Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 10573–10577, December 1991
Biochemistry

Evaluating the effects of a single amino acid substitution on both the native and denatured states of a protein

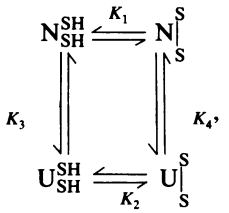
TIAO-YIN LIN* AND PETER S. KIM

Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, MA 02142

Communicated by Robert L. Baldwin, August 19, 1991

ABSTRACT **For proteins that contain a disulfide bond, stability is linked thermodynamically to thiol–disulfide exchange. We use this relationship to obtain unfolding free energies for both the reduced and oxidized forms of *Escherichia coli* thioredoxin from measurements of the effective concentrations of protein thiols. We then evaluate the effect of an amino acid substitution on disulfide bond formation in both the native and denatured states of the protein. Although the Pro-34 → Ser substitution in thioredoxin results in a decrease of the effective concentration of protein thiols in the native state, the effective concentration increases in the denatured state. The net effect of the amino acid substitution is to increase the stability of reduced thioredoxin by ≈2.4 kcal/mol, whereas the stability of the oxidized protein remains the same. By assuming a two-state unfolding equilibrium and a mutation free energy of −7.7 kcal/mol for the Pro-34 → Ser substitution in the reduced, urea-unfolded state (based on estimates of solvation and entropic changes), we obtain relative free energies for the native and denatured states of the mutant and wild-type proteins, in both the reduced and oxidized forms.**

There is increasing interest in evaluating the effect of amino acid substitutions on protein stability (for review, see refs. 1 and 2). Typically, unfolding free energies for the wild-type and mutant proteins are obtained by measuring a spectroscopic parameter as a function of denaturant concentration (for review, see ref. 3). For proteins that contain a disulfide bond, unfolding free energies can be obtained in an alternative manner (e.g., refs. 4 and 5) by taking advantage of the following thermodynamic cycle:



[1]

where N^{S}_{S}, N^{SH}_{SH}, U^{S}_{S}, and U^{SH}_{SH} represent the protein in the native oxidized, native reduced, unfolded oxidized, and unfolded reduced states, respectively; K_3 and K_4 are unfolding equilibrium constants for the reduced and oxidized proteins, and K_1 and K_2 represent pseudo equilibrium constants for the half reactions of disulfide formation in the native and denatured states, respectively. The linkage between conformational stability and disulfide stability is

$$\frac{K_3}{K_4} = \frac{K_1}{K_2}. \qquad [2]$$

The ratio between K_1 and K_2 can be obtained by measuring the effective concentration (C_{eff}) of the protein thiols in the native and unfolded states. Creighton introduced the concept of C_{eff} as a useful way to think about linkage relationships in protein stability (6). The C_{eff} of two interacting groups is the ratio of rate or equilibrium constants of otherwise identical intra- and intermolecular reactions (7). For a protein with two cysteines that can form a disulfide bond, C_{eff} can be obtained by comparing the pseudo equilibrium constant for disulfide bond formation in the protein (K_{intra}) to that for intermolecular disulfide bond formation (K_{inter}) in a reference thiol such as glutathione (5, 6):

$$P^{SH}_{SH} + GSSG \overset{K_{eq}}{\rightleftharpoons} P^{S}_{S} + 2\ GSH \qquad [3]$$

and

$$C_{eff} = \frac{K_{intra}}{K_{inter}} = \frac{[P^{S}_{S}][GSH]^2}{[P^{SH}_{SH}][GSSG]} = K_{eq}, \qquad [4]$$

where P^{SH}_{SH} refers to reduced protein, P^{S}_{S} refers to oxidized protein, GSSG refers to oxidized glutathione, and GSH refers to reduced glutathione. Thiol concentrations are meant to include the thiolate species (cf. ref. 8), and C_{eff} values are expressed in units of molarity.

