Fuentenueva s/n, 18071 Granada, Spain. Tel.: 34-958243189; E-mail: sanchezr@ugr.es.

0006-3495/05/11/3320/12 $2.00

doi: 10.1529/biophysj.105.067025

pattern suggests that mutation effects on protein stability can be interpreted (and potentially predicted) on the basis of evolutionary models (in addition to the usual physical models). By ''evolutionary model'' we specifically mean in this context a model in which certain physical features of the residue environments are evaluated from an assumption regarding how such environments have been selected during protein evolution (as opposed to a purely ''physical model'' in which those features would be derived from some kind of energetic analysis of the protein structural characteristics).

To illustrate this novel approach, we develop in this work a simple model to explain the effect of hydrophobic mutations on protein stability, which takes into account the hydrophobic packing density around the mutation site in the native structure, as well as the kind and degree of evolutionary optimization of the mutated-residue environment. We find such a model to be successful in accounting for the effect of both the cavity-creating I→V mutations and the strain-introducing V→I mutations on stability. From a more general viewpoint, our analysis suggests guidelines as to how the role of evolutionary optimization in protein folding and stability can be systematically investigated.

MATERIALS AND METHODS

Sequence alignments

BLAST 2 (1996–2003, W. Gish, http://blast.wust1.edu) was used to search the UniProt/TrEMBL database (http://www.ebi.ac.uk/trembl/) with the thioredoxin sequence as query and the default options of the search. The sequences found were aligned to the query sequence using the Smith-Waterman algorithm (see Appendix 1 in Supplementary Material) and those with similarity with the query higher than 0.25 were retained (similarities were calculated as the number of matches between the given sequence and the query divided by the number of residues in the latter). The 0.25 similarity cutoff was chosen because it is usually accepted that proteins from various species and having sequence similarity of at least 0.25–0.3 have similar tridimensional structures (see Dokholyan and Shakhnovich (6) and references quoted therein). Therefore, we assumed that most of the sequences selected share the thioredoxin fold. Furthermore, for most of the 491 selected sequences, the parameter p that measures the statistical significance of the alignment with the query was reported by the BLAST program to be $<10^{-6}$ (specifically, among the 491 sequences, there are 428 sequences with $p < 10^{-6}$, 479 sequences with $p < 10^{-4}$ and 489 sequences with $p < 10^{-2}$). Among the 491 selected sequences, 219 belong to bacteria, 249 to eukaryota, and 18 to archaea (for five sequences, the biological classification was not available in the UniProt/TrEMBL database). Within the eukaryota, there are 123 metazoa, 81 viridiplantae, and smaller numbers of others (fungi, alveolata, etc.). Within the bacteria, there are 101 proteobacteria, 51 firmicutes, and smaller numbers of others (cyanobacteria, actinobacteria, etc.). Only 62 sequences of the proteobacteria sequences correspond to γ-proteobacteria (the class to which *E. coli* belongs).

Site-directed mutagenesis, expression, purification, and preliminary characterization of thioredoxin variants

Preparation of wild-type (WT) and variant forms of *E. coli* thioredoxin were carried out as we have recently described (1,7), except for the fact that we have used here plasmid pET30a (into which the thioredoxin gene had been subcloned) and BL21(DE3) supercompetent cells for overexpression.

Stability determinations

Differential scanning calorimetry (DSC) experiments were carried out with a VP-DSC calorimeter from MicroCal (Northhampton, MA) as we have described previously in detail (1,7,8). Protein solutions for the calorimetric experiments were prepared by exhaustive dialysis against the buffer (5 mM HEPES, pH 7.0). A protein concentration dependence for WT thioredoxin denaturation temperature has been reported in the literature and attributed to protein dimerization (9). Therefore, we carried out all the DSC experiments at comparatively low protein concentrations (~0.5 mg/mL or below) and we checked that no protein concentration effects on denaturation occurred within the 0–0.5 mg/mL range. Denaturation of all thioredoxin variants studied was highly reversible and fittings of the two-state equilibrium model to the heat capacity profiles were excellent and similar to that we have previously described for WT thioredoxin (8). It is worth noting that thermal denaturation of thioredoxin has also been shown to conform to a two-macrostate scenario on the basis of the variable-barrier analysis recently proposed by Muñoz and Sanchez-Ruiz (10). Mutation effects on thioredoxin stability ($\Delta\Delta G$ values) were calculated from the mutation effects on denaturation temperature (ΔT_m values) by using the approximate Schellman equation, as we have previously described in detail (1,7). A more rigorous calculation, based on a Gibbs-Helmholz extrapolation, yields $\Delta\Delta G$ values that differed in ~0.1 kJ/mol from those calculated using the Schellman equation. Most $\Delta\Delta G$ and ΔT_m values given in this work are the average of three independent determinations, with typical scatters of the order of 0.1 K for ΔT_m.

Calculation of structural descriptors

Accessible surface areas (ASA) were calculated using a modification of the Shrake-Rupley algorithm (11), which randomly places 2000 points in the expanded van der Waals sphere representing each atom. A radius of 1.4 Å for the solvent probe and the Chothia set (12) for the protein atoms were used. Residue accessibilities were calculated as the ratio between the side-chain ASA in the native structure and that in Gly-X-Gly tripeptide.

Hydrophobic packing densities (η-values) were obtained for all positions with I or V in the WT protein according to the following procedure: we calculate, for each side-chain carbon atom of the residue (I or V) present at a given position, the number of other-residue carbon atoms (N_C) at distances smaller than 6 Å and then we average those numbers over all side-chain atoms. That is, we obtain a side-chain average of the number of carbon atoms within a 6-Å distance (η) as $[N_C(C_\beta) + N_C(C_{\gamma1}) + N_C(C_{\gamma2})]/3$ when V is present and $[N_C(C_\beta) + N_C(C_{\gamma1}) + N_C(C_{\gamma2}) + N_C(C_\delta)]/4$ when I is present.

RESULTS AND DISCUSSION

The effect of I→V and V→I mutations on thioredoxin stability

We have studied the effect on thioredoxin stability of all possible I→V (WT residue = I) and V→I (WT residue = V) mutations in the WT form. All positions mutated had low accessibility to solvent in the native structure, except I5 and V91, which are significantly exposed (Fig. 1). All V→I mutations were found to be destabilizing, as well as most of I→V mutations (Fig. 1). Therefore, despite the similarity in chemical structure (V and I differ only in a methyl group), the environments of most I/V positions in thioredoxin are optimized to a significant extent for stabilizing interactions with the specific residue type (I or V) actually present in the WT form. This is the same type of result as that we recently reported for the surface mutations involving carboxylic acid residues (1).

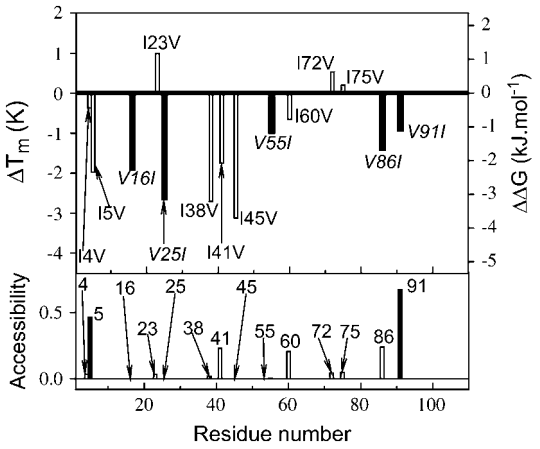

FIGURE 1 (*Upper panel*) Effect of isoleucine→valine and valine→isoleucine mutations on *E. coli* thioredoxin stability. Stability changes are measured by the effect of mutation on denaturation temperature (left axis) and denaturation free energy (right axis). I→V and V→I mutations are shown with open and closed bars, respectively. The numbers alongside the bars indicate the mutated residue in WT thioredoxin. (*Lower panel*) Side-chain accessibility to solvent in the native structure of the residues mutated. ASAs were calculated using a version the Shrake-Rupley algorithm with a radius of 1.4 Å for the solvent probe and the Chothia set for the protein atoms. The accessibilities shown are the ratio between the side-chain ASA values in the native structure and in a Gly-X-Gly tripeptide. Solid bars are used to denote residues with high accessibility (5 and 91).

