Energetics of aliphatic deletions in protein cores

MARTA BUENO, LUIS A. CAMPOS, JORGE ESTRADA, AND JAVIER SANCHO

Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias & Biocomputation, and Complex Systems Physics Institute (BIFI), Universidad de Zaragoza, 50009-Zaragoza, Spain

(RECEIVED April 7, 2006; FINAL REVISION May 3, 2006; ACCEPTED May 3, 2006)

Abstract

Although core residues can sometimes be replaced by shorter ones without introducing significant changes in protein structure, the energetic consequences are typically large and destabilizing. Many efforts have been devoted to understand and predict changes in stability from analysis of the environment of mutated residues, but the relationships proposed for individual proteins have often failed to describe additional data. We report here 17 apoflavodoxin large-to-small mutations that cause overall protein destabilizations of 0.6–3.9 kcal.mol⁻¹. By comparing two-state urea and three-state thermal unfolding data, the overall destabilizations observed are partitioned into effects on the N-to-I and on the I-to-U equilibria. In all cases, the equilibrium intermediate exerts a “buffering” effect that reduces the impact of the overall destabilization on the N-to-I equilibrium. The performance of several structure-energetics relationships, proposed to explain the energetics of hydrophobic shortening mutations, has been evaluated by using an apoflavodoxin data set consisting of 14 mutations involving branching-conservative aliphatic side-chain shortenings and a larger data set, including similar mutations implemented in seven model proteins. Our analysis shows that the stability changes observed for any of the different types of mutations (LA, IA, IV, and VA) in either data set are best explained by a combination of differential hydrophobicity and of the calculated volume of the modeled cavity (as previously observed for LA and IA mutations in lysozyme T4). In contrast, sequence conservation within the flavodoxin family, which is a good predictor for charge-reversal stabilizing mutations, does not perform so well for aliphatic shortening ones.

Keywords: protein stability; hydrophobic interaction; protein intermediate; protein folding; protein cavity

Globular proteins are organized around well-packed hydrophobic cores that play important roles in protein folding and stability (Chan and Dill 1990; Dill 1990). Tight packing optimizes van der Waals interactions and allows for the minimization of the presence of cavities in proteins. Although small perturbations of core packing arising from mutations are often tolerated by proteins and give rise to small structural readjustments, they may result in large energetic effects (Milla et al. 1994; Munson et al. 1996; Vlassi et al. 1999). Understanding the contribution of buried hydrophobic residues to protein

stability is not trivial, and toward that end, considerable experimental work has been carried out to analyze large-to-small side-chain substitutions (Yutani et al. 1987; Matsumura et al. 1988; Kellis et al. 1989; Shortle et al. 1990; Baase et al. 1992; Eriksson et al. 1992b; Serrano et al. 1992; Takano et al. 1997; Xu et al. 1998, 2002; Chen and Stites 2001a). One goal in these studies has been to find out relationships between local protein structural features and the concomitant stability changes observed. For the staphylococcal nuclease, the change in stability upon mutation ($\Delta\Delta G$) was related to local packing density (Shortle et al. 1990); for barnase, a linear relationship was proposed between $\Delta\Delta G$ and an averaged measure of structural perturbations in neighboring atoms (Buckle et al. 1996); for chymotrypsin inhibitor 2 (CI2), $\Delta\Delta G$ was related to the number of methyl and methylene groups within 6 Å from the mutated residues (Otzen et al. 1995). Subsequent

Reprint requests to: Javier Sancho, Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias, Universidad de Zaragoza, E 50009-Zaragoza, Spain; e-mail: jsancho@unizar.es; fax: +34-976-76-21-23.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.062274906>.

studies have nevertheless shown that the correlations proposed were difficult to generalize to other proteins (Xu et al. 1998). A somewhat simpler approach was proposed in a T4 lysozyme study, where $\Delta\Delta G$ was approximated by a constant term, related to differences in hydrophobicity between the exchanged residues, plus a term proportional to cavity size (Eriksson et al. 1992b; Xu et al. 1998). As with the other relationships, this one was later shown not to hold in a different protein (ROP [Vlassi et al. 1999]). It thus seems clear that further investigation and assessment are required to determine whether general relationships can be established between local protein characteristics and the $\Delta\Delta G$ associated with buried mutations.

On the other hand it should be noticed that, as in many other mutational analyses, the contribution of buried hydrophobic residues to protein stability has been mainly studied using two-state proteins, where stability differences can be best determined (Pace and Laurents 1989). While two-state small proteins are very convenient models, they might not be very representative of average proteins, which are larger (i.e., 86% of the human proteins documented in Swiss-Prot [Bairoch and Apweiler 1997] are >15 kDa) and therefore more likely to display equilibrium-folding intermediates. For a three- (or more) state protein, it is useful to distinguish between the free-energy difference of the N-to-I equilibrium, which has been termed relevant stability (Sancho et al. 2002), and the residual stability associated with the I-to-U equilibrium. In previous work on charge-reversal mutations (Campos et al. 2004b), we showed that large overall stabilization of the native state relative to the unfolded one can be obtained using charge-reversal mutations without significantly increasing the relevant stability of the protein. Conversely, it can be anticipated that destabilizations arising from hydrophobic deletions in the core of three-state proteins can be partitioned in various ways, the extremes being the specific destabilization of the native state relative to the intermediate and that of the intermediate relative to the denatured state.

To reassess the contribution of hydrophobic residues to the overall stability of two-state proteins and to investigate how it is partitioned in three-state proteins, we analyze here a set of 17 hydrophobic (Fig. 1; Table 1), large-to-small mutations that have been implemented in the core of the apoflavodoxin from *Anabaena* PCC7119 (169 residues, 19 kDa) (Maldonado et al. 1998; Sancho 2006). The structure, stability, folding, and cofactor binding of this protein are well characterized (Genzor et al. 1996a,b; Fernandez-Recio et al. 1999; Lostao et al. 2000, 2003; Irun et al. 2001b; Langdon et al. 2001; Campos et al. 2005); the protein has been used to investigate different types of side-chain interactions (Fernandez-Recio et al. 1999; Campos et al. 2005); and it displays the interesting feature of behaving as a two-state protein toward chemical denaturation and as a three-state one

Figure 1. C α trace of *Anabaena* apoflavodoxin (PDB code 1FTG) showing, in green, the mutated side-chains.

toward thermal unfolding (Irun et al. 2001a). Our data indicate that all of the mutations tested induce large overall destabilizations that are unevenly distributed between the N-to-I and I-to-U equilibria in a way that can be fully explained on the basis of the available low-resolution structure of the apoflavodoxin thermal intermediate (Campos et al. 2004a). Moreover, 14 of the 17 apoflavodoxin mutants, together with mutants reported in staphylococcal nuclease, barnase, CI2, ROP, and T4 and human lysozyme, have been jointly analyzed to test the overall performance of the various correlations proposed between structural features and stability changes. According to our analysis, the best predictor of the observed stability changes is cavity volume.

Results and Discussion

Spectroscopic characterization of mutants

Structural perturbations associated with mutations may sometimes be detected in the near-UV CD spectra of proteins. For apoflavodoxin, there is a contribution of aromatic residues to the far-UV CD spectra that can also be used to detect departures from the native structure. As Figure 2 shows, the spectra of the mutants (except those of I109A) cannot be distinguished from those of the wild-type protein. This indicates that the tertiary interactions involving aromatic residues are basically the same in all of the proteins, except for I109A. The emission fluorescence spectra of the mutants (data not shown) display the same intensity and wavelength of maximal emission as that of the wild-type protein, except that of I109A, which

Table 1. Summary of the location, solvent accessibility, and contacts made by the replaced groups in the different apoflavodoxin mutants

Mutant	Position	van der Waals contacts (<4 Å)	ASA (%) ^a
L6A	Strand 1	L33/I51/F86	0.0
I22V	Helix 1	V18/F86	0.0
V31A	Strand 2	R23/F26/L33	0.0
L44A	Helix 2	F7/V41/L50/L73/F78	0.1
L50A	Strand 3	G5/F7/L44/Y47/K81	0.0
I51V	Strand 3	I4/F26	0.0
I52V	Strand 3	W66/Y85/L105	0.0
W66F	Helix 3	F7/V36/I52/C54/Q63/L105	0.0
A84G	Strand 4	L50/V117/W159/L163	0.0
L105A	Helix 4	I52/C54/L62/W66/Y70/A101	0.0
I109A	Helix 4	D75/F78/V83/G114	0.4
V117A	Strand 5a	V83/L163/E166/F167	13.2
V139A	Strand 8	I102/E106/T116/Y119/R134	0.0
L143A	Strand 5b	L141/I156/W159	0.1
I156V	Helix 5	L141/A142/L143/T152	1.8
V160A	Helix 5	E25/F26/I156/K164	0.1

^aCalculated with NACCESS (Hubbard et al. 1991).

displays a lower intensity and a red-shifted maximum. The additional mutants used in this work—L6A, V31A, I51V, I52V, A84G, V117A, V139A, L143A, and I156A—display CD and fluorescence emission spectra very similar to those of wild type (Campos et al. 2004a).

Overall (N-to-U) stability of cavity bearing mutants

Apoflavodoxin urea unfolding is two-state (Genzor et al. 1996a). The stability consequences of replacing buried residues by smaller ones have been investigated by urea denaturation equilibrium unfolding (Table 2). Seventeen cavity-creating mutants within the hydrophobic core of *Anabaena* apoflavodoxin were studied (L6A, I22V, V31A, L50A, I51V, I52V, A84G, L105A, I109A, V117A, V139A, L143A, I156V, V160A, W66F, W66F/L44A, and W66F/L105A). Of the eight newly made mutants, four (L105A, L50A, W66F/L105A, and I109A) can be unfolded with mid-concentrations of denaturant lower than 1 M, while three (V160A, I22V, and W66F/L44A) exhibit a urea concentration of mid-denaturation slightly above 1 M, and just one (W66F) is only slightly less stable than wild type. All of the fluorescence transition curves (Fig. 3) can be fitted to two-state transitions using Equation 2. The slopes of the observed transitions (m in Equation 2) are similar except for W66F/L44A, I109A, V160A, and I22V (which display lower m values). Comparison of the calculated ΔG values in water shows that all of the mutants are less stable than the wild-type protein, with destabilizations ranging from 0.6 to 3.9 kcal.mol⁻¹. By far, the I109A mutant is the most destabilized variant. The other nine mutants

(Campos et al. 2004), showed m values similar to that of wild type and are less stable.