If we let C^{N}_{eff} and C^{U}_{eff} represent C_{eff} of the thiols in the native and denatured protein, respectively, we obtain from Eq. 2:

$$\frac{K_3}{K_4} = \frac{C^{N}_{eff}}{C^{U}_{eff}}. \qquad [5]$$

The thermodynamic cycle shown in Eq. 1 has been confirmed experimentally for *Escherichia coli* thioredoxin by measuring independently each of the four equilibrium constants (5). This confirmation indicates that the two-state approximation for unfolding is sufficient to describe the linkage relationship for thioredoxin, even though urea-gradient gel electrophoresis experiments (9) detect a minor native form of the reduced protein, thought to result from isomerization of Pro-76.

Here, we show how unfolding free energies can be derived from the urea dependence of C_{eff}, and we obtain experimentally the stability of reduced and oxidized thioredoxin from a single set of C_{eff} measurements. The same approach is used with a mutant thioredoxin protein, containing the substitution

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: C_{eff}, effective concentration; C^{N}_{eff} and C^{U}_{eff}, C_{eff} of thiols in native and denatured states, respectively; N^{S}_{S}, native oxidized state; N^{SH}_{SH}, native reduced state; U^{S}_{S}, unfolded oxidized state; U^{SH}_{SH}, unfolded reduced state; K_1 and K_2, pseudo equilibrium constants for the half reactions of disulfide formation in the native and denatured states, respectively; K_3 and K_4, unfolding equilibrium constants for the reduced and oxidized proteins, respectively.

*Present address: Graduate Department of Biochemistry, Brandeis University, Waltham, MA 02254.

Pro-34 → Ser. Using these results, we estimate the relative free energies for the native and unfolded states of the mutant and wild-type proteins, in both the reduced and oxidized forms.

MATERIALS AND METHODS

The Pro-34 → Ser mutant of *E. coli* thioredoxin (P34S) was provided by D. M. LeMaster and F. M. Richards (Yale University). The material was judged pure by reverse-phase HPLC. All other materials and methods used have been described by Lin and Kim (5). Briefly, C_{eff} was measured in a buffer containing 0.1 M Tris, 0.2 M KCl, and 1 mM EDTA (pH 8.7). Urea solutions were prepared in the same buffer with a final pH of 8.7. A low concentration of protein in the reaction mixture was used to prevent dimer formation. The concentrations of GSSG and GSH were ≈100-fold in molar excess of the protein, so that the redox potential of the solution was fixed (5, 10). For our experiments, the protein and GSSG concentrations were ≈5 μM and ≈500 μM, respectively, and the GSH concentration ranged from 3 to 70 mM.

All manipulations were done under argon to prevent air oxidation. The reaction mixture was equilibrated at room temperature (23°C) for 2 hr or longer and then quenched with HCl to pH 2. Measurements of C_{eff} as a function of time indicated that equilibrium had been reached. GSH, GSSG, reduced thioredoxin, and oxidized thioredoxin were then separated on a Bio-Rad C_{18} reverse-phase HPLC column with a shallow gradient of acetonitrile in the presence of 0.1% trifluoroacetic acid. Absorbance was monitored simultaneously at 229 nm and 280 nm, and peaks were quantitated by integration. The relative extinction coefficients of reduced and oxidized thioredoxin were taken to be the same at both wavelengths because there was no significant difference in peak area when the same amount of reduced or oxidized protein was injected.

RESULTS

We have shown (5) that the C_{eff} of thiols in a protein can be measured as a function of urea concentration and that, for a structured polypeptide, C_{eff} changes as the polypeptide is unfolded by urea. Assuming a two-state mechanism for the denaturant-induced unfolding, from Eq. **4**, C_{eff} in the unfolding transition region can be expressed as

$$C_{eff} = \frac{([N^S_S] + [U^S_S])\,[GSH]^2}{([N^{SH}_{SH}] + [U^{SH}_{SH}])\,[GSSG]}. \qquad [6]$$

Eq. **6** can be expanded as follows:

$$C_{eff} = \left\{\frac{[N^S_S]}{[N^{SH}_{SH}]}\left(\frac{[N^{SH}_{SH}]}{[N^{SH}_{SH}]+[U^{SH}_{SH}]}\right) + \frac{[U^S_S]}{[U^{SH}_{SH}]}\left(\frac{[U^{SH}_{SH}]}{[N^{SH}_{SH}]+[U^{SH}_{SH}]}\right)\right\}\left(\frac{[GSH]^2}{[GSSG]}\right), \qquad [7]$$

which simplifies to

$$C_{eff} = C^N_{eff}\left(\frac{1}{1+K_3}\right) + C^U_{eff}\left(\frac{K_3}{1+K_3}\right). \qquad [8]$$