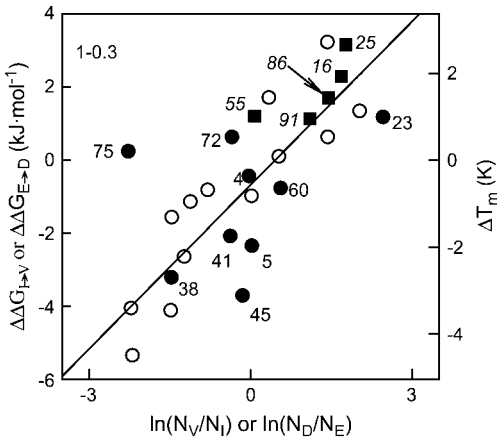

FIGURE 2 Correlation between effects of I→V mutations on thioredoxin stability (*solid symbols*) and the frequency of occurrence of I and V residues in sequence alignments. Stability changes are measured by the effect of mutation on denaturation temperature (left axis) and denaturation free energy (right axis). All the stability changes given describe the effect of I→V mutations. The solid circles correspond to the original I→V mutations in WT thioredoxin. The solid squares are derived from the original V→I mutations, but the sign of the stability changes has been reverted, so that the values shown correspond to I→V effects (for instance, $\Delta T_m = T_m$(WT, V) − T_m(variant, I)). The small numbers alongside the symbols indicate the position mutated. We also include (*open symbols*) the effects of carboxylic acid mutations (E→D) on thioredoxin stability we previously reported (1). The large numbers in both panels indicate the sequence similarity range used in the calculation of the number of residues: N_I, N_V, N_E, and N_D.

Note that, in subsequent analyses in this work, we will consider all mutations in the I→V direction; that is, the sign of the stability effects associated to the V→I mutations (with WT residue = V) is reverted, so that the values correspond to I→V effects. Of course, we will indicate in all relevant figures the type of residue (I or V) present at the given position in the WT form (see Fig. 2). As it will be made clear by the analyses given below, the purpose of considering all mutation effects on a given direction (I→V) is to expose environment optimization. For instance, a positive and sufficiently large value for the mutation effect $\Delta\Delta G_{I\rightarrow V}$ at a given position will be taken to mean that the environment at that position is optimized for V (over I), regardless of the specific residue (I or V) that is present at that position in the WT form (a specific example will be given further below).

Correlating mutation effects on stability with the frequencies of occurrence of residues in sequence alignments: a test of the pseudoequilibrium hypothesis

It has been suggested (1,6,13), mostly on theoretical grounds, that over some evolutionary timescale, sequences are visited according to free energy, in a manner analogous to the Boltzmann distribution of energies for a system at thermodynamic equilibrium. We refer to this proposal as the ''pseudoequilibrium'' or Boltzmann hypothesis. When applied to single, neutral (or quasineutral) mutations, the pseudoequilibrium hypothesis is equivalent to assuming that mutations become fixed in the course of evolution with frequencies that roughly reflect the mutation effect on stability. As we have previously pointed out, the pseudoequilibrium hypothesis is supported by the acceptable level of success of the so-called ''consensus concept'' to protein stabilization (14) that employs statistical analysis of sequence alignments to predict stabilizing mutations.

We recently showed (1) that the effect of glutamate→aspartate and aspartate→glutamate on the stability of *E. coli* thioredoxin show a robust correlation with the frequencies of occurrence of the involved residues in several hundred sequence alignments derived from a BLAST search, a result that is consistent with the pseudoequilibrium hypothesis. Here, we show that the same correlation holds for the hydrophobic mutations reported in this work. Furthermore, we perform in this work (see below) a more detailed characterization and testing of the pseudoequilibrium description of our data.

Our analysis of the correlation between mutation effects on stability and sequence alignments is based on the following equation:

$$\Delta\Delta G_{I(E)\rightarrow V(D)} = A + R\theta \cdot \ln\left(\frac{N_{V(D)}}{N_{I(E)}}\right). \quad (1)$$

In Eq. 1, $\Delta\Delta G_{I\rightarrow V}$ on the left side stands for the effect on stability of an I→V mutation at a given position. On the

right side of Eq.1, A is a constant (whose value is expected to be not too different from 0) and N_I (N_V) is the number of sequences with I (V) at the given position in sequence alignments derived from a BLAST search using the thioredoxin sequence as query (see Materials and Methods for details and for a description of the features of the alignment used; see Appendix 2 in Supplementary Material for the numbers of occurrences of all the 20 amino acids at the positions studied). The term $R\theta\cdot\ln(N_V/N_I)$ is a free-energy-like function that reflects the relative I versus V preference at the given position in the sequence alignments, and the temperature θ could be interpreted as a measure of evolutionary pressure on stability (6). Equation 1 is obviously suggested by the pseudoequilibrium hypothesis (1); that is, according to Eq. 1, if the environment of a given position is strongly optimized for V and consequently the I→V mutation is stabilizing, then during protein evolution V will be fixed at that position more often than I, resulting in a higher frequency of occurrence for V in sequence alignments. Likewise, if the environment is strongly optimized for I, we may expect the I→V mutation to be destabilizing and the frequency of occurrence in sequence alignments to be higher for I (versus V). Note finally that in this work we apply Eq. 1, not only to I→V mutations, but also to the carboxylic acid mutations we previously reported (1). In this latter case, the left side of Eq. 1 is the effect of the E→D mutation on stability ($\Delta\Delta G_{E\to D}$) and the numbers of sequences on the right side are calculated for D and E residues at the given position (N_D and N_E). This is indicated by the subscripts within brackets in Eq. 1.

The plot of $\Delta\Delta G_{I\to V}$ (or $\Delta\Delta G_{E\to D}$) versus $\ln(N_V/N_I)$ (or $\ln(N_D/N_E)$) of Fig. 2 shows a significant correlation between the mutation effects on stability and a simple function of the relative frequencies of occurrence of residues in sequence alignments: there is only one clear outlier (position 75) out of 27 mutation data and, when using for frequency of occurrence calculation the 247 sequences with similarity with the query higher than 0.3 (as in Fig. 2), the correlation coefficient is $r = 0.84$ and the statistical significance (probability that the observed correlation is due to chance) is $p = 6 \times 10^{-8}$. Actually, there is a likely explanation for the outlier: I75 is in close contact with the C32-C35 disulfide bridge in thioredoxin and, therefore, its high conservation could be due to functional (rather than stability-associated) reasons. The correlation is also a robust one; that is, a significant correlation is obtained when using the sequences with similarity with the query higher than 0.25 (491 sequences, $r = 0.85$, $p = 3 \times 10^{-8}$), 0.3 (247 sequences, $r = 0.84$, $p = 6 \times 10^{-8}$), 0.35 (144 sequences, $r = 0.83$, $p = 2 \times 10^{-7}$), or 0.4 (74 sequences, $r = 0.85$, $p = 5 \times 10^{-8}$). Also, significant correlations are obtained when carboxylic acid mutations and hydrophobic mutations are considered separately (for instance, $r = 0.89$ and $p = 6 \times 10^{-5}$ for carboxylic acid mutations and $r = 0.80$ and $p = 1 \times 10^{-3}$ for hydrophobic mutations, using the 491 sequences of similarity with the query higher than 0.25).

The intercept in the plot of Fig. 2 is close to zero (-0.69 ± 0.25 kJ/mol). The slope has units of energy and can be viewed as $R\theta$, where the value of the ''temperature'' θ might be interpreted as a measure of evolutionary pressure on stability or, perhaps, on physical features associated in some way with stability (see ''Concluding remarks''). The latter possibility (see further below) seems to be supported by the value found for θ (180 ± 23 K), which is significantly smaller than the physiological temperature. The above θ-value is calculated using sequences with the query higher than 0.3 (i.e., from the plot in Fig. 2). However, qualitatively similar results are obtained using other similarity cutoffs: 0.25 ($\theta = 211 \pm 26$ K, 491 sequences), 0.35 ($\theta = 150 \pm 21$ K, 144 sequences), and 0.4 ($\theta = 126 \pm 16$ K, 74 sequences).