The structures of the mutants were modeled from the wild-type structure as described in Machicado et al. (2002). They reveal the presence of protein cavities ranging from 14 Å³ for I51V to 118 Å³ for I109A in single aliphatic mutants (Table 3), and of 209 Å³ for the W66F/L44A double mutant (data not shown). In most cases, the modeled cavities are smaller than the theoretical cavities that would have appeared in the mutant protein in the absence of relaxation, but in one case, I156V, the modeled cavity is larger. Incidentally, this mutant shows one of the largest global destabilizations.

Figure 2. Circular dichroism spectra of wild type and some of the mutants analyzed (see Campos et al. 2004a for the rest of the mutants) at 25.0° ± 0.1°C. (A) Near-UV, in 50 mM MOPS (pH 7.0). (B) Far-UV, in 5 mM MOPS (pH 7.0), with 15 mM NaCl. Signals of the W66F mutant (and derived) are shown normalized, for the sake of comparison.

Table 2. Destabilization of wild-type apoflavodoxin by cavity creating mutations, as measured by urea denaturation^a

Protein	m^b (kcal.mol ⁻¹ M ⁻¹)	U_m (M)	$\Delta\Delta G^c$ (kcal.mol ⁻¹ M ⁻¹)	$\Delta\Delta G^d$ (kcal.mol ⁻¹ M ⁻¹)
WT	2.44 ± 0.03	2.07 ± 0.01	—	—
L6A	1.96 ± 0.06	0.77 ± 0.01	3.53 ± 0.09	2.94 ± 0.06
I22V	2.13 ± 0.05	1.46 ± 0.01	1.94 ± 0.10	1.37 ± 0.06
V31A	2.02 ± 0.05	1.31 ± 0.01	2.39 ± 0.10	1.72 ± 0.04
L50A	2.54 ± 0.08	0.88 ± 0.01	2.82 ± 0.10	2.69 ± 0.10
I51V	2.42 ± 0.11	1.27 ± 0.01	1.98 ± 0.16	1.83 ± 0.05
I52V	2.52 ± 0.13	1.45 ± 0.02	1.38 ± 0.21	1.40 ± 0.05
A84G	2.25 ± 0.07	1.28 ± 0.01	2.17 ± 0.11	1.80 ± 0.04
L105A	2.34 ± 0.12	1.00 ± 0.04	2.69 ± 0.17	2.40 ± 0.12
I109A	1.86 ± 0.08 ^e	0.62 ± 0.02	3.89 ± 0.09	3.89 ± 0.09
V117A	2.30 ± 0.03	1.30 ± 0.01	2.05 ± 0.08	1.74 ± 0.04
V139A	2.10 ± 0.09	1.68 ± 0.02	1.51 ± 0.17	0.87 ± 0.04
L143A	2.43 ± 0.12	2.02 ± 0.01	0.14 ± 0.25	0.10 ± 0.03
I156V	1.47 ± 0.06 ^e	0.99 ± 0.02	3.60 ± 0.10	3.60 ± 0.10
V160A	2.10 ± 0.05	1.30 ± 0.02	2.46 ± 0.09	1.88 ± 0.07
W66F	2.34 ± 0.05	1.81 ± 0.01	0.83 ± 0.12	0.59 ± 0.04
W66F/L44A	2.16 ± 0.09	1.32 ± 0.02	2.20 ± 0.14	1.69 ± 0.08
W66F/L105A	2.39 ± 0.15	0.84 ± 0.02	3.05 ± 0.15	2.78 ± 0.11

^aIn 50 mM MOPS (pH 7.0), at 25.0° ± 0.1°C.^b m and U_m values are means of two determinations ± SE.^cCalculated as $\Delta\Delta G = m(\text{wt}) \times U_m(\text{wt}) - m(\text{mut}) \times U_m(\text{mut})$. Errors in $\Delta\Delta G$ are calculated by propagation of the m and U_m errors.^dCalculated as $\Delta\Delta G = m_{\text{av}} \times [U_m(\text{wt}) - U_m(\text{mut})]$, where m_{av} is the mean value of the m slope of the different proteins (2.26 ± 0.07, mean ± SE). Errors in $\Delta\Delta G$ are calculated by propagation of the m_{av} and U_m errors.^eThe m value of these mutants are not within the SD of the other m values, and it is not appropriate to use an averaged m value to calculate $\Delta\Delta G$. Therefore, the two $\Delta\Delta G$ values reported for each mutant have been calculated as described in footnote c.

The W66F, L105A, and W66F/L105A mutants allow evaluation of the cooperativity or structural independence of two mutations that face the same cavity. The sum of the destabilization observed in W66F and L105A (2.99 ± 0.12 kcal.mol⁻¹) is very close to the $\Delta\Delta G$ observed in the double mutant (2.78 ± 0.11 kcal.mol⁻¹). In addition, the sum of the volume of the cavities in the single mutants (111 Å³ and 86 Å³ for W66F and L105A, respectively) is almost the same as that of the cavity calculated for the double mutant (205 Å³). It seems that the two mutations are independent and thus additive, which could be a reflection of their fairly rigid local environment as suggested by the B factors (Genzor et al. 1996b; Eyal et al. 2003).

One of the earliest and most straightforward structure–energetics relationships proposed to account for the destabilization arising from replacing buried aliphatic residues by shorter ones is that of Kellis et al. (1989). From the analysis of two Ile to Val barnase mutants (I88V, I96V) the protein was found to be destabilized by 1.0–1.6 kcal.mol⁻¹/methyl group removed. This value was subsequently refined to 1.5 ± 0.6 kcal.mol⁻¹ (Serrano et al. 1992) by considering two additional barnase mutants (I5V, I76V), one (I72V) from staphylococcal nuclease (Shortle et al. 1990) and one (I47V) from the gene V protein from bacteriophage f1 (Sandberg and Terwilliger 1991). An additional calculation of the $\Delta\Delta G$ value per –CH₂–group was performed by Pace (1992) adding to the

previous set of mutants, four staphylococcal nuclease mutants (I15V, I18V, I92V, and I139V) and a T4 lysozyme (I3V) mutation. A value of 1.3 ± 0.4 kcal.mol⁻¹ was obtained (Pace 1992). An analogous calculation using only four apoflavodoxin Ile-to-Val mutations (Table 2) gives a value of 2.0 ± 1.0 kcal.mol⁻¹/methylene group

Figure 3. Urea denaturation curves of wild-type and mutant apoflavodoxins at 25.0° ± 0.1°C (pH 7.0), 50 mM MOPS. The fluorescence signal of the W66F mutant is shown normalized, for the sake of comparison.

Table 3. Various parameters that describe the environment surrounding the mutated aliphatic residues

Mutation	Theoretical cavity volume (Å ³) ^a	Modeled cavity volume (Å ³) ^b	Thermal factor (Å ²) ^c	ΔASA buried by side chain (Å ²) ^d	Methyl(ene) groups <6 (Å ²) ^e	Cα groups <10 (Å ²) ^f	Δn _H /Δn _{H,wt} (%) ^g
L6A	73	61	9.24	68.22	57	25	−83.9
I22V	40	37	11.13	22.42	27	20	−44.2
V31A	0	0	14.68	47.14	38	22	−62.0
L50A	72	62	9.04	67.76	61	25	−80.6
I51V	33	14	7.88	22	30	24	−43.1
I52V	47	33	7.92	22	22	21	−29.5
A84G	21	0	7.84	28	26	24	N.C. ^h
L105A	99	86	13.63	69.09	71	23	−82.5
I109A	127	118	14.85	70.96	64	22	−92.2
V117A	0	0	9.94	13.44	25	19	−72.5
V139A	0	0	8.51	47	36	27	−70.4
L143A	116	0	9.37	69.11	45	26	N.C. ^h
I156V	20	105	11.70	22.11	18	20	−55.7
V160A	25	30	11.67	46.76	40	18	−66.6

^aCalculated with a probe of a 1.4 Å radius, using the Connolly method (Connolly 1983) on the cavity created by replacing in silico a given side chain without performing any minimization.

^bCavity volume after minimization of the mutant structure, as described in Materials and Methods.

^cResidue thermal factor extracted from the apoflavodoxin X-ray structure (PDB code 1ftg).

^dSolvent-accessible areas of the different residues calculated using the algorithm of Lee and Richards (1971). The change in ASA on mutation was calculated as described in Serrano et al. (1992).

^eMethyl and methylene groups (excluding C_α) at <6 Å of the methyl or methylene groups removed by mutation.

^fNumber of C_α atoms within a sphere of 10 Å radius centered in the C_α of the mutated residue.

^gContacts changed in the mutant relative to the contacts found in the wild-type protein (showed in percentage), as calculated in Vlassi et al. (1999).

^h(N.C.) Not calculated because one of the terms was zero.

removed, and if all of the 15 Ile-to-Val mutations are jointly considered, the estimation of the contribution to protein stability of a buried methyl/methylene group is set at 1.5 ± 0.8 kcal.mol^{−1}.

Relevant (N-to-I) and residual (I-to-U) stability of cavity bearing mutants: An energetic “buffering” effect exerted by protein equilibrium intermediates

The thermal unfolding of apoflavodoxin is three-state (Irun et al. 2001a), which allows for investigation of the specific energetics of N-to-I and I-to-U equilibria (Campos et al. 2004a,b, 2005). The thermal unfolding data corresponding to the hydrophobic apoflavodoxin substitutions is summarized in Table 4. All mutations destabilize the two equilibria. For the N-to-I equilibrium, the fall in T_m values ranges from 0.4° to 14.5° and for the I-to-U equilibrium, from 2.2° to 19.9°. As observed in the chemical denaturation, the I109A mutant appears as the most globally (N-to-U) destabilized protein.