The equilibrium constant for unfolding of reduced polypeptide is therefore

$$K_3 = \frac{C^N_{eff} - C_{eff}}{C_{eff} - C^U_{eff}}. \qquad [9]$$

Finally, the unfolding equilibrium constant of the oxidized polypeptide is obtained from Eqs. **5** and **9**:

$$K_4 = \frac{C^N_{eff}\,C^U_{eff} - C_{eff}\,C^U_{eff}}{C_{eff}\,C^N_{eff} - C^N_{eff}\,C^U_{eff}}. \qquad [10]$$

Fig. 1*A* depicts experimentally determined C_{eff} values for *E. coli* thioredoxin. Within experimental error, at low and high urea concentrations, there is no dependence of C_{eff} on urea. In addition, C_{eff} is independent of urea concentration in a model random coil peptide (5). C_{eff} values for the native and denatured states of thioredoxin are therefore assigned (Fig. 1*A*) to be 10 ± 1 M and 0.024 ± 0.003 M, respectively. By converting the C_{eff} values obtained in the transition zone into unfolding free energies (Eqs. **9** and **10**) and then using a linear extrapolation (3, 11) to zero urea (Fig. 1*B*), we obtain the stabilities of reduced and oxidized thioredoxin as 6.1 kcal/mol and 9.7 kcal/mol, respectively.

The active site thiols of thioredoxin are contained in the sequence Cys-Gly-Pro-Cys that forms a reverse turn in both the reduced and oxidized forms of the native protein (12, 13). Fig. 2*A* depicts C_{eff} for a mutant protein, in which the proline in this turn is replaced by serine (14) as a function of urea concentration, compared to values obtained with the wild-type protein. The striking observation is that, whereas the P34S substitution decreases C_{eff} in the native state, C_{eff} in the

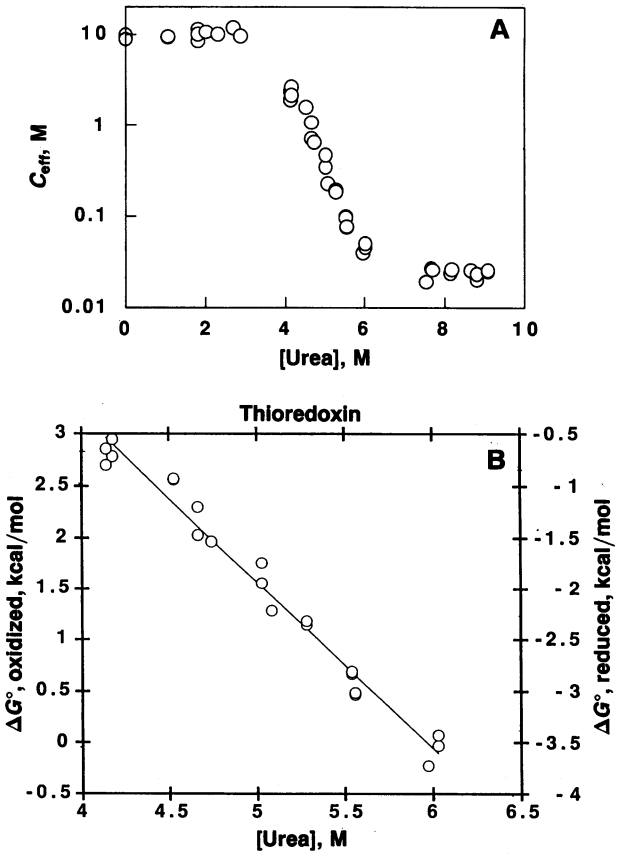


FIG. 1. Determination of unfolding free energies for thioredoxin using C_{eff} measurements. (*A*) Urea dependence of C_{eff} for *E. coli* thioredoxin at pH 8.7 and 23°C. (*B*) Linear extrapolations of the unfolding free energies obtained from C_{eff} measurements. Left and right axes show unfolding free energies in the transition region for oxidized and reduced thioredoxin, respectively, calculated from C_{eff} data by using Eqs. **9** and **10**. Linear extrapolation gives $\Delta G°$ of 6.1 ± 0.3 kcal/mol and 9.7 ± 0.3 kcal/mol for unfolding reduced and oxidized thioredoxin, respectively, without urea (error limits given are the SD of $\Delta G°$ after linear regression).