The above calculations have been carried out using raw numbers of occurrences in the alignment (i.e., the N_I, N_V, N_D, and N_E numbers). In our previous work (1), we used similarity-weighted number of sequences, but because such weighting does not significantly affect the result of the analysis, it has been omitted here for the sake of simplicity. On the other hand, at the request of one reviewer, we have explored in this work the effect of the use of pseudocounts on the studied correlations. Pseudocounts have been previously used to correct raw amino acid frequencies when the number of sequences in the alignment is small and there is a risk of some of the raw frequencies becoming 0, a fact that may cause fatal problems in some algorithms (see Chapter 6 in Ewens and Grant (15)). The use of pseudocounts is based on Bayesian statistics, although the practical choice of the pseudocounts values is often done subjectively (see Chapter 6 in Ewens and Grant (15)). Here, we follow the procedure of Lawrence et al. (16). We define ''corrected'' frequencies as,

$$q_{ij} = \frac{N_{ij} + b_j}{N - 1 + B}, \tag{2}$$

where q_{ij} is the corrected frequency for amino acid j at position i, N_{ij} is the raw number of occurrences of amino acid j at position i, N is the total number of sequences in the alignments, b_j stands for the pseudocounts of amino acid j, and B is the sum of the pseudocount values for the 20 amino acids (i.e., $B = \sum_{j=1}^{20} b_j$). Following Lawrence et al. (16) we made pseudocounts proportional to the background frequencies of the amino acids with a proportionally constant of $\sqrt{N}$. As background frequencies we chose those calculated from the whole alignment (i.e., including all sequences and positions). We then used the calculated corrected frequencies (Eq. 2) instead of the raw numbers of occurrences in the analysis based on Eq. 1. The results obtained, however, were essentially the same as those obtained using raw numbers of occurrences. For instance, using all the sequences in the alignment we obtain $\theta = 216 \pm 27$ K, $r = 0.85$ and $p = 3 \times 10^{-8}$ (versus $\theta = 211 \pm 26$ K, $r = 0.85$, $p = 3 \times 10^{-8}$ using raw numbers of occurrences) and using only the sequences with similarity with the query higher than 0.4 we get $\theta = 143 \pm 19$ K, $r = 0.83$ $p = 1 \times 10^{-7}$ (versus $\theta = 126 \pm 16$,

$r = 0.85$ and $p = 5 \times 10^{-8}$ when using raw numbers of occurrences). This agreement is not surprising given the comparatively large number of sequences in our alignment.

To test further the applicability of the pseudoequilibrium hypothesis, we have carried out the correlation analyses based on Eq. 1 for different subsets of the alignment chosen according to the following two criteria: 1), sequence clustering according to similarity. For this purpose, we created a distance matrix for the 491 sequences in the alignment, using as ''distances'' the values of 1 minus pairwise similarity (i.e., Dayhoff distances), and we performed a K-means clustering analysis based on that matrix. We used the ''kmeans'' program included in the MATLAB suite, with the squared euclidean distance option and 10 replica in each analysis. Runs specifying different numbers of clusters (from two to nine) were performed; for illustration, the results obtained with five clusters are shown in Fig. 3. Note that significant correlations ($p < 10^{-2}$) in the plots of $\Delta\Delta G_{I\rightarrow V}$ (or $\Delta\Delta G_{E\rightarrow D}$) versus $\ln(N_V/N_I)$ (or $\ln(N_D/N_E)$) are obtained in all cases and, for three out of the five clusters, p is on the order of 10^{-5}. This general result is quite robust with respect to the number of clusters specified in the analysis (results not shown), although, for larger numbers of clusters, some of the clusters found are too small for a statistically reliable correlation analysis. 2), Sequence clustering according to the taxa well-represented in the alignment. Thus, plots of $\Delta\Delta G_{I\rightarrow V}$ (or $\Delta\Delta G_{E\rightarrow D}$) versus $\ln(N_V/N_I)$ (or $\ln(N_D/N_E)$) were constructed for the sequences belonging to bacteria (219 sequences), eukaryota (249 sequences), proteobacteria (101 sequences), and metazoa (123 sequences). In all four cases (see Fig. 4), significant correlations were found (p on the order of 10^{-5}–10^{-6}). For viridiplantae (81 sequences) and firmicutes (51 sequences) the correlations were less significant ($p = 1 \times 10^{-2}$ and $p = 7 \times 10^{-3}$, respectively), although this may be a statistical effect associated with the lower number of sequences belonging to these taxa in our alignment.

The above analyses strongly support the pseudoequilibrium hypothesis, at least as first-order description of the probability of mutation during evolution at the positions

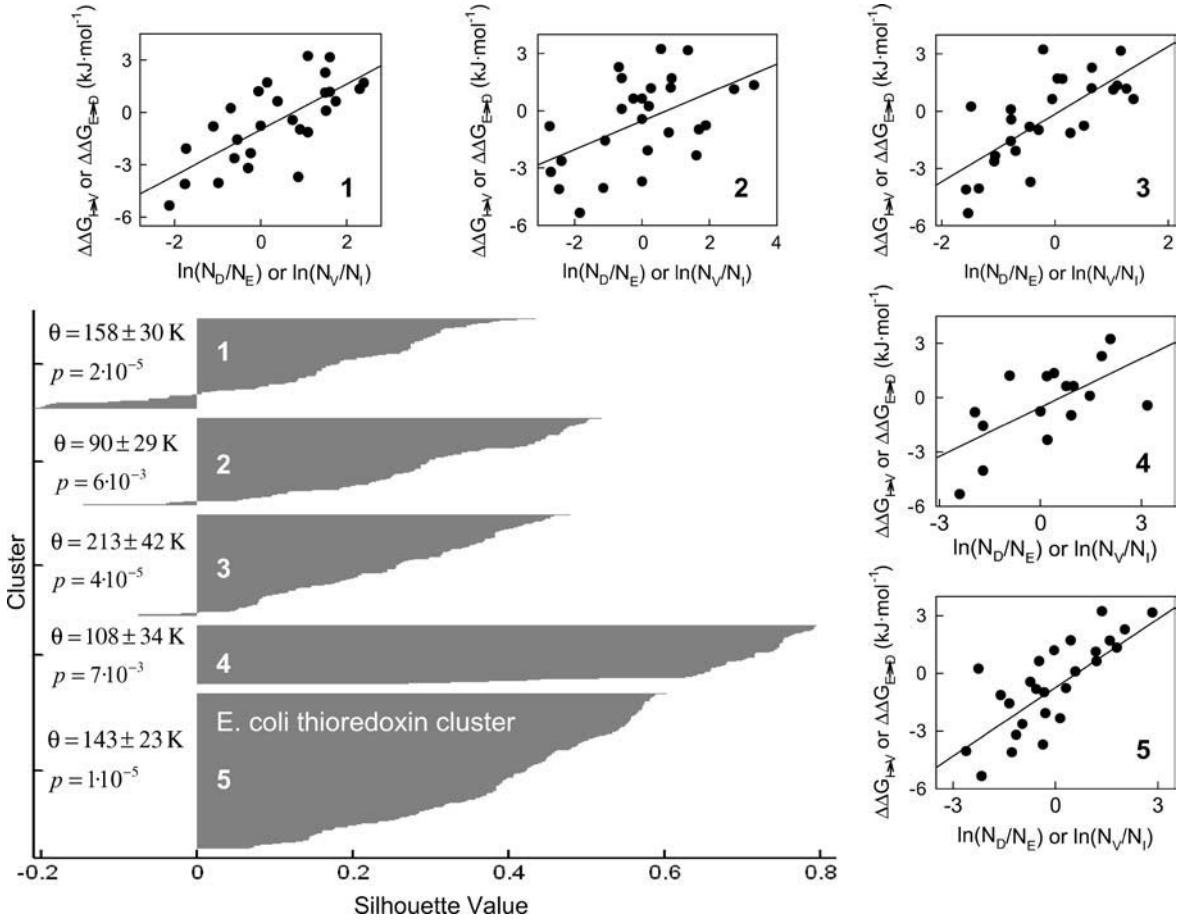

FIGURE 3 Correlation between mutation effects on *E. coli* thioredoxin stability and frequencies of occurrences of the involved amino acids for five sequence subsets obtained from the original alignment by K-means clustering. The larger plot gives the silhouette values for the sequences in each of the five clusters. The silhouette value for each sequence ranges from −1 to 1 and is a measure of how similar that sequence is to sequences in its cluster compared to sequences in other clusters. The numbers in the plots of mutation effect on stability versus logarithm of the ratio of raw occurrences refer to the cluster. The values of θ and p (statistical significance) derived from the five correlation analyses are shown alongside the corresponding cluster in the silhouette plot.