The thermal and chemical denaturation data can be compared. By adding the ΔΔG_{NI} and ΔΔG_{IU} values calculated at the reference temperature of 317.3 K (Table 4), a ΔΔG_{NU} change is obtained at that temperature. On the other hand, an equivalent magnitude is calculated at 298.2 K from the urea unfolding curves (Table 2). If the thermal and urea-denatured states are energetically equiv-

alent and the stability differences between protein variants are retained going from 317.3 K to 298.2 K, the two-data set should be correlated. In fact, the correlation obtained for the available cavity mutants (Fig. 4) is good ($\Delta\Delta G_{\text{NU}}^{\text{thermal}} = 0.3 + 0.9\Delta\Delta G_{\text{NU}}^{\text{urea}}$; $r = 0.9$), as previously found for different sets of apoflavodoxin mutants (Campos et al. 2004).

The three-state global analysis provides information on the stability of the intermediate relative to the native and denatured states, which determines both its temperature of maximal accumulation and its highest molar fraction. The wild-type thermal intermediate accumulates at 323 K, where it constitutes 64% of the population (Campos et al. 2004a,b). The mutant intermediates tend to accumulate at the same or lower temperatures and in a lower percentage (data not shown), reflecting a larger destabilization of the second transition compared with the initial one (Table 4), which narrows the temperature window where the intermediate is stable. This is clearly seen in the changes brought about by the mutations in the relevant and residual stabilities of the protein (ΔΔG_{NI} and ΔΔG_{IU}) (Sancho et al. 2002). While the relevant stability is decreased by 0.5–1.0 kcal.mol^{−1} in most mutants, the thermal intermediates are destabilized, relative to the denatured state by 1.0–2.0 kcal.mol^{−1} in most cases (see Table 4). The differential impact of the mutations in the stability of the native and the intermediate conformations (both likely suffering from the same loss of

Table 4. Thermodynamic parameters of the three-state unfolding equilibrium of wild-type and mutant apoflavodoxins

Protein	ΔH_{NI} (kcal.mol ⁻¹)	T_{mNI} (K)	$\Delta\Delta G_{\text{NI}}$ (kcal.mol ⁻¹)	ΔH_{IU} (kcal.mol ⁻¹)	T_{mIU} (K)	$\Delta\Delta G_{\text{IU}}$ (kcal.mol ⁻¹)	$\Delta\Delta G_{\text{NU}}$ (kcal.mol ⁻¹)	Integrity (Φ_{I}) ^a
WT	34.1 ± 0.1	317.3 ± 0.2	—	52.9 ± 0.6	329.0 ± 0.1	—	—	—
L6A	28.7 ± 0.8	305.5 ± 0.2	1.11 ± 0.04	23.8 ± 0.7	319.1 ± 0.2	1.74 ± 0.03	2.85 ± 0.05	0.61 ± 0.01
I22V	33.2 ± 0.8	311.0 ± 0.2	0.67 ± 0.05	37.5 ± 1.2	324.7 ± 0.3	1.04 ± 0.06	1.71 ± 0.07	0.61 ± 0.04
V31A	30.9 ± 0.41	312.6 ± 0.2	0.46 ± 0.02	41.5 ± 0.8	324.2 ± 0.2	0.99 ± 0.04	1.45 ± 0.04	0.68 ± 0.01
L50A	40.8 ± 0.4	309.8 ± 0.3	0.98 ± 0.06	47.9 ± 1.4	316.7 ± 1.6	1.98 ± 0.24	2.97 ± 0.25	0.67 ± 0.10
I51V	31.2 ± 0.61	313.2 ± 0.2	0.41 ± 0.02	38.8 ± 0.9	324.3 ± 0.2	1.04 ± 0.03	1.45 ± 0.03	0.72 ± 0.01
I52V	39.8 ± 0.3	315.9 ± 0.1	0.18 ± 0.02	54.5 ± 1.2	324.9 ± 0.1	0.60 ± 0.04	0.78 ± 0.05	0.77 ± 0.02
A84G	34.4 ± 0.82	313.0 ± 0.1	0.47 ± 0.02	31.9 ± 0.7	323.6 ± 0.1	1.25 ± 0.03	1.67 ± 0.03	0.73 ± 0.01
L105A	41.8 ± 0.6	313.4 ± 0.4	0.51 ± 0.07	13.6 ± 1.7	313.4 ± 3.1	2.06 ± 0.14	2.57 ± 0.16	0.80 ± 0.07
I109A	20.6 ± 2.0	302.8 ± 1.1	0.99 ± 0.13	13.4 ± 2.1	309.1 ± 3.0	2.25 ± 0.15	3.23 ± 0.20	0.70 ± 0.06
V117A	36.9 ± 0.7	309.3 ± 0.1	0.95 ± 0.03	33.7 ± 0.7	322.6 ± 0.1	1.32 ± 0.03	2.27 ± 0.03	0.58 ± 0.01
V139A	29.0 ± 0.1	310.9 ± 0.2	0.60 ± 0.02	50.8 ± 0.6	327.5 ± 0.1	0.29 ± 0.03	0.89 ± 0.02	0.33 ± 0.02
L143A	33.3 ± 0.61	315.8 ± 0.2	0.15 ± 0.02	50.0 ± 0.5	328.9 ± 0.1	0.11 ± 0.03	0.26 ± 0.04	0.43 ± 0.08
I156V	25.3 ± 0.82	308.8 ± 0.4	0.69 ± 0.04	29.9 ± 0.9	320.3 ± 0.2	1.60 ± 0.03	2.29 ± 0.03	0.70 ± 0.01
V160A	27.8 ± 0.3	309.1 ± 0.6	0.73 ± 0.07	49.4 ± 1.4	322.4 ± 0.2	1.11 ± 0.05	1.84 ± 0.08	0.60 ± 0.04
W66F	35.1 ± 0.8	316.9 ± 1.1	0.04 ± 0.13	31.3 ± 2.1	326.8 ± 3.0	0.98 ± 0.29	1.02 ± 0.31	0.96 ± 0.12
W66F/L44A	29.1 ± 0.8	311.9 ± 1.1	0.50 ± 0.08	36.2 ± 2.0	316.5 ± 0.6	1.99 ± 0.08	2.49 ± 0.11	0.80 ± 0.05
W66F/L105A	41.1 ± 0.8	309.1 ± 0.3	1.08 ± 0.06	28.6 ± 2.1	323.3 ± 1.2	1.35 ± 0.12	2.44 ± 0.13	0.56 ± 0.06

ΔH and T_{m} values are calculated by global fitting of thermal denaturation curves followed by four different techniques to a three-state model. Errors in ΔH and T_{m} are calculated as explained in Materials and Methods. ΔG is calculated as $\Delta H^*[1 - (317.3 / T_{\text{m}})]$ and $\Delta\Delta G = \Delta G(\text{wild type}) - \Delta G(\text{mutant})$. Errors in ΔS , ΔG , and $\Delta\Delta G$ are calculated by propagation of ΔH and T_{m} errors.

^aDefined as $\Phi_{\text{I}} = \Delta\Delta G_{\text{IU}} / \Delta\Delta G_{\text{NU}}$ (see Materials and Methods).

hydrophobic stability relative to wild type because the mutations are located in the native-like region of the intermediate; see below) might be due to collapse (at least partial), in the more flexible intermediate of the cavities appearing in the native state (see discussion below on the relationship between destabilization, hydrophobicity, and cavity size). However, the occurrence of an intermediate in the thermal unfolding of the protein buffers the destabilization of the native state caused by the mutations because—since they also significantly destabilize the intermediate—the destabilizing effect in the first transition is reduced.

This fortunate characteristic, which is likely to be general for proteins displaying equilibrium intermediates, does not come without its negative side. In principle, since proteins are destabilized by hydrophobic cavity-created mutations, they could also be stabilized by the reverse, i.e., by implementing mutations to fill existing cavities. This strategy has been explored in two-state proteins, where it has sometimes led to overall N-to-U stabilization (Ishikawa et al. 1993; Akasako et al. 1997; Lassalle et al. 2001; Ohmura et al. 2001). If the 17 apoflavodoxin mutants are thought of in an inverse manner, as different proteins whose cavities can be filled to yield the largely stabilized wild-type apoflavodoxin, it is clear that most of the stabilization that is achieved (~65% on average) goes to increasing the stability of the intermediate against its full unfolding (Table 5; Fig. 5). Only one-third of the stabilization actually increases the relevant stability of the native state against its partial

unfolding. In practical terms, this means that the overall stabilization of a protein displaying equilibrium intermediates that can be achieved by cavity filling will probably have, on average, a small impact on the specific stabilization of the N-to-I equilibrium. This effect has already been observed in apoflavodoxin using charge-reversal stabilizing mutations that, in most cases, left the

Figure 4. Comparison of $\Delta\Delta G_{\text{NU}}$ values obtained for wild-type apoflavodoxin and mutants thereof by adding the $\Delta\Delta G_{\text{NI}}$ and $\Delta\Delta G_{\text{IU}}$ terms calculated from global analysis of three-state thermal unfolding curves (Table 3) with those directly measured from two-state urea denaturation (Table 1). The filled circle represents data from mutant I156V, excluded for the correlation due to possible structural relaxations (see text). The dotted line is a unity fit.