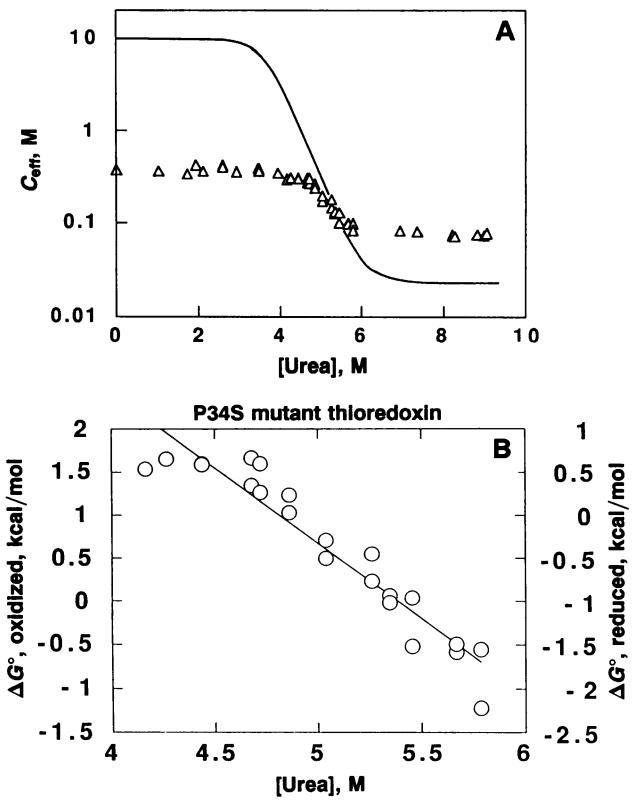


FIG. 2. Determination of unfolding free energies for P34S thioredoxin. (*A*) Urea dependence of C_{eff} for *E. coli* P34S thioredoxin at pH 8.7 and 23°C. Solid line represents values obtained with wild-type protein, interpolated from data in Fig. 1*A*. (*B*) Linear extrapolations of the unfolding free energies for P34S thioredoxin obtained from C_{eff} measurements. Left and right axes show unfolding free energies in the transition region for the oxidized and reduced protein, respectively. Extrapolating linearly to zero molar urea gives $\Delta G°$ of 8.5 ± 0.6 kcal/mol for unfolding reduced P34S thioredoxin and 9.4 ± 0.6 kcal/mol for unfolding oxidized P34S thioredoxin.

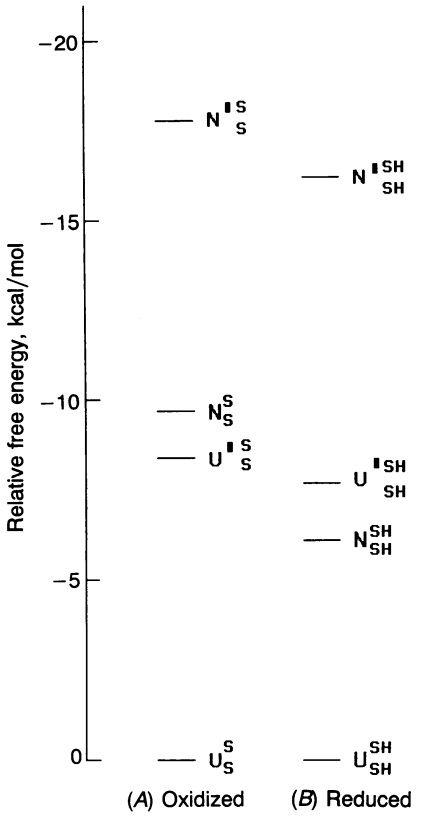


FIG. 3. Relative free energies for native and denatured states of wild-type and mutant (P34S) thioredoxin, obtained by estimating a substitution free energy in the unfolded reduced forms. Primes denote the mutant protein. (*A*) Oxidized forms. (*B*) Reduced forms. Alignment of relative free energies between oxidized and reduced forms (U^{S}_{S} and U^{SH}_{SH} in *A* and *B*) is arbitrary.