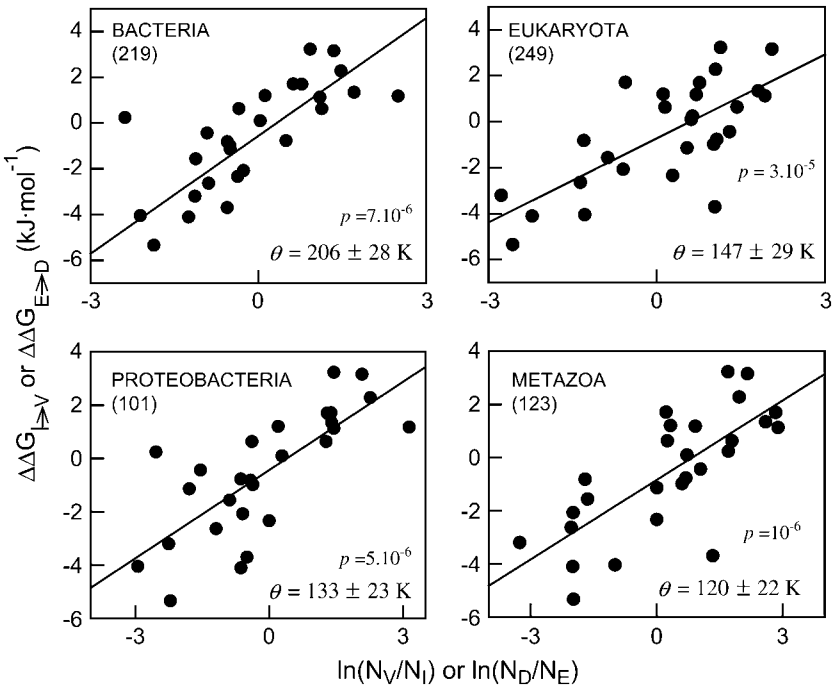

FIGURE 4 Correlation between mutation effects on *E. coli* thioredoxin stability and frequencies of occurrence of the involved amino acids for four sequence subsets corresponding to the taxa well-represented in the original alignment. We show in each plot of mutation effect on stability versus logarithm of the ratio of raw occurrences, the name of the taxa, the number of sequences (in parentheses), and the values of θ and p (statistical significance) derived from the correlation analysis.

studied. We find particularly relevant the correlation between mutation effects on stability of thioredoxin from *E. coli* (a proteobacteria) with the logarithm of the ratio of the number of occurrences of the involved residues in the eukaryota and metazoa subsets of the alignment (see Fig. 4). This clearly suggests that the type of environment optimization at a given position may be highly conserved during evolution (a detailed discussion on this proposal is given further below).

The experimental stability-packing pattern for hydrophobic mutations reflects evolutionary structural optimization

The results summarized in Figs. 1–4 indicate that hydrophobic environments can be optimized for stabilizing interactions to a very high degree of residue specificity (to the point of discriminating between I and V). Most remarkably, this evolutionary structural optimization is clearly reflected in the strikingly simple pattern (Fig. 5 *A*) seen in a plot of mutation effect on stability versus hydrophobic packing density (the side-chain average of the number of carbon atoms within a 6-Å distance; see Materials and Methods for details). For instance, it is clear from Fig. 5 *A* that, for high packing density, environments must be either optimized for V ($\Delta\Delta G_{I\rightarrow V} \gg 0$) or for I ($\Delta\Delta G_{I\rightarrow V} \ll 0$). In structural terms, if the environment of a given position is optimized (i.e., evolutionary selected) for the smaller V, an I in that position will cause overpacking and lead to a decrease in stability. Likewise, if the environment is optimized for I, an V will produce a cavity with the consequent loss of packing interactions leading again to a stability decrease (see, for example, Loladze et al. (17)). The relation between the evolutionary selection of residue environment and the mutation effects on protein stability has also been proposed by Takano et al. (18) on the basis of an analysis on the effect of Thr↔Val mutations on ribonuclease Sa stability.

Furthermore, according to Fig. 5 *A*, the environment capability to discriminate between V and I decreases with decreasing packing density and disappears at a packing density of ~10, where the V-optimized and I-optimized branches join (note that, as was to be expected, the valines in *E. coli* thioredoxin tend to cluster on the V-optimized branch). Only positions 5 and 91 depart significantly from this pattern, an unsurprising result given their high exposure to solvent (see Fig. 5 *A*), which indicates that their environment cannot be solely described in terms of hydrophobic packing density. The slightly negative value of $\Delta\Delta G_{I\rightarrow V}$ for the packing density at which the two branches meet (Fig. 5) likely reflects a general nonspecific preference for I over V associated with the extra methyl group in the former.

Overall, the plot of Fig. 5 *A* indicates a parabolic-like dependence of packing density with $\Delta\Delta G_{I\rightarrow V}$. In the next section we develop a simple model that rationalizes this dependence in terms of its physical and evolutionary origins.

A simple evolutionary model for the effect of hydrophobic mutations on protein stability

The existence of mutation stability patterns that reflect evolutionary optimization (such as that shown in Fig. 5 *A*) suggests the possibility that mutation effects on protein stability can be interpreted (and potentially predicted) on the basis of evolutionary models (in addition to the usual physical models). By ''evolutionary model'' we specifically

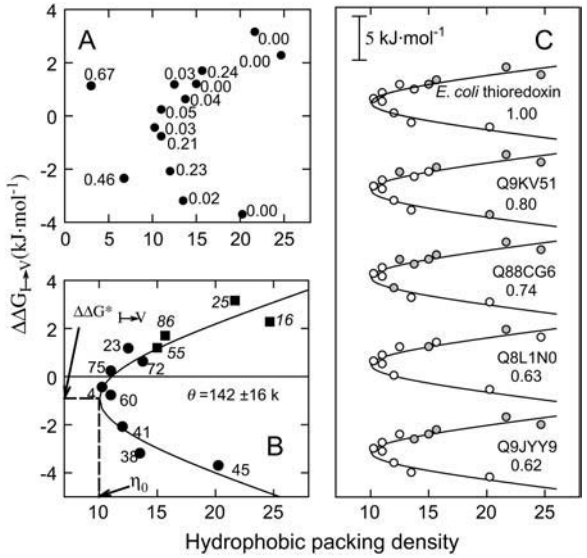

FIGURE 5 (*A*) Effect of hydrophobic mutations on stability versus hydrophobic packing density. The numbers alongside the points stand for the side-chain accessibilities to solvent in the native structure (a value of 0 means ''fully buried'' and a value of 1 means ''fully exposed''). (*B*) Fitting of Eq. 15 to the experimental mutation effects on thioredoxin stability. The numbers alongside the symbols indicate the position mutated. All mutations effects on denaturation free energy are given in the I→V direction and the type of residue present in the WT form is indicated by the symbol used: circles for Ile and squares for Val. The continuous line is the best fit of Eq. 15 to the experimental mutation data for the buried positions. (*C*) Distribution of I/V residues over hydrophobic positions for *E. coli* thioredoxin and four of the sequences in the alignments derived from a BLAST search using *E. coli* thioredoxin as query (see Materials and Methods). From top to bottom those proteins are thioredoxin from *E. coli*, Q9KV51 (thioredoxin from *Vibrio cholerae*), Q88CG6 (thioredoxin from *Pseudomonas putida*), Q8L1N0 (thioredoxin from *Buchnera aphidicola*), and Q9JYY9 (thioredoxin from *Neisseria meningitides*). The hydrophobic positions chosen are those in which I or V residues are present in WT thioredoxin from *E. coli*. In all cases, the I/V distribution (*open symbols/solid symbols*) is displayed over the stability/packing pattern for *E. coli* thioredoxin. The numbers alongside the protein codes stand for the similarity with the query sequence.

refer in this context to a model for mutation effects on stability in which certain physical features of the mutated residue environments are evaluated from an assumption regarding how such environments have been selected during protein evolution (as opposed to a purely ''physical model'' in which those features would be derived from some kind of energetic analysis of the protein structural characteristics). In this section, we illustrate how such models can be constructed using, as a case example, the hydrophobic I↔V mutations studied in this work.