Table 5. Residue distribution and conservation data for the studied flavodoxin positions

Position ^a	Occurrence (%)					fX _i ^b	fM _i ^c	ΔΔG _{cons} (kcal.mol ⁻¹) ^d
	70–100	50–70	30–50	15–30	5–15			
Leu6	—	I	—	L,V	M	0.0	53.7	—
Ile22	I	—	—	—	L,V	10.5	75.8	1.17
Val31	—	V	—	I	A,D,F,S,T	14.2	54.9	0.80
Leu50	—	—	L	V	A,C,I,S	6.2	46.5	1.19
Ile51	—	I	—	L	A,F,M,V	8.5	52.4	1.08
Ile52	—	—	L	F	C,D,I,M,V	9.8	44.0	0.89
Trp66	—	W	—	F	A,G,I,L,M,V	18.7	51.7	0.60
Ala84	A	—	—	G	P,S,V	17.7	71.5	0.83
Leu105	—	—	L	FI	V,W	0.0	40.1	—
Ile109	—	—	—	F,I,L	A,M,S,V,Y	8.4	28.6	0.73
Val117	—	—	V	I	C,E,G,L,P,S,Y	0.0	48.7	—
Val139	—	—	V	—	C,I,L,M,Y	0.0	34.6	—
Leu143	—	I	—	L	F,M,V	0.0	50.1	—
Ile156	—	—	I	L,V	A,C,S	24.7	37.5	0.25
Val160	—	—	—	A,V	C,G,I,L,Q,T	17.5	27.5	0.27

^aResidue present in the sequence of *Anabaena* PCC7119 flavodoxin.
^bFrequency of occurrence of the residue introduced at position *i*.
^cFrequency of occurrence of the most commonly observed residue at position *i*.
^dAs defined in Di Nardo et al. (2003).

relevant stability unchanged (Campos et al. 2004b). However, when the charge-reversal mutations were introduced in the region of the protein that appears unstructured in the thermal intermediate (Campos et al. 2004), significant increases of the relevant stability were obtained. Similarly, a close look at how the destabilizations associated to aliphatic deletions are partitioned into the two equilibria (Fig. 5) shows that in the two instances where the mutations are introduced in the unstructured region of the intermediate (V139A and L143A) (Campos et al. 2004a), the largest destabilization is also associated to the N-to-I transition. In a recent work (Bueno et al. 2006), several cavities in apoflavodoxin have been filled, and one of the mutants that retain the native conformation is significantly more stable than wild type toward full unfolding. However, as the filling mutation is located in the region where the intermediate displays a native-like conformation, the relevant stability of the mutant is not greater than that of wild type.

The apoflavodoxin thermal unfolding intermediate has been investigated in detail, and a low-resolution structure (Fig. 6; Table 1) has been obtained by equilibrium Φ-analysis and confirmed by NMR (Campos et al. 2004). We report here Φ-values of additional interactions (Table 4), which are consistent with the reported structure. They probe the integrity of the packing between the central β-sheet and the layer formed by helices α2, α3, and α4. According to the Φ-values obtained, the N-terminal region of α3 is close to natively packed against the central residues of the β-sheet (W66F); the interface between the central region of α4 and the loop connecting α4 and β4 is mostly formed (I109A); and there is also

significant packing between the parallel β-sheet and helices α3 and α4 (L105A). Double mutants have been used to check the packing between the three helices and the central parallel β-sheet, and their Φ-values indicate substantial packing. All of these mutations confirm that in the thermal intermediate, the interactions of the α2–α3–α4 helical layer onto the central β-sheet are close to native.

Are cavities solvated? Contribution of buried water molecules to apoflavodoxin stability

Despite the efficient packing of protein cores, most proteins contain small cavities that can be empty or

Figure 5. Partitioning of the global destabilization (N-to-U equilibrium) associated to the mutations implemented into the N-to-I equilibrium, relevant destabilization (gray bar), and the I-to-U equilibrium, residual destabilization (white hatched bar).

Figure 6. Apoflavodoxin secondary structure cartoon showing new (continuous thick lines) and previously reported (Campos et al. 2004a) (discontinuous thin lines) ϕ -values reporting on the integrity of native side-chain interactions in the equilibrium thermal unfolding intermediate at 317.3 K. The native-like region of the intermediate is formed by the packing of helices 1–5, and strands 1–4 (where the ϕ -values are >0.6) and possibly strand 5a. In contrast, the long loop splitting strands 5a and 5b (which contains a small three-stranded b-sheet comprising strands 6, 7, and 8), and two additional loops, are markedly weakened (all ϕ -values <0.4) (Campos et al. 2004a). The ϕ -values are colored from dark gray green (native interactions) to red (lost interactions). ID numbers of helices (circles) and strands (rectangles) are marked in light gray, and the positions analyzed in this work are indicated in black.

contain water molecules (Rashin et al. 1986; Hubbard and Argos 1994; Williams et al. 1994). Solvated cavities are larger (mean volume of $39.4 \text{ \AA}^3$) than empty ones ($20.7 \text{ \AA}^3$) (Hubbard et al. 1994). In fact, most empty cavities (66%) are $<20 \text{ \AA}^3$ (Hubbard et al. 1994). The extent and nature of hydration of the interior regions of proteins continues to be a subject of great interest and considerable controversy (Yu et al. 1999). The probability of finding a single water molecule in a small hydrophobic cavity has been estimated to be about one in 20,000 (Wolfenden and Radzicka 1994). Clearly, for a water molecule to be stable within a protein cavity, the environment therein must be energetically favorable relative to liquid water. Two factors that correlate closely with the ordered water content in interior and interfacial protein cavities are the number of polar atoms lining the cavity and cavity size (Adamek et al. 2005). Generally, larger and more polar cavities are more hydrated (Hubbard and Argos 1994). Several studies have focused on the structural and energetic consequences of creating cavities (Eriksson et al. 1992b; Buckle et al. 1993, 1996; Jackson et al. 1993; Xu et al. 1998), but few have reported the hydration of the new cavity even though the cavities are generally lined with some polar atoms (Adamek et al. 2005).

To evaluate whether the cavities created in apoflavodoxin are likely solvated or not, we have selected two

different criteria: first, a volume cutoff, and second, an energetic cutoff. Hubbard et al. (1994) demonstrated that most cavities with volumes $>50 \text{ \AA}^3$ are hydrated. On the other hand, Takano et al. (2003) showed that buried water molecules stabilize a protein structure. Tight packing of buried water molecules in the interior of proteins provides better van der Waals interactions than those in an empty cavity (Takano et al. 1997), and thus a solvated cavity is more stable than an empty one. The reported average destabilization of Ile/Leu to Ala mutations in a solvated cavity is of $1.5 \text{ kcal.mol}^{-1}$, much less than the same type of mutation in empty cavities: $3.4 \text{ kcal.mol}^{-1}$ (Takano et al. 2003). Of all of the apoflavodoxin mutants considered in this work, only mutants L6A, L50A, L105A, I109A, and I156V present cavities $>50 \text{ \AA}^3$ and could be thus solvated, but all of them show destabilizations $>2.5 \text{ kcal.mol}^{-1}$ (see Table 2). Accordingly, the probability of finding water molecules in the apoflavodoxin mutant cavities is low and they will be considered as empty cavities.

Performance of different structure/energetics relationships for hydrophobic large-to-small mutations in protein cores

The relationship between structural characteristics of protein cores and stability changes arising from large-to-small

substitutions is not well understood. Different correlations have been proposed for different data sets corresponding to mutations located in several proteins. Here we reassess, in two ways, the performance of those correlations. First (Fig. 7), we tested them on 14 apoflavodoxin mutants corresponding to single mutations involving aliphatic residues; second (Fig. 8), we tested them on a larger data set comprising mutations from seven different proteins: 14 mutants from apoflavodoxin, two from ROP protein (Vlassi et al. 1999), four from barnase (Serrano et al. 1992), eight from T4 lysozyme (Eriksson et al. 1992a; Xu et al. 1998), three from staphylococcal nuclease (Shortle et al. 1990),

Figure 7. Plots of $\Delta\Delta G$ (wild-type apoflavodoxin free energy of unfolding minus that of mutants) vs. different parameters that describe the environment of the mutated residues. (A) Number of $C\alpha$ groups within a sphere of 10 Å radius centered in the $C\alpha$ of the mutated residue, (B) differences in side-chain solvent-accessible area buried between wild type and mutant, (C) number of methyl or methylene side-chain groups surrounding (within a sphere of 6 Å radius) the methyl or methylene group deleted upon mutation, (D) change of contacts (calculated as contact parameter n_H) relative to the wild-type protein, (E) thermal factors of the mutated residues, as extracted from the PDB file 1FTG, and (F) cavity volume created by the amino acid substitution.

Figure 8. Plots of $\Delta\Delta G$ of unfolding vs. different parameters that describe the surrounding environment of mutated residues in seven different proteins: apoflavodoxin, ROP (Vlassi et al. 1999), barnase (Serrano et al. 1992), T4 lysozyme (Eriksson et al. 1992a; Xu et al. 1998), staphylococcal nuclease (Shortle et al. 1990), CI2 (Otzen and Fersht 1995), and human lysozyme (Takano et al. 1995, 1997). (A) Number of $C\alpha$ groups within a sphere of 10 Å radius centered in the $C\alpha$ of the mutated residue; (B) differences in side-chain solvent-accessible area buried between wild type and mutant; (C) number of methyl or methylene side-chain groups surrounding (within a sphere of 6 Å radius) the methyl or methylene group deleted upon mutation; (D) change of contacts (calculated as contact parameter n_H) relative to the wild-type protein; (E) thermal factors of the mutated residues as extracted from the PDB file of the wild-type protein; (F) cavity volume created by the amino acid substitution.

three from CI2 (Otzen and Fersht 1995), and two from human lysozyme (Takano et al. 1995, 1997). All of these protein variants are also single mutants involving aliphatic residues completely buried from solvent.

For staphylococcal nuclease (Shortle et al. 1990), a correlation was found between $\Delta\Delta G$ and the number of $C\alpha$ groups within a sphere of 10 Å radius centered in the $C\alpha$ of the mutated residue ($r = 0.67$). This correlation does not work either for the apoflavodoxin data set

($r = 0.25$) (Fig. 7A) or for the larger data set of seven proteins ($r = 0.44$) (Fig. 8A). For barnase, two correlations were proposed (Serrano et al. 1992) that related $\Delta\Delta G$ with differences in the side-chain solvent-accessible area between the residue in the wild type and the mutant ($r = 0.83$) or with the number of methyl or methylene side-chain groups surrounding the group deleted ($r = 0.91$). The latter relationship was also reported to perform well ($r = 0.93$) for C12 mutants (Otzen et al. 1995). However, these correlations do not work for the apoflavodoxin data set ($r = 0.34$ for ΔASA ; $r = 0.25$ for number of methyl/ene) (see Fig. 7, B and C, respectively) and do not perform so well in the data set of seven proteins ($r = 0.59$ for ΔASA ; $r = 0.49$ for number of methyl/ene) (see Fig. 8, B and C, respectively). Vlassi et al. (1999) introduced a contact parameter (Δn_H) as a measure of the number of van der Waals contacts lost and for a set of ROP, barnase, and T4 lysozyme mutants, they found a good correlation between Δn_H and $\Delta\Delta G$ ($r = 0.81$). The correlation observed for the 14 apoflavodoxin mutants ($r = 0.52$) (Fig. 7D) and for the data set of seven proteins ($r = 0.68$) (Fig. 8D) are lower. Mutations that substantially destabilized T4 lysozyme were traced to environments with low thermal motions (Alber et al. 1987; Matthews 1993). Takano et al. (1995) showed in human lysozyme an inverse correlation between residue mobility and destabilization ($r = 0.84$). This correlation does not hold either for apoflavodoxin ($r = 0.14$) (Fig. 7E) or for the complete data set ($r = 0.38$) (Fig. 8E). It seems thus that most correlations described for a given protein or set of proteins perform poorly on different or simply larger data sets, which stresses the fact that the energetics of cavity-creating mutations in protein hydrophobic cores are, at least to some extent, site- or type-dependent.