unfolded state increases (Fig. 2*A*). The net effect is that the disulfide contributes only ≈1.0 kcal/mol to stability in the mutant protein (as compared to ≈3.6 kcal/mol in the wild-type protein).

The difference in the stabilizing effect of the disulfide bond could result from changes in the unfolding free energies of the reduced and/or oxidized forms of the proteins. The stabilities for the reduced and oxidized forms of P34S thioredoxin are 8.5 kcal/mol and 9.4 kcal/mol, respectively (Fig. 2*B*). Hence, the Pro → Ser substitution does not affect significantly the unfolding free energy of the oxidized form of protein, but it causes a large increase in the unfolding free energy of the reduced form (Table 1).

DISCUSSION

An amino acid substitution can alter the stability of a protein by affecting the native, denatured, or both states of the protein. Our measurements permit us to evaluate the effect of a substitution on both states of the protein, in terms of a specific interaction (disulfide bond formation). The results show that replacement of a single amino acid can have opposite effects on disulfide bond formation in the native and denatured states of a protein. Substitution of Pro-34 by serine (Fig. 2*A*, Table 1) decreases C_{eff} in native thioredoxin by ≈26 fold (≈ +2.0 kcal/mol), whereas it increases C_{eff} in denatured thioredoxin by ≈3 fold (≈ −0.7 kcal/mol).

Within experimental error, C^{N}_{eff} and C^{U}_{eff} are constant with respect to urea concentration (Fig. 1). This observation, taken together with the linkage relationship between unfolding and disulfide bond formation, leads to the conclusion that the oxidized and reduced forms of thioredoxin have the same dependence of the unfolding free energy on urea concentration (i.e., *m*; see ref. 3). For RNase T1, however, reduction of one or both disulfides leads to significant changes in *m* (15). The *m* values for P34S mutant thioredoxin are similar to those for the wild-type protein and also appear relatively insensitive to reduction of the disulfide (Fig. 2).

C_{eff} measurements can report directly on the effect of a mutation on a local interaction (e.g., the effects of the Pro → Ser mutation on disulfide bond formation in the urea-unfolded state). If, however, the unfolding of the protein is two-state, then the C_{eff} values are thermodynamically linked to global unfolding free energies. It is important to emphasize

Table 1. Effects of Pro-34 → Ser substitution in thioredoxin

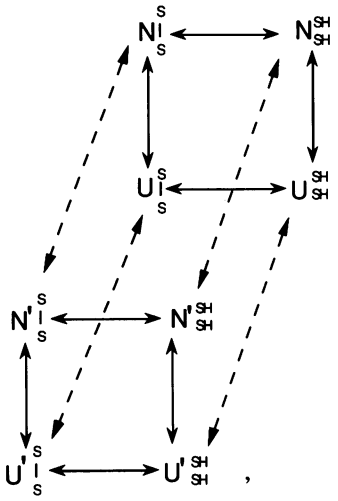
	C^{N}_{eff},* M	C^{U}_{eff},* M	$\Delta G°_{red}$ (N-U), kcal/mol	$\Delta G°_{oxi}$ (N-U), kcal/mol	$\Delta\Delta G°$ (oxi → red), kcal/mol
Thioredoxin	10 ± 1	0.024 ± 0.003	−6.1 ± 0.3	−9.7 ± 0.3	3.6
P34S	0.38 ± 0.03	0.077 ± 0.004	−8.5 ± 0.6	−9.4 ± 0.6	0.9

*Free energy change (Pro → Ser) for C^{N}_{eff} is 2.0 kcal/mol; free energy change for C^{U}_{eff} is −0.7 kcal/mol.
oxi, Oxidized; red, reduced; N, native state; U, denatured state.

in this regard that the two-state linkage relationship (Eq. 5) has been shown to be a good approximation for wild-type thioredoxin by measuring each parameter independently (5).