We start by writing the effect of an I→V mutation at a given position on thermodynamic stability as:

$$\Delta\Delta G_{\mathrm{I}\to\mathrm{V}} = \Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}} - (G_{\mathrm{V}}^{\mathrm{ENV}} - G_{\mathrm{I}}^{\mathrm{ENV}}), \qquad (3)$$

where $G_{\mathrm{V}}^{\mathrm{ENV}}$ and $G_{\mathrm{I}}^{\mathrm{ENV}}$ stand for the free energies associated with the environment in the native structure when V and I are present at the given position, and $\Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}}$ collects all contributions not related to the environment, denatured-state effects among them, for instance. (Actually, as we will discuss below, a small and roughly constant part of the nonspecific environment contribution to $\Delta\Delta G_{\mathrm{I}\to\mathrm{V}}$, is included in $\Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}}$).

We describe the environment of a given hydrophobic residue (I or V) in the native structure in terms of three features:

1. The hydrophobic packing density in the neighborhood of the residue (η, the side-chain average of the number of carbon atoms within a 6-Å distance; see Materials and Methods for details). We note that η is only meant to provide an overall measure of the packing density; finer structural details of the packing, such as those responsible for the optimization of the environment for V or I, are described by the parameter f introduced below.
2. The extent to which the environment is optimized for interactions with I or V. We describe this feature with a single parameter f, which takes a value of 0 when the environment is perfectly optimized for I and a value of unity when the environment is perfectly optimized for V. Thus, we write the free energy associated with the interaction of I with the environment as $\alpha_0 \cdot (\eta_{\mathrm{m}} - \eta) + \alpha \cdot \eta \cdot f$, where η_{m} is the maximum value of η and α_0 and α are constants. That is, we assume a nonspecific contribution related to the overall packing in the region around the I-residue $[\alpha_0 \cdot (\eta_{\mathrm{m}} - \eta)]$ and a contribution related with the environment optimization $[\alpha \cdot \eta \cdot f]$. The optimum I-environment interaction occurs when the overall packing is the maximum possible (i.e., $\eta = \eta_{\mathrm{m}}$) and the environment is fully optimized for I ($f = 0$); in this case, the free energy of interaction takes a minimum possible value of 0. The free energy associated with the interaction of V with the environment is likewise written as $\alpha_0 \cdot (\eta_{\mathrm{m}} - \eta) + \alpha\eta \cdot (1 - f)$ and the optimum V-environment interaction (free energy of interaction equal to zero) occurs when $\eta = \eta_{\mathrm{m}}$ (maximum overall packing) and $f = 1$ (environment fully optimized for V). Note that we use the same expression of the nonspecific interaction term for V and I $[\alpha_0 \cdot (\eta_{\mathrm{m}} - \eta)]$, although this contribution could be expected to be somewhat larger for I (due to the extra methyl group). Actually, in the spirit of proposing a simple model, we are assuming here that the small difference between V and I in nonspecific interaction with the environment is roughly constant and can be included in $\Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}}$ (Eq. 3). Note also that, unlike η, we do not calculate the value of f from the native structure but, in our model, environment optimization is determined on the basis of a hypothesis regarding the effect of natural selection (see further below).
3. The environment ''plasticity''. We do not consider hydrophobic environments to be rigid, but, rather, we take into account that they may change in response to mutation. We thus include a free-energy term associated to environment reorganization and given by $\beta \cdot \eta \cdot (f - f_0)^2$,

where β is a constant. The reorganization free energy is zero when $f = f_0$ and, therefore, the value of f_0 may be considered as a measure of the ''intrinsic'' preference of the environment for I or V. Of course, values of f different from f_0 may be observed if the required reorganization free energy is ''paid'' by the interaction residue environment (see below). Note that this reorganization free energy will be larger the higher the value of η; that is, well-packed environments are assumed in our model to be more ''rigid''.

The environment free-energy terms in Eq. 3 (G_V^{ENV} and G_I^{ENV}) contain, therefore, contributions from the direct interaction of the residue (I or V) with the environment and also from the possible environment reorganization:

$$G_I^{ENV}(f) = \alpha_0 \cdot (\eta_m - \eta) + \alpha \cdot \eta \cdot f + \beta \cdot \eta \cdot (f - f_0)^2 \quad (4)$$

$$G_V^{ENV}(f) = \alpha_0 \cdot (\eta_m - \eta) + \alpha \cdot \eta \cdot (1 - f) + \beta \cdot \eta \cdot (f - f_0)^2. \quad (5)$$

These expressions are taken to be free-energy functionals. That is, the actual values of G_V^{ENV} and G_I^{ENV} are obtained by minimizing Eqs. 4 and 5 with respect to f. Fig. 6 provides graphical illustrations of such minimization. Panels *A*, *B*, and *C* correspond to the case of a rigid (high η) environment that is optimized for I (*A*), for V (*B*), or has no definite intrinsic optimization ($f_0 = 1/2$; *C*); the situation of a ''plastic'' environment (low η) and no definite intrinsic optimization ($f_0 = 1/2$) is shown in panel *D*. Note that the optimum values of f (i.e., the values f_V and f_I that minimize G_V^{ENV} and G_I^{ENV}, respectively) are determined by the interplay between the direct interaction terms and the term associated to environment reorganization. Note also that the optimum f_I and f_V values differ from f_0, as the environment changes in response to the interaction with the residue. General expressions for f_I and f_V can be obtained by solving $\partial G_I^{ENV}/\partial f = 0$ and $\partial G_V^{ENV}/\partial f = 0$, respectively. The results are $f_I = f_0 - \alpha/2\beta$, $f_V = f_0 + \alpha/2\beta$. Note that, because f is in the $0 \le f \le 1$ range, f_0 is restricted in our model to the $\alpha/2\beta \le f_0 \le 1 - \alpha/2\beta$ range. Substitution into Eqs. 4 and 5 yields:

$$G_I^{ENV} = \alpha_0 \cdot (\eta_m - \eta) + \alpha\eta \cdot (f_0 - 1/2) + \frac{\alpha\eta}{2}\left(1 - \frac{\alpha}{2\beta}\right) \quad (6)$$

$$G_V^{ENV} = \alpha_0 \cdot (\eta_m - \eta) - \alpha\eta \cdot (f_0 - 1/2) + \frac{\alpha\eta}{2}\left(1 - \frac{\alpha}{2\beta}\right). \quad (7)$$

We will find it convenient for subsequent derivations to have G_I^{ENV} and G_V^{ENV} referred to their lowest possible values (those corresponding to $f_0 = \alpha/2\beta$ and $f_0 = 1 - \alpha/2\beta$, respectively):

$$\Delta G_I^{ENV} = G_I^{ENV} - G_I^{ENV}(\alpha/2\beta) = \alpha\eta \cdot (f_0 - 1/2) + \frac{\alpha\eta}{2}\left(1 - \frac{\alpha}{\beta}\right) \quad (8)$$