A further attempt to relate the structure and energetics of hydrophobic large-to-small replacements was carried out on T4 lysozyme mutants for which $\Delta\Delta G$ was found to correlate ($r = 0.96$) to the size of the engineered cavities (Eriksson et al. 1992b) with a slope of $22 \text{ cal.mol}^{-1}\text{\AA}^{-3}$ (Xu et al. 1998). This linear correlation was roughly confirmed in other proteins (Takano et al. 1995; Buckle et al. 1996), but more recently it was claimed to be not so general ($r = 0.51$ for ROP [Vlassi et al. 1999]). However, the correlation performs quite well in the apoflavodoxin data set ($r = 0.87$) (Fig. 7F) with a slope of $21 \text{ cal.mol}^{-1}\text{\AA}^{-3}$ and an intercept of $+1.9 \text{ kcal.mol}^{-1}$. For the set of seven proteins, the correlation is still reasonable ($r = 0.72$) (Fig. 8F), with a very similar slope of $22 \text{ cal.mol}^{-1}\text{\AA}^{-3}$ and an intercept of $+1.6 \text{ kcal.mol}^{-1}$.

The work by Matthews' lab (Eriksson et al. 1992b; Xu et al. 1998) showed that the energetic effect of both Leu or Ile to Ala mutations could be correlated to the volume of the created cavity at a rate of $22 \text{ cal.mol}^{-1}\text{\AA}^{-3}$, and

that each type of mutation had, in addition, an intrinsic destabilizing effect that could be attributed to the specific hydrophobicity of the substituted side chain as measured by water/octanol transfer experiments (Fauchere and Pliska 1983). For valine to alanine mutations, the data were more scattered. To investigate whether the same principles can be extended to other proteins and to different mutations, we separated the flavodoxin mutations into three groups and performed specific correlations for Leu to Ala, Val to Ala, and Ile to Val mutations. For the three subsets, the slope of $\Delta\Delta G$ versus cavity volume plots displays values between 20 and 23 $\text{cal.mol}^{-1}\text{\AA}^{-3}$. The intercepts of the plots, however, differ from 1.4 and 1.5 kcal.mol^{-1} for Val to Ala and Leu to Ala mutations to 0.9 kcal.mol^{-1} for Ile to Val ones, which is in qualitative agreement with the corresponding water/octanol difference transfer energies of 1.2, 1.9, and 0.7 kcal.mol^{-1} (Xu et al. 1998). An equivalent analysis performed in the seven-protein database gives slope values ranging from 20 to 24 $\text{cal.mol}^{-1}\text{\AA}^{-3}$ and intercepts that correlate reasonably well with water-to-octanol transfer energies ($r = 0.88$) (Fig. 9). The intercepts also correlate well with the ΔASA of the mutated side chains in the denatured state as calculated using three different types of denatured state models: extended Ala-X-Ala tripeptide ($r = 0.79$) (Shrake and Rupley 1973), extended Gly-X-Gly tripeptide ($r = 0.85$) (Miller et al. 1987), and 15-mer in silico modeled peptide ($r = 0.82$) (Creamer et al. 1995). The data corresponding to the four types of mutations (Ile to Ala, Leu to Ala, Val to Ala, and Ile to Val) investigated in the seven proteins data set were brought to a common, hydrophobicity-independent reference state by subtracting the intercepts (Fig. 10). This raises the correlation between $\Delta\Delta G$ and cavity volume

Figure 9. Correlation between transfer energies and intercepts of linear relationships between cavity volume and experimental $\Delta\Delta G$ for different types of mutations (Leu to Ala, Val to Ala, Ile to Val, and Ile to Ala) implemented in any of the seven proteins of the data set (see Materials and Methods). The dashed line shows a unity fit.

Figure 10. The experimentally determined $\Delta\Delta G$ values shown in Figure 8F have been brought to a common reference state (hydrophobicity-independent) by subtraction of the intercepts of mutation-type separated plots of $\Delta\Delta G_{\text{NU}}$ vs. volume (see Discussion).

from 0.72 (Fig. 8F) to 0.90. Further attempts to improve the correlation considering additional terms accounting for differences in side-chain entropy or secondary structure propensity did not work. Our analyses thus support the proposal (Eriksson et al. 1992b) that the change in free energy upon replacement of Leu or Ile to Ala can be reasonably explained by a combination of differential hydrophobicity plus a term proportional to cavity volume. In addition, the correlation is now extended to Ile to Val and to Val to Ala mutations. All branching-conservative, side-chain shortenings of purely aliphatic residues are thus explained using the same principles.

Does the “consensus approach” work for hydrophobic mutations?

Sequence comparison has been used as a tool to design thermostabilizing substitutions (Serrano et al. 1993; Steipe et al. 1994; Wirtz and Steipe 1999). From an alignment of more than 350 different SH3 domains (Maxwell and Davidson 1998), a correlation was found between residue conservation and thermodynamic stabilities of mutants constructed at two different positions ($r = 0.79$). The term “consensus approach” was coined to refer to the fact that a set of amino acid sequences of homologous proteins contains sufficient information to allow a rapid design of thermal stabilization (Lehmann et al. 2002). The approach has been used in the stabilization of a GroEL minichaperone (Wang et al. 2000) and a fungal pythase (Lehmann et al. 2000). As has been pointed out (Campos et al. 2004b), the consensus approach is a practical application of the known and yet surprising fact that natural protein sequences seem to comply to some extent to Boltzmann’s law (Fernandez-Recio and Sancho 1998; Shortle 2003; Campos et al. 2004; Godoy-Ruiz et al. 2004). Sequence alignments performed on several protein families, including ubiquitin (Finucane

and Woolfson 1999), staphylococcal nuclease (Chen and Stites 2001b), GroES (Taneja and Mande 1999), Ig-like domains (Halaby et al. 1999), and SH3 (Di Nardo et al. 2003) have shown that individual positions in the hydrophobic core generally display distinct preferences for particular types of hydrophobic residues. Di Nardo et al. (2003) have found that experimentally determined $\Delta\Delta G$ values in SH3 domains show a good linear correlation ($r = 0.78$, with a slope of 1) with conservation energies. In this line, we have recently reported that the overall stabilization afforded by charge-reversal mutations in apoflavodoxin (Campos et al. 2004b) correlates linearly with weighted conservation energies ($r = 0.89$), so that the stabilization of charge-reversal stabilizing mutations can be predicted semiquantitatively from simple sequence comparison within the flavodoxin family. To assess whether the energetics of aliphatic deletions can be similarly predicted from sequence comparison, we analyzed 75 flavodoxin sequences from different species (López-Llano et al. 2004) and calculated conservation energies (Di Nardo et al. 2003) for the positions mutated (Table 5). The correlation between conformational stability and conservation energy is not very strong ($r = 0.50$ with a slope of 0.94 if one mutant is excluded). The predictive power of “conservation energies” for aliphatic deletions is thus low, at least in the flavodoxin family.

Conclusion

A total of 14 buried, branching-conservative shortening mutations involving apolar aliphatic residues have been implemented in apoflavodoxin, giving rise to large protein destabilization without introducing major changes in protein structure. In the thermal unfolding, the overall N-to-U destabilization is partitioned into destabilization of the native state relative to an unfolding intermediate (Sancho et al. 2002) and destabilization of the intermediate toward its full unfolding, which lowers the destabilizing impact exerted by the mutations on the native conformation. This suggests that stabilization of three- (or more) state proteins by cavity filling may be of limited success, unless the filled cavity is specifically located in an unfolded region of the intermediate.

An analysis of several structure–energetics relationships proposed to explain mutational effects on protein stability associated to apolar-buried shortening mutations was performed using the apoflavodoxin data set and a larger data set comprising aliphatic-shortening mutations described for seven different proteins. The analysis indicates that modeled cavity volume (Machicado et al. 2002) plus differential hydrophobicity (Eriksson et al. 1992b) are good predictors of the observed energetics, as previously proposed (Xu et al. 1998) for a smaller data set of lysozyme T4 mutants of known X-ray structures. In

contrast, unlike what has been described for SH3 domains (Di Nardo et al. 2003), sequence conservation analysis seems to not work so well.

Materials and methods

Mutagenesis, protein purification and quantitation, and apoprotein preparation

The *Anabaena* PCC7119 flavodoxin gene was mutated using the QuikChange site-directed mutagenesis kit (Stratagene). The following substitutions were made: I22V, L50A, L105A, I109A, V160A, W66F, W66F/L44A, and W66F/L105A (Fig. 1), replacing buried hydrophobic residues by smaller ones. In addition, the mutants L6A, V31A, I51V, I52V, A84G, V117A, V139A, L143A, and I156V (Campos et al. 2004) were used in the analysis (Table 1). The 17 mutants were purified as described (Fillat et al. 1991) and, in all cases, the final preparations were homogeneous as judged by SDS-PAGE.

To obtain the apoflavodoxins, the flavin mononucleotide group was removed from the holoproteins by precipitation with trichloroacetic acid (Edmondson and Tollin 1971). Apoflavodoxin concentration was determined from the absorbance at 280 nm using an extinction coefficient of $34.1 \text{ mM}^{-1} \text{ cm}^{-1}$ (Genzor et al. 1996a), except for W66F and variants containing this mutation, for which an extinction coefficient of $27.6 \text{ mM}^{-1} \text{ cm}^{-1}$ was used (S. Zorrilla, unpubl.).