It is generally thought that substitution of a trans proline by another amino acid residue will decrease the stability of the protein because it will increase the conformational entropy of the unfolded state (e.g., ref. 16). Proton NMR spectroscopy shows that oxidized and reduced thioredoxin have similar structures (13, 17). We, therefore, expect to see a *decrease* in stability for both the oxidized and reduced forms of the serine mutant, assuming that the Pro-34 → Ser substitution does not change the overall structure of the protein. On the contrary, our results show that the Pro-34 → Ser substitution *increases* the stability of the reduced form by ≈2.4 kcal/mol and that it has essentially no effect on the stability of the oxidized form (Table 1).

Free energy simulations have been used to calculate unfolding free energy differences between mutant and wild-type proteins for T4 lysozyme (18), hemoglobin (19), and triose phosphate isomerase (20). To facilitate comparison of our results with theory, we use the following hypothetical cycle:



[11]

where the primes are used to denote the mutant P34S thioredoxin. The thermodynamic cycles for wild-type and mutant thioredoxin (solid lines) are connected by hypothetical amino acid substitution reactions (dashed lines). If we assign a value to any of the four amino acid substitution energies, the remaining three are obtained by the linkage defined in Eq. 11.

From an experimental point of view, the most reasonable place to estimate the substitution free energy is in the reduced, urea-unfolded state because this state is probably closest to a completely unfolded polypeptide. Following Fersht and coworkers, we consider as a first approximation the difference in hydration free energy (cf. refs. 21 and 22) between a proline and a serine residue and the difference in conformational free energy for a proline and serine residue in an unfolded chain. The values used here are only meant to be estimates, and this approach does not consider the free energy change that results from the creation and destruction of atoms in the hypothetical substitution reactions. We have neglected differences in rotational or vibrational energies between the two residues. We have also assumed that, upon mutation in the reduced, urea-unfolded state, there is no change (cf. refs. 21 and 23) in the free energy of (*i*) nonbonded interactions between the mutation site and the rest of the protein, (*ii*) solvation of residues other than the mutated one, and (*iii*) bonded interactions in the protein.

To evaluate the hydration free energies of the side chains of proline and serine, we begin with values obtained by Wolfenden and coworkers (24) for propane (1.99 kcal/mol) and methanol (−5.06 kcal/mol), respectively. Opening of the ring and addition of the amide proton in the replacement of the proline residue is accounted for (ΔG = 0.49 kcal/mol) by using the difference in vapor to water distribution coefficients for *N*-acetylpyrrolidine and *N*-butylacetamide (25). The difference in hydration free energy between proline and serine is, therefore, estimated to be −6.6 kcal/mol. This value compares well with a calculated value of −5.7 kcal/mol, based on the accessible surface area of proline and serine (26). From calculations of the conformational entropy of unfolding of amino acid residues (16, 27), we estimate that the difference in conformational free energy for a Pro → Ser substitution in an unfolded chain is −1.1 kcal/mol. Thus, we estimate a Pro → Ser substitution free energy of −7.7 kcal/mol.

By assuming that this estimate applies to the urea-unfolded reduced state, we obtain the relative free energies of wild-type and mutant thioredoxin in the oxidized (Fig. 3*A*) and reduced (Fig. 3*B*) forms. In this case, the amino acid substitution (P34S) appears to reduce the free energy of both folded and unfolded thioredoxin in both the oxidized and reduced forms (Fig. 3). In the future, it will be interesting to compare our results with those obtained by theoretical free energy simulations (e.g., refs. 18–20).

The method presented here provides a way to evaluate the effect of an amino acid substitution on a specific interaction (i.e., disulfide bond formation) in the native and denatured states of proteins. In the limiting case of two-state equilibrium unfolding, the method also gives the unfolding free energies of both the oxidized and reduced proteins in the same set of measurements.