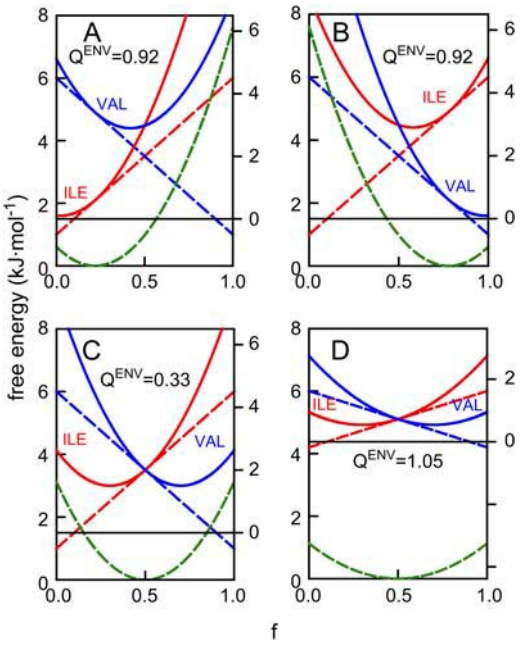

FIGURE 6 Illustrative examples of the free energies associated with the environment when V and I residues are present at a given position, as calculated from the simple model introduced in this work (Eqs. 3–15 in the text). The values of G_I^{ENV} and G_V^{ENV} (Eqs. 4 and 5) are given versus f (the parameter that describes environment optimization) as thick red and blue lines, respectively. The thin-dashed red and blue lines represent the contributions to G_I^{ENV} and G_V^{ENV} associated with the direct residue-environment interaction (that is, the $\alpha_0 \cdot (\eta - \eta_m) + \alpha \cdot \eta \cdot f$ and $\alpha_0 \cdot (\eta - \eta_m) + \alpha \cdot \eta \cdot (1 - f)$ terms in Eqs. 4 and 5, respectively). The thin-dashed green line is the contribution associated to environment reorganization (the term $\beta \cdot (f - f_0)^2$ in both Eqs. 4 and 5). The profiles shown have been calculated with the following values: $\alpha_0 = 0.2$ kJ/mol, $\alpha = 0.2$ kJ/mol, $\beta = 0.5$ kJ/mol, $\eta_m = 30$ (all panels). In panels *A*, *B*, and *C* we have used $\eta = 25$ (a high-packing environment) and, in panel *D*, $\eta = 9$ (a low-packing environment). The values of f_0 are 0.22 in panel *A* (environment with a clear intrinsic preference for I), 0.78 in panel *B* (environment with a clear intrinsic preference for V), and 0.5 in panels *C* and *D* (environment with no clear intrinsic preference). The thick black line represents the lowest possible values of G_I^{ENV} and G_V^{ENV} in each case (note that those lowest possible values are taken as zero for the right-axis free-energy scales). The difference between the minima in the thick red and blue lines and the black line indicates the optimization achieved. Thus, optimization for I is achieved in panel *A*, optimization for V is achieved in panel *B*, and for, both I and V in panel *D*; in these three cases, the Q^{ENV} partition function (Eq. 11) is on the order of unity. In panel *C* no clear optimization for either I or V is achieved and Q^{ENV} is significantly smaller than unity.

$$\Delta G_V^{ENV} = G_V^{ENV} - G_V^{ENV}(1 - \alpha/2\beta) = -\alpha\eta \cdot (f_0 - 1/2) + \frac{\alpha\eta}{2}\left(1 - \frac{\alpha}{\beta}\right). \quad (9)$$

Substitution of Eqs. 6 and 7 into Eq. 3 leads to,

$$\Delta\Delta G_{I\to V} = \Delta\Delta G^*_{I\to V} + 2\alpha\eta \cdot (f_0 - 1/2), \quad (10)$$

which gives the mutation effect on stability in terms of the environment preference for I or V, as measured by the value of f_0.

We now proceed to make a hypothesis regarding the restrictions imposed by natural selection on residue environments. We assume that environments of I/V residues are ''accepted'' during evolution if optimization for at least one kind of residue (I or V) is achieved. According to this hypothesis, the situation depicted in Fig. 6 *A* would be accepted because the environment leads to clear optimization for I (i.e., $G_{\mathrm{I}}^{\mathrm{ENV}}$ is close to its lowest possible value and $\Delta G_{\mathrm{I}}^{\mathrm{ENV}}$ is close to 0); of course, optimization for V is not achieved, but this poses little problem because, according to the pseudoequilibrium hypothesis, V will rarely appear at that position during protein evolution (perhaps, only when stabilizing mutations in other parts of the molecule compensate for the destabilizing effect of having V in a well-packed environment optimized for I). Likewise, the situation shown in Fig. 6 *B* will be accepted (optimization for V is achieved). However, the rigid environment with no clear intrinsic preference for I or V ($f_0 = 1/2$) of Fig. 6 *C* will be rejected, because it cannot be optimized for either I or V. On the other hand, a plastic environment (low-packing) with no clear preference (as in Fig. 6 *D*) will be accepted, because it reorganizes upon interaction with I and V, leading to optimization for both types of residues.

The above qualitative reasoning can be expressed in mathematical terms by stating that the following partition function must have a sufficiently high value:

$$Q^{\mathrm{ENV}} = \exp\left(-\Delta G_{\mathrm{I}}^{\mathrm{ENV}}/R\theta\right) + \exp\left(-\Delta G_{\mathrm{V}}^{\mathrm{ENV}}/R\theta\right) = \left\{\exp\left(-\frac{\alpha\eta}{2R\theta}(1-\alpha/\beta)\right)\right\}\cdot\left\{\exp\left(-\frac{\alpha\eta}{R\theta}(f_0-1/2)\right) + \exp\left(\frac{\alpha\eta}{R\theta}(f_0-1/2)\right)\right\} = 2\cdot\left\{\exp\left(-\frac{\alpha\eta}{2R\theta}(1-\alpha/\beta)\right)\right\}\cosh\left(\frac{\alpha\eta}{R\theta}(f_0-1/2)\right), \quad (11)$$

where, in the first equality we have used Eqs. 8 and 9 and in the second the definition of the hyperbolic cosine: $\cosh(x) = (e^{x} + e^{-x})/2$ (see Chapter 3 in Steiner (19)). Under the assumption that, at least one of the free-energy terms ($\Delta G_{\mathrm{I}}^{\mathrm{ENV}}$ or $\Delta G_{\mathrm{V}}^{\mathrm{ENV}}$) is close to zero, we may expect the value of Q^{ENV} to be close to unity (as illustration, see the Q^{ENV} values in Fig. 6). Of course, we do not state that natural selection keeps $Q^{\mathrm{ENV}} = 1$ strictly; we do expect, however, to obtain a reasonable approximation to the evolutionary accepted environments (as described by the f_0 value) by setting $Q^{\mathrm{ENV}} = 1$ in Eq. 11 and solving for f_0:

$$f_0 = \frac{1}{2} + \frac{R\theta}{\alpha\eta}\cosh^{-1}\left\{\frac{1}{2}\exp\left(\frac{\alpha\eta}{2R\theta}(1-\alpha/\beta)\right)\right\}, \quad (12)$$

where $\cosh^{-1}$ stands for the inverse hyperbolic cosine. This function can be easily calculated (see Chapter 3 in Steiner (19)) from $\cosh^{-1}(x) = \ln\left(x \pm \sqrt{x^2-1}\right)$ and, for $x > 1$, it has two values that only differ in sign. For $x = 1$ it has only one value (zero) and the function is not defined for $x < 1$. This implies that in our model there must be a minimum value for η (we call it η_0), which can be easily found by setting equal to 1 the argument of the $\cosh^{-1}$ in Eq. 12 and solving for η_0:

$$\eta_0 = \frac{2\cdot\ln 2\cdot R\theta}{\alpha\cdot(1-\alpha/\beta)}, \quad (13)$$

and using Eq. 13, Eq. 12 can be written as:

$$f_0 = \frac{1}{2} + \frac{R\theta}{\alpha\eta}\cosh^{-1}\{2^{(\eta-\eta_0)/\eta_0}\}, \quad (14)$$

which, upon substitution into Eq. 10, yields finally:

$$\Delta\Delta G_{\mathrm{I}\to\mathrm{V}} = \Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}} + 2R\theta\cdot\cosh^{-1}\{2^{(\eta-\eta_0)/\eta_0}\}. \quad (15)$$

According to Eqs. 14 and 15, for $\eta = \eta_0$, there is only one possible value of f_0 (1/2) and only one possible value for the mutation effect on stability ($\Delta\Delta G_{\mathrm{I}\to\mathrm{V}} = \Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}}$). Therefore, η_0 may be interpreted as defining a low-packing density for which environment optimization is not possible. On the other hand, for each η-value higher than η_0, there are two possible values of f_0 ($f_0 > 1/2$ and $f_0 < 1/2$) and two possible values of $\Delta\Delta G_{I\to V}$ ($\Delta\Delta G_{\mathrm{I}\to\mathrm{V}} > \Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}}$ and $\Delta\Delta G_{\mathrm{I}\to\mathrm{V}} < \Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}}$), which obviously correspond to environment optimization for V and for I, respectively.

It is important to note that many of the details of our model are not apparent in Eq. 15, which is, in fact, strikingly simple. We take this as indication of the robustness of the equation. For instance, the reader may easily verify that using a general function of η, instead of $\alpha_0\cdot(\eta - \eta_{\mathrm{m}})$, or neglecting environment reorganization (that is, omitting the term $\beta\cdot(f-f_0)^2$ and the subsequent minimization of $G_{\mathrm{I}}^{\mathrm{ENV}}$ and $G_{\mathrm{V}}^{\mathrm{ENV}}$ with respect to f) leads to the same Eq. 15. The latter, however, would have been inconsistent with experimental structural data that show variable degrees of cavity collapse upon mutating a large hydrophobic side chain to a smaller one (20). Therefore, the approach we have taken has been to include environment reorganization in the analysis in a simple but plausible manner (Eqs. 4 and 5) and then to show that the general outcome of the theoretical analysis (Eq. 15) does not actually depend on environment reorganization.