Absorbance, fluorescence, and circular dichroism (CD) spectra

Emission fluorescence were recorded at $25.0^\circ \pm 0.1^\circ\text{C}$ in 50 mM MOPS (pH 7.0), in an Aminco-Bowman Series 2 fluorimeter. CD spectra were recorded in a Jasco 710 spectropolarimeter at $25.0^\circ \pm 0.1^\circ\text{C}$. Near-UV CD spectra were measured in a 1-cm path-length cuvette with 40 μM apoprotein solutions in 50 mM MOPS (pH 7). Far-UV CD spectra were recorded in a 1-mm path-length cuvette using 20 μM apoprotein solutions in 5 mM MOPS (pH 7), with 15 mM NaCl to obtain the same ionic strength of the 50 mM MOPS buffer. Thermal and urea denaturation curves were recorded in the same instruments.

Urea denaturation equilibrium

Protein samples were prepared by mixing 900- μL urea solutions of different concentrations with 100- μL aliquots of 20 μM apoprotein in 500 mM MOPS (pH 7.0). The unfolding at $25.0^\circ \pm 0.1^\circ\text{C}$ was followed, after equilibration for 30 min by measuring a ratio of fluorescence emission (320:360 nm, with excitation at 280 nm; see López-Llano et al. 2006 for details). The urea unfolding of apoflavodoxin at pH 7.0 is two-state (Genzor et al. 1996a). The unfolding curves were analyzed assuming that the free energy of unfolding, ΔG , is a linear function of urea concentration (Pace 1990):

$$\Delta G = \Delta G_w - mD \quad (1)$$

where ΔG_w is the free energy of unfolding in the absence of denaturant, D is the molar concentration of urea, and m is a proportionality constant. The spectroscopic signal of the

folded (S_F) and unfolded (S_U) states are assumed to vary linearly with urea concentration (Bolen and Santoro 1988) with slopes m_F and m_U , respectively. Under this assumption, the observed spectroscopic signal follows:

$$S = \frac{S_F + m_F D + (S_U + m_U D)e^{-\Delta G/RT}}{1 + e^{-\Delta G/RT}} \quad (2)$$

where R is the gas constant and T is the absolute temperature.

Thermal unfolding followed spectroscopically

Four different spectroscopic techniques were used to follow the thermal denaturation of each protein variant. For fluorescence, a ratio of emission (320:360 nm with excitation at 280 nm) was used to minimize the strong temperature dependence of the baselines. Apoprotein concentration was 2 μM in 50 mM MOPS (pH 7). Circular dichroism and absorbance in the near-UV were simultaneously measured at 291 nm using 40 μM apoflavodoxin solution in 50 mM MOPS (pH 7). Circular dichroism in the far-UV (222 nm) was recorded using 1 μM apoprotein solution in 5 mM MOPS (pH 7), with 15 mM NaCl. Near- and far-UV experiments were carried out using a 1-cm path-length cuvette.

Thermal unfolding: Three-state global analysis

Global fitting of the four thermal unfolding curves recorded for each protein (fluorescence, far-UV CD, near-UV CD, and near-UV absorbance) to a three-state model involving native (N), intermediate (I), and unfolded (U) conformations was performed with the program MLAB (Civilized Software, Inc.).

For each protein, the curves were fitted to the following three-state equation, with an adaptation (Campos et al. 2004) of the one previously described for a two-state model (Privalov 1979):

$$Y = \frac{Y_N + m_N T + (Y_I + m_I T)e^{-\Delta G_{NI}/RT} + (Y_U + m_U T)e^{-(\Delta G_{NI} + \Delta G_{IU})/RT}}{1 + e^{-\Delta G_{NI}/RT} + e^{-(\Delta G_{NI} + \Delta G_{IU})/RT}} \quad (3)$$

with

$$\Delta G_{NI} = \Delta H_{NI}(1 - T/T_{mNI}) - \Delta C_{pNI}[(T_{mNI} - T) + T \ln(T/T_{mNI})] \quad (4)$$

$$\Delta G_{IU} = \Delta H_{IU}(1 - T/T_{mIU}) - \Delta C_{pIU}[(T_{mIU} - T) + T \ln(T/T_{mIU})] \quad (5)$$

where Y is the observed spectroscopic signal; Y_N , Y_I , and Y_U are spectroscopic signals at $T = 0 \text{ K}$ of the native, intermediate, and denatured states, respectively; m_N , m_I , and m_U are slopes that describe the temperature dependence of the spectroscopic signals of the native, intermediate, and denatured states, respectively; ΔH_{NI} , T_{mNI} , and ΔC_{pNI} correspond, respectively, to the enthalpy change, the melting temperature, and the change in heat capacity of the first transition; and ΔH_{IU} , T_{mIU} , and ΔC_{pIU} to those of the second transition. For each mutant, the four thermal unfolding curves were roughly normalized in order to obtain property value spans of $\sim 1-0$, which balances the contribution of the different curves in the fitting without having to perform preliminary individual fits. All of the curves were then globally fitted to common ΔH , T_m , and ΔC_p values using MLAB (Civilized Software, Inc.). Due to the typically large errors associated to fitted ΔC_p values, the ΔC_p values were constrained in the fitting to be within the standard deviation

interval around the mean of the ΔC_p values calculated without restrictions for 40 different apoflavodoxin mutants as described in Campos et al. (2004a). Standard deviations have been determined as described in Beechem (1992).

Equilibrium Φ —analysis of the thermal intermediate

Inspired in classical Φ -analysis (Fersht et al. 1992) originally devised to map out side-chain interactions in transition states of protein unfolding and in transient intermediates, we have recently developed equilibrium Φ -analysis as a way to investigate the integrity in equilibrium intermediates of side-chain interactions present in native states (Irun et al. 2001a; Campos et al. 2004; Korkin et al. 2005). In this analysis, Φ values of 1 or 0 represent native interactions that are present or absent, respectively, in the intermediate, while intermediate Φ values, especially those at internal positions, are interpreted as indicative of weakened interactions. Equilibrium Φ -values are calculated as:

$$\Phi_I = \Delta\Delta G_{IU} / \Delta\Delta G_{NU} \quad (6)$$

Although Equations 4 and 5 provide a way to calculate ΔG_{NI} and ΔG_{IU} values, the inaccuracy of the fitted ΔC_p values misadvises it, and we use a simplified equation that allows us to calculate equilibrium Φ values at a precise reference temperature (chosen at 317.3 K). The details of the fitting have been previously described (Campos et al. 2004).

Modeling of protein mutant structure: Cavity volume and solvent accessibility calculations

The modeling of the mutant structures was done as described in Machicado et al. (2002). Energy minimizations were carried out using the CHARMM force field (Brooks et al. 1983). A cutoff distance of 11 Å was used for nonbonded interactions (with a smoothing from 8 Å). Each structure was energy-minimized, without restrictions, using the method of steepest descent (2000 steps). Minimizations were started from the X-ray structure of wild-type *Anabaena* PCC 7119 apoflavodoxin after having implemented the appropriate in silico mutation. Cavity volume was calculated with a probe radius of 1.4 Å using the Connolly method (Connolly 1983) as implemented in SwissPDBViewer (Guex and Peitsch 1996).

The difference in solvent-accessible area buried in wild-type and mutant folded proteins was estimated as described (Serrano et al. 1992). First, the solvent accessibility of the residue of interest in the wild-type structure (calculated using the Lee and Richards algorithm [Lee and Richards 1971]) was subtracted from that calculated for that residue in an extended tripeptide Gly-X-Gly (Miller et al. 1987) to give a value *A*. In parallel, the in silico mutation was performed by deleting the appropriate carbon atoms from the coordinate file of the wild-type protein, and the solvent-accessible area of the mutated residue in the protein was calculated. This value was subtracted from the one calculated in the tripeptide to give a value *B*. The difference between the two, *A*–*B*, is the loss of solvent-accessible area buried in the folded protein upon mutation.

Cavity volumes and differences in solvent-accessible area were also calculated in mutants from ROP (Vlassi et al. 1999), barnase (Serrano et al. 1992), T4 lysozyme (Eriksson et al. 1992a; Xu et al. 1998), staphylococcal nuclease (Shortle et al.

1990), CI2 (Otzen and Fersht 1995), and human lysozyme (Takano et al. 1995, 1997).

Analysis of packing density

The number of contacts between residues located at the site of mutation was calculated on the wild-type structure (PDB code 1ftg) and modeled mutants using the WHATIF software package (Vriend 1990). Similar calculations were made in mutant proteins from ROP (1rop), barnase (1bni), T4 lysozyme (1l63), staphylococcal nuclease (1stn), chymotrypsin inhibitor (2ci2), and human lysozyme (1lz1).

Residue contact parameters, n_H , were calculated as defined (Vlassi et al. 1999). Each atom *i* of a given core residue *h* (site of mutation) is assumed to make N_i contacts to *j* atoms within a 6 Å radius. Each contact is weighted by $1/d_{ij}$, the inverse of the distance between atom *i* and *j*, and by $1/r_{acc,j}$, the inverse of the relative accessibility area (calculated using NACCESS [Hubbard et al. 1991]) of the residue to which atom *j* belongs.

$$n_H = \sum_{i=1}^n \sum_{j=1}^{N_i} \frac{1}{d_{ij}} \frac{1}{r_{acc,j}} \quad (7)$$

Packing densities were calculated in two ways: as the number of methyl and methylene groups within 6 Å of any side-chain atom removed by mutation (Otzen and Fersht 1995) and as the number of C_α carbons within a sphere of 10 Å radius from the C_α of the mutated residue (Shortle et al. 1990). These calculations were performed using the MMTK software (Hinsen 2000).

Flavodoxin sequence alignment

A total of 75 flavodoxins homologs of the *Anabaena* PCC7119 flavodoxin sequence were identified with BLAST (Altschul et al. 1997) and preliminarily aligned with CLUSTALW (Thompson et al. 1994). Incomplete sequences were discarded and the remaining sequences realigned. Each sequence was weighted following an established procedure (Henikoff and Henikoff 1994) so that weighted percentages of conservation of residues at each flavodoxin position mutated in this work were calculated.