In future work, it may be possible to develop further the approach presented here to include intermediate states. The caveats of the method are that the accuracy may not be as high as some optical probes and that it applies only to proteins that contain a disulfide bond. Increasing interest in disulfide engineering (e.g., refs. 28–30), however, may facilitate use of the method.

We thank B. Tidor and T. Alber for advice and discussion in all aspects of this work, and D. M. LeMaster and F. M. Richards for the generous gift of the P34S mutant of thioredoxin. We also thank R. Baldwin, G. Némethy, J. Weissman, R. Wolfenden, and a referee for helpful comments and P. Baud for assistance in manuscript preparation. This work was supported by the Pew Scholars Program in the Biomedical Sciences, the Rita Allen Foundation, and the Whitaker Health Sciences Fund.

1. Alber, T. (1989) *Annu. Rev. Biochem.* **58,** 765–798.
2. Shortle, D. (1989) *J. Biol. Chem.* **264,** 5315–5318.
3. Pace, C. N. (1986) *Methods Enzymol.* **131,** 266–280.
4. Creighton, T. E. (1988) *BioEssays* **8,** 57–63.
5. Lin, T.-Y. & Kim, P. S. (1989) *Biochemistry* **28,** 5282–5287.
6. Creighton, T. E. (1983) *Biopolymers* **22,** 49–58.
7. Page, M. I. & Jencks, W. P. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 1678–1683.
8. Houk, J., Singh, R. & Whitesides, G. M. (1987) *Methods Enzymol.* **143,** 129–140.
9. Langsetmo, K., Fuchs, J. & Woodward, C. (1989) *Biochemistry* **28,** 3211–3220.
10. Snyder, G. H. (1987) *Biochemistry* **26,** 688–694.
11. Schellman, J. A. (1978) *Biopolymers* **17,** 1305–1322.
12. Holmgren, A., Söderberg, B.-O., Eklund, H. & Brändén, C.-I. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 2305–2309.
13. Holmgren, A. & Roberts, G. (1976) *FEBS Lett.* **71,** 261–265.
14. Kelley, R. F. & Richards, F. M. (1987) *Biochemistry* **26,** 6765–6774.
15. Pace, C. N., Grimsley, G. R., Thomson, J. A. & Barnett, B. J. (1988) *J. Biol. Chem.* **263,** 11820–11825.
16. Matthews, B. W., Nicholson, H. & Becktel, W. J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 6663–6667.
17. Dyson, H. J., Holmgren, A. & Wright, P. E. (1989) *Biochemistry* **28,** 7074–7089.

18. Dang, L. X., Merz, K. M. & Kollman, P. A. (1989) *J. Am. Chem. Soc.* **111,** 8505–8508.
19. Gao, J., Kuczera, K., Tidor, B. & Karplus, M. (1989) *Science* **244,** 1069–1072.
20. Daggett, V., Brown, F. & Kollman, P. (1989) *J. Am. Chem. Soc.* **111,** 8247–8256.
21. Kellis, J. T., Jr., Nyberg, K., Sáli, D. & Fersht, A. R. (1988) *Nature (London)* **333,** 784–786.
22. Matouschek, A., Kellis, J. T., Serrano, L. & Fersht, A. R. (1989) *Nature (London)* **340,** 122–126.
23. Tidor, B. & Karplus, M. (1991) *Biochemistry* **30,** 3217–3228.
24. Wolfenden, R., Anderson, L., Cullis, P. M. & Southgate, C. C. B. (1981) *Biochemistry* **20,** 849–855.
25. Gibbs, P., Radzicka, A. & Wolfenden, R. (1991) *J. Am. Chem. Soc.* **113,** 4714–4715.
26. Ooi, T., Oobatake, M., Némethy, G. & Scheraga, H. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 3086–3090.
27. Némethy, G., Leach, S. J. & Scheraga, H. A. (1966) *J. Phys. Chem.* **70,** 998–1004.
28. Wetzel, R., Perry, L. J., Baase, W. A. & Becktel, W. J. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 401–405.
29. Mitchinson, C. & Wells, J. A. (1989) *Biochemistry* **28,** 4807–4815.
30. Matsumura, M., Becktel, W. J., Levitt, M. & Matthews, B. W. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 6562–6566.