Fitting of the model to the thioredoxin mutation data

Equation 15 has only three fitting parameters ($\Delta\Delta G^{*}_{\mathrm{I}\to\mathrm{V}}$, η_0, and θ) and its nonlinear least-squares fitting to the experimental $\Delta\Delta G_{\mathrm{I}\to\mathrm{V}}$ versus η data is straightforward.

We fitted Eq. 15 to the experimental $\Delta\Delta G_{I\rightarrow V}$ data, excluding the mutations at two positions with large exposure to the solvent (5 and 91; see Fig. 5 *A*). The fit to the data for the remaining 12 buried positions is excellent (Fig. 5 *B*). As was to be expected, the mutation data for the two well-exposed positions do not agree with the fitted dependence, which indicates that their environment cannot be solely described in terms of hydrophobic packing density. The fitting is quite robust with respect to the choice of the parameter used to measure packing: although we have used for η the average number of side-chain carbon atoms within a distance of 6 Å (see previous section), we have checked that increasing the distance limit to 8 Å or including all side-chain atoms (not only carbons) yields similar fittings (results not shown).

The values of the fitting parameters derived from the fitting shown in Fig. 5 *B* are $\eta_0 = 10 \pm 0.2$, $\Delta\Delta G^*_{I\rightarrow V} = -0.9 \pm 0.3$ kJ/mol and $\theta = 142 \pm 16$ K (associated errors were calculated by the Monte Carlo method). The value of the temperature θ is in good agreement with those derived from the several correlation analysis reported in this work (see Figs. 2–4). This agreement provides clear support for the validity of the main features of our model.

The small but negative value obtained for $\Delta\Delta G^*_{I\rightarrow V}$ indicates a general preference for I versus V. As pointed out above, the nonspecific (independent of optimization) interaction of the extra methyl group in isoleucine with the environment may contribute to the $\Delta\Delta G^*_{I\rightarrow V}$ value, as well as factors related with solvent exposure of the residues in the denatured state, including residual structure (hydrophobic clustering). In connection with this latter possibility, it must be noted that a strongly ''polarized'' thermally denatured state (i.e., a denatured state with well-defined hydrophobic clustering in certain regions and high exposure to the solvent in others), would have likely led to different $\Delta\Delta G^*_{I\rightarrow V}$ values for positions in different protein regions and the fitting of Eq. 15 to the experimental data with constant $\Delta\Delta G^*_{I\rightarrow V}$ would have failed. To further probe this interpretation, we have employed the approach suggested by Marqusee and co-workers (21) in which residual structure in the denatured state is detected through its disruption (detected by heat capacity determinations) induced by a mutation that introduces a charged residue. We thus prepared four variants of thioredoxin with mutations V16D, L42D, A67D, and L99D, which are located in the buried sides of the α-helices 11–17, 33–49, 67–70, and 94–105. These mutations target hydrophobic regions located at both sides of the central β-sheet in the native structure, but we did not detect any significant increase in denatured-state heat capacity over that of the WT protein (results not shown).

Environment conservation versus residue conservation

The data reported in this work indicate that many of the hydrophobic environments studied are optimized to a significant extent for either V or I but, also, that during the course of protein evolution, I may be accepted in environments optimized for V (and vice versa), with frequencies that roughly reflect the associated free-energy penalties. For instance, the environment of position 23 is optimized for V, as indicated by a positive experimental value for $\Delta\Delta G_{I\rightarrow V}$. In fact, during protein evolution, V appears more often than I at that position (152 V versus 13 I in the sequences with similarity with the query higher than 0.3). This general preference of V over I at position 23 is also observed when analyzing separately the taxa well-represented in the alignment: 107 V versus four I for bacteria, 37 V versus eight I for eukaryota, 52 V versus two I for proteobacteria, and 15 V versus four I for metazoa. Clearly, position 23 is a ''valine-optimized position'', although there is an I at that position in *E. coli* thioredoxin as a result of a recent (in the evolutionary sense) mutation. This interpretation is reminiscent of the recently proposed evolutionary hypothesis for the generation of transmembrane helix kinks (22), according to which nonproline kinks in membrane proteins are created by vestigial prolines and stabilized by second-site mutations that optimize packing around the kink, after which replacement of the proline during the course of evolution occurs leaving the nonproline kink; indeed, at positions of nonproline kinks in given proteins, proline residues were found to occur often in related sequences (22).

What is more remarkable, however, about the results presented in this work is that they indicate that hydrophobic environments can be optimized for stabilizing interactions to a very high degree of residue specificity; i.e, to the point of discriminating between I and V, which only differ in a methyl group (note that the structural consequences of I↔V mutations are expected to be much smaller than those of introducing a proline in a helix). This specificity may be the consequence of the need for good packing in certain protein regions, which requires a tight fit of the involved side chains. Thus, if good packing around a given position was originally (in the evolutionary sense) achieved with, for instance, valine in this position, the environment will be optimized for V. Certainly, the neutral or quasineutral V→I mutation will be occasionally accepted over the course of evolution (as in position 23 in thioredoxin) as a result of random drift. However, the kind of environment optimization is less likely to change (from optimization for V to optimization for I), because this will possibly require the combined effect of several second-site mutations (23).

From all the above, we conclude that, as a first approximation at least, we may consider hydrophobic-residue environments as more evolutionary conserved than residues themselves. This implies, for instance, that the $\Delta\Delta G_{I\rightarrow V}$ versus packing pattern (Fig. 5, *A* and *B*) should be the same (or similar) for evolutionary related proteins, although the distribution of I and V residues may differ to some extent. This interpretation is illustrated in Fig. 5 *C* where we display on the stability/packing pattern for thioredoxin, the I/V

distribution for several of the sequences obtained in the BLAST search using the thioredoxin as query.

On the ruggedness of the stability-sequence landscape (a Japanese rock garden analogy)

Evolutionary optimization of residue environments may perhaps occur to different extents in different proteins (mesophilic versus thermophilic, for instance) and, for the same protein, in different structural regions (due, for instance, to the effect of functional or interaction constraints). This notwithstanding, the results reported here suggest that, in some proteins at least, a significant number of residue environments are evolutionary optimized for stabilizing interactions with specific amino acid residues and that residues not matched with the environment occur rarely (with frequencies related to the associated free-energy penalties, according to the pseudoequilibrium hypothesis). However, the evolutionary optimization of potentially many residue environments in proteins does not mean that overall protein stability is close to its global maximum (which would perhaps seem in conflict with the fact that the thermodynamic stability of natural proteins is marginal). In our view, it only means that the stability-sequence landscape is rugged, in such a way that most single mutations are destabilizing. We believe that an illustrative (albeit imperfect) analogy to such stability-sequence landscape is provided by some Japanese rock gardens (such as those of Ryogen-in and Tofuku-ji in Kyoto; for pictures see http://phototravels.net/japan/photo-gallery/japanese-rock-gardens.html). The sand level represents the minimum stability required for proper folding (for the purpose of this analogy, we disregard the patterns in the sand) and the rocks symbolize ''stability islands'', the top of each rock being a local optimum in stability (note that the top of a small rock will necessarily be close to the sand level). To any given sequence of a natural protein, a definite spot on a rock (between the sand level and the top) may be assigned and single mutations are assumed to move the spot on a given rock. If the WT spot is already close to the top, only a few single mutations will be stabilizing and, in any case, these mutations will not increase stability over the local optimum (the top of the rock). Of course, larger stability enhancements may be achieved, but the sequence spot needs to be moved to a different, taller rock, which requires the combined effect of several mutations (that is, some of those mutations may be destabilizing when they occur individually). Actually, it appears feasible to obtain high stability, severalfold mutant proteins in vitro through the use of adequate protein-design computational procedures (24–26).

Concluding remarks

We show in this work, by providing a simple and clear experimental example, that stability patterns of protein mutations may reflect structural evolutionary optimization. We believe it likely that such kind of stability patterns occur often. They have not been previously detected (to the best of our knowledge) possibly because exhaustive studies on conservative mutations involving two similar residues (for instance, all $I \rightarrow V$ and $V \rightarrow I$ mutations in a given protein) are not usually performed. The existence of these patterns suggests the possibility of rationalizing (and perhaps even predicting) mutation effects on protein stability on the basis of evolutionary models. By ''evolutionary models'' we specifically mean in this context models for mutation effects on stability in which some of the physical characteristics that determine the mutation effect are evaluated from an assumption regarding how the mutated-residue environment has been selected during protein evolution. To illustrate this novel approach and provide general guidelines for its application, we have developed here one such evolutionary model that successfully accounts for the effect of the $I \leftrightarrow V$ mutations on thioredoxin stability. In addition, the model could potentially be used to predict the effect of hydrophobic mutations on other proteins (work in progress); this prediction will likely require a simple procedure to estimate the temperature θ from the sequence alignments and the ability to determine the type of environment optimization from a rough analysis of the native structure (note that the environment optimization of a given position cannot be deduced with absolute certainty from the amino acid present at that position in the protein of interest, because, during evolution, amino acids ''not-matched'' with the environment are expected to be occasionally accepted with probabilities related with the associated free-energy penalties).