Acknowledgments

We acknowledge financial support from grant BFU2004-01411 (Spain). M.B., L.A.C., and J.E. were supported by MEC fellowships.

References

- Adamek, D.H., Guerrero, L., Blaber, M., and Caspar, D.L. 2005. Structural and energetic consequences of mutations in a solvated hydrophobic cavity. *J. Mol. Biol.* **346**: 307–318.
- Akasaka, A., Haruki, M., Oobatake, M., and Kanaya, S. 1997. Conformational stabilities of *Escherichia coli* RNase HI variants with a series of amino acid substitutions at a cavity within the hydrophobic core. *J. Biol. Chem.* **272**: 18686–18693.
- Alber, T., Sun, D.P., Nye, J.A., Muchmore, D.C., and Matthews, B.W. 1987. Temperature-sensitive mutations of bacteriophage T4 lysozyme occur at sites with low mobility and low solvent accessibility in the folded protein. *Biochemistry* **26**: 3754–3758.

- Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**: 3389–3402.
- Baase, W.A., Eriksson, A.E., Zhang, X.J., Heinz, D.W., Sauer, U., Blaber, M., Baldwin, E.P., Wozniak, J.A., and Matthews, B.W. 1992. Dissection of protein structure and folding by directed mutagenesis. *Faraday Discuss.* **93**: 173–181.
- Bairoch, A. and Apweiler, R. 1997. The SWISS-PROT protein sequence database: Its relevance to human molecular medical research. *J. Mol. Med.* **75**: 312–316.
- Beecham, J.M. 1992. Global analysis of biochemical and biophysical data. *Methods Enzymol.* **210**: 37–54.
- Bolen, D.W. and Santoro, M.M. 1988. Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of ΔG^{N-U} values in a thermodynamic cycle. *Biochemistry* **27**: 8069–8074.
- Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S., and Karplus, M. 1983. CHARMM: A program for macromolecular energy, minimisation, and dynamics calculations. *J. Comput. Chem.* **4**: 187–217.
- Buckle, A.M., Henrick, K., and Fersht, A.R. 1993. Crystal structural analysis of mutations in the hydrophobic cores of barnase. *J. Mol. Biol.* **234**: 847–860.
- Buckle, A.M., Cramer, P., and Fersht, A.R. 1996. Structural and energetic responses to cavity-creating mutations in hydrophobic cores: Observation of a buried water molecule and the hydrophilic nature of such hydrophobic cavities. *Biochemistry* **35**: 4298–4305.
- Bueno, M., Cremades, N., Neira, J.L., and Sancho, J. 2006. Filling small, empty protein cavities: Structural and energetic consequences. *J. Mol. Biol.* **358**: 701–712.
- Campos, L.A., Bueno, M., Lopez-Llano, J., Jiménez, M.A., and Sancho, J. 2004a. Structure of stable protein folding intermediates by equilibrium ϕ -analysis: The apoflavodoxin thermal intermediate. *J. Mol. Biol.* **344**: 239–255.
- Campos, L.A., Garcia-Mira, M.M., Godoy-Ruiz, R., Sanchez-Ruiz, J.M., and Sancho, J. 2004b. Do proteins always benefit from a stability increase? Relevant and residual stabilisation in a three-state protein by charge optimisation. *J. Mol. Biol.* **344**: 223–237.
- Campos, L.A., Cuesta-Lopez, S., López-Llano, J., Falo, F., and Sancho, J. 2005. A double-deletion method to quantifying incremental binding energies in proteins from experiment. Example of a destabilizing hydrogen bonding pair. *Biophys. J.* **88**: 1311–1321.
- Chan, H.S. and Dill, K.A. 1990. Origins of structure in globular proteins. *Proc. Natl. Acad. Sci.* **87**: 6388–6392.
- Chen, J. and Stites, W.E. 2001a. Energetics of side chain packing in staphylococcal nuclease assessed by systematic double mutant cycles. *Biochemistry* **40**: 14004–14011.
- . 2001b. Packing is a key selection factor in the evolution of protein hydrophobic cores. *Biochemistry* **40**: 15280–15289.
- Connolly, M.L. 1983. Solvent-accessible surfaces of proteins and nucleic acids. *Science* **221**: 709–713.
- Creamer, T.P., Srinivasan, R., and Rose, G.D. 1995. Modeling unfolded states of peptides and proteins. *Biochemistry* **34**: 16245–16250.
- Dill, K.A. 1990. Dominant forces in protein folding. *Biochemistry* **29**: 7133–7155.
- Di Nardo, A.A., Larson, S.M., and Davidson, A.R. 2003. The relationship between conservation, thermodynamic stability, and function in the SH3 domain hydrophobic core. *J. Mol. Biol.* **333**: 641–655.
- Edmondson, D.E. and Tollin, G. 1971. Flavin-protein interactions and the redox properties of the *Shethna* flavoprotein. *Biochemistry* **10**: 133–145.
- Eriksson, A.E., Baase, W.A., Wozniak, J.A., and Matthews, B.W. 1992a. A cavity-containing mutant of T4 lysozyme is stabilized by buried benzene. *Nature* **355**: 371–373.
- Eriksson, A.E., Baase, W.A., Zhang, X.J., Heinz, D.W., Blaber, M., Baldwin, E.P., and Matthews, B.W. 1992b. Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect. *Science* **255**: 178–183.
- Eyal, E., Najmanovich, R., Edelman, M., and Sobolev, V. 2003. Protein side-chain rearrangement in regions of point mutations. *Proteins* **50**: 272–282.
- Fauchere, J.L. and Pliska, V. 1983. Hydrophobic parameters π of amino-acid side-chains from the partitioning of N-Acetyl-amino-acid amides. *Eur. J. Med. Chem.* **18**: 369–375.
- Fernandez-Recio, J. and Sancho, J. 1998. Intrahelical side chain interactions in α -helices: Poor correlation between energetics and frequency. *FEBS Lett.* **429**: 99–103.
- Fernandez-Recio, J., Romero, A., and Sancho, J. 1999. Energetics of a hydrogen bond (charged and neutral) and of a cation- π interaction in apoflavodoxin. *J. Mol. Biol.* **290**: 319–330.
- Fersht, A.R., Matouschek, A., and Serrano, L. 1992. The folding of an enzyme. I. Theory of protein engineering analysis of stability and pathway of protein folding. *J. Mol. Biol.* **224**: 771–782.
- Fillat, M.F., Borrias, W.E., and Weisbeek, P.J. 1991. Isolation and over-expression in *Escherichia coli* of the flavodoxin gene from *Anabaena* PCC7119. *Biochem. J.* **280**: 187–191.
- Finucane, M.D. and Woolfson, D.N. 1999. Core-directed protein design. II. Rescue of a multiply mutated and destabilized variant of ubiquitin. *Biochemistry* **38**: 11613–11623.
- Genzor, C.G., Beldarrain, A., Gomez-Moreno, C., Lopez-Lacomba, J.L., Cortijo, M., and Sancho, J. 1996a. Conformational stability of apoflavodoxin. *Protein Sci.* **5**: 1376–1388.
- Genzor, C.G., Perales-Alcon, A., Sancho, J., and Romero, A. 1996b. Closure of a tyrosine/tryptophan aromatic gate leads to a compact fold in apoflavodoxin. *Nat. Struct. Biol.* **3**: 329–332.
- Godoy-Ruiz, R., Perez-Jimenez, R., Ibarra-Molero, B., and Sanchez-Ruiz, J.M. 2004. Relation between protein stability, evolution and structure, as probed by carboxylic acid mutations. *J. Mol. Biol.* **336**: 313–318.
- Guex, N. and Peitsch, M. 1996. Swiss-PdbViewer: A fast and easy-to-use PDB viewer for Macintosh and PC. *Protein Data Bank Quarterly Newsletter* **77**: 7.
- Halaby, D.M., Poupon, A., and Mornon, J. 1999. The immunoglobulin fold family: Sequence analysis and 3D structure comparisons. *Protein Eng.* **12**: 563–571.
- Henikoff, S. and Henikoff, J.G. 1994. Position-based sequence weights. *J. Mol. Biol.* **243**: 574–578.
- Hinsen, K. 2000. The Molecular Modeling Toolkit: A new approach to molecular simulations. *J. Comput. Chem.* **21**: 79–85.
- Hubbard, S.J. and Argos, P. 1994. Cavities and packing at protein interfaces. *Protein Sci.* **3**: 2194–2206.
- Hubbard, S.J., Campbell, S.F., and Thornton, J.M. 1991. Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. *J. Mol. Biol.* **220**: 507–530.
- Hubbard, S.J., Gross, K.H., and Argos, P. 1994. Intramolecular cavities in globular proteins. *Protein Eng.* **7**: 613–626.
- Irun, M.P., Garcia-Mira, M.M., Sanchez-Ruiz, J.M., and Sancho, J. 2001a. Native hydrogen bonds in a molten globule: The apoflavodoxin thermal intermediate. *J. Mol. Biol.* **306**: 877–888.
- Irun, M.P., Maldonado, S., and Sancho, J. 2001b. Stabilization of apoflavodoxin by replacing hydrogen-bonded charged Asp or Glu residues by the neutral isosteric Asn or Gln. *Protein Eng.* **14**: 173–181.
- Ishikawa, K., Nakamura, H., Morikawa, K., and Kanaya, S. 1993. Stabilization of *Escherichia coli* ribonuclease HI by cavity-filling mutations within a hydrophobic core. *Biochemistry* **32**: 6171–6178.
- Jackson, S.E., Moracci, M., elMasry, N., Johnson, C.M., and Fersht, A.R. 1993. Effect of cavity-creating mutations in the hydrophobic core of chymotrypsin inhibitor 2. *Biochemistry* **32**: 11259–11269.
- Kellis Jr., J.T., Nyberg, K., and Fersht, A.R. 1989. Energetics of complementary side-chain packing in a protein hydrophobic core. *Biochemistry* **28**: 4914–4922.
- Langdon, G.M., Jiménez, M.A., Genzor, C.G., Maldonado, S., Sancho, J., and Rico, M. 2001. *Anabaena* apoflavodoxin hydrogen exchange: On the stable exchange core of the α/β (21345) flavodoxin-like family. *Proteins* **43**: 476–488.
- Lassalle, M.W., Yamada, H., Morii, H., Ogata, K., Sarai, A., and Akasaka, K. 2001. Filling a cavity dramatically increases pressure stability of the c-Myb R2 subdomain. *Proteins* **45**: 96–101.
- Lee, B. and Richards, F.M. 1971. The interpretation of protein structures: Estimation of static accessibility. *J. Mol. Biol.* **55**: 379–400.
- Lehmann, M., Pasamontes, L., Lassen, S.F., and Wyss, M. 2000. The consensus concept for thermostability engineering of proteins. *Biochim. Biophys. Acta* **1543**: 408–415.
- Lehmann, M., Loch, C., Middendorf, A., Studer, D., Lassen, S.F., Pasamontes, L., van Loon, A.P., and Wyss, M. 2002. The consensus concept for thermostability engineering of proteins: Further proof of concept. *Protein Eng.* **15**: 403–411.
- López-Llano, J., Maldonado, S., Bueno, M., Lostao, A., Angeles-Jimenez, M., Lillo, M.P., and Sancho, J. 2004. The long and short flavodoxins: I. The role of the differentiating loop in apoflavodoxin structure and FMN binding. *J. Biol. Chem.* **279**: 47177–47183.
- López-Llano, J., Campos, L.A., Bueno, M., and Sancho, J. 2006. Equilibrium Φ -analysis of a molten globule: The 1-149 apoflavodoxin fragment. **356**: 354–366.
- Lostao, A., El Harrou, M., Daoudi, F., Romero, A., Parody-Morreale, A., and Sancho, J. 2000. Dissecting the energetics of the apoflavodoxin-FMN complex. *J. Biol. Chem.* **275**: 9518–9526.

- Lostao, A., Daoudi, F., Irun, M.P., Ramon, A., Fernandez-Cabrera, C., Romero, A., and Sancho, J. 2003. How FMN binds to anabaena apoflavodoxin: A hydrophobic encounter at an open binding site. *J. Biol. Chem.* **278**: 24053–24061.
- Machicado, C., Bueno, M., and Sancho, J. 2002. Predicting the structure of protein cavities created by mutation. *Protein Eng.* **15**: 669–675.
- Maldonado, S., Lostao, A., Irun, M.P., Fernandez-Recio, J., Gustavo Genzor, C., Begona Gonzalez, E., Rubio, J.A., Luquita, A., Daoudi, F., and Sancho, J. 1998. Apoflavodoxin: Structure, stability, and FMN binding. *Biochimie* **80**: 813–820.
- Matsumura, M., Becktel, W.J., and Matthews, B.W. 1988. Hydrophobic stabilization in T4 lysozyme determined directly by multiple substitutions of Ile-3. *Nature* **334**: 406–410.
- Matthews, B.W. 1993. Structural and genetic analysis of protein-folding and stability. *Curr. Opin. Struct. Biol.* **3**: 589–593.
- Maxwell, K.L. and Davidson, A.R. 1998. Mutagenesis of a buried polar interaction in an SH3 domain: Sequence conservation provides the best prediction of stability effects. *Biochemistry* **37**: 16172–16182.
- Milla, M.E., Brown, B.M., and Sauer, R.T. 1994. Protein stability effects of a complete set of alanine substitutions in Arc repressor. *Nat. Struct. Biol.* **1**: 518–523.
- Miller, S., Janin, J., Lesk, A.M., and Chothia, C. 1987. Interior and surface of monomeric proteins. *J. Mol. Biol.* **196**: 641–656.
- Munson, M., Balasubramanian, S., Fleming, K.G., Nagi, A.D., O'Brien, R., Sturtevant, J.M., and Regan, L. 1996. What makes a protein a protein? Hydrophobic core designs that specify stability and structural properties. *Protein Sci.* **5**: 1584–1593.
- Ohmura, T., Ueda, T., Ootsuka, K., Saito, M., and Imoto, T. 2001. Stabilization of hen egg white lysozyme by a cavity-filling mutation. *Protein Sci.* **10**: 313–320.
- Otzen, D.E. and Fersht, A.R. 1995. Side-chain determinants of β -sheet stability. *Biochemistry* **34**: 5718–5724.
- Otzen, D.E., Rheinhecker, M., and Fersht, A.R. 1995. Structural factors contributing to the hydrophobic effect: The partly exposed hydrophobic minicore in chymotrypsin inhibitor 2. *Biochemistry* **34**: 13051–13058.
- Pace, C.N. 1990. Conformational stability of globular proteins. *Trends Biochem. Sci.* **15**: 14–17.
- . 1992. Contribution of the hydrophobic effect to globular protein stability. *J. Mol. Biol.* **226**: 29–35.
- Pace, C.N. and Laurents, D.V. 1989. A new method for determining the heat capacity change for protein folding. *Biochemistry* **28**: 2520–2525.
- Privalov, P.L. 1979. Stability of proteins: Small globular proteins. *Adv. Protein Chem.* **33**: 167–241.
- Rashin, A.A., Iofin, M., and Honig, B. 1986. Internal cavities and buried waters in globular proteins. *Biochemistry* **25**: 3619–3625.
- Sancho, J. 2006. Flavodoxin: Sequence, folding, binding, function and beyond. *Cell. Mol. Life Sci.* **63**: 855–864.
- Sancho, J., Bueno, M., Campos, L.A., Fernandez-Recio, J., Irun, M.P., Lopez, J., Machicado, C., Pedrosa, I., and Toja, M. 2002. The 'relevant' stability of proteins with equilibrium intermediates. *ScientificWorldJournal* **2**: 1209–1215.
- Sandberg, W.S. and Terwilliger, T.C. 1991. Energetics of repacking a protein interior. *Proc. Natl. Acad. Sci.* **88**: 1706–1710.
- Serrano, L., Kellis Jr., J.T., Cann, P., Matouschek, A., and Fersht, A.R. 1992. The folding of an enzyme. II. Substructure of barnase and the contribution of different interactions to protein stability. *J. Mol. Biol.* **224**: 783–804.
- Serrano, L., Day, A.G., and Fersht, A.R. 1993. Step-wise mutation of barnase to binase. A procedure for engineering increased stability of proteins and an experimental analysis of the evolution of protein stability. *J. Mol. Biol.* **233**: 305–312.
- Shortle, D. 2003. Propensities, probabilities, and the Boltzmann hypothesis. *Protein Sci.* **12**: 1298–1302.
- Shortle, D., Stites, W.E., and Meeker, A.K. 1990. Contributions of the large hydrophobic amino acids to the stability of staphylococcal nuclease. *Biochemistry* **29**: 8033–8041.
- Shrake, A. and Rupley, J.A. 1973. Environment and exposure to solvent of protein atoms—Lysozyme and insulin. *J. Mol. Biol.* **79**: 351–371.
- Steipe, B., Schiller, B., Pluckthun, A., and Steinbacher, S. 1994. Sequence statistics reliably predict stabilizing mutations in a protein domain. *J. Mol. Biol.* **240**: 188–192.
- Takano, K., Ogasahara, K., Kaneda, H., Yamagata, Y., Fujii, S., Kanaya, E., Kikuchi, M., Oobatake, M., and Yutani, K. 1995. Contribution of hydrophobic residues to the stability of human lysozyme: Calorimetric studies and X-ray structural analysis of the five isoleucine to valine mutants. *J. Mol. Biol.* **254**: 62–76.
- Takano, K., Yamagata, Y., Fujii, S., and Yutani, K. 1997. Contribution of the hydrophobic effect to the stability of human lysozyme: Calorimetric studies and X-ray structural analyses of the nine valine to alanine mutants. *Biochemistry* **36**: 688–698.
- Takano, K., Yamagata, Y., and Yutani, K. 2003. Buried water molecules contribute to the conformational stability of a protein. *Protein Eng.* **16**: 5–9.
- Taneja, B. and Mande, S.C. 1999. Conserved structural features and sequence patterns in the GroES fold family. *Protein Eng.* **12**: 815–818.
- Thompson, J.D., Higgins, D.G., and Gibson, T.J. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
- Vlassi, M., Cesareni, G., and Kokkinidis, M. 1999. A correlation between the loss of hydrophobic core packing interactions and protein stability. *J. Mol. Biol.* **285**: 817–827.
- Vriend, G. 1990. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.* **8**: 52–56.
- Wang, Q.H., Buckle, A.M., and Fersht, A.R. 2000. Stabilization of GroEL minichaperones by core and surface mutations. *J. Mol. Biol.* **298**: 917–926.
- Williams, M.A., Goodfellow, J.M., and Thornton, J.M. 1994. Buried waters and internal cavities in monomeric proteins. *Protein Sci.* **3**: 1224–1235.
- Wirtz, P. and Steipe, B. 1999. Intrabody construction and expression III: Engineering hyperstable V(H) domains. *Protein Sci.* **8**: 2245–2250.
- Wolfenden, R. and Radzicka, A. 1994. On the probability of finding a water molecule in a nonpolar cavity. *Science* **265**: 936–937.
- Xu, J., Baase, W.A., Baldwin, E., and Matthews, B.W. 1998. The response of T4 lysozyme to large-to-small substitutions within the core and its relation to the hydrophobic effect. *Protein Sci.* **7**: 158–177.
- Xu, B., Hua, Q.X., Nakagawa, S.H., Jia, W., Chu, Y.C., Katsoyannis, P.G., and Weiss, M.A. 2002. A cavity-forming mutation in insulin induces segmental unfolding of a surrounding α -helix. *Protein Sci.* **11**: 104–116.
- Yu, B., Blazer, M., Gronenborn, A.M., Clore, G.M., and Caspar, D.L. 1999. Disordered water within a hydrophobic protein cavity visualized by x-ray crystallography. *Proc. Natl. Acad. Sci.* **96**: 103–108.
- Yutani, K., Ogasahara, K., Tsujita, T., and Sugino, Y. 1987. Dependence of conformational stability on hydrophobicity of the amino acid residue in a series of variant proteins substituted at a unique position of tryptophan synthase α subunit. *Proc. Natl. Acad. Sci.* **84**: 4441–4444.