It only remains now to make some comments as to the origin of the energetic optimization of residue environments in proteins. As we have previously pointed out (1), optimization may be a consequence of the fact protein stability is marginal (27), although, in addition, it appears plausible that it may also reflect the natural selection of some other protein physical features associated in some way with stability. For instance, an interesting (although speculative at this stage) possibility is that residue optimization is related with the low level of frustration possibly required for folding efficiency (28) and with the evolutionary design of folding cooperativity. Thus, theoretical work (29) indicates that it may be more difficult for proteins to achieve cooperativity (i.e., a significant folding free-energy barrier) than a stable native structure and, in fact, computer-designed proteins have been found to fold faster than their natural templates, although no selection for a lower folding barrier was included in the design strategy (30). Theoretical analysis of protein models (29) also shows that models with more specific interactions are more cooperative and that cooperativity increases with increasing number of letters in the amino acid alphabet used. Our result that environments in proteins may be optimized to a remarkable degree of residue specificity (to the point of discriminating between V and I or between D and E) seems consistent with the view suggested by these polymer model

analyses and suggests that evolution may actually be taking full advantage of the natural 20-letter amino acid alphabet.

SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

Note added in proof: After this work had been accepted for publication, we became aware of the work of Chen and Stites (*Biochemistry*. 2001. 40:15280–15289), which also addresses the relation between hydrophobic packing and protein evolution.

We thank David Rodriguez-Larrea for helpful suggestions.

This work was supported by Spanish Ministry of Education and Science grant BIO2003-02229 and Feder Funds.

REFERENCES

1. Godoy-Ruiz, R., R. Perez-Jimenez, B. Ibarra-Molero, and J. M. Sanchez-Ruiz. 2004. Relation between protein stability, evolution and structure, as probed by carboxylic acid mutations. *J. Mol. Biol.* 336: 313–318.
2. Sanchez-Ruiz, J. M. 1995. Differential Scanning Calorimetry of Proteins. B. B. R. S. Biswas, editor. Plenum Press, New York, NY.
3. Lumb, K. J., and P. S. Kim. 1995. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry.* 34:8642–8648.
4. Knappenberger, J. E., J. E. Smith, S. H. Thorpe, J. Z. Zitzewitz, and C. R. Matthews. 2002. A buried polar residue in the hydrophobic interface of a coiled-coil peptide, GNC4-p1, plays a thermodynamic, not a kinetic role. *J. Mol. Biol.* 321:1–6.
5. George, D. G., W. C. Baker, and L. T. Hunt. 1990. Mutation data matrix and its uses. *Methods Enzymol.* 183:333–351.
6. Dokholyan, N. V., and E. I. Shakhnovich. 2001. Understanding hierarchical protein evolution from first principles. *J. Mol. Biol.* 312:289–307.
7. Perez-Jimenez, R., R. Godoy-Ruiz, B. Ibarra-Molero, and J. M. Sanchez-Ruiz. 2004. The efficiency of different salts to screen charge interactions in proteins: a Hofmeister effect? *Biophys. J.* 86:2414–2429.
8. Georgescu, R. E., M. M. Garcia-Mira, M. L. Tasayco, and J. M. Sanchez-Ruiz. 2001. Heat capacity analysis of oxidized Escherichia coli thioredoxin fragments (1–73, 74–108) and their noncovalent complex. Evidence for the burial of apolar surface in protein unfolded states. *Eur. J. Biochem.* 268:1477–1485.
9. Ladbury, J. E., R. Wynn, H. W. Hellinga, and J. M. Sturtevant. 1993. Stability of oxidized Escherichia coli thioredoxin and its dependence on protonation of the aspartic acid residue in the 26 position. *Biochemistry.* 32:7526–7530.
10. Munoz, V., and J. M. Sanchez-Ruiz. 2004. Exploring protein-folding ensembles: a variable-barrier model for the analysis of equilibrium unfolding experiments. *Proc. Natl. Acad. Sci. USA.* 101:17646–17651.
11. Shrake, A., and J. A. Rupley. 1973. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. *J. Mol. Biol.* 79: 351–371.
12. Chothia, C. 1975. Structural invariants in protein folding. *Nature.* 254: 304–308.
13. Shortle, D. 2003. Propensities, probabilities, and the Boltzmann hypothesis. *Protein Sci.* 12:1298–1302.
14. Lehmann, M., L. Pasamontes, S. F. Lassen, and M. Wyss. 2000. The consensus concept for thermostability engineering of proteins. *Biochim. Biophys. Acta.* 1543:408–415.
15. Ewens, W. J., and G. R. Grant. 2001. Statistical Methods in Bioinformatics: An introduction. Springer, New York, NY.
16. Lawrence, C. E., S. F. Altschul, M. S. Boguski, J. S. Liu, A. F. Neuwald, and J. C. Wootton. 1993. Detecting subtle sequence signals: a Gibbs sampling strategy for multiple alignment. *Science.* 262: 208–214.
17. Loladze, V. V., D. N. Ermolenko, and G. I. Makhatadze. 2002. Thermodynamic consequences of burial of polar and non-polar amino acid residues in the protein interior. *J. Mol. Biol.* 320:343–357.
18. Takano, K., J. M. Scholtz, J. C. Sacchettini, and C. N. Pace. 2003. The contribution of polar group burial to protein stability is strongly context-dependent. *J. Biol. Chem.* 278:31790–31795.
19. Steiner, E. 1996. The Chemistry Math Book. Oxford University Press, Oxford, UK.
20. Eriksson, A. E., W. A. Baase, X. J. Zhang, D. W. Heinz, M. Blaber, E. P. Baldwin, and B. W. Matthews. 1992. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science.* 255:178–183.
21. Robic, S., M. Guzman-Casado, J. M. Sanchez-Ruiz, and S. Marqusee. 2003. Role of residual structure in the unfolded state of a thermophilic protein. *Proc. Natl. Acad. Sci. USA.* 100:11345–11349.
22. Yohannan, S., S. Faham, D. Yang, J. P. Whitelegge, and J. U. Bowie. 2004. The evolution of transmembrane helix kinks and the structural diversity of G protein-coupled receptors. *Proc. Natl. Acad. Sci. USA.* 101:959–963.
23. Baldwin, E., J. Xu, O. Hajiseyedjavadi, W. A. Baase, and B. W. Matthews. 1996. Thermodynamic and structural compensation in ''size-switch'' core repacking variants of bacteriophage T4 lysozyme. *J. Mol. Biol.* 259:542–559.
24. Malakauskas, S. M., and S. L. Mayo. 1998. Design, structure and stability of a hyperthermophilic protein variant. *Nat. Struct. Biol.* 5: 470–475.
25. Korkegian, A., M. E. Black, D. Baker, and B. L. Stoddard. 2005. Computational thermostabilization of an enzyme. *Science.* 308: 857–860.
26. Ibarra-Molero, B., and J. M. Sanchez-Ruiz. 2002. Genetic algorithm to design stabilizing surface-charge distributions in proteins. *J. Phys. Chem. B.* 106:6609–6613.
27. Taverna, D. M., and R. A. Goldstein. 2002. Why are proteins marginally stable? *Proteins.* 46:105–109.
28. Wolynes, P. G., J. N. Onuchic, and D. Thirumalai. 1995. Navigating the folding routes. *Science.* 267:1619–1620.
29. Chan, H. S., S. Shimizu, and H. Kaya. 2004. Cooperativity principles in protein folding. *Methods Enzymol.* 380:350–379.
30. Scalley-Kim, M., and D. Baker. 2004. Characterization of the folding energy landscapes of computer generated proteins suggests high folding free energy barriers and cooperativity may be consequences of natural selection. *J. Mol. Biol.* 338:573–583.